

APLICACION DE LOS ANALISIS «CLUSTER» AL ESTUDIO DE CARACTERISTICAS FISICO-QUIMICAS DE AGUAS SUBTERRANEAS

S. Moral Callejón (*), A. Pulido Bosch (**) y P. Valenzuela (***)

RESUMEN

Se comentan tres ejemplos de aplicación del análisis cluster referidos a características físico-químicas de aguas subterráneas, dos de evolución espacial y uno temporal. Todos los ejemplos analizados ponen de manifiesto la utilidad de esta herramienta en la descripción de los datos.

PALABRAS CLAVE: Análisis cluster, espacio, tiempo, ciclos hidrogeoquímicos.

ABSTRACT

This paper deals with the application of cluster analysis to the physicochemical characteristics of ground waters. Three examples of such methodology are presented, carried out at two different levels: spatial evolution and time variation. All the cases studied show the utility of clustering technique as a valuable tool in description on this kind of data.

KEY WORDS: Cluster analysis, space, time, hydrogeochemical cycles.

Introducción

Los análisis cluster (o de enjambres —in Ibáñez *et al.*, 1983— o racimos —“analyses de grappes”, en francés—) constituyen una herramienta de descripción de datos cuantitativos y cualitativos que surgió de la mano de los biólogos con fines taxonómicos (Isnard *et al.*, 1972). Permite conocer el grado de relación entre los parámetros observados o medidos en muchos individuos, o agrupar los individuos por la similitud de los parámetros que los caracterizan.

Estas técnicas han sido aplicadas en diversas ramas de la geología con diferentes fines. Brown (*in op. cit.*) lo aplica a la clasificación de yacimientos de cobre y molibdeno; también ha sido utilizado este método para cartografía geoquímica (*op. cit.*) y para el análisis de cavidades kársticas (Blanc y Monteau, 1983).

En el dominio de la Hidrogeología, y concretamente el relativo al tratamiento de las características físico-químicas de las aguas, además de las técnicas usuales de representación y tratamiento, suele ser muy empleado el análisis factorial (Bakalowicz, 1979; Melloul, 1979; Mudry *et al.*, 1979; Benavente *et al.*,

1983, ...); los análisis cluster han sido empleados con esta finalidad, en España, en el trabajo de Ibáñez y Sanchis (1983), sin que conozcamos otros casos de aplicación.

A continuación, y tras una síntesis de los aspectos teóricos, comentamos los resultados obtenidos de la aplicación de estas técnicas de análisis, a tres ejemplos: dos relativos a datos en el espacio y uno a datos hidrogeoquímicos temporales.

Consideraciones teóricas

Vamos a describir brevemente las dos técnicas de análisis de datos utilizadas en este trabajo: cluster de variables y cluster de casos. Lo que hace, en esencia, este tipo de análisis es establecer grupos y clasificar. La ventaja de este procedimiento sobre otros más sofisticados, como el Análisis Factorial, es su simplicidad matemática, que implica pocas hipótesis sobre los datos iniciales, lo que hace que “funcione” bien en la gran mayoría de los casos.

Consideremos m variables aleatorias x_1, x_2, \dots, x_m que se miden en p individuos. Consideremos las observaciones agrupadas en una matriz $p \times m$, $X = (X_{ij})$ donde X_{ij} es el valor de la variable X_j en el i -ésimo individuo. Cada fila está

(*) Licenciado en C. Matemáticas, Dpto. de Estadística e I. O. Univ. de Granada.

(**) Doctor en C. Geológicas, Dpto. de Hidrogeología, Universidad de Granada.

(***) Licenciada en C. Matemáticas.

formada por las m observaciones correspondientes a un individuo, y se puede considerar como un punto en el espacio euclídeo m -dimensional. Análogamente, cada columna está formada por las p -observaciones de una variable, y se puede considerar como un punto del espacio euclídeo p -dimensional. El problema del cluster de casos consiste en determinar n subconjuntos de individuos $\pi_1, \pi_2, \dots, \pi_n$, de tal forma que cada individuo pertenezca a una y sólo una de estas clases, en base a una medida de similitud entre los puntos que representaban cada caso, medida que varía de un caso a otro según sea la naturaleza del problema. De forma similar se define lo que es el cluster de variables. Para un estudio más detallado de lo que es el análisis cluster y los diferentes tipos de distancias, nos remitimos a las obras de Isnard *et al.* (1972), Davis (1973), Durán y Odell (1974).

Vamos a centrarnos en el que hemos realizado aquí, y que se encuentra disponible en la biblioteca de programas estadísticos BMDP (Dixon, 1975). El tipo de cluster realizado es el llamado jerárquico, y el procedimiento es ligeramente distinto según tratemos con variables o casos. El procedimiento es el siguiente en el cluster de casos: una vez fijada la distancia a utilizar se toman los dos puntos más cercanos y se unen en un grupo. Dicho grupo se dice que se ha formado a un nivel que es la distancia a la cual se encuentran los puntos y se considera como un nuevo punto cuyas coordenadas se calculan como la media de las coordenadas de los puntos iniciales. Se repite entonces el procedimiento con los puntos resultantes (uno menos que antes) hasta que todos los puntos estén en el mismo grupo (la media de cada iteración se calcula ponderada con el número de casos de cada uno de los grupos que se unen en un paso dado). La única diferencia de procedimiento en el cluster de variables estriba en que, cuando se unen dos grupos para formar uno mayor, no se considera que las coordenadas del nuevo punto sean la media ponderada de las anteriores, sino que se calculan las distancias directamente del nuevo cluster a los demás por uno de los tres procedimientos disponibles: distancia mínima (distancia entre dos grupos, es la mínima distancia de una variable del primero y una del segundo); distancia promedio (la distancia es la media de las distancias de una variable del primer cluster a una del segundo) y distancia máxima (la distancia es la máxima existente). Se obtiene así para cada nivel, α , una clasificación de los puntos en grupos (se consideran los grupos que se han formado a nivel igual o inferior a α); dicha clasificación evoluciona desde el nivel cero, en que cada individuo es un grupo haciéndose cada vez menos fina, hasta que todos los individuos estén en el mismo grupo. Las distancias utilizadas en nuestro trabajo son la correlación en el caso de cluster de variables y la euclídea $\sqrt{\sum_j (x_{1j} - x_{2j})^2}$

en el caso de cluster de casos. Existen otros tipos de distancias disponibles en los programas del BMDP.

Por último, quisiéramos hacer observar que el cluster es un ejemplo de cómo un procedimiento matemático simple se puede convertir, con ayuda del ordenador, en un potente instrumento de análisis.

Los resultados se suelen presentar en unos gráficos llamados dendrogramas (figs. 2 y sigs.), que permiten una visualización rápida de los resultados. Estos se construyen considerando unos ejes de coordenadas. Se colocan en un eje (abscisas, por ejemplo) los puntos con los cuales vamos a hacer los clusters. Se hace salir de cada uno una línea vertical. Si dos puntos se unen a nivel α_1 las dos líneas verticales se unen mediante un segmento horizontal a dicha altura sobre el eje de abscisas. Del centro de este segmento se saca otra línea vertical que representa el nuevo punto de unión de los anteriores. Los puntos en el eje de abscisas se ordenan de forma que no se intersectan líneas de grupos distintos.

Si queremos saber los grupos a un nivel dado α , tendríamos que trazar una línea horizontal a esa altura. Cada vertical que corte corresponderá a un cluster constituido por los

puntos a los que esté unida. Los gráficos que presenta el BMDP no son exactamente los descritos aquí (ya que, entre otras cosas, tienen forma triangular), pero contienen la misma información y son fácilmente transformables en diagramas similares a los explicados.

Ejemplos estudiados

El acuífero de los Llanos de Carchuna

Características generales del acuífero

Este acuífero se ubica en el litoral de la provincia de Granada, y ocupa una superficie de 4,5 kilómetros cuadrados. El material acuífero está integrado por gravas y arenas cuaternarias con matriz y/o intercalaciones de limos y arcillas; los bordes (fig. 1) están ocupados por filitas con cuarcitas, y calizas. Los aspectos geométricos e hidráulicos fueron estudiados por Benavente (1982) y Fernández San Miguel (1982), y las características físico-químicas de sus aguas por Benavente *et al.* (1983).

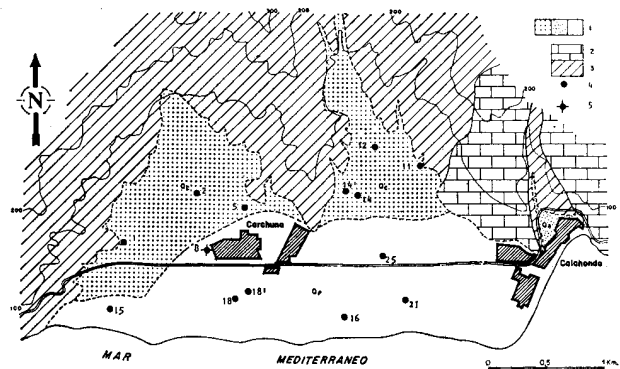


Fig. 1.—Mapa hidrogeológico de los Llanos de Carchuna. 1: depósitos cuaternarios (sin trama; sedimentos de playa). 2: calizas alpujárrides. 3: materiales metapelíticos. 4: pozo o balsa. 5: sondeo. (Tomado de Benavente *et al.*, 1983.)

Gran parte de la franja saturada del acuífero está ocupada por aguas de elevada salinidad, cuyo origen más probable se encuentra en relación con procesos de intrusión marina (Pulido Bosch y Cañada, 1983). El abandono de la mayor parte de los pozos de bombeo, como consecuencia de la traida de aguas para regadío, desde el río Guadalfeo, provoca en la actualidad un "lavado" de las aguas salobres, ayudado por la infiltración en el acuífero de parte de las aguas empleadas en regadío.

Resultados e interpretación de los análisis cluster

Hemos realizado un análisis para diez variables y dieciséis casos (14 pozos, las aguas del canal y el agua del mar —fig. 1—); a continuación exponemos los resultados obtenidos:

Cluster de variables

El dendrograma obtenido se refleja en la figura 2, donde se constata la existencia de correlación extremadamente elevada entre Cl⁻, Na⁺, Cond, Mg⁺⁺, K⁺, So₄⁼ y Ca⁺⁺ (coeficiente de similitud cercano a 100 y r superior a 0,98). Ello expresaría que cualquiera de las variables antes citadas puede obtenerse con el

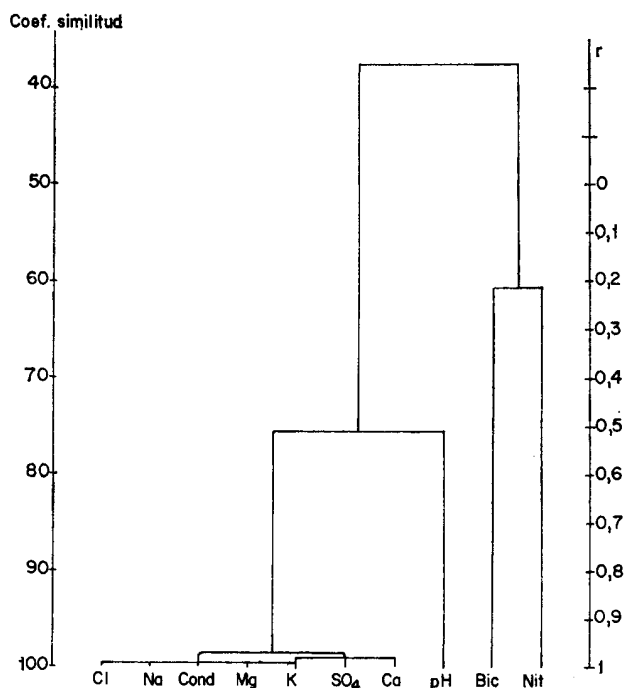


Fig. 2.—Dendrograma de las variables (Carchuna).

simple análisis de una de ellas, por correlación. Con el conjunto de las variables se forma un "cluster" o "taxón", a los que se une el pH (no medido *in situ*, por lo que su representatividad es dudosa), con un coeficiente de similitud de 75 (r ~ 0,51).

Los iones bicarbonato y nitrato no están correlacionados positivamente con el resto de los parámetros medidos, cosa, por otro lado, normal en este tipo de agua; no obstante, ambas variables presentan entre sí un coeficiente de similitud de 61 (r = 0,21), que, aunque no es muy elevado, constituye un criterio que permite afirmar que ambos parámetros guardan una cierta relación, en cualquier caso superior a la que guardarían con los parámetros restantes; en efecto, como es sabido, los procesos de descomposición de la materia orgánica, traen consigo la generación de CO₂, confiriendo agresividad a las aguas, y, por tanto, mayor contenido en CO₃H⁻; por otro lado, estos mismos procesos de descomposición de la materia orgánica van a generar nitratos, y, en menor cuantía, nitratos (Custodio y Díaz, 1976; Pulido-Bosch, 1980).

Sin embargo, ambos cluster presentan una corre-

lación negativa de -0,25 (coeficiente de similitud de 37), lo cual sería indicativo de que el aumento en uno de ellos traería consigo la disminución en el otro, hecho constatable, por ejemplo, en las aguas marinas.

Cluster de casos

Es interesante hacer este análisis (fig. 3) atendiendo a las distancias crecientes entre los casos; así, por ejemplo, a la distancia de 0,2, tan sólo se tienen dos cluster, uno de ellos integrado por las aguas de los puntos acuíferos 1 y 12; y el otro por las del 14' y 8, unidos ambos al 25; el primero de los cluster engloba dos puntos cuyas aguas pondrían de manifiesto la aportación de los bordes metapelíticos poco mineralizados; las aguas del segundo cluster citado serían representativas de la mezcla, en elevada proporción, de las aguas del canal, con las del acuífero.

A una distancia inferior a 0,5 son tres los cluster formados; al integrado por los puntos 1 y 12 se une el punto 15, que indicaría que sus aguas son las "genuinas" del acuífero; al cluster formado por los puntos 25, 14' y 8, se unen las aguas de los puntos 14 y 21, interpretándose el conjunto como característico de la mezcla de aguas "propias" del acuífero, con las del canal; el nuevo cluster formado, está constituido por las aguas de los puntos 5 y 2, relativamente cercanos a los bordes, y cuyo contenido iónico podría tener una cierta similitud con el contenido iónico de las aguas del mar; ello es, en parte, corroborado por el hecho de que a dicho cluster se une, a una distancia de 0,74, el punto número 11, de elevado contenido salino, cuyo origen más probable se

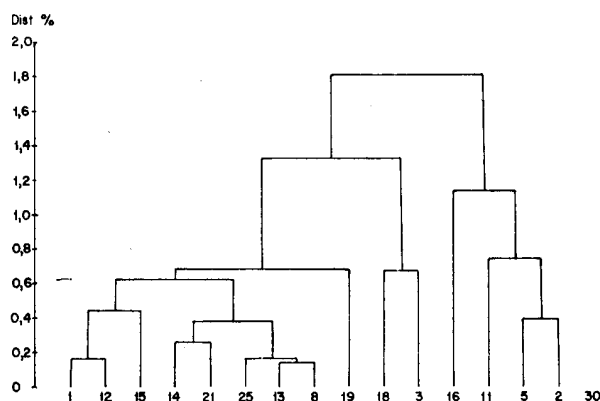


Fig. 3.—Dendrograma de los casos (Carchuna).

encuentra en un proceso de intrusión marina, aunque no se descarta la posibilidad de que las sales tengan su origen en tramos evaporíticos existentes en el borde metapelítico (Benavente, 1982; Benavente *et al.*, 1983).

A ese mismo nivel (< 0,8) son tres los cluster exis-

tentes; además del reseñado, se tiene uno integrado por los puntos 1, 12, 15, 14, 21, 25, 14' y 8', que incluiría todas las aguas con una composición comprendida entre las aguas propias del acuífero y las aguas del canal; el otro cluster incluye las aguas del canal, y las del punto 18, que serían las de mayor similitud y, por tanto, indicarían una acusada recarga en ese sector a partir de las aguas del canal.

A mayor distancia ($\sim 1,4$) el conjunto de los casos (puntos acuíferos) quedan agrupados en dos familias —sin considerar el agua del mar— que corresponderían a las aguas de puntos acuíferos mayormente afectadas por la recarga inducida procedente del canal, cuyo orden creciente de influencia sería: 1, 12, 15, 14, 21, 14', 8, 18' y 18; y la otra, que mostraría mayor afinidad en la composición de sus aguas con la del mar (¿o materiales evaporíticos del borde?), que serían: 16, 11, 5 y 2. Estas dos familias se funden en una sola, a la distancia de 1,8, lo cual permite concluir que existe más afinidad entre las aguas del acuífero, las del canal y las de su mezcla en proporciones diversas, que con las aguas del mar.

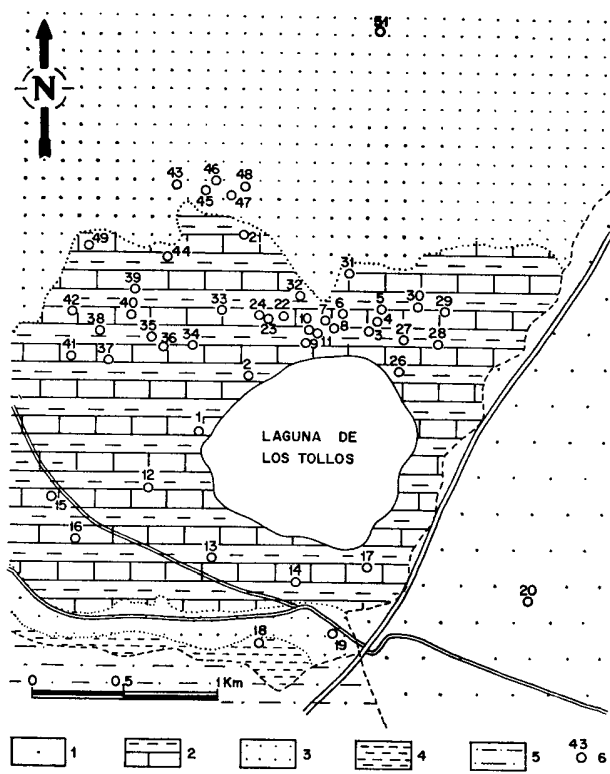


Fig. 4.—Mapa hidrogeológico del entorno de la Laguna de los Tollos. 1: glaciis. 2: sedimentos lacustres. 3: arenas. 4: arcillas margosas. 5: margas blancas miocénicas. 6: pozo o sondeo y su número de orden. (Tomado de Lozano *et al.*, 1983.)

El entorno de la Laguna de los Tollos

Características hidrogeológicas

La Laguna de los Tollos se sitúa en el límite de las provincias de Sevilla y Cádiz, a unos dos kilómetros al este de la localidad de El Cuervo (Sevilla). Las características geológicas del área en cuestión, así como la localización de los puntos acuíferos muestreados se reflejan en la figura 4.

El sector forma parte del “acuífero de Lebrija” (FAO-IGME, 1968), que ocupa una superficie cercana a 75 kilómetros cuadrados. Los materiales acuíferos, en este área, están integrados por arenas del Plioceno y sedimentos detríticos cuaternarios; recubren a las arenas una serie de sedimentos lacustres formados por arcillas y margocalizas con sílex. La síntesis hidrogeológica de este sector, ha sido publicada recientemente por Lozano *et al.* (1983), y a la misma nos remitimos para una mayor información.

El número de puntos acuíferos considerados para los análisis cluster ha sido de 39, correspondientes a aguas de pozos y sondeos ubicados tanto en las arenas pliocenas como en el recubrimiento lacustre o en el “glaciis”. El objetivo principal del análisis ha sido determinar la existencia de familias de aguas que respondiera a situaciones hidrogeológicas afines.

Cluster de variables

El dendrograma obtenido se refleja en la figura 5; se puede observar que es muy diferente al obtenido para el caso de los Llanos de Carchuna. Del mismo se deduce, además, que las variables más correlacionadas son el Cl⁻ y el Na⁺, seguidas muy de cerca por

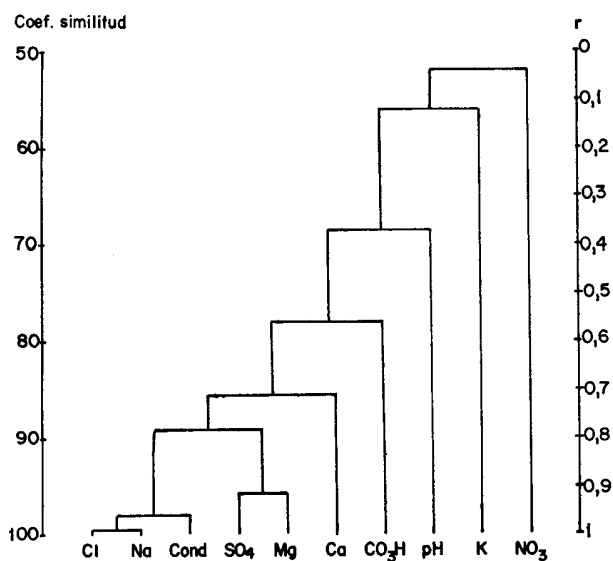


Fig. 5.—Dendrograma de las variables (Tollos).

la conductividad, lo cual permite obtener cualquiera de ellas la partir de la otra, como cabría esperar. La correlación es igualmente muy elevada entre el $SO_4=$ y el Mg^{++} , y éstas, a su vez, con los parámetros anteriormente citados; sin embargo, pH, K^+ y NO_3^- , aunque los coeficientes de similitud no bajan de 50, presentan muy escasa correlación.

Cluster de casos

Se observa (fig. 6) la existencia de numerosas familias y de casos aislados que no se integran en el conjunto hasta distancias muy elevadas. A la distan-

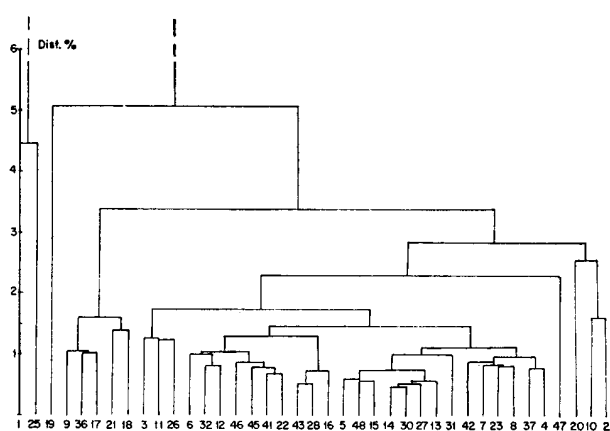


Fig. 6.—Dendrograma de los casos (Tollos).

cia de 1,15 se forma una gran familia, integrada por 14 muestras de agua (5, 48, 15, 14, 30, 27, 13, 31, 42, 7, 23, 8, 37 y 4), las cuales tienen en común el hecho de corresponder a pozos y sondeos que captan el acuífero de las arenas, así como el presentar una salinidad bastante elevada; dentro de esta familia, a su vez, se podrían diferenciar varias subfamilias, que se unen a distancias menores (familia F1).

A 1,3 de distancia se forma otra familia (F2) que agrupa un total de diez muestras de agua (6, 32, 12, 46, 45, 41, 22, 43, 28 y 16), igualmente correspondientes a muestras de pozos que captan las arenas, agrupadas en tres subfamilias, la primera de las cuales incluye pozos equidistantes de la Laguna, y las otras dos, pozos cuya lámina de agua era inferior a cuatro metros en la época de muestreo.

A distancia similar, e incluso algo inferior, se forma otra familia integrada por sólo tres muestras (pozos 3, 11 y 26), que tienen en común el hecho de no alcanzar las arenas, una lámina de agua del orden de 3,5 metros y un elevado contenido salino; esta familia (F3) incluiría los pozos cuya salinidad tiene como origen el lixiviado de las sales acumuladas en algunos sectores de los materiales lacustres que recubren las arenas.

Otra familia (distancia 1,6; F4), integrada por las aguas de muestras de los pozos 9, 36, 17, 21 y 18, tiene en común el registrar una lámina de agua, de contenido salino medio, cercana a 3,5 metros, y el estar muy cercanos a la laguna; ninguno de estos pozos capta las arenas. Corresponderían al sector del recubrimiento de las arenas que se encuentra poco salinizado (sedimentos del borde de la antigua laguna, de menor contenido en sales evaporíticas). Previamente, y a 1,45 de distancia, las familias F1 y F2 se funden en un solo cluster (C1).

A 1,75 de distancia la familia F3 se funde en un cluster con el C1 (C2); a este cluster se añade el pozo 47 (excavado en arenas y de elevado contenido salino de sus aguas), a una distancia de 2,30.

A 2,55 de distancia se diferencia otra familia (F5) constituida por las aguas de los puntos acuíferos 10, 22 (unidos entre sí previamente a una distancia de 1,60) y 20, que corresponden a captaciones excavadas en el relleno lacustre (no penetran en las arenas) cuyas aguas presentan elevado contenido salino.

Esta familia se une al cluster C2 a 2,85 de distancia. A ese nivel permanecen aún sin unir los puntos acuíferos 1, 25 y 19, lo cual pondría de manifiesto que sus aguas tienen una composición muy peculiar.

Aplicación a la evolución temporal

Características del punto seleccionado

Se trata del manantial de las Aguas (Pulido-Bosch, 1975 y 1979), situado en el límite de las provincias de Valencia y Alicante, controlado por el IGME desde 1973, con muestreo de periodicidad variable. Este manantial constituye uno de los principales puntos de drenaje de la unidad Benicadell-Almirante-Gallinera, la cual tiene una superficie aflorante de 240 kilómetros cuadrados.

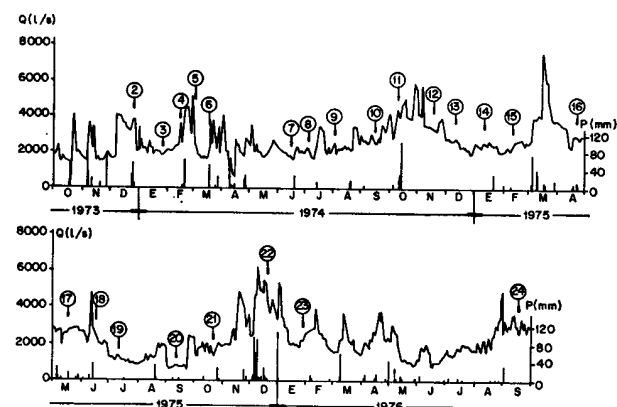


Fig. 7.—Hidrograma de los manantiales de la Sierra de Mustalla (Pulido Bosch, 1979). Los números corresponden a las muestras analizadas.

En la figura 7 se muestra el hidrograma conjunto de la descarga en la sierra de Mustalla (manantiales de las Aguas, Solinar y Salado), con indicación de las precipitaciones ocurridas y localización de las muestras estudiadas, numeradas de 2 a 24, por orden cronológico. Las muestras 1 y 25 corresponden a antes y después del período de registro de caudales, respectivamente.

Hemos realizado dos análisis diferentes: uno con las 25 muestras (una correspondiente al ciclo 72/73, nueve al 73/74, diez al 74/75, cuatro al 75/76 y una al 76/77), y otro con 19 (tan sólo los ciclos 73/74 y 74/75), decidido así a la vista de los resultados obtenidos.

Cluster de variables

A) Con 25 muestras

En este caso se observa, al igual que en los anteriores, que Cl⁻ y Na⁺ están muy altamente correlados (similitud de 96,5 y r=0,93); la conductividad se une al cluster anterior con una similitud de 90 (r=0,8); por lo que cualquiera de estos parámetros puede ser obtenido, con gran fiabilidad, a partir de uno solo de ellos; los restantes parámetros medidos correlacionan peor, en especial CO₃H⁻, Mg⁺⁺, NO₃⁻; curiosamente el potasio presenta mayor similitud con el SO₄⁼ que con el resto de los parámetros; de todos ellos, el CO₃H⁻ correlaciona negativamente con el cluster formado por el resto de los parámetros determinados, aunque con muy bajo coeficiente de correlación (fig. 8).

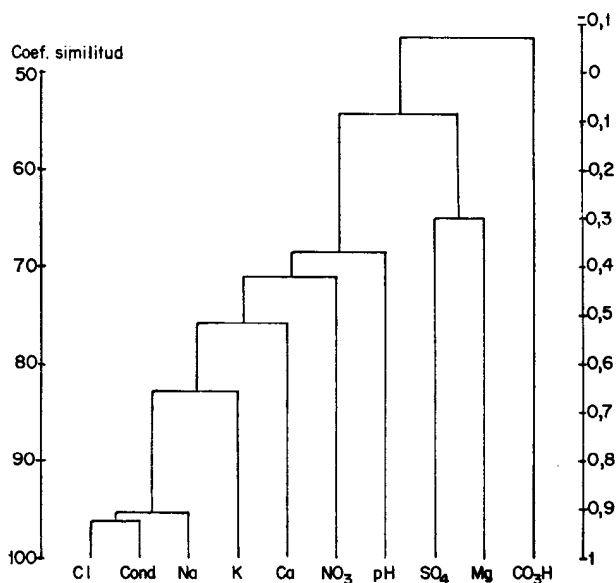


Fig. 8.—Dendrograma de variables (25 muestras).

B) Sólo dos ciclos

El dendrograma obtenido (fig. 9) tras eliminar los datos de los tres ciclos muy incompletos, es sensiblemente diferente; en efecto, se mejora la similitud entre Cl⁻, conductividad y Na⁺; además, SO₄⁼ y Mg⁺⁺ forman un cluster a r=0,3, mientras que en el caso anterior lo formaba con el K⁺; se mantiene la correlación negativa del CO₃H⁻; por último, es constatable el aumento general del coeficiente de similitud entre casi todas las variables.

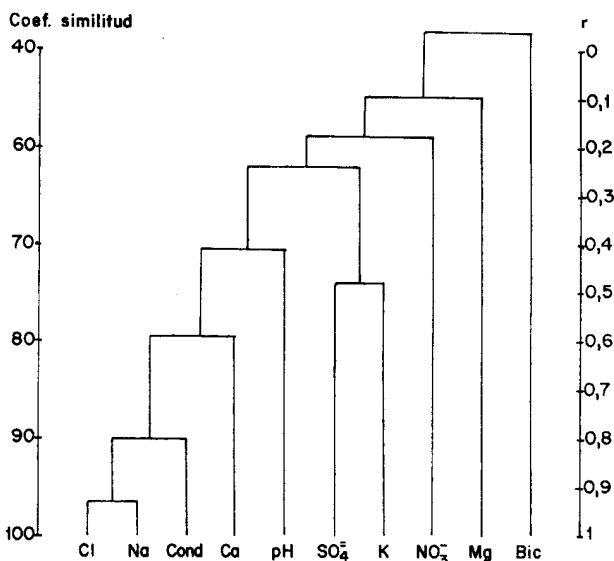


Fig. 9.—Dendrograma de variables (19 muestras).

Cluster de casos

El dendrograma obtenido se refleja en la figura 10; destaca del mismo la existencia de dos grandes familias individualizadas a una distancia de 3; además, quedan al margen tres análisis (1, 25 y 24).

De las dos familias citadas, la situada más a la derecha está integrada por los análisis 2, 3, 4, 5, 6, 7, 8, 9 y 10, todos ellos correspondientes al año hidrológico 1973/74; dentro de esta familia el primer cluster que se forma es el integrado por los análisis 5 y 2 (distancia de 1,15), los cuales tienen en común el hecho de corresponder a fechas con lluvias (2) o inmediatamente después (5); a distancias crecientes se unen los análisis 6 y 10 (principio y final del estiaje, respectivamente), y los números 8, 9 y 7, todos ellos correspondientes al estiaje. Finalmente, se une el 3, y al cluster formado por todos los restantes se une el análisis 4; estos dos últimos corresponden al período invernal.

La segunda gran familia agrupa a los dos años hidrológicos siguientes (1974/75 y 1975/76), aunque

el segundo de ellos incluye tan solo cuatro análisis muy irregularmente repartidos en el tiempo. Dentro de esta segunda familia se observan, a su vez, cuatro subfamilias; la primera está integrada por los análisis 14, 15, 16, 13 y 12, en este orden, y que correspon-

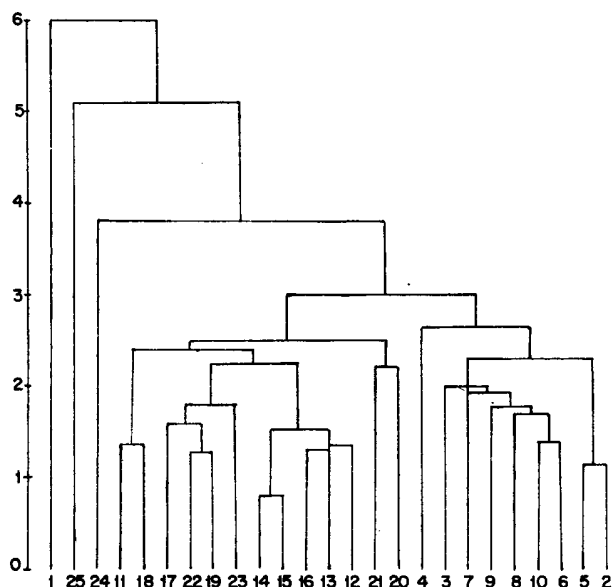


Fig. 10.—Dendrograma de casos (25 muestras).

den al período de recarga (noviembre 1974, abril 1975); otra subfamilia está constituida por los análisis 19, 22, 17 y 23, en donde se mezclan análisis correspondientes al estiaje con los de períodos muy lluviosos; otra subfamilia es la formada por los nú-

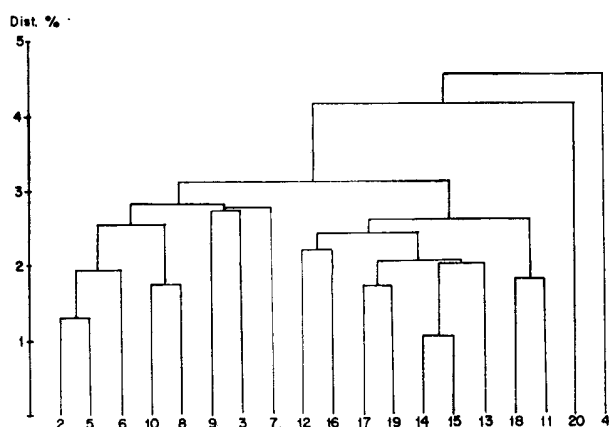


Fig. 11.—Dendrograma de casos (19 muestras)

meros 11 y 18, que corresponden a muestras tomadas tras las lluvias; y, por último, las muestras 21 y 20, que corresponden al final del estiaje.

La existencia de ciclos en surgencias kársticas, en lo relativo a los caudales, es un hecho bien conocido,

reflejo, por otro lado, de la existencia de ciclos en la alimentación. La existencia de ciclos en las características físico-químicas de las aguas del karst, es algo mucho menos evidente, aunque la interpretación de los resultados del cluster apuntan hacia ello. Para poder afirmar esto último, sería preciso contar con un muestreo más denso, y aplicar otras técnicas de tratamiento de datos más sofisticados, como son los análisis de correlación y espectral (Mangin, 1981; Mangin y Pulido-Bosch, 1983).

El resultado tan interesante de este primer análisis, nos indujo a repetir el tratamiento, pero esta vez con solo los dos ciclos mejor representados (73/74 y 74/75), obteniendo el dendrograma de la figura 11. Se puede observar que ambos ciclos vuelven a manifestarse a distancias de 2,6 (74/75) y 2,8 (73/74), respectivamente, aunque una muestra de cada ciclo (20 y 4) permanecen fuera de los cluster formados a las distancias citadas.

Conclusiones

Las técnicas del análisis cluster pueden constituir un buen auxiliar en la descripción de las características físico-químicas de las aguas, tanto al nivel de los parámetros controlados (permite determinar cuáles de ellos pueden ser obtenidos por correlación, sin necesidad de realizar la determinación analítica), como al de los casos a estudiar.

Del estudio de los casos se puede deducir la existencia de familias de aguas, independientemente de su ubicación espacial, que arrojan información sobre la actuación de procesos similares incluso a distancias elevadas.

Dado que estas técnicas trabajan exclusivamente con los datos numéricos como tales, se hace necesario un "filtrado" subjetivo de los casos a estudiar y de los parámetros a tomar en consideración, eliminando previamente aquellos resultados analíticos muy diferentes del resto, pues pueden contribuir a complicar notablemente la interpretación de los resultados; en este sentido, sucede algo similar con el análisis factorial (Mangin, 1981).

Este procedimiento tiene la ventaja sobre el análisis factorial de ser utilizable en una variedad más amplia de aplicaciones; hay muchos casos en que, debido a que el análisis factorial implica unas hipótesis más restrictivas, se obtienen unos resultados difíciles de interpretar. Esto no ocurre así con el análisis cluster, cuyos resultados admiten siempre una interpretación más evidente.

Agradecimientos

Al profesor doctor Gutiérrez Jaimez, Catedrático de Estadística de la Facultad de Ciencias de Granada, por la lectura del manuscrito y sus sugerencias.

Bibliografía

- BAKALOWICZ, M.
1979. *Contribution de la Géochimie des eaux à la connaissance de l'aquifère karstique et de la karstification*. Thèse Univ. Paris VI. Lab. Sout. CNRS, Moulis, 269 págs.
- BENAVENTE, J.
1982. *Contribución al conocimiento hidrogeológico de los acuíferos costeros de la provincia de Granada*. Tesis doctoral. Universidad de Granada, 571 páginas (inérita).
- BENAVENTE, J.; FERNÁNDEZ-RUBIO, R.; FERNÁNDEZ-SAN MIGUEL, M.; PULIDO-BOSCH, A. y SÁNCHEZ-FRESNEDA, V.
1983. *Características físico-químicas de las aguas subterráneas de los Llanos de Carchuna (Granada)*. Hidrogeol. y Rec. Hidrául., 8, 253-262, Madrid.
- BLANC, J. J. y MONTEAU, R.
1983. *Essai sur l'analyse des cavités karstiques du Massif de Marseille-Veyre et des archipels de Riou et du Frioul (Marseille)*. Karstologia, 1, 55-60.
- CUSTODIO, E. y DÍAZ, E.
1976. *Calidad del agua subterránea*, en "Hidrología Subterránea", sección 18, Edit. Omega, Barcelona.
- DAVIS, J. C.
1973. *Statistics and Data Analysis in Geology*. John Wiley & Sons, 550 págs., New York.
- DIXON, W. J. Edit.
1975. *Biomedical Computer Programs (BMDP)*. Univer. Calif. Press., 791 págs., Berkeley.
- DURÁN, S. y ODELL, L.
1974. *Cluster analysis. A survey*. Econometrics, número 100, Springer-Verlag.
- FAO-IGME
1968. *Proyecto piloto de utilización de aguas subterráneas para el desarrollo agrícola de la cuenca del Guadalquivir*. AGL: SF/SPA 16, Madrid.
- FERNÁNDEZ-SANMIGUEL, M.
1982. *Estudio hidrogeológico del sector Llanos de Carchuna-Calahonda*. Dep. Hidrog., Universidad de Granada, 201 págs. (inérito).
- IBÁÑEZ, V. y SANCHIS, E.
1983. *Utilización de modelos estadísticos multivariantes en Hidrogeoquímica*. Hidrogeol. y Rec. Hidráulica, 9, 263-275, Madrid.
- ISNARD, P.; MALLET, J. L.; CAZES, P. y SATTRAN, V.
1972. *Corrélations géologiques. Méthodes statistiques de traitement des données*, en "Traité d'informatique géologique". Laffite, P., edit., Cap. 9, 379-533. Masson & Cie, París.
- LOZANO, A.; FERNÁNDEZ-RUBIO, R.; PULIDO-BOSCH, A.; CASAS, J. y SASTRE, J. L.
Síntesis hidrogeológica del entorno de la Laguna de los Tollos (Lebrija, Sevilla). "Hidrogeol y Rec. Hidrául.", 7, 565-575, Madrid.
- MANGIN, A.
1981a. *Utilisation des analyses corrélatoire et spectrale dans l'approche des systèmes hydrologiques*. C. R. Acad. Sc. París, 293, 401-409.
1981b. *Probabilités et statistiques: leur emploi en Géologie*. Lab. Sout. C.N.R.S., Moulis, 141 págs.
- MANGIN, A. y PULIDO-BOSCH, A.
1983. *Aplicación de los análisis de correlación y espectral en el estudio de acuíferos kársticos*. "Tecniterrae", 51, 53-65.
- MELLOUL, A.
1979. *Contribution à la connaissance hydrogéologique des grands aquifères profonds en milieu poreux à données peu nombreuses*. "Bull. Cent. d'Hydrog.", 3, 55-94.
- MUDRY, J.; KIRALY, L. y MULLER, I.
1979. *Analyse multivariée du chimisme de quelques sources karstiques du Jura Suisse et Franc-Comtois*. "Bull. Cent. d'Hydrog.", 3, 183-221.
- PULIDO-BOSCH, A.
1975. *Los manantiales salinos de la sierra de Mustalla*. Jorn. Min. Metal., Bilbao, 3, 117-128.
1979. *Contribución al conocimiento de la Hidrogeología del Prebético nororiental (provs. de Valencia y Alicante)*. Mem. "IGME", núm. 95, 410 págs., Madrid.
- PULIDO-BOSCH, A. y CAÑADA, P.
1983. *Estudio de acuíferos costeros mediante prospección geofísica eléctrica —Aplicación a los Llanos de Corchuna—*. "Hidrogeol. y Rec. Hidráulica", 9, 363-374, Madrid.
- PULIDO-BOSCH, A.; GARCÍA, J. y BENAVENTE, J.
1980. *Las fuentes potenciales de contaminación en los acuíferos de Salobreña (Granada)*. "Tecniterrae", 33, 31-38.

Recibido el 9 de enero de 1984.
Aceptado el 5 de octubre de 1984.