Method selection to evaluate measurement uncertainty in microflow applications

J A Sousa¹, E Batista¹, O Pellegrino¹, A S Ribeiro², L L Martins²

¹IPQ – Portuguese Institute for Quality, Caparica, Portugal
²LNEC – National Laboratory for Civil Engineering, Lisbon, Portugal

E-mail: jasousa@ipq.pt

Abstract. Microflow applications are gaining increasing importance in medicine, and are indeed crucial in areas such as neonatology and oncology. However, in terms of uncertainty evaluation the very small levels of flow entail problems characterized by a “close to the physical limit” situation. The same happens in nanovolume applications. In these cases the Law of Propagation of Uncertainty framework (GUM) may not to be adequate, since the interval of coverage probability will, in some circumstances, contain negative values, which is physically difficult to explain. Monte Carlo methods may be a good alternative when there are not many negative readings in the application (limit-of-detection problem), but may also prove inadequate otherwise, since the required functional relationship of the model will produce a peak of values centred at the zero value, which is also not very credible to describe satisfactorily the real situation. In the latter, the application of a Bayesian method will be the better approach to deal with the evaluation of measurement uncertainties. The whole process of the revision of the international metrological standardization lies in this principle, of selecting the best method to tackle the particular problem in hand, and not to apply blind recipes to the evaluation of uncertainty in measurement, and there is a new EMPIR project focused on providing examples to BIPM documents to strengthen this concept.

1. Introduction

The overall message conveyed by the new EMPIR (European Metrology Program for Innovation and Research) project 17NMR05 EMUE (Examples of measurement uncertainty evaluation) is that the choice of method to the evaluation of measurement uncertainty must be based on the particular features of the problem in hand, and should not be based on blind recipes [1]. It will contribute to the set of Joint Committee for Guides in Metrology (JCGM) documents related to measurement uncertainties (e.g., JCGM 100, JCGM 101, JCGM 102, JCGM 103, JCGM 106), particularly to JCGM 110 which will collect a number of examples to support this view and help users to select the best method for their particular problem. To provide a flavour of that approach, the present study will interrelate with another EMPIR project, MeDD II (Metrology for drug delivery) [2]. It is aimed at studying the development of new calibration methods requiring traceable measurements of volume, flow and pressure of existing drug delivery devices and inline sensors operating at very low flow rates, which increasingly take place in hospitals and health centres, expanding the existing metrological infrastructure, which has been a matter of applied research funded by other EMPIR projects [3].

The most commonly used form of therapy in health care is infusion therapy [4], which implies that drug delivery is an important topic in this sector. Due to the widespread applications in critical health care, infusion errors are often made, with reported dramatic effects in different applications in the health
sector, especially in neonatology. There are various examples where adverse incidents, morbidity and mortality, can be traced back to poor inaccurate dosing [5, 6].

For instance, the accuracy of the flow rate set point adjustments based on the patient vital signs is insufficient to ensure safe delivery of drugs. Also the increasing implementations of novel microfluidic solutions in medical applications, such as organ-on-a-chip technology, urge the development of metrological infrastructure for validating quality and reproducibility. And a vital part of this infrastructure lies in the ability of producing quality measured values, which can only be achieved if adequate measurement uncertainty values can be obtained.

2. Objectives
This MeDD II health project has the overall goal of enabling traceable measurements of volume, flow and pressure of existing drug delivery devices and inline sensors that work at a flow rate lower than 100 nL/min, in order to minimize inaccurate measurement results, thus improving dosing accuracy in each infusion line.

To that purpose, one specific objective of MeDD II is to develop and validate novel calibration procedures for existing drug delivery devices (e.g. infusion pumps, pain controllers and infusion pump analyzers) with measurement traceability to a primary standard and with a target uncertainty value of 2 % in a 5 nL/min up to 600 mL/min range and in addition to develop a proof-of-concept of an on-chip microfluidic pump used as transfer standard in drug discovery and organ-on-a-chip applications in 5 nL/min to 100 nL/min flows. Another specific objective is to develop new nanoflow metrology techniques based on optical technology to measure flow below 100 nL/min for both steady and fast changing flow. Measurements will be performed for flow sensors and flow generators for Newtonian liquids. The target uncertainty is 1 % (k=2) for steady and 2 % (k=2) for fast changing flow rates.

To exemplify the mathematical treatment of the uncertainty evaluation, experimental values arising from an insulin pump calibration, will be used to prove the clear advantage of using a Bayesian method to quantify measurement uncertainties, when compared to other methods.

In terms of impact, it is known that drug delivery errors account for a significant percentage of medical errors, and depending on the drug type, the patient characteristics and the applied therapy, dosing errors can have severe consequences [7], including a substantial number of fatalities. Important examples can be found in chemotherapy, in oncology, in anaesthesia, in the operating theatre and in nursery wards, especially for the neonates. Therefore, any attempt to prevent adverse events by improving the knowledge on actual doses can already make an enormous difference for the individual patient, especially new born babies, and has a significant impact on the health sector as a whole. And, of course, measurement uncertainty plays a vital role in this equation, since a reliable uncertainty evaluation will lead to an improved confidence in the actual dosing prescribed and provided.

3. Uncertainty budget and possible approaches

3.1. Most readings are positive
A typical uncertainty budget of a flow measurement is illustrated in table 1, resulting from a real experiment with an insulin pump. In this particular test, most of the readings gave positive values as shown in figure 1, where the dispersion of values can be perceived. In this situation, the use of the GUM uncertainty framework (GUF) can first be attempted especially when the repeatability is assumed to have a Gaussian behaviour (which is the case in Table 1). The main uncertainty sources were considered in the mathematical modelling of the flow measurement (Q) expressed by

\[
Q = \frac{1}{T_n-T_1} \left[ \left( M_p - M_t \right) \times (1 - \frac{D_{\text{tube}}}{\rho_{\text{tank}}}^2) \times \frac{1}{\rho_{W}-\rho_A} \times \left( 1 - \frac{\rho_A}{\rho_B} \right) \times [1 - \gamma(t - 20)] \right] + M_{\text{evap}} \tag{1}
\]

with mass measurements (M), densities values of water (\(\rho_W\)), air (\(\rho_A\)) and mass pieces (\(\rho_B\)), evaporation rate (\(M_{\text{evap}}\)), water temperature (\(\iota\)), time(\(T\)), expansion coefficient (\(\gamma\)), diameters (\(D\)), and measurements repeatability.
In this first case study, it is apparent from figure 1 that almost the entire set of readings is non negative, despite having very small magnitude. There are other examples, as will be seen in section 3.2, where this is not the case and the range of values is much larger, including also negative values.

The example of figure 1, though, represents a good response of the insulin pump with a relatively small repeatability, and two possible approaches will be discussed to quantify the corresponding measurement uncertainty, namely the GUF and the Monte Carlo method (MCM).

In the first approach, and for majority of situations, repeatability is treated as having a Gaussian distribution, assuming that data can be well represented by the corresponding mean value and standard deviation [8]. The associated uncertainty is taken as the standard deviation of the mean, which is the standard deviation divided by the square root of the number of readings. The result from this approach can be seen in figure 3, from where it is reasonable to conclude that the GUF approach is acceptable, which was expected since there is only a mild non linearity in the model (1), and there is not a dominant non Gaussian source of uncertainty in table 1. The result of the MCM is displayed in figure 3, with a difference between approaches of about 10 % in terms of the coverage interval for a confidence level of 95 %, thus the GUF approach would be acceptable in most applications.

### Table 1. GUM “uncertainty budget” for the insulin pump.

| Quantity       | PDF  | Best estimate | Standard uncertainty \( u(x) \) \( \sigma_i \) | \( u_i \) |
|----------------|------|---------------|-----------------------------------------------|--------|
| \( T_f \) (s)  | Gaussian | \( 2.65 \times 10^2 \) | \( 7.00 \times 10^{-4} \) | \( 4.62 \times 10^{-11} \) | \( 3.23 \times 10^{-14} \) |
| \( M_f \) (g)  | Combined | \( 2.244425 \) | \( 3.78 \times 10^{-5} \) | \( 3.27 \times 10^{-9} \) | \( 1.23 \times 10^{-13} \) |
| \( \rho_{I} \) (g/ml) | Combined | \( 0.998604 \) | \( 5.57 \times 10^{-4} \) | \( -1.42 \times 10^{-7} \) | \( -7.88 \times 10^{-11} \) |
| \( \rho_{A} \) (g/ml) | Rectangular | \( 0.001208 \) | \( 2.89 \times 10^{-6} \) | \( 1.24 \times 10^{-7} \) | \( 3.58 \times 10^{-13} \) |
| \( \rho_{T} \) (g/ml) | Normal | \( 8.00 \) | \( 2.50 \times 10^{-3} \) | \( 2.66 \times 10^{-12} \) | \( 6.66 \times 10^{-15} \) |
| \( t \) (°C)   | Combined | \( 1.80 \times 10^1 \) | \( 7.22 \times 10^{-1} \) | \( -1.41 \times 10^{-12} \) | \( -1.02 \times 10^{-12} \) |
| \( \gamma \) (°C⁻¹) | Rectangular | \( 1.00 \times 10^{-5} \) | \( 2.89 \times 10^{-7} \) | \( 2.82 \times 10^{-7} \) | \( 8.16 \times 10^{-14} \) |
| \( M_{\text{evap}} \) (ml/s) | Rectangular | \( 1.04 \times 10^{-7} \) | \( 1.47 \times 10^{-8} \) | \( 1.00 \) | \( 1.47 \times 10^{-8} \) |
| \( D_{\text{tube}} \) (cm) | Normal | \( 0.09 \) | \( 0.001 \) | \( -1.38 \times 10^{-8} \) | \( -1.38 \times 10^{-11} \) |
| \( D_{\text{tank}} \) (cm) | Normal | \( 1.36 \) | \( 0.001 \) | \( 9.13 \times 10^{-10} \) | \( 9.13 \times 10^{-13} \) |
| \( \rho_{\text{evap}} \) (ml/s) | Normal | \( 0.00 \) | \( 1.99 \times 10^{-8} \) | \( 1.00 \) | \( 1.99 \times 10^{-8} \) |
| \( M_{F} \) (g) | Combined | \( 2.255857 \) | \( 3.78 \times 10^{-5} \) | \( -3.27 \times 10^{-4} \) | \( -1.23 \times 10^{-8} \) |
| \( T_f \) (s)  | Gaussian | \( 3.32 \times 10^3 \) | \( 7.00 \times 10^{-4} \) | \( -4.62 \times 10^{-11} \) | \( -3.23 \times 10^{-14} \) |

Flow rate (ml/s) \( 2.45 \times 10^{-7} \) \( 3.03 \times 10^{-8} \)

Figure 1. Flowrate readings for the case study 1 when most readings are positive.

However, if it is assumed that data in figure 1 can be represented by a rectangular distribution, based on the fact the scatter of the data only allows using their limits as a valid information [8], then the comparison between the two approaches reveals a striking difference as illustrated in figure 4. The difference between the GUF and MCM is quite apparent in the graph and is accounted as being more...
than 200% in the coverage interval for a level of confidence of 95%. This allows an obvious conclusion that the use of the GUF, under these assumptions, would clearly underestimate the measurement uncertainty associated with the insulin pump.

An interesting feature of the MCM in this situation is the concentration of values at zero shown in figure 4, where a spike near the origin is apparent. In this case study 1 where the concentration of values is not too large, one can reasonably conclude that a greater number of negative values would dramatically increase this effect.

Figure 3. Simulations using the GUM and MCM methods for the case study 1.

3.2. Significant number of negative readings
A different situation arises when one is faced with data as depicted in figure 2. In this case study 2 the number of negative values is significant and the spike of values found in figure 4 would be much enlarged. Therefore neither the GUF nor the MCM method will be adequate to describe this latter situation, as these methods concern the probability density of the measurand estimate. The Bayesian method rather concerns the probability density of the measurand, considering the data produced by the measurement and possible prior information [9]. Indeed, this method should be capable of dealing well with this situation, where the distribution of values is truncated at zero (prior knowledge), and thus redistributes the values over the rest of the integral domain.

The average flow rate $Q$ can be estimated from measurements of mass and time according to a relationship of the form

$$Q = \left( \frac{M_F - M_I}{T_F - T_I} \right) f(a) \quad (2)$$

Where $f(a)$ is a correction factor depending on parameters $b$ of the specific instrument and environmental effects. Given distributions associated with each quantity they can be propagated through to a distribution for $Q$ using a Monte Carlo method, as a measurement equation. An observation equation approach to the analysis of the measurement [10] uses a relationship of the form

$$M_F = (M_I - M_{\text{evap}}) + (T_F - T_I)Qg(a) + \epsilon_F \quad (3)$$

Setting $b = (M_I, T_I, T_F, a^T)^T$ with non-informative priors, we can write (3) as

$$M_F = \phi(Q, b) + \epsilon_F = h(b) + k(b)Q + \epsilon_F \quad (4)$$

The prior for $Q$ should include the constraint that $Q \geq 0$ since there is no physical mechanism for a negative flow. The posterior distribution for the parameters $Q$ and $b$ is such that

$$p(Q, b|M_F) \propto p(M_F|Q, b)p(b)p(Q) \quad (5)$$

where the prior $p(b)$ has been assigned [11]. The form of the observation equation (4) is precisely that of the instrument response model considered in [10] which allows for a particularly simple Markov
chain Monte Carlo (MCMC) algorithm to be implemented in the basis of a GUM supplement [8] sample generated using the measurement model. The posterior distribution \( p(Q | M_F) \) will differ from the distribution derived using the GUMS1 approach since \( \partial \phi / \partial Q = k(b) \) is not constant and the prior distribution \( p(Q) \) must impose the constraint that \( Q \geq 0 \).

The result of the simulation using the MCMC with the constraint indicated above proves the advantage of the Bayesian approach in limit-of-detection problems, as illustrated in figures 5 and 6. In fact, the change of mass is similar to the uncertainty in the mass measurement. The Bayesian posterior is very much like a truncated Gaussian, as was expected. A simple MCM simulation would result in all the negative values allocated at zero creating an unfeasible representation of the flow rate distribution.

**Figure 5.** MCMC simulations truncated at zero compared with MCM simulations without any constraint for the case study 2.

**Figure 6.** MCMC simulations for the case study 2 with constraint \( Q \geq 0 \).

4. Conclusions
The aim of this work is to provide guidance on method selection with respect to the evaluation of measurement uncertainty. Clearly the choice of method depends on the particular problem under evaluation and no blind recipe should be used. Depending on the conditions of the problem, the GUM uncertainty evaluation may be sufficiently adequate, whereas in other circumstances alternative approaches should be applied instead, e.g., the Monte Carlo method or the Bayesian method. What this study indicates is that the MCM approach is a convenient alternative to the GUM uncertainty evaluation approach, with the advantage of having often a simpler application, despite requiring a software implementation. However, this robust alternative method may also prove inadequate in certain circumstances such as in limit-of-detection type of problems. In the latter class of problems, the constraints imposed to the measurand by the use of a prior fits well to the Bayesian approach and is clearly a better method to evaluate measurement uncertainty in a number of problems that have application not only in microflow but also in nanovolume, surface characterization (roughness) or chemical analytical measurements in substances with a high degree of purity.

**Acknowledgments**
Thanks are due to Professor Alistair Forbes of NPL for a thorough revision of the paper and invaluable suggestions for improvement. The EMPIR project “17NRM05-EMUE” is carried out with funding of the European Union under the EMPIR. The EMPIR is jointly funded by the EMPIR participating countries within EURAMET and the European Union.

**References**
[1] EMPIR project 17NRM05 EMUE – Examples of measurement uncertainty evaluation. Program available on www.euramet.org
[2] EMPIR project MeDD II – Metrology for drug delivery. Program available on www.euramet.org
[3] EMPIR project MeDD I – Metrology for drug delivery. Program available on www.euramet.org
[4] Lee P T, Thompson F and Thimbleby H 2012 Analysis of infusion pump error logs and their significance for health care British Journal of Nursing Apr 26-May 9; 21(8) S12, S14, S16-20

[5] Heinemann L, Fleming GA, Petrie JR, Holl RW, Bergenstal RM and Peters AL 2015 Insulin pump risks and benefits: a clinical appraisal of pump safety standards, adverse event reporting, and research needs: a joint statement of the European Association for the Study of Diabetes and the American Diabetes Association Diabetes Technology Working Group Diabetes Care 38(4) 716-22

[6] Ma H, Lovich M A and Peterfreund R A 2011 Quantitative analysis of continuous intravenous infusions in pediatric anesthesia: Safety implications of dead volume, flow rates, and fluid delivery Paediatr. Anaesth. 21 78–86

[7] Husch M, Sullivan C, Rooney D, Barnard C, Fotis M, Clarke J and Noskin G 2005 Insights from the sharp end of intravenous medication errors: implications for infusion pump technology Qual. Saf. Health Care Apr; 14(2) 80-6

[8] JCGM 2008 JCGM 101:2008 Evaluation of Measurement Data – Supplement 1 to the “Guide to the Expression of Uncertainty in Measurement” – Propagation of Distributions Using a Monte Carlo Method (Joint Committee for Guides in Metrology)

[9] Calonico D, Levi F, Lorini L and Mana G 2009 Bayesian inference of a negative quantity from positive measurement results Metrologia 46 267-71

[10] Forbes A B and Sousa J A 2011 The GUM, Bayesian inference and the observation and measurement equations Measurement 44 1422-35

[11] Gelman A, Carlin J B, Stern H S, Dunson D B, Vehari A and Rubin D B 2014 Bayesian Data Analysis third edition (Boca Raton, Florida: Chapman & Hall/CRC)