

Caracterização de processos pontuais marcados pela função de correlação marcada: uma aplicação em dados florestais

Wélson A. Oliveira^{1†}, Wigor D. M. Santos¹, José Márcio de Mello², João D. Scalon³

¹Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, Instituto de Ciências Exatas e Tecnológicas, Universidade Federal de Lavras (UFLA).

²Departamento de Ciências Florestais, Instituto de Ciências Exatas e Tecnológicas, Universidade Federal de Lavras.

³Departamento de Estatística, Escola de Ciências Agrárias de Lavras, Universidade Federal de Lavras.

Resumo: Entende-se um processo pontual espacial como um conjunto de pontos (eventos) distribuídos de forma irregular no espaço, gerados por um mecanismo probabilístico estocástico. Associando um atributo (marca) à coordenada, determina-se um processo pontual marcado. Na análise de processos pontuais marcados, o interesse está na caracterização do padrão de interação entre os processos estocásticos que geraram os pontos e as marcas. As análises iniciam-se com a caracterização de efeitos de primeira ordem para visualização completa das intensidades de ocorrência. Assumindo estacionariedade, pode-se realizar a análise de segunda ordem para caracterizar a dependência espacial presente no fenômeno. Os dados consistem em localizações georreferenciadas de ocorrência de árvores arbóreas em um fragmento de floresta nativa e seus respectivos diâmetros à altura do peito (DAP). Este trabalho tem como objetivo caracterizar os processos pontuais marcados por variáveis contínuas, dados pelo DAP das árvores nativas, através da estimativa da função de correlação marcada. Todas as análises foram realizadas com o software R, desenvolvido pela equipe do R Core Team (2023). A partir dos resultados, foi possível observar o potencial dos métodos utilizados para caracterizar os padrões e a dinâmica da floresta em estudo.

Palavras-chave: Correlação Marcada; Processos Pontuais Marcados; Dependência Espacial; Estimador Kernel.

Characterization of marked point processes by the marked correlation function

Abstract: A spatial point process is understood as a set of points (events) irregularly distributed in space, generated by a stochastic probabilistic mechanism. Associating an attribute (mark) with the coordinate determines a marked point process. In the analysis of marked point processes, the interest lies in characterizing the pattern of interaction between the stochastic processes that generated the points and the marks. Analyses start with the characterization of first-order effects for a complete visualization of occurrence intensities. Assuming stationarity, a second-order analysis can be performed to characterize the spatial dependence present in the phenomenon. The data consist of georeferenced locations of occurrence of tree individuals in a native forest fragment and their respective diameters at breast height (DBH). This work aims to characterize the marked point processes by continuous variables, given by the DBH of native trees, through the estimation of the marked correlation function. All analyses were performed using the R software, developed by the R Core Team (2023). From the results, it was possible to observe the potential of the methods used to characterize the patterns and dynamics of the forest under study.

Keywords: Marked Correlation; Marked Point Processes; Spatial Dependence; Kernel Estimator.

Introdução

Muitos fenômenos aleatórios de origem natural podem ser representados como pontos georreferenciados no espaço e estudados por meio da teoria de processos pontuais. Um processo pontual espacial corresponde a um conjunto de pontos distribuídos de forma irregular em uma área definida, cuja estrutura espacial foi gerada por um mecanismo estocástico.

[†]Autor correspondente: oliveirawelson99@gmail.com.

O objetivo das análises estatísticas de um processo pontual é analisar a estrutura geométrica do padrão de pontos que descrevem as localizações de ocorrência dos eventos ou objetos distribuídos no espaço. Se, além da localização, obtemos informações sobre características dos pontos, estas são chamadas de marcas e definem um processo pontual marcado.

A análise de processos pontuais marcados se concentra em questões como as diferenças numéricas entre as marcas dependentes das distâncias dos pontos correspondentes. Pode haver interesse em analisar se as marcas dos pontos vizinhos tendem a ser menores (ou maiores) que a marca média, podendo haver inibição (ou estimulação) das marcas. Outra questão diz respeito à semelhança numérica das marcas dos pontos vizinhos. Isso significa que esses pontos tendem a ter marcas semelhantes, mas o comportamento oposto também pode ser observado, podendo haver pontos específicos que dominam outros em sua vizinhança, como no caso da dispersão de sementes de uma árvore próxima de sua localidade.

Muitos aspectos diferentes da natureza de um padrão de ponto espacial específico podem ser descritos usando os métodos estatísticos apropriados. Através da análise de primeira ordem, podemos obter a visualização completa das intensidades de ocorrências. As características de segunda ordem para processos pontuais com marcas quantitativas podem ser introduzidas com o interesse de caracterizar a interação entre os processos que geram os pontos e as marcas.

Em processos pontuais marcados de árvores e DAP (diâmetro à altura do peito), a análise de correlação espacial pode ajudar a identificar padrões e tendências espaciais na distribuição dessas variáveis, contribuindo não só para o entendimento do comportamento, mas também para a tomada de decisões relacionadas à gestão, planejamento e preservação ambiental.

Dessa forma, este trabalho tem como objetivo caracterizar os processos pontuais definidos pela ocorrência de árvores nativas, marcados por variáveis contínuas obtidas pelos respectivos DAP de cada planta para duas espécies localizadas em um microfragmento florestal.

Referencial Teórico

Processos Pontuais

Um processo pontual é obtido através de uma coleção de variáveis aleatórias definidas por $X = \{X(\mathbf{s}); \mathbf{s} \in S\}$, sendo $S \subset \mathbb{R}^d$, $d = 2$, $\mathbf{s} = (x_i, y_i) \in \mathbb{R}^2$ e $X(\mathbf{s})$ um mecanismo probabilístico estocástico que rege as distribuições de eventos no espaço. Em geral, as observações serão obtidas apenas de uma região limitada $A \subset S$, contendo um conjunto finito de eventos (ILLIAN, 2008). Em análises de processos pontuais, o interesse está exclusivamente na localização pontual da ocorrência do evento. Quando as ocorrências carregam algum atributo ou informação além da localização, são chamados de marcas e se consideradas nas análises, determina-se um processo pontual marcado (MØLLER; WAAGEPETERSEN, 2017).

Um processo pontual espacial marcado é um processo pontual espacial definido por:

$$X_M = \{(\mathbf{s}, m_{\mathbf{s}}), (\mathbf{s}, m_{\mathbf{s}}) \in X \times M\},$$

sendo que $\mathbf{s} \in X$ corresponde as localizações dos eventos e $m_{\mathbf{s}} = m(\mathbf{s})$ as marcas respectivamente.

A caracterização do fenômeno em estudo pode ser feita pela análise dos efeitos de primeira e segunda ordem (ou primeiro e segundo momentos do processo estocástico). Efeitos de primeira ordem, globais ou de larga escala, correspondem variações do valor médio do processo estocástico no espaço (CRESSIE, 1993). O interesse está em determinar a intensidade ou o número esperado de eventos por unidade de área CÂMARA; DAVIS; MONTEIRO, 2001).

Os efeitos de primeira ordem podem ser definidos pela função de intensidade dada por

$$\lambda_1(s_i, m_i) = \lim_{|ds_i \times dm_i| \rightarrow 0} \left\{ \frac{\mathbb{E}[N(ds_i \times dm_i)]}{|ds_i \times dm_i|} \right\} \quad (1)$$

sendo $d\mathbf{s}_i \times dm_i$ uma pequena região em torno de \mathbf{s}_i marcado por m_i , e $N(d\mathbf{s}, dm_i)$ representa o número de eventos de marcas m_i localizados nessa região.

Considerando um processo pontual estacionário ou homogêneo, o estimador de intensidade equivale ao número médio de eventos observados. Para visualização completa de intensidades, pode-se utilizar de métodos paramétricos e não paramétricos. Neste trabalho utilizará de um método não-paramétrico baseado na estimativa da função Kernel que ajusta uma função bidimensional sobre a área de estudo, criando uma superfície suavizada na região ponderação pela distância (TERRELL; SCOTT, 1992).

Os efeitos de segunda ordem, considerados locais ou de pequena escala, definem a estrutura espacial (dependência espacial) do processo.

As propriedades de segunda ordem podem ser definidas da mesma forma que os efeitos de primeira ordem considerando a intensidade conjunta entre duas regiões infinitesimais, apresentada na equação (2):

$$\lambda_2((\mathbf{s}_i, m_i), (\mathbf{s}_j, m_j)) = \lim_{|d\mathbf{s}_i \times dm_i|, |d\mathbf{s}_j \times dm_j| \rightarrow 0} \left\{ \frac{\mathbb{E}[N(d\mathbf{s}_i \times dm_i)N(d\mathbf{s}_j \times dm_j)]}{|d\mathbf{s}_i \times dm_i||d\mathbf{s}_j \times dm_j|} \right\}, \quad (2)$$

sendo que, $N(d\mathbf{s}_i \times dm_i)$ e $N(d\mathbf{s}_j \times dm_j)$ representam o número de eventos dentro de pequenos círculos centrados nos pontos (\mathbf{s}_i, m_i) e (\mathbf{s}_j, m_j) de área $|d\mathbf{s}_i \times dm_i|$ e $|d\mathbf{s}_j \times dm_j|$.

A não observância de efeitos de segunda ordem em um processo pontual marcado, caracteriza processos de completa independência espacial entre marcas e pontos (CIE). Algumas medidas podem ser utilizadas para testar hipóteses contra CIE como a função de correlação marcada. (DIGGLE, RIBEIRO e CHRISTENSEN, 2003).

Estimador Kernel

Seja $\mathbf{s}_1, \dots, \mathbf{s}_n$ localizações de n eventos de uma região A . Por Terrel e Scott (1992), o estimador de kernel propõe que a estimativa de intensidade para um ponto \mathbf{x} qualquer de A é dado pela Equação (3),

$$\hat{\lambda}_\tau(\mathbf{x}) = \frac{1}{\tau^2} \sum_{i=1}^n k\left(\frac{d(\mathbf{s}_i, \mathbf{x})}{\tau}\right), \quad d(\mathbf{s}_i, \mathbf{x}) \leq \tau. \quad (3)$$

sendo que τ representa um raio de influência ($\tau \geq 0$) que determina quais dos n eventos serão utilizados na estimativa da intensidade de ocorrência na posição de \mathbf{x} e $k(\cdot)$ é uma função de estimação com propriedades de suavização do fenômeno (CÂMARA; DAVIS; MONTEIRO, 2001).

A dificuldade de aplicação do método de kernel, está na definição do raio de influência τ que seja "ideal". Algumas propostas são apresentadas pela literatura baseadas, por exemplo, na relação entre o número de pontos e de dimensões espaciais de Terrel e Scott (1992), minimização do erro quadrático médio Diggle (1985), maximização do critério de validação cruzada de verossimilhança do processo de ponto proposto por Loader (2006). Neste trabalho será utilizado da definição de τ pela fórmula de Campbell, que produz melhores estimativas, sendo mais eficaz na identificação de tendência (CRONIE; LIESHOUT, 2018).

Função de Correlação Marcada

A correlação espacial entre marcas contínuas de um processo pontual marcado pode ser descrita através de uma função de teste $f(m_i, m_j)$ definida em termos das marcas m_i e m_j com $i \neq j$ caracterizando a correlação condicionada a uma distância t entre os pontos. A função de correlação marcada pode ser definida por illian (2008) como:

$$\rho_f(t) = \frac{E[f(m_i, m_j) \times t]}{\mu^2} \quad (4)$$

sendo $f(m_i, m_j)$ uma função teste para duas marcas m_i e m_j , μ é a esperança das marcas consideradas, ($E(m_i) = \mu$) e t é a distância euclidiana entre os pontos em que é observado as marcas m_i e m_j . $E(f(m_i, m_j) \times t)$ representa a média condicionada à distância t e o denominador remove os efeitos de escala de diferentes ordens.

Um estimador não viesado para a função de correlação marcada, usando o suavizador Epanechnikov de kernel e a correção de borda isotrópica, é dado por

$$\hat{\rho}_f(t) = \frac{1}{\mu^2} \sum_{i=1}^n \sum_{j=1}^n f(m_i, m_j) \times A_{ij} \quad (5)$$

sendo, $A_{ij} = \frac{e_\delta(t - |s_i - s_j|)}{W(A_i \cap A_j)}$, e $e_\delta(h) = \begin{cases} \frac{3}{4\delta} \left(1 - \frac{t^2}{\delta^2}\right), & \delta \leq t \leq \delta \\ 0, & c.c. \end{cases}$ o suavizador de Epanechnikov

de kernel com parâmetro de suavidade δ .

Esta função não é uma correlação no sentido habitual em estatística, por assumir qualquer valor real não negativo. Para valores grandes de t ou se as marcas forem não correlacionadas, tem-se que $\hat{\rho}_f(t) = 1$, isto é, a média do produto é igual ao produto das médias. Isso significa que não há dependência entre marcas e pontos do processo pontual. Se $\hat{\rho}_f(t) < 1$, indica uma tendência de inibição das marcas, resultando numa diminuição das marcas ou existência de marcas inferiores a média à uma distância t . Quando $\hat{\rho}_f(t) > 1$, indica tendência de agregação das marcas, resultando no aumento de marcas ou existência de marcas superiores a média à uma distância t definida.

Teste de hipóteses

A definição do teste de hipóteses recorre da seguinte forma: Completa independência espacial entre marcas e pontos como hipótese de nulidade e interação (correlação espacial) positiva ou negativa para hipótese alternativa. Por não trabalhar com distribuições conhecidas para o processo estocástico, recorre-se a algoritmos de simulação para decisão do desvio estatisticamente significativo. Os métodos de Monte Carlo fornecem uma maneira conveniente para decisão do que é estatisticamente significativo e que contorna a teoria da distribuição intratável que surge mesmo com modelos simples de processos de pontos espaciais (ANDERSEN, 1992). Utilizando desse método, a condução do teste de hipótese pode ser realizada considerado tanto testes formais, quanto procedimentos gráficos com a construção dos envelopes de confiança.

Considere uma estatística de teste geral $\Phi(t)$, sendo t conhecido. Ao optar por um teste de hipótese formal, deve-se seguir o seguinte procedimento.

- i Definição de u_i como sendo a medida da discrepância entre a função $\Phi(t)$ estimada $\hat{\Phi}_i(t)$ e a teórica sob hipótese de independência $\Phi(t)$, sobre toda distância $0 \leq t \leq t_0$, dada por:

$$u_i = \int_0^{t_0} \{\hat{\Phi}_i(k) - \Phi(k)\}^2 dk, \quad i = 1, \dots, s. \quad (6)$$

sendo o limite t_0 a distância máxima estabelecida de acordo com a área de estudo ou pelo próprio pesquisador.

- ii Calcula-se u_0 , correspondendo a medida para o processo pontual marcado observado.
- iii Calcula-se $u_i : i = 1, 2, \dots, s$, em que s é o total de simulações e u_i é a medida para os processos simulados sob a hipótese de independência entre marcas e pontos.
- iv Ordena-se u_i .
- vi Se obtivermos u_0 em um dos extremos da distribuição dos $u_i, i = 1, 2, \dots, s$, então existem evidências estatísticas para rejeitar a hipótese nula de independência entre marcas e pontos.

O passo a passo proposto, pode ser utilizado para rejeitar ou não rejeitar a hipótese nula. Porém, não indica o tipo de padrão evidenciado. Para isso utiliza-se da análise de imagens através da construção de envelopes de confiança (ILLIAN, 2008). Este será o tipo de análise utilizada no trabalho, e pode ser obtido através do seguinte procedimento:

1. Considere uma estatística de teste geral $\Phi(t)$.
2. Calcula-se a função $\hat{\Phi}_1(t)$ para a configuração do processo pontual observado.
3. Calcula-se a função $\hat{\Phi}_i(t) : i = 2, 3, \dots, s$, para cada $(s - 1)$ configuração, sob completa aleatoriedade espacial.
4. Calcula-se os envelopes de simulação superiores e inferiores salve as equações (7) e (8) respectivamente.

$$U_i(t) = \min\{\hat{\Phi}_i(t)\}, i = 2, \dots, s. \quad (7)$$

$$U_s(t) = \max\{\hat{\Phi}_i(t)\}, i = 2, \dots, s. \quad (8)$$

5. Constrói um gráfico com os envelopes de simulações definido por $U_i(t)$ e $U_s(t)$ e a estimativa $\hat{\Phi}_1(t)$.
6. Faça a interpretação conveniente para a função $\Phi(t)$ utilizada.

A função $\Phi(t)$ está representando uma família de descritores (J, L , Correlação marcada entre outras), diferindo unicamente pela interpretação (passo 6).

Materiais e Métodos

Apresentação dos dados

Os dados consistem de localizações em coordenadas UTM (Universal Transversa de Mercator) de 1147 árvores nativas de duas espécies distintas. Para isso, a região em estudo foi dividida em parcelas, sendo que nos vértices possuía uma coordenada UTM conhecida. Das árvores de cada parcelas foram retiradas informações do plano cartesiano (x,y) definida na própria parcela. A partir desta coordenada foi feita a transformação em UTM tendo como base a informação do vértice da parcela onde se localizava a árvore. Além da localização, obteve-se informações (marcas) do tipo de espécie e do diâmetro a altura do peito, correspondendo a aproximadamente 1,3 metros do solo. Em estudo têm-se 505 da espécie *Copaifera langsdorffii* e 642 árvores da espécie *Xylopia brasiliensis*, espécies mais populosas encontradas no fragmento pelo inventário de 2017.

A área em estudo corresponde a um fragmento florestal Estacional Semidecidual Montana de domínio Mata Atlântica de aproximadamente 6,48 hectares de extensão sob altitude média de 950 metros acima do nível do mar. Está localizado nas coordenadas $21^{\circ}13'43.4''S$ e $44^{\circ}58'16.1''W$ e abrange uma área localizada dentro do campus da Universidade Federal de Lavras, Lavras, Minas Gerais, Brasil sendo apresentado pela Figura 1.

Métodos Estatísticos

A estatística do trabalho inclui uma análise exploratória inicial utilizado de medidas estatísticas clássicas acerca das marcas como média, mediana, amplitude, desvio padrão, coeficiente de variação. Também serão apresentados histogramas e boxplots para uma melhor visualização da distribuição do DAP das espécies.

Para visualização completa da variação de intensidades para cada uma das espécies, será utilizada a função alisamento de Kernel. O estimador utilizado será o mesmo apresentado pela

Figura 1: Fragmento floresta Semidecidual Montada ("Matinha").



Fonte: Autores.

Equação 3. Para adequação do raio de suavização τ , será utilizado a medida proposta por Cronie e Lieshout (2018).

Para análise de interação entre diâmetro a altura do peito e localização das árvores de cada uma das espécie arbóreas nativas, foi utilizado da função de correlação marcada para dados contínuos estimada dentro do intervalo de 0 a 100 metros. Para isso, utilizou-se da função de teste dada por $f(m_i, m_j) = m_i \times m_j$ e considera-se um estimador não viesado, usando o suavizador Epanechnikov de Kernel e correção de borda, conforme é apresentado pela Equação 5. A construção dos envelopes de confiança foi feita utilizando de 1000 simulações Monte Carlo.

Todas as análises foram realizadas pelo software R Core Team (2023), utilizando funções disponibilizadas na biblioteca *spatstat* (BADDELEY; TURNER, 2005).

Resultados e discussão

Inicialmente foi realizado uma análise descritiva contemplando medidas estatísticas clássicas (Tabela 1), bem como a construção de histogramas e boxplots apresentados na Figura 2 para ilustração do comportamento da distribuição do DAP de cada espécie em estudo.

Tabela 1: Tabela descritiva de diâmetro à altura do peito (cm) por espécie de árvore.

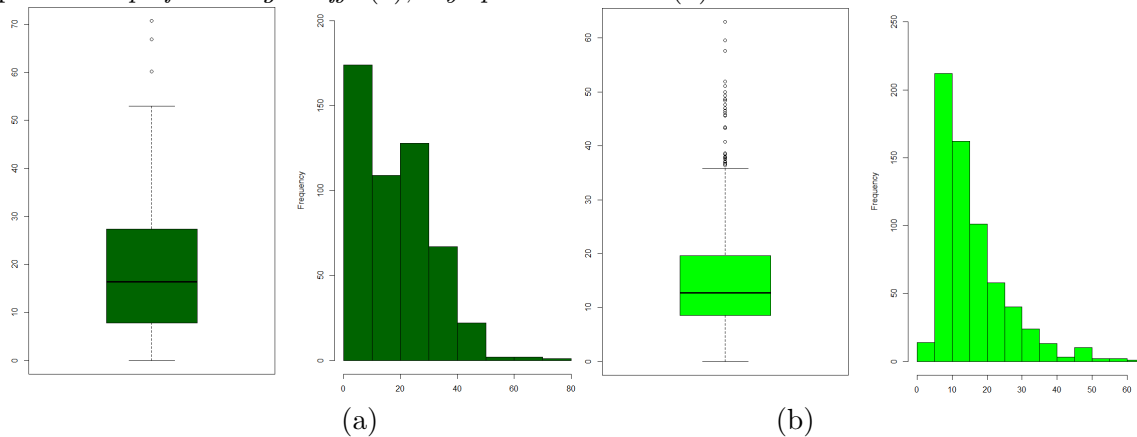
| Espécie | n | Média | Mediana | Desvio Padrão | CV | Amplitude |
|------------------------|-----|-------|---------|---------------|-------|-----------|
| <i>C. Langsdorffii</i> | 505 | 18,83 | 16,40 | 12,25 | 65,07 | 70,69 |
| <i>X. Brasiliensis</i> | 642 | 15,45 | 12,70 | 9,85 | 63,77 | 63,00 |

Fonte: Dos Autores (2023).

Pode-se observar que a variabilidade dos diâmetros da espécie *C. Langsdorffii* é superior a *X. Brasiliensis*. Observa-se ainda para ambas espécies que a média é superior a mediana, indicando uma tendência de assimetria a direita aos dados. Na prática, corresponde um indicativo de alta concentração de árvores jovens.

Pela Figura 2, observa-se que indivíduos que possuem valores menores de DAP, aparecem em uma proporção superior. Nota-se que o gráfico tende a uma forma exponencial negativa (J-invertida), indicando uma alta concentração de árvores jovens, tanto para *Copaifera langsdorffii*, quanto para *Xylopia brasiliensis*. Conforme autores como Schaaf (2006) e Dalmaso (2016), pode-se dizer que tais espécies em estudo estão em um processo de regeneração. Fatos confirmados por Scalon, Oliveira e Mello (2012) que afirmam que no passado, a região em estudo neste trabalho sofreu por uma série de perturbações, como incêndios e exploração em algumas áreas.

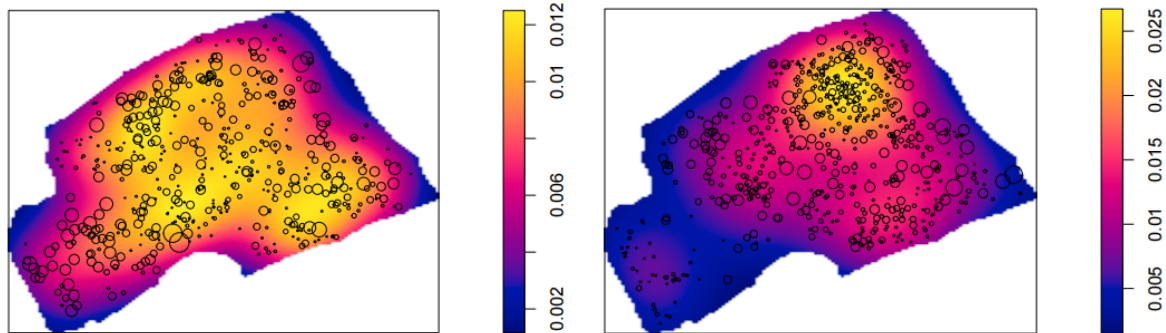
Figura 2: Histogramas de frequências e boxplots para distribuição dos DAPs em (cm) das espécies: *Copaifera langsdorffii*(a), *Xylopia brasiliensis*(b).



Fonte: Autores.

A análise de primeira ordem é o primeiro passo na análise espacial de um fenômeno aleatório (DIGGLE, 2013). Para visualização da variação completa de intensidade na área de estudo de cada uma das espécies, foi realizada a suavização de Kernel, conforme apresentado na Figura 3. Em análise, é possível observar que parece haver uma relação inversa de entre distâncias e diâmetros entre as espécies. Em *Copaifera langsdorffii*, Figura 3 (esquerda), os diâmetros de maior magnitude aparecem mais concentrados enquanto as árvores de menores diâmetros aparecem dispersas na área. Já para *Xylopia brasiliensis* pela Figura 3 (direita), quem aparece em concentração são os diâmetros menores. Tais observações poderão ser comprovadas ou refutadas perante análise de segunda ordem.

Figura 3: Distribuição espacial das estimativas de intensidades por meio do alisamento de kernel para *Copaifera langsdorffii* (esquerda), *Xylopia brasiliensis* (direita).



Fonte: Dos autores.

A definição dos raios de influência para construção dos mapas Kernel foram obtidos pela fórmula de Campbell de Cronie e Lieshout (2018) com estimativas apresentadas na Tabela 2.

Tabela 2: Tabela apresentando o raio de suavização de Kernel para cada uma das espécies.

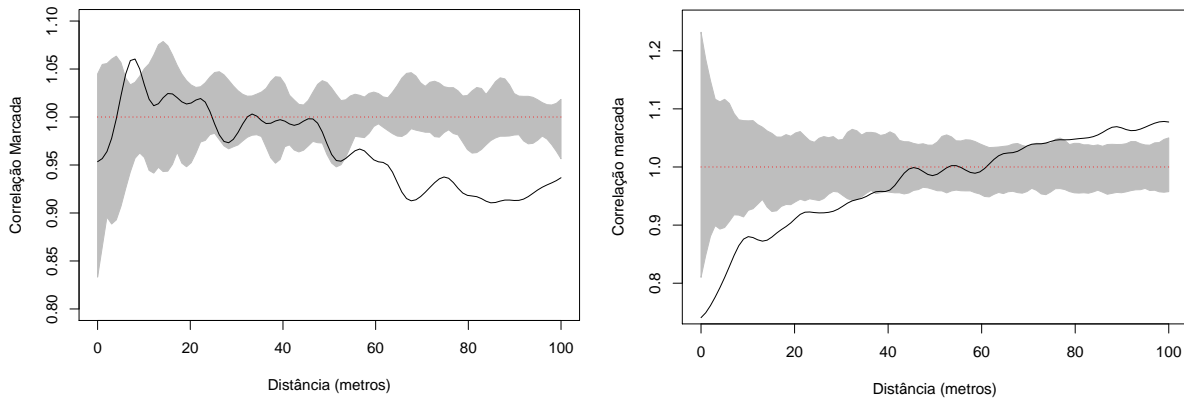
| Espécie | raio τ |
|-------------------------------|--------------|
| <i>Copaifera langsdorffii</i> | 25,63 metros |
| <i>Xylopia brasiliensis</i> | 30,83 metros |

Fonte: Do Autor (2023).

A análise da distribuição espacial do diâmetro à altura do peito das árvores foi conduzida

utilizando a função de correlação marcada $\rho_f(t)$, calculada para raios de 0 a 100 metros. Testou-se hipótese nula de independência entre as localizações das árvores (pontos) e DAPs (marcas). Os resultados são apresentados pela Figura 4 e apresentam evidências para rejeição da hipótese nula. Isso indica uma associação entre marcas e pontos, ou seja, descreve um certo tipo de correlação entre as marcas. Tais comportamentos foram contrários para ambas espécies.

Figura 4: Envelopes de simulações Monte Carlo da função de correlação marcada para *Copaifera langsdorffii* (esquerda), *Xylopia brasiliensis* (direita). As áreas cinzas indicam limites inferior e superior, a linha preta contínua indica a função estimada e a linha vermelha tracejada indica a função teórica.



Fonte: Autores.

O gráfico da Figura 4 (à direita) representa a estimativa da função de correlação marcada para a espécie *Xylopia brasiliensis*. Durante a análise, observa-se a presença de correlação espacial entre as marcas. A função $\rho_f(t)$ assume valores baixos (associação negativa) para raios t menores (0 a 40 metros), valores próximos de 1 (associação nula) para raios intermediários e valores altos (associação positiva) para raios maiores (acima de 80 metros). Esse padrão indica que árvores com diâmetros menores que a média tendem a estar próximas, e à medida que o DAP assume valores superiores à média, tendem a estar mais distantes. De acordo com Stoyan (2000), esse comportamento é comumente observado em análises de correlação de marcas.

Para a espécie *Copaifera langsdorffii* (Figura 4, à esquerda), foram observados resultados diferentes. Na análise, constata-se que a função $\rho_f(t)$ assume valores altos (associação positiva) para raios em torno de 10 metros e baixos (associação negativa) para escalas superiores a 55 metros. Esse comportamento indica que árvores com DAPs maiores que a média estão agrupadas em escalas de aproximadamente 10 metros. À medida que o DAP assume valores menores, as escalas aumentam consideravelmente.

As discrepâncias encontradas na análise gráfica de ambas as espécies podem estar relacionadas ao grau de heterogeneidade ou ao nível de agregação da espécie (DALMASO, 2016). Conforme Machado (2012) destaca, a tendência de agregação de árvores jovens e sua afinidade com árvores adultas são resultados diretos da limitada capacidade de dispersão, fazendo com que a maioria das sementes e frutos produzidos permaneça no solo próximo à árvore de origem.

Considerações Finais

Com base nos resultados obtidos neste estudo, conclui-se que a análise espacial permitiu identificar a dependência entre a localização das árvores e o DAP das mesmas. No entanto, observou-se que as duas espécies apresentam comportamentos distintos nessa dependência. À medida que o DAP das árvores da espécie *Copaifera langsdorffii* aumenta, a associação passa de positiva para negativa, ao passo que na espécie *Xylopia brasiliensis* ocorre o contrário.

É possível que as causas relacionadas à divergência de comportamento das espécies estejam associadas a características intraespecíficas, como dispersão de sementes, níveis de competição, distintos estágios de desenvolvimento e até mesmo ao nível de agregação da espécie. Apesar de essas hipóteses não terem sido verificadas, sugere-se que estudos futuros possam abordar essa perspectiva.

A aplicação bem-sucedida da função de Correlação Marcada demonstra a robustez do método e oferece recursos valiosos para estudos envolvendo a dinâmica espacial de uma floresta de vegetação nativa. Portanto, pode-se concluir que a aplicação da função de Correlação Marcada, aliada aos métodos de simulação Monte Carlo, representa uma ferramenta importante na análise de processos pontuais marcados por variáveis contínuas.

Agradecimentos

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG pela concessão da bolsa de doutorado ao primeiro autor.

References

- BADDELEY, A.; TURNER, R. Spatstat: an r package for analyzing spatial point patterns. *Journal of statistical software*, 2005. 1-42.
- CÂMARA, G.; DAVIS, C.; MONTEIRO, A. M. V. *Introdução à Ciência de geoinformação*. São José dos Campos: INPE, 2001.
- CRESSIE, N. *Statistics for Spatial Data*. Ames: John Wiley & Sons, 1993.
- CRONIE, O.; LIESHOUT, M. N. M. V. A non-model-based approach to bandwidth selection for kernel estimators of spatial intensity functions. *Biometrika*, 105(2), 2018. 455-462.
- DALMASO, C. A. et al. Interações espaciais intraespecíficas de *ocotea odorifera* na floresta nacional de irati. *FLORESTA*, v. 46, n. 1, p. 103-114, 2016.
- DIGGLE, P. A kernel method for smoothing point process data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 34, 1985. 138-147.
- DIGGLE, P. J. *Statistical analysis of spatial and spatio-temporal point patterns*. [S.l.]: CRC press, 2013.
- DIGGLE, P. J.; RIBEIRO, P. J.; CHRISTENSEN, O. F. An introduction to model-based geostatistics. *Spatial Statistics and Computational Methods*, 173, 2003. 43-86.
- ILLIAN, J. *Statistical analysis and modelling of spatial point patterns*. Nova York: *John Wiley & Sons*, v. 70, 2008.
- LOADER, C. *Local regression and likelihood*. *Springer Science & Business Media*, 2006.
- MACHADO, S. d. A. et al. Distribuição espacial de um fragmento de Floresta Ombrófila Mista Montana. [S.l.]: SciELO Brasil, 2012.
- MØLLER, J.; WAAGEPETERSEN, R. Some recent developments in statistics for spatial point patterns. *Annual Review of Statistics and Its Application*, *Annual Reviews*, v. 4, p.

317-342, 2017.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. 2023.

SCALON, J. D.; OLIVEIRA, C. A. P. d.; MELLO, J. M. d. Análise espacial de um fragmento florestal baseada no Mosaico de Dirichlet. *Revista Árvore, SciELO Brasil*, v. 36, p. 733-740, 2012.

SCHAAF, L. B. et al. Alteração na estrutura diamétrica de uma floresta ombrófila mista no período entre 1979 e 2000. *Revista árvore, SciELO Brasil*, v. 30, p. 283-295, 2006.

STOYAN, D.; PENTTINEN, A. *Recent applications of point process methods in forestry statistics*. *Statistical science, JSTOR*, p. 61-78, 2000.

TERRELL, G.; SCOTT, D. *Variable kernel density estimation*. *The Annals of Statistics*, 1992. 1236-1265.