

## DETERMINATION OF MEASUREMENT UNCERTAINTY BY A MONTE CARLO METHOD FOR AN RF POWER SENSOR CALIBRATION SYSTEM USING A VNA

Marek Jaworski, Jarosław Szatkowski, Tomasz Kossek

National Institute of Telecommunications (NIT), Warsaw, Poland (✉ [m.jaworski@il-pib.pl](mailto:m.jaworski@il-pib.pl), [j.szatkowski@il-pib.pl](mailto:j.szatkowski@il-pib.pl), [t.kossek.@il-pib.pl](mailto:t.kossek.@il-pib.pl))

### Abstract

This work proposes a systematic assessment of measuring type A uncertainty (caused by random errors) used in RF power sensor calibration. To reduce A type uncertainty, several successive measurements are repeated. The uncertainty arises from repeatability errors in connectors caused by changes in their electrical properties during repeated mating. The suitability of the METAS UncLib software was analysed and we concluded that software should be developed to take into account the shape of probability density function (PDF) using a Monte Carlo method (MCM), which was lacking in METAS UncLib. The self-developed software was then tested on an example taken from the literature and the superiority of the MCM over the analytical method (GUM) was confirmed. During the calibration of the RF sensor using a vector network analyzer (VNA), a series of repeated measurements were performed and, after applying our MCM software, it was found that the measurement uncertainties calculated by the MCM method were several times larger than those by the GUM. The reason for this was that the correlation between the measured input quantities was not taken into account. When this was done using a covariance matrix and assuming a normal PDF of the input quantities, the results obtained with the GUM and the MCM converged. Our main objective was to investigate the influence of the PDF shape of the input measurement samples on the measurement uncertainty. Taking more than a dozen measurements is too costly, on the other hand, the small sample size prevents a reliable determination of the PDF shape. Finally, to overcome this inconvenience, we have developed a special method that uses the histograms of standardized input data taken at all measurement frequencies under fixed conditions without disconnecting the connectors, to increasing the total number of results which were needed to create the PDF histograms of input quantities.

Keywords: Monte Carlo method (MCM), type A uncertainty, RF power sensor calibration.

© 2023 Polish Academy of Sciences. All rights reserved

## 1. Introduction

In the *Laboratory of Electrical, Electronic and Optoelectronic Metrology* (LMEEiO) at the NIT, we have so far used the analytical method (known as the GUM after the “Guide to the Expression of Uncertainty in Measurement” [1]) for determining high-frequency (RF) measurement

uncertainty as recommended in [2]. However, the document [2] also allows for an alternative Monte Carlo method (MCM), depicted in detail in [3]. The MCM involves determining the PDF of the output quantity  $Y$  by generating sufficiently numerous samples of input quantities  $X_i$  with an assumed PDF and transforming them using measurement function  $Y = f(X_1, \dots, X_N)$ . From the samples  $Y$  thus created, various statistical properties of the output quantity  $Y$  can be determined, for example: the estimate  $mY$ , the standard uncertainty  $sY$  as well as the lower  $Y_{\text{low}}$  and upper  $Y_{\text{high}}$  limits of the uncertainty interval  $dY$ , for an assumed confidence level  $p$ . The MCM has an advantage over the GUM when the estimated uncertainty of measurement consists mainly of measurement errors with a PDF significantly deviating from the normal distribution and/or when it is determined from a small number  $n$  of measurement samples. The low popularity of the MCM was due to the need to use multiple (up to  $10^6$ ) samples of measured quantities with appropriately chosen PDF distributions, but with the currently available computing power of personal computers, the estimation of uncertainty interval at  $p = 95\%$  confidence level usually takes a few seconds, so the MCM is now growing in popularity [4, 5].

We have developed software in the form of Octave scripts to determine the two-dimensional uncertainty (*i.e.* in the domain of complex numbers) of measurements using the MCM [6, 7] according to guidelines [8].

In order to validate the developed software, we repeated simulations carried out in the publication [9] using the MCM and the results obtained were in agreement with those presented in [9]. The part of the calculation related to the GUM was performed using the open-access software METAS UncLib [10]. Section 2 presents a brief description of the operation and evaluation of the suitability of METAS UncLib software for the determination of RF measurement uncertainties at the LMEEiO.

Due to a discrepancy between the results obtained by the GUM and the MCM of estimating the calibration uncertainty of RF power sensors, which was pointed out in [9] but without explaining the reasons for the discrepancy, we carried out a detailed analysis, presented in Section 3, showing that the reason was nonlinearity of the measurement function  $Y = f(X_1, \dots, X_N)$ . Measurements of similar type as presented in [9] are used at the LMEEiO for calibration of RF power sensors.

We used the self-developed software to determine the mismatch uncertainty in the calibration of an RF power sensor using a VNA. Type A uncertainty is due to connector repeatability errors caused by variations in their electrical characteristics during repeated mating. Again, differences between the results obtained by the GUM and the MCM were observed. However, the reason for this was different from the previous case – it was correlations of the measured input quantities, which required further modification of the MCM simulation model by using a covariance matrix, described in Section 4.

After implementing the MCM in the software, which included the influence of correlations on the measurement uncertainty, marked as MCM', the next stage of the work involved determining a PDF of the measured input quantities and their possible impact on the measurement uncertainty described in Section 5. The ability to take into account the shape of the PDF is the main advantage of the final version of the software, marked as MCM'' that distinguishes it from the GUM [3]. The final conclusions are drawn in Section 6.

## 2. Description and evaluation of the usefulness of the METAS UncLib software

METAS UncLib is a software library implemented in C# for determining measurement uncertainty, developed at the *Swiss Federal Institute of Metrology* (METAS) [10, 11]. It supports the creation of objects for determining the uncertainty of measurements and subsequent calculations

using them, as well as storing the results. It is able to handle complex and multidimensional quantities. The implementation is object-oriented and overloaded operators hide the complexity of the computation from the user.

A similar approach is used in the GUM tree calculator [12] used to automatically determine the uncertainty of measurements using a software package Python.

The result in the form of a mean value and measurement uncertainty is automatically calculated. The functions describing the sensitivity factors and components of the combined standard uncertainty are also calculated automatically from the measurement function. Uncertainty values are not stored in objects, but can be computed on demand. METAS UncLib can run in three modes:

1. LinProp – supports linear propagation of uncertainty. This operating mode is fully functional.
2. DistProp – supports the propagation of higher-order uncertainties, *i.e.* higher-order terms of the Taylor expansion of the derivatives of a measurement function used to calculate the sensitivity factors are taken into account. This mode currently supports only normal distributions of the input data and the outputs are also normal distributions. The authors of the software declare that in the future it will be possible to use non-normal PDF distributions and to use higher moments of the probability density function to calculate more realistic uncertainty intervals.
3. MCProp – supports propagation of uncertainty using the MCM, but to a limited extent, as it supports only normal distributions of input data.

The usefulness of the METAS UncLib software for determining the uncertainty of measurement in the calibration of RF measuring instruments was tested, and compared with the analogous software developed at the LMEEiO using the MCM.

Similarly to the MCM, in METAS UncLib there is no need to determine the sensitivity coefficients, since they are calculated automatically from the measurement function. The advantage of the METAS UncLib software is the speed of determining the uncertainty of measurement, which is especially important in calibration using a VNA when we are dealing with a large amount of measurement data in the form of complex values. When the result of the measurement is a complex value, the uncertainty takes the form of an elliptical region, unlike the one-dimensional case where the uncertainty is a range of values. In the process of developing the MCM software for complex values, we encountered a problem with extending the calculation time and the available memory size when calculating the measurement uncertainty area with required reliability. This problem is eliminated in the case of analytical calculation of uncertainty using the METAS UncLib software.

The advantage of the MCM is that the PDF of input quantity can be arbitrarily shaped, hence the MCM can be regarded as a reference method. However, in the measurement practice of the laboratory calibrating measuring instruments, the measurement uncertainty is usually small compared to the measured value and then the central limit theorem works well, which leads to the convergence of the measurand PDF estimation to the normal distribution. This is particularly true for measurements of complex values. In the case of random errors (type A), where it is necessary to average the results of successive measurements, the number of measurements necessary to reliably determine the shape of the two-dimensional PDF is so large that, in practice, a two-dimensional normal distribution that takes into account the correlation of the results is often used, and can therefore be fully characterised in the LinProp mode of the METAS UncLib software.

METAS UncLib has the ability to include correlations between any number of measurement inputs. A suitable covariance matrix should be prepared in advance. The same functionality is implemented in the latest version of the self-developed MCM' software.

The METAS UncLib software has plugins for MATLAB [13] and Python [14], which significantly increases its functionality. One drawback to the practical use of the METAS UncLib software at the LMEEiO is that it cannot be used in the Octave environment, for which scripts have been developed to determine measurement uncertainty. The simplest but expensive alternative is to purchase a MATLAB licence and use the Octave code with minor modifications, or to write these scripts in Python.

In conclusion, the METAS UncLib software is a proven, open-access tool supporting the determination of measurement uncertainty due to its ease of use, high accuracy and speed of calculations. However, it is not possible to run MCM simulations with non-Gaussian PDFs of the input quantities  $X_i$ .

This was the motivation to develop software with such functionality and then test, using a specific example, whether the consideration of the PDF shape of the measured input quantities  $X_i$  is relevant for the determination of the measurement uncertainty  $Y$ .

### 3. Validation of MCM software by comparison with previously published results in [9]

Publication [9] presents a comparison of results of determining the limit values of type A measurement uncertainty (caused by a random error from the mating connectors) for the mismatch factor  $M$  obtained by the GUM and the MCM during the calibration of the RF power sensor. Measurements performed made in the system shown in Fig. 1, consisting of a generator (RF power source) and a power splitter to which the RF source and source power monitoring sensor are permanently connected, and the power sensors *already calibrated* (standard) and undergoing calibration (*device under test* – DUT) are connected interchangeably.

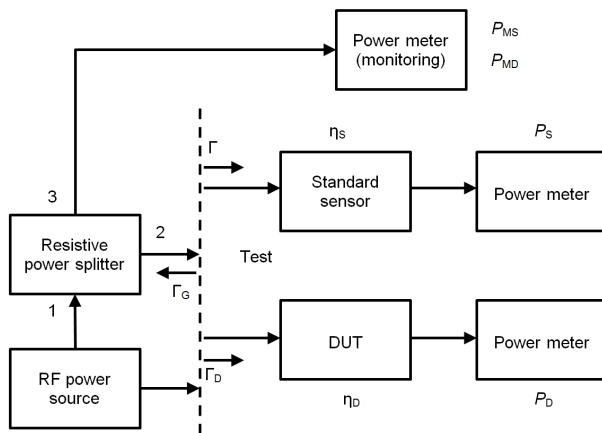


Fig. 1. Block diagram of the RF power sensor calibration system based on direct comparison [9].

The direct comparison method is usually used to calibrate the RF power sensor [15]. The method consists in interchangeably connecting the standard power sensor and the DUT to the RF power source. The effective efficiency of the calibrated RF power sensor is compared to the efficiency of the DUT. In such a comparison, there is a mismatch factor  $M$ , which is the cause of RF power loss [16]. The mismatch is due to reflections from both the standard power sensor  $\Gamma_S$ , the DUT  $\Gamma_D$  and the source  $\Gamma_G$ . The DUT and the standard power sensor are connected interchangeably to the RF source through a resistive power splitter. The mismatch factor is

derived from the measured reflection coefficients  $\Gamma$  and is used to correct the calibration error of the effective efficiency of the sensor. The measurement uncertainties of the coefficients  $\Gamma$  contribute to the measurement uncertainty of the effective efficiency of the DUT through the mismatch factor  $M$ . Therefore, it is important to reliably determine the uncertainty of  $M$  when calibrating the power sensor.

The mismatch factor  $M$  is defined as [9]:

$$M = \frac{(1 - |\Gamma_S|^2)(|1 - \Gamma_D \Gamma_G|^2)}{(1 - |\Gamma_D|^2)(|1 - \Gamma_S \Gamma_G|^2)}, \quad (1)$$

where  $\Gamma_S$ ,  $\Gamma_D$ ,  $\Gamma_G$  are, respectively, the reflection coefficients of the standard sensor, the DUT sensor, and the equivalent source of power splitter, which can be obtained from the conventional one-port VNA calibration method [13, 15], or calculated from the expression [9]:

$$\Gamma_G = S_{22} - \frac{S_{12}S_{23}}{S_{13}}, \quad (2)$$

where  $S_{ij}$  are the elements of the power splitter s-matrix, measured with a two-port VNA.

The  $M$  factor is a component that takes into account the effect of the impedance mismatch when calibrating a RF power sensor using the direct comparison method, according to the function [9]:

$$\eta_D = \eta_S \frac{P_D P_{MS}}{P_S P_{MD}} M \quad (3)$$

in the measurement system shown in Fig. 1.

In order to verify the developed software, the simulations performed in publication [9] using the MCM were repeated and the results were almost identical to those in the publication mentioned above.

The authors of [9] used a PDF consisting of uniform and normal distributions and concluded that the obtained results show, in special cases (*e.g.* for the measurement frequency  $f = 1$  GHz), a large discrepancy (even twofold) in determining the uncertainty interval using the MCM compared to the GUM. However, they did not provide the reason for this discrepancy.

The GUM assumes that the PDF of the measurand  $Y$  is normally distributed. We have checked that even when the simplification is applied, *i.e.* using only a normal PDF for all components listed in (1) and (2), there is still a discrepancy in the results between the GUM and the MCM, even greater than in the case considered in [9].

Figure 2 shows exemplary PDF graphs of the  $S_{22}$  element of the power splitter s-matrix, in the form of a histogram on a logarithmic scale, to highlight the shape differences adopted for the simulation of distributions.

In this case, the cause of the discrepancy cannot be attributed to correlations of the input values, as such correlations do not occur here. The PDF histogram of the  $M$  factor determined from (1) and (2) at 1 GHz, shown in Fig. 3b, differs significantly from the normal PDF, which requires finding the reason for these differences.

It was assumed that the discrepancy was probably due to GUM's analytical model. The METAS UncLib software was used for calculations using the GUM [13]. It automates the determination of sensitivity coefficients, which is particularly helpful for complex measurement functions, *e.g.* here after substituting (2) into (1), and additionally enables the determination of higher-order expansion coefficients – see Section 2 for a more detailed description. It turns

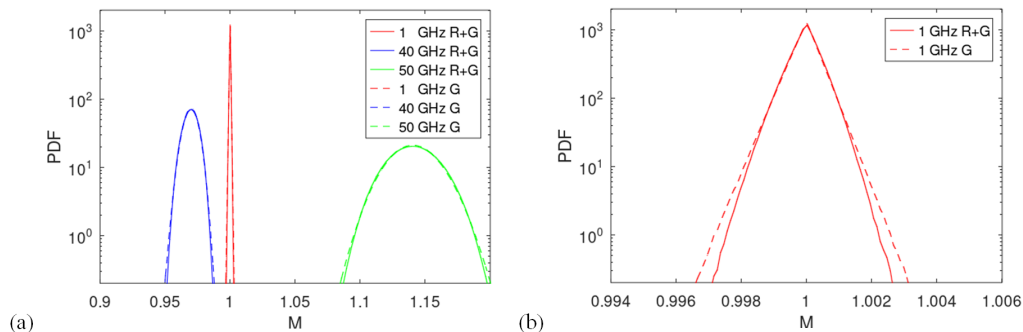


Fig. 2. PDF: (a) real part, (b) imaginary part of the  $S_{22}$  of the power splitter  $s$ - matrix. (Solid line – PDF being the sum of the normal (G) and rectangular (R) distribution, etc., dashed line – normal PDF with the same standard uncertainty as G + R).

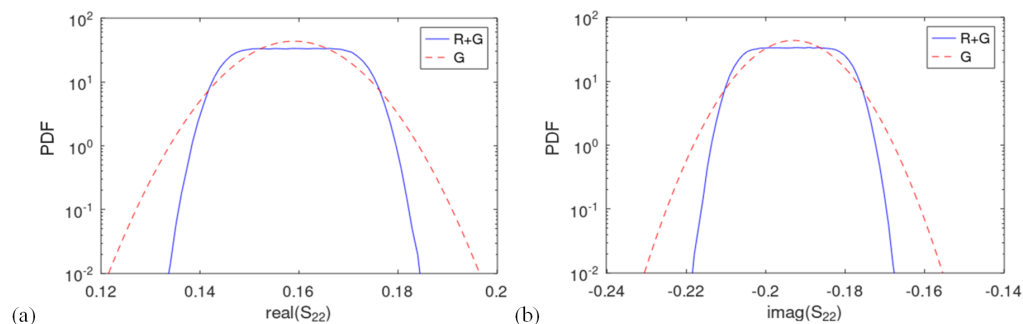


Fig. 3. PDF histogram of  $M$  factor determined by the MCM from (1) and (2) at: (a) 1, 40, 50 GHz, (b) 1 GHz.

out that the reason for the discrepancy between the GUM and the MCM is the nonlinearity of the measurement function. Using first order sensitivity terms (a standard commonly used in measurement procedures), we obtained a large discrepancy of results between the GUM and the MCM, similar to that in the publication [9]. However, when second-order expansion coefficients are used, the results of GUM and MCM almost overlap, as shown in Tables 1 and 2. This demonstrates the superiority of the MCM, which in this case, can be used to validate the GUM, as a reference method.

Table 1. Lower and upper bounds of the uncertainty interval of the  $M$  factor determined by the MCM and GUM for three measurement frequencies.

GUM order	1			2			3		
	1	40	50	1	40	50	1	40	50
$M_{lowGUM}$	0.9996	0.9597	1.0457	0.9989	0.9596	1.0459	0.9989	0.9596	1.0459
$M_{highGUM}$	1.0004	0.9856	1.1314	1.0010	0.9857	1.1317	1.0010	0.9857	1.1317
$M_{lowMCM}$	0.9988	0.9596	1.0460	0.9988	0.9596	1.0461	0.9988	0.9596	1.0462
$M_{highMCM}$	1.0011	0.9856	1.1317	1.0011	0.9856	1.1318	1.0011	0.9857	1.1319

Table 2. Length of the uncertainty interval of the  $M$  factor determined by the MCM and GUM for three measurement frequencies.

GUM order	1			2			3		
$f$ [GHz]	1	40	50	1	40	50	1	40	50
$dM_{\text{GUM}}$	0.0008	0.0259	0.0857	0.0021	0.0260	0.0857	0.0021	0.0260	0.0858
$dM_{\text{MCM}}$	0.0022	0.0261	0.0858	0.0022	0.0261	0.0857	0.0022	0.0260	0.0857
$\delta M_{\text{GUM-MCM}} [\%]$	-64.5	-0.4	-0.1	-7.4	-0.1	0.0	-7.4	-0.1	0.0

Table 1 shows the lower  $M_{\text{low}}$  and upper  $M_{\text{high}}$  limits of the uncertainty interval of the  $M$  factor determined by the GUM and the MCM at the three measurement frequencies  $f$ . Table 2 shows the length  $dM$  of the uncertainty interval of the  $M$  factor for the same parameters as in Table 1. The use of higher order sensitivity coefficients increases the accuracy of the analytical method. The results in the MCM method change slightly in each successive calculation due to the randomness of the process. Tables 1 and 2 highlight a case where there is a discrepancy in the results between the GUM and MCM. This occurs *e.g.* for the measurement frequency of 1 GHz.

In [9], optimization of the position of the uncertainty interval was applied, implemented in the MCM, to obtain the minimum length of this interval. Figure 4a shows the influence of the position of the uncertainty interval on its length for three measurement frequencies.

Figure 4b shows the same relationship but only at 1 GHz. Clearly the asymmetric shape of these curves is due to the asymmetric nature of the corresponding PDF uncertainties. For a normal distribution, the shape of the curve would be symmetric with respect to the expected value. In the case under consideration, this dependence is small as it is about 0.1% of the calculated uncertainty interval. We have also added this type of position optimisation to the developed software.

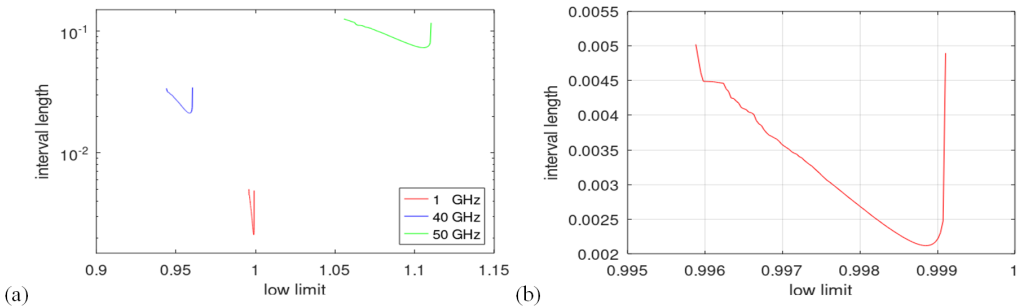


Fig. 4. Influence of the position of the uncertainty interval on its length at: a) 1, 40, 50 GHz, b) 1 GHz.

#### 4. Validation of the MCM software on example of determining type A uncertainty for RF power sensor calibration using a VNA and assuming a Gaussian PDF of input data

The type A uncertainty was determined for the calibration of the RF power sensor using a VNA. This uncertainty is due to the variation in the connection of the electronic calibration module (ECM), standard and DUT sensors to the VNA port during successive measurements. The calibration was carried out using the measuring set-up shown in Fig. 5.





$$M = \frac{\left| 1 - E_S \frac{\Gamma_D - E_D}{E_R - E_D E_S + \Gamma_D E_S} \right|^2}{\left| 1 - E_S \frac{\Gamma_S - E_D}{E_R - E_D E_S + \Gamma_S E_S} \right|^2}, \quad (7)$$

$$\Gamma_D = \frac{A_D}{R_D}, \quad (8)$$

$$\Gamma_S = \frac{A_S}{R_S}, \quad (9)$$

the calibration factor  $K_{D\_GUM}$  was determined with the GUM for all frequencies  $f_{1...k}$  by calculating the mean value  $mK_D$ , the standard deviation  $sK_D$ , followed by the lower and upper limits of the extended uncertainty interval at  $p = 95\%$ .

For comparison, using the MCM, the lower and upper limits of  $K_{D\_MCM}$  of the uncertainty interval were determined by generating samples of length  $N = 10^6$  with a normal PDF for all input quantities  $X_i$  and frequencies  $f_{1...k}$ .

The results shown in Fig. 6a indicate that the uncertainty of  $K_D$  determined with the MCM significantly exceeds those determined with the GUM. There were also discrepancies in the determination of the uncertainty of the mismatch factor  $M$  (Fig. 6b) which are weakly visible in Fig. 6b due to the more than 10 times smaller uncertainty of the mismatch factor  $M$  compared to the uncertainty of the calibration factor  $K_D$ .

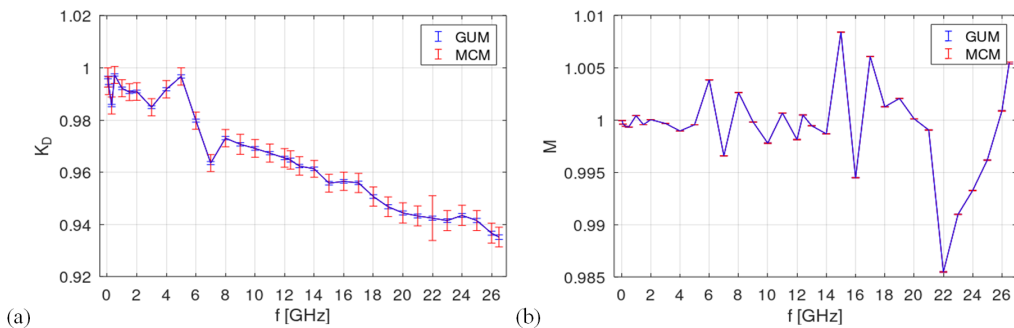


Fig. 6. (a) Calibration factors  $K_D$  (4) and (b) mismatch factors  $M$  (8), respectively, versus frequency, with marked uncertainty intervals, without including correlations between the input values.

As it is crucial to compare the uncertainties obtained by the MCM and GUM methods shown in Fig. 6, the corresponding relative differences in results between the MCM and GUM were calculated for the lower and upper limits of the uncertainty interval, expressed as a percentage of the values calculated by the GUM method, as shown in Fig. 7. Similar to graphs in Fig. 7, containing relative differences between modified MCM calculations and the GUM are presented later in this paper, because they are weakly visible in the format presented in Fig. 6.

A more detailed analysis of the input measurement data  $X_i$  was carried out to explain these discrepancies. Discrepancies occurring in the determination of uncertainty, particularly evident in Fig. 7a, can be eliminated by taking into account the existing correlations between respective  $X_i$ . It was confirmed by calculating and then implementing a covariance matrix in the prepared MCM' software, containing full information about all correlations of the input measurement data.

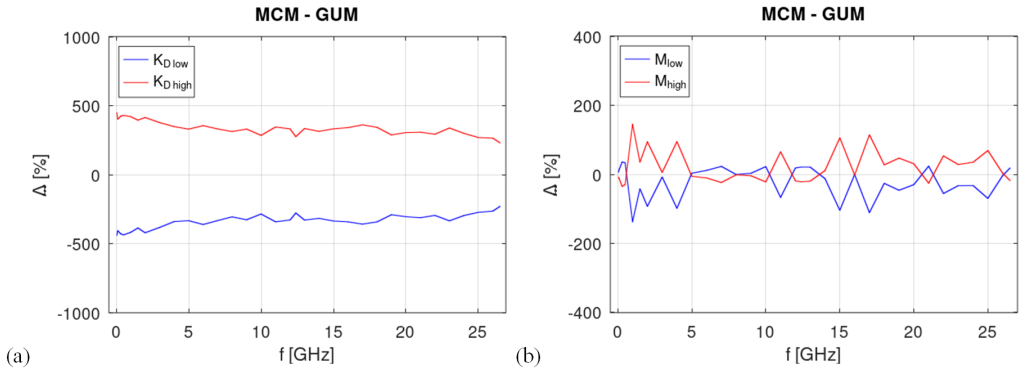


Fig. 7. Relative differences between the MCM and GUM methods calculated for the lower and upper limits of the uncertainty interval, expressed as a percentage of the values calculated by the GUM, for respectively: (a) calibration factors  $K_D$  (4) and (b) mismatch factors  $M$  (7).

The correlation between input quantities was checked by finding the number of instances of statistically significant ( $\alpha < 0.005$ ) correlation using the Octave corrcoef function [18] at frequencies  $f_{1...k}$ . The analysis showed, e.g., a strong correlation (in the range from 0.99 to 0.999) of the measured values  $P_D$  and  $|R_D|^2$ . With such a strong correlation, measurement errors are largely eliminated in this case, which should be taken into account in the MCM software by including the covariance matrix in the calculation of the uncertainty interval.  $P_D$  and  $|R_D|^2$  are measured simultaneously and in the following calculations have been treated here as single quantity  $p_D$  using (5). The same applies to  $P_S$  and  $|R_S|^2$  which leads to  $p_S$  (6). It follows from the above that building the MCM simulation model containing complex quantities requires a good knowledge of the specifics of VNA measurements. For example, treating  $A_S$  and  $R_S$  as independent complex quantities separately and generating independent random samples for them leads to completely erroneous results, as these quantities are measured simultaneously; consequently, instead of  $A_S$  and  $R_S$ , the reflection coefficient  $\Gamma_D$  (8) should be used in simulations, the same applies to  $A_S$ ,  $R_S$  and to the coefficient  $\Gamma_S$  (9).

Finally, we used  $B = 12$  input quantities  $X'_{1...B} = \{p_D, p_S, \text{Re}\Gamma_D, \text{Im}\Gamma_D, \text{Re}\Gamma_S, \text{Im}\Gamma_S, \text{Re}E_R, \text{Im}E_R, \text{Re}E_S, \text{Im}E_S, \text{Re}E_D, \text{Im}E_D\}$  in the measurement function (4). For each  $f_{1...k}$  correlation matrix of size  $B \times B$  was calculated using the Cholesky factorization with the Octave chol function [18], and the correlation degree between input quantities was checked by finding the number of instances of statistically significant ( $\alpha < 0.005$ ) correlation using the Octave corrcoef function [18], which is shown in Table 3 as numbers. When 10 or more instances occurred (from the maximal number  $k = 33$ ), the case is highlighted in yellow.

For all input quantities  $X_i'$ , a corresponding  $12 \times 12$  covariance matrix and random normal samples of length  $N = 10^6$  were generated. The simulation results  $K_{D\_MCM'}$  of the model prepared in this way are shown in Fig. 8.

Relative differences in results between MCM' and the GUM were calculated, taking into account the correlations between the input quantities, for the lower and upper limits of the uncertainty interval, expressed as a percentage of the values calculated by the GUM method. Figure 8a shows the results for calibration factors  $K_D$  and Fig. 8b for the mismatch factors  $M$ , respectively. When the correlation between the measured input data is taken into account, a good convergence between the results obtained by the MCM' and GUM methods is apparent. Compared

Table 3. Number of instances of statistically significant ( $\alpha < 0.005$ ) correlation between input quantities at  $k$  measurement frequencies. Highly correlated quantities highlighted in yellow.

	$p_S$	$Re\Gamma_D$	$Im\Gamma_D$	$Re\Gamma_S$	$Im\Gamma_S$	$ReE_D$	$ImE_D$	$ReE_S$	$ImE_S$	$ReE_R$	$ImE_R$
$p_D$	33	17	11	7	9	6	7	6	5	9	9
$p_S$		19	12	8	9	6	7	4	4	11	10
$Re\Gamma_D$			11	13	9	8	11	8	7	10	5
$Im\Gamma_D$				8	7	7	6	1	6	5	5
$Re\Gamma_S$					11	6	5	4	3	11	8
$Im\Gamma_S$						5	7	1	7	10	7
$ReE_D$							12	11	18	10	5
$ImE_D$								15	19	12	8
$ReE_S$									11	10	5
$ImE_S$										10	9
$ReE_R$											15

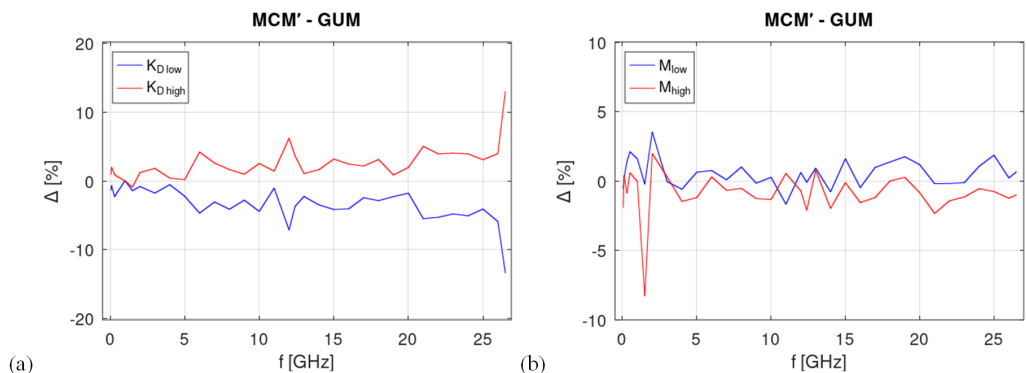


Fig. 8. Relative differences between the MCM' and GUM methods calculated for the lower and upper limits of the uncertainty interval, expressed as a percentage of the values calculated by the GUM, for respectively: (a) calibration factors  $K_D$  (4) and (b) mismatch factors  $M$  (7).

to the results shown in Fig. 7, the discrepancy between the results obtained by the MCM and GUM has decreased by more than 50 times.

### 5. Non-Gaussian PDF of measured input data

After implementing the effect of correlation on measurement uncertainty in MCM', the next task was to determine the PDF of the measured input quantities  $X'_i$  and their possible effect on the  $K_D$  uncertainty interval. The ability to take into account the shape of the PDF is the main advantage of the MCM that distinguishes it from the GUM. Using the GUM, it was *a priori* assumed that the PDF of  $K_D$  uncertainty has the shape of Student's t-distribution with the number of degrees of freedom  $\nu = n - 1$ . In Fig. 9, an example shape of PDF  $K_D$  at 20 GHz is shown using the MCM' approach. The PDF of  $K_D$  determined by the MCM is close to the real one,

provided that the PDFs of the measured input quantities are Student's t PDFs. The graph also shows the limits of the uncertainty interval at  $p = 95\%$ . The shapes and uncertainty intervals are nearly identical. We then tried to find the true shapes of the input quantities.

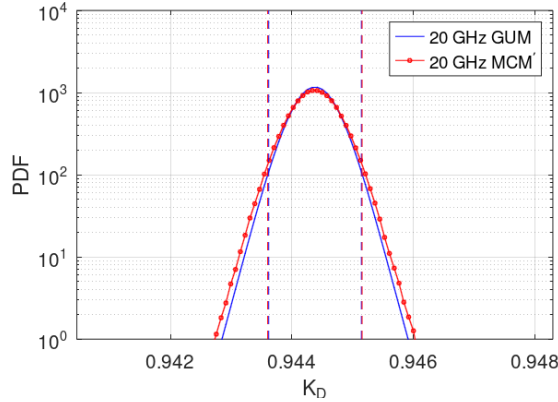


Fig. 9. Uncertainty limits and the PDF of calibration factor  $K_D$  at 20.0 GHz.

Determining the PDF shape requires multiple measurements. Skewness and kurtosis are parameters that determine the PDF shape of measured quantities  $X'_i$ . They are completely independent of the mean value and the standard deviation of the measured quantities.

First, the measurement results of all input quantities  $X'_i$  were analysed as a function of frequency. They are mostly monotonic and change slowly as a function of frequency, as illustrated in Fig. 10 for  $p_S$  where the graphs (a)–(d) show, respectively, the dependence of mean, standard deviation, skewness, and kurtosis on the frequency.

Figure 10e shows a plot with skewness  $\times$  kurtosis coordinates, where each point shows the skewness and kurtosis obtained for a given frequency. Individual PDF distributions (*e.g.*, normal, chi, chi2, gamma, beta, *etc.*) in the skewness  $\times$  kurtosis coordinates take the form of a family of curves. For example, a point with coordinates (0, 0) represents the shape of normal distribution  $N(0, 1)$ .

Unfortunately, the resulting estimates are biased for small sample size non-normal PDFs [19], so the following analysis abandons PDF formation based on sample location in the skewness  $\times$  kurtosis coordinates. Instead, we use a histogram of standardized values to approximate all PDFs of measured input quantities  $X'_i$  shown in Fig. 10f. For  $n$  consecutive measurements of input quantities  $X'_i$ , the resulting random variable was standardized by subtracting its expected value and dividing the difference by its standard deviation:

$$\text{standard}(X'_i) = \frac{X'_i - mX'_i}{sX'_i}. \tag{10}$$

The samples standardized in this way retain PDF shape information expressed in its skewness and kurtosis.

The number of results of independent measurements of the same quantity should be large enough to determine its PDF histogram. In statistics, it is usually assumed that the optimal number of histogram bins is 30. Then the shape of the histogram is sufficiently smoothed and at the same time accurately reflects the PDF of the measured value. In order for the bins of the

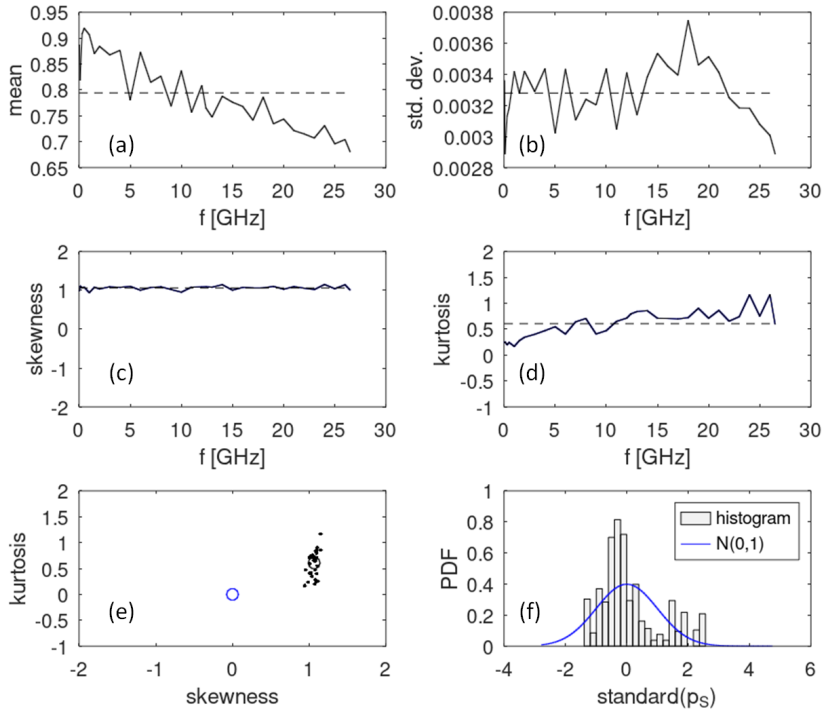


Fig. 10. Statistical properties of  $p_S$  versus frequency.

histogram to be sufficiently filled, on average, they should contain at least 20 results, hence the sample analysed should contain about  $30 \times 20 = 600$  results. This number is very difficult to achieve in practice. A single measurement of the DUT at 33 frequencies, requires at least 30 minutes. In total, it would take 200 man-hours to make 600 measurements. A separate issue is the degradation of measuring connectors when they are repeatedly mated. To overcome this problem, we included standardized values of measurements taken at  $k = 33$  frequencies as reliable ones for determining the histogram. In consequence,  $n = 17$  measurements at  $k$  frequencies were used for statistical analyses, yielding a total of  $n \times k = 561$  results for each input quantity  $X'_i$ . An example of a cumulative histogram of  $n \times k$  measurements determined for standardized  $p_S$  is shown in Fig. 10f.

Based on the calculated  $HX'_i$  histograms (e.g. shown in Fig. 10f), standardized vectors  $VX'_i$  of random values with corresponding PDF shapes and length  $N = 10^6$  were assembled in two steps: (i) initially, a vector was created piece-by-piece from the generated random vectors of length, mean value and standard deviation taken from successive bins of the  $HX'_i$  histogram, (ii) then the position of the elements of  $VX'_i$  vector was randomly changed. This procedure is explained in Figs. 11 and 12, for two cases – using a uniform and normal distribution, respectively, and only 10 histogram bins. The uniform distribution is used to show the idea of generating vectors since it gives worse results than the normal distribution.

The vectors from Fig. 12 sorted in ascending order, as shown in Fig. 13, are the inverse of the *cumulative distribution function* (CDF). A comparison of the shape of the curves in Figs. 13a and 13b shows that a uniform distribution does not produce a smooth final distribution, unlike a normal distribution.

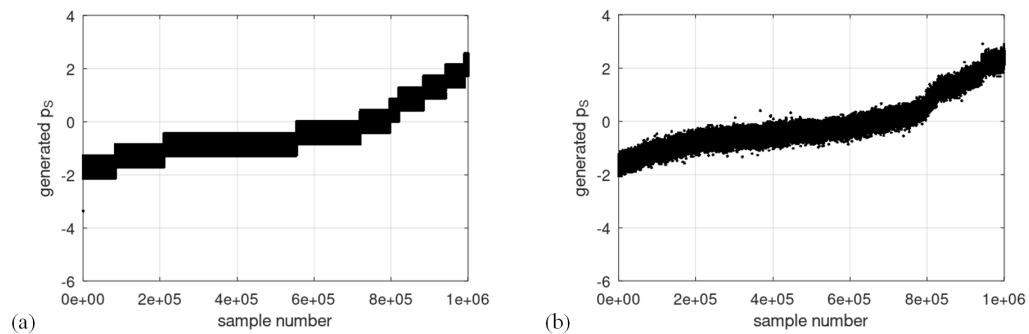


Fig. 11. Vectors created piece-by-piece from random vectors generated according to successive bins of the histogram standard ( $p_S$ ) (10), using: (a) uniform PDF, (b) normal PDF.

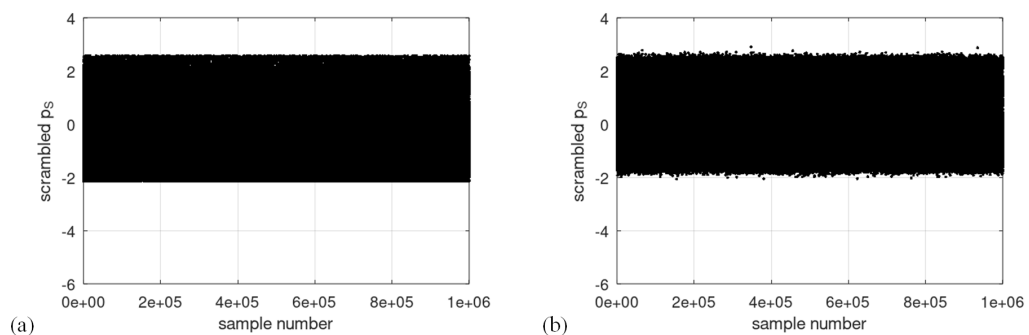


Fig. 12. Vectors from Fig. 11, after randomly changing the position of their elements.

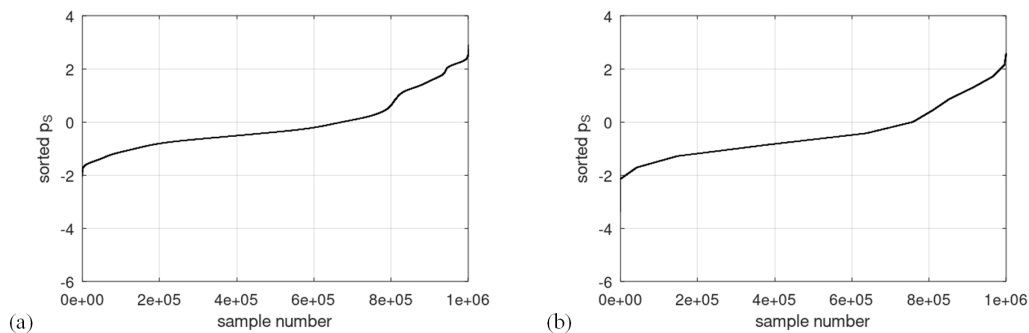


Fig. 13. Ascending sorted vectors from Fig. 12 for: (a) uniform PDF, (b) normal PDF.

The standardized  $VX'_i$  vectors (e.g., generated for  $p_S$  as shown in Fig. 14) were then reversely unstandardized according to:

$$X''_i = \text{standard}(VX'_i) \cdot sX'_i + mX'_i \quad (11)$$

and consequently used for the MCM'' simulation of measurement uncertainty, taking into account, in addition to correlations, the PDF shapes of the input quantities  $X_i''$ , which are arguments of the measurement function  $Y$ .

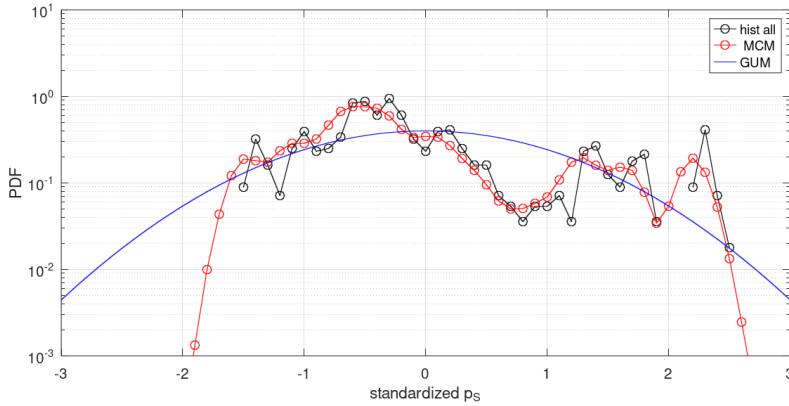


Fig. 14. Example histogram of the generated random sample  $p_s$ .

Finally, after using equations (4)–(9), MCM'' PDFs of the  $K_D$  were obtained (Fig. 15).

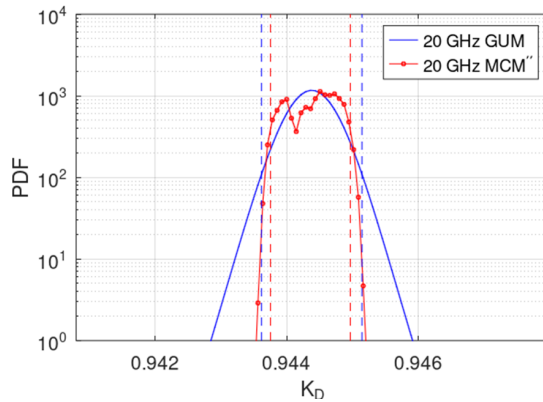


Fig. 15. PDF of the  $K_D$  calibration factor for 20.0 GHz determined with the preservation of information about the shape of the PDF of the measured input quantities.

Relative differences in results between MCM'' and the GUM were calculated, taking into account the correlations and PDF shapes of input quantities, for the lower and upper limits of the uncertainty interval, expressed as a percentage of the values calculated by the GUM method. Figure 16a shows the results for calibration factors  $K_D$  and Fig. 16b for the mismatch factors  $M$ , respectively.

As it is crucial to directly compare results of the MCM'' and MCM' methods, shown in Fig. 8 and Fig. 16, respectively, the differences between them are obtained by using equations (12), (13) and (14):

$$\Delta Y_{\text{high}} = Y''_{\text{high}} - Y'_{\text{high}}, \quad (12)$$

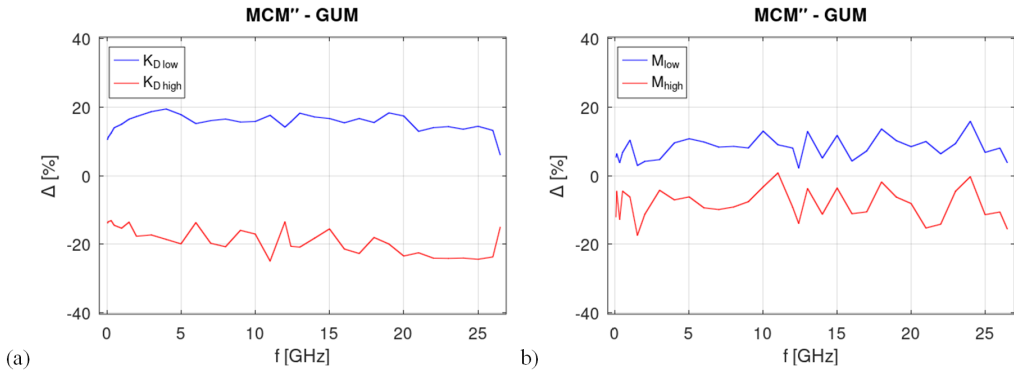


Fig. 16. Relative differences between the MCM'' and GUM methods calculated for the lower and upper limits of the uncertainty interval, expressed as a percentage of the values calculated by the GUM, for respectively: (a) calibration factors  $K_D$  (4) and (b) mismatch factors  $M$  (7).

$$\Delta Y_{low} = Y''_{low} - Y'_{low}, \tag{13}$$

$$\Delta Y_{length} = \Delta Y''_{low} - \Delta Y'_{high}, \tag{14}$$

where  $\Delta Y_{high}$ ,  $\Delta Y_{low}$ , and  $\Delta Y_{length}$  are the differences between the upper and lower limits of the uncertainty interval and its length, respectively, obtained for calculations according to MCM'' and MCM'.  $Y$  denotes the output quantity and is replaced by the uncertainty of the calibration factor  $K_D$  (4) in Fig. 17a and the uncertainty of the mismatch factor  $M$  (7) in Fig. 17b, respectively.

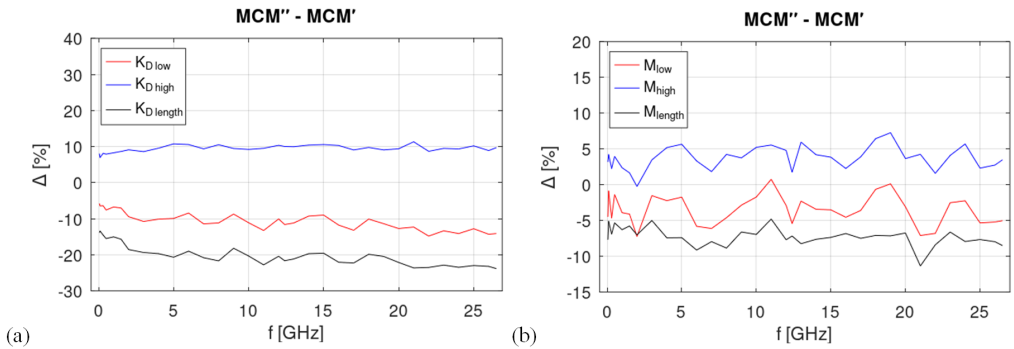


Fig. 17. Differences between the MCM'' and MCM' methods, resulting from the use of the PDF shape information in the determination of the measurement uncertainty parameters, as defined by equations (12)–(14), for the uncertainty of: (a) calibration factors  $K_D$  (4) and (b) factors  $M$  (7), respectively.

The results of this comparison are homogeneous regardless of frequency, as shown in Fig. 17. The differences between the MCM'' and MCM' approaches, resulting from the use of the PDF shape of the input quantities in the latter approach to determine the measurement uncertainty, show that the type A uncertainty of calibration factor  $K_D$  (Fig. 17a) calculated with the MCM' approach gives a larger value than with the MCM'' approach, implying that the inclusion of the PDF shape allows the calibration accuracy to be improved.

The same is true for the type A uncertainty of mismatch factor  $M$  (Fig. 17b), except that the differences between the MCM'' and MCM' approaches are about 2 times smaller.



## 6. Conclusions

It is worth noting that the METAS UncLib software using the GUM approach is, in most cases, faster and more accurate than the software developed at the LMEEiO, as sufficiently long vectors have to be used to ensure adequate accuracy in the MCM'' approach. However, in some cases where we are dealing with specific PDF shapes of the measured values, our software using MCM'' is more accurate and can be used to validate the designated interval or area of measurement uncertainty.

The lower uncertainty of the mismatch factor  $M$  in comparison to the uncertainty of the calibration factor  $K_D$  confirms the advantages of calibrating RF power sensors by direct measurement using the VNA.

Reliable PDF determination of the measured quantities for a small sample of repeated independent measurements is a challenging task. To achieve this, we have developed a special method that uses the histograms of standardized input quantities at all measurement frequencies to increase the total number of results which were needed to create the PDF histogram of measured input quantities.

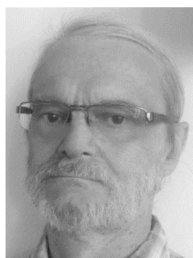
The results presented in Section 5 showed that in our particular case, the influence of the PDF shape on  $K_D$  uncertainty limits is substantial.

It is worth noting that the  $K_D$  calibration coefficients in the 3–7 GHz range (see Fig. 7a) show ripples due to the nonlinearity of the VNA source, which can be observed by lowering the output power of the VNA – these ripples then decrease significantly. As this type of error is systematic, it does not significantly affect the determination of type A uncertainty and is omitted here. However, it is advisable to eliminate this imperfection in order to increase the accuracy of the calibration.

## References

- [1] Joint Committee for Guides in Metrology. (2008). *Evaluation of measurement data – Guide to the expression of uncertainty in measurement* (JCGM 100:2008). [http://www.bipm.org/utis/common/documents/jcgm/JCGM\\_100\\_2008\\_E.pdf](http://www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf)
- [2] EA Laboratory Committee. (2013). *EA-4/02. Evaluation of the uncertainty of measurement in calibration*.
- [3] Joint Committee for Guides in Metrology. (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* (JCGM 101:2008). [https://www.bipm.org/documents/20126/2071204/JCGM\\_101\\_2008\\_E.pdf](https://www.bipm.org/documents/20126/2071204/JCGM_101_2008_E.pdf)
- [4] Coral, R., Flesch, C. A., Penz, C. A., Roisenberg, M., & Pacheco, A. L. (2016). A Monte Carlo-based method for assessing the measurement uncertainty in the training and use of artificial neural networks. *Metrology and Measurement Systems*, 23(2), 281–294. <https://doi.org/10.1515/mms-2016-0015>
- [5] Arkani, M. (2021). Observation probability estimation of dead-time models using Monte Carlo simulations. *Metrology and Measurement Systems*, 28(2), 383–395. <https://doi.org/10.24425/mms.2021.136614>
- [6] Jaworski, M., et al. (2021). Research and development on calibration of measuring instruments in the field of electrical quantities and optoelectronics. *NIT Report*, Warsaw.
- [7] Jaworski, M. (2022, May 5). Estimating the uncertainty of measurement using the Monte Carlo method, *NIT Seminar*, Warsaw.

- [8] JCGM 102 (2011). Evaluation of measurement data – Supplement 2 to the “Guide to the expression of uncertainty in measurement” – Extension to any number of output quantities.
- [9] Wu, T. Y., & Chua, S. W. (2009, May 5-7). Evaluation of mismatch uncertainty in microwave power sensor calibration using Monte Carlo method. *Proceedings of the IEEE International Instrumentation and Measurement Technology Conference (I2MTC 2009)*, Singapore.
- [10] Zeier, M., Hoffmann, J., & Wollensack, M. (2012). Metas. UncLib – A measurement uncertainty calculator for advanced problems. *Metrologia*, 49(6), 809–815. <https://doi.org/10.1088/0026-1394/49/6/809>
- [11] Wollensack, M. (2017). *Introduction to METAS UncLib*. Federal Institute of Metrology METAS
- [12] Hall, B. D. (2022). The GUM tree calculator: A python package for measurement modelling and data processing with automatic evaluation of uncertainty. *Metrology*, 2(1), 128–149. <https://doi.org/10.3390/metrology2010009>
- [13] Wollensack, M. (2022). METAS UncLib MATLAB – User Reference V2.6.0.
- [14] Wollensack, M. (2022). METAS UncLib Python – User Reference V2.6.0.
- [15] Juroshek, J. R. (1997). A direct calibration method for measuring equivalent source mismatch. *Microwave Journal*, 40(10), 106–113.
- [16] Dobbert, M., & Gorin, J. (2011, August). Revisiting mismatch uncertainty with the Rayleigh distribution. In *NCSL International Workshop and Symposium* (pp. 8–9). Agilent Company.
- [17] Szatkowski, J. (2021). Developing RF Power Sensor Calibration Station in Direct Comparison Transfer System using Vector Network Analyzer. *Journal of Telecommunications and Information Technology*. (3), 18–22. <https://doi.org/10.26636/jtit.2021.155021>
- [18] Correlation and Regression Analysis (GNU Octave (version 6.3.0)). (n.d.). <https://docs.octave.org/v6.3.0/Correlation-and-Regression-Analysis.html>
- [19] Krishnamoorthy, K. (2016). *Handbook of Statistical Distributions with Applications*. CRC Press.



**Marek Jaworski** received his Ph.D. degree from the National Institute of Telecommunications (NIT), Warsaw, Poland, in 2001. He is currently an Assistant Professor at the NIT Central Chamber for Telecommunication Metrology. He has been with the NIT since 1982, working on modelling and design of optical fiber transmission systems, measurement methods, and test equipment for optical networks. His current research interests include numerical simulations, advanced modulation formats, and nonlinear photonics.



**Tomasz Kossek** received his M.Sc. in Optoelectronics and his Ph.D. degree from Warsaw University of Technology, Warsaw, Poland, in 1996 and 2002, respectively. He is currently an Assistant Professor at the National Institute of Telecommunications Poland, Warsaw. His current research interests include optoelectronic measurements and their calibration, laser physics, and optical communication.



**Jarosław Szatkowski** received his B.Sc. and M.Sc. in Electronics from Warsaw University of Technology, Poland, in 2015 and 2017 respectively. Since 2017 he has been with the Laboratory of Electrical, Electronic and Optoelectronic Metrology of the National Institute of Telecommunications where he is involved in calibration of RF metrology instrumentation. His research interests focus mainly on microwave metrology such as vector network analysis and RF power measurements.