

Un protocolo de tratamiento estadístico para muestras *polifásicas* de lineaciones de disolución

A procedure for statistical analysis of pliphase samples of pressure-solution lineations

J.L. Simón Gómez

Departamento de Geología, Universidad de Zaragoza. Pedro Cerbuna, 12. 50009 Zaragoza. jsimon@unizar.es

ABSTRACT

Accurate analysis of orientations of linear elements from polyphase populations involves computing statistics (mean, standard deviation) of each sub-population, which needs previous identification of the polymodal character and separation of sub-samples. The procedure here proposed for a regional study involves: (a) distinguishing polyphase from monophasic samples; (b) characterizing mean and deviation of clearly monophasic populations; (c) separating polyphase samples into sub-samples using density diagrams in equal-area projection and standard deviation values; (d) applying a correspondence test based upon the 95% confidence cone, as well as an equality F-test for directional samples, by comparison with monophasic reference samples. An example of application of this procedure to pressure-solution lineations in carbonate pebbles of the Aliaga Tertiary basin (Teruel, Iberian Chain) is shown.

Key words: *pressure-solution, lineation, statistics, polyphase sample.*

Geogaceta, 40 (2006), 67-70
ISSN: 0213683

Introducción

El análisis de datos direccionales en Geología Estructural, dada su lógica variabilidad (tanto natural como debida a errores de observación o medida), debe sustentarse en criterios estadísticos. Las orientaciones de elementos lineales simples e independientes (no asociados genéticamente a planos, no agrupados en sistemas con una simetría particular) son susceptibles de un tratamiento estadístico más o menos convencional, únicamente

con las particularidades propias de la estadística direccional. Pueden incluirse en él tanto los datos estructurales *s.s.* (fábricas tectónicas, lineaciones de presión-disolución...) como los resultados analíticos que toman la forma de líneas en el espacio (ejes de esfuerzos, ejes de deformación...).

La importancia de un tratamiento riguroso de las orientaciones de líneas aumenta en el caso de que las poblaciones sean, o se sospeche puedan ser, polimodales (con un eventual significado

genético de *polifásicas* o *politensoriales*). El cómputo de los parámetros estadísticos básicos de las subpoblaciones (media, desviación) requiere en tal caso una identificación previa de dicho carácter polimodal y una discriminación de los individuos que integran cada submuestra, además de la toma de decisiones sobre la consideración o no de aquellos datos alejados de la media que podrían ser considerados extraños a cualquiera de las subpoblaciones.

Todos estos aspectos son recogidos de forma ordenada en el protocolo analítico aquí descrito. Este ha sido aplicado a un conjunto amplio (más de 1000 datos) de lineaciones de presión-disolución medidas en cantos calcáreos de conglomerados de la cuenca terciaria de Aliaga (Teruel, Cordillera Ibérica). La exposición metodológica contiene referencias al proceso de estudio en dicha zona (aunque sin prestar mayor atención a los resultados finales) y se ilustra con un ejemplo de análisis en una de las estaciones. El conjunto de lineaciones medidas muestra cuatro orientaciones preferentes, cada una de las cuales representa una familia diferenciada de ejes de esfuerzo máximo σ_1 a la que puede atribuirse significado genético, de acuerdo con los criterios cinemáticos y dinámicos expuestos por Simón (2006). Hay que decir que el procedimiento es aplicable asimismo a cualquier lineación cuya génesis y patrón de

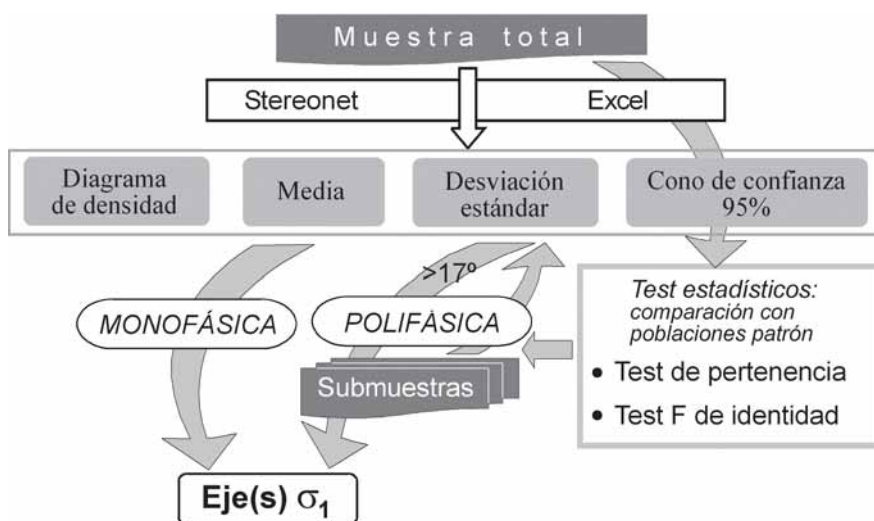


Fig. 1.- Esquema del procedimiento propuesto.

Fig. 1.- Sketch of the proposed procedure

orientación sea independiente de otras estructuras, y cuyo desarrollo evolutivo no dependa significativamente de la historia deformacional previa. Las herramientas estadísticas que utiliza son, por otro lado, totalmente convencionales, y se hallan incluidas dentro de paquetes informáticos de uso común como Stereonet (Allmendinger, 2004) y Excel de Microsoft.

Toma de datos

El análisis de lineaciones de disolución en la cuenca terciaria de Aliaga se ha basado en las observaciones y medidas efectuadas en 51 estaciones, repartidas a lo largo de casi toda la serie sedimentaria. Estas estaciones se seleccionaron de forma que su distribución espacial y estratigráfica fuese lo más homogénea posible, aunque las condiciones litológicas y de afloramiento no siempre lo han hecho posible.

Cada estación consiste en una porción lineal de afloramiento, de entre 10 y 80 m de longitud, en la que se han medido entre 9 y 79 lineaciones. La toma de datos se ha realizado según un procedimiento de muestreo lo más aleatorio posible, aunque eligiendo siempre aquellos cantos en que las marcas son de escala suficiente para ser medidas con precisión aceptable.

Análisis estadístico

El tratamiento de cada muestra de lineaciones parte de su representación en proyección equiareal y sigue los pasos que aparecen esquematizados en la figura 1. Sus objetivos específicos son: (a) caracterizar cada muestra *monofásica*; (b) decidir si son o no tenidos en cuenta los datos más alejados de la media; (c) decidir el posible desdoblamiento en submuestras, y (d) aplicar a estas últimas los pasos a) y b). Se trata, en definitiva, de discriminar entre desviaciones angulares que podemos considerar carentes de significado tectónico y aquellas otras que representan campos de esfuerzos genéticamente diferenciados (Caputo, 2005); dicho de otro modo, pasar de la simple constatación del carácter polimodal de una muestra a su interpretación como *polifásica* o *politensorial*.

Para determinar la orientación media de cada muestra se ha utilizado habitualmente la herramienta de análisis de autovectores que contiene Stereonet (versión 6.3.2 X; Allmendinger, 2004). Este análisis, basado en la distribución axial de Bingham, es adecuado para orientaciones de líneas que no llevan

asociado un sentido (Davis, 1986). Cuando las orientaciones presentan un agrupamiento lineal significativo, esté o no contenido dentro de una guirnalda, el primer autovector representa el eje medio de la muestra. Como elemento de apoyo pueden usarse los diagramas de densidad de puntos en proyección equiareal. No obstante, sólo en casos de distribución claramente asimétrica (lo que supone una moda bien diferenciada de la media), hemos dado prioridad al diagrama de densidad para definir la orientación óptima de s_1 .

La dispersión de orientaciones en cada muestra (o submuestra, en casos *polifásicos*) ha sido caracterizada mediante la desviación estándar de las direcciones y de las inclinaciones, calculadas ambas por separado mediante Excel de Microsoft (versión X para Macintosh). Por otro lado, el margen de error con que se ha calculado la corres-

pondiente orientación media queda expresado mediante el cono de confianza del 95%. En aquellas muestras o submuestras con 25 o más datos, Stereonet proporciona directamente los ángulos de apertura máximo y mínimo de dicho cono asociado al primer autovector. Para muestras más pequeñas, el cálculo de este parámetro basado en la distribución de Bingham se omite. No obstante, puede recurrirse a la distribución vectorial de Fisher (obteniendo en este caso un ángulo de apertura único) si, como ocurre en la muestra estudiada más adelante, todas las líneas tienen el mismo sentido de inclinación.

La distinción entre muestras *monofásicas* y *polifásicas*, y la separación de submuestras dentro de las segundas, se ha realizado en primera aproximación siguiendo dos criterios:

- (a) Diagrama de densidad en proyección equiareal.

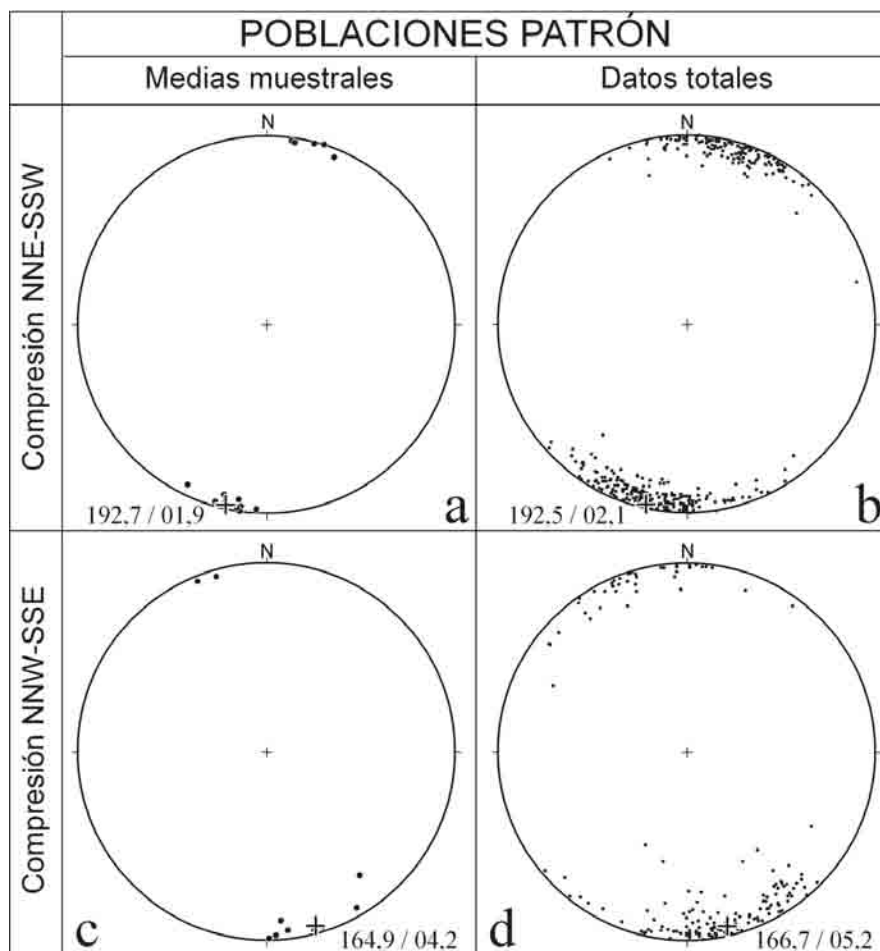


Fig. 2.- (a) y (b) Estereogramas (proyección equiareal, hemisferio inferior) de la población patrón representativa de la compresión NNE-SSW (medias muestrales y datos totales). Cruz: orientación media. (c) y (d) Idem para la población patrón NNW-SSE.

Fig. 2.- (a) y (b) Stereoplots (equal-area projection, lower hemisphere) of the reference population for the NNE-SSW compression (sample means and total data). Cross: mean orientation. (c) and (d) Idem for the reference NNW-SSE population.

(b) Comparación de los valores de desviación estándar con los que presentan las muestras claramente *monofásicas* en el área de estudio. En estas últimas, los valores de desviación oscilan entre 8° y 16° para la dirección y entre 2° y 16° para la inclinación ($11,5^\circ$ y $6,3^\circ$ de promedio, respectivamente). Las muestras que presentaban valores de desviación estándar en dirección o en inclinación superiores a 17° se han considerado como probables *polifásicas* y han sido divididas en submuestras.

En algunas muestras, la aplicación de estos dos criterios no ha sido suficiente para adoptar una decisión fundada sobre su carácter *monofásico* o *polifásico*. Se ha pasado entonces a un segundo estadio de análisis en que las muestras problema han sido comparadas con poblaciones patrón. Estas últimas se han definido a partir de las muestras *monofásicas* analizadas con anterioridad, tras constatar que de ellas se infieren algunas direcciones de compresión persistentes (interpretadas como el registro de campos regionales poco perturbados) que pueden considerarse referencias robustas. De este modo, la consistencia de una parte de los resultados ha servido de apoyo para interpretar otros no tan evidentes. Mediante una marcha analítica “por eliminación”, se ha caracterizado en primer lugar una población que representa la compresión NNE, la más constante en dirección y prácticamente la única representada en las unidades más recientes. A partir de ahí se han podido

aislar submuestras que representan esa misma población en estaciones *polifásicas*, e interpretar direcciones de compresión distintas a partir de los datos restantes de cada muestra. En la medida en que éstas son también persistentes y paralelas a las de otras muestras *monofásicas*, se han llegado a obtener nuevas poblaciones patrón que han servido para completar el análisis de todos los datos.

Las poblaciones patrón han sido definidas teniendo en cuenta las medias muestrales *monofásicas*. Según el teorema del límite central (Till, 1980; Davis, 1986), si un conjunto de muestras pertenecen a una misma población, las medias muestrales tendrán una distribución normal y su media será la misma de la población. Este principio puede aplicarse a datos direccionales y, en el caso de la compresión NNE-SSW, permite corroborar que todas las muestras agrupadas en ella conforman efectivamente una población “natural”. Así, la dirección e inclinación media de las 16 medias muestrales es $192,7 \pm 5,8^\circ / 01,9 \pm 3,6^\circ$, mientras la orientación media de todos los picos de disolución que integran el conjunto de dichas muestras (360 datos) es $192,5 \pm 1,8^\circ / 02,1 \pm 0,9^\circ$ (Fig. 2a,b). De modo análogo se ha definido la media de una población NNW-SSE ($164,9 \pm 11,8^\circ / 04,2 \pm 6,7^\circ$, a partir de 10 medias muestrales; $166,7 \pm 3,6^\circ / 05,2 \pm 1,8^\circ$, a partir de un total de 169 picos; Fig. 2c,d).

Una vez establecidas las poblaciones patrón, la identificación de una muestra o submuestra dada con alguna de ellas se ha realizado sobre la base de dos test estadísticos:

(a) Test de pertenencia basado en el cono de confianza del 95%. Si la orientación media de la población está contenida en el cono de confianza de una muestra, existe el 95% de probabilidades de que dicha muestra pertenezca a la población (Davis, 1986).

(b) Test F de identidad entre muestras direccionales. Este test se aplica únicamente a las direcciones, y sólo es apto por tanto para líneas próximas a la horizontal. Se calcula el parámetro estadístico $F_{1,n-2}$ en función del tamaño de las muestras, de los respectivos módulos de los vectores suma y del factor de concentración k (Davis, 1986). El resultado se compara con los valores tabulados para distintos niveles de significancia y grados de libertad. Si el valor de $F_{1,n-2}$ excede el valor crítico especificado para el 5% de significancia ($\alpha = 0,05$), se concluye que ambas muestras son distintas con un 95% de probabilidad.

Ambos test aportan dos puntos de vista complementarios. En términos prácticos, el primero nos permitirá asignar una muestra concreta de lineaciones a una población estadística definida como *dirección de compresión preferente*, mientras que el segundo nos permitirá descartar dicha posibilidad. Al interpretar los resultados de estos dos test no conviene

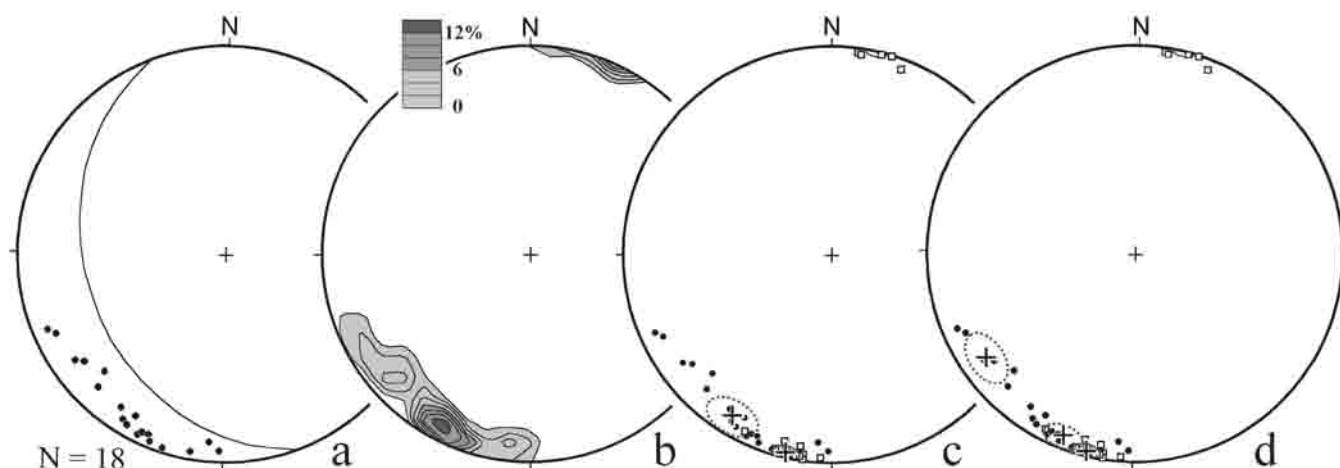


Fig. 3.- Análisis de lineaciones de disolución en la estación 26. (a) Orientaciones de líneas y plano de estratificación. (b) Diagrama de densidad; % de datos en el 1% de área. (c) Valoración estadística como muestra *monofásica* y comparación con la población patrón NNE-SSW. (d) Valoración como muestra *polifásica* y división en submuestras. Cuadrados huecos: medias muestrales de la población patrón; cruces: orientaciones medias de la muestra, submuestras y población patrón; trazo discontinuo: conos de confianza de la muestra total y las submuestras; trazo continuo: cono de confianza de la población patrón.

Fig. 3.- Analysis of solution lineations in site 26. (a) Orientation of lines and bedding. (b) Density diagram; % data in 1% area. (c) Statistical assessment as monophasic sample and comparison with the reference NNE-SSW population. (d) Statistical assessment as polyphasic sample and partition into sub-samples. Squares: sample means of the reference population; crosses: mean orientations of the sample, sub-samples and reference population; dotted line: confidence cones of the total sample and sub-samples; solid line: confidence cone of the reference population.

adoptar un criterio rígido. En el test F, dado que se trata de comparar muestras entre sí y no una muestra con una población, hay que elegir qué muestra se toma como referencia. Una posibilidad es adoptar una "muestra patrón" formada por las medias muestrales (16 y 10 datos para las poblaciones NNE y SSE, respectivamente). Esto supone probablemente una exigencia excesiva: una muestra cuyo "parentesco" es descartado por el test puede representar simplemente una pequeña desviación de las trayectorias de esfuerzos. Otra posibilidad es comparar con muestras monofásicas individuales: si cada una de éstas ha sido atribuida a la población patrón, y la muestra problema presenta una afinidad notable con algunas de ellas, parece razonable suponer que también puede pertenecer a dicha población. Este es un criterio más amplio, que tiende a favorecer el agrupamiento de las muestras y contribuye a evitar un desdoblamiento excesivo que podría llevar a soluciones espurias.

Un ejemplo

Para ilustrar el procedimiento propuesto describimos su aplicación a una de las muestras estudiadas (estación 26). El estereograma de las 18 lineaciones medidas (Fig. 3a) muestra una dispersión amplia de azimut entre 180 y 250. En el diagrama de densidad (Fig. 3b) se aprecia un máximo absoluto próximo a 210, pero también máximos relativos en 185 y 230. Se plantea, por tanto, la duda de si se trata de una muestra *monofásica* o *polifásica*.

Considerada como *monofásica*, la muestra tiene una media 211,3°/9,5° y desviaciones estándar de 19,1° y 5,0° para la dirección y la inclinación, respectivamente. El cono de confianza del 95% (distribución vectorial de Fisher) tiene un ángulo de 8,5° (Fig. 3c). La desviación en dirección supera el umbral de 17°, lo que sugiere la división en dos o más submuestras.

El cono de confianza de la muestra no incluye la media de la población patrón NNE (193/02), ni tiene intersección con el cono de confianza de esta última (Fig. 3c). Dicho cono tan sólo engloba la media de una de las muestras individuales, si bien se trata precisamente de aquella cuya desviación respecto a la media es claramente mayor que el resto. Por tanto, la pertenencia a esta población no queda demostrada.

Para aplicar el test F de identidad, comparamos las direcciones de la muestra total de la estación 26 con la "mues-

tra patrón" construida a partir de las medias muestrales de la población de referencia:

Muestra patrón población NNE			Muestra real estación 26		
192	189	018	184	202	246
188	022	015	189	201	204
193	194	189	197	205	225
008	188	183	209	213	232
196	009	193	181	208	223
206			244	211	234

Para cada una de las 16 direcciones que componen dicha muestra patrón leemos el ángulo α que forma con una dirección de referencia (N-S) y calculamos $\text{sen}\alpha$ y $\text{cos}\alpha$. Los correspondientes sumatorios permiten calcular el módulo del vector resultante (R_1), considerando cada línea individual como un vector unitario:

$$R_1 = (15,534^2 + 3,495^2)^{1/2} = 15,922$$

$$N_1 = 16$$

A continuación realizamos el mismo cálculo para la muestra de la estación 26:

$$R_2 = (14,563^2 + 8,911^2)^{1/2} = 17,073$$

$$N_2 = 18$$

y para la combinación de ambas:

$$R_p = (30,097^2 + 12,406^2)^{1/2} = 32,554$$

$$n = N_1 + N_2 = 34$$

El valor del factor de concentración k para la muestra estudiada se obtiene en la correspondiente tabla (Davis, 1986, p. 323) a partir de la longitud estandarizada del vector suma: $\bar{R}_{2\text{med}} = R_2/N_2 = 0,948 \Rightarrow k \simeq 10$. A continuación, el parámetro F puede calcularse mediante la expresión (Davis, 1986, p. 327):

$$F_{1,n-2} = (1 + 3/8k)(n-2)(R_1 + R_2 - R_p)/(nR_1 - R_2)$$

$$F_{1,n-2} = 14,569$$

El valor crítico de F para $\alpha = 0,05$ y para los grados de libertad de nuestro problema ($n_1 = 1$; $n_2 = 34 - 2 = 32$) es $F_{\text{crítico } 1,32} \simeq 4,15$ (Davis, 1986, p. 70). El valor calculado para la muestra problema es muy superior a este valor crítico, por lo que cabe concluir que ambas muestras son distintas.

Estos resultados sugieren que la muestra no es *monofásica*, a pesar de presentar un máximo absoluto destacado en dirección NNE, lo que plantea la necesidad de subdividirla. A la vista del diagrama de densidad de la figura 3b, la división podría hacerse por el mínimo más acusado, en torno al azimut 220. Las dos submuestras obtenidas de esta subdivisión (12 y 6 datos) presentan orientaciones medias 200,3°/6,8° y 234,1°/13,8°,

con conos de confianza de 5,9° y 8,5°, respectivamente (Fig. 3d). Las desviaciones estándar de direcciones son 10,6° y 9,5°, valores que están dentro de los rangos usuales en poblaciones *monofásicas*.

La aplicación de los test estadísticos a la primera submuestra indica una elevada afinidad con la población NNE. Su cono de confianza, aunque no incluye exactamente la media de la población, sí tiene una intersección notable con el cono de ésta, e incluye tres de las medias muestrales (Fig. 3d). El valor de F obtenido al comparar con la muestra patrón, si bien sigue estando por encima del valor crítico ($F_{1,n-2} = 6,08$ frente a $F_{\text{crítico } 1,26} = 4,23$), se acerca a éste mucho más que cuando comparáramos la muestra total. La comparación con varias muestras monofásicas individuales de la población NNE proporciona valores de $F_{1,n-2}$ por debajo del umbral crítico para cinco de ellas.

Respecto a la segunda solución (submuestra en torno a 234/14), aunque no existen muestras monofásicas de orientación similar con las que poder comparar, hay dos argumentos que apoyan su fiabilidad: (1) La media es similar a la inferida en otra estación vecina (056/03); comparando ambas muestras mediante el test F , obtenemos un valor de $F_{1,n-2} = 0,009$, muy inferior al valor crítico de 4,60. (2) Es similar a la orientación del eje s_1 obtenido del análisis de planos estriados en esta misma estación 26 (229/10).

En conclusión, el conjunto de resultados expuestos aconseja considerar la muestra total como *polifásica* y adoptar dos soluciones distintas de s_1 a partir de los picos de disolución: 200/07 y 234/14.

Agradecimientos

Agradezco a J. González la revisión y sugerencias al manuscrito original.

Referencias

- Allmendinger, R.W. (2004). <http://www.geo.cornell.edu/geology/faculty/RWA/>
- Caputo, R. (2005). *Earth-Science Reviews*, 70, 103-127.
- Davis J.C. (1986). *Statistics and data analysis in Geology*. John Wiley and Sons, 646 p.
- Simón, J.L. (2006). *Revista de la Sociedad Geológica de España*, 19, 153-162.
- Till, R. (1980) *Statistical methods for the Earth scientist*. Macmillan, 154 p.