

Palavras-chave: Padrões Espaciais, Morte Súbita dos Citrus, Censura em Campos Aleatórios Gaussianos

1 O pacote Rcitrus

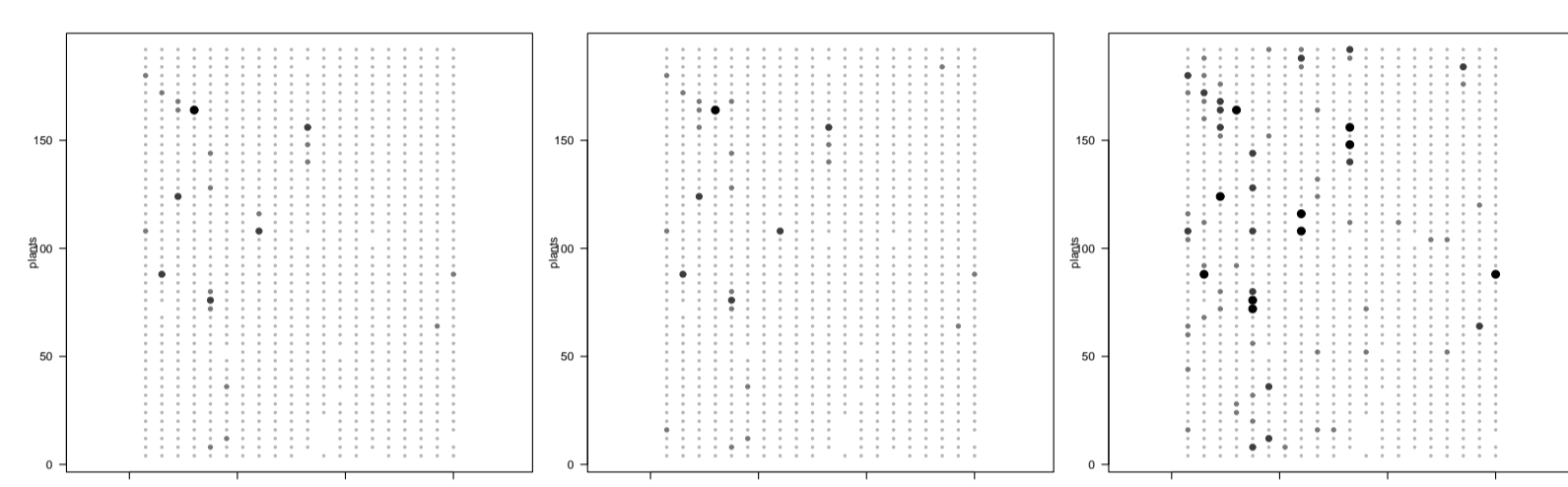
Pacote desenvolvido em R (R DEVELOPMENT CORE TEAM 2006). Adapta e implementa análise estatística para dados de incidência de doenças em plantas. Algumas funções utilizam funções dos pacotes **splancs** (ROWLINGSON, DIGGLE & OTHERS 2006), **geoR** (RIBEIRO JR & DIGGLE 2001), **survival** (original by Terry Therneau & ported by Thomas Lumley 2006), **MASS** (VENABLES & RIPLEY 2002) e **stats4** (R DEVELOPMENT CORE TEAM 2006).

1.1 Manipulação dos dados

- leitura automática de planilhas

```
> d.arr1d.arr1 <- read.citrus("vv303.csv",
  find.form="array", nrow=20,
  row.id=1, n.att=14, sep=";", dec=",")
```

- representação espaço temporal
- conversores para diferentes classes
- validação
- descrição e apresentação gráfica



1.2 Análise por quadrats

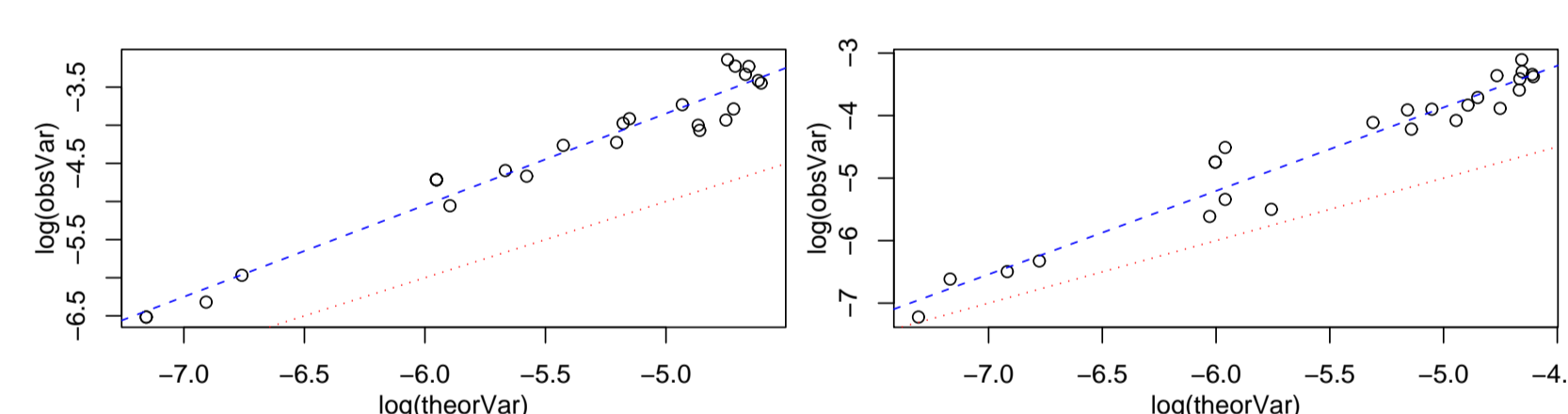
Fixos ou aleatórios e qualquer formato regular

```
> disp.quadrats(d.geo4, dx = 3, dy = 5,
  by = F, eval = 1:3, mod = "beta-binom",
  + usage = "mle")
```

```
$`3x5`
  N n nN prob theta p.value pattern
Av1 70 13 939 0.018 0.051 0.061 Random
Av2 70 13 939 0.021 0.034 0.135 Random
Av3 70 13 917 0.050 0.045 0.073 Random
```

1.3 Lei de Taylor

```
> Tay <- Taylor.citrus(d.geo4, dx = 5,
  + random = FALSE, death = 1:3)
> TayR <- Taylor.citrus(d.geo4, dx = 5,
  + random = TRUE, death = 1:3)
```



1.4 N° ótimo de quadrats aleatórios

A precisão aumenta quanto mais quadrats são analisados.

- realizar n análises com N quadrats aleatórios
- realizar outras n análises com $N + 1$ quadrats
- repetir o passo anterior até um critério ser atingido

```
> set.seed(153)
> Nquad <- sim.N.quadrats(d.geo4, dx = 10,
  + death = 1:3)
> qcc3 <- qcc(Nquad, type = "R")
```

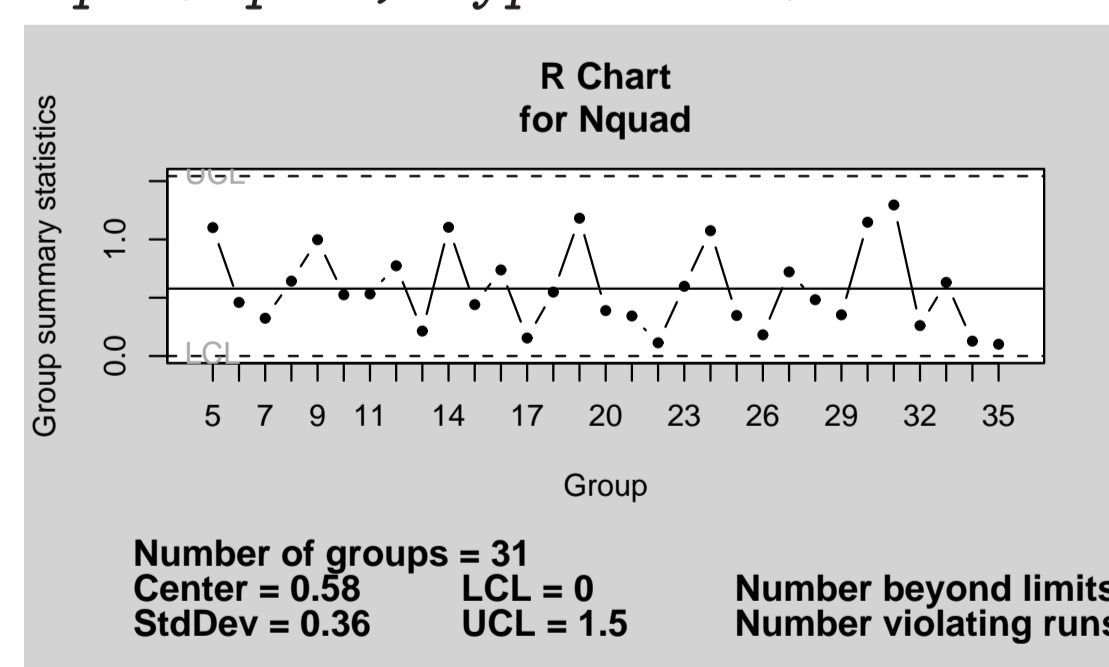
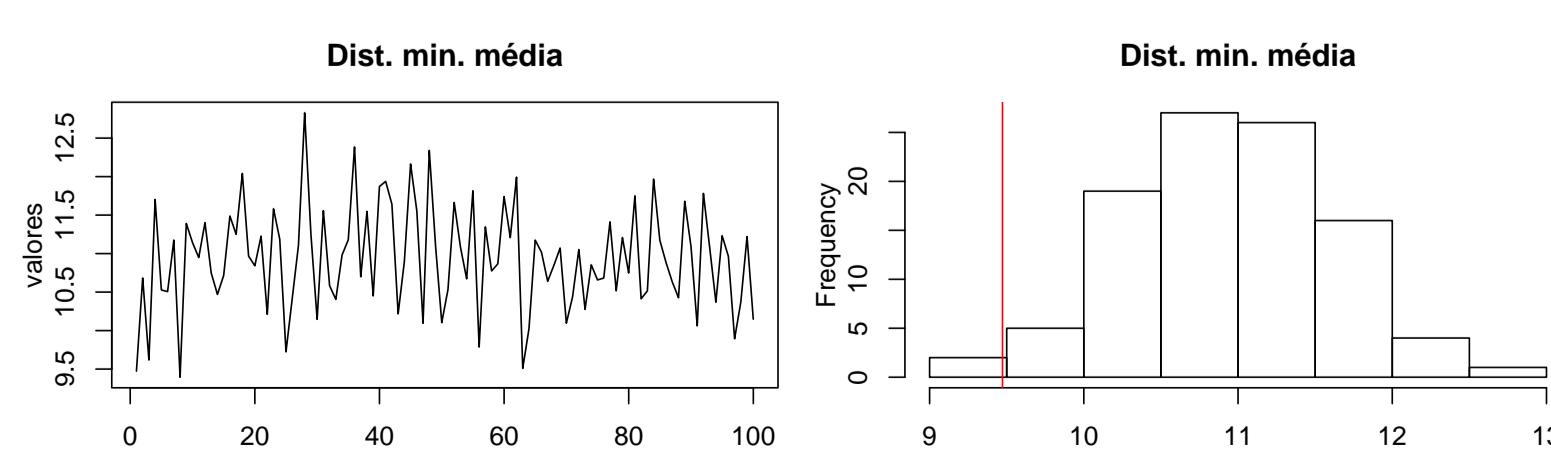


Figura 1: Carta de controle para a amplitude do índice de dispersão em 3 análises para cada N a partir de 5.

1.5 Distância mínima e número de vizinhos

Implementação de testes de Monte Carlo para a distância mínima média e número médio de vizinhos doentes.

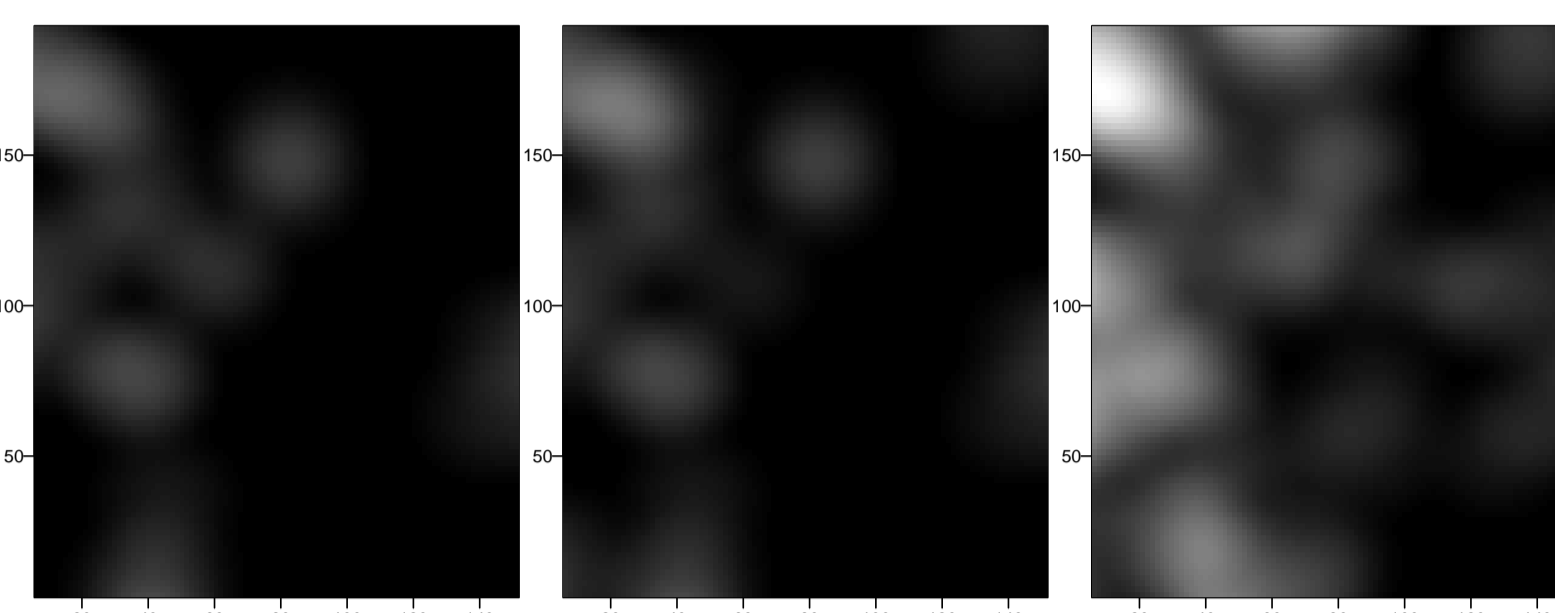
Vizualização gráfica do teste da para distância mínima média



1.6 Suavização por kernel

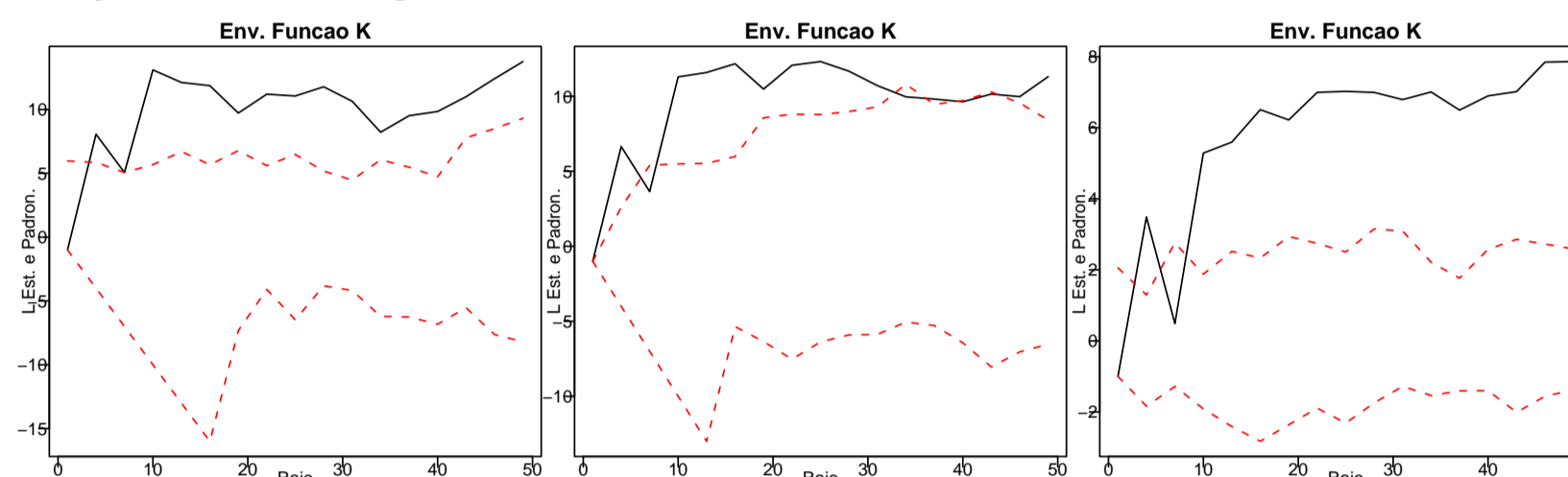
Adaptado do **splancs**

Kernel com escala de cores global.



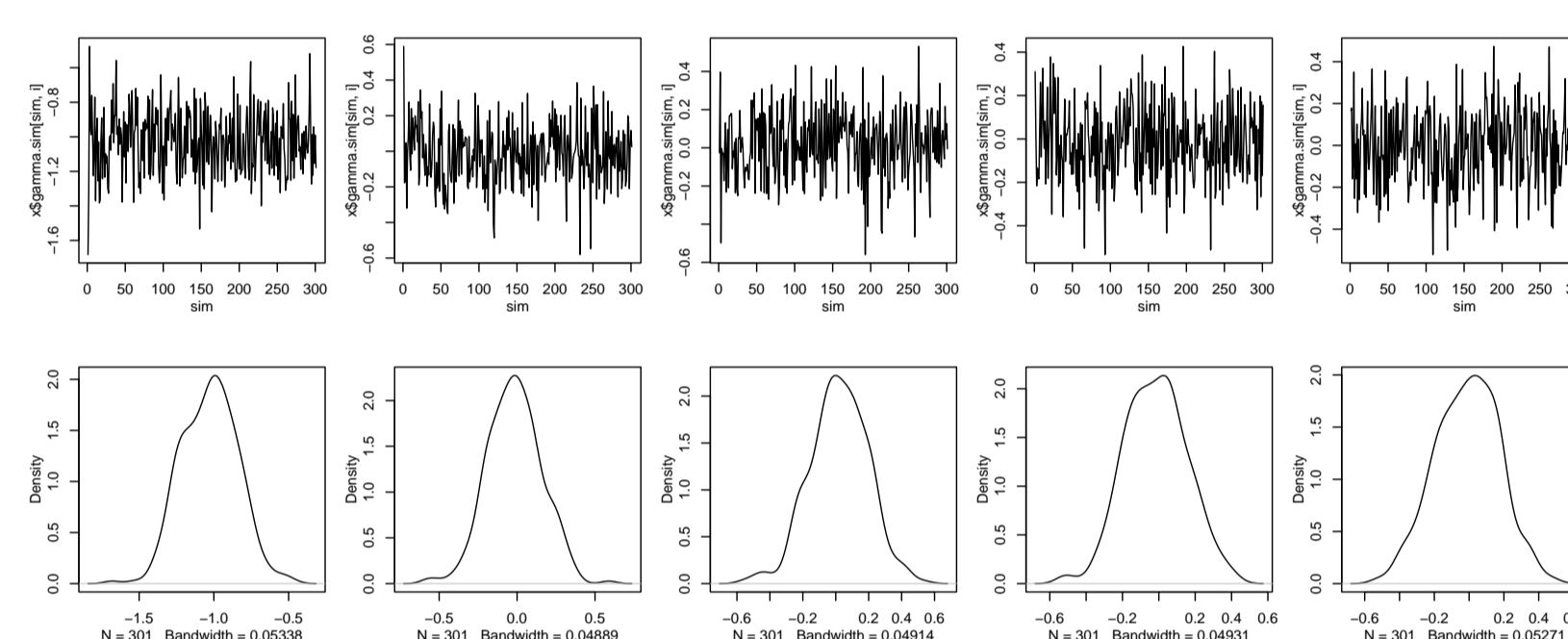
1.7 Função K de Ripley

Adaptada do **splancs**



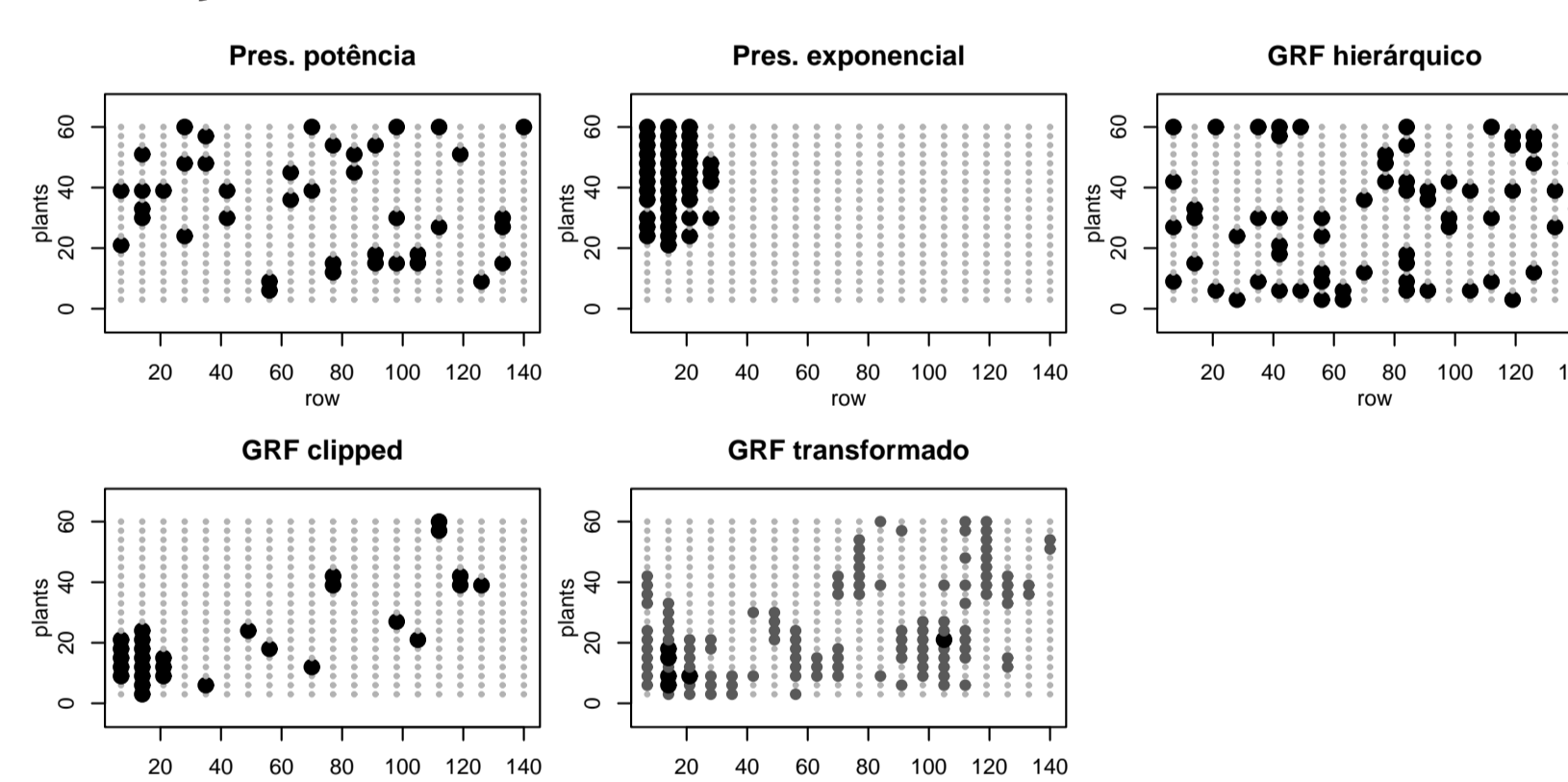
1.8 Modelo autológico

status de vizinhos como covariáveis (GUMPertz & RISTAINO 1997). Estimação: pseudo-verossimilhança com bootstrap ponderado via algoritmo amostrador de Gibbs. Teste: Monte Carlo.

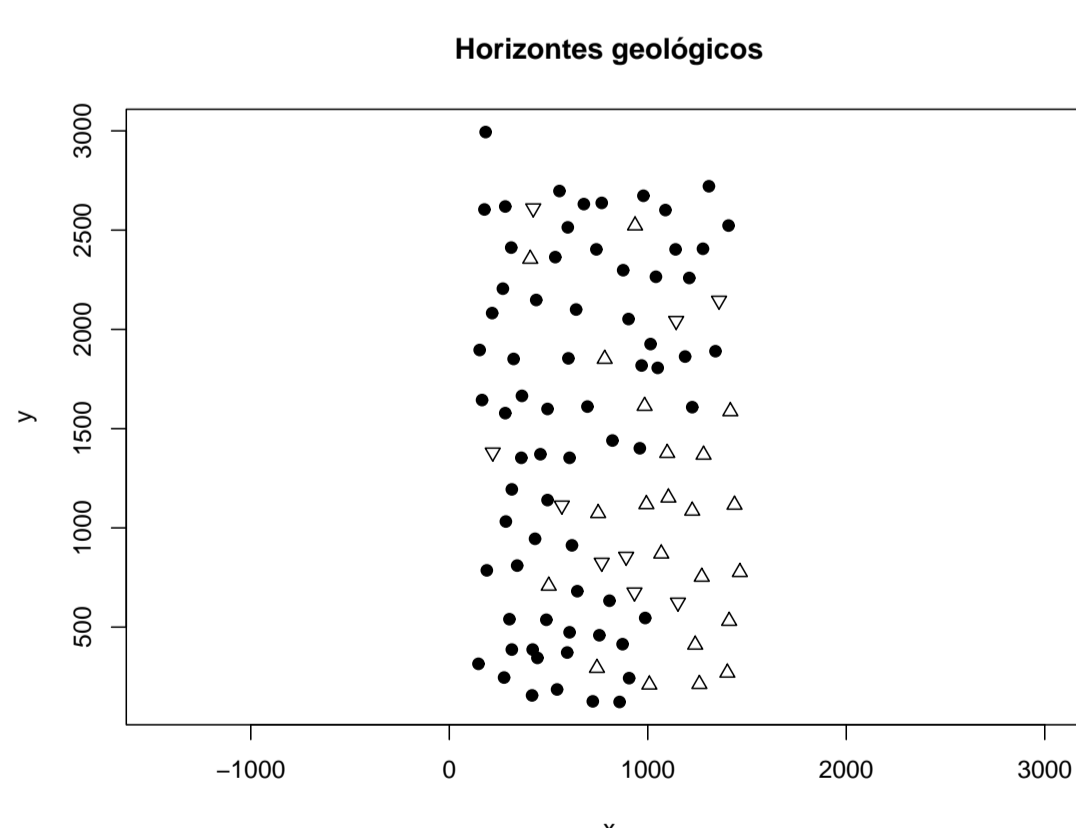


1.9 Simulação padrões espaciais

1. modelo de pressão infectiva potencial,
2. modelo de pressão infectiva exponencial,
3. modelo hierárquico $[Y|S]$, S : campo aleatório gaussiano,
4. grampeamento de um campo aleatório: $Y = 1$ se $S < z$ e $Y = 0$ se $S > z$, onde z é um valor de corte,
5. transformação de S : $Y = f_b(\phi(S))$, f_b é o quantil da distribuição bernoulli.



2 Censura em Campos Aleatórios Gaussianos



$$Y(x_i) = \beta F(x_i) + S(x_i) + Z_i : i = 1, \dots, n.$$

$$D = \{Y(x_j) = u; j \in J\} \cup \{Y(x_i) \in A_i; i \in I\}$$

$L(\theta|D) = [D|\theta]$ é difícil de calcular diretamente, pois

$$L(\theta; D) \equiv [u|\theta] \int [v|\theta] dv \quad (1)$$

2.1 Proposta de Stein (STEIN 1992)

- Uso de um método de MC
- Se a censura é a esquerda $v < b$ então

$$p(y < b|v, u) = \frac{\int_{-\infty, 0]^{n-m} \Phi\left(\frac{b-\alpha-r'v}{\tau}\right) [v|u] dv}{\int_{-\infty, 0]^{n-m} [v|u] dv} \quad (2)$$

onde $\alpha - r'v$ é a esperança marginal

- para aproximar o valor da integral utilizando MC

– Simular o primeiro valor de $[v|u; \theta]$

- Para os demais, simular de $[v|u, \tilde{v}^{q-1}; \theta]$, onde \tilde{v}^q são os q valores já simulados

- v_q é obtido por

$$\tilde{v}_q = \mu_q + \sigma_q \Phi^{-1} \left(U_q \Phi \left(\frac{-\mu_q}{\sigma_q} \right) \right) \quad (3)$$

- em cada simulação calcular

$$d_t = \prod_{q=1}^{n-m} \Phi \left(\frac{-\mu_q t}{\sigma_q} \right) \quad (4)$$

- aproximar 2 por

$$\frac{\sum_{t=1}^N d_t \Phi \left(\frac{t - \mu_{n+1,t}}{\tau} \right)}{\sum_{t=1}^N d_t} \quad (5)$$

- Utilizar θ_0 de referência, e utilizar

$$\frac{\ell(\theta; u, v)}{\ell(\theta_0; u, v)} = \frac{[u, \theta] \int_{-\infty, 0]^{n-m} [v|u; \theta] dv}{[u, \theta_0] \int_{-\infty, 0]^{n-m} [v|u; \theta_0] dv} \quad (6)$$

- aproximada por

$$\frac{[u|\theta] \sum_{t=1}^N d_t(\theta)}{[u|\theta_0] \sum_{t=1}^N d_t(\theta_0)} \quad (7)$$

2.2 (MILITO & UGARTE 1999) - EM

Usar $y_i^* = y_i - y_{(i)} G(\theta)_{(i)}^{-1} g(\theta)$

Passo E:

$$v_{1i}^*(t) = \hat{\mu}_i^*(t) - \hat{\sigma}_i^*(t) \left[\frac{v_i^* - \hat{\mu}_i^*(t)}{\hat{\sigma}_i^*(t)} \right] \Phi \left(\frac{v_i^* - \hat{\mu}_i^*(t)}{\hat{\sigma}_i^*(t)} \right) \quad (8)$$

$$v_{2i}^*(t) = (\hat{\mu}_i^*(t) + (\hat{\sigma}_i^*(t))^2) - \hat{\sigma}_i^*(t) (\hat{\mu}_i^*(t) + v_i^*) \left[\frac{v_i^* - \hat{\mu}_i^*(t)}{\hat{\sigma}_i^*(t)} \right] \Phi \left(\frac{v_i^* - \hat{\mu}_i^*(t)}{\hat{\sigma}_i^*(t)} \right) \quad (9)$$

Passo M:

$$\beta^{(t)} = (F^{*'} W F^*)^{-1} F^{*'} W y_1^* \quad (10)$$

$$(\hat{\sigma}^{(t)}) = \frac{\sum_{i=1}^n w_i (y_{1i}^*(t) - 2\hat{\mu}_i^*(t) + (\hat{\mu}_i^*(t))^2)}{n - p} \quad (11)$$

2.3 (ABRAHAMSEN & BENTH 2001)

- Data Augmentation Algorithm: Inicia com $p_0(y, \beta|u, A_i)$

1. obter $\tilde{v}_1, \dots, \tilde{v}_Q$ de $p^{(t)}(v, \beta|u, A_i)$ em dois passos:

- (a) obter $\tilde{\beta}_{(q)}$ de $p^{(t)}(\beta|u, A_i)$,
- (b) obter $\tilde{v}_{(q)}$ de $p(v|\tilde{\beta}_{(q)}, u, A_i)$ de $p(v|\tilde{\beta}_{(q)}, u)$ enquanto $\tilde{v}_{(q)} \in A_i$.

2. fazer a atualização $p^{(t+1)}(y, \beta|u, A_i) = \frac{1}{Q} \sum_{q=1}^Q p(y, \beta|u, v_q)$

- Iniciar $S = 2$ e dobrar para 10 – 13 primeiras iterações.

2.4 (DE OLIVEIRA 2005)

- análise completamente bayesiana,
- considerar a incerteza dos parâmetros, e
- considerar diferentes níveis de informação

Referências

- ABRAHAMSEN, P. & BENTH, F. E. (2001). Kriging with inequality constraints, *Mathematical Geology* 33(6): 719–744.
- DE OLIVEIRA, V. (2005). Bayesian inference and prediction of gaussian random fields based on censored data, *Journal of Computational and Graphical Statistics* 14(1): 95–115.
- GUMPertz, M. L. ; GRAHAM, J. M. & RISTAINO, J. B. (1997). Autologistic model of spatial pattern of phytophthora epidemic in bell pepper: Effects of soil variables on disease presence, *Journal of Agricultural, Biological and Environmental Statistics* 2(2): 131–156.
- MILITO, A. F. & UGARTE, M. D. (1999). Analyzing censored spatial data, *Mathematical Geology* 31(5): 551–561.
- original by Terry Therneau, S. & ported by Thomas Lumley (2006). *survival: Survival analysis, including penalised likelihood*. R package version 2.26.
- 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. [*http://www.R-project.org](http://www.R-project.org)
- RIBEIRO JR, P. J. & DIGGLE, P. J. (2001). geoR: a package for geostatistical analysis, *R-NEWS* 1(2): 14–18. ISSN 1609-3631. [*http://CRAN.R-project.org/doc/Rnews/](http://CRAN.R-project.org/doc/Rnews/)
- ROWLINGSON, B., DIGGLE, P. & OTHERS (2006). *splancs: Spatial and Space-Time Point Pattern Analysis*. Adapted and packaged for R by Roger Bivand and pcp functions by Giovanni Petris and goodness of fit by Stephen Eglen, R package version 2.01-17. [*http://www.r-project.org](http://www.r-project.org), <http://www.maths.lancs.ac.uk/~rowlings/splancs/>
- STEIN, M. L. (1992). Prediction and inference for truncated spatial data, *American Statistical Association, Institute of Mathematical Statistics and Interface Foundation of North America* 1(1): 91–110.
- VENABLES, W. N. & RIPLEY, B. D. (2002). *Modern Applied Statistics with S*, fourth edn, Springer, New York. ISBN 0-387-95457-0. [*http://www.stats.ox.ac.uk/pub/MASS4](http://www.stats.ox.ac.uk/pub/MASS4)