MATHEMATICAL AND STATISTICAL TOOLS IN METROLOGICAL MEASUREMENTS

Franco Pavese
IMEKO Technical Committee 21 "Mathematical Tools for Measurements"
c/o Istituto Nazionale di Ricerca Metrologica (INRIM),10139 Torino, Italy

Keywords: Metrology, calibration, testing, statistics, data modeling, measurement, uncertainty, error.

Contents

Summary

Measurement science is a very broad field, with already partial contributions to the EOLSS. In this Chapter the contents are restricted to the branch of measurement science called 'metrology', comprising both 'calibration' and 'testing', and concerning only quantitative measurements. However, even after these delimitations, the field remains extremely vast. In general, the aim of this Chapter is to supply the reader an introduction to the essential concepts and to direct them to a selected and ample bibliography. In addition, the reader is directed to the tables of contents of nine books, summing up to more than 360 papers and 3000 pages, dedicated over the last 20 years to the subject matter of advanced mathematical, statistical and computational tools in metrology can be found at http://joomla.imeko.org/index.php/tc21-homepage (AMCTM). In this chapter the reader is assumed to be acquainted with the basic terms used in measurement science.

1. Introduction

A measurement process in physical, chemical and biological science consists initially of experimental determinations of a value of the quantity intended to be measured, the measurand, generally requiring to employ some hardware (measuring instrumentation system).

In the preceding phase of planning the experiment, it is often required to build up a model of the system subject to the measurement for several purposes: to ensure having a sufficiently complete picture of the measurand and understanding of its features; to determine the needs of the instrumentation; to possibly perform a simulation of the experiment, in order to detect possible problems and to optimize the instrumentation and its performance.

Subsequent to the phase of data acquisition from the instruments while performing the experiment, the acquired data will need analysis and the use of a data model in order to obtain the measurement results and an evaluation of its uncertainty, exploiting the full potential of the measurements.

The issues concerning instruments and hardware in general and the instrumental procedures are out of the scope of this Chapter. Also modeling of the systems, a very broad matter, will not be tackled in the following, except by providing some references. The contents will focus instead on data treatment, including data modeling, an issue requiring more and more sophisticated mathematical and statistical tools. Also sophisticated computer-aided tools are necessary today, which will also remain out of the scope of this chapter.

In this chapter double quotes are for literal citations, reported in italics, single quotes for highlighting a term used in this text. See the meaning of the terms and acronyms at the end of the Chapter.

2. Metrology in Measurement Science

Metrology is a branch of measurement science specifically devoted to achieve two main goals:

- At the top level of accuracy, a degree of ‘metrological traceability’ sufficient to obtain nationally and internationally ‘metrological compatibility’ at the needed level;

- At higher levels of uncertainty (lower levels in the metrological hierarchy), metrological traceability of the downward chain (the metrological chain, or pyramid) through an ‘unbroken chain of calibrations’.

In order to ensure the achievement and maintenance in time of these goals, the measurement standards are the competence of metrology, so are the measurements that disseminate the measurement units by means of operations called ‘calibration’ of standards of hierarchically low levels, including a statement about its uncertainty.

For the use of the International System of Units, SI, today adopted by the vast majority of countries all around the world, the BIPM constitutes the top of the metrological chain, followed by the National Metrology Institutes (NMI), and then the secondary laboratories accredited by the NMIs. The latter may be, in turn, organized in ranks at different hierarchical levels, down to the floor level of the workshops and individual users.

At the top level, metrology is a sophisticated and demanding science, as required by the most accurate determination of the values of the fundamental constants of the physical laws. Thus, the state-of-the-art knowledge is adopted or developed at the most sophisticated levels and beyond.

At lower levels, the metrological system relies on more and more simplified procedures and limited need for detailed knowledge, like it happens in most of the ‘testing’ field. In the latter context, the needs are usually covered by written standards, which are promulgated by ISO or similar international and national organizations; for the subject matter of this chapter, standards are mostly those provided by the ISO Technical Committee 69 (TC69).

The specificity of metrology with respect to the general field of measurement science can have a critical effect on the meaning of some basic concepts or on the methods preferred for data treatment.

In particular, some peculiarities of the measurement results more critically affect the decisions to be taken based on them. In this respect, after a general treatment of other aspects in metrology, requiring a careful selection of suitable mathematical and statistical tools, this Chapter will tackle in details the issue of the systematic effects that affect *all* experimental determinations, and are responsible for the possible disagreement of the results that, when occurring in different laboratories, affects the degree of compatibility and, consequently, the integrity of the metrological chain.

3. Modeling and Data Treatment

When top accuracy is aimed at, the experimenter needs to get acquainted with all the known aspects of the physical (chemical, biological) system that will be subjected to measurement.

Based on this knowledge, the experimenter lists the quantities that need to be measured, and those that must be quantitatively evaluated, in both cases because they have an influence on the (numerical) measurement results. They are called in fact ‘influence quantities’.

The required type of “measurement process” (clause 2.1 in (VIM3)) is different depending on the fact that one needs to measure an object (e.g. a mass standard), or an immaterial property (e.g. time), or a functional relationship (e.g. a physical ‘law’) between two or more quantities. Similarly is different when measuring an “extensive” or an “intensive” property.

In addition, measurements are usually grouped into two broad categories:

(a) “direct”, when the kind of quantity (VIM3) that is measured and that of the measurand intended to be measured coincide (e.g. mass for a mass standard):
(b) “indirect”, when the measured quantity is different from the aimed measurand (e.g. electrical resistance for a temperature standard). In the latter instance, the case can be multivariate, when the involved quantities are more than one (see also Section 6.3.1 and 6.3.2)—“corrections” (see Section 8) are excluded from

this distinction.

In both cases, one refers to one, or a few respectively, of the influence quantities. Invariably, the influence quantities are many more, the remaining being normally labeled ‘corrections’. This distinction can cause confusion in some circumstances—see Section 8.4.

In order to proceed, the experimenter needs to ‘model’ the different aspects of his knowledge.

A model can be implicit, as simple as a list of the input quantities:

$$f\left(Y, X_1,\ \ldots, X_i,\ \ldots, X_N\right) = 0 \qquad (1)$$

where Y is the outcome of the measurement process, the output quantity, intended to realize the measurand, and the X_i are the input quantities (see Section 8 for details on this model in (GUM)). This is called the “measurement equation” (VIM3, GUM) and is a relationship between variables, which are random variables whose realizations are the experimental observations. They are written in capital letters, the usual notation convention; the observations are instead written using the same lower-case letter, and an additional subscript index is added in the usual case of replicated measurements: for example, a replicated realization of X_i is written $x_{i,j}$—the term “replicated” should not be confused with the term “repeated”.

Not necessarily all the X_i are subject to measurement in the experiment in question. For some—or even most—of them the values can be obtained from prior information, typically from the literature, from calibration certificates, or in other ways.

Not infrequently, in the case that an implicit model is used, the effect of an influence quantity included in the list of corrections is given as a “sensitivity coefficient” (GUM), i.e. as the value dY / dX_i in the specific range of interest. In addition, this value is often not obtained from a relationship expressed in analytical form, but from a sub-experiment consisting in varying X_i and observing the variation in Y. However, from this subset of observations an empirical law can be subsequently drawn, e.g. a linear relationship or a more complicated one.

Obviously an implicit model cannot be used if the model is intended for studying the behavior of the measurand when the influence quantities change, as required in simulation studies, either performed analytically or using computed-aided methods like Monte Carlo simulation. In these cases the model must be written in an explicit analytical form, i.e. expressed by specifying the analytical relationships between the quantities. Often this results in complicated implicit expressions, e.g. differential or integral equations, not solvable in closed analytical form, but requiring numerical solutions, e.g., by the use of a finite-elements method. The illustration of this type of models, though not infrequent also in metrology, is out of the scope of this Chapter. The reader can consult the modeling literature, since the features for the models developed for metrological applications are not specific, except for the concept of uncertainty—and the requirement for its evaluation. For modeling in metrology the reader can consult, e.g. (Sommer 2008), together with the involved computational tools (Steele and Douglas 2008a).

The model for data treatment differs from the one in (1) in some features. Instead of a model of the variables of the system, the data model is called 'observation model', and is written instead for the observations, or 'input estimates' $x_{i,j}$ of an influence quantity and 'output estimate' y of the output 'realized quantity', which are realizations of the respective random variables X_i and Y, i.e. are members of the respective probability distributions.

Measurement and data models are based on statistical concepts, of critical importance in measurement science, and metrology in particular, which include modeling of uncertainty. They are treated in detail in Section 5. In modern treatments the related computational tools are very important also in data treatment and analysis.

4. Specific Terms and Concepts for Metrology

Before proceeding with the methods and associated tools to be used for exploiting the treatment of metrological data, let us clarify the meaning in metrology of some of the basic term: the Glossary will refer to Section 4 for these terms. In fact, terms are simply a synthetic way to express concepts. It is therefore essential to be acquainted with them and understand the underlying concepts, because the semantic of the terms can change in time and with the language, leading to possible ambiguities or misunderstandings. For this reason, the most recent Vocabulary of Metrology (VIM) includes also "Basic and General Concepts" (VIM3), and will be used as the basis throughout this Chapter.

4.1. Measurand

The measurand is defined as the "*quantity intended to be measured*", because it may happen that a measurement fails the goal to provide a measure of the aimed quantity (VIM3).

The definition requires understanding the concept of "quantity", indicated as "*property of a phenomenon, body, or substance, where the property has a magnitude that can be expressed as a number and a reference*". Therefore, this definition is valid only for quantitative measurements, and "*a reference can be a measurement unit, a measurement procedure, a reference material, or a combination of such*".

What is not explicitly said above is that any measurand can only be defined by a finite list of details. Therefore, should the random component of a measurement even be reduced to almost nil, the uncertainty on any measurement result cannot decrease below what is called the "*definitional uncertainty*". This residual uncertainty is not due to neither random nor systematic effects, the two traditional components of uncertainty (see later). The definitional uncertainty is not of statistical nature, but is due to the multiplicity of the possible definitions of the measurand, something that should better be called the "*non-uniqueness*" of the definition. Though amply used in metrology (White et al. 2009 and 2010) this term is not (yet) included in VIM (VIM3) nor in GUM (GUM).

In top-level metrology, the non-uniqueness emerges very clearly in many fields as the limiting factor, and, in fact, only the addition of further specification factors to the properties of a measurand can allow further progress. A couple of examples are: the definition of the unit of time, the second, where it is now specified that "*In this definition it is understood that the Cs atom at a temperature of* $T = 0$ K *is unperturbed by blackbody radiation*"; the added specification of the isotopic composition of a substance in thermal metrology.

Measurement results aim at accurately representing the value of the measurand. However, although the measurement result is conceptually different from the value of the measurement, "*for any scalar quantity, subtleties arise only when its uncertainty is not symmetric about that quantity's reported value* (Douglas et al. 2005), *so usually no distinction needs to be made between the distribution of measurements (or 'gedanken' measurements) and the distribution of the measurand (to which formal uncertainty distributions refer). The formal justification for this is facilitated by the standard practice for metrologists of using the same value to be the best representation both of the (fully corrected) measurement, and of the measurand. This 'fiducial value'* (Wang and Iyer 2006, Guthrie et al. 2008, ISO12) *simplifies, and in our view strengthens, the fiducial argument*" (Steele and Douglas 2008a).

4.2. Realized Quantity

The GUM [GUM] introduced the concept of "realized quantity": "*Ideally, the quantity realized for measurement would be fully consistent with the definition of the measurand. Often, however, such a quantity cannot be realized and the measurement is performed on a quantity that is an approximation of the measurand … Neither the value of the realized quantity nor the value of the measurand can ever be known exactly; all that can be known is their estimated values*".

4.3. Compatibility

Since the basic aim of metrology is to ensure as much as possible consistent results from independent measurements for the same measurand, compatibility is a pivotal requirement to be achieved. "*Metrological compatibility*" is defined as "*property of a set of measurement results for a specified measurand, such that the absolute value of the difference of any pair of measured quantity values from two different measurement results is smaller than some chosen multiple of the standard measurement uncertainty of that difference*" (VIM3).

Obviously, the definition bound depends on a "*chosen multiple of the standard measurement uncertainty*", where the latter is the standard deviation, a basic statistical parameter related to the second moment of a probability distribution. The choice of the multiple is left to the contingent decision, since it can be different for different circumstances to fit the purpose. Clearly here the "*differences*" in question are not those between measurement results caused by the variability due to random effects, but are those arising from systematic effects (see Section 6).

The importance of compatibility is so great that in the prescriptive document signed by the metrological Institutes to mutually recognize the validity of their calibrations (MRA) (MRA), it is specified: "*If, as a result of a key comparison, a significant unresolved deviation from the key comparison reference value persists for the standard of a particular participating institute, the existence of this deviation is noted in Appendix C. ... In this case, the institute has the choice of either withdrawing from Appendix C one or more of the relevant calibration and measurement services or increasing the corresponding uncertainties given in Appendix C*", where Appendix C is the database of the calibration services of the NMI (the expression "*significant unresolved deviation from the key comparison reference value*" means 'an inconsistency with respect to').

4.4. Traceability

This term is defined as the "*property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty"* (VIM3). It "*requires an established calibration hierarchy*" and, since each calibration contributes to measurement uncertainty, the latter always increases while descending the ladder of metrological hierarchy. In metrology the term does not only mean ensuring to trace back to the origin by knowing the path, but has also a quantitative qualification.

In principle, traceability does not mean that a step of the metrological ladder should necessarily be compatible with the next. However this is desirable.

Two National Metrology Institutes cannot be said to be traceable with each other, since there is no hierarchy among them (for the standards not obtained from another NMI). They ensure, at the top of the "*traceability chain*" (VIM3), the relationships between them called "degree of equivalence" (MRA), by means of *inter*-comparisons, formally setup in the frame of the MRA, called "*key comparisons*"—see Section 4.7.

4.5. The Error and Uncertainty Approaches

An underpinning basic concept of science, and hence of measurement science, is that, due to imperfect knowledge of the observed phenomena, the numerical data that are the outcomes of measurement are affected by *errors*. Irrespective of the reasons that are the causes of these errors, the resulting dispersion of the measured numerical values that is generally observed is interpreted as evidence of the imperfect knowledge.

Thus, the dispersion of the measured values introduces an *uncertainty* in the measure of the observed phenomena. Uncertainty associated with data is specified according to models that are different according to the underpinning assumptions, which must adequately match the characteristics of the observed phenomena or measurement process.

Uncertainty cannot be determined, but only estimated with a certain ‘confidence’ or ‘degree of belief’. This estimate is, together with assigning values to measurement results, the most fundamental aim and task of metrology.

The definitions of uncertainty and their use are studied since the times of Gauß and are the subject matter of statistics. Information in this respect can be found also in several chapters of the EOLSS, e.g. (Viertl 2003a, 2003b, 2003c, Mari 2012). A restricted set of them is also used in metrology, e.g. (Rossi 2008). For example, probability theory is almost exclusively considered, while possibility theory is not (nor data “imprecision” or “measurement inexactness”); interval-related statistical techniques are not either (Kreinovich 2008). On the other hand, different specific approaches have been developed, namely in GUM (GUM).

-

-

-

Bibliography

AMCTM: Advanced Mathematical and Computational Tools in Metrology and Testing, voll. 16, 40, 45, 53, 57, 66, 72, 78, 84, with various Editors (1993–2011), in Series on *Advances in Mathematics for Applied Sciences*, N. Bellomo Ed., Singapore: World Scientific Publ. Co.. See http:// joomla.imeko.org /index.php /tc21-homepage . [Summing up to more than 360 papers and 3000 printed pages, these books are dedicated, over the last 20 years, to the subject matter of advanced mathematical, statistical and computational tools in metrology. The 2011 book is available also as eBook]

Allevard T., Benoit E., and Foulloy L. (2004). Signal processing on fuzzy nominal scales, in Proceedings 10th IMEKO TC7 International Symposium, Saint-Petersburg, Russia, pp. 478-483. [This paper reports a good example of use of fuzzy statistics]

American Association for Laboratory Accreditation (A2LA) (2002) *Guide for the estimation of measurement uncertainty in testing*. [This is a reference Guide provided by one of the National Associations for Accreditation]

Ballico M. (2001). Calculation of key comparison reference values in the presence of non-zero mean uncertainty distributions, using the maximum likelihood technique, *Metrologia* 38, 155–159. [This paper is one supporting Approach A, Section 5.2.1]

Bremser W. and H□asselbarth W. (2008). Approaches to Data Assessment and Uncertainty Estimation in Testing, Chapter 10 in F. Pavese (Editor) and A.B. Forbes (co-Editor), *Advances in data modeling for measurements in metrology and testing*, Series Modeling and Simulation in Science, Editor N. Bellomo, Boston: Birkhauser-Springer, 2008, with attached DVD. Available also as eBook. [This book Chapter is a comprehensive review of uncertainty estimation in testing]

Chunovkina A.G., Chursin A. (2001). GUM and MRA: some problems of data processing and measurement uncertainty evaluation, in Mathematical and Computations Tools in Metrology V (AMCTM V), P. Ciarlini, M.G. Cox, E. Filipe, F. Pavese and D. Richter, eds., (Series on *Advances in mathematics for applied sciences*) 57, Singapore: World Scientific Publ. Co., pp. 55–66. [This paper is about problems in using a ‘mixture’ distribution, Section 6.3.3]

Chunovkina A.G. (2003). Determining reference values and degrees of equivalence in key comparisons, *Measurement Techniques* 46, 406–414. [This paper is about problems in using a 'mixture' distribution, Section 6.3.3]

Chunovkina A.G., Ciarlini P., Cox M. and Pavese F. (2004). Handling and treatment of measurement data from different groups or sources, in Proceedings 10th IMEKO TC7 International Symposium, pp. 176-181. [This paper is one supporting Approach A, Section 5.2.1]

Chunovkina A.G. and Cox M.G. (2003). A model–based approach to key comparison data evaluation, in Proceedings 17th IMEKO World Congress, Cavtat, Slovenia, pp. 1240-43. [This paper is one supporting Approach B, Section 5.2.2]

Chunovkina A.G., Elster C., Lira I., Wöger W. (2008). Analysis of key comparison data and laboratory biases, *Metrologia* 45, 211–216. [This paper treats the issues in a Bayesian context]

Ciarlini P., Pavese P., Regoliosi G. (2001). A bootstrap algorithm for mixture models and interval data in inter-comparisons (Huddersfield Conference, July 2001) in Algorithms for Approximation IV (edited by I. J. Anderson, J. Levesley and J. C. Mason), University of Huddersfield, Huddersfield, UK, 2002, pp. 138–145. See also in: Defense Technical Information Center, Compilation Part Notice, ADP013724, unclassified. [This paper first introduced in the field of metrology the use of bootstrap]

Ciarlini P., Cox M.G., Pavese F., Regoliosi G. (2004). The use of a mixture of probability distributions in temperature interlaboratory comparisons, *Metrologia* 41, 116–121. [This paper is an improved follow up of the previous]

Ciarlini P., Chunovkina A., Pavese F. and Regoliosi G. (2005). Mixture distribution models in temperature metrology and the assessment of the degree of equivalence, in Proceedings TEMPMEKO 2004, LPM, Zabreb, Croatia, pp 1003–1008. [This paper again uses bootstrap in a specific context]

CITAC: IUPAC/CITAC Guide 2010: Selection and use of proficiency testing schemes for a limited number of participants—chemical analytical laboratories (IUPAC Technical Report), Ilya Kuselman and Aleš Fajgelj, *Pure Appl. Chem*., Vol. 82, No. 5, pp. 1099–1135. [This is a Guide on proficiency tests published by an International Organization]

Colclough A.R. (1987). Two theories of experimental error, *J. Res. Natl. Bur. Stand.* 92, 167–185. [This is paper with critical comments to the Bayesian approach]

Cox M.G. (2002). The evaluation of key comparison data. *Metrologia* 39, 589–595. [This paper contains a proposal for the evaluation of the key comparison data]

Cox M.G., Harris P.M. and Woolliams E. (2006). Data evaluation of key comparisons involving several standards, Advanced Mathematical and Computational Tools in Metrology VII (P.Ciarlini, E.Filipe, A.B.Forbes, F.Pavese and B.Siebert, Eds.), vol.7, Series on *Advances in Mathematics for Applied Sciences* vol.72, Singapore: World Scientific, 2006, pp. 23–34. [This is another paper on the evaluation of the key comparison data]

Cox M.G. (2007). The evaluation of key comparison data: determining the largest consistent subset, *Metrologia* 44, 187–200. [on the same subject, this paper proposes to screen out inconsistent data for the computation of the KCRV]

Cox M.G., Forbes A.B., and Harris P. (2007). Distributions associated with measurands, in *PTB–BIPM Workshop on the Impact of Information Technology in Metrology*, Berlin, Germany. Available at http:// www.npl.co.uk . [This paper contains an analysis of distributions found in measurement and their estimation using Monte Carlo method]

Davis C.S. (2002). *Statistical Methods for the Analysis of Repeated Measurements*, New York:Springer. [A basic book on repeated measurements]

Desanfant M. and Priel M. (2006). Road map for measurement uncertainty evaluation, *Measurement* 39, 841–848. [A paper on treatment of *inter*comparison data]

DIN: DIN 1319-1 1995 *Fundamentals of metrology Part I: basic terminology*, Berlin, Beuth, [A basic standard from the German Standardisation Organisation]

Douglas R.J., Steele A.G., Wood B.M. and Hill K.D. (2005). A useful reflection, *Metrologia* 42, L35–L39. [Interesting reflections on 'measurand', Section 4.1]

Duewer D.L. (2007). How to combine results having stated uncertainties, in *Combining and Reporting*

Analytical Results, ed. A. Fajgelj et al, London:RSC, pp 127–142. [A paper on the practical use of ‘mixture’ distributions]

Duewer D.L. (2008). A comparison of location estimators for interlaboratory data contaminated with value and uncertainty outliers. *Accred. Qual. Assur.* 13, 193–216. [Another paper on the practical use of ‘mixture’ distributions]

EA03: European Accreditation 2003 *EA guidelines on the expression of uncertainty in quantitative testing*, EA(4/16, rev00). [A guide on uncertainty in testing issued by the European Accreditation Organisation]

Elster C., Wöger W. and Cox M.G. (2005). Analysis of Key Comparison Data: Unstable Travelling Standards, *Measurement Techniques* 48, 883–893. [A Bayesian approach to the issue]

Elster C. and Toman B. (2010). Analysis of key comparisons: estimating laboratories’ biases by a fixed effects model using Bayesian model averaging, *Metrologia* 47, 113–119. [A Bayesian approach to the issue]

Fellmuth B., Hill K.D., Bloembergen P., de Groot M., Hermier Y., Matveyev M., Pokhodun A., Ripple D., Steur P.P.M. (CCT-WG1) (2005). Methodologies for the estimation of uncertainties and the correction of fixed-point temperatures attributable to the influence of chemical impurities Working document of the Comité Consultatif de Thermométrie, Doc. CCT/05-08, BIPM, Sévres. Available at http:// www.bipm.org . [A public CCT document introducing the OME method, Section 8.1.2]

Ferson S., Kreinovich V., Hajagos J., Oberkampf W., and Ginzburg L. (2007). Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty, Report SANDIA 2007–0939. [A comprehensive Report of the use of interval statistics in measurement]

Filipe E. (2006). Evaluation of standard uncertainties in nested structures, in in Advances in Mathematical and Computations Tools in Metrology VII (AMCTM VII), P. Ciarlini, E. Filipe, A.B. Forbes, F.Pavese and B.R.L. Siebert, eds., (Series on *Advances in mathematics for applied sciences*) 72, Singapore: World Scientific Publ. Co., pp. 151–60. [A paper discussing the use of ANOVA in metrology]

Forbes A.B. (2006a). Measurement system analysis and statistical process control, in in Advances in Mathematical and Computations Tools in Metrology VII (AMCTM VII), P. Ciarlini, E. Filipe, A.B. Forbes, F.Pavese and B.R.L. Siebert, eds., (Series on *Advances in mathematics for applied sciences*) 72, Singapore: World Scientific Publ. Co., pp. 161–170. [A paper discussing the meaning of ‘systematic error’]

Forbes A.B. (2006b). Measurement uncertainty and optimized conformance assessment. *Measurement* 39, 808–814. [A paper discussing the ‘random laboratory-effects model’]

Forbes A.B. and Perruchet C. (2006). Measurement systems analysis: concepts and computational approaches, in Proceedings IMEKO 18th World Congress, Societade Brasileira de Metrologia, Rio de Janeiro, on CD-ROM. [Another paper on ‘systematic error’ and on the treatment of *inter*comparison data]

Forbes A.B. (2008). Parameter estimation based on Least Squares Methods, Chapter 5 in F. Pavese (Editor) and A.B. Forbes (co-Editor), *Advances in data modeling for measurements in metrology and testing*, Series Modeling and Simulation in Science, Editor N. Bellomo, Boston: Birkhauser-Springer, 2008, with attached DVD. Available also as eBook. [A book Chapter dedicated to the well-known Least Squares Method and its potentialities]

Genest C., Zidek J.V. (1986). Combining probability distributions: a critique and an annotated bibliography, *Stat. Sci.* 1, 114–148. [A paper on the ‘linear opinion pool’, Section 6.3.3 on ‘mixture’ distributions]

Giacomo P. (1982). News from the BIPM, *Metrologia* 18, 41–44. [Introductory document on INC-80, the precursor of GUM]

Girao P.S., Postolache O. and Pereira J.M.D. (2008). Data Fusion, Decision-Making, and Risk Analysis: Mathematical Tools and Techniques, Chapter 7 in F. Pavese (Editor) and A.B. Forbes (co-Editor), *Advances in data modeling for measurements in metrology and testing*, Series Modeling and Simulation in Science, Editor N. Bellomo, Boston: Birkhauser-Springer, 2008, with attached DVD. Available also as eBook. [A book Chapter dedicated to a comprehensive illustration of the indicated subjects]

Gräbe M. (2001). Estimation of measurement uncertainties – an alternative to the ISO Guide, *Metrologia* 38, 99–106. [A paper on an alternative approach to GUM]

Gräbe M. (2005). *Measurement uncertainties in science and technology*, Berlin: Springer. [A book developing in a comprehensive way an alternative approach to GUM, namely on the systematic error]

GUM: (and GUM-S1): BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML 1995 *Guide to the expression of uncertainty in measurement* (GUM), BIPM/ISO, JGCM 100:2008, 1995 version with minor corrections, http:// www.bipm.org /en /publications /guides /gum.html . [An International Guide developed within the BIPM by a group of International Organisations about the treatment of uncertainty]

Guthrie W. F., Hung-kung Liu, Rukhin A. L., Toman B., Wang J. C. M., and Nien-fan Zhang (2008). Three Statistical Paradigms for the Assessment and Interpretation of Measurement Uncertainty, Chapter 3 in F. Pavese (Editor) and A.B. Forbes (co-Editor), *Advances in data modeling for measurements in metrology and testing,* Series Modeling and Simulation in Science, Editor N. Bellomo, Boston:Birkhauser-Springer, 2008, with attached DVD. See also an updated version as: ISO-TR 13567:2012, International Standards Organisation, Geneva, Switzerland. Available also as eBook. [A book Chapter comparing the merits of the frequestist, the Bayesian and the 'fiducial' approaches in the interpretation of measurement uncertainty, with examples]

IEC: IEC 60359 (2001) *Electrical and electronic measurement equipment - Expression of performance*, International Electrotechnical Commission. [A basic standard on uncertainty prepared by IEC]

INC: BIPM 1980 Report on the BIPM Enquiry on Error Statements, Rapport BIPM 80/3, Bureau International des Poids et Mesures, Sèvres, http:// www.bipm.org . [This Guide was the precursor of GUM, and is now incorporated in it]

ISO93: ISO 3534-2 (1993) 2nd edn *Statistics – vocabulary and symbols –Part 2: applied statistics*, International Organization for Standardization, Geneva. [An ISO standard on terminology, namely for the testing field]

ISO94: ISO 5725 (1994) 2nd edn *Accuracy (trueness and precision) of measurement methods and results*, International Organization for Standardization, Geneva, Switzerland. New edition (ISO 15725) in progress. [The basic ISO standard on accuracy of measurement methods, presently under revision]

ISO99: ISO 3534-3 (1999) *Statistics – vocabulary and symbols – Part 3: design of experiments*, International Organization for Standardization, Geneva, Switzerland, http:// www.iso.org . [An ISO standard on terminology, namely for the testing field]

ISO04: ISO TS 21748 (2004) *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, International Standards Organisation, Geneva, Switzerland, http://www.iso.org. [An ISO Guide to the use of ISO 5725:1994]

ISO05: ISO 21749 (2005) *Measurement and uncertainty for metrological applications – Repeated measurements and nested experiments*, International Organization for Standardization, Geneva, Switzerland, http://www.iso.org. [An ISO document for guidance specifically for the metrological field]

ISO05b: ISO 13528 (2005) *Statistical methods for use in proficiency testing by interlaboratory comparisons*, International Organization for Standardization, Geneva, Switzerland. New edition in progress. [The basic ISO standard for proficiency testing]

ISO06: ISO 3534-1 (2006) 2nd edn *Statistics – Vocabulary and symbols – Part 1*: General statistical terms and terms used in probability, International Organization for Standardization, Geneva, Switzerland, http:// www.iso.org. [The revision of the ISO standard of 1993]

ISO06b: ISO 3534-2 2006 3rd edn *Statistics – Vocabulary and symbols – Part 2*: Applied statistics, International Organization for Standardization, Geneva, Switzerland, http:// www.iso.org. [The revision of the ISO standard of 1993]

ISO10: ISO/IEC 17043 2010, *Conformity assessment – General requirements for proficiency testing* (previously ISO Guide 43-1) International Organization for Standardization, Geneva, Switzerland, http://www.iso.org. [The other basic ISO standard for proficiency testing]

ISO12: ISO-TR 13567:2012 *Three Statistical Paradigms for the Assessment and Interpretation of Measurement Uncertainty*, International Standards Organisation, Geneva, Switzerland (updated from Guthrie et al. 2008, see above) , http://www.iso.org. [Published by ISO as an update to the indicated book

Chapter]

Iyer H.K., Wang C.M., and Mathew T. (2004) . Models and confidence intervals for true values in interlaboratory trials. *J. Am. Stat. Assoc.* 99, 1060-1071. [A paper on models for *inter*comparisons]

Kaarls R. (1981). Report on the meeting of the BIPM working group on the statement of uncertainties, in Proces Verbaux des séances du Comité International des Poids et Mesures 49 A1–A12 (in French). [Introductory document on INC-80, the precursor of GUM]

Kacker R.N., Datla R.U. and Parr A.C. (2002). Combined result and associated uncertainty from interlaboratory evaluations based on the ISO Guide, *Metrologia* 39, 279–293. [A comprehensive paper on uncertainty in *inter*comparisons]

Kacker R.N. (2004). Combining information from interlaboratory evaluations using a random effects model, *Metrologia* 41, 132–136. [A followup paper focused on random effect models]

Kacker R.N., Datla R.U., Parr A.C. (2004). Statistical analysis of CIPM key comparisons based on the ISO Guide, *Metrologia* 41, 340–252. (Paper on the distinction of standards in Type 1 and Type 2, Section 6.2]

Kacker R.N., Toman B., and Huang D. (2006). Comparison of ISO–GUM, draft GUM Supplement 1 and Bayesian statistics using simple linear calibration, *Metrologia* 43, S167–S177. [Paper on data modeling, based on GUM Supplement 1 and the Bayesian approach]

KCDB: BIPM KC database, available at http:// kcdb.bipm.org .

Kreinovich V. (2008). Interval Computations and Interval-Related Statistical Techniques: Tools for Estimating Uncertainty of the Results of Data Processing and Indirect Measurements, Chapter 4 in F. Pavese (Editor) and A.B. Forbes (co-Editor), *Advances in data modeling for measurements in metrology and testing*, Series Modeling and Simulation in Science, Editor N. Bellomo, Boston: Birkhauser-Springer, 2008, with attached DVD. Available also as eBook. [Book Chapter on the interval statistics for use in metrology]

Levenson M.S., Banks D.L., Eberhart K.R., Gill L.M., Guthrie W.F., Liu H.K., Vangel M.G., Yen J.H. and Zhang N.F. (2000). An approach to combining results from multiple methods motivated by the ISO GUM, *J. Res. Natl. Inst. Stand. Technol.* 105, 571–579. [A paper on evaluating uncertainty for uncorrected measured data]

Lira, I., and Wöger, W. (1998). Evaluation of the uncertainty associated with a measurement result not corrected for systematic effects, *Meas. Sci. Technol.* 9, 1010–1011; The evaluation of the uncertainty in knowing a directly measured quantity, *Meas.Sci.Technol.* 9, 1167–1173. [Two linked papers on the RSS*u* method for using uncorrected data, Section 8.2.1]

Lira I. (2006). Bayesian evaluation of comparison data, *Metrologia* 43, S231–S234. [Paper on a Bayesian approach for evaluating comparison data]

Lira I. and W□oger W. (2006). Evaluation of repeated measurements from the viewpoints of conventional and Bayesian statistics, in Advances in Mathematical and Computations Tools in Metrology VII (AMCTM VII), P. Ciarlini, E. Filipe, A.B. Forbes, F.Pavese and B.R.L. Siebert, eds., (Series on Advances in mathematics for applied sciences) 72, Singapore:World Scientific Publ. Co., pp 73–84. [A comparison of the frequentist and Bayesian approaches]

Lira I. and W□oger W. (2006b). Comparison between the conventional and Bayesian approaches to evaluate measurement data, *Metrologia* 43, S249–S259. [Same argument as above]

Magnusson B., and Ellison S.L.R. (2008). Treatment of uncorrected measurement bias in uncertainty estimation for chemical measurements, *Anal. Bioanal. Chem.* 390, 201–213. [A basic paper comparing methods for using uncorrected data in testing, Section 8.2.1]

Mari, L. (2012),MEASUREMENT SCIENCE AND ITS EVOLUTION, in *Physical Methods, Instruments and Measurements*, [Eds. UNESCO-EOLSS Joint Committee], in *Encyclopedia of Life Support Systems(EOLSS),* Developed under the Auspices of the UNESCO, Eolss Publishers, Oxford ,UK, [http:// www.eolss.net] [Retrieved April 2, 2013]

McLachlan G., Peel D. (2000). *Finite mixture models*, New York: Wiley. [A basic book on mixture models]

MRA: CIPM 1999 Mutual recognition of national measurement standards and of calibration and measurement certificates issued by national metrology institutes, Bureau International des Poids et Mesures, Sèvres. http://www.bipm.org. [Formal document to be signed by all NMIs who agree to recognize reciprocally their calibration certificates]

Neter J., Wasserman W. and Whitmore G.A. (1988). *Applied statistics*, Boston: Allyn and Bacon. [An old but always useful book on applied statistics]

Searle S.R., Casella G. and Mc Culloch C.E. (1992). *Variance components*, New York: Wiley. [An old but always useful book on variance components]

Nielsen L. (2000). Evaluation of measurement intercomparisons by the method of least squares, Report DFM-99-R39, 3208 LN, Danish Institute of Fundamental Metrology, Lungby, Denmark. [Document on using least squares for the data treatment of *inter*comparisons]

NIST06: National Institute of Standards and Technology (NIST) *2006 Engineering statistics handbook* (e-Handbook), NIST, USA. Available online at http://www.nist.gov/stat.handbook . [A comprehensive guide on the use of statistics in measurement, with examples]

Paule R.C. and Mandel J. (1982). Consensus values and weighting factors, *J. Res. Natl. Bur. Stand.* 87, 377–385. [A paper on *inter*comparison data]

Pavese F. and Ciarlini P. (1990). Accurate modeling of translational bias and its application to reduction of thermophysical data series, *Metrologia* 27, 145–152. [A paper using the method LSMFE, Section 5.1.4]

Pavese F. (2000). *Mathematical problems in the definition of standards based on scales: the case of temperature, in Advanced Mathematical and Computational Tools in Metrology* (P. Ciarlini, A.B. Forbes, F. Pavese and D. Richter, Eds.), vol.4, Series on Advances in Mathematics for Applied Sciences vol.53, Singapore: World Scientific, 2000, pp 182–196. [A paper on Class 2 standards]

Pavese F., Ichim I. and Ciarlini P. (2001). An algorithm for on-line outlier rejection by sequence analysis in data acquisition: the equispaced data case, in *Advanced Mathematical and Computational Tools in Metrology V* (P. Ciarlini, M.G. Cox, E. Filipe, F. Pavese and D. Richter, Eds.), vol.5, Series on Advances in Mathematics for Applied Sciences vol. 57, Singapore: World Scientific, 2001, pp. 283–291. [A paper on sequential rejection of outliers, also applicable during data acquisition, with software on http://www.amctm.org]

Pavese F., Ciarlini P. (2003). Classes of inter-comparisons and the case of temperature standards, *PTB Berichte* PTB-IT-10, 63–76. [A paper on Class 1 and Class 2 standards]

Pavese F., Ichim D., Ciarlini P., Balle C. and Casas-Cubillos J. (2003b). Detection of thermometer clustering in the calibration of large batches of industrial thermometers for the LHC by automated data processing, in Temperature, Its Measurement and Control in Science and Industry 8, D Ripple ed, New York:AIP, pp 429–433. [A paper on an application of the method LSMFE]

Pavese F. (2004). Compound modelling of metrological data, in *Advanced Mathematical and Computational Tools in Metrology VI* (P.Ciarlini, M.G.Cox, F.Pavese and G.B.Rossi, Eds.) vol.6, Series on Advances in Mathematics for Applied Sciences vol.66, Singapore: World Scientific, 2004, pp. 240–250. [A paper on data taken ins eries]

Pavese F. (2005). Comments on Statistical analysis of CIPM key comparisons based on the ISO Guide, *Metrologia* 42, L10-L12.

Pavese F. (2006). A metrologist viewpoint on some statistical issues concerning the comparison of non-repeated measurement data, namely MRA Key Comparisons, *Measurement* 39, 821–828. [Comment on the paper of Kacker et al. 2004]

Pavese F. and Filipe E. (2006). Some metrological considerations about replicated measurements on standards, *Metrologia* 43, 419–425. [A comprehensive paper about modeling comparison data]

Pavese .F (2007) Replicated observations in metrology and testing: modelling repeated and non-repeated measurements, *Accred. Qual. Assur.* 12, 525–534. [A comprehensive paper on repeated and nonrepeated measurements]

Pavese F. (2007b). The definition of the measurand in key comparisons: lessons learnt with thermal standards, *Metrologia* 44, 327–339. [A paper on the definition of 'measurand' and Classes of standards]

Pavese F. (2008). An Introduction to Data Modeling Principles in Metrology and Testing, Chapter 1 F. Pavese (Editor) and A.B. Forbes (co-Editor), Advances in data modeling for measurements in metrology and testing, Series Modeling and Simulation in Science, Editor N. Bellomo, Boston:Birkhauser-Springer, 2008, , with attached DVD. Available also as eBook. [A book Chapter devoted to data modeling]

Pavese F. (2008b). On the need to tackle GUM requirement to perform the corrections for recognized systematic effects: the effects of chemical impurities on the realisations of fixed points of the ITS-90, Doc. CCT/24-05 BIPM, Sèvres, at http://www.bipm.org . [A public CCT Document on specific problems affecting correction of measured data]

Pavese F. (2009). On some consequences of the different nature of *within*- and *between*-laboratory data, *Metrologia* 46, L29–L32. [A short by dense paper about the basic difference between *within*- and *between*-laboratory knowledge]

Pavese F. (2009b). Critical review of information relevant to the correction of the effect of chemical impurities in gases used for the realization of ITS-90 fixed points, *Metrologia* 46, 47–61. [A paper on data correction in thermal metrology]

Pavese F. (2011). Methods for the assessment of correction for chemical-impurity effects and related uncertainty in ITS-90 fixed points, namely of e-H_2, Ne, O_2 and Ar, *Metrologia* 48, 268–274. [A paper about the problems in correcting measured values, in a specific domain]

Pavese F. (2012a). On the difference of meanings of 'zero correction': zero value vs. no correction, and of the associated uncertainties, in *Advanced Mathematical and Computational Tools in Metrology and Testing IX* (F. Pavese, M. Bär, J.-R. Filtz, A. B. Forbes, L. Pendrill and K. Shirono, Eds.), vol.9, Series on Advances in Mathematics for Applied Sciences vol. 84, Singapore: World Scientific, 2012, pp. 297–309. [A paper about correcting or non correcting measured data in the specific case when the value of the correction is said to be zero, Section 8]

Pavese F. (2012b). Need for consistency of terminology in international standards and guidelines. A case study: trueness, in *Advanced Mathematical and Computational Tools in Metrology and Testing IX* (F. Pavese, M. Bär, J.-R. Filtz, A. B. Forbes, L. Pendrill and K. Shirono, Eds.), vol.9, Series on Advances in Mathematics for Applied Sciences vol. 84, Singapore: World Scientific, 2012, pp. 310–316. [A paper on terminology, enumerating present inconsistencies in international standards and guides]

Pavese F. (2012c). Corrections and input quantities in GUM-compliant models, in Proceedings IMEKO World Congress 2012, Busan, Korea, September 2012, later published on the new online Journal Acta IMEKO. [A paper on the distinction between 'input quantities' and 'corrections', Section 8.4]

Phillips S.D., Eberhardt K.R. and Parry B. (1997). Guidelines for Expressing the Uncertainty of Measurement Results Containing Uncorrected Bias, *J. Res. Natl. Inst. Stand. Technol.* 102, 577–585. [Paper on the SUM*U* approach for using uncorrected data, Section 8.2.1]

QUAM: Eurachem CITAC 2000 2nd edn Guide CG4. *Quantifying uncertainty in analytical measurements* (QUAM 2000.1). [Guide on quantifying uncertainty issued by the International Organisation Eurachem]

Rossi G.B. (2008). Probability in Metrology, Chapter 2 in F. Pavese (Editor) and A.B. Forbes (co-Editor), *Advances in data modeling for measurements in metrology and testing,* Series Modeling and Simulation in Science, Editor N. Bellomo, Boston: Birkhauser-Springer, 2008, with attached DVD. Available also as eBook. [Book Chapter devoted to a comprehensive examination of the use of probability in metrology]

Ruhm K.H. (2007). Expectation–A fuzzy term in metrology, in Proceedings IEEE International Workshop on Advanced Methods of Uncertainty Estimation in Measurement (AMUEM 2007), IEEE, ISBN-1-4244-0933-0, pp. 53–58. [A paper devoted to the meaning of 'expectation' in metrology]

Rukhin A.L. and Vangel M.G. (1998). Estimation of a common mean and weighted means statistics, *J. Am. Stat. Assoc.* 93, 303–308. [A paper on two common summary statistics in measurement]

Schiller S.B. and Eberhardt K. (1991). Combining data from independent chemical analysis methods, *Spectrochim. Acta* 46B, 1607–1613. [A paper on uncorrecting data and evaluating their uncertainty]

Sommer K.D. and Siebert B.R.L. (2006). Systematic approach to the modelling of measurements for uncertainty evaluation, *Metrologia* 43, S200–S210. [A comprehensive paper on modeling of uncertainty]

Sommer K.D. (2008). Modelling of Measurements, System Theory and Uncertainty Evaluation, Chapter 9 in F. Pavese (Editor) and A.B. Forbes (co-Editor), *Advances in data modeling for measurements in metrology and testing*, Series Modeling and Simulation in Science, Editor N. Bellomo, Boston:Birkhauser-Springer, 2008, with attached DVD. Available also as eBook. [A book Chapter devoted to a comprehensive review of modeling for uncertainty]

Steele A.G., Hill K.D. and R.J. Douglas (2002). Data pooling and key comparison reference values, *Metrologia* 39, 269–277. [A paper on pooling data from several laboratories in *inter*comparisons]

Steele A.G., Hill K.D. and Douglas R.J. (2004) Pooled data distributions: graphical and statistical tools for examining comparison reference values, in *Advances in Mathematical and Computations Tools in Metrology VI* (AMCTM VI), P. Ciarlini, M.G. Cox, F.Pavese and G.B. Rossi, eds., (Series on Advances in mathematics for applied sciences) 66, Singapore: World Scientific Publ. Co., pp. 264–273. [A followup paper on pooling data from several laboratories in *inter*comparisons, with graphical methods]

Steele A.G. and Douglas R.J. (2008a). Monte Carlo Modeling of Randomness, Chapter 11 in F. Pavese (Editor) and A.B. Forbes (co-Editor), *Advances in data modeling for measurements in metrology and testing*, Series Modeling and Simulation in Science, Editor N. Bellomo, Boston:Birkhauser-Springer, 2008, with attached DVD. Available also as eBook. . [A book Chapter devoted to pooling data from several laboratories in *inter*comparisons, treated with the Monte Carlo method]

Toman B. and Possolo A. (2009). Laboratory effects models for interlaboratory comparisons, *Accred. Qual. Assur.* 14, 553–563. [A comprehensive paper in modeling laboratory effects]

Urbanski M.K. and Wasowski J. (2003). Fuzzy approach to the theory of measurement inexactness, *Measurement* 34, 67–74. [A paper of fuzzy statistics]

Viertl, R. (2003a), PROBABILITY AND STATISTICS, in *Probability and Statistics*, [Ed. Reinhard Viertl], in *Encyclopedia of Life Support Systems (EOLSS)*, Developed under the Auspices of the UNESCO, Eolss Publishers, Oxford ,UK, [http://www.eolss.net] [Retrieved April 2, 2013]

Viertl R. (2003b), STATISTICAL INFERENCE WITH IMPRECISE DATA, in *Probability and Statistics*, [Ed. Reinhard Viertl], in *Encyclopedia of Life Support Systems (EOLSS)*, Developed under the Auspices of the UNESCO, Eolss Publishers, Oxford ,UK, [http:// www.eolss.net] [Retrieved April 2, 2013]

Viertl R. (2003c), FOUNDATIONS OF STATISTICS, in *Probability and Statistics*, [Ed. Reinhard Viertl], in *Encyclopedia of Life Support Systems (EOLSS)*, Developed under the Auspices of the UNESCO, Eolss Publishers, Oxford ,UK, [http://www.eolss.net] [Retrieved April 2, 2013]

VIM2: BIPM (1993) *International Vocabulary of Basic and General Terms in Metrology* (VIM), 2nd edn, BIPM/ISO, Bureau International des Poids et Measures, Sèvres. [Previous edition of VIM]

VIM04: BIPM (2004) *International Vocabulary of Basic and General Terms in Metrology* (VIM) (April 2004). [Previous VIM draft of 2004]

VIM3: BIPM (2012) *International Vocabulary of Basic and General Terms in Metrology* (VIM), 3rd edn, BIPM/ISO, JGCM 200:2012. 2008, version with minor corrections, http:// www.bipm.org /en /publications /guides /vim.html. [Current International Vocabulary for metrology developed within the BIPM by a group of International Organisations about the terminology]

Wang C.M. and Iyer H.K. (2006). Uncertainty analysis for vector measurands using fiducial inference, *Metrologia* 43, 486–494. [A paper on the use of the 'fiducial' statistical method]

Wang C.M. and Iyer H.K. (2006b). A generalized confidence interval for a measurand in the presence of type–A and type–B uncertainties, *Measurement* 39, 856–863. [A paper on the meaning of 'confidence interval' in the GUM context]

White D.R. (2000). On the analysis and linking of comparison results, in CPEM Conference digest, CPEM 2000, Sydney, Australia, pp 325–326. [A very early paper on the problem of linking different comaprisons]

White D.R. (2004). On the analysis of measurement comparisons, *Metrologia* 41, 122–131. [A comprehensive analysis of measurement comparisons]

White D.R., Bonnier G., Diril A., Arai M., Ballico M., Chimenti V., Duris S., Filipe E., Ivanova A., Kartal-Dogan A., Mendez-Lango E., Meyer C., Pavese F., Peruzzi A., Renaot E., Rudtsch S., Stock M., and Ugur S. (2009). WG3 Report to CCT, Doc. CCT/05-15rev (2005) and Doc. CCT/08/19rev (2009), Bureau International des Poids et Measures, Sèvres, http://www.bipm.org . [A public CCT document on the treatment of uncertainty, specifically in thermal metrology]

White D.R., Ballico M., del Campo D., Duris S., Filipe E., Ivanova A., Kartal Dogan A., Mendez-Lango E., Meyer C., Pavese F., Peruzzi A., Renaot E., Rudtsch S., Yamazawa K., and Zhang J.T. (2010). Uncertainties in the SPRT sub-ranges of ITS-90: topics for further research, *Int. J. Thermophysics* 31, 1749–1761. [A paper on the previous subject, with the indication of areas requiring more studies]

Willink R. (2006) Meaning and models in key comparisons, with measures of operability and interoperability, *Metrologia* 43, S220–S230. [A paper of a peculiar treatment of models in intercomparisons, Section 5.2.3]

Willink R. (2006b). Forming a comparison reference value from different distributions of believe, *Metrologia* 43,12–20. [Another paper on the problem of the KCRV]

Willink R. (2006c). Principles of probability and statistics for metrology, *Metrologia* 43, S211–S219. [A comprehensive paper on a viewpoint about probability and statistics in metrology –to compare with Rossi 2008]

Biographical Sketch

Franco Pavese, graduated in Engineering, Principal Scientist (Research Director) at the National Research Council of Italy, Institute of Metrology "G. Colonnetti" (IMGC-CNR) from 1967 to 2005, then at the National Institute of Research in Metrology (INRIM) until 2008.

- Chairman of IMEKO Technical Committee 21 (TC21) "Mathematical Tools for Measurements" (2004–)
- Founder and Chairman of the International Conferences "Advanced Mathematical and Computational
- Tools in Metrology and Testing" (AMCTM) (1993–), 9 held until 2011.
- Coordinator of EU Research Projects on the issues of TC21 (Euroconferences Contracts
- ERBCHECCT940254, ERBFMMACT970244; Thematic Network "SoftTools Metronet" G6RT-CT
- 2001-05061) (1995–2005)
- Coordinator of several research Contracts with CERN (Geneva) on statistical treatment of LHC calibration data (1994–2007)
- Coordinator of iMERA/EURAMET Project "Softools" (2005–2009)
- Italian Delegate to ISO Technical Committee 69 "Applications of statistical methods" (2005–)
- Affiliated Scientist of Institute for Low Temperature Research, Poland
- Member of the Comité Consultatif de Thermométrie (CCT) (1980–2010) and occasionally Italian Delegate
- Chairperson of the CCT Task Group "Strategy" (2008–2012; past Chairperson of Working Group 7 "Key
- Comparisons"; Member of Working Group 3 "Uncertainty"
- Member of Commission A2 of the International Association IIF/IIR
- Titular Member, then National Representative, of IUPAC Commission I-1 (1991–)
- President of the International non-profit Society "The evitherm Society", originated from the European
- Project EVITHERM (2006–)
- Coordinator of 3 EU Research Projects (1999–2008) and of several Italian National Research Projects on
- thermal metrology and cryogenic engineering
- Past Italian Representative in VAMAS, Area 15 "Cryogenic Structural Materials" and Area 16

- "Cryogenic Superconducting Materials"
- Editor of 13 multi-author International Books and Journal Special Issues.
- Member of the Editorial Boards of the: Journal of Chemical Thermodynamics (past), Journal
- Measurement and International Journal of Metrology and Quality Engineering.

Author of more than 20 Books, Monographs and book Chapters, of more than 200 scientific papers, most international, and of 3 patents, in the field of thermal metrology, cryogenic engineering, HTC superconductors and mathematics & statistics in measurement science.