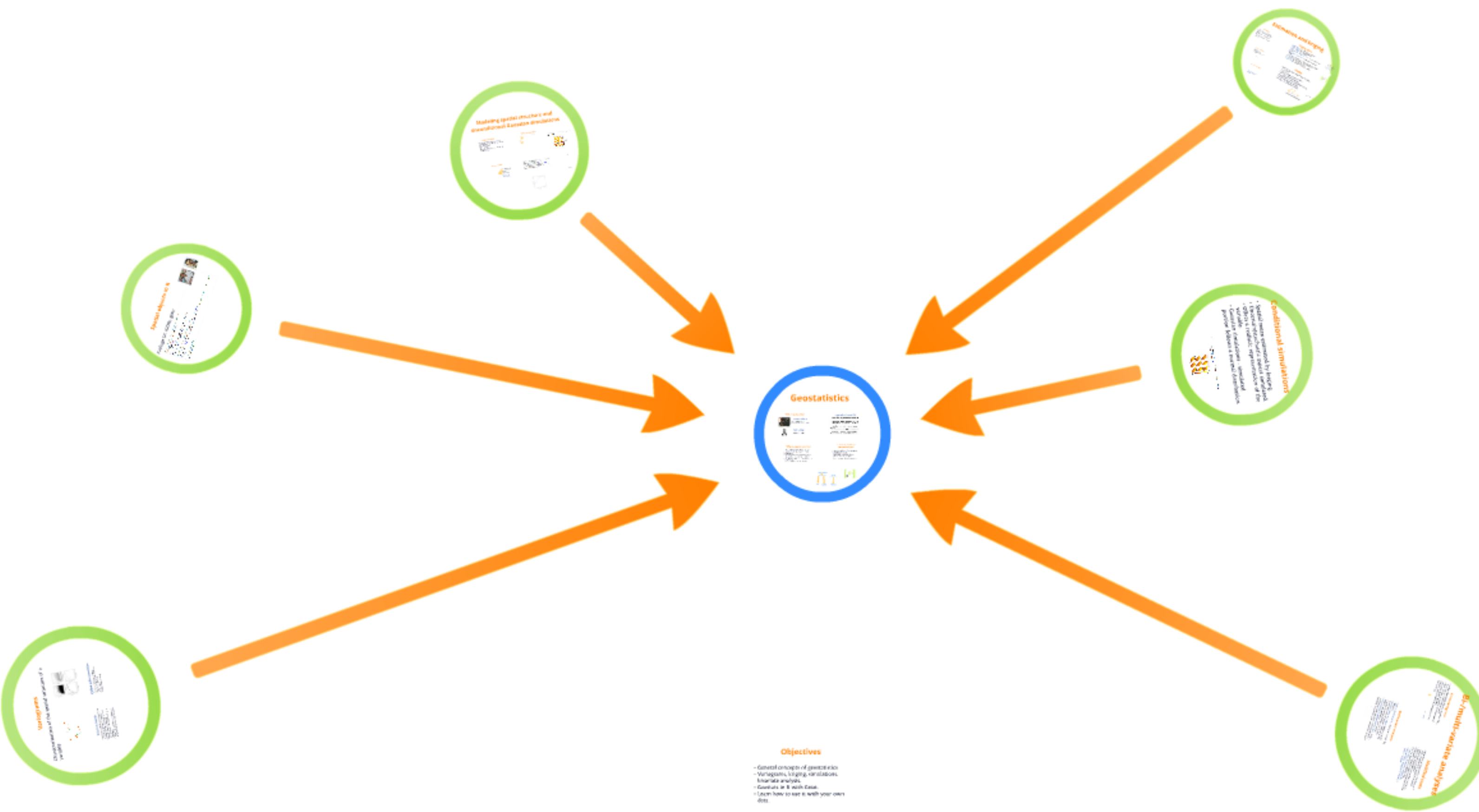
Introduction to geostatistics with R

Guillaume Larocque, research professional

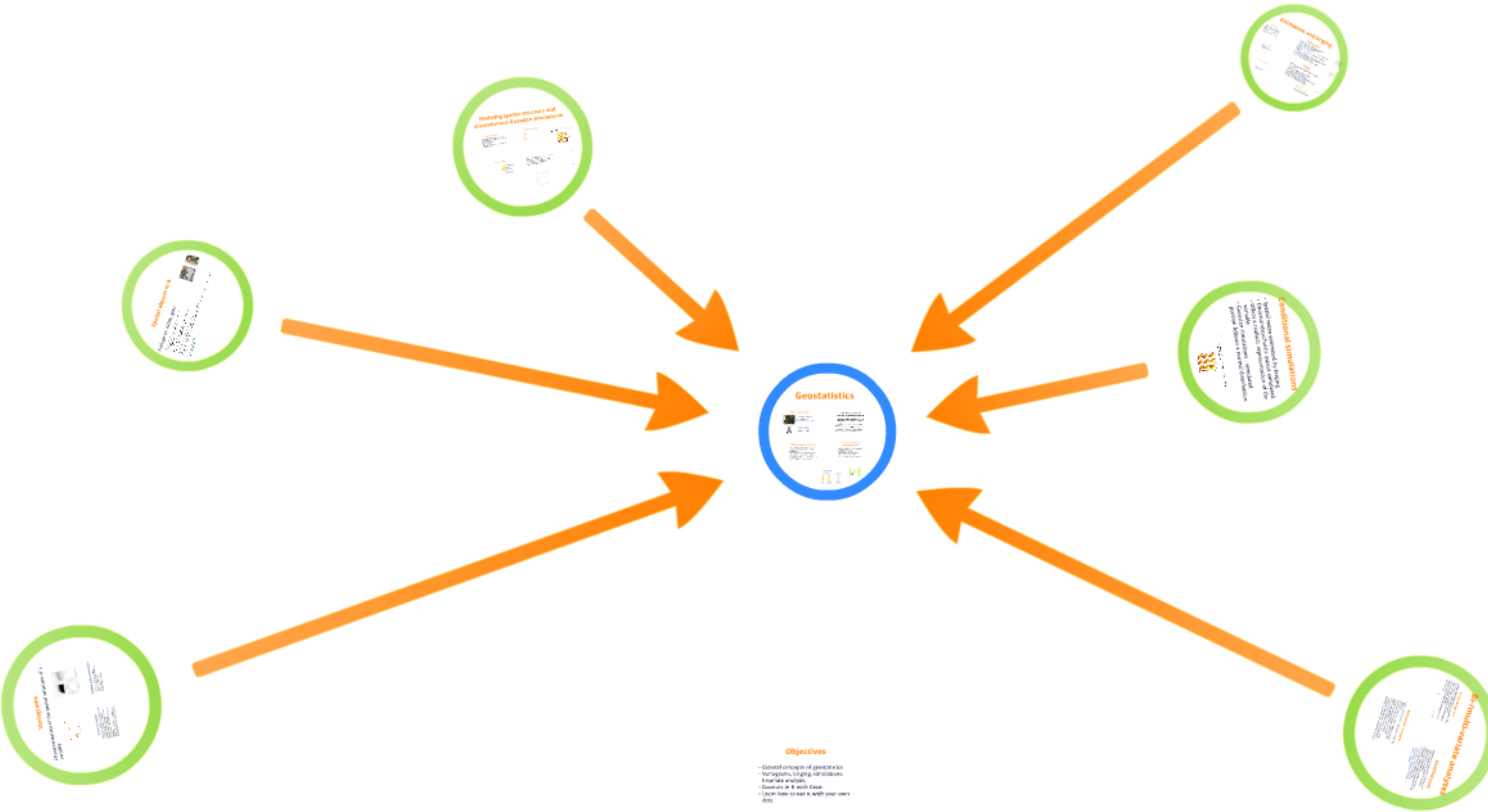


http://qcbs.ca/wiki/geostatistics

CSBQ QCBS
Quebec Centre for Biodiversity Science

Introduction to geostatistics with R

Guillaume Larocque, research professional



http://qcbs.ca/wiki/geostatistics

CSBQ QCBS

Quebec Centre for Biodiversity Science

Objectives

- General concepts of geostatistics
- Variograms, kriging, simulations, bivariate analysis.
- Geostats in R with Gstat.
- Learn how to use it with your own data.

Geostatistics

What are geostatistics?

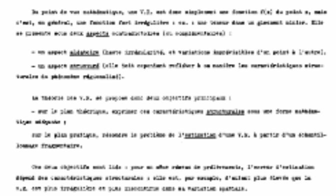
Georges Matheron

"La théorie des variables régionalisées et ses applications"

Danie G. Krige

The father of kriging

Regionalized variables

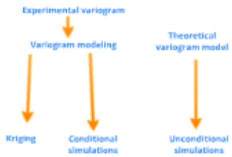


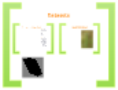
Why use geostatistics?

- Characterization of the spatial structure.
- Spatial relationships between variables.
- Interpolation
- Estimate values at unsampled locaitons (e.g. dense grids, change of support problem).
- Visualization, creation of continuous rasters for GIS analysis.
- Get an idea of the data uncertainty.

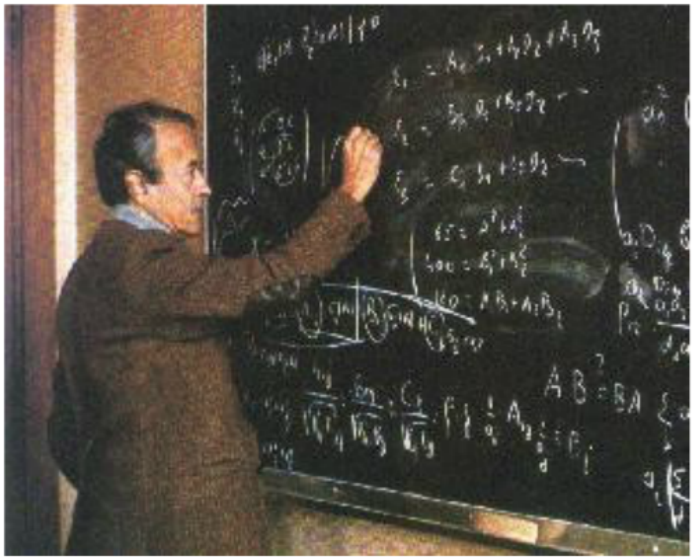
What do we need to use geostatistical tools?

- Spatial sampling. Grid or transect.
- Enough point (>40?).
- Appropriate sampling scheme.
- Covers the extent of interest.
- Representation of small distances.





What are geostatistics?



Georges Matheron

"La théorie des variables régionalisées et ses applications"

Danie G. Krige

The father of kriging

Regionalized variables

Du point de vue mathématique, une V.R. est donc simplement une fonction f(x) du point x, mais c'est, en général, une fonction fort irrégulière : ex. : une teneur dans un gisement minier. Elle se présente sous deux aspects contradictoires (ou complémentaires) :

- un aspect aléatoire (haute irrégularité, et variations imprévisibles d'un point à l'autre).

- un aspect structuré (elle doit cependant refléter à sa manière les caractéristiques structurales du phénomène régionalisé).

La théorie des V.R. se propose donc deux objectifs principaux :

- sur le plan théorique, exprimer ces caractéristiques structurales sous une forme mathématique adéquate ;

sur le plan pratique, résoudre le problème de l'estimation d'une V.R. à partir d'un échantillonnage fragmentaire.

Ces deux objectifs sont liés : pour un même réseau de prélèvements, l'erreur d'estimation dépend des caractéristiques structurales ; elle est, par exemple, d'autant plus élevée que la V.R. est plus irrégulière et plus discontinue dans sa variation spatiale.

Regionalized variables

Du point de vue mathématique, une V.R. est donc simplement une fonction f(x) du point x, mais c'est, en général, une fonction fort irrégulière : ex. : une teneur dans un gisement minier. Elle se présente sous deux aspects contradictoires (ou complémentaires) :

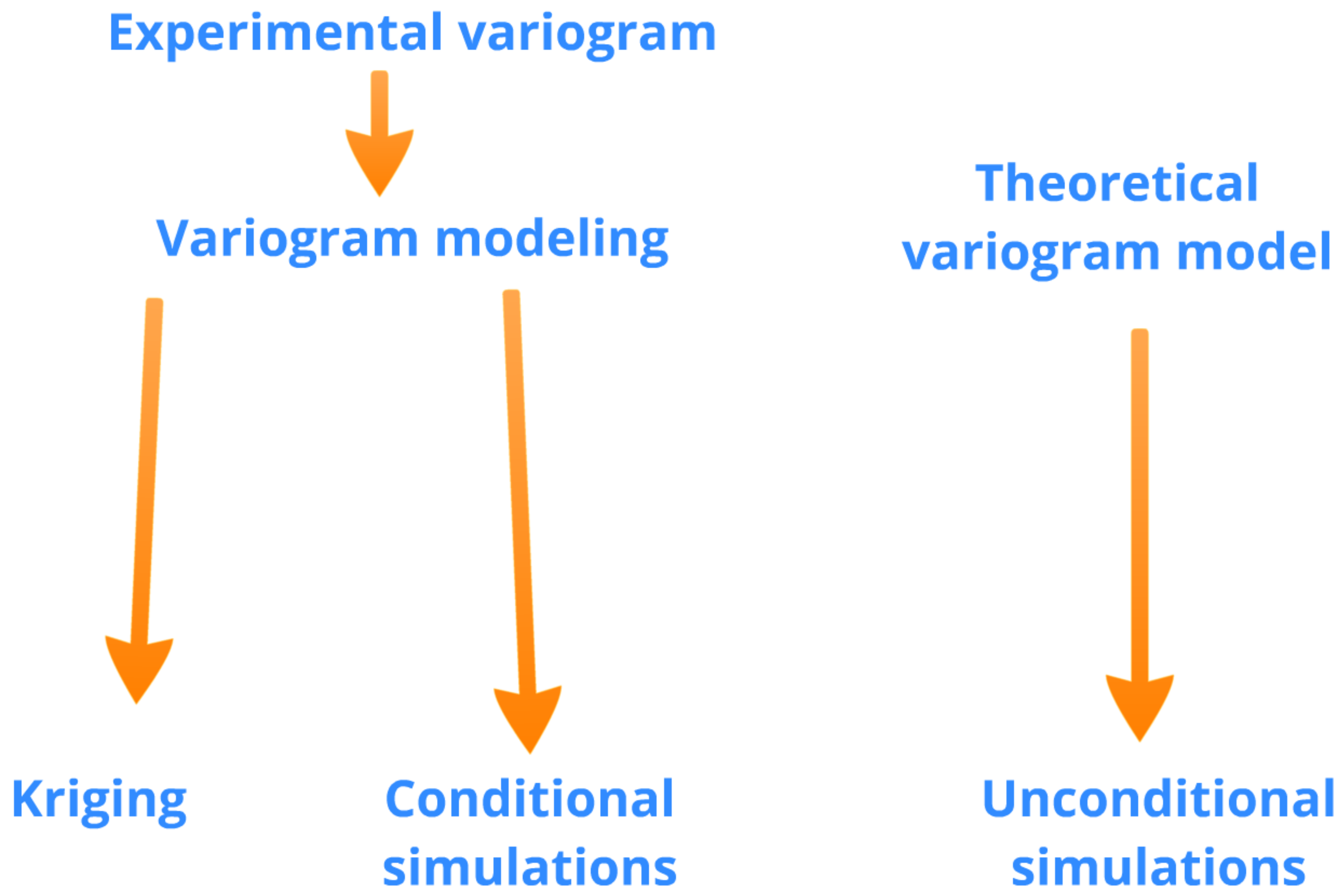
- un aspect aléatoire (haute irrégularité, et variations imprévisibles d'un point à l'autre).
- un aspect structuré (elle doit cependant refléter à sa manière les caractéristiques structurales du phénomène régionalisé).

La théorie des V.R. se propose donc deux objectifs principaux :

- sur le plan théorique, exprimer ces caractéristiques structurales sous une forme mathématique adéquate ;

sur le plan pratique, résoudre le problème de l'estimation d'une V.R. à partir d'un échantillonnage fragmentaire.

Ces deux objectifs sont liés : pour un même réseau de prélèvements, l'erreur d'estimation dépend des caractéristiques structurales ; elle est, par exemple, d'autant plus élevée que la V.R. est plus irrégulière et plus discontinue dans sa variation spatiale.

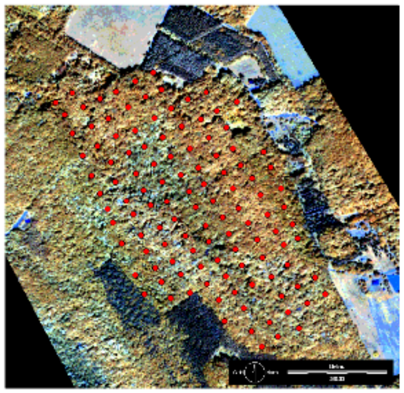

Experimental variogram
Variogram modeling
Theoretical variogram model
Kriging
Conditional simulations
Unconditional simulations

What do we need to use geostatistical tools?

- Spatial sampling. Grid or transect.
- Enough point (›40?).
- Appropriate sampling scheme.
- Covers the extent of interest.
- Representation of small distances.

Datasets

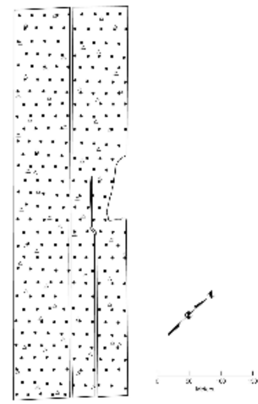
Sampling grid in the Morgan Arboretum



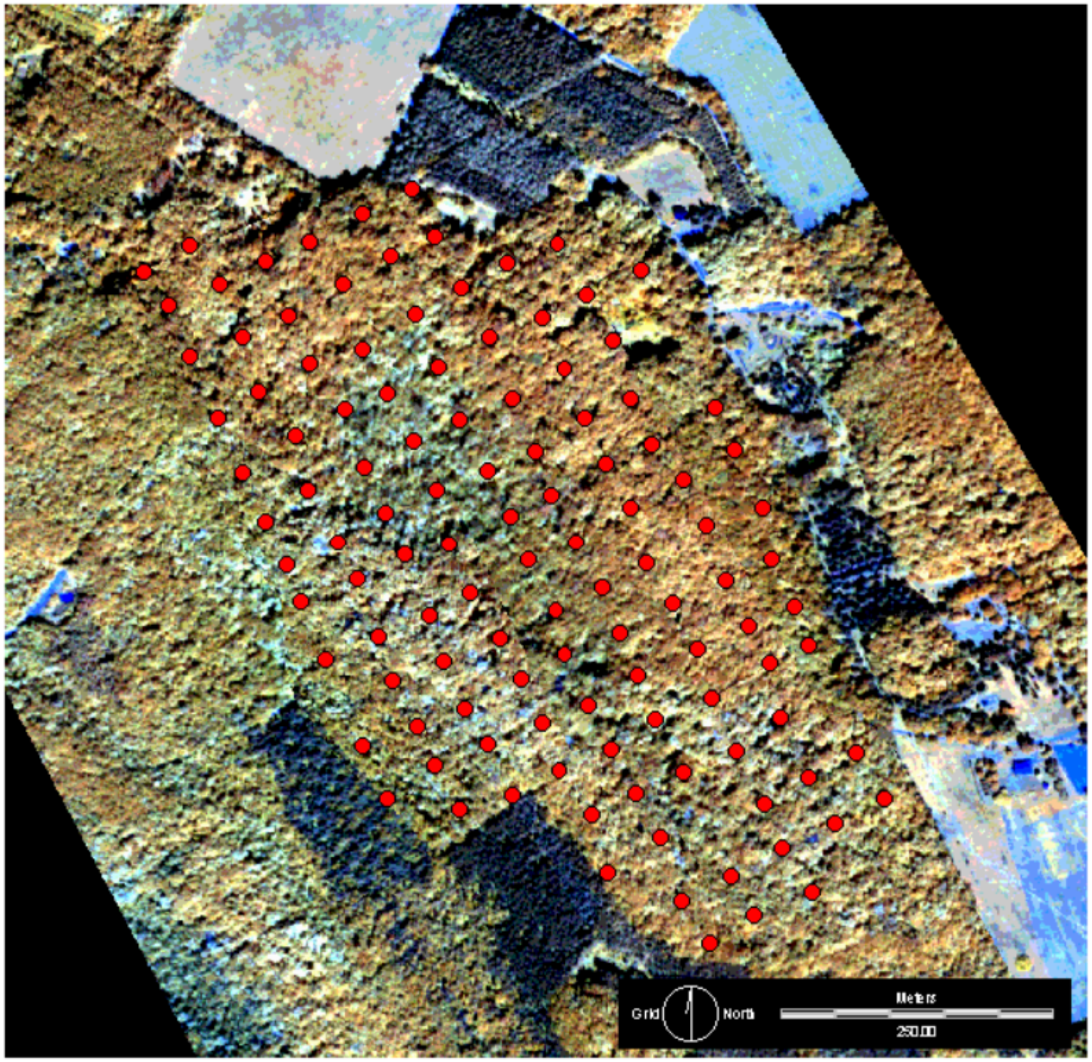
Nutrients
Total C 0-15 cm
Total C 15-30 cm
Total C Forest floor
K
P
Ca
NO3
NH4

Trees
Acer rubrum
Acer saccharum
Acer saccharinum
Fagus grandifolia
Quercus rubra'
Bet. Alleghaniensis
Tilia americana
Carya cordiformis
Carya ovata
Fraxinus nigra
Fraxinus Americana

Agricultural dataset



Sampling grid in the Morgan Arboretum

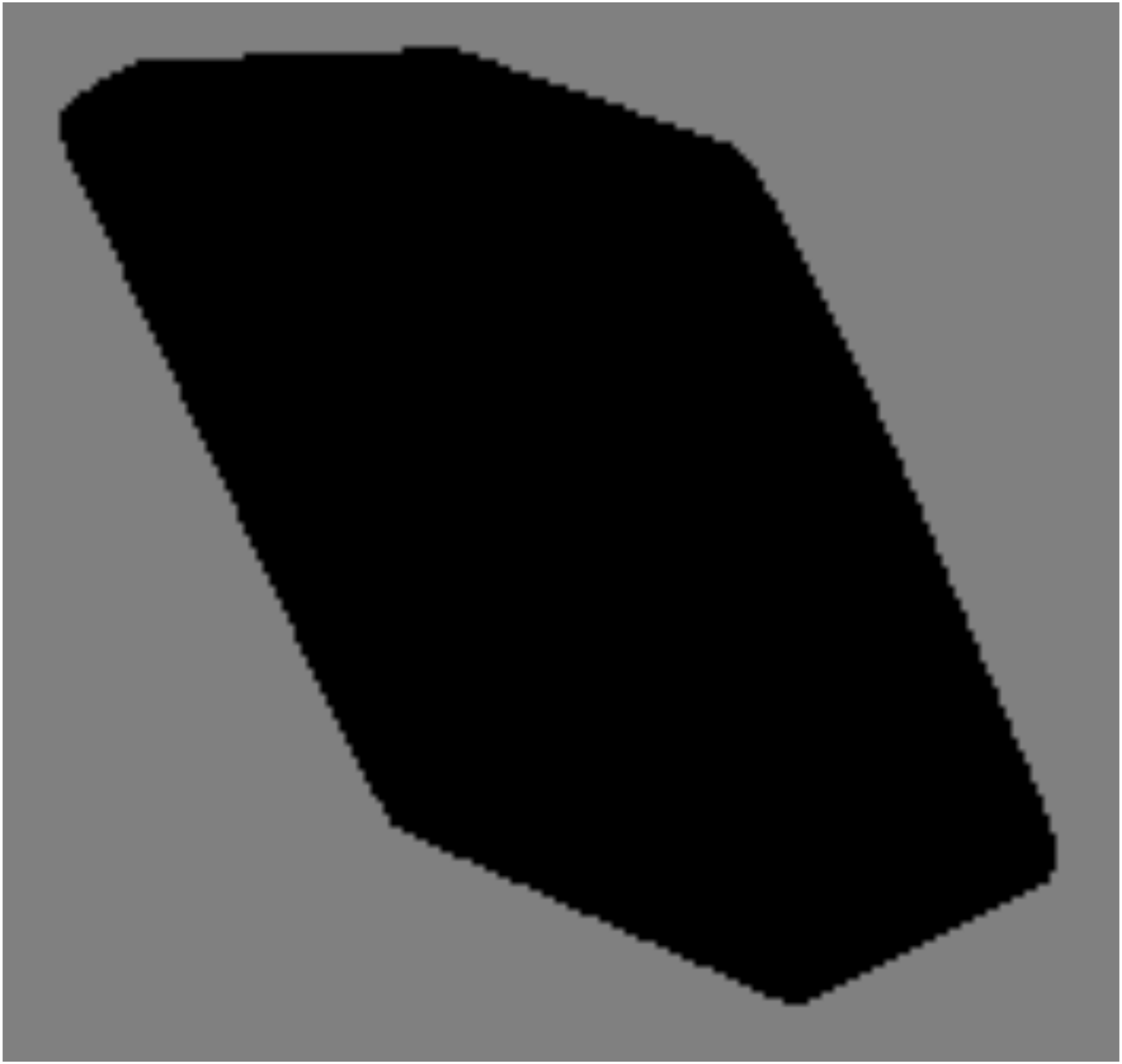


Nutrients

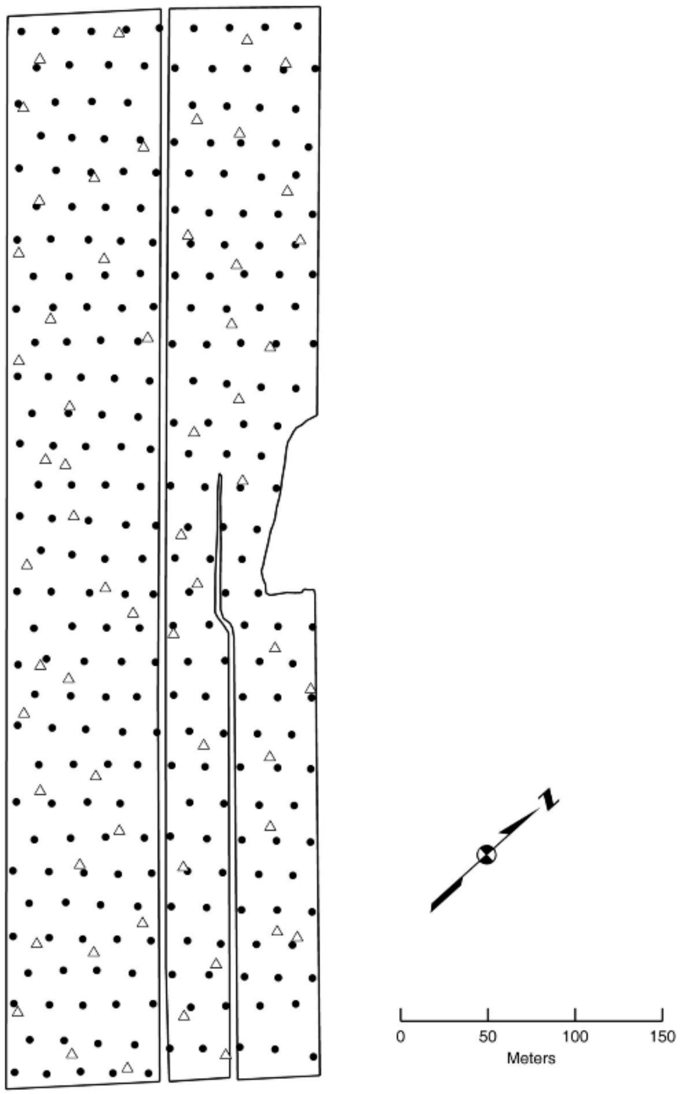
Total C 0-15 cm
Total C 15-30 cm
Total C Forest floor
K
P
Ca
NO3
NH4

Trees

Acer rubrum
Acer saccharum
Acer saccharinum
Fagus grandifolia
Quercus rubra'
Bet. Alleghaniensis
Tilia americana
Carya cordiformis
Carya ovata
Fraxinus nigra
Fraxinus Americana



Agricultural dataset

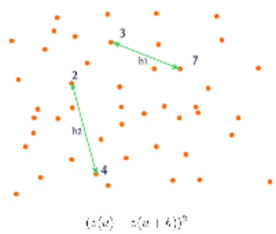


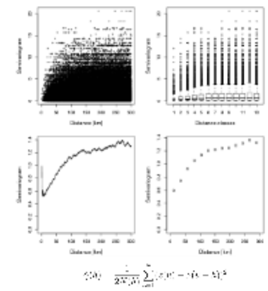
Why use geostatistics?

- Characterization of the spatial structure.
- Spatial relationships between variables.
- Interpolation
- Estimate values at unsampled locaitons (e.g. dense grids, change of support problem).
- Visualization, creation of continuous rasters for GIS analysis.
- Get an idea of the data uncertainty.

Variograms

Characterization of the spatial structure of a variable





Rules of thumb

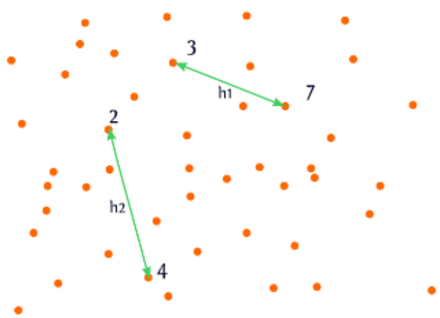
- Maximum width of variogram: 1/2 side of sampling grid or 1/2 sqrt(area).
- Number of lags: make sure to have enough points at each lag.
- Representation of small distances: a challenge.
- Anisotropy - different variogram in different directions. Variogram map.

Spatial autocorrelation

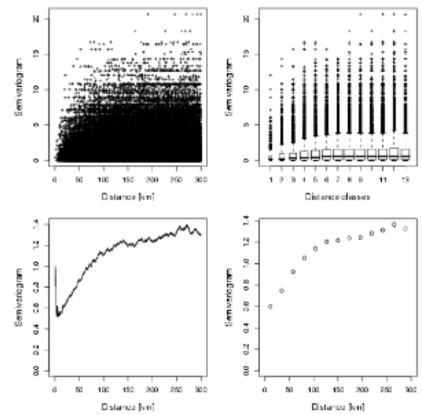
- Geary's C - linked to the variogram
- Moran's I - presence of autocorrelation "globally" or at each lag.
- Package 'spdep'
- Custom functions provided

Variograms

Characterization of the spatial structure of a variable



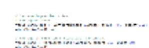
$$(z(u) - z(u+h))^2$$



$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N} (z(u) - z(u+h))^2$$

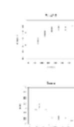
Rules of thumb

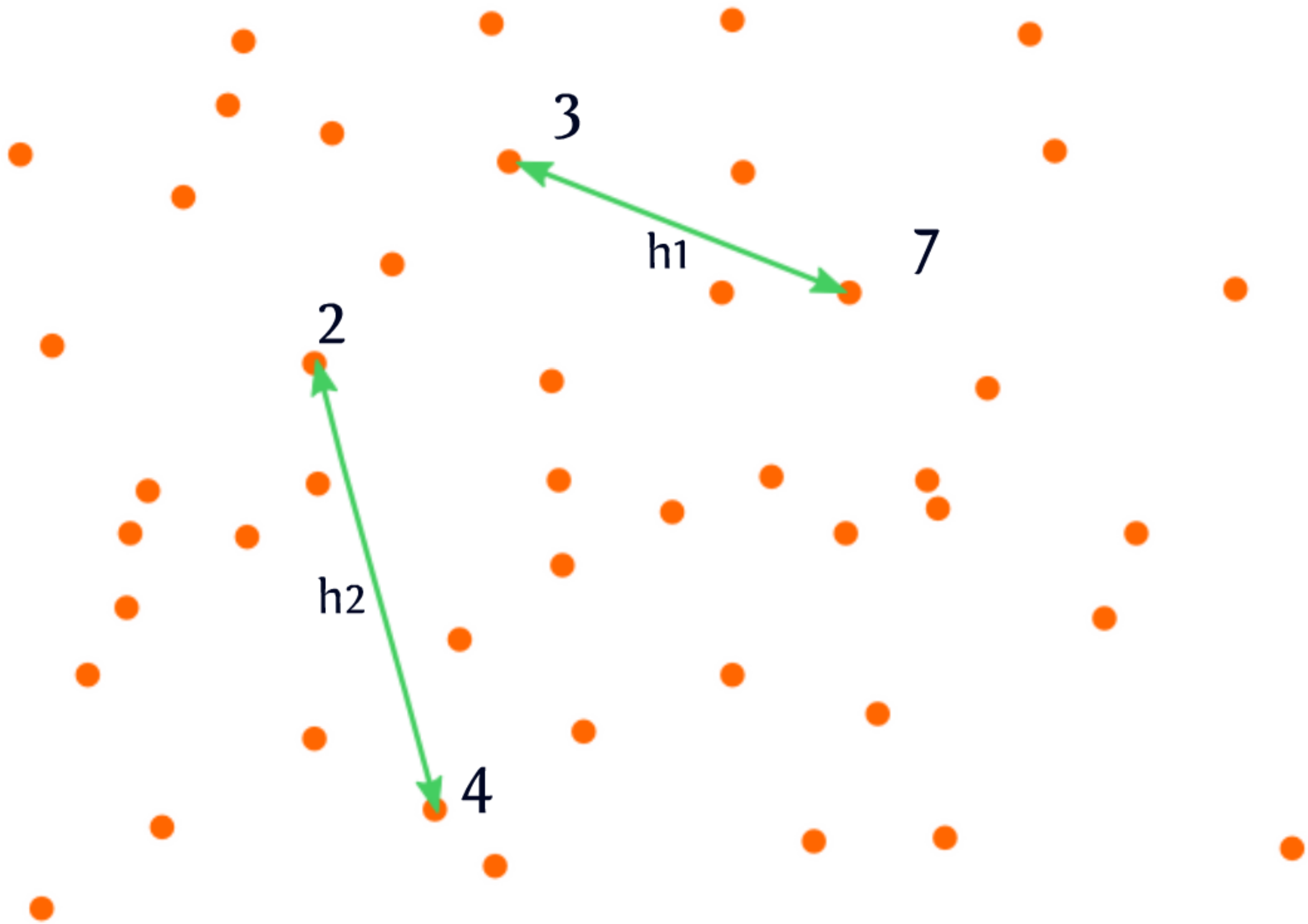
- Maximum width of variogram: 1/2 side of sampling grid or 1/2 sqrt(area).
- Number of lags: make sure to have enough points at each lag.
- Representation of small distances: a challenge.
- Anisotropy - different variogram in different directions. Variogram map.



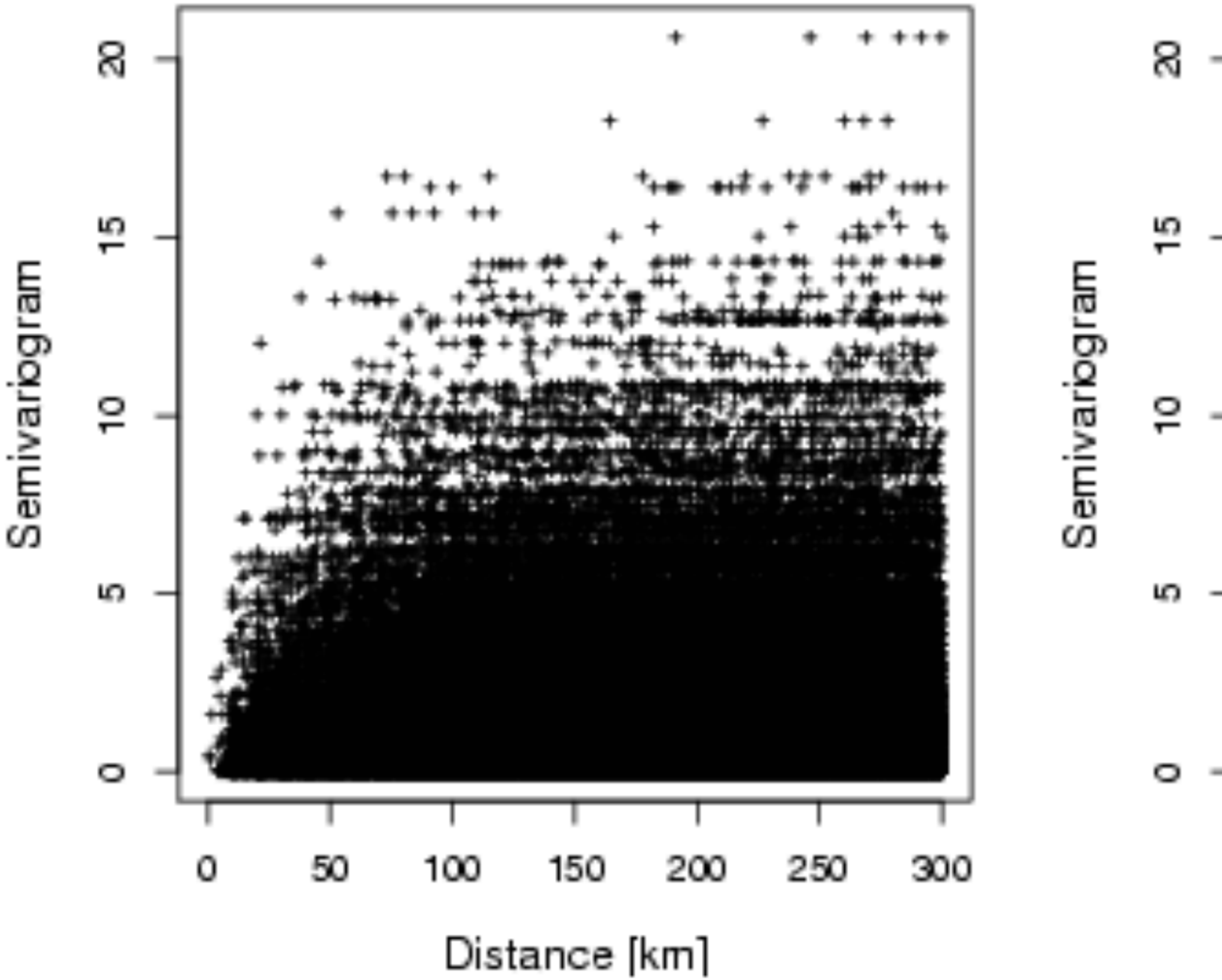
Spatial autocorrelation

- Geary's C - linked to the variogram
- Moran's I - presence of autocorrelation "globally" or at each lag.
- Package 'spdep'
- Custom functions provided



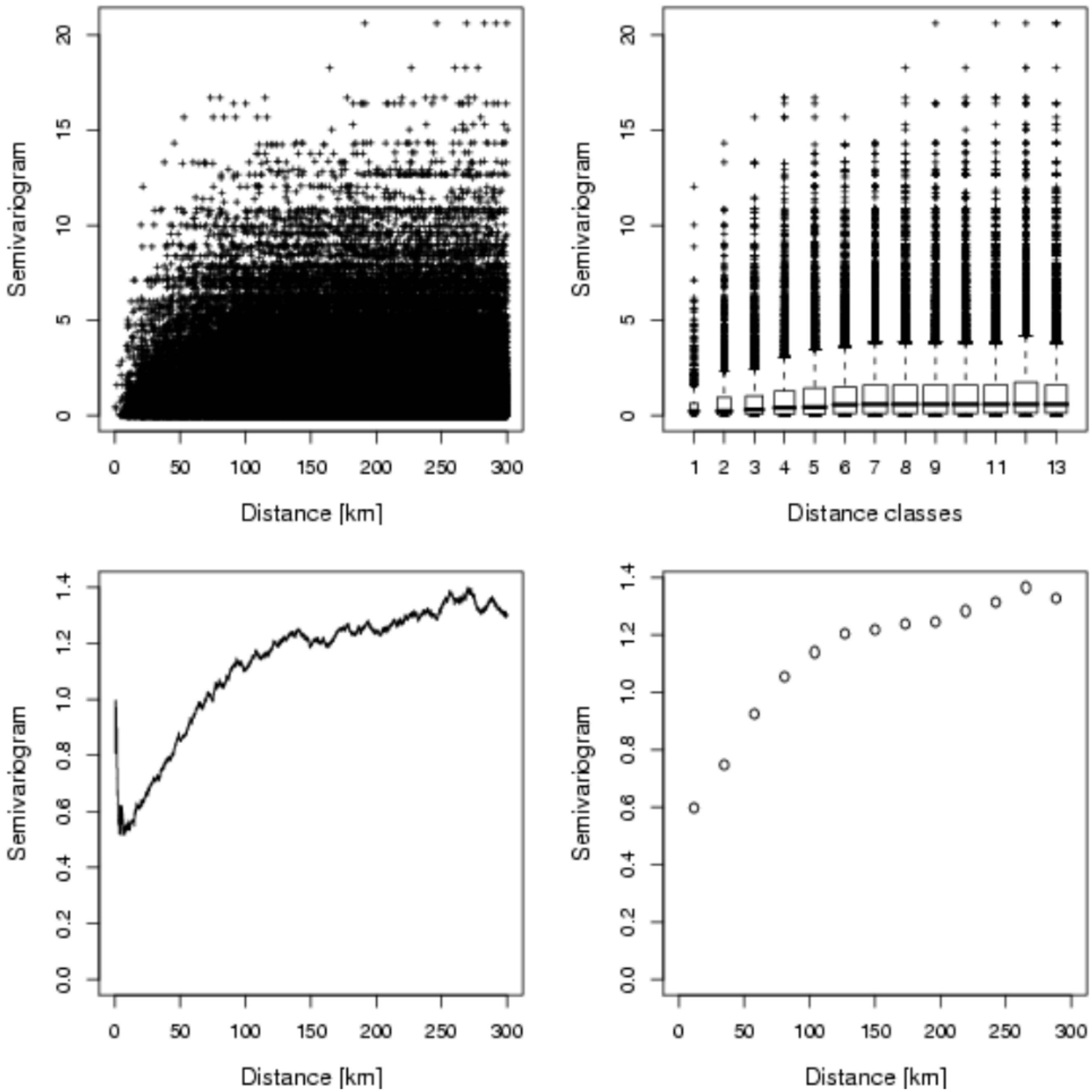


$$(z(u) - z(u + h))^2$$

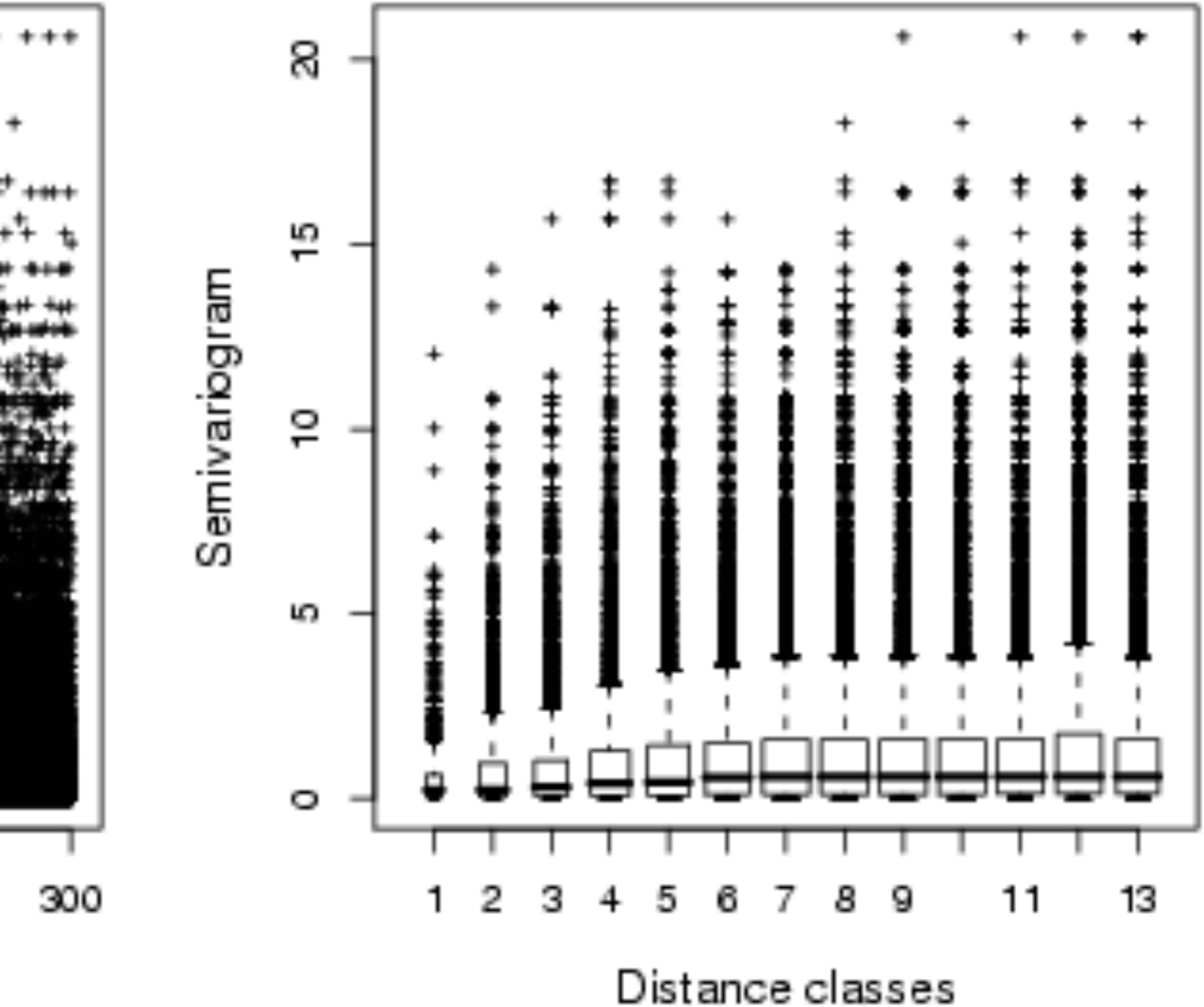


Semivariogram
0
5
10
15
20
0
50
100
150
200
250
300
Distance [km]
Semivariogram
0
5
10
15
20

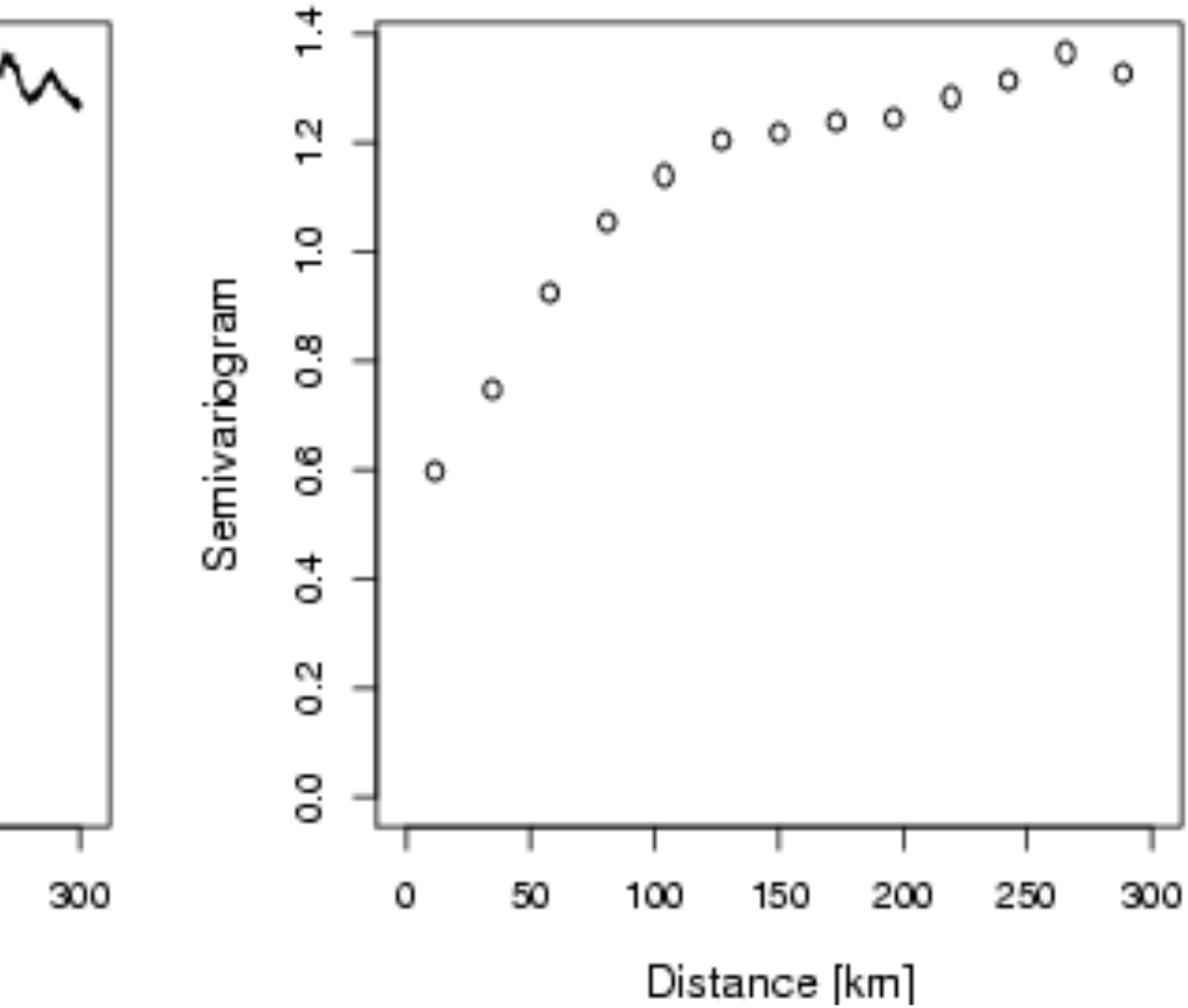
$$\tilde{\gamma}(h) = \frac{1}{2N(h)} \sum_{u=1}^{N} (z(u) - z(u+h))^2$$



$$\tilde{\gamma}(h) = \frac{1}{2N(h)} \sum_{u=1}^{N} (z(u) - z(u+h))^2$$



300
Semivariogram
0
5
10
15
20
1
2
3
4
5
6
7
8
9
11
13
Distance classes



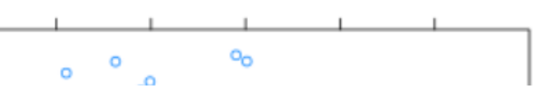
300
Semivariogram
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0
50
100
150
200
250
300
Distance [km]

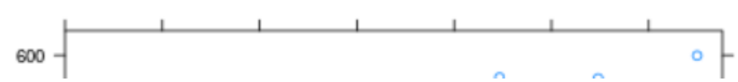
Rules of thumb

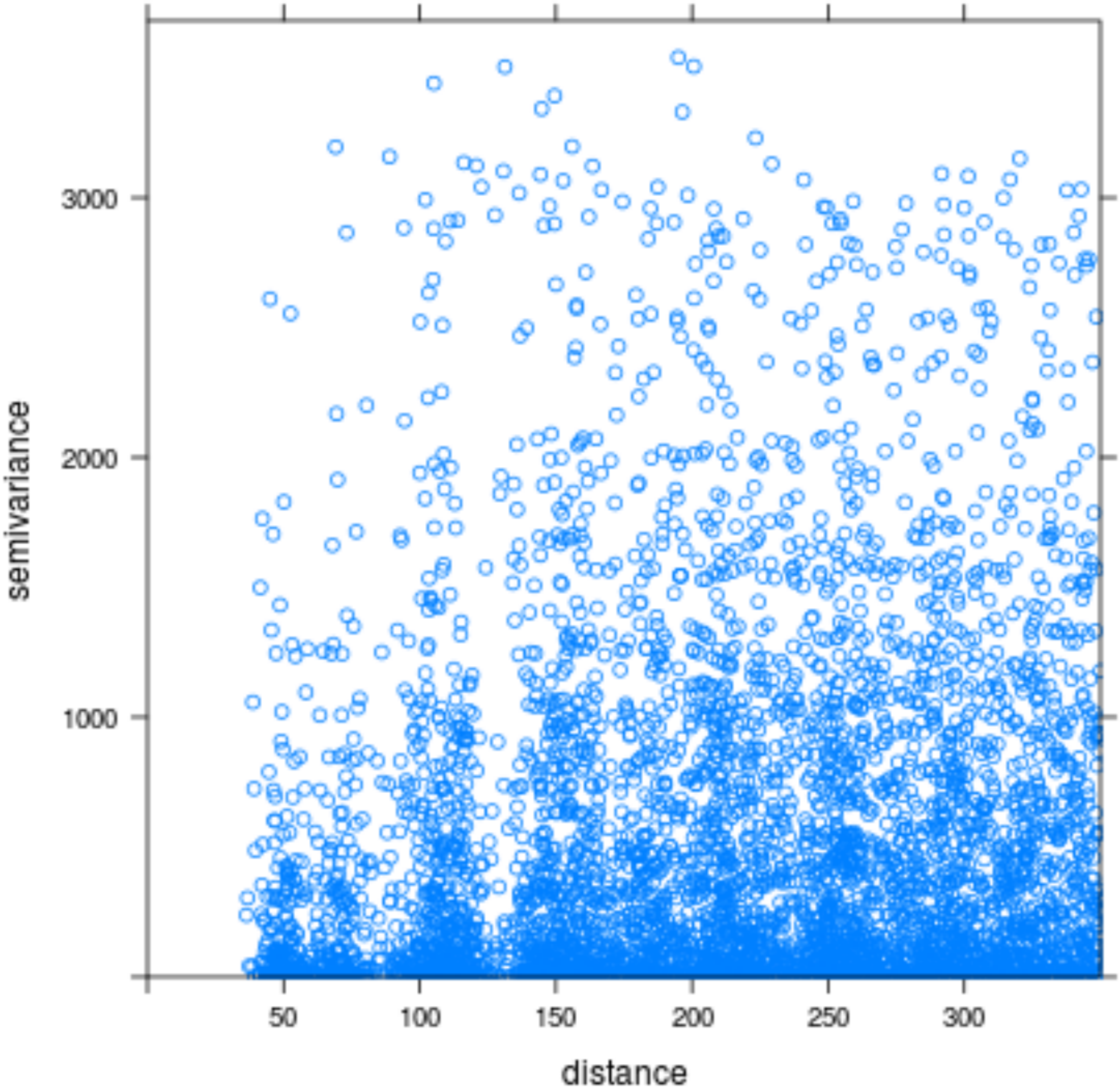
- Maximum width of variogram: 1/2 side of sampling grid or 1/2 sqrt(area).
- Number of lags: make sure to have enough points at each lag.
- Representation of small distances: a challenge.
- Anisotropy - different variogram in different directions. Variogram map.

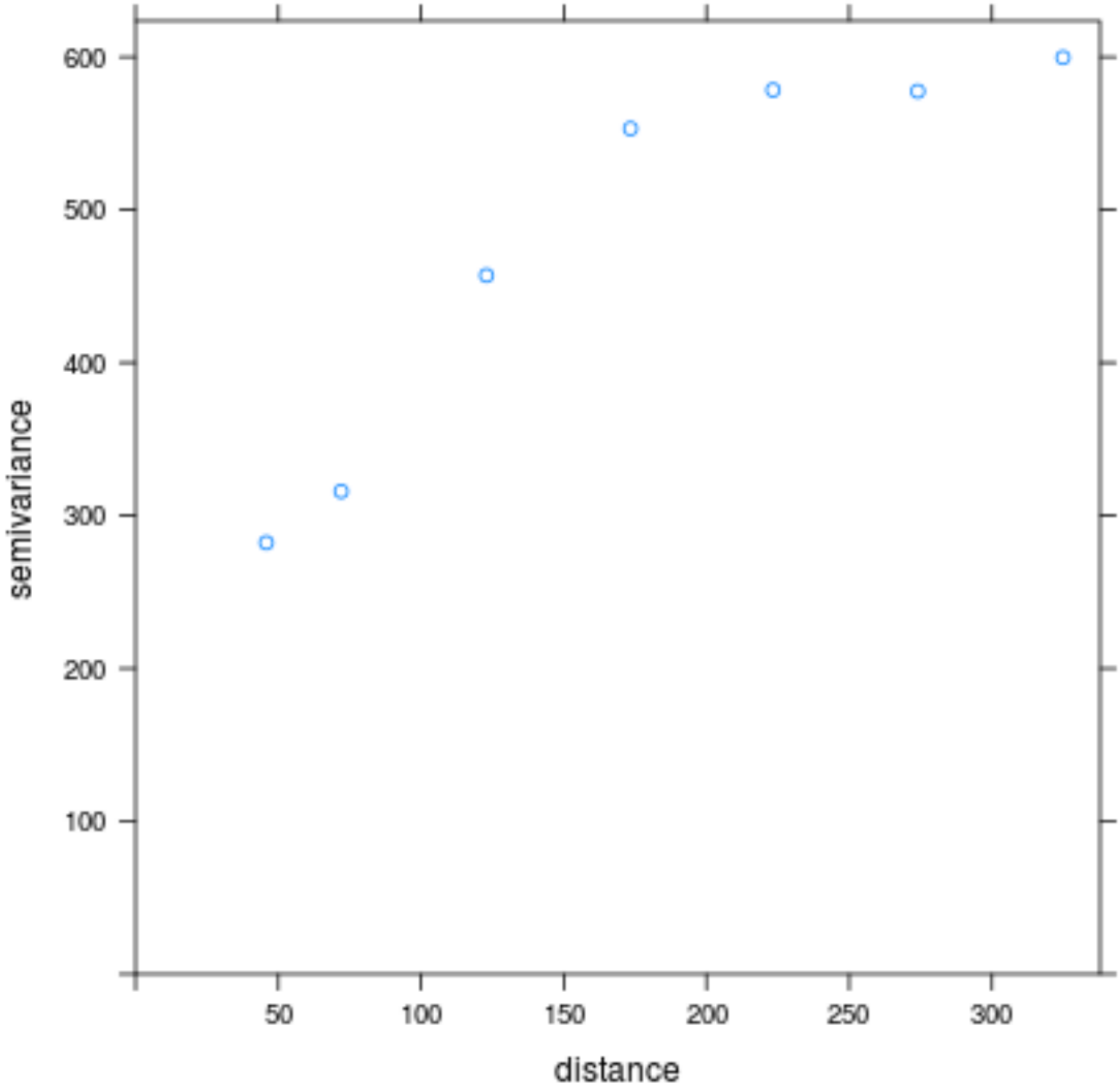
```r
# Semi-variogram for % Sand.
# Variogram cloud
Vario.SandCloud <- variogram(Sand~1,ArboSP,cloud=TRUE, cutoff=350)
plot(Vario.SandCloud)

# Average at each lag ('classical' variogram)
Vario.Sand <- variogram(Sand~1,ArboSP, cutoff=350, width=50)
plot(Vario.Sand)
```

```
# Semi-variogram for % Sand.
# Variogram cloud
Vario.SandCloud <- variogram(Sand~1,ArboSP,cloud=TRUE, cutoff=350)
plot(Vario.SandCloud)

# Average at each lag ('classical' variogram)
Vario.Sand <- variogram(Sand~1,ArboSP, cutoff=350, width=50)
plot(Vario.Sand)
```





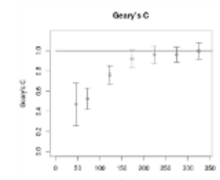

semivariance
3000
2000
1000
50
100
150
200
250
300
distance

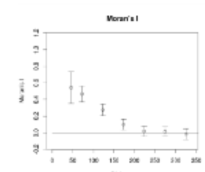

semivariance
600
500
400
300
200
100
50
100
150
200
250
300
distance

Spatial autocorrelation

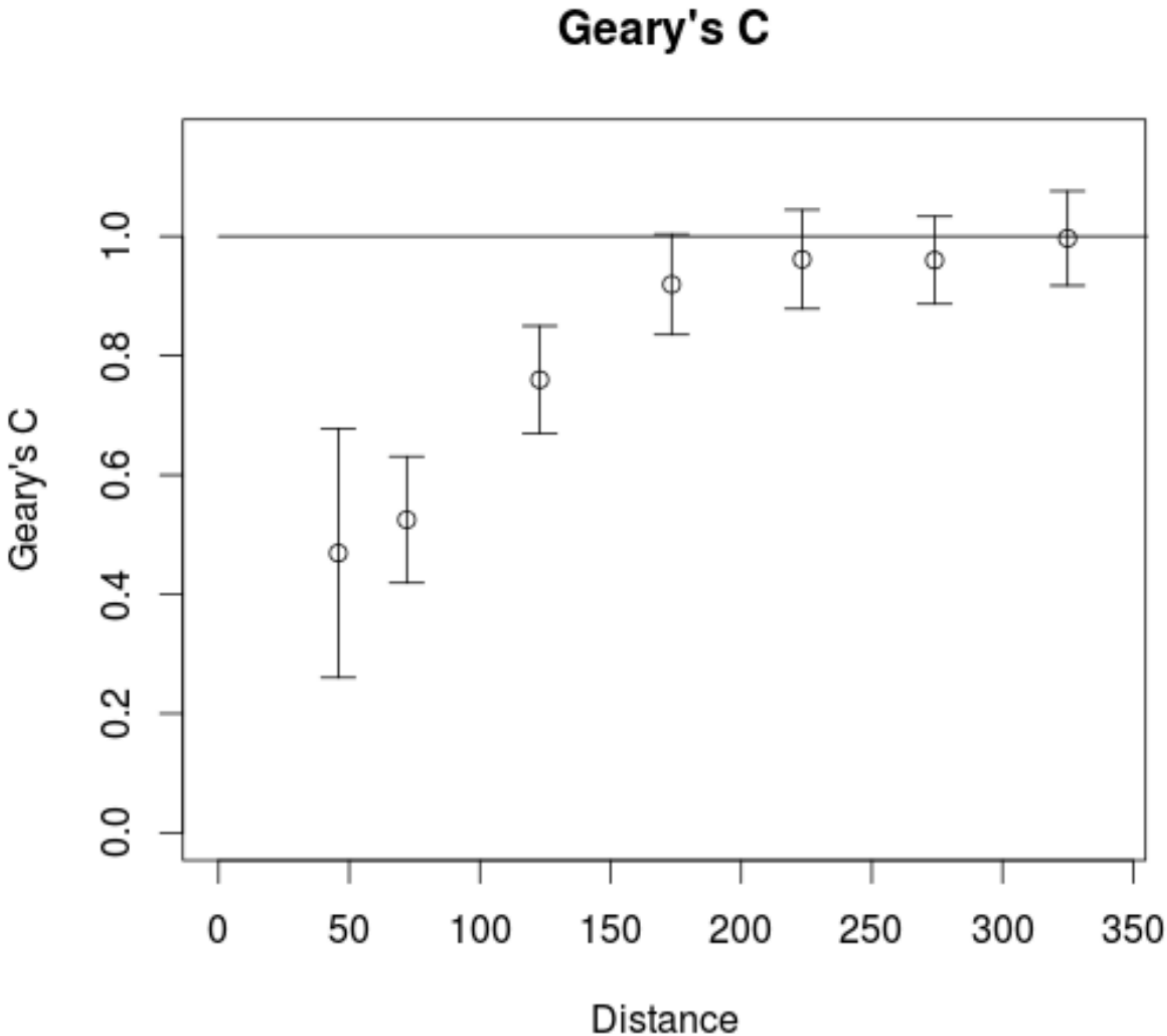
- Geary's C - linked to the variogram
- Moran's I - presence of autocorrelation "globally" or at each lag.
- Package 'spdep'
- Custom functions provided

```
# Geary C
corgramC<-GearyC(cbind(ArboSP$X,ArboSP$Y),ArboSP$Sand,seq(0,350,50))
plot(corgramC)

# Moran's I
corgramI<-MoranI(cbind(ArboSP$X,ArboSP$Y),ArboSP$Sand,seq(0,350,50))
plot(corgramI)
```

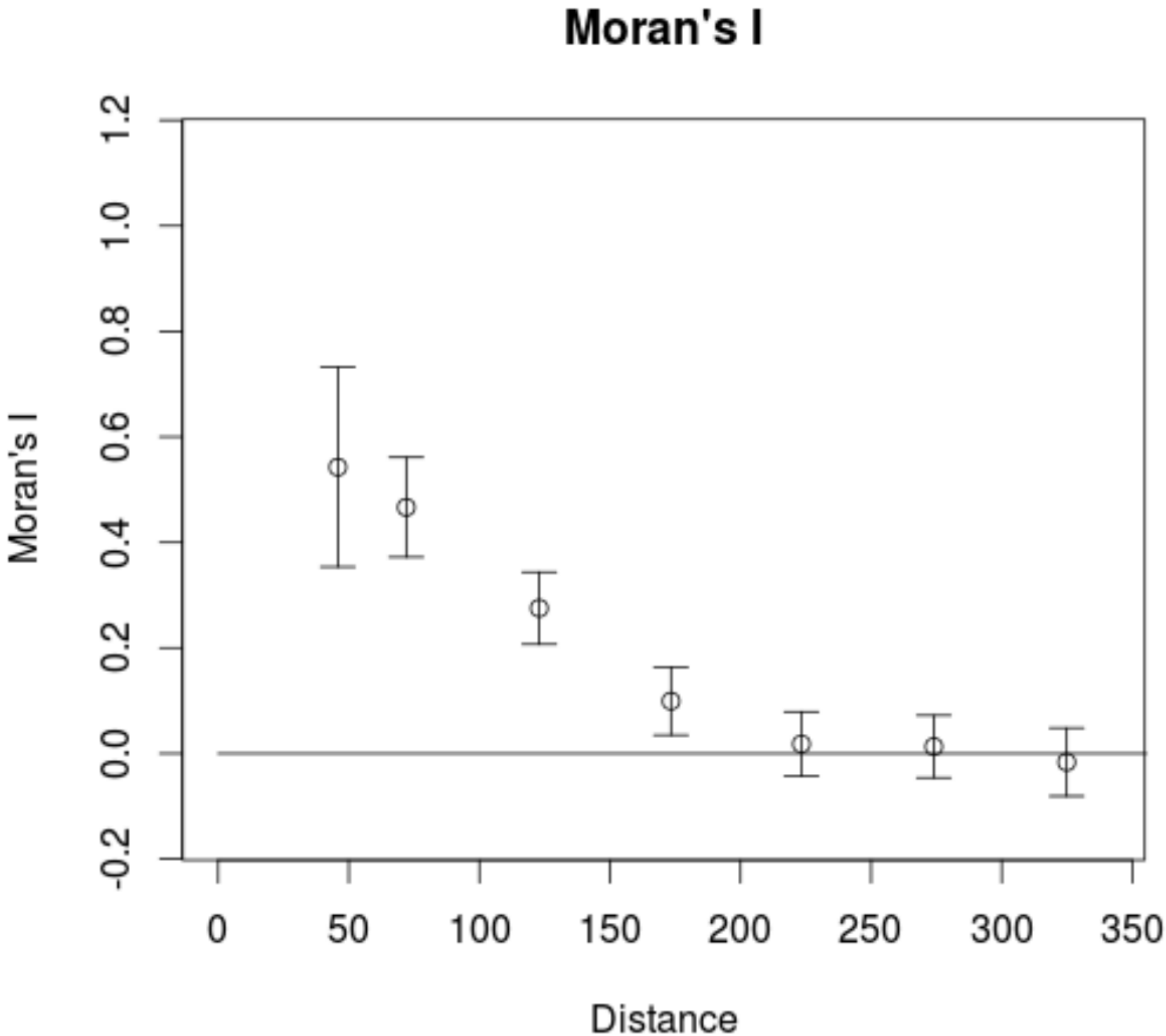




```r
# Geary C
corgramC<-GearyC(cbind(ArboSP$X,ArboSP$Y),ArboSP$Sand,seq(0,350,50))
plot(corgramC)

# Moran's I
corgramI<-MoranI(cbind(ArboSP$X,ArboSP$Y),ArboSP$Sand,seq(0,350,50))
plot(corgramI)
```



Geary's C
Geary's C
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150
200
250
300
350
Distance



Moran's I
Moran's I
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
50
100
150
200
250
300
350
Distance

Spatial objects in R

Package SP, rGDAL, gstat

```
# Select the Arbo_mask.tif file on your computer
Arbo_mask<-readGDAL(file.choose())
xy <- Arbo[1:2]
df <- Arbo[-1:-2]
ArboSP <- SpatialPointsDataFrame(coords=xy, data=df)
proj4string(ArboSP)<-proj4string(Arbo_mask)

# Specify the theme to be used in package SP
sp.theme(set = TRUE, regions = list(col = terrain.colors(100)))
sp.theme(set = TRUE, regions = list(col = colorRampPalette(c("black","brown","orange","yellow","white"
))(50)))

# Visualize some variables
spplot(ArboSP['Ca'])
```

Spatial objects in R

Package SP, rGDAL, gstat

```
# Select the Arbo_mask.tif file on your computer
Arbo_mask<-readGDAL(file.choose())
xy <- Arbo[1:2]
df <- Arbo[-1:-2]
ArboSP <- SpatialPointsDataFrame(coords=xy, data=df)
proj4string(ArboSP)<-proj4string(Arbo_mask)

# Specify the theme to be used in package SP
sp.theme(set = TRUE, regions = list(col = terrain.colors(100)))
sp.theme(set = TRUE, regions = list(col = colorRampPalette(c("black","brown","orange","yellow","white"
))(50)))

# Visualize some variables
spplot(ArboSP['Ca'])
```

Spatial objects in R

ackage SP, rGDAL, gstat

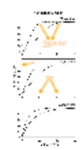
```
# Select the Arbo_mask.tif file on your computer
Arbo_mask<-readGDAL(file.choose())
xy <- Arbo[1:2]
df <- Arbo[-1:-2]
ArboSP <- SpatialPointsDataFrame(coords=xy, data=df)
proj4string(ArboSP)<-proj4string(Arbo_mask)

# Specify the theme to be used in package SP
sp.theme(set = TRUE, regions = list(col = terrain.colors(100)))
sp.theme(set = TRUE, regions = list(col = colorRampPalette(c("black","brown","orange","yellow","white"
))(50)))

# Visualize some variables
spplot(ArboSP['Ca'])
```

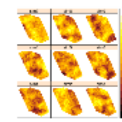
Modeling spatial structure and unconditional Gaussian simulations

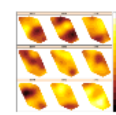
Random function

- Theoretical model characterized by an autocovariance function/variogram model and a mean value.
- Infinite extent.
- Spatially structured but with a random/stochastic aspect.

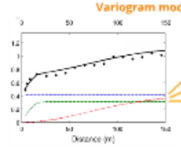
Variogram functions



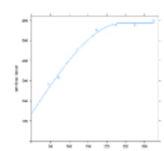




Variogram modeling




```
Vario <- variogram(SuM ~ 1, ~ X + Y, Arbo)
plot(Vario$dist,Vario$gamma)
v <- vgm(500, "Sph", 200, nug = 250)
model = fit.variogram(Vario, model = v)
plot(Vario, model=model)
```

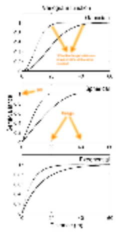


Modeling spatial structure and unconditional Gaussian simulations

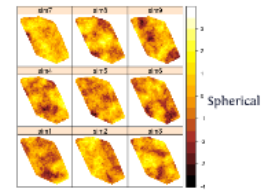
Random function

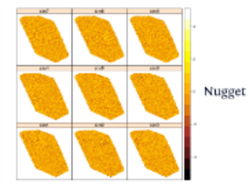
- Theoretical model characterized by an autocovariance function/variogram model and a mean value.
- Infinite extent.
- Spatially structured but with a random/ stochastic aspect.

Variogram functions

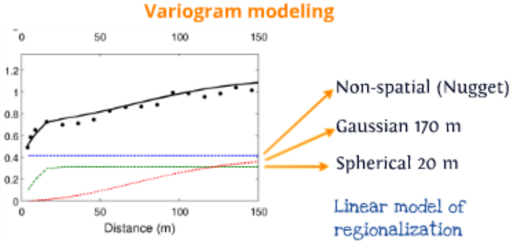


```
v <- vgm(1, "Sph", 200)
g.dummy <- gstat(formula = z~1, locations=ArboSP, dummy = TRUE, beta = 0,model = v, nmax = 20)
g.prd<-predict(g.dummy,Arbo_mask,nsim=9)
spplot(g.prd)
```

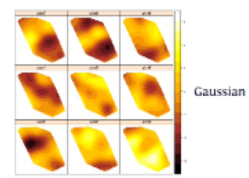


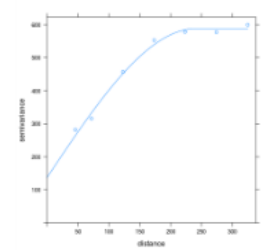


Variogram modeling




```
Vario <- variogram(SuM ~ 1, ~ X + Y, Arbo)
plot(Vario$dist,Vario$gamma)
v <- vgm(500, "Sph", 200, nug = 250)
model = fit.variogram(Vario, model = v)
plot(Vario, model=model)
```

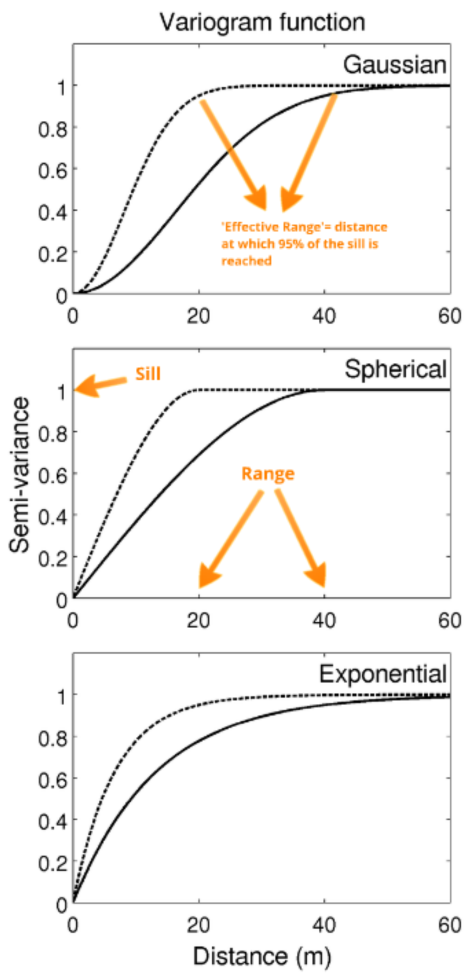


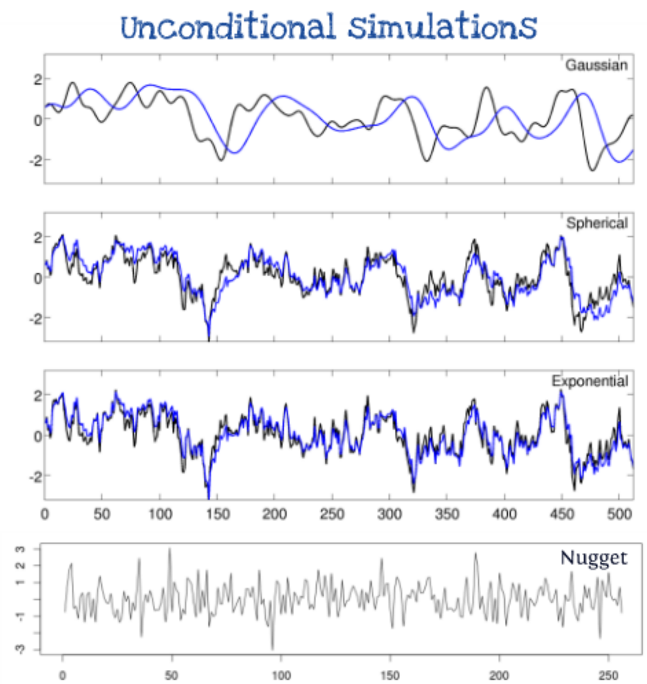


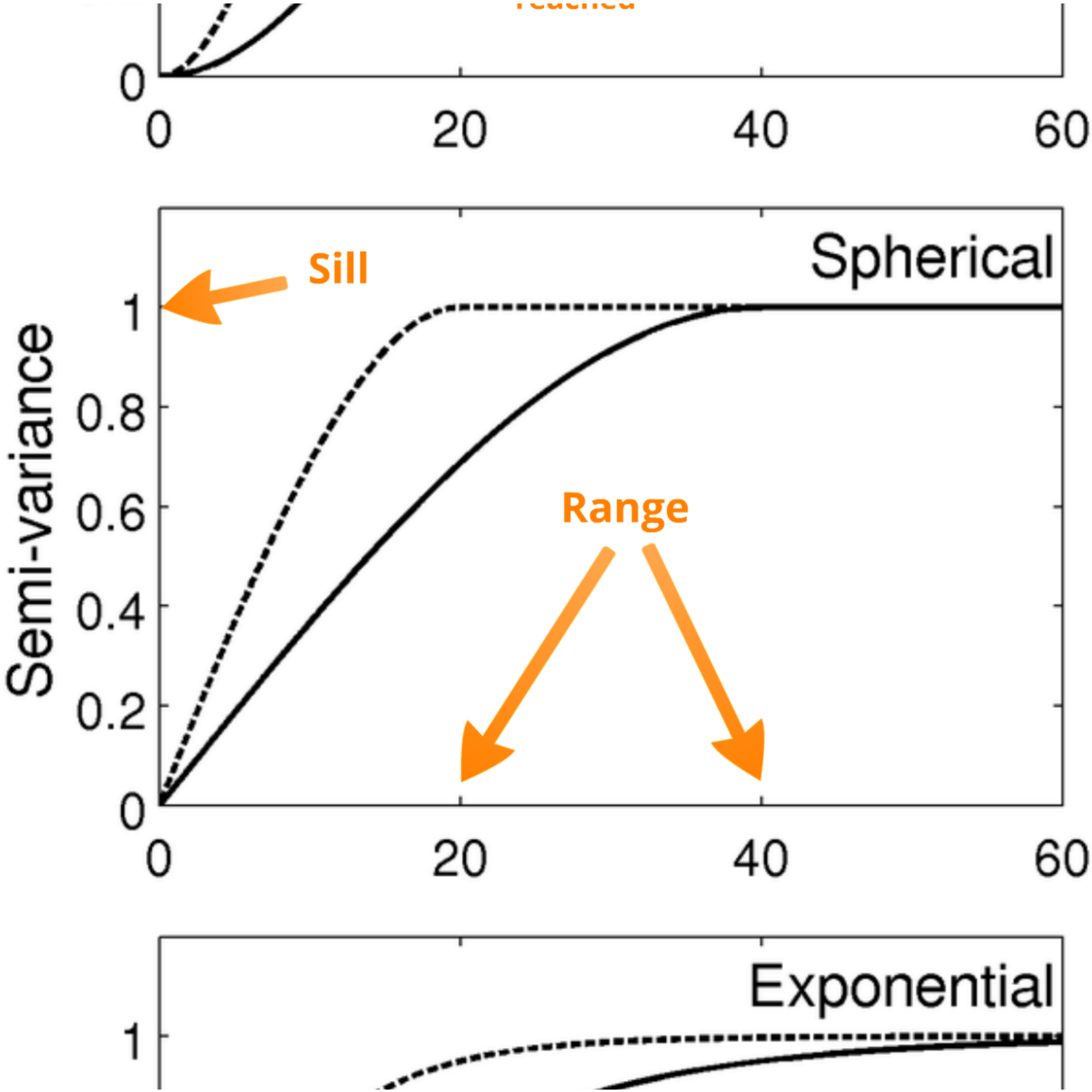
Random function

- Theoretical model characterized by an autocovariance function/variogram model and a mean value.
- Infinite extent.
- Spatially structured but with a random/ stochastic aspect.

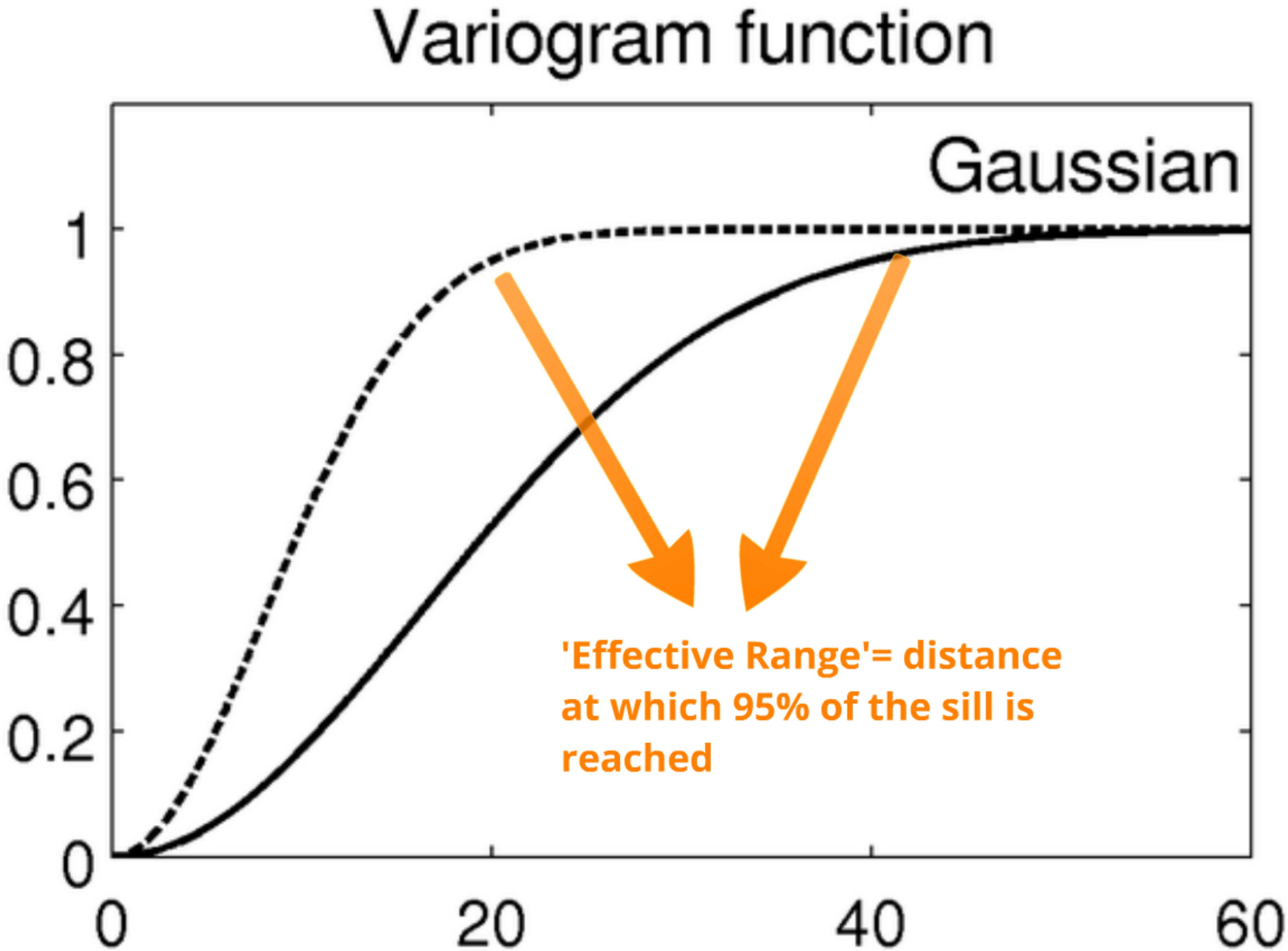
Variogram functions

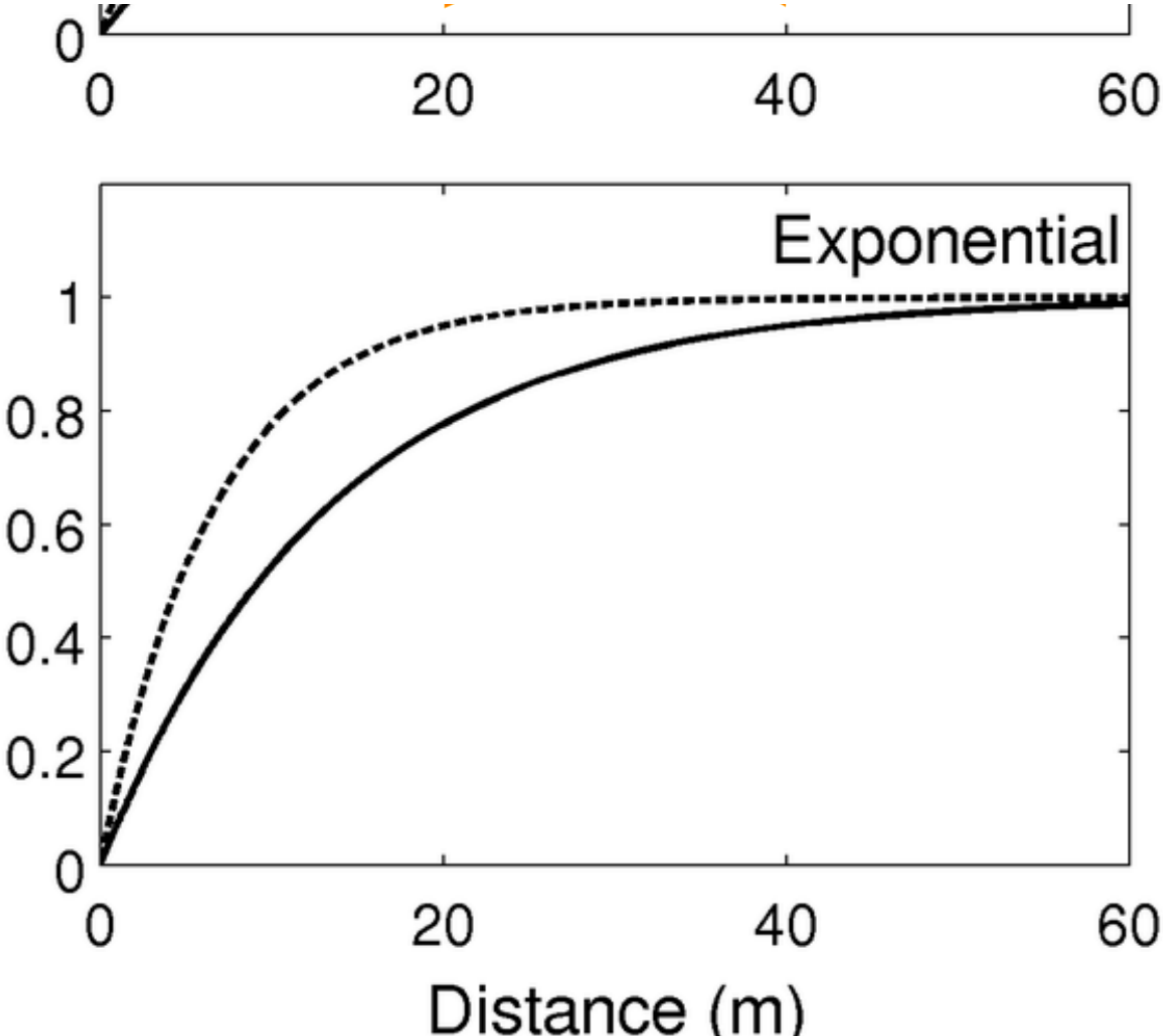




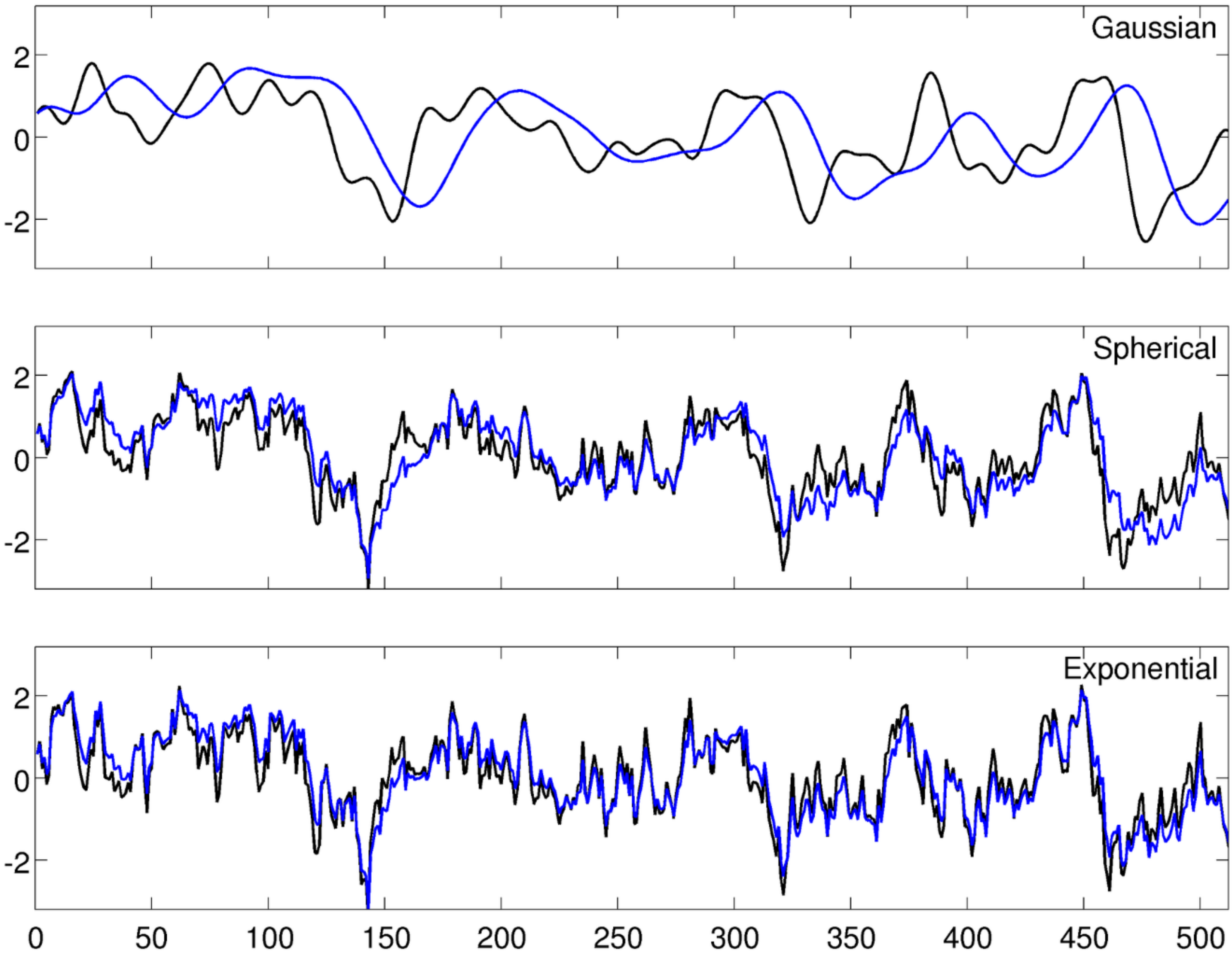


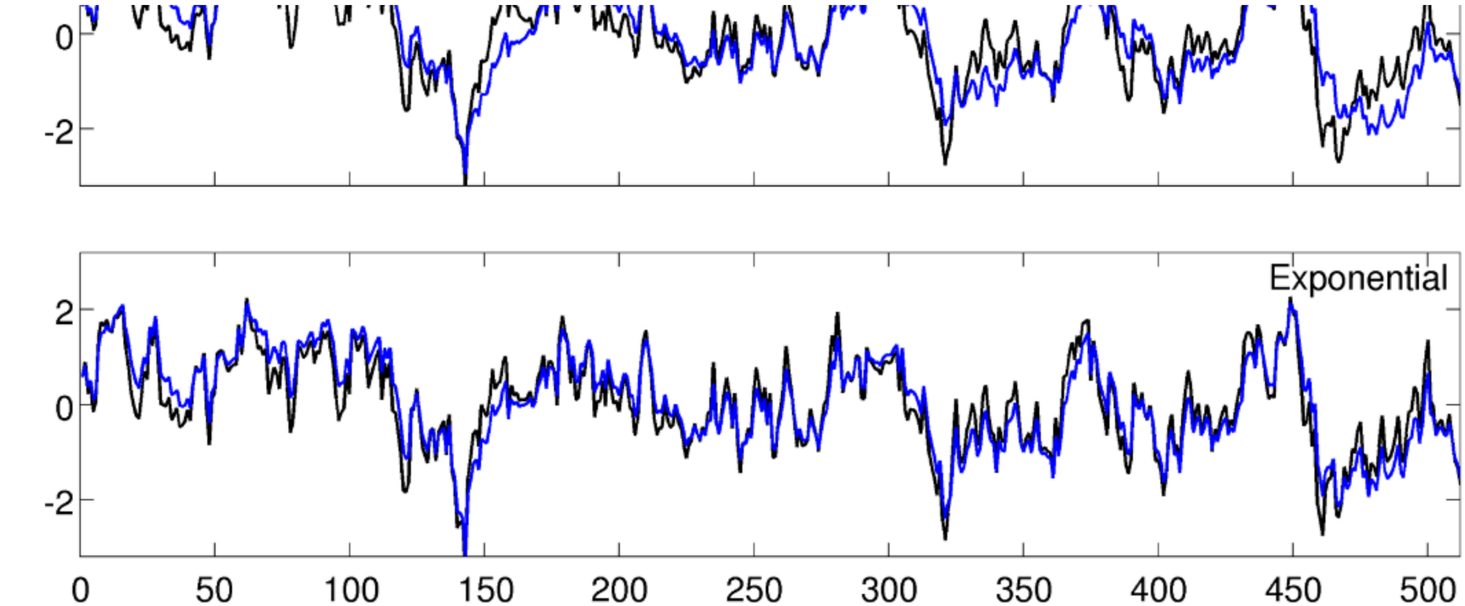
0
20
40
60
Spherical
Sill
1
0.8
0.6
0.4
0.2
0
Semi-variance
Range
Exponential

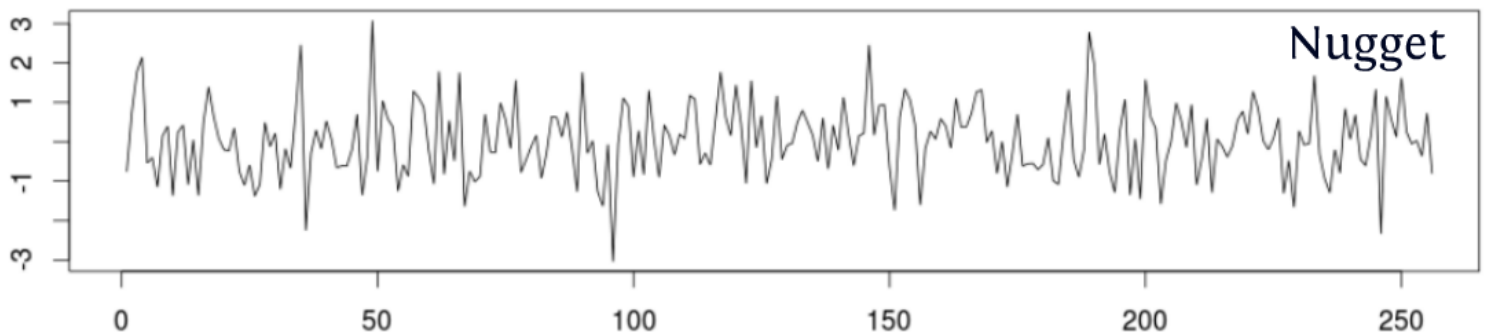

Variogram function
Gaussian
1
0.8
0.6
0.4
0.2
0
0
20
40
60
'Effective Range'= distance at which 95% of the sill is reached


Exponential
0
0.2
0.4
0.6
0.8
1
0
20
40
60
Distance (m)

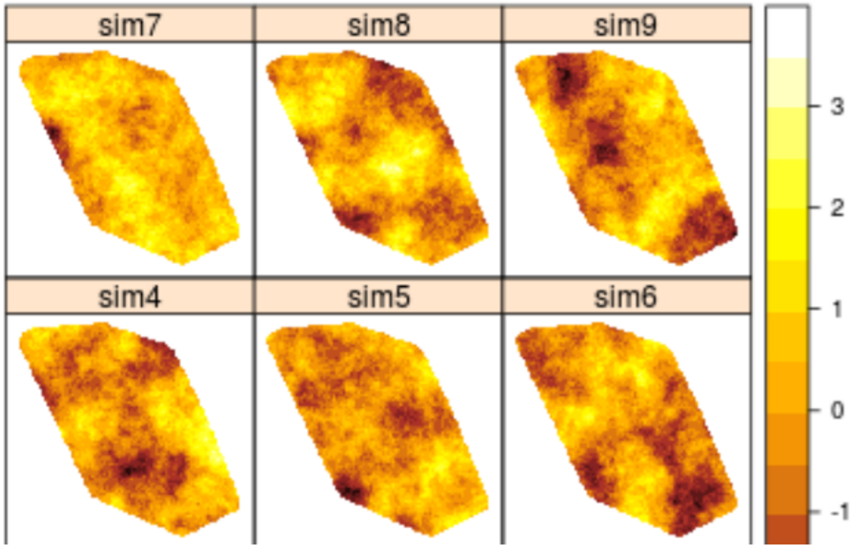
Unconditional Simulations



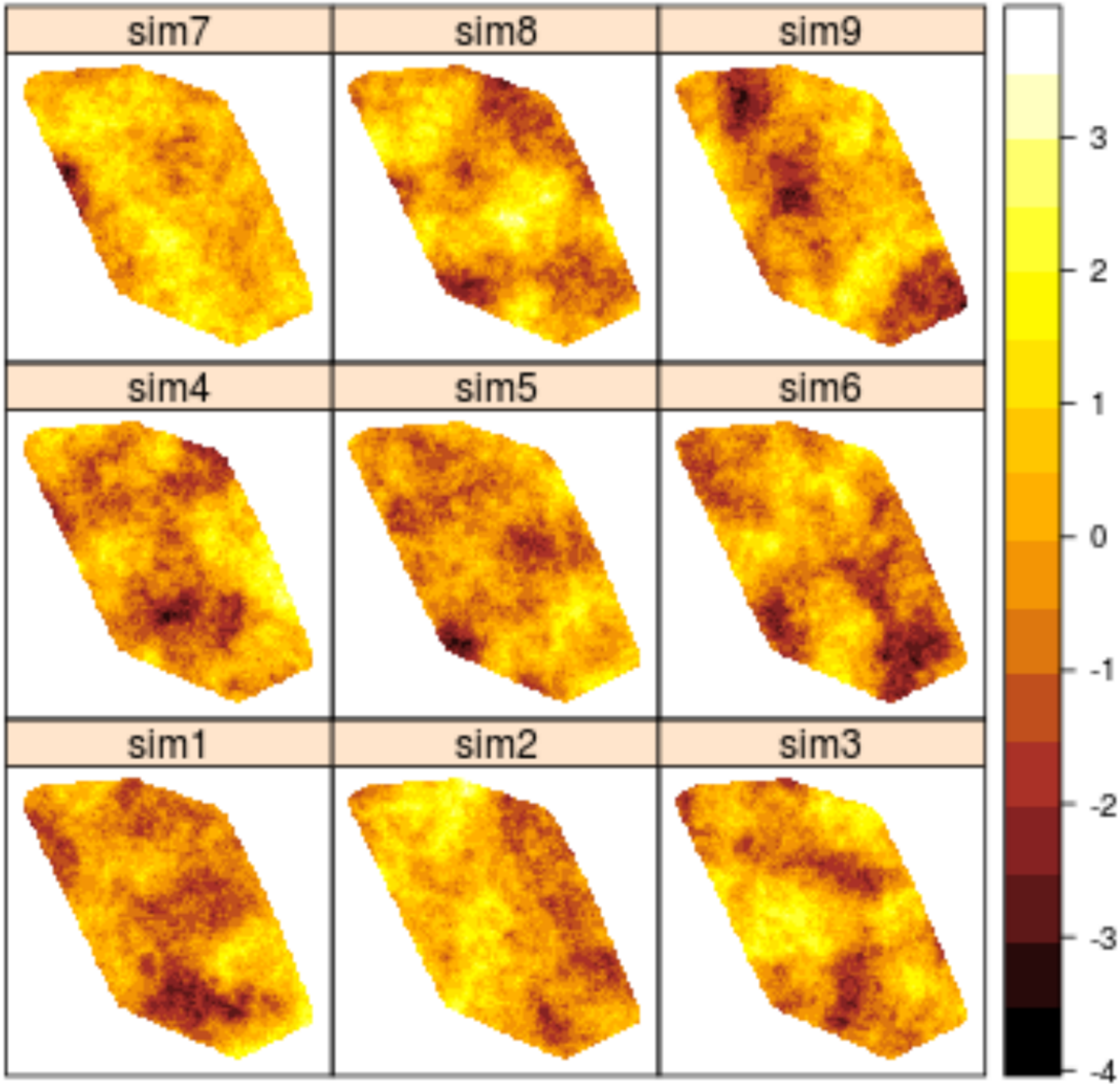

0
-2
Exponential
2
0
-2
0
50
100
150
200
250
300
350
400
450
500


Nugget
3
2
1
-1
-3
0
50
100
150
200
250

```
v <- vgm(1, "Sph", 200)
g.dummy <- gstat(formula = z~1, locations=ArboSP, dummy = TRUE, beta = 0,model = v, nmax = 20)
g.prd<-predict(g.dummy,Arbo_mask,nsim=9)
spplot(g.prd)
```

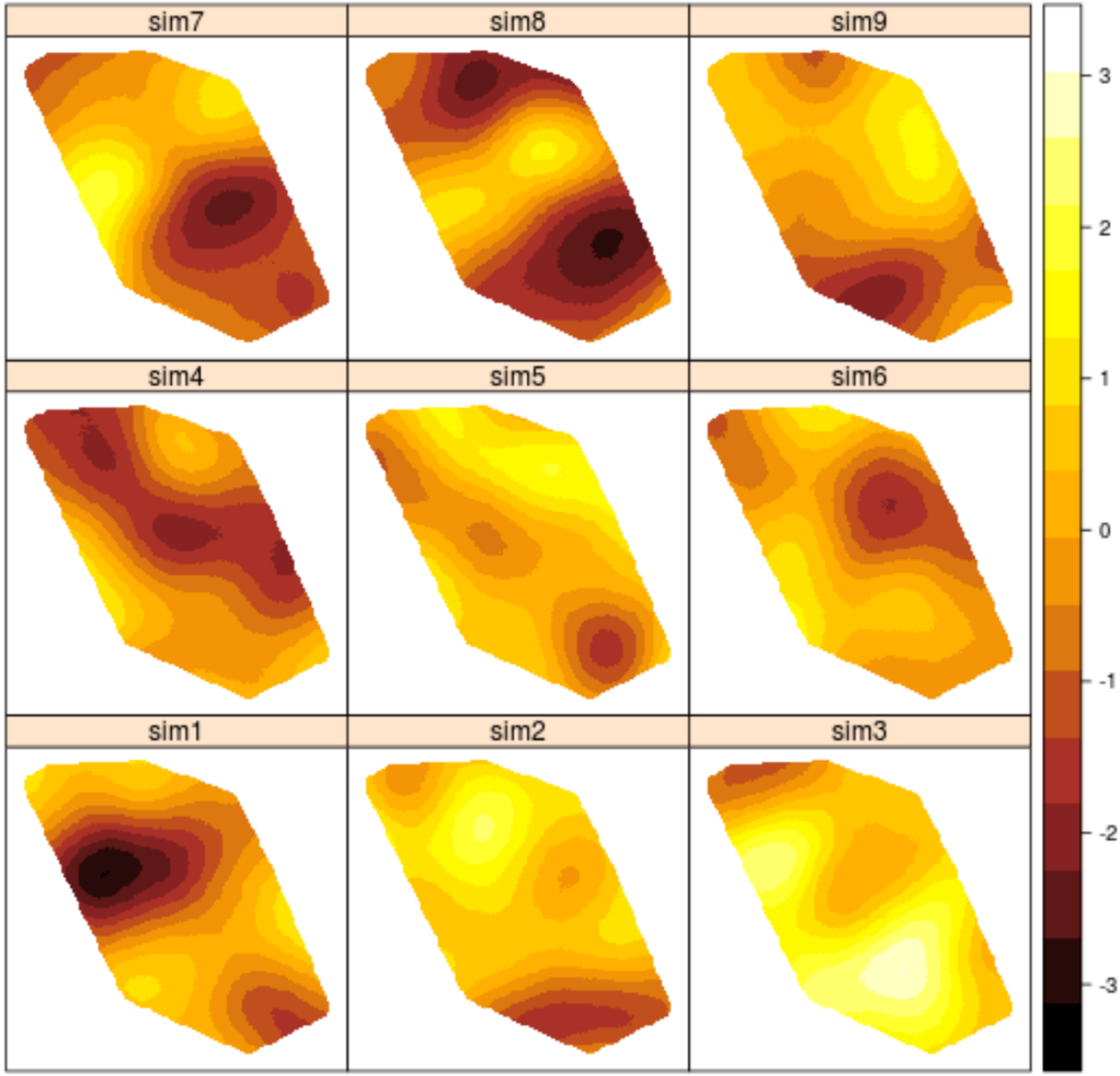


Spherical



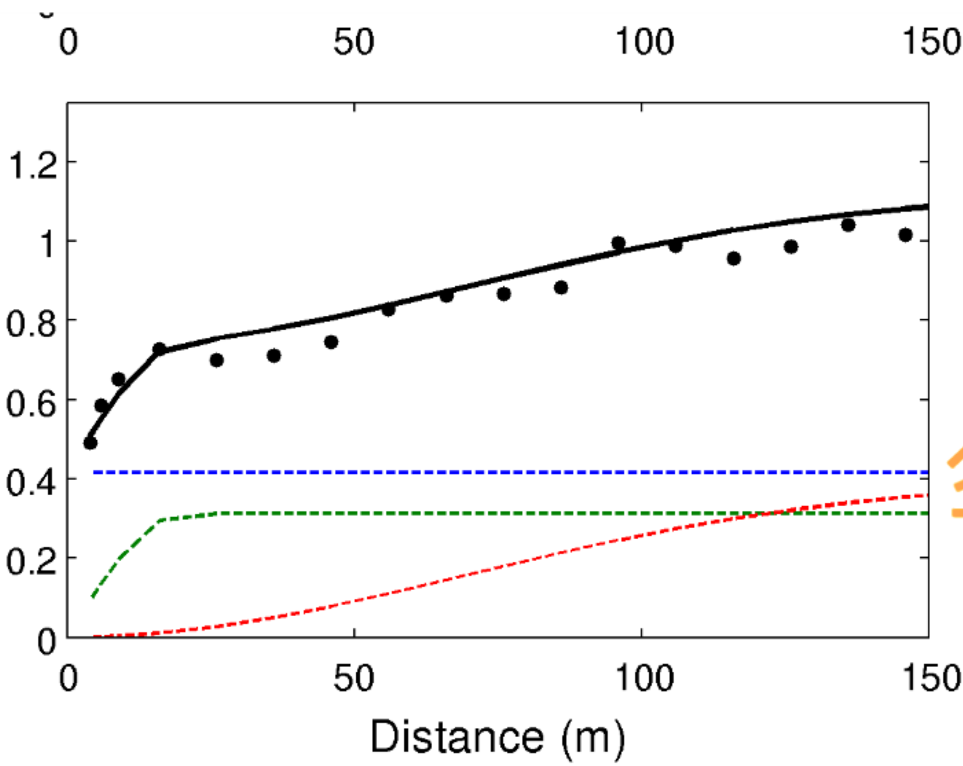
Spheri

Nug



Gau

Variogram modeling



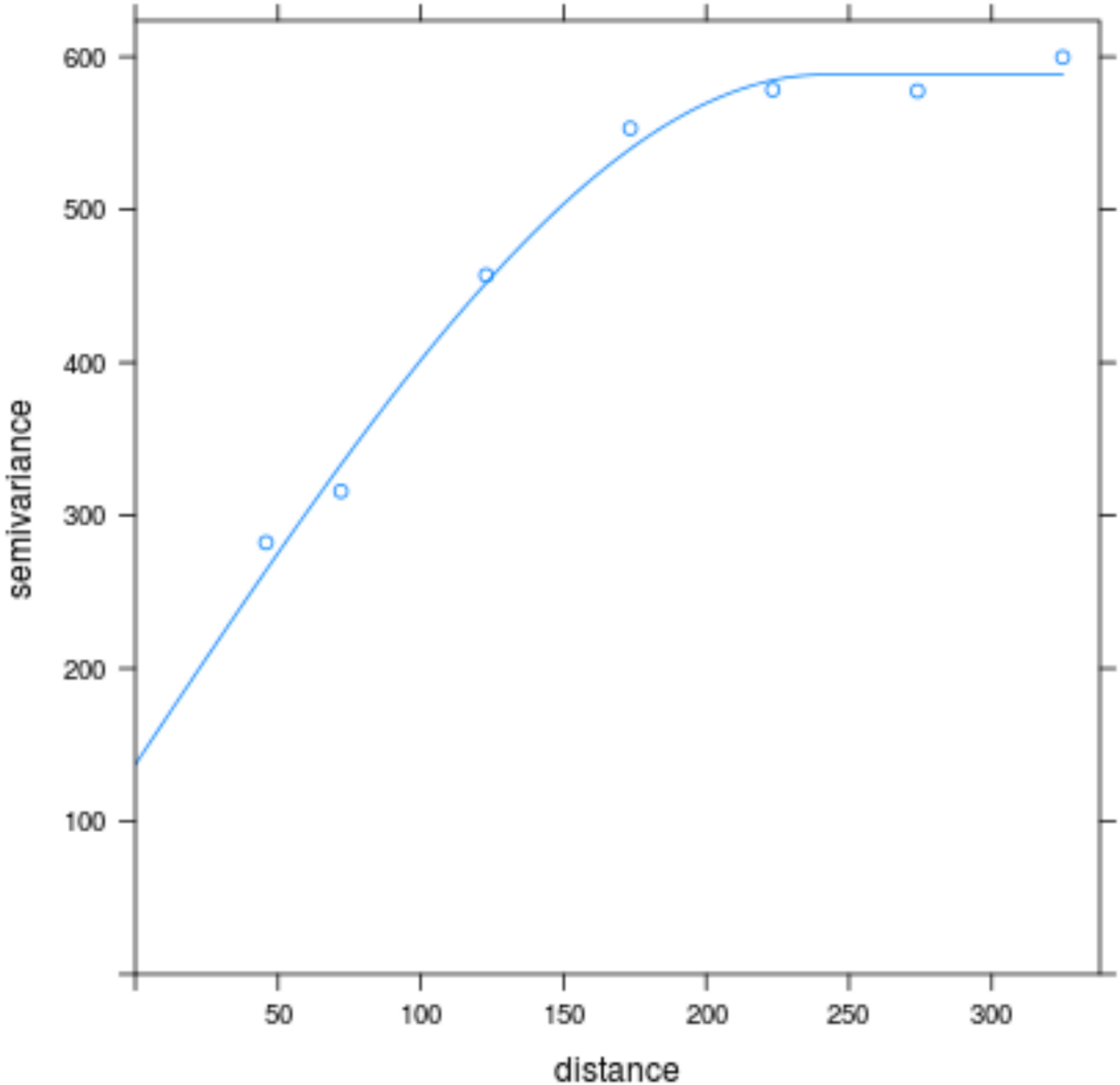
Non-spatial (Nugget)

Gaussian 170 m

Spherical 20 m

Linear model of regionalization

```
Vario <- variogram(SuM ~ 1, ~ X + Y, Arbo)
plot(Vario$dist,Vario$gamma)
v <- vgm(500, "Sph", 200, nug = 250)
model = fit.variogram(Vario, model = v)
plot(Vario, model=model)
```


600
500
400
300
200
100
semivariance
50
100
150
200
250
300
distance

Estimation and kriging

Objectives

- Estimate values at unsampled locations.
- Get an idea of the uncertainty of the estimates.
- Visualization, creation of continous rasters.
- Get an idea of the uncertainty of the estimates.

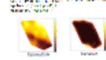
Kriging types

Simple kriging = One global system. Use all points
Ordinary kriging = with local neighbourhood.
Indicator kriging = binary values
Universal kriging, kriging with a trend = Non-stationarity, trend surface
Kriging with an external drift = Drift specified by external variable.
Block kriging = Estimation on a support larger than the distance between sampling points (multiple points within support)
Co-krigeage = Two or more values together.
Factorial kriging = Multivariate, multiscale analysis.

Smoothing effect

- The more pronounced the nugget, the greater the smoothing effect.
- Strong smoothing effect, map is not a good representation of the variable in reality.

Variables without a 'random' aspect: better use spline, TIN, or other deterministic functions

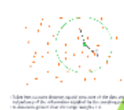


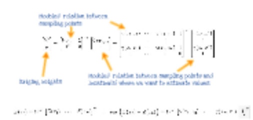
Deterministic functions

- Polynomial functions
- Splines (piecewise polynomials)
- Triangular irregular networks, Delaunay triangulation

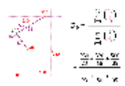
Kriging

- Well adapted to ecological data (random/ stuctured).
- Maps are often visually 'realistic'.
- Get a map of kriging variance.
- Smothing effect more pronounced where then is less spatial structuring.
- Measure of estimation variance.





Inverse distance weighting



- Maps are often non-satisfactory
- No indication of estimation uncertainty

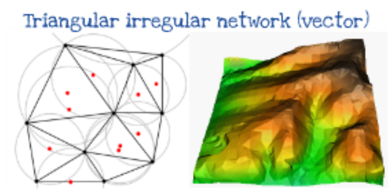
Objectives

- Estimate values at unsampled locations.
- Get an idea of the uncertainty of the estimates.
- Visualization, creation of continous rasters.
- Get an idea of the uncertainty of the estimates.

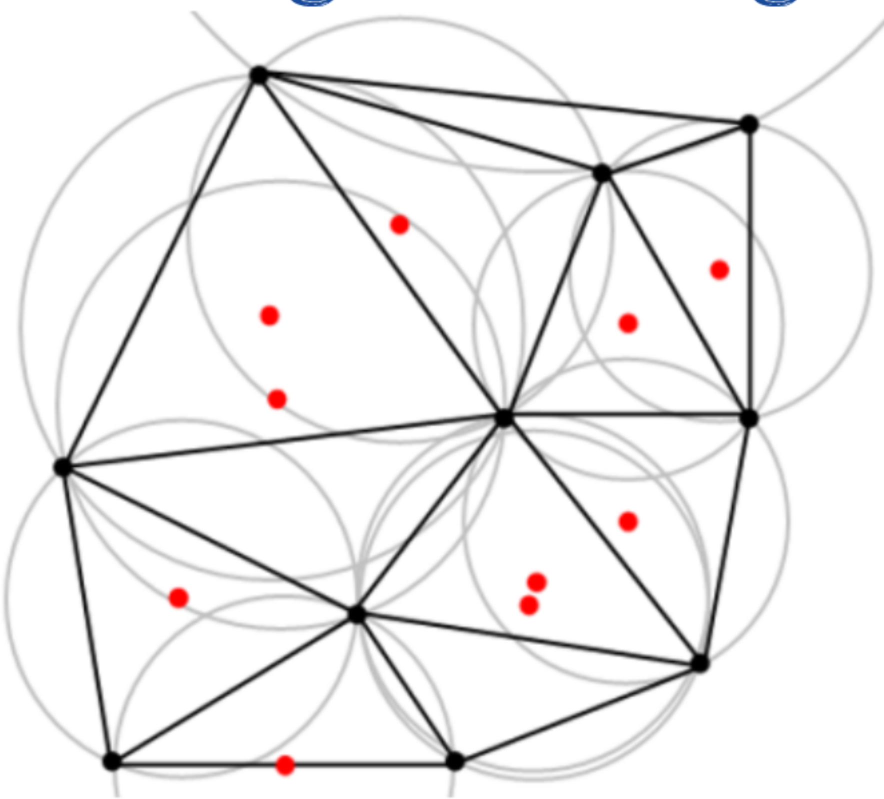
Deterministic functions

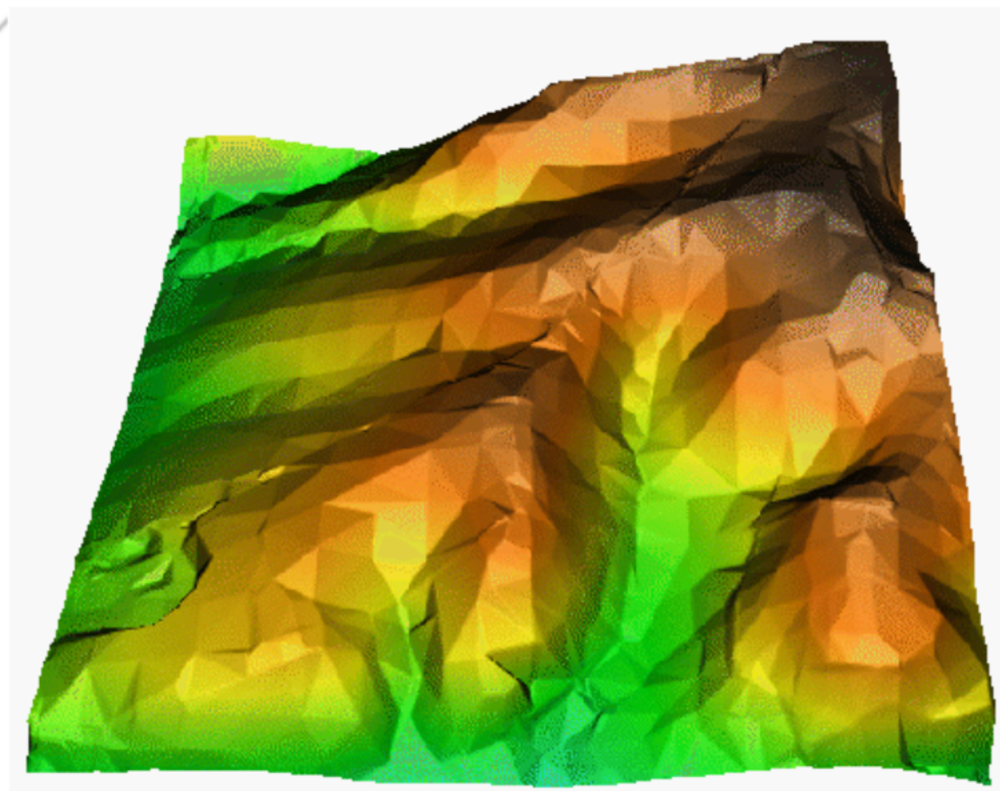
- Polynomial functions
- Splines (piecewise polynomials)
- Triangular irregular networks. Delaunay triangulation

```
# Thin plate spline
tps <- Tps(xy, Arbo$pH)
ras<-raster(Arbo_mask)
spline.pH <- interpolate(ras, tps)
spline.pH <- mask(spline.pH, ras)
spplot(spline.pH)
```



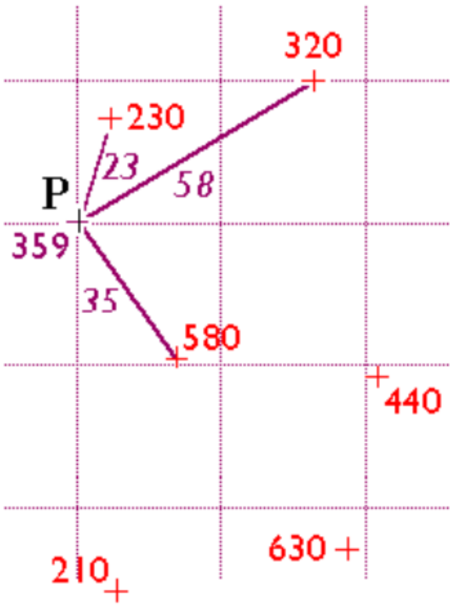
Triangular irregular network (vector)





```r
# Thin plate spline
tps <- Tps(xy, Arbo$pH)
ras<-raster(Arbo_mask)
spline.pH <- interpolate(ras, tps)
spline.pH <- mask(spline.pH, ras)
spplot(spline.pH)
```

Inverse distance weighting



$$z_P = \frac{\sum_{i=1}^{n} \left(\frac{z_i}{d_i}\right)}{\sum_{i=1}^{n} \left(\frac{1}{d_i}\right)}$$

$$= \frac{\frac{230}{23} + \frac{320}{58} + \frac{580}{35}}{\frac{1}{23} + \frac{1}{58} + \frac{1}{35}}$$

- Maps are often non-satisfactory
- No indication of estimation uncertainty

```
# Inverse Distance Weighting
idwr <- idw(pH ~ 1, ArboSP, Arbo_mask,idp=2)
spplot(idwr['var1.pred'])
```

tion

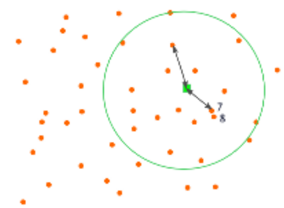
```
# Inverse Distance Weighting
idwr <- idw(pH ~ 1, ArboSP, Arbo_mask,idp=2)
spplot(idwr['var1.pred'])
```

Uncertainty assessm

- Look at kriging variance
- Look at variograms.
- Cross-validation.
- Separate dataset into esti and validation subsets.

Kriging

- Well adapted to ecological data (random/ stuctured).
- Maps are often visually 'realistic'.
- Get a map of kriging variance.
- Smothing effect more pronounced where then is less spatial structuring.
- Measure of estimation variance.



- Takes into account distance, spatial structure of the data and redundancy of the information supplied by the sampling points.
- At distances greater than the range, weights = 0

Modeled relation between sampling points

$$\begin{bmatrix} \hat{W} \\ \mu \end{bmatrix} = \begin{bmatrix} Var_{x_i} & 1 \\ 1^T & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} Cov_{x_i x_0} \\ 1 \end{bmatrix} = \begin{bmatrix} \gamma(x_1, x_1) & \cdots & \gamma(x_1, x_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(x_n, x_1) & \cdots & \gamma(x_n, x_n) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma(x_1, x^*) \\ \vdots \\ \gamma(x_n, x^*) \\ 1 \end{bmatrix}$$

Kriging weights

Modeled relation between sampling points and location(s) where we want to estimate values.

$$\hat{Z}(x_0) = \hat{W}^T \cdot [Z(x_1) \ \cdots \ Z(x_N)]^T \quad ; \quad var\left(\hat{Z}(x_0) - Z(x_0)\right) = \hat{W}^T \cdot [\gamma(x_1, x_0) \ \cdots \ \gamma(x_N, x_0) \ \ 1]^T$$

nation variance.

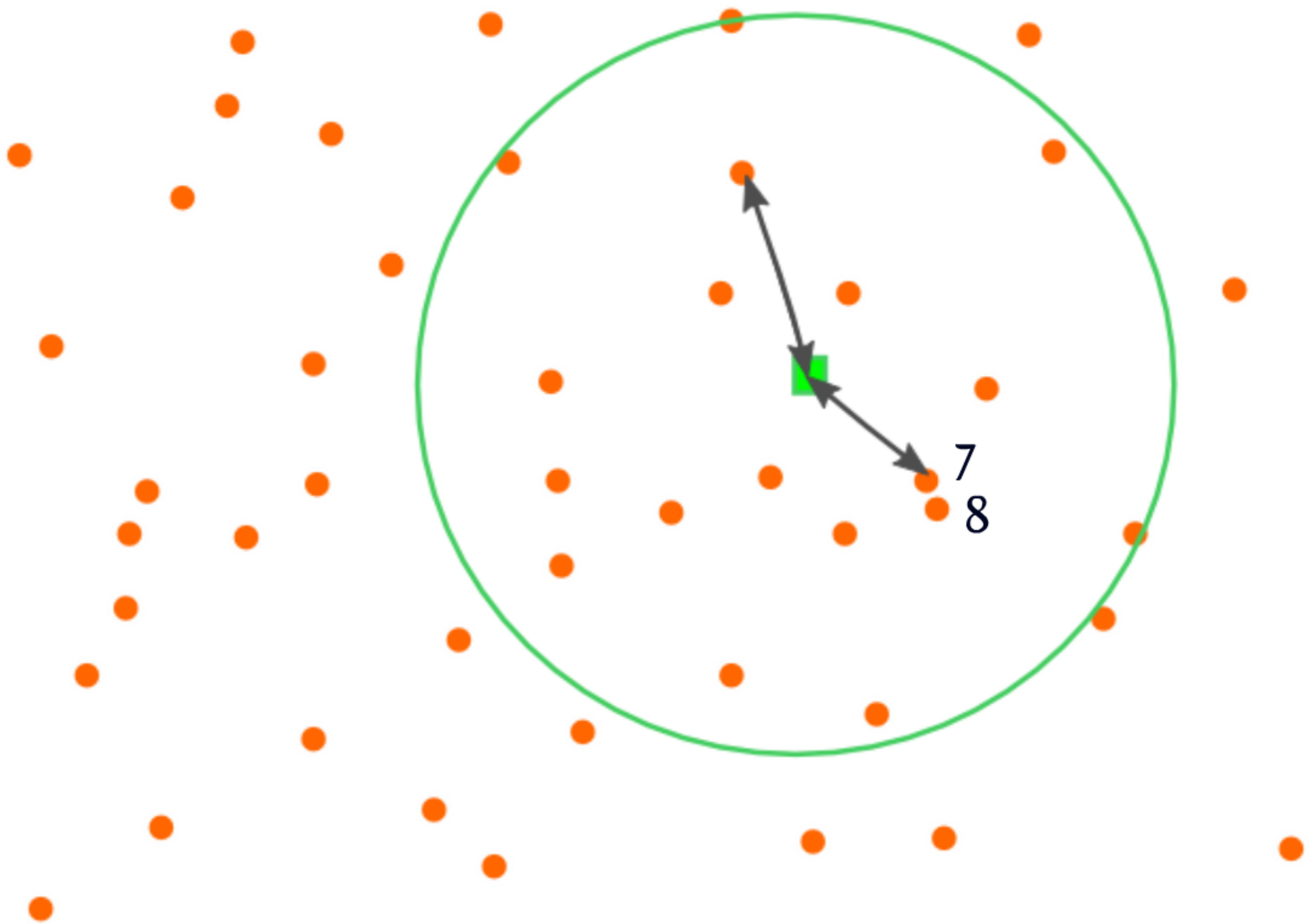
Modeled relation between sampling points

$$\begin{bmatrix}\hat{W}\\ \mu\end{bmatrix} = \begin{bmatrix} Var_{x_i} & \mathbf{1}\\ \mathbf{1}^T & 0\end{bmatrix}^{-1}\cdot\begin{bmatrix}Cov_{x_i x0}\\ 1\end{bmatrix} = \begin{bmatrix}\gamma(x_1,x_1) & \cdots & \gamma(x_1,x_n) & 1\\ \vdots & \ddots & \vdots & \vdots\\ \gamma(x_n,x_1) & \cdots & \gamma(x_n,x_n) & 1\\ 1 & \cdots & 1 & 0\end{bmatrix}^{-1}\begin{bmatrix}\gamma(x_1,x^*)\\ \vdots\\ \gamma(x_n,x^*)\\ 1\end{bmatrix}$$

Kriging weights

Modeled relation between sampling points and location(s) where we want to estimate values.

$$\hat{Z}(x_0) = \hat{W}^T\cdot\left[Z(x_1)\quad\cdots\quad Z(x_N)\right]^T \quad ; \quad var\left(\hat{Z}(x_0) - Z(x_0)\right) = \hat{W}^T\cdot\left[\gamma(x_1,x_0)\quad\cdots\quad\gamma(x_N,x_0)\quad 1\right]^T$$



- Takes into account distance, spatial structure of the data and redundancy of the information supplied by the sampling points.
- At distances greater than the range, weights = 0

on and kriging

Kriging types

Simple kriging = One global system. Use all points
Ordinary kriging = with local neighbourhood.
Indicator kriging = binary values
Universal kriging, kriging with a trend = Non-stationarity, trend surface
Kriging with an external drift = Drift specified by external variable.
Block kriging = Estimation on a support larger than the distance between sampling points (multiple points within support)
Co-krigeage = Two or more values together.
Factorial kriging = Multivariate, multiscale analysis.

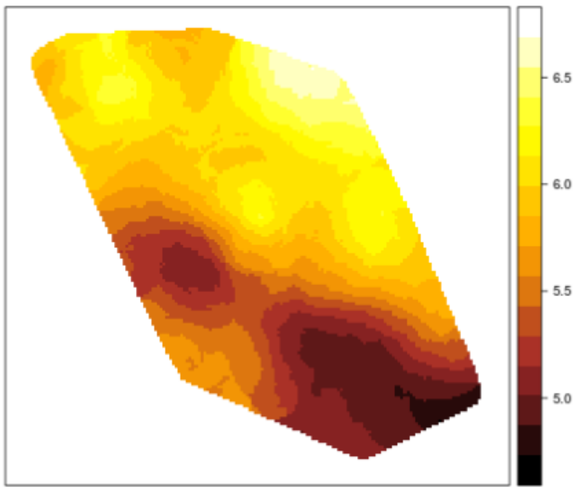
Kriging

Smoothing effect

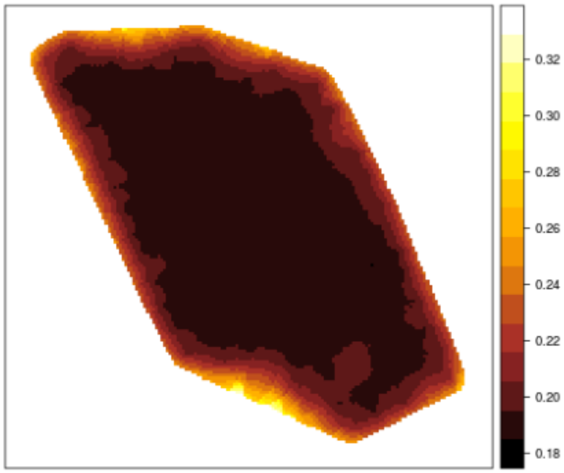
- The more pronounced the nugget, the greater the smoothing effect.
- Strong smoothing effect: map is not a good representation of the variable in reality.

Variables without a 'random' aspect: better use Spline, TIN, or other deterministic functions.

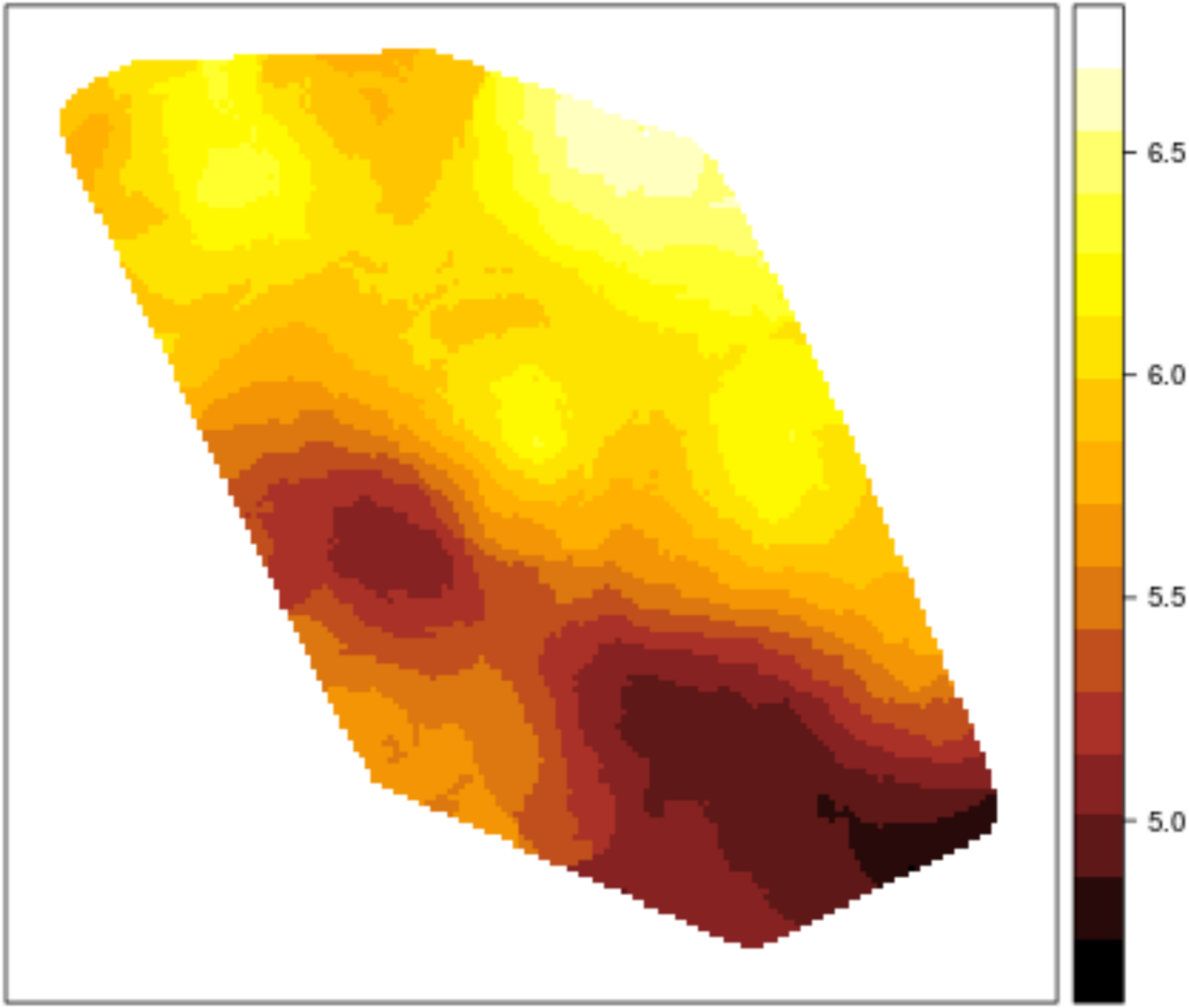
```
# Simple Kriging
kri <- krige(pH ~ 1, ArboSP, Arbo_mask, model = model,beta=1)
spplot(kri['var1.pred'])

# Ordinary Kriging
kri <- krige(pH ~ 1, ArboSP, Arbo_mask, model = model,maxdist=100)
spplot(kri['var1.pred'])
spplot(kri['var1.var'])
```



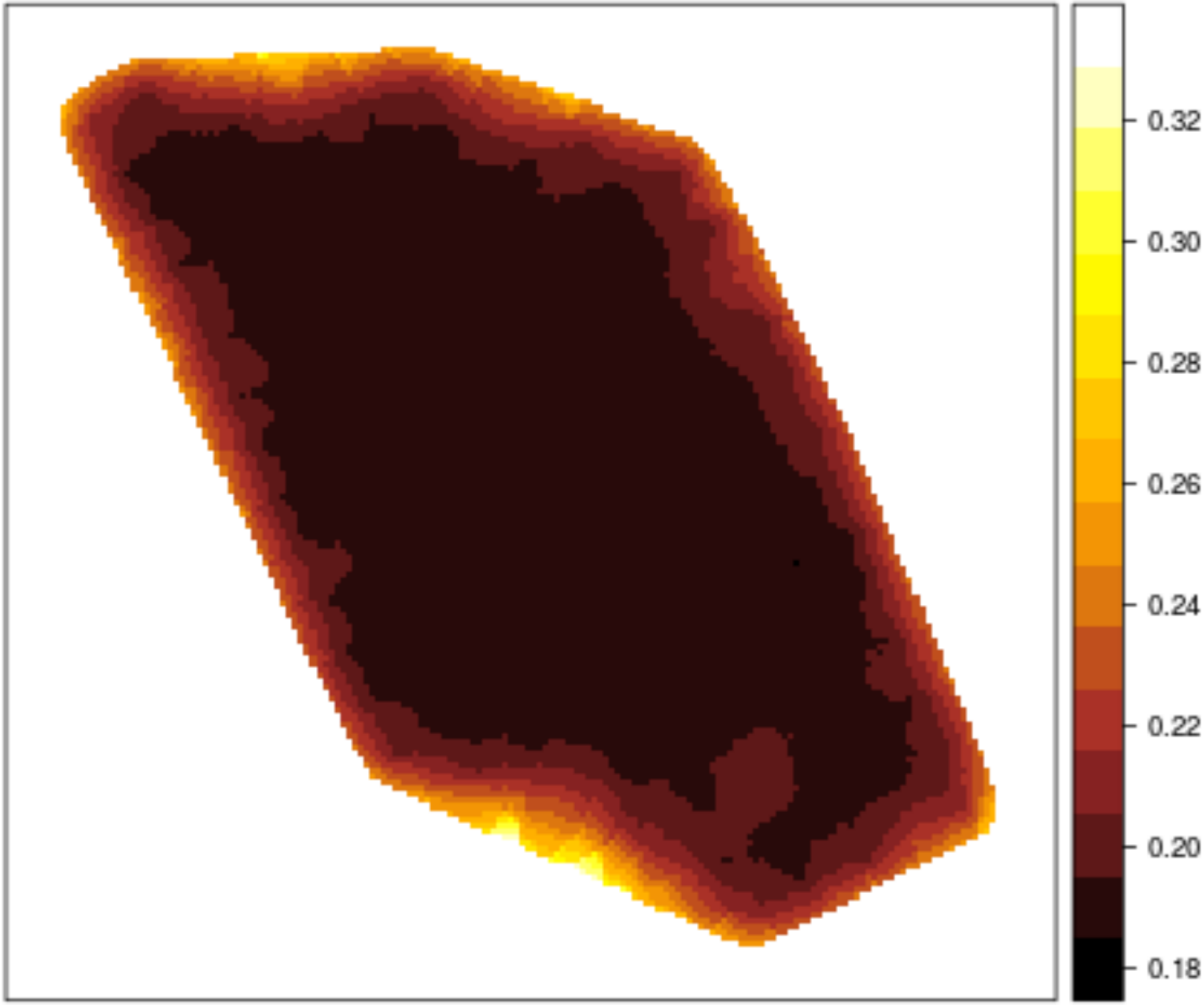
Estimation



Variance



Estimation



Variance

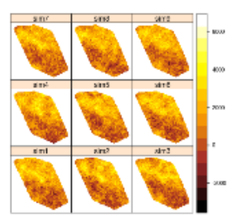
Uncertainty assessment

- Look at kriging variance map.
- Look at variograms.
- Cross-validation.
- Separate dataset into estimation and validation subsets.

Conditional simulations

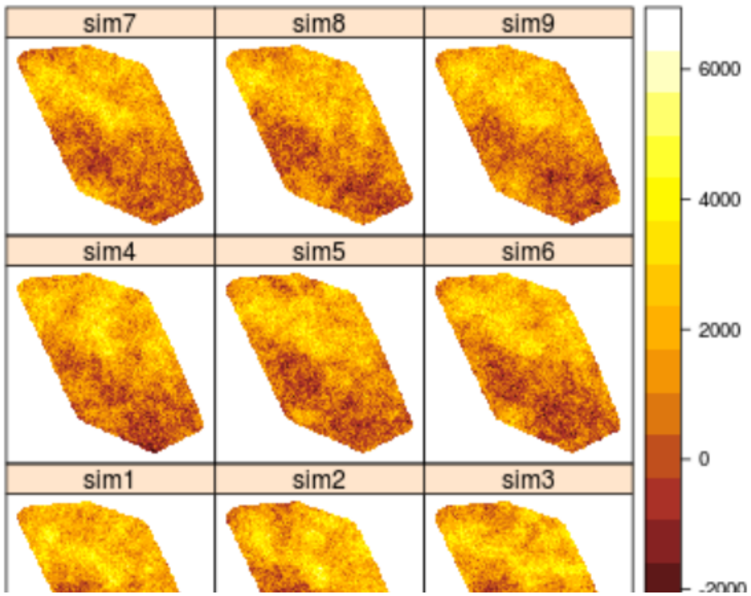
- Spatial mean estimated by kriging.
- Uncertain/stochastic aspect simulated.
- Offers a realistic representation of the variable.
- Gaussian simulations - simulated portion follows a normal distribution.

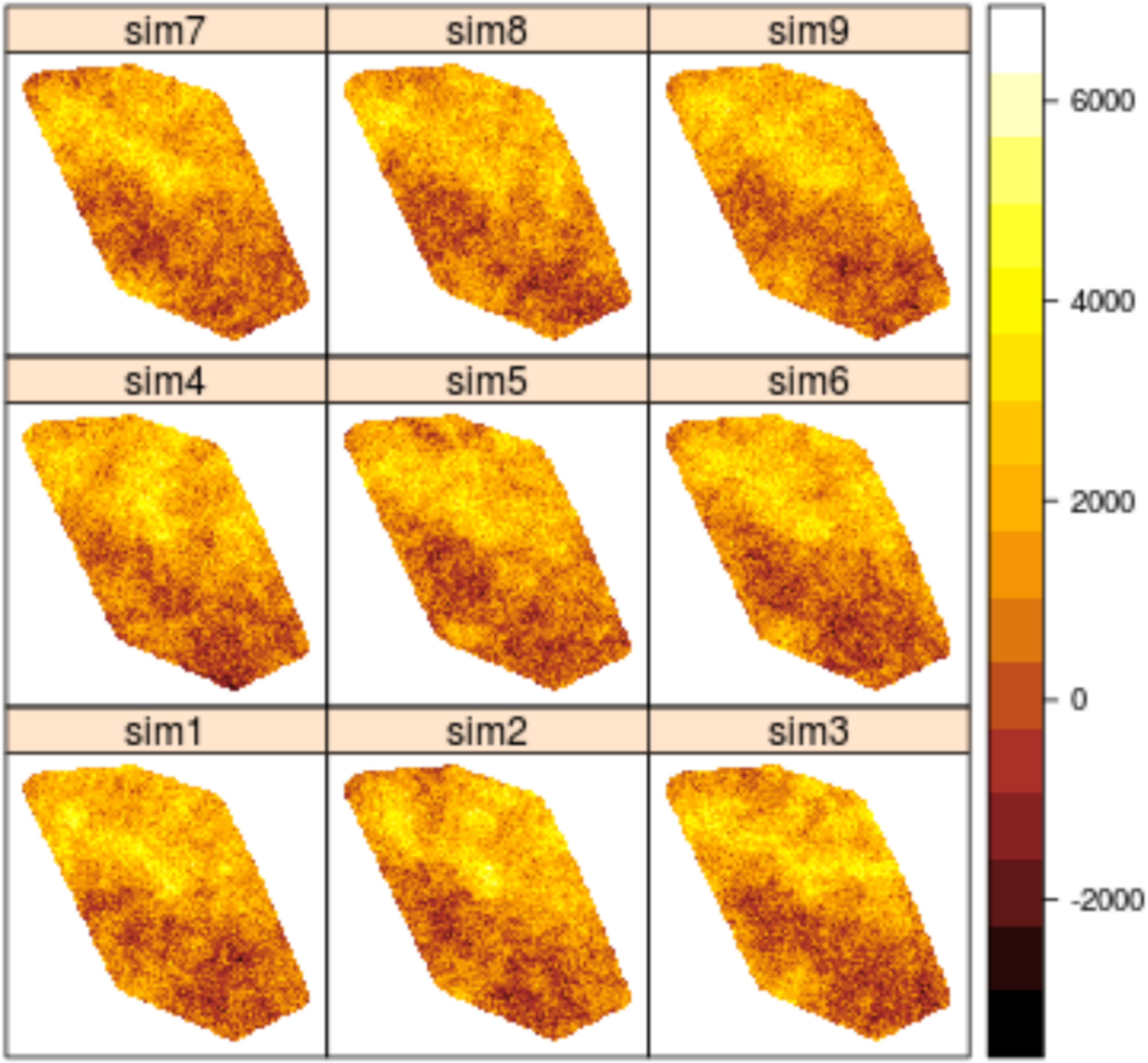
```
Ca.sim <- krige(formula = Ca~1, ArboSP, Arbo_mask, model = Ca.model,
                nmax = 15, nsim = 9)
```



llows a normal distrib

```
Ca.sim <- krige(formula = Ca~1, ArboSP, Arbo_mask, model = Ca.model,
                nmax = 15, nsim = 9)
```

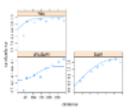


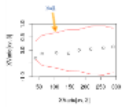

sim7
sim8
sim9
sim4
sim5
sim6
sim1
sim2
sim3
6000
4000
2000
0
-2000

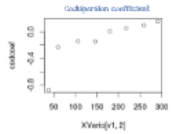
Bi-/multi-variate analyses

Cross-variograms

- Spatial relationships between variables. How it changes as a function of 'scale'.
- Hulls - Perfect correlation, given the spatial structure of each variable.
- Co-dispersion coefficients: between -1 and 1 for each lag.







Modified tests

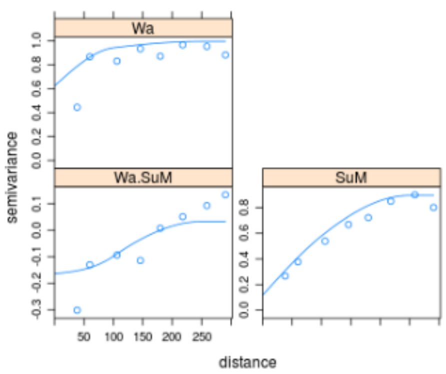
- Hypotheses of independence (IID) for t or F tests non-valid in the presence of spatial structure/autocorrelation.
- Affects the probability (p), not the correlation.
- Take spatial autocorrelation into account in the test (e.g. Dutilleul's modified t-test).
  - Calculates an effective sample size.
  - Uses modeled variogram or Moran's I autocorrelogram.

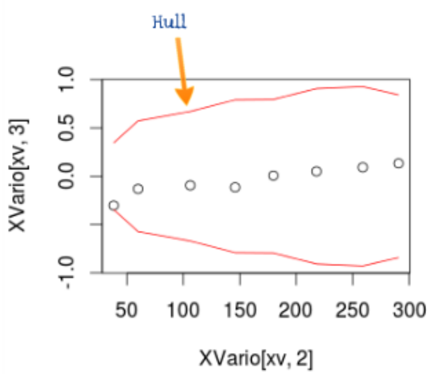
Multivariate analyses

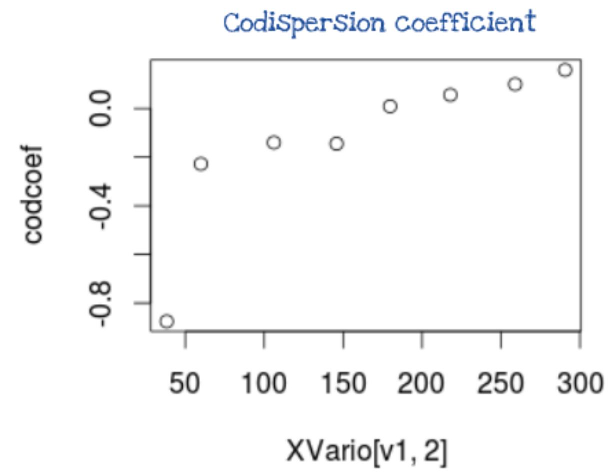
- Multi-scale ordination - Ordination at each lag (Wagner).
- Coregionalization analysis - Modeling of all direct and cross-variograms. Analysis of sill matrices as variance-covariance matrix. Multivariate analysis on each matrix.
- Coregionalization analysis with a drift - deals with large scales structure with 'deterministic' methods. Coregionalization analysis on residuals.

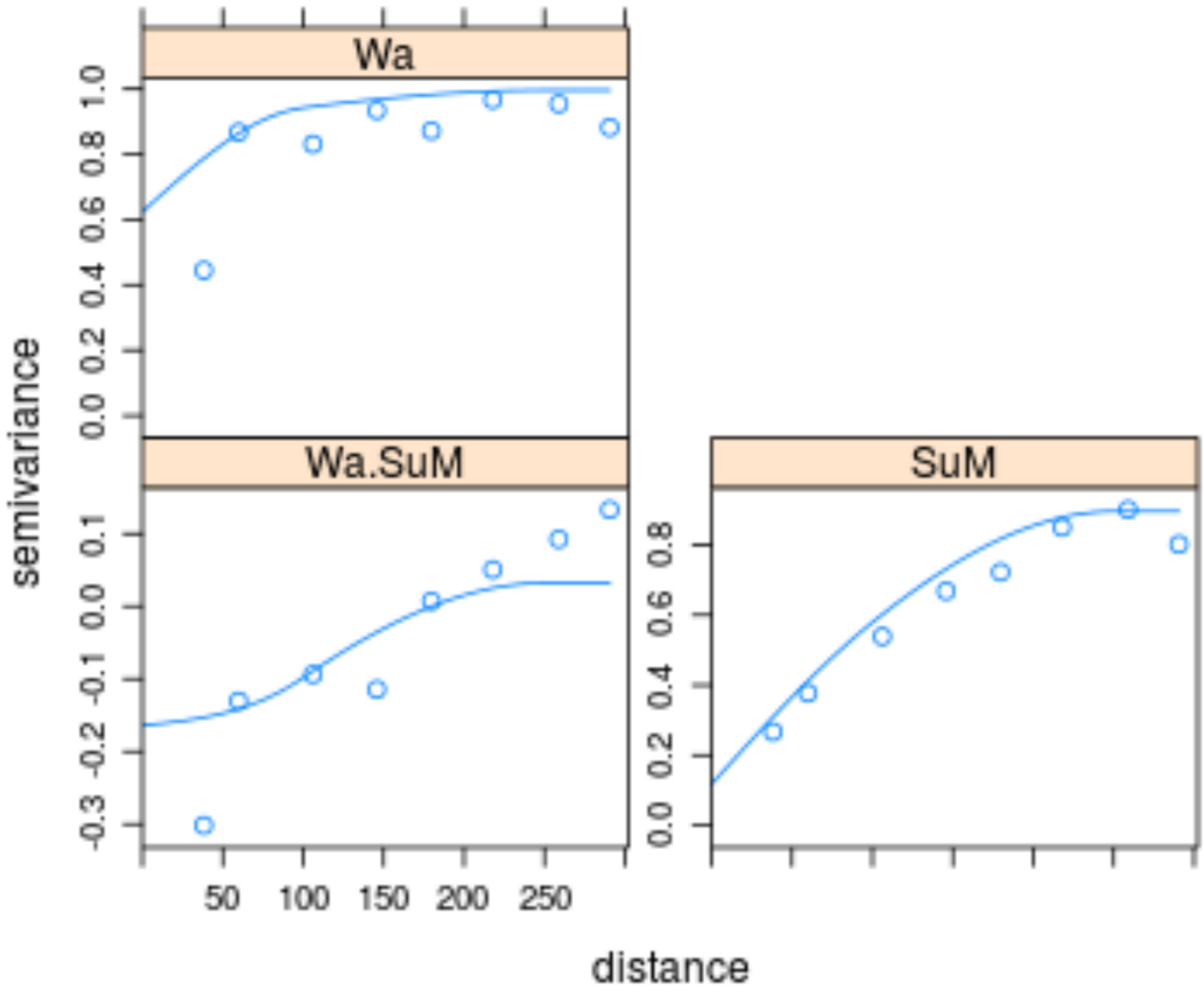
Cross-variograms

- Spatial relationships between variables. How it changes as a function of 'scale'.
- Hulls - Perfect correlation, given the spatial structure of each variable.
- Co-dispersion coefficients: between -1 and 1 for each lag.

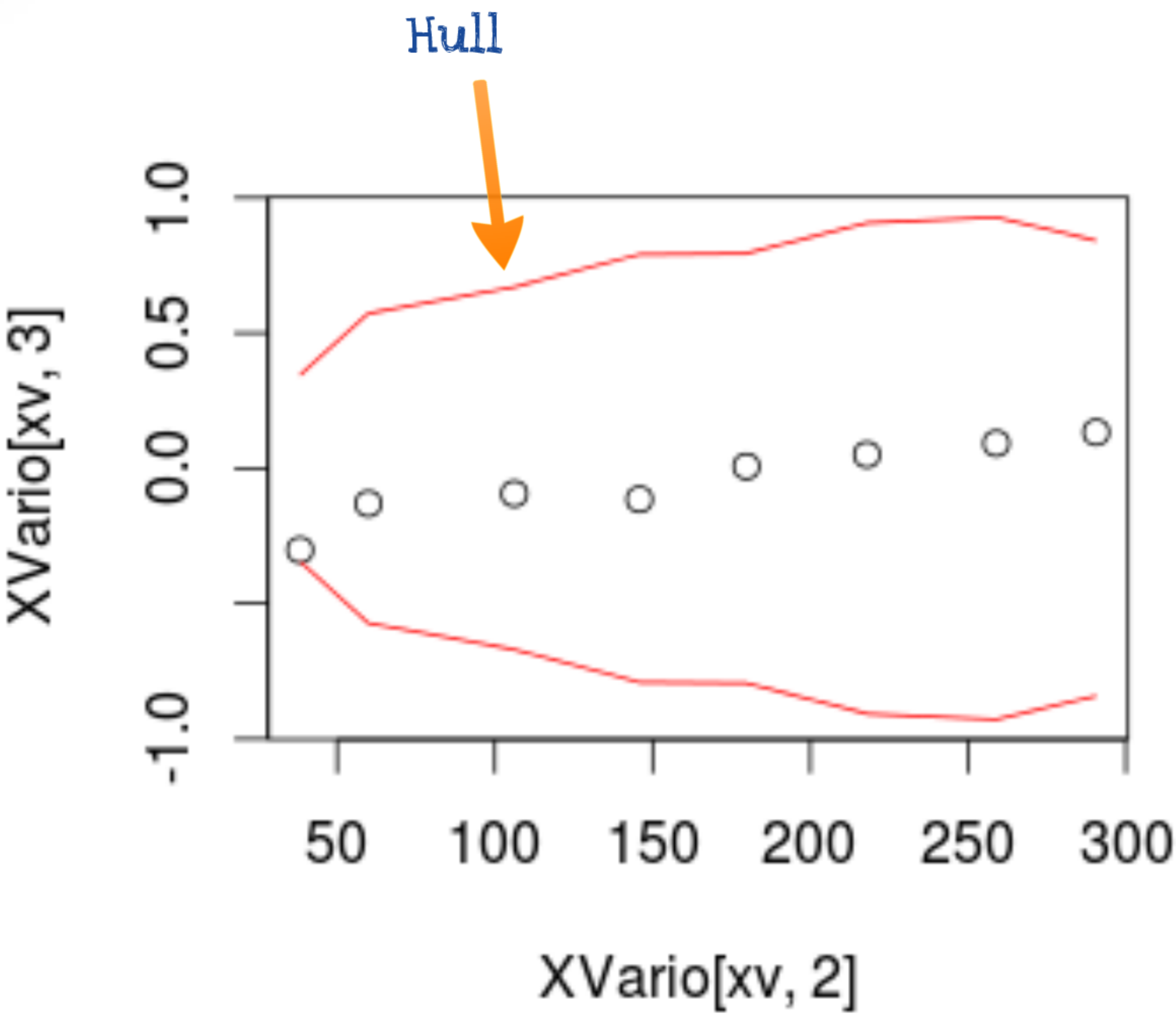




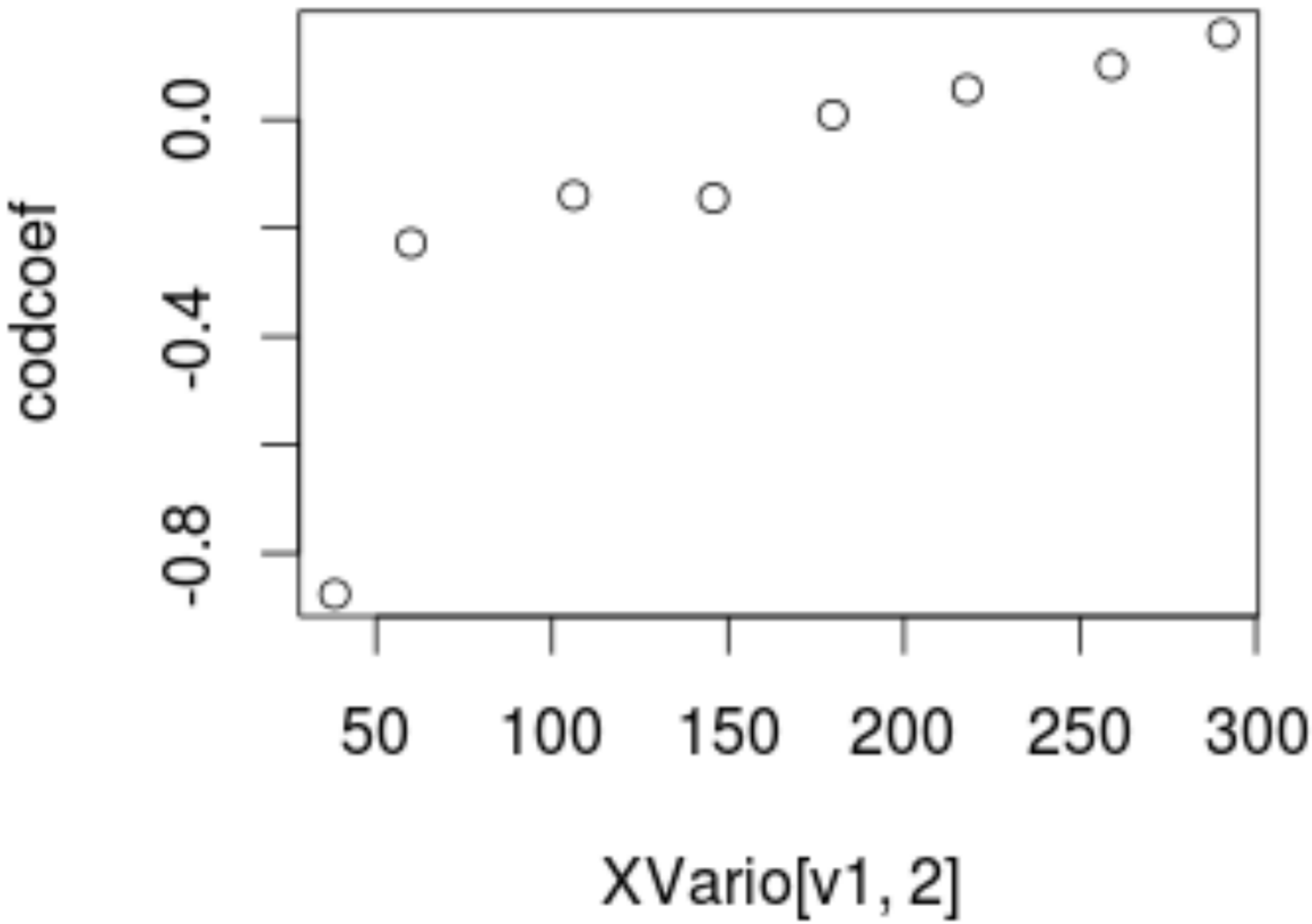




Wa
Wa.SuM
SuM
semivariance
distance
1.0
0.8
0.6
0.4
0.2
0.0
0.1
0.0
-0.1
-0.2
-0.3
0.8
0.6
0.4
0.2
0.0
50
100
150
200
250


Hull
1.0
0.5
0.0
-1.0
XVario[xv, 3]
50
100
150
200
250
300
XVario[xv, 2]

Codispersion coefficient

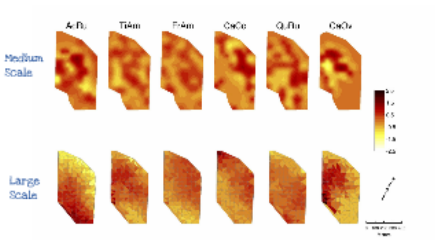


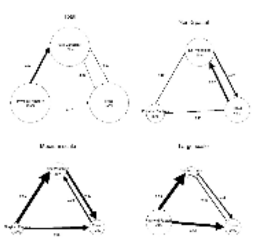
Modified tests

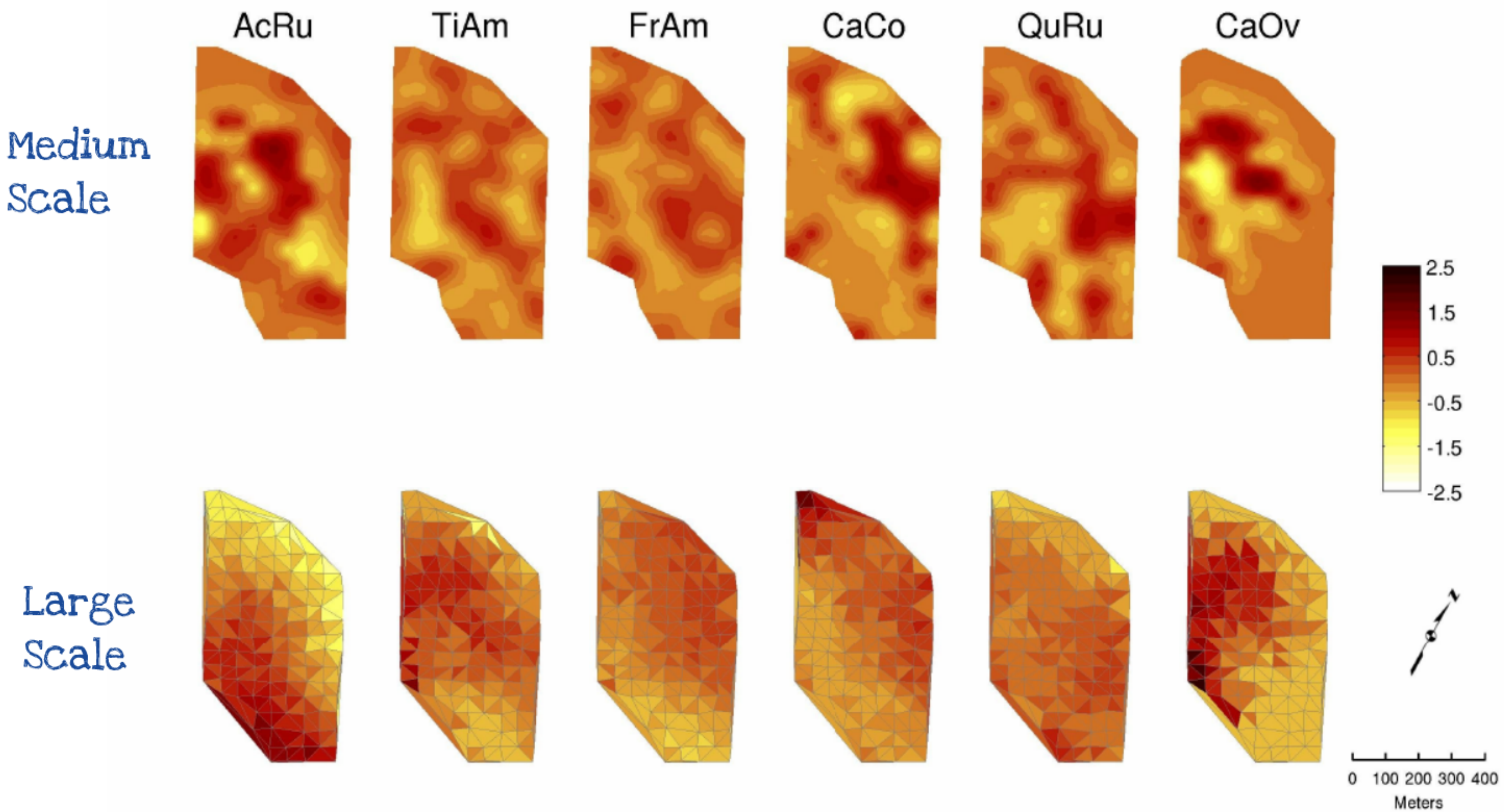
- Hypotheses of independence (IID) for t or F tests non-valid in the presence of spatial structure/autocorrelation.
- Affects the probability (p), not the correlation.
- Take spatial autocorrelation into account in the test (e.g. Dutilleul's modified t-test).
  - Calculates an effective sample size.
  - Uses modeled variogram or Moran's I autocorrelogram.

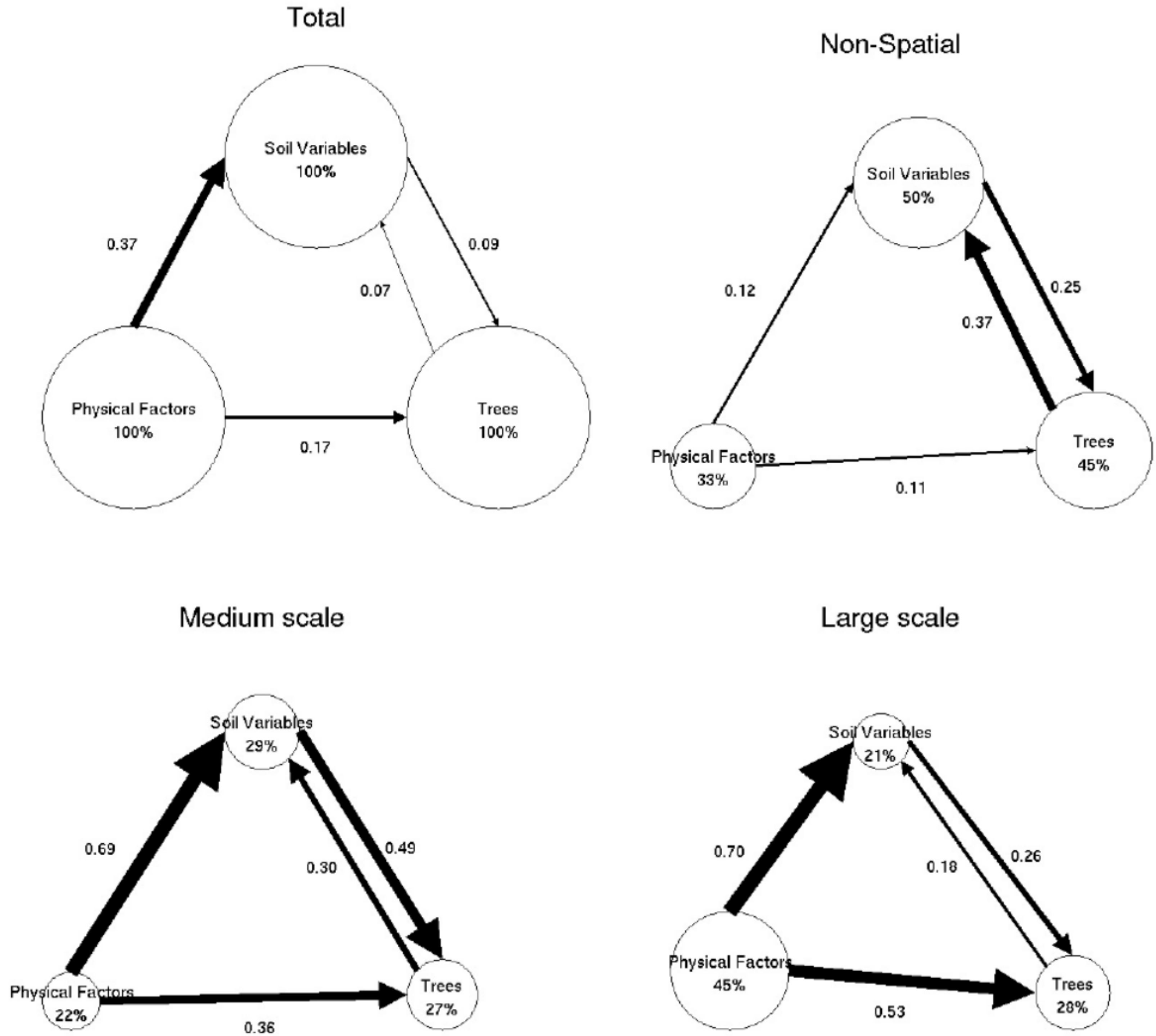
Multivariate analyses

- Multi-scale ordination - Ordination at each lag (Wagner).
- Coregionalization analysis - Modeling of all direct and cross-variograms. Analysis of sill matrices as variance-covariance matrix. Multivariate analysis on each matrix.
- Coregionalization analysis with a drift - deals with large scales structure with 'deterministic' methods. Coregionalization analysis on residuals.






AcRu
TiAm
FrAm
CaCo
QuRu
CaOv
Medium Scale
Large Scale
2.5
1.5
0.5
-0.5
-1.5
-2.5
N
0 100 200 300 400
Meters


Total
Soil Variables
100%
Physical Factors
100%
Trees
100%
0.37
0.09
0.07
0.17
Non-Spatial
Soil Variables
50%
Physical Factors
33%
Trees
45%
0.12
0.25
0.37
0.11
Medium scale
Soil Variables
29%
Physical Factors
22%
Trees
27%
0.69
0.49
0.30
0.36
Large scale
Soil Variables
21%
Physical Factors
45%
Trees
28%
0.70
0.26
0.18
0.53