Accuracy, Precision, Uncertainty, Error, and Confusion

Alban Kellerbauer, CERN

Guide to the expression of uncertainty in measurement (GUM)

Published by ISO / 1993, 1995

Edited by

BIPM	Bureau International des Poids et Mesures
IEC	International Electrotechnical Commission
IFCC	International Federation of Clinical Chemistry
ISO	International Organization for Standardization
IUPAC	International Union of Pure and Applied Chemistry
IUPAP	International Union of Pure and Applied Physics
OIML	Organisation Internationale de Métrologie Légale

Summary: http://physics.nist.gov/Pubs/guidelines/contents.html

Introduction

- Many of the terms used in expressing experimental results mean different things to different people:

  Accuracy & Precision

  Error & Uncertainty

  Statistical & Systematic

  Repeatability & Reproducibility

- This is mainly a problem of terminology (but not only)

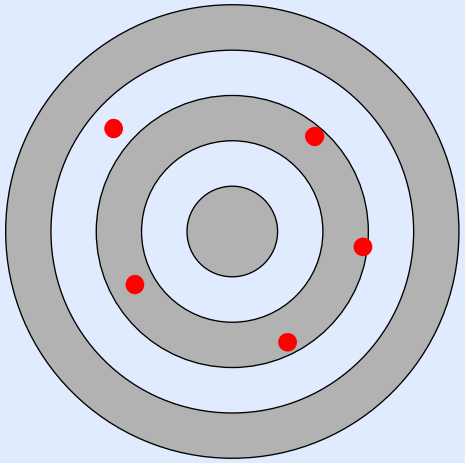
Accuracy vs. precision

- Traditional point of view:

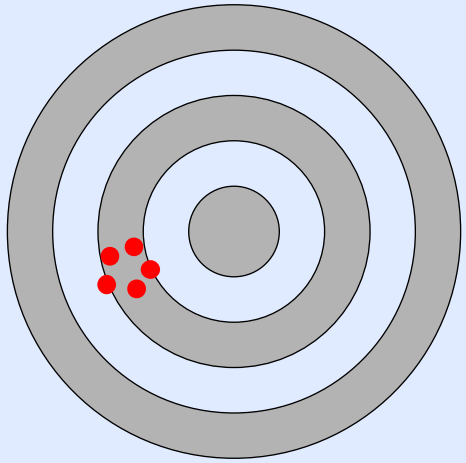
  Accuracy: "Conformity of a measured quantity to its actual (true) value"

  Precision: "Degree to which further measurements will show the same or similar results"

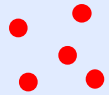
- Target analogy:



accurate, but not precise



precise, but not accurate



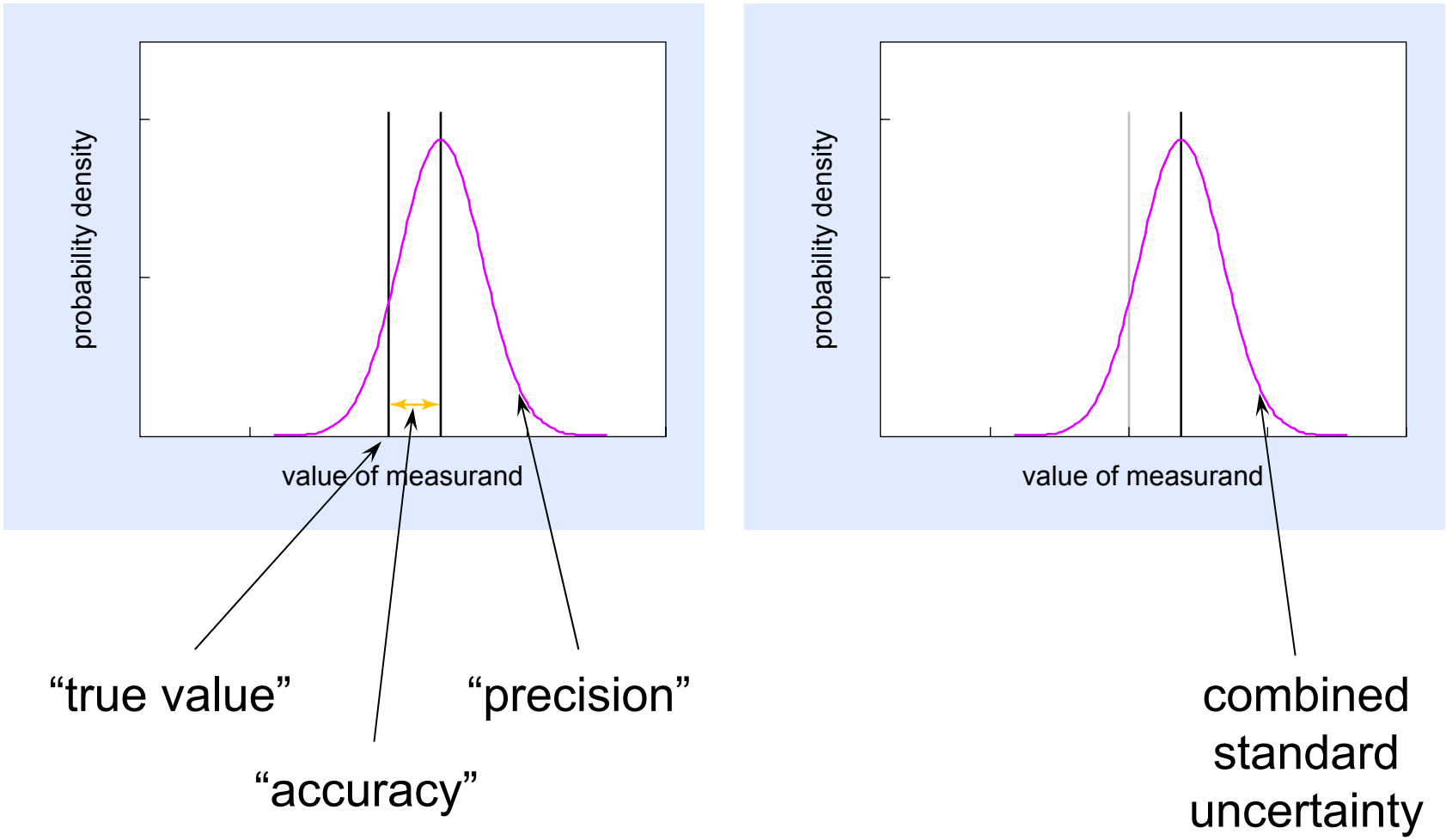
precise and/or accurate ?

In an actual measurement, the “true” value is unknown

Exceptions:

- Calibrations
- In hindsight, as compared with other (better) measurements

- From traditional to GUM terminology:



- Combined uncertainty should contain best estimate of all possible systematic effects

Classification of effects and uncertainties

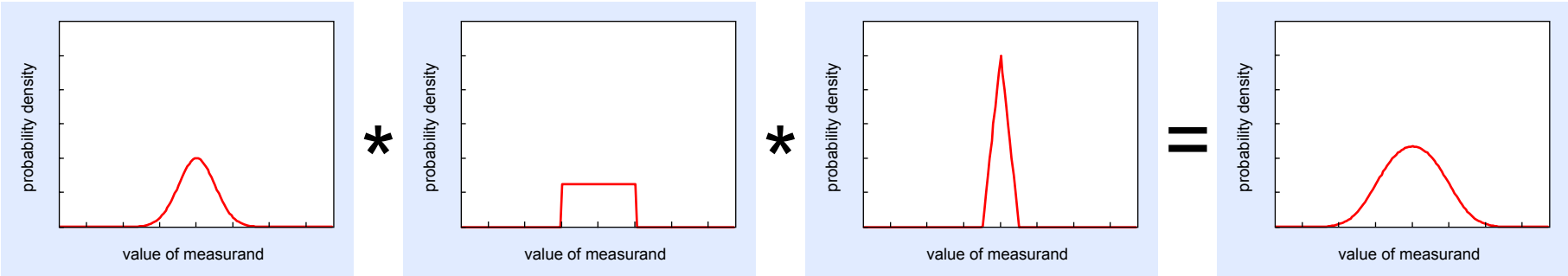
Random effects	• Unpredictable variations of influence quantities • Lead to variations in repeated measurements • Expectation value: 0 • Can be reduced by performing many measurements
Systematic effects	• Recognized variations of influence quantities • Lead to bias in repeated measurements • Expectation value: $\varepsilon(y)$ (*unknown!*) • Can be reduced by applying a correction, which carries an uncertainty

- Distinction between random and syst. effects is blurred
- Most systematic effects behave randomly because the relevant influence quantities are not well controlled
- GUM classifies uncertainties according to practical considerations:

Random effects	Systematic effects
↓	↓
Type A uncertainty	**Type B uncertainty**
Evaluated by statistical analysis of a series of observations	Evaluated by any other method • experience with behavior and properties of instruments • manufacturer's information • calibration

Treatment of uncertainties

- Systematic effects often have several sources
- The convolution of many different probability distributions is generally very close to a Gaussian:



$\Rightarrow$ Most systematic uncertainty contributions follow a normal distribution

- In the GUM, all uncertainties are treated as standard deviations (standard uncertainties)
- Prescriptions for the conversion of *a priori* distributions exist

- Advantages of this approach:
  1. Intermediate results can easily be incorporated into subsequent calculations via the law of propagation of uncertainty:

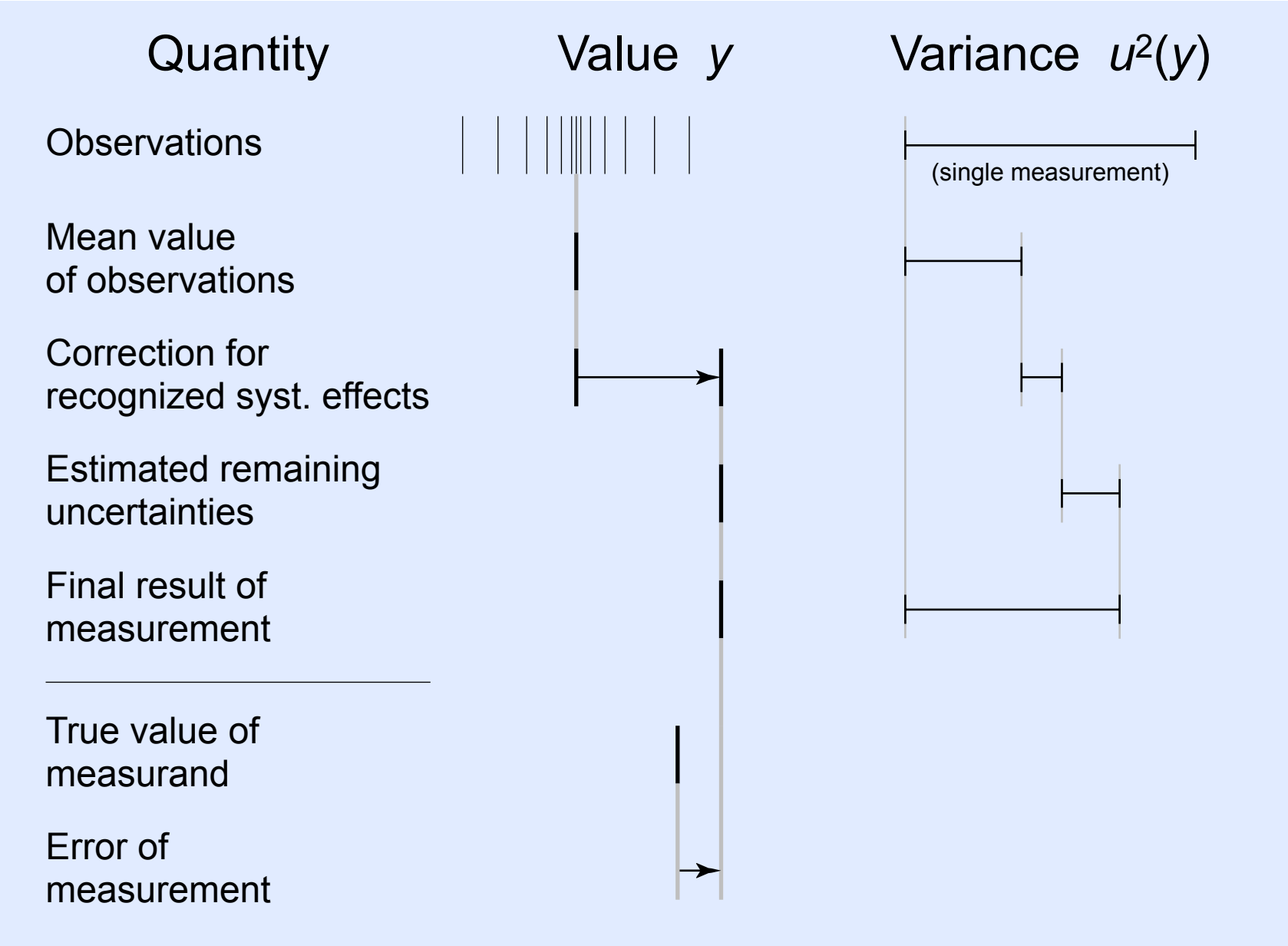
$$u^2(y) = \sum_i \left[\frac{\partial y}{\partial w_i} u(w_i) \right]^2$$

  2. All uncertainties can simply be added quadratically:

$$u^2(y) = \sum_i u_i^2(y)$$

  3. Confidence intervals can easily be calculated

- Graphical illustration:

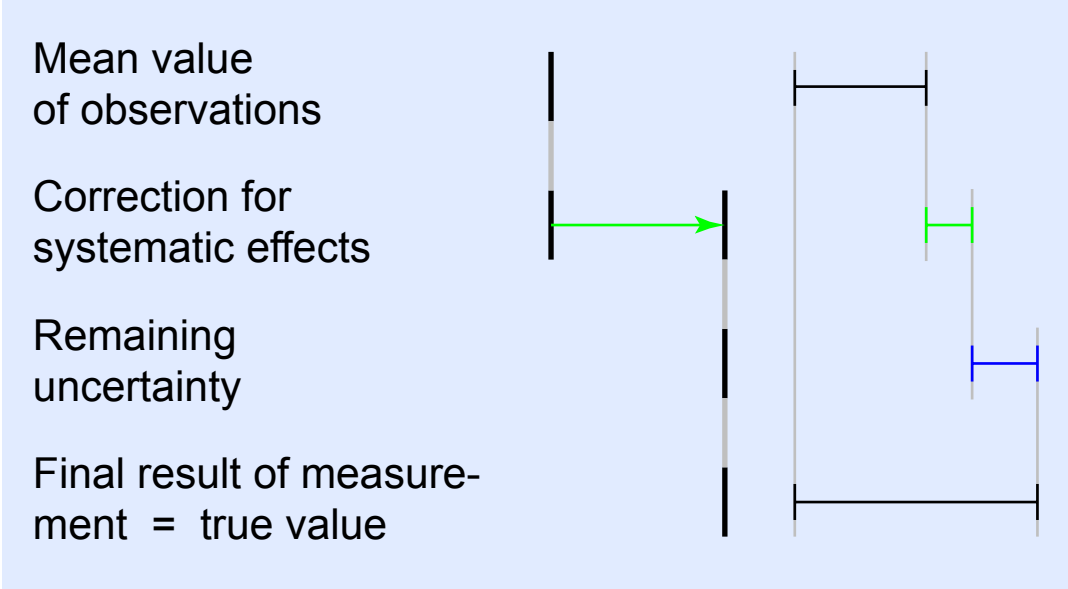


Calibration

- A calibration is a special type of measurement, in which the “true value” of the measurand is known
- This allows to determine

1. Correction for systematic effects
2. Remaining uncertainty

... and use them in actual measurements



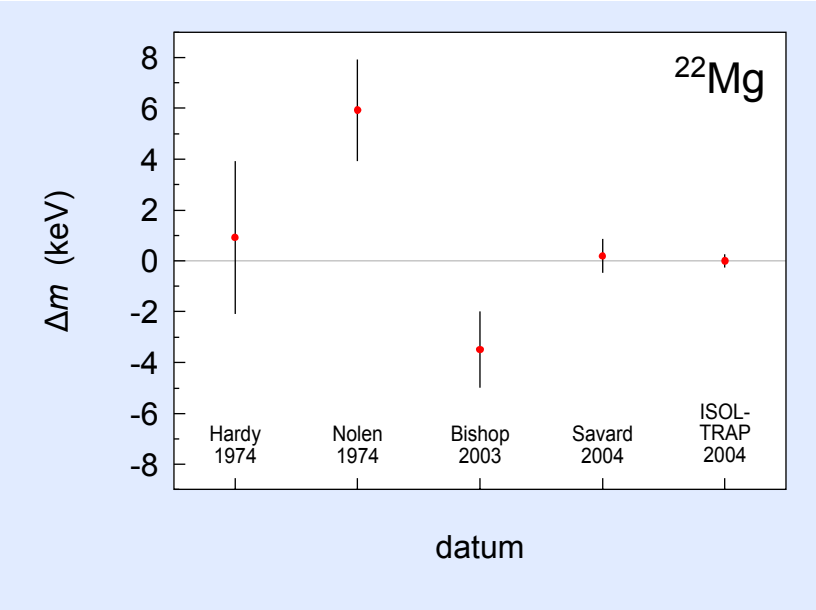
- Caution: Experimental conditions change over time

You could not step twice into the same river
For other waters are ever flowing on to you

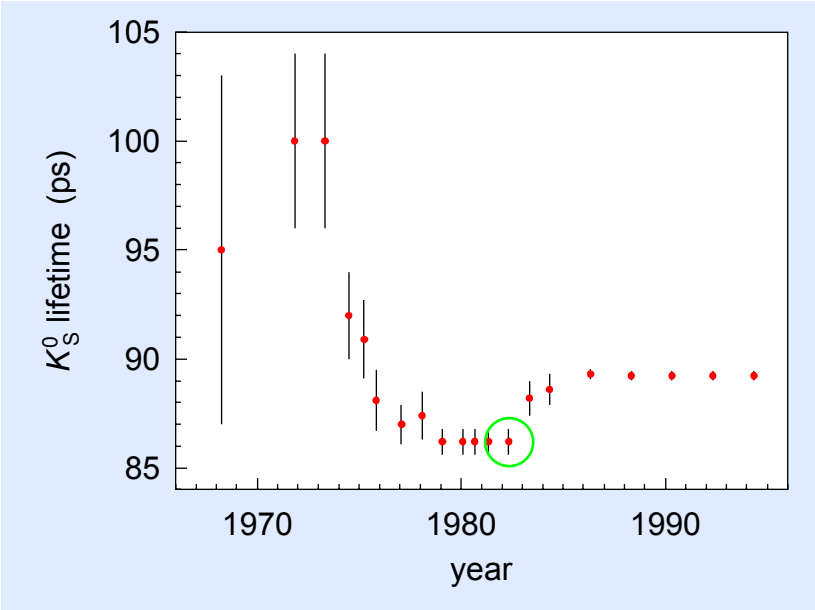
Heraclitus, 535–475 BC

Towards a “true value”

- Increasingly precise measurements approach the true value of a measurand:



[M. Mukherjee *et al.*, PRL **93** (2004) 150801]



[Review of Particle Physics 2004]

- This allows qualitative statements about the accuracy of prior (less precise) measurements
- But beware of surprises!

Summary on terminology

Quantity	Symbol	Note
(Standard) uncertainty	$u(y)$	well-defined quantity
Precision	—	$\approx 1/u(y)$ ambiguous
Error	$\varepsilon(y)$	unknowable, ambiguous
Accuracy [1]	—	$\approx 1/\varepsilon(y)$ unknowable, ambiguous

[1] Except as a statement in hindsight, *i.e.*, in comparisons or evaluations

Conclusions

- The main innovations introduced by GUM are:
  1. A reform of the terminology of expressing uncertainty, *e.g.*, by deprecating the terms *accuracy* and *error*
  2. A simplification of the classification and treatment of *uncertainty* that makes life easier
- This means relatively minor changes for most experimentalists
- Adherence to these rules is not compulsory, but has advantages for both authors and readers