Journal of Physics: Conference Series

PAPER • OPEN ACCESS

Statistical evaluation of hard-to-measure surfaces

To cite this article: M Kubišová *et al* 2024 *J. Phys.: Conf. Ser.* **2712** 012020

Statistical evaluation of hard-to-measure surfaces

M Kubišová[1], J Knedlová[1], H Vrbová[1], V Pata[1] and B Bočáková[2]

[1]Tomas Bata University in Zlín, Faculty of Technology, Vavreckova 5669, 760 01 Zlín, Czech Republic

[2]Slovak University of Technology in Bratislava, Faculty of Materials Science and Technology in Trnava, Ulica Jána Bottu 2781/25, 917 24 Trnava, Slovak Republic

E-mail: mkubisova@utb.cz

Abstract. The main goal of this article will be to find ways to evaluate hard-to-measure surfaces statistically. First, the basic characteristics and rules of surface quality will be described according to the standards ČSN EN 4287, ČSN EN 4288 and ČSN EN ISO 2517-2. Subsequently, the measured values of the roughness parameter Sa (arithmetic average of the height of the measured surface) and Sz (the maximum height of the measured surface) will be compared and evaluated which is the best. These parameters will be described and measured on aluminium plates on which the test surfaces were laser engraved. To evaluate the best surface, statistical methods will be used, such as the EDA methodology (exploratory data analysis), hypothesis testing with normality and outlier tests, and last but not least, cluster or cluster analysis, which compares the similarity of the measured data. This article aims to show the possibilities of surface quality assessment using 3D surface roughness parameters, which are not often used in practice.

1. Introduction

Analyzing the surface texture and structure of a material is crucial for many applications in industry, engineering and scientific research. Standardly, microscopy is used first, which is the classic method for visual analysis of the surface texture and structure of the material. A light microscope can provide high-resolution images of a surface, while an electron microscope (SEM) or atomic force microscope (AFM) allows characters to be observed at the atomic level [1].

This is followed by inspection and measurement using a profilometer when the height profile of the surface is measured. This technique is useful for quantitative evaluation of surface roughness and geometric properties. A profilometer can be performed by contact or non-contact methods [1, 2].

Surface quality assessment using 3D surface roughness parameters is an important process in industry, engineering and scientific research. 3D surface roughness parameters allow for a more detailed description and analysis of the surface texture and structure of the material. There are several different parameters that can be used to evaluate surface roughness. The most frequently mentioned parameter is *R*a - the average value of surface deviations from its mean plane. A higher *R*a value indicates a theoretically rough surface. Another frequently mentioned parameter is *R*z when the height differences between the highest and lowest point on the surface are measured. This value provides information on the height distribution of the roughness. The parameter *S*k (roughness kurtosis) can also be used. Kurtosis measures the shape of the distribution of deviations from the mean. A higher kurtosis value may indicate significant peaks and valleys in the surface. Last but not least, *S*mr (directivity) is used, the directivity shows the direction of preferred deviations on the surface and can be important for specific applications [2].

For surface quality assessment, it is often necessary to combine several of these parameters to provide more complete and accurate information. Surface roughness measurements can be made using various devices such as profilometers, interferometers and scanning microscopes. Analysis of these parameters allows engineers and manufacturers to optimize processes and achieve the desired surface quality for multiple applications [3].

The parameters *S*a (the arithmetic average of the height of the measured surface) and *S*z (the maximum height of the measured surface) are two basic measures that are used to characterize the surface roughness of a material using a profilometer or other surface texture measurement methods. These parameters provide important information about surface roughness and unevenness [1, 3].

*S*a (Arithmetic average of surface height) represents the average height of deviations on the surface of the material from its mean plane. This parameter is calculated as the sum of the absolute values of the height deviations of the surface (regardless of the direction) divided by the area of the measured surface. It expresses the roughness of the surface, where higher values indicate a rougher surface and lower values a smoother surface (Figure 1) [1, 3].

*S*z (Maximum Surface Height) represents the maximum height of deviation on the surface of the material from its mean plane. It is the highest point on the surface (the highest peak) minus the lowest point on the surface (the lowest trough). *S*z is useful for identifying the most prominent surface irregularities and measuring the maximum protrusion or depression (Figure 1) [1, 3].

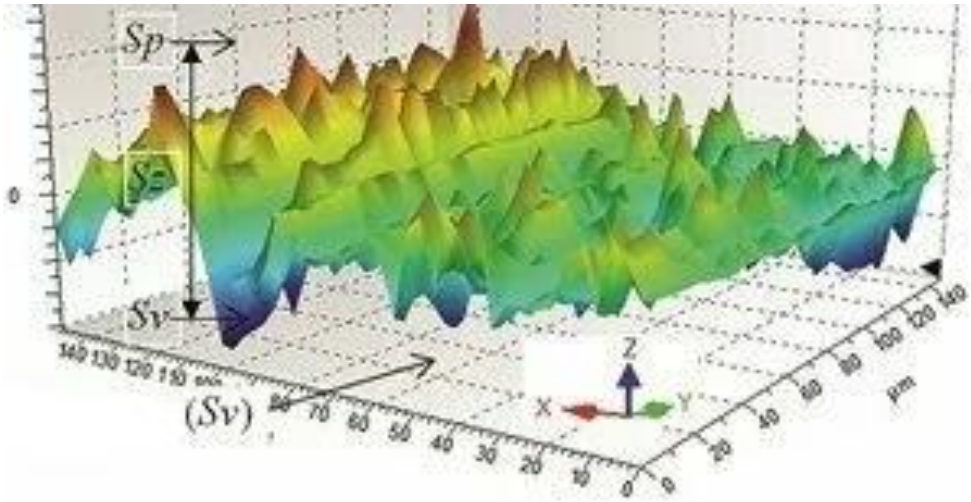

Figure 1. Measurement of parameters *S*a and *S*z.

Both parameters *S*a and *S*z are expressed in the same units as the measured surface (for example, micrometres, nanometers). They are used as part of the quantitative analysis of surface roughness and allow different materials or surfaces to be compared and characterised in different applications. These parameters can be used together with other 3D surface roughness parameters for more detailed analysis and comparison of surface texture and material structure [3].

The EDA methodology (Exploratory Data Analysis) is a statistical approach to data analysis that serves to discover and understand the basic characteristics of data and relationships between variables. The EDA methodology was developed by John Tukey in the 1970s and is still an important tool for data science, statistics and data analysis. Here is a basic overview of the EDA methodology:

Data Acquisition: The first step in the EDA methodology is to acquire the necessary data. This may include collecting data, importing data files or accessing existing data.

Data preprocessing: Data preprocessing is a key step before the actual analysis. It includes activities such as removing missing values, normalizing data, identifying and removing outliers, and transforming variables if needed [4].

Data visualization: Visualization is one of the main elements of the EDA methodology. Charts and graphical techniques can be used to visually display data and identify patterns, trends, and anomalies. Visualization types include histograms, scatter plots, box plots, density plots, and more.

Data Summarization: Data summarization involves creating descriptive statistics such as mean, median, variance, quartiles, and more. This statistic helps to get an idea of the basic characteristics of the data. EDA is an important tool in the data analysis process and helps analysts better understand data and prepare for advanced analysis such as modelling and inferential statistics. It provides insight into which variables are important in the analysis [4, 5].

The applied materials must satisfy the strictest standards in terms of extended durability, wear, or economics in order to maintain the existing production level. The use of unconventional technologies is used to address the high demands on these materials' processability because the usual methods frequently fall short of expectations for production speed and quality. Therefore, procedures that produce better results faster, like laser machining, should be used [6].

A number of things can influence the laser cutting of material. Evaluation of the effects of specific factors is the outcome of the optimization process. To set the laser device's parameters in a way that maximizes the product's capabilities, it is necessary to emphasize the key effects, ignore less significant effects, and select general laser device parameters [7].

This technology has great productivity and manufacturing efficiency and is quick, low-tech, and extremely accurate. A laser is the best tool in this situation because of its characteristics. The use of the laser beam is practically universal in the processing of materials [8, 9].

The purpose of this publication is to assess surfaces that have been etched using various lasers at various settings. The surfaces were evaluated using statistical techniques, making it feasible to determine which surface was the best. Zygo New View 9000, a non-contact 3D scanner, was used to scan the surfaces.

2. Materials, Methods and Results

2.1. Sample preparation

The measured roughness parameters *S*a (arithmetic average of the height of the surface) represent the mean value of the heights of the surface irregularities on the given sample. It is a measure of surface roughness, which is calculated as the average absolute value of the height deviations from the mean value of the surface. *S*z (maximum surface height) represents the highest height of surface irregularities on a given sample. This is the maximum deviation of the surface height from its mean value, marked with ordinal numbers, representing samples, marked with the same ordinal numbers in the test column, which were engraved with different lasers with different settings (Table 1).

Table 1. Marking of samples, types of lasers and machining settings.

Sample	Type of Laser	Engraving time [s]	Engraving Speed [mm/s]	Laser Frequency [kHz]	Laser Power [W]
01	Fiber Fly 30W	18.3	1100	30	20
02	Fiber Fly 30W	17.1	1200	30	15
03	Fiber Fly 30W	23.8	800	30	50
04	Fiber Fly 30W	23.8	800	30	50
05	Fiber Fly 30W	23.8	800	30	50
06	Fiber Fly 30W	23.7	800	30	50
07	Fly Air Green Wave	15.3	1200	10	65
08	Fly Air Green Wave	14.6	1200	10	80
09	Fiber Fly 50W Pico	19.4	1000	500	50
10	Fly Air Green Wave	19.4	1000	500	90

This measurement was performed with a non-contact Zygo new view Nx roughness meter, a device designed to measure surface roughness. The results of *S*a (Table 2) and *S*z (Table 3) measurements can be important in determining whether the surface properties of the samples are suitable for a given application. Figure 2 shows the 3D mapping of two machined samples. First, the form and then the waviness was taken out of the measured surface profile, leaving a 4287-compliant roughness profile that was created in sections Y independently (North-South) and in particular in the X-axis (East-West). The computer-generated a 3D surface profile based on standard 25178. Finally, values for the roughness parameter were produced in accordance with the 4287 requirements.

Table 2. Measured data of parameter *S*a [µm].

Sample	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
*S*a_01	0.777	0.861	0.751	0.813	0.763	0.824	0.864	0.842	0.789	0.842
*S*a_02	0.570	0.570	0.577	0.558	0.595	0.608	0.559	0.598	0.554	0.588
*S*a_03	0.664	0.624	0.713	0.642	0.685	0.695	0.705	0.675	0.697	0.632
*S*a_04	0.391	0.373	0.406	0.488	0.401	0.405	0.381	0.409	0.389	0.405
*S*a_05	0.416	0.442	0.411	0.406	0.328	0.444	0.414	0.426	0.436	0.450
*S*a_06	0.256	0.293	0.299	0.286	0.488	0.317	0.302	0.311	0.323	0.282
*S*a_07	0.397	0.367	0.418	0.353	0.381	0.394	0.417	0.355	0.368	0.370
*S*a_08	0.463	0.492	0.499	0.483	0.463	0.499	0.473	0.479	0.500	0.471
*S*a_09	0.453	0.459	0.465	0.739	0.739	0.462	0.460	0.444	0.460	0.464
*S*a_l0	0.448	0.450	0.449	0.457	0.448	0.477	0.472	0.465	0.468	0.474

Table 3. Measured data of parameter *S*z [µm].

Sample	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
*S*z_01	14.398	14.412	14.413	14.397	14.413	14.400	14.408	14.418	14.401	14.392
*S*z_02	15.212	15.359	15.415	15.485	15.398	15.482	15.359	15.362	15.465	15.296
*S*z_03	10.049	10.168	10.072	10.035	10.056	10.036	10.133	10.091	10.034	10.076
*S*z_04	5.271	5.411	5.347	5.362	5.339	5.345	5.304	5.400	5.327	5.364
*S*z_05	6.788	6.824	6.852	6.880	6.796	6.793	6.814	6.818	6.815	6.820
*S*z_06	5.174	5.169	5.180	5.200	5.188	5.183	5.160	5.185	5.202	5.176
*S*z_07	8.851	8.856	8.859	8.848	8.848	8.856	8.842	8.857	8.853	8.842
*S*z_08	6.181	6.186	6.185	6.185	6.174	6.172	6.186	6.177	6.176	6.182
*S*z_09	5.554	5.552	5.523	5.484	5.498	5.493	5.552	5.523	5.523	5.569
*S*z_l0	8.991	8.916	8.897	8.944	9.085	9.083	9.068	8.890	8.966	8.948

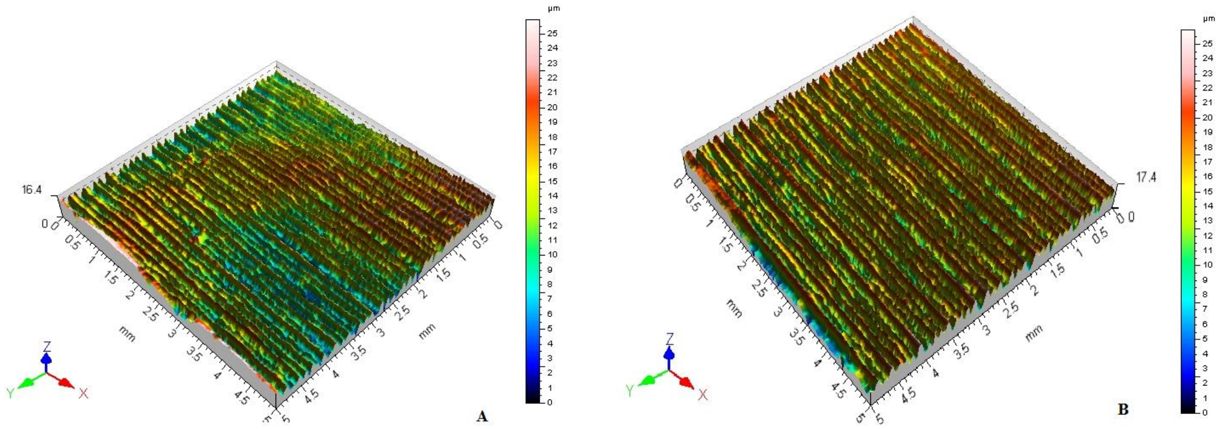

Figure 2. Samples machined a) FiberFly VP 30W; b) FlyAir GreenWave.

2.2. EDA methodology summary report parameter Sa

When the null hypothesis states that the data come from a normal distribution and the alternative hypothesis states that the data do not come from a normal distribution, it is possible to say that the measured data *S*a_1 exhibits a normal distribution. This is the case with the Anderson-Darling test of normality. Because the result of $p = 0.564$ is higher than the margin of error of 0.05, we may say that we do not reject the null hypothesis with a margin of error of 5 %. Figures 3 and 4 show the results of sample 1 for parameter *S*a. Table 4 then shows all the results for the *S*a parameter.

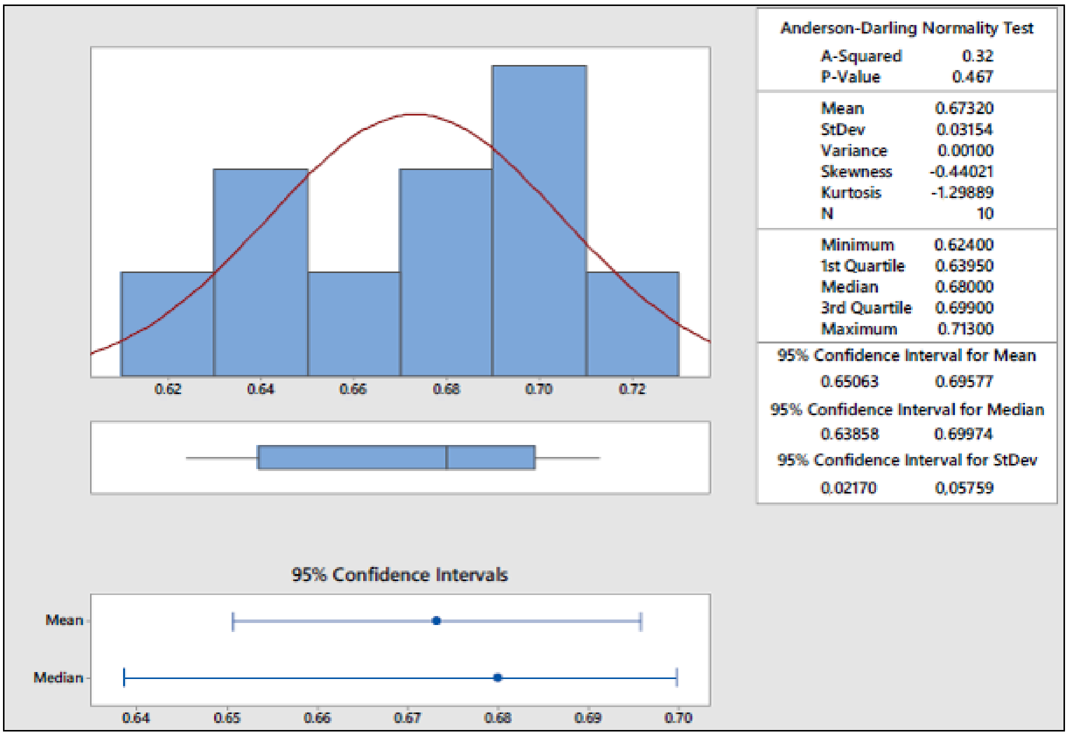

Figure 3. Graphic representation evaluated by EDA for parameter *S*a_01.

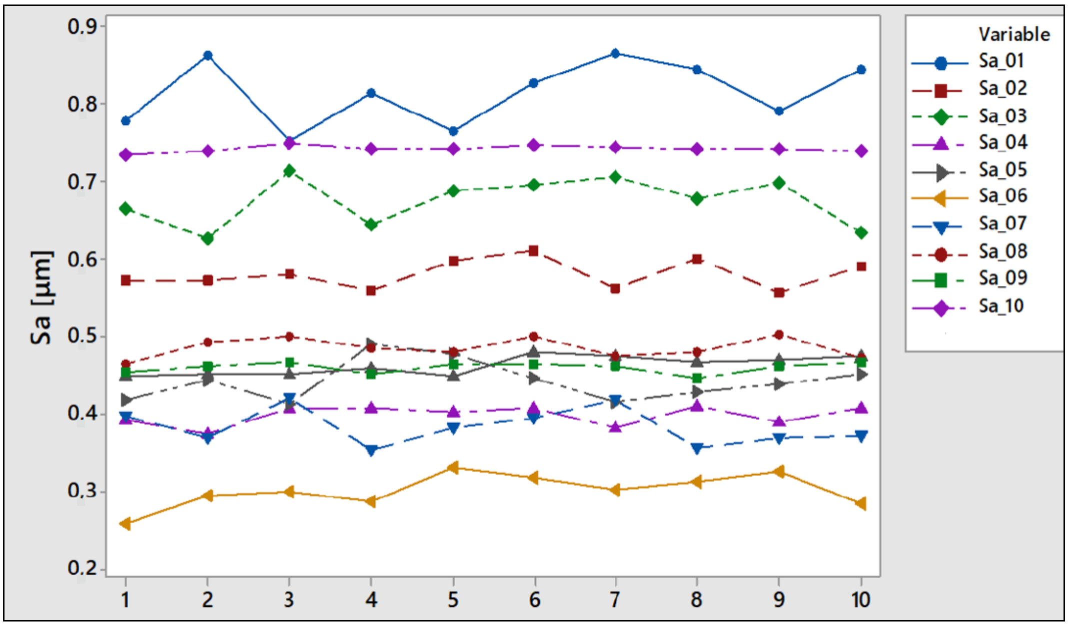

Figure 4. Time series graph of parameter *S*a.

Table 4. The results of the parameter *S*a [μm].

Sample	Ar. Mean	St. Dev.	Min.	Ql	Median	Q3	Max.	Range
*S*a_01	0.813	0.041	0.751	0.774	0.819	0.847	0.864	0.113
*S*a_02	0.578	0.019	0.554	0.559	0.574	0.596	0.608	0.054
*S*a_03	0.673	0.032	0.624	0.640	0.680	0.699	0.713	0.089
*S*a_04	0.397	0.012	0.373	0.387	0.403	0.406	0.409	0.036
*S*a_05	0.440	0.026	0.411	0.415	0.439	0.456	0.488	0.770
*S*a_06	0.300	0.022	0.256	0.285	0.301	0.319	0.328	0.072
*S*a_07	0.382	0.024	0.353	0.364	0.376	0.402	0.418	0.065
*S*a_08	0.484	0.013	0.463	0.473	0.481	0.499	0.500	0.037
*S*a_09	0.458	0.007	0.444	0.452	0.460	0.463	0.465	0.021
*S*a_10	0.740	0.004	0.733	0.738	0.739	0.743	0.747	0.014

A test for outliers was performed for all *S*a parameters, the data were sorted by size, and in this case, the null hypothesis says that all data come from a normal distribution, in contrast, the alternative hypothesis says that the smallest and largest value is an outlier, the significance level is set to 0.05. *P* values that came out more than 0.05 do not reject the null hypothesis.

2.3. EDA methodology summary report parameter Sz

It is possible to state that the measured data of *S*z_01 shows a normal distribution, in the case of the Anderson-Darling normality test, where the null hypothesis says that the data comes from a normal distribution, while the alternative hypothesis stands, which says that the data does not come from a normal distribution. With the possibility of an error of 5 %, it is possible to state that we do not reject the null hypothesis because the value of $p = 0.364$ is greater than the possibility of an error of 0.05. Furthermore, the values of the arithmetic mean, standard deviation, skewness, and kurtosis were calculated, the values of the minimum, first quartile, median, third quartile, and maximum were displayed, and the confidence intervals for the arithmetic mean, median, and standard deviation were also calculated. The findings of sample 1 for parameter *S*z are displayed in Figures 5 and 6. The whole set of *S*a parameter data is then shown in Table 5.

The level of significance was set at 0.05, and all roughness parameters *S*z were checked for outliers. In this case, the null hypothesis states that all data come from a normal distribution, while the alternative hypothesis states that the value with the smallest and largest value is an outlier. *P* values larger than 0.05 are not considered to be a rejection of the null hypothesis.

Table 5. The results of the parameter *S*z [μm].

Sample	Ar. Mean	St. Dev.	Min.	Ql	Median	Q3	Max.	Range
*S*z_01	14.405	0.009	14.392	14.398	14.404	14.413	14.418	0.026
*S*z_02	15.383	0.086	15.212	15.343	15.38	15.469	15.485	0.273
*S*z_03	10.075	0.045	10.034	10.036	10.064	10.101	10.168	0.134
*S*z_04	5.347	0.042	5.271	5.321	5.346	5.373	5.411	0.140
*S*z_05	6.820	0.028	6.788	6.795	6.817	6.831	6.880	0.092
*S*z_06	5.182	0.013	5.160	5.173	5.182	5.191	5.202	0.042
*S*z_07	8.351	0.006	8.842	8.847	8.852	8.856	8.859	0.017
*S*z_08	6.180	0.005	6.172	6.176	6.182	6.182	6.186	0.014
*S*z_09	5.927	0.029	5.484	5.497	5.523	5.523	5.552	0.085
*S*z_l0	8.979	0.075	8.89	8.911	8.957	9.072	9.072	0.195

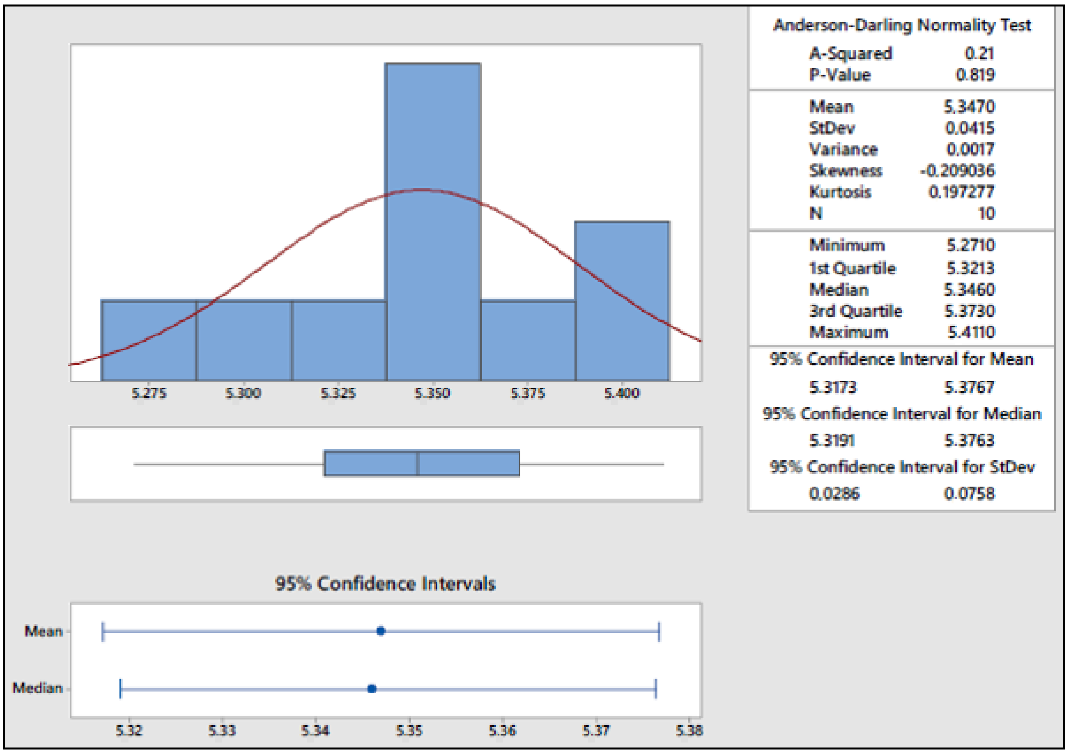

Figure 5. Graphic representation evaluated by EDA for parameter *Sz*_01.

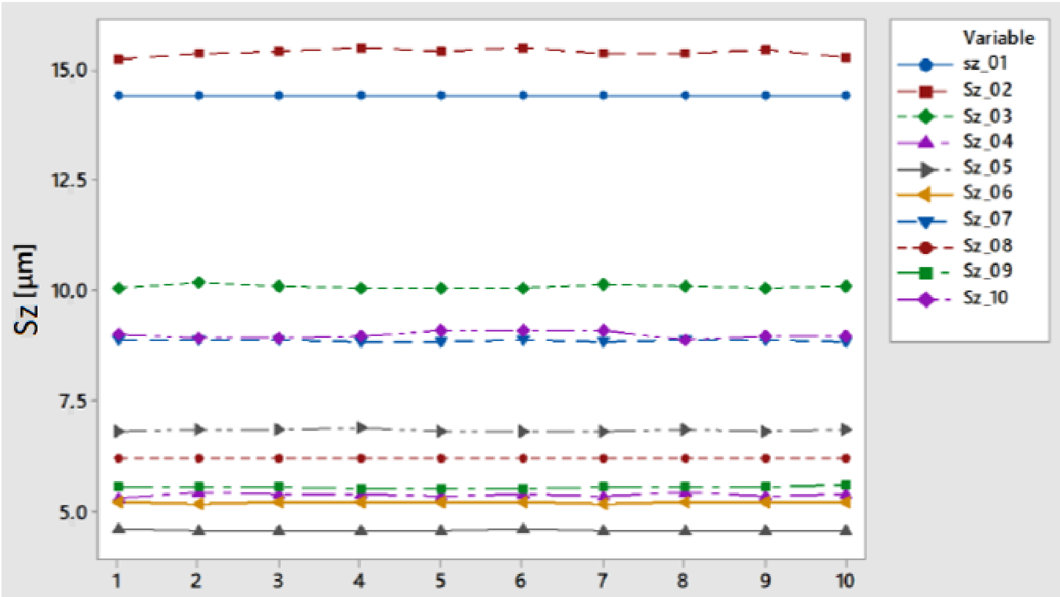

Figure 6. Time series graph of parameter *Sz*.

2.4. Cluster analysis of measured data

From the cluster analysis of the *S*a parameter data (Figure 7), it can be said that *S*a_03 and *S*a_10 have the greatest similarity, which corresponds to the settings of the FiberFlyVp30W laser and its corresponding settings listed in Table 1 and the FiberFly 50W Pico laser and its corresponding settings listed in Table 1. It is also possible to say that the best surface evaluated according to the parameter *S*a according to the previous analysis *S*a_6 corresponding to the laser setting 06 is similar to the 78.08 % roughness of *S*a_08 and the corresponding setting and type of laser.

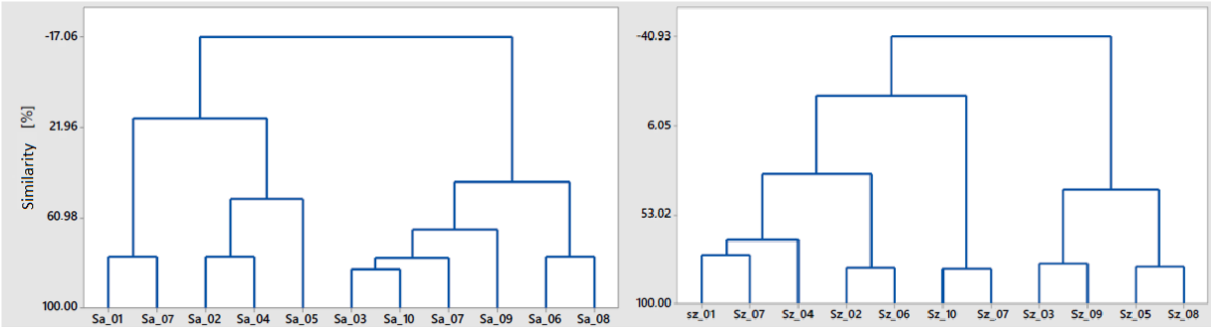

Figure 7. Cluster analysis for all roughness parameters *S*a and Sz.

*S*z_10 and *S*a_07, which correspond to the settings of the FiberFly 50W Pico laser and its related settings stated in Table 1, have the most similarity, according to the cluster analysis of the *S*z parameter data.

3. Conclusions

The aim of this manuscript was to compare the measured values of the roughness parameter *S*a (arithmetic mean of the height of the measured surface of surface) and *S*z (the maximum height of the measured surface) and to evaluate which is the best. These parameters were measured on aluminium plates on which the test surfaces were laser engraved. The test surfaces were engraved with different types of lasers with different settings. Statistical methods were used to evaluate the best surface, such as the EDA methodology (exploratory data analysis), hypothesis testing with normality and outlier tests, and last but not least, cluster analysis, which compared the similarity of the measured data.

For the roughness parameter *S*a, sample 06 performed best, it shows the lowest arithmetic mean of the measured values. Although there is skew in the data for sample 06, the normality test showed that the data came from a normal distribution, and the outlier test showed no outliers. The graph of the time series of the roughness parameter *S*a clearly shows that sample 06 shows the smallest value of all measured parameters *S*a. In cluster analysis, sample 06 most closely resembles the course of sample 08. Sample 06 corresponds to the FiberFly VP 30W laser with the parameters listed in Table 1, and sample 08 corresponds to the FlyAir green Wave laser with the parameters listed in Table 1.

For the roughness parameter *S*z, sample 11 performed best, it shows the lowest arithmetic mean of the measured values. Sample 11 also shows slight skewness, but the normality test showed that the data came from a normal distribution, and the outlier test showed no outliers. The graph of the time series of the roughness parameter *S*z clearly shows that sample 11 shows the smallest value of all measured parameters *S*z. In cluster analysis, sample 10 is most similar to sample 9. Samples 10 and 10 correspond to the FiberFly50WPico laser with the parameters listed in Table 1.

If we assess the engraved surfaces according to the roughness parameter *S*a, then the surface on sample 06 shows the best values. If we assess the engraved surfaces according to the parameter *S*z, then the values from sample 11 and thus also the lasers used according to Table 1 stood the best.

Acknowledgements

This work and the project are realized with the financial support of the internal grant of TBU in Zlin No. IGA/FT/2023/004 funded from the resources of specific university research

References

[1] Whitehouse D J 2011 *Handbook of Surface and Nanometrology 2nd ed*. (Boca Raton: CRC Press) ISBN 978-1-4200-8201-2

[2] Meloun M, Militký M and Forina M 1992 *Chemometrics for Analytical Chemistry* (New York: Ellis Horwood) ISBN 9780131263765

[3] Murat D and Ensarioglu C 2017 Surface roughness analysis of greater cutting depths during hard turning *Mater. Test.* **59** 9 795–802
[4] Abonyi J and Feil B 2007 *Cluster Analysis for Data Mining and System Identification 1st ed* (Basel: Birkhäuser Basel)
[5] Fu G 2019 A deep-learning-based approach for fast and robust steel surface defects classification *Opt. Lasers Eng*. **121** 397–405
[6] Bilodeau M and Brenner D 1999 *Theory of multivariate statistics* (New York: Springer)
[7] Reiss R-D and Thomas M. 2001 *Statistical analysis of extreme values: with applications to insurance, finance, hydrology and other fields 2nd ed* (Basel: Birkhäuser Verlag) ISBN 3-7643-6487-4
[8] Zhang J Feng Ch, Ma Y, Tang W, Wang S and Zhong X 2017 A Mechanistic Model for Prediction of Cutting Parameters in Micro-Scale Milling *Manuf. Technol.* **17** 3 412–418
[9] Hanzl P, Zetková I and Mach J 2017 Optimization of the Pressure Porous Sample and Its Manufacturability by Selective Laser Melting *Manuf. Technol.* **17** 1 34–38