

ANÁLISE DE AGRUPAMENTO (CLUSTER) PARA TIPOLOGIA DE PAISAGENS

<https://doi.org/10.4215/rm2020.e19011>

Adalto Moreira Braz ^{a*} - Ivanilton José de Oliveira ^b - Lucas Costa de Souza Cavalcanti ^c
Antônio Campar de Almeida ^d - Eduardo Salinas Chávez ^e

(a) Doutorando em Geografia. Universidade Federal de Goiás (UFG), Jataí (GO), Brasil.

ORCID: <https://orcid.org/0000-0002-6376-6965>. **LATTES:** <http://lattes.cnpq.br/4717927894301420>.

(b) Dr. em Geografia. Professor da Universidade Federal de Goiás (UFG), Goiânia (GO), Brasil.

ORCID: <https://orcid.org/0000-0002-2718-6947>. **LATTES:** <http://lattes.cnpq.br/4172719252263913>.

(c) Dr. em Geografia. Professor da Universidade Federal de Pernambuco (UFPE), Recife (PE), Brasil.

ORCID: <https://orcid.org/0000-0001-9096-138X>. **LATTES:** <http://lattes.cnpq.br/0571151043430712>.

(d) Dr. em Geografia. Professor da Universidade de Coimbra (UC), Coimbra, Portugal.

ORCID: <https://orcid.org/0000-0002-7616-4023>. **LATTES:** <https://coimbra.academia.edu/Ant%C3%B3nioCamparAlmeida/>.

(e) Dr. em Geografia. Professor da Universidad de La Habana (UH), La Habana, Cuba.

ORCID: <https://orcid.org/0000-0001-6392-4380>. **LATTES:** <http://lattes.cnpq.br/6482925481232424>.

Article history:

Received 23 December, 2019

Accepted 30 March, 2019

Publisher 15 May, 2020

(*) CORRESPONDING AUTHOR

Address: UFG – Campus Cidade Universitária (Jatobá). Rodovia BR 364, Km 195, nº 3800, CEP: 75801-615, Jataí (GO), Brasil. Tel: (+ 55 64) 36068264

E-mail: adaltobraz.geografia@gmail.com

Resumo

O trabalho apresenta uma discussão introdutória sobre o uso do agrupamento estatístico (cluster analysis) para classificação e cartografia de paisagens, tendo como embasamento teórico a tipologia de paisagens, através da noção de zonalidade de L. S. Berg e da teoria dos geossistemas de V. B. Sochava. A utilização desta técnica foi feita através do mapeamento de 272 unidades de paisagens no município de Mineiros, no sudoeste de Goiás, com o objetivo de comparar diferentes formas de agrupamento para estabelecer a tipologia de paisagens utilizando uma abordagem upscaling/bottom-up. As unidades de paisagem foram delimitadas por meio dos elementos de morfoestrutura, geologia, geomorfologia, altitude, declividade, densidade de drenagem, solos e uso e cobertura da terra. Foram avaliados o coeficiente de similaridade Jaccard, a métrica da Distância Euclidiana e o algoritmo k-means. Mesmo que as validações de campo e os testes estatísticos (índices) apontem para um cenário discrepante, considerou-se que as métricas de Jaccard e da Distância Euclidiana apresentaram cenários satisfatórios para representação das paisagens. O agrupamento foi importante na otimização dos processos, embora exista a necessidade de diferenciar a significância estatística da espacial. Portanto, a relevância dessa técnica está na colaboração para agrupar e redefinir grandes quantidades de informações, o que apresentaria excessiva morosidade por vias manuais de análise e reagrupamento espacial.

Palavras-chave: Geossistemas; Taxonomia de Paisagens; Cartografia de Paisagens; Geoinformação, Grupo.

Abstract / Resumen

CLUSTER ANALYSIS FOR LANDSCAPE TYPOLOGY

This paper deals with the classification of landscapes by means of cluster analysis, having as a theoretical basis a typology of landscapes, through the notion of zonality by L. S. Berg and the theory of geosystems by V. B. Sochava. Firstly landscape mapping was performed using geoprocessing techniques resulting in 272 landscape units for the municipality of Mineiros (Goiás State, Brazil). This units were defined by union of different elements of morphostructures, lithology, landforms, altitude, slope degree, drainage density, soils and land use. The objective of this work is to compare different forms of grouping to establish the typology of landscapes using an upscaling/bottom-up approach. The similarity coefficient Jaccard, Euclidean Distance metric and k-means algorithm were evaluated. Even though the field validations and statistical tests point to a different scenario, it was considered that the Jaccard and Euclidean Distance metrics presented satisfactory scenarios for the representation of the landscapes. The grouping was important in the optimization of processes, although there is a need to differentiate between statistical and spatial significance. Therefore, the relevance of this technique is in the collaboration to group and redefine a large amount of information that, by means of manual analysis and spatial regrouping, would present excessive delay.

Keywords: Geosystems; Landscape taxonomy; Landscape mapping; Geoinformation, Cluster.

ANÁLISIS DE CLÚSTER PARA TIPOLOGÍA DE PAISAJE

El trabajo presenta la discusión sobre el uso del agrupamiento estadístico (cluster analysis) para la clasificación y cartografía de los paisajes, teniendo como base teórica la tipología de los paisajes, a través de la noción de zonalidad por L. S. Berg y la teoría de los geosistemas por V. B. Sochava. La utilización de esta técnica se realizó mediante la delimitación de 272 unidades de paisaje en el municipio de Mineiros, estado de Goiás, Brasil. Estas unidades fueron delimitadas por medio de elementos como: morfoestructura, geología, geomorfología, hipsometría, inclinación de la pendiente, densidad de drenaje, suelos y el uso y cobertura de la tierra. El objetivo del trabajo utilizar y comparar diferentes formas de agrupamiento para realizar la clasificación tipológica de los paisajes con un enfoque upscaling/bottom-up. Se evaluaron el coeficiente de similitud Jaccard, la métrica de Distancia Euclidiana y el algoritmo de k-medias. Aunque las validaciones de campo y las pruebas estadísticas (índices) apuntan a un escenario diferente, se consideró que las métricas Jaccard y Euclidean Distance presentaron escenarios satisfactorios para representar los paisajes. La agrupación fue importante en la optimización de los procesos, aunque es necesario diferenciar la significación estadística de la espacial. Por lo tanto, la relevancia de esta técnica radica en la colaboración para agrupar y redefinir una gran cantidad de información que, mediante análisis manual y reagrupación espacial, presentaría un retraso excesivo.

Palabras-clave: Geosistemas; Taxonomia de los Paisajes; Cartografía de los Paisajes; Geoinformación, Grupo.

INTRODUÇÃO

Em Geografia Física, é notável que unidades de paisagem com localizações distintas, possam vir a apresentar características estruturais semelhantes. Num esforço de generalização teórica, é possível classificar tais unidades, criando assim uma tipologia (ISACHENKO, 1991). A tipologia constitui um sistema de classificação (taxonomia) da estrutura das paisagens, baseado em elementos similares que possuem dimensão espacial característica, agrupados por meio de critérios definidos, podendo estar ou não subordinados. Através da tipologia, distinguem-se geossistemas por meio de sua semelhança (homogeneidade) e repetição (espacial), tornando-se fundamental no estudo das paisagens.

De acordo com Rodriguez e Silva (2019, p.34), para entender a esfera da paisagem é essencial a classificação de suas subdivisões. No momento atual, fundamenta-se nos indicadores morfológicos e funcionais, e na subdivisão dos geossistemas.

A construção da tipologia baseia-se em princípios de analogia, homogeneidade, repetibilidade, pertencimento a um mesmo grupo e existência de descontinuidades areais entre seus limites. Numa perspectiva de análise multiescalar, a construção da tipologia pode se dar em dois sentidos (Figura 1), destacados por Cavalcanti e Corrêa (2013, p.153, grifo dos autores):

As análises realizadas em multiescala ainda podem ser classificadas quanto ao sentido da detecção da hierarquia. Partindo-se de grande escala geográfica para as escalas menores, tem-se uma abordagem *downscaling* ou *top-down*. Partindo-se de pequenas escalas geográficas para escalas maiores, tem-se uma abordagem *upscaling* ou *bottom-up*.

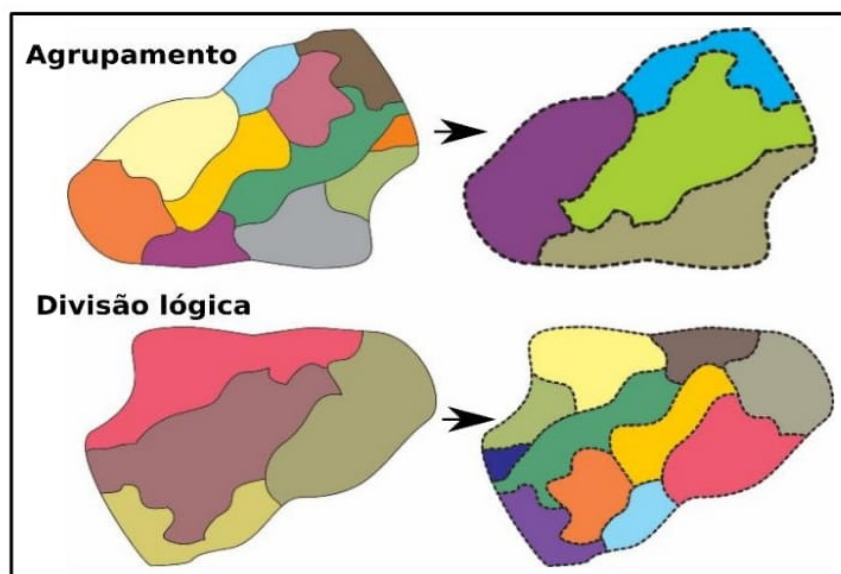


Figura 01 – Agrupamento e divisão lógica para classificação das paisagens. Fonte: Braz, Salinas Chávez e Oliveira (2019).

A classificação é um processo complexo, por vezes envolvendo quantidade elevada de elementos que constituem as paisagens. Uma das possibilidades que emergem neste contexto é a do agrupamento estatístico (cluster) para criar modelos de associações das variáveis, com números elevados de unidades de paisagens¹.

O objetivo deste trabalho é justamente realizar uma avaliação comparativa de diferentes formas de agrupamento (cluster) para estabelecer uma tipologia de paisagens para o município de Mineiros, no estado de Goiás, amparada nos seguintes fatores: morfoestrutura, geologia, geomorfologia, altitude, declividade, densidade de drenagem, solos e uso e cobertura da terra.

REFERENCIAL TEÓRICO

As representações de sínteses das paisagens se dão por meio de ao menos três enfoques, via sistema de classificação taxonômica: tipologia, regionalização e topologia. O que permite confeccionar mapas de unidades de paisagens em diversas escalas e em diferentes condições (ambientais e territoriais), como parte da cartografia de paisagens, além de uma representação espacial das mesmas, possibilita evidenciar também agrupamentos de indivíduos (delimitação de conjuntos espaciais em zonas homogêneas) caracterizados por agrupamentos de atributos ou variáveis, o que comumente se entende através de diferentes unidades de paisagens (ZACHARIAS, 2008; SALINAS CHÁVEZ et al., 2019).

A noção de tipologias surgiu através dos princípios de zonalidade de Berg (1947), e continua a se desenvolver atualmente, com aplicações de análises multivariadas, consideradas um passo fundamental na definição e tratamento lógico de parâmetros setoriais das paisagens (BERUTCHACHVILI e BERTRAND, 1978).

A tipologia representa a classificação das paisagens por meio de sua estrutura, constituída por elementos homogêneos (similares) de acordo com os interesses e a escala de análise do estudo. De acordo com Rodríguez e Silva (2002, p. 98), “A tipologia significa distinguir as unidades pela sua semelhança e repetição, dependendo de determinados parâmetros de homogeneidade”.

Bolòs i Capdevila (1981) indicam que a classificação dos geossistemas consiste em especificar as semelhanças (homogeneidades) entre os indivíduos para reunir aqueles que apresentam possíveis identidades em um táxon. Um táxon é, portanto, um conjunto de indivíduos, neste caso, os geossistemas, que possuem alto grau de semelhança entre si. Além disso, os diferentes taxa também podem ser agrupados em outros níveis (superiores ou inferiores).

À frente do Instituto de Geografia da Sibéria, Sochava (1970; 1975; 1978a; 1978b) conduziu a construção de um sistema taxo-corológico para classificação das paisagens. Este sistema está organizado numa fileira bilateral contendo geômeros, os geossistemas de estrutura homogênea, e geócoros, os geossistemas de estrutura diferenciada. Não obstante, Sochava (1978b, p. 8) destaca que “nenhuma classificação é absoluta; é necessário modificá-la, aperfeiçoá-la”.

A abordagem hierárquica, amplamente discutida por Klijn (1995), levou o autor a considerar que o primeiro e mais importante princípio da hierarquia é que os elementos devem ser baseados em desigualdade em seus relacionamentos. Na hierarquia das paisagens é possível adotar as relações estabelecidas por Klijn (1995), a respeito da homogeneidade ou heterogeneidade, simetria e assimetria, respectivamente (Figura 2).

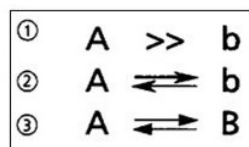


Figura 2 –Relações simétricas (homogêneas) e assimétricas (heterogêneas) nas hierarquias. 1) “A” domina “b” de forma unilateral; 2) “A” domina “b”, mas “b” afeta “A”; 3) “A” e “b” afetam um ao outro de modo comparável. Fonte: Klijn (1995).

Esta sistemática de hierarquizar os geossistemas pressupõe o agrupamento de paisagens em unidades de diferentes classes, escalas ou nível taxonômico. É prática comum que os níveis taxonômicos recebam designação particular: espécies, gênero, tipos de paisagens etc., de acordo com sua especificidade estrutural, genética ou funcional. É neste processo de classificação tipológica (taxonomia) que as propriedades predominantes e também mais importantes das paisagens são estabelecidas (BAYANDINOVA, MAMUTOV e ISSANOVA, 2018; SERRANO GINÉ et al., 2019).

Após o advento da geoinformação, a utilização dos Sistemas de Informações Geográficas (SIG) têm dado um novo impulso às técnicas de mapeamento de unidades de paisagens. Um exemplo que bem representa essa afirmação é o de Isachenko e Reznikov (1995), quando os autores afirmam que

É fácil verificar que, mesmo para o pequeno território na zona de taiga, o número de possíveis cenários

dinâmicos para as paisagens pode chegar a dezenas. O seu mapeamento por via manual tradicional é impossível. Portanto, os futuros passos no caminho da modelagem de paisagens estarão conectados com as possibilidades das tecnologias de SIG (ISACHENKO e REZNIKOV, 1995, p. 804, tradução nossa).

Sobre a possibilidade da modelagem para mapeamento de paisagens, também é possível incluir a utilização de estatística para contribuir nessa tarefa. Em meados da década de 1970, de acordo com Preobrazhenskiy (1983), os primeiros estudos a considerarem tal possibilidade foram os de Aleksandrova (1975), que avaliou paisagens a partir da correlação de um modelo geográfico, matemático e estatístico por meio de técnica automatizada.

Pouco tempo depois, Kuprianova (1977) analisou a correlação geográfica da diferenciação natural da regionalização das paisagens, aplicando procedimentos computadorizados e automáticos em objetos complexos (PREOBRAZHENSKIY, 1983).

PROCEDIMENTOS

O mapeamento se iniciou a partir do cruzamento de 8 elementos: morfoestrutura, unidades geológicas, relevo, hipsometria (altitudes), declividade, densidade de drenagens, solos e uso e cobertura da terra (Tabela 1).

Arquivo	Fonte	Metadado
Unidades geológicas	Moreira et al. (2008) SIEG*	1:500.000
Relevo	Radambrasil (1983) IBGE (2018)**	1:250.000
Solos	Nunes (2015)	1:50.000
Hipsometria (altitude)	USGS / Earth Explorer	Imagem SRTM-X de 30m (resolução espacial) Escala 1:100.000
Declividade	Autores (2018)	Processada a partir de imagem SRTM-X. Escala 1:100.000
Densidade de drenagem	Autores (2018)	Processada a partir dos cursos d'água do Macro ZAEI (Cunha et al., 2014) e disponibilizados pelo SIEG. Escala 1:100.000
Uso e cobertura da terra	Autores (2018)	Processado a partir de imagens Landsat-8 Sensor OLI (USGS/Earth Explorer) de 30m e 15m (resolução espacial) – 13/07/2017. Refinado a partir de imagens CBERS-4 sensor PAN (INPE) de 10m e 5m (resolução espacial) – 07/01/2017; 29/03/2017; 15/06/2017; 11/07/2017. Refinado a partir de dados de flora dos <i>sites</i> de coleta do NeoTropTree (OLIVEIRA-FILHO, 2015; 2017). Escala 1:50.000
Morfoestrutura	Autores (2018)	Elaborada a partir da síntese entre unidades geológicas, relevo, solos, hipsometria e densidade de drenagem Escala 1:100.000

* O Sistema Estadual de Geoinformação de Goiás (SIEG) organiza e disponibiliza os dados, mas não os produz.

** O IBGE (2018) adaptou (compilou e reinterpreto) a base cartográfica de referência do Radambrasil (1983).

Tabela 1 - Características dos dados utilizados. Fonte: Autores (2018).

O cruzamento dos elementos utilizando dados vetoriais (shapefiles) resultou num arquivo com 67.691 feições. O fluxograma da Figura 3 resume as etapas da pesquisa.

O resultado foi convertido para um dado raster (.tiff), reclassificado e submetido à lógica fuzzy (fuzzyficação). Nesse processo de conversão, o raster passa a admitir 245 intervalos (a lógica binária de um raster em 8bits permite um máximo de 255 intervalos).

A classificação difusa [lógica fuzzy] assume o limite entre duas classes vizinhas como uma área contínua sobreposta na qual um objeto possui participação parcial em cada classe. Esse ponto de vista não apenas

reflete a realidade de muitos aplicativos nos quais as categorias têm limites difusos, mas também fornece uma representação simples da partição potencialmente complexa do espaço de recursos (ZHENG e KAINZ, 1999, p. 79, tradução e grifo nosso).

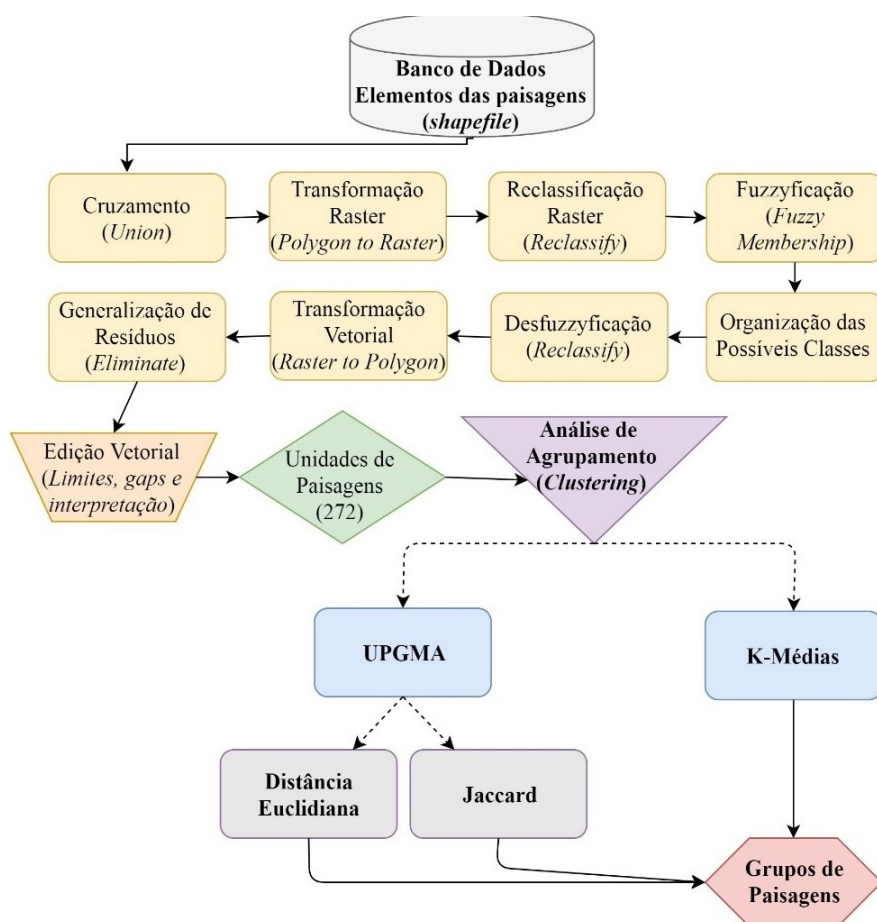


Figura 3 – Fluxograma dos procedimentos para agrupamento das unidades de paisagens. Fonte: autores (2019).

No contexto da cartografia de paisagens, a lógica fuzzy é relevante principalmente na definição dos limites dos geossistemas, pois, de acordo com Marques Neto (2016), os limites entre paisagens podem ser abruptos e bem marcados, mas podem também apresentar passagens de uma unidade para outra de modo difuso e interdigitada que, por vezes, impossibilitam uma representação cartográfica adequada e mais precisa.

O próximo passo constou na desfuzzyficação do raster, e sua transformação em vetor (shapefile), para o tratamento dos resíduos e correção das confusões no agrupamento ou separação das possíveis unidades de paisagens, a partir da generalização cartográfica, optando-se por uma área cartografável de no mínimo 5ha.

Conforme especificado por Salinas Chávez e Ramón Puebla (2013), a área mínima cartografável de 5 ha corresponde a representações cartográficas na escala de 1:50.000. Após os ajustes necessários, as 245 classes foram, então, reorganizadas manualmente (vetorização), revisadas e finalmente obteve-se como resultado 272 unidades de paisagens para o município de Mineiros (GO).

O mapa de unidades de paisagens – e posteriormente de grupos de paisagens – está representado na escala de 1:100.000, com a intenção de valorizar as diferenciações paisagísticas (por um princípio de homogeneidade²). Para os procedimentos anteriormente descritos, a geoinformação foi organizada num Sistema de Informação Geográfica (SIG), utilizando-se o software ArcGIS 10.4.1.

O grande número de unidades de paisagens (272) gerou a necessidade de criar grupos (clusters)

que fossem semelhantes. Esta técnica, comum na estatística para quantidades elevadas de números, é também semelhante aos princípios de mapeamento das paisagens, interessada em determinar elementos homogêneos em sua estrutura, e heterogêneos entre si, passíveis de “agrupamentos” ou “separações” em diferentes hierarquias para, finalmente, simplificar sua representação cartográfica.

A tipologia foi realizada por meio da análise de agrupamentos (clustering), aplicando-se estatística multivariada. A necessidade de se adotar tal procedimento emerge da quantidade de objetos de observação. Desta maneira, os esforços estiveram concentrados no agrupamento dessas unidades em hierarquias superiores com certo grau de similaridade, visando a representação de grupos de paisagens através do clustering.

A análise de agrupamentos (clustering) foi realizada por meio do software estatístico PAST (3.25), e optou-se pelo Método do Grupo de Pares não Ponderados com Média Aritmética (Unweighted Pair Group Method with Arithmetic Mean – UPGMA), usando os coeficientes de similaridades da Distância Euclidiana e Jaccard. Além disso, foi utilizado o algoritmo K-médias (K-means).

O UPGMA calcula as médias das distâncias ou similaridades entre uma unidade de paisagem e cada uma das outras unidades, ao passo que todas elas recebem o mesmo peso, sendo a matriz (de distância ou similaridade) atualizada e reduzida a cada etapa de execução do algoritmo. Trata-se, portanto, de uma estratégia aglomerativa (bottom-up) (LEGENDRE e LEGENDRE, 2012).

Sobre a aplicação do UPGMA, Metz (2006, p.23) destaca que

essa abordagem, assim como as outras, constrói os agrupamentos de modo que exemplos pertencentes ao mesmo cluster possuem alta similaridade e exemplos pertencentes a clusters diferentes possuem baixa similaridade. [...] Entretanto, uma distinção entre essa abordagem e as demais é que o resultado obtido não é constituído apenas de uma partição do conjunto de dados inicial, mas sim de uma hierarquia que descreve um particionamento diferente a cada nível analisado.

Em relação aos coeficientes utilizados neste trabalho, é preciso salientar que a distância euclidiana, que foi “definida pelo matemático grego Euclides e representa a menor distância existente entre dois objetos no plano multidimensional” (MACHADO, 2011, p. 29), é a métrica de distância mais utilizada em análises de agrupamentos (clustering), conforme explica Metz (2006), e é definida pela Equação 1:

$$dist(E_i, E_j) = \sqrt{\sum_{l=1}^M (x_{il} - x_{jl})^2}$$

(1)

O coeficiente de Jaccard (JACCARD, 1901) é uma métrica de similaridade, compreendida como um coeficiente usado para medidas de associação quando as características são descritas por apenas dois valores discretos, por exemplo, 1 ou 0. O coeficiente de Jaccard considera que a correspondência entre os valores 0-0 (não existentes) é menos importante que a dos valores 1-1 (existentes). Isso ocorre pois, na maior parte das aplicações, o valor 1 é usado para atributos com presença da característica descrita e o valor 0 indica a ausência da característica (Equação 2) (METZ, 2006). “Assim, esse algoritmo compara o número de presenças das variáveis comuns e o número total de variáveis envolvidas, excluindo o número de ausências conjuntas” (MEYER, 2002, p. 9).

$$J(E_i, E_j) = \frac{a_{11}}{a_{10} + a_{01} + a_{11}}$$

(2)

Além da utilização de um algoritmo hierárquico (UPGMA), foi utilizado um algoritmo não hierárquico e não supervisionado, o K-means. Este tem suas atribuições de cluster inicialmente aleatórias. Adota um procedimento iterativo, onde os itens são movidos para o cluster que tem a média

mais próxima do agrupamento. O procedimento é repetido continuamente, até que os itens não estejam mais próximos de outros clusters (HAMMER, 2019).

O k-médias (k-means), proposto por MacQueen (1967), é uma heurística, ou seja, quando um algoritmo leva em consideração duas propriedades em sua estrutura de criação. Além disso, o k-médias é considerado um algoritmo de agrupamento não supervisionado, por gerar agrupamentos a partir de números de classes pré-determinadas.

No caso do k-médias, é o seu tempo de execução, a partir de estratégias para tornar a escolha simples e rápida, minimizando a distância dos elementos a um conjunto de k-centros dado por $\square = \{x_1, x_2, \dots, x_k\}$ de forma iterativa. O k-médias depende de um parâmetro (k = número de clusters) pré-estabelecido pelo usuário. Esse é um algoritmo não hierárquico, dado pela Equação 3 (LINDEN, 2009):

$$d(P, X) = \frac{1}{n} \sum_{i=1}^n d(P_i, X)^2$$

(3)

Operacionalmente, a construção da tipologia apropriou-se das 272 unidades e cada uma das 71 classes dos 8 elementos usados para o mapeamento das unidades, organizando-os em uma matriz binária de “presença-ausência” – atribuindo 0 para elementos ausentes na unidade de paisagem e 1 para elementos presentes na unidade. Este procedimento é o que define as semelhanças ou diferenças entre os elementos de entrada (input), a partir de uma distância, neste caso, a distância euclidiana ou coeficiente de similaridade. Sobre isso, Linden (2009, p.18-19) destaca que

O critério baseia-se normalmente em uma função de dissimilaridade, função esta que recebe dois objetos e retorna a distância entre eles. [...] Os grupos determinados por uma métrica de qualidade devem apresentar alta homogeneidade interna e alta separação (heterogeneidade externa). Isto quer dizer que os elementos de um determinado conjunto devem ser mutuamente similares e, preferencialmente, muito diferentes dos elementos de outros conjuntos (LINDEN, 2009, p. 18-19).

Assim, realizou-se o agrupamento das 272 unidades de paisagens em grupos de paisagens, cujo procedimento utilizando UPGMA (Distância Euclidiana e Jaccard) resulta num dendrograma, formado pelo agrupamento das unidades de paisagens em níveis de distância dos clusters formados. Já o k-médias resulta diretamente numa matriz contendo os intervalos dos clusters.

Como forma de validação dos agrupamentos, assumiu-se que os tipos de paisagens (16 tipos) constituem a realidade de campo, tendo em vista que mesmo sendo produtos gerados a partir do cluster, tiveram seu refinamento manual a partir dos trabalhos de campo.

A validação, portanto, consistiu na escolha de duas localidades de maior conhecimento e com ampla averiguação de campo, que foram comparadas entre os limites dos tipos de paisagens, pontos de validação de campo e em relação a cada um dos agrupamentos resultantes do UPGMA e K-means.

Além da aferição manual, baseada em informações de campo, foram incluídos também testes estatísticos para validação dos clusters, processados através da função “evalclusters” do MATLAB, utilizando os índices de Silhouette, Davies-Bouldin e Gap.

RESULTADOS E DISCUSSÃO

A partir do cruzamento dos elementos, foram delimitadas 272 unidades de paisagens. Embora, num primeiro momento, tal situação seja um indicativo da diversidade de paisagens no município de Mineiros (GO), do ponto de vista da representação cartográfica e da intenção de utilizar a informação das paisagens para um planejamento (ambiental e territorial), fez-se necessária a adoção de outras escalas de análise (geográfica e cartográfica)³.

Essa questão, da escala geográfica no estudo das paisagens, é bastante relevante, visto que na própria teoria dos geossistemas de Sochava (1978a) o autor já apontava para as fileiras bilaterais de

geômeros e geócoros, bem como suas subordens e subcategorias⁴. Solntcev (1949; 2006) também havia lançado a preocupação com a delimitação de unidades de paisagens, tanto no que tange às escalas (geográfica e cartográfica), quanto em relação à quantidade de unidades a serem representadas, devido à sua complexidade de análise.

Os agrupamentos pelo método de grupo de pares não ponderados com média aritmética (Unweighted Pair Group Method with Arithmetic Mean – UPGMA) resultaram em dendrogramas, cujos intervalos adotados nas linhas de corte foram de 3,0 para a distância euclidiana, e de 0,25 para Jaccard (Figura 4).

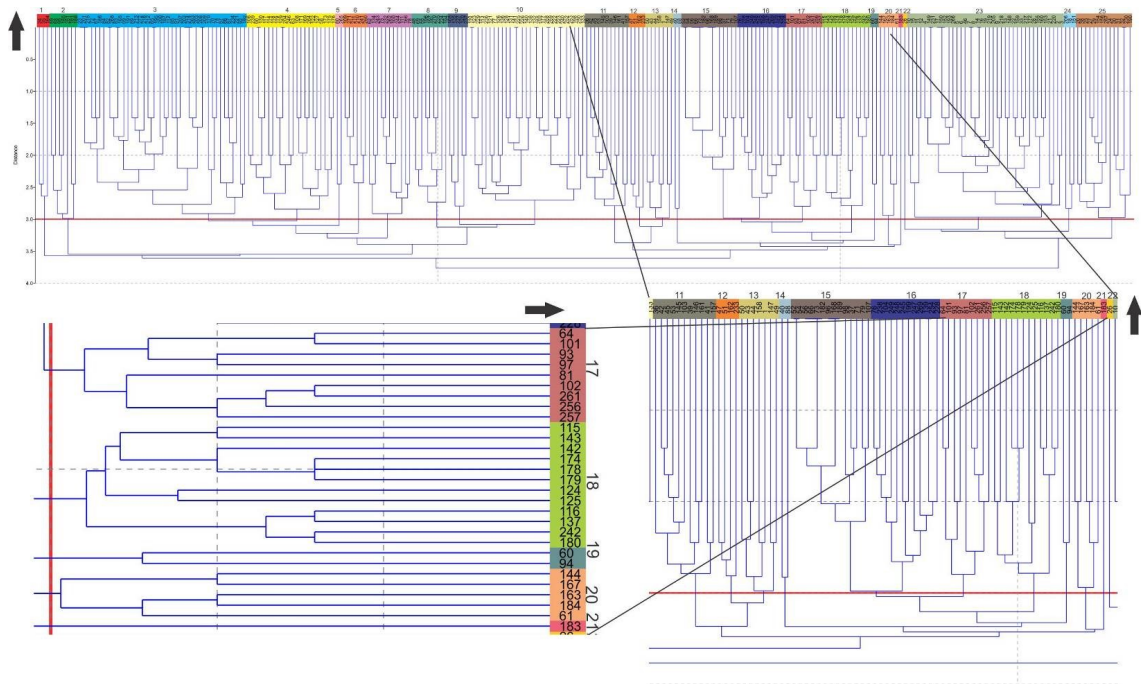


Figura 4 – Dendrograma resultante da análise de cluster através do UPGMA (distância euclidiana).
Fonte: Autores (2019).

Faz-se necessário ressaltar que cada realidade de mapeamento apresentará particularidades quanto à distinção das unidades de paisagens, não havendo uma regra (inflexível) para a escolha de linhas de cortes sempre iguais nesses dendrogramas resultantes. Contudo, a escolha das linhas de corte implica diretamente na quantidade e configuração dos limites das unidades de paisagens (Figura 4).

Com relação ao agrupamento por UPGMA, a determinação de sua qualidade se dá a partir da avaliação do coeficiente de correlação cofenética, sendo que para a distância euclidiana obteve-se 0,7425 e, para Jaccard, 0,759. De acordo com a proposta de Rohlf (1970), correlações cofenéticas > 0,7 são admissíveis para bons agrupamentos.

O resultado para k-médias é dado diretamente por uma matriz indicando os grupos (clusters) das respectivas unidades de paisagens (Tabela 2).

Unidade	Cluster (Grupo)	Unidade	Cluster (Grupo)
1	25	45	2
2	25	46	2
3	25	[...]	[...]
[...]	[...]	270	24
43	17	271	24
44	17	272	24

Tabela 2 – Matriz resultante da análise de cluster por k-médias.

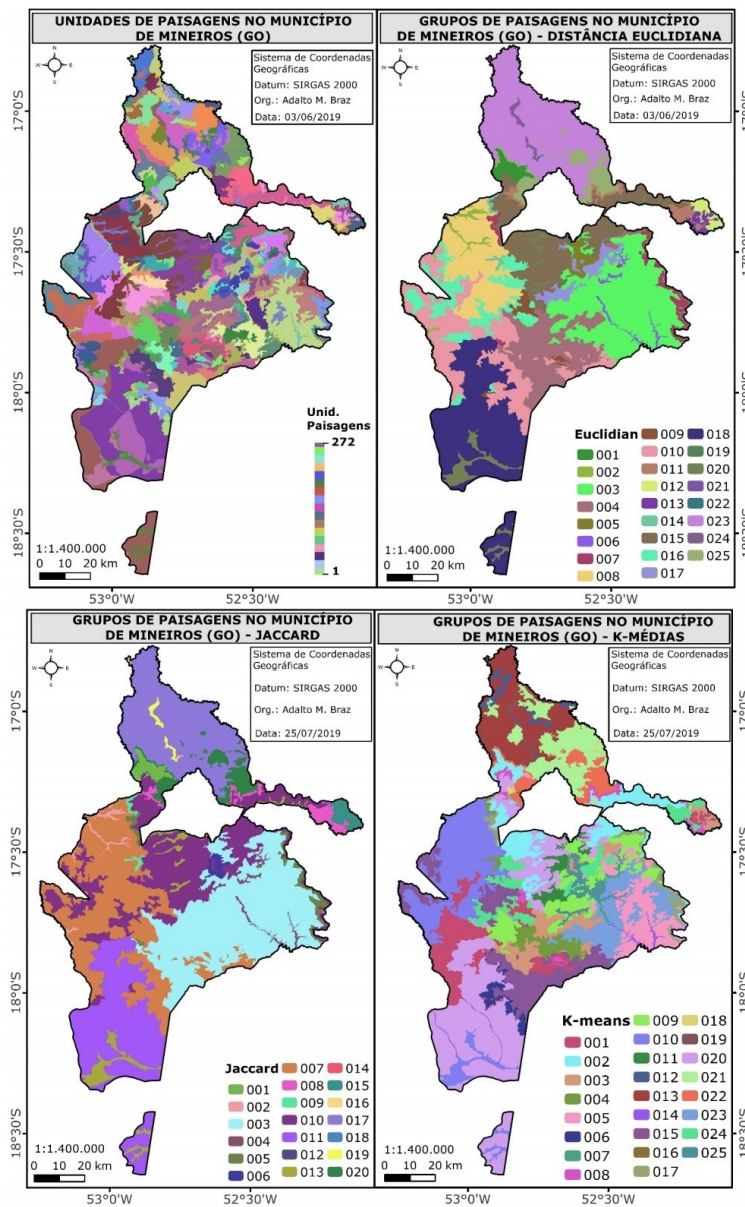


Figura 5 – Mapa de unidades de paisagens e grupos de paisagens por distância euclidiana. Fonte: Autores (2019).

Os clusters são, então, organizados de acordo com as unidades de paisagens, e utilizados para o

mapeamento dos grupos de paisagens (Figura 5).

Gath e Geva (1989) e Silva e Gomide (2004) esclarecem que há pelo menos três requisitos para definir um agrupamento aceitável, sendo: 1) a clara separação entre os grupos resultantes; 2) uma determinada concentração (coesão) de pontos em torno do centro de um grupo; e 3) o menor número possível de grupos, desde que também obedeçam aos requisitos anteriores.

No entanto, tais questões merecem certa cautela quando se trata de agrupamentos envolvendo limites espaciais, pois se reafirma que, neste caso, a significância estatística não considera a significância espacial, atributo de muita importância para a cartografia de paisagens.

Para avaliar os clusters, optou-se por um procedimento manual, ancorado em informações de campo e em uma avaliação estatística, feita com base em índices de validação de clusters.

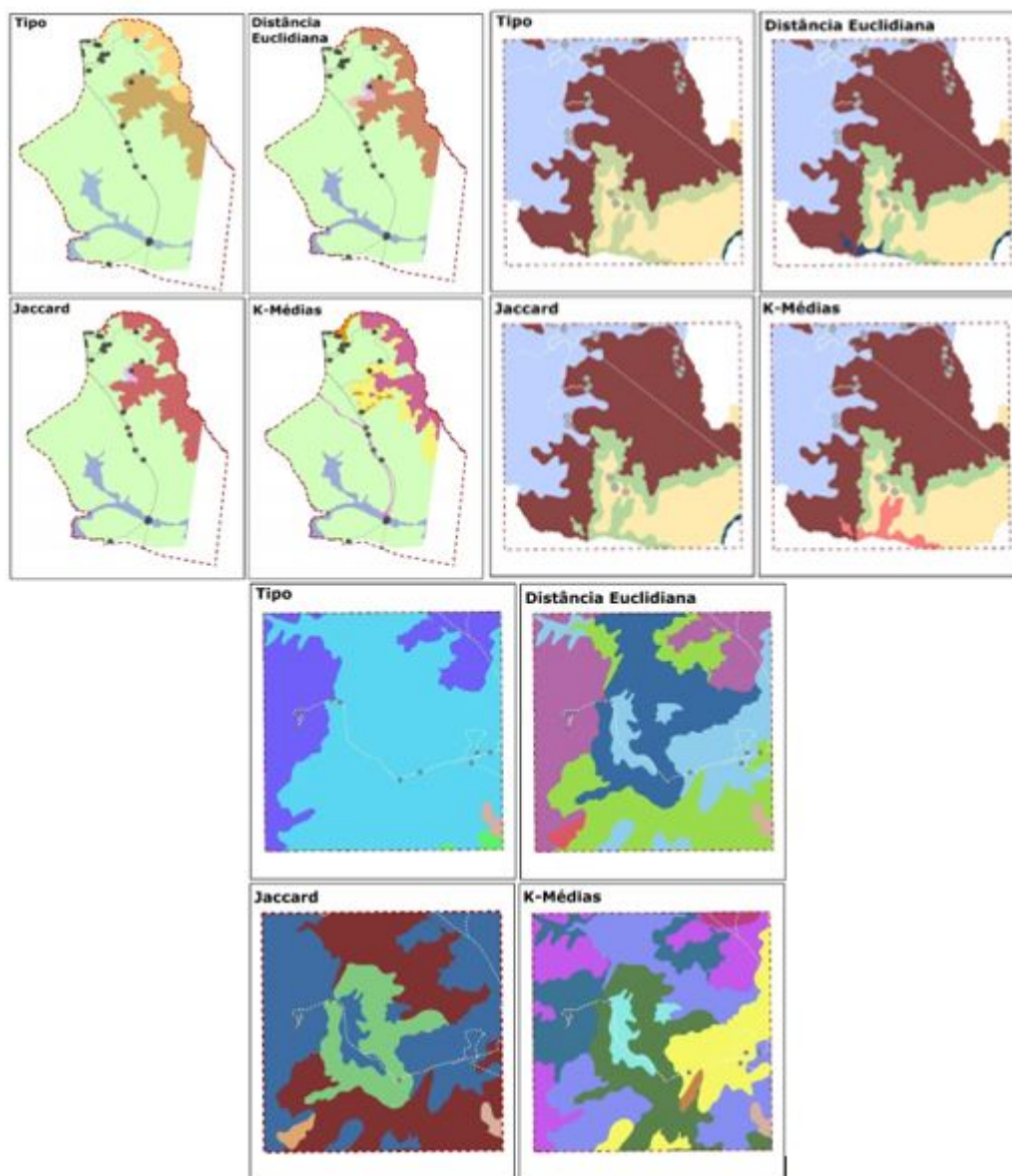


Figura 6 – Validação dos agrupamentos em cada algoritmo e métrica de similaridade do cluster para o sul (Parque Nacional das Emas), norte (Pinga-Fogo) e centro do município de Mineiros (GO).Fonte: Autores (2019).

5

A partir dos resultados da validação manual, considerou-se que o método hierárquico, algoritmo

UPGMA, neste caso, mostrou-se mais vantajoso que o k-means, em termos de agrupamentos confiáveis, do ponto de vista dos limites dos tipos de paisagens, e considerando-se as paisagens de Mineiros. Na validação (Figura 6) foram adotadas três áreas de maior conhecimento de campo e concentração de pontos de aferição, confrontadas aos tipos de paisagens e às paisagens identificadas nas atividades de campo.

A validação estatística utilizou os índices de Silhouette, Davies-Bouldin e Gap. Esses índices procuram avaliar o número ideal de clusters. O índice de Silhouette, apontou para o melhor resultado do cluster com o coeficiente de Jaccard e o k-médias de 25 grupos. O índice de Gap apontou para o cluster mais satisfatório com o uso do algoritmo UPGMA, aplicando-se a métrica de similaridade da distância euclidiana (o que corrobora a avaliação manual, do ponto de vista dos limites das paisagens), seguido do k-médias de 25 grupos. Embora a métrica Davies-Bouldin apontasse para o agrupamento ideal do k-médias de 25 grupos, esse índice não foi capaz de especificar métricas de similaridade (ausentando distância euclidiana e Jaccard), portanto, teve suas considerações descartadas para a avaliação.

Para o procedimento de clustering pela distância euclidiana, adotou-se a linha de corte indicando 25 grupos de paisagens, enquanto por Jaccard, a escolha da linha de corte indicou 20 grupos. Já por k-médias, optou-se pela escolha prévia de 25 grupos de paisagens.

A principal diferença a ser notada nos mapeamentos são dos limites espaciais dos grupos de paisagens resultantes. O agrupamento por Jaccard originou grupos mais extensos, tendendo a ser regulados pelas formações do relevo, litologia e morfoestrutura do município.

A distância euclidiana, embora tenha conservado alguns grupos maiores de paisagens (tal como Jaccard), resultou num conjunto de limites espaciais mais coerente e adequado para a escala desejada do mapeamento (1:100.000). Sendo, por isso, o agrupamento considerado como mais adequado, especificamente para os interesses propostos nesse trabalho.

O agrupamento por k-médias, embora com a mesma quantidade de grupos de paisagens que a distância euclidiana, apresentou maior fragmentação nos limites espaciais, também sem ser possível apontar um dos elementos (solo, relevo etc.) como regulador na formação dos grupos de paisagens.

Desse modo, entende-se que são muito relevantes tanto a questão da escala quanto a significância espacial no estudo das paisagens, devendo-se considerá-las com o mesmo peso atribuído aos procedimentos de agrupamento (cluster) e sua significância estatística.

Por procedimentos tradicionais (manuais), o agrupamento deveria ser realizado de forma analítica, em cada uma das 272 unidades de paisagens, correlacionando-as a partir dos elementos de sua estrutura (relevo, solo, vegetação etc.) e buscando correspondência com unidades análogas (homogêneas).

Entretanto, os algoritmos indicam para o software quais objetos (unidades de paisagens) possuem elementos análogos em sua estrutura; e o software, ao modelar os dados de forma quantitativa (estatística), indica diferentes níveis de agrupamento das unidades de paisagens. Definidos através das linhas de corte sugeridas pelo dendrograma, no caso da distância euclidiana e Jaccard, ou da matriz resultante de k-médias, os níveis de agrupamento também sugerem diferentes escalas geográficas para análise das paisagens.

Ainda que reconhecida a contribuição da análise de cluster para o agrupamento de paisagens, algumas indagações ainda podem ser apontadas.

A primeira questão envolve pensar a necessidade de como podemos obter significância espacial entre as unidades de paisagens a partir dos agrupamentos sugeridos pela classificação do cluster, visto que a significância estatística (cluster) não leva em consideração a significância espacial dos limites das unidades de paisagens.

Outra reflexão se dá a respeito da capacidade da técnica estatística (clustering) determinar elementos preponderantes no agrupamento das paisagens. Por exemplo, a geomorfologia seria, de fato, o principal elemento a condicionar a formação dos grupos de paisagens? Essa situação provoca ainda outra observação: há perspectiva na identificação, a partir da análise de cluster, de possíveis elementos que condicionam os grupos de paisagens que foram definidos?

Uma das problemáticas mais relevantes na cartografia de paisagens é a escala (geográfica e cartográfica). Em vista disso, a análise de cluster por UPGMA (distância euclidiana e Jaccard) indica as linhas de corte, que diferenciarão os níveis dos agrupamentos resultantes.

Logo, a questão perpassa por encontrar uma relação entre os níveis de agrupamento (linhas de corte) com as escalas geográfica e/ou cartográfica que sejam representativas com a complexidade (e realidade) da representação das paisagens. Ou ainda, no caso de k-médias, essa mesma relação junto ao número de clusters pré-determinado pelo usuário.

Assumindo que o agrupamento mais relevante foi, neste caso, o da distância euclidiana, a princípio não se encontrou uma regra, ou mesmo uma tendência, no sentido genético do condicionamento das paisagens que indicasse uma regulação na formação e hierarquia dos geossistemas. Dessa maneira, não foi consenso identificar um ou alguns elementos que fossem preponderantes para a formação ou mesmo agrupamento das paisagens, tendo em vista as três técnicas de agrupamento (distância euclidiana, Jaccard e k-médias).

A consideração a respeito dos resultados apresentados, reforça, por ora, a premissa sobre a organização e formação das paisagens em que todos os elementos dessa estrutura se encontram em inter-relação. O agrupamento obedeceu, portanto, à premissa de um dos modelos teórico-conceituais das paisagens, ao corroborar a definição de Sochava (1978, p. 292), para quem os geossistemas são “uma dimensão do espaço terrestre onde os diversos componentes naturais encontram-se em conexões sistêmicas uns com os outros, apresentando uma integridade definida”. Por isso, considerou-se o resultado satisfatório, com os atributos utilizados exercendo influência descentralizada sobre a formação das unidades de paisagens.

É bastante relevante para os resultados a afirmativa de que houve “ganhos” e “perdas” em cada algoritmo e métrica de similaridade adotados. Ainda assim, a distância euclidiana, através do algoritmo UPGMA, foi a métrica que mais se aproximou da realidade (tipos de paisagens e realidade em campo). Há ainda a afirmação de que a métrica de Jaccard também pode ser considerada adequada para os grupos de paisagens, considerando alguns “ganhos” e “perdas” na representação dos grupos de paisagens⁶.

Já o k-means (com 25 grupos de paisagens) foi o algoritmo considerado menos satisfatório, julgando que seu agrupamento apresentou mais “perdas” para a representação da realidade das paisagens, sobretudo com relação aos limites dos tipos de paisagens do município de Mineiros (GO).

CONCLUSÕES

A utilização de cluster é uma técnica diferencial e ainda pouco explorada em trabalhos no âmbito da Geografia Física, sobretudo no Brasil, mas que demonstrou um considerável potencial em agrupamentos de unidades de paisagens para fins de representação cartográfica.

Considera-se que a utilização de técnicas da estatística (clustering) foi relevante na otimização dos processos. Mesmo assim, é preciso ressaltar que essa é uma possibilidade para subsidiar o mapeamento de paisagens, mas não pode – pelos menos em vista do que foi exposto e discutido nesse trabalho – ser entendida como orientação principal e incontestável na definição de unidades de paisagens. Conforme já apresentado, existe a necessidade de diferenciar a significância estatística da significância espacial nos limites apresentados pelos agrupamentos de paisagens.

A maior vantagem verificada no procedimento foi a da otimização para obter um reagrupamento dos limites de unidades para grupos de paisagens. Com isso, o clustering apresentou, com base na distância euclidiana, uma otimização de 91% dos limites entre as 272 unidades de paisagens e seu agrupamento em 25 grupos de paisagens, num tempo bastante reduzido.

Portanto, a relevância dessa técnica é sua colaboração para o procedimento de cluster, ou seja, conseguir agrupar e redefinir uma grande quantidade de informações que, por vias manuais de análise e reagrupamento espacial, apresentariam excessiva morosidade.

Especificamente sobre os algoritmos de cluster, corroborando as premissas de Linden (2009), identificou-se que o k-médias tende a enfatizar a homogeneidade e ignorar a qualidade na separação dos clusters, o que pode justificar, neste trabalho, a fragmentação nos limites espaciais dos grupos de paisagens e, conseqüentemente, uma menor significância espacial.

Outra fragilidade de k-médias é o problema da escolha prévia do número de clusters a ser definida pelo usuário, visto que comumente não se sabe exatamente quantos clusters, a priori, serão ideais. Isso

pode induzir, também, na disposição dos grupos de paisagens resultantes. Linden (2009) novamente apontou que um número reduzido, de poucos clusters, poderá causar a junção de dois clusters “naturais”, ao passo que um número maior previamente escolhido, influenciará numa quebra exagerada de clusters “naturais”.

Sobre os algoritmos hierárquicos de cluster testados através do UPGMA, distância euclidiana e Jaccard, esses apresentam a vantagem de resultados não apenas na diferenciação dos clusters, mas também na identificação da estrutura e do contexto de agrupamento das unidades de paisagens, através do dendrograma. Com isso, é possível optar por diferentes níveis de agrupamento – numa possível relação com a escala desejada, no caso das paisagens – e obter de modo mais compreensível a hierarquia do agrupamento.

Dessa maneira, é possível considerar que a escolha dos intervalos (linha de corte no dendrograma) pode implicar também, de alguma forma, num “pré-estabelecimento” de grupos de paisagens. Embora não seja uma escolha prévia, como no k-médias, ainda assim é uma decisão que irá impactar em obter mais ou menos grupos.

E essa é, por ora, considerada uma fragilidade do UPGMA, no que tange à determinação de qual linha de corte se adotar em diferentes situações e objetivos para a cartografia de paisagens. O que nos direciona a questionar se, de fato, a linha de corte deverá ser adotada avaliando apenas a quantidades de grupos de paisagens a serem representados posteriormente.

Em se tratando dos resultados da validação estatística, os índices têm por objetivo avaliar o número ideal de clusters. Mesmo assim, essa análise não é inteiramente suficiente para apontar o melhor cluster para o agrupamento de paisagens, tendo em vista que o k-médias foi o algoritmo considerado menos satisfatório para o agrupamento e o mais aceitável do ponto de vista do número ideal de clusters. O índice de Gap, por sua vez, corroborou a escolha da distância euclidiana como melhor métrica de similaridade para agrupamento das paisagens.

O índice Silhouette apontou para Jaccard como uma métrica de similaridade satisfatória, confirmando a ideia de que Jaccard seria uma métrica aceitável para adoção do cluster para agrupamento das paisagens, avaliando-se os “ganhos” e “perdas” de seus resultados para a representação das paisagens.

Conseqüentemente, os resultados ampliam o espectro sobre as perspectivas que ainda podem ser exploradas sobre essa temática, por meio de novas análises que visem superar desafios ainda postos, tais como:

O aprimoramento da asserção sobre os níveis de agrupamento (linhas de corte) e sua possível relação com as escalas geográfica e/ou cartográfica, visando maior rigor na representação cartográfica das paisagens num paralelo com a realidade;

A compreensão sobre quais elementos seriam preponderantes no agrupamento das paisagens, ou seja, quais deles condicionariam a formação de alguns grupos de paisagens;

A complexidade de se obter, de forma concomitante, a significância estatística e espacial entre os agrupamentos sugeridos pela análise estatística (clustering);

A premência de realizar futuros testes e comparações com outros algoritmos e métricas não hierárquicas, a exemplo da Análise de Componentes Principais (Principal Component Analysis – PCA), tendo em conta que o k-médias foi o algoritmo menos satisfatório e a única opção testada sobre um método não hierárquico.

NOTAS

1 - Essa mesma abordagem foi utilizada na elaboração do Atlas Nacional da Espanha, na última década (MATA OLMO, 2009)

2 - Reforça-se aqui que o princípio da homogeneidade nas paisagens não significa, conforme explicado por Luca e Santiago (2015), que toda área seja idêntica, mas sim a existência de um padrão comum que pode ser distinguido por atividades de campo e pela cartografia.

3 - Visto que, conforme explica Castro (2010), a escala cartográfica pode ser entendida como a relação

de proporção entre objetos (ou superfícies) e sua representação nos mapas, cartas, desenhos etc.; enquanto a escala geográfica trata do nível de apreensão do fenômeno geográfico, ou seja, da abrangência de sua ocorrência no espaço geográfico.

4 - Na Geografia Russo-Soviética, um modelo consagrado inclui agrupamentos hierárquicos de fácies. Solntcev (1967; 2006), corroborado por Isachenko (1973; 1991), organiza a morfologia das paisagens através de localidades, tratos, subtratos e fácies, e afirma que as paisagens estão contidas em subunidades do envelope geográfico (hierarquia superior), reconhecendo-as como indivíduos geográficos. Rodríguez e Silva (2002), tratando da geoecologia de paisagens, também sugeriram diferentes nomenclaturas para as ordens (escalas) da tipologia de paisagens, como: tipo, classe, grupo e espécie. Além de três ordens subordinadas ou secundárias: subtipo, subclasse e subgrupo.

5 - As cores dos grupos de paisagens só correspondem quando os limites das paisagens são exatamente coincidentes, alheio a isso, as cores foram definidas aleatoriamente.

6 - Os chamados “ganhos” e “perdas” estão relacionados diretamente com os limites dos grupos de paisagens e sua coincidência com os limites de paisagens encontrados em campo, no município de Mineiros (GO).

AGRADECIMENTOS

O primeiro autor agradece à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de estudos Demanda Social em nível de Doutorado, e à Associação Internacional de Lusitanistas (AIL) pela Bolsa para Jovens Investigadores na Universidade de Coimbra (UC). Os demais autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo financiamento dos projetos “Cartografia das paisagens turísticas das savanas brasileiras e moçambicanas” e “Influência do relevo na estruturação das paisagens em diferentes biomas”. Os autores agradecem à Dra. Rosana Veroneze, pesquisadora de pós-doutorado vinculada à Faculdade de Energia Elétrica e de Computação (FEEC) da Unicamp, e ao Dr. Jean Metz, Machine Learning and Software Engineer na empresa JArchitects (Bélgica), pelo auxílio na execução da validação estatística do cluster e nos esclarecimentos sobre os índices de validação.

REFERENCIAS

- ALEKSANDROVA, T. D. Statistical methods of study of natural complexes. Moscow: Nauka, 1975. (Em russo)
- BAYANDINOVA, S.; MAMUTOV, Z.; ISSANOVA, G. Man-made ecology of East Kazakhstan. Singapore: Springer Nature, 2018.
- BERUTCHACHVILI, N.; BERTRAND, G. Le géosystème ou «système territorial naturel». Revue géographique des Pyrénées et du Sud-Ouest, Toulouse, vol. 49, n. 2, p. 167-180, 1978.
- BOLÒS i CAPDEVILA, M. Problemática actual de los estudios de paisaje integrado. Revista de Geografía, Barcelona; v. 15, n. 1-2, p. 45-68, jan./dez., 1981.
- BRAZ, A. M.; OLIVEIRA, I. J.; SALINAS CHÁVEZ, E. Agrupamento estatístico (cluster) para a determinação hierárquica de unidades tipológicas de paisagens. In: ENCONTRO NACIONAL DA ANPEGE, 13., 2019, São Paulo. Anais... São Paulo: ANPEGE; USP, 2019. p. 1-14.
- CASTRO, I. E. O problema da escala. In: CASTRO, I. E.; GOMES, P. C. C.; CORRÊA, R. L. Geografia: conceitos e temas. 2. ed. Rio de Janeiro: Bertrand Brasil, 2000. p. 117-140).
- CAVALCANTI, L. C. S.; CORRÊA, A. C. B. Problemas de hierarquização espacial e funcional na ecologia da paisagem: uma avaliação a partir da abordagem geossistêmica. Geosul, Florianópolis, v. 28, n. 55, p 143-162, jan./jun. 2013.
- CUNHA, J. G. et al. (Coord.). Macrozoneamento Agroecológico e Econômico do Estado de Goiás. Um novo olhar sobre o território Goiano. Produto I: Sistematização de dados existentes em uma base de

dados georreferenciada em ambiente de Sistema de Informações Geográficas (SIG) e suporte a elaboração das macrozonas homogêneas. Goiânia: SIEG, 2014.

GATH, I.; GEVA, A. B. Unsupervised optimal fuzzy clustering. *IEEE Transactions Pattern Analysis and Machine Intelligence*, Toronto, vol.11, n. 7, p. 773-781, jul. 1989.

Instituto Brasileiro de Geografia e Estatística – IBGE. Mapeamento de recursos naturais do Brasil – Escala 1:250.000. Geomorfologia. Documentação Técnica Geral. Rio de Janeiro: IBGE, 2018.

ISACHENKO, A. G. Principles of landscape science and physical geographic regionalization. Melbourne: Melbourne University Press, 1973.

ISACHENKO, A. G. Ciência da paisagem e regionalização físico-geográfica. Moscou: Vyshaya Shkola. 1991. 370p. (Em russo).

ISACHENKO, G. A.; REZNIKOV, A. I. Landscape-dynamical scenarios simulation and mapping in geographic information systems. In: INTERNATIONAL CARTOGRAPHIC CONFERENCE, 17., 1995, Barcelona. Proceedings... Barcelona: International Cartographic Association, 1995. p. 800-804.

HAMMER, Ø. PAST – PAleontological STatistics: reference manual version 3.25. Oslo: Natural History Museum, University of Oslo, 2019.

JACCARD, P. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin de la Société Vaudoise des Sciences Naturelles*, Lausanne, vol. 37, p. 547-579, 1901.

KLIJN, J. A. Hierarchical concepts in landscape ecology and its underlying disciplines. Report 100. Wageningen: DLO WinandStaring Centre, 1995.

KUPRIANOVA, T. P. Principles and methods of physical geographical computerized regionalization. Moscow: Nauka, 1977. (Em russo).

LEGENDRE, P.; LEGENDRE, L. Numerical Ecology. 3. ed. Oxford: Elsevier, 2012.

LINDEN, R. Técnicas de agrupamento. *Revista de Sistemas de Informação da FSMA, Visconde de Araújo*, n. 4, p. 18-36, jul./dez., 2009.

LUCA, V. G.; SANTIAGO, A. G. Avaliação do caráter da paisagem: abordagens europeias. *Paisagem e Ambiente: Ensaios*, São Paulo, n. 36, p. 37-46, 2015.

MACHADO, R. L. Desenvolvimento de um algoritmo imunológico para agrupamento de dados. 2011. 122 f. Monografia (Bacharelado em Ciência da Computação) – Universidade de Caxias do Sul, Caxias do Sul. 2011.

MACQUEEN, J. B. Some Methods for classification and Analysis of Multivariate Observations. In: BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY, 5., 1967, Berkeley. Proceedings... Berkeley: University of California, 1967. p. 281-297.

MARQUES NETO, R. Geomorfologia e geossistemas: influências do relevo na definição de unidades de paisagem no maciço alcalino do Itatiaia (MG/RJ). *Revista Brasileira de Geomorfologia*, São Paulo, vol. 17, n. 4, p. 729-742, out./dez., 2016.

MATA OLMO, R. El Atlas das Paisaxes de España, In. DÍAZ FIERROS, F. y LÓPEZ SILVESTRE, F. (Coord.). *Olladas críticas sobre a paisaxe*. Santiago de Compostela: Consello da Cultura Galega, 2009. p.137-171.

METZ, J. Interpretação de clusters gerados por algoritmos de clustering hierárquico. 2006. 126 f. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional) do Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional – Instituto de Ciências Matemáticas e de Computação. Universidade de São Paulo (USP), São Carlos, 2006.

MEYER, A. S. Comparação de coeficientes de similaridade usados em análise de agrupamento com dados de marcadores moleculares dominantes. 2002. 106 f. Dissertação (Mestrado em Agronomia) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, São Paulo. 2002.

MOREIRA, M. L. O. et al. (Org.). Geologia do Estado de Goiás e Distrito Federal. Escala 1:500.000. Goiânia: CPRM/SIC - FUNMINERAL, 2008.

NUNES, E. D. Modelagem de processos erosivos hídricos lineares no município de Mineiros – GO. 2015. 242 f. Tese (Doutorado em Geografia) do Programa de Pós-Graduação em Geografia – Instituto de Estudos Socioambientais. Universidade Federal de Goiás (UFG), Goiânia.

OLIVEIRA-FILHO, A. T. Um Sistema de classificação fisionômico-ecológica da vegetação Neotropical. In: EISENLOHR, P. V. et al. (Org.). Fitossociologia no Brasil: métodos e estudos de casos. Volume 2. Viçosa: Editora UFV, 2015. p. 452-473.

OLIVEIRA-FILHO, A. T. NeoTropTree, Flora arbórea da Região Neotropical: um banco de dados envolvendo biogeografia, diversidade e conservação. Universidade Federal de Minas Gerais. (<http://www.neotroptree.info>). Acesso em: 13 ago. 2017

PREOBRAZHENSKIY, V. S. Geosystem as an object of landscape study. *GeoJournal*, vol. 7, n. 2, p. 131-134, 1983.

RADAMBRASIL. Folha SE.22 Goiânia: geologia, geomorfologia, pedologia, vegetação e uso potencial da terra. Rio de Janeiro: MME/SG/Projeto Radambrasil, 1983.

RODRÍGUEZ, J. M. M.; SILVA, E. V. A classificação de paisagens a partir de uma visão geossistêmica. *Mercator*, Fortaleza, v. 1, n. 1, p. 95-112, 2002.

RODRÍGUEZ, J. M. M.; SILVA, E. V. Teoria dos Geossistemas - o legado de V.B. Sochava: Volume 1 Fundamentos Teórico-metodológicos. Fortaleza: Edições UFC, 2019.

SALINAS CHÁVEZ, E.; RAMÓN PUEBLA, A. M. Propuesta metodológica para la delimitación semiautomatizada de unidades de paisaje de nível local. *Revista do Departamento de Geografia – USP*, São Paulo, vol. 25, p. 1-19, 2013.

SALINAS, CHÁVEZ, E.; RODRIGUEZ J. M. M.; CAVALCANTI, L. C. S.; BRAZ, A. M. Cartografía de los Paisajes: teoría y aplicación. *Physis Terrae*, Guimarães, vol. 1, n. 1, p. 7-29, 2019.

SERRANO GINÉ, D.; GARCÍA ROMERO, A.; GARCÍA SÁNCHEZ, L. A.; SALINAS CHÁVEZ, E. Un nuevo método de cartografía del paisaje para altas montañas tropicales, *Cuadernos Geograficos*, Granada, vol. 58, n. 1, p. 83-100, 2019. DOI: <http://dx.doi.org/10.30827/cuadgeo.v58i1.6517>

SILVA, L. R. S.; GOMIDE, F. Um estudo comparativo entre as funções de validação para agrupamento nebuloso de dados. In: CONGRESSO BRASILEIRO DE COMPUTAÇÃO (CBCComp), 4., 2004, Itajaí. Anais... Itajaí: Univali, 2004. p. 266-270.

SOCHAVA, V. B. Introdução à teoria dos geossistemas. Novosibirsk: Nauka, 1978a. (Em russo).

SOCHAVA, V. B. Por uma teoria de classificação dos geossistemas de vida terrestre. *Biogeografia*, São Paulo, n. 14, p. 1-24, 1978b.

SOLNTSEV, N. A. Sobre morfologia de paisagens naturais. *Boletim de Geografia (Geographis)*, v. 16, p. 61-86, 1949. (Em Russo).

SOLNTSEV, N. A. What is the difference between facies and biogeocenosis. *Series Geography*. n. 2. 1967. (Em russo). Disponível em: http://www.landscape.edu.ru/book/book_solncev_2001_184.shtml. acesso em 08 nov

2016.

SOLNETSEV, N.A. The natural geographic landscape and some of its general rules. In: WIENS, J. A. et al. *Foundation papers in landscape ecology*. Columbia: Columbia University Press. 2006. p.19-27.

ZACHARIAS, A. A. As categorias de análise da cartografia no mapeamento e síntese da paisagem. *Revista Geografia e Pesquisa*, vol. 2, n. 1, p. 33-56, jan./jun., 2008.

ZHENG, D.; KAINZ, W. Fuzzy rule extraction from GIS data with a neural fuzzy system for decision

making. In: ACM INTERNATIONAL SYMPOSIUM ON ADVANCES IN GEOGRAPHIC INFORMATION SYSTEMS, 7., 1999, Kansas City. Proceedings... Kansas City: ACM, 1999. p. 79-84.