# PART I

# FUNDAMENTALS

# CHAPTER 1

# MEASUREMENT MODELS AND UNCERTAINTY

ALESSANDRO FERRERO[1] and DARIO PETRI[2]
[1]Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milano, Italy
[2]Dipartimento di Ingegneria Industriale, Università degli Studi di Trento, Povo, Italy

## 1.1 INTRODUCTION

Nowadays, we are surrounded by measuring instruments and we use them several times a day, very often unconsciously and unaware of their complexity and accuracy. To realize how instruments have become a big part of our life, just think of how many times we read the speed indicator when we drive our car, the fuel indicator to know when we have to refill the tank, or, when we eventually refill it, the meter on the fuel pump.

Interestingly enough, we usually don't pay too much attention to the accuracy of the instruments we use, even if we rely on their indication to make important decisions, such as, for instance, driving safely or paying the right amount of money for the quantity of good we purchase. Even more strangely, the only instrument we generally adjust to a reference is our watch, which is probably the most accurate instrument we use in our everyday life: even the cheapest ones don't lag or lead for more than one second a day, which means that their relative accuracy is in the range of $1 \cdot 10^{-5}$!

The above examples give us clear evidence that we do use and read instruments, but they still leave an important question open: are we also making a measurement? Trying to answer this question opens also another fundamental question: which is the difference between reading an instrument and making a measurement?

This chapter is aimed at providing an answer to this question.

*Modern Measurements: Fundamentals and Applications*, First Edition.
Edited by Alessandro Ferrero, Dario Petri, Paolo Carbone, and Marcantonio Catelani.

## 1.2 MEASUREMENT AND METROLOGY

To understand what measuring means, let's start from the definition of measurement, taken from the International Vocabulary of Metrology (VIM) [1].

> *Measurement*
>
> Process of experimentally obtaining one or more *quantity values* that can reasonably be attributed to a *quantity*.

So, a measurement process provides, as a part of the *measurement result*, one or more quantity values that can be attributed to a quantity intended to be measured, that is also called, always according to the VIM [1], *measurand*.

To fully understand this definition, we have to refer to the definition of quantity. We can find it again in the VIM.

> *Quantity*
>
> Property of a phenomenon, body, or substance, where the property has a magnitude that can be expressed as a number and a reference.

The VIM states that a reference can be a *measurement unit*, a *measurement procedure*, a *reference material*, or a combination of such.

When physical properties are considered, the reference is generally a measurement unit, whilst, when chemical measurement are considered, the reference is quite often a reference material.

The quantity values provided by the measurement are therefore a number and a reference together expressing the magnitude of a quantity [1].

Is this the measurement result? Or, better, can a measurement result be expressed only by a number and a reference? As we will see later in section 1.5 of this chapter, a measurement procedure cannot provide the *"true"* value of a measurand, due to a number of factors that we will thoroughly discuss later. This means that a measurement result can only provide a *finite amount of information* about the measurand, and we must know if that amount is enough for the intended use of the measurement result. Otherwise, the measurement result would be meaningless.

Therefore, any measurement result has to be provided with an attribute capable of quantifying how close to the measurand's value the obtained quantity value is. This attribute is called *uncertainty*, and the correct definition of measurement result, as provided by the VIM, is as follows.

> *Measurement result*
>
> Set of *quantity values* being attributed to a *measurand* together with any other available relevant information.

In a note to this general definition, the VIM states that:

> A measurement result is generally expressed as a single *measured quantity value* and a *measurement uncertainty*.

The above general definitions have introduced a number of concepts (*quantity value*, *reference*, *relevant information*, *uncertainty*), that will be covered in the next Sections, and show that a measurement is a definitely more complex procedure than simply reading an instrument.

The science that includes all *theoretical* and *practical* aspects of measurement, regardless to the measurement uncertainty and *field of application*, is called *metrology* [1]. Its definition, as provided by the VIM, is as follows.

> *Metrology*
>
> Science of measurement and its application

As many other sciences that have a deep impact on human life, metrology finds its roots back into the ancient times, and the evolution of the human needs had a significant impact on its development, and its present formulation. Therefore, before analyzing the fundamental concepts of nowadays metrology, let us find an answer to two basic questions: why do we measure? And for what do we use our measurement results?

A quick glance into the past will help us to find the answers to these questions.

## 1.3 MEASUREMENT ALONG THE CENTURIES

At the beginning of the human adventure, the measurement concept did not exist and the experimental activity was confined to direct observations. Instruments did not exist, and our senses were the only available tools to observe and somehow quantify the reality.

The first physical quantity that was barely measurable with our senses and probably started metrology was time. We can only imagine that this interest was related to the impact that the night and day cycle, the lunar period and the rotation of seasons have on human life and the human ability of predict them to optimize all activities aimed at providing food, from hunting to agriculture and cattle breeding.

As a matter of fact, the archeological findings from prehistory show with little doubt that the first conceived measuring instruments were aimed at the measurement of time, through the observation of the stars or the sun. Although the most famous among these instruments is the Stonhenge Circle, its "operating principle" is not yet clear. On the other hand, many remains have been found from a later period, across prehistory and history, that proved how time was measured through the observation of the displacement of the shadow of a vertical device (the *gnomon*) created by the sun on the ground [2].

However, at that time, this activity was more of a religious and prophetic kind, rather than aimed at increasing knowledge. We had to wait until the organized social structures appeared to meet awareness of knowledge and a systematic discussion about its meaning.

### 1.3.1 Measurement in Ancient Greece

There is no doubt that ancient Greece was the cradle of philosophy, and philosophy was considered the most important means of knowledge. We have also evidence, from documents and archeological findings, that instruments were used, mainly for length and capacity measurements. However, the great philosophers of that time did not consider instruments as tools to advance knowledge.

The motivation can be found in the dichotomy between the world of philosophical abstractions and the empirical world of observations, clearly expressed by Socrates in the fifth century BC.

In Socrates' philosophy, we can build our knowledge by defining abstract models that allow us to explain, through suitable logical steps, what happens to us. In this way, we can build our own world of philosophical abstractions.

On the other hand, there is a physical world around us, that shows up through a number of occurrences, facts and events, and that, at Socrates' time, we could perceive only through our senses. This is the so-called empirical world and, according to Socrates, the abstractions were aimed at explaining the true essence of the empirical world. In his philosophy, abstract logical constructions were the only way to knowledge, and observations had the simple role of triggering speculation. The use of observation results as a way to validate abstract models was not part of Socrates' approach, and was not part of the scientific method at least until Galileo's time. The subordinate role assigned to observations with respect to logics has lasted for so many centuries that even now the word "empiricism" implies a negative connotation.

### 1.3.2 Measurement in the Roman Empire

We have to wait until the Roman Empire to discover, from the many available historical documents and texts, that they had a rather modern approach to measurement. Indeed, we know that they had standards (at least for length, mass, and capacity), these standards were approved, kept and maintained by the central government, and secondary standards were disseminated in the provinces of the empire.

Evidence of the importance assigned by Romans to measurement is given by the fact that the primary standards were kept in the temple of Iuno Moneta.[1] Only valuable and important objects were kept in the house of a God. And only money-related objects were kept by the God of Money!

According to these facts, we expect to find several traces, in the Latin literature and in the archeological sites, of the use of their instruments and standards. For instance,

[1] Moneta, in Latin, means money. We cannot probably find better evidence that the economical importance of measurement was quite clear to Romans.

we know that the length standard was the Roman *pes* (Roman foot), and we would expect to find clear evidence of the use of the pes in the remains of imperial Rome and evidence in the dimensions of the many perfectly preserved buildings we can find in the regions of the world that were dominated by the Romans. The big surprise is that this is not true. So, when and for what was the unified length unit used?

The answer to this question comes from the wonderful archeological site of Pompei. Pompei was an important town of the Empire, located South of Rome, not far from the present Naples, close to the Vesuvius volcano. It was destroyed by an eruption in 79 AD. Actually, the explosive eruption killed every form of life in Pompei, but preserved everything under a layer of ash and light pumice stones. When it erupted, it hurled 100,000 tons/s of superheated stones and ash into the stratosphere, and formed a cloud with the classic shape of an umbrella pine tree. The first thing that hit the ground was a hot wind that was estimated to hit Pompei at about 350 km/h and at a temperature of more than 450°C. This wave of hot air killed and dried every form of life in seconds. Then ash came down, and light pumice stones, that buried everything without causing any major damage.

For this reason, many centuries later, archeologists found Pompei almost intact, and have learned so many things about everyday life at that time, and found remains that solve our questions about the use of the unified length standard [3]. Despite the dimensions of the buildings cannot be related to the pes, its use was clear by analyzing some artifacts.

In the Roman houses, water was stored in reservoirs located under the ground floor. All rooms had small wells for access to water, and the opening was protected by small marble curbs. The diameter of the curbs can be traced to the pes quite perfectly. Similarly, the dimensions and the distance between the holes in the iron lockers can be traced to pes submultiples. What is the reason?

Marble artifacts were not produced locally; they came from a distant region. Iron locks were produced by a different craftsman than the one who made doors or coffers on which the locks were installed. The marble devices and the locks had to fit accurately with another part. Since they were not produced locally, their dimensions could not be measured with the same instrument, and therefore they had to be measured according to a unique standard. The Roman Empire was the first global market, and they needed to ensure that parts crafted in different regions of the Empire could be used in all other regions without creating major problems. A need that looks quite similar to our present industrial needs was solved, in principle, in a similar way as nowadays.

The roots of the modern industrial metrology can be found in the use of standards in the Roman Empire. It was not yet a science, but only an application-driven technology. It was not yet a way to knowledge, but it was the first step toward the development of a new science.

### 1.3.3 Measurement in the Renaissance Period

At the fall of the Roman Empire, the economy receded to a local, small scale, and therefore there was no need for universal standards. Many different systems of units

flourished at that time, often according to the convenience of the local squire and the consequences of such a scientific disaster are still suffered in the present days.

Things changed dramatically during the Renaissance, when the modern concept of science began to be elaborated. The experimental activity had a big part in the development of this new concept, so that we can now state that the meaning we assign to the measurement activity today finds its true origin at Renaissance time.

Two scientists, above all, contributed to assert this new concept: Galileo Galilei and Evangelista Torricelli.

Galileo used the new telescope to peer into the sky and proved that we can use an instrument to extend our senses and discover things (planets and satellites, in his case) that are far beyond the range of our "human sensors".

Torricelli invented the barometer, and proved that the atmosphere has a weight, that is something that our senses cannot, once again, show us.

Galileo's observation of the fall of bodies and his observation of the motion of planets proved the flaws of the geocentrical theory and paved the way to the formulation of the general law of gravitation.

Socrates' paradigm received a first, important shock. Abstractions alone cannot build our base of knowledge and observations are not only the starting point of logical constructions. New observations and new experiments are needed to validate theories obtained as the result of a logical construction. Moreover, we cannot totally rely on our senses to observe the physical world: our senses are limited, and keep important phenomena hidden to us, thus preventing us from fully understand them or sometimes leading us to wrong conclusions. The modern science moved its first steps thanks to new instruments and new measurement procedures.

### 1.3.4 Measurement in the Modern Age

A couple of centuries later, another storm upset the traditional scientific belief. During the Enlightenment period, science was trusted as the only gate to knowledge, and the experimental activity was considered as the most important tool for investigating physical phenomena and understanding them.

In this new framework, in which science was, for the first time, considered as a global language for describing physical phenomena, a new issue arose. If science has to be a global language, it must also have a global, unique reference. Therefore, the measurement units must be universally recognized and invariant.

This is a twofold problem. On one side the system of units must be unique, that is universally accepted and adopted. As we saw in Section 1.3.2, this problem was already dealt with at the time of the Roman empire, where conventional standards were defined to solve practical and economical problems. However, the common belief was that scientific problems cannot be solved on the base of conventions. This is the second part of the problem: the unique system of units must also be derived by physical properties of phenomena that can be considered as invariant in space and time.

This was indeed a revolution both in the scientific and everyday habits of that time. Even though the scientific community reached a wide agreement on this point, putting it into practice required another big revolution: the French Revolution. Indeed, these

ideas gained enough political consensus to have the local measurement units banned and replaced by new ones, that were defined as global as possible, according to the scientific knowledge of the eighteenth century.

It is worth reminding that a huge effort was put in place to measure the length of the Earth's meridian, on which the meter definition could be based.

It was another big step in the long process that turned observations into measurements. However, standards defined on properties of our planet, such as the unit of length and the unit of time, were not global and invariant enough. Moreover, a new science was moving its first steps: electromagnetism. It was soon clear that the result of the interactions of electromagnetic quantities was a force. New measurement units had to be considered to serve as a reference for the measurement of electromagnetic quantities, and they had to be coordinated to the already existing ones, so that the physical relationships between the different electromagnetic and mechanical quantities could be reflected into the same relationship between units.

Both problems were clear to James C. Maxwell, the father of electromagnetism, who, in 1871, wrote:

> If we wish to obtain standards of length, time and mass which shall be absolutely permanent, we must seek them not in the dimensions, or the motion, or the mass of our planet, but in the wavelength, the period of vibration, and the absolute mass of these imperishable and unalterable and perfectly similar molecules.

This is the conceptual basis of our present International System of Units, or SI, as it will be shown in the next chapter. It took almost another century, and the work of another great electrical engineer, Giovanni Giorgi, to see these outstanding ideas put into practice.

### 1.3.5 Measurement Today

The long journey that changed the simple activity of observing the empirical world into the measurement science is now accomplished. The experimental activity, of which measurements are the most important part, is now considered an important, essential tool for the advancement of knowledge. Today, any scientific or technical proposal needs to be supported by suitable measurement results.

We have now fully understood what Lord Kelvin stated in his famous speech of May 6, 1886:

> I often say that when you can measure what you are speaking about, and can express it in numbers, you know something about it; but when you cannot express it in numbers your knowledge about it is of meager and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be. So, therefore, if science is measurement, then without metrology there can be no science.

Nowadays, measurement results are also goods that are sold and bought. In a modern, developed country, the economical value of measurements has been estimated to be

the 6% of the Gross Domestic Product. Electrical measurements in industry represent the 4% of sales, and biomedical analysis represent the 5% of the total cost of the health system.

Measurements are essential to determine the quantity of goods in economical transactions, to evaluate cost of goods and services, to control quality of goods and, in general, every other important issue involved in transactions. This application field of measurements is generally known as *legal metrology*.

When technical issues are involved, measurements are employed, for instance, in inspection tests—on materials, components or on the final products—in comparing product characteristics or performances and, in general, in every step of an industrial process. This application field of measurements is generally known as *industrial metrology*.

Last, but not least, measurements are the key point in the experimental validation of theories and in studying and implementing new and more accurate measurement methods and standards. This application field of measurements is generally known as *scientific metrology.*[2]

## 1.4 MEASUREMENT MODEL

The concepts that we have briefly recalled in the previous sections suggest that measurement is a complex activity, far more complex than connecting and reading an instrument. Developing a model for the measurement activity is therefore necessary, to fully understand its meaning and which are the requirements to perform a measurement correctly.

Socrates' model of knowledge, recalled in section 1.3.1, is a good starting point to build the measurement model. According to this approach, our knowledge is built around abstract models that explain phenomena through suitable logical steps. This approach is still valid today: we have models to explain the different phenomena (not only the physical ones), and these models are based on suitable mathematical equations, or logical inferences, in which we assign an abstract symbol to each considered quantity. In this respect, models represent the symbolical world.

On the other hand, as already noted by Socrates, we become aware of the different phenomena because they reveal themselves through occurrences, facts and events that define the empirical world.

The main questions are: how can we validate our abstract model? How can we ensure that it represents the phenomena it models correctly? If we consider that a model is supposed to provide the values taken by the output symbol (or symbols) for a given set of values assigned to the input quantities, we can conclude that a way to validate the model is to observe the phenomenon it represents, measure the quantities

[2]This definition might appear somehow tautological, since metrolgy means "science of measurement," and therefore a science is scientific by itself. However, this definition is generally used whenever we want to refer to the applications of metrology in support of other branches of science.

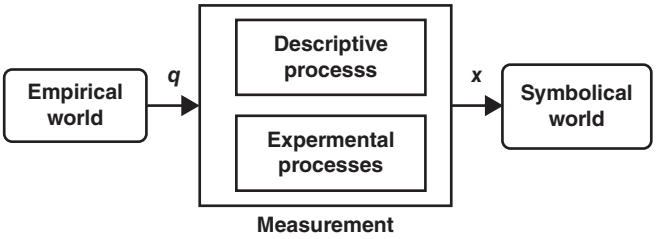


**FIGURE 1.1** Model representing measurement as a bridge between the empirical and symbolical worlds.

that are supposed to affect this phenomenon, measure the quantity that represent the phenomenon itself, and verify whether the measurement result, for this quantity, corresponds to the value provided by the model.

For instance, if we want to validate the Ohm's law, we measure the resistance $R$ of a resistor, we supply it with different voltage values $V_{\mathrm{k}}$ and check whether the corresponding measured values of current $I_{\mathrm{k}}$ flowing through the resistor are such that: $V_{\mathrm{k}} = R \cdot I_{\mathrm{k}}$, for all considered $k$ measurements.

Measurement can be seen, therefore, as the bridge between the empirical world and the symbolical world. The measurement role is that of adding quantitative information to symbols.

### 1.4.1 A First Measurement Model

Defining measurement as a bridge between the empirical and symbolical worlds is a starting point to build a useful measurement model. So let us try to refine this idea, using the graphical representation shown in Figure 1.1.

According to this picture, the measurement process maps a single *manifestation* $q$ of a *property* (for instance weight, length, …) of an object (that is a phenomenon, a body, or a substance) belonging to the empirical world onto a symbol ($x$ in Figure 1.1) that belongs to the symbolical world. In fact, measurement is an objective empirical operation that assigns a symbol $x$ (called the *measured value*) to the manifestation $q$ (called the *measurand*) of the considered property of an object [4, 8, 9]. If we consider again the above example where the current $I$ flowing through a resistor was measured, the considered property of the resistor is the current and, if the measurement process returns "3.5 A", we assign this symbol, whose meaning is assumed to be known, to the current.

The above relationship between an empirical property and a symbol can be represented using a mathematical formalism. Consider a generic property and let $q_1, q_2, \ldots, q_{\mathrm{n}}, \ldots$ represent single manifestations of that property, so that we can define a set of all possible manifestations as

$$Q = \{q_1, q_2, \ldots, q_{\mathrm{n}}, \ldots\}. \tag{1.1}$$

Let $X$ represent a class of symbols $x$. Formally, measurement is an objective empirical operation:

$$\mu : Q \rightarrow X, \tag{1.2}$$

or

$$x = \mu(q), \tag{1.3}$$

where $q$ represents a generic manifestation of the considered property. It is worth noticing that $\mu$ is not a one-to-one mapping, since it maps different but indistinguishable property manifestations to the same symbol.

To better describe this mapping, we need to look more in details into the measurement block in Figure 1.1 and identify the internal processes that contribute to the generation of the measured value. Usually, the first process that comes to mind is the *experimental process*. The above example on the validation of the Ohm's law describes, actually, an experimental process. However, the experimental processes cannot alone realize the mapping structure (1.2). They can only put into practice the *descriptive processes* that allow one to identify and implement (1.2).

So, let us analyze the most important processes that must be considered.

***1.4.1.1 Measured Property Definition*** The first required step is to properly define the manifestation of the property $q$ that we intend to measure. This requires a model for the object that manifests the measurand.

For instance, let us assume that we want to measure the width of a plate which is almost rectangular. This implies that we are describing the plate itself as a rectangle, that is its width is assumed to be the same all along the other side, and hence leads us to measure a single value on a single position, and assign this value to its "width".

If we are not sure that the plate is almost rectangular, we can model it, for instance, with a trapezoid, and measure two width values, on the opposite edges of the plate. In this case, we are assuming that its "width" varies linearly among the two obtained values.

The above examples show that the definition of the manifestation of the property that we intend to measure has a direct impact on the experimental process, and hence cannot be disregarded. We will see later, in section 1.5.1, that this point has a significant impact also on the quantity of information that the measurement process provides about the measurand.

***1.4.1.2 Mapping Definition*** Once the manifestation of the property $q$ that we intend to measure has been suitably defined, the mapping relation $\mu$ in (1.2) has to be defined in such a way that all possible empirical relations between measurands can be mapped into abstract relations between symbols.

To understand this point, let us consider the class of the height of a set of people, and some empirical relations that can be established for property "height". We may have:

- *Unary relations*
  For instance: Charlie is *tall*, Snoopy is *not tall*.
- *Binary relations*
  For instance: Charlie is *taller* than Linus, Charlie is *much taller* than Snoopy.
- *Ternary relations*
  For instance: Snoopy is *taller* than Charlie *if* sitting on Linus' shoulders.

Mapping $\mu$ must preserve these relations on the returned symbols. This means that, for instance, the empirical relation "Charlie is taller than Linus" must be reflected into relation: $\mu(\text{Charlie}) > \mu(\text{Linus})$. Therefore, a measurement procedure that provides the following quantity values: $\mu(\text{Charlie}) = 1.55$ m and $\mu(\text{Linus}) = 1.40$ m is correct, while a measurement procedure that provides the following values: $\mu(\text{Charlie}) = 1.35$ m and $\mu(\text{Linus}) = 1.43$ m is not correct, because it does not preserve the empirical relation set between properties.

The above concept can be expressed using a mathematical formalism. Let us assume that a set $\mathcal{R}_{\text{Q}}$ of empirical relations $R_{\text{Q1}}, R_{\text{Q2}}, \ldots, R_{\text{Qn}}, \ldots$ exists on the set $Q$ of all possible manifestations of the considered property. Let us express this as

$$\mathcal{R}_{\text{Q}} = \{R_{\text{Q1}}, R_{\text{Q2}}, \ldots, R_{\text{Qn}}, \ldots\}$$

and let

$$\mathcal{R}_{\text{X}} = \{R_{\text{X1}}, R_{\text{X2}}, \ldots, R_{\text{Xn}}, \ldots\}$$

be a set of relations defined on the set of symbols $X$.

Then, the property is formally represented by an *empirical relational system* [4]

$$\mathcal{Q} = \langle Q, \mathcal{R}_{\text{Q}} \rangle,$$

which is mapped by the measurement procedure onto the *numerical relational system* [4]

$$\mathcal{X} = \langle X, \mathcal{R}_{\text{X}} \rangle.$$

The so called *representation condition* requires measurement to be a correspondence between property manifestations and symbols in such a way that the relations between the property manifestations imply and are implied by the relations between their images in the symbol set. Thus, $\mathcal{R}_{\text{Q}} \rightarrow \mathcal{R}_{\text{X}}$ must be a one-to-one mapping.

When the above constraint is satisfied, mapping

$$\mu : \mathcal{Q} = \langle Q, \mathcal{R}_{\mathrm{Q}} \rangle \rightarrow \mathcal{X} = \langle X, \mathcal{R}_{\mathrm{X}} \rangle \tag{1.4}$$

is called *homomorphism*.

***1.4.1.3 Scale Definition*** The representation condition defined in the previous section 1.4.1.2 is a required feature of mapping $\mu$, but does not yet explain how a numerical value quantifying property manifestation $q$ can be assigned to symbol $x$. To do so, a suitable *measurement scale* should be defined, through the following steps.

1. A set $\Omega_{\mathrm{S}} = \{\omega_{\mathrm{S_n}}\}$ of *standard objects* is chosen in such a way that it provides a set $Q_{\mathrm{S}} = \{q_{\mathrm{S_n}}\}$ of manifestations of the same property as the measurand.
2. A relation $\mu_{\mathrm{S}} : \Omega_{\mathrm{S}} \rightarrow Q_{\mathrm{S}}$ is defined between the set of standard manifestations of the property and symbols, so that a set $X_{\mathrm{S}} = \{x_{\mathrm{S_n}}\}$ of symbols is obtained, corresponding to each standard manifestation, and the representation condition is satisfied.
3. The measurand is compared with the set of standard manifestations and a single manifestation is selected from set $Q_{\mathrm{S}} = \{q_{\mathrm{S_n}}\}$, as the result of the comparison.
4. Relation $\mu_{\mathrm{S}}$ is applied to the selected standard manifestation and the associated symbol is provided as the measured value.

The above process can be formally described by the following set of relations.

$$\begin{array}{ccc} q & \overset{\mu}{\rightarrow} & x \\ \downarrow & & \uparrow \\ \{q_{\mathrm{S_n}}\} & \overset{\mu_{\mathrm{S}}}{\rightarrow} & \{x_{\mathrm{S_n}}\} \end{array}$$

There are different possible ways to define a measurement scale, and their properties may limit the analysis that can be performed on the measurement results. Five different kinds of scales are generally employed, ordered according to an increasing level of richness in the information they convey. The ordering is such that all operations allowed at a given level are allowed also at the upper levels.

- *Nominal scales*
  When this kind of scale is adopted, each measured object is placed in a *class*, according to the value manifested by the considered property. Symbols are only *class tags* and the classes are not ordered.

  The only admissible operation on these scales is a one-to-one mapping, and the only statistical inference is the mode.[3]

[3] Value of the most commonly occurring item.

When this kind of scale is adopted, the quantity of information provided by the measured value (see Section 2) is associated to the probability of correctness of the assignment of the measurand to one of the possible classes.

- *Ordinal scales*
  When this kind of scale is adopted, each measured object is placed in an *ordered class*, always according to the values of the considered property. The classes are ordered with respect to the property. Properties for which an ordinal scale (or an upper level scale) is defined are called *quantities*, and the related symbols are called *quantity values*.

  An example of ordered scale is the *Mohs* scale of mineral hardness, and the ordering is based on the ability of a mineral to scratch a softer one.

  Increasing mappings (that preserve the order) are additional allowed operations on this scale, and median[4] and percentiles are additional allowed statistical inferences.
- *Interval scales*
  This kind of scale is an ordinal scale that conveys additional information about the *size* of the interval between *adjacent classes*.

  Examples of interval scales are the Celsius and Fahrenheit temperature scales, and the Richter seismic magnitude scale.

  Sums and subtractions are the only allowed mathematical operations. Affine mappings between scales that preserve the order and interval proportions[5] are also allowed. An example of affine mapping is the well-known transformation of a temperature value $\vartheta_C$ given on a Celsius scale into a temperature $\vartheta_F$ given on a Fahrenheit scale: $\vartheta_F = 1.8 \cdot \vartheta_C + 32$.

  Mean and standard deviation are allowed statistical inferences for these scales.
- *Ratio scales*
  This kind of scale is an interval scale to which a *zero* (or *absolute*) element is added, representing the lack of the considered property.

  This scale is the most useful and commonly used scale in physical sciences, since it preserves ordering, size of intervals, and ratio between properties.

  All arithmetic operations are allowed on this scale, and the proportional mapping ($y = a \cdot x, a > 0$) is also allowed.

  The geometrical mean and the coefficient of variation are also allowed statistical inferences.
- *Absolute scale*
  This kind of scale is used to count elements. In this case there is only one possible mapping: the element count. Therefore, the only meaningful transformation is *identity*.

Unless differently stated, in the following we will refer only to measurement ratio scales.

[4] Value of the middle ranked item.

[5] An affine mapping that preserves the order is mathematically expressed as: $y = a \cdot x + b, a > 0$.

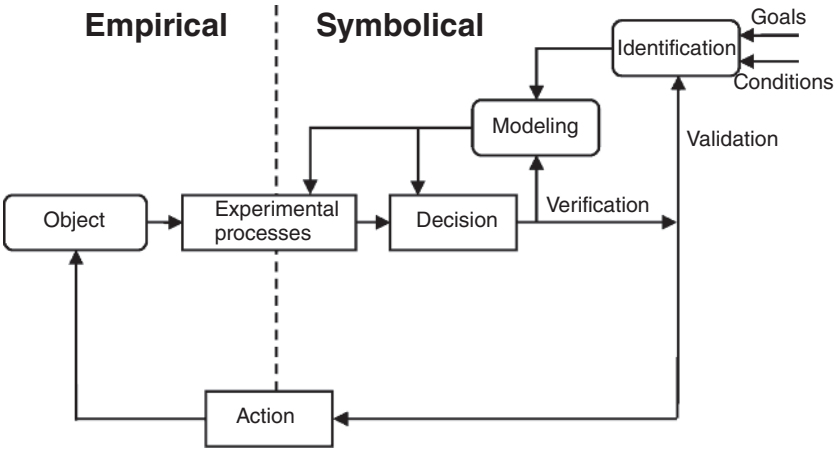


**FIGURE 1.2** A measurement model including different kinds of involved descriptive processes.

### 1.4.2 A More Complex Measurement Model

The model presented in the previous section 1.4.1 describes how measurement activities accomplish the bridge function between the empirical and symbolical worlds by performing both descriptive and experimental processes. Indeed, the former processes are required to ensure a correct implementation of experimental processes and the interpretation of the obtained results.

Even though the model in Figure 1.1 is actually the core of the measurement activity, a more detailed model can be drawn, in which descriptive processes and feedback paths are identified, as shown in Figure 1.2.

The first point that must be considered is that, regardless of the field of application, usually measurement is not a self-motivating process, but is rather a goal-oriented activity. Indeed, the measurement results are quite always employed as relevant input elements in a *decision-making process* aimed at identifying the best *actions* needed to achieve established *goals* while satisfying given *conditions*.

For instance, the speed value provided by the speed indicator in a car is employed to decide whether the car speed has to be decreased or increased (the action) in order to reach destination in a given time, while driving safely (the goals) according to the weather conditions and the speed limits (the conditions).

Obviously, this model still represents measurement as a bridge between the empirical and symbolical worlds, as discussed in the previous section. Moreover, it can be readily verified that the model in Figure 1.2 refines the one in Figure 1.1, by explicitly showing the different descriptive processes that have to be performed to achieve the measurement goals. Let's analyze them.

*Identification* Usually, the result of any measurement activity does not depend only on the measurand, but is affected also by a number of other properties belonging to the empirical environment in which the measurement is performed. Therefore, the

identification of the measurement context, that is the identification of all properties affecting the measurement result and the possible interactions among them, is the first step to be performed.

Four are the critical points in context identification.

- Measurand identification
  At first, it is necessary to identify which property (the measurand) of the object we are analyzing is the best suited one to achieve the goals, under the given conditions.[6] This means that we have to identify which property conveys the most relevant information needed to accomplish the goal.
- Measurement system identification
  Once the measurand has been identified, we need to identify a suitable measurement system. The selected measurement system should be the one that is expected to be the best one to achieve the goals, under the given conditions.
- Empirical environment identification
  During the measurement activity, measurand, measurement system and environment interact. For instance, a temperature probe changes the heat transfer of a body as well as its thermal capacity when it contacts it. Similarly, an electronic instrument dissipates some power in the environment, thus contributing to its temperature variations.

  These interactions manifest themselves through the effect that *variations* in one or more *environment properties* (i.e. temperature, pressure, humidity) and *time* have on the measurement result.

  It is therefore necessary to identify which environment properties, called *influence properties* (or *influence quantities*), affect the measurement result in a significant way,[7] so that they can be (and should be) controlled during the measurement process.

  Moreover, all significant mutual interactions between measurand, measurement system and environment must be identified, in order to achieve an accurate description of the whole measurement process.

  Usually, only a limited number of environmental properties may influence the measurand and the measurement system significantly. It is also worth noting that influence properties can be controlled and kept constant within a given tolerance, while time cannot be controlled or kept constant.

*Modeling* The identification of the aforementioned items is useless unless a mathematical model is defined that describes the measurement context accurately, that is the behavior of each identified interaction among the measurand, the measurement

[6] In order to achieve the goals, the identification of more than only one property might be required. However, in the following, we are considering the measurement of only one property for the sake of simplicity. All conclusions can be easily extended to the case of the measurement of multiple properties.

[7] The meaning of "significant" will be fully understood later, in section 1.5, when measurement uncertainty will be covered. Here, let us only anticipate that the measurement result is affected in a significant way if the uncertainty contribution associated with the considered influence property is non-negligible with respect to all other uncertainty contributions.

system and the environment. In that model, the identified properties are represented by *mathematical variables*, and the mutual interactions are represented by *mathematical relationships*.

In general, these variables and relationships are assumed to be deterministic. However, non-deterministic approaches are employed, such as probabilistic, fuzzy, neural, and neuro-fuzzy models, when either the measurand or the measurement systems, or both, are too complex to be described by a deterministic model in a satisfactory way.

The output of the modeling step is probably the most critical element in the whole measurement process, since it provides the symbolical description of the measurement process itself. All further steps are based on this model, that is therefore affecting the *quality*[8] of the measurement result.

As a matter of fact, this model provides always an *incomplete knowledge* of the measurement context, mainly because of the incomplete knowledge of the measurand, the fact that quantities and interactions that have, singularly taken, negligible effect have not been considered in the model, and the intrinsic approximation of the mathematical representation of variables and relationships.

This incomplete knowledge of the measurement context is reflected in the amount of information that the measurement result provides about the measurand. Due to all approximations performed in the description of the context, only a *limited* amount of information can be obtained by the measurement process, regardless of the accuracy of the employed instrumentation.

It can be stated that the limited amount of information provided by the measurement activity is useful in the subsequent decision step if and only if we can quantify the amount of missing information. This part of the modeling process is called *uncertainty expression and evaluation*. Although this is logically a part of the descriptive processes, it is such an important part of modern metrology that it will be covered as a separate point in Sections 1.5 and 1.6. It is nevertheless important to keep in mind that the need for considering, expressing, and evaluating uncertainty in measurement originates in this step of the measurement process.

*Experimental processes* The *experimental processes* block of the measurement model depicted in Figure 1.2 puts into practice the model defined in the previous step, by performing the processes mentioned in Section 1.4.1.

While the descriptive processes are quite general, the experimental processes are specific for each measurement application, and depend on the given goals and conditions. As it will be mentioned in the next chapters, the modern measurement systems are largely based on the conversion of the input signals into the discrete domain and the subsequent digital signal processing (DSP) of the acquired data. Therefore they apparently allow the measurement model to be implemented in a rather straightforward way.

[8]Here we prefer to refer to the general concept of *quality*, since it gives an immediate perception that the measurement result is not *ideal*, or *perfect*, and therefore its quality can be good or not. The next sections 1.5 and 1.6 will show how this "quality" can be expressed and evaluated

The contribution of such measurement systems to uncertainty is therefore mainly related to the additional approximations introduced by a digital implementation and to imperfections of the field interface (usually sensors and transducers), the electronic front-end and the sampling and analog-to-digital converting devices.

*Decision* As previously mentioned, measurement results are quite always employed as input data in a decision-making process. Generally, this process involves the comparison of the measurement result with one or more threshold values, and a decision is taken according to where it is located, with respect to the given threshold values.

This process is only apparently simple, because, due to measurement uncertainty, the measurement result cannot be represented with a single value. As it will be shown in Section 1.6, any measurement result does always represent an interval of values, with a stated coverage probability [1]. We do not want to enter the mathematical details of a decision-making process in the presence of uncertain values, because this is beyond the scope of this book. We simply wish to draw the readers' attention on the fact that the involved comparison process must be carefully considered, and the final decision does always imply a risk that the decision itself is wrong. The risk of wrong decision depends on measurement uncertainty, and is generally higher, the higher measurement uncertainty is.

It is worth noting that the role of the decision process is not confined to actions that follow the measurement process, but has some interesting implications within the measurement process itself. As shown in Figure 1.2, the normal consequence of the decision-making process is an action on the measured object. This is also, in general, the final goal of the measurement process: data are provided to act on the measured object so that the given goals can be achieved.

Figure 1.2 shows also other connections of the decision block toward the modeling and identification blocks. These connections show a feedback path, inside the measurement model, so that the model can be verified and the whole measurement process can be validated.

If the whole measurement context has been correctly modeled and no significant environmental properties and significant interactions have been neglected, the quality of the final measurement result should meet the expectation. If this is not the case, the model has to be refined, until the expected result quality is met.

Once the model has been verified, it is possible to validate also the whole measurement process, by checking that all goals have been achieved, without violating any of the given conditions. This step allows us also to verify that the whole measurement context has been correctly identified. If the identification is correct, the measurement goals are achieved, otherwise also the identification process has to be refined.

### 1.4.3 Final Remarks

The different processes that take part in any measurement activity have been briefly reported and discussed in this section. The obtained model shows the steps required

to implement a measurement activity that is correct from a methodological point of view, and not only from an experimental (or instrumental) perspective.

The most important conclusion that can be drawn is that any measurement result does provide quantitative information about the measurand, thus achieving the original task of measurement to be a bridge between the empirical and symbolical worlds, but the amount of achieved information is always limited and, generally, never complete.

From a philosophical and ontological perspective, limited information should not be considered an advancement of knowledge, since knowledge itself implies complete and absolute cognition of the considered phenomenon. However, measurement is the only available tool to "know something" of what we "are speaking about", according to Lord Kelvin's famous speech, at least as far as scientific issues are concerned.

Therefore, from a more practical, engineering point of view, a way must be found to quantify also the completeness (or incompleteness) of the information conveyed by a measurement result. If and only if this additional, quantitative information is added to a measurement result, the result itself can be advantageously used in any subsequent decision making process. Otherwise, it is only a meaningless number.

## 1.5 UNCERTAINTY IN MEASUREMENT

The last section has clarified which is the role of the *other available relevant information* that, according to the definition of measurement result given by the VIM [1], must be specified together with the provided quantity values. This kind of knowledge conveyed by the measurement result is aimed at quantifying how complete the information associated with the quantity values provided by the measurement process is.

The problem of qualifying the measurement result with a quantitative attribute capable of characterizing the completeness of the information provided about the measurand has been considered since the beginning of metrology, and has undergone a significant revision during the last two decades of the twentieth century, when the modern concept of *uncertainty* has been introduced and adopted.

Understanding this concept is not immediate. As noted by the Guide to the Expression of Uncertainty in Measurement (GUM) [5], the word *uncertainty* means doubt, and thus in its broadest sense *uncertainty of measurement* means doubt about the information provided by the result of a measurement. The first question that we have to answer to, if we wish to fully understand the uncertainty concept, is where this doubt originates and why. Additional considerations on the model presented in Section 1.4 will be helpful in finding an answer to this question.

However, the concept of doubt is hardly quantifiable. Therefore, the use of the term *uncertainty* to quantify how much we should doubt about the obtained measurement result seems somehow contradictory and tautological, since we apparently try to quantify a doubt with a doubt!

Fortunately, this contradiction is only apparent and is due to a lexical problem. Because of the lack of different words for the *general concept* of uncertainty and the specific quantities that, as we will see in Section 1.6, provide *quantitative measure* of

the concept, the same word "uncertainty" is used, in metrology, in these two different senses [5].

To avoid confusion, Section 1.5 is aimed at introducing and discussing the general concept, while Section 1.6 will consider the quantitative measure of the concept.

### 1.5.1 The Origin of the Doubt

The measurement activity, as noted in Section 1.4.2, is always framed in a context, and this context is generally complex. To identify the best actions required to achieve the measurement goals, according to the given context, a model is needed, as reported in Section 1.4.2.

It is well known that any model, even the best one, provides only an approximated description of the entity it is aimed at representing. The model of the measurement context makes no exception to this general rule, and therefore the origin of the doubt, that is, the sources of the different possible contributions to uncertainty, has to be looked for in this model.

According to the different activities modeled by the different blocks in the diagram of Figure 1.2, the following contributions to uncertainty can be defined.

*Definitional uncertainty* The *identification* and *modeling* blocks of Figure 1.2 have the tasks of identifying and modeling the measurand, the relevant environment properties, and their mutual interactions.

The resulting description of measurand and environment is imperfect. For instance, if we have to measure the external diameter of a pipe, we can model the pipe as a geometrical cylinder. This implies that we are assuming the diameter is the same all along the pipe, and leads to perform a single measurement in a randomly chosen point along the pipe. If the pipe surface is rough, the diameter is not the same along the pipe's length and therefore different measurement results can be obtained, depending on where the measurement is performed. This contribution to uncertainty is called *definitional*,[9] because it is originated by an imperfect definition of the measurand. The consequence of not considering it may lead to incorrect decisions. If, for instance, the diameter has to be used to size a hole into which the pipe has to go through, modeling the pipe as a cylinder makes it possible that the hole is too tight and the pipe cannot be inserted.

Likewise, an incorrect definition of the mutual interaction between measurand and environment may lead to similar problems. If, for instance, we have a metal pipe, and the interaction between temperature and metal thermal expansion is not considered, the same problem as the one mentioned above might be encountered.

Since the identification and modeling steps are the preliminary steps of any measurement activity, and affect all subsequent steps, definitional uncertainty represents the *lower bound* of measurement uncertainty. Its correct evaluation is therefore important, also from an economical point of view, because it makes no sense to reduce the

[9]The VIM definition states that the definitional uncertainty is the *component of measurement uncertainty resulting from the finite amount of detail in the definition of a measurand*.

other contributions to uncertainty at a much lower level than that of the definitional uncertainty. Indeed, the additional cost is not repaid by a lower uncertainty, since the dominant contribution is, in this case, the definitional one.

*Interaction uncertainty* In general, when a measurement system is connected to the measurand, it interacts with the measurand and this interaction may change the measurand value during the measurement process (loading effect). For instance, if a voltmeter is connected to a voltage source to measure its output value, and the input impedance of the voltmeter is of the same order of magnitude as the source internal impedance, the measured voltage will be significantly lower than the desired open-source value.

An imperfect or incomplete description of the interaction between measurand and measurement system originates a contribution to uncertainty that is called *interaction* uncertainty. Actually, the origin of this uncertainty contribution is still in the description and modeling steps, and therefore this contribution to uncertainty can be seen as a definitional uncertainty.[10] However, the interaction between measurand and measurement system is so important that it is generally useful to consider the resulting contribution to uncertainty separately from the other ones.

*Instrumental uncertainty* We call *measurement system* all instruments and procedures employed in the measurement activity. The imperfect behavior of this system and its interactions with the environment give rise to another uncertainty contribution, called *instrumental* uncertainty.[11]

It is worth noting that a significant part of this contribution can be originated outside the employed measurement devices and equipment. In Section 1.4.1 a number of descriptive processes have been discussed, as part of the measurement model. The practical implementation of the scale definition described in Section 1.4.1.3 is the most important of these processes, as far as instrumental uncertainty is concerned.

This process requires the definition of a set of standard objects, with which the measurand is compared during the measurement procedure. A symbol is associated with each standard object, so that the comparison procedure provides the symbol associated with the measurement result (as in Section 1.4.1.3). The implementation of this process requires the practical realization of the standard objects.[12] Thus, due to the unavoidable imperfections, the standard objects do always provide a value that is different from the one assigned to the associated symbols.

[10]For this reason, this contribution to uncertainty is not explicitly defined by the VIM.

[11]Instrumental uncertainty is defined by the VIM as the *component of measurement uncertainty arising from a measuring instrument or measuring system in use*.

[12]When physical quantities are measured, these standard objects are, in principle, the primary standards that realize the involved measurement unit. Since, in practice, primary standards cannot be used in every measurement process, local (or secondary) standards can be used, provided they are *traced* back to the primary ones. In this case, the instrumental uncertainty contains an additional contribution, called *calibration uncertainty*, that reflects the uncertainty of the calibration process employed to trace the local standards back to the primary ones.

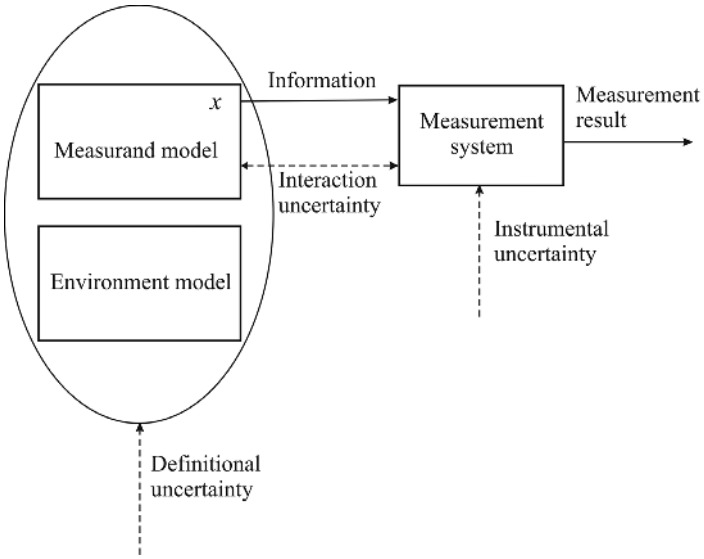


**FIGURE 1.3** Different uncertainty sources affecting the information flow in the measurement procedure.

This is generally a lower bound to instrumental uncertainty. Indeed, taking into account that the calibration process is performed at reference environmental conditions, any deviation, in the actual operative conditions, from these reference conditions increases the instrumental uncertainty. In principle, if the mutual interactions between the instrument and environment are known, a correction can be applied, and the instrumental uncertainty can be reduced back to the calibration one. However, due to the intrinsic approximation of the mathematical model describing the instrument–environment interactions, an additional uncertainty contribution always arises.

The above considerations and the origin of the different uncertainty contributions are graphically sketched in Figure 1.3, where the information flow that originates the measurement result is also shown.

### 1.5.2 The Different Effects on the Measurement Result

In Section 1.5.1, the origin of the different uncertainty contributions has been analyzed and discussed. It can be synthetically concluded that the uncertainty contributions originate in some effects that cannot be properly taken into account by the considered mathematical and logical models.

It can be also intuitively concluded that the different effects generate uncertainty contributions that may affect the measurement result in different ways.

Of course, it is not possible to analyze, in this short survey chapter, all possible different effects. All of them, however, can be grouped into two main classes, each one

characterized by a well-defined impact on the final measurement result: *systematic* effects and *random* effects.

*Systematic effects* Systematic effects are constant in time and reflect, generally, the effect of an influence quantity on the measurement result. Therefore, if the influence quantity does not change its value when repeating the measurement procedure, the resulting effect remains the same, and the measurement result does not change.

Systematic effects are therefore very dangerous, because their detection is very difficult. A way to detect them is to change the values of the influence quantities, or modify the measurement procedure, so that the systematic effect changes, manifesting itself in a variation of the measurement result. Another way is to compare the measurement results obtained by using two independent measuring systems, provided that one is known to give rise to negligible systematic effects with respect to the other.

Thus the systematic effect can be recognized and *corrected*.

Let us consider, for instance, a weighbridge, and let us suppose that the two beams have different lengths. This difference is directly reflected on the measurement result and a systematic effect arises, that cannot be corrected if the length difference is unknown, as it usually is. However, this effect can be compensated for, if a different measurement procedure is adopted. Let us call $W_1$ the obtained measurement result. If the measurement is repeated exchanging the position, on the two pans of the weighbridge, between the unknown weight (the measurand) and the standard weights, a different result $W_2$ is obtained. It is well known that the systematic effect due to the different beam lengths is compensated for if the value: $W = (W_1 + W_2)/2$ is taken as the measurement result.

Let us now consider a strain measurement performed by a resistive strain gauge. It is known that these devices are very sensitive to temperature. However, the resulting effect is fairly linear in a quite wide range of temperature variations. Therefore, if the same strain is measured at different temperatures, the law of variation can be easily obtained through a linear interpolation, thus enabling the correction of the measurement result when the environment temperature is known.

In general, the correction factors for the most significant influence quantities are provided by the instrument manufacturer or by the calibration certificates so that it is no longer necessary to repeat the measurement procedure under different values of the influence quantities. In this case, the systematic effect is identified and corrected by means of a priori information.

The correction of the systematic effects is such an important issue in measurement applications that the GUM [5], in its article 3.2.4, states: *It is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects*.

At last, it is important to consider that, due to the always present imperfections, the corrections can only reduce the impact of the systematic effects on the measurement result, but cannot cancel it totally. Therefore, a residual uncertainty component remains, due to incomplete knowledge of the systematic effects [5].

*Random effects* Random effects are due to unpredictable or stochastic temporal and spatial variations of influence quantities. They give rise to variations in repeated observations of the measurand.

Due to the unpredictable nature of these effects, they cannot be compensated. However, because of their stochastic nature, the statistical *expectation* of their effect on the measurement result is zero.

Therefore, their impact on the measurement result can be usually reduced by repeating the measurement procedure, under the same experimental conditions, a statistically significant number of times, and taking the average value of all measured values. Moreover, the statistical standard deviation of the measured values provides information about the measurement uncertainty, as it will be shown in Section 1.6.

### 1.5.3 The Final Effect

In general, a measurement procedure is affected by both systematic and random effects. Their origin and main characteristics have been discussed in Sections 1.5.1 and 1.5.2. We still have to analyze which is their ultimate effect on the measurement result.

To do this, let us suppose that the value of the measurand is represented by the center of a target, as shown in Figure 1.4. Let us now suppose to repeat the measurement procedure under the same experimental conditions. Each measured value can be graphically represented as a shot on the target.

According to the considerations in the previous sections, the random effects cause a dispersion of the measured values, similar to the effects, on real targets, of unpredictable wind gusts or unconscious movements of the shooter. If only random effects

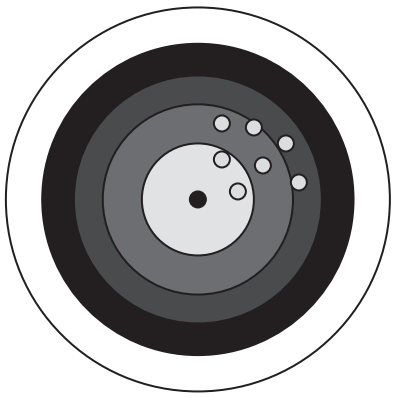

**FIGURE 1.4** Graphical representation of the result of repeated measurements. The black center of the target represents the measurand value. Each small gray dot represents a single measured value. The dispersion of the measured values around their average is due to the random effects. The displacement of all values with respect to the center is due to the systematic effects.

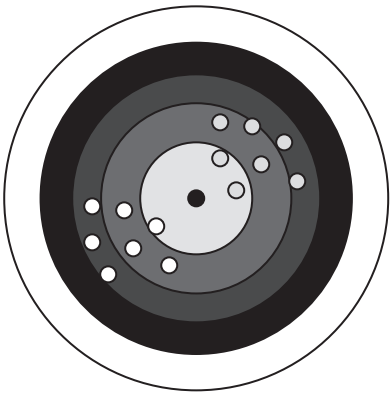

**FIGURE 1.5** Graphical representation of the result of repeated measurements, for the same measurand as that in Figure 1.4, but under different measurement conditions. The new measured values are represented by the small white dots, while the gray dots are the same as in the case of Figure 1.4. The impact of the changed systematic effects is evident in the different position of the white dots with respect to the gray ones.

were present, all shots would be uniformly distributed around the center of the target. Conversely, systematic effects cause their average value (the center of gravity of the shot pattern) to be displaced with respect to the center. This is similar to the effect of an incorrect setting of the sight on the weapon used to shoot the target.

Let us now suppose to repeat the measurement procedure after a period of time, or in another place, or with different instruments, or under different environmental conditions. The results can be now graphically represented by Figure 1.5, where the previous results are shown as well.

The dispersion of the shot positions with respect to their center of gravity is quite similar to the one of the shots in Figure 1.4, showing that the random effects did not change. On the contrary, the center of gravity of the new shots is displaced, with respect to the previous one, thus showing that the systematic effects have changed, as expected, since we assumed that something changed in the measurement conditions. In our example, this effect can be explained with a different setting of the sight.

We can now draw a first significant conclusion. We cannot express a measurement result with a single value, because this is meaningless. This same value will be barely obtained by a new measurement.

Moreover, this single measurement value cannot be used in a comparison, and hence in a decision-making process. Since all measurement results are different, we would be forced to conclude that, for instance, the mass of an object measured in New York today is different from the mass measured in any other place, at the same latitude and height, tomorrow. There is no doubt that this is not acceptable.

However, if we observe again Figures 1.4 and 1.5, we understand that the measurement result is represented by the whole set of shots. Only the complete set is representative of the capabilities of the implemented measurement procedure, exactly

as the whole set of shots is representative of the capabilities of a shooter. We will never judge the ability of a shooter from a single shot: it might be a very lucky shot of a terrible shooter!

It should now be clear that a measurement result is meaningful only if it expresses a distribution of values. If all significant systematic effects have been recognized and have been corrected for, we may expect that the measurand value falls within the distribution of values expressed by the measurement result, close to their average value.

It should also be clear, now, that measurement uncertainty should provide a quantitative measure of this dispersion.

## 1.6 UNCERTAINTY DEFINITION AND EVALUATION

According to the concepts presented in Sections 1.5 and 1.5.3, a measurement result cannot be expressed by a single measurement value. On the other hand, the measurand value should be expected to fall within the set of all possible measurement values provided by an experimental process. Also, it is possible to assume that these values distribute about the measurand value if all systematic effects are compensated for.

Therefore, we can now understand why, according to the definition given by the VIM [1], the result of a measurement cannot be expressed only by a single quantity value, but it requires also some additional information aimed at characterizing the distribution of values that could reasonably be attributed to the measurand.

Before discussing how this additional information can be expressed, it is important to understand the conditions it has to satisfy to be usefully employed in any practical situation. According to the GUM [5], the three following conditions must apply.

- The ideal method for evaluating and expressing the uncertainty of the result of a measurement should be *universal*. This means that 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*. This means that 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*. This means that 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.

The GUM [5] considers also an additional important requirement in many industrial and commercial applications, as well as in the areas of health and safety. In these areas, it is often necessary to provide an interval about the measurement result that may be expected to encompass a large fraction of the distribution of values that

could reasonably be attributed to the quantity subject to measurement. Thus, the GUM states [5] that the ideal method for evaluating and expressing uncertainty in measurement should be capable of readily providing such an interval, in particular, one with a *coverage probability* or *level of confidence* that corresponds in a *realistic way* with that required.

Different approaches may be followed to satisfy the above requirements. Two of them have, until now, prevailed over all possible ones, and will be considered in this chapter.

The first one is the well known, traditional *error* approach, that has been followed until the theoretical considerations that brought to the first edition of the GUM, in 1995, proved that it was flawed from a strict ontological perspective. The second one is, of course, the one followed by the GUM in its current edition [5].

We are here considering both of them, including the old error approach, because its critical analysis is extremely useful to understand the more modern *uncertainty* approach followed by the GUM.

### 1.6.1 The Error Concept and Why it Should be Abandoned

The error approach is quite simple. It is based on the assumption that the true value $x_t$ of measurand $x$ can be somehow known. If a measured value $x_m$ is returned by the measurement process, then the *measurement error* is defined as

$$\varepsilon = x_m - x_t \tag{1.5}$$

and the *relative measurement error* can be defined as

$$\varepsilon_r = \frac{x_m - x_t}{x_t}.$$

If the maximum error affecting the measurement result is denoted by $\varepsilon_M$, the interval

$$(x_m - \varepsilon_M,\ x_m + \varepsilon_M) \tag{1.6}$$

can be built about the measured value, within which the value that can be attributed to the measurand is supposed to lie with 100% confidence.

In many cases, a measurand $y$ is not measured directly, but is determined from $N$ other quantities $x_1, x_2, \ldots, x_N$ through a functional relationship

$$y = f(x_1, x_2, \ldots, x_N). \tag{1.7}$$

The expansion of (1.7) about the true values $x_{1_t}, x_{2_t}, \ldots, x_{N_t}$ in a first order Taylor series yields

$$\varepsilon_y = \frac{\partial f}{\partial x_1}\varepsilon_{x_1} + \frac{\partial f}{\partial x_2}\varepsilon_{x_2} + \cdots + \frac{\partial f}{\partial x_N}\varepsilon_{x_N}, \tag{1.8}$$

where $\varepsilon_y$ is the measurement error affecting $y$, under the assumption that the higher order terms of the Taylor series expansion are negligible for small values of the measurement errors $\varepsilon_{x_i} = x_{m_i} - x_{t_i}, i = 1, 2, \ldots, N$.

It can be readily checked that the error concept satisfies the three requirements given at the beginning of Section 1.6. In particular, (1.8) ensures that errors are transferable, and (1.6) ensures that the error can be employed to define an interval, about the measurement result that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand.

The major flaw of the error approach lies in the fundamental assumption done to define the error itself in (1.5). In fact, this equation implies the knowledge of the measurand true value. However, all considerations in section 1.5 give clear evidence that no instrument and no measurement process can ever provide the true value of the measurand. Therefore, in practice, the true value of the measurand is always unknowable.

Moreover, the very basic concept of measurand can be lost when using a more refined model of the measured object. This is the case, for instance, with the concept of the length of an object at the atomic level, or with other physical quantities when considering the modern quantum physics.

The direct consequence of these statements is that the measurement error concept is meaningless also from a practical point of view, since the error itself cannot be evaluated, being its definition (1.5) based on an unknowable, and sometimes non-exactly definable quantity.

For this reason, the error concept has been abandoned, and replaced by the more modern concept of uncertainty.

### 1.6.2 Uncertainty Definition: The GUM Approach

As discussed in the previous section 1.6.1, the true value of the measurand is unknown and unknowable. Therefore, the task of *characterizing the distribution of values that could reasonably be attributed to the measurand* is made difficult by the fact that, not knowing the measurand, we cannot know whether the measured values belong to that distribution or not. This means that a more complex mathematical approach than the simple error concept is required to define and handle this problem.

In particular, the GUM [5] suggests to model the measured values as realizations of a random variable. Following this approach, the distribution of the measured variables is represented by a *probability density function* (pdf) and the related *standard deviation* provides a quantitative estimate of the dispersion of the possible measured values about their mean value, which is usually assumed to be the best estimate of the measurand value.

According to these considerations, the GUM considers the standard deviation of the pdf of the values that can be attributed to the measurand as a quantitative measure of uncertainty. This quantitative measure of uncertainty is called *standard uncertainty* and is defined as follows [5].

*Standard uncertainty*

Uncertainty of the result of a measurement expressed as a standard deviation.

Thus, standard uncertainty is defined as an estimate of the standard deviation of the distribution of values that can be reasonably assigned to the measurand.

The suggested notation for the standard uncertainty associated to a measured value $x$ is $u(x)$.

Since we assumed to know the pdf of the distribution of measured values, a *coverage probability* (or *level of confidence*) can be assigned to an interval built about its mean value $\bar{x}$ as $[\bar{x} - u(x), \bar{x} + u(x)]$. Of course, the coverage probability depends on the pdf of the measured values. If, for instance, the pdf is normal, the coverage probability is the well known 68.27% value.

It might be then concluded that the additional GUM requirement mentioned at the beginning of Section 1.6 is satisfied, since standard uncertainty allows one to provide an interval [1] that encompasses a fraction of the distribution of values that could reasonably be attributed to the measurand with a known coverage probability.

However, the GUM requires that the coverage interval encompasses a *large* fraction of the distribution of values that could reasonably be attributed to the measurand, and that its coverage probability corresponds in a *realistic way* to that required.

The interval defined using standard uncertainty may not satisfy these requirements. Therefore, the GUM defines the *expanded uncertainty* as follows [5].

*Expanded uncertainty*

Quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand.

The suggested notation for the expanded uncertainty associated to a measured value $x$ is $U(x)$.

The GUM states that the fraction may be viewed as the coverage probability or level of confidence of the defined interval.

It is also extremely important to note that a specific level of confidence can be associated with the interval defined by the expanded uncertainty if and only if the pdf characterized by the measurement result and its standard uncertainty are known, or if *explicit or implicit assumptions* are made on the pdf itself.

Of course, the level of confidence that may be attributed to this interval can be known only to the extent to which such assumptions may be justified.

The expanded uncertainty can be obtained by multiplying the standard uncertainty by a suitable *coverage factor* $K$, defined by the GUM as follows [5].

*Coverage factor*

Numerical factor used as a multiplier of the standard uncertainty in order to obtain an expanded uncertainty.

**TABLE 1.1 Value of the coverage factor $K_p$ that produces an interval having level of confidence $p$ when assuming a normal distribution**

| Level of confidence $p$ (%) | Coverage factor $K_p$ |
|---|---|
| 68.27 | 1 |
| 90 | 1.645 |
| 95 | 1.960 |
| 95.45 | 2 |
| 99 | 2.576 |
| 99.73 | 3 |

If, for instance, a normal pdf is assumed for the distribution of values, the levels of confidence shown in Table 1.1 can be associated with the given coverage factors.

### 1.6.3 Evaluating Standard Uncertainty

The GUM identifies two methods for evaluating standard uncertainty, defined as follows [5].

*Type A evaluation*

Method of evaluation of uncertainty by the statistical analysis of a series of observations.

*Type B evaluation*

Method of evaluation of uncertainty by means other than the statistical analysis of series of observations.

***1.6.3.1 Type A Evaluation Method of Uncertainty*** Under the assumption that the result of a measurement has been corrected for all recognized systematic effects (as done in article 3.2.4 of the GUM [5]), the distribution of the measured values can be characterized through the statistical analysis of repeated observations.

In most cases, the best available estimate of the *expectation* (or *expected value*) $\mu_X$ of a random variable $X$, for which $n$ independent observations $X_k$ have been obtained under the same measurement conditions, is the arithmetic mean of such observations, given by

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k. \tag{1.9}$$

Thus, the arithmetic mean $\bar{X}$ obtained from (1.9), can be taken as the numerical value of the measurement result.

The variability of the individual observations, or more specifically, their dispersion about their mean $\bar{X}$, can be characterized by the *standard deviation* $\sigma$ of the underlying

probability distribution of the random variable $X$, that is the positive square root of its variance $\sigma^2$.

An estimate of $\sigma^2$ is provided by the experimental variance of the observations, given by

$$s^2(X_k) = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \bar{X})^2. \tag{1.10}$$

It is well known that the variance of the mean value (1.9) is given by $\sigma^2(\bar{X}) = \sigma^2/n$. In many situations of practical interest, its best estimate is given by

$$s^2(\bar{X}) = \frac{s^2(X_k)}{n}. \tag{1.11}$$

The experimental standard deviation of the mean $s(\bar{X})$, equal to the positive square root of (1.11), quantifies how well $\bar{X}$ estimates the expectation $\mu_X$ of $X$, and may be used as a measure of the uncertainty of $\bar{X}$. It represents the standard uncertainty evaluated by a Type A evaluation method.

The major drawback of the Type A evaluation method is the need for independent repeated observations, whose number should be large enough to ensure that $\bar{X}$ provides a reliable estimate of the expectation $\mu_X$ of the random variable $X$. This condition cannot be always satisfied, mainly because of time constraints or difficulties in ensuring stable measurement conditions throughout the repeated observations.

*Numerical example* To show how standard uncertainty can be evaluated according to the Type A evaluation method, let us suppose to measure a dc voltage $V$ with a five-digit multimeter, using the 10 V range. Let us also suppose that we repeat the measurement procedure several times, under the same measurement conditions.

In this example we consider only $n = 7$ repeated observations, for the sake of brevity. The obtained measured values are reported in Table 1.2.

The measured value is given by the arithmetic mean $\bar{V} = 7.2583$ V of the observations, and, according to (1.11), its standard uncertainty is: $u(\bar{V}) = 0.35$ mV.

**TABLE 1.2 Measured voltage values for a Type A evaluation of standard uncertainty**

| Observation | Measured value [V] |
|---|---|
| 1 | 7.2587 |
| 2 | 7.2595 |
| 3 | 7.2576 |
| 4 | 7.2568 |
| 5 | 7.2583 |
| 6 | 7.2592 |
| 7 | 7.2581 |

***1.6.3.2 Type B Evaluation Method of Uncertainty*** In many practical situations, independent repeated observations of an input quantity $X$, as required by the Type A evaluation, are difficult to perform, for many different reasons, both technical and economical. For the same reasons, the systematic effects on the measurement result can be barely recognized and corrected. Nevertheless, it is still necessary to evaluate the standard uncertainty associated with the measured value. This can be done, as stated by the GUM [5], by scientific judgment based on all of the available information on the possible variability of $X$.

The pool of information may include [5]

- previous measurement data;
- experience with or general knowledge of the behavior and properties of relevant materials and instruments;
- manufacturer's specifications;
- data provided in calibration and other certificates;
- uncertainties assigned to reference data taken from handbooks.

In order to estimate a standard uncertainty value from the available information, this information should allow the operator to make reasonable assumptions on at least

- the shape of the probability density function that may be reasonably assumed to represent the variability of the considered input quantity $X$: normal, uniform, triangular, U-shaped, Weibull, Poisson, …;
- an interval, about the measured value, that is expected to encompass a fraction of the distribution of values that could be reasonably attributed to the measurand with a given coverage probability.

If such assumptions can be made, then the provided interval defines the expanded uncertainty, and the coverage probability assigned to such interval, together with the assumed shape for the probability density function, allow one to compute the coverage factor and, consequently, the standard uncertainty.

Of course, the obtained uncertainty value is strongly dependent on the reliability of the available information, and the way it is used. It is worth noting [5] that the proper use of the pool of available information for a Type B evaluation of standard uncertainty calls for insight based on experience and general knowledge, and is a skill that can be learned with practice.

It is also worth noting that, when performed by an experienced operator, a Type B evaluation of standard uncertainty can be as reliable as a Type A evaluation, especially in a measurement situation where a Type A evaluation is based on a comparatively small number of statistically independent observations.

*Numerical example* To show how standard uncertainty can be evaluated according to the Type B evaluation method, let us consider again the example considered in Section 1.6.3.1 and let us suppose to measure a dc voltage $V$ with a five-digit

multimeter, using the 10 V range. In this case, only one measurement is performed, and let us suppose that the instrument reading is $V_r = 7.2587$ V.

Let us also suppose that the manufacturer accuracy specifications are available and that, for the considered range, they provide an interval of possible values, whose half-amplitude $a$ is given by 0.02% of the reading + 6 digits.[13] Therefore, according to the reading and the selected range, it is:

$$a = 7.2587 \cdot 2 \cdot 10^{-4} + 6 \cdot 10^{-4} = 2.1 \cdot 10^{-3}\text{V}$$

According to the manufacturer, interval $V_r \pm a$ encompasses all values that can be reasonably attributed to the measurand, and its coverage probability is, therefore, 100%.

In order to evaluate the standard uncertainty, it is necessary to make proper assumptions about the way these values distribute. Since the manufacturer does not provide any additional information about the possible distribution of the measurement results, we can refer to the maximum entropy principle and assume that the probability distribution of the values that can be attributed to the measurand is uniform.

Under this assumption, the standard uncertainty is given by

$$u(V_r) = \frac{a}{\sqrt{3}} = 1.2 \cdot 10^{-3}\text{V}.$$

***1.6.3.3 Combining Uncertainty Values Provided by Type A and Type B Evaluation Methods*** In many practical situations, the two methods for standard uncertainty evaluation—Type A and Type B—quantify different contributions to uncertainty.

For instance, in the dc voltage measurement example considered in Sections 1.6.3.1 and 1.6.3.2, the Type B evaluation method, based on the manufacturer specifications, mainly considers the instrument sensitivity to differences between the values assumed by the influence factors during the actual measurement operations and instrument calibration respectively. Indeed, these differences remain usually constant when we repeat the measurement over a time interval of short duration and manifest themselves as systematic effects.

On the other hand, the Type A evaluation method is supposed to consider the effects, on the instrument reading, of variable noise originated inside the instrument or superimposed to the input signal, which manifests itself in terms of random effects during repeated observations.

Therefore, a better estimate of the standard uncertainty, in the given example, should consider both uncertainty contributions, $u_A(V)$ provided by the Type A evaluation method, and $u_B(V)$ provided by the Type B evaluation method.

[13]It is worth noting that this expression quantifies the effect, on the instrument reading, of variations in the influence factors (such as environmental temperature and instrument aging) with respect to calibration conditions. The coefficients in this expression depend on both the instrument sensitivity to the influence factors and the magnitude of their admissible variations.

Since the effects taken into account by the two evaluation methods are due to different physical phenomena, they can be assumed not correlated. Therefore, the two uncertainty contributions can be composed quadratically, and the final uncertainty value is given by

$$u(V) = \sqrt{u_A^2(V) + u_B^2(V)}. \tag{1.12}$$

In the case of the numerical example considered in Sections 1.6.3.1 and 1.6.3.2, the obtained standard uncertainty is: $u(V) = 1.2 \cdot 10^{-3}$ V.

It is worth noting that, in the considered example, the contribution $u_{\mathrm{A}}(V)$ provided by the Type A evaluation method is negligible. This is usually the case in modern digital instrumentation when the input signal has a small amount of superimposed noise and the instrument is working correctly.

### 1.6.4 The Combined Standard Uncertainty

In most cases, a measurand $Y$ is not measured directly, but is determined from $N$ other quantities $X_1, X_2, \ldots, X_{\mathrm{N}}$, through a functional relationship $f$:

$$Y = f(X_1, X_2, \ldots, X_{\mathrm{N}}). \tag{1.13}$$

In general, the input quantities $X_1, X_2, \ldots, X_{\mathrm{N}}$ may be themselves measurands, and may depend on other quantities. For the sake of clarity, and without losing in generality, these quantities will be considered, in the following, as measurands, but no further dependence on other quantities will be considered.[14]

An estimate $y$ of measurand $Y$ can be obtained by applying (1.13) to the estimates $x_1, x_2, \ldots, x_{\mathrm{N}}$ for the input quantities $X_1, X_2, \ldots, X_{\mathrm{N}}$. Thus, the measured value for $Y$ is given by

$$y = f(x_1, x_2, \ldots, x_{\mathrm{N}}). \tag{1.14}$$

Since the input estimates $x_1, x_2, \ldots, x_{\mathrm{N}}$ in (1.14) are measurement results, each of them has an associated standard uncertainty $u(x_{\mathrm{i}})$, $i = 1, \ldots, n$, which is supposed to contribute to the standard uncertainty of the final measurement result $y$.

The problem that must be dealt with now is how to combine the standard uncertainties $u(x_{\mathrm{i}})$ in order to estimate a standard deviation associated to the measurement result $y$. This estimate is called *combined standard uncertainty* and is denoted by $u_{\mathrm{c}}(y)$.

Let us suppose that function $f(\cdot)$ in (1.14) is fairly linear, about the measured value $y$, at least for small deviations of each input quantity $X_{\mathrm{i}}$ about their estimates $x_{\mathrm{i}}$.

[14] In the case any of these quantities depends on other quantities, the method described in the following part of this section can be iteratively applied.

Under this assumption, function $f(\cdot)$ can be approximated by the first order terms of a Taylor series expansion about the expectations $E(X_\mathrm{i}) = \mu_\mathrm{i}$ of the input quantities, thus providing the following deviation about the expectation $E(Y) = \mu_\mathrm{y}$ of the measurand

$$y - \mu_\mathrm{y} = \sum_{i=1}^{N} \frac{\partial f}{\partial x_\mathrm{i}}(x_\mathrm{i} - \mu_\mathrm{i}), \tag{1.15}$$

where $\mu_\mathrm{y} = f(\mu_1, \mu_2, \ldots, \mu_\mathrm{N})$.

From a strict mathematical point of view, the derivatives in (1.15) should be evaluated in the expected values of the input quantities, which are generally unknown. Thus, in practice, they are approximated by the values obtained in the estimates $x_1, x_2, \ldots, x_\mathrm{N}$ of the input quantities.

Moreover, if the magnitude of all derivatives in (1.15) is close to zero, or the non-linearity of $f(\cdot)$ is significant, higher-order terms must be considered in the Taylor series expansion. However, for the sake of simplicity, this case is not considered here.

Equation (1.15) can be squared, providing

$$(y - \mu_\mathrm{y})^2 = \left[\sum_{i=1}^{N} \frac{\partial f}{\partial x_\mathrm{i}}(x_\mathrm{i} - \mu_\mathrm{i})\right]^2, \tag{1.16}$$

which can be expanded as

$$(y - \mu_\mathrm{y})^2 = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_\mathrm{i}}\right)^2 (x_\mathrm{i} - \mu_\mathrm{i})^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\partial f}{\partial x_\mathrm{i}} \frac{\partial f}{\partial x_\mathrm{j}} (x_\mathrm{i} - \mu_\mathrm{i})(x_\mathrm{j} - \mu_\mathrm{j}). \tag{1.17}$$

Equation (1.17) can be rewritten in terms of expectations, taking into account that

- $E[(y - \mu_\mathrm{y})^2] = \sigma_\mathrm{y}^2$ is the variance of $y$;
- $E[(x_\mathrm{i} - \mu_\mathrm{i})^2] = \sigma_\mathrm{i}^2$ is the variance of $x_\mathrm{i}$;
- $E[(x_\mathrm{i} - \mu_\mathrm{i}) \cdot (x_\mathrm{j} - \mu_\mathrm{j})] = \sigma_{\mathrm{i,j}} = \sigma_{\mathrm{j,i}}$ is the covariance of $x_\mathrm{i}$ and $x_\mathrm{j}$.

Taking also into account that standard uncertainty is a standard deviation by definition, we get

$$u_\mathrm{c}^2(y) = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_\mathrm{i}}\right)^2 u^2(x_\mathrm{i}) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\partial f}{\partial x_\mathrm{i}} \frac{\partial f}{\partial x_\mathrm{j}} u(x_\mathrm{i}, x_\mathrm{j}), \tag{1.18}$$

where $u(x_i, x_j) = u(x_j, x_i)$ is the estimated covariance associated with $x_i$ and $x_j$. The degree of correlation between $x_i$ and $x_j$ is characterized by the estimated *correlation coefficient*, defined as

$$r(x_i, x_j) = \frac{u(x_i, x_j)}{u(x_i)u(x_j)}, \tag{1.19}$$

where $r(x_i, x_j) = r(x_j, x_i)$, and $-1 \leq r(x_i, x_j) \leq +1$.

By replacing (1.19) into (1.18), we get

$$u_c^2(y) = \sum_{i=1}^{N} \left( \frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i) u(x_j) r(x_i, x_j) \tag{1.20}$$

that is also known as the general formulation of the *uncertainty propagation law* [5]. $u_c(y)$ is the *combined standard uncertainty* of $y$.

It can be immediately recognized that, if the estimates $x_i$ and $x_j$ are independent, we have $r(x_i, x_j) = 0$. If this holds for all input variables $x_i$, Equation (1.20) can be simplified as

$$u_c^2(y) = \sum_{i=1}^{N} \left( \frac{\partial f}{\partial x_i} \right)^2 u^2(x_i). \tag{1.21}$$

According to all the above mathematical derivations, Equations (1.20) and (1.21) allows one to evaluate the standard uncertainty of the result of a measurement, when the measurand is not measured directly, but is determined from a number of other quantities, and the standard uncertainties of their estimates are known.

It is worth noting that the uncertainty propagation law cannot provide any information about the distribution of probability associated with the measured value $y$, and hence the obtained standard deviation $u_c(y)$ cannot be used to identify intervals with given coverage probability about $y$.

To do so, we have to refer to the Central Limit Theorem. This theorem states that, if function $f(\cdot)$ in (1.13) is linear, none of the considered $\sigma_i^2$ variances dominates over the others, and $N \to \infty$, then the probability density function of $Y$ exhibits a normal behavior, no matter on the shapes of the single probability density functions associated with $X_i$.

Of course, if $N$ is a finite value, the probability density function of $Y$ has only an approximated normal behavior, and the larger $N$, the better the approximation.

According to all above considerations, the following important conclusion can be drawn.

The uncertainty propagation law (1.20) (or its simplified version (1.21) in case of uncorrelated input quantities) can be used to evaluate the combined standard uncertainty of a measurement result when the measurand is determined from $N$ other quantities, and the obtained combined standard uncertainty can be used to obtain

an expanded uncertainty with given coverage probability if the following conditions are met.

1. Function $f(\cdot)$ in (1.14) is fairly linear about $y$, at least for small deviations of each input quantity $X_i$ about its estimate $x_i$.
2. The partial derivatives of $f(\cdot)$ with respect to all input quantities exist and at least one of them is not nil.
3. The effect of none of the input variables prevails over the others.
4. The number $N$ of considered input quantities is high enough (theoretically infinite, though, in practice, values of $N \geq 5$ are usually high enough).

If all above conditions are satisfied, the Central Limit Theorem can be applied and a normal probability distribution can be assigned to the estimate $y$ of measurand $Y$.

The above conditions cannot be always fully satisfied. The application of the uncertainty propagation law may result in an incorrect estimate of the combined standard uncertainty and, more critically, in an incorrect estimate of the coverage probability assigned to the expanded uncertainty.

In order to overcome this problem, the BIPM issued, in 2008, a Supplement to the GUM [6] that recommends the use of Monte Carlo simulations to estimate the probability distribution of $y$ starting from the probability distributions of the input quantities $x_i$ whenever the uncertainty propagation law cannot be used.

*Numerical examples* Let us consider a dc power measurement, and let us suppose that we can measure voltage $V$ and current $I$ with two multimeters. Of course, the desired power value is obtained as: $P = V \cdot I$.

Let us also suppose that $V_m$ is the measured value for voltage $V$ and $I_m$ is the measured value for current $I$, and that the only available information to evaluate standard uncertainty comes from the manufacturer specifications. Therefore, a Type B evaluation method has to be applied, as shown in the numerical example in Section 1.6.3.2.

Let us suppose that the readings and the associated standard uncertainty values (evaluated as shown in Section 1.6.3.2) are

- $V_m = 8.0125$ V (in the 10 V range), with $u(V) = 1.3 \cdot 10^{-3}$ V;
- $I_m = 50.105$ mA (in the 100 mA range), with $u(I) = 2.6 \cdot 10^{-5}$ A;

The measured power value is, therefore, $P_m = 0.4015$ W.

Let us now suppose that the two multimeters, used to measure voltage and current, are different and so their calibration histories. Under these assumptions, the two measured values $V_m$ and $I_m$ can be assumed to be uncorrelated, and Equation (1.21) can be used to evaluate the combined standard uncertainty.

It is

$$\frac{\partial P}{\partial V} = I = 50.105 \text{ mA}, \tag{1.22}$$

when evaluated for $I = I_m$, and

$$\frac{\partial P}{\partial I} = V = 8.0125 \text{ V}, \tag{1.23}$$

when evaluated for $V = V_m$.

By replacing the above evaluated standard uncertainties $u(V)$ and $u(I)$, and the values provided by (1.22) and (1.23) into (1.21), we get: $u_c(P) = 2.2 \cdot 10^{-4}$ W.

Let us now suppose to employ a single, two-channel multimeter to measure both $V$ and $I$. In this case, a complete and positive correlation can be assumed between the measured values $V_m$ and $I_m$, and hence equation (1.20) with $r(V, I) = 1$ shall be used to evaluate the combined standard uncertainty.

Considering, for the sake of simplicity, the same numerical values as above, we get: $u_c(P) = 2.7 \cdot 10^{-4}$ W. Therefore, in this example, using two independent instruments decreases uncertainty and, therefore, improves the *quality* of the measurement result.

However, this is not a general conclusion. To prove this, let us suppose to measure the resistance of the load connected to the measurement section considered for power measurement. The desired resistance value is: $R = V/I$.

Thus, the instrument readings $V_m$ and $I_m$ and the standard uncertainties $u(V)$ and $u(I)$ are the same as above. The measured resistance value is: $R_m = 159.91\ \Omega$.

The partial derivatives become

$$\frac{\partial R}{\partial V} = \frac{1}{I} = 19.96 \text{ A}^{-1}, \tag{1.24}$$

when evaluated for $I = I_m$, and

$$\frac{\partial R}{\partial I} = \frac{-V}{I^2} = -3191.58\ \Omega\text{A}^{-1}, \tag{1.25}$$

when evaluated for $V = V_m$ and $I = I_m$.

If we use two different, independently calibrated multimeters for voltage and current measurements, then applying (1.21) with the above determined values returns $u_c(R) = 86$ m$\Omega$.

If, on the other hand, we use the same multimeter for both voltage and current measurements, then using (1.20) and the above determined values we have $u_c(R) = 58$ m$\Omega$. In this case, due to the different relationship between the measurand and the input quantities, a positive correlation produces an uncertainty reduction.

## 1.7 CONCLUSIONS

This chapter covered the very fundamentals of the measurement science. First of all it showed how measurement has been an important part of the knowledge processes, and how its importance has evolved along the centuries, as the scientific and human need for knowledge has evolved.

It showed how the measurement activity can be modeled in terms of different descriptive and experimental processes that concur to generate the measured value. A strict analysis of the defined model gave evidence of the different sources of imperfections that cause the measured values to differ from the measurand value. This same analysis proved that the measurand value remains unknown and unknowable, and proved that a quantitative estimate of the deviation between the measured and the measurand values is definitely necessary for every practical measurement application.

The modern uncertainty concept was analyzed as the presently most effective way to provide this quantitative estimate, and the standard recommendations on how to implement this concept were presented.

It is worth noting, as a final remark, that the uncertainty concept, as defined in Section 1.5, has general validity and provides a scientifically sound answer to the need for a quantitative estimate of how "good" a measurement result is. On the other hand, the practical implementation recommended by the presently available standard documents [5, 6] and presented in Section 1.6 has some limitations, mainly due to the considered mathematical framework: the theory of probability.

Different solutions have been recently proposed (and well summarized in Reference 7) to overcome these limitations, by referring to the more general mathematical framework of the theory of evidence. Though the discussion of new proposals is beyond the scope of this book, we wish to conclude this chapter as we started it: with a clear indication that the measurement science, as every other science, is always evolving, and today problems should be a motivation for searching new solutions, even if they may clash against well assessed theories, methods, and tools.

## FURTHER READING

Several books and articles can be found that cover the topics that have been synthetically described in this chapter. In particular, the following books are related to the history of measurement.

Alder, K. *The Measure of All Thing*, Free Press, New York (2002).

Keithley, J.F. *The Story of Electrical and Magnetic Measurements*, IEEE Press, Piscataway, NJ (1999).

Zupko, R. *Revolution in Measurement: Western European Weights and Measures Since the Age of Science*, American Philosophical Society, Philadelphia, PA (1990).

Klein, H.A. *The Science of Measurement: A Historical Survey*, Dover Publications, Mineola, NY (1988).

Favre, A. *Les origines du système métrique*, Les Presses Universitaires de France, Paris (1931).

The following books and articles are related to the measurement model and uncertainty evaluation.

Rossi, G. B. *A Probabilistic Theory of Measurement with Applications*, Springer, New York (2014).

Sydenham, P. H. *Handbook of Measurement Science*, Vol.1, Theoretical fundamentals, John Wiley & Sons Inc, Hoboken, NJ (1992).

Sydenham, P. H., Hancock, N. H., Thorn, R. *Introduction to Measurement Science and Engineering*, John Wiley & Sons Inc, Hoboken, NJ, (1992).

Finkelstein, L. Measurement and instrumentation science – an analytical review, *Measurement*, **14** (1), 3–14 (1994).

Finkelstein, L. State and advances of general principles of measurement and instrumentation science, *Measurement*, **3** (1), 2–6 (1985).

Giordani, A., Mari, L. Measurement, models and uncertainty, *IEEE Transactions on Instrumentation and Measurement*, **61** (8), 2144–2152 (2012).

Ferrero, A., Salicone, S. Uncertainty: only one mathematical approach to its evaluation and expression?, *IEEE Transactions on Instrumentation and Measurement*, **61** (8), 2167–2178 (2012).

## REFERENCES

1. OIML. *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM)*, 3rd edition, (2007). Available at BIPM http://www.bipm.org/en/publications/guides/vim.html
2. Keithley, J.F. *The Story of Electrical and Magnetic Measurements*, IEEE Press, Piscataway, NJ (1999).
3. Di Pasquale, G., Marchis, V. Some considerations about the Roman Pes, *Nuncius*, **11**, 669–675 (1996) (in Italian).
4. Sydenham, P.H. *Handbook of Measurement Science*, *Vol.1: Theoretical Fundamentals*, John Wiley & Sons, Inc., Hoboken, NJ (1992).
5. BIPM JCGM 100:2008. *Evaluation of measurement data – Guide to the expression of uncertainty in measurement* (GUM), GUM 1995 with minor corrections (2008). Available at BIPM http://www.bipm.org/en/publications/guides/gum.html
6. BIPM JCGM 101:2008. *Evaluation of measurement data – Supplement 1 to the Guide to the expression of uncertainty in measurement – Propagation of distributions using a Monte Carlo method*. Available at BIPM http://www.bipm.org/en/publications/guides/gum.html
7. Salicone, S. *Measurement Uncertainty – An Approach via the Mathematical Theory of Evidence*, Springer, New York (2007).
8. Sydenham, P.H., Hancock, N.H., Thorn, R. *Introduction to Measurement Science and Engineering*, John Wiley & Sons, Inc., Hoboken, NJ (1992).
9. Luce, R.D., Krantz, D.H., Suppes, P., Tversky, A. *Foundations of Measurements*, Vol.3: Representation, Axiomatization, and Invariance, Dover Publications, Mineola, NY (2006).

## EXERCISES

**1.1** Why do we measure?

**(a)** To provide a basis for the exchange of goods and services

**(b)** To encourage cooperative activities

**(c)** To control, manage, and improve processes

**(d)** All of the above

**1.2** Typically, any measurement result depends on

**(a)** The value of the measurand

**(b)** The measurand value and other properties pertaining to the measurement context

**(c)** The measurand value and the measurement unit

**(d)** The influence quantities

**1.3** The main components of the measurement context are

**(a)** The environment and the operator

**(b)** The measurand and the measured system

**(c)** The measured system, the measurement system and the environment

**(d)** The measurement system and the influence quantities

**1.4** The main types of measurement uncertainty sources are

**(a)** The definitional and interactional uncertainty

**(b)** The measurement system uncertainty and the operator inexperience

**(c)** The acquisition system uncertainty and instrumental noise sources

**(d)** The definitional, interactional and instrumental uncertainties

**1.5** The measurement uncertainty mainly derive by

**(a)** The definition of the measurand

**(b)** The uncertainty of the employed instruments

**(c)** The uncertainty associated with the adopted model of the context

**(d)** The operator and the empirical procedures performed

**1.6** An optimal model of the measurement context allows to achieve

**(a)** Meaningful information about the measurand with minimum complexity

**(b)** Useful information about the measurand

**(c)** The maximum information about the measurand

**(d)** Enough information about the measurand

**1.7** If the input signal has negligible superimposed noise, the uncertainty of the result provided by an electronic instrument is typically

**(a)** Mainly of Type B

**(b)** Mainly of Type A

**(c)** Negligible if the instrument is correctly used

**(d)** Both of Type A and Type B

**1.8** Generally, any measurement context model should contain

**(a)** The variable that model the measurand and its interactions with the main influence properties

**(b)** The variable that model the measurand, the set of variables modeling the main influence properties and the related allowed ranges, the relationships describing the associated interactions, the set of variables modeling the secondary influence properties and the related allowed ranges

**(c)** The model of the measurand and of the measurement system

**(d)** The variables that model the measurand and the main and secondary influence properties

**1.9** Which of the following statements about measurement uncertainty evaluation is incorrect?

**(a)** Indications provided by the GUM should be disregarded

**(b)** Monte Carlo procedures can be a useful tool to study uncertainty propagation

**(c)** The measurement model inputs should take into account uncertainty arising from transducers

**(d)** The Law of Propagation of Uncertainty suggested by the GUM can be difficult to apply

**1.10** Which of the following statements is true

**(a)** In nominal scales, the classes containing the measurement result must be labeled with numbers

**(b)** In ordinal scales, addition and subtraction are allowed

**(c)** In interval scales, the gap between two different classes is significant

**(d)** In ratio scales, the zero is defined by shared conventions

**1.11** In any measurement, a lower bound to measurement uncertainty is given by

**(a)** Definitional uncertainty

**(b)** Combined uncertainty

**(c)** Instrumental and interaction uncertainty

**(d)** The effect of environmental factors

**1.12** It is enough that any measurement result (explicitly or implicitly) provides the following information

**(a)** Measurement value and uncertainty

**(b)** The pdf of the measurement value

**(c)** Measurement value and measurement unit

**(d)** Measurement value, uncertainty, measurement scale and limits of validity of the information provided

**1.13** The measurement error concept is obsolete and not used any longer to qualify a measurement result because

**(a)** Modern instruments provide the true value of the measurand

**(b)** We do not know how to process it when indirect measurements are considered

**(c)** It refers to the true value, and the true value of the measurand is always unknown and unknowable

**(d)** It does not take into account random effects

**1.14** Standard uncertainty is defined as

**(a)** An interval about the true value of the measurand

**(b)** An estimator of the standard deviation of a probability distribution

**(c)** A confidence interval about the measured value with a given level of confidence

**(d)** An interval about the measured value, within which the value of the measurand is supposed to lie with full certainty

**1.15** Type-B evaluation method of standard uncertainty is used when

**(a)** A systematic effect is present

**(b)** It is not possible to evaluate uncertainty by means of a statistical analysis of a series of observations

**(c)** A great number of experimental data is available

**(d)** A calibration certificate is not available

**1.16** The expanded uncertainty provides a confidence interval about the measurement result with a given confidence level if

**(a)** Explicit or implicit assumptions about the probability distribution of the measurement result can be done

**(b)** A suitable coverage factor can be considered

**(c)** The probability distribution associated with the measurement result is normal

**(d)** The probability distribution associated with the measurement result is uniform

**1.17** The measurement representational condition states that

**(a)** Each measurement result must contain the value of measurement uncertainty

**(b)** Each measurement result must be represented by a set of numbers

**(c)** Any empirical relations between measurands must be mapped to abstract relations between symbols

**(d)** Any empirical relations between measurands must be mapped to arithmetic operations

**1.18** The final goal of a measurement process is

**(a)** Increase the personal knowledge of the measurement expert

**(b)** Provide a quantitative input to a decision-making process

**(c)** Define the experimental set-up

**(d)** Attain an uncertainty value as low as possible

**1.19** Legal metrology covers

**(a)** That part of the measurement science that is used by lawyers in trials

**(b)** Technical expertise to issue laws related to measurements

**(c)** Measurement methods and instruments employed in economical transactions and the application of sanctions

**(d)** Quality measurements to evaluate the performance of courts

**1.20** The coverage factor is

**(a)** A multiplier of standard uncertainty that provides expanded uncertainty

**(b)** The coverage probability associated with the interval defined by the expanded uncertainty

**(c)** A factor that takes always numerical value equal to 2 if the probability distribution assigned to the possible values of the measurand is normal

**(d)** A factor that takes always numerical value equal to $1/\sqrt{3}$ if the probability distribution assigned to the possible values of the measurand is normal