



UNIVERSIDAD NACIONAL AUTÓNOMA DE

05061

1  
?ej-

Unidad Académica de los Ciclos Prof  
y de Posgrado del CCH

Instituto de Investigaciones en  
Matemáticas Aplicadas y en Sistemas

CONSTRUCCION DE UNA RED DE  
MONITOREO OCEANOGRÁFICA  
UTILIZANDO LA TEORÍA DE VARIABLES  
REGIONALIZADAS

TESIS CON  
FALLA LA ORIGEN

E S I S

QUE PARA OBTENER EL GRADO DE:  
MAESTRO EN ESTADÍSTICA E  
INVESTIGACIÓN DE OPERACIONES

P R E S E N T A :

CARLOS DÍAZ AVALOS



Universidad Nacional  
Autónoma de México



## **UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso**

### **DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## INDICE

<b>INTRODUCCION</b>	<b>1</b>
<b>CAPITULO 1. FUNDAMENTOS TEORICOS</b>	<b>3</b>
1.1. MOMENTOS Y ESTACIONARIEDAD	4
1.2. KRIGING	10
1.3. CAMBIO DE SOPORTE	15
1.4. ANALISIS ESTRUCTURAL	16
<b>CAPITULO 2. DISEÑO DE REDES DE MUESTREO</b>	<b>19</b>
2.1. EL ENFOQUE CLASICO	19
2.2. EL ENFOQUE GEOESTADISTICO	20
<b>CAPITULO 3. METODOLOGIA</b>	<b>29</b>
3.1. ANALISIS ESTRUCTURAL	29
3.2. DISEÑO DE LA RED DE MUESTREO	31
<b>CAPITULO 4. RESULTADOS</b>	<b>34</b>
<b>CONCLUSIONES</b>	<b>43</b>
<b>LITERATURA CITADA</b>	<b>45</b>

## INTRODUCCION.

La evaluación para el aprovechamiento de los recursos pesqueros es una labor importante y compleja, sobre todo en el caso de las especies pelágicas consideradas como potenciales, pues su estudio se hace con el método de conteos larvales, que aunque adecuado para este propósito, resulta costoso pues se deben utilizar embarcaciones cuyo costo por día supera los 15 millones de pesos. Por ello no es necesario exagerar la importancia de establecer un patrón de muestreo de tal forma que proporcione la mayor cantidad de información sobre la o las variables de interés.

En el presente trabajo se presenta una aplicación de la teoría de variables regionalizadas al diseño de una red de muestreo oceanográfica, la cual tiene como propósito la estimación del total de larvas de tres especies de atún consideradas como recursos de amplio potencial económico.

En el primer capítulo se presentan los fundamentos de la teoría de variables regionalizadas así como los supuestos teóricos necesarios para hacer inferencias a partir de una sola realización de una función aleatoria y se describe brevemente el desarrollo del método de estimación conocido como "kriging", que es pieza fundamental en los métodos geoestadísticos.

En el segundo capítulo se revisan los métodos propuestos en la literatura para el diseño de redes de muestreo basados en la teoría de variables regionalizadas y se discuten sus posibilidades de aplicación al problema de muestreo oceanográfico, mientras que en el tercer capítulo se describe y justifica la metodología empleada en el diseño de la red de muestreo para las especies de interés.

Finalmente, en el cuarto capítulo se describen y discuten los resultados obtenidos al aplicar algunos de los métodos propuestos para el diseño de redes, haciéndose una comparación detallada entre las disposiciones regulares triangular, cuadrada y hexagonal, y se evalúa el impacto de las redes propuestas en este trabajo sobre el número de días de crucero y sobre la varianza del error de estimación o varianza de kriging.

## CAPITULO 1

### FUNDAMENTOS TEORICOS.

Las variables que caracterizan un fenómeno espacial usualmente tienen un comportamiento demasiado complejo como para poder ser descritas por los métodos matemáticos estándar, ya que su variación espacial es errática y muchas veces impredecible de un punto a otro. Sin embargo, su comportamiento no es totalmente aleatorio, ya que los valores observados en puntos cercanos tienden a ser mas parecidos que los valores de puntos alejados, debido a la existencia de un conjunto de correlaciones que reflejan la estructura del fenómeno bajo estudio.

El término variables regionalizadas fué escogido por Matheron en 1965 para enfatizar las propiedades particulares de estas, ya que caracterizan fenómenos que se distribuyen en el espacio y poseen una cierta estructura, llamada fenómeno regionalizado o regionalización. Ejemplos de este tipo de variables son la cantidad de un contaminante en la atmósfera, la abundancia de una especie, el pH del suelo, etc.. La teoría desarrollada por Matheron intenta proporcionar una herramienta objetiva para reconstruir una variable distribuida espacialmente a partir de un conjunto de observaciones, así como para obtener una medida de la precisión del valor predicho (Gambolati y Volpi, 1979), esto es, la teoría de variables regionalizadas nos da un enfoque coherente a los problemas de análisis estructural y de predicción de la variable de interés. Desde 1957, esta teoría ha sido aplicada extensivamente a la estimación de reservas minerales bajo el nombre de Geoestadística, extendiéndose su uso recientemente a áreas como la ecología, oceanografía, hidrología y meteorología, entre otras.

La teoría de variables regionalizadas esta basada en la teoría de funciones aleatorias. Una variable regionalizada  $z$  puede considerarse entonces como una realización particular de una función aleatoria  $Z$ , esto es,  $z$  es una realización particular de

Z. Esta función puede idealizarse como el resultado de dos contribuciones básicas:

- a) Una componente  $m$  de naturaleza determinística, que representa la tendencia global de  $Z$ , y
- b) Una componente estocástica  $e$  que representa la dispersión natural alrededor de  $m$ , con media cero, varianza constante y una alta autocorrelación (Hughes y Lettenmaier, 1981).

Similarmente, podemos considerar el valor  $z(x)$  tomado por la variable regionalizada en el punto  $x$  como la realización de una variable aleatoria  $Z(x)$ , donde  $x$  es un punto en  $\mathbb{R}^2$  o en  $\mathbb{R}^3$  por lo general. Si consideramos todos los valores  $z(x)$  tomados por  $Z$  en todos los puntos  $x$ , la variable regionalizada puede interpretarse como una realización particular de un conjunto infinito de variables aleatorias  $Z(x)$ . Este conjunto es llamado función aleatoria, ya que asocia una variable aleatoria a cada punto  $x$ .

Puesto que solo tenemos una realización, la naturaleza de la función aleatoria debe ser especificada mediante suposiciones teóricas adicionales, ya que la interpretación probabilística de la variable aleatoria  $Z(x)$  como una realización particular de una función aleatoria  $Z$  solo tiene sentido operativo si es posible inferir toda o parte de la ley de probabilidad que define a esta última. Estas suposiciones teóricas involucran un cierto grado de homogeneidad espacial y son conocidas como hipótesis de estacionariedad, concepto que a continuación se define.

### 1.1. Momentos y estacionariedad

Consideremos la función aleatoria  $Z(x)$ . Para cada conjunto de  $k$  puntos en  $\mathbb{R}^n$  ( $n=1, 2$  ó  $3$  por lo general),  $x_1, \dots, x_k$ , llamados puntos del soporte, corresponde un vector aleatorio  $\{Z(x_1), \dots, Z(x_k)\}$  caracterizado por su función de distribución de

probabilidad

$$F_{x_1, \dots, x_k}(z_1, \dots, z_k) = \text{Prob}\{Z(x_1) \leq z_1, \dots, Z(x_k) \leq z_k\} \quad (1)$$

El conjunto de todas estas funciones de distribución, para todo entero  $k$  y para cualquier elección posible de  $x_1, \dots, x_k$  en  $\mathbb{R}^n$  constituye lo que Journel y Huijbregts (1978) llaman la "ley espacial" de la función aleatoria  $Z$ .

Dado que la cantidad disponible de datos es insuficiente para inferir esta ley espacial, es necesario suponer cierto grado de homogeneidad espacial de la función aleatoria. Así, definiremos a una función aleatoria como *estrictamente estacionaria* si su ley espacial es invariante ante translaciones en el espacio; esto quiere decir que si tenemos dos vectores aleatorios  $\{Z(x_1), \dots, Z(x_k)\}$  y  $\{Z(x_1+h), \dots, Z(x_k+h)\}$ , se cumple que

$$F[z(x_1), \dots, z(x_k)] = F[z(x_1+h), \dots, z(x_k+h)] \quad (2)$$

para cualquier distancia  $h$ .

El cumplimiento de esta condición es difícil de probar, siendo entonces necesario relajarla y solo suponer que la estacionariedad se da hasta el segundo momento. Entonces, una función aleatoria se dice que es estacionaria de segundo orden si:

- 1) El valor esperado de  $Z(x)$  existe y no depende de la posición  $x$ , esto es,

$$E\{Z(x)\} = m \quad \forall x; \quad (3)$$

- 2) Para cada par de variables aleatorias  $Z(x)$  y  $Z(x+h)$  la covarianza existe y solo depende de la separación  $h$  entre ellas,



$$\text{Cov}\{Z(x), Z(x+h)\} = C(h) \quad \forall x \quad (4)$$

donde  $h$  representa un vector en  $\mathbb{R}^n$ ,  $n=1, 2 \text{ ó } 3$ .

Con esta última condición es posible definir a la función de autocorrelación o correlograma como

$$\rho(h) = C(h)/C(0) \quad (5)$$

La hipótesis de estacionariedad de segundo orden asume la existencia de una varianza a priori finita,  $C(0) < \infty$ . En la práctica sin embargo, tanto la estacionariedad estricta como la estacionariedad de segundo orden son muy restrictivas y entonces se utiliza la llamada hipótesis intrínseca. Con ella, en vez de  $Z(x)$  solo son sus incrementos  $[Z(x) - Z(x+h)]$  los que se suponen estacionarios de segundo orden, y el segundo momento de estos define a la función variograma

$$\gamma(h) = (1/2)E\{[Z(x) - Z(x+h)]^2\} \quad (6)$$

Bajo esta hipótesis,  $\gamma(h)$  no depende de la posición sino únicamente de la distancia  $h$ . A las funciones aleatorias que satisfacen ésta hipótesis se les conoce como funciones intrínsecas; una función estacionaria de segundo orden también satisface la hipótesis intrínseca y su variograma satisface la relación

$$\gamma(h) = C(0) - C(h) \quad (7)$$

Sin embargo, una función intrínseca no necesariamente es estacionaria de segundo orden y por lo tanto su covarianza puede no existir. Ambas funciones son simétricas en  $h$ , esto es

$$C(h) = C(-h) ; \quad \gamma(h) = \gamma(-h)$$

La covarianza además cumple la desigualdad de Schwarz

$$|C(h)| \leq C(0)$$

Cuando la variable se distribuye en  $R^2$  o en  $R^3$ , tanto el variograma como la covarianza pueden cambiar con la dirección  $\alpha$ , definiéndose entonces al variograma como función del vector  $h'=[h, \alpha]$ , donde  $h$  es la distancia y  $\alpha$  la dirección, esto es

$$\gamma(h) = \gamma(h, \alpha) \quad ; \quad C(h) = C(h, \alpha)$$

y en este caso se dice que el fenómeno presenta un comportamiento anisotrópico. El variograma indica como aumenta la variabilidad y/o como disminuye la correlación con la distancia. Conociendo el variograma en diferentes direcciones es posible interpretar los rasgos direccionales del fenómeno; de aquí que el variograma sea un potente revelador de la estructura de la variable regionalizada.

En la práctica, la función estructural, ya sea covarianza o variograma solo se aplica a distancias limitadas  $h \leq b$ . Este límite  $b$  representa el radio de una vecindad en la que podemos considerar que la hipótesis de estacionariedad de segundo orden o la intrínseca se cumplen, y que dos variables  $Z(x)$  y  $Z(x+h)$  no se consideran como provenientes de la misma función aleatoria si  $h > b$ . Esta limitación de las hipótesis se conoce como hipótesis de cuasi-estacionariedad.

La cuasi-estacionariedad es en realidad un compromiso entre la escala de homogeneidad del fenómeno y la cantidad de datos disponibles, ya que siempre es posible producir estacionariedad reduciendo  $b$ , pero al hacer esto muchas de las vecindades no tendrán suficientes puntos en ellas como para estimar  $C(h)$  o  $\gamma(h)$  (Journel & Huijbregts, 1978).

Las aplicaciones de la geoestadística son básicamente la estimación de la variable de interés en un punto no muestreado o en un bloque, a partir de información  $Z(x_1), \dots, Z(x_n)$  en los puntos  $x_1, \dots, x_n$ . Dicha estimación, como se verá mas adelante, se hace con base en combinaciones lineales de las  $Z(x_i)$ . Si  $Z$  es una función aleatoria estacionaria con esperanza  $m$  y covarianza  $C(h)$  o variograma  $\gamma(h)$ , y  $Y$  es una combinación lineal finita del tipo

$$Y = \sum_{i=1}^n \lambda_i Z(x_i) \quad (8)$$

para cualesquiera pesos  $\lambda_i$ ,  $Y$  es una variable aleatoria y por lo tanto su varianza debe ser no negativa,  $\text{Var}\{Y\} \geq 0$ . Explícitamente, esta varianza puede escribirse como

$$\text{Var}\{Y\} = \sum_i \sum_j \lambda_i \lambda_j C(x_i - x_j) \geq 0 \quad (9)$$

La función de covarianza debe ser tal que asegura que la condición arriba se alada se cumpla, definiendose entonces como función positiva definida (Christakos, 1984). Usando la relación entre el variograma y la covarianza,

$$\text{Var}\{Y\} = C(0) \sum_i \lambda_i \sum_j \lambda_j - \sum_i \sum_j \lambda_i \lambda_j \gamma(x_i - x_j)$$

Si la varianza  $C(0)$  no existe, solo se asume la validez de la hipótesis intrínseca y en ese caso la varianza de  $Y$  se define bajo la condición

$$\sum_i \lambda_i = 0$$

quedando

$$\text{Var}\{Y\} = -\sum_{i,j} \lambda_i \lambda_j \gamma(x_i - x_j) \quad (10)$$

esto es, la función variograma debe ser tal que  $\text{Var}\{Y\} \geq 0$  sujeta a la condición de que  $\sum \lambda_i = 0$ . Por definición entonces, se dice que  $-\gamma(h)$  es una función condicionalmente positiva definida.

Muchas veces, los datos observados sugieren que la correlación entre dos variables  $Z(x)$  y  $Z(x+h)$  desaparece a medida que  $h$  crece,

$$\rho(h) \rightarrow 0, \text{ cuando } h \rightarrow \infty,$$

y en la práctica podemos poner  $C(h) = 0$  si  $h \geq a$ . La distancia  $a$  mas allá de la cual la correlación es despreciable se conoce como rango de influencia o simplemente rango. En el caso del variograma, conforme  $h$  crece  $\gamma(h)$  también lo hace por lo general, hasta que cuando  $h \geq a$  el variograma se estabiliza alrededor de un valor límite  $\gamma(\infty)$ , llamado meseta, el cual representa la varianza a priori de la función aleatoria, pues si  $\text{Cov}\{Z(x), Z(x+h)\}$  existe,

$$\gamma(\infty) = C(0) = \text{Var}\{Z(x)\}$$

Dichos variogramas son llamados modelos de transición, y corresponden a funciones aleatorias que no solo son intrínsecas, sino que también son estacionarias de segundo orden.

Otros factores importantes de la estructura de un fenómeno son su continuidad y regularidad espacial, los cuales están estrechamente relacionados con el comportamiento del variograma cerca del origen y tienen mucho que ver con la precisión de las estimaciones que se hagan (Chica Olmo, 1988). De acuerdo con Journel y Huijbregts (1978), dicho comportamiento puede ser de cuatro tipos:

- 1) Parabólico
- 2) Lineal
- 3) Dicontinuo en el origen (Efecto pepita)
- 4) Efecto pepita puro

Dichos autores discuten ampliamente las posibles causas y consecuencias de este comportamiento.

Finalmente, usando la definición de que  $\gamma(h)$  es una función positiva definida condicional, es posible demostrar que

$$\lim_{h \rightarrow \infty} \frac{\gamma(h)}{|h^2|} = 0$$

esto es, el variograma necesariamente crece mas lento que  $|h^2|$ . Como consecuencia de esto, un variograma que crezca al menos tan rapido como  $|h^2|$  es incompatible con la hipótesis intrínseca, debido muy probablemente a la existencia de una tendencia o deriva,

$$E\{Z(x)\} = m(x)$$

esto es, el valor esperado de  $Z(x)$  depende de la posición  $x$ .

### 1.2. Kriging

Como ya se mencionó antes, la principal aplicación de la geoestadística es la estimación de los valores de la función aleatoria en puntos donde no se hicieron mediciones. Lo mas

común es asumir que  $Z(x_1), Z(x_2), \dots, Z(x_n)$  son valores conocidos en los puntos  $x_1, \dots, x_n$ , y que el variograma o el correlograma han sido estimados. Entonces, lo que se quiere es estimar el valor

$Z(x_0)$  a través de una combinación lineal  $Z^*$  de las observaciones. El procedimiento de estimación conocido como kriging no solo lleva a un estimador<sup>1/</sup> lineal insesgado de varianza mínima dentro de esta clase, sino que también provee un estimador de  $\text{Var}\{Z-Z^*\}$  que ayuda a evaluar la precisión de la estimación.

Esta limitación a la clase de los estimadores lineales es más o menos natural, ya que solo implica la necesidad de conocer los momentos hasta de segundo orden de la función aleatoria. De esta manera, el estimador de kriging queda

$$Z^*(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (11)$$

El mejor estimador lineal se encuentra seleccionando los pesos  $\lambda_i$  de tal modo que

$$E\{Z^*(x_0) - Z(x_0)\} = 0$$

$$\text{Var}\{Z^*(x_0) - Z(x_0)\} = \text{mínimo}$$

donde  $Z(x_0)$  representa el valor verdadero (pero desconocido) de la función aleatoria en el punto  $x_0$ . Si el valor de  $E\{Z(x)\}$  es desconocido, una condición para que  $Z^*(x_0)$  sea un estimador insesgado de  $Z(x_0)$  es que

$$\sum_{i=1}^n \lambda_i = 1 \quad (12)$$

<sup>1/</sup> En realidad debido al carácter aleatorio de  $Z(x_0)$  la fórmula (11) no es un estimador sino un predictor. Sin embargo, en la literatura del tema se acostumbra llamarlo estimador y en el presente trabajo se seguirá esta costumbre

lo que traduce el problema a minimizar  $\text{Var}\{Z^*(x_0) - Z(x_0)\}$  sujeto a la condición (12). El uso de multiplicadores de Lagrange nos lleva a

$$\min_{\lambda} \text{Var}\{Z^*(x_0) - Z(x_0)\} - 2\mu(\sum \lambda_i - 1)$$

Por la definiciones del apartado anterior, se sigue que si la función aleatoria solo es intrínseca

$$\text{Var}\{Z^*(x_0) - Z(x_0)\} = -\sum_{i,j} \lambda_i \lambda_j \gamma_{ij} + 2 \sum_{j=1}^n \lambda_j \gamma_{0j} \quad (13)$$

donde  $\gamma_{ij} = \gamma(x_i - x_j)$ . Sustituyendo esta expresión para la varianza del error de estimación en la función a minimizar y derivando con respecto a las  $\lambda_i$ , se obtiene el sistema de ecuaciones lineales

$$\sum_{j=1}^n \lambda_j \gamma_{ij} + \mu = \gamma_{i0}, \quad i=1, \dots, n \quad (14)$$

que en notación matricial equivale a

$$Ga = b$$

con

$$G = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1n} & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \gamma_{n1} & \cdot & \cdot & \gamma_{nn} & 1 \\ 1 & \cdot & \cdot & \cdot & 0 \end{bmatrix}; \quad a = \begin{bmatrix} \lambda_1 \\ \cdot \\ \cdot \\ \lambda_n \\ \mu \end{bmatrix}; \quad b = \begin{bmatrix} \gamma_{10} \\ \cdot \\ \cdot \\ \gamma_{n0} \\ 1 \end{bmatrix}$$

De este modo, la solución es

$$a = G^{-1}b$$

En esta misma notación el mínimo del valor esperado del error de estimación o varianza de kriging es

$$\sigma_E^2 = a'b \quad (15)$$

En otras ocasiones lo que se quiere estimar es el valor promedio de  $Z$  en una región  $V$  de área (o volúmen)  $H_v$ , centrada en el punto  $x_0$ , problema de estimación que en la literatura geoestadística se conoce como kriging por bloque. En este caso, el variograma entre los puntos muestrales y el punto a estimar ( $\gamma_{10}$ ) es reemplazado por el variograma promedio entre los puntos muestrales y todos los puntos contenidos en  $V$ , que se define como

$$\bar{\gamma}(x_1, V) = \int_V \gamma(x_1, x) p(x) dx$$

con  $p(x)$  definida como

$$p(x) = \begin{cases} \frac{1}{H_v} & ; \text{ si } x \in V \\ 0 & ; \text{ si } x \notin V \end{cases}$$

y

$$\int_V p(x) dx = 1$$

Los coeficientes  $\lambda_i$  corresponden a la solución de

$$a = G^{-1}s$$

donde  $s' = [\bar{\gamma}(x_1, V), \dots, \bar{\gamma}(x_n, V), 1]$  y la varianza de estimación esta dada por



$$\sigma_v^2 = a's - \bar{\gamma}(V,V) \quad (16)$$

El segundo término del lado derecho representa la varianza promedio entre todos los puntos dentro del bloque  $V$ , algo análogo a la varianza dentro de bloque de la estadística clásica (Burguess & Webster, 1980). Por esta razón, la varianza de estimación de un bloque es menor que la de estimación puntual a menos que el volúmen que se estima sea un punto, en cuyo caso  $\bar{\gamma}(V,V) = 0$ .

Notese que si  $E\{Z(x)\} = m(x)$ ,

$$\begin{aligned} E\{[Z(x) - Z(x+h)]^2\} &= \text{Var}\{Z(x) - Z(x+h)\} + \{[m(x) - m(x+h)]^2\} \\ &= 2\gamma(h) + [m(x) - m(x+h)]^2 \end{aligned}$$

esto es, cuando el proceso presenta una tendencia o deriva, al estimar  $2\gamma(h)$  en realidad lo que se está estimando es  $2\gamma(h)$  distorsionado por una función cuadrática de la diferencia de medias, por lo que para una estimación efectiva del variograma el eliminar esta deriva de los datos es crucial (Cressie et. al., 1990).

Cuando existe deriva, esta puede modelarse como

$$m(x) = \sum_{i=0}^k b_i f_i(x) \quad , \quad (17)$$

donde las  $f_i$  son polinomios para  $i \neq 0$  y  $f_0 = 1$ . En ese caso se asumirá que los residuos

$$Y(x) = Z(x) - m(x)$$

satisfacen alguna de las hipótesis de estacionariedad, con variograma  $\gamma(h)$ . Bajo este contexto, si  $Z^*(x_0)$  es el estimador

puntual de  $Z(x_0)$ , entonces

$$E\{Z^*(x_0)\} = m(x_0) \quad \text{y}$$

$$\sigma_z^2 = a'b \quad (18)$$

donde  $a$  es la solución del sistema de ecuaciones  $Ga=b$ , con

$$a = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} \quad b = \begin{bmatrix} \gamma_{10} \\ \vdots \\ \gamma_{n0} \\ f_1(x_0) \\ \vdots \\ f_m(x_0) \end{bmatrix} \quad G = \begin{bmatrix} 0 & \gamma_{12} & \dots & \gamma_{1n} & 1 & f_1(x_1) & \dots & f_m(x_1) \\ \gamma_{21} & 0 & \dots & \vdots & 1 & \vdots & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 & \vdots & f_1(x_n) & \dots & f_m(x_n) \\ 1 & \dots & \dots & 1 & 0 & 0 & \dots & 0 \\ f_1(x_1) & \dots & f_1(x_n) & 0 & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ f_m(x_1) & \dots & f_m(x_n) & 0 & 0 & \dots & \dots & 0 \end{bmatrix}$$

sistema conocido como de kriging universal.

### 1.3. Cambio de soporte.

Para muchas variables geoestadísticas los valores de las muestras representan un promedio sobre un volumen mas que un valor puntual. Tal es el caso cuando se utilizan muestreadores de perforación como en geología y edafología por ejemplo, en cuyo caso se dice que el soporte de los datos no es puntual, lo que lleva al problema conocido como regularización y desregularización del variograma.

Estos términos se refieren a cambios en las características del variograma debido a diferencias en el soporte. Desregularización se refiere al proceso de estimar un variograma puntual a partir de muestras no puntuales, mientras que regularización se refiere al proceso inverso. Por ejemplo en limnología si en una parte del estudio hidrológico de un lago se utilizó una botella muestreadora de 3 lts. para estimar el variograma del pH digamos, y al mes siguiente se va a utilizar una botella muestreadora de 5 lts., no es posible comparar los

variogramas debido a que están referidos a diferente soporte (volumen), por lo que sería necesario desregularizarlos a variogramas puntuales para poder compararlos.

El cambio más conspicuo de un variograma puntual a uno no puntual es la reducción de la meseta, lo cual es un efecto comparable a la reducción de la varianza que se obtiene en estadística clásica al usar una muestra más grande (Myers, 1987). La metodología para llevar a cabo estos procesos así como las tablas necesarias se pueden consultar en los textos sobre el tema por ejemplo Journel y Huijbregts, (1978); David, (1977); Mc Cuen y Snyder, (1986).

#### 1.4. Análisis estructural.

Por lo que se ha expuesto hasta ahora, la utilidad de la Teoría de Variables Regionalizadas como herramienta para modelar un fenómeno distribuido espacialmente radica en la validez de considerar ya sea la hipótesis de estacionariedad de segundo orden o la intrínseca. Si los datos experimentales sugieren que alguno de estos supuestos puede considerarse como válido, el siguiente paso es la realización de un análisis estructural, con la finalidad de seleccionar un modelo para la función de covarianza o para el variograma, así como la estimación de los parámetros de dicho modelo.

El modelo que se escoja debe cumplir, como ya se mencionó, con ser una función positiva definida en el caso de la covarianza, o con ser condicionalmente positiva definida si se trata del variograma. En el caso de la covarianza, usualmente los modelos que se proponen son de la forma

$$C(h) = C(0)e^{-h/r} \quad (19)$$

aunque existen otras variantes (Ripley, 1981; Rodríguez-Iturbe, 1974).

A diferencia del análisis de series de tiempo, poco se ha hecho en el campo de los procesos espaciales para encontrar métodos de estimación de los parámetros de los modelos de covarianza. El enfoque mas común es tomar

$$\hat{C}(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} \{ [Z(x_i) - \bar{Z}] [Z(x_i+h) - \bar{Z}] \} \quad (20)$$

donde  $\bar{Z}$  es el promedio de las  $Z(x_i)$  y  $N(h)$  es el número de pares  $Z(x)$ ,  $Z(x+h)$  separados una distancia  $h$  (Ripley, 1981; Warrick et. al. 1986).

En el caso del variograma, los modelos mas comunmente usados son:

#### 1.- Esférico

$$\gamma(h) = \begin{cases} 0 & \text{si } h=0 \\ C_0 + C_1 \{ 1.5(h/a) - 0.5(h/a)^3 \} & \text{si } 0 < h < a \\ C_0 + C_1 & \text{si } a \leq h \end{cases}$$

$$C_0 \geq 0, C_1 \geq 0, a > 0$$

#### 2.- Potencial

$$\gamma(h) = \begin{cases} 0 & \text{si } h=0 \\ bh^\alpha & \text{si } 0 < h \end{cases}$$

$$b \geq 0, 0 \leq \alpha < 2$$

### 3.- Gaussiano

$$\gamma(h) = \begin{cases} 0 & \text{si } h=0 \\ C_0 + C_1[1 - \exp(-h^2/a^2)] & \text{si } 0 < h \end{cases}$$

$C_0 \geq 0, C_1 \geq 0, a > 0$

### 4.- Exponencial

$$\gamma(h) = \begin{cases} 0 & \text{si } h=0 \\ C_0 + C_1[1 - \exp(-h/a)] & \text{si } 0 < h \end{cases}$$

$C_0 \geq 0, C_1 \geq 0, a > 0$

La selección del modelo adecuado a los datos es crítica para la solución que se obtenga de las ecuaciones de kriging (Christakos, 1984). Una vez que se ha seleccionado este, el siguiente paso es estimar los parámetros del modelo, para lo cual existen varios caminos propuestos en la literatura (Myers, 1987, Zimmerman y Zimmerman, 1991), y métodos de ajuste para los casos en que ninguno de los modelos arriba mencionados dé un ajuste satisfactorio a los datos experimentales (Shapiro y Botha, 1991).

## CAPITULO 2. DISEÑO DE REDES DE MUESTREO

### 2.1. EL ENFOQUE CLASICO.

Se sabe que el cálculo del tamaño de muestra necesaria para estimar el valor medio de una variable  $X$  debe basarse en un intervalo de confianza para  $\mu$ ; esto requiere seleccionar un nivel de precisión aceptable para el estimador de la media así como un intervalo de confianza. De este modo, el tamaño requerido de muestra es

$$n = \frac{z_{\alpha}^2 \sigma^2}{(x - \mu)^2}$$

donde  $\sigma^2$  y  $\mu$  son la varianza y la media de la población respectivamente,  $|x - \mu|$  el nivel de precisión y  $1 - \alpha$  el nivel de confianza. Este criterio lleva implícitas las suposiciones de que las observaciones son independientes y que  $n$  será lo suficientemente grande como para aplicar el teorema central del límite.

Cuando se requiere reducir el tamaño de muestra, el enfoque clásico consiste en dividir el área de estudio de manera "lógica" en clases homogéneas, resultando el modelo

$$z_{ij} = \mu + \alpha_j + c_{ij} \quad (25)$$

en el que  $z_{ij}$  = valor de  $z$  en el punto  $i$  de la clase  $j$

$\mu$  = media general de  $z$

$\alpha_j = (\mu_j - \mu)$

y  $c_{ij}$  es un componente aleatorio no correlacionado espacialmente con media cero y varianza  $\sigma^2$ . El valor esperado de  $z$  en cualquier punto dentro de la clase  $j$  es  $\mu_j$  y se estima con una muestra de tamaño  $n_j$  utilizando

$$\hat{z}_{1j} = \bar{z}_j = \frac{1}{n_j} \sum_{k=1}^{n_j} z_{kj} \quad (26)$$

Este estimador esta sujeto a error, el cual tiene una varianza dada por

$$\begin{aligned} \sigma_e^2 &= E[(z_{1j} - \hat{z}_{1j})^2] \\ &= \sigma^2 + \sigma^2/n_j \end{aligned} \quad (27)$$

De acuerdo con este enfoque, incrementando  $n_j$  se pueden obtener estimaciones más precisas de  $\mu_j$ , aunque la varianza de la estimación nunca será menor que  $\sigma^2$ . Está implícito en este modelo el supuesto de que la división del área da cuenta de toda la variabilidad espacial de  $z$ , que  $c_{1j}$  es puramente aleatorio y que las líneas divisorias entre las clases son claras.

Bajo estos supuestos, el modelo clásico no puede ser mejorado. Sin embargo y como ya se mencionó en el capítulo anterior, las variables regionalizadas presentan una fuerte dependencia espacial que se manifiesta aún en las divisiones más finas. En estas circunstancias el modelo clásico resulta inadecuado para el diseño de redes de muestreo, pues no toma en cuenta la correlación espacial ni la posición de los puntos de muestreo (Mc Bartney et.al.,1981; Warrick et. al.,1987), por lo que el uso de la teoría de variables regionalizadas resulta ventajoso en este caso.

## 2.2. EL ENFOQUE GEOESTADISTICO

El uso de la teoría de variables regionalizadas en el diseño de redes de muestreo se basa en el hecho de que la varianza de estimación

$$\sigma_o^2 = b'Gb - \bar{y}(V, V) = 2 \sum_i \lambda_i \bar{y}_{i0} - \sum_i \sum_j \lambda_i \lambda_j \tau_{ij} \cdot \mu - \bar{y}(V, V)$$

no depende de los valores  $Z(x_i)$ . Los objetivos que se persigan en el diseño pueden ser la estimación del total de una variable de interés dentro de una región  $\mathcal{R}$ , la estimación ya sea de un bloque, o puntual para contruir un mapa de isopletas. En otras ocasiones el propósito es la selección de puntos adicionales dentro de un conjunto finito de estos para incrementar la precisión de las estimaciones con una red ya implementada; y para cada uno de estos objetivos de diseño la metodología a utilizar es diferente.

Puesto que la finalidad en todos los casos es encontrar una red óptima, siendo el criterio de optimalidad ya sea la minimización de la varianza para un costo fijo o la minimización del costo para una varianza máxima dada, es necesario plantear una función objetivo a minimizar o maximizar según sea el caso, que tome en cuenta tanto a la varianza de kriging como al costo, lo cual lleva el problema de diseño al campo de la investigación de operaciones.

Una solución analítica para la función objetivo que se proponga es intratable, debido a que la varianza de kriging puede cambiar debido a:

- 1) Cambios en el modelo y los parámetros del variograma
- 2) Cambios en la distancia entre los puntos de muestreo  $\tau_{ij}$
- 3) Cambios en la posición del bloque o punto a estimar  $\tau_{i0}$
- 4) Cambios en el número de vecinos usados en la estimación

lo cual se puede ver al analizar la expresión para  $\sigma_o^2$ , por lo que las aproximaciones a la solución del problema tienen que ser numéricas.

El investigador involucrado en un problema de diseño de redes de muestreo está restringido a seleccionar el modelo de



variograma que mejor ajuste a los datos experimentales, y el número de vecinos utilizados en la estimación es considerado un problema de eficiencia computacional, por lo que de los cuatro factores arriba mencionados los únicos manejables son la distancia entre las observaciones y la posición del punto a estimar (Olea, 1984).

La disposición de los puntos de muestreo puede clasificarse de la siguiente manera (Mc Arthur, 1987; Rippley, 1981):

Aleatoria	{	Simple
		Estratificada
Regular	{	Simple
		Preferencial

Dentro de los tipos regulares, las disposiciones más comunes son triangular equilateral, rectangular y hexagonal, pues se consideran las que de manera más eficiente pueden cubrir una región bajo estudio y son más fáciles de implementar en el campo (Burgess y Mc Bartney, 1980, Yftanis et. al, 1987).

Debido a que los variogramas son funciones monótonas no decrecientes, al menos para las distancias donde usualmente se hacen estimaciones por kriging, la varianza de estas tenderá a aumentar a medida que la distancia entre el bloque o punto a estimar  $x_0$  y los puntos observados  $x_i$  aumente. Claramente, el disminuir esta distancia o el añadir nuevos puntos de observación hace que la varianza disminuya y es en este hecho que se basan los criterios geoestadísticos para el diseño de las redes de muestreo.

Para poder estimar el variograma y calcular las varianzas de kriging necesitamos tener una red inicial, misma que

se buscará optimizar de acuerdo a los criterios ya mencionados, y el modo de hacer esto dependerá en gran parte del tipo de instrumentos de medición de la variable  $Z(x)$ . Así, cuando los instrumentos de medición sean móviles o su traslado del sitio actual a otro no implique un costo excesivo, la mejor forma de optimizar la red es escoger un patrón regular e ir cambiando la distancia entre las observaciones hasta obtener una red con las especificaciones de costo y precisión deseadas. Este es el caso por ejemplo de las redes de monitoreo oceanográficas, atmosféricas, limnológicas, etc., en las que los puntos de observación pueden cambiarse fácilmente una vez encontrado el diseño óptimo.

Para estos casos existen varias metodologías para encontrar la red óptima. Una es calcular  $\sigma_o^2$  cuando  $x_o$  esta en el centro de la red para un rango de distancias entre observaciones y hacer una gráfica de distancia vs.  $\sigma_o^2$  (Mc Bartney y Webster, 1981; Mc Bartney et. al., 1981; Burgess y Mc bartney, 1981). Si se desea por ejemplo una red con varianza máxima permitida de  $\sigma_{max}^2$ , bastará localizar en la gráfica este punto y por proyección en el eje de distancias se encontrará la distancia entre puntos. Este proceso puede llevarse a cabo con cualquiera de los patrones regulares y seleccionar el que mas convenga a los fines del muestreo. Yftanis y Flatman (1988) proponen un método similar, solo que la gráfica usa tamaño de muestra ( número de puntos en la red) en lugar de distancia.

Hughes y Lettenmier (1981) proponen minimizar la varianza de kriging con respecto a la posición de los puntos de observación en el caso de la estimación de un promedio de área, aproximandola por los primeros dos términos de una expansión en series de Taylor alrededor de un conjunto de puntos  $X^0 = \{(X_1^0, Y_1^0), \dots, (X_n^0, Y_n^0)\}$ ,

$$F = \sigma_w^2(\lambda) = c + (X_1 - X_1^0) \frac{\partial F}{\partial X_1} + \dots + (Y_n - Y_n^0) \frac{\partial F}{\partial Y_n} \quad (28)$$

$X^0$  son las posiciones iniciales de la red y X son las localidades donde F va a ser evaluada. Esta metodología aunque aparentemente simple, presenta varias desventajas. Una de ellas es que las derivadas no son fáciles de evaluar pues dependen del variograma y por tanto no es fácil encontrar un método numérico que aproxime la solución y otra es el riesgo que se corre de llegar a un mínimo local debido a que la varianza de kriging como ya se mencionó no cumple con las suposiciones de convexidad.

Bogardi et. al (1985) tratan el problema de diseño en el contexto de toma de decisiones con criterio múltiple. En este caso, se considera que una red de muestreo esta caracterizada por la densidad de estaciones o por el área cubierta por una estación (T), y puede representarse por tres variables de decisión:

- T = Area cubierta por estación
- a = Distancia promedio entre estaciones en el eje de orientación de la red
- $\phi$  = Angulo de orientación de la red con respecto al eje N-S

Suponiendo que se tienen k variables de interés, para cada una de ellas la varianza de kriging  $\sigma_e^2$  es función de las variables de decisión T, a y  $\phi$ . Un objetivo es

$$\min_{j=1, \dots, k} \sigma_j^2 (T, a, \phi)$$

Otro conjunto de objetivos es minimizar el costo del muestreo, el cual se incrementa al disminuir T, de aquí que con respecto a este criterio se busque maximizar T. La mejor red es aquella que esté más cerca de una red ideal de acuerdo a alguna métrica  $I_p$ , lo cual nos lleva a la función objetivo

$$CF(\sigma, T) = \left\{ \alpha_1^{p_3} \left[ \sum_{j=1}^k \beta_j^{p_1} \left[ \frac{\sigma_j - \sigma_j^*}{\max \sigma_j - \sigma_j} \right]^{p_1} \right] + \alpha_2^{p_3} \left[ \frac{T^* - T}{T^* - \min T} \right]^{p_3} \right\}^{1/p_3} \quad (29)$$

donde  $\alpha_1$  y  $\alpha_2$  son los pesos que concilian los intereses estadísticos y económicos respectivamente,  $\beta_j$  son los pesos a la conciliación entre los criterios estadísticos,  $\sigma_j^*$  y  $T^*$  son los valores ideales y  $p_1$  y  $p_3$  son parámetros de programación.

Puesto que las  $\sigma_j^2$  no son cantidades físicas, un criterio de valor absoluto es adecuado y entonces  $p_1=1$ ; el equilibrio entre la precisión y el costo se logra desde el punto de vista de mínimos cuadrados y se toma  $p_3=2$ , quedando la función

$$CF(\sigma, T) = \left\{ \alpha_1^2 CF_1^2 + \alpha_2^2 CF_2^2 \right\}^{1/2} \quad (30)$$

donde  $CF_1$  es la función objetivo estadística

$$CF_1 = \sum_{j=1}^k \beta_j \left[ \frac{\sigma_j - \sigma_j^*}{\max \sigma_j - \sigma_j} \right] \quad (31)$$

y  $CF_2$  es la función objetivo del criterio económico, que depende de  $T$

$$CF_2 = \frac{T^* - T}{T^* - \min T} \quad (32)$$

El algoritmo propuesto por estos autores puede resumirse en lo siguiente: para un valor fijo de  $T$ , minimícese la función  $CF_1$  para obtener los parámetros  $\phi(T)$  y  $a(T)$  y asigne pesos  $\alpha_1$  y  $\alpha_2$ . Con estos valores minimice  $CF$  para obtener una solución compuesta  $T^*$ ,  $\phi(T^*)$ ,  $a(T^*)$ ; si esta solución es satisfactoria

terminar y si no, cambiar los valores  $\alpha_i$  y volver a minimizar CF.

En muchas ocasiones una vez que se ha implementado un punto de muestreo, este no conviene o no se puede cambiar de lugar debido a que este simple hecho eleve el costo del muestreo, por ejemplo en el caso de la exploración petrolera, los pozos hidrogeológicos o las estaciones meteorológicas. En estos casos estamos restringidos a únicamente seleccionar puntos de muestreo adicionales a la red ya existente.

Supongamos que para mejorar una red  $S = \{x_1, \dots, x_n\}$  se tiene un conjunto finito de puntos adicionales  $S_p = \{x_{n+1}, \dots, x_{n+m}\}$ , del cual se va a seleccionar un punto adicional. Sea  $\sigma^2(x_0, S)$  la varianza de kriging al estimar el punto  $x_0$  con la red inicial  $S$  y  $\sigma^2(x_0, S_{+j})$  la varianza al estimar el mismo punto con la red aumentada por el punto  $x_j \in S_p$ , esto es,  $S_{+j} = S \cup \{x_j\}$ ;  $j = 1, \dots, m$ .

Cressie et. al (1990) sugieren que de hecho los puntos nuevos a incluir deben estar cercanos a las zonas críticas, que se consideran como aquellas en donde  $Z(x)$  es mayor que una valor  $K$  dado. Con esto en mente, un criterio para seleccionar un punto adicional es escoger aquel que ayude a predecir al resto del conjunto  $S_p$  con menor varianza, sujeto a la condición de que  $Z(x) > K$ . De este modo, la función objetivo a minimizar es

$$\sigma^2(x_j) = \frac{\sum_{\substack{i=1 \\ i \neq j}}^m I(Z(x_i) > K) \sigma^2(x_i; S_{+j})}{\sum_{\substack{i=1 \\ i \neq j}}^m I(Z(x_i) > K)} \quad (33)$$

donde  $\sigma^2(x_j)$  es la varianza de estimación promedio sobre el resto de los sitios potenciales utilizando la red aumentada  $S_{+j}$ , e  $I(A)$

es una función indicadora. Este método puede extenderse al caso de querer incluir más de un punto adicional, pero si este número es muy grande se necesitan algoritmos numéricos. En el caso de que se desee desechar un punto de la red, la sugerencia es quitar aquel que pueda predecirse con menor varianza utilizando la red disminuída  $S_{-i} = S - \{x_i\}$ . En este caso la función a minimizar es  $\sigma^2(x_i; S_{-i})$ .

Otro método de inclusión de observaciones adicionales es el que propone Szidarovszky(1983a, 1983b). Para describirlo, es necesario cambiar la notación un poco. Si se tiene una red previa con n puntos y se desea estimar un área V, sean

$$b = \begin{bmatrix} 1 \\ \bar{y}_{1v} \\ \vdots \\ \bar{y}_{nv} \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 & \dots & 1 \\ 1 & \bar{y}_{11} & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \bar{y}_{n1} & \dots & \bar{y}_{nn} \end{bmatrix}, \quad a = \begin{bmatrix} \mu \\ \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}$$

los vectores y matriz del sistema de ecuaciones de kriging modificados, y sea  $g' = [1, \bar{y}_{1n+1}, \dots, \bar{y}_{nn+1}]$ . El sistema de kriging involucrando al punto adicional  $x_{n+1}$  puede escribirse como

$$\begin{bmatrix} A & g \\ g' & \bar{y}_{n+1, n+1} \end{bmatrix} \begin{bmatrix} a \\ \lambda_{n+1} \end{bmatrix} = \begin{bmatrix} b \\ \bar{y}_{n+1, v} \end{bmatrix} \quad (34)$$

La inversa de la matriz de coeficientes es de la forma

$$\begin{bmatrix} Y & u \\ u' & s \end{bmatrix}; \quad Y = A^{-1} - \mu\mu'; \quad m = A^{-1}g \\ u = -ms; \quad s = (-g'm)^{-1} \quad (35)$$

haciendo un poco de álgebra, se puede demostrar que la relación entre la varianza con la red original y con la red aumentada es

$$\sigma_{n+1}^2 = \sigma_n^2 - \frac{1}{\gamma_{n+1,n+1} + g'A^{-1}g} (\gamma_{n+1,v} - g'A^{-1}b)^2 \quad (36)$$

El criterio es seleccionar aquel punto que minimize  $\sigma_{n+1}^2$ , lo cual equivale a maximizar el segundo término del lado derecho de la expresión. Este, no satisface las condiciones de concavidad necesarias para aplicar métodos de gradiente, de manera que la solución aproximada se obtiene discretizando el problema. En este caso el algoritmo es: Dado un conjunto de alternativas  $S_p$ , para cada punto calcule los valores  $\bar{\gamma}_{n+1,v}, \gamma_{n+1,1} (1 \leq v \leq n)$  y evalúe el segundo término del lado derecho de la expresión. El punto a incorporar será aquél que la maximice.

Para el caso en que se desee incluir varios puntos en  $S_p$ , el sistema de ecuaciones de kriging deberá modificarse acordeamente. Una descripción detallada del algoritmo en este caso puede encontrarse en Szidarovszky (1983b) y en Samper y Carrera (1990).

### CAPITULO 3. METODOLOGIA

Para el diseño de la red de muestreo oceanográfica se tomó como base la red de monitoreo de ictiopláncton del Instituto Nacional de la Pesca, la cual consta de 182 puntos de observación distribuidos en forma cuadrada en la Zona Económica Exclusiva del Golfo de México y Mar Caribe (figura 1), misma que fué implementada con la finalidad de evaluar la biomasa de recursos marinos con potencial de explotación comercial por medio de censos larvales, técnica muy utilizada cuando se desea estimar el tamaño de poblaciones de peces cuando no existen datos de captura comercial. La red esta referida a una cuadrícula tal que la distancia entre puntos en la dirección horizontal y vertical es de 27 millas náuticas, de modo que cada unidad de separación en la malla es 2.7 m.n.

#### 3.1. ANALISIS ESTRUCTURAL.

Es la primera parte de todo análisis geoestadístico pues con él se busca establecer el tipo de correlación espacial existente dentro del área de estudio. Para esto se utilizaron los datos obtenidos en los cruceros efectuados en el verano de 1982, por ser esta campaña en la que se logró visitar el mayor número de puntos, los cuales totalizaron 146. Los valores observados corresponden a las abundancias larvales de atún aleta negra (*Thunnus atlanticus*), atún aleta azul (*T. thynnus*) y barrilete (*Katsuwonus pelamis*), expresadas como cantidad de larvas por 10 m<sup>2</sup> de superficie marina (Olvera et. al, 1988).

Un primer análisis exploratorio de los datos para las tres especies reveló que aunque los valores  $Z(x)$  se distribuían de manera irregular, existía tendencia de la zona NW y SW a presentar valores mas altos que el resto de la región bajo estudio, por lo que sería necesario ya sea suponer la existencia de una deriva en dirección W-E y entonces trabajar con los residuos  $Y(x)=Z(x)-m(x)$ , ó dividir la región en varias subregiones y utilizar la hipótesis



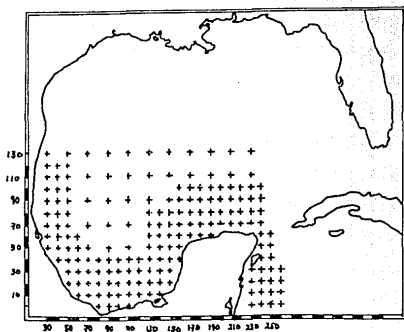


Figura 1. Plan básico de estaciones de monitoreo para ictiopláncton en la Zona Económica Exclusiva del Golfo de México.

cuasiestacionaria en cada una de ellas, optandose por esta última por ser la solución que ofrece menos problemas (Delhome, 1979; Cressie et. al., 1990).

Para dividir la Zona Económica Exclusiva en subregiones se utilizó el método de análisis de conglomerados conocido como K-means, usando como variables de clasificación tanto la posición geográfica como los valores  $Z(x)$  de las tres especies de interés. Se escogieron cinco puntos como centros de conglomerado, aunque los resultados sugirieron solo la existencia de cuatro relativamente bien definidos (figura 2). De estos, en el cuarto, que abarca la zona del Mar Caribe no hubo estaciones con valor positivo lo que hizo imposible la estimación de variogramas en dicho estrato, por lo que se descartó del análisis y solo se trabajó con los restantes.

En cada estrato se procedió a la obtención del variograma experimental de cada especie, utilizando el paquete GEOEAS (Englund y Sparks, 1989), el cual utiliza el estimador de momentos. Este estimador es, como todos los estimadores de momentos, muy sensible a valores extremos por lo que en la obtención de los variogramas experimentales se compararon los resultados obtenidos al utilizar todas las observaciones y quitando los valores más altos (Issaks y Srivastava, 1991), encontrándose poca diferencia, sobre todo en los primeros rezagos.

Los variogramas experimentales se obtuvieron para las direcciones 0, 45, 90 y 135 grados con la finalidad de detectar la existencia de anisotropía en cada especie, obteniéndose resultados negativos por lo que en el resto del trabajo se utilizarán variogramas isotrópicos.

Una vez obtenidos estos, el siguiente paso es la selección de un modelo teórico que ajuste razonablemente los datos experimentales, y los resultados se muestran en la siguiente tabla:

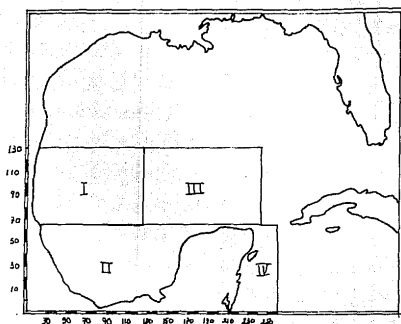


Figura 2. Estratos en que se dividió el área de estudio para aplicar la hipótesis cuasi-estacionaria.

ESTRATO	VARIABLE	VARIOGRAMAS TEORICOS			
		TIPO	C <sub>0</sub>	C <sub>1</sub>	r <sup>1/</sup>
I	T. atlantic.	ESFERICO	33	19	25
I	T. thynnus	ESFERICO	0.4	1.3	30
I	K. pelamis	ESFERICO	5	14	30
-----					
II	T. atlantic.	ESFERICO	60	55	28
II	K. pelamis	ESFERICO	7	6	55
-----					
III	T. atlantic.	ESFERICO	0	17	45
III	T. thynnus	ESFERICO	4	3	45
III	K. pelamis	ESFERICO	4	3	45

1/ EXPRESADO EN UNIDADES DE LA RED

Como se puede apreciar todos son de tipo esférico con pepita, lo cual indica que las abundancias larvales se comportan de manera muy irregular, lo cual es de esperarse debido al comportamiento gregario de los peces pelágicos durante el desove. Asimismo, cabe hacer notar que en el estrato II no se pudo estimar el variograma para atún aleta azul por no haber valores positivos para esta especie en él.

### 3.2. DISEÑO DE LA RED DE MUESTREO.

Dado que para el tipo de variables que se observan en esta clase de estudios los instrumentos de medición son transportados en un buque, los puntos de medición pueden mantenerse fijos y buscar mejorar la red existente ya sea quitando o añadiendo otros, o mejorarla aumentando o disminuyendo la distancia entre puntos, lo que implica cambiarlos de su posición

actual. De entre los métodos existentes en la literatura, se utilizaron los de puntos móviles (Mc Bartney et. al., 1981; Bogardi et. al., 1985; Yftanis y Flatman, 1988) por ser computacionalmente más sencillos y por no presentar el riesgo de llegar a un mínimo local en la minimización de  $\sigma_c^2$  tal como podría ocurrir con los métodos de puntos fijos.

La gráfica de varianza vs distancia entre nodos se obtuvo utilizando un programa hecho ex profeso, llamado OPTIM4, el cual está escrito en FORTRAN IV. En una corrida de este, el usuario escoge un rango de distancias entre nodos, el modelo y parámetros del variograma, el tipo de estimación, ya sea puntual o por bloque y el diseño de la red. La salida consiste en una tabla de la varianza de kriging para diferentes distancias entre nodos, utilizando los 12 vecinos más cercanos. Con este programa se hicieron corridas para todas las especies en los tres estratos con los diseños cuadrado, hexagonal y triangular.

Dado que por este método se obtiene una solución para cada especie, se construyeron otras redes de muestreo utilizando el criterio propuesto por Bogardi et. al (1985), aplicandolo a los tres diseños antes mencionados para obtener la mejor red de criterio múltiple en cada caso. La función objetivo a minimizar fué la expresión (29), tomándose los siguientes valores de los parámetros:

$\sigma^* = 60\%$  del valor máximo de  $\sigma_c^2$  ;

$\sigma$  = efecto pepita menos la varianza dentro del bloque;

$T^*$  = área cuando la distancia del punto  $x_0$  al vecino más cercano es igual al rango del variograma;

$T = 1$  unidad cuadrada de la red original.

Los valores de los coeficientes  $\beta_i$  se tomaron iguales para las tres especies debido a que no existe una preferencia especial por obtener más precisión en la estimación de alguna de ellas.

La selección de los valores obedeció a causas empíricas principalmente, excepto en el caso de las  $\sigma^2$ , la cual se hizo porque en las gráficas de  $\sigma^2$  vs distancia para cada especie se vió que al disminuir la varianza de estimación mas de este valor aumentaba sustantivamente el número de puntos de muestreo. Con las redes así obtenidas se procedió por último a evaluarlas, para lo cual se calculó la varianza de kriging en varios puntos de ellas con el propósito de ver si cumplían con las especificaciones requeridas en cuanto a varianza máxima y menor número de puntos que la red existente se refiere. Esto último solo se hizo con las redes de criterio múltiple pues en las de criterio uniespecífico se espera de antemano que en el peor de los casos la varianza de kriging solo será 60% del límite máximo de esta.

#### 4. RESULTADOS.

El agrupamiento de las 146 observaciones disponibles por el método de K-means sugirió que los datos se agrupaban en cuatro conglomerados o estratos definidos de la siguiente manera:

ESTRATO	MINIMO		MAXIMO		AREA	ZONA	#OBSERV.
	ESTE	NORTE	ESTE	NORTE			
I	30	65	125	130	6175	NW	32
II	30	0	150	65	7800	SUR	53
III	125	65	240	130	7475	NE	49
IV	240	0	260	65	1300	CARIBE	12

NOTA: Los valores en esta tabla están dados en las unidades de la red. Una unidad equivale a 2.7 millas náuticas.

Esta división es similar a la que Yañez et. al. (1986) hicieron basándose en características físico-químicas, geológicas y de composición biótica de las comunidades de la zona costera, por lo que considerando las interrelaciones existentes entre ésta y la zona oceánica podemos tomar los resultados del agrupamiento de las observaciones como adecuado para considerar cierto grado de homogeneidad en cada estrato y poder aplicar la hipótesis cuasi-estacionaria en el resto del análisis.

Como se mencionó en el capítulo anterior, aunque la red de monitoreo existente cubre satisfactoriamente la Zona Económica Exclusiva, el número de puntos con que cuenta hacen sumamente difícil visitarlos todos, pues además del tiempo necesario para que los barcos se trasladan de un punto a otro, el tiempo de trabajo en cada uno es de alrededor de tres horas (Olvera, com. pers.), por lo que el criterio básico de diseño que se tomó fué la disminución del número de puntos en la red sin incrementar

sustantivamente la varianza de las estimaciones con respecto a la red actual.

Cuando se desea una red de muestreo con patrón regular, los diseños sugeridos son el hexagonal, triangular y cuadrado, pues son los que con mayor eficiencia pueden cubrir una región bajo estudio (Gilbert, 1988), siendo superiores a aquellos en que los puntos se colocan de manera aleatoria (Olea, 1984). Sin embargo, estos tres diseños presentan diferencia entre sí, que dependiendo del objetivo del muestreo harán mejor a una red sobre otra y esto se debe a las relaciones geométricas que guardan entre ellos.

Si se tiene una región de área  $A$  que se desee cubrir con una red cuadrada cuya distancia mínima entre puntos es  $a$ , el número de cuadros necesarios será

$$N_c = \frac{A}{a^2}$$

y si de estos cuadros  $k$  de ellos caben en la dirección horizontal y  $m$  en la dirección vertical,  $N_c = km$ , por lo que el total de puntos de muestreo para la región es

$$P_c = (k+1)(m+1) \quad (37)$$

A su vez, si  $N_t$  es el número de triángulos equiláteros que cubren la misma región, cada uno con lado  $a$  y superficie  $a^2\sqrt{3}/4$ , se tiene que

$$N_t = 4N_c/\sqrt{3}$$

por lo que el número de puntos de muestreo para el diseño triangular es

$$P_t = (2/\sqrt{3})(k+1)(m+1) \quad (38)$$



Asimismo, es posible demostrar que el número de puntos necesarios para cubrir la misma región con hexágonos de lado  $a$  es

$$P_h = (4/3\sqrt{3}) (K+1) (m+1) \quad (39)$$

lo cual lleva a las siguientes relaciones entre el número de puntos necesarios para cubrir el área  $A$  con una red cuadrada, triangular o hexagonal:

$$P_t = 1.15P_c \quad ; \quad P_h = 0.77 P_c$$

Como se puede ver, para una misma distancia entre puntos, el diseño hexagonal es el que requiere un menor número de estos y el triangular el que más. Teniendo en cuenta estas relaciones es posible comparar estas tres geometrías en cuanto a su comportamiento para la estimación de la abundancia larval de las especies de interés.

En las figuras 3 a 5 se presentan los resultados obtenidos al evaluar la varianza de kriging por bloque de  $10 \times 10$  unidades para las tres especies en cada diseño y en todos los estratos. En ellas se puede ver que para una distancia entre puntos  $a$ , el diseño triangular es el que nos da una menor varianza de kriging y esto se debe a que la distancia del centro del triángulo a cualquier vértice es menor que en el caso de un cuadrado o de un hexágono, de manera que en estas circunstancias el diseño triangular es más eficiente en este sentido; sin embargo, como se mencionó en el párrafo anterior, requiere 15% más puntos que el diseño cuadrado, lo que debe considerarse para decidir por alguno de los dos. También es notorio que aunque el diseño hexagonal requiere una cantidad menor de puntos de muestreo es el que daría una mayor varianza de kriging, quedando en desventaja con respecto a los anteriores en términos de precisión.

En el diseño de redes de monitoreo un caso común es aquél en que por cuestiones económicas solo se dispone de recursos

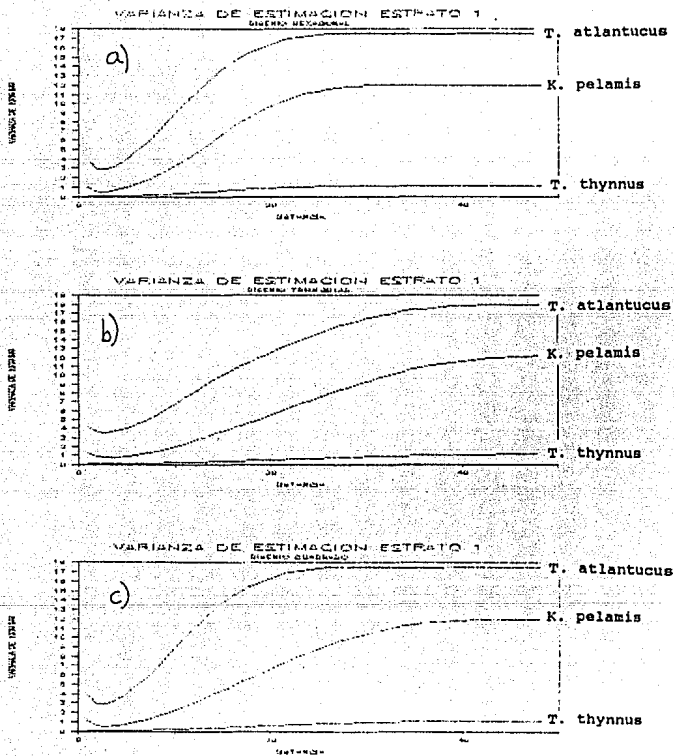


Figura 3. Comportamiento de la varianza de kriging para los tres diseños y las tres especies de interés en el Estrato 1: a) Diseño Hexagonal; b) diseño triangular; c) diseño cuadrado.

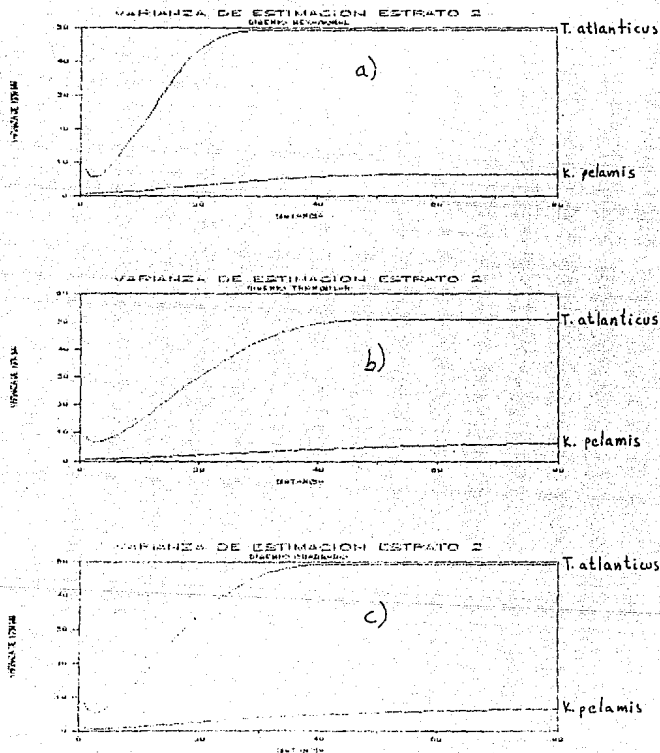


Figura 4. Comportamiento de la varianza de kriging para los tres diseños y las tres especies de interés en el Estrato 2: a) Diseño Hexagonal; b) diseño triangular; c) diseño cuadrado.

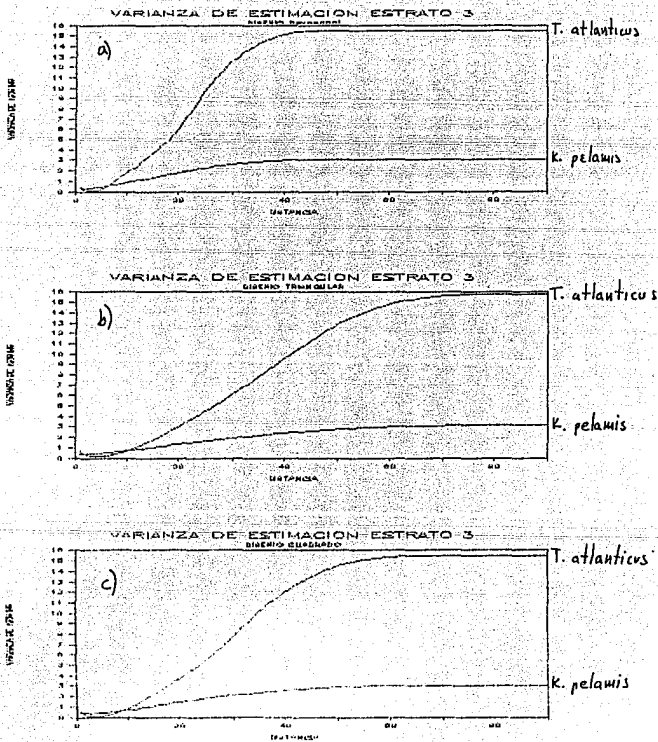


Figura 5. Comportamiento de la varianza de kriging para los tres diseños y las tres especies de interés en el Estrato 3: a) Diseño Hexagonal; b) diseño triangular; c) diseño cuadrado.

para implementar una cantidad determinada de puntos de muestreo, digamos  $r$ , y entonces se deberá escoger aquella disposición espacial que nos minimice la varianza bajo esta restricción.

Haciendo un poco de álgebra, se puede demostrar que para cubrir un área  $A$  con  $r$  puntos en forma de cuadros, la distancia mínima entre estos debe ser

$$d_c = \left[ \frac{A}{r} \right]^{1/2}$$

y por la relación existente entre el número de puntos con los casos triangular y hexagonal, la distancia mínima entre puntos será

$$d_{tr} = \left[ \frac{1.15 A}{r} \right]^{1/2} ; d_{h} = \left[ \frac{0.77 A}{r} \right]^{1/2}$$

respectivamente por lo que las distancias para un número dado de puntos guardan la relación  $d_h < d_c < d_r$ .

Si para el estrato I por ejemplo tuviésemos solo dinero para 24 puntos, las distancias serían  $d_c=16$ ,  $d_h=14$  y  $d_r=17$  unidades. Las varianzas de kriging como se observa en las graficas guardan entonces un orden inverso, i.e.,  $\sigma_h^2 > \sigma_c^2 > \sigma_r^2$ , lo cual hace superior al diseño triangular sobre los otros dos y el único factor que podría descalificarlo es la existencia de dificultades técnicas para implementarlo en el campo, cosa poco probable en un muestreo oceanográfico.

Aparte de las ventajas aquí mencionadas, el diseño triangular permite la obtención del variograma experimental en tres direcciones para rezagos cortos, lo cual resulta deseable cuando se pretende detectar anisotropías (Yftanis et. al., 1987; Russo, 1984).

Quando el objetivo es encontrar una red cuya varianza de kriging no exceda un valor  $\sigma_{\max}^2$ , de las gráficas podemos ver que al proyectar este valor en el eje de las distancias obtenemos un valor para cada diseño, en el mismo orden que cuando se tiene un número fijo de puntos de muestreo. Sin embargo en este caso, al convertir las distancias a número de puntos necesarios para cubrir el área A se obtiene que la menor cantidad de estos se da con la configuración cuadrada y por tanto será mejor en este caso. Así, podemos decir que los mejores diseños regulares son el cuadrado cuando se tiene una cota para la varianza de kriging y el triangular cuando las restricciones son en el costo del muestreo, el cual es proporcional al número de puntos.

La varianza de kriging en todos los casos alcanza un valor  $\max \sigma^2$  después de que la distancia entre los nodos de la red sobrepasa un cierto valor, debido a que ya la distancia del punto  $x_0$  al vecino más cercano es mayor que el rango del variograma. Si  $d(x_0, x)$  es la distancia entre el punto  $x_0$  y su vecino más cercano, se puede demostrar que la varianza de kriging alcanza su cota  $\max \sigma^2$  cuando  $d(x_0, x) \geq \sqrt{2} r$  para el diseño cuadrado,  $d(x_0, x) \geq \sqrt{3} r$  en el triangular y  $d(x_0, x) \geq r$  para el hexagonal.

Quando se hace estimación por bloque a medida que la distancia entre los nodos decrece, también lo hace la varianza de kriging, hasta que la separación entre los nodos es tan pequeña que algunos de ellos comienzan a caer dentro del bloque a estimar y por ello es que en las gráficas se ve que existe un valor a partir del cual, el acercar los puntos de muestreo incrementa la varianza (Mc Bartney y Webster, 1981).

La aplicación de estos criterios a los datos para las tres especies de atún nos lleva a diferentes "óptimos" para cada especie, tal como se muestra en la siguiente tabla

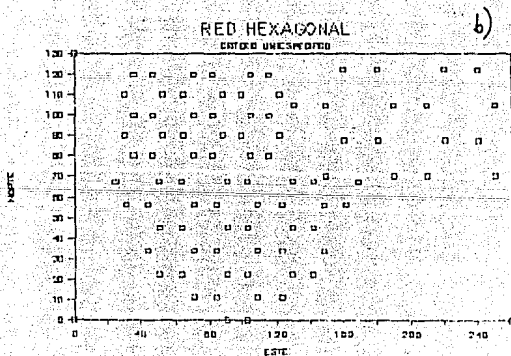
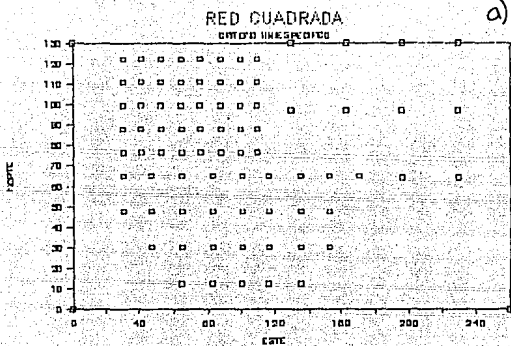
ESTRATO	ESPECIE	CUADRADO		HEXAGONAL		TRIANGULAR	
		DIST*	# PTS.	DIST*	# PTS.	DIST*	#PTS.
I	T. atlantic.	11.5	40	11.5	30	16.0	31
I	T. thynnus	21.0	20	16.0	23	24.5	13
I	K. pelamis	21.5	20	16.0	23	24.0	14
II	T. atlantic.	17.5	28	13.5	40	19.5	25
II	K. pelamis	31.0	8	25.0	9	36.0	7
III	T. atlantic.	33.0	11	24.5	17	39.0	14
III	T. thynnus	25.0	12	20.0	14	28.5	10
III	K. pelamis	25.0	12	20.0	14	28.5	10

\* Tomando  $\sigma_{\max}^2 = 0.6(\max\sigma^2)$

Se puede apreciar que si se escoge un diseño específico resulta aún complicado decidir a qué distancia se pondrán los nodos de la red, pues cada especie nos sugiere un valor distinto. Una opción es tomar la menor de las distancias dentro de cada estrato y esto nos garantiza que en el peor de los casos la máxima varianza de kriging será 60% de  $\max\sigma^2$  para alguna especie y menor en las otras dos. Así, si por ejemplo deseamos un diseño cuadrado, la distancia entre puntos sería de 11.5 en el estrato I, 17.5 en el II y 33 en el III, dándonos un total de 79 puntos de muestreo repartidos 40, 28 y 11 en los estratos I, II y III respectivamente. La red de estaciones con cada diseño se presenta en la figura 6.

La metodología de construcción de la red de monitoreo descrita en el párrafo anterior puede no ser eficiente en cuanto a costo, por lo que se probó el método propuesto por Bogardi et. al (1985), el cual busca conciliar los criterios estadísticos y económicos a través de la minimización de la función (29).

En las figuras 7 a 9 se muestra el comportamiento de esta función con respecto a la distancia para diferentes valores de  $\alpha_1$  por estrato y para las tres disposiciones regulares entre puntos





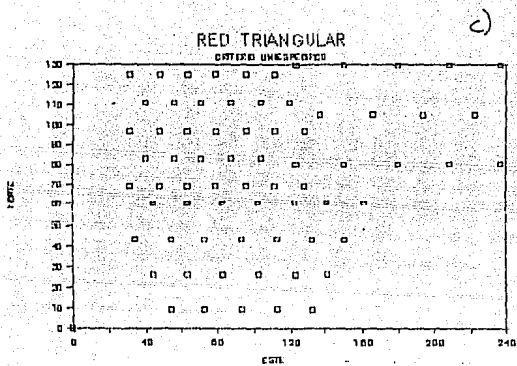


Figura 6. Redes obtenidas utilizando el criterio unespecifico:  
a) Red cuadrada; b) red hexagonal; c) red triangular.

de muestreo, notando que a medida que el valor de  $\alpha$  decrece la distancia óptima aumenta tal como es de esperarse. Con este enfoque, aún cuando se de todo el peso al criterio estadístico la distancia entre puntos que se obtiene para cualquier diseño resulta en una disminución significativa de puntos con respecto a la red original, tal como se presenta en la siguiente tabla:

ESTRATO	DISEÑO	DISTANCIA	No. PTS.
I	CUADRADO	18	15
	TRIANGULAR	24	16
	HEXAGONAL	16	21
II	CUADRADO	16	36
	TRIANGULAR	17	33
	HEXAGONAL	13	40
III	CUADRADO	35	6
	TRIANGULAR	39	8
	HEXAGONAL	23	16

De esta tabla se tiene que el número de nodos en la red de monitoreo sería 57, 77 y 57 para los diseños cuadrado, hexagonal y triangular respectivamente, que comparados con los 182 nodos de la red original significan una reducción aproximada del 68, 57 y 68%.

Comparando los resultados de esta tabla con los de la anterior, se ve que existe cierta concordancia sobre todo en los estratos II y III, y las diferencias se deben a que este método toma en cuenta el costo del muestreo a través del área de influencia de una estación, así como qué tanto se aleje uno del valor ideal para cada variable ( $\sigma_k^*$  y  $T^*$ ).

Al usar este método de diseño, la varianza resultante nos da una reducción porcentual de  $\max \sigma^2$  diferente para cada especie, corriendo el riesgo de submuestrear algunas variables y

# ESTRATO 1 ESTIMACION POR BLOQUE

DISEÑO CUADRADO ERROR MÚLTIPLE

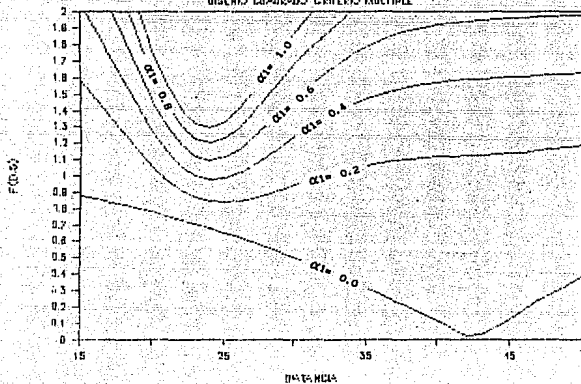


Figura 7. Comportamiento de la función  $CF(\sigma, T)$  con respecto a la distancia entre nodos, para distintos valores de  $\alpha_1$ , cuando el diseño es cuadrado.

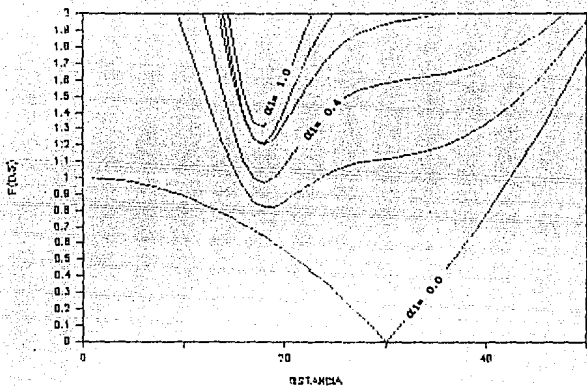


Figura 8. Comportamiento de la función  $CF(\sigma, T)$  con respecto a la distancia entre nodos, para distintos valores de  $\alpha_1$ , cuando el diseño es hexagonal.

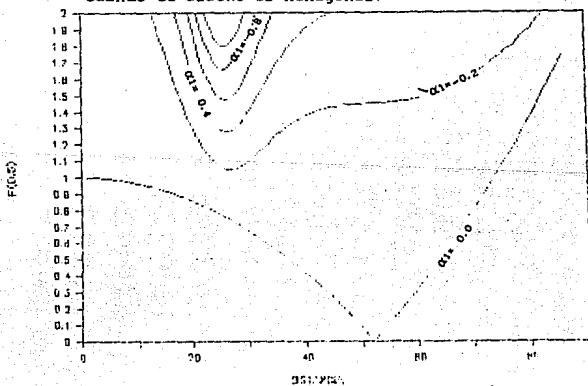


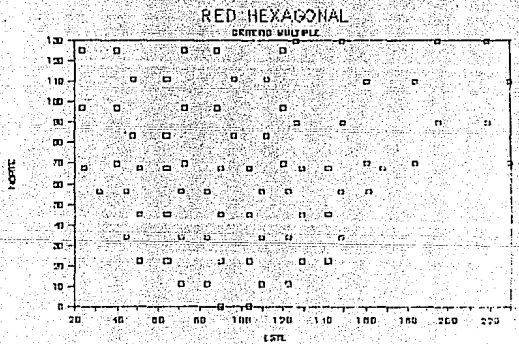
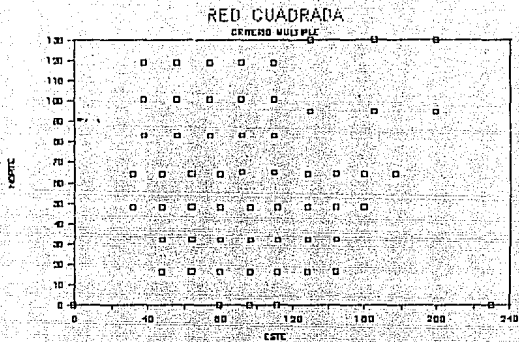
Figura 9. Comportamiento de la función  $CF(\sigma, T)$  con respecto a la distancia entre nodos, para distintos valores de  $\alpha_1$ , cuando el diseño es triangular.

sobremuestrear otras, aunque este problema puede corregirse cambiando los valores de los coeficientes  $\beta_i$  de la expresión para  $CF(\sigma, T)$ , dando mayor peso a las variables que se consideren más importantes, pero para el objetivo de la red que aquí se pretende no existe preferencia por ninguna. Las redes de monitoreo con cada diseño bajo este criterio de construcción se presentan en la figura 10.

Las figuras 11 a 13 muestran el comportamiento de la varianza de kriging para *T. atlanticus* con los tres diseños en cada uno de los tres estratos, así como con la red original para efectos de comparación, y solo se presenta esta especie por ser la que menor reducción de la varianza presentaba en todos los casos.

En general estas son mayores que las que se tienen con la red original pero esto se compensa por el menor número de puntos necesarios para cubrir cada estrato. Cabe aclarar que la varianza es mayor en las orillas debido al efecto de borde, pues para estimar en estos sitios se dispone de menos observaciones. Al comparar los resultados obtenidos con este criterio y con el uniespecífico, se ve que se obtiene un menor número de puntos de muestreo aunque la varianza de kriging en el caso del atún aleta negra solo se reduce a un 80% de su valor máximo, resultando en estimaciones poco precisas de la abundancia de esta especie comparadas con las que se obtendrían utilizando el primer método.

A diferencia del método uniespecífico, en este todos los diseños sugieren que el muestreo sea más intenso en la parte sur del Golfo de México, lo cual tiene cierto sentido intuitivo, pues es en esta zona donde desemboca la mayor parte de los ríos de esta cuenca y se espera a priori que la dinámica del pláncton en esta zona sea mayor, requiriéndose un muestreo más denso para detectarla (Smith y Richardson, 1977), sin perder de vista que las redes aquí diseñadas tienen como propósito la estimación del total de larvas de atún de las especies mencionadas y no la detección de áreas de desove, problema que se enfrenta de otra manera (Gilbert, 1988).



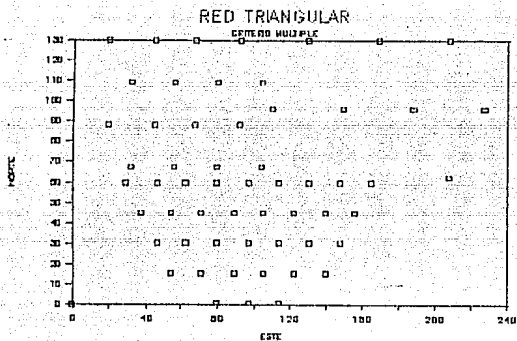


Figura 10. Redes obtenidas utilizando el criterio múltiple:  
a) Red cuadrada; b) red hexagonal; c) red triangular.

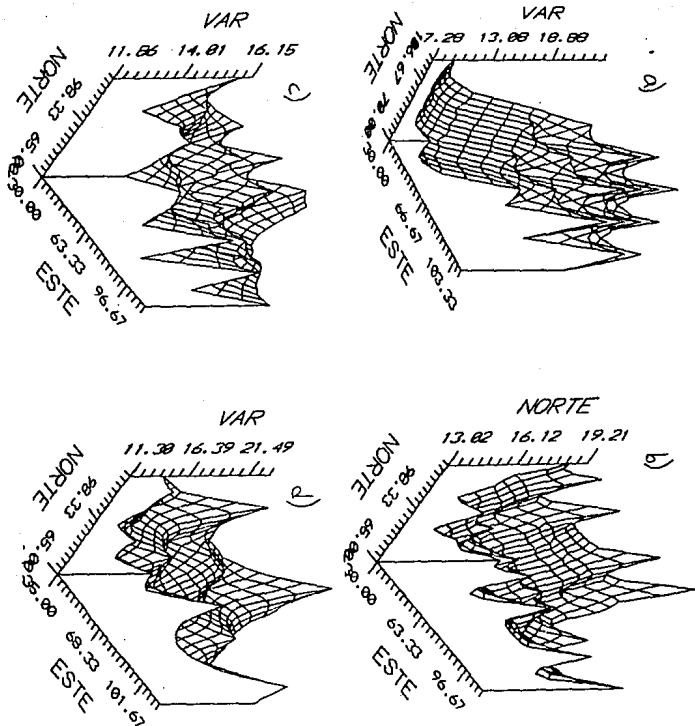


Figura 11. Comportamiento de la varianza de kriging para *T. atlanticus* con la disposición original de los puntos de muestreo y con las redes obtenidas en el estrato 1, usando el criterio múltiple: a) Red original; b) red cuadrada; c) red triangular; d) red hexagonal.



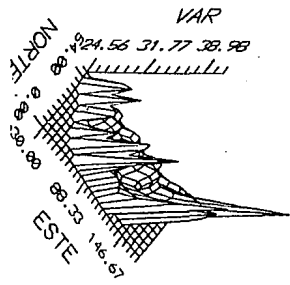
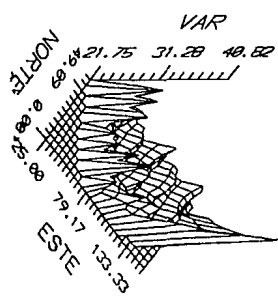
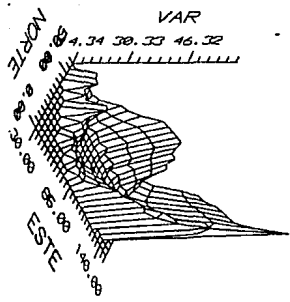
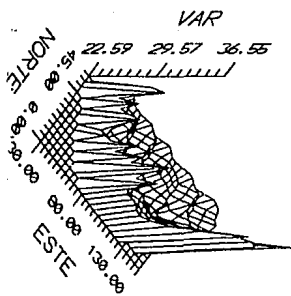


Figura 12. Comportamiento de la varianza de kriging para *T. atlanticus* con la disposición original de los puntos de muestreo y con las redes obtenidas en el estrato 2, usando el criterio múltiple: a) Red original; b) red cuadrada; c) red triangular; d) red hexagonal.

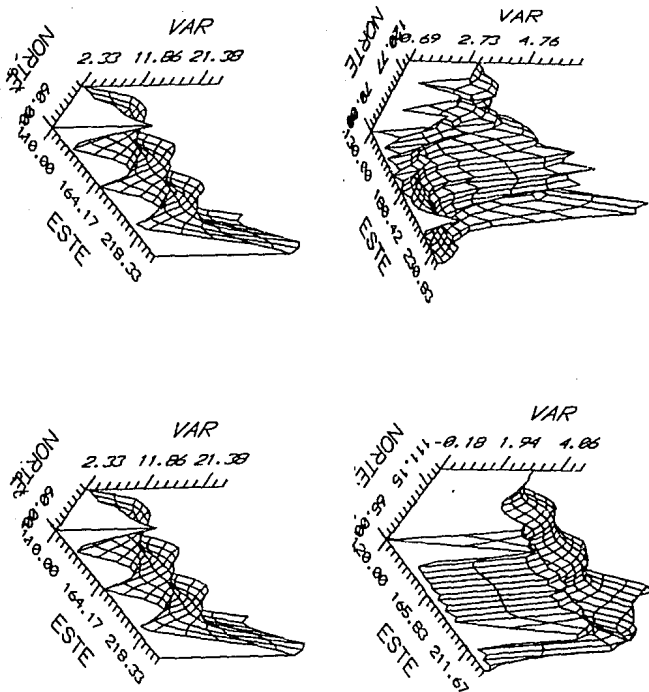


Figura 13. Comportamiento de la varianza de kriging para *T. atlanticus* con la disposición original de los puntos de muestreo y con las redes obtenidas en el estrato 3, usando el criterio múltiple: a) Red original; b) red cuadrada; c) red triangular; d) red hexagonal.

Al utilizar el criterio uniespecifico y el de objetivo múltiple se obtienen redes con menor cantidad de nodos que la que actualmente se tiene implementada, reducciones que van desde un 51 hasta un 68% en este rubro. No es muy fácil evaluar el impacto real que esto tiene sobre el costo total del muestreo porque existen gastos que tienen que hacerse independientemente del número de puntos de muestreo, tal como seguros para la nave y el personal a bordo, pago de tarifas por uso de muelles, etc. Sin embargo, considerando que la toma de datos en cada punto lleva alrededor de tres horas, por cada ocho puntos que se eliminan se estará ahorrando un día de crucero, lo que traducido al número de nodos de las redes aquí obtenidas significan ahorros de entre cuatro y cinco días de crucero, que entre pago de salarios al personal y avituallamiento representan un ahorro nada despreciable de dinero, aparte del menor trabajo que se tendrá en laboratorio al procesarse menos muestras, quedando por verse si en realidad el monto de estos justifica la pérdida de precisión que significó el aumentar la distancia entre puntos.

## CONCLUSIONES.

Del análisis del problema tratado en este trabajo y de los resultados obtenidos al explorar las alternativas de solución, se puede concluir lo siguiente:

El diseño de redes de monitoreo utilizando la teoría de variables regionalizadas presenta la ventaja sobre el enfoque clásico, de considerar la existencia de correlación espacial en los datos, permitiendo no solo encontrar el número necesario de puntos de muestreo sino también la posición y geometría más adecuada.

Los métodos de diseño que se basan en los cambios de posición de los puntos muestrales resultan adecuados al problema que aquí se trató debido, sobre todo, a que son mucho más fáciles de programar y a que permiten la evaluación de diferentes disposiciones de las estaciones de muestreo.

Dentro de las redes regulares, las más eficientes son la triangular cuando se tiene un número preespecificado de puntos y la cuadrada cuando se desea una red cuya varianza máxima de estimación no rebase un valor predefinido. Para el caso específico de la red de monitoreo oceanográfica de este trabajo, la sugerencia es tomar una red triangular pues para el caso uniespecífico es la que da una menor cantidad de nodos y para el caso del diseño de criterio múltiple es la que da una menor varianza.

Los datos utilizados en el presente trabajo corresponden al Verano y no se puede esperar que las condiciones en que se

obtuvieron se mantengan constantes en el tiempo, pero el método de construcción de la red de muestreo aquí presentado bien puede aplicarse al resto de las estaciones del año para obtener una red de monitoreo para primavera, etc., aunque otra posible solución es calcular un variograma promedio anual y en base a él diseñar una red única.

#### LITERATURA CITADA

- Bogardi, I., A. Bardossy and L. Duckstein. 1985. Multicriterion network design using geostatistics. Water Res. Research, 21(2):199-208.
- Burgess, T.M., and A.B. Mc Bartney. 1981. Optimal interpolation and isarithmic mapping of soil properties IV. Sampling strategy. Jour. Soil Sci., 32:643-659.
- Burgess, T.M., and R. Webster. 1980. Optimal interpolation and isarithmic mapping of soil properties II. Block kriging. Jour. Soil Sci., 31:333-341.
- Cressie, N., C.A. Gotway and M. Grondona. 1990. Spatial prediction from networks. Chemom. and Intell. Lab. Sys., 7:251-271.
- Chica Olmo, M. 1988. Análisis Geostatístico en el Estudio de la Explotación de los Recursos Minerales. Tesis doctoral. Universidad de Granada. Granada, España.
- Christakos, G. 1984. On the problem of permissible covariance and variogram models. Water Res. Research, 20(2):251-265.
- David, M. 1977. Geostatistical Ore Reserve Estimation. Elsevier, Amsterdam.
- Delhomme, J.P. 1979. Spatial variability and uncertainty in groundwater flow parameters: A geostatistical approach. Water Res. Research, 15(2):269-280.
- Englund E. and A. Sparks. 1988. GEO-EAS User's Guide. Environmental Protection Agency. Las Vegas, Nevada.

ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA

- Gambolati G. and G. Volpi. 1979. Groundwater contour mapping in Venice by stochastic interpolators 1. Theory. Water Res. Research, 15(2):281-290.
- Gilbert, R.O. 1987. Statistical Methods for Environmental Pollution Monitoring. Van Nostrand Reinhold, N.Y.
- Hughes, J.P. and D.P. Lettenmaier. 1981. Data requirements for kriging: Estimation and network design. Water Res. Research, 17(6):1641-1650.
- Issaks, E.H. and R.M. Srivastava. 1990. An Introduction to Applied Geostatistics. Oxford University Press, N.Y.
- Journel A.G. and Ch.J. Huijbregts. 1978. Mining Geostatistics. Academic Press, London.
- Mc Arthur, R.D. 1987. An evaluation of sample designs for estimating a locally concentrated pollutant. Commun. Statist. Simula., 16(3):735-759.
- Mc Bartney A.B. and R. Webster. 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables II. Program and examples. Comp. and Geosciences, 7(4):336-365.
- Mc Bartney, A.B., R. Webster and T.M. Burgess. 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables I. Theory and method. Comp. and Geosciences, 7(4):331- 334.
- Mc Cuen R.H. and W.M. Snyder. 1986. Hydrologic Modelling: Statistical Methods and Applications. Prentice Hall, N.Y.

- Myers, D.E. 1987. On variogram estimation. Proceeding of the First Int. Conf. on Statistical Computing, Cesme, Izmin, Turkey.
- Olea, R. 1984. Sampling design optimization for spatial functions. Math. Geol.,16(4):369-392.
- Olvera, R.M., J.L. Cerecedo y G. A. Compean. 1988. Distribución de larvas de tónidos en el Golfo de México y Mar Caribe: Abundancia y biomasa de tres especies en la Zona Económica Exclusiva. Ciencia Pesquera,6:119-140.
- Ripley, B.D. 1981. Spatial Statistics. Wiley, N.Y.
- Rodríguez-Iturbe, I., and J.M. Mejía. 1974. The design of rainfall networks in time and space. Water Res. Research,10(4):713-728.
- Russo, D. 1984. Design of an optimal sampling network for estimating the variogram. Soil. Sci. Soc. Am. 48:708-716.
- Samper, F.J. y Carrera J. 1990. Geostatística. Aplicaciones a la Hidrología Subterránea. CIMNE, Barcelona, España.
- Shapiro A. and J.D. Botha. Variogram fitting with a general class of conditionally nonnegative definite functions. Comp. Stat. and data analysis.11:87-96.
- Smith, P.E. and S.L. Richardson. 1977. Standard techniques for pelagic fish egg and larva surveys. FAO Fisheries Tech. Paper No.175.
- Szidarovszky, F. 1983a. Multiobjective observation network designs for regionalized variables. Int. Jour. of Min. Eng.,1:331-342.



- Szidarovszky, F. 1983b. Optimal observation network in geostatistics and underground hydrology. App. Math. Modelling,7:25-32.
- Warrick A.W., D.E. Myers and D.R. Nielsen. 1986. Geostatistical methods applied to soil science. pp 53-82. En Methods for soil analysis, Part 1. Physical and mineralogical methods. Monograph No. 9. 2nd Edition. Am. Soc. of America, Soil Sci. Soc. of America. Madison, WI. USA.
- Yañez, A. 1986. Ecología de la Zona Costera. AGT Editor, México.
- Yftanis, G.A., and G.T. Flatman. 1988. On sampling nonstationary spatial autocorrelated data. Comp. and Geosciences, 14(5):667-686
- Yftanis, G.A., G.T. Flatman and J.V. Behar. 1987. Efficiency of kriging estimation for square, triangular and hexagonal grids. Math. Geol.,19(3):183-205.
- Zimmerman, D.L. and M.B. Zimmerman. 1991. A comparison of spatial semivariogram estimators and corresponding ordinary kriging predictors. Technometrics,33(1):77-91.