RPD	Revista Produção e Desenvolvimento Research in Production and Development	eISSN: 2446-9580 Doi: https://doi.org/10.32358/rpd.2023.v9.660

UNDERSTANDING THE USAGE OF LAND USE AND LAND COVER CLASSIFIERS IN SCIENTIFIC RESEARCH

Vitor da Silva Gonçalves[1], **vitorsgoncalves@protonmail.com**, **https://orcid.org/0000-0003-2783-3865**
Italo de Oliveira Matias[1], **italo@ucam-campos.br**, **https://orcid.org/0000-0001-6683-5287**
Aldo Shimoya[1], aldoshimoya@yahoo.com.br, https://orcid.org/0000-0003-1098-9580

1. Universidade Cândido Mendes, Campos dos Goytacazes/ RJ - Brasil

Submitted: 13/11/2023. Accepted: 22/12/2023
Published: 29/12/2023

ABSTRACT
Purpose: The objective of this study is to assess the primary methods utilized in land use and land cover classification research to determine the most frequently applied techniques and potential trends

Methodology/Approach: A bibliometric study was carried out in the Scopus scientific article databases, counting the use of classification methods between 2012 and 2022. The obtained data was converted into SQL database tables and processed using queries, looking for articles whose abstracts contains keywords related to land cover and land use methods.

Findings: A general growth trend in the studies of this area was discovered, with an increase in the use of machine learning-based methodologies and stabilization in the use of other methods. Statistical methods were also heavily used, while others were used less frequently.

Research Limitation/implication: It's important to note that keywords present in abstract sections does not necessarily correspond to the mentioned methods being applied in the studies. This leads to some expected degree of imprecision, but the data remains a reasonable representation of the popularity of each method.

Originality/Value of paper: Considering that the scientific literature lack a quantitative understanding of the usage of land use and land cover classifiers, this work aims to fill that gap by providing usable data for other studies and/or decision taking.

KEYWORDS: land use, land cover, classification, bibliometrics.

UTILIZAÇÃO DE CLASSIFICADORES DE USO E COBERTURA DO SOLO EM PESQUISAS CIENTÍFICAS

RESUMO
Objetivo: O objetivo deste estudo é avaliar os principais métodos utilizados na investigação sobre a classificação da utilização e da ocupação do solo, a fim de determinar as técnicas mais frequentemente aplicadas e as tendências potenciais

Metodologia/Abordagem: Foi realizado um estudo bibliométrico na base de dados de artigos científicos Scopus, contabilizando o uso de métodos de classificação entre 2012 e 2022. Os dados obtidos foram convertidos em tabelas de banco de dados SQL e processados por meio de consultas, procurando artigos cujos resumos contenham palavras-chave relacionadas com métodos de classificação de uso e cobertura do solo.

Resultados: Foi identificada uma tendência geral de crescimento nos estudos desta área, com um aumento expressivo no uso de metodologias baseadas em aprendizado de máquina e estabilização no uso de outros métodos. Os métodos estatísticos também se demonstraram frequentemente utilizados, enquanto outros foram utilizados com menor frequência.

Limitações da pesquisa/implicações: É importante notar que as palavras-chave presentes nas secções dos resumos não correspondem necessariamente aos métodos aplicados nos estudos. Isto leva a um certo grau de imprecisão, mas os dados obtidos permanecem uma representação razoável da popularidade de cada método.

Originalidade/Valor do artigo: Considerando que a literatura científica carece de uma compreensão quantitativa da utilização de classificadores de uso e ocupação do solo, este trabalho pretende preencher essa lacuna, fornecendo dados utilizáveis para outros estudos e/ou tomadas de decisão.

KEYWORDS: uso do solo, cobertura do solo, classificação, bibliometria.

1. INTRODUCTION

Land use and land cover mapping has become an important tool for geoprocessing studies, with the potential to provide information on changes in natural, urban and other areas (Chaves et al., 2020). This type of mapping facilitates the quantification and temporal analysis of complex phenomena, and, together with other types of studies, can lead to a deeper understanding of such phenomena.

Wulder et al. (2018) described land cover and land use as a set of hierarchical classes, each representing the most common biotic and abiotic groups, which constitute themselves as critical descriptors of the Earth's surface. This type of mapping can also be understood, in a simpler way, as the labeling of the elements of Earth's surface into categories. In this way, classifying the elements is the main task in order to produce a land cover and land use map, whose characteristics, quality and fidelity of the results depend on the applied method.

Several methods have been developed in order to perform this type of classification, called classifiers. Historically, classifiers relied on manual analysis, based on the experience of specialists and/or the application of numerical methods, but with the advents of computing, new techniques have been created, allowing larger areas to be analyzed, with higher resolution, and faster (Phiri & Morgenrorth, 2017). Given the large amount of land use and land cover classifiers in use, it is important to understand how often they are applied in the scientific literature, which ones are the most used, and to identify possible trends, whether to understand the state of the art, to help researchers develop their methodologies, or to help professionals in the field make decisions. In this context, the present work aims to identify the classifiers used in scientific research on land cover and land use, cataloged in the scientific database Scopus, in the period between 2012 and 2022, as well as quantify their usage proportions and observe them for possible trends.

2. LITERATURE REVIEW

Several land use and land cover classification methods have been developed over the years, supported by technological advances, in both image processing and image acquisition. Phiri and Morgenroth (2017) pointed out that prior to the 1970s, such studies were carried out by experts, using visual and manual methods, with the first digital analyses beginning in that decade, supported by the use of filtering and numerical classification techniques. According to the same authors, the recognition of numerical patterns was an important development and served as the basis for modern methods, which allowed the use of software approaches, implementing different types of classifiers.

Jozdani et al. (2019) consider that selecting a classifier is a critical aspect of mapping land use and land cover. The authors distinguished two categories of classifiers that are widely used today: those based on machine learning, including methods such as random forest, bagging trees and boosted trees, and those based on deep learning, which include the use of deep neural networks. The text offers an objective overview of the different options available, with a focus on technical terms that are explained clearly.

Although a diverse range of techniques can be found in the scientific literature, they fall mainly into 5 categories, according to their primary operational mode: clustering, statistical approaches, machine learning, image segmentation, and manual methods. Some techniques do not fit into these categories, but they have more limited applications. It is also possible to use multiple approaches in the same work, either through a combination of methods or by using different methods at different stages of the same work.

Clustering is an unsupervised classification technique wherein the algorithm groups pixels from an image into a predefined number of natural clusters, based on their proximity to features and/or statistical similarity (Borra et al., 2019). Despite being an unsupervised methodology, there are clustering techniques that allow some data input, a method known as semi-supervised classification (Mai & Ngo, 2018). It is noteworthy that the term clustering is sometimes utilized in the literature as a synonym for classification, leading to potential confusion.

Statistical methods have been utilized for land use and land cover classification since their early years of development (Frazier & Shovic 1980), and are still widely used. These methods involve generating statistical models from the training data in order to identify the category best corresponds to each pixel in the image.

Machine learning is a subfield of computer science that explores algorithms and techniques for automating complex solutions, that might be difficult to achieve through conventional programming, and whose design may be difficult or impracticable to execute (Rebala et al., 2019). According to Alzubi et al., (2018), machine learning approaches have demonstrated their effectiveness in solving data science problems, such as those relating to classification and clustering. Thus, machine learning presents itself as a viable classification tool, that can be applied into different kinds of research and is regarded by Wang et al. (2022), as a central part of land use and cover change studies utilizing the technique. Its application addresses problems of single image processing and pattern recognition.

The generic machine learning model comprises six components: collection and preparation of data, feature selection, algorithm choice, selection of the model and parameters, training, and performance evaluation (Alzubi et al., 2018). In the domain of land use and land cover classification, this process is often simplified through the utilization of software tools, allowing the researcher to focus on creating a good training model and, when necessary, adjusting parameters.

Image segmentation is a technique in image processing that separates an image into distinct segments or regions, according to the differentiation between pixels. It should be noted that segmentation is not a classification technique per se, but it can be used for this purpose in some contexts (Horning, 2004). Compared to image classification techniques, segmentation categorizes entire regions, instead of individual pixels. However, it is worth noting that the literature does not present a consensus on the definitions, and that a distinction is not always made between the terms, which some studies use interchangeably. Some authors, such as Lei and Nandi (2022), differentiate between the two types of procedures, but approach segmentation as a technique of image classification.

Manual methods are the earliest to be developed, employed even before the use of satellites for image acquisition (Phiri & Morgenroth, 2017). In order to use such techniques, it is necessary to apply the knowledge of experts in the topography of the studied region, and it can take a long time to interpret and output results. While these techniques may be considered dated, they remain in use and can be applied along with GIS, serving usually as a means to assist in the accuracy assessment of images classified using other techniques.

Land use and land cover classification has undergone significant changes over time, due to technological advances in image processing and data acquisition technologies. Analyzes have shifted from manual and visual to digital techniques using filters and numerical classification. Some studies in the literature discuss the historical development of various methodologies (Phiri & Morgenroth, 2017), while others suggest a growing trend in the adoption of a particular technique (Talukdar et al., 2020; Maxwell et al., 2018; Mora et al., 2014). In this sense, the scientific literature lacks quantification on the utilization of land use and land cover classification techniques. This quantification is essential in understanding the evolution in applied methods in studies in the area, which the present works aims to fulfill.

3. MATERIALS AND METHODS

A search for scientific articles was conducted in the Scopus database, with the aim of obtaining journal articles that employed land use and land cover classification methods during the period between 2012 and 2022. The adopted strategy used the search term “TITLE-ABS -KEY (land AND (cover OR use) AND classification) AND PUBYEAR > 2011 AND PUBYEAR < 2023 AND (LIMIT-TO (DOCTYPE, "ar"))”, resulting in a total of 5406 documents, which were exported in CSV format, separated year by year, due to platform limitations.

The CSV files were then converted into tables within a SQLite database, employing the SQLite DB Browser tool. The data was analyzed utilizing queries in the “Author Keywords”, “Index Keywords and “Abstract” fields.

The study utilized search terms consisting of method names and their abbreviations extracted from relevant literature to query the database. The results were recorded for each query, and the corresponding keywords and methods have been compiled in Table 1.

Table 1 - List of keywords searched and their associated methods

Category	Method	Keywords
Clustering		clustering
	ISODATA	isodata
	K-Means	kmeans, k-means, k means
	Semantic Clustering	semantic clustering
Machine learning		machine learning
	Deep Learning	deep learning
	Random Forest	random forest, decision forest
	Decision Tree	decision tree
	Artificial Neural Networks	neural network
	Support Vector Machine	support vector, svm
	K-Nearest Neighbor	nearest neighbo, k-nearest, k nearest, knn
	Classification and Regression Tree	cart, classification and regression tree
	Autoencoder	autoencoder, auto encoder
Statistical methods		
	Maximum likelihood	maximum likelihood, mle
	Markov chain	markov
	Mahalanobis distance	mahalanobis
	Principal Component Analysis	principal component, pca
Land use/land cover Indices	Not applicable	normalized difference, ndvi, ndbi, indice, index, rvi, ndwi,, dbi, dbsi
Others		
	Overlay analysis	overlay analysis
	Cellular automata	cellular automata
	Image segmentation	image segmentation, thresholds
	Analytic Hierarchy Process	analytic hierarchy
Manual analysis		visual interpretation, visual analysis
		manual interpretation, manual analysis

After completing the initial searches, some articles were identified for manual analysis to locate classifiers types not included in the original search terms. This process was repeated until no new classifiers were found. The results were then tabulated and quantitatively analyzed, in order to determine the amount of published articles in the field, as well as the most commonly used classifiers and their possible usage trends. Charts were also created to help visualize and discuss the results.

4. RESULTS

When analyzing the results, the first piece of relevant information is the total number of articles using land use and land cover classification techniques, shown in Figure 1. There is a noticeable uptick in publication volume from 2012 to 2022, with quantitative values rising from below 500 articles in 2013 to more than 1000 articles 2020 and beyond. This increase aligns with predictions based on the widespread adoption of techniques and tools, resulting in a growing number of researchers from a variety of fields conducting these types of studies.

Figure 1 - Number of articles published between 2012 and 2022 on the subject of "land use and land cover classification" on the Scopus database

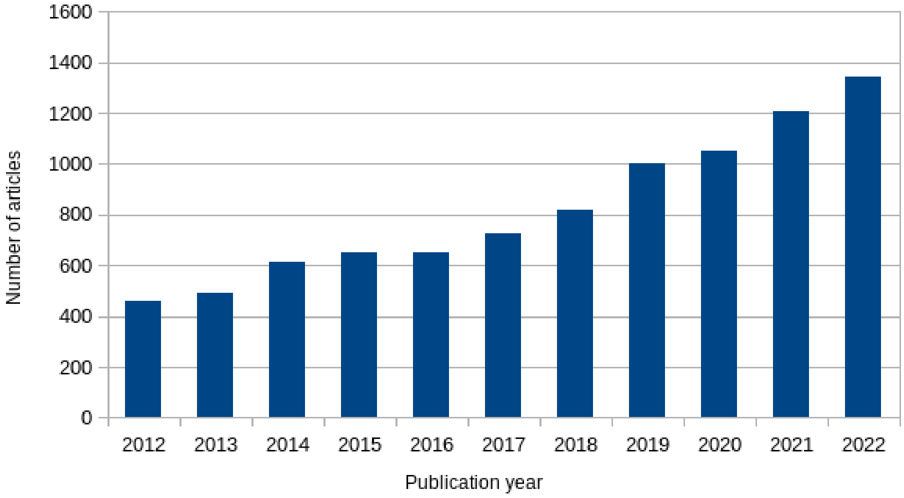
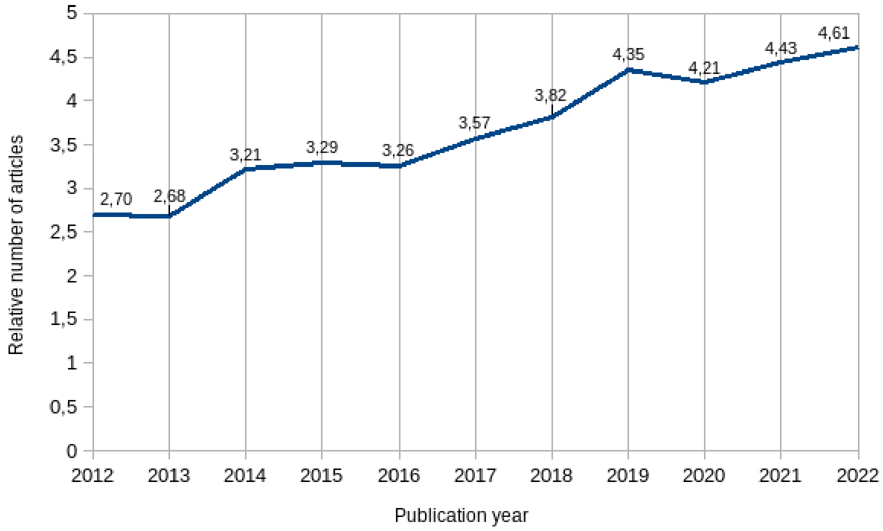


Figure 2 provides additional support for this conclusion. The analysis of article publications indicates a consistent increase in the topic's coverage. From 2012 to 2022, the number of articles per 10,000 publications rose from 460 to 1341, which signifies an upsurge in interest in this research area. This increase in publications demonstrates a growing significance of the specific topic in contrast to all topics in the database.

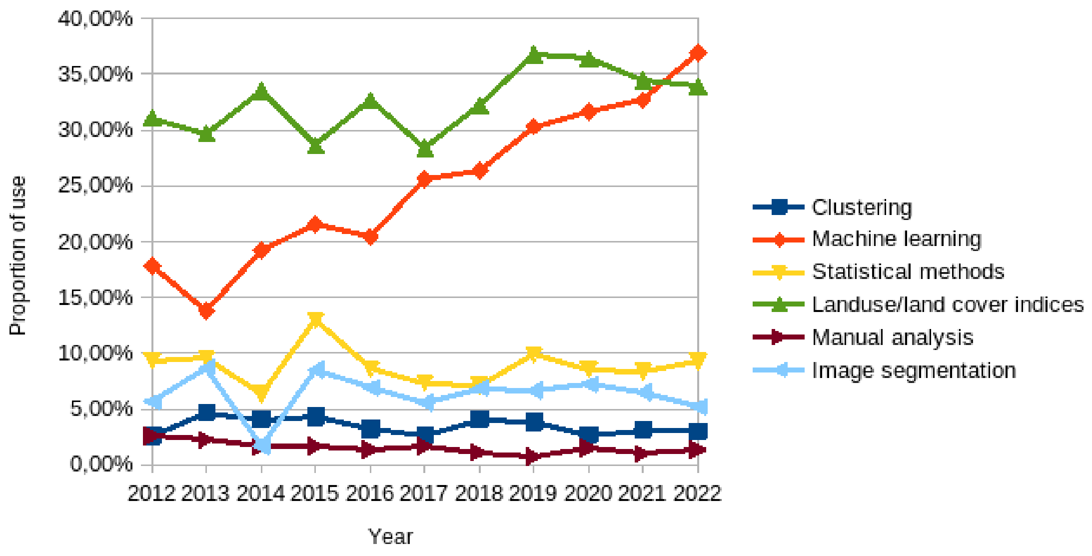
Figure 2 - Number of articles published on the topic for every 10,000 articles published on all topics



Several factors have likely contributed to the rise in research interest. These factors include the increased accessibility of data sources and tools, which makes research in the field more facilitated and widespread. Additionally, researchers are making greater use of computational techniques and there has been an expansion in the field, propelled by the increased interest in scholars from multiple disciplines who are intrigued by the methods utilized in land cover and land use studies. In this regard, land use and land cover classification presents itself as a growing field of study, with potential for growth and development of improvements.

Figure 3 illustrates the primary categories of land use and cover classification methods, including the number of articles containing relevant keywords discovered, for each year studied. It is worth noting that the "Indexes" category refers to the application of indexes computed from the radiometric values of the source data, which, although not a classification technique, is often utilized for this purpose. Hence, the inclusion of this category was considered relevant.

Figure 3 - Usage percentages of the different techniques from 2012 to 2022



Machine learning techniques were predominant, with their percentage of use more than doubling over the period studied, starting from a presence in less than 20% of the total studies until 2014, to a presence in more than 30% of the same, from 2019. No growth or reduction is observed in other categories, but rather a stability in their proportion of use, except for a growth of about 5% in the use of normalized difference indices, while the use of the other techniques presented a pattern of stability.

Upon comparing Figures 1 and 3, it becomes apparent that the growth curve in the utilization of machine learning methods mirrors the growth curve in land use and land cover studies. In other words, the overall growth in the field has been driven primarily by the popularization and application of machine learning techniques, so that there has been no decrease in the use of other methods, but rather a significant increase in the utilization of machine learning. Consequently, this discovery suggests that a sizeable portion of the recent surge in research on land use and land cover is attributable to scholars from disparate disciplines. The observed growth in these methods may have been the main driver of the increased number of publications in this area.

To identify variations for each method, along with all data obtained, it is possible to observe them from Table 2.

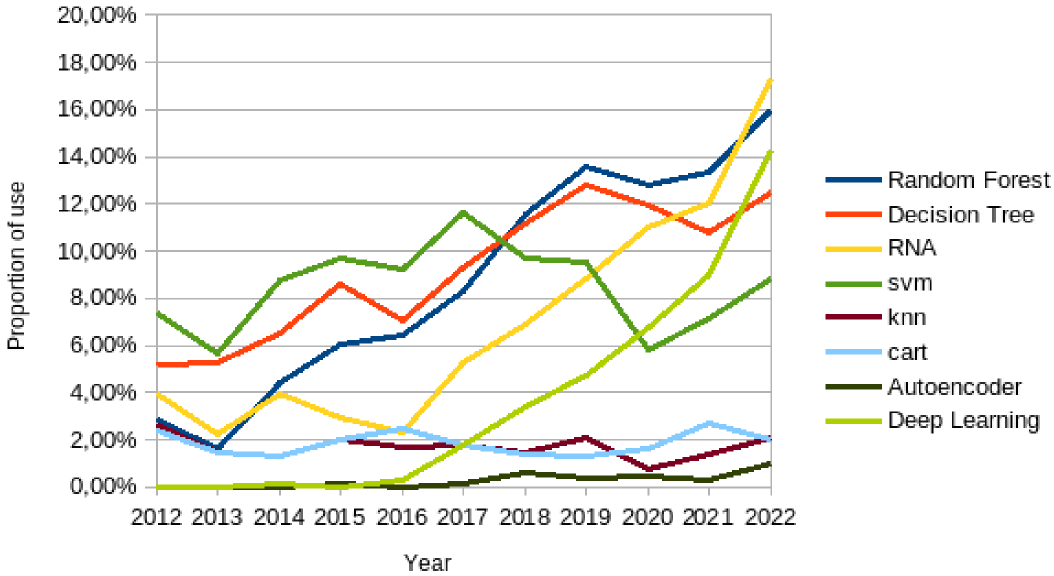
Table 2 – Percentage of articles containing each searched keyword

Category	Subcategory	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Clustering		2,61%	4,67%	4,07%	4,31%	3,22%	2,63%	4,04%	3,83%	2,66%	3,13%	3,68%
	ISODATA	0,00%	1,22%	0,65%	0,92%	0,46%	0,42%	0,25%	0,50%	0,10%	0,08%	0,09%
	K-Means	0,43%	0,81%	1,14%	0,46%	1,23%	0,97%	1,10%	0,50%	0,67%	0,99%	0,64%
	Semantic Clustering	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Machine learning		17,83%	13,82%	19,22%	21,57%	20,40%	25,62%	26,35%	30,31%	31,65%	32,70%	45,26%
	Random Forest	2,83%	1,63%	4,40%	6,01%	6,44%	8,31%	11,52%	13,60%	12,83%	13,34%	16,01%
	Deep Learning	0,00%	0,00%	0,16%	0,00%	0,31%	1,80%	3,43%	4,73%	6,75%	8,98%	14,26%
	Decision Tree	5,22%	5,28%	6,51%	8,63%	7,06%	9,28%	11,15%	12,79%	11,98%	10,79%	12,51%
	Artificial neural networks	3,91%	2,24%	3,91%	2,93%	2,30%	5,26%	6,86%	8,86%	11,03%	12,03%	17,30%
	SVM	7,39%	5,69%	8,79%	9,71%	9,20%	11,63%	9,68%	9,57%	5,80%	7,17%	8,83%
	K-Nearest Neighbor	2,61%	1,42%	1,30%	2,00%	1,69%	1,80%	1,47%	2,11%	0,76%	1,40%	2,12%
	CART	2,39%	1,42%	1,30%	2,00%	2,45%	1,80%	1,35%	1,31%	1,62%	2,72%	2,02%
	Autoencoder	0,00%	0,00%	0,00%	0,15%	0,00%	0,14%	0,61%	0,40%	0,48%	0,33%	1,01%
Statistical methods		9,35%	9,55%	6,35%	12,94%	8,59%	7,34%	7,11%	9,87%	8,56%	8,32%	11,41%
	Maximum Likelihood	4,13%	6,50%	2,93%	7,86%	6,60%	4,85%	4,53%	5,64%	4,94%	4,78%	8,46%
	Markov chain	1,74%	1,63%	2,12%	3,24%	1,38%	1,11%	1,35%	2,42%	2,19%	2,80%	1,75%
	Mahalanobis distance	0,43%	0,41%	0,65%	0,15%	0,00%	0,42%	0,49%	0,30%	0,19%	0,25%	0,37%
	PCA	0,43%	0,81%	0,98%	1,54%	0,46%	1,11%	0,98%	1,11%	1,14%	0,91%	1,84%
Land use/land cover indices		31,09%	29,67%	33,55%	28,66%	32,67%	28,39%	32,23%	36,76%	36,41%	34,43%	41,58%
Manual analysis		2,61%	2,24%	1,63%	1,69%	1,38%	1,66%	1,10%	0,70%	1,52%	0,99%	1,66%
Image segmentation		5,65%	8,74%	7,82%	8,47%	6,90%	5,54%	6,86%	6,65%	7,22%	6,51%	6,35%
Others		0,22%	0,41%	0,33%	0,77%	0,46%	0,55%	0,74%	0,91%	0,76%	1,65%	1,84%

Table 2 indicates that in the field of machine learning, not all methods have seen an increased in usage. Some have decreased, including the SVM method, which was previously the most used until 2017, but undergone a significant decrease in recent years. On the other hand, methods based on artificial neural networks experienced a notable increase in the same period, ultimately becoming the most used approach in the last analyzed year. To better visualize this data, Figure 4 was created, which shows the variation in machine learning techniques identified over the period studied.

Figure 4 – Usage percentage of different machine learning techniques

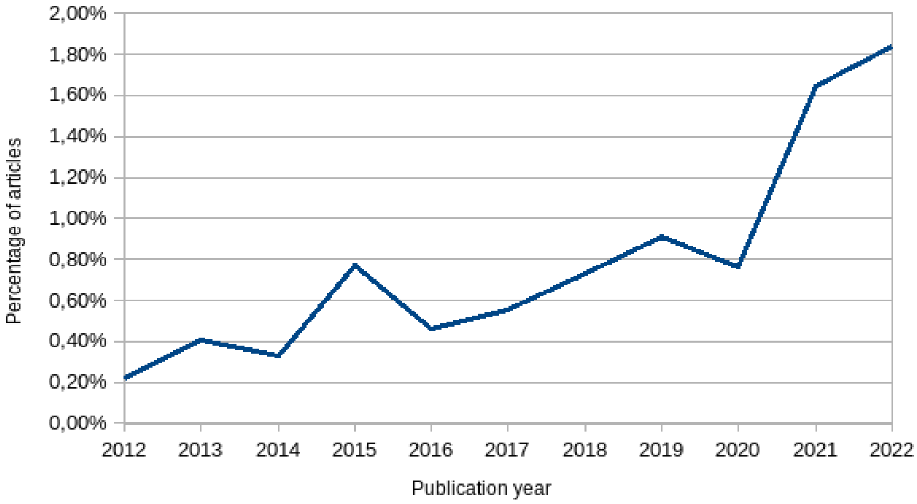
A potential explanation for the surge in the utilization of artificial neural networks since 2016,



although there was no previous trend, is the use of deep learning techniques in more recent studies. These techniques are based on artificial neural networks that employ multiple layers in the network, as opposed to other approaches, deemed "shallow".

Still in Table 2, there are no discernible patterns of growth or reduction in the utilization of individual techniques within statistical or clustering methods, highlighting their stable use. However, it is possible to observe an increase in the "other" category, as illustrated in Figure 5.

Figure 5 - Percentage of articles applying uncategorized techniques



The "Others" category comprises articles with lesser used methods, which could not be grouped, as well as articles containing keywords related to land use and land cover classification, but without referring to the researched methods. The articles in this category underwent manual analysis, aiming to find methods not included in the search and to create new categories, leaving those that did not have a relevant number of occurrences for an individual category. Thus, the increase in this category may be correlated with the scientific community's heightened interest in land use and land cover research, resulting in more efforts towards the development of new classifiers, as well as greater usage of less commonly used methods and experimentation. Thus, the increase may suggest a positive outlook for this field of study, with future prospects for the availability of new methods for classification, which, in turn, could lead to improvements in study results or lower barriers of entry for new researchers. interested parties.

5. CONCLUSION

The present study analysed scientific publications on the classification of land use and land cover from the last ten years, spanning from 2012 to 2022. The research identified a growing trend in this field and mapped the usage of various methodologies during this time frame.

It was found that the number of publications focused on this topic surpassed the number of publications on all other topics in the database. One possible explanation for the increase in studies within the field is the ease facilitated by modern data sources, techniques, and equipment, which decrease entry barriers for inexperienced researchers and attract those from other fields.

Since 2013, there has been a marked rise in the prevalence of machine learning-based methodologies, which currently feature in over 35% of studies. During the same period, other techniques did not experience a significant decrease or increase. This implies that the upsurge in studies regarding land use and cover classification resulted from a rise in the utilization of machine learning techniques, rather than a decline or decrease of interest in other methodologies.

The study of land use and cover has gained popularity in recent years, prompting the development of new methodologies, techniques, and computing improvements. However, the dearth of quantitative data on this topic poses a challenge to meaningful discourse, often relying solely on expert opinions and personal perceptions. The current work achieved its objective of gathering scientific research data and identifying the utilization profile of various techniques for land use and cover classification in an objective manner. The obtained information broadens our comprehension of the field's advancements and supports future research.

6. REFERENCES

Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. *Journal of Physics: Conference Series*, *1142*(1), 012012. https://doi.org/10.1088/1742-6596/1142/1/012012

Belgiu, M., & Csillik, O. (2018). Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sensing of Environment*, *204*, 509–523. https://doi.org/10.1016/j.rse.2017.10.005

Borra, S., Thanki, R., & Dey, N. (2019). *Satellite Image Analysis: Clustering and Classification*. Springer.

Chaves, M. E. D., Picoli, M. C. A., & Sanches, I. D. (2020). Recent Applications of Landsat 8/OLI and Sentinel-2/MSI for Land Use and Land Cover Mapping: A Systematic Review. *Remote Sensing*, *12*(18), Article 18. https://doi.org/10.3390/rs12183062

Frazier, B. E., & Shovic, H. (1980). Statistical methods for determining land-use change with aerial photographs. *Photogrammetric Engineering and Remote Sensing*, *46*(8), 1067–1077.

Horning, N. (2004). *Land cover classification methods, Version 1.0.* American Museum of Natural History, Center for Biodiversity and Conservation. https://www.amnh.org/content/download/74344/1391366/file/LandCoverClassification_Final.pdf

Jozdani, S. E., Johnson, B. A., & Chen, D. (2019). Comparing Deep Neural Networks, Ensemble Classifiers, and Support Vector Machine Algorithms for Object-Based Urban Land Use/Land Cover Classification. *Remote Sensing*, *11*(14), Article 14. https://doi.org/10.3390/rs11141713

Lei, T., & Nandi, A. K. (2022). *Image Segmentation: Principles, Techniques, and Applications*. Wiley.

Mai, S. D., & Ngo, L. T. (2018). Multiple kernel approach to semi-supervised fuzzy clustering algorithm for land-cover classification. *Engineering Applications of Artificial Intelligence*, *68*, 205–213. https://doi.org/10.1016/j.engappai.2017.11.007

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, *39*(9), 2784–2817. https://doi.org/10.1080/01431161.2018.1433343

Mora, B., Tsendbazar, N.-E., Herold, M., & Arino, O. (2014). Global Land Cover Mapping: Current Status and Future Trends. In I. Manakos & M. Braun (Eds.), *Land Use and Land Cover Mapping in Europe: Practices & Trends* (pp. 11–30). Springer Netherlands. https://doi.org/10.1007/978-94-007-7969-3_2

Phiri, D., & Morgenroth, J. (2017). Developments in Landsat Land Cover Classification Methods: A Review. *Remote Sensing*, *9*(9), Article 9. https://doi.org/10.3390/rs9090967

Rebala, G., Ravi, A., & Churiwala, S. (2019). Machine Learning Definition and Basics. In G. Rebala, A. Ravi, & S. Churiwala (Eds.), *An Introduction to Machine Learning* (pp. 1–17). Springer International Publishing. https://doi.org/10.1007/978-3-030-15729-6_1

Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y.-A., & Rahman, A. (2020). Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review. *Remote Sensing*, *12*. https://doi.org/10.3390/rs12071135

Wang, J., Bretz, M., Dewan, M. A. A., & Delavar, M. A. (2022). Machine learning in modelling land-use and land cover-change (LULCC): Current status, challenges and prospects. *Science of The Total Environment*, *822*, 153559. https://doi.org/10.1016/j.scitotenv.2022.153559

Wulder, M. A., Coops, N. C., Roy, D. P., White, J. C., & Hermosilla, T. (2018). Land cover 2.0. *International Journal of Remote Sensing*, *39*(12), 4254–4284. https://doi.org/10.1080/01431161.2018.1452075

DECLARATION OF CONTRIBUTIONS TO THE ARTICLE - CRediT

ROLE	Gonçalves	Matias	Shimoya
Conceptualization – Ideas; formulation or evolution of overarching research goals and aims.	X	X	
Data curation – Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.	X		
Formal analysis – Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data.	X	X	
Funding acquisition - Acquisition of the financial support for the project leading to this publication.			
Investigation – Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection.	X		
Methodology – Development or design of methodology; creation of models.	X		
Project administration – Management and coordination responsibility for the research activity planning and execution.	X	X	X
Resources – Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.	X	X	
Software – Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.	X		
Supervision – Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team.		X	X
Validation – Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs.	X	X	X
Visualization – Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.	X		X
Writing – original draft – Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).	X		
Writing – review & editing – Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages.	X		X