POLITECNICO DI TORINO
Repository ISTITUZIONALE

Residual flatness and scale calibration for a point autofocus surface topography measuring instrument

Original
Residual flatness and scale calibration for a point autofocus surface topography measuring instrument / Maculotti, G., Feng, X., Su, R., Galetto, M., Leach, R.. - In: MEASUREMENT SCIENCE & TECHNOLOGY. - ISSN 0957-0233. - STAMPA. - 30:(2019), p. 075005. [10.1088/1361-6501/ab188f]

Availability:
This version is available at: 11583/2743010 since: 2019-07-22T12:20:50Z

Publisher:
IOP Science

Published
DOI:10.1088/1361-6501/ab188f

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IOP postprint/Author's Accepted Manuscript

“This is the accepted manuscript version of an article accepted for publication in MEASUREMENT SCIENCE & TECHNOLOGY. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The Version of Record is available online at http://dx.doi.org/10.1088/1361-6501/ab188f

(Article begins on next page)

28 November 2026

Measurement Science and Technology

PAPER • OPEN ACCESS

Residual flatness and scale calibration for a point autofocus surface topography measuring instrument

To cite this article: Giacomo Maculotti *et al* 2019 *Meas. Sci. Technol.* **30** 075005

View the article online for updates and enhancements.

This content was downloaded from IP address 130.192.6.32 on 12/06/2019 at 08:40

OPEN ACCESS
IOP Publishing
Measurement Science and Technology
Meas. Sci. Technol. **30** (2019) 075005 (15pp)
https://doi.org/10.1088/1361-6501/ab188f

Residual flatness and scale calibration for a point autofocus surface topography measuring instrument

Giacomo Maculotti[1], Xiaobing Feng[2,3], Rong Su[2], Maurizio Galetto[1] and Richard Leach[2,4]

[1] Department of Management and Production Engineering, Politecnico di Torino, Torino, Italy
[2] Manufacturing Metrology Team, Faculty of Engineering, University of Nottingham, Nottingham, United Kingdom
[3] School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, People's Republic of China

E-mail: richard.leach@nottingham.ac.uk

Received 1 February 2019, revised 25 March 2019
Accepted for publication 11 April 2019
Published 4 June 2019

Abstract

Point autofocus instruments are often used for measuring the surface topography of objects with complex geometry. Determining the metrological characteristics of the instrument is key to ensuring a traceable areal surface topography measurement. In this work, several metrological characteristics, as outlined in ISO/FDIS 25178-600, are determined for a commercial point autofocus instrument, including flatness deviation, the amplification and linearity of the lateral and vertical axes, and the perpendicularity between the axes. Calibrated material measures including an optical flat, step heights and areal cross gratings are used to determine the metrological characteristics. The impact of the point autofocus operating principle and the evaluation method on the metrological characteristics is discussed.

Keywords: metrological characteristics, flatness deviation, linearity, axis calibration, point autofocus instrument

(Some figures may appear in colour only in the online journal)

1. Introduction

Accurate surface topography measurement is required to evaluate the mechanical and optical behaviour of engineered surfaces, and is often an integral part of process control [1]. Several technologies are commercially available for areal surface topography measurement, such as contact stylus instruments, confocal microscopes, focus variation microscopes and coherence scanning interferometry (CSI) [2]. The need to standardise the specification of measurement accuracy for the various instruments is being addressed by ISO technical committee 213 working group 16, with the drafting of two new specification standards. ISO/FDIS 25178-600 [3] introduces a general set of metrological characteristics (MCs), defined as 'characteristics that may influence the result of measurement and have an immediate contribution to measurement uncertainty'. These characteristics include measurement noise, flatness deviation, linearity deviation, amplification coefficient, *x*-*y* mapping deviations, topographic spatial resolution and topography fidelity [3]. Whilst part 600 defines the MCs, ISO/WD 25178-700 [4] specifies default material measures and methods to determine their values. Further specifications are given in the ISO 25178-6xx series [5–8] for instruments based on various types of technologies. Note that whilst part 700 is still only at draft stage, part 600 is currently in its final stages of standardisation and the list of MCs is not expected to change.

Determination of MCs is used to accurately characterise the performance of a surface topography measuring

Original content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

[4] Author to whom any correspondence should be addressed.

 © 2019 IOP Publishing Ltd Printed in the UK

instrument, which installs confidence in the measurement result and provides a meaningful and standardised route to compare the performance of different instruments. The infrastructure for the calibration and verification of areal surface topography measuring instruments is reviewed elsewhere [1], including ISO 25178 geometrical product specifications and material measures commonly used to determine MCs. In addition, the operating principles and bandwidth limitations of different types of instruments need to be taken into account when comparing areal topography measurements [1, 9]. de Groot [10] highlighted a lack of consensus in industry and academia alike regarding the determination of vertical resolution, a frequently cited performance indicator, which is closely related to instrument noise. Given that random instrument noise can be reduced by averaging, de Groot proposed that specifying instrument noise along with data acquisition speed would provide a more meaningful way of comparing different instruments, and examples of performance comparison are presented by Leach and Haitjema [11]. Giusca *et al* [12–14] outlined pipelines for evaluating measurement noise, flatness deviation, amplification coefficient, linearity and perpendicularity deviation, and lateral period limit for a contact stylus instrument, a CSI and an imaging confocal microscope, which include the use of a type AFL flat plane, a type ACG cross grid and type ASG star-shaped grooves. Seewig and Eifler [15, 16] developed a calibration sample that includes multiple types of material measure geometries for the calibration of MCs, enabling calibration of all MCs using a single material measure. Alburayt *et al* [17] developed an alternative material measure for the calibration of the lateral scales of focus variation microscopes. The material measure consisted of an array of hemispherical grooves with sufficient surface texture required for a focus variation microscope. Ekberg *et al* [18] proposed a method to determine and correct for the lateral distortion in a CSI using an arbitrary surface and the method is applicable to other wide-field optical imaging systems.

The point autofocus instrument (PAI) is a non-contact instrument that measures areal surface topography through raster scanning, and is often used to measure optics, cutting tools and complex three-dimensional (3D) geometries [19, 20]. MCs relevant to areal surface topography measurement using PAI are introduced in ISO 25178-600 [3]. Unlike image-based measurement technologies, such as CSI and focus variation microscopy, a PAI relies on a focused laser beam to detect the height at a point on the surface by mechanically bringing the surface point into focus with the optical probe (i.e. operating akin to a contact stylus tip). Due to its point focusing operating principle, evaluation of MCs of this type of instrument has not been attempted to the best knowledge of the authors. In a previous paper [21], the authors investigated the measurement noise of a PAI and discussed the influence of environmental disturbance on areal surface topography measurement. The current work further evaluates other MCs of the PAI, including flatness deviation and the amplification, linearity and perpendicularity of the lateral and vertical scales of the instrument as specified by ISO 25178-600, and their contributions to measurement uncertainty. The influence of evaluation methods, such as filtering and outlier removal, on the determined MCs is discussed, as well as aspects of the operating principle that need to be considered when performing the evaluation.

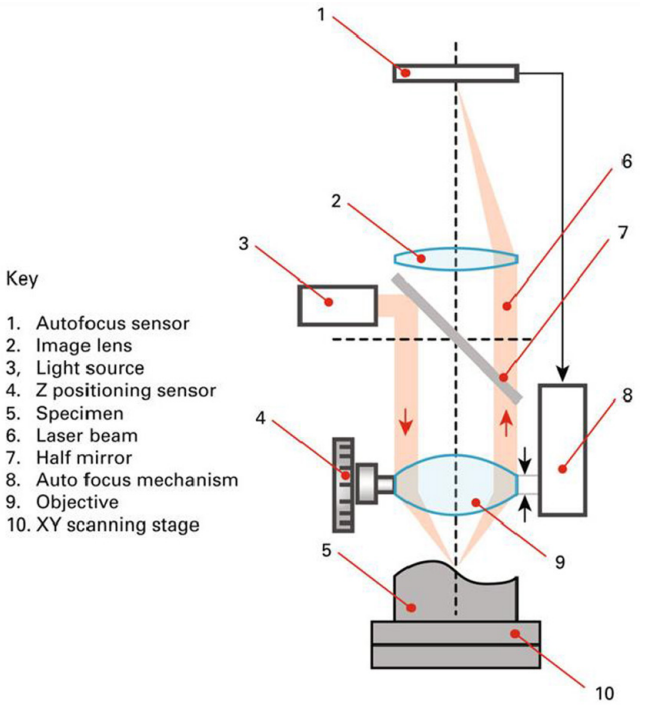

Figure 1. Schema of a point autofocus instrument based on the beam-offset method [22] (2011) (© Springer-Verlag Berlin Heidelberg 2011). With permission of Springer.

The rest of this paper is structured as follows: section 2 explains the working principle of the PAI in order to provide context for discussions in the following sections; section 3 describes the methods used to determine the MCs as specified in ISO 25178-600; section 4 presents the results and discussions; and section 5 draws conclusions of the study.

2. Instrumentation

The instrument being evaluated is a commercial point autofocus surface topography measuring instrument (MLP-3SP, Mitaka-Kohki) housed in the Manufacturing Metrology Team of the University of Nottingham. Although its working principle is described in detail elsewhere [8, 22], this section will highlight a number of the main attributes to provide context for the observations and discussion in section 4.

The PAI has a nominal measurement volume of 120 mm × 120 mm × 40 mm for areal surface topography measurement, enabled by three Cartesian moving axes. Nominal resolutions of the linear encoders are 10 nm for both the lateral scales (i.e. x- and y-axis) and the vertical scale (z-axis) and 1 nm for the autofocus sensor stage (AF). The instrument is equipped with a 100 × objective with numerical aperture (N_A) of 0.8, which focuses a laser beam to a single point on the surface and raster scans an area of interest [8]. The tracing and stepping axes in the raster scan are equivalent to the fast and slow axes in a contact stylus instrument. Autofocus is achieved with the beam-offset method [2], as illustrated in

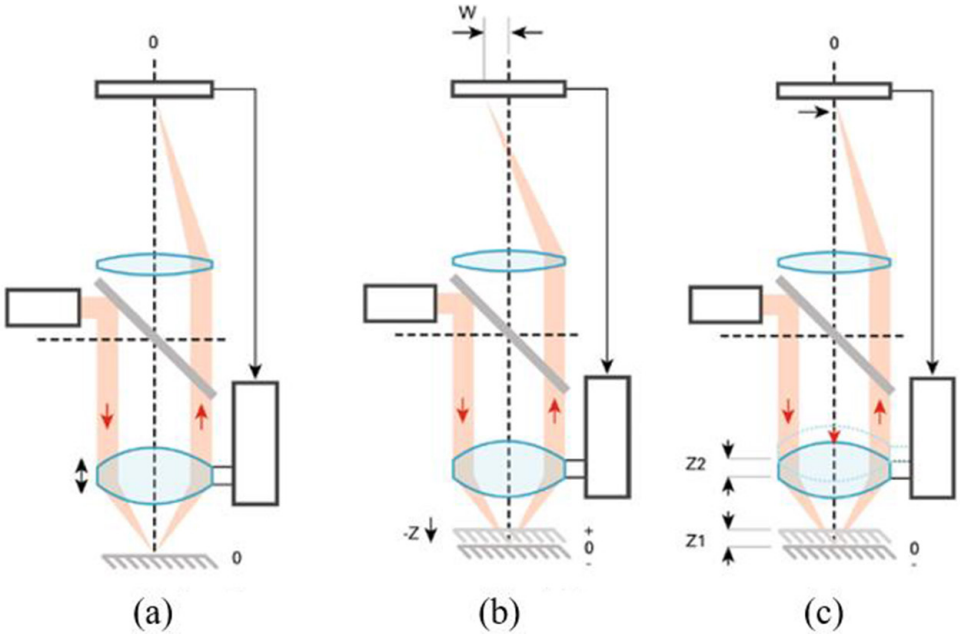

Figure 2. Illustration of a PAI at different sensor positions: (a) in-focus, (b) defocus and (c) autofocus [22] (2011) (© Springer-Verlag Berlin Heidelberg 2011). With permission of Springer.

figure 1. The incident beam is reflected by a half mirror (element 7), to pass through one side of an objective lens (element 9), and is focused on the sample surface (element 5). The reflected beam passes through the opposite side of the objective lens, through an imaging lens (element 2), and is received by the autofocus sensor (element 1). The autofocus sensor detects the laser spot displacement and feeds back the information to the autofocus mechanism (element 8) to automatically bring the surface point to an in-focus position. Figure 2(a) shows the instrument in an in-focus status, where the laser beam returns to the centre of the autofocus sensor. In figure 2(b), when the surface is moved out of the focus position, the laser beam returns to the autofocus sensor at a displacement w from the centre of the sensor. In figure 2(c), the objective lens is automatically moved until $w = 0$, thus effectively bringing the surface point back to an in-focus position. The distance moved by the objective lens is equal to the distance the surface has moved.

As surface height is determined only when the instrument is at an in-focus position, it becomes evident that the laser path length remains unchanged for all measured points. The laser beam passes through each optical component at the same location. As a result, PAI is not subject to field-dependent distortions in the optical components which, when considering MCs, is a fundamental difference from image-based areal surface topography measuring instruments [14, 15].

The PAI also incorporates a drift compensation function to compensate for drifting in the z-axis due to changes in the environmental temperature [21], which is recommended by the manufacturer for areal topography measurement. Compensation is achieved by regularly monitoring the height of a reference point during measurement and compensating the measured topography with the height deviation at the reference point.

3. Methodology

Determination of the MCs was performed using a type AFL flat plane and three type ACG cross grids (part of the calibrated set from the National Physical Laboratory, NPL-BNT 019 [23]) as material measures. The material measures are made of nickel and manufactured by electroforming [24–26]. NPL-BNT 019 includes a type AFL optical flat sample, one type ASG areal star pattern sample, two type AIR irregular samples and three type ACG cross grid samples each with a different nominal step height, i.e. 0.5 μm, 1.2 μm and 2 μm. Each type ACG sample consisted of five grids with pitches of 16 μm, 40 μm, 100 μm, 160 μm, 400 μm, and each grid contains 13 periods. The type AFL flat plane was calibrated to have a Sz value of 4 nm, with an expanded uncertainty of 10 nm ($k = 2$, with a confidence interval of 95%), which can be interpreted to have a lower bound of 0 nm and an upper bound of 14 nm.. The type ACG cross grids were calibrated with an expanded uncertainty ($k = 2$, with a confidence interval of 95%) of 4.1 nm in the step height, and 1.1 μm in the lateral centre-of-gravity (CoG) locations of the grid features, for the grid with pitch 40 μm. When cross grids were involved, the shallowest and largest grid available were used to alleviate edge effects, which are discussed in detail in section 4.2.3. Surface processing was performed using MountainsMap® version 7.4.9.

3.1. Flatness deviation

The flatness deviation z_{FLT} describes the quality of the areal reference of an instrument as the departure from the ideal flat reference, which results in measurement error in the z-axis [3, 26]. In optical surface measuring instruments that are

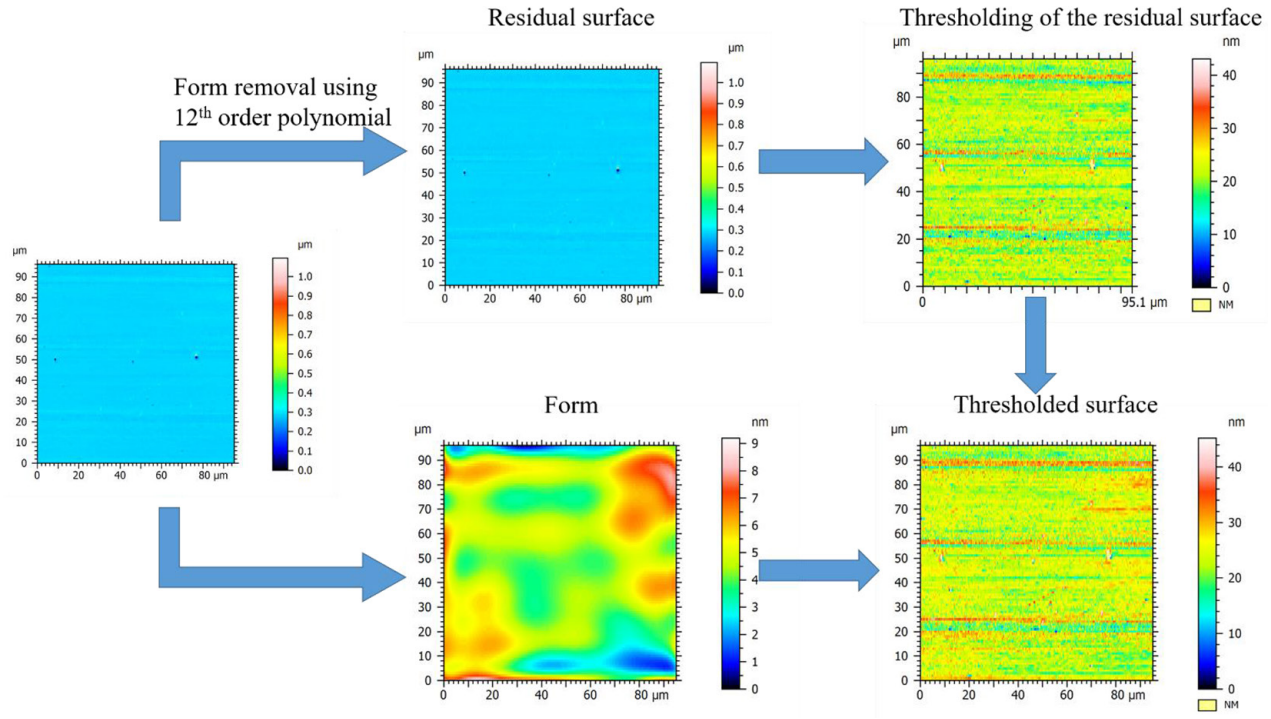

Figure 3. Flow chart for outlier removal during flatness deviation evaluation.

based on wide-field imaging system, the flatness deviation is mainly caused by optical aberrations. The PAI relies on point measurement and raster scanning, of which the motion errors in the lateral stages are expected to be the main source of flatness deviation. Another potential source of error is drift in the z-axis due to variation in the environmental temperature during the measurement. The in-built drift compensation function was enabled to compensate for drift during the evaluation. To assess the effectiveness of compensation, measurement without applying drift compensation was also performed and is discussed in section 4.1.2.

Flatness deviation was evaluated by computing the Sz parameter of the measured topographies on a calibrated type AFL flat plane [2, 12, 27–29]. An area of 100 μm × 100 μm was measured with a sampling interval of 0.1 μm along the tracing axis, and 1 μm along the stepping axis. In order to suppress the influence of the topography of the optical flat and measurement noise on z_{FLT}, ten measurements of different regions of the optical flat were averaged. Additionally, the presence of imperfections, e.g. scratches and dirt particles, on the optical flat was eliminated by an outlier removal procedure illustrated in figure 3. The form of the surface was first removed from the topography using a high order polynomial (following the procedure in [12]). The residual surface was then subject to outlier removal by removing peaks and valleys outside the threshold of three times the Sq of the residual surface. Finally, the residual surface after thresholding was added to the form to obtain the thresholded surface topography. Thresholding is effective if the Sz of the surface resulting from the averaging is smaller than the maximum value of the Sz of the non-averaged surfaces. This condition is necessary for the evaluation process to be stable and produce converging results, and is henceforth referred to as the first stability criterion (SC#1), quantitatively:

$$Sz_{\mathrm{FLT}} < \max\{Sz_i\}, i = 1 \ldots n. \tag{1}$$

where i is the number of topographies used for averaging, Sz_{FLT} is the Sz of the averaged topography after thresholding, and Sz_i is the Sz of the ith topography after thresholding.

In practice, less than ten measurements may be required if the Sz parameter of the averaged topography converges. A second stability criterion (SC#2) is met if Sz_{FLT} does not fluctuate (fluctuation is denoted by $D\%$) by more than 5% when additional topographies are averaged. Once SC#2 is met for three consecutive increments of i, no further repeated measurement will be required.

Once averaging of the thresholded topographies has been carried out, the S-filter, levelling and L-filter are applied prior to computing the Sz of the averaged topography [30, 31]. The Sz parameter of the resulting S-L surface is z_{FLT} [12]. An alternative method that could be applied to evaluate z_{FLT} is proposed by Evans [32], with an amplitude parameter that is robust to outliers on the surface on the surface topography. The choice of filter type and nesting indices will impact on the derived z_{FLT} value. ISO 25178-3 [30] and ISO 25178-2 [31] provide guidelines on applying appropriate filtration. In this work, three filter types were considered: Gaussian filter, robust Gaussian filter, both introduced in ISO 16610-21 [33], and cubic spline filter, introduced in ISO 16610-22 [34]. Both S- and L-nesting indices were applied using the same filter type. ISO 25178-3 recommends that, for an optical surface, the S-nesting index be set at 3:1 ratio with the maximum

Table 1. Nominal dimensions of type ACG samples used for calibrating amplification coefficient and linearity deviation.

ACG	Axis to be calibrated	Step height/μm	Grid pitch/μm
1		0.5	
2	Z	1.2	160
3		2.1	
4	x and y	1.2	40

sampling distance. The L-nesting index was selected based on the S-nesting index and on the size of the coarsest structure on the S-F surface (the F-operator is levelling by subtraction of the least-squares plane) in a ratio of at least 5:1, and the index value was rounded according to [30]. The selection of nesting indices also considers the scale of structures on the measured topography, which is discussed in section 4.1.1.

The uncertainty contribution of the flatness deviation is propagated in the form of a rectangular distribution [35], and is determined using equation (2).

$$u_{\text{flatness}} = \frac{Z_{\text{FLT}}}{\sqrt{12}}. \tag{2}$$

As both flatness deviation and measurement noise are indications of error in the z-axis over a measured region, their contributions to measurement uncertainty are often reported together [12]. In section 4.1.3, the uncertainty contribution of measurement noise determined in the previous study [21] was combined with that of flatness deviation through the propagation of uncertainty [35], according to equation (3). In addition, using an alternative method proposed by Giusca *et al* [12], the overall contribution of measurement noise and flatness deviation to measurement uncertainty can be determined using the same topographies from which Z_{FLT} was determined. This method follows a similar workflow as that illustrated in figure 3, except that the threshold limit is set as two times the Sq of the residual surface instead of three times [12]. Their combined uncertainty contribution, which is considered to propagate in the form of a trapezoidal distribution, can be determined using equation (4).

$$u_{\text{NF}} = \sqrt{u_{\text{noise}}^2 + u_{\text{flatness}}^2}, \tag{3}$$

$$u_{\text{NF}} = \sqrt{\frac{\mu_{Sz}^2}{12} + \frac{\sigma_{Sz}^2}{9}} \tag{4}$$

where μ_{Sz} is the mean of the Sz parameter of the topographies, and σ_{Sz} is the standard deviation of Sz.

The overall contribution of measurement noise and residual flatness to measurement uncertainty was determined using both methods and is discussed in section 4.1.3.

3.2. Amplification coefficient and linearity deviation

The amplification coefficient α and the linearity deviation l are MCs that quantify the difference between an ideal response curve and the actual system response [13, 28]. They can be calibrated on the three Cartesian axes, by measuring calibrated step heights for the z-axis, and cross grid coordinates for the x- and y-axes. In this work, three type ACG cross grid samples were used as material measures, which included both step height and areal cross grid features. The nominal dimensions of the type ACG cross grid samples are listed in table 1. Measurement and evaluation of the step heights were performed according to ISO 5436-1 [35]. Outlier removal (maximum acceptable slope of 30° with the soft removal option in MountainsMap) and levelling (by subtraction of the least-squares plane) were performed on measured topographies.

Table 2. Sz value of averaged and individual topographies.

n	Sz_{FLT}/nm	Sz_i/nm	D%
1	—	371	—
2	350	360	—
3	347	358	0.86
4	345	357	0.58
5	343	370	0.58
6	343	357	0
7	340	341	0.87
8	338	343	0.59
9	337	339	0.30
10	340	387	0.89

3.2.1. z-axis. To account for the reproducibility of the measurement, samples were measured at 10%, 30%, 50%, 70% and 90% of the z-axis range. At each z-axis position, five repeated measurements were performed. Amplification α for the z-axis was determined using equation (5).

$$\alpha_z = \frac{\sum_i M_i C_i}{\sum_i C_i^2} \tag{5}$$

where M is the average step height value derived from repeated measurements performed at five different z-axis height levels, C is the calibrated step height value and subscript $i = 1, 2, 3$ refers to three different step height samples used.

Linearity deviation l was determined as the maximum average residual after correcting the slope of the response curve with amplification α.

3.2.2. x- and y-axis. Calibration of amplification coefficients and linearity deviations for the x- and y-axis was performed similarly, using the calibrated CoG locations of the grid features in the type ACG cross grid samples. As discussed elsewhere [13, 28], repeated measurements at different z-axis levels or using multiple step heights was not necessary. A grid of twelve-by-twelve features with a nominal height of 1.2 μm and pitch of 40 μm was measured, providing an evaluation range of 440 μm per axis. The 40 μm pitch cross grid was chosen as it was the largest grid in which CoG coordinates were calibrated.

The cross grid patterns were segmented using the histogram of the measured surface height. In the histogram, two peaks, separated by the step height of the cross grid, each corresponding to the top and bottom surface, were located.

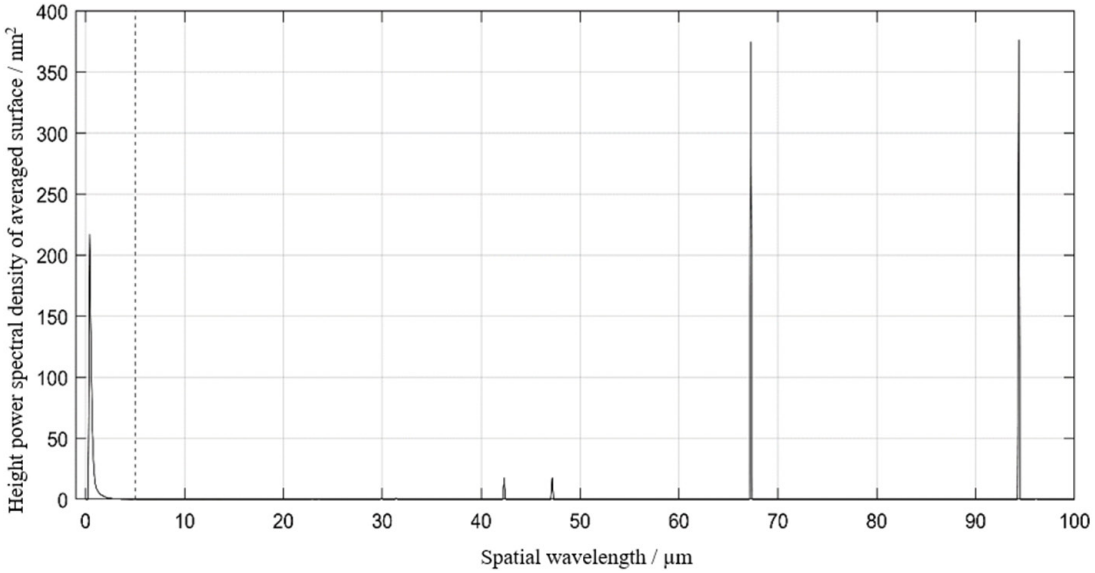

Figure 4. Height power spectral density of the averaged surface (unfiltered). Dashed line indicates the S-filter nesting index is set at 5 µm.

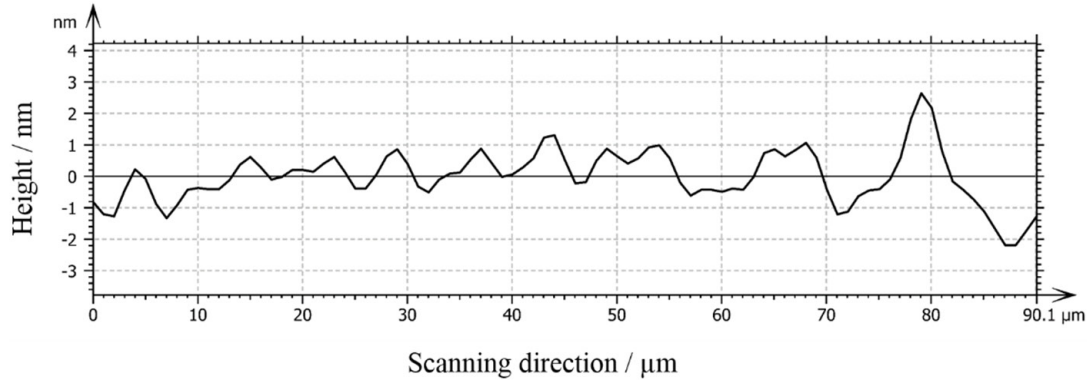

Figure 5. Average profile of the S-F surface. The coarsest structure is around 15 µm.

The midpoint between the two peaks was used as the threshold height to segment the top and bottom surfaces. After segmentation, the missing area (i.e. holes) in the top surface was used to determine the CoGs of the grid features. The top surface was used during segmentation because (i) dust tends to settle in the bottom surface, which is also difficult to clean; and (ii) the edges on the top surface are better measured than the bottom surface due to multiple reflections of light at the bottom surface. The segmentation threshold was further shifted by 0.1 µm from the midpoint towards the top surface to address edge effects.

Calibration of the lateral axis is performed by the determination of mapping deviations, $\Delta x\,(x,y)$ and $\Delta y\,(x,y)$ [3, 4], which characterises lateral distortion in the measured topography. For a point sensing instrument such as the PAI, mapping deviations are mainly caused by the straightness and linearity deviations in the lateral axes of the motion stage. Mapping deviations can also be used to derive amplification coefficients and linear deviations in the x- and y-axis. The overall contribution of α and l along each lateral axis to measurement uncertainty u can be determined with equation(6), accounting for reproducibility,$u_{\text{reproducibility}}$, repeatability, $u_{\text{repeatability}}$, and traceability, u_{t}, which are assumed to distribute according to a normal distribution [13],

$$u = \max\left\{\left(\sqrt{u(E_i)^2 + u_{\text{reproducibility}_i^2} + u_{\text{repeatability}_i^2} + u_{\text{t}_i^2}}\right)\right\} \tag{6}$$

$$u(E_i)^2 = \frac{{E_i}^2}{3} \tag{7}$$

where E is the difference between the measured value M and calibrated value C, namely $\Delta x\,(x, y)$ for the x-axis and $\Delta y\,(x, y)$ for the y-axis; and i refers to three different step heights for the z-axis, and to the twelve rows and columns of grid features for the x- and y-axis. E distributes according to a uniform distribution, from which equation (7) follows [13].

Due to the instrument operating principle and the possibility that the laser path direction may affect measured surface topographies, which in turn affect the determination of α_z and l_z, type ACG cross grids with pitch values of 16 µm, 40 µm and 160 µm were used to investigate this effect, and an explanation of this is provided in section 4.2.3.

The perpendicularity between the x-axis and the y-axis can also be derived from the mapping deviations. Perpendicularity

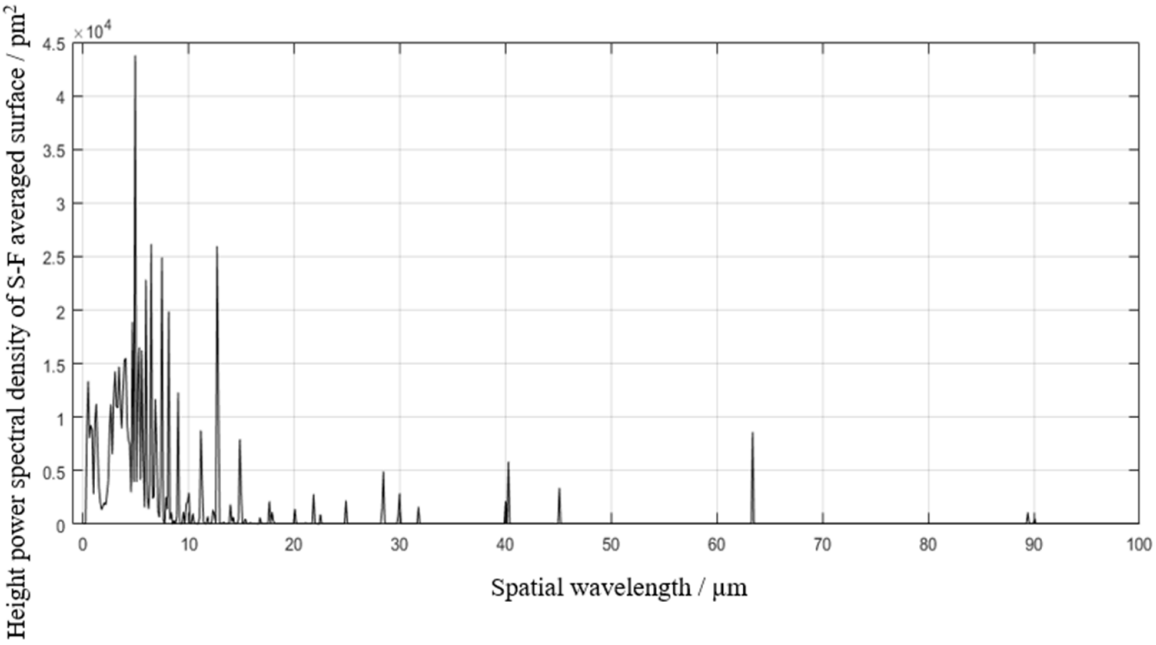

Figure 6. Height power spectral density of the S-F surface.

deviation Δ_{PERxy} has been removed from the list of MCs in the latest draft of ISO/FDIS 25178-600. Given that the PAI relies on the motions in the lateral axes for raster scanning, perpendicularity deviation is also determined and briefly discussed. Δ_{PERxy} was determined by comparing the angle of the intersecting column and row of grid features against the calibrated angle. The orientation of each axis was calculated by fitting a straight line through the CoGs of the column and row of features.

4. Results and discussion

In this section, the MCs are determined, and the influences of the evaluation methods and the operating principle of the instrument are discussed.

4.1. Flatness deviation

The optical flat was measured at ten different locations and the areal surface topographies with 100 μm × 100 μm area were evaluated as shown in table 2. Both stability criteria SC#1 and SC#2 established in section 3.1 were satisfied. Stabilisation was reached almost immediately and, therefore, five measurements would have been sufficient and resulted in a flatness deviation of 9 nm after applying robust Gaussian filter and nesting indices of 5 μm and 80 μm for S- and L-filter, respectively. The choice of filters and nesting indices was compliant with ISO 25178-2, however, further options are available and their effect is assessed in section 4.1.1.

4.1.1. Influence of filtering. According to ISO 25178-2, *Sz* evaluation requires the application of standard filtering [31]. As discussed in section 3.1, several types of filters and nesting indices are available. In this section, three types of filters, i.e. Gaussian filter, spline filter and robust Gaussian filter, were used and a combination of nesting indices were selected based on the scale of the structures on the topography.

Table 3. Flatness deviation comparison for different filter types.

Filter type	S-filter nesting index/μm	L-filter nesting index/μm	z_{FLT}/nm
Gaussian			9
Cubic spline	5	80	20
Robust Gaussian			9

Figure 4 shows the power spectral density of the unfiltered averaged surface, indicating a harmonic peak at approximately 0.4 μm. As a result, the S-filter nesting index was set to be 5 μm to remove this noise component.

The choice of L-filter nesting index depends on the dimension of the coarsest topography structure on the S-F surface, which can be evaluated on the average surface profile, as shown in figure 5. According to the criteria presented in section 3.1, an index value of 80 μm was selected, which was at least five times the coarsest structure size (15 μm wide as shown in figure 5) and took into account the harmonics at larger scales, as shown in figure 6.

Flatness deviation values derived after applying the filters and nesting index values discussed above are listed in table 3. Gaussian and robust Gaussian filters resulted in similar z_{FLT} values, whilst the cubic spline filter systematically produced significantly higher values. The high z_{FLT} values resulted from cubic spline filter may be attributed to badly managed end-effect, which is shown in figure 7. Given the above observations, the robust Gaussian filter, a S-filter nesting index of 5 μm and a L-filter nesting index of 80 μm were chosen. The flatness deviation was found to be 9 nm.

4.1.2. Influence of drift compensation. As discussed in section 3.1, drift of the environmental temperature is also a

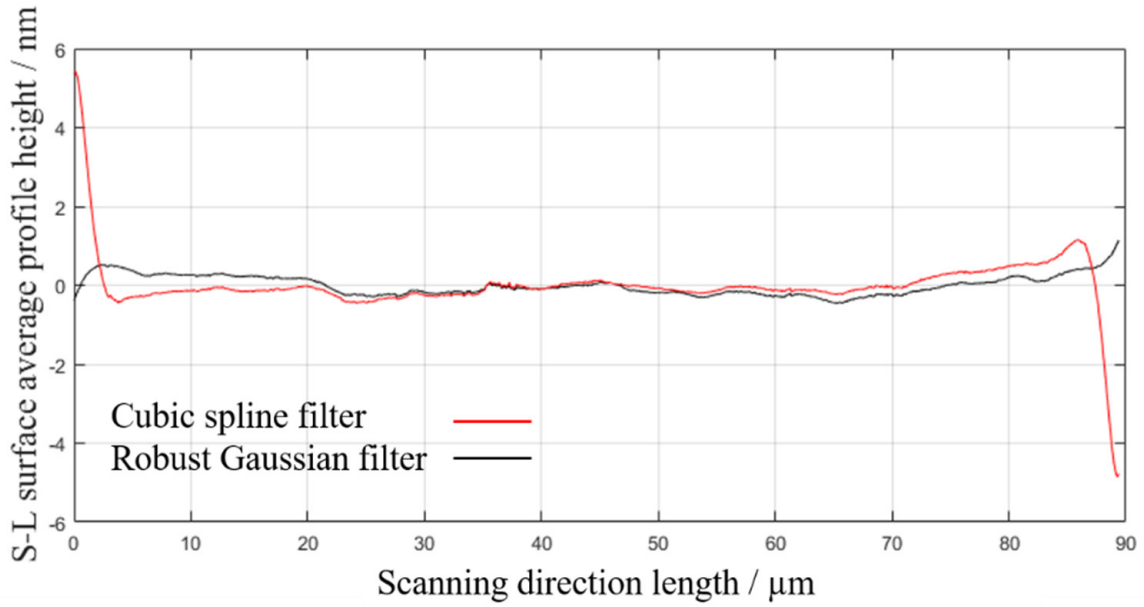

Figure 7. Average profile of the S-L surface after applying robust Gaussian (black) and cubic spline (red) filters.

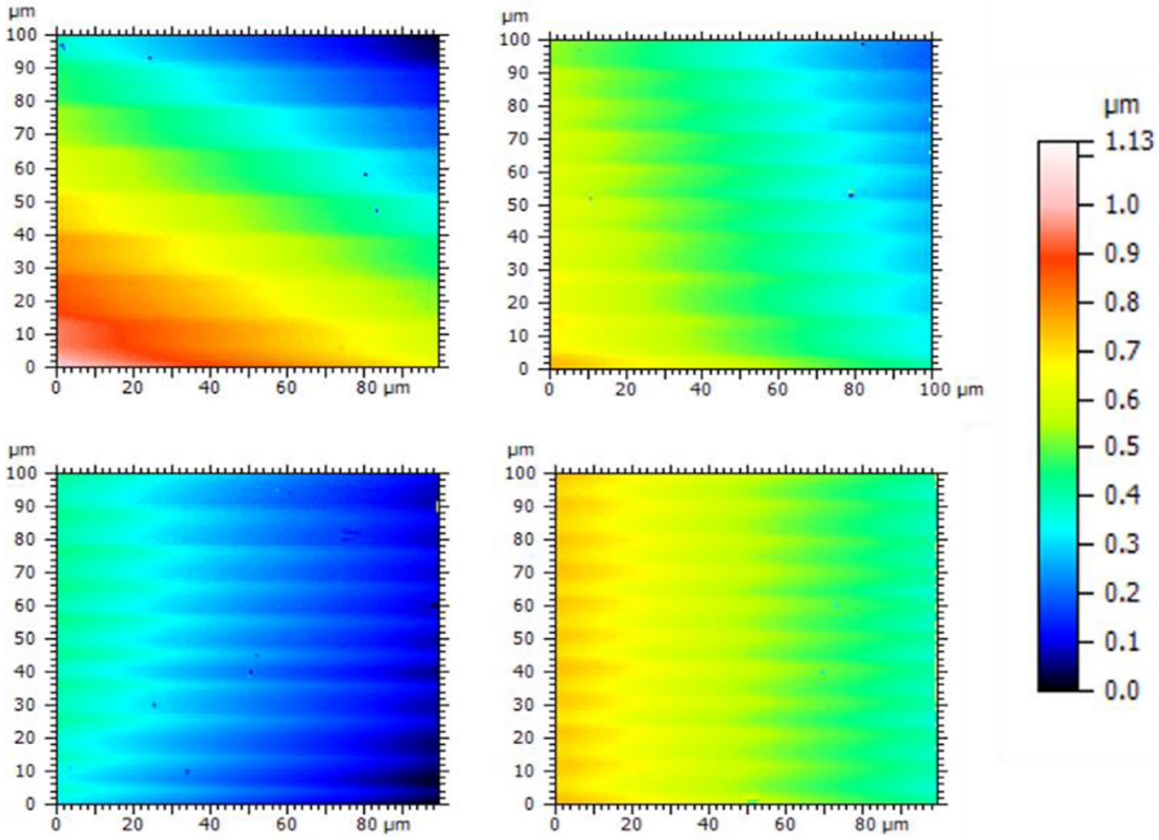

Figure 8. Four measured surface topographies on the optical flat while drift compensation was not applied.

source of flatness deviation due to the relatively long period required to raster scan an area. The application of the in-built drift compensation function has previously been shown to effectively alleviate drift and reduce measurement noise [21]. In this section, the influence of drift and drift compensation is also investigated. Figure 8 shows four measured topographies on the optical flat without applying drift compensation, where each measured topography appeared to have a different form. The differences in the measured topographies were over two orders of magnitude higher than the *Sz* of the optical flat and reflected the rate of variation in the environmental temperature during the measurement. The step like effects in the surface topographies were due to the oscillation of temperature during measurement, which has been reported in a previous study

Table 4. Sz value of averaged and individual topographies without applying drift compensation.

n	Sz_{FLT}/nm	Sz_i/nm	D%
1	—	959	—
2	875	795	—
3	761	548	13.03
4	694	502	8.80
5	702	737	1.15
6	664	487	5.41
7	603	424	9.19
8	568	361	5.80
9	519	539	8.63
10	493	465	5.01

[21]. The peak-to-valley height of such steps was approximately 100 nm when drift compensation was not applied and was reduced by an order of magnitude after applying drift compensation. The Sz values of the measured topographies, which are shown in table 4, were significantly higher when drift compensation was applied (see table 2) and inconsistent among repeated measurements. As a result, the Sz value of the averaged topographies could not converge after ten measurements. Using all ten measurements, flatness deviation was found to be 24 nm, with an uncertainty contribution of 7 nm. The application of drift compensation was able to reduce the flatness deviation by over 60%, and should be enabled whenever possible.

4.1.3. Overall contribution of measurement noise and flatness deviation to uncertainty. With the filtering settings determined in section 4.1.1, the flatness deviation was found to be 9 nm, entailing an uncertainty contribution of 3 nm when evaluated individually according to equation (2). The overall contribution of measurement noise (determined in [21]) and flatness deviation to measurement uncertainty was 4 nm determined through propagation of uncertainty, according to equation (3), and 5 nm determined using the alternative method proposed in equation (4) [12].

4.2. Amplification coefficient and linearity deviation

4.2.1. z-axis. The measured step height values for three sets of material measures were compared to the calibrated values, and the differences are shown in figure 9(a). The amplification coefficient of z-axis, α_z, was found to be 0.9786, suggesting an underestimation of height. Correction of the response curve by the amplification coefficient yielded a linearity deviation of 4 nm, as shown in figure 9(b). Uncertainty contribution u_z, determined according to equation (6), was found to be 24 nm, with its contributors summarised in table 5. If amplification correction were to be applied, uncertainty contribution would have decreased to 8 nm (see residual error and u_z after amplification correction in table 5).

It can be observed in table 5 that the reproducibility contribution was smaller than the repeatability contribution. This is because it was derived as the standard deviation of the mean values measured at each of the five z-axis levels [13], instead of how reproducibility is typically derived (i.e. standard deviation of individual values). Therefore, the term excluded repeatability and was an indication of the homogeneity of the z-axis. Once all 25 individual topographies were included, reproducibility became comparable to repeatability, as shown in table 6. This confirmed that measurements performed at different z-axis positions were highly homogenous.

4.2.2. x- and y-axis. The measured CoG coordinates of a grid of 12 × 12 features were compared with the calibrated coordinates in figure 10, where mapping deviations $\Delta x\,(x, y)$ and $\Delta y\,(x, y)$ are exaggerated for visual clarity (see the scale bar). Maximum deviations in the x- and y-axis were found to be 0.370 μm and 0.596 μm respectively and the root mean square (RMS) deviations were 0.181 μm and 0.287 μm, respectively, which were comparable to the calibration uncertainty of the material measure (0.55 μm, $k = 1$). Another main contributor to the determined mapping deviations, besides the calibration uncertainty of the material measure, was the measurement artefacts in the measured topographies due to the edge effect when measuring surface features with vertical walls, which is discussed in further details in section 4.2.3. The perpendicularity deviation between the lateral axes was 0.025°, which would, in the worst case, result in an absolute length error of 0.133 nm in a measurement area of (1 × 1) mm. Considering that the specified perpendicularity error of the x-y stage itself is an order of magnitude smaller than the perpendicularity deviation determined from the measured topographies, the two above mentioned factors (i.e. calibration uncertainty of the material measure and edge effect) have also contributed to the perpendicularity deviation. Nonetheless, the cosine error due to the perpendicularity deviation was three orders of magnitude smaller than the linearity deviation of the axes, and therefore could be neglected.

Using the mapping deviations, amplification coefficients and linear deviations in the x- and y-axis were derived and shown figures 11 and 12 respectively. Figures 11(b) and 12(b) show the residual errors in the axes after correcting for amplification coefficients. The contributions of amplification coefficient and linear deviation in each axis to measurement uncertainty are summarised in table 7. Amplification coefficients for both lateral axes were very close to unity, indicating a response curve similar to ideal at the investigated scale. The linearity deviation in the lateral axes was two orders of magnitude larger than that in the z-axis. One reason for the differences in the residuals is that the resolution of the linear scale in the z-axis is ten times better than that of the lateral axes. Furthermore, the calibration uncertainty of the CoG coordinates was over two orders of magnitude larger than that of the step height, thus being dominant with respect to other contributions.

4.2.3. Edge effect and the influence of grid height and pitch. As the grid features in the type ACG cross grid samples consist of almost vertical edges, edge effect resulted in

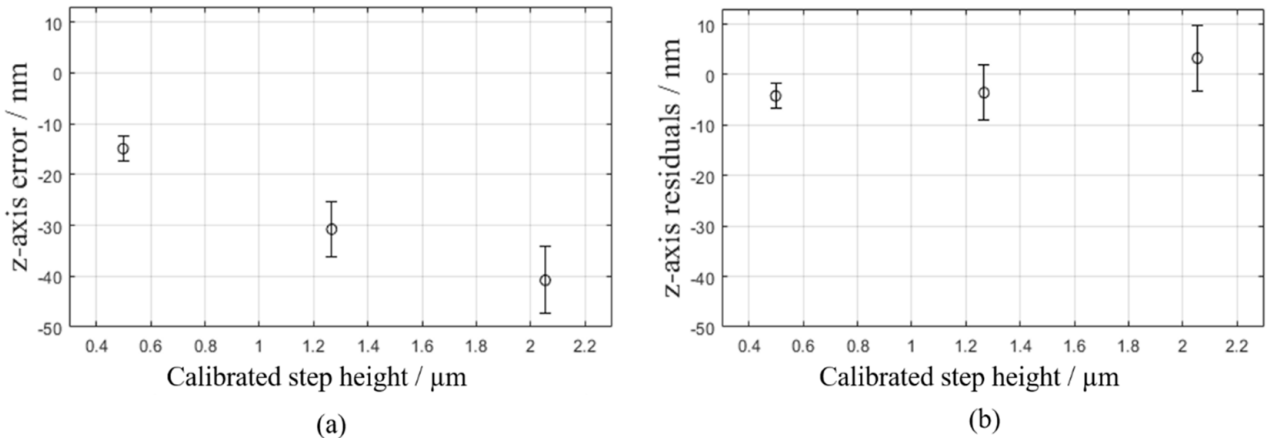

Figure 9. (a) Error plot of vertical response and (b) residuals after amplification correction. Error bars show standard uncertainty including contribution of repeatability, reproducibility and traceability.

Table 5. Summary of uncertainty contributions for u_z. U is expanded uncertainty of depth ($k = 2$). Error, repeatability and reproducibility are truncated according to instrument digitisation step.

	ACG 0.5	ACG 1.2	ACG 2.1
Nominal depth/nm	500	1200	2100
Calibrated depth (C)/nm	498.6	1267	2053.6
Expanded uncertainty of depth calibration (U)/nm	4.1	4.2	4.1
Error (E)/nm	−15	−31	−41
$u_{\text{reproducibility}}$/nm	1	5	5
$u_{\text{repeatability}}$/nm	1	1	4
u_t/nm	2.05	2.10	2.05
u_z/nm	9	19	24
Residual error/nm	3	5	7
u_z after amplification correction/nm	3	6	8

Table 6. Instrument step height measurement reproducibility.

	ACG 0.5	ACG 1.2	ACG 2.1
Reproducibility/nm	1	5	5

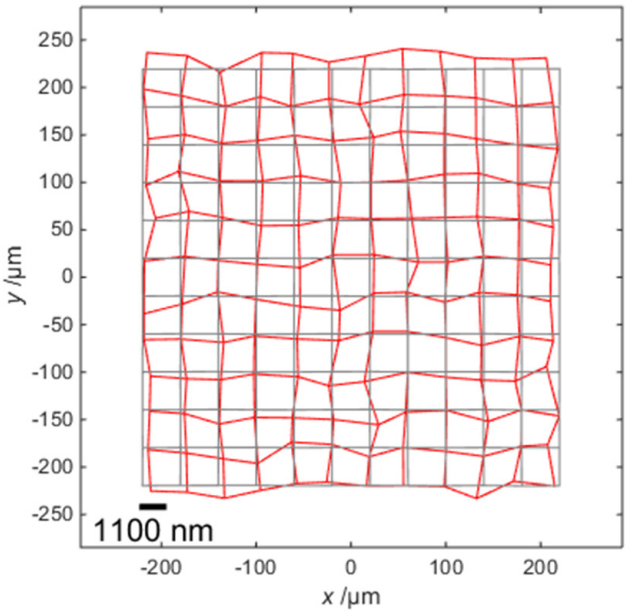

Figure 10. CoG coordinates of measured grid features (red) with Δx (x, y) and Δy (x, y) magnified by 20 times. Calibrated coordinates are shown in grey. Scale bar corresponds to magnified deviations.

measurement artefacts in the form of spikes near the edges, which is a common issue with optical measurement instruments [37–39]. The PAI also suffers from this effect, as shown in figure 13. However, due to the beam-offset operating principle, the edge effect was most severe on one of the four edges, more specifically the edge that was perpendicular to the plane of laser path, where multiple reflection of the laser beam occurred. This is evident in figures 13(a) and (b) where the laser path direction was altered while the scanning direction remained unchanged.

As the probe scans toward a spike-prone edge, a condition occurs which likely leads the instrument to falsely believe that the surface is at a height position different from the real height position, as illustrated in figure 14. In the figure, the instrument scans from left to right towards a vertical edge, with the plane of laser path parallel to the scanning direction. In figure 14(a) the instrument is operating at the normal condition without multiple reflection and the laser beam is returned to the centre of the autofocus sensor, a criterion indicating that the laser beam is focused onto the surface. The bottom surface is correctly measured until multiple reflection occurs on the edge, as shown in figure 14(b). As a result of the additional reflection, the returning laser beam passes through the same side of the objective lens as the incident beam, along an unexpected path. As the laser beam is returned at an offset from the centre of the sensor, the instrument is not in a state of focus. In order to regain focus, the objective lens is

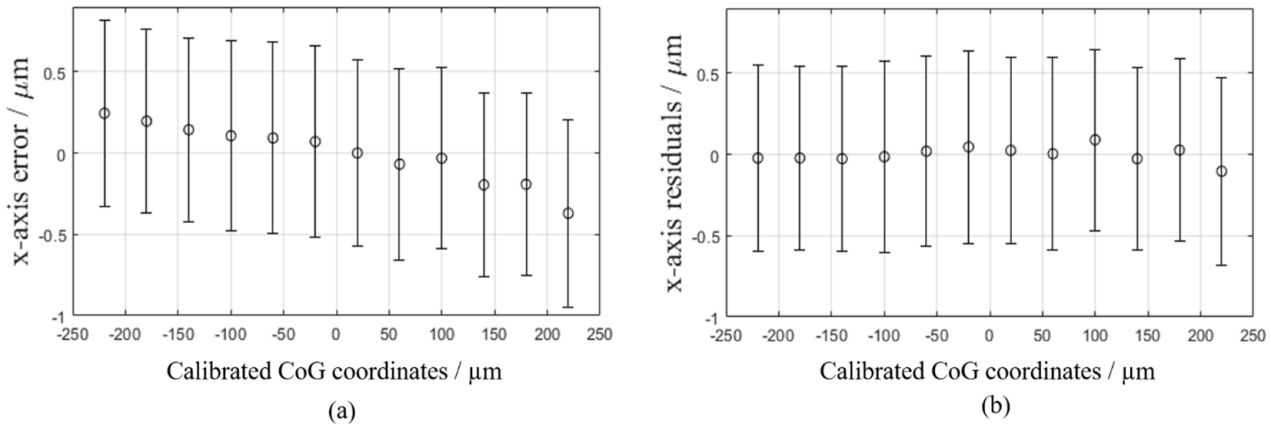

Figure 11. Linear deviation in the *x*-axis: (a) error and (b) residual error after correcting for amplification coefficient. Error bars show uncertainty contribution of reproducibility and traceability.

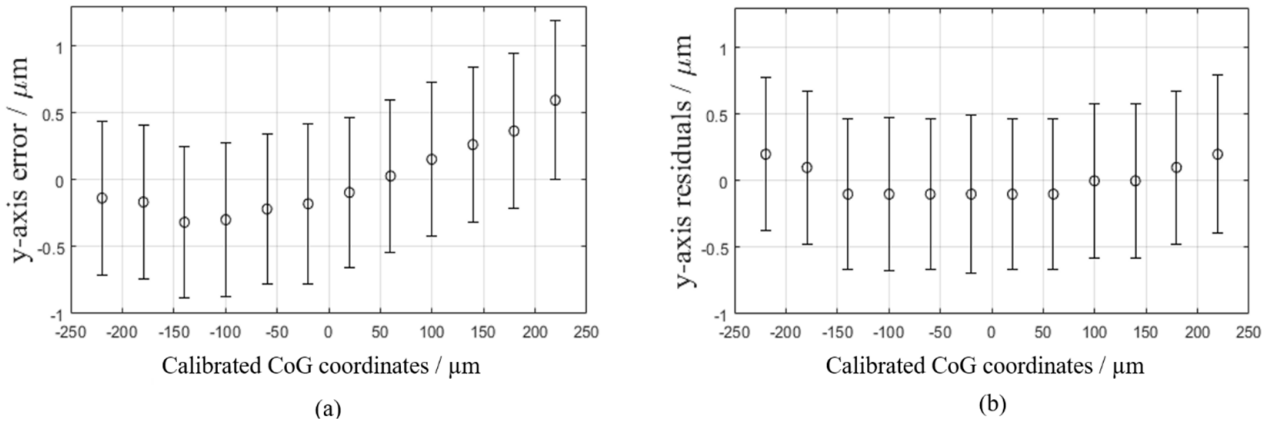

Figure 12. Linear deviation in the *y*-axis: (a) error and (b) residual error after correcting for amplification coefficient. Error bars show uncertainty contribution of reproducibility and traceability.

Table 7. Amplification coefficients, linearity deviations of the *x*- and *y*-axis and their contributions to measurement uncertainty.

	x-axis	*y*-axis
α	0.9988	1.0018
l/nm	103	248
u/μm	0.6	0.6

moved away from the surface until the laser beam is returned to the centre of the sensor again, as shown in figure 14(c). As the additional reflection is not included in the instrument model, a false focus condition is reached where the surface is 'believed' to be higher than it is. As the probed point continues to move towards the vertical edge, the objective will be forced to move further away from the surface to reach a false focus state. Very likely, this rapid and falsely perceived increase of surface height resulted in the measurement artefact observed figure 13. It was found that such edge effects can be minimised by aligning the scanning direction orthogonal to the plane of laser path, as indicated in figure 13(b).

The edge effect was also observed to be more severe when the vertical step was higher, since there was more space (along the scanning direction) for the spikes to grow. Figure 15 shows the measured profiles on two type ACG samples with the same pitch distance and different step height values. The edge effect was significantly worse on the sample with the deeper edge.

In the evaluation of the *z*-axis, the edge effect was a significant source of error in determining measured step height. In table 5, error (*E*) in the measured step height was the largest contributor to measurement uncertainty, which was shown in figure 9(a) to increase with the edge depth. Grid feature size, which is proportional to the pitch between features, was also found to affect the severity of edge effect. As ISO 5436-1 [36] requires at least five parallel profiles to be averaged, on smaller grids it was sometimes inevitable to include profiles affected by the edge effect. Table 8 provides a comprehensive summary of the reproducibility of step height measurement on cross grids with three different step height values and three different grid pitch values. Reproducibility rather than uncertainty contribution is considered because step height calibration was only provided for the material measure with nominal step height of 1.2 μm. It can be seen from table 8 that amongst features of the

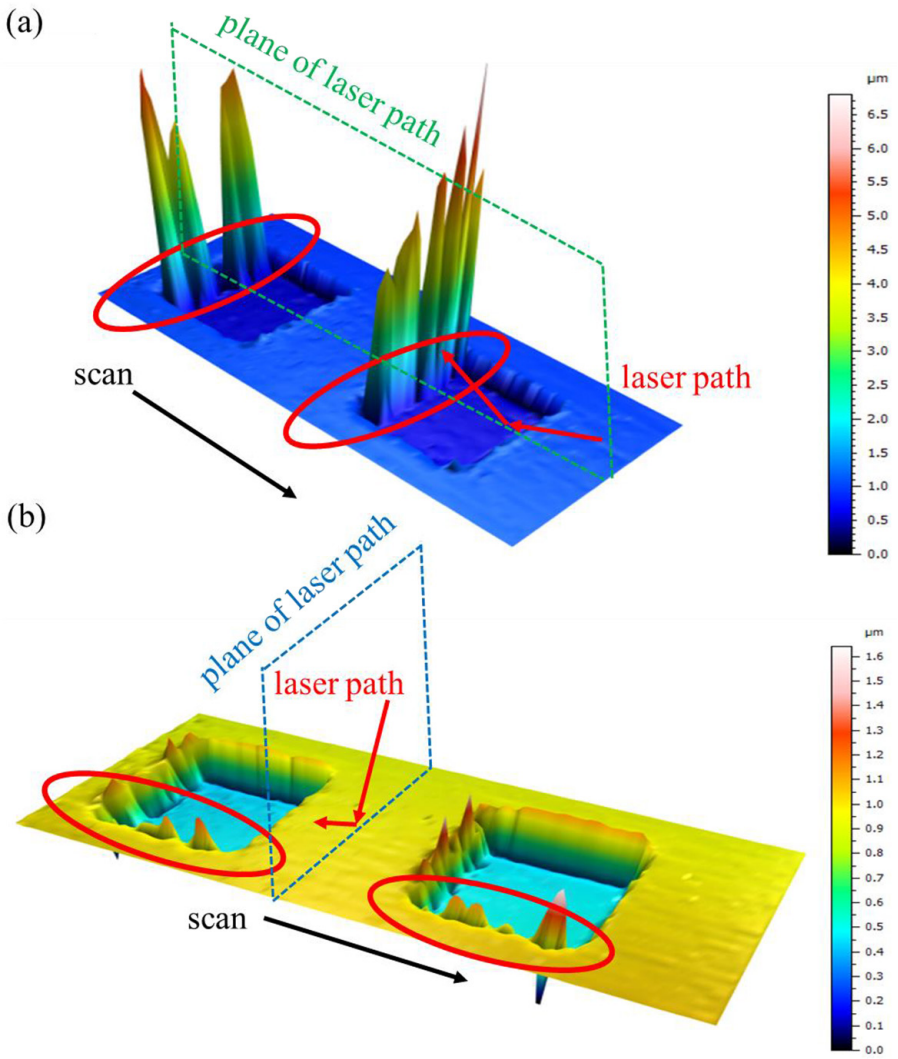

Figure 13. Measured surface topographies on a type ACG sample with nominal height of 0.5 µm and nominal grid pitch of 40 µm: (a) plane of laser path is parallel to the scanning direction, and (b) plane of laser path is orthogonal to the scanning direction.

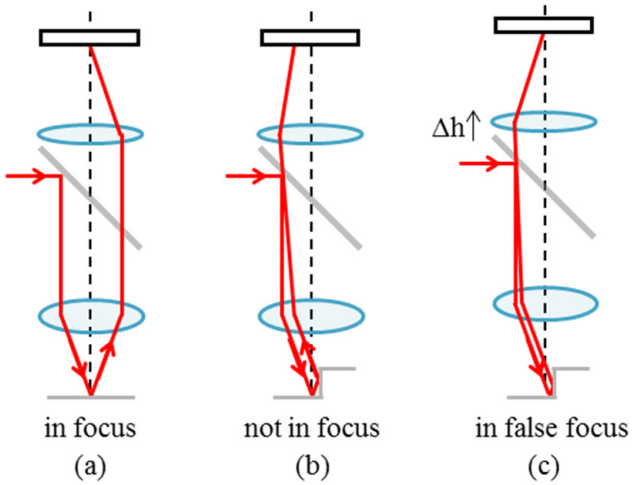

Figure 14. Illustration of edge effect as PAI scans toward a vertical edge, with the instrument (a) in focus, (b) out-of-focus and (c) in false focus.

same grid pitch (consequently the same size of the feature), deeper features tend to result in worse reprodu-cibility. Similarly, amongst features of the same height, smaller grid features tend to result in worse reproducibility. To summarise, measurement was consistently less reproducible with deeper and smaller grids.

Similarly, the estimation of the CoG of the grid features was also affected by the edge effect, as parts of the spikes may be included as part of the surface during the segmentation. The calibration of the lateral axes was repeated on the type ACG cross grid with a nominal step height of 1.2 µm and grid pitch of 16 µm. Two sets of grids with grid pitches of 16 µm and 40 µm were measured, the results of which are shown in table 9. Measurement uncertainty was significantly worse when using the smaller grids. Therefore, it is recommended to use the shallowest and largest grid features possible when calibrating α and l for the axes and mapping deviations. Alternatively, features that do not contain vertical edges may

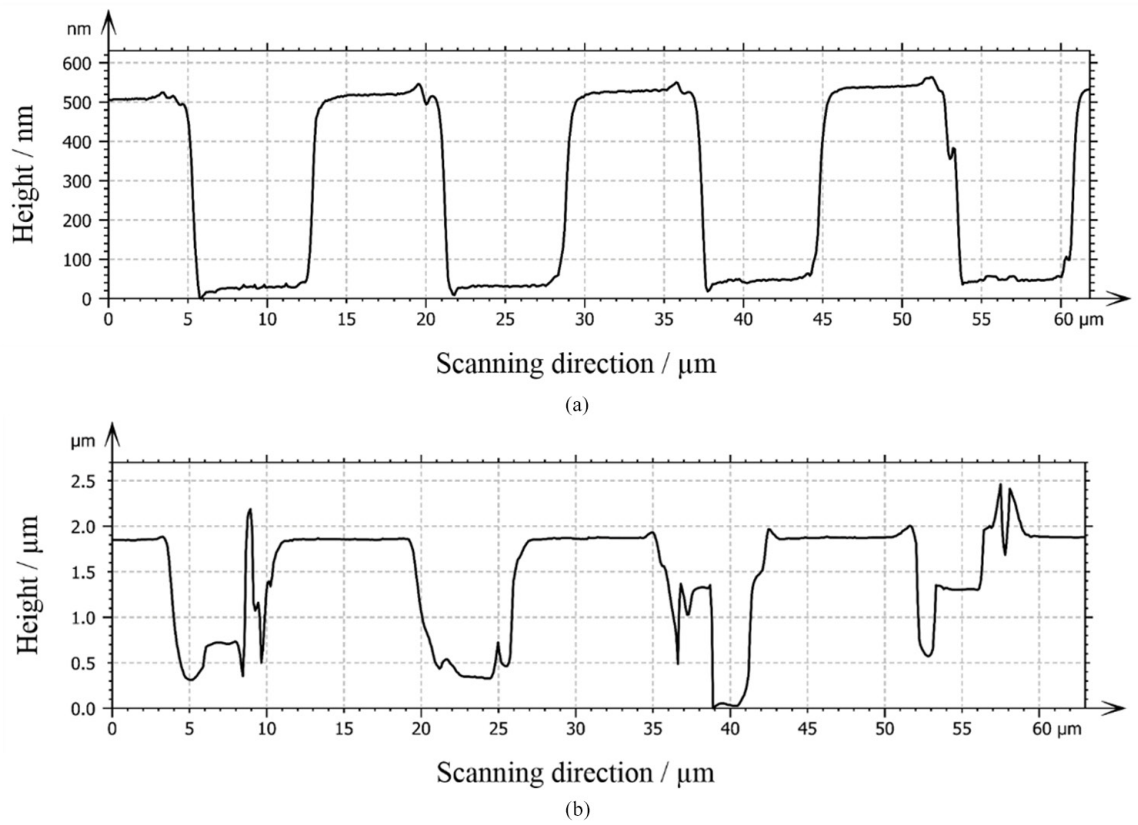

Figure 15. Profiles across the centre of grid features on two sets of type ACG cross grids with 16 µm pitch and (a) 0.5 µm step height and (b) 1.2 µm step height.

Table 8. Reproducibility of step height measurement on grid features of various dimensions.

Nominal height/µm	0.5	1.2	2.1	0.5	1.2	2.1	0.5	1.2	2.1
Grid pitch/µm		160			40			16	
Reproducibility/nm	1	5	5	2	6	5	4	24	44

Table 9. Amplification coefficients, linearity deviations of the *x*- and *y*-axis and their contributions to measurement uncertainty calibrated on grids of various dimensions of the type ACG with features deep 1.2 µm.

	40		16	
Grid pitch/µm	*x*-axis	*y*-axis	*x*-axis	*y*-axis
α	0.9988	1.0018	0.9989	1.0039
l/nm	103	248	400	356
u/µm	0.6	0.6	1.1	1

be used, such as a grid of spherical features similar to those used elsewhere [17].

5. Conclusions

The point autofocus instrument is an optical probe whose working principle enables measurements of optics, cutting tools and complex three-dimensional (3D) geometries. Consequently, the assessment of its metrological capabilities and resulting measurement uncertainty is core to provide users with confidence in exploiting this technology; however, this is still unreported in literature. Therefore, the main contribution of this paper is that it presents a procedure to assess the metrological characteristics as specified in ISO 25178-600 for a point auto focus instrument.

In this work, several metrological characteristics, specified in ISO/FDIS 25178-600, have been determined for a point autofocus surface topography measuring instrument, including flatness deviation (9 nm), amplification coefficients (0.9786, 0.9988, and 1.0018 respectively for the *z*-, *x*- and *y-axis*), linearity deviation (4 nm, 103 nm and 248 nm respectively for the *z*-, *x*- and *y-axis*) and *x*-*y* mapping deviations.

Most measurement instruments, like the instrument involved in this paper, are designed with thermal stability considerations, and operated in a temperature-controlled environment. Therefore, the authors expect the metrological characteristics to remain stable over the long term. In the absence of specified calibration frequency in the current ISO 25178, the authors recommend the metrological characteristics be re-verified once a year or after any major modification to the instrument. The contribution of these metrological characteristics to measurement uncertainty has been determined and metrological performance of point autofocus probe has been assessed to be compatible with other surface topography measuring instruments characterised in the literature. The characterisation of the *z-axis* allowed to detect a biased

measurement of step height, which, though, can be corrected by means of the amplification coefficient.

Furthermore, multiple reflections may cause significant measurement errors especially when measuring high steps, which is a common problem for most optical surface measuring techniques. Further investigation of this effect will be addressed in future work to provide guide for practical use of the instrument and improved instrument design. Moreover, filtering pipelines are shown to significantly impact on characterisation results along with adopted calibrated material measures, which are relevant as far as the measured feature and their calibration standard uncertainty are concerned.

Acknowledgments

This work was supported by the Engineering and Physical Sciences Research Council (Grant Number EP/M008983/1) (EPSRC), and the authors would like to thank Mitaka Kohki for collaboration.

ORCID iDs

Giacomo Maculotti https://orcid.org/0000-0002-2467-0426
Xiaobing Feng https://orcid.org/0000-0002-3328-5247

References

[1] Leach R K, Giusca C L, Haitjema H, Evans C and Jiang X 2015 Calibration and verification of areal surface texture measuring instruments *CIRP Ann.—Manuf. Technol.* **64** 797–813
[2] Leach R K 2011 *Optical Measurement of Surface Topography* (Berlin: Springer)
[3] ISO/FDIS 25178-600 2018 *Geometrical Product Specification (GPS)—Surface Texture: Areal—Part 600: Metrological Characteristics for Areal Topography Measuring Methods* (International Organisation for Standardization)
[4] ISO/CD 25178-700 2018 *Geometrical Product Specification (GPS)—Surface Texture: Areal—Part 700: Calibration, Adjustment and Verification of Areal Topography Measuring Instruments* (International Organisation for Standardization)
[5] ISO 25178-601 2010 *Geometrical Product Specification (GPS)—Surface Texture: Areal—Part 601: Nominal Characteristics of Contact (stylus) Instruments* (International Organisation for Standardization)
[6] ISO 25178-602 2010 *Geometrical Product Specification (GPS)—Surface Texture: Areal—Part 602: Nominal Characteristics of Non-Contact (Confocal Chromatic Probe) Instruments* (International Organisation for Standardization)
[7] ISO 25178-604 2013 *Geometrical Product Specification (GPS)—Surface Texture: Areal—Part 604: Nominal Characteristics of Non-Contact (Coherence Scanning Interferometry) Instruments* (International Organisation for Standardization)
[8] ISO 25178-605 2014 *Geometrical Product Specifications (GPS)—Surface Texture: Areal—Part 605 : Nominal Characteristics of Non-Contact (Point Autofocus Probe) Instrument* (International Organisation for Standardization)
[9] Wiora G, Bauer W and Seewig J 2016 *Definition of a Comparable Datasheet for Optical Surface Measurement Devices* (www.optassyst.de/fairesdatenblatt) (Stuttgart: Messe Stuttgart)
[10] de Groot P 2017 The meaning and measure of vertical resolution in optical surface topography measurement *Appl. Sci.* **7** 1–6
[11] Leach R K and Haitjema H 2010 Bandwidth characteristics and comparisons of surface texture measuring instruments *Meas. Sci. Technol.* **21** 032001
[12] Giusca C L, Leach R K, Helary F, Gutauskas T and Nimishakavi L 2012 Calibration of the scales of areal surface topography measuring instruments: part 1. Measurement noise and residual flatness *Meas. Sci. Technol.* **23** 035008
[13] Giusca C L, Leach R K and Helery F 2012 Calibration of the scales of areal surface topography measuring instruments: part 2. Amplification, linearity and squareness *Meas. Sci. Technol.* **23** 065005
[14] Giusca C L and Leach R K 2013 Calibration of the scales of areal surface topography measuring instruments: part 3. Resolution *Meas. Sci. Technol.* **24** 105010
[15] Seewig J and Eifler M 2017 Calibration of areal surface topography measuring instruments *Proc. SPIE* **10449** 1044911
[16] Eifler M, Hering J, von Freymann G and Seewig J 2018 Calibration for arbitrary metrological characteristics of optical topography measuring instruments *Opt. Express* **26** 16609–23
[17] Alburayt A, Syam W P and Leach R K 2018 Lateral scale calibration for focus variation microscopy *Meas. Sci. Technol.* **29** 065012
[18] Ekberg P, Su R and Leach R K 2017 High-precision lateral distortion measurement and correction in coherence scanning interferometry using an arbitrary surface *Opt. Express* **25** 18703–12
[19] Miura K, Nose A, Suzuki H and Okada M 2017 Cutting tool edge and textured surface measurements with a point autofocus probe *Int. J. Automation Technol.* **11** 761–5
[20] Miura K, Tsukamoto T, Nose A, Takeda R and Ueda S 2018 Accuracy verification of gear measurement with a point autofocus probe *Proc. Euspen's 18th Int. Conf. & Exhibition (Venice, 4–8 June)* pp 121–2
[21] Maculotti G, Feng X, Galetto M and Leach R K 2018 Noise evaluation of point autofocus surface topography measuring instrument *Meas. Sci. Technol.* **29** 065008
[22] Miura K and Nose A 2011 Point autofocus instruments *Optical Measurements of Surface Topography* ed R K Leach (Berlin: Springer)
[23] Leach R, Giusca C and Rubert P 2013 A single set of material measures for the calibration of areal surface topography measuring instruments: the NPL areal bento box *Proc. of the 14th Int. Con. MPES* pp 406–13
[24] Song J F, Vorberger T V V and Rubert P 1982 Comparison between precision roughness master specimens and their electroformed replicas *Precis. Eng.* **14** 84–90
[25] Leach R, Giusca C, Rickens K, Riemer O and Rubert P 2014 Development of material measures for performance verifying serface topography measuring instruments *Surf. Topogr.: Metrol. Prop.* **2** 025002–7
[26] Leach R, Giusca C, Guttmann M, Jakobs P J and Rubert P 2015 Development of low-cost material measures for calibration of the metrological characteristics of areal surface texture instruments *CIRP Ann.—Manuf. Technol.* **64** 545–8
[27] Leach R K 2013 *Characterization of Areal Surface Texture* (Berlin: Springer)

[28] Giusca C L and Leach R K 2013 *Calibration of the Metrological Characteristics of Coherence Scanning Interferometers (CSI) and Phase Shifting Interferometers (PSI). NPL Good Practice Guide No. 127* (Teddington, UK: National Physical Laboratory)
[29] VDI/VDE 2655 Part 1.1 2008 *Optical Measurement and Microtopographies; Calibration of Interference Microscopes and Depth Setting Standards for Roughness Measurement* (Berlin: Beuth–Verlag)
[30] ISO 25178-3 2012 *Geometrical Product Specifications (GPS)—Surface Texture: Areal—Part 3: Specification Operators* (International Organisation for Standardization)
[31] ISO 25178-2 2012 *Geometrical Product Specifications (GPS)—Surface Texture: Areal—Part 2: Terms, Definitions and Surface Texture Parameters* (International Organisation for Standardization)
[32] Evans C J 2009 PVr-a robust amplitude parameter for optical surface specification *Opt. Eng.* **48** 043605–12
[33] ISO 16610-21 2011 *Geometrical Product Specification (GPS)—Filtration—Part 21: Linear Profile Filters: Gaussian Filters* (International Organisation for Standardization)
[34] ISO 16610-22 2015 *Geometrical Product Specification (GPS)—Filtration—Part 22: Linear Profile Filters: Spline Filters* (International Organisation for Standardization)
[35] BIPM, IEC, IFCC, ISO, IUPAC and IUPAP 2008 *Guide to the Expression of Uncertainty in Measurement (GUM)* (Sèvres: Bureau Interantional des Poides et Mesures) (JCGM 100)
[36] ISO 5436-1 2000 *Geometrical Product Specifications (GPS)—Surface Texture: Profile Method; Measurement Standards—Part 1: Material Measures* (International Organisation for Standardization)
[37] Vorburger T V, Rhee H-G, Renegar T B, Song J-F and Zheng A 2007 Comparison of optical and stylus methods for measurement of surface texture *Int. J. Adv. Manuf. Technol.* **33** 110–8
[38] Gao F, Leach R K, Petzing J and Coupland J M 2008 Surface measurement errors using commercial scanning white light interferometers *Meas. Sci. Technol.* **19** 015303
[39] Spencer A, Dobryden I, Almqvist N, Almqvist A and Larsson R 2013 The influence of AFM and VSI techniques on the accurate calculation of tribological surface roughness parameters *Tribol. Intern.* **57** 242–50