



MODELO NO PARAMÉTRICO PARA EL DISEÑO Y OPTIMIZACIÓN DE REDES DE MONITOREO DE AGUAS SUBTERRÁNEAS

NON-PARAMETRIC MODEL FOR THE DESIGN AND OPTIMIZATION OF GROUNDWATER MONITORING NETWORKS

LESLIE F. MOLERO-LEÓN

INVERSIONES GAMMA, S.A. PO Box 6219, CP 10600, Habana 6, La Habana, Cuba. E-mail: especialistaprincipal@gmail.com

Palabras clave: Resumen

aguas subterráneas karst monitoreo optimización Se describen los elementos básicos de un Modelo No Paramétrico para el diseño y optimización de redes de monitoreo de aguas subterráneas desarrollado por el autor y aplicado desde 1978 sistemáticamente en diferentes entornos geológicos (karst, rocas fisuradas no cársicas, acuíferos granulares) del Trópico Húmedo (regiones continentales, pequeñas islas, acuíferos costeros) y con objetivos específicos en El Caribe, Latinoamérica, África y Oceanía. El procedimiento se basa en la identificación mediante técnicas geomatemáticas de las características dominantes de las componentes geométricas de la red (detalles constructivos y distribución espacial), contenido y frecuencia de muestreo mediante la combinación de técnicas de análisis factorial, clasificación numérica, autocorrelatorio, espectral y de varianza, heterogeneidad y anisotropía de las series temporales de las variables que caracterizan el régimen de las aguas subterráneas en un entorno dado. La tecnología puede utilizarse para el diseño inicial de una red de monitoreo como para optimizar, por el cambio en las condiciones de contorno y operación, la composición, distribución espacial, contenido y frecuencia de monitoreo en una red establecida.

Key words: Abstract

groundwater karst monitoring optimization The paper describes the basic elements of a Non-Parametric Model for the design and optimization of groundwater monitoring networks developed by the author. The procedure has been systematically applied since 1978 in different geological environments (karst, non-karstic fissured rocks, granular aquifers) of the Humid Tropics (continental regions, small islands, coastal aquifers) in the Caribbean are described. Latin America, Africa and Oceania. The technology is based on the identification by geomathematical techniques of the dominant characteristics of the geometric components of the network (construction specifications and spatial distribution), network composition, selected variables and sampling frequency through the combination of techniques of factor analysis, numerical, auto and cross-correlations, spectral and variance analysis of the time series, clustering, classification of heterogeneity and anisotropy. The technology can be used for the initial design of a monitoring network as well as to optimize, by changing boundary and operating conditions, the composition, spatial distribution, content and frequency of monitoring in an established network.

Introducción

La primera Red Nacional de Monitoreo del Régimen y la Calidad de las Aguas Subterráneas con alcance nacional en Cuba comenzó a implementarse en la década de 1970 bajo la dirección técnica de un grupo de colaboradores búlgaros en el entonces Grupo Hidráulico Nacional del DAP (Desarrollo Agropecuario del País, una entidad aglutinadora de centros de gestión y construcción agropecuaria e hidráulica que, entre

otras, asumió funciones y recursos del Instituto Nacional de Recursos Hidráulicos al que absorbió en 1968).

En la Escuela Hidrogeológica Soviética añadir “*calidad de las aguas*” a la denominación de la red resulta redundante, en tanto “*por régimen de las aguas subterráneas se entienden las variaciones en el espacio y en función del tiempo de los recursos, las propiedades y su composición (comprendidos los niveles, los caudales, las velocidades, la temperatura, la composición química, gaseosa y bacteriológica que reflejan*

Recibido: 23 de enero de 2025

Aceptado: 10 de febrero de 2025

Conflicto de intereses: El autor declara no tener ningún conflicto de intereses con respecto a la publicación de este artículo. Este artículo se encuentra bajo los términos de la licencia Creative Commons Attribution-NonCommercial (CC BY-NC 4.0). <https://creativecommons.org/licenses/by-nc/4.0/>



el proceso de formación de las aguas subterráneas” (Kliméntov y Kónonov, 1982). Un concepto debido a Kamenski, que en 1938 lo definió como “la combinación de fenómenos asociados con los cambios temporales (fluctuaciones del nivel, descara y velocidad de flujo) así como cambios en la composición química y propiedades físicas” (Altovsky, 1959; Bochever, 1968; Galabov, 1983, 1985; Mijailov, 1989).

Como era común en aquella época, el diseño de la Red consideraba la construcción de pozos alineados regularmente en trayectos de pozos paralelos a la dirección de flujo, distribuidos en retículos equidistantes para facilitar la interpolación a partir de las técnicas disponibles entonces y que requerían de intervalos espaciales regulares para ello. De modo semejante aun antes, en la década de 1940, fueron construidos los perfiles de pozos de monitoreo de la intrusión marina TS y TG, por la Comisión de Fomento Nacional en la Cuenca Sur de La Habana y, distribuidos de manera no tan semejante en las de Vento y Ariguanabo, que databan de principios del siglo XX. Dominó en aquellos años iniciales del establecimiento de la Red Nacional, una aproximación cualitativa del diseño, sustentada en la interpretación no formalizada matemáticamente de los fundamentos teóricos establecidos en lo fundamental por la literatura metodológica soviética. La retícula regular permitía no solamente la interpolación estricta mediante triangulación, sino resolver algunos problemas relacionados con la aplicación de los métodos hidráulicos e hidrodinámicos del cálculo de los recursos de agua subterránea o la infiltración efectiva mediante modelos deterministas, como el de Kamiensky para el cálculo de la infiltración efectiva (De Miguel, 2012).

Salvo muy contadas excepciones se tomaron en cuenta los indicadores de **heterogeneidad y anisotropía** de los sistemas acuíferos, que en Cuba son predominantemente cársicos (66,1 % de la superficie del país y poco más del 80% de los acuíferos) propiedades que, por añadidura, se fueron cuantificando poco a poco, pero los modelos se simplificaron en demasía bajo una concepción que generalizó el concepto de la dinámica de medios porosos homogéneos e isotrópicos al dominante karst del país. El desarrollo de soportes automatizados y nuevas técnicas de interpolación (aun no generalizadas, como el kriging. Funciones Spline, interpolaciones bilineales, interpolación con espaciamiento variable de Lagrange, interpolación multivariable, entre otras) permitió, no obstante, resolver el problema de la irregularidad de la distribución espacial en muchos casos pero aún hoy se mantiene, sin embargo, una fuerte componente de incertidumbre, sobre todo geológica que afecta la gestión eficiente de los recursos hídricos subterráneos y, en no poca medida, de los superficiales en tanto en el karst, sobre todo, son un único y mismo recurso (Molerio y Sardiñas, 2024).

Resultó fundamental tomar en cuenta el eventualmente distorsionador **efecto del factor de escala** (Fig. 1) sobre

la estructura del campo de propiedades físicas; sobre todo, cuando en la red se combinan pozos y manantiales (particularmente en el karst) por la diferente magnitud acuífera que interesan y el valor intrínseco y regionalizado de la propiedad medida. Lo propio ocurre al tratar la homogeneidad relativa (Fig. 2; Tablas 1 y 2) y su relación con la anisotropía del sistema acuífero (Bear, Zaslavsky e Irmay, 1968; Kiraly, 1975, 1978; Borevskii, Samsonov y Yazvin, 1979; Molerio, 1984, 1985; Bear y Cheng, 2010; Bear et al., 1999; Beat y Yavuz, 1984; Bear y Bachmat, 1990).

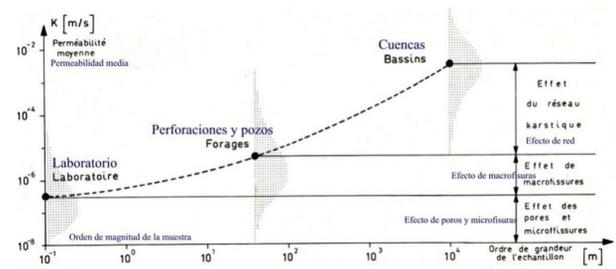


Figura 1. Curva de efecto de escala (Kiraly, 1975, 1978)

Figure 1. Scale Factor Curve (Kiraly, 1975, 1978)

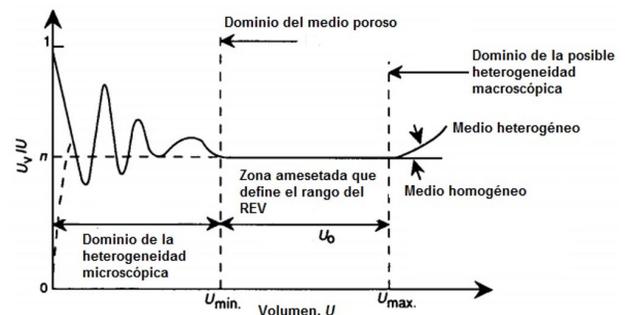


Figura 2. Volúmenes crecientes de muestreo y variación de la homogeneidad relativa (Molerio & Portuondo, 1997)

Figure 2. Increasing sampling volumes and variation of the relative homogeneity (Molerio & Portuondo, 1997)

$$\log \sigma = \sqrt{\frac{\sum_{i=1}^n \log x_i^2 - \frac{\sum_{i=1}^n \log x_i}{n}}{n-1}}$$

$$CV = \sqrt{\frac{\sum_{i=1}^n \left(\frac{x_i}{x_0} - 1\right)^2}{n-1}}$$

La **variabilidad espacial** de las propiedades que caracteriza el sistema de aguas subterráneas introduce un número importante de consecuencias negativas en la evaluación de los recursos disponibles y en la toma de decisiones respecto a la gestión del recurso hídrico. A la

Tabla 1. Grados de homogeneidad relativa de los sistemas acuíferos a partir de la estructura del campo de propiedades físicas, según Borevskii, Samsonov y Yazvin (1979) y Piñera, Molerio y March (1982, 1984) (σ , desviación estándar; C.V., Coeficiente de variación)

Grado de homogeneidad	$\log \sigma$	C.V.
Homogéneos	< 0,2	< 0,4
Heterogéneos	0,2 - 0,4	0,4 - 0,8
Muy heterogéneos	0,4 - 0,75	0,8 - 1,5
Extremadamente heterogéneos	> 0,75	> 1,5

Tabla 2. Número necesario en series de transmisividad para la valoración del grado de certidumbre según Borevskii, Samsonov y Yazvin (1979) y Piñera, Molerio y March (1982) (σ , desviación estándar)

Table 2. Necessary terms in Tansmissivity series to evaluate the certainty degree after Borevskii, Samsonov & Yazvin (1979) and Piñera, Molerio & March (1982) (σ , desviación estándar)

$\log \sigma$	95%	90%	80%	70%	50%
0,2	22	15	9	6	3
0,3	45	32	20	13	6
0,4	79	55	34	22	10
0,5	119	84	51	34	15
0,6	171	120	74	49	21
0,8	306	215	131	86	36

variabilidad espacial, que en principio es una consecuencia directa de la **heterogeneidad** de las rocas y del sistema de colectores conductores de las aguas que circulan por el macizo, debe añadirse la que es provocada por la **anisotropía** de las propiedades físicas. También es de esperar una cierta **dependencia del tiempo** de algunas de las propiedades del campo físico y, obviamente, las que caracterizan la composición química o la calidad de las aguas de los acuíferos son variables en las que el tiempo influye de manera decisiva.

En la década de 1990, el autor, en su posición de Director del Servicio Hidrogeológico de Cuba recibió la encomienda de intentar mantener la informatividad del sistema de monitoreo del régimen de las aguas subterráneas del país con una reducción considerable de los costos de operación. Se desarrolló entonces un recurso tecnológico que permitió remodelar la red en los acuíferos principales del país aplicando técnicas geomatemáticas -sobre todo de Estadística No Paramétrica- y de Teoría del Información combinadas. Algunas de estas técnicas ya habían sido aplicadas por nuestro equipo de trabajo para resolver problemas puntuales relacionados con la dinámica del régimen de las aguas subterráneas específicos por lo que se contaba con algunos intentos exitosos que requerían de una evaluación a mayor escala. El trabajo de optimización de las redes hidrogeológicas se interrumpió bruscamente hacia 1998 y sigue siendo una tarea no resuelta que merece continuidad, sobre todo, por los nuevos desafíos que enfrenta la gestión eficiente de los recursos hídricos subterráneo ante la crisis económica, la demanda creciente de agua y el Cambio Climático.

Este artículo resume las acciones básicas de la tecnología de optimización para la distribución espacial, detalles constructivos, contenido y frecuencia de muestreo aplicando técnicas estadísticas no paramétricas para el caso de una red ya construida y en operación. (Molerio, 1992, 1998, 2013; Molerio y Portuondo, 1997; Molerio, Portuondo y Torres, 1998; Yera y Molerio, 1998; Molerio y Torres, 2002; Molerio et al., 2002a, 2002b; Molerio y Toujague, 2004; Toujague y Molerio, 2013).

Diseño y optimización de redes de monitoreo

El diseño y optimización de una red de monitoreo de aguas subterráneas se fundamenta en los siguientes criterios:

- Aprovechamiento óptimo de la información de archivo y, caso de existir, de la red de observación hidrogeológica y de calidad de las aguas subterráneas en operación
- Reducción del número de puntos a monitorear y de la frecuencia de las observaciones y muestreos con relación al diseño actual tomando en cuenta: la representatividad de los puntos de la red; la calidad del dato primario; las perspectivas de uso de los recursos hidráulicos superficiales y subterráneos del territorio; y los objetivos de la red de monitoreo
- Elevación máxima de la cantidad de información primaria derivable del levantamiento hidrogeológico reduciendo al mínimo los trabajos de perforación y aforos, que elevan sustancialmente los costos de investigación y

prolongan los trabajos de campo y de adquisición del dato básico, sustituyéndolos - en lo posible- por la documentación minuciosa del relieve y la estructura geológica, apoyados en el empleo de métodos indirectos tales como el análisis morfométrico y la documentación de imágenes aeroespaciales

La práctica ha confirmado que puede lograrse un incremento en la eficiencia económica de las investigaciones hidrogeológicas siempre que se disponga de métodos adecuados para evaluar, continuamente, el valor de uso de las redes de monitoreo de las aguas subterráneas.

La optimización, aquí entendida, **se define como la transformación de una red existente que permite tener igual o mayor informatividad, al menor costo de operación**. Las componentes del costo más importantes son: el número de estaciones de monitoreo, las variables a muestrear, la frecuencia de muestreo y aquellos costos derivados del almacenamiento, recuperación, generalización y divulgación de la información. *Díaz (2002)* ha examinado el caso de diseño de redes con poca información

El diseño u optimización, deben satisfacer los principios siguientes:

- Los objetivos del monitoreo deben ser identificados y cuantificados - en lo posible- para cada uno de los sistemas hidrológicos, lo que incluye la definición de cierta medida de efectividad. Este es el aspecto más complejo en la optimización de las redes, ya que se complica por los diferentes objetivos que, para distintos usuarios, puede poseer la red. Estos objetivos incluyen la estimación del estado actual de calidad, la detección de tendencias a largo plazo, violaciones en las normas de calidad de las aguas subterráneas o de los diseños de explotación, y la simulación matemática de acuíferos o del uso conjunto de las aguas superficiales y subterráneas
- Identificación de los procesos hidrodinámicos e hidrogeoquímicos más importantes, toda vez que ellos indican, en cada caso, la forma y los métodos de investigación y de análisis de los datos.
- Determinación de la efectividad de la información derivada del análisis de los datos que proporciona la red o de la data colectada para conformar la matriz inicial de evidencias. Muchas veces, tal efectividad puede relacionarse con conceptos estadísticos tales como la varianza de las muestras, la varianza explicada mediante el análisis factorial, la probabilidad de ocurrencia de un evento cualquiera o el error de interpolación, entre otros
- Cálculo del costo de la red y del programa de monitoreo en su conjunto
- Análisis de la relación costo-beneficio o costo-efectividad, aspecto en extremo complejo e, incluso, con una componente subjetiva muy importante. Debe tomar en

consideración la importancia relativa de las estaciones de muestreo y las variables a observar en cada una de ellas. En la práctica, sin embargo, se requiere de un nivel mínimo de efectividad que se basa en criterios subjetivos, consideraciones económicas y aun políticas que, muchas veces, no toman en cuenta el óptimo mínimo del monitoreo, sobredimensionan la red, o se ignora o sobrestima la necesidad real de adquisición del dato primario

En el caso de redes preexistentes, para proceder a la optimización, las series cronológicas de niveles, composición fisico-química, de indicadores de calidad de las aguas subterráneas, se someten a un:

- Procesamiento y análisis crítico de la informatividad y costo de adquisición y procesamiento del dato básico
- Procesamiento estadístico de las series cronológicas de cada una de las estaciones de la red existente y validación de los resultados de su operación
- Análisis multivariado y de clasificación numérica de los indicadores geométricos del acuífero asociados a los puntos de la Red y de la serie cronológica de cada uno de ellos, a fin de determinar los factores que condicionan el régimen, los parámetros de diseño de las estaciones de monitoreo y el nivel de informatividad de cada uno de ellos
- Comparación de los niveles de informatividad de cada grupo de series de acuerdo con la frecuencia de muestreo
- Automatización del procesamiento del dato básico del régimen de niveles de las aguas subterráneas e incremento de su efectividad para la operación de las redes y de los recursos de explotación de los acuíferos
- Categorización de la importancia socio - económica, política y ambiental de los sistemas acuíferos a monitorear
- Validación de los resultados del diseño para la regionalización más precisa del sistema acuífero
- Perfeccionamiento del modelo conceptual del sistema hidrológico

El caso más complejo es, obviamente, aquel en que ha de diseñarse, por vez primera, una red de monitoreo y no existe información previa de los indicadores hidroambientales. Para resolver este problema, se requiere la preparación y procesamiento de una matriz de evidencias derivable de la prospección y la cartografía Hidrogeológica y del procesamiento digital de imágenes aeroespaciales, a la que se debe añadir información que caracterice o identifique los posibles impactos sobre el régimen y calidad de las aguas que han decidido el diseño de tal red.

Este caso, que será tratado en otra oportunidad se aproxima según el siguiente grupo de operaciones básicas:

- Elaboración del Modelo Conceptual del Sistema Hidrológico

- Procesamiento digital de imágenes aeroespaciales y elaboración de la plataforma geológica, geomorfológica e hidrográfica del territorio. Definición de las condiciones de accesibilidad, definición de impactos sobre el régimen y la calidad de las aguas subterráneas y evaluación preliminar del costo de adquisición y procesamiento del dato básico
- Documentación hidrogeológica de campo, que incluye inventario, muestreo y aforo de puntos de agua, cartografía geológica y geomorfológica
- Definir la estacionalidad de los estímulos sobre el sistema (lluvia, evapotranspiración, escurrimiento superficial) y la respuesta de éste mediante técnicas combinadas de procesamiento estadístico y cartografía de campo
- Análisis multivariado y de clasificación numérica de los indicadores geométricos del acuífero a fin de determinar los factores que condicionan el régimen, los parámetros de diseño de las estaciones de monitoreo y el nivel de informatividad de cada uno de ellos incluyendo decidir si la red responderá a un modelo Euleriano, Lagrangiano o mixto
- Definición de un ciclo de muestreo acorde a los mismos intervalos de la estacionalidad de los estímulos sobre el sistema de mayor importancia hidrológica o ambiental o ambas y validarlo durante un ciclo completo
- Comparación de los niveles de informatividad de cada grupo de series de acuerdo con la frecuencia de muestreo
- Validación de los resultados del diseño para la regionalización más precisa del sistema acuífero
- Perfeccionamiento del Modelo Conceptual del Sistema Hidrológico
- Optimización de la Red.

Las técnicas geomatemáticas conducen a la reducción de las fuentes de incertidumbre en la representación y descripción de las propiedades del sistema (Matheron, 1965, 1970; Agterberg, 1974; Journel y Huijbregts, 1978; Brower y Defize, 1983; Delhomme, 1978; Molerio, 1992, 1997; Portuondo, March y Guerra, 1997; Obdam, 1983; Peck et al., 1988; Pshenichny y Dech, 2015; Quintana et al., 1999; Reymont y Joreskog, 1993).

Etapas de aplicación del modelo

El análisis no paramétrico considera las siguientes etapas de trabajo:

1. Evaluación de la representatividad de la red y la data
2. Identificación de la efectividad de la información
3. Procesamiento de las series temporales:
 - a. Dominio temporal

- b. Dominio frecuencial
 - c. Dominio espacial
4. Eficiencia de la red
 - a. Indicadores constructivos
 - b. Distribución espacial

Efectividad de la información

La efectividad de la información derivable de la red de monitoreo puede evaluarse (Schilperoort y de Groot, 1983), por una medida E, función de la frecuencia de muestreo f, el número y localización de las estaciones de muestreo y el número y tipo de las variables a observar. La función E, en consecuencia, puede derivarse conociendo la estructura de autocorrelación de cada variable, en tiempo y espacio, y la estructura de la correlación cruzada entre todas las variables, también en tiempo y espacio. La coherencia, consistencia y representatividad de la red se evalúa en términos de la geometría y detalles constructivos de las estaciones de monitoreo así como de su posición en el espacio.

Ya en 1946, Bayley y Hammersley establecieron que "la cantidad de información que puede extraerse de una serie cronológica ergódica $x(t)$ univariada y unidimensional, muestreada durante un intervalo de tiempo T, puede relacionarse con el 'número efectivo de observaciones independientes N^* '".

Este indicador es una función del intervalo de muestreo D, el período de observación T, y la función de autocorrelación $\rho_x(\Delta)$, de manera que,

$$N^*(\Delta, T) = N \left[1 + 2 \sum_{i=1}^{N-1} \left(1 - \frac{i}{N} \right) \rho_x(i\Delta) \right]^{-1}$$

Donde, $N=T/D$, número de observaciones con varianza sobre un valor medio:

$$m(x) = \frac{1}{N} \sum_{x=1}^N x(i\Delta)$$

Tal que,

$$\sigma_m^2(\Delta, T) = \frac{\sigma_x^2}{N^*}(\Delta, T)$$

Las relaciones anteriores son útiles para el caso de optimización de la red de una variable unidimensional, toda vez que ellas permiten derivar la relación entre una cierta medida de efectividad E y la frecuencia de muestreo (f) para variables fijas en localidades también fijas.

La efectividad y, por ello, la calidad de la red, puede evaluarse como una función de (excluyendo mediciones reales en la red de monitoreo pero, por otro lado, se requieren datos históricos confiables para estimar la estructura de correlación o para ajustar el modelo de ésta.

Procesamiento de series temporales y dominios del tiempo y la frecuencia

Para evaluar la incertidumbre de los datos de variables del campo de propiedades físicas son útiles, especialmente, las **funciones de autocorrelación, el semivariograma, la distribución de frecuencia y el kriging.**

Para identificar las **componentes determinísticas y estocásticas** de la serie se aplican las técnicas del análisis autocorrelatorio y espectral a las series cronológicas disponibles de piezometría, mineralización e hidrocarburos en las aguas subterráneas, por ser las variables de discriminación de calidad de las aguas más importantes del sistema bajo observación.

El análisis correlatorio ilustra la variación en el dominio del tiempo mientras que el análisis espectral lo hace en el dominio de la frecuencia. El análisis espectral no puede realizarse sin conocer la autocorrelación de las series, toda vez que la densidad espectral no es más que la transformada de Fourier de la función de autocorrelación, pero el inverso no es siempre cierto.

Optimización en el dominio temporal: Función de autocorrelación

La **autocorrelación de una serie temporal discreta** de un proceso X_t es simplemente la correlación de dicho proceso con una versión desplazada en el tiempo de la propia serie temporal. Si X_t representa un proceso estacionario de segundo orden con un valor principal de μ , la autocorrelación se define entonces, como:

$$R(k) = \frac{E[(X_i - \mu)(X_{i+k} - \mu)]}{\sigma^2}$$

Donde E es el valor esperado y k el desplazamiento temporal considerado (normalmente denominado desfase). Esta función varía dentro del rango $[-1, 1]$, donde 1 indica una correlación perfecta (la señal se superpone perfectamente tras un desplazamiento temporal de k) y -1 indica una autocorrelación perfecta. Es una práctica común abandonar la normalización por σ^2 y utilizar los términos autocorrelación y autocovarianza de manera intercambiable.

En el procesamiento de señales, dada una señal temporal $f(t)$, la autocorrelación continua $R_f(\tau)$ es la correlación continua cruzada de $f(t)$ consigo mismo tras un desfase τ , y se define como:

$$R_f(\tau) = f^*(-\tau)$$

O, lo que es lo mismo,

$$f(\tau) = \int_{-\infty}^{\infty} f(t+\tau)f^*(t)dt = \int_{-\infty}^{\infty} f(t)f^*(t-\tau)dt$$

Donde f^* representa el conjugado complejo y el círculo representa una convolución. Para una función real, $f^* = f$.

Formalmente, la autocorrelación discreta R con un desfase j para una señal x_n es:

$$R(j) = \sum_n (x_n - m)(x_{n-j} - m)$$

Donde m es el valor esperado de x_n .

Frecuentemente las autocorrelaciones se calculan para señales centradas alrededor del cero, es decir con un valor principal de cero. En ese caso la definición de la autocorrelación viene dada por:

$$R(j) = \sum_n x_n x_{n-j}$$

Las autocorrelaciones multidimensionales pueden definirse de manera similar. Por ejemplo, en tres dimensiones puede definirse la autocorrelación de una función como:

$$R(j, k, l) = \sum_{n, q, r} (x_{n, q, r} - m)(x_{n-j, q-k, r-l} - m)$$

La mayoría de las propiedades de la autocorrelación unidimensional son extensibles fácilmente a los casos multidimensionales; esto es,

- Simetría: $R(i) = R(-i)$,
- La función de autocorrelación alcanza un valor máximo en el origen, donde adquiere un valor real. El mismo resultado puede encontrarse en el caso discreto.
- Como la autocorrelación es un tipo específico de correlación mantiene todas las propiedades de ésta
- La autocorrelación de una señal de ruido blanco tendrá un fuerte pico en $\tau = 0$ y tendrá valores cercanos a 0 para cualquier otro τ . Esto muestra que el ruido blanco carece de periodicidad.
- El efecto de memoria, tomado como equivalente al Límite de Andersen ($r = 0,2$) en el autocorrelograma es un indicador de la heterogeneidad del acuífero cársico y, en particular del diferente modo en que se estructura el campo de propiedades físicas del sistema. Obviamente existe una correspondencia entre el efecto de memoria del sistema y el ancho de la banda espectral, que equivale al tiempo de regulación

La función de autocorrelación define la pérdida de información en la trayectoria, por lo que siendo,

$$\langle z \rangle = \sum_{k=1}^n \frac{z_k}{n}$$

$$z'_i = z_i - \langle z \rangle$$

Entonces se define la autocorrelación como:

$$C(m\Delta t) \equiv C_m = \sum_{i=1}^n \frac{1}{n}$$

Una serie estacionaria, entonces, puede ser caracterizada por su función de autocorrelación $\rho(k)$ definida por:

$$\rho(k) = \frac{\text{cov}[x(t), x(t+k)]}{\text{var } x(t)}$$

La autocorrelación de la población $\rho(k)$ se estima por el coeficiente de autocorrelación serial $r(k)$ mediante la expresión:

$$r(k) = \frac{\sum x(t) \cdot x(t+k) - \frac{1}{n-k} \sum x(t) \cdot x(t+k)}{\left[\sum x^2(t) - \frac{1}{n-k} (\sum x(t))^2 \right]^{1/2} \left[\sum x^2(t+k) - \frac{1}{n-k} (\sum x(t+k))^2 \right]^{1/2}}$$

Donde todas las sumatorias se llevan desde $t=1$ hasta $t=n-k$.

El análisis de las series estacionarias se complica por el hecho de que las observaciones no son independientes. Ello implica que $\rho(k) \neq 0$ cuando $k \geq 0$. Para $k = 0$, entonces $\rho(0) = r(0) = 1$. Cuando se incrementa el lag k , el número de pares de los elementos que se usan para calcular $r(k)$ disminuye. Por ello, es práctica común fijar el límite superior del lag k en $0,1n$ a $0,25n$, dependiendo de la longitud (n) de la serie. Si la autocorrelación de la población $\rho(k) = 0$ para todos los lags diferentes de cero, el proceso es completamente aleatorio puro. Todas las observaciones, en este caso, son independientes.

El conjunto de coeficientes de autocorrelación $\rho(1), \rho(2), \dots$, graficados como ordenadas respecto a los lags k , constituyen el llamado autocorrelograma (Fig. 3). Los niveles de significación α se definen mediante las expresiones siguientes:

$$(r_k)_u = \frac{1}{n-k} (-1 + z_{1-\alpha/2} \sqrt{n-k-1})$$

$$(r_k)_l = \frac{1}{n-k} (-1 + z_{1-\alpha/2} \sqrt{n-k-1})$$

El correlograma cruzado lluvia-caudal, lluvia-carga hidráulica reproduce, en tanto la lluvia es un fenómeno cuasi-aleatorio la forma del hidrograma y permite obtener una imagen de la respuesta impulsional del sistema. Los hidrogramas de sistemas acuíferos mal drenados en la terminología de Mangin (1975) y con reservas importantes presentan hidrogramas muy extendidos; en tanto aquellos bien drenados y con reservas escasas su hidrograma será estrecho y con picos notables. Este análisis es válido para los casos de determinados iones o especies químicas con independencia de su interpretación formal (Spiridonov y Lopatkin, 1973). En estudios de intrusión marina la movilidad del ión cloruro en el sistema acuífero puede identificarse con éxito de esta manera (Molerio, 2022).

Optimización en el dominio frecuencial: Espectro de varianza

El objetivo del análisis espectral es la descomposición de la serie temporal original en funciones seno y coseno de diferentes frecuencias a fin de identificar aquellas que parezcan más fuertes o importantes (Fig. 4). El análisis espectral y la teoría del filtrado de las señales se han desarrollado especialmente para el análisis de funciones dependientes de tiempo y de series temporales. Las señales pueden corresponder a estados y procesos determinísticos

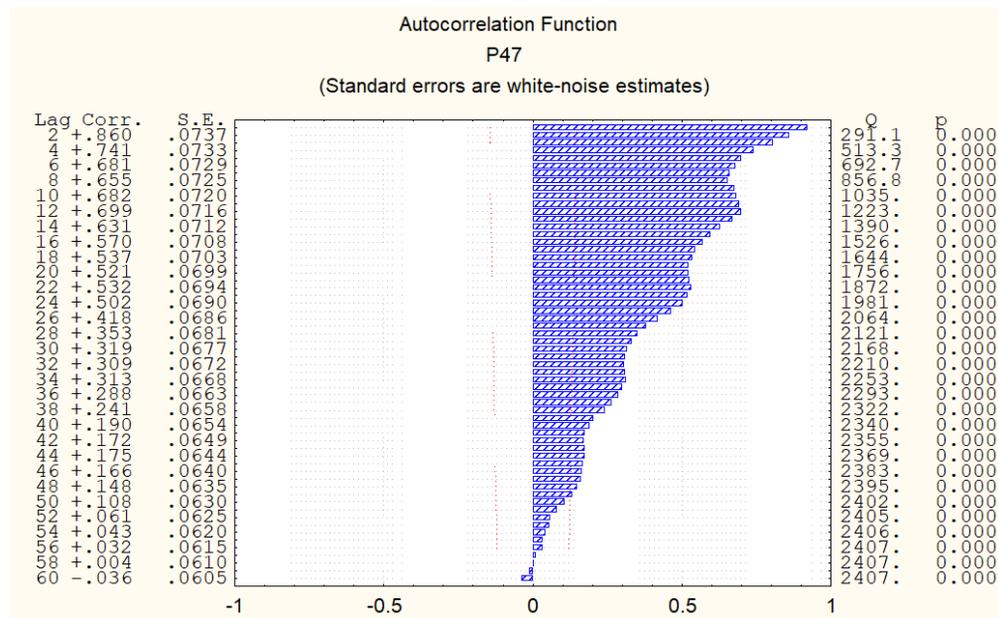


Figura 3. Función de autocorrelación con alto efecto de memoria y respuesta estacional

Figure 3. Autocorrelation function with high memory effect and seasonal response

y aleatorios. Los determinísticos son periódicos y no-periódicos, en tanto los aleatorios son estacionarios, que tienen propiedades espaciales y temporales invariantes, y los no-estacionarios (Buttkus, 2000). Una forma de lograr esto consiste en producir un problema de regresión lineal múltiple en donde las variables dependientes consistan en la serie cronológica observada y las variables independientes sean las funciones seno (discretas) de todas las frecuencias posibles. Tal modelo de regresión lineal múltiple puede expresarse como:

$$x_t = a_0 + \sum_{k=1}^q [a_k \cos(\lambda t) + b_k \text{sen}(\lambda t)]$$

El Teorema de Wiener-Khinchin relaciona la función de autocorrelación con la densidad espectral a través de la Transformada de Fourier:

$$R(\tau) = \int_{-\infty}^{\infty} S(f) e^{j2\pi f \tau} d\tau$$

$$S(f) = \int_{-\infty}^{\infty} R(\tau) e^{-j2\pi f \tau} d\tau$$

Aplicando la notación común del análisis armónico clásico, λ es la frecuencia expresada en radianes por unidad de tiempo ($\lambda=2\pi\nu$), en donde ν es la longitud de onda. Los parámetros a_k y b_k son los coeficientes de regresión que indican el grado en que tales funciones se correlacionan con la data. Por ello, el análisis espectral permite identificar la correlación que existe entre las funciones seno y coseno de diferentes frecuencias, con los datos observados. Si tales coeficientes son altos, en la serie existe una fuerte periodicidad de la frecuencia dada.

Para definir la frecuencia con las mayores densidades espectrales, es decir, las regiones de frecuencia que más contribuyen al comportamiento periódico total de la serie, ha resultado práctico aplicar la ventana de Hamming donde, para cada frecuencia, los pesos para la media móvil ponderada del periodograma se obtienen a partir de la expresión siguiente:

$$w_j = 0,54 + 0,46\cos(\pi j/p)$$

Desde $j=0$ hasta p , donde $w_j = w_{j_0}$, para j_0 , siendo $p=(m-1)/2$ y m , el ancho de la ventana.

La respuesta cíclica del sistema, así como el tiempo en que se filtra completamente la señal de entrada, es muy variable. La señal, en algunos casos, responde a variaciones estacionales, acusando ciclos de 5 y 6 meses perfectamente asociados a la ciclicidad de las precipitaciones. En otros casos, se observan ciclicidades inferiores, del orden de los 2-3 meses, en los que el sistema local de flujo responde rápidamente a la señal de entrada, vinculado a zonas de recarga concentrada y drenaje rápido, con poco almacenamiento, muy probablemente vinculado a sectores altamente transmisivos, quizás hasta de flujo concentrado.

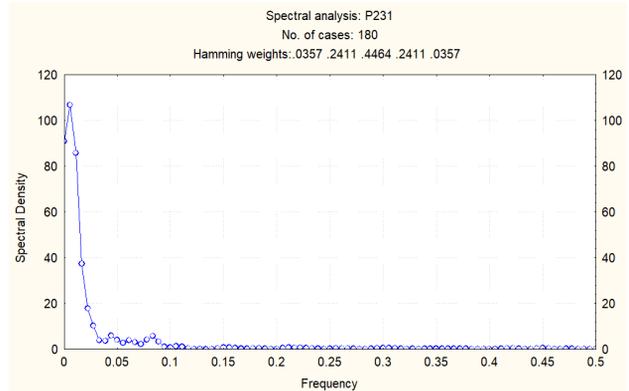


Figura 4. Espectro de varianza de un sistema lineal estable de alta autoregulación (Molerio et al, 2002b)

Figure 4. Variance spectra of a lineal system highly self regulated (Molerio et al, 2002b)

Un caso extremo se reconoce en aquellos puntos que no filtran la señal de entrada. Se trata de sectores locales de flujo rápido y recarga muy concentrada, con muy poco almacenamiento, que la señal de salida muestra oscilaciones muy acusadas en el rango de las altas frecuencias y cuya banda espectral es muy ancha, de manera que las respuestas no se amortiguan antes de los dos meses.

El otro caso extremo está representado por sectores de flujo difuso, de respuesta muy lenta en el dominio de las bajas frecuencias. La banda espectral es muy estrecha y la respuestas del sistema resultan absolutamente independientes de la variación estacional de las precipitaciones (señal de entrada). En tales casos suele reconocerse un filtrado muy lento que coincide, en general, con zonas de alto almacenamiento, flujo difuso e importantes reservas.

En fecha tan temprana como 1959, Altovskii (1959) llamó la atención sobre el uso del periodograma como método para procesar las observaciones históricas del nivel de las aguas subterráneas considerándolo “de suprema importancia para la economía nacional...permitiendo un pronóstico a largo plazo para planear ofensivas contra la salinización y para compilar cronogramas y tasas de riego”, siendo su uso limitado en tanto no exista “suficiente información sobre las variaciones de largo término del nivel de las aguas subterráneas y los métodos para tratar esos datos”. Los trabajos ya clásicos de Mangin (1975, 1981, 1984a, 1984b) son especialmente importantes para identificar las características hidrodinámicas de los sistemas cársicos.

Optimización en el dominio espacial: Semivariograma y complemento del semivariograma

El variograma o semivariograma (Fig. 5) es la representación gráfica de la varianza (la mitad de la diferencia cuadrática) de una muestra de mediciones apareadas en

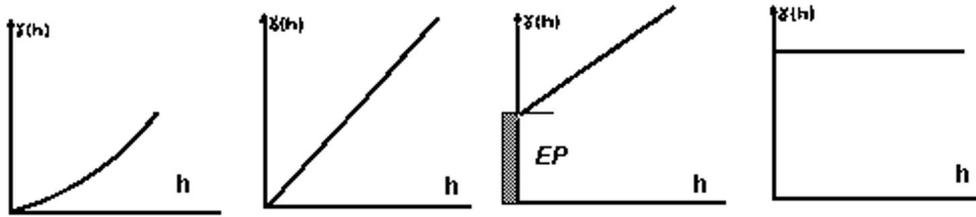


Figura 5. Comportamiento de los variogramas en el origen: a/ derivable; b/ continuo; c/ con efecto de pepita; d/ aleatorio puro (según Delhomme, 1978:348)

Figure 5. Behaviour of variograms at the origin: a/ derivable; b/continuos; c/ with nugger effect; d/pure ramdon (after Delhomme, 1978:348)

función de la distancia y, opcionalmente, de la dirección entre los puntos de la muestra. Todos los posibles pares de muestra se examinan y se agrupan en clases (lags) como veremos el tratar el modo de estimación del variograma, de dirección y distancia aproximadamente iguales. De este modo, el variograma suministra un modo de cuantificar la relación comúnmente observada de que las muestras más cercanas exhiben la tendencia de parecerse más entre sí que aquellas que se encuentran más distantes.

Existe, también, el denominado **variograma no-ergódico**, término aplicado para aquel variograma determinado como resultado de la resta de las covarianzas del lag de la varianza de la muestra. Este tipo de variograma compensa el caso de los variogramas direccionales con fuertes diferencias en las medias en dos direcciones diferentes. Este tipo de variogramas pueden ser modelados y utilizados en el kriging del mismo modo que los variogramas ordinarios.

Por otro lado, el **variograma relativo** es aquel en el cual el valor de cada lag en el variograma ordinario se divide por el cuadrado de la media de la muestra empleada en el cálculo del lag. Es útil, sobre todo, cuando se encuentran **“efectos de proporcionalidad”**; es decir, **cuando áreas con concentraciones más altas que la media tienen también varianzas superiores a la media**. Al utilizar estos variogramas en el kriging, las desviaciones estándar del kriging (KSD) representan fracciones decimales de los valores estimados.

El **semivariograma** permite examinar la varianza de las diferencias entre los valores de una variable espacial, medida en diferentes estaciones, con un mismo lag de tiempo, en función de ese lag (Fig. 6).

Obdam (1983) define la relación entre el correlograma $r(h)$ y el semivariograma $\gamma(h)$ de la forma siguiente:

$$\gamma(h) = 1/2N(h) \left[\sum_i^{N(h)} F_{rj} - F(r_i - h) \right]^2$$

En la que,

$N(h)$ es el número de pares a una distancia relativa o un cierto lag h ; f , es la variable espacial; r , la i -ésima posición de la variable espacial medida. La esperanza matemática $E()$ de una variable cualquiera para $1/N \sum()$ puede plantearse como:

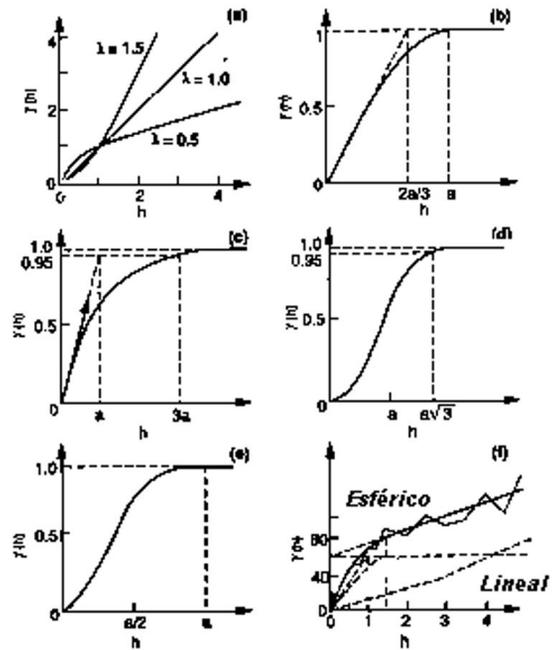


Figura 6. Modelos de Variograma (según Delhomme, 1976, de Marsily, 1986, Peck et al., 1988)

Figure 6. Variogram models (after Delhomme, 1976; de Marsily, 1986; Peck et al., 1988)

$$\gamma(h) = \frac{1}{2} E[F(r) - F(r-h)]^2$$

De manera que,

$$\gamma(h) = var(F) - cov ar[F(h)]$$

en la que $var(F)$ es la varianza de F y $cov ar[F(h)]$ la covarianza de F .

Del mismo modo, como la función de autocorrelación:

$$\rho(h) = \frac{cov arF(h)}{var(F)}$$

La expresión anterior equivale a:

$$\frac{\gamma(h)}{var(F)} = 1 - \rho(h)$$

En las ecuaciones anteriores, el aumento de h conduce a $F(h)$ a cero y $\gamma(h)$ tiende a $var(F)$, de ahí que, para esta última, la función de covarianza puede graficarse en función de lag h como en el semi variograma. Sin tendencia, la función de

covarianza es el complemento del semivariograma respecto a la varianza de F.

Asumiendo cuatro escalas de heterogeneidad, una para cada espacio constitutivo del curso, de manera que cada nivel de escala i- contenga N_i unidades del nivel i, y el número total de muestras sea $N_1 \times N_2 \times N_3 \times N_4$, puede realizarse el análisis de varianza de cualquier variable del campo físico en el karst, incluidas las componentes de varianza de las propiedades de cada espacio (matriz, poros, grietas y cavernas; Fig. 7). El test de Fisher (F) puede emplearse para comprobar la significación de las varianzas diferentes. La Tabla 3 muestra el esquema de cálculo correspondiente.

El test de significación de V_i sobre V_{i+1} , en caso de $V_i > V_{i+1}$, es:

$$\frac{V_i}{V_{i+1}} < F_{1-\alpha}[GL_i, GL_{i+1} + 1]$$

y, en el caso de $V_{i+1} > V_i$, el test es de V_{i+1} sobre V_i .

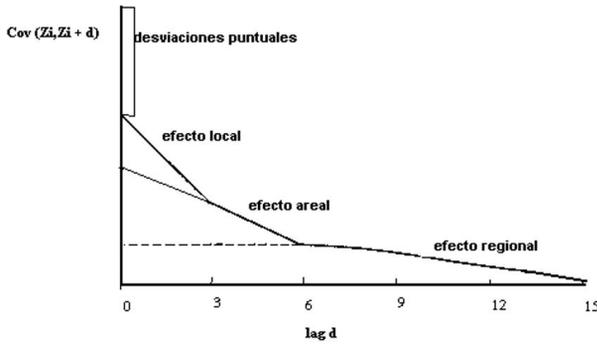


Figura 7. Función de covarianza con efecto de factor de escala (Molerio & Portuondo, 1997)

Figure 7. Covariance function with scale factor effect (Molerio & Portuondo, 1997)

Eficiencia de la red: Patrones constructivos y espaciales

Análisis Factorial (AF)

Para este análisis es necesario crear una matriz observacional compuesta por elementos que caractericen el sistema con el objetivo de lograr identificar y cuantificar similitudes entre las variables. En este caso, se aplican técnicas de Análisis Factorial (AF) y de Clasificación Numérica. Mediante el AF pretende encontrar, con un mínimo de pérdida de información, “factores”; esto es, combinaciones lineales de aquellas variables que resultan casi linealmente dependientes y están correlacionadas y cuyos efectos en la matriz observacional depende, por lo tanto, unos de otros (Spath, 1980). El Análisis Factorial tiene muchos puntos en común con el Análisis en Componentes Principales (ACP), y busca esencialmente nuevas variables o factores que expliquen los datos. En el ACP, en realidad, sólo se hacen transformaciones ortogonales de las variables originales,

Tabla 3. Análisis de Varianza. Test de significación de V_i respecto a V_{i+1} : $V_i/V_{i+1} < F_{1-\alpha}(DGF_i, DGF_{i+1})$ en el caso de $V_i > V_{i+1}$; para $V_i > V_{i+1}$ el test es de V_{i+1} respecto a V_i . La varianza total $\sigma_{tot}^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2$ (según Obdam, 1983)

Table 3. Variance analysis. Significance test of V_i respect V_{i+1} : $V_i/V_{i+1} < F_{1-\alpha}(DGF_i, DGF_{i+1})$ for the case $V_i > V_{i+1}$; for $V_i > V_{i+1}$ the test becomes V_{i+1} respect V_i . Total variance $\sigma_{tot}^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2$ (after Obdam, 1983)

Nivel de la subdivisión	Grados de libertad (DGF)	Suma de los cuadrados	Media de los cuadrados (estimados óptimos)	Componentes de la varianza estimados mediante la media de los cuadrados
1	$N_1 - 1$	$S_1 = N_2 N_3 N_4 \sum (x_i - \bar{x})^2$	$V_1 = S_1 / (N_1 - 1)$	$\sigma_1^2 + N_4 \sigma_3^2 + N_3 N_4 \sigma_2^2 + N_2 N_3 N_4 \sigma_1^2$
2	$N_1 (N_2 - 1)$	$S_2 = N_3 N_4 \sum_i \sum_j (x_{ij} - \bar{x}_i)^2$	$V_2 = S_2 / (N_1 (N_2 - 1))$	$\sigma_2^2 + N_4 \sigma_3^2 + N_3 N_4 \sigma_2^2$
3	$N_1 N_2 (N_3 - 1)$	$S_3 = N_4 \sum_i \sum_j (x_{ijk} - \bar{x}_{ij})^2$	$V_3 = S_3 / (N_1 N_2 (N_3 - 1))$	$\sigma_3^2 + N_4 \sigma_3^2$
4	$N_1 N_2 N_3 (N_4 - 1)$	$S_4 = N_4 \sum_i \sum_j \sum_k \sum_l (x_{ijkl} - \bar{x}_{ijk})^2$	$V_4 = S_4 / (N_1 N_2 N_3 (N_4 - 1))$	σ_4^2
	$N_1 N_2 N_3 N_4 - 1$	$S = \sum_l \sum_j \sum_k \sum_l (x_{ijkl} - \bar{x})^2$	$S / N_1 N_2 N_3 N_4 - 1$	

haciendo hincapié en la varianza de las nuevas variables. En el Análisis Factorial, por el contrario, interesa más explicar la estructura de las covarianzas entre las variables.

De este modo, un conjunto de variables puede ser reducida a un pequeño número de factores capaces de expresar las propiedades del conjunto pero bajo la presunción de la existencia de relaciones lineales entre ellos. Existen varios modos del Análisis Factorial (Tabla 4). El hecho de tomar un número adecuado de factores k para representar las covarianzas observadas es muy importante: entre una solución con k ó con $k + 1$ factores se pueden encontrar pesos factoriales muy diferentes, al contrario que en el método de componentes principales, donde los primeros k componentes son siempre iguales. Una ventaja del método de máxima verosimilitud es que lleva asociado un test estadístico para estimar el número de factores. Nosotros hemos aplicado con éxito marcar la varianza deseada en primera aproximación para la identificación de los factores.

Para lograr mayor claridad en la identificación de los patrones de las cargas factoriales; es decir, aquellos factores marcados por los valores más altos en algunos casos y bajos en otros en la estructura simple. Las estrategias de rotación son de varios tipos según el caso: Varimax, Biquartimax, Quartimax y Equamax. La Rotación Varimax fue propuesta por Kaiser (1958), y trata de que los factores tengan unas pocas saturaciones altas y muchas casi nulas en las variables y es la que usamos preferentemente. Esto hace que haya factores con correlaciones altas con un número pequeño de variables y correlaciones nulas en el resto, quedando así redistribuida la varianza de los factores. En la Rotación Quartimax se trata que una variable dada esté muy correlacionada con un factor y muy poco correlacionada con el resto de factores.

Las propiedades que debe cumplir la rotación son las siguientes:

- Cada fila de la matriz factorial de pesos debe contener, al menos, un cero.
- Cada columna de la matriz factorial de pesos debe contener, al menos, k ceros.
- Cada par de columnas de la matriz factorial de pesos debe contener varias variables cuyos pesos sean nulos en una columna pero no en la otra.
- Si hay más de cuatro factores cada par de columnas de la matriz factorial de pesos debe contener un número elevado de variables con pesos nulos en ambas columnas.

- De manera recíproca, si hay más de cuatro factores, en cada par de columnas de la matriz factorial de pesos sólo un número pequeño de variables debe contener pesos no nulos.

Las características del sistema que más comúnmente no deben faltar en la conformación de las matrices son:

- Espesor de acuífero cortado.
- Cota piezométrica media.
- Profundidad del pozo.
- Distancia a la costa.
- Distancia al acueducto/captaciones más cercanas.
- Distancia a un sistema hidráulico/hidrológico importante (presa, río).
- Rumbo a un sistema hidráulico/hidrológico importante.

En un caso de intrusión marina (Molerio y Portuondo, 1997) incluyeron en la matriz observacional las siguientes variables, cuya distribución en el plano factorial 1,2 se presenta en la Fig. 8.

- D_COSTA: Distancia a la costa (km)
- PROF: Profundidad total del pozo (m)
- CP: Cota absoluta del nivel piezométrico (m)
- CP_G_L: Altura absoluta de la posición de las aguas con 1 g/L de mineralización total CE: Conductividad Eléctrica Específica del agua ($\mu\text{S}/\text{cm}$)
- CL: Concentración de iones Cl en el agua subterránea (mg/l)
- HCO3: Concentración de iones HCO3 en las aguas subterráneas (mg/l)
- LL_1_2: Lluvia media anual (mm)
- DIST_AC: Distancia al campo de pozos más cercano (km)
- QAC: Caudal de bombeo de los pozos del campo más cercano (m^3/s)
- CPAC: Cota piezométrica de las aguas subterráneas en el campo de pozos mas cercano (m)
- EAD: Espesor de agua dulce en el pozo de observación (m)

Las matrices deben ser en **modo R**, donde las variables están en las columnas y los objetos (pozos) en las filas y

Tabla 4. Modos del Análisis Factorial (Molerio, 1998)

Table 4. Modes of Factor Analysis (Molerio, 1998)

RELACIÓN	OBJETO	VARIABLES	OCASIONES	MODOS	EJEMPLO
Objeto/Variable	Todos	Todas	Una	R - Variable O - Objeto	$\text{HCO}_3, \text{SO}_4, \text{Cl} \dots n \text{ T}_1, \text{T}_2, \text{T}_3, \dots \text{T}_n$
Objeto/Ocasión	Todos	Una	Todas	T - Ocasiones S - Objetos	$\text{Cl}_1, \text{Cl}_2, \text{Cl}_3, \dots \text{Cl}_n \text{ T}_1, \text{T}_2, \text{T}_3, \dots \text{T}_n$
Variable/Ocasión	Uno	Todas	Todas	O - Ocasiones P - Variables	$\text{HCO}_3, \text{SO}_4, \text{Cl} \dots n \text{ T}_{4(1)}, \text{T}_{4(2)}, \text{T}_{4(n)}$

modo Q, donde los objetos se encuentran por columnas y las variables en las filas, aunque este caso no siempre es aconsejable y no necesariamente representa una semejanza o sustituto del análisis de clúster. La matriz en modo R tiene como objetivo establecer similitudes entre las variables con el objetivo para el diseño del sistema. Por su parte, la matriz en modo Q tiene como objetivo establecer las similitudes entre los objetos, es decir, obtener grupos de pozos con características similares para poder seleccionar los representativos entre ellos

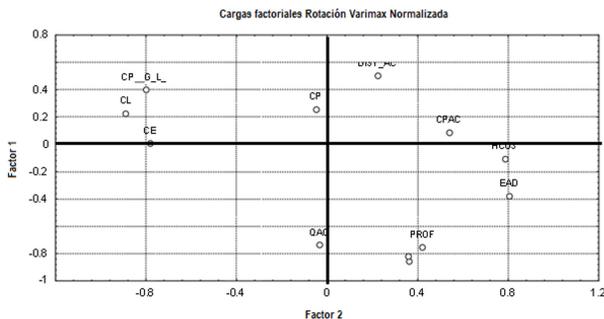


Figura 8. Distribución de factores en el plano factorial euclidiano (Yera y Molerio, 1998)

Figure 8. Factor distribution on the Euclidean Factorial Plane (Yera y Molerio, 1998)

Las variables decisivas para el diseño de la red son las siguientes

- Concentración de HCO₃
- Conductividad Eléctrica Específica de las aguas
- Distancia a la costa
- Concentración de cloruros en las aguas
- Altura absoluta de las aguas con una mineralización de 1 g/L
- Espesor de agua dulce
- Lluvia media anual
- Profundidad del pozo

Clasificación numérica

Esta es una técnica exploratoria que también reduce significativamente la subjetividad del observador y permite clasificar objetos de acuerdo con procesos interdependientes descritos por conjuntos de valores de diferentes variables. Según Spath (1980) el rango confiable de aplicación considera objetos m=50 y hasta M=5000 y variables entre l=2 a l=10. El método es especialmente apropiado cuando deben tratarse más de dos variables simultáneamente tanto por métodos de segmentación heurística como jerárquicos y los problemas básicos que se presentan son: a) la selección de la medida de distancia; b) la selección del algoritmo; c) el

número de clústeres y d) la selección de las variables. Al trabajar con datos mezclados se debe evitar utilizar el uso de valores promedio tan pronto aparezcan los valores nominales u ordinales.

Las medidas usuales de distancia para cuantificar la proximidad de los objetos son la Euclideana; a saber,

$$d_M(C_p, C_q) = \left| \bar{x}_p^T - \bar{x}_q^T \right|$$

$$d_M''(C_p, C_q) = \sqrt{\frac{m_p m_q}{m_p + m_q}} \left| \bar{x}_p^T - \bar{x}_q^T \right|$$

Y la de Mahalanobis generalizada,

$$d_S(C_p, C_q) = \sqrt{(\bar{x}_{qp} - \bar{x}_q) S^{-1} (\bar{x}_{qp} - \bar{x}_q)^T}$$

Donde,

$$S = \frac{1}{m} \bar{X}^T \bar{X}$$

d_M , es la distancia Euclideana a los centroides del cluster, d_M'' , es lo mismo pero multiplicado por una función del número de elementos y d_S la Distancia Generalizada de Mahalanobis.

Para el caso anterior de Molerio y Portuondo, el agrupamiento de los casos semejantes se presenta en el dendograma de la Fig. 9.

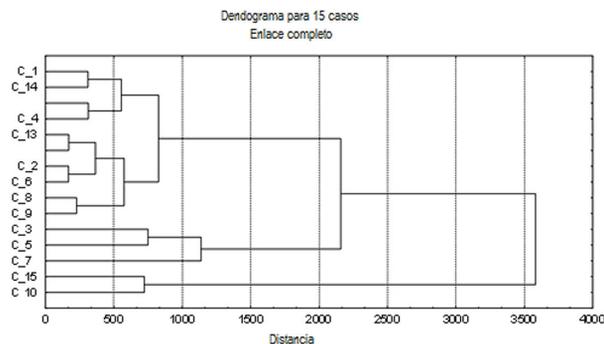


Figura 9. Dendograma de enlace completo para un modelo de ajuste de red para la vigilancia de la intrusión marina (Yera y Molerio, 1998)

Figure 9. Linkage dendrogram of an adjusted model for a sea water intrusion monitoring network (Yera y Molerio, 1998)

En el caso, anterior, a los efectos de diseño se obtuvieron las conclusiones de la Tabla 5.

Nota final

Las dos ventajas principales del procedimiento aquí resumido son: a) la reducción de la subjetividad del observador en la identificación de las leyes o reglas que rigen la hidrodinámica de los procesos y caracterizan las componentes, contenidos, distribución espacial y frecuencia de la red y de sus observaciones; b) la posibilidad de explicar y cuantificar la varianza y, por tanto, la certidumbre de los factores de control de los procesos. También los análisis individuales, por sí mismo contribuyen mucho a precisar

Tabla 5. Resultados de la aplicación conjunta de técnicas de Análisis Factorial y Clasificación Numérica en un caso de optimización de una red de monitoreo de la intrusión marina (Molerio y Portundo, 1997)

Table 5. Results of the joint application of Factor and Clustering Analysis in a case of optimization of a sea water intrusion network (Molerio y Portundo, 1997)

Variable	Indicadores de diseño óptimo
Distancia a la costa	Entre 5 y 15 km. Colocados a mayores o menores distancias no ofrecen información significativa. A menos de 5 km las concentraciones de cloruros y la mineralización son prácticamente constantes y a más de 15 km no tiene sentido práctico
Profundidad de los pozos	Óptimo 50 m (entre 40 y 100 m)
Espesor de agua dulce	Los espesores máximos de agua dulce respecto a la distancia a la costa son: Entre 5 - 9 km de la costa: 10 m 10-15 km: 50 m >15 km: 60 - 120 m
Mineralización de las aguas	1400 μ S/cm como indicador de presencia de 250 mg/L de cloruros en las aguas
Frecuencia de monitoreo	Anual. No hay ganancia notable de información en intervalos más cortos. La señal de entrada de cloruros se filtra completamente a bajas frecuencias, cercanas a los 12 meses lo que confirma que no hay ganancia de información con frecuencias de muestreo menores de un año.

aspectos cruciales relacionados directamente con la operación de la red y los costos de su gestión, como la búsqueda de la frecuencia óptima de monitoreo o el contenido de muestreo.

La flexibilidad del procedimiento descrito para reajustar redes preexistentes por cambio de objetivos o introducción de otros nuevos y de restricciones como la distancia, la “simultaneidad” que eventualmente se requiere para caracterizar estados del sistema o la medida del cambio de las reservas de agua subterránea son ganancias de informatividad y de eficiencia del sistema así diseñado para también evaluar casos específicos de detección de tendencias, violaciones en la calidad de las aguas, de la omisión de puntos o muestreos e incluso de otro tipo de dato faltante. En ciertos casos la optimización e ha acompañado de técnicas complementarias como el Observador de Luenberger, un estimador de estado que proporciona información del estado interno del sistema a partir de mediciones de entradas y salidas en el sistema real (Luenberger, 1966) o la tunelización (Gómez et al., 2001, 2003) para la caracterización de valores mínimos sin que se pierda el rigor del procesamiento de la data y su inmediata conversión en información.

Reconocimientos

Enriquecedoras experiencias en el diseño, operación, procesamiento y generalización de la data y optimización de redes y en el intercambio de experiencias debo, históricamente, a numerosos colegas por lo que expreso mi agradecimiento a Humberto Álvarez, Dionisio Amor, Antonio Armesto, Osvaldo Barros, Juan Bobes Consuegra, Yamilé Bustamante, Luz Graciela Caldevilla, Diana de Centanaro, Antonio Eduardo Cordeiro Neto, Constantino de Miguel, Andrés Díaz Arenas, Martín Díaz, Willie Doemo, Pedro Luis Dorticós del Río, Ernesto Flores, Mijail Galabov, Jorge L. Gelabert, Susana Gómez; Arturo González Báez, Mario Guerra, Svetanka Guneva, Rebeca Hernández, Ismael Herrera, Flor Jiménez, Lászlo Kiraly, Agnia León, Enriqueito López, Didier Louvat, Alain Mangin, Cecilia March,

Ambar Menéndez, John Neckless, Jesús Piñera, Eduardo Planos, Ibrahim Plaza, Yoemí Portuondo, Alberto E. Rivero, Ernesto Rocamora, José Francisco Santiago, Konstantin Spassov, Aftandil Tzirekitze, Ernesto Varela, Jaroslav Vrba, Peter Yankov y Gladys Yera. Muy especialmente a Ana, mi compañera.

Bibliografía

- Agterberg, F.P. (1974): *Geomathematics*. Elsevier, 595:
 Altovsky, M.E. (1959). Concept of the term “Regime of Underground Waters”. in Altovsky, M. y A. Konoplyantsev (1959). *Manual for the systematic study of the regime of underground waters*. Foreign Languages Edit., Moscú:14-16
 Bear, J., Y. Bachmat (1990). *Introduction to Modeling of Transport Phenomena in Porous Media*. Kluwer Academic Publishers, The Netherlands, 553:
 Bear, J., A. H.-D. Cheng (2010). *Modeling Groundwater Flow and Contaminant Transport*. Jacob Bear [Ed]: Theory and Applications of Transport in Porous Media. Volume 23. Springer Dordrecht, Heidelberg, London, New York, 834:
 Bear, J., A. H.-D. Cheng, S. Sorek, D. Ouazar, I. Herrera (1999). *Seawater Intrusion in Coastal Aquifers. Concepts, Methods and Practices*. Springer Science + Business Media Dordrecht. Theory and Applications of Transport in Porous Media, Vol 14, 625:
 Bear, J., M. Yavuz Corapcioglu (1984). *Fundamentals of Transport Phenomena in Porous Media*. Jacob Bear [Ed]. Proceedings of the NATO Advanced Study Institute on Mechanics of Fluids in Porous Media, Newark, Delaware, USA, July 18-27, 1982. Martinus Nijhoff Publishers, Dordrecht-Boston-Lancaster, 1003:
 Bear, J., D. Zaslavsky, S. Irmay (1968). *Physical Principles of Water Percolation and Seepage*. Unesco, Ariz Zone Research, XXIX, Paris, 465:

- Bochever, F.M. (1968). *Métodos hidrogeológicos teóricos y prácticos para el cálculo de las reservas de explotación de las aguas subterráneas*. Edit. Nedra, Moscú, 322:
- Borevskii, B.; B. Samsonov & L. Yazvin (1979). *Metódica para la Determinación de los Parámetros de los Acuíferos por Datos de Aforos*. Edit. Nedra, Moscú, 328: (en ruso)
- Brower G. K., P. R. Defize (1983). Spatial Uncertainty in Piezometric Head. *Internatl. Symp. MIIGS*, The Netherlands, Unesco, TNO:118-128
- Buttkus, Burkhard (2000). *Spectral analysis and Filter Theory in Applied Geophysics*. Springer, 667:
- Delhomme, J. P. (1978). Kriging in the Hydrosciences. *Adv. in Water Resourc.* 1(5):251-261
- De Miguel Fernández, Constantino (2012). *Hidrogeología Aplicada con Aspectos Ambientales*. Editorial Digital Universitaria Moa, 374:
- Díaz Viera, M.A. (2002). *Una metodología alternativa para el diseño de una red de monitoreo nueva en aguas subterráneas para fines de manejo con casi ninguna información*. UNAM, IIMAS. 8:
- Galabov, Mijail (1983). *Dinámica de las aguas subterráneas*. Edit. Técnica, Sofía, 319: (en búlgaro)
- Galabov, Mijail (1985). *Hidrodinámica de las aguas subterráneas y evaluación de recursos*. Edit. Técnica, Sofía, 319: (en búlgaro)
- Gómez, Susana, Nelson del Castillo, Longina Castellanos y Julio Solano (2003): The parallel tunneling method. *Parallel Computing* 29 (2003). 523-53
- Gomez, S., J. Solorzano, L. Castellanos, M.I. Quintana (2001). *Tunneling and genetic algorithms for global optimization*, in: N. Hadjisavvas, P. Pardalos (Eds.), *Advances in Convex analysis and Global Optimization. Non Convex Optimization and its Applications*, Kluwer Academic Publishers: 553-567
- Journel, A. & C. Huijbregts (1978). *Mining Geostatistics*. Acad. Press, New York, 600:
- Kiraly, Lászlo (1975). Rapport sur l'état actuel des connaissances dans le domaine des caracteres physiques des roches karstiques. In/ Burger, A. y L. Dubertret: *Hydrogeology of karsic terraines*. *Internatl. Ass. Hydrogeol.*, Paris:53-67
- Kiraly, Lászlo (1978). La notion d'unité hydrogeologique. Essai de definition. Thèse. Bull Centr. Hydrogeokl. Univ. Neuchatel:82-216
- Kliméntov, P.P., V.M. Kónonov (1983). *Metodología de las investigaciones hidrogeológicas*. Edit. Mir, Moscú, 447:
- Luenberger, D.E. (1966). Observers for multivariable systems. *IEEE Trans on automatic control* AC-11(2):190-197
- Mangin, A. (1975). *Contribution a l'étude hydrodynamique des aquifers karstiques*. Thèse. Univ.- Dijon., 124:
- Mangin, A. (1981). Utilisation des analyses corrélatrice et spectrale dans l'approche des systèmes hydrologiques. *C.R. Acad. Sci. Paris* 293:401-404
- Mangin, A. (1984a). Pour un meilleure connaissance des systems hydrologiques a partir des analyses corrélatrice et spectral. *Jour. Hydrol.* 67:25-43
- Mangin, A. (1984b). incidence du choix du pas de discrétisation d'un processus sur sa simulation a partir de modeles. *Hydrogeo.-Geol de l'Ingénieur.* 3:317-321
- Matheron, G. (1965). *Les Variables Regionalisees et leur Estimation*. Masson ed., Paris, 300:
- Matheron, G. (1970). La theorie des variables regionalisees et ses applications. *Cah. Cent. Morph. Math.* (5)
- Mijailov, L. (1989). *Hidrogeología*. Edit. Mir, Moscú, 286:
- Molerio León, Leslie F. (1984). El Efecto del Factor de Escala en la Interpretación del Campo Físico de las rocas carsificadas. *Resumen. XXVII Internatl. Geol. Congr.*, Abril 4-14, 1984, Moscú, Vol VII, Secc. 16,:468-469. Véase, *Voluntad Hidráulica* 75:19-26
- Molerio León, Leslie F. (1985). El Área Elemental Representativa (AER) para la Evaluación de las Propiedades Físicas del Carso. *Modelo Teórico. Simp. XLV Aniv.Soc. Espel. Cuba*, La Habana :45; In/Núñez Jiménez, A. (1990). Medio siglo explorando a Cuba. Historia documentada de la Sociedad Espeleológica de Cuba. Tomo II, Imprenta Central de las FAR, La Habana, :323
- Molerio León, Leslie F. (1992). Técnicas Estadísticas y de la Teoría de la Información en el Diseño y control de las redes de Monitoreo de las Aguas Subterráneas en el Carso. *1er Taller Iberoamericano de Informática y Geociencias, Acad. Cienc. Cuba*, La Habana, : 32
- Molerio León, Leslie F.(1998). Redes de monitoreo de las aguas subterráneas. *Notas de clase. Inst. Superior de Ciencias y Tecnología Nuclear*, La Habana, 132: <https://www.researchgate.net/publication/380876084>
- Molerio-León, L.F. (2022). Cloruros como indicadores de la dinámica de las aguas subterráneas en el karst de La Cachurra (Matanzas, Cuba): Problemas Teóricos y Aplicaciones. *Argentina subterránea* 22(52): octubre: 67-112 <https://www.researchgate.net/publication/365534489>
- Molerio León, L.F. y Ana M. Sardiñas Gómez (2024). Fuentes de incertidumbre en la evaluación de los recursos de agua subterránea en el karst cubano. *Revista Maya de Geociencias*. Suplemento abril 2024:33-54. Universidad Autónoma de Baja California Sur. Depto Académico de Ciencias de la Tierra. <https://www.researchgate.net/publication/380375516>
- Molerio León, L.F. (2013). Modelos genéricos de diseño de redes de monitoreo hidrogeológico en yacimientos de lateritas ferroniquelíferas: aplicación a las saprolitas y serpentinitas blandas del horizonte inferior de las cortezas

- lateríticas de Moa, Cuba Oriental. Resumen. *Cub@: Medio Ambiente y Desarrollo*. 13, 24, 2012
- Molerio León, L.F. & Y. Portuondo López (1997). *Design of Sea Water Intrusion Early Warning Systems in Coastal Karstic Aquifers: A Multivariate Approach*. Draft Report. UNESCO IHP-V Project 3.2. Monitoring Strategies for Detecting Groundwater Quality Problems, CONAPH, La Habana, 27:
- Molerio León, L.F.; Y. Portuondo López y J.C. Torres Rodríguez (1998). Métodos Geomatemáticos de Diseño y Optimización de Redes de Monitoreo de Aguas Subterráneas. *XII Seminario Internacional del CIRA (Centro Interamericano de Recursos del Agua) y III Taller Internacional sobre Gestión y Tecnologías de Suministro de Agua Potable y Saneamiento Ambiental*. La Habana, 17-20 Noviembre, 1988, 10:
- Molerio León, L.F.; R. Toujague de la Rosa (2004). Arsenic in a hard rock aquifer: multivariate optimisation of the groundwater monitoring network. 32nd Internatl. *Geol. Congr.* <http://www.alhsud.com>, 9:
- Molerio León, L.F., J.C. Torres Rodríguez (2002): Métodos Geomatemáticos de diseño y optimización de redes de monitoreo de aguas subterráneas. *Ing. Hidr. y Ambiental*, La Habana, XXIII (2),38-45
- Molerio León, L.F., J.C. Torres Rodríguez, E., Rocamora Álvarez, M.G. Guerra Oliva, K. del Rosario (2002a). Optimización de la red de monitoreo del régimen de las aguas subterráneas en la Cuenca Sur de Matanzas, Cuba. I. Dominio espacial. *Ing. Hidr. y Ambiental*, La Habana, XXIII (2):10-22
- Molerio León, L.F., J.C. Torres Rodríguez, E., Rocamora Álvarez, M.G. Guerra Oliva, K. del Rosario (2002b). Optimización de la red de monitoreo del régimen de las aguas subterráneas en la Cuenca Sur de Matanzas, Cuba. II. Dominio frecuencial. *Ing. Hidr. y Ambiental*, La Habana, XXIII (2):23-37
- Obdam, A. N. M. (1983). Statistical Treatment of Sedimentary Structures in the Underground and its use for Geohydrological Sampling.. *Internatl. Symp. MIIGS*, The Netherlands, Unesco, TNO:55-72
- Peck, A. et al. (1988). Consequences of Spatial Variability in Aquifer Properties and Data Limitations for Groundwater Modelling Practice. *IAHS Publ.* 175, Oxfordshire 272:
- Piñera, J., L.F. Molerio, C. March (1982). Análisis de un modelo estadístico para la regionalización de las transmisividades en un poje. Col. Internac. *Hidrología del Carso de la Región del Caribe*, UNESCO, La Habana: 302-321
- Piñera, J., L.F. Molerio, C. March (1984). Carácter probabilístico de las propiedades hidrogeológicas de las rocas acuíferas. *Ing. Hidráulica* (3), 344-345
- Portuondo López, Y.; C. March Delgado & M. G. Guerra Oliva (1997). *Optimización de la Red de Monitoreo de Niveles de Agua Subterránea de la Cuenca de Vento*, Ciudad de La Habana. Ofic. Nac. Hidrogeología (CENHICA, Inst Nac. Rec. Hidr. 31:
- Pshenichny, C., V. Dech, (2015). *Mathematical Geology and Knowledge Engineering: View From St. Petersburg*. Russian Academy of Sciences, St. Petersburg, Russian Federation, 4:
- Quintana Garmendía, A., B. Ballagas Flores, L.F. Molerio León, J.C. Torres Rodríguez, E. Rocamora Alvarez, M.G. Guerra Oliva (1999). Estructura autocorrelatoria y espectral de los niveles piezométricos de la cuenca Ariguanabo. *VIII Jor. Cient. Inst. Geof.Astron.*, Dic. 20-21, La Habana, :35
- Reyment, R.A.& K.G. Joreskog (1993). *Applied Factor Analysis in the Natural Sciences*. Cambridge Univ. Press 371:
- Schilperoort, T., S. Groot (1983). Design and Optimization of Water Quality Monitoring Networks. *Internatl. Symp. MIIGS*, The Netherlands, Unesco, TNO:86-100
- Späth, Helmuth (1980). *Cluster algorithms for data reduction and classification of objects*. Ellis Horwood, Ltd, John Wley & sons, 226:
- Spiridonov, V.P., y A.A. Lopatkin (1973). *Tratamiento matemático de datos fisico-químicos*. Edit. Mir, Moscú, 207:
- Toujague de la Rosa, R. y L.F. Molerio León (2013): Análisis multivariado de la contaminación por arsénico en las aguas subterráneas de mina Delita, Cuba. *Cub@: Medio Ambiente y Desarrollo; Revista electrónica de la Agencia de Medio Ambiente*. 13, (24).
- Yera Digat, G. Y L.F. Molerio León (1998). *Combinaciones Lineales Estandarizadas (SLC) entre los Factores de Control de Acuosidad de las Rocas de Baja Permeabilidad*. XII Seminario Internacional del CIRA (Centro Interamericano de Recursos del Agua) y III Taller Internacional sobre Gestión y Tecnologías de Suministro de Agua Potable y Saneamiento Ambiental, La Habana, 17-20 Noviembre, 1988, 9: