

<b>1 ESTUDO DA DISPERSÃO DO TRIPES DO PRATEAMENTO NA CULTURA DA CEBOLA</b>	<b>2</b>
Resumo . . . . .	2
Abstract . . . . .	2
1.1 Introdução . . . . .	3
1.2 Descrição dos dados . . . . .	4
1.3 <b>Métodos</b> . . . . .	7
1.4 <b>Detecção de padrão espacial através do teste de aleatorização de Mantel</b> . . . . .	7
1.4.1 <b>Resultados e discussão</b> . . . . .	13
1.5 Conclusões . . . . .	20
1.6 Referências . . . . .	21
<b>2 ANEXO A</b>	<b>23</b>

# 1 ESTUDO DA DISPERSÃO DO TRIPES DO PRATEAMENTO NA CULTURA DA CEBOLA

Denise N. Viola (LCE-ESALQ/USP - DES-UFBa), Clarice G. B. Demétrio (LCE-ESALQ/USP), Paulo J. Ribeiro Júnior (DES-UFPR), Bryan F. Manly (Western EcoSystems Technology), Odair A. Fernandes (FCAV-UNESP)

## Resumo

A cebola é uma das hortaliças mais cultivadas e consumidas no Brasil e sua importância social se deve à grande demanda por mão-de-obra. Uma das principais pragas que afeta essa cultura é o tripses do prateamento (*Thrips tabaci*) e sua distribuição espacial, embora importante, não tem sido considerada nas recomendações de manejo da cultura, planejamento de experimentos ou estudos amostrais. Para caracterizar o padrão espacial da dispersão do tripses do prateamento da cebola foi feito um levantamento anotando-se o número de insetos por fase de desenvolvimento em folhas de plantas de cebola, em diferentes datas e pontos amostrais dentro de quatro propriedades com fazendas vizinhas apresentando diferentes níveis de infestação e métodos de plantio. O teste de aleatorização de Mantel foi utilizado para testar a presença de padrão espacial, que quando detectado foi descrito por um modelo de Poisson misto espacial com componente aleatório geostatístico. Tal modelo possibilitou a caracterização do padrão espacial bem como a obtenção de mapas de predição dos níveis de susceptibilidade à infestação na área.

Palavras-chave: Tripes do prateamento, cebola, testes de aleatorização, geostatística, distribuição de Poisson, ...

## Abstract

The onion is one of the cultivated hortaliças more and consumed in Brazil and its social importance if it must to the great demand for man power. One of the main plagues that affect this culture is tripses of the prateamento (*Thrips tabaci*) and its space distribution, even so important, has not been considered in the recommendations of handling of the culture, planning of experiments or studies you show. To characterize the space standard of the dispersion of tripses of the prateamento of the onion the number of insects for phase of development in leves of onion plants was made a survey writing down itself, in diferentes dates and points you show inside of four properties with neighboring farms presenting different levels of infestation and methods of plantation. The test of randomization of Mantel was used to test the presence of space standard, that when detected he was described for a model of space mixing Poisson with geostatístico random component. Such model made possible the caracterização of the space standard as well as the attainment of maps of prediction of the levels of susceptibilidade to the infestation in the area.

Key-words: AAAAAAAA

## 1.1 Introdução

A cebola é uma das hortaliças mais cultivadas e consumidas no Brasil e sua importância social se deve à grande demanda por mão-de-obra. Estima-se que 70% da produção seja proveniente da agricultura familiar, pois é típica de pequenas e médias propriedades. É uma planta anual para a produção de bulbos e bianual para a produção de sementes. Seu plantio pode ser feito por semeadura direta, por bulbinhos ou por mudas cultivadas em canteiros e transplantadas para o campo.

Uma das principais pragas da cebola é o trips do prateamento (*Thrips tabaci*) que em altos níveis de infestação prejudica a colheita (WORKMAN; MARTIN, 2002) que pode ser reduzida em até 80% nas épocas quentes e secas (COSTA; MEDEIROS, 1950 e SATO, 1989).

Esse inseto é encontrado na base das folhas e se alimenta da seiva e dos grãos de clorofila provocando manchas cinzentas que depois se tornam prateadas como resultado da destruição dos tecidos externos das folhas. Ataques severos da parte aérea da planta causam perda na produção dos bulbos, que têm tamanho diminuído e pior qualidade, o que prejudica a comercialização e dificulta sua exportação (COSTA; MEDEIROS, 1950). Quando o ataque é muito intenso, as folhas ficam amareladas, secas e com a extremidade retorcida causando a murcha e a morte da planta (SATO, 1989) e facilitando a entrada de água até o bulbo, que apodrece. O inseto é considerado ainda vetor de agente fitopatogênico por ter a capacidade de transmitir virose para a planta.

O desenvolvimento desse inseto ocorre em quatro fases: ovo, ninfa, pupa e adulto, sendo que a ninfa e o adulto são os que de fato causam prejuízos à produção, pois a fase de pupa desenvolve-se no solo. A ninfa apresenta baixa mobilidade, enquanto o adulto, embora alado, tem movimentação restrita. Tipicamente o ciclo total de vida varia de 14 a 30 dias passando para 10 a 11 dias quando a temperatura é superior a 30°C.

A distribuição espacial de tripses em campos comerciais é importante, por exemplo, para aplicação racional de inseticidas. Entretanto, não tem sido considerada nas recomendações de manejo da cultura, planejamento de experimentos e estudos amostrais, sendo escassa literatura a esse respeito. Acredita-se que o principal fator de dispersão do tripses, e potencialmente determinante na ocorrência de padrões espaciais é o vento, pois as ninfas não são aladas e os adultos, embora alados, possuem asas rudimentares.

Numa caracterização inicial, o padrão espacial de dispersão de insetos pode ser classificado como aleatório, agregado ou regular. O aleatório ocorre quando, em condições naturais, há oportunidades iguais de infestação para todas as plantas, enquanto que o padrão agregado está associado à pequena mobilidade do inseto. O padrão regular raramente ocorre de forma natural, mas pode ser induzido pelo plantio alternado de plantas resistentes e susceptíveis. Para estudar se o padrão de dispersão de leucemia infantil era aleatório, Mantel (1967) propôs um teste de aleatorização, baseado nas matrizes de distância entre as observações. Entretanto, não foram encontradas aplicações no estudo de dispersão de insetos e, em particular, do tripses do prateamento.

Por outro lado, é muito comum, no estudo da dispersão de insetos, encontrar o uso de índices baseados na relação entre variância e média, tais como o índice de David & More, a lei de poder de Taylor, os índices de agregação de Lloyd e Iwao, dentre outros (RUIZ et al., 2003). Entretanto, esses índices ignoram a localização espacial das amostras, tendo capacidade limitada de descrição de padrões espaciais, além de dependerem fortemente do tamanho das unidades amostrais.

Métodos geoestatísticos (ISAAKS E SRISVARTAVA; 1989; GOOVAERTS; 1997) têm sido usados para descrever os padrões espaciais de insetos como por exemplo em Greco, Vieira e Lourenco, 2006. Tais métodos foram originalmente desenvolvidos para variáveis aleatórias con-

tínuas com diversas implementações computacionais disponíveis. A variável contagem de insetos é discreta e tipicamente distribuída em aglomerados e com excesso de valores iguais a zero. Logo, os dados de contagens podem não ter a estrutura de covariância assumida pelos métodos tradicionais da análise geoestatística que tem uma estrutura de covariância espacial estacionária na área (RUIZ, 2002). Portanto, é recomendável adotar modelos que incorporem explicitamente o mecanismo gerador dos dados, tais como, neste caso, modelos com distribuição de Poisson para dados de contagem combinados com estruturas que descrevam o padrão espacial. Modelos desse tipo propostos na literatura estatística (e.g. DIGGLE, TAWN E MOYEED, 1998) ainda encontram poucas aplicações na análise de dados.

Neste artigo, faz-se um estudo da distribuição espacial do tripses do prateamento da cebola a partir de dados provenientes de um levantamento realizado em quatro propriedades, considerando-se diferentes níveis de infestação e métodos de plantio nas fazendas vizinhas. O teste de aleatorização de Mantel (MANLY, 2006) foi utilizado para testar a presença de padrão espacial, que quando detectado foi descrito por um modelo de Poisson misto espacial com componente aleatório geoestatístico. Tal modelo possibilitou a caracterização do padrão espacial bem como a obtenção de mapas de predição dos níveis de susceptibilidade à infestação na área.

O restante deste artigo está organizado como se segue. A Seção ?? descreve os dados a serem utilizados, enquanto que a Seção 3 traz uma revisão do teste de aleatorização de Mantel para a detecção do padrão espacial e a Seção 4 apresenta um modelo de Poisson misto espacial com componente aleatório geoestatístico. Resultados obtidos são apresentados e discutidos na Seção 5. Finalmente, são feitas algumas considerações finais na Seção 1.5.

## 1.2 Descrição dos dados

A motivação para este trabalho surgiu a partir de um conjunto de dados provenientes de um estudo por amostragem do tripses do prateamento em cultura de cebola, em quatro propriedades, localizadas no Município de São José do Rio Pardo, interior do Estado de São Paulo, no período de junho a setembro de 1996. As quatro propriedades escolhidas usaram o híbrido Granex 33 de cebola e o método de plantio por mudas. As áreas experimentais foram escolhidas com vizinhos que adotavam diferentes tipos de plantio e tinham diferentes níveis de infestação. Teve como objetivo estudar a distribuição espacial e temporal do tripses.

Na Tabela 1.1, podem ser vistos os detalhes referentes a tipo de plantio vizinho e data das coletas e número de amostras por coletas nas diferentes propriedades. A Fazenda São Paulo está localizada em um ponto alto da região e o plantio de cebola mais próximo se encontrava a uma distância de pelo menos um quilômetro. Nas áreas vizinhas à Estância Bela Vista já havia culturas atacadas pela praga do tripses do prateamento.

A unidade amostral foi um círculo de 1m de raio com uma estaca no centro sendo sorteada uma planta em cada data de amostragem. A Figura 1.1 mostra a posição das estacas nas quatro propriedades estudadas, em geral, com um espaçamento de referência nas linhas e colunas de 10 em 10 metros, sendo que na Fazenda São Paulo há variações. As variáveis medidas foram a localização das estacas no eixo das coordenadas, o número de ninfas, o número de insetos adultos e o número de folhas por planta. A listagem completa do número de insetos (ninfas + adultos) por folha, encontra-se no Apêndice A.

Tabela 1.1 – Detalhes sobre a origem dos dados, tipos de plantios vizinhos, datas de coletas e números de amostras

Propriedade	Plantios vizinhos	Datas das Coletas	Nº de Amostras
Fazenda São Paulo	isolada de outros plantios	10/7, 24/7, 31/7, 7/8, 14/8, 21/8, 28/8, 04/9	100, 100, 100, 98
Estância Bela Vista	bulbinhos	11/7, 1/8, 8/8, 14/8, 9/9	100, 100, 84
Sítio Rosário	mudas	21/6, 29/6, 7/7, 14/7, 21/7, 28/7, 4/8, 11/8, 18/8, 25/8, 3/9	50, 50, 48, 50, 50, 50, 50, 50, 50, 50, 50
Sítio Novo II	mudas	4/6, 19/6, 27/6, 28/6, 4/7, 11/7, 24/7, 31/7, 7/8	100, 100, 100, 100, 100, 100, 100, 100

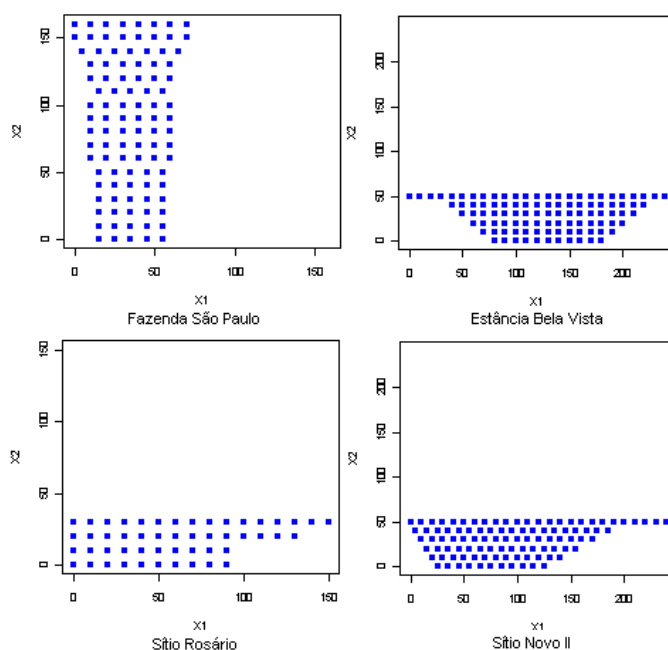


Figura 1.1 – Localização das estacas centrais dos círculos de 1m de raio, nas quatro propriedades.

Pode-se ver pela Tabela 1.1 que as coletas dos insetos foram feitas em números e épocas distintas nas quatro propriedades e que as variáveis respostas são discretas (contagem de ninfas e adultos). Existiram problemas na coleta dos dados, tais como: em alguns casos, as contagens de insetos foram registradas apenas como múltiplos de 5 ou 10 e as contagens acima de 100 foram registrados iguais a 100 e, também, há dados faltantes.

A Figura 1.2 mostra o gráfico *box-plot* para o número médio de insetos por folha, nas quatro propriedades, para as diversas coletas, enquanto que estatísticas descritivas estão nas Tabelas 2 a 2 do Anexo A. Vê-se, em todos os casos, que há uma variabilidade grande e aparentemente não existe um padrão ao longo do tempo, com exceção da Fazenda São Paulo, em que aparentemente o número

médio de insetos e a variabilidade aumentam com o tempo, já nos outros casos, a média aumenta e depois diminui. Nota-se ainda que as observações estão mais dispersas acima da mediana em todos os casos, mostrando assimetria positiva e apresentando valores discrepantes na maioria deles.

Na Fazenda São Paulo, observa-se que o dia 21/07 apresentou média menor de insetos por folha e também menor variância, tendo um inseto por folha como valor máximo, enquanto que o dia 04/09 apresentou uma maior média de insetos por folha e também a maior variabilidade. Verifica-se ainda, que essa propriedade apresenta menor proporção de plantas infestadas, variando de 35% a 89%, com exceção do dia 04/09 em que todas as plantas estavam infestadas.

Para a Estância Bela Vista, observa-se que as menores médias de insetos por folha ocorreram nos dias 11/07 e 14/08, e essa propriedade foi a que mostrou maior proporção de plantas infestadas, variando de 89,00% a 100,00%, sendo que em três datas de coleta, todas as plantas estavam infestadas.

Observa-se também que o Sítio Rosário foi o que apresentou menor variabilidade, apresentando também a maior média de insetos por folha nos dias 11 e 18 de agosto, vale ressaltar também que nessa propriedade foram analisados apenas 50 pontos.

No Sítio Novo II, observa-se a menor variabilidade nas médias e também, as menores médias para o número de insetos por folha.

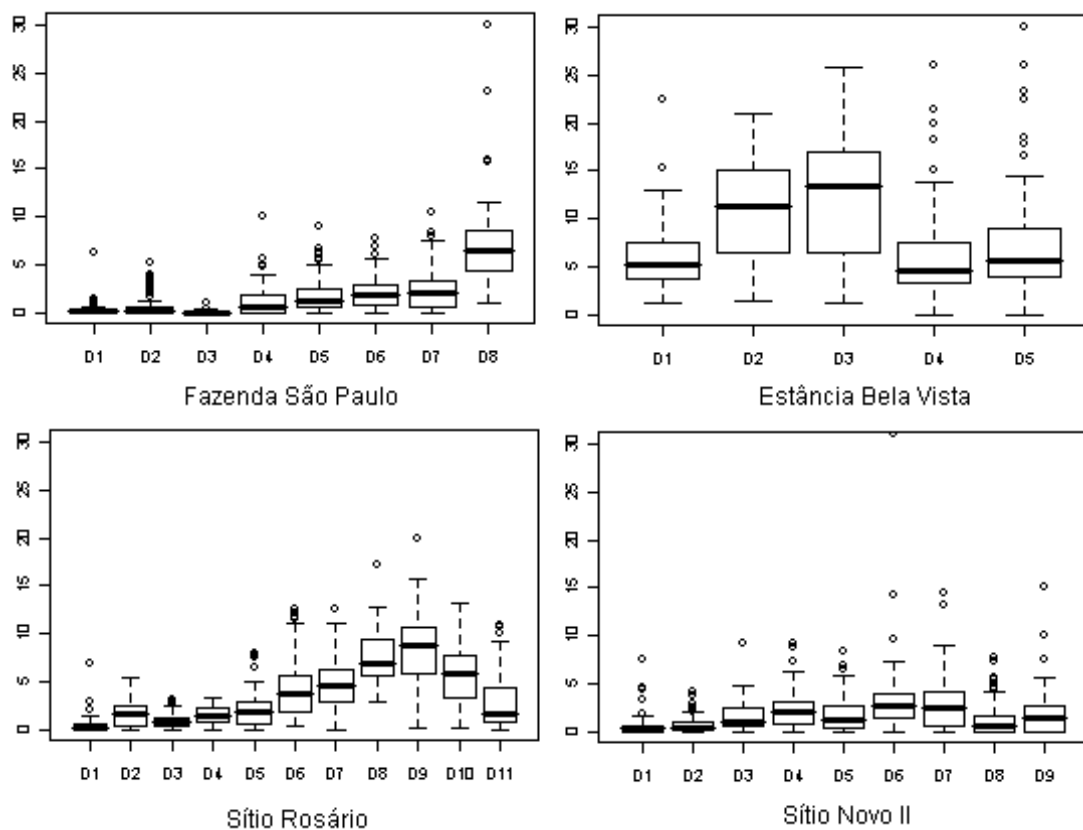


Figura 1.2 – Box-Plot do número médio de insetos por folha.

### 1.3 Métodos

#### 1.4 Detecção de padrão espacial através do teste de aleatorização de Mantel

A hipótese de não existência de padrão espacial na dispersão dos insetos pode ser considerada como hipótese de aleatoriedade, pois pode sugerir a existência de um padrão nos dados que é afetado quando se aleatorizam as ordens dos dados observados (MANLY, 2006).

Os testes de aleatorização têm como base o fato que se a hipótese de nulidade é verdadeira, todas as possíveis ordens para os dados têm a mesma chance de ocorrer. Assim, para um conjunto de observações calcula-se o valor  $e_o$  de uma estatística  $E$  e a seguir, faz-se um número grande (em geral, 10.000) de aleatorizações (reordenações aleatórias) dos dados, obtendo-se valores  $e_a$  que irão gerar uma distribuição para  $E$ . A interpretação usada é a mesma dos testes clássicos, isto é, considera-se  $p$  como a proporção dos valores  $e_a$  que são maiores do que ou iguais a  $e_o$  e, por exemplo, se  $p < 0,05$ , conclui-se que existe uma forte evidência de que a hipótese de nulidade não seja verdadeira ao nível de 5% de probabilidade (MANLY, 2006).

Como vantagens dos testes de aleatorização, em relação aos testes padrões, é que têm a mesma validade para amostras não aleatórias, são relativamente fáceis de serem calculados, são baseados em estatísticas não padronizadas e não necessitam de informações prévias a respeito da população da qual a amostra foi retirada. Na realidade, as amostragens nem sempre são aleatórias, consistindo simplesmente de itens que precisam ser analisados (MANLY, 2006). Entretanto, os testes de aleatorização são mais facilmente justificados quando as amostras analisadas são aleatórias ou o próprio delineamento experimental exige um teste de aleatorização.

A principal desvantagem em se usarem os testes de aleatorização é que à primeira vista, parecem ser pouco rigorosos, pois nem sempre é possível generalizar as conclusões obtidas, sendo específicos para cada conjunto de dados. O que um teste de aleatorização diz é que um certo padrão nos dados pode, ou não, ter aparecido por acaso (MANLY, 2006).

No contexto de dados espaciais, deseja-se testar as hipóteses

$H_0$ : Padrão espacial aleatório *versus*  $H_a$ : Padrão espacial não aleatório

Um teste para essas hipóteses foi proposto por Mantel (1967) e uma forma de se verificar se o padrão aleatório explica a dispersão dos insetos é baseada no estudo de dispersão entre pontos. Assim, considerando-se a comparação de duas variáveis, duas matrizes  $n \times n$ , simétricas, de distâncias podem ser obtidas, isto é,

$$A = \begin{pmatrix} 0 & a_{21} & \dots & a_{n1} \\ a_{21} & 0 & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & 0 \end{pmatrix} \quad e \quad B = \begin{pmatrix} 0 & b_{21} & \dots & b_{n1} \\ b_{21} & 0 & \dots & b_{n2} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & 0 \end{pmatrix}$$

sendo que,  $A$  é a matriz das distâncias Euclidianas entre os pontos  $(x_{1i}, x_{2i})$  e  $(x_{1j}, x_{2j})$ , isto é, com elementos  $a_{ij} = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2}$  e  $B$  é a matriz com elementos  $b_{ij} = \sqrt{(z_i - z_j)^2}$ ,

em que  $Z$  é a variável número médio de insetos por folha. A estatística teste é dada pelo coeficiente de correlação de Pearson entre os elementos correspondentes de  $A$  e  $B$ , isto é,

$$r = \frac{m \sum_{i < j} a_{ij} b_{ij} - \sum_{i < j} a_{ij} \sum_{i < j} b_{ij}}{\sqrt{[m \sum_{i < j} a_{ij}^2 - (\sum_{i < j} a_{ij})^2][m \sum_{i < j} b_{ij}^2 - (\sum_{i < j} b_{ij})^2]}}, \quad (1.1)$$

que é calculada para os valores observados ( $r_o$ ). A seguir permutam-se linhas e colunas de uma das matrizes um número suficientemente grande e obtêm-se os valores  $r_{ak}$ , para  $k = 1, 2, \dots, 10.000$  e  $p$ . A proporção  $p$  de vezes que  $r_a \geq r_o$ , é computada e comparada com um valor pré-fixado  $\alpha$  (por exemplo, 0,05). Rejeita-se  $H_o$  se  $p < \alpha$  (MANLY, 2006).

Como as matrizes  $A$  e  $B$  são simétricas, então a correlação entre todos os elementos fora da diagonal principal é a mesma que a correlação de  $m = n(n - 1)/2$  elementos na parte triangular superior da matriz. É importante ressaltar que o único termo de (1.1) que é alterado pela mudança da ordem dos elementos em uma das matrizes  $A$  ou  $B$  é a soma de produtos  $Z = \sum a_{ij} b_{ij}$ .

O teste de aleatorização para padrões espaciais foi calculado considerando-se as observações em cada data de coleta.

Outras métricas utilizadas para o cálculo das distâncias podem ser Euclidiana com dados padronizados, Euclidiana quadrática, Euclidiana quadrática com dados padronizados, distância proporcional e diferença amostral para procedimentos de permutação de respostas múltiplas. Esse teste, entretanto tem como suposição forte que a correlação é linear e testes alternativos de aleatorização propostos podem ser citados tais como de Besag-Diggle, de Edgington, Monte Carlo, dentre outros (MANLY, 2006). Uma outra proposta é dada por Ribeiro Jr. (200?) que usaram os modelos generalizados aditivos.

Quando o teste de Mantel rejeita a hipótese de nulidade, pode haver interesse em se detectar qual o tipo de associação existente entre as variáveis e isso pode ser mostrado pelo gráfico de  $b_{ij}$  versus  $a_{ij}$ . Um dos possíveis modelos de associação é a regressão linear simples, em que os elementos da matriz  $A$  entram como variável explanatória e os elementos da matriz  $B$  como variável resposta, isto é,

$$b_{ij} = \beta_0 + \beta_1 a_{ij} + \epsilon_{ij}$$

em que  $\beta_0$  e  $\beta_1$  são parâmetros a serem estimados e  $\epsilon_{ij}$  é o erro associado à resposta. Entretanto, formas mais complexas de dependência espacial podem ocorrer.

A variável resposta utilizada foi a distância Euclidiana do número médio de insetos por folha e a covariável foi a distância Euclidiana dos pontos amostrados.

De forma semelhante, foi usado o teste de aleatorização para detecção de padrões temporais, considerando-se as observações em cada ponto amostrado para cada data estudada. Para a realização do teste de Mantel foram consideradas as distâncias simples e as distâncias recíprocas (inverso da distância) dos tempos amostrados. Para a regressão linear, a variável resposta utilizada foi a distância Euclidiana do número de insetos por folha e a covariável foi a distância Euclidiana dos tempos amostrados. É importante enfatizar que o uso do teste de Mantel para a dependência temporal levanta problemas de como combinar a informação de um número grande de plantas.



### Modelagem do padrão espacial

Foram feitos gráficos de pontos com a localização das plantas *versus* o número médio de insetos por folha, para uma idéia visual da existência de padrão espacial.

Existem diversos métodos para estudar a distribuição espacial de objetos ou indivíduos, sendo comum o uso das técnicas da Geoestatística, que podem estimar, a partir de amostras, seu grau de associação (correlação), usando medidas de distância entre eles e da direção deles. A dependência espacial (autocorrelação espacial) mostra que os pontos amostrados mais próximos são mais parecidos do que pontos mais distantes (MONTAGNA, 2001). O termo geoestatística é utilizado para identificar uma parte da estatística espacial, sendo utilizado quando existe uma variação contínua das observações no espaço (DIGGLE; RIBEIRO JR; CHRISTENSEN, 2003).

O formato básico de dados geoestatísticos univariados é dado por  $(x_i, y_i), i = 1, 2, \dots, n$ , em que  $x_i = (x_{1i}, x_{2i})$  identifica a localização espacial (espaço bidimensional) e  $y_i$  é a medida escalar na posição  $x_i$  da  $i$ -ésima observação. Tem-se ainda que  $y$  pode estar localizado a princípio em qualquer lugar da região estudada (DIGGLE; RIBEIRO JR., 2006).

O modelo geoestatístico é representado por:  $[S, Y] = [S][Y|S]$ , em que:  $Y(x) : x \in A$  é o processo de medida;  $y_i$  são os dados observados;  $S = \{S(x) : x \in R^2\}$  é um processo gaussiano com média  $\mu$  e variância  $\sigma^2$ ,  $S(x)$  é chamado de sinal e seus valores não são diretamente observados,  $x$  é a posição da observação no espaço;  $A$  é um subconjunto fixo de pontos no  $R^2$  (DIGGLE; MOYEED; TAWN, 1998; DIGGLE; RIBEIRO JR., 2006). Tem-se ainda que  $Y_i$  é uma versão de “ruído” de  $S(x_i)$  para um conjunto de localizações  $x_i$  (DIGGLE; MOYEED; TAWN, 1998). Geralmente, assume-se que o delineamento de amostragem para a posição  $x_i$  é estocasticamente independente do processo que gera as medidas de  $y_i$ . O modelo, pode então ser escrito como:  $Y_i = S(x_i) + Z_i$ , em que  $Z_i$  são mutuamente independentes e seguem distribuição normal, com média 0 e variância  $\tau^2$  (DIGGLE; RIBEIRO JR., 2006).

Logo, pode-se dizer que

$$\mathbf{Y} \sim N(\mu \mathbf{1}, \sigma^2 R + \tau^2 I)$$

em que  $\mu$  é a média geral, mas no caso de existirem covariáveis ou tendência,  $\mu$  pode ser uma função.

Os métodos de análise geoestatística tradicionais são baseados em modelos Gaussianos, isto é, em situações em que as variáveis respostas são contínuas, como por exemplo: (i) no estudo da concentração de ouro, Krige (1951), verificou que apenas a informação da variância seria insuficiente para explicar o fenômeno estudado e que seria necessário levar em consideração a localização geográfica e a dependência espacial das observações; (ii) em um estudo da distribuição da precipitação pluviométrica medida em 143 postos meteorológicos do Estado do Paraná no período de seca, maio/junho (RIBEIRO JR, 2004, ZAMBOTI, 2001); (iii) no estudo da densidade e da umidade do solo, medidas em 250 pontos distribuídos em uma malha regular, utilizando sonda de neutrons em seis diferentes profundidades (RIBEIRO JR, 1995); (iv) no estudo do teor de argila de 85 amostras de solos da Fazenda Canchim, localizada em São Carlos, SP para fins de mapeamento (CAMARGO), (v) no estudo para obter os métodos de estimação dos parâmetros em modelos geoestatísticos, considerando diferentes estruturas de covariância para diferentes métodos de estimação, avaliando características morfológicas, físicas e químicas dos solos em diferentes

camadas (OLIVEIRA, 2003); (vi) no estudo para modelar a profundidade *secchi* e a concentração de clorofila A em nove pontos de coleta no Rio Anil, localizado em São Luís - MA, que pode estar associada a covariáveis, tais como salinidade, pH e temperatura (ALCÂNTARA et al., 2004) .

Como uma análise exploratória, um gráfico muito utilizado na análise geoestatística, quando a variável resposta é contínua, é o semi-variograma, isto é, o gráfico dos valores da semivariância,  $\gamma(h)$ , de pares de observações separadas pela mesma distância  $h$  versus a distância  $h = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2}$  entre os pontos amostrados. A semivariância é dada por:

$$\gamma^*(h) = \frac{1}{2N(h)} \sum_{Y_i+h, Y_i \in N(h)} \{Y(x+h) - Y(x)\}^2$$

em que  $Y(x)$  e  $Y(x+h)$  são as medidas nas posições  $x$  e  $x+h$  e  $N(h)$  é o número de pares de pontos separados pela distância  $h$ . Se não houver tendência, o variograma é feito com a variável resposta e no caso de existir tendência, retira-se a tendência e o variograma então é feito com os resíduos.

Esse gráfico pode mostrar o grau de similaridade entre pontos mais próximos, tendo como vantagens em relação à covariância, a não necessidade de conhecer a  $E[Y(x)]$  e poder incluir funções aleatórias com variância infinita. Sua forma é crescente, pois a variabilidade aumenta à medida que cresce a distância entre os dados. Uma forma alternativa, é o variograma que é obtido multiplicando o semi-variograma por 2. (BRAGA, 1990, MONTAGNA, 2001, ORTIZ, 2002; PILON, 2004, SCHTZKO; KNUDSEN, 1992).

Existem três tipos de variogramas, o experimental, o verdadeiro e o teórico. O variograma experimental é obtido a partir de observações feitas por um processo de amostragem sobre as coordenadas geográficas. O verdadeiro é desconhecido e representa a situação real. A análise estrutural é utilizada para verificar qual o variograma teórico que melhor se ajusta ao variograma experimental e, então, fazer inferências sobre o verdadeiro variograma (ORTIZ, 2002, OLIVEIRA, 2003, PILON, 2004).

Como complemento, ao variograma, o gráfico do envelope simulado também é muito utilizado para verificar a existência de padrões espaciais. Os limites do envelope são obtidos considerando os valores máximos e mínimos dos variogramas ajustados, aleatorizando-se as ordens das posições das observações e considerando-se a variável resposta fixa, (GeoR <http://www.est.ufpr.br/geoR>). Recomendam-se 99 aleatorizações para uma confiança razoável nos envelopes (ORTIZ, 2003). Se existirem pontos fora desses limites, é uma evidência da existência de algum padrão espacial.

Entretanto, em alguns casos, as variáveis respostas são contagens, proporções, ou então, assumem apenas valores positivos ou muitos iguais a zeros, e a distribuição condicionada às variáveis de localização espacial, não é normal, isto é,  $Y|S$  não é um processo gaussiano.

Como exemplo de variáveis respostas de contagem podem ser citados (i) um estudo realizado na Ilha de Rongelap, que faz parte da República Marshal e está localizada no Oceano Pacífico, sobre a localização de níveis de alta concentração de césio radioativo, um resíduo de contaminação decorrente de testes de armas nucleares na década de 50, em que a variável resposta é o número de *photon emissions* na escala de energia discreta (DIGGLE; MOYEED; TAWN, 1998, DIGGLE; RIBEIRO JR; CHRISTENSEN, 2003); (ii) um estudo da distribuição espacial das espécies de cigarrinhas vedoras da *Xylella fastidiosa*, agente que causa a *Clorose Variiegada* dos Citros em um

pomar de laranjas “Pera”, com o objetivo de estabelecer meios para controlar melhor a doença, em que a variável resposta é o número dos indivíduos coletado no período de dezembro 1996 a dezembro 1999, agrupados por estações (FARIAS et al., 2004).

Como exemplo de variáveis respostas binárias podem ser citados (i) um estudo da distribuição espacial de espécies de líquens em que a variável resposta é presença, ou não, de líquens em troncos de árvores, que pode estar associada a covariáveis tais como diâmetro, sombreamento, cobertura do tronco, umidade etc; (ii) um estudo da existência de padrão espacial da incidência de malária em crianças de Gambia, África, em que a variável resposta presença, ou não, do agente causador, pode estar associada a covariáveis tais como localização das vilas, presença de centro de saúde, índice de vegetação, idade da criança, uso e tratamento de mosquiteiros, etc (DIGGLE; RIBEIRO JR., 2006); (iii) um estudo do levantamento da infestação da broca do café na Colômbia em um período de quatro anos em que a variável resposta é presença, ou não, da broca e, no caso de encontrar pelo menos um fruto com broca, foi calculada a proporção de frutos infestados na planta, que pode estar associada a covariáveis, tais como temperatura (máxima, mínima e média), insolação, umidade relativa e precipitação pluviométrica (RUIZ, 2002); (iv) um estudo realizado em dezenove setores censitários na região do município de Duque de Caxias, RJ para modelar a distribuição espacial da epidemiologia de *ascariase* em que a variável resposta é presença, ou não, da doença (FORTES et al., 2004).

Nesses casos, uma solução para a análise dos dados seria usar uma transformação do tipo permitido pela família de Box-Cox e depois o modelo Gaussiano, como proposto por Christensen; Diggle; Ribeiro Jr., (2000). Entretanto, há limitações e os autores sugerem que isso seja usado apenas para uma análise inicial dos dados. A melhor alternativa, é usar os modelos espaciais lineares generalizados.

No modelo linear generalizado padrão, tem-se que os  $Y_i, i = 1, 2, \dots, n$  são mutuamente independentes com distribuição pertencente à família exponencial na forma canônica e  $\mu_i = E[Y_i]$  especificado pelo preditor linear  $h(\mu_i) = \eta_i$  para alguma função de ligação  $h(\cdot)$  conhecida.

O modelo linear generalizado geoestatístico é um caso particular de modelo linear generalizado misto em que os  $Y_i, i = 1, 2, \dots, n$  condicionais a  $S(x)$  são mutuamente independentes com esperança  $E[Y_i|S(x)] = \mu_i$  e preditor linear  $h(\mu_i) = S(x_i), i = 1, 2, \dots, n$  para uma função de ligação  $h(\cdot)$  conhecida, sendo que o sinal do processo é  $\{h^{-1}(S(x)) : x \in A\}$  e  $S(x_i) = S(x) + d(x)^T \beta$ , em que  $\beta$  é o vetor dos parâmetros de regressão (DIGGLE; MOYEED; TAWN, 1998, DIGGLE; RIBEIRO JR.; CHRISTENSEN, 2003).

Supondo-se que o número total de insetos  $Y(x_i)|S(x_i)$  tem distribuição de Poisson com média  $t_i \exp(S(x_i)), i = 1, 2, \dots, n$ , em que  $t_i$  representa os números de folhas, tem-se que

$$P[Y(x)|S(x) = y(x_i)] = \frac{e^{-t_i e^{S(x_i)}} (t_i e^{S(x_i)})^{y(x_i)}}{y(x_i)!}.$$

No caso de modelos lineares generalizados espaciais, a função de verossimilhança, geralmente, não é expressa de forma fechada, sendo descrita pela integral

$$L = \int \prod_{i=1}^n \frac{e^{-t_i e^{S(x_i)}} (t_i e^{S(x_i)})^{y(x_i)}}{y(x_i)!} \cdot \frac{1}{\sqrt{2\pi} |\sigma^2 R|} e^{-\frac{1}{2}(S(x_i) - \mu_i)' \sigma^2 R (S(x_i) - \mu_i)} dS, \quad (1.2)$$

que tem dimensionalidade grande, não tendo solução explícita na maior parte dos casos. Métodos padrões para resolver a integral (1.2) são desconhecidos, porém, uma possível solução é utilizar Cadeias de Markov via Monte Carlo (MCMC), o que já está implementado no pacote GeoRGLM do R (<http://www.R-project.org>).

No caso de variáveis aleatórias discretas, o variograma é um resumo menos natural dos dados, mas pode ser útil como ferramenta para diagnóstico, após o ajuste do modelo linear generalizado misto (DIGGLE; RIBEIRO JR, 2006). Nesse caso, compara-se o variograma ajustado com o variograma experimental, obtido através dos resíduos. Os variogramas experimental e ajustado são dados, respectivamente por:

$$\gamma_Y(h) = \frac{1}{2} Var\{Y(x)\} + \frac{1}{2} Var\{Y(x+h)\} - Cov\{Y(x), Y(x+h)\}.$$

e

$$\gamma_Y(h) = \exp\left(\beta + \frac{\sigma^2}{2}\right) + \exp(2\beta + \sigma^2) [\exp(\sigma^2) - \exp\{\sigma^2 \rho(u)\}].$$

Para se fazer essa análise, as estimativas dos parâmetros do modelo ajustado são supostas verdadeiras. Simulam-se 99 amostras espaciais independentes com as mesmas localizações dos dados observados. O variograma experimental é calculado para cada uma das simulações e o envelope é construído através dos quantis dos variogramas para os dados simulados e, finalmente, verifica-se se o intervalo de confiança contém o variograma ajustado aos dados observados.

Porém, deve-se tomar cuidado pois o variograma é errático, uma vez que é gerado por dados assimétricos, sendo portanto, mais aconselhado para orientar as análises do que para orientar nas estimativas dos parâmetros.

Após a escolha de um determinado modelo, em geral, há interesse na construção de um mapa de predição. Assumindo, inicialmente, que se conhecem os parâmetros e que o interesse está na intensidade de insetos dada por  $\lambda(x_0) = \exp(\beta + S(x_0))$ , para o local  $x_0 = (x_{10}, x_{20})$ , com a distribuição marginal de  $S$  e a distribuição condicional de  $Y|S$  pode-se simular a distribuição condicional de  $[S|y]$ , através do método MCMC. A superfície predita é dada por (DIGGLE; MOYEED; TAWN, 1998):

$$\hat{\beta} + \hat{S}(x) + \frac{Var(x)}{2},$$

sendo que,  $\hat{\beta}$  é a média do processo, nesse caso, pois não existem covariáveis explanatórias nem tendência;  $\hat{S}(x)$  é o preditor linear de krigagem e  $Var(x)$  é a variância da predição.

#### 1.4.1 Resultados e discussão

### Detecção de padrão espacial através do teste de aleatorização de Mantel

#### Fazenda São Paulo

Os valores de  $p$  obtidos para o teste de Mantel feito para cada uma das coletas foram **0,0228**, **0,0022**, **0,0235**, 0,0588, 0,1005, 0,0749, 0,1297 e 0,5540. Pode-se observar, portanto, que apenas nos dias 10, 24 e 31 de julho existe evidência de rejeição da hipótese de padrão aleatório, mostrando uma relação entre a localização da planta e o número de insetos por folha. Isso é confirmado quando se observa o gráfico de pontos (Figura 1.3) em que os pontos mais próximos têm cores semelhantes. O gráfico de pontos é dividido por quartis, em que azul representa as cores das observações menores do que o primeiro quartil, verde são as observações entre o primeiro quartil e a mediana, amarelo as cores entre a mediana e o terceiro quartil e as maiores do que o terceiro quartil são vermelhas. Nesse caso, os gráficos são representados por menos cores pois existem pontos com os mesmos valores pertencentes a dois intervalos diferentes. Observa-se também que a distribuição de número médio de insetos por folha é assimétrica e aparentemente não apresenta tendência.

Fazendo-se a regressão linear simples das distâncias do número de insetos por folhas em relação às distâncias da localização das estacas, vê-se que existe evidência de que à medida que aumenta a distância entre pontos, aumenta a distância entre o número de insetos por folha (ver Tabela 1.2).

Tabela 1.2 – Modelos de regressão - Testes de Aleatorização

Propriedade	Data	Modelo	p
	10/07	Insetos/folha = 0,2102 + 0,002325loc	0,0205
Fazenda São Paulo	24/07	Insetos/folha = 0,6024 + 0,004216loc	0,0022
	31/07	Insetos/folha = 0,09317 + 0,0004174	0,0264
Estância Bela Vista	08/08	Insetos/folha = 6,218 + 0,0090371loc	0,0334
	04/06	Insetos/folha = 0,3035 + 0,007206loc	0,0012
Sítio Novo II	27/06	Insetos/folha = 1,181 + 0,004034loc	0,0258
	04/07	Insetos/folha = 1,524 + 0,003371loc	0,0455

#### Estância Bela Vista

Os valores de  $p$  obtidos para o teste de Mantel feito para cada uma das coletas foram 0,8986, 0,0902, **0,0338**, 0,7880 e 0,6224. Nota-se, portanto, que apenas no dia 08 de agosto existe evidência de rejeição da hipótese de padrão aleatório, mostrando uma relação entre a localização da planta

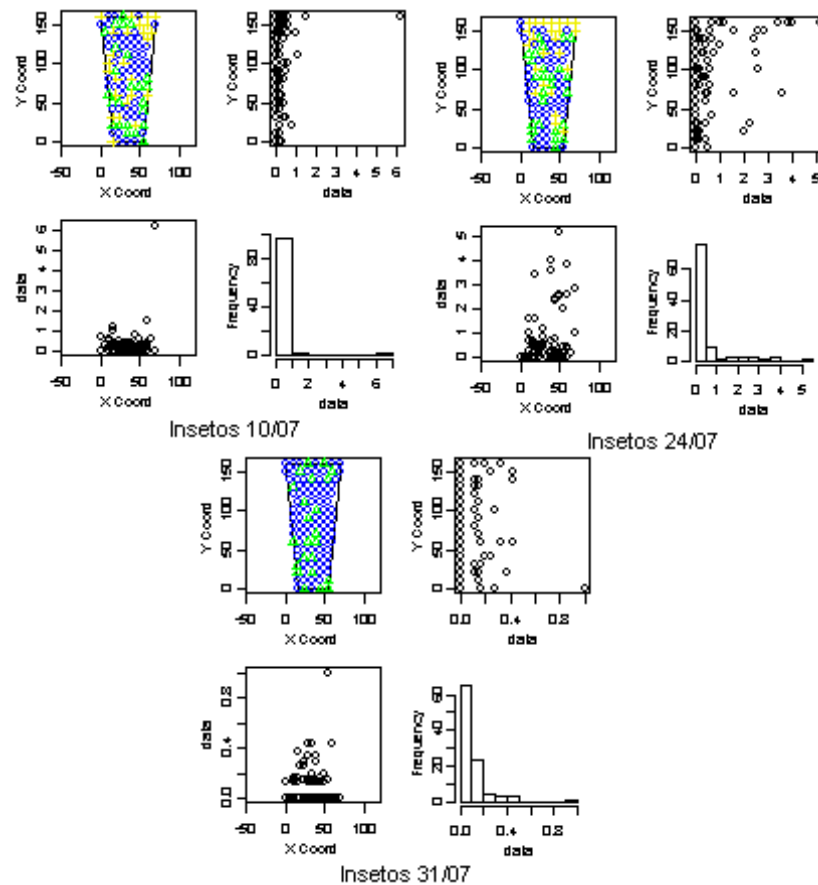


Figura 1.3 – Gráficos de dispersão e histogramas para o número médio de insetos - Fazenda São Paulo.

e o número de insetos por folha. A Figura 1.4 confirma isso, pois observa-se um padrão nas cores, sendo que o vermelho mostra uma maior concentração de insetos enquanto verde e azul indicam baixa infestação. Nota-se também que a distribuição de insetos por folha é assimétrica e aparentemente não apresenta tendência.

Fazendo-se a regressão linear simples das distâncias do número de insetos por folhas em relação às distâncias da localização das estacas, observa-se que existe evidência de que à medida que aumenta a distância entre pontos, aumenta a distância entre o número de insetos por folha (ver Tabela 1.2).

### Sítio Rosário

Os valores de  $p$  obtidos para o teste de Mantel foram 0,5309, 0,5961, 0,9236, 0,9512, 0,9412, 0,7297, 0,3223, 0,1897, 0,2771, 0,1771 e 0,7020. Nota-se que não existe evidência de rejeição da hipótese de padrão aleatório, portanto, o comportamento dos insetos é aleatório quando relacionado com a posição da planta.

### Sítio Novo II

Os valores de  $p$  obtidos para o teste de Mantel feito para cada uma das coletas foram **0,0006**,

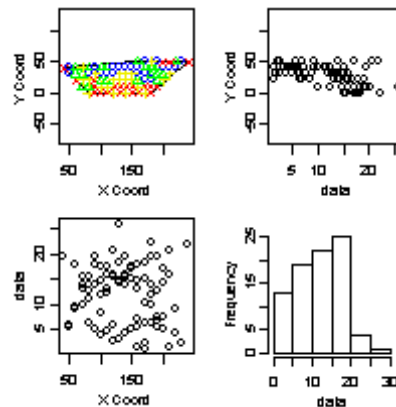


Figura 1.4 – Gráficos de dispersão e histogramas para o número médio de insetos - Estância Bela Vista.

0,6127, **0,0251**, 0,7326, **0,0478**, 0,4781, 0,0651, 0,6084 e 0,4264. Pode-se observar, portanto, que apenas nos dias 04 e 27 de junho e 04 de julho existe evidência de rejeição da hipótese de padrão aleatório, mostrando relação entre a localização da planta e o número de insetos por folha. Isso é confirmado quando se observa o gráfico de pontos (Figura 1.5) em que há grupos de pontos cores semelhantes.

Fazendo-se a regressão linear simples das distâncias do número de insetos por folhas em relação às distâncias da localização das estacas, nota-se que existe evidência de que à medida que aumenta a distância entre pontos, aumenta a distância entre o número de insetos por folha (ver Tabela 1.2).

### Modelagem do padrão espacial

Após verificar-se a existência de evidência de padrão espacial nos dados foi feita a modelagem espacial, considerando-se técnicas de modelos lineares generalizados para a distribuição de Poisson, com função de ligação logarítmica. Foram feitos também intervalos de confiança para os parâmetros estimados através de MCMC, considerando-se uma cadeia de 50.000 iterações, com ciclo de pré convergência (*burn in*) 10.000. Foram guardadas 2.500 (uma amostra a cada vinte amostras geradas) para garantir a independência destas. Esse processo foi repetido 1.000 vezes e, através das 1.000 estimativas obtidas, foram calculados os intervalos de confiança para os quantis 2,5% e 97,5%. Os resultados obtidos estão na Tabela 1.3.

As estimativas obtidas na Tabela 1.3 estão dentro dos intervalos de confiança, porém, nota-se que  $\phi$  apresenta grande amplitude no intervalo, o que já era esperado, uma vez que esse parâmetro fornece informações do alcance do modelo, ou seja, a distância a partir da qual as observações não apresentam mais dependência espacial. Para analisar esse parâmetro, deve-se obter informações da distância em cada propriedade. No caso da Fazenda São Paulo, essa distância foi de 10 a 170 metros, na Estância Bela Vista, essa distância foi de 10 a 200 metros e no Sítio Novo II, essa distância foi de 10 a 204 metros. Logo, nota-se que esses intervalos estão dentro do esperado. Observa-se ainda que existem três casos em que a estimativa é menor do que a distância mínima.

Observa-se também que a amplitude dos intervalos para  $\beta$  e  $\sigma^2$  são pequenas. Tem-se ainda que  $\beta$  é o parâmetro associado à função de ligação e  $\sigma^2$ ,  $\phi$  e  $\tau^2$  são parâmetros associados à superfície  $S(x)$ .

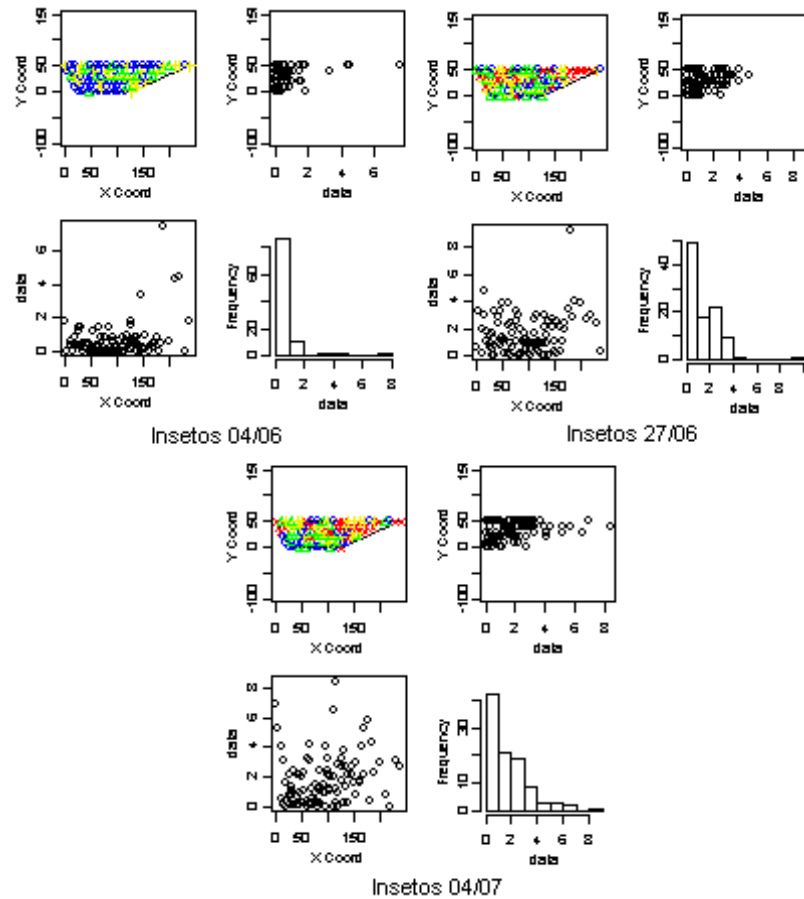


Figura 1.5 – Gráficos de dispersão e histogramas para o número médio de insetos - Sítio Novo II.

Considerando o parâmetro  $\beta$ , observa-se que na Fazenda São Paulo seus valores foram negativos, o que já era esperado, uma vez que essa Fazenda estava isolada de outros plantios de cebola, resultando em baixas médias de infestação. Já na Estância Bela Vista que estava rodeada por plantio de cebola já infestados pelo tripses essas estimativas foram altas. O Sítio Novo II obteve estimativas próximas a zero, porém, como já foi visto anteriormente, essa propriedade forneceu baixas médias para número de insetos por folha.

O resumo e o gráfico de convergência dos parâmetros estão no Anexo A. Pode-se observar que houve uma boa convergência dos parâmetros, apesar desses apresentarem alta variabilidade em alguns casos.



Tabela 1.3 – Modelagem geoestatística e intervalos para os parâmetros

Propriedade	Data	$\beta$	$\sigma^2$	$\phi$	$\tau^2$	LogMV
Fazenda São Paulo	10/07	-0,6302	0,2763	3,7852	0,0000	231,2
		(-0,57; -0,46)	(0,03; 0,34)	(0,57; 36,86)	(0,00; 7,04)	
	24/07	-0,5123	0,3773	13,1430	0,5781	186,5
		(-0,81; -0,62)	(0,16; 1,01)	(5,67; 82,86)	(0,00; 5,40)	
	31/07	-0,9424	0,2550	1,0000	0,0000	320,6
		(-0,97; -0,88)	(0,04; 0,30)	(0,56; 16,04)	(0,00; 5,31)	
Estância Bela Vista	08/08	2,3522	0,1994	14,2512	0,9220	127,5
		(2,34; 2,38)	(0,18; 0,28)	(10,72; 21,19)	(0,39; 1,28)	
Sítio Novo II	04/06	-0,4212	0,5535	1,0002	0,0000	160,7
		(-0,45; -0,15)	(0,07; 0,58)	(0,57; 110,18)	(0,00; 5,94)	
	27/06	0,2341	0,2233	23,6466	2,0880	55,53
		(0,14; 0,25)	(0,14; 0,73)	(5,60; 40,50)	(0,00; 4,28)	
	04/07	0,2697	0,2660	23,5905	2,3366	51,63
		(0,20; 0,30)	(0,22; 0,46)	(13,68; 37,71)	(0,90; 3,08)	

A seguir, foram construídos os mapas de predição para a área estudada considerando os parâmetros estimados da Tabela 1.3.

Comparando o mapa de predição (Figura 1.6) com o gráfico de pontos (Figura 1.3), observa-se que seu comportamento é semelhante ao observado, uma vez que as áreas de alta e baixa infestação são as mesmas. Observa-se ainda que as tonalidades mais claras indicam baixa infestação enquanto que as tonalidades mais escuras indicam maior infestação.

O mapa de predição 1.7 e o gráfico de pontos (Figura 1.4) apresentam comportamento semelhante uma vez que as áreas de alta e baixa infestação são as mesmas.

O comportamento do mapa de predição 1.8 e do gráfico de pontos (Figura 1.5) é semelhante uma vez que as áreas de alta e baixa infestação são as mesmas.

### Testes de Aleatorização para Padrão Temporal Aleatório

A Tabela 1.4 é referente aos resultados obtidos para os teste de aleatorização para o número de insetos por folha ao longo do tempo.

A influência do tempo foi observada apenas no Sítio Rosário, no qual não houve relação do número médio de insetos por folha e a localização da planta. Provavelmente, a ausência de relação entre o número de insetos por folha em relação à amplitude do intervalo de tempo se deva às amplitudes das coletas. Espera-se que, se as observações fossem coletadas em intervalos menores, talvez essa relação fosse mais evidente.

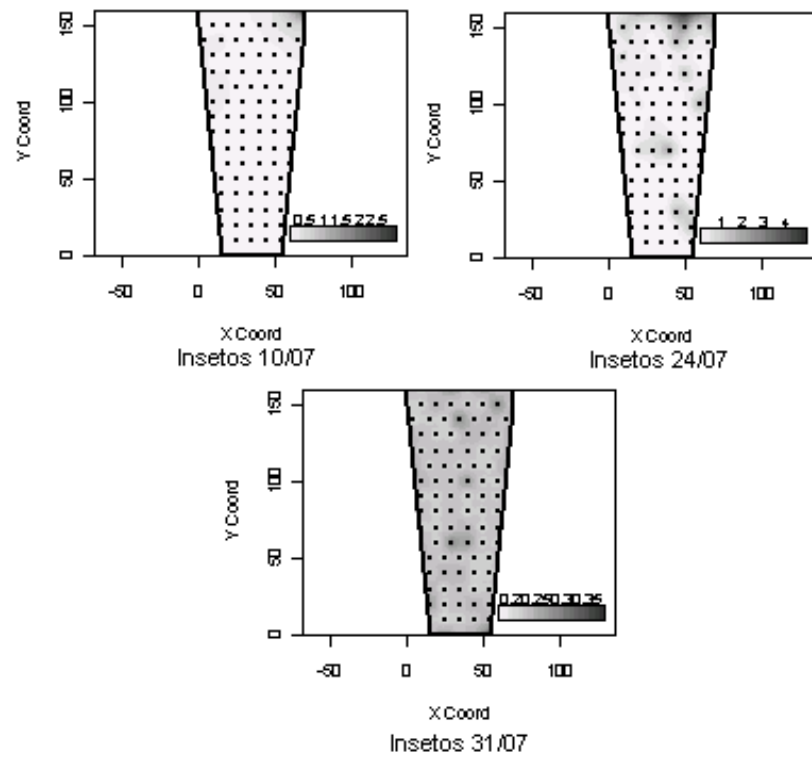


Figura 1.6 – Mapa de predição - I

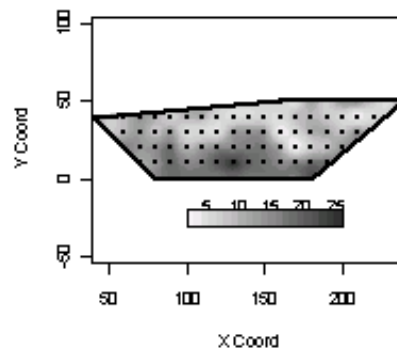


Figura 1.7 – Mapa de predição - Estância Bela Vista.

Como análise exploratória da relação entre o número de insetos e a distância Euclidiana no tempo os gráficos de dispersão apresentados na Figura 1.9, para o Sítio Rosário, confirmam evidência de dependência temporal.

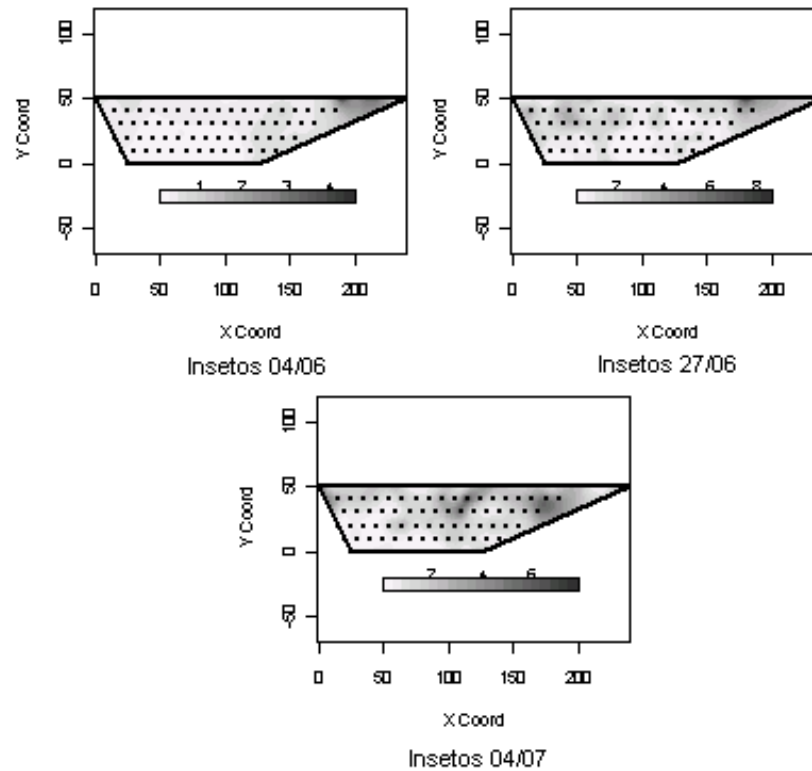


Figura 1.8 – Mapa de predição - Sítio Novo II.

Tabela 1.4 – Valor de p para testes de aleatorização para padrão ao longo do tempo

Propriedade	Distância	Insetos
Fazenda	Distância	0,1220
São Paulo	Distância recíproca	0,1270
Estância	Distância	0,3334
Bela Vista	Distância recíproca	0,4321
Sítio	Distância	<b>0,0267</b>
Rosário	Distância recíproca	<b>0,0025</b>
Sítio	Distância	0,8685
Novo II	Distância recíproca	0,5766

O modelo de regressão linear simples para o número médio de insetos por folha no Sítio Rosário considerando a distância euclidiana foi:  $Insetos = 25,36 + 0,2584t$  e para a distância recíproca foi:  $Insetos = 40,42 - 135,1t$ , confirmando evidência de relação entre a distância do número de insetos por folha e a distância do tempo.

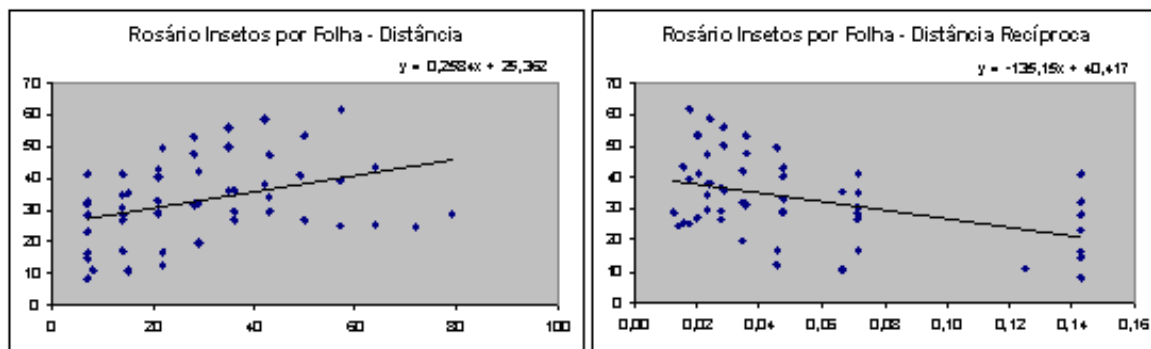


Figura 1.9 – Relação das distâncias

## 1.5 Conclusões

Quando se consideram os dados para cada data, pode-se observar um padrão de dependência espacial em alguns casos, mas em outros, isso não é tão evidente, o que se deve, possivelmente, ao fato da variabilidade de outras covariáveis não medidas, gerando condições heterogêneas de coleta.

Aparentemente, existe influência do tipo de plantio na vizinhança sobre o número de insetos por folha na planta, pois a Estância Bela Vista, que tinha na vizinhança área plantada no sistema bulbinho, já infestada de tripses, foi a que apresentou maiores médias para o número de insetos por folha e maiores proporções de plantas infestadas e a Fazenda São Paulo, isolada de outros plantios de cebola, foi a que apresentou a menor proporção de plantas infestadas, sendo que, essa proporção, no geral, foi aumentando no decorrer do tempo, indicando que possa realmente haver influência da infestação das plantas vizinhas.

O teste de aleatorização de Mantel mostrou-se bastante satisfatório para a verificação do padrão espacial aleatório na dispersão do tripses do prateamento.

Recomenda-se que a amostragem seja realizada para distâncias menores, assim, seria mais fácil de observar a existência de padrão espacial, uma vez que o inseto não apresenta muita mobilidade.

## 1.6 Referências

- ALCÂNTARA et al. - Modelagem da profundidade secchi e da concentração de clorofila A no estuário do Rio Anil, São Luís-MA. **Caminhos de geografia**, 2(13), p.19 – 40, 2004
- BRAGA, L. P. V. - Geoestatística e aplicações. **9º SINAPE**, 1990.
- CAMARGO, E. C. G. [http://www.dpi.inpe.br/gilberto/tutoriais/gis\\_ambiente/5geoest.pdf](http://www.dpi.inpe.br/gilberto/tutoriais/gis_ambiente/5geoest.pdf)
- CHRISTENSEN, O. F.; RIBEIRO JR., P.J. geoRglm: A package for generalised linear spatial models. **R-NEWS**, Vol 2, No 2, 26-28, 2002.
- COSTA, A. A. & MEDEIROS, Z. P. - O piolho da cebola (Thrips Tabaci Lindeman). **Agronomia**, 9(1), 77 – 86, 1950.
- DIGGLE, P. J.; RIBEIRO JR., P. J. Model-Based geostatistics. **Springer**, New York, 228p. 2007.
- DIGGLE, P. J.; TAWN, J. A.; MOYEED, R. A. Model-Based geostatistics. **Applied Statistics**, v. 47, p.299-350, 1998.
- DIGGLE, P. J.; RIBEIRO JR, P. J.; CHRISTENSEN, O. F. An introduction to Model-Based Geostatistics **Spatial statistics and computational methods**, 2003.
- FARIAS, P. R. S. et al., Geostatistical Characterization of the Spatial Distribution of Xylella fastidiosa Sharpshooter Vectors on Citrus **Neotropical Entomology**, p.13-30, 2004.
- FORTES et al., Modelagem geoestatística por *Ascaris lumbricóides*. **Caderno da saúde pública**, Rio de Janeiro, 20(3): p.727-734, 2004.
- Krige, D. G. **A statistical approach to some basic mine evaluation problems on the Witwatersrand**, Journal South African. Institute Mining Metall. n. 52. p. 119-139. 1951.
- MANLY, B. F. J. - Randomization, Bootstrap and Monte Carlo Methods in Biology, 2006 - Chapman & Hall - Flórida, 460 p.
- \* Mantel (1967)
- MONTAGNA, M. A. - Distribuição espacial e amostragem sequencial da mosca-branca *Bemisia Tabaci Raça B* (Homoptera: Aleyrodidae) no agroecossistema do melão, 2001 - Ribeirão Preto - USP - Dissertação de mestrado.
- OLIVEIRA, M. C. N. - Métodos de estimação de parâmetros em modelos geoestatísticos com diferentes estruturas de covariâncias: uma aplicação ao teor de cálcio no solo, 2003 - Piracicaba - SP - Tese de doutorado
- ORTIZ, G. C. - Aplicação de métodos geoestatísticos para identificar a magnitude e a estrutura da variabilidade espacial de variáveis físicas do solo, 2002 - ESALQ/USP - Piracicaba - SP - Dissertação de mestrado
- PILON, A. A. - Métodos para incorporação da dependência espacial na análise de dados experimentais, 2004 - ESALQ/USP - Piracicaba - SP - Dissertação de mestrado
- R Development Core Team (2006). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- RIBEIRO JR., P. J. - Abordagem geoestatística em um estudo de variabilidade espacial, 1995 - ESALQ/USP - Piracicaba - SP - Dissertação de mestrado
- \* (RIBEIRO JR, 2004,
- RIBEIRO JR., P. J. ; DIGGLE, P. J. geoR: a package for geostatistical analysis **R-NEWS**, 1(2):15-18. June, 2001.

RUIZ, R. C. - Modelagem da Distribuição Espaço-Temporal da Broca do Café (*Hypothenemus hampei Ferrari*) em uma Cultura de Região Central Colombiana, 2002 - Piracicaba - SP - Dissertação de mestrado.

SATO, M. E. - Avaliação do Dano e Controle do *Thrips tabaci* Lindeman, 1988, na Cultura da Cebola (*Allium cepaL.*), 1989 - Piracicaba - SP - Dissertação de mestrado.

SCHOTZKO, D. J. & KNUDSE N. G. R. - Use of Geostatistics to evaluate a spatial simulation of russian wheat aphid (Homoptera: Aphedidae) movement behavior on preferred and nonpreferred hosts. **Environmental of Entomology**, V21, n.6, p.1271 – 1282, 1992.

ZAMBOTI, J. L. **Mapas pluviométricos mensais no verão para o Estado do Paraná, utilizando métodos geoestatísticos**, 2001. 134f. Tese (Doutorado em Agronomia/Energia da Agricultura - Faculdade de Ciências Agronômicas, UNESP/Botucatu, 2001

(LEMBRAR DE INCLUIR AGRADECIMENTOS AO FINAL)

GREGO, Célia Regina, VIEIRA, Sidney Rosa and LOURENCAO, André Luiz. Spatial distribution of *Pseudaletia sequax* Franclemont in triticale under no-till management. *Sci. agric.* (Piracicaba, Braz.), July/Aug. 2006, vol.63, no.4, p.321-327. ISSN 0103-9016.

@Bookisaaks+srivastava:1989, author = Isaaks, E. H. and Srisvastava, R. M., title = An Introduction to Applied Geostatistics, publisher = Oxford University Press, address = New York, year = 1989,

@Bookgoovaerts:1997, author = Goovaerts, P., title = Geostatistics for Natural Resources Evaluation, publisher = Oxford University Press, address = New York, year = 1997,

REESCREVER AQUI: 2. DADOS 3. MANTEL 4. MODELO GEOEST. 5. RESULTADOS E DISCUSSAO

## **2 ANEXO A**

Tabela 2.1 – Estatísticas Descritivas - Fazenda São Paulo

Datas	Média	Var.	Mín.	Máx.	N	p
10/07	0,25	0,43	0,00	6,20	100	0,56
24/07	0,56	1,08	0,00	5,17	100	0,51
21/07	0,08	0,02	0,00	1,00	100	0,35
07/08	1,18	2,56	0,00	10,00	98	0,69
14/08	1,87	3,49	0,00	9,00	100	0,84
21/08	2,09	2,94	0,00	7,78	100	0,89
28/08	2,32	4,50	0,00	10,50	100	0,83
04/09	6,94	15,61	1,11	30,00	100	1,00

Tabela 2.2 – Estatísticas Descritivas - Estância Bela Vista

Datas	Média	Var.	Mín.	Máx.	N	p
11/07	6,03	11,25	1,25	22,50	100	0,89
01/08	11,00	25,76	1,33	21,00	100	1,00
08/08	12,14	34,89	1,17	25,83	84	1,00
14/08	6,03	20,53	0,00	26,00	99	1,00
09/09	7,15	34,12	0,00	30,00	99	0,98

Tabela 2.3 – Estatísticas Descritivas - Sítio Rosário

Datas	Média	Var.	Mín.	Máx.	N	p
21/06	0,52	1,20	0,00	7,00	50	0,56
29/06	1,67	1,73	0,00	5,40	50	0,90
07/07	0,98	0,70	0,00	3,20	48	0,90
14/07	1,44	0,80	0,00	3,40	50	0,98
21/07	2,30	4,14	0,00	8,00	50	0,94
28/07	4,42	10,68	0,33	12,67	50	1,00
04/08	4,72	6,24	0,00	12,67	50	0,98
11/08	7,63	8,81	2,86	17,20	50	1,00
18/08	8,61	16,25	0,14	20,00	50	1,00
25/08	5,80	11,13	0,20	13,29	50	1,00
03/09	2,95	10,21	0,00	10,83	50	0,98



Tabela 2.4 – Estatísticas Descritivas - Sítio Novo II

Datas	Média	Var.	Mín.	Máx.	N	p
04/06	0,60	1,11	0,00	7,50	100	0,69
19/06	0,67	0,60	0,00	4,14	100	0,78
27/06	1,53	2,00	0,00	9,29	100	0,88
28/06	2,22	3,33	0,00	9,13	100	0,90
04/07	1,76	2,79	0,00	8,43	100	0,93
11/07	3,15	12,32	0,00	31,00	100	0,96
24/07	2,80	6,89	0,00	14,50	100	0,84
31/07	1,25	2,77	0,00	7,73	100	0,71
07/08	1,85	5,17	0,00	15,00	100	0,74

Tabela 2.5 – Resumo estatístico para  $\beta$ 

Propriedade	Data	Mín.	Q1	Md	$\bar{X}$	Q3	Máx.	Var.
	10/07	-0,60	-0,55	-0,53	-0,53	-0,52	-0,34	0,001
Fazenda São Paulo	24/07	-0,86	-0,76	-0,73	-0,73	-0,70	-0,52	0,002
	31/07	-1,00	-0,94	-0,92	-0,92	-0,91	-0,84	0,001
Estância Bela Vista	08/08	2,33	2,35	2,36	2,36	2,36	2,39	0,0001
	04/06	-0,48	-0,41	-0,30	-0,31	-0,23	-0,05	0,009
Sítio Novo II	27/06	0,11	0,19	0,21	0,21	0,23	0,30	0,001
	04/07	0,18	0,25	0,26	0,26	0,28	0,37	0,001

Tabela 2.6 – Resumo estatístico para  $\sigma^2$ 

Propriedade	Data	Mín.	Q1	Md	$\bar{X}$	Q3	Máx.	Var.
	10/07	0,01	0,12	0,25	0,22	0,29	0,42	0,01
Fazenda São Paulo	24/07	0,09	0,26	0,33	0,39	0,44	1,20	0,04
	31/07	0,01	0,19	0,22	0,21	0,24	0,39	0,004
Estância Bela Vista	08/08	0,15	0,20	0,21	0,21	0,22	0,41	0,001
	04/06	0,05	0,14	0,19	0,29	0,49	0,68	0,03
Sítio Novo II	27/06	0,08	0,18	0,20	0,23	0,23	0,87	0,01
	04/07	0,15	0,27	0,30	0,31	0,33	0,98	0,01

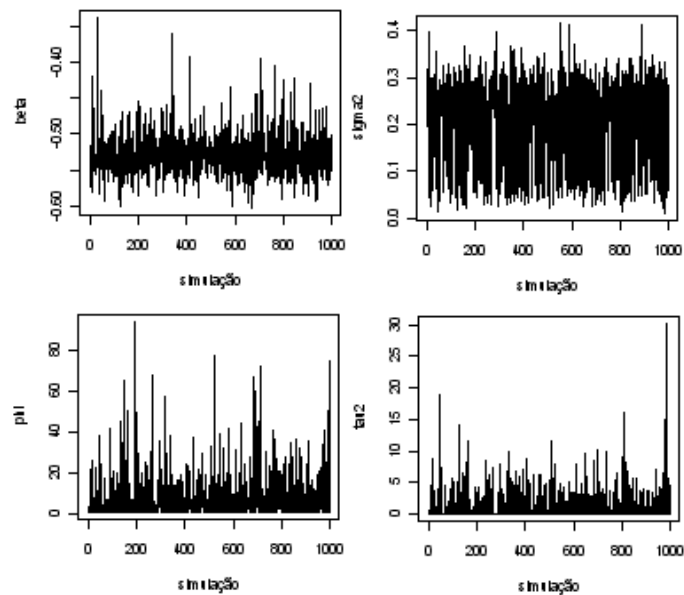


Figura 2.1 – Gráficos da convergência dos parâmetros - Fazenda São Paulo - 10/07.

Tabela 2.7 – Resumo estatístico para  $\phi$ 

Propriedade	Data	Mín.	Q1	Md	$\bar{X}$	Q3	Máx.	Var.
Fazenda São Paulo	10/07	0,54	1,00	3,84	7,30	8,24	93,86	110,22
	24/07	0,57	15,18	27,83	32,65	46,34	111,20	470,72
	31/07	0,54	0,60	1,00	2,87	3,37	67,15	27,93
Estância Bela Vista	08/08	1,03	14,38	15,72	15,79	17,12	28,54	6,51
	04/06	0,54	2,48	44,70	41,22	67,61	201,30	1283,98
Sítio Novo II	27/06	3,89	23,06	26,88	26,23	30,44	56,54	55,17
	04/07	4,85	21,61	24,51	24,86	27,88	49,41	33,48

Tabela 2.8 – Resumo estatístico para  $\tau^2$ 

Propriedade	Data	Mín.	Q1	Md	$\bar{X}$	Q3	Máx.	Var.
Fazenda São Paulo	10/07	0,00	0,00	0,01	1,07	1,30	30,23	5,33
	24/07	0,00	1,25	2,01	2,10	2,78	10,030	1,87
	31/07	0,00	0,00	0,01	0,47	0,01	25,57	3,19
Estância Bela Vista	08/08	0,00	0,83	0,95	0,93	1,05	1,60	0,047
	04/06	0,00	0,01	1,99	1,90	3,01	10,75	3,23
Sítio Novo II	27/06	0,00	2,12	2,54	2,51	2,95	10,09	0,85
	04/07	0,00	1,68	1,97	1,97	2,25	5,00	0,30

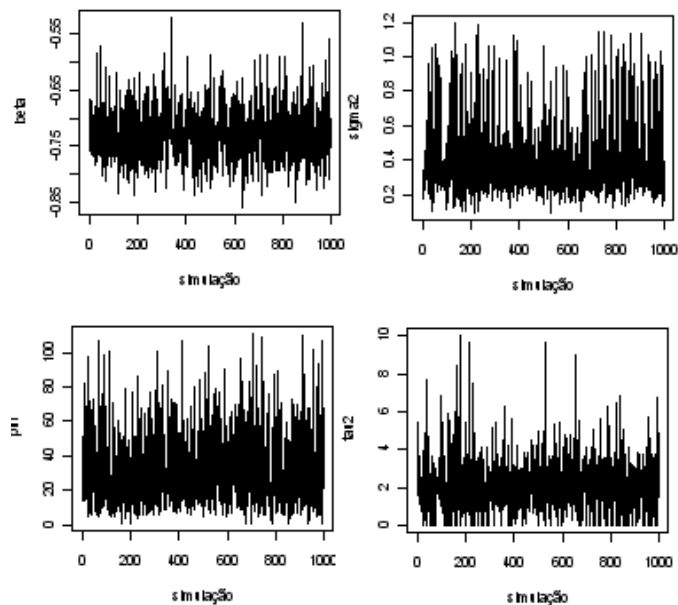


Figura 2.2 – Gráficos da convergência dos parâmetros - Fazenda São Paulo - 24/07.

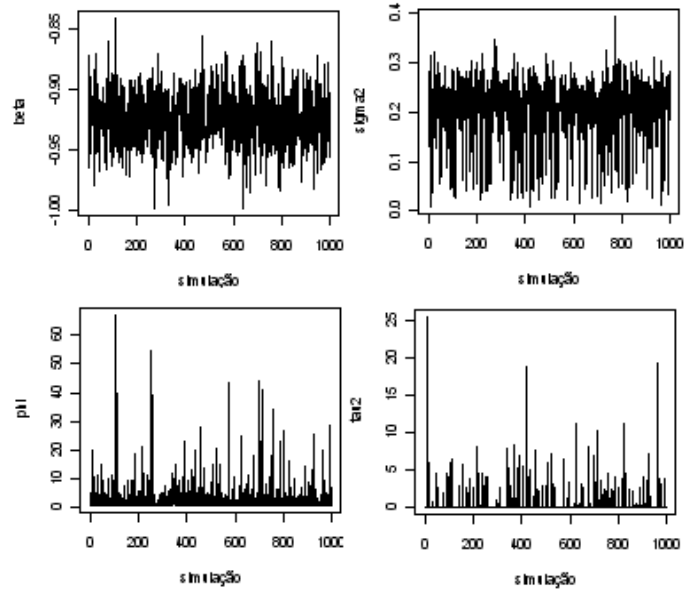


Figura 2.3 – Gráficos da convergência dos parâmetros - Fazenda São Paulo - 31/07.

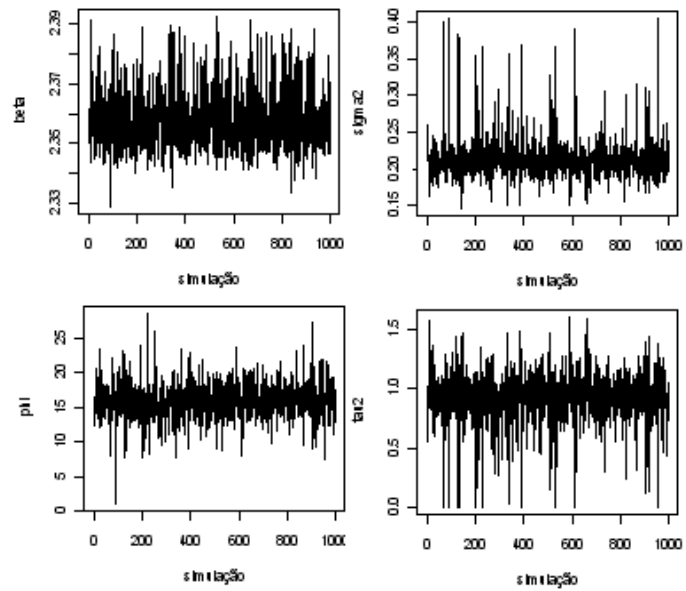


Figura 2.4 – Gráficos da convergência dos parâmetros - Estância Bela Vista - 08/08.

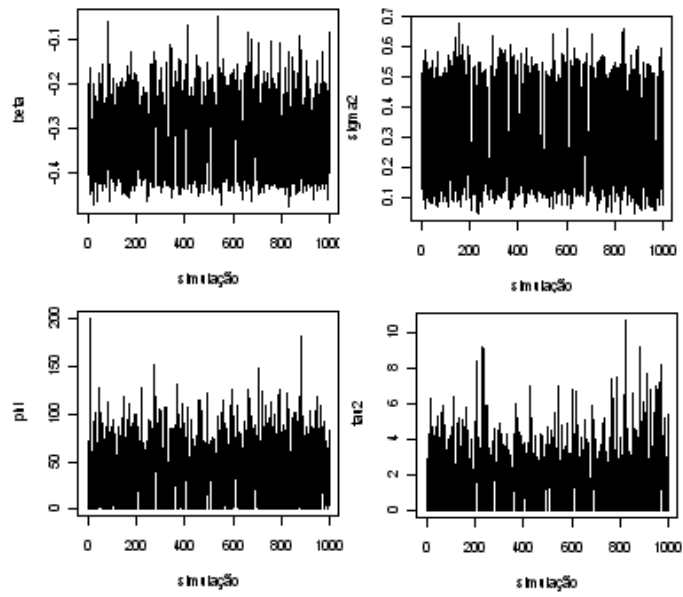


Figura 2.5 – Gráficos da convergência dos parâmetros - Sítio Novo II - 04/06.

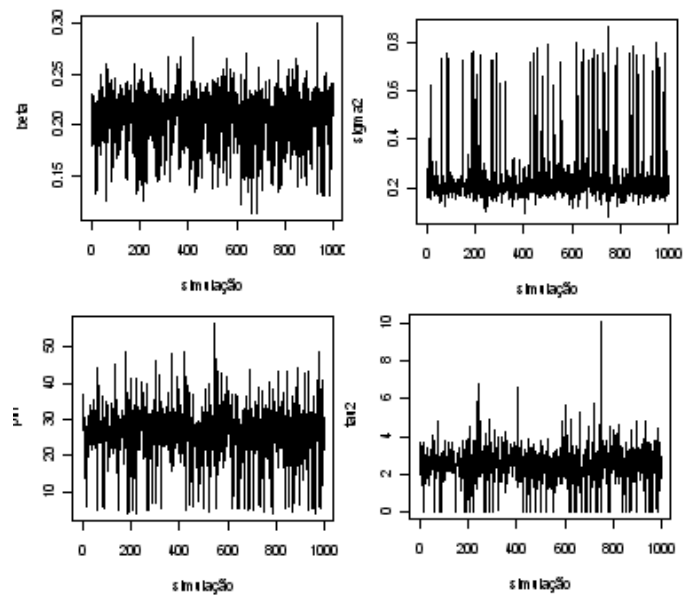


Figura 2.6 – Gráficos da convergência dos parâmetros - Sítio Novo II - 27/06.

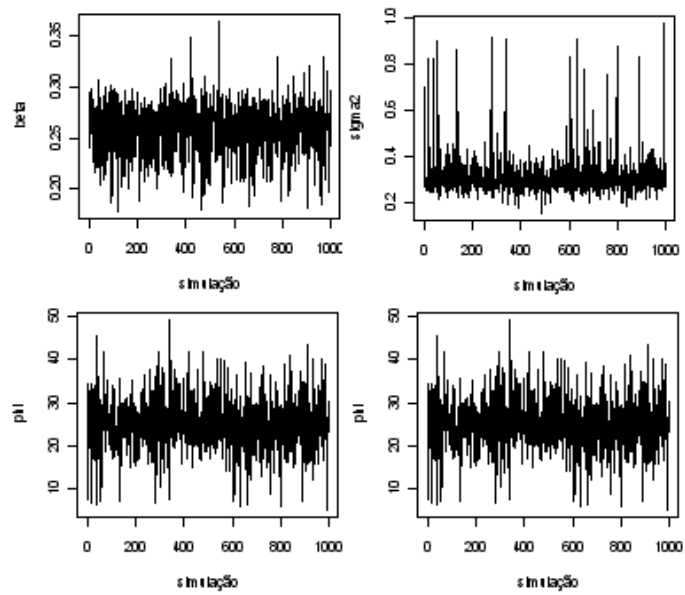


Figura 2.7 – Gráficos da convergência dos parâmetros - Sítio Novo II - 04/07.