

Capítulo 12

Sistemas de Informação Geográfica (SIG) e sua Aplicações na Análise de Políticas Públicas

Ricardo Ceneviva

Abril de 2025

Introdução

Os Sistemas de Informação Geográfica (SIG) são ferramentas computacionais que permitem coletar, armazenar, analisar e visualizar dados com referência geográfica (Brasil Escola, s.d.). Em termos simples, um SIG combina software, hardware, dados espaciais e procedimentos analíticos para representar fenômenos na superfície terrestre de forma georreferenciada. Esses sistemas possibilitam mapear elementos do território e espacializar fenômenos sociais, facilitando a identificação de relações, padrões e tendências no espaço (Esri Portugal, s.d.). Por exemplo, pesquisadores podem usar SIG para mapear a distribuição de serviços públicos em uma cidade ou visualizar índices socioeconômicos por bairro, contribuindo para o planejamento urbano e a formulação de políticas.

Um conceito-chave em SIG é o de camadas temáticas. Os dados geográficos são organizados em camadas sobrepostas, cada uma contendo um tipo de informação (por exemplo, uma camada de ruas, outra de bairros e outra de dados populacionais) (Wikipédia, s.d.-a; Wikipédia, s.d.-b). Essa estrutura em camadas permite comparar e combinar informações distintas referentes a um mesmo local, através de procedimentos de sobreposição de mapas (map overlay). Assim, um pesquisador pode, por exemplo, sobrepor um mapa de índices de criminalidade com outro de distribuição de renda para examinar visualmente possíveis coincidências espaciais entre pobreza e violência. Para que outro pesquisador possa reproduzir um mapa feito em SIG, é imprescindível explicitar certas decisões cartográficas: quais camadas de dados foram utilizadas e suas fontes, qual o sistema de projeção e coordenadas empregado, qual método de classificação dos dados foi aplicado na simbologia do mapa e qual a escala ou resolução espacial dos dados. Documentar essas escolhas assegura que os resultados sejam replicáveis e compreendidos em seu contexto técnico.

Modelos de dados geográficos: Em SIG, trabalham-se principalmente com dois modelos de representação de dados espaciais: vetorial e raster (matricial) (Brasil

Escola, s.d.). No modelo vetorial, as entidades geográficas do mundo real são representadas por pontos, linhas ou polígonos, focando na precisão de sua localização exata no espaço (Wikipédia, s.d.-a; Wikipédia, s.d.-b). Por exemplo, a localização de escolas pode ser representada por pontos com coordenadas, estradas por linhas e os limites de bairros por polígonos. Já no modelo raster, o espaço é dividido em uma grade regular de células (pixels), onde cada célula armazena um valor representando uma característica naquele local (como altitude, temperatura ou densidade populacional) (Wikipédia, s.d.-c). Quanto maior o tamanho da célula (menor resolução), menos detalhes espaciais são capturados (Wikipédia, s.d.-d). Dados raster são apropriados para fenômenos contínuos (e.g. uso do solo, imagens de satélite), enquanto dados vetoriais são indicados para objetos discretos (e.g. localização de equipamentos urbanos). Saber distinguir e escolher entre dados vetoriais e raster é fundamental: na pesquisa sociológica, dados vetoriais costumam modelar unidades como municípios, setores censitários ou regiões administrativas, enquanto dados raster podem representar superfícies como densidade populacional estimada ou índices ambientais que variam continuamente no território.

Projeções cartográficas e sistemas de referência: Todo dado geográfico utiliza um Sistema de Referência de Coordenadas (SRC ou CRS), que define como o globo terrestre (tridimensional) é projetado numa superfície plana (mapa) e como as coordenadas numéricas correspondem a locais no mundo real. A escolha da projeção é uma decisão cartográfica crucial, pois diferentes projeções introduzem diferentes distorções de área, forma, distância ou ângulo. Por exemplo, a projeção de Mercator preserva ângulos (conformidade) mas distorce áreas (regiões próximas aos polos aparecem exageradamente grandes), enquanto projeções equivalentes (como a de Albers equal-area) preservam proporções de área ao custo de distorcer formas. Em análises sociológicas que envolvem cálculo de distâncias ou áreas, a seleção do CRS apropriado afeta diretamente as medidas: um CRS equidistante pode ser preferível se o interesse for medir distâncias exatas (por exemplo, acesso a serviços em quilômetros), já um CRS equivalente é indicado se for importante comparar áreas (por exemplo, comparar densidade demográfica em distintas regiões sem viés de projeção). Portanto, ao elaborar mapas ou análises espaciais rigorosas, deve-se justificar a escolha do CRS, documentando-o (p.ex., “SIG em SAD69 datum, projeção policônica”) e compreender suas implicações métricas – garantindo que comparações de distância e área façam sentido no contexto do fenômeno estudado.

Unidades de análise e MAUP: Na pesquisa espacial, definir claramente a unidade de análise geográfica é essencial. A unidade de análise é o delineamento espacial sobre o qual os dados são agregados ou analisados – pode ser uma divisão adminis-

trativa (cidades, estados), uma malha regular (grade quadrada), unidades naturais (bacias hidrográficas) ou zonas definidas pelo pesquisador. Essa escolha deve ser adequada à escala do fenômeno social investigado. Por exemplo, um estudo sobre segregação socioeconômica pode usar setores censitários ou bairros como unidades, pois capturam variações intraurbanas; já um estudo comparando políticas educacionais pode optar por municípios ou estados. Entretanto, é importante reconhecer o Problema da Unidade de Área Modificável (MAUP): resultados estatísticos podem variar conforme a unidade espacial adotada e o nível de agregação dos dados. Em outras palavras, diferentes delineamentos das unidades espaciais podem levar a conclusões diferentes a partir do mesmo dado subjacente (Ávila e Monasterio, 2006). Esse problema ocorre em duas dimensões: o efeito de escala, em que análises em escalas maiores ou menores (por exemplo, bairros vs. cidades) produzem resultados distintos; e o efeito de zoneamento, em que diferentes formas de agrupar zonas na mesma escala (por exemplo, redefinir os limites regionais) também alteram os resultados (Ávila e Monasterio, 2006). Uma implicação do MAUP é a falácia ecológica, que consiste em inferir indevidamente comportamentos individuais a partir de dados agregados (Ávila e Monasterio, 2006). Por exemplo, encontrar correlação entre violência e baixa renda em municípios não significa que os indivíduos de baixa renda sejam mais violentos – pode haver outras explicações, e a correlação observada pode ser artefato do nível de agregação (Ávila e Monasterio, 2006). Para mitigar o MAUP, pesquisadores devem testar a robustez das análises em diferentes escalas ou delimitações, além de justificar por que a unidade escolhida é apropriada ao fenômeno (por exemplo, municípios podem ser adequados para analisar políticas públicas que são implementadas municipalmente, ao passo que vizinhanças podem ser melhores para fenômenos comunitários). Sempre que possível, recomenda-se também verificar se padrões detectados persistem com unidades alternativas ou métodos multi-escalas, reforçando a validade dos insights espaciais obtidos.

Vizinhança espacial e matrizes de pesos: Grande parte das análises espaciais quantitativas depende do conceito de vizinhança espacial, que define como cada unidade geográfica relaciona-se com as demais no espaço. Para formalizar essa relação, constrói-se uma matriz de contiguidade ou de pesos espaciais (spatial weights matrix), usualmente denotada W . Essa matriz W especifica, para cada par de unidades i, j , um peso que representa o grau de proximidade ou influência entre elas – muitas vezes pesos binários 0/1 indicando se as unidades são vizinhas (1) ou não (0). Existem diferentes critérios para definir “vizinhos”. Os dois critérios clássicos de contiguidade em áreas poligonais são conhecidos como padrão Torre (Rook) e Rainha (Queen) – em alusão ao movimento das peças de xadrez. Na contiguidade

Rook, duas áreas são vizinhas apenas se compartilham um lado comum (fronteira linear); já na contiguidade Queen, considera-se vizinho qualquer área que toque a outra em pelo menos um ponto, seja lado ou vértice (ou seja, inclui também contatos diagonais) (Anselin, 1995). Assim, a vizinhança do tipo rainha tende a incluir mais vizinhos para cada unidade (até 8 numa grade regular, como os movimentos de uma rainha no tabuleiro), enquanto a do tipo torre inclui os adjacentes ortogonais (até 4, como movimenta a torre). Além da contiguidade direta, outro critério comum é a distância geográfica: define-se um raio (e.g. 50 km) ou um número k de vizinhos mais próximos, e unidades dentro desse limiar recebem peso como vizinhas. Há também matrizes de vizinhança padronizadas (row-standardized), em que os pesos de cada linha somam 1, e variantes que atribuem pesos graduados inversamente proporcionais à distância (p.ex., $1/d$ ou $1/d^2$). A construção cuidadosa da matriz de pesos é crucial, pois ela formaliza a suposição de como o espaço importa no fenômeno: por contiguidade administrativa? Por proximidade física? Por similaridade de contexto urbano? Por exemplo, no estudo de difusão de epidemias, pode fazer sentido usar vizinhança por distância (proximidade geográfica), enquanto num estudo de interações entre municípios, a contiguidade administrativa (compartilhar fronteira) pode capturar contatos socioeconômicos regionais. Em todos os casos, é fundamental explicitar o critério de vizinhança escolhido e justificar sua adequação. Essa decisão afeta diretamente métricas como autocorrelação espacial e modelos espaciais – e portanto análises de robustez com diferentes matrizes de pesos podem ser necessárias para verificar se os resultados não são altamente sensíveis a essa escolha.

Exemplo de mapa temático e integração de dados geográficos: Mapa do bioma Pantanal e municípios correspondentes, ilustrando a sobreposição de camadas – limites territoriais e dados populacionais municipais – em um único mapa. Este exemplo, produzido em QGIS e utilizando dados abertos do IBGE, mostra como um SIG permite combinar dados de diversas fontes (limites oficiais, estatísticas demográficas) para comunicar informação espacial de forma clara (Gonçalves *et al.*, 2025). No contexto de pesquisa social, mapas temáticos como este podem revelar padrões regionais (por exemplo, distribuição de população) e servir de base para análises mais complexas.

Métodos Exploratórios de Análise Espacial e Diagnóstico Socioterritorial

Uma vez compreendidos os fundamentos de representação espacial, os pesquisadores em ciências sociais empregam métodos exploratórios de análise espacial para identificar padrões e associações nos dados geográficos. Esses métodos, reunidos sob o rótulo de Análise Exploratória de Dados Espaciais (ESDA), oferecem ferramentas para diagnóstico espacial – isto é, para descobrir clusters, outliers e dependências espaciais que possam existir em fenômenos sociais distribuídos no território. Diferentemente de técnicas confirmatórias, os métodos exploratórios visam principalmente a descrição e a geração de hipóteses, não testes causais definitivos. Abaixo, destacamos alguns conceitos e técnicas introdutórias importantes: cartografia temática e classificação, autocorrelação espacial global (Moran I) e indicadores locais de associação espacial (LISA).

Cartografia temática e escolhas de classificação: A confecção de mapas temáticos é geralmente o primeiro passo na exploração de dados espaciais. Um mapa coroplético – aquele em que regiões são preenchidas com cores de acordo com valores de uma variável – pode revelar de imediato desigualdades territoriais ou padrões socioespaciais. Entretanto, um aspecto frequentemente subestimado é que a forma de classificar os dados num mapa coroplético pode alterar significativamente a percepção do padrão. Ao criar um mapa, o analista precisa decidir em quantas classes agrupar os valores e quais pontos de corte utilizar para definir essas classes (por exemplo, quartis, média etc.). Por exemplo, imagine que estamos mapeando a taxa de analfabetismo nos municípios. Se utilizarmos quebras naturais (método de Jenks), os limites entre classes serão escolhidos para maximizar as diferenças entre classes e minimizar dentro delas, possivelmente destacando regiões com valores excepcionais. Se, em vez disso, usarmos quintis (quantis) – de modo que 20% dos municípios caibam em cada classe – o mapa exibirá uma distribuição mais equilibrada de cores, mas valores extremos podem ser “suavizados” dentro de classes amplas. Já uma classificação por intervalos iguais (p.ex. 0–5%, 5–10%, 10–15%, etc.) facilita a comparação com metas ou faixas absolutas, porém pode resultar em muitas regiões na mesma classe se os dados estiverem concentrados numa faixa estreita. Cada técnica de classificação pinta uma história diferente para o leitor do mapa usando os mesmos dados subjacentes (Geo Sem Fronteiras, 2022). Portanto, encontrar um cluster alto-alto (regiões de valor alto cercadas por outras de valor alto) pode depender da paleta e das quebras utilizadas: escolhas arbitrárias podem exagerar ou ocultar um agrupamento. O analista deve então experimentar diferentes métodos, observar a

estabilidade dos padrões e, idealmente, informar qual classificação foi usada e por quê. Além disso, boas práticas incluem normalizar indicadores (e.g. usar taxas em vez de contagens absolutas) e aplicar cores intuitivas (por exemplo, do claro para o escuro indicando do baixo para o alto). Em suma, a cartografia temática é tanto arte quanto ciência, e escolhas conscientes de classificação aumentam a transparência e confiabilidade do diagnóstico espacial.

Autocorrelação espacial global – Índice I de Moran: Um conceito central em análise espacial é o de autocorrelação espacial, que avalia em que medida a distribuição de um fenômeno no espaço apresenta padrões de aglomeração (cluster) ou dispersão. O Índice I de Moran é uma estatística clássica para medir a autocorrelação espacial global de uma variável em todo o mapa. Seus valores variam tipicamente de -1 a +1 (descontando aspectos de amostragem): valores positivos indicam autocorrelação positiva (unidades vizinhas tendem a ter valores similares, ou seja, padrão em cluster), valor zero indica ausência de correlação espacial (padrão aleatório), e valores negativos indicam autocorrelação negativa (vizinhos tendem a ser muito diferentes, ou seja, padrão de segregação ou tabuleiro de xadrez). O Moran I é frequentemente expresso por uma fórmula como:

$$I = \frac{n}{W} \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2},$$

onde x_i é o valor da variável na unidade i , \bar{x} é a média global, w_{ij} são os elementos da matriz de pesos espaciais W (normalizados ou não) e W é a soma de todos w_{ij} . Intuitivamente, trata-se de comparar a covariância espacial (entre cada observação e seus vizinhos) com a variância total.

Uma maneira didática de interpretar o sinal e a magnitude do Moran I é através do diagrama de dispersão de Moran. Nesse diagrama, coloca-se no eixo horizontal o valor padronizado da variável em cada unidade (desviando da média) e no eixo vertical o valor médio (ou ponderado) dessa variável nas unidades vizinhas – essencialmente $W \times x$, a defasagem espacial. Cada ponto no gráfico representa uma unidade geográfica. Uma reta de regressão ajustada aos pontos tem coeficiente angular igual ao Moran I (Druck *et al.*, 2004; Anselin, 1995). Assim, se muitas unidades têm valor alto cercadas por vizinhos de valor alto e unidades de valor baixo cercadas por vizinhos de valor baixo, os pontos se concentrarão no quadrante alto-alto e baixo-baixo, respectivamente, resultando numa inclinação positiva ($I > 0$). Isso sugere a presença de clusters de semelhança – por exemplo, municípios ricos próximos de municípios também ricos, e municípios pobres próximos de outros pobres, indicando disparidades regionais marcantes. Por outro lado, se for comum que unidades al-

tas estejam cercadas de baixas ou vice-versa (pontos nos quadrantes alto-baixo e baixo-alto), a inclinação será negativa ($I < 0$), sugerindo um padrão espacial compensatório ou intercalado (ex.: bolsões de riqueza em meio a pobreza ou outliers espaciais). Sempre é necessário calcular a significância estatística do Moran I – usualmente via teste de permutação aleatória dos valores geográficos – para verificar se o padrão observado difere do acaso. Um Moran I significativamente positivo autoriza a inferência de que há dependência espacial global (o espaço importa na distribuição daquela variável), enquanto um Moran não significativo indica que, no nível global, não se pode rejeitar a hipótese de distribuição aleatória espacialmente. Importante notar: um Moran significativo não estabelece causas, apenas levanta hipóteses de processos geradores. Por exemplo, ao encontrar autocorrelação espacial positiva em índices de desemprego, pode-se hipotetizar processos de difusão espacial de condições econômicas ou efeitos de contexto regional; no entanto, isso não prova causalidade (pode haver variáveis omitidas influenciando regiões próximas de forma similar). O diagnóstico sugerido por um Moran I alto é “há clusters – investigue-os!”, mas não autoriza afirmar que um fenômeno causa o outro sem análise adicional.

Diagrama de dispersão de Moran: Ilustração de um Moran Scatterplot típico, onde cada ponto representa uma unidade geográfica. As linhas tracejadas dividem os quatro quadrantes (alto-alto, baixo-baixo, alto-baixo, baixo-alto) relativos à média. O coeficiente I de Moran corresponde à inclinação da reta ajustada aos pontos (Druck *et al.*, 2004; Anselin, 1995). No exemplo esquemático acima (dados de criminalidade em distritos), observa-se muitos distritos no quadrante alto-alto (valor acima da média e vizinhança também acima da média) e baixo-baixo, refletindo autocorrelação espacial positiva – distritos similares tendem a se agrupar. Esse diagrama auxilia a identificar outliers espaciais: pontos em quadrantes alto-baixo ou baixo-alto (valores destoantes do entorno). Em análise exploratória, mapas e Moran scatterplots são usados em conjunto: o índice global resume o padrão, enquanto a inspeção dos quadrantes e casos individuais permite perguntar que locais estão puxando o padrão e se eles fazem sentido substantivo (por exemplo, uma área de renda alta cercada por áreas de renda baixa pode indicar um enclave de riqueza em região pobre, merecendo atenção específica).

Clusters locais e mapas LISA: Complementando o Moran I global, que resume o padrão em todo o mapa, é muitas vezes útil identificar onde exatamente estão os clusters ou valores discrepantes. Para isso servem os indicadores locais de autocorrelação espacial (LISA), introduzidos por Luc Anselin (1995). O LISA de Moran local (às vezes chamado I_i local) calcula, para cada unidade i , uma estatística que reflete a associação daquele valor com os valores nos vizinhos de i . Em termos simples,

podemos pensar que o LISA de i indica quanto a unidade i contribui para o Moran I global. Mais praticamente, os LISAs permitem decompor o padrão: eles identificam clusters de Alto–Alto, clusters de Baixo–Baixo, outliers do tipo Alto–Baixo e Baixo–Alto, de forma estatisticamente significativa. Por exemplo, suponha a taxa de desemprego por município. Um cluster Alto–Alto significa um município com desemprego alto rodeado por municípios também com desemprego alto – isso poderia delinear uma região problemática, talvez um cinturão de estagnação econômica. Um cluster Baixo–Baixo, por outro lado, indica um bolsão de boas condições (baixa taxa rodeada de baixa taxa). Já um outlier Alto–Baixo pode apontar um município destoante positivamente (ilha de desemprego alto numa área de baixo desemprego, sugerindo um caso isolado a investigar), enquanto um Baixo–Alto seria o inverso (um “oásis” de desemprego baixo circundado por área de alto desemprego). Esses padrões locais são tipicamente representados em um mapa LISA, onde cada área é colorida segundo seu quadrante (alto-alto, etc.) e geralmente marca-se com símbolos as áreas cujo LISA é estatisticamente significativo a um nível escolhido (por exemplo, $p < 0,05$ após correção pelo número de testes). A leitura de um mapa LISA deve ser sempre conceitual: encontrar um cluster alto–alto para uma variável social sugere possíveis processos de homofilia espacial ou influências de fatores regionais comuns; porém, não autoriza imediatamente concluir causalidade. Por exemplo, um cluster de alta mortalidade infantil em municípios adjacentes (alto–alto) certamente aponta a existência de um fenômeno regional a ser explicado – talvez pobreza estrutural ou acesso precário a serviços de saúde numa sub-região – mas o LISA em si não diz qual é a causa, nem se os municípios interagem entre si ou apenas compartilham características. Ele sugere hipóteses: ex. “será que há uma política pública ausente nessa área?” ou “será uma zona socioeconômica historicamente desfavorecida?”. Ao mesmo tempo, o diagnóstico local também não prova que todos municípios do cluster alto-alto sejam iguais – eles podem ter diferenças internas, e o cluster mostra uma tendência geral. Portanto, a função dos métodos exploratórios (Moran global e locais) é diagnosticar padrões espaciais de maneira rigorosa, oferecendo insights que conduzem a perguntas de pesquisa mais refinadas ou apontam para a necessidade de modelos espaciais em etapas subsequentes. Em suma, encontrar um cluster no diagnóstico é ponto de partida, não de chegada: deve-se integrar esse achado ao conhecimento substantivo (por que ali?) e eventualmente testá-lo com dados adicionais ou métodos confirmatórios, evitando a armadilha de confundir padrão com causalidade.

Por fim, vale destacar que as escolhas feitas na etapa exploratória – como o tipo de mapa, método de classificação, definição da vizinhança e tratamento de outliers

– podem alterar a percepção dos padrões. Um exemplo concreto: se dividimos os dados de renda em poucas classes (digamos, apenas “alto” e “baixo”), podemos mascarar variações importantes e até alterar quais clusters aparecem significativos localmente. Já a decisão de normalizar variáveis (por exemplo, casos de doença *per capita* em vez de números brutos) é essencial para evitar que áreas simplesmente maiores ou mais populosas dominem o mapa. Portanto, o pesquisador deve refletir: “Que escolhas de visualização e medida estou fazendo e como elas influenciam o que vejo?” – isso faz parte do rigor no diagnóstico espacial. Ferramentas computacionais modernas (como GeoDa, ArcGIS, QGIS, R espacial) facilitam a experimentação interativa com esses parâmetros, permitindo checar como as aparências mudam. Essa prática de “*sensitivity analysis*” informal é altamente recomendada antes de fixar interpretações.

Em síntese, os métodos exploratórios fornecem um olhar inicial estruturado sobre dados espaciais sociológicos: revelam desigualdades regionais, identificam clusters quentes (*hotspots*) ou frios, sugerem regularidades espaciais que requerem explicação teórica. Na próxima seção, veremos como estudos empíricos concretos empregaram esses conceitos – como Moran e vizinhança – para diagnosticar problemas sociais e embasar políticas públicas, ilustrando a integração entre análise espacial e questões sociológicas substanciais.

Aplicações Empíricas e Comparação de Estudos Exemplares

Para tornar os conceitos mais palpáveis, examinaremos dois estudos que utilizaram SIG e lógica de vizinhança de forma exemplar em pesquisas sociais, destacando como formularam perguntas claras, documentaram escolhas espaciais e conectaram seus diagnósticos a implicações de política. Em seguida, faremos uma comparação narrativa entre esses estudos em quatro eixos: (1) pergunta de pesquisa, (2) dados e construção espacial, (3) estratégia de análise e diagnóstico, (4) implicações e limites. Os estudos selecionados são:

- Albertus, Brambor e Ceneviva (2018) – que investigou a relação entre desigualdade fundiária e conflito rural no Brasil, explorando como a presença de programas de reforma agrária em municípios vizinhos afeta a ocorrência de invasões de terra.
- Gonçalves *et al.* (2025) – que avaliou os efeitos das queimadas no Pantanal

sobre internações hospitalares por doenças respiratórias e cardiovasculares, integrando dados ambientais (focos de incêndio, clima, poluição) com dados de saúde pública.

Estudo 1: Desigualdade fundiária e conflito rural no Brasil (Albertus *et al.*, 2018) *Pergunta e contexto:* O estudo de Albertus, Brambor e Ceneviva (2018) parte de uma questão clássica de sociologia política: qual é a relação entre concentração de terras e ocorrência de conflitos rurais (como invasões de terra por movimentos sociais)? Além disso, pergunta por que programas de reforma agrária – que teoricamente atenuariam conflitos ao redistribuir terras – às vezes agravam os conflitos em vez de mitigá-los (Albertus, Brambor e Ceneviva, 2018). Os autores focam no caso do Brasil entre 1988 e 2013, período marcado por intensa mobilização por reforma agrária e respostas heterogêneas das elites agrárias. A hipótese inovadora deles é que não basta olhar a desigualdade fundiária isoladamente; é preciso entender o contexto espacial de ações de reforma agrária próximas. Em outras palavras, eles investigam se a presença de assentamentos de reforma agrária em municípios vizinhos altera o comportamento de proprietários de terras e invasores em um dado município. Isso já indica uma abordagem espacial: a consideração de eventos em localidades vizinhas como parte da explicação para conflitos locais.

Dados e construção espacial: A unidade de análise escolhida foi o município brasileiro, adequado ao fenômeno pois invasões de terra e assentamentos são contabilizados nessa escala (dados administrativos e de movimentos sociais costumam ser agregados municipalmente). Os autores compilaram um conjunto de dados municipais abrangendo: medidas de desigualdade fundiária (por exemplo, Gini de distribuição de terras ou percentuais de área em grandes propriedades), ocorrência de invasões de terra (eventos de conflito, possivelmente dos registros da Comissão Pastoral da Terra ou INCRA) e informação sobre assentamentos de reforma agrária implementados. Uma característica importante é que eles introduzem variáveis para captar a “vizinhança” ou contexto espacial: em particular, uma medida de ameaça local vinda de perto, operacionalizada como algo como o número (ou existência) de projetos de reforma agrária em municípios adjacentes (“*nearby land reforms*”) (Albertus, Brambor e Ceneviva, 2018). Ou seja, para cada município, verificaram se nos municípios contíguos (vizinhos diretos) houve implementação de reforma agrária em dado ano, interpretando isso como um choque externo percebido pelas elites fundiárias locais. Em termos de matriz de vizinhança, é plausível o uso de contiguidade do tipo *queen* ou *rook* para definir quem é vizinho de quem. A análise considerou também a dimensão temporal (1988–2013), possivelmente usando modelos de painel, e incluiu controles socioeconômicos regionais. Para robustez, os autores relatam ter feito

testes com defasagens espaciais adicionais e controle de efeitos regionais, o que indica preocupação em captar autocorrelação residual e dependência entre municípios próximos (Albertus, Brambor e Ceneviva, 2018).

Estratégia de análise e diagnóstico espacial: Embora o trabalho seja quantitativo e inferencial (modelos estatísticos), podemos interpretá-lo sob a luz de diagnóstico espacial. Os autores notaram algo contraintuitivo: alguns dos municípios menos desiguais em terras eram bastante conflituosos, o que sugeria que não havia uma relação linear simples entre concentração fundiária e invasões. Ao invés disso, a análise exploratória indicou padrões condicionais: conflitos pareciam ocorrer onde reformas agrárias aconteceram próximo, alterando a dinâmica local. Em termos de diagnóstico, isso é detectar um padrão de interação espacial – a variável “assentamentos em vizinhos” parecia “catalisar” respostas diferentes conforme o contexto local de desigualdade. A formalização veio nos modelos: eles incluíram um termo de interação entre desigualdade fundiária local e indicador de reforma agrária em vizinhos. Os resultados mostraram um efeito condicional: onde não havia ameaças próximas (sem reforma agrária em torno), municípios com maior desigualdade de terras apresentavam mais invasões; entretanto, onde havia reforma agrária em municípios vizinhos, o padrão se invertia: nesses locais, uma alta desigualdade estava associada a menos invasões, sugerindo que as elites fundiárias, ao verem reformas acontecendo ao redor, se organizaram coletivamente (no âmbito regional) para reprimir ou prevenir invasões em seu território (Albertus, Brambor e Ceneviva, 2018). Ou seja, desencadeou-se um *spillover* territorial: as reformas agrárias vizinhas atuaram como gatilho para mobilização defensiva, “vacinando” ou blindando aquelas localidades contra invasões. Os autores verificaram a robustez disso com vários testes, inclusive adicionando defasagens espaciais das variáveis dependentes para checar se a dependência espacial direta influenciava – confirmando que o mecanismo proposto se mantinha (Albertus, Brambor e Ceneviva, 2018).

Em termos de limites, os autores reconhecem que seus resultados são condicionais ao contexto brasileiro e período específico – outros países com outros arranjos políticos poderiam apresentar dinâmicas diferentes. Além disso, como todo estudo observacional, não há prova definitiva de causalidade micro; eles mitigaram isso incluindo efeitos fixos no tempo e espaço e testando outros mecanismos, mas a ressalva permanece.

Estudo 2: Queimadas e internações hospitalares no Pantanal (Gonçalves *et al.*, 2025) *Pergunta e contexto:* O estudo de Gonçalves e colaboradores (2025) examina um problema socioambiental agudo: as queimadas florestais no Pantanal

aumentam as internações hospitalares por problemas respiratórios e cardiovasculares? O Pantanal, maior área úmida tropical do mundo, passou por incêndios severos nos últimos anos, levantando preocupações sobre impactos sanitários para as populações locais. A pergunta tem relevância de saúde pública e ambiental – especialmente num contexto de mudanças climáticas e políticas ambientais em discussão. Diferente do estudo anterior, que era espacial em termos regionais, aqui a ênfase é em integrar dados ambientais e de saúde no espaço geográfico e no tempo para quantificar uma relação.

Dados e construção espacial: Este trabalho ilustra a integração de múltiplas bases de dados espaciais. Primeiro, dados de queimadas: contagens de focos ativos de incêndio detectados por satélites (INPE), agregados por dia e por município. Segundo, dados de internações hospitalares: extraídos do SIH/SUS (DATASUS), contendo registros diários de internações por causas definidas (doenças respiratórias e cardiovasculares). Para controle de fatores de confusão ambientais, incorporaram dados de qualidade do ar e clima (temperatura, talvez umidade) como variáveis explicativas. Em termos espaciais, a unidade básica de observação foi município–dia na região do Pantanal (21 municípios de MT e MS), com mapas contextuais da área de estudo (Gonçalves *et al.*, 2025).

Estratégia de análise e diagnóstico: A abordagem foi quantitativa com modelos lineares generalizados (contagem), precedida de diagnóstico exploratório (tendências, mapas de focos, poluição e mortalidade). Resultado principal: um aumento de 10 focos ativos de queimada associa-se a aumento significativo nas internações diárias; estimou-se que, ao longo de 10 anos, cada 10 focos adicionais por dia elevaram em cerca de 23% o risco de internação respiratória e 22% o de internação cardiovascular (Gonçalves *et al.*, 2025). A plausibilidade biológica e espacial (relevo plano que retém fumaça) sustenta a interpretação.

Implicações e limites: O estudo fornece insumo acionável para políticas: sistemas de alerta com base em detecção de focos para preparação do sistema de saúde; integração entre meio ambiente e saúde; e ações preventivas de queimadas. Limites: dados de saúde captam casos que chegam ao hospital (subregistro possível), e a inferência é agregada; contudo, a robustez temporal e a coerência espacial reforçam os achados (Gonçalves *et al.*, 2025).

Comparação crítica dos dois estudos

Perguntas e dados: Albertus *et al.* (2018) tratam de sociologia política rural (desigualdade e conflito) com ênfase em contexto vizinho; Gonçalves *et al.* (2025) abordam saúde ambiental com integração de dados ambientais e hospitalares. **Estratégia:** O primeiro enfatiza interação espacial (vizinhança de reformas), o segundo quantifica associação espaço-temporal com GLM. **Implicações:** O primeiro alerta para transbordamentos regionais em políticas agrárias; o segundo quantifica custos humanos das queimadas, orientando prevenção e preparo da rede de saúde.

Ética, Privacidade e Governança de Dados Espaciais

O uso de dados espaciais em ciências sociais traz responsabilidades éticas e legais. Sob a LGPD brasileira, dados de geolocalização podem ser sensíveis, exigindo anonimização/pseudonimização, agregação espacial e controle de acesso quando se trabalha com indivíduos (Imagem Geosistemas, 2025; Mattos Filho, 2025). É crucial evitar estigmatização territorial (contextualizar mapas de “hotspots” sem culpabilizar comunidades) e adotar transparência/reprodutibilidade: compartilhar dados abertos quando possível, scripts e metadados, seguindo princípios FAIR e padrões abertos (Esri, 2025; INDE, 2018; INDE, 2025; CIDACS, 2025; Murrieta, 2024).

Inovações e Desafios em Países de Renda Média e Baixa

Destacam-se integrações entre bases administrativas/censitárias e dados digitais (imagens de satélite, celular), plataformas abertas e *open source* (QGIS, R, Python), além de iniciativas de ciência aberta e reprodutibilidade (Murrieta, 2024; Esri, 2025; CIDACS, 2025). Persistem desafios de infraestrutura, custo de dados e capacitação, bem como sustentabilidade de projetos e abertura governamental (INDE, 2025; INDE, 2018).

Limites e Próximos Passos da Análise Espacial Sociológica

Padrões descritivos não provam causalidade; escolhas de escala (MAUP) influenciam resultados; e bordas administrativas podem ocultar transbordamentos. Próximos

passos incluem desenhos causais, análises multi-escala, triangulação com métodos qualitativos e testes de sensibilidade (Ávila e Monasterio, 2006; Druck *et al.*, 2004).

Apêndice de Transparência: Decisões Técnicas do Estudo

(Unidade de análise; preparo dos dados; CRS; classificação cartográfica; desenho de vizinhança; privacidade; metadados e reprodutibilidade; variáveis e controles; documentação de scripts — conforme descrito no texto principal, com ênfase em boas práticas e princípios FAIR.)

Referências

ALBERTUS, M.; BRAMBOR, T.; CENEVIVA, R. Land Inequality and Rural Unrest: Theory and Evidence from Brazil. *Journal of Conflict Resolution*, v.62, n.3, p.557–596, 2018.

ANSELIN, L. Local Indicators of Spatial Association — LISA. *Geographical Analysis*, v.27, n.2, p.93–115, 1995.

ÁVILA, R. P.; MONASTERIO, L. M. O MAUP e a análise espacial: um estudo de caso para o Rio Grande do Sul (1991–2000). *Anais do 3º Encontro de Economia Gaúcha*, 2006. Disponível em: <https://arquivofee.rs.gov.br/3eeg/Artigos/m23t01.pdf>. Acesso em: 2025.

BRASIL ESCOLA. Sistema de Informações Geográficas (SIG). *Brasil Escola — Geografia*, s.d. Disponível em: <https://brasilecola.uol.com.br/geografia/sig.htm>. Acesso em: 2025.

CIDACS — Centro de Integração de Dados e Conhecimentos para Saúde. *Dados geoespaciais*. 2025. Disponível em: <https://cidacs.bahia.fiocruz.br/tag/dados-geoespaciais/>. Acesso em: 2025.

DRUCK, S.; CARVALHO, M. S.; CÂMARA, G.; MONTEIRO, A. M. V. (orgs.). *Análise espacial de dados geográficos*. Brasília: EMBRAPA, 2004.

ESRI. Recursos de Ciência Aberta — Iniciativas de ciência aberta da Esri. 2025. Disponível em: <https://www.esri.com/pt-br/about/science/initiatives/open-science>. Acesso em: 2025.

ESRI PORTUGAL. O que são SIG? *Esri Portugal*, s.d. Disponível em: <https://www.esri-portugal.pt/pt-pt/o-que-sao-os-sig/overview>. Acesso em: 2025.

GEO SEM FRONTEIRAS. Mapas coropléticos — Um guia para classificação de dados.

Blog Geo Sem Fronteiras, 15 jun. 2022. Disponível em: <https://geosemfronteiras.org/blog/mapas-coropleticos-um-guia-para-classificacao-de-dados/>. Acesso em: 2025.

GONÇALVES, A. C. *et al.* Additive effect of wildfires on hospital admission in the Pantanal wetland, Brazil. *Scientific Reports*, v.15, art.27572, 2025.

IMAGEM GEOSSISTEMAS (Esri Brasil). Geoprivacidade e ética de dados espaciais na era da LGPD. *Blog da Imagem*, 10 out. 2025. Disponível em: <https://blog.img.com.br/arcgis/geoprivacidade-gis-lgpd-etica-2/>. Acesso em: 2025.

INDE — Infraestrutura Nacional de Dados Espaciais. *Portal INDE*. 2025. Disponível em: <https://inde.gov.br/>. Acesso em: 2025.

INDE — Infraestrutura Nacional de Dados Espaciais. *Perfil de Metadados Geoespaciais do Brasil*. Brasília, 2018. Disponível em: <https://www.inde.gov.br/pdf/liv101802.pdf>. Acesso em: 2025.

MATTOS FILHO. ANPD abre consultas sobre anonimização, pseudonimização e direitos dos titulares. *Portal Mattos Filho*, 2025. Disponível em: <https://www.mattosfilho.com.br/unico/anonimizacao-pseudonimizacao-direitos-titulares/>. Acesso em: 2025.

MURRIETA, M. Ciência aberta: um novo modo de fazer ciência — mais colaborativo, transparente e sustentável. *Controladoria-Geral da União*, 15 ago. 2024.

WIKIPÉDIA. Sistema de informação geográfica. *Wikipédia, a enciclopédia livre*, s.d.-a. Disponível em: https://pt.wikipedia.org/wiki/Sistema_de_informa%C3%A7%C3%A3o_geogr%C3%A1fica. Acesso em: 2025.

WIKIPÉDIA. Dados vetoriais. *Wikipédia, a enciclopédia livre*, s.d.-b.

WIKIPÉDIA. Dados raster. *Wikipédia, a enciclopédia livre*, s.d.-c.

WIKIPÉDIA. Resolução espacial. *Wikipédia, a enciclopédia livre*, s.d.-d.