Bayesian Analysis of a Simple Measurement Model Distinguishing between Types of Information

Ignacio Lira[1] and Dieter Grientschnig[2]

[1]Department of Mechanical and Metallurgical Engineering, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Santiago, Chile, ilira@ing.puc.cl
[2]Metrolytik, Kernstockstr. 4, 8600 Bruck an der Mur, Austria

Let a quantity of interest, *Y*, be modeled in terms of a quantity *X* and a set of other quantities *Z*. Suppose that for *Z* there is type B information, by which we mean that it leads directly to a joint state-of-knowledge probability density function (PDF) for that set, without reference to likelihoods. Suppose also that for *X* there is type A information, which signifies that a likelihood is available. The posterior for *X* is then obtained by updating its prior with said likelihood by means of Bayes' rule, where the prior encodes whatever type B information there may be available for *X*. If there is no such information, an appropriate non-informative prior should be used. Once the PDFs for *X* and *Z* have been constructed, they can be propagated through the measurement model to obtain the PDF for *Y*, either analytically or numerically. But suppose that, at the same time, there is also information of type A, type B or both types together for the quantity *Y*. By processing such information in the manner described above we obtain *another* PDF for *Y*. Which one is right? Should both PDFs be merged somehow? Is there another way of applying Bayes' rule such that a single PDF for *Y* is obtained that encodes all existing information? In this paper we examine what we believe should be the proper ways of dealing with such a (not uncommon) situation.

Keywords: Measurement uncertainty, probability density function, Bayesian statistics, information types, data processing.

1. INTRODUCTION

IN THE EVALUATION of measurement uncertainty, the dispute over the Bayesian versus frequentist interpretations of probability seems to be resolving itself in favor of the former view. One argument supporting this assertion is the upcoming revision of the *Guide to the Expression of the Uncertainty in Measurement* (GUM) [1] – currently in preparation by the Joint Committee for Guides in Metrology, JCGM – which is expected to follow an entirely Bayesian approach [2, 3]. This implies that no distinction will be made between the 'type A' and 'type B' evaluation methods.

As is well known, these methods were established by Recommendation INC-1 (1980) of the Working Group on the Statement of Uncertainties, from which the GUM originated, to differentiate between what at the time were considered to be the two ways of obtaining the standard uncertainty associated with an estimate for a certain quantity. In the type A evaluation, uncertainty components were to be characterized by estimated variances (or estimated standard deviations) together with their corresponding number of degrees of freedom. It follows from this terminology that this method was intended to be strictly frequentist in nature. In turn, the 'type B' evaluation was defined as that proceeding 'by other means', that is, in ways other than the (frequentist) statistical analysis of the measurement information. The uncertainty components in this category were to be characterized by terms that could be considered as approximations to the corresponding variances.

This brief outline was augmented in the GUM [1], wherein detailed guidance was provided for implementing the two evaluation methods. For example, the type B method would be used if one only knew that the possible values of a quantity are contained within a given finite interval, in which case a rectangular distribution with support on that interval would be assigned [1, clause 4.3.7]. However, such a distribution is not to be interpreted in a frequentist sense, but rather as a probability density function (PDF) encoding the existing 'state of knowledge' about the quantity. It follows that Bayesian concepts do appear in the current GUM, albeit implicitly.

Along with Bich [3], we believe that in metrological applications the time is ripe to forgo the frequentist view altogether in favor of the Bayesian one, and to drop the distinction between *methods* of uncertainty evaluation. The reason is that, in the Bayesian framework, there is just one single and general evaluation methodology: it consists in using probability theory for deriving a state of knowledge PDF for the quantity of interest. Summary features of this distribution can then be provided: its mean (for the best estimate of the quantity) and its standard deviation (for the standard uncertainty associated with the best estimate). Various coverage intervals having specified probabilities of containing the value of the quantity can also be computed. This value is assumed to be fixed – at the time of measurement – but indeterminable; it could only be made known exactly by means of a perfect measurement.

Nevertheless, we suggest it would be useful to retain the notion of *types* of information about the quantity whose PDF is to be constructed, say *X*. The type B information, designated as $\mathscr{I}_X^B$, can be defined as that leading directly to a PDF $f(x \mid \mathscr{I}_X^B)$ without reference to a likelihood, as in the example above for the rectangular PDF [1].

---

[1]Throughout we shall use upper case letters to denote quantities and lower case letters for their possible values. The 'given' part of the PDF's arguments shall indicate on which information they rely.

To obtain a PDF based on type A information about X, application of the well-known Bayes rule is required. We define this type of information, denoted by $\mathscr{I}_X^A$, as that imparted by direct observations, which are typically regarded as being independent draws from an appropriate sampling distribution of which X and other quantities $\boldsymbol{P}_X$ are parameters. It is assumed that this information is sufficient for setting up a likelihood $\ell\left(x, \boldsymbol{p}_X; \mathscr{I}_X^A\right)$, where $\boldsymbol{p}_X$ is the set of variables for the possible values of the parameters $\boldsymbol{P}_X$ (see (22) in section 3 for an example.). Once the likelihood has been established, it is used to update the prior for the parameters X and $\boldsymbol{P}_X$. This prior encodes any other relevant information $\mathscr{I}$ about them, if there is any, otherwise a non-informative prior needs to be chosen. The result is the posterior PDF for these parameters, out of which the PDF $f(x \,|\, \mathscr{I}_X^A, \mathscr{I})$ would be obtained by marginalization (see below for further details).

Note that this procedure does not involve the notion of degrees of freedom, is not frequentist, and differs markedly from the type A evaluation as outlined in the GUM. Note also that it does not imply unique results in all situations. In particular, since non-informative priors may be established by different rules, minor variations of results may be expected, see e.g. [4, 5]. Hence, despite the statement above that 'just one single and general evaluation methodology' exists, slightly differing results may nonetheless ensue.

In metrology, however, a quantity of interest not directly accessible by measurement is usually modeled as a function of one or more input quantities, about which information of one or both types should be available. Two ways of processing that information are then possible. The first is the so-called GUM uncertainty framework, which consists in applying the 'law of propagation of standard uncertainties' associated with the best estimates of the input quantities, to obtain the standard uncertainty associated with the best estimate of the measurand. Naturally, the best estimates and the standard uncertainties of the input quantities should be the means and standard deviations of the corresponding PDFs, obtained as described above. Alternatively, the PDFs for the input quantities may be 'propagated' through the measurement model in order to obtain the PDF for the output quantity. This second option is preferable, because the resulting PDF fully utilizes all information at hand.

An example is the model $Y = X/Z$, where Y is the length of a gauge block at a reference temperature, X is its length at the calibration temperature, and Z is a thermal correction factor [6, 7]. If the measurand is Y, some information on both X and Z is needed. That on the former might be of type A, as acquired with an interferometer, say, while knowledge about the latter might be taken from the scientific literature and be of type B. The analysis would yield PDF $f(y \,|\, \mathscr{M}, \mathscr{I}_X^A, \mathscr{I}_Z^B)$ for Y, where the letter $\mathscr{M}$ symbolizes that the model is also included in the information that has been utilized. However, suppose that at the same time there were type B information for Y, as taken from a previous calibration certificate of the gauge block. This information would lead directly to PDF

Scenario	Quantity	
	X	Y
1	A	N
2	B	N
3	A & B	N
4	N	A
5	A	A
6	B	A
7	A & B	A
8	N	B
9	A	B
10	B	B
11	A & B	B
12	N	A & B
13	A	A & B
14	B	A & B
15	A & B	A & B

Table 1: Combinations of type A, type B and no information (N) for the quantities X and Y to be considered in the model $Y = F(X, \boldsymbol{Z})$. In all scenarios type B information about $\boldsymbol{Z}$ is assumed to have already been encoded by a given PDF.

$f(y \,|\, \mathscr{I}_Y^B)$, with no reference to the model at all. How should a PDF $f(y \,|\, \mathscr{M}, \mathscr{I}_X^A, \mathscr{I}_Y^B, \mathscr{I}_Z^B)$ based on the model and on all three pieces of information be obtained? Would the method for so doing depend on the type of information at hand for the three quantities? Would such a PDF have a smaller or larger standard uncertainty than the two previous ones?

So far, the documents produced by the JCGM have been silent about such questions, although they are obviously relevant to practical problems in metrology. A general method to produce a PDF based on information of both types in a simple measurement model will be proposed in section 2 below. Section 3 presents a short discussion on the non-informative priors to be used in the case of Gaussian likelihoods. An example for illustrating the method is given in section 4 and conclusions are drawn in section 5.

2. ANALYSIS

We shall consider a measurement model of the type

$$Y = F(X, \boldsymbol{Z}), \tag{1}$$

where Y is the measurand. We shall analyze scenarios where for both X and Y there may be information of type A, type B, both types together or no information at all, whereas for $\boldsymbol{Z}$ we shall consider type B information only, see table 1.

It will be assumed that, by itself, acquisition of knowledge about any one of these quantities will not provide new knowledge about the other quantities *before considering the model*. For example, in scenario 9, information pieces $\mathscr{I}_X^A$ and $\mathscr{I}_Y^B$ lead to independent PDFs $f(x \,|\, \mathscr{I}_X^A)$ and $f(y \,|\, \mathscr{I}_Y^B)$. However,

by action of the model in conjunction with the information on $\boldsymbol{Z}$, information $\mathscr{I}_X^A$ gives rise to a joint PDF for $\{Y, \boldsymbol{Z}\}$, while information $\mathscr{I}_Y^B$ produces a joint PDF for $\{X, \boldsymbol{Z}\}$. From these PDFs the marginals $f\left(y | \mathscr{M}, \mathscr{I}_X^A\right)$ and $f\left(x | \mathscr{M}, \mathscr{I}_Y^B\right)$ can be obtained. In other words, the model plus information on X results in information on Y and vice versa.

For completeness, all scenarios in table 1 will be studied. The analysis of those for which there is no information about X or Y (i.e. scenarios 1 – 4, 8, and 12) is straightforward. Notwithstanding, we shall begin with them because they are the basis of the rest of the scenarios.

2.1. Scenarios 1, 2 and 3: No information about Y

The case in which there is information about the input quantities X and $\boldsymbol{Z}$ but none about the measurand Y presents no difficulties. Consider first information of type B only about X, which leads to PDF $f(x | \mathscr{I}_X^B)$. To obtain the PDF for Y, we shall assume that for given $\boldsymbol{Z}$ the relation between X and Y is one-to-one, at least within the domains of interest, so model (1) can be written equivalently as $X = G(Y, \boldsymbol{Z})$. Thus, changing variables yields

$$f\left(y, \boldsymbol{z} | \mathscr{M}, \mathscr{I}_X^B\right) = \left|\frac{\partial G}{\partial y}\right| f\left(G | \mathscr{I}_X^B\right) f(\boldsymbol{z}), \qquad (2)$$

from which $f\left(y | \mathscr{M}, \mathscr{I}_X^B\right)$ can be obtained by marginalization. Note that, for conciseness, we have written G instead of $G(y, z)$ and $f(\boldsymbol{z})$ instead of $f(\boldsymbol{z} | \mathscr{I}_{\boldsymbol{Z}}^B)$.

Suppose now that there is information of types A and B about X but again no information about Y. As was mentioned in the Introduction, information $\mathscr{I}_X^A$ leads to the likelihood $\ell\left(x, \boldsymbol{p}_X; \mathscr{I}_X^A\right)$, while information $\mathscr{I}_X^B$ is encoded in the prior $f\left(x, \boldsymbol{p}_X | \mathscr{I}_X^B\right)$. The term $\mathscr{I}_X^B$ may or may not include information about the parameters $\boldsymbol{P}_X$. If not, some kind of formalism may be needed to construct this prior, see e.g. [8]. From Bayes' rule we get

$$f\left(x, \boldsymbol{p}_X | \mathscr{I}_X^{A+B}\right) \propto \ell\left(x, \boldsymbol{p}_X; \mathscr{I}_X^A\right) f\left(x, \boldsymbol{p}_X | \mathscr{I}_X^B\right), \qquad (3)$$

and by transformation into the $\{Y, \boldsymbol{Z}\}$ coordinates we obtain

$$f\left(y, \boldsymbol{z}, \boldsymbol{p}_X | \mathscr{M}, \mathscr{I}_X^{A+B}\right) \propto \\ \left|\frac{\partial G}{\partial y}\right| \ell\left(G, \boldsymbol{p}_X; \mathscr{I}_X^A\right) f\left(G, \boldsymbol{p}_X | \mathscr{I}_X^B\right) f(\boldsymbol{z}). \quad (4)$$

This time, to find $f\left(y | \mathscr{M}, \mathscr{I}_X^{A+B}\right)$ we need not only marginalize but also compute the normalization constant.

An alternative approach to treat scenario 3 is to apply Bayes' rule in the form [9]

$$f\left(y, \boldsymbol{z}, \boldsymbol{p}_X | \mathscr{M}, \mathscr{I}_X^{A+B}\right) \propto \ell\left(G, \boldsymbol{p}_X; \mathscr{I}_X^A\right) f\left(y, \boldsymbol{z}, \boldsymbol{p}_X | \mathscr{I}_X^B\right). \qquad (5)$$

This so-called *observation equation* approach [10] needs a prior $f\left(y, \boldsymbol{z}, \boldsymbol{p}_X | \mathscr{I}_X^B\right)$. By comparing (4) and (5) we see that, for the sake of consistency, this prior must be the transformed

version of $f\left(x, \boldsymbol{p}_X | \mathscr{I}_X^B\right)$, that is,

$$f\left(y, \boldsymbol{z}, \boldsymbol{p}_X | \mathscr{I}_X^B\right) = \left|\frac{\partial G}{\partial y}\right| f\left(G, \boldsymbol{p}_X | \mathscr{I}_X^B\right) f(\boldsymbol{z}). \qquad (6)$$

The analysis for the case of type A information only about X (and no information about Y) follows from (4) and (5) which become, respectively,

$$f\left(y, \boldsymbol{z}, \boldsymbol{p}_X | \mathscr{M}, \mathscr{I}_X^A\right) \propto \left|\frac{\partial G}{\partial y}\right| \ell\left(G, \boldsymbol{p}_X; \mathscr{I}_X^A\right) f\left(G, \boldsymbol{p}_X\right) f(\boldsymbol{z}) \qquad (7)$$

and

$$f\left(y, \boldsymbol{z}, \boldsymbol{p}_X | \mathscr{M}, \mathscr{I}_X^A\right) \propto \ell\left(G, \boldsymbol{p}_X; \mathscr{I}_X^A\right) f\left(y, \boldsymbol{z}, \boldsymbol{p}_X\right), \qquad (8)$$

where $f\left(G, \boldsymbol{p}_X\right)$ in (7) must be an appropriate non-informative prior, whose transformation yields the one to be used in (8).

2.2. Scenarios 4, 8 and 12 : No information about X

Formulas for those cases where there is no information about X and any type of information about Y follow immediately from the analysis above. Thus, if only $\mathscr{I}_Y^B$ is available, PDF $f\left(y | \mathscr{I}_Y^B\right)$ can be obtained directly.

If both $\mathscr{I}_Y^A$ and $\mathscr{I}_Y^B$ are present, the analog of (3) gives

$$f\left(y, \boldsymbol{p}_Y | \mathscr{I}_Y^{A+B}\right) \propto \ell\left(y, \boldsymbol{p}_Y; \mathscr{I}_Y^A\right) f\left(y, \boldsymbol{p}_Y | \mathscr{I}_Y^B\right), \qquad (9)$$

where $\boldsymbol{p}_Y$ are the possible values of the parameters $\boldsymbol{P}_Y$ of the sampling distribution from which the data $\mathscr{I}_Y^A$ are assumed to be drawn. Again, note that the prior $f\left(y, \boldsymbol{p}_Y | \mathscr{I}_Y^B\right)$ may need to be determined by using the procedure in [8] if $\mathscr{I}_Y^B$ does not include information about all parameters $\boldsymbol{P}_Y$.

Finally, if only $\mathscr{I}_Y^A$ exists we use

$$f\left(y, \boldsymbol{p}_Y | \mathscr{I}_Y^A\right) \propto \ell\left(y, \boldsymbol{p}_Y; \mathscr{I}_Y^A\right) f\left(y, \boldsymbol{p}_Y\right), \qquad (10)$$

where $f\left(y, \boldsymbol{p}_Y\right)$ is a non-informative prior. In all these instances, the measurement model is not needed. It would only be used if the PDF for X were desired.

2.3. Scenarios 6 and 9 : type B information about X plus type A information about Y, and vice versa

We are now ready to address the more interesting scenarios for which there is information on all quantities in the model. Consider first scenario 6, in which there is type B information about X plus type A information about Y. The former is encoded by $f(x | \mathscr{I}_X^B)$, while the latter is expressed by the likelihood $\ell\left(y, \boldsymbol{p}_Y; \mathscr{I}_Y^A\right)$. To use Bayes' rule we need a prior for Y and $\boldsymbol{P}_Y$. We shall assume that this prior can be written as $f(y, \boldsymbol{p}_Y) = f(y) f(\boldsymbol{p}_Y)$, implying that the parameters $\boldsymbol{P}_Y$ (if any) are independent from Y. This will usually be a fitting premise, because in most cases there will be no logical connection between Y and the parameters $\boldsymbol{P}_Y$ of the sampling

distribution from which information $\mathscr{I}_Y^A$ is assumed to be obtained. Use of (2) for the prior $f\left(y,z\,|\,\mathscr{M},\mathscr{I}_X^B\right)$ then yields

$$f\left(y,z,\boldsymbol{p}_Y\,|\,\mathscr{M},\mathscr{I}_X^B,\mathscr{I}_Y^A\right) \propto$$
$$\left|\frac{\partial G}{\partial y}\right| \ell\left(y,\boldsymbol{p}_Y;\mathscr{I}_Y^A\right) f\left(G\,|\,\mathscr{I}_X^B\right) f(z) f(\boldsymbol{p}_Y), \quad (11)$$

where $f(\boldsymbol{p}_Y)$ must be non-informative.

It should be mentioned that, since there is no information about Y and $\boldsymbol{P}_Y$, one might be tempted to write (11) as

$$f\left(y,z,\boldsymbol{p}_Y\,|\,\mathscr{M},\mathscr{I}_X^B,\mathscr{I}_Y^A\right) \propto \ell\left(y,\boldsymbol{p}_Y;\mathscr{I}_Y^A\right) f(y,\boldsymbol{p}_Y) f(z),$$
$$(12)$$

with a non-informative prior $f(y,\boldsymbol{p}_Y)$ but without utilizing $f\left(x\,|\,\mathscr{I}_X^B\right)$ at all. However, the inconsistency between the 'given' part of the arguments of the PDFs on either side of this equation makes evident that, as remarked in [6], 'it would not be sensible to perform a Bayesian analysis with a non-informative prior when in fact information is available'.

The situation in scenario 9 is opposite to that in scenario 6, for we now have type A information about X and type B information about Y. (See [11] for a recent example of such scenario.) In this case compute first

$$f\left(x,z\,|\,\mathscr{M},\mathscr{I}_Y^B\right) = \left|\frac{\partial F}{\partial x}\right| f\left(F\,|\,\mathscr{I}_Y^B\right) f(z), \quad (13)$$

where F is short for $F(x,z)$. Then insert this prior in

$$f\left(x,z,\boldsymbol{p}_X\,|\,\mathscr{M},\mathscr{I}_X^A,\mathscr{I}_Y^B\right) \propto$$
$$\ell\left(x,\boldsymbol{p}_X;\mathscr{I}_X^A\right) f\left(x,z\,|\,\mathscr{M},\mathscr{I}_Y^B\right) f(\boldsymbol{p}_X), \quad (14)$$

with $f(\boldsymbol{p}_X)$ being non-informative. Finally, transform this PDF back to the $\{Y,\boldsymbol{Z}\}$ coordinates. These operations yield

$$f\left(y,z,\boldsymbol{p}_X\,|\,\mathscr{M},\mathscr{I}_X^A,\mathscr{I}_Y^B\right) \propto$$
$$\ell\left(G,\boldsymbol{p}_X;\mathscr{I}_X^A\right) f\left(y\,|\,\mathscr{I}_Y^B\right) f(z) f(\boldsymbol{p}_X), \quad (15)$$

which can be regarded as an observation equation approach.

2.4. Scenarios 7 and 13 : types A and B information about X plus type A information about Y, and vice versa

The analysis when there are both types of information about X together with type A information about Y follows by simply updating the PDF $f\left(y,z\,|\,\mathscr{M},\mathscr{I}_X^{A+B}\right)$, obtained from (4), with the likelihood $\ell\left(y,\boldsymbol{p}_Y;\mathscr{I}_Y^A\right)$, that is

$$f\left(y,z,\boldsymbol{p}_Y\,|\,\mathscr{M},\mathscr{I}_X^{A+B},\mathscr{I}_Y^A\right) \propto$$
$$\ell\left(y,\boldsymbol{p}_Y;\mathscr{I}_Y^A\right) f\left(y,z\,|\,\mathscr{M},\mathscr{I}_X^{A+B}\right) f(\boldsymbol{p}_Y). \quad (16)$$

In the opposite case (type A information about X together with types A and B information about Y) the PDF $f\left(y\,|\,\mathscr{I}_Y^{A+B}\right)$, obtained from (9), is updated by the likelihood $\ell\left(x,\boldsymbol{p}_X;\mathscr{I}_X^A\right)$. This is done by writing

$$f\left(y,z,\boldsymbol{p}_X\,|\,\mathscr{M},\mathscr{I}_X^A,\mathscr{I}_Y^{A+B}\right) \propto$$
$$\ell\left(G,\boldsymbol{p}_X;\mathscr{I}_X^A\right) f\left(y\,|\,\mathscr{I}_Y^{A+B}\right) f(z) f(\boldsymbol{p}_X), \quad (17)$$

which is a simple extension of (15).

2.5. Scenarios 10, 11, 14 and 15 : type B information about X and Y

In all scenarios considered so far, the existing information is such that it leads to an unambiguous PDF for the measurand Y. This does not happen in the remaining scenarios. Consider first scenario 10, in which there is only type B information about both X and Y. The reason for the ambiguity is that information $\mathscr{I}_Y^B$ allows $f\left(y\,|\,\mathscr{I}_Y^B\right)$ to be obtained directly, but marginalizing (2) yields $f\left(y\,|\,\mathscr{M},\mathscr{I}_X^B\right)$. Naturally, only under very unusual circumstances will these two PDFs coincide.

The simplest way out of this seemingly conflicting situation would be to discard one of the two pieces of information. But if both of them are to be retained, $f\left(y\,|\,\mathscr{I}_Y^B\right)$ and $f\left(y\,|\,\mathscr{M},\mathscr{I}_X^B\right)$ ought to be merged. In [12] we reviewed two methods for doing so: *logarithmic pooling*

$$f\left(y\,|\,\mathscr{M},\mathscr{I}_X^B,\mathscr{I}_Y^B\right) \propto \left[f\left(y\,|\,\mathscr{M},\mathscr{I}_X^B\right)\right]^{w_X} \left[f\left(y\,|\,\mathscr{I}_Y^B\right)\right]^{w_Y},$$
$$(18)$$

and *linear pooling*

$$f\left(y\,|\,\mathscr{M},\mathscr{I}_X^B,\mathscr{I}_Y^B\right) = w_X f\left(y\,|\,\mathscr{M},\mathscr{I}_X^B\right) + w_Y f\left(y\,|\,\mathscr{I}_Y^B\right),$$
$$(19)$$

where in both cases the pooling weights w_X and w_Y should add up to one. They should be selected by subjectively judging the reliabilities of the corresponding pieces of information.

If additionally there were type A information, the PDF $f\left(y\,|\,\mathscr{M},\mathscr{I}_X^B,\mathscr{I}_Y^B\right)$ obtained by either method should be updated by the likelihoods $\ell\left(x,\boldsymbol{p}_X;\mathscr{I}_X^A\right)$ in scenario 11, $\ell\left(y,\boldsymbol{p}_Y;\mathscr{I}_Y^A\right)$ in scenario 14 or both of them in scenario 15.

2.6. Scenario 5 : type A information about X and Y

The remaining scenario is that of type A information only about both X and Y. Two approaches are then possible. In approach a) the PDF $f\left(y,z,\boldsymbol{p}_X\,|\,\mathscr{M},\mathscr{I}_X^A\right)$, obtained from the equivalent equations (7) or (8), serves as prior to be updated by the likelihood $\ell\left(y,\boldsymbol{p}_Y;\mathscr{I}_Y^A\right)$. This approach yields

$$f_a\left(y,z,\boldsymbol{p}_X,\boldsymbol{p}_Y\,|\,\mathscr{M},\mathscr{I}_X^A,\mathscr{I}_Y^A\right) \propto \left|\frac{\partial G}{\partial y}\right| \times$$
$$\ell\left(G,\boldsymbol{p}_X;\mathscr{I}_X^A\right) \ell\left(y,\boldsymbol{p}_Y;\mathscr{I}_Y^A\right) f(G,\boldsymbol{p}_X) f(\boldsymbol{p}_Y) f(z). \quad (20)$$

In approach b) transform $f\left(y,\boldsymbol{p}_Y\,|\,\mathscr{I}_Y^A\right)$ obtained from (10) into the $\{X,\boldsymbol{Z}\}$ coordinates and update the result with the likelihood $\ell\left(x,\boldsymbol{p}_X;\mathscr{I}_X^A\right)$. Then transform the result back to the $\{Y,\boldsymbol{Z}\}$ coordinates. This approach leads to

$$f_b\left(y,z,\boldsymbol{p}_X,\boldsymbol{p}_Y\,|\,\mathscr{M},\mathscr{I}_X^A,\mathscr{I}_Y^A\right) \propto$$
$$\ell\left(G,\boldsymbol{p}_X;\mathscr{I}_X^A\right) \ell\left(y,\boldsymbol{p}_Y;\mathscr{I}_Y^A\right) f(\boldsymbol{p}_X) f(y,\boldsymbol{p}_Y) f(z). \quad (21)$$

Thus, two different PDFs for Y based on the model and on information $\mathscr{I}_X^A$ and $\mathscr{I}_Y^A$ can be obtained. Their difference lies in the non-informative priors used: for X in approach a) and for Y in approach b). Is this an important ambiguity and is there a way out of it?

To answer these questions, we note that the PDF derived by approach a) is the limit of the one in (16) with $f(x, \boldsymbol{p}_X)$ taking the place of $f(x, \boldsymbol{p}_X | \mathscr{I}_X^{\mathrm{B}})$, and that the PDF derived by approach b) is the limit of the one in (17) with $f(y, \boldsymbol{p}_Y)$ taking the place of $f(y, \boldsymbol{p}_Y | \mathscr{I}_Y^{\mathrm{B}})$. Therefore, since scenario 5 may be considered as the limiting case of scenario 15 (with less and less information $\mathscr{I}_X^{\mathrm{B}}$ and $\mathscr{I}_Y^{\mathrm{B}}$) one might think of adopting the approach of the latter, i.e. pooling of the two non-informative priors. Since the same pair of likelihoods is used for successive updating in (20) and (21), logarithmic pooling of the priors followed by updating with the two likelihoods would lead to the same outcome as if (20) and (21) were directly logarithmically pooled. This property, which additive pooling does not have, is called 'external Bayesianity' [13] and was considered in a metrological context in [14]. However, given that non-informative priors normally have only a minor impact on the outcome, pooling them will in most cases be an unnecessary hassle. In fact, as we shall see in the example below, the marginals $f_{\mathrm{a}}(y | \mathscr{M}, \mathscr{I}_X^{\mathrm{A}}, \mathscr{I}_Y^{\mathrm{A}})$ and $f_{\mathrm{b}}(y | \mathscr{M}, \mathscr{I}_X^{\mathrm{A}}, \mathscr{I}_Y^{\mathrm{A}})$ will usually have nearly equal standard deviations, their means being only slightly displaced from one another. Thus, simply choosing the one PDF $f_{\mathrm{a}}(y | \mathscr{M}, \mathscr{I}_X^{\mathrm{A}}, \mathscr{I}_Y^{\mathrm{A}})$ or $f_{\mathrm{b}}(y | \mathscr{M}, \mathscr{I}_X^{\mathrm{A}}, \mathscr{I}_Y^{\mathrm{A}})$ that better suits the application at hand may suffice.

3. GAUSSIAN LIKELIHOODS

Non-informative priors were mentioned many times above. Their construction may be effected by different principles [15] and will not be addressed in this article, except for the brief discussion that follows. It refers to the case of Gaussian likelihoods, which appear frequently in metrology. These likelihoods arise when type A information about a certain quantity, say X, consists of n_X samples collected from a Gaussian sampling distribution of mean X and standard deviation Σ_X. If the mean of these samples is $\bar{x}$ and their standard deviation is s_X, the likelihood takes the form [16, eq. 1.4.2]

$$\ell(x, \sigma_X; \mathscr{I}_X^{\mathrm{A}}) \propto \frac{1}{\sigma_X^{n_X}} \exp\left[-\frac{(n_X - 1)s_X^2 + n_X(x - \bar{x})^2}{2\sigma_X^2}\right], \tag{22}$$

where σ_X represents the possible values of Σ_X.

Suppose information $\mathscr{I}_X^{\mathrm{B}}$ is available. It may or may not include that about Σ_X. If it did not, and if X and Σ_X were taken as independent, one would use

$$f(x, \sigma_X | \mathscr{I}_X^{\mathrm{B}}) \propto \frac{1}{\sigma_X} f(x | \mathscr{I}_X^{\mathrm{B}}), \tag{23}$$

as the prior to be updated by the likelihood (22) [8, 15].

The case when there is no type B information about X (nor about Σ_X) has been amply studied [6, 17]. The accepted non-informative prior is in this case [16]

$$f(x, \sigma_X) \propto \frac{1}{\sigma_X}, \tag{24}$$

which corresponds to (23) with $f(x | \mathscr{I}_X^{\mathrm{B}}) = \text{const.}$

In contrast, the appropriate non-informative prior $f(y, \boldsymbol{z}, \sigma_X)$ to be updated by a likelihood of the form $\ell(G, \boldsymbol{p}_X; \mathscr{I}_X^{\mathrm{A}})$ (equation 8) has been the subject of some controversy. In [17], Bodnar et al studied the following three options for this prior:

$$f(y, \boldsymbol{z}, \sigma_X) \propto \frac{1}{\sigma_X} f(\boldsymbol{z}), \tag{25}$$

$$f(y, \boldsymbol{z}, \sigma_X | \mathscr{M}) \propto \frac{1}{\sigma_X} \left|\frac{\partial G}{\partial y}\right| \phi(\boldsymbol{z}) f(\boldsymbol{z}), \tag{26}$$

and

$$f(y, \boldsymbol{z}, \sigma_X | \mathscr{M}) \propto \frac{1}{\sigma_X} \left|\frac{\partial G}{\partial y}\right| f(\boldsymbol{z}), \tag{27}$$

where function $\phi(\boldsymbol{z})$ in (26) is given by equation (7) in [17].

Bodnar et al found that prior (25) may fail to produce a proper posterior $f(y | \mathscr{M}, \mathscr{I}_X^{\mathrm{A}})$, so it is definitively inadvisable. They also found that in some cases prior (26) fails as well to produce proper posteriors. That prior, which had been proposed by us [18], was alas based on an inadvisable sequence of subsets in the general calculation procedure developed in [8]. We detected this mishap and corrected it [19], arriving at prior (27). Unfortunately, our revision came too late for the authors of [17], otherwise prior (26) would not have been considered at all. Last, Bodnar et al showed that the marginal posterior $f(y | \mathscr{M}, \mathscr{I}_X^{\mathrm{A}})$ is always proper if prior (27) is used. This result was to be expected, because said prior is the transformed version of (24), which is known not to cause any difficulties in the case of a Gaussian likelihood.

4. EXAMPLE

As a simple illustration of the analysis in section 2, consider the determination of the average diameter Y of metallic-coated hollow glass micro-spheres for use in a particle image velocimetry (PIV) application. Two methods were applied: direct measurement of the diameter of these objects through microscope image processing, and observation of their terminal velocity X when falling in tap water at 20 °C. Under the Stokes' regime, this velocity is given by

$$X = \frac{gY^2(\rho - \rho_{\mathrm{w}})}{18\mu_{\mathrm{w}}}, \tag{28}$$

where g is the acceleration due to gravity, μ_{w} and ρ_{w} are the dynamic viscosity and density of water, and ρ is the apparent density of the hollow spheres [20, eq. 4.9.20]. This model, which is valid if the Reynolds number $\mathrm{Re} = XY\rho_{\mathrm{w}}/\mu_{\mathrm{w}}$ is much less than 1, may be written as

$$Y = Z\sqrt{X}, \tag{29}$$

where

$$Z = 3\sqrt{\frac{2\mu_{\mathrm{w}}}{g(\rho - \rho_{\mathrm{w}})}}. \tag{30}$$
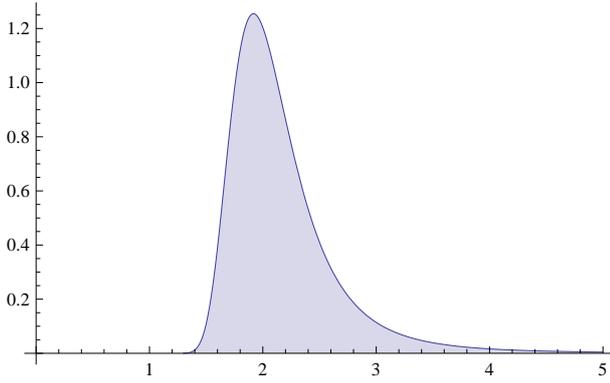
Fig. 1: PDF for the quantity Z defined by (30). The horizontal axis represents the possible values z of Z, in $(\mu m\ s)^{1/2}$.

Although the properties of pure water as functions of temperature are known quite precisely, in this example their uncertainty should be considered due to the influence of chlorine, fluorides, dissolved minerals and other impurities, and also because of the uncertainty in the measurement of the water temperature. For simplicity, however, we shall use the accepted values at 20 °C, $\mu_w = 1.00 \times 10^{-3}$ Pa s and $\rho_w = 998$ kg/m^3, and shall assume that they are exactly known. The uncertainty on the value of g will also be neglected.

Information $\mathscr{I}_X^A$ consists of $n_X = 10$ measurements of the terminal velocity with mean $\bar{x} = 22.5$ μm/s and standard deviation $s_X = 4.6$ μm/s. In turn, information $\mathscr{I}_Y^A$ comprises $n_Y = 7$ direct measurements of the diameter, giving a mean $\bar{y} = 10.5$ μm and standard deviation $s_Y = 2.3$ μm. From these results, the Reynolds number is of the order of 10^{-4}, so model (29) can be applied safely.

In its brochure, the manufacturer states that the diameter of the spheres is within (12 ± 3) μm, so for $f\left(y \mid \mathscr{I}_Y^B\right)$ a rectangular PDF with support on this interval will be used.

Type B information for X exists in the form of the opinion of an expert according to whom measurements of the terminal velocity outside of the interval ranging from $x_l = 17$ μm/s to $x_u = 25$ μm/s could be taken as being definitely wrong. Therefore, we shall assume that $f\left(x \mid \mathscr{I}_X^B\right)$ is a rectangular distribution supported on the interval bounded by those limits.

Finally, the manufacturer's certificate for the batch of micro-spheres specifies that the best estimate of their apparent density is 1430 kg/m^3 with an associated standard uncertainty of 150 kg/m^3. Hence, it is reasonable to use for ρ a Gaussian PDF whose mean and standard deviation correspond to these values, respectively. Transforming variables, we get for Z the PDF $f(z)$ depicted in figure 1. Its left tail goes to zero very rapidly, so its support can be assumed to be the positive axis.

4.1. PDFs in scenarios 4 and 8

Let us start the analysis of this example by considering separately both types of information for Y. We have already said that $f_8\left(y \mid \mathscr{I}_Y^B\right)$ is a rectangular distribution, where for clarity a subscript has been introduced for denoting correspondence
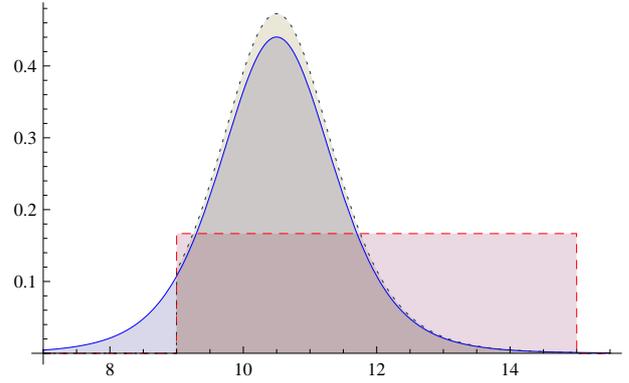
Fig. 2: PDFs for the measurand Y in scenarios 4 (solid line), 8 (dashed line) and 12 (dotted line). In this and all remaining plots, the horizontal axis represents the possible values y of Y, in micrometers.
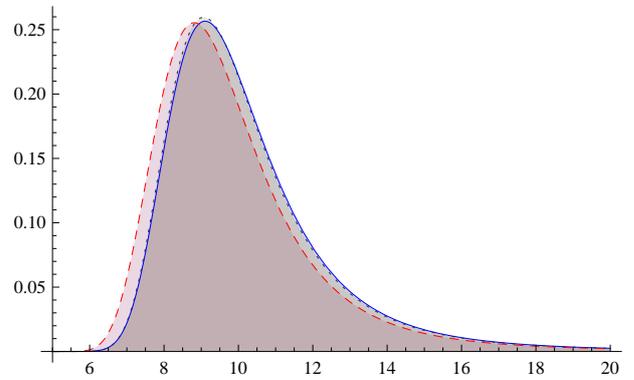
Fig. 3: PDFs for Y in scenarios 1 (solid line), 2 (dashed line) and 3 (dotted line).

with scenario 8 (the same will be done for the rest of the scenarios). This PDF is plotted in figure 2 (dashed line).

If only $\mathscr{I}_Y^A$ is taken into account we use (10) in the form

$$f_4\left(y \mid \mathscr{I}_Y^A\right) \propto \int_0^\infty \ell\left(y, \sigma_Y; \mathscr{I}_Y^A\right) \frac{d\sigma_Y}{\sigma_Y}, \qquad (31)$$

where prior (24) has been used. Since the likelihood is Gaussian, the integral yields for f_4 a t-distribution with $n_Y - 1$ degrees of freedom, location parameter $\bar{y}$ and scale parameter $s_Y / \sqrt{n_Y}$ [16, p. 97], see figure 2 (solid line).

4.2. PDFs in scenarios 1 and 2

Consider now separately both types of information for X. When only $\mathscr{I}_X^B$ is taken into account, (2) gives

$$f_2\left(y \mid \mathscr{M}, \mathscr{I}_X^B\right) \propto y \int_{a(y)}^{b(y)} \frac{f(z)}{z^2}\, dz, \qquad (32)$$

where $a(y) = y/\sqrt{x_u}$ and $b(y) = y/\sqrt{x_l}$. The dashed line in figure 3 shows PDF f_2.
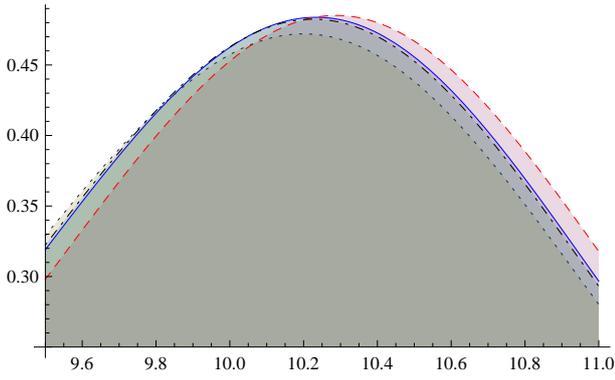
Fig. 4: PDFs for Y in scenarios 5a (solid line), 5b (dashed line), 6 (dotted line) and 7 (dot-dashed line).
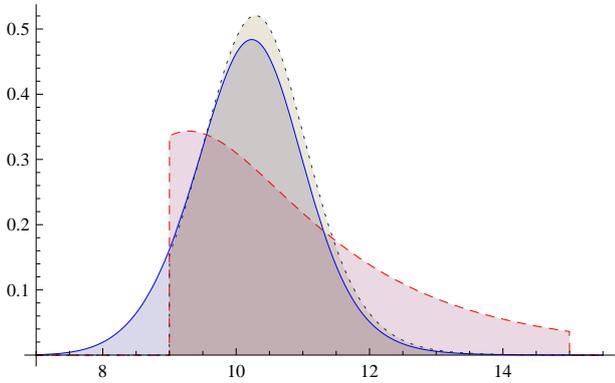
Fig. 5: PDFs for Y in scenarios 5a (solid line), 9 (dashed line) and 13 (dotted line).

If only $\mathscr{I}_X^A$ is considered we use (8) and (27), giving

$$
f\left(y, z \mid \mathscr{M}, \mathscr{I}_X^A\right) \propto \\
\frac{y}{z^2} f(z) \int_0^\infty \ell\left(\frac{y^2}{z^2}, \sigma_X; \mathscr{I}_X^A\right) \frac{\mathrm{d}\sigma_X}{\sigma_X}, \quad (33)
$$

from which

$$
f_1\left(y \mid \mathscr{M}, \mathscr{I}_X^A\right) \propto y \int_0^\infty f\left(\frac{y^2}{z^2} \mid \mathscr{I}_X^A\right) \frac{f(z)}{z^2} \, \mathrm{d}z, \quad (34)
$$

where $f\left(x \mid \mathscr{I}_X^A\right)$, of which ample use will be made in the following, is a t-distribution with $n_X - 1$ degrees of freedom, location parameter $\bar{x}$ and scale parameter $s_X / \sqrt{n_X}$. Figure 3 depicts f_1 as a solid line.

4.3. PDFs in scenarios 12 and 13

If information $\mathscr{I}_X^A$ is considered together with $\mathscr{I}_Y^B$, from (23) we see that the prior to be updated by $\ell\left(y, \sigma_Y; \mathscr{I}_Y^A\right)$ is f_8 / σ_Y. It is then immediate that $f_{12}\left(y \mid \mathscr{I}_Y^{A+B}\right)$ is equal to the normalized product of the t-distribution f_4 with the rectangular distribution f_8, see figure 2 (dotted line).

In scenario 13 we use (17) with $f(\sigma_X) \propto 1/\sigma_X$. Integrating

out σ_X and z, we get

$$
f_{13}\left(y \mid \mathscr{M}, \mathscr{I}_X^A, \mathscr{I}_Y^{A+B}\right) \propto \\
f_{12}\left(y \mid \mathscr{I}_Y^{A+B}\right) \int_0^\infty f\left(\frac{y^2}{z^2} \mid \mathscr{I}_X^A\right) f(z) \, \mathrm{d}z, \quad (35)
$$

which appears as a dotted line in figure 5.

4.4. PDFs in scenario 5

Evidently, PDFs f_{12} and f_{13} are much influenced by the type B information on Y, which causes the cut-off at $y = 9$ μm (and the imperceptible one at $y = 15$ μm). Therefore, it is of interest to explore scenario 5, which includes information $\mathscr{I}_X^A$ and $\mathscr{I}_Y^A$ only. As discussed in subsection 2.6, in this case there are two possible approaches. In approach a) we start from the limit of (16) as $\mathscr{I}_X^B$ vanishes, which gives

$$
f\left(y, z \mid \mathscr{M}, \mathscr{I}_X^A, \mathscr{I}_Y^A\right) \propto \\
f\left(y, z \mid \mathscr{M}, \mathscr{I}_X^A\right) \int_0^\infty \ell\left(y, \sigma_Y; \mathscr{I}_Y^A\right) \frac{\mathrm{d}\sigma_Y}{\sigma_Y}. \quad (36)
$$

The expression for $f\left(y, z \mid \mathscr{M}, \mathscr{I}_X^A\right)$ is given by (33) and the integral on the right-hand side is just f_4. Thus,

$$
f_{5a}\left(y \mid \mathscr{M}, \mathscr{I}_X^A, \mathscr{I}_Y^A\right) \propto \\
y f_4\left(y \mid \mathscr{I}_Y^A\right) \int_0^\infty f\left(\frac{y^2}{z^2} \mid \mathscr{I}_X^A\right) \frac{f(z)}{z^2} \, \mathrm{d}z, \quad (37)
$$

which is portrayed in figure 4 (solid line).

In approach b), (17) with vanishing $\mathscr{I}_Y^B$ gives directly

$$
f\left(y, z \mid \mathscr{M}, \mathscr{I}_X^A, \mathscr{I}_Y^A\right) \propto \\
f_4\left(y \mid \mathscr{I}_Y^A\right) f(z) \int_0^\infty \ell\left(\frac{y^2}{z^2}, \sigma_X; \mathscr{I}_X^A\right) \frac{\mathrm{d}\sigma_X}{\sigma_X}, \quad (38)
$$

from which

$$
f_{5b}\left(y \mid \mathscr{M}, \mathscr{I}_X^A, \mathscr{I}_Y^A\right) \propto \\
f_4\left(y \mid \mathscr{I}_Y^A\right) \int_0^\infty f\left(\frac{y^2}{z^2} \mid \mathscr{I}_X^A\right) f(z) \, \mathrm{d}z. \quad (39)
$$

This alternative PDF is also shown in figure 4 (dashed line).

4.5. PDFs in scenarios 3 and 6

We have already mentioned that f_{12} is equal to the normalized product of f_4 and f_8. For the same reason

$$
f\left(x \mid \mathscr{I}_X^{A+B}\right) \propto f\left(x \mid \mathscr{I}_X^A\right) f\left(x \mid \mathscr{I}_X^B\right) \quad (40)
$$

and therefore, since $f\left(x \mid \mathscr{I}_X^B\right)$ is a uniform distribution,

$$
f_3\left(y \mid \mathscr{M}, \mathscr{I}_X^{A+B}\right) \propto y \int_{a(y)}^{b(y)} f\left(\frac{y^2}{z^2} \mid \mathscr{I}_X^A\right) \frac{f(z)}{z^2} \, \mathrm{d}z, \quad (41)
$$

which appears in figure 3 as a dotted line.
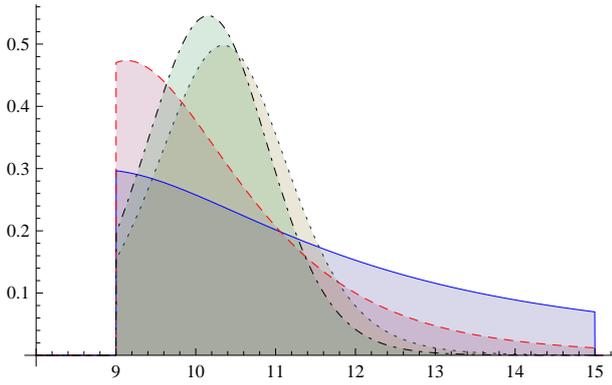
Fig. 6: PDFs for Y in scenarios 10 (solid line), 11 (dashed line), 14 (dotted line) and 15 (dot-dashed line).

In scenario 6, we use the likelihood $\ell\left(y, \sigma_Y; \mathscr{I}_Y^A\right)$ for updating the prior f_2 and so obtain

$$f_6\left(y \,|\, \mathscr{M}, \mathscr{I}_X^B, \mathscr{I}_Y^A\right) \propto y\, f_4\left(y \,|\, \mathscr{I}_Y^A\right) \int_{a(y)}^{b(y)} \frac{f(z)}{z^2}\, \mathrm{d}z, \quad (42)$$

which is shown as a dotted line in figure 4.

4.6. PDFs in scenarios 7 and 9

Consider now disregarding the manufacturer's information $\mathscr{I}_Y^B$ while retaining information $\mathscr{I}_X^{A+B}$ together with $\mathscr{I}_Y^A$. Then, (16) yields simply

$$f_7\left(y \,|\, \mathscr{M}, \mathscr{I}_X^{A+B}, \mathscr{I}_Y^A\right) \propto f_4\left(y \,|\, \mathscr{I}_Y^A\right) f_3\left(y \,|\, \mathscr{M}, \mathscr{I}_X^{A+B}\right), \quad (43)$$

which is depicted as a dot-dashed line in figure 4.

If the measurement information $\mathscr{I}_Y^A$ is disregarded and the manufacturer's information $\mathscr{I}_Y^B$ is kept together with the measurement information $\mathscr{I}_X^A$, (15) gives

$$f\left(y, z \,|\, \mathscr{M}, \mathscr{I}_X^A, \mathscr{I}_Y^B\right) \propto \\ f_8\left(y \,|\, \mathscr{I}_Y^B\right) f(z) \int_0^\infty \ell\left(\frac{y^2}{z^2}, \sigma_X; \mathscr{I}_X^A\right) \frac{\mathrm{d}\sigma_X}{\sigma_X}, \quad (44)$$

from which

$$f_9\left(y \,|\, \mathscr{M}, \mathscr{I}_X^A, \mathscr{I}_Y^B\right) \propto \\ f_8\left(y \,|\, \mathscr{I}_Y^B\right) \int_0^\infty f\left(\frac{y^2}{z^2} \,|\, \mathscr{I}_X^A\right) f(z)\, \mathrm{d}z. \quad (45)$$

This PDF is shown in figure 5 (dashed line), where we see that the cut-off at the right is now much more visible than that for scenario 12 (figure 2, dotted line).

4.7. PDFs in scenarios 10, 11, 14 and 15

As mentioned in subsection 2.5, in those scenarios for which there is type B information about both X and Y we should start by merging the two PDFs that encode just this information, possibly with different pooling weights if one judges that, say, the manufacturer's information is more reliable than

the expert's. But one still has to choose between logarithmic and linear pooling. Here the decision is easy, because the support of f_8 (figure 2, dashed line) is contained within that of f_2 (figure 3, dashed line), implying that linear pooling would produce an unreasonable discontinuity at the boundaries of the former PDF. Hence, since logarithmic pooling appears to be the better choice, we shall use

$$f_{10}\left(y \,|\, \mathscr{M}, \mathscr{I}_X^B, \mathscr{I}_Y^B\right) \propto \left[f_2\left(y \,|\, \mathscr{M}, \mathscr{I}_X^B\right)\right]^w \left[f_8\left(y \,|\, \mathscr{I}_Y^B\right)\right]^{1-w}. \quad (46)$$

The solid line in figure 6 illustrates this PDF with $w = 0.5$, since we judge that there is no reason to assign more importance to the manufacturer's information than to the expert's.

In scenario 11, f_{10} is updated using the likelihood $\ell\left(x, \sigma_X; \mathscr{I}_X^A\right)$. By following the same steps that led to (45) we then get

$$f_{11}\left(y \,|\, \mathscr{M}, \mathscr{I}_X^{A+B}, \mathscr{I}_Y^B\right) \propto \\ f_{10}\left(y \,|\, \mathscr{M}, \mathscr{I}_X^B, \mathscr{I}_Y^B\right) \int_0^\infty f\left(\frac{y^2}{z^2} \,|\, \mathscr{I}_X^A\right) f(z)\, \mathrm{d}z, \quad (47)$$

which appears as a dashed line in figure 6.

In scenario 14 we use the likelihood $\ell\left(y, \sigma_Y; \mathscr{I}_Y^A\right)$ for updating f_{10}. As in scenario 12, we obtain in this way

$$f_{14}\left(y \,|\, \mathscr{M}, \mathscr{I}_X^B, \mathscr{I}_Y^{A+B}\right) \propto f_{10}\left(y \,|\, \mathscr{M}, \mathscr{I}_X^B, \mathscr{I}_Y^B\right) f_4\left(y \,|\, \mathscr{I}_Y^A\right), \quad (48)$$

which is shown as a dotted line in figure 6.

Finally, in scenario 15, either this last PDF is updated by means of $\ell\left(x, \sigma_X; \mathscr{I}_X^A\right)$, giving

$$f_{15}\left(y \,|\, \mathscr{M}, \mathscr{I}_X^{A+B}, \mathscr{I}_Y^{A+B}\right) \propto \\ f_{14}\left(y \,|\, \mathscr{M}, \mathscr{I}_X^B, \mathscr{I}_Y^{A+B}\right) \int_0^\infty f\left(\frac{y^2}{z^2} \,|\, \mathscr{I}_X^A\right) f(z)\, \mathrm{d}z, \quad (49)$$

or the PDF obtained for scenario 11 is updated by means of $\ell\left(y, \sigma_Y; \mathscr{I}_Y^A\right)$, which yields

$$f_{15}\left(y \,|\, \mathscr{M}, \mathscr{I}_X^{A+B}, \mathscr{I}_Y^{A+B}\right) \propto \\ f_{11}\left(y \,|\, \mathscr{M}, \mathscr{I}_X^{A+B}, \mathscr{I}_Y^B\right) f_4\left(y \,|\, \mathscr{I}_Y^A\right). \quad (50)$$

Equations (49) and (50) are seen to be identical by substituting (48) in the former and (47) in the latter. This reflects the well-known fact that it is irrelevant in which order updating takes place. The PDF encoded by (49) or (50) is depicted as a dot-dashed line in figure 6.

4.8. Discussion

Table 2 lists the means and standard deviations of the various PDFs that result from the analysis. The first group of scenarios in this table are those that take into account information about Y only, so to obtain the corresponding PDFs the measurement model was not considered and the information about X and Z was disregarded. These PDFs are plotted in figure 2: a t-distribution for scenario 4, a rectangular distribution for

| Scenario | Information type | | Mean | Std. dev. |
	on X	on Y		
4		A	10.50	1.06
8		B	12.00	1.73
12		A & B	10.65	0.88
1	A		10.44	4.28
2	B		10.08	4.44
3	A & B		10.40	4.54
5a	A	A	10.22	0.88
5b	A	A	10.29	0.88
6	B	A	10.15	0.90
7	A & B	A	10.20	0.88
9	A	B	10.93	1.46
13	A	A & B	10.41	0.77
10	B	B	10.26	1.62
11	A & B	B	10.48	1.22
14	B	A & B	10.48	0.81
15	A & B	A & B	10.28	0.72

Table 2: Means and standard uncertainties of the PDFs for Y, in micrometers, derived in the various scenarios.

scenario 8 and the normalized product of the first two for scenario 12. Therefore, since f_{12} is the truncated version of f_4, it is natural to expect that the standard deviation of the former should be less than that of the latter.

The next group of scenarios, 1, 2 and 3, are those of the GUM type, that is, those that include information on the input quantities X and Z (or rather, X and ρ) but not on Y, so they were obtained by propagating analytically the PDFs for the input quantities through the measurement model. The corresponding PDFs, depicted in figure 3, are all very similar. It turns out that their standard deviations are much larger than those in the other scenarios, which take into account direct information on the output quantity. This is a consequence of the vagueness of the information on X and Z provided. If the information on these quantities were more precise than that on Y, an opposite picture would have resulted. It is interesting to note that (34) for f_1 and (41) for f_3 differ only in the limits of the integrals, which in the latter case arise from the expert's information $\mathscr{I}_X^B$. The fact that these two PDFs have similar means and standard deviations shows that the expert's opinion is consistent with the rest of the information. It is rather imprecise, however, so his contribution is not determinant.

Scenarios 5, 6, and 7 make up the group of those that share type A information on Y but differ in the type of information considered for X. We note that in this example the two ways of treating scenario 5 yield almost indistinguishable PDFs, for this reason, only their central portions are plotted in figure 4. This figure depicts also the PDFs in scenarios 6 and 7. The fact that these two PDFs are quite similar to that for scenario 5 again reveals that information pieces $\mathscr{I}_X^A$ and $\mathscr{I}_X^B$ in

conjunction with information on Z are unimportant in comparison with $\mathscr{I}_Y^A$.

The PDFs for scenarios 9 and 13, depicted in figure 5, are both restricted to the interval indicated by the manufacturer for the diameter of the spheres. These PDFs differ in that the latter includes the type A information for Y, which causes a marked change in shape (from the dashed line to the dotted one) and a substantial reduction of the standard deviation.

Finally, the PDFs for scenarios 10, 11, 14 and 15 are shown in figure 6. They are all based on type B information about both X and Y. In scenario 10 there is no type A information, and the corresponding PDF, f_{10}, was obtained by merging those for scenarios 2 and 8. Since for the reason explained above we chose the logarithmic pooling technique, the support of f_{10} gets restricted to that of f_8. This fact is determinant, for then the PDFs for scenarios 11, 14 and 15, which use f_{10} as prior, share the same restriction. Thus, f_{11} was obtained by updating this prior with the likelihood for X, and for obtaining f_{14} we used the likelihood for Y. Either of these PDFs was updated with the remaining likelihood to produce f_{15}.

This last PDF, which incorporates all information, exhibits the smallest standard deviation. However, this is due to the sharp drop to zero at $y = 9$ μm. So unless one were completely certain of the limits given by the manufacturer – especially the lower one – it would appear more sensible to drop that information altogether, which is equivalent to assigning zero weight to $\mathscr{I}_Y^B$. The PDF for scenario 15 would then change abruptly to that for scenario 7 (or to one of those for scenario 5 if the expert's opinion were also to be disregarded). On the other hand, it might be that the manufacturer is right and that one or both instruments used to obtain the type A information are biased toward the low side. But to decide on this possibility, more information would be required.

In short, to produce a PDF for the measurand, the basic principle is that one should utilize all existing information, unless there are good reasons to consider some elements of that information as doubtful.

5. CONCLUSIONS

When evaluating physical models where a quantity of interest is expressed as a function of other quantities, it may happen that information on the former exists before the model is taken into account. Neither the GUM nor its supplements address such situations although they are important in practice. We have done so in this paper through an example involving just three quantities, X, Y and Z, where the second one was assumed to be that of interest.

It was found that in obtaining the PDF for Y it is essential to distinguish between the two types of information that may be available on any of these quantities: that arising from direct measurement data (type A) and that originating from other sources, such as some expert's opinion (type B).

We assumed that only type B information existed for Z, and that either type or both together were available for X and

Y. The analysis required extensive change of variables and – only in case type A information was present – of Bayes' rule. As is well known, the latter allows updating from prior to posterior distributions by means of likelihoods, either one at a time (the first measurement, then the next one, and so on) or by constructing the product of likelihoods for a given number of measurements from the outset.

In scenarios 1, 3, 4, 9 and 11 there was just one such product to contemplate. But in scenarios 5, 7, 13 and 15 it was necessary to treat likelihoods arising from distinct sources. In these cases our approach was, simply, to update a prior with one likelihood and then use the result to be updated with the other likelihood. To our knowledge, this kind of procedure has not been described in metrological applications yet, although it is accepted practice in statistics [21].

Finally, in scenarios 10, 11, 14 and 15, type B information was at hand for all quantities in the model. It was shown that this situation requires merging two concurrent PDFs obtained from that information, and that the PDF that results from the merging procedure can then be updated with whatever likelihoods exist, if any.

References

[1] Joint Committee for Guides in Metrology. (2008). *Evaluation of measurement data Guide to the expression of uncertainty in measurement*. JCGM 100:2008 (GUM 1995 with minor corrections).

[2] Bich, W., Cox, M., Dybkaer, R., Elster, C., Estler, W., Hibbert, B., Imai, H., Kool, W., Michotte, C., Nielsen, L., Pendrill, L., Sidney, S., van der Veen, A., Woger, W. (2012). Revision of the 'Guide to the expression of uncertainty in measurement'. *Metrologia*, 49, 702-705.

[3] Bich, W. (2014). Revision of the 'Guide to the expression of uncertainty in measurement'. Why and how. *Metrologia*, 51, S155–S158.

[4] Lira, I., Grientschnig, D. (2011). Non-informative priors in GUM Supplement 1. *Measurement* 44, 1790–1791.

[5] Lira, I., Grientschnig, D. (2012). Assignment of a non-informative prior when using a calibration function. *Measurement Science and Technology* 23, 712–719.

[6] Elster, C., Toman, B. (2009). Bayesian uncertainty analysis under prior ignorance of the measurand versus analysis using the Supplement 1 to the Guide: A comparison. *Metrologia* 46, 261–266.

[7] Lira, I., Grientschnig, D. (2013). A formalism for expressing the probability density functions of interrelated quantities. *Measurement Science Review* 13, 50–55.

[8] Sun, D., Berger, J. O. (1998). Reference priors with partial information. *Biometrika* 85, 55–71.

[9] Lira, I., Grientschnig, D. (2010). Equivalence of alternative Bayesian procedures for evaluating measurement uncertainty. *Metrologia* 47, 334–336.

[10] Possolo, A., Toman, B. (2007). Assessment of measurement uncertainty via observation equations. *Metrologia* 44, 464–475.

[11] Elster, C. (2014). Bayesian uncertainty analysis compared with the application of the GUM and its supplements. *Metrologia* 51, S159–S166.

[12] Lira, I., Grientschnig, D. (2014). Deriving PDFs for interrelated quantities: What to do if there is 'more than enough' information? *IEEE Transactions on Instrumentation and Measurement* 63, 1937–1946.

[13] Clemen, R. T., Winkler, R. L. (1999). Combining probability distributions from experts in risk analysis. *Risk Analysis* 19, 187–203.

[14] Grientschnig, D., Lira, I. (2014). Combining probability distributions by multiplication in metrology: A viable method? *International Statistical Review* 82, 392–410.

[15] Berger, J. O., Bernardo, J. M., Sun, D. (2009). The formal definition of reference priors. *Annals of Statistics* 37, 905–938.

[16] Box, G. E. P., Tiao, G. C. (1973). *Bayesian Inference in Statistical Analysis*. (Reprinted by Wiley in 1992 in the Wiley Classics Library Edition.)

[17] Bodnar, O., Wübbeler, G., Elster, C. (2011). On the application of Supplement 1 to the GUM to non-linear problems. *Metrologia* 48, 333–342.

[18] Grientschnig, D., Lira, I. (2011). Reassessment of a calibration model by Bayesian reference analysis. *Metrologia* 48, L7–L11.

[19] Grientschnig, D., Lira, I. (2012). Revision of 'Reassessment of a calibration model by Bayesian reference analysis'. *Metrologia* 49, L1–L3.

[20] Batchelor, G. K. (1981). *An Introduction to Fluid Dynamics*. Cambridge University Press.

[21] Caticha, A. (2008). Lectures on probability, entropy, and statistical physics. *arXiv:0808.0012 [physics.data-an]*.