

**PADRÕES ESPACIAIS NA EPIDEMIOLOGIA DA MORTE SÚBITA DOS
CITROS**

RELATÓRIO DAS ANÁLISES ESTATÍSTICAS DE UM TALHÃO

**Responsável técnico: Elias Teixeira Krainski
Supervisão: Dr. Paulo Justiniano Ribeiro Jr.**

**Curitiba
Estado do Paraná - Brasil
Janeiro de 2005**

SUMÁRIO

1	Introdução	3
2	Descrição e Validação dos dados	4
3	Análise por <i>quadrat counts</i>	6
3.1	Índice de Dispersão	6
3.2	Lei de Taylor	11
4	Métodos de Análise dos Vizinhos Próximos	13
4.1	Distância Mínima Média	13
4.2	Número Médio de Vizinhos Doentes	16
5	Análise de Processos Pontuais	19
5.1	Suavização por Kernel	19
5.2	Função K de Ripley	22
6	Conclusões	23

1 Introdução

Segundo [2] a Morte Súbita dos Citrus - MSC, provoca rápido definhamento e morte de plantas. A MSC ocorre, praticamente, apenas nas plantas enxertadas em limoeiro *Cravo*.

O Estado de São Paulo sozinho, responde por 80% da produção citrícola nacional, sendo que seus pomares apresentam baixa variabilidade com aproximadamente 85% das laranjeiras doces enxertadas sobre limoeiro cravo, o que gera elevada vulnerabilidade da cultura à ocorrência de novas epidemias. A importância desta doença é devida a representatividade desse porta-enxerto na citricultura brasileira, pela rusticidade, vigor à copa e resistência às deficiências hídrica.

O primeiro relato oficial da doença foi realizado em fevereiro de 2001 no município de Comendador Gomes, MG. Em 2002 a MSC atingiu o estado de São Paulo e vários trabalhos tem sido conduzidos para a descoberta de informações a respeito.

O *Citrus Sudden Death Virus* - CSDV, novo vírus da família Tymoviridae, tem sido associado à MSC em pesquisas feitas pela empresa de biotecnologia Alellyx. Em abril de 2004 a mesma empresa anunciou a descoberta da presença do patógeno da MSC. O CSDV foi encontrado em três insetos, dois deles sendo capazes de transmitir o vírus para as plantas, [1].

A análise de agregação é feita para avaliar o padrão da incidência de uma determinada doença no espaço, tendo-se padrões agregação regular, aleatório ou agregado. Segundo [3] esta análise é preliminar à uma modelagem estatística espacial.

Este relatório traz resultados de análises para um talhão com MSC, ilustrando métodos para avaliação dos padrões da doença que auxiliam o entendimento do comportamento de expansão da epidemia. São apresentados alguns dos métodos implementados e utilizados até o momento no escopo da colaboração entre o Departamento de Estatística da UFPR e do FUNDECITRUS. Os métodos e resultados obtidos tem como objetivo subsidiar discussões entre as equipes envolvidas e deverão ser estendidos para análise de todos os talhões monitorados pelo FUNDECITRUS.

As análises dos dados contidas neste relatório foram produzidas utilizando-se o pacote **Rcitrus**, que foi desenvolvido em R [5] para automatizar as análises dos dados de incidência de doenças em citrus. Desta forma no decorrer do texto são referenciados as funções do **Rcitrus** utilizadas nas análises.

2 Descrição e Validação dos dados

Os dados analisados são do talhão 303 da fazenda Vale Verde. Neste talhão há 960 plantas e foram feitas 25 avaliações, entre as os dias 01/08/2001 e 09/05/2003.

Os códigos utilizados para identificar o *status* da doença foram 0, 1, 2, e 3, indicando, respectivamente: planta sadia, estado inicial da doença, estado avançado da doença e morte. Nos dados apareceram outros códigos não relacionados ao *status* da MSC, tais como: F - Falha, R - Replante, G - Gomose e O - Erro de digitação. Na Tabela 1, podemos ver o número de plantas e respectivo *status* em cada avaliação.

Tabela 1: Número de Plantas com Código Atribuído em Cada Avaliação

Data da Avaliação	0	1	2	3	F	G	O	R
01/08/2001	923	17	5	1	4	0	0	10
08/08/2001	920	20	5	1	4	0	0	10
16/08/2001	873	45	18	10	4	0	0	10
22/08/2001	856	59	18	13	4	0	0	10
29/08/2001	847	68	18	13	4	0	0	10
06/09/2001	794	119	18	15	4	0	0	10
12/09/2001	707	205	19	15	4	0	0	10
25/09/2001	698	193	35	15	4	0	5	10
06/10/2001	646	251	34	15	4	0	0	10
10/10/2001	626	268	36	15	4	0	1	10
07/11/2001	478	403	44	19	4	0	2	10
06/12/2001	424	438	55	29	4	0	0	10
08/01/2002	369	477	55	45	4	0	0	10
12/02/2002	357	460	69	60	4	0	0	10
15/03/2002	326	488	70	62	4	0	0	10
05/04/2002	308	492	80	66	4	0	0	10
24/04/2002	225	500	108	113	4	0	0	10
08/05/2002	168	554	108	116	4	0	0	10
03/06/2002	164	554	112	116	4	0	0	10
05/08/2002	121	385	108	333	3	0	0	10
10/01/2003	69	223	110	544	3	1	0	10
10/02/2003	68	217	98	562	3	2	0	10
06/03/2003	69	211	100	566	3	0	0	11
14/04/2003	27	125	20	774	3	0	0	11
09/05/2003	27	125	20	774	3	0	0	11

Às plantas com códigos *G* e *O* foram atribuídos o *status* de sadias, usando a função `change.code()`. As plantas replantadas e os locais onde ocorreram falhas foram desconsiderados das análises. Para identificá-las e removê-las utilizou-se a função `select.code()`.

O *status* da MSC não regride no tempo e essa informação foi utilizada para fazer uma validação temporal nos dados. Foram encontrados nove erros que foram corrigidos usando a função `valid.time.citrus()`. Os dados originais com erros e após validados estão na Tabela 2.

Na primeira avaliação 23 plantas (2.43%) estavam infectadas com MSC, sendo que apenas uma planta estava morta (0.001%). Em menos de dois anos, 646 dias, 918 plantas estavam infectadas (97.14%) com 773 já mortas pela MSC (81.80%).

Tabela 2: Dados das nove plantas com sequência temporal errada, antes e depois da validação temporal.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
0	0	0	0	0	0	1	1	1	1	1	1	1	3	1	1	1	1	1	2	2	2	2	3	3
0	0	0	0	0	1	0	0	1	1	1	1	1	1	1	2	2	2	2	3	3	3	3	3	3
0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	3	3
0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	3	3
0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	1	1	1	2	3	3	3	3	3
0	0	0	0	0	1	1	2	1	1	1	1	1	1	1	1	2	2	2	3	3	3	3	3	3
1	0	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
0	1	0	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3
0	0	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	3	3	3	3
0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	3	3
0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	2	2	2	2	3	3	3	3	3	3
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	3	3
0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	3	3
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	2	3	3	3	3	3
0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	2	2	2	3	3	3	3	3	3
1	1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
0	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	3	3	3	3

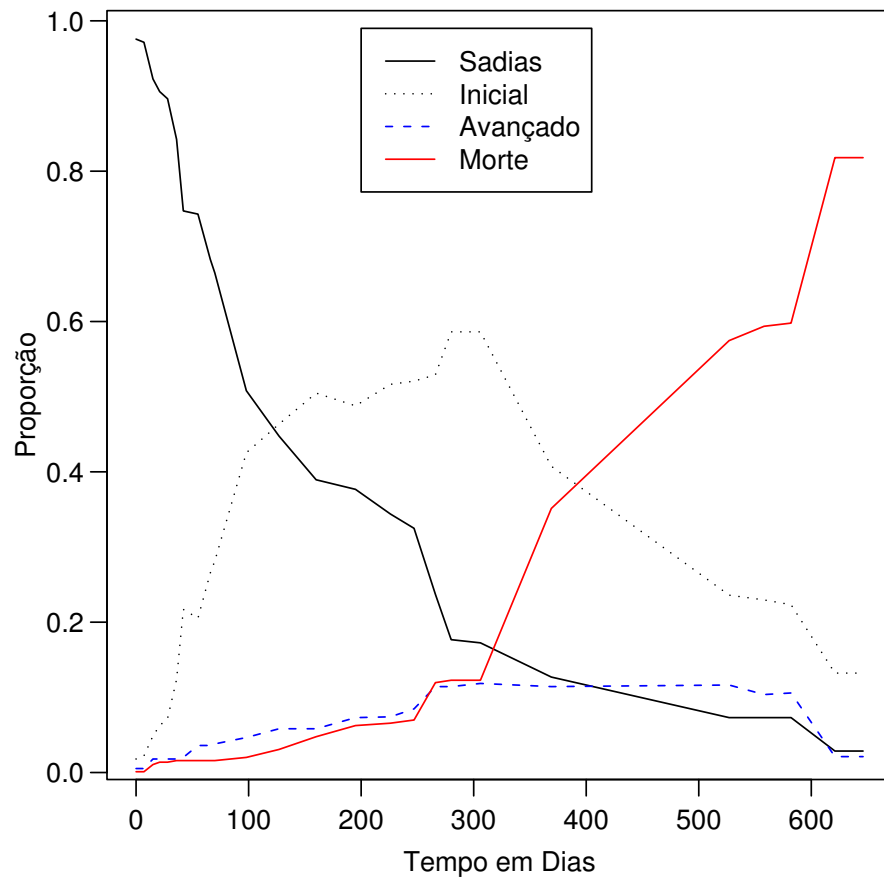


Figura 1: Evolução temporal do estágio da Morte Súbita no talhão

3 Análise por *quadrat counts*

Define-se *quadrat* uma área retangular de plantas abrangendo determinado número de linhas l e de plantas na linha c , o número de plantas em um *quadrat* é $n = l * c$.

Na análise por *quadrat* podemos ter *quadrats fixos* ou *quadrats aleatórios*. Em *quadrats fixos*, o talhão é subdividido em *quadrat* e é feita a análise todos os dados, tendo-se um número determinístico de q . Usando *quadrats aleatórios* os *quadrat* são selecionados de qualquer parte do talhão e é especificado um número de *quadrat*.

O Índice de Dispersão e a Lei de Taylor, são análises do padrão espacial feitas a partir da variância observada em um conjunto de dados e da variância esperada segundo um modelo [4].

A variância observada para o número de plantas doentes por *quadrat* é dada por

$$V_o(Y) = \sum_i (Y_i - np)^2 / (N - 1) \quad (1)$$

onde Y_i é o número de plantas doentes no i -ésimo *quadrat*, n é o número de plantas por *quadrat*, p é a proporção de plantas doentes no talhão e N o número de *quadrats* avaliados.

A variância teórica do número de plantas doentes por *quadrat* considerando um modelo Binomial é dada por

$$V_t(Y) = np(1 - p) \quad (2)$$

No modelo Poisson, a variância teórica é a média ponderada, dada por:

$$V_t(Y) = E(Y) = \sum_i Y_i n_i / \sum_i n_i, \quad (3)$$

onde n_i é o número de plantas no i -ésimo *quadrat*.

No modelo binomial há uma restrição no tamanho dos *quadrat*, que devem ter o mesmo número de plantas. No modelo Poisson a restrição é que o número de plantas deve ser suficientemente grande de forma que a aproximação da binomial pela Poisson seja satisfatória e portanto para que os resultados sejam válidos.

3.1 Índice de Dispersão

O índice de dispersão é dado por

$$D = V_o / V_t, \quad (4)$$

e a quantidade $(N - 1)D$ tem distribuição χ^2_{N-1} .

Para a análise do índice de dispersão por *quadrat counts*, foi utilizado a função `disp.quadrats()`. Esta função usa internamente as funções `disp.index()`, para o cálculo do índice de dispersão baseado em um vetor de contagens; e a função `quadrat.count()`, utilizada para efetuar a contagem de plantas doentes nos *quadrat*, podendo ser utilizada em outro métodos de análises baseados em *quadrat counts*.

A função `disp.quadrats()` foi utilizada para a análise por *quadrat counts* para os modelos binomial e de Poisson. Pode-se selecionar todos os *quadrat* possíveis ou selecionar aleatoriamente com reposição, sendo possível, descartar os *quadrat* incompletos. A análise pode ser feita para todas as avaliações feitas no talhão e para diversos tamanhos de *quadrat* ao mesmo tempo, retornando-se uma lista de tabelas com os resultados.

Na Tabela 3, tem-se os resultados da avaliação do padrão espacial usando o Índice de Dispersão binomial considerando *quadrat* fixos de tamanho 3×3 .

Na Tabela 4, temos os resultados da avaliação do padrão espacial usando o Índice de Dispersão binomial considerando 100 *quadrat* aleatórios de tamanho 3×3 .

Na Tabela 5, temos os resultados da avaliação do padrão espacial usando o Índice de Dispersão Poisson considerando *quadrat* fixos de tamanho 3×3 .

Na Tabela 6, temos os resultados da avaliação do padrão espacial usando o Índice de Dispersão Poisson considerando 100 *quadrat* aleatórios de tamanho 3×3 .

Observa-se que o Índice de Dispersão binomial detectou mais agregação enquanto que o Índice de Dispersão Poisson detecta menos o padrão agregado, detectando mais o padrão regular. Isso pode ser apontado como uma falha do método, que é baseado em aproximações.

Tabela 3: Análise do Índice de Dispersão usando o modelo binomial com quadrats fixos.

Avaliação	n	N	nN	p	obs.var	theor.var	index	p.value	conc
Av1	9	85	765	0.02614	0.00372	0.00283	1.3141	0.02839	Agregado
Av2	9	85	765	0.03137	0.00430	0.00338	1.2720	0.04697	Agregado
Av3	9	85	765	0.08105	0.01158	0.00828	1.3991	0.00925	Agregado
Av4	9	85	765	0.10065	0.01723	0.01006	1.7133	0.00005	Agregado
Av5	9	85	765	0.11111	0.01764	0.01097	1.6071	0.00035	Agregado
Av6	9	85	765	0.16340	0.02575	0.01519	1.6951	0.00007	Agregado
Av7	9	85	765	0.26405	0.02880	0.02159	1.3338	0.02217	Agregado
Av8	9	85	765	0.26667	0.02916	0.02173	1.3420	0.01995	Agregado
Av9	9	85	765	0.32680	0.03508	0.02444	1.4352	0.00552	Agregado
Av10	9	85	765	0.34510	0.03587	0.02511	1.4283	0.00610	Agregado
Av11	9	85	765	0.49804	0.04515	0.02778	1.6256	0.00026	Agregado
Av12	9	85	765	0.55817	0.05055	0.02740	1.8448	0.00000	Agregado
Av13	9	85	765	0.62092	0.05476	0.02615	2.0938	0.00000	Agregado
Av14	9	85	765	0.63399	0.05109	0.02578	1.9817	0.00000	Agregado
Av15	9	85	765	0.67320	0.05272	0.02444	2.1567	0.00000	Agregado
Av16	9	85	765	0.69412	0.05318	0.02359	2.2541	0.00000	Agregado
Av17	9	85	765	0.76209	0.04090	0.02015	2.0304	0.00000	Agregado
Av18	9	85	765	0.82876	0.02926	0.01577	1.8558	0.00000	Agregado
Av19	9	85	765	0.83399	0.02840	0.01538	1.8462	0.00000	Agregado
Av20	9	85	765	0.87843	0.02370	0.01187	1.9973	0.00000	Agregado
Av21	9	85	765	0.93464	0.01420	0.00679	2.0915	0.00000	Agregado
Av22	9	85	765	0.93464	0.01420	0.00679	2.0915	0.00000	Agregado
Av23	9	85	765	0.93464	0.01420	0.00679	2.0915	0.00000	Agregado
Av24	9	85	765	0.97516	0.00423	0.00269	1.5703	0.00066	Agregado
Av25	9	85	765	0.97516	0.00423	0.00269	1.5703	0.00066	Agregado

Tabela 4: Análise do Índice de Dispersão usando o modelo binomial com 100 quadrats aleatórios.

Avaliação	n	N	nN	p	obs.var	theor.var	index	p.value	conc
Av1	9	100	900	0.02556	0.00321	0.00277	1.1587	0.13363	Aleatorio
Av2	9	100	900	0.02556	0.00495	0.00277	1.7897	0.00000	Agregado
Av3	9	100	900	0.07000	0.01313	0.00723	1.8156	0.00000	Agregado
Av4	9	100	900	0.07889	0.01604	0.00807	1.9861	0.00000	Agregado
Av5	9	100	900	0.11556	0.01968	0.01136	1.7333	0.00001	Agregado
Av6	9	100	900	0.14333	0.02501	0.01364	1.8335	0.00000	Agregado
Av7	9	100	900	0.26222	0.03979	0.02150	1.8509	0.00000	Agregado
Av8	9	100	900	0.27778	0.03404	0.02229	1.5273	0.00057	Agregado
Av9	9	100	900	0.35333	0.04324	0.02539	1.7033	0.00002	Agregado
Av10	9	100	900	0.33111	0.03915	0.02461	1.5910	0.00017	Agregado
Av11	9	100	900	0.45444	0.04367	0.02755	1.5853	0.00019	Agregado
Av12	9	100	900	0.55111	0.06258	0.02749	2.2767	0.00000	Agregado
Av13	9	100	900	0.61333	0.05973	0.02635	2.2667	0.00000	Agregado
Av14	9	100	900	0.61889	0.04795	0.02621	1.8296	0.00000	Agregado
Av15	9	100	900	0.64889	0.04507	0.02531	1.7805	0.00000	Agregado
Av16	9	100	900	0.66667	0.06135	0.02469	2.4848	0.00000	Agregado
Av17	9	100	900	0.75889	0.03792	0.02033	1.8653	0.00000	Agregado
Av18	9	100	900	0.82667	0.03225	0.01592	2.0258	0.00000	Agregado
Av19	9	100	900	0.83333	0.03280	0.01543	2.1252	0.00000	Agregado
Av20	9	100	900	0.86667	0.02394	0.01284	1.8648	0.00000	Agregado
Av21	9	100	900	0.92333	0.01539	0.00787	1.9563	0.00000	Agregado
Av22	9	100	900	0.94333	0.01284	0.00594	2.1623	0.00000	Agregado
Av23	9	100	900	0.93333	0.01247	0.00691	1.8037	0.00000	Agregado
Av24	9	100	900	0.98222	0.00267	0.00194	1.3780	0.00756	Agregado
Av25	9	100	900	0.97667	0.00332	0.00253	1.3095	0.02103	Agregado

Tabela 5: Análise do Índice de Dispersão usando o modelo Poisson com quadrats fixos.

Avaliação	n	np	p	obs.var	theor.var	index	p.value	conc
Av1	96	850	0.02471	0.27797	0.22118	1.25676	0.04601	Agregado
Av2	96	850	0.02941	0.32095	0.26353	1.21790	0.07321	Aleatorio
Av3	96	850	0.08235	0.93641	0.73059	1.28171	0.03350	Agregado
Av4	96	850	0.10118	1.33640	0.89647	1.49074	0.00137	Agregado
Av5	96	850	0.11059	1.36799	0.98118	1.39423	0.00672	Agregado
Av6	96	850	0.16471	1.93509	1.46000	1.32540	0.01856	Agregado
Av7	96	850	0.26118	2.21718	2.32118	0.95520	0.60447	Aleatorio
Av8	96	850	0.26353	2.24570	2.34235	0.95874	0.59478	Aleatorio
Av9	96	850	0.32824	2.67534	2.90941	0.91955	0.69895	Aleatorio
Av10	96	850	0.34706	2.74201	3.07647	0.89128	0.76714	Aleatorio
Av11	96	850	0.50471	3.55692	4.47176	0.79542	0.92926	Aleatorio
Av12	96	850	0.56471	3.89475	5.00353	0.77840	0.94609	Aleatorio
Av13	96	850	0.62706	4.12359	5.55647	0.74212	0.97185	Regular
Av14	96	850	0.63882	3.87010	5.66235	0.68348	0.99217	Regular
Av15	96	850	0.67412	4.03072	5.98000	0.67403	0.99381	Regular
Av16	96	850	0.69412	4.04184	6.15882	0.65627	0.99612	Regular
Av17	96	850	0.76471	3.16818	6.78471	0.46696	1.00000	Regular
Av18	96	850	0.82824	2.37223	7.35059	0.32273	1.00000	Regular
Av19	96	850	0.83294	2.32138	7.39294	0.31400	1.00000	Regular
Av20	96	850	0.87882	1.89933	7.79882	0.24354	1.00000	Regular
Av21	96	850	0.93294	1.26880	8.28235	0.15319	1.00000	Regular
Av22	96	850	0.93294	1.26880	8.28235	0.15319	1.00000	Regular
Av23	96	850	0.93294	1.26880	8.28235	0.15319	1.00000	Regular
Av24	96	850	0.97059	0.62339	8.62000	0.07232	1.00000	Regular
Av25	96	850	0.97059	0.62339	8.62000	0.07232	1.00000	Regular

Tabela 6: Análise do Índice de Dispersão usando o modelo Poisson com 100 quadrats aleatórios.

Avaliação	n	np	p	obs.var	theor.var	index	p.value	conc
Av1	100	900	0.02556	0.25970	0.23	1.12912	0.17906	Aleatorio
Av2	100	900	0.02556	0.40111	0.23	1.74396	0.00001	Agregado
Av3	100	900	0.07000	1.06374	0.63	1.68847	0.00002	Agregado
Av4	100	900	0.07889	1.29889	0.71	1.82942	0.00000	Agregado
Av5	100	900	0.11556	1.59434	1.04	1.53302	0.00052	Agregado
Av6	100	900	0.14333	2.02616	1.29	1.57067	0.00025	Agregado
Av7	100	900	0.26222	3.22263	2.36	1.36552	0.00919	Agregado
Av8	100	900	0.27778	2.75758	2.50	1.10303	0.22711	Aleatorio
Av9	100	900	0.35333	3.50263	3.18	1.10145	0.23026	Aleatorio
Av10	100	900	0.33111	3.17131	2.98	1.06420	0.31224	Aleatorio
Av11	100	900	0.45444	3.53727	4.09	0.86486	0.82879	Aleatorio
Av12	100	900	0.55111	5.06909	4.96	1.02199	0.42037	Aleatorio
Av13	100	900	0.61333	4.83798	5.52	0.87645	0.80511	Aleatorio
Av14	100	900	0.61889	3.88394	5.57	0.69730	0.99047	Regular
Av15	100	900	0.64889	3.65091	5.84	0.62516	0.99872	Regular
Av16	100	900	0.66667	4.96970	6.00	0.82828	0.89197	Aleatorio
Av17	100	900	0.75889	3.07182	6.83	0.44975	1.00000	Regular
Av18	100	900	0.82667	2.61253	7.44	0.35115	1.00000	Regular
Av19	100	900	0.83333	2.65657	7.50	0.35421	1.00000	Regular
Av20	100	900	0.86667	1.93939	7.80	0.24864	1.00000	Regular
Av21	100	900	0.92333	1.24636	8.31	0.14998	1.00000	Regular
Av22	100	900	0.94333	1.04030	8.49	0.12253	1.00000	Regular
Av23	100	900	0.93333	1.01010	8.40	0.12025	1.00000	Regular
Av24	100	900	0.98222	0.21657	8.84	0.02450	1.00000	Regular
Av25	100	900	0.97667	0.26859	8.79	0.03056	1.00000	Regular

3.2 Lei de Taylor

Assume-se que a variância de dados de contagem de plantas doentes é proporcional a uma potência da média [4]:

$$v = a.m^b, \quad (5)$$

onde v é a variância observada, m é a média da densidade populacional e a e b são parâmetros.

As estimativas de a e b são obtidas através do modelo de regressão linear simples, fazendo;

$$\log(v) = \log(a) + b.\log(m) \quad (6)$$

Usando a função `Taylor.citrus()`, os coeficientes a e b foram estimados considerando as 25 avaliações feitas, usando *quadrat* 3×3 , fixos e aleatórios com 100 *quadrats*.

Na Tabela 7 pode-se observar as estimativas obtidas para os parâmetros da Lei de Taylor. O modelo binomial detectou agregação enquanto o modelo Poisson não detectou padrão agregado.

Tabela 7: Estimativas e Intervalo de Confiança dos Parâmetros da Lei de Taylor para Quadrats Fixos e Aleatórios.

Modelo/Coeficiente	Inferior	Estimado	Superior
Binomial Quadrats Fixos			
a	0.39942	0.82293	1.24643
b	0.96915	1.06383	1.15852
Binomial Quadrats Aleatorios			
a	0.50590	0.84997	1.19405
b	0.98406	1.06065	1.13723
Poisson Quadrats Fixos			
a	-0.31325	0.09282	0.49888
b	0.13324	0.38706	0.64088
Poisson Quadrats Aleatorios			
a	-0.28622	0.20942	0.70505
b	0.00537	0.31538	0.62539

Na Figura 3.2, produzida utilizando-se o método `plot` implementado no **Rcitrus** para o resultado da função `Taylor.citrus`, visualiza-se o logaritmo da variância esperada *versus* o logaritmo da variância observada. A linha tracejada em azul é o modelo ajustado e a linha pontilhada em vermelho é a reta sob padrão aleatório. Vê-se que há uma associação linear forte usando o modelo binomial. Usando o modelo de Poisson observa-se que não é adequado utilizar a Lei de Taylor e por isso os resultados para este modelo estão comprometidos.

Figura 2: Logaritmo da variância observada *versus* o logaritmo da variância esperada, considerando os modelos binomial com quadrats fixos (superior esquerdo), binomial com quadrats aleatórios (superior direito), Poisson com quadrats fixos (inferior esquerdo) e Poisson com quadrats aleatórios (inferior direito).

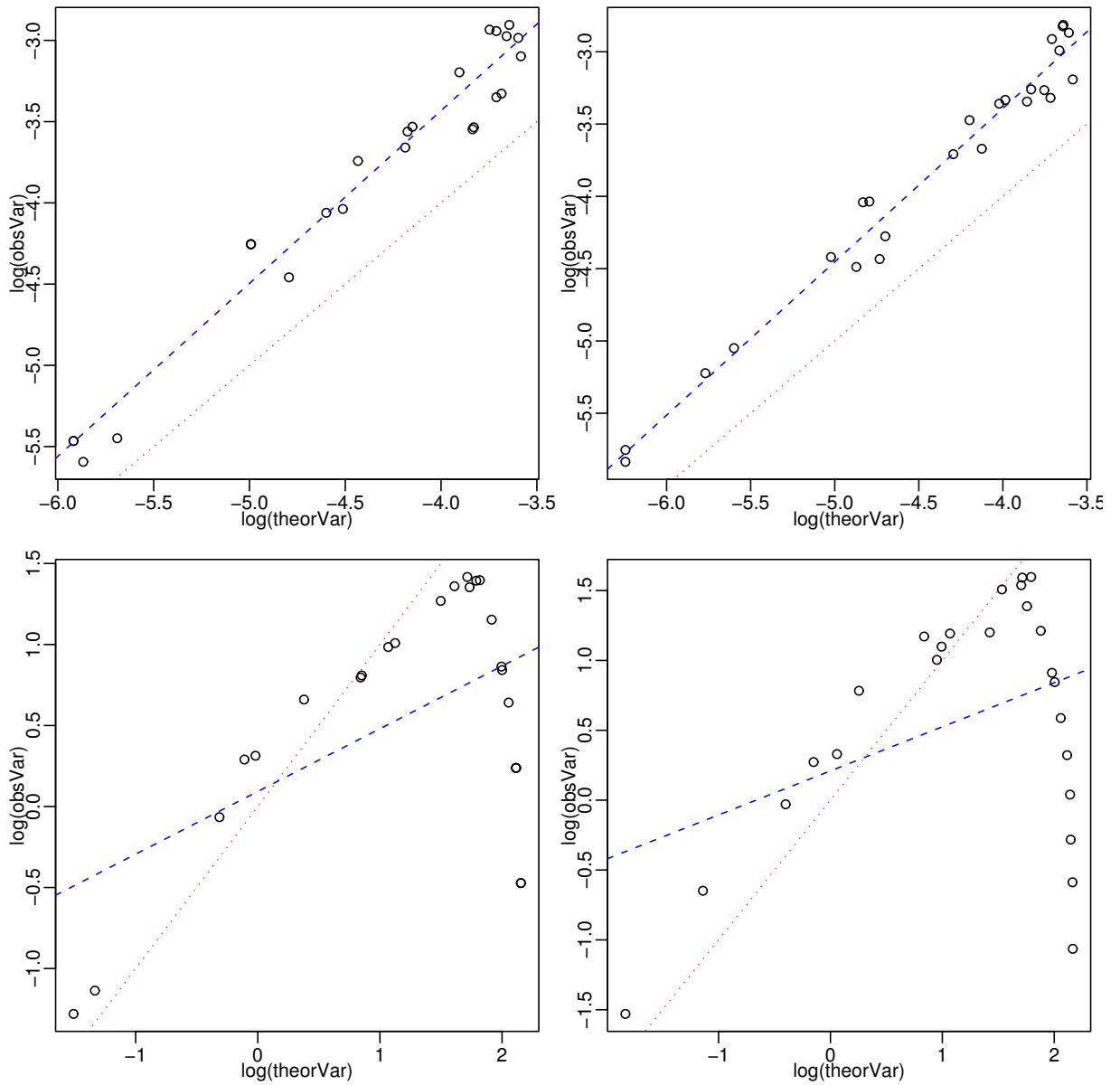


Tabela 8: Resultados do Método da Distância Mínima Média

Avaliação	Incidência	Observado	P.valor	Padrao
Av1	0.024339	11.6532	0.005	Agregado
Av2	0.028571	13.8915	0.025	Agregado
Av3	0.077249	9.6440	0.025	Agregado
Av4	0.094180	8.1599	0.005	Agregado
Av5	0.103704	7.8237	0.005	Agregado
Av6	0.157672	6.7249	0.005	Agregado
Av7	0.252910	5.9412	0.005	Agregado
Av8	0.257143	5.9383	0.005	Agregado
Av9	0.317460	5.5567	0.005	Agregado
Av10	0.335450	5.4860	0.005	Agregado
Av11	0.492063	4.9031	0.045	Agregado
Av12	0.552381	4.7239	0.125	Aleatorio
Av13	0.610582	4.5198	0.065	Aleatorio
Av14	0.623280	4.4897	0.060	Aleatorio
Av15	0.656085	4.4209	0.095	Aleatorio
Av16	0.675132	4.3753	0.090	Aleatorio
Av17	0.762963	4.2323	0.385	Aleatorio
Av18	0.823280	4.1440	0.450	Aleatorio
Av19	0.827513	4.1439	0.640	Aleatorio
Av20	0.873016	4.0976	0.795	Aleatorio
Av21	0.926984	4.0679	0.995	Aleatorio
Av22	0.926984	4.0679	1.000	Aleatorio
Av23	0.926984	4.0679	1.000	Aleatorio
Av24	0.971429	4.0267	0.990	Aleatorio
Av25	0.971429	4.0267	1.000	Aleatorio

4 Métodos de Análise dos Vizinhos Próximos

O padrão espacial pode ser detectado analisando as propriedades dos eventos próximos a cada evento. Pode-se avaliar uma estatística que incorpore a dependência espacial entre os eventos.

Para testar a hipótese de aleatoriedade espacial, é feito um teste de Monte Carlo, comparando o valor da estatística de teste observado nos dados com valores obtidos de dados simulados sob hipótese de aleatoriedade espacial completa.

Seja t_1 a estatística observada e $t_i : i = 2, \dots, s$ a i -ésima estatística obtida sob a hipótese de aleatoriedade. Seja $t_{(j)}$ a j -ésima estatística de ordem, então

$$P[t_1 = t_{(j)}] = s^{-1} : j = 1, \dots, s. \quad (7)$$

e para o cálculo do p-valor, seja $t_1 = t_{(k)}$, então o p-valor será k/s .

Para um teste ao nível de 5% de significância, $s = 100$ é adequado, [3].

4.1 Distância Mínima Média

A distância mínima é a distância entre uma planta doente e a planta doente mais próxima a ela. Calculando a distância mínima para cada planta doente em uma avaliação feita no talhão, pode-se obter a distância mínima média. No teste de Monte Carlo, rejeita-se a hipótese de padrão aleatório em favor da hipótese de padrão agregado quando a distância mínima média observada for menor que k/s distâncias simuladas.

O teste da distância mínima pode ser aplicado utilizando-se a função `mmdist.test` e os resultados podem ser vistos utilizando-se os métodos `summary`, `plot` e `hist`, implementados para os resultados dessa função.

O p-valor obtido pelo teste de Monte Carlo com 199 simulações mostra padrão agregado em todas as avaliações, Tabela 8.

Na Figura 4.1 observa-se os valores da distância mínima média sob hipótese de CSR.

Na Figura 4.1 em cada gráfico estão o histograma dos valores da distância mínima média observada simuladas. A linha vertical indica o valor observado.

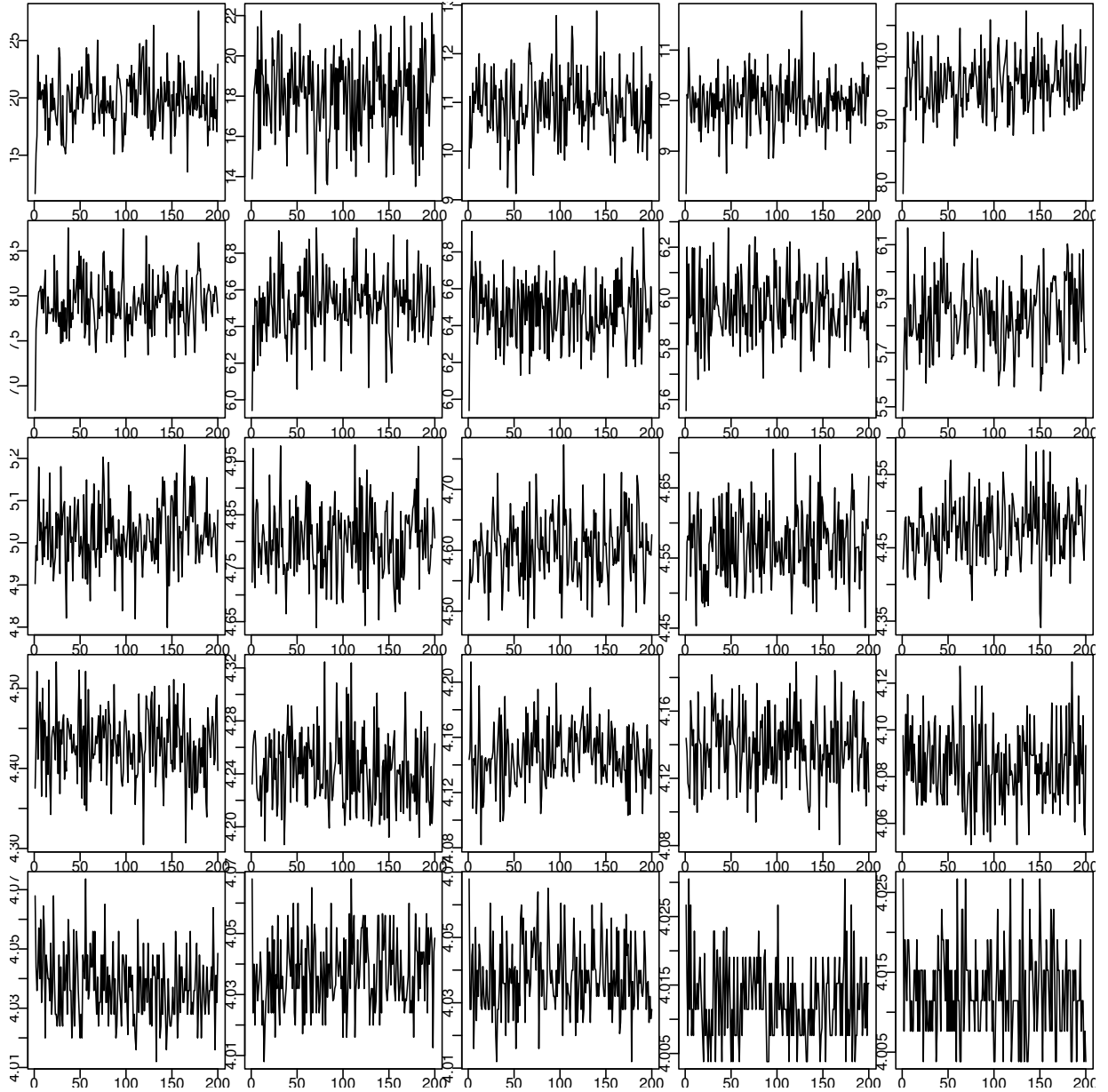


Figura 3: Distância Mínima Média Sob Hipótese de Aleatoriedade Espacial Completa

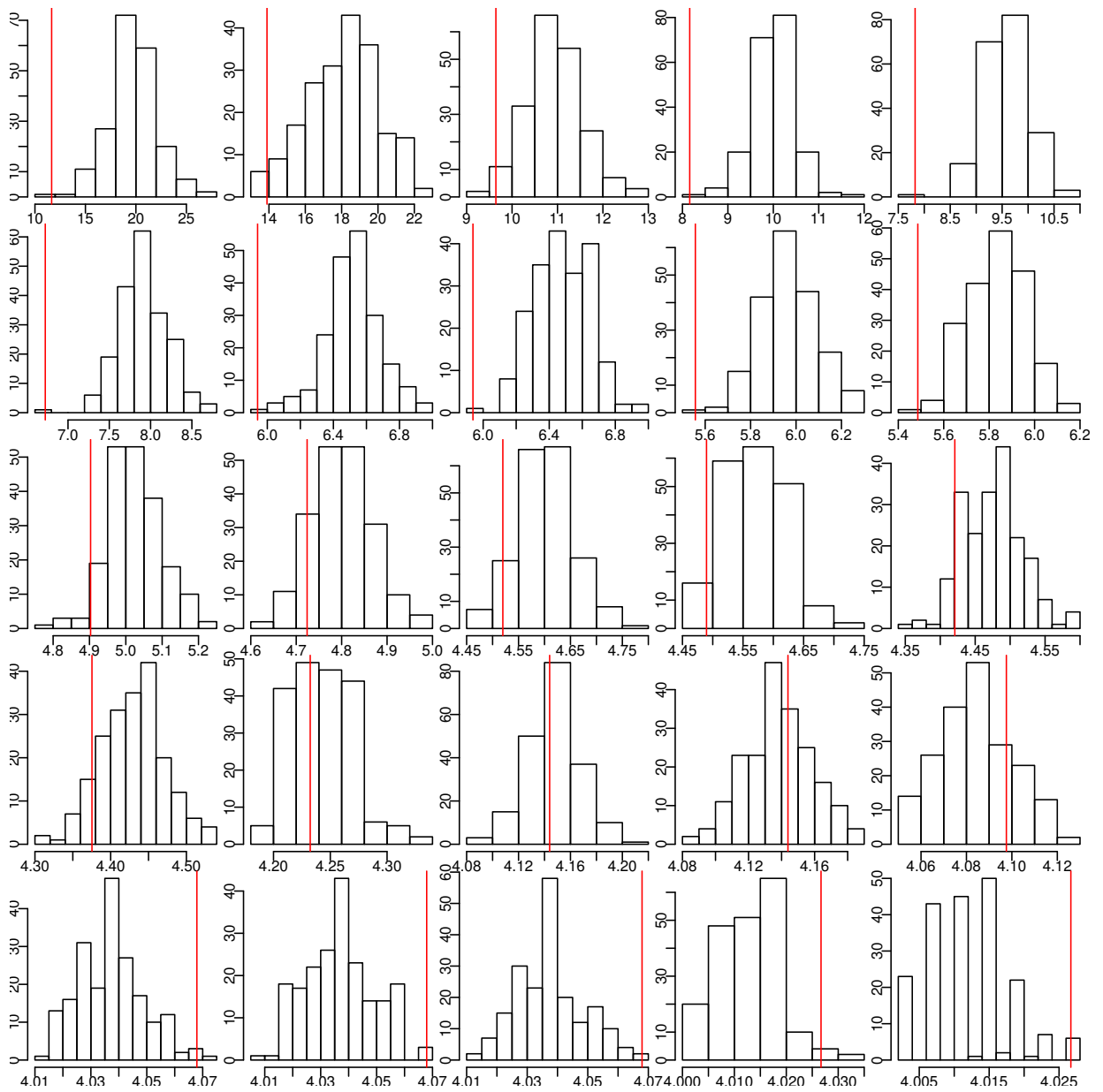


Figura 4: "Visualização dos Valores Simulados Sob Hipótese de Aleatoriedade Espacial Completa"

O padrão foi agregado na maioria das avaliações, Tabela 8, concordando com os resultados da análise do índice de dispersão do modelo binomial na análise por *quadrat*.

4.2 Número Médio de Vizinhos Doentes

O número de vizinhos doentes é o número de plantas doentes dentro de um raio em torno de cada planta doente. O número médio de vizinhos doentes é a média do número de vizinhos doentes para cada planta doente.

Aplicando-se um teste de Monte Carlo, se o padrão é agregado, espera-se que o número médio de vizinhos doentes seja maior nos dados observados que o número obtido em dados simulados sob a hipótese de CSR.

O teste do número de vizinhos doentes pode ser aplicado utilizando-se a função `neigh.test` e a melhor visualização dos resultados pode ser obtida utilizando-se os métodos `summary`, `plot` e `hist`, implementados para os resultados dessa função.

O número médio de vizinhos foi calculado para cada avaliação. O p-valor do teste de Monte Carlo com 199 simulações mostra padrão agregado em todas as avaliações, Tabela 9.

Tabela 9: Resultados do Método do Número Médio de Vizinhos Próximos

Avaliação	Incidencia	Observado	P.valor	Padrao
Av1	0.024339	2.8696	0.010	Agregado
Av2	0.028571	3.3333	0.005	Agregado
Av3	0.077249	7.6712	0.005	Agregado
Av4	0.094180	9.8652	0.005	Agregado
Av5	0.103704	9.6327	0.005	Agregado
Av6	0.157672	15.4631	0.005	Agregado
Av7	0.252910	20.9623	0.005	Agregado
Av8	0.257143	21.0370	0.005	Agregado
Av9	0.317460	26.0933	0.005	Agregado
Av10	0.335450	27.4069	0.005	Agregado
Av11	0.492063	39.8280	0.005	Agregado
Av12	0.552381	44.9119	0.005	Agregado
Av13	0.610582	49.9445	0.005	Agregado
Av14	0.623280	50.7199	0.005	Agregado
Av15	0.656085	52.9452	0.005	Agregado
Av16	0.675132	54.2947	0.005	Agregado
Av17	0.762963	58.5076	0.005	Agregado
Av18	0.823280	63.0283	0.005	Agregado
Av19	0.827513	63.3913	0.005	Agregado
Av20	0.873016	66.6279	0.005	Agregado
Av21	0.926984	70.7694	0.005	Agregado
Av22	0.926984	70.7694	0.005	Agregado
Av23	0.926984	70.7694	0.005	Agregado
Av24	0.971429	73.8126	0.005	Agregado
Av25	0.971429	73.8126	0.005	Agregado

Na Figura 4.2 estão plotados os valores do número de vizinhos doentes observado e os valores simulados sob hipótese CSR para cada avaliação, com 199 simulações.

Na Figura 4.2 observa-se que em todas as avaliações o número médio de vizinhos doentes observado é geralmente maior.

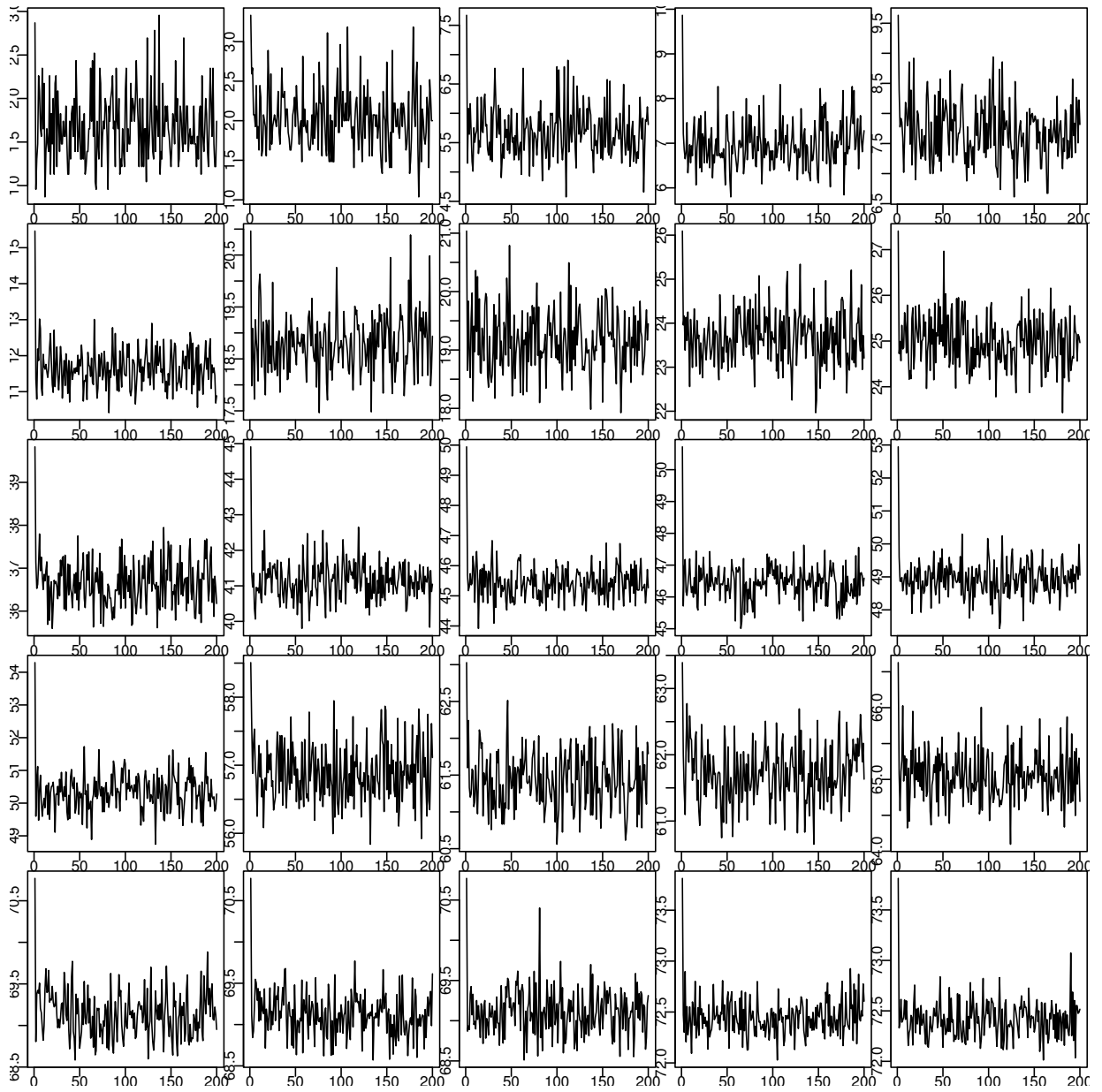


Figura 5: Número Médio de Vizinhos Doentes Sob Hipótese de Aleatoriedade Espacial Completa

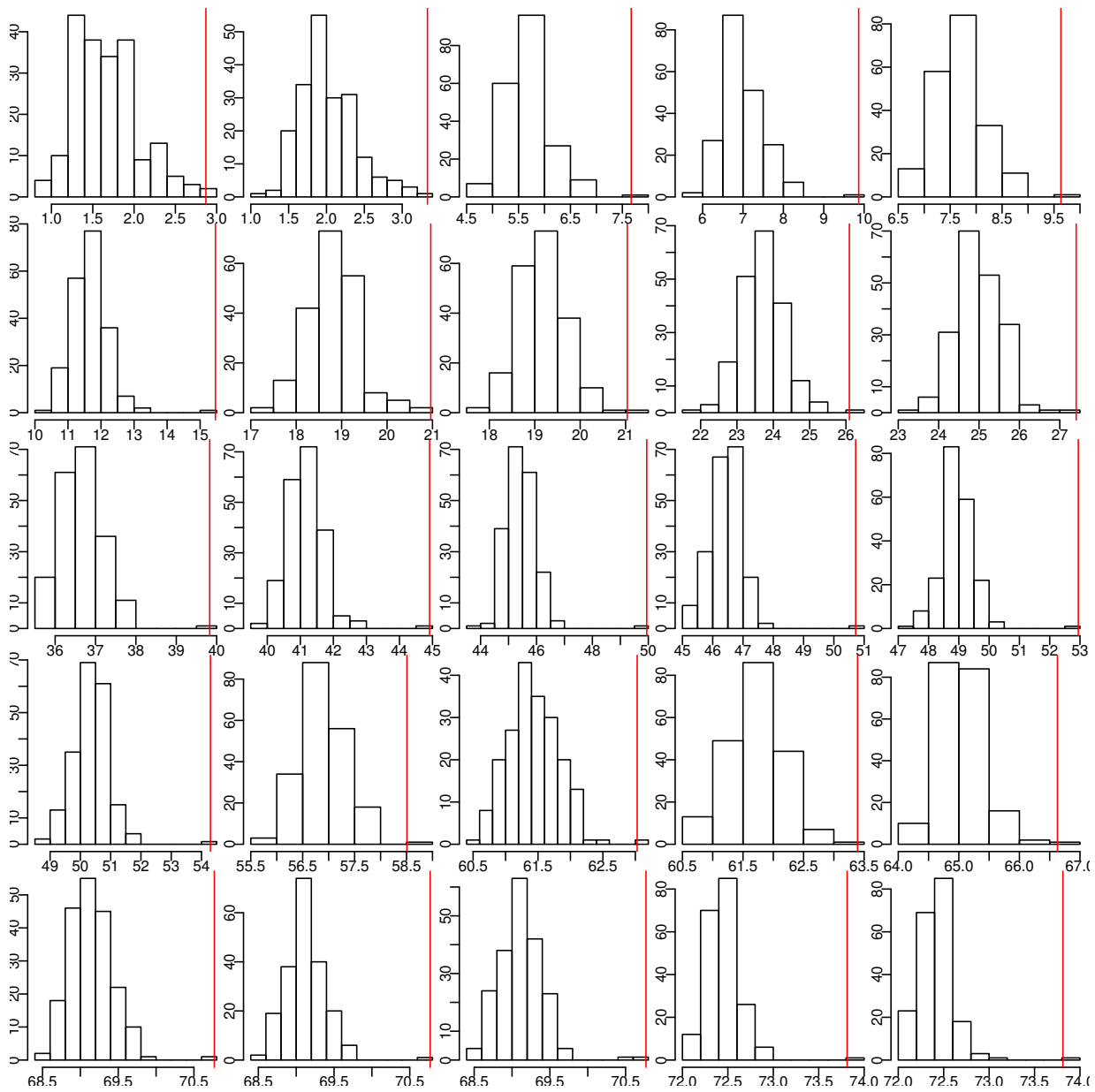


Figura 6: Visualização dos Valores Simulados Sob Hipótese de Aleatoriedade Espacial Completa

5 Análise de Processos Pontuais

A análise de processos pontuais é o estudo da ocorrência de eventos no espaço. Nessa abordagem, a localização do evento é estocástica. Na análise da MSC, o evento de interesse considerado foi a incidência no estado inicial e avançado e a morte.

Para detectar padrões espaciais, podemos inicialmente avaliar se a intensidade do processo varia no espaço. Neste trabalho utilizamos a suavização por *kernel*, para visualizar a intensidade do processo no talhão em vários momentos, e a função K de Ripley, para avaliar o padrão espacial.

5.1 Suavização por Kernel

A suavização consiste em ajustar uma função bidimensional com valores proporcionais ao número de eventos na área. O valor da função em cada ponto da área é o número de pontos dentro de uma região de influência ponderado pela distância, também chamado de intensidade do processo.

A partir de uma n eventos com coordenadas u_1, u_2, \dots, u_n , o estimador de intensidade *kernel estimator* é dado por:

$$\hat{\lambda}_\tau(u) = \frac{1}{\tau^2} \sum_{i=1}^n k\left(\frac{d(u_i, u)}{\tau}\right), d(u_i, u) \leq \tau, \quad (8)$$

onde τ é o raio de influência e k uma função de estimação com propriedades de suavização.

É comum usar funções de terceira ou quarta ordem, como

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

ou o estimador gaussiano:

$$k(h) = \frac{1}{2\pi\tau} \exp\left(-\frac{h^2}{2\tau^2}\right), \quad (9)$$

onde h é a distância entre o ponto e o evento observado.

Nos dados de MSC, pode-se aplicar o `kernel2d` utilizando a função `kernel2d.citrus` do **Rcitrus**. Na Figura 5.1 observamos a suavização da intensidade da ocorrência de MSC nas 25 avaliações. A escala de cores é única para todas as avaliações, permitindo observar melhor a evolução da incidência no talhão.

Na Figura 5.1 observamos a suavização da intensidade da ocorrência de MSC nas 25 avaliações, com escala de cores individual para cada uma das avaliações, permitindo observar melhor a evolução da incidência dentro do talhão e observar os focos de maior intensidade.

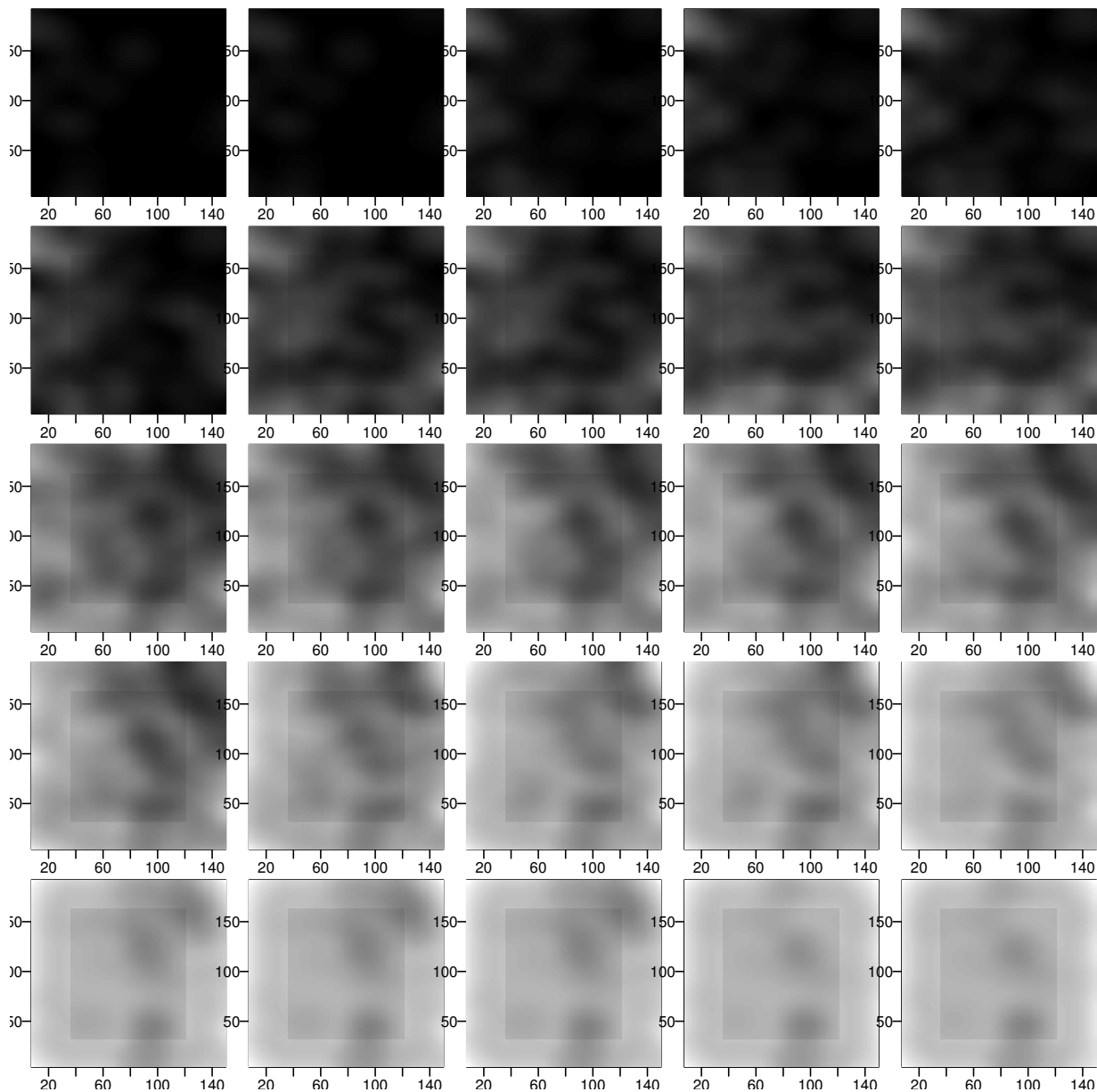


Figura 7: Intensidade da MSC nas 25 avaliações com escala global de cor.

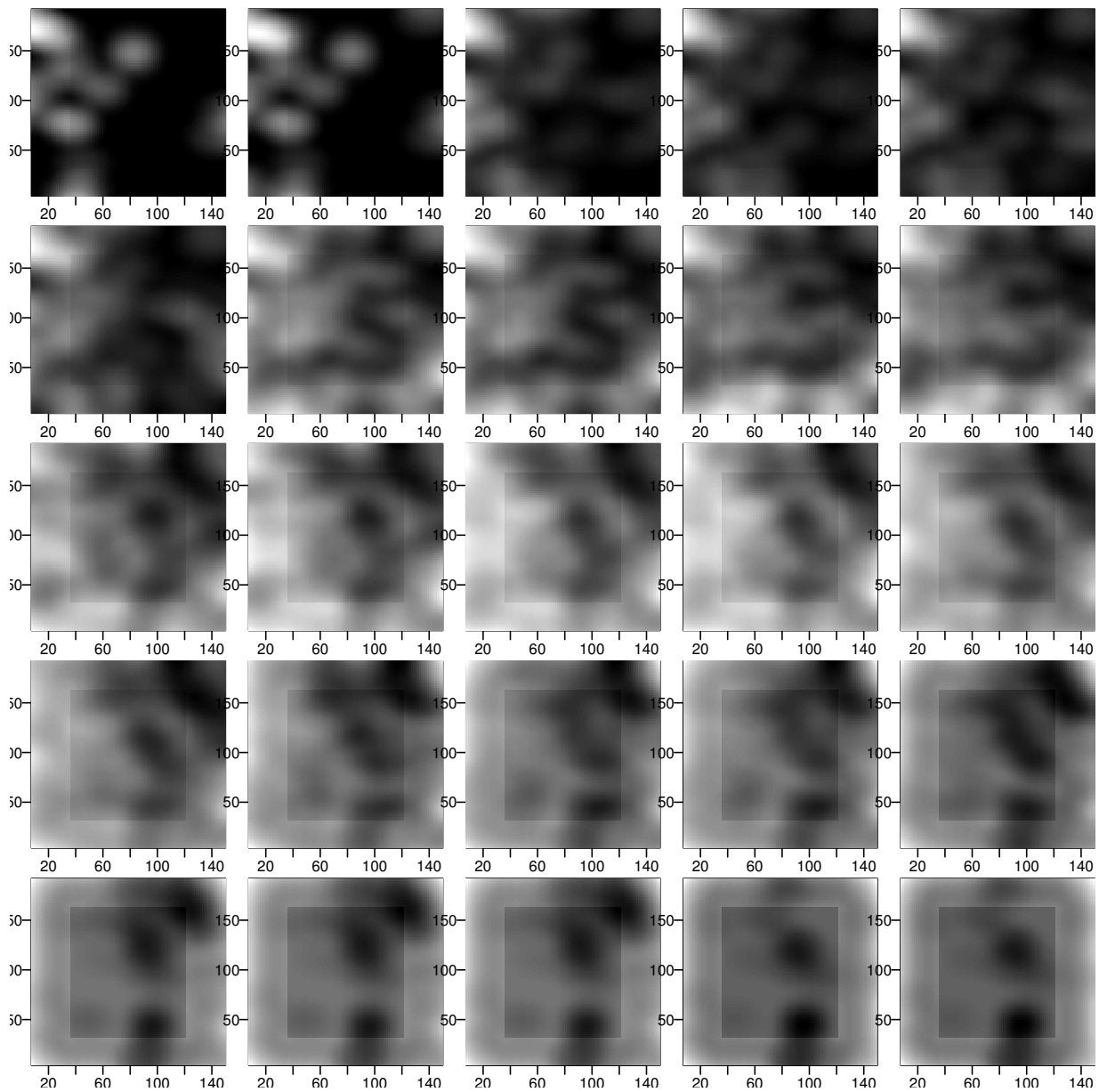


Figura 8: Intensidade da MSC nas 25 avaliações com escala individual de cor.

5.2 Função K de Ripley

A função K é uma medida de momento de segunda ordem reduzido. Para um processo univariado é definida como:

$$\lambda K(t) = E[N_0(t)] , \quad (10)$$

onde λ é o número médio de eventos por unidade de área, ou seja, o número de esperado de eventos a menos de uma distância t de um evento arbitrário.

O estimador não viesado, considerando o efeito de borda e substituindo o parâmetro desconhecido λ por $(n - 1) / |A|$ obtido por Ripley é:

$$\widehat{K}(t) = n(n - 1)^{-1} |A| \sum_{i=1}^n \sum_{j \neq i} w_{ij}^{-1} I_t(u_{ij}) \quad (11)$$

Para a estimação da função K para os dados de MSC, pode-se usar a função `Kenv.csr.citrus`.

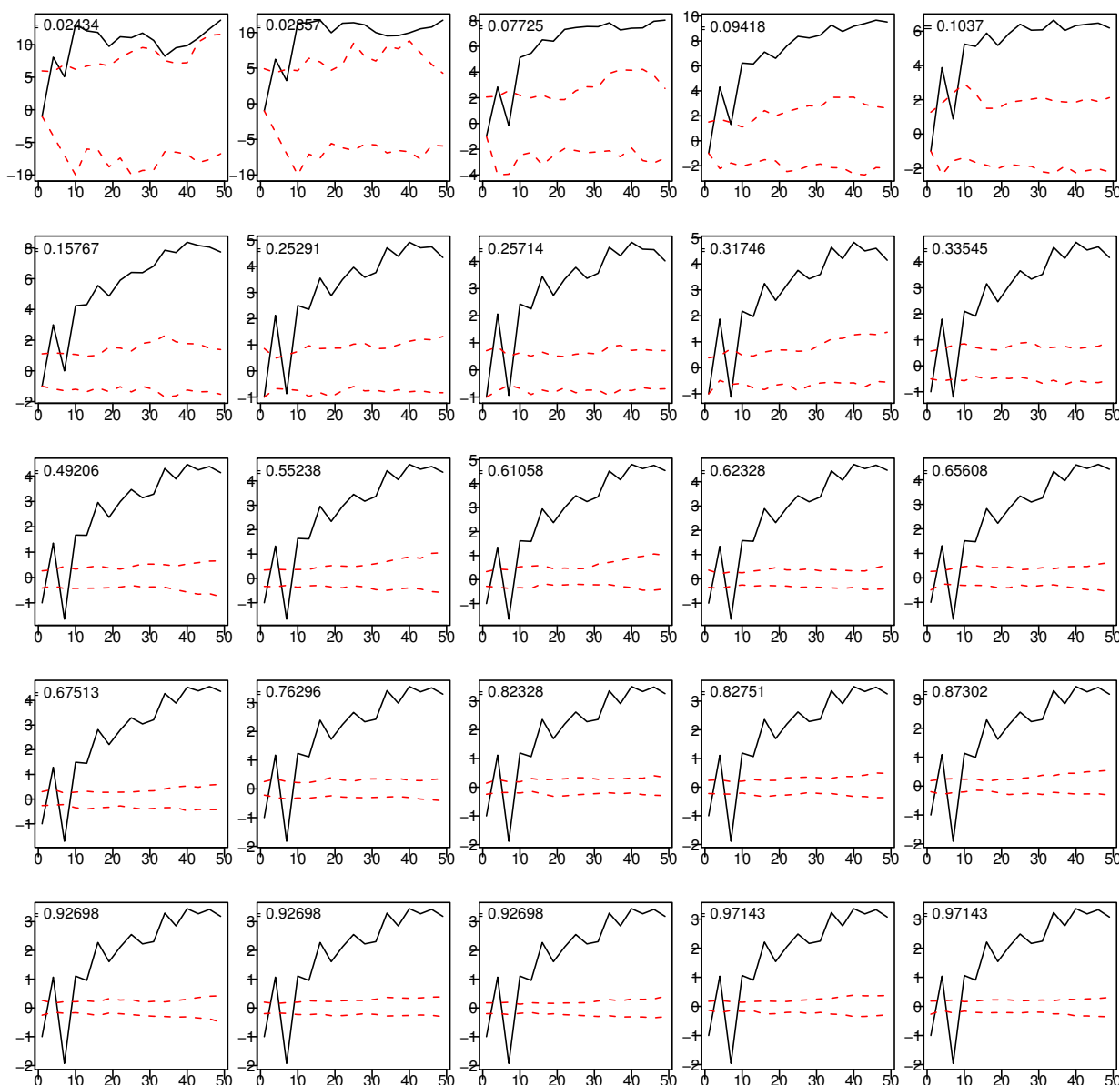


Figura 9: Função K de Ripley e envelopes simulados para as 25 avaliações.

6 Conclusões

Foi analisado o padrão espacial da MSC em 25 avaliações de um talhão usando-se de vários métodos estatísticos e constatou-se agregação pela maioria dos métodos em todas as avaliações.

O modelo Poisson na análise por *quadrat counts* foi o que mais detectou padrão regular e em outro extremo, o número de vizinhos foi o que detectou mais padrões agregados. O modelo binomial na análise por *quadrat counts* e a distância mínima média tiveram resultados parecidos.

Na abordagem por processos pontuais, o kernel permitiu visualizar o padrão da incidência de MSC em cada avaliação no talhão. O envelope simulado para a função K de Ripley permitiu concluir que o padrão espacial é agregado observando-se uma escala razoável.

Agradecimentos

Este trabalho foi desenvolvido como parte das atividades do convênio firmado entre o Fundo de Defesa da Citricultura (FUNDECITRUS) e o Departamento de Estatística da Universidade Federal do Paraná e financiado pelo FUNDECITRUS.

Referências Bibliográficas

- [1] Alellyx. Alellyx identifica vetor da morte súbita dos citros. *Alellyx*, 2004. URL: http://www.gravena.com.br/dicas_alellyx_vetor_MSC.htm.
- [2] R. B. BASSANEZI, N. G. FERNANDES, and P. T. YAMMAMOTO. Morte súbita do citros. Technical report, Fundecitrus, 2003.
- [3] P. J. DIGGLE. *Statistical Analysis of Spatial Point Patterns*. Oxford University Press Inc, 2003.
- [4] L. V. MADDEN and G. HUGHES. Plant disease incidence: Distributions, heterogeneity, and temporal analysis. *Phytopathology*, 1995.
- [5] R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2004. ISBN 3-900051-00-3.