



Evaluating land cover types from Landsat TM using SAGA GIS for vegetation mapping based on ISODATA and K-means clustering

Polina Lemenkova^{1*}

¹ Université Libre de Bruxelles, École polytechnique de Bruxelles (Brussels Faculty of Engineering), Laboratory of Image Synthesis and Analysis (LISA). Building L, Campus de Solbosch, Avenue Franklin Roosevelt 50, Brussels 1000, Belgium. ORCID ID: <https://orcid.org/0000-0002-5759-1089>

*Corresponding author: polina.lemenkova@ulb.be or pauline.lemenkova@gmail.com

Received 31 March 2021; Accepted 9 November 2021

ABSTRACT

The paper presents the cartographic processing of the Landsat TM image by the two unsupervised classification methods of SAGA GIS: ISODATA and K-means clustering. The approaches were tested and compared for land cover type mapping. Vegetation areas were detected and separated from other land cover types in the study area of southwestern Iceland. The number of clusters was set to ten classes. The processing of the satellite image by SAGA GIS was achieved using Imagery Classification tools in the Geoprocessing menu of SAGA GIS. Unsupervised classification performed effectively in the unlabeled pixels for the land cover types using machine learning in GIS. Following an iterative approach of clustering, the pixels were grouped in each step of the algorithm and the clusters were reassigned as centroids. The paper contributes to the technical development of the application of machine learning in cartography by demonstrating the effectiveness of SAGA GIS in remote sensing data processing applied for vegetation and environmental mapping.

Keywords: SAGA GIS, mapping, vegetation, K-means, ISODATA, clustering, cartography, machine learning.

ИЗВОД

У раду је представљена картографска обрада Landsat TM слике помоћу две ненадгледане методе класификације SAGA GIS: ISODATA и K-means груписање. Приступили су тестирани и упоређени за мапирање типова земљишног покривача. Подручја вегетације откривена су и одвојена од осталих врста земљишног покривача у истраживаном подручју југозападнoг Исланда. Број кластера постављен је на десет класа. Обрада сателитске слике помоћу SAGA GIS постигнута је помоћу алата за класификацију слика у менију за геопроцесирање SAGA GIS. Ненадгледана класификација је била ефикасна у необележеним пикселима за типове покривача земљишта користећи машинско учење у GIS-у. Следећи итеративни приступ кластерисања, пиксели су груписани у сваком кораку алгоритма и кластери су поново додељени као центроиди. Рад доприноси техничком развоју примене машинског учења у картографији показујући ефикасност SAGA GIS у обради података даљинским испитивањем примењеним за мапирање вегетације и животне средине.

Кључне речи: SAGA GIS, мапирање, вегетација, K-means, ISODATA, груписање, картографија, машинско учење.

1. Introduction

Vegetation mapping is one of the most important tools for environmental monitoring. Using remote sensing data processed by GIS is the fastest way that helps land cover types to be visualized and assessed. There are various GIS applications for thematic vegetation mapping (Klaučo et al., 2013a, 2013b; Lemenkova, 2011). The specific geologic setting including volcanism in the southwestern part of Iceland (Figure 1) resulted in the development of erosion prone soils and fragile vegetation (Eckert and Engesser, 2013; Eddudottir et al., 2020). Together with climate impact, ice cover change (Blauvelt et al., 2020; Cabedo-Sanz et al., 2016) and overgrazing, this affected Arctic landscapes and land cover (Lehnhart-Barnett and Waldron, 2020).

The goal of this paper is to present the processing of the Landsat TM image covering the study area i.e. Iceland (Figure 2). Landsat TM images are widely used in environmental studies due to the accessibility, repeatability of survey and coverage (Bryant et al. 2002; Lymburner et al. 2013). The aim is to highlight the distribution of various land cover types. Technical approaches include ISODATA i.e. Iterative Self-Organizing Data Analysis (Memarsadeghi et al., 2007) and K-means image classification (Fard et al., 2020; Peña et al., 1999; Zhao et al., 2020; Bottou and Bengio, 1995), which aim at grouping image pixels into classes of similar properties representing land cover types.

Cartographic processing is based on SAGA GIS (Conrad, 2006; Hengl et al., 2009). Clustering methods are widely used in geosciences for grouping data using similarity properties (Davies and Bouldin, 1979; Filippone et al., 2008; Forgy, 1965; Lemenkova,

2020a). The methods are based on the unsupervised approaches of pixel-based classification, which implies the machine learning approach in remote sensing data processing. These classification techniques can be used to monitor environmental changes such as land degradation or deforestation.

2. Materials and methods

The pixel-based classification approach involves the determination of a spectral response (that is, a digital number, DN) for each pixel of a satellite image. Using the selected mathematical algorithm (ISODATA or K-means), it automatically groups pixels into a class based on the similarities of their DN values. Both methods of unsupervised classification are referred to as cluster analysis in SAGA GIS. The theoretical background of cluster analysis is based on the

principle of data grouping and sorting by mathematical algorithms (Rubin, 1967). The methodology of this work consists in the following workflow.

2.1. Image Destriping

A Landsat TM image was loaded to SAGA GIS and noise corrected using the 'Destriping' procedure by the path 'Geoprocessing > Grid > Filter > Destriping'. The destriping filter removed the straight parallel lines in the raw Landsat TM raster scene by using two low-pass filters elongated in the stripes direction: the one with a pixel unit wide and the one with a striping wavelength wide (Oimoen, 2000). The difference indicated a striping error, which was removed from the original Landsat TM image.

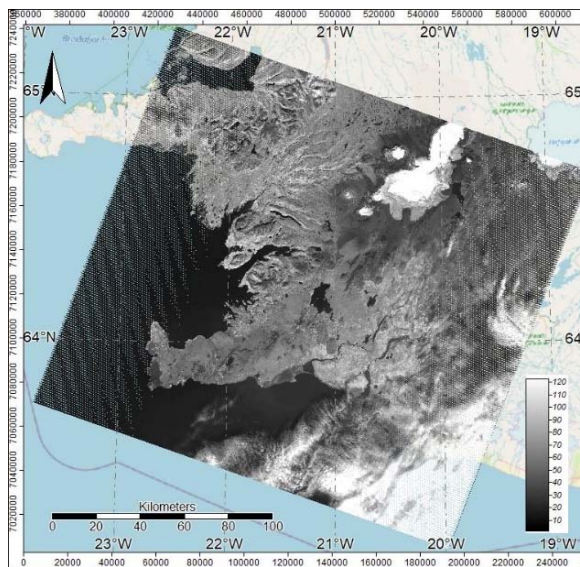


Figure 1. Monochrome image, SAGA GIS.

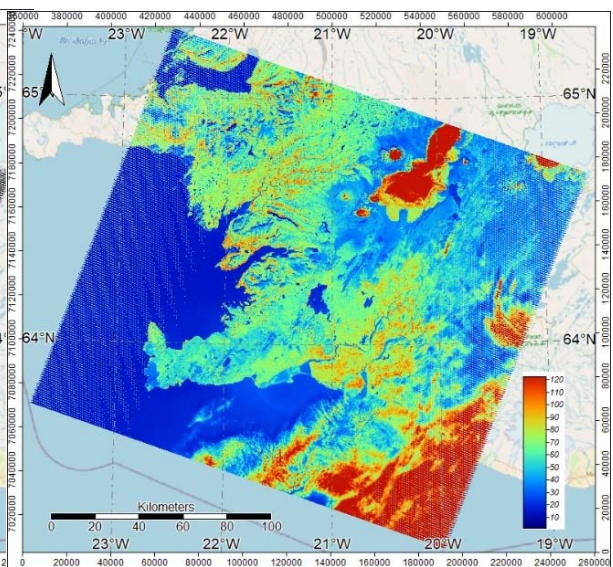


Figure 2. RGB color composite image, SAGA GIS.

2.2. Color Band Composites

The next step included loading Landsat TM bands and generating synthetic images based on the

available Landsat bands. The menu of the SAGA GIS used for testing various combinations of the image is presented in Figure 3.

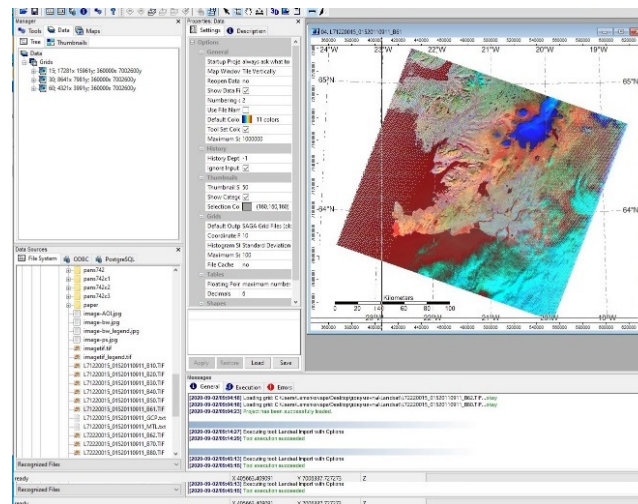


Figure 3. Menu of SAGA GIS and visualized image (Here: Bands R6-G5-B4)

The workflow included the creation of a false color composite (Figure 4) and a natural color composite (Figure 5). For the Landsat TM multiband imagery, three bands in R, G, and B were displayed in color combination from various monochrome bands. A true (natural) color image is composed of the RGB combinations. Using near-infrared bands (NIR), more information (e.g., land-water contrast, vegetation)

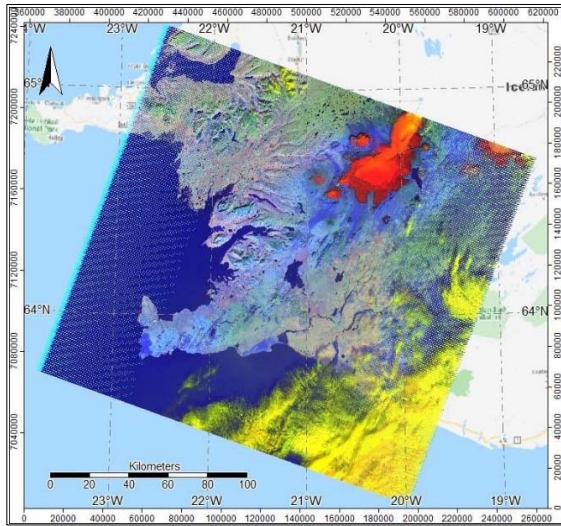


Figure 4. False color composite

2.3. ISODATA Clustering

Unsupervised classification was performed using the SAGA GIS path: Geoprocessing> Imagery> Classification> Unsupervised> ISODATA Clustering for grids (Figure 6). The selected bands were 3,4,5,7 as input bands. Afterwards, the K-means cluster grid was reclassified to land cover classes using the SAGA GIS path: Grid> Values> Reclassify Grid Values. Finally, the statistics on the land cover classes was visualized.

ISODATA clustering, an unsupervised pixel classification method by SAGA GIS, was used for detecting and mapping the land cover classes of southwestern Iceland. Image bands were selected to be used in the assignment of bands for ISODATA clustering and the number of eventual output classes was defined as 10.

The SAGA GIS solution of ISODATA clustering was used to solve the problem of a large amount of unlabeled pixels for land cover types. Since the training pixels of supervised classification require

was added. The blue channel was used for a false color composite. The combination presented in Figure 5 shows bright cyan-colored ice and glacier areas, dark (black) colors for water and natural looking landscapes (green for vegetation areas and brown for bare soils). The false color composite in Figure 4 shows ice covered areas as bright red, useful for glacier mapping.

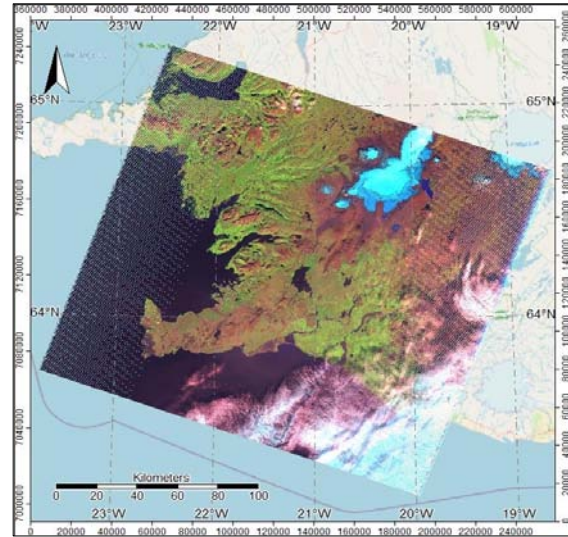


Figure 5. Natural color composite. Bands: R7-G4-B2

fieldwork data observations, unsupervised classification performs better for a distance based data analysis.

Understanding the meaning of land cover types behind the pool of pixels on a Landsat TM scene requires a machine learning algorithm that classifies these pixels into groups based on the patterns it finds. The unsupervised learning of ISODATA conducts an iterative process, analyzing pixels without the intervention of a cartographer. The key approach of ISODATA relies on the assumption that each class has a multivariate normal distribution. Therefore, it uses class means and a covariance matrix for each class. In case of complicated landscapes, there can be many variables in vegetation and mixed land cover types, which can make hand-made supervised classification a difficult process. Instead, the machine-learning classifiers ISODATA and K-means, based on the clustering and association of pixels on a Landsat TM scene, are applied to identify land cover classes automatically.

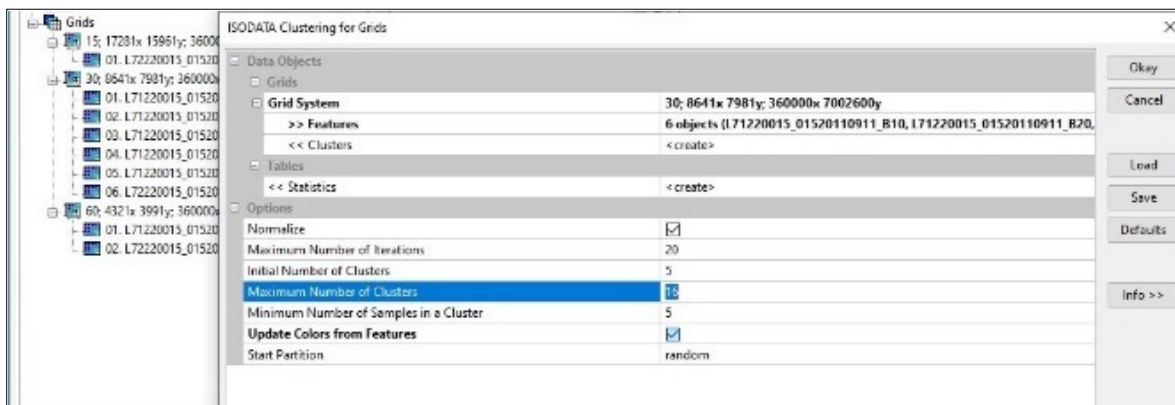


Figure 6. Menu of ISODATA clustering, SAGA GIS

2.4. K-means Clustering

The K-means method is another widely used approach of unsupervised classification. These classes were assigned as 10 machine-defined classes and then reclassified as post-processing as 'land cover classes' by reference to ground data (Figure 7). In contrast to supervised classification, both ISODATA and K-means do not require defining 'training sites' of the land cover type. This presents a higher level of machine learning in a cartographic workflow. The automation

and methods of machine learning are commonly used in geosciences.

The iterative process of image processing by K-means clustering in SAGA GIS leads to an improvement in the associations of pixels into cluster groups by machine learning. The SAGA GIS divides the pixels from the Landsat TM image and groups them into the assigned number of classes. The K-means based classifications were fast and efficient. However, some drawbacks include errors from misclassified pixels caused by cloudiness.

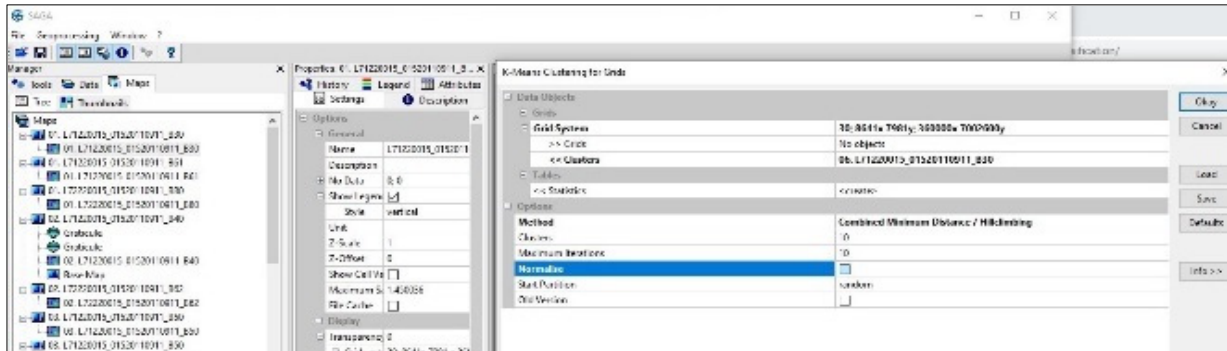


Figure 7. Menu of K-means clustering, SAGA GIS

3. Results and discussions

The resulting maps show the unsupervised techniques of ISODATA (Figure 8) and K-means (Figure 9). The results determine 10 classes within

the Landsat TM for optimal classification performance. The maps keep spatial resolution and texture in the image. K-means clustering was reclassified (Figure 10) for the land cover classes with the assigned types.

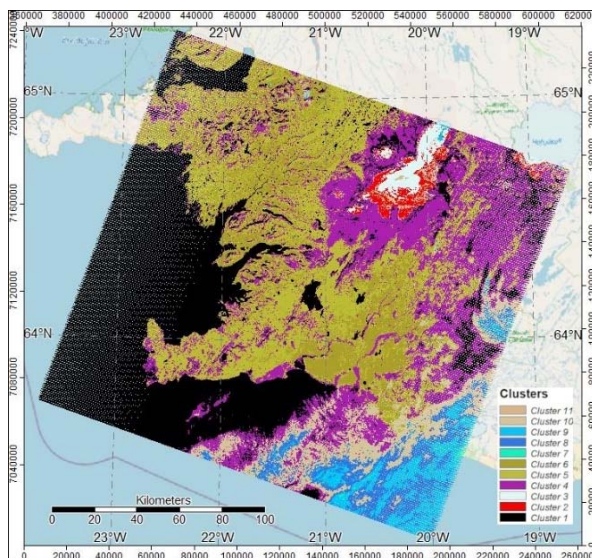


Figure 8. ISODATA cluster analysis for grids

The difference in performance consists in the mathematical approach of the k-means algorithm, which aims at the placement of the centers that minimizes the average squared distance of each point to its nearest center (Likas et al. 2003). ISODATA tries to treat each class in a multivariate normal distribution, and computes class means and a covariance matrix for each class (Figure 11).

The algorithm of the K-means is implemented in the most straightforward manner, assuming that the number of clusters, k , is much smaller than the total

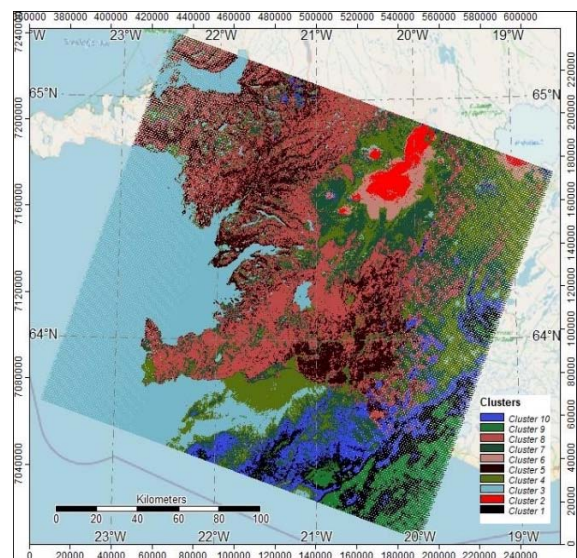


Figure 9. K-means cluster analysis for grids

number of pixels in a scene (Pollard, 1982) (Figure 12). Hence, the performance of the algorithm is more time-consuming compared to ISODATA. As with ISODATA, the time of algorithm performance is explained by the time required to compute the cluster center nearest to each pixel in an image (Jain and Dubes, 1988).

The principles of ISODATA classification based on the iterative approach have the following nature: pixels are being grouped in each iteration of the algorithm: pixels in an image are assigned to their

closest (nearest) cluster centers. Afterwards, cluster centers are reassigned to be the centroid of these associated points.

The next resulting step includes clusters with very few deleted points. Finally, larger clusters

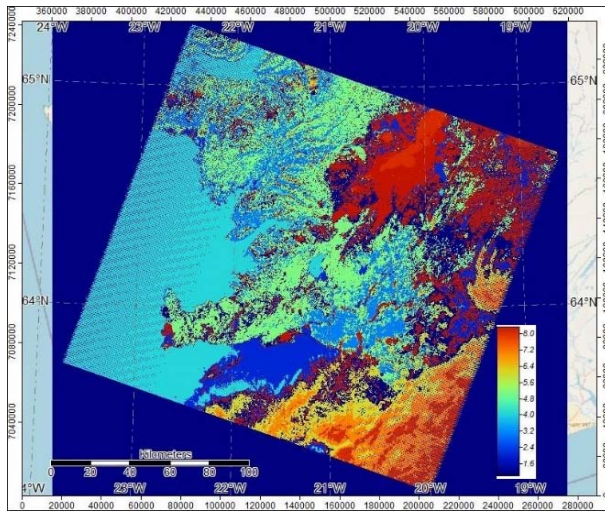


Figure 10. Reclassified map of K-means clustering

4. Conclusions

The demonstrated advantage of machine learning in cartographic data processing over supervised classification in image analysis consists in the automation of data processing. Due to the geometric complexity of contours and fragmentation of many landscapes, the patterns and associations of land cover types may be easily overlooked by a human eye.

Better accuracy and reliability result from automation, which is part of machine learning. The machine learning applied in cartographic workflow enables GIS to train on datasets before data processing (Lemenkova, 2019a, 2019b; Schenke and Lemenkova, 2008). Machine-learning approaches in geosciences include advanced statistical analysis, numerical data processing and applied programming (Chavez et al. 1982; Lemenkov and Lemenkova, 2021a, 2021b; Clark and Roush, 1984; Klaučo et al., 2014, 2017; Lemenkova, 2019c, 2019d, 2021a; Jensen, 2005).

The comparison between the ISODATA and K-means approaches showed that ISODATA operates more slowly, particularly with several processed bands, while the K-means algorithm is a faster method. Both algorithms are central to studies on Landsat TM image processing, classification and environmental applications (Esche and Franklin, 2002; Lemenkova, 2020b; Chen et al., 2020; Xu and Wunsch, 2005). Both ISODATA and K-means are popular and widely used unsupervised classification methods (Kanungo et al., 2002; Forgey, 1965; Arya et al., 2004; Murariu et al., 2018) both in general data analysis and in remote sensing applications and can be recommended in further studies.

The presented work revealed that the use of Landsat TM satellite imagery and various approaches of remote sensing data processing also continues previous applications of Landsat TM (Liu et al. 2010; Zhao et al. 2016; Mondal et al., 2020; Zerrouki et al. 2021). Moreover, clustering is an effective method of

satisfying new conditions are split again, and smaller clusters reassigned according to the proximity of pixels are merged (Tou and Gonzalez, 1974). The algorithms continue until the number of iterations reaches the number of defined classes.

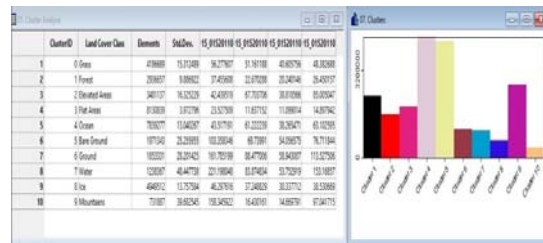


Figure 11. K-means cluster analysis and a histogram



Figure 12. ISODATA cluster analysis and a histogram

image processing in environmental mapping and vegetation monitoring using SAGA GIS, producing both cartographic maps and statistical graphs and tables. However, various approaches in classification techniques (both supervised and unsupervised) applied for satellite images may lead to ambiguities and difficulties to accurately interpret the land cover types. Noise, such as atmospheric effects, cloudiness or technical stripes, is still a challenge in the Landsat scenes with existing applications on solving these problems (Mitchell et al., 1977; Iikura, 2002; Deng et al. 2016).

For the selected study area, the final map includes the following land cover types: grass, forest, elevated areas, flat areas, ocean, bare ground, ground, water, ice, mountains. A noted discrepancy between the two approaches can be a recommended direction in further studies on Landsat TM image processing in the optimization of cartographic techniques using remote sensing data by SAGA GIS.

An accurate method for mapping vegetation and detecting land cover types aims to harmonize the existing methods of geospatial data processing using the advanced GIS software combined with various mathematical algorithms of machine learning. Compared to the traditional GIS mapping (Săvulescu and Mihai, 2011; Suetova et al., 2005; Anny et al. 2014; Vilček and Koco, 2018), the machine learning approach enables more rapid, accurate and precise mapping (Lemenkova, 2020c, 2020d, 2021b, 2021c). An outcome of the machine methods in environmental cartography aims at optimizing the techniques of remote sensing data processing for forest and vegetation monitoring (Zaimes et al., 2019). This study contributed to the existing research by presenting methods of machine learning cartographic techniques in environmental agricultural studies.

Acknowledgment

This study has been supported by the project No. 0144-2019-0011, Schmidt Institute of Physics of the Earth, Russian Academy of Sciences.

Declaration of competing interest

The author declares no conflict of interest.

References

- Anny, K., Frankl, A., Spalević, V., Čurović, M., Borota, D., Nyssen, J. (2014) Geomorphology of the Durmitor Mountains and surrounding plateau Jezerska Površ (Montenegro). *Journal of Maps*, 10(4), 600-611. doi: 10.1080/17445647.2014.909338
- Arya, V., Garg, N., Khandekar, R., Pandit, V., Meyerson, A., Munagala, K. (2004). Local search heuristics for k-median and facility location problems. *SIAM Journal of Computing*, 33(3), 544-562. doi: 10.1137/S0097539702416402
- Blauvelt, D. J., Russell, A. J., Large, A. R.G., Tweed, F. S., Hiemstra, J. F., Kulesa, B., Evans, D.J.A., Waller, R.I. (2020). Controls on jökulhlaup-transported buried ice melt-out at Skeiðarársandur, Iceland: Implications for the evolution of ice-marginal environments. *Geomorphology*, 360, 107164. doi: 10.1016/j.geomorph.2020.107164
- Bottou, L., Bengio, Y. (1995). Convergence properties of the k-means algorithms. *Advances in Neural Information Processing Systems* 7, 585-592.
- Bryant, R., Moran, M. S., McElroy, S., Holifield, C., Thome, K., Miura, T. (2002). Data continuity of Landsat-4 TM, Landsat-5 TM, Landsat-7 ETM+, and Advanced Land Imager (ALI) sensors," *IEEE International Geoscience and Remote Sensing Symposium*, 584-586 vol.1, doi: 10.1109/IGARSS.2002.1025112.
- Cabedo-Sanz, P., Belt, S.T., Jennings, A.E., Andrews, J.T., Geirsdottir, A. (2016). Variability in drift ice export from the Arctic Ocean to the North Icelandic Shelf over the last 8000 years: a multi-proxy evaluation. *Quaternary Science Reviews*, 146, 99-115. doi: 10.1016/j.quascirev.2016.06.012
- Chavez, P.L., Berlin, G.L., Sowers, L.B. (1982). Statistical Method for Selecting Landsat MSS Ratios. *Journal of Applied Photographic Engineering*, 8(1), 23-30.
- Chen, J., Qi, X., Chen, L., Chen, F., Cheng, G. (2020). Quantum-inspired ant lion optimized hybrid k-means for cluster analysis and intrusion detection. *Knowledge-Based Systems*, 203, 106167. doi: 10.1016/j.knsys.2020.106167
- Clark, R.N., Roush, T.L. (1984). Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research*, 89, 6329-6340. doi: 10.1029/JB089iB07p06329
- Conrad, O. (2006). SAGA – program structure and current state of implementation. SAGA – Analysis and Modelling Applications. *Göttinger Geographische Abhandlungen*, 115, 39-52.
- Davies, D.L., Bouldin, D.W. (1979). A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell. PAMI-1(2)* 224-227. doi: 10.1109/TPAMI.1979.4766909
- Deng, J., Wang, H., Ma, J. (2016). An automatic cloud detection algorithm for Landsat remote sensing image. *4th International Workshop on Earth Observation and Remote Sensing Applications (EORSA)*, 395-399. doi: 10.1109/EORSA.2016.7552837
- Eckert, S., Engesser, M. (2013). Assessing vegetation cover and biomass in restored erosion areas in Iceland using SPOT satellite data. *Applied Geography*, 40, 179-190. doi: 10.1016/j.apgeog.2013.02.015
- Eddudottir, S. D., Erlendsson, E., Gísladottir, G. (2020). Landscape change in the Icelandic highland: A long-term record of the impacts of land use, climate and volcanism. *Quaternary Science Reviews*, 240, 106363. doi: 10.1016/j.quascirev.2020.106363
- Esche, H. A., Franklin, S. E. (2002). Assessing cloud contamination effects on K-means unsupervised classifications of Landsat data. *IEEE International Geoscience and Remote Sensing Symposium*, 6, 3387-3389, doi: 10.1109/IGARSS.2002.1027191
- Fard, M. M., Thonet, T., Gaussier, E. (2020). Deep k-Means: Jointly clustering with k-Means and learning representations. *Pattern Recognition Letters*, 138, 185-192. doi: 10.1016/j.patrec.2020.07.028
- Filippone, M., Camastra, F., Masulli, F., Rovetta, S. (2008). A survey of kernel and spectral methods for clustering. *Pattern Recognit.*, 41(1), 176-190. doi: 10.1016/j.patcog.2007.05.018
- Forgy, E. (1965). Cluster analysis of multivariate data: efficiency vs. interpretability of classifications. *Biometrics*, 21, 768.
- Hengl, T., Grohmann, C.H., Bivand, R.S., Conrad, O., Lobo, A. (2009). SAGA vs GRASS: A Comparative Analysis of the Two Open Source Desktop GIS for the Automated Analysis of Elevation Data. In: Purves, R., Gruber, S., Straumann, R., Hengl, T. (Eds.), *Geomorphometry Conference Proceedings*, 22-27.
- Ikura, Y. (2002). Analysis of scan stripe noise in Landsat TM images. *Proceedings of the 41st SICE Annual Conference. SICE 2002*, 103-105 vol.1, doi: 10.1109/SICE.2002.1195192.
- Jainand, A.K., Dubes, R.C. (1988). *Algorithms for Clustering Data*. Prentice Hall, Englewood Cliffs, NJ.
- Jensen JR. (2005). *Thematic map accuracy assessment. In Introductory Digital Image Processing – A Remote Sensing Perspective*. (3rd ed.) Keith, C.C., Prentice Hall Series in Geographic Information Science: Saddle River, NJ, USA, pp. 495-515.
- Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R., Wu, A.Y. (2002). An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24, 881-892. doi: 10.1109/TPAMI.2002.1017616
- Klaučo M., Gregorová B., Stankov U., Marković V., Lemenkova P. (2013a). Determination of ecological significance based on geostatistical assessment: a case study from the Slovak Natura 2000 protected area. *Open Geosciences*, 5(1), 28-42. doi: 10.2478/s13533-012-0120-0
- Klaučo, M., Gregorová, B., Stankov, U., Marković, V., Lemenkova, P. (2013b). Interpretation of Landscape Values, Typology and Quality Using Methods of Spatial Metrics for Ecological Planning. 54th International Scientific Conference *Environmental and Climate Technologies*, October 14, 2013. Riga, Latvia. doi: 10.13140/RG.2.2.23026.96963
- Klaučo, M., Gregorová, B., Stankov, U., Marković, V., Lemenkova, P. (2014). Landscape metrics as indicator for ecological significance: assessment of Sitno Natura 2000 sites, Slovakia. *Ecology and Environmental Protection*, 85-90. doi: 10.6084/m9.figshare.7434200
- Klaučo, M., Gregorová, B., Koleda, P., Stankov, U., Marković, V., Lemenkova, P. (2017). Land planning as a support for sustainable development based on tourism: A case study of Slovak Rural Region. *Environmental Engineering and Management Journal*, 2(16), 449-458. doi: 10.30638/eemj.2017.045
- Lehnhart-Barnett, H., Waldron, S. (2020). The influence of land cover, including Nootka lupin, on organic carbon exports in east Icelandic rivers. *Catena*, 184: 104245. doi: 10.1016/j.catena.2019.104245
- Lemenkov, V., Lemenkova, P. (2021a). Using TeX Markup Language for 3D and 2D Geological Plotting. *Foundations of Computing and Decision Sciences*, 46(3), 43-69. doi: 10.2478/fcds-2021-0004
- Lemenkov, V., Lemenkova, P. (2021b). Measuring Equivalent Cohesion C_{eq} of the Frozen Soils by Compression Strength Using Kriolab Equipment. *Civil and Environmental Engineering Reports*, 31(2), 63-84. doi: 10.2478/ceer-2021-0020

- Lemenkova P. (2011). *Seagrass Mapping and Monitoring Along the Coasts of Crete, Greece*. M.Sc. Thesis. University of Twente, Enschede, Netherlands. doi: 10.13140/RG.2.2.16945.22881
- Lemenkova P. (2019a). Processing oceanographic data by Python libraries NumPy, SciPy and Pandas. *Aquatic Research*, 2(2), 73–91. doi: 10.3153/AR19009
- Lemenkova, P. (2019b). Automatic Data Processing for Visualising Yap and Palau Trenches by Generic Mapping Tools. *Cartographic Letters*, 27(2), 72–89. doi: 10.6084/m9.figshare.11544048
- Lemenkova P. (2019c). AWK and GNU Octave Programming Languages Integrated with Generic Mapping Tools for Geomorphological Analysis. *GeoScience Engineering*, 65(4), 1–22. doi: 10.35180/gse-2019-0020
- Lemenkova P. (2019d). Statistical Analysis of the Mariana Trench Geomorphology Using R Programming Language. *Geodesy and Cartography*, 45(2), 57–84. doi: 10.3846/gac.2019.3785
- Lemenkova, P. (2020a). R Libraries {dendextend} and {magrittr} and Clustering Package scippy. cluster of Python For Modelling Diagrams of Dendrogram Trees. *Carpathian Journal of Electronic and Computer Engineering*, 13(1), 5–12. doi: 10.2478/cjece-2020-0002
- Lemenkova, P. (2020b). SAGA GIS for information extraction on presence and conditions of vegetation of northern coast of Iceland based on the Landsat TM. *Acta Biologica Marisiensis*, 3(2), 10–21. doi: 10.2478/abmj-2020-0007
- Lemenkova, P. (2020c). Sentinel-2 for High Resolution Mapping of Slope-Based Vegetation Indices Using Machine Learning by SAGA GIS. *Transylvanian Review of Systematical and Ecological Research*, 22(3), 17–34. doi: 10.2478/trser-2020-0015
- Lemenkova, P. (2020d). Using R packages 'tmap', 'raster' and 'ggmap' for cartographic visualization: An example of dem-based terrain modelling of Italy, Apennine Peninsula. *Zbornik radova – Geografski fakultet Univerziteta u Beogradu*, 68, 99–116. doi: 10.5937/zrgfub2068099L
- Lemenkova, P. (2021a). Scripting Languages for Geomorphological Modelling and Topographic Visualization of Serbia. *Zbornik Matice srpske za prirodnu nauku*, 140, 113–125. doi: 10.2298/ZMSPN2140113L
- Lemenkova, P. (2021b). Data-driven insights into correlation among geophysical setting, topography and seafloor sediments in the Ross Sea, Antarctic. *Caderno de Geografia*, 31(64), 1–20. doi: 10.5752/P.2318-2962.2021v31n64p1
- Lemenkova, P. (2021c). Geodynamic setting of Scotia Sea and its effects on geomorphology of South Sandwich Trench, Southern Ocean. *Polish Polar Research*, 42(1), 1–23. doi: 10.24425/ppr.2021.136510
- Likas, A., Vlassis, N., Verbeek, J.J. (2003). The global k-means clustering algorithm. *Pattern Recognition*, 36 (2), 451–461. doi: 10.1016/S0031-3203(02)00060-2
- Liu, L., Xu, Q., Yang, S. (2010). Identification of river ice on the Yellow river using LANDSAT images. *18th International Conference on Geoinformatics*, 1-4, doi: 10.1109/GEOINFORMATICS.2010.5567714
- Lymburner, L., McIntyre, A., Li, F., Ip, A., Thankappan, M., Sixsmith, J. (2013). Creating multi-sensor time series using data from Landsat-5 TM and Landsat-7 ETM+ to characterise vegetation dynamics. *2013 IEEE International Geoscience and Remote Sensing Symposium – IGARSS*, 961-963, doi: 10.1109/IGARSS.2013.6721321.
- Memarsadeghi, N., Mount, D. M., Netanyahu, N. S., Le Moigne, J. (2007). A Fast Implementation of the ISODATA Clustering Algorithm. *International Journal of Computational Geometry and Applications*, 17, 71–103. doi: 10.1142/S0218195907002252
- Mitchell, O. R., Delp, E. J., Chen, P. L. (1977). Filtering to remove cloud cover in satellite imagery. *IEEE Transactions on Geoscience Electronics*, 15(3), 137-141. doi: 10.1109/TGE.1977.6498971
- Mondal, R., Mukherjee, J., Mukhopadhyay, J. (2020). Automated Coastline Detection from Landsat 8 Oli/Tirs Images with the Presence of Inland Water Bodies in Andaman. *IGARSS 2020 – 2020 IEEE International Geoscience and Remote Sensing Symposium*, 57-60, doi: 10.1109/IGARSS39084.2020.9324366.
- Murariu, G., Munteanu, D., Georgescu, L.P., Murariu, A., Popa, I., Hahuie, V., Dragu, M.D., Iticescu, C. (2018). The forest composition monitoring system using k-means algorithms on satellite imagery. Case study – Independenta Forest. *22nd International Conference on System Theory, Control and Computing (ICSTCC)*, 1-8, doi: 10.1109/ICSTCC.2018.8540776.
- Oimoen, M.J. (2000). An Effective Filter for Removal of Production Artifacts. In U.S. Geological Survey 7.5-Minute Digital Elevation Models. *Proceedings of the 14th International Conference on Applied Geologic Remote Sensing*, 6-8 November, Las Vegas, NV.
- Peña, J.M., Lozano, J.A., Larrañaga, P. (1999). An empirical comparison of four initialization methods for the k-means algorithm. *Pattern Recognition Letters* 20, 1027–1040. doi: 10.1016/S0167-8655(99)00069-0
- Pollard, D. (1982). A central limit theorem for k-means clustering. *Annals of Probability*, 10, 919–926.
- Rubin, J. (1967). Optimal classification into groups: an approach for solving the taxonomy problem. *J. Theoretical Biology*, 15, 103–144. doi: 10.1016/0022-5193(67)90046-X
- Săvulescu, I., Mihai, B. (2011) Mapping forest landscape change in Iezer Mountains, Romanian Carpathians. AGIS approach based on cartographic heritage, forestry data and remote sensing imagery. *Journal of Maps*, 7,1, 429-446. doi: 10.4113/jom.2011.1170
- Schenke H.W., Lemenkova P. (2008). Zur Frage der Meeresboden-Kartographie: Die Nutzung von AutoTrace Digitizer für die Vektorisierung der Bathymetrischen Daten in der Petschora-See. *Hydrographische Nachrichten*, 81, 16–21. doi: 10.6084/m9.figshare.7435538
- Suetova I.A., Ushakova L.A., Lemenkova P. (2005). Geoinformation mapping of the Barents and Pechora Seas. *Geography and Natural Resources*, 4, 138–142. doi: 10.6084/m9.figshare.7435535
- Tou, J. T., Gonzalez, R. C. (1974). *Pattern Recognition Principles*. Addison-Wesley, London. doi: 10.1002/zamm.19770570626
- Vilček, J., Koco, Š. (2018). Integrated index of agricultural soil quality in Slovakia. *Journal of Maps*, 14(2), 68-76. doi: 10.1080/17445647.2018.1428233
- Xu, R., Wunsch, D.C. (2005). Survey of clustering algorithms. *IEEE Transactions Neural Networks*, 16(3), 645–678. doi: 10.1109/TNN.2005.845141
- Zaimes, G.N., Gounaridis, D., Symeonakis, E. (2019). Assessing the impact of dams on riparian and deltaic vegetation using remotely-sensed vegetation indices and Random Forests modelling. *Ecological Indicators*, 103, 630–641. doi: 10.1016/j.ecolind.2019.04.047
- Zhao, Y., Lin, L., Lu, W., Meng, Y. (2016). Landsat time series clustering under modified Dynamic Time Warping. *4th International Workshop on Earth Observation and Remote Sensing Applications (EORSAA)*, 62-66. doi: 10.1109/EORSAA.2016.7552767.
- Zhao, F., Hung, D.L.S., Wu, S. (2020). K-means clustering-driven detection of time-resolved vortex patterns and cyclic variations inside a direct injection engine. *Applied Thermal Engineering*, 180, 115810. doi: 10.1016/j.applthermaleng.2020.115810
- Zerrouki, Y., Harrou, F., Zerrouki, N., Dairi, A., Sun, Y. (2021). Desertification Detection Using an Improved Variational Autoencoder-Based Approach Through ETM-Landsat Satellite Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 202-213. doi: 10.1109/JSTARS.2020.3042760