EMPIRICAL ANALYSIS OF MEASUREMENT UNCERTAINTY FOR THE INSPECTION OF MICRO-FEATURES USING A CONFOCAL MICROSCOPE

Teresa WERNER[1,2], Fengzhou FANG[1]

1 Centre of MicroNano Manufacturing Technology (MNMT), Tianjin University, 92 Weijin Lu, Nankai District, Tianjin 300072, China; teresa.werner@yahoo.de
2 Metrodata GmbH, Im Winkel 15-1, 79576 Weil am Rhein, Germany

Abstract:

While for measurements on the macro-scale the applied measurement systems and typical influences are generally well understood, this does not hold true for measurements on the micro-scale. This limited knowledge results in a rather high value of measurement uncertainty. Thus, to support a suitable analysis of measurement uncertainty and at the same time enable a targeted improvement of the implemented measurement strategy, an empirical method for the determination of measurement uncertainty based on a systematic approach for system analysis is proposed.

On the example of inspecting parts with micro-features by a confocal microscope, a strategy has been developed to identify and quantify relevant influence factors on the measurement result, applying methods of Design of Experiments (DoE). By statistical evaluation of the results gathered there, an empirical description of the measurement system was synthesized. Major influence factors have been described and the influence of several options for the measurement strategy on the reliability of the measurement results have been quantified.

Based on this characterization, a value for measurement uncertainty was determined. Also, recommendations for ideal conditions of the measurement and the applied strategy have been derived. By applying such findings to the actual performance of measurements as well as to the according uncertainty budget, the overall uncertainty of the measurement results can be effectively reduced. Thus, an experimental strategy for the analysis of measurement uncertainty can at the same time support the determination of a realistic value for measurement uncertainty and the improvement of applied measurement strategies.

Keywords: Measurement Strategy, Reliability, Micro-Metrology, Confocal Microscope

1. INTRODUCTION

Advances in manufacturing technology enable the production of parts with complex geometry on the micro- or even nano-scale. This includes dimensional aspects of the overall form as well as surface properties like texture and roughness [1]. Along with developing these new possibilities for manufacturing, produced parts are necessary, ensuring reproducible, comparable and traceable measurement results.

Due to the small size of the inspected features, conventional tactile measurement methods cannot successfully be used. For example, the size of a probe might be so large compared to the measured geometry that the property of interest cannot be detected at all or is subjected to the mechanical filter of the touching element [2], [3]. Specific tactile methods for the nano-scale on the other hand often require an extensive measuring time and also put high demands on the measuring environment. Thus, they are not suitable for typical inspection tasks related to manufacturing. [4]

Thus, for the inspection of parts in the area of micro- and nano-range, optical methods are widely used, that often origin from methods of surface inspection [5]. Yet, the planning, implementation and evaluation of these measurements currently is a big challenge for the user. Whereas generally in metrology the application of commonly agreed standards and recommendations of Good Practice is the most efficient way to achieve reliable measurement results with low uncertainty, this is very difficult for the measurement of micro- or nano features.

As the experience on these comparatively new tasks is still limited, there are not yet sufficient commonly accepted standards or Good Practice guidelines to cover the scope of applications. Although standards have been under development for optical systems [6] in the last years, they are not yet sufficient for the planning of measurement tasks. Also, the various optical measurement techniques differ from each other widely. Thus, experiences gained for one method and recommendations derived from them (e.g. [7], [8]) cannot easily be transferred to another method. Accordingly, during the conduction of such measurements, the individual operator has to take many decisions in own responsibility.

To be able to use a measurement value for decisions in manufacturing, it has to be reliable enough. To quantify the reliability of a given result, the measurement uncertainty is stated [9], [10]. Very often, this value is also used as an indicator of the suitability of the underlying measurement procedure [11]. If the uncertainty of the measurement result is too high for the intended use, the measurement can be optimized, following the “Procedure for Uncertainty Management (PUMA)” [12].

To enable the determination and subsequent optimization of the measurement uncertainty, there has to be sufficient knowledge about the underlying measurement system to describe and quantify relevant influences, respectively how to reduce them if required. A comprehensive description and analysis can be done by a simulation of the measurement instrument, known as a virtual instrument. Although some virtual instruments have been developed for various method [13], [14], due to the wide variety of optical instrument they are still not available for many cases. Thus, a method is proposed that enables a practice-oriented determination of measurement uncertainty, flexible enough to be applied for any given measurement task.

2. METHODICAL APPROACH

Reliable measurement results have to fulfil four closely interrelated requirements: Repeatability, reproducibility, comparability and traceability.

Repeatability is the most basic requirement for measurements. It means that the variation of a measurement result is acceptably small, if the measurement is repeated under identical conditions, typically same measurement device with same settings, same environment and same operator, within a short time frame. Obviously, if a measurement method is not capable to produce sufficiently repeatable results, the obtained values are useless as they are random.

Considering a larger range of influences on the measurement, reproducibility describes, how well a given measurement result can be confirmed, when a measurement is repeated. To achieve a sufficient reproducibility, it is necessary to know which factors might influence the gathered result. These can be chosen adequately for the task at hand, documented and passed along with the measurement result in a protocol. This documentation allows to verify the current result if necessary by repeating the measurement even under slightly different conditions, e.g. using the same type of measurement instrument but from another supplier or at a different site, and thus enables the exchange of measurement results as an important information between different parties.

While repeatability and reproducibility consider only the single result obtained for a specific measurement task, comparability also includes the possibility to compare and combine it with results gathered from other methods or for different tasks. Thus, comparability describes if a given value can be used for quantitative or qualitative comparisons. This ensures that two parts with the same result are similar concerning the inspected property, whereas parts with different results differ from each other and accordingly can be put into an order.

Considering an even wider frame, traceability is the highest claim. It states that a given measurement result is connected to the definition of the relevant unit. This means that there is not only a relative difference between different values, but states that the realisation of the measurand, under consideration of measurement uncertainty, is related to an absolute value resulting from the definition of the respective unit. Thus, only traceability allows for a complete measurement result made up of value, unit and assigned uncertainty to be interpreted as a globally valid statement of information.

The ultimate goal for reliability always must be traceability of measurement results with an acceptably small measurement uncertainty. But, especially for new measurement techniques, this aim often can’t be achieved since the required metrological infrastructure is not yet established. This demands a suitable calibration chain with adequate standards to link the applied device and cross-check measurement results.

Yet, to set up suitable methods of product inspection as a part of quality management for the support of innovative manufacturing technologies, repeatability and reproducibility of measurement results will be sufficient. They enable to control the properties of products compared to a fixed qualitative standard which has be defined by experience or other empirical methods. By that, the stability of manufacturing processes can be assured and changes to prior results can be detected and targeted. Comparability would then be desirable as a next step, enabling a more informative evaluation of measurement results.

To efficiently facilitate the inspection of products with freeform surfaces in the micro range, the current project therefore focuses on the enhancement of repeatability and reproducibility, thus complementing ongoing more generalized research on the description of measurements in micro- and nano-range by other institutions. To this aim, a comprehensive case study was conducted, allowing to identify influence factors on the application of several 2,5D measurement instruments with different measurement principles available at the institution.

For the case study, two approaches were combined: On the one hand, related recommendations from standards and guidelines were identified and their suitability for the transfer on measurement tasks in the micro-range was tested. On the other hand, an analysis of possible influence factors oriented on the measurement process at hand has been conducted. The collected variables were then combined in an experimental plan according to methods of Design of Experiments, thus enabling a full statistical analysis of their significance as well as the identification of relevant interrelations.

Based on these findings, recommendations have been derived to improve the repeatability and reproducibility of measurement results obtained by these devices and thus ensure their reliability for the application in quality inspection.

3. CONDUCTION OF EXPERIMENTS REGARDING THE MEASUREMENT STRATEGY

To find out, how possible factors influence the measurement results and define an optimized approach, experiments have been conducted. There, exemplary geometric properties have been measured repeatedly under varying conditions. By a comprehensive statistical analysis of the data and by comparing measured values to calibrated values, a suitable measurement strategy can be defined, and the effect of changes can be estimated.

3.1 Measurement system

For the experiments, a confocal microscope from Olympus has been used. Two sets of experiments have been conducted: On the one hand on a typical, real part, featuring a variety of geometrical quantities (micro-geometry in lateral and vertical direction, form deviations, surface roughness). There, the influence of different strategies on the various features can be investigated. Yet, as the part is uncalibrated and also there are no specifications given, it is not possible to conclude, which setting is better. On the other hand, experiments have been performed using a calibrated step height artefact. In this case, the target value of the measurement is known so one can conclude on the optimal setting for this specific task. Yet, the calibration artefact features only one quantity to be measured, i.e. a micro-geometry in vertical direction in the order of 10 µm.

So, the possible conclusion are limited to this specific application.

3.2 Experimental setup

In a first round, experiments with basic settings have been conducted, verifying the application of recommendations given in existing standards. This contains mainly the selection of an appropriate objective lens and the adaptation of basic parameters for the measurement, such as positioning, focusing and light settings.

With the real-life workpiece, then various measurements have been conducted to check the applicability of rules given in according standards. It was found, that the application of the existing recommendations given in standards, although they do not fully cover the scope of possible settings and are partially adapted from measurements on the macro-range, enable a suitable basic set-up for the measurements. Whereas a deviation from these rules results in strong deviations and thus unreliable measurement results. Yet, the range of possible measurement strategies within the existing limitations nevertheless results in a quite high range of measurement results. This highlights the need for further regulations together with the necessity for comprehensive documentation of a gathered measurement result in order to reproduce it and thus be able to use the information contained therein.

For a second series of measurement experiments, a calibrated step-height artefact with a nominal value for the height of the step of 9,976 µm has been used. As the target value of the measurement is known, this artefact can be used to check the accuracy of the measurement device and also find out, which settings enable the reduction of systematic deviation.

As preparatory experiments, possible basic settings have been tested to enable suitable measurements. There, the step height has been repeatedly measured while adapting basic influence factors:

- Changing of objective lens, applying different magnifications (5x, 10x, 20x, 50x)
- Subsequent adaptation of other relevant parameters (positioning, focusing, light settings) to the currently effective objective lens
- Conduction of measurement in areal detection mode and profile detection mode
- Evaluation of data in different modes (step height, profile, roughness, geometry)

Based on that, a suitable configuration was defined under consideration of settings recommended in good practice (50x magnification, areal detection mode, evaluation in step height mode after applying automatic orientation correction). Then, the actual experiments to check the properties of the measurement device were conducted. Three influence factors were considered:

- Position of the artefact on the plate, i.e. position of the moving plate during measurement (center, lower left corner, upper left corner, lower right corner, upper right corner)
- Rotational position of the artefact in relation to plate main axis (0° - writing on the artefact in reading direction, 90°, 180°, 270° - turning clockwise from 0° position)
- Position of evaluated measurement field on the artefact (orienting on the long sides of the rectangular step height area, three fields spread over the length)

A partial-factorial experimental plan based on the rules of Design of Experiments (DoE) was set up (Table 1). This method allows to observe the influence of each factor on the system as well as possible interdependencies between two or more factors with a minimum number of experiments. For each experiment, five measurement values were taken to enable a proper subsequent statistical analysis.

By this experimental setup, the suitability and accuracy of the measurement device itself has been tested, namely the dependence of measurement results on the position and orientation of the artefact in the measurement area of the device. It has been found, that there are deviations between measurements at different positions, yet they remain within the specification of the instrument. This nevertheless suggests, that it is useful to keep parameters regarding the hardware, e.g. the position of the measured workpiece, constant, when performing measurements for a comparison of different products.

Based on these results, in a follow up experimental series, then the effect of possible settings and their interdependency has been tested. Specifically, the influence of normal vs. high accuracy registration of the picture and different methods of preparation and evaluation has been investigated. Thus, specific options of the measurement device can be checked for their influence on the measurement result and recommendations for an optimized strategy on the specific combination of measuring task and measurement instrument can be defined. It has been found, that all investigated parameters have significant influence on the achieved result and the option enabling a more accurate result has been identified.

3.3 Statistical evaluation of experimental results

The results of the experiments have been analyzed for outliers, indicating unusual combinations or problems. For this, display as a time series has been chosen. Additionally, this enables the easy comparison with specifications of accuracy for the device given by the manufacturer (Figure 1). For the utilized instrument, the specified accuracy for the measurement tasks at hand, i.e. the height of a step artefact of nominally 9.976µm is a < 0.2 µm +L/100 µm=0.3 µm. This graphical display of measurement results over time, both of summarized samples for the 5 values taken in each experiments and of original values, additionally enables the identification of possible drifts or similar slow influences as well as the identification of samples with atypical results, i.e. outside of 95%/2σ-spreading region.

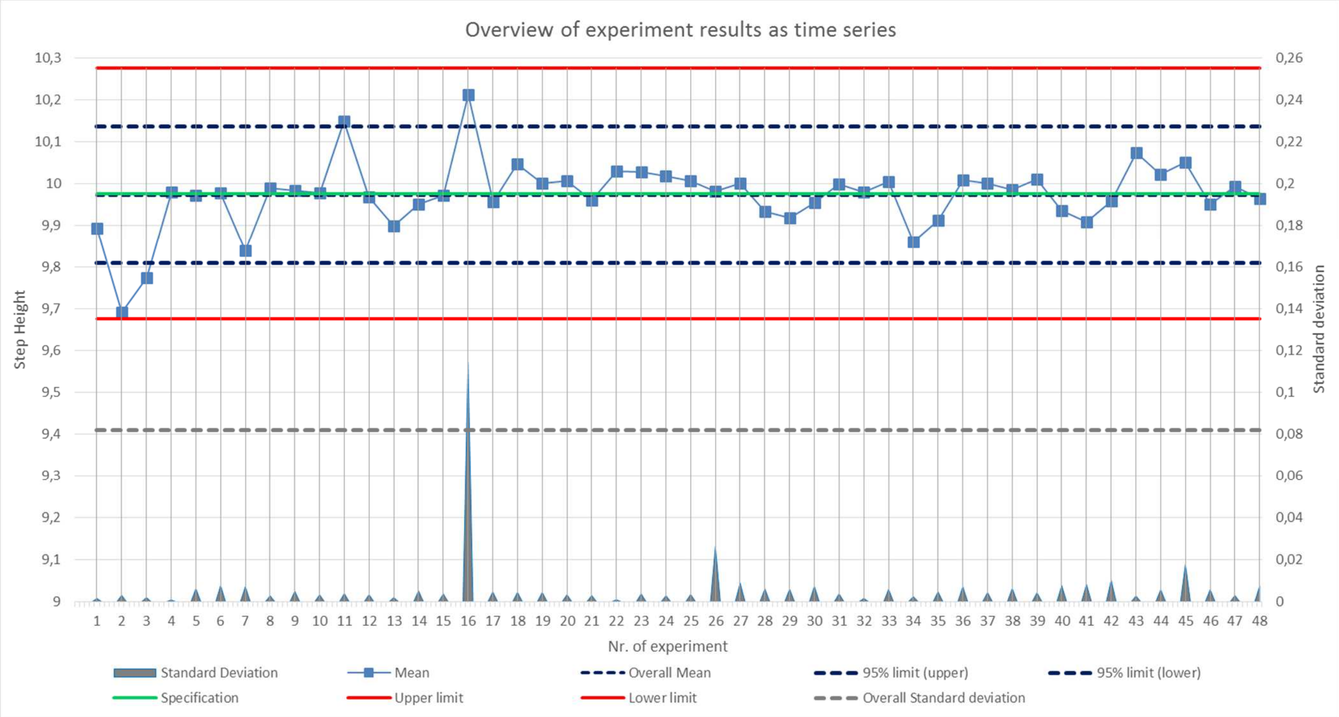


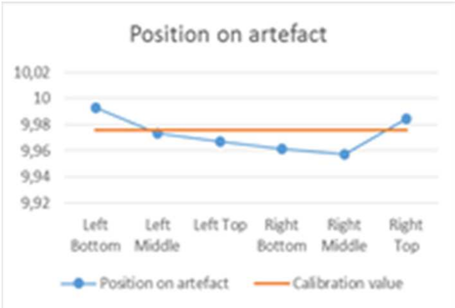
Fig. 1: Graphical evaluation of results: Time series of sample results for measurements on a step-height artefact (blue dots: mean; gray peaks: standard deviation; red continuous line: specification of accuracy for the device by manufacturer; green continuous central line: specified value of the artefact; blue / gray interrupted lines: limit of confidence interval at 95% level resp. mean of means)

From these results, it can by concluded that the Olympus confocal microscope works within the accuracy specifications stated by the manufacturer, yet the allowed range is fully used. Also, the specifications for repeatability have been kept for most samples, as indicated by the standard deviation related to the individual samples taken under continuous settings.

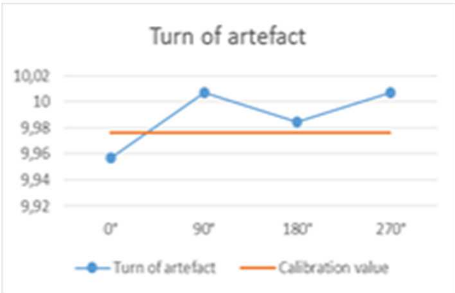
Nevertheless, some samples contain unusual values close to the lower (2 and 3) or upper limit (11 and 16). Especially for sample 16, in addition the standard deviation therein is also unusually high. A re-evaluation of the raw measurement data maintained the values for sample 2, 3 and 11. As possible reason for the deviation, non-optimal light settings can be considered as the measurement data appear rather dark (2, 3) resp. bright (11) compared to other measurement shots, although the limiting constraints for illumination are kept. Regarding sample 16, the values could not be confirmed by re-evaluation of the measurement data, thus showing an example of poor reproducibility under consideration of the so far identified influence factors.

To enable a quantitative analysis of the influences, the statistical data has been analysed based on the method of ANOVA (Analysis of Variances) which is targeted towards systems with multiple interacting influences.

For the experiments on the calibrated artefact, the effect of the three controlled influence factors was evaluated based on the partial-factorial experiment design and graphically displayed in effect diagrams (Fig. 2). For that, data of the measurement samples was accordingly grouped and summarized.



a)



b)

c)

Fig. 2: Display of effects by controlled influence factors compared to nominal value of artefact

Then, the significance of the observed influences was calculated based on Fisher distribution (F-test). Also, the η^2-coefficient has been determined to estimate the overall amount of variation that can be explained by that factor:

- Position on artefact: confidence level $1-\alpha=0.98$; $\eta^2\approx3\%$
- Turn of artefact: confidence level $1-\alpha=0.99$; $\eta^2\approx10\%$
- Position of artefact on plate: confidence level $1-\alpha=0.99$; $\eta^2\approx7\%$

All observed influence factors are significant, yet the ratio of unexplained variation of results is still 80%. As this is likely to result from imperfections of the measurement system as such, emphasis has to be put on the correct handling of the measurement device and improving the repeatability of results gathered there, especially considering the optimal choice of settings for lighting as they may strongly influence both the accuracy and the repeatability of the measurement result.

The influence of the position of the artefact on the plate is significant, yet overall low. A difference is to be expected due to mechanical imperfections; the low effect might indicate a good state of instrument's overall mechanical behaviour. The comparatively high influence of rotational position of artefact should be subjected to further investigation; this might possibly indicate differences between x- and y-direction of data processing. As the ratio of unexplained variation still is very high, further analysis of influence factors on the measurement instrument will have to be conducted. Especially, the conformity to repeatability specification has to be checked in addition to the now confirmed accuracy specification.

Additionally, a series of experiments have been conducted to explore the influence of additional settings in the software on the measurement result. To enable the comparison with the experiments from the main experimental schedule, the basic set-up has been kept the same. Six factors related to the data detection as well as the data evaluation have been analysed:

- Measurement with or without additional collection of a "colour shot", i.e. an evaluation of height information based on a conventional confocal image rather than on the laser induced image
- Measurement with usual or "high resolution" setting
- Evaluation of data with or without correction of data for orientation
- Evaluation of data with different averaging zones

By comparing the results with the target value given for the calibrated artefact, it can be concluded, which operational conditions enable a better measurement result. Here, it has been found that taking a picture with high resolution, using only height information instead of additional colour information and a manual control of light settings provide more accurate results.

Overall, the effects of these settings is very small compared to deviations that can be caused by not following existing recommendations given in standards and guidelines. Thus, it has to be noticed that the foremost necessity on planning geometrical measurements for micro-scale features is to follow the existing recommendations in standards and guidelines. Further improvement of the measurement strategy is desirable as well as a proper documentation of whatever strategy was actually implemented. For, the way how the measurement is conducted, may greatly influence the results gathered.

Based on these findings, recommendations for the conduction of measurement at the confocal microscope have been derived, considering both the regulations given in standards and the additional information found through the experiments.

5. SUMMARY AND OUTLOOK

So far, experiments have been conducted mainly for the confocal microscope. Also, the scope of measured quantities is covering only micro-height for calibrated features, whereas for other features only information based on experimental measurements of uncalibrated workpieces are available. Due to this narrow scope of experimentation, the reliability and transferability of the gathered results are limited.

In order to enable a broader interpretation of the statistical results and a higher level of trust in the gathered conclusions, additional experiments should be performed. These will either confirm the results gathered so far, or will show up the boundaries of application for the current findings. Thus, it is necessary to perform additional experiments, including different measured quantities, different measurement instruments and a different range of measured values.

Such a broadened data base can provide a reliable starting point to derive guidelines for Good Practice for a wide variety of measurement tasks related to geometrical features on the micro- and nano-scale. Thus, they will help to truly establish experiences about good measurement strategies in order to gather reliable, traceable and reproducible measurement results.

In addition, there are several possibilities for the further use of the current results in other areas. Remaining in the field of metrology, for this case study only a special group of measurement instruments have been considered. Yet, measurement of free-form elements and other micro-features can also be conducted by several other types of measuring instruments, such as fringe projection devices, micro-coordinate measuring machines equipped with opto-tactile fiber probes or electric probes or computer tomographs. The method implemented in this case study can easily be transferred to application for such instruments, then considering the influence factors of specific relevance for the respective measurement principle.

On a long term, research on product specification also should be broadened to include function-oriented specification, as the new features also put new requirements on the way how to define the required attributes of the manufactured parts. Thus, a close collaboration between various disciplines of engineering, from product design over manufacturing up to inspection, is necessary to master the challenges and opportunities provided by this area. So far, experiments have been conducted mainly for the confocal microscope. Also, the scope of measured quantities is covering only micro-height for calibrated features, whereas for other features only information based on experimental measurements of uncalibrated workpieces are available. Due to this narrow scope of experimentation, the reliability and transferability of the gathered results are limited.

In order to enable a broader interpretation of the statistical results and a higher level of trust in the gathered conclusions, additional experiments should be performed. These will either confirm the results gathered so far, or will show up the boundaries of application for the current findings. Thus, it is necessary to perform additional experiments, including different measured quantities, different measurement instruments and a different range of measured values.

Such a broadened data base can provide a reliable starting point to derive guidelines for Good Practice for a wide variety of measurement tasks related to geometrical features on the micro- and nano-scale. Thus, they will help to truly establish experiences about good measurement strategies in order to gather reliable, traceable and reproducible measurement results.

In addition, there are several possibilities for the further use of the current results in other areas. Remaining in the field of metrology, for this case study only a special group of measurement instruments have been considered. Yet, measurement of free-form elements and other micro-features can also be conducted by several other types of measuring instruments, such as fringe projection devices, micro-coordinate measuring machines equipped with opto-tactile fiber probes or electric probes or computer tomographs. The method implemented in this case study can easily be transferred to application for such instruments, then considering the influence factors of specific relevance for the respective measurement principle.

On a long term, research on product specification also should be broadened to include function-oriented specification, as the new features also put new requirements on the way how to define the required attributes of the manufactured parts. Thus, a close collaboration between various disciplines of engineering, from product design over manufacturing up to inspection, is necessary to master the challenges and opportunities provided by this area.

ACKNOWLEDGEMENTS

The underlying research for the development of a prototypical application of the described assistance system was gratefully supported by the National Natural Science Foundation of China (NSFC) within the framework of the program „Research Fellowship for International Young Scientists“.

REFERENCES

[1] DIN 4760: Form deviations; Concepts; Classification system. 1982.

[2] T. Pfeifer, R. Schmitt: Fertigungsmesstechnik. Oldenbourg Wissenschaftsverlag, München 2010.

[3] A. Weckenmann: Koordinatenmesstechnik – Flexible Strategien für funktions- und fertigungsgerechtes Prüfen. Carl Hanser Verlag, München 2012

[4] D.J. Whitehouse: Handbook of Surface and Nanometrology. CRC Press, Boca Raton 2011.

[5] Rahlves, M., Seewig, J.: Optisches Messen technischer Oberflächen: Messprinzipien und Begriffe. Beuth Verlag, Berlin 2009.

[6] ISO 25178-3: Geometrical product specifications (GPS) - Surface texture: Areal - Part 3: Specification operators, 2012.

[7] R. Leach, L. Brown, X. Jiang, R. Blunt, M. Conroy and D. Mauger: Measurement Good Practice Guide No. 108 – Guide to the Measurement of Smooth Surface Topography using Coherence Scanning Interferometry. NPL, 2008.

[8] J. Petzing, J. Coupland and R. Leach: Good Practice Guide No. 116 - The Measurement of Rough Surface Topography using Coherence Scanning Interferometry. NPL, 2010.

[9] JCGM 100:2008: GUM 1995 with minor corrections. Evaluation of measurement data — Guide to the expression of uncertainty in measurement. JCGM/WG 1, 2010, corrected version. (via: http://www.bipm.org/)

[10] JCGM 101:2008: Evaluation of measurement data — Supplement 1 to the “Guide to the expression of uncertainty in measurement” — Propagation of distributions using a Monte Carlo method. JCGM/WG 1, 2008. (via: http://www.bipm.org/)

[11] VDA: Band 5 Prüfprozesseignung, Eignung von Messsystemen, Mess- und Prüfprozessen, Erweiterte Messunsicherheit, Konformitätsbewertung. VDA, 2010.

[12] ISO 14253-1: Geometrical Product Specifications (GPS) - Inspection by measurement of workpieces and measuring equipments - Part 1: Decision rules for proving conformance or non-conformance with specifications. ISO, 1998.

[13] R. Schmitt; F. Koerfer; O. Sawodny; R. Krüger-Sehm; G. Goch; S. Simon; C. Bellon; A. Staude; P. Woias; F. Goldschmidtböing; M. Rabold: Virtuelle Messgeräte - Definition und Stand der Entwicklung. tm - Technisches Messen 75/5 (2008). pp. 298-310.

[14] H. K. Mischo; T. Pfeifer; F. Bitte: Model-based optimization of interferometers for testing aspherical surfaces. In: SPIE 45th Annual Meeting: Laser Interferometry X (San Diego, USA, August 2000) SPIE vol 4101. pp. 497 – 510