Climate Variability and Change—Hydrological Impacts (Proceedings of the Fifth FRIEND World Conference held at Havana, Cuba, November 2006), IAHS Publ. 308, 2006.

Evaluación de la incertidumbre en la variabilidad espacial de la precipitación usando un modelamiento multiGaussiano

SOLANGE DUSSAUBAT[1], XIMENA VARGAS[1] & JULIAN ORTIZ[2]

1 *Departamento de Ingeniería Civil, Universidad de Chile, Blanco Encalada 2002, Santiago, Chile*
sdussaub@ing.uchile.cl

2 *Departamento de Ingeniería de Minas, Universidad de Chile, Av. Tupper 2069, Santiago, Chile*

Resumen En este trabajo se estudia el modelamiento de la incertidumbre, proponiendo el uso de distribuciones de probabilidad condicional a través del modelo multiGaussiano en una zona cordillerana de la VII Región del Maule, Chile. Para esto se utilizó información extraída del modelo de pronóstico del tiempo MM5-DGF de dos tormentas ocurridas en el año 2002. Los resultados de la media condicional y la varianza condicional en el área de estudio evidencian el carácter heterocedástico de la variable, lo que permite argumentar que la varianza no sólo depende de la configuración geométrica de los datos, sino que varía de acuerdo al valor estimado en cada sitio.

Palabras claves incertidumbre; modelo multiGaussiano; kriging; variabilidad espacial de la precipitación; varianza de estimación; heterocedasticidad; homocedasticidad

Evaluation of the uncertainty of spatial precipitation variability using a multiGaussian model

Abstract. The modelling of the uncertainty of spatial precipitation variability is studied in a mountainous zone at the Maule Region, Chile. Forecast data for two storms that occurred in 2002, which were extracted from the results of the meteorological model MM5-DGF, are used to recommend the use of a conditional probability distribution through the multiGaussian model. The heteroscedasticity of the variable is demonstrated through the results obtained for the conditional mean and the conditional variance in the region, which shows that the variance depends on the geometric configuration of the data and, also, on the estimated value at each site.

Key words uncertainty; multiGaussian model; kriging; precipitation spatial variability; estimation variance; heteroscedasticity; homoscedasticity

INTRODUCCION

En la actualidad, resulta de gran interés contar con valores confiables de la precipitación, puesto que constituyen el cimiento de múltiples propósitos en la gestión misma del recurso hídrico y, por ende, es interesante evaluar la incertidumbre existente en su determinación.

Al realizar un proceso de estimación a través de métodos geoestadísticos se trata de predecir el valor de una variable partiendo de una malla de observaciones. Debido a que la varianza de estimación del kriging depende de la configuración geométrica de los sitios con información y no del valor registrado, muchos autores plantean que ésta es una excelente medida de incertidumbre, y que es una buena herramienta para el mejoramiento, diseño u optimización de una red de estaciones. Sin embargo, en este trabajo se cuestiona su uso como la mejor forma para cuantificar la incertidumbre, esencialmente porque la homocedasticidad es una situación raramente encontrada en la práctica.

El índice de incertidumbre más común es el intervalo de confianza, sin embargo, difícilmente puede ser usado directamente para toma de decisiones, principalmente porque es expresado como un intervalo simétrico y porque, además, proporciona información muy limitada. En este sentido, la función de distribución de probabilidad proporciona considerablemente más información que un simple intervalo de confianza, ya que, entre otras cosas, puede ser usada para describir intervalos de confianza asimétricos o para describir la probabilidad de exceder algún valor umbral (Isaaks & Srivastava, 1989). Como modelo de incertidumbre permite contar con una herramienta de múltiples aplicaciones para la toma de decisión según se requiera en cada caso.

En este trabajo se estudia el modelamiento de la incertidumbre, propiamente tal, proponiendo el uso de distribuciones de probabilidad condicional a través del modelo multiGaussiano en una zona cordillerana de la VII Región del Maule, Chile.

EL MODELO MULTIGAUSSIANO

Se tienen N+1 variables $Z(x_0)$, $Z(x_1)$, …, $Z(x_N)$, y se transforman a $N + 1$ variables normales estandarizadas $Y(x_0)$, $Y(x_1)$, …, $Y(x_N)$, mediante una transformación adecuada. Sobre las variables estandarizadas, se establece la hipótesis básica del modelo, que consiste en suponer que el vector $(Y(x_0), Y(x_1), \ldots, Y(x_N))'$ tiene una distribución Gaussiana multivariada, y es caracterizado por la matriz de covarianzas σ_{ij}, $(i,j = 0, 1,\ldots, N)$. La distribución condicional de $Y(x_0)$ es, por ende, Gaussiana (Chilès & Delfiner, 1999):

$$Y(x_0) \mid \{Y(x_a) : \alpha = 1,\ldots,N\} = Y_{SK}^{*} + \sigma_{SK} K \tag{1}$$

donde Y_{SK}* es el estimador del kriging simple de $Y(x_0)$ a partir $\{Y(x_\alpha):\alpha = 1, \ldots N\}$; σ_{SK} es la desviación estándar del kriging asociada; K es una desviación normal estándar independiente (error del kriging estandarizado).

La información condicionante se concentra en el estimador de kriging simple, Y_{SK}*. El supuesto del modelo multiGaussiano es prácticamente incomprobable, por lo que se verifica que las distribuciones bivariadas sean Gaussianas (Chilès & Delfiner, 1999).

Las etapas para aplicar el kriging multiGaussiano son las siguientes:

1. Distribución representativa: se debe contar con un histograma que sea representativo, vale decir, que no exista un efecto de clusters o grupos de datos.
2. Transformación de los datos originales a una distribución normal estandarizada: la primera condición para que la función aleatoria $Y(x)$ sea normalmente multivariada, es que su función distribución acumulada univariada (histograma) sea normal estandarizada, vale decir, de media cero y varianza uno (Deutsch & Journel, 1998).
3. Verificación de la hipótesis multiGaussiana: la transformación de los datos a una distribución normal estandarizada es una condición necesaria pero no suficiente para que los valores $Y(x)$, distribuidos espacialmente sean normales multivariados. Luego, se debe verificar que la función de distribución acumulada de cualquier par de valores $Y(x)$, $Y(x+h)$, $\forall x$, $\forall h$, sea bivariada normal (Deutsch & Journel, 1998).

   Hay varias maneras de chequear que la distribución de dos puntos (bivariada) de datos sea normal $\{y(x_\alpha),\alpha = 1,\ldots N\}$. Un método consiste en verificar que el valor de la función de distribución bivariada experimental de cualquier conjunto de pares de datos separados por el mismo vector h $\{y(x_\alpha),y(x_\alpha + h),\alpha = 1, \ldots N(h)\}$ sea comparable con el modelo teórico biGaussiano (ver Goovaerts, 1997). Lo anterior se traduce en comparar variogramas de indicadores de los datos con los variogramas de indicadores de una variable multiGaussiana con la misma función de covarianza.

   Otro test consiste en verificar que bajo la hipótesis biGaussiana, el variograma de orden 1 o madograma, $\gamma_Y^{(1)}(h)$, es proporcional a la raíz del variograma usual $\gamma_Y(h)$, vale decir, $\sqrt{\gamma_y(h)}\big/\gamma_Y^{(1)}(h) = \sqrt{\pi}$, independiente del valor de h (Emery, 2000).

   Finalmente, se chequea gráficamente que la función de distribución acumulada (bivariada) es biGaussiana. Para esto, se grafican las características de una función de densidad conjunta para dos variables aleatorias, X e Y, y las marginales asociadas.
4. Estimación para determinar las distribuciones condicionales: en el caso multiGaussiano, la distribución es Gaussiana, el estimador de kriging simple se identifica con la media condicional y la varianza de kriging simple con la varianza condicional.
5. Transformación inversa de las distribuciones condicionales: para obtener las distribuciones condicionales de la variable original es necesario realizar numéricamente la transformación inversa de la distribución condicional Gaussiana.

APLICACION

Área de estudio

La zona escogida para evaluar la incertidumbre en la variabilidad espacial de la precipitación, es la zona cordillerana de la Séptima Región de Chile, específicamente entre los paralelos 35.5° y 36.7°

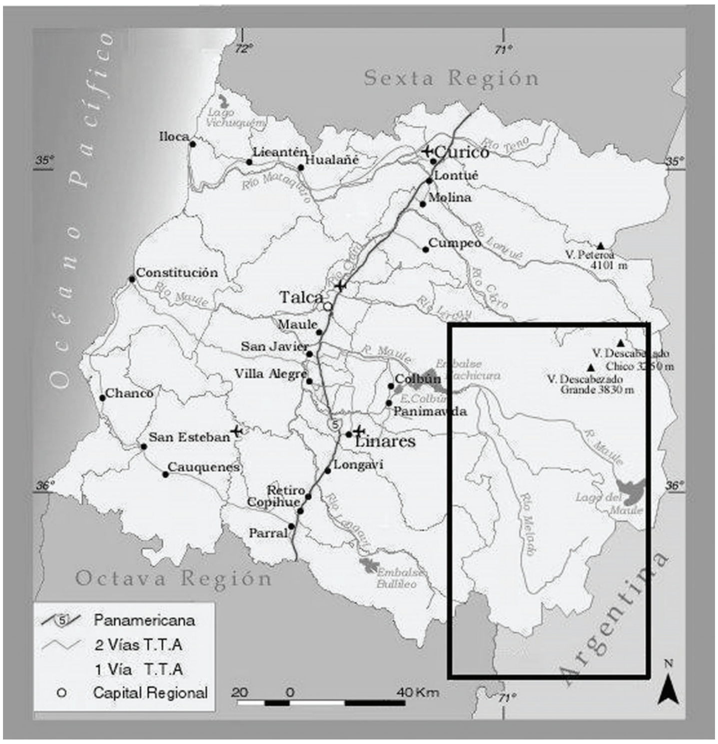

Fig. 1 Mapa de la Séptima Región del Maule, Chile. Se destaca zona de estudio.

Latitud Sur, y los meridianos 70.4° y 71.2° Longitud Oeste. Esta región abarca la mayor parte de la zona cordillerana de la hoya del río Maule, cubriendo un área aproximada de 11 880 km^2 (ver Fig. 1).

Información disponible

Debido a que dentro del área de estudio existen sólo dos estaciones pluviográficas y una de ruta de nieve, lo que resulta insuficiente para caracterizar en forma adecuada la variabilidad espacial de la precipitación, se optó por utilizar datos generados con el modelo regional de pronóstico del tiempo MM5. Este modelo numérico, desarrollado en la Universidad de Pensilvania, ha sido implementado por el departamento de Geofísica de la Universidad de Chile (DGF) en forma operacional desde Abril del 2002 como un apoyo al pronóstico del tiempo en la región central del país. Este modelo, inicializado diariamente, pronostica la circulación atmosférica en Chile central tomando en cuenta un horizonte temporal de tres días. El modelo opera con tres dominios anidados, denominados Dominio 1, 2 y 3, con resoluciones de 135, 45 y 15 [km]. El MM5-DGF permite conocer la distribución temporal y espacial de una serie de variables meteorológicas, de las cuales se obtiene como subproducto la precipitación.

Según la resolución del modelo de pronóstico MM5-DGF en el Dominio 3, el área de estudio corresponde a una grilla de 84 elementos MM5-DGF, siete en dirección Este por doce en dirección Norte.

Para evaluar la incertidumbre en la variabilidad espacial de la precipitación se utilizó información disponible de dos tormentas ocurridas durante el año 2002. Para contemplar sólo la variable espacial se decidió utilizar los datos de precipitación total ocurrida por evento.

Uso de kriging MultiGaussiano: Tormenta del 13 al 18 de Septiembre del año 2002

La transformación de los datos originales a una distribución normal estandarizada se realizó directamente a través del módulo "nscore.exe" del paquete geoestadístico GSLIB (Deutsch & Journel, 1998), que calcula la probabilidad acumulada de los datos usando la fórmula de Hazer (ver Fig. 2). Los resultados gráficos de los tests biGaussianos (Fig. 3) indican que es adecuado aceptar la hipótesis multiGaussiana.

Luego, se procede a aplicar un kriging simple a los datos transformados. Como se mencionó anteriormente, el estimador de kriging simple se identifica con la esperanza condicional y la varianza

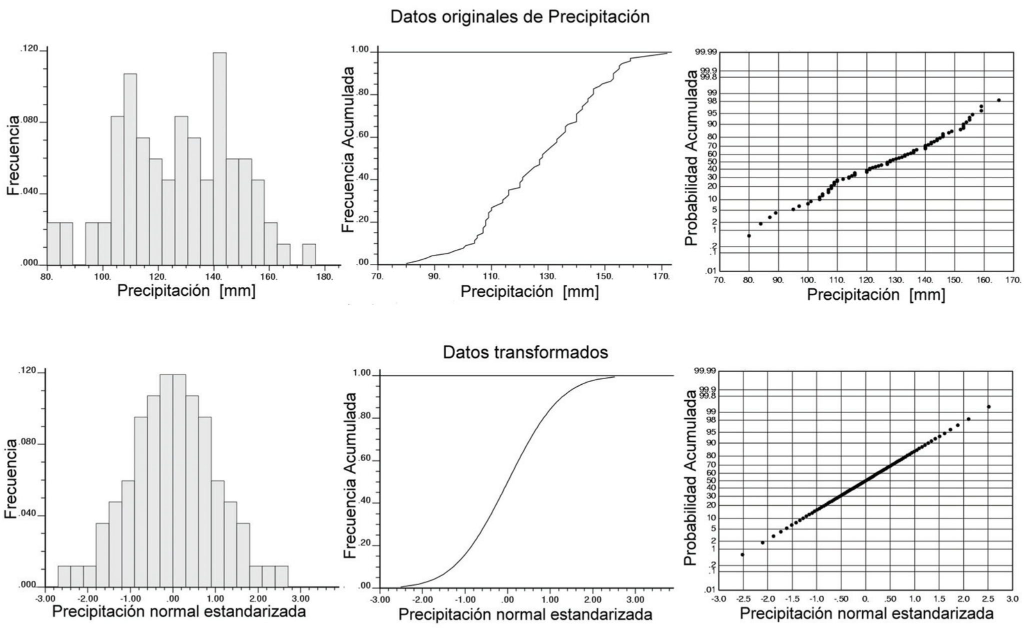

Fig. 2 Transformación de los datos originales a una normal estandarizada. Tormenta del 13 al 18 de Septiembre del 2002.

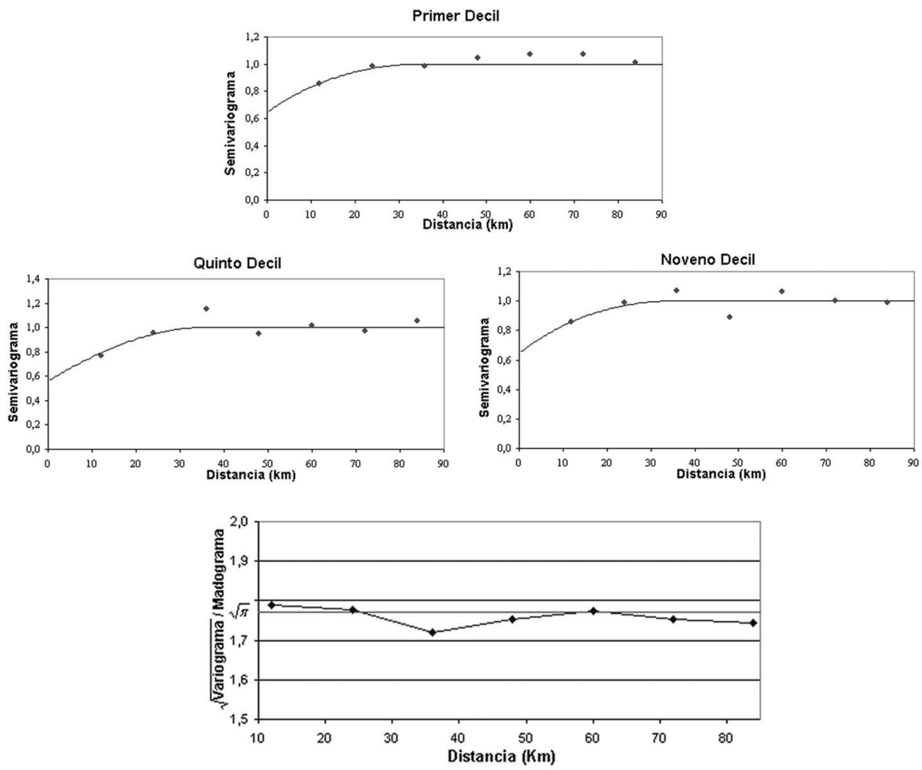

Fig. 3 Tests de binormalidad – Tormenta del 13 al 18 de Septiembre del 2002.

del kriging se identifica con la varianza condicional. Los resultados obtenidos son presentados en la Fig. 4. Aquí se destaca el carácter homocedástico de la varianza condicional, la cual no depende de los valores medidos, sino que exclusivamente de la configuración geométrica de los datos y del modelo de covarianza adoptado.

Para obtener las distribuciones condicionales de la variable original (precipitación), se debe realizar la transformación inversa de la distribución condicional Gaussiana definida en el espacio *Y*. Para esta transformación se utilizó el programa "postMG" (Ortiz *et al*., 2004), el cual realiza un

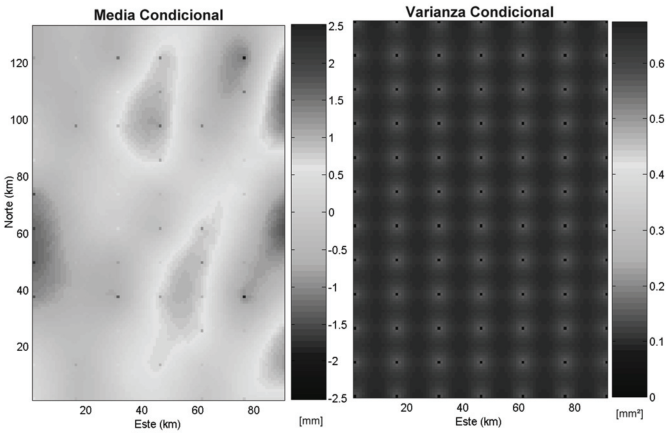

Fig. 4 Media condicional y varianza condicional de la variable transformada. Tormenta del 13 al 18 de Septiembre del 2002.

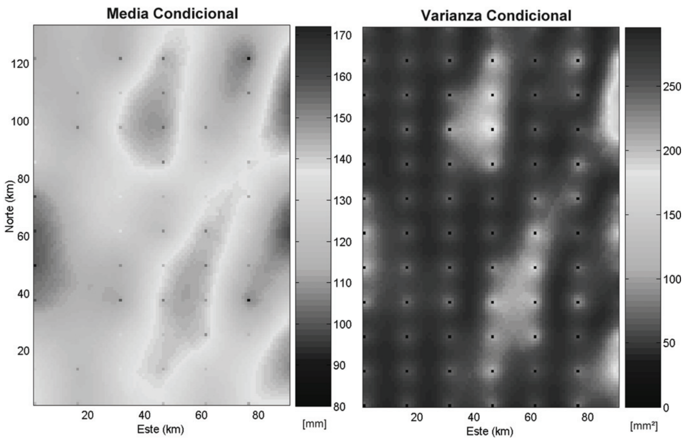

Fig. 5 Media condicional y varianza condicional de la variable original (precipitación).Tormenta del 13 al 18 de Septiembre del 2002.

muestreo regular de los cuantiles con una transformación inversa de cada uno, usando la relación entre las distribuciones globales (a priori).

Finalmente, después de la transformación respectiva se obtiene la media condicional y la varianza condicional de la función de distribución acumulada en el espacio de la variable original (precipitación), siendo ambas presentadas de forma “local” en la Fig. 5. En esta imagen se ve que la varianza condicional es heterocedástica, con una clara variación de acuerdo a los datos registrados por el modelo MM5-DGF. La incertidumbre tiende a ser menor en las zonas extremas, vale decir, donde se registran los valores más altos y más bajos de precipitación. Por el contrario, la incertidumbre aumenta en las zonas en donde los valores estimados son cercanos a la media global de la muestra.

Uso de kriging multiGaussiano: Tormenta del 19 al 24 de Julio del 2002

Procediendo en forma análoga al caso anterior, se verifica la hipótesis multiGaussiana y se aplica el modelamiento respectivo. Los principales resultados se muestran en la Fig. 6 en donde se observa la media condicional y la varianza condicional de la variable original (precipitación).

En la Fig. 7 se presenta la varianza condicional en función de la media en el espacio de la variable original para las dos tormentas analizadas. En esta figura se observa que la varianza condicional en ambas tormentas es mayor para los valores cercanos al promedio.

CONCLUSIONES

Los resultados obtenidos ponen en evidencia el carácter heterocedástico de la variable precipitación. Los lugares que presentan mayor incertidumbre son los sitios intermedios entre zonas homogéneas de mayor o menor precipitación.

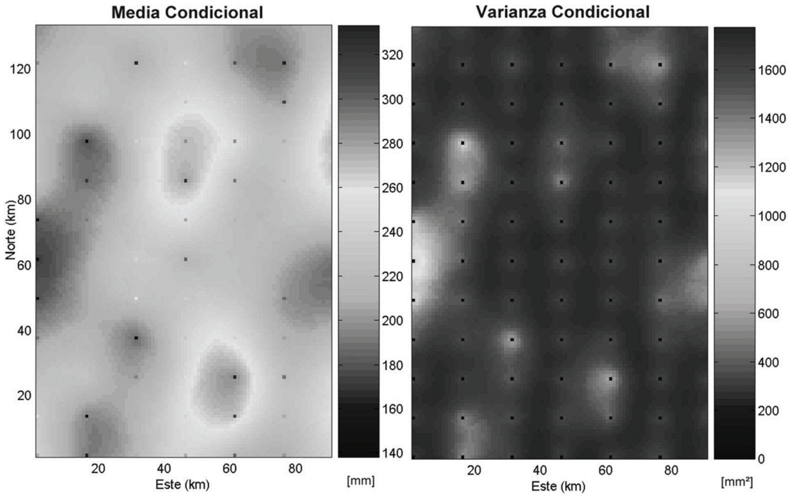

Fig. 6 Media condicional y varianza condicional de la variable original (precipitación). Tormenta del 19 al 24 de Julio del 2002.

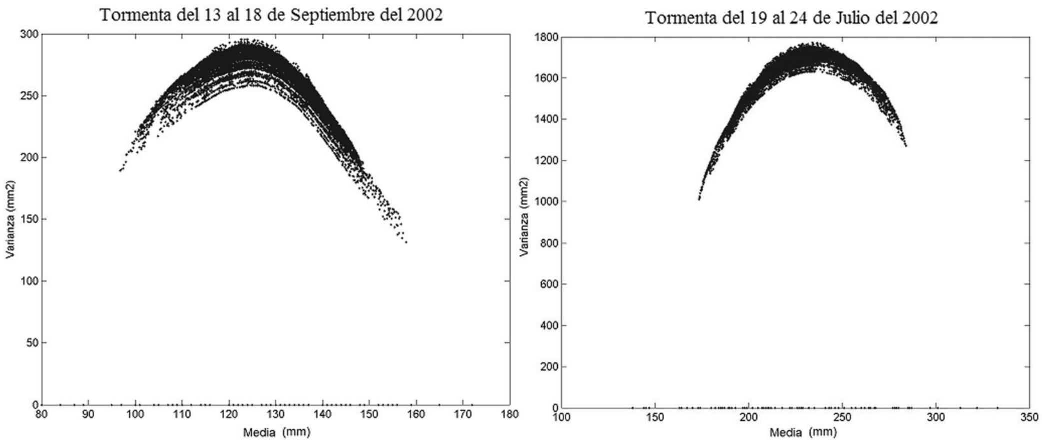

Fig. 7 Media condicional versus varianza condicional. Izquierda: Tormenta del 13 al 18 de Septiembre del 2002. Derecha: Tormenta del 19 al 24 de Julio del 2002.

Todo lo anterior permite argumentar que la incertidumbre no sólo depende de la configuración geométrica de los sitios con información, sino que varía de acuerdo al valor estimado en cada punto.

Es importante destacar, que antes de aplicar el kriging multiGaussiano, es fundamental verificar que los datos transformados cumplan con la hipótesis biGaussiana. Se puede destacar que las dos tormentas analizadas presentan un comportamiento binormal que permite aceptar la hipótesis planteada.

Se concluye finalmente que el uso de funciones de distribución de probabilidad condicional a través de un modelamiento multiGaussiano constituye una poderosa herramienta que permite evaluar la incertidumbre de la precipitación dentro de un área.

REFERENCIAS

Chilès, J. P. & Delfiner, P. (1999) *Geostatistics: Modeling Spatial Uncertainty*. Wiley-Interscience, New York, USA.

Deutsch, C. V. & Journel, A.G. (1998) *GSLIB: Geostatistical Software Library and User's Guide* (second edn). Oxford University Press, New York, USA.

Emery, X. (2000) *Geoestadística Lineal.* Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, Departamento Ingeniería de Minas.

Goovaerts, P. (1997) *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New-York, USA.

Isaaks, E. H., & Srivastava, R. M. (1989) *An Introduction to Applied Geostatistics*. Oxford University Press, New York, USA

Ortiz, J. M., Leuangthong, O. & Deutsch, C. V. (2004) A multiGaussian approach to assess block grade uncertainty. CIM Conference and Exhibition, Edmonton 2004, May 9 to 12.