

Comparing Qualitative Raster Maps

WIECZOREK Małgorzata ^{1, a} and PRZYBYŁ Wojciech ^{2, b}

¹ University of Wrocław, Plac Uniwersytecki 1, 50-137 Wrocław, Poland

² Military Institute of Engineer Technology, ul. Obornicka 136, 50-961 Wrocław, Poland

^a malgorzata.wieczorek@uwr.edu.pl, ^b przybyl@witi.wroc.pl

Keywords: Raster Maps, Validation, Filters

Abstract. Spatial diversity of the natural environment can be presented using raster qualitative data. They can be the result of collecting field data or be the result of stochastic modelling of a certain complexity of the environment. One example of such modelling is the determination of similar elements of relief forms on the basis of morphometric variables. The use of unsupervised methods for clustering raster data in modelling can produce different maps. It is necessary to assess the compatibility of the obtained maps in order to assess how significant the differences between them are. The article presents selected stochastic and deterministic methods for assessing the spatial distribution of data. Exemplary methods of measuring variability at the global and local level were discussed.

Introduction

In the era of relatively easy access to numerical data, including raster data, both qualitative and quantitative, the attention of scientists is directed not so much to the ways of obtaining data, but to the ways of using them, analyzing, modelling, and then assessing the results of their validation. Currently, the application of image analysis and classification (i.e. rasters) is used in many fields. The most common methods include the analysis of satellite and meteorological images, morphometric classification, medical image analysis (including tomograms, roentgenograms, ultrasound images, etc.), monitoring of product wear and quality (such as metal products, fabrics, paper, and foodstuffs), as well as analysis of surface images and microscopic images of material structures to monitor tool wear.

Algorithms have already been developed for recognizing objects in satellite images or diagnostic images used in medicine, often using neural networks or fuzzy logic. Comparing two images and assessing whether they are statistically [1] the same when these are modeling results and there is no pattern to relate to is still a challenge, although the topic has already been addressed in the literature [2]. This paper presents two such research situations along with a proposal of measures to assess the similarity of two images. The obtained results and limitations resulting from the measures used were discussed. In the article, the authors focused on exemplary global and local indicators that allow the assessment of differences between two images.

The methodology presented in this article goes beyond its applications solely in the field of topographic maps, as image analysis, particularly for images featuring stochastic fields, can be widely applied in industry and management, significantly impacting product quality enhancement [3-5], as well as modifying management methods [6,7] and related analytics [8,9]. The applications of such analyses encompass various issues, ranging from the assessment of corrosion phenomena [10], both in reference to typical metallic construction materials [11-13] and special alloys [14], to the visual evaluation of separator performance [15], and even to the automatic assessment of surfaces [16], coatings [17-19], and structural welds [20]. Currently, this significantly influences the organization of enterprises that have surpassed the level of Industry 3.0 and are reaching Industry 4.0, impacting the means of automatic control, for instance, in the application of DLC



[21-23] and ESD [24-26] coatings, as well as in the automation of thermal insulation evaluation using infrared cameras [27-29]. It is also a significant tool supporting the detection of machine damage [30,31], particularly hydraulic power equipment [32-34].

Automated methods for processing image maps are also becoming a vital tool in process optimization [35-37], including the application of qualitative and/or subjective factors [38-40], although this typically requires prior strong dimensionality reduction [41,42] to avoid unwanted correlations.

Area of Study

A fragment of the Owl Mountains (a mountain range in the Sudetes in central-southern Poland, located in the Lower Silesian Voivodeship) was selected for the research area with the intervention introduced so as to obtain a flat area. It is an area of 8,725 m in the E-W direction and 6,725 m in the N-S direction (269 rows x 349 columns). Using a digital elevation model (DEM) with a field resolution of 25 m, a k-median clustering with Manhattan metrics method was developed. This grouping was based on morphometric variables calculated from DEM, such as elevation, slope, aspect, curvature.

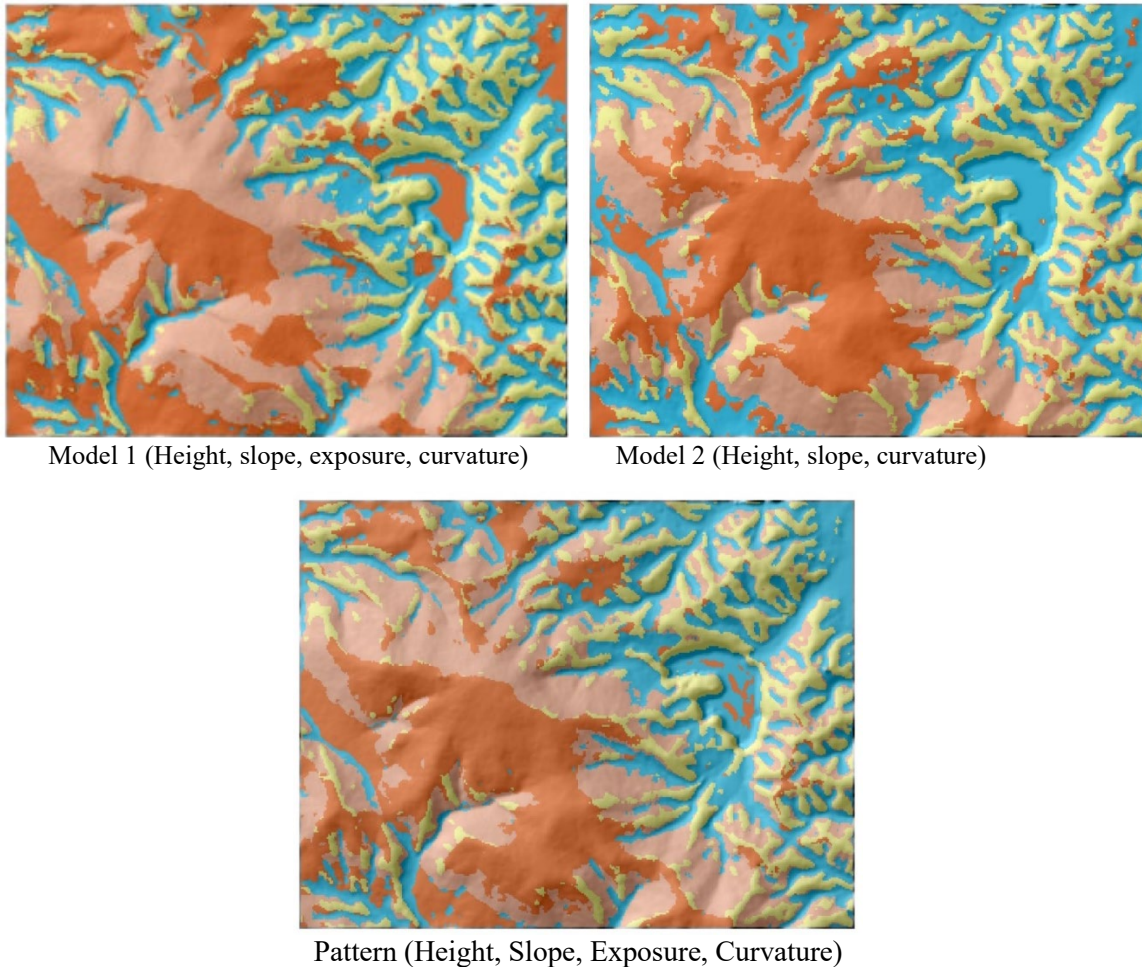


Fig.1. Models and pattern of the area subject to classification, divided into classes

Three images (Fig.1) were selected for the calculations, broken down into 4 classes (clusters) of landform elements: Model 1 – developed on the basis of all 4 morphometric variables, Model 2 – obtained on the basis of height, inclination and curvature, and the third considered as a Pattern - model obtained similarly to model 1, but characterized by the greatest intergroup variability (sum of square between clusters). Exposure is not a morphometric variable that is commonly included

in modelling. Having a tool (apart from the expert's assessment) to assess to what extent the inclusion of an exposure gives a significantly different modelling effect, it would be helpful in assessing the legitimacy of its inclusion in the model. An expert assessment of the advantages and disadvantages of modelling elements of sculpture forms with exposure was presented by Wiczorek and Migoń [43], but advanced similarity measures were not used in this assessment.

Problem to Solve

In this case, we are interested in the answer to the question: How to assess which model fits the pattern better? Although they are not identical, is the scale of similarity large enough to consider them identical models? Where to put this limit of assumed similarity? Obtaining answers to such questions would allow us to decide whether it is worth dividing into classes using 3 or 4 morphometric variables as input.

Another situation in which a measure of statistical significance would be needed in determining the differences between two rasters with qualitative data is the processing of maps by means of filtering. The key question is: After how many filters do we get an image that is too smooth? Fig.2 shows the pattern and its modification by carrying out the filtering procedure 5 times. The agreement between the two images is 92.6% (explained later in the article). The parameters set during filtering (the size of the filtering window and the filtering method), the number of filters and the homogeneity of the image are also important. The smaller information noise, the greater the differences between the original and the final image – however, we leave these considerations for another research study.

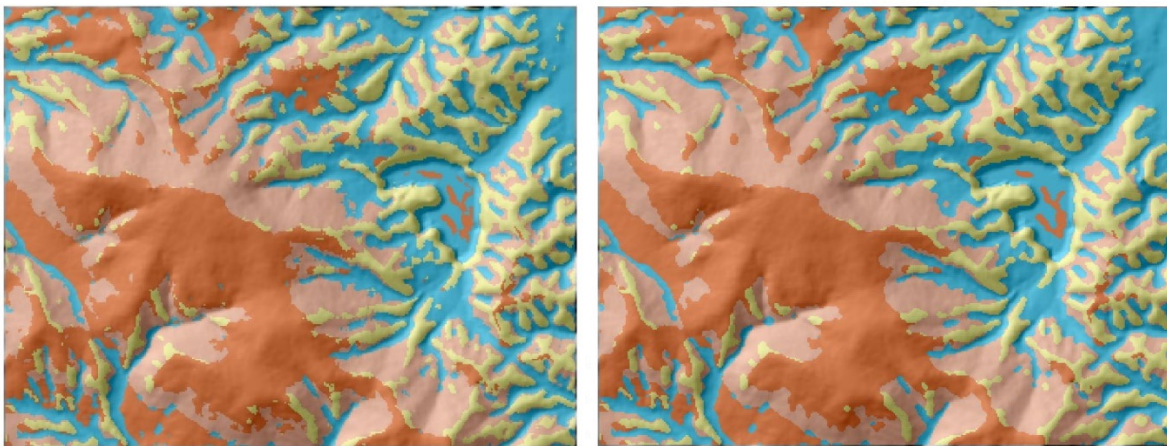


Fig.2. Pattern and its modification after 5-fold majority filtration (majority in 3x3 window).

Research Procedure

Map comparison procedures are derived from four spatial data analysis traditions:

1. An accuracy rating that characterizes a match (or mismatch) between a reference map considered to be accurate and one or more of its approximations.
2. Detection of changes that are interpreted as a function of time.
3. Model comparison, where the (predicted or simulated) model results are compared with observed data and/or with other model results (pixel-to-pixel consistency is not expected here).
4. The last tradition is to compare the landscape. The key feature of this approach is that the comparison is undertaken against one or more spatial metrics computed from the maps.

In practice, there are a number of methods used to compare and measure differences between raster images. They can be divided into: objective methods - using image descriptors based on mathematical models, and subjective methods – based on observations made by expert observers.

Three methods were selected in the article: differential, using Cohen's kappa coefficient, and based on the Haralick event matrix (Co-occurrence Matrix). These methods were applied globally – to the entire area and locally – to selected fragments of areas. Different areas and adjacent areas were deliberately selected to test the effectiveness of the measures for different types of terrain.

Global Measures

The simplest approach to finding differences between two raster maps is to subtract one raster from the other and mark the differences. The results of such subtraction are shown in Fig.3, where the differences between the pattern and Model1 are marked in green, and the differences between the pattern and Model2 in yellow. The overall agreement was 62% for Model1 and 76% for Model2, respectively (Table 1). Setting a threshold for the allowable number of differing pixels is a half-way solution, because there is still the question of distributing these differences in a spatial context.

Table 1. Value of differences between the models and the standard based on Cohen's kappa coefficient

Model 1		Model 2	
Cohen's kappa agreement coefficient	agreement	Cohen's kappa agreement coefficient	agreement
0.71	62%	0.82	76%

Table 2. Values of textural properties for Pattern, Model 1 and Model 2

Property	Model 1	Standard	Model 2
Angular Second Moment	0.1913	0.1895	0.1864
Contrast	0.4885	0.3595	0.3797
Correlation	0.8067	0.8563	0.8496
Sum of Squares	1.2638	1.2513	1.262
Inverse Difference Moment	0.9036	0.9104	0.9073
Sum Average	2.9844	2.9951	3.1949
Sum Variance	4.5665	4.6457	4.6683
Sum Entropy	2.3805	2.4063	2.421
Entropy	2.7962	2.7874	2.8102
Difference Variance	0.1228	0.1228	0.121
Difference Entropy	0.8035	0.7724	0.7928
Information Measures of Correlation I	-0.5898	-0.5966	-0.5861
Information Measures of Correlation II	0.9498	0.9513	0.9492

The second proposed global measure is Cohen's Kappa Coefficient [44] – determined according to the equation [1]. This coefficient takes values from -1 to 1. The closer the value is to 1, the greater the agreement between the model and the standard. For the study area, Cohen's kappa coefficient was 0.71 for Model1 and 0.82 for Model2 (Table 1). Based on these two global measures, Model2 is closer to the benchmark, but is this close enough to conclude that the maps are consistent at a certain level of statistical significance?

$$K = \frac{\sum_{i=1}^n c_{ji} - \frac{\sum_{i=1}^n c_{ji} c_{ii}}{c}}{c_{ji} - \frac{\sum_{i=1}^n c_{ji} c_{ii}}{c}} \quad [1]$$

where c_{ij} – values in the i -row and j -column, dots mean summing up e.g. c_i – summing up over all columns in row i , c – total sum, n – number of classes.

The third method used to determine the differences was the Cartesian distance from the textural properties according to Haralick [45], determined on the basis of the Co-occurrence Matrix. Table 2 contains the values of individual properties for the pattern and models.

Table 3. *The value of differences (taking into account the classes) between the models and the standard based on the Cartesian distance from the textural properties according to Haralick after averaging for angles*

Model	Class 1	Class 2	Class 3	Class 4
Model1	0.173	0.057	0.189	0.071
Model2	0.115	0.003	0.040	0.167
The best fitting	model 2	model 2	model 2	model 1

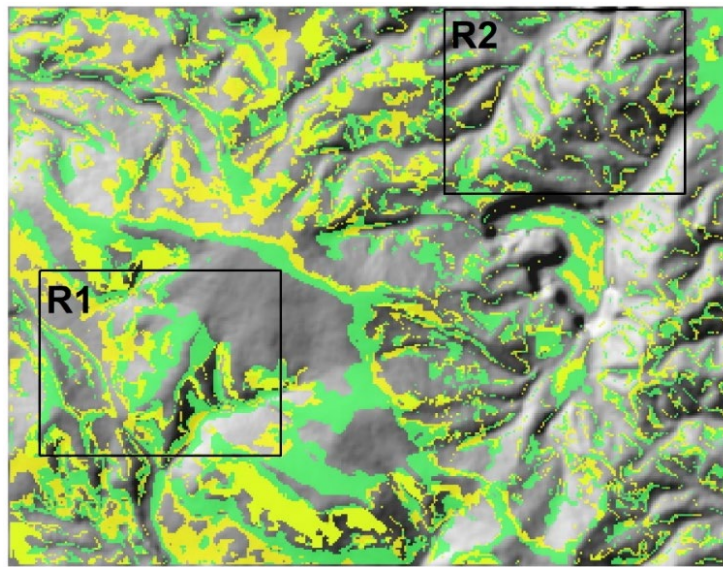


Fig.3. The concept of locality in similarity measures can be understood as a similarity within subclusters or a measure of global similarity calculated for a smaller area (subregion - R1, R2).

Fig.3 shows that greater differences occur in the western part of the study area (R1 rectangle). Difference pixels are larger clusters and there are more of them. In the eastern part of the area, the differences are minor and do not constitute compact patches (R2 rectangle). This experiment shows that differences (Table 3) can be distributed very differently in space, so one global measure of this nature is not sufficient. In addition to the spatial distribution of differences, there is also the issue of their share in a given class, which was already indicated in their research by Pontius et al. [46] and Boots and Csillag [42]. Hence, a local approach seems to be more promising.

Local Measures

In order to determine local measures, 2 sets of sub-areas were designated. One resulted from the simple division of the entire area into equal adjacent sub-areas (Fig.4). The second method of division was to designate two sub-areas that differed according to the field expert (Fig.5).

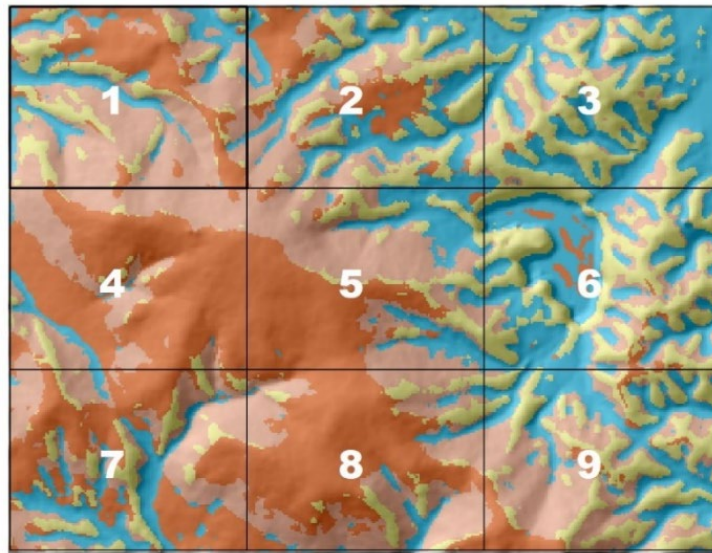


Fig.4. Scheme of division of the area into equal, adjacent sub-areas

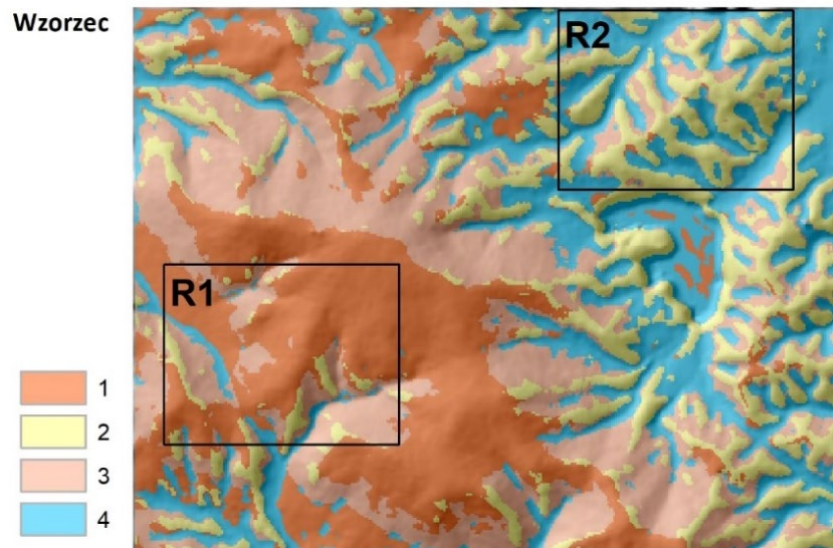


Fig.5. Designation of two sub-areas with large geomorphological differences

Therefore, the determination of local measures consisted in determining global measures for 9 modules (tiles being sub-areas of the entire study area) and for two selected frames, for which the largest and smallest differences between the Pattern and Models 1 and 2 were recorded (Table 4). It is worth noting that the Haralick distance does not indicate Model 2 as closer to the standard for all sub-areas, which is the case for Cohen's kappa coefficient. For Cohen's kappa coefficient, a value of 0.7 was adopted as the limit for a satisfactory fit. Although model 2 has a better match to the pattern everywhere, the critical value of 0.7 was exceeded only for 4 out of 9 sub-regions

Table 4. Summary of differences based on Haralick distance and Cohen's Kappa coefficient for Region 1 (R1) and Region 2 (R2)

Module	Distances according to Haralick		Cohen's Kappa Coefficient	
	Model 1	Model 2	Model 1	Model 2
1	0.54	0.70	0.59	0.69
2	1.48	0.90	0.63	0.65
3	1.66	0.19	0.64	0.84
4	0.95	1.17	0.41	0.64
5	0.46	0.29	0.65	0.82
6	1.47	0.30	0.65	0.82
7	0.32	0.87	0.50	0.59
8	1.40	0.42	0.34	0.63
9	0.64	0.13	0.66	0.83
R1	1.01	0.84	0.42	0.65
R2	0.92	0.05	0.75	0.83

Conclusions

The number of pixels in which two raster maps differ is very specific information, but not sufficient to consider two maps similar or not, because there is no information about the spatial distribution of these differences. The presented measures and indicators are an attempt to develop a tool to assess the significance of differences between qualitative raster maps in the quantitative and spatial context. Even with expert knowledge, it is difficult to indicate the boundary between similar and dissimilar depiction of a given phenomenon. None of the presented measures is sufficient to unequivocally indicate the critical value allowing two maps to be considered similar. The experiment indicates that you should look for a combination of different indicators. The second approach is to determine similarity measures for subregions and then combine them. However, at this stage it is difficult to indicate the optimal number of subregions. The question remains how to combine the obtained measures for subregions into one summary indicator, which requires further research.

References

[1] Z. Rudnicki. Investigation of some discrimination features of texture images, AGH, Automatyka, 2009, T. 13, z. 3/1, 959-969.

[2] B. Boots, F. Csillag. Categorical maps, comparisons, and confidence, J. Geogr. Syst 8 (2006) 109-118. <https://doi.org/10.1007/s10109-006-0018-9>

[3] S. Borkowski et al. The use of 3x3 matrix to evaluation of ribbed wire manufacturing technology, METAL 2012 - 21st Int. Conf. Metallurgy and Materials (2012), Ostrava, Tanger 1722-1728.

[4] R. Ulewicz. Outsourcing quality control in the automotive industry, MATEC Web of Conf. 183 (2018) art. 03001. <https://doi.org/10.1051/matecconf/201818303001>

[5] K. Czerwinska et al. Improving quality control of siluminial castings used in the automotive industry, METAL 2020 29th Int. Conf. Metall. Mater. (2020) 1382-1387. <https://doi.org/10.37904/metal.2020.3661>

[6] P. Fobel, A. Kuzior. The future (Industry 4.0) is closer than we think. Will it also be ethical? AIP Conference Proceedings 2186 (2019) art.80003. <https://doi.org/10.1063/1.5137987>

- [7] A. Kuzior, J. Zozul'ak. Adaptation of the Idea of Phronesis in Contemporary Approach to Innovation, *Manag. Sys. Prod. Eng.* 27 (2019) 84-87. <https://doi.org/10.1515/mspe-2019-0014>
- [8] A. Kuzior, A. Kwilinski. Cognitive Technologies and Artificial Intelligence in Social Perception, *Manag. Sys. Prod. Eng.* 30 (2022) 109-115. <https://doi.org/10.2478/mspe-2022-0014>
- [9] L. Cedro. Model parameter on-line identification with nonlinear parametrization – manipulator model, *Technical Transactions* 119 (2022) art. e2022007. <https://doi.org/10.37705/TechTrans/e2022007>
- [10] T. Lipinski, J. Pietraszek. Influence of animal slurry on carbon C35 steel with different microstructure at room temperature, *Engineering for Rural Development* 21 (2022) 344-350. <https://doi.org/10.22616/ERDev.2022.21.TF115>
- [11] A. Szczotok et al. The Impact of the Thickness of the Ceramic Shell Mould on the ($\gamma + \gamma'$) Eutectic in the IN713C Superalloy Airfoil Blade Casting, *Arch. Metall. Mater.* 62 (2017) 587-593. <https://doi.org/10.1515/amm-2017-0087>
- [12] D. Klimecka-Tatar, M. Ingaldi. Assessment of the technological position of a selected enterprise in the metallurgical industry, *Mater. Res. Proc.* 17 (2020) 72-78. <https://doi.org/10.21741/9781644901038-11>
- [13] P. Jonšta et al. The effect of rare earth metals alloying on the internal quality of industrially produced heavy steel forgings, *Materials* 14 (2021) art. 5160. <https://doi.org/10.3390/ma14185160>
- [14] A. Dudek et al. The effect of alloying method on the structure and properties of sintered stainless steel, *Arch. Metall. Mater.* 62 (2017) 281-287. <https://doi.org/10.1515/amm-2017-0042>
- [15] M. Zenkiewicz et al. Electrostatic separation of binary mixtures of some biodegradable polymers and poly(vinyl chloride) or poly(ethylene terephthalate), *Polimery* 61 (2016) 835-843. <https://doi.org/10.14314/polimery.2016.835>
- [16] N. Radek et al. The influence of plasma cutting parameters on the geometric structure of cut surfaces, *Mater. Res. Proc.* 17 (2020) 132-137. <https://doi.org/10.21741/9781644901038-20>
- [17] N. Radek et al. Technology and application of anti-graffiti coating systems for rolling stock, *METAL 2019 28th Int. Conf. Metall. Mater.* (2019) 1127-1132. ISBN 978-8087294925
- [18] N. Radek et al. The effect of laser beam processing on the properties of WC-Co coatings deposited on steel. *Materials* 14 (2021) art. 538. <https://doi.org/10.3390/ma14030538>
- [19] N. Radek et al. Formation of coatings with technologies using concentrated energy stream, *Prod. Eng. Arch.* 28 (2022) 117-122. <https://doi.org/10.30657/pea.2022.28.13>
- [20] N. Radek et al. The impact of laser welding parameters on the mechanical properties of the weld, *AIP Conf. Proc.* 2017 (2018) art.20025. <https://doi.org/10.1063/1.5056288>
- [21] N. Radek et al. Microstructure and tribological properties of DLC coatings, *Mater. Res. Proc.* 17 (2020) 171-176. <https://doi.org/10.21741/9781644901038-26>
- [22] N. Radek et al. Influence of laser texturing on tribological properties of DLC coatings, *Prod. Eng. Arch.* 27 (2021) 119-123. <https://doi.org/10.30657/pea.2021.27.15>

- [23] N. Radek et al. Operational properties of DLC coatings and their potential application, METAL 2022 – 31st Int. Conf. Metall. Mater. (2022) 531-536. <https://doi.org/10.37904/metal.2022.4491>
- [24] N. Radek. Determining the operational properties of steel beaters after electrospark deposition, Eksploatacja i Niezawodnosc 44 (2009) 10-16.
- [25] N. Radek, J. Konstanty. Cermet ESD coatings modified by laser treatment, Arch. Metall. Mater. 57 (2012) 665-670. <https://doi.org/10.2478/v10172-012-0071-y>
- [26] N. Radek, K. Bartkowiak. Laser Treatment of Electro-Spark Coatings Deposited in the Carbon Steel Substrate with using Nanostructured WC-Cu Electrodes, Physics Procedia 39 (2012) 295-301. <https://doi.org/10.1016/j.phpro.2012.10.041>
- [27] A. Bakowski et al. Frequency analysis of urban traffic noise, ICCO 2019 20th Int. Carpathian Contr. Conf. (2019) 1660-1670. <https://doi.org/10.1109/CarpathianCC.2019.8766012>
- [28] J.M. Djoković et al. Selection of the Optimal Window Type and Orientation for the Two Cities in Serbia and One in Slovakia, Energies 15 (2022) art.323. <https://doi.org/10.3390/en15010323>
- [29] Ł.J. Orman et al. Analysis of Thermal Comfort in Intelligent and Traditional Buildings, Energies 15 (2022) art.6522. <https://doi.org/10.3390/en15186522>
- [30] R. Ulewicz, M. Mazur. Economic aspects of robotization of production processes by example of a car semi-trailers manufacturer, Manuf. Technol. 19 (2019) 1054-1059. <https://doi.org/10.21062/ujep/408.2019/a/1213-2489/MT/19/6/1054>
- [31] S. Blasiak et al. Rapid prototyping of pneumatic directional control valves, Polymers 13 (2021) art.1458. <https://doi.org/10.3390/polym13091458>
- [32] G. Filo et al. Modelling of pressure pulse generator with the use of a flow control valve and a fuzzy logic controller, AIP Conf. Proc. 2029 (2018) art. 20015. <https://doi.org/10.1063/1.5066477>
- [33] E. Lisowski et al. Flow Analysis of a 2URED6C Cartridge Valve, Lecture Notes in Mechanical Engineering 24 (2021) 40-49. https://doi.org/10.1007/978-3-030-59509-8_4
- [34] M. Domagala et al. CFD Estimation of a Resistance Coefficient for an Egg-Shaped Geometric Dome, Appl. Sci. 12 (2022) art.10780. <https://doi.org/10.3390/app122110780>
- [35] R. Dwornicka, J. Pietraszek. The outline of the expert system for the design of experiment, Prod. Eng. Arch. 20 (2018) 43-48. <https://doi.org/10.30657/pea.2018.20.09>
- [36] J. Pietraszek et al. Challenges for the DOE methodology related to the introduction of Industry 4.0. Prod. Eng. Arch. 26 (2020) 190-194. <https://doi.org/10.30657/pea.2020.26.33>
- [37] B. Jasiewicz et al. Inter-observer and intra-observer reliability in the radiographic measurements of paediatric forefoot alignment, Foot Ankle Surg. 27 (2021) 371-376. <https://doi.org/10.1016/j.fas.2020.04.015>
- [38] J. Pietraszek. The modified sequential-binary approach for fuzzy operations on correlated assessments, LNAI 7894 (2013) 353-364. https://doi.org/10.1007/978-3-642-38658-9_32

- [39] J. Pietraszek et al. Non-parametric assessment of the uncertainty in the analysis of the airfoil blade traces, METAL 2017 – 26th Int. Conf. Metall. Mater. (2017) 1412-1418. ISBN 978-8087294796
- [40] J. Pietraszek et al. The non-parametric approach to the quantification of the uncertainty in the design of experiments modelling, UNCECOMP 2017 Proc. 2nd Int. Conf. Uncert. Quant. Comput. Sci. Eng. (2017) 598-604. <https://doi.org/10.7712/120217.5395.17225>
- [41] J. Pietraszek et al. The principal component analysis of tribological tests of surface layers modified with IF-WS2 nanoparticles, Solid State Phenom. 235 (2015) 9-15. <https://doi.org/10.4028/www.scientific.net/SSP.235.9>
- [42] J. Pietraszek, E. Skrzypczak-Pietraszek. The uncertainty and robustness of the principal component analysis as a tool for the dimensionality reduction. Solid State Phenom. 235 (2015) 1-8. <https://doi.org/10.4028/www.scientific.net/SSP.235.1>
- [43] M. Wieczorek, P. Migoń, 2014. Automatic relief classification versus expert and field-based landform classification for the medium-altitude mountain range, the Sudetes, SW Poland, Geomorphology 206 (2014) 133–146. <https://doi.org/10.1016/j.geomorph.2013.10.005>
- [44] P.A. Longley, M.F. Goodchild, D.J. Maguire, D.W. Rhind. Geographical Information Systems: Principles, Techniques, Management and Applications, 2nd Ed., Abridged, Wiley, Hoboken, 2005. ISBN 978-0-471-73545-8
- [45] X. Huang, X. Liu, L. Zhang. A Multichannel Gray Level Co-Occurrence Matrix for Multi/Hyperspectral Image Texture Representation, Remote Sens. 6 (2014) 8424-8445. <https://doi.org/10.3390/rs6098424>
- [46] R.G. Pontius Jr., N.R. Malizia. Effect of category aggregation on map comparison, LNCS 3234 (2004) 251-268. https://doi.org/10.1007/978-3-540-30231-5_17