ANÁLISE ESPACIAL DA PRODUTIVIDADE DA SOJA VIA MAPEAMENTO COROPLÉTICO PELO MÉTODO DO DESVIO PADRÃO

Lucas Ramos Marchetti[1], Gustavo Henrique Dalposso[2], Jones Erni Schmitz[2], Marcio Paulo de Oliveira[3]

[1]Estudante do Ensino Médio, Instituto Federal do Paraná, Toledo-PR, Brasil (lucasramosmarchetti2@gmail.com)
[2]PPGBIO – Programa de Pós-Graduação em Tecnologias em Biociências, UTFPR, Toledo-PR, Brasil
[3]COMAT – Coordenação do Curso de Licenciatura em Matemática, UTFPR, Toledo-PR, Brasil

Resumo: Este estudo analisa a produtividade da soja na região Geográfica Intermediária de Cascavel (PR) na safra 2023/2024. Utilizaram-se estatística descritiva e a técnica de mapeamento coroplético fundamentada no método de classificação por desvio padrão. Os resultados evidenciam a eficácia do método em identificar *clusters* produtivos e contrastes regionais, conferindo rigor estatístico à visualização de dados agrícolas, que podem subsidiar estratégias de manejo e políticas públicas.

Palavras-chave: Mapa coroplético; Produtividade agrícola; Análise espacial; Desvio padrão

INTRODUÇÃO

A soja destaca-se como uma das culturas agrícolas mais importantes do mundo (Mikhailova et al., 2025), exercendo papel estratégico tanto na alimentação humana e animal quanto na economia global, devido ao seu elevado teor de proteína e óleo e à ampla diversidade de aplicações industriais e alimentícias (Uskutoğlu, 2025). Nas últimas décadas, a produção mundial de soja apresentou crescimento expressivo, impulsionado pelo aumento da demanda por óleo vegetal, farelo proteico e biocombustíveis, sendo o Brasil, os Estados Unidos e a Argentina responsáveis por grande parte da produção global (Addae-Frimpomaah et al., 2026). Nesse contexto, a produtividade agrícola da soja apresenta significativa variação espacial, influenciada por fatores ambientais, tecnológicos, climáticos e de manejo, o que reforça a importância de estudos voltados à compreensão da distribuição espacial da produtividade e ao aprimoramento das práticas agrícolas (Hamza et al., 2024).

Os mapas coropléticos desempenham papel fundamental em estudos geográficos e agrícolas, pois permitem visualizar a distribuição espacial dos fenômenos e identificar, de forma clara e objetiva, padrões regionais. Esse tipo de mapa representa magnitudes estatísticas por meio de cores ou tonalidades aplicadas a unidades espaciais, facilitando a interpretação de concentrações e diferenças entre regiões (Torguson, 2017). Além disso, a variação de cores e tonalidades influencia diretamente a percepção visual dos dados, contribuindo para a identificação de semelhanças e contrastes espaciais entre as áreas analisadas (Lloyd e Steinke, 1977). Apesar de sua ampla utilização e importância na análise espacial, os mapas coropléticos exigem cuidados metodológicos na classificação e na interpretação das informações representadas (Moser, 2025). Nesse contexto, avanços recentes envolvendo inteligência artificial e sistemas automatizados de cartografia demonstram o crescente potencial das tecnologias digitais na produção e análise de mapas coropléticos (Pannoon e Netek, 2025).

Para que a distribuição espacial da produtividade seja interpretada de forma clara, é necessário organizar os dados em classes, reduzindo a complexidade da informação e facilitando a leitura cartográfica (Zhang e Maciejewski, 2017). O método do desvio padrão fundamenta essa classificação em medidas de tendência central e dispersão, o que permite uma análise baseada na variabilidade dos dados (Huang et al., 2020). Na prática, a técnica define os limites das categorias com base em intervalos fixos acima e abaixo da média aritmética, resultando em uma representação estatisticamente equilibrada (Milić et al., 2019). Essa característica torna o método especialmente eficaz para evidenciar áreas de

produtividade muito alta ou muito baixa, destacando visualmente os extremos e os contrastes em relação ao comportamento médio da região (Cantarino et al., 2019).

Dessa forma, o presente estudo tem como objetivo elaborar, apresentar e discutir um mapa coroplético da produtividade da soja na região geográfica intermediária de Cascavel (PR), utilizando o método do desvio padrão, buscando identificar padrões espaciais e diferenças regionais na produtividade agrícola.

MATERIAL E MÉTODOS

A área de estudo compreende a Região Geográfica Intermediária de Cascavel, localizada no oeste do estado do Paraná, Brasil (Figura 1). Definida pelo Instituto Brasileiro de Geografia e Estatística (IBGE) em 2017, a região é composta por 100 municípios. Climaticamente, a região enquadra-se no tipo subtropical úmido (Cfa, Köppen), com temperaturas médias anuais entre 18 °C e 22 °C e precipitação anual entre 1,600 mm e 2,000 mm, distribuída ao longo do ano (Alvarez et al., 2013).

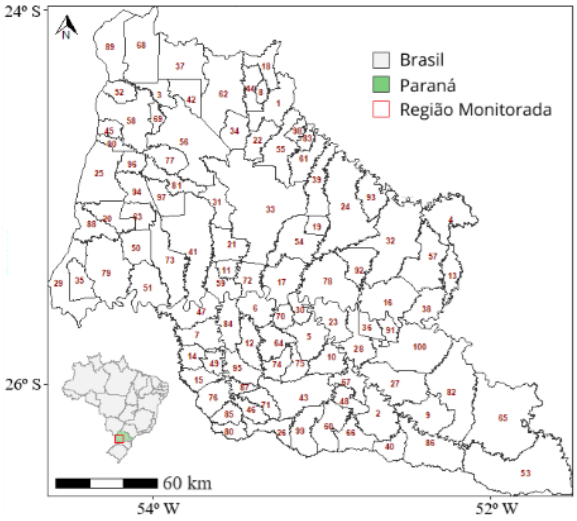

Figura 1. Região geográfica intermediária de Cascavel.

Os dados oficiais da produtividade da soja (t ha^{-1}) na safra 2023/2024 foram obtidos junto ao Departamento de Economia Rural (DERAL) da Secretaria de Estado da Agricultura e do Abastecimento do Paraná, Brasil (Seab, 2025).

Análise exploratória

A fase inicial do estudo consistiu em uma análise exploratória fundamentada na aplicação de estatísticas descritivas para sintetizar o volume de dados brutos e estruturar a interpretação da produtividade da soja na região monitorada. Para identificar o centro da distribuição, foram calculadas as medidas de tendência central: a média e a mediana. A compreensão da heterogeneidade e do espalhamento dos dados foi obtida por meio de medidas de variabilidade, incluindo o mínimo, o máximo, o desvio padrão, o coeficiente de variação (CV%) e os limites interquartílicos (primeiro e terceiro quartis).

Complementarmente, avaliou-se a assimetria para caracterizar a forma do conjunto de dados, permitindo identificar tendências de concentração de valores e desvios em relação à normalidade. Esta abordagem metodológica foi essencial para transformar observações isoladas em informações estruturadas e fundamentadas sobre o comportamento das variáveis analisadas (Bussab; Morettin, 2017).

Posteriormente, elaborou-se um *boxplot*, gráfico de análise exploratória utilizado para caracterizar a distribuição e a variabilidade dos dados de produtividade. O *boxplot* descreve a distribuição de uma variável numérica por meio do resumo de cinco números: valor mínimo, primeiro quartil (Q_1), mediana, terceiro quartil (Q_3) e valor máximo (Martinez et al., 2017). A estrutura do gráfico é formada por uma caixa delimitada entre Q_1 e Q_3, com a mediana em seu interior, e por extensões denominadas *whiskers*, que indicam a dispersão dos dados além dos limites da caixa. Sua construção parte da ordenação dos dados e do cálculo dos quartis, sendo o intervalo interquartil (IQR) obtido pela diferença $IQR = Q_3 - Q_1$, constituindo uma medida robusta de dispersão por não ser influenciada por valores extremos. Os *whiskers* se estendem a partir das bordas da caixa até o valor observado mais distante dentro dos limites definidos por $Q_1 - 1{,}5 \times IQR$ e $Q_3 + 1{,}5 \times IQR$. Observações que ultrapassem esses limiares são classificadas como *outliers* e representadas individualmente no gráfico. Do ponto de vista estatístico, o *boxplot* permite identificar a posição central, a variabilidade e a assimetria da distribuição, uma vez que a posição relativa da mediana na caixa e o comprimento desigual dos *whiskers* fornecem indicações visuais sobre o comportamento da distribuição.

Segundo Martinez et al. (2017), essa representação fornece medidas robustas de localização e dispersão, sendo especialmente útil na comparação entre diferentes conjuntos de dados. Entretanto, ressalta-se que o *boxplot* sintetiza a distribuição, sem evidenciar detalhes como lacunas ou agrupamentos específicos. Dessa forma, sua utilização é recomendada como ferramenta complementar em análises estatísticas exploratórias.

Método de classificação por desvio padrão

Para a espacialização da produtividade, utilizou-se o método de classificação por desvio padrão, que fundamenta a delimitação das classes na dispersão dos dados em torno da média aritmética. Nesse método, os limites das classes são estabelecidos a partir da média ($\bar{x}$), utilizando múltiplos do desvio padrão (s) como intervalos de corte. Inicialmente define-se o intervalo de classe central $[\bar{x}, \bar{x} + s)$, e todos os municípios que apresentam valores de produtividade neste intervalo devem ser representados com uma mesma cor no mapa coroplético. Posteriormente, são criadas duas novas classes, uma abaixo $[\bar{x} - s, \bar{x})$ e uma acima $[\bar{x} + s, \bar{x} + 2s)$, sendo representadas por duas novas cores no mapa coroplético. Este procedimento é repetido até que todos os municípios fiquem alocados dentro das classes.

Esta técnica permite que o mapa coroplético evidencie, de forma estatisticamente rigorosa, o grau de afastamento da produtividade de cada unidade espacial em relação ao comportamento médio da região, destacando visualmente as áreas de excepcionalidade produtiva (extremos positivos e negativos).

Recursos computacionais

Foram utilizados os softwares R (R CORE TEAM, 2025) e QGIS (QGIS DEVELOPMENT TEAM, 2009) no desenvolvimento deste trabalho. O R e o QGIS são recursos computacionais amplamente utilizados em análises estatísticas e modelagem espacial devido à sua capacidade de processar, visualizar e analisar dados geográficos. O QGIS destaca-se pelas funcionalidades de geoprocessamento, elaboração de mapas temáticos e manipulação de dados espaciais em um ambiente gráfico intuitivo. O R foi empregado nos cálculos estatísticos e na definição dos limites de classe pelo método do desvio padrão, enquanto o QGIS foi utilizado na elaboração do mapa coroplético. A integração dessas ferramentas possibilitou maior eficiência no processamento, na análise e na visualização dos dados.

RESULTADOS E DISCUSSÃO

A Tabela 1 evidencia a variabilidade espacial da produtividade da soja na Região Geográfica Intermediária de Cascavel durante a safra 2023/2024, com valores variando entre 2,000 t ha^{-1} e 4,500 t ha^{-1}. A produtividade média observada foi de 3,478 t ha^{-1}, valor ligeiramente inferior à média nacional da safra 2023/2024, estimada em 3,507 t ha^{-1}, e também abaixo da produtividade média do estado do Paraná, que atingiu 3,860 t ha^{-1} (Conab, 2024). A proximidade entre média (3,478 t ha^{-1}) e mediana (3,461 t ha^{-1}), associada ao baixo valor de assimetria (-0,160), indica uma distribuição relativamente simétrica dos dados. Além disso, o coeficiente de variação de 13,18% sugere variabilidade moderada da produtividade entre os municípios analisados.

Tabela 1. Estatísticas descritivas da produtividade da soja (t ha^{-1}).

Mínimo	2,000
Máximo	4,500
Primeiro quartil	3,200
Média	3,478
Mediana	3,461
Terceiro quartil	3,811
Desvio padrão	0,458
Assimetria	-0,160
Coeficiente de variação (CV%)	13,18

Os quartis mostram que 50% das observações concentraram-se entre 3,200 t ha^{-1} e 3,811 t ha^{-1}, evidenciando predominância de produtividades intermediárias na região. Esses resultados reforçam a existência de diferenças espaciais na produtividade da soja, possivelmente associadas a fatores climáticos, ambientais e de manejo agrícola, conforme observado em escala nacional na safra 2023/2024 (Conab, 2024).

A Figura 2, que apresenta o gráfico *boxplot* da produtividade da soja na safra 2023/2024 na região monitorada, evidencia a presença de uma observação atípica inferior (*outlier*). Por meio do *boxplot* é possível observar também que os valores variam entre 2,0 e 4,5 (t ha^{-1}), sendo que os maiores valores de produtividade foram observados nos municípios que possuem uma vocação para a cultura da soja, contribuindo de forma relevante na produção da soja no estado do Paraná, e no consequente desenvolvimento econômico. A elevada produtividade registrada nesses municípios indica a existência de uma infraestrutura consolidada de suporte à produção agrícola, incluindo silos, armazéns e cooperativas, que contribuem para a eficiência das operações de armazenamento, comercialização e escoamento da produção.

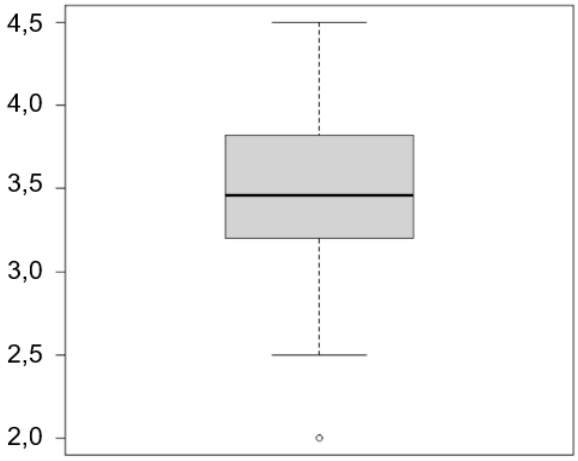

Figura 2. *Boxplot* da produtividade da soja (t ha^{-1}).

O *outlier* corresponde ao município de Terra Roxa, identificado pelo código 68 na Figura 1, localizado na porção noroeste da Região Geográfica Intermediária de Cascavel. O município apresentou produtividade consideravelmente inferior em relação aos demais e a identificação desse comportamento discrepante é relevante, pois demonstra que existem diferenças espaciais significativas na produtividade da soja na região, possivelmente influenciadas por condições climáticas, características do solo ou diferenças nas práticas agrícolas utilizadas pelos produtores. A Figura 3 apresenta o mapa coroplético da produtividade da soja na Região Geográfica Intermediária de Cascavel, cujas classes foram elaboradas pelo método de classificação pelo desvio padrão.

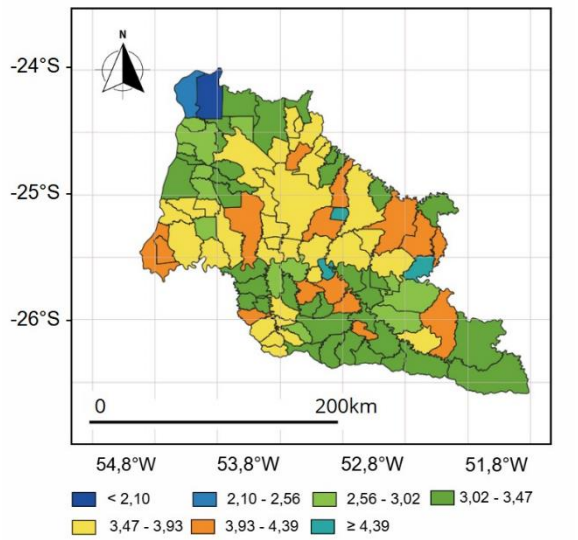

Figura 3. Mapa coroplético da produtividade da soja (t ha^{-1}) na Região Geográfica Intermediária de Cascavel (safra 2023/2024) elaborado pelo método de classificação por desvio padrão.

A análise do mapa coroplético revela expressiva variabilidade espacial da produtividade da soja na Região Geográfica Intermediária de Cascavel durante a safra 2023/2024. A porção central da região, predominantemente classificada nos intervalos de 3,02 a 3,47 t ha^{-1} e de 3,47 a 3,93 t ha^{-1}, reflete condições de solo e clima moderadamente favoráveis, associadas à fertilidade natural dos Latossolos Vermelhos do oeste paranaense e ao regime pluviométrico subtropical úmido (Alvarez et al., 2013), além de tecnificação agrícola relativamente difundida entre os produtores.

Em contraste, o município de Terra Roxa, representado pela classe inferior a 2,10 t ha^{-1} e identificado como *outlier* no *boxplot*, situa-se mais de três desvios padrão abaixo da média regional, sugerindo a convergência de fatores desfavoráveis como maior vulnerabilidade climática ou possíveis déficits hídricos em períodos fenologicamente críticos. Além disso, segundo Pinto et al. (2025), o solo de Terra Roxa é predominantemente um argissolo vermelho-amarelo distrófico com textura arenosa na superfície, o que proporciona menor fertilidade em comparação com os latossolos vermelhos predominantes na maior parte da região.

Os municípios classificados nas categorias superiores, de 3,93 a 4,39 t ha^{-1} e acima de 4,39 t ha^{-1}, tendem a reunir condições de solo e clima mais favoráveis, maior nível tecnológico e melhor estabilidade hídrica ao longo do ciclo da cultura, fatores que elevam seu desempenho produtivo acima da média regional.

Do ponto de vista metodológico, o método do desvio padrão mostrou-se eficaz para evidenciar os contrastes espaciais da produtividade, e conforme indicam Salim et al. (2023), o método do desvio padrão é capaz de criar subgrupos coesos ao fundamentar a classificação na dispersão dos dados em relação à média aritmética. Essa característica torna o método eficaz para evidenciar contrastes espaciais, conferindo rigor estatístico à representação cartográfica da produtividade.

CONCLUSÃO

Diante do exposto, o método de classificação por desvio padrão revelou-se uma alternativa eficaz para a análise espacial da produtividade da soja. Ao fundamentar os limites de classe na dispersão dos dados em torno da média aritmética, essa técnica permite evidenciar contrastes espaciais significativos, conferindo maior rigor estatístico à representação cartográfica. A aplicação deste método permitiu observar claramente a formação de *clusters* regionais de produtividade, evidenciando áreas de maior desempenho que se distanciam positivamente da

média regional, bem como zonas que apresentam resiliência ou vulnerabilidade produtiva específica. Dessa forma, o método demonstra ser uma ferramenta valiosa para a interpretação de fenômenos geográficos, facilitando a identificação de padrões e a compreensão da heterogeneidade espacial a partir da distribuição estatística dos valores observados.

AGRADECIMENTOS

Os autores agradecem à Universidade Tecnológica Federal do Paraná (UTFPR), Campus Toledo, pelo suporte acadêmico e tecnológico oferecido para o desenvolvimento deste projeto. Agradecem, também, ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de pesquisa.

REFERÊNCIAS

ADDAE-FRIMPOMAAH, F. *et al*. Distribution, Cultivation, and Economic Importance of Soybean. *In*: RAJENDRAN, R. A. *et al*. (ed.). **Soybeans**: **history, breeding, and product profile**. 1. ed. New York: Apple Academic Press, 2026. v. 1, cap. 4.

ALVAREZ, C.A.; STAPE, J.L.; SENTELHAS, P.C.; DE MORAES GONÇALVES, J.L.; SPAROVEK, G. G. Köppen's climate classification map for Brazil, **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711-728, 2013.

BUSSAB, W.; MORETTIN, P. A. **Estatística básica**. 9. ed. São Paulo: Saraiva, 2017.

CANTARINO, I. *et al*. A ROC analysis-based classification method for landslide susceptibility maps. **Landslides**, v. 16, p. 265-282, 2019.

CONAB – COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira de grãos: safra 2023/24 – 6º levantamento**. Brasília: CONAB, 2024. Disponível em: https://www.conab.gov.br

HAMZA, M. *et al*. Global Impact of Soybean Production: A Review. **Asian Journal of Biochemistry, Genetics and Molecular Biology**, [s. l.], v. 16, n. 2, p. 12-20, 2024.

HUANG, Z.; LU, Y.; MACK, E.A.; CHEN, W.; MACIEJEWSKI, R. Exploring the Sensitivity of Choropleths under Attribute Uncertainty. **IEEE Transactions on Visualization and Computer Graphics**, v. 26, n. 8, p. 2576-2590, 2020.

LLOYD, R.; STEINKE, T. Visual and statistical comparison of choropleth maps. **Annals of the Association of American Geographers**, [s. l.], v. 67, n. 3, p. 429-436, 1976.

MIKHAILOVA, M. P.; GRETCHENKO, A. E.; BLINOVA, A. A. Reaction of short-season soybean variety to pre-sowing UV treatment of seeds. **BIO Web of Conferences**, Les Ulis, v. 179, art. 08004, 2025.

MILIĆ, N. *et al*. The Influence of Data Classification Methods on Predictive Accuracy of Kernel Density Estimation Hotspot Maps**. The International Arab Journal of Information Technology**, v. 16, n. 6, p. 1053-1061, 2019.

MARTINEZ, W. L.; MARTINEZ, A. R.; SOLKA, JEFFREY. **Exploratory data analysis with MATLAB**. Chapman and Hall/CRC, 2017.

MOSER, J. Choropleth Maps: Critical Reflections on the Frequently Used Representation Method in an Educational Context. *KN J.* **Cartography and Geographic Information Science,** *v.*75, p. 191–204, 2025.

PANNON, P.; NETEK, R. Creating choropleth maps by artificial intelligence—case study on ChatGPT-4. **ISPRS International Journal of Geo-Information**, Basel, v. 14, n. 12, p. 486, 2025.

PINTO, L.A.S.R. *et al*. Management systems in the subtropical region of Brazil: aggregation in the assessment of the chemical quality of fragile soils. **Experimental Agriculture**, v. 61, p. 1-20, 2025.

QGIS DEVELOPMENT TEAM. **QGIS Geographic Information System**: Open Source Geospatial Foundation, 2009. Disponível em: http://qgis.org.

R CORE TEAM. R: **A Language and Environment for Statistical Computing**. Vienna, Austria: R Foundation for Statistical Computing, 2025.

SALIM, A. B.; RAHMAN, M. B.; DAS, A.; ROY, S. Defining the Regions Based on Remittance on Household Welfare in Bangladesh. In: **International Conference on Planning, Architecture and Civil Engineering**, 2023, Rajshahi. **Proceedings** [...]. Rajshahi: Rajshahi University of Engineering & Technology, 2023. p. 1-6

SEAB/DERAL (2025) Secretaria de Estado da Agricultura e do Abastecimento. Departamento de Economia Rural. **Diagnóstico Agropecuário Paranaense: Safra 2023/2024**. Available from: https://www.agricultura.pr.gov.br

TORGUSON, J.S. (2017). Choropleth Map. In **International Encyclopedia of Geography: People, the Earth, Environment and Technology** (eds D. Richardson, N. Castree, M.F. Goodchild, A. Kobayashi, W. Liu and R.A. Marston).

USKUTOĞLU, D. The Multifunctional Role of Soybean (*Glycine max* L.): Its Nutritional Value, Industrial Potential, and Future Projections. **Bozok Tarım ve Doğa Bilimleri Dergisi**, [s. l.], v. 4, n. 2, p. 127-135, 2025.

ZHANG, Y.; MACIEJEWSKI, R. Quantifying the Visual Impact of Classification Boundaries in Choropleth Maps. **IEEE Transactions on Visualization and Computer Graphics**, v. 23, n. 1, p. 371-380, 2017.