NPL Report CISE 18/99

to the National

Measurement System

Policy Unit, Department

of Trade & Industry

INITIAL REPORT ON STATUS OF SOFTWARE AND MATHEMATICS IN EACH METROLOGY AREA

FROM THE SOFTWARE SUPPORT FOR METROLOGY PROGRAMME

BY

D. RAYNER, NPL

February 1999

ISSN 1361-407X

Authorised by Dr Dave Rayner,
Head of the Centre for Information Systems Engineering

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

Contents

1. Introduction

The purpose of this report is to inform all those interested in the UK National Measurement System (NMS) Software Support for Metrology (SSfM) programme about the status of the mathematics and software used to support each metrology area in the UK NMS, concentrating on the status within the respective NMS Programmes. The metrology areas covered are Foundation Metrology, Length, Mass, Electrical Metrology, Photonics, Optical Radiation, Ionising Radiation, Thermal Metrology, Time and Frequency, Acoustics, Flow, and Valid Analytical Measurements (VAM). The NMS Programmes on Legal Metrology and the National Measurement Partnership have been excluded from the scope, because the NMS Policy Unit (NMSPU) regards it as being inappropriate to extend the study to cover these Programmes at this time.

This report (together with supporting detail made available to NMSPU and its Measurement Advisory Committee (MAC) Software Working Group) and the companion status report on world-wide activity on software support for metrology will inform the first annual review of the SSfM programme and will help in the identification of best-practice to be disseminated, and suitable topics for case studies and feasibility studies. It should also lead to appropriate linkage between the specific NMS Programmes and the SSfM programme.

The SSfM programme is an underpinning programme that provides generic support in the use of software and mathematics to the NMS Programmes for each metrology area. Those NMS Programmes for specific metrology areas provide metrological support to industry. The SSfM Programme in contrast has relatively little direct impact upon industry, although there is some as evidenced by the SSfM Club membership. This relationship is depicted in Figure 1. It is because of this relationship that this Status Report concentrates on the use of software and mathematics in the other NMS Programmes.

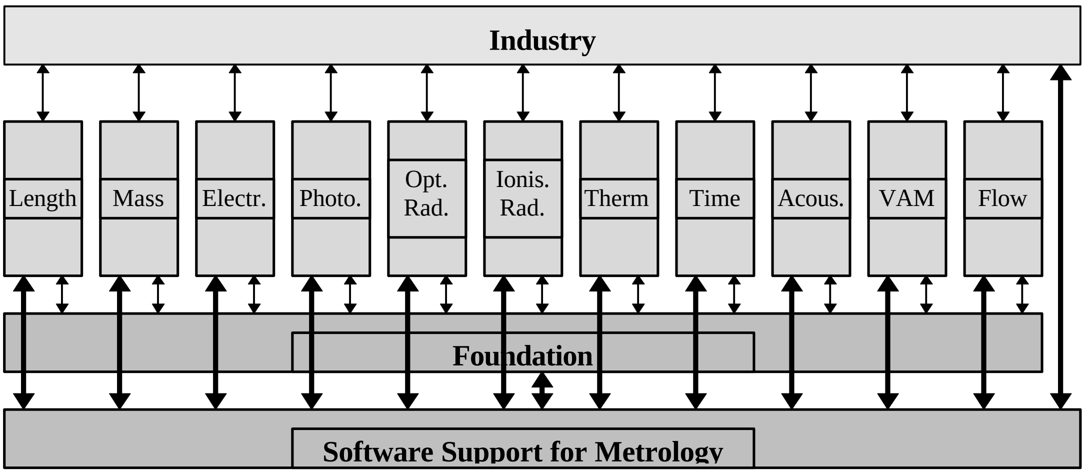

Figure 1. Relationship of SSfM to other NMS Programmes and Industry

In particular, this report addresses each of the themes within the SSfM programme and describes the status concerning the topics covered by each of the relevant projects. It also considers whether there are any important software or mathematics issues in the specific NMS Programmes which are not addressed by the current SSfM programme.

2. Scope of the Areas Covered

2.1 Foundation Programme

The NMS Foundation Programme covers the underpinning research to support the top level infrastructure of the NMS and its ability to provide traceability to internationally agreed primary SI (International System) units of measurement. It aims to provide innovative investigations into those aspects of fundamental physical phenomena which have long term potential impact on the NMS. It sustains a range of expertise on the science of measurement at the highest level of accuracy in order to meet the evolving NMS needs, contributes to the international work on fundamental constants and SI units, and exploits atomic and quantum phenomena and related developments in materials and processes for the next generation of measurement standards.

The Foundation Programme includes work on new primary standards which will move into the mainstream NMS in 3 to 5 years, like the caesium fountain, and more speculative research with a 10 to 15 year lead time, like the quantum roulette noise thermometer and the calculation of spectral transitions in the hydrogen atom.

This is an area in which there are, of course, no measurement services.

The area spans most of the different metrology fields, particularly length, mass, electrical metrology, time and frequency, optical radiation, and thermal metrology. Therefore, in some respects there is a lack of homogeneity, with each project using software and mathematics in ways that are independent of the other projects in the Foundation Programme, but more dependent upon what is appropriate for use within the specific metrology area ("customer" programme) into which the project will feed; on the other hand, the fact that it is all leading edge research gives rise to some commonality in the requirements for use of software and mathematics.

2.2 Length

Length metrology is of direct importance to manufacturing industry. It is also a main building block of the SI metrology system. However, length metrology has the additional feature of dimensionality which is missing from other disciplines. Thus, the Length area encompasses measurement in 1-, 2- and 3-dimensions involving a large number of geometrical properties such as angle measurement, roundness/circularity, cylindricity, and screw and gear metrology. Length metrology also covers a vast range of scales, from X-ray interferometry at approximately 10^{-10} m to surveying at 10^{3} m. (In astronomy, units such as the light-year involve distances of the order of 10^{16} m.). In two or three dimensions, it is not unusual for different axes to involve different scales, e.g. millimetres in one or two dimensions, micrometres in the remaining dimension.

Corresponding to the vast range of measurement activity in the Length area, there are a large number of technologies and instruments used in length metrology, many of which are mature and sophisticated, including x-ray and optical interferometry, atomic force microscopy, co-ordinate measuring machines (CMMs), roundness measuring machines (e.g. Talyrond), surface characterisation (e.g. Talysurf), theodolite and photogrammetric systems.

The current NMS Length Programme is organised into a number of Projects, the main technical projects are listed below. As indicated, these projects cover a wide range of activity and technologies.

The metre and wavelength standards

Frequency stabilisation of gas, diode and solid state lasers, trapped ion standards, measurement of wavelength/optical frequency, calibration of laser interferometric systems

Maintenance and development of dimensional metrology standards

Calibration of end standards, gauge blocks, length bars, step gauges and reference length artefacts, calibration of internal and external diameter standards and reference screw threads, including American Petroleum Institute (API) standards, calibration of angle standards, nanometrology and surface texture measurement, optical profilometry, atomic force and scanning tunnelling microscopy, X-ray interferometry

Micro-and optical metrology

Form and deformation using Moiré and photogrammetry, 1- and 2-dimensional micrometrology, linewidth standards, photomasks and wafer metrology, graticules and metrology gratings, displacement transducers, gas and liquid refractometry, optical fibre metrology, diffracting optics and gratings

Gear metrology

National Gear Metrology Centre, University of Newcastle-upon-Tyne

Nanorond-2 roundness instrument

Collaborative project with Taylor-Hobson Ltd to develop an accurate roundness measuring facility

Small angle generator

Collaborative project with Renishaw plc.

National primary co-ordinate measuring machine

Collaborative project with Brown & Sharpe and Cranfield University to develop two CMMs based on interferometry.

2.3 Mass

The central aim of the NMS Mass Programme is to provide a national measurement infrastructure for mass and related mechanical quantities in support of trade, quality assurance in industry, innovation and a range of important social activities. The first requirement for this is provision of a UK standard for the unit of mass, the kilogram, which is a base unit of the International System (SI) of Units.

The metrology of mass and derived mechanical quantities is a broad and challenging field which extends from routine, but essential, measurement of small masses with balances to a spectrum of innovative activity involving basic science and advanced engineering. It calls for an understanding of surface adsorption on primary mass standards, of molecular dynamics at very low pressures in vacuum standards where gases no longer behave as fluids, and of properties of materials to be used in force transfer standards for heavy engineering applications up to 30 MN (3000 tonnes weight) and in pressure balances used for standardising extreme pressures. Mechanical standards also provide essential support to the standardisation of other quantities including some of those found in dimensional, electrical, thermal, acoustical, optical and ionising radiation metrology.

The programme includes three technical themes that are outlined below.

Standards for mass and density

This theme underpins the whole programme in that it provides the UK's primary standard of mass, copy No 18 of the International Prototype Kilogram in terms of which the SI unit is defined internationally. Provision is made for occasional intercomparisons with the International Prototype held at the BIPM (International Bureau of Weights and Measures) at Sèvres near Paris and for realising a consistent mass scale from 0.5 mg to 2.7 tonnes. In support of this fundamental work, special attention is given to the correction for air buoyancy in weighing, for which a capability in density measurement also needs to be maintained, to surface contamination of weights and the use of weights in vacuum.

Standards for force, torque, hardness and extensometry

The customer base for force calibrations and consultancy is one of the largest with numerous United Kingdom Accreditation Service (UKAS) laboratories and over 300 companies drawn from the mechanical and civil engineering, aerospace, defence, nuclear power, iron and steel, shipping, off-shore and motor vehicle industries. The core of the theme is the provision and maintenance of a suite of dead-weight and hydraulic force machines used to calibrate force transfer standards. The theme also addresses the calibration of extensometers used for tensile testing of materials and standards for the measurement of torque. In addition, the theme aims to re-establish hardness standards in the UK on which there is wide industrial dependence.

Pressure and vacuum standards

Pressure measurement standards are provided in support of UKAS laboratories and a wide range of industries, including the instrumentation, aerospace, defence, nuclear, power, gas, petro-chemical, biological, pharmaceutical, meteorological, environmental, semi-conductor, optical, ventilation, filtration and process industries in general. The standards provided cover absolute, gauge or differential pressures over nearly 17 orders of magnitude, from ultra-high vacua at approximately 10^{-8} Pa to ultra-high pressures at approximately 10^{9} Pa.

2.4 Electrical Metrology

The NMS Electrical programme covers the development, maintenance and dissemination of a wide range of measurement parameters across the electromagnetic spectrum. The programme is divided into four themes:

DC and Low Frequency

This theme provides facilities for the measurement of electrical and magnetic quantities from DC up to around 1 MHz. The theme covers voltage and resistance standards, impedance standards, AC standards, and magnetic standards and current transformer measurements.

RF and Microwave Free Field Metrology

This theme underpins the NMS for the measurement of electromagnetic fields in the Radio Frequency (RF) and Microwave regions. Most activities are in the few megahertz to about 50 GHz range but some fall in the range from around 10 Hz to beyond 100 GHz. The main drivers are telecommunications, aerospace, defence, electromagnetic compatibility (EMC) testing and health and safety. The theme covers antenna metrology development, EMC measurement facilities development, field measurement and mapping, and electromagnetic materials characterisation.

RF and Microwave Guided Wave Metrology

This theme covers RF and Microwave measurement standards in transmission media such as hollow waveguides, coaxial lines, dielectric waveguides and planar lines in high frequency integrated circuits. The four main measurement parameters, power, attenuation, impedance and noise, are highly interdependent, and progress in the development of measurement standards in this field requires a predominantly integrated approach. The main application areas for guided-wave standards are broadly those served by the free field RF and Microwave standards, i.e. telecommunications, aerospace, defence, radar, EMC issues and health and safety, and free-field standards depend strongly on guided-wave standards for traceability.

Ultrafast Measurements

This theme is aimed at providing measurement standards to support traceable measurement in fast, high bandwidth systems, not just in the field of communications and broadcasting but in many areas of electronics and computing where there are demanding requirements. For example, measurement standards are required for the characterisation of electrical pulses, the calibration of high bandwidth oscilloscopes as well as new types of test instrumentation, and the electro-optical characterisation of monolithic microwave integrated circuits (MMICs). The theme also encompasses laser power and energy standards and traceable calibration to support optical communications.

The Electrical programme aims to provide:

- measurement standards, and related techniques and expertise to ensure an effective and timely technical capability at UK national level;
- access to standards through effective technical services made available by the NPL and other NMS contractors to the direct beneficiary community in the UK;
- awareness and availability of the expertise and knowledge generated by NMS-supported work to the beneficiary communities, and measures to ensure responsiveness of the NMS work to evolving need;
- activities to promote best practice and exploitation of the UK's expertise in the relevant measurement fields.

The major route for the transfer of measurement standards from the NMS programmes to industry and other beneficiaries is through the provision of calibration and measurement services. The maintenance of such services is part of the Electrical programme. A major part of maintenance is concerned with the realisation of primary standards and the calibration of secondary and working standards. However, activities also include the comparison of UK standards with those of our trading partners world-wide to assure international agreement of standards.

2.5 Photonics

The NMS Photonics Programme is a newly established programme that brings together all Department of Trade and Industry (DTI) funded measurement activities relating to the photonics industry. The programme includes three technical themes that are outlined below:

Optical Fibres and Components

Work in this area includes:

- Provision of measurement services for optical fibre parameters;
- Development of measurement techniques for distributed measurements in fibres;
- Development of measurement techniques for fibre components and gratings;
- Fibre coating geometry.

Test Equipment, Amplifiers, Sources and Detectors

Work in this area includes:

- Provision of calibration services for fibre optic test equipment including optical time domain reflectometers (OTDRs), power meters and spectrum analysers;
- Development of measurement techniques for fibre optic amplifiers;
- Provision of wavelength standards for wavelength-division multiplexing applications.

Optical Properties of Materials

Work in this area includes:

- Optical properties of displays;
- Refractometry of photonic devices.

2.6 Optical Radiation

The NMS Optical Radiation Programme covers the following technical themes:

a) improved underpinning of primary optical radiation scales, including those for visible light and colour;

b) maintenance and dissemination of optical radiation scales, including those for visible light and colour;

c) development of improved transfer standards;

d) development of new or improved measurement techniques and facilities.

Optical Radiation at NPL involves the following services:

i) measurement of visual and optical properties of materials that transmit, reflect or absorb light;

ii) measurement of visual and optical properties of sources of optical radiation;

iii) measurement of detectors of optical radiation;

iv) calibration and production of certificates;

v) consultancy and advisory services.

Software is used extensively, chiefly for data manipulation, and mathematical problem solving. Typical applications are as follows:

1) interpolation between points using simple mathematical routines and models;
2) least-squares fitting algorithms to determine best fits to data points;
3) mathematical manipulation, weighting and analysis of spectral measurements on sources and materials;
4) simple statistical analysis;
5) uncertainty analysis (involving correlated quantities);
6) modelling of integral transforms for determination of diffraction patterns produced by apertures, ray tracing within cavities, etc.

2.7 Ionising Radiation

The current NMS Ionising Radiation Programme covers the following themes:

a) radiation dosimetry;

b) radioactivity measurements;

c) neutron measurements.

Much of the requirement for the work arises in connection with health and safety in the uses of x-rays, gamma-rays, beta-rays, electrons, neutrons and radionuclides, and also for environmental and health monitoring. As part of this work NPL manages three large accelerator facilities.

The NMS Ionising Radiation Programme carried out at NPL provides the basis for the following services :

1. maintenance of primary standards for:
   a) beta-ray, X-ray and gamma-ray dosimetry at environmental, protection, therapy, sterilisation and industrial processing dose rates;
   b) the standardisation of a wide range of radionuclides and the provision of standardised radioactive source material at environmental, protection and therapy activities;
   c) neutrons with energies from thermal to 20 MeV, primarily for radiation protection purposes;
2. development of new techniques;
3. calibration and measurement services.

Typically, the DTI acts as proxy customer for 1) and 2) whereas 3) is offered directly to customers, both in the UK and overseas.

Software is used extensively in the Ionising Radiation Programme ranging from instrument control, through data manipulation, to mathematical problem solving.

2.8 Thermal Metrology

The NMS Thermal Metrology Programme covers the following themes

a) Temperature;

b) Thermophysical properties;

c) Humidity.

The last of these is shared with the Mass programme.

Thermal metrology at NPL is based on the following services:

i) Maintenance of primary standards;

ii) Development of new techniques;

iii) Calibration and bespoke measurement services.

Typically, the DTI and/or CEU (Commission of the European Union) act as proxy customers for i) and parts of ii); the remaining parts of ii) and iii) are offered directly to customers, in the UK and elsewhere.

Software is used extensively, from data acquisition and instrument control of digital voltmeters (DVMs) and resistance bridges, through data manipulation in spreadsheets and similar packages, to production of certificates (using WordPerfect and Word) to mathematical problem solving, e.g.,

a) data approximation (discrete modelling);

b) finite element analysis (continuous modelling) for problems in heat flow and designing black bodies (in particular, for predicting the performance of a guarded hot plate).

New techniques (e.g., using lasers) for accurate measurement of thermal properties of gases are being developed, which will require extensive software development both at the design (mainly continuous modelling) and implementation (mainly discrete modelling) stages.

In the temperature measurement services area, the main emphasis is on data approximation (discrete modelling), particularly when realising the International Temperature Scale, ITS-90.

Most of this work is fixed-price and routine, but there are exceptions. The thermophysical properties work is more likely to take the form of research investigations and consultancy.

In a sense, humidity straddles both Mass and Thermal programmes, since it requires mass and temperature measurement of gases. The humidity work relates to continued development of instruments such as the standard dew-point generators and the provision of measurement services. Often, this continued development is achieved by building on existing systems, though a recent CEU project is to develop hardware and software for humidity measurement using wavelength measurement, which is a new principle in humidity measurement.

2.9 Time and Frequency

Time and Frequency metrology depends on models both for the physics on which the SI definition of the second is based, and for the propagation of time and frequency signals around the globe. The latter includes models for the atmosphere, models for mapping the Earth, models for satellite orbit determination, models for relativistic effects, models for clock noise, etc. Most of these examples are well understood, although subject to refinement as more precise measurements become available. Note that the clocks and time transfer links are routinely characterised using statistical analysis (Allan variances - 2nd order differences - to characterise frequency stability and 1st order differences for accuracy).

As far as modelling is concerned, the current NMS Time and Frequency Measurement Programme focuses on those aspects which promise immediate improvements for time and frequency metrology. Signal delays through the ionosphere and troposphere are of particular interest for international time comparisons at the moment, while models for characterising Global Positioning System (GPS) disciplined oscillators (GPSDOs) are of prime importance for disseminating standards in the UK. The latter issue is particularly complex because manufacturers of GPSDOs are unwilling to provide any information on the disciplining algorithms used. NPL models are being developed to help identify tests for faults occurring above well defined uncertainty levels.

At the present time, error bars relating to Time and Frequency data are not published, because of difficulties in estimating the uncertainties (particularly the non-stochastic Type B kind). This represents a significant challenge in view of the imminent publication of guideline documents and the requirement for UKAS accreditation.

2.10 Acoustics

The NMS Acoustics Programme includes air-borne sound and noise, and underwater and medical acoustics. The technology of acoustic sensors (generally comprising, in their simplest form, sensitive piezoelectric transducers coupled with diaphragms which serve to convert acoustic energy to electrical energy) is still under development. Therefore, although microphones and piezoelectric transducers have been in existence for many decades, devices such as fibre-optic hydrophones are recent developments. In addition, laser interferometry to measure acoustic variables is less mature (NPL uses lasers to measure acoustic particle displacement and acoustic particle velocity) and some of the issues which arise here, such as the nature of the acousto-optic interaction in nonlinear fields, are not well-understood.

Another sensor used in acoustics is the thermocouple, to measure temperature rises in tissue mimics due to ultrasound fields. NPL has been active in modelling these effects. For example, finite element analysis has been employed to model the errors which arise in thermal measurements from heat conduction by thermocouple wires.

Much of the work done on these latter topics (acoustic particle velocity laser and thermocouple measurements) is not directly for the DTI, but is covered by NPL's strategic research programme or third-party contracts.

Generally, acoustic sensors are built on the principle of energy or momentum transfer, e.g., for measuring acoustic power in ultrasonics the transfer of momentum from the acoustic wave to a large target (either reflecting or absorbing) is the manner in which the primary standard for acoustic power is realised.

In summary, the challenges lie in

a) improving the existing measurement capability (increasing accuracy and reducing uncertainty),

b) extending the capability (e.g., to new devices and/or frequency ranges), and

c) delivering customised solutions to industry (particularly in respect of ultrasonics).

In the Noise field, a further consideration is the performance of the human auditory system, whose response is non-uniform, e.g., in frequency and level. More importantly, Noise Section deals with psychoacoustics, i.e., human reaction to sound and hence the measurand is indirect: the human perception of “annoyance” rather than direct measurements of sound pressure.

2.11 Flow Measurement

Within the NMS Flow Programme, the National Engineering Laboratory (NEL) provides the UK National Standard for Flow Measurement. It provides some of the most comprehensive flow measurement facilities in the world and provides expertise to industry and other parts of the NMS. It also has contacts with Government laboratories throughout the world through personal contacts, Standards committees and EUROMET. By its nature flow measurement is a secondary standard, all measurements being traceable to primary standards of length, mass and time. Unlike many areas of metrology the size and complexity of flow measurement facilities are significant. Hence the cost of calibration, maintenance etc., is proportionally large.

Flow is a dynamic measurement. The interaction of the flowing medium with the artefact measuring it, and the variations of the flowing fluid through viscosity and density changes make the measurement relatively imprecise compared with mass and length. Whereas uncertainties in parts per million are possible in other areas of metrology, the best flow measurement uncertainty will be around 0.05%. Tests of a flow measurement device have to take place across a range of flowrates and conditions and a twenty-flowrate test may take up to a day to accomplish.

The software used can be divided into a number of distinct areas: firstly the SCADA software designed to control and record data from a test; secondly the processing software designed to take this data, process it and produce reports; thirdly the statistical software to analyse the reduced data, produce curve fits and uncertainty analyses.

In support of these processes lies another disparate group of software products: the mathematical modelling software used to predict flowing fluid behaviour and hence better explain and predict the performance of flow measurement devices (Computational Fluid Dynamics (CFD)); the equations and algorithms used to predict the physical properties of the fluids being measured; lastly the embedded software in commercial instruments and flow computers which must be understood, evaluated and possibly verified for industrial use.

The Flow aspects of this report have been compiled from a knowledge of NEL practice, practices in other UKAS accredited laboratories and from information from industry. EUROMET partners were also consulted and six countries responded with information of their practice.

2.12 Valid Analytical Measurements

Analytical measurements are those made in the pursuit of analytical science; they are obtained using a broad range of techniques and are spread across several disciplines. The term *Valid Analytical Measurements* was originally applied to the field of analytical chemistry and its sub-disciplines such as clinical chemistry. However, it should be noted that the boundaries of this discipline are diffuse and frequent use is made of techniques from other disciplines, for example, the biochemical and physical sciences. Although analytical chemistry has had the longest exposure to the concept of Valid Analytical Measurements, the VAM area in this report covers the whole scope of the current NMS VAM Programme, including both analytical chemistry and the physical analysis of gases and surfaces.

Analytical chemistry is that branch of applied chemistry concerned with investigating the composition of materials. It comprises, traditionally, two main strands: qualitative analysis and quantitative analysis. Qualitative analysis consists of activities designed to reveal the identity of a material or to show that a stated substance is a component of a given material. It often involves making quantitative measurements in order to establish whether the value of a given property is above or below a reference level. Quantitative analysis consists of activities designed to determine the amount of a particular component in a given material or the magnitude of a physical property.

In order to quantify the amount of a given substance it is necessary to measure an appropriate property of the substance and relate this, via a calibration, to a similar measurement on a known amount of the same substance. For the purposes of this report measurements are considered to be made on chemical, biochemical or physical properties of substances.

The physical analysis of gases also has quantitative and qualitative aspects. Quantitative analysis is based upon the use of primary gas standards, prepared using absolute gravimetric techniques so that concentrations are traceable to the primary standard of mass. In addition, there is quantitative analysis of particulates in gases. Qualitative analysis concerns the analysis of odours using novel techniques with the potential, eventually, to identify an odour measurement scale.

The physical analysis of surfaces of solids is a relatively new subject for which reference materials, as the basis for calibration of measurement instruments, are largely inappropriate. Instead, it is necessary to work from first principles in analysing the measurement process as a whole.

The software, and where appropriate the mathematics, employed in the VAM area falls into three main categories of use:

- data acquisition - including instrument control;
- data processing - including modelling, fitting, optimization, and evaluation of uncertainties;
- report generation.

Application areas covered by analytical chemistry include, but are not limited to, the following:

- forensic science;
- food and nutrition;
- environmental monitoring;
- health & safety;
- patent protection and litigation;
- process control;
- R&D support.

This report does not cover analytical chemistry R&D *per se* and hence excludes such topics as molecular modelling, computational (theoretical) chemistry and structure/activity relationship modelling. These areas make intensive use of complicated mathematical models and sophisticated software. The issues here are extremely diverse and beyond the scope of the present report.

3. Modelling Techniques

3.1 Methods for Modelling Measurement Data

3.1.1 Types of data

There are three categories of measurement data sets:

1. discrete measurement data, which means:
   a) data obtained by sampling a discrete variable: e.g., a variable that may take only integer values such as in counting processes, and
   b) data obtained by sampling a continuous variable;
2. continuous measurement signal, which means an analogue signal prior to any analogue to digital conversion, which would result in data of type 1 b);
3. hybrid measurement data composed of both discrete and continuous data sets.

3.1.2 Types of model

There are three categories of model:

1. discrete models in which the outputs of the measurement system are related to measurements of its inputs by a system of algebraic equations;
2. continuous models in which the outputs of the measurement system are related to measurements of its inputs by a system of differential equations;
3. hybrid models in which the outputs of the measurement system are related to measurements of its inputs by a system composed of both algebraic and differential equations.

It is appropriate to distinguish here between the nature of the model and the approach to solving the model (which for a continuous model typically also involves solving algebraic equations).

Given infinite precision arithmetic, the set of equations derived from a discrete model can be solved exactly. This is not the case for a continuous model which typically requires approximations to be made, for example, in terms of domain discretisation.

3.1.3 Models used in metrology

Most quantitative measurement data relates to essentially continuous variables. However, such data are often collected and recorded as discrete data by digital sampling, so that the measurement data that are subsequently processed consists of sets of discrete values. Most measurement data are real-valued, but some are complex-valued. Some measurement data are collected for a single variable (univariate data) whilst other data are collected for a set of variables (multivariate data).

Discrete modelling is used in all areas of metrology, whereas the use of continuous modelling is concentrated mostly in flow, thermal and electrical metrology, and hybrid modelling is currently applied mainly to specific problems in length metrology.

The models range widely in complexity, some having a simple relationship between input and output, whereas others need pages of equations to describe the system. Some have only a small number of parameters, whereas others have a large number of parameters with complicated interrelationships. Sometimes the significant measurement errors are limited to just a few measured quantities, whereas in other cases many variables are subject to significant measurement error. Often, measurement errors are consistent with a Gaussian distribution, although there is an important minority of cases in which non-Gaussian behaviour occurs, with skewed distributions, outliers etc. Usually the input quantities are well controlled and well understood, although in some cases inputs have to be estimated.

The maturity of modelling varies across the metrology areas. In some areas, the physical theory is well understood and accepted, whereas in other areas this is not so and there may be competing models and model development is still a research topic. Sometimes the theory is comprehensive, taking into account all main influence factors, whereas in other cases approximations are used and/or terms are omitted to simplify the model.

Measurement uncertainty may be modelled for all input quantities, or modelled carefully for only a few input quantities, with other quantities being assumed to be exact or being assigned uncertainty contributions based on experience. It is often the case that correlations among input quantities are ignored, although in a number of areas it is recognised that more work needs to be done to take account of such correlations.

For the more straightforward models, the solution methods usually solve the computational problem posed by the model, but in other situations approximations are used, with some terms being simplified, linearised or even ignored. The solution method may take into account all the sources of measurement error, variances and correlations amongst the input quantities, but often some simplification is made, particularly treating data as uncorrelated.

Despite approximations and simplifications made, the models and methods of solution used are very often effective, giving solutions consistent with those predicted by theory or established by other methods. The solution methods usually produce realistic estimates, but in some cases the methods do not converge to a solution or the solutions are not stable — small changes to the inputs can sometimes produce relatively large changes in the solution.

3.1.3.1 Foundation Programme

There is a wide range of modelling requirements in the Foundation Programme. In many cases the need is to model the basic physics of various processes, and this tends to imply the use of continuous modelling. Nevertheless, there are examples of discrete modelling, and possibly also some hybrid modelling.

3.1.3.2 Length

In the Length area, a wide variety of discrete models is used. In addition, there is some use of both continuous and hybrid modelling.

Because of the rich interplay between Length metrology (and co-ordinate metrology in particular) and geometry, many of the modelling issues that arise in Length are naturally the subject of past and current applied mathematics research. There is a long history of collaboration between co-ordinate metrologists and numerical mathematicians at NPL on modelling and software development. However there are areas where further work on modelling and algorithm development is required or would be of potential benefit. These include:

Geometric tolerance assessment. There are a large number of unresolved modelling problems associated with geometric tolerance assessment:

- Development of algorithms that successfully take into account form and measurement error. Chebyshev methods take into account form error only, least squares methods usually only take measurement error into account.
- Development of algorithms that account for the correlation in the co-ordinate estimates that arises in laser tracking, theodolite and photogrammetry systems.
- Data approximation using parametric curves and surfaces, important in tolerance assessment using Computer Aided Design (CAD) models.
- Robust (outlier resistant) data approximation methods.

Error separation techniques. In the calibration of reference artefacts it is important to separate the systematic error behaviour of the measuring instrument from the deviation of the artefact from its nominal geometric form. In many circumstances, measurement strategies can be determined to isolate both sets of errors. Further work could be done on developing the general principles and in their application, for instance, to determining the form error in the reference spherical wave front in interferometry.

Empirical models in higher dimensions. While polynomials, splines, etc., can be used efficiently and effectively to model empirical functions in one dimension, they are less well adapted to two and higher dimensions. For example, the refractive index of air can be represented by a real valued function of three spatial co-ordinates. It may be that technology based on radial basis functions, when sufficiently developed, will provide a better approach.

Image analysis. In the optical measurement of gratings, pin holes, etc., the system response depends in a complicated way on the optical and geometrical properties of the sample. This means that determining unbiased estimates and their uncertainties is difficult.

3.1.3.3 Mass

Most quantitative data in the Mass area relate to essentially continuous variables. Examples include air temperature, pressure and humidity. However, such data are often collected and recorded as discrete data by digital sampling, so that the measurement data that are subsequently processed consist of sets of discrete values.

It is noted that pressure presents some particular problems. Pressures cannot be added, subtracted, multiplied or divided in the way that lengths and masses can. Consequently, it is necessary to realise pressure independently over its entire range which amounts to of the order of seventeen decades. This situation contrasts with length and mass for which it is possible, in principle, to start with a well defined reference standard and generate multiples and sub-multiples.

Both *empirical* and *physical* models are used to describe the measurement systems and data arising in the Mass area. Furthermore, there is considerable variation in the complexity of these models. Most of the modelling is discrete but there are also examples of continuous modelling in which finite difference and finite element techniques are applied to the solution of systems of differential equations.

Linear regression, as in the fitting of linear, polynomial and spline models to data with error in the response variable only, is widely used. There is awareness in the Mass area that there exist a number of numerical schemes for solving these regression problems, but general guidance concerning the choice of an appropriate scheme is required.

In the Mass Programme there is on-going debate with other National Measurement Institutes (NMIs) concerning the precise form of some of the modelling undertaken. In the Mass and Density area, there are differences in the way that prior information about calibration standards is incorporated into the analysis of data from weighing schemes, and the way this information affects uncertainty estimation. However, because the differences in the results obtained from the various approaches are not significant, this is not regarded as presenting a problem, but instead serves to provide validation of the methods used.

In the Pressure and Vacuum area, there are fundamental differences in the approaches taken by NMIs to processing the data arising from the calibration of pressure balances using the cross-floating method. Work has been undertaken to understand the consequences of these differences, and to show the greater robustness of the NPL approach. Recently, new approaches have been investigated that attempt to account properly for the errors in the measured quantities, and to feed this information into the estimation of uncertainties. However, these approaches lead to difficult optimisation problems requiring more sophisticated numerical software. It is anticipated that detailed validation of the methods will be required before these approaches are generally accepted.

3.1.3.4 Electrical Metrology

Most quantitative data in the Electrical area relate to essentially continuous variables. Examples include voltage, current and power, which vary continuously with time. However, such data are often collected and recorded as discrete data by digital sampling. Furthermore, the recorded output of instruments used in the Electrical area, such as DVMs, may result from the processing of raw measurement data. For example, digital filtering may be applied, so that the value output by the instrument is the arithmetic mean of data collected over a short time period.

Measurement data arising in the Electrical area may be

- *real-valued*, as in the measurement of time and voltage, or *complex-valued*, as in the measurement of the magnitude and phase of a reflection coefficient, and
- *univariate*, as in the collection of repeated values of a quantity, or *multivariate*, as in the collection of values of control (independent) and response (dependent) variables in calibration.

Both *empirical* and *physical* models are used to describe the measurement systems and data arising in this area. Both discrete and continuous modelling are used. Linear regression, as in the fitting of linear, polynomial and polynomial spline models to data with error in the response variable only, is widely used and well understood, although in the DC and low frequency theme, linear regression where there are several independent variables is regarded as outside the scope of many commercial packages. Challenging regression problems arise, however, where there are

- errors in more than one variable,
- nonlinear models, and
- complex-valued quantities.

For example, in the Ultrafast measurement theme, the presence of jitter can lead to errors in the recorded values of time, and it is necessary to account for this both in the data modelling as well as in the associated uncertainty estimation. Furthermore, the calibration of automated network analysers in the RF and microwave guided wave theme, leads to regression problems involving nonlinear models, complex-valued quantities and errors in more than one variable. Similarly, such problems arise when analysing electro-optical measurements of monolithic microwave integrated circuits in the Ultrafast theme.

Deconvolution of digital signals is also seen to be a challenging problem, particularly with regard to providing estimates of associated uncertainties.

3.1.3.5 Photonics

Wavelength is usually considered to be a continuous property, in the sense that it varies across the spectrum. However, wavelength data are collected and recorded as discrete data by digital sampling, so that the measurement data that are subsequently processed will consist of discrete values. Another example of this type of consideration is the measurement of the geometric parameters of fibres which can also be considered as continuous, but which are often recorded as discrete quantities with interpolation between points.

Both discrete and continuous modelling are used with some success, in certain well-defined areas of Photonics, specifically where the physics of the underlying models is well-understood. The success may be hampered if the mathematics of the underlying physical model is highly complicated and/or non-linear.

There has also been success in understanding processes using models where the physics has not been understood initially. This has been achieved by the development and validation of empirical models which may then feed into improved knowledge of the physical model. For example, this has been seen in the determination of field distributions in fibres under bending conditions, the results of which have a direct impact on measurements of certain fibre parameters.

3.1.3.6 Optical Radiation

In the Optical Radiation area, as in Photonics, many of the measurements made are expressed in terms of values at discrete wavelength points. This is again a convenient way of expressing what are actually continuous (but not necessarily smoothly or predictably varying) quantities which vary with wavelength across the spectrum. Sometimes, when spectral techniques are used, the raw data is an analogue signal. This analogue signal may be used directly as a continuous signal or may be fed into an A/D converter, resulting in large amounts of discrete data. In addition, occasionally, hybrid data may be needed for some applications.

Both discrete and continuous modelling have been used with some success, particularly in areas where the physics is well understood.

Where there is a need to interpolate between data points, this is frequently done empirically, rather than being based upon a validated physical model. There is concern that as systematic errors are reduced, problems with interpolation will become significant and a better approach towards interpolation will be required.

In the past, the measurement results have tended to fit smooth distributions, but in future there will be more irregular distributions (e.g. from fluorescent tubes) and there is likely to be more need to fit the more flexible spline functions to the data.

There is a serious R&D problem regarding the uncertainties in some types of Optical Radiation measurement (e.g. colour measurements). The issue is how to integrate an array of spectral data with a weighting function. Three numbers are obtained, for the three primary colours, rather than many spectral points, but an understanding of the relationships among these numbers requires an understanding of how to model the responses of the human eye.

3.1.3.7 Ionising Radiation

Typically, data in Ionising Radiation arises from observations of continuous phenomena and hence continuous modelling is important, using Monte Carlo simulations and finite element and finite difference methods. There is also some use of discrete modelling, for example for data approximation.

3.1.3.8 Thermal Metrology

Both heat and temperature (and derived properties) are continuous quantities, since they are usually functions of space and time. The design of novel measurement systems, whether generic (e.g., to measure the temperature of hot gases) or specific (e.g., to measure a given thermophysical property of an artefact, material or system) requires continuous modelling. Typically, the metrologist is faced with time-dependent differential equations, e.g., those governing heat flow and/or electromagnetism (for laser systems) and/or fluid mechanics (for liquids and gases). In practice, the data obtained will be discrete (sampled at fixed intervals in time or space), and approximation problems need to be solved (discrete modelling). Discrete modelling is often required at the design stage to take account of the effects of sampling (discretisation), and to assist in the estimation of the uncertainties of output quantities. There is scope for hybrid modelling if data fusion techniques are used to combine discrete (measured) data and continuous (simulated) data.

The techniques used in continuous modelling include finite element and finite difference methods. Particular problems include the design of apparatus, where the metrologist is required to model the steady-state response (typically governed by a Poisson partial differential equation) and transient responses (often with nonlinear boundary conditions) in problems involving diffusion and hot wires.

The techniques used in discrete modelling include Monte Carlo simulation and data approximation.

Recently, there has been considerable growth of interest in multi-property measurement. This type of measurement is characterised by augmenting continuous modelling of the governing differential equations (e.g., heat conduction) with filtering (discrete modelling) of measurement data, with optimisation (effectively a discrete process). The drivers for this development of interest include:

i) potential cost savings;
ii) improved uncertainty estimation, since it may be possible to take greater account of correlated input quantities.

The success of multi-property measurement will depend on

a) building, solving and validating comprehensive models of the “signal”;
b) building, solving and validating comprehensive statistical models of the “noise”, if the uncertainty estimates are to be reliable and sharp;
c) visual modelling, to support a) and b);
d) data fusion techniques to combine all the data obtained (discrete and continuous) into a coherent whole.

Modelling has had considerable success, to the extent that many temperature and thermophysical property measurements (partly those which are steady-state and have simple geometries) are now relatively routine, and the remaining issues relate to effective implementation of the software solutions.

For more difficult problems, further modelling effort is needed. Their numerical solution commands more resources and raises the question of which package to use. Easy to use packages are insufficiently flexible, while comprehensive packages are more expensive and require more training.

In contrast, in Humidity work, the basic model is often quite simple, often laid down in the relevant standard. Hence calibration curve fitting is not normally required and there is little need to develop models further; they are generally used to evaluate results from given measured inputs. Some error correction is implemented in software, e.g., relating to the dew-point temperature.

3.1.3.9 Time and Frequency

There is a modest but growing requirement for modelling in the Time and Frequency area. In the timescale comparison work based on GPS, a model has been developed to aid in the understanding of the physical processes (such as the influence of humidity), but more work is needed on the correlations related to the tropospheric effects. In GPS disciplined oscillator work, the emphasis of modelling is on statistical analysis (Allan variances - 2nd order differences - to characterise the GPS disciplined oscillator frequency stability and 1st order differences for accuracy). At some time in the future a service to calibrate and characterise disciplined oscillators is planned and consequentially some additional modelling work will be needed.

Uncertainties relating to Time and Frequency data are not published, because of difficulties in estimating their values (particularly the non-stochastic Type B evaluation). This difficulty represents a significant challenge in view of the imminent publication of guideline documents and the requirement for UKAS accreditation.

3.1.3.10 Acoustics

The use of modelling in the Acoustics area is advanced and effective. Both discrete (fast Fourier transforms and regression methods) and continuous (finite element analysis) modelling are performed.

The goal of the signal analysis is to choose suitable drive (input) signals and determine transfer functions. The signal processing includes optimal (Wiener) filtering, and hence the analysis of errors which exhibit structure. Such filter methods employ convolution techniques. In general, both time and frequency domain analyses are used in processing the signals.

NPL is investigating novel techniques, e.g., pulse methods in underwater acoustics, that overcome some of the difficulties experienced with long duration signals when there is interference between the source and waves returning from the walls of the tank. Other new approaches include the use of wavelet decomposition methods in addition to traditional Fourier methods. The former are more appropriate for nonstationary signals (e.g., where the frequency components are time-dependent).

Precision requirements are high, particularly because of health and safety requirements in the use of scanning equipment. These requirements are expressed in relevant standards. For example, IEC 61157 [9] (Medical Ultrasound) requires corrections based on curve fitting to parts of the waveform.

Mathematical modelling is also important in the Ultrasound area. Ultrasonics in water generally involves high frequencies in linear and nonlinear acoustics propagation algorithms and hence fine discretisation in the time domain in finite element simulations. There are performance issues and trade-offs have to be made, the results of which need to be validated.

Linear acoustic theory is generally well-understood following the pioneering work of Rayleigh. Analysis of acoustic systems (as met in practice) can be difficult, and hence there can be discrepancies i) between theory and practice, and ii) between practitioners. NPL has started to look at acoustic cavitation phenomena and is continuing to study nonlinear behaviour (wave propagation, interactions, etc.) in general. NPL's work on nonlinear ultrasonics began in the early 1980s and has progressed to the stage that it underpins the existing UKAS-accredited hydrophone calibration service.

The main areas of concern, and hence research, in modelling terms relate to

1) modelling the entire system;
2) removal of assumptions in continuous models, e.g., to include nonlinear wave phenomena;
3) extracting more information from signals, by novel spectrometry and wavelet basis techniques;
4) statistical modelling, particularly of correlated and/or complex-valued quantities.

Generally, the more speculative the model, the less comprehensive it is. However, using model validation techniques, it is possible to identify the key features of the model. This is the subject of on-going research in the Acoustics Programme.

3.1.3.11 Flow Measurement

In the measurement of flow all models describing the calibration and assessment processes are based on discrete data. Measurement of time, mass, volume and meter output are all discrete data sets. This assumes that analogue readings from meters or instruments are classified as discrete data sets. This is a valid assumption as any analogue signal from a measurement transducer or transmitter is inherently averaged over a time period and the average result usually defines a discrete value in the subsequent model. The models used to calculate the end result from these discrete data are all inherently simple.

The definition of algorithms to define the volume correction factors for fluids will require the application of appropriate software procedures. Many equations of state are complicated continuous functions defining phase boundaries and density at variable conditions. The inputs of temperature and pressure may be discrete but the underlying models are continuous and complicated. The provision and verification of these equations of state and their application in a non-specialist world needs to be addressed.

Continuous data arises in the research and modelling area of measurement. CFD uses the equations governing the behaviour of fluids to calculate velocity vectors and local pressure conditions within a flowing stream. CFD is used extensively to predict the effect of changing flow conditions on measurement devices. It is in this area that continuous modelling techniques are used successfully to model and predict performance of measuring devices and to predict changes in performance in areas where testing would be impossible. By contrast, the industrial flow measurement community is primarily interested in the measurement of flow volume for which CFD is little used.

As continuous modelling is by nature a process by which a picture of the flow characteristics and interaction with the flow devices or containment is produced, little application in real time is possible. It is also clear that although the processes are becoming well understood, it is still vital to test models against experimental measurements. Currently, levels of agreement are still poor relative to experimental calibration of flow devices. The techniques are however very useful in predicting and explaining trends and shifts in flow device performance and to indicate levels of performance change. This allows extrapolation of experimental data and significant cost reduction compared with experiment where facilities may not exist for the conditions of interest.

Explicit hybrids of discrete and continuous data types are almost non-existent in flow measurement, especially in the standards and calibration area. Within industry the advent of intelligent flowmeters where the performance is dependent on a microprocessor measuring a flow dependent parameter and then calculating and correcting this measurement to derive an accurate output of flow or quantity across variable conditions may include hybrid algorithms.

Other simple models are sometimes used to predict flowmeter behaviour over a range of temperatures, viscosity and pressures. In general these are simple models based on experimental data and apply only to the meter tested. Owing to the extreme sensitivity of most flowmeter performance to fine manufacturing tolerance, even the most sophisticated models may be deficient in fully accounting for generic behaviour. Performance curves of many meters vary in magnitude, slope and shape depending on the fluid. Consequently, extrapolation between the calibration curves at different conditions is very difficult to carry out accurately, and assumptions are hence nearly always made. There is a requirement for mathematical modelling to provide easy-to-use tools to determine these multivariate calibration surfaces.

The most common technique beyond straight-line regression is curve fitting by polynomials or splines; the former is most often used in the UK but other European laboratories frequently employ splines. Unfortunately the nature of most flowmeter performance curves does not lend itself to polynomial curve fits and extrapolation is very dangerous. Sometimes different curve fits are applied over different parts of the characteristic to follow the data better but this must be considered an undesirable practice.

3.1.3.12 Valid Analytical Measurements

In the VAM area, although some of the quantities for which data are collected are continuous variables (such as temperature and time), the data that are collected are nearly always discrete. Sometimes the raw data are discrete. However, where spectral techniques are used, the raw data are analogue and hence continuous, but this analogue data are fed into an A/D converter, resulting in large amounts of discrete data. In addition, occasionally, hybrid data may arise in some applications.

Discrete data can usefully be categorised further into univariate (or scalar) data and multivariate, array or vector data. In some cases, analogue data are collected and used without A/D conversion as continuous data.

Hybrid data are not commonly employed in analytical measurements. Some interpretative modelling may use combinations of discrete and continuous data, but in most cases the continuous data will be converted to discrete. In exceptional cases, classifications may be based on hybrid data; for example, in categorising the country of origin of drugs of abuse, the presence or absence of a particular “cutting agent” may be combined with the concentration values of drug metabolites to classify samples.

It is useful to distinguish between models used to obtain values from raw data, models formed for calibration and prediction of measurement values, models used in interpretation, and measurement system models used for uncertainty estimation:

- **Obtaining values from data.** In many cases, simple procedures are used which require no explicit modelling. Nevertheless, elements of model building or model validation are implicit or assumed. For example, selecting a local maximum from a spectrum as a measure of peak location or intensity implies a model based on peaks with defined maxima. Such simple implied models are typically validated by experiment, for example by adjusting sampling interval and observing the effect.

  In a few cases, including spectroscopic integration in the presence of overlapping peaks, curve fitting may be used to obtain reliable integrated intensities. Analysts are more likely to avoid the necessity of curve fitting by choosing a system in which overlap is small. Where fitting to a model is done, it usually follows least squares fitting based on well established iterative numerical procedures. It is almost always possible to find an appropriate and sufficiently robust algorithm for the applications.

- **Calibration models.** The majority of calibration models in the VAM area are linear relationships between scalar input and output quantities. Where possible, measurement systems are chosen or developed to give linear calibration models. There are exceptions, however. Immunoassay techniques, for example, produce sigmoidal response curves which are fitted by physical (e.g., kinetic) or empirical (e.g., log-logistic) models. Because these models are non-linear in their parameters, iterative methods (see below) are used to obtain the fit. Comparatively large statistical variability in the responses, together with exponential terms in the underlying model, can result in instability. Another class of calibration model is multivariate calibration, used increasingly in practical field measurements but not generally in standards calibration. Multivariate calibration establishes relationships between vector or multivariate observed data (usually spectra) and one or more properties of interest.

- **Interpretation.** Modelling may be used to assist in interpretation of data, for example to aid classification. Classification is common in regulatory and other applications of analysis, and may require construction of classification models based on multivariate data.

- **Measurement process modelling.** In the VAM programme work on surface analysis, the complete measurement process is being modelled; these models are very much the subject of current research and development. In all other areas of the VAM programme, measurement processes are not generally modelled prior to routine development, except in that the basic relationship required to calculate the result from input parameters is written down. The approach to development is almost invariably practical and based on experience rather than on prior modelling. There are well developed and applicable theories supporting the physical methods used, but these are not generally useful in practical situations except to indicate limitations of a given technique. These theories are inappropriate largely because the major influences on analytical measurement — effect of sample matrix, operator, interpretation of methods etc. — are not readily amenable to predictive mathematical modelling. There is a resulting profound impact on studies of reliability and uncertainty. While good progress has been made in interpreting the ISO Guide to the expression of Uncertainty in Measurement (GUM) [1] for chemistry via a EURACHEM document [5], measurement process models in the VAM area can usefully be based on measurable input quantities only in developing quite specific methods for individual test materials and determinands; typically, this applies only in reference measurements, and very rarely to routine measurement.

3.2 Uncertainties and Statistical Modelling

The estimation of uncertainty is important in all areas of metrology.

In the Foundation Programme, apart from when uncertainties are quoted with published results in scientific papers, the main reason for given attention to uncertainties is the need to minimise systematic errors (type B evaluation of uncertainty) in measuring fundamental constants.

In the Length Metrology area, there are uncertainty problems which have only recently begun to be tackled. One concerns how to assess the uncertainties associated with a given choice of probing strategy. Another concerns the proper handling of correlations between estimates of target co-ordinates in connection with multi-station measuring systems, such as theodolite, photogrammetry, and laser tracking systems.

In the Mass area, there are problems with the robustness of the calculation of an uncertainty for the estimate of the distortion coefficient for a pressure balance in the high-pressure range.

In the Electrical Metrology area, there is a particular need to handle measurements which are complex-valued. There are difficulties in quantifying type B contributions and a tendency to be pessimistic when estimating these.

There are some legal issues related to uncertainties in the Optical Radiation area, where the measurements have safety and cost implications.

In the Ionising Radiation area, there is a problem with spectrum unfolding involving a covariance matrix of order 50 which is underdetermined because of an appreciable number of perfect correlations.

In the Thermal Metrology area, where signal processing is concerned, the type A and B classification is considered unhelpful, but it can be reinterpreted as relating to stationary noise and non-stationary noise, respectively.

In the Time and Frequency area, NPL provides the NPL Truetime service but that is not a measurement service requiring uncertainty estimation. NPL does not currently offer any proper measurement service in the Time and Frequency area, but in the future there are expected to be clock characterisation services. It will then become necessary to reconcile the GUM with the body of work describing how to handle uncertainties in the Allan Variance (used to characterise frequency instabilities) for various noise types.

In Acoustics, there is a major concern related to statistical tolerancing, particularly regarding the uncertainties of meeting a specification. There is relevant OIML guidance, but it advocates the use of four possible results (passes, probably passes, probably does not pass, does not pass) which is difficult for users to understand.

In the Flow Measurement area, in the UK, there is agreement on expressing the uncertainty of the quantity of fluid passing through the meter under test as the prime factor given. Further uncertainties are then given for any secondary instruments measuring the flowmeter output (e.g. pressure, pulse counts, milliamps etc.). Having achieved this consistency in the expression of uncertainty, the flowmeter calibration laboratories are now about to embark on the next stage of how, consistently, to give the uncertainty on the measured parameter of the flowmeter. This is not a trivial task. Consistency between laboratories is the first objective to ensure the customers can understand and compare uncertainties. Secondly, the problems of how to express uncertainty from the different numbers of test points, the different flow rates at which repeat points are taken, and the weaknesses of curve fits and expressing uncertainty of curve fits have all to be addressed. If this change in presentation is achieved, an estimate of the uncertainty the client can expect by taking the meter from the laboratory to his process may be possible. This will be the future approach in the UK.

In the VAM area, there is considerable knowledge and experience of estimating measurement uncertainties in the gas analysis field; whereas in other fields, the production of measurement uncertainty information is a current area of concern for those developing new analytical methodologies. Apart from in the gas analysis field, measurement uncertainty is a relatively new area for analysts and so its use in accordance with the GUM is not yet widespread. Analytical laboratories typically operate many different methods (frequently hundreds) and determining uncertainty estimates for results from these methods, based on ISO principles, can be extremely time consuming. In addition, it is commonly found that the main influences on analytical measurements are not associated with traceable measured input quantities. In the VAM area, reliability has historically been established through experimental studies of method performance ("method validation") with no formal intent to estimate a single uncertainty figure. Indeed, the range of method performance parameters determined during method validation studies is substantially more informative to a trained analyst. This kind of study still represents the mainstay of reliability assessment in analytical work, and that is unlikely to change quickly. Nonetheless, the recent emergence of measurement uncertainty as a key requirement of accreditation bodies has generated substantial effort to apply the concept in the VAM area.

3.2.1 Derivation of statistical models

In some areas the emphasis is more on systematic errors (type B evaluation of uncertainty) than on statistical modelling (type A). This is particularly true of the Foundation Programme.

Measurement uncertainty estimation usually follows the principles of the GUM [1]. Deficiencies have been identified in the following cases:

- dealing correctly with numbers of degrees of freedom;
- calibration problems where there are errors in more than one variable;
- application to cases where the measurement system model is complicated, e.g. where it is not possible to express the output(s) of the measurement system as an explicit function of its inputs;
- where it is necessary to use the standard deviation statistic as well as the mean statistic;
- where measurements are complex-valued;
- the estimation of uncertainties associated with deconvolution operations;
- where measurements are highly correlated;
- where distributions are non-Gaussian, including asymmetric ones;
- where the presence of outliers indicates the appropriateness of using robust statistics, such as trimmed means and median estimators;
- when uncertainty intervals are to be expressed near to fundamental limits;
- statistical analysis of data arising from interlaboratory comparisons;
- when Monte Carlo simulations are used;
- when continuous modelling is used (e.g. in Thermal Metrology or Flow Measurement).

Measurement accuracy varies from area to area. In the Optical Radiation area it is typically 1% or smaller. In Acoustics, it can be several percent for the primary standard in underwater acoustics but an order of magnitude better in air acoustics. Generally, the uncertainties determined within NMS Programmes are comparable to those quoted by other national standards laboratories. However, there are some problems with international comparisons where the uncertainty intervals quoted by participating laboratories are markedly inconsistent in some cases. Advice is needed on how to determine reference values and accompanying uncertainties for such comparisons which account for these discrepancies.

In some areas, there is no rigorous treatment of the errors in measurements and the possible correlation that might exist. There may, as a result, be an overestimate or an underestimate of the uncertainties in such areas. A more rigorous treatment of correlations would provide more reliable uncertainties.

There is also some recognition of a need for more robust statistical treatment.

3.2.2 Determination of the statistical distribution of measurement results

In most areas, the statistical distributions of measurement results are generally considered to be Gaussian, apart from a few cases where a uniform distribution is assumed. These assumptions, which are consistent with the GUM, are generally not thought to cause major problems. In some cases, evidence of non-Gaussian distributions gives clues to the presence of systematic errors. In the Photonics area non-Gaussian statistics feature strongly; for example, there are applications where it is necessary to

- take account of the Maxwellian nature of the distributions;
- establish and characterise the type of distribution from which a histogram originates.

In Thermal Metrology, noise in laser-based measurement sometimes is such that an assumption of Gaussian noise is inappropriate.

In Acoustics, the Gaussian assumption is often made despite the fact that it is known to be untenable in some cases.

In Flow Measurement, most input parameters to the calibration process are taken to have uniform distributions, in order to give simple conservative estimates.

In the VAM area, many of the distributions are non-Gaussian and there is a strong demand for guidance complementary to the GUM to deal with non-Gaussian distributions.

3.2.3 Levels of confidence other than 95%

Where a level of confidence is required, 95% is usually considered to be acceptable, but this is often quoted in terms of two standard deviations (which is true for a Gaussian distribution). Some research and development work, however, uses only one standard deviation (equivalent to 68% confidence for a Gaussian distribution), while some theoretical work uses standard uncertainty which is a measure of dispersion that is independent of the underlying statistical error distribution.

For some Humidity work for the UK Ministry of Defence (MoD), 99% confidence is required, but this is rare.

For some VAM work, in certain forensic applications, a nominal 3σ allowance is made in assessing breach of a statutory limit, corresponding to approximately 99.7% confidence assuming normality. In critical cases, however, it is substantially more useful to obtain independent verification of a value than to express a wider confidence interval. This preference follows from the typical rate of control failures; expressing a 99% or 99.7% based on a normal assumption is futile when typical measurement performance leads to gross outliers in up to 3% of cases.

3.2.4 Supporting software

In some areas (Length, Photonics, Optical Radiation, Flow, VAM), statistical calculations are made using spreadsheets, typically Excel, and their in-built functions, although these are not always fit for the purpose. In other areas (e.g. in the Foundation Programme and in some Flow laboratories), MathCad is used in preference. In yet other areas (e.g. Acoustics, gas and surface analysis within VAM), home-grown software is used, perhaps written in FORTRAN or using Matlab. In some laboratories (e.g. in the analytical chemistry area), statistical packages such as Statistica, Unistat, SAS, and S+ are used. In the Flow Measurement area, there is seen to be a need for easy-to-use curve fitting routines for the complicated functions they use in uncertainty estimation.

3.2.5 Supporting documentation

The GUM [1] is the base document in the area of uncertainty estimation. UKAS document M 3003 [2] is also used in the areas where measurement and calibration services are provided, and is often preferred to the GUM. In several areas, the GUM's distinction between Type A and B evaluation of uncertainty is felt to be less helpful that the old "random" and "systematic" distinction.

In addition, in the Photonics area an IEC document gives guidelines on how to establish and characterise the type of distribution from which a histogram originates by comparison with the procedure for Maxwellian or Gaussian distributions. There is soon to be a new appendix to M 3003 devoted to the Ionising Radiation area, replacing UKAS NIS 0825 which has been withdrawn.

In the Flow Measurement area, BS ISO TR 5168 [6] has historically been the main document used and is much preferred to the GUM. Nevertheless, attitudes are changing; laboratories are now generally adopting the approaches in the GUM and reassessing their uncertainties accordingly.

In the analytical chemistry area of VAM, the GUM is considered to be too theoretical and has been interpreted by EURACHEM in the document *Quantifying Uncertainty in Analytical Measurement*, but some find this to be too detailed for routine application. However, this interpretation, the GUM itself and M 3003 all fail substantially in one crucial respect. They appear to require a comprehensive model based entirely on measurable input quantities. Though practical and applicable for reference measurements, this requirement is grossly unrealistic for most chemical testing. In the chemical analysis area, most routine testing is accordingly supported by interlaboratory studies leading to statistical evaluations of whole method performance under stated conditions. The GUM and its derivative interpretations give no guidance on the use of this type of study, nor the data arising therefrom. It follows that the documents do not provide usable guidance for most of the routine chemical testing area. However, the provision of "GUM-compliant" guidance on uncertainty estimation using data from method evaluation studies is being addressed under the present NMS VAM programme, both through theoretical development and via input to EURACHEM and, more recently, ISO documents.

3.2.6 Training

Basic training in uncertainty estimation, such as that provided by UKAS, is sufficient for most metrologists. It has been suggested by some that it would be useful to have follow-up training to relate to how uncertainty estimation works in practice for a given metrological task. To this end, the Laboratory of the Government Chemist (LGC) provides a course specific to the VAM area. Furthermore, there is a significant minority for whom training in robust statistics, non-parametric methods, and techniques to handle correlations and non-Gaussian distributions would be useful in future. In addition, those who use continuous modelling may well require some additional training to appreciate the difficulties of applying the GUM in relation to such models. Two courses on uncertainties are to be offered as a result of the Ionising Radiation Programme and these will need to be co-ordinated with the course to be developed under the SSfM programme.

3.3 Visual Modelling and Data Visualisation

In many metrology areas, visual modelling and data visualisation have not yet been applied other than through the use of common packages to present data graphically. In some areas, visual modelling is generally seen to have limited application. In such areas, the models encountered are invariably simple and pose few, if any, problems with their interpretation. In other areas, the potential value of visual modelling and data visualisation are becoming recognised but no serious effort has yet been put into their application. In contrast, there are a few areas in which data visualisation is well advanced. The following subsections identify the existing and potential application areas that have been identified.

3.3.1 Foundation Programme

In the moving coil kilogram work, data visualisation is particularly important for the understanding of the various effects and processes disturbing the system. At present visualisation is in the form of 2D graphs, each giving a 2D slice through multi-dimensional data. Origin, which has been chosen for its speed, enabling the scientist to switch quickly between different 2D views of the data, is used for this purpose. There is a need for major improvements in several aspects: data formatting, fusing data from a large number of sensors, and 3D visualisation. Any proposed improved system would need to be able to cope with very noisy data, whilst retaining good usability, connectability, speed and the ability to handle real-time data.

There is also potential for visual modelling to be used to visualise microwave fields, for example, to give help in improving the design of resonators and chokes.

3.3.2 Length

Data visualisation is used extensively in graphing sensor outputs, roundness profiles, residual errors, etc., and is important in monitoring system performance and understanding system behaviour. In analysing the measurement strategies for determining CMM systematic errors, visualisation was used to analyse the modes of artefact and system distortion that were undetectable from the data. This led to a clearer understanding of the systems and ultimately to the design of better measurement strategies. Visualisation is also important in understanding artefact deflections estimated from finite element modelling. Many optical systems involve imaging directly.

Commercial tolerance assessment packages use visualisation to display the CAD data. There are opportunities to extend these to portray tolerance information, recommended measurement strategies, actual measurements taken and tolerance assessment results.

In multi-station systems such as multilateration, theodolites, etc., given a CAD representation of the object to be measured, visual modelling could be used to design system configuration and measurement strategies to take into account line-of-sight and other configuration constraints.

3.3.3 Mass

In Pressure and Vacuum the visualisation of measurement data is used to inform the operator undertaking a pressure balance calibration. There is potential for applying both visual modelling and data visualisation to the problem of understanding the deformation of the piston-cylinder assembly of a high pressure balance.

3.3.4 Electrical Metrology

There is potential for visual modelling and data visualisation in the following areas:

- The presentation of large quantities of data. For example, in the RF and microwave guided wave theme, calibration certificates can contain between 5000 and 10000 individual results, and there may be some benefit in a graphical presentation of these results.

- The visualisation of a number of correlated complex-valued quantities requiring dynamic visualisation within many dimensions.
- The electro-optic mapping of fields for MMIC measurements requires 2D presentation.

3.3.5 Photonics

Visual modelling and data visualisation are seen as having potential application in, for example, the reconstruction of three-dimensional surfaces from sets of two-dimensional measurements.

3.3.6 Optical Radiation

Visualisation is seen as potentially valuable in relation to the measurement of appearance properties, such as texture, in order to deal with 4^{th}, 5^{th} and 6^{th} dimensions. The work on appearance is very new, so it is too early to develop a case study in this area, but those working on appearance may well benefit from the output of the SSfM Visual Modelling and Data Visualisation project or related projects in the next SSfM programme. The timescale for this potential application of visualisation is about 5 or 6 years.

3.3.7 Ionising Radiation

Collation and presentation of results from Monte Carlo simulations is considered an area where improved visualisation techniques might be of benefit. For this purpose, there is some software currently available and more in preparation.

3.3.8 Thermal Metrology

Visual modelling is seen as very valuable in the Thermal Metrology area. Most of the continuous modelling packages, e.g., Fluent, offer visualisation of the results to aid interpretation and analysis. However, the graphics are not truly interactive. Hence, it is generally not easy to use them for design of the measurement system.

Visual modelling is one of the leading items on a “wish list” for the Thermophysical Properties work. It is seen as an enabling technology to assist metrologists in modelling more difficult materials (composites, etc.). The main concern is that visual modelling should be implemented in such a way as to exploit the mathematical strengths of the application packages. There is also a concern about validating a visual model.

In the Humidity area, the primary gravimetric hygrometer requires knowledge of flow rate, pressure, mass of moisture and mass of dry gas in addition to temperature. If these are varying, e.g., to test the apparatus, more sophisticated data visualisation would be desirable.

3.3.9 Time and Frequency

One possible application lies in the collection of streams of data from various different systems. At present these streams are viewed as lists of numbers. There would be benefit in converting them to a visual form where it would assist analysis (such as identifying outliers).

3.3.10 Acoustics

Some quantities require 3D plots, in which case packages are used to display directional response, by frequency. Scaling is important for interpretation. Similarly, sound intensity is a vector quantity, requiring multidimensional plots. Intensities are always obtained on the assumption of plane or spherical waves — typically being derived from squared pressure. Quiver plots are already in use for sound power output visualisation.

There is scope to apply visual modelling in this area of metrology, e.g., in optimising the parameters of the transducer or the various acoustic transmission mechanisms.

3.3.11 Flow Measurement

For intercomparison work, Youden plots are used extensively with mixed opinions of effectiveness. They are seldom used without accompanying calibration curves. In some areas of multiphase flow measurement, three-dimensional surface plotting is being used effectively and several contouring software tools are employed.

3.3.12 Valid Analytical Measurements

Visualisation is used extensively in surface analysis, in general using commercial packages. In-house plotting and imaging software mainly uses Matlab.

Applications such as surface modelling appear in post-measurement interpretation. For example, the distribution of contaminants on development sites uses purpose-designed surface modelling software. It has been suggested that a visual indication of the uncertainties in such surfaces would be a valuable extension of such software.

3.4 Data Fusion

3.4.1 Applications

The list below identifies the typical areas of application of data fusion, with an indication of the extent to which each is applicable to the different metrology areas.

While the examples given for the Length area can legitimately be viewed as data fusion, it is fair to say that the data fusion *per* se is not particularly emphasised within the Length community. Many models involve some sort of data fusion. One can view co-ordinate metrology as “the data fusion of tactile sensors”. However, many of the problems do have common characteristics, e.g., involve structured matrix systems, and a generic approach, whether labelled data fusion or not, could bring benefits in terms of efficiency, improved understanding and, ultimately, more accurate measurement systems.

Although the specific examples given below do not include any from the Flow Measurement area, data fusion is becoming a feature of many modern flowmeters and flow sensors. In general the number of inputs is small and the data processing relatively simple. Even with the simple systems making the software interact in real time with the user is a problem. Many flow applications are batching systems and the dynamics of this seem to be very difficult to control.

Data Fusion was recognised in the Foresight Programme as one of the most important topics in industrial measurement affecting flow measurement and process control. As a result an NMS project was formulated and completed in 1998. Sira was the main contractor and the final report has been produced [10].

In the area of multiphase flow, data fusion, complicated algorithms, neural networks and many other techniques are being researched and investigated. They have been investigated within the Flow Programme by Cranfield University and a report produced [11]. Further work has been proposed for the 1999-2002 programme and awaits assessment and prioritisation.

In the VAM area, data fusion is best identified with the discipline of *chemometrics*, the application of multivariate statistical methods to chemical data. This is now a well established discipline with a strong academic and industrial research base. New applications of chemometrics techniques are reported weekly in the analytical literature and, in addition, there are regular meetings of special interest groups such as that of the Royal Society of Chemistry. Instrument manufacturers routinely incorporate chemometric techniques in their data processing software, either as explicit functions called by the user or as hidden routines of which the user may be unaware.

Some of the more common chemometric techniques employed in this area are Principal Components Analysis, Factor Analysis, Linear Discriminant Analysis, Cluster Analysis and, in particular for calibration, Principal Components Regression and Partial Least Squares. Artificial Neural Networks (ANNs) are often employed for classification purposes although, where comparisons have been carried out, it is not clear that the results produced are significantly better than those obtained using the more traditional multivariate statistical techniques. ANNs are also being used for quantitative measurements and an application of an ANN for the quantification of proton NMR spectra has recently been described.

Accordingly, the examples from the VAM area which are identified below focus on applications of chemometrics. It should be noted, however, that data fusion in principle covers a broader field; there is, though, no evidence at present of applications in the VAM area other than the chemometrics applications described below.

3.4.1.1 Processing and reconciling measurements from an array of sensors

Processing and reconciling measurements from an array of sensors is the current use of data fusion in the moving coil kilogram work in the Foundation Programme. The data are fused by using time indicators in the data based on the modified Julian date. Data reduction is performed by hand, using the knowledge that the scientist has regarding which data should be ignored (e.g. because there were unusual factors corrupting the data in a given time interval). It is considered that the automation of this data reduction would be too difficult, due to the difficulty of capturing the expert knowledge of what was happening during the experiment at specific times. However, there is scope for automated data fusion to look for hidden patterns in the data.

There are circumstances in some Foundation Programme projects in which data from a large number of thermometers and time sequences has to be reduced. At present this reduction is carried out by looking for visible correlations and using intuition rather than mathematical techniques. There seems to be scope for investigating whether data fusion techniques could be used to perform more effective data reduction in such applications. However, there remains some understandable scepticism from the physicists that any form of algorithm could replace scientific understanding in areas of fundamental research.

In the Length area, multi-station measuring systems require fusing sensor data from more than one station in order to determine target co-ordinate estimates in a common frame of reference.

In the Electrical Programme’s Ultrafast theme, the matching of measurement data sets with different origins by appropriate shifting and scaling of those sets is undertaken. Generally, *ad hoc* methods are used at present for this purpose. Data fusion techniques can be expected to be appropriate

In the Photonics area, data fusion has potential in the analysis of distributed stress measurements in fibres, for example.

In the Thermal Metrology area, data fusion is used for processing and reconciling measurements from an array of sensors.

In the Acoustics area, different analysis techniques are used over different regions of the frequency scale. Generally there will be some regions of overlap, which eases the task of smoothly fusing the data. One method of fusing such data takes account of the relative uncertainties and uses weighted means.

3.4.1.2 Fusing measurements corresponding to different sets of parameters

In the Length area, in the use of the comparator principle, measurements of an artefact to be calibrated are fused with measurements and calibration information associated with a similar, high quality artefact (possibly obtained using a different system) in order to reduce the uncertainties arising from systematic machine errors. In theodolite and photogrammetry systems measurements of a calibrated length artefact have to be fused with the sensor information in order to determine the scale for the target co-ordinates. In optical interferometry, temperature, pressure and humidity measurements are fused with fringe counts via refractive index corrections in order to determine estimates of displacement.

In the VAM area, combining different sets of parameters in order to generate new information is a well developed application area. In the field of environmental analysis, for example, unknown crude oils (from spillages, for instance) can be classified using chemometric techniques.

3.4.1.3 Fault and error detection/diagnosis by multivariate statistical process control

In the Pressure and Vacuum part of the Mass area there is scope for applying data fusion in the maintenance of the high pressure scale. The scale is established at each end by different calibration techniques (cross-floating at one end and dimensioning at the other), and using a variety of different instruments in between that cover (possibly overlapping) sub-ranges of the scale. There is a requirement to develop a mathematical framework for the high pressure scale into which experimental and theoretical data from these various sources can be inserted, and to be able to investigate the effect on its associated uncertainties of changing the relative accuracies of the information provided.

In the Thermal Metrology area, data fusion is used for fault and error detection/diagnosis by intercomparison of measurements.

In the VAM area, fault and error detection/diagnosis using chemometrics is not a mainstream activity for analysts. An emerging field, however, is multivariate statistical process control, which combines data from many different measurements to provide a rapid indication of process failure.

3.4.1.4 Combining results from several different measurement processes or systems

Combining results from several different measurement processes or systems is applicable in the Foundation Programme in that it applies to the combination of data from two very different measurement processes in the moving coil kilogram work.

In the Length area, in the combined optical and X-ray interferometry instrument, X-ray fringe counts are fused with optical fringe fractioning data.

In Mass and Density, there is a potential application of data fusion to remove any “mismatch” between the measurement data arising from the use of different mass balances.

In the Thermal Metrology area, there is scope to apply data fusion to combine results from several very different measurement processes, because the target measurement was infeasible.

In the Time and Frequency area, it is hoped that combining time and frequency data from GPS, GLONAS (Global Space Navigation System, Russian equivalent of GPS) and the 2-way Satellite system will result in lower uncertainties associated with time and frequency transfer.

In the Acoustics area, one use of data fusion is in characterising noise in terms of various measured attributes. The statistical package SPSS is then used to analyse the individual data streams and fuse them to gauge how the noise would be perceived.

In the VAM area, chemometric techniques are employed in industrial products for, among other things, this type of application. An often quoted example is the “Warwick Nose” (developed at Warwick University) which consists of an array of sensors designed to characterise the chemical components of gases and vapours. This system has been commercialised and one of its uses is to detect the presence of off-flavours in beers and other beverages. The topic of electronic noses is also being considered in the work on odours within the gas analysis part of the VAM programme.

The use of chemometric techniques to generate information from proxy analytes has been investigated. The typical motivation is where the target measurement is difficult and hence expensive to make, but a number of cheaper measurements with which it can be shown to be correlated can be made more easily. For example, the total amount of pollutants entering a river from an industrial site has been estimated from a subset of all the parameters measured, and, as another example, the number of different measurements needed to characterise certain classes of hydrocarbon fuels has been reduced by up to 60% by identifying correlations between specific parameters.

3.4.1.5 Dealing with predictive data from models, or historical or archive data, in tandem with current measurement data

Dealing with predictive data from models, or historical or archive data, in tandem with current measurement data is applicable in the moving coil kilogram work in the Foundation Programme. In studying the acceleration due to gravity, the effects of the sun and moon are predictable and corrections are made by using the appropriate model. However, the effects of the mass of air above the experiment and the water table underneath are not predictable and hence are not modelled. Some reconciliation and correction are necessary, and to which data fusion has potential application.

In the Length area, in determining CMM behaviour, prior calibration information about an artefact has to be fused with current CMM measurements. Once the systematic errors have been estimated, this information can be fused with ongoing measurements in order to monitor system stability.

There has been and will continue to be some work in the Optical Radiation area which compares results from colour specification models with others obtained by human experimentation.

Data fusion, in a conceptually simple form, is used in the Thermal Metrology area in the ITS-90 scale, which employs a composite discrete model, the individual models applying over finite regions of the temperature scale. There is also scope to apply data fusion for dealing with predictive data from models, or historical or archive data, in tandem with current measurement data (e.g., by Kalman filtering).

In the Time and Frequency area, in the longer term, GPS disciplined oscillator developments will lead to a requirement for data fusion techniques to combine data.

In the Acoustics area, their is potential for integrating predictive data from (continuous) models with measured data, using updating strategies based on powerful statistical techniques, such that the overall uncertainty is reduced.

3.4.1.6 Filtering out various systematic effects to reveal underlying trends

In Ionising Radiation, an area that might benefit from data fusion is the examination of the low dose radiation of a pellet with large background noise.

3.4.1.7 Handling numerical, image, and/or qualitative data within a single framework

Handling numerical, image, and/or qualitative data within a single framework is potentially applicable in Optical Radiation to the analysis of appearance parameters, for example to measure how healthy hair is from its appearance.

Such data handling is also potentially applicable in the Thermal Metrology area.

In the VAM area, combining numeric and non-numeric data within a single framework is most likely to be handled via expert-systems approaches, though it is possible to use “qualitative” data in chemometrics models.

3.4.2 INTErSECT awareness

Data Fusion is one of three main themes being pursued by the INTErSECT Faraday Partnership run jointly by NPL and Sira. The other two themes are Structural Modelling and Sensors. Data fusion is the main link between SSfM and INTErSECT.

There is good awareness of INTErSECT within NPL. There are already planned collaborations in several areas within INTErSECT (e.g. Optical Radiation, Thermal Metrology, Acoustics). Thermal programme staff are involved in the LUMITHERM Sensors theme Flagship project, but have fewer links with the Data Fusion theme. Acoustics programme staff are involved in the AETSAD Flagship project, in collaboration with staff from the Materials and SSfM areas.

In contrast, there is little awareness of the INTErSECT Faraday Partnership in NEL and the Flow Measurement community.

There is little, if any, awareness of the INTErSECT Faraday Partnership in the VAM area, apart from in the gas analysis area. Those working in the other areas of the VAM programme would find it useful to have information on this activity. There are no known collaborations with INTErSECT projects in the VAM area. The only identified possibility of an INTErSECT sponsored project in the VAM area is the development of novel sensors and related data processing techniques for use in gas analysis.

4. Validation and Testing

4.1 Spreadsheets and Other Mathematical Software Packages

4.1.1 “Off the shelf” software

The “off the shelf” mathematical software packages used most commonly in the NMS metrology programmes are as follows (the number of metrology areas in which each is used is given in parentheses):

1. Excel (10) - Used for the calculation of uncertainty budgets, the calculation of basic summary statistics, modelling, data analysis and curve fitting, data visualisation.
2. MathCad (8) - Used for routine data analysis, more complicated data analysis, modelling, general mathematics, and algebraic manipulation.
3. Matlab (7) - Used in a range of modelling and application software development, data visualisation, general mathematics and simple statistics.
4. LabView (6) - Used for instrument interfacing and control, and for related modelling.
5. Mathematica (6) - Used in a range of modelling and application software development, data visualisation, and the control systems toolbox is used to design and model control systems.
6. Origin (3) - Used for data analysis and graphics.
7. The NAG Library (3) - A general-purpose numerical subroutine library, the use of which is decreasing.
8. PAFEC (3) - Used for finite element modelling.

The most commonly used language processing software (compilers and interpreters) are those for HT-Basic, HP-Basic, Visual Basic, and FORTRAN.

In addition, the Foundation Programme uses Maple, Mafia, and Measure.

The Length area also uses co-ordinate measuring system in-built functions, e.g. Quindos, and some image analysis packages.

The Electrical and Photonics areas also use Lotus 1-2-3.

The Ionising Radiation area also uses the following:

- Sigmaplot (for graphics, calibration curve fitting and curve fitting).
- Instat (based on the statistical software package Minitab), used in the analysis of primary standard measurements.
- Staysl (least squares code for fitting spectrum unfolding for ill-determined problems).
- The FORTRAN related SLAC language MORTRAN.
- A collection of software developed at NPL with some software contributed from external sources (such as SLAC and the Physikalisch-Technische Bundesanstalt (PTB)). This software tends to be written in FORTRAN. FORTRAN calling routines are written for the Monte Carlo simulation software.

The Thermal Metrology area also uses the following:

- FEA packages such as Ansys.

- Opticad and Zemax for ray tracing and optical modelling in the advanced measurement techniques area.
- Thermal modelling packages (often research/non-commercial code).
- Fluent, for CFD modelling of high temperature gases and combustion.
- Sigmaplot (for measurement services data analysis).
- Software from hygrometer manufacturers which relates dew-point to vapour pressure.

The Acoustics area also uses SPSS (for more advanced statistics), NPLFIT (for calibration curve fitting), IDES (for integral and differential equations, and associated calculations, e.g., evaluation of Bessel functions), and LabWindows (data acquisition, instrument design and control package). NPLFIT and IDES are NPL written packages/libraries.

The VAM area also makes use of the following:

- older DOS based spreadsheets such as SuperCalc;
- statistics packages including BMDP (orientated towards the medical sciences), SPSS (orientated towards the social sciences) STATISTICA (which has a large user base in the chemical and physical sciences) and SAS (which is heavily used in the pharmaceutical sector).

4.1.2 Spreadsheets

Spreadsheets, particularly Excel, are commonly used for statistical modelling for uncertainty estimations in most metrology areas. Generally, the spreadsheets used are not particularly complicated, but rely on in-built statistical functions, such as mean and standard deviations, other in-built functions such as max and min, as well as linear fitting routines such as LINEST. Processing is carried out by way of cell formulae and manipulation by macros. Pre-processing of raw measurement data, including scaling and normalisation of the data, is sometimes used prior to the application of in-built spreadsheet functions for regression, and this helps to ensure the integrity of the results returned by these functions.

Spreadsheets are not generally used for modelling or other mathematical computations in the Foundation Programme. However, on some projects data are recorded in Excel spreadsheets for archival purposes. Only simple calculations, such as averaging, are performed on the data.

In the Electrical Programme’s RF and microwave guided wave theme, the use of commercial spreadsheet packages for data manipulation and presentation is widespread. As a consequence, there is a real requirement in this area for assurance to be provided concerning spreadsheet quality including information about the parameter ranges over which specific in-built functions are accurate and fit for purpose. The use of spreadsheets is seen to be unavoidable, and work in this area needs to reflect this and to concentrate on providing guidelines and information to underpin the use of such packages.

In the Electrical Programme’s Ultrafast theme, spreadsheets are increasingly being used as part of research and development projects to investigate “what if” scenarios. Examples of such projects include the processing of experimental data in electrical waveform metrology and the analysis of calorimeter outputs from laser energy measurements. HT-Basic and LabVIEW software has been written to output data to files in a format that is suitable for direct input into Excel where further analysis is carried out. The Excel “Solver” function is used to apply various functions to the data, including sine waves (to determine the period of a cycle to calibrate timebases) or multiple exponentials (calorimeter responses).

In the Humidity theme of the Thermal Metrology Programme, determination of saturated vapour pressure requires a polynomial with a minimum of four terms, implemented in Excel.

In the Acoustics area, although attractive to users, the policy is to discourage the use of spreadsheets, because of the difficulty of configuration control of spreadsheet solutions. It is also noted that spreadsheets are difficult to make generic (which militates against re-use), but easy to alter.

In several metrology areas, particularly the Flow Measurement area, the feature of Excel which gives most attraction is the ability to write functions and routines in Visual Basic for Applications, hence replicating traditional programming but keeping the power and friendliness of a spreadsheet. This approach eases validation, encourages software reuse, and simplifies many complicated calculations and routines. It is suspected that this feature is underused by many laboratories and staff through fear of the complications.

Extensive use is made of spreadsheets in the VAM area with computations at several levels of complexity being performed. As well as being used to evaluate algebraic measurement models of the type $y = f(x_1, x_2, \ldots)$ and regression models, they are also used for statistical calculations (means, standard deviations, t and F tests, Analysis of Variance (ANOVA), etc.), processing measurement uncertainty data and experimental design. In-built functions of Excel that are used include average, stdev, sumsq, sum, var, covar, max, min, sqrt, power, log, ln, exp, abs, and, if, ceiling, floor, rand and normdist. Most uses involve applying the in-built functions, but some use is also made of these functions within user- or professionally-developed “macros”. Some users apply features such as Direct Data Exchange (DDE) or Object Linking and Embedding (OLE) to help automate data acquisition, by driving instrument software from the spreadsheet application.

4.1.3 Awareness of limitations

Within all metrology areas, there is some awareness of the potential limitations in spreadsheets and other software packages arising from inaccurate or unstable numerical algorithms and useful parameter ranges. The level of awareness seems to be highest in the major metrology laboratories (NPL, NEL, LGC) but rather lower in the measurement community in industry.

Users tend to become aware of limitations in spreadsheets and other software packages when an obvious inconsistency arises, for example, when the results obtained with one package differ from those obtained with another package for the same input data. Limitations are also indicated when a package produces garbage output; this can be as a result of a “bug” (coding error) or, not infrequently, through inappropriate use of a function, for example, attempting to fit a polynomial of too high a degree, such as the highest degree available, to a data set.

4.1.4 Testing and validation

In most metrology areas, relatively little testing and validation are generally performed on off-the-shelf software packages, except in the case of the purchase of a new expensive package when some testing will be undertaken, varying in its thoroughness according to how critical it is. Usually, there is a tendency to rely upon well established packages and to check a few results by using alternative methods and spot-checks. In contrast, bespoke software, possibly written using these packages, is usually tested in accordance with software development quality procedures, for example, by comparing calculated results with reference results or results determined using other software.

In the case of Origin in the moving coil kilogram work within the Foundation Programme, validation is performed by using a simulation of the moving coil system. This showed that despite double precision arithmetic, it was only accurate to 1 part in 10^9 which is just adequate.

Many measurements in the Length area rely on complicated in-built functions associated with measurement systems, CMMs in particular. These functions determine best-fit associated features such as spheres and cylinders to co-ordinate data and calculate the parameters such as radius associated with these features. These calculations are a key component in the chain from measurement data to measurement result. In the 1980s, it became apparent that not all software performing these calculations gave reliable results and two CEU projects undertook a test of such software. The results of the first exercise showed that many implementations produced inaccurate results. The second exercise, a few years later, demonstrated that for least squares calculations (Gaussian substitute features), most software (but not all) performed satisfactorily. These exercises lead to the draft of an ISO standard for testing substitute feature calculations and also third party testing services at PTB, the National Institute of Standards and Technology (NIST) and NPL and elsewhere.

In view of the large number of different calculations associated with co-ordinate metrology, it is clear that further work on validating the software performing these calculations is necessary, covering, for example,

- Calculation of Chebyshev associated features and more general tolerance assessment.
- Target location estimation algorithms in laser tracking, photogrammetry, theodolite and portable (flexible arm) CMMs.
- Image analysis software.

In Ionising Radiation, external software, such as public domain software, is tested as far as possible, as the majority has safety aspects and dosimetry errors in radiotherapy and in radiation protection have a high public profile. Independent checks are performed on the software. Results are compared with experimental data and independent calculations are made (for example graphs are checked and spreadsheet models are verified by hand).

In Acoustics, reference data files for reciprocity calibrations have been developed jointly between NPL, DTU (Technical University of Denmark), DPLA (Danish Primary Laboratory of Acoustics) and PTB in a EUROMET project.

In the VAM area, when using a mathematical function within a package for the first time users will very often test the routines by providing input data for which a result has already been obtained manually or with another system. This does not of course constitute a full validation but it does provide assurance that the function is at least capable of performing correctly on the test data.

In the case of bespoke software in the VAM area, the software developer is generally required to demonstrate that the application performs satisfactorily on test data sets supplied by the customer. The test data sets are selected to cover the range of values likely to be encountered and may even be more severe. Despite such testing, users can often recount catastrophic failures arising from unusual data sets, for example, where the standard deviation of a data set forms the divisor in an expression and happens (correctly) to equate to zero. Such pathological cases are, fortunately, rare but illustrate the difficulty of trying to anticipate all possibilities.

Also in the VAM area, the amount of testing and validation performed depends on the type of package and its scale of application. At one extreme (typical within gas and surface analysis areas) users will test mean and standard deviation results, say, from a spreadsheet by repeating the calculations using a calculator or a different software application having the same functionality. At the other extreme, in the case of a Laboratory Information Management System (LIMS) and commercial surface analysis data reduction software, extensive and highly structured testing and validation are performed. This testing includes the use of data sets for which results are known analytically.

4.1.5 Supplier advice

In general, there is little evidence that metrologists demand assurances from software suppliers that their software is fit for purpose. In particular, it is not common for a user of a commercial package to seek assurances from software suppliers that mathematical functions included in the package employ algorithms that are valid, accurate and stable over the likely range of values of the user's input data. The claims of the suppliers are often taken on trust, possibly backed up by expert reviews, prior experience of others or the user's own experience of a demonstration version. If, however, users could be persuaded to press for assurances from suppliers, it might push suppliers to seek objective validation.

If problems arise the supplier ought to be contacted for advice. For low cost products, advice from suppliers is variable. The general impression is that major suppliers, in particular, are often reluctant to do anything about reported faults other than to thank the customer and say that the information will be passed to the development team. Such customer reports are certainly not always acted upon. Smaller or niche market suppliers of moderate- to high-cost products are usually more inclined to respond positively to user concerns.

In the case of software/hardware systems representing large capital investments, the software will be scrutinised to some extent by the user. Where mathematical performance is a critical issue, this will be tested to a level deemed appropriate by the user. There will normally be a contractual obligation on the supplier to co-operate in such testing and it will normally be able to offer much useful advice and information about an established product.

For example, in the Acoustics area, some enhancements to PAFEC have recently been commissioned by NPL. A test plan has been drawn up and reviewed by appropriate software experts, such that these extensions to the functionality of PAFEC will be accepted only when they pass the tests at NPL.

4.2 Model Validation

There is a high level of awareness of the importance of model validation in all areas of metrology, and an acceptance that it is vital to the success of modelling. In particular, it is an increasingly important issue as new models and methods of modelling are considered. Where possible, the models used are validated. Such validation is carried out, for example, by comparing quantities predicted by a computed model with results that can be measured independently or calculated from other considerations. A common approach in discrete modelling is to examine the residual errors associated with the computed fit. The model is (partially) validated if these errors appear random and free from systematic behaviour. Where physical models are used, these are underpinned by theoretical analysis, and are generally accepted on this basis. In addition, validation of empirical models by applying statistical tests to the computed models taken from within a family of functions is sometimes undertaken.

Experimental design is often the purpose of projects in the Foundation Programme. Designs are usually checked by the intuition and understanding of the scientists involved. Predictions are made and compared with experimental results. However, in the moving coil kilogram work validation of experimental design is performed against a simple simulation model of the physics involved. In some cases, models are validated by using limit data for which the answer is known.

The kinds of issues that arise in model validation can be illustrated by the following examples taken from the Length area:

- **Selection from a family of empirical models.** In the use of empirical functions such as polynomial or spline curves and surfaces to model profiles and surfaces, it is important to select the degree or number of terms that produces an adequate fit to the data. Statistical techniques such as examining the root-mean square (RMS) residual as a function of degree can be used more widely to supplement more subjective approaches. For example, the kinematic error model for CMMs uses empirical functions to describe the 18 error components. The choice of degree in the polynomial representation of these components is often determined on the basis of custom, experience, constraints on the number of variables the software can handle or the number of measurements required to determine the model parameters, rather than from strict model validation principles. The kinematic model itself requires validation as it makes assumptions about the independence of the error components. One validation approach is to embed the kinematic error model in a more comprehensive empirical model and determine if the larger model provides a better prediction of behaviour.

- **Measurement strategy validation.** To validate a model from measurement data, it is necessary to have enough data in order to make a sensible judgement. It follows that the measurement strategy used in the model validation must itself be validated. The calculation of form error is heavily dependent on measurement strategy. In determining roundness (circularity) a measurement strategy consisting of just three points produces an (invalid) estimate of zero form error. While this example is clearly absurd, in more complicated situations it is not so simple to specify a measurement strategy that is guaranteed to test the validity of a model and there are model validation or verification procedures to implement that provide only partial information. Much of the focus of some recent and current projects is providing a theoretical and numerical basis for measurement strategies that adequately test the model assumptions. The main approach is the use of estimates of the uncertainties of the fitted parameters as a measure of adequacy of the measurement strategy.
- **Estimator validation.** It is appropriate to consider estimator validation as a component of model validation. In determining system parameters from measured data, it is often the case that the computational aim (for example, the minimisation of an error function) is correctly formulated but that an estimate is determined by solving a simplified computational problem. In determining the Gaussian associated circle or sphere, an approximation is often used to convert the intrinsically nonlinear problem into a simpler, linear problem. It has been shown that this is generally a safe approximation for data distributed uniformly around a circle but introduces significant bias for data on a small arc of the circle. In the determination of form error, the Gaussian associated feature can be calculated instead of the more difficult Chebyshev associated feature, generally without introducing significant bias into the subsequent analysis. In co-ordinate metrology using interferometric transducers, the correction for refractive index effects strictly leads to a generalised regression problem, but an ordinary regression estimator is implemented instead. In implementing approximate estimators, it is important to be able to quantify the resulting bias, over the applicable range of problem domains. In recent years, the null space techniques developed by NPL to generate reference data and results to test software implementations have been successfully adapted to validating estimators. They have been used in a number of Length applications, including the analysis of theodolite data, National Primary CMM and geometric tolerance assessment.

Another example of model validation applied to a discrete model comes from the Acoustics area, where a hydrophone response application was developed in Matlab and tested in an “evolutionary” fashion, by three methods:

a) inputting extreme data and checking robustness;

b) providing data with “known” results and checking accuracy;

c) verifying behaviour of the complete system and checking reliability.

In Flow Measurement, the same high level of awareness of model validation as in other metrology areas exists where competent mathematicians are producing good models to explain behaviour of sensors such as orifice plates, or modelling flows using CFD. However, at the lower end of the scale, models based on curve fitting are employed to describe or extrapolate flowmeter performance and in this area the tools and awareness are found wanting. Staff who are not mathematical experts frequently use curve fitting and statistical tools to describe results with limited understanding of the consequences.

In the VAM area, validation of a measurement model occurs during analytical method development as a matter of course. When methods are put to use they usually involve a calibration stage in which the response of the analytical system, for a set of known standards, is recorded. Analysts strive, and are frequently able, to arrange experimental conditions so that a straight line calibration is adequate. The calibration model is traditionally solved using unweighted linear regression. For some methods, the analytical technique employed leads to non-linear calibrations and here polynomials of degree two or occasionally three are sufficient.

In the special case of multivariate modelling (including ANNs), several techniques for model validation are employed in the VAM area, particularly different approaches to cross-validation, including “leave-one-out” recalculation, and, increasingly, techniques which subdivide the data set at random into “training” and “validation” sets. Models are built on the “training set” and comparison of predicted and calculated values on the “validation” set gives an effective indication of under- or over-fitting.

Experimental designs usually employed in the VAM area are standard designs such as factorial, central composite, Taguchi, etc. These are provided by a statistician working in collaboration with the user and so they are assumed to be appropriate. The Analysis of Variance applied to the resulting experimental data is an effective technique for determining the statistical significance, and hence usefulness, of the model.

In the VAM area of surface analysis, model validation is an important area for further work. There are few calibration stages, but rather the measurement process is analysed from first principles. It is not possible to compare results with other methods of calculation, since often no other method exists.

4.3 Measurement System Validation

4.3.1 System validation

The usual approach taken to measurement system validation in all areas of metrology is to apply black-box testing to the system as a whole. The users expect to notice anomalous results if the system is not performing correctly. The results may be compared against previous results and data from international agreements. They are also usually checked for self-consistency.

Sometimes defensive design techniques are used, in which the measurement system is chosen so as to reduce the sensitivity to all sources of error, including software errors. Traceability and validation are then normally available for the simplest instrument functions, but research-based activity will require more sophisticated functions, which are often difficult to validate. An example of such defensive design, from the Acoustics area, is where, if possible, the key calibration is based on voltage ratios, making the measurement system more robust.

Generally, where data analysis software is separate from the measurement system, such software can be validated separately using existing data and results.

Complete measurement systems are often verified through measurements of calibrated artefacts, calibrated reference instruments or reference materials, as appropriate to the quantity being measured. For example, ISO 10360-2 [7] describes a procedure for verifying the length measuring capabilities of CMMs from measurements of calibrated length artefacts. Also in the Length area, composite artefacts have been designed and used to test a system’s capabilities on a range of standard features such as cylinders and cones. These tests are an indirect test of element fitting software. In Thermal Metrology, some measurement systems are validated by comparing the results of a measurement of a known (i.e., fully characterised) black body with what would be expected.

Novel measurement systems are usually validated by checking their results against results from old systems using established methods. New measurements of fundamental constants are compared against the previous best practice or established values.

Sometimes, the results from direct methods of measurement are compared against those from indirect methods. Occasionally, measurement systems validation is made by comparing results with those produced by a model. For example, in the moving coil kilogram work, in the Foundation Programme, the validation is performed against a simple simulation model of the physics involved. In an example concerning an open-loop positioning system in Acoustics, it was discovered that a certain linear position encoder had a problem, which was identified by validating against a geometric model. In this instance, the Acoustics work benefited from links with those in the Length area.

Systems are also tested against each other. For example, a ability of a CMM to measure circularity can be tested against a roundness instrument, involving a partial test of the embedded software.

Certificates produced at the end of the calibration process are often just checked by eye, whether the certificate has been produced by the software or produced manually (i.e. copied by hand).

Accredited test laboratories in the VAM area are not generally required to validate measurement apparatus, but are required to check correct operation using appropriate test materials. This check may involve as little as running a single test sample through prior to use, though method validation requirements would also have generated sufficient data to validate the precision and accuracy and to show that the instrument was performing adequately in context.

In the pharmaceutical sector, validation is a formal (legal) requirement under Good Manufacturing Practice (GMP) [12], and this sector has accordingly developed a formal approach to instrument validation, termed “equipment qualification”. This approach requires formal testing against specifications set out at the purchase stage, and will typically verify instrument control parameters and check overall performance using test materials. Where software forms part of the function, appropriate tests are used to verify the software performance; it is considered best practice to examine extreme cases where possible.

4.3.2 Embedded software

Most modern instrumentation contains embedded software irrespective of the area of metrology. Furthermore, the sophistication of that embedded software is increasing, making the values output by the instruments highly dependent upon the software. Errors in such embedded software can be very difficult to detect, especially if they cause only small regions of non-linearity in the outputs.

Upgrades to embedded software or firmware can have subtle consequences in respect of timing, etc., and these may compromise consistency.

Even in areas where software is present in most measurement systems and instruments, metrologists will often try to avoid instruments with embedded software for critical applications, for example, this situation applies in the Thermal Metrology area with spectrum analysers.

The complexity of embedded software varies from area to area. For example, a large number of instruments used in the Length area have significant amounts of software performing nontrivial calculations, a main example being CMMs. In contrast, in the Mass and Density and Force areas of the Mass Programme, the software embedded in instruments is generally simple. Where there is the need to extend the functionality of the software controlling these instruments, this is often undertaken “in-house”, and separate from the instrument. One exception to the simplicity of the embedded software is in hardness machines used in the Force area which includes software for obtaining dimensional information from images.

In the Electrical area, even "simple" instruments, such as digital voltmeters, contain software for operating on the raw measurement data. For example, software is often included to undertake digital filtering. Typically, details of the algorithms used are not documented. In addition, there are many instruments, such as network analysers, that contain software for performing significant numerical computation.

Increasingly in the Flow Measurement area, flow sensors have microprocessors built in. In the simplest form, the turbine or rotary meter may have a lineariser which will take the calibration errors of the meter at different flow-rates and apply them to correct the output. Invariably these are five point and interpolation type corrections. Increasing sophistication adds second sensors to provide temperature and pressure correction and of course communication. At the higher level, some modern flow meters, for example Coriolis and ultrasonic meters, rely on sophisticated software to operate by converting and correcting the sensor information to engineering units or producing pulsed and current signals.

In the Pressure and Vacuum area of the Mass Programme, some problems with the software embedded in instruments have been experienced, for example:

- for a particular vacuum gauge, the software embedded in the gauge introduces a "kink" into the measured characteristic, and
- some instruments include procedures for calibration where the calibration points are fixed by the manufacturer, even though there is evidence for non-linear behaviour between these points; in this case, the calibration of the instrument is driven by the manufacturer's procedure rather than good metrology.

In the surface analysis area of the VAM Programme, a diagnosis system has shown software bugs and the use of inadequate algorithms in commercial data reduction software.

Within the Foundation Programme, in the moving coil kilogram work, embedded software is designed and developed within the projects concerned using an unusual approach. The instruments are connected to the processor via a fibre-optic ring using the General Purpose Interface Bus (GPIB) specification, which is a protocol-free bus, giving low-noise high-speed connectivity, whilst allowing the controlling software to be considered both as embedded and non-embedded. There is also some completely embedded code in the form of programmable logic.

4.3.3 Supplier assurances

Metrologists, as a rule, have to rely on the claims of manufacturers that embedded software has been adequately tested against supplier specifications. Users are nonetheless considered responsible for ensuring that purchased products meet specific needs. In many areas, assurances are not requested from suppliers of commercial software packages, but assurances of adequate testing are required from suppliers of bespoke software.

There is a concern among many metrologists that instrument vendors may run "obvious" tests which are insufficiently rigorous to catch all the errors and hence such instruments must be viewed as unpredictable. Many metrologists are sceptical of the manufacturer's claims. Of course, the continuing successful use of the software confirms and increases the confidence that users have in it.

Within NPL, software supplied by one part of NPL to another will normally be produced to a software quality standard forming part of the NPL quality system. This ensures that adequate testing is performed before the software is put into use. However, in areas of more advanced research, the software will often be developed by the scientists themselves for their own use and continual adaptation; such software is not generally tested as thoroughly (in quality system terms it is prototype software rather than production software).

In the Time and Frequency area, NPL receives statements on Year 2000 (Y2K) and GPS week number rollover, but does not receive guarantees on code integrity.

4.3.4 Supplier validation of embedded software

Metrologists expect that suppliers should validate embedded software as part of the design and implementation process. Current recommendations on software qualification require a "design qualification" to verify that the design meets requirements; this is necessary but not considered sufficient. Most suppliers use a specification-design-code-test process through implementation; in conjunction with a design qualification, the specification/test pair constitutes validation. In addition, software needs to be tested in situ and under normal performance conditions.

In some sectors, particularly the pharmaceutical sector, common guidelines exist. However, different sectors have different requirements, making general guidelines hard to provide. Development and testing in general are almost certainly best carried out within recognised software development systems, such as TickIT, which allow suppliers to choose their own processes to achieve fitness for purpose, accommodating sectoral guidance as appropriate.

Applicable software certification or validation schemes for suppliers to use are rare. One exception is the C source code validation service run by NPL.

4.3.5 Access

Metrologists should ideally have sufficient access to embedded software to satisfy themselves that it meets their own requirements. It is not often necessary to require access to software performing a purely control function, for example. However, where the software is operating on actual measurement data, it would be very useful for the metrologist to be able to test this independently of the supplier. Unfortunately, suppliers do not generally provide such access. Indeed, it is usually considered that embedded software is inaccessible by definition. This means that the user has no access to the source code and may have limited means of inputting test data and intercepting intermediate calculations. Since much of the intellectual property associated with many of these systems is bound up in the software, there is no likelihood that users will have access to source code in the future. Furthermore, details of the algorithms used in embedded software are usually not documented.

The approach used in the moving coil kilogram work (as described in 4.3.2) may be applicable to other areas of metrology where separation of the software from the hardware would be a significant advantage, without compromising speed or control.

4.3.6 Validating embedded software

Given the lack of access to embedded software, most modern instruments containing embedded software tend to be validated as complete systems.

In those cases where the software is developed in-house, there is the potential to validate the software in isolation from the hardware, but this is not always done; the software is instead sometimes only tested together with the hardware in full system validation. This approach would still be consistent with quality system requirements if the software is classed as prototype software for internal use only and which needs to be under continual development. However, if suspicious results indicate a problem with such software, then it can and should be tested in isolation from the hardware.

In the case of the control software in the Foundation Programme's moving coil kilogram work, some parts of it need to be very thoroughly tested and validated using techniques appropriate to high-integrity software. The fact that the software is readily separable from the instruments makes this possible.

In the Length area, both users and suppliers recognise the need for validation of such software and there is a draft ISO standard in place for testing software for calculating Gaussian associated features. This standard addresses a significant area but there are many other areas for which a similar approach could be applied. The NIST algorithm testing service also includes testing Chebyshev associated features (but uses an approach that could be problematical). NPL's Jigless Aerospace Manufacturing project includes a work package to develop a methodology (based on black-box testing) to validate software for calculating target locations from sensor reading for theodolites and similar systems.

In the Mass and Density area of the Mass Programme, the in-house extensions to the functionality of the software controlling instruments such as mass balances can be validated in isolation from the instruments. Similarly, in the Force area, the software for obtaining dimensional information from images, which is included in hardness machines, is independently, as well as in context when the instrument is tested as a whole.

In the Ultrafast theme of the Electrical Programme, embedded software for performing numerical data processing is validated by comparing the results returned by the instrument with results calculated using software written at NPL. This approach is akin to using reference software to generate reference data sets and results. In some cases, it is possible to inject the data into the instrument, but generally this is not done and the measurement system is tested in its entirety.

For instruments used in the Electrical area, such as network analysers, that contain software for performing significant numerical computation, the only feasible approach to validating them is to use calibrated components and artefacts. Ideally, these instruments should record their raw inputs and outputs so that the processing undertaken by the instrument can be independently validated.

In the Acoustics area, embedded software implements functions such as Fast Fourier Transforms and spectrum analysis. In Noise and Audiometry, such embedded software is supplemented by off-line calculations in a bought-in package which implements the relevant ISO Standard, with prescribed validation procedures. The validation procedures supplied with embedded software in the Acoustics area are often based on inputting a known signal and testing against predicted output. Indeed, this is seen as part of the measurement audit process. As an example of this process, microphones with a full calibration history can be validated experimentally with new data.

In the Flow Measurement area, the software in some modern flowmeters, for example, Coriolis and ultrasonic meters, is not independently checked and in fact could not easily be checked as too much proprietary information on performance is included. For this type of meter, most errors are found to be due to inadequate allowance for changes of fluid parameters or conditions. This inadequacy could be due to deficient algorithms or the use of the device outside the original design range of the software. Most sophisticated meters of this type are tested by calibration and the software does not require separate validation by the calibration or regulatory authority for normal use. It is interesting to note that the operating speeds and update times can frequently be the cause of poor calibration, owing to a lack of understanding of the interaction of the microprocessor, the software and the calibration process.

Some Smart instruments designed for flow measurement now have many flow computer routines built in and validation of these should be possible. It is relevant that most sophisticated flowmeters have many diagnostics built in which if they can be accessed can tell much about the behaviour of the meter in service. Access to these and their meanings is often not well documented by the supplier to the user test laboratory or even the front line service engineers.

In the surface analysis area of the VAM Programme, the development of standard import and export facilities allows reference data sets to be used for validating aspects of data reduction software, but not the instrument control software.

4.3.7 Reference data

Reference data is sometimes used in the validation of small subsystems. More often, typical data or simulated data (for which the right answers have been calculated) are used instead of reference data in measurement system validation.

Reference data sets are provided with embedded software in some areas, for example the Photonics and Optical Radiation areas, and these are used for validation purposes. Good reference data sets are collected by NPL in the Time and Frequency area, so that receivers can be tested for transient effects such as cut-off angles, reflections and atmospheric. In the VAM gas analysis area, there is some regression testing performed using previously captured data used as a reference data set. Also in the VAM surface analysis area, reference data sets are generated to validate commercial software used for data reduction.

In Thermal Metrology, reference data are used in a sense, in that results from measuring a fully characterised black body are used, but generally software is not formally tested with third-party reference data in that way.

In other areas, for example the Ionising Radiation and VAM chemical analysis areas, reference data sets are not often used.

4.4 Validation of Simulated Instruments

There are many metrology areas in which there is no current explicit use of simulated instruments or any projected use, although it is fair to say that the concept is not yet well understood in most of these areas. Thus, the SSfM project on Validation of Simulated Instruments will need to focus on those few areas in which there is current use of simulated instruments or in which the potential use is recognised. As the concept and its applicability become better understood through the technology transfer activities of SSfM, it is possible that other areas of metrology may also find value in applying the concept.

Numerical simulation has been an important tool in Length metrology, used in instrument design and configuration, the design of control systems, the selection of measurement strategies and the evaluation of measurement uncertainties. A large European project led by PTB has as its focus the implementation of an online Virtual CMM which takes actual measurement data and uses a comprehensive numerical simulation of the CMM and its main sources of uncertainty to estimate, through Monte Carlo and other techniques, the measurement uncertainty associated with the measurement task. Similarly, there is a Virtual Multilateration CMM produced by NPL using Matlab. However, there are many issues associated with the approach adopted in these simulations, some of which will be discussed in the SSfM Validation of Simulated Instruments project.

Monte Carlo simulations are used in some projects in the Foundation Programme. In one project, simulation software that was provided with a vacuum system is used for guidance purposes only. There is also a relatively simple simulation used in the moving coil kilogram work. It would be appropriate to have a more comprehensive simulation, but it would need to be highly sophisticated and it is unclear that it would be cost-effective to produce. There may well be more scope for simulation to be used in other Foundation Programme projects, but only in situations where the physics is well understood.

Continuous modelling in Thermal Metrology is often used to simulate the operation of an instrument or measurement technique using Monte Carlo simulations. However, there is no evidence that this includes statistical modelling of the errors and the ability to configure such models to simulate actual instruments, which is the case with the Virtual CMM concept.

Monte Carlo simulations are also used with evident success in Ionising Radiation. For example, the work with the long counter could well prove to be useful to the SSfM Validation of Simulated Instruments project.

Considerable simulation is undertaken in the Acoustics area, particularly in Underwater and Medical Ultrasound. The simulations, using continuous models, tend to focus on key features, such as the diaphragm and piston, that have a profound effect on the measurement results. The work on calculable response systems has potential to be developed into simulated instruments.

5. Metrology Software Development Techniques

5.1 Software development methods

5.1.1 Software Quality

5.1.1.1 Quality systems

NPL is certified to ISO 9001 [3] plus TickIT, so any new software NPL develops for sale or for use in providing a service is produced in accordance with the requirements of TickIT (and classified as production software, *NPL software integrity level 3*). Thus, it is properly engineered and produced in accordance with documented procedures. In such cases, the discipline of using the quality system has improved documentation, particularly the specifications. However, much of the bespoke software developed within NMS Programmes is only for use in research and as prototypes for internal use only; hence, such software need not meet the requirements of TickIT, but only the lesser requirements appropriate to such software (the *NPL software integrity level 2* requirements for recent software). Furthermore, there is still in use a lot of production software produced before the introduction of a formal quality system, but it has been validated in practice over several years of operation; such software is classified as *NPL software integrity level 3 extant.*

LGC reports that all the VAM instrument supply respondents contacted operate under an ISO 9001 Quality Management System. In contrast, NEL reports that for many involved in Flow Measurement, quality assurance through TickIT is but a dream!

Sometimes, new software is developed by external software specialists, many of whom are TickIT certified. Although this has often produced good results in terms of cost, timeliness and quality, some have found that considerable effort needs to be put into writing the functional specification to ensure that the requirements are tightly specified; others report that there is a point in developing a functional specification beyond which further refinement does not reduce the amount of debugging that will be necessary. In any case, many people need training in how to write more effective functional specifications.

5.1.1.2 Quality attributes

The key quality attributes for metrology software are reliability, numerical accuracy, usability, maintainability, configurability and (particularly in the research areas) adaptability. Of these, reliability is the one most frequently identified by metrologists. Commercial developers, on the other hand, cite usability first. Usability also impacts on reliability; as well designed software interfaces should reduce human error and alert the user to other problems. Conversely, ease of reconfiguration may lead to practical difficulties in configuration control.

In addition, for key software used in measurement services, it may be necessary for the software to provide a statement of the processing undertaken, as well as to enable the measurement chain linking raw data to measurement result to be recreated. For example, the particular use of instruments, and the way they are linked within the measurement chain, may change from day to day, but traceability from measurement data to measurement results may need to be maintained.

5.1.1.3 Reliability records

Where the reliability of software impacts the quality of a measurement service, software reliability records are generally kept. For example, such software in the VAM area is subjected to a Software Quality Audit before it is released and, on an ongoing basis, a system of recording and tracking faults is used. Faults, as entered in a fault tracking database, are reviewed regularly and attention is drawn to any reliability problems.

Formal records of software reliability are, however, not normally kept for software used in research and development activities. Although project-specific records are kept in individual lab books, no programme-wide or organisation-wide monitoring or control takes place.

5.1.1.4 Reliability

It is generally considered that current software development methods and the methods used for checking software reliability provide sufficient reliability for research purposes, but it is recognised that the reliability can always be improved.

In some NMS programmes, such as Time and Frequency, there is heavy reliance on the fact that data are compared with those from outside groups, and there is the expectation that any significant errors will show up through the comparisons.

5.1.1.5 Version Control

Within NPL, configuration and version control of software are used consistent with the requirements of the NPL quality system. This is important if results need to be reproduced. Control information is put in the headers of data files to ensure that the data can be correctly traced to the exercise in which it was generated.

Special attention is paid to version control in the moving coil kilogram work, using the SourceSafe package. At present this is run as a single-user system on a single machine, but it would be better to have a system which could be run over the network whilst maintaining a high level of security, and providing a fast and clear visual differencer.

5.1.1.6 Quality system concerns

It is clear that attitudes vary towards the effects of quality system requirements on software development. The variations seem to reflect more the different ways that different people interpret the requirements, rather than different needs of different metrology areas. Some people, but apparently only a minority, express the following concerns about the quality system requirements:

1) home-grown legacy software (often classified as extant) is difficult and costly to modify and hence the user/developer is inclined to live with bugs rather than correct them;

2) the design stage is made unnecessarily cumbersome, to the extent that fewer options may be tried than would have been the case otherwise;

3) stringent requirements for regression testing soak up too many resources;

4) quality system requirements assume that the developer is completely separate from the user, whereas in many cases, especially in research activities, the developer and the user are the same person;

5) people engage in retrospective efforts to satisfy the quality system with very little apparent added value.

On the other hand, it is widely agreed that much had been learnt from implementing ISO 9001 [3] and TickIT, with significant improvements in documentation and a growing appreciation of a professional approach towards metrology software development.

It seems that there is a need for more guidance on appropriate interpretation of the quality system requirements, making sure that the effort expended is in proportion to the risks faced and benefits obtained. There also needs to be a positive promotion of the benefits of adopting a thorough approach to design, documentation, testing and version control in software development.

5.1.2 Management issues

5.1.2.1 Maintenance

Software developed in-house for research and prototyping purposes often does not so much need maintenance as continual refinement as the research progresses; the development and refinement of such software can only be done effectively by the scientist or by a programmer working very closely with that scientist.

Legacy software developed for instrument control of old hardware needs to be maintained for the life of the hardware. This and other legacy software is prone to the following problems: inadequate documentation, loss of contact with the original software developer, use of obsolete languages for which training is no longer available, use of an unreliable or obsolete hardware platform. However, in some cases, good quality support has been available for maintaining legacy software; instead, the main problem found with continued use of such legacy software is its slow speed.

For commercial software developers, however, the position is generally much better and the large investments made in reliability measures mean that maintenance costs can be kept relatively low.

Designing for software reuse is an effective means of reducing software maintenance costs (see Section 5.2).

5.1.2.2 Development costs

In most metrology areas, current software development costs are considered to be reasonable. However, the quality of service provided by the software developer is considered to be more important than the initial development cost, because it is recognised that the effectiveness, reliability and flexibility of the software all have a major impact on the lifetime costs of the software. As a result, some software is developed externally, some by software professionals within NPL, and some by the scientists themselves, depending on the skills of the scientists themselves and on the relationships they have developed with software developers. In cases where most of the development work involves learning about the capabilities of the equipment, and where the software needs to be developed in an iterative manner, it is often felt to be cost-effective to have a member of the scientific staff carrying out the programming tasks.

There is an interesting difference of opinion among the different metrology areas concerning whether or not it is cost-effective to rewrite old instrument control software, often written in HT-Basic or HP-Basic, in LabView. In the Ultrafast theme of the Electrical programme, while the initial results of such rewriting have been satisfactory, and the developers and users like the LabView software, the training cost has been very substantial. It is felt that the cost of transferring all the software within the theme would not be justified, particularly given the existing investment and expertise in HT-Basic. In contrast, the conversion of more than 70 programs in the Thermal programme to LabView is being supported on cost grounds, because of the high cost of maintenance of old programs including the cost of training new staff to use old languages. Thus, it can be seen that the costs are very dependent upon which languages and programming environments the staff are trained in.

In other areas, such as Acoustics, there are concerns about cost of software development, particularly balancing the potential cost of software failure (which could be enormous, harming the laboratory’s reputation for reliable primary standards) against the actual cost of developing the software.

In the Flow Measurement area, the small scientific based laboratories have, in general, tended toward self-generation of software to allow the flexible response required for the wide variation of meters requiring test. The financial burden of professional programming techniques is seen as prohibitively expensive.

In contrast, commercial developers seem to feel that, given the level of complexity of their products, the development costs are quite reasonable compared to the return on investment.

5.1.2.3 Other problems

A major problem comes from the pace of IT developments: metrologists often find it hard to keep pace with developments and find it difficult to make the time to climb the learning curve to mastery of each new language and package that would be useful. Some of the software used requires too much specialist training. It is considered important that the metrologists should be able to make effective use of software without having to continually call upon the services of professional software engineers. To obtain more user-friendly software, there will need to be better requirements capture at the outset, and more attention will need to be given by software developers to usability. Where the use of professional software engineers is appropriate, it is important that they work closely with the metrologists in effective multi-disciplinary teams.

Another problem is that bought-in bespoke software is sometimes unreliable and/or not delivered to schedule. The root cause of these problems is often either inadequate requirements capture at the outset or an insistence on the developer enhancing existing poorly engineered and poorly documented software, in the mistaken belief that this will save time and money.

Commercial developers and in-house software professionals alike complain of projects not being sufficiently well defined in time for project start-up.

Users find configuration control difficult on multi-purpose systems; the prevalence of Windows systems is useful, but it is widely recognised that installation of apparently unrelated software can have an impact on the measurement software due to changes in versions of shared executables.

5.1.2.4 Risks

In many metrology research and development areas, the risks from software errors are not high. The measurement systems tend to be heavily validated as part of the development process. The main costs of software errors would be project delays and additional effort costs.

In contrast, in measurement services areas, risks from software errors are relatively high and the importance of software risk analysis is recognised. In this respect, the most mission critical software is the LIMS or other software to automate the measurement process. Failure of software used to provide any service to customers results in down time and the potential for embarrassment and loss of business for the laboratory. Access control failures may compromise confidentiality, leading to commercial losses for the laboratory or customer. The risk of software errors causing incorrect results to be reported to customers is also serious. In some areas, such as Ionising Radiation and VAM, the risks include safety as well as commercial repercussions for the users of the results. In such cases, the risks are minimised by use of thorough review/test procedures, for example as mandated by the NPL quality system.

Commercial developers are also concerned about the loss of revenue that would arise as a result of the late introduction of a piece of software.

5.1.3 Technical issues for software developers

5.1.3.1 Uses of metrology software

Almost all metrology software can be classified into the following categories of use:

- instrument control and data acquisition,
- data analysis and other processing,
- presentation or visualisation of results.

5.1.3.2 Languages

Leaving aside the use of packages (see Section 4.1.1), the programming languages cited in more than one metrology area are as follows (the number of metrology areas in which each is cited as being used is given in parentheses):

1. Visual Basic (9) - much in use for current software developments;
2. HT Basic (9) - mostly in older instrument control software, but still in use;
3. FORTRAN (9) - includes both FORTRAN 77 and FORTRAN 90, often in conjunction with the NAG Library or other packages;
4. Pascal (8);
5. HP Basic (7) - in old instrument control software, largely superseded;
6. C (5) - not extensively used;
7. C++ (4) - used in some recent developments;
8. Quick Basic (4) - only in legacy software;
9. Turbo Pascal (3) - cited only in the Foundation Programme, Length and Electrical Metrology areas;
10. Modula 2 (2) - used only in the Foundation Programme and the Time and Frequency area;
11. Delphi (2) - beginning to take off in the Acoustics and Time and Frequency areas, for example, to modernise the Pascal code.

Other languages used in individual areas include other variants of Basic in legacy software, Perl, Forth, Visual C++ and MORTRAN.

Metrologists are all too often expected to pick up the rudiments of a new language/software environment as part of project work, with little formal training. It is not unusual for someone to have to program in three or more languages at any given time.

Note that in many cases, the use of conventional programming languages is being replaced by the use of LabView, spreadsheets and other packages. Two other trends affect the choice of programming language: the move away from mainframes and HP systems to PC Windows based software; and the training in object oriented programming that an increasing number of new staff have received before they begin work in metrology.

5.1.3.3 Development methods

The size of metrology software can run from a few lines to over 1 million (e.g. the Flapack software in the Length area). It is not uncommon for programs to be of the order of 20,000 or 30,000 lines.

There is no formal development methodology in widespread use for metrology software. Software development methods reflect the application, programming language, operating system, hardware, personal experience, training and preference of the staff involved.

It is recognised that the software development process used should reflect the risk associated with the software. This is, for example, reflected in the NPL quality procedures for software development, identifying four software integrity levels to which different requirements apply.

The development methodologies employed for bespoke software developed in-house for prototyping in a research context (software integrity level 2) have tended to be informal, somewhat quick and dirty, aimed at getting some working code in order to demonstrate feasibility. Such software is, in practice, often written before the design documentation and the documentation is not comprehensive.

For software that needs to be better engineered (software integrity level 3), developers in most metrology areas use traditional structured programming techniques with the waterfall model (specify, design, code, and test). Developers in a number of areas report using a combination of both the waterfall model and prototype refinement. The waterfall model is used to develop a working prototype and then that prototype is refined as requirements change. The design may well get improved as a result of the code and testing phases. In some projects, there is an emphasis on good module design with simple interfaces, and on sufficiently thorough testing to give high reliability. There is in several areas a growing tendency to use object oriented design. There is also some use of data flow diagrams.

A recent innovation in the Ionising Radiation area is the use of HTML (Hypertext Mark-up Language) and web browsers to develop and test both functional specifications and user requirements.

Benefits of formal development methodology. The use of many software languages is probably inevitable. Thus, there is a large potential benefit of having a language independent description of the software design as part of a formal development methodology, helping users familiar with one language to understand software in other languages and aiding the upgrade of application modules from one language to a later release or to a completely different language as hardware and/or operating systems become obsolete.

Benefits of object oriented approaches. Visual C++ and other object oriented systems have software development tools that in theory help the user design and implement software in a structured and re-usable way, providing another method of capturing a high level description of the software design.

Good external contractors employ some type of formal development methodology and have quality assurance programs in place. An approach of rapid prototyping followed by formal specification and re-engineering is used by one major LIMS developer.

5.1.3.4 Software testing methods

The extent of software testing performed will depend on the risks associated with the software (i.e. the software integrity level). For software used in research (software integrity level 2), verification that the software is giving correct results is often achieved by checking results by hand and/or using a spreadsheet. Software for experimental control, for example using multi-channel analogue and digital measurements, may be tested by direct measurements, checking the results by using another instrument.

For software that needs to be well tested (software integrity level 3), a range of methods are used, including checking the results it produces against published results, hand calculations, a mathematical tool, other software (or previous versions of the same software), another instrument or measurement system, and reference results for reference data, as appropriate. The Flapack software in the Length area was specifically designed to allow all components to be tested using test data, a good example of best practice. In some projects, each command of experimental control software is tested in use with the apparatus. Some software is checked by isolating parts of the software at places which form a natural break or interface in the calculation.

In addition, the user interface and functionality are subjected to test plans executed independently of the development team by software quality assurance engineers. Algorithm checking is performed in the course of software validation which is carried out against published standards.

5.1.3.5 Development problems

The metrologists contacted are generally satisfied with their current software development methods, reporting that they are not aware of any significant problems.

In the moving coil kilogram work in the Foundation Programme, the Modula 2 compiler for the Transputer gives problems that constantly require work-arounds. The solution must be to move away from the Transputer and indeed a new real-time platform is under evaluation.

For instrument manufacturers one of the biggest problems is created by the very large number of environments and configurations in which their software must work (e.g. different PCs, operating systems, and interfaces), which makes testing all combinations virtually impossible.

5.1.4 NPL Measurement Good Practice Guide 5, Software in Scientific Instruments

NPL Measurement Good Practice Guide 5 [4] is widely available within NPL and has been extensively distributed outside through Competing Precisely events and now through the Software Support for Metrology Club. Feedback on its use has been received, as a result of which it has just been revised.

Nevertheless, not all the scientists working in NPL are aware of NPL Guide 5, but it would appear that many of the software development practices are consistent with that document. The software professionals within NPL are very much aware of it and do use it as appropriate.

LGC and NEL have copies of NPL Guide 5, but not many of their scientists are aware of it and there is no requirement for it to be used in any of their software projects.

5.2 *Software reuse*

5.2.1 Management issues

5.2.1.1 Importance

Software reuse is seen as an important means of reducing software maintenance costs. However, in many areas, re-engineering of software, rather than reuse, is the emphasis; existing software is re-engineered for use in a similar context or to solve a similar problem. In other areas, software is reused extensively.

While software reuse is seen as worthwhile, there are a number of factors that work against it. The move from mainframe to work-station to PC often implies a change in supported languages, leaving large amounts of software that have to be re-implemented or discarded.

In some areas, instruments support or have supported non-standard languages leading to incompatibility, limited opportunities for software reuse and hence high maintenance costs. However, most instrument control software is written in HP-Basic, HT-Basic or LabView. In many cases, only minor modifications to old code are required to control different instruments. In some areas, externally-supplied instrument drivers or virtual instruments are used where appropriate and reused as much as possible.

5.2.1.2 Responsibility

The details of software reuse and mixed language programming tend to be left to the scientists concerned or their suppliers, although general guidelines and policy may be provided by management. The use of LabView is, for example, a management policy within several metrology areas for new instrument control software. In other areas, there is a policy to use Visual Basic for new software developments.

There is a formal library of LabWindows instrument drivers in one section of the Acoustics Programme, with a specific quality procedure for its maintenance and a software manager appointed to ensure that the library's use is co-ordinated across the 12 or so machines using LabWindows in the section.

5.2.1.3 Awareness

One issue is the awareness and accessibility of good quality implementations of key functions. There are many instances of key functions being used in different areas in different applications. For example, ellipse fitting is required in the dimensional assessment of fibre optics, co-ordinate metrology and interferometry. At present, there is no obvious way that a metrologist working in one sub-area will know what key functions have been implemented in another sub-area.

Metrologists would welcome more information on the software that is available more generally, but they would need to be assured of or check its compatibility and quality.

5.2.1.4 Legacy software

There are, in most metrology areas, programs in use that were developed before software quality procedures were in place (for example, code relevant to the time dissemination service broadcast from Rugby has been running since 1970). Such legacy software is generally not supported, in part because support is no longer available from the suppliers, but also because of "technology pull" and a management desire to migrate all applications to modern systems. In some cases, legacy software continues to be operated on a "no support" basis, while in other cases it is abandoned altogether when licences are not renewed.

Thus, there is generally a trend to re-write these programs to take advantage of modern platforms, enhance the user interface and generally improve the quality. For example, legacy software is being translated into modern languages or packages, for example Visual Basic or LabView, but with the realisation that further updating may be required as new versions of supporting software are adopted.

Some legacy software suffers from hardware dependencies and will not operate on modern systems. In these cases, it can be difficult to reuse or re-engineer software, and consequently there can be significant overheads and costs involved in upgrading these systems. The NAG library and the IDES library (developed over many years by Prof. George Symm, NPL, now retired) are regarded by many as legacy software because they are available on a central mainframe, but not on the user's PC.

However, most metrology areas have a considerable amount of instrument control software, for example written in HP or HT Basic, which is used to control old instruments which are unchanging. There is accordingly little need for maintenance of the software, despite the fact that it may be in regular use.

In some areas, macros for the production of certificates are written in Word Perfect. This causes a problem because NPL corporate support has switched to Word. It is often not appropriate to try to convert these macros to Word ones, leaving the problem of the maintenance of this legacy software.

5.2.1.5 Year 2000

Most metrologists report no known Y2K problems. However, the Y2K problem has been recognised by management for some time. There is a potential problem in the date-stamping of data and in the preparation of certificates and a concern about the impact on operating systems and major software packages. Thus, NPL and LGC have task groups working on this issue, testing the software on every piece of appropriate equipment. Where systems are found or suspected not to be millennium compliant, a risk analysis is performed. The result may be to replace the system, to make minimal modification to the system to fix the problem, or leave the system unchanged because it is believed that no serious consequences will result.

With the shrink-wrapped software that is used, the scientists concerned are relying upon the suppliers to fix any Y2K problems that those packages contain, whilst the management are pressing for statements of millennium compliance from suppliers. Informal contact with suppliers indicates that the issue has generated substantial administrative costs, simply to respond to the volume of requests for information. As far as can be ascertained, most major companies have made a substantial investment in identifying and solving Y2K problems.

NEL reports that the Y2K problem does not seem to be a big problem in the Flow Measurement area and appears to be under control.

It is not yet known what impact Y2K will have on the delivery of NMS milestones since, in part, it depends on how the relevant suppliers and subcontractors will be affected.

In the moving coil kilogram work in the Foundation Programme, the use of the modified Julian date makes all the important software millennium compliant.

Note that NPL has developed a Y2K tester for radio-controlled clocks which use the time dissemination service broadcast from Rugby.

GPS systems do not suffer from the Y2K problem, but unfortunately there is a similar problem that will affect some receivers. Counters in the receivers roll over at midnight 21st August 1999 and, unless care is taken, the date will appear as 6 January 1980. The problem is well known.

5.2.2 Technical issues for software developers

5.2.2.1 Software reuse strategies

In some instances, software *reuse* takes the form of software *re-engineering*. When it is necessary to produce a program very similar to an existing one, any of the following approaches may be used depending upon circumstances:

a) copy the existing program and make the changes needed;

b) add an additional mode to the existing program to cater for small changes;

c) completely re-design, for example using LabView, instrument control software which was previously produced in a now obsolete development environment.

For example, in the Ultrafast measurement theme of the Electrical Programme, software development revolves around a small number of programs. The approach here is to *upgrade* the software as necessary to include new functionality, rather than to *customise* the software to each measurement task. The approach works well provided appropriate software control mechanisms are put in place.

In the RF and microwave guided wave theme of the Electrical Programme, Visual Basic is used for data acquisition and instrument control software, with the numerical processing of the data performed in FORTRAN. This approach has enabled substantial reuse of high-quality mathematical software routines that have traditionally been developed in FORTRAN. Indeed, the FORTRAN NAG library is widely used within the Electrical area.

Similarly, in several areas, for example in the VAM area, there is a preference for reusing existing compiled subroutines when producing a program very similar to an existing one. It is generally too expensive to re-code from scratch and little value is seen in copying the source code. The old program on which the new one is based is updated as necessary to minimise the software maintenance costs.

Some feel that greater reuse of object code would be possible if there was a more formalised way of organising software libraries.

In the Length area, software libraries are used mainly through functions in packages such as Matlab and Mathematica and CMM packages such as Quindos. Object oriented languages such as C++ are also used, partly because they promote reusability. Many of the objects within the Flapack software (such as temperature measurement data acquisition) are completely reusable. FORTRAN 90 is used to facilitate reuse of FORTRAN 77 code in Dynamically Linked Libraries (DLLs) that can be interfaced with Visual Basic and to re-implement Matlab software to run it on systems that do not have Matlab installed.

In the Time and Frequency area at NPL, it has been decided to use Delphi to facilitate maximum use of existing Pascal code (Delphi is Pascal based).

As more instrument control software becomes written in LabView, the scope for software reuse will increase. Common Virtual Instruments are readily available for downloading from the Internet. There are separate sites for virtual instruments produced by National Instruments and those produced by third parties. There is an implication, but no guarantee, that those produced by National Instruments have been validated and hence should be preferred.

In the Acoustics area, a similar approach is used with instrument control software written in LabWindows (the forerunner of LabView), by adding to a common repository (a formal library) of instrument drivers.

Also in the Acoustics area, some software is being re-engineered to cater for changes in hardware platform. Beam calibration software is being converted from Forth to Delphi and inherited Pascal software is being replaced by Delphi or Matlab, where appropriate.

It is clear that now Excel has become the dominant package in Flow Measurement, a return to software reuse is possible by sharing macros, templates and functions in Visual Basic for Applications. Functions for unit conversion, polynomial curve fitting and calculation of physical properties are beginning to be reused as either functions or add-ins for Excel. Some are of poor quality in programming terms but are reused in many spreadsheets. Others are well constructed, fully verified, algorithms.

Experimentation is about to be carried out at NEL to maintain all standard programs, functions etc., on a central server. This will provide a secure and controlled base copy of vital programs, spreadsheets and templates. A user, on logging on, would download the software required to the PC and, if required, modify it as appropriate; if this is in a way specific to a particular problem, then the version would be destroyed at the end of the project. If, however, the modification was generally useful, it would be properly added to the base version, or archived for possible future use. It would be necessary to log the software change to allow reproduction in the future, should results require re-calculation.

Finally, it is important to remember that reuse requires good documentation and version control; in some projects these have rather a low priority at the moment.

5.2.2.2 Mixed language programming

Mixed language programming is generally not used in the Mass, Thermal Metrology, and Time and Frequency areas. Similarly, mixed language programming is not greatly used in the Foundation Programme, but it is likely to be used in the future in some projects, for example where Modula 2 will need to be used together with C.

A number of projects in the Length area, including the primary CMM projects, use a Matlab-FORTRAN 90-C++ interface to allow for rapid prototyping and efficient implementation.

Mixed language programming is beginning to be used in software development undertaken in the RF and microwave guided wave theme of the Electrical Programme. As described in section 5.2.2.1 above, Visual Basic is used in combination with FORTRAN.

In the Optical Radiation and Photonics areas, much of the programming can be regarded as sequential mixed-language, in the sense that it involves development of results in one language, followed by processing of the results in another. Commonly, the second “language” is actually Excel, although sometimes LabView is used for the processing. There are two reasons for this approach:

a) the pre-processing is often done by trusted legacy software written in a language which is now considered obsolete (e.g. HT-Basic); or

b) the hardware may be provided with its own programming language which is not suitable for the further processing of the results.

In the Acoustics area, mixed language programming occurs when graphical user interfaces, typically written in Visual Basic, are added to existing software.

In the Flow Measurement area, mixed language programming is inherent in the way some packages are used. Numerical calculations are often codes in FORTRAN with the user interface provided through Excel with Visual Basic for Applications. This approach is proving to be very successful in physical properties applications where the calculations are well established in FORTRAN but need to be evoked from a spreadsheet.

In the VAM area, LGC tends not to use mixed-language programming for in-house projects but such programming is employed at NPL and by commercial developers where, typically, the user interface is programmed in Visual Basic, while C/C++, FORTRAN or Matlab are used for lower level functions.

5.2.2.3 Problems

A problem arises with the use of Matlab, for example for developing prototype software to test concepts and help with algorithm design. Matlab is ideal for writing sophisticated and reliable software quickly. Unfortunately, it is generally impossible to develop stand-alone executables or DLLs from Matlab. Furthermore, Matlab software will only run on a system with Matlab installed. This means that for many applications, the Matlab software has to be translated into another language. Experience with the Matlab “compiler” which converts Matlab software into C has not been positive. A similar situation exists with the large amount of reliable FORTRAN library software for which an interface is required to languages and packages such as Visual Basic, LabView and Excel. These two problems are now potentially less troublesome with the advent of FORTRAN 90 because

a) FORTRAN 90 supports many of the features of Matlab, so that translation to FORTRAN 90 is very straight forward,

b) FORTRAN 90 DLLs can be interfaced with Visual Basic, etc., and

c) FORTRAN 90 is compatible with FORTRAN 77, hence facilitating reuse of FORTRAN 77 software.

This solution of using FORTRAN 90 is being applied in some areas, particularly the Length area. However, in some other metrology areas, such as the Flow Measurement area, mainframe-based FORTRAN subroutine libraries are regarded as obsolete and hence they are increasingly being replaced by either the functions in standard packages or by bespoke calculations embedded in spreadsheets.

5.3 Virtual Instruments

The main package used which employs Virtual Instruments is LabView. The only other package employing this concept that has been identified as in use in any metrology area is the Flapack package in the Length area.

LabView is corporately supported by NPL, but is not used by all NMS Programmes run by NPL. Nevertheless, its use is on the increase in those areas where it is used and is under consideration in several other areas. It is currently used in the Foundation Programme, the Length, Electrical Metrology, Optical Radiation, and Thermal Metrology areas, and the gas analysis part of the VAM area. It is not significantly used either by NEL or others in the Flow Measurement area, nor by LGC or others in the chemical analysis part of the VAM area.

In the Foundation Programme, several people have been on LabView training courses within the last year, but they are still not yet as efficient at creating instrument control software as they were with HT or HP Basic. Nevertheless, the use of LabView is sometimes a necessity because of interfacing problems. One difficulty with LabView is that it is poor at real time control, which gives problems in correlating data collected from two parallel systems monitored by separate computers. Precision real-time control requires the use of hardware clocks. Problems experienced in this area arise from LabView either not supporting or else duplicating the functionality of the relevant board software. In principle real-time control should be no more difficult with LabView than with other software packages. Further tips and hints would probably be helpful, although the scientists concerned doubt that they need a formal training course.

Both LabView and Flapack are used in the Length area. While Flapack is designed to work with a number of hardware systems, it can also work independently of any hardware, reading measurement data from files instead of acquiring it from sensors. While this is an example of a virtual instrument package in the sense that it has a purely software representation, it does not properly embody the attributes of visual programming.

In the Ultrafast theme of the Electrical Programme, LabView has been used for the development of production software (software integrity level 3 in the NPL Quality System) within a number of projects. However, although the results are good and developers and users like it, the training costs have been too high to justify converting all the instrument control software used to LabView.

In the Photonics area, with a view to possible future use of LabView, it is considered that there is a need for guidelines in its use and possibly a training course.

In the Optical Radiation area, there are no perceived problems with the use of LabView and it is expected that it will soon become the dominant approach used for instrument control software. Some of the LabView software development is outsourced. The specification is developed in-house and this is passed to the developers who produce the code. The software is tested in-house at the level of individual components of virtual instruments ("sub-VIs"). The software is designed so that the components are stand-alone and can be tested and demonstrated in isolation. Bug-fixes and enhancements may be either done in-house or passed back to the developers, as appropriate. This method of operation relies on the developers gaining an understanding of the experimental application so they can interpret the requirements reliably. The developer must also know how the LabView instrument drivers are accessed, especially as the drivers may have been developed in-house. This has proved to be very cost-effective where the developer has achieved a close working relationship with the scientists involved. Current LabView training courses, run at NPL or in Reading, are considered to be quite good and sufficient for the needs of the staff involved.

LabView is increasingly used in the Thermal area. However, instruments such as lock-in amplifiers present special difficulties for testing and it is not clear whether these are made any easier using the virtual instruments model. Being relatively new to LabView, Thermal Metrology staff are just becoming aware of the issues involved and would welcome guidance. It is unclear to them at this stage whether developing virtual instruments is more or less efficient than developing conventional software modules. For example, there is concern that some efficiency may be lost, since HT Basic was ideal for system control using the IEEE bus.

The largest software development task facing the Thermal Metrology Programme is that of converting more than 70 programs to LabView. The reasons for this conversion include

1) the old programs are becoming increasingly difficult to maintain;
2) new staff and students have experience in LabView, hence new software is being developed in such languages/environments;
3) NPL is giving corporate support to LabView and Visual Basic.

Although LabView is not currently used in the Time and Frequency Programme, it has been used in the development of the caesium fountain which is soon to be inherited from the Foundation Programme.

In the Acoustics area, LabView has not been used until recently. However, LabWindows (a conventional set of modules with a similar functionality) has been in use in the area for 10 years. Some initial observations of the staff involved are as follows:

1) the virtual instrument concept implicitly assumes only one instrument of a given type is attached to the control PC, but this is often not the case at NPL;
2) an advantage is that software design is difficult to separate from system design, hence the virtual instrument concept is a good fit with measurement software in Acoustics;
3) training in the development of virtual instruments would be welcome.

Within the VAM area, LabView is only being used in the gas analysis area, but there it is used for most new instrument control applications. No particular problems are reported with design, documentation or testing.

6. Support for Measurement and Calibration Processes

6.1 Automation of Measurement and Calibration Processes

6.1.1 Foundation Programme

There are no measurement or calibration services within the Foundation Programme, hence the automation of the processes through LIMSs or other means is irrelevant.

6.1.2 Length

There is no significant use of LIMSs in the automation of measurement and calibration processes. There is a noticeable contrast between the sophistication of much of the metrology software (Flapack, Quindos, etc.) and the hand calculation and checking of certificated results and uncertainties. There is a clear need for automation in this area but there are obviously resource constraints. Moreover, simple approaches based, for example, on importing Excel spreadsheets into Word documents are fragile, often with unforeseen side effects. There are also concerns about data integrity using such systems as there are plenty of documented cases of data being wrongly interpreted or corrupted.

6.1.3 Mass

The extent to which measurement and calibration processes are automated within the Mass area is varied. The various approaches are described below.

In Mass and Density, a paper-based system with manual transfer of data between systems is currently used. This is recognised as being time-consuming and expensive to operate, but the costs associated with installing a LIMS and configuring the system to the particular calibration services offered are regarded as considerable. Nevertheless, it is anticipated that more efficient electronic data gathering and transfer between the various software packages and systems used will form an important part of the activity in this area in the future.

In Pressure and Vacuum, there is in operation a successful LIMS which was developed “in-house”, not based upon a LIMS package.

In Force, there is very little software used in the calibration services offered and, in this sense, this area of the programme lags well behind the others. At present, much of the process from data measurement to calibration certificate involves the manual completion of forms. Work is currently underway to introduce a LIMS to help with data management aspects and eliminate much of this paperwork. However, it is anticipated that there will be deficiencies in the system: for example, the inability to automatically prompt customers for calibrations, and to manage scheduling. In the interim, a system called SCIRT (Small Contracts Information Recording Tool), based on an Excel spreadsheet, has been implemented and is used for booking-in and scheduling test work. Guidance (provided by someone experienced in the provision of comparable services in other metrological areas) is seen to be paramount.

6.1.4 Electrical Metrology

It is generally recognised that it is a costly exercise to generate a calibration certificate. A LIMS is seen to provide a possible solution, but *reliability* of the system is an essential requirement. This is particularly so in the RF and microwave free field theme where a very large number of calibrations for customers is undertaken each year.

Within the DC and low frequency theme, a LIMS specific to that activity is currently being used. Elsewhere, a pilot LIMS is being implemented in the RF and microwave free field theme, with the intention of using the experience gained in this area to inform the other themes as well as other metrology programmes at NPL.

6.1.5 Photonics

At present, LIMSs are not used widely in the measurement services within the Photonics area, though this is under review, especially with NPL's corporate study of LIMSs. At present, there are concerns about flexibility and how the system would handle exceptions. There is already some automation of the production of calibration certificates.

6.1.6 Optical Radiation

Over the past two years, a LIMS has been developed within the Optical Radiation Metrology programme, but its use in the measurement and calibration services has only recently begun, on a trial basis in parallel with the existing system which it will replace. This development has been independent of the NPL corporate study of LIMSs, even though both developments are based on the same package, Matrix Plus. The Optical Radiation LIMS development has tried to accommodate more complex requirements. Further refinement of the LIMS is planned within the current Optical Radiation Metrology programme. The experience to date is that it is important to keep field definitions flexible, avoiding over-constraining them. In contrast, the NPL corporate LIMS is seen as being too simplistic for use with Optical Radiation measurement services. A remaining concern with the Optical Radiation LIMS is how to handle non-standard (bespoke) measurement services. It seems likely that the conclusion will be that it is inappropriate to try to use the full functionality of a LIMS for non-standard measurement services, although less ambitious automation can still be used to beneficial effect with such services.

6.1.7 Ionising Radiation

Each calibration and production of a certificate is often specific to the particular case in hand, although there has been some attempt to automate the process, for example in the interrogation of databases and merging of information.

The calibration service based upon the linear accelerator involves disparate measurements and a wide variety of data; it is very labour intensive to track data through summary sheets to certificates. NPL staff involved have tried to use WordPerfect merge but this has proved expensive to maintain. A LIMS might prove to be of use in this area.

6.1.8 Thermal Metrology

At present, LIMSs are not used in the Thermal calibration services are, though this is under review, especially with corporate NPL's study of LIMSs. At present, there are concerns about flexibility, in particular, how the system would handle exceptions. There is already some automation of the production of calibration certificates.

The perception is that existing LIMSs are geared towards high-volume test work. The existing system is manual and based on documented procedures and forms. Needless to say, any system which improved efficiency would be welcome, but there were concerns about cost and training.

Although a manual forms-based system is in place to track measurement service work, the actual measurement is as automatic as possible, e.g., in Humidity, cameras can be used to record digital displays.

6.1.9 Time and Frequency

Three new calibration services are being established and this may lead to a requirement for a LIMS, hence guidance on selection of LIMSs would be timely.

6.1.10 Acoustics

At present, the acoustic measurement service is administered manually via paper forms. This is probably adequate in that it is relatively low-volume but any efficiency savings would be welcome.

Acoustics was reviewing the use of LIMSs and awaiting developments elsewhere at NPL. The key requirement of a LIMS system would be that it should link directly to the production of calibration certificates. A secondary requirement would be that it would be sufficiently easy to use that errors would not be introduced, though the existing manual process also carries the risk of errors.

6.1.11 Flow Measurement

LIMSs are, as far as can be ascertained, not used to any great extent in the flow measurement laboratories in UK or Europe. Moves to automate report and data handling in all laboratories are more a gradual change rather than any purposeful move to establish automation. It is a common feeling that either automating the calibration process or reducing the operator skill level because of automation is not necessarily the correct route as many meter calibrations depend on the operator knowing the process, experiencing the meters being tested, and being able to interpret data as it is produced rather than rely on an automated process where gross errors of curve shape etc. would only be seen on completion and after the meter has probably been removed.

Although automation of data reduction and report preparation is generally seen as a good thing and leads towards reduction in errors, the investment in software to cover the wide range of meter options is definitely the reason for slower than expected progress in this area.

Initiatives from the US led by NIST are being watched with regard to “Teleworking”. This involves allowing access by operators, managers or clients, via the Internet, to the calibration process in real time.

6.1.12 Valid Analytical Measurements

There is no experience of using LIMSs in the gas and surface analysis areas. In the gas analysis area, databases of historical data are used to aid in running the measurement and calibration services, but the intention is to bring in a simple LIMS quite soon. Currently, measurement and calibration services are only just starting in the surface analysis area. Hence, the rest of this section will report on the chemical analysis area in which there is much more experience of using LIMSs.

A LIMS is in essence a database and associated management system designed to store and process laboratory data. This apparently simple definition belies the power and potential of a LIMS which can bring many benefits to a laboratory. As with any major IT system, however, it can prove to be an expensive failure if not planned carefully or if it does not receive the full commitment and support of senior management.

Some of the features a LIMS can provide are:

- Acceptance of sample data entered manually at a terminal;
- Acquisition of data automatically from laboratory instrumentation;
- Storage of data;
- Processing of data (performing calculations on it);
- Controlling access to data;
- Report generation;
- Assistance with work scheduling;

- Provision of information on the status of an analysis (not started, in progress, complete, etc.);
- Tracking of samples through different locations within a laboratory;
- Maintenance of instrument service and calibration records;
- Utilisation of information from, and in return supporting, a corporate Management Information System.

The above list is not intended to be exhaustive but reflects the main uses to which LIMSs are put.

LIMSs started appearing in the mid 1980s. These early systems were mostly based on mini-computers but their appearance at that time was due to the increasing cheapness and availability of PCs which fuelled a growth in laboratory automation and a resulting data explosion. The main idea behind the concept of a LIMS was the desire to integrate the data storage and processing functions being performed, in a typical laboratory, by many disparate systems. For example, the process of checking in samples was often accomplished with manual registers or card boxes; sample reporting was starting to be done using word processors and price information may have been generated manually or via a separate computerised finance system. By bringing such functions together under one computerised information management system tailored to the needs of the laboratory, the expectation was (and still is) that information flows would be speeded up and operational efficiency thereby greatly improved.

Over the years commercial LIMSs have steadily evolved and today a wide range is available to suit almost every need. Networked PC based systems are popular for the small laboratory whilst, for larger systems, there is a move to more fully integrate the traditional LIMS functions with corporate computing functions to produce an enterprise-wide “solution” to a company’s data management needs.

6.1.12.1 LIMSs in use

The main laboratory-wide LIMS in use at LGC is a Beckman CALS (current version 8.4) client server. There are, in addition, two smaller systems, a William Woodard dedicated to urine testing work and one developed in-house using dBase for tobacco work.

6.1.12.2 Other processing methods

Most of the work carried out in the VAM area is concerned with test data, although some calibration data necessarily accompanies this. Data may be processed in a number of ways other than through a LIMS, the commonest of which are:

1. Using a hand calculator;
2. Using a spreadsheet (possibly with the aid of user written macros);
   a) with manual data entry;
   b) with automatic data entry;
3. Using a proprietary data processing computer program;
   a) with manual data entry;
   b) with automatic data entry.

Methods 2b and 3b are often found in association with instrumentation supplied to implement a specific analytical technique. A variant of 1 involves the use of software calculators such as those provided by Windows or some Operating Systems such as Novell Netware.

6.1.12.3 Duration of use of LIMS/other systems

The current Beckman system has been in use for about three years. This system replaced an older Hewlett-Packard LIMS which had been in use for about six years.

6.1.12.4 Advantages and disadvantages

In addition to the general advantages provided by the features outlined above there are other specific benefits. For example, there is a real benefit in areas like DNA testing where there are large numbers of samples and it is necessary to keep track of the chain of custody. In the case of the DNA work, the system has been designed to make sure that the steps of the analysis are carried out in the correct order. It is necessary to ensure, e.g., that a sample is of the correct type, in the correct location and is available for testing. There is a lot of automation to reduce transcription errors, and triggered events to check the validity of the data entered. The database employed is the widely used Oracle system and so there is good potential for linking to other compatible systems.

The main disadvantage is the lack of speed of the system and in some cases data entry cannot be made flexible enough without making it too complex for the user. The system can accept input via terminals or instrument interfaces, but all input is handled through a single event queue and this can lead to bottlenecks occurring. Other disadvantages include the high cost of upgrades and the fact that the software is tightly coupled to Oracle versions. Though effective, the security features provided could also be improved. The system permits 16 levels of access, but these are not really adaptable enough for LGC's purposes.

6.1.12.5 Finding a suitable LIMS

There are many ways of obtaining information on LIMSs and their potential suitability for a particular application. The most fruitful would be:

- Consult the literature - *e.g.* journals such as *Scientific Computing and Automation*, *LIMS/Letter*; trade journals such as *Chemistry in Britain* are also useful sources of information, as are advertising publications like *Labmate* and *International Labmate*.
- Approach other organisations known to be using a LIMS and seek their opinions/comments. This option is best reserved for when a contact is already available in an appropriate organisation and the organisation is not a direct competitor.
- Attend seminars and conferences such as *Pittcon* or the *International LIMS Conference*; the latter is held annually and alternates between different locations in Europe and the USA.
- Search the WWW; all LIMS manufacturers have a presence on the Web and much useful information is available (see for example the *LIMS/Letter* site at http:/www.LIMSource.com - contains links to over 80 LIMS vendors).
- Employ a consultant. This will almost certainly be a necessity in order to confirm the suitability of a potential system, particularly if an organisation has no previous experience of commissioning and installing a LIMS.

6.1.12.6 In-house development

LGC's main LIMS is a commercial product but was partially tailored in-house using SSADM for template development. Given the ready availability of commercial systems, few organisations would now contemplate building their own one in-house. The small LIMS used for LGC's tobaccos work has been developed over many years to meet a particular need not catered for at the time by commercial systems. It is built around dBase, but no particular methodology was used in its development.

6.1.12.7 Adherence to appropriate standards and procedures

This question really concerns the issue of validation and this is something that needs to be given careful thought when the specifications requirements document is being drawn up. Commercial suppliers of LIMSs are well aware of the demands made by regulatory agencies such as the US Food and Drugs Administration and the UK Department of Health. These demands generally involve demonstrating compliance with an international protocol such as GMP [12], Good Automated Manufacturing Practice (GAMP) [13,14], Good Laboratory Practice (GLP) [15,16], or Good Automated Laboratory Practices (GALP) [17].

The software for the Beckman system was designed under an ISO 9001 [3] regimen. Before accepting the system, LGC ran a large number of checks to ensure that the software manipulated data in accordance with its design, i.e., the correct tables were updated, calculations were consistent and reports pulled data from the correct fields. Every screen, entry and customised code module was exercised during testing.

LIMS suppliers are often able to provide a basic validation service but the onus is still on the user to ensure that a validation meets their full requirements. Issues such as the provision of an adequate audit trail facility and security arrangements appropriate to their needs will have to be addressed.

6.2 *Format Standards for Measurement Data*

6.2.1 Format Standards for Length

There are a number of data standards, relevant to Length, including the Dimensional Measurement Industry Standard (DMIS) for CAD and CMM data. This is a major activity promising the easy exchange between CAD, Computer Aided Manufacturing (CAM) and Computer Aided Inspection systems as well as graphics and visual modelling systems.

At the laboratory level, there are problems in converting CMM data files into suitable ASCII (American Standard Code for Information Interchange) files for use by other systems and special software has to be bought or developed in order to perform the conversion. Post-editing required in order to extract the required co-ordinates, a process subject to human error.

There are also issues concerning the specification of geometric elements, with different enterprises using different methods of specifying a cone, for example. This means that two sets of completely different numbers could specify the same geometric object according to different conventions and to compare these numbers properly, it is necessary to derive the correct conversion formulae, significant and error-prone activity.

6.2.2 Format Standards for Time and Frequency

Since there are currently no calibration services in the Time and Frequency area, there is no requirement for standard formats for calibration data. Data relating to work on GPS time transfer standards is, however, exchanged with BIPM in Paris and the format is rigidly defined. In addition, there are industries that specify their own time and frequency data formats.

6.2.3 Format Standards for Flow Measurement

Standardised data formats have always been a problem within NEL and remain so. A form of data standard is now being introduced in the liquid meter calibration areas, but will not find acceptance until the older software packages can be retired. Generally between laboratories no standard data format is used. The different requirements for data output and the different methods which laboratories use to record the data seems to make a common format very unlikely to come to fruition. Is temperature measured upstream and downstream, only downstream, or at the meter? Different laboratories use different techniques and these do not lend themselves to common data formats.

For intercomparison work, it is usual for the organising laboratory to specify the data it requires and frequently to supply a spreadsheet template. A table of reduced data is normally all that is required, but providing a template gives the common format and avoids overloading the analyses while still retaining enough information to estimate influence factors.

6.2.4 Format Standards for Valid Analytical Measurements

Standard data formats for measurement data are only needed in a few specific areas within the VAM field.

JCAMP-DX is a human-readable data transfer format developed and frequently used in interchange of mid-infra-red and other chemical spectroscopic and chromatographic data.

In the surface analysis area, a standard format has been devised for storing and exporting data from the manufacturers' instruments. This standard format, developed by NPL, has now been standardized as ISO 14976 [8]. This standard has found support among most manufacturers and has facilitated software development by them. It has shown many benefits of having software which works with data from many different instruments.

There is generally no need in the VAM field to share extensive amounts of raw measurement data with other organisations, except in the area of surface analysis. For some purposes, such as in Proficiency Testing schemes and for inter-company collaborations, there is a requirement to share processed data or small amounts of raw data. Where this is the case, data are usually stored and transmitted in the form of ASCII files or as spreadsheet files.

There is a perceived need in some quarters for a common format for electronic exchange of measurement data. Much work has been done, particularly in the USA, on the development of a common exchange format for chromatographic data. A protocol known as netCDF has been produced which is based on the Common Data Format of the National Centre for Atmospheric Research at Boulder in Colorado. This protocol is not as yet in common use.

In most parts of the VAM area, it is felt that the existence of different proprietary data formats in different measuring instruments is not a factor which would inhibit their use. A LIMS would circumvent this apparent problem through the use of data translators provided by the LIMS supplier. In addition, many instruments offer users a choice of alternative formats such as ASCII or CSV specifically to enhance portability of data.

6.2.5 Format Standards for Other Metrology Areas

There is no apparent need for any standard formats for measurement data used in any other metrology area. There is no apparent demand for a common format for calibration records, although there might be some benefits in such a format, nor is there any apparent need for a common format for electronic exchange of measurement data. Furthermore, there are no identified problems with the different proprietary data formats found in different measuring instruments.

The use of Excel for processing measurement data is seen as diminishing the need for format standards for the data. In other cases, the formats used are determined either by the instrument manufacturers or by the software packages which operate on the data.

Measurement data are shared with other NMIs particularly in key comparison exercises. In these cases, simple text files (usually with headers to explain what each contains) and Excel spreadsheets are used at present. There is doubt that a generic solution exists.

7. Suggestions for Future Activities

7.1 Foundation Programme

7.1.1 Expected Benefits of the SSfM Programme

The areas of the SSfM programme which are most likely to benefit the Foundation Programme are modelling techniques (particularly continuous modelling), model validation, visual modelling and data visualisation, data fusion, the results of testing relevant software packages, and virtual instruments.

It is possible that some improvements to the modelling techniques or model validation used in some Foundation Programme projects may be identified. Visual modelling of microwave fields could benefit the design of resonators and chokes. The moving coil kilogram work could benefit from improved data visualisation and from automated data fusion techniques. If significant errors are found in relevant software packages then these can be avoided or overcome in future. Further guidance on the development of virtual instruments should assist with the future use of LabView in instrument control.

As regards delivery mechanisms, there is doubt about the need for formal training courses. On the other hand, seminars, workshops, club meetings and awareness articles, plus intranet activities such as the web site and news groups, are considered to be effective means of providing the tips, hints and other guidance that are needed.

7.1.2 Case Studies and Feasibility Studies

The continuous modelling undertaken within the Foundation Programme should be investigated to see whether it could provide an example of best practice to be discussed in the report on continuous modelling, although there are some concerns that it may be too specific to provide any useful transferable lessons.

In the moving coil kilogram work, there are potential case studies in model validation, measurement system validation, and uncertainty estimation.

7.2 Length

7.2.1 Expected Benefits of the SSfM Programme

Benefits for the Length Programme arising from the SSfM programme are expected in the following areas:

- Modelling: for example, the development of error separation techniques, empirical models in two or more dimensions, orthogonal distance regression;
- Data fusion: for example, the development of generic approaches to fusing prior calibration information with current measurement data;
- Model validation: for example, rules for selection from a family of empirical models, measurement strategy validation, estimator validation;
- Metrology software development methods;
- Metrology software reuse and the creation of libraries of key functions;
- Automation of measurement and calibration processes, particularly in the preparation of certificates.

7.2.2 Case Studies and Feasibility Studies

Possible case studies and feasibility studies in the following areas have been identified:

- Analysis of roundness calculations from the point of view of model validation, measurement strategy validation, instrument validation, estimator validation and uncertainty evaluation;
- Analysis of refractive index correction from the point of view of generalised regression, data fusion, high dimensional empirical modelling, estimator validation and uncertainty evaluation;
- Visual modelling in designing measurement strategies that take into account line-of-sight constraints in multilateration systems;
- Data fusion relating to multi-station co-ordinate measuring systems and the incorporation of prior calibration information;
- Substitute feature calculations, refractive index calculations, fringe analysis as key functions in a metrology software library.

7.3 Mass

7.3.1 Expected Benefits of the SSfM Programme

Benefits for the Mass Programme arising from the SSfM programme are expected in the following areas:

- Modelling: for example, the development of robust and reliable estimators for analysing measurement data;
- Uncertainty estimation: particularly covering correlation, and the analysis of data from interlaboratory comparisons;
- Data fusion: for example, in the analysis of data provided by a variety of instruments;
- Support to software development by way of guidance and tools to help the development process;
- Support to system development for the automation of measurement and calibration processes, particularly with regard to the integration of different software systems.

7.3.2 Case Studies and Feasibility Studies

Possible case studies and feasibility studies in the following areas have been identified:

- The application of robust methods for the statistical analysis of measurement data including that obtained from interlaboratory comparisons;
- The application of visual modelling and/or data visualisation to understanding piston-cylinder deformation in pressure balances;
- The application of data fusion techniques to the maintenance of a unified high pressure scale;
- The application of LIMS technology to the calibration services provided in the Force area.

7.4 Electrical Metrology

7.4.1 Expected Benefits of the SSfM Programme

Benefits for the Electrical Programme arising from the SSfM programme are expected in the following areas:

- Modelling: for example, with regard to handling errors in all variables;

- Uncertainty estimation: particularly covering complex-valued variables, correlation, robust estimation, and the analysis of data from interlaboratory comparisons;
- Validation of mathematical software packages: particularly, their in-built functions;
- Model validation: for example, the use of refinement in validating finite element model solutions;
- Support to software development by way of guidance and tools to help the development process.

7.4.2 Case Studies and Feasibility Studies

Possible case studies and feasibility studies in the following areas have been identified:

- The application of robust methods for the statistical analysis of measurement data including that obtained from interlaboratory comparisons;
- The application of visual modelling techniques to the presentation of large quantities of data representing correlated complex-valued quantities, including the presentation of such models on interactive Web pages;
- The transfer of LIMS experience from the DC and low frequency and RF and microwave free field themes to the other themes in the Electrical area and possibly other metrology fields.

7.5 Photonics

7.5.1 Expected Benefits of the SSfM Programme

The key benefits that the Photonics Programme might be expected to gain from the SSfM programme include:

- Guidance and seminars on modelling and validation, particularly of continuous/hybrid models;
- Guidance on uncertainty estimation, in particular for correlated data;
- Guidance on testing third-party packages;
- Guidance on visual modelling.

7.5.2 Case Studies and Feasibility Studies

It is clear that within the Photonics area there is valuable experience in the practical use of statistics, with emphasis on non-Gaussian statistics, and the automated use of Word and Excel in the production of calibration certificates. These experiences are likely to benefit the SSfM programme in general, and are expected to feed into the uncertainty project and the competitively tendered project on “Automation of Measurement and Calibration Processes”.

7.6 Optical Radiation

7.6.1 Expected Benefits of the SSfM Programme

The areas of the SSfM Programme which are most likely to benefit the Optical Radiation Programme are modelling techniques (interpolation of data values based upon modelling the physics), guidance on uncertainty estimation (for handling correlations and interpolation between data points), visual modelling and data visualisation (for use with appearance), data fusion (also for use with appearance), the results of testing relevant software packages, and guidance and training on developing virtual instruments. However, it is recognised that the benefits expected in each of these areas are either relatively slight or else several years away.

7.6.2 Case Studies and Feasibility Studies

There is a problem in how to deal with arrays of equally spaced spectral data that have to be integrated with a defined weighting function, particularly relating to the evaluation of uncertainties. The problem is compounded by issues of error correlation. The problem is made more difficult when a detector is produced which is designed to follow the defined weighting function (e.g. a photometer matching the response of the eye, or a colorimeter seeking to measure primary colours), but where the match is not perfect. This area seems to be a good candidate for a case study.

In addition, it is clear that, within the Optical Radiation Programme, there is valuable experience in the use of LabView and in the development of a LIMS. These experiences are likely to benefit the SSfM competitively tendered projects on “Virtual Instruments” and “Automation of Measurement and Calibration Processes”, possibly providing useful case studies.

7.7 Ionising Radiation

7.7.1 Expected Benefits of the SSfM Programme

Although those involved in the Ionising Radiation area are very conscious of the need for validation of continuous and discrete modelling, it is possible that the SSfM programme may lead to some improvement in the techniques used for continuous modelling and the associated model validation.

As LabView is about to be considered for instrument control in this area, some guidance and training in the use of LabView might be of benefit.

7.7.2 Case Studies and Feasibility Studies

Potential feasibility or case studies have been identified in some aspects of visualisation of Monte Carlo simulation results for radiation transport, the provision of an efficient data management for the calibration service and the conversion of WordPerfect macros for certificate preparation to a more efficient alternative.

In addition, SSfM work on the validation of simulated instruments may benefit from consideration of the work done on the long counter. The recent use of HTML and web browsers for software development may also be of wider interest.

7.8 Thermal Metrology

7.8.1 Expected Benefits of the SSfM Programme

The key benefits that the Thermal Programme might be expected to gain from the SSfM programme include:

- Guidance and seminars on modelling and validation, particularly of continuous/hybrid models;
- Guidance on uncertainty estimation for continuous/hybrid models;
- Guidance on testing third-party packages, e.g., spreadsheets, unit conversion and FEA packages;
- Guidance on visual modelling.

7.8.2 Case Studies and Feasibility Studies

Possible case studies and feasibility studies that have been identified are as follows:

- Validation of spreadsheets used in measurement services work. Spreadsheets used in measurement services require validation and configuration control.

- Procurement of metrology software. This study would examine the technical and cost implications of combining “best of breed” software versus choosing a single acceptable package which is not optimal for all requirements.
- Modelling and validation. This would be a major case study, including aspects of modelling of measurement data, uncertainties and statistical modelling, visual modelling, model validation, validation of numerical software packages.

  Heat is instantaneously applied (by laser flash) uniformly to one side of a plane specimen disc at time $t = 0$. For time $t > 0$ it is radiating in vacuum, and conducting, though very little since it is mounted on pin-tips. The change in surface temperature on one plane surface is monitored and from that and the absorbed laser energy the specimen's thermal diffusivity and specific heat capacity are derived. The present model is one-dimensional and adiabatic (no heat loss from the specimen) and corrections for heat conduction and radiation losses are estimated as part of the model.

  The aim of the case study would be to reduce the present relative uncertainty of several percent by an order of magnitude. This would be achieved by

  a) improved modelling: 2-dimensional (axisymmetric) geometry and including additional physical effects, as indicated by model validation;

  b) improved design of the apparatus (supported by visual modelling to provide near real-time exploration of the available parameter space);

  c) improved uncertainty modelling, taking better account of uncertainties in input quantities, such as laser power, dimensions, etc.

  In addition, the apparatus would need to be applied to more exotic materials, e.g., those that are thermally anisotropic (have directional thermal properties), or are composite (typically, fibre reinforced), have thin film coverings or are fibrous. Such materials pose greater modelling challenges, e.g., fibrous materials have more complicated geometry and convection properties.

  The instrument can be designed to be more flexible, such that it can be used for multi-property measurement. The potential cost benefits need to be weighed against traditional approaches.

  The modelling would form part of the NPL Thermophysical Properties Section’s proposed enhancement of an apparatus that is used world-wide for thermal diffusivity measurement, so it would be important for both metrological and industrial applications.

  In addition, other NMS programmes have interests in the modelling of heat flow and dissipation, particularly in Ionising Radiation, where a project is under way to model heat flow from an irradiated graphite plate, such that thermistor measurements can be corrected for indirect heating by surrounding parts of the graphite block.

7.9 Time and Frequency

7.9.1 Expected Benefits of the SSfM Programme

The areas of the SSfM programme which are most likely to benefit the T&F area are modelling techniques, model validation, visual modelling (possibly), software development methodology and guidance on LIMS. Some thought is being given to the dissemination of data through the web, so the SSfM tasks related to Java and the web are also relevant.

7.9.2 Case Studies and Feasibility Studies

Development of a GPSDO Common View time transfer service and a possible web application for distributing time corrections could be considered as case studies. The problem of visualising large amounts of data could provide a case study for the Visual Modelling and Data Visualisation SSfM project.

7.10 Acoustics

7.10.1 Expected Benefits of the SSfM Programme

Generally, the Acoustics Programme makes sophisticated use of mathematics and software. It has particular strengths in modelling and measurement systems validation, each of which could provide exemplars of existing good practice.

The direction of technology transfer might well be reversed in the case of

1) visual modelling;
2) data fusion;
3) validation of simulated instruments;
4) virtual instruments;
5) automation of measurement and calibration processes.

Taking these in turn:

1) Noise and Audiometry would be expected to benefit from visual modelling, e.g., determining the spatial or temporal distributions of sound. Usability of packages such as PAFEC would be improved if the user could *interact* visually.
2) There is considerable scope for data fusion, e.g., from arrays of microphones and when combining historical data with measurements.
3) The simulation of parts could be integrated to simulate entire measurement systems, including the effect of introducing different artefacts, generating example data, etc.
4) Acoustics is relatively new to virtual instruments and might benefit from the experience of others.
5) The same applies to LIMS, where it would appear that the full potential of LIMS has yet to be appreciated in physical metrology.

7.10.2 Case Studies and Feasibility Studies

Case studies might include:

- Validating FE models;
- Simulating acoustic measurement systems;
- Updating software and moving platform: estimating and controlling the project time and cost.

7.11 Flow Measurement

7.11.1 Expected Benefits of the SSfM Programme

It is believed the Flow Programme should benefit from the SSfM programme through the uncertainty project, the spreadsheet management techniques and perhaps the software control areas. It may also benefit from the investigation into format standards.

7.11.2 Case Studies and Feasibility Studies

The feasibility of the programme identifying software techniques and tools for data handling or analyses used in some areas which could be packaged for use in other areas is attractive.

Packages from one area could be passed on or expertise identified to "polish" in-house metrology and data routines for other areas to use. For instance, NEL has functions for density of water (three methods), API oil tables as Excel functions, and tables of unit conversions. These could be made available to other laboratories if time was available under the programme to complete these and other tools to the quality required for third party use. The advantages of such an approach are considerable. Standardisation of calculation routines and the use of common fluid properties are vital, and the availability of a library of routines would aid this process.

7.12 Valid Analytical Measurements

7.12.1 Expected Benefits of the SSfM Programme

If the current SSfM programme leads to improved testability of commercial software supplied with laboratory instrumentation – that is, if the user can inject his/her own data sets to test the underlying algorithms – then this will be a significant benefit. This approach has already had significant results in surface analysis.

While many analysts would wish to check the reasoning and arithmetic behind results produced manually by others they tend to believe that results produced by a computer program are correct. The famous old adage "garbage in, garbage out" is widely appreciated but the possibility that "quality in, garbage out" is not. If the current SSfM programme leads to a wider appreciation among analysts of the limitations of the algorithms employed in commercial data processing software then this too will be of benefit.

Particularly in the gas analysis area, benefits are sought in terms of more guidance on uncertainty estimation, particularly covering non-Gaussian distributions, correlations, using multivariate error analysis (generalised distance regression — i.e., handling errors in all variables). It is also hoped that SSfM will provide support to software development by way of guidance and tools to help the development process.

7.12.2 Case Studies and Feasibility Studies

Two possible areas for case study have been identified:

- the transfer of LIMS experience from the chemical analysis area to the gas analysis area and possibly other metrology fields;
- the transfer of format standards experience from the surface analysis area to other metrology fields.

There is also the possibility of a feasibility study for applying data fusion to electronic noses in the odour work within the gas analysis area.

7.13 Future SSfM Topics

It is suggested that there is scope for further work in the next SSfM programme on uncertainties, visual modelling and data visualisation, data fusion, and validation of embedded software. In particular, there is the suggestion that feasibility studies could be undertaken to apply both data visualisation and data fusion to the measurement of appearance within the Optical Radiation area. The validation of embedded software could be applied to GPS timing receivers, as there is currently little explicit validation of the base data or algorithms used in such equipment.

It is also suggested that the SSfM programme should make a major input in the international effort to provide a sound basis for the analysis of key comparison data in underpinning the equivalence of measurement standards.

Two new ideas for possible future SSfM topics have arisen from the use of advanced mathematical and software techniques, principally in design and modelling, in the Thermophysical Properties activity within the Thermal Programme. These are as follows:

1) multi-objective design, e.g., multiple property measurement;
2) numerical performance of large numerical simulations, e.g., of laser-based gas thermometry and combustion.

Another new idea for a future SSfM topic has arisen from the Time and Frequency area, as follows:

3) There are a number of Time and Frequency metrics used to quantify the stability of time and frequency signals. In the Time Metrology community, the Allan variance is widely used, while in the telecommunications world, concepts such as “jitter” and “wander” are more common. There may be value in the SSfM Programme examining such statistical measures, to identify ways of incorporating these terms into future versions of the GUM.

Two new ideas for SSfM projects on “*approximation methods taking account of scale information*” come from the Acoustics are. These are as follows:

4) Perhaps the next technological step to be taken is to realise that traditional approximation techniques fail to take direct account of data scaling in the independent variable. As an example, PAFEC uses traditional basis functions and hence has difficulty in representing short-scale information such as that which is associated with high spatial frequencies. The result is a very fine discretisation and high computational cost. *Multigrid methods* offer an alternative set of basis functions which use controlled extrapolation to reduce the computational effort.
5) Awareness of scale can also improve discrete modelling. For instance, traditional time series analysis methods fall into two categories: time-based and frequency-based. Some signals e.g., those that are non-stationary, are best represented in a mixed format. *Wavelet methods* utilise a flexible set of basis functions which can be tuned to relatively economical representations of signals which are not easily represented by other methods.

Two further ideas for topics have come from the Flow Measurement area. These are as follows:

6) The development of software tools to fit mathematically sound curve fits and compute associated uncertainties would be extremely useful These tools should be user friendly and give the relatively inexperienced user the ability to fit more complicated functions to data and give confidence that these are sound mathematically and the uncertainty figure is credible. Most packages commercially available for this purpose are written for the experienced mathematician, but what is needed is a much more comprehensible lower level of help for the non-expert.
7) The establishment of a register and validation of Excel based functions and routines associated with metrology would be a useful initiative. Perhaps a database of metrology users’ experience of problems/faults and shortcomings would be possible.

8. Acknowledgements

The author gratefully acknowledges the contributions of Alistair Forbes, Peter Harris, Mark Dainton, Bernard Chorley, Gabriel Lord and Bernard Butler, from NPL, Steve Ellison and Bill Hardcastle, from LGC, and Richard Paton from NEL, in assembling most of the material on which this report has been based. He also acknowledges the review work undertaken by Maurice Cox, Alistair Forbes, Peter Harris, Bernard Butler, Robin Barker, Bernard Chorley, Gavin Kelly and the respective NMS Programme Managers.

Appendix 1: Glossary of Abbreviations

Abbreviation	Expansion
2D	2 Dimensional
3D	3 Dimensional
AC	Alternating Current
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
API	American Petroleum Institute
ASCII	American Standard Code for Information Interchange
BIPM	International Bureau of Weights and Measures
BS	British Standard
CAD	Computer Aided Design
CAM	Computer Aided Manufacturing
CEU	Commission of the European Union
CFD	Computational Fluid Dynamics
CMM	Co-ordinate Measuring Machine
DC	Direct Current
DLL	Dynamically Linked Library
DMIS	Dimensional Measurement Industry Standard
DPLA	Danish Primary Laboratory of Acoustics
DTI	Department of Trade and Industry
DTU	Technical University of Denmark
DVM	Digital Voltmeter
EMC	Electromagnetic Compatibility
GALP	Good Automated Laboratory Practices
GAMP	Good Automated Manufacturing Practice
GLONAS	Global Space Navigation System, Russian equivalent of GPS
GLP	Good Laboratory Practice
GMP	Good Manufacturing Practice
GPS	Global Positioning System
GPSDO	GPS Disciplined Oscillator
GUM	Guide to the expression of Uncertainty in Measurement
HTML	Hypertext Mark-up Language
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronic Engineers

ISO	International Organization for Standardization
IT	Information Technology
ITS	International Temperature Scale
LGC	Laboratory of the Government Chemist
LIMS	Laboratory Information Management System
MAC	Measurement Advisory Committee
MMIC	Monolithic Microwave Integrated Circuit
MoD	UK Ministry of Defence
NAMAS	National Measurement Accreditation Service
NEL	National Engineering Laboratory
NIS	NAMAS Information Sheet
NIST	National Institute of Standards and Technology
NMI	National Measurement Institute
NMS	National Measurement System
NPL	National Physical Laboratory
OTDR	Optical Time Domain Reflectometers
PTB	Physikalisch-Technische Bundesanstalt
R&D	Research and Development
RF	Radio Frequency
RMS	Root Mean Square
SCIRT	Small Contracts Information Recording Tool
SI	International System (of Units)
SSfM	Software Support for Metrology
TR	Technical Report
UKAS	United Kingdom Accreditation Service
VAM	Valid Analytical Measurements
Y2K	Year 2000

Appendix 2: Bibliography

1. *Guide to the Expression of Uncertainty in Measurement*, International Organization for Standardization, 1993, ISBN 92-67-10188-9.
2. *M3003: The Expression of Uncertainty and Confidence in Measurement*, United Kingdom Accreditation Service, 1997.
3. *ISO 9001: Quality systems - Model for quality assurance in design, development, production, installation and servicing*, 1994, Geneva, Switzerland.
4. *Measurement Good Practice Guide 5, Software in Scientific Instruments*, NPL, 1997.
5. *Quantifying Uncertainty in Analytical Measurement*, EURACHEM, 1995, ISBN 0-948926-08-2.
6. *BS ISO TR 5168: Measurement of Fluid Flow. - evaluation of uncertainties*. ISBN 0580 2755 31.
7. *ISO 10360-2. Co-ordinate metrology Part 2: Performance assessment of co-ordinate measuring machines*, 1995, Geneva, Switzerland.
8. *BS ISO 14976: Surface Chemical Analysis - Data Transfer Format*, 1998, Geneva, Switzerland.
9. *IEC 61157: Requirements for the declaration of the acoustic output of medical diagnostic ultrasonic equipment*, 1992, Geneva, Switzerland.
10. *The impact of field bus and sensor fusion on flow Measurement*, Flow Measurement Guidance Note No. 12, 1998.
11. *Intelligent systems for multiphase flow measurements*, Flow Measurement Guidance Note No. 14, 1998.
12. *Rules and Guidance for Pharmaceutical Manufacturers and Distributors*, ISBN 0-11-321995-4.
13. *Good Automated Manufacturing Practice, GAMP3: For Validation of Automated Systems in Pharmaceutical Manufacture*. Available from: ISPE, 3816 W. Linebaugh Ave., Suite 412, Tampa, FL 33624, USA.
14. *Good Automated Manufacturing Practice, GAMP96: Supplier Guide for Validation of Automated Systems in Pharmaceutical Manufacture*. Available from: ISPE, 3816 W. Linebaugh Ave., Suite 412, Tampa, FL 33624, USA.
15. *OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, Number 1: OECD Principles of Good Laboratory Practice* (as revised in 1997), Environmental Directorate, OECD, Paris, 1998.
16. *The Good Laboratory Practice Regulations 1997*, Statutory Instrument 1997:654, HMSO, ISBN 0-11-065105-1.
17. *EPA Directive 2185, Good Automated Laboratory Practices: Principles and Guidance to Regulations for Ensuring Data Integrity in Automated Laboratory Operations*, 1995. Available from: United States Environmental Protection Agency, Office of Information Resources Management, Research Triangle Park, NC 27711, USA.

Document History

Title: Status Report Synthesis

Issue	Date	Author	Description of change	Approved
0.1	22/1/99	D Rayner	First draft for comment by SSfM Project Managers	DR
1.0	31/1/99	D Rayner	Nearly final draft for review by NMSPU	DR
1.1	1/2/99	D Rayner	Final draft for review by NMS Programme Managers, after application of comments from SSfM Project Managers	DR
2.0	5/2/99	D. Rayner	Comments from Programme Managers, Maurice Cox, Bernard Chorley, Robin Barker	DR

Theme name:	Status Report Synthesis
Path and filename:	f:\ssfm\5-status\initial\synthesis.doc

Document control This is a controlled document. It is an authorised version if and only if it is signed or initialled by the project manager or an appointed deputy. Named holders will receive updated copies whenever a new issue is released. Earlier versions of this document must be destroyed or stamped "superseded".

Distribution list

Holder	Organisation	Signed
Dave Rayner	NPL, CISE	
Graham Reed	NMSPU	
Lee Vousden	NMSPU	
Steve Ellison	LGC	
Richard Paton	NEL	
Maurice Cox	NPL, CISE	
Alistair Forbes	NPL, CISE	
Peter Harris	NPL, CISE	
Bernard Chorley	NPL, CISE	
Bernard Butler	NPL, CISE	
Robin Barker	NPL, CISE	
Gavin Kelly	NPL, CISE	
NPL File	ISE/XSP/510	