

# ***THE AUSTRALIAN METROLOGIST***

A publication of the Metrology Society of Australia. ISSN 1321-6082



DOUBLE ISSUE NO 26 AUGUST/NOVEMBER 2001



**A NOVEL PENDULUM**

## ***The Australian Metrologist***

is published four times per year by the Metrology Society of Australia Inc., an Association representing the interests of metrologists of all disciplines throughout Australia.

Normal dates of publication are mid-month in February, May, August and November.

All editorial copy should be sent to the editor by the middle of the month prior to publication.

Further information regarding the Metrology Society of Australia may be found on the Information Page.

## **CONTENTS**

|                                                            |    |
|------------------------------------------------------------|----|
| Advertising rates                                          | 1  |
| President's report to AGM                                  | 2  |
| And the winner is...                                       | 2  |
| Impact of Metrology on the Economy and International trade | 3  |
| Conference Calendar                                        | 11 |
| IMEKO News                                                 | 12 |
| Uncertainty estimation for a Novel Pendulum                | 14 |
| News from NML                                              | 18 |
| Mass and Balance Course                                    | 19 |
| Annual Financial Report 2000-01                            | 20 |
| MSA 2001 in pictures                                       | 23 |
| IIZUKA Prize                                               | 24 |
| MSA Golf Day                                               | 24 |
| From the Dinner                                            | 24 |
| Bayesian Statistics                                        | 25 |
| MSA Information Page                                       | 39 |

## **From the Editor**

In a Conference year there seems to be too many things happening, making it difficult to obtain enough material to keep the TAM issues on schedule.

This issue is a double-sized one, to make up for the missing August issue.

There are two major articles from the conference - from Hratch Semerjian and Ron Cook.

Reports pertaining to the AGM held during the conference are also included.

I have finally printed Bob Frenkel's paper on Bayesian Statistics - it is in 'facsimile' form due to the technical difficulties encountered in trying to change file formats.

Please send plenty of contributions for the next issue as soon as possible to allow me to get the TAM schedule back on track.

Letters to the Editor are always welcome.

I don't have details of the new committee, so the back page information is not up to date.

- Maurie Hooper

## **2001/02 Advertising Rates for The Australian Metrologist**

| Space<br>A4 page | One issue<br>issue | Two issues<br>issues | Three/Four<br>issues |
|------------------|--------------------|----------------------|----------------------|
| Full page        | \$400              | \$750                | \$1050               |
| 1/2 page         | \$225              | \$425                | \$600                |
| 1/3 page         | \$150              | \$130                | \$400                |
| 1/4 page         | \$115              | \$215                | \$290                |
| 1/8 page         | \$ 60              | \$110                | \$150                |
| <b>Colour</b>    |                    |                      |                      |
| Full page        | \$800              | \$1,500              | \$2,100              |

***Closing date for copy to be received for TAM is the 15th of the month preceding publication.***

**Contact the TAM editor for further details.**

Camera ready artwork is to be supplied. Size and specifications are available from the editor. If extra typesetting etc is required an extra charge will apply. MSA members receive a 10% discount when they place advertisements in TAM.

## President's Report to the MSA AGM October 2001

This year has been one of mostly business-as-usual for the MSA, as befits a society that is now well established. A major advance has been the use of teleconferencing, enabling the full participation of interstate members in the committee affairs.

It remains a concern that the society is still top-heavy, with a strong representation from the national bodies such as NML and NSC and less from the private metrology firms further down the calibration chain.

One aspect of the current conference is an appeal to those who work in calibration areas but don't think of themselves as metrologists. The conference is an experiment in that much of the organisation has been handled at a central level interstate, with a local committee drumming up publicity and providing guidelines on the needs of local members and potential members.

It is hoped that this style of organisation will provide a successful model whereby the smaller states can host a MSA conference.

The role of the MSA as a peak body has seen Swinburn University seek and be given endorsement by the MSA for its IRIS graduate metrology course, up and running with the order of 10 students.

MSA interaction was strong during the period of formation of this course.

MSA external interaction has also seen a memorandum of understanding signed with the Brazilian Society of Metrology, and links forged with the newly-formed Society of African Metrologists.

A formal proposal from the MSA Pressure Group on calibration intervals was submitted to and accepted by NATA. The committee have made a firm commitment to the international measurement confederation IMEKO, agreeing to provide significant funding for a 3 year period for subscription as the Australian representative and to provide support for MSA members to attend meetings and be active in the IMEKO technical committees, to be reviewed at the end of that period.

The committee will undergo a significant change at this AGM, with the main operation shifting back to Melbourne.

I'd like to thank all the members of the committee for their support and efforts during this past year. Special thanks also go to the MSA conference committee for their progress and success with this new experimental model for MSA conferences.

- Jim Gardner

## And the winner is ...

The MSA Award, given two-yearly to coincide with the bi-ennial conference, goes this year not to an individual, but to the MSA/IRIS Education Team. The award is given to MSA members for excellence in Metrology.

One of the goals of the society is to foster education in Metrology - an education sub-committee was formed early in the MSA's life.

Strong interaction with Swinburne University in Melbourne has led to the running of a graduate program in Metrology, with a number of people at various stages of completion of the course - the award recognises the foresight and effort involved in this important segment of metrology.

MSA members whose contributions were recognised are Laurie Besley, Walter Giardini, Marian Haire, Stuart McDonald, John Miles, Denis Sexton, and Jane Warne - all will receive certificates. A former member, Helmi Salem, was a significant contributor. On the Industrial Research Institute Swinburne (IRIS) side, important contributors are Brian Costello, Kishor Dabke and Bill Swinson.

During the society's Annual General Meeting, Denis Sexton accepted the award on behalf of the team members. The framed certificate of the award will be passed to Brian Costello for display at Swinburne University.

- Jim Gardner



Jim Gardner

**MSA 2001 - 4th Biennial Conference - October, 2001  
Plenary Lecture****Impact of Metrology on the Economy and International Trade**

**Hratch G. Semerjian and Ellyn S. Beary**  
**Chemical Science and Technology Laboratory**  
**National Institute of Standards and Technology, Gaithersburg, MD, USA**

**Abstract**

*A robust metrology infrastructure is critical for each nation to reap the benefits of a growing global economy. While international trade continues to grow at a rate of 15% per year, tariff-based have been supplanted with technical barriers to trade. Mutual recognition of measurements and standards are essential to overcome these barriers. Advanced technology is another factor promoting economic growth. In the United States alone it is estimated that more than 50% of the economic growth can be attributed to technological advances. This technology requires sophisticated measurement science and standards that are globally recognized. Accurate and reliable measurements are also critical for policy and decision making; they impact the quality of our lives, from the environment we live in, to the food and water we consume, and the healthcare we receive. Retrospective economic impact studies conducted by NIST have demonstrated a high rate of return on past investments, and helped articulate the benefits of metrology programs. Specific examples of studies on the economic impact of physical and chemical metrology programs are presented.*

**Introduction**

Metrology is of critical importance for economic growth, international trade and the quality of life for our global community. The federal role in metrology is becoming ever more important as the health of domestic industries becomes increasingly dependent on global trade. International trade is growing at approximately 15% each year. The metrology infrastructure within each nation is becoming increasingly important to facilitate this trade. In addition, tariff-based barriers are being replaced with technical barriers to trade. To overcome these barriers, measurements performed to characterize products and for conformity assessment must be globally recognized.

Technology is the main driving force for economic growth; according to a recent study, in the United States alone, more than 50% of the economic growth is attributed to technological advances [1]. Advanced technologies such as microelectronics, biotechnology, and nanotechnology, require measurements of higher spatial resolution, sensitivity and selectivity. In addition,

the deregulation of many mature industries, such as natural gas and electricity, require more frequent and accurate measurements to ensure equity in trade.

While metrology impacts equity in trade, industrial competitiveness and international trade, it also affects nearly every aspect of our daily life:

- ◆ Ensuring the quality of the air we breathe requires accurate characterization of the ambient atmosphere, and monitoring of gaseous and particulate emissions from stationary sources as well as transportation systems.
- ◆ Public health is impacted by the purity of the water we drink; therefore, each municipality conducts measurements of inorganic and organic impurities in drinking water on an on-going basis.
- ◆ The quality of the foods we eat and the actual content of packaged foods we buy are of critical importance for us all; regulations about nutritional labeling and pesticide and herbicide content of food require reliable measurements.
- ◆ The quality of our global environment, and critical policy decisions made by the international community, e.g., on ozone depleting CFCs, global warming gases, acid rain, etc., require high quality data to provide the scientific basis for these decisions.
- ◆ The increasing cost of quality health care requires accurate and reliable measurements, both to ensure accurate diagnosis, and to avoid unnecessary spending on treatment.
- ◆ Accurate and reliable metering for utilities and services such as electricity, natural gas, and water to our homes, and the gasoline we buy at the gas station is essential to ensure that fees are fair and equitable.

All of these emphasize the importance of metrology, and the necessity of accurate and reliable measurements that are traceable to national standards, and are recognized in the global marketplace.

Globally Recognized National Standards: National Metrology Institutes (NMIs) throughout the world are responsible for providing their respective national standards. In addition, NMIs play a unique and pivotal role in ensuring the comparability of physical and chemical measurements worldwide. While NMIs have been col-

laborating and carrying out international comparisons of their national measurement standards for more than one hundred years, recent years have seen the development of formal mechanisms to demonstrate comparability of measurements and standards among nations. NMIs must continue to advance the state of measurement science to support the needs of their own domestic industry, as well as to interact on a global basis with other NMIs to be recognized as being among the world's leading measurements and standards institutions. These efforts facilitate the harmonization of systems of measurement and standards and eventually lead to mutual recognition of these systems among trading partners.

***Metrology Investments in the United States:*** In the U.S., the federal government has the constitutional responsibility to provide the weights and measures for the nation, and this responsibility is vested in NIST. NIST invests approximately \$500 million per year (0.7% of the federal R&D funds) in metrology programs, and supports a \$10 billion private sector investment in measurements and standards. More than half of the \$7.6 trillion per year U.S. GDP in sales is supported by this measurement infrastructure.

The measurement and standards infrastructure developed by NIST is designed to address needs of rapidly growing, high technology industries, as well as mature industries. NIST provides measurements and standards for these industries not only to provide better tools to characterize products, but also tools to improve processes, enhance process yields, and promote competitiveness in the global market.

While 0.7% of U.S. R&D investment is a small fraction of the total budget, it is not a small investment; on the other hand, it is not enough to meet all the needs of the U.S. economy. Therefore, NIST continually evaluates its programs and products to be sure that they are effective, address important segments of the economy, and provide a significant return on the nation's investment.

A formal way to measure the results of NIST metrology R&D is through an ongoing program of economic impact studies. Evaluations of the economic impact of NIST's R&D in specific technical areas are carried out through NIST-commissioned studies, mostly performed by external contractors. These studies provide both qualitative assessments and quantitative estimates of the economic impacts resulting from the several categories of technology infrastructure that NIST provides to U.S. industry. Quantitative estimates are provided both as benefit-cost ratios (BCR) and as rates of return to the nation or social rate of return (SRR). The BCR an indication of the investment industry would have had to make if NIST did not perform the work. A list of all formal studies conducted by NIST is shown in the table. Additional information is provided on the

NIST webpages at [www.nist.gov/director/planning/studies.htm](http://www.nist.gov/director/planning/studies.htm).

The results of studies conducted to date have consistently shown high rates of return from NIST research, and help quantify the impact of NIST programs. While by their nature these studies are retrospective, they also contribute to future strategic directions and planning. Several benefits of NIST activities to industry are quantitatively evaluated including: lower transaction costs, lower compliance costs, energy conservation, increased R&D efficiency, increased product quality, and enabling new markets. In addition to these formal studies, NIST also collects informal or anecdotal information on the economic impact of its programs. Examples of both formal and informal studies, categorized by industry sector, are given in the following sections [2].

### ***Environmental Measurements and Standards***

An increase in global environmental concerns has prompted the development of several new standards to serve as quality assurance tools for environmental measurement and monitoring worldwide. The global market for environmental technologies is projected to reach \$600 billion by 2010 [3].

***Sediment Standard Reference Materials (SRMs):*** A New York/New Jersey Sediment SRM is the first NIST natural matrix SRM with values assigned for polychlorinated dibenzo-p-dioxins and dibenzofurans, in addition to 52 PAHs, 29 PCB congener, 11 chlorinated pesticides, and 28 inorganic elements including the toxic heavy metals. This standard provides an accuracy benchmark for reliable environmental decision-making. Analytical inaccuracies can result in inappropriate disposal of contaminated material and the resulting adverse health and environmental effects, while unnecessary remediation results in wasted dollars. The U.S. Army Corps of Engineers in the New York District estimate that in the New York/New Jersey area alone the quantity of dredged material is more than  $4 \times 10^6 \text{ m}^3$  per year, with disposal costs of \$30 million/year and containment costs from \$150 million and up to \$600 million per year. About \$1 million per year is associated with testing costs for each federal project in the U.S.

***NIST Traceable Reference Materials (NTRMs):*** In order to respond to the requirements of the 1990 Clean Air Act, regulators, specialty gas producers, and end-users of gas standards generated such a demand for gas SRMs that over 15% of NIST SRM production resources were devoted to this area each year. Therefore, NIST worked with ten commercial specialty gas vendors to develop a protocol for the production and value-assignment of NIST traceable gas mixtures used to implement the "emissions trading" provisions of the Act. The protocol specifies that vendors manufacture a

batch of gas cylinders and measure them against a NIST gas standard. Each vendor sends both the data and one cylinder from the batch to NIST. NIST measures the batch standard against a NIST primary standard and uses the vendor data to provide both quality assurance and value assignment to each cylinder in the batch [4]. Vendors can then use the certified batch to prepare and analyze similar mixtures for sale to end-users. Compared to an average output of 300 SRM cylinders per year, the NTRM program leverages fewer than 100 batch measurements per year at NIST into a total of over 8000 cylinders traceable to NIST. In turn, these NTRMs have been used to produce about 500,000 commercial gas standards, traceable to NIST at a value of over \$100 million. In addition, the timeliness of NIST response has been significantly improved to about three months, compared to an average of two years or more that are required to produce a new gas SRM.

Proficiency Testing for Environmental Laboratories:

Since the 1970's, EPA has conducted semiannual proficiency testing (PT) to assess the competence of over 6,000 public and private sector laboratories to conduct analyses required by the Clean Water and the Safe Drinking Water Acts. In 1998, the cost-free provision of these services was phased out, and replaced by a multi-provider system in which interested states and private companies provide these PT services on a fee-basis. In a government-private sector partnership, NIST's Chemical Science and Technology Laboratory and NIST's National Voluntary Laboratory Accreditation Program (NVLAP) worked with the EPA, States, and other public and commercial entities in order to establish appropriate oversight of this new effort to externalize and improve the nation's environmental laboratory PT programs. NIST provided the mechanisms and tools that are required for PT providers' use in measurement quality assurance and to enable appropriate government oversight of these programs. Released in 1998, the NIST Handbook 150-19, Chemical Calibration: Providers of Proficiency Testing Handbook, describes the technical requirements of this accreditation program [5]. The first group of accredited providers was announced by NIST/NVLAP in October 1999 and there are currently 12 accredited providers covering all 48 chemistry and microbiology PT program fields. This program leverages NIST effort in that the 12 accredited PT laboratories service approximately 6000 environmental laboratories.

Reference Data for FTIR Measurements: Over the last decade, growing concerns about the environment in general and air quality in particular have stimulated the development of cost-effective field monitoring methods. The U.S. market for air pollution control equipment and related services is estimated to be more than \$16 billion by 2001. With FT infrared-based technologies multiple air-borne chemical contaminants can

be measured simultaneously, *in situ* and in real time. However, the accuracy of these methods is dependent upon the availability of high-quality spectral data from a definitive source. The newly released NIST Quantitative Infrared Database is a validated quantitative database that is traceable to NIST primary gas standards. Early evaluation indicates interest from diverse market sectors including: defense related and other government agencies, national laboratories, universities, gas companies, instrument manufacturers, and environmental laboratories.

Reference Data for Alternative Refrigerants: Occasionally, an accelerated R&D program must be undertaken to respond to industry needs that are constrained by set deadlines. Such was the case for NIST's program on the chemical and physical properties of alternative refrigerants used to replace chlorofluorocarbon (CFC)-based refrigerants. Until the past decade, most refrigerants used throughout the world were made up of CFCs. But as a result of research findings on the deleterious effects of CFCs on the earth's ozone layer, a global agreement to phase out the production and consumption of CFCs and replace them with alternative refrigerants was signed in 1987 (the Montreal Protocol).

With the timetable imposed by the Protocol as an incentive to develop new alternatives to CFCs, NIST engaged in research that would allow industry to make the switch to alternative refrigerants in a timely and economic fashion. NIST began by identifying the basic requirements for new refrigerants according to the new rules, and then started research on determining the physical properties of such candidate alternatives. NIST's most effective form of information dissemination has been the REFPROP program, a computer software package that is available through NIST's Standard Reference Data Program. The REFPROP program continues to enable manufacturers and users of alternative refrigerants to model the behavior of refrigerant mixtures in their respective manufacturing processes. The NIST database has been key in developing CFC replacements seven years ahead of the original schedule proposed in the 1987 Montreal Protocol. A comparison of industry benefits with the funding stream of NIST's research program estimated a social rate of return of at least 433%, and a benefit to cost ratio of 4 to 1 [6]. This study did not include the benefit realized through the reduction of damage to the ozone layer that has in turn resulted in the slowing of global warming, and the myriad of environmental consequences.

Reference Photometer for Ozone Measurements: The concentration of ozone in the atmosphere remains a significant issue from both scientific and political perspectives. Ozone, at tropospheric levels, is a health concern and contributes to climate change as a green-

house gas, while stratospheric ozone protects earth from harmful UV radiation. These concerns have led to the establishment of air quality standards and international protocols to reduce the emissions of pollutants that either contribute to tropospheric ozone formation or deplete stratospheric ozone. Since 1983, NIST has provided Standard Reference Photometers (SRPs) based on UV photometry to ten EPA facilities to provide an infrastructure for the calibration and traceability of ozone measurements within the US. The international interest in ozone measurements has prompted eleven national laboratories to acquire NIST SRPs, thus underpinning the accuracy of ozone measurements worldwide.

### **Health Care Measurements and Standards**

Chemical metrology is at the heart of accurate medical diagnosis and the development of measures to improve our health and ensure long life. In the U.S. about 1.5 trillion dollars are spent each year on health care, which is over 14% of the U.S. GDP [2]. More than 25% of these expenditures are for measurements. It is estimated that over one third of these measurements are performed for non-diagnostic purposes, such as retesting, error prevention, and detection limitations.

**Cholesterol Standards:** In the area of cholesterol measurements alone, it has been estimated that measurement uncertainty was on the order of ( 18 % relative in 1969, before any reference materials were available. Over the last three decades, NIST, in cooperation with the College of American Pathologists (CAP), has developed a series of highly accurate and precise methods for a number of clinically important serum constituents, including cholesterol. These methods are recognized by the international clinical laboratory community as "definitive" and have been used to certify a series of cholesterol SRMs. The first pure crystalline cholesterol (SRM 911) was introduced in 1967. Using the definitive method, serum cholesterol SRMs were developed in 1981 (SRM 909) and again in 1988 (SRMs 1951 and 1952).

A formal economic impact study, completed in September 2000, quantifies a portion of the economic benefits associated with these SRMs beginning in 1986. The economic consequences of NIST's Cholesterol Standards Program are experienced at several levels of the supply chain from manufacturers, to network laboratories, to clinical laboratories that ultimately deliver medical services to the consumer. The benefits to industry resulting from the NIST investment have changed over the more than three decades of NIST involvement. However, this analysis timeframe was limited to 1986-1999 covering only part of the program's life cycle, thus biasing the measured impacts downward. The results indicate that NIST has played

an important economic role in support of a national effort to monitor, measure, and control cholesterol levels, thereby contributing to the reduced level of heart disease. In addition, these SRMs have led to a steady decrease in the number of false positives and negatives resulting from clinical laboratory results for cholesterol in blood to between ( 5.5% to 7.2%, relative [7]. This economic impact study estimates a benefit-to-cost ratio of 4.5, and a social rate of return of 154%. The Net Present Value was calculated to be more than \$3.6 million [8]. The formal impact study does not include the impact of misdiagnosis; improved measurements have been estimated to represent a savings of almost \$100 million per year in unnecessary treatment costs, in addition to the lives saved through timely and accurate diagnosis.

**Measurements and Standards for DNA Testing:** NIST is collaborating with this high-technology community by providing the measurements and standards required for powerful, low-cost, hand held devices for convenient, rapid DNA analysis based on capillary and DNA microarray technologies. These devices facilitate the management of diseases such as cancer and AIDS, contribute to new drug discovery, and help reduce costs in the trillion dollar U.S. healthcare industry. In addition, NIST's expertise in MALDI-TOF mass spectrometry supported the commercial development of a rapid DNA sequencer that fosters drug discovery and reduces testing costs. This novel sequencer incorporates new chemistries, robotics, instrumentation and software that reduces the analysis time of a sample from 3 hours to 5 seconds, and a cost of \$300 - \$5000 to a few dollars.

NIST recently released the first Mitochondrial DNA (mtDNA) Standard, SRM 2392. This SRM is used in quality control, sequencing, medical diagnostics, mutation detection and forensic identification. Several genetic diseases such as Chronic Progressive External Ophthalmoplegia, Kearns Sayre Syndrome, and Pearson's Marrow Syndrome have been linked to mutations in mtDNA, and passed on through the generations through the mother's DNA. The NIST standard supports the rapid advances in the field of mitochondrial genetics.

The SRM includes extracted DNA and all information for performing: PCR amplification process, cycle sequencing steps, gel separation, data analyses to determine DNA sequence, and materials to assess the accuracy of the results. In addition, sequences of 58 sets of unique primers are also included to allow any area or all mtDNA to be amplified and sequenced.

In addition to the use of DNA standards in healthcare, they are also used by the forensics community. NIST SRM 2390 and SRM 2391, based on the newer polymerase chain reaction (PCR) technology, have significantly reduced forensic testing costs, while improv-

ing accuracy, and providing legally defensible traceability for this measurement system nationwide. The Department of Justice's ability to implement DNA testing has benefited greatly from NIST reference materials and measurements. As of January 1, 1990, DNA tests had been admitted into evidence in at least 185 cases in over 35 states. By the late 1990s the FBI was performing over 600,000 DNA evidence examinations per year [9]. As of October 1998, all laboratories that receive federal funding for DNA testing are required to follow the DNA Quality Assurance Standards and use a NIST Standard Reference Material (SRM) or a reference material traceable to NIST.

The Y chromosome is in many ways the final frontier of forensic DNA analysis, having the potential to lead to more rapid identification of the perpetrator of a crime, when mixed fluids are present. To keep abreast with this rapidly growing technology, NIST is developing SRM 2395, Y chromosome DNA markers.

### **Measurements and Standards for Manufacturing**

NIST supports U.S. manufacturers in the global market by providing internationally recognized measurements and standards. Information technologies are the mainstay for growth and capital spending. The continued strength of computers, semiconductors, and communications equipment reflect the ongoing digital revolution, which is transforming the ways in which goods and services are produced, and the kinds of goods consumed. The U.S. is highly competitive worldwide in high technology industries, including aerospace in which the growth of exports continues to be strong.

Standards for Semiconductor Wafer Processing: NIST has established the National Semiconductor Metrology Program to focus its broad metrology expertise to aid this critical, high technology industry meet its increasingly stringent requirements. The Semiconductor Industry Association (SIA) reports that from 1990 to 1995 U.S. companies spent an average of 12% of annual sales revenues on R&D. Industry observers predict that semiconductor markets worldwide will grow at about 15% annually from 1998 to 2002, when the total market value should reach about \$300 billion [10]

NIST projects support industrial research and range from improvements in the accuracy of temperature measurements required in the rapid thermal processing of wafers, to the characterization of properties of optical materials needed for the design and implementation of 248 and 197 nm lithographic systems. Mass Flow Controllers (MFCs) are critical to the control of the wide-range of process chemistries used in manufacturing integrated circuits. They are used to deliver process gases for CVD and other processes such as plasma etching. NIST has developed a properties da-

tabase for process gases (such as boron trichloride) that can be used to validate the procedures used to calibrate MFCs [11].

The NIST Low frost-point humidity generator (LFPG) was developed in response to the semiconductor industry's need for accurate water vapor concentration measurement standards. Water vapor is a primary contaminant of process gases used in integrated circuit fabrication. The LFPG provides gas streams with well-characterized moisture content over 5 orders of magnitude from 1 nmol/mol to 100  $\mu$ mol/mol, extending NIST moisture-in-gases standards capabilities by 3 orders of magnitude. With this system, the vapor pressure of water is calculated based on high accuracy temperature measurements. To extend that capability to the pmol/mol range, optical methods that measure vapor pressure directly are under investigation. Methods such as wavelength modulation spectroscopy and single-mode cavity ring down spectroscopy (CRDS) are being developed for quantitative measurements of partial pressure. Thus far, CRDS has been demonstrated to be a robust technique with precisions of 0.3%.

Standards for the Aerospace Industry: Recently developed low level Sulfur in Nickel Alloy SRMs support the approximately \$24B aircraft engine industry by providing the necessary measurement infrastructure to confirm less than 1 ppm sulfur content in high technology materials. Low sulfur content is critical to the material strength, thus reducing failure rate in gas turbines while allowing higher temperature operation and efficiency.

A NIST/industry consortium developed automated processing technology for producing rapidly solidified metal (RSM) powders. These powders are used to make special high-performance materials for jet engines and advanced magnets. Instead of casting these parts by pouring molten material into molds, or first forming billets that have to be machined, the RSM process converts the molten material into a fine powder using inert gas atomization. The powder can then be loaded into molds and, if it is fine and uniform enough, hot isostatic pressing with very little final machining is all that is needed to produce the parts. After six years of work the process was perfected resulting in increased efficiencies and lower costs for this production method. One crucial part of the work was a computer model developed at NIST. This model enabled engineers to simulate accurately the production process defining the relationships between processing parameters and powder characteristics. Model development and validation was possible through in-process sensors and real-time measurements of particle size distribution. One of the consortium member companies was able to re-design its gas delivery system based on the NIST model thereby increasing its output of usable atomized powder by over 40%. At the same time, the new system

consumed less argon gas, reducing operating costs.

**Standards for the Automotive Industry:** NIST low concentration gas standards allow verification of next generation vehicle emission thus supporting R&D in the U.S. automotive industry's long-term goal to develop affordable mid-sized cars that will travel the equivalent of 80 miles per gallon

Another consequence of the Clean Air Act is the development of new gasoline formulations to reduce the emissions from automobile exhausts. Federal law mandates the sale of reformulated gasoline in nine U.S. metropolitan areas with the worst ozone smog: Baltimore, Chicago, Hartford, Houston, Los Angeles, Milwaukee, New York, Philadelphia, and San Diego. Some other cities voluntarily require reformulated gasoline. About 100 billion liters of oxygenated and reformulated gasoline are produced in the United States each year at an added cost of \$825 million over that of regular gasoline. EPA claims that its use has the same effect on the environment as taking seven million cars off the road. NIST has certified twelve gasoline SRMs for various oxygenates and formulations. Reducing the uncertainties of oxygenate determinations from (15% relative, typical of the performance of standard methods, to the SRM uncertainties of a few percent represents an annual savings of over \$150 million based on the material cost of one oxygenate, methyl tert-butyl ether (MTBE) alone.

NIST has also developed specialized measurement methods and produced SRMs for use in determining the precious metal content of automobile catalysts. The availability of SRM 3144 (Rhodium Spectrometric Solution) enabled long-standing disagreements among commercial rhodium standards of up to ( 7% relative to be resolved. Currently the annual value of rhodium demand is nearly \$300 million dollars [12,13]. The discrepancies represented over \$21 million, but the NIST calibration material has reduced that by a factor of 20.

NIST also developed and transferred to the automotive catalyst industry a high-accuracy method to determine platinum-group elements in new and spent catalyst material. Two used-catalyst SRMs were certified for platinum, palladium, rhodium, and lead to uncertainties of less than 1% relative. Results from previously used methods often disagreed or had relative uncertainties of a few to several percent.

The U.S. lubricants industry is estimated to be worth \$5 billion with a projected growth of about 5% per year over the next few years [14]. Industry representatives have identified a standard certified for additive elements in passenger car motor oil as a top priority. The new Lubricant Additive Package SRM provides traceable standards for testing engine wear, and is needed for linking measurements made on finished

products to national or international standards to comply with ISO 9000 and QS 9000 requirements.

**Standards for the Fluid Power Industry:** The Fluid Power Industry is a \$13 billion industry in the United States, with \$1.1 billion in exports. Particulate contamination of lubricants in hydraulic systems is a serious and costly problem common to mobile equipment, vehicles of all types, and manufacturing machines. It is estimated that 85% of hydraulic wear is caused by particles. Optical particle counters are used by this industry to monitor particle contamination levels of fluids, and to test the capacities and characteristics of filters used in hydraulic systems. In the mid-1990's the National Fluid Power Association (NFPA) requested that NIST develop an SRM to replace the existing material used by the industry, and to provide national and international traceability for particle contamination measurements. NIST responded with the development and production of the first SRM and two associated Reference Materials (RMs) for calibrating particle contamination monitoring devices. SRM 2806 consists of a silica material dust suspended at a known concentration in clean hydraulic fluid. The material is certified for the total number of particles greater than a specified size, per milliliter of fluid. This SRM is the basis for calibrations and accepted measurement protocols for NFPA and the International Standards Organization (ISO). Since its release, SRM 2806 is the material required by ISO standard 11171, and is widely used by the Aerospace Industry as well as the Hydraulic Fluid Power Industry for which the standard was designed.

**Standards for the Construction Industry:** A suite of ten new cement SRMs, certified for chemical composition, is being produced by NIST in response to a continuing and increasing demand for cement standards. In the U.S., 45 companies and 118 plants in 37 states produce more than 80 million tons of cement [3]. These ten SRMs comprise six different Portland cements, one white cement, one blended cement, and two different calcium aluminate cements. The SRMs are used by the industry primarily for production quality control in order to meet manufacturing specification. Industry laboratories must demonstrate their competence to perform American Society for Testing and Materials (ASTM) Standard Test Methods, which require high levels of accuracy and repeatability, better than 1% relative for major components. Standards are also used for research and development related to strength, stability and durability of new products, and associated environmental concerns. As the variety of cement products grow, standards must cover a wide range of composition. One of the new SRMs contains slag and fly ash and addresses some of the environmental issues where cement production is viewed as an avenue for productive use of waste materials.

### **Measurements and Standards for Energy Production and Distribution**

NIST has a long history of working with the Natural Gas and Petroleum Industries to provide them with the tools for accurate characterization and distribution of natural gas and petroleum products. NIST has completed a series of projects that have significantly reduced the uncertainty of orifice meter measurements in the metering and sale of natural gas. The NIST work directly impacts the metering of all natural gas pipeline operations in the U.S. (about \$60 billion per year), the processing of natural gas into products such as propane and butane, and the sale and use of supercritical carbon dioxide from natural gas fields. The industry segments benefiting from the research include gas producers, national and local distributors, processors, and the millions of customers of natural gas.

NIST flow measurements have resulted in a database of orifice discharge coefficients developed in cooperation with the American Petroleum Institute, correcting a long-standing error of about 0.25% in the value of orifice discharge coefficient over their range of use [15]. Based on the value of U.S. annual natural gas consumption, the impact of resolving this error in gas metering is estimated to translate into annual savings of about \$200 million in the natural gas industry.

Chemical composition of petroleum and coal has significant impact on fossil fuel costs; accurate characterization of chemical and physical properties is critical to ensure equity in trade and equitable pricing for the consumer. Sulfur content, moisture and calorific value of petroleum products and coal are some of the factors affecting price.

Standard Reference Materials for Sulfur in Fossil Fuels: Sulfur dioxide emissions from stationary sources are carefully regulated because of their environmental impact; hence, the sulfur content of fossil fuels is one of the most important intrinsic factors that determine their price. Environmental regulations require increasingly lower limits on the sulfur content of fossil fuels, and impose large fines for non-compliance. At every stage in the process (mining, transportation, buying and selling, and combustion) the sulfur content of both oil and coal must be determined in order to meet buyer and seller specifications established to meet environmental regulations. The efficient and cost effective movement of coal and oil from the mine and well to power plants and refineries requires precise and accurate determination of sulfur content in two or more laboratories. For equity in trade and the efficient production of energy, it is mandatory that instrumentation in these laboratories be calibrated using accurate standards.

NIST has developed a primary method based on isotope dilution mass spectrometry (IDMS) to determine

sulfur in fossil fuels to an accuracy of better than (0.1% relative). Using the IDMS method, NIST has certified the sulfur content in about 30 of coal and fuel oil SRMs. In 1984, this technique was recognized as one of that year's top 100 technological advances with a prestigious R&D 100 Award. These SRMs provide industry with the primary calibration materials needed for instrumentation used in routine measurements. SRMs also provide industry with a strong traceability link to NIST for such measurements, whether they are for setting the price of fuel or for demonstrating compliance with environmental regulations.

A formal impact study completed in February 2000, performed by independent economic analysts, quantifies a portion of the economic benefits associated with these SRMs beginning in 1984, and projected through 2003. Surveyed industry representatives indicated that NIST SRMs have decreased the level of uncertainty associated with their measurements of sulfur content. This reduction has led to economic benefits throughout the supply chain. Included in the measures of economic benefits are improvements in product quality, production efficiency, and reductions in transaction costs and sulfur emissions to the environment. This study estimates a benefit/cost ratio of 113, and a social rate of return of 1,056%. The Net Present Value was calculated to be more than \$400 million [16].

### **Standards for Instrumentation**

NIST supports analytical instrument manufacturers by providing software packages and databases, in addition to artifacts designed for instrument calibration. In 1998 the value of laboratory instruments shipments was estimated to be nearly \$12 billion, and include gas chromatographs, mass spectrometers, and spectrophotometers [3].

NIST Mass Spectral Database: This database, NIST 98, has approximately 130,000 evaluated spectra for nearly 108,000 compounds, with tested and documented search algorithms. The NIST mass spectral library is installed on >3000 GC/MS instruments per year, over 50% of all the GC/MS instruments sold worldwide. Each instrument is valued at about \$100,000. Between 1988 and 1999, 38 companies, virtually all the manufacturers of GC/MS instruments, have distributed more than 31,000 mass spectral libraries. In 2000, the ability to add user-drawn structures and synonyms to the user data has been implemented.

Optical Filter NTRMs: Over several decades, NIST has produced optical filter SRMs for calibrating the wavelength scale and verifying the absorbance accuracy of UV/visible spectrophotometers. The primary end user of this instrumentation is the pharmaceutical industry, in which QA/QC procedures must show measurement

traceability in order to bring new pharmaceuticals to market. This booming industry, with a worldwide market estimated at over \$300 billion, has placed a strain on the NIST capacity of filter production and calibration. The successful NTRM program for the NIST gas standards, previously described, has been applied to the optical filter standards. Thus far, four commercial producers of optical filters are producing the first NTRMs for chemical spectrophotometry. NIST's active participation in testing and value assignment assures measurement traceability, and compliance with FDA regulations.

**NIST Thermometry Program:** NIST provides both calibration services for and research on thermocouples. Thermocouples are among the most commonly used sensors for monitoring and control of manufacturing processes. The annual sales of thermocouple products sold by the U.S. thermocouple industry (suppliers of wire and thermocouple assemblies) into the U.S. market are approximately \$280 million. The incorporation of these devices into higher levels of product structures across a broad base of domestic industries affects a much larger portion of the manufacturing sector, estimated to be on the order of \$81 billion.

Benefits were estimated based on surveys and interviews of the thermocouple industry. Participants were asked to estimate the additional expenses that would have been incurred if NIST were to cease to provide primary calibration services [17]. NIST's expenditures in the TCP from 1990 to 1993 included support for research on the basic physical properties that underlie the measurement science to incorporate the change from IPTS-68 to ITS-90. For this effort, NIST led the development of the ITS-90 update, and shouldered 60% of the costs with eight other national standards laboratories. Costs over the whole period of the study (1990 - 1996) also include support for R&D on test methods as well as the calibration services themselves. Therefore, there was a significant time during which NIST expenditures on the fundamental and infrastructural aspects of thermocouple principles, measurement, and test methods did not result in immediate benefits to industry. But once benefits are realized, they are substantial, and these estimates do not include the much larger, though diffuse, community of device users. This study, greatly affected by both the short time-line and limited scope, conservatively estimated the social rate of return to be 32%, and a cost-benefit ratio of 2.95 [17].

## Conclusions

A strong measurement and standards infrastructure is critical to ensure equity in trade and a high quality of life, and to facilitate global recognition of measurements to promote international trade and the economic

growth of a nation. In the U.S., NIST works closely with industry, national and international standards organizations, and other NMIs worldwide, to ensure that products and services from the U.S. are internationally accepted. NIST continually assesses the impact of its work by conducting both formal and informal impact studies. To date, the results of formal studies conducted have consistently shown high rates of return from NIST research, relative to both private investments in technology and other public technology investments. Methodologies and approaches developed for assessment of the economic impact of metrology programs may be used as models for articulating the critical importance of metrology for economic growth.

## References:

1. M. Boskin, L. Lau, *Generalized Solow-Neutral Technical Progress and Postwar Economic Growth*, NBER Working Paper, W8023, December 2000.
2. G. Tasse, *R&D Trends in the U.S. Economy: Strategies and Policy Implications*, NIST Planning Report 99-2, U.S. Department of Commerce, April 1999.
3. *U.S. Industry and Trade Outlook '99*, U.S. Department of Commerce/International Trade Administration, McGraw-Hill Companies, 1999.
4. F.R. Guenther, W.D. Dorko, W.R. Miller, G.C. Rhoderick, *Standard Reference Materials: The NIST traceable Reference Materials Program for Gas Standards*, NIST Special Publication 260-126, U.S. Department of Commerce/Technology Administration, July 1996.
5. C.D. Faison, R.M. Parris, S.D. Rasberry, *NVLAP Chemical Calibration: Providers for Proficiency Testing*, NIST Handbook 150-19 U.S. Government Printing Office, Washington DC, June 1999.
6. M. Shedlick, A. Link, J. Scott, *Economic Assessment of the NIST Alternative Refrigerants Research Program*, NIST Planning Report 98-1, U.S. Department of Commerce, January 1998.
7. *Cholesterol Measurement - Test Accuracy and Factors that Influence Cholesterol Levels*, General Accounting Office Report GAO/PEMD-95-8, December 1994.
8. Leech, D.P, Belmont, P.A, *The Economic Impacts of NIST's Cholesterol Standards Program*, NIST Planning Report 00-4, U.S. Department of Commerce, September, 2000.
9. Freeh, L.J., 1997 Federal Judges Conference, Washington, DC May 13, 1997.
10. Finan, W.F., *Metrology-Related Costs in the U.S. Semiconductor Industry, 1990, 1996 and 2001*, NIST Planning Report 98-4, May 1998.

11. Hurly J.J., *Thermophysical Properties of Gaseous HBr and BCl<sub>3</sub> from Speed-of-Sound Measurements*, Int. J. Thermophys., Vol 21, No. 4, pp 805-829, 2000.
12. Johnson and Matthey, Summary and Outlook - Rhodium, 1997.
13. Kitco, Inc., Gold and Precious Metal Price Update, May 1998.
14. *Hart's Lubricants World*, news reports of December 1997 and January 1998.
15. Whetstone, J.R., Cleveland, W.G., Bateman, B.R., and Sindt, C.F., *Measurements of Coefficients of Discharge for Concentric Flange-Tapped Square-Edged Orifice Meters in Natural Gas Over the Reynolds Number Range 25,000 to 16,000,000*, NIST Technical Note 1270, U.S. Department of Commerce, September 1989.
16. Martin, S., Gallaher, M., O'Connor, A., *Economic Impact of Standard Reference Materials for Sulfur in Fossil Fuels*, NIST Planning Report 00-1, U.S. Department of Commerce, February 2000.
17. Marx, M., Link, A., and Scott, J., *Economic Assessment of the NIST Thermocouple Calibration Program*, NIST Planning Report 97-1, U.S. Department of Commerce, July 1997.
18. *Mutual Recognition of National Measurement Standards and of Calibration and Measurement Certificates Issued by National Metrology Institutes*, Edited by the BIPM, Paris, France, September 1999.

## Upcoming Conference Calendar

July 7-12, 2002

15th Biennial Conference of the Aust Institute of Physics

University of NSW, Sydney, NSW, Australia

July 21-25, 2002

Interact 2002: analytical chemistry, chemical metrology, chemometrics, ecotoxicology, environmental chemistry, pharmaceutical chemistry  
Univ. Tech. Sydney, Sydney, New South Wales, Australia

Website: [www.pco.com.au/interact2002](http://www.pco.com.au/interact2002)

September 24-26, 2002

IMEKO Conference on Force, Mass, Torque, Hardness and Civil Engineering Technology in the Age of Globalization

Celle, Germany

Website: [www.imeko.org](http://www.imeko.org)

October 14-18, 2002

1st International IMEKO Conference on Environmental Measurements

Budapest, Hungary

Website: [www.imeko2002.mtesz.hu](http://www.imeko2002.mtesz.hu)

May 13-15, 2002

Sensor 2003, 11th International Trade fair and Conference

Nuremberg, Germany

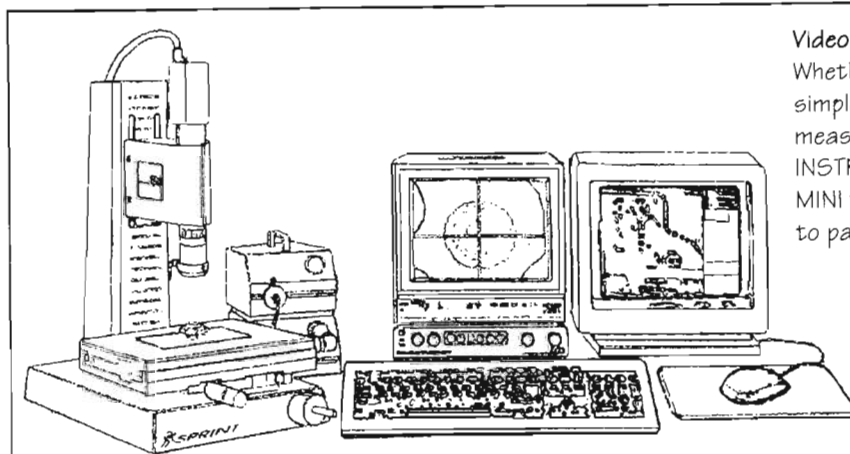
Website: [www.sensorfairs.de](http://www.sensorfairs.de)

June 22-26, 2003

XVII IMEKO World Congress

Dubrovnik, Croatia

Website: [www.hmd.hr/imeko](http://www.hmd.hr/imeko)



**Video Based Co-ordinate Measuring Systems**  
Whether your parts are highly complex or fairly simple you can expect fast, accurate & reliable measuring results with all RAM OPTICAL INSTRUMENTATION systems. From the OMISS MINI to the OMISS III. The Video CMM will start to pay for itself the day it is installed.

Available from:

TESTEQUIP 2000 P/L

Ph: 03 9748 8547

Fax: 03 9748 8086

Email: [te2@ozemail.com.au](mailto:te2@ozemail.com.au)



## News from IMEKO

in Celle, Germany

### Report of the IMEKO Secretary-General, August, 2001

- ◆ Excellent World Congress in Vienna, Austria in June, 2001
- ◆ IMEKO presence at conferences in Dubai, Germany, Hungary, Croatia
- ◆ Internet web address is [www.imeko.org](http://www.imeko.org)
- ◆ Applications or expression of interest in membership from South Africa, Thailand, Iran and Malta
- ◆ Resignation from member organisation in Denmark

### Report of the Secretary of the Technical Board

Activities of the various technical committees since September, 2000

#### **TC1: Education and Training in Measurement and Instrumentation**

- ◆ Jun, 2001: 1st Virtual Workshop on Tools for Education in Measurement & Instrumentation, in Tampere, Finland
- ◆ Sept, 2001: Symposium on Virtual & Real Tools for Education in Measurement, in Enschede, Netherlands

#### **TC2: Photonic Measurements**

- ◆ June, 2002: Symposium in Munich, Germany

#### **TC3: Measurement of Force, Mass and Torque**

- ◆ Nov 2000: co-sponsor of APMP symposium in Tsukuba, Japan
- ◆ Sept 2001: 17th IMEKO Conference on F, M & T in Istanbul, Turkey
- ◆ Sept 2002: Joint symposium with TC5 in Celle, Germany

#### **TC4: Measurement of Electrical Quantities**

- ◆ Sept, 2001: Symposium in Lisbon, Portugal
- ◆ June 2002: Workshop in Prague, Czech Republic

#### **TC5: Hardness Measurement**

- ◆ Sept, 2002: Joint Symposium with TC3, TC20

#### **TC6: Vocabulary Committee**

#### **TC7: Measurement Science**

- ◆ May, 2001: Conference in Smolenice, Slovakia
- ◆ June 2002: Cracow, Poland
- ◆ 2004: Tomsk, Russia

#### **TC8: Traceability**

- ◆ 22-25 Oct, 2001: Workshop with 10th International Metrology Congress in Saint-Louis, France
- ◆ Sept, 2002: Egypt

#### **TC9: Flow Measurement**

- ◆ May, 2003: 11th Conference FLOMEKO in Groningen, The Netherlands
- ◆ 2005: FLOMEKO Conference in UK
- ◆ 2007: FLOMEKO Conference in USA
- ◆ 2009: FLOMEKO Conference in Croatia

#### **TC10: Technical Diagnostics**

- ◆ Sept 2002: 10th Symposium in Budapest, Hungary

#### **TC11: Metrological Infrastructures**

- ◆ No activity in the reported period.

#### **TC12: Temperature and Thermal Measurements**

- ◆ June, 2001: Symposium in Berlin, Germany
- ◆ 2004: Symposium in Zagreb, Croatia

#### **TC13: Measurements in Biology and Medicine**

- ◆ No activity in reported period.

#### **TC14: Measurements of Geometrical Quantities**

- ◆ Sept. 2001: 7th Symposium on Surface Metrology for Quality Assurance in Cairo, Egypt.
- ◆ Sept. 2002: 7th Symposium on Laser Metrol-

ogy in Novosibirsk, Russia

### TC15: *Experimental Mechanics*

- ♦ Aug. 2001: Youth Symposium cancelled
- ♦ March 2002: Youth Symposium to be held
- ♦ Oct. 2001: 3rd International Conference on Experimental Mechanics in Beijing, China

### TC16: *Pressure Measurement*

- ♦ April/May 2002: 1st Conference on Pressure Metrology in Japan

### TC17: *Measurement in Robotics*

- ♦ June. 2001: International Conference on Field and Service Robotics, Helsinki, Finland
- ♦ Aug, 2001: IEEE Conference of Mechatronics and Machine Vision in Practice - M2VIP in Hongkong
- ♦ Sept. 2001: 11th International Symposium on Measurement and Control Robotics, London, UK
- ♦ June 2002: ISMCR Symposium, Bourges, France

### TC18: *Measurement in Human Functions*

- ♦ Sept, 2001: 1st Symposium of the Committee on Measurement, Analysis and Modelling of Human Functions, Sapporo, Japan

### TC19: *Environmental Measurements*

- ♦ Oct, 2002: 1st Conference in Budapest, Hungary

### TC20: *Measurement Techniques for the Construction Industry*

- ♦ Committee is still soliciting membership and will determine its working plan.

### "Measurement"

The latest copies of the IMEKO journal *Measurement* contain the following articles. If any member would like a copy of any article please contact the IMEKO Subcommittee chairman, Laurie Besley (02 9413 7770), who will be pleased to supply it.

Vol. 28, No 1, July 2000 - (Special issue on vibration measurements by laser techniques)

Traceability of vibration and shock measurements by laser interferometry

Laser vibration measurements through combustive flows: application to an industrial burner in working conditions.

Non-invasive measurements of damage of frescoes paintings and icons by laser scanning vibrometer: experimental results on artificial samples and real works of art.

Positional calibration of galvanometric scanners used in laser Doppler vibrometers.

Vol. 28, No 2, September 2000 - (Special issue on ADC modelling and testing).

Third workshop on ADC modelling and testing

Model for the spectral effects of ADC nonlinearity

A 12-bit 125-MHz ADC using direct interpolation

Various scale errors in dithered quantizers: visualisation and reduction

Digital compensation of analog circuit imperfections in a 2-stage 6th-order  $(-)$  modulator.

Issues in the design of a test set-up for high-speed A/D converters.

Influence of disturbance on measurement precision using AD plug-in boards.

Modelling and realisation of high-accuracy, high-speed, current-steering CMOS D/A converters.

Vol. 28, No 3, October 2000

Measurement of extremely long microbores by application of laser metrology.

Fault diagnosis of electronic analog circuits using a radial basis function network classifier.

Measurements of sandy bed scour processes in an oscillating flow by using structured light.

Vision-based measurement of temperature distribution in a 500-kW model furnace using the two-colour method.

Computer vision applied to the automatic calibration of measuring instruments.

On windowing effects in estimating averaged periodograms of noisy signals.

Reliable tool wear monitoring by optimised image and illumination control in machine vision.

Water quality monitoring using a smart sensing system.

## UNCERTAINTY ESTIMATION FOR A NOVEL PENDULUM

R.R. Cook and W.J. Giardini,  
CSIRO NML Melbourne Branch,  
Locked bag 33 Clayton South MDC,  
CLAYTON, VIC., 3169.

## ABSTRACT

In the NML ISO GUM training courses, learning the techniques of uncertainty estimation is enhanced by participation in classroom exercises. Measuring the local value of "g" with a simple pendulum has been used as an example with considerable success for several years. The need for a robust and compact device with well-defined but simply measured parameters led to the development of a novel pendulum. This paper describes the pendulum and a comprehensive uncertainty analysis is presented.

## 1 INTRODUCTION

Galileo Galilei discovered the principle of the pendulum in 1581 from observations he made of the great hanging lamp in the cathedral at Pisa [1]. The traditional pendulum clock used a bob on a slim rod to form a 1 metre long pendulum with a two second period. When fully developed these instruments kept time to better than 0.01 sec per day or approximately 1 part in  $10^7$  [3].

One of the authors, (WJG), realised that a pendulum could be used as a teaching aid for courses on uncertainty of measurement. Instead of a mere sketch on a whiteboard, here was an instrument that was highly visible to the class, easy to set up, had some novelty, and made use of the excellent time interval measuring capability of most wrist watches. The whole class could participate in making the period measurement and a volunteer could measure the effective pendulum length using a tape. The exercise was to measure the value of local g and calculate the uncertainty of the measurement.

## 2 THE CLASS EXERCISE

For a simple pendulum, [2], the period,  $T$ , in seconds, is approximately given by:

$$T = 2\pi \sqrt{\frac{l}{g}} \quad (1)$$

where  $l$  is the pendulum length in metres and  $g$  is the local value of gravitational acceleration,  $\text{ms}^{-2}$ .

This can be rewritten as:

$$g = \frac{4\pi^2 l}{T^2} \quad (2)$$

In the uncertainty estimation, no account is taken of:

the mass of the string, and string-related effects,

the correction for non-simple harmonic motion,

other effects due to the motion of the pendulum through air.

After the class measures the pendulum length and an average period it is soon apparent that the largest source of uncertainty is in the time measurement.

## 3 AN ALTERNATIVE PENDULUM

In reviewing the course, we came to the conclusion that it would be helpful if we had a robust pendulum that did not require being able to tie a string to the ceiling of the lecture room. Further, if the pendulum could measure the value of g to better than 50 ppm it would be useful for calibrators of pressure gauges who use dead weight testers. Of course the international formula for g, given altitude and latitude, already allows the calculation of g to better than this. It does assume a good knowledge of position and height and no local geological anomalies. The experimental verification of g is therefore a valuable quantitative support for the calculation.

Consideration was given to various means of construction. A stiff 1 metre long support for the bob was not practical for air travel, so a more compact solution was sought.

A pendulum in the shape of a thick ring was selected and preliminary calculations showed that even a modest sized ring could have a period that was almost a second. The position of the centre of mass would be determined by the geometry of the ring. It was expected that the ring could be very accurately machined if required. A razor blade was selected for a knife-edge pivot, being readily available and having one of the finest strong edges made. These were somewhat arbitrary decisions.

## 4 THE COMPOUND PENDULUM

A compound pendulum, [2], [3], [4], is one where the mass is distributed and some of it is above the pivot point. The same equation applies as for a simple pendulum except that the equivalent pendulum length,  $l_E$ , is given by:

$$l_E = \frac{I}{Mr} \quad (3)$$

where:  $I$  = moment of inertia of the pendulum,

$M$  = mass of pendulum and

$r$  = distance from axis of rotation to centre of mass.

For an annular body, external diameter =  $2R_1$  and inner diameter =  $2R_2$ , the moment of inertia  $I$  about an axis through the centre and perpendicular to the plane of the ring [4], [5], [6], is:

$$I = \frac{1}{2} M(R_1^2 + R_2^2) \quad (4)$$

This gives a value for the moment of inertia about the knife-edge of

$$I = \frac{1}{2} M(R_1^2 + R_2^2) + MR_2^2 \quad (5)$$

See references [4] and [5]. Here  $R_2 = r$ . Thus the effective pendulum length is:

$$l_E = \frac{1}{2} \frac{(R_1^2 + R_2^2)}{R_2} + R_2 \quad (6)$$

For the prototype,  $R_1 = 82.5$  mm and  $R_2 = 62.5$  mm.

Thus  $l_E = 148.2$  mm.

Note that the effective pendulum length is more than the inner diameter of the ring.

The period of this pendulum is a little less than 0.8 seconds. A measure of 100 swings by the class will yield a good estimate of the period. For the class exercise a set of vernier callipers provides a useable measure of the diameters.

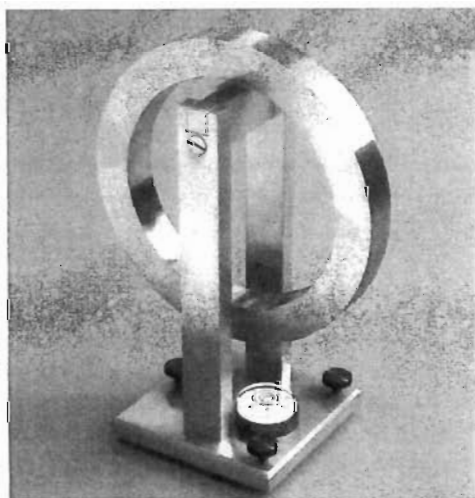


Fig 1. Photograph of prototype pendulum.

## 5 THE NEW CLASS EXERCISE

The rest of this paper will cover the estimation of uncertainty [7] of the value of  $g$  measured using the ring pendulum, using values typical of those obtained in a class. Reference will be made to the best laboratory measurements and the improvements that could be obtained.

The measurand is the local value of  $g$ .

### 5.1 The Uncertainty Model.

The model is given by the equation:

$$g = \frac{4\pi^2 l_E}{T^2} + C \quad (7)$$

where  $l_E = \frac{1}{2} \frac{(R_1^2 + R_2^2)}{R_2} + R_2$ , as in eqn. (6) and

$T$  is the period and

$C$  is the correction for non-harmonic motion.

The effective length,  $l_E$  is calculated from the measured values of the two radii, which are obtained from diameter measurements. The period  $T$  is conveniently measured by timing a large number of swings and dividing this time by the number of swings. The correction  $C$  is amplitude dependent and is calculated using the published formula. We will now consider each term and its uncertainty.

### 5.2 Length, $l_E$ .

This can be readily measured and is not affected by reassembly, relative humidity etc. There is a temperature coefficient to be considered. The string pendulum was ten times longer and hence could be measured with a ten times worse uncertainty for the same proportional uncertainty.

However, compared to a piece of string with its lack of reproducibility, ill defined pivot point and inability to define the centre of mass of the bob particularly well the new arrangement is a significant improvement.

The new shape has many advantages. In class, a pair of callipers can be used to measure each diameter. Because of the varying expertise of the participants there will be significant scatter in the results. In addition to the uncertainty from the calibration certificate we will need to include the ESDM for the mean of each radius. For this paper we will assume the diameter uncertainty components arise mainly from the resolution and scale errors of the callipers and that an adequate diameter uncertainty estimate is  $\pm 0.03$  mm. This is based on manufacturer's claims and experience in making these measurements. The uncertainty in the radius will be half of this or  $\pm 0.015$  mm. The coverage factor will be assumed to be 2 with 50 degrees of freedom. This corresponds to 10% relative uncertainty in the semi-range estimate.

Using the highest grade ring measuring equipment, the diameters can be measured with an expanded uncertainty of about 2 micrometer or about 30 ppm.

There will be two uncertainty elements for the pendulum length, each of  $\pm 0.015$  mm. The sensitivity coefficients, calculated by the numeric differentiation process, are  $87.4 \text{ m.s}^{-2}/\text{m}$  for  $R_1$  and  $41.6 \text{ m.s}^{-2}/\text{m}$  for  $R_2$ .

If the calliper and the ring have stabilised at the ambient temperature, then only the difference in the temperature

coefficients will need to be considered. It is estimated that the difference would not exceed 3 ppm/°C and the temperature change from the calibration temperature of the callipers would not exceed 5 °C. The uncertainty semi-range would thus be 15 ppm or 0.001 24 mm for  $R_1$  and 0.000 938 mm for  $R_2$ . This is small enough to be neglected for the class exercise, however, we have included it here. The limits will be taken as equivalent to two standard deviations with 50 degrees of freedom.

### 5.3 Period

We can estimate the uncertainty using some initial runs as follows. Firstly the time for 100 swings of the pendulum was measured six times. In class many students independently measure the same time. The average time for 100 swings was 77.557 seconds. In a class we would take the mean of all measurements made. The period is then,

$$\bar{T} = 0.775 \text{ 6 seconds.}$$

In class the ESDM is usually about 0.003 seconds with  $n$  the number of measurements typically 10. The values obtained in this trial were ESDM = 0.000 322 seconds and  $n = 6$ .

There is an uncorrected systematic error due to the difference in reaction time in starting and stopping a stopwatch. The worst reaction time is estimated to be 0.3 seconds. If we assume that the variation is up to 50% then there is an uncertainty component of 0.15 seconds. This has to be divided by 100 of course, giving 0.001 5 seconds as the semi-range in this uncertainty component. If we assume this is equivalent to two standard deviations and has 50 degrees of freedom we have all the information required for this component other than the sensitivity coefficient. This can be calculated using the numeric method, which gives  $-25.4 \text{ ms}^{-2}/\text{s}$ .

It is worth commenting that there is correlation between the start and stop timing errors because they are both the reaction time of an individual. We are concerned with the difference in these times and any variability of that difference, so we avoid the need to involve a correlation term.

Using common electronic instrumentation the period could be measured to a resolution of 1  $\mu\text{sec}$ , or 1.3 ppm, for a single measurement, with a better resolution for multiple swings. The uncertainty in the time base of the time interval meter may be as high as 10 ppm for a low cost portable instrument, so this would become the dominant term for the period uncertainty.

### 5.4 Corrections

The motion of a simple pendulum is only an approximation to true simple harmonic motion. In fact the complete equation for the pendulum, while fairly easy

to derive using Newtonian principles, is more difficult to solve.

An approximate formula [3] for a simple pendulum with the first two terms of non-simple harmonic motion correction included is:

$$T = 2\pi \sqrt{\frac{l}{g}} \left\{ 1 + \frac{\alpha^2}{16} + \frac{\alpha^4}{3072} \right\} \quad (8)$$

where  $\alpha$  = the angle from the vertical through which the pendulum swings, in radian.

It can be seen that the period is modified by a correcting factor, which is called the circular correction. The first two correction terms are therefore:

$$-\alpha^2/16 \text{ and } -\alpha^4/3072$$

The formula is adequate up to 20 degrees. Substituting an angular swing of 5 degrees gives a correction of about -0.05% and for 20 degrees the correction is about -0.77 % for the time interval. Figure 2 below shows the correction for up to a 15 degree swing.

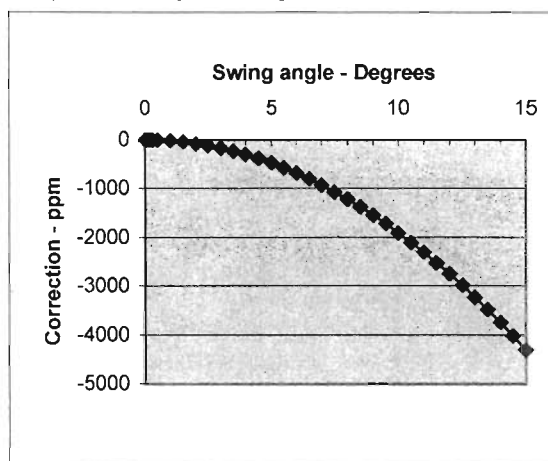


Figure 2. Graph of correction vs swing angle (measured from the vertical)

We need to measure the angle of swing and estimate the uncertainty of the angle. This needs to be done carefully as otherwise we will have a major error source to contend with.

For the simulated class exercise the swing angle was estimated to be 15 degrees with an uncertainty of  $\pm 2$  degrees in that estimation. The correction for 13 degrees is -3,227 ppm, for 15 degrees -4,300 ppm and for 17 degrees -5,529 ppm. That is a range of +1,073 ppm, -1,229 ppm from the 15 degree value. This is one of those interesting asymmetric cases. We will assume a correction of 4,300 ppm with an associated symmetric semi-range of 1,229 ppm or 0.000 953 sec. In this case we will assume a rectangular distribution again with 50 degrees of freedom.

The uncertainty estimate is summarised in the following table.

| Uncertainty Component         | Units | Magnitude<br>U or a | Divisor | Standard uncertainty<br>$u(x_i)$ | Degrees of<br>freedom,<br>$\nu$ | Sensitivity<br>coefficient<br>$c_i$ | $u(y) = c_i u(x_i)$ | $(u_i(y))^2$ | $(u_i(y))^4/\nu$ |
|-------------------------------|-------|---------------------|---------|----------------------------------|---------------------------------|-------------------------------------|---------------------|--------------|------------------|
| Radius $R_1$                  | m     | 1.500E-05           | 2       | 7.500E-06                        | 50                              | 87.4                                | 6.555E-04           | 4.297E-07    | 3.693E-15        |
| Radius $R_2$                  | m     | 1.500E-05           | 2       | 7.500E-06                        | 50                              | 41.6                                | 3.120E-04           | 9.734E-08    | 1.895E-16        |
| Temp Coeff $R_1$              | m     | 1.240E-06           | 2       | 6.200E-07                        | 50                              | 87.4                                | 5.419E-05           | 2.936E-09    | 1.724E-19        |
| Temp Coeff $R_2$              | m     | 9.400E-07           | 2       | 4.700E-07                        | 50                              | 41.6                                | 1.955E-05           | 3.823E-10    | 2.923E-21        |
| Period ESDM                   | sec   | 3.221E-04           | 1       | 3.221E-04                        | 5                               | -25.4                               | 8.182E-03           | 6.695E-05    | 8.965E-10        |
| Period timing                 | sec   | 1.500E-03           | 2       | 7.500E-04                        | 50                              | -25.4                               | 1.905E-02           | 3.629E-04    | 2.634E-09        |
| Correction                    | sec   | 9.530E-04           | 1.732   | 5.502E-04                        | 50                              | -25.4                               | 1.398E-02           | 1.953E-04    | 7.630E-10        |
| SUMS                          |       |                     |         |                                  |                                 |                                     |                     | 6.257E-04    | 4.294E-09        |
| Combined standard uncertainty |       |                     |         |                                  |                                 |                                     | 0.025014            |              |                  |
| Effective degrees of freedom  |       |                     |         |                                  | 91.19                           |                                     |                     |              |                  |
| Coverage factor, k            |       |                     | 1.986   |                                  |                                 |                                     |                     |              |                  |
| Expanded uncertainty, U       |       | 0.0497              |         |                                  |                                 |                                     |                     |              |                  |

Figure 3. Pendulum uncertainty budget.

The measured value of  $g$  is  $9.81 \text{ ms}^{-2}$  with an uncertainty of  $0.05 \text{ ms}^{-2}$  at the 95% confidence level. The coverage factor is 1.99.

The value of  $g$  near where the initial tests were made is  $9.799 \text{ ms}^{-2}$ . The difference is just over  $0.010 \text{ ms}^{-2}$  or 0.1%. This is similar to the results obtained with the simple pendulum.

## DISCUSSION

The expanded uncertainty is about 0.5% of the value of  $g$  and in this example the difference from the "true" value is about 0.1%. It may be that the assumptions in the uncertainty estimate were too pessimistic, but until more measurements are made it cannot be decided if that is so or the result was fortuitous.

Corrections for air and other frictional damping were ignored, as were aerodynamic effects and interactions of the steel pendulum with the earth's and other magnetic fields. These would increase the period, so our calculated value of  $g$  is likely to be biased a little high.

The bending of the pivot point is detectable, but no attempt has been made to analyse the effect. Even when pendulum technology reached its peak, this error source was minimised but not corrected for.

In classroom measurements the uncertainty of the period timing is dominant. When the best possible measurements are made the uncertainty in the correction (or the swing angle) is the dominant component. Even if the correction uncertainty can be reduced to around 10 ppm by using a small amplitude of oscillation and measuring the swing angle carefully, the estimated uncertainty is unlikely to be reduced much below about 50 ppm as the geometric deficiencies are unlikely to allow the effective pendulum length to be established to better than 2  $\mu\text{m}$ , or about 30 ppm.

In class measurements with the simple pendulum, most measurements were within 0.5% of the expected value of  $g$ . The new pendulum solves some problems but at the expense of a more complex measurement model. There is scope for modifications and improvements.

## CONCLUSION

A novel pendulum has been designed, built and tested. As it is more complex than the simple string and bob pendulum it is perhaps less suitable for an initial teaching example. Nevertheless it is quite suitable for use in a training situation, but is not a precision instrument in its current form.

## REFERENCES

- [1] F.K. Richtmyer and E.H. Kennard, *Introduction to Modern Physics*, McGraw-Hill, Ch. 1 p 12, 1947.
- [2] J. Duncan and S.G. Starling, *A Text Book of Physics*, Macmillan and Co., Ch. XVI, pp 224 - 229g, 1939.
- [3] J. Thewlis, *Encyclopaedic Dictionary of Physics*, pp 319 - 326, Pergamon, 1962.
- [4] J.A. Richards, F.W. Sears, M.R. Wehr and M.W. Zemansky, *Modern University Physics*, Ch. 11, p 234, Addison-Wesley, 1960.
- [5] J.A. Richards, F.W. Sears, M.R. Wehr and M.W. Zemansky, *Modern University Physics*, Ch. 9, p 189, Addison-Wesley, 1960.
- [6] H.J. Halstead and D.A. Harris, *A Course in Pure and Applied Mathematics*, Macmillan, Ch. 5, pp 196 - 199, 1963.
- [7] ISO TAG4/WG3, *Guide to the Expression of Uncertainty in Measurement*, 1993.

## NEWS FROM NATIONAL MEASUREMENT LABORATORY

### Progress on the NML trapped ion frequency standard

For several years NML has been developing a new atomic frequency standard based on extremely cold, electromagnetically trapped Ytterbium ions (electrically charged atoms). This has been a long and difficult process, involving evaluating, with very high accuracy, all the uncertainties in the value of the 12.6 GHz microwave resonance in Ytterbium. This microwave resonance frequency is analogous to the RF vibration frequency of a quartz crystal in a crystal oscillator, and to the 9 GHz microwave resonance in Cesium which presently defines the SI second.

In 1995 and 1997 NML made (and published) determinations of the absolute frequency with respect to the SI second of the 12.6 GHz NMR "clock" transition in  $^{171}\text{Yb}^+$  with the results:

(1995) 12 642 812 118.4680  $\pm$  0.0016 Hz

(1997) 12 642 812 118.4664  $\pm$  0.0002 Hz

The uncertainty is mainly due to a complex theoretical model of the thermodynamics of the hot (400 K) cloud of trapped ions and the relativistic time dilation shift due to their thermal motion. The magnitude of this shift is about 1 part in  $10^{12}$  or 0.013 Hz. The uncertainty takes no account of whether the model accurately represents the ion cloud, since at the time there was no way to independently test this.

Recent work has been aimed at eliminating the need for this model, and the consequent uncertainty, by laser-cooling the ions to temperatures of less than 1 K, effectively slowing the ions and setting the magnitude of the time dilation effect to a value smaller than a few parts in  $10^{15}$ .

Two determinations of this frequency, based on a laser-cooled ion cloud at a temperature of less than 0.5 K, have recently been completed. The results are:

Determination 1: 12 642 812 118.4690  $\pm$  0.0010 Hz

Determination 2: 12 642 812 118.4684  $\pm$  0.0009 Hz

These numbers are in good agreement with the previous results, vindicating the theoretical model of the hot cloud. We hope to refine the present uncertainty of 8 parts in  $10^{14}$  to about 5 parts in  $10^{15}$  in the foreseeable future. NML will then have a frequency standard 400 times more accurate than their present Cesium standards.

### Quality systems

NML's RF and Microwave group underwent NATA assessment in July for accreditation under ISO 17025 in the areas of RF power measurement, RF voltage measurement and RF attenuation measurement.

### Two-Way Satellite Time Transfer (TWSTT)

The United States Naval Observatory (USNO) has commissioned a TWSTT station at Vandenberg Air Force Base in California, for the purpose of TWSTT with Asia-Pacific nations. As the initial use of this system, a regular TWSTT link between NML and USNO was established in August, for an initial period of six weeks, with a view to continuing on a permanent basis if funding for satellite usage fees can be arranged.

### Support of regional metrology laboratories

NML staff member Dr Peter Fisk visited the Telecommunications Laboratory, Taiwan, between Sept 4 and Sept, to serve as an external assessor in the ISO 17025 accreditation process of their time and frequency laboratory.

### International metrology committee service

NML staff member Dr Mark Ballico attended a meeting of the BIPM Consultative Committee for Temperature in September 2001. One of the outcomes of this meeting was substantial progress towards the standardization of uncertainty calculations for temperature measurements. NML has made a significant contribution to this work.

### Awards

A paper entitled "Precision Digital Filters for High Voltage Impulse Measurement Systems" written by NML staff members Dr Yi Li and Dr Juris Rungis has received the IEEE Power Engineering Society's Power System Instrumentation and Measurement Committee's Prize Paper Award for 2000.

The International Council on Large Electric Systems (CIGRE) presented NML staff member Dr Vic Morgan with an Award of Merit for prolonged meritorious service to the electricity industry.

### NML Scientist Awarded Japanese Fellowship for Research Collaboration in Electrical Standards

Dr Ilya Budovsky, NML's Project Leader in AC-DC Transfer Standards, was awarded a prestigious three-month Japanese Fellowship to undertake joint research with counterparts at the National Metrology Institute of Japan (NMIJ). Dr Budovsky took up the scholarship over July-September 2001, developing a system for the comparison of the new ac Josephson voltage standards with thermal voltage converters. It is expected that measurement uncertainties with the new system will be 100 times smaller than obtainable with currently available equipment.

### Mass and Balance Calibration Course, 21 & 22 March 2002

The CSIRO National Measurement Laboratory (NML) is offering a two-day intensive course in the theory and practice of mass and balance calibration on 21 and 22 March 2002, to be held at NML, West Lindfield, Sydney. The course will cover general principles of measurement as well as specific techniques in mass measurement. It consists of lectures and a laboratory tour, and will be of value to technicians, engineers, scientists and others involved in or responsible for work in laboratories in which the measurement of mass is important.

The scope of the course includes the calibration of balances and weighing machines, weights (mass standards) upon which these balance measurements depend, mass traceability, the effects of buoyancy, true and conventional mass and the estimation of uncertainties at all levels. The course fees include course participation, two NML Monographs: "Mass and Balance Calibration" and "Uncertainty in Measurement: the ISO Guide", and lunch and morning and afternoon teas.

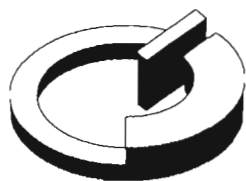
#### Fees:

If received before 4th Feb 2002: AU\$880 (incl. GST)

If received after 4th Feb 2002: AU\$1,100 (incl. GST)

Registrations should be received no later than 1st March 2002.

For more information and registration forms for NML's Mass and Balance Calibration Course or other training programs and courses run by NML, please contact Mr Robin Bentley (Ph: +61 2 9413 7764; Fax: +61 2 7474; E-mail: [robin.bentley@csiro.au](mailto:robin.bentley@csiro.au)).



# METROLOGY SOCIETY OF AUSTRALIA

ABN 802 123 257 48

## Annual Financial Report 2000 – 01

### Treasurer's Report

This financial report represents a period of 12 months from July 1, 2000 to June 30, 2001. The Balance Sheet reflects a healthy financial state showing the society has assets of \$ 59,170.95. The Statement of Receipts and Expenditure for the period shows a surplus of \$ 1,136.85. The small size of the surplus is due to the fact that \$8,000 was spent on the 2001 conference in the financial year to 30 June 2001.

Subscription fees reduced slightly from the previous year with membership reducing from \$14,615 to \$12,305. The cost of producing TAM was less as one less issue was produced during the year. I believe we have reached a critical stage in the societies development. Members need to be vocal about what they expect from the society to ensure it continues to grow and mature while serving the needs of its members.

This is the first year IMEKO fees have appeared on the books of the society. The committee has undertaken to work to ensure this membership is useful to the society. An amount of \$4,500 has been allocated on an annual basis to support MSA involvement in IMEKO. This will cover annual fees and support for representatives at meetings.

I would like to thank the MSA executive committee for their support during the past year. Special thanks to our honorary auditor Bryce Thornton, who is guiding the development of sound financial reporting systems that will stand the Society in good stead as it grows in strength.

Marian Haire

27 September 2001

### Metrology Society of Australia

#### Balance Sheet as of June 30, 200

| ASSETS                                | 2000-2001               | 1999-2000               |
|---------------------------------------|-------------------------|-------------------------|
| <b>ASSETS Current Assets</b>          |                         |                         |
| Cash management                       | 15,430.35               |                         |
| MSA Conference                        | 5,826.24                | 5,776.93                |
| MSA No.1                              | 9,490.75                | 25,372.68               |
| Term Deposit 1                        | 17,532.76               | 16,618.45               |
| Term Deposit 2                        | 10,890.85               | 10,266.04               |
| <b>TOTAL ASSETS</b>                   | <b><u>59,170.95</u></b> | <b><u>58,034.10</u></b> |
| <b>LIABILITIES &amp; EQUITY</b>       |                         |                         |
| Equity                                |                         |                         |
| Opening Bal Equity                    | 35,150.46               | 35,150.46               |
| Retained Earnings                     | 22,883.64               | 26,745.21               |
| Net Income                            | 1,136.85                | -3,861.57               |
| <b>TOTAL LIABILITIES &amp; EQUITY</b> | <b><u>59,170.95</u></b> | <b><u>58,034.10</u></b> |

**Statement of Receipts and Expenditure for year ended 30 June 2001**

|                                 | <b>2000-2001</b>        | <b>1999-2000</b>         |
|---------------------------------|-------------------------|--------------------------|
| <b>INCOME</b>                   |                         |                          |
| <b>CMM Activities</b>           | <u>8,223.72</u>         | <u>0.00</u>              |
| <b>Fees</b>                     |                         |                          |
| Annual                          | 11,640.00               | 13,345.00                |
| Debts                           | 0                       | 30.00                    |
| Nominating                      | 665.00                  | 1,240.00                 |
| <b>Total Fees</b>               | <u>12,305.00</u>        | <u>14,615.00</u>         |
| <b>Interest</b>                 | <u>2,110.33</u>         | <u>1,119.72</u>          |
| MSA 2001 income                 |                         |                          |
| Interest                        | 20.42                   | 5.17                     |
| <b>Total MSA 2001 income</b>    |                         | <u>5.17</u>              |
| <b>Total MSA 99 Income</b>      | <u>0.00</u>             | <u>51,647.70</u>         |
| <b>Other</b>                    |                         | <u>35.00</u>             |
| <b>TAM income</b>               |                         |                          |
| Advertising TAM                 | 290.00                  | 799.00                   |
| Subscriptions TAM               | <u>0.00</u>             | 60.00                    |
| <b>Total TAM income</b>         | <u>290.00</u>           | <u>859.00</u>            |
| <b>TOTAL INCOME</b>             | <u><b>23,174.17</b></u> | <u><b>68,281.59</b></u>  |
| <b>EXPENSE</b>                  |                         |                          |
| <b>MSA 2001</b>                 |                         |                          |
| Bank Fees                       | (13.99)                 | 12.12                    |
| MSA 2001 - Other                | 8,000                   | 724.70                   |
| <b>Total MSA 2001</b>           | <u><b>7,986.01</b></u>  | <u><b>736.82</b></u>     |
| MSA99 Exp                       |                         |                          |
| <b>Total MSA 99 Exp</b>         | <u><b>0.00</b></u>      | <u><b>58,137.39</b></u>  |
| <b>Society General</b>          |                         |                          |
| Bank Fees                       | 252.12                  | 255.64                   |
| Disbursements                   | 1,283.04                | 178.33                   |
| IMEKO Fees                      | 2,500.00                |                          |
| Meetings                        | 475.77                  | 1,149.38                 |
| Office                          | 1,108.99                | 1,087.60                 |
| Other                           | 1,934.91                | 2,069.43                 |
| <b>Total Society General</b>    | <u><b>7,554.83</b></u>  | <u><b>4,740.38</b></u>   |
| <b>TAM</b>                      | <u><b>6,496.48</b></u>  | <u><b>8,528.57</b></u>   |
| <b>TOTAL EXPENSE</b>            | <u><b>22,037.32</b></u> | <u><b>72,143.16</b></u>  |
| <b>Net Surplus or (Deficit)</b> | <u><b>1,136.85</b></u>  | <u><b>(3,861.57)</b></u> |

**INDEPENDENT AUDIT REPORT**

To the members of Metrology Society of Australia

**Scope**

I have audited the attached annual financial report comprising the Statement of Cash Balances and the Statement of Cash Payments and Cash Receipts of Metrology Society of Australia for the year ended 30 June 2001. The Committee of Management is responsible for the financial report and has determined that the accounting policies used are consistent with the financial reporting requirements of the Metrology Society's constitution and are appropriate to meet the needs of the members. I have conducted an independent audit of the financial report in order to express an opinion on it to the members of Metrology Society of Australia. No opinion is expressed as to whether the accounting policies used are appropriate to the needs of the members.

The financial report has been prepared for distribution to members for the purpose of fulfilling the Committee of Management's financial reporting requirements under the Society's constitution. I disclaim any assumption of responsibility for any reliance on this report or on the financial report to which it relates to any person other than the members, or for any purpose other than that for which it was prepared.

My audit has been conducted in accordance with Australian Auditing Standards. My procedures included examination, on a test basis, of evidence supporting the amounts and other disclosures in the financial report. These procedures have been undertaken to form an opinion whether, in all material respects, the financial report is presented fairly in accordance with the cash basis of accounting whereby revenue is recorded when it is received, expenses are recorded when they are paid, and no assets or liabilities, other than cash and bank balances, are recorded. Accounting Standards and other professional reporting requirements (UIG Consensus Views) are not applicable to the cash basis of accounting adopted by Metrology Society of Australia.

The audit opinion expressed in this report has been formed on the above basis.

**Audit Opinion**

In my opinion the financial reports presents fairly in accordance with the cash basis of accounting, as described above, the payments and receipts of the Metrology Society of Australia for the year ended 30 June 2001 and its cash and bank balances as at that 30 June 2001.

Date 27 September 2001  
Address

BF Thornton CPA  
26 Penrith Avenue, WHEELER HEIGHTS NSW 2097

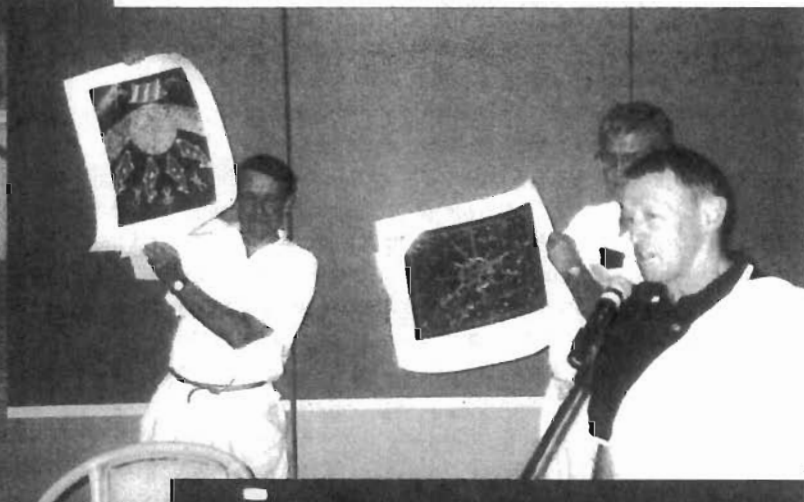
**MOTIONS for AGM**

1. Fees for 2001/ 2002 remain as they are:

|                                               |      |
|-----------------------------------------------|------|
| Associate member                              | \$35 |
| Full member                                   | \$40 |
| Fellow                                        | \$45 |
| Joining fee is equal to 1 year's subscription |      |

2. That Bryce Thornton is elected as the auditor for the financial year July 2001 to June 2002.

## MSA 2001



## IIZUKA PRIZE AWARDED TO ESA JAATINEN

NML's Dr Esa Jaatinen has been awarded the inaugural IIZUKA Prize, together with Cheong Tak Leong of PSB (Singapore Productivity & Standards Board).

The Iizuka Prize was set up by Dr Kozo Iizuka, former Director-General of the Agency of Industrial Science and Technology (AIST), Japan, member and Vice-President of CIPM for 15 years, and adviser to the APMP (Asia Pacific Metrology Programme) Secretariat. The Prize, awarded to two people every year, is worth US\$1000. It is awarded to individuals aged 35 years or under in recognition of their outstanding contribution to research activity related to metrology standards in the Asia Pacific region.

The presentation ceremony will take place on 7 November 2001 during the 17th APMP General Assembly in Tsukuba, Japan.

## MSA Golf Day 2001

The inaugural biennial MSA golf day was run on Friday the 5th of October 2001 (during the MSA 2001 4th Biennial Conference) at the Paradise Springs Golf Club on the Gold Coast.

The course is one of Queensland's most spectacular courses, with a Queensland colonial style club house, beautiful lakes and undulating manicured fairways.

From a golfers point of view, it was deceptively difficult with a lot of water and well placed fairway bunkers. This was reflected in the scores recorded. Etiquette was maintained throughout the day. Players were often heard warning other players of the water and sand bunkers prior to hitting. This didn't seem to help much, so the 'encouragement' continued.

The event was a single stableford. Handicaps used were normal handicaps for the club golfers and 27 awarded to social players. Although the option of 'dynamic' handicapping was reserved in case of 'social burglars', it wasn't needed.

### Results

Winner: Geoff Barnier (17) 31pts

Runner up: Shane Brann (19) 30pts

A rather large perpetual trophy was kindly donated by VMS International and presented by Shane Brann (VMS Managing Director) to the winner over a few cool ales in the club house at the end of the event.

Overall, the day was a lot of fun although expensive

with a lot of wet balls. It is hoped that further golf days will be run during each MSA conference with the trophy again being competed for. There is plenty of room for engraving the winners names!

## From the Dinner...

From the photos elsewhere in this issue, you will have gathered the obvious impression that it was not a formal affair!

Organiser Tony instructed each table to provide an impromptu musical item. The winners were the table led by budding songwriter/guitarist Ilya Budovsky, who produced the following:

*When I get older, loosing my hair  
Many years from now.  
Will they still be thinking of uncertainties,  
Error budgets, cal intervals?  
When you find out I'm with MSA,  
Would you lock the door?  
Will I have to measure at my own leisure  
When I'm 64?*

*Maybe I can simply buy a clever instrument from Fluke  
If it's not too dear.  
I will turn it on.  
Ooo and if I say a word  
All errors disappear.*

*No one's uncertain, everyone's right,  
What a point of view!  
Figures mean precisely what they mean to say  
And I'm not sure what I'm gonna do.  
No calibration, no accreditation,  
Who could ask for more!  
I'll have to measure at my own leisure  
When I'm 64.*

*I hope I don't measure at my own leisure  
When I'm 64.*

Music: John Lennon

Words: John Lennon/Ilya Budovsky

**Footnote:** There were some other excellent wittily conceived contributions by the other tables, but unfortunately there could only be one winner.

## Bayesian Statistics

Bob Frenkel

National Measurement Laboratory

Thomas Bayes was born in London in 1702, and became a Nonconformist minister at the Chapel of Tunbridge Wells in Kent. He was also a mathematician of some considerable repute, and was elected Fellow of the Royal Society of London in 1742. One of his writings was a paper entitled: 'Essay towards Solving a Problem in the Doctrine of Chances', which was sent for publication to the Royal Society only after his death in 1761. This is the paper that has spawned the branch of statistics known as Bayesian inference, or simply as Bayesian statistics.

Some of this is uncontroversially part of probability theory, but nevertheless leads to some tricky and non-intuitive results. Other equally interesting, but rather more controversial, aspects of Bayesian statistics have important implications for the calculation of results and their uncertainties in metrology. Although at the present time none of these have achieved formal recommendation, and the ISO Guide to the Expression of Uncertainty in Measurement (GUM) makes no explicit mention of Bayesian statistics, the substantial increase in papers on Bayesian topics in Metrologia and other journals in recent years attests to the growing interest in this subject.

The problem that Bayes tackled is the problem of inferring from newly observed data the probability of an event or hypothesis. In all cases there must be an initial so-called 'prior' probability that precedes the data. This prior probability may follow from data collected in the past, but also equally respectably from nothing more than an educated guess. (As will be discussed later, this is not as risky as may seem). So this problem is a very general one of scientific inference: how do we modify our beliefs in the light of new evidence? Moreover, conventional statistical analysis has little to say in this respect, because it pays no attention to prior probabilities and regards the new data as merely one sample out of very many possible samples - hence the common name 'sampling' statistics to distinguish it from Bayesian statistics. Another description of sampling statistics is 'frequentist', because in this conventional approach confidence intervals are obtained whose interpretation is the relative frequency with which 'true' values of parameters, such as population mean

and population variance, fall within such intervals during imaginary repeated sampling.

Bayesian statistics uses a different approach. The set of newly obtained data is taken as a fixed or 'given' quantity, and 'true' values of parameters do not exist - only their range of variation corresponding to a selected level of probability. The prior probability, combined with a feature of the newly observed set of data called the 'likelihood', provides the desired 'posterior' probability of the event or hypothesis. In fact the posterior probability is essentially the prior probability multiplied by the likelihood.

### Conditional Probabilities and Bayes' Theorem

As a preliminary to deriving Bayes' theorem (which is almost a mathematical one-liner!) we write  $P(A|B)$  to denote the 'conditional probability' of event A happening, *given* that event B has been observed to have happened. The vertical bar, pronounced 'given', is now in universal use in the Bayesian context. If  $P(A, B)$  is the probability of both events A and B happening (more precisely, the probability of any feature common to both A and B happening), we note (as part of this preliminary) that

$$P(A|B) = \frac{P(A, B)}{P(B)}. \quad (1)$$

Because  $P(B)$  (being a probability) is a fraction between 0 (impossible) and 1 (certain),  $P(A|B)$  is always larger than  $P(A, B)$ . This makes sense: if B is given, then some uncertainty has been dispelled and  $P(A|B)$  must surely be larger than  $P(A, B)$ . Moreover  $P(A, B)$  must always be less than or equal to  $P(B)$ . (This is a very general rule, and is fairly obvious if we remember that  $(A, B)$  is the set of events *common* to both A and B. If A and B are represented by enclosed areas of any shape - they can overlap, they can be separate or one can enclose the other - then  $(A, B)$  is the area of overlap, if any, and is always less than or equal to  $(B)$ , the area of B). So  $P(A|B)$  in (1) must be between 0 and 1, and so qualifies as a probability.

If events A and B are independent, then it is always true that  $P(A, B) = P(A)P(B)$ . Then (1) gives  $P(A|B) = P(A)$ , and this also makes sense: if A and B are independent, then whether B happens or not is immaterial to A.

As an illustration of (1), suppose that the proportion of full-time tertiary students in science is 14%. Suppose also that these science students are split 50-50 male-female. In the general student population, suppose that 56% are female. What is the proportion of female students in science? Equ. (1) gives:

$$P(\text{science}|\text{female}) = \frac{P(\text{science, female})}{P(\text{female})} = \frac{7/100}{56/100} = 1/8.$$

So although the proportion of students who are both female and doing science is

7%, the probability that any randomly selected female student is doing science is quite a lot higher at around 12%.

From (1) Bayes' Theorem can be readily obtained using the fact that, by symmetry,  $P(A, B) = P(B, A)$ . But (1) can equally well be written

$$P(B|A) = \frac{P(B, A)}{P(A)}, \quad (2)$$

so combining (1) and (2) gives:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \quad (3)$$

and (3) is essentially Bayes' Theorem. There are several ways of writing it: in many cases when probabilities of discrete events are being calculated,  $P(B)$  on the right side can be more explicitly written

$$P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A})$$

where  $\bar{A}$  is 'not-A'. This simply says that if B is made up of two mutually exclusive components, A and of course  $\bar{A}$ , then B is the sum of the overlap areas of itself with A and with  $\bar{A}$ . So Bayes' Theorem in a commonly used form is

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})}. \quad (4)$$

$P(A)$  in the numerator of (4) is the prior probability of A, before new data or new information have been made available.  $P(A|B)$  on the left of (4) is then the posterior probability of A, after the new data or new information have arrived. The term  $P(B|A)$  in the numerator on the right of A is proportional to the likelihood, as will be discussed in more detail later.

More generally, if  $A_1, A_2, A_3, \dots$  are mutually exclusive components of B,  $P(B)$  on the right side of (3) is equal to  $P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3) \dots$ , and this becomes the expression in the denominator of (4).

### Bayes' Theorem with Discrete Probabilities

Here are some examples of Bayes' Theorem in operation. They are examples concerning 'discrete' probabilities rather than the probability distributions that we are accustomed to handling in metrology. However, it is worth spending some time on the discrete case, since some of the essential issues involved in the Bayesian approach can be clearly seen.

A rather rare and serious disease is observed to attack 1 in 2000 people in the

general population. Your doctor tests you for this disease, using a test that she says is very reliable: in fact, it has a false-positive rate of only 1%, meaning that 1% of people tested appear to have the disease (they test positive) when in fact they are healthy. Your test indicates that you have the disease. Should you be alarmed? Most people would be, because they have this conditional probability in mind:

$$P(\text{positive test}|\text{disease is present}) = 0.99.$$

This is a correctly stated conditional probability of 99% - hence the understandable alarm. But the conditional probability that you ought to have in mind is the other way around: it is

$$P(\text{disease is present}|\text{positive test}).$$

One of the virtues of Bayes' Theorem is that it brings into sharp focus the conceptual (and numerical) difference between two such conditional probabilities. Using (4) we have for the second and more relevant conditional probability (and abbreviating):

$$\begin{aligned} P(\text{disease}|\text{pos}) &= \frac{P(\text{pos}|\text{disease})P(\text{disease})}{P(\text{pos}|\text{disease})P(\text{disease}) + P(\text{pos}|\text{no disease})P(\text{no disease})} \quad (5) \\ &= \frac{0.99 \times (1/2000)}{0.99 \times (1/2000) + 0.01 \times (1999/2000)} \\ &\sim 0.047. \end{aligned}$$

The relevant conditional probability is only about 5%! So you should respond to your positive test result with (cautious) optimism. One can see, in fact, that Bayes' Theorem emphasises the need for a *super*-reliable test when the disease is a rare one. The reliability of the test should be at least commensurate with the rareness of the disease; in this example, the apparently high figure of 99% reliability actually falls short of 100% by a margin twenty times wider than the frequency of the disease.

There is a quick and rough way to check the above result. Out of 2000 people there will be 1 person with the disease (on average), and 1999 without the disease. Of these 1999, about 20 people (1%) will test positive. So if a particular person's test has yielded a positive result, the probability that he or she is that one who actually has the disease is about 1 in 20.

A closely related calculation in the same example provides a reason for reducing the level of anxiety in the medical profession as well. Such anxiety may arise from the possibility that a diagnostic test has erroneously - and therefore extremely regrettably - returned a *negative* result. As reported recently in the Sydney Morning Herald (5 February 2001) regarding one category of test, medical practitioners avoid

this perceived risk by recommending frequent and expensive re-testing for their patients. Let us calculate the probability that a negative test result is erroneous and that the patient has the disease. Suppose that the false-negative rate is 1% (it need not be the same as the false-positive rate; the argument is not affected). Then

$$P(\text{negative test}|\text{disease is present}) = 0.01.$$

But this is not the conditional probability on which the medical profession should base its re-testing strategy. The correct conditional probability is:

$$P(\text{disease is present}|\text{negative test}),$$

and this is given by (again abbreviating):

$$\begin{aligned} P(\text{disease}|\text{neg}) &= \frac{P(\text{neg}|\text{disease})P(\text{disease})}{P(\text{neg}|\text{disease})P(\text{disease}) + P(\text{neg}|\text{no disease})P(\text{no disease})} \quad (6) \\ &= \frac{0.01 \times (1/2000)}{0.01 \times (1/2000) + 0.99 \times (1999/2000)} \\ &\sim 5 \times 10^{-6}. \end{aligned}$$

So only one in two hundred thousand patients who test negative will actually have the disease! Again, this result can be quickly checked. Out of two hundred thousand people, one hundred will have the disease (because its rate is 1 in 2000) and of these, one person will test negative. The number of people without the disease is two hundred thousand less one hundred, or still two hundred thousand as near as makes no difference. So if a particular person's test has yielded a negative result, the probability that he or she actually has the disease is 1 in two hundred thousand.

What is to be emphasised is not this ultralow probability as such, but the contrast between it and the illusory probability that implies that because the false-negative rate for the test is 1%, therefore one in a hundred patients who test negative will have the disease.

There are other areas than the medical where we need to be careful with probability arguments. When a jury is required to decide whether a defendant is guilty of a crime 'beyond reasonable doubt', incorrect probabilities may be presented to the jury. Consider the scenario - of increasing relevance nowadays - where a suspect has been charged with the offence on the evidence of a DNA match. The forensic value of DNA testing can be summarised by a conditional probability such as the following:

$$P(\text{match}|\text{innocent}) = 10^{-6}.$$

If I am innocent of the crime, then the probability that my DNA sample will match the sample taken from the crime scene is one in a million (or some such ultralow probability). Suppose that in a city of two million people a suspect has been charged on the evidence of a DNA match. At the trial, the counsel for the prosecution is likely to remind the jury emphatically of the above tiny conditional probability. But this would be committing the so-called 'prosecutor's fallacy' [1]. The conditional probability that the jury should consider is

$$P(\text{innocent}|\text{match})$$

and this is (abbreviating with gty for guilty, and remembering that  $P(\text{match}|\text{gty}) \sim 1$ ):

$$P(\text{inn}|\text{match}) = \frac{P(\text{match}|\text{inn})P(\text{inn})}{P(\text{match}|\text{inn})P(\text{inn}) + P(\text{match}|\text{gty})P(\text{gty})} \quad (7)$$

$$= \frac{10^{-6} \times (1999999/2000000)}{10^{-6} \times (1999999/2000000) + 1 \times (1/2000000)}$$

$$\sim 2/3.$$

The odds of the suspect being innocent are in fact two to one. (Given that there is a match, two possibilities follow: the suspect is in fact the criminal, or the suspect is one of the expected two innocent people who will match by chance). In a city with a larger population the odds in favour of innocence would be increased still further, whereas in a smaller community they would be reduced. The main point is that the probability of innocence, when a match is found, is nothing like the ultralow figure regularly quoted in this context, and DNA testing should always be supplemented by further evidence. The prosecutor's fallacy also has a corresponding defender's fallacy that acts the opposite way: the defending counsel offers the jury a conditional probability which, again, is 'back-to-front' and which incorrectly implies a very low probability of guilt. A case in point occurred in the US some years ago during the trial of a high-profile public figure [2].

### Bayes' Theorem with Probability Distributions

Equ. (4) is readily adapted to the very common case in metrology, or indeed in any branch of science, where the gathering of data is followed by an estimation of one or several parameters of interest. The probabilities involved are generally calculated from continuous distributions such as the normal. In (4), we identify A with some parameter of interest, and this is commonly given the symbol  $\theta$ . B is then the data, often denoted by  $y$ . The left-hand side of (4) is then  $P(\theta|y)$ , expressing the probability of the parameter taking a particular value, given an available set of data  $y$ . This probability is called the 'posterior' probability of  $\theta$ , implying that it follows from, and is based on, the available data. In the numerator of the right-hand side

of (4) the term  $P(A)$  is now  $P(\theta)$ , and this is the 'prior' probability of  $\theta$ , based on our state of knowledge or ignorance of  $\theta$  prior to the gathering of the data.

In (4), the term  $P(B|A)$  becomes  $P(y|\theta)$  and is known as the 'likelihood' of  $\theta$  and  $y$ . Computing the likelihood is the 'guts' of Bayesian analysis and will be illustrated below. So (4) now reads:

$$P(\theta|y) \propto P(y|\theta)P(\theta), \quad (8)$$

or: 'posterior is proportional to prior multiplied by likelihood'. With continuous distributions the denominator of (4) takes on the role of a normalising factor. Bayesian analysis can proceed quite comfortably with the proportionality symbol in (8), because at its conclusion the normalising factor is then determined simply by requiring the integral of  $P(\theta|y)$  over all  $\theta$  to be 1 (it is certain that any parameter must take some value over its entire permitted range!).

The presence of the prior probability  $P(\theta)$  in (8) is one of the crucial distinctions between conventional sampling statistics and the Bayesian approach when analysing scientific data. It is interesting that in discrete-probability cases such as those discussed earlier, prior probabilities are obviously needed - for example, the prior probability of a disease before any tests for the disease are made. Without prior probabilities the correct answers cannot be found. In conventional sampling statistics that we use in, for example, calculating the mean and standard deviation of a population, no prior probabilities are considered - but the Bayesian enthusiast will claim that they are still present but not explicitly stated. This can easily be shown to be so, and moreover there is a gain in philosophical consistency by requiring prior probabilities to be considered in all cases.

Prior probabilities are generally assigned in one of two ways. One way is by appeal to results obtained in the past, so that by a kind of 'chain rule' the posterior probability obtained in the past now becomes the prior probability for the present measurement. This choice of prior has interesting and potentially radical metrological implications, as will be discussed below. The other way (when, for example, the past is unknown or cannot provide sufficiently reliable information for the present) is to use so-called 'non-informative' or 'uniform' or 'flat' priors - all different names for the same thing. These state in mathematical terms that we know the parameters of interest only roughly, or not at all, *until* we have obtained new data. These priors are 'flat' because they have an almost uniform distribution over the feasible range of each parameter of interest. They may represent, therefore, no more than an educated guess, but it is interesting that in many cases flat priors yield identical results to conventional sampling statistics.

Historically, using flat priors was the most controversial aspect of the Bayesian approach ('if you have no idea what's going to happen, then all the possibilities are equally likely'). Bayes himself was not happy with this aspect, and this is one reason

why his paper was published only after his death.

When the parameter may, '*a priori*' - before the data - take any value from plus to minus infinity, the common choice for the prior is simply a constant, so that  $P(\theta) = C$ . The mean of a distribution is an obvious parameter that *a priori* has this range. The value of the constant  $C$  does not matter, since as mentioned above the final step in the Bayesian analysis is the appropriate choice of normalising factor, into which that constant can be absorbed. A probability like  $P(\theta) = C$  (more precisely, a probability *density*) is certainly flat, and it may be correctly objected that its integral over the entire range of  $\theta$  is then infinite. In practice this, again, does not matter, since it is only necessary for the prior to be approximately flat near its 'most likely' range of values. This range of values is, unsurprisingly, that for which the likelihood as determined by the newly obtained data is large. There is therefore a contrast between the 'peakiness' of the likelihood and the relative flatness of the prior in the region of the most likely range of values for the parameter.

Other parameters of interest may be restricted to zero or positive values, with a range 0 to plus infinity. Among these are, of course, the standard deviation and variance of the population. For these parameters the commonly used flat prior is not  $P(\theta) = C$  but rather  $P(\theta) \propto 1/\theta$ . This dependence on  $1/\theta$  can be argued as follows, using the fact that the integral of the function  $1/\theta$  with respect to  $\theta$  is  $\log \theta$ . If the permitted range of  $\theta$  is 0 to plus infinity, then  $\log \theta$  can range from minus infinity to plus infinity, and, by analogy with the previous example of the mean, we have not  $P(\theta) = C$  but  $P(\log \theta) = C$ . So  $P(\log \theta)$  is then our flat prior. Consider its range from (say)  $\log \theta = 3$  to  $\log \theta = 7$ . Since the prior is flat, the total probability within this rectangular 'box' is obviously  $(7 - 3)C = 4C$ . But the value of  $\theta$  corresponding to  $\log \theta = 3$  is of course  $\theta = \exp 3$ , and similarly the other bound for  $\theta$  is  $\theta = \exp 7$ . Which functional relation  $f(\theta)$  of  $\theta$ , integrated between  $\theta = \exp 3$  and  $\theta = \exp 7$ , gives the same total probability  $4C$ ? The answer is  $f(\theta) = C/\theta$ , since its integral over that range is  $C [\log \exp 7 - \log \exp 3] = (7 - 3)C = 4C$ .

The prior  $P(\theta) \propto 1/\theta$  is infinite at  $\theta = 0$ , but this has not caused problems, either because integrating over  $\theta$  excludes the point  $\theta = 0$  in a particular application, or because (as in the case of normal variables) the prior is multiplied by a likelihood and the product is zero or finite at  $\theta = 0$ .

Here is a quick look to see how some of this works in practice in a simple case. Suppose we measure  $n$  quantities  $y_1, y_2, \dots, y_n$  from a normal population with mean  $\mu$  and variance  $\sigma^2$ . From the standard expression for the normal probability function, the probability of getting each  $y_i$  is proportional to

$$(1/\sigma) \exp \left[ (-1/2)(y_i - \mu)^2/\sigma^2 \right].$$

The likelihood of getting all the  $n$   $y$ 's, influenced by the mean  $\mu$  and variance  $\sigma^2$ , is the product of  $n$  terms like the above, and since the product of exponentials is the

exponential of the sum, the likelihood is proportional to

$$(1/\sigma)^n \exp \left[ (-1/2) \sum_{i=1}^n (y_i - \mu)^2 / \sigma^2 \right].$$

Suppose the variance is known but the mean is not. The first term  $1/\sigma^n$  ceases to be of interest and can be dropped. So the likelihood is now proportional to

$$\exp \left[ -(1/2) \sum_{i=1}^n (y_i - \mu)^2 / \sigma^2 \right].$$

If  $\bar{y}$  is the mean  $(1/n) \sum_{i=1}^n y_i$  of the sample, and  $s^2 = (1/n) \sum_{i=1}^n (y_i - \bar{y})^2$  is the variance of the sample (both of these being *known* and *fixed* quantities by virtue of the sample), then by writing  $y_i - \mu = (y_i - \bar{y}) - (\mu - \bar{y})$  one can show that

$$\sum_{i=1}^n (y_i - \mu)^2 = ns^2 + n(\bar{y} - \mu)^2.$$

So the likelihood is now proportional to

$$\begin{aligned} & \exp \left[ -(n/2\sigma^2) \{s^2 + (\bar{y} - \mu)^2\} \right] \\ &= \exp(-ns^2/2\sigma^2) \exp \left[ -n(\bar{y} - \mu)^2/2\sigma^2 \right], \end{aligned}$$

and now the first term can be dropped, since  $s^2$  and  $\sigma^2$  are both known. The likelihood is now proportional to

$$\exp \left[ -n(\bar{y} - \mu)^2/2\sigma^2 \right].$$

The flat prior for  $\mu$  is the constant  $C$ . So the posterior probability for  $\mu$  is

$$C \exp \left[ -n(\bar{y} - \mu)^2/2\sigma^2 \right].$$

$\mu$  is therefore distributed normally with mean  $\bar{y}$  and variance  $\sigma^2/n$ , or standard deviation  $\sigma/\sqrt{n}$ .

We notice that in this Bayesian approach it is the population mean  $\mu$ , *not* the sample mean  $\bar{y}$ , whose distribution is inferred from the single fixed sample that provides  $\bar{y}$  and  $s$ .  $\mu$  is now effectively a random variable. In sampling statistics things are the other way round: the single sample that provides  $\bar{y}$  and  $s$  is assumed to be reliably representative of many imaginary repeated samples, and all of these together serve to locate the *fixed* constant  $\mu$  at  $\mu = \bar{y}$  within a confidence interval determined by  $s/\sqrt{n}$ .

### Some differences between Bayesian and sampling statistics

The result obtained above for the posterior distribution of  $\mu$  and its standard deviation is of course very like the more familiar result from sampling statistics. But if Bayesian and sampling statistics always gave the same results, there would be little practical point in adopting the Bayesian viewpoint of 'prior probabilities, fixed data, variable parameters' that contrasts with the conventional practice of 'variable data, fixed parameters' (and no talk of prior probabilities). However, results from sampling statistics can be shown to be approximations to Bayesian results, these approximations becoming more accurate with increasing sample sizes. With small samples the two methods may diverge markedly. A case in point is the  $t$ -distribution, which in the Bayesian formulation yields a result for the standard deviation of the posterior distribution of the population mean that forbids sample sizes of 3 or less. In the conventional formulation, such small sample sizes are permitted (although discouraged). Below are several further metrological issues where the Bayesian approach is decidedly different from the conventional, or where some benefits of the Bayesian approach are evident:

1. For metrologists who regularly calibrate artefact standards such as standard masses and standard resistors, the values to be reported for these artefacts, and the uncertainties in these values, will be affected by the need in Bayesian statistics to consider prior probabilities. It is here that the distinction between Bayesian and sampling statistics is undoubtedly sharpest, and will provoke lively debate among metrologists!

Suppose that an artefact was calibrated for the first time in January 2000 and was reported to have the value  $x_1$  with standard uncertainty  $u_1$ . It was next calibrated in January 2001, when the calibration process assigned it a value  $x_2$  with standard uncertainty  $u_2$ . With sampling statistics, the values to be reported for this second calibration would be simply  $x_2$  and  $u_2$ . But in the Bayesian approach, the values that should be reported for the January 2001 calibration are the posterior values for the artefact, given the prior values from the calibration in January 2000 and the likelihood (in the technical sense discussed above) of the new data obtained in January 2001. If it can be assumed that the artefact had no significant drift over the twelve-month interval, then one can easily show (assuming normally distributed prior values and likelihood) that the posterior values for January 2001 are given by

$$x_2(\text{post}) = \left[ (x_1/u_1^2) + (x_2/u_2^2) \right] / \left[ (1/u_1^2) + (1/u_2^2) \right],$$

and

$$u_2(\text{post}) = u_1 u_2 / \sqrt{u_1^2 + u_2^2}.$$

$x_2(\text{post})$  is therefore simply the weighted mean of  $x_1$  and  $x_2$ , the weights being the inverse squares of the two standard uncertainties. (Taking weighted means, and

using the inverse squares of standard uncertainties as the weights, is itself a familiar enough practice). So if the second calibration is more accurate than the first,  $x_2(\text{post})$  will be closer to  $x_2$  than to  $x_1$ . It may also be shown that  $u_2(\text{post})$  is less than both  $u_1$  and  $u_2$ , and that a correlation exists between  $x_1$  and  $x_2(\text{post})$ , the value of the correlation coefficient being  $u_2/\sqrt{u_1^2 + u_2^2}$ .

If, for example, the artefact is a 1 ohm standard resistor reported to have the correction  $x_1 = +10$  microhms with  $u_1 = 1$  microhm in January 2000 (its first calibration) and then measured as having a correction  $x_2 = +8.0$  microhms with  $u_2 = 0.3$  microhm in January 2001, then the reported values for January 2001 will be  $x_2(\text{post}) = 8.2$  microhms,  $u_2(\text{post}) = 0.29$  microhm. The report on the second calibration would state the correction as 8.2 microhms, even though it was measured in January 2001 as having the correction 8.0 microhms. The standard uncertainty of the second calibration would be likely to be quoted as 0.3 microhm, this being the sensible rounding of 0.29 microhm.

Of course, standard resistors cannot be generally assumed to have zero drift. So this example of a Bayesian analysis applied to just two calibrations is unrealistic. However, it is possible to attempt a Bayesian analysis of a drifting artefact. Thus the existence and value of the drift could be at least roughly assessed after, say, five calibrations. During this period the uncertainty in the performance of the artefact would effectively create flat priors for each successive calibration, so that no 'memory' of the preceding calibrations was present for each new calibration. After sufficient measurements were taken to enable the drift to be estimated, the uncertainty in the value of resistance extrapolated along the drift line from previous measurements to the present (this uncertainty of course increasing with the extrapolation interval), combined with the overall root-mean-square (rms) scatter in the results, would constitute the prior uncertainty for the present measurement.

For a shorter period of repeated measurements, the originally estimated standard uncertainty for each measurement could be added in quadrature to an estimate of the (known) overall rms scatter to provide a prior value and its standard uncertainty for each successive measurement. With the priors for each measurement taken as originating only from the previous measurement (so that the 'memory' extends only as far back as the previous measurement, but no further), the illustration shows the effect on the reported values for a voltage standard measured over 16 days, where the known overall rms scatter is taken as 0.049 microvolts. There is a small improvement in stability and reduction in uncertainty for each measurement except for the first (as indicated by the uncertainty bars). (A complication of voltage standards has been neglected here, in that successive values of measurements taken days apart may be correlated because of low-frequency  $1/f$  noise).

The obvious objection to all this is: I have just measured the artefact as having

a particular value with an associated uncertainty. What on earth does its previously measured value have to do with the value I have obtained now? The short answer is - the previous value would indeed be irrelevant to the present value, *if* my present uncertainty was a lot smaller than the uncertainty for the previous value. Then the weighting described above would greatly favour my present value. But if the two uncertainties were roughly equal, or my present uncertainty was significantly degraded for some reason, then there would be a 'memory' of the previous value which would need to be taken into account, according to the Bayesian approach. Taking a slightly different (and perhaps more illuminating) point of view, the previous and present calibrations could be regarded as two parts of a single greatly prolonged calibration, the results of the two parts being subsequently combined and weighted as indicated to yield a single reported value.

But now suppose the present and previous values have roughly equal uncertainties, and the previous value was last measured twenty years ago - would there still be a memory of it? The answer is - very likely not, because the change in the artefact (due to linear and non-linear drift, scatter and so forth) would not be known, and so the prior probability needed for the present measurement would be flattened almost completely by the uncertainty induced by twenty years of non-measurement. To these very reasonable objections, therefore, there appear to be reasonable Bayesian answers! In fact, any passage of time, no matter how small, between successive calibrations will tend to flatten the prior.

2. As remarked by several authors (for example [3], [4]) the ISO GUM in its definitions of variables and several measures of uncertainty does not consistently adhere to either sampling or Bayesian statistics. Thus section G3.2 refers to the *t*-distribution from the sampling point of view, where the population mean is taken as fixed. However, the definition of expanded uncertainty (2.3.5 and 6.2), which quotes the 'values that could reasonably be attributed to the measurand', is a Bayesian definition.

The GUM has recommended a unified treatment of Type A and Type B uncertainties, in the sense that the desired combined standard uncertainty is the root-sum-square of these two components, and can be reported as a single quantity. But the definitions of Type A and Type B uncertainties are: Type A, evaluated by statistical analysis (2.3.2), and Type B, by other means (2.3.3). Such a distinction - not always helpful, as even seasoned GUM-users will admit - is not needed in the Bayesian approach, where the theoretical handling of the two components is identical when the prior probabilities are being formulated ([4],[5]).

3. The normal distribution is in popular use in sampling statistics partly because of the tendency towards normality of the net result of many independent random effects. But another reason for its use is its simplicity, and the convenient feature that sample values such as the mean and variance can represent and summarise individual data for the purpose of estimating the corresponding population values

(the sample mean and variance are 'sufficient', to use the technical term). However, some measurements may well need to be described by other more complicated distributions than the normal, and such distributions may not possess these 'sufficient' attributes. Unlike the case in sampling statistics, the Bayesian approach allows such distributions to be handled in a straightforward way.

4. The Welch-Satterthwaite formula, obtained from sampling statistics, has been recommended by the GUM for the purpose of estimating the effective number of degrees of freedom attaching to the standard uncertainty evaluated from several independent component standard uncertainties. Each of these components has its own number of degrees of freedom, but a defect of the Welch-Satterthwaite formula ([6],[7]) is that when, among these components, there are one or several dominant components with very few degrees of freedom, then the coverage factor obtained from the effective number may be unreasonably small, leading to an unacceptably small expanded uncertainty. The same class of problem approached via Bayesian statistics results in the so-called Behrens-Fisher distribution which is free of this defect, and of which more will probably be heard in the future.

In any case, the concept of 'degrees of freedom' itself has little relevance in the Bayesian scheme of things, since parameters are random variables and are determined from their posterior distributions. Bayesian 'uncertainty budgets' would not need to include degrees-of-freedom entries (although more computation might be involved). To end on a speculative note, the metrological community may offer a hearty welcome to this freedom from degrees of freedom.

### Further reading

General articles on Bayesian statistics:

'Faith, Hope and Statistics', Robert Matthews, *New Scientist*, 22 November 1997.

'Bayes Offers a 'New' Way to Make Sense of Numbers', David Malakoff, *Science*, 19 November 1999.

### References

1. 'How Convincing is DNA evidence?', David Balding and Peter Donnelly, *Nature*, vol. 368, 24 March 1994.
2. 'When Batterer Turns Murderer', I.J. Good, *Nature*, vol. 375, 15 June 1995.
3. 'An Interpretation of the Guide to the Expression of Uncertainty in Measurement', Raghu Kacker, Monograph 25 May 2000, National Institute of Standards and Technology, Gaithersburg, Maryland US.

4. 'A Bayesian Theory of Measurement Uncertainty', K. Weise and W. Woger, *Measurement Science and Technology*, vol. 3, 1992, pp. 1-11.
5. 'Bayesian Inference from Measurement Information', I. Lira and G. Kyriazis, *Metrologia*, vol. 36, 1999, pp. 163-169.
6. Report on Workshop 11-12 November 1999, 'Statistical Analysis of Interlaboratory Comparisons', Maurice Cox, National Physical Laboratory, UK, December 1999.
7. 'Limitations of the Welch-Satterthwaite Approximation for Measurement Uncertainty Calculations', M. Ballico, *Metrologia*, vol. 37, No. 1, 2000, pp 61-64.

Other articles and books:

- 'Bayesian Approach to Recalibration', V. Tuninsky and W. Woger, *Metrologia*, vol. 34, 1997, pp. 459-465.
- 'Calculation of Measurement Uncertainty Using Prior Information', S.D.Phillips, W.T.Estler, M.S.Levenson and K.R.Eberhardt, *Journal of Research of the National Institute of Standards and Technology*, vol. 103, No. 6, Nov-Dec 1998.
- 'The Value of Additional Knowledge in Measurement', J. Beyerer, *Measurement*, vol. 25, 1999, pp. 1-7.
- 'Bayesian Statistics - an Introduction', P. Lee, London, Edward Arnold, 1989.
- 'Bayesian Inference in Statistical Analysis', G. Box and G. Tiao, Addison-Wesley Publishing Company, 1973.

**The Australian Metrologist** is published four times per year by the Metrology Society of Australia Inc., an Association representing the interests of metrologists of all disciplines throughout Australia. Membership is available to all appropriately qualified and experienced individuals. Associate membership is also available.

#### Membership Enquiries

Contact either your State Coordinators or the Secretary, Dr. Laurie Besley on (02) 9413 7770 or fax (02) 9413 7202, e-mail address [laurieb@tip.csiro.au](mailto:laurieb@tip.csiro.au) or write to:

The Secretary, Metrology Society of Australia  
c/o CSIRO National Measurement Laboratory  
PO Box 218  
LINDFIELD NSW 2070

The **MSA website** address is [www.metrology.asn.au](http://www.metrology.asn.au)

**Webmaster:** Mark Thomas (03) 9244 4042 (wk)

#### Membership Fees

|            |                          |
|------------|--------------------------|
| Fellows    | \$45 Joining Fee         |
|            | \$45 Annual Subscription |
| Members    | \$40 Joining Fee         |
|            | \$40 Annual Subscription |
| Associates | \$35 Joining Fee         |
|            | \$35 Annual Subscription |

#### Contributions

Articles, news, papers and letters, either via e-mail, disk or hard copy, should be sent to:

The Editor  
The Australian Metrologist  
11 Richland Road  
NEWTON SA 5074  
Phone: (08) 8365 2451  
Fax: by arrangement only  
E-mail: [maurieh@ozemail.com.au](mailto:maurieh@ozemail.com.au)

**The deadline for the next issue is 16th January 2002.**

#### Sponsorship/Advertising

Would you or your company be interested in sponsoring a future issue of The Australian Metrologist? If you are a Member or your company is in the metrology business, a contribution of \$400 permits the sponsor to include a relevant insert (up to A4 in size) in the mail-out. If you wish to place an advertisement in TAM, contact the Editor for current pricing.

#### Positions Wanted/Vacant

Need a Position?

Write or e-mail the Editor with your details including years of experience and qualifications. This service is offered free of charge.

#### Need a Metrologist?

If you have a position vacant, write or e-mail the Editor with the details. A charge of \$20 for up to 10 lines applies. (The circulation may be small but it is well targeted.)

The deadline for positions wanted/vacant is as above.

#### Letters to the Editor

Letters should normally be limited to about 200 words. Writers will be contacted if significant editorial changes are considered necessary.

#### Editorial Policy

The Editor welcomes all material relevant to the practice of Metrology. Non-original material submitted must identify the source and contact details of the author and publisher. The editor reserves the right to refuse material that may compromise the Metrology Society of Australia. Contributors may be contacted regarding verification of material.

Opinions expressed in *The Australian Metrologist* do not necessarily represent those of the Metrology Society of Australia. Material in this journal is ©Metrology Society of Australia Inc. but may be reproduced with prior approval of the Editor.

**Editor:** Maurie Hooper

#### Management Committee

**President** Dr Jim Gardner (02) 9413 7323  
CSIRO (NML)

**Vice-president** Dr Ilya Budovsky (02) 9413 7201  
CSIRO (NML)

**Secretary** Dr Laurie Besley (02) 9413 7770  
CSIRO (NML)

**Treasurer** Ms Marian Haire (02) 9888-3922  
Nat. Standards Commission

**Members** Mr Barry Deeth (02) 9562 2778  
ADI NSW

Mr Frederick Emms (02) 9742 8724

Telstra

Mr Tony Jackson (02) 4724 4984

Workcover NSW

Mr Patrick McLain (02) 9869 3310

Agility Measurement Services

Mr Jim Miles (02) 9760 6575

TAFE Commission

Mr Brian Phillips (07) 3372 0430

Weigh-Tech Qld Pty Ltd

Ms Mary Ryan (02) 9736 8217

NATA

Mr Jeffrey Tapping (08) 8363 3602

#### National Liaison Officer

Jim Miles (02) 9760 6575

#### Marketing

Horst Sieker (03) 9295 8700

#### State Contacts

**NSW** Mr Brian Pritchard (02) 9413 7732 (wk)  
CSIRO National Measurement Laboratory  
PO Box 218  
Lindfield NSW 2070  
Fax (wk) (02) 9413 7202  
e-mail [Brian.Pritchard@tip.csiro.au](mailto:Brian.Pritchard@tip.csiro.au)

**NT** Bill Deusien (089) 413 382  
12 Dwyer Court  
Driver NT 0830  
Fax: (089) 411 951

**Qld** Mr Brian Phillips (07) 3372 0430(wk)  
Weigh-Tech Qld Pty Ltd  
e-mail [bztphil@technet2000.com.au](mailto:bztphil@technet2000.com.au)

**SA** Mr Jeffrey Tapping (08) 8363 3602  
102A Phillis St  
Maylands SA 5069  
Fax (08) 8362 1240  
e-mail [tapping@ozemail.com.au](mailto:tapping@ozemail.com.au)

**Tas** Mr Phil Wilde (03) 6324 4613 (wk)  
ACL Bearing Company  
PO Box 1088  
Launceston Tas 7250  
Fax: (03) 6326 6600  
e-mail [phil\\_wilde@acl.com.au](mailto:phil_wilde@acl.com.au)

**Vic** Mr Mark Thomas (03) 9244 4042 (wk)  
10 Wilton Close  
Wyndhamvale Vic 3024  
Fax (wk) (03) 9244 4004  
e-mail [mthomas@netspace.net.au](mailto:mthomas@netspace.net.au)

**WA** No rep.