urn:nbn:de:gbv:ilm1-2011imeko-009:1

Joint International IMEKO TC1+ TC7+ TC13 Symposium
August 31st – September 2nd, 2011, Jena, Germany
urn:nbn:de:gbv:ilm1-2011imeko:2

MEASUREMENT MODELLING: FOUNDATIONS AND PROBABILISTIC APPROACH

Giovanni Battista Rossi

Facoltà di Ingegneria, Università degli Studi di Genova, Genova, Italy

Abstract –The role of modelling in science and in measurement is first discussed. This is followed by a section on modelling in measurement, either for founding the measurability of a characteristic or for performing a measurement. A general probabilistic model of the measurement process, capable of addressing calibration, static and dynamic measurement, is then examined. To demonstrate the application of the model, a working example is developed in detail. Lastly, the current status of measurement modelling is reviewed and commented.

Keywords : measurement science, model, probability

1. FOUNDATIONS OF MEASUREMENT MODELLING

1.1. Background: realism versus instrumentalism in science

The notion of model is central in science and engineering since it is related to the epistemological value of scientific theories and the effectiveness of engineering activities.

A controversy about the value of scientific theories accompanied the birth of modern science. When they first appeared, Galileo's theories were challenged as being potentially aversive for the then dominating view of the world which was based on the authority of the Scriptures and their authentic interpretation by the religious authorities [11, 25][1].

In defence of the new approach, Galileo claimed that scientific investigation set out to find out how nature operates: to understand "the way the heavens go" rather than "the way to go to heaven". This is possible – in his view – since the book of nature is written in the same language as science, that is, a mathematical language whose "characters are triangles, circles and other geometrical figures".

By contrast, his opponent, Cardinal Bellarmino, made a clear distinction between the quality of the knowledge provided by the holy scriptures and that provided by the new theories. These are developed *ex suppositione* and are based on a mathematical description of things that is instrumental to predicting events, but incapable of grasping the real basis of what exists as this may only be revealed by the holy scriptures.

Galileo's and Bellarmino's arguments are the precursors of two views that are still debated today, *realism* versus *instrumentalism* [12, 13]. Both have evolved over time, nonetheless. Modern instrumentalism, which is quite popular especially in the area of computer science, is somewhat related to Anglo-American pragmatism. Yes, models are essentially good for making predictions and for supporting decision-making processes, but *this is what really matters* – a pragmatist would argue. In fact, modelling has grown greatly as regards the development of computer technology, and related speculation areas include information and system theory.

On the other hand, today Galileo's realism sounds rather extreme although there are still scientist that adhere to it. He believed that science can give a "true" description of reality but, as we are now aware, especially after the criticisms of Kuhn, in a given historical period scientific theories are strongly influenced by leading *paradigms* [18, 25]. When such paradigms change, science undergoes profound transformations which Kuhn called scientific revolutions. So how can we believe that our knowledge in a given period is *the* true one? This *strong*, so to speak, *realism* is hard to support. Yet, if models are good at making predictions and supporting decision-making processes - and most people, at least in industrialised countries, now agree on this – they must somehow "bite into reality", as Barone used to say [2]. This is a form of *moderate realism* which I personally agree with.

In this perspective, I propose to consider a model as *an abstract system setting out to represent, to some extent and from a certain viewpoint, a real system* (or a class of real systems). I do not make a sharp distinction between a model and a theory, rather I consider a theory as a very general model. Simply put, a model (or a theory) is "good" if it succeeds in its task, which consists in giving a proper representation of a piece of reality. The ultimate decision on the acceptance of a model is up to the scientific community and is subject to the dialectics of that community and may change over the course of time.

In the process of controlling the consistency of a model with reality, measurement comes into play.

[1] References are linked in chronological order, to give a feeling of the development of the subject in the last twenty years.

Measurement thus enjoys a peculiar position in science [6, 14] since it is, on the one hand, the key tool for assessing the validity of scientific models, yet, being a scientific activity, it is subject to modelling itself. This peculiar position is the reason for the intrinsic difficulty in measurement modelling which will emerge at a later stage[2]. Before proceeding, we must establish some language. I will call *characteristic* what we (want to) measure, whilst I will use the term *property* for denoting the empirical properties of a characteristic that allows its measurement; a measurable characteristic will be called *quantity*. I will use the term *object* for conventionally denoting what carries (expresses, manifests) the characteristic under investigation, irrespective of whether it is a material object, an event or even a person. Furthermore, since objects manifest the characteristic of interest in different ways (levels, degrees), I will call the way in which an object manifests a characteristic a *state*.

1.2. Foreground: measurement modelling

Measurement modelling concerns modelling both *the measurand* and *the measurement process*. In both cases three types of models can be considered: concerned with the internal properties of a quantity, dealing with influence quantities or related to derived quantities.

Concerning the *internal properties* of a characteristic, we should remember that, according to *representational theory*, measurement may be seen as the mapping of objects into numbers (measures) [6]. Measures reproduce in a numerical domain the relations that the objects empirically hold in respect of the characteristic under investigation. So if object *a* is longer that object *b*, its length measure, *m(a)*, will be greater than that of *b*, *m(b)*, and so on. Consequently, a characteristic is *measurable*, i.e., it is a *quantity* in our language, if it satisfies proper *empirical properties*. In the case of length, for example, the required empirical properties are order (one object being longer than another) and additivity (e.g. similar blocks may be piled to form a block whose length is the sum of the length of individual blocks). Typical solutions, as concerns length, are the geometrical models of "segment" or "pair of parallel planes" which may be used for objects, such as rods or blocks, respectively. The models guarantee measurability in that they incorporate the property of order and the operation of addition. As long as the model can give a good description of the empirical reality, similar properties may be assumed for the real objects (rods, blocks) and for the corresponding abstract ones (segments, parallel planes).

In other cases, we do not have a well-established model that can guarantee measurability. This is often the case, for example, when perception is involved. In such cases, measurability can be proved by empirically assessing that the required properties are satisfied. Here the representational framework itself constitutes the reference model [27, 28].

Modelling the effect of *influence quantities* is of major concern when designing a specific measurement task. In general, a real object can be characterised by a set of quantities, linked by equations. Even though measurement aims at just one of them, the others must also be accounted for if they affect the measuring process. The model here is essential for evaluating whether the effect of influence quantities is acceptable for the target uncertainty. If it is not, the model may support the design of corrections. Suppose, for example, that we are considering length, *l*, and wish to account for the influence of temperature, *t*. Consider two rods, *a, b,* and suppose that if they are at the same temperature, *a* is longer than *b*, that is $a \succ b$. Suppose now that the temperature may differ from one object to another. Then we must use the model to evaluate whether $b \succ a$ can be observed due to temperature variations. The model will enable us to prevent this from happening and/or to evaluate the uncertainty related to this thermal effect.

Lastly, *derived quantities* are those that are measured on the basis of some "natural laws" that link them to other quantities. For example, density, ρ, may be indirectly measured by measuring mass, *m*, and volume, *V*, through the relation $\rho = m/V$. In the case of loudness, *L*, we may measure its related physical intensity, *I*, and then apply a psychophysical law, such as Stevens law, $L = \alpha I^{\beta}$.

The second reason why we use a measurement model is to perform measurements. In this case, we must consider not only the measurand object but also the measuring system and the interaction between the two. The measurement process may be seen as the concatenation of two phases [4, 20][3],

- *Observation*, in which the object to be measured is entered to the measuring system that produces indications, and
- *Restitution*, where the measurement value is obtained on the basis of instrumental indications thanks to a previously established calibration function.

If *x* denotes the unknown value of the measurand the quantity to be measured, *a* the object to be measured

[2] In some academic environments a division has historical aroused between "modellers" and "experimenters", the latter including measurement experts. This division has confined measurement in the purely experimental disciplines, with noxious consequences, including a biased attitude in the funding of research projects. I think, on the contrary, that theoretical and modelling challenges in measurement are not less demanding than in other, apparently more theoretically-oriented, disciplines.

[3] The term "restitution", which I learnt from my mentor, Professor Michelini, is routinely used in photogrammetry. "Reconstruction" (Morawski) and "gauging" (Aumala & Sobolev) are synonymous. "Measurement evaluation" is sometimes used (Woeger) with a similar but more vague meaning.

and $\hat{x}$ the measurement value, the overall measurement process may be seen either as a mapping from things to numbers, $a \mapsto \hat{x}$, or from numbers to numbers, $x \mapsto \hat{x}$. The latter is the usual perspective in science and engineering and the one we will also develop here. Yet in measurement, the former is also important and needs to be considered, especially from a foundational viewpoint. Further on, I will concentrate on the measurement process and will present and discuss a general probabilistic model. But before doing that, we must briefly discuss the notion of probability in this modelling perspective.

1.3. Notes on the philosophy of probability

The discussion on the nature of probability was part of nineteenth century epistemological debate [16]. According to Hacking [7], probability may be understood either as a *relative frequency* or as a *degree of belief*. This is also related to seeing the objects of probability as *events* or *statements* and the two perspectives are sometimes convertible: if I am interested in the probability that tomorrow it will rain, I may see "tomorrow it will rain" as an event, that may happen or not, or as a statement, that may be true or false.

Historically, both positions were pursued up to their extreme consequences. Von Mises, for example, fully investigated the frequentistic approach which admitted the possibility of a probabilistic estimation only when it refers to a "collective"[4] of realisations [16]. On the other hand, De Finetti, reached a subjectivistic vision. Pursued up to its extreme consequences, this approach leads to considering probabilistic statements as reflecting the vision of a single person, and, as such, having limited scientific value since scientific knowledge is understood to be intersubjective.

In my opinion, these two views are related to the general question in science of whether the properties that we attribute to objects have an *ontic* or an *epistemic* character [14]. Ontic means that they are inherent in the object, epistemic means that they strongly dependent on our cognitive categories. Again, I think that the notion of model is essential for solving this dilemma. My view is illustrated in Figure 1.

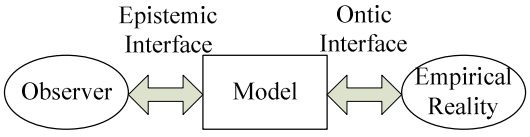

Fig. 1. The role of the model in the scientific method of observing the world.

In science we do not look at things directly but through the intermediation of a model [11]. This model interfaces with us on one side and with the empirical reality on the other. As long as it is a model of ours, it must comply with our cognitive categories: hence its epistemic character. Yet, as it "bites into reality" it must take on some ontic aspects. So these two perspectives are not, in my opinion, irreconcilable: rather they are two faces of the same coin, in general not separable.

Coming back to probability, as this is used to develop probabilistic models, it belongs to this pattern. Probabilistic models are those that are expressed in terms of probabilistic relations and/or variables, and, as models, they usually give a description of things which is both epistemic and, to some extent, ontic. Thus, concerning the "nature" of probability, I basically regard it as a *primitive notion* – everyone intuitively understands the term "probability" as they do terms such as "line" or "plane" – *mathematically characterised by a set of axioms*. In science, probability is used as a kind of "logic" that allows models to be developed and inferences made, whose validity is subjected to evaluation by the scientific community, as is the case for any scientific construct. The adherence of the scientist who formulated the model to some credo, Bayesian, frequentistic or any other, does not, in my opinion, add to or subtract from the validity of the model.

2. A PROBABILISTIC MODEL OF THE MEASUREMENT PROCESS

2.1. General framework

We are now ready to present and discuss the proposed probabilistic model of the measurement process [21]. More in general, we consider all the quantities and parameters involved as *vectors*. Let $\mathbf{x}$ denote the measurand, $\mathbf{y}$ the instrument indications and $\boldsymbol{\theta}$ the influence quantities or parameters.

Observation may then be described by the conditional distribution

$$p(\mathbf{y} \mid \mathbf{x}, \boldsymbol{\theta}). \tag{1}$$

This distribution is *the core of the model* and, in practice, its formulation is the main modelling task. We will thus call it the *characteristic distribution* of the measurement process. Once it has been obtained, what follows is simply calculation. Restitution may be seen as the probabilistic inversion of observation and may be obtained by the Bayes-Laplace rule, followed by probabilistic de-conditioning from influence quantities/parameters:

$$p(\mathbf{x} \mid \mathbf{y}) = \int_{\Theta} \frac{p(\mathbf{y} \mid \mathbf{x}, \boldsymbol{\theta})}{\int_{\mathrm{X}} p(\mathbf{y} \mid \mathbf{x}, \boldsymbol{\theta}) \mathrm{d}\mathbf{x}} p(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta}. \tag{2}$$

A probabilistic description of the overall measurement process may be obtained by combining the above two transformations:

[4] A "collective" is an infinite series of outcomes in which each attribute has a limiting relative frequency that is insensitive to place selection.

$$p(\hat{\mathbf{x}}\,|\,\mathbf{x}) = \int_{\mathbf{Y}} \delta\big(\hat{\mathbf{x}} - \mathrm{E}(\mathbf{x}\,|\,\mathbf{y})\big)\Big(\int_{\Theta} p(\mathbf{y}\,|\,\mathbf{x},\boldsymbol{\theta})\, p(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}\Big)\,\mathrm{d}\mathbf{y}, \tag{3}$$

Where δ is the Dirac-delta operator.

Let us now apply this general framework to calibration and to static and dynamic measurement.

2.2. Calibration

Calibration is an experiment performed in order to characterise a measuring system and provides vital information for the measurement processes that are based on it. In terms of our model, we may simply say that calibration sets out to obtain the characteristic distribution (1). This may be done – and usually is – by assuming a parametrical expression for the distribution (1) and by estimating the relevant parameters. Calibration may be either static (steady-state) or dynamic; here we shall focus on the former. The experiment consists in inputting the system with a series of standard objects that materialise a series of values of the quantity, and in recording the corresponding steady-state indications of the instrument. The experimental data thus consist in a series of pairs, $\{(x_i, y_i), i = 1,...,n\}$. The reference distribution is as in (1), *where* $\boldsymbol{\theta}$ *is now the vector of the parameters to be estimated*. Since in this case $\mathbf{x}$ and $\mathbf{y}$ are known, we *may apply the Bayes-Laplace rule to obtain* $\boldsymbol{\theta}$:

$$p(\boldsymbol{\theta}\,|\,\mathbf{x},\mathbf{y}) = \frac{p(\mathbf{y}\,|\,\mathbf{x},\boldsymbol{\theta})\, p(\boldsymbol{\theta})}{\int_{\Theta} p(\mathbf{y}\,|\,\mathbf{x},\boldsymbol{\theta})\, p(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}}, \tag{4}$$

Where $p(\boldsymbol{\theta})$ may be a vague prior [1]. Again, the main modelling task is to obtain a proper expression for the characteristic distribution. This may be done by assuming an input-output model for the measuring system, on the basis of our knowledge of its behaviour. For a single static measurement, this may be expressed in quite general terms as

$$y = f(x, \boldsymbol{\theta}, w), \tag{5}$$

Where w is a random variable accounting for observation noise. Assuming that the realisations of this noise are uncorrelated along the calibration process, one obtains

$$p(\mathbf{y}\,|\,\mathbf{x},\boldsymbol{\theta}) = \prod_i \int_{\mathbf{W}_i} \delta\big(y_i - f(x_i,\boldsymbol{\theta},w_i)\big)\, p(w_i)\,\mathrm{d}w_i, \tag{6}$$

Which allows us to solve our problem by applying formula (4). An explicit analytical example will be provided at a later stage.

2.3. Static measurement

Static measurement is a measurement in which the measurand is assumed to be constant, at least for a time span that includes the duration of the experiment. Observation is based on the steady-state indication of the measuring system; it may be single or repeated. In the case of repeated observations, we may observe variations that are generally attributed to random noise. The final result is usually obtained by averaging the indications and this method is considered to give a better quality than the one based on a single observation, since the averaging operation allows the noise to be filtered to some extent. Variations are usually assumed to be uncorrelated, especially if sufficient time elapses between subsequent observations and the initial conditions of the measuring device are each time restored. In our model, the reference description is still provided by the characteristic distribution (1), where the measurand is now a scalar, x, and influence parameters define the behaviour of the device and the effect of influence quantities. For some parameters, we must totally rely on information obtained, for example, during calibration or, in general, outside the measurement experiment. For others, we may also obtain information from the measurement process itself, from repeated observations. An example of the latter is the variance of the random variations which may be estimated during the measurement process.

We will thus distinguish between unobservable and observable parameters, and denote the former by $\boldsymbol{\theta}$, the latter by $\boldsymbol{\sigma}$. The input-output model for the single observation may be still expressed by formula (5)[5].

If we also assume that the random variable describing the random variations is always the same, denoted by w, the characteristic distribution may be obtained by

$$p(\mathbf{y}\,|\,\mathbf{x},\boldsymbol{\theta},\boldsymbol{\sigma}) = \prod_i \int_{\mathbf{W}} \delta\big(y_i - f(x,\boldsymbol{\theta},w)\big)\, p(w\,|\,\sigma)\,\mathrm{d}w, \tag{7}$$

Whilst restitution is provided by

$$p(x\,|\,\mathbf{y}) = \int_{\Theta}\int_{\Sigma} \frac{p(\mathbf{y}\,|\,x,\boldsymbol{\theta},\sigma)}{\int_{\mathbf{X}} p(\mathbf{y}\,|\,x,\boldsymbol{\theta},\sigma)\,\mathrm{d}\mathbf{x}}\, p(\sigma\,|\,\mathbf{y})\, p(\boldsymbol{\theta})\,\mathrm{d}\sigma\mathrm{d}\boldsymbol{\theta}. \tag{8}$$

Note the difference between $\boldsymbol{\theta}$ and σ: for the former, the reference distribution is $p(\boldsymbol{\theta})$, which does not depend upon $\mathbf{y}$, since for $\boldsymbol{\theta}$ we do not obtain information from the indications $\mathbf{y}$; for σ, instead, we use the distribution $p(\sigma\,|\,\mathbf{y})$, which is a "poste-

[5] The parameters in vector $\boldsymbol{\sigma}$ appear at a higher level than that expressed by function f and, for this reason, they are sometimes called hyper-parameters.

rior" distribution that follows the acquisition of $\mathbf{y}$. This case will be also treated in the illustrative example in the next section.

2.4. Dynamic measurement

Dynamic measurement sets out to measure the variations in the values of a quantity over time [4, 19, 23]. Such measurement is typically affected by the dynamic characteristic of the measuring device and by noise in the measuring chain. Although continuous variations in time are often assumed for the quantities involved, a discrete-time representation is also appropriate, provided that the sampling rate is sufficiently high (Nyquist condition). Therefore the measurand is now a vector, $\mathbf{x}$, that collects the values assumed by the quantity x_t in a time interval $T = N\Delta t$, where Δt is the sampling interval and $\mathbf{y}$ is the corresponding vector of instrument indications. Let us discuss the structure of the characteristic distribution in this case. Let *t* be a discrete time index and let us define

$$\begin{aligned} x^t &\triangleq (x_1, x_2, \ldots x_t), \\ y^t &\triangleq (y_1, y_2, \ldots y_t). \end{aligned} \tag{9}$$

The characteristic distribution (1) may be factorised as (ignoring, for the sake of simplicity, the dependence on $\boldsymbol{\theta}$):

$$\begin{aligned} &p(\mathbf{y} \mid \mathbf{x}) = \\ &p(y_1 \mid \mathbf{x}) \cdot \ldots p\left(y_t \mid y^{t-1}, \mathbf{x}\right) \cdot \ldots p\left(y_N \mid y^{N-1}, \mathbf{x}\right). \end{aligned} \tag{10}$$

Furthermore, given the *causality* condition (the cause must precede the effect), the indication at instant *t* does not depend on the measurand at instants following *t*. Thus we obtain

$$\begin{aligned} &p(\mathbf{y} \mid \mathbf{x}) = \\ &p(y_1 \mid x_1) \ldots p\left(y_t \mid y^{t-1}, x^{t-1}\right) \ldots p\left(y_N \mid y^{N-1}, x^{N-1}\right). \end{aligned} \tag{11}$$

Proceeding further, we describe the measuring system using the following, very general, *stochastic model*:

$$\begin{aligned} \mathbf{z}_{t+1} &= g(\mathbf{z}_t, \boldsymbol{\xi}_t, x_t), \\ y_t &= f(\mathbf{z}_t, w_t), \end{aligned} \tag{12}$$

Where $\mathbf{z}$ is a m-dimensional (stochastic) *state vector*. Interestingly, the *t*-eth term of formula (11), $p\left(y_t \mid y^{t-1}, x^{t-1}\right)$ *may be calculated iteratively*. This is possible through the mediation of the state vector, $\mathbf{z}_t$, that in a sense summarises the relevant information at each instant. We need to calculate $p\left(\mathbf{z}_t \mid y^t, x^{t-1}\right)$ first: at step $t = 1$, by setting $x_0 = 0$, we have

$$\begin{aligned} &p\left(\mathbf{z}_1 \mid y^1, x^0\right) \\ &= p(\mathbf{z}_1 \mid y_1, x_0) \propto p(y_1 \mid \mathbf{z}_1)\, p(\mathbf{z}_1), \end{aligned} \tag{13}$$

Where

$$p(y_1 \mid \mathbf{z}_1) = \int \delta\left(y_1 - f(\mathbf{z}_1, w_1)\right) p(w_1)\, \mathrm{d}w_1 . \tag{14}$$

At step t, instead we obtain,

$$\begin{aligned} &p\left(\mathbf{z}_t \mid y^t, x^{t-1}\right) = p\left(\mathbf{z}_t \mid y_t, y^{t-1}, x^{t-1}\right) \\ &\propto p\left(y_t \mid \mathbf{z}_t, y^{t-1}, x^{t-1}\right) p\left(\mathbf{z}_t \mid y^{t-1}, x^{t-1}\right) \\ &= p\left(y_t \mid \mathbf{z}_t\right) \times \\ &\int p\left(\mathbf{z}_t \mid \mathbf{z}_{t-1}, x_{t-1}\right) p\left(\mathbf{z}_{t-1} \mid y^{t-1}, x^{t-2}\right) \mathrm{d}\mathbf{z}_{t-1}. \end{aligned} \tag{15}$$

Now, since

$$p(y_t \mid \mathbf{z}_t) = \int \delta\left[y_t - f(\mathbf{z}_t, w_t)\right] p(w_t)\, \mathrm{d}w_t \tag{16}$$

And

$$\begin{aligned} &p\left(\mathbf{z}_t \mid \mathbf{z}_{t-1}, x_{t-1}\right) = \\ &\int \delta\left[\mathbf{z}_t - g\left(\mathbf{z}_{t-1}, \xi_{t-1}\right)\right] p\left(\xi_{t-1}\right) \mathrm{d}\xi_{t-1}, \end{aligned} \tag{17}$$

Formula (15) allows $p\left(\mathbf{z}_t \mid y^t, x^{t-1}\right)$ to be calculated once $p\left(\mathbf{z}_t \mid y^t, x^{t-1}\right)$ is known.

Finally, to calculate $p\left(y_t \mid y^{t-1}, x^{t-1}\right)$, we observe that

$$\begin{aligned} &p\left(y_t \mid y^{t-1}, x^{t-1}\right) = \\ &\int p\left(y_t \mid \mathbf{z}_t, y^{t-1}, x^{t-1}\right) p\left(\mathbf{z}_t \mid y^{t-1}, x^{t-1}\right) \mathrm{d}\mathbf{z}_t. \end{aligned} \tag{18}$$

The right side of this equation is simply the integral, in respect of $\mathrm{d}\mathbf{z}_t$, of the right side of (15). Considering what has already been observed, we finally obtain

$$\begin{aligned} &p\left(y_t \mid y^{t-1}, x^{t-1}\right) = \int \Big[p\left(y_t \mid \mathbf{z}_t\right) \times \\ &\int p\left(\mathbf{z}_t \mid \mathbf{z}_{t-1}, x_{t-1}\right) p\left(\mathbf{z}_{t-1} \mid y^{t-1}, x^{t-2}\right) \mathrm{d}\mathbf{z}_{t-1} \Big] \mathrm{d}\mathbf{z}_t, \end{aligned} \tag{19}$$

Which, together with (15-17), provides the required result.

2.5 Uncertainty evaluation

The probabilistic approach allows uncertainty to be expressed in the most complete way, since the final results are probability distributions. Any desired uncertainty figure can be immediately obtained from such distributions. For example, the Guide to the expression of uncertainty in measurement (GUM) [3] considers *standard* and *expanded uncertainty*, usually denoted by the symbols *u* and *U* respectively.

When we perform a static measurement, we obtain, as result of restitution, the distribution $p(x\,|\,\mathbf{y})$, from which we obtain the *measurement value*

$$\hat{x} = \mathrm{E}(x\,|\,\mathbf{y}) = \int_X x p(x\,|\,\mathbf{y})\,\mathrm{d}x\,, \tag{20}$$

The standard uncertainty

$$\begin{aligned} u &= \sqrt{\mathrm{E}\left[(\hat{x}-x)^2\,|\,\mathbf{y}\right]} \\ &= \sqrt{\int_X (\hat{x}-x)^2 p(x\,|\,\mathbf{y})\,\mathrm{d}x} \end{aligned} \tag{21}$$

And the expanded uncertainty, at coverage level p_0, as the value U such that

$$\int_{\hat{x}-U}^{\hat{x}+U} p(x\,|\,\mathbf{y})\,\mathrm{d}x = p_0\,. \tag{22}$$

The probabilistic approach is particularly useful for an accurate evaluation of expanded uncertainty.

Instead, if we want to characterise the overall performance of a measuring system for any measurement in its measuring range, we should refer to the distribution $p(\hat{x}\,|\,x)$ and the standard uncertainty for each possible value x of the measurand, as given by

$$\begin{aligned} u(x) &= \sqrt{\mathrm{E}\left[(\hat{x}-x)^2\,|\,x\right]} \\ &= \sqrt{\int_X (\hat{x}-x)^2 p(\hat{x}\,|\,x)\,\mathrm{d}x} \end{aligned} \tag{23}$$

Usually, in order to simplify communication, the maximum value is only provided:

$$u_0 = \max_x \left[u(x)\right]. \tag{24}$$

The expanded uncertainty can be obtained in a similar way.

The above formulae generalise in the case of vector measurement, such as dynamic measurement, and calibration. For example, in the case of dynamic measurement, we obtain the covariance matrix

$$\mathbf{\Sigma} = \mathrm{E}\left((\hat{\mathbf{x}}-\mathbf{x})(\hat{\mathbf{x}}-\mathbf{x})^{\mathrm{T}}\,|\,\mathbf{y}\right), \tag{25}$$

Where the vectors involved are column and "T" denotes transposition. The standard uncertainty of each individual value can be obtained by the square root of the corresponding diagonal term in the matrix. Information on correlation between the measurement values is also provided. It may be noted that, at present, the GUM does not explicitly address vector and dynamic measurement, which could be included in a forthcoming new edition. The results presented in this paper anticipate such a development, however.

Let us now discuss the application of the above model in an example.

3. A WORKING EXAMPLE

Consider the case of a contact linear thermometer [8], whose steady state behaviour is described by the input-output equation:

$$y = kx + w\,, \tag{26}$$

Where k is the sensitivity and w is a normal random variable that accounts for noise in the measuring chain. Suppose, furthermore, that the dynamic behaviour can be described by the first order equation:

$$\begin{aligned} &\tau\dot{z} + z = kx \\ &y = z + w \end{aligned} \tag{27}$$

Where z is a state variable. The system may be calibrated by putting the sensor in a bath where a series of different thermal states are achieved, the temperatures of which are accurately measured by a reference platinum thermometer. The data set $\{(x_i, y_i), i = 1,...,n\}$ is then acquired. Introducing the standard Gaussian distribution

$$\varphi(\xi) = (2\pi)^{-1/2} \exp\left(-\xi^2/2\right), \tag{28}$$

We can express the characteristic distribution as:

$$p(\mathbf{y}\,|\,\mathbf{x}, k, \sigma_w) = \prod_i \sigma_w^{-1} \varphi\left[\sigma_w^{-1}(y_i - kx_i)\right], \tag{29}$$

Where the dependence upon the dispersion parameter σ_w is apparent. Let us rearrange this expression by including the (sufficient) statistics:

$$\begin{aligned} &\hat{k} = \textstyle\sum_i x_i y_i \,/\, \sum_i x_i^2\,, \\ &\hat{\sigma}_w^2 = (n-1)^{-1} \textstyle\sum_i \left(y_i - \hat{k}x_i\right)^2. \end{aligned} \tag{30}$$

After some calculation, we obtain

$$\begin{aligned} &p(\mathbf{y}\,|\,\mathbf{x}, k, \sigma_w) = (2\pi)^{-n/2}\, \sigma_w^{-n} \\ &\exp\left\{-\frac{1}{2\sigma_w^2}\left[(N-1)\hat{\sigma}_w^2 + \left(k-\hat{k}\right)^2 \textstyle\sum_i x_i^2\right]\right\}. \end{aligned} \tag{31}$$

After assuming non informative priors for k and σ_w, we reach the final joint distribution:

$$\begin{aligned} &p(k, \sigma_w\,|\,\mathbf{y}, \mathbf{x}) \propto \sigma_w^{-(n+1)} \\ &\exp\left\{-\frac{1}{2\sigma_w^2}\left[(n-1)\hat{\sigma}_w^2 + \left(k-\hat{k}\right)^2 \textstyle\sum_i x_i^2\right]\right\}, \end{aligned} \tag{32}$$

From which the marginal distributions for k and σ_w can be obtained. Recalling the *t-Student* distribution, with ν degrees of freedom:

$$\psi(\xi;\nu) \propto \left(1+\frac{\xi^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \tag{33}$$

And the *inverse gamma:*

$$\zeta(\xi;\alpha,\beta) \propto \xi^{-(\alpha+1)} \exp(-\beta/\xi), \tag{34}$$

We finally obtain, ignoring from now on the dependence on $(\mathbf{x},\mathbf{y})$ [9],

$$\begin{aligned} p(k) &= \psi\left(\frac{k-\hat{k}}{\hat{\sigma}_w^2/\sum_i x_i^2};n-1\right), \\ p(\sigma_w) &= \zeta\left(\hat{\sigma}_w^2;\frac{n}{2}-1,\frac{1}{2}(n-1)\hat{\sigma}_w^2\right). \end{aligned} \tag{35}$$

The calibration result provides all the information required to use the instrument in equivalent operating and environmental conditions to the calibration ones, i.e., in the same or a similar laboratory and with equivalent definition-uncertainty for the measurand. In this case, the characteristic distribution is (again omitting dependence on the calibration data $(\mathbf{x},\mathbf{y})$):

$$\begin{aligned} p(y\mid x) &= \int_{K\Sigma} p(y\mid x,k,\sigma_w)\, p(k,\sigma_w)\,\mathrm{d}k\mathrm{d}\sigma_w \\ &= \psi\left[(y-kx)\Big/\hat{\sigma}_w\left(1+x^2/\sum_i x_i^2\right)^{1/2};n-1\right]. \end{aligned} \tag{36}$$

On the other hand, a more widespread case, whereas measurement conditions are different from calibration ones, they may be treated by properly combining information from different sources.

Consider, for example, the case in which we have a different, usually higher, level of random noise and a limited resolution. Indeed, operators working with variable-resolution devices often prefer to decrease resolution in order to obtain more stable readings. Then a proper model is

$$y = Q(kx+w'), \tag{37}$$

Where *w'* is the current random effect and Q is a quantisation operator, defined by

$$Q(\xi) = iq \Leftrightarrow (i-1)q \le \xi < iq, \tag{38}$$

Where q is the quantisation interval. Note that the observation y is now a *discrete* random variable: we will then denote its distribution with P, to underlie that this function now is a probability rather than a probability-density. Suppose we take M observations. Then the characteristic distribution is:

$$\begin{aligned} &P(\mathbf{y}\mid x,k,\sigma_{w'}) = \\ &\prod_i \int_{-q/2}^{+q/2} \sigma_{w'}^{-1}\varphi\left[\sigma_{w'}^{-1}(y_i-kx+\xi)\right]\mathrm{d}\xi \\ &= \prod_i \Phi\left(\frac{y_i-kx+q/2}{\sigma_{w'}}\right) - \Phi\left(\frac{y_i-kx-q/2}{\sigma_{w'}}\right), \end{aligned} \tag{39}$$

Where Φ is the cumulative normal distribution. The corresponding restitution equation is, according to the general formula (2):

$$\begin{aligned} &p(x\mid y) = \\ &\int_K \frac{P(\mathbf{y}\mid x,k,\sigma_{w'})\sigma_{w'}^{-1}}{\int_{X,\Sigma_{w'}} P(\mathbf{y}\mid x,k,\sigma_{w'})\sigma_{w'}^{-1}\mathrm{d}\sigma_{w'}\mathrm{d}x} p(k)\,\mathrm{d}k, \end{aligned} \tag{40}$$

Where $p(k)$ is the first distribution in formula (35), obtained by calibration. Expression (40) may be numerically calculated by general-purpose prototype software, such as the UNCERT package developed in our Laboratory [15], or by an *ad hoc* code. On the other hand, an approximate solution can also be obtained if standard uncertainty is the sole concern. Our simulation studies have shown [5] that the effect of resolution is negligible for $\sigma_u > q/2$, in which case it may be ignored. When this condition is not satisfied, remembering that the variance of a t-Student variable with ν degrees of freedom is $\nu/(\nu-2)$, the standard uncertainty may be approximately evaluated by

$$u = \sqrt{\frac{N-1}{N-3}\frac{\hat{\sigma}_{w'}^2}{N} + \frac{q^2}{3} + \hat{\sigma}_k^2\hat{x}^2}, \tag{41}$$

Where $\hat{x}$ is the measurement value and $\hat{\sigma}_k^2$ has been provided by calibration, according to

$$\begin{aligned} x &= \left(\hat{k}M\right)^{-1}\sum_i y_i, \\ \hat{\sigma}_k^2 &= \frac{n-1}{n-3}\frac{\hat{\sigma}_w^2}{\sum_i x_i^2}. \end{aligned} \tag{42}$$

The model thus provides both the complete solution in terms of probability distributions, or simple practical approximate solutions for less demanding applications.

Lastly, consider the case of dynamic measurement, governed in this example by equation (27).

We must first obtain an equivalent discrete-time representation. This may be done, in this case, by using the criterion of having the same impulse response to a stepwise constant input. This yields the system

$$z_{t+1} = az_t + bx_t$$
$$y_t = z_t + \varepsilon_t \tag{43}$$

Where Δt is the sampling interval, $a = \exp(-\Delta t / \tau)$, $b = (1-a)k$ and ε_t is the discrete-time version of the continuous-time random process *w*. By introducing the process

$$v_{t+1} \triangleq \varepsilon_{t+1} - a\varepsilon_t, \tag{44}$$

We obtain

$$y_{t+1} = ay_t + bx_t + v_{t+1} \tag{45}$$

And

$$p(y_1 \mid x_1) = p(y_1)$$
$$p(y_{i+1} \mid y^t, x^t) = p(y_{i+1} \mid y_t, x_t) \tag{46}$$
$$= p_v(y_{t+1} - ay_t - bx_t).$$

Observation is now characterised by

$$p(\mathbf{y} \mid \mathbf{x}) = p(y_1) \prod_{t=1}^{N-1} p_v(y_{t+1} - ay_t - bx_t) \tag{47}$$

And restitution by

$$p(\mathbf{x} \mid \mathbf{y}) \propto \prod_{t=1}^{N-1} p_v(y_{t+1} - ay_t - bx_t). \tag{48}$$

The marginal distribution for each time value x_t may be explicitly expressed by:

$$p(x_t \mid \mathbf{y}) = \sigma_c^{-1} \varphi\left(\sigma_c^{-1}(x_t - \hat{x}_t)\right) \tag{49}$$

where

$$\hat{x}_t = \frac{y_{t+1} - ay_t}{b} \tag{50}$$

And

$$c_t \triangleq x_t - \hat{x}_t = -\frac{\varepsilon_{t+1} - a\varepsilon_t}{b}, \tag{51}$$

$$\sigma_c = b^{-1}\sqrt{1+a^2}\,\sigma_\varepsilon. \tag{52}$$

Note that formula (51) defines a (non-causal) probabilistic de-convolution filter. The final distribution will derive from a combination of the marginal found and the distribution for *k*, (35), provided by static calibration. Again we have an example of the combination of pieces of information coming from different experiments.

3. FINAL REMARKS

Mathematical modelling is essential for measurement as well as for any other scientific activity. In measurement, modelling concerns both the object to be measured and the process of measuring. In both cases, uncertainty is an inherent feature that needs to be accounted for at some stage.

I think that, at a high level of abstraction, modelling may be somewhat discipline-independent, since measurement follows a common pattern in physics, in engineering and, say, in psychology [27, 28]. Here I have mainly discussed the modelling of the measuring process which, in my opinion, can be addressed in fairly general terms.

In fact, measurement may be seen as a mapping of states of objects, in respect to some characteristic that interests us, into measurement values, that is, numbers. This requires an intermediate step, consisting in acquiring indications from a measuring system, that must be processed in order to obtain the final result. Uncertainty may be accounted for by adopting a probabilistic approach and, in this case, the main part of the model is a conditional probabilistic distribution that I have called the characteristic distribution of the process. Interestingly enough, this distribution addresses both calibration and measurement, and is applicable in both the static and the dynamic case. Since the approach is probabilistic, the final results are always expressed as probabilistic distributions from which the desired uncertainty features, such as the standard or the expanded uncertainty, may be directly obtained. Starting from the general formulation, it is also possible to introduce simplifications which may provide acceptable results, in practical cases, using a simpler form of mathematics. To illustrate the application of the model, I have discussed in some analytical detail the case of a linear measuring device, with first order dynamics, such as a thermometer, and I have shown how the same model can be used to address calibration and static and dynamic measurement. I have also shown how it is possible to combine calibration information with that derived from the measuring process.

At present, current approaches to uncertainty evaluation range from a pragmatic application of agreed good practices [3] to more theoretical formulations [10, 22, 24]. The latter often refer to a “Bayesian” approach. Yet, as I have discussed elsewhere [26], although the restitution process naturally “embeds” a Bayesian inference, it is not appropriate, in my opinion, to claim that it is totally reducible to a Bayesian inference. In fact, part of the restitution process, the one that deals with systematic effects, does not fit in with a Bayesian scheme since, simply, for systematic effects is not possible to learn from experience during the measurement process (whilst it is in some

side verification processes). So, in my opinion, the approach proposed in this paper is more satisfactory in that it provides a general framework based on an analysis of how the measuring system works rather than on generic adherence to some school of thought (such as the Bayesian one). Once the general framework - which consists of just three formulae (1-3)! - is accepted, no additional ad hoc assumption is required: the framework has simply to be explicated for the case under consideration, as illustrated in the above example. Moreover, there is another, even more stringent aspect in favour of this approach. Although, for the sake of brevity, I have mainly discussed the measuring process, elsewhere I have shown that this framework is just part of an overall probabilistic and interdisciplinary theory of measurement [17, 21]. The statements made in this paper are part of a general coherent framework. In fact, for any measurement theory, the first question to be asked concerns the meaning of measurement and the related issue of measurability. This leads to the notion of measurement scale and of measuring system that have to be correctly integrated.

The GUM, for example, does not consider this kind of problem, and this is good, provided the GUM is correctly interpreted as a *practical pragmatic guide*. But this is no longer satisfactory, in my opinion, when a general theoretical framework is searched. In order to find a simple approach to measurement evaluation, the GUM starts from the paradigm of indirect measurement. This makes sense to some extent in that it provides practical computation guidelines, but it is no longer appropriate when the goal is to put measurement on a solid foundation. In this case, it is apparent that the indirect measurement paradigm is inadequate since it requires direct measurement to be defined first. In a system of quantities, one of them at least needs to be defined directly.

In conclusion, I would say that modelling is essential in measurement and that it may be addressed in general terms within a coherent framework for the overall theory of measurement. One such framework, which uses a probabilistic logic for expressing uncertainty, has been discussed in this paper. Alternative solutions could and should be explored. I am firmly convinced that the richness of a discipline lies in its openness and not in any kind of homologation. But a measurement theory should in any case consider the overall building and not just a single problem if measurement is to play a key role in science.

ACKNOWLEDGEMENT

I am grateful to Roman Morawski for his precious hints and suggestions in the development of this work.

REFERENCES

[1] S. J. Press, *Bayesian statistics*, Wiley, New York, 1989
[2] F. Barone, “I problemi epistemologici della misurazione”, in: M. Cunietti, L. Mari (Eds) *Atti della X Giornata della Misurazione*, CLUP, Milano, 1992.
[3] BIPM, IEC, IFCC, ISO, IUPC, IUPAP, OIMIL, *Guide to the expression of uncertainty in measurement*, ISO, Geneva, 1993
[4] R. Morawski, “Unified approach to measurand reconstruction”, *IEEE Trans IM*, vol. 43, no.2, pp. 226-231, 1994
[5] R. C. Michelini, G. B. Rossi, “Assessing measurement uncertainty in quality engineering”, IEEE IM Technology Conference, Brussels, Belgium, June 4-6, 1996, pp. 1217-1221
[6] L. Finkelstein, “Foundational problems in measurement”, in K. Kariya, L. Finkelstein (Eds): *Measurement Science - A Discussion*, Ohmsa Press, Amsterdam, 2000.
[7] I. Hacking, *An introduction to probability and inductive logic*, Cambridge University Press, 2001
[8] J. V. Nicholas, D.R. White, *Traceable temperature*, Wiley, 2001
[9] J. Gill, *Bayesian methods*, Chapman & Hall/CRC, Boca Raton, 2002
[10] I. Lira, *Evaluating the measurement uncertainty*, IOP, Bristol, 2002
[11] P. Rossi, *La nascita della scienza moderna in Europa*, Laterza, Torino, 2002
[12] M. Marsonet (Ed.) *The problem of realism*, Ashgate, Aldershot, 2002
[13] R. Morawski, “Realists’ versus instrumentalists’ understanding of measurement”, XVII World Congress, 2003.
[14] L. Mari, “Epistemology of measurement”, *Measurement*, vol. 34, no. 1, pp. 17-30, 2003
[15] G. B. Rossi, F. Crenna, M. Codda, “Metrology software for the expression of measurement results by direct calculation of probability distributions”, in P. Ciarlini, M. G. Cox, F. Pavese, D. Richter and G. B. Rossi (eds.): *Advanced Mathematical Tools In Metrology VI*, World Scientific, Singapore, 2004
[16] D. Costantini, *I fondamenti storico filosofici delle discipline statistico-probabilistiche*, Bollati Boringhieri, Torino, 2004
[17] G. B. Rossi, “A probabilistic theory of measurement”, Measurement, vol. 39, pp. 34-50, 2006
[18] G. B. Rossi, “An attempt to interpret some problems in measurement science on the basis of Kuhn’s theory of paradigms”, Measurement, vol. 39, pp. 512-521, 2006
[19] J. P. Hessling, “Dynamic metrology”, Measurement Science and Technology, vol 19, 084008 (7p), 2008
[20] M. G. Cox, G. B. Rossi, P. M. Harris, A. Forbes, A probabilistic approach to the analysis of measurement processes, Metrologia, vol. 45, pp. 493-502, 2008
[21] G.B. Rossi, “Probability in Metrology”, in: F. Pavese, A. Forbes (Eds) *Data modeling for metrology and testing in measurement science*, Birkhauser-Springer, Boston, 2009.
[22] K. D. Sommer, “Modelling of Measurements, System Theory, and Uncertanty Evaluation”, in: F. Pavese, A. Forbes (Eds) *Data modeling for metrology and testing in measurement science*, Birkhauser-Springer, Boston, 2009.
[23] K. D. Sommer, “Modelling of dynamic measurements for uncertainty analysis by means of discretised state-

space forms", XIX IMEKO World Congress, Lisbon, 2009
[24] I. Lira, "Bayesian assessment of uncertainty in metrology", Metrologia, Vol. 47, R1-R14, 2010.
[25] G. Reale, D. Antiseri , *Storia della filosofia*, 14 Vols, Bompiani, 2010.
[26] G.B. Rossi, "Probabilistic inferences related to the measurement process", Journal of Physics: Conference Series 238 (2010)
[27] G.B. Rossi, B. Berglund, "Measurement of quantities involving human perception and interpretation", *Measurement*, vol. 44, pp. 815-822, 2011
[28] B. Berglund, G. B. Rossi, J. Townsend, L. Pendrill (Eds.), *Measurement with persons*, Taylor and Francis, 2011, in press

Author: Prof. Giovanni B. Rossi, Università degli Studi di Genova, Italy, gb.rossi@dimec.unige.it.