

# Geología 94

## II CONGRESO CUBANO DE GEOLOGIA Y MINERIA GEOLOGIA Y MINERIA'94

Durante los días 22, 23, y 24 de noviembre se desarrollará en la ciudad de Santiago de Cuba el Segundo Congreso Cubano de Geología y Minería, auspiciado por la Sociedad Cubana de Geología y el Ministerio de la Industria Básica.

El Congreso estará dedicado al intercambio de información acerca de todos los aspectos de la geología y minería de Cuba y del Caribe, aunque se aceptarán también trabajos de otras partes del mundo.

Las temáticas principales serán:

- Geología regional, estratigrafía, tectónica, paleobiogeografía, sedimentología, petrología, metamorfismo, paleontología, geofísica, paleomagnetismo.
- Metalogenia y mineragenia, prospección de petróleo y gas, minerales metálicos y no metálicos, materiales de construcción.
- Hidrogeología e ingeniería geológica.
- Sismología y movimientos recientes de la corteza terrestre.
- Cortezas de intemperismo lateríticas.
- Geología del carso.
- Geología ambiental, sensores remotos, y técnicas de computación aplicadas a la geología.
- Minería.
- Sesiones especiales.

Se ofrecerán actividades colaterales que permitirán conocer la geología de la región oriental de Cuba, la ciudad de Santiago de Cuba, su vida cultural y social.

Se espera la participación de 250 especialistas cubanos y 150 extranjeros, fundamentalmente de países latinoamericanos y caribeños.

Para cualquier información sobre el congreso pueden dirigirse a:

Geología'94  
Apartado 370, Habana 10100  
Ciudad Habana, Cuba.

Geología'94  
Palacio de las Convenciones de Cuba  
Apartado 16046, La Habana, Cuba  
Télex: 511609 palco cu  
Fax: 22-8382 y 33-1657  
Correo electrónico: palco ext@cenial.cu



## UTILIZACION DE TECNICAS ESTADISTICAS MULTIVARIADAS EN LA VALORACION DE LA CALIDAD DE LAS AGUAS QUE ABASTECEN A LA POBLACION DE MOA

Ing. Joaquín Proenza Fernández  
Ing. José Batista Rodríguez  
Ing. Adela Caridad Terrero Abella

Instituto Superior Minero Metalúrgico

**RESUMEN:** El análisis de componentes principales y factores, el análisis cluster y la regresión múltiple son métodos de la estadística multivariada que han alcanzado una gran aceptación en los estudios hidroquímicos debido a los buenos resultados que se han obtenido con su aplicación. En este trabajo se hace una valoración del uso de estas técnicas y se muestra su utilización en el estudio de la calidad de las aguas que abastecen a la población de Moa. Se clasifica el acuífero en diferentes zonas en dependencia de grado de contaminación, se confirman los criterios que se tenían sobre la dirección predominante y foco de la contaminación, se comparan estos resultados con otros anteriores y finalmente se demuestra la validación del uso de estos métodos.

**ABSTRACT:** The analysis of principle components and factors, the cluster analysis and the multiple regression are methods of multivariate statistics that have gained a great utility in the hydrochemical studies due to the good results obtained during its application. This work evaluates the use of these techniques and shows its utilization in the quality study of waters that supply the Moa population. The aquifer has been classified into different zones according to the grade of its contamination. Confirmed are the criteria of determining the predominating direction of the contamination, and its source. Shown in the work also are the previous results. In addition this work shows the validation of the usage of these methods.

### INTRODUCCION

En el estudio de contaminación de las aguas, el investigador debe analizar e interpretar de forma integrada un conjunto de atributos (pH, mineralización, dureza, etc.) que caracterizan los distintos puntos muestreados. Es probable que algunos de estos atributos sean redundantes, por estar relacionados en menor o mayor grado entre sí y ser portadores de la misma información.

También se hace necesaria la clasificación del área estudiada en grupos homogéneos en cuanto a los parámetros que la caracterizan, definiendo sectores con diferentes grados de contaminación.

En este trabajo se muestra la potencialidad del uso de técnicas estadísticas multivariantes para la solución de los problemas anteriormente citados. Se seleccionó el principal acuífero que abastece a la población de Moa del que se poseen datos considerables sobre la composición química de las aguas.

Estudios precedentes sobre la calidad de las aguas en el principal acuífero del municipio Moa, han reportado la contaminación que presentan las aguas subterráneas y la tendencia al aumento de la misma, teniendo como causa principal la presencia de una presa de residuales (colas) de una industria de níquel.

### FUNDAMENTOS BASICOS DEL ANALISIS MULTIVARIADO

La estadística multivariada es aquella que estudia, interpreta y elabora el material estadístico sobre la base de un conjunto  $n > 1$  variables.

Análisis de componentes principales: consiste en transformar un conjunto inicial de  $m$  variables en un nuevo conjunto también de  $m$  variables, de forma que las nuevas se obtengan como combinación lineal de las variables originales. El proceso se lleva a cabo de manera que cada nueva variable contenga la mayor proporción posible de la

varianza total del conjunto original, con ello se consigue un resultado en que las variables no están correlacionadas y se ordenan según un grado decreciente de varianza.

Este análisis se emplea para reducir la cantidad de variables observadas sin pérdidas notables de la información contenida en los datos.

En la literatura se ha reportado el uso de las primeras componentes principales como una vía para la determinación de la tendencia regional en el procesamiento de datos

tridimensionales y su aplicación con éxito en geomorfología, sedimentología, hidrogeología, meteorología y geofísica.

**Análisis de factores:** su principio básico consiste en postular que sobre las variables originales actúan sólo un número *q* de factores generales, que son denominados factores comunes. El resto de *p-q* factores se denominan no comunes o singulares y se asume que la influencia de

### TECNICAS DE CLASIFICACION QUE NO REQUIEREN INFORMACION A PRIORI

Entre ellas se encuentran las llamadas técnicas de análisis de agrupaciones (Cluster Analysis) en la literatura especializada. Entre las técnicas de agrupación se encuentra la clasificación jerárquica según el algoritmo de Ward-Wishart, método que tiene como objetivo representar los vínculos entre los puntos observados en el espacio de las variables, de manera que refleje una estructura jerárquica, en la que cada grupo formado al nivel *i* forma parte de un grupo mayor al nivel *i + 1*. Estos grupos se van unificando unos con otros con el paso a niveles superiores, de tal manera que todos forman parte de un solo conjunto en el nivel jerárquico máximo.

El uso exitoso de estas técnicas se ha reportado en sedimentología, problemas de taxonomía evolutiva en paleontología, en geoquímica para el estudio de la distribución de los elementos del grupo de las tierras raras, estudios hidroquímicos, entre otros.

El algoritmo de Ward-Wishart fue diseñado para obtener grupos con la mínima varianza en cada ciclo; es decir, grupos con la máxima homogeneidad interna.

### UBICACION GEOGRAFICA Y MARCO ECONOMICO

El municipio de Moa se encuentra ubicado al noreste de la provincia de Holguín en la parte oriental de Cuba; en él se encuentran importantes reservas de mineral laterítico por lo que se ha desarrollado un amplio proceso construc-

los mismos sobre el sistema de variables no presenta una regularidad definida en sentido general. Se puede utilizar para revelar e interpretar la existencia de factores no evidentes que controlan las relaciones entre los puntos o las variables observadas en un sistema.

Los resultados de todos los métodos jerárquicos de clasificación se representan en forma de tablas llamadas de orden de enlaces o en sus equivalentes gráficos llamados dendogramas.

La aplicación de métodos multivariados a estudios hidroquímicos debe convertirse en una herramienta habitual para el procesamiento de los datos, pero todavía prevalecen los métodos clásicos de representación gráfica como el diagrama de Piper, logarítmicos, triangulares, etc.

La aplicación de métodos multivariantes para el estudio de contaminación no está muy extendida, se deben mencionar los trabajos de Thomas (1986) sobre metodología de control de calidad de aguas subterráneas, el de Font Cistero y A. Navarro Flores (1991) sobre aplicación de métodos estadísticos uni y multivariantes a estudios de contaminación, el caso de la Cubeta de la Ilagosta, (A. Navarro Flores, 1989), que utilizan análisis de componentes principales (ACP) y análisis de factores (AF) para estudiar la contaminación y evaluación geoquímica de las aguas subterráneas del acuífero aluvial del valle de Congost (Barcelona).

tivo e industrial. Tiene en funcionamiento dos grandes fábricas para el procesamiento de los yacimientos ferroniquelíferos y una en construcción, todo lo cual tiene gran impacto sobre el medio ambiente.

### DATOS UTILIZADOS EN EL ANALISIS

Fueron utilizados los datos sobre la composición físico-química de las aguas subterráneas obtenidos en el quinquenio 85-90, para los cuales se tomaron un total de 160 muestras distribuidas en 17 puntos, concentradas en 4

pozos, 4 calas, 6 calicatas, además de 3 puntos en el río Moa. Los parámetros determinados fueron: mineralización (M), dureza (D), pH, SO<sub>4</sub>, Fe, Mn, Cr y Ni.

Tabla 1. Composición físico química promedio de las aguas subterráneas en el período 85-90

Ptos	M g/l	PH	D m-ell	SO <sub>4</sub> mg/l	Fe mg/l	Mn mg/l	Cr	Ni
1	0.48	7.6	2.3	410	0.5	2.3	0.02	0.01
2	0.6	6.8	2.4	480	1.4	2.75	0.03	0.022
3	0.5	7.2	2.3	428	0.68	2.5	0.02	0.01
4	0.4	7.3	2.3	422	0.54	2.3	0.02	0.01
5	0.5	6.8	2.3	422	0.5	2.62	0.02	0.01
6	0.8	6.8	2.4	422	1.8	3.2	0.04	0.03
7	1.0	6.3	2.7	958	3.2	5.8	0.06	0.04
8	2.8	5.9	5.4	3490	3.8	7.8	1.43	0.07
9	0.4	6.8	2.4	424	0.5	2.62	0.02	0.01
10	0.9	6.7	2.6	570	2.3	4.1	0.053	0.038
11	1.8	6.1	3.0	1847	4.2	6.3	0.08	0.051
12	1.8	6.1	3.0	2560	4.3	6.8	0.092	0.052

Tabla 1. (continuación).

Ptos	M g/l	PH	D m-ell	SO <sub>4</sub> mg/l	Fe mg/l	Mn mg/l	Cr	Ni
13	2.0	6.0	4.2	2830	4.3	7.2	1.2	0.068
14	3.3	5.8	6.0	3648	5.1	8.3	1.67	0.086
15	0.2	7.7	1.8	90	1.84	2.3	0.2	0.001
16	0.28	7.8	2.2	91	1.83	2.2	0.2	0.001
17	0.27	7.5	1.9	91	1.83	2.2	0.2	0.001

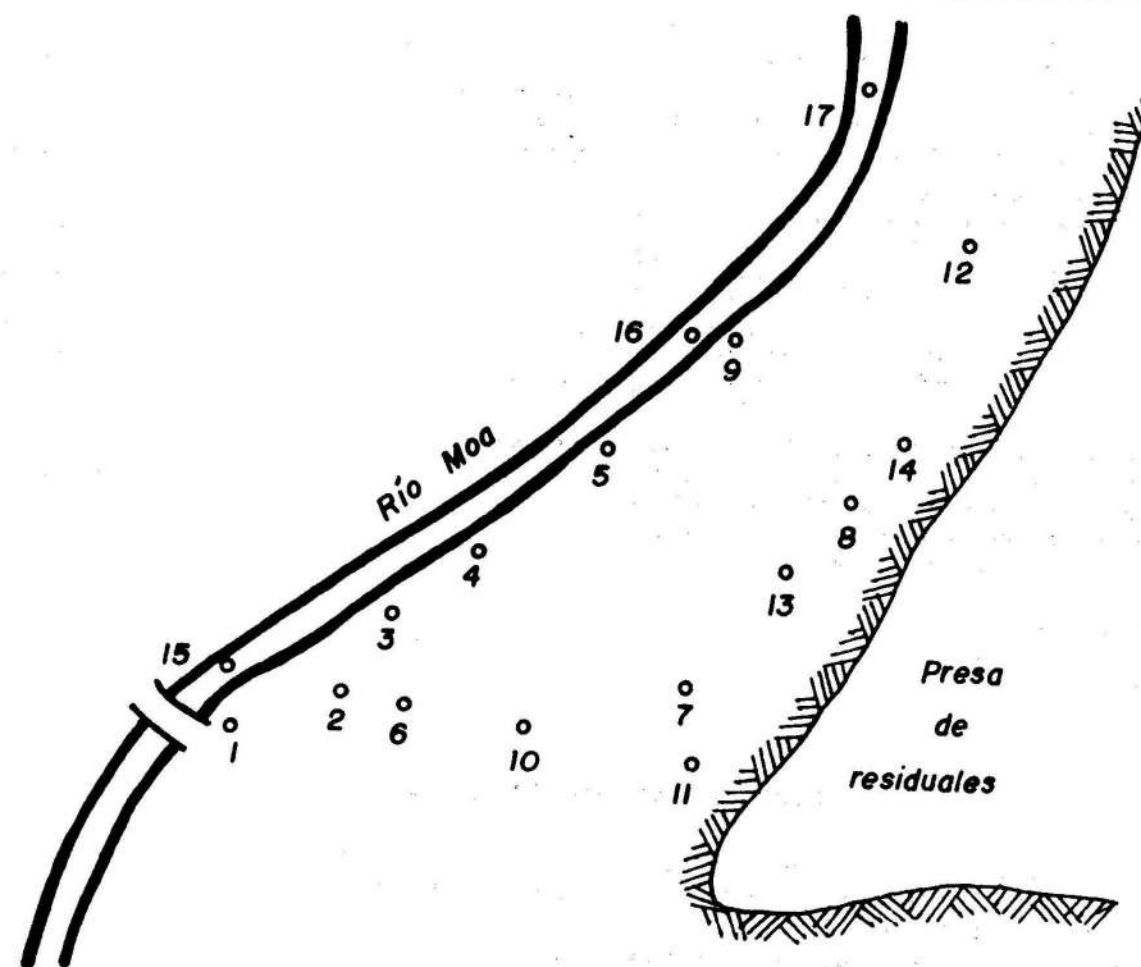


FIGURA 1. Toma de muestras.

### CARACTERISTICAS DEL OBJETO GEOLOGICO

El complejo acuífero de los sedimentos aluviales del Cuaternario se encuentra ubicado en las terrazas del río Moa. Presenta una permeabilidad por intersticios; litológicamente está constituido por depósitos areno-gravosos con intercalaciones de material arcilloso, tiene una potencia de 12-15 m, le subyace el complejo de rocas serpentínicas que le sirve de impermeable.

Las aguas subterráneas son de tipo freáticas y se encuentran en interrelación hidráulica con las aguas del río Moa; en el tiempo de lluvia el acuífero descarga en el río, y en tiempo de seca el río alimenta el acuífero.

Actualmente se encuentran en explotación los pozos 49, 50, 51 y 53, a los que se les extrae un gasto aproximado de 400 l/s, los cuales constituyen la fuente fundamental de abasto de agua a la ciudad de Moa; la profundidad de yacencia de los niveles tiene una variación en el área, con niveles que oscilan entre 2,70 y 5,26 m.

Las precipitaciones atmosféricas constituyen la fuente de alimentación fundamental y en período de seca, las aguas superficiales.



## ANÁLISIS DE LOS RESULTADOS RESULTADOS DEL ANÁLISIS DE COMPONENTES PRINCIPALES Y FACTORES

En el tratamiento numérico de los análisis se utilizaron 17 muestras y para su procesamiento el sistema estadístico MULVAR, elaborado por profesores del Departamento de Geofísica del ISPJAE.

El conjunto de datos tratados abarca un total de 7 variables, cuya relación es la siguiente: mineralización (M), pH, SO<sub>4</sub><sup>-</sup>, Fe, Mn, Cr y Ni, de ésta fue excluida la dureza por su poca variabilidad.

Mediante este sistema se han obtenido los parámetros estadísticos más corrientes en función de 7 variables

**TABLA 2. Parámetros estadísticos**

Variables	X media	Dispersión
M	1.060588	0.9400562
pH	6.77647	0.6628814
SO <sub>4</sub> <sup>-</sup>	1128.412	1234.368
Fe	2.271765	1.574243
Mn	4.193529	2.268323
Cr	0.315	0.543872
Ni	0.029841	0.0274027

Las componentes principales calculadas son 7; sin embargo, las dos primeras explican el 95.22% de la varianza total, es decir, es posible reducir las 7 variables iniciales a 2 nuevas componentes que explican casi toda la variabilidad de la población.

En la matriz de ejes principales, tomando como significativos los valores señalados, se definen estas 2 nuevas componentes relacionadas con las variables primitivas del modo siguiente:

**Componente 1:** está asociado a las variables M, SO<sub>4</sub><sup>-</sup>, Mn, Ni, Fe, Cr y con el pH en sentido contrario. Representa la contaminación de las aguas subterráneas (factor de contaminación), mostrando que un aumento de la M, el contenido de SO<sub>4</sub><sup>-</sup>, Mn, Ni, y Cr trae consigo una disminución del pH, es decir, las aguas son más ácidas.

**Componente 2:** está asociado al Cr y al pH.

**TABLA 3. Resultados del análisis de componentes principales**

Comp. princ. #	V. Propio	% que explica	% acumulativo
1	6.2022	88.6036	88.6036
2	0.4637	6.6241	95.2277
3	0.2147	3.0677	98.2954
4	0.0589	0.8409	99.1363
5	0.0355	0.5075	99.6439
6	0.0162	0.2312	99.8751
7	0.0087	0.1250	100.00

**TABLA 4. Matriz de ejes principales**

Variables	Vector #1	Vector #2
M	0.39564	0.10976
pH	-0.36434	0.47866
SO <sub>4</sub> <sup>-</sup>	0.39297	0.12609
Fe	0.36692	-0.18018
Mn	0.39587	-0.17338
Cr	0.33093	0.81379
Ni	0.39425	-0.13476

### REPRESENTACIONES EN EL PLANO

Utilizando las componentes principales como ejes de coordenadas y el coeficiente de cada muestra respecto al eje, se ha representado la totalidad de las muestras.

En la siguiente figura se han representado las muestras en función de las componentes principales I y II, factores que explican la mayoría de la varianza de la población.

En la misma se pueden observar 4 poblaciones:

**Población 1:** formada por las muestras 8, 13 y 14, tomadas en la zona más próxima a la presa de residuales.

**Población 2:** formada por las muestras 2, 6, 7, 10, 11 y 12, tomadas entre la presa y el río.

**Población 3:** constituida por las muestras 1, 3, 4, 5 y 9, tomadas entre el río y la presa pero en una zona con un grado menor de contaminación que la población anterior.

**Población 4:** constituida por las muestras 15, 16 y 17, tomadas en el río, presentando la mejor calidad de las aguas.

Desde el punto de vista de su composición estas poblaciones pueden clasificarse en:

- aguas sulfatadas hidrocarbonatadas magnésicas cálcicas (Poblaciones 1 y 2).
- aguas hidrocarbonatadas sulfatadas magnésicas cálcicas (Poblaciones 3 y 4).

La población 1 corresponde al foco contaminante, este puede asociarse con el primer componente principal. En la figura 2 se puede observar el buen ajuste de esta población a los ejes I y II. El resto de las poblaciones se encuentran en los cuadrantes negativos del gráfico, observándose en el sentido de las poblaciones 2, 3, y 4 una disminución en el contenido de los atributos que se analizan, esto se justifica por una disminución en los valores de sus coeficientes con respecto al factor de contaminación, es decir, que el contenido disminuye de la población 1 a la 4.

Del análisis de factores se obtuvo la matriz factorial reducida y la matriz de correlación reproducida.

**TABLA 5. Matriz factorial reducida**

Variables	Factor #1	Factor #2	Factor #3	Factor #4
1	0.985	0.075	-0.048	0.088
2	-0.907	0.326	0.236	0.091
3	0.979	0.086	-0.052	0.161
4	0.914	-0.123	0.382	-0.046
5	0.986	-0.118	0.055	0.023
6	0.824	0.554	-0.038	-0.105
7	0.982	-0.092	-0.061	-0.056

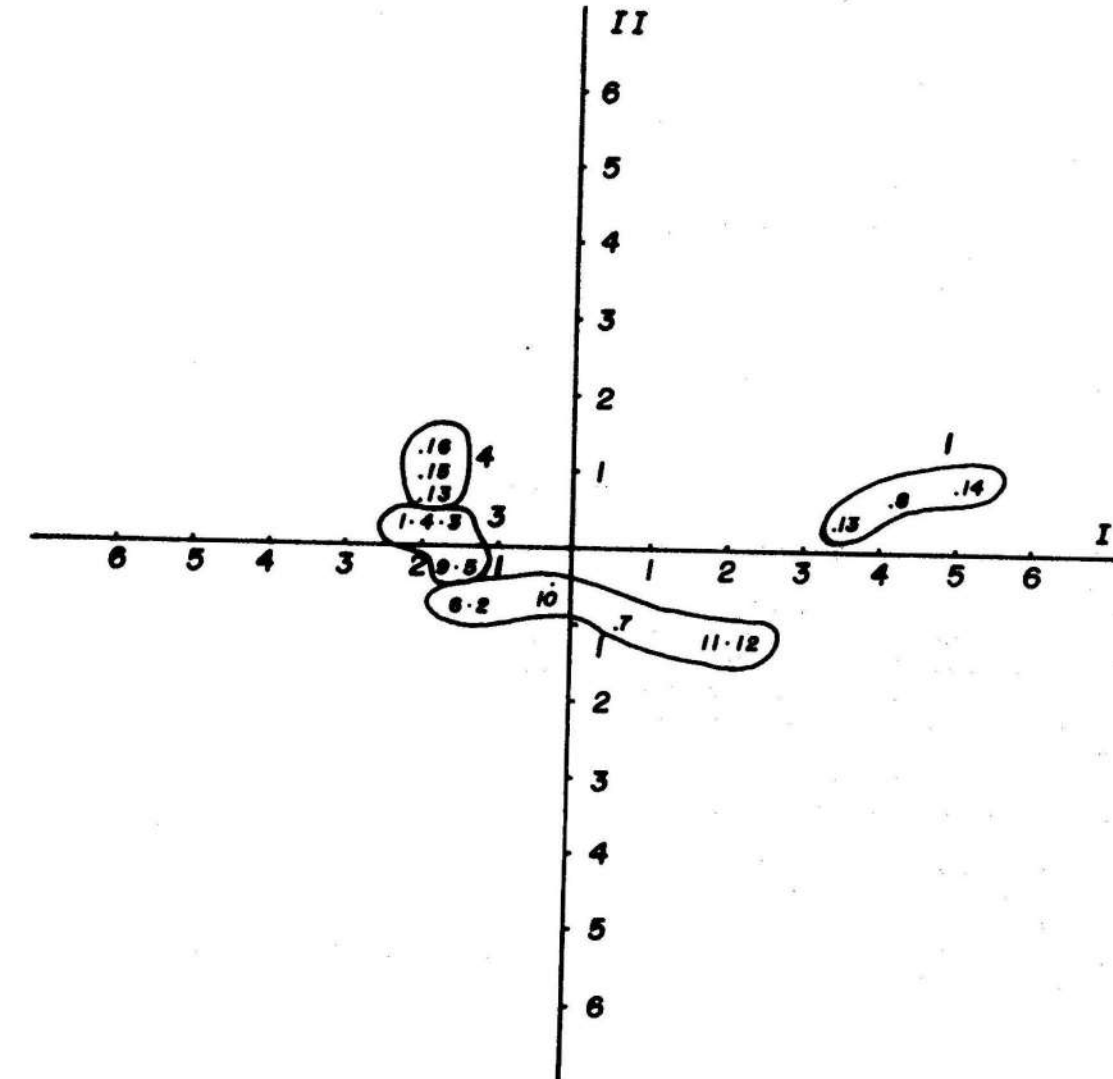


FIGURA 2. Representación de las muestras en función de las componentes I y II.

**TABLA 6. Matriz de correlación reproducida**

0.986	-0.873	0.987	0.869	0.962	0.846	0.959
-0.873	0.994	-0.857	-0.783	-0.918	-0.586	-0.940
0.987	-0.857	0.994	0.857	0.956	0.839	0.947
0.869	-0.783	0.857	0.998	0.935	0.675	0.888
0.962	-0.918	0.956	0.935	0.989	0.743	0.974
0.846	-0.586	0.839	0.675	0.743	0.999	0.767
0.959	-0.940	0.947	0.888	0.974	0.767	0.979

De aquí se obtiene la siguiente información: el primer factor controla la M, SO<sub>4</sub><sup>-</sup>, Fe, Mn, Cr y Ni y además el pH con signo negativo. Esto implica que bajo el efecto del primer componente principal un aumento de M, SO<sub>4</sub><sup>-</sup>, Fe,

Mn, Cr y Ni va acompañado, en términos estadísticos, de una disminución del pH. Esto también se observa en la matriz de ejes principales.

En este caso el análisis factorial no reportó información nueva, sino la misma información que el análisis de componentes principales, la matriz factorial reducida se transformó a una estructura simple empleando el algoritmo varimax (algoritmo de rotación ortogonal más usado en la práctica). La matriz factorial rotada no reveló información nueva de interés, situaciones como éstas ya han sido reportadas en otros estudios de contaminación [4]. Compartimos la opinión de que este criterio de cálculo no es recomendable para aguas subterráneas.

### RESULTADOS DE LA CLASIFICACION JERARQUICA SEGUN EL ALGORITMO DE WARD-WISHART

Utilizando la información de la Tabla 1, se conformó una matriz, la cual se sometió al método de clasificación jerárquica de Ward-Wishart, con los datos tipificados. El

dendograma obtenido de la aplicación del método se muestra a continuación:

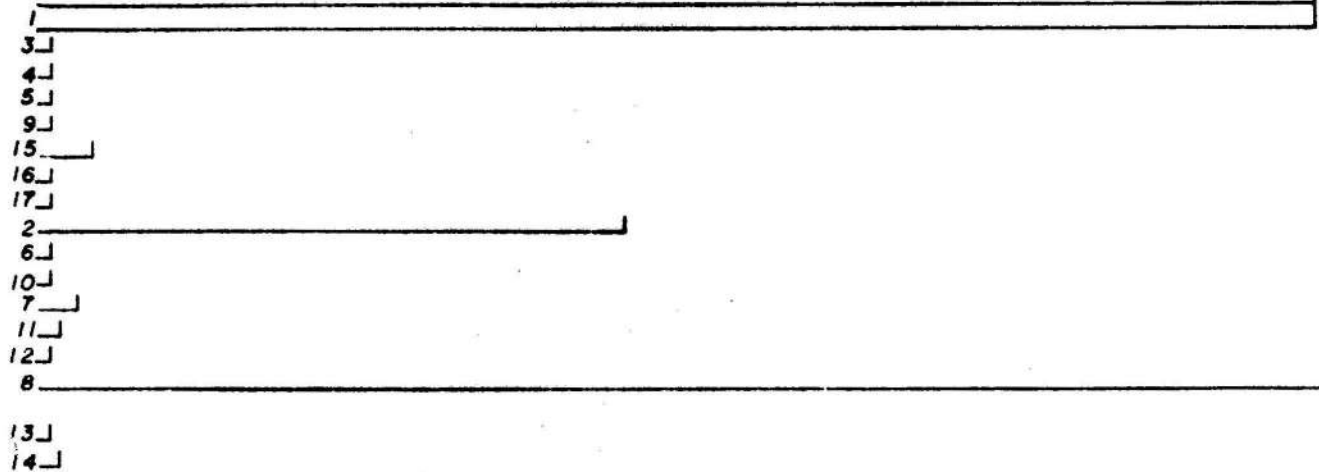


FIGURA 3. Dendrograma.

Del análisis de este dendrograma pueden extraerse los siguientes grupos:

Grupo 1: constituido por las muestras 8, 13 y 14, las cuales al proceder de la zona más próxima a la presa de residuales presentan un grado de contaminación máximo.

Grupo 2: constituido por las muestras 2, 6, 7, 10, 11 y 12, correspondientes a las aguas con nivel de contaminación intermedio.

Grupo 3: formado por las muestras 1, 3, 4, 5, 9, 15, 16 y 17, tomadas más próximo al río. Dentro de ellas pueden diferenciarse dos subgrupos, el primero formado por las muestras 1, 3, 4, 5 y 9 y el segundo por las muestras 15, 16 y 17. De forma general este grupo corresponde a las aguas menos contaminadas.

Los valores de los principales parámetros físico-químicos de cada uno de estos grupos y las marcadas diferen-

cias entre cada uno de ellos pueden apreciarse en la siguiente tabla:

TABLA 7. Valores promedios de los principales parámetros físico-químico de cada grupo

Parámetros	Grupo I (1,3,4,5,9,15, 16 y 17)	Grupo II (2,6,10,7,11 y 12)	Grupo III (8,13 y 14)
M	0.38	1.15	2.7
pH	7.33	6.46	5.9
SO <sub>4</sub>	297	862	3322
Fe	1.06	2.86	4.4
Mn	2.39	4.78	7.76
Cr	0.02	0.06	1.43
Ni	0.01	0.04	0.074

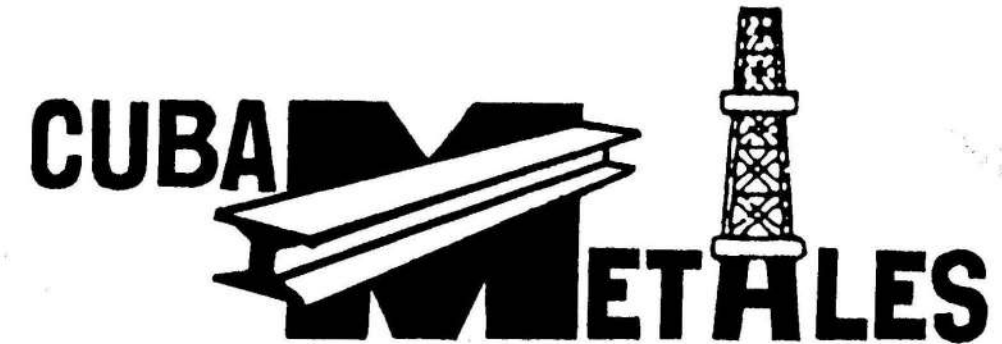
#### CONCLUSIONES

1. Se pone de manifiesto cómo mediante un tratamiento estadístico multivariado aplicado al acuífero que abastece a la población de Moa, es posible establecer diferentes zonas de acuerdo al grado de contaminación, así como precisar la dirección y foco de la misma.
2. Tanto el análisis de clasificación empleado (algoritmo de Ward-Wishart), como el de componentes principales y factores, muestran la presencia de 4 poblaciones o grupos con diferente grado de contaminación.
3. Las técnicas multivariadas usadas mayoritariamente en estudios geofísicos, geoquímicos, etc, han demostrado su eficiencia para estudios de contaminación.
4. En el mundo de hoy donde la contaminación de las aguas crece constantemente, la utilización de técnicas como éstas permiten un estudio integral y más eficiente de la misma, permitiendo revelar información que no puede ser ofrecida por otro tipo de método.

#### REFERENCIAS

1. ALFONSO ROCHE, J.: Estadística en las Ciencias Geológicas. Ed. ISPJAE, Ciudad de La Habana, tomo II, 1989.
2. FONT, X. y A. NAVARRO: "Aplicación de métodos estadísticos uni y multivariantes a estudios de contaminación. El caso de la cubeta de la lagosta", en Boletín Geológico y Minero, sept-oct, pp. 71-91, 1991.
3. LOPEZ, A. y B. CALVO: "Aplicación de algunas técnicas estadísticas al estudio de la distribución de los elementos del grupo de las tierras raras en el macizo granítico de Lugo", en Boletín Geológico y Minero, mayo-junio, pp. 82-93, 1990.
4. NAVARRO, A.: "Contaminación y evaluación geoquímica del agua subterránea del acuífero aluvial del valle del Congost (Barcelona)", en Boletín Geológico y Minero, enero-febrero, pp. 111-133, 1989.

5. PAJON, J. y otros: "Caracterización geoquímica y geomatemática de formaciones geológicas y sedimentos de la cuenca del río Cuyaguaje. Relación con las características hidroquímicas de los acuíferos", en Voluntad Hidráulica, vol. 72, pp. 43-49, 1986.
6. RIAZA, A. y A. BEL-LAN: "Evaluación de una técnica de tratamiento digital de imágenes. Aplicación del análisis de componentes principales a una imagen TM en un área granítica del macizo Hispérico central", en Boletín Geológico y Minero, julio-agosto, pp. 25-37, 1987.
7. TERRERO, A.: "Estudio de los elementos antropogénicos que alteran la calidad de las aguas subterráneas de los sedimentos aluviales del río Moa". (Inédito.)
8. Manual del usuario del sistema MULVAR, versión 2.0, 1989, Departamento de Geofísica del ISPJAE.



**ACERO  
PARA EL DESARROLLO  
ENTRE EN CONTACTO  
CON NOSOTROS**

EMPRESA IMPORTADORA  
DE METALES, COMBUSTIBLES Y LUBRICANTES

**STEEL  
FOR DEVELOPMENT  
GET IN TOUCH  
WITH US**

METAL, FUEL AND LUBRICANTS  
IMPORTING ENTERPRISE