



CAROLINA CRISTINA BICALHO

**MÉTODOS NÃO PARAMÉTRICOS EM PROCESSOS
PONTUAIS REPLICADOS MARCADOS**

LAVRAS – MG

2018

CAROLINA CRISTINA BICALHO

**MÉTODOS NÃO PARAMÉTRICOS EM PROCESSOS PONTUAIS REPLICADOS
MARCADOS**

Tese 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 Doutor.

Prof. Dr. 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).**

Bicalho, Carolina Cristina.

Métodos não paramétricos em processos pontuais replicados
marcados / Carolina Cristina Bicalho. - 2018.

95 p. : il.

Orientador(a): João Domingos Scalon.

.

Tese (doutorado) - Universidade Federal de Lavras, 2018.

Bibliografia.

1. Estatística espacial. 2. Panificação. 3. Análise de variância. I.
Scalon, João Domingos. . II. Título.

CAROLINA CRISTINA BICALHO

**MÉTODOS NÃO PARAMÉTRICOS EM PROCESSOS PONTUAIS REPLICADOS
MARCADOS**

Tese 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 Doutor.

APROVADA em 16 de abril de 2018.

Dra. Joelma Pereira	DCA-UFLA
Dr. Paulo Henrique Sales Guimarães	DES-UFLA
Dr. Marcelo Silva de Oliveira	DES-UFLA
Dr. Ricardo Alves de Olinda	CCT-UEPB



Prof. Dr. João Domingos Scalon
Orientador

**LAVRAS – MG
2018**

*A toda minha família.
Aos meus pais Antônio e Arlete, meu porto seguro,
minhas irmãs, sobrinhos e meu namorado Elias.
Vocês foram a força que me impulsionou,
me fazendo acreditar que a realização do meu sonho era possível.
A vocês pertence grande parte desta vitória.
Grande parte desta vitória é de vocês.*

DEDICO

AGRADECIMENTOS

A Deus que sempre esteve presente na minha vida, me iluminando e dando-me forças e coragem para realizar este sonho de cursar este doutorado.

Ao meu orientador, o querido professor Dr. João Domingos Scalon pelas suas orientações, por compartilhar seu conhecimento e pela sua paciência frente às minhas incertezas, e ainda, por ter acreditado muitas vezes mais do que eu, na realização desse trabalho.

Aos professores do Departamento de Estatística da UFLA pela dedicação ao ensino com afinho e carinho, incentivando a busca pelo conhecimento.

Aos funcionários do Departamento de Estatística da UFLA, em especial a Nádia, por sempre estar disposta a ajudar e descontraír com conversas boas durante esse período difícil.

Ao Raul pela ajuda e incentivo durante todo esse período.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de doutorado.

Aos meus pais, Antônio e Arlete pelo amor e apoio, aqueles que são os maiores mestres para mim, pois me ensinaram os valores para a formação do meu caráter.

Às minhas irmãs, Camila e Karine, pelo incentivo, carinho e amizade.

Ao meu namorado e melhor amigo Elias, pela compreensão e apoio incondicional estando sempre ao meu lado colaborando de maneira tão singela e amorosa para que eu possa realizar o meu sonho.

Aos meus sobrinhos, Beatriz, Mariana e Gabriel, por trazerem tanta alegria para minha vida, principalmente nos momentos mais complicados.

Aos amigos de Bambuí e Lavras, que fazem parte da minha história há algum tempo ou aqueles que encontrei durante esta jornada, pelos momentos de aprendizado juntos.

Em especial as minhas amigas Jackélya, Hortência e Rosana por estarem sempre por perto.

E a todos que sempre torceram por mim, desejando-me sucesso e felicidade.

MUITO OBRIGADA!

*A maior recompensa pelo nosso trabalho não é o que nos
pagam por ele, mas aquilo em que ele nos transforma*
(John Ruskin)

RESUMO GERAL

Em diversas áreas do conhecimento é de grande importância analisar a estrutura de dependência espacial de uma variável aleatória. Diante disso, o objetivo dessa tese é apresentar métodos estatísticos para analisar a estrutura de dependência espacial em dados de processos pontuais marcados e não marcados. Neste sentido, o trabalho apresenta diversos métodos de processos pontuais no plano, que estão consolidados na literatura e propõe um novo método para testar a hipótese de igualdade entre grupos de processos pontuais replicados com marcas quantitativas. O método proposto é avaliado realizando as análises do erro do tipo I. Todos os métodos apresentados são aplicados para caracterizar a distribuição espacial de alvéolos, localizados na massa congelada, sob diferentes temperaturas do pão tipo francês. Os resultados obtidos permitem afirmar que os métodos de processos pontuais replicados são ferramentas indispensáveis para caracterizar a dependência espacial de eventos localizados em uma região.

Palavras-chave: Função K de Ripley. Replicação. Teste de permutação. Simulação. Panificação.

ABSTRACT

In several areas of knowledge, it is of great significance, to analyze the spatial dependency structure of a random variable. Therefore, the aim of this thesis is to present statistical methods to analyze the structure of spatial dependence in data of marked and unmarked point patterns. By this way, the work presents several methods of point patterns in the plan; those are consolidated in the literature and propose a new method to test the equality hypothesis between groups of marked (quantitative) replicated point process. The suggested method is evaluated by performing type I error analysis. All presented methods are applied to characterize the spatial distribution of alveoli located in the frozen dough under different temperatures of French type bread. The obtained results allow us to affirm that the methods of replicated point patterns indispensable tools characterize the spatial dependence of events located in a region.

Keywords: Ripley's K function. Replication. Permutation test. Simulation. Baking.

SUMÁRIO

	PRIMEIRA PARTE	10
1	INTRODUÇÃO	10
2	REFERENCIAL TEÓRICO	12
2.1	Processos pontuais espaciais	12
2.1.1	Propriedade de primeira ordem	13
2.1.2	Propriedade de segunda ordem	15
2.1.2.1	A Função F para processos pontuais homogêneos	17
2.1.2.2	A Função G para processos pontuais homogêneos	19
2.1.2.3	A função K para processos pontuais homogêneos	20
2.1.3	Testes contra hipótese de Completa Aleatoriedade Espacial	24
2.2	Processos pontuais replicados não marcados	26
2.2.1	Propriedade de primeira ordem para dados replicados	27
2.2.2	Propriedade de segunda ordem para dados replicados	27
2.2.3	Teste para comparação de grupos em padrão de pontos	28
2.3	Processos pontuais marcados estacionários	29
2.3.1	Propriedade de segunda ordem processos pontuais marcados	31
2.4	Processos pontuais replicados com marcas quantitativas	32
2.5	Panificação	33
	REFERÊNCIAS	35
	SEGUNDA PARTE- ARTIGOS	38
	ARTIGO 1 - ALVEOLAR DISTRIBUTION OF FRENCH TYPE BREAD USING TARO MUCILAGE	38
	ARTIGO 2 - O EFEITO DO CONGELAMENTO NA DISTRIBUIÇÃO ESPACIAL DOS ALVÉOLOS NA MASSA DO PÃO: UMA ANÁLISE BASEADA EM PADRÕES DE PONTOS ESPACIAL REPLICADOS	61
1	INTRODUÇÃO	63
2	MATERIAL E MÉTODOS	64
3	RESULTADOS E DISCUSSÃO	68
4	CONCLUSÃO	74
	REFERÊNCIAS	74

	ARTIGO 3 - ANÁLISE DE VARIÂNCIA NÃO PARAMÉTRICA DE PADRÕES DE PONTOS ESPACIAIS REPLICADOS COM MARCAS QUANTITATIVAS	77
1	INTRODUÇÃO	79
2	MATERIAL E MÉTODOS	82
3	RESULTADOS E DISCUSSÃO	87
4	CONCLUSÃO	92
	REFERÊNCIAS	93

PRIMEIRA PARTE

1 INTRODUÇÃO

O pão é um dos alimentos mais consumidos e uma das principais fontes calóricas da dieta da população de muitos países, inclusive no Brasil. Por esse motivo, o pão sempre foi alvo de muitos estudos nutricionais e também de manufatura, pois existe a necessidade de diminuir os custos operacionais na tentativa de expandir o mercado da panificação.

Atualmente, observa-se o desenvolvimento de novos métodos e tecnologias para produção de massas, como o congelamento e a preservação, com o objetivo de retardar alterações na qualidade dos alimentos.

Para obter uma massa de qualidade, várias características devem ser observadas durante o processo de produção. Por exemplo, espera-se que um pão de alta qualidade apresente uma massa com porosidade adequada e células ovaladas (alvéolos) de volumes uniformes. Assim, a presença de grandes alvéolos e ou a ausência deles na massa não são bons indicadores para um pão de qualidade. Nesse sentido, vários estudos estão sendo realizados para avaliar a distribuição dos alvéolos na massa e o tempo de armazenamento da massa congelada. Os dados gerados por esses estudos são conduzidos utilizando diversos métodos estatísticos, tais como gráficos, medidas descritivas, testes de hipóteses, análises de regressão, dentre outros.

Esta tese defende a ideia que ao incluir a informação da localização dos alvéolos na análise estatística dos dados, será possível trazer contribuições importantes para um melhor entendimento sobre a microestrutura e, conseqüentemente, para a qualidade do pão. Essa abordagem exige a utilização de métodos de uma área da estatística conhecida como estatística espacial. Como o principal interesse da análise será a localização e o tamanho dos alvéolos do miolo do pão francês, mais especificamente, serão utilizados métodos de processos pontuais.

Dados espaciais, em geral, não formam um conjunto de dados independentes. Espera-se que dados que estão mais próximos no espaço sejam mais parecidos que aqueles que estão mais distantes e, portanto, deve-se analisar esta correlação. A motivação para o desenvolvimento do presente trabalho consiste tanto no fato de que, até o momento, os métodos de processos pontuais mencionados anteriormente nunca foram utilizados em dados de panificação, quanto na possibilidade de desenvolver novos métodos para análise de configurações pontuais replicadas.

A presente tese é organizada em duas partes: na primeira parte é apresentado o referencial teórico, explorando aspectos relevantes para o entendimento da panificação, estatística espacial, processos pontuais replicados e processos pontuais marcados.

A segunda parte trata dos resultados em formato de três artigos. Inicialmente no Artigo 1 considerou-se na análise estatística apenas a coordenada de cada alvéolo localizado em fatias de pão francês, com objetivo de testar a hipótese de que os alvéolos apresentam uma distribuição completamente aleatória na massa do pão. Os métodos de processos pontuais para esse tipo de análise são bem desenvolvidos na literatura de estatística espacial.

Uma característica interessante dos dados de panificação é que, com o avanço da tecnologia é relativamente fácil obter não somente a localização dos alvéolos, mas também os seus tamanhos. Assim, esses dados podem ser obtidos repetidas vezes em muitas fatias de pão, sob diversas condições experimentais tais como: tempo de cozimento, temperatura do forno, tempo de congelamento etc. Essa replicação de dados possibilita explorar e desenvolver novos métodos dentro da área de processos pontuais, que é denominada de processos pontuais replicados, os quais podem ser não marcados (não considera o tamanho do alvéolo na análise) e marcados (considera o tamanho do alvéolo na análise). Esses métodos podem ser paramétricos e não-paramétricos.

Métodos de processos pontuais replicados não marcados paramétricos e não-paramétricos estão sendo desenvolvidos e começam a ser utilizados em diversas áreas do conhecimento. Assim, no Artigo 2, os métodos não paramétricos de processos pontuais replicados não marcados, foram aplicados para analisar a distribuição espacial de alvéolos em amostras de pães que foram submetidas a diferentes tempos de congelamento. O objetivo é testar a hipótese de que o tempo de congelamento não altera a distribuição espacial dos alvéolos na massa de pão.

Existem alguns métodos propostos de processos pontuais replicados com marcas categóricas. Entretanto, esses métodos ainda não estão bem desenvolvidos para marcas quantitativas. Por último, no artigo 3 é proposto um novo método não paramétrico para análise de dados de processos pontuais replicados com marcas quantitativas. O novo método é aplicado para testar diferenças e semelhanças nas variações da distribuição alveolar em diferentes tempos de congelamento.

2 REFERENCIAL TEÓRICO

2.1 Processos pontuais espaciais

A estatística espacial é uma área da estatística que, segundo Assunção (2001), estuda as metodologias de coleta, descrição, visualização e análise de dados associados a posições e que podem ser modelados como um processo estocástico.

Um processo estocástico, segundo Schabenberger e Gotway (2004), é uma coleção de variáveis aleatórias que são indexadas por algum conjunto $D \subset \mathbb{R}^d$ contendo coordenadas espaciais, ou seja,

$$\{Z(\mathbf{s}) : \mathbf{s} \in D \subset \mathbb{R}^d\}, \quad (2.1)$$

em que D é um conjunto de índices e $Z(\mathbf{s})$ é o atributo de interesse na localização \mathbf{s} , em que $\mathbf{s} = [s_1, s_2, \dots, s_d]'$.

Neste trabalho, considera-se $d = 2$, isto é, um processo no plano em que as coordenadas de longitude e latitude são definidas como $\mathbf{s} = [x, y]'$. A natureza do conjunto D permite a definição de três principais tipos de dados espaciais, no que diz respeito à natureza estocástica da observação (CRESSIE, 1993):

Dados geoestatísticos: Seja $Z(\mathbf{s})$ uma variável aleatória observada nas localizações $\mathbf{s} \in D$, em que D é fixo e contínuo. Exemplos: medições do volume de chuva em estações meteorológicas em uma determinada região de estudo, medições do nível de um poluente atmosférico em pontos de uma cidade.

Dados de área: Seja $Z(\mathbf{s})$ uma variável aleatória observada nas áreas $A(\mathbf{s})$, centradas nas localizações $\mathbf{s} \in D$, em que D é fixo e discreto. Exemplos: número de casos de uma doença por município de um estado, número de furtos de veículos por bairro de uma cidade.

Dados de pontos: Seja $Z(\mathbf{s})$ uma variável aleatória observada nas localizações $\mathbf{s} \in D$, em que D é um conjunto aleatório de índices. Exemplos: as localizações de focos de incêndio em uma floresta, localização de alvéolos em pão.

Em geral, para cada tipo de dado existem metodologias específicas para a análise espacial. Os métodos estatísticos utilizados neste trabalho foram desenvolvidos, exclusivamente, para a análise de dados de pontos.

Dados de pontos, são descritos formalmente, como uma coleção de eventos aleatórios $(s_i) : i = 1, 2, \dots$, em que $s_i \in D \subset \mathbb{R}^2$. A realização desse processo gera um conjunto finito de eventos de localizações em $A \subset D \subset \mathbb{R}^2$ chamado de configuração pontual (WALLER; GOTWAY, 2004). Cada localização de s é uma observação de interesse em processos pontuais e A é a área delimitada para o estudo.

Um dos principais objetivos da análise é estudar a distribuição espacial desses eventos e testar a hipótese de Completa Aleatoriedade Espacial (CAE) ou *Complete Spatial Randomness* (CSR), do padrão de pontos, aquele em que não há efeitos de primeira nem de segunda ordem: a intensidade é constante no plano e os eventos não interagem espacialmente (Processo de Poisson Homogêneo).

Uma variável aleatória pode ser descrita por meio de suas propriedades de primeira (esperança) e segunda (variância) ordem. As variáveis aleatórias oriundas de um processo estocástico espacial também podem ser descritas por essas propriedades. As propriedades de primeira ordem, consideradas globais ou de larga escala, correspondem a variações no valor médio do processo no espaço. Nesse caso, o interesse está na intensidade do processo, isto é, no número de eventos por unidade de área. As propriedades de segunda ordem são denominadas locais ou de pequena escala e representam a dependência espacial do processo, proveniente da estrutura de correlação espacial (BAILEY; GATRELL, 1995).

O estimador de kernel é uma das técnicas mais utilizadas para estimar as propriedades de primeira ordem. Já as funções empíricas baseadas em vizinhos mais próximos (funções F e G) e a função K , são utilizadas para caracterizar os efeitos de segunda ordem. Esses métodos serão descritos a seguir.

2.1.1 Propriedade de primeira ordem

Representados pela função intensidade $\lambda(s)$, os efeitos de primeira ordem correspondem à densidade dos pontos. Ou seja, em processos de Poisson homogêneo, a intensidade é dita uniforme e representa a variação do número de eventos do mapa por unidade de área no espaço ($\lambda = \frac{N(s)}{|A|}$). Entretanto, em processos não homogêneos de Poisson, a intensidade é não uniforme e pode ser definida como um limite matemático:

$$\lambda(s) = \lim_{|ds| \rightarrow 0} \left\{ \frac{\mathbb{E}[N(ds)]}{|ds|} \right\}, \quad (2.2)$$

em que $N(ds)$ representa o número de eventos em uma região ao redor do ponto s ; $|d(s)|$ representa a área dessa região.

Existem vários métodos de investigação e estimação da intensidade, seja mediante métodos paramétricos (modelos de regressão) ou não paramétricos (Kernel).

O método não paramétrico Kernel é o mais utilizado na literatura para estimar a intensidade em toda a região espacial de estudo, obtendo-se assim, uma superfície contínua. Um estimador proposto por Diggle (1985) é dado por

$$\hat{\lambda}(s_0) = \frac{1}{\delta_\tau(s_0)} \sum_{i=1}^n \frac{1}{\tau^2} k\left(\frac{s_0 - s_i}{\tau}\right), \quad (2.3)$$

em que s_0 é um ponto espacial o qual deseja-se estimar a intensidade; s_i representa a localização espacial dos n eventos observados; $k(\cdot)$ é a função kernel escolhida; τ é conhecido como uma largura de banda ($\tau > 0$) que controla o alisamento do kernel na superfície espacial; $\delta(s_0)$ é um fator para correção de efeito de borda, ou seja, a proporção do volume sob o kernel encontrado em s_0 que está contido em A .

Em relação aos efeitos de borda, estimativas da intensidade em localizações próximas às delimitações da região em estudo podem ser viesadas. Com intuito de corrigir este viés, Diggle (1985) propôs um fator de correção de borda que considera a proporção do volume da kernel que se encontra dentro da região espacial, que é expresso por

$$\delta_\tau(s_0) = \int_A \frac{1}{\tau^2} k\left(\frac{s_0 - \mathbf{u}}{\tau}\right) d\mathbf{u}. \quad (2.4)$$

A função $k(\cdot)$ assume a forma de uma função densidade de probabilidade bivariada escolhida. Para os autores Schabenberger e Gotway (2004), esta função pode apresentar várias formas, tais como a kernel Gaussiana, triangular, quártica entre outras. A função quártica utilizada neste trabalho é dada por

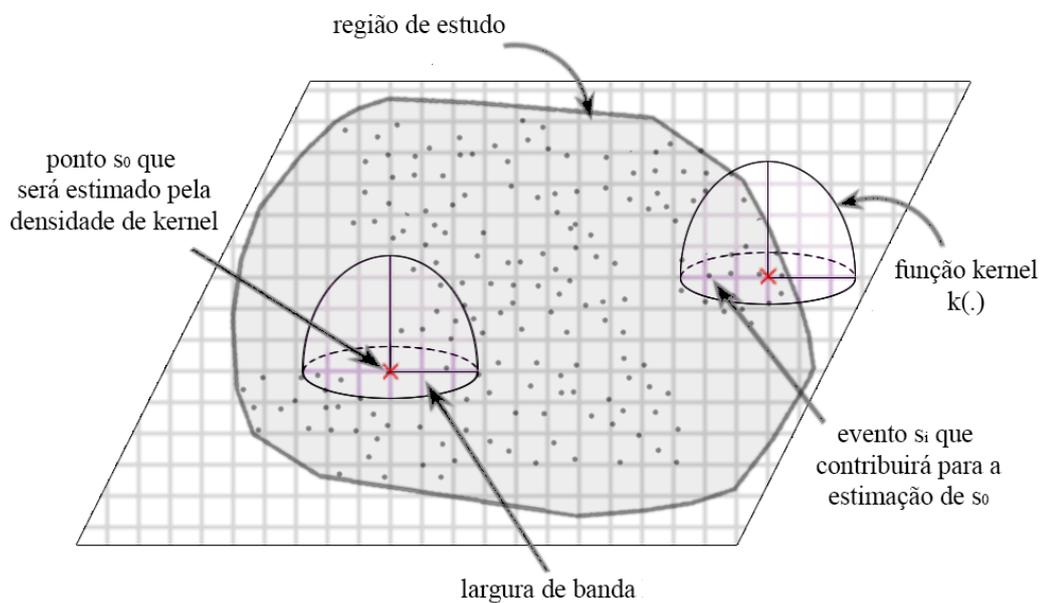
$$k(\mathbf{u}) = \begin{cases} \frac{3}{\pi}(1 - \mathbf{u}^T \mathbf{u})^2, & \text{para } \mathbf{u}^T \mathbf{u} \leq 1. \\ 0, & \text{caso contrário.} \end{cases} \quad (2.5)$$

A função quártica tem a propriedade de atribuir pesos maiores quanto mais próximo o evento se encontrar do local da estimativa. Assim, o decréscimo da intensidade é gradual à medida que o evento se afasta do local da estimativa.

Sabe-se que a escolha da função kernel pouco influencia nas estimativas se comparada com a escolha da largura de banda τ . Valores altos para τ geram estimativas de intensidade mais homogêneas, ao passo que valores mais baixos para τ apresentam estimativas mais heterogêneas (DIGGLE, 2013).

Para uma melhor compreensão deste resultado, a Figura 2.1 apresenta uma esquematização sobre o funcionamento do estimador de kernel, ilustrado em apenas dois pontos localizados na região espacial.

Figura 2.1 – Ilustração do estimador da intensidade por kernel

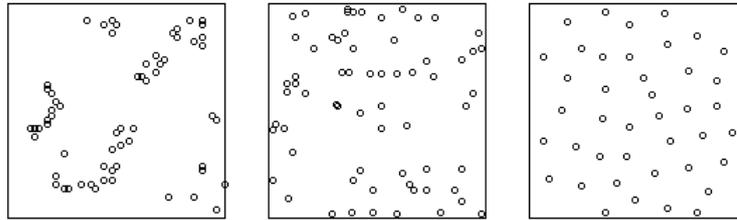


Fonte: Do autor (2017).

2.1.2 Propriedade de segunda ordem

Na análise dos efeitos de segunda ordem, o interesse é identificar a existência de interações entre eventos e como elas ocorrem (DIGGLE, 2013). De maneira geral, uma configuração pontual espacial pode apresentar interações com agrupamento ou regularidade, ou ainda não apresentar interações (aleatória). A Figura (2.2), mostra uma realização simulada de cada um destes tipos.

Figura 2.2 – Tipos de distribuição dos pontos. Da esquerda para a direita: agrupado, aleatório e regular



Fonte: Do autor (2017)

A intensidade de segunda ordem descreve formalmente as propriedades de segunda ordem, medindo a estrutura de dependência entre as localizações \mathbf{s}_i e \mathbf{s}_j e é definida por:

$$\lambda_2(\mathbf{s}_i, \mathbf{s}_j) = \lim_{|ds_i| \rightarrow 0, |ds_j| \rightarrow 0} \frac{\mathbb{E}[N(ds_i)N(ds_j)]}{|ds_i||ds_j|}, \quad (2.6)$$

em que $\mathbb{E}[N(ds_i)N(ds_j)]$ representa o número esperado de eventos dentro de pequenos círculos de área $|ds_i|$ e $|ds_j|$.

A função densidade de covariância e a função de correlação entre pares de eventos, são outras medidas estatísticas equivalentes, que podem descrever as propriedades de segunda ordem. Com o propósito de facilitar o processo de análise, se faz duas suposições:

Estacionariedade: o processo é estacionário quando sua intensidade de primeira ordem não varia sobre a região no espaço d -dimensional e que a relação entre dois eventos depende somente da distância relativa entre eles, isto é, o processo é invariante sobre translação. Assim, temos que $\lambda(\mathbf{s}) = \lambda$ e $\lambda_2(\mathbf{s}_i, \mathbf{s}_j) = \lambda_2(\|\mathbf{s}_i - \mathbf{s}_j\|)$, sendo $\|\cdot\|$ a distância euclidiana entre dois eventos no plano.

Isotropia: é quando a estrutura de dependência espacial entre os eventos for invariante sob rotação (WALLER; GOTWAY, 2004). Nesse caso, a relação entre eventos dependerá somente de sua distância, independente de suas localizações e direções no espaço.

Assim, sob a suposição de estacionariedade e isotropia, tem-se que:

$$\lambda(\mathbf{s}) = \lambda \quad \text{e} \quad \lambda_2(\mathbf{s}_i, \mathbf{s}_j) = \lambda_2(\|\mathbf{s}_i - \mathbf{s}_j\|) = \lambda_2(h), \quad (2.7)$$

em que $h = \|\mathbf{s}_i - \mathbf{s}_j\|$, é a distância euclidiana entre dois pontos.

Trabalhos de Ripley (1976,1977) apresentaram alternativas para descrever as propriedades de segunda ordem para processos estacionários e isotrópicos. O autor apresenta funções que verificam relações de distâncias entre eventos vizinhos mais próximos e entre pares de eventos. As funções F , G e K são algumas dessas funções.

2.1.2.1 A Função F para processos pontuais homogêneos

Seja uma configuração de eventos, gerada por um processo estacionário em uma região plana de área $|A|$. A função F é a função de distribuição acumulada da distância entre um ponto aleatoriamente escolhido e o evento mais próximo, conhecida como a função dos espaços vazios. O mais simples estimador para esta função F pode ser obtido pela equação:

$$\hat{F}(x) = \hat{\lambda}^{-1} \sum_{i=1}^n I_x(x_i), \quad (2.8)$$

em que n é o número de pontos arbitrários e aleatórios no mapa; $\hat{\lambda}$ é o estimador de intensidade; x_i é a distância do i -ésimo ponto aleatório para o evento mais próximo no mapa; $I_x(x_i)$ é uma função indicadora igual a 1 quando $x_i \leq x$ e 0 caso contrário.

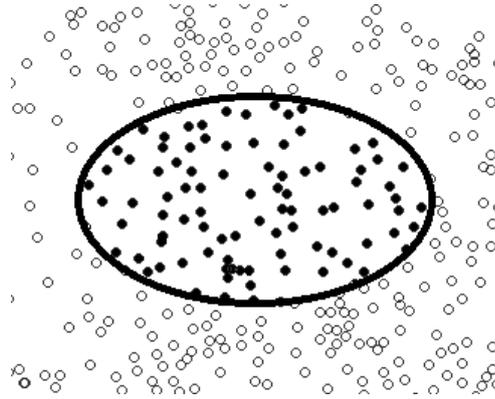
No entanto, o estimador apresentado na Equação (2.8) é viesado, isto é, pontos aleatórios localizados nas bordas do mapa terão uma probabilidade menor de ter um evento próximo do que os pontos localizados ao centro do mapa, para uma mesma distância. Na prática, as regiões estudadas são de área finita 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 esta esteja menor que qualquer distância entre um ponto aleatório e um evento localizado dentro do mapa (Figura 2.3).

Para minimizar esse efeito, existem diversos estimadores que corrigem o efeito de bordas. A amostra reduzida é um estimador não viesado proposto por Ripley (1977), dado pela equação abaixo:

$$\hat{F}(x) = \frac{\sum_{i=1}^n I_x(x_i, r_i)}{\sum_{i=1}^n I_x(r_i)}, \quad x > 0, \quad (2.9)$$

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

Figura 2.3 – Região de estudo (pontos cheios) de área finita, fixada e conhecida.



Fonte: Do autor (2017)

Na prática, o estimador da amostra reduzida não considera na análise os eventos localizados próximos das bordas. Outros estimadores de correção de efeito de bordas podem ser vistos em Baddeley e Gill (1997) e Chiu e Stoyan (1998).

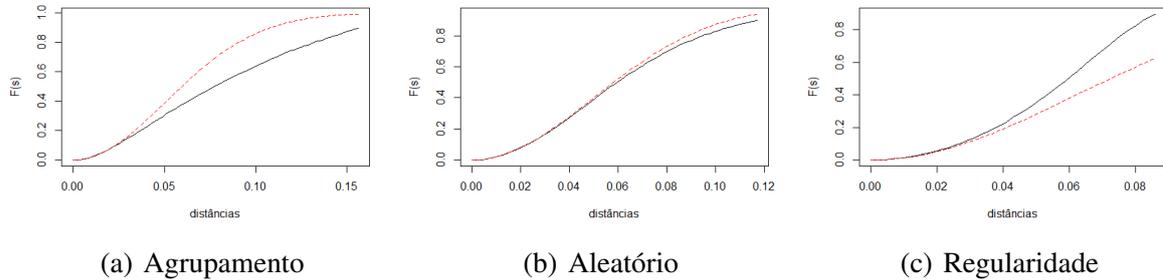
A análise estatística dos padrões de distribuição de pontos requer um modelo teórico de referência. O modelo teórico mais simples é o modelo de Completa Aleatoriedade Espacial que corresponde a uma situação onde é igualmente provável que um evento aconteça em qualquer local dentro da área de estudo, ou seja, os eventos são realizações de um processo de Poisson homogêneo com intensidade λ .

Diggle (2013) demonstra que a função F teórica para um processo de Poisson homogêneo com área infinita é dada pela equação:

$$F(x) = 1 - \exp(-\lambda \pi x^2), \quad \text{com } x > 0. \quad (2.10)$$

Uma maneira simples de interpretar a função F é por meio de gráfico. A Figura (2.4), representa a função F para os três tipos de distribuição dos pontos. A linha pontilhada (vermelha) representa a função teórica descrita em (2.10) e a linha contínua (em preto) representa a função estimada $\hat{F}(x)$ com correção de bordas. Se $\hat{F}(x) < F(x)$ implica agrupamento e se $\hat{F}(x) > F(x)$ representa regularidade. Tendo como base estudos de simulação Diggle (2013), a função F apresenta grande poder para detectar agrupamentos de eventos.

Figura 2.4 – Função $\hat{F}(x)$ empírica (linhas contínuas) para cada um dos três padrões de pontos e a função $F(x)$ teórica para uma processo de Poisson (linhas tracejadas)



(a) Aggrupamento

(b) Aleatório

(c) Regularidade

Fonte: Do autor (2017)

2.1.2.2 A Função G para processos pontuais homogêneos

A função G proposta por Ripley (1977, 1976) é uma função de distribuição acumulada da distância entre eventos, ou seja, a distância de um evento ao seu vizinho mais próximo (y_i). O estimador mais simples, sem correção do efeito de borda para a função G , é definido por:

$$\hat{G}(y) = \hat{\lambda}^{-1} \sum_{i=1}^n I_y(y_i), \quad (2.11)$$

em que n é o número de eventos arbitrários e aleatórios do mapa; $\hat{\lambda}$ é o estimador de intensidade; y_i é a distância do i -ésimo evento aleatório para o mais próximo dos n eventos analisado no mapa; $I_y(y_i)$ é uma função indicadora igual a 1 quando y_i é menor ou igual a y e 0 caso contrário.

Ripley(1977) propôs um estimador não viesado que corrige o efeito de bordas dado pela equação:

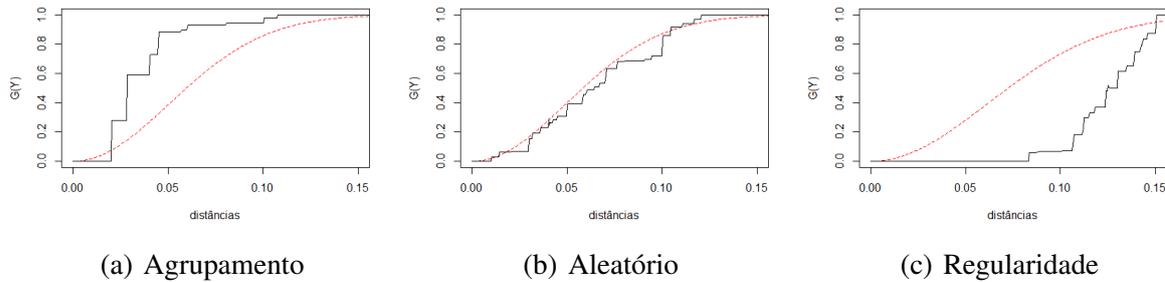
$$\hat{G}(y) = \frac{\sum_{i=1}^n I_y(y_i, r_i)}{\sum_{i=1}^n I_y(r_i)} \quad \text{com } y > 0, \quad (2.12)$$

em que r_i é a distância de um evento até o mais próximo na borda do mapa; $I_y(y_i, r_i)$ é uma função indicadora. Sob a hipótese de CAE, a função G é dada por $G(y) = 1 - \exp(-\lambda \pi y^2)$. Tendo como base estudos de simulação Diggle (2013), a função G apresenta grande poder para detectar configurações com eventos distribuídos regularmente.

O gráfico da função G é simples de interpretação. A Figura (2.5), representa a função G para os três tipos de distribuição dos pontos. A linha pontilhada (vermelha) representa a função

teórica e a linha cheia (preta) representa a função estimada $\hat{G}(y)$ com correção de bordas (2.12). Se $\hat{G}(y) > G(y)$ representa agrupamento e se $\hat{G}(y) < G(y)$ implica regularidade.

Figura 2.5 – Função $\hat{G}(y)$ empírica (linhas contínuas) para cada um dos três padrões de pontos e a função $G(y)$ teórica para uma processo de Poisson (linhas tracejadas)



Fonte: Do autor (2017)

2.1.2.3 A função K para processos pontuais homogêneos

Um dos problemas das funções F e G , descritas anteriormente, é que elas desconsideram padrões de escalas maiores, perdendo-se informações para grandes distâncias, porque as abordagens são realizadas com eventos vizinhos. Para obter um panorama geral da dependência espacial, sob diversas escalas e igualmente poderoso para detectar tanto os padrões com agrupamento como regularidade, recomenda-se utilizar a função K , também proposta por Ripley (1977).

Segundo Cressie (1993) a finalidade da função K está em capturar a dependência espacial a partir de diferentes escalas de distâncias entre eventos de uma configuração pontual. Considerando os conceitos de estacionariedade e isotropia, a dependência espacial é obtida exclusivamente pelas distâncias entre eventos e não depende das suas localizações relativas e suas direções. Atendendo essas condições, a função K é definida por

$$K(\mathbf{h}) = \lambda^{-1} E[N(\mathbf{h})], \quad (2.13)$$

em que $E[N(\mathbf{h})]$ é o valor esperado do número de pares eventos dentro de uma distância \mathbf{h} em torno de um evento arbitrário; λ é a intensidade do processo pontual.

Baddeley, Møller e Waagepetersen (2000) discutem intensidade ponderada (de segunda ordem), definida por

$$\frac{\lambda_2(\mathbf{s}_i, \mathbf{s}_j)}{\lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j)} = \frac{\lambda_2(h)}{\lambda^2} = g(\mathbf{h}), \quad (2.14)$$

em que $g(\mathbf{h})$ é a função de correlação de pares.

Diggle (2013) define a função teórica K de um processo pontual estacionário em termos da função de correlação entre pares de eventos, isto é:

$$K(\mathbf{h}) = 2\pi \int_0^h g(x)xdx. \quad (2.15)$$

Para representar o caso de CAE sob a suposição de independência e estacionariedade, a função K é dada por

$$\begin{aligned} K(\mathbf{h}) &= 2\pi \int_0^h g(s_i, s_j)xdx \\ &= 2\pi \int_0^h \frac{\lambda_2(s_i, s_j)}{\lambda(s_i)\lambda(s_j)}xdx && \text{(independência)} \\ &= 2\pi \int_0^h \frac{\lambda(s_i)\lambda(s_j)}{\lambda(s_i)\lambda(s_j)}xdx && \text{(estacionariedade)} \\ &= 2\pi \int_0^h \frac{\lambda^2}{\lambda^2}xdx \\ &= 2\pi \int_0^h xdx \\ &= \pi h^2. \end{aligned} \quad (2.16)$$

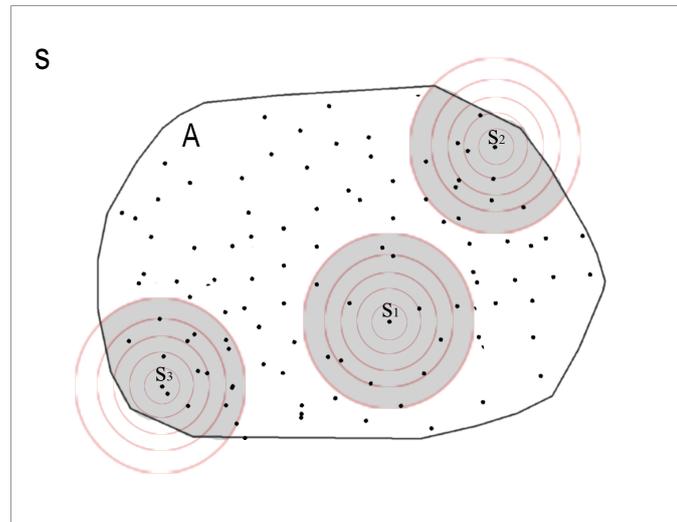
O estimador da função K proposto por Ripley (1977) é dado por

$$\hat{K}(\mathbf{h}) = \hat{\lambda}^{-1} \sum_{i=1}^n \sum_{i \neq j} \frac{I(\|\mathbf{s}_i - \mathbf{s}_j\| \leq h)}{n-1}, \quad (2.17)$$

em que $\hat{\lambda}$ é um estimador da intensidade; $I(\cdot)$ é a função que indica se existe, ou não, evento com distância menor ou igual a \mathbf{h} .

Uma alternativa em compreender melhor o estimador da função K é estimar o número médio de eventos dentro de um círculo de raio h centrado no evento \mathbf{s}_i e dividi-lo pelo estimador da intensidade homogênea do processo. Na Figura (2.6), ilustra-se a função K estimando três eventos para diferentes distâncias de raio h .

A ideia é que seja feito o procedimento para cada $\mathbf{s}_i \in A$. Cada uma das circunferências em torno dos eventos \mathbf{s}_i representa as várias distâncias espaciais a serem utilizadas, a fim de

Figura 2.6 – Processo de estimação da função K para três eventos

Fonte: Do autor (2017)

que a dependência espacial possa ser obtida em diferentes escalas. Os eventos s_2 e s_3 estão próximos às delimitações da região de estudo e, ao serem estimados, pode ocorrer um viés (efeito de borda). Assim, o estimador da função K dado em (2.17) não corrige esse problema.

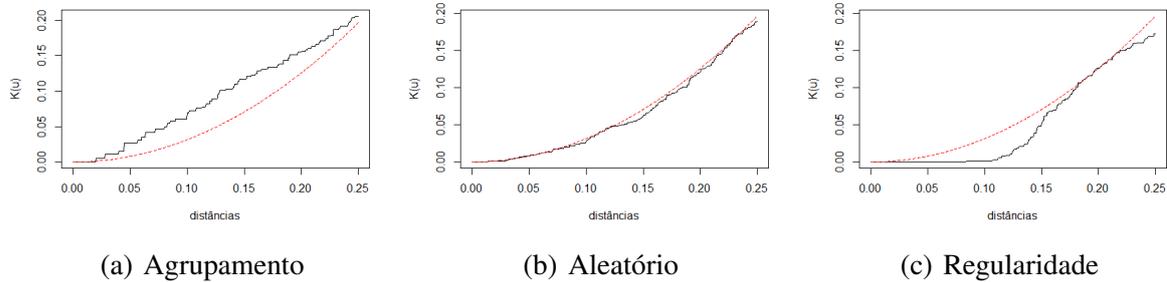
Existem diversas propostas para a correção desses efeitos de borda. A mais utilizada consiste em ponderar o estimador pela proporção da circunferência, de raio h , que se encontra dentro da região de estudo A . Utilizando este conceito, Ripley (1977) propôs o seguinte estimador para a função K com correção para efeito de borda:

$$\hat{K}(\mathbf{h}) = \hat{\lambda}^{-1} \sum_{i=1}^n \sum_{i \neq j} \frac{1}{a_{ij}} \frac{I(\|\mathbf{s}_i - \mathbf{s}_j\| \leq \mathbf{h})}{n-1}, \quad (2.18)$$

em que a_{ij} é a proporção da área da circunferência que se encontra dentro da região A , centrado em um evento arbitrário e os demais são análogos a (2.17).

Essa função pode ser utilizada para testar a hipótese de CAE. Na Figura (2.7) é possível observar que para um processo aleatório tem-se $K(\mathbf{h}) = \pi \mathbf{h}^2$ (linha pontilhada vermelha), para $\hat{K}(\mathbf{h}) > \pi \mathbf{h}^2$ o processo é agregado, isto é, os eventos têm em média mais vizinhos do que o esperado sob a hipótese de CAE e $K(\mathbf{h}) < \pi \mathbf{h}^2$ os eventos têm em média menos vizinhos do que o esperado sob a hipótese de CAE, ou seja, o processo é regular.

Figura 2.7 – Função $\hat{K}(h)$ empírica (linhas contínuas) para cada um dos três padrões de pontos e a função $K(h)$ teórica para uma processo de Poisson (linhas tracejadas)



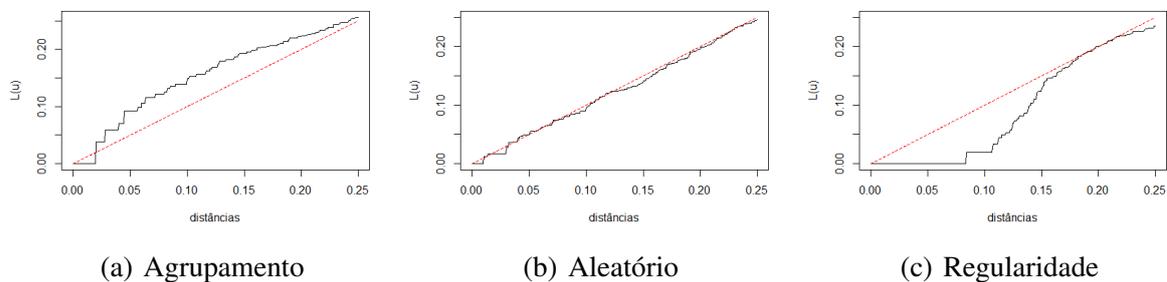
Fonte: Do autor (2017)

Com o objetivo de facilitar a interpretação da função K , estão presentes na literatura várias transformações dessa função. Em geral, essas transformações são utilizadas para linearizá-la e estabilizar a variância da função K . A função L é uma transformação proposta por Besag (1977),

$$\hat{L}(\mathbf{h}) = \sqrt{\frac{\hat{K}(\mathbf{h})}{\pi}}. \quad (2.19)$$

A interpretação gráfica para a função L é análoga à função K . A Figura (2.8) apresenta o gráfico das Funções L empírica (curva em vermelho) e teórica (curva em preto) para os possíveis cenários.

Figura 2.8 – Função $\hat{L}(h)$ empírica (linhas contínuas) para cada um dos três padrões de pontos e a função $L(h)$ teórica para um processo de Poisson (linhas tracejadas)



Fonte: Do autor (2017)

2.1.3 Testes contra hipótese de Completa Aleatoriedade Espacial

Uma abordagem baseada em simulações de Monte Carlo, para realizar um teste contra a hipótese nula de CAE foi proposta por Diggle (2013). Esta abordagem compara a função de distribuição observada e a função de distribuição esperada, sob a hipótese de CAE. Utiliza-se tanto testes de hipóteses formais quanto procedimentos gráficos (conhecidos como envelopes de confiança) para avaliar se essa comparação entre as duas funções é estatisticamente significativo.

É necessário definir uma medida de discrepância entre as funções estimada e teórica. Para uma função escolhida (F , G ou K), a estatística de teste é dada por

$$u_i = \int_0^R [\hat{S}(r) - S_{teo}(r)]^2 dr, \quad (2.20)$$

em que R é um limite superior escolhido para o intervalo de distâncias de interesse; $\hat{S}(r)$ é a função estimada escolhida; $S_{teo}(r)$ é a função teórica escolhida.

Entretanto a estatística u não têm distribuição amostral conhecida, a qual pode ser obtida por simulações de Monte Carlo, para realizar o teste de hipótese.

- i) Calcula-se a estatística de u_1 para os dados observados;
- ii) Simula-se t padrões de pontos espaciais independentes sob hipótese nula de CAE com a mesma intensidade que o padrão observado;
- iii) Calcula-se u_i para cada padrão simulado ($i = 2, \dots, t$);
- iv) Compara-se o valor de u_1 com os valores u_i com ($i = 2, \dots, t$) obtidos a partir das configurações simuladas;
- v) Se u_1 ficar entre o maiores valores de u_2, \dots, u_t , então, tem-se indicação que a configuração observada não apresenta CAE.

Suponha $u_1 = u_{(j)}$ para algum $j \in \{1, \dots, t\}$ então rejeita-se a hipótese de aleatoriedade espacial se valor-p = $\left(\frac{t+1-j}{t}\right) \leq \alpha$, em que α é o nível de significância escolhido. Todo este procedimento pode ser conduzido para as funções F , G e K .

O teste não fornece informações sobre a causa da rejeição da CAE e, portanto, não é possível decidir se a rejeição é em direção à regularidade ou agrupamento. Assim, a causa da rejeição da hipótese de CAE pode ser explicada pela abordagem gráfica, utilizando envelopes de simulação.

Os envelopes de simulação inferior e superior são construídos a partir de t simulações Monte Carlo, sob a hipótese nula de CAE, com n eventos na região $|A|$, dados por

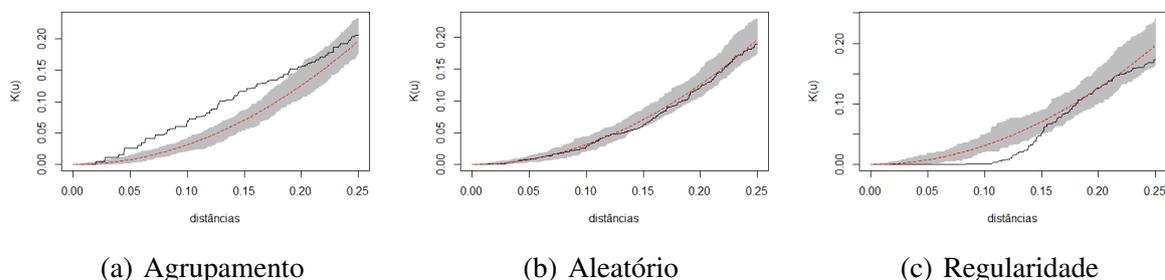
$$L(h) = \min_{i=1, \dots, t} \{\hat{S}_i(h)\} \quad \text{e} \quad U(h) = \max_{i=1, \dots, t} \{\hat{S}_i(h)\},$$

em que $S(\cdot)$ representa a função que está em estudo; $L(h)$ é os mínimos da função $S(\cdot)$; $U(h)$ é os máximos da função $S(\cdot)$, calculadas a partir de t simulações independentes de processo aleatórios com n eventos.

A análise gráfica das funções apresentadas anteriormente poderá ser utilizada como método formal para se comparar estatisticamente a distribuição dos eventos observados, com o que se esperaria na hipótese de CAE. Para isso, são construídos envelopes de simulação para a distribuição de CAE. Os envelopes de simulação inferior e superior são construídos a partir de t simulações Monte Carlo, sob a hipótese nula de CAE, com n eventos na região $|A|$, dados por

Na análise do gráfico com a distribuição e os envelopes para a função $F(\cdot)$, valores acima do envelope de simulação revelam configurações pontuais regulares e, abaixo do envelope, configurações pontuais agrupadas. Para as funções $G(\cdot)$, $K(\cdot)$ e $L(\cdot)$, a análise é interpretada de maneira contrária à função F , ou seja, valores acima do envelope de simulação representam agrupamento e, abaixo, configurações pontuais com regularidade. A Figura (2.9) apresenta um exemplo da função K com os envelopes de simulação (cinza), para os três tipos de distribuição dos pontos.

Figura 2.9 – Função $\hat{K}(h)$ empírica (linhas contínuas) para cada um dos três padrões de pontos e a função $K(h)$ teórica para uma processo de Poisson (linhas tracejadas), envelopes de simulação (cinza)



Fonte: Do autor (2017)

A metodologia da análise para um único padrão de pontos é bastante extensa. No entanto, com os avanços relativamente recentes em microscopia e tecnologia, os cientistas estão

descobrimos que seus dados podem consistir em vários padrões de pontos espaciais, os quais denominam-se de processos pontuais replicados.

2.2 Processos pontuais replicados não marcados

Um experimento que pode ser repetido várias vezes sob condições idênticas é denominado de experimento com observações replicadas (réplicas). É de suma importância replicações na estatística, pois elas permitem observar a variabilidade e tentar separar várias fontes de variabilidade e estudá-las. Em configurações pontuais espaciais isso não é diferente, sendo que nessa situação tem-se o padrão de pontos replicados.

Os autores Baddeley, Rubak e Turner (2015) definem dados de configurações pontuais replicadas como o conjunto de padrões de pontos que podem ser considerados como realizações do mesmo processo estocástico pontual. A replicação é obtida quando um padrão de pontos observado é o resultado de uma experiência que pode ser repetida em condições idênticas.

Formalmente, Illian et al. (2008) definem v configurações pontuais de áreas A_1, \dots, A_v , podendo ser congruentes, ou não. Os v padrões de pontos em A_i , com $i = 1, \dots, v$, são considerados como amostras independentes e identicamente distribuídas (i.i.d.), estacionários e isotrópicos. O objetivo dessa análise é encontrar estimativas para características conjuntas dos dados, tais como λ e funções: G, K . Usualmente, estimam-se essas características separadamente para cada uma das áreas A_i e, em seguida, agregam-se essas estimativas.

Os métodos para analisar padrões de pontos replicados são classificados em métodos paramétricos e não paramétricos (LANDAU; EVERALL, 2008). A abordagem não paramétrica usa estimativas separadas para cada padrão e incorpora a replicação de padrões com métodos *ad-hoc*, que são apropriados quando as configurações pontuais são decorrentes de um processo estocástico, cujo modelo não é assumido explicitamente na análise.

Neste trabalho serão utilizados métodos não paramétricos, que avaliam as diferenças nas funções de resumo (principalmente a função K de Ripley) entre grupos experimentais. Entre as abordagens desenvolvidas dentro dos métodos não-paramétricos, o trabalho de Diggle, Lange e Benes (1991) foi o principal motivador para análise de processos pontuais replicados; os autores analisaram 30 configurações pontuais de localizações de neurônios dividido em três grupos. O interesse foi comparar o arranjo espacial dos neurônios, a fim de identificar as diferenças anatômicas entre grupos de pessoas normais, esquizoafetivo e esquizofrênicas. Resumindo cada padrão por uma estimativa da função K e comparando com a média das funções K entre

grupos experimentais, os autores propuseram um procedimento análogo à análise de variância (ANOVA) para um delineamento inteiramente casualizado (DIC). Configurações pontuais replicadas não marcadas também podem ser analisadas em termos de propriedades de primeira e segunda ordem.

2.2.1 Propriedade de primeira ordem para dados replicados

Para dados replicados temos v configurações pontuais de áreas A_1, \dots, A_v . Se essas áreas forem iguais, assumem-se intensidades específicas para cada configuração pontual, $\hat{\lambda}_i$ para $i = 1, \dots, v$. Neste caso, o estimador já foi visto anteriormente e uma alternativa é fazer a média aritmética das intensidades.

Se as áreas estudadas forem de tamanhos diferentes, existem vários métodos de agregação que podem ser aplicados. Illian et al. (2008) apresentam uma opção de uma intensidade comum que pode ser estimada $\bar{\lambda}$, como descrito a seguir: Seja a intensidade $\hat{\lambda}_1, \dots, \hat{\lambda}_v$ estimadores de λ , dados por $\hat{\lambda}_i = \frac{n_i}{|A_i|}$, para as áreas A_1, \dots, A_v . Assim, o estimador agregado é dado por:

$$\bar{\lambda} = \sum_{i=1}^v \hat{\lambda}_i \frac{A_i}{\omega}, \quad (2.21)$$

em que A_i é a área de cada v configuração pontual; ω é definida como a soma de todas as áreas de cada v configuração pontual, isto é, $\omega = \sum_{i=1}^v A_i$.

2.2.2 Propriedade de segunda ordem para dados replicados

Se estiverem disponíveis as repetições, é possível estimar a função K a partir de cada repetição e, posteriormente, agrupar as funções K individuais em uma estimativa que, em seguida, representa todo o grupo do estudo. Quando todos os padrões são repetições advindas de um mesmo processo estocástico, as estimativas correspondentes da função K são distribuídas de forma idêntica e uma estimativa geral pode ser obtida por meio de uma média aritmética simples.

Seja v configurações pontuais, cada uma observada em uma região A , conforme menciona Diggle (2013), a função $K_{ij}(h)$ com $i = 1, \dots, g$ e $j = 1, \dots, v_i$ consistindo em g grupos de tamanhos v_1, \dots, v_g . A função K do i -ésimo grupo é obtida por ponderação das estimativas individuais $\hat{K}_{ij}(h)$ dada por

$$\bar{K}_i(h) = \frac{1}{n_i} \sum_{j=1}^{v_i} n_{ij} \hat{K}_{ij}(h). \quad (2.22)$$

O estimador da função K geral é dado por pela média dos grupos

$$\bar{K}(h) = \frac{1}{n} \sum_{i=1}^g n_i \bar{K}_i(h), \quad (2.23)$$

em que $n_i = \sum_{j=1}^{v_i} n_{ij}$ e $n = \sum_{i=1}^g n_i$.

A função $\bar{K}_i(h)$ é definida de modo que não dependa da intensidade dos eventos e esta é interpretada de maneira análoga a função $\hat{K}(h)$.

Quando se tem grupos, o principal interesse é verificar se existem evidências estatísticas de diferença entre os grupos. Para tal, foram propostos alguns métodos para detectar essas diferenças. O primeiro teste proposto foi elaborado por Diggle, Lange e Benes (1991), sugerindo o uso de método de *bootstrap* para determinar a distribuição amostral da estatística do teste em conjuntos de dados, consistindo em vários padrões pontuais amostrados independentemente. Esse mesmo teste foi estudado por Baddeley et al. (1993) e aplicado na análise de repetições independentes de padrões pontuais tridimensionais.

Diggle, Mateu e Clough (2000) apresentaram uma versão mais recente desse teste de *bootstrap*, que faz a diferença entre as médias de grupos por meio da função K de repetições independentes. Este teste foi utilizado também em Ramón et al.(2016). Porém, para amostras pequenas este teste apresenta uma não uniformidade dos valores- p sob a hipótese nula.

2.2.3 Teste para comparação de grupos em padrão de pontos

Hahn (2012) criou uma alternativa para contornar esse problema dos valores- p , elaborando um novo teste para a hipótese nula. O autor partiu da ideia de que duas ou mais configurações pontuais são realizações de um mesmo processo pontual, ou seja, sem diferença entre grupos. Além disso, supôs que os padrões de pontos observados são aleatórios, independentes e identicamente distribuídos, independente do grupo ao qual pertençam.

Suponha que tenha g grupos de padrões de pontos, contendo v_1, \dots, v_g padrões, respectivamente. Os padrões de pontos no grupo 1 são denominados b_{11}, \dots, b_{1v_1} , etc. Uma estimativa da função K é calculada para cada configuração. É realizada uma comparação por meio de um teste de permutação utilizando uma estatística de teste estudentizada dado por

$$T = \sum_{1 \leq i < j \leq g} \int_0^{u_0} \frac{(\bar{K}_i(h) - \bar{K}_j(h))^2}{\frac{1}{v_i} s_i^2(h) + \frac{1}{v_j} s_j^2(h)} dh, \quad (2.24)$$

em que $\bar{K}_i(h) = \frac{1}{v_i} \sum_{j=1}^{v_i} \hat{K}_{ij}(h)$ é a média dos grupos; $s_i^2(h) = \frac{1}{v_i-1} \sum_{j=1}^{v_i} (\hat{K}_{ij}(h) - \bar{K}_i(h))^2$ é a variância amostral dentro do i -ésimo grupo para uma determinada distância h .

A estatística dada em (2.24) é muito semelhante à estatística utilizada no teste Monte Carlo de Diggle, Mateu e Clough (2000), quando a comparação for entre dois grupos. Porém os testes baseados na estatística T são sensíveis quando há diferenças dos tamanhos das áreas ou às densidades pontuais. Recomenda-se substituir a variância $s_i^2(h)$ por uma versão suavizada \bar{s}_i^2 e usar a estatística \bar{T} , dada por

$$\bar{T} = \sum_{1 \leq i < j \leq g} \int_0^{h_0} \frac{(\bar{K}_i(h) - \bar{K}_j(h))^2}{\frac{1}{v_i} \bar{s}_i^2(h) + \frac{1}{v_j} \bar{s}_j^2(h)} dh, \quad (2.25)$$

com

$$\bar{s}_i^2(h) = h^2 \frac{1}{h_0} \int_0^{h_0} \frac{s_i^2(t)}{t^2} dt.$$

O uso de \bar{T} , em vez de T , melhora a robustez do teste no caso em que a intensidade difere entre os grupos, mas pode levar a uma perda de poder para padrões de pontos regulares (BADDELEY; RUBAK; TURNER, 2015).

A estatística T não têm distribuição amostral conhecida. Assim, utiliza-se um método de computação intensiva para obtê-la. Seja T_1 a estatística do teste e T_2, \dots, T_p as estatísticas calculadas a partir de p permutações. Calcula-se a estatística de T_1 . Compara-se o valor de T_1 com os valores T_2, \dots, T_p obtidos a partir das estatísticas permutadas. Se T_1 ficar entre o maiores valores de T_2, \dots, T_p , então, tem-se indicação que as configurações pontuais não são realizações de um mesmo processo pontual, isto é, existe diferença entre os grupos. Suponha $T_1 = T_{(j)}$ para algum $j \in \{1, \dots, p\}$ então rejeita-se a hipótese nula se o valor-p = $\frac{(p+1-j)}{p} \leq \alpha$.

2.3 Processos pontuais marcados estacionários

Penttinen, Stoyan e Henttonen (1992) definem um processo pontual marcado como sendo um processo estocástico $\Psi = [s_i; m(s_i)]$ que gera os eventos (s_i) e as marcas associadas aos eventos $m(s_i)$. Uma realização de Ψ é um conjunto de posições juntamente com as marcas associadas, em que as marcas podem ser quantitativas ou qualitativas.

As marcas qualitativas são variáveis categóricas, que descrevem diferentes tipos de eventos, tais como espécies de plantas. O processo é chamado de bivariado se o evento pode ter apenas dois tipos de marcas; caso contrário, é multivariado. As marcas quantitativas são valores reais que descrevem, por exemplo, altura da árvore, diâmetro de células. Claramente, as

marcas quantitativas podem ser reduzidas a marcas qualitativas, dividindo as marcas em classes discretas, como pequeno, médio e grande.

O processo pontual marcado é definido por meio de um processo estacionário e isotrópico, semelhante à ideia de processo pontuais não marcados. Illian et al. (2008) definem um processo $\Psi = [s_1; m(s_1)], [s_2; m(s_2)], \dots$ juntamente com um processo transladado $\Psi_s = [s_1 + s; m(s_1)], [s_2 + s; m(s_2), \dots]$. Isto significa que no processo pontual marcado transladado, as marcas permanecem as mesmas, somente os pontos são transladados. Assim, o processo de pontos marcados Ψ é dito estacionário se os efeitos de primeira e segunda ordem forem constantes em toda a região A estudada.

A definição de isotropia é análoga aos processos pontuais não marcados, vistos anteriormente. Considera-se o processo Ψ estacionário, o conceito de isotropia, se além de ser estacionário, no processo Ψ as covariâncias dos pontos s_i e s_j dependem apenas da distância euclidiana entre eles, e não da direção.

Um processo pontual marcado é descrito em termos de propriedades de primeira e segunda ordem. No caso estacionário são utilizados dois tipos diferentes de características de primeira ordem: a que diz respeito aos eventos, isto é, a intensidade λ , e a outra que descreve as marcas (ILLIAN et al., 2008). As características de segunda ordem são úteis para caracterizar não apenas a variabilidade da distribuição dos eventos, mas também a variabilidade das marcas e, ainda, descrever correlações entre marcas e pontos (WÄLDER; STOYAN, 1996).

Assim, o objetivo da análise de processos pontuais marcados é verificar se existe dependência entre marcas e pontos, mas isso vai depender da escolha do tipo da marca. Se as marcas forem qualitativas, o interesse é estudar as posições relativas dos diferentes tipos de pontos, verificando a existência de agregação ou repulsão dentro do tipo e entre os tipos.

Já para marcas quantitativas, a análise avalia se as marcas variam continuamente verificando assim, independência entre as marcas, isto é, até que ponto a interação entre as marcas tem influência sobre as mesmas. Por exemplo, as marcas de pontos vizinhos podem tender a ser menores (ou maiores) do que a marca média, isto é, pode haver inibição ou estimulação das marcas. Outra questão diz respeito à semelhança numérica das marcas de eventos vizinhos. Isto significa que estes pontos tendem a ter marcas semelhantes, mas o comportamento oposto também pode ser observado, ou seja, pontos específicos podem dominar os pontos em sua vizinhança e, portanto, ter uma marca grande, enquanto os outros pontos têm marcas pequenas.

Frequentemente, os eventos e as marcas em um padrão de pontos marcados são correlacionados e para isso é necessário testar a hipótese de independência. Se as marcas forem independentes da localização do ponto, basta estudar simplesmente a distribuição da marca na região de estudo.

Existem alguns métodos para analisar a independência entre a localização do ponto e sua marca; destaca-se aqui a função K ponderada por marcas, que é uma generalização da função K de Ripley para padrões de pontos marcados.

2.3.1 Propriedade de segunda ordem processos pontuais marcados

A função K ponderada por marcas, de acordo com Penttinen, Stoyan e Henttonen (1992), é uma extensão da função K de Ripley, na qual para cada par de eventos existe uma ponderação por uma função teste $f(m_1, m_2)$ de suas respectivas marcas.

Um estimador não-viesado, que corrige o efeito de borda para a função K_f é dado por

$$\hat{K}_f(h) = \frac{|A|}{n^2} \sum_{k=1}^n \sum_{l=1, l \neq k}^n f(m(\mathbf{s}_k), m(\mathbf{s}_l)) 1(\|\mathbf{s}_k - \mathbf{s}_l\| \leq h) \delta(\mathbf{s}_k, \mathbf{s}_l), \quad (2.26)$$

em que n é o número de pontos marcados na região de estudo A de área $|A|$; f é uma função que atribui um peso a um par dos pontos \mathbf{s}_k e \mathbf{s}_l dependendo do tipo de suas marcas $m(\mathbf{s}_k)$ e $m(\mathbf{s}_l)$; $1(\cdot)$ é a função que é igual a 1 se o argumento for verdadeiro e 0 caso contrário; $\delta(\mathbf{s}_k, \mathbf{s}_l)$ é um termo de correção de borda.

A escolha de f se dá de acordo com o tipo das marcas, no caso das marcas qualitativas, defini-se a função f como

$$f(m(\mathbf{s}_k), m(\mathbf{s}_l)) = \frac{n^2}{n_i n_j} 1(m(\mathbf{s}_k) = i) 1(m(\mathbf{s}_l) = j), \quad (2.27)$$

com n_i e n_j denotando os números de i -ésimo e j -ésimo pontos, conhecida como a função K_{ij} bivariada (DIGGLE, 2013).

Para a função K_f com marcas quantitativas utiliza-se a função f como sendo

$$f(m(\mathbf{s}_k), m(\mathbf{s}_l)) = \frac{m(\mathbf{s}_k)m(\mathbf{s}_l)}{\hat{m}^2}, \quad (2.28)$$

em que $\hat{m} = \frac{1}{n} \sum m(s_i)$ é a média das marcas.

No caso de marcas quantitativas, a hipótese nula estabelece marcas independentes dos pontos, o que caracteriza a hipótese de Completa Aleatoriedade Espacial na configuração de pontos não marcados. Quando a hipótese nula é rejeitada, pode-se ter o indicativo de estimulação ou inibição entre marcas. A estimulação indica que eventos vizinhos têm as marcas maiores do que a marca média, enquanto que a inibição indica que as marcas são menores do que a marca média para pontos próximos.

2.4 Processos pontuais replicados com marcas quantitativas

A análise de processos pontuais replicados marcados vem sendo aplicada em algumas áreas do conhecimento. Essas análises são realizadas exclusivamente em dados com marcas qualitativas bivariadas ou multivariadas. Os métodos de análise de processos pontuais replicados marcados podem ser paramétricos e não paramétricos.

O trabalho de Landau, Rabe-Hesketh e Everall (2004) discute o problema da análise de padrões de pontos espaciais marcados bivariados replicados sob a abordagem não-paramétrica. Este trabalho é uma extensão do trabalho proposto por Diggle, Lange e Benes (1991) para configurações replicadas não marcadas. Ambos os trabalhos tem como objetivo utilizar a função K para comparar respostas de grupos submetidos a diferentes condições experimentais. Como em uma análise de variância clássica, o objetivo é testar a hipótese nula de que não existe diferença entre as médias dos grupos, em termos de função K.

Para análise paramétrica para configurações replicadas de pontos marcados destacam-se os trabalhos de Eckel et al. (2009) e Gervini e Baur (2017). Não foram encontrado na literatura trabalhos que utilizem a análise não paramétrica de padrão de pontos replicados com marcas quantitativas. Assim, um dos propósitos desta tese é a criação de um teste de hipótese que identifique diferenças entre grupos de configurações pontuais replicadas com marcas quantitativas (Artigo 3).

Os métodos estatísticos apresentados anteriormente para análise de processos pontuais replicados marcados ou não marcados não estão sendo utilizados para análise de dados de panificação. Assim, esta tese explora o potencial desses métodos nesse tipo de análise. Para tal exploração, estudou-se o problema da distribuição alveolar na massa do pão francês. Neste sentido apresenta-se uma breve revisão bibliográfica sobre a panificação.

2.5 Panificação

O pão francês é considerado um alimento universal, fonte de carboidratos e proteínas e consumido pela população infantil, adolescentes e adultos, fazendo parte de lanches ou como acompanhamento nas refeições principais. A fim de atender as exigências do consumidor final com pães frescos ou recém-assados, novas tecnologias para a fabricação e a escolha do tipo de armazenamento da massa de alta qualidade de pães crus estão cada vez mais sendo estudadas e desenvolvidas e, conseqüentemente, o congelamento é uma ótima alternativa.

O tempo em que o alimento é conservado sob determinadas condições de temperatura, umidade relativa, luz e outras é chamado vida útil; neste período ocorrem algumas alterações no produto que, até certo ponto, podem ser consideradas aceitáveis pelo fabricante, pelo consumidor e pela legislação vigente. O congelamento é, atualmente, um mecanismo que otimiza a vida útil, por se tratar de uma solução viável economicamente. Porém os chamados de *flaking* são um dos principais defeitos de qualidade em pães congelados, que é o desprendimento da casca do miolo. Esse fenômeno pode ser entendido, por um lado, quando ocorre a concentração de cristais de gelo entre o miolo e a casca devido à migração de umidade do centro da peça, que provoca uma zona de fragilidade. Por outro lado, ocorre um processo de contração do pão durante o congelamento, devido ao choque termomecânico. Assim, a qualidade sensorial do pão deve ser estudada durante o armazenamento.

Para os autores Ribotta et al. (2004), Bárcenas e Rosell (2007) e Almeida e Chang (2014) diversas outras alterações podem ocorrer durante o armazenamento dos pães sob congelamento e, à medida que o tempo de armazenamento aumenta: a redução do volume específico, diminuição no teor de umidade, aumento da dureza do miolo e mudanças no aroma, resultando em um produto deteriorado.

Os pães assados e, posteriormente, congelados eram mais macios que os pães provindos de massa congelada, devido à massa possuir alvéolos grandes, uniformes e bem distribuídos no miolo do pão, além de terem fatores como mobilidade da água, temperatura de armazenamento e formulação influenciando a estrutura do miolo (ECKARDT et al., 2013). Diante disso, o congelamento afeta diretamente a distribuição alveolar e têm sido um grande foco de estudo na tecnologia de panificação. Atualmente, a análise da distribuição alveolar é realizada por meio de métodos estatísticos clássicos.

O trabalho de Esteller e Lannes (2005) propôs a reavaliação da legislação brasileira para técnicas e aparelhagens, caracterizando melhor produtos derivados do trigo. Foram analisadas

diferentes marcas de pães e torradas nos parâmetros de porosidade da massa (alvéolos), textura, cor, volume específico, densidade e umidade. Foram obtidos valores médios para cada grupo de produtos, concluindo que é perfeitamente viável uma complementação na fixação da identidade e qualidade de produtos panificados.

Já os autores *Zavareze et al. (2008)* investigaram o efeito do soro de leite no teor proteico e na qualidade tecnológica e sensorial de pães. Foram avaliadas várias características, como cor do miolo, estrutura da célula (alvéolos), textura do miolo entre outras. Os resultados foram obtidos mediante uma análise de variância e teste de Tukey e concluíram que os pães com soro de leite apresentaram maior teor protéico e notas tecnológicas superiores ao pão padrão, com exceção do pão com soro desidratado que apresentou menor volume específico. Na avaliação sensorial, os pães com adição de soro não apresentaram diferença em relação ao pão padrão.

Outro trabalho foi o estudo da relação entre o tamanho dos alvéolos e aceitação de pães de forma enriquecidos com farelo de mandioca. Os autores *Castiglioni et al. (2014)* realizaram as análises físico-química e microbiológica, avaliadas por meio do diâmetro médio dos alvéolos da massa, análise de imagem e teste de aceitação. Concluiu-se que o uso da farinha de mandioca seco em pão de forma é viável, do ponto de vista tecnológico, microbiológico e sensorial.

REFERÊNCIAS

- ALMEIDA, E. L.; CHANG, Y. K. Influence of different enzymes during the frozen storage of pre-baked french bread elaborated with whole-wheat flour. **Journal of Food Processing and Preservation**, v. 38, n. 3, p. 737–748, 2014.
- ASSUNÇÃO, R. M. Estatística espacial com aplicações em epidemiologia, economia e sociologia. **São Carlos: Associação Brasileira de Estatística**, v. 131, 2001.
- BADDELEY, A.; GILL, R. D. Kaplan-meier estimators of distance distributions for spatial point processes. **The Annals of Statistics**, p. 263–292, 1997.
- BADDELEY, A. et al. Analysis of a three-dimensional point pattern with replication. **Applied Statistics**, p. 641–668, 1993.
- BADDELEY, A.; RUBAK, E.; TURNER, R. **Spatial point patterns: methodology and applications with R**. London: CRC Press, 2015.
- BADDELEY, A. J.; MØLLER, J.; WAAGEPETERSEN, R. Non-and semi-parametric estimation of interaction in inhomogeneous point patterns. **Statistica Neerlandica**, Wiley Online Library, v. 54, n. 3, p. 329–350, 2000.
- BAILEY, T. C.; GATRELL, A. C. **Interactive spatial data analysis**. Harlow:Longman: Longman Scientific & Technical Essex, 1995. v. 413.
- BÁRCENAS, M. E.; ROSELL, C. M. Different approaches for increasing the shelf life of partially baked bread: low temperatures and hydrocolloid addition. **Food Chemistry**, v. 100, n. 4, p. 1594–1601, 2007.
- BELL, M. L.; GRUNWALD, G. K. Mixed models for the analysis of replicated spatial point patterns. **Biostatistics**, Oxford University Press, v. 5, n. 4, p. 633–648, 2004.
- BESAG, J. Contribution to the discussion of dr. ripleys paper. **Journal of the Royal Statistical Society. Series B**, v. 39, n. 2, p. 193–195, 1977.
- BESBES, E. et al. Effect of baking conditions and storage with crust on the moisture profile, local textural properties and staling kinetics of pan bread. **LWT-Food Science and Technology**, Elsevier, v. 58, n. 2, p. 658–666, 2014.
- BICALHO, C. C. et al. Alveolar distribution of french type bread using taro mucilage. **Brazilian Journal of Food Technology**, artigo submetido, 2018.
- CASTIGLIONI, G. L. et al. Tamanho dos alvéolos e aceitação de pães de forma enriquecidos com farelo de mandioca. **Pesquisa Agropecuária Tropical (Agricultural Research in the Tropics)**, v. 2, n. 1, p. 10–1590, 2014.
- CHIU, S.; STOYAN, D. Estimators of distance distributions for spatial patterns. **Statistica Neerlandica**, v. 52, n. 2, p. 239–246, 1998.
- CRESSIE, N. **Statistics for spatial data**. Canada: John Wiley & Sons, 1993.

- CURTI, E. et al. Bread staling: effect of gluten on physico-chemical properties and molecular mobility. **LWT-Food Science and Technology**, v. 59, n. 1, p. 418–425, 2014.
- DIGGLE, P. A kernel method for smoothing point process data. **Applied statistics**, p. 138–147, 1985.
- DIGGLE, P. J. **Statistical analysis of spatial and spatio-temporal point patterns**. England, UK: CRC Press, 2013.
- DIGGLE, P. J.; LANGE, N.; BENES, F. M. Analysis of variance for replicated spatial point patterns in clinical neuroanatomy. **Journal of the American Statistical Association**, v. 86, n. 415, p. 618–625, 1991.
- DIGGLE, P. J.; MATEU, J.; CLOUGH, H. E. A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns. **Advances in Applied Probability**, v. 32, n. 02, p. 331–343, 2000.
- ECKARDT, J. et al. Long-term frozen storage of wheat bread and dough—effect of time, temperature and fibre on sensory quality, microstructure and state of water. **Journal of Cereal Science**, v. 57, n. 1, p. 125–133, 2013.
- ECKEL, S. et al. Modelling tree roots in mixed forest stands by inhomogeneous marked gibbs point processes. **Biometrical Journal**, Wiley Online Library, v. 51, n. 3, p. 522–539, 2009.
- ESTELLER, M. S.; LANNES, S. C. da S. Parâmetros complementares para fixação de identidade e qualidade de produtos panificados. **Ciênc. Tecnol. Aliment**, v. 25, n. 4, p. 802–806, 2005.
- GERVINI, D.; BAUR, T. J. Regression models for replicated marked point processes. **arXiv preprint arXiv:1705.06259**, 2017.
- GRABARNIK, P.; MYLLYMÄKI, M.; STOYAN, D. Correct testing of mark independence for marked point patterns. **Ecological Modelling**, Elsevier, v. 222, n. 23, p. 3888–3894, 2011.
- GUAN, Y. Tests for independence between marks and points of a marked point process. **Biometrics**, Wiley Online Library, v. 62, n. 1, p. 126–134, 2006.
- GUAN, Y.; AFSHARTOUS, D. R. Test for independence between marks and points of marked point processes: a subsampling approach. **Environmental and Ecological Statistics**, Springer, v. 14, n. 2, p. 101–111, 2007.
- GUTKOSKI, L. C. et al. Efeito de ingredientes na qualidade da massa de pão de forma congelada não fermentada durante o armazenamento. **Ciência e Tecnologia de Alimentos**, v. 25, n. 3, p. 460–467, 2005.
- HAHN, U. A studentized permutation test for the comparison of spatial point patterns. **Journal of the American Statistical Association**, v. 107, n. 498, p. 754–764, 2012.
- HUEN, J. et al. Confocal raman microscopy of frozen bread dough. **Journal of Cereal Science**, Elsevier, v. 60, n. 3, p. 555–560, 2014.
- ILLIAN, J. et al. **Statistical analysis and modelling of spatial point patterns**. England: John Wiley & Sons, 2008. v. 70.

- LANDAU, S.; EVERALL, I. P. Nonparametric bootstrap for k-functions arising from mixed-effects models with applications in neuropathology. **Statistica Sinica**, JSTOR, p. 1375–1393, 2008.
- LANDAU, S.; RABE-HESKETH, S.; EVERALL, I. P. Nonparametric one-way analysis of variance of replicated bivariate spatial point patterns. **Biometrical journal**, Wiley Online Library, v. 46, n. 1, p. 19–34, 2004.
- LASSOUED, N. et al. Granulometry of bread crumb grain: Contributions of 2d and 3d image analysis at different scale. **Food Research International**, Elsevier, v. 40, n. 8, p. 1087–1097, 2007.
- LIESHOUT, M. V. A j-function for marked point patterns. **Annals of the Institute of Statistical Mathematics**, Springer, v. 58, n. 2, p. 235, 2006.
- MORI, R. Tecnologia para p congelados. **Food Ingredients Brasil, Puratos**, n. 05, p. 45–46, 2008.
- MYLLYMÄKI, M.; SÄRKKÄ, A.; VEHTARI, A. Hierarchical second-order analysis of replicated spatial point patterns with non-spatial covariates. **Spatial Statistics**, Elsevier, v. 8, p. 104–121, 2014.
- PENTTINEN, A.; STOYAN, D.; HENTTONEN, H. M. Marked point processes in forest statistics. **Forest science**, v. 38, n. 4, p. 806–824, 1992.
- POLAKI, A. et al. Fiber and hydrocolloid content affect the microstructural and sensory characteristics of fresh and frozen stored bread. **Journal of Food Engineering**, v. 97, n. 1, p. 1–7, 2010.
- R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2017. Disponível em: <<https://www.R-project.org/>>.
- RAMÓN, P. et al. A new non-parametric method for analyzing replicated point patterns in ecology. **Ecography**, Wiley Online Library, v. 39, n. 11, p. 1109–1117, 2016.
- REDENBACH, C. et al. Anisotropy analysis of pressed point processes. **AStA Advances in Statistical Analysis**, Springer, v. 93, n. 3, p. 237–261, 2009.
- RIBOTTA, P. et al. Effect of emulsifier and guar gum on micro structural, rheological and baking performance of frozen bread dough. **Food Hydrocolloids**, v. 18, n. 2, p. 305–313, 2004.
- RIPLEY, B. D. The second-order analysis of stationary point processes. **Journal of applied probability**, v. 13, n. 02, p. 255–266, 1976.
- RIPLEY, B. D. Modelling spatial patterns. **Journal of the Royal Statistical Society.**, p. 172–212, 1977.
- SCANLON, M.; ZGHAL, M. Bread properties and crumb structure. **Food Research International**, Elsevier, v. 34, n. 10, p. 841–864, 2001.
- SCHABENBERGER, O.; GOTWAY, C. A. **Statistical methods for spatial data analysis**. U.S.A.: CRC press, 2004.

SCHEUER, P. M. et al. Effects of fat replacement on properties of whole wheat bread. **Brazilian Journal of Pharmaceutical Sciences**, v. 50, n. 4, p. 703–712, 2014.

SCHLADITZ, K. et al. Statistical analysis of intramembranous particles using freeze fracture specimens. **Journal of microscopy**, Wiley Online Library, v. 211, n. 2, p. 137–153, 2003.

SCHLATHER, M.; RIBEIRO, P. J.; DIGGLE, P. J. Detecting dependence between marks and locations of marked point processes. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 66, n. 1, p. 79–93, 2004.

WÄLDER, O.; STOYAN, D. On variograms in point process statistics. **Biometrical Journal**, Wiley Online Library, v. 38, n. 8, p. 895–905, 1996.

WALLER, L. A.; GOTWAY, C. A. **Applied spatial statistics for public health data**. New Jersey: John Wiley & Sons, 2004.

ZAVAREZE, E. da R. et al. Efeito do soro de leite no teor protéico e na qualidade tecnológica e sensorial de pães. **Revista do Instituto de Laticínios Cândido Tostes**, v. 63, n. 363, p. 44–50, 2008.

ZOUNIS, S. et al. Effect of final dough temperature on the microstructure of frozen bread dough. **Journal of Cereal Science**, Elsevier, v. 36, n. 2, p. 135–146, 2002.

SEGUNDA PARTE - ARTIGOS

ARTIGO 1

Alveolar distribution of French type bread using taro mucilage

**Artigo redigido conforme as normas da
revista Brazilian Journal of Food Technology
(versão submetida)**

ALVEOLAR DISTRIBUTION OF FRENCH TYPE BREAD USING TARO MUCILAGE

DISTRIBUIÇÃO ALVEOLAR DO PÃO TIPO FRANCÊS USANDO A MUCILAGEM DO TARO

ALVEOLAR DISTRIBUTION OF FRENCH TYPE BREAD

Carolina Cristina Bicalho^{1*}, Raul Antônio Viana Madeira², Joelma Pereira³, João Domingos Scalon⁴

^{1*}Corresponding Author

Carolina Cristina Bicalho,

Universidade Federal de Lavras (UFLA),

Departamento de Estatística, Campus Universitário, CEP: 37200-000, Lavras/MG – Brasil,

Telefone: 55(35)99140-0266

e-mail: carolinabicalho@gmail.com

2- Raul Antônio Viana Madeira

Departamento de Ciências de Alimentos, Universidade Federal de Lavras, Campus Universitário, CEP: 37200-000, Lavras/MG – Brasi

e-mail: raulmadeira@yahoo.com.br

3 Joelma Pereira

Departamento de Ciências de Alimentos, Universidade Federal de Lavras, Campus Universitário, CEP: 37200-000, Lavras/MG – Brasi

e-mail: joper@dca.ufla.br

4 João Domingos Scalon

Departamento de Estatística, Universidade Federal de Lavras, Campus Universitário, CEP: 37200-000, Lavras/MG – Brasil

scalon@des.ufla.br

ALVEOLAR DISTRIBUTION OF FRENCH TYPE BREAD USING TARO MUCILAGE

DISTRIBUIÇÃO ALVEOLAR DO PÃO TIPO FRANCÊS USANDO A MUCILAGEM DO TARO

ALVEOLAR DISTRIBUTION OF FRENCH TYPE BREAD

ABSTRACT

Emulsifiers are responsible for forming a mass that contributes to the retention of gluten in alveoli, the gas produced by yeast, thus increasing the volume and softness of bread. There are indications that hydrocolloids, like taro mucilage (*colocasia esculenta*), can present similar characteristic as emulsifiers. The presence of arabinose and galactose in the taro mucilage is responsible for the formation of a complex with proteins, called Proteoglycan arabinogalactan, that could present substantial emulsifying power, consequently favoring gas retention in the alveoli. Thus, the quantity, distribution and size of the alveoli must be analyzed for effective quality control in the bread making process, and this is therefore recommended. Conventional statistical methods already are regularly used in analyses of those variables, although spatial statistics have not been used for this purpose. In this paper, we defend the application of methods of univariate point processes that have not been applied to the analysis of spatial distribution of alveoli located in French bread with taro mucilage. The results showed that these methods made it possible to characterize the spatial distribution of alveoli in French bread, demonstrating that the use of taro mucilage provides a good alveolar distribution to the French type bread.

Keyword: kernel estimation, Function K, Bread making, Simulation, Frozen storage, microstructural.

RESUMO

Os emulsionantes são responsáveis pela formação de uma massa que contribui para a retenção de glúten em alvéolos, o gás produzido por levedura, aumentando assim o volume e a suavidade do pão. Há indícios de que os hidrocolóides, como a mucosa de taro (*colocasia esculenta*), podem apresentar características semelhantes

aos emulsionantes. A presença de arabinose e galactose na mucosa de taro é responsável pela formação de um complexo com proteínas, chamado Proteoglycan arabinogalactan, que poderia apresentar um poder emulsionante substancial, conseqüentemente favorecendo a retenção de gás nos alvéolos. Assim, a quantidade, a distribuição e o tamanho dos alvéolos devem ser analisados para um controle de qualidade efetivo no processo de fabricação de pão e, portanto, é recomendado. Métodos estatísticos convencionais já são usados regularmente nas análises dessas variáveis, embora as estatísticas espaciais não tenham sido utilizadas para esse fim. Neste artigo, defendemos a aplicação de métodos de processos pontuais univariados que não tenham sido aplicados na análise de distribuição espacial de alvéolos localizados em pão francês com mucilagem de taro. Os resultados mostraram que esses métodos possibilitaram caracterizar a distribuição espacial dos alvéolos no pão francês, demonstrando que o uso da mucilagem de taro fornece uma boa distribuição alveolar ao pão tipo francês.

Palavra-chave: estimativa do kernel, função K, fabricação de pão, simulação, armazenamento congelado, microestrutural.

1. INTRODUCTION

In bread making, various emulsifying agents are used in the mass in order to reduce humidity loss, increase softness in the product and, consequently, slow aging. Thus, many additions are added to bread in order to improve the potential of their sensory characteristics, such as hydrocolloids (BÁRCENAS; ROSELL, 2007), enzymes and emulsifiers (GOESAERT et al., 2005; GOMES-RUFFI et al., 2012).

Emulsifiers are capable of forming films in the bread mass interface between gluten and starch, increasing the ability of the mass to retain the gas produced by yeasts and, as a consequence, the volume and softness of the bread (GOMES-RUFFI et al., 2012; KROG, 1981).

Some Hydrocolloids have an effect similar to an emulsifier (BÁRCENAS; ROSELL 2007; RIBOTTA et al., 2004), and taro mucilage is an example of a hydrocolloid with emulsifying power (NAGATA et al., 2015). Mucilage is a gummy substance found in vegetables and is usually present in tropical roots and tubers; it is particularly abundant in taro (HUANG et al., 2010; NAGATA et al., 2015).

The emulsifying property of the mucilage results mainly from the protein content, with the presence of amino acids weakly polar, in particular leucine, isoleucine and tryptophan, and the presence of lipids possibly providing the lipophilic fraction (HUANG et al., 2010; NJINTANG et al., 2014). Galactose and arabinose are the main monosaccharides present in taro mucilage (HUANG et al., 2010; JIANG; RAMSDEN, 1999; NGUIMBOU et al., 2014), and these are capable of forming a complex called Arabinogalactan that, as a consequence of the high percentage of protein, forms a proteoglycan arabinogalactan (AGP) with emulsifying power (JIANG; RAMSDEN, 1999).

The addition of emulsifiers contributes to increasing mass volume, which positively affects the softness of the bread crumbs. In addition, bread crumb characteristics, alveolar distribution and texture in the sensory analysis were considered better than bread control (GOMES-RUFFI et al., 2012). Thus, the distribution and shape of the alveoli of the bread, in a general way, is important to the product sensory qualities.

Good quality bread must have the proper porosity, i.e., oval, uniform cells and no holes. Changes in these parameters demonstrate formulation or process errors such as: excess improvers, improper mixing, very hard mass, poor molding, excessive fermentation and inadequate heat distribution in the furnace. In addition, a very closed and compact crumb makes the bread dark and dull, (FERREIRA et al., 2001). In this case, baked and later frozen

bread was softer than frozen breads. This characteristic would be due to the fact that pasta has large, uniform alveoli that are well distributed in the bread crumb, in addition to having factors such as water mobility, storage temperature and formulation, which influence the structure of the bread crumb (ECKARDT et al., 2013). It is believed that the use of freeze-dried taro mucilage in the bread mass entails a better distribution of the alveoli, which improves the mass texture and, consequently, market acceptance of the product.

Thus, the aim of the present work is to present methods to verify if the alveoli are uniformly well distributed in the bread crumbs with lyophilized taro mucilage; that is, the hypothesis that the alveoli present in the bread crumb present complete spatial randomness (CSR) distribution is tested. Several statistical methods have been proposed, including testing the null hypothesis of CSR in point configurations, e.g., methods based on event counting by quadrant, distances between nearer neighbor events and descriptor functions as shown by (CRESSIE, 1993; DIGGLE, 2013). In this work, we defend the use of methods to characterize the first-order and second-order properties of point processes in space. Analysis of the first-order properties is performed by using a kernel estimator of local intensity (DIGGLE, 2013). For the analysis of second-order properties, the descriptor F-, G-, and K-functions are used in conjunction with Monte Carlo simulation (DIGGLE, 2013). These methods are expected to characterize whether the lyophilized taro mucilage bread sockets are distributed in the mass in a regular, random or agglomerated distribution. It has been observed that, until now, studies with these characteristics have not been found in the literature.

2. MATERIALS AND METHODS

2.1 Description of experimental data or material or raw material

French bread types of 50 g were prepared according to the following formulation: 40 g of fresh biological yeast; 20 g of salt and 564 g of potable water at 10 °C based on 1 kg of wheat flour according to pharynography; wheat flour contained moisture of 13.8 g 100 g⁻¹; 10.62 g 100 g⁻¹ protein; ash content of 0.45 g 100 g⁻¹; falling number of 369 s, stability greater than 20 minutes, water absorption of 56.4 g 100⁻¹; 26.4 g 100⁻¹ wet gluten and 9.4 g 100⁻¹ dry gluten, plus the addition of hydrogenated fat (15.8 g kg⁻¹) and lyophilized taro mucilage (7.3 g kg⁻¹) following the recommendation by (NAGATA et al., 2015).

The mucilage of the Taro (*colocasia esculenta*) was extracted according to recommendations made by (Andrade et al., 2015). The centesimal composition of the lyophilized taro mucilage is: 0.28 g 100 g⁻¹ ash; 0.638 g 100 g⁻¹ ethereal extract; 8.038 g 100 g⁻¹ of protein; 0.358 g 100 100 g⁻¹ fiber; 6.798 g 100 g⁻¹ of moisture and 84.128 g 100 g⁻¹ of non-nitrogenous extract.

We applied statistical analysis on three samples of French bread (S1-S3) produced at the Food Department of the Federal University of Lavras, Brazil. For each bread sample, a cross section cut was made for the production of 3 slices. The middle slice was scanned into an HP deskjet 3040 at 200 dpi color scanner. The images obtained were treated by the ImageJ software using the following macro: enhance contrast 0.4%, smoth, adjust color threshold set in black and white. The samples were then analyzed using Image-Pro Plus 4.5 image software, Media Cybernetics, Inc., USA. These images were obtained in pixels, to be transformed in (cm²), and the conversion was done using a rectangular image of 8 cm² that has 53735 pixels.

Alveolar parameters were determined: mean number of alveoli, alveolar area (cm²), mean diameter of the alveoli (mm) and mass area (cm²). For the application of point process methods, each alveolus of the bread is treated as a point defined by its coordinates. So we have a map, in a flat region, with the location of all the alveoli present in the slice of bread.

Statistical analyses were performed in R environment (R CORE TEAM, 2016) using the spatstat package (BADDELEY; TURNER, 2005).

2.2 Kernel Estimator

Let (x_1, x_2, \dots, x_n) be the localizations (coordinates) of n alveoli in a slice of bread, considered flat, with area $|A|$. According to Diggle (2013), the estimator of the local intensity of alveolus $\lambda(x)$ in is given by the Equation 1.

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

where $\tau > 0$ is the parameter that determines the bandwidth (smoothing amount) and $k(\cdot)$ is a bivariate probability density function, known as the kernel, which is symmetric in relation to the origin. A quartic function can be used as an alternative to $k(\cdot)$, and then the intensity estimator $\hat{\lambda}(x)$ can be expressed as

$$\hat{\lambda}_\tau(x) = \sum_{i=1}^n \frac{3}{\pi\tau^2} \left(1 - \frac{h_i^2}{\tau^2}\right)^2, \quad (2)$$

where h_i is the distance between the point and the location of the observed alveolus (Equation 2).

2.3 K Function

The K function was proposed by Ripley (1977) as a tool to evaluate the second-order properties of a point process. The theoretical K function of a stationary point process is defined by $K(t) = \lambda^{-1}E$ (number of new events within a distance t from an arbitrary event). Under the CSR hypothesis, the function K is given by $K(t) = \pi t^2$. An estimator for $K(t)$ proposed by (Ripley, 1977) that corrects edge effects is given by

$$\hat{K}(t) = \frac{|A|}{n^2} \sum_{i=1}^n \sum_{j=1}^n \frac{I_t(u_{ij})}{w_{ij}}. \quad (3)$$

In Equation 3, u_{ij} is the distance between the alveoli i and j ($i \neq j$), $I_t(u_{ij})$ is an indicator function that assumes the value 1 when u_{ij} is less than the distance t , n is the number of alveoli in the bread slice (Area $|A|$), e w_{ij} is a weighting factor that represents the ratio of the circumference of the circle around the alveolus i , passing through the alveolus j , which is found inside the area of the bread slice.

One way to analyze the estimates of the K function of an observed point congruence is to use an auxiliary function L proposed by Besag (1977), which linearizes the K -function, stabilizes the variance and, consequently, facilitates its graphic interpretation. One of the estimators of the L function is given by: $\hat{L}(t) = \sqrt{\hat{K}(t)/\pi}$.

2.4 The F and G Functions

The function F was proposed by Ripley (1977) and is based on the distance x , that is, the distance between an arbitrary point (not the location of an alveolus) and the nearest alveolus. The function F is denoted by the distribution function $F(x) = P\{$ (the distance from

an arbitrary point to the nearest event)). Under the null hypothesis of CSR, the theoretical F function is given by:

$$F(x) = 1 - \exp(-\lambda\pi x^2), \quad x > 0. \quad (4)$$

The simplest estimator of the F function with corrected edge effect is given by equation:

$$\hat{F}(x) = \frac{\sum_{i=1}^n I_x(x_i, r_i)}{\sum_{i=1}^n I_x(r_i)}, \quad x > 0, \quad (5)$$

where m is the number of alveoli in the bread slice, x_i denotes the distances of the i th arbitrary point (between m available points) to the nearest alveolus, between the n alveoli observed, $I_x(r_i)$ is an indicator function that assumes a value of 1 when $x_i \leq x$, r_i is the distance from each alveolus to the point nearest to the edge of $|A|$ and $I_x(x_i, r_i)$ is an indicator function that assumes the value 1 when $x_i \leq x$ e $r_i \geq x$. To choose arbitrary m points, one can use a regular grid $k \times k$, where $k \cong \sqrt{n}$.

The function G is based on the distance y , that is, the distance between two nearest alveoli. It is given by the expression $G(y) = P\{\text{distance from an arbitrary event to another nearest event at most } y\}$. The theoretical G function under the null hypothesis of CSR and the corrected edge estimator of $G(y)$ are given by equations analogous to Equation 4 and Equation 5, respectively, by substituting Distance x by distance y .

2.5 Formal method for testing against hypothesis of Complete Spatial Randomness

An approach based on Monte Carlo simulations to perform a test against the null hypothesis of CSR is proposed by Diggle (2013). This approach compares the observed

distribution function and the expected distribution function under the CSR hypothesis. For the F function, the test statistic is given by:

$$FI_i = \int_0^{x_0} \{\hat{F}_i(x) - [1 - \exp(-\lambda\pi x^2)]\}^2 dx; \quad i = 2, \dots, s. \quad (6)$$

The FI statistic has no known sample distribution, and thus this distribution can be obtained by Monte Carlo simulation. Let $\hat{F}_1(x)$ be the function F of a punctual observed configuration with n events, as defined in Equation 6 and $\hat{F}_2(x), \dots, \hat{F}_s(x)$ the F functions calculated from the s Monte Carlo simulations under the CSR hypothesis, each one with n events.

The statistics of FI_i are calculated for the configuration observed and simulated. The value of FI_i is compared with the values $\hat{FI}_2(x), \dots, \hat{FI}_s(x)$ obtained from the simulated configurations. If FI_1 is between the highest values of $\hat{FI}_2(x), \dots, \hat{FI}_s(x)$, then it is indicated that configuration does not have CSR. Assume $F_1 = F_{(j)}$ for some $j \in \{1, \dots, s\}$, and then reject the CSR hypothesis if $P_{\text{value}} = ((s + 1 - j)/2) \leq \alpha$.

This procedure can be conducted anomalously for functions K and G .

2.6 Graphical method for testing against the hypothesis of complete spatial randomness

The formal test presented above does not indicate the direction (regularity or grouping alveoli) of the rejection of the null hypothesis. The graphical method, presented below, fills this gap.

Let $\hat{F}_1(x)$ be a function F of a pattern of observed points with n events. By simulation, $s - 1$ sets of n events are generated assuming CSR. For these $s - 1$ simulated data sets, the upper and lower simulation envelopes, respectively, are calculated as:

$$U(x) = \max \hat{F}_i(x), \quad i = 2, \dots, s$$

e

$$L(x) = \min \hat{F}_i(x), \quad i = 2, \dots, s.$$

Constructing a graph of $\hat{F}_1(x)$, $U(x)$, $L(x)$ (ordinates) against the distances x . If the curve represented by $\hat{F}_1(x)$ lies inside the simulated envelopes, the CSR hypothesis is not rejected. If $\hat{F}_1(x)$ falls below the smallest envelope of simulation, for any distance, there are indications that the alveoli present a distribution that is typical of a pattern that exhibits clusters. Otherwise, if $\hat{F}_1(x)$ lies above the simulation envelope at the top, for any distance, there is evidence that the alveoli exhibit a regular pattern.

It is possible to construct the simulation envelopes for the functions G and K in a manner analogous to that used for the function F ; however, the interpretations that detect regularity and grouping are performed in a way that is inverse to the function F .

3. RESULTS AND DISCUSSIONS

The first requirement in any data analysis is to observe the data that are to be analyzed. The visualization of the spatial pattern means mapping the alveoli in the bread slice. The plotting of the location of each alveolus as a point map is a natural tool for visualizing the spatial distribution. This map provides an idea of alveolar distribution and may

show some obvious pattern of alveolar distribution in samples. The point maps of the samples (S1, S2, S3) are shown in Figure 1.

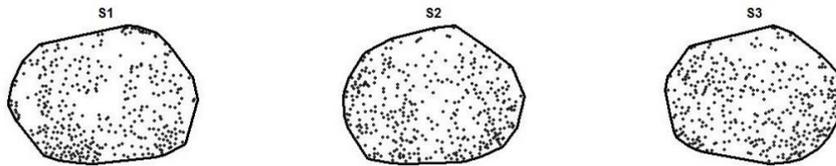


Figure 1 Localization map of the alveoli of the French bread slice of the samples (S1, S2, S3).

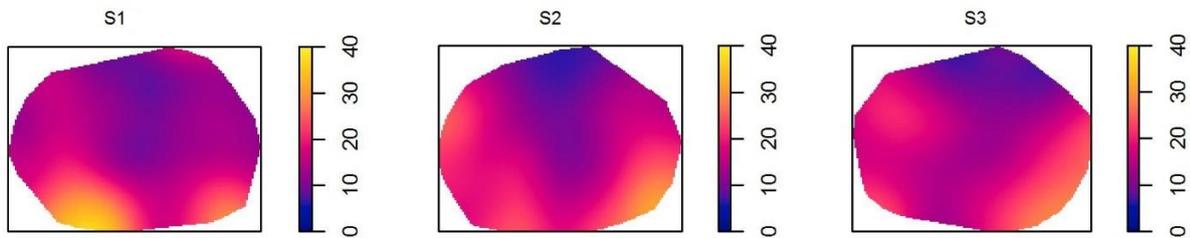
The visual inspection in Figure 1 shows that it is very difficult to assess the nature of the pattern of points based on this information, in other words, whether there is a tendency for clustering or regularity of the alveoli in the slices of bread.

Table 1 shows the overall intensity λ for the three samples, indicating the presence of approximately 16 alveoli on average per cm^2 of bread slice. However, Figure 1 seems to show us that using only the global mean to characterize the distribution of alveoli on bread seems to lead to an inaccurate interpretation. A more advanced statistical tool for analyzing the expected number of alveoli per unit area is the use of an estimate of the local alveoli intensity using, for example, the kernel intensity estimator (section 2.2).

Table1: Summarized data from the samples.

Sample	Number de alveoli	Area (cm^2)	Intensity (λ)
S1	339	21.04	16.111
S2	348	21.17	16.435
S3	346	21.24	16.286

Estimating local intensity, the most important is to choose the appropriate bandwidth (τ). If the bandwidth tends to zero, the intensity is estimated as a series of peaks in individual



observations, while if the bandwidth becomes large, all details will be obfuscated. (Diggle, 2013) suggests adopting $\tau = 0.68n^{-0.2}$ for unitary areas. As the bread slice areas do not have unit areas, we adopt a value that minimizes the likelihood function of the point process obtained by cross-validation, as suggested by Loader (2006). Diggle (2013) states that choice of kernel function has secondary importance since any kernel function (e.g., Gaussian, quartic) leads to the same results. In this work, we use the Quartic kernel (Equation 2). The maps smoothed by the kernel intensity estimator $\hat{\lambda}_\tau$ are presented in Figure 2.

Figure 2 Estimates of local intensity using the quartic kernel estimator.

The analysis of the local intensity is important since the alveolar distribution in the bread crumbs tells us a lot about the action of the emulsifiers. For example, works by Scheuer et al. (2014) and Polaki et al. (2010) show that when the fat was replaced by some emulsifying agent, in this case enzymatically modified corn starch and polydextrose, respectively, the alveolar intensity was larger and more homogeneous, in addition to leading to smaller alveolar sizes.

It is possible to observe in the samples of Figure 2 not only areas with a presence of local concentrations of alveoli, in which the local intensity of alveoli is much larger than the

global intensity, but also areas with small local concentrations of alveoli, where local intensity is much lower than the overall mean of alveoli. The upper parts of bread slices appear to have a tendency to have a lower alveoli intensity, so there is evidence that the distribution of alveoli along the bread slice may not be random.

However, it is not recommended to draw a conclusion on the distribution of alveoli in breads by analyzing only the intensities (first-order properties). It is also recommended to analyze possible interactions between alveoli (second-order properties). For this, formal tests were performed against the hypothesis of complete spatial randomness using the functions G , G and K ($s = 100$); the results are presented in Table 2.

Table 2: P-values of the tests against the hypothesis of complete spatial randomness based on the integral of the quadratic difference of the functions F , G and K .

Test	Sample		
	S1	S2	S3
F-function	0.35	0.48	0.25
G-function	0.01	0.01	0.01
K-function	0.01	0.01	0.01

The p-values were attached based 99 on Monte Carlo simulations ($s = 100$) of a homogeneous Poisson process.

The results presented in Table 2 show that the tests based on the G and K functions lead to an emphatic rejection of the CSR hypothesis in all the samples ($P < 0.05$), while the F -function suggests the acceptance of CSR in all the samples. The G -function is able to more powerfully detect regularity, while the F -function is able to more powerfully detect clusters of events (Diggle, 2013). Thus, it is probable that the distribution of the alveoli in the bread dough can present regularity. In order to be sure of the nature of the alveoli distribution, the

simulation envelopes (section 2.2) must be adopted for the F, G and K functions. The envelopes are shown in Figures 3, 4 and 5.

We can see that the empirical F function is practically all of the simulation envelopes, while the empirical functions G and K are below the simulation envelopes, indicating a rejection of the hypothesis of complete spatial randomness in the direction of regularity.

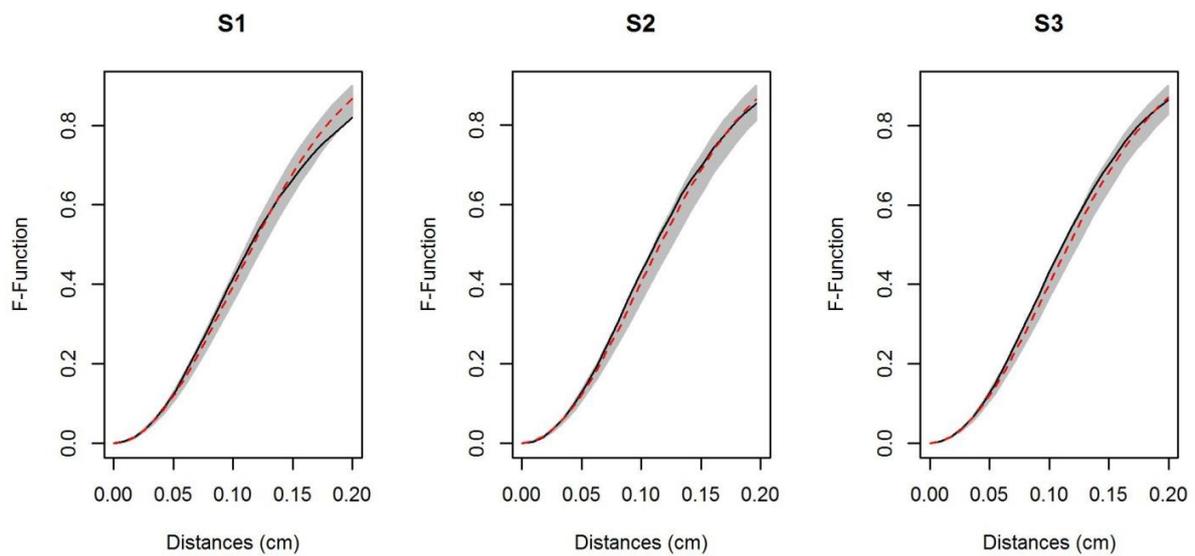


Figure 3 Envelopes with 99 simulations (gray color) against the CSR hypothesis of a Poisson process (red dotted lines) for the empirical F (solid line) functions in samples S1, S2, and S3.

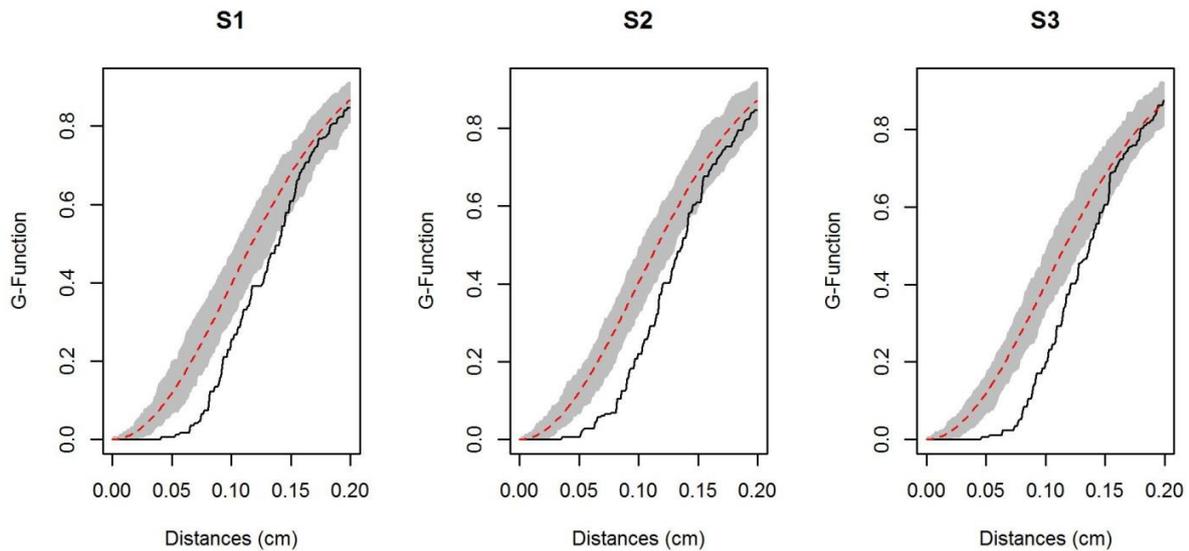


Figure 4 Envelopes with 99 simulations (gray color) against the CSR hypothesis of a Poisson process (red dotted lines) for the empirical G functions (solid line) in samples S1, S2, and S3.

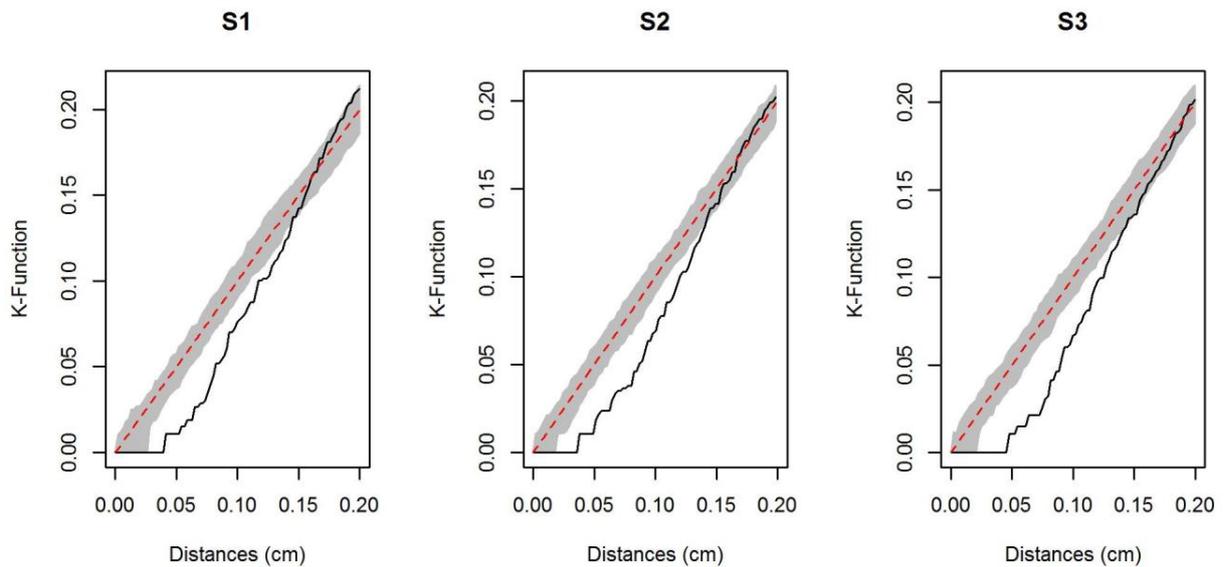


Figure 5 Envelopes with 99 simulations (gray color) against the CSR hypothesis of a Poisson process (red dotted lines) for the empirical L functions (solid line) in samples S1, S2, and S3.

An interesting feature observed in the envelopes of the G function is absence of neighboring alveoli for small distances up to 0.05 cm. This inhibitory mechanism can be explained by the fact that in the analysis, only the positions (centers) of the alveoli were considered, the areas of the alveoli being ignored. Thus, there will always be a minimum distance corresponding to the sizes of the alveoli ray that were not considered in the analyses. The same explanation can be given for the L function.

The results above show the importance of using the three functions for analysis of the interaction between the alveoli since each of these functions has greater power in the detection of certain characteristics of the standard, and therefore the use of all functions leads to a more sensitive analysis than when used alone. For example, analysis based on the G function is more sensitive to regularity, unlike analysis with the F function from CSR, which is more sensitive to clustering detection. This behavior explains the differences observed between the results provided by the F and G functions. The K function is considered equally powerful in terms of the ability to detect configurations with both clusters and regularity (Diggle, 2013).

The presented analysis is valuable in suggesting the spatial structure that may be present in the data, as the use of emulsifiers in baking products contributes to the improvement of the structure of the kernel with smaller and nearer alveoli, brighter kernels and greater uniformity in the Size of cells (Stampfli; Nersten, 1995); adding the formulation of the breads together with the fats has the function of retaining the air in the mass during mixing and roasting. The grease crystals absorbed by the air-water interface melt and allow the water-fat interface to expand into a bubble. The size of these air bubbles is related to the different amounts of air incorporated in the mass. The presence of a large number of

adsorbed crystals allows the bubbles to expand without rupture, and this allows the formation of a kernel with different textures (Polaki et al., 2010).

In fact, the previous analysis indicated wide regularity on a small scale. Thus, it is believed that the use of taro mucilage provides a mass with regularly distributed alveoli, as the complexes formed between carbohydrates and proteins present in the mucilage act as tensoactive compounds in the aqueous phase that permeates the alveolar wall, allowing the formation of an interface between the phases and producing a larger number of small alveoli. According to (Polaki et al., 2010), the pore distribution reveals important information about the bread structure. In their research, they classified the pores into small (area < 4 mm²), medium (area < 4 - 8 mm²) and large (area > 8 mm²) and associated them with sensory characteristics; in other words, the size of the alveoli is directly connected to the quality of the mass and thus the taste of the bread. In this study, the size of the alveoli was not used in the analysis (the marked dot pattern), which may lead to a comprehensive conclusion regarding the data.

4. CONCLUSION

This paper puts forward a comprehensive analysis of the spatial distribution of the breadcrumb alveoli from the mass with an emulsifier provided by the taro. The analysis clearly showed that the functional pattern of the descriptors described different properties of spatial point patterns. The F, G, and K functions are powerful tools for summarizing essential pattern properties and therefore ideal for testing against CRS. The results based on three images of bread slices showed a deviation that was statistically significant in CRS toward regularity. Thus, it is believed that the use of the taro mucilage provides a mass with regularly distributed

alveoli indicating an emulsifying effect that will be beneficial to the texture and consequently the quality of the bread.

5. ACKNOWLEDGEMENTS

This work was conducted during a scholarship supported by the International Cooperation Program CAPES at the University of Federal of Lavras. Financed by CAPES - Brazilian Federal Agency for Support and Evaluation of Graduate Education within the Ministry of Education of Brazil.

REFERENCES

- ANDRADE, Luan Alberto; NUNES, Cleiton Antônio; PEREIRA, Joelma. Relationship between the chemical components of taro rhizome mucilage and its emulsifying property. **Food chemistry**, v. 178, p. 331-338, 2015.
- BADDELEY, Adrian et al. Spatstat: an R package for analyzing spatial point patterns. **Journal of statistical software**, v. 12, n. 6, p. 1-42, 2005.
- BÁRCENAS, María Eugenia; ROSELL, Cristina M. Different approaches for increasing the shelf life of partially baked bread: Low temperatures and hydrocolloid addition. **Food Chemistry**, v. 100, n. 4, p. 1594-1601, 2007.
- BESAG, J. E. Comments on Ripley's paper: Royal Statistical Society. **Journal**, v. 39, p. 193-195, 1977.
- CRESSIE, Noel; WIKLE, Christopher K. **Statistics for spatio-temporal data**. John Wiley & Sons, 1993.
- DIGGLE, Peter J. **Statistical analysis of spatial and spatio-temporal point patterns**. CRC Press, 2013.
- ECKARDT, J., ÖHGREN, C., ALP, A., EKMAN, S., ÅSTRÖM, A., CHEN, G., SWENSON, J., JOHANSSON, D., LANGTON, M. . Long-term frozen storage of wheat bread and dough— Effect of time, temperature and fibre on sensory quality, microstructure and state of water. **Journal of cereal science**, v. 57, n. 1, p. 125-133, 2013.
- FERREIRA, Sila Mary Rodrigues; DE OLIVEIRA, PATRÍCIA V.; PRETTO, Daniela. Parâmetros de qualidade do pão francês. **Boletim do Centro de Pesquisa de Processamento de Alimentos**, v. 19, n. 2, 2001.

- GOESAERT, H., BRIJS, K., VERAVERBEKE, W.S., COURTIN, C.M., GEBRUERS, K., DELCOUR, J.A. Wheat flour constituents: how they impact bread quality, and how to impact their functionality. **Trends in food science & technology**, v. 16, n. 1, p. 12-30, 2005.
- GOMES-RUFFI, C.R., DA CUNHA, R.H., ALMEIDA, E.L., CHANG, Y.K. & STEEL, C.J., Effect of the emulsifier sodium stearyl lactylate and of the enzyme maltogenic amylase on the quality of pan bread during storage. **LWT-Food Science and Technology**, v. 49, n. 1, p. 96-101, 2012..
- HUANG, C.C., LAI, P., CHEN, I.H., LIU, Y.F. & WANG, C.C.. HUANG, Chien-Chun et al. Effects of mucilage on the thermal and pasting properties of yam, taro, and sweet potato starches. **LWT-Food Science and Technology**, v. 43, n. 6, p. 849-855, 2010.
- JIANG, Gaosong; RAMSDEN, Lawrence. Characterisation and yield of the arabinogalactan–protein mucilage of taro corms. **Journal of the Science of Food and Agriculture**, v. 79, n. 5, p. 671-674, 1999.
- KROG, N. Theoretical aspects of surfactants in relation to their use in breadmaking. **Cereal Chemistry**, 1981.
- LOADER, Clive. **Local regression and likelihood**. Springer Science & Business Media, 2006.
- NAGATA, Carina Lumie Pereira; ANDRADE, Luan Alberto; PEREIRA, Joelma. Optimization of taro mucilage and fat levels in sliced breads. **Journal of food science and technology**, v. 52, n. 9, p. 5890-5897, 2015.
- NGUIMBOU, Richard Marcel et al. Mucilage chemical profile and antioxidant properties of giant swamp taro tubers. **Journal of food science and technology**, v. 51, n. 12, p. 3559-3567, 2014.
- NJINTANG, Nicolas Yanou et al. Compositional, spectroscopic and rheological analyses of mucilage isolated from taro (*Colocasia esculenta* L. Schott) corms. **Journal of food science and technology**, v. 51, n. 5, p. 900-907, 2014.
- POLAKI, A. et al. Fiber and hydrocolloid content affect the microstructural and sensory characteristics of fresh and frozen stored bread. **Journal of Food Engineering**, v. 97, n. 1, p. 1-7, 2010.
- R CORE TEAM: **A Language and Environment for Statistical Computing**. Vienna, Austria: 2016, <https://www.R-project.org/>.
- RIBOTTA, P. D. et al. Effect of emulsifier and guar gum on micro structural, rheological and baking performance of frozen bread dough. **Food Hydrocolloids**, v. 18, n. 2, p. 305-313, 2004.
- RIPLEY, Brian D. Modelling spatial patterns. **Journal of the Royal Statistical Society. Series B (Methodological)**, p. 172-212, 1977.
- SCHEUER, Patrícia Matos et al. Effects of fat replacement on properties of whole wheat bread. **Brazilian Journal of Pharmaceutical Sciences**, v. 50, n. 4, p. 703-712, 2014.

STAMPFLI, L.; NERSTEN, B. Emulsifiers in bread making. **Food Chemistry**, v. 52, n. 4, p. 353-360, 1995.

ARTIGO 2**O EFEITO DO CONGELAMENTO NA DISTRIBUIÇÃO ESPACIAL DOS
ALVÉOLOS NA MASSA DO PÃO: UMA ANÁLISE BASEADA EM PADRÕES DE
PONTOS ESPACIAL REPLICADOS**

**Artigo será traduzido e submetido à
revista *Brazilian Journal of Food Technology*
(versão preliminar)**

O EFEITO DO CONGELAMENTO NA DISTRIBUIÇÃO ESPACIAL DOS ALVÉOLOS NA MASSA DO PÃO: UMA ANÁLISE BASEADA EM PADRÕES DE PONTOS ESPACIAL REPLICADOS

RESUMO

A qualidade dos pães produzidos a partir de massa congelada é afetada, principalmente, por alterações provocadas pelos cristais de gelo. Pretende-se estudar o efeito do congelamento por meio da distribuição dos alvéolos do miolo do pão francês em função do tempo de armazenamento. No presente estudo, o interesse está em determinar se o tempo de congelamento altera a distribuição dos alvéolos na massa do pão. Foi utilizada uma amostra de 30 pães divididos em 5 grupos, com tempos de congelamento iguais a 0, 20, 40, 60 e 80 dias. Assim, cada grupo foi composto por 6 pães. Após o tempo de congelamento, os pães foram assados e uma fatia central foi retirada para que os alvéolos fossem identificados e tivessem suas coordenadas de localização determinadas. Com isso, obteve-se 5 grupos com 6 de configurações pontuais por grupo. Para estudar esses dados utilizou-se a análise de variância não paramétrica para processos pontuais replicados não marcados. Os resultados mostraram que a distribuição alveolar em todos os tempos de congelamento é regular, ou seja, existe uma distância mínima entre os alvéolos. Além disso, os resultados também mostraram que existem diferenças estatisticamente significativas nessa distribuição regular dos alvéolos entre os tempos de congelamento. Pode-se inferir que o tempo de congelamento para uma massa de boa qualidade é de 20 dias. Este tempo de congelamento foi o que apresentou a distribuição alveolar mais próxima da completa aleatoriedade espacial.

Palavras-chave: Panificação; Distribuição espacial de alvéolos; Configurações pontuais; Função K de Ripley; Análise de variância.

1 INTRODUÇÃO

O congelamento é um dos melhores métodos de preservação da qualidade dos alimentos (GUTKOSKI et al., 2005). Atendendo as necessidades de regiões distantes e mantendo a qualidade os alimentos.

De acordo com Polaki et al. (2010), a distribuição alveolar revela informações importantes sobre a estrutura do pão.

A análise de imagens, em geral, é utilizada como uma ferramenta para a observar estruturas espumosas, como a do pão, que pode gerar um grande número de informações sobre a estrutura da massa do pão, tais como quantidade, localização, área, diâmetro, volume dos alvéolos, dentre outros. A análise de grandes quantidades de dados somente pode ser realizada por meio de métodos estatísticos. Neste sentido, pesquisadores têm recorrido, em geral, à métodos usuais clássicos tais como: análise de variância com teste de Tukey (SCHEUER et al., 2014); teste de Fisher, análise de componentes principais, correlação (POLAKI et al., 2010; BESBES et al., 2014) entre outros.

A possibilidade de usar explicitamente a localização dos alvéolos na análise de dados abre novas possibilidades para o entendimento da influência do congelamento na distribuição dos alvéolos na massa do pão. Por exemplo, qual é o comportamento da distribuição espacial dos alvéolos na massa após o congelamento? Esta questão foi abordada pelo trabalho pioneiro de Bicalho et al. (2018) que usaram métodos baseados na teoria de processos pontuais, em amostras individuais de pão, para descobrir que a distribuição dos alvéolos na massa tende a apresentar um padrão regular, ou seja, os alvéolos tendem a localizarem-se mais afastados, uns dos outros, do que seria esperado para uma distribuição completamente aleatória.

A facilidade de se obter múltiplas amostras de pão sob as mesmas condições experimentais (replicação) permite estender os métodos estatísticos utilizados por Bicalho et al. (2018) para amostras replicadas e, conseqüentemente, fornecer respostas mais precisas para questões chaves, como por exemplo: (a) existe diferença na distribuição espacial dos alvéolos entre os diferentes tempos de congelamento? e (b) qual o tempo de congelamento ótimo que propicia a melhor distribuição espacial dos alvéolos (Completa Aleatoriedade Espacial) na massa de pão?

Assim, o objetivo deste trabalho é explorar a abordagem de métodos estatísticos baseados em processos pontuais replicados para analisar a distribuição dos alvéolos, além de verificar se existe diferença na distribuição espacial dos alvéolos em tempos de congelamento diferentes.

2 MATERIAL E MÉTODOS

2.1 Preparação dos pães francêss e obtenção dos alvéolos

A preparação dos pães francêss e obtenção dos dados dos alvéolos foram realizadas no Departamento de Ciências de Alimentos, da Universidade Federal de Lavras, Brasil, seguindo as etapas apresentadas a seguir.

A partir da massa do pão francês foram produzidos 60 pães, utilizando a farinha com as seguintes características: farinha de trigo continha umidade de $13,8\text{g } 100\text{g}^{-1}$; $10,62\text{g } 100\text{g}^{-1}$ de proteína; teor de cinzas de $0,45\text{ g } 100\text{g}^{-1}$; *falling number* de 369 segundos, estabilidade maior que 20 minutos, absorção de água de $56,4\text{g}100^{-1}$; $26,4\text{g}100^{-1}$ de glúten úmido e $9,4\text{g}100^{-1}$ de glúten seco.

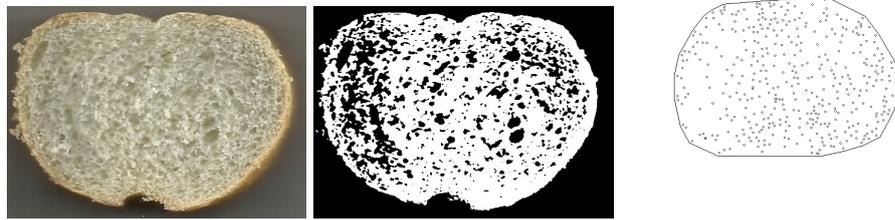
Os pães foram divididos e congelados por 0, 20, 40, 60 e 80 dias em *frezzer* doméstico a -18°C . No final de cada período de congelamento, foram escolhidos aleatoriamente 6 pães que foram assados em forno elétrico industrial a uma temperatura de 190°C durante quinze minutos, o que caracteriza um delineamento inteiramente casualizado (DIC), contendo cinco tratamentos e seis repetições.

Em cada pão francês foram feitos cortes transversais para produção de três fatias. A fatia do meio, unidade experimental, foi escolhida e escaneada em *scanner HP deskjet 3040* a 200 dpi colorido. As imagens obtidas foram tratadas pelo software ImageJ usando a seguinte macro: *enhance contrast 0.4%*, *smoth*, *adjust color threshold* ajustado em *black & white*.

Para cada tempo de congelamento foram obtidas seis imagens das fatias de pão francês que foram processadas utilizando o *software* Image-Pro Plus 4.5, Media Cybernetics, Inc., USA. Essas imagens foram obtidas em pixels e posteriormente convertidas para cm^2 . Foram determinados a localização e o diâmetro dos alvéolos, a área de cada fatia (cm^2) e a intensidade de alvéolos por cm^2 (veja Tabela 3.1).

Para a aplicação dos métodos de processos pontuais, cada evento a ser estudado é um alvéolo do pão definido pelas suas coordenadas. Assim tem-se um mapa em uma região plana, com a localização de todos os alvéolos presentes na fatia do pão conforme mostra a Figura 2.1.

Figura 2.1 – Exemplo típico da posição dos alvéolos em uma fatia de pão francês



Fonte: Do autor (2017).

2.2. Configurações pontuais espaciais replicadas

Os modelos matemáticos que descrevem o arranjo de elementos aleatoriamente (ou irregularmente) distribuídos no espaço são chamados de processos pontuais espaciais. A realização desse processo é denominada de configuração pontual. Os eventos da configuração pontual são representados por coordenadas de pontos na dimensão apropriada (ILLIAN et al., 2008). Nesse estudo, os eventos são as localizações dos alvéolos na fatia do pão.

Configurações pontuais replicadas ocorrem quando é possível obter diversas configurações pontuais sob a mesma condição experimental. No presente trabalho, cada grupo dos 5 tempos de congelamento apresenta 6 réplicas de configurações pontuais. Na Figura 2.2 apresenta todas as posições dos alvéolos das fatias dos pães. Foram estudados aproximadamente 14.000 alvéolos em uma área total de 558,94 cm².

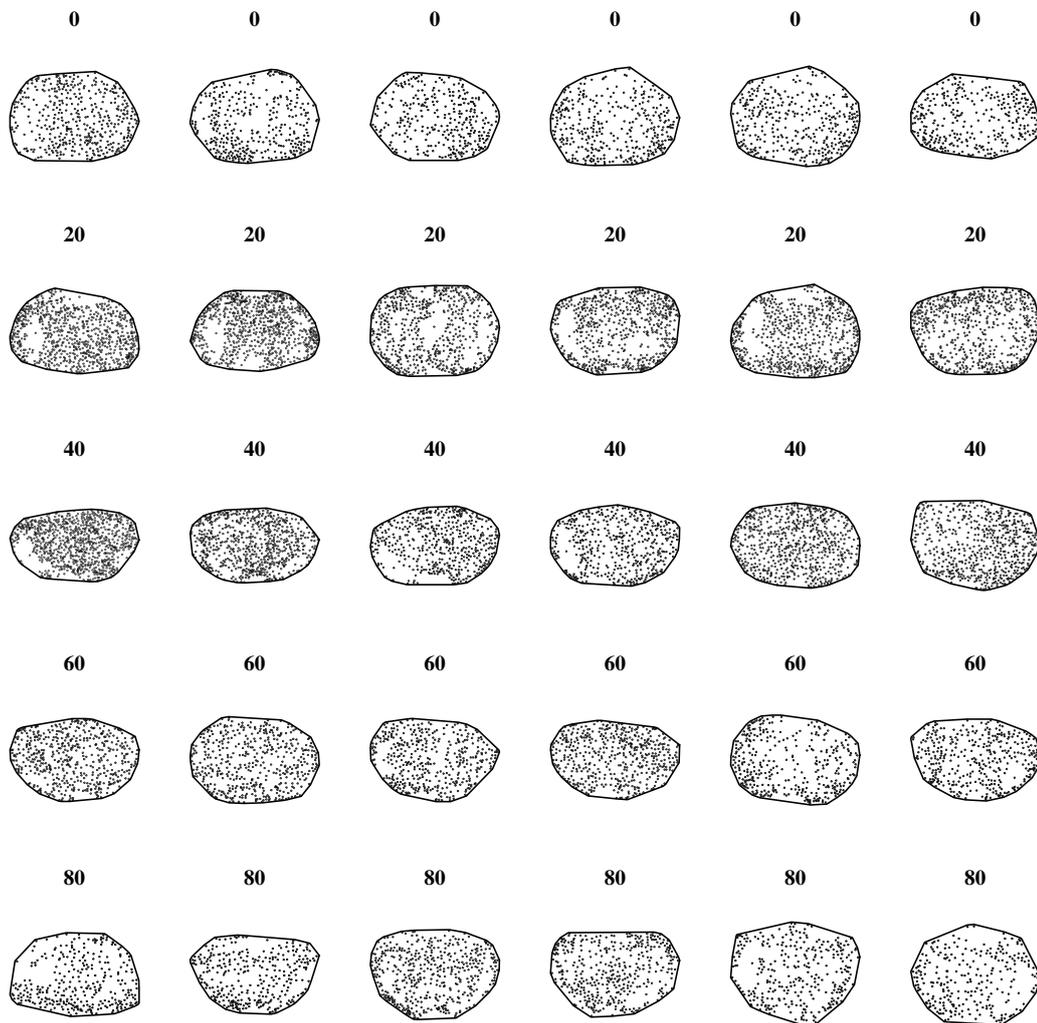
Assim, seja m_{ij} a configuração pontual da fatia de pão do i -ésimo grupo ($i = 1, \dots, 5$) e j -ésima repetição ($j = 1, \dots, 6$). Primeiramente, obteve-se a estimativa individual da função $\hat{K}_{ij}(h)$ de Ripley para cada configuração pontual dada por:

$$\hat{K}(h) = \hat{\lambda}^{-1} \sum_{i=1}^n \sum_{i \neq j} \frac{1}{a_{ij}} \frac{I(\|\mathbf{s}_i - \mathbf{s}_j\| \leq h)}{n-1}, \quad (2.1)$$

em que $\hat{\lambda}$ é um estimador da intensidade; $I(\cdot)$ é a função que indica se existe, ou não, evento com distância menor ou igual a h ; a_{ij} é a proporção da área da circunferência que se encontra dentro da região A , centrado em um evento arbitrário.

Calculamos a função $\bar{K}_i(h)$ do i -ésimo grupo, ponderando as estimativas individuais dada por

Figura 2.2 – Posições dos alvéolos digitalizados para os cinco tempos de congelamento dos pães franceses.



$$\bar{K}_i(h) = \frac{1}{n_i} \sum_{j=1}^{m_i} n_{ij} \hat{K}_{ij}(h). \quad (2.2)$$

Desse modo, a média dos grupos é dada pelo estimador da função K geral como apresentado a seguir:

$$\bar{K}(h) = \frac{1}{n} \sum_{i=1}^g n_i \bar{K}_i(h), \quad (2.3)$$

em que $n_i = \sum_{j=1}^{m_i} n_{ij}$ e $n = \sum_{i=1}^g n_i$.

As funções apresentadas podem ser analisadas por meio da abordagem gráfica. São esboçadas a função teórica $K(h)$ em conjunto com $\bar{K}(h)$ em função da distância h . Para a função $\bar{K}(h)$ são apresentados bandas de confiança inferior e superior, construídas a partir do

agrupamento das j -ésima repetição de cada i -ésimo grupo, dados por

$$L(h)_i = \min_{j=1,\dots,6} \{\hat{K}_{ij}(h)\} \quad \text{e} \quad U(h)_i = \max_{j=1,\dots,6} \{\hat{K}_{ij}(h)\},$$

$L(h)_i$ é os mínimos da função $\hat{K}_{ij}(h)$ para i -ésimo grupo; $U(h)_i$ é os máximos da função $\hat{K}_{ij}(h)$ para i -ésimo grupo.

Na análise do gráfico, distâncias em que a função teórica $K(h)$ não está contida na banda de confiança de $\bar{K}(h)$ rejeita-se a hipótese de que as configurações pontuais são completamente aleatórios. Valores acima da função teórica representam agrupamento entre os eventos e, abaixo, configurações pontuais com regularidade.

Para investigar a diferença e semelhança entre os grupos, Hahn (2012) propôs um teste em que a hipótese nula é que dois ou mais grupos de configurações pontuais são realizações de um mesmo processo pontual. Em outras palavras, se a distribuição dos alvéolos apresenta as mesmas características independente do tempo de congelamento. A estatística de teste é definida por

$$\bar{T} = \sum_{1 \leq i < j \leq g} \int_0^{h_0} \frac{(\bar{K}_i(h) - \bar{K}_j(h))^2}{\frac{1}{m_i} \bar{s}_i^2(h) + \frac{1}{m_j} \bar{s}_j^2(h)} dh, \quad (2.4)$$

com

$$\bar{s}_i^2(h) = r^2 \frac{1}{h_0} \int_0^{h_0} \frac{s_i^2(t)}{t^2} dt.$$

em que

$\bar{K}_i(r) = \frac{1}{m_i} \sum_{j=1}^{m_i} \hat{K}_{ij}(r)$ é a média dos grupos; $s_i^2(r) = \frac{1}{m_i - 1} \sum_{j=1}^{m_i} (\hat{K}_{ij}(r) - \bar{K}_i(r))^2$ é a variância amostral dentro do i -ésimo grupo para uma determinada distância r .

A estatística \bar{T} não têm distribuição amostral conhecida sob a hipótese nula. Este problema pode ser resolvido utilizando computação intensiva. Assim, permutam-se as configurações pontuais entre os grupos, mantendo inalterado o número de configurações por grupo e, a cada permutação, recalcula-se a estatística \bar{T} . Assim, T_1 é a estatística do teste é calculada e T_2, \dots, T_p são as estatísticas calculadas a partir de p permutações; neste trabalho a diferença entre grupos foi estudada usando 999 permutações. Compara-se o valor de T_1 com os valores T_2, \dots, T_p obtidos a partir das estatísticas permutadas. Se T_1 ficar entre o maiores valores de T_2, \dots, T_p , então, tem-se indicação que as configurações pontuais são realizações de um mesmo processo pon-

tual, isto é, não existe diferença entre os grupos. Suponha que $T_1 = T_{(j)}$ para algum $j \in \{1, \dots, p\}$ e considerando $\alpha \leq 0,05$ então rejeita-se a hipótese nula se o valor-p $= \frac{(p+1-j)}{p} \leq \alpha$.

As análises estatísticas das configurações pontuais replicadas foram realizada no software (R Core Team, 2017) utilizando o pacote Spatstat (BADDELEY; RUBAK; TURNER, 2015).

3 RESULTADOS E DISCUSSÃO

Obteve-se para cada tempo de congelamento, todas as posições dos alvéolos em cada fatia dos pães. O tempo de congelamento (grupo) 40 dias apresentou a maior média da intensidade, isto é, a maior quantidade de alvéolos por unidade de área, seguido do tempo de congelamento 20, 60, 80, 0. As intensidades individuais podem ser vista na Tabela 3.1.

A área do tempo de congelamento 80 diminuiu consideravelmente em relação ao tempo 0, porém os dois tempos apresentaram quase a mesma quantidade de alvéolos. Contudo, quando se leva em consideração a área de cada fatia do pão, o tempo 0 apresenta maior área que no tempo 80, indicando que apesar de haver um número de alvéolos relativamente igual, o tamanho deles e sua distribuição na área foram diferentes entre os tempos.

Reforçando esta hipótese tem-se a área das fatias que diminuiu à medida que o tempo de armazenamento aumentou. Várias alterações podem ocorrer durante o armazenamento dos pães sob congelamento, por exemplo, à medida que o tempo de armazenamento aumenta ocorre a redução do volume específico, diminuição no teor de umidade, aumento da dureza do miolo e mudanças no aroma (RIBOTTA et al., 2004; ALMEIDA; CHANG, 2014; BÁRCENAS; ROSELL, 2007).

Com o congelamento, além da formação de cristais de gelo, ocorrem processos de crescimento e recristalização de cristais que danificam significativamente a rede de glúten (MORI, 2008). Alguns trabalhos associam o enfraquecimento da massa a substâncias redutoras liberadas pelas leveduras durante o armazenamento congelado (CURTI et al., 2014; BESBES et al., 2014; MORI, 2008; RIBOTTA et al., 2004).

Os tempos de congelamento 20 e 40 dias apresentaram as maiores intensidades médias, 31.89 e 34.25 alvéolos/cm², respectivamente. Apesar da maior intensidade de alvéolos nestes tempos o volume dos pães foi relativamente menor. Isto indica que os pães nos tempos 20 e 40 dias apresentaram grande número de alvéolos menores por unidade de área. Sendo assim, torna-se necessário aprofundar o estudo nesse fato a fim de verificar sua significância.

Tabela 3.1 – Identificação da amostra, número de alvéolos, área e intensidade (por amostra de cada tempo de congelamento) e média do grupo

	Amostra	Número de alvéolos	Área (cm^2)	Intensidade (λ) (alvéolos/cm^2)
Tempo 0	1	413	18,53	22,29
	2	339	21,04	16,11
	3	301	20,53	14,66
	4	348	21,75	16,43
	5	346	21,24	16,29
	6	315	19,84	15,87
Tempo 20	7	723	19,46	37,14
	8	744	19,88	37,42
	9	554	19,48	28,43
	10	577	19,35	29,82
	11	596	20,17	29,55
	12	578	19,95	28,97
Tempo 40	13	855	17,68	48,34
	14	651	16,16	40,28
	15	457	18,63	24,53
	16	426	16,89	25,22
	17	608	16,28	37,34
	18	551	18,50	29,80
Tempo 60	19	493	18,56	26,56
	20	453	17,29	26,19
	21	392	18,78	20,87
	22	471	18,59	25,33
	23	346	17,85	19,38
	24	358	17,25	20,75
Tempo 80	25	317	18,06	17,55
	26	310	17,01	18,23
	27	449	19,03	23,59
	28	405	17,37	23,32
	29	313	15,44	20,27
	30	264	18,35	14,38
Todas as amostras		13953	558,94	754,91
Médias	Tempo 0	343,67	20,49	16,94
	Tempo 20	628,66	19,71	31,89
	Tempo 40	591,33	17,36	34,25
	Tempo 60	418,83	18,05	23,18
	Tempo 80	343	17,54	19,56

Aproveitando o fato de se ter 30 amostras de pães, estende-se a ideia da proposta de Bicalho et al. (2018) e foi verificada a dependência entre os alvéolos, sob a hipótese de aleatoriedade espacial aplicando o teste de Diggle (2013). Em todos os tempos de congelamentos há indícios para rejeitar a hipótese nula (valor- $p < 0,05$). Considerando este resultado, os alvéolos

do pão não estão distribuídos de forma completamente aleatória, ou seja, eles estabelecem um padrão regular ou se agrupam em algum momento, o que para o pão isso pode se tornar um problema. (ver a Tabela 3.2).

Tabela 3.2 – Estatística e valor-p do teste formal contra hipótese de completa aleatoriedade espacial para a função K

Tempo de congelamento	U	valor-p
0	$2,732 \times 10^{-5}$	0,01
20	$1,119 \times 10^{-5}$	0,01
40	$1,244 \times 10^{-5}$	0,01
60	$2,317 \times 10^{-5}$	0,01
80	$1,699 \times 10^{-5}$	0,01

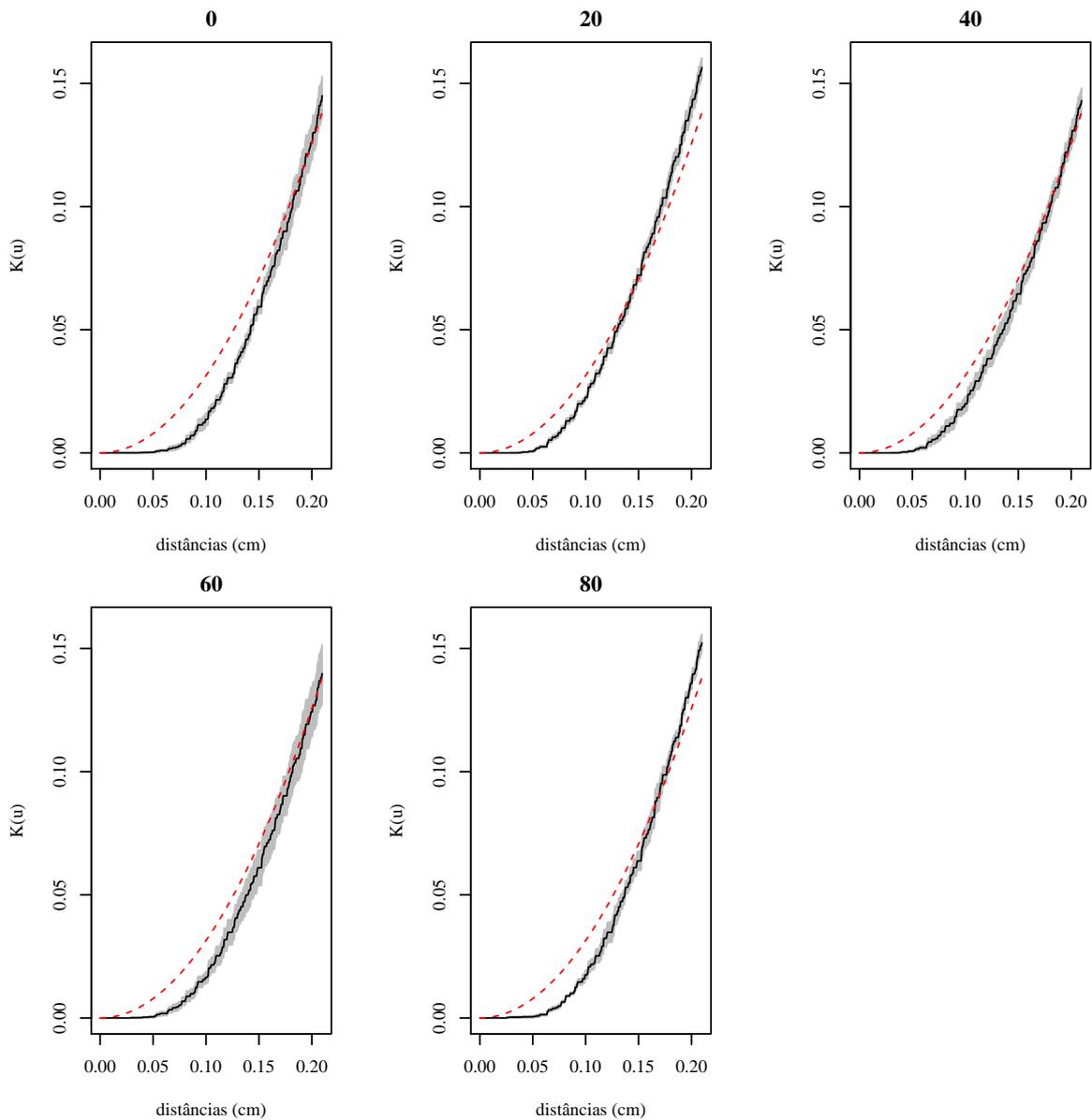
Em panificação, alvéolos agrupados podem estar relacionados com a baixa qualidade da farinha ou do fermento, ou ainda a problemas relacionados ao processamento do pão. Em produtos congelados, pães com *flaking* teriam alvéolos agrupados, por exemplo. Já um pão que apresenta distribuição alveolar regular indica que estão presentes menos alvéolos do que se esperaria em uma pão de boa qualidade.

Não se sabe em qual direção a hipótese nula foi rejeitada, então foi utilizada a análise gráfica. Como visto, as intensidades dos alvéolos eram diferentes; sendo assim, a função K para padrão de pontos replicados foi utilizada, porque esta não depende da intensidade para averiguar o comportamento da distribuição alveolar do pão.

A Figura (3.1) mostra a função K teórica na linha vermelha (pontilhada) e na linha cheia a função K média estimada para cada tempo de congelamento. Quando as configurações pontuais forem completamente aleatórias, as funções K teóricas serão iguais a πh^2 . Nota-se uma forte indicação de regularidade em pequena escala para todos os tempos de congelamento, pois a função estimada está abaixo da função teórica, rejeitando a hipótese nula de Completa Aleatoriedade Espacial.

Percebe-se que o comportamento das cinco funções K são muito semelhantes entre si, em todas as distâncias, para todas as amostras. Além disso, para distâncias curtas (0,05 até 0,20 cm), as funções estão abaixo da linha da função teórica (linha pontilhada vermelha), apresentando assim uma regularidade entre os alvéolos para todos os tempos de congelamento. A regularidade em pequena escala é suspeita quando uma distância mínima de inibição ou exclusão é evidente, já que não existe um alvéolo em cima de outro alvéolo. Assim, distâncias

Figura 3.1 – Função K teórica em vermelho (linha pontilhada) e a função K média em preto (linha cheia) de cada grupo de congelamento.

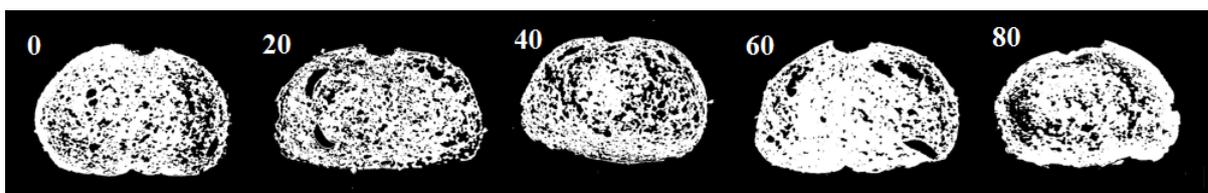


de até aproximadamente 0,05 cm não apresentam nenhum alvéolo, que é a distância mínima correspondente aos tamanhos dos raios dos alvéolos que não foram considerados nas análises.

As funções estimadas aproximam da função teórica: para o tempo 0 em 0,18 cm; tempo 20 em 0,13 cm; tempo 40 em 0,17 cm; para tempo 60 em 0,20 cm; tempo 80 em 0,16 cm. Estes valores apresentam a distância na qual se pode considerar que a distribuição alveolar é completamente aleatória. Quanto menor for esta distância, mais próximos da distribuição aleatoriamente estão os alvéolos. Então o tempo 20 é aparentemente o tempo que atende a esta afirmação e os tempos 60 e 80 são aqueles com tendência contrária a esta afirmação.

Na Figura (3.2), estão apresentadas cinco imagens das fatias dos pães em função do tempo de armazenamento, onde, claramente, observa-se o grau de aeração da massa durante os primeiros 20 dias de armazenamento e os efeitos do congelamento sobre a estrutura da massa após os 40 dias de armazenamento. No tempo 80, observa-se uma grande concentração de alvéolos ao redor da casca do pão, que vai ao oposto do observado na imagem do tempo 20 e 40, que tem alvéolos bem distribuídos em toda a superfície avaliada.

Figura 3.2 – Imagens do interior de 5 pães após cocção em função do tempo de armazenamento



Fonte: Do autor (2017).

Quando se tem grupos, o principal interesse é estudar semelhanças e diferenças entre os grupos de dados replicados. Então criaram-se testes para detectar essa diferença, como o teste de Diggle, Lange e Benes (1991), sugerindo o uso de método de *bootstrap*; uma versão mais recente deste teste foi criada por Diggle, Mateu e Clough (2000). Outro teste mais recente foi proposto por Hahn (2012) para comparar diferentes grupos de funções K .

Mesmo com a regularidade presente em todos os grupos, rejeitando a hipótese de aleatoriedade espacial, há também o interesse em determinar se esta regularidade apresenta semelhanças e diferenças entre os tempos de congelamento. No caso estudado, aplicou-se o teste do Hahn, a estatística do teste foi $\bar{T} = 13,714$ e o valor- $p = 0,038$, rejeitando a hipótese nula, detectando-se diferenças estatisticamente significativas da distribuição alveolar entre os tempos de congelamento. Essa diferença pode estar diretamente relacionada com o grau de regularidade em pequena escala.

Para entender melhor a distribuição alveolar, aprimou-se a análise na tentativa de detectar as diferenças mais específicas entre os tempos de congelamento. Aplicou-se o teste de Hahn 10 vezes com tratamentos em dupla ($g = 2$). Nesta análise foram verificadas quais tratamentos diferem entre si.

Nesta análise, os tempos de congelamento 0 – 20, 0 – 40, 0 – 80, 20 – 40, 20 – 60, 20 – 80 diferem entre si (valor- $p < 0,05$). A distribuição alveolar dos pães congelados por 20 dias é estatisticamente diferente dos demais pães. As maiores diferenças de acordo com a

estatística foram 0 – 20 e 20 – 40. Além disso, os tempos de congelamento 0 – 60, 40 – 60, 40 – 80 e 60 – 80 são considerados iguais, ou seja, o teste Hanh não encontrou diferenças significativas entre a distribuição dos alvéolos para os tempos (ver Tabela 3.3).

Tabela 3.3 – Estatística do teste de Hanh e seus respectivos valores-*p* para comparação 2 a 2

Comparação entre grupos	Estatística do Teste \bar{T}	Valor - <i>p</i>
0-20	4,7057	0,001
0-40	0,7684	0,039
0-60	0,1127	0,544
0-80	0,9889	0,035
20-40	2,2737	0,011
20-60	1,8288	0,017
20-80	1,9768	0,004
40-60	0,14176	0,464
40-80	0,50399	0,101
60-80	0,41232	0,172

Por meio do teste de Diggle (2013), analisou-se separadamente cada tempo de congelamento que foi detectado como estatisticamente significativo. O resultado da estatística do teste para os tempos de congelamentos 20 e 40 foi bem próximo, porém sabe-se que, quanto menor é este resultado, mais próximo da completa aleatoriedade espacial e acredita-se que melhor será a característica de qualidade da massa avaliada. Assim os pães armazenados e congelados por 20 dias foram aqueles que apresentaram menor estatística (Tabela 3.2).

Este resultado indica que os pães franceses armazenados congelados por 20 dias apresentavam-se mais aerados e com uma distribuição homogênea de alvéolos na massa, este efeito pode estar relacionado com a emulsificação provocada pela adição de mucilagem à massa. Contudo, à medida que o tempo de armazenamento aumentou, observou-se o aparecimento de zonas de fragilidade entre a casca e o miolo. Possivelmente o enfraquecimento do glúten e desestabilização da emulsão proporcionou a formação das bolhas maiores na região próxima a casca (Figura 3.2).

Espera-se que os resultados obtidos neste trabalho possa contribuir para área alimentícia, utilizando uma nova abordagem na análise da distribuição alveolar. Em trabalhos futuros,

pretende-se analisar, além da posição do alvéolos, alguma característica relacionada a essa posição, como por exemplo, o diâmetro.

4 CONCLUSÃO

A aplicação das metodologias de processos pontuais replicados torna-se efetiva para capturar variações da distribuição alveolar que justifiquem as diferenças entre os grupos (tempos de congelamento). Em todos os tempos de congelamento, os alvéolos apresentaram padrão regular e por meio do teste do Hanh, foi detectada diferença nessa regularidade. Pelo teste proposto por Diggle, o tempo de congelamento por 20 dias foi o que apresenta-se mais próximo da completa aleatoriedade espacial. Podendo-se sugerir que este seja o tempo limite para o congelamento de uma massa de qualidade.

Agradecimentos

Os autores agradecem à CAPES (Comissão de Aperfeiçoamento de Pessoal do Nível Superior), pelo suporte financeiro.

REFERÊNCIAS

- ALMEIDA, E. L.; CHANG, Y. K. Influence of different enzymes during the frozen storage of pre-baked french bread elaborated with whole-wheat flour. **Journal of Food Processing and Preservation**, v. 38, n. 3, p. 737–748, 2014.
- ASSUNÇÃO, R. M. Estatística espacial com aplicações em epidemiologia, economia e sociologia. **São Carlos: Associação Brasileira de Estatística**, v. 131, 2001.
- BADDELEY, A.; GILL, R. D. Kaplan-meier estimators of distance distributions for spatial point processes. **The Annals of Statistics**, p. 263–292, 1997.
- BADDELEY, A. et al. Analysis of a three-dimensional point pattern with replication. **Applied Statistics**, p. 641–668, 1993.
- BADDELEY, A.; RUBAK, E.; TURNER, R. **Spatial point patterns: methodology and applications with R**. London: CRC Press, 2015.
- BADDELEY, A. J.; MØLLER, J.; WAAGEPETERSEN, R. Non-and semi-parametric estimation of interaction in inhomogeneous point patterns. **Statistica Neerlandica**, Wiley Online Library, v. 54, n. 3, p. 329–350, 2000.
- BAILEY, T. C.; GATRELL, A. C. **Interactive spatial data analysis**. Harlow:Longman: Longman Scientific & Technical Essex, 1995. v. 413.

- BÁRCENAS, M. E.; ROSELL, C. M. Different approaches for increasing the shelf life of partially baked bread: low temperatures and hydrocolloid addition. **Food Chemistry**, v. 100, n. 4, p. 1594–1601, 2007.
- BELL, M. L.; GRUNWALD, G. K. Mixed models for the analysis of replicated spatial point patterns. **Biostatistics**, Oxford University Press, v. 5, n. 4, p. 633–648, 2004.
- BESAG, J. Contribution to the discussion of dr. ripleys paper. **Journal of the Royal Statistical Society. Series B**, v. 39, n. 2, p. 193–195, 1977.
- BESBES, E. et al. Effect of baking conditions and storage with crust on the moisture profile, local textural properties and staling kinetics of pan bread. **LWT-Food Science and Technology**, Elsevier, v. 58, n. 2, p. 658–666, 2014.
- BICALHO, C. C. et al. Alveolar distribution of french type bread using taro mucilage. **Brazilian Journal of Food Technology**, artigo submetido, 2018.
- CASTIGLIONI, G. L. et al. Tamanho dos alvéolos e aceitação de pães de forma enriquecidos com farelo de mandioca. **Pesquisa Agropecuária Tropical (Agricultural Research in the Tropics)**, v. 2, n. 1, p. 10–1590, 2014.
- CHIU, S.; STOYAN, D. Estimators of distance distributions for spatial patterns. **Statistica Neerlandica**, v. 52, n. 2, p. 239–246, 1998.
- CRESSIE, N. **Statistics for spatial data**. Canada: John Wiley & Sons, 1993.
- CURTI, E. et al. Bread staling: effect of gluten on physico-chemical properties and molecular mobility. **LWT-Food Science and Technology**, v. 59, n. 1, p. 418–425, 2014.
- DIGGLE, P. A kernel method for smoothing point process data. **Applied statistics**, p. 138–147, 1985.
- DIGGLE, P. J. **Statistical analysis of spatial and spatio-temporal point patterns**. England, UK: CRC Press, 2013.
- DIGGLE, P. J.; LANGE, N.; BENES, F. M. Analysis of variance for replicated spatial point patterns in clinical neuroanatomy. **Journal of the American Statistical Association**, v. 86, n. 415, p. 618–625, 1991.
- DIGGLE, P. J.; MATEU, J.; CLOUGH, H. E. A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns. **Advances in Applied Probability**, v. 32, n. 02, p. 331–343, 2000.
- ECKARDT, J. et al. Long-term frozen storage of wheat bread and dough—effect of time, temperature and fibre on sensory quality, microstructure and state of water. **Journal of Cereal Science**, v. 57, n. 1, p. 125–133, 2013.
- ECKEL, S. et al. Modelling tree roots in mixed forest stands by inhomogeneous marked gibbs point processes. **Biometrical Journal**, Wiley Online Library, v. 51, n. 3, p. 522–539, 2009.
- ESTELLER, M. S.; LANNES, S. C. da S. Parâmetros complementares para fixação de identidade e qualidade de produtos panificados. **Ciênc. Tecnol. Aliment**, v. 25, n. 4, p. 802–806, 2005.

- GERVINI, D.; BAUR, T. J. Regression models for replicated marked point processes. **arXiv preprint arXiv:1705.06259**, 2017.
- GRABARNIK, P.; MYLLYMÄKI, M.; STOYAN, D. Correct testing of mark independence for marked point patterns. **Ecological Modelling**, Elsevier, v. 222, n. 23, p. 3888–3894, 2011.
- GUAN, Y. Tests for independence between marks and points of a marked point process. **Biometrics**, Wiley Online Library, v. 62, n. 1, p. 126–134, 2006.
- GUAN, Y.; AFSHARTOUS, D. R. Test for independence between marks and points of marked point processes: a subsampling approach. **Environmental and Ecological Statistics**, Springer, v. 14, n. 2, p. 101–111, 2007.
- GUTKOSKI, L. C. et al. Efeito de ingredientes na qualidade da massa de pão de forma congelada não fermentada durante o armazenamento. **Ciência e Tecnologia de Alimentos**, v. 25, n. 3, p. 460–467, 2005.
- HAHN, U. A studentized permutation test for the comparison of spatial point patterns. **Journal of the American Statistical Association**, v. 107, n. 498, p. 754–764, 2012.
- HUEN, J. et al. Confocal raman microscopy of frozen bread dough. **Journal of Cereal Science**, Elsevier, v. 60, n. 3, p. 555–560, 2014.
- ILLIAN, J. et al. **Statistical analysis and modelling of spatial point patterns**. England: John Wiley & Sons, 2008. v. 70.
- LANDAU, S.; EVERALL, I. P. Nonparametric bootstrap for k-functions arising from mixed-effects models with applications in neuropathology. **Statistica Sinica**, JSTOR, p. 1375–1393, 2008.
- LANDAU, S.; RABE-HESKETH, S.; EVERALL, I. P. Nonparametric one-way analysis of variance of replicated bivariate spatial point patterns. **Biometrical journal**, Wiley Online Library, v. 46, n. 1, p. 19–34, 2004.
- LASSOUED, N. et al. Granulometry of bread crumb grain: Contributions of 2d and 3d image analysis at different scale. **Food Research International**, Elsevier, v. 40, n. 8, p. 1087–1097, 2007.
- LIESHOUT, M. V. A j-function for marked point patterns. **Annals of the Institute of Statistical Mathematics**, Springer, v. 58, n. 2, p. 235, 2006.
- MORI, R. Tecnologia para p congelados. **Food Ingredients Brasil, Puratos**, n. 05, p. 45–46, 2008.
- MYLLYMÄKI, M.; SÄRKKÄ, A.; VEHTARI, A. Hierarchical second-order analysis of replicated spatial point patterns with non-spatial covariates. **Spatial Statistics**, Elsevier, v. 8, p. 104–121, 2014.
- PENTTINEN, A.; STOYAN, D.; HENTTONEN, H. M. Marked point processes in forest statistics. **Forest science**, v. 38, n. 4, p. 806–824, 1992.
- POLAKI, A. et al. Fiber and hydrocolloid content affect the microstructural and sensory characteristics of fresh and frozen stored bread. **Journal of Food Engineering**, v. 97, n. 1, p. 1–7, 2010.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2017. Disponível em: <<https://www.R-project.org/>>.

RAMÓN, P. et al. A new non-parametric method for analyzing replicated point patterns in ecology. **Ecography**, Wiley Online Library, v. 39, n. 11, p. 1109–1117, 2016.

REDENBACH, C. et al. Anisotropy analysis of pressed point processes. **AStA Advances in Statistical Analysis**, Springer, v. 93, n. 3, p. 237–261, 2009.

RIBOTTA, P. et al. Effect of emulsifier and guar gum on micro structural, rheological and baking performance of frozen bread dough. **Food Hydrocolloids**, v. 18, n. 2, p. 305–313, 2004.

RIPLEY, B. D. The second-order analysis of stationary point processes. **Journal of applied probability**, v. 13, n. 02, p. 255–266, 1976.

RIPLEY, B. D. Modelling spatial patterns. **Journal of the Royal Statistical Society.**, p. 172–212, 1977.

SCANLON, M.; ZGHAL, M. Bread properties and crumb structure. **Food Research International**, Elsevier, v. 34, n. 10, p. 841–864, 2001.

SCHABENBERGER, O.; GOTWAY, C. A. **Statistical methods for spatial data analysis**. U.S.A.: CRC press, 2004.

SCHEUER, P. M. et al. Effects of fat replacement on properties of whole wheat bread. **Brazilian Journal of Pharmaceutical Sciences**, v. 50, n. 4, p. 703–712, 2014.

SCHLADITZ, K. et al. Statistical analysis of intramembranous particles using freeze fracture specimens. **Journal of microscopy**, Wiley Online Library, v. 211, n. 2, p. 137–153, 2003.

SCHLATHER, M.; RIBEIRO, P. J.; DIGGLE, P. J. Detecting dependence between marks and locations of marked point processes. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 66, n. 1, p. 79–93, 2004.

WÄLDER, O.; STOYAN, D. On variograms in point process statistics. **Biometrical Journal**, Wiley Online Library, v. 38, n. 8, p. 895–905, 1996.

WALLER, L. A.; GOTWAY, C. A. **Applied spatial statistics for public health data**. New Jersey: John Wiley & Sons, 2004.

ZAVAREZE, E. da R. et al. Efeito do soro de leite no teor protéico e na qualidade tecnológica e sensorial de pães. **Revista do Instituto de Laticínios Cândido Tostes**, v. 63, n. 363, p. 44–50, 2008.

ZOUNIS, S. et al. Effect of final dough temperature on the microstructure of frozen bread dough. **Journal of Cereal Science**, Elsevier, v. 36, n. 2, p. 135–146, 2002.

ARTIGO 3**ANÁLISE DE VARIÂNCIA NÃO PARAMÉTRICA DE PADRÕES DE PONTOS
ESPACIAIS REPLICADOS COM MARCAS QUANTITATIVAS**

**O artigo será traduzido e submetido
à uma revista na área de estatística
(versão preliminar)**

ANÁLISE DE VARIÂNCIA NÃO PARAMÉTRICA DE PADRÕES DE PONTOS ESPACIAIS REPLICADOS COM MARCAS QUANTITATIVAS

RESUMO

Estudos na área de panificação investigam a microestrutura dos produtos, que são obtidos através de imagens, na maioria das vezes, é caracterizada analisando somente valores médios. O estudo de cada alvéolo da massa é importante quando se tem a localização do alvéolo e alguma característica dessa localização (marca quantitativa), isso possibilita utilizar a análise espacial de processos pontuais marcados. Se esses dados forem obtidos a partir de grupos (tempo de congelamento), uma boa alternativa é usar a análise espacial de processos pontuais replicados com marcas quantitativas, uma questão desafiadora para a qual acredita-se que nenhum método foi proposto até agora. O método do presente trabalho é baseado na abordagem de replicação bivariada (marcas bivariadas), propondo um teste estatístico que compara grupos com marcas quantitativas, usando a função K ponderada por marcas. Dados simulados foram utilizados para caracterizar e validar essa abordagem, através do erro do tipo I. O método foi então aplicado em diâmetros de alvéolos (marca), em tempos de congelamento (grupos) do pão para avaliar se o tempo de congelamento influencia na distribuição espacial dos alvéolos. Realizou-se uma análise de variância não-paramétrica com 30 amostras de fatias de pães, divididos em 5 tempos de congelamento (0, 20, 40, 60, 80 dias), com 6 repetições cada. Avaliaram-se grupos inteiros contra a completa aleatoriedade espacial e independência das marcas. Os grupos foram comparados pelo novo teste proposto para a hipótese de que dois (ou mais) padrões de pontos são realizações do mesmo processo de pontos com marca. Com os resultados foi possível caracterizar os alvéolos em pães congelados, permitindo estabelecer os tempos de congelamento que apresentam diferenças significativas nos tamanhos de alvéolos.

Palavras-chave: Processo do ponto espacial marcados; Replicação; Permutação; Comparação de grupo.

1 INTRODUÇÃO

Existe uma diversidade de métodos para analisar e modelar padrões de pontos espaciais (por exemplo: Diggle (2013); Illian et al. (2008); Cressie (1993)). Na verdade, a maioria dos estudos envolvendo padrão de pontos são análises baseadas em apenas um conjunto de dados ou em estudos em que os vários padrões são analisados individualmente.

Além do estudo da localização do ponto, pode-se estar interessado em algumas características relacionadas a este ponto, a qual é chamada de marca, caracteriza-se, assim, o processo pontual marcado. Processo pontual marcado é um processo estocástico $\Psi = [s_i; m(s_i)]$, que gera os eventos (s_i) e as marcas associadas aos eventos $m(s_i)$. Uma realização de Ψ é um conjunto de posições juntamente com as marcas associadas, sendo que essas marcas podem ser quantitativas ou qualitativas (PENTTINEN; STOYAN; HENTTONEN, 1992). Quando se tem marcas numéricas (quantitativas), o estudo deve ser iniciado com gráficos exploratórios e teste de inde-

pendência espacial entre marcas e pontos. Considerando um padrão homogêneo, vários tipos de testes entre marca e pontos estão presentes na literatura: pode-se utilizar qualquer função de resumo como a função de correlação de marcas, a função K ponderada por marcas, variograma marcado, entre outros (BADDELEY; RUBAK; TURNER, 2015), além de testes propostos por: Schlather, Ribeiro e Diggle (2004); Guan (2006) e Guan e Afshartous (2007). No caso de marcas quantitativas, considera-se a hipótese nula como o resultado de independência na atribuição da marca aos pontos; se essa hipótese for rejeitada pode-se ter indicativo de estimulação ou inibição entre marcas. A estimulação ocorre quando eventos vizinhos têm as marcas maiores do que a marca média de todo o processo, enquanto que a inibição indica que as marcas são menores do que a marca média para pontos próximos (GRABARNIK; MYLLYMÄKI; STOYAN, 2011).

Quando se tem duas ou mais configurações pontuais disponibilizadas em grupos, a análise vem acarretar estudos mais elaborados, denominada de análise de padrão de pontos replicados. Os métodos para analisar padrões de pontos replicados estão disponíveis na literatura estatística e podem ser classificados em métodos paramétricos e não paramétricos (LANDAU; EVERALL, 2008). Os trabalhos de Eckel et al. (2009); Bell e Grunwald (2004); Diggle, Mateu e Clough (2000), é possível apreciar métodos paramétricos baseados em modelos de padrões de pontos replicados.

Já para métodos não paramétricos, Diggle, Lange e Benes (1991) e Baddeley et al. (1993) introduziram em padrões de pontos algumas estatísticas de resumo que, ao serem agrupadas e estimadas, podem ser utilizadas para testar diferenças em grupos dos padrões espaciais. Esses métodos já foram aplicados nos trabalhos de Schladitz et al. (2003); Redenbach et al. (2009); Bicalho et al. (2018); Hahn (2012), além de aplicar esses métodos, propôs um teste não paramétrico para comparação de dois ou mais padrões de pontos. Já Myllymäki, Särkkä e Vehtari (2014) propôs um modelo não paramétrico para estudar a dependência entre as funções de resumo e as covariáveis.

Abordagens para analisar padrões de pontos replicados marcados são mais restritas. Landau, Rabe-Hesketh e Everall (2004) foram os precursores nesse tipo de análise, estendendo a abordagem baseada em replicação de Diggle, Lange e Benes (1991), permitindo a comparação da interação espacial de dois tipos de células entre grupos, incluindo comparações planejadas e avaliando grupos inteiros contra a independência espacial e a Completa Aleatoriedade Espacial.

As relações das marcas são a principal tarefa de investigação em padrões de pontos marcados. Se as marcas forem independentes não faz sentido estudar utilizando técnicas de padrões de pontos espacial marcados, bastando estudar simplesmente a distribuição da marca independente das posições pontuais.

Portanto, os testes de independência de marca em padrões de pontos desempenham um papel fundamental na estatística. Neste trabalho, uma extensão do teste proposto por Hahn (2012) é estudado. Concentrou-se em marcas quantitativas em padrão de pontos replicados marcados, com o intuito de testar a hipótese nula de igualdade entre as funções K_f (PENTTINEN; STOYAN; HENTTONEN, 1992), ou seja, comparar a dependência entre a marca e o ponto em processos de pontos replicados marcados.

O crescimento gradativo da indústria alimentícia proporciona pesquisas que devem continuar melhorando a variedade, qualidade e disponibilidade de produtos alimentares, particularmente o pão, por ser um alimento mundial básico. Neste trabalho, as configurações pontuais utilizadas são amostras de pão francês que tiveram sua massa congelada antes do assamento.

Muitos trabalhos com avaliações qualitativas de produtos congelados apontam considerações exploratórias sobre a microestrutura formada antes ou após congelamento ou assamento. A exemplo, tem-se o estudo que avaliou o efeito do tempo de armazenamento congelado sobre a microestrutura da massa (BÁRCENAS; ROSELL, 2007); ou que avaliou o efeito da temperatura final da massa sobre a microestrutura da massa congelada (ZOUNIS et al., 2002); ou que avaliou o uso de fibra e hidrocolóides sobre a microestrutura do miolo (POLAKI et al., 2010); ou ainda mais recentemente, que avaliou a interação da matriz dos ingredientes do pão na tentativa de caracterizar fenômenos de cristalização em produtos assados congelados, permitindo compreender melhor os motivos das perdas de qualidade e desenvolver estratégias para evitar perdas (HUEN et al., 2014).

De um modo geral, muitos dados são obtidos nas avaliações exploratórias feitas sobre a microestrutura do miolo através da microscopia. Utilizou-se a localização e o diâmetro dos alvéolos (marcas quantitativas). Testou-se a existência de evidências de dependência do tamanho dos alvéolos, em grupos diferentes de tempos de armazenamento congelado, a fim de avaliar os eventos gerados pela interação da matriz alimentar de forma quantitativa, que por sua vez, parece ser uma forma interessante e inovadora.

Contabilizando os efeitos dos diâmetros dos alvéolos em cada amostra e entre os grupos de congelamento da massa espera-se que os alvéolos fossem distribuídos de forma semelhante,

independentemente do tempo de congelamento. Os métodos propostos neste trabalho foram submetidos ao estudo de simulação para verificar o comportamento do erro tipo I. Até onde se sabe, é a primeira vez que tais testes são utilizados para averiguar dependência entre marcas quantitativas em padrões de pontos replicados.

2 MATERIAL E MÉTODOS

2.1 - A função K univariada e a função K ponderada por marcas

A função K univariada (RIPLEY, 1977) mede a distribuição espacial dos eventos de um processo unitário e isotrópico e é definida como

$$K(h) = \lambda^{-1} E[N(h)], \quad (2.1)$$

em que $E[N(h)]$ é o valor esperado do número de eventos dentro de uma distância h em torno de um evento arbitrário; λ é a intensidade do processo pontual.

Assim, função $K(h)$ reflete a relação entre os eventos, a completa aleatoriedade espacial (CAE); afirma que os n eventos dado em um processo na região A , as posições espaciais desses eventos são uma amostra aleatória independente de uma distribuição uniforme sobre a região A . O modelo é equivalente a um processo de Poisson (DIGGLE, 2013). Sob CAE temos $K(h) = \pi h^2$. Além disso, o valor πh^2 pode ser mostrado para definir uma referência para categorizar a agregação espacial de um processo univariado, com $K(h) > \pi h^2$ indicando agrupamento e $K(h) < \pi h^2$ indicando regularidade até a distância h .

O estimador da função K proposto por Ripley (1977) que corrige efeito de borda é dado por

$$\hat{K}(h) = \hat{\lambda}^{-1} \sum_{i=1}^n \sum_{i \neq j} \frac{1}{w_{ij}} \frac{I(\|\mathbf{s}_i - \mathbf{s}_j\| \leq h)}{n-1}, \quad (2.2)$$

em que $\hat{\lambda}$ é um estimador da intensidade; $I(\cdot)$ é a função que indica se existe ou não evento com distância menor ou igual a h . w_{ij} é a proporção da área da circunferência que se encontra dentro da região A , centrado em um evento arbitrário.

A função K ponderada por marcas, de acordo com Penttinen, Stoyan e Henttonen (1992) é uma extensão da função K de Ripley, na qual para cada par de eventos existe uma ponderação

por uma função teste $f(m_1, m_2)$ de suas respectivas marcas. Um estimador não-viesado, que corrige o efeito de borda para a função K_f é dado por

$$\hat{K}_f(h) = \frac{|A|}{n^2} \sum_{k=1}^n \sum_{l=1, l \neq k}^n f(m(s_k), m(s_l)) \mathbf{1}(\|s_k - s_l\| \leq h) \delta(s_k, s_l), \quad (2.3)$$

em que n é o número de pontos marcados na região de estudo A de área $|A|$; f é uma função que atribui um peso a um par dos pontos s_k e s_l dependendo do tipo de suas marcas $m(s_k)$ e $m(s_l)$; $\mathbf{1}(\cdot)$ é a função que é igual a 1 se o argumento for verdadeiro e 0, caso contrário; $\delta(s_k, s_l)$ é um termo de correção de borda. Sob hipótese de completa aleatoriedade espacial (CAE) tem-se que $K(h) = K_f(h)$. A escolha de f se dá de acordo com o tipo das marcas, para marcas quantitativas utilizaremos a função f como sendo

$$f[m(s_k), m(s_l)] = \frac{m(s_k)m(s_l)}{\hat{m}^2}, \quad (2.4)$$

em que $\hat{m} = \frac{1}{n} \sum m(s_i)$ é a média das marcas.

Neste caso a função (2.4) é denominada de função de correlação de marcas ($\kappa_{mm}(\mathbf{h})$) e a sua interpretação é se $\kappa_{mm}(h) < 1$, sugere-se inibição mútua entre as marcas à uma distância de tamanho h ; se $\kappa_{mm}(h) > 1$ estimulação mútua para uma certa distância h e se as marcas não serem correlacionadas (independência) $\kappa_{mm}(h) \equiv 1$.

2.2 Análise de variância com a função K ponderada por marcas

Propôs-se um teste não paramétrico para verificar a independência espacial entre a marcas e pontos nos grupos de réplicas independentes, que pode ser denominado de análise de variância unidirecional com funções K para marcas quantitativas. Uma extensão para o caso de análise de variância em padrão de pontos replicado bivariado (LANDAU; RABE-HESKETH; EVERALL, 2004).

Seja $\hat{K}_{f,ij}(h)$ a função K ponderada por marca para a j -ésima ($j = 1, \dots, v_i$) configuração espacial no i -ésimo grupo ($i = 1, \dots, g$) definida por,

$$\hat{K}_{f,ij}(h) = \frac{|A_{ij}|}{n_{ij}^2} \sum_{k=1}^n \sum_{l=1, l \neq k}^n f_{ij}(m(s_k), m(s_l)) \mathbf{1}(\|s_k - s_l\| \leq h) \delta(s_k, s_l), \quad (2.5)$$

em que n_{ij} indica o número de eventos em cada configuração pontual na região de estudo A de área $|A_{ij}|$; f_{ij} é uma função que atribui um peso a um par dos pontos s_k e s_l .

O primeiro passo é obter todas as $\hat{K}_{f,ij}(h)$, em seguida, calcula-se a função $\bar{K}_{f,i}(h)$, a média dos grupos, que é definida como

$$\bar{K}_{f,i}(h) = \frac{1}{n_i} \sum_{j=1}^{m_i} n_{ij} \hat{K}_{f,ij}(h), \quad (2.6)$$

em que n_i é o número de eventos do i -ésimo grupo e n_{ij} indica o número de eventos em cada configuração pontual de um processo espacial aleatório.

No caso replicado com marcas quantitativas, o primeiro interesse está na comparação geral do grupo com a independência das marcas, isto é, se as funções $K_{f,i}$ consistentes com a função K_f teórica de um processo estacionário sob a independência espacial.

$$\begin{cases} H_0 : K_{f,i}(h) = \pi h^2 & \text{para todos } h \\ H_1 : K_{f,i}(h) \neq \pi h^2 & \text{para algum } h, \end{cases}$$

em que $K_{f,i}(h) = E[\hat{K}_{f,ij}(h)]$.

Realiza-se a interpretação gráfica da função $\bar{K}_{f,i}(h)$ em função das distâncias h escolhida, em conjunto com a banda de confiança inferior e superior de $\bar{K}_{f,i}(h)$, construídas a partir do agrupamento das j -ésima repetição de cada i -ésimo grupo, dados por

$$L(h)_i = \min_{j=1, \dots, v_i} \{\hat{K}_{f,ij}(h)\} \quad \text{e} \quad U(h)_i = \max_{j=1, \dots, v_i} \{\hat{K}_{f,ij}(h)\},$$

$L(h)_i$ é os mínimos da função $\hat{K}_{f,ij}(h)$ para i -ésimo grupo; $U(h)_i$ é os máximos da função $\hat{K}_{f,ij}(h)$ para i -ésimo grupo.

Se para distâncias h a banda de confiança não incluir a função teórica K_f são consideradas estatisticamente significativas, recusando a hipótese nula. Caso contrário, a função $\bar{K}_{f,i}(h)$ não é estatisticamente significativa, logo $\bar{K}_{f,i}(h) = \pi h^2$ para todos h .

Foi proposto um teste para funções K_f de padrões de pontos replicados com marcas quantitativas, realizando uma comparação global da igualdade das funções entre os grupos. Assim definiu-se a hipótese nula como,

$$H_0 : K_{f,1}(h) = K_{f,2}(h) = \dots = K_{f,g}(h) \quad \text{para todo } h,$$

contra a hipótese alternativa de que uma diferença possa ser detectada.

Para tal, considere que exista b configurações pontuais, divididas em g grupos, de tamanhos v_1, \dots, v_g . Para as configurações no grupo 1 denominou-se b_{11}, \dots, b_{1v_1} , etc.

Realizados os cálculos da estimativa das funções $K_{f,i,j}$ e a média de cada grupo $\bar{K}_{f,i}$, comparou-se a igualdade das funções utilizando a seguinte estatística de teste dada por

$$\bar{T}_f = \sum_{1 \leq i < j \leq g} \int_0^{h_0} \frac{(\bar{K}_{f,i}(h) - \bar{K}_{f,j}(h))^2}{\frac{1}{m_i} \bar{s}_i^2(h) + \frac{1}{m_j} \bar{s}_j^2(h)} dh, \quad (2.7)$$

com

$$\bar{s}_i^2(h) = h^2 \frac{1}{h_0} \int_0^{h_0} \frac{s_i^2(t)}{t^2} dt, \quad s_i^2(h) = \frac{1}{v_i - 1} \sum_{j=1}^{v_i} (\hat{K}_{ij}(h) - \bar{K}_i(h))^2,$$

em que $s_i^2(h)$ é a variância amostral dentro do i -ésimo grupo para uma determinada distância h . A estatística \bar{T}_f não têm distribuição amostral conhecida sob a hipótese nula e como solução utiliza-se o método de permutação. Este método é descrito da forma:

- i** Permutam-se as configurações pontuais entre os grupos, mantendo inalterado o número de configurações por grupo;
- ii** a cada permutação recalcula-se a estatística \bar{T}_f ;
- iii** Por meio da estatística de ordem calcula-se a estatística do teste T_1 ;
- iv** A partir de p permutações calculam-se as estatísticas T_2, \dots, T_p (neste trabalho foi utilizado $p = 999$);
- v** Compara-se o valor de T_1 com os valores T_2, \dots, T_p ;
- vi** Se T_1 ficar entre o maiores valores de T_2, \dots, T_p , então, tem-se indicação que as marcas são dependentes nas configurações pontuais, isto é, há indícios para rejeitar H_0 ;
- vii** Suponha que $T_1 = T_{(j)}$ para algum $j \in \{1, \dots, p\}$ e considerando $\alpha \leq 0,05$ há indícios de rejeitar a hipótese nula se o valor-p = $\frac{(p+1-j)}{p} \leq \alpha$, pelo menos um grupo difere dos demais ao que se refere dependência de marcas.

2.3 Estudo de simulações

Com o intuito de investigar o desempenho do teste proposto neste trabalho (\bar{T}_f), realizaram-se simulações sob H_0 para estudar o erro tipo I , o qual consiste na probabilidade de rejeitar a

$H_0 \leq \alpha$ quando H_0 é verdadeira,

$$H_0 : K_{f,1}(h) = K_{f,2}(h) = \dots = K_{f,g}(h) \quad \text{para todo } h.$$

Consideraram-se 1000 situações de estruturas de padrões de pontos replicados marcados simulados sob a Completa Aleatoriedade Espacial com marcas independentes. Para cada experimento simulado, geraram-se 30 configurações pontuais divididas em 5 grupos, totalizando 30000 configurações pontuais com marcas.

Cada configuração pontual foi gerada em janelas retangulares, com área variando de 15,0 a 24,0 cm², em que a largura e altura foram geradas por uma distribuição uniforme [5, 6] e [3, 4], respectivamente; a quantidade total de eventos foi gerada através de uma distribuição de Poisson com intensidade de 24,6 eventos/cm², reproduzindo condições reais.

Aplicaram-se 1000 vezes o teste \bar{T}_f , para obter o valor- p com 999 permutações em cada experimento. Calculou-se a proporção de vezes que o teste indicou a rejeição de H_0 , a quantidades de vezes em que o teste detectou alguma diferença entre as funções $\bar{K}_f(h)$ para alguma distância h , isto é, cometeu-se o erro do tipo I . Devido à complexidade da simulação da hipótese alternativa, o poder do teste será estudado em trabalhos futuros.

2.4 Estudo de caso

Aplicou-se a abordagem em imagens de fatias do pão do tipo francês, as quais podem ser usadas para caracterização de alvéolos, evitando assim a necessidade de análise sensorial (LASSOUED et al., 2007).

A preparação das fatias dos pães e obtenção dos dados dos alvéolos foram realizadas no Departamento de Ciência de Alimentos, da Universidade Federal de Lavras, Brasil. O experimento foi conduzido em um delineamento inteiramente casualizado (DIC) contendo cinco tratamentos e seis repetições. Os tratamentos foram os tempos de armazenamento da massa do pão congelada em 0, 20, 40, 60 e 80 dias. A fatia do meio, unidade experimental, foi escolhida e escaneada em *scanner HP deskjet 3040* a 200 dpi colorido. As imagens obtidas foram tratadas pelo software ImageJ usando a seguinte macro: *enhance contrast 0.4%, smoth, adjust color threshold* ajustado em *black & white*.

Primeiramente, realizou-se uma análise visual das amostras dos alvéolos dos pães. As 30 amostras das fatias dos pães tiveram áreas de tamanhos diferentes e todas as medidas desse

estudo foram dadas em centímetros. Para a aplicação dos métodos de processos pontuais marcados, cada evento estudado foi o diâmetro do alvéolo do pão definido pelas suas coordenadas. Assim, estimaram-se as 30 funções K_f , em seguida, agrupou-se as K_f por tempo de congelamento e testou a hipótese nula $K_{f,i}(h) = \pi h^2$ para todas distâncias h verificando a independência entre marcas e eventos.

Após a rejeição de independência, executa-se o teste proposto neste trabalho, testando a hipótese nula de igualdade das $K_{f,i}(h)$ na tentativa de encontrar diferenças e semelhanças na estrutura espacial entre diâmetro (marca) e os alvéolos (eventos) em tempos de congelamento diferentes (grupos), obtendo \bar{T}_f e o valor- p . Rejeitado essa hipótese, o próximo passo da análise consiste em identificar quais os grupos estão diferindo, realiza-se uma análise 2 a 2.

Nas análises exploratórias, nos testes da hipótese, foram usadas funções desenvolvidas e/ou disponíveis no software R (R Core Team, 2017) utilizando a biblioteca *Spatstat* (BADDELEY; RUBAK; TURNER, 2015). Para realizar a análise de processos pontuais replicados marcados foi elaborada funções utilizando o software R (R Core Team, 2017).

3 RESULTADOS E DISCUSSÃO

3.1 Estudos de simulações

Realizou-se um estudo de simulação para validar o teste proposto neste trabalho. Investigou-se o erro do tipo I que foi exemplificado com base nos dados simulados sob a hipótese nula de Completa Aleatoriedade Espacial e marcas independentes (H_0 verdadeira).

A Tabela (3.1) relata os três níveis de significância (0,01, 0,05 e 0,10). Os resultados apresentados foram a proporção de rejeição de H_0 verdadeira em 1000 situações simuladas.

Tabela 3.1 – Proporção de rejeição H_0 sendo H_0 verdadeira

Teste proposto	Nível de significância		
	$\alpha = 0,01$	$\alpha = 0,05$	$\alpha = 0,10$
\bar{T}_f	0,012	0,052	0,105

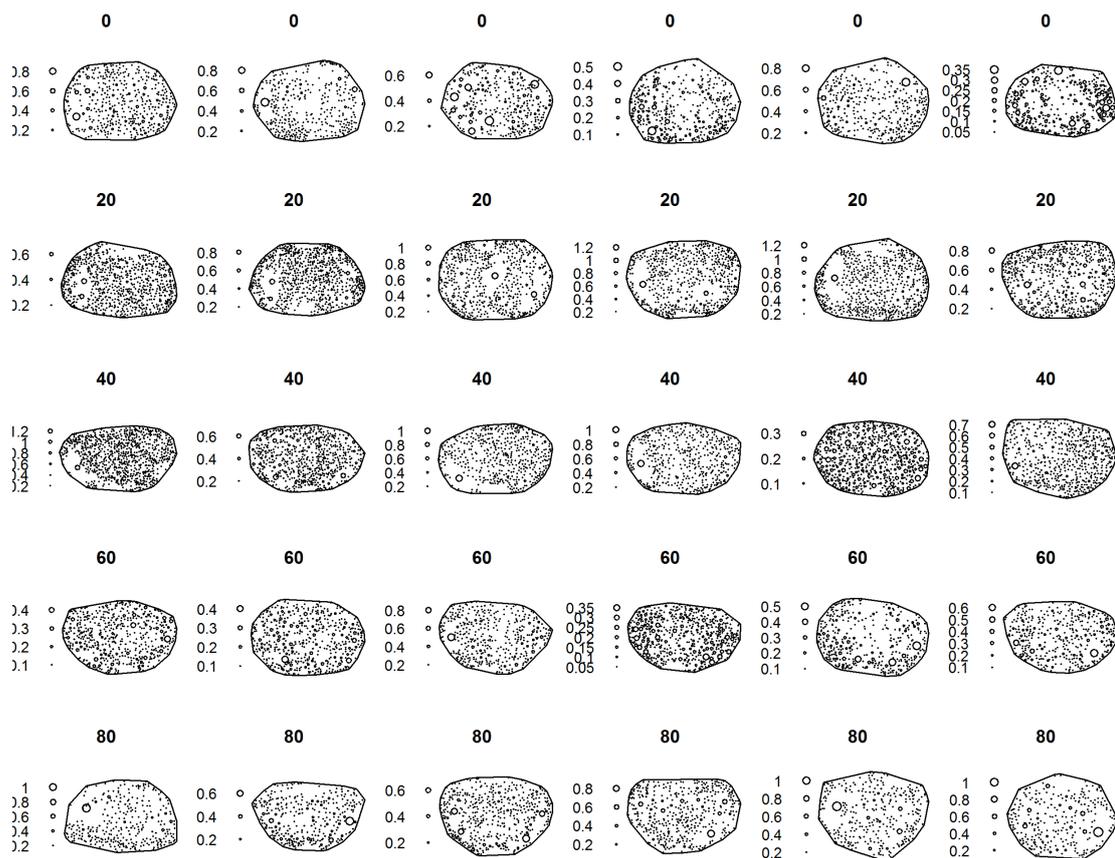
Para todos os níveis de significância observou-se que o teste proposto aderiu uma baixa taxa de rejeição de H_0 . Com base nesta evidência, os resultados da simulação incluídos neste artigo sugerem que o teste proposto pode detectar diferença entre grupos para padrões de pontos replicados com marcas quantitativas.

3.2 Análise descritiva das marcas

A composição e o armazenamento da massa influencia diretamente a estrutura do pão (POLAKI et al., 2010). Estudos que analisem a estrutura do pão podem ser usados para prever outros atributos de qualidade, como: textura, volume de pão ou características sensoriais e informações sobre o papel dos ingredientes na estrutura e aparência da estrutura dos alvéolos (SCANLON; ZGHAL, 2001).

Aplicou-se o teste proposto em um estudo de tamanho do alvéolos no pão. Na Figura (3.1) apresenta as 30 fatias de pães francês com o tamanho dos diâmetros de todos os alvéolos. Nota-se que alvéolos de tamanho pequeno tendem a ser predominantes em todos os tempos de congelamento.

Figura 3.1 – Distribuição dos alvéolos com 6 repetições para cada tempo de congelamento da massa.



A quantidade de alvéolos a serem estudados é diferente para cada tempo de congelamento. O tamanho mínimo do diâmetro do alvéolo é de 0,01433 cm, os maiores alvéolos medem aproximadamente 1,32 cm de diâmetro e estão presentes no tempo de congelamento

20, que foi também o tempo com maior quantidade de alvéolos. Já a menor quantidade de alvéolos estiveram presentes no tempo de congelamento 80 dias (ver tabela 3.2).

Tabela 3.2 – Número de alvéolos e medida mínima e máxima em centímetros dos alvéolos presentes em cada tempo de congelamento de fatias de pães franceses.

Tempo de congelamento	Número de alvéolos	diâmetro em centímetro [menor-maior]
0	2062	[0,01433-0,98932]
20	3772	[0,01433-1,32306]
40	3548	[0,01433-1,30344]
60	2513	[0,01433-0,99212]
80	2058	[0,01433-1,18033]

Os histogramas (Figura 3.2) das marcas exibem claramente uma assimetria à direita, mostrando que em todos os tempos de congelamento os alvéolos pequenos são a maioria, medindo em torno de de 0,014 até 0,04 cm aproximadamente. Os tempos de congelamento 20 e 40, além de apresentarem os maiores alvéolos, foram também os que apresentaram maior quantidade de alvéolos pequenos (cerca de 800). Aparentemente isso pode ser considerado bom, pois a quantidade de alvéolos está diretamente relacionada com a maciez da massa; quanto mais alvéolos e mais uniformes, melhor é considerada a massa e, conseqüentemente, a textura.

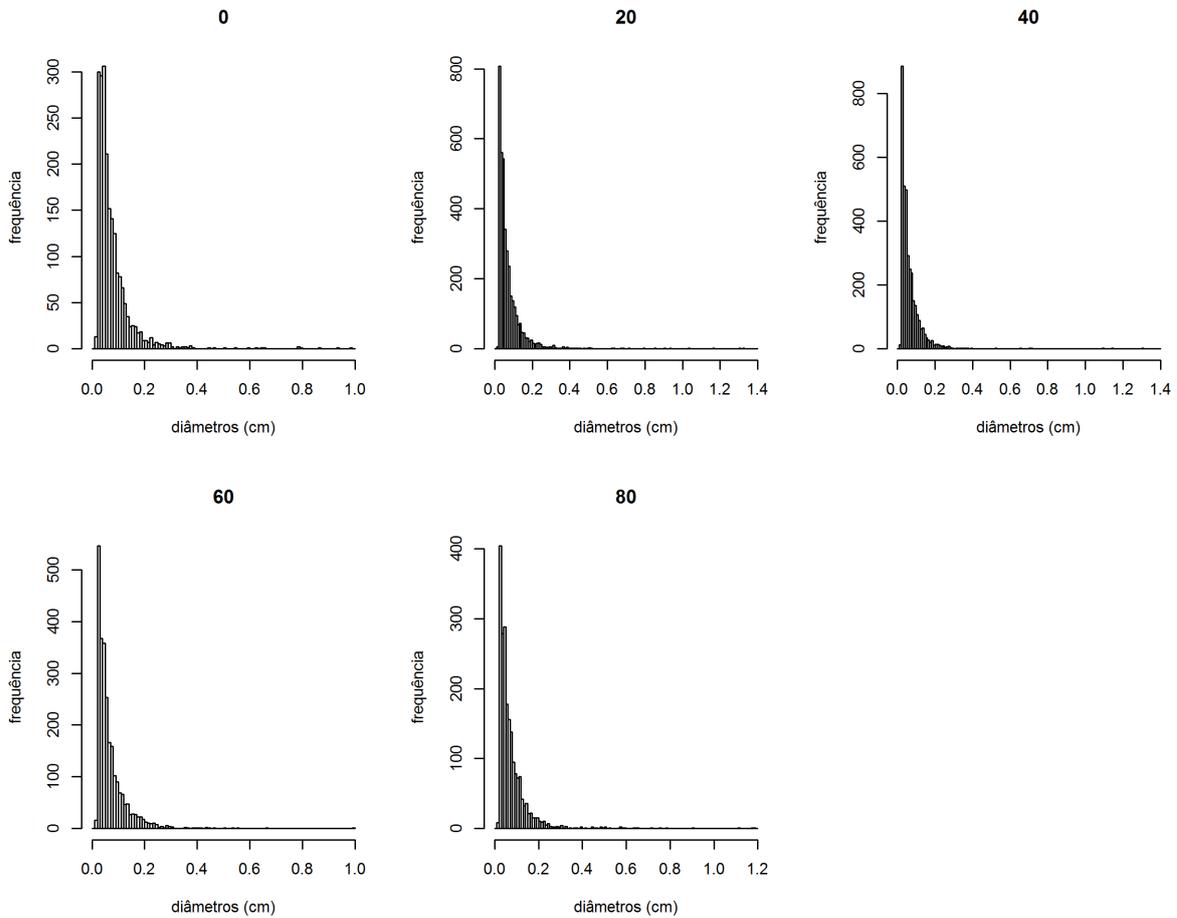
3.3 Resultado da análise de variância não paramétrica

Para detectar a independência entre as marcas e o evento em padrão de ponto, várias abordagens gráficas têm sido propostas. Estas incluem o uso da função de correlação marcada/variograma (WÄLDER; STOYAN, 1996); as funções E e V (SCHLATHER; RIBEIRO; DIGGLE, 2004); a função J (LIESHOUT, 2006); as duas funções propostas por Guan (2006) que desenvolveram tanto uma abordagem gráfica quanto uma abordagem formal de testes.

Seria possível utilizar qualquer um desses testes, mas expandindo o teste proposto por Hahn (2012) para processos marcados, foi oportuna a escolha pela função K_f (PENTTINEN; STOYAN; HENTTONEN, 1992) que também é considerada uma extensão da função K (RIPLEY, 1977).

Na Figura (3.3), a linha vermelha (pontilhada) representa a função K_f teórica, a linha preta (contínua), juntamente com as bandas de confiança (cinza) representam a função K_f estimada para cada grupo. Os valores que a banda de confiança não englobar, a função teórica fornece evidências contra a hipótese nula ($H_0 : K_{f,i} = \pi h^2$). Para distâncias de até 0,1 cm, esta

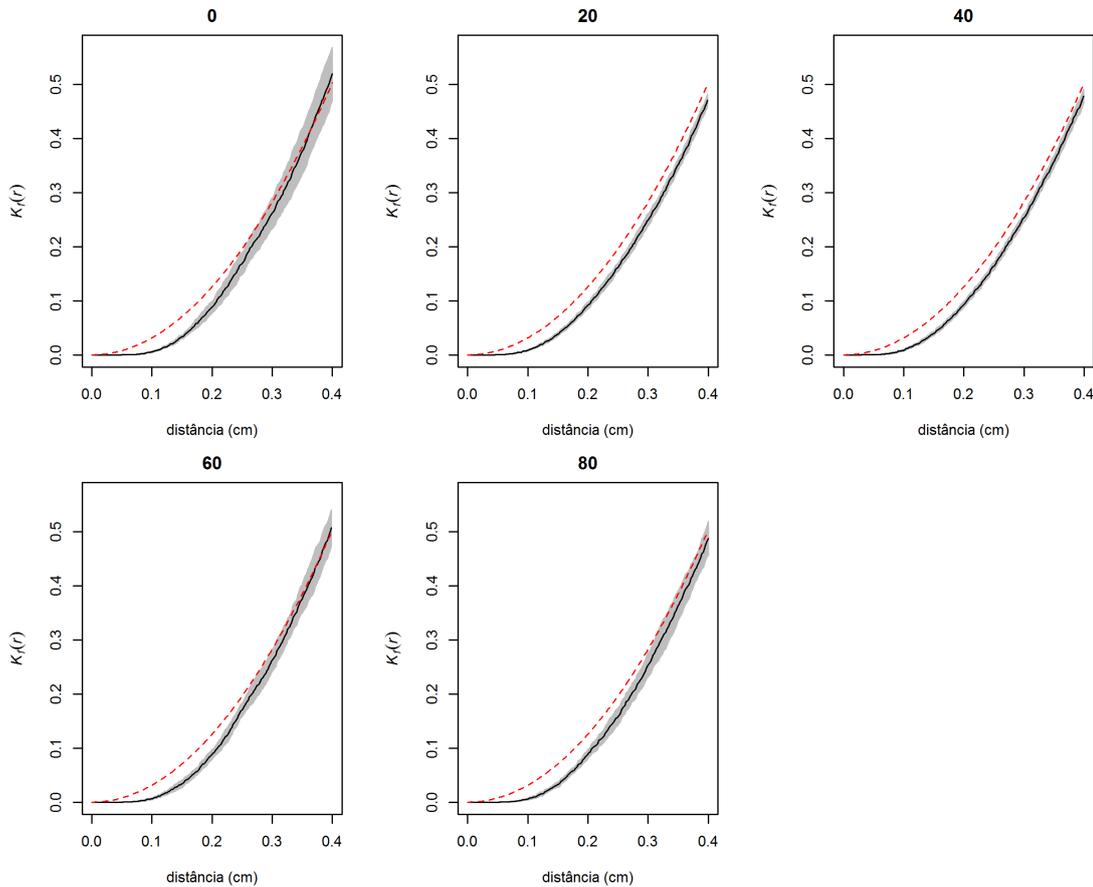
Figura 3.2 – Histogramas referente aos diâmetros dos alvéolos de pães franceses por meio dos grupos



função é (ou quase é) zero para todos os grupos; em decorrência do tamanho mínimo para se encontrar um alvéolo, parece haver uma certa regularidade presente. A partir de 0,1 cm até aproximadamente 0,4 cm há evidências para rejeitar a hipótese nula H_0 , as funções diferem da aleatoriedade espacial para algum h . A função K_f estimada devido ao fato de estar abaixo da função teórica, representa a existência de inibição das marcas, ou seja, os diâmetros dos alvéolos são menores que o diâmetro médio para alvéolos próximos.

Após o diagnóstico da presença da dependência entre a marca e o ponto em todos os grupos, pode-se continuar a análise, em que o principal interesse está em realizar uma comparação global desta dependência. Assim, a partir da metodologia proposta neste trabalho, testou-se a hipótese de igualdade das funções K_f , contra a hipótese alternativa de que uma diferença possa ser detectada. A estatística do teste $\bar{T}_f = 26,551$, obteve-se o valor- $p = 0,037$ para uma distância de até 0,86 centímetros. Portanto, há indícios para rejeitar H_0 , ou seja, pelo menos um grupo difere dos demais ao que se refere a dependência de marcas.

Figura 3.3 – Função K_f agrupada para cada grupo de tempo de congelamento de pães franceses



Com intuito de aprimorar a análise e tentar detectar qual tempo difere significativamente dos demais, aplicou-se novamente o teste proposto com todas as combinações possíveis dos grupos. Nesta análise verificou-se quais grupos diferem entre si 2 a 2 (Tabela 3.3).

Tabela 3.3 – Estatística do teste \bar{T}_f e seus respectivos valores- p para comparação 2 a 2 com 99 permutações

Comparação entre grupos	\bar{T}_f	Valor - p
0-20	8,155	0,03
0-40	5,063	0,04
0-60	0,868	0,45
0-80	2,860	0,12
20-40	0,498	0,60
20-60	5,365	0,01
20-80	2,085	0,19
40-60	3,526	0,02
40-80	0,911	0,33
60-80	0,951	0,35

Os tempos de congelamento 0 – 20; 0 – 40; 20 – 60 e 40 – 60 diferem entre si, com relação a dependência das marcas (valor- $p < 0,05$). Os tempos 0 e 60 diferem dos tempos 20 e 40. A maior diferença detectada foi entre os tempos 0 e 20 com $\bar{T}_f = 8,155$. Porém, os tempos 20 e 40 são considerados estatisticamente iguais no que diz respeito à dependência da marca com o evento.

Após 60 e 80 dias de armazenamento da massa congelada os pães apresentam alterações bioquímicas e perdem a capacidade de manter a estrutura alveolar (ALMEIDA; CHANG, 2014; BÁRCENAS; ROSELL, 2007; RIBOTTA et al., 2004).

Um dos motivos que levam os pesquisadores a utilizarem a microscopia é a possibilidade de contabilizar todos os eventos independentes do seu tamanho. Porém, limita-se a avaliação a uma área específica da superfície estudada. No método utilizado optou-se por avaliar toda a área do miolo e, com isso, sujeitou-se à não detecção de valores muito pequenos pela baixa detecção do software e, por isso, foram ignorados. Porém, observa-se uma boa capacidade de interpretação do teste usado, pois ele detectou aumento dos alvéolos e da sua uniformidade após 20 e 40 dias de armazenamento, além da desuniformidade dos alvéolos nos demais tempos.

Motivados por aplicações em panificação, para o presente trabalho, foram desenvolvidos métodos no contexto das posições de alvéolos em massa de pão congelada. No entanto, os processos são geralmente aplicáveis sempre que a questão para investigar for colocada em termos de padrões de pontos replicados com marcas quantitativas. Por exemplo altura de árvores, diâmetros de células, etc. Para melhor avaliação do método proposto, sugere-se que em trabalhos futuros sejam realizados estudos do poder do teste proposto e o desenvolvimento de um teste que apresente o melhor tempo de congelamento.

4 CONCLUSÃO

A partir do erro do tipo I pode-se dizer que o teste proposto apresenta bom desempenho na situação simulada, o que pode-se considerar um bom indicativo, ao ser aplicado para verificar diferenças entre grupos de padrões de pontos replicados com marcas quantitativas. A metodologia proposta é capaz de detectar diferenças e semelhanças nas variações da distribuição alveolar que justifiquem as diferenças entre os grupos (tempos de congelamento). Os alvéolos apresentam uma inibição da marca em todos os grupos, e por meio do teste proposto, é detectada diferença nessa inibição através da abordagem gráfica.

Agradecimentos

Os autores agradecem à CAPES (Comissão de Aperfeiçoamento de Pessoal do Nível Superior), pelo suporte financeiro.

REFERÊNCIAS

- ALMEIDA, E. L.; CHANG, Y. K. Influence of different enzymes during the frozen storage of pre-baked french bread elaborated with whole-wheat flour. **Journal of Food Processing and Preservation**, v. 38, n. 3, p. 737–748, 2014.
- ASSUNÇÃO, R. M. Estatística espacial com aplicações em epidemiologia, economia e sociologia. **São Carlos: Associação Brasileira de Estatística**, v. 131, 2001.
- BADDELEY, A.; GILL, R. D. Kaplan-meier estimators of distance distributions for spatial point processes. **The Annals of Statistics**, p. 263–292, 1997.
- BADDELEY, A. et al. Analysis of a three-dimensional point pattern with replication. **Applied Statistics**, p. 641–668, 1993.
- BADDELEY, A.; RUBAK, E.; TURNER, R. **Spatial point patterns: methodology and applications with R**. London: CRC Press, 2015.
- BADDELEY, A. J.; MØLLER, J.; WAAGEPETERSEN, R. Non-and semi-parametric estimation of interaction in inhomogeneous point patterns. **Statistica Neerlandica**, Wiley Online Library, v. 54, n. 3, p. 329–350, 2000.
- BAILEY, T. C.; GATRELL, A. C. **Interactive spatial data analysis**. Harlow:Longman: Longman Scientific & Technical Essex, 1995. v. 413.
- BÁRCENAS, M. E.; ROSELL, C. M. Different approaches for increasing the shelf life of partially baked bread: low temperatures and hydrocolloid addition. **Food Chemistry**, v. 100, n. 4, p. 1594–1601, 2007.
- BELL, M. L.; GRUNWALD, G. K. Mixed models for the analysis of replicated spatial point patterns. **Biostatistics**, Oxford University Press, v. 5, n. 4, p. 633–648, 2004.
- BESAG, J. Contribution to the discussion of dr. ripleys paper. **Journal of the Royal Statistical Society. Series B**, v. 39, n. 2, p. 193–195, 1977.
- BESBES, E. et al. Effect of baking conditions and storage with crust on the moisture profile, local textural properties and staling kinetics of pan bread. **LWT-Food Science and Technology**, Elsevier, v. 58, n. 2, p. 658–666, 2014.
- BICALHO, C. C. et al. Alveolar distribution of french type bread using taro mucilage. **Brazilian Journal of Food Technology**, artigo submetido, 2018.
- CASTIGLIONI, G. L. et al. Tamanho dos alvéolos e aceitação de pães de forma enriquecidos com farelo de mandioca. **Pesquisa Agropecuária Tropical (Agricultural Research in the Tropics)**, v. 2, n. 1, p. 10–1590, 2014.
- CHIU, S.; STOYAN, D. Estimators of distance distributions for spatial patterns. **Statistica Neerlandica**, v. 52, n. 2, p. 239–246, 1998.

- CRESSIE, N. **Statistics for spatial data**. Canada: John Wiley & Sons, 1993.
- CURTI, E. et al. Bread staling: effect of gluten on physico-chemical properties and molecular mobility. **LWT-Food Science and Technology**, v. 59, n. 1, p. 418–425, 2014.
- DIGGLE, P. A kernel method for smoothing point process data. **Applied statistics**, p. 138–147, 1985.
- DIGGLE, P. J. **Statistical analysis of spatial and spatio-temporal point patterns**. England, UK: CRC Press, 2013.
- DIGGLE, P. J.; LANGE, N.; BENES, F. M. Analysis of variance for replicated spatial point patterns in clinical neuroanatomy. **Journal of the American Statistical Association**, v. 86, n. 415, p. 618–625, 1991.
- DIGGLE, P. J.; MATEU, J.; CLOUGH, H. E. A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns. **Advances in Applied Probability**, v. 32, n. 02, p. 331–343, 2000.
- ECKARDT, J. et al. Long-term frozen storage of wheat bread and dough—effect of time, temperature and fibre on sensory quality, microstructure and state of water. **Journal of Cereal Science**, v. 57, n. 1, p. 125–133, 2013.
- ECKEL, S. et al. Modelling tree roots in mixed forest stands by inhomogeneous marked gibbs point processes. **Biometrical Journal**, Wiley Online Library, v. 51, n. 3, p. 522–539, 2009.
- ESTELLER, M. S.; LANNES, S. C. da S. Parâmetros complementares para fixação de identidade e qualidade de produtos panificados. **Ciênc. Tecnol. Aliment**, v. 25, n. 4, p. 802–806, 2005.
- GERVINI, D.; BAUR, T. J. Regression models for replicated marked point processes. **arXiv preprint arXiv:1705.06259**, 2017.
- GRABARNIK, P.; MYLLYMÄKI, M.; STOYAN, D. Correct testing of mark independence for marked point patterns. **Ecological Modelling**, Elsevier, v. 222, n. 23, p. 3888–3894, 2011.
- GUAN, Y. Tests for independence between marks and points of a marked point process. **Biometrics**, Wiley Online Library, v. 62, n. 1, p. 126–134, 2006.
- GUAN, Y.; AFSHARTOUS, D. R. Test for independence between marks and points of marked point processes: a subsampling approach. **Environmental and Ecological Statistics**, Springer, v. 14, n. 2, p. 101–111, 2007.
- GUTKOSKI, L. C. et al. Efeito de ingredientes na qualidade da massa de pão de forma congelada não fermentada durante o armazenamento. **Ciência e Tecnologia de Alimentos**, v. 25, n. 3, p. 460–467, 2005.
- HAHN, U. A studentized permutation test for the comparison of spatial point patterns. **Journal of the American Statistical Association**, v. 107, n. 498, p. 754–764, 2012.
- HUEN, J. et al. Confocal raman microscopy of frozen bread dough. **Journal of Cereal Science**, Elsevier, v. 60, n. 3, p. 555–560, 2014.

- ILLIAN, J. et al. **Statistical analysis and modelling of spatial point patterns**. England: John Wiley & Sons, 2008. v. 70.
- LANDAU, S.; EVERALL, I. P. Nonparametric bootstrap for k-functions arising from mixed-effects models with applications in neuropathology. **Statistica Sinica**, JSTOR, p. 1375–1393, 2008.
- LANDAU, S.; RABE-HESKETH, S.; EVERALL, I. P. Nonparametric one-way analysis of variance of replicated bivariate spatial point patterns. **Biometrical journal**, Wiley Online Library, v. 46, n. 1, p. 19–34, 2004.
- LASSOUED, N. et al. Granulometry of bread crumb grain: Contributions of 2d and 3d image analysis at different scale. **Food Research International**, Elsevier, v. 40, n. 8, p. 1087–1097, 2007.
- LIESHOUT, M. V. A j-function for marked point patterns. **Annals of the Institute of Statistical Mathematics**, Springer, v. 58, n. 2, p. 235, 2006.
- MORI, R. Tecnologia para p congelados. **Food Ingredients Brasil, Puratos**, n. 05, p. 45–46, 2008.
- MYLLYMÄKI, M.; SÄRKKÄ, A.; VEHTARI, A. Hierarchical second-order analysis of replicated spatial point patterns with non-spatial covariates. **Spatial Statistics**, Elsevier, v. 8, p. 104–121, 2014.
- PENTTINEN, A.; STOYAN, D.; HENTTONEN, H. M. Marked point processes in forest statistics. **Forest science**, v. 38, n. 4, p. 806–824, 1992.
- POLAKI, A. et al. Fiber and hydrocolloid content affect the microstructural and sensory characteristics of fresh and frozen stored bread. **Journal of Food Engineering**, v. 97, n. 1, p. 1–7, 2010.
- R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2017. Disponível em: <<https://www.R-project.org/>>.
- RAMÓN, P. et al. A new non-parametric method for analyzing replicated point patterns in ecology. **Ecography**, Wiley Online Library, v. 39, n. 11, p. 1109–1117, 2016.
- REDENBACH, C. et al. Anisotropy analysis of pressed point processes. **AStA Advances in Statistical Analysis**, Springer, v. 93, n. 3, p. 237–261, 2009.
- RIBOTTA, P. et al. Effect of emulsifier and guar gum on micro structural, rheological and baking performance of frozen bread dough. **Food Hydrocolloids**, v. 18, n. 2, p. 305–313, 2004.
- RIPLEY, B. D. The second-order analysis of stationary point processes. **Journal of applied probability**, v. 13, n. 02, p. 255–266, 1976.
- RIPLEY, B. D. Modelling spatial patterns. **Journal of the Royal Statistical Society.**, p. 172–212, 1977.
- SCANLON, M.; ZGHAL, M. Bread properties and crumb structure. **Food Research International**, Elsevier, v. 34, n. 10, p. 841–864, 2001.