Project Title

Measurement of Areal Roughness by Optical Microscopes

Coordinator, Institute, Country

Sai Gao, Ludger Koenders, PTB (Germany)

EURAMET Registration No.

1242

Subject Field

Length

Date

2023-11-30

Physikalisch-Technische Bundesanstalt (PTB),
Braunschweig, Germany

Report on Euramet – Project No. 1242
Measurement of Areal Roughness
by Optical Microscopes

Final Report

(Concluded without using the results from VNIIMS (participant from Russia) for the analysis following the EURAMET recommendation on April 7, 2022)

A. Felgner (PTB), D. Hüser (PTB), T. Dziomba (PTB), L. Koenders (PTB), S. Gao (PTB),
P-E. Hansen (DFM), G-B. Picotto (INRIM), R. Bellotti (INRIM), M. Zucco (INRIM),
A. Bossen (METAS), F. Meli (METAS), Christian Kottler (METAS),
V. Heikkinen (VTT MIKES), A. Lassila (VTT MIKES).

Braunschweig, November 2023

Glossary

CSI: Coherence Scanning Interferometry
PSI: Phase Shifting Interferometry
CF: Confocal Microscope
VSI: Vertical Scanning Interferometry
WLI: White Light Interferometry

Abbreviations of all instruments and measurement modes used by participants*:

PTB L CF: Olympus LEXT OLS 4100 Laser scanning confocal microscope
PTB S CF: Sensofar S neox in confocal mode
PTB S VSI: Sensofar S neox in CSI mode
DFM Plμ CF: Sensofar PLμ neox in confocal mode
DFM Plμ VSI: Sensofar PLμ neox in CSI mode
INRIM Plμ CF: Sensofar PLμ 2300 in confocal mode
INRIM Plμ CF 50ELWD: Sensofar PLμ 2300 in confocal mode with extra long working distance objective 50x
INRIM Plμ CF 125ELWD: Sensofar PLμ 2300 in confocal mode with extra long working distance objective 125x
INRIM Plμ PSI: Sensofar PLμ 2300 in interferometry PSI mode
INRIM Plμ VSI: Sensofar PLμ 2300 in interferometry CSI mode
VTT MIKES VSI: Bruker contour GT-K coherence scanning interferometer
METAS WLI:Home built microscope in white light interferometry
METAS PSI:Home built microscope in phase shift interferometry

*Objectives and the corresponding specifications used for this comparison are listed in Table 26 in Appendix A

Contents

1 Document control

Version Draft A.1 Issued on 20 January 2020.

Version Draft A.2 Issued on 31 October 2022.

Version Draft A.3 issued on 31 January 2023.

Version Draft A.4 issued on 19 April 2023.

Version Draft B issued on 11 July 2023.

Version Final issued on 30 November 2023.

2 Introduction

The metrological equivalence of national measurement standards and of calibration certificates issued by national metrology institutes is established by a set of key and supplementary comparisons chosen and organized by the Consultative Committees of the CIPM or by the regional metrology organizations, like EURAMET, in collaboration with the Consultative Committees.

At its meeting in 2012, the EURAMET TC-L decided to start a pilot comparison on measurement of areal roughness with PTB as the pilot laboratory. The comparison was later registered as EURAMET project No 1242, the artefact circulation started in March 2017 and was completed in May 2018.

The report is **confidential** to the participants. Copies of the whole report or parts of it should not be handed over to non-participants. Subsequently, the procedure outlined in the BIPM guidelines (2) should be followed.

3 Organization

3.1 Participants

The participants of this pilot comparison are listed in table 1.

Table 1. List of participant laboratories and their contacts.

Laboratory Code	Contact person, Laboratory	Phone, Fax, email
CEM	Laura Carcedo CENTRO ESPAÑOL DE METROLOGÍA (CEM) Laboratorio de Calidad Superficial Length Area Alfar, 2- Tres Cantos 28760 Madrid Spain	+34 91 8074 795 / 801 lcarcedo@cem.minetur.es
DFM	Paul-Erik Hansen Danish Fundamental Metrology (DFM) 307 Matematiktorvet 2800 Kgs. Lyngby Denmark	+45 4525 5879 peh@dfm.dk
INRIM	Massimo Zucco (Previous: Gian Bartolo Picotto) Istituto Nazionale di Ricerca Metrologica (INRIM) Strada delle Cacce 91 10135 Torino Italy	+39 011 3919 968 m.zucco@inrim.it (+39 011 3919 969 or 973 g.picotto@inrim.it)

METAS	Christian Kottler (previous: Felix Meli) Federal Institute of Metrology (METAS) Division Length, Optics and Time Lindenweg 50 3003 Bern-Wabern SWITZERLAND	Tel: Christian.Kottler@metas.ch (+41 58 387 03 46 felix.meli@metas.ch)
NPL	Andrew Lewis, Dave Gunn, Lakshmi Nimishakavi National Physical Laboratory (NPL) Hampton Road Teddington Middlesex TW11 0LW United Kingdom	+44 20 8943 6074 andrew.lewis@npl.co.uk dave.gunn@npl.co.uk lakshmi.nimishakavi@npl.co.uk
VTT MIKES	Antti Lassila, Ville Heikkinen VTT Technical Research Centre of Finland - Centre for Metrology (MIKES) Tekniikantie 1 02150 Espoo Finland	+358504155980 Ville.Heikkinen@vtt.fi +358 40 514 8658 antti.lassila@vtt.fi
VNIIMS	Vladimir Kosteev Russian Research Institute for Metrological Service (VNIIMS) 46, Ozwenaya st. Moscow 119 631 Russia	+7 (495) 781 4506 vkosteev@vniims.ru
PTB (Pilot)	Sai Gao, Thorsten Dziomba (previous: Ludger Koenders) Physikalische-Technische Bundesanstalt (PTB) Bundesallee 100 38116 Braunschweig Germany	+49 531 592-5167 sai.gao@ptb.de +49 531 592-5122 thorsten.dziomba@ptb.de

During the comparison some changes with the participants occurred:

NPL withdrew due to other activities with higher priority
CEM withdrew due to unexpected problems with the instrument
In this version the measurement results from VNIIMS are excluded from the analysis according to the recommendation of EURAMET on April 7, 2022 due to the Russian invasion in Ukraine.

3.2 Schedule

The participating laboratories agreed to measure the samples during the timetable given in table 2. Each laboratory had four weeks that include customs clearance, measurement and transportation to the following participant. With its confirmation to participate, each laboratory agreed to perform the measurements in the allocated period and to allow enough time in advance for transportation so that the following participant received the samples in time.

NPL was originally scheduled for Oct. 2017, withdrew on 29. August 2017.
CEM was originally scheduled for July 2017, asked for postponement and finally withdrew on 23. February 2018.
METAS was scheduled for June 2017, asked for postponement and finally measured in April 2018.

Table 2. Schedule of the comparison.

Institute	Original	Received	Sent to next	Data received	Comment
PTB	Mar. 2017		13.03.2017		
DFM	Apr. 2017	20.04.2017	23.05.2017	29.08.2018	
INRIM	May 2017	23.05.2017	23.06.2017	11.06.2018	

METAS	Jun. 2017	26.06.2017	11.07.2017		No measurements due to delayed samples, shift to March 2018, cleaning on the samples Fts and ARS F1
CEM	Jul. 2017	17.07.2017	27.07.2017		No measurements, shift to April 2018
PTB	Aug. 2017	03.08.2017	08.09.2017		1. check 1.delivery 25.08.2017 Return 30.08.2017 2. delivery 08.09.2017
VNIIMS	Sep. 2017	18.09.2017		28.11.2017	Contract for the comparison and sending of samples
PTB		22.11.2017	11.01.2018		2. check
NPL	Oct. 2017				Participation withdrawn
VTT MIKES	Nov. 2017	17.01.2018	28.02.2018	21.08.2018	
METAS	Mar. 2018	05.03.2018	12.04.2018	25.06.2018	
CEM	Apr. 2018				Equipment not ready, withdrawn
PTB	May 2018				final check
INRIM	Dec. 2018	15.12.2018	28.01.2019		B1000 only, for re-measurement

4 Artefacts

4.1 Description of artefacts

A total of seven samples was to be measured in this comparison (Table 3).

On some samples, a dot on the metal or glass plate marks the “North” in order to allow convenient mounting in the right orientation. It is generally assumed that the samples are mounted and measured in the same orientation as shown here.

Table 3. List of artefacts.

No.	Type	Manufacturer	Identification	Nominal value
1	Step height	Fraunhofer Institute of Microelectronics Stuttgart (ims)	SH B 1000	d ~ 1000 nm
2	Flat	SiMetricS	FtS	Sa < 10 nm
3	Pitch	SiMetricS	RS-N	300 nm < p < 6 µm d < 190 nm
4	Roughness (2D-FIB)	m2c / PointElectronic	UFRS145-100-20 ID 005	Sa ~ 10 nm
5	Roughness (AIR)	SiMetricS	ARS-F2	Sa ~ 45 nm
6	Roughness (AIR)	SiMetricS	ARS F1	Sa ~ 80 nm
7	Roughness (AIR)	Rubert & Co	B40 VP04	Sa ~ 800 nm

4.2 Stability of artefacts

The pilot laboratory has made the stability measurements on the artefacts before and after the comparison by confocal microscopy. The results of these stability check of the seven samples are listed in Table 4 to Table 10.

Table 4. Stability of artefact No. 1 step height (SH B 1000) during comparison, measured by Lext CF Mode

Step height B1000	**Profile *h* / nm**	**Histogram *h* / nm**
Feb. 17	1017.1	1017.3
Jun. 18	1016.9	1017.1
Difference	0.2	0.2

Table 5. Stability of artefact No. 2 Flat (FtS) during comparison, measured by Lext CF Mode

FtS	**S-Filter/*µm***	**L-Filter/*µm***	***Sq /nm***
Feb. 17	none	none	3.40
Jun. 18	none	none	7.11
Difference			3.71

We assume that the relatively large difference is caused by increasing surface contamination over time.

Table 6. Stability of artefact No. 3 Pitch (RS-N) during comparison, measured by Lext CF Mode

RS-N		**Pitch / µm**	**h / µm**
3 µm pitch	Feb. 17	2.981	173.6
	Jun. 18	2.968	169.7
	Difference	0.014	3.9
4 µm pitch	Feb. 17	3.974	181.2
	Jun. 18	3.965	175.0
	Difference	0.008	6.2
6 µm pitch	Feb. 17	5.929	178.8
	Jun. 18	5.947	177.4
	Difference	0.017	1.5

Table 7. Stability of artefact No. 4 Roughness (UFRS145-100-20 ID 005) during comparison, measured by Lext CF Mode

UFRS Roughness (2D FIB)	**S-Filter/*µm***	**L-Filter/*µm***	***Sq /nm***
Feb. 17	1.0	none	7.0
Jun. 18	1.0	none	8.1
Difference			1.1

The relatively large difference is caused by the increasing contamination during the circulation of the sample.

Table 8. Stability of artefact No. 5 Roughness (ARS-F2) during comparison, measured by Lext CF Mode

ARS-F2	**S-Filter/*µm***	**L-Filter/*µm***	***Sq /nm***
Feb. 17	1.0	none	61.7
Jun. 18	1.0	none	61.3
Difference			0.4

Table 9. Stability of artefact No. 6 Roughness (ARS-F1) during comparison, measured by Lext CF Mode

ARS-F1	**S-Filter/*µm***	**L-Filter/*µm***	***Sq /nm***
Feb. 17	1.0	none	109.1

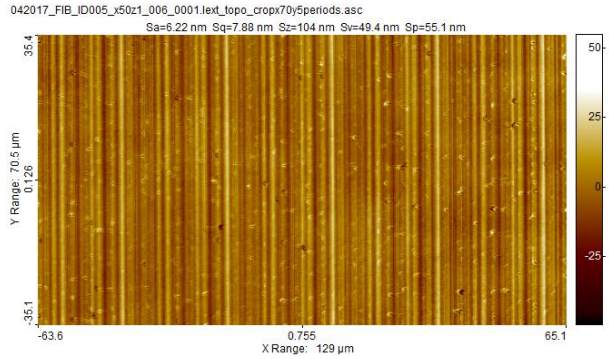
Jun. 18	1.0	none	108.7
Difference			0.4

Table 10. Stability of artefact No. 7 Roughness (B40 VP04) during comparison, measured by Lext CF Mode

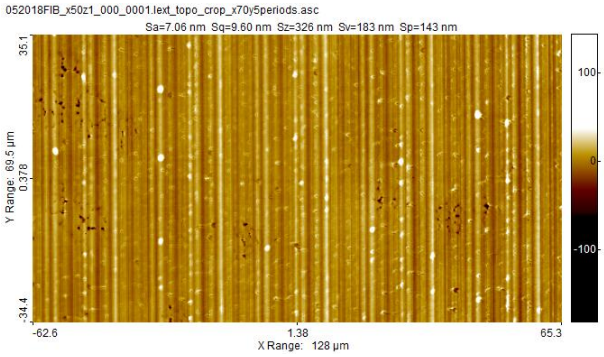
ARI B40	**S-Filter/*µm***	**L-Filter/*µm***	***Sq /nm***
Feb. 17	1.0	500.0	1048
Jun. 18	1.0	500.0	1015
Difference			33

4.3 Condition of artefacts at start/end of comparison

The condition of the measurement surfaces of the artefacts were measured by optical microscopy (confocal microscope) before the start of the comparison and after the comparison. Also, the participants were asked to check the condition of the artefact surfaces prior to their measurements. No apparent damages have been reported. Some changes on the surface of artefact UFRS (FIB 2D) have been observed. Figure 1 (a) and (b) show the topography images of the 2D FIB UFRS measured by the pilot laboratory using the LEXT confocal microscope at the beginning and by the end of the comparison, respectively. By the end of the comparison, more spikes appeared randomly on the surface, and also more apparent holes showed up on the samples compared to the measurement at the beginning. From the stability evaluation it can be seen there is a change of the *Sq* (with 1 µm S filter) value from 7.0 nm to 8.1 nm taken from LEXT measurements. A similar slight increase was observed by AFM control measurements on this sample.



(a) Topography image (LEXT Confocal) of the 2D FIB UFRS at the beginning of the comparison measurement in April 2017, unfiltered



(b) Topography image (LEXT Confocal) of the 2D FIB UFRS by the end of the comparison in June 2018 (PTB final check), unfiltered

Figure 1 Condition of the artefact No. 4 Roughness UFRS145-100-20 ID 005 at start / end of the comparison

5 Measuring instructions

5.1 Measurands

Step height standard

Description of the sample

The step height standard used for this comparison was manufactured by the Fraunhofer Institute of Microelectronics (ims), Stuttgart, for PTB. These standards are available with nominal step heights between 7 nm and 2000 nm in two different layouts A and B. The structures featuring the step height consist of SiO_2 on a Silicon substrate. The whole surface is coated by ~ 100 nm of Chromium.

The standard to be measured in this comparison is of "PTB Layout B" with a nominal step height of 1000 nm (B1000). The Si chip is sized approx. 5 mm x 5 mm and glued onto a steel disk of approx. 12 mm in diameter. The four single bars in the centre of the chip (Figure 2) of approx. 20 µm (left one), 3 µm, 100 µm and 6 µm (right one) in width are interrupted three times to define measurement positions more easily. Orientation structures, frame, logo, serial number and square-shaped may have a step height significantly different from that of the bars intended for calibration.

Measurement positions

The reference area is located on the bar of nom. 20 µm width on the B1000 as indicated in Figure 2, "above" the central interruption of the bar. The first 20 µm of the lower part of the bar should not be used for the analysis.

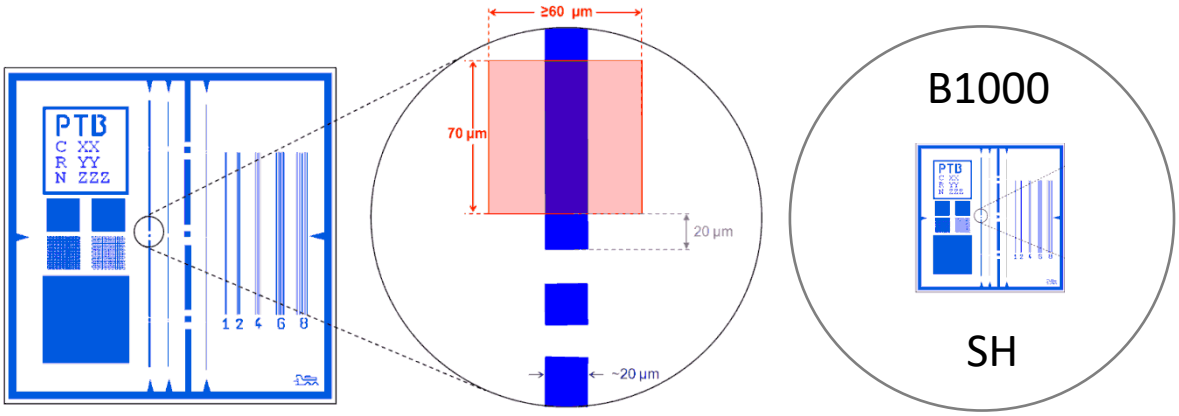


Figure 2 Step height standard with the nominal height of 1000 nm. The inset shows the location of the reference area to be measured, sized approx. 60 µm x 70 µm on the bar of approx. 20 µm in width, starting around 20 µm from the lower end of this bar section.

Measurement conditions[1]

Depending on instrument and mode, the objective and field of view as declared in table 26 had to be used. The centre of the reference area should be in the centre of the field of view. The sample's surface should be aligned as well as possible to avoid large tilting corrections.

Measurand

The field used for the evaluation should be 60 µm x 70 µm. Step height *d* from profile evaluation should be determined according to ISO 25178-700 and using the histogram technique (see 5.2.2).

Silicon flat sample

Description of the sample

The block of polished Silicon (SiMetricS, Typ FtS) of 12 mm x 12 mm x 6 mm (Figure 3) is glued onto a glass with ~ 25.4 mm in diameter. The orientation is given by the letters on the glass. This very flat and smooth surface is meant to determine the quality of the reference plane in optical microscopes or the out-of-plane motion in AFM. Furthermore, it allows to study the noise of instruments.

[1] During a measurement run of all samples (step height, …, roughness) the objective and the zoom factor should be kept constant.

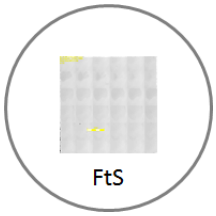


Figure 3 (a) Sketch of the polished Si chip of approx. 12 mm x 12 mm glued onto quartz glass

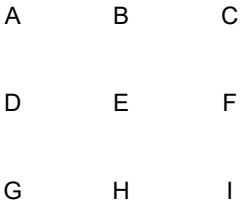


Figure 3 (b) Measurement scheme for the determination of the flatness deviation.

Measurement positions

The participants should measure at the approximate centre of the sample and should check whether the selected region is free from contaminations, scratches and other apparent irregularities.

Measurement conditions

Depending on instrument and mode, the objective and field of view as declared in table 26 had to be used. The mode, objective and zoom factor should be kept constant during the following measurements on other samples. It should be tried to align the sample surface as well as possible to avoid large tilting corrections.

Measurands

Noise

Two measurements should be performed immediately back-to-back. The difference between both data sets should be used as an indication of the noise of the instruments. This noise should be calculated without any alignment of each image and without filtering of the measured data. The *S*q-value of the difference divided by √2 is the measurement noise $N_M = 1/\sqrt{2} * Sq$

Reference plane deviations

For the evaluation of quality of the reference plane used in the following measurements, nine measurements A to I near the centre as sketched in scheme in Figure 3 (b) should be performed. Between each measurement the sample needed to be moved by at least 150 µm.

The reference plane deviation should be obtained by calculation of the average image from all the measurements at positions A to I. From the average image the *S*z and the *S*q values should be determined without any additional filtering.

Multi-pitch resolution sample

Description of the sample

The resolution standard RS-N of SiMetricS company is a Si chip of approx. 10 mm x 10 mm in size with nine 1D gratings arranged in a 3x3 pattern. The nominal pitch p_i of these gratings ranges from 300 nm to 6 µm (Figure 4). The gratings in the diagonal fields from upper left to lower right have the largest pitch with nominally 4 µm, 3 µm and 6 µm and are sized 90 µm x 150 µm. All other six gratings of finer pitch are larger with a size of 110 µm x 150 µm each.

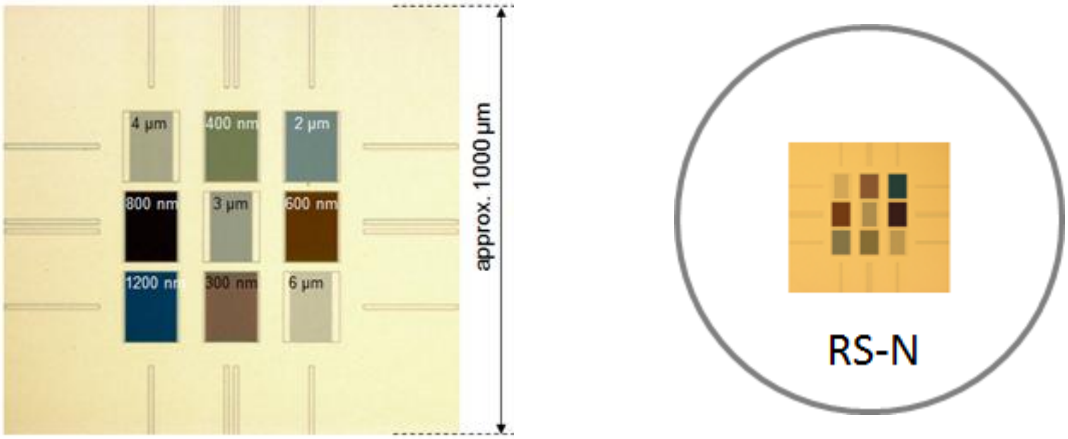


Figure 4 Optical micrograph (approx. 1 mm x 1 mm) of SiMetricS resolution standard of type RS-N with nine 1D-gratings.

Measurement positions

The sample is glued onto a glass plate of 25.4 mm in diameter. The letters on the substrate give the orientation. In this comparison, all fields should be measured to determine the Width limit for full height transmission parameter W_l of the instrument.

Measurement condition

Depending on instrument and mode, the objective and field of view as declared in table 26 had to be used. The mode, objective and zoom factor should be kept constant during the following measurements on other samples. It should be tried to align the sample surface as well as possible to avoid large tilting corrections. All fields should be measured.

Measurands

From the profiles obtained on each grating, the height h_i should be determined as function of the nominal pitch value p_i. The Width-for-full-height-transmission parameter W_l should be determined.

FIB produced 2D roughness pattern

Description of the sample

Silicon chip produced in cooperation of m2c (now merged with point electronics, Halle/Saale, Germany) and PTB by using focussed ion beam technique to generate a given 2D profile pattern (see Figure 5). The nominal value is *S*a ~10 nm. The sample is glued onto an Al chip of approx. 15 mm x 15 mm x 3 mm.

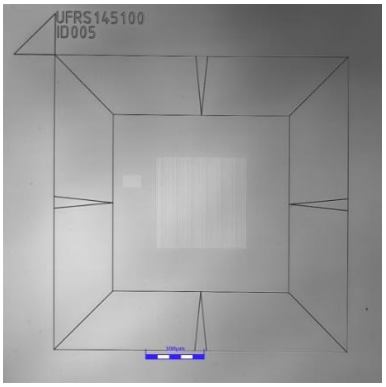


Figure 5 (a) Image of the Si sample with orientation marks and sample identification.

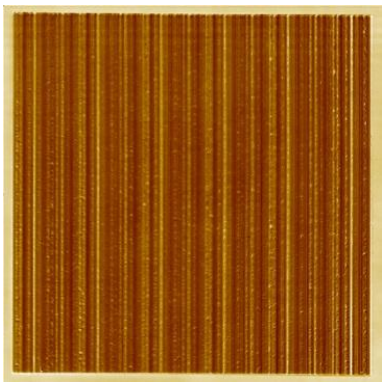


Figure 5 (b) Image 152 µm x 152 µm showing the reference field

Measurement positions

The identification number is in the upper left part of the surrounding rim. This should be used as orientation of the sample to the "Top". There are two fields on the sample. The field close to the lower boundary of the sample was to be taken. Arrows guide to the field of interest. The central part of this field should be used for the measurements. The pattern is a profile which is more than five times repeated. The advantage of such a periodic pattern is that it is less sensitive to the initial starting point. However, the condition to be fulfilled is that the evaluation length is sufficiently accurate.

Measurement conditions

Depending on instrument and mode, the objective and field of view as declared in table 26 had to be used. The mode, objective and zoom factor should be kept constant during the following measurements on other samples. The sample surface should be aligned as well as possible to avoid large tilting corrections.

Measurands

The S-filter should be applied and the L-Filter should be selected depending on your field of view (see table 12). Afterwards an evaluation field in the x-direction with five repeated periods and in the y-direction with a range of 70 µm should be chosen. All roughness parameters as defined in table 13 should be determined by using unfiltered (without S- and L-filters) and filtered (with S- and L-filters applied) data.

Silicon roughness sample

Description of the samples SiMetricS RnS

The silicon roughness sample is a prototype of the second generation of isotropic roughness samples in the ARS product series of the company SiMetricS. The ARS product series has been developed in a joint project with PTB[2]. The samples of the ARS series are based on a polished surface of a (100) Si wafer on which a rough surface is produced by lapping with aqueous or oily suspensions containing grains of SiC or diamond on a plane support plate.

The ARS chips are generally sized approx. 10 mm x 10 mm. This roughened Si chip is glued onto a glass plate approx. 25.4 mm in diameter. The letters on the substrate holder give the direction to the "Top" apart from a simplified SiMetricS logo in the lower left corner of the chip.

Fine roughness samples of type f1 and f2

The surface of the fine roughness samples f1 and f2 is divided into 4x4 quadratic fields of 1760 µm x 1760 µm nominal size, marked by dashed grooves visible already by eye (Figure 6 left). In the diagonal fields from upper left to lower right finer dots can be seen with magnifying glasses or optical microscopes, enumerated 1 (by means of one square-shaped dot) in the upper left to 4 (by four square-shaped dots) in the lower right.

Inside the fields (Figure 6 (c)) along the diagonal axis five quadratic measurement fields of different sizes can be seen, starting with nominally 32 µm x 32 µm inner size in the North, then doubling in edge length from field to field, and ending in the South with 512 µm x 512 µm inner size.

Measurement positions

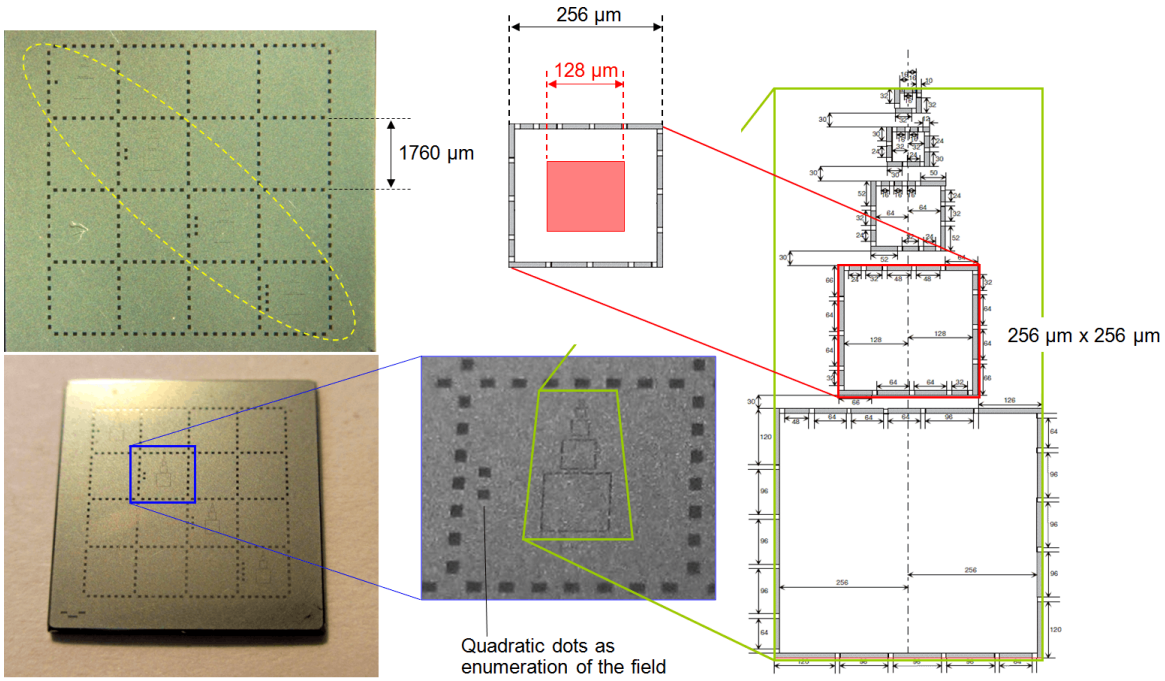
In each of the four diagonal fields (upper left to lower right, marked within the yellow area (Figure 6 (a)) the fourth inner field from the top which has a size of 256 µm x 256 mm has to be measured. The centre of the field of view should be aligned to the centre of each marked 256 µm x 256 µm area.

[2] *Frühauf, J., Krüger-Sehm, R., Felgner, A., Dziomba, T., "Areal roughness standards", Proc. 12th International Conference of EUSPEN, Vol. 1: pp. 133 - 136 (2012)*

Aligning to the marks

If the field of view is smaller than that of the area of 256 µm x 256 µm, it should be tried to align the center of the microscope field of view to the center of the marked area.

It is helpful to focus on the bottom area of the trenches (see Figure 7), as the reflectivity of the lower plateau is much higher, and also the edges are more defined, which eases the alignment. The left marks need to be aligned firstly (x-position, perpendicular to the boundary), then the x-position of the right marks. The central field of view of your instrument should be moved to the center in the x-direction. With the same method the y-positions of the lower and upper marks can be determined and the center in y-direction should be adjusted.



(a) Top view (dashed yellow: the main diagonal) and titled view (with simplified Simetrics logo in the lower left, the two-dot field marked by blue rectangle) of the Si chip

(b) The 256 µm x 256 µm field with the demanded measurement of 128 µm x 100 µm in its center

(c) Sketch of the measurement fields with an inner area ranging from 32 µm x 32 µm to 512 µm x 512 µm; the red frame marks the 256 µm x 256 µm field to be measured

Figure 6 Fine roughness sample ARS

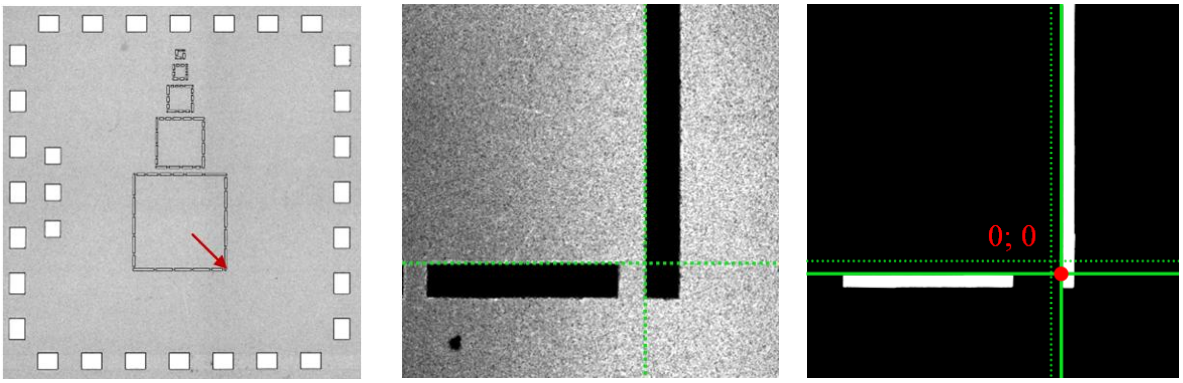


Figure 7 (a) Example shows the alignment to the mark at edge

Figure 7 (b) focus plane at the top surface

Figure 7 (c) focus plane at the trench bottom

Measurement conditions

Depending on instrument and mode, the objective and field of view as declared in table 26 had to be used. The mode, objective and zoom factor should be kept constant during the following measurements on other samples. It should be tried to align the sample surface as precisely as possible to avoid large tilting corrections.

Measurands

The S-filter should be applied and the L-Filter should be chosen depending on your field of view (see table 12). For the parameter calculation a central field of 128 µm by 100 µm should be selected. All roughness parameters as defined in table 13 should be determined from this data by using unfiltered (without S- and L-filters) and filtered (with S- and L-Filter applied) data. The mean value and the standard deviation for each roughness parameter should be calculated from all four fields.

Areal irregular roughness standard

Description of the sample Rubert & Co. AIR B40

This is an areal irregular roughness sample from Rubert & Co (see Figure 8). The size is ~ 20 mm x 20 mm x 2.5 mm.

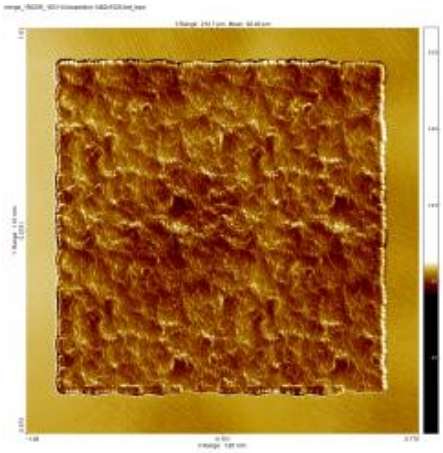


Figure 8 (a) Image of sample AIR B40.

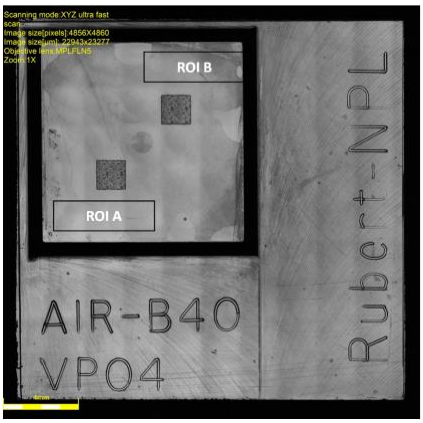


Figure 8 (b) Layout of the sample with Region of Interest (ROI) A and B. In the comparison region B is used.

Measurement positions

There are two regions of interest A and B on the sample. Region B as indicated in Figure 8 (b) should be chosen. Due to the size it is necessary to record a larger area and, consequently, stitch some images. A stitching in the x-direction should be used. Therefore, within the region around the centre of the y-axis should be measured.

Measurement conditions

The left and right boundaries should be aligned perpendicular to x-direction and parallel to the y-direction. The boundary is a little bit rough; it should be averaged as far as possible. The centre of the field of view in the y-direction should be aligned with the centre of the marked area B in y-direction. A number of at least 7 images with an appropriate overlap should be taken and stitched together.

Depending on instrument and mode, the objective and field of view as declared in table 26 was to be used. The mode, objective and zoom factor should be kept constant during the following measurements on other samples. It should be tried to align the sample surface as precisely as possible to avoid large tilting corrections.

Measurands

The S-filter need to be applied and the L-Filter (table 12) should be chosen depending on your stitched field. For the parameter calculation a central field of 800 µm by 100 µm should be selected. All roughness parameters as defined in table 13 should be determined from this data by using unfiltered (without S- and L-filters) and filtered (with S- and L-Filter applied) data.

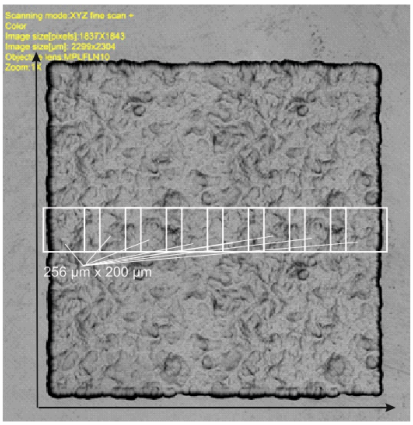


Figure 9 Region B (AIR B40) with a field size of 1560 µm x 1560 µm. Example of image in the x-direction shows 7 stitched image frames.

Flat sample and power spectral density

After finishing of all measurements with one mode and one objective, the flat sample FtS should be measured again to determine the power spectral density function of the instrument for various magnifications.

Measurement positions

An area near to the centre which is free of dust should be chosen again. The sample surface should be aligned as precisely as possible to avoid large tilting corrections. This measurement is used to calculate the power-spectral-density function (see 5.2.5.).

Measurement conditions

Measurements in one mode at the same position should be performed with different magnifications, i.e. 50x, 20x and 10x, and if possible 5x and 100x magnification.

Measurand

The line averaged 1D and if possible, the 2D power-spectral-density function (see 5.2.5.) from each image should be calculated and the result as function of the spatial frequency should be plotted for all magnifications used.

5.2 Parameters and evaluation

General part

All samples used in this comparison are mainly flat. Therefore a straight-line or plane fit for tilt correction was necessary, but no other form operator. With respect to the evaluation of roughness parameters,

other correction methods are allowed for the filtered data. Please see for defined application in passage 5.2.3.

Discrepancies between surface roughness values of different labs might originate both from the measurement itself and from the subsequent data processing and analysis. For a better understanding of the influencing factors, it is therfore vital to distinguish between measurement-realted and analysis-related effects. Therefore, in additon to the analysis done by the participants themseleves, the pilot lab also evaluated the raw topography data sets submitted by the participants by using the analysis software developed at PTB. This allows a central analysis according to uniform procedures and is thus expected to largely eliminate analysis-related discrepancies, so that differences in the measurements can be revealed. The participants were requested to write their raw measurement data on a CD, DVD or USB stick and sent it to the pilot lab by mail. A folder with the individual institute's name and the sample name included in all file names should be created. If different instruments had been used, respective subfolders and the name of the instrument should be included additionally in all file names to avoid any confusion.

The pilot lab assumed the submitted topography to be correctly scaled according to the best knowledge of the participants, unless the participant explicitly informed the pilot lab that the data on the storage media still needed to be corrected by multiplication with a stated correction factor; this information needed to be provided to the pilot lab upon upload of the data. Most data formats used by the major manufacturers are supported by the analysis software envisaged by the pilot lab, unless they were new or had been modified individually. No major difficulties with the submitted data files were observed by the pilot lab.

Step height standards: determination of the step heights

The step height *d* (Figure 10) is defined as in ISO 25178-700 (which is used in analogy to ISO 5436-1, but with more freedom on lengths). Raw data should be used for the evaluation without any form operator or S-filter. If a filter must be applied before, the type of filter and its parameter(s) should be indicated. All values had to be given for the reference temperature of 20° C.

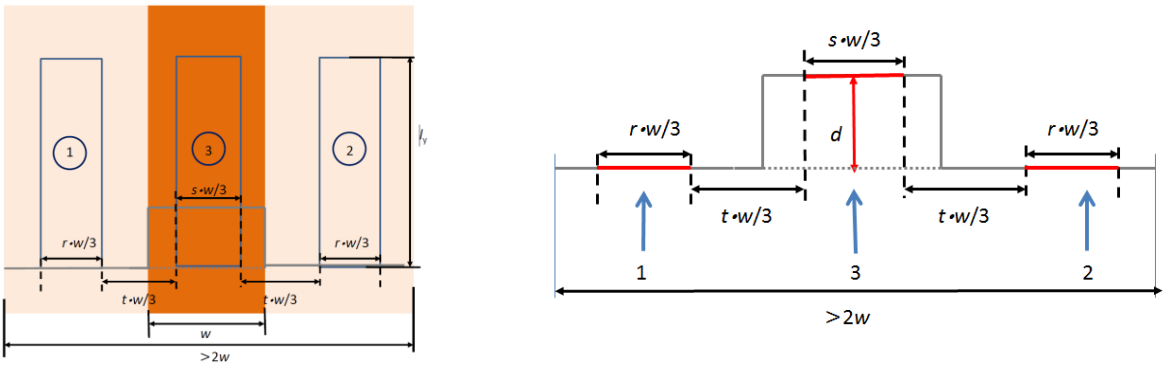


Figure 10 Example of step height d in the top view (left) and profile analysis (right). The values to be used are: w = 20 µm, s*w/3 = 10 µm, r*w/3 = 10 µm, t*w/3 = 15 µm and l_x= 60 µm. The evaluation length in the y-direction (i.e. along the bar) is l_y = 60 µm.

Areal analysis based on mean profile analysis

For the evaluation of the step height/groove depth *d* by profile measurements, a set of at least 10 profiles with parallel total least squares profile line evaluations according to ISO 25178-70 shall be performed. The distance of the profiles shall be chosen so that the calibrated profile rsp. area is covered, including possible positioning error on the specimen area. The step height/groove depth *d* is defined as the perpendicular distance of the mean of the field 3 to the line through the mean of field 1 and the mean of field 2. Evaluation on samples is done by using

$$z(x,y) = ax + by + c + \delta * h/2 \quad (1)$$

Therein

x, *y*	are the lateral coordinates of the measured points
z	is the measured topography as function of x and y
a, *b*, *c*	are unknown parameters
h	is the projected distance (parallel to the z-axis) of the line through regions 1 and 2 and to the line through region 3
d	is the depth/height of the groove/step height to be determined
δ	is a function defined which is

in case of groove(s) given by:

$$\boldsymbol{\delta_{\mathbf{groove}}} = \begin{cases} +\mathbf{1}, & \mathbf{x},\mathbf{y} \in \mathbf{area\ 1} \vee \mathbf{area\ 2} \\ -\mathbf{1}, & \mathbf{x},\mathbf{y} \in \mathbf{area\ 3} \end{cases}$$

and in case of step height(s) given by:

$$\boldsymbol{\delta_{step\ height}} = \begin{cases} -\mathbf{1}, & \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{area\ 1} \vee \boldsymbol{area\ 2} \\ +\mathbf{1}, & \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{area\ 3} \end{cases}.$$

The step height/groove depth *d* is obtained from

$$d = \frac{h}{\sqrt{(1+a^2)(1+b^2)}} \tag{2}$$

Areal analysis based on histogram technique (not standardized)

Besides the ISO 25178-700 procedure we also applied the histogram technique, because it is very often used for step height or groove depth analysis, e. g. in AFM according to ISO 11952:2019. However, it requires a good alignment (levelling) of the surface. The histogram (Figure 11) is the graphical representation of the distribution of measured numerical data above the vertical scale.

The number of data points will be counted in an appropriate z-interval and plotted as function of the vertical axis. In this case only data points *z(x, y)* are used in which the pair (*x*, *y*) is within the area 1, area 2 and area 3. Depending on the quality of the sample surface this will yield two peaks which represent the upper and lower level of the surface. The difference between the maximum peak position of the upper and lower is the step height or groove depth, respectively. A refined extraction of the step height/groove depth *d* is based on the difference between the centers of gravity of the two peaks representing the two height levels of the step height standard, with possibly a threshold (such as the FWHM) applied to exclude extended tails of peaks. A more detailed description is found in ISO 11952:2019.

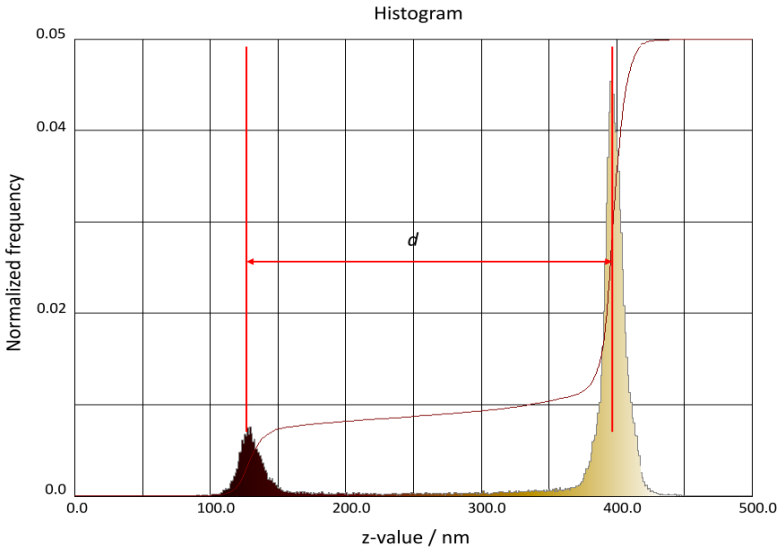


Figure 11 Histogram of distribution of z(x, y) numerical data selected from the areas 1, 2 and 3 in Figure 10 (left) of a step height or groove depth.

Areal filters

For optical measurements the S-filter nesting index is listed in ISO 25178-3. Table 11 displays an excerpt which is related to the samples of the comparison. The nesting index value should be stated in mm.

The roughness measurement data should be filtered firstly to remove noise and high spatial frequency components. If possible, filter described in ISO-16610-71 “Robust areal filter: Gaussian” should be chosen to avoid end effects[3]. If it is not possible, filter described in ISO 16610-61 “Linear areal filter: Gaussian” should be used.

Table 11. Relationships between S-filter nesting index value, sampling distance, and the lateral period limit for optical surface measurements

S-filter nesting index value in mm	**Maximum sampling distance in mm**	**Maximum lateral period limit in mm**
…	…	…
0.000 1	0.000 03	0.000 1
0.000 2	0.000 06	0.000 2
0.000 25	0.000 08	0.000 25
0.000 5	0.000 15	0.000 5
0.000 8	0.000 25	0.000 8
0.001	0.000 3	0.001
0.002	0.000 6	0.002
…	…	…

Table 12. Relationship between nesting index value for L-Filter, S-Filter and bandwidth ratio

[3] Chose it as a linear filter, i.e. $\delta_{ij}^{0} = 1$ and without using M-estimator

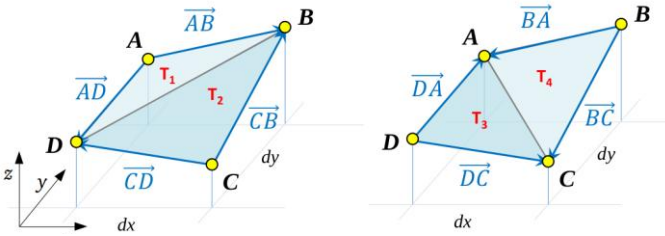
L-Filter nesting index value in mm	S-Filer nesting index in mm	Approach bandwidth ratio between L- and S-Filter nesting index value
…	…	…
0.08	0.000 08	1000:1
	0.000 25	300:1
	0.000 8	100:1
0.1	0.000 1	1000:1
	0.000 2	500:1
	0.000 5	200:1
	0.001	100:1
0.2	0.0002	1000:1
	0.0005	400:1
	0.001	200:1
	0.002	100:1
0.25	0.000 25	1000:1
	0.000 8	300:1
	0.002 5	100:1
0.5	0.000 5	1000:1
	0.001	500:1
	0.002	250:1
	0.005	100:1
0.8	0.000 8	1000:1
	0.002 5	300:1
	0.008	100:1
…	…	…

Terms and definition used in ISO 25178-2

In this comparison the following areal roughness parameters should be determined according to the definition in Table 13. Therein *z*(x, *y*) are scale-limited values, i.e. a S-Filter and/or a L-Filter have already been applied.

Table 13. Areal roughness parameter to be evaluated

Name	Symbol	Equation	Ref.
Terms			ISO 25178-2 :2012
Height	*d*	signed normal distance from the reference surface to the scale limited surface (SLS)	3.2.5
Peak		point on the surface which is higher than all other points within a neighbourhood to that point	3.3.1
Pit		point on the surface which is lower than all other points within a neighbourhood of that point	3.3.2
Amplitude parameter		All height parameters are defined over the definition area.	ISO 25178-2 :2012
Root mean square	*S*q	$Sq = \sqrt{\frac{1}{A}\iint_A z^2(x,y)dx\,dy}$	4.1.1
Skewness of SLS	*S*sk	$Ssk = \frac{1}{S_q^3}\left[\frac{1}{A}\iint_A z^3(x,y)dx\,dy\right]$	4.1.2

Kurtosis of SLS	*Sku*	$Sku = \frac{1}{S_q^4}\left[\frac{1}{A}\iint_A z^4(x,y)dx\,dy\right]$	4.1.3
Maximum peak height of SLS	*Sp*	largest peak height value within a definition area	4.1.4
Maximum pit height of SLS	*Sv*	minus the smallest pit height value within a definition area	4.1.5
Maximum height of SLS	*Sz*	sum of the maximum peak height value and the maximum pit height value within a definition area	4.1.6
Roughness average	*Sa*	$Sa = \frac{1}{A}\iint_A \lvert z(x,y)\rvert\,dx\,dy$	4.1.7
Spatial		All height parameters are defined over the definition area.	
Auto-correlation function	$f_{ACF}(t_x,t_y$	$f_{ACF}(t_x,t_y) = \frac{\iint_A z(x,y) * z(x-t_x, y-t_y)dxdy}{\iint_A z(x,y) * z(x,y)\,dxdy}$	3.2.8
Auto-correlation length	*Sal*	$S_{al} = \min_{t_x,t_y \in R}\sqrt{(t_x^2+t_y^2)}$ where $R = \{(t_x,t_y): f_{ACF}(t_x,t_y) \le s\}$	4.2.1
Texture aspect ratio	*Str*	$S_{tr} = \frac{\min_{t_x,t_y\in R}\sqrt{(t_x^2+t_y^2)}}{\max_{t_x,t_y\in Q}\sqrt{(t_x^2+t_y^2)}}$ where $R = \{(t_x,t_y): f_{ACF}(t_x,t_y) \le s\}$ $Q = \{(t_x,t_y): f_{ACF}(t_x,t_y) \ge s \text{ and } **\}$ where ** is the property that the ACF ≥ s on the straight line connecting the point (tx, ty) to the origin.	4.2.2
Hybrid		All spatial parameters are defined over the definition area.	
Root mean square gradient	*Sdq*	$s_{dq} = \sqrt{\frac{1}{A}\iint_A \left[\left(\frac{\partial z(x,y)}{\partial x}\right)^2 + \left(\frac{\partial z(x,y)}{\partial y}\right)^2\right]dx\,dy}$	4.3.1
Developed interfacial area ratio	*Sdr*	 $T_1 = \frac{1}{2}\sqrt{(z_B - z_A)^2\,dy^2 + (z_D - z_A)^2\,dx^2 + dx^2\,dy^2}$ $T_2 = \frac{1}{2}\sqrt{(z_D - z_C)^2\,dy^2 + (z_B - z_C)^2\,dx^2 + dx^2\,dy^2}$ $T_3 = \frac{1}{2}\sqrt{(z_C - z_D)^2\,dy^2 + (z_A - z_D)^2\,dx^2 + dx^2\,dy^2}$ $T_1 = \frac{1}{2}\sqrt{(z_A - z_B)^2\,dy^2 + (z_C - z_B)^2\,dx^2 + dx^2\,dy^2}$ with $D = (i,j)$, $C = (i+1,j)$, $B = (i+1,j+1)$, and $A = (i,j+1)$ $A_{i,j} = \frac{1}{2}(T_1 + T_2 + T_3 + T_4)$ $S_{\mathrm{dr}} = \left[\frac{1}{(N-1)(M-1)\,dx\,dy}\left(\sum_{i=1}^{N-1}\sum_{j=1}^{M-1} A_{i,j}\right) - 1\right]\cdot 100\%$	4.3.2 – Note: Here we use a corrected version!

2D Power Spectral Density (PSD) function

In addition to the areal roughness parameter, the power spectral density (PSD) function should be calculated for some samples, e.g. flatness. This gives the strength of different roughness components (if calculated for data obtained on rough surfaces) or of the frequency components of the instrument noise (if flat clean surfaces are measured) as function of the spatial frequency and is calculated from the absolute value of the square of the Fourier transform of the height values $z(x, y)$[4]:

The areal Fourier transform is given by

$$F_z(f_x, f_y) = \int_{L_y/2}^{L_y/2} \int_{-L_x/2}^{L_x/2} z(x,y) exp[-i2\pi(f_x\ x + f_y\ y]dxdy \quad (3)$$

The Power Spectral Density (PSD) is calculated from

$$PSD(f_x, f_y) := \lim_{L_x\to\infty} \lim_{L_y\to\infty} \frac{1}{L_x L_y} F_z^*(f_x, f_y) F_z(f_x, f_y) \quad (4)$$

where

- F^* is the complex conjugate function of F
- f_x, f_y are the spatial frequencies of the surface profile with an area of l_x x l_y.

The boundaries f_{min} and f_{max} depend on the inherent bandwidth of the measurement instrument or the task.

5.3 Measurement uncertainty

The uncertainty of measurement shall be estimated according to the ISO Guide to the Expression of Uncertainty in Measurement. The participating laboratories were encouraged to use their usual model for the uncertainty calculation.

All measurement uncertainties shall be stated as standard uncertainties. If appropriate, the corresponding effective degree of freedom might be stated by the participants. If none is given, ∞ is assumed. For efficient evaluation and subsequent assessment of CMC claims, an uncertainty statement in the following functional form (5) is preferred:

$$u(e_c) = Q[a, b \cdot l_n] = \sqrt{a^2 + (b \cdot l_n)^2} \quad (5)$$

6 Results

6.1 General procedures

The analysis strategy in this comparison was twofold:

Firstly, the participants processed their measurement data by using their own software tools and determined the requested measurands. The results and their uncertainty were reported in the documents.

Secondly, the obtained topographical data of all valid measurements should be sent to the pilot lab for uniform analysis by the evaluation software developed at PTB.

Results obtained in both ways have been compared in the comparison report.

[4] A. *Duparré, J. Ferre-Borrull, S. Gliech, G. Notni, J. Steinert, J.M. Bennett: Surface characterization techniques for determining the root-mean-square roughness and power spectral densities of optical components. In: Appl. Opt. 41. 2002, S. 159-160*

6.2 Results and standard uncertainties as reported by participants

The participants' reports should contain:

- the measurement set-up and the conditions
- details on the measurement procedure
- evaluation methods or deviation from the protocol
- the result(s) of the measurements
- incl. standard deviations where applicable
- the combined standard uncertainty for *d*, *S*a, *S*q, *S*z, …
- the degrees of freedom *v*
- the complete uncertainty budget
- the results have to be stated for the reference temperature at 20 °C
- use the thermal expansion coefficient of the silicon of $(2.5554 \pm 0.0002) \cdot 10^{-6}/K$ [5].

For the optical instruments (ISO 25178-604, 607, …) used, the uncertainty contributions as listed in Table 14 should be considered.

Table 14. Influence quantities affecting metrological characteristics

Component/element	**Influence quantity**	**Affecting metrological element(s)**
Light source	wavelength, bandwidth, polarization	α_z
Interferometer imagining system (CSI)	numerical aperture, magnification, wavefront distortion, quality of the optics (aberration, …) lateral distortion, illumination uniformity	α_x, α_y, α_z, z_{FLT}, l_x, l_y, l_z, W_R, Δ_{PERxy}
Microscope imaging system (CF)	Numerical aperture (pinhole), magnification between object and image size, optical aberrations, general quality of optical components, lateral distortion, illumination uniformity	α_x, α_y, α_z, z_{FLT}, l_x, l_y, l_z, W_R, Δ_{PERxy}
Camera or photodetector	x- and y-pixel spacing, cross-talk	α_x, α_y, W_R
Controller		
- acquisition software	scan speed v, scan length l_z, increment Δz, linearity Δz-lin, integration time, height digitization step	α_z, l_z
- profile analysis software	evaluation of CSI or confocal signal, z-cal-factor	α_z, l_z
Instrument overall	lateral sampling interval, instrument noise, environmental vibration, hysteresis	N_M, l_x, l_y, l_z
Sample	tilt between optical axis and surface normal, relative phase shift upon reflection, thickness of transparent and semi-transparent films, unresolved features	α_x, α_y, α_z,

The measurement report forms in appendix C of the technical protocol were sent by e-mail (Word document) to all participating laboratories. It was appreciated if the report forms (in particular the results sheet) could be completed by computer and sent back electronically to the pilot. In any case, the signed

5 *R. Schödel and G. Bönsch, Precise interferometric measurements at single crystal silicon yielding thermal expansion coefficients from 12 °C to 28 °C and compressibility, in Recent Development in Traceble Dimensional Measurements, Jennifer E. Decker, Nichilas Brown, Editors, Proceedings of SPIE Vol. 4401 (2001)*

report was also to be sent in paper form by mail or electronically as a scanned pdf document. In case of any differences, the signed forms were considered to be the definitive version.

The uncertainty of the measurement should be estimated according to the Guide to the Expression of Uncertainty in Measurement (GUM).

The measurand *y* is expressed as a function of the input quantities x_i

$$y = f(x_i) \tag{6}$$

The combined standard uncertainty $u_c(y)$ is the square mean of the standard uncertainties of the input quantities $u(xi)$, each weighted by a sensitivity coefficient C_i

$$u_c^2(y) = \sum_i c_i^2 u^2(x_i) \text{ with } c_i = \frac{\partial y}{\partial x_i} \tag{7}$$

The participants are requested to report their measurement uncertainty budget in a table whose format corresponds with the scheme below:

Quantity X_i	Estimated x_i	Uncertainty $u(x_i)$	Probability distribution [6] {N, R, T, U}	Sensitivity coefficient c_i	Uncertainty contribution $u_i(y)$	Degree of freedom ν_{eff}

Following receipt of all measurement reports from the participating laboratories, the pilot laboratory has analysed the results and prepared this first draft A.1 report on the comparison. If a participant has clearly reported an anomalous result the participant was invited to check his results for numerical errors within a short time scale but he was not informed about the magnitude or sign of the apparent anomaly.

After the first data checking by the pilot laboratory, INRIM has been notified to check the measured step height values. Therefore, the sample B1000 was sent to INRIM again in Jan. 2019. The revised report was received in February 2019. The corrected step height values from INRIM are used for the reference value calculation.

7 Analysis

7.1 Calculation of the key comparison reference value (KCRV)

The weighted mean was chosen as the method for calculation of the KCRV for each measurand. Before the weights could be assigned and the mean taken, it was necessary to exclude any clear outliers from the analysis. The analysis for each measurand proceeds as follows:

We assume the total number of participants submitting a result is I.

Each laboratory reports a measured value, x_i, and its associated standard uncertainty $u(x_i)$.

<ITERATION START> Compute the normalised weight, w_i, for the result x_i is given by:

$$w_i = C \cdot \frac{1}{[u(x_i)]^2} \tag{8}$$

where the normalising factor, C, is given by:

$$C = \frac{1}{\sum_{i=1}^{I} \left(\frac{1}{u(x_i)}\right)^2} \tag{9}$$

Then calculate the weighted mean, $\bar{x}_w$, which is given by:

$$\bar{x}_w = \sum_{i=1}^{I} w_i \cdot x_i \tag{10}$$

6 For the type of probability distribution please use: N = normal; R = rectangular; T = triangular; U = U-shaped.

The uncertainty of the weighted mean is calculated by:

$$u(\bar{x}_w) = \sqrt{\frac{1}{\sum_{i=1}^{I}\left(\frac{1}{u(x_i)}\right)^2}} = \sqrt{C} \tag{11}$$

After deriving the weighted mean and its associated standard uncertainty, the deviation of each laboratory's result from the weighted mean is determined simply as $x_i - \bar{x}_w$. The uncertainty of this deviation is calculated as a combination of the uncertainties of the result, $u(x_i)$, and the uncertainty of the weighted mean $u(\bar{x}_w)$. The uncertainty of the deviation from the weighted mean is given by equation (12), which includes a minus sign to take into account the correlation between the two uncertainties (it would be a plus sign if dealing with uncorrelated uncertainties, such as when comparing data from two separate laboratories).

$$u(x_i - \bar{x}_w) = \sqrt{[u(x_i)]^2 - [u(\bar{x}_w)]^2} \tag{12}$$

For the determination of the key comparison reference value KCRV, statistical consistency of the results contributing to the KCRV is required. A check for statistical consistency of the results with their associated uncertainties can be made by calculating the E_n value for each laboratory's result, where E_n is defined as the ratio of the deviation from the weighted mean to the expanded uncertainty of this deviation:

$$E_n = \frac{x_i - \bar{x}_w}{2\sqrt{[u(x_i)]^2 - [u(\bar{x}_w)]^2}} \tag{13}$$

The results are examined and any for which $|E_n| > 1$ is considered as inconsistent result. Identify the result with the largest $|E_n|$, go back to <ITERATION START> and repeat the analysis, but excluding this result from contributing to the weighted mean – *i.e.* they have a weighting of zero. Because inconsistent results are no longer correlated with the weighted mean, when calculating their deviation from the weighted mean, and when calculating their E_n value, a positive sign is used in equation (12) and consequently in the denominator of equation (13).

This process is iterated until there are no inconsistent results contributing to the weighted mean. After reaching consistency, the calculated weighted mean is the KCRV (see Table 15 -Table 25).

To calculate the KCRV, all the measurement results, except PLμ confocal 125 ELWD from INRIM, from all participants are taken into account. The reason for the exclusion of PLμ confocal 125 ELWD is that we did not receive a corrected new step height value. And INRIM would like to indicate that PLμ confocal 50x ELWD is suitable for samples of complex shape and geometry but is not currently used to calibrate fine roughness samples such as the four from Simetrics and that from m2c / PointElectronic.

In the case of the step height standard and the UFRS, the E_n criterion was applied, while for the fine roughness standards ARS F1, ARS F2 and Rubert B40 the scattering of reported results needs to be discussed first. The scattering seems to depend on the measurement principle and instrument properties.

In this report we concentrate on the analysis of Sq values, although more parameters were calculated by the participants. We feel that the Sq parameter deserves in-depth analysis and discussion first, before we can reasonably expand the analysis to the other parameters.

Table 15 – Table 25 and Figure 12 – Figure 19 show all the measurement results determined and reported by the participants or evaluated centrally by PTB in comparison of the KCRV. Central data analysis by the pilot lab are shown in Figure 14 (PSD), Table 25 and Figure 19 (b) (Rubert AIR B40).

The detailed information of the instruments and the corresponding specifications used by participants for the comparison is summarized in Table 26 in Appendix A of this document.

Additional AFM reference measurements on the roughness samples 2D-FIB UFRS, ARS F1 and ARS F2 have been realized for comparison with the results obtained by the optical measurement instruments. An area of 106 μm x 107 μm at 1024x1024 pixels, scanned with 0.07 lines/s, approximately in the center of the

field of 256 µm x 256 µm of the samples have been measured by the SIS Nanostation II AFM in the PTB cleanroom center. The AFM data was filtered by S = 1 µm similar to the optical data. The evaluated AFM roughness values are also shown in Figure 16, Figure 17 and Figure 18 in a dotted green line.

Artefact No.1: Step height SH B 1000

Table 15. Step height SH B1000 (artefact No. 1) – profile evaluation: Measurement values, standard and expanded uncertainties, E_n values, degrees of equivalence $\boldsymbol{x_i - \bar{x}_w}$, associated expanded uncertainties $\boldsymbol{U(x_i - \bar{x}_w)}$ and contribution to the weighted mean.

SH B 1000	x_i	$u(x_i)$	$U(x_i)$	$\lvert E_n \rvert$	$(x_i - \bar{x}_w)$	$U(x_i - \bar{x}_w)$	Contribution to ref. val.
	nm	*nm*	*nm*		*nm*	*nm*	
PTB L CF	1017.1	5.6	11.2	0.26	2.8	10.7	Y
PTB S CF	1014.9	4.1	8.2	0.08	0.6	7.5	Y
PTB S VSI	1014.3	3.1	6.2	0.00	0.0	5.3	Y
DFM Plµ CF	1021.1	6.3	12.6	0.56	6.8	12.2	Y
DFM Plµ VSI	1019.4	6.4	12.9	0.41	5.1	12.4	Y
INRIM Plµ CF	1023.3	7.2	14.4	0.64	9.0	14.0	Y
INRIM Plµ VSI	1010.7	5.2	10.4	0.37	-3.6	9.9	Y
INRIM Plµ CF 50ELWD	1007.7	8.2	16.4	0.41	-6.6	16.1	Y
VTT MIKES VSI	1008.3	4.1	8.3	0.80	-6.0	7.5	Y
METAS WLI	1015.0	31.0	61.0	0.01	0.7	61.9	Y
Mean value	1015.2						
Stand. dev.	5.2						
Ref. value	1014.3						
Uncertainty	3.3						

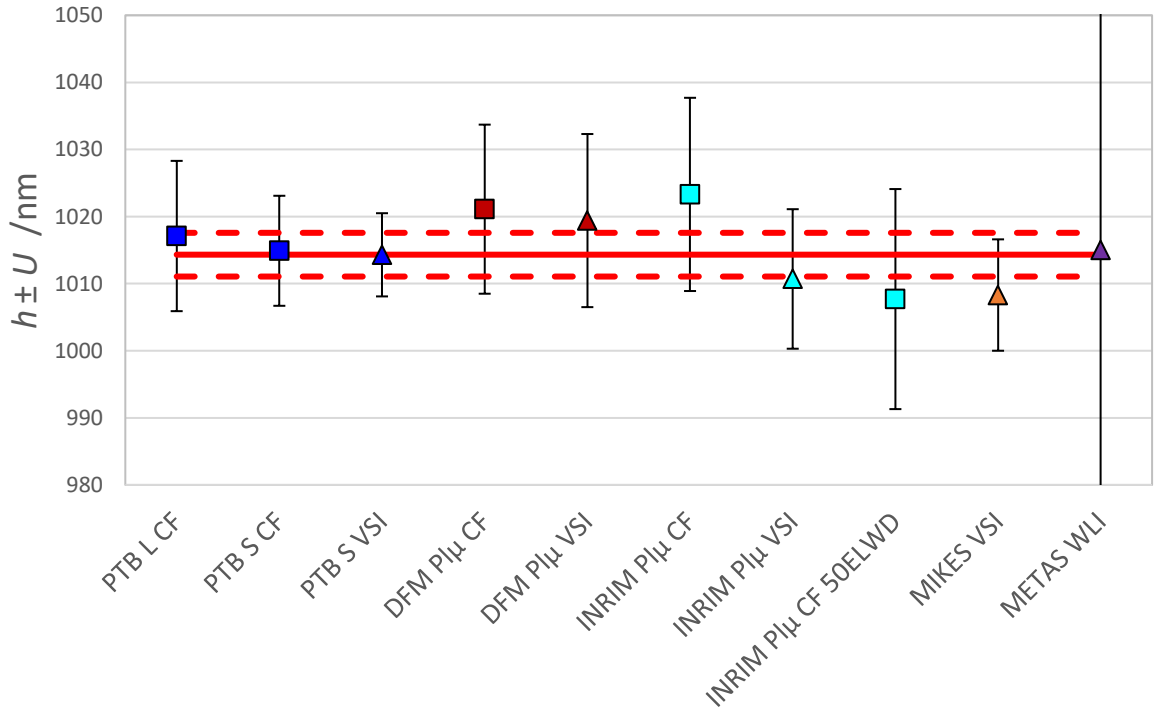


Figure 12 (a). Measured step height values (profile evaluation) of the participants for the step height sample B1000 (red line: reference value, dotted red lines: expanded uncertainty of the reference value).

Table 16. Step height SH B1000 (artefact No. 1) – histogram evaluation: Measurement values, standard and expanded uncertainties, E_n values, degrees of equivalence $\boldsymbol{x_i - \bar{x}_w}$, associated expanded uncertainties $\boldsymbol{U(x_i - \bar{x}_w)}$ and contribution to the weighted mean.

SH B 1000	x_i	$u(x_i)$	$U(x_i)$	$\lvert E_n \rvert$	$(x_i - \bar{x}_w)$	$U(x_i - \bar{x}_w)$	Contribution to ref. val.
	nm	*nm*	*nm*		*nm*	*nm*	
PTB L CF	1017.3	5.6	11.2	0.24	2.6	10.7	Y
PTB S CF	1015.2	4.1	8.2	0.06	0.5	7.5	Y
PTB S VSI	1014.1	3.1	6.2	0.12	-0.6	5.3	Y
DFM Plμ CF	1021.5	7.1	14.2	0.49	6.8	13.8	Y
DFM Plμ VSI	1019.5	8.7	17.4	0.28	4.8	17.1	Y
INRIM Plμ CF	1020.0	7.0	14.0	0.39	5.3	13.6	Y
INRIM Plμ VSI	1011.5	5.2	10.4	0.33	-3.2	9.9	Y
INRIM Plμ CF 50ELWD	1010.1	8.2	16.4	0.29	-4.6	16.1	Y
VTT MIKES VSI	1012.5	3.8	7.6	0.33	-2.2	6.9	Y
METAS WLI	1014.0	31.0	61.0	0.01	-0.7	61.9	Y
Mean value	1015.6						
Stand. dev.	3.9						
Ref. value	1014.7						
Uncertainty	3.3						

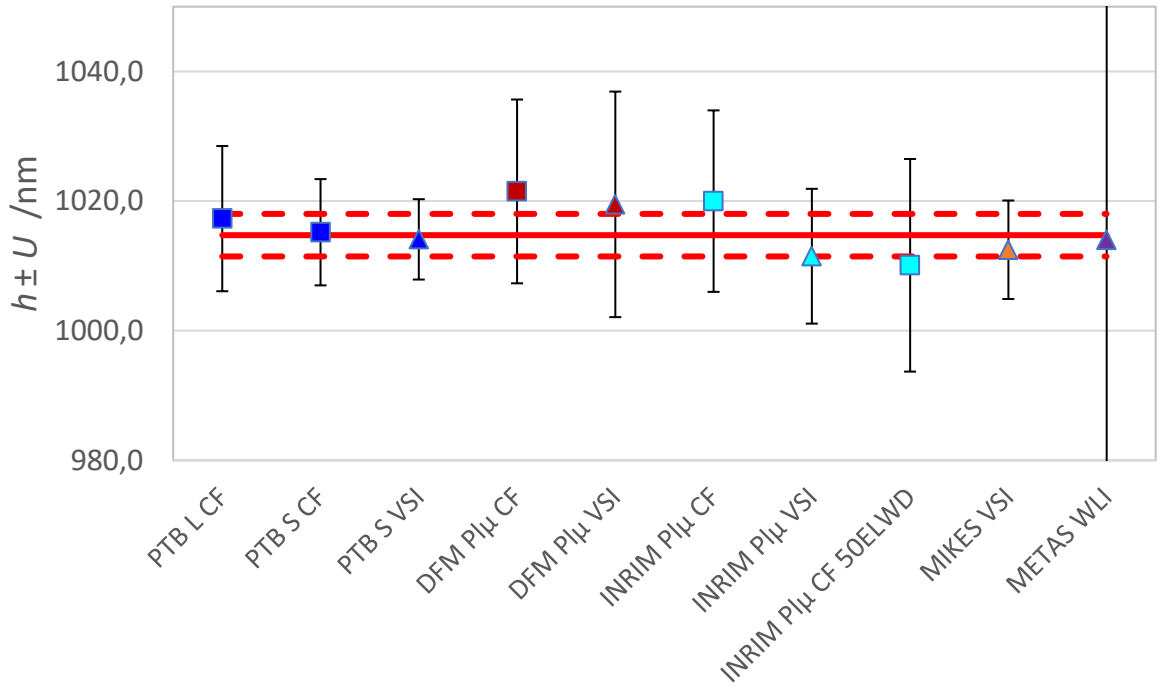


Figure 12 (b). Measured step height values (histogram evaluation) of the participants for the step height sample B1000 (red line: reference value, dotted red lines: expanded uncertainty of the reference value).

Artefact No.2: Flat FtS (mainly used for the calculation of uncertainty)

Table 17. Flat FtS (artefact No. 2) - noise: Measurement values, standard deviations.

FtS	x_i	$\sigma(x_i)$
	nm	*nm*
PTB L CF	1.35	0.16
PTB S CF	1.20	0.08
PTB S VSI	1.37	0.43
DFM Plμ CF	1.36	0.22
DFM Plμ VSI	1.91	0.20
INRIM Plμ CF	1.20	0.20
INRIM Plμ PSI	0.30	0.20
INRIM Plμ CF 50ELWD	3.00	0.80
VTT MIKES VSI	1.45	2.1
METAS WLI	2.91	0.45
METAS PSI	0.59	0.09
Mean value	1.51	
Stand. dev.	0.83	

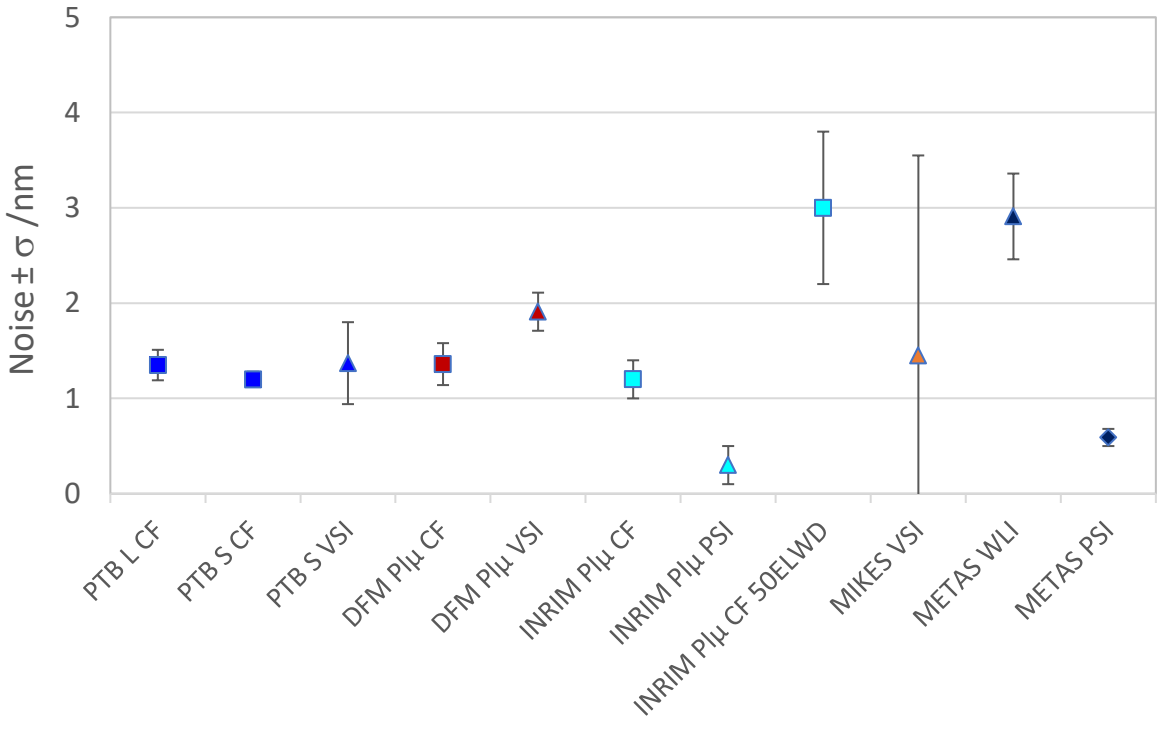


Figure 13 (a). Measured noise values of the participants for the silicon flat sample FtS.

Table 18. Flat FtS (artefact No. 2) – flatness *Sz*: Measurement values, expanded uncertainties.

FtS	x_i	$U(x_i)$

	nm	*nm*
PTB L CF	39.5	8.0
PTB S CF	27.0	8.0
PTB S VSI	16.6	10.0
DFM Plμ CF	51.9	30.1
DFM Plμ VSI	32.5	19.1
INRIM Plμ CF	47.8	21.0
INRIM Plμ PSI	19.4	7.1
INRIM Plμ CF 50ELWD	186.7	64.3
VTT MIKES VSI	13.8	14.2
METAS WLI	11.1	11.7
METAS PSI	3.8	3.0
Mean value	40.9	
Stand. dev.	50.7	

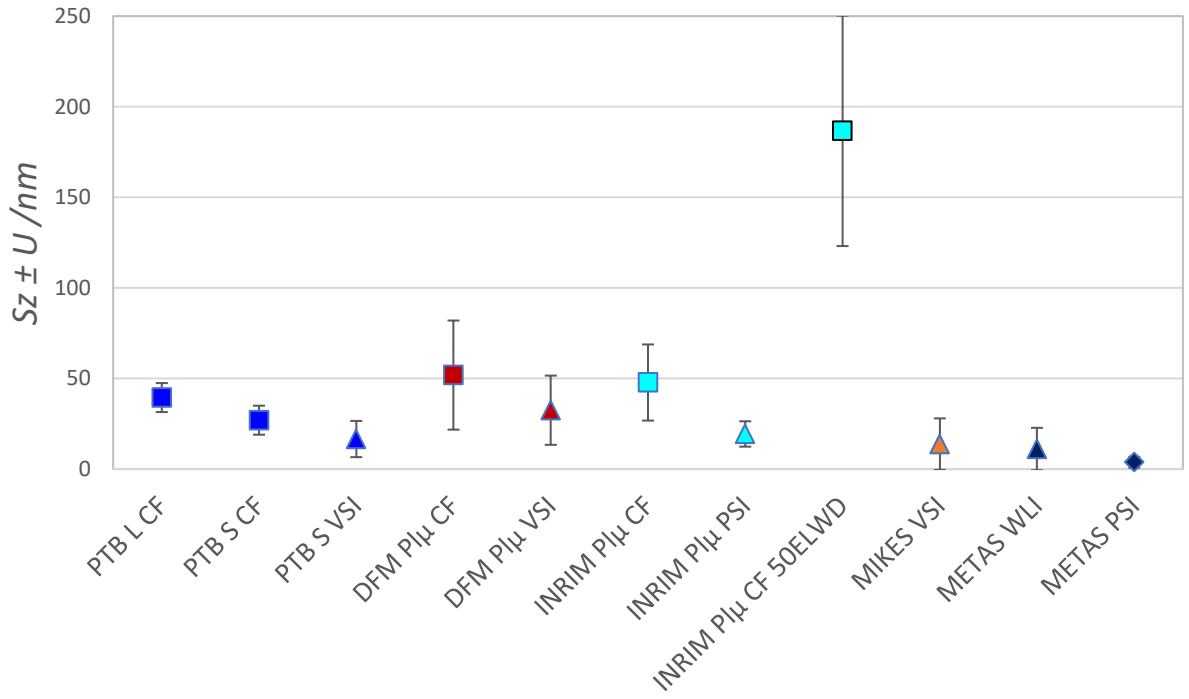


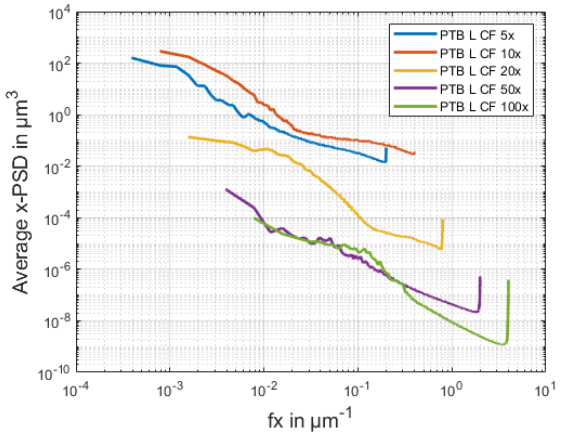
Figure 13 (b). Measured flatness values of the participants for the silicon flat sample FtS.

Most of the instruments show a noise level between 1 and 2 nm. Only two instruments show larger noise. With respect to the flatness deviation, most of the instruments got Sz values of 50 nm or below.

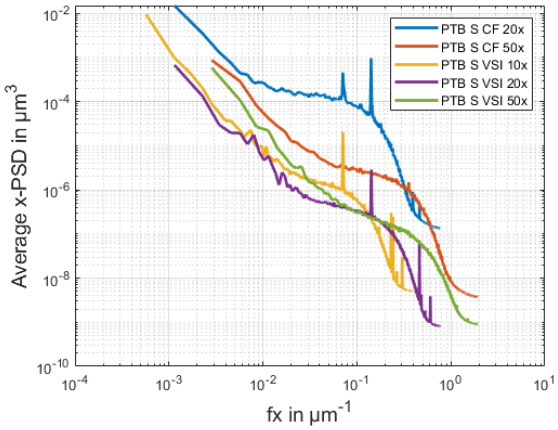
PSD comparison

The PSD reveals the strength of different roughness or noise components as function of the spatial frequency. Some data files and diagrams were received from the participants, however, these diagrams do not allow us to get a good comparison of the instruments. The reason maybe lies in the fact that different software packages are used which do not use the same algorithm to calculate PSD or apply different normalization/scaling of the PSD amplitudes.

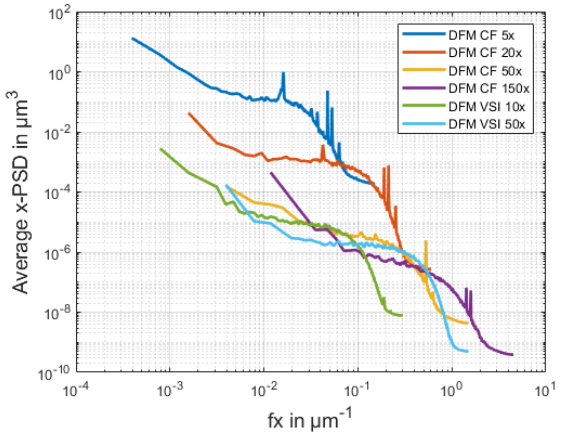
Line averaged 1D and 2D power spectral density functions from the images provided by participants are calculated by PTB. Figure 14(a) shows the average x-PSD diagrams, each for one instrument with different objectives. Figure 14(b) shows averaged x-PSDs obtained by different instruments but the 50x or 40x objectives used in this comparison. The PSDs are evaluated by PTB.



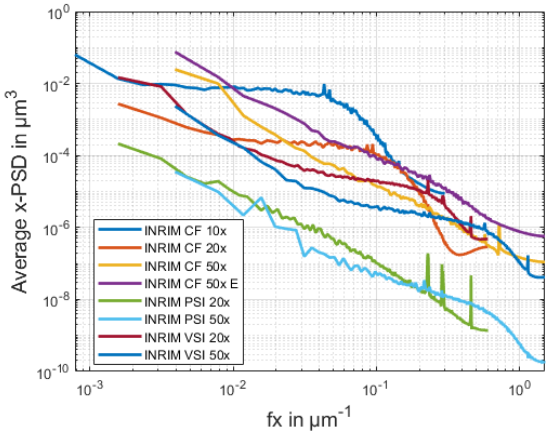
(a) PTB Lext CF



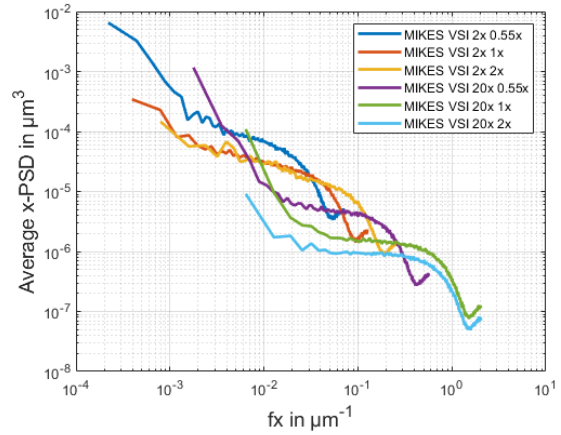
(b) PTB S Neox CF/VSI



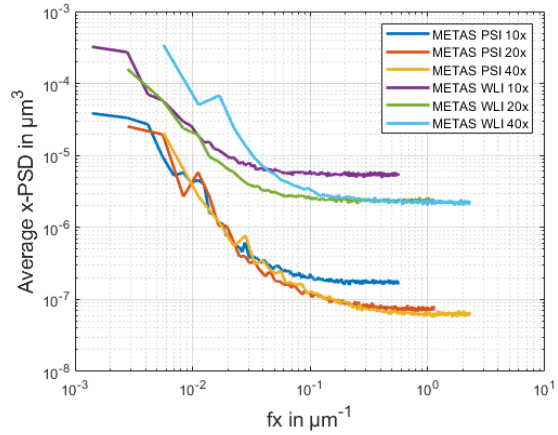
(c) DFM Plµ CF/VSI



(d) INRIM Plµ CF, PSI and VSI

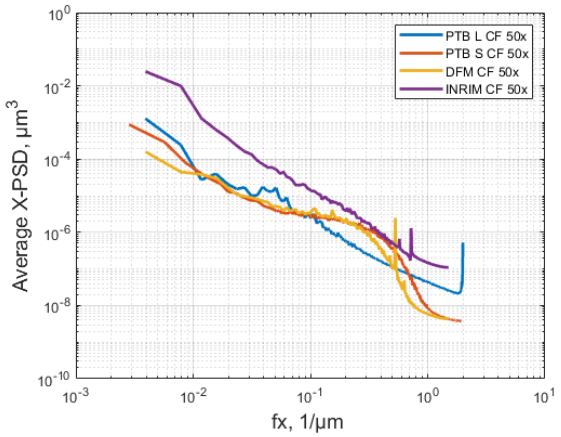


(e) VTT MIKES VSI

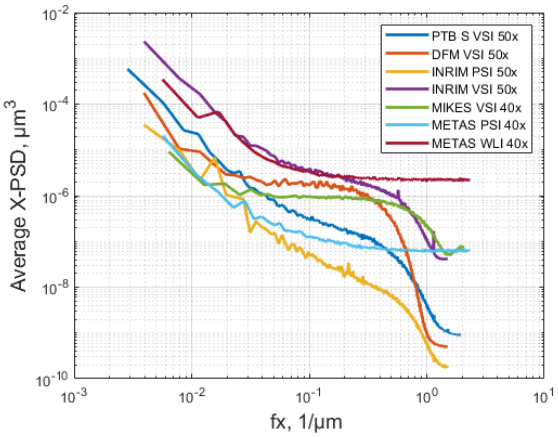


(f) METAS PSI / WLI

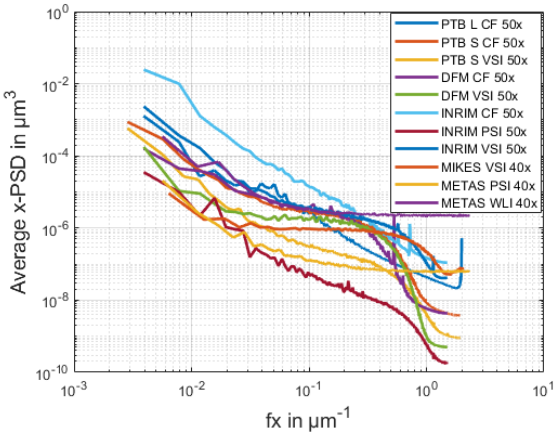
Figure 14 (a) Averaged x-PSDs of the silicon flat sample FtS, measured by participants with different objectives, evaluated by PTB



All 50x CF



All 40x or 50x PSI/VSI/WLI



All 40x or 50x CF/PSI/CSI/WLI

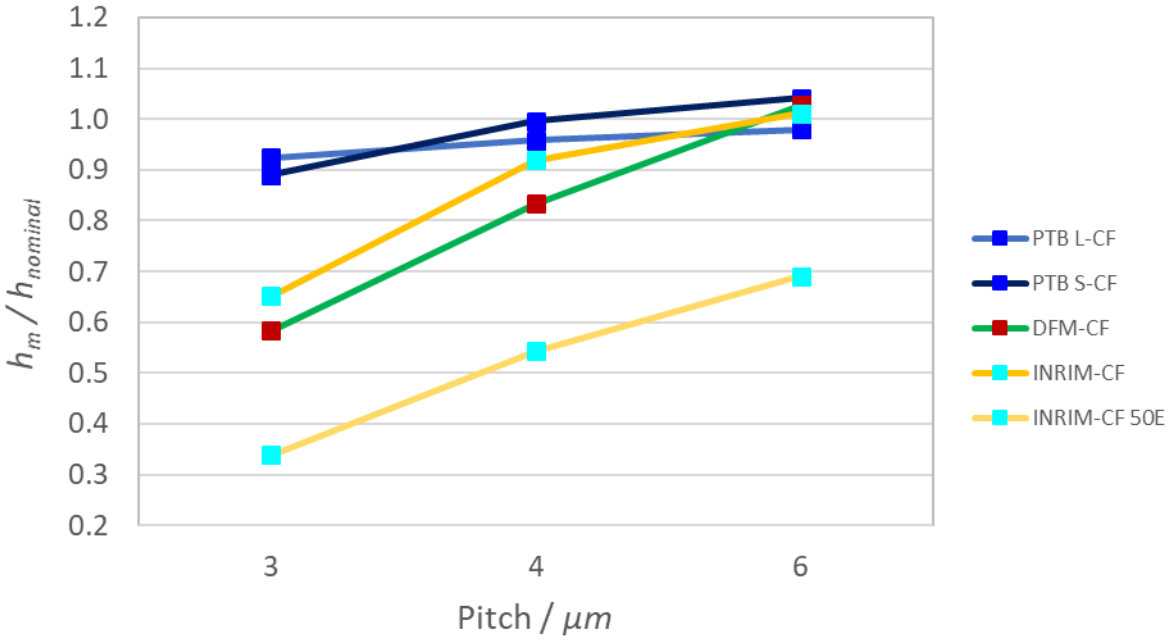
Figure 14 (b) Averaged x-PSDs of the silicon flat sample FtS, measured by participants with 50x or 40x objectives used in this comparison, evaluated by PTB

Artefact No.3: Pitch RS-N

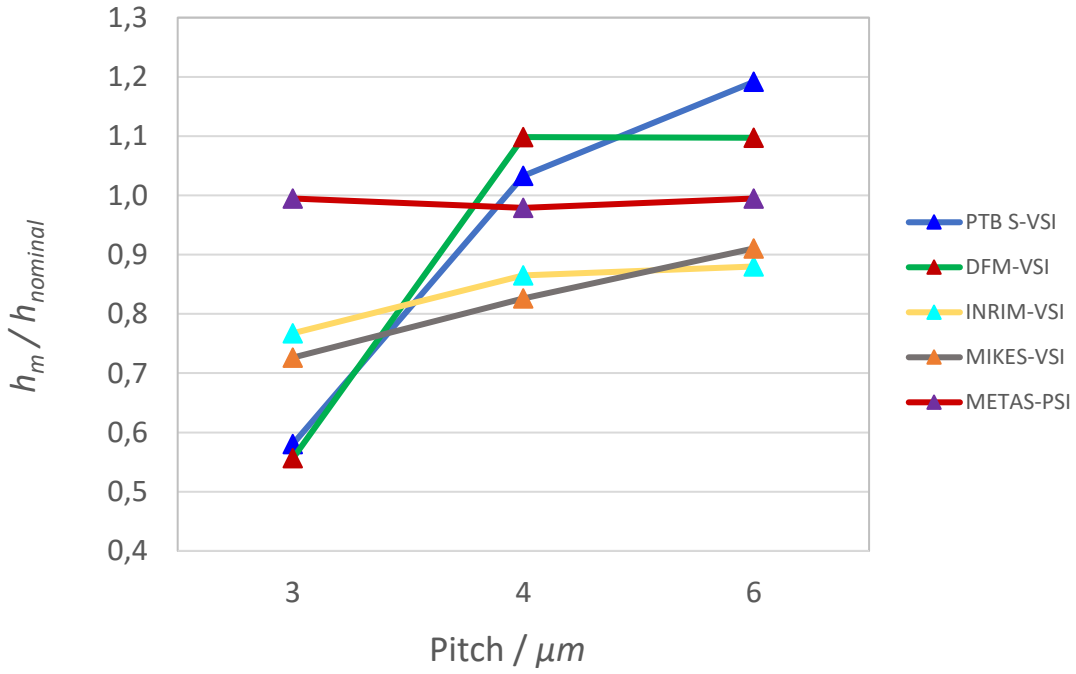
Table 19. Pitch RS-N (artefact No. 3) - $h_m/h_{nominal}$: evaluated by participants

pitch	3 µm		4 µm		6 µm	
	h_m	$h_m/h_{nom.}$	h_m	$h_m/h_{nom.}$	h_m	$h_m/h_{nom.}$
	nm		*nm*		*nm*	
confocal						
PTB L CF	175.7	0.92	182.0	0.96	186.2	0.98
PTB S CF	169.2	0.89	189.3	1.00	197.9	1.04
DFM CF	110.9	0.58	158.3	0.83	195.3	1.03
INRIM CF	123.9	0.65	174.7	0.92	192.2	1.01
INRIM CF 50ELWD	64.3	0.34	103.2	0.54	131.3	0.69
WLI						
PTB S WLI	110.3	0.58	196.3	1.03	226.4	1.19
DFM VSI	105.7	0.56	208.7	1.10	208.5	1.10
INRIM VSI	145.8	0.77	164.3	0.86	167.2	0.88

VTT MIKES VSI	138.0	0.73	157.0	0.83	173.0	0.91
METAS PSI	189.0	0.99	186.0	0.98	189.0	0.99



(a) Measured by confocal microscopes, evaluated by participants.



(b) measured by interference microscopes, evaluated by participants.

Figure 15 Measured height and nominal height (190 nm) rations of the participants for the multipitch resolution sample RS-N.

Table 20. Pitch RS-N (artefact No. 3) - $h_m/h_{nominal}$: evaluated by PTB

pitch	3 µm			4 µm			6 µm		
	h_m	$\frac{h_m}{h_{nom.}}$	$\frac{h_{m,PTB}}{h_{m,Part.}}$	h_m	$\frac{h_m}{h_{nom.}}$	$\frac{h_{m,PTB}}{h_{m,Part.}}$	h_m	$\frac{h_m}{h_{nom.}}$	$\frac{h_{m,PTB}}{h_{m,Part.}}$
	nm			*nm*			*nm*		
PTB L CF	175.7	0.92	1.00	182.0	0.96	1.00	186.2	0.98	1.00
PTB S CF	169.2	0.89	1.00	189.3	1.00	1.00	197.9	1.04	1.00
DFM CF	113.0	0.59	1.02	156.8	0.83	0.99	184.2	0.97	0.94
INRIM CF	133.0	0.70	1.07	178.2	0.94	1.02	195.4	1.03	1.02
INRIM CF 50ELWD	32.2	0.17	0.50	76.9	0.40	0.75	121.4	0.64	0.92
PTB S WLI	110.3	0.58	1.00	196.3	1.03	1.00	226.4	1.19	1.00
INRIM VSI	148.8	0.78	1.02	150.2	0.79	0.91	158.6	0.83	0.95
VTT MIKES VSI	181.9	0.96	1.32	163.4	0.86	1.04	174.2	0.92	1.01
METAS PSI	183.8	0.97	0.97	189.3	1.00	1.02	187.2	0.99	0.99

In the following the results obtained on roughness samples UFRS, ARS-F1 and ARS-F2, and AIR will be shown. As can be seen there are differences in the results depending on the principle of microscope used. Therefore, AFM measurements have been realized to get an independent value, shown by the green dotted lines in the following diagrams, for the samples UFRS, ARS-F1 and ARS-F2. This is NOT a reference value, but it should help to understand differences between different optical microscopes on ARS-F1 and ARS-F2.

Artefact No.4: Roughness 2D-FIB UFRS

Table 21. Roughness 2D-FIB UFRS (artefact No. 4) – *Sq*: Measurement values, standard and expanded uncertainties, E_n values, degrees of equivalence $\boldsymbol{x_i - \bar{x}_w}$, associated expanded uncertainties $\boldsymbol{U(x_i - \bar{x}_w)}$ and contribution to the weighted mean.

2D-FIB UFRS	x_i	$u(x_i)$	$U(x_i)$	$\lvert E_n \rvert$	$(x_i - \bar{x}_w)$	$U(x_i - \bar{x}_w)$	Contribution to ref. val.
	nm	*nm*	*nm*		*nm*	*nm*	
PTB L CF	7.0	1.9	3.8	0.81	3.0	3.7	Y
PTB S CF	5.2	1.7	3.3	0.37	1.2	3.3	Y
PTB S VSI	9.7	1.7	3.3	1.67	5.7	3.4	N
DFM Plµ CF	4.2	1.7	3.4	0.07	0.2	3.3	Y
DFM Plµ VSI	4.5	2.2	4.4	0.12	0.5	4.3	Y
INRIM Plµ CF	4.0	1.2	2.4	0.01	0	2.3	Y
INRIM Plµ PSI	3.3	0.5	1.0	0.97	-0.7	0.7	Y
INRIM Plµ CF 50ELWD	12.4	3.5	6.9	1.21	8.4	6.9	N
VTT MIKES VSI	8.6	1.5	2.9	1.55	4.6	3.0	N
METAS PSI	4.6	0.7	1.4	0.51	0.6	1.2	Y
Mean value	6.4						
Stand. dev.	3.0						
Ref. value	4.0						
Uncertainty	0.7						
PTB-AFM value	9.7						N

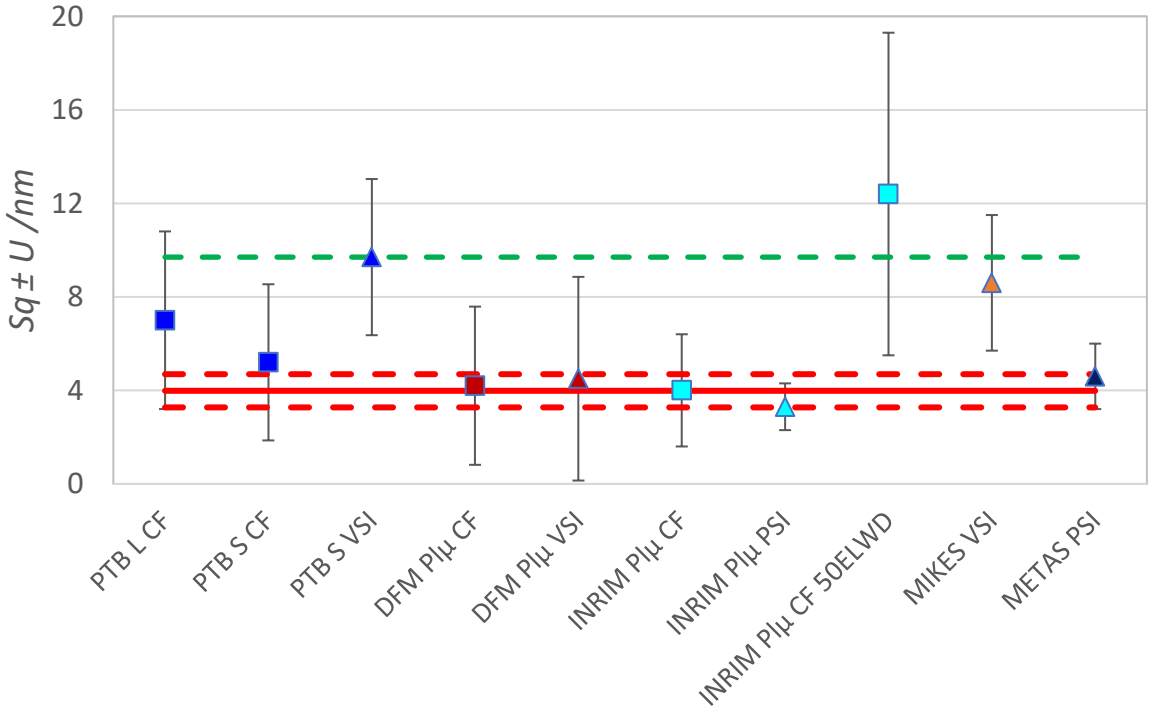


Figure 16. Measured roughness values *Sq* of the participants for the FIB produced 2D roughness pattern (red line: reference value, dotted red lines: expanded uncertainty of the reference value; green line: PTB-AFM value).

Artefact No.5: Roughness ARS F2

Table 22. Roughness ARS-F2 (artefact No. 5) – *Sq*: Measurement values, standard and expanded uncertainties, E_n values*, degrees of equivalence $\boldsymbol{x_i - \bar{x}_w}$, associated expanded uncertainties $\boldsymbol{U(x_i - \bar{x}_w)}$ and contribution to the weighted mean.

ARS F2	x_i	$u(x_i)$	$U(x_i)$	$\lvert E_n \rvert$*	$(x_i - \bar{x}_w)$	$U(x_i - \bar{x}_w)$	Contribution to ref. val.
	nm	*nm*	*nm*		*nm*	*nm*	
PTB L CF	61.7	1.9	3.7	0.3	-1.1	3.6	Y
PTB S CF	37.2	1.7	3.3	8.0	-25.6	3.2	Y
PTB S VSI	95.5	1.9	3.7	9.1	32.7	3.6	Y
DFM Plμ CF	30.6	1.7	3.5	9.6	-32.2	3.3	Y
INRIM Plμ CF	38.4	1.2	2.4	10.9	-24.4	2.2	Y
INRIM Plμ VSI	70.6	0.6	1.3	8.4	7.8	0.9	Y
INRIM Plμ CF 50ELWD	91.7	3.3	6.7	4.4	28.9	6.6	Y
VTT MIKES VSI	103.1	1.5	3.0	14.0	40.3	2.9	Y
METAS PSI	32.0	1.6	3.2	10.0	-30.8	3.1	Y
Mean value	62.3						
Stand. dev.	29.2						
Ref. value	62.8						
Uncertainty	0.9						
PTB-AFM value	58						

only one iteration has been done. The detailed discussion please see section 7.3 “discussion of results”.

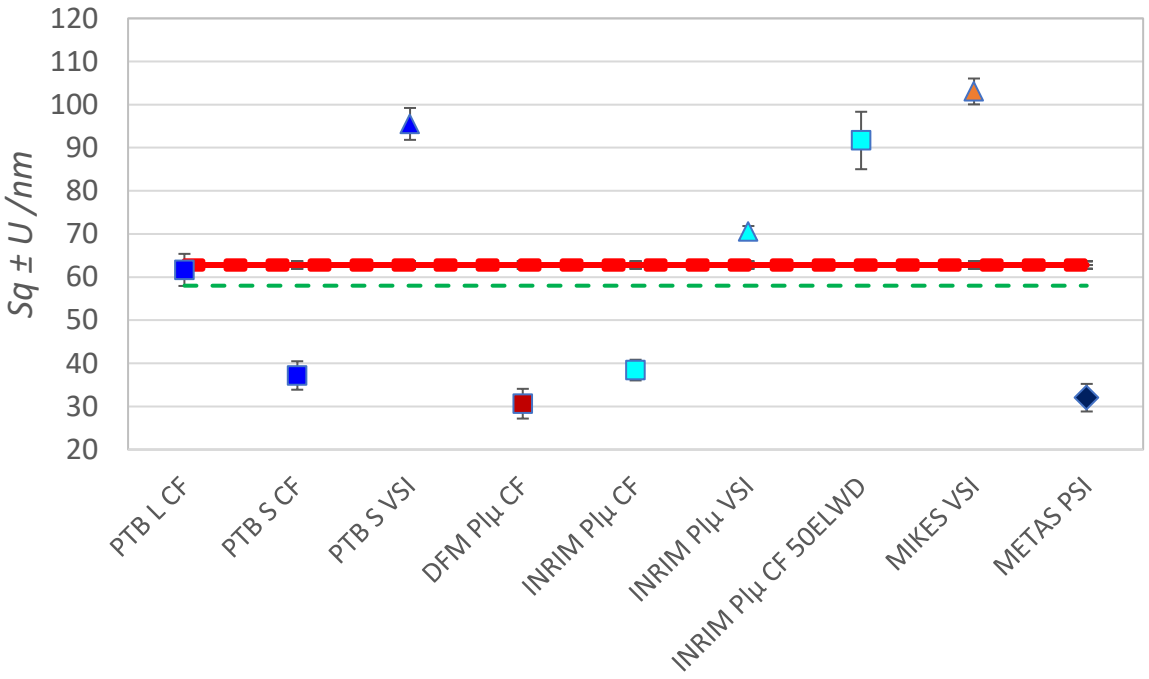


Figure 17 (a). Measured roughness values of the participants for the silicon fine roughness sample ARS F2 (red line: reference value, dotted red lines: expanded uncertainty of the reference value; green line: PTB-AFM value).

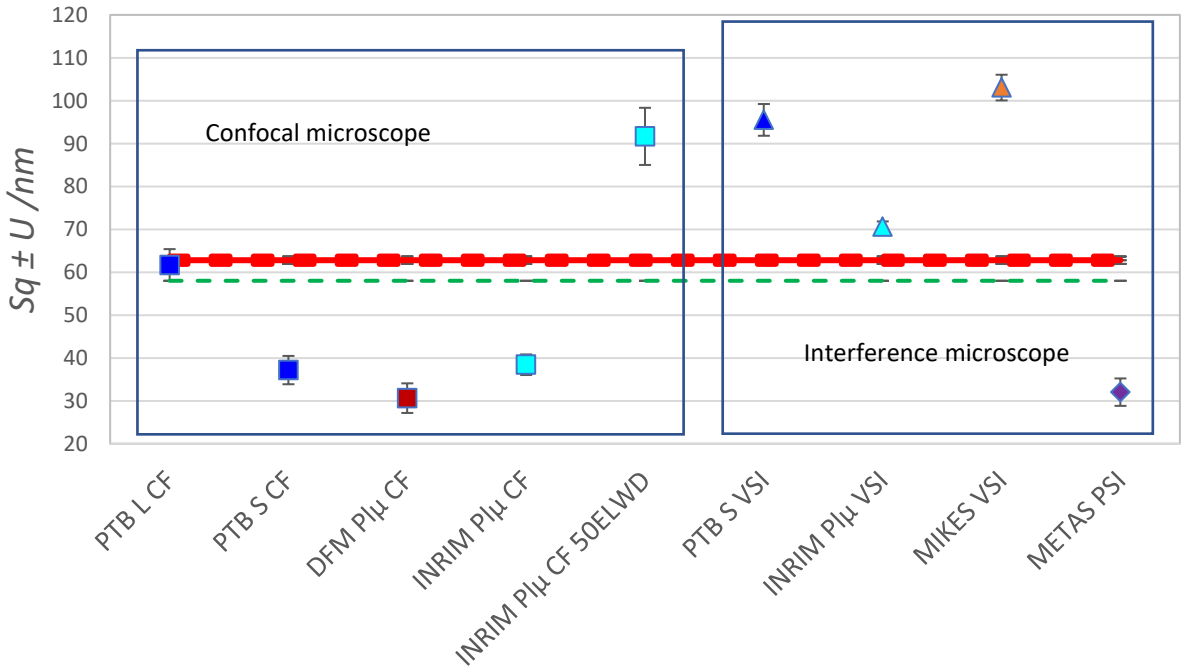


Figure 17 (b). Measured roughness values of the participants for the silicon fine roughness sample ARS F2, separated into two groups: confocal and interference measurement parameters. (red line: reference value, dotted red lines: expanded uncertainty of the reference value; green line: PTB-AFM value).

Artefact No.6: Roughness ARS F1

Table 23. Roughness ARS-F1 (artefact No. 6) – *Sq*: Measurement values, standard and expanded uncertainties, E_n values*, degrees of equivalence $\boldsymbol{x_i - \bar{x}_w}$, associated expanded uncertainties $\boldsymbol{U(x_i - \bar{x}_w)}$ and contribution to the weighted mean.

ARS F1	x_i	$u(x_i)$	$U(x_i)$	$\lvert E_n \rvert$*	$(x_i - \bar{x}_w)$	$U(x_i - \bar{x}_w)$	Contribution to ref. val.
	nm	*nm*	*nm*		*nm*	*nm*	
PTB L CF	109.1	1.9	3.8	0.87	3.2	3.7	Y
PTB S CF	78.0	1.7	3.3	8.78	-28.0	3.2	Y
PTB S VSI	165.3	1.7	3.3	18.66	59.4	3.2	Y
DFM Plμ CF	68.9	1.9	3.8	10.17	-37.0	3.6	Y
INRIM Plμ CF	78.4	1.3	2.5	11.90	-27.6	2.3	Y
INRIM Plμ VSI	104.4	0.7	1.4	1.48	-1.5	1.0	Y
INRIM Plμ CF 50ELWD	138.8	3.3	6.6	5.03	32.8	6.5	Y
VTT MIKES VSI	141.8	1.5	3.0	12.60	35.9	2.8	Y
Mean value	110.6						
Stand. dev.	35.1						
Ref. value	105.9						
Uncertainty	0.9						
PTB-AFM value	103						

only one iteration has been done. The detailed discussion in section 7.3 “discussion of results”.

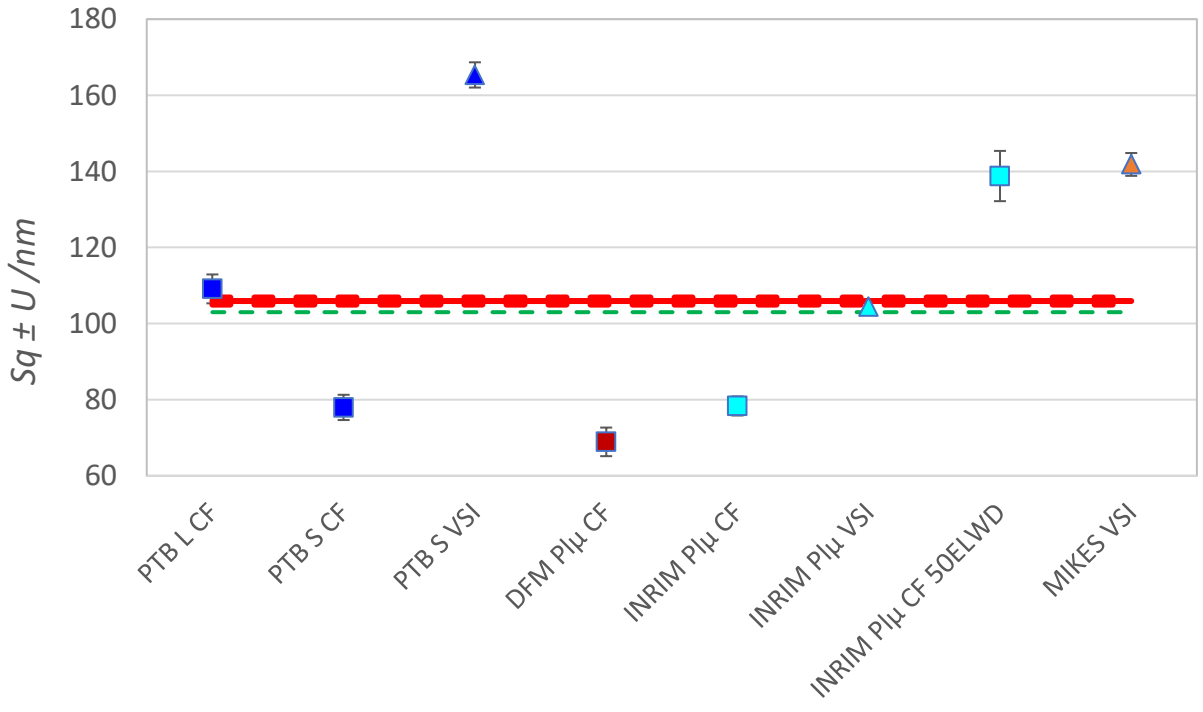


Figure 18. (a) Measured roughness values of the participants for the silicon fine roughness sample ARS F1 (red line: reference value, dotted red lines: expanded uncertainty of the reference value; green line: PTB-AFM value).

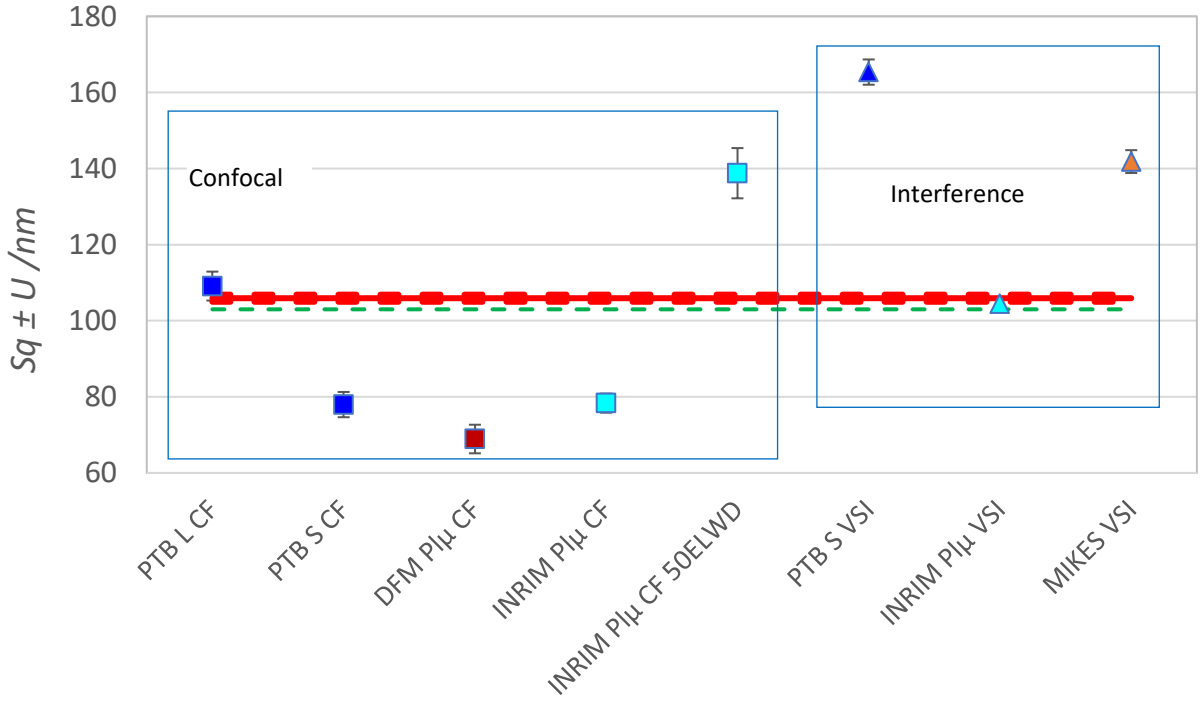


Figure 18. (b) Measured roughness values of the participants for the silicon fine roughness sample ARS F1, separated into two groups: confocal and interference measurement parameters (red line: reference value, dotted red lines: expanded uncertainty of the reference value; green line: PTB-AFM value).

Artefact No.7: Roughness Rubert AIR-B40

Table 24. Rubert AIR B40 (artefact No. 7) - *Sq*: Measurement values, expanded uncertainties, evaluated by participants

Rubert AIR B40	x_i	$U(x_i)$
	nm	*nm*
PTB L CF	1047.0[a]	56.7
PTB S CF	1023.0[a]	61.0
PTB S VSI	1027.0[a]	72.0
DFM Plμ VSI	596.0[b]	53.6
INRIM Plμ CF*	857.0[c]	30.0
INRIM Plμ CF 50ELWD*	930.0[c]	48.0
METAS WLI	772.0[d]	81

a evaluated by using the first method described in Appendix I in the technical protocol, S filter 1 µm, L filter 0.5 mm
b evaluated by using the second method described in Appendix I in the technical protocol, S filter 1 µm, L filter 0.25 mm
c Stitching with MountainsMap, S filter 1 µm, L filter 0.1 mm
d S filter 1 µm, L filter 0.5 mm

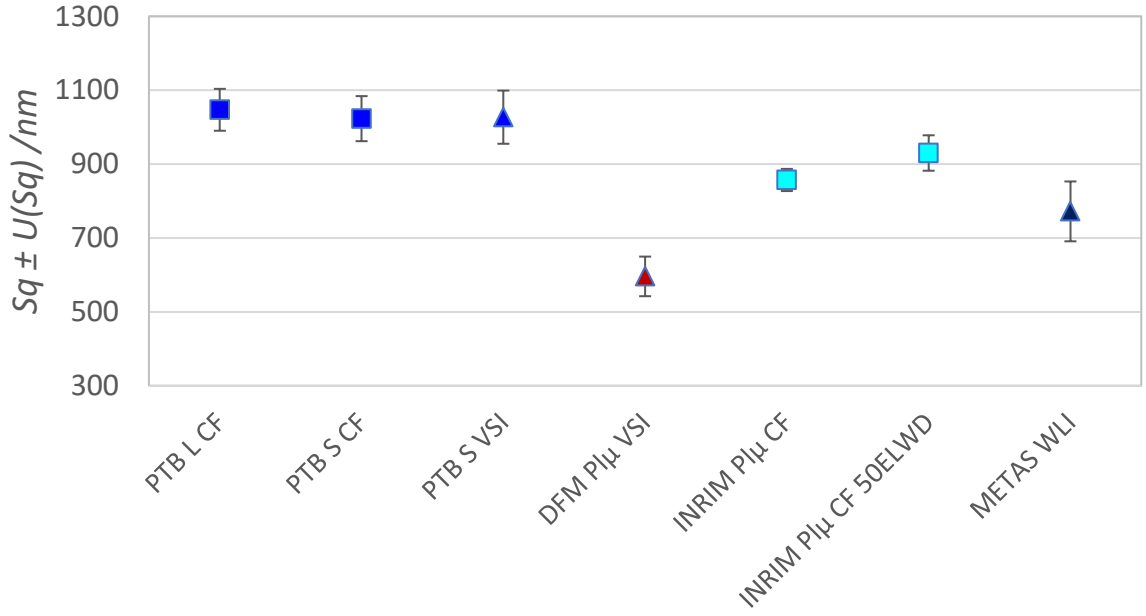


Figure 19 (a). Measured roughness values of the participants from the areal irregular roughness sample AIR B40, evaluated by the participants (red line: reference value, dotted red lines: expanded uncertainty of the reference value).

Table 25. Rubert AIR B40 (artefact No. 7) – *Sq*: Measurement values, standard and expanded uncertainties, E_n values, degrees of equivalence $\boldsymbol{x_i - \bar{x}_w}$, associated expanded uncertainties $\boldsymbol{U(x_i - \bar{x}_w)}$ and contribution to the weighted mean, evaluated by PTB using the first method described in Appendix I in the technical protocol.

Rubert AIR B40	x_i	$u(x_i)$	$U(x_i)$	$\lvert E_n \rvert$	$(x_i - \bar{x}_w)$	$U(x_i - \bar{x}_w)$	Contribution to ref. val.
	nm	*nm*	*nm*		*nm*	*nm*	
PTB L CF	1047	28.4	56.7	0.57	30.1	53.0	Y
PTB S CF	1023	30.5	61.0	0.11	6.1	57.5	Y
PTB S VSI	1027	36	72.0	0.15	10.1	69.0	Y
DFM Plμ VSI	1017	26.8	53.6	0	0.1	49.5	Y
INRIM Plμ CF	1014	15	30	0.13	-2.9	21.9	Y
INRIM Plμ CF 50ELWD	1076	24	48	1.13	59.1	52.2	N
METAS WLI	953	40.5	81	0.82	-63.9	78.4	Y
Mean value	1022.4						
Stand. dev.	37.5						
Ref. value	1016.9						
Uncertainty	20.5						

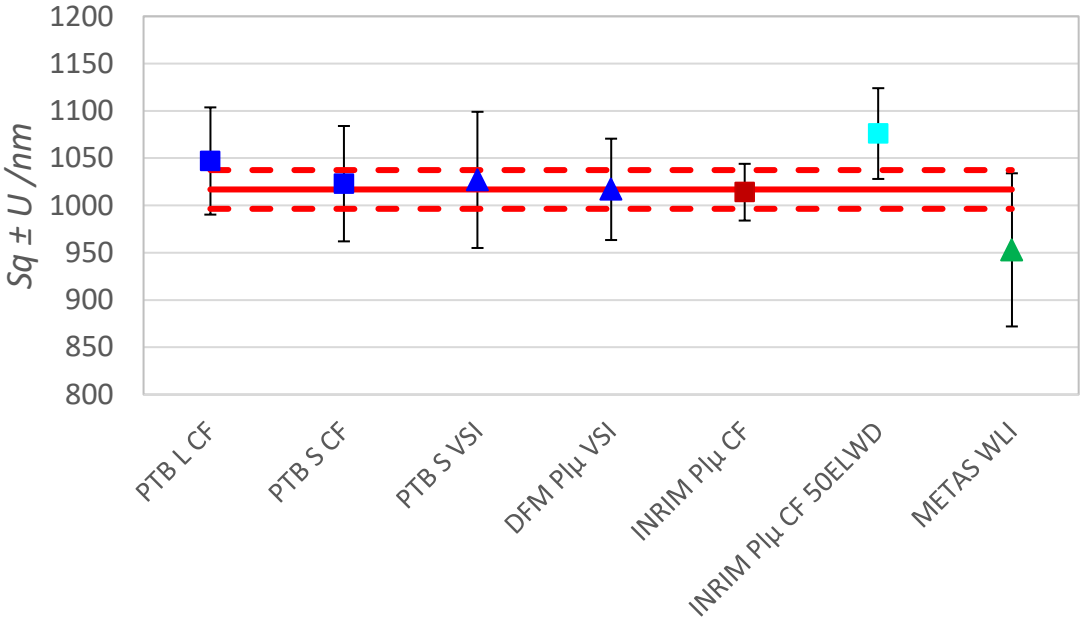


Figure 19 (b). Measured roughness values of the participants from the areal irregular roughness sample AIR B40, evaluated by PTB (red line: reference value, dotted red lines: expanded uncertainty of the reference value).

7.2 Software comparison

The participants have provided some raw data files for most of the samples. These data files have been used to calculate roughness parameters by PTB's software. By comparison with the results reported by the participants, deviations would indicate that there are differences in the algorithms of the various software packages used, or that some analysis procedures/settings have been chosen differently by the individual labs or operators. For the various parameters calculated here we show the results of *Sq* of FIB and ARS F2, only, as an example. Figure 20 (a) and (b) show the differences of the parameter of *Sq* for the FIB produced 2D roughness pattern and for the silicon fine roughness sample ARS F2, as determined and reported by the participants themselves and by the software developed at PTB, respectively.

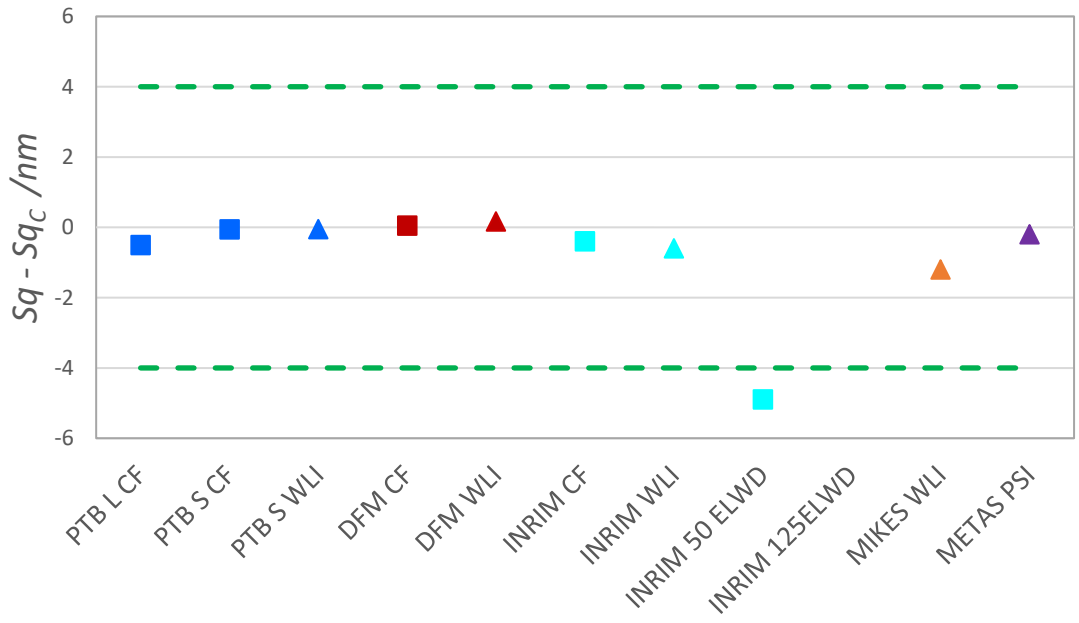


Figure 20 (a). Differences of the *Sq* parameter of the FIB produced 2D roughness pattern evaluated by participants and by PTB's software (Sqc: *Sq* parameters evaluated by PTB's software, green lines indicate the uncertainty range of typical instruments for comparison).

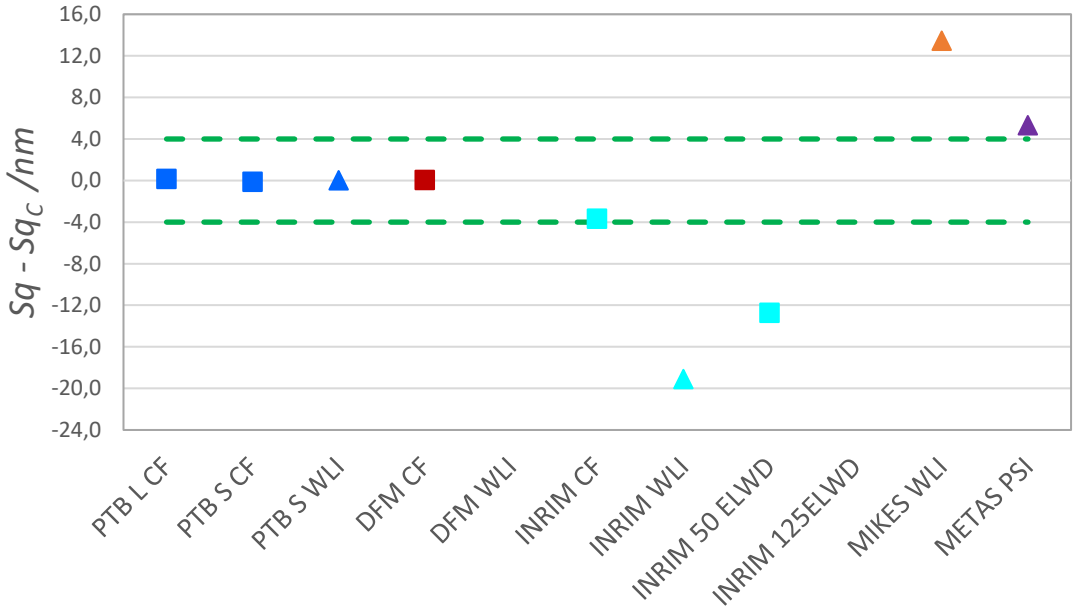


Figure 20 (b). Differences of the *Sq* parameter of the silicon fine roughness sample ARS F2 evaluated by participants and by PTB's software (Sqc: *Sq* parameters evaluated by PTB's software, green lines indicate the uncertainty range of typical instruments for comparison).

For all the parameters we will provide the excel data sheet file with which the participants can check their calculations. If a large deviation occurred, an in-depth analysis is recommended to find differences in software or in the data pre-treatment/procedures/settings chosen.

7.3 Calculation of Degrees of Equivalence

The Degree of Equivalence, DoE, for a laboratory result x_i is calculated simply as $x_i - \bar{x}_w$. The uncertainty of the DoE is calculated using either

$u(x_i - \bar{x}_w) = \sqrt{[u(x_i)]^2 - [u(\bar{x}_w)]^2}$ for results which contributed to the weighted mean

or $u(x_i - \bar{x}_w) = \sqrt{[u(x_i)]^2 + [u(\bar{x}_w)]^2}$ for results which made no contribution. The tables 15-25 show the measurement values, standard and expanded uncertainties, E_n values, degrees of equivalence $x_i - \bar{x}_w$, associated expanded uncertainties $U(x_i - \bar{x}_w)$ and the contribution to the mean value and to the weighted mean for all seven artefacts.

7.4 Discussion of results

1. Step height B1000 (Artefact No.1)

   After INRIM had been notified of an anomaly, INRIM has made a correction of the measurement results, and the step height values measured by PLμ confocal 125x ELWD are withdrawn. After the correction the measurement results of all participants show a good agreement (Figure 12).

2. Multipitch resolution artefact RS-N (Artefact No.3)

   Measurements on the RS-N sample, which has rectangular structures, with the height of approximately 190 nm, pitch values in the range of 300 nm to 6000 nm, show that a strong height deviation from the amplitude of 190 nm already occurs when the pitch is smaller than 3000 nm. This

effect (instrument transfer function) is strongly device- and process-dependent and should be considered more carefully in the instrument characterization, as it has a large effect on topography measurements, especially roughness.

3. FIB produced 2D roughness pattern UFRS (Artefact No.4)

   The calculated roughness parameters Sq, delivered by the participants, show good agreement (Figure 16). However, the roughness parameters Sq, Sa are in a range, where noise and uncertainty are approximately at the same level.

4. Silicon fine roughness artefacts ARS F1 (Artefact No.6) and ARS F2 (Artefact No.5)

   The calculated roughness parameters show measurement differences that are dependent on the instrument/principle. DFM has reported that it is difficult to measure the ARS F1 and ARS F2 with VSI, therefore they reported only the measurement results for the confocal mode. Take ARS F1 as an example, Figure 18 shows the roughness parameter *Sq* of all participants with different instruments or different measurement principles, Figure 21(a) shows the parameters measured by different instruments but all with confocal principle, Figure 21(b) shows the parameters measured by different participants but all with confocal principle from the same manufacturer, Figure 21(c) shows those measured by the different instruments but with white light interference (WLI), or coherence scanning interference (CSI) principles. Better consistency can be seen in Figure 21(a), (b) or (c). Nevertheless, even for the same measurement principle, the $|En|$ values are still (much) larger than 1. The measurement of such finer, i.e higher frequency, roughness by optical surface measurement instruments thus requires further in-depth investigations to achieve a better understanding of dependencies and influencing factors, and to finally allow to better estimate the uncertainty or, in some cases, to conclude which instruments/objectives, etc. do not allow to measure a given roughness adequately.

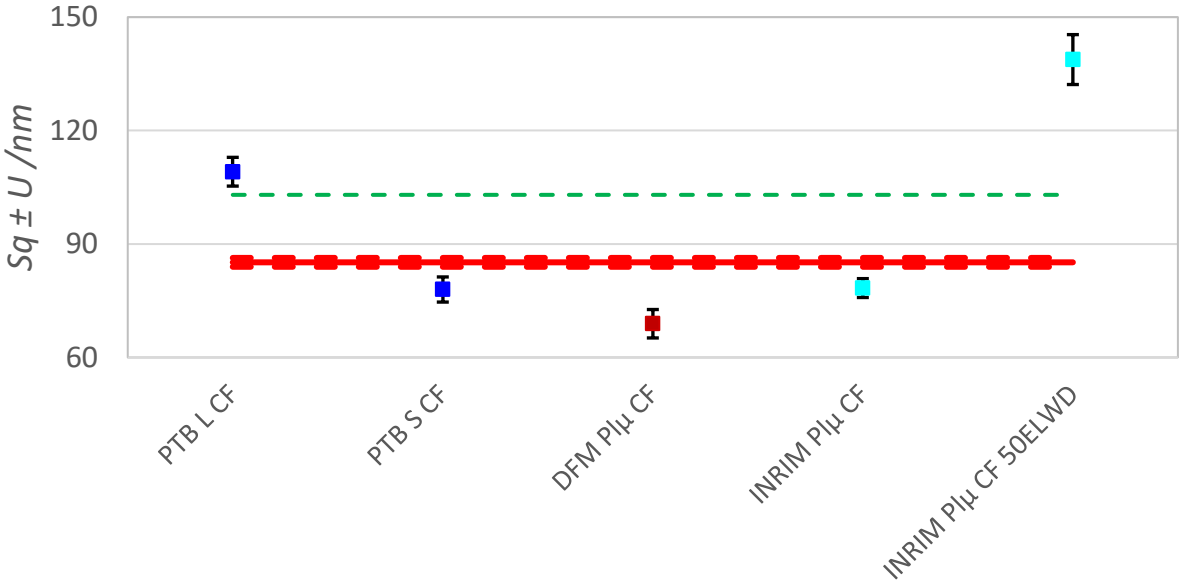


Figure 21 (a) *Sq* measured by all confocal microscopes of the participants for the silicon fine roughness sample ARS F1 (green line: AFM value).

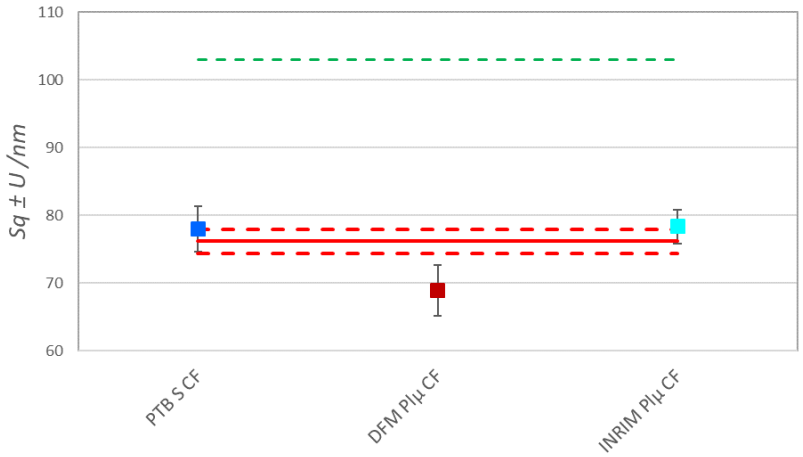


Figure 21 (b) *Sq* measured by confocal microscopes from the same manufacturer and using the same implementation of a confocal mode (red line: reference value, dotted red lines: expanded uncertainty of the reference value, green line: AFM value).

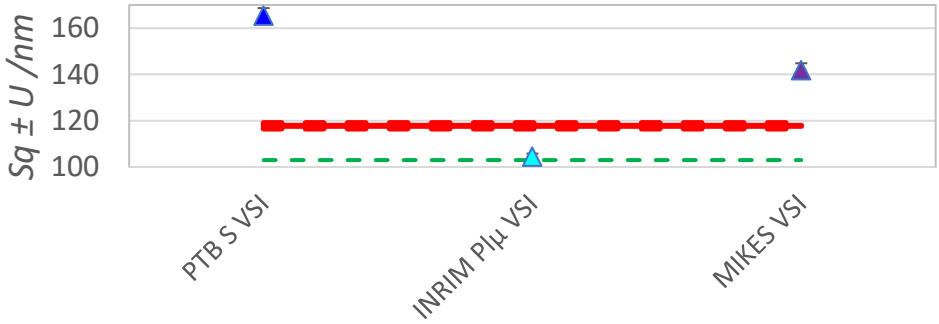


Figure 21 (c) *Sq* measured by interference microscope of the participants for the silicon fine roughness sample ARS F1 (red line: reference value, dotted red lines: expanded uncertainty of the reference value, green line: AFM value).

The first indicator is the measured slope distribution of the samples. These would allow in the easier way to check the quality of the measurements. In the following we show some examples as obtained from the ARS F1 sample.

For the confocal microscope measurements, the numerical aperture of the microscope objective has an important influence on the measured roughness parameters. Measurements with AFM have confirmed the measurements of a high-resolution confocal system (confocal microscope I, with objective 50x, very high numerical aperture $A_N = 0.95$). But measurement results of the other confocal microscopes (confocal microscope II, III and IV) show that with a slightly smaller numerical aperture, much smaller roughness values have been obtained. Calculation shows that the histogram of the measurable slope distributions of such microscopes are much narrower than expected, as shown in Figure 22.

In the pilot laboratory at PTB, a 300 µm diameter ruby sphere has been used as a test object to investigate the maximum measurable surface slopes of the microscope objectives, which were used for the comparison. Measurement results show that the maximum measurable surface slopes of the objective (50x, $A_N = 0.95$) of confocal microscope I is about 64°, while that of the objective (50x, $A_N = 0.8$) of the confocal microscope II is with about 34°, which are much smaller than the theoretical values of 71.8° and 53.1° calculated from the A_N, respectively. Further investigations need to be done to find out the reason for the difference. To understand this effect, it would be necessary to have

more data about the “transfer” from true height values to measured values as function of height and width, i.e. some information about the instrument fidelity on sinusoidal curves. The use of the rectangular pattern indicates a damping (see RS-N sample), but provides no information about slope dependencies, which seems to be much more important.

Also, as the reported uncertainties show, it is totally unclear up to now to reasonably calculate the measurement uncertainty. This means, we need to investigate the instrument properties and implementation of the mode much more in detail, because the results depend on it.

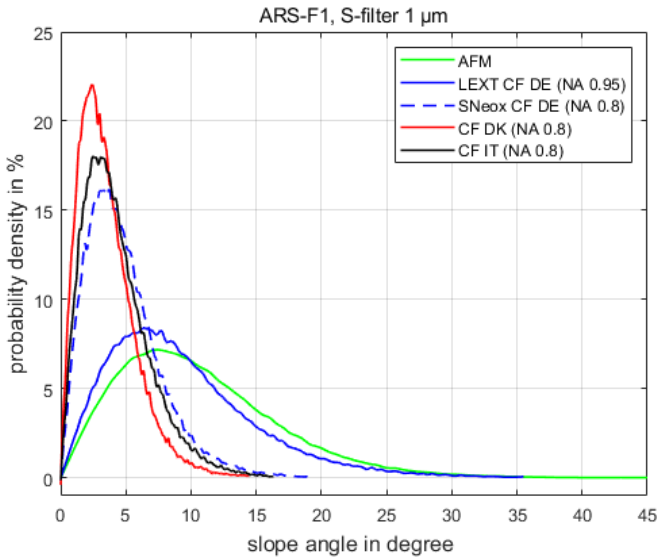


Figure 22 Slope distribution of the sample ARS F1 measured by different confocal microscopes

In the case of confocal instruments, the implementation of the particular confocal mode depends on the manufacturer. For example, one manufacturer uses a matrix camera and additional software filter of approximately 1 µm to realize the confocal mode. The user should have the knowledge about the details of the implementation and their effect on the transfer function. In Figure 21 (b) the data points from PTB, DFM and INRIM of confocal instruments of the same manufacturer are similar, however, there are still unsolved differences. As for SNEOX and Plµ, an additional point has to be taken into detailed consideration. That is the use of an internal shortwave filter at approximately 1 µm, which in turn means that an additional filtering of the “measured data” with the 1 µm filter would not be necessary but would reduce the amplitudes. For example, the *Sq*-values obtained from the ”measured data” without additional S-Filter of 1 µm and with additional S-filter are 39.6 nm (“as measured”) to 37.6 nm (as measured and filtered). Since we have got information for our instrument only, each institute may ask the manufacturer about “internal” used filters or algorithms which could have an impact on measured amplitudes. Similarly, the identification and subsequent elimination of apparent outliers might be an issue.

For the optical instruments operating on the coherent interference principle, the nominal numerical apertures of the objectives used are significantly smaller than those used in the confocal systems, however, as shown in Figure 23, the histograms of the measured slope distribution of the sample ARS F1 show that “larger slopes” have been obtained, even larger than by AFM. That means the instrument has measured slopes which are not present on the sample!

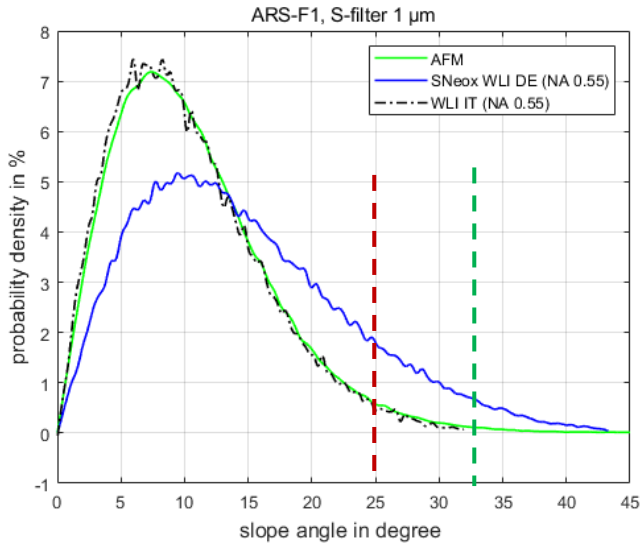


Figure 23 Slope distribution of the same ARS F1 measured by different interference microscopes (green line: theoretical maximum slope that can be measured, red line: the experimentally determined maximum slope investigated by a 300 µm diameter ruby sphere).

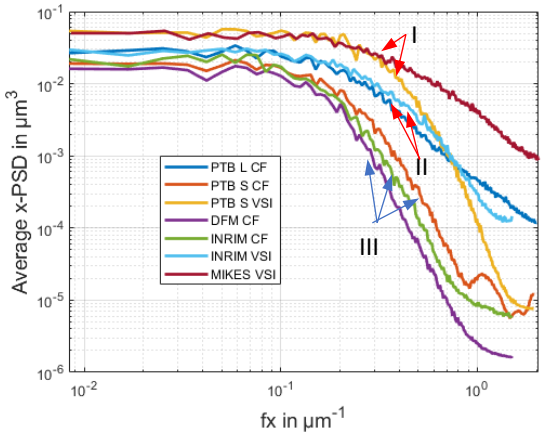
One possible reason is that the locally occurring "bat wings" strongly influence the measured topography and among others, introduce erroneous steep slopes. The outlier issue is consequently much more prominent here than for the confocal mode discussed before. This shows that, unlike in the confocal case discussed above, slope histograms alone do not help to understand which slope angles are actually measured correctly by interference microscopes. Moreover, if angles occur in the slope histogram which are larger than allowed by the numerical aperture, the slope histogram is an indicator that, the measurement is not correct!

The second (SNEOX, DE) and the third (WLI, IT) histograms display a bit wide slope distribution than the AFM slope histogram for ARS-F1. In the case of ARS-F1 only the third histogram (WLI, IT) is similar to that of the AFM. However, only the WLI(IT) value is close to the AFM value for *Sq*. The SNEOX values seems much larger, instead!

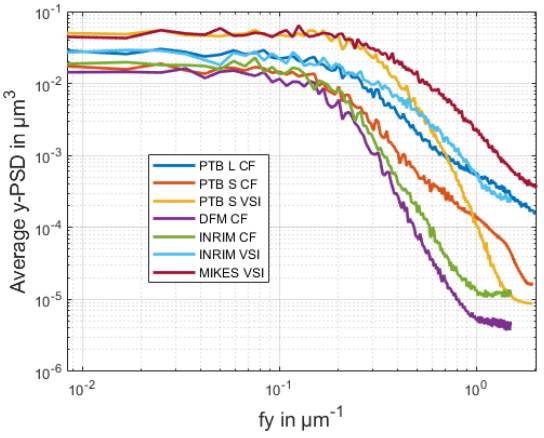
The second indicator of different transfer behaviour of the instruments is the PSD calculated from the measurement result of the sample. Averaged x- and y-PSD diagrams of the first measurement of ARS-F1 (dot1) from participants are shown in Figure 24.

As shown in Figure 24(a) the x-PSD diagrams can be coarsely divided into three groups according to the similarity of the PSDs. Group I: PTB S VSI and VTT MIKES VSI (by $fx < 3\times10^{-1}\,\mu m^{-1}$), whose Sq values are 165 nm and 141 nm respectively; Group II: PTB L CF and INRIM VSI, the corresponding Sq values are 109 nm and 104 nm respectively; Group III: PTB S CF, INRIM CF and DFM CF, the corresponding Sq values are 78 nm, 78 nm and 69 nm respectively.

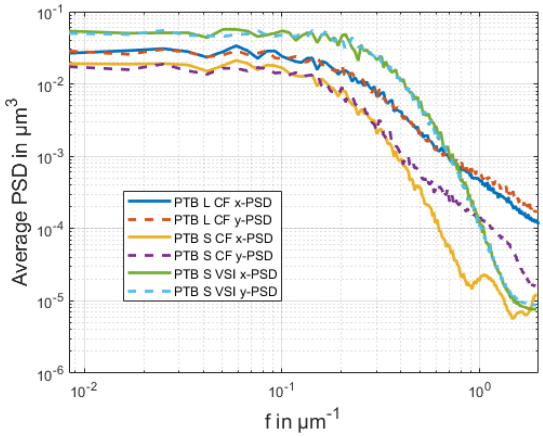
Figure 24 (c) – (e) show the x- and y- PSDs measured by each participant.



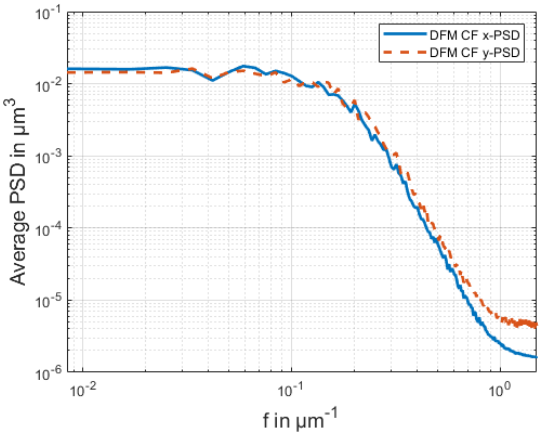
(a) Averaged x-PSDs of the roughness sample ARS F1. The corresponding Sq values are 109 nm, 78 nm, 165 nm, 69 nm, 78 nm, 104 nm and 141 nm respectively, in the same order as listed in the legend.



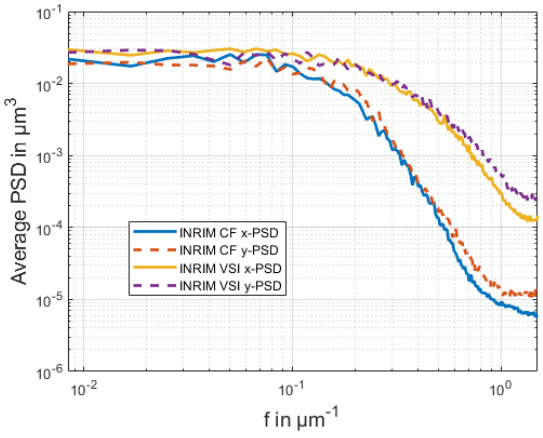
(b) Averaged y-PSDs of the roughness sample ARS F1. The corresponding Sq values are 109 nm, 78 nm, 165 nm, 69 nm, 78 nm, 104 nm and 141 nm respectively, in the same order as listed in the legend.



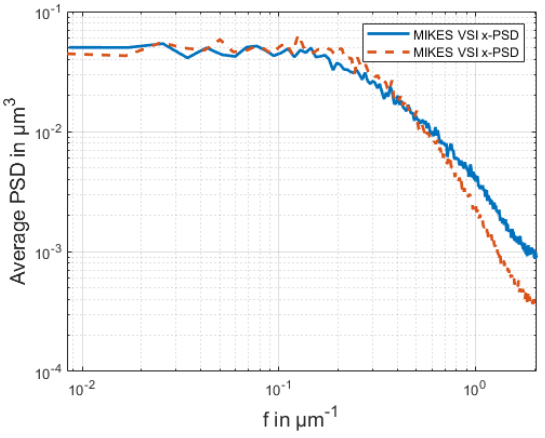
(c) Averaged x- and y-PSDs of the roughness sample ARS F1, measured by PTB



(d) Averaged x- and y-PSDs of the roughness sample ARS F1, measured by DFM



(e) Averaged x- and y-PSDs of the roughness sample ARS F1, measured by INRIM



(f) Averaged x- and y-PSDs of the roughness sample ARS F1, measured by VTT MIKES

Figure 24 Averaged x- and y-PSDs of the roughness sample ARS F1, measured by participants, evaluated by PTB.

5. Areal irregular roughness artefact Rubert AIR B40 (Artefact No.7)

As shown in Figure 19(a), there is no good consistency of the roughness parameters evaluated by the participants. However, evaluation of the raw data delivered by participants using the same evaluation software developed by PTB shows that the measurement results of *Sq* agree with each other very well (as shown in Figure 19(b)).

The reason for the less consistency of the roughness parameters delivered by the participants should lie in the different evaluation algorithms / settings or analysis software chosen by the individual labs. Due to the size of the field to be measured, a stitching technique has been used in the x-direction which may lead to different strategies to evaluate the roughness parameters. Single-image levelling, stitching and stitched-image levelling as well as filtering may have a significant influence on the roughness results, and larger uncertainty contributions might be inherent in this procedure. With regard to the analysis software, for example, INRIM reported that two different procedures were followed to perform the stitching of images measured by Plμ confocal microscope with 50x objective. The first is based on taking independent images to be later stitched by a commercial image analysis soft tool (MountainsMap), the second is based on the stitching performed by the instrument itself, which provides at the end a single stitched image. The evaluated Sq values are 857 nm and 635 nm respectively for the same measured images.

The evaluation algorithm used by PTB (see Appendix I in the technical protocol) works as follows:

(1) stitch all images together
(2) tilt correction of the whole data set
(3) calculate the average x-profile of all lines <z(x)>
(4) apply a *Lc* Filter of *Lc* = 500 µm to the averaged profile yielding the waviness profile <w(x)>
(5) subtract the waviness profile from each of the lines in the image to obtain the roughness
(6) calculate the areal parameters *S*x(Wp) of the roughness data set

According to the comparison results the following steps are suggested:

1. For the artefacts of step height SH B1000 (Artefact No.1), multipitch RS-N (Artefact No.3), 2D-FIB UFRS (Artefact No.4), no further actions are needed
2. For roughness artefact Rubert AIR B40 (Artefact No.7) the participants need to check the algorithms they used. The participants are therefore asked to send a detailed report on the sequence of processing and analysis steps to the pilot.
3. With respect to the samples ARS F1 and ARS F2, a reasonable comparison for confocal systems can be obtained only if the slope distribution of the surface to be measured has been characterized. It can be concluded: 1) if the histograms of the slope distribution of the measured surfaces are similar, the calculated roughness parameters are comparable; 2) the narrower the histogram of the slope distribution of the measured surfaces is, the smaller the calculated roughness parameters can be obtained.
4. For the samples ARS F1 and ARS F2, no reference value will be delivered for the interference instruments, since there is a large divergency of the roughness values due to the overshooting or "transfer of non-detected gradients". It is currently not possible to characterize such devices in a similar way as typically applied for stylus devices. Thus, comparisons are reasonably possible only for samples with small amplitudes (flatness, UFRS-FIB *S*q <= 10 nm) or pure large spatial-wavelength structures (such as AIR-B40). Surfaces with finer, small spatial wavelength structures and large slopes are only conditionally measurable and comparable if certain prerequisites are fulfilled by the instruments.
5. The PSDs have been calculated by the software developed at PTB. The relationship between the PSD and the surface texture parameters needs to be further investigated.

6. A EURAMET follow-up joint research project has been started (EMPIR 20IND07 TracOptic, June 2021 -May 2023) with the aim to enable traceable areal surface texture and dimensional measurements using optical 3D microscopy and optical distance sensors, with a special focus on providing guidelines for the selection of the most suitable instrumentation for a given purpose.

8 (Optional) Appendix A
Equipment and measuring processes of the participants

If a wide range of equipment or techniques has been used, or if participants have deviated from a standard measuring process, include an appendix describing the measurement methods and equipment. The participants can be asked to supply this information verbatim in a format ready for inclusion – often this is requested in the protocol document.

DFM

Make and type of instrument(s):

The microscope used in this report is a combined interference and confocal microscope (Sensofar Plu Neox) manufactured by Sensofar. The microscope is equipped with 6 Nikon objectives, four confocal (5x, 20x, 50x and 150x) and two interferences (DI10 and DI50). The microscope is calibrated annularly using lateral and vertical transfer standards. The calibration procedures are part of DFMs quality system. The 50x confocal objective with a 0.8 numerical aperture and the DI 50 interference objective with a 0.55 numerical aperture are part of the EURAMET 1242 comparison. The measurement was also performed with the 150x objective, which has a NA of 0.95.

Light sources / wavelengths used or traceability path:

λ_{white} = 550nm LED Lightsource interferometric measurements and λ_{Blue} = 460nm for confocal measurements (wavelength of illumination according to manufacturer's specification)

Z-axis Scanning:

Vertical scanning via close loop z-piezo for both confocal and interferometric setup.

Calibration:

The microscope is calibrated laterally and vertically using traceable standards. The vertical calibration is used in the step height measurements.

Environment:

The microscope stand on a passive isolated Thorlabs science desk.

Temperature during measurement 20 to 24 degrees Celsius

INRIM:

Make and type of instrument(s):

Sensofar PLu2300 optical microscope equipped with a MCL moving objective (NANO-F100). The optical microscope stands on a Halcyonics Micro 60 antivibration table, this last placed on an Newport Research Series Plus high performance optical table with laminar flow isolator supports. The microscope makes use of objectives from 5X to about 100X magnification and operates in confocal (intensity), vertical scanning interferometry (VSI) and phase-shifting interferometry (PSI) modes driven by the own-instrument control software (SensoScan). The measurements were performed with the objectives:

- 50X, NA 0.8, WD 1.0 (Nikon);
- 50X DI, NA 0.55, WD 3.4 (Nikon);
- 50X ELWD, NA 0.55, WD 10.1 (Nikon)
- 125X ELWD, NA 0.8 (Leitz)

Light sources / wavelengths used or traceability path:

The optical profiler makes use of a monochromatic LED light source (blue) whose wavelength is calibrated by a high-resolution spectrometer. An interferometer set-up is used for in-situ calibration of the vertical axis of the optical profiler operating with a MCL moving objective (NANO-F100). A retroreflector is mounted on the objective to detect the optical path changes while scanning the vertical movements of the system. As well, the interferometer set-up with a different optical path configuration is used for calibrating the lateral XY axes of the moving stage and sizes of the image/pixel. A 2D grating with a sub-micrometer pitch is imaged by positioning a selected reference line at different XY lateral positions on the image while recording the associated XY pixels of the image and the XY displacements of the moving table, as measured by the interferometer.

Description of measuring technique (including any corrections such as phase correction & platen material, vertical to horizontal corrections, etc.):

Measurements were performed as requested by the technical protocol and with the recommended objectives, namely:

- sample Step Height: confocal , interferometric VSI;
- sample flat FtS: confocal, interferometric PSI;
- sample RS-N: confocal, interferometric PSI;
- sample 2D roughness pattern: confocal, interferometric PSI;
- sample Fine roughness sample type F1: confocal, interferometric VSI;
- sample Fine roughness sample type F2: confocal, interferometric VSI;
- sample Aerial irregular roughness standard: stitching, confocal, interferometric VSI;

No corrections such as phase correction, platen material, & vertical to horizontal were made.

Range of sample temperature during measurements & description of temperature measurement method:

Temperature of the ambient air close to the sample and of the moving stage supporting the sample is taken by a thermometer Systemteknik S1220 operating with PT100 probes.

Range of temperature during measurements:

Ambient air (20 ± 0.2) °C;

Sample support (20 ± 0.2) °C

METAS:

Make and type of instrument(s):

Nikon microscope with Mireau objectives and piezo driven objective z-stage with capacitive feedback. Basler Ace USB sensor module. METAS homebuilt control and evaluation software (LabView).

Light sources / wavelengths used or traceability path:

White light interferometry (WLI): Halogen lamp, grey filter, mean wavelength: 572.8 mm; calibrated capacitive sensor. Traceability: height standards (VLSI and SIM) calibrated at METAS with tactile methods. Magnification by line scales.

Phase shift interferometry (PSI): halogen lamp with red interference filter, wavelength 633.5 nm, traceability: lamp and filter spectrum, aperture corrected wavelength. Validated with SIM height standards, calibrated at METAS with tactile methods (profiler and metrology AFM). Magnification by line scales.

Description of measuring technique (including any corrections such as phase correction & platen material, vertical to horizontal corrections, etc.):

WLI: Tilt adjustment of the sample, camera acquisition time and gain adjustment, acquisition of N images with lambda/8 steps, for each pixel in the stack, calculation of height with largest modulation Data evaluation by export to SPIP.

PSI: Tilt adjustment of the sample, camera acquisition time and gain adjustment, and acquisition of 6 images with lambda/8 steps, calculation of phase, unwrapping of phase, height calculation with aperture corrected wavelength, subtraction of reference plane image.

Range of sample temperature during measurements & description of temperature measurement method:

Material temperature range during all measurement series: (19.95 – 20.35)°C

Temperature sensors: Calibrated ntc temperature probe

PTB I:

Make and type of instrument(s): Olympus Lext – OLS 4100, Software version: 3.1.9

Light sources / wavelengths used or traceability path:

Semiconductor – "LASER" LED, λ = 405nm (wavelength of illumination according to manufacturer's specification)

Description of measuring technique (including any corrections such as phase correction & platen material, vertical to horizontal corrections, etc.):

confocal scanning via MEM Scanner and Galvo – mirror, Magnification: x50 Objective (wavelength corrected) NA = 0,95, 12bit – Gray-level - Photomultiplier as capturing element, Zoomlevel 1.0 = 1024 x 1024 Pixel bei 258um x 258um

Corrections:

Lateral: manufacturer calibration by line scale in lateral axes

Vertical: additional instrument calibration by measuring calibrated step height samples (PTB - 70nm, 300nm & 1000nm)

vertical correction factor c'= 0,99946, based on an average of 40 Measurements on B1000 (C09 R15 N292)

The measurement principle is responsible for the instrument behaviour of different scanning speed for both lateral axes (fast axis x (4 kHz), slow axis y (4 kHz / 1024).

For precise measurements the recording and evaluation of the images is done by several images in 0° and 90° (cw - rotation) orientation. The calculation of the correction factor considered images of both orientations with equal weight.

Range of sample temperature during measurements & description of temperature measurement method:

Temperature capture during the measuring by mercury – thermometer beside the instrument (Amarell +17°C - +35°C; increment = 0,05 °K). Measuring in parts during several days, individual temperature change, captured by hand for each measurement before and after measurement

Step: 20,2°C – 21,0°C;

Flat: 20,9°C – 21,2°C & 20,9°C – 21,8°C (p1-Goniometer);

RSN: 21,5°C-21,7°C & 21,5°C – 22,1°C (90°Rotation);

Roughness F1: 19,5°C – 20,5°C & 20,4°C – 20,9°C (90°Rotation);

Roughness F2: 20,7°C – 21,0°C & 20,8°C – 21,1°C (90°Rotation);

Rubert: 20,8°C – 21,8°C; FIB 20,0°C & 20,0°C – 20,2°C (90°Rotation)

PTB II

Make and type of instrument(s): Sensofar - S-Neox (combined interference and confocal microscope), Software version: 5.3.21

Light sources / wavelengths used or traceability path:

λ_{white} = 550nm LED Lightsource (used for confocal as well interferometric measurement / wavelength of illumination according to manufacturer's specification)

Description of measuring technique (including any corrections such as phase correction & platen material, vertical to horizontal corrections, etc.):

Confocal measurement: confocal scanning via micro-display in star-heaven mode, vertical scanning via close loop z-piezo, distance increment of vertical stack-layer = 1um

Interferometric measurement: vertical scanning via close loop z-piezo, white light mode

Objective and resulting image field size:

Magnification: x50 Objective;

1360 x 1024 Pixel bei 350,88um x 264,19um;

NA = 0,55 (interferometric); NA = 0,80 (confocal)

Corrections:

Reference plane: calibration of objective by individual preference plane correction measurement

Lateral: manufacturer calibration by line scale in lateral axes

Vertical: additional instrument calibration by measuring calibrated step height samples (PTB - 70nm, 300nm & 1000nm)

vertical correction factor for interference measurement cI'= 1,00195 based on an average of 20 Measurements on B1000 (C09 R15 N292)

vertical correction factor for confocal measurement cC'= 1,00016 based on an average of 20 Measurements on B1000 (C09 R15 N292)

These factors are applied for step height measurement only.

Environment:

active vibration table (Accurion Micro60), housing for the instrument

For precise measurements the recording and evaluation of the images is done by several images in 0° and 90° (cw - rotation) orientation. The calculation of the correction factor considered images of both orientations with equal weight.

Range of sample temperature during measurements & description of temperature measurement method:

Temperature aquisition during the measurements by mercury – thermometer besides the instrument (Amarell +17°C - +35°C; increment = 0,05 °K).

Measured in parts during several days → individual temperature inside the microscope housing, captured by hand for each measurement.

Due to presence of the active vibration insulation table inside of the housing, the equilibrium temperature of the system is higher than the temperature outside.

Step height: 25,2°C;

Flat: 24,9°C;

RSN: 25,1° C – 25,2°C;

Roughness F1: 24,9°C – 24,8°C – 24,9°C;

Roughness F2: 24,8°C – 24,9°C;

Rubert: 24,6°C

VTT MIKES:

Make and type of instrument(s): Bruker contour GT-K coherence scanning interferometer

Light sources / wavelengths used or traceability path: White LED light source was used in all measurements in this work.

Traceability came from a step height standard calibrated by AFM.

Description of measuring technique (including any corrections such as phase correction & platen material, vertical to horizontal corrections, etc.):

The device scans the objective to sample distance while taking several images of the sample. As the interference fringes created by the Mirau objective run across the sample a white light interferograms are recorded across the field of view. Sample height is determined by the position of the peak of the envelope function of the interferogram for each pixel individually. Height information for the images comes from the scanner moving the objective. Height scanner was calibrated by measuring 5 µm step height standard at several heights around the height used for this measurement. The heights for measurements of the step height standard were selected to eliminate periodic non-linearity of the scanner scale (periodicity 10 µm).

Range of sample temperature during measurements & description of temperature measurement method:

Measurement room temperature is monitored automatically using P1000 sensors. The samples were left untouched for long enough time to ensure the temperature is stabilized close to the room temperature. The sample temperature was within 20.0 °C - 20.5 °C.

Some more information on the instruments used in the comparison are listed in the table 26 – table 29.

Table 26: Instrument used and selected specifications

Institute	Instrument	Mode	Objective	N_A	Field of view		Pixel size	
					x-size /µm	y-size /µm	dx /nm	dy /nm
PTB	Lext	CF	50x	0.95	258	258	252	252
PTB	SNeox	CF.	50x	0.8	350.9	264.2	258	258
PTB	SNeox	VSI	50x	0.55	350.9	264.2	258	258
DFM	PLµ Neox	CF	50x	0.8	254.6	190.9	332	331
DFM	PLµ Neox	VSI	50x DI	0.55	254.6	190.9	332	331
INRIM	PLµ 2300	CF	50x	0.8	254.6	190.9	332	331
INRIM	PLµ 2300	VSI	50x DI	0.55	254.6	190.9	332	331
INRIM	PLµ 2300	CF	50x ELWD	0.55	254.6	190.9	332	331
INRIM	PLµ 2300	CF	125x ELWD	0.8	127.3	95.5	166	166
VNIIMS	CCI 6000	CSI	50x	0.4	360	360	700	700
VTT MIKES	Bruker Contour GT-K	VSI	20x 2x	0.4	155.6	117.4	245	245
METAS	Home built	PSI/WLI	40x	0.5	179	136	220	220

Table 27. Part of ISO standard 25178-60x related to areal measuring instruments

ISO 25178	**Standard**	**Status @May 2023**
Part 600	Metrological characteristics for areal-topography measuring methods	published 2019-3-1
Part 602	Nominal characteristics of non-contact (confocal chromatic probe) instruments [1)]	to be revised, 2022-10-4
Part 603	Nominal characteristics of non-contact (phase shifting interferometric microscopy) instruments [1)]	to be revised, 2022-10-4
Part 604	Nominal characteristics of non-contact (coherence scanning interferometric microscopy) instruments [1)]	to be revised, 2022-10-43
Part 605	Nominal characteristics of non-contact (point autofocus probe) instruments [1)]	to be revised, 2022-10-4
Part 606	Nominal characteristics of non-contact (focus variation microscopy) instruments [1)]	confirmed 2022-12-12
Part 607	Nominal characteristics of non-contact (confocal microscopy) instruments [1)]	published: 2019-03-05

Table 28. Part of ISO standard 25178-70x related to areal measuring instruments

ISO 25178	**Standard**	**Status** **@May 2023**
Part 700	Calibration and verification of metrological characteristics of areal-topography measuring instruments (under development)	Published, 2022-12-21
	Comment to the other 70x	

Table 29. List of metrological characteristics of instruments (ISO 25178-600: 2019)

Metrological characteristic	**Symbol**	**Main potential error along**
Amplification coefficient	α_x, α_y, α_z	*x, y, z*
Linearity deviation	l_x, l_y, l_z	*x, y, z*
Flatness deviation	z_{FLT}	*z*
Measurement noise	N_M	*z*
Topographic spatial resolution	W_R	*z*
x-y mapping deviation	Δ_{PERxy}	*x, y*
Topography fidelity	T_{Fi}	*x, y, z*