



VICTOR FERREIRA DA SILVA

**MÉTODOS PARA ANÁLISE DE CONFIGURAÇÕES
PONTUAIS EM REDES LINEARES**

LAVRAS – MG

2018

VICTOR FERREIRA DA SILVA

**MÉTODOS PARA ANÁLISE DE CONFIGURAÇÕES PONTUAIS EM REDES
LINEARES**

Dissertação apresentada à Universidade Federal de Lavras como parte integrante da avaliação para obtenção do título de mestre no Programa de Pós-Graduação em Estatística e Experimentação Agropecuária.

Prof. DSc. João Domingos Scalon
Orientador

**LAVRAS – MG
2018**

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

Silva, Victor Ferreira da.

Métodos para análise de configurações pontuais em redes
lineares / Victor Ferreira da Silva. - 2018.

49 p. : il.

Orientador(a): João Domingos Scalon.

.
Dissertação (mestrado acadêmico) - Universidade Federal de
Lavras, 2018.

Bibliografia.

1. Redes Lineares. 2. Processos Pontuais. 3. Manejo Florestal.
I. Scalon, João Domingos. . II. Título.

VICTOR FERREIRA DA SILVA

**MÉTODOS PARA ANÁLISE DE CONFIGURAÇÕES PONTUAIS EM REDES
LINEARES**

Dissertação apresentada à Universidade Federal de Lavras como parte integrante da avaliação para obtenção do título de mestre no Programa de Pós-Graduação em Estatística e Experimentação Agropecuária.

APROVADA em 18 de Maio de 2018.

Prof.Dr. João Domingos Scalon UFLA
Prof. Dr. Marcelo Silva de Oliveira UFLA
Prof. Dr. Deive Ciro de Oliveira UNIFAL

Prof. DSc. João Domingos Scalon
Orientador

**LAVRAS – MG
2018**

*A todos que compartilham da vontade do fogo, que batalham contra as adversidades com coragem para mostrar que diante de qualquer sistema falido ou adversidade, basta força de vontade e fé para se conquistar o improvável.
Dedico.*

AGRADECIMENTOS

Agradeço a Deus, por me dar paciência, forças e sabedoria para continuar caminhando mesmo quando exausto.

Minha mãe, Delfina e as outras duas Maria Madalena e Maria do Carmo por serem motivação, o carinho e o zelo de continuar vivendo e lutando para mostrar para as mesma o quanto são capazes de criar bem um filho. A minha irmã e meu pai por estarem sempre ao meu lado, e meus primos por estarem comigo sem e sendo mais que irmãos.

Aos professores Scalon, Deive, Joel e Daniel, principalmente pela paciência, dedicação, ensinamentos e conselhos.

Aos professores do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, pelo ensinamento de qualidade.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudos.

Ao meus amigos do DES por fazerem os dias na ufla mais interessantes e sempre de bom humor.

A Nádia, por fazer um excelente trabalho e simpatia mesmo quando debilitada.

A galera do "Gases Nobres" pelas festas, almoços, jantas e principalmente pela amizade.

A Kelly, por ser uma das pessoas especial que me ajudou muito.

A Fernanda por ter sido importante nesta etapa da minha vida e mostrar o quão brilhantes e frente do seu tempo as pessoas podem ser. A toda minha turma de mestrado, Ariana, Luciano, Mohamed, Lina, Marcela, Pablo, Tati e Allana pelo companheirismo e ajuda nas disciplinas.

Ao Édipo, Henrique Felipe pela amizade, companheirismo, concelhos e por dividir os dias dentro de casa.

Por fim, agradecer a todos que de forma direta ou indireta fizeram parte dessa caminhada.

MUITO OBRIGADO!!!

*“Nada pode ser obtido sem alguma espécie de sacrifício. É preciso oferecer em troca algo de
valor equivalente”
(Edward Elric - Fullmetal Alchemist)*

RESUMO

Com o intuito de compreender padrões utilizando processos pontuais em um cenário diferente do de costume onde temos uma área de estudo espacial na maioria das vezes bivariado, ou seja, as observações espalhadas por uma superfície onde os pontos observados podem se localizar em qualquer lugar nessa área, podendo ser georeferenciadas por latitude e longitude, tendo em vista como objetivo deste apresentar e descrever a metodologia em estudo quando os pontos estão posicionados sobre uma seção de reta, que quando unidas é chamada de rede linear. Sabendo que o Brasil possui a maior floresta tropical do mundo e nela é também situada a maior bacia hidrográfica do mundo. Será usado como rede linear um rio, que corta a Fazenda Canário situada no estado do Acre, que estará fazendo o papel de rede linear e os eventos georreferenciados observados nas margens desse rio será a espécie de árvore *Heura Crepitans* (açacú) com DAP acima da medida estabelecida para o abate. Sendo proposto analisar como se distribuem nesse tipo de cenário espacial de redes. A partir das análises de efeitos de primeira ordem, utilizando alisamento Kernel linear, onde mostrou-se melhor que o Kernel planar que por sua vez subestimou as taxas de intensidade. Usando função K-linear para efeitos de segunda ordem, os resultados mostram não haver agrupamento presente ao longo do rio, eficaz comparado a abordagem planar que verificou os padrões de agrupamento, assim o método se mostra eficaz para o estudo e importante para a análise em um censo florestal fluvial da Floresta Amazônica, definindo seu comportamento.

Palavras-chave: Redes Lineares. Processos Pontuais. Mata Ciliar. Manejo Florestal

ABSTRACT

In order to understand patterns using punctual processes in a scenario different from the usual one where we have a spatial area of study most often bivariate, that is, the observations scattered over a surface where the observed points can be located anywhere in that area, being able to be georeferenced by latitude and longitude, aiming to present and describe the methodology under study when the points are positioned on a section of line, that when united is called a linear network. Knowing that Brazil has the largest rainforest in the world and in it is also situated the largest hydrographic basin in the world. It will be used as a linear network a river, which cuts off the Fazenda Canário located in the state of Acre, which will be playing the role of linear network and the georeferenced event observed along the banks of this river will be the *Heura Crepitans* with properties above the established slaughtering measure. It is proposed to analyze how they are distributed in this type of network space scenario. From the analysis of first-order effects, using linear kernel smoothing, where it was shown to be better than the planar kernel which in turn underestimated the intensity rates. Using K-linear function for second-order effects, the results show no clustering present along the river, effective compared to the planar approach that verified the clustering patterns, thus the method is effective for the study and important for the analysis in a forest river census of the Amazon Forest, defining its patterns.

Keywords: Linear Network. Punctual Processes. Riparian Forest. Forest Management

LISTA DE FIGURAS

Figura 2.1 – Kernel Planar	14
Figura 2.2 – Função K planar	19
Figura 2.3 – Rede formada	21
Figura 2.4 – Estimador Kernel Linear	23
Figura 3.1 – (a) Área (b) Rios (c) APP	27
Figura 3.2 – Área de estudo	28
Figura 3.3 – Pontos ajustados para a rede	29
Figura 3.4 – Rede leste	30
Figura 3.5 – Rede linear Leste com as ocorrências	31
Figura 3.6 – Rede Oeste	32
Figura 3.7 – Rede linear Oeste com as ocorrências	33
Figura 4.1 – Alisamento Kernel Leste em tons de cinza	35
Figura 4.2 – Alisamento kernel Leste	36
Figura 4.3 – Alisamento kernel planar Leste	37
Figura 4.4 – Alisamento kernel linear Oeste em tons de cinza	38
Figura 4.5 – Alisamento kernel linear Oeste	39
Figura 4.6 – Alisamento kernel Oeste	40
Figura 4.7 – Função K-linear Leste	41
Figura 4.8 – Função K planar Leste	42
Figura 4.9 – Função K-linear Oeste	43
Figura 4.10 – Função K planar Oeste	44

SUMÁRIO

1	INTRODUÇÃO	9
2	REFERENCIAL TEÓRICO	12
2.1	Processos Pontuais	12
2.2	Processos pontuais espaciais	13
2.3	Efeitos de primeira ordem	13
2.4	Efeitos de segunda ordem	15
2.5	Processos Pontuais em redes	19
2.6	Efeitos de primeira ordem em um processo pontual em redes	21
2.7	Efeitos de segunda ordem em um processo pontual em redes	23
2.8	A Floresta Amazônica e seus rios	25
2.9	A Hura crepitans e suas peculiaridades	26
3	MATERIAL E MÉTODOS	27
3.1	O conjunto de dados	27
3.2	Métodos estatísticos	28
4	RESULTADOS E DISCUSSÃO	34
4.1	Propriedades de primeira ordem	34
4.2	Propriedades de segunda ordem	40
5	CONCLUSÃO	45
	REFERÊNCIAS	46
	APENDICE A – Script	48

1 INTRODUÇÃO

O estudo de Processos Pontuais apresentado por Diggle (1985), é uma das principais ferramentas usadas na análise estatística de dados georreferenciados e indexados por um sistema de coordenadas bem definidos, sendo amplamente discutida e conhecida. As análises espaciais são realizadas por meio de processos estocásticos que são definidos como uma coleção ou família de variáveis aleatórias definidas no mesmo espaço de probabilidade e indexadas por um conjunto de índices. Um processo espacial é um caso especial de um processo estocástico cuja indexação é a posição do evento.

Ao se analisar um conjunto de dados, muitas das vezes depara-se com observações que foram coletadas com suas respectivas coordenadas geográficas ou coordenadas indexadas por um sistema previamente definido. Nesse tipo de dado uma pergunta principal é se a ocorrência dessas observações são correlacionadas (interagem entre si) ou não. Caso as observações sejam espacialmente correlacionadas, o uso de métodos da estatística usual podem ser inapropriados. Nesse cenário, recomenda-se utilizar métodos da estatística espacial que, conforme Bailey e Gatrell (1995), pode ser definida como um conjunto de técnicas que busca descrever os padrões existentes em que os dados são espacialmente localizados e se considera explicitamente a possível importância de seu arranjo espacial na análise e/ou interpretação dos resultados.

O objetivo deste trabalho é analisar o padrão da distribuição espacial das árvores da espécie *Hura crepitans* na Floresta Amazônica. Mais especificamente, o interesse está em analisar o padrão da distribuição espacial das árvores dessa espécie que estão localizadas nas margens dos rios. Esse tipo de análise exige métodos específicos da área de processos pontuais espaciais que são denominados de processos pontuais espaciais em rede.

Utilizando a abordagem de processos pontuais em redes proposta por Okabe e Yamada (2001), tratando o rio como sendo uma rede linear, cuja metodologia mostrada por Baddeley, Rubak e Turner (2015), em seu livro, que é um compilado de técnicas a serem utilizadas em diversos cenários espacialmente referenciados, não apenas geograficamente. A abordagem em rede pode ser utilizada em rede de estradas, ruas, redes de neurônios entre outros cenários. Tendo uma melhoria significativa no desenvolvimento de novos estimadores exclusivos para esse tipo de ambiente, o estudo de processos pontuais em redes lineares abre um novo leque de opções para explicação de ocorrência de eventos, ou para alocação de recursos, de maneira mais eficaz e lucrativa. Este trabalho então apresenta como principal objetivo mostrar a metodologia de análise de processos pontuais em redes lineares, neste trabalho será mostrado as peculiaridades

ridades e formas de se utilizar as ferramentas que compõe essa forma de análise de processos pontuais.

Sabendo que, para o desenvolvimento sustentável de qualquer ambiente, a aplicação correta do manejo florestal, como o escrito por Gonçalves e Santos (2008), pode ser uma alternativa viável para a manutenção da Floresta Amazônica que, segundo o mesmo representa 50% da biomassa florestal do mundo. As floresta tropicais se destacam por possuírem uma diversidade de espécies arbóreas muito maior do que em outras florestas, sendo o maior exemplo a floresta amazônica(CHESSON, 2000).

Tendo em vista que muitos pesquisadores desconhecem os métodos utilizados para análise de configurações pontuais espaciais em rede, eles acabam utilizando os métodos usuais (planar) para análise de configurações pontuais no espaço. Nesta dissertação defende-se a tese que esta utilização inapropriada pode acarretar uma caracterização inapropriada das propriedades de primeira e segunda ordem da configuração pontual. Assim, a seguir, apresenta-se os resultados e discussão da análise da distribuição espacial das árvores da espécie *Hura crepitans*, em condições de abate, nas áreas de proteção permanente as margens do rio, pois, como descrito em Martinelli e Moraes (2013), esta espécie é abundante em rios e igarapés, existem em água pura. As margens desses rios devem ser bem cuidadas e protegidas, sendo assim, árvores oriundas dessas florestas são usadas para vários fins como extração de óleos, uso medicinal e de sua madeira para confecção de objetos. Serão aplicadas as ferramentas de análise de processos pontuais em redes lineares, para definir o comportamento da incidência das árvores. Utilizando os métodos oriundos das duas abordagens (planar e rede), para caracterizar os efeitos de primeira e segunda ordem e, conseqüentemente, testar a hipótese de que a distribuição das árvores dessa espécie segue a completa aleatoriedade espacial.

Deste modo, este trabalho terá a seguinte organização:

- Capítulo 2- Referencial Teórico: onde está apresentada uma revisão de literatura contendo uma breve introdução sobre a estatística espacial e as diferentes áreas de estudo; uma abordagem da teoria de processos pontuais sob os aspectos puramente espacial e em redes lineares e informações sobre as características do ambiente de estudo ;
- Capítulo 3- Material e Métodos: Uma explicação sobre o ambiente de estudo e os eventos considerados, informações sobre o conjunto de dados e a forma como foram tratados, e os softwares necessários para a execução das análises;

- Capítulo 4- Resultados e discussão: estão expostos os resultados obtidos em cada análise executada e uma discussão a respeito.
- Capítulo 5 - Conclusão: descreve a conclusão e as considerações a respeito do trabalho.
- Referências
- Apêndices

2 REFERENCIAL TEÓRICO

2.1 Processos Pontuais

Segundo Diggle (2003), a abordagem de processo pontual espacial pode ser visto como uma produção estocástica (mecanismo gerador), que por sua vez gera um conjunto de eventos contáveis (x_i, y_j) no plano, podendo ser estendido para qualquer número de dimensões. Assim, os dados provenientes de um processos pontual são pontos localizados no espaço e identificados por coordenadas do fenômeno em estudo. Em geral, o objetivo da análise está em determinar se a ocorrência dos pontos na área em estudo apresenta algum padrão espacial, como por exemplo, agrupamentos. Entre as áreas de aplicações, incluem Segurança Pública, Epidemiologia, Ciências Florestais, Ecologia, etc.

Diggle (2013) e Schabenberger e Gotway (2005) afirmam que dados de um processo pontual podem ser obtidos por amostragem ou mapeamento. Na amostragem, apenas um subconjunto dos eventos gerados pelo processo estocástico são observados e considerados na análise. Se todos os eventos gerados pelo processo estocástico forem observados e considerados na análise, os dados do processo pontual são denominados de dados mapeados ou mapa de pontos. No presente trabalho será considerado apenas mapa de pontos.

Em geral, processos pontuais são divididos em duas categorias:

- Não marcados: Considera apenas um tipo de evento, por exemplo, árvores de uma única espécie ou ocorrências de uma doença.
- Marcados: Considera uma variável aleatória associada a cada observação(ponto) do processo, por exemplo, altura ou espécie de uma árvore.

Também temos os processos divididos em:

- Planar: Considera os eventos (pontos) ocorrendo em qualquer posição do espaço.
- Redes lineares: Considera os eventos (pontos) ocorrendo em uma rede linear.

Como objeto de estudo no presente trabalho, analisaremos apenas a categoria dos processos pontuais em redes lineares não marcados.

2.2 Processos pontuais espaciais

Um processo pontual espacial é, segundo Diggle (2013) definido formalmente, como um mecanismo estocástico que gera um conjunto contável finito ou infinito de pontos (denominados eventos) $\{\mathbf{s}_i : i = 1, 2, \dots\}$ tais que $\mathbf{s}_i \in S \subset \mathbb{R}^d$, ou seja, cada variável aleatória denota a localização de um evento específico no espaço d-dimensional.

Como descrito em Baddeley, Bárány e Schneider (2007), pode-se descrever os processos pontuais em termos de momentos como o valor esperado (efeitos de primeira ordem) e variância ou covariância (efeitos de segunda ordem), de modo análogo ao que é feito para uma variável aleatória. Essas quantidades, segundo o autor, são úteis no estudo teórico de processos pontuais e na inferência estatística sobre padrões pontuais.

2.3 Efeitos de primeira ordem

Importantes aspectos no comportamento de um processo estocástico espacial podem ser caracterizados em termos de efeitos de primeira ordem e segunda ordem. Os efeitos de primeira ordem, também conhecidos como globais ou de larga escala, descrevem como o valor esperado (média) varia dentro do espaço. O objetivo da análise está em estimar a intensidade do processo, ou seja, o número de eventos por unidade de área.

As propriedades de primeira ordem do processo pontual espacial podem ser definidas pela função intensidade dada por

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

em que ds define uma pequena região em torno da localização \mathbf{s} , $|ds|$ é a sua área e $N(ds)$ é número de eventos localizados dentro de ds .

Estimador de intensidade Kernel

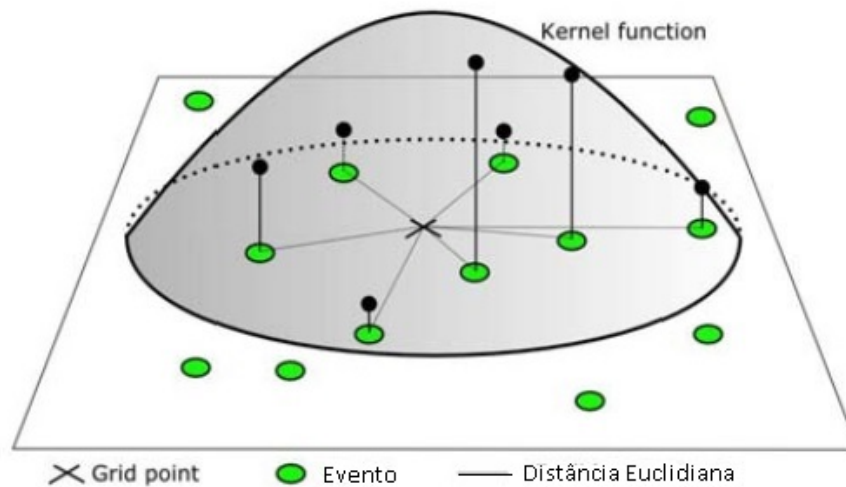
Neste caso para o, considere \mathbf{s}_i uma localização qualquer em uma área A e s_1, \dots, s_n são as localizações dos n eventos observados nesta área. Um estimador de intensidade $\lambda(\mathbf{s})$ na posição \mathbf{s} , com correção do efeito de borda (que nada mais é que um ajuste para situação de presença de ocorrências nos limites da área estudada), proposto por Diggle (1985), pode ser expressado por:

$$\hat{\lambda}(s) = \frac{1}{p_h(s)} \sum_{i=1}^n \frac{1}{h^2} k\left(\frac{s-s_i}{h}\right), \quad (2.1)$$

em que h é denotado como raio de influência ou largura de banda, nesse trabalho usaremos o termo *bandwith* para o mesmo.

A função Kernel planar $k(\cdot)$ pode ser geometricamente vista como:

Figura 2.1 – Kernel Planar



Pontuando temos que $k\left(\frac{s-s_i}{h}\right) du$ é um fator de correção do efeito de bordas e $k(\cdot)$ não necessariamente precisa ser uma função de densidade de probabilidade, por ser usada como um fator ponderador. Segundo Schabenberger e Gotway (2005), as das funções de densidade mais aplicada é:

- Kernel quadrática:

$$k(t) = \begin{cases} \frac{3}{4} (1 - t^2), & \text{se } |t| \leq 1; \\ 0, & \text{caso contrário.} \end{cases} \quad (2.2)$$

De acordo com Schabenberger e Gotway (2005), a escolha de $K(\cdot)$ tem menos importância que a escolha da largura da banda. Se h é pequena, as estimativas quase não apresenta viés, no mais apresentarão variabilidade alta. No entanto, com o aumento da largura da janela h , a estimativa é suavizada, variando menos, entretanto, ficando mais tendenciosa. Para definir um valor “adequad” para h existem várias possibilidades (ex. erro quadrático médio como validação cruzada) conforme podem ser vistos em Diggle (2013) e Schabenberger e Gotway (2005).

2.4 Efeitos de segunda ordem

Os efeitos de segunda ordem, ainda de acordo Câmara e Carvalho (2004), são aqueles “denominados locais ou de pequena escala, representam a dependência espacial no processo, proveniente da estrutura de correlação espacial”.

Na análise dos efeitos de segunda ordem procura-se responder se os eventos estão mais próximos ou mais distantes entre si, exibindo algum grau de relação do que seria esperado no caso de completa aleatoriedade espacial (CAE).

Segundo Schabenberger e Gotway (2005), a CAE ocorre se o padrão pontual atende aos seguintes critérios:

- O número médio de eventos por unidade de área, ou seja, a intensidade $\lambda(\mathbf{s})$ é homogênea em S ,
- o número de eventos em duas sub-regiões sem sobreposição A_1 e A_2 são independentes,
- o número de eventos em qualquer sub-região segue distribuição Poisson.

Desse modo, os eventos que distribuem-se uniformemente e independentemente ao longo do domínio e a manifestação matemática da completa aleatoriedade espacial é o chamado processo de Poisson homogêneo. Este é um processo que não exhibe qualquer estrutura espacial e serve como hipótese nula para muitas investigações estatísticas em padrões pontuais.

Os efeitos de segunda ordem, considerados como locais ou de pequena escala, representam a correlação entre valores dos eventos do processo em diferentes regiões do espaço.

As propriedades de segunda ordem podem ser formalmente definidas, conforme Diggle (2013), pela função intensidade de segunda ordem dada por:

$$\lambda_2(\mathbf{s}_i, \mathbf{s}_j) = \lim_{|ds_i|, |ds_j| \rightarrow 0} \frac{E[N(ds_i)N(ds_j)]}{|ds_i||ds_j|},$$

em que $N(ds_i)$ e $N(ds_j)$ representam o número de eventos dentro pequenos círculos centrados nas localizações \mathbf{s}_i e \mathbf{s}_j com áreas iguais a $|ds_i|$ e $|ds_j|$, respectivamente.

Uma medida fortemente relacionada a essa propriedade é a intensidade condicional $\lambda_c(s_i, s_j) = \frac{\lambda_2(s_i, s_j)}{\lambda(s_j)}$, a qual corresponde a intensidade no ponto s_i condicional a informação de que existe um evento em s_j . Outra medida é a função de correlação de pares, a qual é baseada no produto de densidade de segunda ordem e, pode ser expressa por

$$g(s_i, s_j) = \frac{\lambda_2(s_i, s_j)}{\lambda(s_i)\lambda(s_j)}, \quad (2.3)$$

Dois conceitos que são importantes para a análise estatística de um processo pontual e que estão intimamente relacionados aos efeitos de primeira e segunda ordem são os conceitos de estacionariedade e isotropia. Um processo é dito estacionário ou homogêneo se os efeitos de primeira e segunda ordem são invariantes sob a translação, ou seja, a intensidade de primeira ordem $\lambda(\mathbf{s}) = \lambda$ não varia no espaço e a intensidade de segunda ordem $\lambda_2(s_i, s_j) = \lambda_2(s_i - s_j)$ depende apenas da distância relativa entre os pontos. Pode ser visto em Cressie (1993), um processo que é tanto estacionário como isotrópico é chamado de “movimento invariante”, dado que a covariância varia apenas com a distância entre dois pontos e, portanto, os efeitos direcionais são ausentes, assim tem-se que

$$\lambda_2(\mathbf{s}_i, \mathbf{s}_j) = \lambda_2(t),$$

em que $t = \|s_i - s_j\|$ é a distância euclidiana entre dois pontos, e como consequência da independência de um processo de CAE, $\lambda_2(t) = \lambda^2$.

A hipótese nula de Completa Aleatoriedade Espacial

O mais simples modelo de padrão pontual teórico denomina-se Completa Aleatoriedade Espacial (CAE) que apresenta como característica não apresentar efeitos de primeira e segunda ordem. Sob este modelo, a configuração apresenta intensidade de primeira ordem constante e independência entre os eventos. Em outras palavras, tomando (CRESSIE, 1993), diz que: "processos espaciais pontuais que exibem uma CAE são caracterizados pela ausência de estrutura nos dados." O termo CAE é usado como sinônimo de um processo homogêneo de Poisson com parâmetro (intensidade) λ . Assim, este modelo é utilizado na maioria das análises como a hipótese nula em um teste cujas hipóteses alternativas estabelecem algum tipo de estrutura espacial (regularidade e agrupamentos).

Função K Planar

A função K de Ripley (1976) e Ripley (1977), também conhecida como "função de segunda ordem reduzida" e que para fins de melhor entendimento será dita neste trabalho por "Função K Planar" está inteiramente ligada a interação dos diferentes eventos do processo pontual em análise.

Segundo Cressie (1993), a função K tem como vantagem detectar o padrão espacial pontual em diferentes escalas de distâncias simultaneamente. Além disso, a função K possibilita analisar o grau de dependência entre dois ou mais processos.

A função K univariada é definida por Ripley (1976) e Ripley (1977) como sendo:

$$K(t) = \lambda^{-1}E[N(t)], \quad (2.4)$$

- $N(t)$ é o número de eventos encontrados em uma distância t ;
- λ é a intensidade de primeira ordem do processo pontual.

Com intenção de estabelecer uma relação entre $K(t)$ e $\lambda_2(t)$, segundo Diggle (2013), assume-se que o processo seja ordenado. Sendo assim, não deverá ocorrer múltiplos eventos sobrepostos, ou seja, $P\{N(ds) > 1\}$ é uma ordem de grandeza menor que $|ds|$. Significando que $E[N(ds)] \sim P\{N(ds) = 1\}$ no sentido que a relação destas duas quantidades tende para 1 quando $|ds_i| \rightarrow 0$. De maneira análoga, $E[N(ds_i)N(ds_j)] \sim P\{N(ds_i) = N(ds_j) = 1\}$. Observando essas condições, o número esperado de eventos dentro da distância t de um evento arbitrário pode ser obtido integrando a função intensidade condicional sobre um disco com centro na origem e raio t . Assim,

$$K(t) = \lambda^{-1} \int_0^{2\pi} \int_0^t \lambda_c(s|o) s ds d\theta.$$

Considerando que, $\lambda_c(s|o) = \lambda_2(s)/\lambda$, conseqüentemente,

$$K(t) = 2\pi\lambda^{-2} \int_0^t \lambda_2(s) s ds. \quad (2.5)$$

Conforme definido anteriormente, a função $K(t)$ representa o número esperado de pontos dentro de uma distância t de um ponto arbitrário. Assim, para um processo de CAE, $\lambda_2(s) = \lambda^2$ e, portanto, a equação (2.5) se reduz a

$$\begin{aligned} K(t) &= 2\pi\lambda^{-2} \int_0^t \lambda_2(s) s ds \\ &= 2\pi\lambda^{-2} \lambda^2 \frac{t^2}{2} \\ &= \pi t^2. \end{aligned}$$

O estimador mais simples para λ utilizado na função K univariada definida na equação (2.4) é o número observado de eventos por unidade de área, ou seja, $\hat{\lambda} = n/|A|$. Segundo Diggle (2013), considerando que $E(t) = E[N(t)]$, segue que a interpretação é a mesma, ou seja, o número esperado de eventos dentro de uma distância t de um evento arbitrário. Assim, um estimador para $E(t)$ pode ser expresso por

$$\tilde{E}(t) = n^{-1} \sum_{i=1}^n \sum_{j \neq i}^n I(t_{ij} \leq t), \quad (2.6)$$

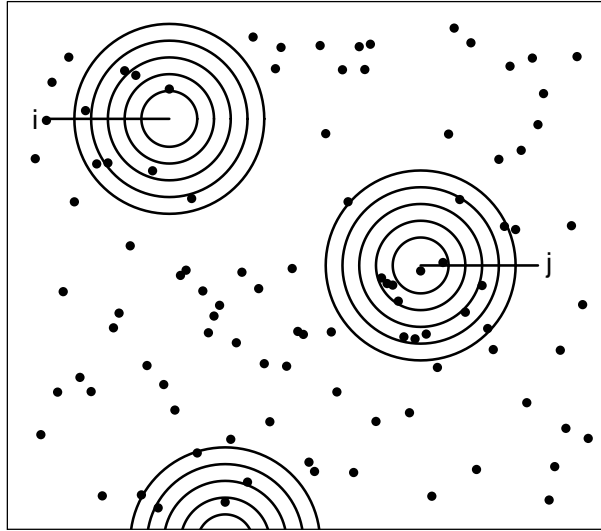
em que $I(\cdot)$ representa a função indicadora e $t_{ij} = \|s_i - s_j\|$.

Portanto, um estimador não tendencioso (corrigido para efeito de bordas) para a função K , proposto por Ripley (1976), é dado por

$$\hat{K}(t) = \frac{|A|}{n^2} \sum_{i=1}^n \sum_{j=1}^n \left(\frac{I(t_{ij})}{w_{ij}} \right). \quad (2.7)$$

Deve-se observar que $K(t)$ é uma função acumulada, pois são traçados círculos com origem iguais em um evento arbitrário até a uma distância máxima t , em que são computados todos os eventos dentro desses círculos. Uma ilustração da estimação da função K pode ser observada na Figura 2.2. Para detectar um padrão são considerados outros pontos arbitrários e esse procedimento é simulado várias vezes.

Figura 2.2 – Função K planar



Fonte: Adaptado de Bailey e Gatrell (1995)

2.5 Processos Pontuais em redes

Baddeley, Rubak e Turner (2015) definem uma situação particular da anterior, sendo ela mais restrita, em que o processo pontual em rede L é tratado como um mecanismo que gera um número de pontos e localizações finitas e aleatórias.

Seja $\mathbf{x} = [x_1, \dots, x_n]$, $x_i \in L$, $n \geq 1$ uma realização do processo pontual \mathbf{X} na rede L .

Seguindo, temos por definição que se \mathbf{X} é um processo pontual em uma rede L e $\mathbf{X} \subseteq L$, então $N_{\mathbf{X}}$ é definido como número de pontos de \mathbf{X} em r , logo temos que:

- r = comprimento da janela tomada para estudo chamada de *bandwidth* (largura de banda). Uma medida usada para definir o tamanho da janela de onde será feito os cálculos, ficando mais claro seu sua a medida que avançarmos nos detalhes metodológicos.
- $N_{\mathbf{X}}$ = número de pontos em $\mathbf{X} \cap S$, ou seja, ocorrências dos eventos em estudo.

Uma característica de um processo pontual em rede que nos dá a ideia básica do seu padrão é a intensidade que mede a taxa média de ocorrência de pontos por comprimento da unidade. A intensidade pode variar em para diferentes locais.

Um processo pontual \mathbf{X} em uma rede \mathbf{L} é dito ter *Função de intensidade [de primeira ordem]* $\lambda(\mathbf{u}), \mathbf{u} \in L$ se, para todos segmento $\mathbf{S} \subseteq \mathbf{L}$,

$$\mathbb{E}[\mathbf{N}_{\mathbf{X}}] = \int_{\mathbf{S}} \lambda(\mathbf{u}) d\mathbf{u}, \quad (2.8)$$

em que $\lambda(\mathbf{u})$ pode ser interpretado, intuitivamente, como o número esperado de pontos por unidade de comprimento em um vizinhança ou intervalo de \mathbf{u} .

Definições básicas de uma Rede

Para a análise proposta nesse trabalho, precisamos de uma condição específica das nossas observações, que é estarem posicionadas sobre uma rede. Sendo assim tomaremos o cuidado de ao longo do trabalho nomear todos estimadores e funções como *espacial*, quando se tratar de um processo pontual normal em \mathbb{R}^2 e *linear* quando for abordado o processo pontual em rede.

Começaremos com algumas definições que tornam interseções de segmentos de retas em uma rede, conforme apresentadas por Fuhrmann e Helmke (2015):

Tomando uma seção linear $l \subset \mathbb{R}^2$ a mesma é chamada de segmento de reta em um plano, com limites u e v inferior e superior, respectivamente, que pode ser expressa formalmente por:

$$l = l_{u,v} = \{tu + (1-t)v : 0 \leq t \leq 1\},$$

para alguns pontos $u, v \in \mathbb{R}^2, \forall u \neq v$ é tido como segmento de reta entre os pontos u e v , onde o módulo (sua amplitude) é chamado de distância euclidiana em \mathbb{R}^2 dada por:

$$|l| = |l_{u,v}| = \|u - v\|$$

Logo podemos definir uma rede \mathbf{L} em \mathbb{R}^2 como a união finita de segmentos de reta l sendo:

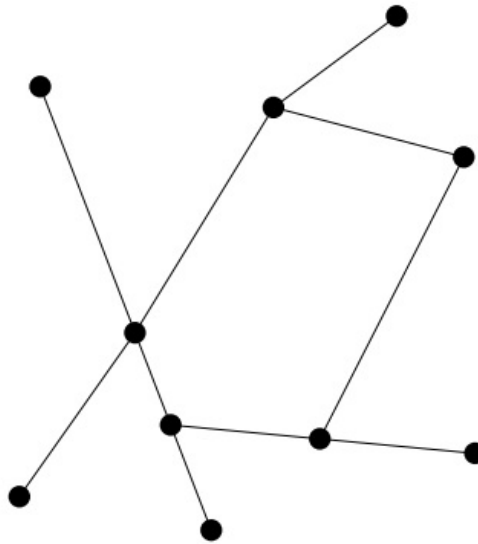
$$\mathbf{L} = \cup_{i=1}^n l_i.$$

Sabendo que um segmento de reta $l_i = l_{u_i, v_i}$, temos $l_i \cap l_j = \emptyset, \forall i \neq j$, o comprimento total de \mathbf{L} pode ser escrito como:

$$|\mathbf{L}| = \sum_{i=1}^n |l_i|$$

Seja \mathbf{L} uma rede, então um ponto $v \in \mathbf{L}$ é chamado de vértice de interseção se $v = l_i \cap l_j$ para algum $i, j, i \neq j$ e este será chamado de vértice livre se $l_i = l_{x,v}$ (ou $l_{v,x}$), para algum i, x , e $l_i \cap l_j \neq v, \forall i \neq j$. Será chamado de *interseção* a união dos vértices de interseção e os vértices livres. Temos então, ao fim dos conceitos definidos acima, a seguinte representação geométrica para uma rede:

Figura 2.3 – Rede formada



(●) são as interseções e (-) são segmentos de reta

2.6 Efeitos de primeira ordem em um processo pontual em redes

O comportamento de um processo estocástico em rede podem ser caracterizados em termos de efeitos de primeira ordem. Os efeitos de primeira ordem, também conhecidos como globais ou de larga escala, descrevem como o valor esperado (média) varia dentro da rede. O objetivo da análise está em estimar a intensidade do processo, ou seja, o número de eventos por unidade de comprimento.

Os efeitos primeira ordem do processo pontual em rede \mathbf{X} , podem ser definidos pela função intensidade dada por:

$$\lambda(u) = \lim_{|du| \rightarrow 0} \left\{ \frac{\mathbb{E}[\mathbf{N}_{\mathbf{X}}(du)]}{|du|} \right\},$$

em que du define um pequeno intervalo em torno de um ponto u situado na rede, onde o número esperado de pontos por unidade de comprimento é constante e independente da localização. Assim, quando a intensidade é igual a algum λ constante diz-se que é a intensidade de primeira ordem é estacionária (homogênea), com função de intensidade $\lambda(u) = \lambda$, em que λ pode ser interpretada como o número esperado de pontos por unidade de comprimento.

Com isso podemos definir L sendo uma rede e \mathbf{X} um processo pontual em L . Então podemos dizer que um \mathbf{X} é estacionário de primeira ordem para qualquer $du \subseteq L$.

Estimador de intensidade global

Nas análises de dados pontuais em geral é relevante explorar o efeito de primeira ordem de um processo estacionário e não estacionário. Nesse contexto, deve-se estimar a intensidade. Para um processo de Poisson homogêneo, o estimador de intensidade, como vimos em Baddeley, Rubak e Turner (2015), é o número observado de eventos por comprimento de rede que pode ser expresso por:

$$\hat{\lambda} = \frac{\mathbf{N}_{\mathbf{X}}(L)}{|L|}. \quad (2.9)$$

Em que: $\mathbf{N}_{\mathbf{X}}(L)$ é o número de eventos observado na rede, $|L|$ comprimento total da rede e $\hat{\lambda}$ a intensidade global da rede.

Pensando em método não paramétrico o mais usual para se estimar a intensidade de um processo não estacionário, é o alisamento kernel que cria uma relação entre as estimativas de densidade e intensidade. Esta função realiza uma contagem de todos os pontos dentro de uma região de influência, ponderando-os pela distância de cada um à localização de interesse (CÂMARA; MONTEIRO; MEDEIROS, 2003).

Estimador de local intensidade Kernel Linear

Quando tratamos de uma ambiente não planar ou seja em uma rede, os estimadores espaciais costumam apresentar uma estimativa super estimada, como foi discutido em Borruso (2008) que propõe um estimado de densidade para ambientes lineares, Timothée et al. (2010) em seu trabalho propõe uma variação da função kernel quadrática, sendo descrita como:

- A kernel de redes (**NetKDE**):

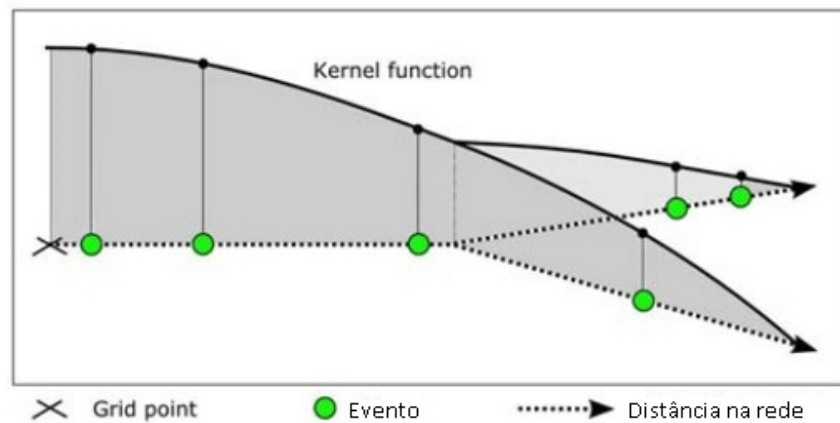
$$k_{net}(t_{net,i}) = \begin{cases} \frac{1}{3\pi} (1 - t_{net,i}^2), & \text{se } t_{net,i}^2 \leq 1; \\ 0, & \text{caso contrário.} \end{cases} \quad (2.10)$$

Com $t_{net,i} = d_{net,ij}/h$ em que r é a largura de banda definida e $d_{net,ij}$ a distância ponto do *grid* x_j ao evento observado n_i medido sobre a rede. Podemos então denotar x_j no caso Kernel linear por:

$$NetKDE(x_j) = \frac{1}{nr^2} \sum_{i=1}^n K_{net}(t_{net,i}) \quad (2.11)$$

em que n é o número de eventos na largura de banda h , a figura 2.4 ilustra geometricamente o comportamento do estimador, com a rede em vista lateral, onde a altura é a intensidade kernel de cada ocorrência :

Figura 2.4 – Estimador Kernel Linear



Fonte: Adaptado de Timothée (2010)

Vista do estimador em perspectiva, onde os pontos verdes são os eventos observados, e a altura até a linha em negrito sua intensidade estimada pela função.

2.7 Efeitos de segunda ordem em um processo pontual em redes

Podemos definir a intensidade do segundo momento (ou densidade do produto) $\lambda_2(u, v)$. Heuristicamente, é apresentado por Baddeley, Rubak e Turner (2015) que dado dois locais distintos u, v na rede, consideramos um segmento muito curto da rede em torno da localização u do comprimento d_1u , e de forma semelhante um curto pedaço de rede em torno de v de comprimento d_1v . A probabilidade de $p(u, v)$ que ambos estes curtos os segmentos contêm pelo

menos um ponto aleatório. ser $p(u, v) = \lambda_2(u, v)d_1ud_1v$. Então, para quaisquer segmentos de duas linhas $A, B \subset L$ que sejam disjuntos, temos:

$$\mathbb{E}[n(\mathbf{X} \cap \mathbf{A}), n(\mathbf{X} \cap \mathbf{B})] = \lambda^{-1} \int_A \int_B \lambda_2(u, v)d_1vd_1u \quad (2.12)$$

Por exemplo, o processo pontual em rede, com função de intensidade $\lambda(u)$, tem intensidade segundo momento $\lambda_2(u, v) = \lambda(u)\lambda(v)$.

Função K-Linear

Como já comentado anteriormente, o ambiente de estudo é linear, portanto, temos que ter um cuidado especial ao tomar nossos estimadores para não ocorrer nenhum problema de super estimação. Okabe e Yamada (2001) utilizaram a função K de Ripley (1976) como base para propor uma nova função para analisar os padrões espaciais e interações quando o processo ocorrem em uma configuração de redes lineares:

- Função K -linear (K_{OY}):

$$\hat{K}_{OY}(r) = \frac{1}{\hat{\lambda}^2 |L|} \sum_i \sum_{j \neq i} \mathbf{1}(d_L(x_i, x_j) \leq r). \quad (2.13)$$

onde $\hat{\lambda}$ é a intensidade global do processo, $|L|$ comprimento da rede, $d_L(x_i, x_j)$ distância entre uma ocorrência x_i e x_j em uma distância r e $\mathbf{1}(\cdot)$, função indicadora.

É explicitado em Capretz et al. (2012) que uma das principais estatísticas descritivas na análise de processos pontuais é a Função K . Sua principal vantagem é permitir a detecção do padrão espacial em diferentes escalas de distâncias simultaneamente.

- A função K -linear pode ser utilizada para testar a hipótese nula CAE:

Tipo de Processo	$\hat{K}_{OY}(r)$
Aleatório	$= r$
Regularidade	$< r$
Agrupamento	$> r$

sendo testadas contra a hipótese de completa aleatoriedade espacial, para rede.

Análise contra a hipótese CAE em redes lineares

Para a análise da interação entre os eventos de uma configuração pontual observada, usualmente, compara-se $\hat{K}_{OY}(r)$ com a função teórica K sob a suposição de hipótese nula de

CAE, ou seja, $K_{OY}(r) = r$, para um conjunto de distâncias r . Para avaliar se a diferença entre essas duas funções é estatisticamente significativa, pode-se utilizar tanto testes de hipóteses formais quanto procedimentos gráficos conhecidos como envelopes de confiança. Em ambos os casos utiliza-se simulações de Monte Carlo como vista em Scalon e Silva (2006) e adapta par as redes lineares em Baddeley, Rubak e Turner (2015) para o caso de rede lineares.

Os métodos de Monte Carlo para a construção de envelopes simulados que fornecem uma maneira conveniente de enfrentar este problema. Assim, são estabelecidas as seguintes etapas:

- i) Calcula-se o valor de $\hat{K}_{OY_1}(r)$ para o processo pontual observado;
- ii) Calcula-se $\hat{K}_{OY_i}(r) : i = 2, 3, \dots, s$ para cada $s - 1$ simulações independentes e identicamente distribuídas na rede linear \mathbf{L} sob a hipótese de CAE;
- iii) São construídos os envelope de simulações superior e inferior, em que são considerados os valores máximos e mínimos de $\hat{K}_{OY_i}(r)$, definidos como:

$$U(r) = \max\{\hat{K}_{OY_i}(r)\}, i = 1, \dots, s. \quad (2.14)$$

$$L(r) = \min\{\hat{K}_{OY_i}(r)\}, i = 1, \dots, s. \quad (2.15)$$

- iv) Constrói-se um gráfico, em que os valores de $\hat{K}_{OY_i}(r)$, $U(r)$ e $L(r)$ são colocados no eixo das coordenadas e as distâncias r no eixo das abcissas.
- v) Cálculo do nível de significância dado por $2(k + 1)/(s + 1)$, onde k é o rank escolhido para os limites e s o número de simulações.

O teste descrito anteriormente determina qual direção pela interpretação gráfica, para qualquer distância r , se a linha de $\hat{K}_{OY_i}(r)$ estiver acima do envelope, há evidências de um padrão de agrupamento. Caso $\hat{K}_{OY_i}(r)$ estiver abaixo do envelope, há evidências do padrão de regularidade.

2.8 A Floresta Amazônica e seus rios

A Amazônia é representada por uma região geográfica que ocupa quase todo centro oeste da América do Sul, constituindo-se no maior ecossistema florestal do mundo. Em termos

de Brasil, essa região ocupa uma superfície de $38160 h^{-1}$, dos quais uma grande parcela é ocupada por floresta de terra firme. Face as suas dimensões, bem com a o que representa a região, em termos de ambiente e recursos naturais para o País, torna-se evidente que toda análise do ecossistema amazônico e de importância significativa, tendo-se por objetivo a sua exploração racional e adequada (LEOPOLDO et al., 1982).

Pode ser encontrado em FILIZOLA (2006) uma breve amostra do cenário da hidrografia amazônica, na qual mostra que é está composta por todos os afluentes e rios formadores do Rio Amazonas. Ela drena sete países e corresponde a quase 40% da América do Sul, com uma área de 6,6 milhões de quilômetros quadrados (incluindo as macrobacias consideradas parcialmente amazônicas, como as do Araguaia e do Tocantins). A Bacia Amazônica é o maior compartimento de água doce superficial do planeta, com cerca de 15% do total disponível desse recurso.

2.9 A *Hura crepitans* e suas peculiaridades

Ocorre em toda a Amazônia, em matas de várzea, de solo argiloso e alagadiço, margeando os rios. Distribui-se nas Guianas, Antilhas, América Central, Peru e Bolívia. A (*Hura crepitans* L. *Euphorbiaceae*), conhecida regionalmente como Assacú, uma das árvores cuja madeira está entre as mais exportadas desde 1998 Parolin (2000). É uma planta de grande porte, que pode chegar até 40 metros de altura, possui casca lisa, parda e repleta de espinhos. Suas folhas são compostas e pecioladas; além de possuir pelos na nervura principal. Suas flores masculinas são em formato de espigas e as femininas em forma de pedúnculo, aparecem de maio a junho.

O seu emprego comercial por Iwakiri et al. (2011) se dá na confecção de compensados, obras internas, caixotaria leve, brinquedos, tamancos, artefatos de madeira, miolo de painéis, de portas, para contraplacados e chapas de partículas. Sua aplicação deve ser precedida de tratamento preservante. Os troncos das grandes árvores são utilizados pelos índios para a construção de canoas e os exploradores de madeiras usam as grandes toras para sustentar peças de madeira mais pesadas que a água, quando baixam os rios (chamada de boieira).

Na medicina popular a espécie também se destaca, o sincretismo da etnia Apurinã do Amazonas utiliza o suco da *Hura crepitans* para envenenamento das pontas das flechas para caçadas. Já foi muito comum no estado do Pará a utilização de seu látex para cura da lepra. Seu látex é bem prejudicial às mucosas. Suas cascas são usadas para alívio dos terríveis furúnculos.

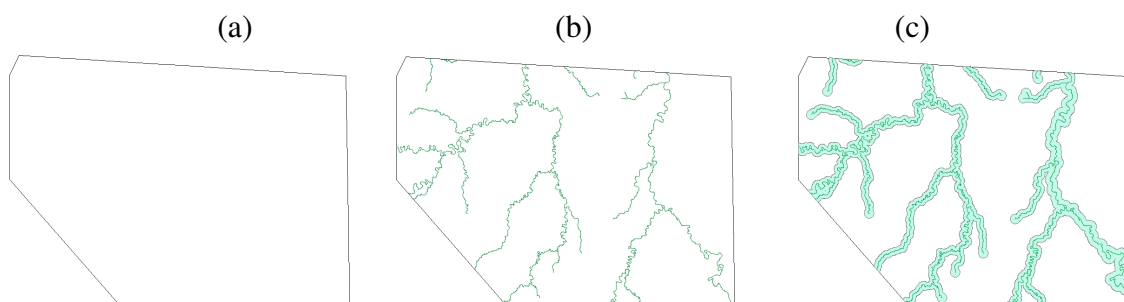
3 MATERIAL E MÉTODOS

3.1 O conjunto de dados

O presente estudo é parte de um projeto feito para o censo florestal foi realizado na Fazenda Canary, localizada no município de Bujari-Acre. A propriedade detém uma área total de 15.810,14 ha, sendo 8.446,80 ha de reserva legal, deixando uma área efetiva de manejo florestal de 4465,89 ha. Entretanto, no estudo em proposta será utilizada uma área menor pois será avaliada apenas regiões com concentração fluvial, utilizando como rede a extensão da área de proteção permanente (APP). A vegetação existente é constituída de Floresta Amazônica nativa com bambu, palmeira e espécies comerciais (ACRE, 2006).

Tomamos para o presente estudo uma área de 785 ha, uma APP de 160,40 ha, e os rios que tem aproximadamente 27,25 km de comprimento serão tratados como os segmentos de retas da rede. Como descrito na figura 3.1:

Figura 3.1 – (a) Área (b) Rios (c) APP

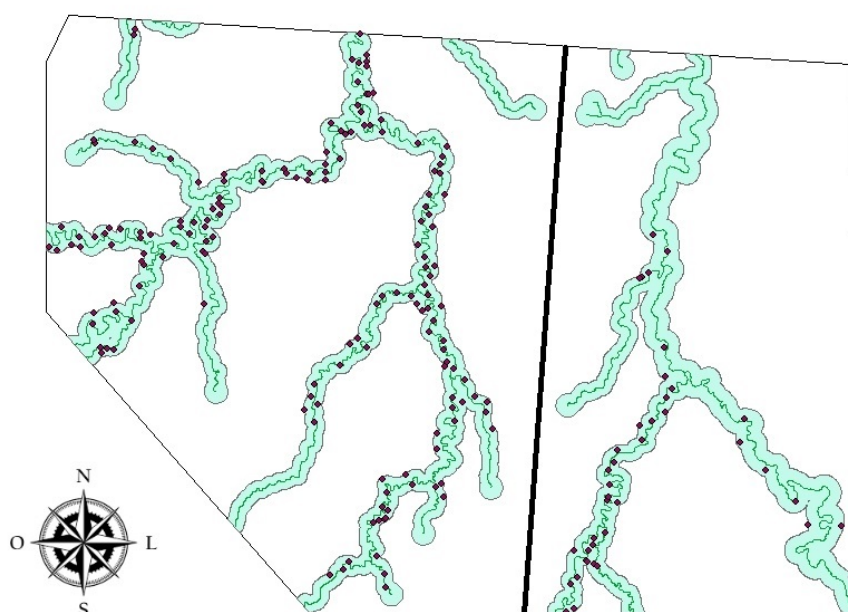


O inventário florestal procedeu por meio do georreferenciamento das árvores. No levantamento feito e publicado em Junior et al. (2014), a coleta dos dados iniciou-se com um inventário florestal realizado na área efetiva de manejo, onde foram mensuradas as 56 espécies de árvores comerciais a partir de um diâmetro mínimo de corte ($DMC > 50cm$), sendo registrado as informações de diâmetro a altura do peito (DAP), que é tomado a 1,30m de altura verticalmente medidos, utilizando uma fita métrica para tirar a medida da seção circular, trena laser para altura comercial, fazer o cálculo do volume e identificação botânica e classes de qualidade do fuste. Na sequência, a mesma foi marcada com uma placa de alumínio contendo informações relacionadas a seu número de registro. As árvores foram georreferenciadas em coordenadas UTM, utilizando um aparelho GPS. Posteriormente, as informações geradas foram convertidas em um banco de dados espaciais.

Sendo marcadas com placas de alumínio contendo suas devidas numerações, as árvores selecionadas para exploração foram estabelecidas conforme o plano de manejo florestal e a Instrução Normativa 05 de 11 de dezembro de 2006, resolução do CONAMA número 406 de 2 de fevereiro de 2009.

Neste Estudo foram observadas 201 árvores de *Hura crepitans* dentro da APP da área total tomada para estudo, como ilustrado na figura 3.2, que foi dividida por uma linha mais espessa, como mostrado :

Figura 3.2 – Área de estudo



Como não temos uma rede completamente conectada, podemos dividir em duas partes o banco de dados em oeste e leste, para assim prosseguir as análises, sendo assim com essa separação foi perdido duas observações.

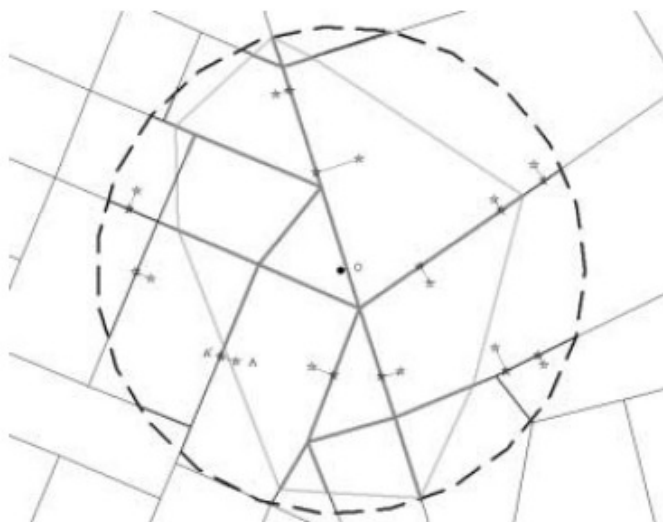
3.2 Métodos estatísticos

A análise descritiva dos dados e a estimação da intensidade global, mapa Kernel, e função K para ambas abordagens, tanto planar quanto em redes, foram feitas utilizando o pacote *spatstat* Baddeley, Turner et al. (2005) do *software* R (2017), versão 3.3.3.

Uma observação importante a ser comentada ao iniciar as análises é que será abordagem sugerida por Borruo (2008) que é de criar *mirror points* (pontos espelhados) na rede, por os eventos não ocorrem necessariamente sobre o traçado linear da rede e sim em suas imediação

logo para efeito de estimação os eventos serão trazidos para cima da rede como no exemplo da figura 3.3:

Figura 3.3 – Pontos ajustados para a rede



Fonte: Borruso (2008)

Dado que os eventos, no caso exposto, são a ocorrência das árvores da espécie escolhida, não é pontualmente feita dentro do rio, logo é necessária uma interpolação. Foi usado o pacote *mapprools* desenvolvido por Bivand (2016).

A distribuição empírica das estatísticas para esses testes sob a hipótese nula e o valor p serão obtidas a partir do procedimento de permutação de Monte Carlo considerando $s = 79$ repetições, rank $k = 1$ obtendo um nível de significância de 5%, como sugerida em (BADDELEY; RUBAK; TURNER, 2015).

Para o alisamento Kernel linear foi tomado uma distancia de $10m$ como sugerida em (BORRUSO, 2008), para exploração podendo ser mudado a pleno gosto ou informação a priori.

Para as análises de efeitos de primeira ordem via kernel planar será tomada distância igual a da $10m$, e para análise de seus efeitos de segunda ordem via função K de Ripley, com rank $k = 5$ e número de simulações $s = 99$ como sugerido por (BADDELEY; TURNER et al., 2005), obtendo se um nível de significância de 5%, no processo de permutação Monte Carlo.

A seguir temos exposta na figura 3.4 a parte leste da rede de rios e as respectivas observações contidas nos seus limites:

- Rede Leste:

Figura 3.4 – Rede leste



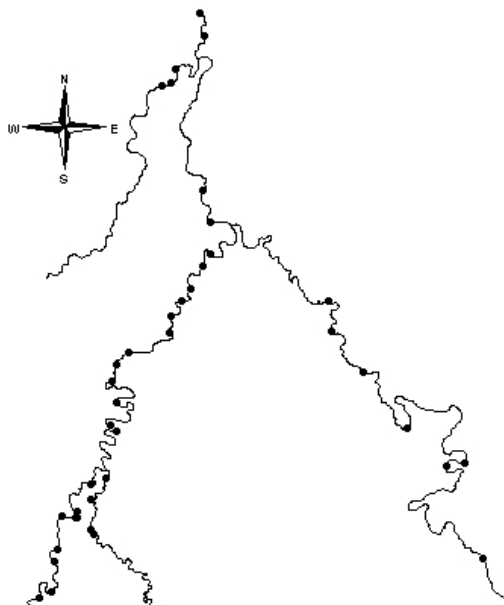
(a) Rede (b) Pontos

Como se apresenta na figura 3.4 temos em:

- (a) a rede linear dos rios
- (b) seus pontos associados

É de se salientar que em estudos desse tipo onde naturalmente a configuração da rede não é linear e os pontos não estão necessariamente posicionados sobre a rede, tem-se que linearizar a rede em toda sua extensão e interpolar os seus pontos de forma ortogonal para cima da rede. Na figura a seguir podemos notar como a rede onde o estudo será realizado se configura:

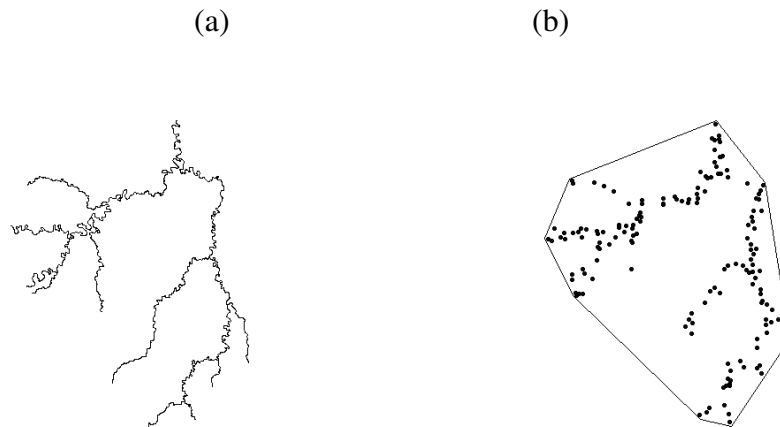
Figura 3.5 – Rede linear Leste com as ocorrências



Após configurarmos a rede com seus respectivos pontos utilizando o pacote *maptools*, estamos aptos a aplicar as ferramentas descritas anteriormente. Nesta rede temos um total de 40 observações em um comprimento de rio de 7.408m.

- Rede Oeste:

Figura 3.6 – Rede Oeste



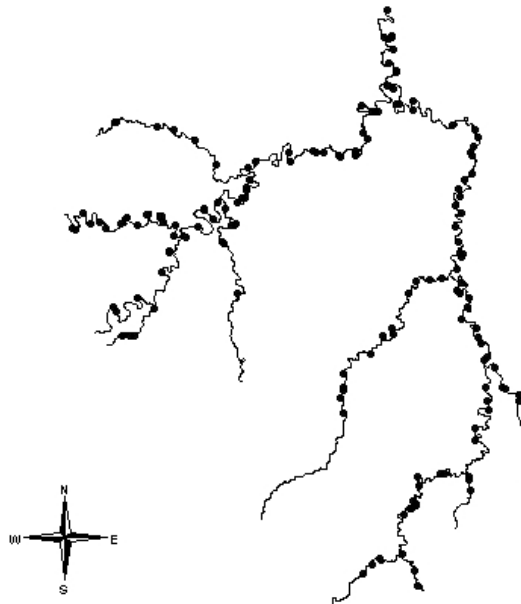
(a) Rede (b) Pontos

Como se apresenta na figura 3.6 temos em:

- (a) a rede linear dos rios
- (b) seus pontos associados

Na figura 3.7 podemos notar como a rede onde o estudo será realizado se configura:

Figura 3.7 – Rede linear Oeste com as ocorrências



Neste lado oeste da rede temos um total de 149 observações em um comprimento de linear de rio de 16,003km.

Podemos esperar por possuir um extensão grande linear ambas as redes, comparado com o número de eventos em cada uma, é esperado uma baixa intensidade global , também por se tratar de uma estudo em um ambiente de floresta natural. Com o toda a rede dividida e configurada podemos assim partir para a análise de Processos pontuais em rede e responder a pergunta se ocorre ou não agrupamento na disposição das árvores em estudo.

4 RESULTADOS E DISCUSSÃO

Os resultados obtidos a partir de uma partição feita no banco de dados original, pois como já observado anteriormente apresenta uma descontinuidade na sua rede de rios, impossibilitando metodologicamente nesse nível proposto a análise conjunta. Sendo assim será feita analogamente a aplicação das ferramentas e metodologia para ambas as partes oeste e leste do mapa, a análise de propriedades de efeitos de primeira ordem seguidos dos de segunda ordem.

4.1 Propriedades de primeira ordem

Iniciou-se a análise com o cálculo das estimativas de intensidade global para as duas regiões, utilizando os dois estimadores (planar e rede). Para a região leste, utilizando uma janela de $1.299km^2$, obteve-se uma intensidade planar, estimada de $\hat{\lambda} = 0,0003$ que pode ser interpretado como uma presença média de uma árvore a cada $32,20km^2$, aproximadamente, enquanto a intensidade em rede estimada foi de $\hat{\lambda} = 0,005$ que pode ser interpretado como uma presença média de uma árvore a cada $185m$, aproximadamente.

Para a região oeste, obteve-se uma intensidade planar estimada de $\hat{\lambda} = 0,00005$ que pode ser interpretado como uma presença média de uma árvore em cada $20,32km^2$, utilizando uma janela de $3.034km^2$, aproximadamente, enquanto a intensidade em rede estimada foi de $\hat{\lambda} = 0.009$ que pode ser interpretado como uma presença média de uma árvore a cada $1,064km$, aproximadamente.

Nota-se uma sub-estimação em ambos os lados quando usado o método planar, ou seja apresentando uma intensidade global $\hat{\lambda}$ menos que no caso do estimador para rede, uma aproximação mais refinada para o estimador planar, pode ser dada se disponível a janela nominal da área real da APP e não a automática utilizada, mesmo que essa melhora na estimação não seja garantida. Logo temos uma estimativa mais precisa tomando os resultados nesse caso do estimador linear de intensidade global.

A intensidade global $\hat{\lambda}$ trás uma boa noção descritiva da média do processo. Entretanto, pode-se obter uma estimativa mais refinada usando o alisamento Kernel que é uma espécie de estimativa de intensidade local.

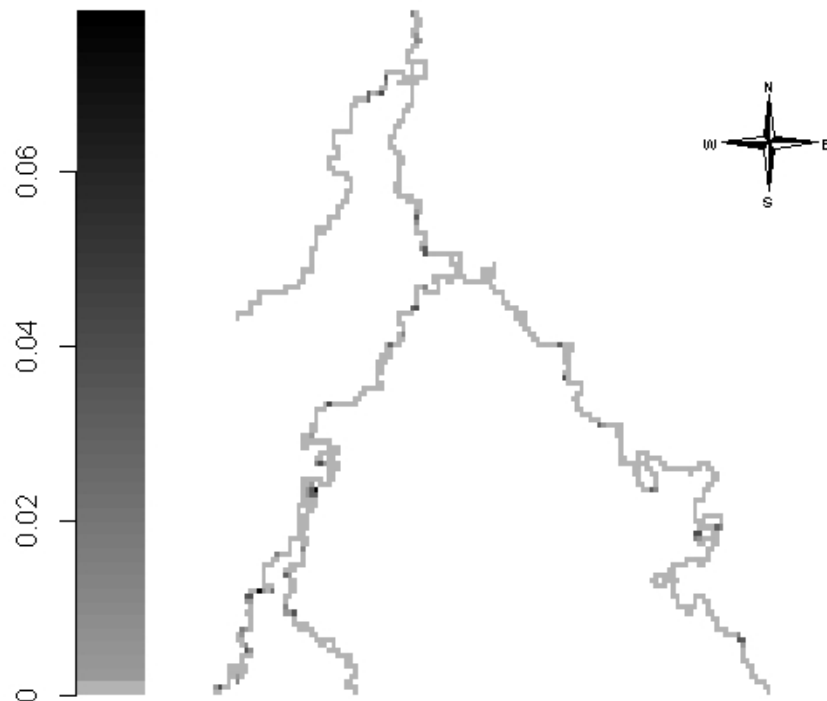
Alisamento Kernel

Um estimativa mais geral como a intensidade global $\hat{\lambda}$, trás uma boa noção descritiva do cenário que pode ser refinada usando o alisamento Kernel, uma espécie de intensidade local, onde escaneamos toda a rede a partir de pontos sorteados no *grid*, como descrito em Borruco (2008). Foi usado uma janela para escaneamento de 10m. Para melhor visualização serão apresentados os gráficos do alisamento de cada região separadamente de forma subsequente.

Como pode-se notar na figura 4.1 onde em estão destacados com tons mais fortes de cinza são regiões do rio com maior incidência via Kernel linear de árvores da espécie em questão:

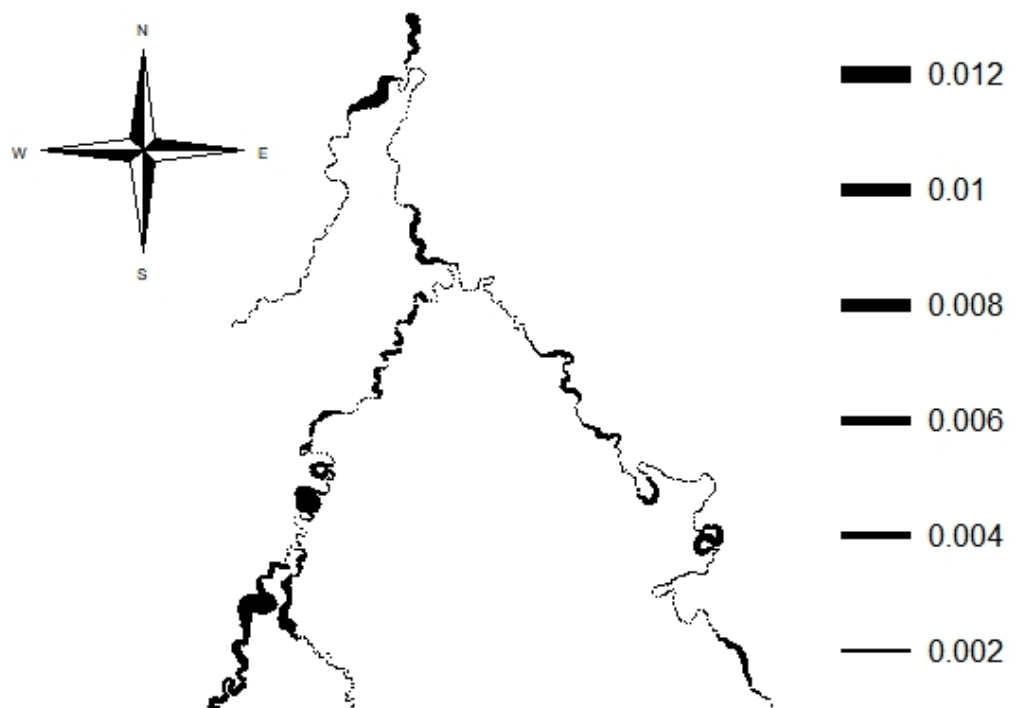
- Rede Leste:

Figura 4.1 – Alisamento Kernel Leste em tons de cinza



Na figura 4.2 podemos ver de maneira mais destacada a intensidade descrita pelo o alisamento kernel linear como proposta em (TIMOTHÉE et al., 2010) e usada, com esse formato gráfico, por Baddeley, Jammalamadaka e Nair (2014) para análise de dendritos por se posicionarem em poucos locais da rede neural cerebral, tornando a difícil visualização no gráfico kernel tradicional.

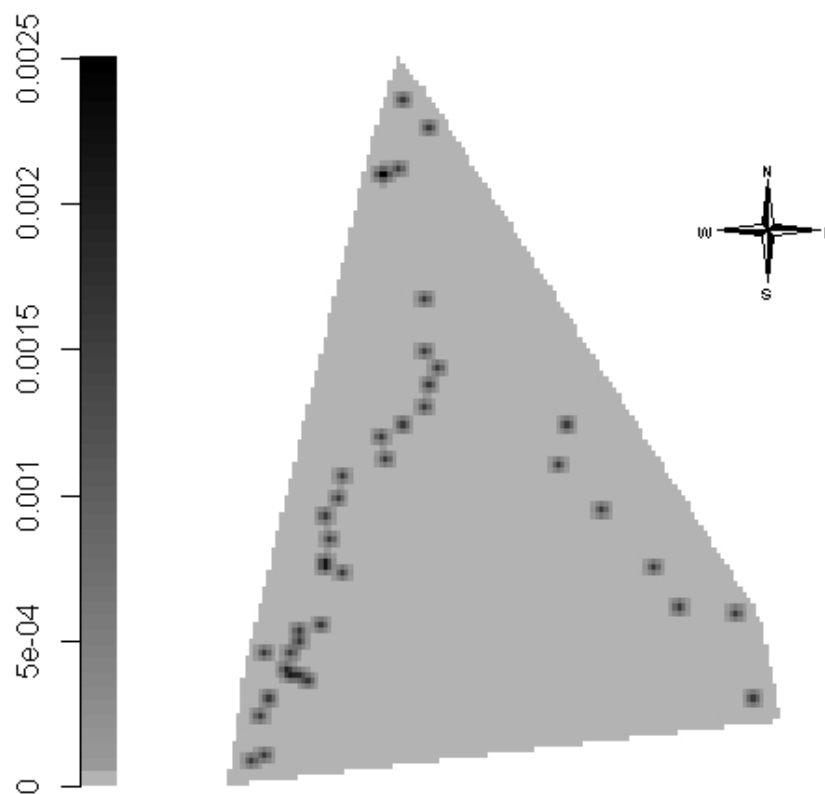
Figura 4.2 – Alisamento kernel Leste



Temos em cunho de comparação na figura 4.3 a aplicação do alisamento Kernel planar, nas mesmas condições de área que foi aplicada o cálculo de $\hat{\lambda}$ anteriormente, com é apresentado

em tons mais escuros de cinza também ocorre um menos valor quando da incidência quando comparado a abordagem em redes.

Figura 4.3 – Alisamento kernel planar Leste

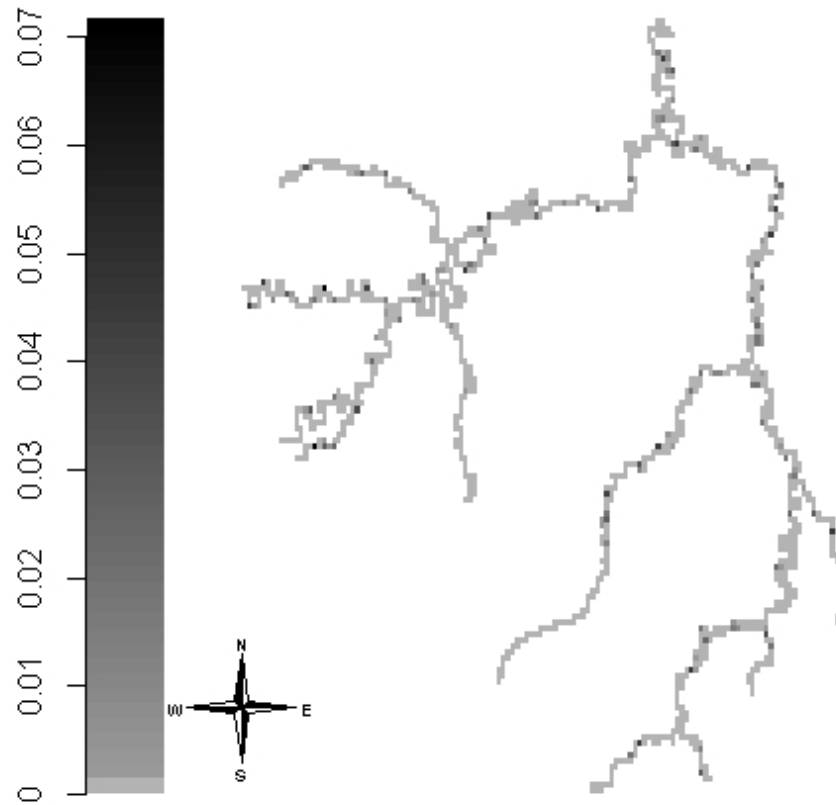


De forma geral é dito em Okabe, Satoh e Sugihara (2009) que o uso da adaptação da kernel linear é necessário para uma melhor suavização das intensidades das ocorrências principalmente próximas aos nós (entrocamento dos rios).

- Rede Oeste:

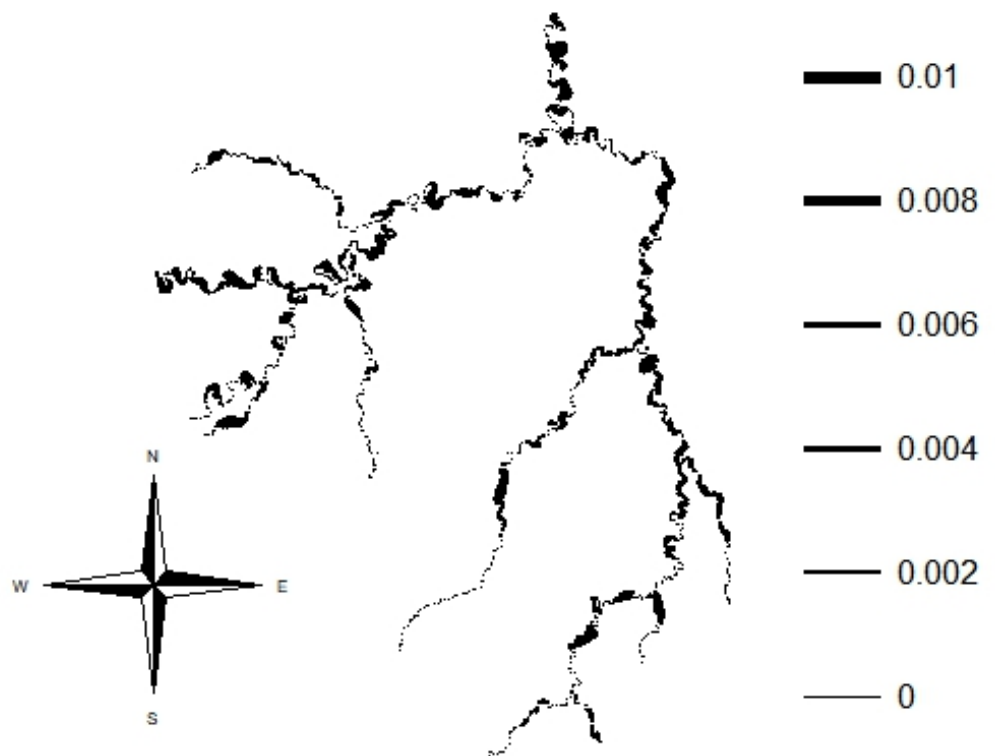
Como observado que na rede leste, é apresentado na 4.4, onde podemos de forma análoga interpretar, os tons mais escuros de cinza como regiões da rede com maior ocorrência das árvores em estudo.

Figura 4.4 – Alisamento kernel linear Oeste em tons de cinza



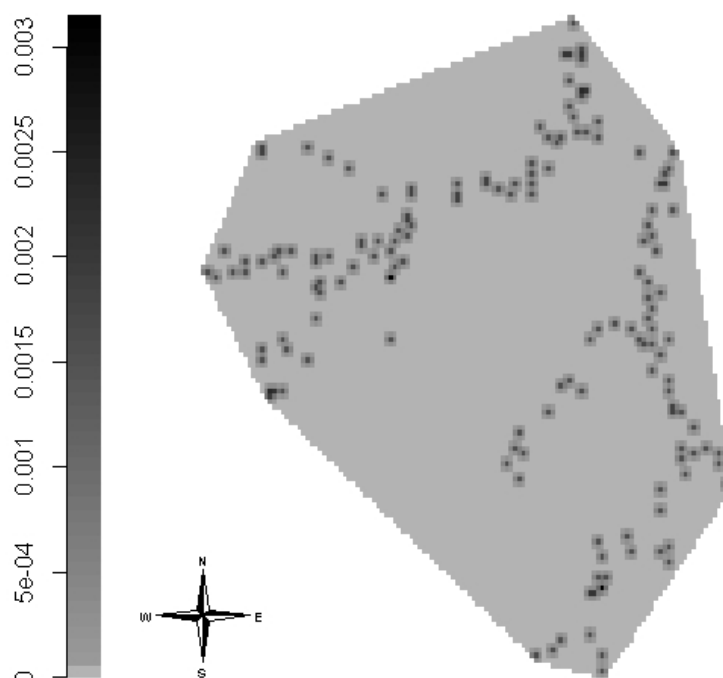
Para melhor visualização temos a figura 4.5, que nos mostra uma maior intensidade de ocorrência da espécie em estudo onde as linhas são mais grossas como mostrado anteriormente para o lado leste, onde detectamos regiões em que a grossura da linha é mais espessa exatamente a maior intensidade, até mesmo ligeiramente maior que no lado leste .

Figura 4.5 – Alisamento kernel linear Oeste



A figura 4.6, apresenta a aplicação do lado oeste da alisamento kernel planar, onde podemos notar que em tons mais escuros de cinza a intensidade máxima não chega a metade do valor obtido utilizando a abordagem adaptada para redes, corroborando com a subestimação apresentada na $\hat{\lambda}$.

Figura 4.6 – Alisamento kernel Oeste



A partir das análises descritas devemos tomar nota que a abordagem proposta para redes é a mais indicada por não apresentar subestimação quando comparada com o alisamento kernel planar. Lembrando que o alisamento Kernel linear, apesar de usar estimadores não viesados em sua construção, é uma ferramenta inteiramente descritiva, mas que nos guia para uma averiguação da presença ou não de agrupamento via função K-linear.

4.2 Propriedades de segunda ordem

O uso da função K-linear é exclusivamente necessário para responder a pergunta de presença ou não de agrupamento, mas além disso, podemos a partir dela também inferir se a disposição dos eventos em estudo na rede tem um caráter aleatório ou regularidade. Essa abordagem também faz-se necessário assumir a estacionariedade de primeira ordem.

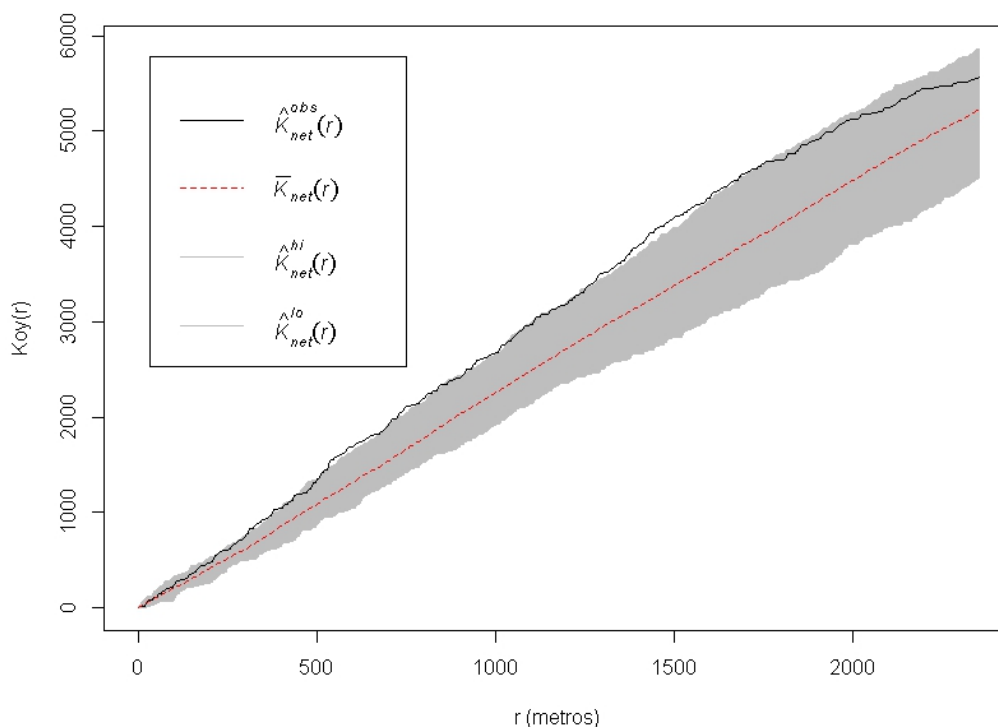
Aplicação da função K-linear

Sendo usado em ambos os lados da rede em estudo a Função K-linear temos os seguintes resultados:

- Rede Leste:

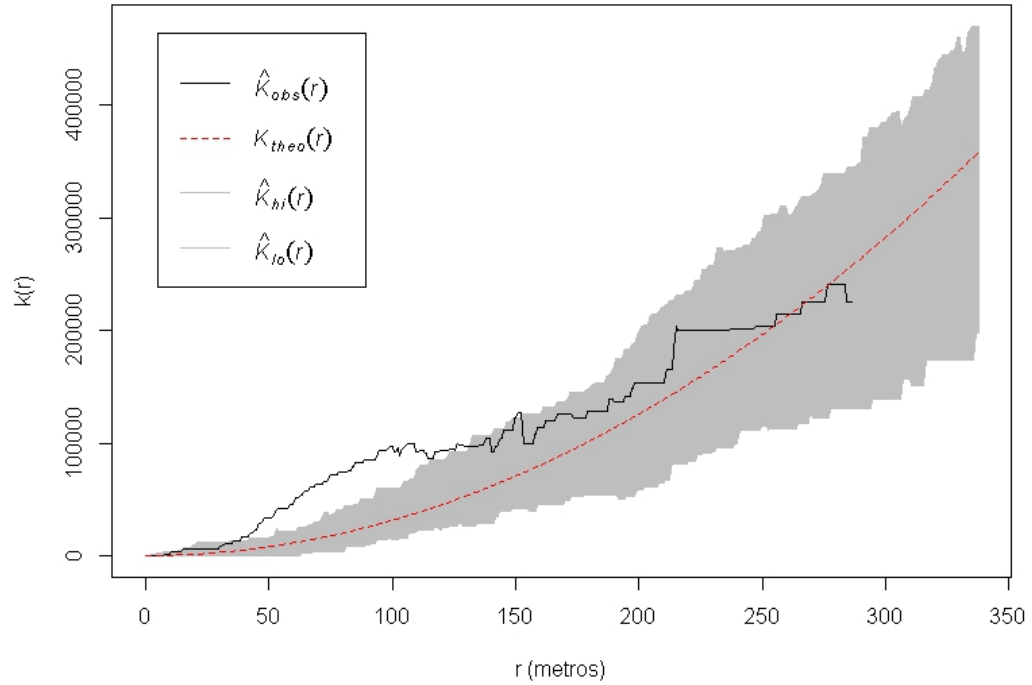
Temos pela figura 4.7 da função K-linear feito o envelope de confiança via simulação de Monte-Carlo de 79 iterações ao nível de significância $\alpha = 0,05$, em que, até a amplitude de aproximadamente 500m a rede apresenta um comportamento aleatória na disposição da árvores, dado que a curva esta alocada dentro do envelope, dos 500m aos 2000m de amplitude aproximadamente, sendo que a curva está acima da reta (que determina a hipótese nula) e fora do envelope, temos um caráter de agrupamento, o que, segundo Baddeley, Turner et al. (2005) é natural por se acumular um número grande de ocorrências dos eventos em estudo.

Figura 4.7 – Função K-linear Leste



Na figura 4.8 apresentada a seguir, nos mostra a justificativa dada em Baddeley, Rubak e Turner (2015) para o uso da K-linear de Okabe e Yamada (2001), e não a tradicional função K de (RIPLEY, 1976), por haver superestimação, onde antes dos 50m já é observado comportamento de agrupamento.

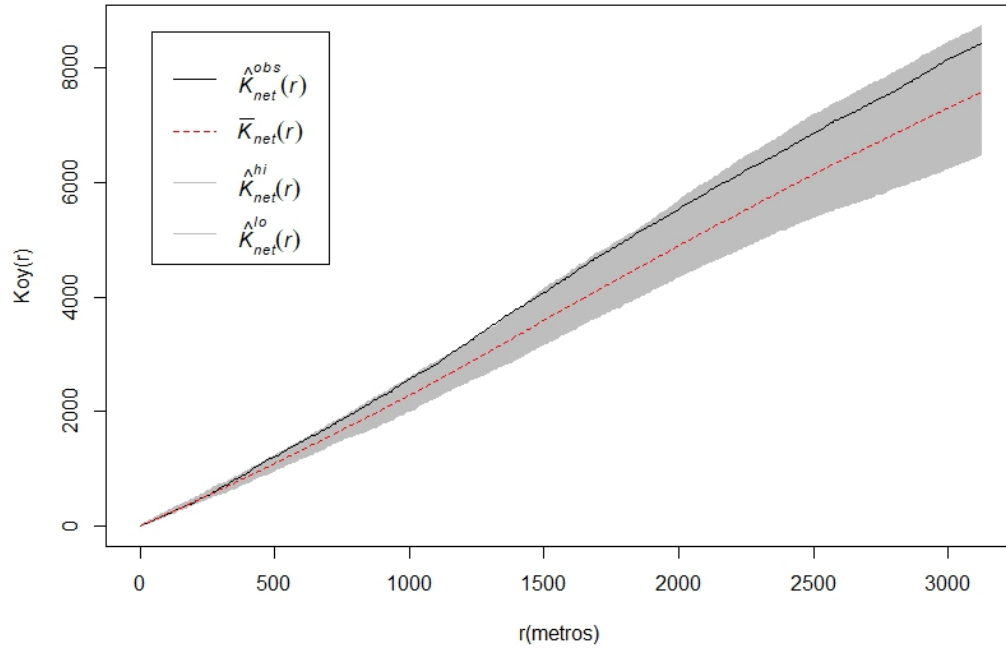
Figura 4.8 – Função K planar Leste



- Rede Oeste:

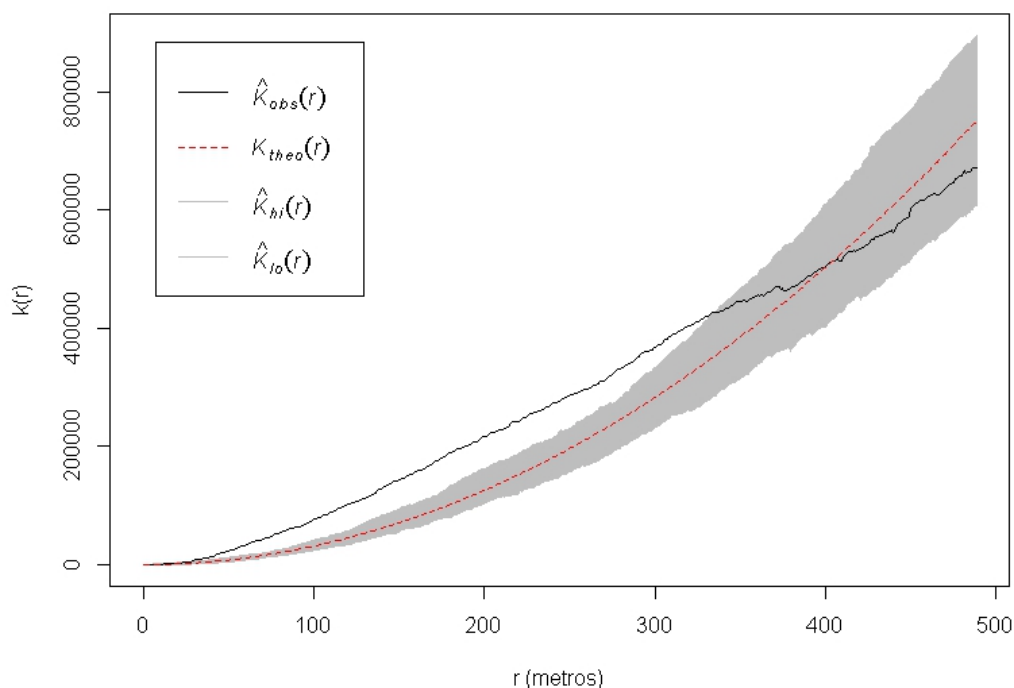
Podemos notar para o lado oeste na figura 4.9 da função K-linear, que o padrão de agrupamento não é observado, levando em consideração os envelopes construídos, ou seja, para lado oeste da rede as árvores de *Hura Crepitans* tem um comportamento de completa aleatoriedade.

Figura 4.9 – Função K-linear Oeste



Temos na figura 4.10 que assim com para o lado leste, o oeste quando usada a função K planar, apresenta um caráter de agrupamento a partir de uma distância próxima da inicial.

Figura 4.10 – Função K planar Oeste



Utilizando as ferramentas de análise de efeitos de primeira e segunda ordem, para abordagem de processos pontuais em redes lineares, não rejeitamos em ambos os lados a hipótese de completa aleatoriedade.

Podemos observar em trabalhos que usaram em suas análises estruturas de redes lineares, as quais, se assemelham a redes neurais cerebrais, tendo o caráter de agrupamento é pouco observado como apresentado em Jammalamadaka et al. (2013) e Baddeley, Jammalamadaka e Nair (2014), onde se trata de fato com redes neurais cerebrais em forma de redes lineares, em ambos os casos, assim como no tratado nesse trabalho, quando aumentada a distância temos a completa aleatoriedade. Observando por um olhar técnico dentro da área florestal é descrito por Martinez, Mourão e Junior (2011) em seu estudo afirma que, por suas propriedades morfofisiológicas, a *Hura Crepitans* tem como características não se agruparem.

5 CONCLUSÃO

O objetivo deste trabalho foi apresentar o uso básico da abordagem de processos pontuais em redes e expor de forma clara sua metodologia e ferramentas computacionais, este trabalho mostrou a partir desse princípio, uma aplicação ambiental sustentável, em uma rede formada por rios. Onde os eventos de interesse foram as árvores de *Hura crepitans* situadas as margens em sua APP, definidos os eventos de estudo e suas particularidades foi explicitado com clareza o mapa descritivo de alisamento kernel e o mais importante via função k-linear foi verificado que em ambas as redes o processo tem um caráter de completa aleatoriedade. Podemos assim verificar que a abordagem é sim eficaz e para o reconhecimento de padrões no ambiente de redes lineares. Nesta pesquisa uma dificuldade enfrentada foi o extensivo tempo computacional necessário para a execução das análises, devido ao alto número de combinações possíveis em cada análise e o tratamento das redes para análise. Para futuros trabalhos pode-se tentar explorar a distribuição dos dados de outros modos, como por exemplo, estudar como a utilização de covariáveis e marcas quantitativas.

REFERÊNCIAS

- ACRE, A. Governo do Estado do. **Zoneamento Ecologico-Economico do Acre, Fase II: documento sintese-escala 1: 250.000.** : Sema Rio Branco, 2006.
- BADDELEY, A.; BÁRÁNY, I.; SCHNEIDER, R. Spatial point processes and their applications. **Stochastic Geometry: Lectures given at the CIME Summer School held in Martina Franca, Italy, September 13–18, 2004**, Springer, p. 1–75, 2007.
- BADDELEY, A.; JAMMALAMADAKA, A.; NAIR, G. Multitype point process analysis of spines on the dendrite network of a neuron. **Journal of the Royal Statistical Society: Series C (Applied Statistics)**, Wiley Online Library, v. 63, n. 5, p. 673–694, 2014.
- BADDELEY, A.; RUBAK, E.; TURNER, R. **Spatial point patterns: methodology and applications with R.** [S.l.]: CRC Press, 2015.
- BADDELEY, A.; TURNER, R. et al. Spatstat: an r package for analyzing spatial point patterns. **Journal of statistical software**, v. 12, n. 6, p. 1–42, 2005.
- BAILEY, T. C.; GATRELL, A. C. **Interactive spatial data analysis.** [S.l.]: Longman Scientific & Technical Essex, 1995. v. 413.
- BIVAND, R. **Lewin-Koh N. maptools: Tools for Reading and Handling Spatial Objects. R package version 0.8–30. 2014.** 2016.
- BORRUSO, G. Network density estimation: a gis approach for analysing point patterns in a network space. **Transactions in GIS**, Wiley Online Library, v. 12, n. 3, p. 377–402, 2008.
- CÂMARA, G.; CARVALHO, M. S. Análise espacial de eventos. **Análise espacial de dados geográficos. Brasília: Embrapa**, 2004.
- CÂMARA, G.; MONTEIRO, A. M. V.; MEDEIROS, J. d. Representações computacionais do espaço: fundamentos epistemológicos da ciência da geoinformação. **Geografia, Rio Claro**, v. 28, n. 1, p. 83–96, 2003.
- CAPRETZ, R. L. et al. Padrão espacial de quatro formações florestais do estado de são paulo, através da função k de ripley. **Ciência Florestal**, Universidade Federal de Santa Maria, v. 22, n. 3, 2012.
- CHESSON, P. General theory of competitive coexistence in spatially-varying environments. **Theoretical population biology**, Elsevier, v. 58, n. 3, p. 211–237, 2000.
- CRESSIE, N. A. **Statistics for spatial data.** [S.l.]: Wiley Online Library, 1993. 900 p.
- DIGGLE, P. A kernel method for smoothing point process data. **Applied statistics**, JSTOR, p. 138–147, 1985.
- DIGGLE, P. **Statistical analysis of spatial point patterns. Second Edition.**: London: Arnold, 2003. 160 p.
- DIGGLE, P. J. **Statistical analysis of spatial and spatio-temporal point patterns.**: CRC Press, 2013. 268 p.
- FILIZOLA, A. Caderno regional–região hidrográfica amazônica. **Ministério do Meio Ambiente, Secretaria de Recursos. Hídricos**, Brasília, v. 2, 2006.

- FUHRMANN, P. A.; HELMKE, U. **The mathematics of networks of linear systems**. New York: Springer, 2015. 680 p.
- GONÇALVES, F. G.; SANTOS, J. d. Composição florística e estrutura de uma unidade de manejo florestal sustentável na floresta nacional do tapajós, Pará. **Acta Amazonica**, SciELO Brasil, v. 38, n. 2, p. 229–244, 2008.
- IWAKIRI, S. et al. Avaliação da qualidade de painéis compensados produzidos com lâminas de madeira de schizolobium amazonicum. **Floresta**, v. 41, n. 3, 2011.
- JAMMALAMADAKA, A. et al. Statistical analysis of dendritic spine distributions in rat hippocampal cultures. **BMC bioinformatics**, BioMed Central, v. 14, n. 1, p. 287, 2013.
- JUNIOR, M. et al. Alocação de pátios de armazenamento de madeira em um plano de manejo florestal na amazônia ocidental. **Anais do XLVI SBPO-Simpósio Bras. Pesqui. Oper**, v. 46, p. 704–713, 2014.
- LEOPOLDO, P. R. et al. Estimativa de evapotranspiração de floresta amazônica de terra firme. **Acta Amazonica**, SciELO Brasil, v. 12, p. 23–28, 1982.
- MARTINELLI, G.; MORAES, M. A. **Livro vermelho da flora do Brasil**.: CNCFlora, Centro Nacional de Conservação da Flora, Rio de Janeiro, 2013. 1089 p.
- MARTINEZ, G. B.; MOURÃO, M.; JUNIOR, S. B. Respostas morfofisiológicas de plantas de açacu (hura crepitans L.) provenientes de várzeas do rio Amazonas: efeito da anoxia do solo. **Revista Árvore**, Universidade Federal de Viçosa, v. 35, n. 6, 2011.
- OKABE, A.; SATOH, T.; SUGIHARA, K. A kernel density estimation method for networks, its computational method and a GIS-based tool. **International Journal of Geographical Information Science**, Taylor & Francis, v. 23, n. 1, p. 7–32, 2009.
- OKABE, A.; YAMADA, I. The k-function method on a network and its computational implementation. **Geographical Analysis**, Wiley Online Library, v. 33, n. 3, p. 271–290, 2001.
- PAROLIN, P. O uso de árvores nas florestas inundadas por água branca na amazônia central. **Amazoniana**, Amazonas, v. 16, n. 1/2, p. 241–248, 2000.
- RIPLEY, B. D. The second-order analysis of stationary point processes. **Journal of applied probability**, Cambridge University Press, v. 13, n. 2, p. 255–266, 1976.
- RIPLEY, B. D. Modelling spatial patterns. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, London, p. 172–212, 1977.
- SCALON, J. D.; SILVA, F. M. Power of tests for spatial randomness in patterns with small number of events. **Revista de Ciência e Tecnologia**, Rio de Janeiro, v. 12, n. 2, p. 7–14, 2006.
- SCHABENBERGER, O.; GOTWAY, C. **Statistical methods for spatial data analysis**. London: Chapman & Hall/CRC, 2005. 488 p.
- TIMOTHÉE, P. et al. A network based kernel density estimator applied to barcelona economic activities. p. 32–45, 2010.

APÊNDICE A – Script

O script comentado usado para a análise segue:

```
# Carregando os pacotes necessários

library("spatstat")
library("mapproj")
library("sp")

# Selecionando o diretório onde estão os dados e a rede
setwd("C:/Users/seu diretório")

# Entrando com os dados, ou seja os pontos
dados01<-read.table("teste.txt",sep = ",", dec = ".", header=TRUE)

#janela automática, sendo as coordenadas em UTM
w2 = ripras(dados01$EASTING,dados01$NORTHING)
head(dados01)

# Convertendo para o formato ppp os seus pontos
dadoslppp = ppp(x = dados01$EASTING, y = dados01$NORTHING,
marks = NULL, window=w2)

#Plotando os pontos observados
plot(dadoslppp, main = "Configuração pontual \n Homogênea")

# Função que constroi a rede linear a partir de um shape
if(require(spatstat, quietly=TRUE)) {
dname <- system.file("shapes", package="mapproj")
fname <- file.path(dname, "New_Shapefile.shp")
fylk <- readShapeSpatial(fname)
L <- as(fylk, "linnet")
```

```
print(max(vertexdegree(L)))
L0 <- as.linnet.SpatialLines(fylk, fuse=FALSE)
print(max(vertexdegree(L0)))
}

print(L)#informações da rede

#Plotando a rede sem os pontos
plot(L)

# Fazendo a interpolação dos pontos para a rede lienar
X=lpp(dadoslppp,L)
#Plotando a rede com os pontos(objeto que será feito as análises)
plot(X)

#Estimador global de intensidade
intensity.lpp(X)

# Alisamento Kernel linear
DSD <- density(X, sigma=10)
plot(DSD, main="Alisamento Kernel")

#Função K-linear
klin=envelope(X, linearK, correction="ang", nsim=39)
plot(klin)
```