ACTA IMEKO
ISSN: 2221-870X
March 2026, Volume 15, Number 1, 1 - 6

Monte Carlo analysis of gate-time-resolved uncertainty and oscillator noise in frequency measurements

Assaf Alassaf[1], Waleed M. Al Harbi[1], Khalid S. Al Dawood[1], Ramiz Hamid[1]

[1] *National Measurement and Calibration Center (NMCC), Saudi Standards, Metrology and Quality Organization (SASO), Riyadh, Kingdom of Saudi Arabia*

ABSTRACT
This study presents a gate-time-resolved evaluation of frequency-measurement uncertainty using the GUM analytical framework and large-scale Monte Carlo simulations. The experiment employed a time-interval counter (TIC) disciplined by a 10 MHz reference from a Cesium (Cs) clock to measure phase differences against a Rubidium (Rb) oscillator. Phase data were collected continuously over two days to analyse oscillator stability, and additional frequency datasets at 1 ms, 100 ms, and 1 s gate times, each spanning ten minutes, were examined to assess the effect of averaging time on measurement uncertainty. Uncertainty was evaluated using the GUM model and a Python-based parametric Monte Carlo simulation with one million iterations. A moving-block bootstrap (MBB) method was additionally applied to the same data to assess the influence of time-correlated noise. The results show that GUM slightly overestimated expanded uncertainty at short gate times (−8.38 % at 1 ms, −3.37 % at 100 ms, 0 % at 1 s). Kurtosis analysis revealed non-Gaussian behaviour at shorter averaging times, while Allan-variance analysis identified transitions between white, flicker, and drift noise regimes. These findings demonstrate that simulation-based approaches can capture temporal correlations more effectively, enabling realistic and automated uncertainty evaluation in time-and-frequency metrology under the Metrology 4.0 framework.

Section: RESEARCH PAPER

Keywords: gate time analysis; Monte Carlo bootstrap; oscillator noise process; time and frequency metrology

Citation: A. Alassaf, W. M. Al Harbi, K. S. Al Dawood, R. Hamid, Monte Carlo analysis of gate-time-resolved uncertainty and oscillator noise in frequency measurements, Acta IMEKO, vol. 15 (2026) no. 1, pp. 1-6. DOI: 10.21014/actaimeko.v15i1.2135

Section Editor: Leonardo Iannucci, Politecnico di Torino, Italy

Received June 18, 2025; **In final form** December 14, 2025; **Published** March 2026

Corresponding author: Assaf N. Alassaf, e-mail: a.assaf@saso.gov.sa

1. INTRODUCTION

In metrology, evaluating measurement uncertainty is fundamental to reliable practice. The Guide to the Expression of Uncertainty in Measurement (GUM) [1] has long provided the standard analytical framework, with Monte Carlo methods later introduced to address cases in which data deviate from Gaussian behaviour. Previous studies, such as those on hardness measurement [2] and on perspiration systems [3], confirmed the usefulness of Monte Carlo methods, but focused on relatively simple scenarios in which results remained consistent with the GUM. Time and frequency metrology poses greater challenges: datasets may span days or use millisecond-scale gate times, where Gaussian assumptions often fail and noise processes dominate. This study applies a Python-based [4] Monte Carlo framework [5] to evaluate uncertainty at gate times of 1 ms, 100 ms, and 1 s. In parallel, oscillator noise is characterized using Allan variance, providing insight into how noise processes and instrumentation affect distributional behaviour and how simulation can capture these effects in practice.

2. EXPERIMENTAL SETUP

In this study, a time interval counter (TIC) was disciplined by a 10 MHz signal derived from a Cesium (Cs) clock, while phase differences were measured against a Rubidium (Rb) clock through their 1 PPS outputs (Figure 1). Phase data between the Cs and Rb clocks was collected continuously for approximately two days to investigate oscillator noise characteristics and stability behaviour across a wide range of averaging times.

In addition, three frequency datasets were acquired at gate times of 1 ms, 100 ms, and 1 s, each spanning ten minutes, to examine how averaging time influences measurement uncertainty and stability.

The analysis combines Allan variance and Monte Carlo methods to link frequency stability with numerical uncertainty evaluation. Allan variance is used to characterize noise processes and identify temporal correlations in the data, while Monte Carlo propagates these effects through the measurement model to estimate the resulting uncertainty. The detailed computational procedures are described in Section 3.1.

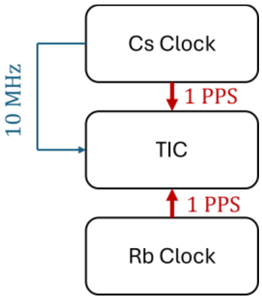

Figure 1. Experimental setup for frequency comparison. The time-interval counter (TIC) was disciplined by a 10 MHz reference derived from a Cesium (Cs) clock. Phase differences between the Cs clock and a Rubidium (Rb) clock were measured through their 1 PPS outputs.

3. RESULTS AND DISCUSSION

3.1. Computational conditions

3.1.1. Monte Carlo framework

The Monte Carlo method estimates measurement uncertainty by repeatedly sampling input quantities from their probability distributions and propagating them through the measurement model to obtain a distribution of output values. In this work, the simulation was implemented in Python using NumPy, SciPy, Matplotlib, and Pandas, with a fixed random seed (42) for reproducibility [5]–[9]. Each run consisted of 1,000,000 iterations, and the measurement model is given by:

$$f = f_0 + \varepsilon_{\mathrm{A}} + \varepsilon_{\mathrm{random}} + \varepsilon_{\mathrm{tb}} + \varepsilon_{\mathrm{sys}}\,, \tag{1}$$

where

- f_0 - Nominal frequency (10 MHz)
- ε_{A} - Type A error
- $\varepsilon_{\mathrm{random}}$ - Random jitter error
- $\varepsilon_{\mathrm{tb}}$ - Time-base error
- $\varepsilon_{\mathrm{sys}}$ - Systematic error.

Each term ε_i in the model represents an individual error component, while the corresponding standard deviation u_i defines its associated uncertainty. Because each term contributes directly and linearly to the measured frequency, the sensitivity coefficients are unity by definition. The statistical distributions assigned to these quantities and used as sampling parameters in the Monte Carlo simulation are summarized in Table 1.

3.1.2. Allan-variance analysis

The Allan variance is a time-domain statistical tool used to quantify the frequency stability of oscillators and other time-series signals. It measures how the fractional-frequency fluctuations y vary as a function of averaging time τ. For a discrete sequence sampled at intervals (gate time) τ_0, the overlapping Allan variance $\sigma_y^2(\tau)$ is defined as

$$\sigma_y^2(\tau) = \frac{1}{2(N-2\,m+1)} \sum_{i=1}^{N-2\,m+1} \left(\overline{y}_{i+m} - \overline{y}_i\right)^2\,, \tag{2}$$

where $\tau = m\,\tau_0$, N is the total number of samples, and $\bar{y}_i$ is the mean of m consecutive fractional-frequency values. The overlapping form uses all possible adjacent windows, rather than disjoint ones, increasing the number of samples contributing to each τ and thereby reducing the estimator's variance without changing its expected value.

Table 1. Characteristics of the uncertainty components, including their classification under the GUM framework, employed symbols (Source), (Type), and assumed statistical distributions (Distribution).

Source	Type	Distribution
u_{A}	A	Normal
u_{random}	B	Rectangular
u_{tb}	B	Rectangular
u_{sys}	B	Rectangular

In this work, the overlapping Allan deviation $\sigma_y(\tau)$ was computed from the measured fractional-frequency data using the open-source allantools library in Python. The resulting Allan deviation plots were used to identify the dominant noise regimes—white, flicker, drift—and to interpret how frequency stability evolves with averaging time.

3.2. Gate-time analysis

Frequency counters determine the input frequency by counting the number of signal cycles within a fixed gate time τ_0, the interval during which the counter's gate remains open. The choice of gate time has a direct effect on both resolution and measurement uncertainty. Short gate times yield faster updates and allow tracking of rapid frequency variations, but they introduce higher statistical noise due to the limited number of counted cycles. Conversely, longer gate times improve frequency resolution by averaging out random fluctuations, though at the cost of slower acquisition and reduced temporal responsiveness.

The gate time primarily affects noise and uncertainty introduced by the counter and measurement process, such as quantization and random jitter, rather than the intrinsic noise of the oscillator itself. These instrumental effects scale differently with averaging: quantization and jitter decrease approximately as $1/\sqrt{\tau}$, while time-base and systematic contributions remain largely independent of gate duration. In contrast, the oscillator's internal noise characterized later by Allan variance originates from the source's physical stability and is independent of the counter's gate setting. The scaling behaviour of each uncertainty component used in this work is summarized in Table 2.

3.2.1. Gate Time Results

To investigate the influence of gate time on measurement uncertainty, the evaluation was first carried out analytically using the GUM framework, as summarized in Table 2, and then verified through the Monte Carlo framework described in Section 3.1.1. The Monte Carlo approach provides a complementary means of verification and further enables the generation of additional graphical insights, which can be used to assess the statistical nature of the measurement data, as illustrated in Figure 2.

Table 2. Individual standard uncertainty values for each source at different gate times (1 ms, 100 ms, and 1 s), along with the calculated combined standard (u_c) and expanded ($U = 1.96 \cdot u_c$) uncertainties.

Source	At 1 ms (Hz)	At 100 ms (Hz)	At 1 s (Hz)
u_{A}	10.04×10^{-5}	4.36×10^{-5}	1.23×10^{-5}
u_{random}	5.02×10^{-2}	5.02×10^{-4}	5.02×10^{-5}
u_{tb}	2.10×10^{-3}	2.10×10^{-4}	6.70×10^{-5}
u_{sys}	2.00×10^{-2}	2.00×10^{-4}	2.00×10^{-5}
u_{c}	3.13×10^{-2}	3.79×10^{-4}	7.46×10^{-5}
$U = 1.96.u_{\mathrm{c}}$	6.13×10^{-2}	7.42×10^{-4}	1.46×10^{-4}

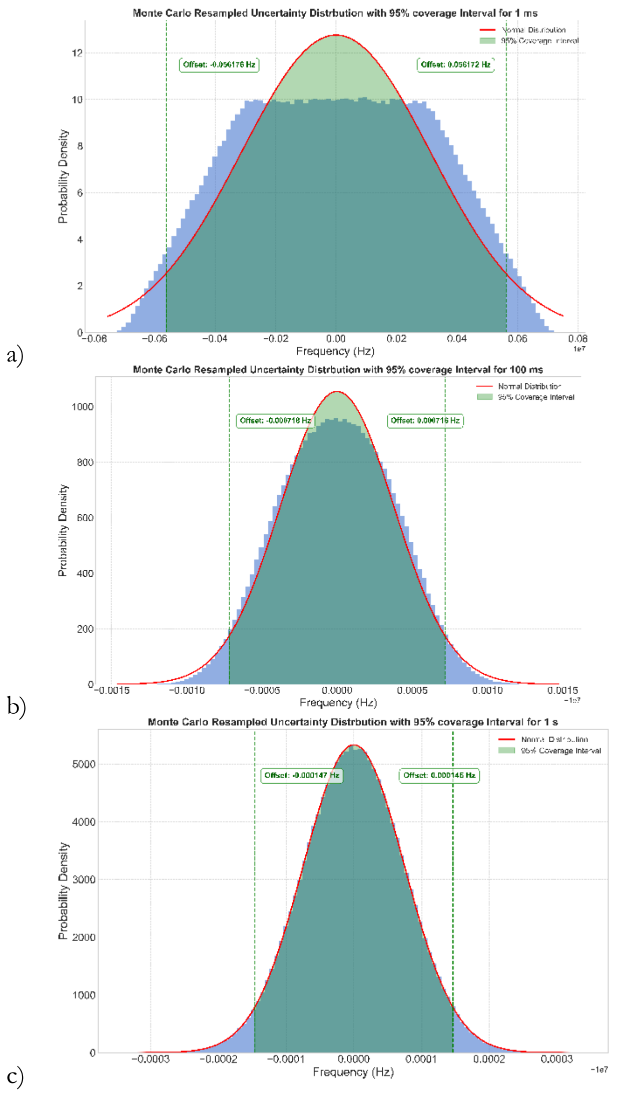

Figure 2. Monte Carlo resampled uncertainty distributions with normal distribution overlays for different gate times: a) 1 ms, b) 100 ms, and c) 1 s.

To better understand the differences between the GUM and Monte Carlo results, kurtosis was employed as a measure of normality. Excess kurtosis, defined as the deviation of kurtosis from the Gaussian reference value of zero, provides insight into whether a distribution is more peaked (leptokurtic, positive values) or flatter (platykurtic, negative values) than the normal distribution:

$$kurtosis = \frac{E(X-\mu)^4}{s^4}, \tag{3}$$

where E denotes the expected value, μ is the mean of the distribution, and s is the standard deviation, where X represents the random variable. As summarized in Table 3, the measurement results obtained from both the GUM framework and the Monte Carlo simulations are presented alongside the corresponding kurtosis values, which highlight deviations from Gaussian behaviour at shorter gate times. In addition, the dominant contributors to measurement uncertainty were found to vary with gate time: at 1 ms, systematic uncertainty accounted for approximately 99.5 % of the total; at 100 ms, random uncertainty dominated with about 75 %; and at 1 s, time base accuracy became the primary contributor with roughly 70 %. This shift in dominant components underscores the importance of gate-time-resolved analysis when characterizing frequency measurement uncertainty. Analytical results were derived using (k = 1.96). At shorter gate times, deviations from Gaussian behaviour are quantitatively confirmed by kurtosis analysis. As shown in Table 3, the kurtosis values are negative across all gate times, indicating platykurtic distributions. The effect is strongest at 1 ms (***kurtosis*** = – 0.907), reflecting a pronounced departure from normality. At 100 ms, the distribution is still somewhat platykurtic (***kurtosis*** = – 0.427), while at 1 s, the distribution approaches Gaussian behaviour with a kurtosis close to zero (***kurtosis*** = – 0.03). These results demonstrate that the assumption of normality is less valid at short gate times.

Table 3. Comparison of expanded uncertainties obtained by the GUM method and coverage intervals obtained by the Monte Carlo method for three gate times (1 ms, 100 ms, and 1 s). The relative and absolute differences are calculated with respect to the GUM values.

Gate Time	GUM Expanded Uncertainty (Hz)	Monte Carlo 95 % Coverage (Hz)	Relative Difference (%)	Absolute Difference (Hz)	Kurtosis
1 ms	± 0.061314	± 0.056174	− 8.38%	0.005140	-0.907
100 ms	± 0.000742	± 0.000717	− 3.37%	0.000025	-0.427
1 s	± 0.000146	± 0.000146	0.00%	0.000000	-0.030

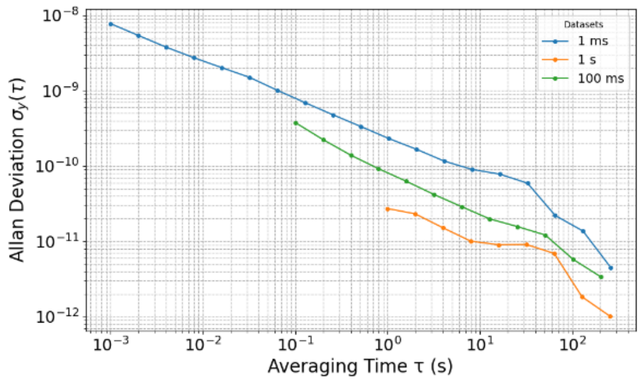

Figure 3. Allan deviation of frequency measurements at gate times of 1 ms (blue), 100 ms (green), and 1 s (orange).

3.2.2. Gate-time effects on stability

To further assess the stability and accuracy of the frequency measurements, the Allan deviation of the Rb clock data was analysed for gate times of 1 ms, 100 ms, and 1 s, as shown in Figure 3. Although all datasets originated from the same oscillator, the time-interval counter (TIC) noticeably influences the observed stability. At shorter gate times, the TIC's quantization and feedback limitations increase the apparent instability by introducing high-frequency noise and measurement jitter.

As evident in Figure 3, stability improves progressively with longer gate times. Random short-term fluctuations are averaged out over wider intervals, reducing the Allan deviation. At 1 ms, the measurements are dominated by high-frequency noise and counter quantization effects; at 100 ms, this influence is reduced but still visible; and at 1 s, the intrinsic oscillator stability becomes dominant, yielding the lowest Allan deviation.

These results confirm that extending the gate time enhances measurement stability by suppressing short-term fluctuations and revealing the underlying performance of the oscillator.

3.3. Characterization of Oscillator Noise Processes

The deviations from Gaussian behaviour discussed earlier can be traced back to two main sources: the instrumental effects from the time interval counter (TIC) and the intrinsic noise of the oscillator itself. The TIC contributes quantization and feedback artefacts that mainly affect short-term statistics, whereas the oscillator introduces fundamental noise processes that dominate over longer averaging times.

Cesium (Cs) clocks typically exhibit white and flicker frequency noise, leading to excellent long-term stability [10], [11]. Rubidium (Rb) clocks, however, display a broader combination of white noise, flicker noise, and slow drift [11], which makes them ideal for illustrating typical oscillator noise behaviour. The overlapping Allan deviation of the Rb data (Figure 4) reflects these regimes, with distinct slopes corresponding to white-noise, flicker-noise, and drift-dominated intervals.

Figure 5 further visualizes how these noise types manifest in the data. At short averaging times, rapid uncorrelated fluctuations produce nearly Gaussian histograms, consistent with white noise. As the averaging time increases, correlations accumulate and the distribution broadens, revealing flicker-like behaviour. At the longest times, slow monotonic trends dominate, and the distribution approaches a rectangular shape associated with drift.

Understanding these transitions helps relate the statistical behaviour of the data to the underlying noise sources.

These correlated fluctuations directly influence how measurement uncertainty evolves with averaging time. Under the analytical GUM framework, the standard deviation of the mean is expressed as $s_y = \frac{s}{\sqrt{m}} = \frac{s}{\sqrt{\tau/\tau_0}}$ where s is the standard deviation of the individual frequency samples, τ_0 is the sampling interval (gate time), and $m = \tau/\tau_0$ is the number of samples averaged over the time τ. This analytical model assumes statistically independent samples and predicts $\tau^{-1/2}$ decrease in uncertainty with increasing averaging time. However, as Figure 4 and Figure 5 demonstrate, real oscillator data are time-correlated due to colored noise processes such as flicker and drift.

To capture these correlations, a moving-block bootstrap (MBB) analysis was implemented on the same dataset. For each averaging time τ, the frequency record was divided into non-overlapping segments of length τ, and the mean of each segment was computed to form a τ-averaged series. The MBB then generated 5000 synthetic realizations by randomly resampling overlapping blocks of consecutive τ-averaged data (each block containing $L \approx K^{1/3}$ samples, where K is the number of averages), thereby preserving the temporal correlation structure of the noise [9]. Each realization was truncated to the original length, and its standard deviation computed. The resulting distribution provided the bootstrap mean and 95 % confidence interval, shown in Figure 6. The bootstrap-derived standard deviation decays more slowly than the idealized GUM prediction, reflecting the influence of time-correlated noise on the uncertainty evolution.

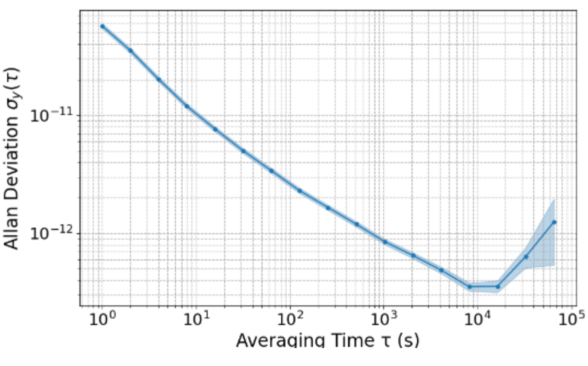

Figure 4. Overlapping Allan deviation of approximately two days of the Rb clock data, where distinct regions can be identified: at short averaging times, the slope follows $\tau^{-1/2}$, indicating white FM noise; at intermediate times, the slope flattens to τ^0, characteristic of flicker FM noise; and at longer times, the slope increases with $\tau^{+1/2}$ reflecting drift.

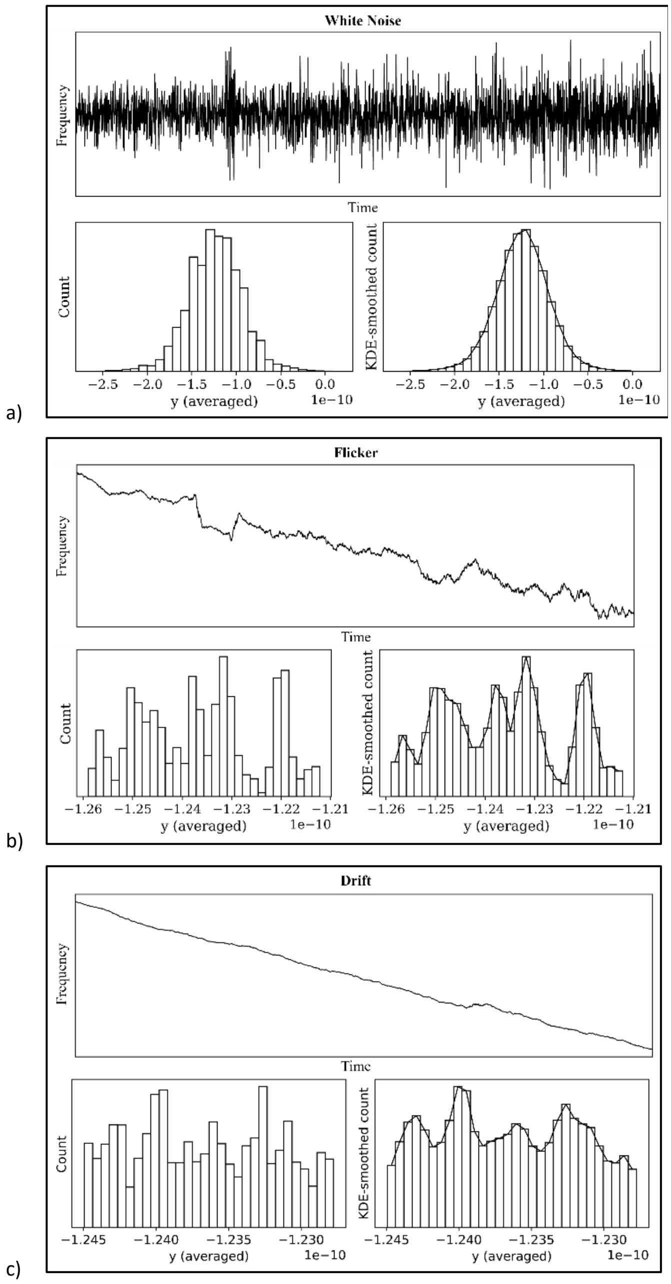

Figure 5. Frequency data at different averaging times showing the effect of dominant noise processes in a Rubidium oscillator.
a) White noise — rapid short-term fluctuations with nearly Gaussian histograms;
b) Flicker noise — slow correlated variations causing departures from Gaussian shape;
c) Drift — gradual long-term trend leading to a rectangular-like distribution.
These panels illustrate how different noise types influence the statistical behaviour of frequency measurements.

3.4. Allan Variance and Monte Carlo complementary analysis for time series measurements

As discussed above, stability is first evaluated from raw time-series data using Allan variance analysis, which characterizes the noise type as a function of averaging time. At short gate times, deviations from Gaussian behaviour may arise from instrumental limitations of the time-interval counter (TIC), such as quantization and feedback effects. These effects influence short-term stability but are distinct from the intrinsic noise processes of the oscillator that dominate at longer averaging times.

When the data exhibits predominantly white-noise behaviour, the uncertainty can be reliably evaluated using the GUM approach, which assumes statistical independence between samples. However, as the Allan variance shows, real oscillator data often transition to correlated noise regimes, such as flicker or drift, where the GUM assumption no longer holds. In such cases, the Monte Carlo method provides a complementary solution by accounting for temporal correlations and deviations from normality, leading to a more realistic estimate of uncertainty.

From this complementary perspective, Allan variance acts as a diagnostic tool, revealing the noise regime and potential non-Gaussian behaviour, while Monte Carlo analysis refines the uncertainty evaluation of when correlations or instrumental effects are significant. In practice, Allan variance should first be used to assess the dominant noise type and verify the presence of correlations, followed by examining the distribution of the average data at each gate time. Based on these findings, the appropriate uncertainty approach, either GUM or Monte Carlo, can then be selected. This integrated strategy ensures that both instrumental effects and intrinsic noise processes are properly represented in the final uncertainty evaluation.

4. CONCLUSION

This study demonstrated that, in time and frequency metrology, the assumption of Gaussian-distributed data does not always hold. At short gate times, such as 1 ms, quantization and feedback limitations of the time-interval counter (TIC) cause the distribution to deviate strongly from normality, as reflected by a kurtosis of – 0.907. These effects also reduce the apparent stability of the measurement, underscoring the need for statistical methods that can more reliably capture the underlying behaviour. At longer gate times, the distributions narrow and become closer to Gaussian, with kurtosis approaching zero (– 0.030 at 1 s), highlighting how averaging suppresses short-term fluctuations.

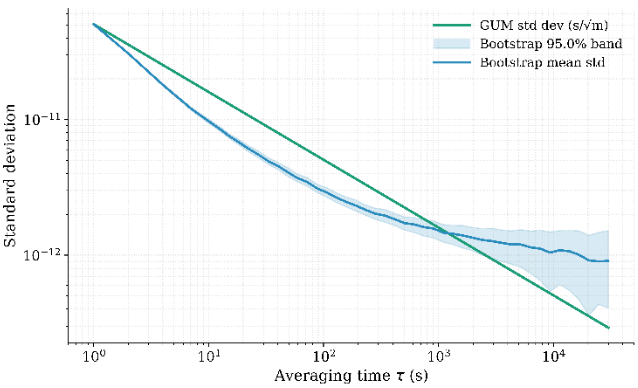

Figure 6. Standard deviation of Rb data: GUM prediction $s/\sqrt{m}$ versus Monte Carlo bootstrap. Bootstrap accounts for data correlations, showing slower decay and deviation from GUM. The shaded region is the 95 % confidence interval.

The numerical comparison further showed that GUM can overestimate expanded uncertainty by up to 8.38 % at 1 ms, while differences decrease at longer gate times (3.37 % at 100 ms, 0 % at 1 s). In parallel, the physical nature of atomic oscillators introduces characteristic noise processes (white noise, flicker noise, drift) that reinforce this departure from Gaussian statistics. These results indicate that methods assuming uncorrelated normal samples may yield misleading outcomes, while simulation-based approaches such as Monte Carlo and bootstrap resampling provide more realistic uncertainty estimates by preserving data correlations. It is worth noting that the GUM Supplements already acknowledge such limitations of the classical framework and recommend Monte Carlo methods for cases involving non-linear models or non-Gaussian inputs [5]. Our results are therefore consistent with these recommendations and provide a practical demonstration of their relevance in time and frequency metrology.

Beyond methodological insights, this work supports the digital transformation of metrology. By providing openly available Python code and datasets, the study enables the validation of results, offers value for education and training, and facilitates the integration of automated uncertainty evaluation into laboratory workflows. This aligns with the objectives of Metrology 4.0, which extends Industry 4.0 principles to measurement science by promoting digitalization, automation, and connectivity in calibration practices. Under Metrology 4.0, reliable, automated, and reproducible data analysis becomes central to ensuring traceability and efficiency in modern laboratories [12]–[14]. The full code and datasets, and all supporting tools used in this work are available at: https://github.com/Assaf-SASO/monte-carlo-frequency-analysis (MIT licensed). Altogether, this work not only clarifies the statistical and physical origins of uncertainty in frequency measurements but also provides practical tools and open resources to advance modern metrology practices.

REFERENCES

[1] BIPM, JCGM 100:2008, Evaluation of measurement data – Guide to the expression of uncertainty in measurement, 2008. DOI: 10.59161/JCGM100-2008E

[2] G. M. Mahmoud, R. S. Hegazy, Comparison of GUM and Monte Carlo methods for the uncertainty estimation in hardness measurements, International Journal of Metrology and Quality Engineering vol. 8 (2017) 14. DOI: 10.1051/ijmqe/2017014

[3] A. Chen, C. Chen, Comparison of GUM and Monte Carlo methods for evaluating measurement uncertainty of perspiration measurement systems, Measurement vol. 87 (2016), pp. 27-37. DOI: 10.1016/j.measurement.2016.03.007

[4] S. Eichstädt, C. Elster, I. M. Smith, T. J. Esward, Evaluation of dynamic measurement uncertainty – an open-source software package to bridge theory and practice, Journal of Sensors and Sensor Systems vol. 6 (2017) no. 1, pp. 97-105. DOI: 10.5194/jsss-6-97-2017

[5] BIPM, 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, 2008. DOI: 10.59161/JCGM101-2008

[6] J. C. Wu, A. F. Martin, G. Sanders, R. N. Kacker, Bootstrap method versus analytical approach for estimating uncertainty of measure in ROC analysis on large datasets, NIST Interagency/Internal Report (NISTIR), National Institute of

Standards and Technology, Gaithersburg, MD.
DOI: 10.6028/NIST.IR.8218
[7] A. M. H. Van Der Veen, M. Cox, J. Greenwood, A. Bosjnakovic, V. Karahodzic, S. Martens, K. Klauenberg, C. Elster, (+ 63 more authors), Good practice in evaluating measurement uncertainty – Compendium of examples, Bureau International des Poids et Mesures (BIPM), 2021.
https://zenodo.org/records/5142180
[8] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, (+ 18 more authors), Array programming with NumPy, Nature vol. 585 (2020), pp. 357-362.
DOI: 10.1038/s41586-020-2649-2
[9] D. J. Sheskin, Handbook of Parametric and Nonparametric Statistical Procedures, Fifth Edition. Chapman and Hall/CRC, New York, 2011.
DOI: 10.1201/9780429186196
[10] D. W. Allan, N. Ashby, C. C. Hodge, The Science of Timekeeping. Hewlett Packard Application Note 1289, 1997.
[11] J. Vanier, C. Audoin, The Quantum Physics of Atomic Frequency Standards. CRC Press, New York, 1989.
DOI: 10.1201/9781003041085
[12] M. Wieczorowski, J. Trojanowska, Towards Metrology 4.0 in Dimensional Measurements, Journal of Machine Engineering vol. 23 (2023) no. 1, pp. 100-113.
DOI: 10.36897/jme/161717
[13] P. J. De Groot, M. Schmidt, Metrology & Industry 4.0, PhotonicsViews vol. 18 (2021) no. 4, pp. 73-75.
DOI: 10.1002/phvs.202100053.
[14] J. M. Nzumile, V. Mahabi, I. W. R. Taifa, Towards Metrology 4.0 in Developing Countries' Manufacturing Industries, Tanzania Journal of Engineering and Technology vol. 44 (2025) no. 2, pp. 181-198. Online [Accessed 12 February 2026].
https://journals.udsm.ac.tz/index.php/tjet/article/view/9209