



## Herramientas de Análisis Estadístico y Espacial aplicadas en la separación Línea Base – Anomalía Geoquímica. Implicaciones en Exploración de Recursos Naturales. Parte 1: Revisión metodológica

## Spatial and Statistical Analysis Tools applied in Geochemical Baseline – Anomaly discrimination. Implications for Natural Resources Exploration. Part 1: Methodological review

OSCAR JAVIER GARCÍA – CABREJO<sup>1</sup>

GIOVANNI MORENO-SÁNCHEZ<sup>2</sup>

<sup>1</sup>Departamento de Geociencias, Universidad Nacional de Colombia, Bogotá.

E-mail: [oscar.garciac@gmail.com](mailto:oscar.garciac@gmail.com)

<sup>2</sup>Departamento de Geociencias, Universidad Nacional de Colombia, Bogotá.

Instituto Colombiano de Geología y Minería – INGEOMINAS

E-mail: [gmorenos@unal.edu.co](mailto:gmorenos@unal.edu.co)

GARCÍA-CABREJO, O. & MORENO-SÁNCHEZ, G. (2006): Herramientas de Análisis Estadístico y Espacial aplicadas en la separación Línea Base – Anomalía Geoquímica. Implicaciones en Exploración de Recursos Naturales. Parte 1: Revisión metodológica.- GEOLOGÍA COLOMBIANA, 31, pp. 27-38, 10 Figs., Bogotá.

### RESUMEN

Uno de los principales problemas relacionados con la exploración geoquímica consiste en la determinación del valor de umbral, con el cual se discrimina tanto la línea base como la anomalía geoquímica. En el presente trabajo se realiza una revisión metodológica de diferentes aproximaciones para la discriminación de línea base – anomalía en estudios geoquímicos y se explican los fundamentos de los mismos, así como sus requerimientos, ventajas y desventajas. Los métodos presentados en este trabajo incluyen: Métodos estadísticos: Promedio  $\pm$  2 desviaciones estándar, cuantil 90 y Estadística espacial U; Métodos gráficos: Diagramas cuantil – cuantil y papeles de probabilidad; Método fractal: concentración – área; Métodos geoestadísticos: krigueaje factorial y krigueaje robusto.

Palabras Clave: *Anomalía geoquímica, correlación espacial, distribución de probabilidad, exploración geoquímica, fractal, geoestadística, muestreo, línea base.*

### ABSTRACT

A key point in geochemical exploration concerns the determination of threshold values which in turn defines the geochemical baseline – anomaly values. Here we deal a methodological review of different points of view in order to gather discrimination between baseline – anomaly in geochemistry with an explanation of them, his requirements, advantages and disadvantages. The methods here included are; Statistical methods: Average  $\pm$  2 standard deviation, 90 quantile and U spatial statistics; Graphical methods: Q – Q plot, P – P plot; Fractal method: Concentration – area; Geostatistical methods: factorial – robust kriging.

Key words: *Baseline, fractal, geochemical anomaly, geochemical exploration, geostatistics, probability distribution, sampling, spatial correlation.*

## INTRODUCCION

Uno de los principales problemas relacionados con la exploración geoquímica consiste en la determinación del valor de umbral, con el cual se discrimina tanto la línea base como la anomalía geoquímica. Esta labor es de vital importancia en el caso de la búsqueda de mineralizaciones para la definición de los denominados índices de favorabilidad de mineralización (Luo & DIMITRAKOPoulos 2003) o la detección de tendencias inusuales en aplicaciones ambientales. Para la determinación de este valor se realizan muestreos geoquímicos (sedimentos activos, suelos, roca) para adquirir información sobre la variación de las variables geoquímicas de interés y efectuar el correspondiente análisis de los resultados. Sin embargo cuando se procede a realizar este análisis el geólogo se enfrenta con el problema que no hay un único método para la determinación de la línea base y por lo tanto puede existir más de un valor para el mismo conjunto de datos geoquímicos.

El presente trabajo se enfoca en una revisión metodológica de diferentes aproximaciones para la discriminación de líneas base – anomalías en estudios geoquímicos y se explican los fundamentos de los mismos, así como sus requerimientos, ventajas y desventajas. Los métodos presentados en este trabajo incluyen:

- Métodos estadísticos. Promedio  $\pm$  2 desviaciones estándar y cuantil 90.
- Métodos gráficos. Diagramas de cuantil – cuantil y papeles de probabilidad.
- Métodos fractales. Concentración – área.
- Métodos geoestadísticos. Krigeaje factorial y Krigeaje robusto.
- Estadística espacial U.

También se propone una estrategia sencilla para la definición de las áreas de interés, que integra los resultados de estos métodos usando un Sistema de Información Georeferenciada. Estas aproximaciones se aplican en un caso de estudio de la plancha 5-09 (escala 1:500000) la cual cubre el sector central de Colombia. Otro de los problemas abordados en el presente trabajo corresponde a la variación con la escala de los valores de línea base – anomalía determinados por estos métodos. Para resolver esta inquietud se aplicaron estas metodologías en una zona con un muestreo más detallado dentro de la plancha 5-09 (Plancha 208 – Villetá, escala 1:100000).

La estructura del presente trabajo es la siguiente. En la primera parte se revisan algunos conceptos fundamentales sobre exploración geoquímica y el modelo que se asume para los datos geoquímicos. Posteriormente se pasa a una explicación sencilla pero completa de los métodos de discriminación de líneas base – anomalías, para seguir con los casos de aplicación, la discusión de resultados y las conclusiones preliminares.

## EXPLORACIÓN GEOQUÍMICA

La línea base se define como la concentración promedio que tiene un elemento en una roca de un tipo particular. Este valor se puede determinar a partir de los resultados de los análisis químicos de muestras de roca. El objetivo central de la exploración geoquímica consiste en la determinación de valores que sean más altos que la línea base ya que estos indicarían un enriquecimiento con respecto a esta y por lo tanto la presencia de un cuerpo mineralizado. A estos valores se les denomina anomalía geoquímica y resulta entonces crítica su correcta determinación en un estudio geoquímico. El valor que separa la línea base de la anomalía se denomina umbral y generalmente es definido a partir de un cuidadoso análisis estadístico de la información geoquímica disponible.

A nivel geológico, la anomalía geoquímica aparece cuando se adiciona algún componente a las rocas a partir de una fuente externa (Fig. 1). Dependiendo de las características del proceso geológico esto ocasiona que se presenten valores contrastantes entre las concentraciones originales del elemento en esa litología y las producidas por el nuevo evento, hecho que se ve reflejado en la presencia de distribuciones estadísticas de tipo multimodal para los elementos estudiados (Fig. 2). El objetivo del análisis de la información obtenida durante la exploración geoquímica consiste en detectar estas distribuciones superpuestas, con lo cual se identificarían los diferentes eventos geológicos que han ocurrido y que puedan haber contribuido al enriquecimiento de algunos elementos de interés. La identificación de dichas distribuciones requiere de un análisis estadístico detallado y de allí que la gran

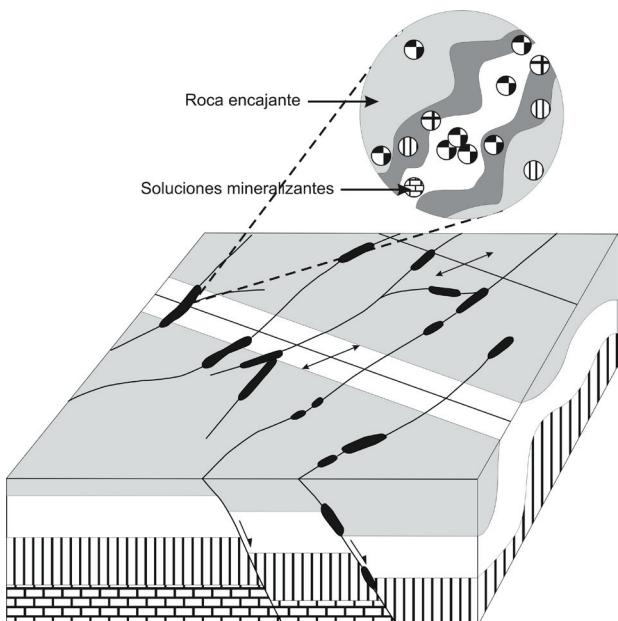
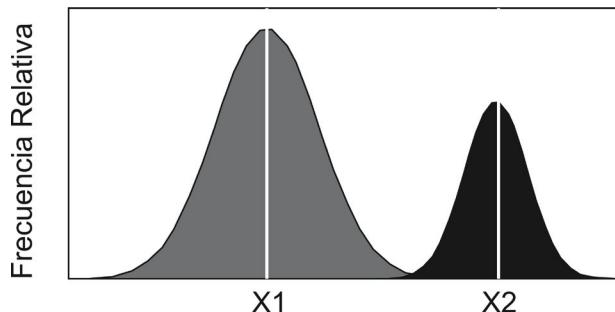


Fig. 1. Concepto de línea base y anomalía geoquímica (Modificado de PETERS 2001).

mayoría de métodos de separación de anomalías geoquímicas sean estadísticos por naturaleza.



**Fig. 2. Distribución multimodal.  $X_1$ : promedio de la distribución de probabilidad de la población 1.  $X_2$ : promedio de la distribución de probabilidad de la población 2.**

## MÉTODOS DE SEPARACIÓN DE ANOMALÍAS GEOQUÍMICAS

### Métodos Estadísticos

En el caso que se tengan pocos datos de la variable de interés, el procedimiento de separación de la anomalía puede dificultarse, con lo cual se hace necesario que el analista aplique criterios adicionales. En este caso es donde resultan útiles los métodos estadísticos. Estos procedimientos permiten llevar a cabo la definición de la anomalía en términos de alguna característica de la distribución de los datos como tal, evitando de esta forma la determinación del umbral por inspección visual:

$$\text{Umbral} = \text{Característica estadística} \quad (1)$$

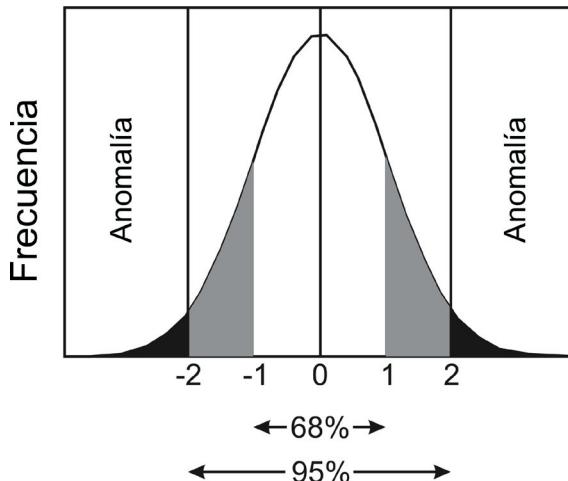
Dentro de esta categoría se pueden mencionar dos criterios. El primero de ellos define el umbral en términos del promedio y la desviación estándar de la variable de interés  $Z$  (LEPELTIER 1969):

$$\text{Umbral} = \bar{Z} + 2 \cdot \sigma_Z \quad (2)$$

Donde  $\bar{Z}$  es el promedio y  $\sigma_Z$  es la desviación estándar de la variable  $Z$  respectivamente, con lo cual se definen los rangos (Fig. 3):

$$\begin{aligned} Z \leq \bar{Z} + 2 \cdot \sigma_Z & \text{ Para la línea base} \\ Z \geq \bar{Z} + 2 \cdot \sigma_Z & \text{ Para la anomalía} \end{aligned} \quad (3)$$

El valor de umbral definido por la ecuación 2 y los rangos asociados (ecuación 3) se basan en la suposición que la variable geoquímica sigue una distribución normal y que los valores mayores al umbral definido por la ecuación 2 estarían controlados por procesos diferentes a los que originan los valores más cercanos al promedio. De



**Fig. 3. Criterio estadístico para la definición de anomalía.**

acuerdo con KOTTEGODA & Rozzo (1993) el valor de umbral determina como muestras anómalas aproximadamente el 2.5% de los datos, que son los que resultan de interés en las labores de exploración geoquímica.

El otro criterio estadístico que se utiliza habitualmente corresponde al del cuantil 90 o cuantil 95. Los cuantiles son los valores de la variable que corresponden a un valor específico de probabilidad acumulada.

$$\text{Umbral} = \begin{cases} z_{0,90} = F^{-1}(0,90) \\ z_{0,95} = F^{-1}(0,95) \end{cases} \quad (4)$$

Donde  $F^{-1}$  corresponde a la función acumulada de probabilidad inversa, la cual se determina a partir de la información en estudio. A partir de la ecuación 4 y dependiendo del cuantil que se haya seleccionado se obtienen los intervalos:

$$\begin{aligned} Z < z_{\text{cuantil}} & \text{ Para la línea base} \\ Z \geq z_{\text{cuantil}} & \text{ Para la anomalía geoquímica} \end{aligned} \quad (5)$$

En este caso el método no asume ninguna forma paramétrica para la distribución de la variable geoquímica, pero se sigue suponiendo que los valores altos son controlados por procesos geológicos diferentes a los valores medios y bajos.

La gran ventaja de estas metodologías estadísticas corresponde al hecho que se basan en criterios objetivos de la distribución del elemento particular, así como que se pueden emplear en todos los casos (aún si se tienen pocas muestras). Sin embargo como desventaja se puede mencionar que los valores obtenidos para el umbral y que separan la línea base de la anomalía no tiene correspondencia directa con procesos geológicos particulares y en consecuencia se corre el riesgo que en muchos casos no se separen grupos geológicos diferentes.

## Métodos Gráficos

Estos métodos se basan en la inspección visual de gráficos estadísticos en donde es posible diferenciar grupos con características definidas. El criterio que se emplea en este caso es la identificación de quiebres de pendiente, con lo cual quedaría definido el valor de umbral y que separaría la línea base de la anomalía. Generalmente para cumplir este propósito se emplean diagramas cuantil – cuantil, de probabilidad acumulada o papeles de probabilidad. Este tipo de diagramas se emplean para determinar si una distribución se ajusta a un modelo pre-determinado (normal, lognormal, gumbel, entre otros) o para comparar distribuciones de probabilidad (KOTTEGODA & Rozzo 1993). En los diagramas cuantil – cuantil se grafican las posiciones de los cuantiles de la distribución de datos con respecto a los de otra distribución, la cual en este caso corresponde a una distribución teórica (normal o lognormal). En el caso de los papeles de probabilidad se grafican los valores de la variable de interés contra su probabilidad acumulada y si se ubican en una línea recta de pendiente positiva, la distribución de los datos se ajusta al modelo correspondiente al papel (normal, lognormal, gumbel, etc.). Los quiebres de pendiente en los diagramas representan valores que no se ajustan a la distribución global y que pueden estar caracterizados por parámetros diferentes (Fig. 4).

Estos métodos tienen la ventaja de ser de fácil utilización, ya que solo requiere el manejo de un paquete de software que permita la elaboración del diagrama y además el valor de umbral se puede identificar fácilmente con lo cual deja de ser subjetivo o interpretativo. Sin embargo como desventaja se puede mencionar, que pueden existir ocasiones en las que la separación línea base – anom-

alia no se pueda realizar ya sea por un bajo número de muestras o porque no existen grupos con características estadísticas contrastantes que se vean reflejados en los diagramas.

## Estadística Espacial U

En esta metodología se combinan tanto elementos estadísticos de la variable de estudio como las características espaciales de la misma. Este método se basa en la suposición que en el área de estudio existen por lo menos dos zonas, y en cada una de ellas la distribución del elemento en estudio es diferente. Estas zonas corresponden directamente a la línea base y a la anomalía. Así mismo la discriminación entre línea base y anomalía se puede mejorar si se incorpora la componente espacial y se hace comparación entre una muestra y sus vecinas, ya que se pueden resaltar condiciones locales anómalas.

El procedimiento para la determinación de la anomalía geoquímica usando esta metodología es el siguiente (Fig. 5):

Para cada muestra  $\alpha$  ubicada en el punto  $u_\alpha$  y que presenta una concentración de un elemento particular denominada  $Z(u_\alpha)$  se define un vecindario circular de radio  $r$  y se determinan las muestras cercanas que caen en este.

Para cada una de las muestras cercanas se determina el valor de la función:

$$\mu_\alpha(u_\alpha) = \frac{r - d(u_\alpha, u_\beta)}{r} \quad (6)$$

Donde  $d(u_\alpha, u_\beta) \leq r$  corresponde a la distancia de sepa-

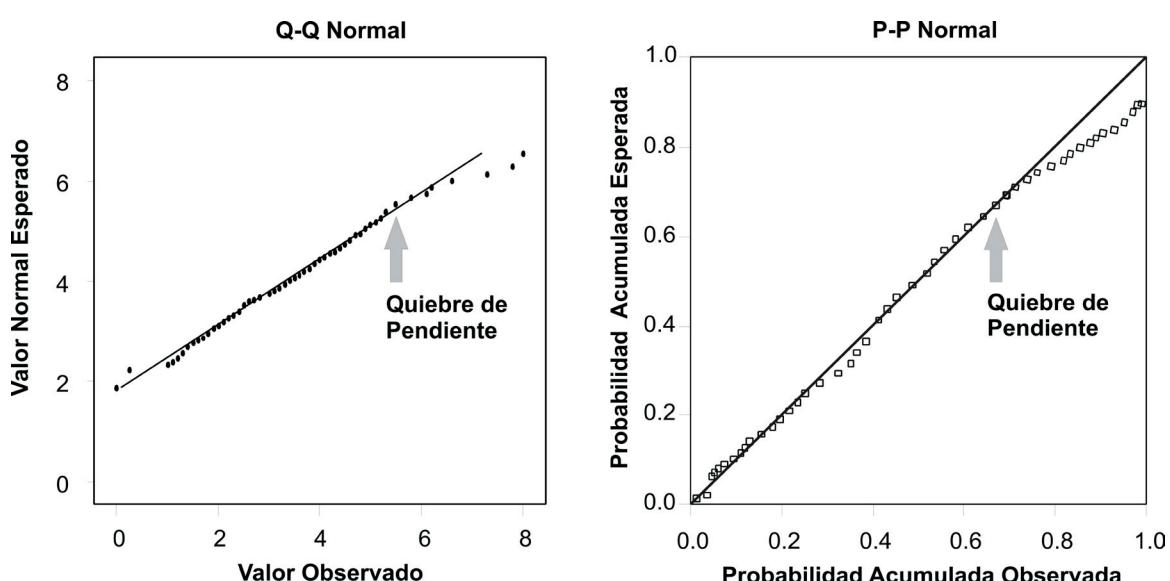


Fig. 4. Diagrama cuantil – cuantil y papel de probabilidad.

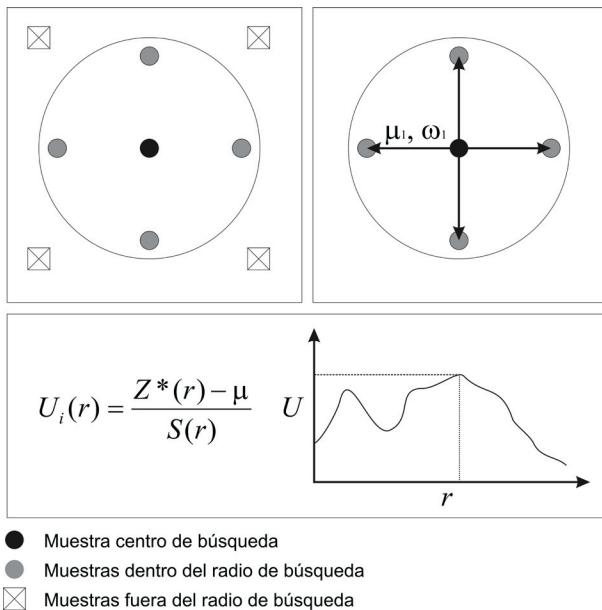


Fig. 5. Estadística espacial U.

ración entre la ubicación  $u_\alpha$  y  $u_\beta$ . Esta función asigna una ponderación de acuerdo a la distancia de separación, de tal forma que las muestras más cercanas reciben un mayor peso que las lejanas. Esta función de ponderación tiene las siguientes características:

- $0 \leq \mu(u_\alpha) \leq 1$
- $\mu(u_\alpha) = 1$  Si  $\alpha_i = \alpha_j$
- $\mu_{\alpha_j} = 0$  Si  $\alpha_j \in N(\alpha, r)$

Para cada muestra  $\alpha$  se determina la estadística U como un promedio ponderado de las concentraciones

$$U_i(r) = \frac{\bar{Z}(r) - \mu}{S(r)} \quad (7)$$

$$\begin{aligned} \bar{Z}(r) &= \sum_{\alpha=1}^N w_j(r) Z(u_\alpha) \\ S(r) &= \sqrt{\sum_{\alpha=1}^N w_j^2(r)} \\ w_j(r) &= \frac{\mu(\alpha_j)}{\sum_{j=1}^N \mu(\alpha_j)} \end{aligned} \quad (8)$$

Donde  $\mu$  corresponde a la media de la concentración del elemento en particular para toda el área de estudio (usando todos los valores  $Z_i$ ). Para cada muestra  $i$  se puede definir la estadística  $U_i^* = U_i(r_0)$  para un valor específico de  $r_0$  de tal forma que cumpla la condición:

$$|U_i(r_0)| = \max |U_i(r)| \quad (9)$$

De acuerdo con (CHENG *et al.* 1996) el valor de  $U_i^*$  para la muestra  $i$  tiene una distribución normal y para el caso del grupo A:

$$U_i \sim N \left( \frac{\mu_A - \mu}{S_i(r_0)}, \frac{\sigma_A}{\sigma} \right) \quad (10)$$

Mientras que para el grupo B:

$$U_i \sim N \left[ \frac{\mu_B - \mu}{S_i(r_0)}, \frac{\sigma_B}{\sigma} \right] \quad (11)$$

Donde  $\mu$  y  $\sigma$  corresponden a los valores de media y desviación estándar para cada uno de los grupos. En el caso que el grupo A se encuentre caracterizado por un valor promedio  $\mu_A$  mayor que el promedio global, el valor de la estadística U será positivo y este grupo se puede asociar directamente con la anomalía geoquímica. En el caso que el valor de la estadística U sea negativo, el grupo se puede asociar con la línea base y de esta forma se pueden clasificar las muestras como pertenecientes a uno u otro grupo. Es de aclarar que este criterio no es definitivo para la discriminación de línea base y anomalía y que la selección del umbral de la estadística U que permite realizar dicha separación se debe realizar de acuerdo a un criterio estadístico o gráfico como los explicados anteriormente.

La ventaja del método de la estadística U corresponde al hecho de tener en cuenta en el procedimiento de reconocimiento y separación de la anomalía geoquímica la relación entre una muestra y las muestras cercanas. Esto es un aspecto muy importante si se tiene en cuenta que las zonas anómalas tienen una geometría particular dada por la geología y que pueden ser detectadas cuando las muestras se comparan con los valores medidos en litologías de características geoquímicas diferentes. Sin embargo, la metodología adolece de criterios objetivos que permitan la separación (ya que deben usarse criterios gráficos o estadísticos) y de acuerdo con CHENG *et al.* (1996) el error de clasificación de las zonas anómalas disminuye a medida que aumenta el número de muestras, por lo que resulta prioritario evaluar su aplicabilidad en situaciones con pocas muestras. Así mismo la distribución de la estadística U (ecuación 7) tiende a ser normal pero de forma asintótica, con lo cual esta característica se reproduce mejor si los datos se encuentran convenientemente transformados.

### Métodos Fractales

Los conjuntos fractales fueron creados por Benoit Mandelbrot a principios de los años 60, como una alternativa a la geometría tradicional o euclidiana para la representación de los patrones geométricos naturales. De acuerdo con MANDELBROT (1983) un conjunto fractal se define como un conjunto geométrico para el cual se cumplen dos condiciones:

- Poseen una dimensión que tiene un valor que no necesariamente es entero. Puede existir un conjunto o patrón geométrico que presente un valor de dimensión entre 1 y 2.
- Presentan invarianza bajo cambios de escala. Si se toma una parte del conjunto geométrico, esta de cierta forma es igual a todo el conjunto.

La primera condición resulta extraña si se tiene en cuenta que el sentido común indica que los objetos tienen dimensiones que son enteras. Un punto tiene una dimensión igual a 0, una línea, un cuadrado y un cubo tienen dimensiones iguales a 1, 2 y 3 respectivamente (Fig. 6). Sin embargo la dimensión tal y como se maneja habitualmente es un parámetro que mide la tendencia de un objeto a llenar el espacio. Si se tiene en cuenta esta definición resulta lógico que una línea como la presentada en la figura 7 tenga una dimensión mayor que la asignada a una línea recta, ya que tiende a llenar más espacio que esta última por su grado de irregularidad. Visto de otro modo la dimensión fractal es un parámetro que permite medir la irregularidad de los objetos naturales y por lo tanto resulta de interés en muchos campos del conocimiento ya que brinda una herramienta sencilla pero poderosa para caracterizar la forma de patrones que distan mucho de ser figuras regulares (TURCOTTE 2002).

Para un conjunto fractal se cumple la siguiente relación (DAVIS 2002):

$$N_i = Cr_i^{-D} = \frac{C}{r_i^D} \quad (12)$$

Donde  $N_i$  es el número de objetos con una dimensión característica  $r_i$ ,  $C$  es una constante de proporcionalidad y  $D$  es la dimensión fractal. El procedimiento para determinar el valor de la dimensión fractal consiste en superponer al conjunto de interés una cuadrícula y contar el número de cuadrados en los que hay presencia de una parte del conjunto. Cada cuadrado tiene un lado de longitud igual a  $r_i$ . Repetir este procedimiento para diferentes valores de  $r_i$ , graficar los resultados y obtener el valor de pendiente por regresión lineal simple (Fig. 7).

En el caso de los depósitos minerales se ha encontrado que la relación entre el tonelaje acumulado  $M$  para un tenor promedio  $C^*$  esta dado por (CARGILL 1981):

$$M = aC^{-E} \quad (13)$$

Donde  $E$  es un exponente dado. Nótese la similitud de la forma matemática entre las ecuaciones 12 y 13, con la diferencia que la última no involucra características geométricas del depósito sino características técnicas. En este caso se tiene una relación fractal entre dos propiedades que no son geométricas y la cual en la literatura especializada se denomina una relación o ley de potencia. Así mismo existen otras características asociadas a las ocurrencias minerales tales como su distribución espacial

que son leyes de potencia (CARLSON 1991, LI *et al.* 2003).

La justificación para la ocurrencia de este tipo de relaciones en la distribución de elementos en la corteza terrestre se puede explicar desde los modelos de cristalización fraccionada tales como el modelo de destilación de Rayleigh y el modelo cromatográfico (ALLEGRE & LEWIN 1995), y cuyas variantes se pueden aplicar para el entendimiento del origen de mineralizaciones. Para más detalles de la aplicación de la geometría fractal en Geología y

	Número (n)	Factor de Escala (s)	Relación
SEGMENTO	3	3	$3^1=3$
CUADRADO	9	3	$3^2=9$
CUBO	27	3	$3^3=27$

Fig. 6. Concepto de dimensión para figuras geométricas regulares.

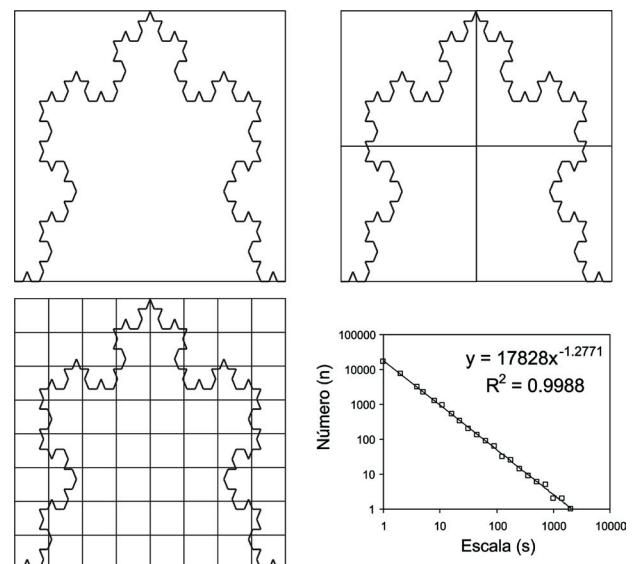


Fig. 7. Método de conteo de cajas. En este caso la curva de Koch presenta un  $D=1.2771$ .

Geofísica, incluyendo depósitos minerales ver TURCOTTE (1992).

### Método de la concentración - área

Esta metodología se basa en la determinación del área acumulada por encima de valores dados de concentración obtenidos a partir de mapas de contornos (CHENG 1999, SHEN & COHEN 2005). Cuando esta información se grafica en una escala doblemente logarítmica aparecen segmentos de línea recta con pendientes constantes (conjuntos fractales). En consecuencia la relación entre la concentración y el área acumulada es de tipo fractal y se encuentra descrita de la siguiente forma:

$$\begin{aligned} A(\rho < v) &\sim \rho^{-\alpha_1} & \text{Para la línea base} \\ A(\rho \geq v) &\sim \rho^{-\alpha_2} & \text{Para la anomalía geoquímica} \end{aligned} \quad (14)$$

Donde  $A(\rho)$  denota el área con valores de concentración mayor que el contorno de valor  $\rho$ ,  $v$  representa el valor de punto de quiebre y  $\alpha_1$  y  $\alpha_2$  son los exponentes de escalamiento. Por lo tanto la determinación del umbral y por ende la separación de la línea base y la anomalía geoquímica se puede realizar por la determinación de los puntos de quiebre en los gráficos de concentración contra área acumulada (Fig. 8). Generalmente se elige el quiebre de pendiente que ocurra para el valor más alto de concentración y se asume este como el valor umbral. Se puede emplear cualquier metodología para la elaboración de los mapas de contornos, ya sea Krigeaje (GONÇALVES *et al.* 2001) o inverso ponderado de la distancia (LIMA *et al.* 2003).

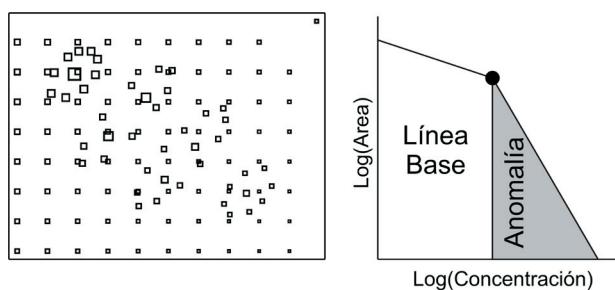


Fig. 8. Método fractal concentración – área.

Como ventaja de este método se puede mencionar que brinda una definición objetiva del valor de umbral y que en teoría funciona sin importar el tamaño de la muestra que se emplee, tanto como esta sea suficiente para la elaboración de un mapa de contornos. CARR (1995) sugiere que si se utiliza Krigeaje se deben tener al menos 30 muestras para la determinación del semivariograma y el modelo obtenido mediante inverso de la distancia ponderada con este nivel de información también se puede considerar satisfactorio. La desventaja principal de este

método es que requiere de la realización de post procesamiento del mapa de contornos, para lo cual se hace necesario contar con herramientas adicionales al software convencional de interpolación.

### Métodos Geoestadísticos

La Geoestadística se define como la aplicación de los métodos probabilísticos a fenómenos que varían en el espacio (ISAAKS & SRIVASTAVA 1989). Generalmente la variación espacial de los fenómenos es muy compleja para ser descrita por medio de modelos matemáticos sencillos y aunque pareciera que es completamente aleatoria, estos fenómenos tienen cierto grado de correlación espacial. La geoestadística brinda una herramienta para la caracterización de esta correlación espacial, la cual se denomina semivariograma (CARR 1995). El semivariograma es una función que mide el grado de disimilitud de una variable en función de la distancia de separación (CLARK 1979). Para determinar esta disimilitud de usan las diferencias de los valores elevadas al cuadrado, con lo cual el semivariograma se define:

$$\gamma(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^n [Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})]^2 \quad (15)$$

Otro de los problemas con el análisis de información espacial consiste en la reconstrucción de la variación espacial del fenómeno a partir de medidas puntuales, es decir la elaboración de mapas a partir de muestras. A nivel geostadístico la metodología del Krigeaje permite cumplir este objetivo, asegurando que las estimaciones en cada punto tienen la mínima varianza de error posible (ISAAKS & SRIVASTAVA 1989). El valor en un punto donde no se tenga muestra  $\mathbf{u}$  se puede determinar a partir de los valores de las muestras cercanas  $Z_i$ , por medio de un estimador:

$$Z^*(\mathbf{u}) = \sum_{i=1}^n \lambda_i Z_i \quad (16)$$

Donde  $\lambda_i$  son unos pesos de ponderación. Estos pesos se determinan por medio de la solución de un sistema de ecuaciones dado por:

$$\begin{cases} C(\mathbf{u}_\alpha - \mathbf{u}_\beta) \lambda + \mu = C(\mathbf{u} - \mathbf{u}_\alpha) \\ \sum_{i=1}^n \lambda_i = 1 \end{cases} \quad (17)$$

En donde  $C(\mathbf{u} - \mathbf{u}_\alpha)$  es la covarianza entre el punto de estimación y las muestras cercanas  $\alpha$ ,  $C(\mathbf{u} - \mathbf{u}_\beta)$  es la covarianza entre las muestras cercanas y  $\mu$  es el multiplicador de Lagrange. Estos valores de covarianza se determinan a partir del semivariograma (ecuación 15) por medio de:

$$C(\mathbf{h}) = \text{VAR} - \gamma(\mathbf{h}) \quad (18)$$

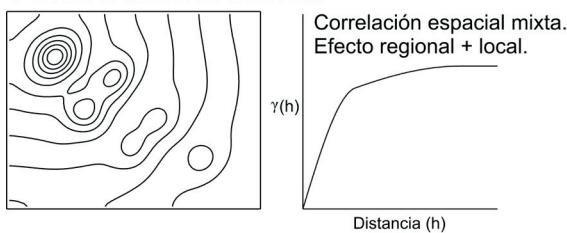
Donde VAR es la varianza de los datos. Para más detalle del procedimiento de Krigeaje ver las explicaciones dadas por ISAAKS & SRIVASTAVA (1989) y CARR (1995). Estas metodologías han sido aplicadas ampliamente en exploración geoquímica para la separación de anomalías por

JIMENEZ & CHICA-OLMO (1999), BATISTA *et al.* (2001), REIS *et al.* (2003, 2004) y GRUNFELD (2005).

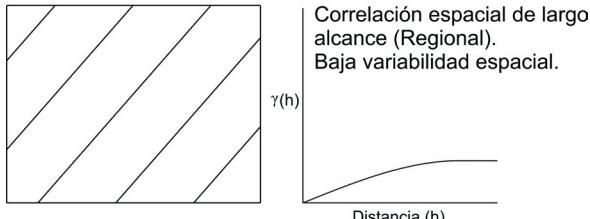
### Krigueaje Factorial

En muchas oportunidades los fenómenos espaciales se encuentran formados por una superposición de componentes que son independientes entre si, y cuya separación resultaría de interés (Fig. 9):

#### Concentración Elemental



#### Línea Base



#### Anomalía

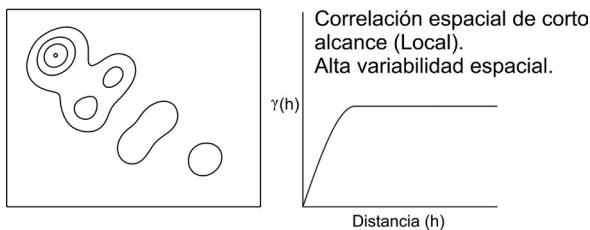


Fig. 9. Krigueaje Factorial.

$$Z(\mathbf{u}) = Z_0(\mathbf{u}) + Z_1(\mathbf{u}) + \dots + Z_K(\mathbf{u}) \quad (19)$$

Generalmente estas componentes se asocian a la señal y al ruido, o a componentes regionales y locales. Esta metodología se ha aplicado ampliamente en el filtrado de información sísmica (YAO *et al.* 1999) y en el procesamiento de imágenes (BOURGAULT 1994, WEN & SINDING-LARSEN 1997).

En el caso de exploración geoquímica estas componentes son la línea base y la anomalía y su separación se puede realizar por medio del método denominado Krigueaje factorial. Para realizar esto es necesario que el modelo de correlación espacial de la variable geoquímica se encuentre formado por la suma de los modelos de correlación de cada uno de los componentes, es decir:

$$C_Z(\mathbf{h}) = \sum_{k=0}^K C_k(\mathbf{h}) \quad (20)$$

Generalmente se supone que la línea base se encuentra asociada a la geología regional y por lo tanto para esta se espera que los rangos de influencia en el modelo de semivariograma sean mayores, mientras que la anomalía se encuentra relacionada con fenómenos más localizados y por lo tanto con distancias de correlación menores y con efecto pepita.

En el caso que se quiera separar la componente de anomalía el sistema de ecuaciones del Krigueaje (ecuación 17) queda expresado como:

$$\begin{aligned} \sum_{\beta=1}^n d_{\beta}^{k_0}(\mathbf{u}) C_Z(\mathbf{u}_{\beta} - \mathbf{u}_{\alpha}) + \mu^{(k_0)}(\mathbf{u}) &= \sum_{k=0}^{k_0-1} C_k(\mathbf{u} - \mathbf{u}_{\alpha}) \\ \sum_{\beta=1}^n d_{\beta}^{k_0}(\mathbf{u}) &= 0 \end{aligned} \quad (21)$$

Donde  $\mu^{(k_0)}(\mathbf{u})$  es el denominado multiplicador de la grange. En este caso se asume que esta componente tiene una media igual a 0, y de ahí la restricción en los pesos que aparece en la parte inferior de la ecuación 21. El valor estimado de la componente de anomalía estaría dado por:

$$[Z^{k_0}(\mathbf{u})]_{KO}^* = \sum_{\alpha=1}^n d_{\alpha}^{(k_0)}(\mathbf{u}) Z(\mathbf{u}_{\alpha}) \quad (22)$$

Donde  $d_{\alpha}^{(k_0)}$  son los pesos obtenidos del sistema de ecuaciones (ecuación 21). Para el caso de la línea base el sistema de ecuaciones del Krigueaje (ecuación 17) queda expresado:

$$\begin{aligned} \sum_{\beta=1}^n d_{\beta}^{K-k_0}(\mathbf{u}) C_Z(\mathbf{u}_{\beta} - \mathbf{u}_{\alpha}) + \mu^{K-k_0}(\mathbf{u}) &= \sum_{k=k_0}^K C_k(\mathbf{u} - \mathbf{u}_{\alpha}) \\ \sum_{\beta=1}^n d_{\beta}^{K-k_0}(\mathbf{u}) &= 1 \end{aligned} \quad (23)$$

Y en el cual se asume que la línea base tiene una media igual a la de la variable geoquímica. El valor estimado de la línea base queda dado por:

$$[Z^{(K-k_0)}(\mathbf{u})]_{KO}^* = \sum_{\alpha=1}^n d_{\alpha}^{(K-k_0)}(\mathbf{u}) Z(\mathbf{u}_{\alpha}) \quad (24)$$

El resultado de aplicar este procedimiento sobre un conjunto de datos geoquímico corresponde a un mapa de línea base (obtenido mediante la ecuación 23) y otro de anomalía geoquímica (ecuación 21). Sin embargo en el caso que no se desee elaborar el mapa, es posible aplicar este procedimiento de Krigueaje factorial sobre las muestras directamente y visualizar los resultados como un mapa de puntos. Una ventaja de este método corresponde al hecho que permite la determinación de la línea base y anomalía para cada muestra por separado, en lugar de definir un único valor de umbral para todo el conjunto de datos. Así mismo la incorporación del semivariograma en el procedimiento de separación reconoce la natura-

leza espacial del fenómeno estudiado. Una desventaja potencial de este método consiste en que requiere por parte del usuario la determinación y el modelamiento de los semivariogramas, así como de software especializado para la implementación de los procedimientos. También la discriminación entre línea base y anomalía a partir de las estructuras de correlación espacial dadas por el semivariograma puede ser un poco arbitraria, ya que el analista no puede estar completamente seguro que la anomalía geoquímica presente correlación espacial y que dicha correlación se pueda inferir. Además la correlación espacial de la anomalía puede quedar enmascarada por la variación de la línea base.

### Krigueaje Robusto

Uno de los grandes problemas del estimador del Krigueaje (ecuación 16) es que al ser de tipo lineal es sensible a valores extremos muy grandes o muy pequeños ("outliers" o valores extremos). La presencia de estos valores en un conjunto de datos hace que los resultados de los diferentes análisis estadísticos sean obscurecidos y de difícil interpretación, por lo cual resulta vital su temprana identificación por medio de las técnicas del análisis exploratorio de datos. Una vez se han detectado estos valores se procede a su exclusión y se repiten los análisis respectivos. Sin embargo es de aclarar que en muchas oportunidades, como en la exploración geoquímica, los "outliers" de valores altos son los que resultan de especial importancia, ya que son una evidencia de enriquecimiento con respecto a las concentraciones pertenecientes a la línea base. Así mismo en los programas de exploración el objetivo final no es detectar si una muestra determinada pertenece a la anomalía, sino determinar la extensión espacial de esas zonas anómalas. Para esto se puede emplear la técnica conocida como Krigueaje robusto (CRESSIE 1991), la cual fue diseñada para trabajar en situaciones en donde se tengan al menos dos grupos de observaciones, uno de los cuales presenta valores más alto que la mayoría de los datos. Esta metodología trata de determinar la consistencia de cada muestra y en el caso que esta no se respete procede a realizar una corrección del valor medido. Esto se logra definiendo una distribución local por medio de Krigueaje (ecuación 17) y determinando si el dato medido cae en la parte central de esta. En este caso particular se asume que la anomalía geoquímica se encuentra asociada a los "outliers" o muestras anómalas de la distribución de los datos y que la línea base corresponde a las muestras normales de la distribución (normales en el sentido que no están contaminadas).

El procedimiento para realizar krigueaje robusto es el siguiente (Fig. 10):

- Calcular y modelar el semivariograma robusto.
- Determinar los pesos de krigueaje para cada muestra usando las muestras cercanas a partir del semivariograma robusto determinado en el paso anterior.

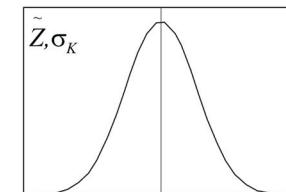
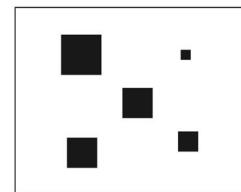
- Determinar el valor estimado  $Z^*$  y la desviación estándar del krigueaje  $\sigma_{K_i}$  en cada punto. Como valor estimado considere la mediana ponderada  $\tilde{Z}$ .

Editar el valor de cada muestra  $Z_i$  y reemplazarla por el valor corregido dado por:

$$Z_i^{(c)}(u_\alpha) = \begin{cases} \tilde{Z}(u_\alpha) + c\sigma_{K_i} & \text{Si } Z_i(u_\alpha) - \tilde{Z}(u_\alpha) > c\sigma_{K_i} \\ Z_i(u_\alpha) & \text{Si } |Z_i(u_\alpha) - \tilde{Z}(u_\alpha)| \leq c\sigma_{K_i} \\ \tilde{Z}(u_\alpha) - c\sigma_{K_i} & \text{Si } Z_i(u_\alpha) - \tilde{Z}(u_\alpha) < c\sigma_{K_i} \end{cases} \quad (25)$$

Donde  $c$  es un parámetro que controla el grado de corrección (y cuyo valor varía entre 1.5 y 2.5).

Población original



Sustracción de muestra

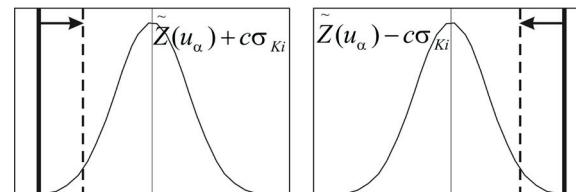
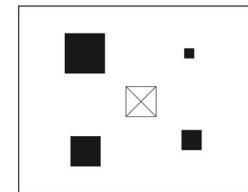


Fig. 10. Procedimiento de Krigueaje robusto.

- Con los datos corregidos  $Z_i^{(c)}(u_\alpha)$  realizar el procedimiento de krigueaje usando el semivariograma robusto.

En el procedimiento anterior CRESSIE (1991) recomienda la utilización de los semivariogramas robustos, sin embargo los resultados obtenidos por COSTA (1997) sugieren que si el semivariograma se estima de forma conveniente ya sea utilizando semivariogramas relativos (ISAACS & SRIVASTAVA 1989) o trabajando sobre el conjunto de datos depurado (KOTTEGODA & ROZZO 1993) y se puede realizar un ajuste del mismo, entonces este semivariograma se puede utilizar sin afectar el procedimiento de Krigueaje robusto.

El valor corregido obtenido anteriormente  $Z_i^{(c)}(u_\alpha)$  se relaciona con la línea base y las diferencias entre los valores corregidos y los valores reales  $Z_i - Z_i^{(c)}$  corresponde-

rian a la anomalía geoquímica. Con los valores corregidos anteriormente es posible ahora definir el valor de la línea base y la anomalía para cada muestra corregida de la siguiente forma:

$$\begin{aligned} \text{Línea base} &= Z_i^{(c)}(\mathbf{u}_\alpha) \\ \text{Anomalía} &= Z_i(\mathbf{u}_\alpha) - Z_i^{(c)}(\mathbf{u}_\alpha) \end{aligned} \quad (26)$$

Los valores obtenidos de la ecuación 26 pueden ser visualizados ya sea a través de mapas de puntos o mapas de contornos, según sea el caso.

Las ventajas de este método son las mismas que para el Krigeaje factorial: incorporación de la estructura de correlación espacial y la capacidad de realizar la separación por muestra. De acuerdo a los autores este método presenta dos desventajas serias. La primera se relaciona con la determinación del valor apropiado del factor de corrección  $c$  que aparece en la ecuación 25, ya que todo el procedimiento depende de este. CRESSIE (1991) propone determinar el valor de  $c$  que permita a los datos corregidos reproducir la varianza original de los datos. La otra desventaja hace referencia a la misma estabilidad de las correcciones obtenidas, ya que una muestra puede ser clasificada como anómala si hay más de una muestra anómala cerca y por lo tanto se estarían cometiendo errores significativos en el procedimiento. Así mismo se hace necesario contar con software especializado que implemente este procedimiento.

## CONCLUSIONES

Un aspecto importante a resaltar de estos métodos de discriminación de línea base/anomalía es que tienen una fuerte componente estadística y que el entendimiento de la misma es de vital importancia para tener criterios claros con los que cada uno de estos métodos se pueda aplicar en situaciones concretas. Estos criterios pueden estar relacionados con el número de datos, facilidad de aplicación de la metodología particular, complejidad del procedimiento a emplear y conocimiento por parte del usuario para la aplicación de la metodología concreta.

La gran suposición que hacen todos los métodos de discriminación explicado en el presente trabajo, corresponde a que la distribución geoquímica de elementos es muy compleja, y en donde coexisten múltiples distribuciones/grupos (multimodal). El objetivo consiste en delimitar de forma clara cada una de las distribuciones/grupos que conforman la distribución global del elemento, cada una de las cuales ha sido producida por un evento geológico particular, y en especial identificar aquellos grupos que hayan sido condicionados por procesos de enriquecimiento del elemento particular de estudio.

En el caso de los métodos estadísticos (promedio + 2desviación y cuantil95) se supone que los valores mayores a estos umbrales se encuentran controlados por proce-

sos geológicos diferentes a los del resto de la distribución. Esta suposición se basa en un criterio estadístico válido para una distribución normal, pero no tiene nada que ver con los diferentes procesos geológicos que controlan la distribución de elementos en la naturaleza. Sin embargo esto no quiere decir que sea correcto y la utilización de estas metodologías es recomendable como método inicial o si se emplea en conjunto con otros métodos.

En los métodos gráficos se asume que los cambios en las pendientes de ciertos gráficos estadísticos revelan la ocurrencia de grupos de muestras con características estadísticas propias. La determinación de estos quiebres de pendiente se puede hacer mediante la inspección visual de los gráficos estadísticos y por ende requiere la interpretación del analista de la información. Esto los hace fáciles de aplicar aunque con un cierto grado de subjetividad, ya que es necesario determinar cual es el quiebre con el que hace la discriminación.

El método de concentración – área para la discriminación de línea base/anomalía geoquímica tiene en cuenta la variación de las características fractales de la distribución del elemento particular. Con los datos puntuales se elabora el mapa de contornos de la concentración del elemento y se determina el área acumulada de los contornos en función de la concentración. En este diagrama se pueden apreciar si existen variaciones en las propiedades fractales de la distribución de concentración, y se puede seleccionar un valor de umbral con el que se realiza finalmente la separación. Las características fractales en la distribuciones geoquímicas tiene su explicación en la física de los procesos de diferenciación magmática y de difusión en medios heterogéneos, tal como lo propone ALLEGRE & LEWIN (1995) y por lo tanto esta metodología es la única, de las explicada en este trabajo, que tiene un fundamento físico para su aplicación.

La estadística espacial U tiene la ventaja de reconocer las características espaciales de las anomalías, ya que estas se concentran en zonas particulares debido a la ocurrencia de procesos geológicos particulares en contraposición a la delimitación de muestras que se da en los métodos estadísticos y gráficos. Sin embargo cabe mencionar el hecho, que esta metodología se basa todavía en suposiciones de normalidad y que adicionalmente el valor del umbral se debe determinar a partir del análisis de la distribución de los valores de U, ya sea por un método estadístico y gráfico.

En el caso de los métodos geoestadísticos estos involucran un aspecto importante de cualquier fenómeno geológico como lo es la estructura de correlación espacial. Los fenómenos geológicos no ocurren de forma aleatoria en el espacio ni en el tiempo, y por lo tanto el estudio de los patrones de ocurrencia espacial y su relación con la geología resulta de especial importancia en estudios de este tipo. La correlación espacial se estudia por medio del semivariograma y generalmente este se puede explicar

en términos de procesos geológicos particulares. Cada proceso geológico tiene sus características espaciales propias y este hecho es la base del krigueaje factorial, en el cual la variación espacial de la concentración del elemento de interés se descompone en sus partes constituyentes. La suposición fuerte que se hace en este método corresponde al hecho que cada proceso geológico presenta una variación espacial independiente de los restantes eventos, y que además cada fenómeno de interés presenta una variación espacial propia; lo cual muchas veces no es correcto. En el caso del krigueaje robusto, se asume que las muestras anómalas son aquellas que presentan unas características contrastantes con las muestras de la línea base desde el punto de vista de la estimación espacial. Se consideran como anómalas aquellas muestras que no se pueden estimar de forma correcta a partir de las muestras cercanas. Esta es una forma interesante de realizar la discriminación de anomalía/línea base que reconoce la presencia de al menos dos grupos con características estadísticas contrastantes y que además incorpora la estructura de correlación espacial del fenómeno de interés.

## AGRADECIMIENTOS

Al Instituto Colombiano de Geología y Minería (INGEOMINAS) por facilitar los datos obtenidos en sus proyectos de investigación. A las Químicas Luz Myriam González y Gloria Prieto (INGEOMINAS, Bogotá) quienes amablemente realizaron valiosos aportes en etapas iniciales del presente proyecto.

GMS reconoce el apoyo recibido mediante el programa “Becas para Estudiantes Sobresalientes de Postgrado” (Vicerrectoría Académica, Universidad Nacional de Colombia) durante sus estudios en el programa curricular “Maestría en Ciencias – Geología” del Departamento de Geociencias de la Universidad Nacional de Colombia.

## REFERENCIAS

- ALLEGRE, C. & LEWIN, E. (1995): Scaling laws and geochemical distribution. *Earth and Planetary Sciences Letters* 132 (1), 1 – 13.
- BATISTA, A., SOUSA, A., BATISTA, M., & VIEGAS, L. (2001): Factorial kriging with external drift: a case study on the Penedono Region, Portugal. *Applied Geochemistry* 16 (8), 921 – 929.
- BOURGAULT, G. (1994): Robustness of noise filtering by kriging analysis. *Mathematical Geology* 26 (6), 733 – 752.
- CARGILL, S. (1981): United States gold resource profile. *Economic Geology* 76 (8), 937 – 943.
- CARLSON, C. (1991): Spatial distribution of ore deposits. *Geology* 19 (2), 111 – 114.
- CARR, J. (1995): Numerical analysis for the Geological Sciences. Prentice-Hall Ltd. 592p. New Jersey.
- CHENG, Q. (1999): Spatial and scaling modelling for geochemical anomaly separation. *Journal of Geochemical Exploration* 65 (2), 175 – 194.
- CHENG, Q., AGTERBERG, F., & BONHAM-CARTER, G. (1996): A spatial analysis method for geochemical anomaly separation. *Journal of Geochemical Exploration*, 56 (2), 183 – 195.
- CLARK, I. (1979): Practical Geostatistics. Blackwell Science. 120p. Scotland.
- COSTA, J. (1997): Developments in recoverable reserves estimation and ore body modelling. Tesis de Doctorado, Queensland University. 233p. Australia.
- CRESSIE, N. (1991): Statistics for spatial data. John Wiley & Sons, 1 edition. 900p. New York.
- DAVIS, J. (2002): Statistics and Data analysis in Geology 3 edition John Wiley & Sons. 638p. New York.
- FEDER, J. (1988): Fractals. Plenum Press. 283p. New York.
- GONÇALVES, M., MATEUS, A., & OLIVEIRA, V. (2001): Geochemical anomaly separation by multifractal modelling. *Journal of Geochemical Exploration* 72 (1), 91 – 114.
- GRUNFELD, K. (2005): Visualization, integration and analysis of multi-element geochemical data. Tesis de Doctorado, Universidad de Lund. 54p. Suecia.
- ISAAKS, E. & SRIVASTAVA, R. (1989): Applied Geostatistics. Oxford University Press. 561p. New York.
- JIMENEZ, R. & CHICA-OLMO, M. (1999): Application of geostatistics to identify gold-rich areas in the Finisterre – Fervenza region, NW Spain. *Applied Geochmistry* 14 (2), 133 – 145.
- KOTTEGODA, N. & ROZZO, R. (1993): Statistics, Probability and Reliability for Civil and Environmental Engineers. McGraw Hill. 735p. Singapore.
- LEPELTIER, C. (1969): A simplified statistical treatment of geochemical data by graphical representation. *Economic Geology* 64 (4), 538 – 550.
- LI, C., MA, T., & SHI, J. (2003): Application of a fractal method relating concentrations and distances for separation of geochemical anomalies from background. *Journal of Geochemical Exploration* 77 (1), 167 – 175.
- LIMA, A., DE VIVO, B., CICCHELLA, D., CORTINI, M., & ALBANESE, S. (2003): Multifractal IDW interpolation and fractal filtering method in environmental studies: an application on regional stream sediments of Italy Campania region. *Applied Geochemistry* 18 (10), 1853 – 1865.
- LUO, X. & DIMITRAKOPOULOS, R. (2003): Data-driven fuzzy analysis in quantitative mineral resource assessment. *Computers & Geosciences* 29 (1), 3 – 13.
- MANDELBROT, B. (1983): The fractal geometry of nature. W. H. Freeman and Company. 468p. New York.
- OLEA, R. (1994): Fundamentals on Semivariogram Estimation, Modelling, and Usage. In: Yarus, J. & Chambers, R. (Eds). Stochastic Modeling and Geostatistics. AAPG Computer Applications in Geology 3, 27 – 36. Oklahoma.

- PETERS, G. (2001): Use of Structural Geology in Exploration for and Mining of Sedimentary Rock – Hosted Au Deposits. USGS Open – File Report 01 – 151. 40p. Reno.
- REIS, A., SOUSA, A., FERREIRA DA SILVA, E., PATINHA, C. & FONSECA, E. (2004): Combining multiple correspondence analysis with factorial kriging analysis for geochemical mapping of the gold-silver deposit at Marrancos (Portugal). *Applied Geochemistry* 19 (3), 623 – 631.
- REIS, A., SOUZA, A. & CARDOSO-FONSECA, E. (2003): Application of geostatistical methods in gold geochemical anomalies identification (Montemor-o-Novo, Portugal). *Journal of Geochemical Exploration* 77 (1), 45 – 63.
- SHEN, W. & COHEN, D. (2005): Fractally Invariant Distributions and an Application in Geochemical Exploration. *Mathematical Geology* 37 (8), 895 – 913.
- TURCOTTE, D. (1992): *Fractals and Chaos in Geology and Geophysics* 1 edition. Cambridge University Press. 221p. Cambridge.
- TURCOTTE D. (2002): Fractals in petrology. *Lithos* 65, 261 – 271.
- WEN, R. & SINDING-LARSEN, R. (1997): Image filtering by factorial kriging: sensitivity analysis and application to gloria side – scan sonar images. *Mathematical Geology* 29 (4), 433 – 468.
- YAO, T., MUKERJI, T., JOURNEL, A., & MAVKO, G. (1999): Scale matching with factorial kriging for improved porosity estimation from seismic data. *Mathematical Geology* 31 (1), 23 – 46.

*Manuscrito recibido Junio 2005; aceptado Octubre 2006.*