



WÉLSON ANTÔNIO DE OLIVEIRA

**MÉTODOS DE PROCESSOS PONTUAIS PARA ANÁLISE
ESPACIAL DE ESPÉCIES ARBÓREAS NATIVAS DE UM
FRAGMENTO FLORESTAL DA REGIÃO DE LAVRAS-MG**

LAVRAS – MG

2022

WÉLSON ANTÔNIO DE OLIVEIRA

**MÉTODOS DE PROCESSOS PONTUAIS PARA ANÁLISE ESPACIAL DE ESPÉCIES
ARBÓREAS NATIVAS DE UM FRAGMENTO FLORESTAL DA REGIÃO DE
LAVRAS-MG**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

Prof. Dr. João Domingos Scalon
Orientador

**LAVRAS – MG
2022**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Oliveira, Wélson Antônio de.

Métodos de processos pontuais para análise espacial de espécies arbóreas nativas de um fragmento florestal da região de Lavras-MG / Wélson Antônio de Oliveira. - 2022.

77 p. : il.

Orientador(a): João Domingos Scalon.

Dissertação (mestrado acadêmico) - Universidade Federal de Lavras, 2022.

Bibliografia.

1. Configuração espacial. 2. Tendência espacial. 3. Método Monte Carlo. 4. Manejo Florestal. I. Scalon, João Domingos. II. Título.

WÉLSON ANTÔNIO DE OLIVEIRA

**MÉTODOS DE PROCESSOS PONTUAIS PARA ANÁLISE ESPACIAL DE ESPÉCIES
ARBÓREAS NATIVAS DE UM FRAGMENTO FLORESTAL DA REGIÃO DE
LAVRAS-MG
POINT PROCESS METHODS FOR SPATIAL ANALYSIS OF NATIVE TREE SPECIES
IN A FOREST FRAGMENT IN THE REGION OF LAVRAS-MG**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

APROVADA em 29 de setembro de 2022.

Prof. Dr. José Márcio de Mello UFLA
Prof. Dr. Ricardo Alves de Olinda UFPB
Prof. Dr. Renato Ribeiro de Lima UFLA

Prof. Dr. João Domingos Scalon
Orientador

**LAVRAS – MG
2022**

AGRADECIMENTOS

A Deus por permitir que grandes realizações tornasse possível.

Aos meus pais, José Adair e Izaura e meu irmão Weider pelo suporte e incentivo e apoio de sempre.

Ao professor João Domingos Scalon pela orientação na elaboração deste trabalho.

Aos meus colegas Miguel, Rafaela, Lucas e Natiele pela amizade.

Aos meus amigos de república Rodrigo, Gabriel, Fábio e Bruno pela parceria.

A todos professores da Pós-Graduação em Estatística e Experimentação Agropecuária, em especial aos Professores Renato Ribeiro e Tales Fernandes.

Aos integrantes do Grupo de Pesquisa em Estatística Espacial - GPS.

À Universidade Federal de Lavras - UFLA, pela oportunidade.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, pela concessão da bolsa de mestrado.

RESUMO

Muitos fenômenos aleatórios podem ser expressos por ocorrências identificadas por coordenadas localizadas no espaço. Esses fenômenos são denominados processos pontuais. Quando algum atributo (marca) é associado à coordenada, determina-se um processo pontual marcado. No manejo florestal de espécies nativas é de extrema importância caracterizar não somente a configuração espacial de cada espécie de árvores, mas também a interação existente entre diversas espécies. Essa caracterização pode auxiliar o pesquisador fornecendo-lhe informações que incluem competição, distribuição, crescimento, mortalidade e coexistência de espécies naquele espaço. Assim, o objetivo deste trabalho foi utilizar de métodos de processos pontuais para análise espacial de espécies arbóreas nativas de um fragmento de floresta Semidecidual Montana localizada na região de Lavras – MG. Para isso, serão utilizados alguns métodos de processos pontuais marcados e não marcados não homogêneos para analisar os efeitos de primeira e segunda ordem, permitindo uma análise de distribuição, interação e dependência espacial entre pontos (árvores) e marcas (DAP e espécie). Os efeitos de primeira ordem (ou efeitos globais) caracterizam o número esperado de ocorrências por unidade de área (tendência), enquanto os efeitos de segunda ordem (ou efeitos locais de pequena escala) caracterizam a estrutura de dependência espacial do processo. Pelos resultados é possível observar o potencial dos métodos para análise de processos pontuais espaciais para responder questões que auxiliem no manejo, controle e desenvolvimento de espécies arbóreas nativas.

Palavras-chave: Configuração espacial. Tendência espacial. Método Monte Carlo. Manejo Florestal.

ABSTRACT

Many random phenomena can be expressed by occurrences identified by coordinates located in space. These phenomena are called point processes. When some attribute (mark) is associated with the coordinate, a marked point process is determined. In forest management of native species, it is extremely important to characterize not only the spatial configuration of each tree species, but also the interaction between different species. This characterization can help the researcher by providing information that includes competition, distribution, growth, mortality and coexistence of species in that space. Thus, the objective of this work was to use point process methods for spatial analysis of native tree species from a fragment of Montana semideciduous forest located in the region of Lavras - MG. For this, some methods of non-homogeneous marked and unmarked point processes will be used to analyze the first and second order effects, allowing an analysis of distribution, interaction and spatial dependence between points (trees) and marks (DAP and species). First-order effects (or global effects) characterize the expected number of occurrences per unit area (trend), while second-order effects (or small-scale local effects) characterize the spatial dependence structure of the process. Based on the results, it is possible to observe the potential of methods for analyzing punctual spatial processes to answer questions that help in the management, control and development of native tree species.

Keywords: Spatial configuration. Spatial trend. Monte Carlo Method. Forest Management.

LISTA DE FIGURAS

Figura 3.1 – Padrões Espaciais: Regular (esquerda), Aleatório (centro) e Agregado (direita).	17
Figura 3.2 – Representação gráfica de um processo pontual não homogêneo, com intensidade variando nas distintas localidades da região.	22
Figura 3.3 – Estimador de intensidade pontual do processo em toda região de estudo (Kernel)	24
Figura 3.4 – Região de estudo apresentada pelos pontos pretos, e demais ocorrências (pontos brancos) que poderiam influenciar a análise mas que não estão sendo considerados por estar fora da área de estudo.	26
Figura 3.5 – Representação gráfica da função $K(t)$ de Ripley	30
Figura 3.6 – Representação gráfica de um processo pontual marcado com as marcas qualitativas definidas pela diferença (com preenchimento e sem preenchimento) do ponto.	35
Figura 3.7 – Representação gráfica de um processo pontual marcado com as marcas quantitativas definidas pela diferença de tamanho do ponto.	41
Figura 4.1 – Fragmento floresta Semidecidual Montada ("Matinha").	47
Figura 4.2 – <i>Copaifera langsdorffii</i>	49
Figura 4.3 – <i>Xylopia brasiliensis</i>	50
Figura 5.1 – Distribuição espacial e classificação das árvores quanto a espécie.	54
Figura 5.2 – Histogramas com frequências de DAP em (cm) para as espécies: <i>Copaifera langsdorffii</i> (a), <i>Xylopia brasiliensis</i> (b).	55
Figura 5.3 – Distribuição espacial das estimativas de intensidades por meio do alisamento de kernel para <i>Copaifera langsdorffii</i> (esquerda), <i>Xylopia brasiliensis</i> (direita).	56
Figura 5.4 – Envelopes de simulações Monte Carlo das funções univariadas não homogêneas $L_{inhom}(t)$ (esquerda) e $J_{inhom}(t)$ (direita) para <i>Copaifera langsdorffii</i> . 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.	57
Figura 5.5 – Envelopes de simulações Monte Carlo das funções univariadas não homogêneas $L_{inhom}(t)$ (esquerda) e $J_{inhom}(t)$ (direita) para <i>Xylopia brasiliensis</i> . 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.	58

Figura 5.6 – Envelopes de simulações Monte Carlo das funções $\hat{J}_{12}^{inhom}(t)$ (esquerda) e $\hat{L}_{12}^{inhom}(t)$ (direita) não homogênea bivariadas para interação *Copaifera langsdorffii* e *Xylopiã brasiliensis*. 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. 59

Figura 5.7 – Envelopes de simulações Monte Carlo da função de correlação marcada para *Copaifera langsdorffii* (esquerda), *Xylopiã 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. 61

LISTA DE TABELAS

- Tabela 5.1 – Tabela descritiva de diâmetro à altura do peito (cm) por espécie de árvore. . . . 55
- Tabela 5.2 – Tabela apresentando o raio de suavização de Kernel para cada uma das espécies. 56

SUMÁRIO

1	INTRODUÇÃO	12
2	OBJETIVOS	14
2.1	Objetivo Geral	14
2.2	Objetivos Específicos	14
3	REFERENCIAL TEÓRICO	15
3.1	Estatística Espacial	15
3.2	Processos Pontuais	16
3.2.1	Estacionariedade e Isotropia	18
3.2.2	Efeitos de Primeira e segunda ordem	19
3.2.3	Processos pontuais homogêneos	20
3.2.4	Processos pontuais não homogêneos	21
3.2.5	Estimador de Intensidade	22
3.2.6	Estimadores de Dependência Espacial	25
3.2.6.1	Função $J(t)$ homogênea	25
3.2.6.2	Função $J_{inhom}(t)$ não homogênea	28
3.2.6.3	Função $L(t)$ homogênea	29
3.2.6.4	Função $L_{inhom}(t)$ não homogênea	32
3.3	Processos pontuais marcados	33
3.3.1	Marcas Qualitativas	34
3.3.1.1	Efeitos de primeira e segunda ordem (caso bivariado)	35
3.3.1.2	Função $J_{12}(t)$ bivariada	36
3.3.1.3	Função $J_{12}^{inhom}(t)$ bivariada não homogênea	37
3.3.1.4	Função $L_{12}(t)$ bivariada	39
3.3.1.5	Função $L_{12}^{inhom}(t)$ bivariada não homogênea	40
3.3.2	Marcas Quantitativas	41
3.3.2.1	Função de Correlação Marcada	41
3.4	Teste de Monte Carlo	42
3.5	Processos pontuais em ciências florestais	44
4	MATERIAIS E MÉTODOS	47

4.1	Dados	47
4.2	<i>Copaifera langsdorffii</i>	48
4.3	<i>Xylopia brasiliensis</i>	48
4.4	Métodos Estatísticos	49
4.4.1	Análise Exploratória	49
4.4.2	Análise de primeira ordem	50
4.4.3	Análise de segunda ordem (Univariada)	50
4.4.4	Análise de segunda ordem (Bivariada)	51
4.4.5	Análise de segunda ordem (Marcas Contínuas)	51
4.4.6	Recursos Computacionais	52
5	RESULTADOS E DISCUSSÕES	53
5.1	Análise Exploratória	53
5.2	Análise de primeira ordem	54
5.3	Análise de segunda ordem caso não-homogêneo (Unviaridade)	56
5.4	Análise de segunda ordem caso não homogêneo (Bivaridade)	59
5.5	Análise de segunda ordem (marcas contínuas)	60
6	CONCLUSÕES	63
	REFERÊNCIAS BIBLIOGRÁFICAS	67
A	ANEXOS - Códigos em R	68
A.1	Leitura dos dados	68
A.2	Códigos da análise de primeira ordem	69
A.3	Códigos da análise de segunda ordem (univariado)	70
A.4	Códigos da análise de segunda ordem (bivariado)	71
A.5	Códigos da análise de Correlação Marcada	75

Muitos fenômenos aleatórios de origem natural podem ser expressos por eventos e identificadas como pontos. Exemplos incluem casos de doenças, ocorrências de crimes, acidentes, localização de espécies vegetais e entre outras. Esses fenômenos, portanto, podem ser representados e estudados através da teoria de processos pontuais. Um processo pontual espacial pode ser entendido como um conjunto de pontos irregularmente distribuídos em uma área definida, cuja estrutura espacial, foi gerada por um mecanismo estocástico. O interesse da análise de processos pontuais está em analisar a intensidade de ocorrências dos eventos e determinar o padrão de distribuição espacial dos pontos na área de estudo.

A análise consiste na estimativa dos efeitos de primeira e segunda ordem. Quando ambos se ausentam em um padrão de pontos, determina-se um processo de completa aleatoriedade espacial (CAE). Na prática, se tratando de fenômenos naturais, um padrão de CAE é pouco detectado pois, é comum a presença de uma certa heterogeneidade na intensidade de ocorrências ao longo da região em estudo bem como a presença de uma certa dependência espacial, sendo necessário utilizar de recursos apropriados para se analisar como é o caso dos métodos de processos pontuais. A utilização desses recursos é justificado como sendo uma forma conveniente para se tirar conclusões coerentes e confiáveis acerca do fenômenos em estudo, que não podem ser feitos de forma visual. Além disso, a disposição de pacotes em R Core Team viabiliza a realização das análises.

Quando se tem informações, atributos ou características vinculadas aos pontos, são chamadas de marcas. Elas podem ser classificadas como qualitativas ou quantitativas e quando acrescentadas nas análises determinam um processo pontual marcado. Na análise de processos pontuais marcados, o foco está em investigar sobre a existência de dependência entre marcas e pontos. Tais possibilidades de análises é interessante em diversas áreas do conhecimento, tais como: Epidemiologia, Sociologia, Biologia, Demografia e entre outras. Em Ciências Florestais, grande parte das aplicações estão na análise da distribuição e interação espacial de espécies vegetais nativas. Ecologistas têm estudado e se interessam pelo padrão espacial de espécies vegetais para inferir sobre a existência de processos subjacentes que explicam o padrão observado.

Conhecer o padrão de distribuição e interação espacial de espécies vegetais nativas de algum fragmento florestal ou área de conservação, torna-se de grande importância para o estudo das espécies. Isso pois, a configuração espacial de árvores afeta um grande número de processos fi-

siológicos e ecológicos na vegetação, incluindo existência, coexistência, competição, distribuição, tamanho, crescimento, mortalidade e heterogeneidade ambiental. Assim, trabalhos que se propõem levantar discussões acerca de análise espacial de espécies arbóreas nativas contribuem no entendimento não só do comportamento, mas também na tomada de decisão sobre o manejo a ser utilizado nas áreas de conservação de remanescentes florestais. Dessa forma, o presente trabalho teve por objetivo utilizar métodos de processos pontuais para análise espacial de espécies arbóreas nativas de um fragmento florestal da região de Lavras-MG.

2 OBJETIVOS

2.1 Objetivo Geral

Utilizar de métodos de processos pontuais para análise espacial de espécies arbóreas nativas localizadas em um fragmento florestal da região de Lavras – MG.

2.2 Objetivos Específicos

Os objetivos específicos são:

- a) caracterizar o padrão de distribuição espacial (regular, aleatório, agrupado) dos indivíduos de cada uma das espécies arbóreas em estudo (*Copaifera langsdorffii* e *Xylopia brasiliensis*);
- b) investigar a interação espacial entre as espécies arbóreas nativas (*Copaifera langsdorffii* vs *Xylopia brasiliensis*);
- c) Caracterizar a dependência entre o diâmetro à altura do peito (DAP) e a localização das árvores para cada uma das duas espécies (*Copaifera langsdorffii* e *Xylopia brasiliensis*).

3 REFERENCIAL TEÓRICO

3.1 Estatística Espacial

Na abordagem tradicional de tratamento e análise de dados, como a análises de variância, modelos lineares entre outros, apesar de consolidados e bastante utilizados em diversas áreas, não permitem considerar o espaço (localização) das observações. Entretanto, em diversos ambientes de pesquisa é comum instigar sobre as relações de dependência entre ocorrências próximas. Waldo Tolber em 1970 apud Monteiro et al. (2004, p.33) expõe uma ideia fundamental na introdução da estatística espacial, dizendo que "*Todas as coisas são parecidas, mas coisas próximas se parecem mais que coisas distantes*".

A estatística espacial é composta por um conjunto de procedimentos envolvendo a parte exploratória (apresentação, visualização e identificação de padrões espaciais) e inferencial (ajuste de modelo que explicam os padrões espaciais) fazendo, explicitamente, o uso da localização do fenômeno estudado (MONTEIRO et al., 2004). A ênfase para a análise espacial está em suas técnicas que dão destaque e enfoque para o arranjo espacial dos objetos ou fenômenos que estão sendo analisados. (GATRELL; BAILEY, 1996).

Para Cressie (1993), os dados espaciais podem ser divididos da seguinte forma:

- a) *Dados de Superfícies Contínuas*: Obtidos a partir de fenômenos ou regiões contínuas e geralmente adquiridos de amostras coletadas sob levantamento de recursos naturais. Exemplos: Dados obtidos no campo para a criação de mapas topográficos, estudo da cama de um sistema de confinamento de gado compost-barn, etc (YAMAMOTO; LANDIM, 2015). Na análise deste tipo de dados, observa-se pontos de uma superfície contínua com objetivo de reconstruir a região, inferindo sobre as regiões não amostradas. O modelo inferencial utiliza-se de técnicas geoestatísticas cuja hipótese central é o conceito de estacionariedade. Métodos como a krigagem são devidamente utilizados (DIGGLE et al., 2003);
- b) *Dados de Áreas Agregadas*: São dados identificados geralmente por localização pontual e obtidos de forma agrupada. Envolve levantamentos populacionais, cuja localização está associada a áreas delimitadas por polígonos fechados. Exemplos incluem casos de uma determinada doença, nascimentos de crianças e entre outros, organizados por estados, municípios, bairros e afins. (MONTEIRO et al., 2004). Para a análise de dados de área, além

da parte exploratória utilizada através de mapas, é feita a escolha de modelos inferenciais. A hipótese testada é de homogeneidade entre os polígonos e que a distribuição de probabilidade seja distinta para cada área. São utilizados frequentemente em dados de saúde pública, economia, sociologia entre outros (CRESSIE, 1993).

c) *Dados de Padrões Pontuais*: São dados obtidos a partir de fenômenos expressos através de pontos. O interesse está necessariamente na localização da ocorrência dos eventos. As análises consistem na caracterização da distribuição espacial com base na estimação dos efeitos de primeira ordem ou determinação da intensidade por unidade de área e segunda ordem, pela caracterização do padrão espacial de distribuição dos pontos (DIGGLE, 2013);

O presente trabalho, está centrado em métodos de análise para dados pontuais, com isso, conceitos de processos pontuais serão apresentados e detalhados nos próximos capítulos.

3.2 Processos Pontuais

Muitos dos fenômenos naturais aleatórios de ocorrências motivadas ou não pela ação humana podem ser representados através de pontos. Exemplos incluem casos de doenças, ocorrências de crimes, acidentes, localização de espécies vegetais e entre outros. Estes, podem estar situados em uma reta (caso unidimensional), em uma região do plano (caso bidimensional), em superfícies (caso tridimensional) ou até em espaços multidimensionais. Este trabalho trata particularmente do caso bidimensional, envolvendo teoria e análises de *Processos Pontuais*.

Um processo pontual é obtido através de uma coleção de variáveis aleatórias definidas por

$$X = \{X(\mathbf{s}); \mathbf{s} \in S\}, \quad (3.1)$$

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, recebendo 1 com a ocorrência do evento em \mathbf{s} e 0 caso contrário.

Conforme Cressie (1993), o conjunto X poderia ser qualquer espaço Hausdorff¹ localmente compacto cuja topologia tenha uma base contável. Em geral, as observações serão obtidas apenas de uma região limitada $A \subset S$, contendo um conjunto finito de eventos (BADDELEY et al., 2008).

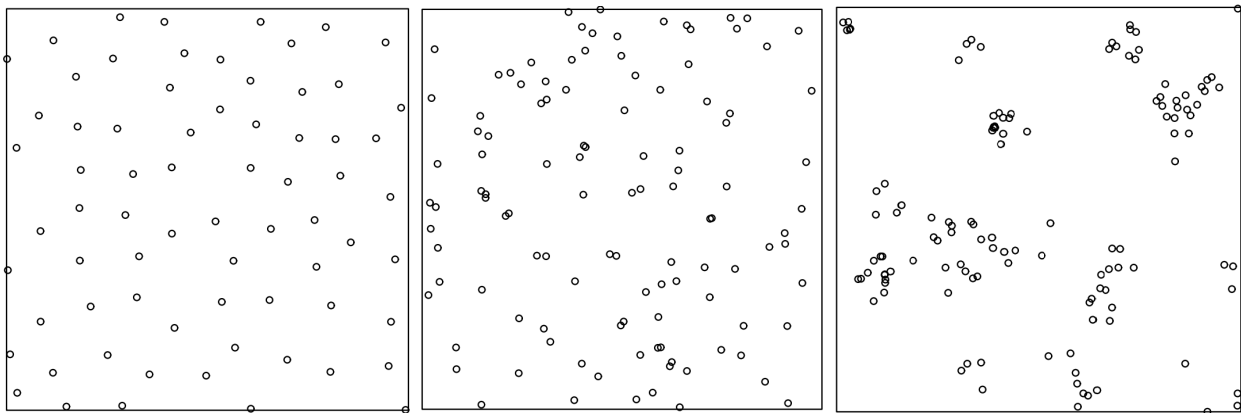
¹ Um espaço de Hausdorff é um espaço topológico X com a propriedade de que para todo $x, y \in X$ com $x \neq y$, existem conjuntos abertos A e B , tais que, $x \in A$ e $y \in B$ e $A \cap B = \emptyset$.

As análises de processos pontuais possuem as seguintes características:

- i. O interesse está exclusivamente na localização pontual da ocorrência do evento.
- ii. Não se leva em consideração a área do evento. Mesmo que estes ocupem espaço, como em localizações de cidades ou árvores densas de uma floresta, considera-se unicamente um ponto podendo ser central ou localizado de maneira conveniente.
- iii. 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.

O objetivo da análise de processos pontuais está no estudo do padrão de distribuição espacial dos pontos, considerando basicamente três padrões referência: *Aleatório* caracterizados por serem distribuídos totalmente ao acaso no espaço, espalhados sem nenhum tipo de regularidade. *Regular* no qual é possível prever uma distância média entre os pontos e *Agregado* com a presença de agrupamentos locais ou aglomeração de pontos, como mostra a Figura 3.1.

Figura 3.1 – Padrões Espaciais: Regular (esquerda), Aleatório (centro) e Agregado (direita).



Fonte: Baddeley et al. (2008, p.7)

Os padrões de distribuição pontual possuem características específicas com relação a média e variância. Padrões regulares apresentam variância menor do que a média, indicando subdispersão. Padrões aleatórios possuem variância equivalente a média, indicando aleatoriedade. E padrões agregados possuem variância maior que a média, indicando superdispersão dos pontos (CRESSIE, 1993).

3.2.1 Estacionariedade e Isotropia

O termo estacionariedade é usado para se referir a processos que possuem propriedades similares em todas as localizações na região de estudo. Incluem efeitos de média e variância constantes e covariância entre dois pontos dependendo apenas da localização relativa entre eles (OLINDA; SCALON, 2010). Elevando a formalidade, entende-se que um processo pontual é estacionário quando suas propriedades são invariantes sob translação (DIGGLE et al., 2003). Dado um processo pontual X e um outro processo X_d transladado por d , em Illian et al. (2008) tem-se,

$$X \equiv X_d, \quad (3.2)$$

sendo $X = \{x_1, x_2, \dots\}$ e $X_d = \{x_1 + d, x_2 + d, \dots\}$.

Considere uma região de estudo A subdividida em regiões $B_i; i = 1, \dots, n$. A contagem dos eventos em uma região de estudo é determinada pela realização de $X(x \in B_i) = N(B_i)$. Um processo é estacionário se e somente se, para qualquer inteiro n e regiões $B_i : i = 1, \dots, n$, temos que

$$\begin{aligned} P(N(B_1) = n_1, \dots, N(B_k) = n_k) &= P(N_d(B_1) = n_1, \dots, N_d(B_k) = n_k) \\ &= P(N(B_1 - d) = n_1, \dots, N(B_k - d) = n_k) \end{aligned} \quad (3.3)$$

ou seja, a distribuição conjunta não varia sob a translação ou deslocamento d .

O conceito de isotropia está condicionado a invariância com base na rotação ao invés da translação conforme Illian et al. (2008). As descrições das rotações é dada por um ângulo α entre 0° e 360° . Sendo $\mathbf{s} = (x, y)$, então as coordenadas do ponto transladado são dadas por

$$\mathbf{s}_\alpha = (x_\alpha, y_\alpha) = (x \cos(\alpha) + y \sin(\alpha), -x \sin(\alpha) + y \cos(\alpha)) \quad (3.4)$$

Com isso, tem-se que um processo pontual estacionário se a distribuição conjunta é invariante sob rotação por meio de um ângulo α arbitrário.

3.2.2 Efeitos de Primeira e segunda ordem

A caracterização do processo estocástico pode ser explicado através da análise dos efeitos de primeira e segunda ordem. Efeitos de primeira ordem são considerados globais ou de larga escala e correspondem a variações do valor médio do processo estocástico no espaço (CRESSIE, 1993). Neste caso, o interesse está em determinar a intensidade ou o número esperado de eventos por unidade de área (MONTEIRO et al., 2004).

Os efeitos de primeira ordem podem ser definidos pela função de intensidade $\lambda(\mathbf{s})$, dada pela equação (3.5)

$$\lambda(\mathbf{s}) = \lim_{|d\mathbf{s}| \rightarrow 0} \left\{ \frac{\mathbb{E}[N(d\mathbf{s})]}{|d\mathbf{s}|} \right\}, \quad (3.5)$$

sendo $d\mathbf{s}$ uma pequena região em torno de \mathbf{s} , e $N(d\mathbf{s})$ representa a quantidade de eventos localizados em $d\mathbf{s}$. Para um processo pontual estacionário, a medida da intensidade não varia sob translação, tem-se $\lambda(\mathbf{s}) = \lambda$ constante (DIGGLE et al., 2003).

Os efeitos de segunda ordem são considerados locais ou de pequena escala. Definem a estrutura espacial (dependência espacial) ou a relação entre a localização dos eventos do processo em diferentes regiões do espaço (CRESSIE, 1993).

Conforme Diggle et al. (2003), as propriedades de segunda ordem podem ser definidas da mesma forma que os efeitos de primeira ordem considerando a intensidade conjunta $\lambda(\mathbf{s}_i, \mathbf{s}_j)$ entre duas regiões infinitesimais $|d\mathbf{s}_i|$ e $|d\mathbf{s}_j|$ contendo os pontos \mathbf{s}_i e \mathbf{s}_j como na equação (3.6):

$$\lambda_2(\mathbf{s}_i, \mathbf{s}_j) = \lim_{|d\mathbf{s}_i|, |d\mathbf{s}_j| \rightarrow 0} \left\{ \frac{\mathbb{E}[N(d\mathbf{s}_i) \times N(d\mathbf{s}_j)]}{|d\mathbf{s}_i| |d\mathbf{s}_j|} \right\}, \quad \mathbf{s}_i, \mathbf{s}_j \in S. \quad (3.6)$$

Quando o processo é estacionário e também isotrópico, $\lambda(d\mathbf{s}_i, d\mathbf{s}_j) = \lambda(|h|) = h$, o efeito de segunda ordem ou a dependência espacial do fenômeno se reduz na distância entre os eventos. (MONTEIRO et al., 2004).

Algumas medidas podem ser utilizadas para relacionar as intensidades de primeira e segunda ordem, no qual duas são apresentadas. A função de intensidade condicional que corresponde a intensidade no ponto \mathbf{s}_i condicionada a informação de existência do ponto \mathbf{s}_j , é dada pela Equação

(3.7)

$$\lambda_c(\mathbf{s}_i|\mathbf{s}_j) = \frac{\lambda_2(\mathbf{s}_i, \mathbf{s}_j)}{\lambda(\mathbf{s}_j)}. \quad (3.7)$$

A função de correlação de pares é outra medida que relaciona as intensidades de primeira e segunda ordem, definida por Cressie (1993) pela Equação (3.8)

$$g(\mathbf{s}_i|\mathbf{s}_j) = \frac{\lambda_2(\mathbf{s}_i, \mathbf{s}_j)}{\lambda(\mathbf{s}_j) \times \lambda(\mathbf{s}_j)}. \quad (3.8)$$

As propriedades de estacionariedade e isotropia, estão diretamente ligados com a intensidade do processo pontual. Diggle (2013) expõe que um processo estacionário, $\lambda_2(\mathbf{s}_i, \mathbf{s}_j) \equiv \lambda_2(\mathbf{s}_i - \mathbf{s}_j)$ é invariante sobre translação em \mathbb{R}^d e se um processo estacionário é também isotrópico, sua distribuição é invariante sobre rotações na origem de \mathbb{R}^d , com isso, $\lambda_2(\mathbf{s}_i - \mathbf{s}_j)$ se reduz a $\lambda_2(t)$, onde $t = \|\mathbf{s}_i - \mathbf{s}_j\|$ equivale a distância euclidiana entre \mathbf{s}_i e \mathbf{s}_j .

3.2.3 Processos pontuais homogêneos

Um processo pontual é homogêneo quando detém da propriedade de estacionariedade não havendo, pois, efeitos de primeira ordem. Pode ser caracterizado pelas seguintes propriedades (DIGGLE, 2013):

- i. O número de pontos em qualquer conjunto limitado B_i segue uma distribuição de Poisson com média $\lambda|B_i|$.
- ii. Os n pontos de A formam um evento aleatório independente em A com função densidade de probabilidade proporcional à intensidade $\lambda(\mathbf{s})$.

Processos que também não apresentam efeitos de segunda ordem são denominados de completa aleatoriedade espacial (CAE). Os modelos teóricos de CAE são os mais simples correspondendo uma mesma probabilidade de ocorrência de algum evento dentro da área de estudo (CRESSIE, 1993). Isto é, dividindo a área de estudo A em sub-regiões B_i com $i = 1, \dots, n$, no processo de CAE, a cada $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n\}$, tem-se

$$P[\mathbf{s}_1 \in B_1, \mathbf{s}_2 \in B_2, \dots, \mathbf{s}_n \in B_n] = \prod_{i=1}^n \left[\frac{|B_i|}{|A|} \right], \quad (3.9)$$

sendo $|B_i| \equiv \int_{B_i} d\mathbf{s}$ e $|A|$ a área de A .

Em processos de CAE, os eventos em cada sub-região são não correlacionados e estão associados a uma mesma distribuição de probabilidade de Poisson. Em outras palavras, a posição de cada evento é independente e os eventos têm a mesma probabilidade de ocorrência em toda a região A .

Illian et al. (2008) caracteriza um processo pontual de CAE de acordo com as seguintes propriedades:

- i Um processo de CAE é um processo pontual homogêneo, ou seja, para qualquer região A de um plano, $N(A)$ segue uma distribuição de Poisson com média $\lambda|A| > 0$.
- ii Em um processo de CAE, tem-se que os n eventos formam um evento aleatórios e independentes em A .
- iii Dado um processo de CAE, tem-se que o número de pontos $N(B_i)$ e $N(B_j)$ são independentes para quaisquer regiões B_i e B_j com $i \neq j$ e $B_i \cap B_j = \emptyset$.

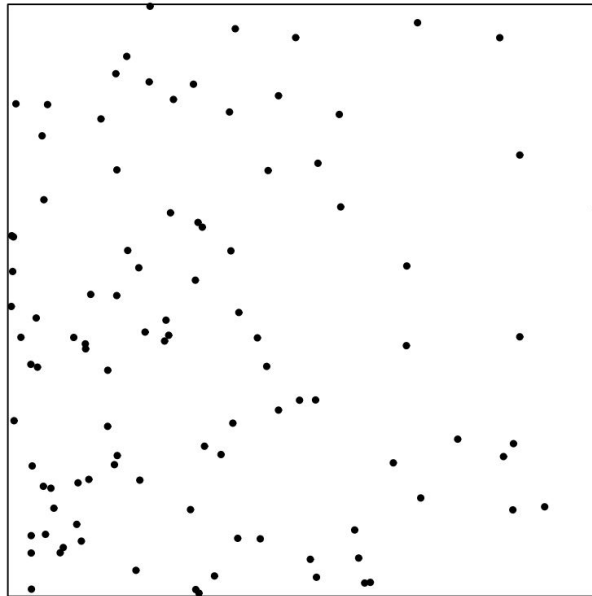
Conforme Diggle (2013), pela propriedade (i), a intensidade dos eventos de um processo de CAE não varia ao longo do plano. Pelas propriedades (ii) e (iii), observa-se que em um processo de CAE, não há interações entre os eventos. Por exemplo, a suposição de independência seria violada se a existência de um evento encorajasse ou inibisse a ocorrência de outros eventos em sua vizinhança.

3.2.4 Processos pontuais não homogêneos

De modo geral, é bastante conveniente considerar a estacionariedade ou homogeneidade de um processo pontual. Contudo, tal consideração é inviável diante de muitos fenômenos, exigindo de recursos mais sofisticados para a análise. O processo homogêneo pode ser generalizado de maneira direta pela introdução de heterogeneidade nas análises, o que prevê também, da utilização de processos de Poisson não homogêneos (ILLIAN et al., 2008). Conforme Diggle (2013), a intensidade constante λ do processo de Poisson é substituído por uma função de intensidade espacialmente variável $\lambda(x)$ que define a classe de processos de Poisson não homogêneos, para os quais:

- i. O número de pontos em qualquer conjunto limitado B_i segue uma distribuição de Poisson com média $\int_{B_i} \lambda(x) dx$.
- ii. Os n pontos de A formam um evento aleatório independente em A com função densidade de probabilidade proporcional a intensidade $\lambda(x)$.

Figura 3.2 – Representação gráfica de um processo pontual não homogêneo, com intensidade variando nas distintas localidades da região.



Fonte: Diggle (2013, p.105)

O processo de poisson não homogêneo fornece uma estrutura possível para a introdução de covariáveis ou fatores na análise de padrões de pontos espaciais. Isso pode ser possível por meio de uma função de intensidade dada por $\lambda(x) = \lambda\{z_1(x), z_2(x), \dots, z_p(x)\}$ sendo $z_i(x)$ com $i = 1, \dots, p$ cada um dos fatores. Na prática, a inclusão de heterogeneidade da área em estudo está relacionado com as variáveis aleatórias adicionais, como por exemplo a variação de altitude, quantidade nutrientes, recursos hídricos, etc. Isso permite analisar se tais fatores podem ser determinísticos e motivadores do padrão identificado (DIGGLE, 2013).

3.2.5 Estimador de Intensidade

A análise descritiva de dados pontuais se inicia com a caracterização da distribuição espacial dos pontos ou eventos do processo através da estimação das propriedades de primeira e

segunda ordem (CRESSIE, 1993). Considerando um processo pontual homogêneo, o estimador de intensidade equivale ao número médio de eventos observados conforme a Equação (3.10):

$$\hat{\lambda} = \frac{n}{|A|}, \quad (3.10)$$

correspondendo a razão entre o número observado de eventos por unidades de área.

Havendo indícios de heterogeneidade no processo, ou interesse em analisar as intensidades em cada localidade, deve-se estimar a função intensidade. Para isso pode-se utilizar de métodos paramétricos ou não-paramétricos.

"Ao utilizar de modelos paramétricos para estimar a função de intensidade, os parâmetros desse modelo devem ser estimados. O caso mais simples é um tanto trivial de um padrão de ponto uniforme, que leva ao processo binomial. Exemplos não triviais incluem padrões de pontos cuja função de intensidade mostra uma tendência, que podem ser determinadas por regressão" (ILLIAN et al., 2008, p.111).

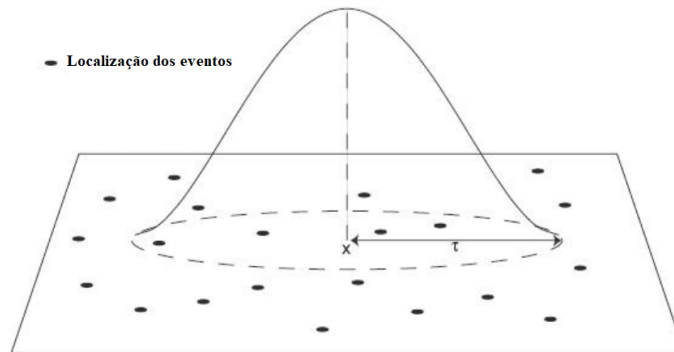
Frequentemente, utiliza-se de modelos não paramétricos para estimativa de $\hat{\lambda}(x)$. Neste trabalho, o enfoque na estimativa da função intensidade está na utilização de um desses modelos, baseados especificamente nos métodos de kernel com suavização gaussiana. Outros bastante utilizados são a uniforme, quártica, triangular, Epanechnikov, normal, entre outras. A suavização do kernel fornece uma maneira simples de encontrar estrutura nos dados conjuntos sem a imposição de um modelo paramétrico, dependendo unicamente do raio de τ , parâmetro que define o nível de suavização da superfície gerada sobre o fenômeno (KIM; SCOTT, 2012). O interesse está no ajuste de uma função bidimensional sobre a área de estudo, criando uma superfície cujo valor será proporcional à intensidade do processo e suavizada de acordo com ponderação pela distância (WAND; JONES, 1994) como ilustra a Figura 3.3:

Considere $\mathbf{s}_1, \dots, \mathbf{s}_n$ localizações de n eventos de uma região A . Por Terrell 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 (3.11),

$$\hat{\lambda}_\tau(\mathbf{x}) = \frac{1}{\tau^2} \sum_{i=1}^n k\left(\frac{d(\mathbf{s}_i, \mathbf{x})}{\tau}\right), \quad d(\mathbf{s}_i, \mathbf{x}) \leq \tau. \quad (3.11)$$

sendo que τ representa um raio de influência ($\tau \geq 0$) e $k(\cdot)$ é uma função de estimação com propriedades de suavização do fenômeno.

Figura 3.3 – Estimador de intensidade pontual do processo em toda região de estudo (Kernel)



Fonte: Olinda e Scalon (2010)

Conforme apresenta Monteiro et al. (2004), o raio de suavização τ determina quais dos n eventos serão utilizados na estimativa da intensidade de ocorrência na posição de \mathbf{x} . Para a função de interpolação $k(\cdot)$, é comum usar funções de terceira ou quarta ordem, como apresentada na Equação (3.12):

$$k(h) = \frac{3}{\pi}(1 - h^2), \quad (3.12)$$

sendo h a distância euclidiana entre a localização. Das Equações (3.11) e (3.12), tem-se que o estimador de intensidade para se obter uma visão geral da distribuição de primeira ordem dos eventos, pode ser expresso pela equação (3.13):

$$\hat{\lambda}_{\tau}(\mathbf{x}) = \sum_{h_i \leq \tau} \frac{3}{\pi \tau^2} \left(1 - \frac{h_i^2}{\tau^2}\right)^2. \quad (3.13)$$

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 (TERRELL; SCOTT, 1992), minimização do erro quadrático médio (DIGGLE, 1985), maximização do critério de validação cruzada de verossimilhança do processo de ponto (LOADER, 2006) e entre outros. Este trabalho, em particular, utilizará a definição de τ baseado na fórmula de Campbell, que relaciona médias de padrões a médias espaciais ponderadas por função de intensidade (CRONIE; LIESHOUT, 2018). Para os mesmos autores, além da visualização completa de intensidades, esta medida torna a estimativa da função confiável para correção de heterogeneidade espacial. Tal característica torna-se muito

importante para este estudo visto que, não se teve informações disponíveis sobre covariáveis que poderiam estar presentes na área em estudo.

3.2.6 Estimadores de Dependência Espacial

As técnicas mais utilizadas para caracterizar a dependência espacial ou as propriedades de segunda ordem do fenômeno, são baseadas na estimação de algumas funções descritoras. Para o caso homogêneo, utiliza-se frequentemente das funções $J(t)$ e $L(t)$. Para o caso não homogêneo (não estacionário), pode-se utilizar das funções $J_{inhom}(t)$ e $L_{inhom}(t)$ não homogêneas.

3.2.6.1 Função $J(t)$ homogênea

Para se obter a função $J(t)$, será necessário apresentar as funções $F(t)$ e $G(t)$. A função $F(t)$ corresponde a distribuição de probabilidade acumulada da distância entre um ponto escolhido aleatoriamente e o evento mais próximo. Seu estimador mais simples, pode ser obtido pela Equação (3.14):

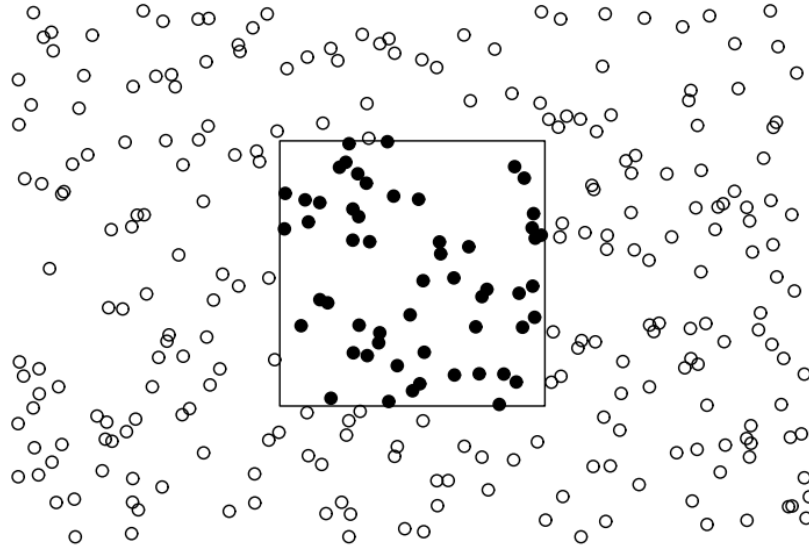
$$\hat{F}(t) = \lambda^{-1} \sum_{i=1}^m I_t(t_i), \quad (3.14)$$

sendo m o número de pontos arbitrários aleatórios gerados na região, t_i representa a distância do i -ésimo ponto aleatório para o mais próximo dos n eventos no mapa analisado e $I_t(t_i)$ é uma função indicadora que recebe 1 quando t_i é menor ou igual a t , e 0 caso contrário.

É necessário realizar uma correção de borda, tendo em vista que apesar da função ser definida para áreas infinitas, na prática, utiliza-se regiões de área finita. Para essas, o estimador dado na equação (3.14) apresenta viés. Isso pois, pontos aleatórios localizados nas bordas do mapa terão uma probabilidade menor de ter um evento próximo do que os pontos localizados no centro do mapa, sendo que eventos próximos às bordas externas a área de estudo são descartados como ilustra a Figura 3.4.

Um estimador não viesado proposto por Ripley (1977), que corrige o efeito de bordas é dado pela Equação (3.15),

Figura 3.4 – Região de estudo apresentada pelos pontos pretos, e demais ocorrências (pontos brancos) que poderiam influenciar a análise mas que não estão sendo considerados por estar fora da área de estudo.



Fonte: Baddeley et al. (2008, p.13)

$$\hat{F}(t) = \frac{\sum_{i=1}^m I_t(t_i, r_i)}{\sum_{i=1}^m I_t(r_i)}, \quad t > 0, \quad (3.15)$$

em que r_i corresponde a distância de um evento até o ponto mais próximo na borda do mapa e $I_t(t_i, r_i)$ é uma função indicadora igual a 1 quando t_i é menor ou igual a t e r_i é maior ou igual a t concomitante e 0, caso contrário.

A função $G(t)$, também conhecida como método do vizinho mais próximo, define a distribuição acumulada de distâncias entre um dado evento (fixo) e um evento (arbitrário) mais próximo. O estimador mais simples para a função $G(t)$ proposta por Ripley (1977), é apresentado na Equação (3.16):

$$\hat{G}(t) = \lambda^{-1} \sum_{i=1}^n I_t(t_i), \quad (3.16)$$

sendo n o número de pontos arbitrários no mapa, t_i a distância do i -ésimo ponto aleatório para o mais próximo dos n eventos analisado no mapa e $I_t(t_i)$ é uma função indicadora igual a 1 quando t_i é menor ou igual a t e 0 caso contrário.

A correção de borda também é necessária, e assim, um estimador proposto por Ripley (1977) que faz tal correção é dado pela Equação (3.17):

$$\hat{G}(t) = \frac{\sum_{i=1}^n I_t(t_i, r_i)}{\sum_{i=1}^n I_t(r_i)}, \quad t > 0, \quad (3.17)$$

sendo r_i a distância de um evento até o mais próximo corrigido na borda do mapa e $I_t(t_i, r_i)$ corresponde uma função indicadora que assume 1 quando t_i é menor ou igual t e r_i é maior ou igual a t e 0, caso contrário .

Com isso podemos apresentar a função $J(t)$, definida pelo quociente entre as funções $G(t)$ (distância de um evento para o evento mais próximo) e $F(t)$ (distância de um ponto arbitrário até o evento mais próximo). Conforme definida por Baddeley e Lieshout (1993), a função $J(t)$ é dada pela Equação (3.18)

$$J(t) = \frac{1 - G(t)}{1 - F(t)}, \quad (3.18)$$

para toda distância $t > 0$, tal que $F(t) < 1$.

A função $J(t)$ pode ser usada para inferir tanto na identificação ou caracterização do padrão, quanto para determinar o alcance da dependência espacial (LIESHOUT; BADDELEY, 1996). Quando utilizada em uma análise de padrão pontual, ela retorna um coeficiente numérico. Conforme Lieshout e Baddeley (1996), se $J(t)$ assumir valor identicamente igual a 1, indica detecção de um processo de poisson homogêneo ou de completa aleatoriedade espacial (CAE). Se $J(t)$ assumir valor menor ou maior que 1, indicam a presença de um padrão de agrupamento ou de regularidade respectivamente (BADDELEY et al., 2000).

A partir de exemplos apresentados em Berg (1997), é possível perceber que existem processos pontuais que não seguem a distribuição poisson, mas que $J(t) \equiv 1$. Isso, pois, a afirmação de que $J(t) \equiv 1$ sob CAE, recorre da ideia de que o viés resultante dos efeitos de borda é, aproximadamente, igual para as estimativas de $1 - F(t)$ e $1 - G(t)$, de modo que se cancelem no estimador de $J(t)$ (BADDELEY et al., 2000).

Na prática, o grande problema nas análises de processos pontuais está na restrição por meio da fronteira de uma determinada região de estudos A . Como consequência, a estimação das funções empíricas são baseadas na média de várias distâncias no processo pontual. Com isso, segue-se que, quando não corrigido, o estimador de $J(t)$ pode ser usado para uma avaliação direta da divergência

de um processo de Poisson homogêneo, comportando-se bem sobre a suposição de CAE. Porém algo bem inviável com as estimativas não corrigidas de $F(t)$ e $G(t)$ (BERG, 1997). Um estimador para a função $J(t)$ é obtido por Baddeley et al. (2000) pela Equação (3.19).

$$\hat{J}(t) = \frac{1 - \hat{G}(t)}{1 - \hat{F}(t)}, \quad (3.19)$$

em que $\hat{G}(t)$ e $\hat{F}(t)$ são estimadores das funções $F(t)$ e $G(t)$ respectivamente.

3.2.6.2 Função $J_{inhom}(t)$ não homogênea

A função $J_{inhom}(t)$ não homogênea é uma extensão da função $J(t)$ apresentada na Equação (3.19). Está definida de forma a ser capaz de lidar com a heterogeneidade espacial que possa estar presente no processo (LIESHOUT, 2010). A ideia conceitual implícita da função $J_{inhom}(t)$ é comparar o padrão de pontos em torno de um ponto típico no mapa com aqueles em torno de uma origem no espaço escolhida arbitrariamente, ponderando com a variação de intensidades presentes no processo. A partir de então, as informações sobre a estrutura do processo de pontos podem ser obtidas de forma mais precisa (BADDELEY; RUBAK; TURNER, 2015).

Lieshout (2010) define a estimativa da função $J_{inhom}(t)$ por:

$$\hat{J}_{inhom}(t) = \frac{1 - \hat{G}_{inhom}(t)}{1 - \hat{F}_{inhom}(t)}, \quad (3.20)$$

sendo $\hat{G}_{inhom}(t)$ a estimativa da função não homogênea de distância de um evento para o evento mais próximo e $\hat{F}_{inhom}(t)$ a estimativa da função não homogênea da distância de um ponto arbitrário até o evento mais próximo. Tais funções são apresentadas a seguir:

A função $\hat{G}_{inhom}(t)$ representa a estimativa da função $G(t)$ para o caso não homogêneo, apresentada por Lieshout (2010) pela Equação (3.21):

$$\hat{G}_{inhom}(t) = \frac{\sum_{i=1}^n I_t(t_i, r_i)}{\sum_{i=1}^n I_t(r_i) \hat{\delta}(t_i)}, \quad t > 0, \quad (3.21)$$

sendo $I_t(t_i, r_i)$ uma função indicadora que assume 1 quando $t_i \leq t$ e $r_i > t$ simultaneamente e 0 caso contrário. $I_t(r_i)$ é uma função indicadora que assume 1 quando $r_i > t$ e 0 caso contrário e

$\hat{\delta}(y_i)$ é a ponderação da estimativa de intensidade do processo não homogêneo. A função $G_{inhom}(t)$ apresenta uma proporção, ponderada pela intensidade variável, das distâncias dos vizinhos mais próximos, dos que estão a pelo menos uma distância t da fronteira de A (DIGGLE, 2013).

Por Lieshout (2010), a estimativa da função $\hat{F}_{inhom}(t)$, é dada pela Equação (3.22):

$$\hat{F}_{inhom}(t) = \frac{\sum_{i=1}^m I_t(t_i, r_i)}{\sum_{i=1}^m I_t(r_i) \hat{\delta}(t_i)}, \quad t > 0, \quad (3.22)$$

sendo m o número de pontos aleatórios $x_i, i = 1, \dots, m$ amostrados, $I_t(t_i, r_i)$ é uma função indicadora que assume 1 quando $t_i < t$ e $r_i > t$ concomitantemente e 0 caso contrário. $I_t(r_i)$ é uma função indicadora que assume 1 quando $r_i > t$ e 0 caso contrário e $\hat{\delta}(t_i)$ equivale a ponderação de intensidade como na Equação (3.21). A função $F_{inhom}(t)$ representa uma proporção das distâncias entre pontos gerados aleatórios e eventos, condicionado a pelo menos uma distância t da fronteira da área de estudo (DIGGLE, 2013).

3.2.6.3 Função $L(t)$ homogênea

Para apresentar a função $L(t)$, é necessário introduzir os conceitos da função $K(t)$ de Ripley (1977). Conforme Gatrell e Bailey (1996), a função $K(t)$ é capaz de compreender a dependência espacial entre diferentes regiões do processo pontual. Isto pois, se relaciona diretamente com a intensidade do processo. Em Diggle et al. (2003), recomenda-se a utilização da função $K(t)$ por sua eficiência em detectar padrões em diferentes escalas.

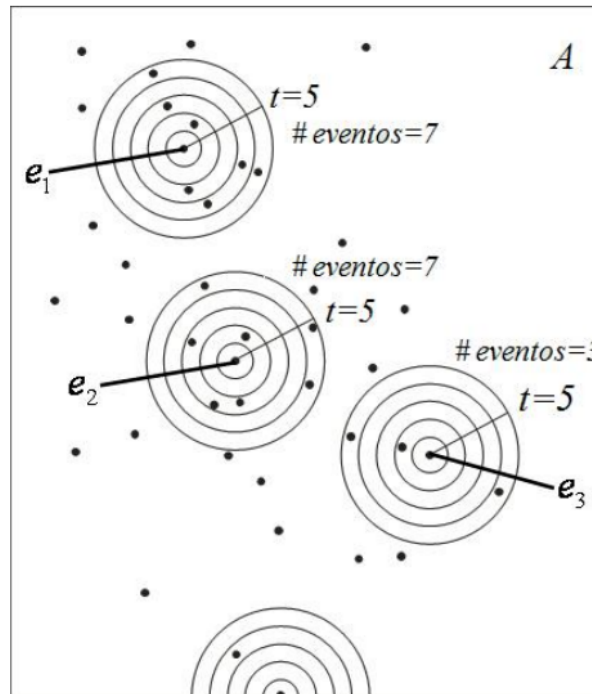
Conforme apresenta Illian et al. (2008), a ideia por trás da função $K(t)$ de Ripley é analisar o número médio de pontos encontrados dentro da distância t do ponto de referência. Para a construção da função $K(t)$, considere $\lambda K(t)$ o número médio de pontos em uma circunferência (caso bidimensional) de raio t centrado no ponto x_i (que não será contado). Variando o ponto de referência, pode-se estimar $K(t)$ como a seguir:

Seja n o número de pontos obtidos na área de estudo A , têm-se que a função $K(t)$ é obtida pela Equação (3.23):

$$K(t) = \lambda^{-1} \mathbb{E} [N(b(x_i, t) \setminus \{x_i\})], \quad (3.23)$$

onde λ corresponde a intensidade do processo, t equivale ao raio e $N(b(x_i, t) \setminus \{x_i\})$ é o número de pontos encontrados dentro da distância t do ponto x_i excluindo o próprio x_i . A ideia gráfica do cálculo de da função $K(t)$ é apresentada pela Figura (3.5).

Figura 3.5 – Representação gráfica da função $K(t)$ de Ripley



Fonte: Olinda e Scalon (2010)

Formalmente, a função teórica $K(t)$ de um processo pontual estacionário é definida pela Equação (3.24).

$$K(t) = \frac{2\pi}{\lambda^2} \int_0^t \lambda(x)^2 x dx \quad (3.24)$$

sendo t o valor do raio e $\lambda(x)$ a intensidade calculada no ponto x .

Quando se estuda o padrão espacial em uma determinada área, a aleatoriedade é definida pela CAE, ou seja, o processo gerado por um processo aleatório. Sob pressuposto de CAE, as intensidades são constantes como é apresentado nas Equações (3.25) e (3.26):

$$\lambda_1(x) = \lambda \quad (3.25)$$

$$\lambda_2(x, y) = \lambda_2(|x - y|) = \lambda_2(x) \equiv \lambda^2 \quad (3.26)$$

sendo $|x - y|$ a distância euclidiana.

De (3.24) temos que sob CAE, a função $K(t)$ será:

$$\begin{aligned} K(t) &= \frac{2\pi}{\lambda^2} \int_0^t \lambda_2(x) x \, dx \\ &= \frac{2\pi}{\lambda^2} \int_0^t \lambda^2 x \, dx \\ &= \pi t^2. \end{aligned} \quad (3.27)$$

Com isso têm-se que, para os padrões espaciais agregados, a função $K(t)$ assume valores $k(t) > \pi t^2$ e para regularidade, $k(t) < \pi t^2$.

Conforme Cressie (1993), o estimador mais simples e que corrige o efeito de bordas, é dado pela Equação (3.28):

$$\hat{K}(t) = \frac{|A|}{n^2} \sum_{i=1}^n \sum_{j=1}^n \frac{I_t(t_{ij})}{w_{ij}}, \quad t > 0, \quad (3.28)$$

sendo t_{ij} a distância entre eventos u_i e u_j com $i \neq j$, $I_t(t_{ij})$ é uma função indicadora que assume 1 quando t_{ij} é menor que a distância t e 0 caso contrário. A variável n representa o número de eventos no mapa observados em A e w_{ij} é um fator de correção que representa a proporção da circunferência ao redor de um evento u_i passando sobre o evento u_j que está dentro de A .

A função $L(t)$ representa uma transformação da função $K(t)$ dada da seguinte forma:

$$L(t) = \sqrt{\frac{K(t)}{\pi}} - t, \quad (3.29)$$

onde $K(t)$ é a função apresentada por (3.24). Para um processo de completa aleatoriedade espacial, tem-se:

$$L(t) = \sqrt{\frac{K(t)}{\pi}} - t = \sqrt{\frac{\pi t^2}{\pi}} - t = 0. \quad (3.30)$$

Em interpretação, têm-se que um padrão aleatório é caracterizado por $L(t) = 0$. Quando $L(t) > 0$ indica indícios de agregação e quando $L(t) < 0$, regularidade.

De acordo com Illian et al. (2008), a função $L(t)$ pode ser estimada pela Equação 3.31,

$$\hat{L}(t) = \sqrt{\frac{\hat{K}(t)}{\pi}} - t, \quad (3.31)$$

onde $\hat{K}(t)$ é a função apresentada em (3.28). Para Illian et al. (2008) há vantagens de se utilizar a função $L(t)$ em relação a $K(t)$. Isso pois, a interpretação de $L(t)$ obtida de forma linear na reta horizontal em 0 se torna mais fácil que o caso potencial obtido na estimativa de $K(t)$. Além disso, $L(t)$ tende a obter variâncias reduzidas para quando t toma valores maiores.

3.2.6.4 Função $L_{inhom}(t)$ não homogênea

Havendo indícios de não homogeneidade, pode-se utilizar da função $L_{inhom}(t)$ não homogênea. A função $L_{inhom}(t)$ não homogênea ou $K_{inhom}(t)$ de Ripley não homogênea transformada, é uma generalização da função $K(t)$ homogênea apresentada na subseção 3.2.6.3 para processos estacionários. A modificação para o caso não homogêneo consiste em redimensionar as distâncias entre eventos pelo produto das intensidades de primeira ordem nos dois locais correspondentes (DIGGLE, 2013).

Autores como Cressie (1993), Baddeley et al. (2000), Diggle et al. (2003) defendem a utilização de descritores não homogêneos, como o caso da função $L_{inhom}(t)$, que acomodem a heterogeneidade espacial presente na área em estudo. Apesar de um poder menor de detecção de interações (agrupamentos, regularidade), as conclusões obtidas por descritores que não levam em consideração a variação de intensidade podem acarretar em padrões tendenciosos (SOUSA; SCALON, 2018).

Por Baddeley et al. (2000), a estimativa da função $K_{inhom}(t)$ que corrige o efeito de borda, é dada por

$$\hat{K}_{inhom}(t) = \frac{1}{|A|} \sum_{i=1}^n \sum_{(j \neq i)=1}^n \frac{I_t(t_{ij})}{w_{ij} \hat{\lambda}(u_i) \hat{\lambda}(u_j)}, \quad t \geq 0, \quad (3.32)$$

sendo $|A|$ a área da região A , n o número de ocorrências observadas, w_{ij} corresponde o fator de correção dos efeitos de borda que equivale a proporção da circunferência do círculo centrado no evento u_i contendo u_j que está dentro da região de estudo e $I_t(t_{ij})$ é uma função indicadora que

assume 1 quando $t_{ij} \leq t$ e $\hat{\lambda}(u_i)$, $\hat{\lambda}(u_j)$ correspondem aos valores de intensidade estimados. Note que a única diferença entre (3.28) e (3.32) está nas correções feitas pelas estimativas $\hat{\lambda}(u_i)$ e $\hat{\lambda}(u_j)$ obtidas pelas variações de intensidade.

Na prática, $\hat{\lambda}(u_i)$ e $\hat{\lambda}(u_j)$ podem ser obtidos pela suavização de Kernel estritamente positiva, ou pela estimativa de alguma função paramétrica. Os resultados experimentais mostram que, para processos de Poisson não homogêneos, o estimador de $K_{inhom}(t)$ não homogêneo é ligeiramente enviesado para baixo quando $\hat{\lambda}(u)$ é usado. Uma discussão mais detalhada pode ser encontrada em Baddeley et al. (2000).

A estimativa da função $\hat{L}_{inhom}(t)$ é equivalente ao caso homogêneo, obtém-se através da transformação da função $\hat{K}_{inhom}(t)$, e obtida pela equação (3.33):

$$\hat{L}_{inhom}(t) = \sqrt{\frac{\hat{K}_{inhom}(t)}{\pi}} - t, \quad (3.33)$$

sendo $\hat{K}_{inhom}(t)$ a estimativa da função $K(t)$ não homogênea (BADDELEY; RUBAK; TURNER, 2015). Assim como no caso não homogêneo, a justifica para utilização da transformação $L_{inhom}(t)$ também para casos não homogêneos se dá por facilidades na interpretação (BADDELEY et al., 2000).

3.3 Processos pontuais marcados

Até o momento, o interesse esteve concentrado especificamente na distribuição dos eventos ao longo da área de estudo ou do ambiente de domínio espacial. Ainda não foi feita referência a nenhum tipo de informação além da posição do evento, mesmo sabendo que elas existem. Entretanto, conforme discute Schabenberger e Gotway (2005), a maioria dos silvicultores ou pesquisadores da área florestal, por exemplo, não ficariam satisfeitos em saber apenas sobre a distribuição espacial (regular, aleatório e agregado) das árvores. A possibilidade de acrescentar atributos ou covariáveis, tais como: diâmetro da altura do peito, idade, altura e tipo de espécies nas análises e de entender a relação destes com a configuração espacial das árvores é algo muito importante e almejado.

Adicionando-se informações aos pontos, defini-se um processo pontual marcado. A marca representa a característica ou atributo, podendo ser classificadas em categóricas (espécies de ani-

mais, categoria genética de plantas, entre outros) e contínuas (altura de árvores e diâmetro de troncos de plantas, entre outros) (BADDELEY et al., 2008).

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. Na análise de processos pontuais marcados, o interesse não está somente na distribuição espacial dos pontos, como no caso de processos pontuais não-marcados, tendo como objetivo básico de verificar se há interação ou dependência entre a distribuição dos pontos e das marcas (DIGGLE, 2013).

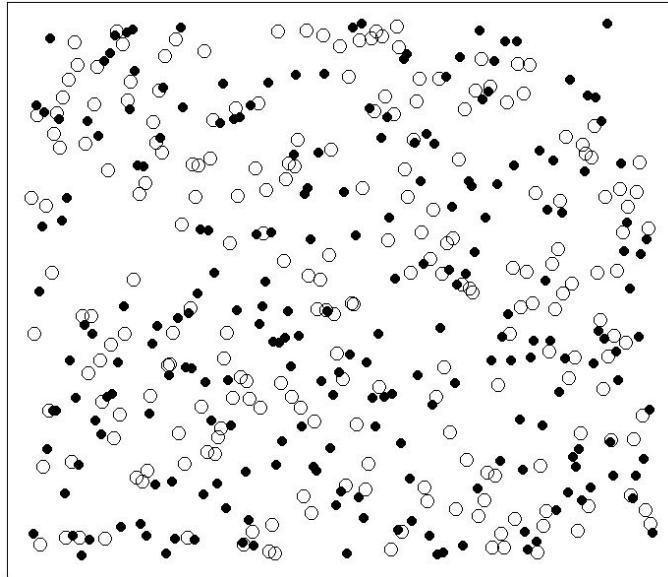
3.3.1 Marcas Qualitativas

Processos pontuais com marcas qualitativas é um caso particular em que as marcas são dos tipos categóricos ou nominais. Exemplos incluem o tipo, espécie, família, característica fisiológica ou física do fenômeno de estudo, conforme ilustrado pela Figura 3.6. Pode-se dividir em dois casos (marcas *aleatórias* ou *simultâneas*) dependendo de como foram geradas.

Processos pontuais marcados, com marcas qualitativas e aleatórias, são obtidos quando são geradas por um processo estocástico. Exemplos que sempre aparecem na literatura decorrem da observação de árvores nativas contendo algum tipo de doença. Há portanto, um processo estocástico gerando a localização dos pontos, e um segundo processo estocástico gerando as marcas indicando a condição física da árvore (estar ou não doente). Pela Figura 3.6, pode-se pensar que um único processo estocástico gerou todos os pontos (árvores), e o segundo definiu as marcas (cores correspondentes) de acordo com a condição da planta.

A hipótese nula é definida como sendo a ausência de interação (independência) entre o processo pontual que gerou os pontos e o processo que gerou as marcas. Na análise, o interesse está em responder se há relação entre a geração das marcas em relação a distribuição dos pontos (DIGGLE, 2013). Ou seja, pelo exemplo, o interesse poderia estar em saber se há tendência (aglomerado ou regular) de árvores infectadas ao longo de uma floresta.

Figura 3.6 – Representação gráfica de um processo pontual marcado com as marcas qualitativas definidas pela diferença (com preenchimento e sem preenchimento) do ponto.



Fonte: do Autor (2022)

Para um processo pontual marcado, com marcas qualitativas e obtidas de forma simultânea (não sendo por uma realização de uma variável aleatória), o processo pontual que gerou todos os pontos é interpretado como um composto de sub-processos definidos por cada tipo de marca (ILLIAN et al., 2008). Exemplificando com a Figura 3.6, pode-se entender e observar a existência de dois sub-processos, um formado pelos pontos de cor branca e outro, pelos pontos de cor preta.

Neste trabalho, a análise de processos pontuais marcados com marcas qualitativas é feita entre pontos (árvores) e marcas (espécie). Dessa forma, é necessário introduzir conceitos de processo pontual espacial bivariado. As estimativas dos efeitos de primeira e segunda ordem definidos na sequência.

3.3.1.1 Efeitos de primeira e segunda ordem (caso bivariado)

Os efeitos de primeira e segunda ordem permanecem para o caso bivariado. Em analogia ao processo pontual não marcado apresentado na seção 3.2.2, Cressie (1993) define o efeito de primeira ordem para o i -ésimo subprocesso com $i \in \{1, 2\}$ é dado pela Equação (3.34):

$$\lambda_i(\mathbf{s}) = \lim_{|d\mathbf{s}| \rightarrow 0} \left\{ \frac{\mathbb{E}[N_i(d\mathbf{s})]}{|d\mathbf{s}|} \right\} \quad (3.34)$$

sendo $d\mathbf{s}$ uma pequena região em torno de \mathbf{s} , e $N_i(d\mathbf{s})$ representa o número de eventos de marcas $i \in \{1, 2\}$ localizados nessa região.

O efeito de segunda ordem é apresentado pela Equação (3.35):

$$\lambda_{12}(d\mathbf{s}_i, d\mathbf{s}_j) = \lim_{|d\mathbf{s}_i|, |d\mathbf{s}_j| \rightarrow 0} \left\{ \frac{\mathbb{E} [N_1(d\mathbf{s}_i)N_2(d\mathbf{s}_j)]}{|d\mathbf{s}_i||d\mathbf{s}_j|} \right\}, \quad (3.35)$$

sendo que, $N_1(d\mathbf{s}_i)$ e $N_2(d\mathbf{s}_j)$ representam o número de eventos dentro de pequenos círculos centrados nos pontos \mathbf{s}_i e \mathbf{s}_j , $|d\mathbf{s}_i|$ e $|d\mathbf{s}_j|$ corresponde às áreas dos círculos $d\mathbf{s}_i$ e $d\mathbf{s}_j$ respectivamente.

3.3.1.2 Função $J_{12}(t)$ bivariada

Um padrão de pontos bivariado é entendido como um processo $\mathbf{X} \subset \mathbb{R}^d$ marcado com marcas em $\{1, 2\}$, podendo ser representado por $\mathbf{X} = \{X_1, X_2\}$ sendo X_1 e X_2 padrões de pontos univariados. De acordo com Lieshout e Baddeley (1999), pode-se definir uma função $J_{12}(t)$ para um padrão de pontos marcados bivariado obtida a partir de uma extensão da função $J(t)$ univariada definida por Lieshout e Baddeley (1996). Intuitivamente, a função fará a comparação entre a distribuição de distâncias de um evento de marca 2 para o ponto de marca 1 mais próximo e a distribuição de distâncias de um ponto arbitrário em relação a um ponto típico do evento de marca 2.

Para Lieshout e Baddeley (1999), considerando um processo de pontos \mathbf{X} estacionário, a estimativa da função $J_{12}(t)$ bivariada é obtida pela Equação (3.36).

$$\hat{J}_{12}(t) = \frac{1 - \hat{G}_{12}(t)}{1 - \hat{F}_2(t)}, \quad (3.36)$$

para todo $t \leq 0$ e $\hat{F}_2(t) < 1$.

A função $\hat{F}_2(t)$ corresponde a estimativa da distribuição de distâncias t de um ponto arbitrário a um evento do processo de ponto 2. Baddeley et al. (2000) apresenta a estimativa da função $\hat{F}_2(t)$ bivariada com correção dos efeitos de borda pela Equação (3.37):

$$\hat{F}_2(t) = \frac{\sum_{i=1}^m I_t(z_i, s_2) I_t(z_i)}{\sum_{i=1}^m I_t(z_i)}, \quad (3.37)$$

sendo m o número de pontos $z_i, i = 1, \dots, m$ gerados de um processo aleatório e s_2 é o conjunto de pontos do tipo 2. $I_t(z_i, s_2)$ corresponde a uma função indicadora que assume 1 quando a distância entre z_i e s_2 é menor que t e assume 0 caso contrário. $I_t(z_i)$ corresponde a uma função indicadora que assume 1 quando z_i está a uma distância superior a t da borda da área de estudo.

A função presente no numerador da parte direita da equação (3.36), corresponde a função do vizinho mais próximo par ao caso bivariado e descreve a distribuição de distâncias dos eventos do fenômeno tipo 2 aos pontos do fenômeno do tipo 1. Por Baddeley et al. (2000), tem-se que a estimativa da função $\hat{G}_{12}(t)$ bivariada, com correção de efeitos de bordas é dada pela Equação (3.38):

$$\hat{G}_{12}(t) = \frac{\sum_{i=1}^n I_t(s_{1i}, s_2) I_t(s_{1i})}{\sum_{i=1}^n I_t(s_{1i})}, \quad (3.38)$$

sendo n o número de eventos do fenômeno 1, s_2 corresponde ao conjunto de eventos do fenômeno 2, s_{1i} representa o i -ésimo evento do tipo 1. $I_t(s_{1i}, s_2)$ é uma função indicadora que recebe 1 quando a distância entre os eventos s_{1i} e s_2 é menor que t e 0 caso contrário e $I_t(s_{1i})$ é uma função indicadora que recebe 1 quando s_{1i} está a uma distância igual ou superior a t da borda da área de estudo A e 0 caso contrário.

Os valores de $J_{12}(t) > 1$ são interpretados como indicativos de regularidade de pontos do tipo 2 por pontos do tipo 1. Isso pois, é equivalente a dizer que $G_{12}(t) < F_2(t)$, ou seja, existe uma tendência de haver menos pontos do tipo 2 próximos aos pontos do tipo 1, do que pontos gerados arbitrariamente em relação aos eventos do tipo 1. O mesmo pode-se pensar $J_{12}(t) < 1$ indicando agrupamentos entre os pontos do tipo 1 e 2.

3.3.1.3 Função $J_{12}^{inhom}(t)$ bivariada não homogênea

Apesar de conveniente supor a estacionariedade, a heterogeneidade espacial se faz presente em muitos fenômenos. A função $J_{12}^{inhom}(t)$ bivariada não homogênea é uma generalização da função $J_{12}(t)$ bivariada homogênea, capaz de lidar com a heterogeneidade da dispersão dos pontos. Assim como as outras versões da função $J(t)$, na versão bivariada não homogênea, ela compara o padrão de pontos em torno de um ponto típico no mapa com aquele em torno de um ponto arbitrário de origem escolhida no espaço levando em consideração a variação de intensidades ou heteroge-

neidade da área em estudo, a fim de obter informações sobre a estrutura de interação do processo de ponto que gerou os dados Cronie e Lieshout (2016).

Para os mesmos autores, a estimativa da função $\hat{J}_{12}^{inhom}(t)$ é obtida através da razão entre as funções $\hat{G}_{12}^{inhom}(t)$ e $\hat{F}_{12}^{inhom}(t)$ conforme apresentado pela equação (3.39),

$$\hat{J}_{12}^{inhom}(t) = \frac{1 - \hat{G}_{12}^{inhom}(t)}{1 - \hat{F}_{12}^{inhom}(t)}. \quad (3.39)$$

A estimativa da função de $\hat{G}_{12}^{inhom}(t)$, representa a distribuição de distâncias ponderadas por intensidade de um ponto do tipo 1 para os eventos do tipo 2 mais próximo. Com correção dos efeitos de borda, a mesma pode ser definida por:

$$\hat{G}_{12}^{inhom}(t) = \frac{\sum_{i=1}^n I_t(s_{1i}, s_2) I_t(s_{1i})}{\sum_{i=1}^n I_t(s_{1i}) \hat{\delta}_1(t_i)}, \quad (3.40)$$

dados que n é o número de eventos do fenômeno 1, s_2 corresponde ao conjunto de eventos do fenômeno 2, s_{1i} representa o i -ésimo evento do tipo 1. $I_t(s_{1i}, s_2)$ é uma função indicadora que recebe 1 quando a distância entre os eventos s_{1i} e s_2 é menor que t e 0 caso contrário e $I_t(s_{1i})$ é uma função indicadora que recebe 1 quando s_{1i} está a uma distância igual ou superior a t da borda da área de estudo e 0 caso contrário e $\hat{\delta}_1(t_i)$ representa a ponderação de intensidade do fenômeno 1.

A distribuição de distâncias dos eventos do tipo 2 por pontos gerados aleatórios ponderados pela intensidade relativa do mesmo fenômeno, é estimada pela função $\hat{F}_2^{inhom}(t)$ com correção dos efeitos de borda, conforme apresentado por (3.41):

$$\hat{F}_2^{inhom}(t) = \frac{\sum_{i=1}^m I_t(z_i, s_2) I_t(z_i)}{\sum_{i=1}^m I_t(z_i) \hat{\delta}_2(t_i)}, \quad (3.41)$$

sendo m o número de pontos $z_i, i = 1, \dots, m$ gerados de um processo aleatório e s_2 é o conjunto de pontos do tipo 2. $I_t(z_i, s_2)$ corresponde a uma função indicadora que assume 1 quando a distância entre z_i e s_2 é menor que t e assume 0 caso contrário. $I_t(z_i)$ corresponde a uma função indicadora que assume 1 quando z_i está a uma distância superior a t da borda da área de estudo e $\hat{\delta}_2(t_i)$ representa a ponderação de intensidades através do fenômeno 2.

3.3.1.4 Função $L_{12}(t)$ bivariada

A função $L_{12}(t)$ bivariada é uma transformação da função $K_{12}(t)$ bivariada, correspondendo a mais um descritor que auxilia na estimação do nível de dependência entre dois tipos de eventos ou fenômenos (ILLIAN et al., 2008). Por Diggle et al. (2003), a função $K_{12}(t)$ bivariada pode ser definida de acordo com a Equação (3.42):

$$K_{12}(t) = \lambda_2^{-1} E[N(t)], \quad (3.42)$$

sendo t o raio obtido para contagem, $N(t)$ representa o número de eventos do tipo 2 encontrados dentro de uma distância t de uma dada ocorrência arbitrária do tipo 1. λ_2^{-1} corresponde o inverso da intensidade de primeira ordem do evento do tipo 2.

Por Cressie (1993), o estimador da função $K(t)$ bivariada é feito através da média ponderada dos dois estimadores obtidos pelas Equações (3.43) e (3.44):

$$\tilde{K}_{12}(t) = \hat{\lambda}_2^{-1} n_1^{-1} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} w_{ij} I(t_{ij} \leq t), \quad (3.43)$$

$$\tilde{K}_{21}(t) = \hat{\lambda}_1^{-1} n_2^{-1} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} w_{ij} I(t_{ij} \leq t), \quad (3.44)$$

sendo que, $t_{ij} = |s_{1i} - s_{2j}|$, corresponde a distância entre o i -ésimo evento do tipo 1 ao j -ésimo evento do tipo 2. w_{ij} corresponde a correção de borda, e n_1 e n_2 corresponde ao número de eventos do tipo 1 e 2 respectivamente.

De (3.43) e (3.44) tem-se que o estimador de $K_{12}(t)$ é dado pela Equação (3.45) apresentada a seguir.

$$\begin{aligned} \hat{K}_{12}(t) &= (n_1 n_2)^{-1} |A| \left\{ \frac{n_1 \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} w_{ij} I(d_{ij} \leq t) + n_2 \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} w_{ij} I(d_{ij} \leq t)}{n_1 + n_2} \right\} \\ &= (n_1 n_2)^{-1} |A| \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} W_{ij}^* I(d_{ij} \leq t), \end{aligned} \quad (3.45)$$

sendo $W_{ij}^* = \frac{n_1 w_{ij} + n_2 w_{ij}}{n_1 + n_2}$ a correção de borda.

Assim como para o caso univariado, a função $L_{12}(t)$ bivariada representa uma transformação da função $K_{12}(t)$ bivariada. Por Illian et al. (2008), A função $L_{12}(t)$ bivariada estimada é dada Equação 3.46:

$$\hat{L}_{12}(t) = \sqrt{\frac{\hat{K}_{12}(t)}{\pi}} - t \quad (3.46)$$

sendo, $\hat{K}_{12}(t)$ é a estimativa da função $K_{12}(t)$ bivariada.

3.3.1.5 Função $L_{12}^{inhom}(t)$ bivariada não homogênea

A função $L_{12}^{inhom}(t)$ bivariada não homogênea é obtida através de uma transformação da função $K_{12}^{inhom}(t)$ marcada bivariada não homogênea. A função $K_{12}^{inhom}(t)$ bivariada de Ripley não homogênea é uma generalização da função $L_{12}(t)$ marcada bivariada, que permite analisar a correlação espacial da localização dos eventos ao redor de uma escala de interesse sendo capaz de lidar com a distribuição não estacionária dos eventos. Desse modo, a detecção do padrão de distribuição espacial torna-se mais preciso levando em consideração a heterogeneidade natural da dispersão dos eventos.

O estimador de $\hat{K}_{12}^{inhom}(t)$ é dado pela Equação (3.47) apresentada a seguir.

$$\hat{K}_{12}^{inhom}(t) = \frac{1}{|A|} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \frac{w_{ij} I(t_{ij} \leq t)}{\hat{\lambda}(u_{1i}) \hat{\lambda}(u_{2j})}, \quad (3.47)$$

sendo $t_{ij} = |u_{1i} - u_{2j}|$, corresponde a distância entre o i -ésimo evento do tipo 1 ao j -ésimo evento do tipo 2. w_{ij} corresponde a correção de borda, e n_1 e n_2 corresponde ao número de eventos do tipo 1 e 2 respectivamente e $\hat{\lambda}(u_{1i})$, $\hat{\lambda}(u_{2j})$ correspondem aos valores de intensidade estimados.

A estimativa da função $\hat{L}_{12}^{inhom}(t)$ é dada por,

$$\hat{L}_{12}^{inhom}(t) = \sqrt{\frac{\hat{K}_{12}^{inhom}(t)}{\pi}} - t \quad (3.48)$$

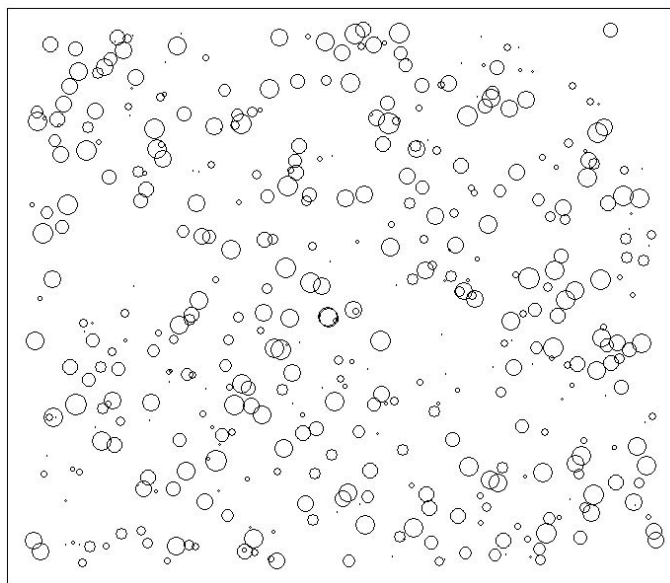
sendo, $\hat{K}_{12}^{inhom}(t)$ é a estimativa da função $K_{12}^{inhom}(t)$ bivariada não homogênea.

3.3.2 Marcas Quantitativas

Os processos pontuais marcados com marcas quantitativas é obtida quando se obtém variáveis numéricas (contínuas). Pode-se pensar em altura, magnitude ou diâmetro do fenômeno estudado, como ilustra a figura 3.7. O interesse está em verificar se há dependência entre pontos e marcas, ou seja, se as marcas dos pontos vizinhos tendem a ser mais parecidas. A análise concentra-se em averiguar se há relação de dependência entre as diferenças numéricas das marcas, considerando a distância (posição) dos pontos correspondentes Diggle (2013).

Existem várias formas para análise de hipótese de independência para estes casos conforme apresenta Illian et al. (2008). Este trabalho utilizará a função de correlação marcada conforme Olinda e Scalon (2010).

Figura 3.7 – Representação gráfica de um processo pontual marcado com as marcas quantitativas definidas pela diferença de tamanho do ponto.



Fonte: do Autor (2022)

3.3.2.1 Função de Correlação Marcada

Para testar a hipótese de independência existe um grande número de medidas, sendo apresentado neste trabalho apenas na função de correlação marcada definida por Illian et al. (2008)

por:

$$\rho_f(t) = \frac{E[f(m_i, m_j) \times t]}{\mu^2} \quad (3.49)$$

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

$$\hat{\rho}_f(t) = \frac{1}{\mu^2} \sum_{i=1}^n \sum_{j=1}^n f(m_i, m_j) \times A_{ij} \quad (3.50)$$

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.

3.4 Teste de Monte Carlo

A condução de um teste de hipótese pode ser realizada considerado tanto testes formais, quanto procedimentos gráficos com a construção dos envelopes de confiança. Ambos utilizam simulações Monte Carlo. 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).

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:

$$u_i = \int_0^{t_0} \{\hat{\Phi}_i(k) - \Phi(k)\}^2 dk, \quad i = 1, \dots, s. \quad (3.51)$$

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 et al., 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 (3.52) e (3.53) respectivamente.

$$U_i(t) = \min\{\hat{\Phi}_i(t)\}, i = 2, \dots, s. \quad (3.52)$$

$$U_s(t) = \max\{\hat{\Phi}_i(t)\}, i = 2, \dots, s. \quad (3.53)$$

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).

3.5 Processos pontuais em ciências florestais

Existem diversos trabalhos onde pesquisadores aplicam métodos de processos pontuais para análise de configurações pontuais de espécies de árvores. Vejamos alguns exemplos:

Pelo trabalho de Stoyan e Penttinen (2000), utilizou-se das funções como a $J(t)$, $L(t)$ e função de correlação marcada em aplicação e propostas de análises para dados florestais, bem como análise de distribuição e correlação entre as marcas. Para os mesmos autores, fica evidente que processos pontuais e processos de pontos marcados são ferramentas poderosas em estatísticas florestais. Os autores ainda ressaltam sobre a possibilidade de incluir fatores (covariáveis) que geralmente levam a supostas heterogeneidade na área em estudo.

Nos estudos de Anjos et al. (2004), na análise do padrão espacial da araucária em áreas do estado do Paraná, observa-se a eficácia da função $K(t)$ de Ripley, podendo ser um método capaz de fornecer muita informação sobre a estrutura espacial de uma espécie florestal nativa.

Capretz (2004) utilizou das funções $K(t)$ e $K_{12}(t)$ de Ripley para analisar o padrão de distribuição e interação espacial duas a duas de quatro espécies mais abundantes de Unidades de Conservação localizadas estado de São Paulo. Para o mesmo autor, é evidente o potencial da Função K de Ripley para investigar uma série de questões relevantes e de interesse da ecologia florestal.

Em Olinda e Scalon (2010), utilizou-se das funções $F(t)$, $G(t)$, $J(t)$ e $K(t)$ e correlação marcada na análise de distribuição de marcadores genéticos em espécies vegetais. O interesse era verificar se os marcadores são determinados pela estrutura espacial das árvores. Para o mesmo autor são tais funções são poderosas quando unidas à simulação Monte Carlo.

Machado et al. (2012) utilizaram das funções $K(t)$ (univariada) e $K_{12}(t)$ (bivariada) de Ripley na análise de distribuição e interação espacial de espécies localizadas em um fragmento de Floresta Ombrófila Mista Montana. Em seu trabalho ainda recomendou-se a utilização dessas funções para análises da dinâmica de florestas.

Scalon, Oliveira e Mello (2012) utilizou de métodos de processos pontuais baseados na função $D(t)$ de Dirichlet e $K(t)$ de Ripley para descrever a distribuição espacial de árvores em um fragmento de floresta Semidecidual Montana localizada em Lavras, MG. Os mesmos autores ainda sugeriram a utilização de ambas funções como complementares na análise.

No trabalho de Cunha (2013), os métodos de processos pontuais foram utilizados para analisar a estrutura espacial de uma floresta na Amazônia. Com o estudo foi possível observar que a utilização da função K de Ripley apresentou resultados consistentes permitindo estabelecer parâmetros de referência mínimas nos estudos de fragmentos densos de florestas.

No trabalho de Frade (2014), foi possível observar evidências estatísticas de fortes associações entre as localizações das árvores (pontos) e os fatores ambientais (marcas) em uma área de conservação ecológica localizado no estado de São Paulo, utilizando das funções $K(t)$ (homogênea) e $K_{inhom}(t)$ (não homogênea) de Ripley.

O trabalho de Dalmaso et al. (2016) utilizou da função $K_{12}(t)$ de Ripley bivariada para analisar e descrever padrões de interações espaciais de espécies intraespecíficas de *Ocotea odorifera* e *Rohwer Sassafrás* de um fragmento florestal de três hectares. A hipótese levantada que não existem evidências da ocorrência de interações espaciais entre os indivíduos de canela-sassafrás foi rejeitada, evidenciando que *Ocotea odorifera* apresentou interações negativas de indivíduos adultos com a regeneração em distâncias curtas e interações positivas entre os indivíduos regenerantes.

No trabalho de Pottker et al. (2016) foi realizado um estudo dos processos intrínsecos da distribuição espacial da espécie *Ocotea odorifera* em uma floresta em Irati, Paraná. Para análise, foi utilizado do estimador de intensidade de kernel para avaliação de padrões globais e dispersão de árvores na região amostrada, e para evidenciação de padrões espaciais, utilizou-se da função $K_{inhom}(t)$ de Ripley não homogênea. Os resultados obtidos no estudo permitiram investigar as preferências ambientais das espécies avaliadas, bem como complementar as informações obtidas em análises tradicionalmente utilizadas. Verificou-se agrupamentos da espécie nas áreas mais elevadas da região amostrada, possivelmente em solos mais bem drenados.

Almeida (2018) realizou um estudo de interação de espécies nativas na Amazônia, apresentando resultados que mostram a importância de analisar as propriedades de primeira e de segunda ordem na análise de configurações pontuais. Recomendou-se ainda, a utilização de pelo menos dois

descritores para avaliação das propriedades de segunda ordem. Como exemplo, pode-se pensar nas funções $J(t)$ e $L(t)$ que serão apresentadas e utilizadas neste trabalho.

Duarte et al. (2019) utilizou da função $L(t)$ univariada com correção isotrópica de borda para caracterizar a estrutura espacial de uma espécie arbórea localizada em um fragmento de Floresta Ombrófila Mista Alto-Montana em Urubici, Santa Catarina. Os resultados obtidos pela função $L(t)$ apontaram um padrão de distribuição agregado da espécie.

As funções $F(t)$ e $K(t)$ de Ripley foram utilizadas no trabalho de Souza (2020) para identificar o padrão espacial das árvores de castanha-da-amazônia na floresta nativa no sul do Amapá. Pelos resultados, obteve-se que as árvores de *Bertholletia excelsa* apresentaram distribuição espacial aleatória tanto em floresta ombrófila aberta quanto em transição cerrado-floresta, Não havendo portanto, diferença no padrão de distribuição nas áreas estudadas.

No trabalho de Vieira et al. (2021), utilizou-se da função $K(t)$ de Ripley com raios equiespaçados em 5 metros para avaliar a estrutura diamétrica e distribuição espacial da espécie *Dipteryx odorata* no estado do Pará, Brasil. Através da análise gráfica, constatou-se que a distribuição espacial da espécie é predominantemente aleatória. As conclusões obtidas no trabalho descreveu que a aleatoriedade seja decorrência do ciclo de vida da espécie, sendo que, a medida que os indivíduos crescem sua distribuição torna-se mais aleatória.

4 MATERIAIS E MÉTODOS

4.1 Dados

Os dados utilizados no trabalho correspondem à informações de duas das espécies mais populosas encontradas no fragmento de floresta de classificação Semidecidual Montana do bioma Mata Atlântica, localizado na Universidade Federal de Lavras (UFLA) pelo inventário de 2017. O remanescente florestal, conhecido como Matinha, cobre uma área de aproximadamente 6.48 hectares localizada nas coordenadas $21^{\circ}13'43.4''S$ e $44^{\circ}58'16.1''W$ com altitude média de 950 metros acima do nível do mar conforme apresentado no mapa da Figura 4.1. O clima de Lavras, segundo a classificação climática de Köppen-Geiger é caracterizado como subtropical com inverno seco e verão chuvoso (DANTAS; CARVALHO; FERREIRA, 2007). De acordo com (NUNES et al., 2003), os solos são Latossolos Vermelhos Distroféricos típicos, bem drenados e com textura argilosa.

Figura 4.1 – Fragmento floresta Semidecidual Montada ("Matinha").



Fonte: Do Autor (2022)

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*.

4.2 *Copaifera langsdorffii*

A *Copaifera langsdorffii*, Figura 4.2, é popularmente conhecida por copaíba ou óleo-capaíba. A espécie pertence à família *Fabaceae*, na ordem *Fabales* da classe *Dicotyledonae* e divisão *Angiospermae*. É uma árvore que atinge até 15 m de altura e pode ser encontrada em todo território brasileiro (FREITAS; OLIVEIRA, 2002). Apresenta uma alta compatibilidade á condições edáficas, aparecendo tanto em solos drenados e de maneira geral, cresce melhor em solos de matas ciliares e semidecíduas (CARVALHO, 2014). A espécie possui o tronco áspero, de coloração escura com folhas pecioladas e penuladas. As flores são pequenas, apétalas, hermafroditas e arranjadas em panículos axilaresos e os frutos contêm uma semente ovoide envolvida por um arilo abundante e colorido. A dispersão de sementes da espécie ocorre geralmente por aves ou outros animais que consomem seus frutos (JUNIOR; PINTO, 2002).

4.3 *Xylopia brasiliensis*

A *Xylopia brasiliensis* (Figura 4.3) é conhecida popularmente em Minas Gerais pelos nomes de pidaíba, pindaíba-boca-seca, pindaibinha e entre outros. A espécie possui a taxonomia hierarquizada pela divisão *Magnoliophyta* (*Angiospermae*), classe *Magnoliopsida* (*Dicotyledonae*), ordem *Magnoliales*, família *Annonaceae* e gênero *Xylopia*.

De acordo com Carvalho (2014), a espécie está presente em vários estados brasileiros incluindo Minas Gerais, São Paulo, Rio de Janeiro, Paraná e Santa Catarina. Ocorre naturalmente em planície litorânea e solos rasos de rápida drenagem, de fertilidade química alta e arenosa. Sua madeira possui cerne branco acinzentada, com densidade média de $0,70g/cm^3$ e utilizada para construção de caixotaria, tamancaria, caibros, vigas e entre outros.

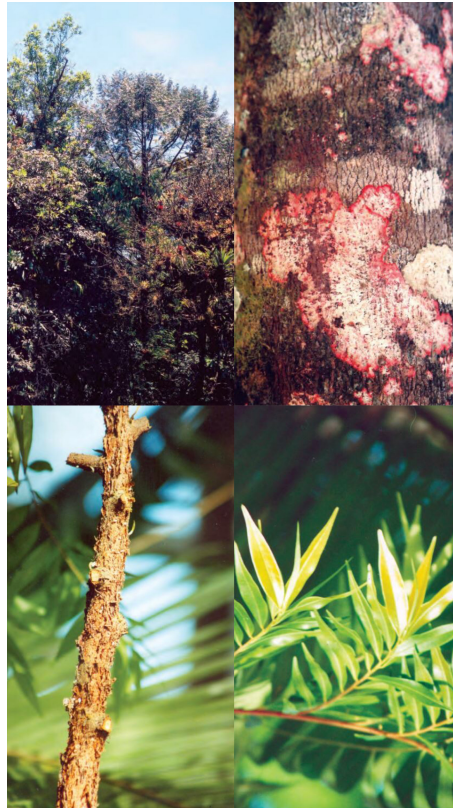
Figura 4.2 – *Copaifera langsdorffii*

Fonte: Lorenzi (1992, p.152)

4.4 Métodos Estatísticos

4.4.1 Análise Exploratória

A estatística do trabalho inclui uma análise exploratória inicial. Consiste na construção de gráficos de distribuição de pontos no sistema de coordenadas (x,y) para visualização dos pontos ou dispersão de árvores de cada espécie. Ademais, será apresentada uma tabela com as principais estatísticas descritivas contendo a contagem de árvores por espécie e informações das marcas quantitativas (diâmetro a altura do peito), sendo calculadas as médias, desvio-padrão e amplitude. Ainda das marcas, será apresentado um histograma com a distribuição dos DAP de cada uma das espécies.

Figura 4.3 – *Xylopia brasiliensis*

Fonte: Carvalho (2014, p.450)

4.4.2 Análise de primeira ordem

No trabalho, consta a análise dos efeitos de primeira ordem ou globais. Consiste em determinar a intensidade de ocorrências do processo. Para visualização completa da variação de intensidades para cada uma das espécies, será utilizada a função alisamento de Kernel. O estimador de Kernel utilizado será o mesmo apresentado pela Equação (3.13) que utiliza da função de interpolação apresentada na Equação (3.12). Para adequação do raio de suavização τ , será utilizado a medida proposta por (CRONIE; LIESHOUT, 2018).

4.4.3 Análise de segunda ordem (Univariada)

A análise de segunda ordem para o caso univariado foi feita utilizando das funções $J_{inhom}(t)$ e $L_{inhom}(t)$ não homogêneas. Hipóteses de completa aleatoriedade espacial foram testadas para

ambos casos. Para correção dos efeitos de borda, foi levado em consideração os estimadores não viesados definidos nas Equações (3.20) e (3.33).

Foram estabelecidas distâncias compreendidas entre 0 e 12 metros para a função $J_{inhom}(t)$ não homogênea e 0 a 70 metros para a função $L_{inhom}(t)$ não homogêneas. Para a construção dos envelopes de confiança, foram realizadas 1000 simulações Monte Carlo. A análise foi realizada para cada uma das duas espécies em estudo.

4.4.4 Análise de segunda ordem (Bivariada)

Os efeitos de segunda ordem (bivariada) foram estimados para analisar o padrão espacial de interação entre as espécies. Para isso foram utilizadas as funções $J_{12}^{inhom}(t)$ e $L_{12}^{inhom}(t)$ bivariadas não homogênea. As ponderações de intensidades foram obtidas através da estimativa de uma suavização kernel (CRONIE; LIESHOUT, 2018). A hipótese nula é definida como a independência entre as espécies. Não rejeitá-la implica que não há interação (dependência) espacial entre as duas espécies. As funções $J_{12}^{inhom}(t)$ e $L_{12}^{inhom}(t)$ bivariadas não homogêneas foram estimadas com correção dos efeitos de borda, salve às Equações (3.39) e (3.48) respectivamente.

Foram definidas distâncias máximas de 12 metros para a função $J_{12}^{inhom}(t)$ bivariada não homogênea e de 70 metros para a $L_{12}^{inhom}(t)$ bivariada não homogênea. Para construção dos envelopes foram utilizadas de 1000 simulações Monte Carlo.

4.4.5 Análise de segunda ordem (Marcas Contínuas)

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. 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 (3.50). Neste teste, considera-se como hipótese nula, como a independência entre pontos (localização das árvores) e marcas (DAP). A função de correlação marcada foi estimada dentro do intervalo de 0 a 100 metros. A construção dos envelopes de confiança foi feita utilizando de 1000 simulações Monte Carlo.

4.4.6 Recursos Computacionais

As análises foram realizadas pelo software R Core Team (2022), utilizando funções disponibilizadas na biblioteca *spatstat* (BADDELEY et al., 2005).

5 RESULTADOS E DISCUSSÕES

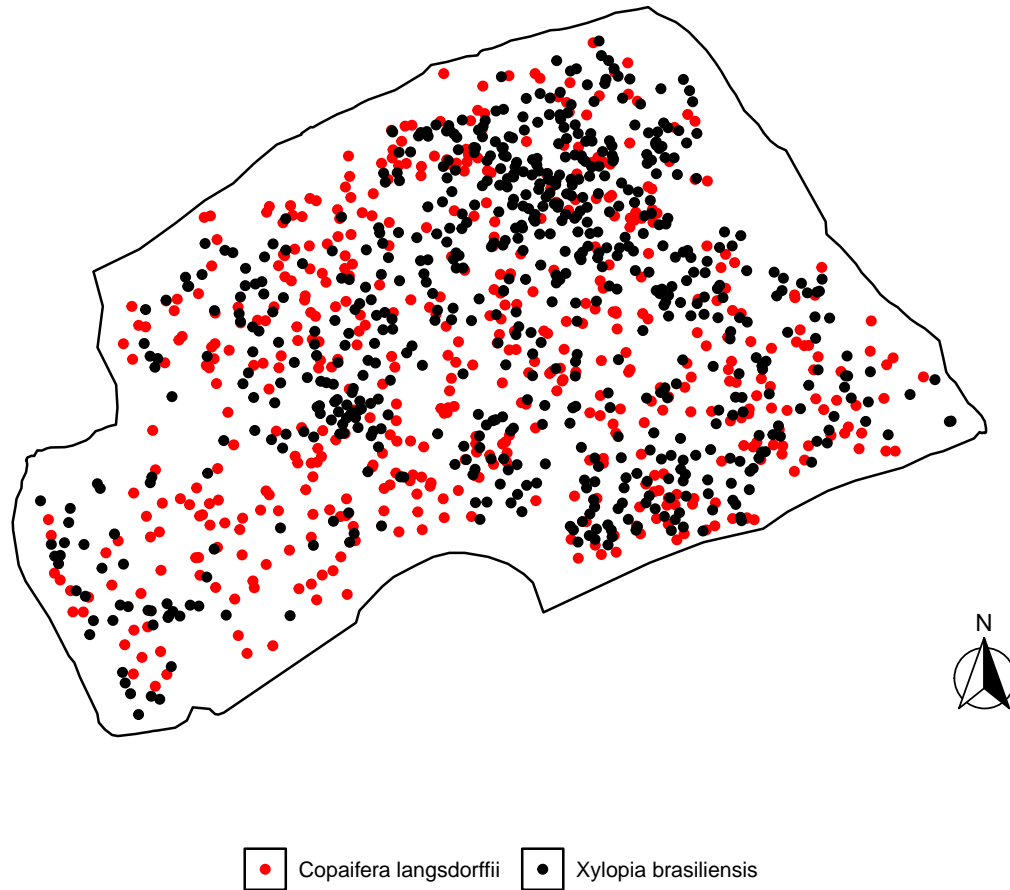
5.1 Análise Exploratória

A distribuição espacial das espécies de árvores nativas em estudo neste trabalho são apresentadas pela Figura 5.1. São elas: *Copaifera langsdorffi* e *Xylopia brasiliensis*. Pode-se observar que ambas espécies se distribuem em toda área de estudo com indícios de heterogeneidade espacial, sendo que, algumas regiões apresentam maior concentração de árvores. Este fato foi mais evidente com a análise de primeira ordem apresentada na seção 5.2. A baixa intensidade de eventos nas bordas se dá ao fato de que as parcelas de coleta dos dados não contemplou efetivamente a região total apresentada pela Figura 5.1. Com isso, algumas árvores que se encontram nas extremidades da região não foram amostradas. Apesar de algumas tendências de aglomerados, a Figura 5.1 não possibilita maiores comentários sobre o padrão de distribuição e interação das espécies, tornando fundamental a utilização dos métodos de processos pontuais para análises dos efeitos de primeira e segunda ordem (MONTEIRO et al., 2004).

A Figura 5.2 apresenta histogramas de distribuições dos diâmetros à altura do peito de cada uma das espécies. É possível observar assimetrias em suas distribuições para ambas as espécies. Complementando às estatísticas descritivas, a Tabela 5.1 apresenta frequência, valores médios, desvio padrão e amplitude para os DAP.

Pela Figura 5.2, 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 et al. (2006), Dalmaso et al. (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 5.1 – Distribuição espacial e classificação das árvores quanto a espécie.



Fonte: Do Autor (2022)

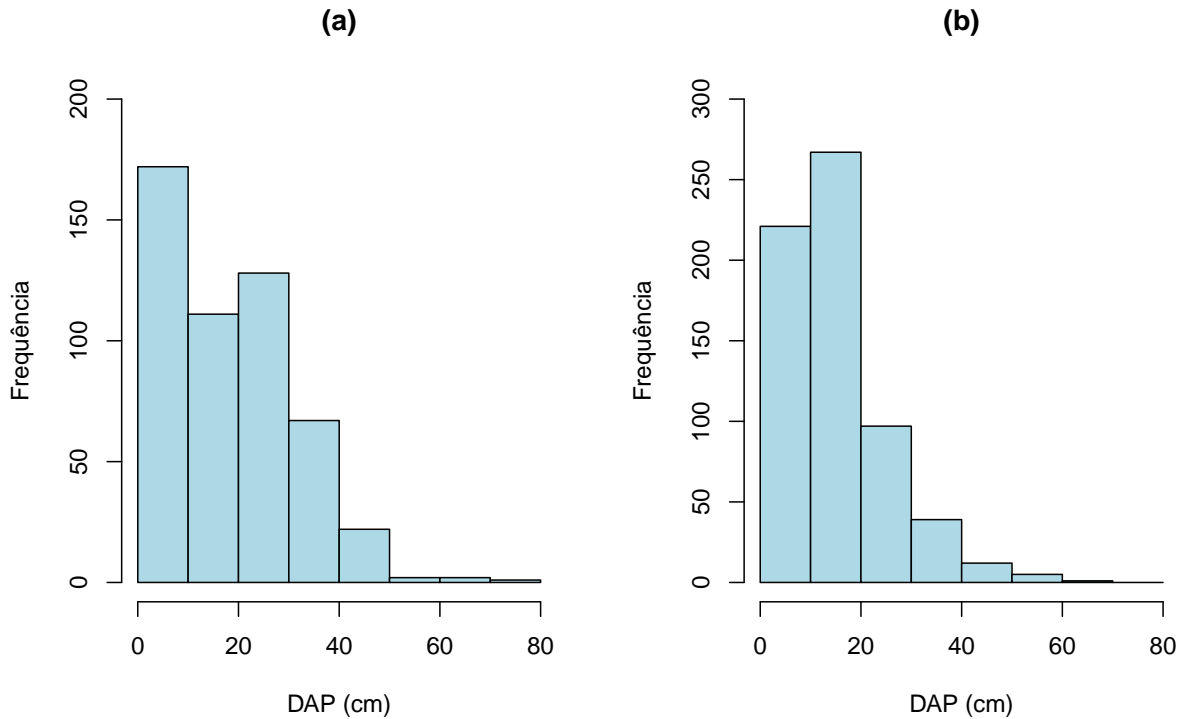
5.2 Análise de primeira ordem

A análise de primeira ordem é o primeiro passo na análise espacial de um fenômeno aleatório (MONTEIRO et al., 2004). 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 5.3.

A definição dos raios de influência para os mapas apresentados pela Figura 5.3, foram obtidos pela fórmula de Campbell conforme (CRONIE; LIESHOUT, 2018). Os raios obtidos para suavização na estimativa da função de Kernel é apresentada na Tabela 5.2.

Pela análise dos mapas kernel (Figura 5.3), é possível observar que ambas espécies arbóreas se distribuem ao longo de toda região de estudo. Fica evidente a presença da heterogeneidade

Figura 5.2 – Histogramas com frequências de DAP em (cm) para as espécies: *Copaifera langsdorffii* (a), *Xylopiya brasiliensis* (b).



Fonte: Do Autor (2022)

Tabela 5.1 – Tabela descritiva de diâmetro à altura do peito (cm) por espécie de árvore.

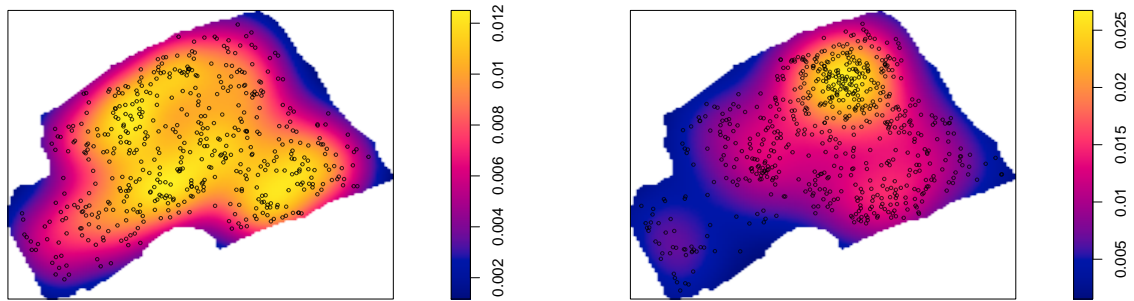
Espécie	n	Média	Desvio Padrão	Amplitude
<i>Copaifera langsdorffii</i>	505	18,83	12,25	70,69
<i>Xylopiya brasiliensis</i>	642	15,45	9,85	63,00

Fonte: Do Autor (2022)

pela observação da diferença de intensidade em diferentes regiões da área de estudo para ambas espécies. Na distribuição da *Copaifera langsdorffii*, foi identificado uma variação de 0,002 a 0,012 árvores/m², enquanto *Xylopiya brasiliensis*, com uma variação ainda mais severa de 0,005 a 0,25 árvores/m². Tais observações são importantes na condução da estimação dos efeitos de segunda ordem. Nas próximas análises, serão realizadas utilizando de descritores que permitam considerar a heterogeneidade presente na região em estudo.

Apesar de não se ter informações de possíveis fatores abióticos presentes na região, na prática, considera-se que os mapas da Figura 5.3 indicam a aparição de regiões que favorecem a existência de árvores. Tais fatores está relacionado com a presença de: rios ou lagos, variação da

Figura 5.3 – Distribuição espacial das estimativas de intensidades por meio do alisamento de kernel para *Copaifera langsdorffii* (esquerda), *Xylopia brasiliensis* (direita).



Fonte: Do Autor (2022)

Tabela 5.2 – 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 (2022)

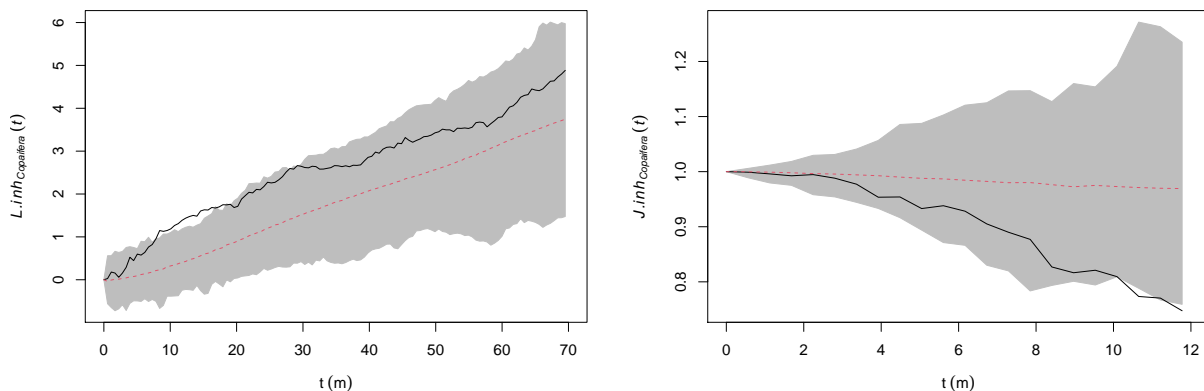
altitude ou relevo e entre outros fatores que poderiam influenciar a forma com que as árvores estão distribuídas.

5.3 Análise de segunda ordem caso não-homogêneo (Unviaridade)

A caracterização da dependência espacial entre as árvores, foi feita através da análise dos efeitos de segunda ordem obtida pela estimação das funções não homogêneas $L_{inhom}(t)$ e $J_{inhom}(t)$. Tais funções também foram utilizadas e recomendadas por Lieshout (2010), Baddeley, Rubak e Turner (2015), Diggle (2013), Sousa e Scalon (2018), Pottker et al. (2016), Frade (2014), Møller e Waagepetersen (2017). Conforme Marcon e Puech (2009), a dificuldade aparece na estimativa das intensidades locais. Matematicamente, pode ser indistinguível um fenômeno agrupado de um processo heterogêneo. Tal discussão pode ser encontrada em Diggle (1985), Cressie (1993), Sousa e Scalon (2018). No trabalho foi utilizado da suavização de kernel obtida na Figura 5.3 (CRONIE; LIESHOUT, 2018).

Pela análise dos gráficos da Figura 5.4 é possível observar que a hipótese nula de completa aleatoriedade espacial foi rejeitada para *Copaifera langsdorffii*. Através da análise do comportamento da função $L_{inhom}(t)$, representada pela Figura 5.4(esquerda), observa-se que a rejeição da hipótese nula acontece para distâncias variando aproximadamente de 10 a 20 metros, identificando um padrão agregado. Pela função $J_{inhom}(t)$, observa-se que a rejeição da hipótese nula acontece para distâncias acima de aproximadamente, 10 metros. Identificou-se também, um padrão agregado para esta espécie (Figura 5.4(direita)).

Figura 5.4 – Envelopes de simulações Monte Carlo das funções univariadas não homogêneas $L_{inhom}(t)$ (esquerda) e $J_{inhom}(t)$ (direita) para *Copaifera langsdorffii*. 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.

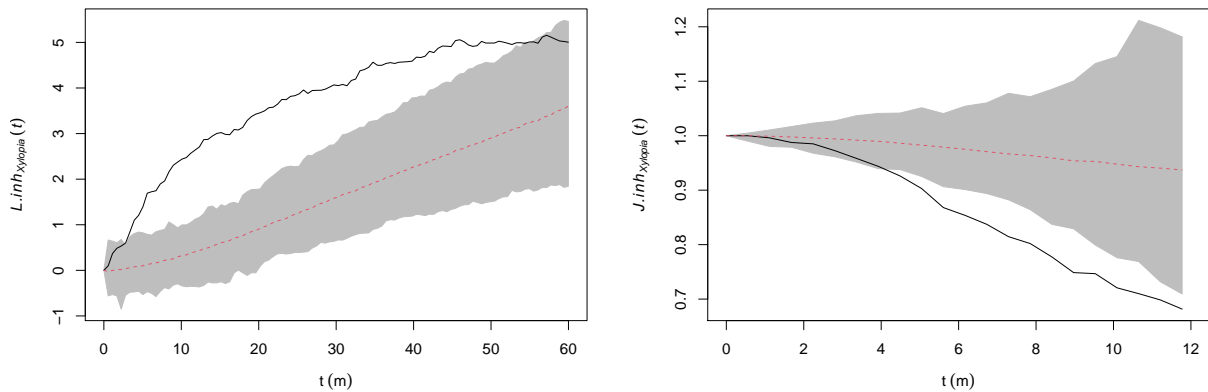


Fonte: Do Autor (2022)

Em alguns estudos com a espécie *Copaifera langsdorffii*, os resultados corroboram com este trabalho, identificando também o padrão agregado. Tais como: Na caracterização da estrutura, dispersão e distribuição espacial na floresta nacional do Araripe (VIEIRA et al., 2021). Na análise de um remanescente de Floresta Estacional Semidecidual em Lavras, MG Arruda e Daniel (2007). Na investigação de padrões espaciais em um fragmento do bioma cerrado Souza et al. (2020).

O padrão agregado também é identificado para espécie *Xylopia brasiliensis*. Pela função $L_{inhom}(t)$, Figura 5.5(esquerda) deve-se rejeitar a hipótese de CAE para distâncias entre aproximadamente 5 e 55 metros. Para as demais distâncias analisadas identificou-se aleatoriedade. Já pela função $L_{inhom}(t)$, Figura 5.5(direita), os agrupamentos acontecem para distâncias acima de 4 metros.

Figura 5.5 – Envelopes de simulações Monte Carlo das funções univariadas não homogêneas $L_{inhom}(t)$ (esquerda) e $J_{inhom}(t)$ (direita) para *Xylopia brasiliensis*. 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: Do Autor (2022)

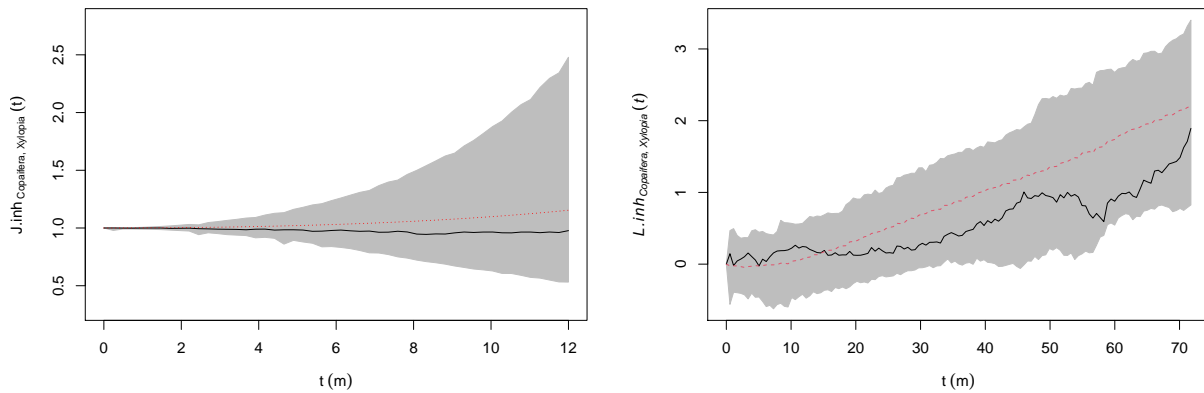
Os resultados obtidos neste trabalho corroboram com estudos de Higuchi et al. (2010) que detectou um padrão agregado no estudo da distribuição espacial de uma população de *Xylopia brasiliensis* em uma área experimental de floresta Semidecidual Montana localizada em Lavras, MG.

Com base nos resultados obtidos em ambas análises para caracterização espacial, pode-se dizer que o padrão agregado parece ser uma característica de ambas espécies em estudo. Tais considerações corroboram com Legendre e Fortin (1989), que destaca o padrão agregado como um dos mais recorrentes na distribuição espacial de espécies vegetais nativas. Diversas podem ser as causas que levam neste padrão. Scalon, Oliveira e Mello (2012) afirmam que tais conclusões podem ter relações com o tamanho e localização do fragmento florestal em estudo. Áreas pequenas e isoladas, torna-se vulnerável a exploração ilegal e portanto pode ser um dos fatores responsáveis pelo padrão agregado detectado. Autores como Vieira et al. (2021), Pottker et al. (2016), Souza (2020), entre outros, defendem que tal padrão remete a persistência da heterogeneidade microambientais dos recursos disponíveis.

5.4 Análise de segunda ordem caso não homogêneo (Bivariadado)

A análise de interação espacial entre as espécies *Copaifera langsdorffii* e *Xylopia brasiliensis* foi realizada pelas estimativas das funções $\hat{J}_{12}^{inhom}(t)$ e $\hat{L}_{12}^{inhom}(t)$ bivariadas não homogêneas, conforme apresentado pelas Equações (3.39) e (3.48) respectivamente. A hipótese testada na análise, refere-se a independência entre os subprocessos (as duas espécies) em questão. A ponderação de intensidades, em correção da heterogeneidade espacial para ambas espécies, foi realizada utilizando das estimativas de kernel obtidas pela Figura 5.3 (CRONIE; LIESHOUT, 2018). Os resultados são apresentados pela Figura 5.6.

Figura 5.6 – Envelopes de simulações Monte Carlo das funções $\hat{J}_{12}^{inhom}(t)$ (esquerda) e $\hat{L}_{12}^{inhom}(t)$ (direita) não homogênea bivariadas para interação *Copaifera langsdorffii* e *Xylopia brasiliensis*. 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: Do Autor (2022)

Os resultados apontam que não se deve rejeitar a hipótese de independência entre as espécies para todas escalas testadas. Tais resultados são contrários aos obtidos por Machado et al. (2012), que analisou a relação existente entre o padrão espacial das espécies nativas *Araucaria Angustifolia*, *Casearia Sylvestris* e *Cedrela Fissilis* de um fragmento floresta Ombrófila Mista Montana localizada no Campus Jardim Botânico da UFPR, Curitiba-PR. Para os mesmos autores, a disponibilidade de luz do sol e de nutrientes do solo são fatores que podem afetar a interação espacial de espécies que possuem características ecológicas semelhantes.

Pelo trabalho de Almeida (2018), também rejeitou-se a hipótese de independência espacial entre as espécies vegetais nativas *Dipteryx Odorata*, *Apuleia Leiocarpa* e *Ceiba Samauma*, indicando uma tendência de repulsão. A autora complementa que tais resultados podem estar associados à competição por recursos abaixo e acima do solo, mas que não se deve basear somente em tal análise para tirar conclusões. Os resultados deste trabalho ainda não corroboram com os padrões agregados encontrados na maioria dos resultados de Capretz (2004), na análise da interação espacial duas a duas entre as espécies arbóreas nativas *Copaifera Langsdorffi*, *Vochysia Tucanorum*, *Xylopia Aromatica* e *Ocotea Corymbosa*, localizadas em Unidades de Conservação do Estado de São Paulo. Para este autor, o grau de heterogeneidade do ambiente na disponibilidade de recursos pode influenciar a distribuição das espécies na formações de agrupamentos.

Vale lembrar que, diferentemente das análises realizadas neste trabalho, em muitos estudos encontrados na literatura, considera-se que as regiões de análise são homogêneas. Porém este é um fato contestável. Condit et al. (2002) afirma que é natural que a composição de floresta tropical apresente estruturas heterogêneas. Assim, ao valer-se da não-homogeneidade na análise de interações planta-planta, é comum que padrões não aleatórios seja detectado com menor frequência ou em escalas bastante locais (WIEGAND et al., 2007).

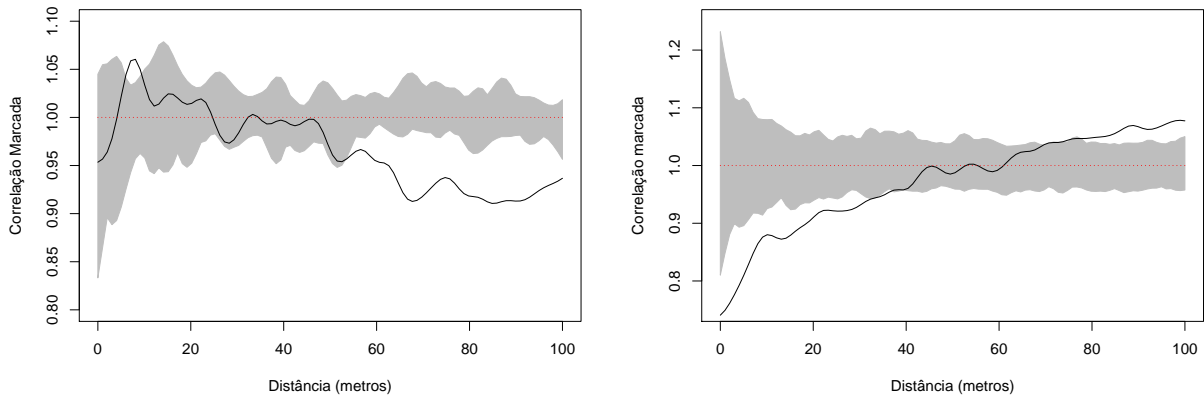
Apesar de apresentar resultados discordantes com grande parte de estudos de interação espacial de espécies de vegetação nativa encontrados na literatura, não é incomum a presença da aleatoriedade para o tipo de análise proposta por este trabalho. Em associação, pode-se dizer que a falta de interação entre tais espécies pode estar relacionada com a abundância de recursos obtidos na região. Assim, pode ser que a competição interespecífica não esteja ocorrendo ou então ocorre em baixas escalas (RODRIGUES, 2018).

5.5 Análise de segunda ordem (marcas contínuas)

A análise da distribuição espacial do diâmetro à altura do peito das árvores foi conduzida utilizando a função de correlação marcada $\rho_f(t)$ apresentada na seção 3.3.2.1, calculada para raios de 0 a 100 metros. Utilizou-se do procedimento visto em 3.4, testando hipótese nula de independência entre as localizações das árvores (pontos) e DAPs (marcas). Os resultados são apresentados pela Figura 5.7 e apresentam evidências para rejeição da hipótese nula. Isso indica uma associ-

ação entre marcas e pontos, ou seja, descreve um certo tipo de correlação entre as marcas. Tais comportamentos foram contrários para ambas espécies.

Figura 5.7 – 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: Do Autor (2022)

O gráfico da Figura 5.7 (direita) refere-se a estimativa da função de correlação marcada para a espécie *Xylopia brasiliensis*. Em 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 próximo de 1 (associação nula) para raios intermediários e para raios maiores (acima de 80 metros), $\rho_f(t)$ assume valores altos (associação positiva). Tal comportamento indica que árvores com diâmetros menores que a média tendem estar próximas e a medida que o DAP assume valores superiores que a média, tendem estar distantes. De acordo com Stoyan e Penttinen (2000), este comportamento é comum e frequentemente observado em análises de correlação de marcas.

Entretanto, para a espécie *Copaifera langsdorffii* (Figura 5.7esquerda), observou-se resultados diferentes. Em análise, têm-se que a função $\rho_f(t)$ assume valores altos (associação positiva) para raios entorno de 10 metros e baixos (associação negativa) para escalas superiores a 55 metros. Este comportamento indica que árvores que possuem DAPs maiores que a média estão agrupadas em escalas de aproximadamente 10 metros. A medida que o DAP assume valores menores, as escalas aumentam consideravelmente.

As diferenças encontradas na análise gráfica de ambas espécies podem ter relação com o grau de heterogeneidade ou nível de agregação da espécie (DALMASO et al., 2016). De acordo com Machado et al. (2012), a tendência de agregação de árvores jovens e sua afinidade com árvores adultas é resultado direto da limitada capacidade de dispersão, o que faz com que a maioria das sementes e frutos produzidos fique no solo próximo à árvore de origem.

6 CONCLUSÕES

Os resultados obtidos neste trabalho mostram que os métodos de processos pontuais para análises dos efeitos de primeira e segunda ordem, atrelados aos métodos Monte Carlo, permitiram uma análise detalhada da distribuição e interação espacial das árvores das espécies *Copaifera langsdorffii* e *Xylopia brasiliensis* no fragmento florestal.

As análises univariadas mostraram que as duas espécies apresentam árvores que estão distribuídas de forma não homogênea e apresentam dependência espacial (agregação) no fragmento florestal. Pela análise bivariada constatou-se que não houve indícios para rejeitar a hipótese nula de independência entre as duas espécies. Indicando que não há competição entre tais espécies.

Através da função de Correlação Marcada, foi possível detectar a dependência entre a localização das árvores e os DAPs das mesmas. Entretanto, as duas espécies apresentam comportamentos diferentes nessa dependência. A medida que os DAPs das árvores da espécie *Copaifera langsdorffii* aumentam, a associação passa de positiva para negativa, enquanto para a espécie *Xylopia brasiliensis* ocorre o contrário.

Por fim, é válido ressaltar que existem muitas lacunas que ainda devem ser preenchidas sobre o padrão de distribuição e interação das espécies estudadas. Para tais, propõe-se a realizações de estudos futuros que permitam obter maior riqueza nos detalhes.

- ALMEIDA, E. **Métodos de Processos Pontuais para análise de Interação entre Árvores de Espécies Nativas da Amazônia**. 65 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) — Universidade Federal de Lavras, 2018.
- ANDERSEN, M. Spatial analysis of two-species interactions. **Oecologia**, Springer, v. 91, n. 1, p. 134–140, 1992.
- ANJOS, A. d. et al. Análise do padrão de distribuição espacial da araucária (*Araucaria Angustifolia*) em algumas áreas no Estado do Paraná, utilizando a função K de Ripley. **Embrapa Florestas-Artigo em periódico indexado (ALICE)**, Scientia Forestalis, Piracicaba, n. 66, p. 38-45, 2004.
- ARRUDA, L.; DANIEL, O. Florística e diversidade em um fragmento de floresta estacional semidecidual aluvial em dourados, ms. **Floresta**, v. 37, n. 2, 2007.
- BADDELEY, A. et al. **Analysing spatial point patterns in R**. [S.l.]: CSIRO, 2008.
- BADDELEY, A.; RUBAK, E.; TURNER, R. **Spatial point patterns: methodology and applications with R**. [S.l.]: CRC press, 2015.
- BADDELEY, A. et al. spatstat: Spatial Point Pattern analysis, model fitting and simulation. R package version. Citeseer, 2005.
- BADDELEY, A. J. et al. Estimating the J function without edge correction. **Statistica Neerlandica**, Wiley Online Library, v. 54, n. 3, p. 315–328, 2000.
- BADDELEY, A. J.; LIESHOUT, M. N. M. **A nonparametric measure of spatial interaction in point patterns**. [S.l.]: Vrije Universiteit, Faculteit der Wiskunde en Informatica, 1993.
- BERG, J. V. D. A remark on the van lieshout and baddeley j-function for point processes. 1997.
- CAPRETZ, R. L. **Análise dos padrões espaciais de árvores em quatro formações florestais do estado de São Paulo, através de análises de segunda ordem, como a função K de Ripley**. Tese (Doutorado) — Universidade de São Paulo, 2004.
- CARVALHO, P. E. R. **Espécies arbóreas brasileiras**. Brasília, DF: Embrapa Informação Tecnológica; Colombo: Embrapa Florestas, 2014.
- CONDIT, R. et al. Beta-diversity in tropical forest trees. **Science**, American Association for the Advancement of Science, v. 295, n. 5555, p. 666–669, 2002.
- CRESSIE, N. **Statistics for Spatial Data: Wiley Series in Probability and Statistics**. [S.l.]: Wiley-Interscience, 1993. 908 p.
- CRONIE, O.; LIESHOUT, M. N. M. V. A non-model-based approach to bandwidth selection for kernel estimators of spatial intensity functions. **Biometrika**, Oxford University Press, v. 105, n. 2, p. 455–462, 2018.

- CRONIE, O.; LIESHOUT, M. N. M. van. Summary statistics for inhomogeneous marked point processes. **Annals of the Institute of Statistical Mathematics**, Springer, v. 68, n. 4, p. 905–928, 2016.
- CUNHA, U. S. d. **Análise da estrutura espacial horizontal de uma floresta de terra firme da Amazônia**. Dissertação (Doutorado) — Universidade Federal do Paraná, 2013.
- 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.
- DANTAS, A. A. A.; CARVALHO, L. G. d.; FERREIRA, E. Classificação e tendências climáticas em lavras, mg. **Ciência e Agrotecnologia**, SciELO Brasil, v. 31, p. 1862–1866, 2007.
- DIGGLE, P. A kernel method for smoothing point process data. **Journal of the Royal Statistical Society: Series C (Applied Statistics)**, Wiley Online Library, v. 34, n. 2, p. 138–147, 1985.
- DIGGLE, P. J. **Statistical analysis of spatial and spatio-temporal point patterns**. [S.l.]: CRC press, 2013.
- DIGGLE, P. J. et al. An introduction to model-based geostatistics. In: **Spatial statistics and computational methods**. [S.l.]: Springer, 2003. p. 43–86.
- DUARTE, E. et al. Estrutura populacional de drimys angustifolia miers em um fragmento de floresta ombrófila mista alto-montana, urubici, sc. **Revista de Ciências Agroveterinárias**, v. 18, p. 30–37, 2019.
- FRADE, D. D. R. **Relações entre fatores ambientais e espécies florestais por metodologias de processos pontuais**. Tese (Doutorado) — Universidade de São Paulo, 2014.
- FREITAS, C. V.; OLIVEIRA, P. E. Biologia reprodutiva de copaifera langsdorffii desf.(leguminosae, caesalpinioideae. **Brazilian Journal of Botany**, SciELO Brasil, v. 25, p. 311–321, 2002.
- GATRELL, A. C.; BAILEY, T. Interactive spatial data analysis in medical geography. **Social science & medicine**, Elsevier, v. 42, n. 6, p. 843–855, 1996.
- HIGUCHI, P. et al. Spatial patterns of a tropical tree species growing under an eucalyptus plantation in south-east brazil. **Brazilian Journal of Biology**, SciELO Brasil, v. 70, p. 271–277, 2010.
- ILLIAN, J. et al. **Statistical analysis and modelling of spatial point patterns**. [S.l.]: John Wiley & Sons, 2008. v. 70.
- JUNIOR, V. F. V.; PINTO, A. C. O gênero copaifera l. **Química nova**, SciELO Brasil, v. 25, p. 273–286, 2002.
- KIM, J.; SCOTT, C. D. Robust kernel density estimation. **The Journal of Machine Learning Research**, JMLR. org, v. 13, n. 1, p. 2529–2565, 2012.

- LEGENDRE, P.; FORTIN, M. J. Spatial pattern and ecological analysis. **Vegetatio**, Springer, v. 80, n. 2, p. 107–138, 1989.
- LIESHOUT, M. V. A j -function for inhomogeneous point processes. **Statistica Neerlandica**, Wiley Online Library, v. 65, n. 2, p. 183–201, 2010.
- LIESHOUT, M. V.; BADDELEY, A. A nonparametric measure of spatial interaction in point patterns. **Statistica Neerlandica**, Wiley Online Library, v. 50, n. 3, p. 344–361, 1996.
- LIESHOUT, M. V.; BADDELEY, A. J. Indices of dependence between types in multivariate point patterns. **Scandinavian Journal of Statistics**, Wiley Online Library, v. 26, n. 4, p. 511–532, 1999.
- LOADER, C. **Local regression and likelihood**. [S.l.]: Springer Science & Business Media, 2006.
- LORENZI, H. Árvores brasileiras: manual de identificação e cultivo de plantas arbóreas nativas do Brasil. Plantarum Nova Odessa, 1992.
- MACHADO, S. d. A. et al. **Distribuição espacial de um fragmento de Floresta Ombrófila Mista Montana**. [S.l.]: SciELO Brasil, 2012.
- MARCON, E.; PUECH, F. Generalizing Ripley's k function to inhomogeneous populations. 2009.
- 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.
- MONTEIRO, A. M. V. et al. Análise espacial de dados geográficos. **Brasília: Embrapa**, 2004.
- NUNES, Y. R. F. et al. Variações da fisionomia, diversidade e composição de guildas da comunidade arbórea em um fragmento de floresta semidecidual em lavras, mg. **Acta botânica brasilica**, SciELO Brasil, v. 17, p. 213–229, 2003.
- OLINDA, R. A. de; SCALON, J. D. Métodos de Monte Carlo para análise de processos pontuais marcados. **Rev. Bras. Biom**, v. 28, n. 1, p. 39–56, 2010.
- POTTKER, G. S. et al. Padrão espacial de espécies florestais: estudo de caso com *Ocotea odorifera* (Vell.) Rohrer. **Ciência Florestal**, SciELO Brasil, v. 26, p. 1097–1106, 2016.
- R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2022. Disponível em: <<https://www.R-project.org/>>.
- RIPLEY, B. D. Modelling spatial patterns. **Journal of the Royal Statistical Society: Series B (Methodological)**, Wiley Online Library, v. 39, n. 2, p. 172–192, 1977.
- RODRIGUES, B. L. Padrões espaciais e quantificação de árvores comerciais na Amazônia. UFVJM, 2018.
- 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.

SCHABENBERGER, O.; GOTWAY, C. A. **Statistical methods for spatial data analysis**. [S.l.]: Chapman and Hall/CRC, 2005.

SOUSA, S. B. de; SCALON, J. D. Funções I homogênea e não homogênea em testes de hipótese de completa aleatoriedade espacial. **Revista da Estatística da Universidade Federal de Ouro Preto**, v. 7, 2018.

SOUZA, P. B. de et al. Population structure and spatial distribution of 'copaifera langsdorffii' desf. **Australian Journal of Crop Science**, Southern Cross Publishers Lismore, NSW, v. 14, n. 3, p. 3000–3005, 2020.

SOUZA, R. L. F. de. Padrão espacial da castanheira em floresta no sul do amapá, amazônia oriental. **Revista Arquivos Científicos (IMMES)**, v. 3, n. 1, p. 67–72, 2020.

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

TERRELL, G. R.; SCOTT, D. W. Variable kernel density estimation. **The Annals of Statistics**, JSTOR, p. 1236–1265, 1992.

VIEIRA, D. dos S. et al. Estrutura diamétrica e distribuição espacial de dipteryx odorata (aubl.) willd. no oeste do estado do Pará, Brasil. **Scientia Florestalis**, 2021.

WAND, M. P.; JONES, M. C. **Kernel smoothing**. [S.l.]: CRC press, 1994.

WIEGAND, T. et al. Species associations in a heterogeneous sri lankan dipterocarp forest. **The American Naturalist**, The University of Chicago Press, v. 170, n. 4, p. E77–E95, 2007.

YAMAMOTO, J. K.; LANDIM, P. M. B. **Geostatística: conceitos e aplicações**. [S.l.]: Oficina de textos, 2015.

A ANEXOS - CÓDIGOS EM R

A.1 Leitura dos dados

```
### Pacotes utilizados ###
library(spatstat)
library(readxl)
library(dplyr)
library(geobr)
library(tmap)
library(sf)
library(mapttools)
library(ggplot2)
library(sp)

### Lendo dados ###
Dados <- read_xlsx("C:\\Users\\Desktop\\MESTRADO\\DADOS_MATINHA.xlsx")

### Filtrando dados por especie ####
Dadosc <- filter(Dados, Especie=="Copaifera langsdorffii")
Dadosx <- filter(Dados, Especie=="Xylopiya brasiliensis")

### Janela de estudo ###
M= rgdal::readOGR("C:\\Users\\Desktop\\Shape matinha\\Matinha.shp")
SP=as(M, "SpatialPolygons")
janela=as(SP, "owin")
janelal=as(SP, "sf")
```

A.2 Códigos da análise de primeira ordem

```
### Kernel de cada uma das especies ###
### obtendo shapes
sh.cop <- ppp ( Dadosc$X , Dadosc$Y , window = janela)
sh.xyl <- ppp ( Dadosx$X , Dadosx$Y , window = janela)

### Suavizacao de Kernel para Copaifera
Kernelc <- density(sh.cop, sigma = bw.CvL, kernel="gaussian",
diggle =TRUE, positive = TRUE)
plot (Kernelc, main = NULL)
plot ( sh.cop, add = TRUE, cex = 0.5)

### Suavizacao de Kernel para Xylophia
Kernelx <- density(sh.xyl, sigma=bw.CvL, kernel="gaussian",
diggle =TRUE, positive = TRUE )
plot ( Kernelx , main = NULL)
plot ( sh.xyl , add = TRUE, cex = 0.5)
```

A.3 Códigos da análise de segunda ordem (univariado)

```

### simulando J nao homogenea
Jinhomc <- envelope ( shape.cop, Jinhom, nsim =1000, correction = "border",
simulate = expression(rpoispp(Kernelc)))

Jinhomx <- envelope ( shape.xyl , Jinhom , nsim =1000, correction = "border",
simulate = expression(rpoispp(Kernelx)))

### graficos J nao homogenea
plot (Jinhomc, legend=F, xlim=c(0,12), main=NULL,
xlab=expression(t~("m")), ylab=expression(J.inh["Copaifera"]~(t)))

plot (Jinhomx, legend=F, xlim =c(0,12), main=NULL,
xlab=expression(t~("m")), ylab=expression(J.inh["Xylopia"]~(t)))

### simulando L nao homogenea
Linhomc <- envelope ( shape.cop , Linhom , nsim =1000, correction = "border",
simulate = expression(rpoispp(Kernelc)))

Linhomx <- envelope ( shape.xyl , Linhom , nsim =1000, correction = "border",
simulate = expression ( rpoispp (Kernelx)))

### grafico L nao homogenea
plot (Linhomc, legend=F, xlim=c(0,70), .-r~r, main = NULL,
xlab=expression(t~("m")), ylab=expression(L.inh["Copaifera"]~(t)))
plot (Linhomx, legend=F, xlim=c(0,60), .-r~r, main = NULL,
xlab=expression(t~("m")), ylab=expression(L.inh["Xylopia"]~(t)))

```

A.4 Códigos da análise de segunda ordem (bivariado)

```

### Lendo shapes ###
Dadoscx <- rbind(Dadosc, Dadosx)
shapecx <- ppp(Dadoscx$X,Dadoscx$Y, window = janela,
              marks = as.factor(Dadoscx$Especie))

Pm.xyl <- ppp(Dadosx$X,Dadosx$Y, window = janela,
             marks = as.factor(Dadosx$Especie))
Pm.cop <- ppp(Dadosc$X,Dadosc$Y, window = janela,
             marks = as.factor(Dadosc$Especie))

### Kernel
Kernelcx<-density(split(shapecx), sigma=bw.CvL, diggle =TRUE
kernel="gaussian", positive = TRUE )

### L marcada nao homogenea
Lmarc.inh.cx<-envelope(Pm,Lcross.inhom, nsim = 1000, correction = "border",
i="Copaifera langsdorffii", j="Xylopia brasiliensis",
simulate = expression(rmpoispp(Kernelcx,types=names(Kernelcx))))

plot(Lmarc.inh.cx,.-r~r,legend=F, main = " ",xlab = expression (t~(m)),
ylab = expression (L.inh["Copaifera, Xylopia"] ~( t )))

### J marcada nao homogenea
# Lembrando que: Jinh=(1-Ginh)/(1-Finh).

## F mult inhomogeneous
r<- seq(0,12,length = 50)
y<- seq(0,1,length=50)

```



```

jx <- (marks(Pm.xyl) == "Xylophia brasiliensis")

lambda.xyl<-density(Pm.xyl,at = "point", sigma=bw.CvL, kernel="gaussian",
diggle =TRUE, positive = TRUE )
lambdax.min<- min(lambda.xyl)*0.9

fs<-FmultiInhom(Pm.xyl, jx, lambda=lambda.xyl, lambdamin=lambdax.min, r=r)

## G mult inhomogeneous
lambda.cx<- density(split(shapecx), at = "point",sigma=bw.CvL,
kernel="gaussian", diggle =TRUE, positive = TRUE )

lambdaj.cx<- c(lambda.cx$`Copaifera langsdorffii`,
               lambda.cx$`Xylophia brasiliensis`)
lminimo<- min(lambdaj.cx)*0.9

icx <- (marks(shapecx) == "Copaifera langsdorffii")
jcx <- (marks(shapecx) == "Xylophia brasiliensis")

gs<-GmultiInhom(shapecx, icx, jcx, lambda=lambdaj.cx, lambdamin=lminimo, r=r)

### J mult inhemogeneous

jobs<-(1-gs$bord)/(1-fs$bord)
jteo<-(1-gs$theo)/(1-fs$theo)

# Bandas de confianca com 1000 simulacoes

nsim<-1000

```

```

jmc<- matrix(0,length(r), nsim)
jmcteo<- matrix(0,length(r), nsim)

for(i in (1:nsim)) {
ker<-density(split.ppp(Pm.xyl), sigma=bw.CvL, kernel="gaussian",
diggle =TRUE, positive = TRUE )
pontos9<-rmpoispp ( ker, types=names(ker))

l7<-density(split(pontos9), at = "point",sigma=bw.CvL, kernel="gaussian",
diggle =TRUE, positive = TRUE )
l9<-c(l7$`Xylophia brasiliensis `)

lmin<- min(l9)*0.9

J <- (marks(pontos9) == "Xylophia brasiliensis")

f<-FmultiInhom(pontos9, J, lambda=19, lambdamin=lmin, r=r)

ker2<-density(split(Pm),sigma=bw.CvL, kernel="gaussian",
diggle =TRUE, positive = TRUE )
pontos92<-rmpoispp ( ker2, types=names(ker2))

lambdaj2<- density(split(pontos92), at = "point",sigma=bw.CvL,
kernel="gaussian",diggle =TRUE, positive = TRUE )

l52<- c(lambdaj2$`Copaifera langsdorffii`, lambdaj2$`Xylophia brasiliensis `)
lminimo2<- min(l52)*0.9

i2 <- (marks(pontos92) == "Copaifera langsdorffii")
j2 <- (marks(pontos92) == "Xylophia brasiliensis")

```

```

g<-GmultiInhom(pontos92, i2, j2, lambda=152, lambdamin=lminimo2, r=r)

#####

jmc[,i]<-(1-g$bord)/(1-f$bord)
jmcteo[,i]<-(1-g$theo)/(1-f$theo)

}
mn<- apply(jmcteo,1,mean)

limax<- apply(jmc,1,max)
limin<- apply(jmc,1,min)

y<- seq(0.3,2.8, length=length(r))

plot(r,y, type = "n", xlim=c(0,12), xlab = expression (t~( m)),
ylab = expression (J.inh [ " Copaifera, Xylophia" ] ~( t )))

polygon(c(rev(r),r),c(rev(limax),limin),col="gray", border = "gray")
lines(r, jobs)
lines(r,mn, lty=3, col="red")

```

A.5 Códigos da análise de Correlação Marcada

```
##### SIMULACAO DE FUNCAO DE CORRELACAO MARCADA #####
#### shape de pontos
marcasc<- c(Dadosc$DAP2017)
shapec <- ppp ( Dadosc$X , Dadosc$Y , window = janela , marks = marcasc)

marcasx<- c(Dadosx$DAP2017)
shapex <- ppp ( Dadosx$X , Dadosx$Y , window = janela , marks = marcasx)

### Simulando a funcao de correlacao marcada para copaifera
markcor<- function(event=shapec, nsim = 1000){
r<- seq(0, 100, length=100)
area<- area(janela)
lengt<- length(r)
np<- length(event$x)
y<- seq(0,1.5,length=lengt)
mmin<- min(event$marks)
mmax<- max(event$marks)

ko<- markcorr(event, f = function(m1,m2){ m1 * m2},
correction= "Ripley", r=r, method= "density")

#obtendo as simulacoes monte carlo
holdc<- matrix(0, nsim,lengt)

for(i in (1:nsim)) {X<- rpoispp(np/area(janela),win=janela)
spts<- X %mark% runif(X$n, min=mmin, max=mmax)
k <- markcorr(spts, f = function(m1,m2) { m1 * m2 },
correction="Ripley", r=r, method="density")
holdc[i,] <- k$iso}}
```

```

mn<- apply(hold,2,mean)
up<- apply(hold,2,max)
dow<- apply(hold,2,min)

### construindo o grafico
plot(r, y, type = "n", xlab = "Distancia (metros)", ylim = c(0.8,1.1),
ylab = "Correlacao Marcada")

polygon(c(rev(r),r),c(rev(up),dow),col="gray", border = "gray")
lines(r, ko$iso)
lines(r, ko$theo, lty=3, col="red")
#####

#### simulando correlacao marcada para xylophia

markcor<- function(event=shapex, nsim = 1000){
r<- seq(0, 100, length=100)
area<- area(janela)
lengt<- length(r)
np<- length(event$x)
y<- seq(0,1.5,length=lengt)
mmin<- min(event$marks)
mmax<- max(event$marks)

ko<- markcorr(event, f = function(m1,m2){ m1 * m2},
correction= "Ripley", r=r, method= "density")

#obtendo simulacao monte carlo
holdx<- matrix(0, nsim,length)
for(i in (1:nsim)) {X<- rpoispp(np/area(janela),win=janela)

```

```
spts<- X %mark% runif(X$n, min=mmin, max=mmax)
k <- markcorr(spts,
f = function(m1,m2) { m1 * m2 },
correction="Ripley", r=r, method="density")
holdx[i, ] <- k$iso}}

mn<- apply(holdx,2,mean)
up<- apply(holdx,2,max)
dow<- apply(holdx,2,min)
#criando o grafico
plot(r, y, type = "n", ylim = c(0.75,1.25), xlab = "Distancia (metros)",
ylab = "Correlacao marcada")
polygon(c(rev(r),r),c(rev(up),dow),col="gray", border = "gray")
lines(r, ko$iso)
lines(r, ko$theo, lty=3, col="red")
```