

OBJETIVO DE LA PRÁCTICA: Introducir conceptos básicos de la geoestadística.

English

## EJEMPLO 1. Interpolación ponderada de distancia inversa, IDW

Suponer que se tiene un problema de estimación local de un valor en el punto A de la figura 1.1, (Clark, 2001).

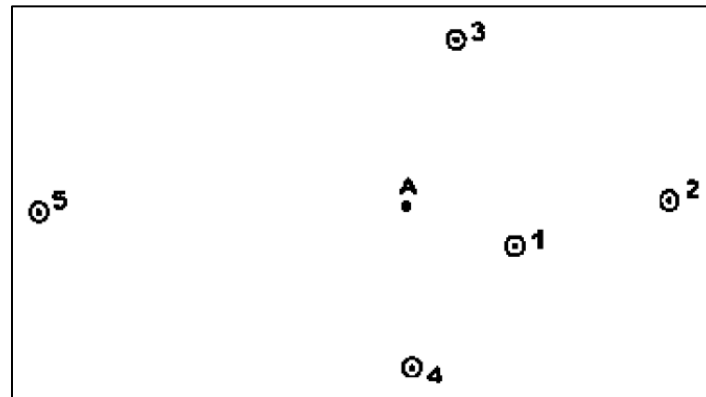


Figura 1.1. Situación hipotética de muestreo y estimación, (Clark, 2001)

Para resolver este ejemplo aplicando el método no geoestadístico IDW, se proponen unos esquemas como los mostrados en las figuras 1.2 y 1.3.

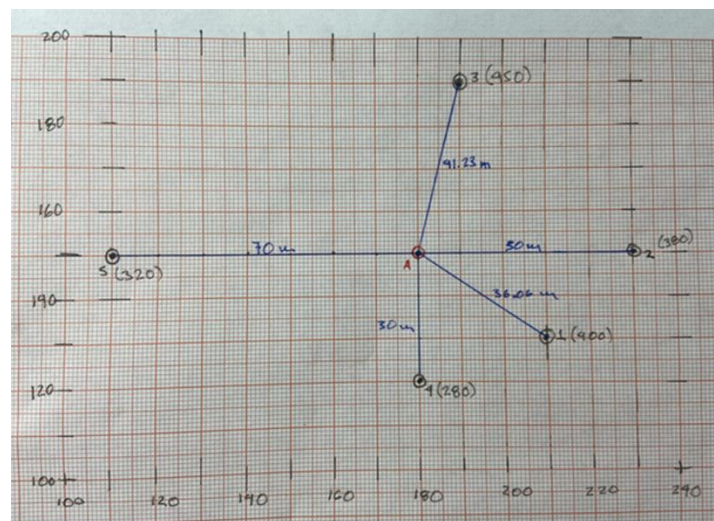


Figura 1.2. Esquema dimensionado de la figura 1.1, en m

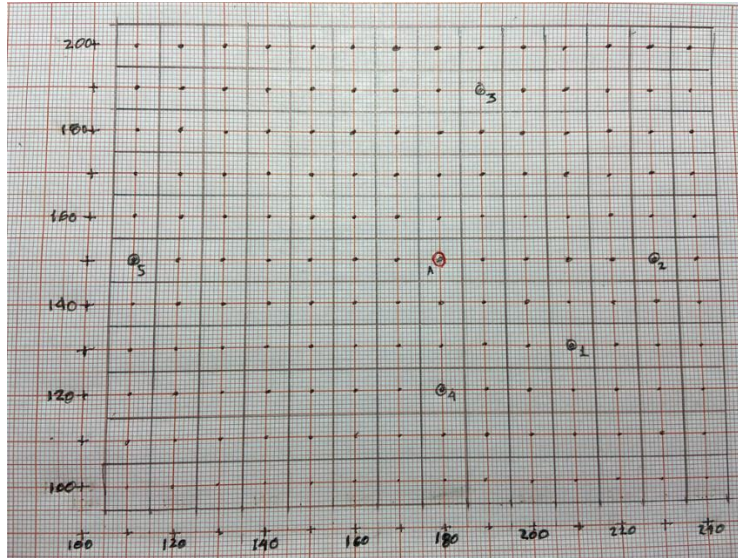


Figura 1.3. Esquema de la malla para la figura 1.1.

## Interpolación ponderada de distancia inversa, IDW. (Bivand, 2013)

La interpolación ponderada basada en la distancia inversa (IDW) calcula un promedio ponderado

$$\hat{Z}(s_0) = \frac{\sum_{i=1}^n w(s_i) Z(s_i)}{\sum_{i=1}^n w(s_i)},$$

donde los pesos de las observaciones se calculan de acuerdo con su distancia a la ubicación de interpolación,

$$w(s_i) = ||s_i - s_0||^{-p},$$

con  $|| \cdot ||$  indica la distancia euclidiana y  $p$  una potencia de ponderación de la distancia inversa, con un valor predeterminado de 2. Si  $s_0$  coincide con una ubicación de observación, el valor observado se devuelve para evitar pesos infinitos.

La potencia de la distancia inversa determina el grado en que se prefieren los puntos más cercanos a los puntos más distantes; para valores grandes, IDW converge a la interpolación de un vecino más cercano. Se puede ajustar, por ejemplo, mediante validación cruzada. IDW también se puede utilizar dentro de las vecindades de búsqueda locales.

Entonces,

$p = 1$ ;

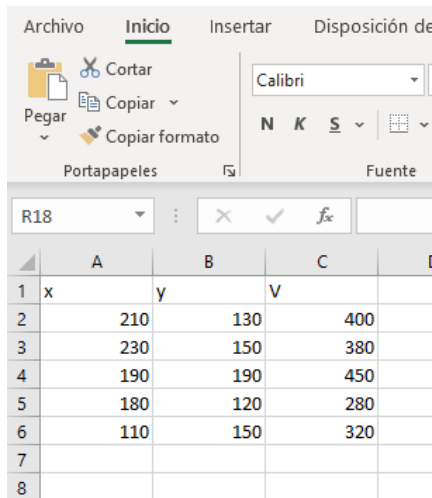
$$\bar{A} = \frac{\left(\frac{1}{36.06} \cdot 400\right) + \left(\frac{1}{50} \cdot 380\right) + \left(\frac{1}{41.23} \cdot 450\right) + \left(\frac{1}{30} \cdot 280\right) + \left(\frac{1}{70} \cdot 3200\right)}{36.06^{-1} + 50^{-1} + 41.23^{-1} + 30^{-1} + 70^{-1}} = 363.80$$

$p = 2$ ;

$$\bar{A} = \frac{\left(\frac{1}{36.06} \cdot 400\right) + \left(\frac{1}{50} \cdot 380\right) + \left(\frac{1}{41.23} \cdot 450\right) + \left(\frac{1}{30} \cdot 280\right) + \left(\frac{1}{70} \cdot 3200\right)}{36.06^{-2} + 50^{-2} + 41.23^{-2} + 30^{-2} + 70^{-2}} = 358.26$$

## Resolviendo con R

Ejemplo\_0.csv



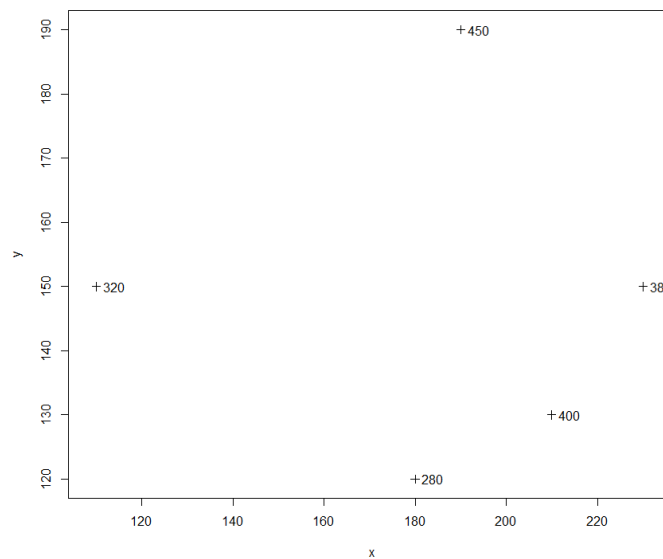
	A	B	C
1	x	y	V
2	210	130	400
3	230	150	380
4	190	190	450
5	180	120	280
6	110	150	320
7			
8			

Malla\_Ejemplo\_0.csv

x	y
110	100
110	110
110	120
110	130
110	140
110	150
110	160
110	170
110	180
110	190
120	100
120	110
120	120
120	130

230	140
230	150
230	160
230	170
230	180
230	190
240	100
240	110
240	120
240	130
240	140
240	150
240	160
240	170
240	180
240	190

```
library(sp)
library(gstat)
setwd("C:/Geoestadística/Mi carpeta")
# CARGA LOS DATOS
#
datos <- read.csv("Ejemplo_0.csv", header=TRUE, sep=";", dec=".")
class(datos)
str(datos)
attach(datos)
coordinates(datos) <- c("x", "y") # crea un objeto espacial
# Ver los datos V
plot(x,y,xlim=c(min(x)*0.99,max(x)*1.01),ylim=c(min(y)*.999,max(y)*1.001),pch=3)
text(x+3,y,V)
```



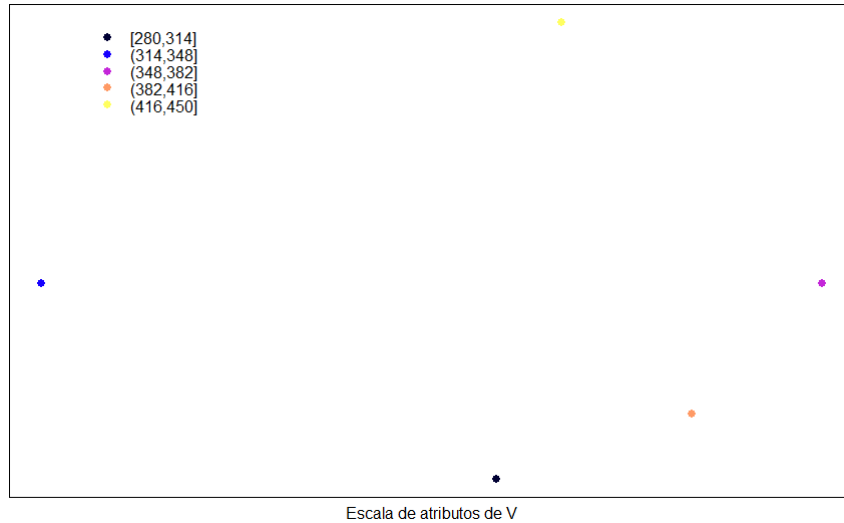
```
# Rechazar la hipótesis nula si p <= alfa=0.05
# H0: la muestra procede de una población normalmente distribuida
shapiro.test(V) # *** Hay normalidad ***
qqnorm(V)
qqline(V)
```

```
> # Rechazar la hipótesis nula si p <= alfa=0.05
> # H0: la muestra procede de una población normalmente distribuida
> shapiro.test(V) # *** Hay normalidad ***

Shapiro-Wilk normality test

data:  V
W = 0.97671, p-value = 0.9163
```

```
# grafica métodos para datos espaciales con atributos
# en caso de no normalidad: spplot(datos, "V", do.log = T, colorkey = TRUE)
spplot(datos, "V", xlab="Escala de atributos de V", colorkey = F,
       key.space=list(x=0.1,y=.95))
```



#### # Lectura de la malla

```
datosred <- read.csv("Malla_Ejemplo_0.csv", header=TRUE, sep=",", dec=".")
class(datosred)
str(datosred)
attach(datosred)
coordinates(datosred) <- c("x", "y")
#datosred <- as(datosred, "SpatialPixelsDataFrame")
datosred <- as(datosred, "SpatialPixels") # Corrección
```

```
idw.out <- idw(V~1, datos, datosred, idp=1)
as.data.frame(idw.out)[1:100,]
idwdata <- as.data.frame(idw.out)
write.csv(idw.out, file = "idw_res.csv") # Guarda en un archivo .csv
plot(idw.out, axes=T)
idw.out
class(idw.out)
spplot(idw.out, "var1.pred", do.log = F, colorkey = TRUE)
```

```
> idw.out <- idw(V~1, datos, datosred, idp=1)
[inverse distance weighted interpolation]
> as.data.frame(idw.out)[1:100,]
      x    y var1.pred var1.var
1  110 100  349.6394      NA
2  110 110  348.0423      NA
3    110 120  345.7471      NA
...
75  180 140  351.5289      NA
76  180 150  363.7968      NA
77  180 160  374.1768      NA
...
```

```
> idw.out <- idw(V~1,datos,datosred,idp=2)
[inverse distance weighted interpolation]
> as.data.frame(idw.out)[1:100,]
      x    y var1.pred var1.var
1   110  100  335.2222      NA
2   110  110  332.8073      NA
3   110  120  329.6830      NA
...
75  180  140  331.8761      NA
76  180  150  358.2615      NA
77  180  160  382.3909      NA
...
```

Visto en un plano, con  $p=2$ ,

```
plot(x,y,xlim=c(min(x)*0.99,max(x)*1.01),ylim=c(min(y)*.999,max(y)*1.001),pch=3)
text(x,y-2,round(idw.out$var1.pred,2),cex = 0.8)
```



## Conclusiones

Se confirma la semejanza entre los resultados del inciso anterior y los actuales.

## EJEMPLO 2. Introducción al Variograma

En la figura 2.1 se tiene un muestreo del valor promedio de la ley para el Fe (% en peso) en un conjunto de perforaciones. Se trata de obtener el semivariograma calculado o experimental en donde  $h$  depende de la distancia entre el par de muestras y de la orientación relativa en un plano bidimensional. (Clark, 2001).

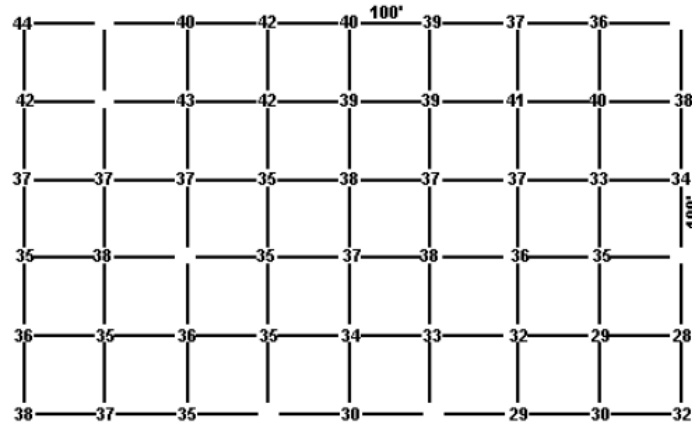


Figura 2.1 Ejemplo de datos en una malla para el cálculo de un semivariograma experimental - mineral de hierro, tomado de (Clark, 2001)

Entonces, para obtener el semivariograma, o variograma, se aplica la ecuación

$$2\gamma^*(h) = \frac{1}{n} \sum [g(x) - g(x+h)]^2$$

donde  $\gamma^*(h)$  se denomina precisamente semivariograma experimental.

Iniciando el cálculo de este a oeste, E-O, y con  $h = 100$  pies:

$$\begin{aligned} \gamma^*(100) = & [(40 - 42)^2 + (42 - 40)^2 + (40 - 39)^2 + (39 - 37)^2 + (37 - 36)^2 + (43 - 42)^2 \\ & + (42 - 39)^2 + (39 - 39)^2 + (39 - 41)^2 + (41 - 40)^2 + (40 - 38)^2 + (37 - 37)^2 \\ & + (37 - 37)^2 + (37 - 35)^2 + (35 - 38)^2 + (38 - 37)^2 + (37 - 37)^2 + (37 - 33)^2 \\ & + (33 - 34)^2 + (35 - 38)^2 + (35 - 37)^2 + (37 - 38)^2 + (38 - 36)^2 + (36 - 35)^2 \\ & + (36 - 35)^2 + (35 - 36)^2 + (36 - 35)^2 + (35 - 34)^2 + (34 - 33)^2 + (33 - 32)^2 \\ & + (32 - 29)^2 + (29 - 28)^2 + (38 - 37)^2 + (37 - 35)^2 + (29 - 30)^2 + (30 - 32)^2] \div (2 \\ & + 36) = 1.42 (\%)^2 \end{aligned}$$

Esto proporciona un punto que podemos trazar en una gráfica del semivariograma experimental ( $\gamma^*$ ) versus la distancia entre las muestras ( $h$ ), es decir  $[100 \text{ pies}, 1.42 (\%)^2]$ . En la figura 2.2 se esquematiza el procedimiento de cálculo.

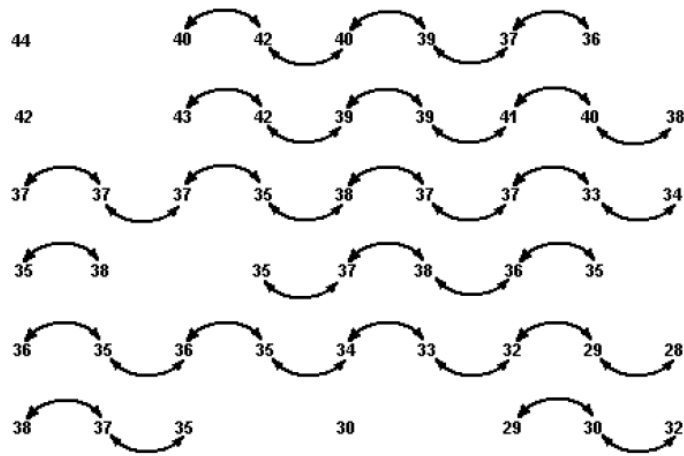


Figura 2.2. Identificación de todos los pares de puntos de muestreo a 100 pies de distancia, en dirección este-oeste, E-O. (Clark, 2001)

Ahora con  $h = 200$  pies,

$$\begin{aligned} \gamma^*(200) = & [(44 - 40)^2 + (40 - 40)^2 + (42 - 39)^2 + (40 - 37)^2 + (39 - 36)^2 + (42 - 43)^2 \\ & + (43 - 39)^2 + (42 - 39)^2 + (39 - 41)^2 + (39 - 40)^2 + (41 - 38)^2 + (37 - 37)^2 \\ & + (37 - 35)^2 + (37 - 38)^2 + (35 - 37)^2 + (38 - 37)^2 + (37 - 33)^2 + (37 - 34)^2 \\ & + (38 - 35)^2 + (35 - 38)^2 + (37 - 36)^2 + (36 - 35)^2 + (36 - 36)^2 + (35 - 35)^2 \\ & + (36 - 34)^2 + (35 - 33)^2 + (34 - 32)^2 + (33 - 29)^2 + (32 - 28)^2 + (38 - 35)^2 \\ & + (35 - 30)^2 + (30 - 29)^2 + (29 - 32)^2] \div (2 + 33) = 3.55 (\%)^2 \end{aligned}$$

En la figura 2.3 se muestra el procedimiento de cálculo esquematizado.

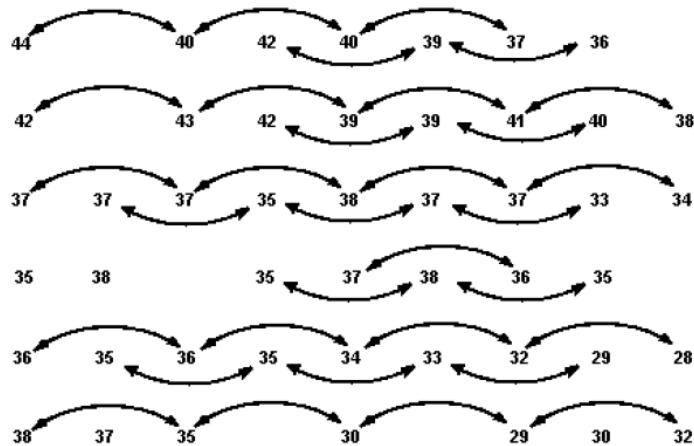


Figura 2.3. Identificación de todos los pares de puntos de muestreo a 200 pies de distancia, en dirección este-oeste, E-O. (Clark, 2001)



En la tabla 2.1 se resumen los resultados obtenidos para las dos direcciones principales, considerando distancias hasta la mitad de la extensión máxima en ambas direcciones y en la figura 2.4 la gráfica de los semivariogramas correspondientes.

Dirección	Distancia entre muestras (pies)	Semivariograma experimental	Número de pares
E-O	100	1.42	36
	200	3.55	33
	300	3.54	26
	400	6.61	23
N-S	100	5.26	35
	200	9.06	27
	300	16.21	21
	400	25.54	13

Tabla 2.1 Cálculo de valores de semivariograma experimental en dos direcciones principales para el ejemplo de mineral de hierro en cuadrícula cuadrada

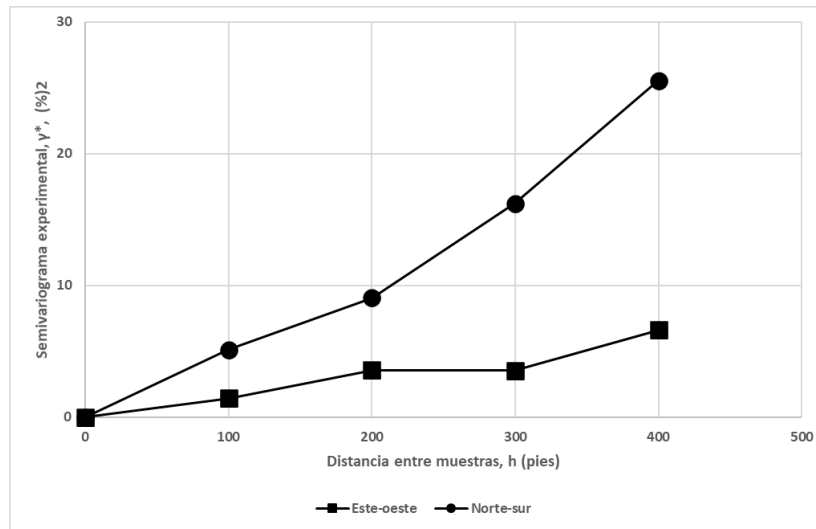


Figura 2.4 Semivariogramas experimentales en las dos direcciones principales para el ejemplo del mineral de hierro.

## Conclusiones

Como se aprecia, parece haber una diferencia distinta en la estructura en las dos direcciones (anisotropía). El semivariograma N-S aumenta mucho más bruscamente que el E-O, lo que sugiere una mayor continuidad en la dirección E-O. Esto se puede verificar calculando el semivariograma en direcciones diagonales, por ejemplo Noroeste-sureste, NO-SE, y Noreste-suroeste, NE-SO. Para ello la  $h$  tomará valores de 141.42, 282.84 y 424.26 pies. En la tabla 2.2 se muestran los resultados y en la figura 2.5 se aprecia el comportamiento de todos los semivariogramas calculados E-O, N-S y diagonales NO-SE y NE-SE.

El  $\gamma^*$  en ambas direcciones diagonales parece confirmar la diferencia entre los otros dos principales, ya que se encuentran entre ellos; aunque parecen estar más cerca la N-S de la NO-SE y la E-O de la NE-SO. Además, considerando que entre menos pares se tengan, menos confiable son los cálculos, habría que evaluar la confiabilidad de los semivariogramas experimentales en dirección de las diagonales ya que se calcularon con 14 pares solamente en 424 pies.

Dirección	Distancia entre muestras (pies)	Semivariograma experimental	Número de pares
Diagonal	141	7.31	32
Noroeste-Sureste	283	13.14	21
Diagonal	424	28.21	14
Diagonal	141	3.60	31
Noreste-Suroeste	283	5.43	22
Diagonal	424	7.64	14

Tabla 2.2. Cálculo del semivariograma en dirección diagonal para el mineral hierro.

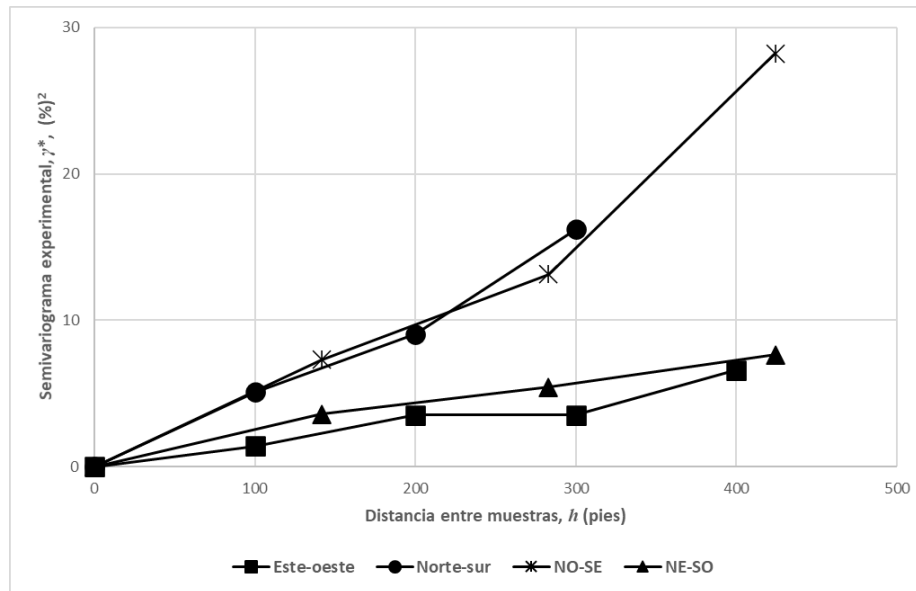


Figura 2.5. Semivariogramas experimentales incluyendo en dirección diagonal para el mineral hierro.

También, si se ajusta una recta a los semivariogramas, la pendiente E-O sería  $0.0153 (\%)^2 / \text{pie}$ , es decir, aproximadamente  $\gamma(h) = 0.0153h (\%)^2$ . En dirección N-S sería  $0.0524 (\%)^2 / \text{pie}$ ,  $\gamma(h) = 0.0524h (\%)^2$ , en dirección NO-SE  $0.064 (\%)^2 / \text{pie}$ ,  $\gamma(h) = 0.064h (\%)^2$  y en dirección NE-SO  $0.0175 (\%)^2 / \text{pie}$ ,  $\gamma(h) = 0.0175h (\%)^2$ . Es decir, en dirección E-O "se espera" una diferencia al cuadrado de  $0.0153 (\%)^2$  por cada pie entre las muestras. Dicho de otra manera, se espera una diferencia en la ley de  $0.0153 = 0.1237\%$  Fe para dos muestras separadas por 1 pie, con una orientación relativa de E-O; en la dirección N-S, la cifra correspondiente es  $0.2289\%$  Fe. Para muestras separadas por 100 pies, se esperarían diferencias de  $1.237\%$  Fe (E-O) y  $2.289\%$  Fe (N-S) y así sucesivamente. Con estas observaciones se puede verificar que existe similitud del comportamiento de la ley del Fe en las direcciones N-S y NO-SE, así como en las direcciones E-O y NE-SO. Por lo tanto, se ha creado una imagen de las fluctuaciones de la ley dentro de

esta sección del depósito y se tiene un modelo bastante simple para describir las diferencias en la ley en las direcciones seleccionadas.

## Resolviendo con R

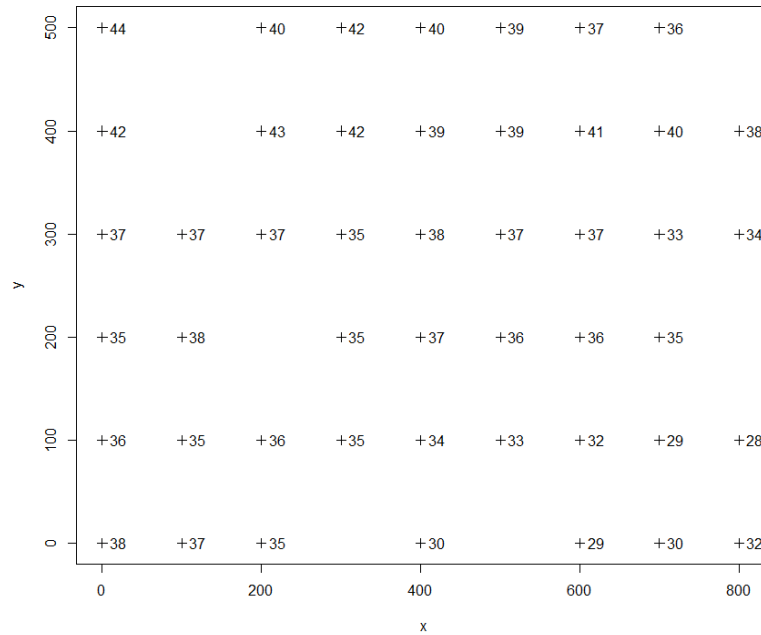
```
library(sp)
library(gstat)
setwd("C:/Geoestadística/Mi carpeta")
# CARGA LOS DATOS
#
datos <- read.csv("Ejemplo_11.csv", header=TRUE, sep=";", dec=".")
```

*Ejemplo\_11.csv*

x	y	z
0	500	44
0	400	42
0	300	37
0	200	35
0	100	36
0	0	38
100	300	37
100	200	38
100	100	35
100	0	37
200	500	40
200	400	43
200	300	37
200	100	36
200	0	35
300	500	42
300	400	42
300	300	35
300	200	35
300	100	35
400	500	40
400	400	39
400	300	38
400	200	37
400	100	34
400	0	30

500	500	39
500	400	39
500	300	37
500	200	36
500	100	33
600	500	37
600	400	41
600	300	37
600	200	36
600	100	32
600	0	29
700	500	36
700	400	40
700	300	33
700	200	35
700	100	29
700	0	30
800	400	38
800	300	34
800	100	28
800	0	32

```
#Ver los datos z
plot(x,y,xlim=c(min(x)*0.99,max(x)*1.01),ylim=c(min(y)*.999,max(y)*1.001),pch=3)
text(x+20,y,z)
```



### # EXPLORACIÓN DE LOS ATRIBUTOS

#

hist(z,breaks=10) # Distribución algo simétrica

summary(z) # Mediana y media similares

# H0: la muestra procede de una población normalmente distribuida

# Rechazar la hipótesis nula si  $p \leq \alpha=0.05$

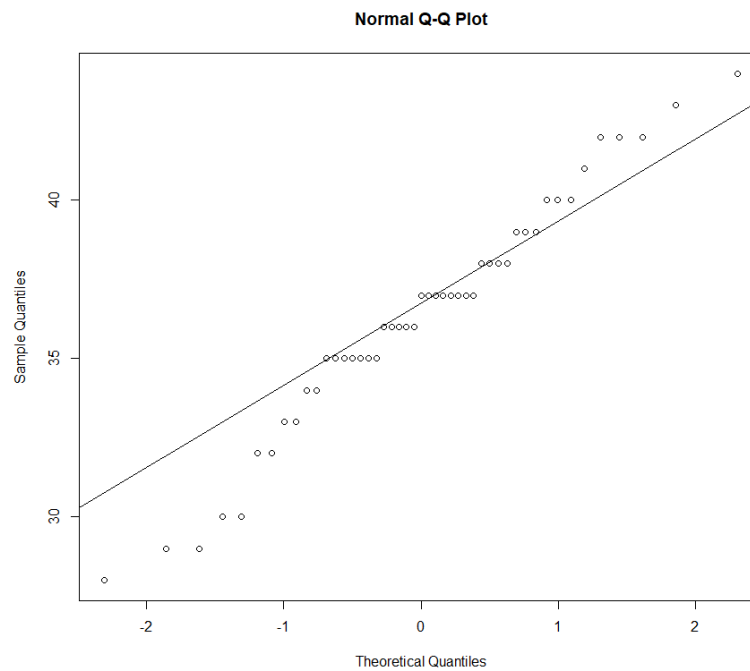
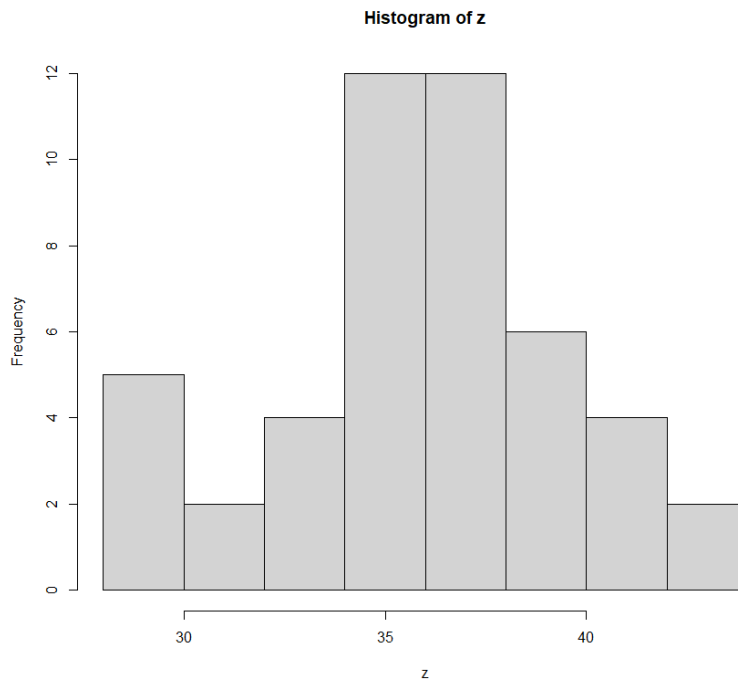
shapiro.test(z) # \*\*\* Hay normalidad \*\*\*

qqnorm(z)

```
> summary(z) # Mediana y media similares
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 28.00  35.00   37.00  36.34  38.50   44.00
> # H0: la muestra procede de una población normalmente distribuida
> # Rechazar la hipótesis nula si p <= alfa=0.05
> shapiro.test(z) # *** Hay normalidad ***
```

Shapiro-Wilk normality test

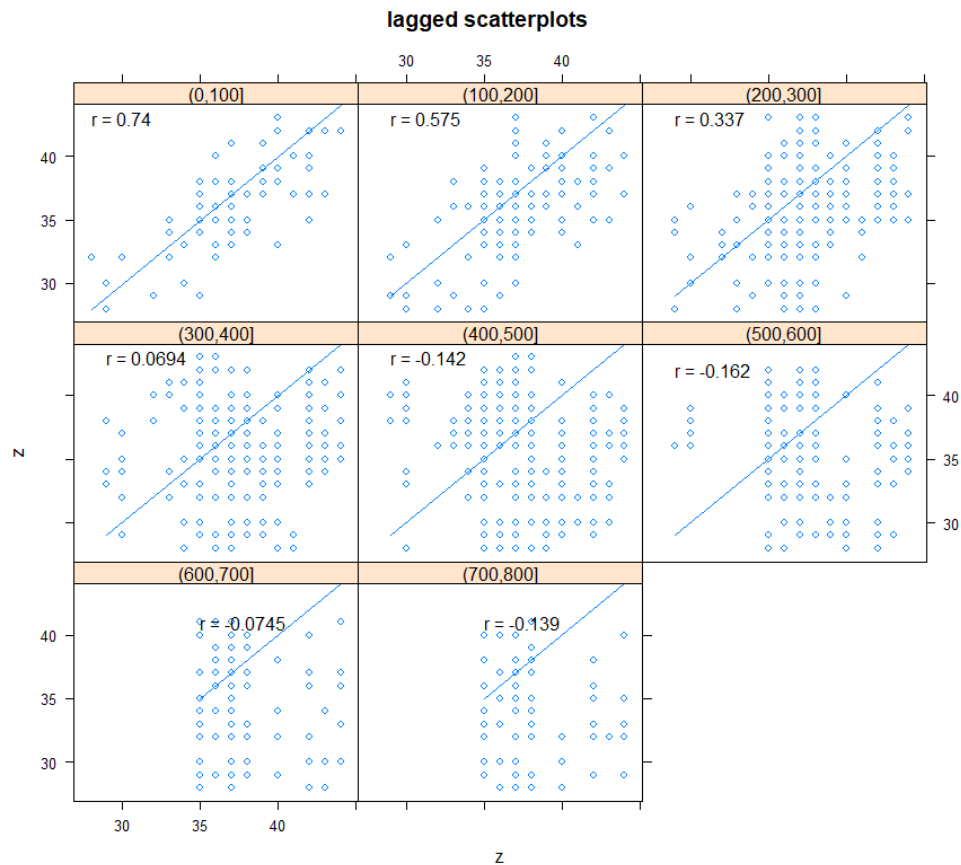
```
data: z
W = 0.97355, p-value = 0.3593
```



# Estima la correlación espacial: variograma

coordinates(datos) <- c("x", "y") # crear spatial data frame

hscat(z~1,datos,(0:8)\*100)qqline(z) # generalmente entre más pequeño es h, mejor es la correlación

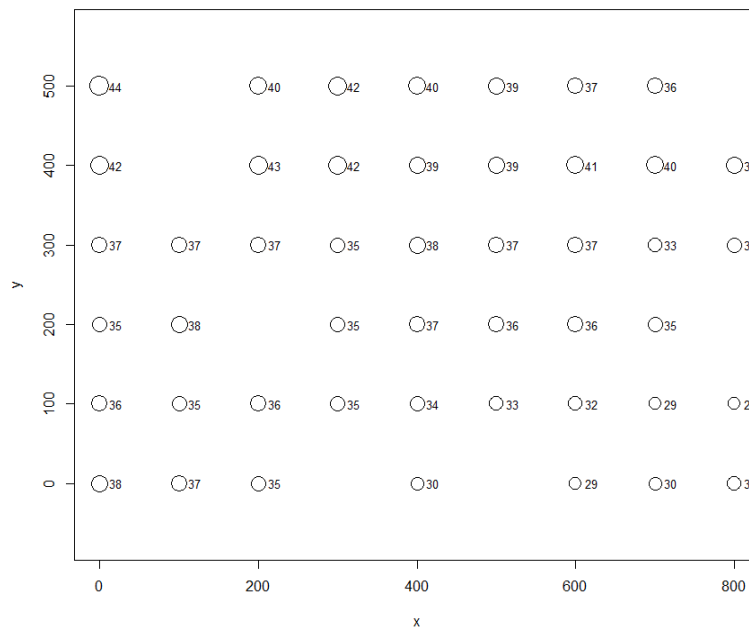


# Otra visualización con símbolos proporcionales.

# Gráfico de burbujas

plot(datos, asp=1, cex = 3\*z/max(z), pch=1, axes=T)

text(x+20,y,z, cex=0.8)



```
# calcula cuántos pares de puntos hay en datos.
```

```
n <- length(z)
```

```
n
```

```
n*(n+1)/2
```

```
[1] 47
```

```
> n*(n+1)/2
```

```
[1] 1128
```

```
# ESTRUCTURA ESPACIAL LOCAL: AUTOCORRELACIÓN
```

```
#
```

```
class(datos)
```

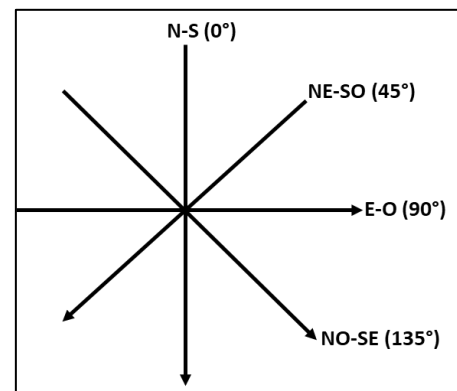
```
str(datos) # SLot @proj4string no tiene asignado el sistema de referencia.
```

```
> # ESTRUCTURA ESPACIAL LOCAL: AUTOCORRELACIÓN
> #
> class(datos)
[1] "SpatialPointsDataFrame"
attr(,"package")
[1] "sp"
> str(datos) # Slot @proj4string no tiene asignado el sistema de referencia.
Formal class 'SpatialPointsDataFrame' [package "sp"] with 5 slots
..@ data      : 'data.frame': 47 obs. of 1 variable:
.. ..$ z: int [1:47] 44 42 37 35 36 38 37 38 35 37 ...
..@ coords.nrs : int [1:2] 1 2
..@ coords     : num [1:47, 1:2] 0 0 0 0 0 0 100 100 100 100 ...
.. ..- attr(*, "dimnames")=List of 2
.. .. ..$ : NULL
.. .. ..$ : chr [1:2] "x" "y"
..@ bbox      : num [1:2, 1:2] 0 0 800 500
.. ..- attr(*, "dimnames")=List of 2
.. .. ..$ : chr [1:2] "x" "y"
.. .. ..$ : chr [1:2] "min" "max"
..@ proj4string: Formal class 'CRS' [package "sp"] with 1 slot
.. .. ..@ projargs: chr NA
```

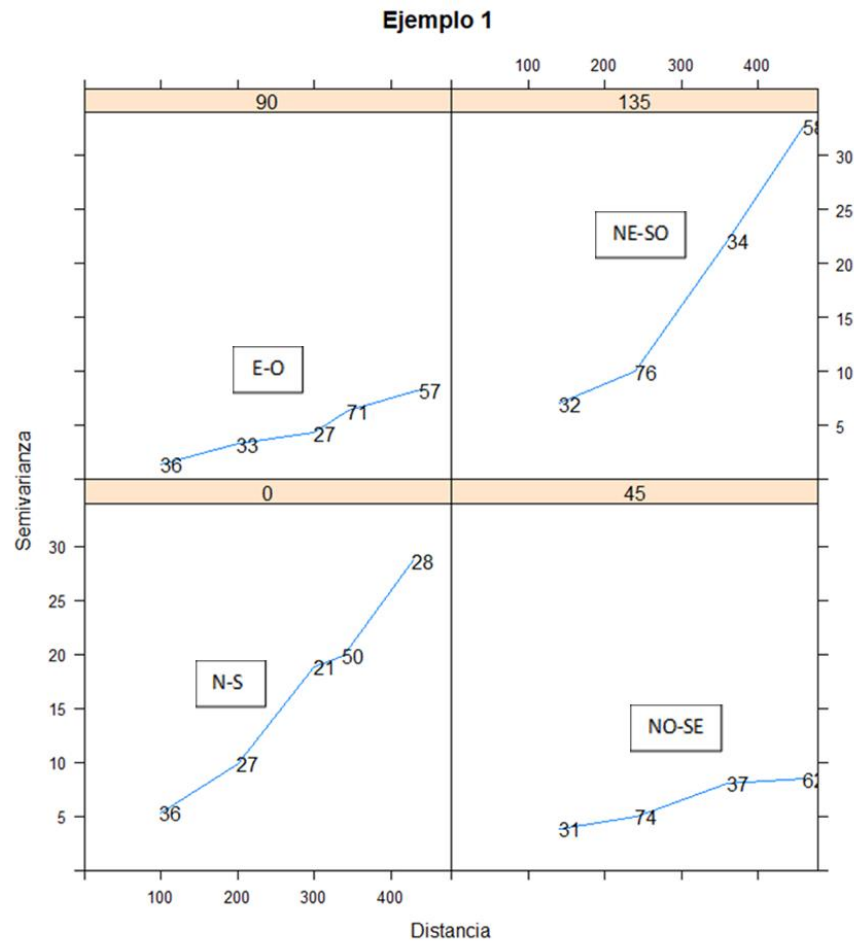
Para comparar con los resultados de la figura 2.5,

```
# VARIOGRAMA EMPÍRICO EXPERIMENTAL
#
ve <- variogram(z~1,datos,alpha=c(0,45,90,135),cutoff=500,width=100)
ve
plot(ve, plot.numbers = T, main = "Ejemplo 1", xlab = "Distancia",
     ylab = "Semivarianza", asp=1, type="l")
```

```
> ve
  np    dist    gamma dir.hor dir.ver  id
1  36 100.0000  5.347222      0      0 var1
2  27 200.0000  9.870370      0      0 var1
3  21 300.0000 18.880952      0      0 var1
4  50 338.0085 19.900000      0      0 var1
5  28 427.9694 28.660714      0      0 var1
6  31 141.4214  3.887097     45      0 var1
7  74 241.2175  5.013514     45      0 var1
8  37 360.5551  8.094595     45      0 var1
9  62 459.9107  8.443548     45      0 var1
10 36 100.0000  1.458333     90      0 var1
11 33 200.0000  3.303030     90      0 var1
12 27 300.0000  4.314815     90      0 var1
13 71 343.3653  6.408451     90      0 var1
14 57 438.4636  8.315789     90      0 var1
15 32 141.4214  7.062500    135      0 var1
16 76 239.9746 10.006579    135      0 var1
17 34 360.5551 22.132353    135      0 var1
18 58 459.3618 32.655172    135      0 var1
```







## Conclusiones

Se confirma la similitud con lo descrito en las conclusiones anteriormente: comportamiento similar en las direcciones N-S y NE-SO, así como en las direcciones E-O y NO-SE.

## Referencias

---

Bivand, R. P. (2013). *Applied Spatial Data Analysis with R*. New York: Springer-Science+Business.

Clark, I. (2001). *Practical Geostatistics*. Central Scotland: Alloa Business Centre.

**OBJECTIVE OF THE PRACTICE:** Introduce basic concepts of geostatistics.

## EXAMPLE 1. Inverse Distance Weighted Interpolation, IDW

Suppose we have a local estimation problem of a value at point A in figure 1.1, (Clark, 2001)

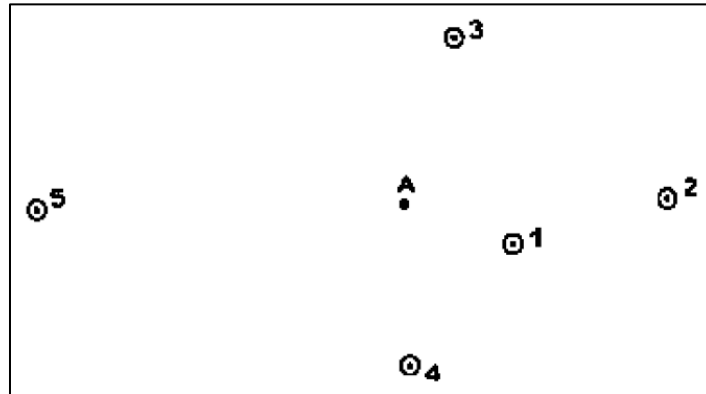


Figure 1.1. Hypothetical situation of sampling and estimation, (Clark, 2001)

To solve this example by applying the non-geostatistical IDW method, schemes such as those shown in figures 1.2 and 1.3 are proposed.

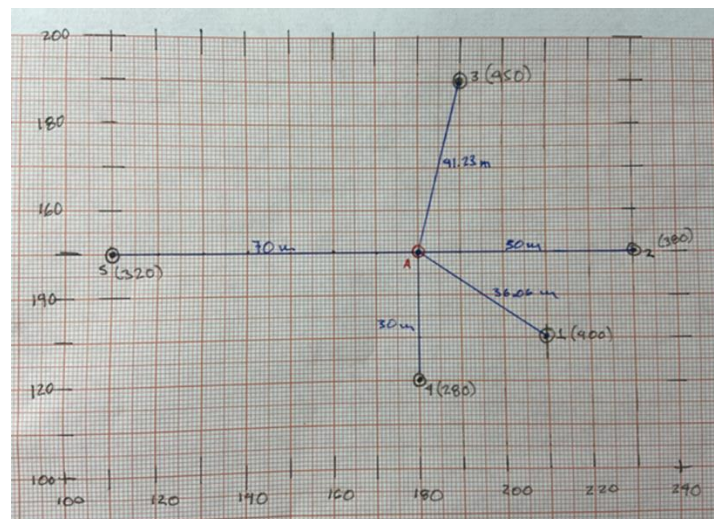


Figure 1.2. Dimensioned diagram of figure 1.1, in m

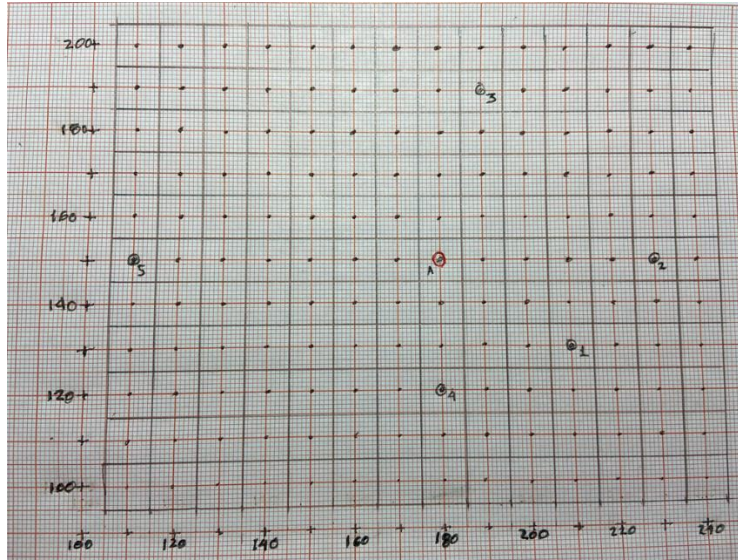


Figure 1.3. Scheme of the mesh for figure 1.1

## Inverse Distance Weighted Interpolation, IDW. (Bivand, 2013)

Inverse distance-based weighted interpolation (IDW) computes a weighted average

$$\hat{Z}(s_0) = \frac{\sum_{i=1}^n w(s_i) Z(s_i)}{\sum_{i=1}^n w(s_i)},$$

where the weights of the observations are calculated according to their distance from the interpolation location,

$$w(s_i) = ||s_i - s_0||^{-p},$$

where  $|| \cdot ||$  denotes the Euclidean distance and  $p$  an inverse distance weighting power, with a default value of 2. If  $s_0$  matches an observation location, the observed value is returned to avoid infinite weights.

The power of the inverse distance determines the degree to which closer points are preferred over more distant points; for large values, IDW converges to the nearest neighbor interpolation. It can be adjusted, for example, by cross-validation. IDW can also be used within local search neighborhoods.

Therefore,

$p = 1$ ;

$$\bar{A} = \frac{\left(\frac{1}{36.06} \cdot 400\right) + \left(\frac{1}{50} \cdot 380\right) + \left(\frac{1}{41.23} \cdot 450\right) + \left(\frac{1}{30} \cdot 280\right) + \left(\frac{1}{70} \cdot 3200\right)}{36.06^{-1} + 50^{-1} + 41.23^{-1} + 30^{-1} + 70^{-1}} = 363.80$$

$p = 2$ ;

$$\bar{A} = \frac{\left(\frac{1}{36.06} \cdot 400\right) + \left(\frac{1}{50} \cdot 380\right) + \left(\frac{1}{41.23} \cdot 450\right) + \left(\frac{1}{30} \cdot 280\right) + \left(\frac{1}{70} \cdot 3200\right)}{36.06^{-2} + 50^{-2} + 41.23^{-2} + 30^{-2} + 70^{-2}} = 358.26$$

## Solving with R

Example\_0.csv

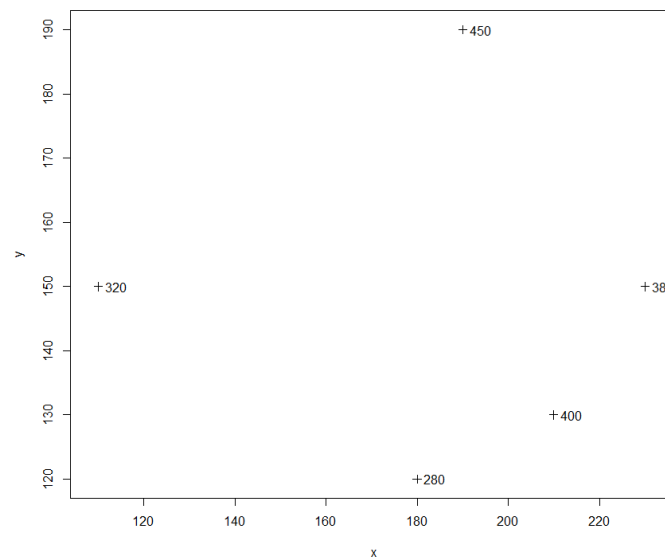
	A	B	C	D
1	x	y	V	
2	210	130	400	
3	230	150	380	
4	190	190	450	
5	180	120	280	
6	110	150	320	
7				
8				

Grid\_Example\_0.csv

x	y
110	100
110	110
110	120
110	130
110	140
110	150
110	160
110	170
110	180
110	190
120	100
120	110
120	120
120	130

230	140
230	150
230	160
230	170
230	180
230	190
240	100
240	110
240	120
240	130
240	140
240	150
240	160
240	170
240	180
240	190

```
library(sp)
library(gstat)
setwd("C:/Geostatistics/My directory")
# Read data
#
data <- read.csv("Example_0.csv", header=TRUE, sep=",", dec=".")
class(data)
str(data)
attach(data)
coordinates(data) <- c("x", "y") # create a spatial object
# Show data V
plot(x,y,xlim=c(min(x)*0.99,max(x)*1.01),ylim=c(min(y)*.999,max(y)*1.001),pch=3)
text(x+3,y,V)
```



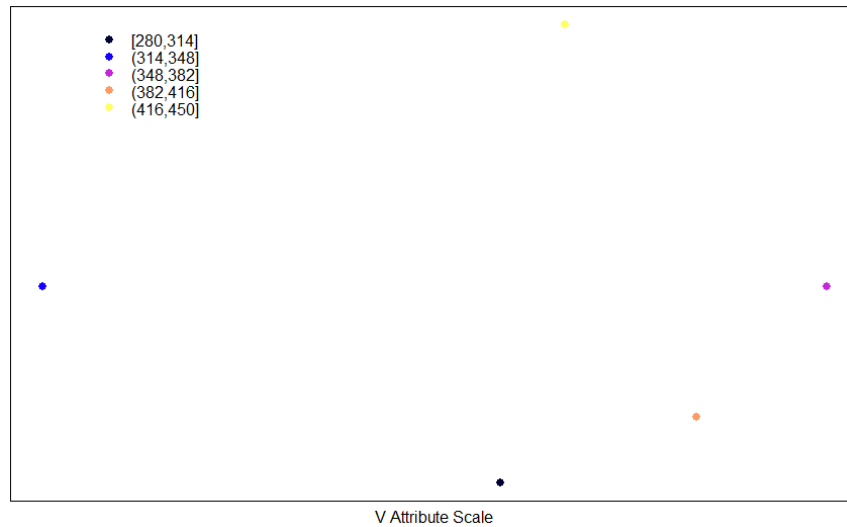
```
# Reject the null hypothesis if p <= alpha=0.05
# H0: the sample comes from a normally distributed population
shapiro.test(V) # *** Normality ***
qqnorm(V)
qqline(V)
```

```
> # Reject the null hypothesis if p <= alpha=0.05
> # H0: the sample comes from a normally distributed population
> shapiro.test(V) # *** Normality ***

Shapiro-Wilk normality test

data:  V
W = 0.97671, p-value = 0.9163
```

```
# plot methods for spatial data with attributes
# in case of non-normality: spplot(datos, "V", do.log = T, colorkey = TRUE)
spplot(datos, "V", xlab=" V Attribute Scale ", colorkey = F,
       key.space=list(x=0.1,y=.95))
```



```
# grid reading
griddata <- read.csv("Grid_Example_0.csv", header=TRUE, sep=";", dec=".")
class(griddata)
str(griddata)
attach(griddata)
coordinates(griddata) <- c("x", "y")
griddata <- as(griddata, "SpatialPixels")
```

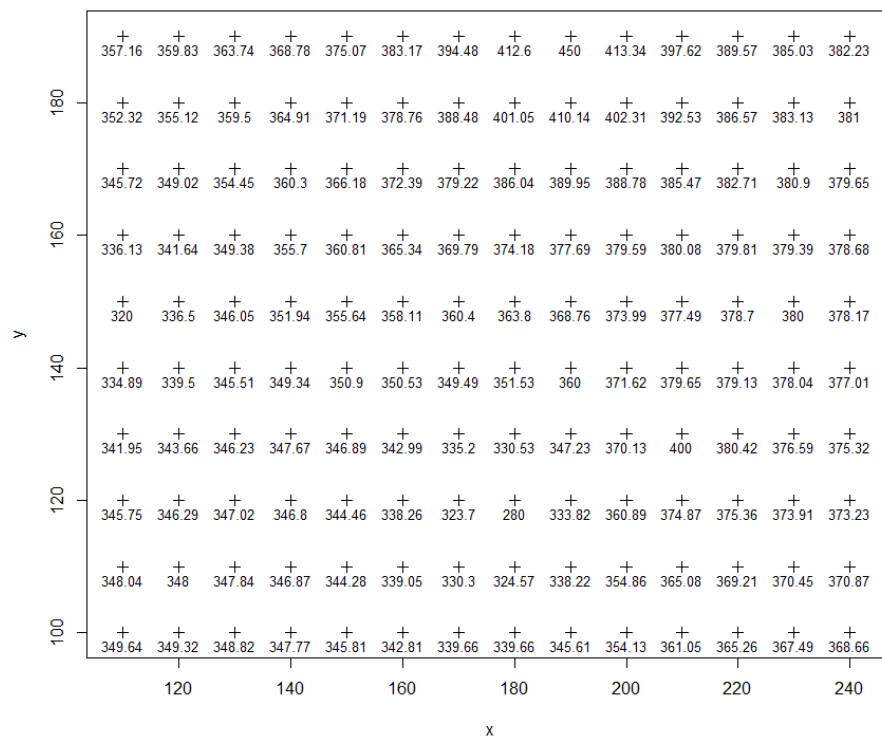
```
idw.out <- idw(V~1, data, griddata, idp=1)
as.data.frame(idw.out)[1:100,]
idwdata <- as.data.frame(idw.out)
write.csv(idw.out, file = "idw_res.csv") # Save a file .csv
plot(idw.out, axes=T)
idw.out
class(idw.out)
spplot(idw.out, "var1.pred", do.log = F, colorkey = TRUE)
```

```
> idw.out <- idw(V~1, data, griddata, idp=1)
[inverse distance weighted interpolation]
> as.data.frame(idw.out)[1:100,]
      x    y var1.pred var1.var
1  110  100  349.6394      NA
2  110  110  348.0423      NA
...
4   110  120  345.7471      NA
...
75  180  140  351.5289      NA
76  180  150  363.7968      NA
77  180  160  374.1768      NA
...
```

```
> idw.out <- idw(V~1,data,griddata,idp=2)
[inverse distance weighted interpolation]
> as.data.frame(idw.out)[1:100,]
      x    y var1.pred var1.var
1   110  100  335.2222      NA
2   110  110  332.8073      NA
3   110  120  329.6830      NA
...
75  180  140  331.8761      NA
76  180  150  358.2615      NA
77  180  160  382.3909      NA
...
```

Seen in a plane, with p=2,

```
plot(x,y,xlim=c(min(x)*0.99,max(x)*1.01),ylim=c(min(y)*.999,max(y)*1.001),pch=3)
text(x,y-2,round(idw.out$var1.pred,2),cex = 0.8)
```



## Conclusions

The similarity between the results of the previous section and the current ones is confirmed.



## EXAMPLE 2. Introduction to the Variogram

Figure 2.1 shows a sampling of the average value of the law for Fe (% by weight) in a set of perforations. The aim is to obtain the calculated or experimental semivariogram where  $h$  depends on the distance between the pair of samples and the relative orientation in a two-dimensional plane. (Clarks, 2001).

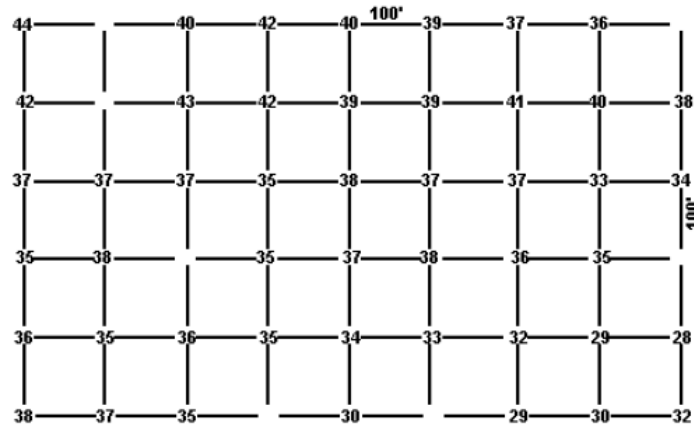


Figure 2.1 Example of data in a grid for the calculation of an experimental semivariogram - iron ore, taken from (Clark, 2001)

Then, to obtain the semivariogram, or variogram, the equation

$$2\gamma^*(h) = \frac{1}{n} \sum [g(x) - g(x+h)]^2$$

where  $\gamma^*(h)$  is precisely called the experimental semivariogram.

Starting the calculation from east to west, E-W, and with  $h = 100$  feet:

$$\begin{aligned} \gamma^*(100) = & [(40 - 42)^2 + (42 - 40)^2 + (40 - 39)^2 + (39 - 37)^2 + (37 - 36)^2 + (43 - 42)^2 \\ & + (42 - 39)^2 + (39 - 39)^2 + (39 - 41)^2 + (41 - 40)^2 + (40 - 38)^2 + (37 - 37)^2 \\ & + (37 - 37)^2 + (37 - 35)^2 + (35 - 38)^2 + (38 - 37)^2 + (37 - 37)^2 + (37 - 33)^2 \\ & + (33 - 34)^2 + (35 - 38)^2 + (35 - 37)^2 + (37 - 38)^2 + (38 - 36)^2 + (36 - 35)^2 \\ & + (36 - 35)^2 + (35 - 36)^2 + (36 - 35)^2 + (35 - 34)^2 + (34 - 33)^2 + (33 - 32)^2 \\ & + (32 - 29)^2 + (29 - 28)^2 + (38 - 37)^2 + (37 - 35)^2 + (29 - 30)^2 + (30 - 32)^2] \div (2 \\ & + 36) = 1.42 (\%)^2 \end{aligned}$$

This provides a point that we can plot on a plot of the experimental semivariogram ( $\gamma^*$ ) versus the distance between samples ( $h$ ), ie [100 ft, 1.42 (%)<sup>2</sup>]. Figure 2.2 outlines the calculation procedure.

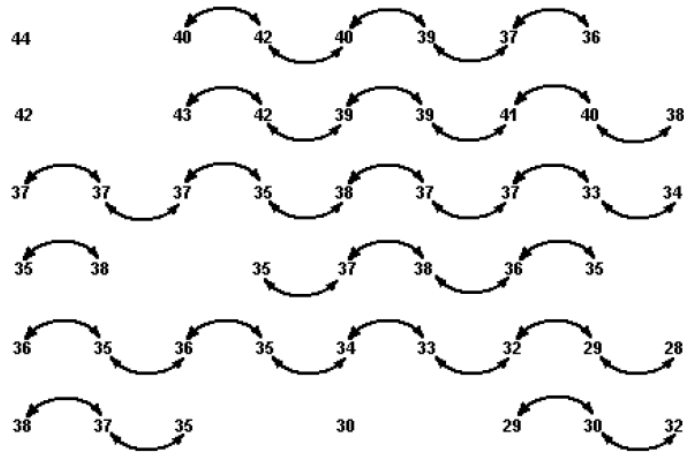


Figure 2.2. Identification of all pairs of sampling points within 100 feet of each other, in an east-west, E-W direction. (Clarks, 2001)

Now with  $h = 200$  feet,

$$\begin{aligned} \gamma^*(200) = & [(44 - 40)^2 + (40 - 40)^2 + (42 - 39)^2 + (40 - 37)^2 + (39 - 36)^2 + (42 - 43)^2 \\ & + (43 - 39)^2 + (42 - 39)^2 + (39 - 41)^2 + (39 - 40)^2 + (41 - 38)^2 + (37 - 37)^2 \\ & + (37 - 35)^2 + (37 - 38)^2 + (35 - 37)^2 + (38 - 37)^2 + (37 - 33)^2 + (37 - 34)^2 \\ & + (38 - 35)^2 + (35 - 38)^2 + (37 - 36)^2 + (36 - 35)^2 + (36 - 36)^2 + (35 - 35)^2 \\ & + (36 - 34)^2 + (35 - 33)^2 + (34 - 32)^2 + (33 - 29)^2 + (32 - 28)^2 + (38 - 35)^2 \\ & + (35 - 30)^2 + (30 - 29)^2 + (29 - 32)^2] \div (2 + 33) = 3.55 (\%)^2 \end{aligned}$$

Figure 2.3 shows the schematic calculation procedure.

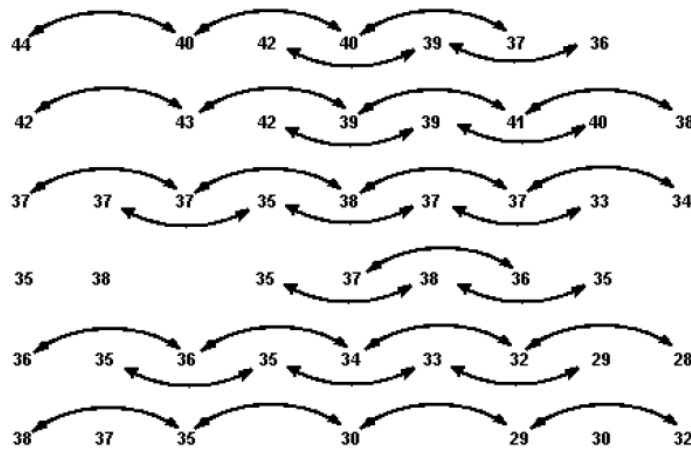


Figure 2.3. Identification of all pairs of sampling points within 200 feet of each other, in an east-west, E-W direction. (Clarks, 2001)

Table 2.1 summarizes the results obtained for the two main directions, considering distances up to half of the maximum extension in both directions, and figure 2.4 shows the graph of the corresponding semivariograms.

Dirección	Distancia entre muestras (feet)	Semivariograma experimental	Número de pares
E-O	100	1.42	36
	200	3.55	33
	300	3.54	26
	400	6.61	23
N-S	100	5.26	35
	200	9.06	27
	300	16.21	21
	400	25.54	13

Table 2.1 Calculation of Experimental Semivariogram Values in Two Principal Directions for the Square Grid Iron Ore Example

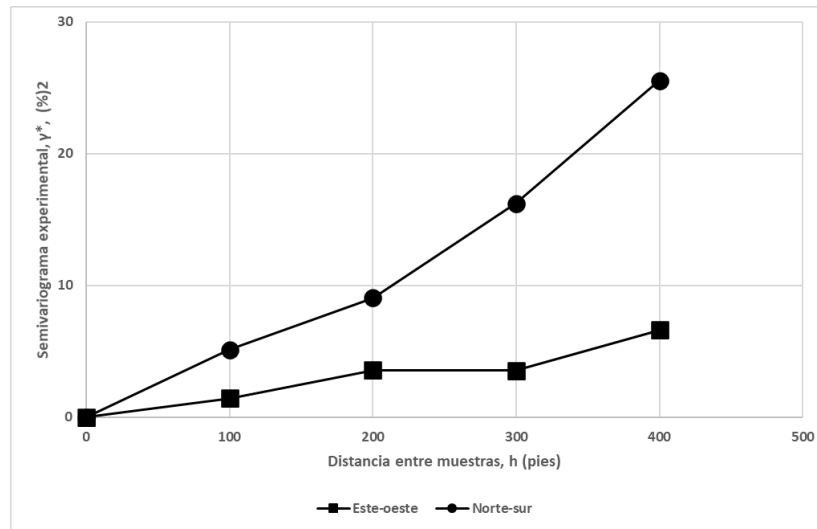


Figure 2.4 Experimental semivariograms in the two principal directions for the iron ore example.

## Conclusions

As can be seen, there appears to be a distinct difference in structure in the two directions (anisotropy). The N-S semivariogram rises much more sharply than the E-W, suggesting greater continuity in the E-W direction. This can be verified by calculating the semivariogram in diagonal directions, for example Northwest-Southeast, NW-SE, and Northeast-Southwest, NE-SW. For this, the  $h$  will take values of 141.42, 282.84 and 424.26 feet. Table 2.2 shows the results and figure 2.5 shows the behavior of all the calculated semivariograms E-W, N-S and diagonals NW-SE and NE-SE.

The  $\gamma^*$  in both diagonal directions seems to confirm the difference between the other two principals, since they lie between them; although the N-S of the NW-SE and the E-W of the NE-SW seem

to be closer. In addition, considering that the fewer pairs there are, the less reliable the calculations are, it would be necessary to evaluate the reliability of the experimental semivariograms in the direction of the diagonals, since they were calculated with 14 pairs only in 424 feet.

Direction	Distance between samples (feet)	Experimental semivariogram	Number of pairs
Northwest-Southeast	141	7.31	32
Diagonal	283	13.14	21
Diagonal	424	28.21	14
Northeast-Southwest	141	3.60	31
Diagonal	283	5.43	22
Diagonal	424	7.64	14

Table 2.2. Calculation of the semivariogram in the diagonal direction for iron ore.

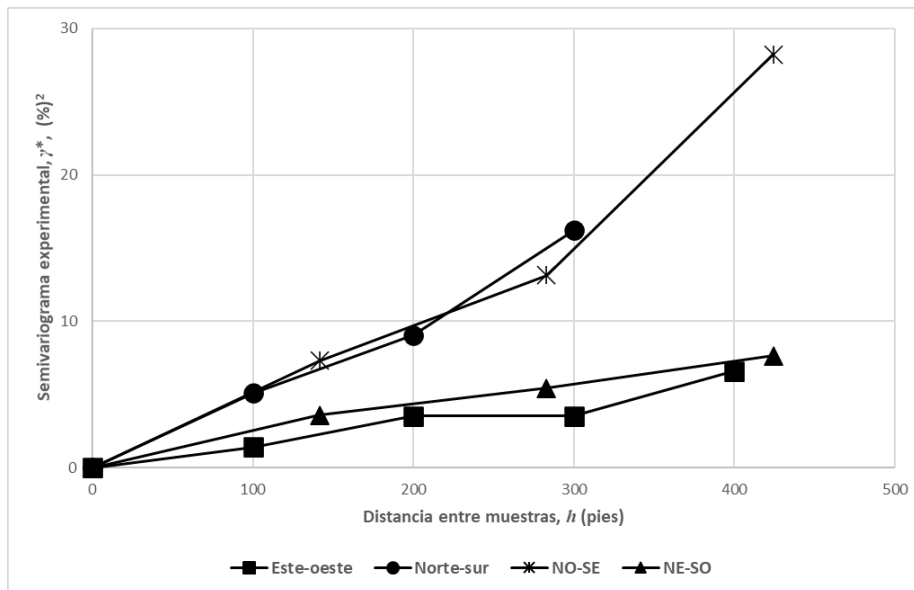


Figure 2.5. Experimental semivariograms including in diagonal direction for iron ore

Also, if a straight line is fitted to the semivariograms, the E-W slope would be  $0.0153 (\%)^2/\text{ft}$ , that is, approximately  $\gamma(h) = 0.0153h (\%)^2$ . In N-S direction it would be  $0.0524 (\%)^2/\text{ft}$ ,  $\gamma(h) = 0.0524h (\%)^2$ , in NW-SE direction  $0.064 (\%)^2/\text{ft}$ ,  $\gamma(h) = 0.064h (\%)^2$  and in NE-SW direction  $0.0175 (\%)^2/\text{ft}$ ,  $\gamma(h) = 0.0175h (\%)^2$ . That is, in the E-W direction, a squared difference of  $0.0153 (\%)^2$  is "expected" for each foot between the samples. Stated another way, a difference in grade of  $0.0153 = 0.1237\%$  Fe is expected for two samples 1 foot apart, with an E-W relative orientation; in the N-S direction, the corresponding figure is  $0.2289\%$  Fe. For samples 100 feet apart, differences of  $1.237\%$  Fe (E-O) and  $2.2289\%$  Fe (N-S) would be expected, and so on. With these observations it can be verified that there is similarity in the behavior of the Fe law in the N-S and NW-SE directions, as well as in the E-W and NE-SO directions. Therefore, a picture of the grade fluctuations within this section of the deposit has been created and a fairly simple model has been created to describe the differences in grade in the selected directions.

## Solving with R

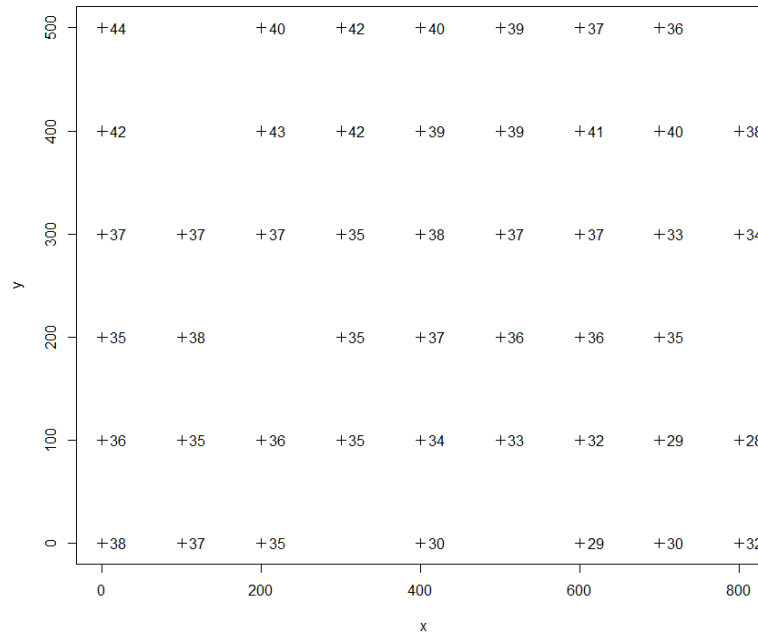
```
library(sp)
library(gstat)
setwd("C:/Geoestadistics/My Directory")
# Read the data
#
datos <- read.csv("Example_11.csv", header=TRUE, sep=";", dec=".")
```

*Example\_11.csv*

x	y	z
0	500	44
0	400	42
0	300	37
0	200	35
0	100	36
0	0	38
100	300	37
100	200	38
100	100	35
100	0	37
200	500	40
200	400	43
200	300	37
200	100	36
200	0	35
300	500	42
300	400	42
300	300	35
300	200	35
300	100	35
400	500	40
400	400	39
400	300	38
400	200	37
400	100	34
400	0	30

500	500	39
500	400	39
500	300	37
500	200	36
500	100	33
600	500	37
600	400	41
600	300	37
600	200	36
600	100	32
600	0	29
700	500	36
700	400	40
700	300	33
700	200	35
700	100	29
700	0	30
800	400	38
800	300	34
800	100	28
800	0	32

```
# view z-data
plot(x,y,xlim=c(min(x)*0.99,max(x)*1.01),ylim=c(min(y)*.999,max(y)*1.001),pch=3)
text(x+20,y,z)
```



# Estimate spatial correlation: variogram

coordinates(datos) <- c("x", "y") # create spatial data frame

hscat(z~1,datos,(0:8)\*100)qqline(z) # generally the smaller h, the better the correlation

# H0: the sample comes from a normally distributed population

# Reject the null hypothesis if  $p \leq \alpha=0.05$

shapiro.test(z) # \*\*\* Normality \*\*\*

qqnorm(z)

```
> summary(z) # Similar median and mean
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
28.00	35.00	37.00	36.34	38.50	44.00

```
> # H0: the sample comes from a normally distributed population
```

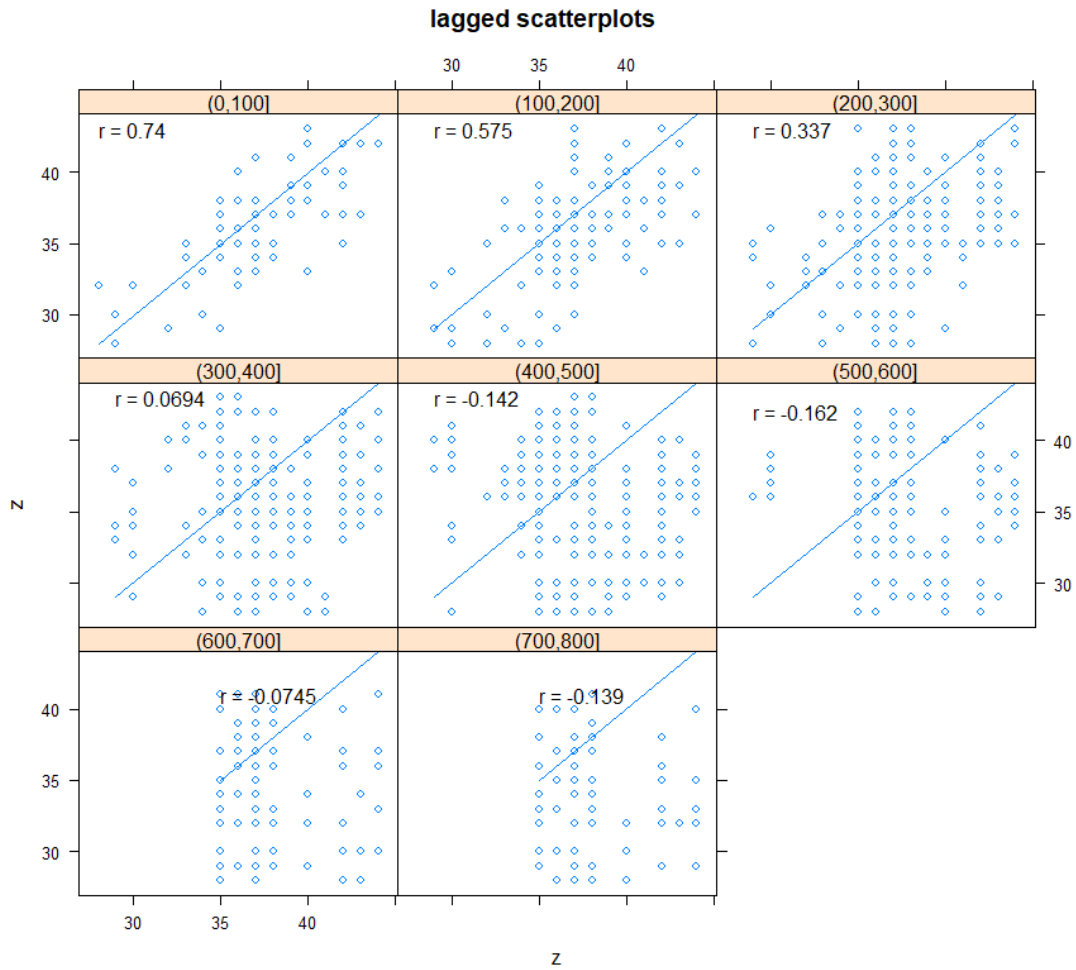
```
> # Reject the null hypothesis if  $p \leq \alpha=0.05$ 
```

```
> shapiro.test(z) # *** Normality ***
```

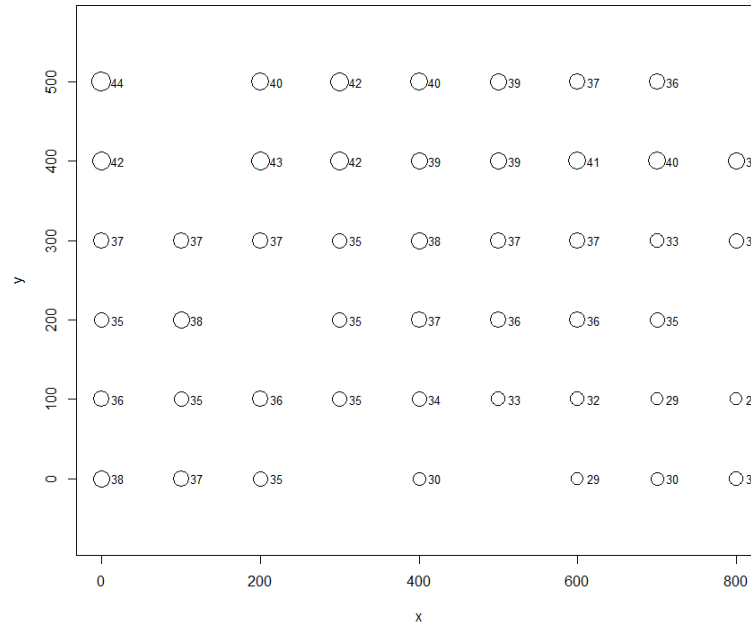
Shapiro-Wilk normality test

data: z

W = 0.97355, p-value = 0.3593



```
# Another visualization with proportional symbols.
# bubble chart
plot(datos, asp=1, cex = 3*z/max(z), pch=1, axes=T)
text(x+20,y,z, cex=0.8)
```



```
# calculates how many pairs of points there are in dataset data.
```

```
n <- length(z)
```

```
n
```

```
n*(n+1)/2
```

```
[1] 47
> n*(n+1)/2
[1] 1128
```

```
# LOCAL SPATIAL STRUCTURE: AUTOCORELATION
```

```
#
```

```
class(datos)
```

```
str(datos) # Slot @proj4string does not have the reference system assigned.
```

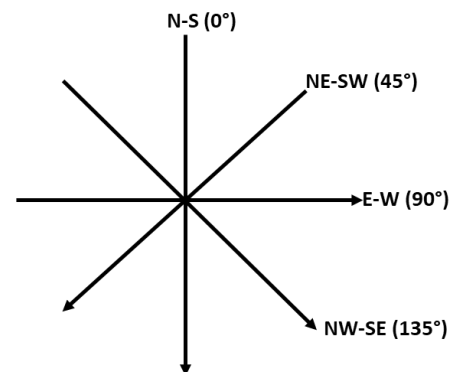


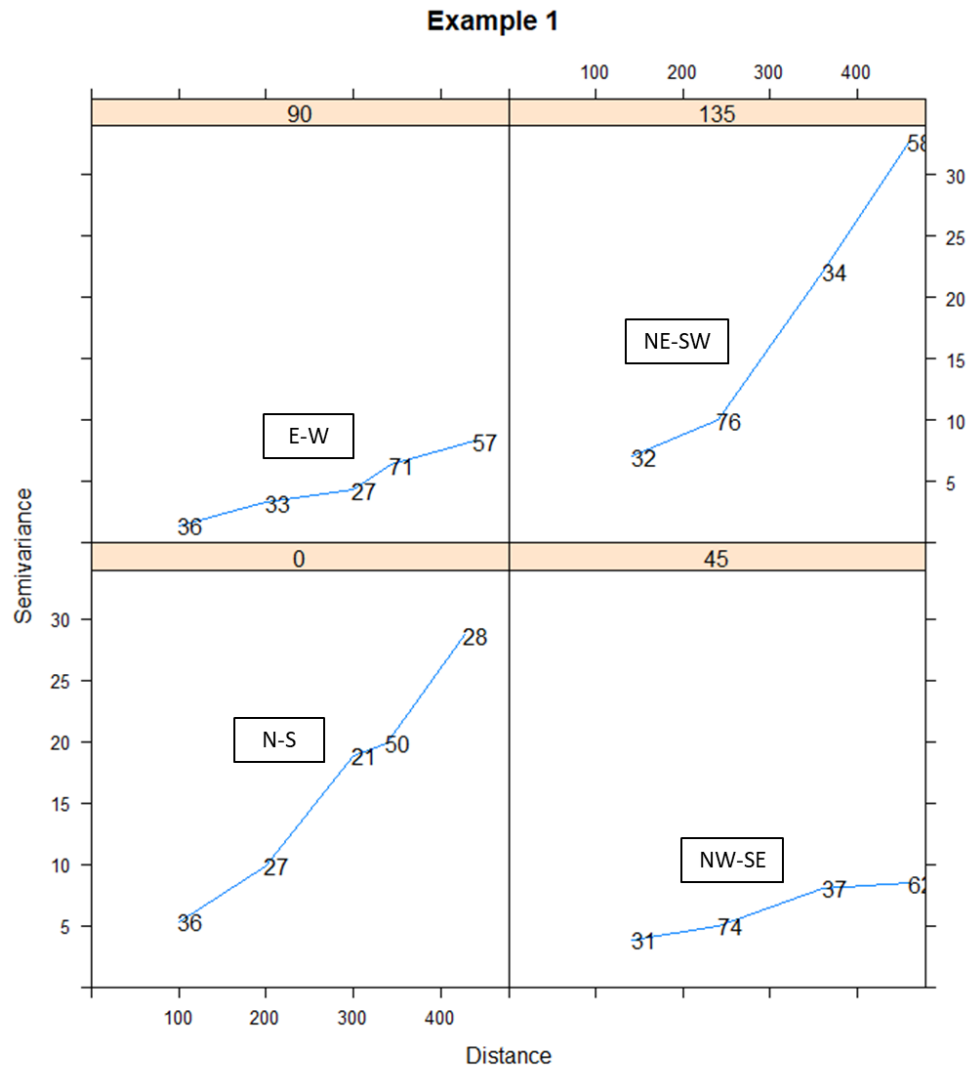
```
> # LOCAL SPATIAL STRUCTURE: AUTOCORELATION
> #
> class(datos)
[1] "SpatialPointsDataFrame"
attr(,"package")
[1] "sp"
> str(datos) # Slot @proj4string does not have reference system assigned.
Formal class 'SpatialPointsDataFrame' [package "sp"] with 5 slots
..@ data      : 'data.frame': 47 obs. of 1 variable:
.. ..$ z: int [1:47] 44 42 37 35 36 38 37 38 35 37 ...
..@ coords.nrs : int [1:2] 1 2
..@ coords     : num [1:47, 1:2] 0 0 0 0 0 0 100 100 100 100 ...
.. ..- attr(*, "dimnames")=List of 2
.. .. ..$ : NULL
.. .. ..$ : chr [1:2] "x" "y"
..@ bbox      : num [1:2, 1:2] 0 0 800 500
.. ..- attr(*, "dimnames")=List of 2
.. .. ..$ : chr [1:2] "x" "y"
.. .. ..$ : chr [1:2] "min" "max"
..@ proj4string: Formal class 'CRS' [package "sp"] with 1 slot
.. .. ..@ projargs: chr NA
```

To compare with the results in Figure 2.5,

```
# EXPERIMENTAL EMPIRICAL VARIOGRAM
#
ve <- variogram(z~1,data,alpha=c(0,45,90,135),cutoff=500,width=100)
ve
plot(ve, plot.numbers = T, main = "Example 1", xlab = "Distance",
     ylab = "Semivariance", asp=1, type="l")
```

```
> ve
  np    dist    gamma dir.hor dir.ver  id
1  36 100.0000  5.347222      0      0 var1
2  27 200.0000  9.870370      0      0 var1
3  21 300.0000 18.880952      0      0 var1
4  50 338.0085 19.900000      0      0 var1
5  28 427.9694 28.660714      0      0 var1
6  31 141.4214  3.887097     45      0 var1
7  74 241.2175  5.013514     45      0 var1
8  37 360.5551  8.094595     45      0 var1
9  62 459.9107  8.443548     45      0 var1
10 36 100.0000  1.458333     90      0 var1
11 33 200.0000  3.303030     90      0 var1
12 27 300.0000  4.314815     90      0 var1
13 71 343.3653  6.408451     90      0 var1
14 57 438.4636  8.315789     90      0 var1
15 32 141.4214  7.062500    135      0 var1
16 76 239.9746 10.006579    135      0 var1
17 34 360.5551 22.132353    135      0 var1
18 58 459.3618 32.655172    135      0 var1
```





## Conclusions

The similarity with what was described in the previous conclusions is confirmed: similar behavior in the N-S and NE-SW directions, as well as in the E-W and NW-SE directions.

## References

---

Bivand, R. P. (2013). *Applied Spatial Data Analysis with R*. New York: Springer-Science+Business.

Clark, I. (2001). *Practical Geostatistics*. Central Scotland: Alloa Business Centre.