Geoestatística aplicada à AP

José P. Molin
ESALQ/USP
jpmolin@usp.br

www.agriculturadeprecisao.org.br

Objetivo

Abordar os conceitos fundamentais relacionados à geoestatistica aplicada à agricultura de precisão visando à modelagem de dependência espacial e interpolações

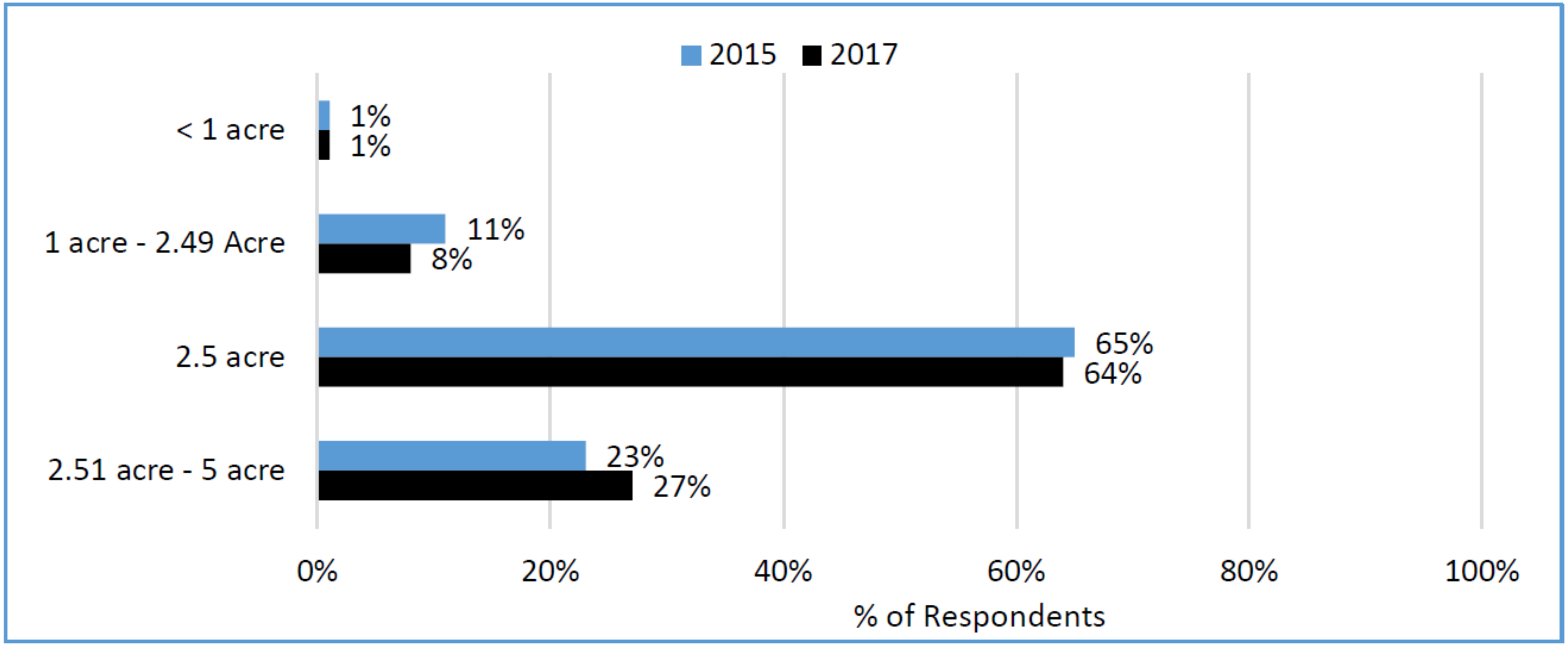
Como se define o “tamanho” mais correto de uma grade amostral?

ou

Como se produz mapas confiáveis?

A resposta certamente passa pela Geoestatística

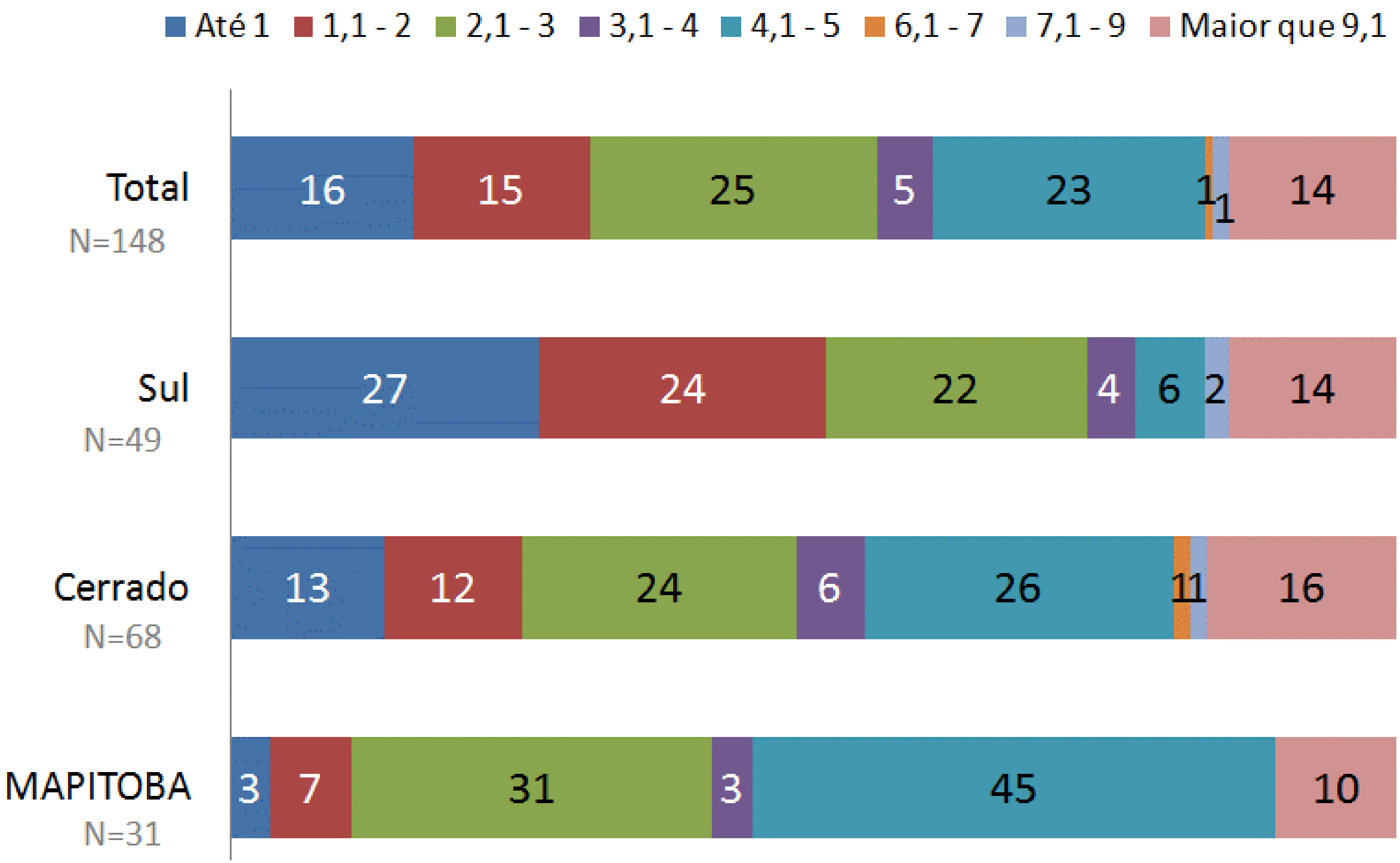
Tamanhos de grades amostrais praticadas (EUA)



B.Erickson, J. L. Lowenberg-DeBoer & J. Bradford (2017)

Tamanhos de grades amostrais praticadas (BR)

Todas as indicações em %. Base: Entrevistados que realizam o mapeamento da fertilidade do solo por grades N = 148



J. P. Molin (2017)

A estatística...

- Permite a caracterização de parâmetros da população a partir de parâmetros da amostra

- Pressupõem:
  - independência entre as observações;
  - distribuição normal dos dados;
  - variância e coeficiente de variação constantes.

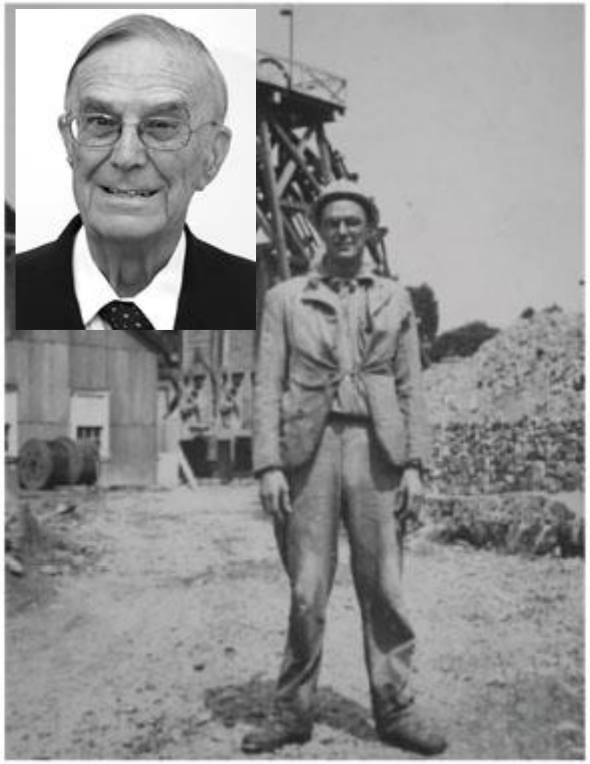
E a geoestatística?

Tem por objetivos:

- identificar, na aparente desordem das amostras, uma medida da correlação espacial;
- modelar e quantificar a dependência espacial;
- identificar padrões de amostragem adequados;
- estimar valores em locais não amostrados a partir das amostras existentes (interpolação por krigagem).

A origem...

- Daniel G. Krige (1951), na África do Sul, estudou teor de ouro em amostras;
- importância de ligar as variâncias com as distâncias entre as amostras;

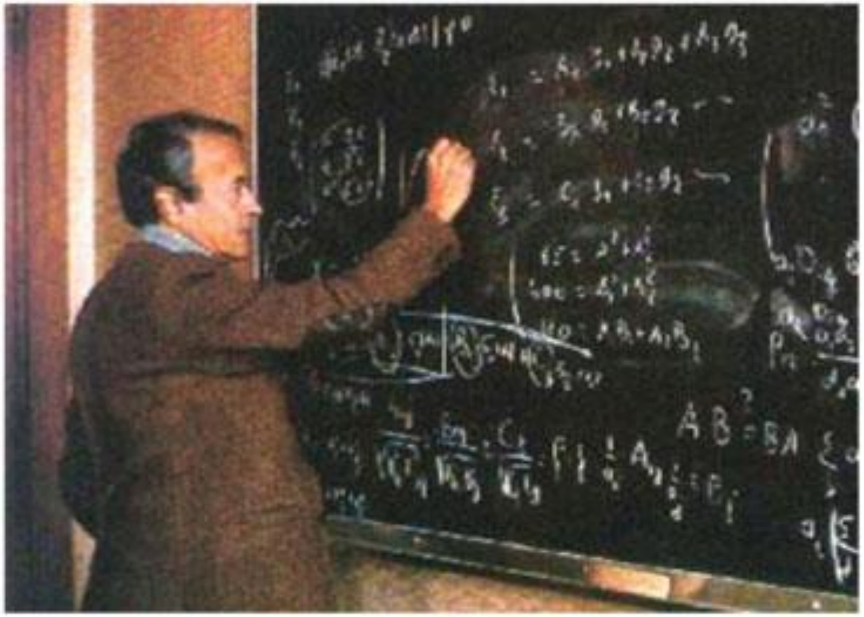


Danie Krige learning the trade, 1939.

Minnitt & Assibey-Bonsu (2014)

A origem...

- Matheron (1961), na França, elaborou a Teoria das Variáveis Regionalizadas: "esperança de que, na média, as amostras próximas, no tempo e espaço, sejam mais similares entre si do que as que estiverem mais distantes"



Professor Georges Matheron.

Minnitt & Assibey-Bonsu (2014)

Geoestatística e dependência espacial

- Valores de uma mesma variável, próximos um ao outro, têm a probabilidade de serem mais semelhantes do que valores mais distantes entre si.
- Como investigar essa dependência espacial?
- Calculando a variância entre pares de pontos e representando-a em um gráfico como função da distância entre os dois pontos – o semivariograma

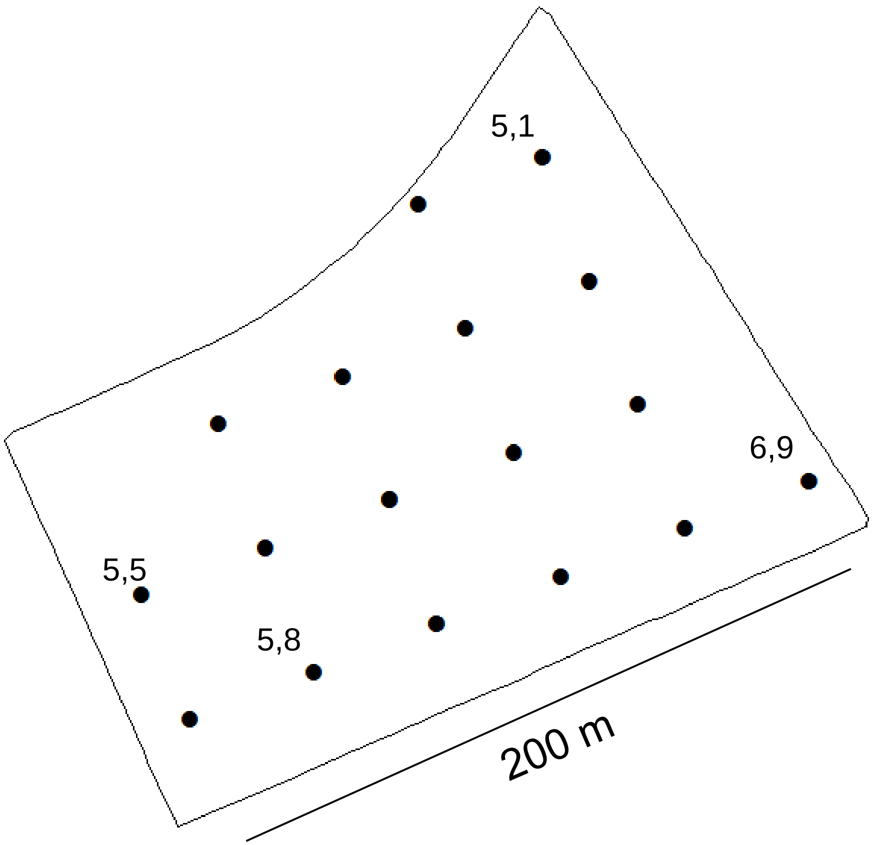
O que é o semivariograma?

- é um modelo que mede e descreve a dependência espacial entre as amostras.
- gráfico que representa a semivariância dos dados $\gamma(h)$ em relação à distância que os separa (h).

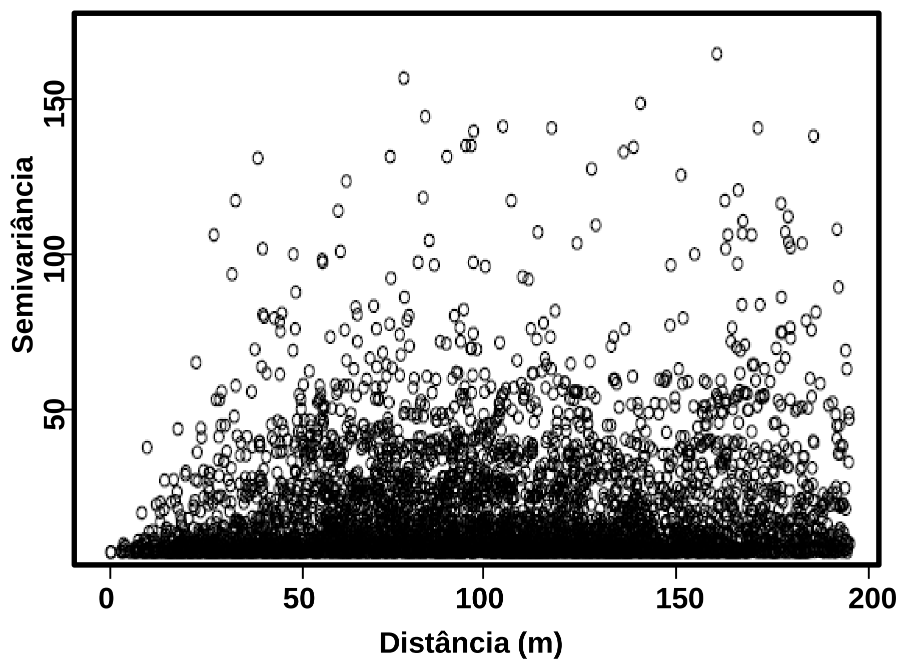
$$\gamma(h) = \frac{1}{2N(h)} \sum_{N(h)} [z(s+h) - z(s)]^2$$

$\gamma(h)$ é a semivariância
z é a variável em estudo
z(s+h) e z(s) são os valores separados por um vetor h
N(h) é o número de pares de valores [z(s+h)-z(s)] separados por um vetor h

Cálculo das semivariâncias:

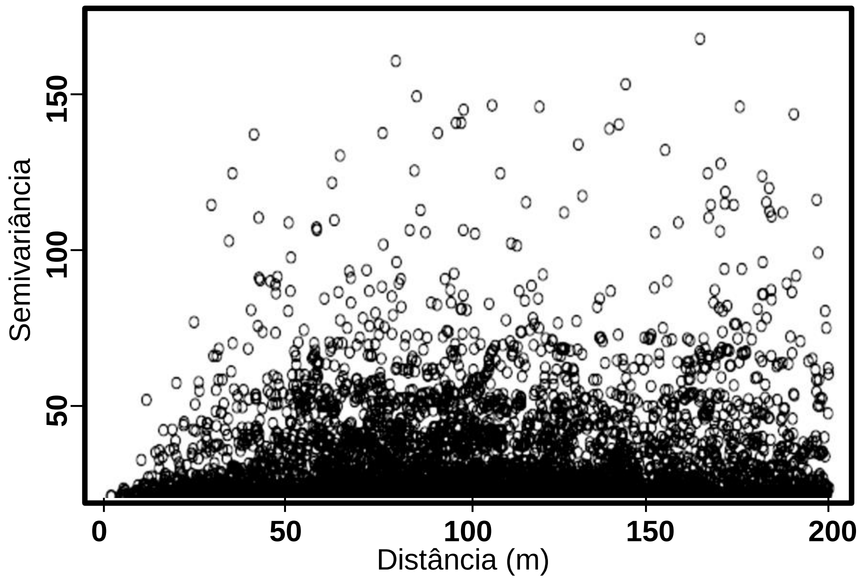


Semivariancia = $(z_2 - z_1)^2/2$

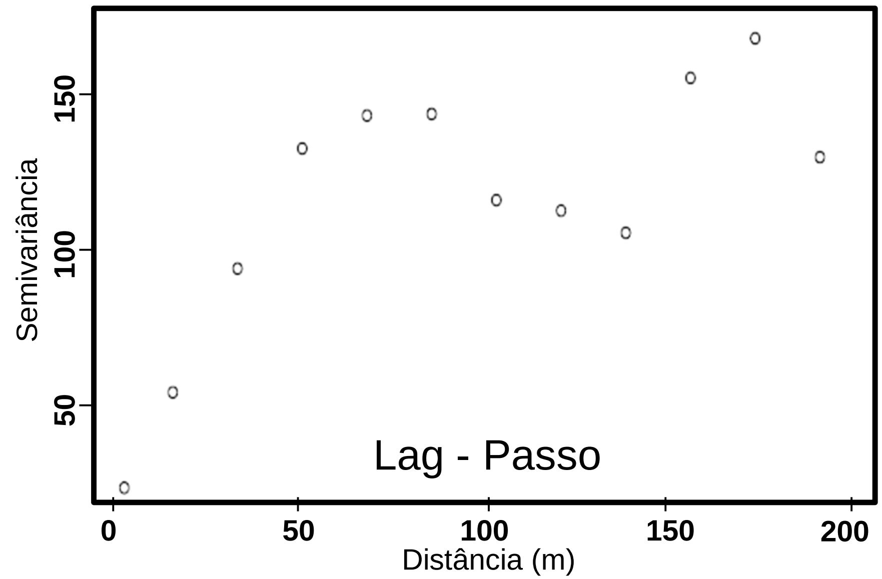


L.R. AMARAL, 2014

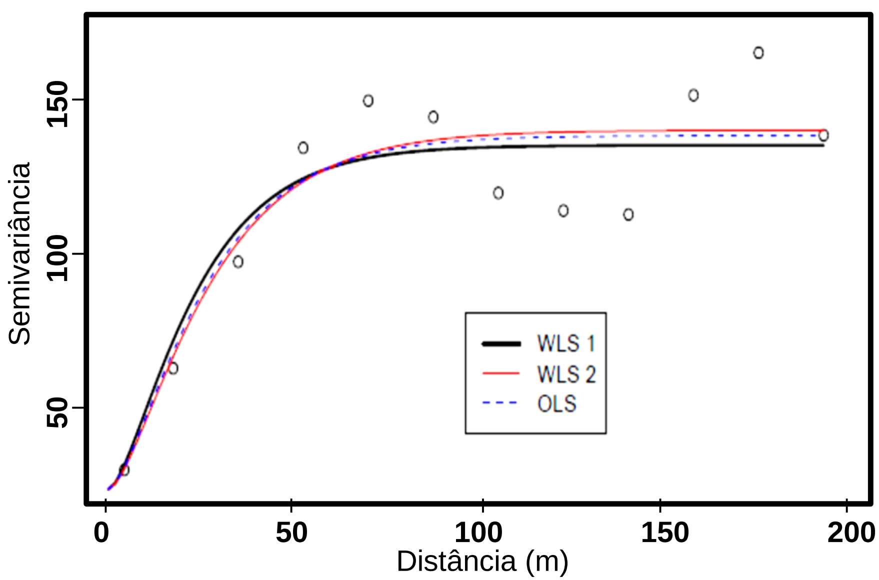
Nuvem variográfica



Semivariograma experimental



Semivariograma teórico (modelado)

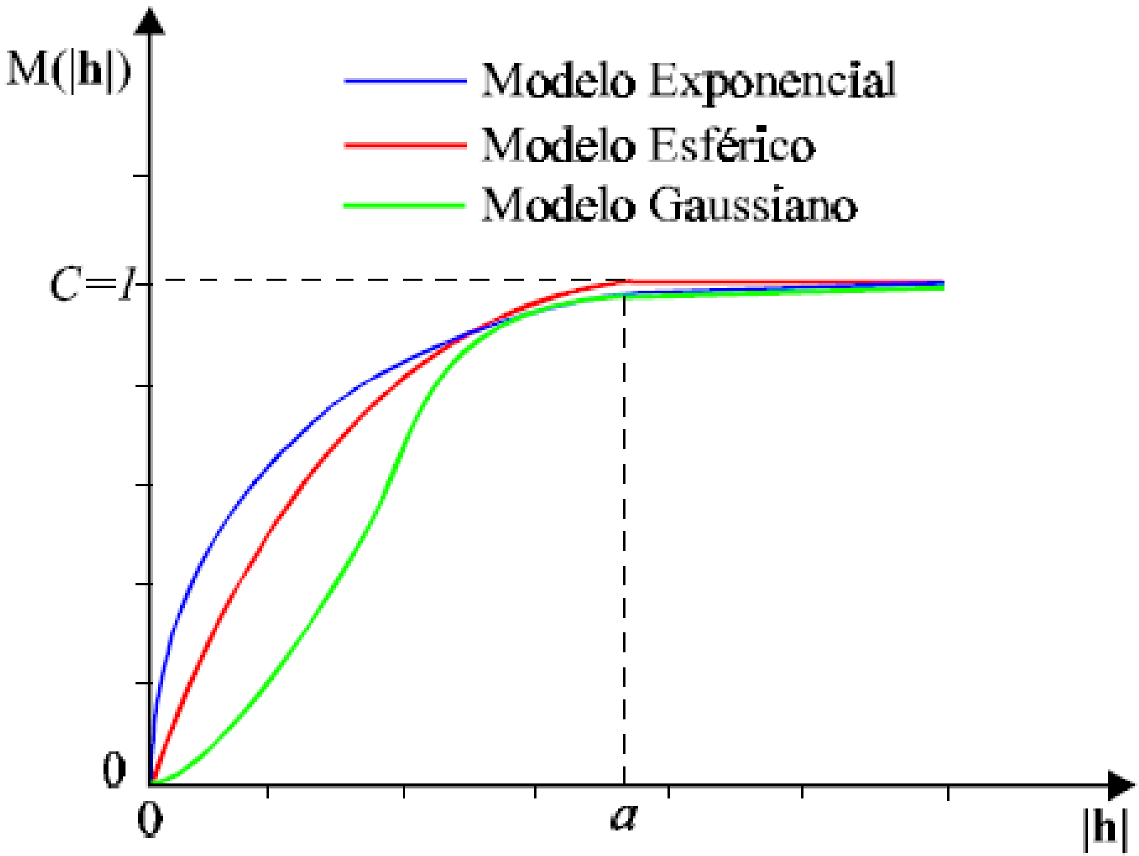


L.R. AMARAL, 2014

Ajuste de modelo a um semivariograma

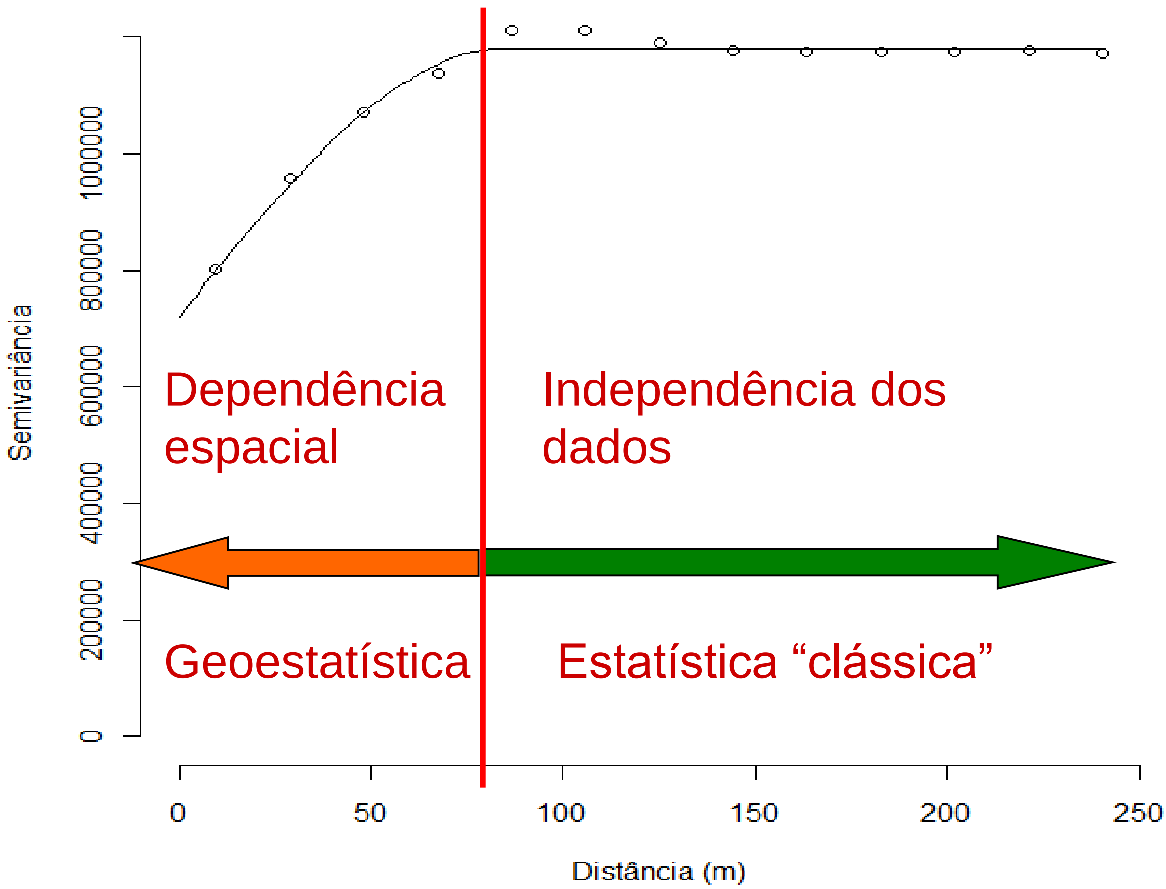
- Existem métodos para ajuste do modelo do semivariograma, dentre eles:
  - ajuste visual
  - minimização da soma dos erros quadrados
- Esse ajuste é uma etapa importante pois na interpolação por krigagem será necessário ter disponíveis valores de semivariância para qualquer distância dentro do limite do alcance.

Principais modelos de ajuste matemáticos ao semivariograma

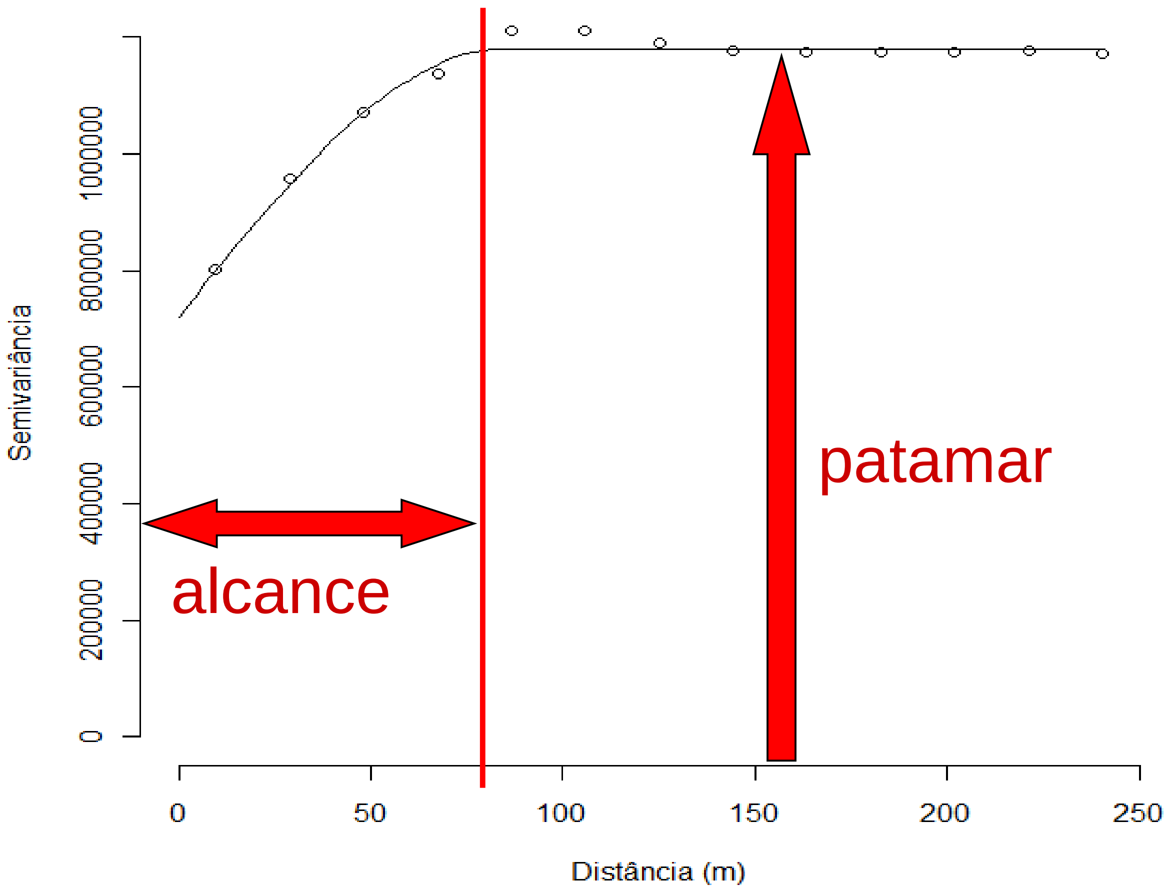


Camargo (1998)

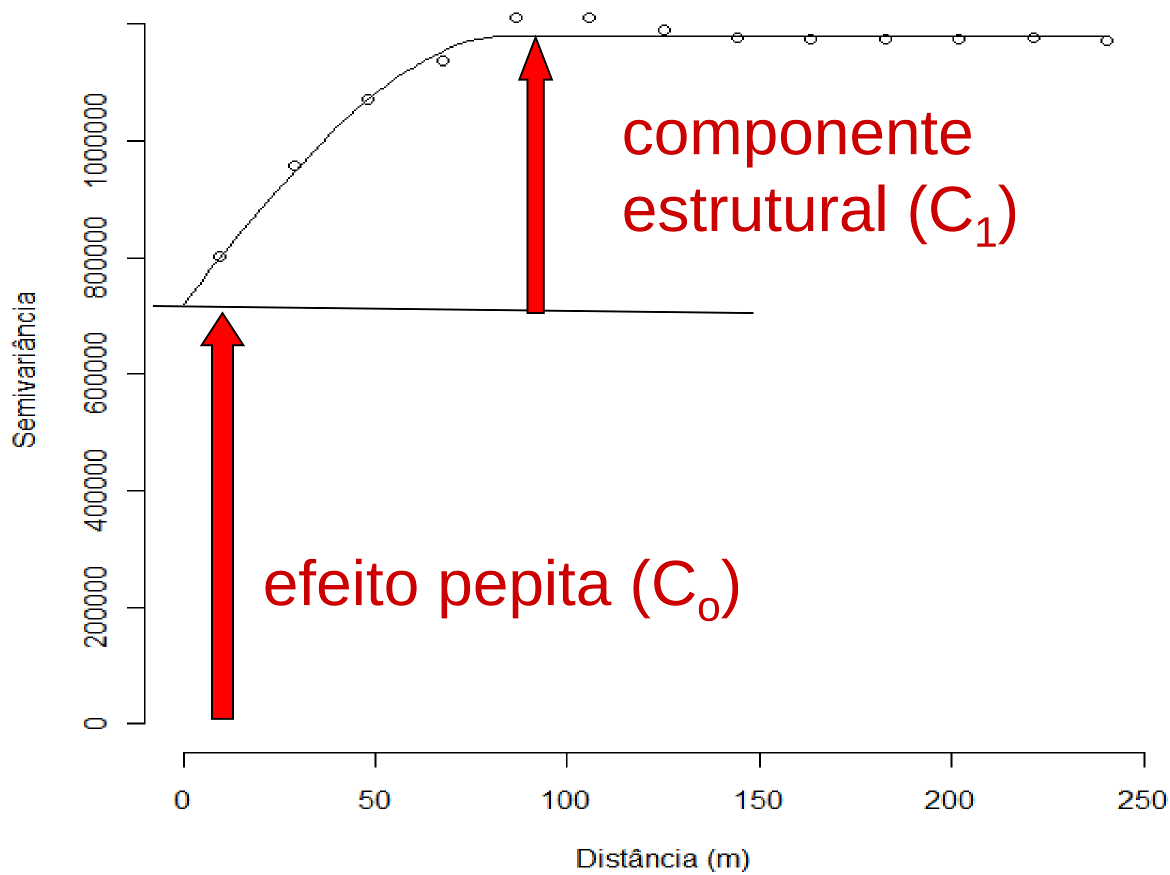
Componentes do semivariograma



Componentes do semivariograma



Componentes do semivariograma

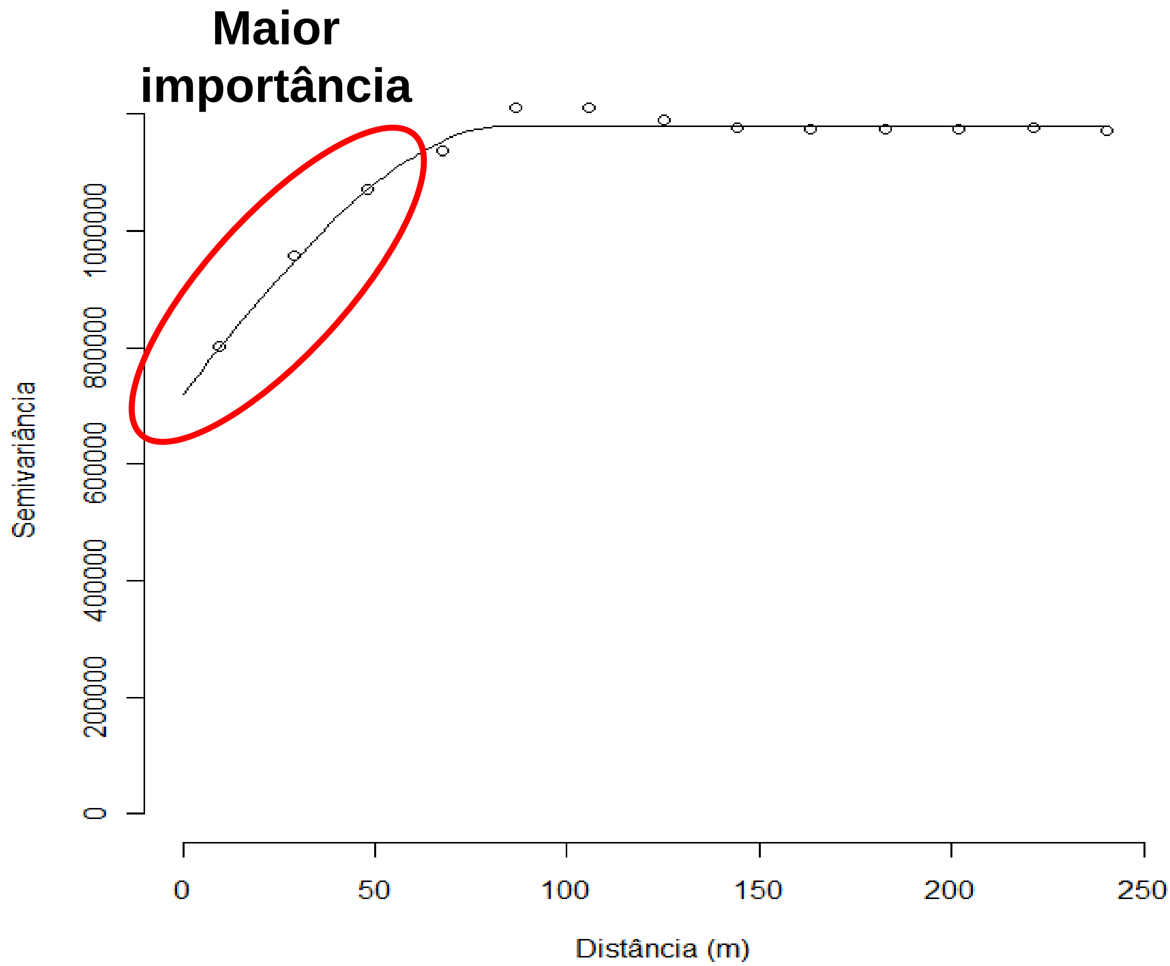


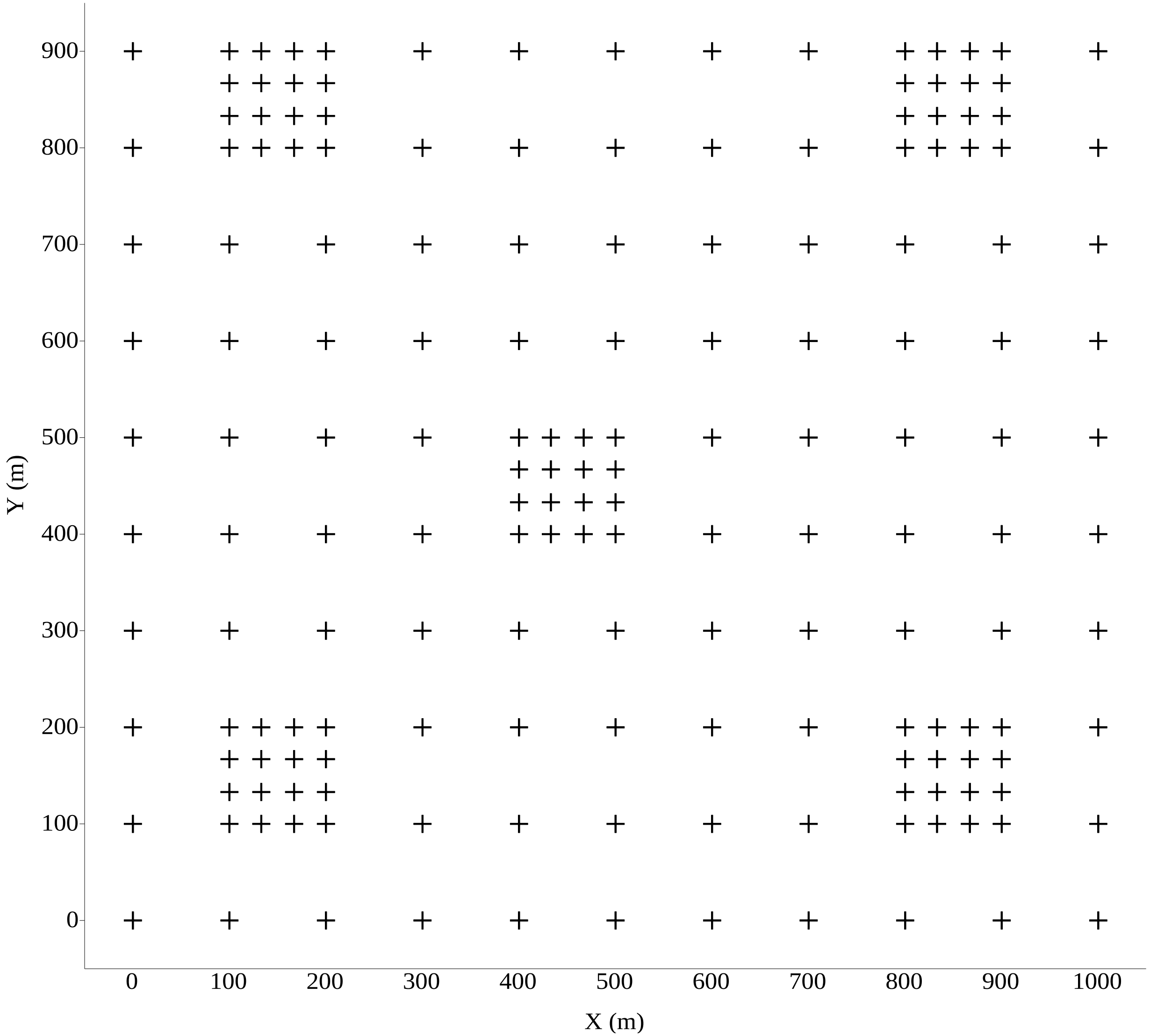
$C_o + C_1$ = patamar

componente estrutural – indicador da dependência especial - [C1/(C0+C1)] Cambardella et al. (1994)

acima de 0,75 – elevada
entre 0,25 e 0,75 – média
abaixo de 0,25 - baixa

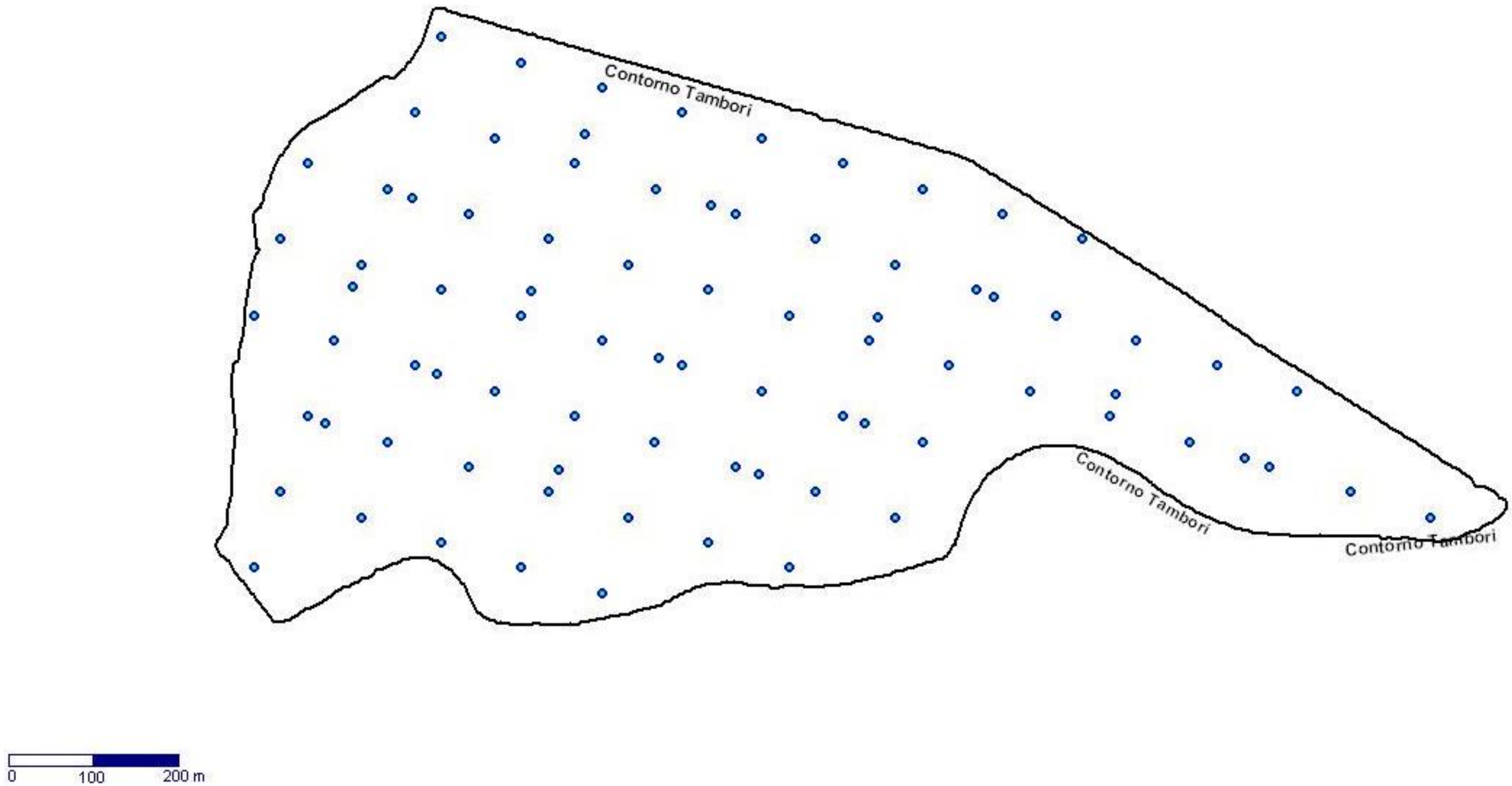
Ajuste de modelo a um semivariograma





Esquema amostral em grade regular acrescida de “ilhas” com espaçamento adensado

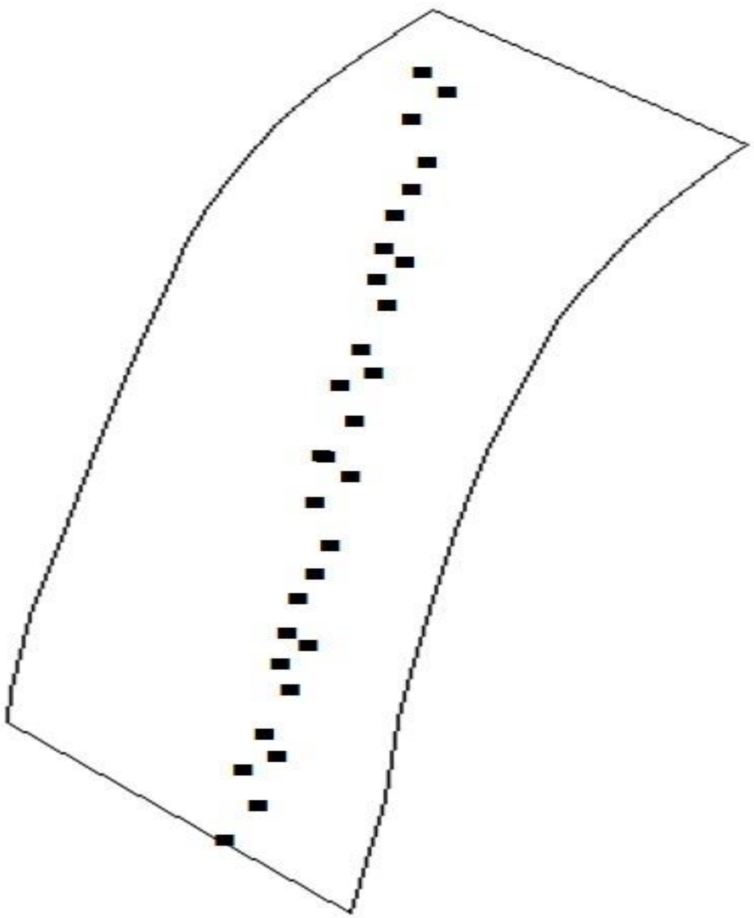
MOTOMIYA (2007)



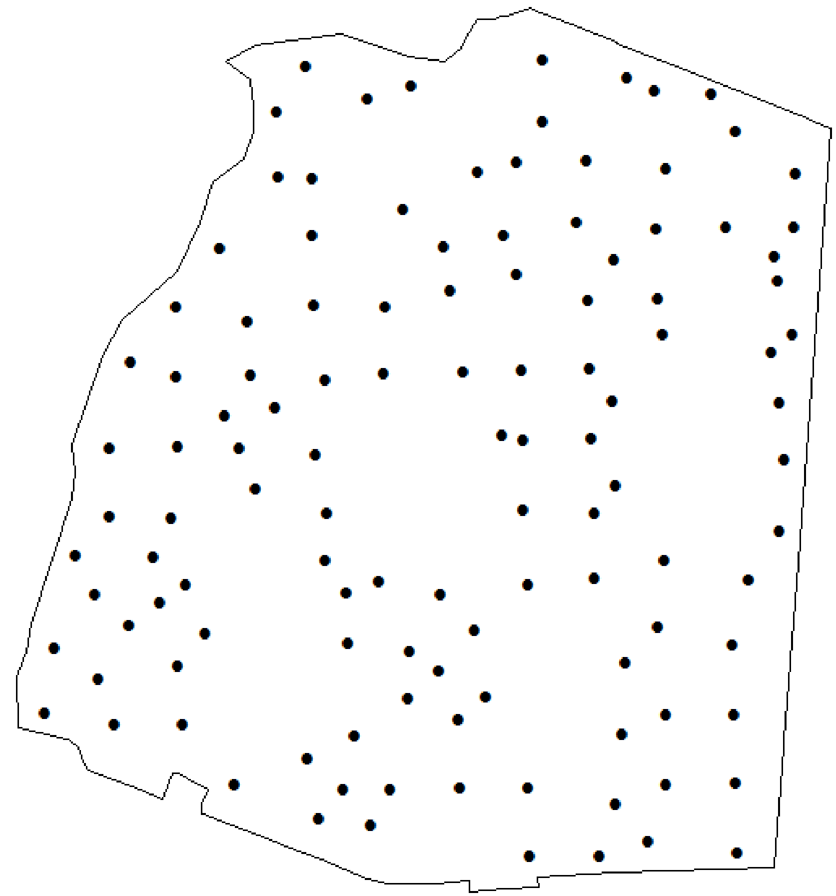
Esquema amostral em grade regular acrescida de pontos com espaçamento adensado

gMAP (2010)

Transecto

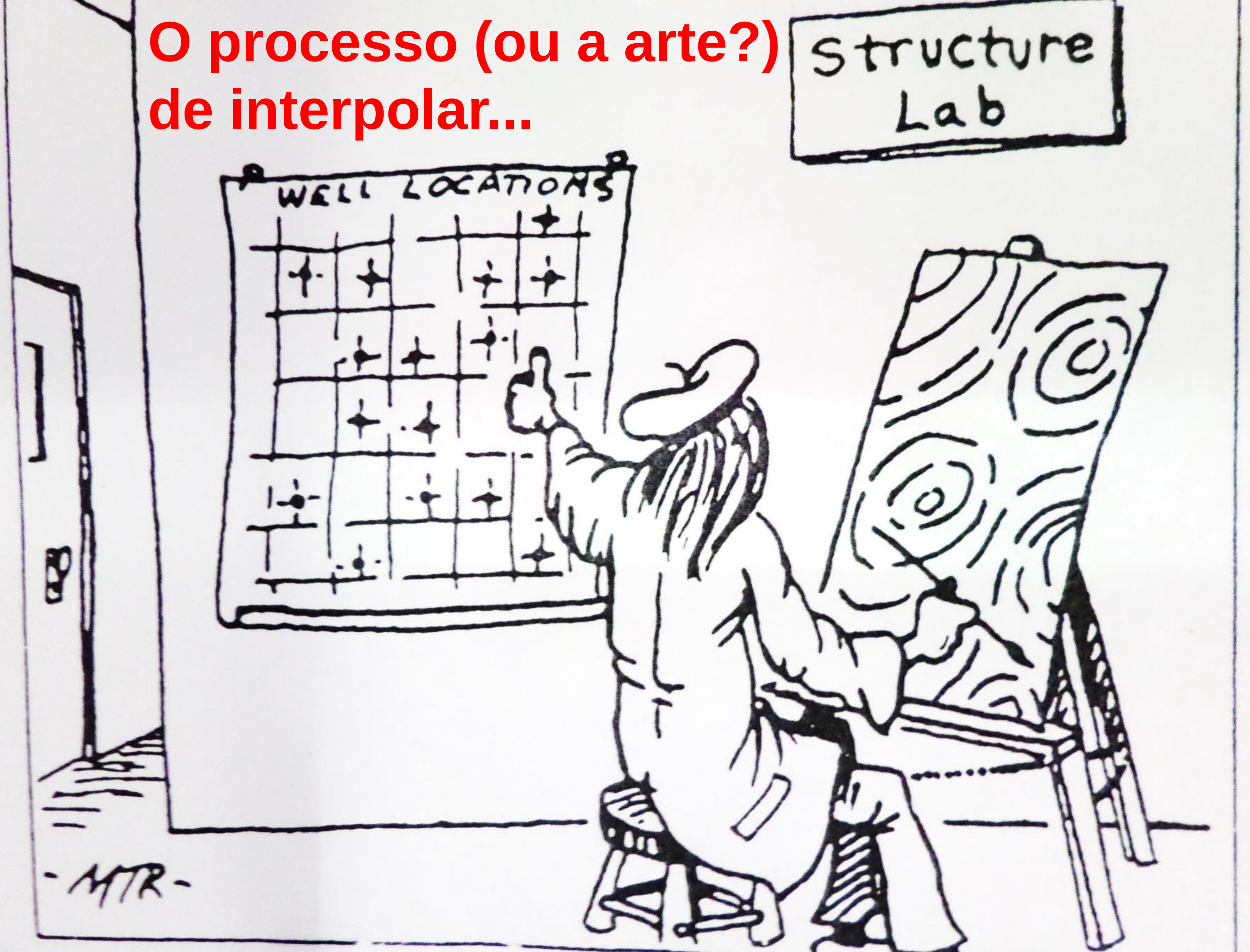


Aleatorização

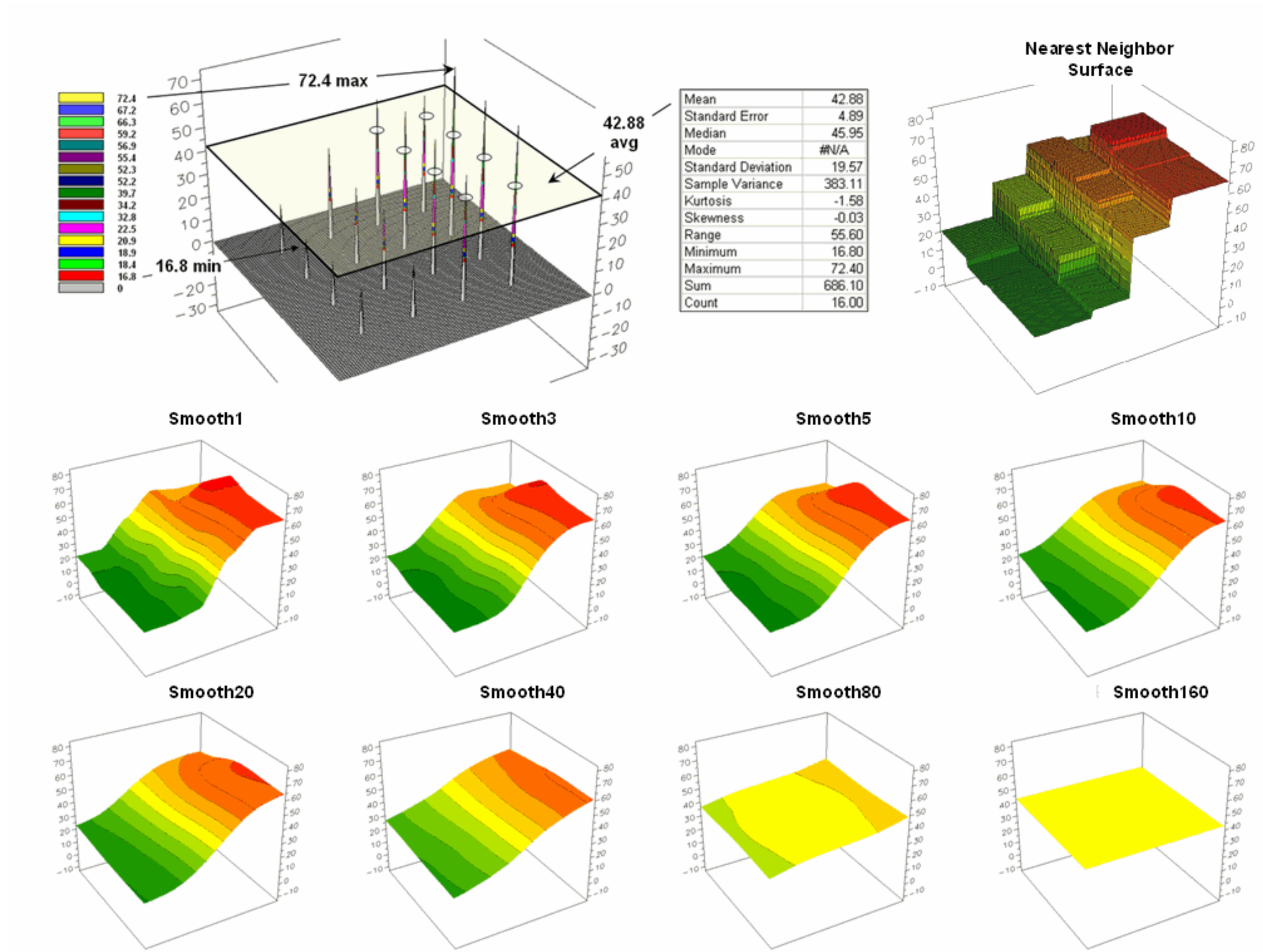


L.R. AMARAL, 2014

O processo (ou a arte?) de interpolar...



O processo de interpolação



Mean	42.88
Standard Error	4.89
Median	45.95
Mode	#N/A
Standard Deviation	19.57
Sample Variance	383.11
Kurtosis	-1.58
Skewness	-0.03
Range	55.60
Minimum	16.80
Maximum	72.40
Sum	686.10
Count	16.00

The spatial distribution implied by a set of discrete sample points can be estimated by iterative smoothing of the point values.

O processo de interpolação

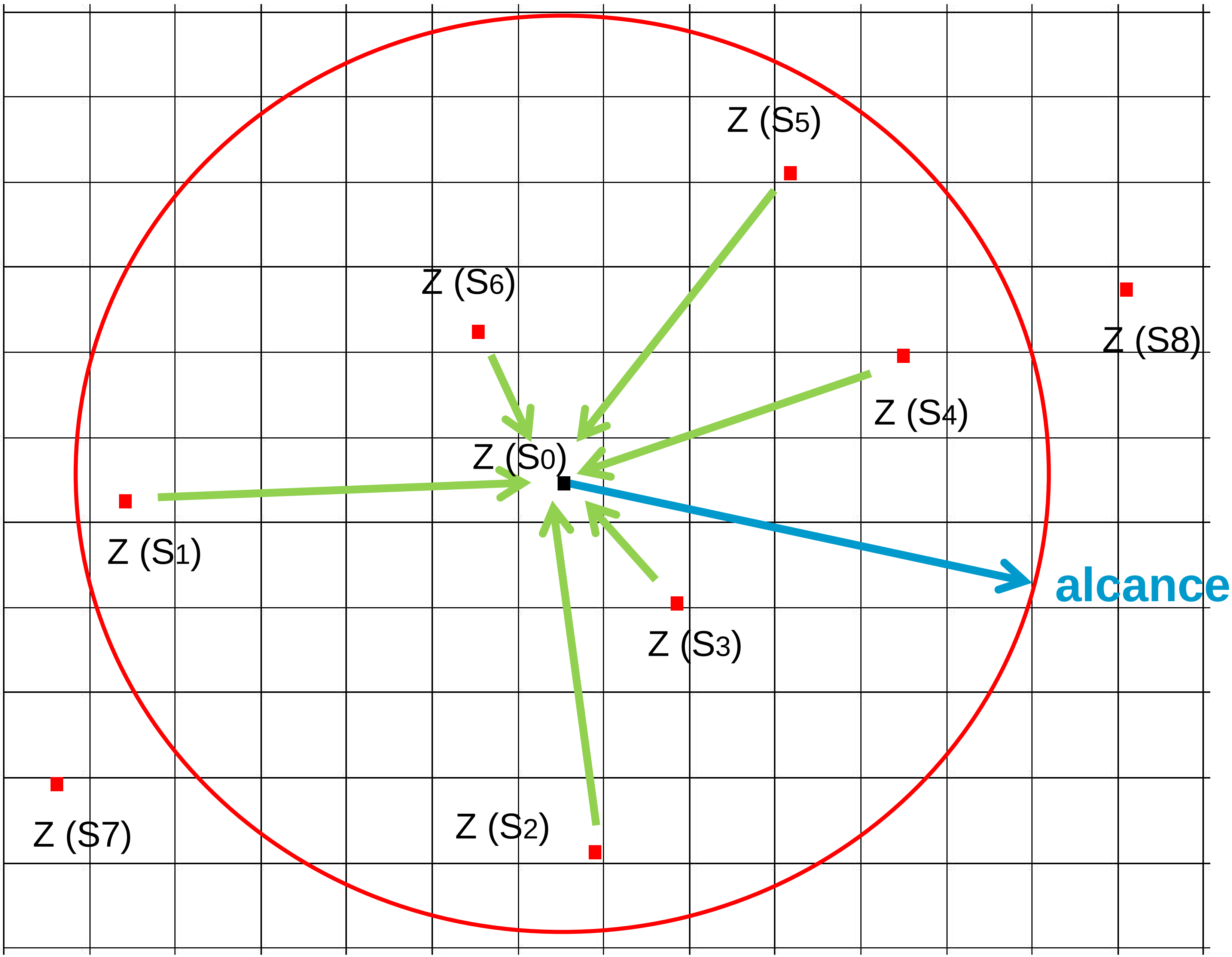
Métodos determinísticos - utilizam funções matemáticas para prever os valores (inverso da distância, triangulação, ...)

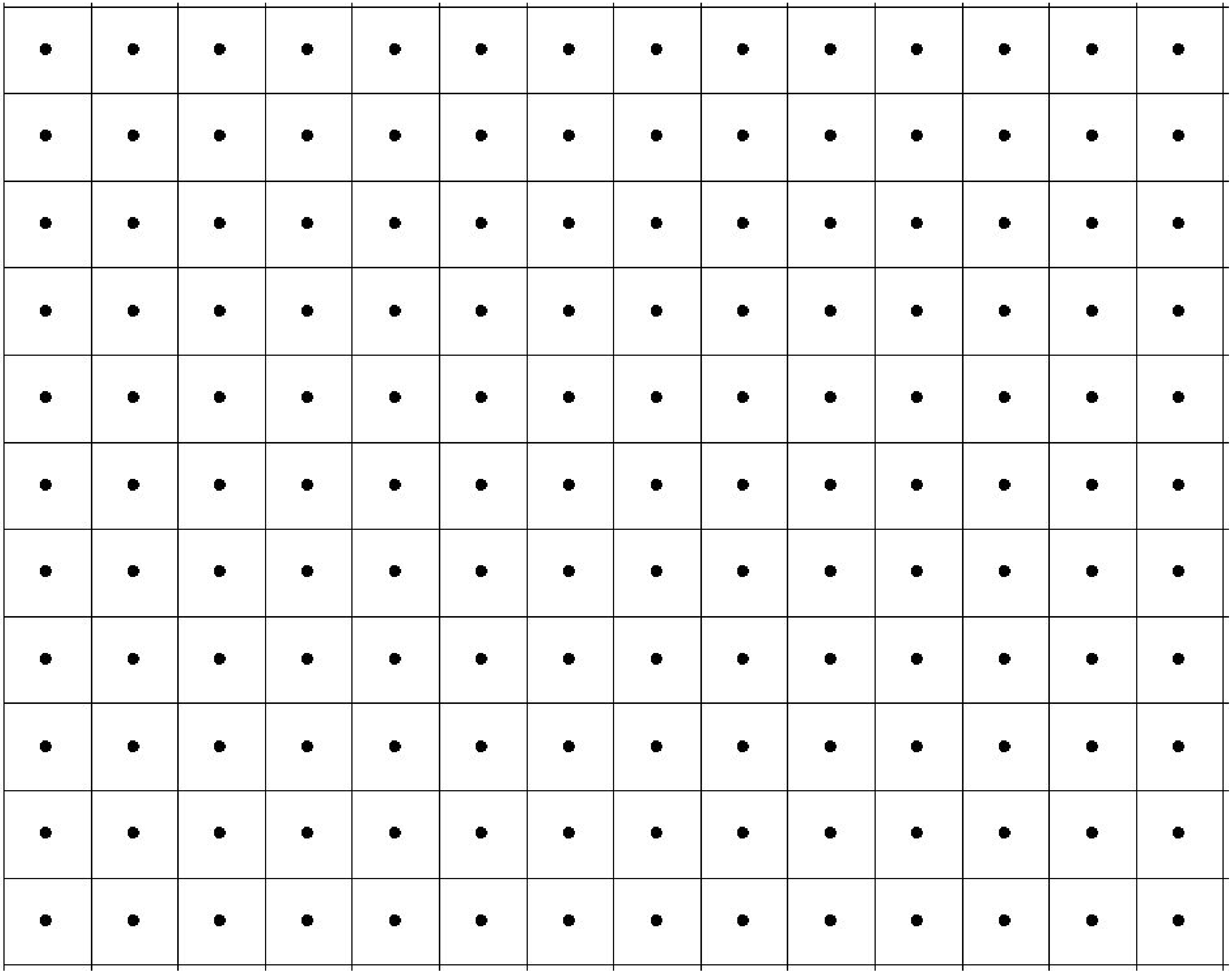
Métodos geoestatísticos – matemática e estatística para criar superfícies e avaliar a incerteza das previsões (krigagem, cokrigagem)

Johnston et al. (2001)

Krigagem

- É um estimador geoestatístico
- Leva em consideração a continuidade espacial da variável no processo de estimação da mesma.


Z (S5)
Z (S6)
Z (S8)
Z (S4)
Z (S0)
Z (S1)
alcance
Z (S3)
Z (S7)
Z (S2)



Krigagem

A krigagem resolve um sistema de matrizes do tipo

$$[\lambda]=[b]*[\gamma]^{-1}$$

para calcular o peso que será dado a cada amostra para a estimativa do ponto não amostrado

- $[\lambda]$ é a matriz com os valores dos pesos a serem aplicados em cada amostra
- [b] é a matriz com os valores da semivariância de cada amostra em relação ao ponto a ser estimado
- $[\gamma]^{-1}$ é a matriz que contêm o valor da semivariância entre as amostras disponíveis na área

Como saber se a interpolação foi acertada?

Validação cruzada - permite comparar valores preditos com os amostrados.

O valor da amostra, no local Z(si) é temporariamente descartado do conjunto de dados e é feita uma previsão por krigagem ($\hat{Z}(s_{(i)})$) no local usando-se as amostras restantes.

Assim é possível estimar o Erro Médio (EM):

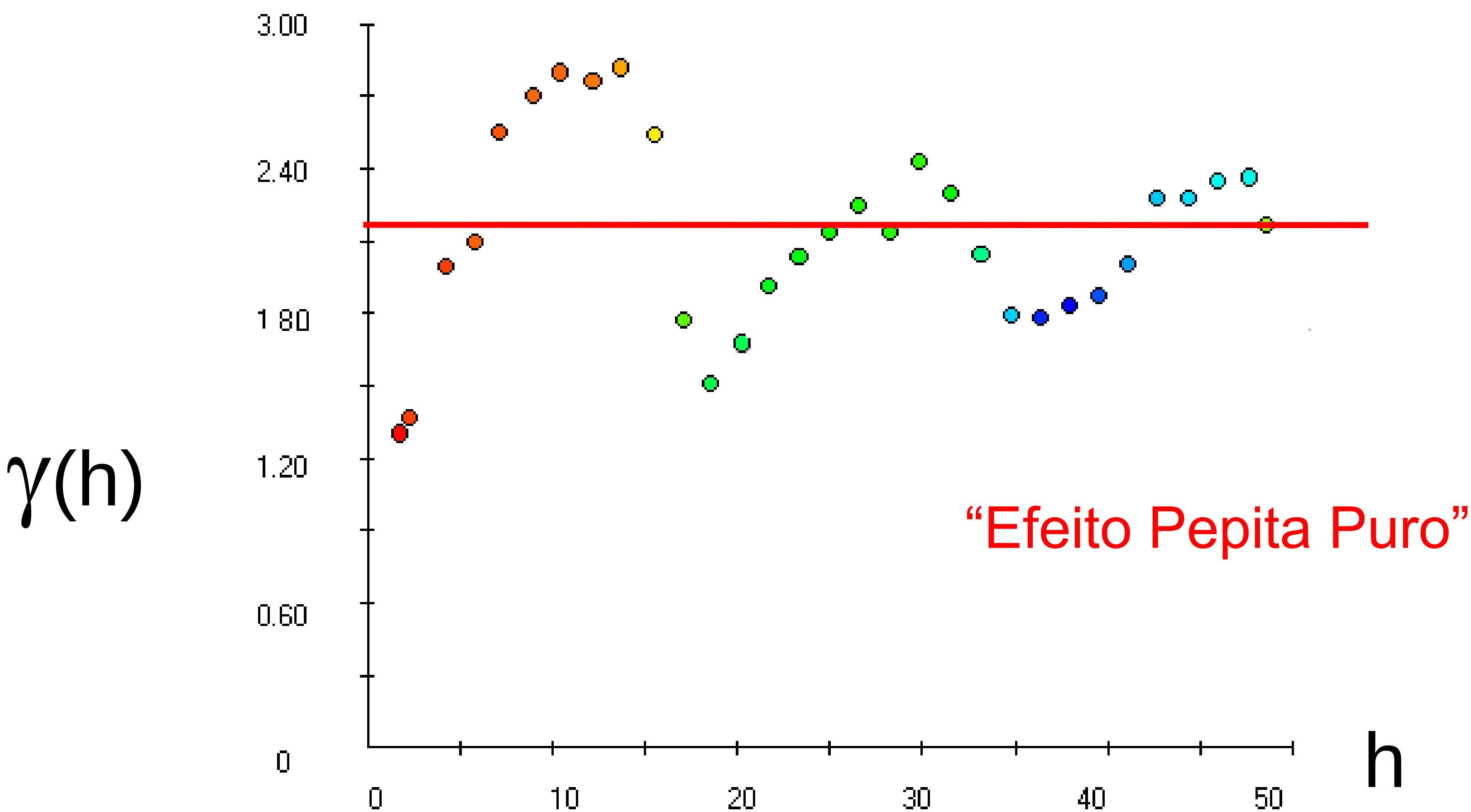
$$EM = \frac{1}{n}\sum_{i=1}^{1}\left(Z(s_i) - \hat{Z}(s_{(i)})\right)$$

n é o número de dados

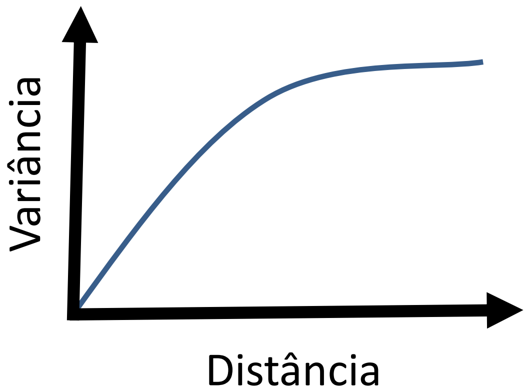
$Z(s_i)$, valor observado no ponto s_i

$\hat{Z}\,(s_{(i)})$, valor predito por krigagem ordinária no ponto s_i, sem considerar a observação $Z(s_i)$

E quando não é possível obter os parâmetros do semivariograma?

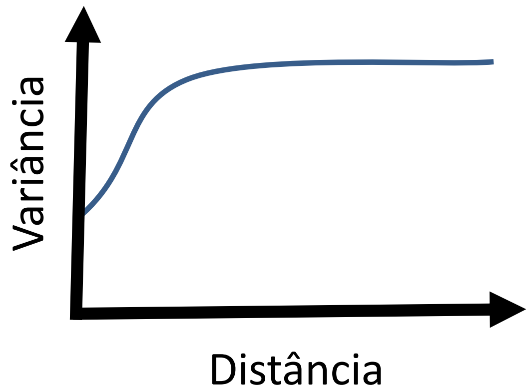


Argila



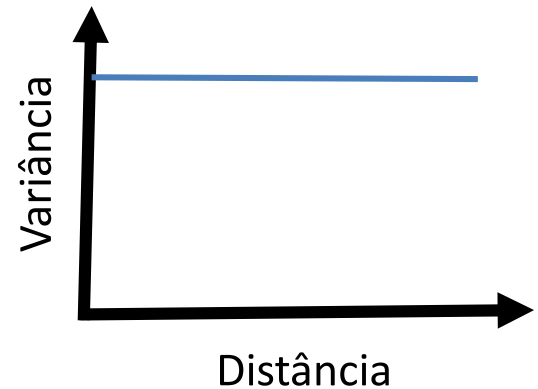
Alta dependência espacial

pH



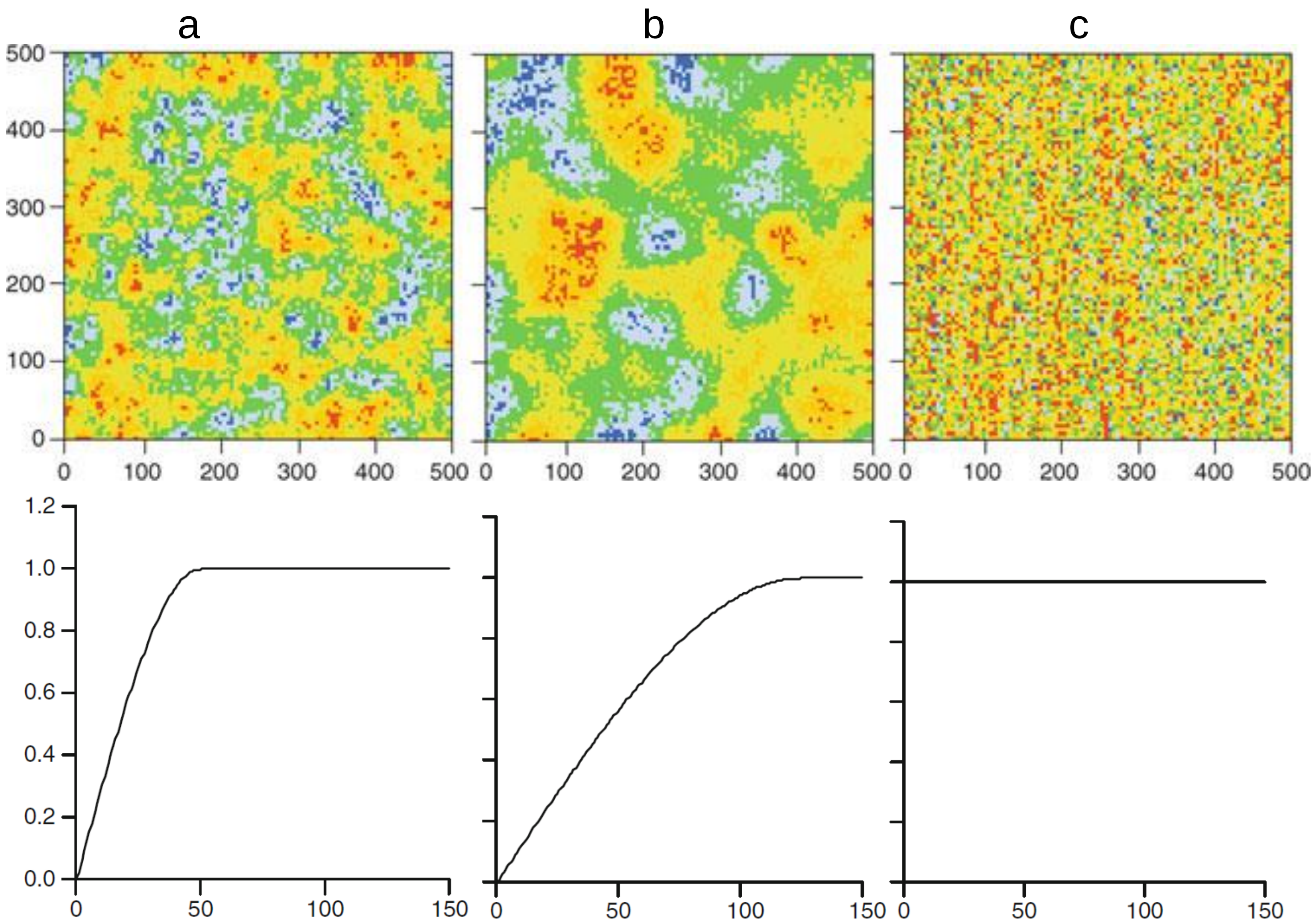
Dependência espacial
\+
Efeito pepita

Fósforo



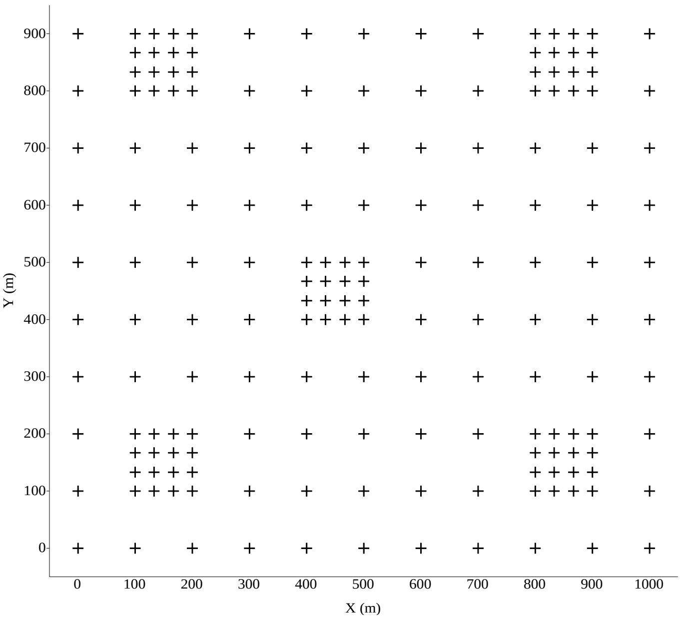
Efeito pepita puro

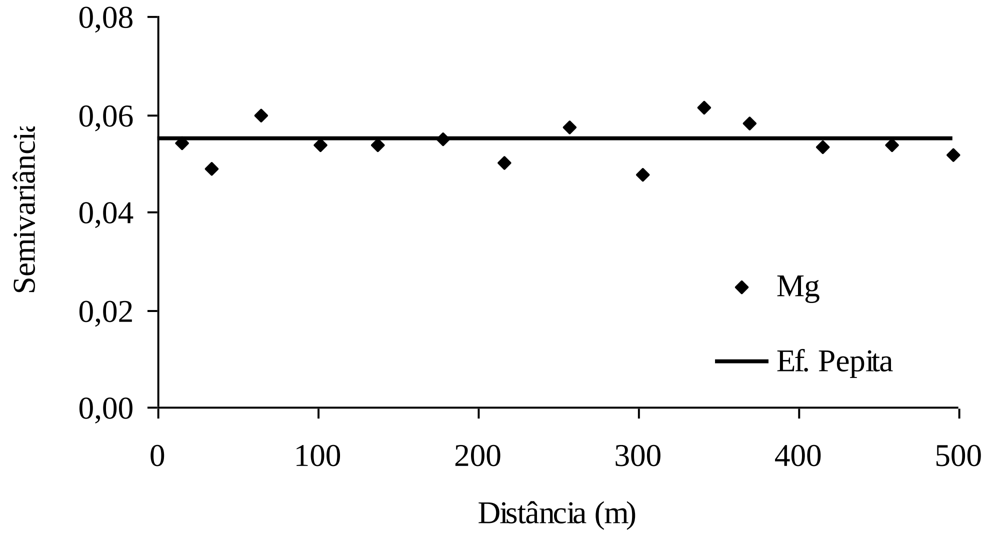
Eitelwein, 2017

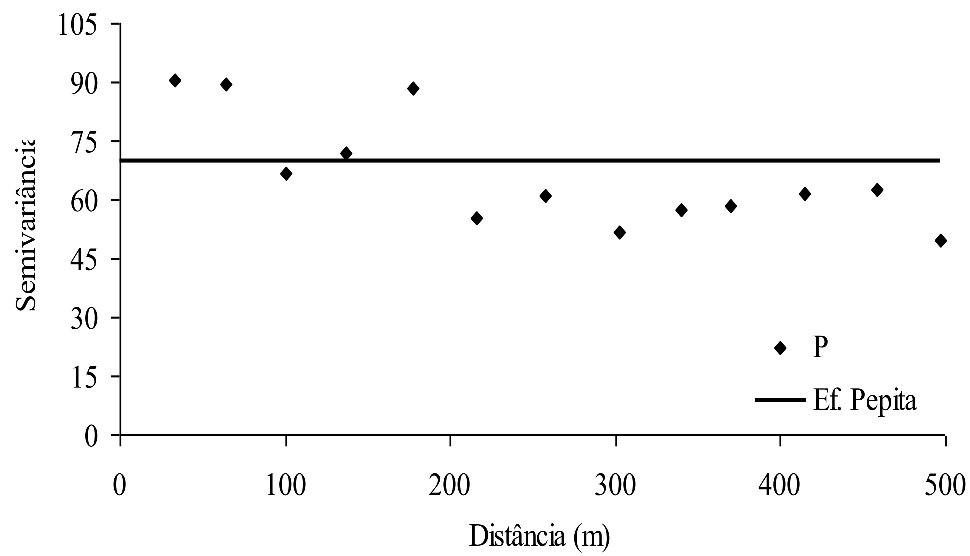


Distância entre amostras menor do que “manchas”

Talhões simulados com funções esféricas, com efeito pepita zero, patamar 1 e alcances de: (**a**) 50 m e (**b**) 125m; (**c**) variograma com efeito pepita puro (Kerry et al., 2010)







MOTOMIYA (2007)