

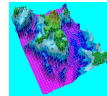
GEOESTADÍSTICA APLICADA

Tema: Análisis Variográfico

Dr. Martín A. Díaz Viera,
Dr. Ricardo Casar González



UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO
mdiazv@imp.mx



Contenido I

- 1 Introducción
- 2 El semivariograma
 - El semivariograma. Definición
 - Estimación del semivariograma
 - Forma general del variograma
 - Consideraciones para el cómputo del variograma
- 3 Modelos de variogramas
 - Modelos válidos o autorizados
 - Modelos transitivos o acotados
 - Modelos no acotados
 - Variogramas anisotrópicos
- 4 Modelación del variograma
 - Modelación del variograma

Contenido II

- Ajuste con mínimos cuadrados
- Ajuste con visual
- Validación del modelo

Introducción

- El análisis variográfico o estructural es uno de los tópicos más importantes de la geoestadística puesto que se encarga de la caracterización de la estructura espacial de una propiedad o fenómeno.
- Es el proceso en el marco del cual se obtiene un modelo geoestadístico para la función aleatoria que se estudia.

Introducción

- El análisis variográfico consiste en estimar y modelar una función que refleje la correlación espacial de una muestra de datos a partir de la adopción razonada de la hipótesis más adecuada acerca de su estacionaridad.
- En dependencia del grado de estacionaridad se modelará la función de covarianzas o la de semivarianzas.
- Por su generalidad estudiaremos el proceso de estimación y modelación de la función de semivarianzas o semivariograma.

El semivariograma. Definición

- El **semivariograma**, conocido también como **variograma**, es la herramienta central de la geoestadística.
- Dada una FA $\mathbf{Z}(\underline{x})$ que cumpla al menos la **Hipótesis Intrínseca** entonces la función de semivarianzas es estacionaria y se define como sigue:

$$\gamma(\underline{h}) = \frac{1}{2} \text{Var} [Z(\underline{x}) - Z(\underline{x} + \underline{h})] = \frac{1}{2} E \left[\{Z(\underline{x}) - Z(\underline{x} + \underline{h})\}^2 \right] \quad (1)$$

Estimación del semivariograma

- Existen varios estimadores del semivariograma pero el estimador más común, conocido como estimador de Matheron, está dado por

$$\gamma^*(\underline{h}) = \frac{1}{2N(\underline{h})} \sum_{i=1}^{N(\underline{h})} [Z(\underline{x}_i + \underline{h}) - Z(\underline{x}_i)]^2 \quad (2)$$

- $N(\underline{h})$ es el número de pares $Z(\underline{x}_i)$ y $Z(\underline{x}_i + \underline{h})$ separados a una distancia $h = |\underline{h}|$

Estimación del semivariograma

- Debido a que el estimador es esencialmente una media muestral (promedio), tiene todas las desventajas comúnmente asociadas a este tipo de estimadores como es la no robustez.
- A menudo el empleo de este estimador produce variogramas experimentales erráticos, lo cuál se debe a desviaciones del caso ideal para la aplicación del mismo.

Estimación del semivariograma

- Estas desviaciones pueden ser enumeradas por su importancia en el orden siguiente:
 - 1 Desviaciones en la distribución (asimetría)
 - 2 No homogeneidad de la varianza
 - 3 Desviaciones en el muestreo
 - 4 Existencia de valores atípicos (outliers) tanto distribucionales como espaciales.

Forma general del variograma

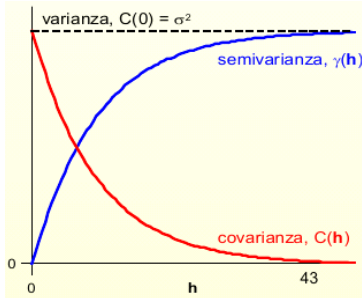


Figura 1: Relación de las funciones de covarianzas y semivarianzas ($\sigma_{ij} = \sigma^2 - \gamma_{ij}$).

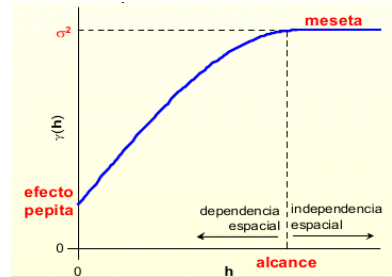


Figura 2: Parámetros del variograma: meseta (sill), alcance (range) y efecto pepita (nugget effect).

Consideraciones para el cómputo del variograma

- Los pares de todas las observaciones se agrupan según su distancia de separación dentro de un **intervalo** o **lag** h y sus múltiplos $2h, 3h, \dots, nh$, con una tolerancia de $\pm h/2$ y en una dirección con ángulo ϕ con una tolerancia angular.
- Se calcula para intervalos menores que la mitad de la distancia máxima de separación de todos los puntos.
- Se considera que un número máximo de 25 intervalos es suficiente para cualquier propósito y un mínimo de 10 debe ser usado para determinar con precisión el alcance y la meseta del variograma.
- Se considera que debe haber entre 30 y 50 pares de puntos como mínimo por intervalo.
- Los valores estimados para cada intervalo se deben graficar contra la distancia promedio de todos los pares que se encuentran dentro de dicho intervalo.

Consideraciones para el cómputo del variograma

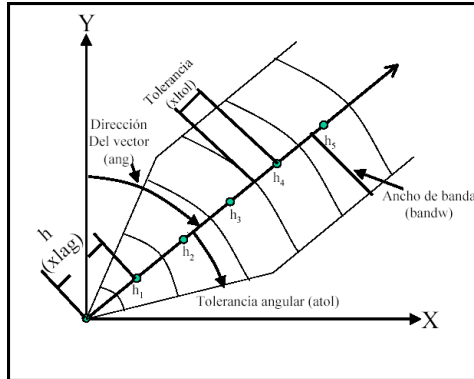


Figura 3: Diagrama del cómputo del variograma.

Cómputo del variograma



$$\begin{aligned} h = 1: & (10-7)^2 + (11-10)^2 + (13-11)^2 \\ & + (12-13)^2 + (14-12)^2 + (12-14)^2 \\ & + (13-12)^2 + (10-13)^2 + (11-10)^2 \\ & + (9-11)^2 + (8-9)^2 = 39 \end{aligned}$$

$$N(1) = 11$$

$$\hat{\gamma}(1) = 39/22 = 1.78$$

Figura 4: Ejemplo del cómputo del variograma en 1D.

Distribución Espacial

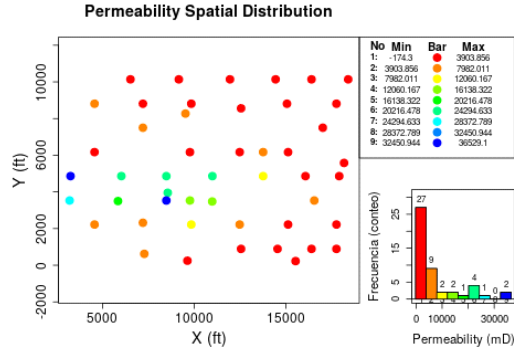


Figura 5: Distribución espacial.

Consideraciones para el cómputo del variograma

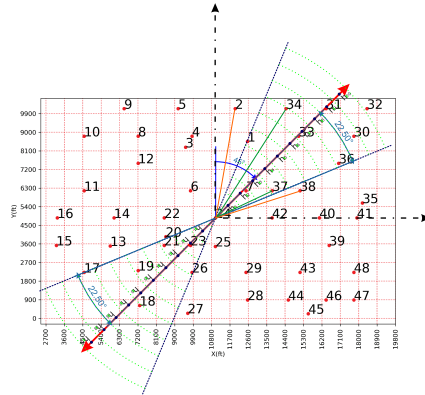
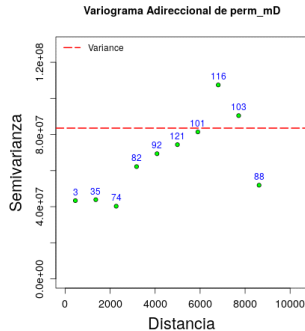


Figura 6: Diagrama del cómputo del variograma.

Ejemplo de variograma estimado: $d_{min}=428$, $d_{max}=16,575$, $h=828$



Estadígrafo	Valor
Muestras	48
Mínimo	7.4
1º cuartil	1002.45
Mediana	3482.205
Media	6818.24521
3º cuartil	7743.625
Máximo	36347.4
Rango	36340
Rango intercuartil	6741.175
Varianza	83532706.36
Desviación estándar	9139.62288
Simetría	1.83579
Curtosis	5.62603

Tabla 1: Estadística básica.

Figura 7: Variograma estimado.

Ejemplo de variograma estimado

Dirección = 0 grados, Tolerancia = 90.0 grados, Intervalo = 828 ft.

No. lag	No. pares	Intervalos (lags)	Semivarianzas
1	3	453.71	43,357,868.50
2	35	1361.14	43,869,273.09
3	74	2268.56	40,249,566.62
4	82	3175.98	62,199,450.34
5	92	4083.41	69,343,355.94
6	121	4990.83	74,381,397.02
7	101	5898.25	81,429,116.65
8	116	6805.68	107,465,654.01
9	103	7713.10	90,459,560.73
10	88	8620.52	51,904,218.42

Tabla 2: Ejemplo de variograma estimado.

Modelos válidos o autorizados

- Sólo ciertas funciones pueden ser consideradas modelos válidos de variogramas y se dicen que son **modelos autorizados**.
- El hecho de probar si una función dada es aceptable o no, está relacionado con el examen de su Transformada de Fourier. [Christakos, 1984] obtuvo las condiciones que el espectro de la función debe reunir para ser un modelo autorizado.
- Como una propiedad importante se debe destacar que cualquier combinación lineal de modelos autorizados es un modelo autorizado.

Modelos transitivos o acotados

- Este grupo de modelos se deriva a partir de la noción de autocorrelación entre los valores promedios dentro de los bloques.
- La idea es que la función aleatoria, de la cual la propiedad medida es una realización, depende del grado de traslape de los dos bloques.

Modelo Esférico

$$\gamma(h) = \begin{cases} \frac{S}{2} \left\{ 3 \left(\frac{h}{a} \right) - \left(\frac{h}{a} \right)^3 \right\} & \text{para } 0 \leq h \leq a \\ S & \text{para } h > a \end{cases} \quad (3)$$

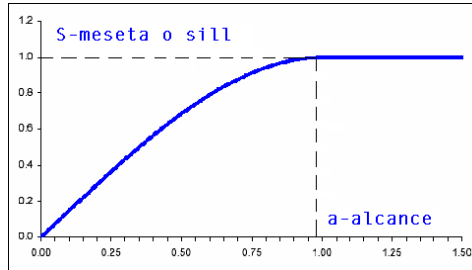


Figura 8: Modelo Esférico.

Modelo Exponencial

$$\gamma(h) = S \left[1 - \exp \left(-\frac{3h}{a} \right) \right] \text{ para } h \geq 0 \quad (4)$$

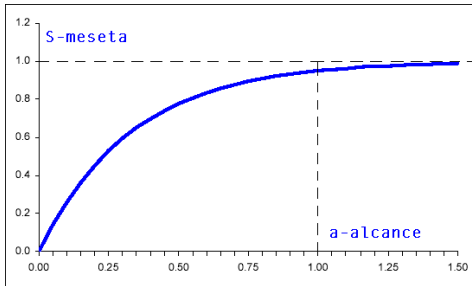


Figura 9: Modelo Exponencial.

Modelo Gaussiano

$$\gamma(h) = S \left[1 - \exp \left(- \left(\frac{3h}{a} \right)^2 \right) \right] \text{ para } h \geq 0 \quad (5)$$

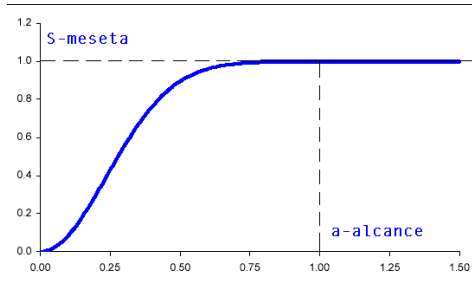


Figura 10: Modelo Gaussiano.

Modelo Efecto Agujero (Hole)

$$\gamma(h) = S \left(1 - \cos \left(\frac{\pi h}{a} \right) \right) \text{ para } h \geq 0 \quad (6)$$

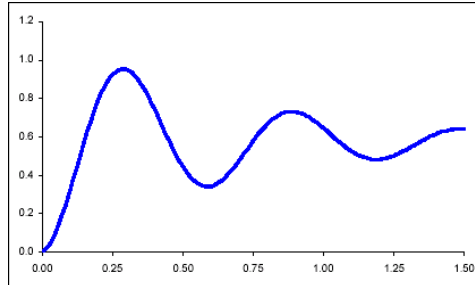


Figura 11: Modelo Efecto Agujero.

Modelo Efecto Pepita (Nugget)

$$\gamma(h) = S(1 - \delta(h)) \quad (7)$$

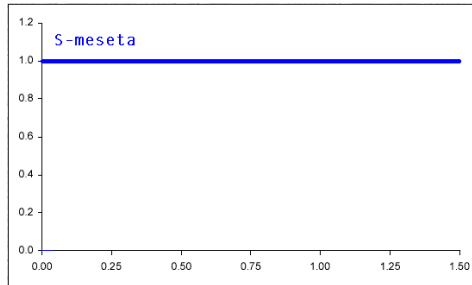


Figura 12: Modelo Efecto Pepita (Nugget).

Influencia del Alcance

Dominio de 100x100 metros. Modelo de variograma con diferentes alcances

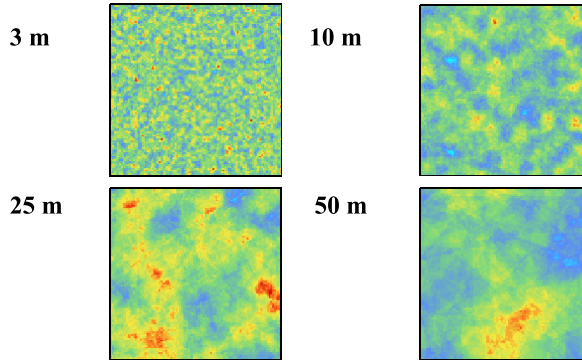
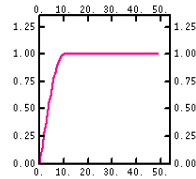
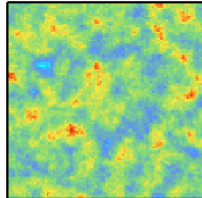


Figura 13: Influencia del Alcance.

Influencia del Efecto Pepita

**Sin Efecto
Nugget**



**33%
Nugget**

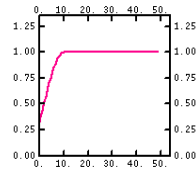
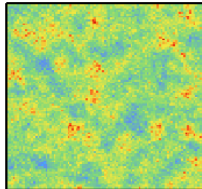
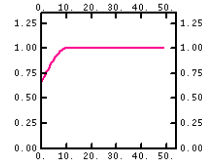
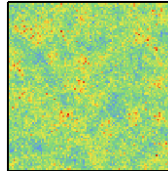


Figura 14: Influencia del Efecto Pepita.

Influencia del Efecto Pepita

**66%
Nugget**



**100%
Nugget**

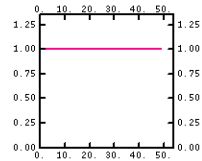
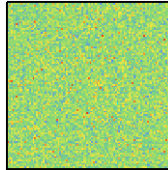


Figura 15: Influencia del Efecto Pepita.

Modelos no acotados

- Existen casos en que la varianza aparenta incrementarse indefinidamente.
- Si se toma cada vez un menor intervalo de muestreo, siempre existe alguna variación que queda sin resolver.
- [Mandelbrot et al., 1983] llamó al resultado de tales procesos fractales.
- Ejemplos: El ruido Gaussiano fraccional (fGn) y el movimiento Browniano fraccional (fBm).
- Para conocer más detalles ver la presentación CG8a sobre Geoestadística Fractal.

Modelos no acotados

- Modelo Potencia

$$\gamma(h) = \frac{1}{2}h^\theta \text{ para } 0 < \theta < 2 \quad (8)$$

- Relación entre la dimensión fractal **D** (Hausdorff-Besicovitch) y el parámetro θ

$$D = 2 - \left(\frac{\theta}{2}\right) \quad (9)$$

Casos Extremos:

- 1 $\theta = 2$ y **D**=1, es una parábola, no representa un proceso aleatorio.
- 2 $\theta = 0$ y **D**=2, ruido puro (efecto nugget).

Combinación de Modelos

- Una combinación lineal de modelos de variogramas autorizados con coeficientes positivos representa un modelo de variograma válido.
- Usualmente los modelos anteriores los encontramos como combinaciones del tipo:

$$\gamma(h) = \gamma_0(h) + \gamma_1(h) \quad (10)$$

Donde $\gamma_0(h)$ es el efecto nugget y $\gamma_1(h)$ es otro modelo.

Combinación de Modelos

- La combinación formada por variogramas con diferentes alcances se conoce como **estructura anidada**.
- Describen variaciones espaciales a diferentes escalas y que se deben por lo tanto, a factores de naturaleza diferente.
- Por ejemplo:

$$\gamma(h) = \gamma_0(h) + \gamma_1(h) + \gamma_2(h) \quad (11)$$

Donde $\gamma_0(h)$ es el efecto nugget, $\gamma_1(h)$ es el modelo esférico 1 y $\gamma_2(h)$ es el modelo esférico 2.

Combinación de Modelos

$$\gamma(h) = \gamma_0(h) + \gamma_1(h) + \gamma_2(h)$$

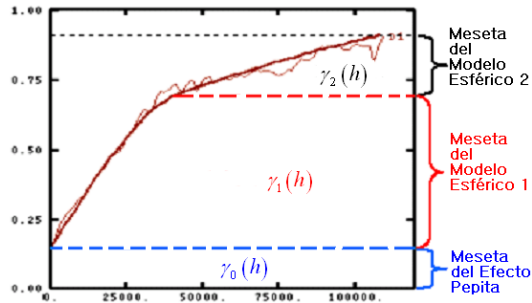


Figura 16: Combinación de Modelos.

Anisotropía Geométrica

- Existen numerosas situaciones en que la variación es anisotrópica, es decir depende de la dirección.
- Cuando la anisotropía se puede tener en cuenta mediante una transformación lineal simple de las coordenadas, entonces se dice que la **anisotropía es geométrica** o afín.
- La anisotropía se refleja en diferentes alcances según la dirección. El gráfico direccional de los alcances forma una elipse.

Variogramas Anisotrópicos

- Fórmula de la transformación:

$$\Omega(\theta) = (\mathbf{A}^2 \cos^2(\theta - \phi) + \mathbf{B}^2 \sin^2(\theta - \phi))^{\frac{1}{2}} \quad (12)$$

A es el eje mayor (variabilidad es más lenta).

B es el eje menor (variabilidad es más rápida)

ϕ es la dirección (ángulo) del eje mayor.

- Se aplica como un factor al argumento h del variograma en los modelos acotados o al gradiente en los modelos no acotados.
- $\lambda = \mathbf{A}/\mathbf{B}$ es una medida de la anisotropía.

Anisotropía Geométrica

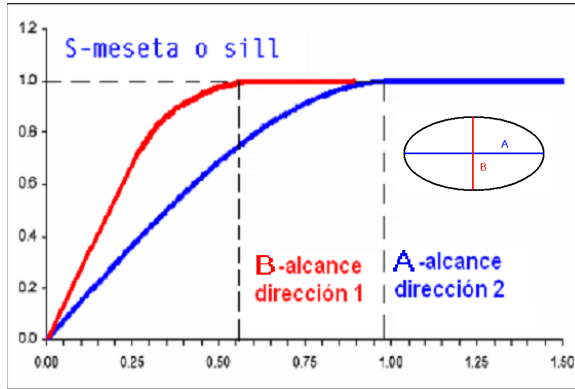


Figura 17: Anisotropía Geométrica.

Anisotropía Geométrica

- En la práctica se estiman los variogramas en 4 direcciones principales (0, 45, 90 y 135 grados).
- Se determinan los alcances para cada dirección.
- Luego se construye el gráfico direccional de los alcances para decidir si hay anisotropía geométrica presente o no.
- Finalmente, se determinan A, B y la dirección (ángulo) de mayor alcance.

Anisotropía Geométrica

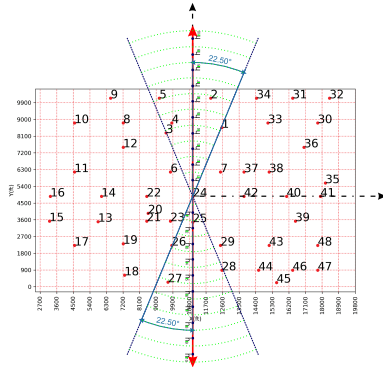


Figura 18: Dirección de 0 grados.

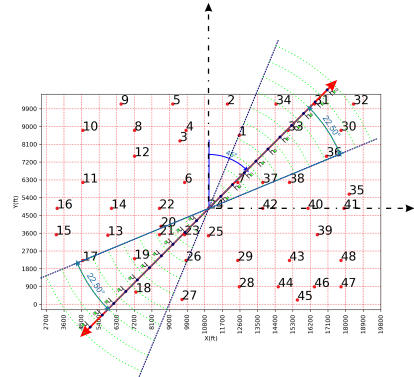


Figura 19: Dirección de 45 grados.

Anisotropía Geométrica

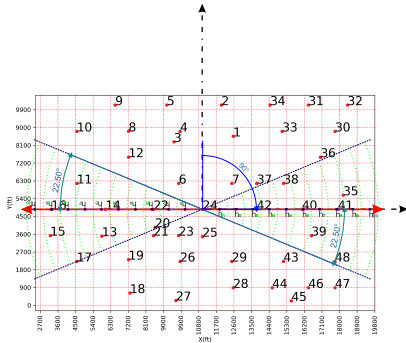


Figura 20: Dirección de 90 grados.

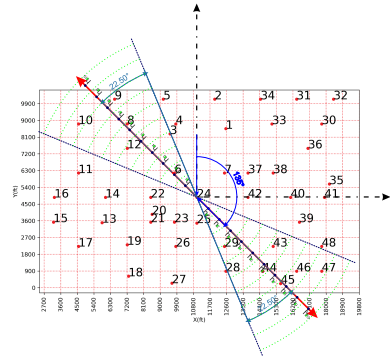


Figura 21: Dirección de 135 grados.

Anisotropía Geométrica

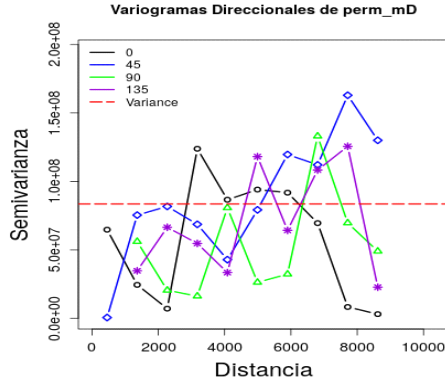
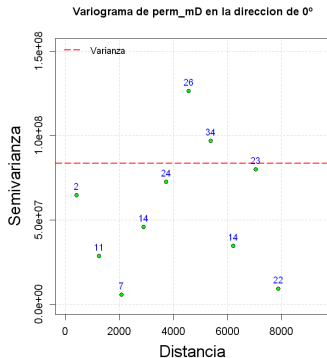


Figura 22: Variogramas direccionales en cuatro direcciones: 0, 45, 90 y 135 grados.

Anisotropía Geométrica

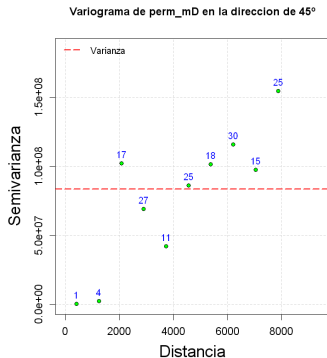


Intervalo	Pares	Lags	Semivarianzas
1	2	414.37	64750363.71
2	11	1243.12	28755244.6
3	7	2071.88	5724445.13
4	14	2900.63	45843992.7
5	24	3729.38	72759015.99
6	26	4558.13	126347699.3
7	34	5386.88	97059733.6
8	14	6215.64	34739560.48
9	23	7044.39	80081235.43
10	22	7873.14	9397970.11

Tabla 3: Variograma direccional 0 grados

Figura 23: Variograma direccional 0 grados.

Anisotropía Geométrica

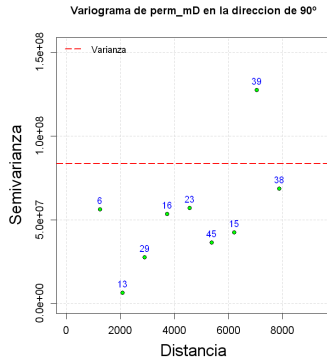


Intervalo	Pares	Lags	Semivarianzas
1	1	414.37	572878.08
2	4	1243.12	2320975.358
3	17	2071.88	102360387.4
4	27	2900.63	69134525.37
5	11	3729.38	42216734.23
6	25	4558.13	86186198.57
7	18	5386.88	101636450.7
8	30	6215.64	115825532.8
9	15	7044.3	97763871.59
10	25	7873.14	154618969.6

Tabla 4: Variograma direccional 45 grados

Figura 24: Variograma direccional 45 grados.

Anisotropía Geométrica

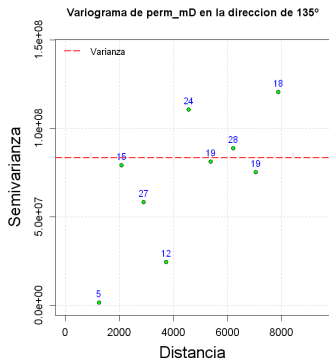


Intervalo	Pares	Lags	Semivarianzas
1	6	1243.12	56094073.24
2	13	2071.88	6430924.343
3	29	2900.63	27468937.94
4	16	3729.38	53356753.33
5	23	4558.13	56949656.45
6	45	5386.88	36504982.62
7	15	6215.64	42374284.98
8	39	7044.39	127421339.8
9	38	7873.14	68497416.97
10			

Tabla 5: Variograma direccional 90 grados

Figura 25: Variograma direccional 90 grados.

Anisotropía Geométrica



Intervalo	Pares	Lags	Semivarianzas
1	5	1243.128337	1743103.834
2	15	2071.88	79321110.63
3	27	2900.63	58488612.68
4	12	3729.38	24627239.99
5	24	4558.13	110669453.9
6	19	5386.88	81342908.24
7	28	6215.64	88835739.01
8	19	7044.39	75397314.35
9	18	7873.14	120586555.3
10			

Tabla 6: Variograma direccional 135 grados

Figura 26: Variograma direccional 135 grados.

Anisotropía Zonal

- Cuando la anisotropía se refleja en la meseta, es decir, en dependencia de la dirección el variograma presenta diferentes mesetas.
- Un ejemplo típico es la situación cuando tenemos medida cierta propiedad en diferentes pozos.
- Por lo general el variograma en la vertical (en profundidad) presenta una mayor variabilidad (meseta) que en la dirección horizontal.

Anisotropía Zonal

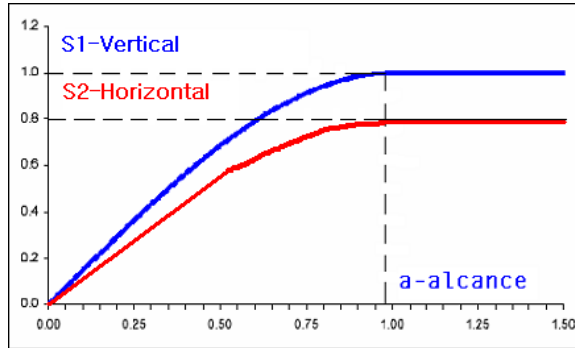


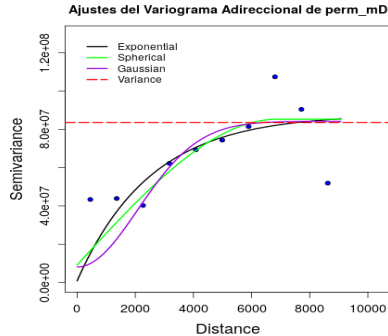
Figura 27: Anisotropía Zonal.

Modelación del variograma

- Algunos geoestadísticos ajustan los modelos de forma visual.
- Es recomendable auxiliarse con algún procedimiento estadístico.
- Por ejemplo, un ajuste con un Método de Mínimos Cuadrados.
- Se requiere de un criterio que considere la bondad del ajuste y la complejidad del modelo.
- El criterio de información de Akaike (AIC), que se define como

$$AIC = -2 \ln(\text{máxima verosimilitud}) + 2 (\text{número de parámetros}) \quad (13)$$

Ajuste con mínimos cuadrados



Modelo	Nugget	Meseta+Nugget	Alcance	SCE
Exponential	814,495.7	87,992,879	2,533.95	2.79 e+15
Spherical	9,113,830.9	85,317,292	6,882.69	2.56 e+15
Gaussian	8,103,578.2	84,238,551	2,952.40	3.26 e+15

Tabla 7: Ajuste con mínimos cuadrados.

Figura 28: Ajuste con mínimos cuadrados.

Mejor ajuste con mínimos cuadrados

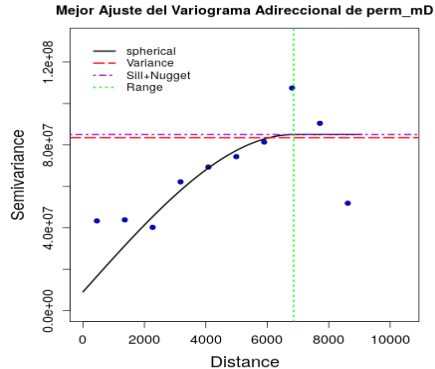


Figura 29: Mejor ajuste con mínimos cuadrados.

Ajuste con visual

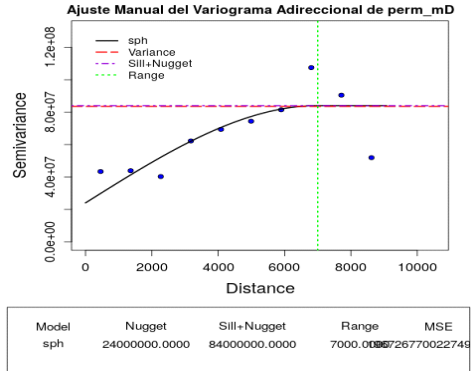


Figura 30: Mejor ajuste visual.

Validación cruzada del modelo

- Un método sencillo y eficiente es el método de **validación cruzada** conocido como **leave one out**.
- Consiste en sacar un punto de la muestra y estimar con **Kriging** el valor en ese punto usando el modelo de variograma obtenido.
- De forma análoga se actúa para el resto de los elementos de la muestra.
- Como resultado se obtiene un mapa y la estadística de las diferencias $(Z(\underline{x}_i) - Z^*(\underline{x}_i))$ entre el valor observado $Z(\underline{x}_i)$ y el valor estimado $Z^*(\underline{x}_i)$ para cada ubicación \underline{x}_i .

Validación cruzada del modelo

- Si el modelo del variograma refleja adecuadamente la estructura espacial del conjunto de datos, entonces los valores estimados $Z^*(\underline{x}_i)$ deben ser cercanos a los valores observados $Z(\underline{x}_i)$.
- Esta **cercanía** puede ser caracterizada según los siguientes criterios:

$$\frac{1}{n} \sum_{i=1}^n \{Z(\underline{x}_i) - Z^*(\underline{x}_i)\} \quad \text{cercano a 0} \quad (14)$$

$$\frac{1}{n} \sum_{i=1}^n \{Z(\underline{x}_i) - Z^*(\underline{x}_i)\}^2 \quad \text{pequeño} \quad (15)$$

- El histograma de las diferencias $Z(\underline{x}_i) - Z^*(\underline{x}_i)$ permite identificar **valores atípicos espaciales**, datos sospechosos o anomalías de otra naturaleza.

Validación cruzada del modelo

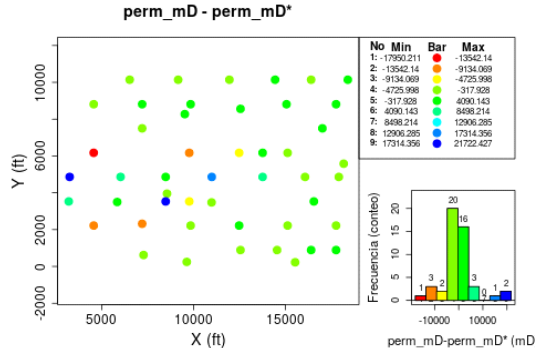


Figura 31: Mapa de la validación cruzada del modelo.

Validación cruzada del modelo

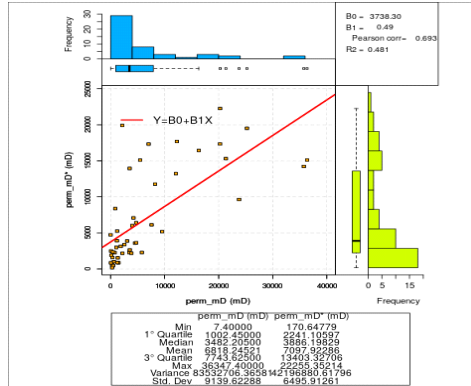


Figura 32: Valores estimados vs. valores reales.

Validación cruzada del modelo

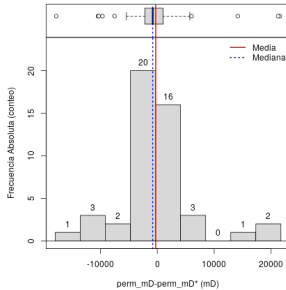


Figura 33: Histograma de las diferencias.

Estadígrafo	Valor
Muestras	48
Mínimo	-17753.81
1º cuartil	-2101.11
Mediana	-799.22
Media	-279.68
3º cuartil	985.77
Máximo	21526.03
Rango	39279.84
Rango intercuartil	3086.88
Varianza	43409713.43
Desviación estándar	6588.60
Simetría	1.13
Curtosis	7.07

Tabla 8: Estadística de las diferencias.

Referencias

[Christakos, 1984] Christakos, G. (1984).

On the problem of permissible covariance and variogram models.

Water Resources Research, 20(2):251–265.

[Mandelbrot et al., 1983] Mandelbrot, B., Mandelbrot, B., Freeman, W. H., and Company (1983).

The Fractal Geometry of Nature.

Einaudi paperbacks. Henry Holt and Company.

Agradecimiento especial

Al estudiante de doctorado M. en C. Daniel Vázquez Ramírez, por su desinteresado apoyo en la conversión de esta presentación del curso de Powerpoint a Latex con Beamer.

Siguiente tema: Estimación Univariada