Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 118 (2023) 1084–1089

www.elsevier.com/locate/procedia

16th CIRP Conference on Intelligent Computation in Manufacturing Engineering, CIRP ICME '22, Italy

Comparison of methods for management of measurement errors in surface topography measurements

Giacomo Maculotti[a,*], Gianfranco Genta[a], Danilo Quagliotti[b], Hans N. Hansen[b], Maurizio Galetto[a]

[a]*Department of Management and Production Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy*
[b]*Department of Mechanical Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark*

* Corresponding author. Tel.: +39-011-090-7236; *E-mail address:* giacomo.maculotti@polito.it

Abstract

Surface technology is essential to engineer surface properties by topologically optimised design or machining and finishing treatments. Optical surface topography measuring instruments represent state-of-the-art solution to characterise technological surfaces. Topographies measured by optical instruments are affected by errors (non-measured points and spikes), due to complex interactions between the measurand (the topography) and the instrument, liable of poor measurement quality and biasing characterisation. The literature proposes several approaches to manage measurement errors basing on empirical approaches (thresholding, interpolation) and machine learning modelling. This work compares the methods performances applied to industrially relevant case studies (highly polished and native additive manufacturing surfaces).

Keywords: Topography; Surface analysis; Machine learning

1. Introduction

Surface topography is the boundary of a component that separates it from the surrounding environment and through which all interaction between the component and the surroundings take place [1–3]. Surface technology consists in modifying technological surfaces to obtain a modification and control of their properties [4,5]. Several properties can be controlled by coating [6], mechanical processes and energy-based process, e.g. laser ablation, to achieve a micro-texturing by moving or removing material [7]. The numerous surface modification technologies are relevant in several fields, such as optics, biomedical engineering, mechanical engineering. In fact, a wide range of properties can be engineered by well-established relationship between surface topography and surface functionality. Amongst the other, optical properties, absorption and adsorption properties, essential in biological application, technological mechanical, hydrodynamical and tribological properties, e.g. friction resistance, stiction, and physical properties, e.g. wetting and adhesion [7].

Similarly, not only final product characteristics but also manufacturing and finishing processes can be studied and optimized on the basis of resulting surface topography, for this bears a uniquely defined manufacturing signature characteristic of the process [8]. This concept has found an extensive application in literature to study high precision processes, e.g., hard metals and cermet finishing [9], micro-injection molding [10,11], innovative manufacturing processes, e.g. additive manufacturing [12,13], as well as to address complex interaction of mechanical application, e.g., for tribology [14–17] and machining [18,19].

Therefore, the measurement of surface topography is necessary to enable these applications. This can be achieved by surface topography measurement instruments [2,20]. Several alternatives are available, ranging from contact probe [21], optical probes [22,23] to several working principles of actual topographical microscopes [2,20]. Although the former class is

10.1016/j.procir.2023.06.186

traditionally and still largely adopted in industry, its representativeness, time-cost and metrological performances, in terms of accuracy and precision, is significantly worse than those of topographical microscopes.

Surface topography measuring microscopes are the state-of-the-art to measure and characterize surfaces. These rely on several working principles. Amongst the most extensively adopted, we can find coherence scanning interferometry (CSI), confocal microscopy (CM) and focus variation (FV). Indeed, they all come with related strengths and weaknesses and none can be considered as the absolute best option when considering metrological performances and measurement efficiency [24]. These can be typically estimated by metrological characteristics [25], the amount of measurement errors and the measurement time, respectively.

CSI typically presents the best resolution and measurement reproducibility, respectively in the order of 0.1 nm and 1 nm [26]. It provides best performances on highly polished, reflective surfaces, it can measure transparent and translucent materials, but measurements are typically affected by measurement errors when highly rough surfaces are targeted and despite measurement approaches can be devised and deployed to partially overcome the issue, these typically result in extremely long measurement time [27].

FV presents worse resolution and measurement noise than CSI, i.e., 10 nm and 50 nm. Best performances are obtained in measuring rough surfaces and shape, typically by very fast measurements. Conversely, high reflective surfaces and very low roughness induce several errors [28].

CM shows intermediate performances of resolution and reproducibility, i.e., 1 nm and 10 nm, respectively. Similarly, to CSI best performances are obtained on polished and smooth surfaces. However, the contemporary presence of low roughness, i.e. little light scatter and the presence of a significative shape, mostly high slopes, on the topography greatly limits performances, despite the possibility of large numerical aperture [2].

As any other measurement, surface topography measurement instruments may be affected by measurement errors. These are the result of complex interaction between the measurement system and the physical surface of the measurand. They can be sourced by random factors, most typically environmental noise, electromagnetic noise, and systematic factor, e.g., light scatter, drift, optical path limitation due to, for example, the objective numerical aperture, and system working principle, as detailed above, in the worst working condition [24,29–31].

Measurement errors, in the most common cases, result in two categories of errors: *non-measured points*, or *voids*, and *spikes* [30,32]. These significantly affect the subsequent characterization of surfaces, performed by evaluating surface topography parameters, e.g. *Sa*, *Sq*, *Sdq*, etc [30–32]. Therefore, the identification and correction, i.e., the *management*, of measurement errors is crucial to obtain unbiased surface characterization. Several approaches have been proposed in the literature. However, the comparison of their performances has not been addressed, if not sparsely.

This work aims at comparing the performances of the available management methods by considering state-of-the-art surface topography measuring instrument in challenging working condition, which is unreported in literature. Section 2 presents the methodology, first summarizing the available management method and then describing the experimental procedure. Section 3 presents and discusses the obtained results. Section 4 finally draws the conclusions.

2. Methodology

This section first discusses the methods for identification and correction of measurement errors that are available in literature and here considered, then describes the experimental setup exploited in this work, and last addressed the methodology to compare the mentioned methods. In general the output of a topographical measurement is a micrography, consisting of *z(x,y)*, i.e., vertical deviation *z* of the measured surface with respect to a reference plane discretized, at a certain lateral resolution equal to the pixel size, in a *(x,y)* grid.

2.1. State-of-the-art measurement error management methods

Three main alternative spike identification methods are available: Sq-thresholding, median filtering and gaussian process regression [30,33]. Typically, literature and state-of-the-art measurement management software [30,33,34] proposes first to identify spikes, for non-measured points are inherently identified, and then set them as voids. Indeed, the Sq-thresholding does not set identified spikes to non-measured points. Subsequently, voids are filled in exploiting median filtering or gaussian process regression, most typically [30,33,34].

2.1.1. Sq thresholding

This is a statistical method assuming spikes as outliers of the micrography. In particular, the method states that a measured point *z* is a spike if $z \notin [Sa - 3Sq, Sa + 3Sq]$, and its height is reduced accordingly to the limit of the interval. This approach, although resembling outlier detection method and relying on rule of thumb and experience, may induce a biased identification of the spikes and often require operators to tailor the threshold to suit best the application [30,33].

2.1.2. Median filtering

Median filtering applies a median-based filter to the micrographs and detects the spikes as the outliers of the residuals. Although several filters are available, e.g., gaussian and modifications, spline, erosions filters, etc., median is the most commonly adopted and implemented in software packages thanks to its inherent robustness to extreme values, upon definition of the filter [34]. The main drawback in (median) filtering is the liability of removing spatially relevant scales from the topography or apply an unwanted denoising [33]. Voids management is performed by replacing the non-measured point with the theoretical value that the median filter assumes in the missing measurement.

2.1.3. Gaussian Process Regression

Gaussian process regression (GPR) is a supervised machine learning technique recently proposed to manage measurement

errors [30]. The procedure is exactly the same as median filtering, i.e., spikes are identified as outliers on the residuals and, once set to non-measured points, all the voids are replaced by the GPR model predicted response. The procedure offers several advantages. Firstly, by means of a kernel, it caters for spatial correlation that is possibly present in the measured surface, which is often sourced by manufacturing signature or measurement process, i.e., CCD cameras of the sensor. Secondly, it includes a self-adjusting capability thanks to the presence in the prediction estimator of weights proportional to the residuals, so that high accuracy can be achieved [35]. The procedure assumes a constant model, to generalize the surface modelling [30], and selects the kernel to minimizes the residual RMSE by Bayesian optimization [36].

2.2. Measurement set-up

According to the discussion presented in the Section 1, three most commonly adopted and state-of-the-art surface topography measuring instruments were considered: a CSI Zygo NewView 9000, a FV Alicona Infinite Focus Sensor R25, both hosted at the Mind4Lab at Politecnico di Torino and a CM Olympus LEXT 4100, hosted at DTU, as shown in Fig. 1. Each instrument measured a challenging surface for their working principle and representative of industrially highly relevant case studies. In particular, the CSI performed a measurement on surface by electron beam melting (EBM), for this is one of the most largely adopted additive manufacturing technologies and present high roughness [13], the FV on a lapped and polished stainless steel (SS) nano-hardness reference block [37], and the CM a replica of a 528 Rubert specimen, i.e. a material measure for the calibration of surface topography measuring instruments [38]. The 528 replica was achieved by impression of a silicon-based medium, whose replication fidelity is often a critical aspect. Measurements were set up according to best practices to minimize measurement errors, each sample was measured 10 times in repeatability condition, and only a single field of view (FOV) was acquired. CSI featured a Mirau 20× objective, with numerical aperture 0.4, pixel size of 0.429 µm × 0.429 µm and FOV of 1000 × 1000 pixels. FV mounted a long-distance 10× objective, with numerical aperture 0.28, pixel size of 1 µm × 1 µm and FOV of 2040 × 2040 pixels mounted on a Yaskawa cobot, being a system for freeform surface topography measurement. CM was equipped with a 50× lens objective, with numerical aperture 0.95, FOV of 260 µm × 260 µm with 1024 × 1024 pixels.

2.3. Measurement error method effect comparison

Measured surfaces are characterized independently by estimating surface topography parameters. No standard filtering was applied prior the estimation of surface topography parameters to avoid superimposing the effect of this operation on the error correction. Levelling is performed by least-square plane correction prior the evaluation of surface topography parameters. The evaluated set of surface topography parameters includes areal height parameters, namely the average height *Sa*, the root mean square height *Sq*, the skewness *Ssk*, the kurtosis *Sku*, and the maximum height *Sz*, and the hybrid parameter root mean square gradient of the height *Sdq* [39]. Plane correction, Sq-thresholding and median filtering measurement error correction are performed exploiting a commercial state-of-the-art software MountainsLab v8.2 [34]. GPR based measurement error correction is performed by a self-developed script in Matlab 2021b. Topographical parameters are evaluated in MountainsLab v8.2.

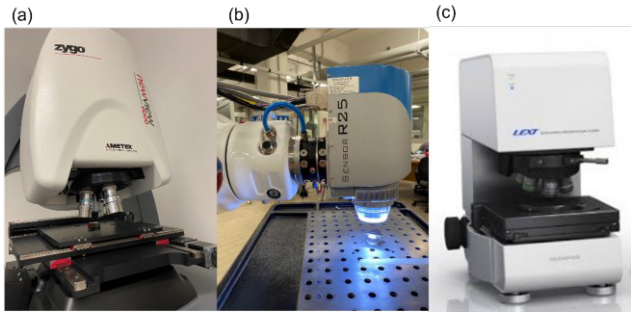

Fig. 1. (a) CSI Zygo NewView 9000; (b) FV Alicona IF SR25 on Yaskawa cobot; (c) CM Olympus LEXT 4100.

The effect of different measurement error management is assessed performing an ANOVA with one factor having three levels, i.e., the measurement error management method, on the evaluated parameters. Indeed, different instruments, i.e., different surfaces, are considered separately. Additionally, corrected surfaces by the three methods discussed in Section 2.1 are pairwise compared pixel-by-pixel. This allows comparing the point-wise correction effect. The comparison consists in performing the difference of the pixels height in correspondence of the extended set of voids, i.e., the set containing non-measured points and void pixels set in that condition after the spikes identification. Per each pair of compared micrograph, the larger extended set of voids is considered. The result of the difference is tested for normality by means of graphical test, i.e., normal probability plot (NPP), and quantitative chi-squared test [40]. Hypothesis testing on the difference with null hypothesis of normality is relevant, for ideal condition would introduce only random errors in the effect of different measurement errors management methods.

3. Results and Discussion

3.1. CSI measurements of EBM surface

The EBM surface topography presented a typical rough texture with presence of several hills, dales, and local discontinuities, as shown in Fig. 2. Several non-measured points affect the measurement performed by the CSI. No spikes can be appreciated. Accordingly, only the median and GPR correction method were applied.

Qualitatively, no differences can be appreciated, as shown in Fig. 2. Quantitative analysis, performed by paired hypothesis test on the average of the evaluated parameters, for it ANOVA cannot be run on one factor with two level design, shows that, with a risk of error of 5%, the two measurement error correction methods introduce a systematic difference on the considered parameters, as shown in the interval plot in Fig. 3. The statistical significance of the methods' diversity in the

correction, is confirmed also by the NPP of the corrected points height, as shown in Fig. 4(a).

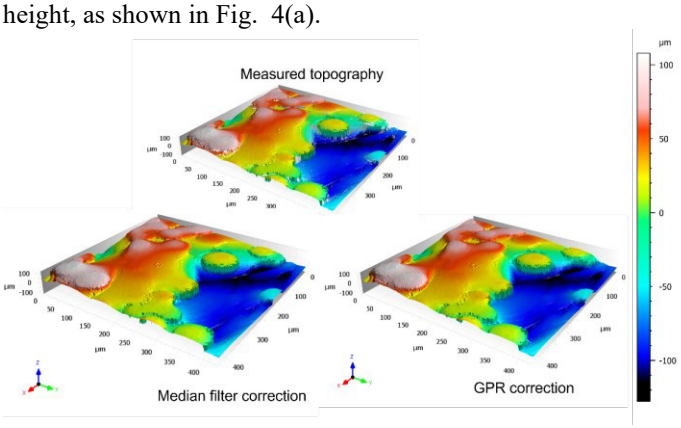

Fig. 2. EBM topography measured by CSI. Original topography with only voids and corrected surfaces: qualitatively no differences can be appreciated.

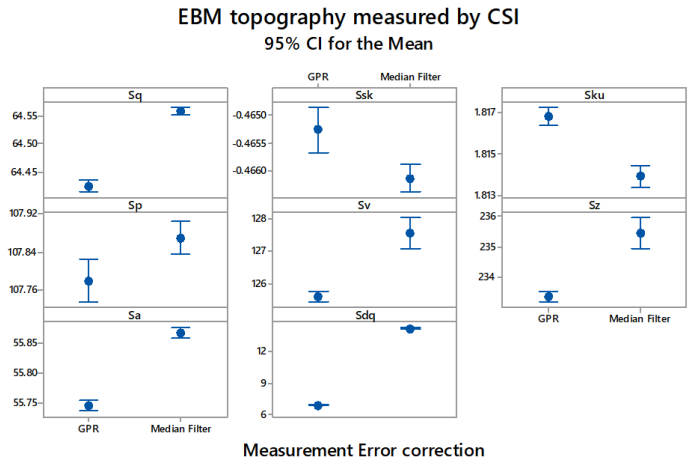

Fig. 3. Interval plot of the considered parameters of the EBM topography measured by the CSI after the voids correction by GPR and median filter.

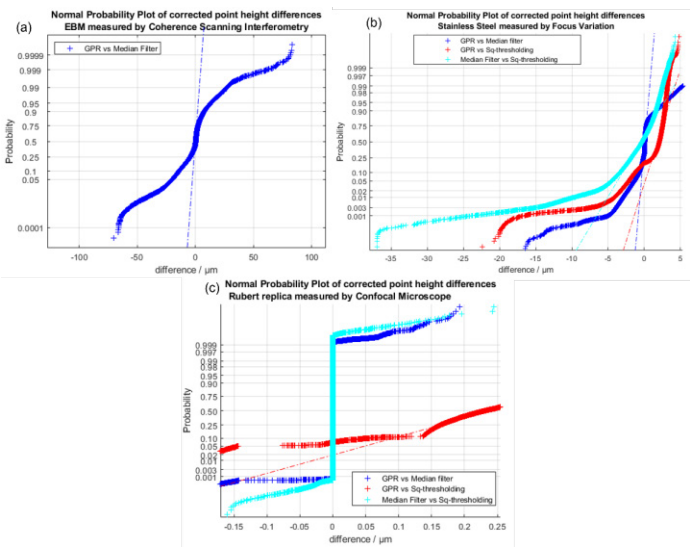

Fig. 4. Normal Probability Plot of the point heigh differences of the three case studies: (a) EBM by CSI, (b) SS by FV and (c) Rubert by CM. Due to space, only the NPP of the difference of the corrected points of first topography (per each case study) is shown.

3.2. FV measurement of polished surface

Fig. 5 shows the results of the measurement of the stainless-steel specimen by the FV. Several spikes can be appreciated in the raw data, due to the high reflectivity of the surface. Also, some non-measured points (0.1%) were present.

The three considered measurement error correction methods show, from Fig. 5, a relevant difference in the measured height range, i.e., *Sz*. ANOVA demonstrated that the correction methods introduce a significative difference in the considered topographical parameters for all the cases, but for *Sp*, i.e., the deepest peak, with a risk of error of 5%. Interval plot of the parameters allows a qualitative appreciation of ANOVA results, and is reported in Fig. 6. It motivates that the higher dispersion of *Sp* after the median filter does not allow to see systematic differences amongst the correction methods. Consistently with the ANOVA results, also NPPs of the difference of the corrected points show a trend significantly different from a normal distribution, as shown in Fig. 4(b).

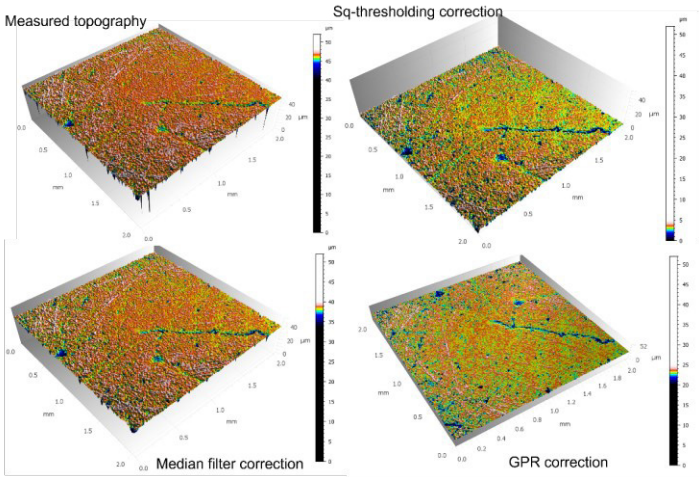

Fig. 5. SS topography measured by FV. Original topography with several spikes and voids due to high reflectivity. Correction shows relevant impact on height ranges.

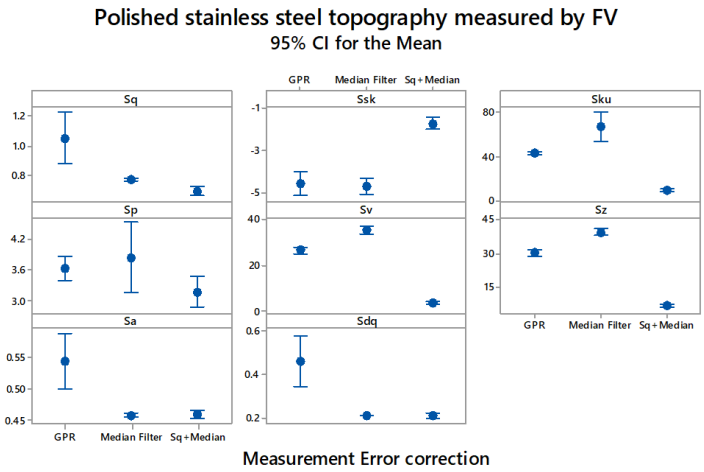

Fig. 6. Interval plot of the considered parameters of the highly polished stainless steel sample topography measured by the FV after the spikes and voids correction.

3.3. CM measurement of reflective and highly-sloped surface

The Rubert specimen replica obtained by impression was measured by the CM and results are shown in Fig. 7. No voids were generated, but several spikes affect the topography due to spurious scatter prompted by the high reflectivity, high average local slope, i.e., the inherent sinusoidal shape, and the texture, due to the manufacturing signature.

Qualitatively, Sq-thresholding proved to perform worse than alternative approaches in correcting the spikes. Also, to enable application of the thresholding approach, while catering for the topography shape, the thresholding method was applied on the topography resulting from the difference between the average measured micrograph and the considered measurement.

ANOVA, and related interval plot represented in Fig. 8, show that, with a risk of error of 5%, the three considered measurement error correction method introduce significative difference in all the considered topographical parameters, but for *Ssk*.

Accordingly, also the NPPs of the pairwise difference of the corrected points heights show a trend significantly different from a normal distribution, for all the possible comparisons.

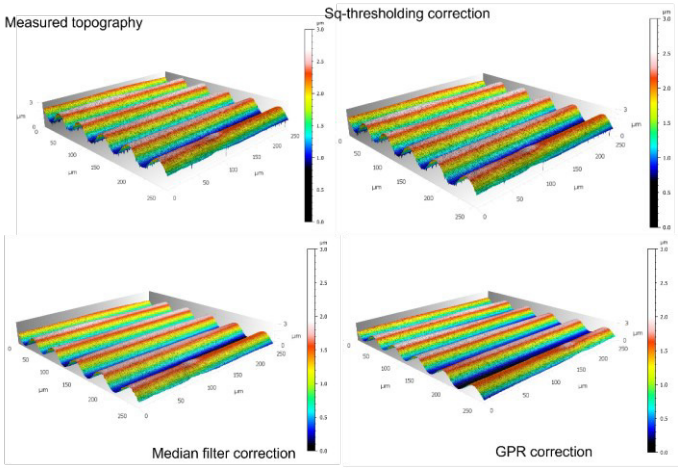

Fig. 7. Rubert specimen replica topography measured by CM. Original topography with several spikes. Notice poor performances of the thresholding correction.

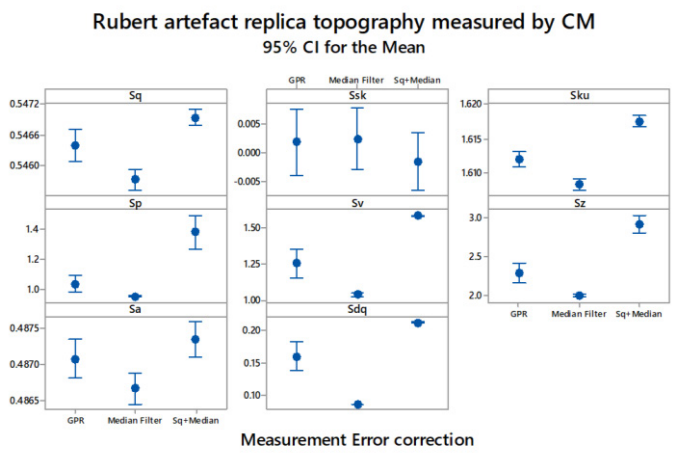

Fig. 8. Interval plot of the considered parameters of the Rubert artefact replica topography measured by the CM after the spikes and voids correction.

3.4. Discussion

As formerly partially reported in literature [30,32], different available methods to identify and correct measurement errors of surface topography measurements yield statistically different corrections, which significantly affect at a different severe extent the considered set of topographical parameters regardless of the instrument and the surface.

In particular, Sq-thresholding requires an operator-based implementation and performs worse in all the considered cases. In fact, when applied to a flat surface it over corrects spikes, yielding very small *Sz*. Conversely, when applied to a non-flat surface it tends to neglect smaller spikes resulting in excessively large *Sv* and *Sp*. This might seem in contrast with the previous case, but Sq-thresholding application to non-flat surface requires the correction to be carried out on the residual of the average. Correction is performed, and average added back. This operation cannot identify systematic spikes, thus resulting in the larger *Sz*. This is confirmed by the NPPs of Fig. 8(b) and (c) which provide insight on the distribution of the differences. In all cases, very relevant tails leading to hyponormal or multimodal distribution are present when the comparison is with the Sq-thresholding.

Median filtering and interpolation results in a correction less severe than the thresholding and more similar to GPR. However, it is liable of filtering relevant length scales, e.g. as shown by the systematically lower *Sdq*.

Conversely, GPR, as formerly validated in the literature [30], is capable of unbiased corrections, and its behavior yields satisfactory results in all the considered cases.

4. Conclusions

Measurement of surface topography is essential in several field and to verify the quality of surface technologies apt to engineer surface properties. Measurement errors, i.e., spikes and non-measured points, are generated due to critical and complex interactions between the measurand surface and the measuring instrument working principle. The literature has proposed several methods and recipes to manages measurement errors. This work addressed an unprecedented comparison of the effect of these methods considering different instruments and surface topographies. Amongst the several considered available methods, this work showed that machine learning based on gaussian process regression provides a formal, robust, automatic and operator-independent tool to correct measurement errors with satisfactory results independent on the measurement principle and measured surface topography.

Acknowledgements

This work was partially supported by supported by “Ministero dell’Istruzione, dell’Università e della Ricerca” Award “TESUN-83486178370409 finanziamento dipartimenti di eccellenza CAP. 1694 TIT. 232 ART. 6”.

References

[1] Leach R K Characterisation of Areal Surface Texture. Berlin: Springer; 2013.

[2] Leach R K Optical Measurement of Surface Topography. Berlin: Springer; 2011.

[3] Leach R K, Giusca C L, Haitjema H, Evans C, Jiang X Calibration and verification of areal surface texture measuring instruments CIRP Ann - Manuf Technol 2015 64:797–813.

[4] De Chiffre L, Kunzmann H, Peggs G N, Lucca D A Surfaces in precision engineering, microengineering and nanotechnology CIRP Ann - Manuf Technol 2003 52:561–77.

[5] Evans C J, Bryan J B `Structured', `textured' or `engineered' surfaces CIRP Ann - Manuf Technol 1999 48:541–56.

[6] Bewilogua K, Bräuer G, Dietz A, Gäbler J, Goch G, Karpuschewski B, et al. Surface technology for automotive engineering CIRP Ann - Manuf Technol 2009 58:608–27.

[7] Bruzzone A A G, Costa H L, Lonardo P M, Lucca D A Advances in engineered surfaces for functional performance CIRP Ann - Manuf Technol 2008 57:750–69.

[8] Polini W, Moroni G Manufacturing Signature for Tolerance Analysis J Comput Inf Sci Eng 2015 15:1–5.

[9] Maculotti G, Senin N, Oyelola O, Galetto M, Clare A, Leach R Multi-sensor data fusion for the characterisation of laser cladded cermet coatings. Eur. Soc. Precis. Eng. Nanotechnology, Conf. Proc. - 19th Int. Conf. Exhib. EUSPEN 2019, 2019.

[10] Loaldi D, Quagliotti D, Calaon M, Parenti P, Annoni M, Tosello G Manufacturing signatures of injection molding and injection compression molding for micro-structured polymer fresnel lens production 2018 9.

[11] Regi F, Doest M, Loaldi D, Li D, Frisvad J R, Tosello G, et al. Functionality characterization of injection moulded micro-structured surfaces Precis Eng 2019 60:594–601.

[12] Newton L, Senin N, Chatzivagiannis E, Smith B, Leach R Feature-based characterisation of Ti6Al4V electron beam powder bed fusion surfaces fabricated at different surface orientations Addit Manuf 2020 35:101273.

[13] Maculotti G, Piscopo G, Marchiandi G, Atzeni E, Salmi A, Iuliano L Build orientation effect on Ti6Al4V thin-wall topography by electron beam powder bed fusion Procedia CIRP 2022 Accepted for publication.

[14] Gu P, Zhu C, Mura A, Maculotti G, Goti E Grinding performance and theoretical analysis for a high volume fraction SiCp/Al composite J Manuf Process 2022 76:796–811.

[15] Maculotti G, Goti E, Genta G, Marchiandi G, Mura A, Mazza L, et al. Effect of track geometry on the measurement uncertainty of wear in pin-on-disc tribological test Proc 21st Int Conf Exhib EUSPEN 2021.

[16] Maculotti G, Goti E, Genta G, Mazza L, Galetto M Uncertainty-based comparison of conventional and surface topography-based methods for wear volume evaluation in pin-on-disc tribological test Tribol Int 2022 165:107260.

[17] Genta G, Maculotti G Uncertainty evaluation of small wear measurements on complex technological surfaces by machine vision-aided topographical methods CIRP Ann - Manuf Technol 2021 70.

[18] Eppinger S D, Huber C D, Pham V H A methodology for manufacturing process signature analysis J Manuf Syst 1995 14:20–34.

[19] Fang T, Jafari M A, Danforth S C, Safari A Signature analysis and defect detection in layered manufacturing of ceramic sensors and actuators Mach Vis Appl 2003 15:63–75.

[20] ISO 25178-6:2010 Geometrical product specifications (GPS) — Surface texture : Areal Part 6 : Classification of methods for measuring surface texture. ISO, Genève.

[21] Galati M, Minetola P, Rizza G Surface Roughness Characterisation and Analysis of the Electron Beam Melting (EBM) Process Materials (Basel) 2019 12:2211.

[22] Maculotti G, Feng X, Galetto M, Leach R Noise evaluation of point autofocus surface topography measuring instrument Meas Sci Technol 2018 29.

[23] Maculotti G, Feng X, Su R, Galetto M, Leach R Residual flatness and scale calibration for a point autofocus surface topography measuring instrument Meas Sci Technol 2019 30.

[24] Thompson A, Senin N, Giusca C, Leach R Topography of selectively laser melted surfaces: A comparison of different measurement methods CIRP Ann - Manuf Technol 2017 66:543–6.

[25] ISO 25178-600:2019 Geometrical product specification (GPS) - Surface texture: Areal Part 600: Metrological characteristics for areal-topography measuring. ISO, Genève.

[26] Giusca C L, Leach R K, Helary F, Gutauskas T, Nimishakavi L Calibration of the scales of areal surface topography measuring instruments: part 1. Measurement noise and residual flatness Meas Sci Technol 2012 23:065005.

[27] Gomez C, Su R, Thompson A, DiSciacca J, Lawes S, Leach R Optimization of surface measurement for metal additive manufacturing using coherence scanning interferometry Opt Eng 2017 56:111714.

[28] Alburayt A, Syam W P, Leach R Lateral scale calibration for focus variation microscopy Meas Sci Technol 2018 29.

[29] Tosello G, Haitjema H, Leach R K, Quagliotti D, Gasparin S, Hansen H N An international comparison of surface texture parameters quantification on polymer artefacts using optical instruments CIRP Ann - Manuf Technol 2016 65:529–32.

[30] Maculotti G, Genta G, Quagliotti D, Galetto M, Hansen H N Gaussian process regression-based detection and correction of disturbances in surface topography measurements Qual Reliab Eng Int 2021:1–18.

[31] Quagliotti D Modeling the systematic behavior at the micro and nano length scales Surf Topogr Metrol Prop 2022 10:015011.

[32] Medeossi F, Carmignato S, Lucchetta G, Savio E Effect of void pixels on the quantification of surface topography parameters Proc 16th Int Conf Eur Soc Precis Eng Nanotechnology, EUSPEN 2016 2016:2–3.

[33] Podulka P, Pawlus P, Dobrzański P, Lenart A Spikes removal in surface measurement J Phys Conf Ser 2014 483.

[34] Mountains Map. www.digitalsurf.com.

[35] Maculotti G, Pistone G, Vicario G Inference on errors in industrial parts: Kriging and variogram vs geometrical product specification standard Appl Stoch Model Bus Ind 2021 37:839–58.

[36] Gelbart M A, Snoek J, Adams R P Bayesian optimization with unknown constraints Uncertain Artif Intell - Proc 30th Conf UAI 2014 2014:250–9.

[37] Galetto M, Genta G, Maculotti G Single-step calibration method for nano indentation testing machines CIRP Ann 2020 69:429–32.

[38] ISO ISO 5436-1:2001 Geometrical product specifications (GPS) — Surface texture: Profile method; Measurement standards — Part 2: Software Measurement Standards 2012. ISO, Genève.

[39] ISO 25178-2:2012 Geometrical product specifications (GPS) — Surface texture : Areal Part 2 : Terms , definitions and surface. ISO, Genève.

[40] Montgomery D, Runger G, Hubele N Engineering statistics. New York: John Wiley & Sons Inc.; 2010.