

KRIGEAJE DE INDICATRICES COMO METODO EN LA CLASIFICACION DIGITAL DE IMAGENES TM

F. Abarca y M. Chica-Olmo

G.I. "Geoestadística, Teledetección y GIS". Departamento de Geodinámica
Facultad de Ciencias. Universidad de Granada. Av. Fuentenueva S/N. 18071 GRANADA
Tfno/Fax: (958) 243363. Email: grsg@carpanta.ugr.es.

RESUMEN.- Uno de los aspectos más importantes en Teledetección es la clasificación de imágenes, proceso en el cual los píxeles que integran la imagen son asignados a un grupo o clase de referencia establecido o no a priori. Dentro de este campo el procesado matemático-estadístico ocupa un lugar destacado. Las técnicas geoestadísticas han demostrado ser una herramienta útil para el tratamiento de datos digitales de imágenes de satélite. En este trabajo se presenta una aplicación del método de estimación de Krigeaje de Indicatrices (KI) a la clasificación de imágenes TM en el contexto de una aplicación de exploración minera en el área de Cabo de Gata (Almería).

ABSTRACT.- One of the most important aspects in Remote Sensing is the image classification, process in which the pixels that integrate the image are assigned to an established group. In Remote Sensing the mathematical statistical processing occupies an important position. The geostatistic techniques have proved to be useful tool in the processing of satellite image digital information. In this paper we give an example of the application of the Kriging Indicator Estimating method (IK) to TM image classification in the field of mining exploration, in the area of Cabo de Gata (Almería).

1.- INTRODUCCION

Desde el punto de vista geoestadístico el valor digital (ND) de una imagen de satélite puede ser considerado como una variable regionalizada (Chica-Olmo, 1991). Sus características son, por un lado, un aspecto aleatorio de la información numérica y, por otro lado, la existencia de una estructura de correlación o variograma. Se han desarrollado numerosos trabajos basados en una metodología geoestadística enmarcados en el campo de la Teledetección (Van Der Meer, 1992; Atkinson, 1995). En este trabajo se ha aplicado el Krigeaje de Indicatrices (KI) como un método para la clasificación supervisada de imágenes de satélite.

La metodología que se propone se ha aplicado en el área ubicada en la región de Cabo de Gata, próxima a las explotaciones mineras de Rodalquilar. Para el mismo se ha dispuesto de una imagen LANDSAT, de la cual se han seleccionado 4 bandas (TM2, TM4, TM5 y TM7). La zona de estudio viene definida por las siguientes coordenadas UTM: (581660, 4083701) y (58870, 4073741), correspondientes al punto superior izquierdo e inferior derecho respectivamente. La imagen cubre un área de 72 Km², con un tamaño en píxeles de 241x332. La geología del área la definen rocas volcánicas y sedimentarias de edad neógena, si se exceptúan los recubrimientos cuaternarios. Las rocas volcánicas son de carácter calcoalcalino, variando entre andesitas y riolitas, (Arribas et al, 1989). Una gran parte de las rocas volcánicas del área están afectadas por fenómenos de alteración hidrotermal más o menos intensos. Para esta aplicación se ha seleccionado una clase, rocas volcánicas poco alteradas,

que es la que se pretende clasificar, Figura 1.

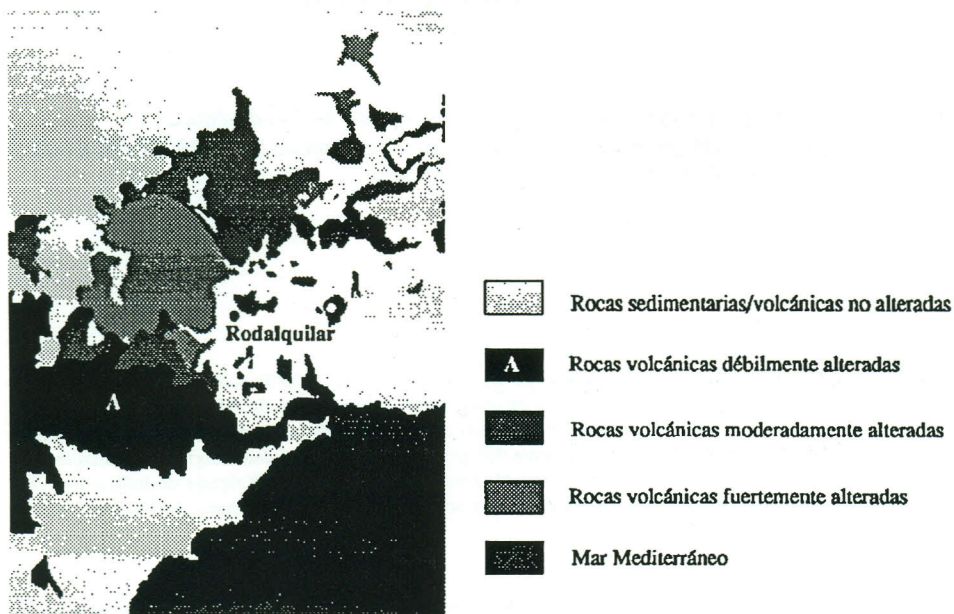


Figura 1.- Distribución de las rocas volcánicas alteradas. Clase A: rocas volcánicas débilmente alteradas.

2.- OBJETIVOS

El objetivo que se plantea en este trabajo es la aplicación del KI en la clasificación supervisada de imágenes. Para ello, se siguen los mismos pasos que en los otros métodos supervisados, se define un área de entrenamiento y se analizan geoestadísticamente los ND que la integran considerándolos, después de su transformación, como una variable indicatriz $[0,1]$. El método permite calcular en términos de probabilidad el grado de pertenencia de un pixel a una clase definida.

Es difícil comparar los resultados de este método con cualquier otro, sabiendo que lo que se obtiene es un mapa de probabilidades y no temático como es el caso de la gran mayoría de los clasificadores. Es posible generar un mapa temático, para ello es necesario establecer un valor de corte de los valores obtenidos en la clasificación para la inclusión o exclusión de un pixel a una clase, obteniendo finalmente la distribución de la clase para una determinada probabilidad. Este mapa temático puede verse modificado dependiendo de los valores que se hayan tomado en la definición de la firma espectral y de la probabilidad de corte que se considere.

3.- ASPECTOS METODOLOGICOS

Uno de los cambios más importantes registrados en Geoestadística en la pasada década ha sido la aparición de métodos de estimación de funciones de probabilidad a nivel local de tipo no paramétrico, mediante la transformación de los datos experimentales según una función

de tipo indicatriz. Esta función es de tipo binario, en la que los valores experimentales se transforman en 0 y 1, según superen o no el límite de corte impuesto (z_c).

$$I(x_i) = 1 \quad \text{si } Z(x) \leq z_c$$

$$I(x_i) = 0 \quad \text{si } Z(x) > z_c$$

El Krigeaje de Indicatrizes (KI) se puede definir como un Krigeaje Ordinario aplicado a los datos transformados indicatrizes en base a un determinado valor límite. Por tanto, los datos y los modelos de covarianza se derivan directamente de los experimentales sin utilizar un modelo previo. Su objetivo es proporcionar un modelo de distribución condicional de una variable desconocida en un soporte no muestreado.

Este método proporciona la estimación de la distribución espacial de la variable. Para ello se utiliza la transformación previa de los datos experimentales en función del límite impuesto. Una propiedad muy importante de los datos transformados es que si se modifica el valor de corte, la función indicatriz varía, pudiendo considerarse, en este sentido, como una función de distribución. Esta es la base del método, la estimación mediante krigeaje simple u ordinario de un número suficiente de valores de corte, utilizándose para cada uno de los sistemas la función variograma indicatriz definida como:

$$\gamma_I(h, z_c) = \frac{1}{2} E [(I(x+h, z_c) - I(x, z_c))^2]$$

El sistema de ecuaciones de krigeaje de indicatrizes es:

$$\sum_{j=1}^n \lambda_j(z_c) \gamma_I(x_\alpha - x_\beta, z_c) + \mu(z_c) = \gamma_I(x_\beta, V, z_c) \quad \beta = 1 \dots n$$

$$\sum_{j=1}^n \lambda_j(z_c) = 1$$

siendo su error de estimación:

$$\sigma_{KI}^2 = \sum_{i=1}^n \lambda_i(z_c) \gamma_I(x_\alpha, V, z_c) + \mu(z_c) - \gamma_I(V, V, z_c)$$

La función de distribución estimada para un determinado valor límite (z_c) viene dada por la expresión:

$$\phi^*(V, z_c) = \sum_{i=1}^n \lambda_i(z_c) I(x_i, z_c) = I^*(x)$$

El valor estimado $I^*(x)$ representa la probabilidad de que la variable no supere un determinado valor límite en un determinado soporte o punto x .

Las etapas de elaboración de un estudio de Krigeaje de Indicatrizes pueden resumirse de la siguiente forma:

- a) Transformación previa de los datos, en base a los valores límites elegidos (z_c).
- b) Estimación de la función variograma de indicatriz para cada uno de los valores; aunque en la práctica se puede optar por el cálculo de una única función para el valor límite correspondiente a la mediana de los valores experimentales ($\gamma_I(h)$).
- c) Planteamiento y resolución de los sistemas de krigeaje de indicatrizes. Estos sistemas son similares a los sistemas de krigeaje lineal, sustituyéndose los

valores experimentales y variogramas correspondientes, por sus respectivos transformados.

- d) Cálculo de la función de densidad de probabilidad a nivel local para cada uno de los píxeles estimados.

La aplicación del método se ha hecho considerando dos límites de corte, z_c , inferior (ndi) y superior (nds), para cada banda TM.

4.- APLICACION DEL METODO

Para llevar a cabo esta aplicación se ha seleccionado un área de entrenamiento de 20 x 20 píxeles y cuatro bandas del sensor TM (TM2, TM4, TM5 y TM7) correspondiente a materiales volcánicos débilmente alterados. Por tanto, se pretende obtener la distribución de probabilidades de que un píxel pertenezca a esta clase. En el área de estudio todos los materiales volcánicos aflorante son dacitas y riolitas afectados por fenómenos de alteración hidrotermal, más o menos intensos, por este motivo se persigue discriminar grados de alteración y no tipos litológicos.

Un primer paso consiste en transformar los ND del área de entrenamiento a una variable indicatriz $i(x; z_c)$, siendo 1 si $ND < z_c$ y 0 si no. El valor de corte z_c se ha fijado como la mediana de los ND (Tabla I). Sobre los datos transformados se calcula la función variograma para cada banda y se ajusta a un modelo teórico (Chica-Olmo et al., 1995). El modelo teórico que se ha empleado en el ajuste del variograma experimental es el esférico, cuya ecuación se expresa a continuación:

$$\gamma(h) = C_0 + C \left[\frac{1.5h}{A} - \frac{0.5h^3}{A^3} \right] \quad \text{si } h < A$$

$$\gamma(h) = C_0 + C \quad \text{si } h \geq A$$

Los parámetros que lo definen son: alcance (A), distancia para la cual se pierde la correlación espacial; meseta (C), correspondiente a la varianza de los datos y, efecto pepita (C_0). Existe un comportamiento diferencial de la variabilidad espacial como pone de manifiesto los ponderadores de Krigeaje y el factor de anisotropía, cociente de los alcances de dos direcciones ortogonales (A_{E-W}/A_{N-S}), una de máxima variabilidad y otra de mínima. La variabilidad espacial en las cuatro bandas es similar (Chica-Olmo et al., 1995) que se traduce en valores de alcance del mismo orden de magnitud (Tabla II). En la Figura 2 se han dibujado los variogramas experimental y teórico de las bandas TM4 y TM7 calculados en la dirección de mayor variabilidad, E-W.

TM	media	σ^2	mediana	min	max
2	25	6	25	19	30
4	57	37	57	45	73
5	108	114	110	80	129
7	52	36	53	37	65

Tabla I.- Estadística básica del área de entrenamiento

TM	M	C_0	C	A	Anis
2	Esf.	0.03	0.22	4.9	0.35
4	Esf.	0.02	0.23	6.0	0.35
5	Esf.	0.01	0.24	4.8	0.80
7	Esf.	0.02	0.28	6.0	0.80

Tabla II.- Ajuste de los variogramas calculados en el área de entrenamiento indicatriz

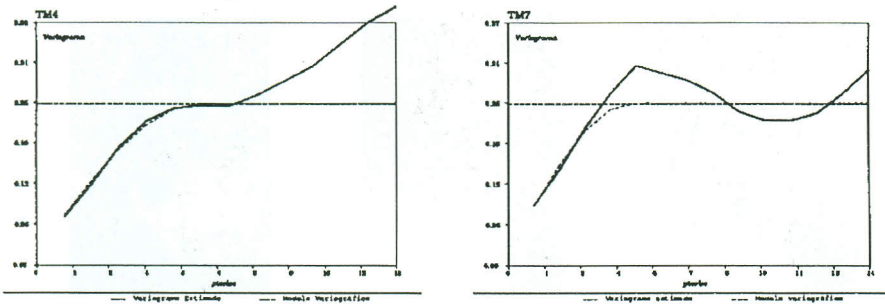


Figura 2.- Variogramas experimental y teórico de las bandas TM4 y TM7 (dirección E-W)

A partir de los modelos variográficos y del diseño de la vecindad que interviene en el proceso de estimación mediante Krigeaje Ordinario (KO) se han calculado los ponderadores (λ_i) correspondientes a cada píxel vecino (Figura 4) resolviendo el sistema de ecuaciones del krigeaje.

		0.0267		
	0.0152	0.4777	0.0152	
-0.0169	0.0355		0.0355	-0.0169
	0.0152	0.4777	0.0152	
		-0.0267		

Figura 4.- Vecindad de estimación y ponderadores del KI

Llegado este punto se han establecido los valores de corte para los ND de la imagen para obtener así la imagen indicatriz. Para el análisis variográfico en el área de entrenamiento se fijó como valor de corte la mediana. Ahora el objetivo es distinto, lo que se pretende es encontrar valores espectrales que definan estadísticamente el área de entrenamiento. Para ello hemos considerado la media $\pm 1\sigma$ (ndi,nds). Estos dos valores son útiles para definir los cortes a fin de generar las imágenes indicatrices, las cuales se van a estimar según el modelo que hemos definido. El resultado son dos imágenes de probabilidad, por un lado, P_1 , tenemos la probabilidad de que un ND sea menor que el valor de corte mínimo (ndi), véase Figura 5c, y por otro, P_s , de que sea menor que el valor de corte máximo (nds), Figura 5d. Por tanto, es fácil obtener la imagen P_i de probabilidad de que el ND pertenezca al intervalo es $P_s - P_1$, Figura 5e. Este procedimiento tendrá que ser repetido para cada banda. Para esta aplicación se han seleccionado aquellas bandas que aportaban más información, por ese motivo del visible sólo se ha elegido TM3, con la consiguiente reducción de tiempo de interpretación y procesamiento de datos.

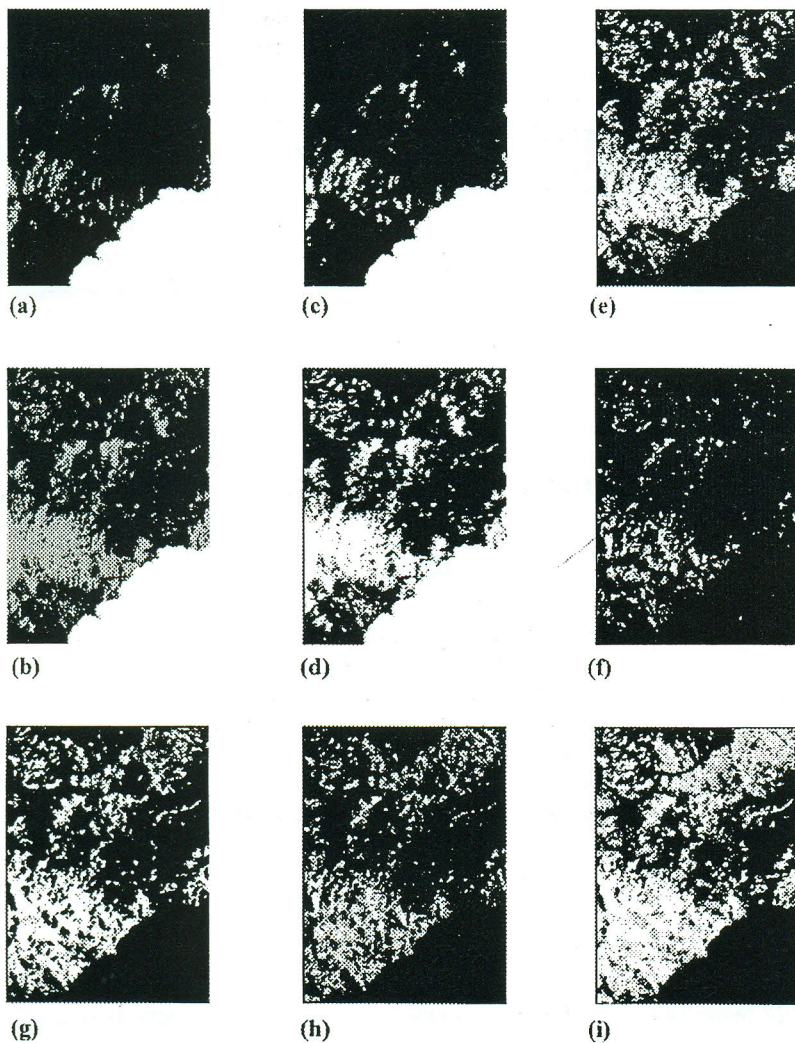


Figura 5.- Etapas en la clasificación de imágenes por KI. (a) Imagen indicatriz $I_i(x)$; corte = media - σ , (z_{ci}). (b) Imagen indicatriz $I_i(x)$; corte = media + σ , (z_{cs}). (c) y (d) Imágenes de probabilidad: $I_i \leq z_{ci}$ y $I_s \leq z_{cs}$ respectivamente. (e) Diferencia de las imágenes de probabilidad (d) - (c) para la banda TM2. (f) Resultado final integrando las 4 bandas: TM2, TM4, TM5 y TM7. (g) Resultado final considerando inicialmente $z_{ci} = \text{media} - 2\sigma$ y $z_{cs} = \text{media} + 2\sigma$. (h) Suma de los mapas temáticos obtenidos para cada banda con los píxeles clasificados con una probabilidad mayor del 50%. (i) Idem con una probabilidad del 85% con $z_c = \text{media} \pm 2\sigma$.

Los resultados que se han expuesto se han obtenido a partir de dos situaciones iniciales distintas, la primera fijando los valores de corte para la transformación indicatriz como la media $\pm \sigma$ y la segunda tomando como valores de corte la media $\pm 2\sigma$. Otros valores pueden ser el mínimo y máximo del área de entrenamiento. Dependiendo del intervalo escogido la clasificación final será más o menos restrictiva.

El mapa de probabilidad espacial representado en la Figura 5f puede presentar alguna dificultad para su interpretación por lo que es más cómodo fijar un umbral de probabilidad para el cual un pixel se asigne o no a la clase. Este proceso repetido para cada banda genera una colección de imágenes binarias que indican inclusión o exclusión del pixel a la clase. En la Figura 5h se presenta la resultante de la suma de las cuatro imágenes obtenidas para una probabilidad superior al 50% y, la Figura 5i superior al 85% pero partiendo de unos valores de corte inferior y superior correspondiente a la media $\pm 2\sigma$. En estos mapas, el valor del pixel viene determinado por el número de veces que ha sido asignado por superar la probabilidad de corte.

5.- CONCLUSIONES

El método de Krigeaje de Indicatrices (KI) proporciona una metodología de interés para la clasificación de imágenes de satélite al introducir el concepto de distribución espacial de probabilidades. La distribución espacial de la clase se obtiene fijando un nivel de probabilidad en la imagen estimada. Además, el método es versátil en el sentido de que permite establecer valores de corte más o menos restrictivos.

No obstante, el método presenta algunos inconvenientes en lo referente a la complejidad que presenta su aplicación. Precisa un análisis variográfico para cada área de entrenamiento, y son necesarias dos estimaciones por banda para llegar al mapa de probabilidad final, lo cual puede consumir gran tiempo de computación. Otro desventaja frente a otros métodos es la imposibilidad de clasificar simultáneamente varias clases.

Consideramos que la aplicación de este método, conjuntamente con otros clásicos, permite mejorar los resultados de la clasificación de imágenes introduciendo un término probabilístico de la distribución espacial de las clases, como se ha deducido en este trabajo preliminar.

6.- AGRADECIMIENTOS

La realización de este trabajo ha sido posible gracias al Proyecto de I+D "DARSTIMEX" del Programa BRITE/EURAM de la CE y a la financiación recibida por la DGICYT correspondiente al proyecto PB92-0925-CO2-02.

7.- BIBLIOGRAFIA

- Van Der Meer, F.D. (1992). Classification of high spectral resolution imagery using an indicator kriging. *Geostatistics Tróia '92*. pp. 829-840.
- Chica-Olmo, M., y Abarca, F. (1995). Análisis geoestadístico de la variabilidad espacial de datos multispectrales TM y SPOT. VI Reunión Científica asociación Española de Teledetección.
- Atkinson, Peter M., Harrison, Andrew R. 1995. *An improved method of estimating measurement error in remotely sensed*. Sensors and Environmental applications of Remote Sensing. Pp 399-406. Askne (ed.)
- Chica-Olmo, M., Vives, E. y Abarca, F. (1991). Métodos geoestadísticos aplicados al análisis de variables espaciales en Teledetección. IV Reunión Científica asociación Española de Teledetección.
- Arribas, A.Jr., Rytuba, JJ, Rye, R.O., Cunningham, C.G., Podwysocki, M.H., Kelly, W.C., Arribas, A., McKee, E.H. and Smith, J.G. (1989) *Preliminary study of the ore deposits and hydrothermal alteration in the Rodalquilar caldera complex, southeastern Spain*. U.S. Geol. Survey Open-File Report 89-327, 39 p.