Workshops

- Wednesday, November 20, 12:00 PM to 2:00 PM:
  Models of measurement: the general structure

- Thursday, November 21, 9:00 AM to 11:00 AM:
  Models of measurement: measuring systems and metrological infrastructure

- Thursday, November 21, 2:00 PM to 4:00 PM:
  An overview on measurement uncertainty: from the standpoint of the Guide to the Expression of Uncertainty in Measurement (GUM)

- Friday, November 22, 10:00 AM to noon:
  Is the body of knowledge on measurement worth to be a ‘science’, and what may be the scope of a measurement science?

Workshop 3

An overview on measurement uncertainty: from the standpoint of the Guide to the Expression of Uncertainty in Measurement (GUM)

Luca Mari

Università Cattaneo – LIUC, Italy

University of California, Berkeley
Thursday, November 21, 2013

Abstract

The concept of measurement uncertainty offers some new connotations with respect to the traditional way the quality of measurement results has been represented, in a more and more encompassing path from ontology (true value and error), to epistemology (degree of belief), to pragmatics (target measurement uncertainty). The workshop presents a conceptual framework in which measurement uncertainty is interpreted as an overall property, synthesizing both instrumental and definitional contributions.

My profile

Luca Mari (M.Sc. in physics; Ph.D. in measurement science) is full professor of measurement science at the Cattaneo University – LIUC, Castellanza (VA), Italy, where he teaches courses on measurement science, statistical data analysis, system theory.

He is currently the chairman of the TC1 (Terminology) and the secretary of the TC25 (Quantities and Units) of the International Electrotechnical Commission (IEC), and an IEC expert in the WG2 (VIM) of the Joint Committee for Guides in Metrology (JCGM). He has been the chairman of the TC7 (Measurement Science) of the International Measurement Confederation (IMEKO). He is the author or coauthor of several scientific papers published in international journals and international conference proceedings. His research interests include measurement science and system theory.

Some of my recent publications

LM, **A quest for the definition of measurement**, *Measurement*, 2013

LM, A.Giordani, **Quantity and quantity value**, *Metrologia*, 2012

LM, P.Carbone, D.Petri, **Measurement fundamentals: a pragmatic view**, *IEEE Trans. Instr. Meas.*, 2012

A.Giordani, LM, **Measurement, models, uncertainty**, *IEEE Trans. Instr. Meas.*, 2012

A.Giordani, LM, **Property evaluation types**, *Measurement*, 2012

A.Frigerio, A.Giordani, LM, **Outline of a general model of measurement**, *Synthese*, 2010

D.Macii, LM, D.Petri, **Comparison of measured quantity value estimators in nonlinear models**, *IEEE Trans. Instr. Meas.*, 2010

LM, V.Lazzarotti, R.Manzini, **Measurement in soft systems: epistemological framework and a case study**, *Measurement*, 2009

LM, **A computational system for uncertainty propagation of measurement results**, *Measurement*, 2009

LM, **On (kinds of) quantities**, *Metrologia*, 2009

LM, **The problem of foundations of measurement**, *Measurement*, 2005

LM, **Epistemology of measurement**, *Measurement*, 2003

LM, **Beyond the representational viewpoint: a new formalization of measurement**, *Measurement*, 2000

Stereotypes

«If your experiment needs statistics
you ought to have done a better experiment.»

Ernest Rutherford, ~1910

«The more important fundamental laws and facts
of physical science have all been discovered.
Our future discoveries must be looked for
in the sixth place of decimals.»

Albert Michelson, 1899

More stereotypes

[Dealing with uncertainties]

«is interesting and important
but is not part of the foundations of measurement *per se*»

P. Suppes et al. 1989

Joint Committee for Guides in Metrology (JCGM)

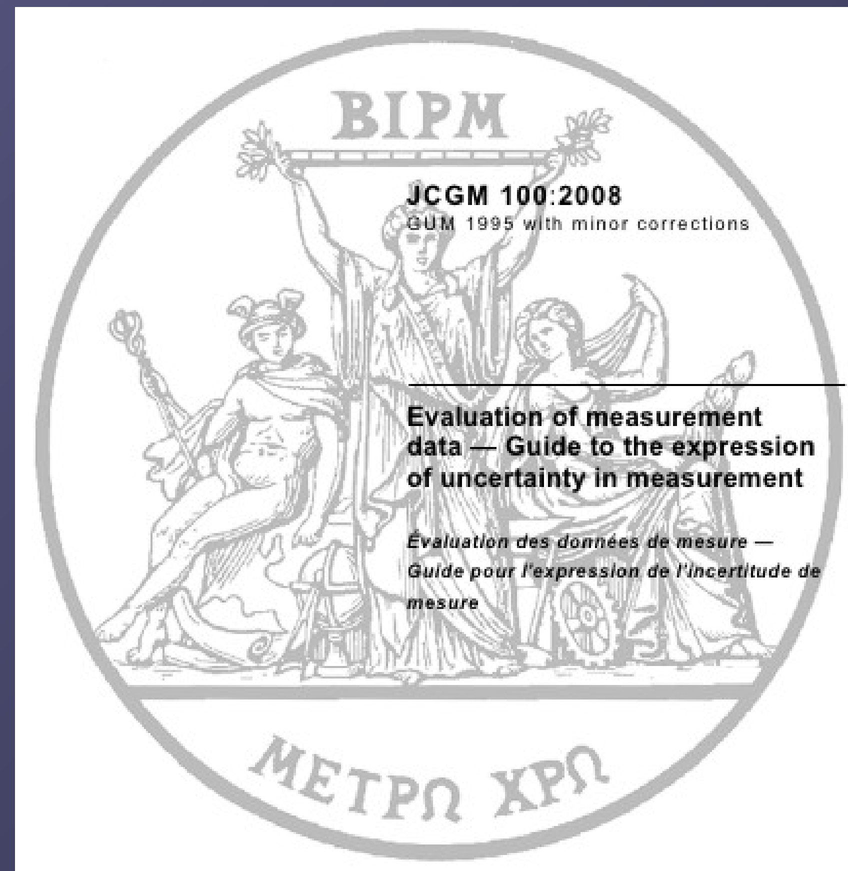
(established in 1997)

(BIPM) Int.l Bureau of Weights and Measures
(IEC) Int.l Electrotechnical Commission
(IFCC) Int.l Federation of Clinical Chemistry and Laboratory Medicine
(ILAC) Int.l Laboratory Accreditation Cooperation
(ISO) Int.l Organization for Standardization
(IUPAC) Int.l Union of Pure and Applied Chemistry
(IUPAP) Int.l Union of Pure and Applied Physics
(OIML) Int.l Organization of Legal Metrology

JCGM guidance docs

the "VIM"

the "GUM"



http://www.bipm.org/en/publications/guides/ vim.html
gum.html

A pragmatic justification

«When reporting the result of a measurement of a physical quantity, it is obligatory that **some quantitative indication of the quality of the result** be given so that those who use it can assess its reliability.
Without such an indication, measurement results cannot be compared, either among themselves or with reference values given in a specification or standard.»

[GUM, Introduction]

Error and uncertainty

«All measurements, however careful and scientific, are subject to some **uncertainties**. **Error analysis** is the study and evaluation of these **uncertainties**, its two main functions being to allow the scientist to estimate how large his **uncertainties** are, and to help him to reduce them when necessary. The **analysis of uncertainties**, or **errors**, is a vital part of any scientific experiment, and **error analysis** is therefore an important part of any college course in experimental science.»

quoted from a (well known) university book

(2nd edition, 1997)

Hypothesis: **uncertainty ≈ error**

Consequence: **nothing really new under the sun**

A customary situation

Having variously interacted, with the support of various devices, with an object ω, someone states:

"the length of ω is 1.234 m"

and writes:

"length(ω) = 1.234 m"

that is:

- the **measurand** («quantity intended to be measured» [VIM3]):

length(ω)

is (equal to / represented by)

- the **quantity value** («number and reference together expressing magnitude of a quantity» [VIM3]):

1.234 m

The customary interpretation

Does it mean: length(ω) = 1.2340000000000000000... m ?

No, of course: measured quantity values can only have a finite number of “significant digits”

so that 1.234 m should be interpreted as 1.234±0.0005 m,
i.e., as the interval [1.2335, 1.2345] m

This is traditionally accounted for in terms of:

- **finite precision** of measuring instruments
- **error** affecting measurement results

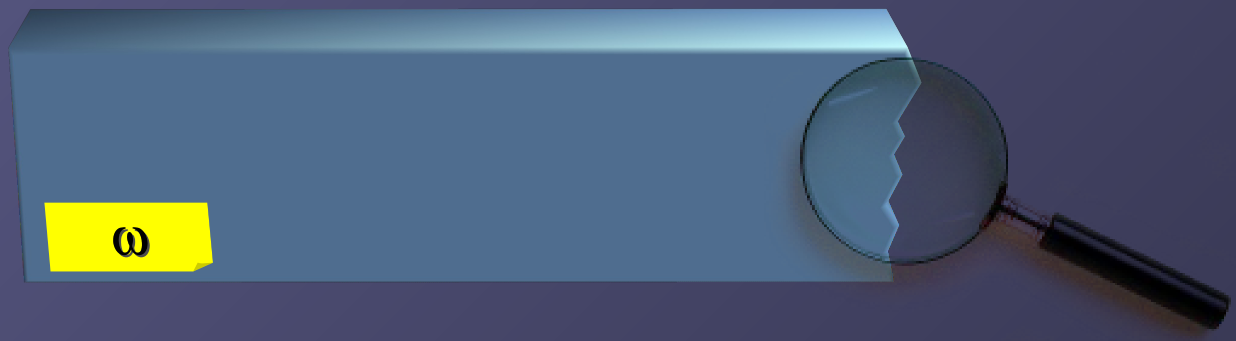
under the hypotheses that:

1. the measurand is a quantity given independently of measurement
2. it has a **true value**, the one that would be produced by an ideal measuring system operating in the best measurement conditions

True value?

Talking about length(ω) implies assuming that ω has a (and usually one) length, and therefore adopting a (geometric) model of X such that, e.g., its shape is assumed to be rectangular

→ length(ω) is defined only within a **model of the measurand**



Truth is primarily asserted in reference to a model, not to "reality"

True value? (2)

The model of the measurand includes specifications on the dependence of the measurand on environmental conditions

→ length(ω) is acknowledged as depending on **influence properties**

What do you want to measure? Three options (the measurand is...):

1. what is measured here and now
2. what is measured in the current conditions to be assessed
3. what would be measured in some reference conditions

True value? (3)

The measurand is:

1. **what is measured here and now**

→ simplest option but information from measurement is of limited use

2. **what is measured in the current conditions to be assessed**

→ influence properties must be measured

3. **what would be measured in some reference conditions**

→ most general option but influence properties must be measured
and a “correction” model must be (known and) applied

(if influence properties must be measured, they become measurands of new measurement problems...)

True value? (4)

The measurand is:

VIM2 (1993): «quantity subject to measurement»

VIM3 (2008): «quantity intended to be measured»

so that a new concept arises:

definitional uncertainty: «component of measurement uncertainty resulting from the finite amount of detail in the definition of a measurand» [VIM3]

A short story

«In 1977, recognizing the lack of international consensus on the expression of uncertainty in measurement, the world's highest authority in metrology, the *Comité International des Poids et Mesures* (CIPM), requested the *Bureau International des Poids et Mesures* (BIPM) to address the problem in conjunction with the national standards laboratories and to make a recommendation.»

[GUM, Foreword]

Relevant standards

Together with the GUM and the VIM,

ISO 3534: Statistics – Vocabulary and symbols

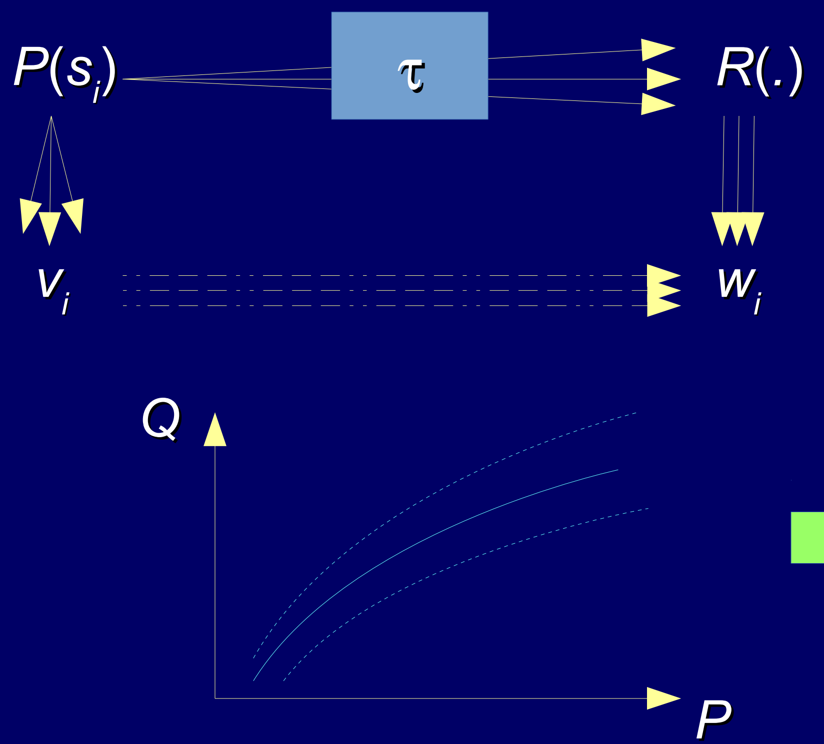
ISO 5725: Accuracy (trueness and precision) of measurement methods and results

A general model

	Features
Measurement results	**Measurement uncertainty**
Measuring system	**accuracy**
Measuring instrument	**trueness (*location statistic*), precision (*dispersion statistic*)**
Transducer	**sensitivity, selectivity, repeatability, resolution, …**

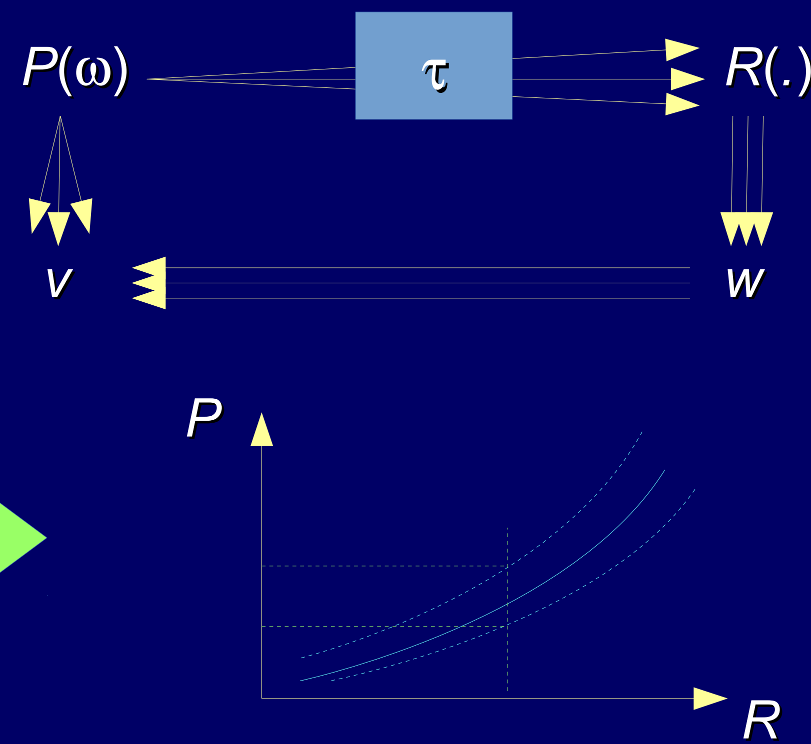
A general model (2)

Calibration



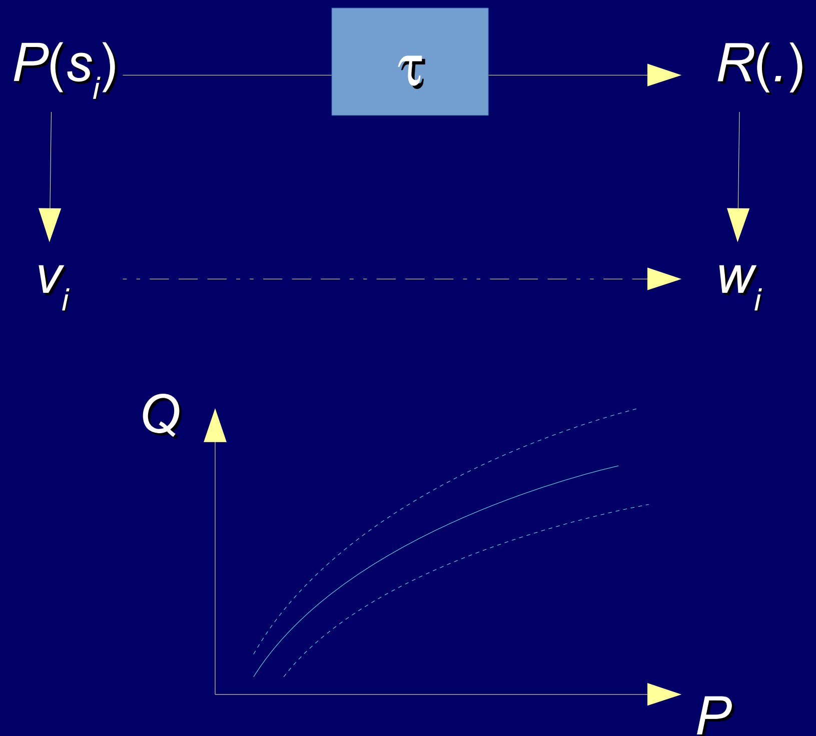
given some *P*-values
and their uncertainties,
find the corresponding *R*-values
and their uncertainties
and then construct the **strip**

Measurement



given an *R*-value,
find the corresponding *P*-value
and its uncertainty
in the inverted **strip**

Calibration, once again



given some *P*-values
and their uncertainties,
find the corresponding *R*-values
and their uncertainties
and then construct the **strip**

«operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication» [VIM3]

The traditional problem

How to combine

a **location** statistic (such as mean),
conveying information on instrument **trueness**,

and

a **dispersion** statistic (such as standard deviation),
conveying information on instrument **precision**,

into

an **overall** statistic,
conveying information on system **accuracy**?

Sources of uncertainty

«There are many possible sources of uncertainty in a measurement, including:

a) incomplete definition of the measurand;

b) imperfect realization of the definition of the measurand;

c) nonrepresentative sampling – the sample measured may not represent the defined measurand;

d) inadequate knowledge of the effects of environmental conditions on the measurement or imperfect measurement of environmental conditions;

e) personal bias in reading analogue instruments;

f) finite instrument resolution or discrimination threshold;

g) inexact values of measurement standards and reference materials;

h) inexact values of constants and other parameters obtained from external sources and used in the data-reduction algorithm;

i) approximations and assumptions incorporated in the measurement method and procedure;

j) variations in repeated observations of the measurand under apparently identical conditions.

These sources are not necessarily independent, and some of sources a) to i) may contribute to source j).» [GUM]

Requirements

«The ideal method for evaluating and expressing the uncertainty of the result of a measurement should be:

- **universal**: the method should be applicable to all kinds of measurements and to all types of input data used in measurements.

The actual quantity used to express uncertainty should be:

- **internally consistent**: it should be directly derivable from the components that contribute to it, as well as independent of how these components are grouped and of the decomposition of the components into subcomponents;
- **transferable**: it should be possible to use directly the uncertainty evaluated for one result as a component in evaluating the uncertainty of another measurement in which the first result is used.»

[GUM]

Recommendation [BIPM 1980]

«The uncertainty in the result of a measurement generally consists of several components which may be grouped into two categories according to the way in which their numerical value is estimated:

A. those which are evaluated by statistical methods,

B. those which are evaluated by other means.

There is not always a simple correspondence between the classification into categories A or B and the previously used classification into “random” and “systematic” uncertainties. The term “systematic uncertainty” can be misleading and should be avoided.

Any detailed report of the uncertainty should consist of a complete list of the components, specifying for each the method used to obtain its numerical value.»

The distinction is about the evaluation method, not the source (or the “kind”) of uncertainty

Recommendation (2) [BIPM 1980]

«The components in category A are characterized by the estimated variances s_i^2 and the number of degrees of freedom v_i. Where appropriate, the covariances should be given.

The components in category B should be characterized by quantities u_j^2, which may be considered as approximations to the corresponding variances, the existence of which is assumed. The quantities u_j^2 may be treated like variances and the quantities u_j like standard deviations. Where appropriate, the covariances should be treated in a similar way.

The combined uncertainty should be characterized by the numerical value obtained by applying the usual method for the combination of variances. The combined uncertainty and its components should be expressed in the form of standard deviations.»

All components of uncertainty are represented as variances and combined as such

Type A – Type B, both standard deviations

«Uncertainty of measurement comprises, in general, many components.

Some of these components may be evaluated from the statistical distribution of the results of series of measurements and can be characterized by experimental standard deviations [**Type A**].

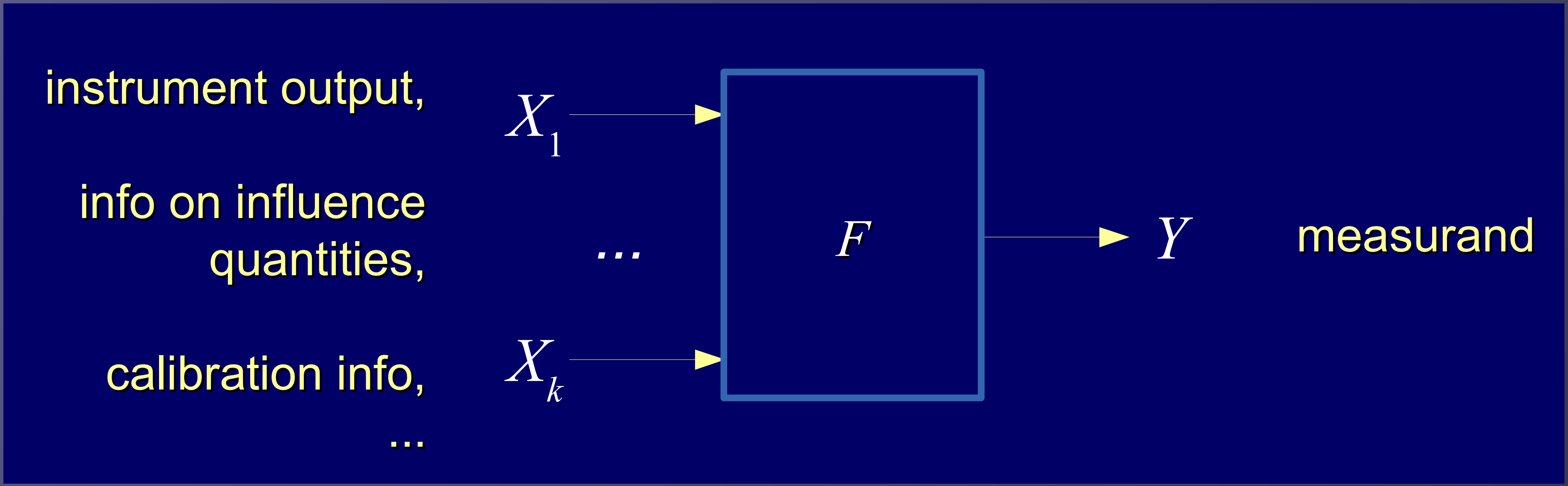
The other components, which also can be characterized by standard deviations, are evaluated from assumed probability distributions based on experience or other information [**Type B**].»

«Thus a Type A standard uncertainty is obtained from a probability density function derived from an **observed frequency distribution**, while a Type B standard uncertainty is obtained from an assumed probability density function based on the **degree of belief** that an event will occur.»

[GUM]

The mathematical model

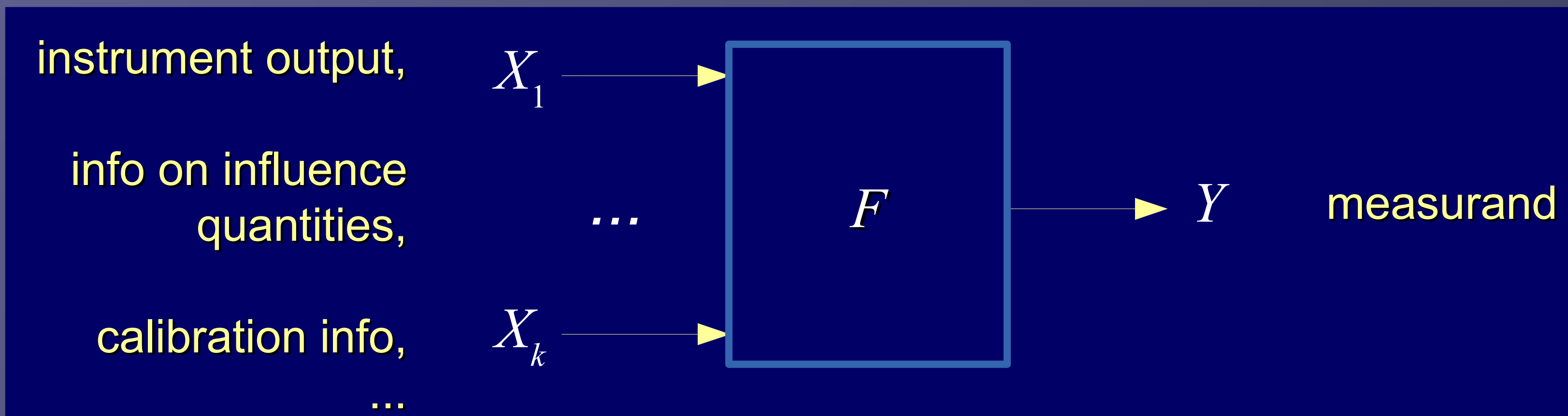
The measurement result is obtained by combining several components through a **measurement function**



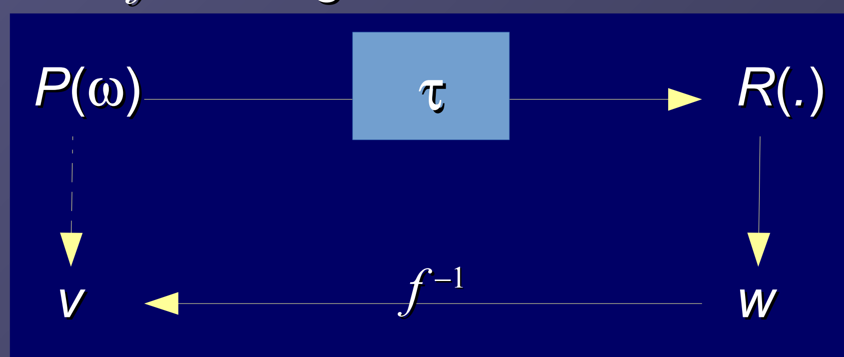
Any measurement is a model-based, inferential process

The mathematical model (2)

The model:



under the condition $F = f^{-1}$ is a generalization of:



Measurement functions are the generalized inverse of transduction / observation / calibration functions

Basic version [GUM]

- Each input component X_i is represented by a couple:

$$\langle x_i, u(x_i)\rangle = \langle \text{input quantity value, standard uncertainty}\rangle$$

- Analogously, the measurand Y is represented by a couple:

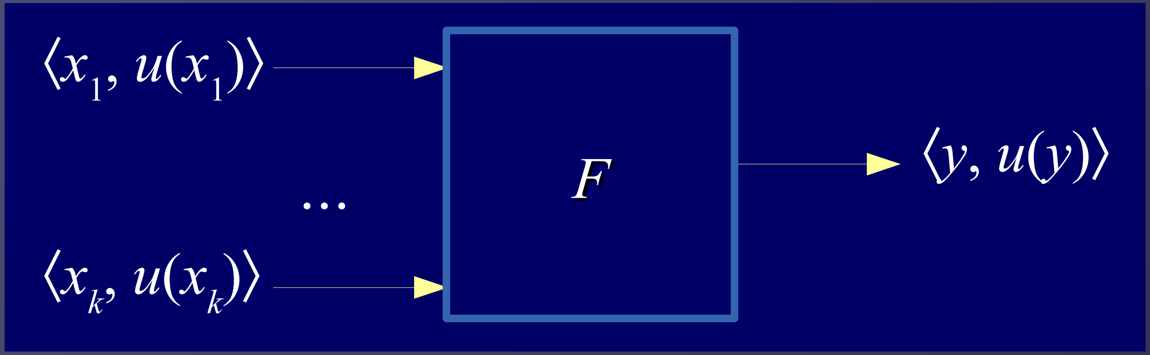
$\langle y, u(y)\rangle$ = ⟨measured value, standard uncertainty⟩ = measurement result

Under the hypothesis that F is analytically known:

- the measured value is computed as usual:

$$y = F(x_1, \ldots, x_k)$$

- uncertainties are "propagated" as weighed sum of variances



Uncertainty propagation [GUM]

Under the hypothesis that F is analytically known and is derivable:

1 input quantity

$Y = F(X)$

$$u(y) = \left|\frac{df}{dx}\right| u(x)$$

$K > 1$ input uncorrelated quantities

$Y = F(X_1, \ldots, X_K)$

$$u(y) = \sqrt{\sum_{i=1}^{K} \left[\frac{\partial f}{\partial x_i}\right]^2 u(x_i)^2}$$

$K > 1$ input correlated quantities

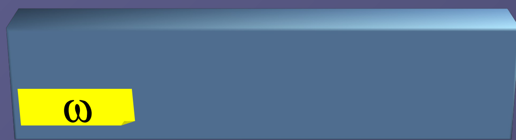
$Y = F(X_1, \ldots, X_K)$

$$u(y) = \sqrt{\sum_{i=1}^{K} \sum_{j=1}^{K} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j)}$$

(first-order approximation of Taylor series of F about $\langle x_1, \ldots, x_K \rangle$)

Uncertainty propagation: an example

The length l of



is measured at the current temperature t, the measurand being l_{ref} at $t_{\mathrm{ref}} = 20\ °\mathrm{C}$

Let us suppose that:

- temperature is the only influence quantity
- the thermal expansion is linear with given coefficient, α

  → $$l_{\mathrm{ref}} = \frac{l}{1 + \alpha\left(t - t_{\mathrm{ref}}\right)}$$

- measurements of length and temperature produce uncorrelated values

  → the simpler version of the propagation rule can be used

Uncertainty propagation: example (2)

If, e.g.:

- $\alpha = 1.2 \cdot 10^{-3}$ °C^{-1} [specifications (*)]
- $l = 0.4567$ m; $u(l) = 0.0002$ m [experiment]
- $t = 60.5$ °C; $u(t) = 0.4$ °C [experiment]

then,
given the measurement equation:

$$l_{\text{ref}} = \frac{l}{1+\alpha\left(t-t_{\text{ref}}\right)}$$

α	0.0012		
t_{ref}	20		
l; $u(l)$	0.4567	0.0002	
t; $u(t)$	60.5	0.4	

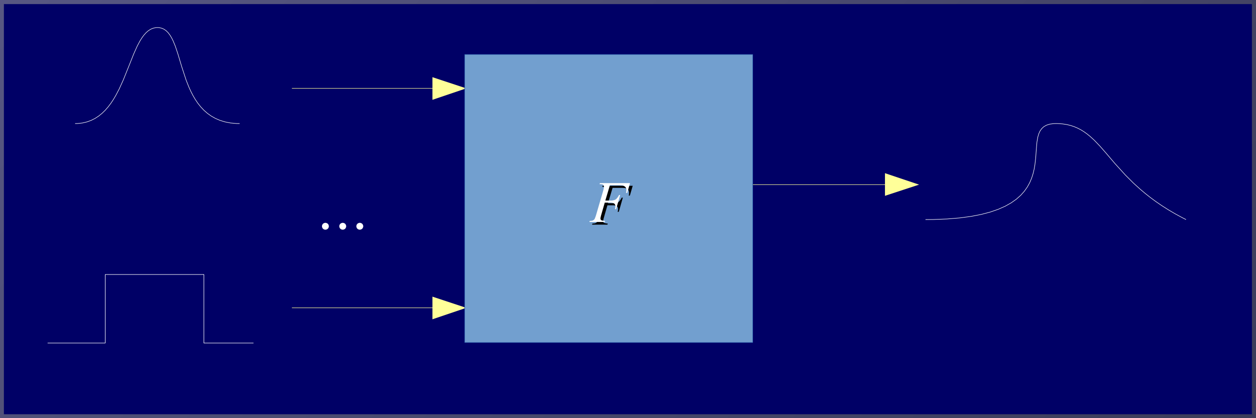
	$l/(1+\alpha(t-t_{\text{ref}}))$		
l_{ref}	0.4355		
	$1/(1+\alpha(t-t_{\text{ref}}))$		contribution of l
$dF\,/\,dl$	0.9537		0.0002
	$-\alpha l/(1+\alpha(t-t_{\text{ref}}))^2$		contribution of t
$dF\,/\,dt$	-0.0004984		0.0002
$u(l_{\text{ref}})$	0.0003		

(*) no uncertainty;
2 orders of magnitude greater on purpose

Again on the mathematical model

Nothing in this model implies that $X_i = \langle x_i, u(x_i) \rangle$

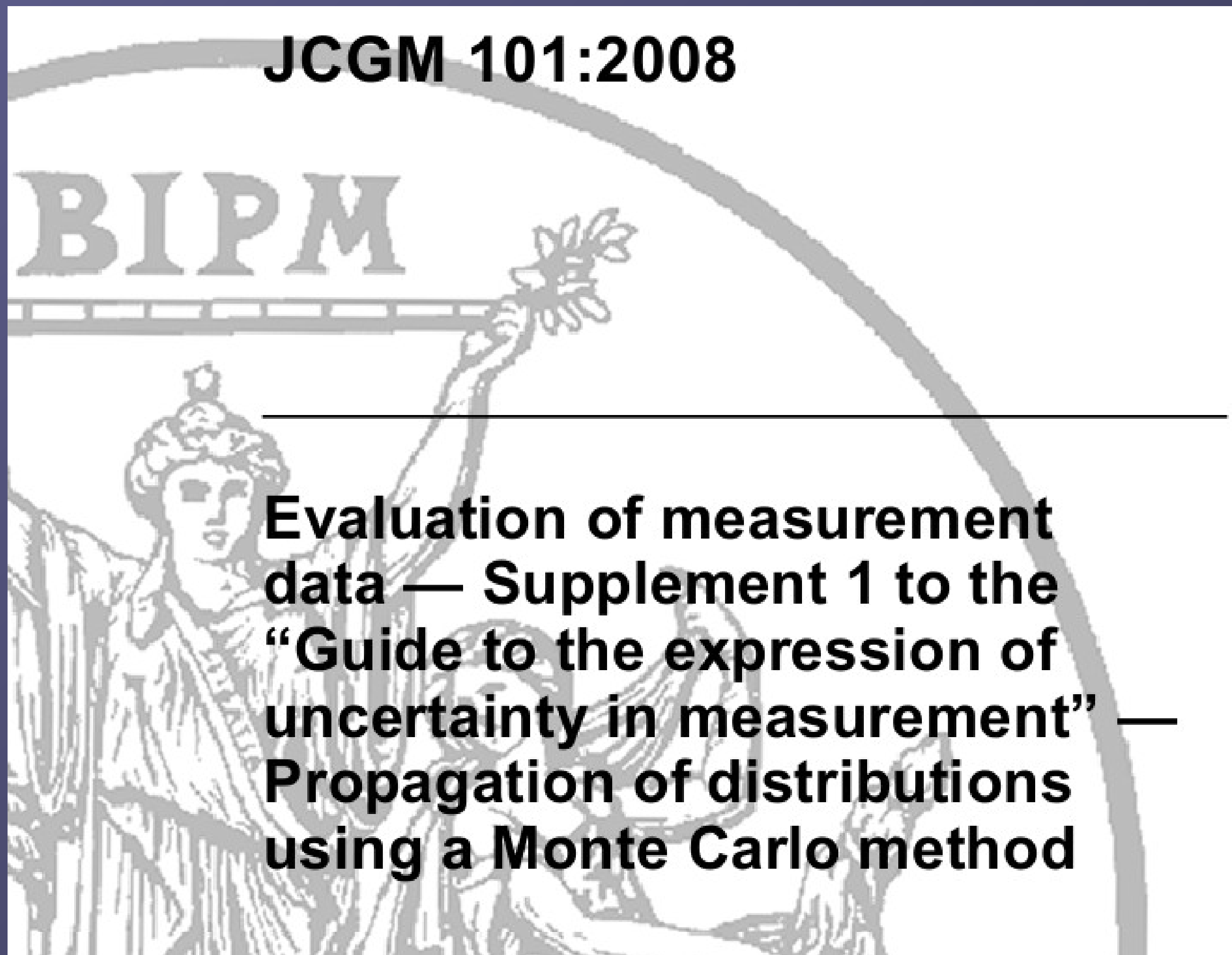
More generally, each quantity could be represented as a pdf



Propagation of uncertainties becomes **propagation of distributions**

But the problem is generally not analytically solvable...

Fifteen years later...



JCGM 101:2008

BIPM

Evaluation of measurement data — Supplement 1 to the “Guide to the expression of uncertainty in measurement” — Propagation of distributions using a Monte Carlo method

Epistemological note

A previous consideration:

truth is primarily asserted in reference to a model, not to "reality"

i.e., from:

property value $\xrightarrow{\text{is the (possibly true) value of}}$ property

to:

property value $\xrightarrow{\text{is referred to}}$ property

property value $\xrightarrow{\text{is the (possibly true) value of}}$ modeled property $\xrightarrow{\text{is the (more or less certain) model of}}$ property

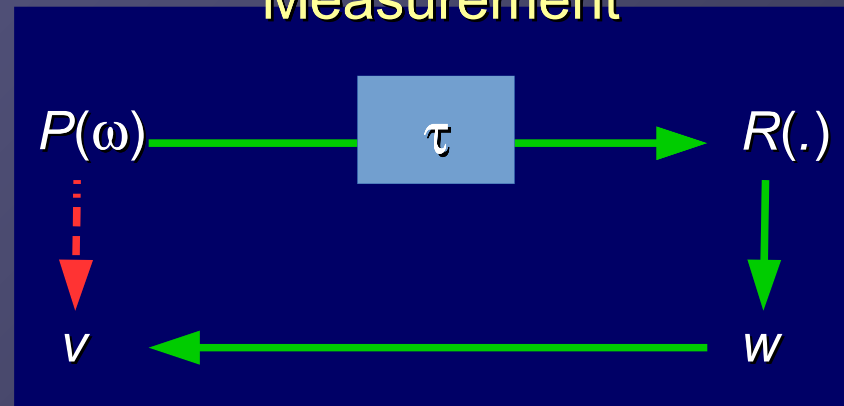
Does it mean that there is no more room for true values (and error) in measurement?

Epistemological note (2)

Let us recall:

Calibration

Measurement



where: unknown, to be found

assumed to be known

$P(s_i) \longrightarrow v_i$ known by calibration: truth is in principle available

$R(.) \longrightarrow w_i$ “operatively true” value

Pragmatic note

target measurement uncertainty: «measurement uncertainty specified as an upper limit and decided on the basis of the intended use of measurement results» [VIM3]

Measurement uncertainty as tool for decision making

specifications (tolerance)

worst case, acceptable situation

target uncertainty

measurement uncertainty

from the information on the experiment

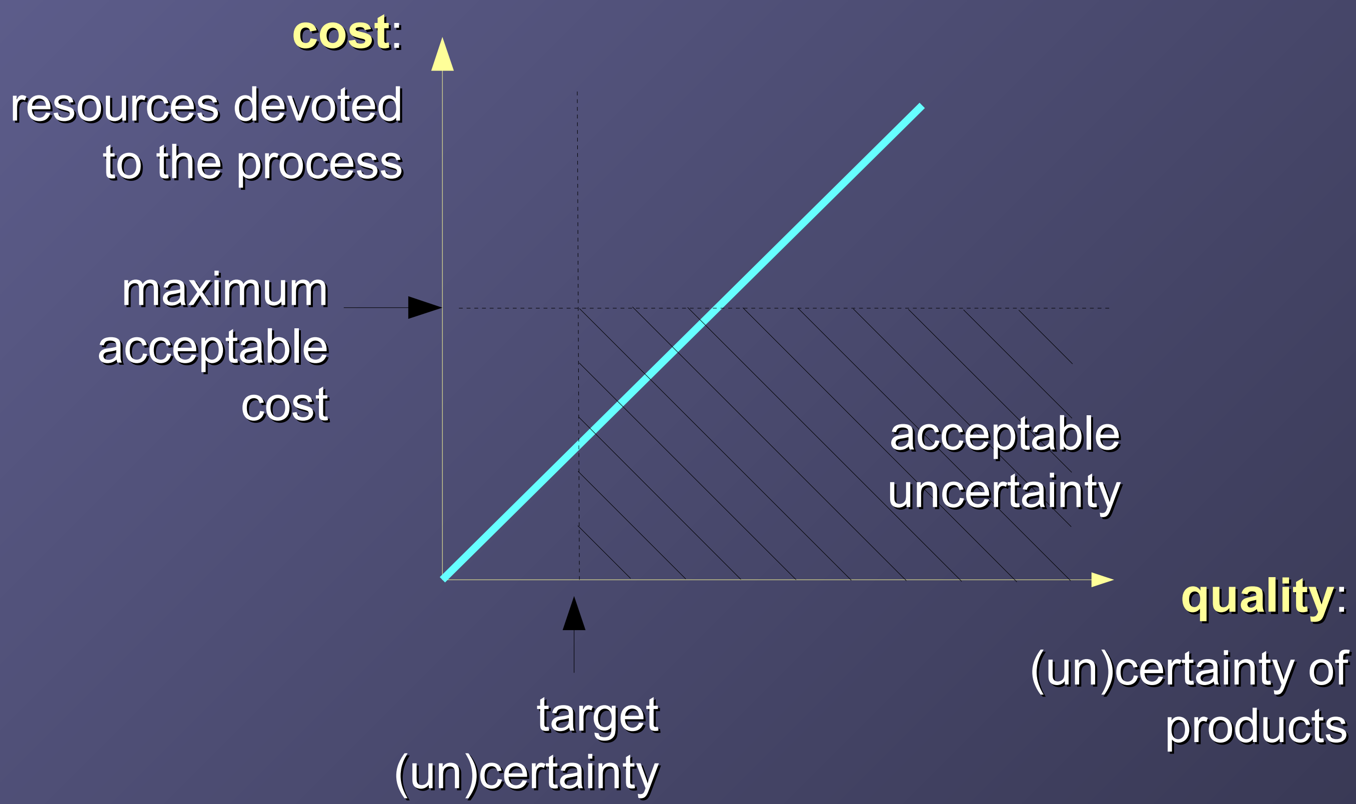
definitional uncertainty

from the model of the measurand

«Definitional uncertainty is the practical minimum measurement uncertainty achievable in any measurement of a given measurand.» [VIM3]

... in conformity to the so-called “GIGO principle”...

Measurement as production process



measurement uncertainty = target uncertainty: pragmatically the best case

Measurement vs. forecast?

There are «three distinct ways in which a forecast can be good:

in the type 1 sense if it corresponds to the forecaster's best judgment derived from her knowledge base;

in the type 2 sense if the forecast conditions correspond closely to the observed conditions at (or during) the valid time of the forecast;

in the type 3 sense if the forecast, when employed by one or more users as an input into their decision-making processes, results in incremental economic and/or other benefits.»

[A.H. Murphy, What is a good forecast?
An essay on the nature of goodness in weather forecasting, 1993]

Measurement in the bigger picture?

TABLE 1. Names and short definitions of three types of goodness.

Type	Name	Definition
1	Consistency	Correspondence between forecasts and judgments
2	Quality	Correspondence between forecasts and observations
3	Value	Incremental benefits of forecasts to users

[A.H. Murphy, What is a good forecast?
An essay on the nature of goodness in weather forecasting, 1993]

**Do analogous types
apply also to measurement?**

More than math...

«Although this Guide provides a framework for assessing uncertainty, it cannot substitute for critical thinking, intellectual honesty and professional skill.

The evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of the measurement.

The quality and utility of the uncertainty quoted for the result of a measurement therefore ultimately depend on the understanding, critical analysis, and integrity of those who contribute to the assignment of its value.»

[GUM]

A comment of the proposed definition of ‘measurement’

Measurement is a both conceptual and experimental process
implementing a v-assignment
able to produce information on a predefined property
with a specified and provable level
of objectivity and intersubjectivity

The provability of the specified level
of objectivity and intersubjectivity
is mainly an open issue

THANK YOU
FOR YOUR KIND ATTENTION

Luca Mari

lmari@liuc.it