XIX IMEKO World Congress
Fundamental and Applied Metrology
September 6–11, 2009, Lisbon, Portugal

THE CHOICE OF METHOD TO THE EVALUATION OF MEASUREMENT UNCERTAINTY IN METROLOGY

J Alves Sousa [1] and *A Silva Ribeiro* [2]

[1] Laboratório Regional de Engenharia Civil, 9000-264 Funchal, Portugal, jasousa@lrec.pt
[2] Laboratório Nacional de Engenharia Civil, 1700-066 Lisboa, Portugal, asribeiro@lnec.pt

Abstract – This paper discusses the selection of appropriate uncertainty framework in metrology related to the class of problem to be solved.

Keywords: solution method, measurement uncertainty, metrology

1. INTRODUCTION

This paper is concerned with the need to ensure that an appropriate uncertainty framework is selected when determining the measurement uncertainty of a particular problem in metrology. The GUM uncertainty framework [1] is indubitably the most widely used method and thus its adequacy will always be tested against more elaborate methods. It will be attempted to establish some simple guidance rules based on the selection parameters.

Although the theoretical grounds for the application of the mainstream GUM are well defined, they are often overlooked and will result in inadequate applications. On the other hand, situations exist where it is known that GUM provides accurate results despite the fact that not all requirements for its application are met. It would therefore be useful to have some knowledge on the factors that influence mostly the outcome and adequacy of GUM applications. This methodology requires proper validation tools and a Monte Carlo method (MCM) will generally be used for that purpose.

Another important approach to the evaluation of measurement uncertainty is based on Bayesian methods. Its fundamentals will concisely be explained and the merits of its application will be discussed.

The differences between approaches will be explored and a comparison between GUM, MCM and Bayesian methods will be drawn, based on examples of different classes of typical metrology problems. The objective of generic method selection guidelines will be attempted.

2. APPROACH

Different approaches can be used to provide a best estimate of the measurand and the associated measurement uncertainty, and a coverage interval for the measurand for a prescribed coverage probability.

This is the whole set of information that GUM uncertainty framework can provide. It operates with summary information derived from the probability density functions (PDFs) for the input quantities, such as best estimates of the input quantities and the associated standard uncertainties, and through a Taylor expansion of a functional relationship will provide the required parameters – best estimate and coverage interval – associated with the measurand.

A Monte Carlo method (MCM) implements the propagation of distributions [2] by sampling from the PDFs for the input quantities to provide an output PDF for the measurand. From this PDF the statistics parameters associated with the measurand can readily be obtained. This latter distinction in comparison with the mainstream GUM, represents an important advantage as it will be shown later.

Bayesian methods, on the other hand, can also incorporate a prior PDF for the measurand in its probabilistic formulation, accounting for previous knowledge, e.g., physical knowledge on the output quantity, which can be relevant when, for example, physical limitations to the outcome result are known. As will be illustrated with examples, this feature of the method can determine its selection as the best suitable approach for some classes of problem.

Considering that all methods have a process based on two stages, called formulation stage and calculation stage, and that they share similar requirements on the information needed for the formulation stage (the mathematical model and the PDFs of the input variables), the main differences that can define its suitability to each metrological problem are necessarily connected with the calculation stage requirements.

In this way, a main task is to identify the relevant characteristics of metrological problems and the constraints of the evaluation methods, taking this information as a basis to aggregate these metrological problems under similar conditions to allow a classification suitable to act as guidance to the metrologist.

3. DISCUSSION

The selection of an appropriate methodology for the evaluation of measurement uncertainties is, in certain circumstances, preponderant for the correctness of that evaluation with respect to the physical reality it intends to represent [3].

ISBN 978-963-88410-0-1 © 2009 IMEKO

The mathematical models used as the support of that representation may differ in the number of variables and its combinations, some of which are particularly common in metrology, such as ratio, power and exponential expressions, by themselves or in some sort of combination. They will all predictably introduce some degree of non linearity or asymmetry in the output quantity whose influence needs to be studied.

However, the particular mathematical model will not define alone the best suited approach to its evaluation. Rather, the order of magnitude between uncertainties and the PDF associated with each of those input quantities will also have a very important role to play.

Generically, it can be stated that the analytical approach is appropriate to validate other methods, and should be applied whenever possible. Its main shortcoming lies in the scope of its applicability which is limited, in practice, to simple models. Therefore, its application in real life experiments is almost never considered.

The GUM uncertainty framework, on the other hand, is particularly suited to differentiable linear models, or with mild non linearity, symmetric input PDFs, and Central Limit Theorem conditions, or the level of approximation provided will be difficult to estimate.

Finally, the methods based on numerical simulations have a broader application, even to strongly non linear models, provide more information due to access to the outcome PDF and can converge rapidly to near exact solutions.

3.1 Ratios

As a first example we can look into a fairly simple problem of determining the measurement uncertainty associated with the estimate of a volumetric flow rate, where the measurand Q_v is given by

$$Q_v = \frac{a * b * h}{\Delta t} \tag{1}$$

Variables *a* and *b* are the width and length of the weighing tank, respectively, with assigned rectangular PDFs, and *h* is the liquid height in the weighing tank, having a Gaussian PDF. Lets assume that *a* and *b* have both the same limits [0,3495 – 0,3505] m, whereas *h* has a mean value of 0,08 m and an associated standard deviation of the mean of 0,0023 m. The time interval taken to fill the weighing tank is represented by Δ*t* and this variable can be crucial to the shape of the output PDF and thus to the validity of the GUM approach.

If one considers first that *Δt* is well represented by a Gaussian PDF with mean = 6,0 s and σ = 0,6 s the resulting output has a Gaussian shape as expected and the validity of the GUM is apparently unquestionable (see Figure 1). The validity holds for better (lower) values of uncertainty (standard deviation). However, as σ increases, the output PDF will collapse into a very narrow strip around zero, as Figures 2 and 3 illustrate for values of $\sigma = 1.05$ and $\sigma = 1.15$, and the assumption of Gaussian shape for the output PDF will not hold. The coverage interval may be overestimated when applying GUM in these conditions.

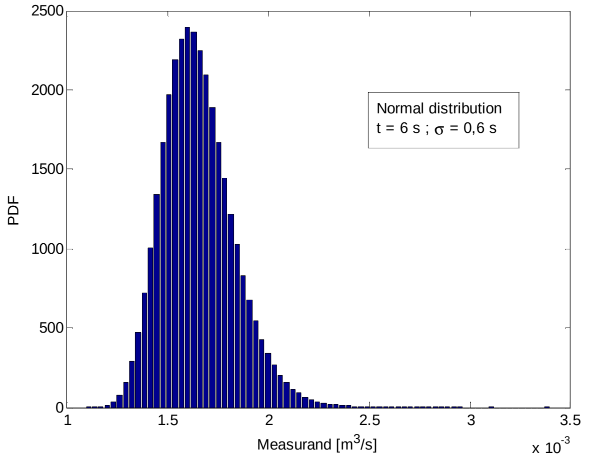

Figure 1 – Output PDF for input Δ*t* with Gaussian PDF.

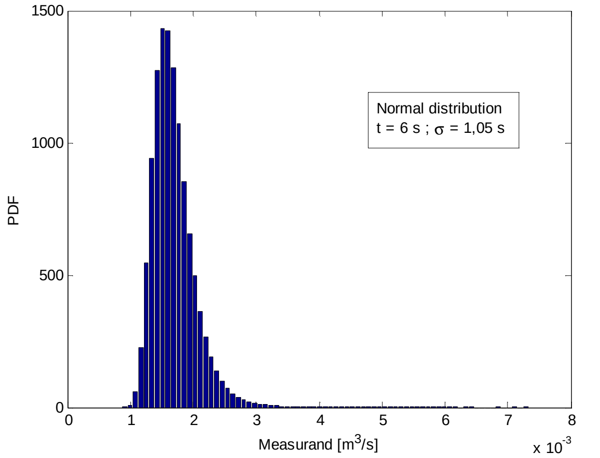

Figure 2 – Output PDF for input Δ*t* with Gaussian PDF.

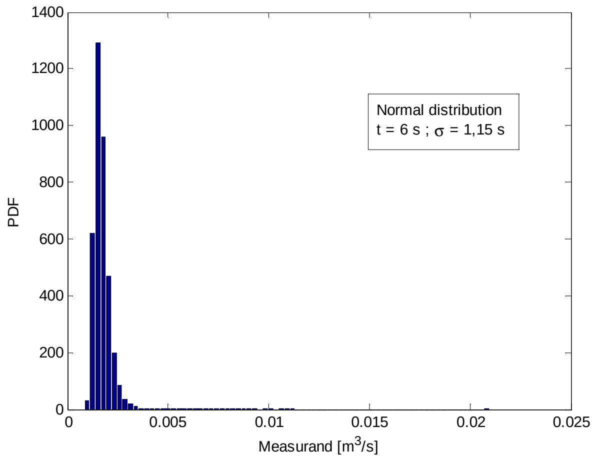

Figure 3 – Output PDF for input Δ*t* with Gaussian PDF.

Changing the shape and magnitude of this variable, e.g., considering that a rectangular PDF is instead assigned to it, will produce a much greater effect. In fact, for the same relative values of uncertainty, that is, with limits set by

[5,4 – 6,6] s and [4,85 – 7,15] s, the resulting output PDF departs considerably from a Gaussian distribution as Figures 4 and 5 show, and the uncertainty evaluation associated with the corresponding volumetric flow rate, using GUM or a MCM (Figure 2) approach are likely to produce rather different results. A further increase in the uncertainty associated with t, up to 50 % of its mean value, say, will lead to an exponential shape of the output PDF (Figure 6).

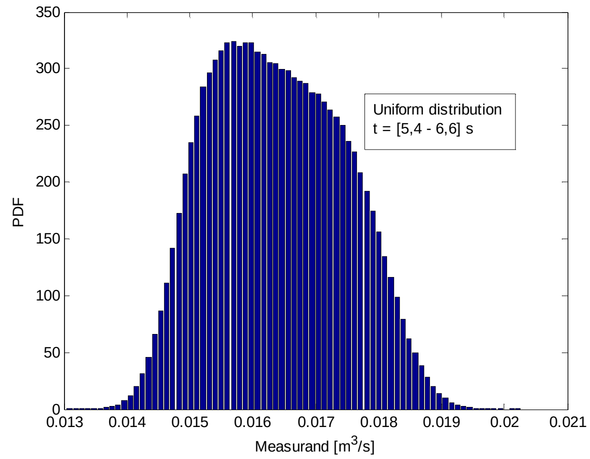

Figure 4 – Output PDF for input *t* with uniform PDF.

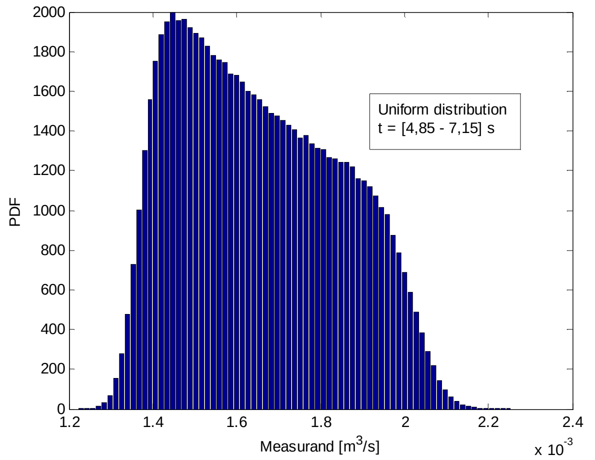

Figure 5 – Output PDF for input *t* with uniform PDF.

The uncomplicated nature of this example should not be a reason for a lesser impact. On the contrary, it is its broad range of applications, in areas such as chemical, volume and thermometry, that makes it a good example of how in a simple problem, things can easily go wrong.

In this particular problem, the factors that seem to influence mostly the output PDF are the input PDF of the variable on the denominator and the relative value of its measurement uncertainty. For relative values of uncertainty, in the referred variable, smaller or equal to 5 %, even having a uniform PDF will not produce an output PDF much different than the assumed Gaussian PDF.

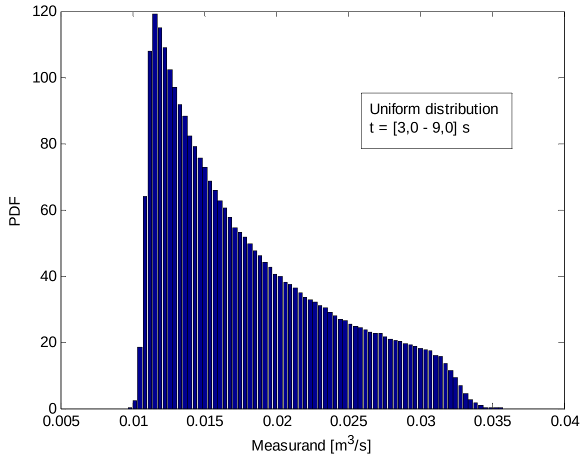

Figure 6 – Output PDF for input *t* with uniform PDF.

In this context, and bearing in mind this example, one is confronted with the fact that a number of variables can influence the uncertainty evaluation, so that ideally criteria to support the decision of method selection should be established.

3.2 Close to the physical limit

To illustrate the concept behind the section title, an example concerning the determination of an analyte concentration in chemical metrology is used. It will be the basis for a discussion concerning the various approaches available to deal with such constraints [4].

The problem is that it might not be possible, when using a conventional method for solving such a problem, to guarantee feasibility of the solution. For instance, an application of the GUM uncertainty framework [1] might provide 0.30 % as an estimate of a concentration and 0.25 % as the associated standard uncertainty. If the quantity concerned were characterized by a Gaussian distribution, the expanded uncertainty corresponding to a 95 % coverage probability would be 2 × 0.25 % = 0.50 %, and hence a 95 % coverage interval for the concentration would be (0.30 ± 0.50) %. Since the part of this interval that is below 0 % is infeasible, it is difficult to interpret this result in a meaningful way for an application. A correctly computed coverage interval would have no negative values and thus a lower limit equal or greater than 0 %.

This prior knowledge of the feasible interval for the output quantity can prove valuable in many instances. Bayes' theorem takes the form

$$g(\eta \mid x) = K\, l(x \mid \eta)\, g(\eta) \qquad (2)$$

where $g(\eta)$ is a prior PDF for *Y*, $l(x \mid \eta)$ is the likelihood function for the data *x*, $g(\eta \mid x)$ is the posterior PDF for *Y*, and *K* constitutes a normalization factor. In words, the degree of belief for a given value η of the measurand *Y*, expressed as the posterior PDF for *Y* given data *x*, is proportional to (a) the likelihood that η will produce the observed data *x*, and (b) the degree of belief attributed to η

before the observation, the so-called prior PDF for Y, expressed as $g(\eta)$.

The posterior PDF may be used to provide summary information about the measurand Y, such as its expectation (mean) $\mathrm{E}(Y)$ and variance $\mathrm{V}(Y)$, defined by

$$\mathrm{E}(Y) = \int \eta\, g(\eta|x)\,\mathrm{d}\eta, \quad \mathrm{V}(Y) = \mathrm{E}(Y - \mathrm{E}(Y))^2, \tag{6}$$

The prior distribution represents the information about the values η available before the measurement x was taken, while the posterior represents an aggregation of the prior information and that supplied by the data. In "data-rich" experiments, the information supplied by the data is much more comprehensive than the prior information, so that the posterior is essentially proportional to the likelihood. In other circumstances, the prior distribution can contain information that the data cannot supply.

Figure 7 shows, for the case x = 0.1 and $u(x)$ = 0.2, the solution PDFs provided by a Bayesian treatment and an application of MCM as a numerical implementation of the propagation of distributions. The height of the left-most bin is in fact greater than 20 rather than as shown. For purposes of illustration, the chosen scale was used.

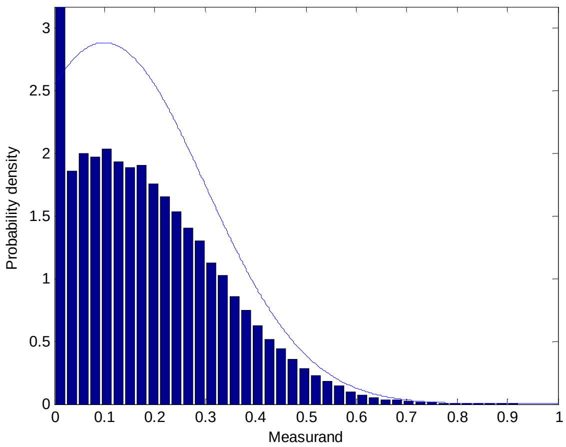

Figure 7 – Solution PDFs obtained from a Bayesian treatment (continuous curve) and a Monte Carlo method (scaled frequency distribution) corresponding to an estimate x = 0.1 and associated standard uncertainty $u(x)$ = 0.2.

Finally, Figure 8 shows, and compares, for the case x = 0.1 and $u(x)$ = 0.1, the results obtained from three of the approaches considered. The solution PDF provided by the GUM uncertainty framework is shown as the Gaussian distribution (thin continuous curve), that provided by the Bayesian treatment as the thicker curve, and that provided by the Monte Carlo method as a scaled frequency distribution. For each distribution the endpoints of the (shortest) 95 % coverage interval are shown, as broken, dotted and continuous vertical lines, respectively.

Both the Bayesian approach to the problem and the use of a Monte Carlo method deliver solution PDFs that are feasible. Furthermore, unlike the GUM uncertainty framework, neither the Bayesian approach nor the Monte Carlo method make an assumption about x or $u(x)$, e.g., that x is itself feasible.

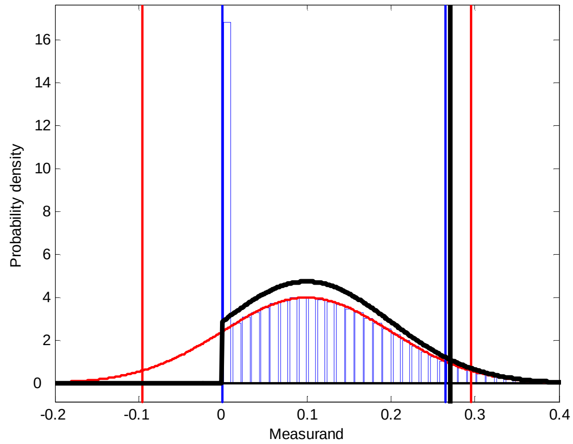

Figure 8 – Comparison of solution PDFs provided by the GUM uncertainty framework (thin continuous curve), Bayesian approach (thick continuous curve) and a Monte Carlo method (scaled frequency distribution), for the case x = 0.1 and $u(x)$ = 0.1. The (shortest) 95 % coverage intervals are also shown as broken and continuous vertical lines, respectively.

The Bayesian approach and the Monte Carlo method treat the physical knowledge about the problem differently. In the Bayesian approach the knowledge is treated *probabilistically*. The prior PDF for real analyte concentration Y encapsulates the knowledge about Y independently of any measurement, and through the likelihood function a negative value of measured concentration may arise for a positive value of real analyte concentration with non-zero probability. In the use of a Monte Carlo method (and generally in the application of the propagation of distributions with the proposed functional model) the knowledge is treated *functionally*: real analyte concentration regarded as a quantity can never be negative, even though measured analyte concentration also regarded as a quantity can be positive or negative.

The modelling part of the solution approach is generally a crucial stage. Although both functional and probabilistic modelling have roles to play in the determination of feasible solutions, the choice of model is paramount, and to this particular class of problems the Bayesian approach, apparently, has the edge. Looking into Figure 7 it is obvious why: the Monte Carlo model aggregates all the possible negative values into zero leading to a shorter distribution curve for the rest of values, whereas the Bayesian approach in its own formulation imposes the range of possible values for the measurand and thus provides a more "believable" distribution associated with it.

Examples from dimensional metrology, e.g., surface texture profiles, could have also been used to demonstrate the relevance of incorporating prior knowledge into the model uncertainty framework.

3.3 Implicit functions

The example used to illustrate this class of problems is the calibration of a pressure balance (deadweight machine). The methodology applied is based on a general formulation [5] and the results are supported on experimental data [6]. The expression for the determination of the required pressure value is:

$$p = \frac{M(1-(\rho_a/\rho_w))g}{A_0(1+\lambda p)(1+\alpha(T-20))} \quad (3)$$

The variable M represents the total mass used and the corresponding ρ_w also applies to the total mass applied. The others variables apply to the effective area of the piston (A_0), the distortion coefficient of the piston-cylinder (λ), the temperature coefficient (α) and mean temperature during the calibration test (T). It is relevant to point out that expression (3) has neglected others influencing terms such as the height difference and the piston angle to the vertical, and is therefore a simplification.

It can be noted that the required pressure value term, p, appears in both sides of the equation, forcing the use of zero finding numerical scheme. Under the GUM uncertainty framework derivatives have to be calculated, for the sensivity coefficients, including the partial derivative of p in the implicit equation. It is advisable to use a compact matrix formulation as suggested [5] to determine the uncertainty associated with the pressure value p,

$$u^2(y)\,c_y^2 = \boldsymbol{c}_x{}^T \boldsymbol{V}_x \boldsymbol{c}_x \quad (3)$$

since it helps in the calculation stages, by allowing the covariances ($\boldsymbol{V_x}$) and sensitivity coefficients ($\boldsymbol{c_x}$ and c_y) to be accounted for in a systematic formulation.

Even with the compact formulation the mathematics involved in the solution of this type of problems can become quite intricate. An obvious advantage of the Monte Carlo method (MCM) is precisely the simplification that its use can provide, and here is a good example. For the input data in Table 1 below,

Table 1. Input data for the pressure balance problem

Symbol	Best Estimate	Variation of limits	Dist	Units
M	32,9000	$\pm$ 3,6×10^{-4}	N	[kg]
A_0	8,0661×10^{-5}	$\pm$ 3,2×10^{-8}	N	[m^2]
λ	3,3×10^{-11}	$\pm$ 7,0×10^{-13}	N	[Pa^{-1}]
g	9,8007	$\pm$ 1,0×10^{-5}	N	[m.s^{-2}]
α	4,6×10^{-5}	$\pm$ 4,2×10^{-6}	T	[ºC^{-1}]
ρ_a	1,2	$\pm$ 2,0×10^{-2}	T	[kg/m^3]
ρ_w	7850,0	$\pm$ 5,0×10^{-1}	T	[kg/m^3]
T	30,0	$\pm$ 1,0	R	[ºC]

where the letters in the 4th column refers to distributions assigned to the input quantities, namely normal (N), triangular (T) and rectangular (R), the best estimate for the measurand is about 3,995 MPa. Using the GUM approach, the value for the standard uncertainty is about ± 1,6 kPa whereas the same quantity determined through the Monte Carlo method has a lower value of ± 0,92 kPa. This, however, is not the always the case, being perhaps more common the opposite situation where a GUM application underestimates the uncertainty evaluation. In this particular problem the non linearity and the ratio involved might explain the situation, bearing also in mind that the use of second order derivatives might have diminished the final result of the standard uncertainty. A better understanding of the reasons behind these differences would require further sensivity analysis, which is out of the scope of this paper.

3.4 Iterative processes

Some measurands are related to equilibrium states (e.g. thermodynamic) that are established under specific conditions for a set of measurands. In some cases, the complexity of the phenomena under study leads to the use of mathematical models based on iterative processes as the only way to achieve adequate solutions.

A typical example is found in humidity related measurements, where the dew point temperature evaluation uses a reference two-pressure and two-temperature generator. In this case, the conditions are generated in a controlled chamber containing an air moisture sampling, being measured the pressure and the temperature of the chamber and of a saturator. The mathematical model that relates the measured quantities with the dew point temperature, T_d, is given by:

$$f_\mathrm{ws}(p_\mathrm{c},T_\mathrm{d})\,p_\mathrm{ws}(T_\mathrm{d}) = f_\mathrm{ws}(p_\mathrm{s},T_\mathrm{s})\,p_\mathrm{ws}(T_\mathrm{s})\cdot\frac{p_\mathrm{c}}{p_\mathrm{s}}\cdot\eta \quad (4)$$

where,

p_s , T_s - saturator pressure and temperature;

p_c - Chamber pressure;

$f_\mathrm{ws}(p,T)$ - Enhancement factor [7];

$p_\mathrm{ws}(T)$ - Saturation vapor pressure [7];

η – saturator eficciency.

In this model, two of the input variables, $f_\mathrm{ws}(p,T)$ and $p_\mathrm{ws}(T)$, are obtained using relations with a number of coefficients, given in [7], imposing a two-stage approach to the evaluation of uncertainties problem.

The iterative formula that gives the dew point is readily obtained from (4)

$$g(T_{\mathrm{d}_\mathrm{n}}) = \frac{f_\mathrm{ws}(p_\mathrm{s},T_\mathrm{s})\,p_\mathrm{ws}(T_\mathrm{s})\,p_\mathrm{c}}{f_\mathrm{ws}(p_\mathrm{c},T_{\mathrm{d}_\mathrm{n}})\,p_\mathrm{ws}(T_{\mathrm{d}_\mathrm{n}})\,p_\mathrm{s}}\eta - 1 \quad (5)$$

and a solution for the dew point temperature can then be found applying a Newton-Raphson method (6), provided a seed for the iterative process, T_{d_0}, a maximum number of iterations, n, and a numerical criterion

$$T_{\mathrm{d}_{\mathrm{n+1}}} = T_{\mathrm{d}_\mathrm{n}} - \frac{g(T_{\mathrm{d}_\mathrm{n}})}{g'(T_{\mathrm{d}_\mathrm{n}})} \quad (6)$$

Two approaches can be recommended for this type of metrological problem: MCM or Bayesian Inference.

In the first case, it is required to know the parameters of the input quantities PDFs (pressures and temperatures of the chamber and of the saturator), the numerical simulation is developed under an iterative process where the solution is tested every time a cycle of calculations ends, being stopped successfully if the numerical criterion is reached, or not successfully if the maximum number of iterations is performed without reaching the numerical criterion. A study using this approach is found in [8].

Regarding the use of Bayesian inference, the general approach can be seen at [9], describing the main procedure of an inverse uncertainty evaluation approach, the so called inverse Monte Carlo (IMC) method. This technique is being applied in a study [10] concerned with the evaluation of measurement uncertainties in a two-pressure humidity generator. The preliminary results show that IMC provides a similar outcome in comparison with MCM, being the selection of the tolerance parameter a crucial factor in IMC.

3.5 Final remarks

Summing up the studied cases, and taking into account more general classes of problem well known and widely published, permits the construction of the following Table,

Table 2. Summary of model choice cases.

Model	GUM	MCM	Bayesian
Linear	☑	☑	☑
Non linear	☐	☑	☑
Ratio	☐	☑	–
Physical limit	☒	☐	☑
Implicit	☐	☑	–
Iterative	☒	☑	☑

where the inserted symbols refer to situations of adequate (☑), conditional (☐) and inadequate (☒) use. In the present work, two of the cases where not studied using a Bayesian approach.

4. CONCLUSIONS

The solution approaches considered are capable of treating functional or probabilistic models to the degree of approximation typically required in practice. However, the modelling itself constitutes a critical stage. The choice of model dictates the solution.

Examples have shown that, depending on the mathematical model, the value and assigned distributions for each input variable will influence the validity of the approach taken to the evaluation of the measurement uncertainty.

This paper was set out to establish simple, general rules, to decide upon the correct choice of method to perform uncertainty calculations.

REFERENCES

[1] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML *Guide to the Expression of Uncertainty in Measurement* 2nd edn (Geneva: International Organization for Standardization), 1995.

[2] JCGM *Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method*, JCGM 101 :2008.

[3] M. G. Cox and P. M. Harris, "Measurement uncertainty and traceability", *Meas. Sci. Technol.* **17** 533-40, 2006.

[4] Maurice Cox, Alistair Forbes, Peter Harris and João Sousa, "Accounting for physical knowledge in obtaining measurement results and associated uncertainties", *XVIII IMEKO World Congress*, 17-22 September, Rio de Janeiro, Brazil, 2006.

[5] M. G. Cox and P. M. Harris, "Uncertainty Evaluation", Software Support for Metrology Best Practice Guide No. 6, NPL Report, DEM-ES-011, September, 2006.

[6] Álvaro Silva Ribeiro, *Evaluation of measurement uncertainties in complex linear and nonlinear systems*, PhD thesis (in Portuguese), Edition LNEC, ISBN 978-972-49-2116-7, 2007.

[7] Hardy, B., "ITS-90 formulations for vapor pressure, frost point temperature, dew point temperature and enhancement factors in the range -100 ºC to +100 ºC", *in* Proceedings of the Third International Symposium on Humidity & Moisture, Teddington (UK), 1998.

[8] Ribeiro, A. e Martins, L., "Mathematical Models Comparative Analysis Applied to the Dew Point Temperature Measurement in Reference Conditions" (in Portuguese), 2nd SPMet National Metrology Conference, Funchal, Portugal, 2007.

[9] Forbes, A. and Sousa, J. "The GUM, Bayesian Inference and Forward and Inverse Uncertainty Evaluation", 2nd SPMet National Metrology Conference, Funchal, Portugal, 2007.

[10] L. Lages Martins, A. Silva Ribeiro and J. Alves e Sousa "Dew point temperature measurement uncertainties in a two-pressure humidity generator", to be presented at the TEMPMEKO & ISHM Conference to be held in Slovenia, May 31 – June 4, 2010.