



**EDERSON DE SOUZA**

**ANÁLISE DE TERCEIRA ORDEM EM CONFIGURAÇÕES  
PONTUAIS**

**LAVRAS-MG  
2024**

**EDERSON DE SOUZA**

**ANÁLISE DE TERCEIRA ORDEM EM CONFIGURAÇÕES PONTUAIS**

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  
2024**

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

Souza, Ederson de.

Análise de terceira ordem em configurações pontuais / Ederson  
de Souza. – 2018.

50 p. : il.

Orientador: João Domingos Scalon.

Dissertação (Mestrado acadêmico) – Universidade Federal de  
Lavras, 2018.

Bibliografia.

1. Função *T*. 2. Terceira ordem. 3. Processos pontuais. I. Scalon,  
João Domingos. II. Título.

**EDERSON DE SOUZA**

**ANÁLISE DE TERCEIRA ORDEM EM CONFIGURAÇÕES PONTUAIS**

**THIRD-ORDER ANALYSIS IN SPATIAL POINT PATTERNS**

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 27 de novembro de 2018.

Dr. Anderson Castro Soares de Oliveira UFMT

Dr. Renato Ribeiro de Lima UFLA

Prof. Dr. João Domingos Scalon  
Orientador

**LAVRAS – MG  
2024**

*Aos meus pais José (in memoriam) e Valdair, que dignamente me apresentaram a importância da família e o caminho da honestidade e persistência.*

## **AGRADECIMENTOS**

A DEUS, por me guiar durante esses dois anos de mestrado.

Aos meus pais José e Valdair, por todo amor e incentivo.

Ao meu irmão, Richardson pela força carinho.

Ao meu orientador, professor Dr. João Domingos Scalon, pelo apoio, pela confiança e pela paciência.

Ao professor Anderson do Departamento de Estatística da UFMT.

Aos amigos da UFLA, em especial, Cristian Tiago Mendes, Lucas Panissa, Mateus Aquino Gonçalves e Paulo Roberto Almeida.

Ao Pesquisador Evandro Orfano Figueiredo da EMBRAPA do Acre e ao Isaac, por conceder os dados e ajudar na confecção dos mapas.

Aos colegas de mestrado, pela amizade.

À Universidade Federal de Lavras e ao Departamento de Estatística, pela oportunidade.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, e CAPES pelo suporte financeiro que é a concessão da bolsa de mestrado.

E a todos que contribuíram direta e indiretamente no meu trabalho.

**MUITO OBRIGADO!**

*Sem a matemática não nos seria possível compreender  
muitas passagens das santas escrituras  
(Santo Agostinho).*

## RESUMO

A análise dos padrões de configurações pontuais no espaço têm como foco, em geral, caracterizar as propriedades de primeira e segunda ordem do processo estocástico que gerou os pontos. No entanto, para uma análise mais rigorosa, em algumas situações, é necessário analisar características de ordem superior. Este trabalho tem como finalidade apresentar um estudo sobre a função  $T$  como método para caracterizar a propriedade de terceira ordem em configurações pontuais no espaço. Essa caracterização pode ser útil para diferenciar processos pontuais que exibem as mesmas propriedades de primeira e segunda ordem, mas podem apresentar comportamentos diferentes na terceira ordem. Além disso, a função  $T$  pode ser usada para testar a independência entre eventos tendo como base a localização simultânea de três eventos. Os resultados mostram que a função  $T$  pode ser útil para caracterizar a distribuição espacial de árvores de espécies nativas da Floresta Amazônica.

**Palavras-chave:** Função  $T$ ; terceira ordem; processos pontuais.



## ABSTRACT

The analysis of spatial point patterns generally focuses on characterizing the first- and second-order properties of the stochastic process that generated the points. However, for a more rigorous analysis, in some situations, it is necessary to analyze higher order characteristics. This work aims to present a study on the  $T$ -function as a method to characterize the third-order property in spatial point patterns. This characterization can be useful to differentiate point processes that exhibit the same first- and second-order properties, but may exhibit different behaviors in the third-order. Furthermore, the  $T$  function can be used to test the independence between events based on the simultaneous location of three events. The results show that the  $T$ -function can be useful to characterize the spatial distribution of trees of native species in the Amazon Rain Forest.

**Keywords:** T-function; third-order; point processes.

## INDICADORES DE IMPACTO

Nos biomas brasileiros existem árvores de variadas espécies, sendo que algumas nascem mais próximas umas das outras, exibindo configurações aglomeradas, enquanto outras tendem a manter uma distância mínima, uma das outras, exibindo configurações regulares e outras que não apresentam um padrão bem definido e, portanto, são denominadas de configurações aleatórias. O grande problema é caracterizar estatisticamente esses tipos de configurações entre árvores de uma determinada espécie. Existem diferentes métodos estatísticos para realizar a caracterização da configuração espacial de árvores de espécies nativas. A grande maioria desses métodos tem como base distâncias entre pares de eventos. No entanto, para uma análise mais rigorosa da configuração pontual, em algumas situações, é necessário utilizar alguma medida de ordem superior, tendo como base as distâncias entre três eventos. Esse tipo de análise é raramente empregado na análise de configurações pontuais no espaço. Esta dissertação apresenta a função  $T$  como uma estatística importante para caracterizar efeitos de terceira ordem em configurações pontuais no espaço. Métodos baseados na função  $T$  foram aplicados para analisar as configurações pontuais duas espécies nativas da Floresta Amazônica: *Antrocaryon amazonicum* e *Bombax humile*. Foram utilizadas as localizações de 66 árvores da espécie *Antrocaryon amazonicum* e de 293 árvores da espécie *Bombax humile*. Os resultados mostram que a análise de terceira ordem baseada na função  $T$  pode ser extremamente útil para determinar as distâncias mínimas de plantio de árvores em trabalhos de reflorestamento de florestas nativas, o que acarreta um ganho de informação em comparação as análises baseadas em distâncias de pares de árvores.

## IMPACT INDICATORS

In Brazilian biomes there are trees of different species, with some growing closer to each other, exhibiting clustered patterns, while others tend to maintain a minimum distance from each other, exhibiting regular patterns and others that do not present a well-defined pattern and, therefore, they are called random patterns. The big problem is to statistically characterize these types of patterns between trees of a given species. There are different statistical methods to characterize the spatial patterns of native species trees. The vast majority of these methods are based on distances between pairs of events. However, for a more rigorous analysis of the spatial point pattern, in some situations, it is necessary to use some higher order measure, based on the distances between three events. This type of analysis is rarely used in spatial point pattern analysis. This dissertation presents the  $T$ -function as an important statistic for characterizing third-order effects in spatial point patterns. Methods based on the  $T$ -function were applied to analyze the spatial point patterns of two species native to the Amazon Rainforest: *Antrocaryon amazonicum* and *Bombax humile*. The locations of 66 trees of the species *Antrocaryon amazonicum* and 293 trees of the species *Bombax humile* were used. The results show that the third-order analysis based on the  $T$ -function can be extremely useful for determining the minimum tree planting distances in native forest reforestation work, which results in a gain in information compared to analyzes based on pairwise distances of trees

## LISTA DE FIGURAS

Figura 1 – Tipos de configurações pontuais no espaço. ....	21
Figura 2 – Fazenda Canary situada no município de Bujari, Acre. ....	32
Figura 3 – Distribuição das árvores na fazenda Canary situado no município de Bujari no estado do Acre da espécie <i>Antrocaryon amazonicum</i> . ....	35
Figura 4 – Distribuição das árvores na fazenda Canary situado no município de Bujari no estado do Acre da espécie <i>Bombaxhumile</i> . ....	35
Figura 5 – Distribuição das árvores na região de estudo (a) e gráfico de calor (b) para a espécie <i>Antrocaryon amazonicum</i> . ....	36
Figura 6 – Distribuição dos pontos na região de estudo (a) e gráfico de calor (b) para a espécie <i>Bombaxhumile</i> . ....	36
Figura 7 – Gráfico em 3D da densidade dos pontos na região de estudo. ....	37
Figura 8 – Gráfico dos quantis e o teste de Kolmogorov–Smirnov. ....	37
Figura 9 – Envelope de aleatoriedade para a função $G$ . ....	38
Figura 10 – Envelope de aleatoriedade para a função $F$ . ....	39
Figura 11 – Envelope de aleatoriedade para a função $J$ . ....	39
Figura 12 – Envelope de aleatoriedade para a função $K$ . ....	40
Figura 13 – Envelope de aleatoriedade para a função $L$ . ....	40
Figura 14 – Envelope de aleatoriedade para a função $T$ . ....	41

## SUMÁRIO

1	INTRODUÇÃO.....	13
2	OBJETIVOS .....	15
3	REFERENCIAL TEÓRICO .....	16
3.1	Análise espacial .....	16
3.2	Configuração pontual espacial .....	17
3.3	Propriedade de primeira ordem.....	18
3.3.1	Intensidade homogênea .....	18
3.3.2	O teste de Kolmogorov–Smirnov.....	18
3.3.3	Intensidade local .....	19
3.4	Propriedades de segunda ordem .....	20
3.4.1	Função $F$ .....	21
3.4.2	Função $G$ .....	22
3.4.3	Função $J$ .....	23
3.4.4	Função $K$ .....	24
3.4.5	Função $L$ .....	25
3.5	Envelopes de simulação .....	26
3.6	Propriedade de terceira ordem.....	27
3.6.1	Função $T$ .....	27
3.6.2	Valor teórico da função $T$ .....	29
3.6.3	Estimador da função $T$ .....	29
3.6.4	Método da fronteira.....	29
4	MATERIAIS E MÉTODOS .....	31
4.1	Dados.....	31
4.2	Análise estatística.....	32
5	RESULTADOS E DISCUSSÃO.....	34
6	CONCLUSÕES.....	42
	REFERÊNCIAS.....	43
	APÊNDICE A – códigos para as análises do conjunto das árvores, utilizando o software R.....	45

## 1 INTRODUÇÃO

A Estatística é uma ciência dedicada à coleta, análise e interpretação dos dados e, cada vez mais, está ganhando espaço nas áreas aplicadas, sendo um instrumento importante ao se tirar conclusões e tomar decisões. Para realizar determinados estudos, a Estatística pode ser dividida em diversas áreas, sendo uma delas a Estatística espacial.

A Estatística Espacial considera explicitamente a localização espacial do fenômeno em estudo na análise. Assim fornece ao pesquisador a capacidade de estabelecer critérios quantitativos para determinar o grau de tendência e dependência espacial entre as observações.

Muitos desses fenômenos são identificados como pontos localizados no espaço e são denominados de processos pontuais espaciais. São exemplos: localização de crimes em um bairro, ocorrência de doenças em uma determinada região, localização de espécies vegetais em uma floresta nativa, entre outras. Nesses fenômenos, o objeto da análise é caracterizar o processo estocástico que gerou a configuração pontual.

O primeiro passo à análise de uma configuração pontual é, geralmente, obter estimativas de várias estatísticas de resumo que caracterizam as propriedades de primeira e segunda ordem do processo estocástico. As propriedades de primeira ordem podem ser caracterizadas pela medida de resumo intensidade (número de pontos por unidade de área). Já as propriedades de segunda ordem podem ser caracterizadas por diversas funções que descrevem as interações existentes entre os pontos no espaço.

Um dos métodos mais utilizados para caracterizar a propriedade de primeira ordem é o estimador de intensidade não paramétrico do tipo kernel.

Uma das estatísticas de resumo mais popular, para caracterizar as propriedades de segunda ordem, é a função  $K$  de Ripley (Ripley, 1977), também conhecida como função reduzida de segundo momento. Essa função mede o número médio de pares de pontos, que se encontram dentro de uma particular distância, por unidade de área. Outras estatísticas de resumo utilizam a função de distribuição de distâncias entre vizinhos mais próximos, como as funções  $F$ ,  $G$  e  $J$  (Cressie, 1983; Diggle, 2003)

Não é difícil de esperar que apenas as estatísticas de primeira e segunda ordem podem não descrever completamente um processo pontual. Baddeley e Silverman (1984) construíram um exemplo de uma realização de um processo pontual que apresentava uma função  $K$  compatível com um processo de Poisson homogêneo. Entretanto o processo pontual, utilizado para gerar a realização da configuração pontual, era muito diferente do processo de Poisson

homogêneo. Esse resultado contraditório pode indicar que o processo apresenta um padrão de ordem superior a dois.

Não existe um método padrão, para caracterizar configurações de ordem superior a dois, embora estatísticas de terceira ordem tenham sido adaptadas para alguns problemas específicos. Hanisch (1983) e Hanisch e Stoyan (1984) usaram uma estatística de terceira ordem para detectar linearidades internas em configurações pontuais. Müller, Syversveen e Waagepetersen (1998) propuseram uma estatística de terceira ordem que distinguisse Processos log Cox–Gaussianos de outros tipos de processos pontuais. Ayala e Simó (2000) introduziram medidas de correlação, para processos pontuais trivariados, baseados em contagens de triplas de pontos.

Schladitz e Baddeley (1997) propuseram uma estatística de resumo, mais geral, para caracterizar propriedades de terceira ordem de uma configuração pontual. Essa estatística, denominada de função  $T$  é baseada no número médio de triplas que ocorrem dentro de uma particular distância. Esses autores determinaram a função  $T$  para o processo homogêneo de Poisson e para o processo de Gauss–Poisson. Além disso, eles também propuseram estimadores simples, corrigidos, para o efeito de borda, para a função  $T$ . Os autores ainda mostraram, por estudos de simulação, a capacidade da função  $T$  para discriminar entre diferentes tipos de processos pontuais.

## 2 OBJETIVOS

Este trabalho defende que, para uma completa caracterização do processo pontual espacial, além de explorar as propriedades de primeira e segunda ordem, deve-se explorar também as propriedades de terceira ordem. Assim, este trabalho tem três objetivos bem definidos:

- a) apresentar a função  $T$  para pontos triplos em processos pontuais;
- b) mostrar como a função  $T$  pode ser utilizada, de forma análoga às funções, para caracterizar as propriedades de segunda ordem, para analisar as propriedades de terceira ordem de configurações pontuais realizadas por diferentes processos pontuais espaciais; A hipótese a ser testada é que existem configurações pontuais que, apesar de apresentarem as mesmas propriedades de primeira e segunda ordem, podem ter diferentes propriedades de terceira ordem;
- c) aplicar a função  $T$  para determinar se existem efeitos de terceira ordem na distribuição territorial de árvores de espécies nativas da Floresta Amazônica.



### 3 REFERENCIAL TEÓRICO

A Estatística Espacial consiste no estudo, caracterização e modelagem de variáveis aleatórias que apresentam uma estrutura espacial ou espaço-temporal.

#### 3.1 Análise espacial

O objetivo da análise espacial é mensurar características e relações entre eventos, considerando sua localização espacial de forma explícita. Assim, a análise espacial pode ser definida como o estudo, exploração e modelagem de processos estocásticos que geram eventos distribuídos no espaço, usualmente, no  $R^2$  (Bailey; Gatrell, 1995).

A análise espacial pode ser aplicada, em distintas áreas da ciência, como Geologia, Ciência do Solo, Processamento de Imagens, Epidemiologia, Agronomia, Ecologia, Engenharia Florestal, Astronomia e, essencialmente, em qualquer disciplina que colete dados em diferentes localidades (Cressie, 1993). A análise de dados espaciais abrange a definição precisa dos dados que estão relacionados a um processo que opera no espaço. Portanto trabalha na exploração de padrões e relacionamentos em tais dados e na procura por explicações dos padrões e relações encontradas (Bailey; Gatrell, 1995). Para tanto, a estatística espacial abrange uma coleção de métodos de análise voltados à caracterização de padrões e relações espaciais dos dados.

Dada a grande abundância de áreas de conhecimento e aplicações, existe também uma grande quantidade de tipos de dados espaciais (Schabenberger; Gotway, 2005). Cressie (1993) considera os dados espaciais, em três principais tipos de categorias, dependendo da natureza estocástica da observação: dados geoestatísticos, dados de área e padrões pontuais.

- a) **padrões pontuais:** esse tipo de dado pode representar fenômenos identificados como um conjunto de localizações pontuais, em que os eventos do fenômeno acontecem em uma área de estudo  $R$  com dimensão  $d$ . Apesar de ser possível analisar dados em dimensões  $d$  maior que 2, como, por exemplo, a representação pontual de galáxias no espaço, a maior parte se limita a casos bidimensionais. Como exemplos temos as localizações de espécies de plantas, crimes, doenças, etc. Ao conjunto de dados dá-se o nome de padrão (configuração) pontual espacial. As localidades dos fenômenos são chamadas de eventos, para distingui-los de pontos arbitrários localizados na região em questão (Diggle, 2003);

- b) **dados geoestatísticos:** nesse caso, o domínio  $D$  é um conjunto fixo, contínuo. Por conjunto contínuo entende-se que o valor do fenômeno analisado pode ser encontrado em qualquer local dentro de  $D$ , ou seja, entre duas localidades amostradas existe um número infinito de outras amostras. Por conjunto fixo entende-se que os pontos amostrados em  $D$  são não estocásticos. Outra ressalva é que a continuidade está relacionada ao domínio e não ao atributo que está sendo mensurado (Schabenberger; Gotway, 2005). Exemplos de fenômenos que envolvem dados geoestatísticos, coletados em locais pré-especificados dentro de uma região, tais como, temperatura, pressão atmosférica, teor de cálcio, etc. Observa-se que, em todos os casos, a mensuração da variável de interesse pode ser realizada em qualquer ponto dentro da região de estudo;
- c) **dados de área:** esse tipo de dado refere-se a fenômenos que, apesar de poderem se difundir de forma contínua no espaço são analisados em sub-regiões que têm uma composição da partição na região de interesse. Nesse caso, o pesquisador dispõe de avaliações individuais agregadas em cada sub-região e tem como interesse detectar padrões subjacentes ou quantificar o grau de correlação espacial do fenômeno dentro da região de estudo. Número de casos de doenças ou crimes agregados por bairro, municípios ou estados são exemplos de dados de área.

### 3.2 Configuração pontual espacial

Uma configuração pontual  $s$  é um conjunto de coordenadas  $s_i$  geradas por um mecanismo estocástico que podem estar associadas a informações adicionais (marcas) dentro de uma região plana de estudo  $A$ . A configuração pontual observada  $s$  é tratada como uma realização de um processo estocástico pontual  $S$  no espaço bidimensional.

A análise estatística de configurações pontuais requer um modelo teórico de referência. Esse modelo é denominado de Completa Aleatoriedade Espacial (CAE) (*Complete Spatial Randomness (CSR)*). Ele representa um caso em que é igualmente possível que um evento ocorra em qualquer local dentro da área de estudo. Quer dizer, eventos são distribuídos uniformemente sobre a área de estudo e são independentes um do outro. No modelo de CAE, considera-se que os eventos, em cada subárea, não são correlacionados, apresentam intensidade constante e estão associadas à mesma distribuição de probabilidade Poisson com intensidade  $\lambda$ . Intuitivamente, pode-se considerar que a posição dos eventos é independente e

os eventos têm igual probabilidade de ocorrer em qualquer local dentro da região  $B$  (Cressie, 1993; Diggle, 2003).

### 3.3 Propriedade de primeira ordem

Propriedade de primeira ordem de um processo pontual está relacionada à média do processo. Pode ser representada, por meio da intensidade do processo, isto é, o número de eventos por unidade de área. A intensidade de um processo pontual é associada ao valor esperado de uma variável aleatória. Ela pode ser constante (uniforme ou homogênea) ou pode variar de localização para localização (não uniforme ou não homogênea). A intensidade é sempre dada em termos do número de eventos por unidade de área (Ex.: dois eventos por metro quadrado). Segundo Diggle (2003), a investigação da intensidade deve ser uma das primeiras etapas na análise das configurações pontuais.

#### 3.3.1 Intensidade homogênea

Se um processo pontual  $S$  é homogêneo, então, para qualquer sub-região  $B$  do espaço bidimensional, o número esperado de pontos em  $B$  é proporcional à sua área:

$$E[N(X \cap B)] = \lambda \cdot \text{área}B \quad (1)$$

em que  $\lambda$  é a constante de proporcionalidade (intensidade).

Um estimador não tendencioso da verdadeira intensidade do processo pontual  $X$  é dado por:

$$\hat{\lambda} = \frac{n(s)}{\text{área}(B)}, \quad (2)$$

em que  $n(s)$  é o número de eventos dentro de uma região  $A$ .

#### 3.3.2 O teste de Kolmogorov–Smirnov

Uma importante classe de teste não paramétrico é constituída pelos chamados testes de aderência, em que a hipótese nula testada  $H_0$  refere-se à forma da distribuição da população.

Nesses testes, admite-se, por hipótese, que a distribuição da variável de interesse na população seja descrita por determinado modelo de distribuição de probabilidade e testa-se esse modelo, ou seja verifica-se a boa ou a má aderência dos dados da amostra ao modelo. Se o modelo tiver uma boa aderência e a amostra for razoavelmente grande, pode-se, em princípio, admitir que ele forneça uma idealização da distribuição populacional. Inversamente, a rejeição de  $H_0$ , em um dado nível de significância, indica que o modelo testado é inadequado para representar a distribuição da população.

No caso de configurações pontuais, o teste de Kolmogorov–Smirnov compara a máxima distância entre a ECDF (função de distribuição acumulada empírica) dos eventos em um eixo e a distribuição esperada sob a hipótese de completa aleatoriedade espacial, ou seja, o processo de Poisson homogêneo com intensidade constante  $\lambda$ . O valor  $p$  é a probabilidade de se obter uma estatística de teste (como a estatística de Kolmogorov–Smirnov) que seja pelo menos tão extrema quanto o valor que é calculado, a partir da amostra, quando os dados são realizações de um processo de Poisson com intensidade  $\lambda$ . Valores maiores, para a estatística de Kolmogorov–Smirnov, indicam que os dados não seguem a distribuição de Poisson e, portanto pode-se dizer que a configuração pontual é não homogênea. O valor  $p$  é uma probabilidade que mede a evidência contra a hipótese nula. Um valor  $p$  menor que uma nível de significância estabelecido a priori fornece uma evidência mais forte contra a hipótese nula (Justel; Peña; Zamar, 1994).

### 3.3.3 Intensidade local

Se um processo pontual  $\mathcal{S}$  é não homogêneo, o estimador kernel é um dos métodos mais utilizados, para estimar a intensidade local e, conseqüentemente, analisar as propriedades de primeira ordem.

Segundo Bailey e Gatrell (1995) e Diggle (2003), o estimador de kernel foi desenvolvido com o objetivo de obter uma estimativa "suave" de uma densidade de probabilidade univariada ou multivariada de observações amostrais.

Essa função faz uma contagem de todos os eventos dentro de uma região de influência, ponderando-os pela distância de cada um a uma localização de interesse. Seu estimador é dado por:

$$\hat{\lambda}(s) = \frac{1}{\delta_r} \sum_{i=1}^n \frac{1}{\tau} k \left( \frac{d(s_i; s)}{\tau} \right), \quad d(s_i; s) \leq \tau, \quad (3)$$

em que  $\tau \geq 0$  é o raio de influência que define a vizinhança do ponto a ser interpolado e controla a "suavização" da superfície gerada;  $k(\cdot)$  é uma função de densidade de probabilidade bivariada escolhida, a qual é simétrica em relação à origem. O raio de influência define a área centrada no ponto de estimação  $s$ , que indica quantos eventos  $s_i$  contribuem para estimar a função intensidade  $\lambda$ . No caso da função de interpolação  $k(\cdot)$ , é comum usar funções de terceira ou quarta ordem.

As funções Normal ou Quártica são as mais usualmente utilizadas em  $k(\cdot)$ . Entretanto saber qual a função de estimação kernel que será utilizada não é um ponto crítico e, sim, a escolha do raio de influência ( $\tau$ ) é decisiva, pois pode alterar as estimativas finais (Bailey; Gatrell, 1995). Existem vários métodos propostos para a escolha  $\tau$ , todos com prós e contras.

A função kernel quártica é dada por:

$$k(d) = \frac{3}{\pi} (1 - d^2)^2 \quad (4)$$

em que  $d$  representa a distância euclidiana entre a localização  $s$ , em que se deseja calcular a função e o evento observado  $s_i$ . Assim, um estimador da intensidade para uma localização arbitrária  $s$  utilizando a função kernel quártica é dado por:

$$\hat{\lambda}(s) = \sum_{i=1}^n \frac{1}{\pi\tau^2} \left(1 - \frac{d_i^2}{\tau^2}\right)^2 \quad (5)$$

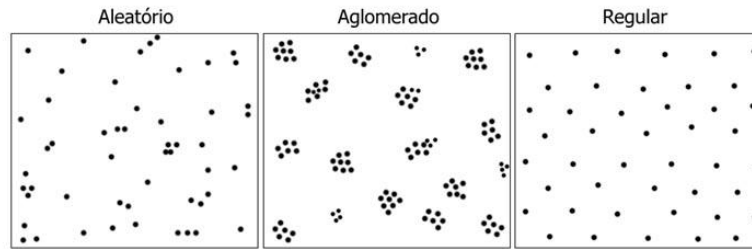
Uma maneira de apresentar as estimativas kernel é através do gráfico de calor que destaca as maiores intensidades das menores.

### 3.4 Propriedades de segunda ordem

A propriedade de segunda ordem de um processo pontual está relacionada à variância do processo, caracterizando a dependência espacial do processo pontual, ou seja, a estrutura de correlação espacial existente entre os eventos. Na análise dos efeitos de segunda ordem, estamos interessados na existência de interações entre os eventos, identificando como eles estão correlacionados (dependentes uns dos outros). Assim, podem ocorrer interações com agrupamentos (aglomerado), regularidade ou aleatoriedade (Cressie, 1993; Diggle, 2003).

A Figura 1 mostra um exemplo de cada um dos tipos de configurações pontuais no espaço.

Figura 1 – Tipo de configurações pontuais no espaço.



Fonte: Do autor (2018).

Existem várias formas para descrever as propriedades de segunda ordem. As técnicas mais utilizadas, para estimar as propriedades de segunda ordem, são baseadas nas funções empíricas  $F$ ,  $G$ ,  $J$ ,  $K$  e  $L$  que serão descritas a seguir.

### 3.4.1 Função $F$

Seja uma configuração de eventos gerados por um processo estacionário em uma região plana de área  $|A|$  e seja  $x_i = \|s - s_i\|$  é a distância entre um ponto aleatório ( $s$ ) e o  $i$ -ésimo evento mais próximo ( $s_i$ ). A função  $F$ , ou a função das distâncias dos espaços vazios, introduzida por Ripley (1977), é a função de distribuição acumulada da distância entre um ponto aleatoriamente escolhido ( $s$ ) e o evento mais próximo ( $s_i$ ). Essa função é definida por

$$F(x) = P(x_i \leq x). \quad (6)$$

O estimador mais simples, para a função  $F$ , sem correção do efeito de borda, pode ser obtido pela equação

$$\hat{F}(x) = \sum_{i=1}^m \frac{I(x_i \leq x)}{m}, \quad (7)$$

em que  $m$  é o número de pontos arbitrários e aleatórios do mapa,  $s_i$  representa a distância do  $i$ -ésimo ponto aleatório para o evento mais próximo no mapa analisado,  $I(\cdot)$  é uma função indicadora igual a 1 quando  $x_i$  é menor ou igual a uma distância  $x$  e 0 caso contrário.

Contudo o estimador apresentado, na Equação (7), é viesado (tendencioso), uma vez que, para uma mesma distância, pontos aleatórios, localizados nas bordas do mapa, terão uma probabilidade menor de ter um evento próximo que os pontos localizados no centro do mapa.

Na prática, esse efeito ocorre, porque se trabalha com regiões de área finita enquanto teoricamente a área é considerada infinita. Portanto a distância entre um ponto aleatório e um evento que se encontra fora do mapa é ignorada, mesmo que ela esteja menor que qualquer distância entre um ponto aleatório e um evento localizado dentro do mapa.

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

$$\hat{F}(x) = \frac{\sum_{i=1}^n I_x(x_i, r_i)}{\sum_{i=1}^n I_x(x_i)} \quad x > 0, \quad (8)$$

em que  $r_i$  é a distância de um evento até o ponto mais próximo na borda do mapa e  $I_x(x_i, r_i)$  é uma função indicadora igual a 1, quando  $s_i$  é menor ou igual  $x$  e  $r_i$  é maior ou igual a  $r$  e 0, caso contrário.

Diggle (2003) mostra que a função  $F$  teórica, para um processo de Poisson homogêneo, com intensidade  $\lambda$ , sob a hipótese de completa aleatoriedade espacial e área infinita, é dada pela equação

$$F(x) = 1 - e^{-\lambda\pi s^2}, \quad x > 0. \quad (9)$$

Uma maneira simples de se estudar a função  $F$  é por meio de um gráfico. Se  $\hat{F}(x) > F(x)$  significa que a função apresenta regularidade. Se  $\hat{F}(x) < F(x)$  significa que a função apresenta agrupamento. Diggle (2003) mostra que, empiricamente, a Função  $F$  apresenta grande poder para detectar agrupamentos de eventos.

### 3.4.2 Função $G$

A função  $G$ , também definida por Ripley (1977), é a função de distribuição acumulada da distância entre um evento e o seu vizinho mais próximo. Seja  $y_i = \|s_i - s_j\|$  a distância entre o  $i$ -ésimo evento e o seu  $j$ -ésimo vizinho mais próximo dentro de uma região com área  $|A|$ . Assumindo que o processo pontual é estacionário, a função  $G$  é definida como sendo a probabilidade de encontrar um evento que tenha um vizinho com distância menor ou igual a  $y$ . Assim temos que

$$G(y) = P(y_i \leq y). \quad (10)$$

$G(y)$  pode ser estimada pela função de distribuição empírica das distâncias observadas entre vizinhos mais próximos. O estimador simples, sem correção do efeito de borda para a função  $G$ , é definido por:

$$\hat{G}(y) = \sum_{i=1}^n \frac{I(y_i \leq y)}{n} \quad (11)$$

em que  $n$  é o número de eventos arbitrários e aleatórios do mapa,  $I(\cdot)$  é uma função indicadora se  $y_i \leq y$ , temos  $I(\cdot) = 1$  ou se  $y_i > y$ , temos  $I(\cdot) = 0$ .

A função  $G$  teórica, para um processo de Poisson homogêneo, com intensidade  $\lambda$  sob a hipótese de completa aleatoriedade espacial, em uma região de área  $|A|$ , é dada pela equação

$$G(y) = 1 - e^{-\lambda\pi y^2}. \quad (12)$$

Na prática,  $\lambda$  é não conhecido sob a suposição de estacionariedade,  $\lambda$  pode ser estimado por:  $\hat{\lambda} = \frac{n}{|A|}$ . Um estimador não viesado, proposto por Ripley (1977), corrige o efeito de bordas, é dado pela equação

$$\hat{G}(y) = \frac{\sum_{i=1}^n I_y(y_i, r_i)}{\sum_{i=1}^n I_y(y_i)}, \quad y > 0. \quad (13)$$

Uma maneira simples de se estudar a função  $G$  é por meio de gráfico. Se  $\hat{G}(y) > G(y)$  significa que a função apresenta agrupamento. Se  $\hat{G}(y) < G(y)$  significa que a função apresenta regularidade.

Diggle (2003) afirma que a Função  $G$  é mais poderosa para detectar eventos regularmente distribuídos.

### 3.4.3 Função $J$

As funções  $F$  e  $G$  descrevem a interação existente na configuração de maneiras “opostas”. Entretanto, sob a suposição de completa aleatoriedade espacial, ambas são iguais. Segundo Baddeley *et al.* (2000), um estimador da função  $J$  é obtido, a partir dos estimadores das funções  $F$  e  $G$ , para toda distância  $d > 0$ , tal que  $F(d) < 1$  dado por



$$F(d) = G(d) = 1 - e^{-\lambda\pi d^2}. \quad (14)$$

Observando esse comportamento, Van Leishout e Baddeley (1996) propuseram a função  $J$ , definida por:

$$J(d) = \frac{1-G(d)}{1-F(d)}, \quad \text{para todo } F(d) \neq 1 \quad (15)$$

Sob a hipótese de completa aleatoriedade espacial,  $J(d) \equiv 1$ . Se  $J(d) > 1$  indica regularidade. No caso de  $J(d) < 1$ , indica que há agrupamentos. Um estimador da função  $J$  é obtido diretamente dos estimadores das funções  $G$  e  $F$ , ou seja,

$$\hat{J}(d) = \frac{1-\hat{G}(d)}{1-\hat{F}(d)} \quad (16)$$

A função  $J$ , em geral, apresenta o mesmo poder para detectar configurações com agrupamentos e regularidades. Entretanto pode ter um poder menor que a função  $G$ , para regularidade e a função  $F$  para agrupamentos. Podem-se utilizar os mesmos estimadores para correção do efeito de borda utilizados nas funções  $F$  e  $G$ . Entretanto, Lieshout e Baddeley (1996) argumentam que não é necessário fazer correção para o efeito de borda na função  $J$ .

Segundo Van Lieshout e Baddeley (1996), valores de  $J(d)$  menores que 1, para uma particular distância  $d$ , indicam agrupamentos e valores maiores que 1 indicam regularidades.

#### 3.4.4 Função $K$

Para caracterizar a propriedade de segunda ordem de uma configuração pontual, em distintas escalas e que seja igualmente poderosa, para detectar tanto os padrões com agrupamento como regularidade foram as motivações para que Ripley (1977) desenvolvesse a Função  $K$  que por pode ser definida como

$$k(d) = \frac{2\pi}{\lambda^2} \int_0^d \lambda^2(x) x dx. \quad (17)$$

Para Ripley (1977), a função  $K$ , ao combinar as funções de intensidade de primeira e segunda ordem, resulta numa função descritiva do padrão espacial com interpretações mais diretas.

Diversos estimadores têm sido propostos para a função  $K$ . Um estimador, com correção para o efeito de bordas para uma região de área  $|A|$  foi proposto por Ripley (1977) como sendo

$$\hat{K}(d) = \frac{|A|}{n^2} \sum_{i=1}^n \sum_{j=1}^n \frac{I(d_{ij})}{e_{ij}}, \quad (18)$$

em que  $n$  é o número de eventos,  $I(\cdot)$  é uma função indicadora se  $d_i \leq d$ , temos  $I(\cdot) = 1$  ou se  $d_i > d$ , temos  $I(\cdot) = 0$  e  $e_{ij}$  é o corretor de borda que representa a proporção de uma circunferência ao redor de um evento  $s_i$  passando sobre o evento  $s_j$  que está dentro de  $A$ .

Ripley (1977) mostra que para um processo de Poisson homogêneo com intensidade constante (completa aleatoriedade espacial), a Função  $K$  é dada por

$$k(d) = \frac{2\pi}{\lambda^2} \int_0^d \lambda^2(x) x dx = k(d) = \pi d^2 \quad (19)$$

Essa função pode ser empregada para o teste da hipótese de CAE. Portanto, para um processo aleatório, teríamos  $k(d) = \pi d^2$ ; para um processo agregado, os pontos têm, em média, mais vizinhos que o esperado sob a hipótese de CAE, assim,  $k(d) > \pi d^2$ . Para um processo regular, os pontos têm, em média, menos pontos que o esperado sob a hipótese de CAE, ou seja,  $k(d) < \pi d^2$ .

### 3.4.5 Função $L$

Existem várias transformações que podem ser feitas na função  $K$  para facilitar a sua interpretação. Em geral, essas transformações são utilizadas, para linearizar e estabilizar a variância da função  $K$  e recebem o nome de função  $L$ . Uma dessas formas que foi proposta por Besag e Diggle (1977) é dada por:

$$\hat{L}(t) = \sqrt{\frac{\hat{K}(d)}{\pi}} - d \quad (20)$$

A função  $K$  (ou  $L$ ) é aproximadamente não tendenciosa para  $d$  fixo. O viés da função  $K$  aumenta com o aumento de  $d$ . Existem várias propostas, para estabelecer o valor máximo de  $d$  (Diggle, 2003).

A função  $L(d)$  é mais simples de interpretar que  $K(d)$ , porque a linha obtida para a hipótese nula é uma reta com abscissa igual a zero, ou seja,  $L(d) = 0$  em todas as distâncias  $d$ . Observe, na Tabela 1, um comparativo entre os valores de  $K(d)$  e  $L(d)$ .

Tabela 1 – Comparação entre os valores de  $L(h)$  e  $K(h)$ .

Tipo de Processo	$L(d)$	$K(d)$
Aleatório	..... = 0	..... = $\pi d^2$
Regular	..... < 0	..... < $\pi d^2$
Aglomerado	..... > 0	..... > $\pi d^2$

Fonte: Do autor (2018).

### 3.5 Envelopes de simulação

Segundo Diggle (2003) e Câmara *et al.* (2004), é aceitável utilizar a função  $\hat{L}(d)$ , como método formal, para se confrontar estatisticamente a distribuição dos eventos analisados com o que se confiaria na hipótese de completa aleatoriedade espacial (CAE). Para isso, são construídos envelopes de simulação, para a distribuição de CAE, a fim de acessar o valor dos desvios. Os envelopes de simulação inferior e superior são construídos a partir de  $k$  simulações independentes de  $n$  eventos na região  $A$ .

$$L(h) = \min_{i=1, \dots, k} \{ \hat{L}_i(d) \} \tag{21}$$

$$U(h) = \max_{i=1, \dots, k} \{ \hat{L}_i(d) \}$$

$L(d)$  e  $U(d)$  são os mínimos e os máximos da função  $L$ , respectivamente avaliadas a partir de  $k$  simulações independentes de processo aleatório com  $n$  eventos. No diagnóstico do gráfico com a distribuição e os envelopes (FIGURA 5), picos positivos na função estimada  $\hat{L}(d)$  que estão acima do envelope superior, confirmam fatos de agrupamento na escala analisada, portanto, se todos os valores da função  $\hat{L}(d)$  estiverem acima do envelope superior

e com valores positivos, teremos agrupamentos em todas as escalas. Valores negativos na função estimada  $\hat{L}(d)$ , que estiverem abaixo do envelope inferior, confirmam regularidade nessa escala, portanto, se todos os valores de  $\hat{L}(d)$  estiverem abaixo do envelope inferior e com valores negativos, tem-se regularidade em todas as escalas.

### 3.6 Propriedade de terceira ordem

A propriedade de segunda ordem, conforme visto nos itens anteriores é baseada nas distâncias entre pares de eventos. Uma alternativa seria considerar distâncias entre três eventos localizados dentro de uma particular distância em uma região. Essa é a idéia básica por trás das funções utilizadas para caracterizar a propriedade de terceira ordem que tem como objetivo discriminar melhor os tipos de interações entre os eventos (agrupamento, regularidade ou completa aleatoriedade espacial) quando comparado com as funções de segunda ordem.

As propriedades de terceira ordem, em um processo espacial de pontos, envolvem o relacionamento entre os eventos localizados em uma região. Uma abordagem prática é procurar relações de distâncias entre eventos e propriedades de terceira ordem. Os métodos de distância ao vizinho mais próximo e a funções  $T$  são exemplos de procedimentos que consideram a distância entre eventos.

#### 3.6.1 Função $T$

Para caracterizar essa propriedade, pode-se utilizar a função  $T$  introduzida por Schladitz e Baddeley (1997). Neste tópico apresentam-se os aspectos teóricos dessa função.

Seja  $\phi$  um processo pontual estacionário com intensidade  $\lambda$ . A função  $T$  associada a  $\phi$  é definida como

$$T(r) = \frac{1}{2\lambda^2} E^0! \sum_{x,y \in b(0,r)} clr(s_i, s_j) \text{ se } r \geq 0 \quad (22)$$

em que  $s_i = (x_1; x_2)$  e  $s_j = (y_1; y_2)$  em que os  $s_i$  e  $s_j$  são as coordenadas dos eventos. Foi denotada  $E^0!$  a esperança em relação à distribuição de Palm reduzida  $\phi$  na origem 0. Observe que a intensidade é multiplicada por 2 para evitar a contagem dupla de conjuntos de pontos.

Desejamos contar o número médio de pares triplos dentro do raio  $r$ . Fixa um valor  $r > 0$ .

Denote  $\|s_i - s_j\| = \sqrt{(x_1 - y_1)^2 + (x_n - y_n)^2}$  a distância euclidiana em  $\mathbb{R}^d$  e  $b(0, r)$ , a esfera de raio  $r$  centrada na origem em  $\mathbb{R}^d$ . Para dois pontos  $x, y, z \in \mathbb{R}^d$  define:

$$clr(x, z) = \begin{cases} 1 & \text{se } 0 < \|s_i - s_j\| < r, \\ 0 & \text{caso contrário} \end{cases} \quad (23)$$

Para o indicador  $x$  e  $y$  são pontos distintos da região de estudo. Ao aplicar a fórmula de Campbell–Mecke (Mecke, 1967; Stoyan; Kendall; Mecke, 1995) (23) pode ser reescrita como:

$$\mathbf{T}(r) = \frac{1}{2\lambda^3|B|} \mathbb{E} \sum_{x \in \phi \cap B} \sum_{y, z \in \phi} clr(s_i, s_j) clr(s_i, s_k) clr(s_j, s_k) \quad (24)$$

em que:  $|B| = (\|s_i - s_j\|)(\|s_i - s_k\|)(\|s_j - s_k\|) \quad |B| > 0$

$$clr(s_i; s_j) = \begin{cases} 1 & \text{se } 0 < \|s_i - s_j\| < r \\ 0 & \text{caso contrário} \end{cases}$$

$$clr(s_j; s_k) = \begin{cases} 1 & \text{se } 0 < \|s_j - s_k\| < r \\ 0 & \text{caso contrário} \end{cases}$$

$$clr(s_i; s_k) = \begin{cases} 1 & \text{se } 0 < \|s_i - s_k\| < r \\ 0 & \text{caso contrário} \end{cases}$$

em que:  $s_i = (x_1; x_2)$ ;  $s_j = (y_1; y_2)$  e  $s_k = (z_1; z_2)$

$$\|s_i - s_j\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

$$\|s_j - s_k\| = \sqrt{(y_1 - k_1)^2 + (y_2 - k_2)^2}$$

$$\|s_i - s_k\| = \sqrt{(x_1 - k_1)^2 + (x_2 - k_2)^2}$$

Observe que o processo em qualquer região delimitada  $W \subset \mathbb{R}^d$  é  $|B| > 0$ .

### 3.6.2 Valor teórico da função $T$

Schladitz e Baddeley (1997) desenvolvem a função  $T$  para diversos tipos de processos pontuais no espaço tais como o processo de Poisson homogêneo, processo de Poisson com agrupamentos, processo de Neyman-Scott e processo de Gauss-Poisson

A equação (24) serve de base para a construção de estimadores, também serve para encontrar um valor ao  $\phi$  processo de Poisson homogêneo em  $\mathbb{R}^2$ . Assim, Schladitz e Baddeley (1997) demonstram que a função  $T$  para o processo de Poisson homogêneo é dado por:

$$T(r) = \frac{1}{2\lambda^3|B|} E \sum_{x \in \phi \cap B} \sum_{y, z \in \phi} clr(s_i, s_j) clr(s_i, s_k) clr(s_j, s_k) \quad (25)$$

$$T(r) = \frac{1}{2}\pi(\pi - 3/4\sqrt{3}) r^4 \quad \text{se } r \geq 0 \quad (26)$$

### 3.6.3 Estimador da função $T$

A função  $T$  deve ser estimada pela observação (realização) do processo pontual em uma região  $W \subseteq \mathbb{R}^2$ . A seguir, constroem-se estimadores imparciais para  $\lambda^3 T(r)$  e uma relação-imparcial do estimador para  $\lambda^2 T(r)$ . A base para a construção desses estimadores é a equação (25).

Como no caso das funções  $F$ ,  $G$  e  $K$ , o estimador de  $T$  é influenciado pelos efeitos de bordas que devem ser corrigidos. Os estimadores com correção de borda são, basicamente, os mesmos utilizados na função  $K$ , ou seja, amostra reduzida (fronteira), translação e isotrópico conforme pode ser visto em Schladitz e Baddeley (1997).

### 3.6.4 Método da fronteira

A solução mais fácil é o método de fronteira (ou da amostra reduzida), no qual reduzimos os pontos de  $x$ , de modo que todo o círculo de raio  $r$  centrada em  $x$  pode ser observado.

Seja

$$W_{-r} = W_{\theta} b(0, r) = \{x \in W : \inf \|s_i - s_j\| \geq r\} \quad (27)$$

uma janela representada por uma esfera de raio  $r$ .

No método da fronteira são consideradas no estimador apenas as triplas em que pelo menos um dos seus pontos está mais longe que  $r$  da borda da janela  $W$ . Assim, Schladitz e Baddeley (1997) mostram que o estimador para  $\lambda^2 T(r)$  pelo método de fronteira pode ser obtido pela equação (28):

$$\widehat{2\lambda^2 T}(r) = \frac{1}{2\phi(W_{-r})} E \sum_{s_i \in \phi \cap B} \sum_{s_j, s_k \in \phi} cl_r s_i cl_r s_j cl_r s_k \quad (28)$$

em que:  $s_i = (x_1; x_2)$ ;  $s_j = (y_1; y_2)$  e  $s_k = (z_1; z_2)$  e  $cl$  foram definidos anteriormente.

Esta é um estimador não tendencioso para  $\lambda^2 T(r)$  para cada  $r$  (isto é, a razão entre as expectativas de o numerador e o denominador é igual ao valor real  $\lambda^2 T(r)$ ). A sua divisão por  $(\phi(W)(\phi(W) - 1)) / |W|^2$  (um estimador para  $\lambda^2$ ) leva ao estimador pelo método fronteira  $\widehat{T}_{\phi, b}(r)$  para  $T(r)$  sempre que  $\phi(W_{-r}) > 0$ . A desvantagem desse estimador é que, aumentando o raio  $r$ , usa-se cada vez menos pontos e, portanto, tornando o estimador pouco confiável.

Schladitz e Baddeley (1997) desenvolvem estimadores para  $T(r)$  utilizando os demais corretores de borda.

## 4 MATERIAIS E MÉTODOS

O estudo de campo foi feita com uma coleta amostral de mais de 50 espécies de árvores, das quais duas delas foram selecionadas para a realização de análises estatísticas do trabalho.

### 4.1 Dados

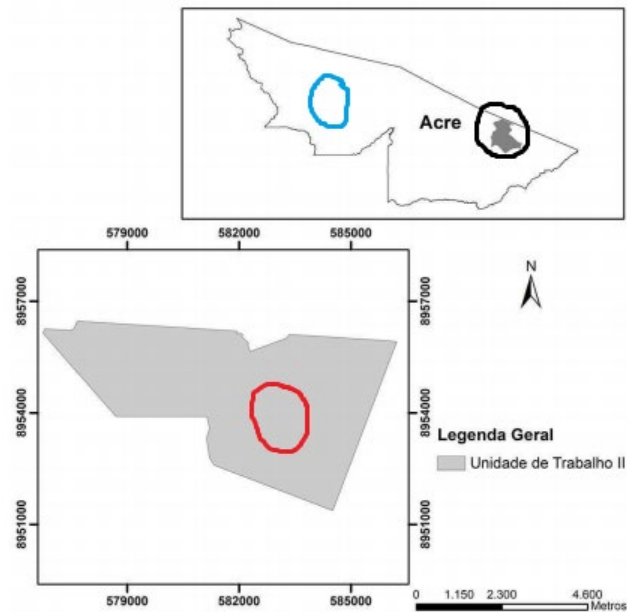
Os dados utilizados neste trabalho são provenientes de um inventário florestal realizado na Fazenda Canary, localizada no município de Bujari, Acre.

A vegetação existente é constituída de floresta aberta com bambu e palmeira, bem como espécies comerciais e potenciais para a produção de madeira (Acre, 2006). O clima é do tipo Ami (sistema Köppen), o qual apresenta, anualmente, elevado índice pluviométrico, com breve período seco, em que se pode obter um índice pluviométrico inferior a 60 mm em, pelo menos, um mês do ano. A temperatura média anual é em torno de 24,5° C, sendo os valores máximos e mínimos médios 32°C e 18°C, respectivamente. A umidade relativa do ar apresenta médias anuais entre 80% e 90% (Mesquita, 1996).

A propriedade consta com uma área total de aproximadamente 15.810 ha, sendo a reserva legal de aproximadamente 8.446 ha e uma área efetiva de manejo florestal de aproximadamente 4.465 ha. Entretanto, no presente estudo, foi utilizada uma área parcial de aproximadamente 2.601 há, conforme Figura 2.



Figura 2 – Fazenda Canary situada no município de Bujari, Acre.



Fonte: Do autor (2018).

O inventário florestal foi realizado, por meio do georreferenciamento das árvores em coordenadas UTM, por meio de um aparelho GPS, pelo qual foram mensuradas as 56 espécies de árvores comerciais, a partir de um diâmetro mínimo de corte ( $DMC \geq 30$  cm), sendo marcadas com placas de alumínio contendo suas devidas numerações. Nas árvores, foram mensurados o diâmetro a 1,30 m do solo (DAP), altura comercial da árvore, identificação botânica e a determinação da qualidade do fuste. As árvores selecionadas, para a exploração, foram estabelecidas, conforme o plano de manejo florestal e a Instrução Normativa, resolução conjunta CEMACT/CFE número 003 de 12 de agosto de 2008. Posteriormente, as informações geradas foram convertidas em um banco de dados espaciais.

Neste estudo, para as análises espaciais, foram utilizadas as espécies: Cedro Taperebá (*Antrocaryon amazonicum*) e Embiruçu (*Bombax humile*).

#### 4.2 Análise estatística

As análises estatísticas das configurações pontuais foram realizadas com base nos dados de coordenadas das árvores das duas espécies: *Antrocaryon amazonicum* e *Bombax humile*.

Para a análise da intensidade kernel, utilizamos uma base de 66 pontos para a espécie *Antrocaryon amazonicum* e 293 pontos para a espécie *Bombax humile*. Para obter as

estimativas da intensidade, Diggle (2003) considera que poderíamos escolher, a partir dos vértices, uma grade regular fina posicionada sobre a janela  $W$ . Quanto mais fina for a grade, melhor a resolução da intensidade e, conseqüentemente, maior o trabalho computacional. Foi utilizada uma kernel quártico, indicado por Bailey e Gatrell (1995), por apresentar a propriedade de que quanto mais próximo do ponto onde será feita estimativa estiver o evento, maior será o seu peso na função intensidade. As intensidades estimadas são apresentadas em um gráfico de calor que destaca as maiores e menores intensidades.

Para o teste de Kolmogorov–Smirnov, usamos o eixo do  $x$  dentro de uma distância de 8952000m a 8956000m para fazermos uma comparação se teste segue um processo de Poisson Homogêneo

Para a função  $F$ , utilizamos uma distância de 0 – 500m, para a espécie *Antrocaryon amazonicum*; enquanto para o caso da espécie *Bombax humil*, utilizamos de 0 – 250m.

Para a função  $G$ , utilizamos uma distância de 0 – 400m para a espécie *Antrocaryon amazonicum*; enquanto para a espécie *Bombax humile* utilizamos de 0 – 250m.

Para as funções  $K$  e  $L$  utilizamos uma distância de 0 – 1200m para a espécie *Antrocaryon amazonicum*; enquanto para a espécie *Bombax humile* utilizamos de 0 – 1200m.

Os estimador de todas as funções podem ser influenciados por eventos próximos das bordas, assim, utilizou-se o corretor de borda de fonteiras (border)

Para a função  $J$  podem-se foram utilizados os mesmos estimadores das funções  $F$  e  $G$ . Entretanto os autores argumentam que, na função  $J$ , não é necessário fazer correção para o efeito de borda.

Para a função  $T$  utilizamos uma distância de 0 – 1200m, para a espécie *Antrocaryon amazonicum*; enquanto para o caso da espécie *Bombax humile*, utilizamos de 0 – 1200m. Mais uma vez utilizou o método da fronteira para a correção do efeito de bordas.

Para construir os envelopes, foram feitas 999 simulações Monte Carlo sob a suposição de completa aleatoriedade espacial (processo de Poisson homogêneo) para as funções  $G$ ,  $J$ ,  $L$ ,  $F$ ,  $K$  e  $T$  conforme descrito em Scalon *et al.* (2003).

As análises usando o estimador de Kernel e as funções  $F$ ,  $G$ ,  $J$ ,  $K$ ,  $L$  e  $T$  com seus respectivos envelopes de simulação (intervalo de confiança) foram realizadas utilizando funções disponíveis e/ou desenvolvidas *software* R (R version 3.4.0, 2017) e na biblioteca *spatstat* (Baddeley e Turner, 2005).

## 5 RESULTADOS E DISCUSSÃO

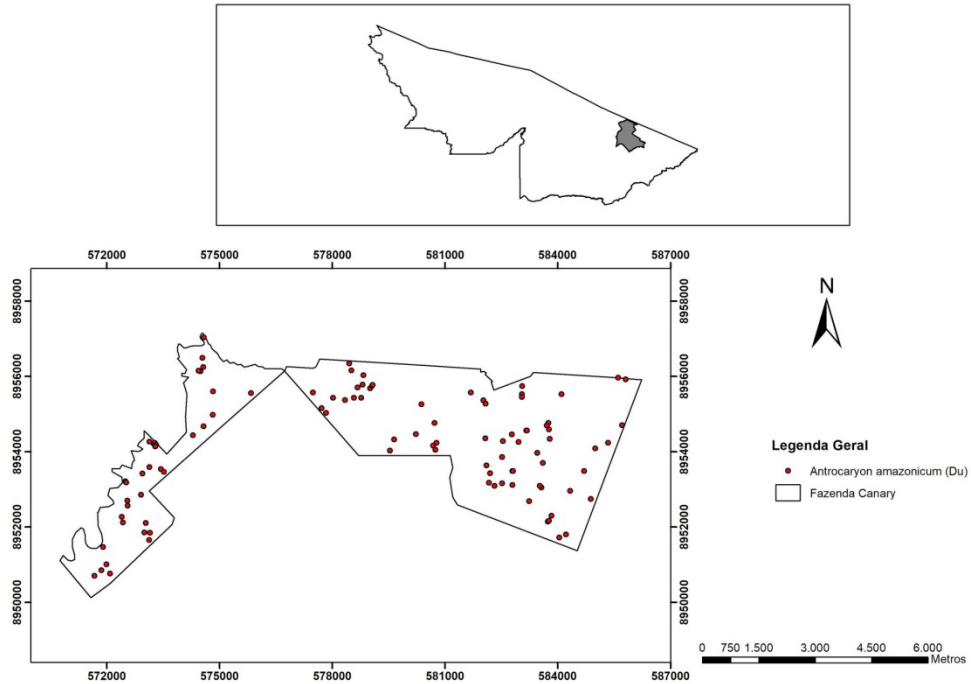
A grande maioria das análises de configurações pontuais no espaço é focada na caracterização da propriedade de primeira ordem (em geral, usando um alisador kernel) e da propriedade de segunda ordem (utilizando alguma função descritora como a função  $K$ ). Não são realizadas análises para caracterizar a propriedade de terceira ordem. Não existe um método padrão, para caracterizar configurações de ordem superior a dois, embora estatísticas de terceira ordem tenham sido adaptadas para alguns problemas específicos conforme pode ser visto nos trabalhos de Hanisch (1983), Hanisch e Stoyan (1984), Muller, Syversveen e Waagepetersen (1998) e Ayala e Simó (2000).

Schladitz e Baddeley (1997) propuseram uma estatística, mais geral, para caracterizar a propriedade de terceira ordem de uma configuração pontual tendo como base a contagens de triplas de pontos. Essa estatística, denominada de função  $T$  é baseada no número médio de triplas que ocorrem dentro de uma particular distância  $e$ , de certa forma, guarda alguma semelhança com a função  $K$  (Ripley, 1977) que é baseada no número médio de pares de eventos que ocorrem dentro de uma particular distância.

Esta dissertação apresenta e utiliza a função  $T$  para caracterizar a propriedade de terceira ordem das configurações pontuais de duas espécies (*Antrocaryon amazonicum* e *Bombax humile*) de árvores nativas da Floresta Amazônica.

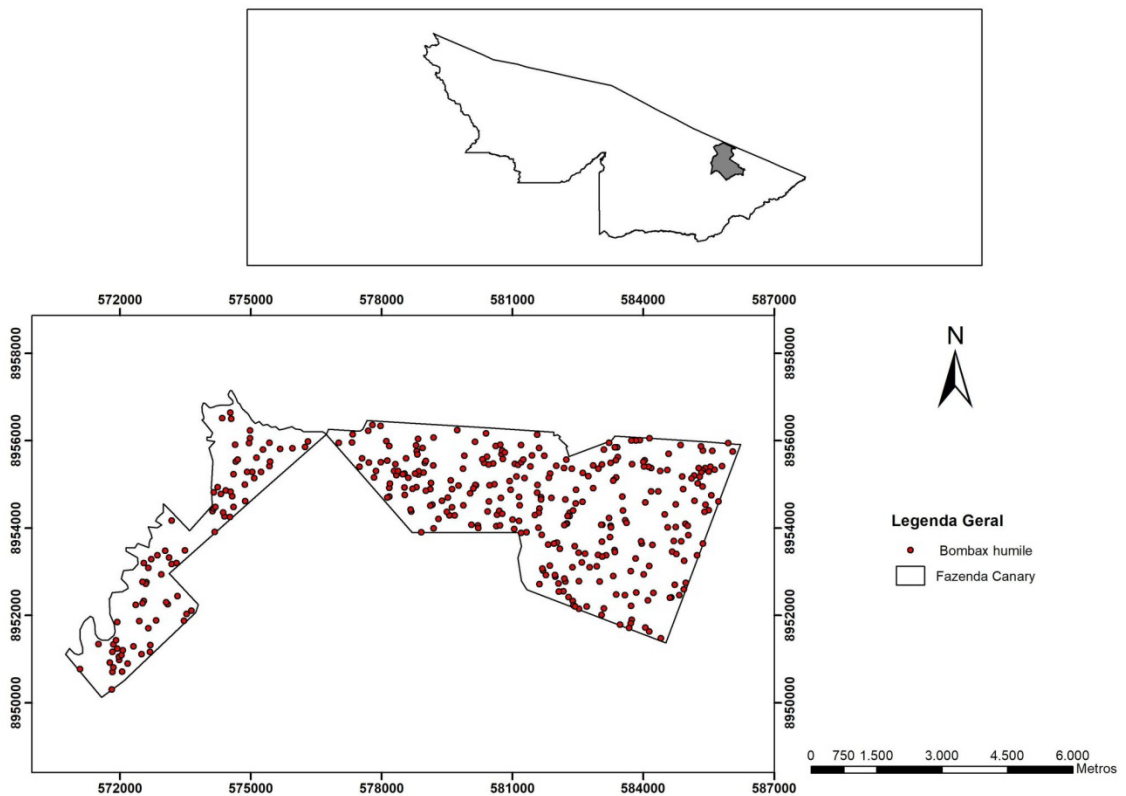
Primeiramente, foram obtidos os gráficos da distribuição das árvores na região de estudo. A Figura 2 mostra a localização da região de estudo, as Figuras 3 e 4 mostram de forma detalhada como as árvores (pontos) das duas espécies estão distribuídos na região de estudo, enquanto as Figuras de 5 e 6 mostram as configurações pontuais, juntamente com os mapas de calor obtidos pelo método kernel.

Figura 3 – Distribuição das árvores na fazenda Canary situado no município de Bujari no estado do Acre da espécie *Antrocaryon amazonicum*.



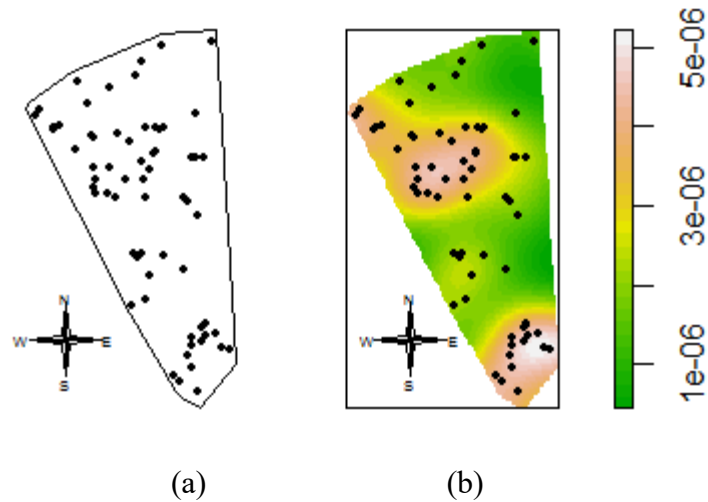
Fonte: Do autor (2018).

Figura 4 – Distribuição das árvores na fazenda Canary situado no município de Bujari no estado do Acre da espécie *Bombaxhumile*.



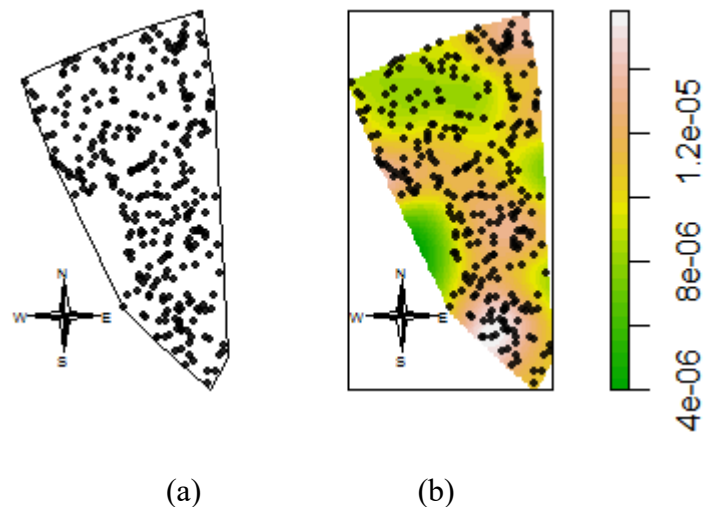
Fonte: Do autor (2018).

Figura 5 – Distribuição das árvores na região de estudo (a) e gráfico de calor (b) para a espécie *Antrocaryon amazonicum*.



Fonte: Do autor (2018).

Figura 6 – Distribuição dos pontos na região de estudo (a) e gráfico de calor (b) para a espécie *Bombaxhumile*.

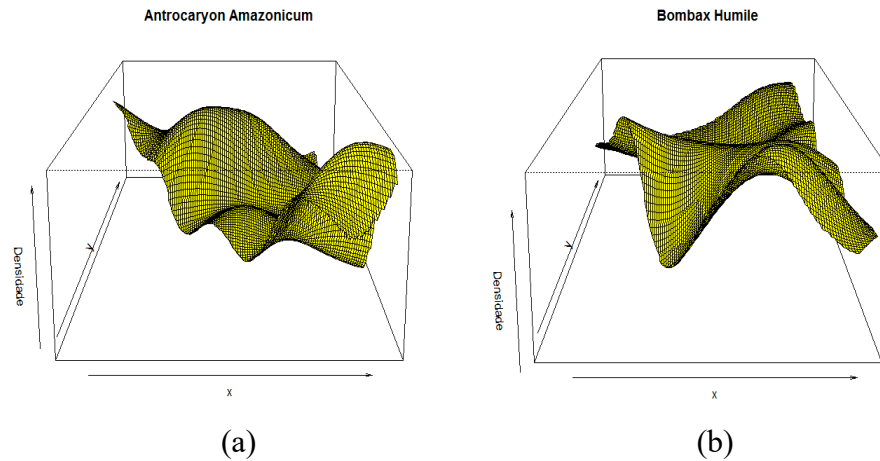


Fonte: Do autor (2018).

Dos dados da espécie *Antrocaryon amazonicum* foi observado uma intensidade global igual a 0,006045657 pontos por metro quadrado, ou seja, equivale dizer que existe aproximadamente seis árvores por hectare. Já considerando os dados da espécie *Bombaxhumile*, foi estimada uma intensidade global igual a 0,007107665 pontos por metro quadrado, ou seja, aproximadamente, sete árvores por hectare.

A Figura 7 apresenta os gráficos em 3D das intensidades dos pontos na região do estudo. Vale salientar que os pontos mais altos do gráfico são onde há uma maior concentração (intensidade) de árvores na região do estudo.

Figura 7 – Gráfico em 3D da densidade dos pontos na região de estudo.

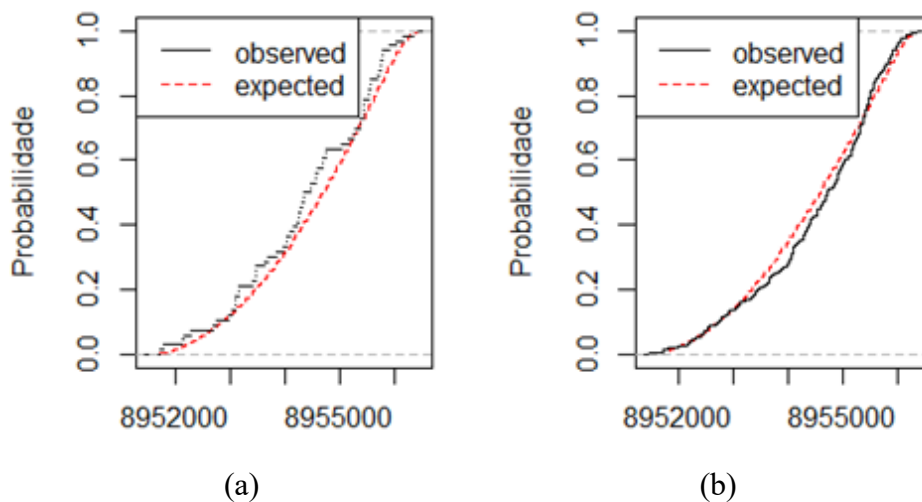


Fonte: Do autor (2018).

As Figuras 5, 6 e 7 parecem não indicar regiões aparentes com altas e baixas intensidades, o que poderia indicar que as configurações pontuais apresentam intensidades homogêneas.

Para verificar se a suposição de que a distribuição dos pontos segue um processo de Poisson homogêneo foi feito o gráfico dos quantis, apresentado na Figura 8 e o teste de Kolmogorov–Smirnov.

Figura 8 – Gráfico dos quantis e o teste de Kolmogorov–Smirnov.



Fonte: Do autor (2018).

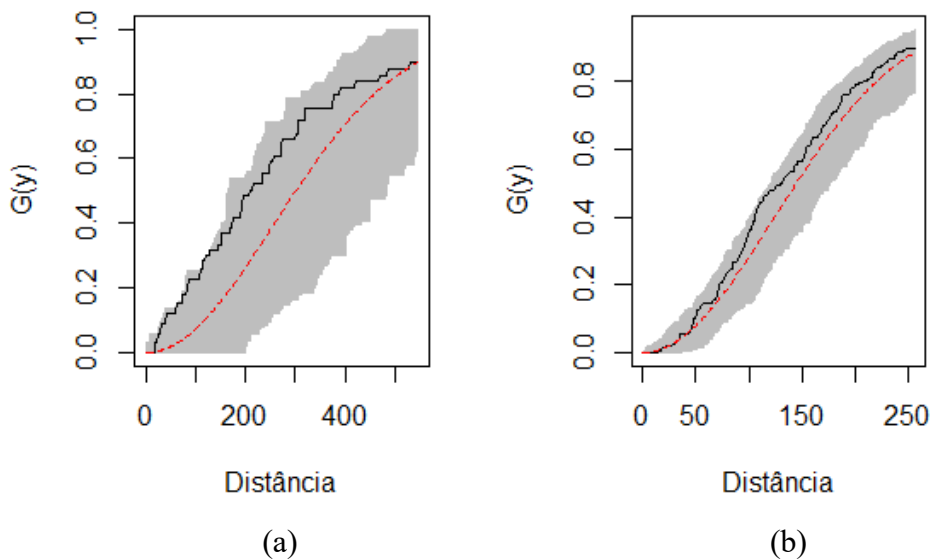
De acordo com a Figura 8, podemos perceber que os gráficos das distribuições observadas e esperadas estão bem próximos, o que nos dá indícios de que os pontos se distribuem segundo um processo de Poisson homogêneo. Além disso, ao realizarmos o teste

de Kolmorov–Smirnov, obtivemos um valor  $-p = 0,3784$ , para a espécie *Antrocaryon amazonicum* e um valor  $-p = 0,0782$ , para a espécie *Bombaxvhumile*. Assim, considerando um nível de 5% de significância, podemos concluir que os pontos se distribuem segundo um processo de Poisson homogêneo, tanto para a espécie *Antrocaryon amazonicum* quanto para a *Bombax humile*.

Ripley (1981) garante que esta análise de primeira ordem realizada anteriormente é um passo preliminar, importante e indispensável para verificar que tipo de padrão espacial segue a configuração pontual dos eventos.

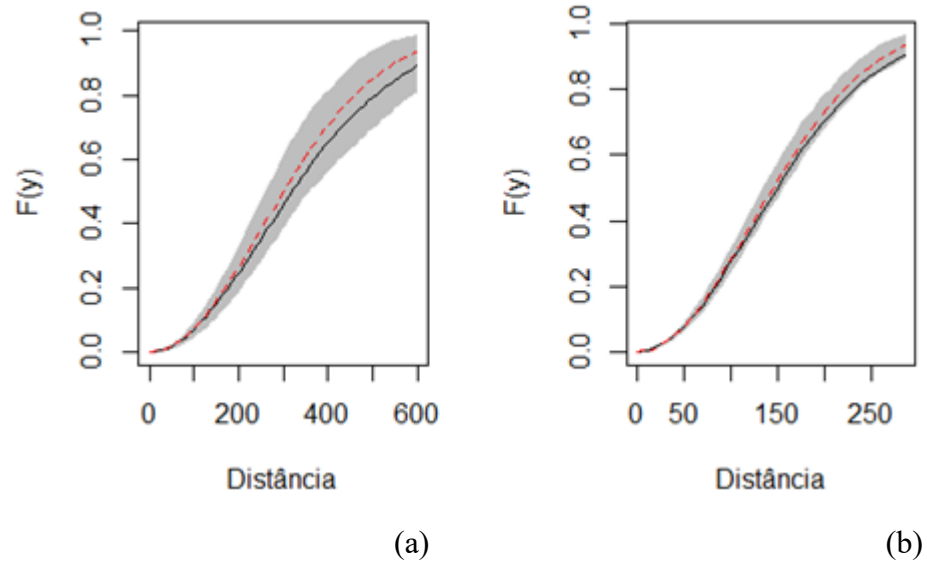
Com o objetivo de verificar se os pontos apresentam dependência espacial foram utilizadas as funções  $G$ ,  $F$ ,  $J$ ,  $K$  e  $L$ , juntamente com os respectivos envelopes de confiança, conforme pode ser observado nas Figuras de 9 a 13. Em todos os gráficos, a figura da esquerda é referente a espécie *Antrocaryon amazonicum*, enquanto a figura da direita refere-se a espécie *Bombax humile*

Figura 9 – Envelope de aleatoriedade para a função  $G$ .



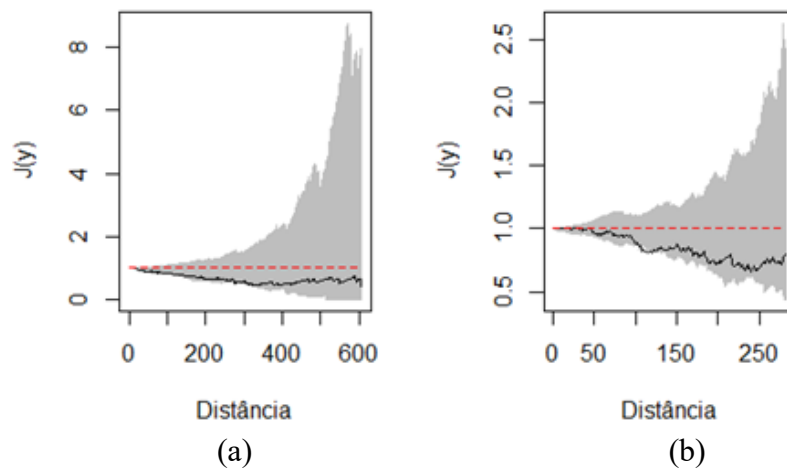
Fonte: Do autor (2018).

Figura 10 – Envelope de aleatoriedade para a função  $F$ .



Fonte: Do autor (2018).

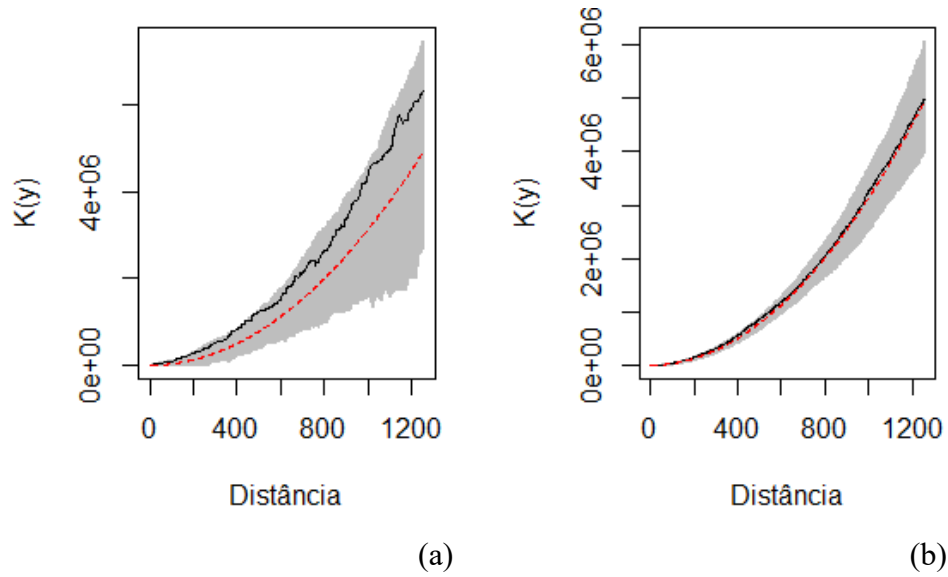
Figura 11 – Envelope de aleatoriedade para a função  $J$ .



Fonte: Do autor (2018).

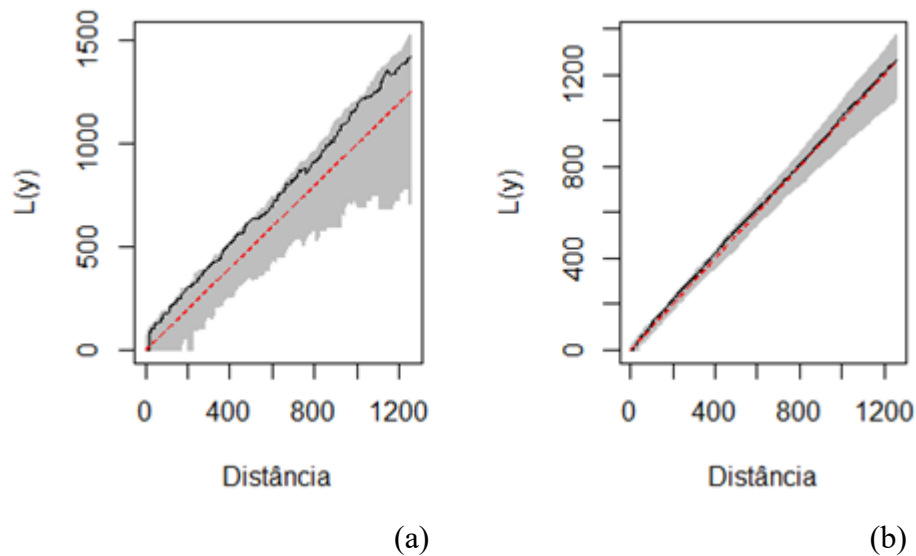


Figura 12 – Envelope de aleatoriedade para a função  $K$ .



Fonte: Do autor (2018).

Figura 13 – Envelope de aleatoriedade para a função  $L$ .



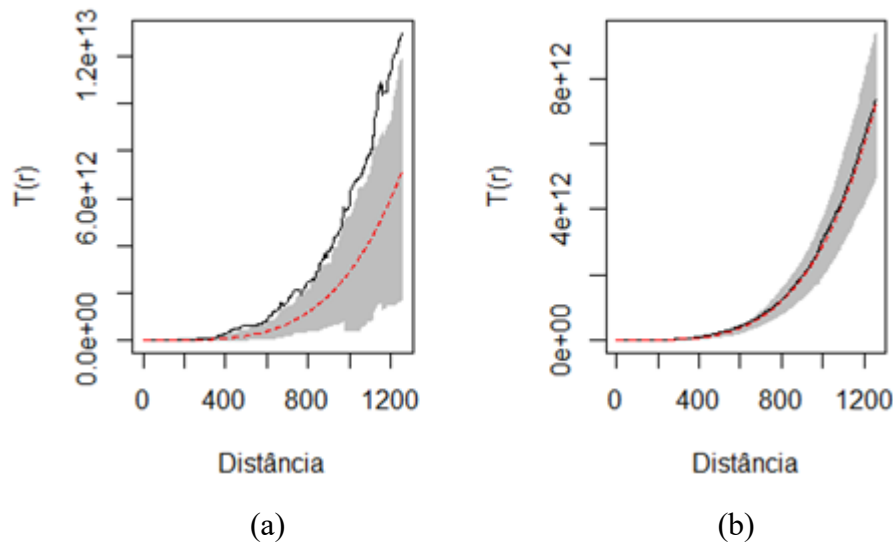
Fonte: Do autor (2018).

Pode-se observar pelas Figuras de 9 a 13 que as árvores da espécie *Antrocaryon amazonicum* e da espécie *Bombax humile* apresentam um padrão espacial com independência entre as árvores, uma vez que a curva observada se encontra totalmente dentro das bandas de confiança.

A Figura 14 apresenta os envelopes sob a suposição de ausência de efeitos de terceira ordem para as árvores da espécie *Antrocaryon amazonicum* (esquerda-a) e da espécie *Bombax*

*humile* (direita-b) utilizando a função  $T$ . Limite superior e inferior em cinza; função teórica (linha pontilhada em vermelho); função estimada (linha contínua preta).

Figura 14 – Envelope de aleatoriedade para a função  $T$ .



Fonte: Do autor (2018).

Os resultados apresentados na Figura 14-a mostram que a espécie *Antrocaryon amazonicum* apresentou efeitos de terceira ordem para as distâncias acima de 380 metros, aproximadamente, pois o gráfico observado não se encontra dentro das bandas de confiança para essas distâncias. Esse comportamento indica que as árvores dessa espécie apresentam um comportamento que exhibe dependência de terceira ordem na região do estudo.

Já no caso da espécie *Bombax humile* (Figura 14b) temos que o gráfico observado se encontra totalmente dentro das bandas de confiança, indicando que o comportamento dos pontos, para essa espécie de árvore, não apresenta efeitos de terceira ordem.

Os resultados encontrados neste trabalho mostram a importância de analisar as propriedades de terceira ordem na análise de configurações pontuais, pois é possível que uma configuração não apresente nem efeitos de primeira ordem nem efeitos de segunda ordem e, mesmo assim, podem apresentar efeitos de terceira ordem. Assim, ausência de efeitos de primeira e segunda ordem não é garantia que a configuração pontual seja uma realização de um processo de Poisson homogêneo.

## 6 CONCLUSÕES

A função  $T$ , que considera triplas de pontos ao invés de pares de pontos (como a função  $K$ ) mostrou-se adequada para caracterizar as propriedades de terceira ordem de uma configuração pontual no espaço. O uso da função  $T$  para analisar as configurações pontuais das duas espécies nativas da Floresta Amazônica (*Antrocaryon amazonicum* e *Bombax humile*) mostrou que a configuração da espécie *Antrocaryon amazonicum* apresenta dependência espacial de terceira ordem, enquanto as árvores da espécie *Bombax humile* apresenta uma configuração aleatória. A análise de terceira ordem baseada na função  $T$  pode ser extremamente útil para determinar as distâncias mínimas de plantio de árvores em trabalhos de reflorestamento de florestas nativas.

## REFERÊNCIAS

- ACRE. **Programa Estadual de Zoneamento Ecológico–Econômico do Estado do Acre:** zoneamento ecológico–econômico do Acre fase II: documento Síntese: escala 1:250.000. Rio Branco: Sema, 2006.
- AYALA, G.; SIMÓ, A. **Measuring interaction in bivariate point patterns with applications.** Jaume: Mathematics Department, 2000. (Preprint, 30).
- BADDELEY, A. *et al.* Estimating the J function without edge correction. **Statistica Neerlandica**, Amsterdam, v. 543, p. 315–328, 2000.
- BADDELEY, A.; SILVERMAN, B. A cautionary example on the use of second–order methods for analyzing point patterns. **Biometrics**, Washington, DC, v. 40, n. 4, p. 1089–1094, Dec. 1984.
- BADDELEY, A. J.; TURNER, R. Spatstat: an R package for analyzing spatial point patterns. **Journal of Statistical Software**, New York, v. 12, n. 6, p. 1–42, June 2005.
- BAILEY, T. C.; GATRELL, A. C. **Interactive spatial data analysis.** New York: Longman Scientific & Technical, 1995. 413 p.
- BESAG, J. E.; DIGGLE, P. Simple Monte Carlo tests for spatial pattern. **Journal of Applied Statistics**, Abingdon, v. 26, n. 3, p. 327–333, 1977.
- CAMARA, G. *et al.* **Análise espacial de dados geográficos.** Planaltina: Embrapa, 2004. 208 p.
- CRESSIE, N. A. C. **Statistics for spatial data.** New York: John Wiley, 1993. 900 p.
- DIGGLE, P. J. **Statistical analysis of spatial point patterns.** 2<sup>nd</sup> ed. London: Arnold, 2003. 159 p.
- HANISCH, K. H. Reduction of n–th moment measures and the special case of the third moment measure of stationary and isotropic planar point processes. **Statistics**, London, v. 14, n. 3, p. 421–435, 1983.
- HANISCH, K. H.; STOYAN, D. Once more on orientations in point processes. **Elektronisch Informations Verarbeitung und Kybernetik**, Berlin, v. 20, p. 279–284, 1984.
- JUSTEL, A.; PEÑA, D.; ZAMAR, D. A. **Multivariate Kolmogorov–Smirnov test of goodness of fit.** Madrid: Universidad Carlos III, 1994. (Statistics and econometrics series, 13).
- MECKE, J. Stationäre zufällige Maße auf lokalkompakten Abelschen Gruppen. **Z. Wahrscheinlichkeitstheorie und Verwandte Gebiete**, Berlin, v. 9, n. 1, p. 36–58, 1967.
- MESQUITA, C. C. de. **O clima do estado do Acre.** Rio Branco: IMAC, 1996. 53 p.

MÚLLER, J.; SYVERSVEEN, A. R.; WAAGEPETERSEN, R. P. Log gaussian cox processes. **Scandinavian Journal of Statistics**, Oxford, v. 25, n. 3, p. 451–482, Sept. 1998.

RIPLEY, B. D. Modelling spatial patterns. **Journal of the Royal Statistical Society. Series B. Methodological**, Hoboken, v. 39, n. 2, p. 172–212, June 1977.

RIPLEY, B. D. **Spatial statistics**. London: John Wiley, 1981. 340 p.

SCALON, J. D. *et al.* Spatial pattern analysis of second–phase particles in composite materials. **Materials Science e Engineering**, Lausanne, v. 356, n. 1/2, p. 245–257, Sept. 2003.

SCHABENBERGER, O.; GOTWAY, C. A. **Statistical methods for spatial data analysis**. Boca Raton: Chapman & Hall, 2005. 488 p.

SCHLADITZ, K.; BADDELEY, A. A third order point process characteristic. **Technical Report**, Western Australia, n. 1997/20, p. 1–17, 1997.

STOYAN, D.; KENDALL, W. S.; MECKE, J. **Stochastic geometry and its applications**. 2<sup>nd</sup> ed. Chichester: John Wiley, 1995. 436 p.

VAN LIESHOUT, M. N. M.; BADDELEY, A. J. A nonparametric measure of spatial interaction in point patterns. **Statistica Neerlandica**, Amsterdam, v. 50, p. 344–361, 1996.

## APÊNDICE A – CÓDIGOS PARA AS ANÁLISES DO CONJUNTO DAS ÁRVORES, UTILIZANDO O SOFTWARE R

Neste apêndice, são apresentados os códigos, para as análises do conjunto das árvores, utilizando o *software* R.

```
### Carregando o pacote Spatstat
```

```
library(spatstat)
```

```
setwd("D:/Minhas Pastas/DeltaStat/Consultorias/Ederson de Sousa/Confronto (Função T)")
```

```
### Entrada de dados
```

```
dados.antro<- read.csv("dados_antrocaryon.csv", header = T, sep = ";")
```

```
dados.antro
```

```
dados.bombax<- read.csv("dados_bombax.csv", header = T, sep = ";")
```

```
dados.bombax
```

```
### Escolher a janela para o estudo
```

```
w.antro<- ripras(dados.antro$Latitude, dados.antro$Longitude)
```

```
w.antro
```

```
w.bombax<- ripras(dados.bombax$Latitude, dados.bombax$Longitude)
```

```
w.bombax
```

```
### Função que transforma os dados em padrões de pontos
```

```
dados.antro.ppp<- ppp(x = dados.antro$Latitude, y = dados.antro$Longitude, window =  
w.antro, check = T)
```

```
dados.antro.ppp
```

```
dados.bombax.ppp<- ppp(x = dados.bombax$Latitude, y = dados.bombax$Longitude,  
window = w.bombax, check = T)
```

```
dados.bombax.ppp
```

```
#####
```

```
##
```

```
par(mfrow = c(1, 2))
```

```
compassRose<- dget("rosa_dos_ventos.R")
```

```
### Plots para a espécie Antrocaryon Amazonicum
```

```
plot(dados.antro.ppp, cex = 0.50, pch = 19, main = "")
```

```
compassRose(8952432, 578700.3, rot=0, cex=0.45) # Rosa dos ventos
```

```
densidade.antro<- density.ppp(dados.antro.ppp)
```

```
plot(densidade.antro, main = "", fil = T, col = terrain.colors(30))
```

```
plot(dados.antro.ppp, add = T, cex = 0.50, lty=4, pch=16, use.marks=TRUE)
```

```
compassRose(8953000, 578700.3, rot=0, cex=0.45) # Rosa dos ventos
```

```
### Plots para a espécie BombaxHumile
```

```
plot(dados.bombax.ppp, cex = 0.50, pch = 19, main = "")
```

```
compassRose(8952432, 578700.3, rot=0, cex=0.45) # Rosa dos ventos
```

```
densidade.bombax<- density.ppp(dados.bombax.ppp)
```

```
plot(densidade.bombax, main = "", fil = T, col = terrain.colors(30))
```

```
plot(dados.bombax.ppp, add = T, cex = 0.50, lty=4, pch=16, use.marks=TRUE)
```

```
compassRose(8952700, 578700.3, rot=0, cex=0.45) # Rosa dos ventos
```

```
par(mfrow = c(1, 1))
```

```
### Lambda para a espécie AntrocaryonAmazonicum
```

```
lambda.antro<- summary(dados.antro.ppp)$intensity
lambda.antro
```

```
### Lambda para a espécie BombaxHumile
```

```
lambda.bombax<- summary(dados.bombax.ppp)$intensity
lambda.bombax
```

```
#####
##
```

```
par(mfrow = c(1, 2))
```

```
persp(densidade.antro, col = "yellow", main = "AntrocaryonAmazonicum", zlab =
"Densidade")
```

```
persp(densidade.bombax, col = "yellow", main = "BombaxHumile", zlab = "Densidade")
```

```
#####
##
```

```
set.seed(4)
```

```
ks.antro<- cdf.test(dados.antro.ppp, function(x, y) {
```

```
x
```

```
})
```

```
plot(ks.antro, main = "AntrocaryonAmazonicum", xlab = "", ylab = "Probabilidade")
```

```
pvalor.antro<- ks.antro$p.value
```

```
pvalor.antro
```

```
set.seed(4)
```

```
ks.bombax<- cdf.test(dados.bombax.ppp, function(x, y) {
```

```
x
```



```

})
plot(ks.bombax, main = "BombaxHumile", xlab = "", ylab = "Probabilidade")
pvalor.bombax<- ks.bombax$p.value
pvalor.bombax

#####
##

### Gráficos da Função G

set.seed(7483)

G.antro<- envelope(dados.antro.ppp, Gest, correction = "border", nsim = 999)
plot(G.antro, main = "AntrocaryonAmazonicum", xlab = "Distância", ylab="G(y)", legend =
F)

set.seed(7483)

G.bombax<- envelope(dados.bombax.ppp, Gest, correction = "border", nsim = 999)
plot(G.bombax, main = "BombaxHumile", xlab = "Distância", ylab="G(y)", legend = F)

#####
##

### Gráficos da Função F

set.seed(2046)

F.antro<- envelope(dados.antro.ppp, Fest, correction = "border", nsim = 99)
plot(F.antro, main ="AntrocaryonAmazonicum", xlab = "Distância", ylab="F(y)", legend = F)

set.seed(2046)

F.bombax<- envelope(dados.bombax.ppp, Fest, correction = "border", nsim = 99)

```

```
plot(F.bombax, main = "BombaxHumile", xlab = "Distância", ylab="F(y)", legend = F)
```

```
#####
```

```
##
```

```
### Gráficos da Função J
```

```
set.seed(2658)
```

```
J.antro<- envelope(dados.antro.ppp, Jest, correction = "border", nsim = 99)
```

```
plot(J.antro, main = "AntrocaryonAmazonicum", xlab = "Distância", ylab="J(y)", legend = F)
```

```
set.seed(2658)
```

```
J.bombax<- envelope(dados.bombax.ppp, Jest, correction = "border", nsim = 99)
```

```
plot(J.bombax, main = "BombaxHumile", xlab = "Distância", ylab="J(y)", legend = F)
```

```
#####
```

```
##
```

```
### Gráficos da Função K
```

```
set.seed(4326)
```

```
K.antro<- envelope(dados.antro.ppp, Kest, correction = "border", nsim = 999)
```

```
plot(K.antro, main = "AntrocaryonAmazonicum", xlab = "Distância", ylab="K(y)", legend =  
F)
```

```
set.seed(4326)
```

```
K.bombax<- envelope(dados.bombax.ppp, Kest, correction = "border", nsim = 999)
```

```
plot(K.bombax, main = "BombaxHumile", xlab = "Distância", ylab="K(y)", legend = F)
```

```
#####  
##
```

```
### Gráficos da Função L
```

```
set.seed(535)
```

```
L.antro<- envelope(dados.antro.ppp, Lest, correction = "border", nsim = 999)  
plot(L.antro, main = "AntrocaryonAmazonicum", xlab = "Distância", ylab="L(y)", legend = F)
```

```
set.seed(535)
```

```
L.bombax<- envelope(dados.bombax.ppp, Lest, correction = "border", nsim = 999)  
plot(L.bombax, main = "BombaxHumile", xlab = "Distância", ylab="L(y)", legend = F)
```

```
#####  
##
```

```
### Gráficos da Função T
```

```
set.seed(1525)
```

```
T.antro<- envelope(dados.antro.ppp, Tstat, correction = "border", nsim = 99)  
plot(T.antro, main = "AntrocaryonAmazonicum", xlab = "Distância", ylab="T(r)", legend = F)
```

```
set.seed(1525)
```

```
T.bombax<- envelope(dados.bombax.ppp, Tstat, correction = "border", nsim = 99)  
plot(T.bombax, main = "BombaxHumile", xlab = "Distância", ylab="T(r)", legend = F)
```

```
par(mfrow = c(1, 1))
```