Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Uncertainty calculation methodologies in microflow measurements: Comparison of GUM, GUM-S1 and Bayesian approach

J.A. Sousa a,*, E. Batista a, S. Demeyer b, N. Fischer b, O. Pellegrino a, A.S. Ribeiro c, L.L. Martins c

a IPQ – Portuguese Institute for Quality, 2829-513 Caparica, Portugal
b LNE – Laboratoire National de Métrologie et d’Essais, 75724 Paris, France
c LNEC – National Laboratory for Civil Engineering, 1700-513 Lisboa, Portugal

ARTICLE INFO

Keywords:
Uncertainty evaluation
Accuracy
Microflow measurement
Syringe pump
Gravimetric method
Bayesian approach

ABSTRACT

The importance of measurement quality cannot be over emphasized in medical applications, as one is dealing with life issues and the wellbeing of society, from oncology to new-borns, and more recently to patients of the COVID-19 pandemic. In all these dire situations, the accuracy of fluid delivered according to a prescribed dose can be critical.

Microflow applications are growing in importance for a wide variety of scientific fields, namely drug development and administration, Organ-on-a-Chip, or bioanalysis, but accurate and reliable measurements are a tough challenge in micro-to-femto flow operating ranges, from $2.78 \times 10^{-4}$ mL/s down to $2.78 \times 10^{-7}$ mL/s (1000 µL/h down to 1 µL/h). Several sources of error have been established such as the mass measurement, the fluid evaporation dependent on the gravimetric methodology implemented, the tube adsorption and the repeatability, believed to be closely related to the operating mode of the stepper motor and drive screw pitch of a syringe pump. In addition, the difficulty in dealing with microflow applications extends to the evaluation of measurement uncertainty which will qualify the quality of measurement. This is due to the conditions entailed when measuring very small values, close to zero, of a quantity such as the flow rate which is inherently positive. Alternative methods able to handle these features were developed and implemented, and their suitability will be discussed.

1. Introduction

The most commonly used form of therapy in health care, at hospitals, is infusion therapy [1], which implies that drug delivery devices are very important instruments 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. There are various examples where adverse incidents, morbidity and mortality, can be traced back to poor inaccurate dosing [23]. Important examples can be found in chemotherapy, in oncology, in anaesthesia, in the operating theatre and in nursery wards, especially for the neonates. This situation is even more critical at very small flow rates, e.g., in neonatology, where any variation to the normal infusion procedure can lead to very large errors. Also, the unprecedented conditions Public Health Institutions experience due to COVID-19 pandemic crisis has forced the hospital administrations to take measures outside the usual work practices in order to manage several challenges:

- Reduction of health care staff exposure to COVID-19
- Preservation of personal protective equipment (PPE)
- Manage shortages of equipment
- Manage of wastage of now-very-scarce critical medications needed for COVID-19 care and other critical drugs or substances
- The measures taken by some hospitals outside the usual work practices include:
  - To use the drug delivery devices outside the patient room
  - To use the drug delivery devices with associated clinical risks and lower accuracy
  - To postpone maintenance and calibration of equipment deadlines in instruments with critical use.

These practices, if not performed under very controlled conditions, will lead to large dosing errors [4]. 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-
aims at developing and improving the required metrology tools to assure the traceability of clinical data, allowing the metrological comparability involving NMIs/DIs and the health industry, was approved. This JRP, coordinated by the Portuguese Institute for Quality (IPQ), with 15 partners involved in primary standards for liquid flow rates from at least 1 mL/min down to 100 nL/min (17 μL/s down to 1.7 nL/s) within the framework of the EMRP project “HLT07 Metrology for Drug Delivery – MeDD” [11] which ended in June 2015. However, as pointed out above, the requirement now is more stringent and new advances [5] point to flowrates below 100 nL/min (1.7 nL/s).

Several National Metrology Institutes have developed primary standards for liquid flow rates from at least 1 mL/min down to 100 nL/min (17 μL/s down to 1.7 nL/s) within the framework of the EMRP project “HLT07 Metrology for Drug Delivery – MeDD” [11] which ended in June 2015. However, as pointed out above, the requirement now is more stringent and new advances [5] point to flowrates below 100 nL/min (1.7 nL/s).

Primary standards are based on the gravimetric method that measures the mass of a liquid as a function of time on an analytical balance. The flow rate is in general determined by the quotient of the mass difference, \( \Delta m \), and time difference including appropriate corrections, \( \Delta t \). This method, however, has obvious limitations such as evaporation and other related to the existing link with a weighing instrument required to measure the mass difference. If one thinks on the limits imposed by the number of digits of such measuring instrument (MI) and the lower limits (weight) of the artefacts used to calibrate those MI, it can readily be concluded that the proportion of uncertainty associated with the measuring standard will increase as the flowrate diminishes. The general working principle is the same for all gravimetric setups and consists of a flow generation, connection to a device under test and connections to a water collection on top of a precision balance or an analytical balance as shown in Fig. 1 and Fig. 2. This balance reads the mass, which can be combined with timing equipment to calculate the \( \Delta m/\Delta t \) quotient. In this example, the mass of ultra-pure and degassed water delivered by the instrument under test is weighed at a set time and converted to volume at a reference temperature [9]. The volumetric flow rate is determined

- Fig. 1. Simplified Flow circuit. The fluid goes from a glass syringe (B) of the Nexus 3000 pump (A), to the balance (E) through tubing (C) that is immersed in the weighing vessel (F) that is inside an evaporation trap (D).

- Fig. 2. Nexus 3000 syringe complete calibration setup for the gravimetric setup.
The volumetric flow rate \( Q \) is then converted from mass to volume at a reference temperature [9]. The gravimetric method, that become apparent at very low flow rates (due to very low mass differences), other possibilities have been explored to enable the measurement down to 1.6 nL/min (0.27 pL/s) with relative uncertainties close to 2%, e.g., in a recent work [17] an interferometer was used to monitor the distance travelled by a pusher block of a flow generator connected to a glass syringe in order to determine the flow rate (Fig. 3).

The use of interferometry for flow measurement involved the use of the following components: a laser unit with a detector incorporated, an optical arrangement composed by two retroreflector cubes (one of which with a beam splitter attached), a Control Unit, a pusher block, a Nexus 3000 pump (flow generator) and a glass syringe. In practice, the generation of flow was accomplished by a stepper motor which drove a screw connected to a pusher block that itself pushed the syringe piston. What is important to emphasize is that along with the specific problems facing microflow measurement, which have led to the development of alternative methods to overcome existing limitations, as detailed above, there are also problems facing the evaluation of measurement uncertainty associated with very small flow rates (close to zero) as will be addressed in the next chapters.

3. Measurement model and uncertainty budget

A typical uncertainty budget of a flow measurement is illustrated in Table 1, resulting from a real experiment with a syringe pump calibrated by the gravimetric method (Figs. 1 and 2) at a flow rate of 2.70 × 10^{-7} mL/s. This measuring principle is based on weighing the mass of the

\[
Q = \frac{1}{T_f - T_i} \left[ \left( M_f - M_i \right) \times \left( 1 - \frac{D_{obs}}{D_{sk}} \right) \right] \times \frac{1}{\rho_w - \rho_A} \times \left( 1 - \frac{\rho_A}{\rho_h} \right) \times [1 - \gamma(t - 20)] + Q_{\text{evap}} \quad (1)
\]

working fluid delivered by the instrument under test at a set time, which is then converted from mass to volume at a reference temperature [9]. The volumetric flow rate \( Q \) is determined by the quotient of the volume of the reference liquid and the time interval, with the corrections included in Eq. (1) [10].

A LabVIEW application is used for data acquisition, validation, online visualization of measured data and flow rate calculation [11]. Data acquisition starts approximately 10 min after steady flow is reached. Two tests were carried out, with 35 readings (case study 1) and 76 readings (case study 2), for different sets of conditions, mainly the used tube is from polypropylene in case 2. The test duration and number of acquisition points were increased in order to improve the knowledge of the experience.

The measurement model associated with the syringe pump experiment can be described by Eq. (1) below [10]

### Table 1

GUM “uncertainty budget” for the syringe pump (case study 1).

| Quantity/Unit | PDF     | Best estimate (\( u_i \)) | Standard uncertainty \( u_i(\%) \) | Sensitivity coefficients (\( c_i \)) | \( u(Q) \) |
|---------------|---------|--------------------------|----------------------------------|-----------------------------------|----------|
| \( T_f/s \)   | Gaussian| 2.65 × 10^{-7} \( 7.00 \times 10^{-4} \) | 1.15 × 10^{-10} \( 1.47 \times 10^{-8} \) | 8.03 \( 1.07 \) | \( u(Q) \) |
| \( M_f/g \)   | Combined| 3.799544 \( 2.86 \times 10^{-5} \) | 6.90 × 10^{-8} \( 1.66 \times 10^{-6} \) | 1.97 \( 8.03 \) | \( u(Q) \) |
| \( \rho_w/(g/mL) \) | Combined| 0.997615 \( 6.26 \times 10^{-4} \) | \(-1.67 \times 10^{-7} \) | 1.04 \( 6.90 \) | \( u(Q) \) |
| \( \rho_A/(g/mL) \) | Rectangular| 0.001181 \( 2.89 \times 10^{-6} \) | 1.46 × 10^{-7} \( 1.66 \times 10^{-8} \) | 4.21 \( 8.03 \) | \( u(Q) \) |
| \( \rho_h/(g/mL) \) | Normal| 8.00 \( 2.50 \times 10^{-3} \) | 3.07 × 10^{-12} \( 1.07 \times 10^{-11} \) | 7.67 \( 4.00 \) | \( u(Q) \) |
| \( \gamma/\degree C \) | Combined| 2.27 × 10^{-3} \( 5.17 \times 10^{-1} \) | \(-1.66 \times 10^{-12} \) | 8.60 \( 8.03 \) | \( u(Q) \) |
| \( \rho_{air}/(g/cm^3) \) | Rectangular| 1.00 × 10^{-7} \( 2.89 \times 10^{-5} \) | \(-4.42 \times 10^{-7} \) | 1.28 \( 1.07 \) | \( u(Q) \) |
| \( Q_{\text{evap}}/(mL/s) \) | Rectangular| 1.04 × 10^{-7} \( 1.47 \times 10^{-8} \) | 1.00 \( 1.07 \times 10^{-9} \) | 1.47 \( 4.00 \) | \( u(Q) \) |
| \( D_{\text{tank/cm}} \) | Normal| 0.09 \( 0.001 \) | \(-1.62 \times 10^{-8} \) | 1.62 \( 8.03 \) | \( u(Q) \) |
| \( D_{\text{tank/cm}} \) | Normal| 1.36 \( 0.001 \) | 1.07 \( 1.07 \times 10^{-9} \) | 1.07 \( 8.03 \) | \( u(Q) \) |
| \( \delta Q_{\text{evap}}/(mL/s) \) | Normal| 0.00 \( 4.00 \times 10^{-8} \) | 1.00 \( 1.07 \times 10^{-9} \) | 4.00 \( 4.00 \) | \( u(Q) \) |
| \( M_f/g \) | Combined| 3.799785 \( 2.86 \times 10^{-5} \) | \(-6.90 \times 10^{-8} \) | 1.97 \( 8.03 \) | \( u(Q) \) |
| \( T_f/s \) | Gaussian| 1.71 × 10^{7} \( 7.00 \times 10^{-4} \) | \(-1.15 \times 10^{-10} \) | 8.03 \( 8.03 \) | \( u(Q) \) |

Flow rate, \( Q/(mL/s) \): 2.70 \( 1.07 \times 10^{-1} \) Combined standard uncertainty \( u(y) \): 5.09 \( 1.07 \times 10^{-1} \)
expansion coefficient ($\gamma$), diameters ($D$). A typical set of uncertainty contributions, used for case study 1, is illustrated in Table 1 from which the best estimate and standard uncertainties can be used to evaluate the measurement uncertainty. In here, as for most situations, repeatability ($\delta Q_{\text{rep}}$ in Table 1) is treated as having a centred Gaussian distribution (with mean 0), 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. The model is only mildly nonlinear and there is not any non-Gaussian dominant source of uncertainty.

4. Experimental data from two case studies

The experiments with the syringe pumps have highlighted two different patterns that may occur when dealing with measurements close to the physical limit of a quantity. Depending on the measuring instrument (e.g., precision, accuracy), the conditions of the experiment (e.g., temperature, operator, fluid), and the nominal flow rate targeted, the readings may show a proportion of negative values since in microflow experiments the range of values is very close to zero. Bearing in mind that negative flow rates do not have a physical interpretation but reflect the actual behavior of the equipment and setup, a dilemma exists on how to handle this phenomenon in uncertainty evaluation.

Data from the first case study are displayed in Fig. 4, where most readings are of very small flow rates, as expected, but only a very small proportion of them are negative. It is worth noting that the first 7 readings were ignored, since the stabilization time mentioned above was only partially observed.

The example of Fig. 4, though, represents a good response of the syringe pump with a relatively small repeatability.
However, in the second case study illustrated in Fig. 5, there is a clear increase in the proportion of negative values, and a higher dispersion of values. This situation may arise from different factors such as evaporation, adsorption in the polypropylene tube or air bubbles, thus should entail a different type of problem in the evaluation of measurement uncertainty, as it will be seen in the discussion of results section. It is important to point out that the negative values of flow rate are the result of working very close to the physical limit of a system and as a consequence the intrinsic noise and lack of precision will, at some instances, provide negative outputs, which have no physical interpretation but exist and reflect the actual behaviour of the equipment and the setup and it is not a systematic effect.

For case study 1, the GUM is expected to perform well, whereas for case study 2, the conditions for the applicability of the GUM uncertainty framework are not met since part of the data is not meaningful, and the expanded uncertainty interval is expected to include some of those cases. For each case study, the GUM, GUM-S1 (Monte Carlo) and Bayesian approach will be compared to ascertain which of them can better handle the presence of a significant number of negative values in the readings and, as a consequence provide a more reliable estimate of the flow rate and associated standard uncertainty.

5. The GUM uncertainty framework and the GUM-S1 (Monte Carlo Method)

5.1. The GUM uncertainty framework

In the more common explicit univariate measurement model, a single real output quantity \( Y \) is related to a number of input quantities \( X = (X_1, ..., X_N) \) by a functional relationship \( f \) in the form \( Y = f(X) \), as stated in the GUM [7]. The estimate of the output quantity is taken as \( y = f(x) \).

The standard uncertainty \( u(y) \) associated with \( y \) is evaluated from

\[
u^2(y) = \sum_{i=1}^{N} \sum_{j=1}^{N} c_i u(x_i) c_j
\]

where \( c_i \) is the partial derivative \( \partial f / \partial X_i \) evaluated at \( X = x \) and is known as the \( i \)th sensitivity coefficient, \( u(x) \) is the standard uncertainty associated with \( X_i \), i.e. \( u(x_i) = \sqrt{u^2(x_i, x)} \), and \( u(x_i, x_j) \) the covariance associated with \( x_i \) and \( x_j \).

A compact way of writing the sum in expression above, often used for software programming sustaining matrix formulation, e.g., MATLAB, is

\[
u^2(y) = \lambda^T \lambda
\]

where \( \lambda \) is the covariance matrix of dimension \( N \times N \) containing the covariances \( u(x_i, x_j) \),

\[
\lambda = \begin{bmatrix} u(x_1, x_1) & \cdots & u(x_1, x_N) \\ \vdots & \ddots & \vdots \\ u(x_N, x_1) & \cdots & u(x_N, x_N) \end{bmatrix}
\]

and the (row) vector \( \lambda^T = [c_1, ..., c_N] \) of dimension \( 1 \times N \) contains the sensitivity coefficients. Both expressions on \( u^2(y) \) above are equivalent representations of LPU (Law of Propagation of Uncertainties).

For independent input quantities, we would obtain the better-known simplified expression (equivalent to using \( \lambda \)) with its off-diagonal elements replaced by zeros

\[
u^2(y) = \sum_{i=1}^{N} c_i u(x_i) = \sum_{i=1}^{N} u_i^2(y) \equiv |c_i| u(x_i)
\]

The \( u_i(y) \) are often used in uncertainty budgets to identify which input quantities, with respect to their corresponding standard uncertainties, have significant influence on the standard uncertainty \( u(y) \) associated with the estimate \( y \) of the output quantity (e.g., see Table 1).

The application of the GUM assumes several prerequisites that must be taken into consideration. They include, but are not limited to, the number of input quantities, the PDFs of the input quantities, the linearity of the model and the central limit theorem. When the conditions of linearity or normality are not met, it is usually recommended to turn to the Monte Carlo method as implemented in the GUM-S1 [8].

5.2. GUM-S1 (Monte Carlo Method)

The use of Monte Carlo method in metrology is widespread and formalized in a supplement to the GUM [8]. Monte Carlo method is usually performed after a traditional GUM analysis to confirm the analysis, or to improve the reliability of the result when the conditions of applicability of the GUM are not met (non-linearity of the measurement model, non-normality of the measurement result, non-applicability of the central limit theorem, etc.). Monte Carlo can also be performed alone to avoid tedious computation of partial derivatives (at the first and/or second order) in complex models (in which case results can be compared with a GUM approach of a simplified model).

The heart of the Monte Carlo method of GUM-S1 can be summarized as follows [8, clause 7]. Given a measurement model of the form \( Y = f(X) \), as above, and probability density functions assigned to each of the input quantities \( X = (X_1, ..., X_N) \), it is required to generate \( M \) sets of input quantities \( X_1, ..., X_N \) (\( r = 1, ..., M \)) and use the measurement model to compute the corresponding value for \( Y_r \). The number of sets of input quantities \( M \) should be chosen to be sufficiently large so that a representative sample of the probability density function of the output quantity \( Y \) is obtained. The approach here applies to independent input quantities and a scalar output quantity \( Y \), although the extension to dependent input variables is also covered in [8].

In this method (also called “propagation of distributions”), the PDFs for the input quantities are propagated through the measurement model to provide the PDF for the output quantity. The expectation of this PDF is then used as the estimate of the measurand and the standard deviation of the PDF is used as the standard uncertainty associated with that estimate. The calculation of coverage intervals is based on the actual PDF and not on an assumption of normality as in the GUM.

Monte Carlo method should provide valid results, provided an adequate number of samples is drawn, usually \( 10^8 \) samples are required.

For most applications, GUM and GUM-S1 uncertainty evaluations would prove sufficient. However, limitations to the Monte Carlo method may appear when constraints on the measurand need to be introduced to guarantee the physical interpretation of the result, as in this paper. In such situations, a Bayesian approach is usually recommended as an alternative.

6. The Bayesian method

The Bayesian method [12] determines a posterior distribution for the measurand (and for additional parameters, if any) given observations and a prior information on the measurand (and the additional parameters, if any). For linear measurement models and so-called non informative prior distributions, the posterior distribution coincides with the GUM-S1 PDF. Here, the Bayesian approach is used to impose a positivity constraint on the measurand in the presence of a proportion of negative flow rates readings obtained during calibration (for reasons mentioned earlier).

According to Table 1, we use the following error-in-variables formulation to represent the random variables \( M_p = m_p + \xi_m, M_t = m_t + \xi_m, T_p = t_p, T_t = t_t + \xi_t \), where \( (m_p, m_t, t_p, t_t) \) are the measurements of the final mass, the initial mass, the final time and the initial time respectively (given in the column “Best estimate” of Table 1) and \( (\xi_{M_p}, \xi_{M_t}, \xi_{T_p}, \xi_{T_t}) \) are centered, normally distributed error terms.

Let \( q_{obs} \) denote the observed flow rate computed from measurements
\[ q_{\text{obs}} = \frac{m_T - m_l}{T_T - T_l} \]

and \( Q_{\text{obs}} \) the associated random variable defined by

\[ Q_{\text{obs}} = \frac{M_T - M_l}{T_T - T_T} \sim N(q_{\text{obs}}, \sigma_{q_{\text{obs}}}) \]

We show with Monte Carlo sampling in the distributions of \( (M_T, M_l, T_T, T_l) \) that \( Q_{\text{obs}} \) is normally distributed, with parameters values displayed in the Appendix A for case study 1 and case study 2.

The corresponding error-in-variables formulation writes

\[ Q_{\text{obs}} = q_{\text{obs}} + \eta \sim N(0, \sigma_{q_{\text{obs}}}) \]

Replacing \( Q_{\text{obs}} \) in the measurement model gives

\[ Q = Q_{\text{obs}} \left[ \left( 1 - \frac{D_{\text{evap}}}{D_{\text{tank}}} \right)^2 \frac{1}{\rho_W - \rho_A} \left( \frac{1 - \gamma}{\gamma} \right) \right] + Q_{\text{rep}} \]  

(6)

The statistical model explaining the measurement \( Q_{\text{obs}} \) is obtained by introducing a repeatability term \( \delta Q_{\text{rep}} \sim N(0, \sigma_{Q_{\text{rep}}}) \) defined in Section 3

\[ Q_{\text{obs}} = \frac{Q - \sigma_{Q_{\text{rep}}}}{\sigma_Q} + \delta Q_{\text{rep}} \]

where

\[ C(\theta) = \left[ \left( 1 - \frac{D_{\text{evap}}}{D_{\text{tank}}} \right)^2 \frac{1}{\rho_W - \rho_A} \left( \frac{1 - \gamma}{\gamma} \right) \right] \]

\[ \theta = (D_{\text{evap}}, D_{\text{tank}}, \rho_W, \rho_A, \gamma, t) \]

By replacing the left-hand term, we obtain

\[ q_{\text{obs}} + \eta = \frac{Q - \sigma_{Q_{\text{rep}}}}{\sigma_Q} + \delta Q_{\text{rep}} \]

The statistical model associated with the measurement model writes

\[ q_{\text{obs}} = \frac{Q - \sigma_{Q_{\text{rep}}}}{\sigma_Q} + \zeta + \delta Q_{\text{rep}} \]  

(with \( \zeta = -\eta \))

\[ \zeta \sim N(0, \sigma_{q_{\text{obs}}}) \]

\[ \delta Q_{\text{rep}} \sim N(0, \frac{s}{\sqrt{n}}) \]

The quantity of interest is \( \pi(Q|q_{\text{obs}}) \) obtained from the joint posterior distribution

\[ \pi(Q, \theta|q_{\text{obs}}) \propto \pi(q_{\text{obs}}|Q, \theta) \pi(Q) \pi(\theta) \]

Fig. 6. GUM (line) and Monte Carlo (bars) for case study 1, with vertical lines representing the limits for the 95% confidence interval of both approaches.

| Approach | Best Estimate \( Q/(\text{mL/s}) \) | Standard Uncertainty \( u(Q)/(\text{mL/s}) \) | Coverage Interval \( (95\%), l_q/(\text{mL/s}) \) |
|----------|-------------------------------|----------------|---------------------|
| GUM      | 2.6975 \times 10^{-7}         | 5.0924 \times 10^{-8} | [1.6791 \times 10^{-7}, 3.7160 \times 10^{-7}] |
| GUM-S1   | 2.6972 \times 10^{-7}         | 5.0906 \times 10^{-8} | [1.6898 \times 10^{-7}, 3.6815 \times 10^{-7}] |
| Bayes    | 2.6996 \times 10^{-7}         | 5.1009 \times 10^{-8} | [1.7180 \times 10^{-7}, 3.6970 \times 10^{-7}] |

where \( l(q_{\text{obs}}|Q, \theta) \) is the likelihood of the data, \( \pi(Q) \) is the prior distribution of the measurand and \( \pi(\theta) \) is the joint distribution of the vector \( \theta \) given by the type B distributions associated with the uncertainty sources in Table 1.

In order to impose the positivity of the measurand, the prior distribution should have a support excluding negative values such as \( \pi(Q) \sim \text{Unif}(0, +\infty) \).

Since the posterior distribution has no closed form, Markov Chain Monte Carlo methods are usually employed to sample from the posterior distribution [13-15,18,19,21]. These methods construct a sequence of dependent values which form a Markov chain with stationary distribution equal to the sought distribution. The Metropolis-Hastings algorithm constitutes a popular class of MCMC methods as it only requires to know the right hand part of (3) to sample from the posterior distribution. The sequence of values is usually considered only after a first period of burn-in”, and often the chains are thinned (i.e. only each 10th value is used, say) in order to reduce the correlation between successive values. Different criteria have been suggested for assessing the convergence of MCMC methods.

We refer to [14,20] for a general introduction to these methods and to [15] for an introductory example in metrology. In this paper, results are obtained with Python 3 using probabilistic programming with PyMC3 [16].

7. Discussion of results

7.1. Case study 1

A comparison between the GUM and the Monte Carlo method from GUM-S1 is illustrated in Fig. 6 and estimates are provided in Table 2. For most purposes, it would be reasonable to conclude that the GUM
conventional approach is acceptable, which was expected since there is only a mild nonlinearity in the model (1), and there is not a dominant non-Gaussian source of uncertainty in Table 1 (in conformity assessment, these differences should be taken into account). The coverage interval for GUM was taken using a coverage factor of \( k = 2 \) whereas for the GUM-S1 the percentiles of 2.5% and 97.5%, applied to the sorted sample of output values, were used to obtain the same information.

It is important to underline that the conclusions drawn are only valid for the specific set of conditions defined in Table 1. Just for exemplifying purposes, in the case study 1, increasing by a factor of 5 the uncertainty associated with the evaporation quantity will change the output considerably and the suitability of the GUM may now be questionable, as illustrated in Fig. 7, showing a noticeable difference between coverage intervals. It is always important to validate the suitability of the GUM uncertainty framework whenever the measurement model departs from linearity, or the non-Gaussian input quantities may have a significant impact and interfere with the conditions of the central limit theorem.

Results obtained with the Bayesian method under the hypothesis that \( Q > 0 \) are displayed in Table 3. Bayesian approach provides results similar to those obtained by the other two methods.

Fig. 7. GUM (line) and Monte Carlo (bars) comparison for an increased percentage of the rectangular distribution associated with \( Q_{evap} \).

Fig. 8. Measurand PDF for all approaches applied to case study 1, with bars for GUM-S1 and Bayes \( Q > 0 \) and the line curve for GUM.

| Approach       | Best Estimate, \( Q/(\text{mL/s}) \) | Standard uncertainty \( u(Q)/(\text{mL/s}) \) | Coverage Interval (95%), \( I_Q/(\text{mL/s}) \) |
|----------------|---------------------------------|---------------------------------|---------------------------------|
| GUM            | \( 2.5325 \times 10^{-8} \)     | \( 3.6951 \times 10^{-8} \)     | \([-4.8577 \times 10^{-8}, 9.2227 \times 10^{-8}]\) |
| GUM-S1 (Monte Carlo) | \( 2.5392 \times 10^{-8} \)     | \( 3.6937 \times 10^{-8} \)     | \([-4.6647 \times 10^{-8}, 9.7517 \times 10^{-8}]\) |
| Bayes (Q > 0)  | \( 4.1893 \times 10^{-8} \)     | \( 2.7728 \times 10^{-8} \)     | \([1.1994 \times 10^{-12}, 9.3596 \times 10^{-8}]\) |

Table 3
GUM, Monte Carlo and Bayesian results (case study 2).
impose a positive output but without noticeable influence in this case study since most flowrate values were positive.

7.2. Case study 2

A different situation arises when one is faced with data as depicted in Fig. 5. In this second case study, the number of negative values is significant, which means that negative values are expected to be found in the 95% coverage interval, which is a problem. The Monte Carlo method could only deal with this situation by concentrating all negative values at zero, but this would distort the PDF of the measurand. Therefore, neither the GUM nor the MCM method will be adequate to describe the latter situation, in terms of naturally representing the probability density function for the measurand. The Bayesian method, however, rather concerns the probability density of the measurand, considering the data produced by the measurement and possible prior information. Indeed, this method should handle 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.

A comparison between the GUM, the Monte Carlo method from GUM-S1 and the Bayesian approach under the hypothesis that \( Q \) is positive is given in Table 3 and illustrated in Fig. 9. The advantage and suitability of the Bayesian approach in this case is evident and Fig. 9 illustrates this situation quite clearly by including only positive flowrate values and redistributing them in the positive part of the graph, so that the output PDF is smoothly adapted to accommodate the case study constraints.

The result of the simulation using the Bayesian method with the positivity constraint indicated above proves the advantage of the Bayesian approach in limit-of-detection problems. The Bayesian posterior, as expected, is very much like a truncated Gaussian.

As hinted before, the Monte Carlo method does not handle particularly well problems “close-to-the-physical-limit” as are microflow rates or nanovolumes which are widely used in health applications. It is therefore important to characterize errors and be able to qualify all measurements in this area through the correct evaluation of the measurement uncertainty. There are several similar problems in engineering, such as roughness measurements of very smooth surfaces or in chemistry when dealing with detection of components, see also [22,23].

8. 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 prove to be 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 for cases where the number of negative values is not significant, with the advantage of having often a simpler application, despite requiring a software implementation.

However, this robust alternative method may also prove inadequate in other circumstances, e.g., in cases too close to the physical limit of a system when the number of out of bound values is significant. 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 similar problems where that evaluation is not a trivial matter.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

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.
Appendix A

This appendix displays the density of \( Q_{\text{obs}} \) obtained with Monte Carlo sampling in the distributions of \((M_F, M_I, T_F, T_I)\) given in Table 1, for case study 1 in Fig. A1 and for case study 2 in Fig. A2.

Fig. A1. Plot of the distribution of the random variable \( Q_{\text{obs}} \) associated with the observed flowrate for case study 1.

Fig. A2. Plot of the distribution of the random variable \( Q_{\text{obs}} \) associated with the observed flowrate for case study 2.

References

[1] P.T. Lee, F. Thompson, H. Thimbleby, Analysis of infusion pump error logs and their significance for health care, Br. J. Nurs. 21 (8) (2012). S12, S14, S16-20.
[2] L. Heinemann, G.A. Fleming, J.R. Petrie, R.W. HolI, R.M. Bergenstal, A.L. Peters, 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) (2015) 716–722.
[3] H. Ma, M.A. Lovich, R.A. Peterfreund, Quantitative analysis of continuous intravenous infusions in pediatric anesthesia: Safety implications of dead volume, flow rates, and fluid delivery, Paediatr. Anaesth. 21 (2011) 79–86.
[4] M. Huensch, C. Sullivan, D. Rooney, C. Barnard, M. Fotis, J. Clarke, G. Noskin, Insights from the sharp end of intravenous medication errors: implications for infusion pump technology, Qual Saf Health Care. 14 (2) (2005 Apr) 80–86.
[5] E. Batista, et al., New EMPIR project – Metrology for Drug Delivery, Flow Measur. Instrum. 72 (2020). 103176.
[6] EMPIR project 17NRM05 EMUE – Examples of measurement uncertainty evaluation. Program available on www.euramet.org.
[7] JCGM 100:2008 GUM 1995 with minor corrections Evaluation of measurement data – Guide to the expression of uncertainty in measurement. BIPM.
[8] 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. BIPM.
[9] ISO 4787:2010 – Laboratory Glassware – Volumetric Glassware – Methods for use and Testing of Capacity.
[10] E. Batista, N. Almeida, I. Godinho, E. Filipe, Uncertainty Calculation in Gravimetric Microflow Measurements, AMCTM X 86 (2015) 58–104.
[11] H. Bissig, H.T. Petter, P. Lucas, E. Batista, et al., Primary standards for measuring flow rates from 100 nL/min to 1 mL/min – gravimetric principle, Biomedical Eng. 60 (4) (2015).
[12] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, D.B. Rubin. Bayesian Data Analysis, third ed., Chapman & Hall/CRC, Boca Raton, Fl, 2014.
[13] S. Demeyere, N. Fischer, C. Elster, Guidance on Bayesian uncertainty evaluation for a class of GUM measurement models, 2021 Metrologia, vol. 58, 014001.
[14] C. Robert, G. Casella, Monte Carlo Statistical Methods, Springer Science & Business Media, 2013.
[15] K. Klausenberg, C. Elster, Markov chain Monte Carlo methods: an introductory example, Metrologia 53 (1) (2016) S32.
[16] J. Salvatier, T.V. Wiecki, C. Fennelbeck, Probabilistic programming in Python using PyMC3, Peer J. Computer Sci. 2 (2016), e55, https://doi.org/10.7717/peerj-cs.55.
[17] E. Batista, I. Godinho, R. Martim, R. Mendes, J. Robarts, Development of an experimental setup for microflow measurement using interferometry, Flow Measur. Instrum. 75 (2020).
[18] M.C. Ausin, Markov Chain Monte Carlo methods: an introductory example, Metrologia 53 (1) (2016) S32.
[19] J.M. Bernardo, A.F.M. Smith, Bayesian Theory, Wiley & sons, 1994.
[20] L. Martino, D. Luengo, J. Míguez, Independent Random Sampling Methods, Springer, 2018.
[21] R. Willink, Measurement of small quantities: further observations of Bayesian methodology, Accreditation and Quality Assurance, 15, pp. 521-527, 2010.
[22] EURACHEM/CITAC Guide CG 4 Quantifying Uncertainty in Analytical Measurement, third ed., 2012.