ÁREAS POTENCIAS DE OCORRÊNCIA DE PRODUÇÃO AGROALIMENTAR CAMPONESA NA TRÍPLICE FRONTEIRA ENTRE OS ESTADOS DE MINAS GERAIS, RIO DE JANEIRO E SÃO PAULO

Beatriz Davida da Silva[1], Tathiane Mayumi Anazawa [2], Bruno Vargas Adorno [1], Antônio Miguel Vieira Monteiro [1]

[1]Instituto Nacional de Pesquisas Espaciais - INPE, São José dos Campos, São Paulo, Brasil. {beatriz.silva, miguel.monteiro, bruno.adorno}@inpe.br; [2]Universidade Estadual de Campinas - UNICAMP, Campinas, São Paulo, Brasil. tathiane@unicamp.br

Resumo

A produção agroalimentar camponesa apresenta pouca presença territorializada nas estatísticas agrícolas. Este trabalho desenvolve e testa estratégias metodológicas para identificar áreas potenciais de ocorrência de produção agroalimentar camponesa na região da tríplice fronteira entre os estados de Minas Gerais, Rio de Janeiro e São Paulo, com a utilização de dados secundários (Censo Agropecuário, Setores Censitários, CAR, CNEFE, CNAE/CNPJ, SOS Mata Atlântica e MapBiomas) integrados em uma grade com células de [1x1]km e classificação não supervisionada (k-medoids). De cinco grupos identificados, três (1, 3, 4) apresentaram características dos modos de produção de agentes camponeses. Uma avaliação qualitativa das áreas associadas a esses grupos foi realizada. Os agrupamentos obtidos irão apoiar a seleção de locais para a realização de visitas de campo e missões de drone. Os resultados mostraram o potencial dos dados e da técnica de agrupamento para uma primeira aproximação na produção de cartografias para estas áreas.

Palavras-chave *— modo de produzir camponês, análise de agrupamentos, dados secundários, k-medoids*

Abstract

Peasant agri-food production has limited territorial presence in agricultural statistics. This study develops and tests methodological strategies to identify potential areas of peasant agri-food production in the tri-border region between the states of Minas Gerais, Rio de Janeiro, and São Paulo. It uses secondary data sources (Agricultural Census, Census tracts, CAR, CNEFE, CNAE/CNPJ, SOS Mata Atlântica, and MapBiomas) integrated into a grid with [1x1] km cells and unsupervised classification (k-medoids). Out of five groups that were identified, three had characteristics associated with peasant production. A qualitative assessment of areas related to these groups was conducted. The resulting clusters will support site selection for field visits and drone missions. The findings demonstrated the potential of data and clustering techniques for an initial approach for mapping these areas.

Key words *— peasant mode of production, cluster analysis, secondary data, k-medoids*

1. INTRODUÇÃO

A agricultura camponesa é representada pela maior parcela da população agrícola mundial, com uma contribuição fundamental para a produção de alimentos e a geração de empregos e renda, além de envolver questões relacionadas à segurança alimentar, agroecologia e inclusão social [1]. Entretanto, a produção agroalimentar camponesa apresenta pouca presença territorializada nas estatísticas agrícolas, como é o caso de todo o território brasileiro.

Quando se abordam os mapeamentos extensivos de uso e cobertura da terra realizados com imagens de satélite, as produções camponesas, caracterizadas por pequenas áreas de agricultura, não atingem o tamanho mínimo de mapeamento e/ou possuem usos variados não são contempladas [2,3]. Isso ocorre pois o nível de detalhamento desses mapeamentos depende da resolução espacial e espectral dos sensores a bordo dos satélites. Esta falta de representações cartográficas implica na invisibilidade da produção camponesa diante das pautas relacionadas às políticas públicas. Esse contexto também esconde a diversidade desses agricultores, com suas práticas e dinâmicas variadas [4].

Este trabalho objetiva desenvolver e testar estratégias metodológicas para identificar áreas potenciais de ocorrência de produção agroalimentar camponesa na região da tríplice fronteira entre os estados de Minas Gerais, Rio de Janeiro e São Paulo. A partir da compreensão de características desse modo de produção, são utilizadas informações indicadoras da ocorrência dessas áreas. Para isso, foram utilizados dados secundários, Sistemas de Informações Geográficas (SIGs) e técnica de clusterização não supervisionada (k-medoids). O mosaico de grupos obtido é um dos instrumentos para apoiar a seleção de áreas para a realização de visitas de campo e a definição das missões de drone.

2. PROCEDIMENTOS METODOLÓGICOS

2.1 Área de estudo

A área de estudo abrange 43 municípios da tríplice fronteira entre os estados de Minas Gerais, Rio de Janeiro e São Paulo. A definição desta área resultou da interface entre o Laboratório de Investigação de Sistemas Socioambientais (LiSS) do INPE e a Silo – Arte e Latitude Rural, uma Organização da Sociedade Civil (OSC) da Serrinha do Alambari, município de Resende, Rio de Janeiro. Desde 2020, a Silo tem identificado locais de produção agroalimentar e sistematizado dados sobre esses produtores.

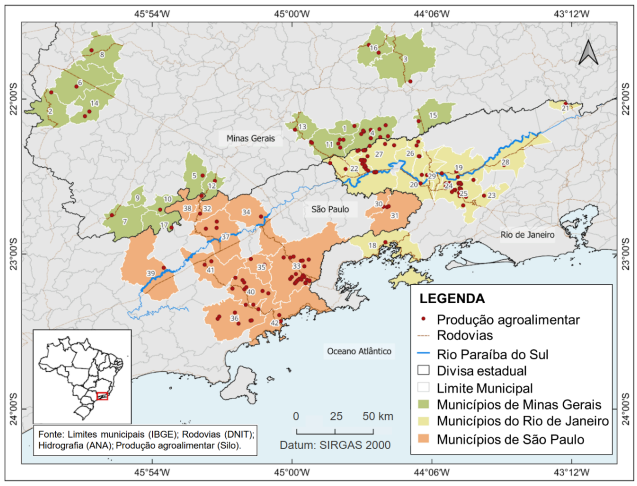


Figura 1. Área de estudo

Dados do Censo Agropecuário de 2017 indicam que a área de estudo possui 23.721 estabelecimentos agropecuários, sendo 16.826 da agricultura familiar [5]. Em Minas Gerais, dos 13.152 estabelecimentos, 80,86% são da agricultura familiar. No Rio de Janeiro, de 4.345 estabelecimentos, 51,69% são de agricultores familiares, enquanto em São Paulo, dos 6.224 estabelecimentos, 63,38% são da agricultura familiar.

2.2 Materiais e métodos

A seleção das variáveis desta pesquisa fundamenta-se em premissas relacionadas ao modo de produção camponês. Dados secundários disponibilizados para todo Brasil podem ser indicadores da presença de produção camponesa.

Os dados do Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE) do Censo Agropecuário de 2017 e do Censo Demográfico de 2022 foram utilizados como indicativo da presença de estabelecimentos agropecuários, que se caracterizam como "unidades de produção ou exploração dedicada, total ou parcialmente, a atividades agropecuárias, florestais e aquícolas" [5]. Para uma melhor caracterização dos estabelecimentos em relação à escala de produção, foi utilizada a classificação por tamanho dos imóveis rurais (minifúndio, pequeno, médio e grande). Assim, os estabelecimentos localizados em imóveis rurais minifúndios e pequenos possuem uma maior possibilidade de estarem relacionados ao modo de produzir camponês.

A partir da Classificação Nacional de Atividades Econômicas (CNAE), foram selecionadas as atividades econômicas agropecuárias primárias: agricultura, pecuária, apicultura, pesca e extrativismo. Em sequência, foi realizada a consulta sistematizada ao banco de dados do Cadastro Nacional de Pessoas Jurídicas (CNPJs) da Receita Federal, a partir da plataforma BigQuery do Google. Essa consulta considerou os seguintes aspectos: a) CNPJs ativos; b) Porte Microempresa; c) CNAES selecionados; d) Municípios da área de estudo. Após a organização e tratamento dos dados, os CNPJs foram georreferenciados a partir da informação do logradouro pela plataforma Google My Maps. Alguns CNPJs não localizados por essa ferramenta foram identificados manualmente. Devido à falta de informações do logradouro, de 111 CNPJs georreferenciados manualmente, 30 foram localizados no local exato, 23 na rua/estrada/avenida, 18 no bairro/distrito e 40 no município.

Os dados dos imóveis rurais do Cadastro Ambiental Rural (CAR) classificados por tamanho, auxiliaram na identificação de potenciais locais de produção camponesa, visto que as atividades são geralmente desenvolvidas em locais essencialmente pequenos [6,7]. Embora não tenha sido confirmada a ocorrência de atividades agrícolas nesses imóveis, o dado pode indicar os seus locais de ocorrência [8].

A Malha dos Setores Censitários de 2021 foi utilizada para reagrupar os setores em rurais, periurbanos e urbanos, a partir da classificação de situação desses setores [9]. Esse reagrupamento pretende analisar o território além da dicotomia urbano-rural [8].

As classes de uso e cobertura da terra utilizados foram: mosaico de usos (áreas de uso agropecuário em que pastagem e agricultura são indistinguíveis) [10], mata ("formações florestais naturais equivalentes às matas primárias e secundárias em estágios médio e avançado de regeneração") e vegetação natural [11]. Essas duas últimas classes foram selecionadas pela possibilidade de identificação de atividades extrativistas.

Após o pré-processamento dos dados, as variáveis geradas foram integradas em uma grade com células regulares de [1x1]km. A mudança do suporte geométrico dos dados de origem procura homogeneizar informações provenientes de diferentes fontes e com distintas geometrias (dados vetoriais, matriciais e também outros espaços celulares), em uma mesma base espaço-temporal [12]. Para o preenchimento das células, foi utilizado o operador estatístico "número total de valores", para agregar os dados do CNEFE de 2017 e 2022 e dos CNPJs. Pelo fato de nem todos os CNPJs terem sido localizados no local exato, foi atribuído um valor de confiança de localização a cada um dos CNPJs considerando sua localização exata, rua/estrada/rodovia, bairro/distrito e município. Já para os dados referentes ao tipo do setor censitário, dos imóveis rurais classificados por tamanho e das classes de uso e

cobertura da terra, foi aplicado o operador estatístico "percentual de cada classe pela área da célula". Essas etapas foram realizadas no *software* TerraView, versão 5.6.5.

Para a análise de agrupamento, foi utilizado o algoritmo *k-medoid Partitioning Around Medoids* (PAM) [13], uma abordagem relacionada ao *k-means*, entretanto mais adequada para se trabalhar com conjunto de dados que apresenta *outliers* e ruídos, como observado nas variáveis selecionadas. Neste método, para caracterizar os grupos que serão formados, busca-se um objeto no conjunto de dados, chamado de medoide, para o qual a dissimilaridade média entre ele e todos os outros membros do *cluster* é mínima, correspondendo ao ponto mais centralmente localizado no *cluster* [14]. Optou-se por utilizar a distância de Manhattan, por ser mais adequada quando comparada à distância euclidiana, ao utilizar dados em uma grade regular. Para determinar o número k de grupos, foram realizados testes com k=[3 a 6]. A avaliação de k foi feita com o uso da análise do gráfico de Silhueta [15]. A Figura 3 mostra o índice de Silhueta que levou a escolha de um k=5.

3. RESULTADOS E DISCUSSÃO

O resultado da análise de agrupamento pelo k-medoids encontra-se na Figura 2, com o mapa da distribuição espacial dos grupos formados. A descrição das principais características de cada grupo encontra-se a seguir:

Grupo 1 (3.714 células): Os setores censitários predominantes são os urbanos e periurbanos. Do total de 318 CNPJs, 59% estão localizados neste grupo. A classe mosaico de usos está presente em 82% das células.

Grupo 2 (3.194 células): Todas as células apresentam setores censitários rurais. Os estabelecimentos agropecuários médios e grandes são 17% do total (Censo 2017) e 19% do total (Censo 2022). Os imóveis rurais grandes estão presentes em 91% delas.

Grupo 3 (2.821 células): Todas as células apresentam setores censitários rurais e a classe mata. A quantidade de CNPJs é de apenas 6 e de estabelecimentos agropecuários minifúndios e pequenos é de 0,6% (Censo 2017) e de 0,5% (Censo 2022).

Grupo 4 (8.576 células): Todas as células apresentam setores censitários rurais. De 16.421 estabelecimentos agropecuários minifúndios e pequenos do Censo de 2022, 92% se encontram neste grupo. Em relação ao Censo de 2017, de um total de 16.959 estabelecimentos agropecuários minifúndios e pequenos, 89% se localizam neste grupo. De 318 CNPJs, 31% (99) se encontram neste grupo. Todas as células apresentam imóveis pequenos.

Grupo 5 (5.080 células): Todas as células apresentam setores censitários rurais. Os estabelecimentos agropecuários médios e grandes são os que predominam nas células, em relação ao Censo de 2017 são 62,78% (2233), e do Censo de 2022, são 61,82% (2402). Os imóveis rurais que predominam são os médios, em 91% das células.

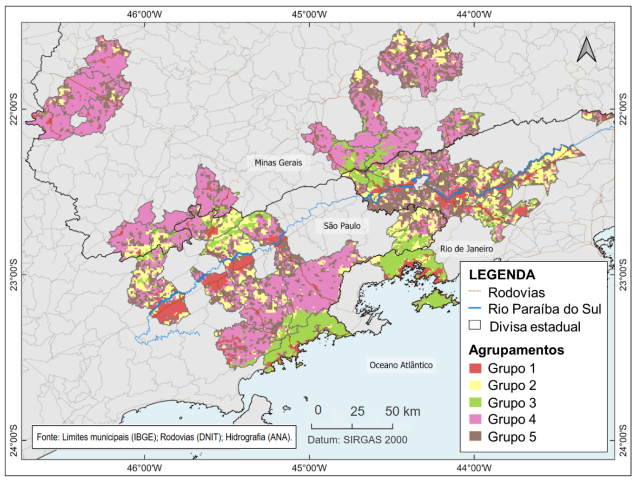


Figura 2. Grupos formados a partir da classificação não supervisionada - Método K-Medoids (k =5).

A Figura 3 apresenta a avaliação dos grupos por meio da análise de Silhueta, em que os valores próximos a 1 indicam que o objeto está bem agrupado. Verificou-se que a avaliação geral dos agrupamentos foi de 0,28, indicando sobreposição em boa parte dos grupos formados. Ao analisar os grupos isoladamente, percebe-se que o Grupo 3 apresentou o valor mais elevado, de 0,49, seguido do Grupo 1 com 0,35, Grupo 4 com 0,30, Grupo 5 com 0,18, Grupo 2 com 0,14. Esses resultados podem ser explicados, em parte, pela similaridade entre o perfil de variáveis que formam os grupos, muitos deles divergindo em poucas variáveis.

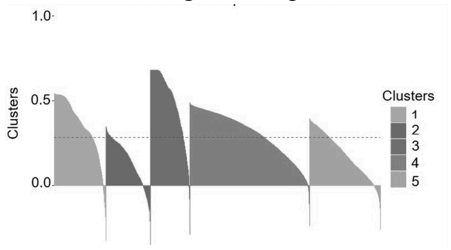


Figura 3. Gráfico de silhueta para os clusters

A partir da análise dos grupos formados, verificou-se que o grupo 4 apresenta as características associadas à produção camponesa e, portanto, representa áreas potenciais de ocorrência de produção agroalimentar. Entretanto, os grupos 1 e 3, também apresentaram atributos importantes, associados à lógica da produção camponesa, como uma quantidade de CNPJs adequados a esta produção e a porcentagem das classes de uso e cobertura, sendo áreas importantes para serem amostradas e melhor investigadas em visitas em campo e para definição dos voos de drone. Para avaliar qualitativamente os grupos que possuem uma maior potencialidade relacionada ao modo de produzir camponês, foram selecionadas três áreas (A, B, C). Na Tabela 1 é apresentada a descrição das características das variáveis encontradas em cada uma dessas áreas.

ÁREA A (Grupo 1) - está localizada no município de Andradas/MG em setor censitário periurbano, com a presença de imóveis rurais minifúndios e pequenos. Apresenta estabelecimentos agropecuários relacionados a esses imóveis. Apresenta a classe mosaico de usos.  Fonte: Google Earth, 2024.
ÁREA B (Grupo 3) - está localizada no município de Piraí/RJ, no assentamento Roseli Nunes, em setor censitário rural, sem presença de imóveis rurais e estabelecimentos agropecuários. Apresenta fragmentos de mata e vegetação natural.  Fonte: Google Earth, 2024.
ÁREA C (Grupo 4) - localiza-se no município de Cunha/SP em setor censitário rural, com a predominância da classe mosaico de usos. Apresenta imóveis rurais pequenos, com a existência de estabelecimentos agropecuários relacionados a esses imóveis.  Fonte: Google Earth, 2024.

Tabela 1. Descrição de áreas selecionadas pertencentes ao grupo 1, 3 e 4

4. CONSIDERAÇÕES FINAIS

A aplicação do método de classificação não supervisionada para a categorização das células, contribuiu para uma análise exploratória das características dos grupos formados, que apoiarão a seleção de locais para a realização de visitas de campo e missões com drones. Os resultados mostraram o potencial dos dados e da técnica de agrupamento para uma primeira aproximação na construção de novas cartografias para estas áreas de produção.

5. AGRADECIMENTOS

À FAPESP, pelo financiamento da pesquisa pelo processo nº 2023/18219-0.

6. REFERÊNCIAS

[1] J. D. van der Ploeg. Sete teses sobre a agricultura camponesa. In: P. Peterson (Org.). *Agricultura familiar camponesa na construção do futuro*. AS-PTA, Rio de Janeiro, 2009. p. 17-31.

[2] A. R. Souza, M. I. S. Escada, R. F. B. Marujo, A. M. V. Monteiro. Cartografia do invisível: revelando a agricultura de pequena escala com imagens Rapideye na região do baixo Tocantins. *Revista do Departamento de Geografia*, v. 38, pp. 137-153, 2019.

[3] F. D. Pacheco, M. R. Matias, G. M. Silva, A. R. Souza, Y. E. Shimabukuro, M. I. S. Escada. Image classification methods assessment for identification of small-scale agriculture in brazilian amazon. *In The Thirteenth International Conference on Advanced Geographic Information Systems, Applications, and Services*, pp. 12–19, 2021.

[4] H. M. Carvalho. Prefácio. In: F. A. Costa. *Economia camponesa nas fronteiras do capitalismo*: teoria e prática nos EUA e na Amazônia Brasileira. NAEA, Belém, 2012.

[5] Instituto Brasileiro de Geografia e Estatística (IBGE). Censo Agropecuário 2017, 2017.

[6] J. D. van der Ploeg. *Camponeses e a arte da agricultura: um manifesto Chayanoviano*. Editora Unesp, São Paulo/Porto Alegre, 2016.

[7] M. N. B. Wanderley. Raízes históricas do campesinato brasileiro. *Agricultura familiar*: realidades e perspectivas, v. 3, pp. 21-55, 1996.

[8] M. R. Matias. Cartografias da agricultura urbana: contribuições ao planejamento territorial na região metropolitana do Vale do Paraíba e Litoral Norte. 2022. 115 p. (INPE - 08.22.15.45-TDI). Dissertação (Mestrado em Sensoriamento Remoto) - INPE, São José dos Campos, 2022.

[9] Instituto Brasileiro de Geografia e Estatística. Malha dos setores censitários - Acesso ao produto, 2021.

[10] Mapbiomas. Descrição da legenda. 2024.

[11] SOS Mata Atlântica. *Atlas dos Remanescentes Florestais da Mata Atlântica*: Período 2021-2022. SOS Mata Atlântica/ Instituto Nacional de Pesquisas Espaciais, São Paulo, 2023.

[12] S. Amaral, A. Gavlak, M. I. S. Escada, A. M. V. Monteiro. Using remote sensing and census tract data to improve representation of population spatial distribution: case studies in the Brazilian Amazon. *Population and Environment*, v. 34, n. 1, pp. 142–170, 2012.

[13] L. Kaufman, P. J. Rousseew. *Finding groups in data: an introduction to cluster analysis*. Wiley, New York, 1990.

[14] A. Kassambra. *Practical guide to cluster analysis in R: unsupervised machine learning*. Sthda, 2017.

[15] P. J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, v. 20, pp. 53-65, 1987.