

Interlaboratory Comparison of RF Power Measurements Made by Automatic Power Measurement Software

Downloaded from: https://research.chalmers.se, 2026-11-24 09:51 UTC

Citation for the original published paper (version of record):

Cetinkaya, A., Dogan, A., Danaci, E. et al (2026). Interlaboratory Comparison of RF Power Measurements Made by Automatic Power Measurement Software. Metrology, 6(3). http://dx.doi.org/10.3390/metrology6030047

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

metrology

MDPI

Article

Interlaboratory Comparison of RF Power Measurements Made by Automatic Power Measurement Software

Anil Cetinkaya [1,2], Aliye Kartal Dogan [1,*], Erkan Danaci [1], Martin Hudlicka [3], Jan Grajciar [3], Osman Sibonjic [4], Torsten Lippert [5], Søren Mortensen [5], Sean Prendergast [6], Łukasz Usydus [7], Marko Berginc [8], Emre Çetin [9], Muhammed Cagri Kaya [10] and Halit Oğuztüzün [11]

1 National Metrology Institute (UME), The Scientific and Technological Research Council of Türkiye (TÜBİTAK), 41470 Kocaeli, Türkiye; anil.cetinkaya@tubitak.gov.tr (A.C.); erkan.danaci@tubitak.gov.tr (E.D.)
2 Department of Computer Engineering, Faculty of Engineering and Natural Sciences, Iskenderun Technical University, 31200 Hatay, Türkiye
3 Czech Metrology Institute (CMI), Praha 10, CZ-10200 Prague, Czech Republic; martin.hudlicka@cmi.gov.cz (M.H.); jan.grajciar@cmi.gov.cz (J.G.)
4 Institute of Metrology of Bosnia and Herzegovina (IMBIH), 71000 Sarajevo, Bosnia and Herzegovina; osman.sibonjic@met.gov.ba
5 Trescal A/S, 8660 Silkeborg, Denmark; torsten.lippert@trescal.com (T.L.); soren.mortensen@trescal.com (S.M.)
6 National Standards Authority of Ireland (NSAI), D09 A0E4 Dublin, Ireland; sean.prendergast@nsai.ie
7 Central Office of Measures (GUM), 00-139 Warsaw, Poland; lukasz.usydus@gum.gov.pl
8 Slovenian Institute of Quality and Metrology (SIQ), SI-1000 Ljubljana, Slovenia; marko.berginc@siq.si
9 Spark Calibration Services Inc. (SPARK), 06530 Ankara, Türkiye; emre_cetin@sparkmeasure.com
10 Department of Computer Science and Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden; cagri.kaya@chalmers.se
11 Department of Computer Engineering, Middle East Technical University, 06800 Ankara, Türkiye; oguztuzn@ceng.metu.edu.tr
* Correspondence: aliye.dogan@tubitak.gov.tr

Abstract

An interlaboratory comparison of RF power measurements using automatic power measurement software was carried out among partners of the European Partnership on Metrology project titled "Development of RF and microwave metrology capability II" (RFMicrowave2). The absolute RF power output of a travelling signal generator was measured at six frequencies between 2 GHz and 18 GHz and at three nominal power levels of −10 dBm, 0 dBm, and 10 dBm. Comparison results were evaluated in terms of comparison reference values, degrees of equivalence, and normalised errors. Uncertainty evaluations were performed using both the Guide to the Expression of Uncertainty in Measurement and Monte Carlo simulation methods. The results demonstrate an acceptable level of agreement among participating laboratories and confirm the applicability of automatic power measurement software for interlaboratory RF power comparisons.

Keywords: RF power measurement; interlaboratory comparison; automatic measurement; uncertainty evaluation; Monte Carlo simulation; metrology

Check for updates

Academic Editor: Patrice Salzenstein

Received: 28 April 2026
Revised: 4 June 2026
Accepted: 11 June 2026
Published: 6 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The integrity of the global metrology system is based fundamentally on international comparisons to achieve measurement accuracy, consistency, and trust [1]. Mostly administered by International Committee for Weights and Measures (CIPM) Mutual Recognition Arrangement (MRA), these comparisons provide National Metrology Institutes (NMIs) with the opportunity to show the equivalence of their national standards and validate their

Calibration Measurement Capabilities (CMCs). Therefore, interlaboratory comparisons allow institutions to ensure that their measurements are traceable to the International System of Units (SI) and ensure international recognition by reducing technical barriers. For example, in the RF and microwave domain, international comparisons have been carried out for quantities such as RF power and complex reflection coefficients, establishing the equivalence of national standards across participating institutes [2,3].

Reliable measurement of radio-frequency (RF) and microwave (MW) power is a fundamental requirement for calibration, testing, communications, electromagnetic compatibility (EMC), and a wide range of industrial and regulatory applications [4]. In RF and MW metrology, power is one of the core quantities underpinning traceability chains for related measurements such as attenuation and sensor calibration, and its reliable determination directly affects the quality of higher-level measurement and calibration services [5,6].

RF power calibration is often time-consuming because measurements are repeated over wide frequency and power ranges and require careful management of sensor behaviour, mismatch effects, and instrumental stability. For this reason, automated measurement software offers clear practical advantages by reducing operator dependence, improving repeatability, and enabling more systematic handling of uncertainty budgets [7]. These aspects are particularly relevant for developing laboratories aiming to establish robust and reproducible services for industrial stakeholders.

The evaluation of measurement uncertainty is equally important. The Guide to the Expression of Uncertainty in Measurement (GUM) provides the internationally accepted framework for uncertainty evaluation [8]. The Monte Carlo simulation (MCS) method described in Supplement S1 to the GUM provides a distribution-based alternative that is especially useful when linearisation assumptions are weak or when the output distribution deviates from normality [9]. In RF power transfer measurements, uncertainty contributions associated with mismatch, transfer stability, sensor characteristics, and measurement system configuration can strongly influence the reported results, making rigorous and comparable uncertainty treatment essential [5,6].

At the European level, these needs have motivated coordinated capability-building efforts for RF and MW metrology. The European Partnership on Metrology (EPM) project 22RPT04, "Development of RF and microwave metrology capability II" (RFMicrowave2), aims to improve measurement and research capability in RF and MW metrology, particularly in less experienced NMIs and DIs, and to strengthen future cooperation among European institutes and stakeholders. The project addresses several strategically important areas, including RF and MW power, attenuation, reflection and transmission measurements, antenna parameters, and electromagnetic field measurements, together with a unified approach to measurement uncertainty evaluation [10].

Within this framework, the project focuses on advanced measurement techniques for RF power and power-related quantities. Its objectives include developing traceable capabilities for attenuation and reflection/transmission measurements while establishing a harmonised approach to evaluating their uncertainties. In particular, automatic RF power measurement and uncertainty calculation with different methods is addressed within the framework of the RFMicrowave2 project. The goals of the task are to establish automated RF power measurement capability in participating laboratories, incorporate both traditional law of propagation (GUM-based) and Monte Carlo simulation (MCS) based uncertainty evaluation approaches into the software environment, and validate the established capabilities through an interlaboratory comparison of direct RF power measurement.

Our previous research has been focused systematically on digitalisation and automation of metrological processes, providing the foundational components for the current study. Initially, the "Internet of Measurement Things" (IoMT) architecture was proposed [11],

establishing the communication and networking requirements for automated calibration in the Industry 4.0 era. Within this framework, the AutoRFPower software was specifically developed and introduced [12] as a specialised desktop application for automated RF power measurements, designed to ensure metrological rigour and flexibility; subsequent research [13] explored the potential of "Uncertainty Calculation as a Service" (UCaaS) through cloud-based microservices as a proof of concept, the present study focuses on the practical validation of the established AutoRFPower desktop tool. This paper demonstrates the software's performance in a large-scale interlaboratory comparison involving eight international partners, confirming its reliability as a standardised tool for high-precision RF power metrology.

This study, conducted within the framework of the 22RPT04 RFMicrowave2 project, presents an interlaboratory comparison of RF power measurements made by automatic power measurement software. The comparison was designed to assess the consistency of RF power measurements performed by different laboratories using a common travelling standard and to evaluate the applicability of automated measurement software in conjunction with both GUM-based and MCS-based uncertainty evaluation approaches. Eight laboratories participated in the comparison according to a technical protocol prepared by the National Metrology Institute (UME) of the Scientific and Technological Research Council of Türkiye (TÜBİTAK) and approved by all of the comparison participants.

Section 2 of this article presents a review of related work. Sections 3 and 4 describe the travelling standard used in the comparison and the measurement methodology, focusing on the automated workflow of the AutoRFPower software and the implementation of GUM-based and MCS-based uncertainty evaluations, respectively. The interlaboratory comparison results, including Comparison Reference Values (CRVs) and Degrees of Equivalence (DoEs), are presented in Section 5. Section 6 discusses the findings, the impact of automated tools on measurement harmony, and potential refinements. Finally, Section 7 concludes the paper with possible future directions.

2. Related Work

Evaluation of interlaboratory comparisons and attainment of measurement equivalence rely on established metrological principles and international guidelines. Salzenstein et al. [1] provide an overview of the global metrology infrastructure and emphasise the critical role of interlaboratory comparisons in validating CMCs. However, the metrology sector is undergoing a profound digital transformation in which the focus is shifting toward establishing robust e-infrastructures and digital quality frameworks. To this end, Neyezhmakov et al. [14] highlight that paving the way for digital transformation requires the development of metrological e-infrastructures and the digitalisation of calibration data to ensure seamless data exchange. Moreover, Verma [15] discusses the potential of cutting-edge technologies such as AI and Big Data in creating a digital quality infrastructure that enhances international trade and measurement reliability. In line with this trend, our work presents a specific case study utilising the AutoRFPower software to implement a digitalised and automated workflow for an interlaboratory comparison.

The mathematical model used to analyse the comparison data follows the procedures set forth by Cox [16]. In our study, we employed the weighted mean approach introduced in [16] when calculating CRVs. This process is automated using the AutoRFPower software. The used statistical framework is supported by other studies in the literature as well: Rukhin and Sedransk [17] provide a rigorous basis for using meta-analysis models in international key comparisons. By building these core stochastic ideas into the process, our automated method ensures that both the CRV and the related DoE come from a solid consensus value that reflects the different uncertainty levels of all participating laboratories.

Reliability also depends heavily on how measurement uncertainty is handled in RF power metrology and sensor comparison. For example, Koepke et al. [18] advocate the use of advanced model-based approaches, such as hierarchical Bayesian models and MCS, to account for "dark uncertainty". Their work provides the theoretical foundation for the MCS methods that are integrated into our automated software environment. While they focus on the statistical rigour of mathematical models, our study bridges the gap between theory and practice by incorporating these models (both GUM-based and MCS-based) into an automated measurement framework. Thus, our method reduces operator dependence and improves the repeatability of results across different laboratory settings.

3. Travelling Standard and Measurement Conditions

The travelling standard, supplied by Trescal A/S, was an Agilent model 83640A signal generator with serial number 3250A00717. The signal generator operates in the frequency range from 10 MHz to 40 GHz and provides output power levels between −110 dBm and 15 dBm.

A total of eight laboratories participated in the interlaboratory comparison. Measurements were carried out under controlled environmental conditions, with temperatures maintained around (23 ± 2) °C and relative humidity around (45 ± 15) %rh. The participating laboratories and their environmental conditions are listed in Table 1.

Table 1. Laboratories participating in the interlaboratory comparison of RF power.

Country	Institute	Temperature (°C)	Humidity (%rh)
Bosnia and Herzegovina	Institute of Metrology of Bosnia and Herzegovina (IMBIH)	23 ± 2	50 ± 20
Czech Republic	Czech Metrology Institute (CMI)	23 ± 2	40 ± 10
Denmark	Trescal A/S	23 ± 1	45 ± 15
Ireland	National Standards Authority of Ireland (NSAI)	23 ± 2	45 ± 20
Poland	Central Office of Measures (GUM)	23.6	45.7
Slovenia	Slovenian Institute of Quality and Metrology (SIQ)	23 ± 2	50 ± 20
Türkiye	Spark Calibration Services Inc.	23 ± 2	50 ± 20
Türkiye	TÜBİTAK National Metrology Institute (TÜBİTAK UME)	23 ± 1	45 ± 15

Measurements were taken at six frequencies, namely, 2 GHz, 5 GHz, 8 GHz, 12 GHz, 15 GHz, and 18 GHz, and at nominal power levels of 10 dBm, 0 dBm, and −10 dBm. Each participating institute carried out the comparison measurements using its own measurement system, including a power meter, a power sensor and the proper adaptors.

Each participant measured the absolute RF power output of the travelling standard using the AutoRFPower software, which was used to manage the complete measurement workflow, including data acquisition and uncertainty evaluation. The software incorporates both Standing Wave Ratio (SWR) values provided by the manufacturer and complex reflection coefficient (RCC) values measured by selected laboratories.

For one participating laboratory, incompatibility issues prevented the use of AutoRFPower for direct data acquisition. In this case, measurement data were recorded using an alternative data logging approach. However, the recorded data were subsequently processed using the AutoRFPower software, particularly its uncertainty evaluation modules, in the same manner as for the other participants. Therefore, consistency in data processing and uncertainty evaluation was maintained across all laboratories.

The SWR values represent a scalar description of the mismatch behaviour between the source and the power sensor and are typically provided by the manufacturer as part of the instrument specifications. In contrast, the RCC is a complex quantity describing both the magnitude and phase of the reflected signal, usually obtained from vector network analyser (VNA) measurements.

Prior to the comparison, the vector RCC of the adaptor connected to the signal generator to be used in the comparison as a travelling standard was measured by TÜBİTAK UME with the help of Spark Calibration Services Inc., Ankara, Türkiye (Spark) using a VNA-based vector reflection coefficient measurement technique [19]. The measured vector RCC values were shared with the comparison participants.

Using the complex reflection coefficient allows a more complete vector description of the mismatch interaction between the adaptor and the power sensor, which can improve the accuracy of mismatch corrections and uncertainty evaluation [20]. For this reason, both SWR-based and RCC-based approaches were considered in the software to assess their influence on the calculated RF power and associated measurement uncertainties.

4. Measurement Methodology and Evaluation

The circulation of the travelling standard was organised to follow a ring-style path. The circulation began in June 2024, and comparison measurements were completed in January 2025. Each laboratory had a week to carry out the measurements.

Each participant institute used the AutoRFPower software to measure the absolute RF power of the travelling standard against its own traceable reference standards, namely, power sensors and power meters. At each test point, the measurement was completed by taking at least 10 successive readings. Further details of this comparison are provided in the associated technical report [21].

Uncertainty evaluations were performed according to two different approaches included in the AutoRFPower software:

- the Guide to the Expression of Uncertainty in Measurement (GUM);
- Monte Carlo Simulation (MCS).

The MCS-based evaluation is implemented in the developed software following Supplement S1 to the Joint Committee for Guides in Metrology (JCGM) 101 [9]. Its application to RF power measurements was described and validated against a commercial Monte Carlo tool in [22]. In contrast to the law of propagation approach, in which every contribution is reduced to the equivalent standard uncertainty of a normal distribution, the MCS module assigns a probability density function (PDF) to each input quantity of the measurement model and propagates these PDFs through the model by repeated random sampling, so that the actual shape of each input distribution is retained in the output distribution [22].

The measurement model relates the power indicated by the sensor to the absolute output power of the travelling standard through the calibration factor of the standard power sensor, the (complex) forward transmission coefficient of the adaptor/attenuator where used, and the impedance-mismatch factor M between the source and the sensor [22]. For each measurement point, the input quantities were assigned PDFs according to their nature. The repeatability of the power readings, the calibration factor of the standard power sensor, the forward transmission coefficient of the adaptor/attenuator, and the magnitude and phase of the reflection coefficients were modelled with Gaussian (normal) PDFs, whereas the power-meter accuracy and the power-meter resolution were modelled with rectangular (uniform) PDFs. Each PDF is characterised by a location parameter (the best estimate of the quantity) and a scale parameter (its standard uncertainty); $M_{\mathrm{MCS}} = 2 \times 10^5$ samples were then drawn from each parameterised PDF and propagated through the

model, and the mean and standard deviation of the resulting output distribution were taken as the estimated power and its standard uncertainty, respectively.

The treatment of the mismatch contribution distinguishes the two cases considered in this work and is consistent with the mismatch modelling described by Danaci et al. [22]. In the RCC case, the complex reflection coefficients are available; the mismatch factor is then evaluated from the magnitudes and phases of Γ (each modelled with a Gaussian PDF), so that the resulting mismatch contribution is approximately normal. In the SWR case, only the magnitude of the reflection coefficient is known (derived from the manufacturer's standing-wave ratio); the mismatch factor is bounded by its extreme values and is therefore modelled with a U-shaped (arcsine) PDF [23]. This difference in the mismatch PDF is one of the main reasons for the slightly different uncertainties obtained for the RCC and SWR cases, particularly at higher frequencies, where mismatch effects become more pronounced.

The comparison results and combined uncertainties were used to determine the Comparison Reference Values (x_{CRV}).

The degree of equivalence of each participant and the expanded uncertainty of the degree of equivalence were calculated [16].

The E_n value is stated in ISO/IEC 17043 "Conformity assessment—General requirements for proficiency testing" Standard [24]. The participant results were regarded as satisfactory if $\mid E_n \mid < 1$.

The x_{CRV} was calculated for each measurement point using the weighted mean with normalised weights. The x_{CRV} is given by

$$x_{\mathrm{CRV}} = \frac{\sum_{i=1}^{N} \frac{x_i}{u_{x_i}^2}}{\sum_{i=1}^{N} \frac{1}{u_{x_i}^2}} \tag{1}$$

and the uncertainty of the Comparison Reference Values (x_{CRV}) was calculated as

$$u(x_{\mathrm{CRV}}) = \sqrt{\frac{1}{\sum_{i=1}^{N} \frac{1}{u_{x_i}^2}}} \tag{2}$$

The expanded uncertainty of the Comparison Reference Values (x_{CRV}) was calculated as

$$U(x_{\mathrm{CRV}}) = 2\, u(x_{\mathrm{CRV}}) \tag{3}$$

5. Comparison Results

The x_{CRV} values were calculated for both RCC and SWR cases using GUM and MCS methods at three nominal power levels (−10 dBm, 0 dBm, and 10 dBm). The Comparison Reference Values (CRVs) were determined as weighted means of the participant results, where the weights were derived from the reported uncertainties of each laboratory. The results at 0 dBm, −10 dBm, and 10 dBm are presented in Tables 2, 3, and 4, respectively.

Degrees of equivalence (DoEs) with expanded uncertainties at 0 dBm for the RCC and SWR cases based on the GUM method are presented in Figure 1 and Figure 2, respectively, while the corresponding results based on the MCS method are shown in Figures 3 and 4.

Similarly, the DoEs at −10 dBm for the RCC and SWR cases based on the GUM method are presented in Figure 5 and Figure 6, respectively, while the corresponding results based on the MCS method are shown in Figures 7 and 8.

Likewise, the DoEs at 10 dBm for the RCC and SWR cases based on the GUM method are presented in Figure 9 and Figure 10, respectively, while the corresponding results based on the MCS method are shown in Figures 11 and 12.

Table 2. x_{CRV} and expanded uncertainties $U(x_{CRV})$ (k = 2) at 0 dBm.

Freq. (GHz)	GUM-Based				MCS-Based			
	RCC		SWR		RCC		SWR	
	x	U	x	U	x	U	x	U
2	0.041	0.027	0.006	0.034	0.005	0.029	−0.003	0.033
5	0.022	0.032	0.002	0.039	−0.019	0.037	−0.026	0.044
8	0.068	0.032	0.031	0.040	0.035	0.037	0.034	0.043
12	0.027	0.036	0.003	0.046	0.012	0.041	0.017	0.050
15	0.014	0.040	0.010	0.054	0.005	0.048	−0.003	0.060
18	−0.051	0.042	−0.028	0.056	−0.027	0.049	−0.024	0.061

Table 3. x_{CRV} and expanded uncertainties $U(x_{CRV})$ (k = 2) at −10 dBm.

Freq. (GHz)	GUM-Based				MCS-Based			
	RCC		SWR		RCC		SWR	
	x	U	x	U	x	U	x	U
2	−9.896	0.029	−9.950	0.036	−9.929	0.032	−9.919	0.036
5	−9.917	0.032	−9.939	0.040	−9.949	0.038	−9.941	0.045
8	−9.901	0.033	−9.941	0.043	−9.921	0.039	−9.909	0.044
12	−9.920	0.037	−9.942	0.049	−9.919	0.043	−9.893	0.051
15	−9.949	0.042	−9.958	0.059	−9.942	0.051	−9.928	0.060
18	−9.984	0.043	−9.980	0.060	−9.956	0.050	−9.925	0.062

Table 4. x_{CRV} and expanded uncertainties $U(x_{CRV})$ (k = 2) at 10 dBm.

Freq. (GHz)	GUM-Based				MCS-Based			
	RCC		SWR		RCC		SWR	
	x	U	x	U	x	U	x	U
2	10.086	0.026	10.051	0.034	10.050	0.029	10.041	0.033
5	10.055	0.031	10.037	0.039	10.014	0.036	10.007	0.041
8	10.101	0.032	10.063	0.041	10.068	0.038	10.066	0.043
12	10.064	0.036	10.036	0.048	10.046	0.042	10.049	0.050
15	10.047	0.041	10.040	0.057	10.035	0.049	10.025	0.060
18	9.979	0.042	10.002	0.059	9.999	0.049	10.002	0.061

The laboratories shown on the x-axis are presented in a randomised order to avoid any implicit ranking or bias in visual interpretation. Each marker represents the DoE of a participating laboratory at a given frequency, while the vertical error bars indicate the associated expanded uncertainty $U(D_i)$. The blue-dotted horizontal line in the Figure 1 to Figure 12 indicates the degrees of equivalence, which is the difference between the result of the participants of the travelling standard, and x_{CRV}. The different marker styles correspond to different frequency points. Measurement results with $E_n \geq 1$ are highlighted with red dotted circles, indicating potential inconsistency with the comparison reference value.

The results showed that most participant measurements were in good agreement with the CRV values across all frequencies and power levels. The majority of participant results satisfied the acceptance criterion $|E_n| < 1$. This indicates a high level of consistency among participating laboratories and confirms the robustness of the measurement methodology.

It should be noted that for one participating laboratory, the uncertainty component associated with measurement non-reproducibility was not fully included in the uncertainty evaluation performed using the AutoRFPower software. As a result, the reported uncertainties at the −10 dBm level were significantly underestimated. For example,

at 2 GHz and −10 dBm, the expanded uncertainty obtained from the software was approximately four times smaller than the value estimated by the laboratory when including non-reproducibility effects.

A comparison between RCC- and SWR-based evaluations reveals that the use of vector RCC information generally provides slightly improved agreement compared to scalar SWR-based modelling, particularly at higher frequencies where mismatch effects are more pronounced. Similarly, the agreement between GUM and MCS results indicates that the linearisation assumptions of the GUM approach are valid for most cases, while the MCS approach provides additional robustness in scenarios involving non-linear effects.

Some deviations were observed for specific laboratories and frequency points. In particular, significant deviations were identified for one participant, which may be related to instrument zeroing issues or differences in measurement procedures. These deviations are more noticeable at lower power levels and higher frequencies, where measurement sensitivity and mismatch-related effects become more critical.

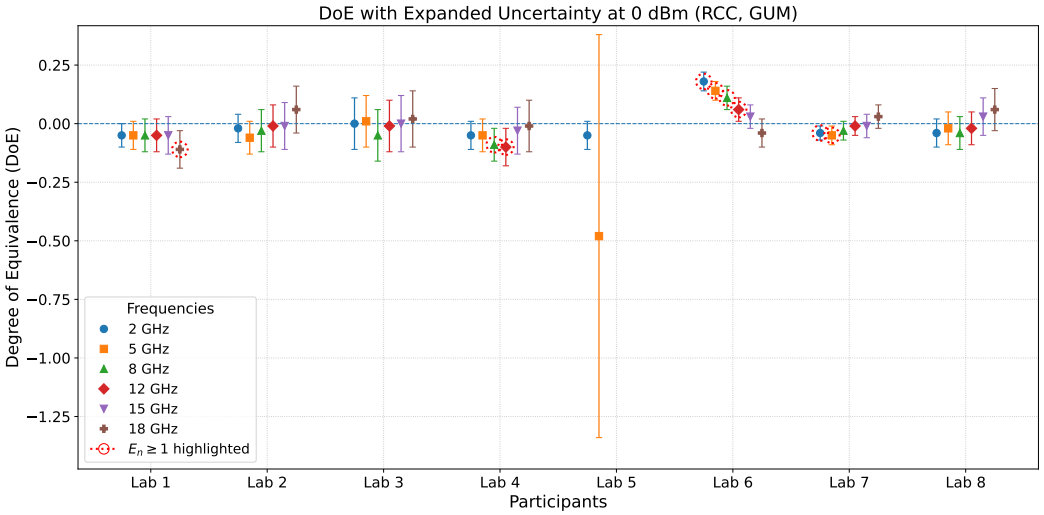

Figure 1. Degrees of equivalence with expanded uncertainties at 0 dBm for the RCC case according to the GUM method. Red dotted circles indicate points with $E_n \geq 1$.

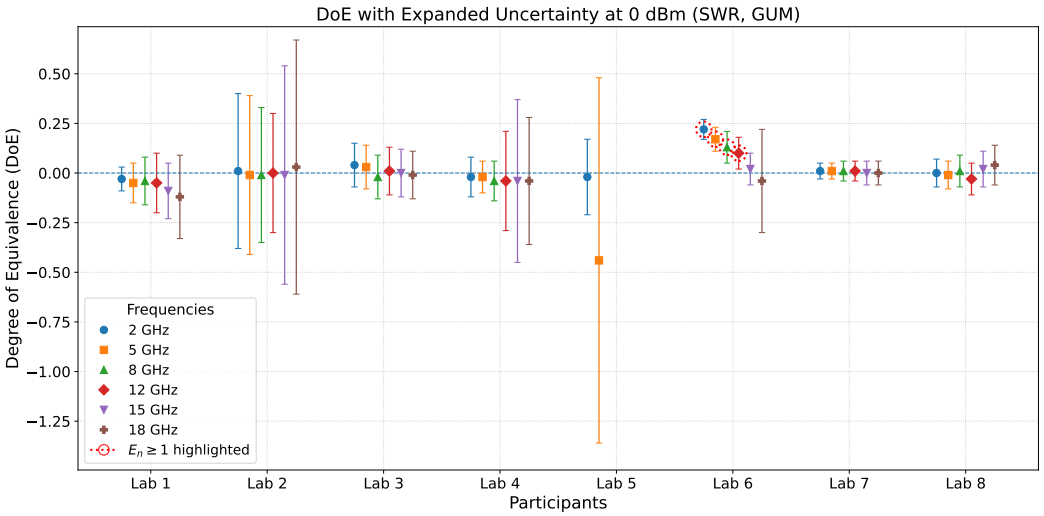

Figure 2. Degrees of equivalence with expanded uncertainties at 0 dBm for the SWR case according to the GUM method. Red dotted circles indicate points with $E_n \geq 1$.

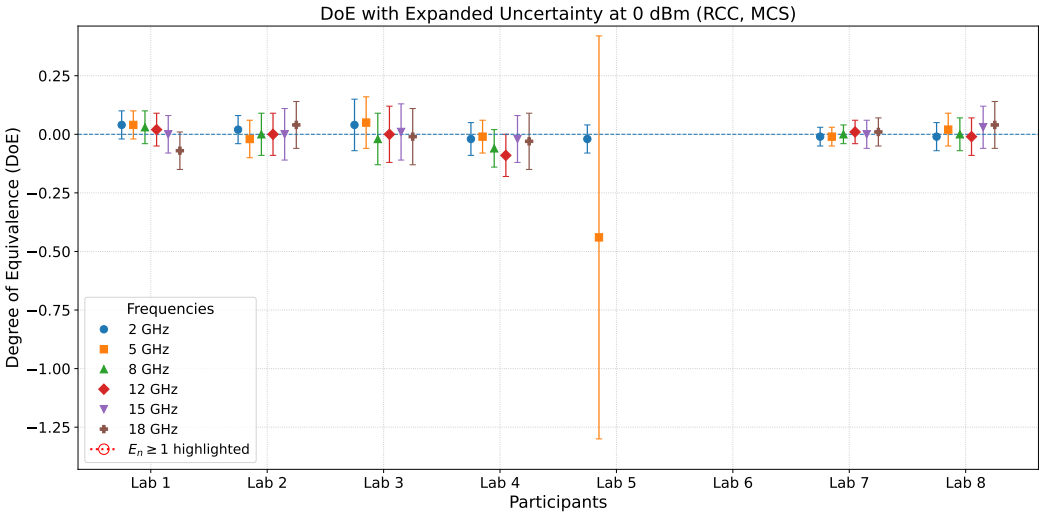

Figure 3. Degrees of equivalence with expanded uncertainties at 0 dBm for the RCC case according to the Monte Carlo simulation (MCS) method. Red dotted circles indicate points with $E_n \geq 1$.

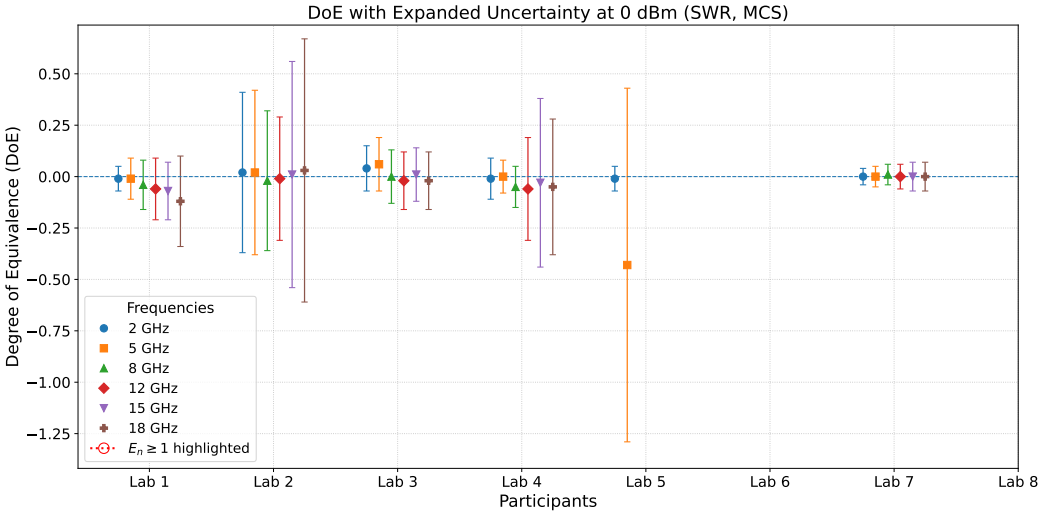

Figure 4. Degrees of equivalence with expanded uncertainties at 0 dBm for the SWR case according to the Monte Carlo simulation (MCS) method. Red dotted circles indicate points with $E_n \geq 1$.

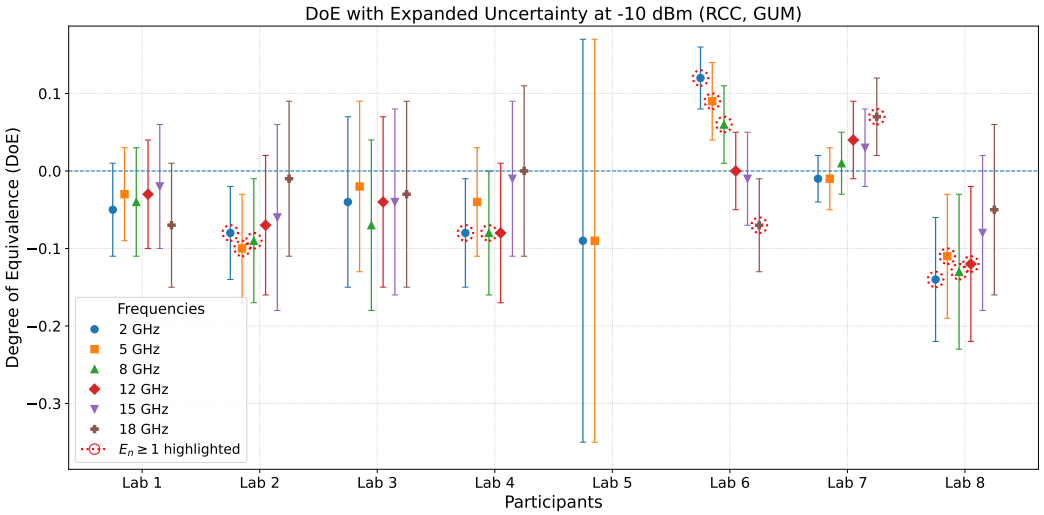

Figure 5. Degrees of equivalence with expanded uncertainties at −10 dBm for the RCC case according to the GUM method. Red dotted circles indicate points with $E_n \geq 1$.

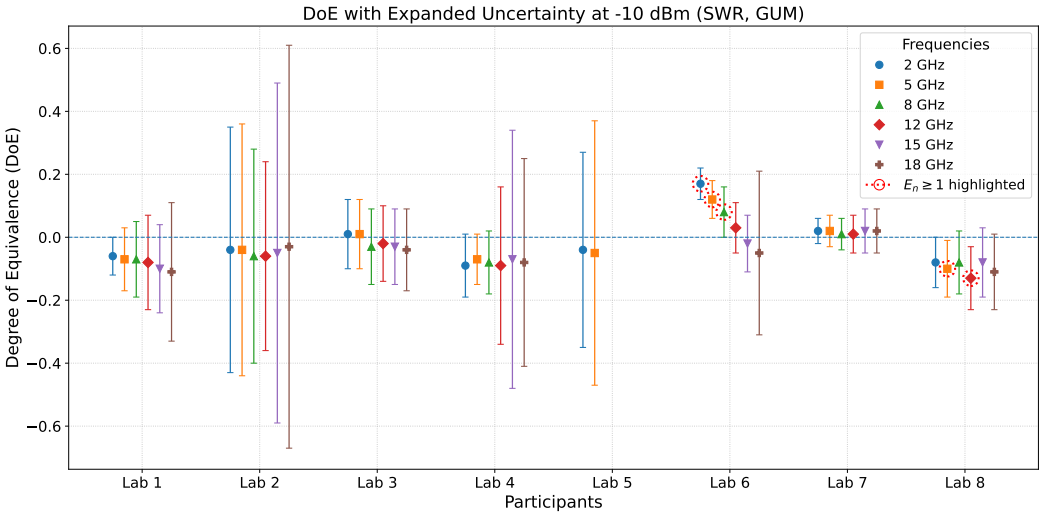

Figure 6. Degrees of equivalence with expanded uncertainties at −10 dBm for the SWR case according to the GUM method. Red dotted circles indicate points with $E_n \geq 1$.

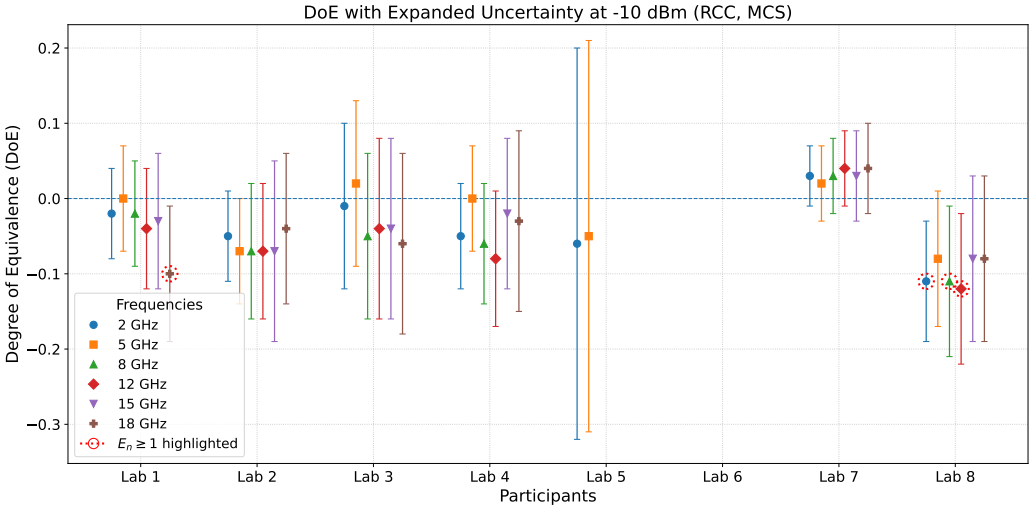

Figure 7. Degrees of equivalence with expanded uncertainties at −10 dBm for the RCC case according to the Monte Carlo simulation (MCS) method. Red dotted circles indicate points with $E_n \geq 1$.

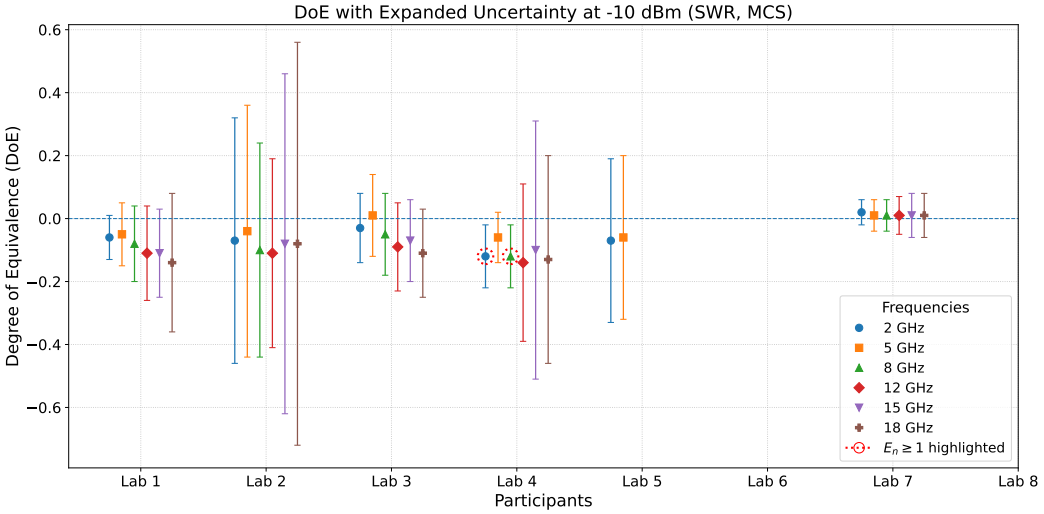

Figure 8. Degrees of equivalence with expanded uncertainties at −10 dBm for the SWR case according to the Monte Carlo simulation (MCS) method. Red dotted circles indicate points with $E_n \geq 1$.

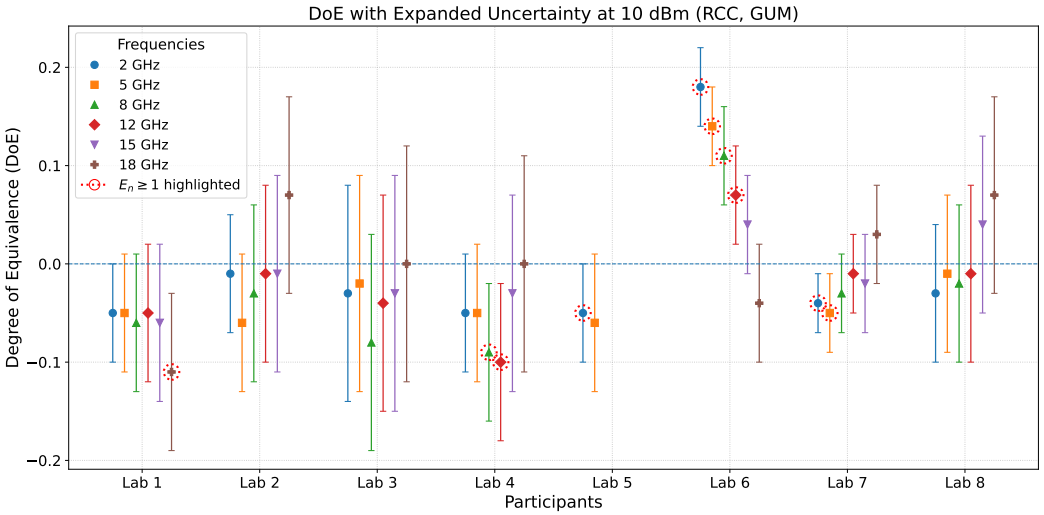

Figure 9. Degrees of equivalence with expanded uncertainties at 10 dBm for the RCC case according to the GUM method. Red dotted circles indicate points with $E_n \geq 1$.

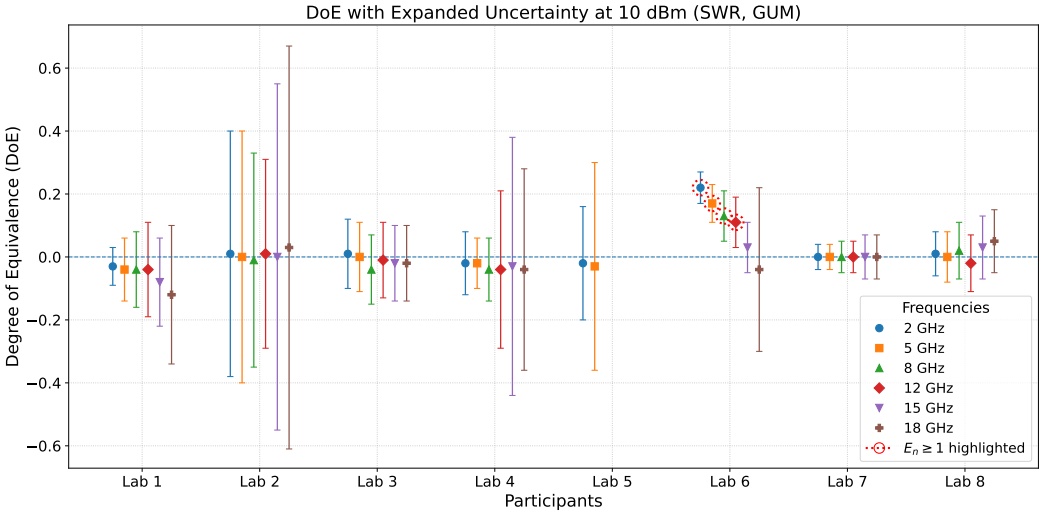

Figure 10. Degrees of equivalence with expanded uncertainties at 10 dBm for the SWR case according to the GUM method. Red dotted circles indicate points with $E_n \geq 1$.

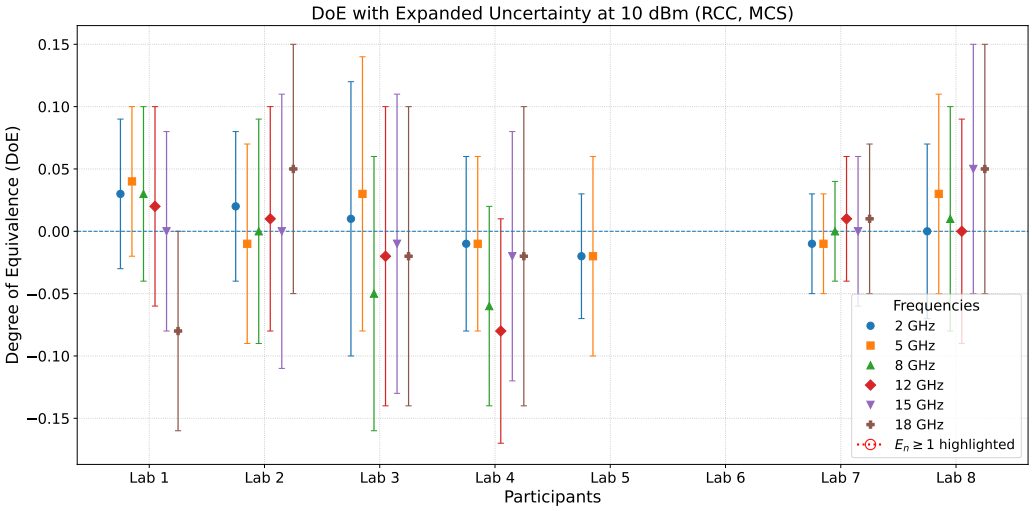

Figure 11. Degrees of equivalence with expanded uncertainties at 10 dBm for the RCC case according to the Monte Carlo simulation (MCS) method. Red dotted circles indicate points with $E_n \geq 1$.

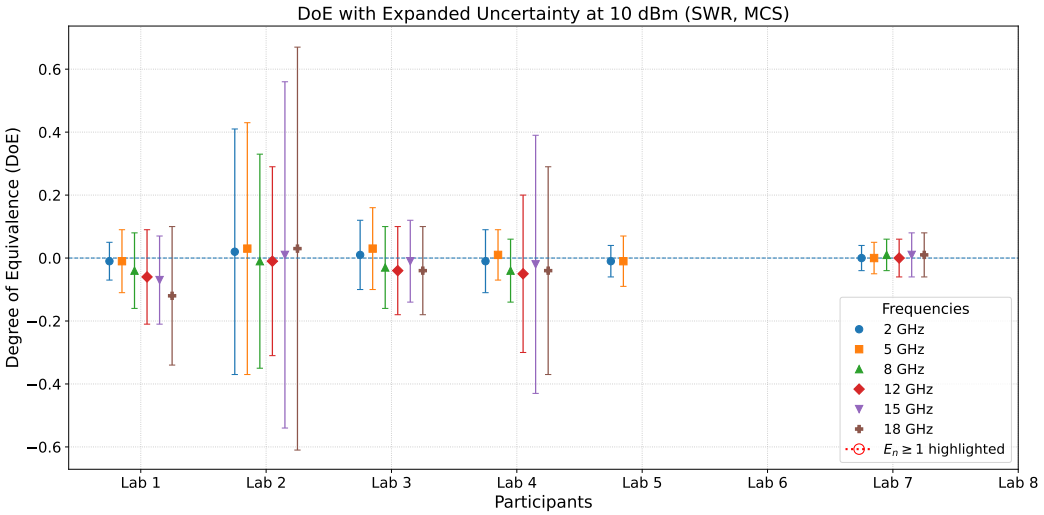

Figure 12. Degrees of equivalence with expanded uncertainties at 10 dBm for the SWR case according to the Monte Carlo simulation (MCS) method. Red dotted circles indicate points with $E_n \geq 1$.

The uncertainty values obtained using GUM and MCS approaches were generally consistent, although slight differences in uncertainty distributions were observed in certain cases.

6. Discussion

The comparison results highlight the influence of reflection coefficient modelling on RF power measurements. Differences between RCC- and SWR-based uncertainty evaluations demonstrate that modelling assumptions can significantly affect comparison outcomes.

The study also shows that the AutoRFPower software can harmonise measurement procedures and uncertainty evaluation across laboratories in conducted interlaboratory comparisons. Thus, the utilisation of such software helps to improve measurement process and uncertainty calculations.

Furthermore, the comparison provides insight into the applicability of automated measurement and uncertainty evaluation tools in RF power metrology. The results indicate that such tools can support reliable interlaboratory comparisons and contribute to the development of consistent RF power measurement methodologies.

The current version of AutoRFPower follows the standard inclusive approach, in which each participant's result also contributes to x_{CRV}. This can introduce a self-correlation bias, since a laboratory's own value pulls the reference value towards itself and may make performance metrics such as E_n appear more favourable. To mitigate this risk, incorporating the "exclusive statistics" proposed by Steele et al. [25], in which each participant is excluded from the reference value used to assess it, represents an important direction for future software refinements aimed at unbiased laboratory evaluations.

7. Conclusions

The comparison was successfully conducted among eight laboratories participating in the EPM project 22RPT04, "RFMicrowave2". The comparison results demonstrate generally good agreement among participants and confirm the feasibility of automated RF power measurement approaches. The high level of agreement observed across different frequencies and power levels confirms the validity of both the measurement methodology and the uncertainty evaluation approaches.

The scope of this project encompassed an interlaboratory comparison of RF power measurements using automatic power measurement software (AutoRFPower), which yielded

favourable results. This allowed project participants to gain capability for automatic RF power measurement and uncertainty calculation. The study demonstrates that automated measurement workflows can be successfully implemented in RF power metrology without compromising accuracy or reliability.

The findings of this study support the integration of automated measurement and uncertainty evaluation tools into RF metrology practices and provide a basis for future harmonised comparison studies. In addition, the results highlight the importance of using vector-based mismatch modelling (RCC) for improved accuracy in high-frequency measurements.

Based on our findings, we suggest that previously measured vector RCCs of signal generators can be used as a travelling standard for RF power comparisons. This approach offers a practical and scalable alternative for future interlaboratory comparisons, reducing the need for mismatch characterisation.

Future work will focus on extending the automation capabilities demonstrated in this study, in order to design smarter and more adaptive metrology systems. Comprehensive reviews in recent literature have demonstrated that machine learning algorithms can successfully optimise complex measurement procedures and enable robust, data-driven metrology applications [26]. Building upon this, one particularly promising avenue is the integration of artificial intelligence (AI) techniques for the automatic extraction and interpretation of calibration procedures from manufacturers' documentation, thus enabling the generation of harmonised and machine-interpretable calibration workflows and helping less experienced laboratories implement complex measurement procedures in a consistent and traceable manner.

Furthermore, combining automated measurement systems with AI-based decision support mechanisms has the potential to improve uncertainty assessment and method selection processes. By integrating domain knowledge, historical measurement data, and formalised calibration requirements, future systems will be able to help operators choose optimal measurement strategies and identify potential sources of error. These advancements are likely to contribute to ongoing digital transformation efforts in the field of metrology.

Author Contributions: Conceptualisation, A.C., A.K.D., E.D., M.H., J.G., O.S., T.L., S.M., S.P., Ł.U., M.B. and E.Ç.; methodology, A.C., A.K.D., E.D., M.H., J.G., O.S., T.L., S.M., S.P., Ł.U., M.B., E.Ç. and H.O.; software, A.C., M.C.K. and H.O.; validation, A.C., A.K.D. and E.D.; formal analysis, A.K.D., E.D. and J.G.; investigation, A.K.D., E.D., M.H., J.G., O.S., T.L., S.M., S.P., Ł.U., M.B. and E.Ç.; resources, T.L.; data curation, A.K.D.; writing—original draft preparation, A.C. and A.K.D.; writing—review and editing, A.C., A.K.D., M.C.K. and E.D.; visualisation, A.C. and A.K.D.; supervision, H.O., M.H. and E.D.; project administration, E.D. and M.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the EPM project titled "RFMicrowave2" with grant number 22RPT04.

Data Availability Statement: No new data were created or analyzed in this study.

Acknowledgments: This study was conducted within the framework of the 22RPT04 RFMicrowave2 project, which has received funding from the European Partnership on Metrology, co-financed by the European Union's Horizon Europe Research and Innovation Programme and by the Participating States.

Conflicts of Interest: Spark Calibration Services and Trescal A/S have no commercial interest in this work. Spark Calibration Services used related software in the article to collect, analyse and interpret datas obtained from the measurements for the purpose of comparison. The software is not fully designed by Spark Calibration Services, it is developed by the participating states (Middle East Technical University, TUBITAK UME and Spark Calibration Services) within the 22RPT04 project.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
CIPM	International Committee for Weights and Measures
CMC	Calibration and Measurement Capability
CRV	Comparison Reference Value
DI	Designated Institute
DoE	Degree of Equivalence
EMC	Electromagnetic Compatibility
En	Normalised Error
EPM	European Partnership on Metrology
GUM	Guide to the Expression of Uncertainty in Measurement
IoMT	Internet of Measurement Things
JCGM	Joint Committee for Guides in Metrology
MCS	Monte Carlo Simulation
MRA	Mutual Recognition Arrangement
MW	Microwave
NMI	National Metrology Institute
PDF	Probability Density Function
RCC	Complex Reflection Coefficient
RF	Radio Frequency
SI	International System of Units
SWR	Standing Wave Ratio
UCaaS	Uncertainty Calculation as a Service
VNA	Vector Network Analyser

References

1. Salzenstein, P.; Wu, T.Y.; Pavlyuchenko, E. The Critical Role of International Comparisons in Global Metrology System: An Overview. *Metrology* **2025**, *5*, 74 . [CrossRef]
2. Janik, D.; Inoue, T.; Michaud, A. Final report on CCEM key comparison CCEM. RF-K10. CL (GT-RF/99-2)'Power in 50 Ω coaxial lines, frequency: 50 MHz to 26 GHz'–measurement techniques and results. *Metrologia* **2006**, *43*, 01009.
3. Patel, K.; Negi, P.; Swarup, R.; Kim, J.H.; Kang, J.S.; Chow, M.W.; Tsui, C.; Dressler, E.; Shay, W.T.; Li, H.; et al. Final report on APMP supplementary comparison P1-APMP. EM. RF-S3: APMP comparison of 50 ohm coaxial mismatches. *Metrologia* **2012**, *49*, 01013.
4. Pozar, D.M. *Microwave Engineering*, 4th ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2012; pp. 26–30.
5. Weidman, M.P. *Direct Comparison Transfer of Microwave Power Sensor Calibrations*; U.S. Government Printing Office: Washington, DC, USA, 1996.
6. Jargon, J.A.; Gu, D.; Long, C.J.; Hagerstrom, A.M.; Stelson, A.C.; Monke, A.F. *Theoretical Basis of the Direct-Comparison System for Power Calibration Including Equivalent Source Mismatch*; Technical Report; National Institute of Standards and Technology (NIST): Gaithersburg, MD, USA, 2019.
7. Jia, J.; Kuang, J.; He, Z.; Fang, J. Design of automated test system based on GPIB. In *Proceedings of the 2009 9th International Conference on Electronic Measurement & Instruments*; IEEE: New York, NY, USA, 2009; pp. 1-943–1-948. [CrossRef]
8. *JCGM 100:2008*; Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement (GUM). Technical Report; Joint Committee for Guides in Metrology (JCGM): Geneva, Switzerland, 2008.
9. *JCGM 101:2008*; Evaluation of Measurement Data—Supplement 1 to the GUM: Propagation of Distributions Using a Monte Carlo Method. Technical Report; Joint Committee for Guides in Metrology (JCGM): Geneva, Switzerland, 2008.
10. RFMicrowave2 Project. Development of RF and Microwave Metrology Capability II. 2026. Available online: https://rfmwii.cmi.gov.cz/index.php (accessed on 11 March 2026).
11. Kaya, M.C.; Saeedi Nikoo, M.; Schwartz, M.L.; Oguztuzun, H. Internet of measurement things architecture: Proof of concept with scope of accreditation. *Sensors* **2020**, *20*, 503. [CrossRef] [PubMed]
12. Çetinkaya, A.; Doğan, A.K.; Danaci, E.; Oğuztüzün, H. Autorfpower: Automatic rf power measurement software for metrological applications. In *Proceedings of the 2021 2nd International Informatics and Software Engineering Conference (IISEC)*; IEEE: New York, NY, USA, 2021; pp. 1–4.

13. Cetinkaya, A.; Kaya, M.C.; Danaci, E.; Oguztuzun, H. Uncertainty calculation as a service: Integrating cloud-based microservices for enhanced calibration and DCC generation. *Sensors* **2024**, *24*, 5651. [CrossRef] [PubMed]
14. Neyezhmakov, P.; Zub, S.; Pivnenko, S. Preliminary e-infrastructure for digital metrology. *Meas. Sens.* **2022**, *23*, 100398. [CrossRef]
15. Verma, P. Potential role of metrology in digital transformation for quality infrastructure. *MAPAN* **2024**, *39*, 111–118. [CrossRef]
16. Cox, M.G. The evaluation of key comparison data. *Metrologia* **2002**, *39*, 589–595. [CrossRef]
17. Rukhin, A.L.; Sedransk, N. Statistics in metrology: International key comparisons and interlaboratory studies. *J. Data Sci.* **2007**, *5*, 393–412. [CrossRef]
18. Koepke, A.; Lafarge, T.; Possolo, A.; Toman, B. Consensus building for interlaboratory studies, key comparisons, and meta-analysis. *Metrologia* **2017**, *54*, S34–S62. [CrossRef]
19. Cetin, E.; Doğan, A.K.; Cetinkaya, A.; Danaci, E. VNA-Based Vector Reflection Coefficient Measurement Technique for Powered RF Signal Generators. *Sensors* **2026**, *26*, 3590. [CrossRef] [PubMed]
20. Guldbrandsen, T. Uncertainty contributions from mismatch in microwave measurements. *IEE Proc.-Microwaves Antennas Propag.* **2001**, *148*, 393–397. [CrossRef]
21. Kartal Dogan, A.; Berginc, M.; Cetin, E.; Danacı, E.; Grajciar, J.; Lippert, T.; Prendergast, S.; Sibonjic, O. *Comparison on RF Power Measurements Using Automatic Power Measurement Software*; Zenodo: Genève, Switzerland, 2026. [CrossRef]
22. Danacı, E.; Kartal Doğan, A.; Çiçek, E.C.; Çetinkaya, A.; Kaya, M.Ç.; Oğuztüzün, M.S.H.; Tunay, G. Uncertainty Evaluation Using Law of Propagation and Monte Carlo Simulation Methods with the AutoRFPower Measurement Software. *Konya J. Eng. Sci.* **2024**, *12*, 596–607. [CrossRef]
23. Harris, I.; Warner, F. Re-examination of mismatch uncertainty when measuring microwave power and attenuation. In *Proceedings of the IEE Proceedings H (Microwaves, Optics and Antennas)*; IET: Stevenage, UK, 1981; Volume 128, pp. 35–41.
24. *ISO/IEC 17043*; Conformity Assessment—General Requirements for Proficiency Testing. International Organization for Standardization: Geneva, Switzerland, 2023.
25. Steele, A.; Wood, B.; Douglas, R. Exclusive statistics: Simple treatment of the unavoidable correlations from key comparison reference values. *Metrologia* **2001**, *38*, 483–488. [CrossRef]
26. Vallejo, M.; De La Espriella, C.; Gómez-Santamaría, J.; Ramírez-Barrera, A.F.; Delgado-Trejos, E. Soft metrology based on machine learning: A review. *Meas. Sci. Technol.* **2020**, *31*, 032001.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.