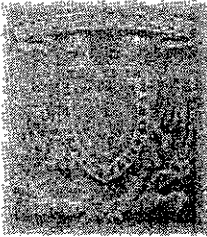


03061

1



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
Unidad Académica de los Ciclos Profesional y de Posgrado
Instituto de Investigación en Matemáticas Aplicadas y
en Sistemas

ANALISIS GEOESTADISTICO DE LA DIFUSION DE
AGUAS DE DESECHO EN ALTAMIRA, TAMAULIPAS

TESIS

2017

Que para obtener el grado de
MAESTRIA EN ESTADISTICA E
INVESTIGACION DE OPERACIONES
Presenta

CLAUDIA LARA PEREZ SOTO

Mayo, 2001



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 esta 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.

A MARIO

que dentro de las muchas
cosas que me has enseñado ha
sido respetar a la naturaleza
y amar la vida intensamente.

A Eduardo, Diego, Claudio, Irene y Andrés.

Agradecimientos

A Belem y Carlos a quienes con mucho esfuerzo y entusiasmo me ayudaron a elaborar y terminar este trabajo, no dejándome desistir.

A Alberto, Dr. Ignacio Méndez y Salvador por sus comentarios a este trabajo.

A Dennis Hurley y a ECA por todo el apoyo y tiempo brindados para concluir esta tesis.

A Rubén Hernández quien siempre ha sido una persona muy importante para mí.

A Martha, Arturo y Sergio que me contagiaron de su entusiasmo y de ganas para terminar.

A los amigos que siempre han estado conmigo: Ana, Lucy, Fernando, Luis Alfonso, Martín y Víctor.

A mi familia Lara Pérez Soto porque a pesar de las adversidades seguimos siendo una sola.

A Klaus por enseñarme la belleza de una relación llena de amor, paciencia y entrega.

Índice

Resumen	I
CAPITULO 1: Planteamiento del Problema.....	1
1.1 Antecedentes	1
1.2 Objetivos.....	4
1.3 Area de Estudio.....	5
CAPITULO 2: Teoría De Variables Regionalizadas	9
2.1 Introducción.....	9
2.2 Teoría de Variables Regionalizadas	10
2.2.1 Hipótesis de Estacionariedad.....	12
2.2.2 Propiedades de la Covarianza y del Variograma.	14
2.2.3 Análisis Gráfico del Variograma	16
2.2.4 Modelos Teóricos del Variograma	20
2.4 Anisotropía.....	23
2.5 Kriging.	25
2.5.1 Kriging Ordinario.....	26
2.5.2 Kriging como un Modelo de Tendencia	30
2.5.3 Median Polish	34
2.5.4 Validación Cruzada	37
2.5.5 Cokriging.....	38
CAPÍTULO 3: Análisis Geoestadístico.....	41
3.1 Generalidades.....	41
3.2 Características Generales de la Zona Muestreada durante las Diferentes Épocas del Año.	42

3.3 Análisis Exploratorio.....	43
3.3.1 Transformaciones	58
3.4 Análisis Geoestadístico en \mathcal{R}^2	63
3.4.1 Descripción de Datos en 2 Dimensiones.....	63
3.4.2 Descripción General del Análisis Estructural.....	69
3.4.3 Variogramas.....	74
3.4.4 Kriging.....	81
3.4.5 Validación Cruzada.....	101
3.5 Análisis Geoestadístico en \mathcal{R}^3	110
3.5.1 Descripción General del Análisis Estructural.....	110
3.5.3 Kriging.....	118
3.5.4 Validación Cruzada.....	128
3.6 Discusión.....	157
3.6.1 Antecedentes.....	157
3.6.2 Resultados.....	159
Conclusiones.....	161
Comentarios Adicionales.....	162
Apéndice 1 Especificaciones en Gslib.....	163
A.1 Especificaciones del Modelo para Variograma.....	163
A.2. Estrategias de Búsqueda.....	166
Bibliografía.....	169

Resumen

La naturaleza es complicada. En cualquier sistema ecológico, los patrones que se observan presentan un número de interacciones de procesos ecológicos que varían en función del tiempo y del espacio. La distribución, la intensidad o el desarrollo de un fenómeno son y se ven afectados por muchos factores simultáneos de diferentes características. La multiplicidad y la interacción de factores causales o casuales hacen que estos fenómenos sean aún más complicados. En la última década los estudios de impacto ambiental han tomado mucha importancia debido a que los sistemas ecológicos son una fuente de ingreso muy fuerte para el país, no solo económico o social sino también por la preservación de los seres vivos que habitamos en ella.

Dentro de los muchos estudios que se realizan en impacto ambiental se encuentran los relacionados a contaminantes de diversas clases, que pueden ser de aire, tierra o agua y la forma de analizarlos podría ser similar pues depende mucho de las variables o factores a observar. Muchas veces lo que interesa en estos estudios es ver como son afectadas las zonas aledañas a la fuente de contaminación y la situación de distribución espacial del fenómeno resulta de gran importancia para evaluar y conservar el ecosistema.

El caso que se presenta en este trabajo tiene que ver con de la industria química en las zonas costeras porque gran parte de los desperdicios que se generan van a dar al mar afectándolo de manera importante si no se hace de manera adecuada.

En las ciencias de la tierra la información recolectada comunmente se distribuye en el espacio y/o tiempo como lo son, en este caso, las aguas de desecho de la planta química y es así como el interés de la localización y del tiempo sean medidos y contemplados dentro del análisis. La Geoestadística provee un conjunto de herramientas estadísticas que incorporan

las coordenadas espaciales y temporales de las observaciones en el procesamiento de la información y que serán explicadas a lo largo de este trabajo como una propuesta metodológica para la evaluación de las zonas afectadas por diferentes causas.

En el Capítulo 1 se presentan los antecedentes del problema a trabajar junto con los objetivos del proyecto global.

En el Capítulo 2 se resume la teoría de variables regionalizadas que son la parte fundamental de la Geoestadística, la hipótesis de estacionariedad, los tipos y características del variograma y los tipos de estimación o predicción mas comunes de kriging.

En el Capítulo 3 se presentan los resultados de manera descriptiva, en dos y tres dimensiones y la validación de los modelos elegidos.

Por último se presentan las conclusiones de trabajo.

CAPITULO 1: Planteamiento del Problema

1.1 Antecedentes

En las últimas décadas, la región sur del estado de Tamaulipas ha registrado un crecimiento acelerado al instalarse una serie de plantas industriales en el Municipio de Altamira como parte de un programa nacional de desarrollo industrial para incrementar la capacidad productiva del país. No obstante, en muchos casos el establecimiento de estas plantas ha contribuido paulatinamente al deterioro ambiental de la zona debido a la falta de un tratamiento efectivo de sus aguas de desecho y otros desperdicios industriales.

Algunas plantas de importancia en la costa de Altamira son Humex, Novaquim, Petrocel y Pigmentos y Productos Químicos (PPQ), que difieren entre sí por la especificidad de sus procesos y productos.

La Planta de PPQ, perteneciente a Du Pont, S.A. se encuentra operando en la zona desde 1960. En ella se produce un pigmento blanco a partir de Bióxido de Titanio. El proceso empleado originalmente para obtener este pigmento se basaba en la extracción con ácido sulfúrico y fue sustituido en 1976 por el cloruro de ilmenita - mineral de alto contenido de titanio y fierro -, cloro y coque de petróleo. El nuevo proceso consiste en la reacción del cloruro con el cloro; con la formación subsecuente de tetracloruros de titanio y fierro, entre otros compuestos. Durante el proceso, la planta extrae agua del río Tamesí con un gasto máximo proyectado de 3,500 galones por minuto (GPM). Antes del proceso industrial el agua se desinfecta, clarifica, filtra, suaviza y desmineraliza.

Las aguas de residuo que descarga la planta, como resultado del proceso de producción, son depuestas en el Golfo de México a través de una línea de aproximadamente 7.5 km de

longitud, que incluye un emisor y difusor submarino de 3 km de largo. En 1988, la Secretaría de Desarrollo Urbano y Ecología (SEDUE, actualmente Secretaría de Medio Ambiente y Recursos Naturales) estimó que el gasto residuo promedio de esta última empresa es de 69 lt/seg, aportando cerca del 34% de los desechos que se vierten al medio marino (SEDEMAR, 1992). De acuerdo con SEDUE, la tasa del efluente se estimó en $6.375 \text{ m}^3 / \text{min}$. Antes de su expulsión en el Golfo, se aplica un tratamiento de sedimentación, para la cual la planta tiene tres fosas. En el diseño de la planta se consideró al hierro como el elemento clave para ser diluido, en virtud de ser el más abundante en el efluente (Rangel, 1988).

Se sabe que la circulación de corrientes marinas afecta a las poblaciones biológicas, a las condiciones meteorológicas, al clima, a la distribución de las propiedades químicas en el agua, a los procesos sedimentarios y a la dispersión de los contaminantes (Capurro, 1972). Por esta razón se considera importante evaluar si la manera de dispersar los contaminantes alteran lo menos posible el ecosistema marino.

Las características particulares de la descarga de la planta son una fuerte acidez, con valores de pH de 1.0, una concentración de sólidos disueltos del orden de 145,000 partes por millón (ppm), la cual es superior a la de las descargas de las otras compañías aledañas, y una alta concentración de hierro total (26,600 ppm). Dichos compuestos al entrar en contacto con el agua marina se hidrolizan y el hierro en estado ferroso es oxidado por el oxígeno disuelto en el agua, a hierro trivalente, como hidróxido férrico. Este compuesto por su baja solubilidad permanece como coloide insoluble y provoca una coloración pardo-amarilla muy notoria (SEDEMAR, 1992).

En 1981, Du Pont realizó estudios referentes a la efectividad de dispersión del emisor y al efecto de la descarga en la flora y fauna regional. De este estudio se infirió que para que el agua residual alcance un valor de pH dentro del intervalo encontrado en el agua marina, se requiere una dilución superior a 2,000 X.

En 1985, se construyó una escollera de 2.5 km de longitud, 7 km al norte del difusor, la cual se ubicó a 19 km de la escollera del Puerto de Tampico y paralela a ésta. Esta estructura

modificó el comportamiento de las corrientes marinas en la costa de Tamaulipas. Bajo estas circunstancias y ante la necesidad de incrementar la dilución de la descarga a fin de cumplir con los requerimientos establecidos, la empresa Du Pont aplicó el modelo DKHPLM, desarrollado por Environmental Protection Agency, para realizar una evaluación de la dilución de la descarga.

El modelo DKHPLM aplicado indicó que un difusor con 80 boquillas de descarga horizontal de $1\frac{3}{8}$ pulgadas de diámetro y 6 metros (m) de separación entre cada una de éstas, proporcionaría una dilución de 2,300 X. La evaluación se realizó bajo los siguientes supuestos: una velocidad de la corriente de 15 cm/seg a una profundidad promedio de 10 m, una profundidad efectiva de 15 m en el área de descarga y un volumen de descarga de 3,200 GPM, condiciones que se consideró se presentaban a unos 3,000 m de la costa.

El proyecto ejecutivo para la instalación del nuevo difusor de mayor longitud y otras características, fue presentado por Du Pont ante SEDESOL en julio de 1991. La empresa procedió a construir el nuevo difusor a 3 km de la costa, e inició sus operaciones de descarga en septiembre del mismo año. La dependencia antes mencionada manifestó que la empresa debía llevar a cabo evaluaciones del ecosistema acuático en la zona de impacto de la descarga, a fin de conocer las características originales del ambiente y de determinar la influencia ejercida por el efluente descargado.

La empresa PPQ de Du Pont solicitó los servicios de la Secretaría de Marina, la cual, por medio de la Dirección General de Oceanografía Naval, a través de su Estación Oceanográfica de Tampico y de su Dirección Centro de Datos y Servicios Océánicos, realizó un estudio de caracterización ambiental.

A su vez, La Secretaria de Marina estableció un convenio de colaboración con el Centro Científico de IBM de México, para realizar el análisis geoestadístico de la zona, la representación gráfica por medio de modelos tridimensionales y el análisis espacial por medio de sensores remotos.

En este trabajo se presentan la metodología para la parte del análisis geoestadístico realizado en la zona. Los resultados presentados en el reporte acerca de este tema solo fueron de carácter informativo y descriptivo. La intención de esta tesis es completar el trabajo a fin de tener el análisis lo más completo posible.

1.2 Objetivos

Los Objetivos Generales del proyecto son:

i) Realizar la caracterización ambiental del ecosistema marino en la zona de influencia de la descarga del emisor submarino de PPQ en la Costa de Altamira, Tamaulipas y evaluar el efecto que produce dicha descarga en el área de estudio, con base en los resultados obtenidos en la determinación de parámetros físicos, químicos, geológicos y biológicos, tanto en el área de mayor influencia, como en puntos testigo.

ii) Representar gráficamente las características hidrológicas generales de una porción de la costa y de la zona marina de Altamira, Tamaulipas, mediante el análisis de imágenes de satélite, a partir de procesamiento digital y del uso de un paquete de visualización de datos científicos.

Como objetivo específico de este trabajo, se presenta una propuesta de análisis geoestadístico como una metodología adecuada para el proceso de los datos, dada la situación espacial de los mismos datos. Este trabajo forma parte del convenio de colaboración que tuvo el Centro Científico de IBM de México con la Secretaría de Marina y cuyo propósito es el de analizar los factores que son afectados por la descarga del difusor

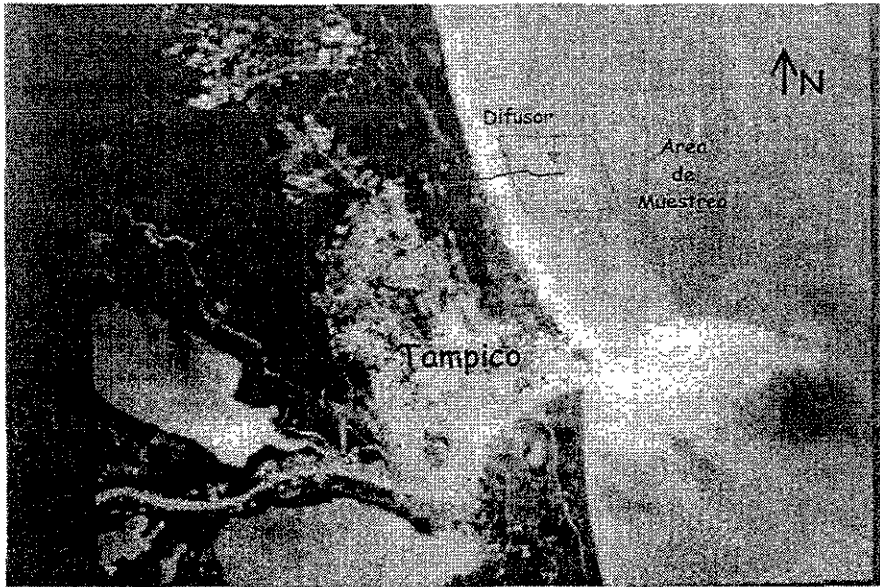


Figura 1.1 Imagen TM de Satélite Lansat 5, Octubre de 1992.

debido a la descarga de los contaminantes, en este caso fierro (Fe) y manganeso (Mn).

1.3 Area de Estudio

El área de estudio se encuentra ubicada en la costa Sureste del Estado de Tamaulipas, frente al municipio de Altamira, entre los $22^{\circ} 18' 36''$ y $22^{\circ} 22' 11''$ de Latitud Norte y los $97^{\circ} 46' 33''$ y $97^{\circ} 50' 08''$ de Longitud Oeste (Figura 1.1).

La red principal de localidades de muestreo abarcó un área de 25 km^2 y formó una retícula con distancias de 1 km entre cada punto, tomando como referencia el transecto central, ubicado en el tendido del nuevo difusor. Las estaciones de muestreo se distribuyeron en

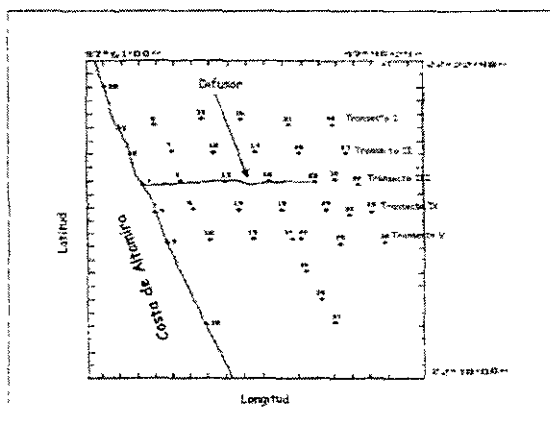


Figura 1.2 Distribución de las principales estaciones de muestreo

cinco transectos perpendiculares a la línea de costa de la siguiente manera: transecto I, estaciones 1, 6, 11, 16, 21 y 26; transecto II estaciones 2, 7, 12, 17, 22 y 27; transecto III estaciones 3, 8, 13, 18, 23 y 28; transecto IV estaciones 4, 9, 14, 19, 24 y 29; transecto V estaciones 5, 10, 20, 25 y 30. Fue trazado un sexto transecto paralelo a la costa, a la mitad de la longitud total del difusor, cuyas estaciones se localizaron a 0.05, 0.10, 0.5, 1, 2, 3, 4 y 5 km de éste, tanto hacia el Norte, como al Sur; la posición de este transecto varió en función de la dirección predominante de las corrientes marinas. En mayo fue hacia el sur, en julio al norte y en noviembre en ambas direcciones, con la finalidad de cubrir representativamente el área de influencia del difusor, según se señala en la Figura 1.2. Las estaciones 1, 2, 3, 4 y 5, así como aquellas ubicadas al norte y sur sobre la costa, se marcaron de forma permanente y a partir de ellas el posicionamiento se efectuó mediante triangulación empleando teodolitos. El posicionamiento marino se efectuó considerando las señales de las estaciones de tierra, realizando triangulación hasta a 1 km de la costa y con el empleo de un Loran.

Las variables que se consideraron para caracterizar el ecosistema marino fueron las siguientes:

- i) Biológicas: Nutrientes (Nitratos, Nitritos, Nitrógeno Amoniacal, Fósforo Total)
- ii) Físicas: Temperatura, Condiciones Meteorológicas (secas, lluvias y nortes)

iii) Químicas: Hierro (Fe), Manganeso (Mn), Cromo Total (Cr^{total}), Cromo Hexavalente (Cr^6), Aluminio (Al), Titanio (Ti), Vanadio (V), Oxígeno Disuelto, pH, Salinidad.

Se realizaron tres campañas de muestreo durante las épocas de secas, lluvias y nortes, en los meses de mayo, julio y noviembre de 1992 respectivamente. Estas se desarrollaron en dos embarcaciones con motor fuera de borda, equipadas una con ecosonda y la otra con un posicionador Loran.

La obtención de muestras de agua para el análisis de los parámetros de la columna de agua, se realizó mediante el empleo de botellas Niskin y Van Dorn, equipadas con termómetros reversibles. Las calas se efectuaron a 5 niveles: 0, 4, 8, 12 y 16 m, determinados a partir de la profundidad de cada punto.

La medición de las características hidrológicas se realizó de la siguiente manera: la temperatura del agua con termómetros reversibles; la salinidad con un salinómetro de inducción; la densidad por el método de polinomio propuesto por Cox et al. (1970); el pH con un potenciómetro; el oxígeno disuelto se determinó por el método de Winkler modificado; los nitritos aplicando el método de Shin; los nitratos con la técnica de reducción de nitratos a nitritos por medio de una columna de cadmio-cobre; el nitrógeno amoniacal a partir de la técnica de fenol hipoclorito descrita en Strickland y Parsons (1972) y el fósforo total por medio de una digestión ácida de acuerdo a APHA (1985).

Los sólidos sedimentables se determinaron utilizando conos Imhoff y los no sedimentables por los métodos descritos por APHA (1985).

En la determinación de metales pesados en la columna de agua (Fe, Mn, Cr^{total} , Cr^6 y Al) se utilizó la técnica de quelación-extracción con metil isobutil cetona y pirrolidín-ditio-carbamato de amonio (APDC) y a partir de 8-hidroxiquinoleína y APDC. Se empleó una matriz de agua marina para la preparación de los estándares y se tomó la lectura de la concentración del extracto en un espectrofotómetro de absorción atómica varían. El cromo hexavalente se determinó aplicando el método colorimétrico de la difenilcarbazida. El vanadio y el titanio se

determinaron mediante la adición de estándares y la lectura en el espectrofotómetro de absorción atómica.

Una vez recolectados los datos para cada uno de los muestreos, se llevó a cabo el análisis geoestadístico y la visualización de los datos para cada una de las variables.

CAPITULO 2: Teoría de Variables Regionalizadas

2.1 Introducción

Un fenómeno natural puede ser caracterizado por el comportamiento en el espacio de una o varias partes que lo conforman. Por ejemplo, la cantidad de minerales que contienen 10 m² de tierra, en este caso el fenómeno a estudiar es la riqueza de minerales que contiene ese pedazo de tierra. El cuantificar cada uno de los minerales que la constituyen es una forma de contar cuánto y qué contiene e irá variando de acuerdo al lugar de donde se obtengan las mediciones.

Existen áreas en donde este tipo de problemas se estudian de manera cotidiana como por ejemplo Geología, Astronomía y Ecología. En ellas se ha creado metodología para analizar la información desde diferentes puntos de vista. Uno de estos puntos de vista es el estadístico espacial, en donde se considera a la o las variables en estudio como variables aleatorias cuyos valores dependen de su posición tanto en espacio como en el tiempo. Al estudio de dichos problemas se les denomina modelos espaciales o datos espaciales y tienen una forma muy particular de ser tratados estadísticamente.

Una variable regionalizada es una variable aleatoria que toma diferentes valores de acuerdo a su posición dentro de alguna región. Por ejemplo, en la estimación de una reserva mineral, la mina es la región y el grado, densidad mineral o cualquier otra medida es una variable regionalizada (Royle, 1980).

El término de variable regionalizada fue escogido por Matheron en 1965, para enfatizar las propiedades particulares de éstas, ya que caracterizan fenómenos que se distribuyen en el espacio y poseen una estructura llamada regionalización.

La Geoestadística es la parte de la estadística que se dedica a estudiar a este tipo de variables, siendo su base conceptual la Teoría de Variables Regionalizadas (TVR). El objetivo principal de dicha teoría, es la resolución práctica de problemas de estimación y simulación de fenómenos naturales, aplicando herramientas probabilísticas adecuadas. Propone un conjunto de modelos probabilísticos y métodos prácticos, entre los que hay que saber elegir aquellos que en cada caso dan la mejor representación de la variable, así como resultados confiables.

2.2 Teoría de Variables Regionalizadas

La teoría de las variables regionalizadas, como su nombre lo indica, es la fundamentación matemática de este tipo de variables, pero antes de definir qué es una variable regionalizada matemáticamente hablando, es necesario definir qué es una variable aleatoria.

Una variable aleatoria es aquella que toma valores numéricos de acuerdo a una función de distribución. Si se considera el rango de valores que toma una variable z en el punto x_i en una área delimitada, este rango $z(x_i)$ puede ser considerado como el valor que toma la variable aleatoria.

Desde el punto de vista matemático, una variable regionalizada es una función $f(x)$ la cual toma valor en cada punto de coordenadas $x \in \mathfrak{R}^k$, $k \geq 1$, es decir, depende de su posición espacial.

La variable regionalizada posee 2 componentes básicos:

i) Un componente m local y determinístico que representa la tendencia global de z .

ii) y una componente aleatoria e que representa la dispersión natural alrededor de m .

Con estos dos componentes la variabilidad espacial queda representada de manera sencilla. Esto último obedece a una de las muchas interpretaciones probabilísticas de una función aleatoria.

Una función aleatoria representa en muchos aspectos a una variable regionalizada, es decir:

i) localmente $Z[x_i]$ es una variable aleatoria.

ii) $Z[x]$ es una función aleatoria en el sentido de que para cada par de puntos x_i y $x_i + h$, las correspondientes variables aleatorias $Z[x_i]$ y $Z[x_i + h]$ están relacionadas por un patrón de correlación en su estructura espacial de variable regionalizada $Z[x]$.

La interpretación probabilística de la variable regionalizada $z[x_i]$ tiene sentido operativo cuando es posible inferir todo o parte de la ley de probabilidad que la define.

Para poder aplicar los resultados de la TVR, es necesario conocer parte de la ley de distribución de la función aleatoria $Z[x]$. Dos caminos diferentes pueden seguirse:

i) Si se conoce un número suficiente de realizaciones de la función aleatoria $Z[x]$, se puede entonces estimar las características de dichas funciones con ayuda de las medias obtenidas sobre el conjunto de las diferentes realizaciones.

ii) Si se conoce una única realización, es decir que solamente se cuenta con un único conjunto de observaciones con la idea de obtener el mayor provecho de ello es necesario establecer supuestos.

No es posible adoptar el primer punto, puesto que en la práctica no se cuenta con un número suficiente de realizaciones, por lo que es mejor considerar la segunda opción e introducir una serie de hipótesis restrictivas, que permitan encontrar una solución a la inferencia estadística. Estas suposiciones teóricas involucran un cierto grado de homogeneidad espacial y son conocidas como Hipótesis de Estacionariedad (Journel & Huijbregts, 1991).

2.2.1 Hipótesis de Estacionariedad

Estacionariedad Estricta: Una función aleatoria se dice estacionaria, en el sentido estricto, cuando sus distribuciones de dimensión finita son invariantes ante traslaciones, es decir, la distribución de $(Z[x_1], \dots, Z[x_k])$ y la de $(Z[x_1 + h], \dots, Z[x_k + h])$ son idénticas para toda $h \in \mathfrak{R}^k$.

Por el tipo de aplicaciones, es necesario considerar un tipo de estacionariedad más débil:

Estacionariedad de Segundo Orden: Una función aleatoria es estacionaria de segundo orden si

i) La esperanza de $Z[x]$ no depende de x , esto es

$$E[Z[x]] = m \quad \forall x$$

ii) La covarianza entre $Z[x+h]$ y $Z[x]$ solo depende de h por tanto

$$C(h) = E[Z[x+h]Z[x]] - m^2$$

Esta propiedad implica varianza constante, es decir,

$$\text{Var}[Z[x]] = E\{[Z[x] - m]^2\} = C(0) \quad \forall x$$

y estacionaria en el llamado variograma definido como

$$\gamma(h) = \frac{1}{2} \text{Var}[Z[x+h] - Z[x]] = C(0) - C(h) \quad \forall x.$$

De igual forma, la covarianza y el variograma resultan ser equivalentes para caracterizar la autocorrelación entre dos variables, por lo que se puede definir el correlograma como

$$\rho(h) = \frac{C(h)}{C(0)} = 1 - \frac{\gamma(h)}{C(0)}$$

Hipótesis Intrínseca: Una función aleatoria $Z[x]$ se dice que es intrínseca cuando:

i) $E[Z[x]] = m \quad \forall x$

ii) $\text{Var}[Z[x+h] - Z[x]] = E\{[Z[x+h] - Z[x]]^2\} = 2\gamma(h) \quad \forall x.$

En la práctica, ocurre que el concepto de estacionariedad está ligado a la distancia de un punto a otro, de forma tal que se refleja un comportamiento no estacionario, pero puede hacerse a distancias pequeñas, es decir los incrementos son estacionarios de segundo orden y por tanto

$$i) \quad \Rightarrow i') E[Z(x) - Z(x+h)] = m(h)$$

por ello se tiene la siguiente hipótesis:

Quasi-Estacionariedad: La función estructural, covarianza o variograma, se utiliza únicamente en distancias limitadas $|h| \leq b$, donde b representa el límite del diámetro de vecindad de estimación. Por otro lado b puede ser la extensión de la zona homogénea y las variables $z(x_k)$ y $z(x_k + h)$ pueden ser consideradas como provenientes de diferentes regiones no homogéneas si $|h| > b$. En tales casos la función estructural se considera como una función que cumple la estacionariedad local.

Como la mayor parte de las hipótesis dependen del comportamiento de la covarianza y/o el variograma es necesario mencionar algunas propiedades de estos y las condiciones que deben cumplir.

2.2.2 Propiedades de la Covarianza y del Variograma.

Condiciones: Las aplicaciones en la Geoestadística son básicamente la estimación de la variable de interés en un punto no muestreado o en un bloque, a partir de la información que se conoce. Dicha estimación se hace en base a combinaciones lineales de la variable observada. Sea $Z(x)$ una función aleatoria estacionaria y sea Z^* una combinación lineal finita, es decir

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

Esta combinación lineal es una función aleatoria y su varianza está dada por

$$V(Z^*) = \sum_{ij} \lambda_i \lambda_j C(x_i - x_j).$$

Para que $V(Z^*)$ sea no negativa se necesita que la función de covarianza sea una función definida positiva. Por otro lado se tiene la siguiente relación:

$$V(Z^*) = C(0) \sum_i \lambda_i \sum_j \lambda_j - \sum_i \sum_j \lambda_i \lambda_j \gamma(x_i - x_j).$$

Si $C(0) = V(x)$, entonces se aplica la hipótesis de segundo orden. Si $C(0)$ no existe, se supondrá la hipótesis intrínseca y la varianza estará definida a través de la condición

$$\sum_i \lambda_i = 0$$

de aquí se implica que

$$V(Z^*) = - \sum_i \sum_j \lambda_i \lambda_j \gamma(x_i - x_j) > 0$$

Es necesario entonces verificar que el variograma sea una función definida positiva.

Propiedades:

El carácter de definida positiva, establece las siguientes propiedades de la función covarianza:

$$i) C(0) = V\{z(x)\} \geq 0$$

$$ii) C(h) = C(-h)$$

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

La definición de variograma como la varianza de los incrementos establece las siguientes propiedades:

$$i) \gamma(0) = 0$$

$$ii) \gamma(h) = \gamma(-h) \geq 0$$

2.2.3 Análisis Gráfico del Variograma

El estudio del comportamiento gráfico del variograma es de interés primordial porque proporciona una descripción sintetizada de la variación del fenómeno, permite conocer la relación que existe en la estructura y la precisión de la estimación.

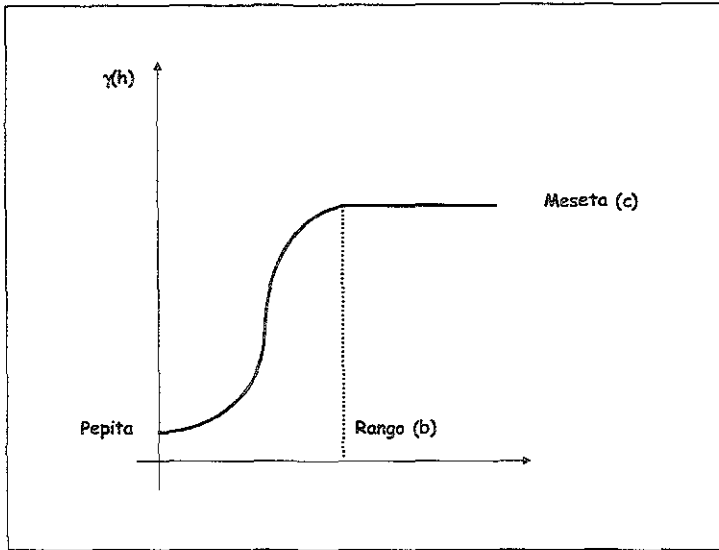


Figura 2.1 Partes del Variograma

Se considera que el variograma proporciona información acerca de 4 puntos (Figura 2.1) que son importantes:

1) Fenómeno de Transición

El semivariograma cuando deja de crecer a una distancia determinada (*rango*), se estabiliza alrededor de un valor límite, este valor límite se denomina *meseta*, la cual resulta ser la varianza a priori de la función aleatoria.

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

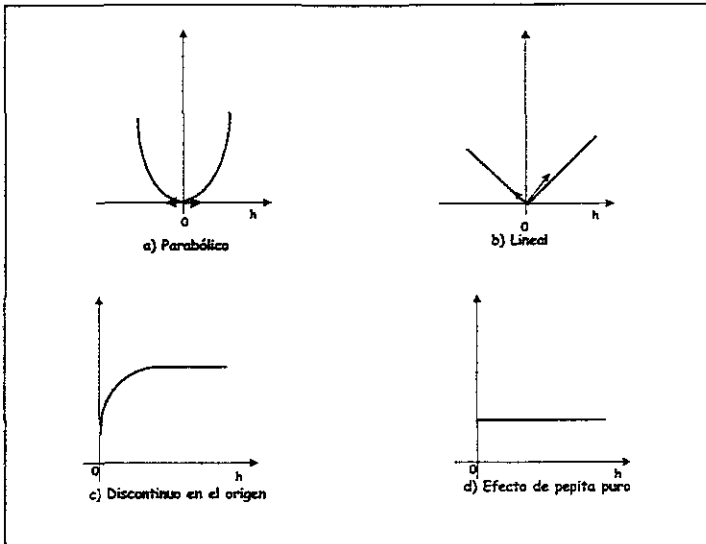


Figura2.2 Comportamiento del variograma cercano al origen.

Estos tipos de variogramas que se caracterizan por tener valor de meseta y rango, son denominados modelos de transición y corresponde no sólo a la función aleatoria que cumple la hipótesis intrínseca sino también es estacionaria de segundo orden.

ii) Zona de Influencia

En un fenómeno de transición, cualquier valor $Z(x)$ debe estar correlacionado con cualquier otro que caiga dentro de un radio b de x . Esta correlación decrece cuando la distancia entre los puntos crece. Esta distancia es el rango y la región comprendida se denomina zona de influencia de la variable regionalizada.

iii) Comportamiento en el Origen

Otros factores importantes de la estructura del fenómeno son su continuidad y regularidad espacial, las cuales están relacionadas con el comportamiento del variograma en el origen. Es frecuente distinguir, según el grado de continuidad decreciente, 4 tipos:

- a) Parabólico (Figura 2.2 (a))
- b) Lineal (Figura 2.2 (b))
- c) Discontinuo en el origen: Si $\gamma(h) \neq 0$ cuando $h \rightarrow 0$ quiere decir que existe una discontinuidad en el variograma. A esta discontinuidad se le denomina *efecto pepita* (Figura 2.2 (c)).
- d) Efecto Pepita Puro: El efecto pepita puro corresponde a la ausencia total de autocorrelación espacial entre variable (Figura 2.2 (d)).

iv) Comportamiento en el Infinito

Utilizando la definición de que $-\gamma(h)$ es una función positiva definida condicional a la zona de estudio, es posible demostrar que

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

esto es, el variograma necesariamente crece más lento que $\|h\|^2$. El hecho de que esto no suceda es incompatible con la hipótesis intrínseca y se debe a que existe una tendencia que hace que

$$E[Z(x)] = m(x)$$

lo que quiere decir que el valor esperado de $Z(x)$ depende de la posición de x .

2.2.4 Modelos Teóricos del Variograma

Para propósitos de estimación o simulación de variables regionalizadas, no es posible utilizar el variograma experimental de manera directa, por lo que se hace un ajuste previo del variograma empírico a un modelo teórico. Esto con la finalidad de que el variograma cumpla con las condiciones de consistencia matemática, básicamente con que sea una función positiva definida.

Los modelos básicos que se utilizan con mayor frecuencia se dividen en 2 tipos, los que alcanzan su valor de meseta y los que no. El valor de c_0 corresponde al valor de pepita, y los valores b_f , c_s , c_e , c_g y b_p son constantes positivas y a_s , a_e y a_g son los valores de rangos calculados para cada modelo. Estos se presentan a continuación:

i) Lineal (Figura 2.3(a)).

$$\gamma(h; \theta) = \begin{cases} c_0, & h = 0 \\ c_0 + b_f \|h\|, & h \neq 0 \end{cases}$$

para $\theta = (c_0, b_f)$, donde $c_0 \geq 0$ y $b_f \geq 0$.

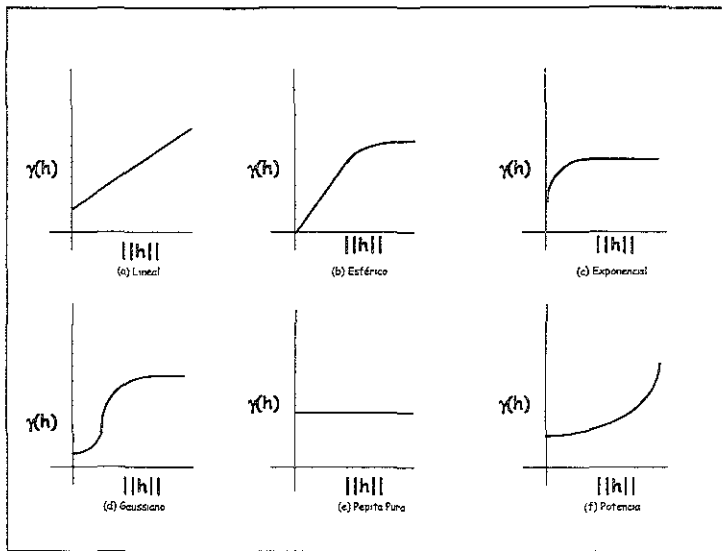


Figura 2.3 Variogramas más comunes

ii) Esférico (Figura 2.3(b))

$$\gamma(h; \theta) = \begin{cases} c_0, & h = 0 \\ c_0 + c_s \left\{ k_1 \left(\|h\|/a_s \right) - k_2 \left(\|h\|/a_s \right)^3 \right\}, & 0 < \|h\| \leq a_s \\ c_0 + c_s, & \|h\| \geq a_s \end{cases}$$

con $\theta = (c_0, c_s, a_s)$, donde $c_0 \geq 0$, $c_s \geq 0$ y $a_s \geq 0$. k_1 y k_2 son constantes positivas que en la práctica toman valores de $3/2$ y $1/2$ respectivamente.

iii) Exponencial (Figura 2.3(c)):

$$\gamma(h; \theta) = \begin{cases} c_0, & h = 0 \\ c_0 + c_e \{1 - \exp(-\|h\|/a_e)\}, & h \neq 0 \end{cases}$$

con $\theta = (c_0, c_e, a_e)$, donde $c_0 \geq 0$, $c_e \geq 0$ y $a_e \geq 0$.

iv) Gaussiano (Figura 2.3(d)):

$$\gamma(h; \theta) = \begin{cases} c_0, & h = 0 \\ c_0 + c_g \{1 - \exp(-k\|h\|^2/a_g^2)\}, & h \neq 0 \end{cases}$$

con $\theta = (c_0, c_g, a_g)$, donde $c_0 \geq 0$, $c_g \geq 0$ y $a_g \geq 0$. k toma el valor de 3 en la práctica

v) Pepita Puro (Figura 2.3(e)):

$$\gamma(h; \theta) = c_0, \quad \forall h$$

donde $c_0 > 0$.

vi) Potencia (Figura 2.3(f)):

$$\gamma(h; \theta) = \begin{cases} c_0, & h = 0 \\ c_0 + b_p \|h\|^\lambda, & h \neq 0 \end{cases}$$

con $\theta = (c_0, b_p, \lambda)$, donde $c_0 \geq 0$, $b_p > 0$ y $0 \leq \lambda < 2$.

Estos modelos son isotrópicos, es decir, sólo dependen de la distancia y no de la dirección.

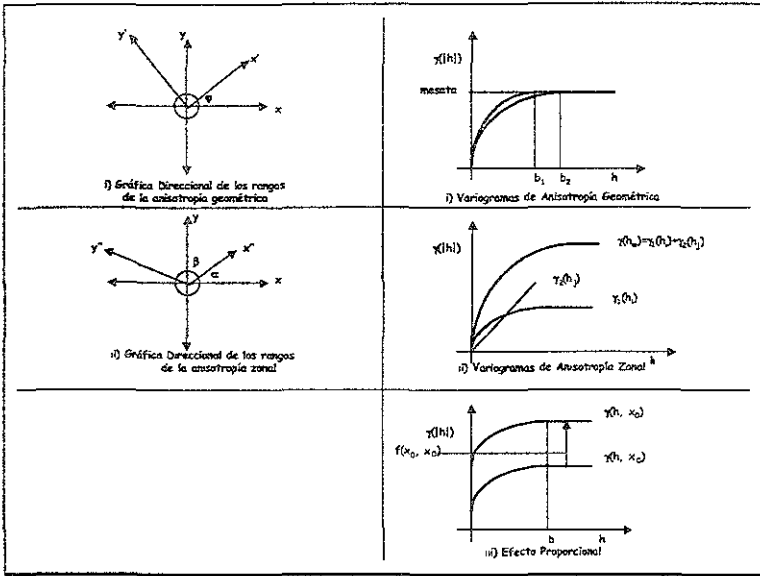


Figura 2.4 Tipos de Anisotropía.

2.4 Anisotropía

El fenómeno se considera de comportamiento anisotrópico cuando su variabilidad no es la misma para todas las direcciones. La función $\gamma(\|h\|, \alpha, \varphi)$ que caracteriza la variabilidad espacial depende ahora de los parámetros de dirección α y φ , estos parámetros corresponde a los valores de latitud y longitud.

En la práctica, la anisotropía de la función $\gamma(\|h\|, \alpha, \varphi)$ corresponde a la existencia de direcciones preferenciales del fenómeno en estudio. Estas direcciones algunas veces son conocidas con anterioridad y el variograma cuantifica sus respectivas variabilidades.

La anisotropía debe ser representada por una sola función variograma, es decir, debe ser llevada al caso isotrópico, lo que implica hacer una transformación lineal de las coordenadas

rectangulares o, ser expresadas de manera individual para cada dirección, según sea el caso.

Existen 3 tipos de anisotropía que son :

i) Anisotropía Geométrica.

Si se tienen r semivariogramas direccionales $\{\gamma_{\alpha_i}(h), i = 1, \dots, r\}$, y pueden ser representados por el mismo modelo de transición, con la misma meseta y n rangos con los cuales se puede formar una elipse para \mathfrak{R}^2 , o un elipsoide para \mathfrak{R}^3 , se dice que la anisotropía es de tipo geométrico (Figura 2.4(i)).

ii) Anisotropía de Zona.

Sea $\gamma(h)$ un modelo anidado que caracteriza la variabilidad en \mathfrak{R}^3 , es decir:

$$\gamma(h) = \sum_i \gamma_i(h).$$

Cada componente del modelo puede ser anisotrópico para h . Más aún, la anisotropía de $\gamma_i(h)$ puede ser diferente a la de $\gamma_j(h)$.

En pocas palabras la anisotropía de zona se define como una estructura anidada en la cual cada componente puede tener su propia anisotropía (Figura 2.4 (ii)).

iii) Efecto Proporcional.

Puede ocurrir que los variogramas sea el mismo modelo y que en sus respectivas mesetas sean diferentes pero presentando el mismo rango. A este tipo de comportamiento se le denomina efecto proporcional.

En la práctica se debe utilizar un único variograma $\gamma_0(h)$, a partir del cual se deducen los variogramas direccionales $\gamma_i(h)$ mediante una corrección lineal, es decir:

$$\gamma_i(h) = f(x_0, x'_0) \gamma_0(h)$$

en donde $f(\dots)$, por lo general es una función cuadrática de la media experimental (Figura 2.4(iii)) y (x_0, x'_0) es el vector de mesetas.

2.5 Kriging.

El propósito de contar con un modelo de variograma es el de predecir el valor de la variable en puntos donde no se tuvo medición. Esto se hace por medio de *kriging* que es un conjunto de métodos para hacer predicción espacial. El predictor de kriging se basa en la minimización del error cuadrático medio del estimador de kriging y por lo general, utiliza las propiedades de segundo orden.

En la predicción espacial se requiere construir una función $g(z(\cdot))$ que dependa de los datos. Esta función puede ser encontrada de diferentes maneras, pero todas estas tienen un objetivo común, minimizar el error cuadrático medio de $g(z(\cdot)) - z(\cdot)$.

El kriging es un interpolador exacto, es decir, si el punto x_0 se confunde con alguno de los puntos x_i de la muestra, el estimador $z^*(x_0)$ coincide el valor de $z(x_i)$; el kriging pasa por los puntos experimentales.

2.5.1 Kriging Ordinario

Este método tiene como característica principal obtener estimadores lineales insesgados y de varianza mínima para $g(z(\cdot)) - z(\cdot)$. Para la construcción del estimador se tiene que dada una muestra z_1, z_2, \dots, z_n , se desea predecir el valor de z_0 en el punto x_0 , es decir

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

en donde se pide que $E[z^*(x_0)] = z(x_0)$ por lo que se debe cumplir la siguiente condición sobre las ponderaciones

$$\sum_{i=1}^n \lambda_i = 1,$$

lo que implica que

$$E[z^*(x_0) - z(x_0)] = 0.$$

Además se pide que estas ponderaciones sean tales que minimicen el valor de la varianza del error de predicción, esto es, minimicen

$$\text{Var}[z^*(x_0) - z(x_0)] = \text{Cov}[z^*(x_0), z^*(x_0)] - 2\text{Cov}[z^*(x_0), z(x_0)] + \text{Cov}[z(x_0), z(x_0)]$$

Utilizando el método de Lagrange para encontrar los valores de las λ_i , se construye la función

$$L(\lambda, \mu) = \text{Var}[z^*(x_0) - z(x_0)] - 2\mu \left(\sum_{i=1}^n \lambda_i - 1 \right).$$

Esta función genera el siguiente sistema de ecuaciones denominadas ecuaciones de Kriging Ordinario

$$\sum_{j=1}^n \lambda_j \tilde{C}_{ij} + \mu = \tilde{C}_{i0} \quad \forall \quad i = 1, \dots, n$$

$$\sum_{i=1}^n \lambda_i = 1$$

en donde $\tilde{C}_{ij} = \tilde{Cov}[z(x_i), z(x_j)]$. La varianza de Kriging Ordinario se define como

$$\hat{\sigma}_{KO}^2 = \text{Var}[z^*(x_0)] - \left(\sum_{i=1}^n \lambda_i C_{i0} + \mu \right).$$

El sistema de ecuaciones de Kriging Ordinario también puede escribirse de manera matricial como

$$A\lambda = D$$

donde

$$C = \begin{bmatrix} \tilde{C}_{11} & \dots & \tilde{C}_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \tilde{C}_{ni} & \dots & \tilde{C}_{nn} & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix}, \quad \lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} \tilde{C}_{10} \\ \vdots \\ \tilde{C}_{n0} \\ 1 \end{bmatrix}$$

con el cual se tiene la siguiente solución para λ

$$\lambda = C^{-1}D$$

si C es invertible y la varianza de Kriging Ordinario puede escribirse como

$$\tilde{\sigma}_{KO}^2 = \text{Var}[z^*(x_0)] - \lambda D$$

Muchas veces lo que se quiere estimar es el valor promedio de $z(x)$ en una región o en un volumen y para obtener tal estimación se suele discretizar la región en varios cuadros o bloques y promediar los valores estimados. Esto implica un exceso computacional porque es necesario calcular cada uno de los valores a estimar y promediarlos, pero es sencillo simplificar esto haciéndolo totalmente equivalente a utilizar las ecuaciones del Kriging Ordinario, a esto se le denomina Kriging por Bloque. Para ello suponga que se desea estimar el valor medio para alguna variable en una área determinada. Dentro de la estructura del modelo de la función aleatoria, el valor medio es el promedio que es una combinación lineal, de todos los puntos en el área. El valor medio del área está dado por

$$z_A = \frac{1}{|A|} \sum_{j \in A} z_j$$

donde z_A es el valor medio en el área A y z_j es el valor de la variable aleatoria dentro de A .

El sistema de ecuaciones de Kriging por Bloque es similar al de Kriging Ordinario, es decir, la única matriz que cambia es D porque esta consiste de los valores de covarianza entre las variables aleatorias de la muestra y el valor de la variable aleatoria que se está tratando de estimar, es decir

$$\tilde{C}_{iA} = \text{Cov}[z_A, z_i] = \frac{1}{|A|} \sum_{j \in A} \text{Cov}[z_i, z_j]$$

La covarianza entre la variable aleatoria i -ésima y la variable z_A , es la misma que el promedio de las covarianzas puntuales entre z_i y las variables aleatorias de todos los puntos dentro de A . Por tanto el sistema de ecuaciones de Kriging por Bloque puede escribirse como

$$C\alpha = D$$

en donde ahora

$$D = \begin{bmatrix} \tilde{C}_{1A} \\ \vdots \\ \tilde{C}_{nA} \\ 1 \end{bmatrix}$$

con

$$\tilde{C}_{iA} = \frac{1}{|A|} \sum_{j \in A} \tilde{C}_{ij} \quad i=1, \dots, n.$$

La varianza de Kriging por Bloque está dada por

$$\hat{\sigma}_{KO}^2 = \bar{C}_{AA} - \left(\sum_{i=1}^n \lambda_i \bar{C}_{iA} + \mu \right)$$

en donde

$$\bar{C}_{AA} = \frac{1}{|A|^2} \sum_{i \in A} \sum_{j \in A} \tilde{C}_{ij}$$

2.5.2 Kriging como un Modelo de Tendencia

Considere la siguiente situación, suponga una región en la cual se tienen mediciones para la variable $z(\cdot)$ en los puntos x_1, x_2, \dots, x_n distribuidos como se presenta en la Figura 2.5 y que en la dirección *SE* estos valores son mayores, por tanto se podría pensar que existe mayor cantidad de $z(\cdot)$ en esa dirección y que existe tendencia en los datos, la cual debería ser tomada en cuenta para la predicción espacial.

El *Kriging como un modelo de tendencia* es el método de predicción espacial que toma en cuenta la tendencia de los datos al construir la función $g(z(\cdot))$. Esta función puede ser expresada como una combinación lineal de funciones, es decir

$$m(x) = \sum_{i=1}^n \alpha_i f_i(x_i) \quad \forall \quad \ell = 1, \dots, l$$

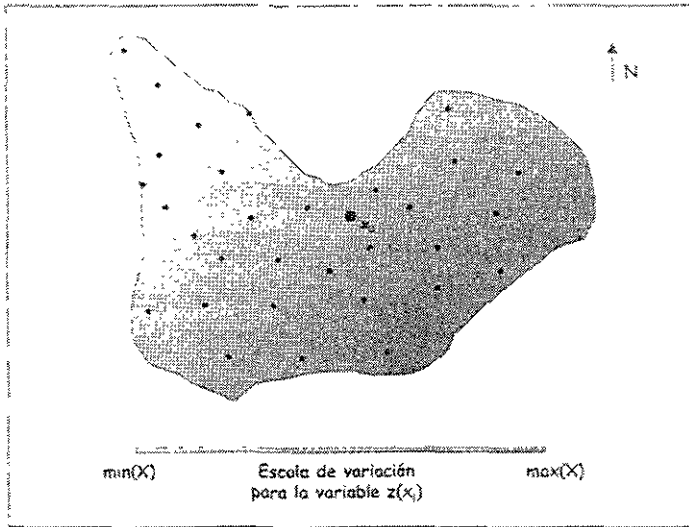


Figura 2.5 Región en Estudio

en donde α_i es el factor de ponderación y $f_i(x_i)$ es la función potencia definida sobre las coordenadas de x cuyo exponente varía de 0 a k , por ejemplo si $k = 2$ y $x_0 = (u, v)$, se tiene a lo más las 6 funciones que estarían dadas como

$$f_1 = 1, f_2 = u, f_3 = v, f_4 = u^2, f_5 = v^2 \text{ y } f_6 = uv$$

y

$$m(u, v) = \alpha_0 + \alpha_1 u + \alpha_2 v + \alpha_3 u^2 + \alpha_4 v^2 + \alpha_5 uv$$

El número de funciones a considerar en la combinación lineal depende asimismo de la dirección en donde se presenta la tendencia. Si esta es en ambas direcciones, entonces estará dada como en la función anterior. Si es el caso de 3 dimensiones con $k = 2$ el predictor estará dado

$$m(U, V, W) = \alpha_1 + \alpha_2 U + \alpha_3 V + \alpha_4 W + \alpha_5 U^2 + \alpha_6 V^2 + \alpha_7 W^2 + \alpha_8 UV + \alpha_9 UW + \alpha_{10} VW$$

En este caso el valor de la media no se puede considerar constante en toda la zona. Este hecho hace que se tome en cuenta que $E\{z(x)\} = m(x)$ y se considera la hipótesis de cuasi-estacionariedad.

Al igual que en Kriging por Bloque, se desea estimar el valor de la variable $z_V(x_0)$ en el área $V(x_0)$ por tanto el estimador estará dado como

$$z_V^*(x_0) = \sum_{i=1}^n \lambda_i z(x_i)$$

en donde se pide que $E\{z_V^*(x_0)\} = z_V(x_0)$, es decir que sea insesgado y que los valores de los λ sean tales que minimicen a su variabilidad, esto es calcular

$$E\left\{z_V^*(x_0) - z_V(x_0)\right\}^2 = \bar{C}[z_V(x_0), z_V(x_0)] - 2 \sum_{i=1}^n \lambda_i \bar{C}[z_V(x_0), z_V(x_i)] + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \bar{C}[z_V(x_i), z_V(x_j)]$$

Utilizando multiplicadores de Lagrange, se tiene la siguiente función

$$L(\lambda, \mu) = E\left\{z_V^*(x_0) - z_V(x_0)\right\}^2 - 2 \sum_{i=1}^k \mu_i \left(\sum_{j=1}^n \lambda_j b_j^i - b^i \right)$$

en donde $b_j^i = \frac{1}{V_j} \int_{V_j} f_j(x) dx$ que es el valor medio de f_j en V_j y $b^i = \sum_{j=1}^n b_j^i$ es el

valor medio de toda la región V .

Esta función genera el sistema de ecuaciones para el Kriging como un modelo de tendencia que es

$$\sum_{j=1}^n \lambda_j \bar{C}(z_v^*(x_i), z_v^*(x_j)) - \sum_{\ell=1}^k \mu_\ell b_\ell = \bar{C}[z_v^*(x_i), z_v(x_0)]$$

$$\sum_{j=1}^n \lambda_j b_j = b^\ell \quad \forall \ell = 1, \dots, k$$

y la varianza es

$$\hat{\sigma}_{kV}^2 = \bar{C}[z_v(x_0), z_v(x_0)] + \sum_{\ell=1}^k \mu_\ell b_\ell - \sum_{i=1}^n \lambda_i \bar{C}[z_v^*(x_i), z_v(x_0)]$$

Tanto el sistema como la varianza pueden escribirse en forma matricial de la siguiente manera

$$K\lambda = M$$

cuya solución está dada por

$$\lambda = K^{-1} M$$

y $\hat{\sigma}_k^2 = \bar{C}(z_v(x_0), z_v(x_0)) - \lambda' M$, en donde

$$\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ -\mu_1 \\ -\mu_2 \\ \vdots \\ -\mu_k \end{bmatrix}, \quad M = \begin{bmatrix} \bar{C}(v_1, V) \\ \bar{C}(v_2, V) \\ \vdots \\ \bar{C}(v_n, V) \\ 1 \\ b_v^2 \\ \vdots \\ b_v^k \end{bmatrix}$$

$$K = \begin{bmatrix} \bar{C}(v_1, v_1) & \dots & \bar{C}(v_1, v_n) & 1 & b_{v_1}^2 & \dots & b_{v_1}^k \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \bar{C}(v_n, v_1) & \dots & \bar{C}(v_n, v_n) & 1 & b_{v_n}^2 & \dots & b_{v_n}^k \\ 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ b_{v_1}^2 & \dots & b_{v_1}^k & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{v_n}^2 & \dots & b_{v_n}^k & 0 & 0 & \dots & 0 \end{bmatrix}.$$

2.5.3 Median Polish

Esta manera de hacer predicción espacial también toma en cuenta la tendencia de los datos pero tiene una manera muy particular de hacerlo y es la que se describe a continuación.

Suponga que la malla de datos es una tabla de contingencia de $p \times q$, por ejemplo si se estuviera trabajando en el plano cartesiano, las coordenadas en X corresponderían a las columnas de la tabla y las coordenadas en Y a los renglones tal que cada nodo es una entrada de la tabla, ver Figura 2.6.

Como se ha venido mencionando la predicción espacial requiere de una función, la cual puede ser una combinación lineal de los datos, entonces

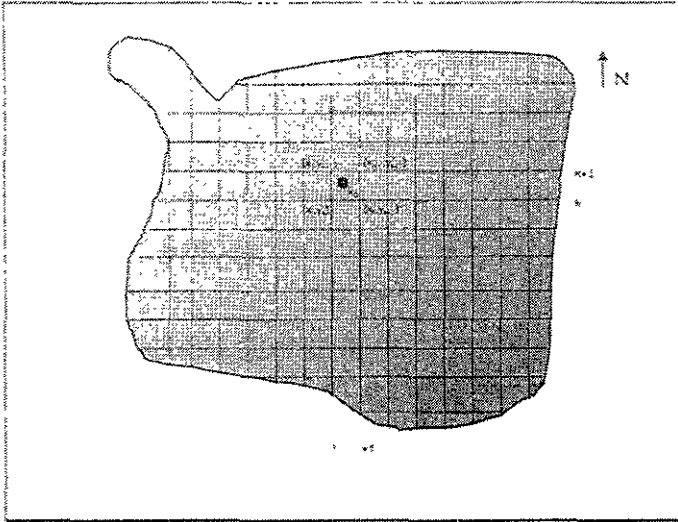


Figura 2.6 Región Median Polish

$$z^*(x) = \mu + c(x) + r(y) \\ = \mu + c_l + r_k$$

en donde r_k es el efecto del k -ésimo renglón, c_l el efecto de la l -ésima columna y μ es la media general de los datos.

Considerando que una observación por cada nodo, los valores de r_k , c_l y μ pueden ser estimados por mínimos cuadrados y estarán dados por:

$$\mu^* = \frac{\sum_{i=1}^n z(x_i)}{n}, \quad r_k^* = \left(\sum_{(x_i, y_i)} z(x_i, y_i) \right) - \mu^*, \quad c_l^* = \left(\sum_{(x_i, y_i)} z(x_i, y_i) \right) - \mu^*$$

para $k = 1, \dots, q$, $l = 1, \dots, p$ y donde $N(y_k) = \{i : x_{i-1}, y_k\} \cup \{i, \dots, n\}$ y $M(x_l) = \{j : x_l = (x_j, y_j), j = 1, \dots, n\}$

Ahora suponga que el punto que se quiere interpolar es $x_0 = (x, y)$ (ver Figura 2.6), el cual se encuentra entre los puntos (x_1, y_k) , (x_{l-1}, y_k) , (x_l, y_{k-1}) y (x_{l-1}, y_{k-1}) en donde $x_l < x_{l-1}$ y $y_k < y_{k-1}$ definen el plano interpolado por

$$z^*(x) = \mu^* + c_l^* + \left(\frac{x - x_l}{x_{l-1} - x_l} \right) (c_{l-1}^* - c_l^*) - r_k^* + \left(\frac{y - y_k}{y_{k-1} - y_k} \right) (r_{k-1}^* - r_k^*)$$

para $k = 1, \dots, p-1$ y $l = 1, \dots, q-1$.

Esta manera de estimar el valor del punto x_0 tiene la ventaja de ser un estimador no paramétrico, espacialmente escalable y continuo. Su desventaja es que sus residuos $\{z(x_0) - z^*(x_0)\}$ conducen a estimadores sesgados de la dependencia espacial en el proceso de error.

La idea de Median-Polish es que se puede construir estimadores no paramétricos de μ^* , c_l^* y r_k^* que sean lo más sensibles posible para reducir el sesgo, esto se puede hacer bajo la idea de considerar que la distribución de los errores es simétrica. Con esta idea se tiene lo siguiente.

$$E[\text{med}\{z(x_i) : i \in A\}] = E[\text{med}\{z(x_i) : i \in A\}]$$

en donde $med\{\}$ y $medi\{\}$ son la media y la mediana de los valores de $z(x_i)$ en el área A . Esto tiene la ventaja de tomar en cuenta algunas propiedades de la mediana como son que sea resistente a "outliers", la homoscedasticidad, etc. Al establecer la analogía con la mediana, se tendrán residuos que harán que el sesgo sea pequeño, lo que implica que los valores estimados de μ^* , c_i^* y r_k^* serán menos sesgados.

La principal propiedad que tiene este método es la forma que tiene de representar la variación de gran escala, es decir, de la tendencia de los datos. El algoritmo de Median Polish se presenta en el Apéndice 1.

2.5.4 Validación Cruzada

Kriging es un interpolado exacto con el cual se puede evaluar la bondad de ajuste del variograma elegido. Esto es, una vez elegido el variograma, se calculan las ecuaciones de Kriging y se predice el valor del punto elegido. Si ese punto elegido fuera alguno de los puntos muestreados y con el resto se predijera el valor que le correspondería, se podría calcular la diferencia entre el verdadero valor y el estimado, la cual se esperaría fuera mínima si la estimación fuese buena. A este procedimiento se llama *Validación Cruzada*. La Validación debe ser realizada de manera secuencial para cada uno de los puntos y con los diferentes modelos, con la idea de determinar cual es el mas indicado para representar la correlación espacial.

Con la diferencia entre el verdadero valor y el esperado se obtienen n errores de validación y con ellos se calculan medidas que cumplan con criterios que ayuden a determinar si un modelo es óptimo o no. Algunos de los criterios son:

Mínimo Error Cuadrático Medio. Se elige el modelo del variograma que proporcione el menor valor de

$$XVAL = \frac{1}{n} \sum_{i=1}^n [z^*(x_i) - z(x_i)]^2$$

Mínimo Valor Medio de las Varianzas. Se escoge el modelo del variograma asociado a la menor medida

$$VKM = \frac{1}{n} \sum_{i=1}^n \sigma_e^2$$

2.5.5 Cokriging

Las diferentes maneras de calcular el Kriging han sido considerando una sola variable. Sin embargo es útil pensar que cuando se llevan a cabo mediciones, no solamente se toma en cuenta una variable sino varias y que es posible que estas variables estén relacionadas, por ejemplo la salinidad del agua marina puede estar en función de su acidez. Esta correlación que existe puede servir para mejorar el valor de estimación y por tanto para minimizar la varianza del error de predicción.

La idea es extender Kriging pero para n-variables tomando en cuenta la correlación que existe entre ellas y minimizando su variabilidad. La forma de hacerlo se presenta a continuación y se le denomina *Cokriging*.

Es importante definir cuál es la variable que se desea predecir, pues las otras sólo se tomarán en cuenta la relación que existe con la primera. En ese orden, las variables serán denominadas variables primaria y secundarias. Por facilidad y sin pérdida de generalidad, se considerará el caso para dos variables.

Con el Cokriging también se desea estimar valores de puntos no muestreados mediante la combinación lineal de las variables muestreadas y podría ser definida como

$$z^*(x_0) = \sum_{i=1}^n \lambda_i z(x_i) + \sum_{j=1}^m \eta_j w(y_j)$$

donde $z^*(x_0)$ es el valor de la variable primaria que se desea estimar en el punto x_0 , $z(x_1), \dots, z(x_n)$ y $w(x_1), \dots, w(x_m)$ son los valores de las variables primaria y secundaria respectivamente, λ_i y η_j son los factores de ponderación fijos pero desconocidos.

La construcción del conjunto de ecuaciones de Cokriging es idéntica a la de Kriging Ordinario, es decir, se quiere un estimador insesgado por tanto

$$E[z^*(x_0)] = z(x_0)$$

lo cual lleva a establecer las condiciones de insesgamiento y deben ser tales que minimicen la variabilidad.

Para cumplir con que el estimador sea insesgado se tiene

$$\begin{aligned} E[\hat{z}(x_0)] &= E\left[\sum_{i=1}^n \lambda_i z(x_i) + \sum_{j=1}^m \eta_j w(y_j)\right] \\ &= \tilde{m}_z \sum_{i=1}^n \lambda_i + \tilde{m}_w \sum_{j=1}^m \eta_j \end{aligned}$$

donde $E[z(x_i)] = \tilde{m}_z$ y $E[w(x_j)] = \tilde{m}_w$.

Una primera forma de inesgamiento es la siguiente

$$\sum_{i=1}^n \lambda_i = 1 \quad \text{y} \quad \sum_{j=1}^m \eta_j = 0$$

con las cuales se tendría la siguiente función de Lagrange para calcular los valores de λ_i y η_j

$$L(\lambda, \eta, \mu) = E\left[(z^*(x_0) - z(x_0))^2\right] + 2\mu_1 \left(\sum_{i=1}^n \lambda_i - 1\right) + 2\mu_2 \left(\sum_{j=1}^m \eta_j\right)$$

El sistema de ecuaciones de Cokriging es

$$\sum_{i=1}^n \lambda_i \text{Cov}[z(x_i), z(x_j)] + \sum_{j=1}^m \eta_j \text{Cov}[z(x_i), w(x_j)] + \mu_1 = \text{Cov}[z(x_0), z(x_j)] \quad \text{para } j=1, \dots, n$$

$$\sum_{i=1}^n \lambda_i \text{Cov}[z(x_i), w(x_j)] + \sum_{j=1}^m \eta_j \text{Cov}[w(x_i), w(x_j)] + \mu_2 = \text{Cov}[z(x_0), w(x_j)] \quad \text{para } j=1, \dots, m$$

$$\sum_{i=1}^n \lambda_i = 1 \quad \text{y} \quad \sum_{j=1}^m \eta_j = 0.$$

Si se establece otro tipo de condiciones de inesgamiento este sistema cambiaría.

Capítulo 3: Análisis Geoestadístico

3.1 Generalidades

En el capítulo 1 se mencionó que este trabajo tiene como objetivo el de presentar una propuesta de Análisis Geoestadístico como una metodología adecuada para el procesamiento de los datos considerando la situación espacial de ellos.

Los conjuntos de datos que se trabajaron en esta tesis corresponden a 3 muestreos, cada uno realizado en diferente época del año. Estos muestreos se realizaron en las temporadas de secas (marzo a junio), lluvias (julio a septiembre) y nortes (septiembre a febrero). La localización geográfica del área muestreada se presenta en la Figura 1.1.

El análisis que aquí se presenta es únicamente para las variables de Hierro (Fe , 10^{-3} $\mu\text{g/g}$) y Manganeso (Mn , 10^{-3} $\mu\text{g/g}$), elementos que, al no ser distribuidos o diluidos de manera adecuada, son los que más afectan al medio marino. El conjunto de datos observados se presenta en las Figuras 3.2.1 a 3.2.6, en donde algunos valores no son reportados debido a que la batimetría de la zona no llegaba a la profundidad medida, principalmente para las estaciones cercanas a la costa.

El orden en que se presenta el análisis será inicialmente las características que describen la zona durante los periodos de muestreo en términos de temperatura, salinidad y dirección de la corriente y después se hablará del proceso de análisis de los datos desde el punto de vista geoestadístico. Dentro de este, se hicieron los tres tipos de análisis que se describen a continuación:

Análisis Exploratorio. En este apartado, se analiza el comportamiento de los datos por medio de histogramas, medidas de tendencia central, de dispersión, y gráficas en papel normal, por nivel de profundidad, global y sin tomar en cuenta la posición espacial.

Análisis Geoestadístico en \mathfrak{R}^2 . En esta parte para cada contaminante el análisis estructural o variográfico para cada nivel de profundidad, es decir, ahora no sólo se tomará en cuenta el valor de la variable sino también su localización espacial. En función de estas dos partes, se propondrán modelos de variograma que indiquen la continuidad espacial de los datos. Posteriormente, esta información se incorporará dentro del proceso de estimación por medio del kriging, el cual será evaluado por el método de validación cruzada.

Análisis Geoestadístico en \mathfrak{R}^3 . En esta parte se consideran las 3 dimensiones (latitud, longitud y profundidad). Al igual que en el caso de \mathfrak{R}^2 , se calcula el variograma empírico pero ahora tomando en cuenta también la profundidad, para identificar donde se da la continuidad espacial. Posteriormente, se genera el volumen de valores predictivos por medio de kriging, el cual será nuevamente evaluado por validación cruzada.

La presentación de la información para cada uno de los apartados se hará primero para Fe, luego Mn y posteriormente para algunas variables transformadas. El análisis geoestadístico se realizó con la ayuda del programa GEOEAS y las bibliotecas de GSLIB editadas por Deutsch y Journel (1992, 1998).

3.2 Características generales de la zona muestreada durante las diferentes épocas del año.

Las condiciones que se presentaron durante cada muestreo, en términos generales son las siguientes (SEDEMAR, 1992):

- **Periodo de Secas:** Comprende de los meses de marzo a junio. Durante el mes de mayo la corriente predominante fue con dirección Noroeste y fue cambiando a través del tiempo,

de manera bidireccional, con componente alta hacia el Sureste. La temperatura promedio varió entre 22.20 °C a 22.80 °C, siendo mayor en la parte cercana a la costa. La salinidad media varió entre 34.31 ‰ y 35.40 ‰ siendo mayor a la altura del difusor.

- Periodo de Lluvias: Comprende los meses de julio a septiembre. La corriente predominante inició de manera bidireccional con componente alta hacia el Sureste y cambiando a través del tiempo hacia el noreste. La temperatura promedio osciló entre 23.02 °C a 23.73 °C con variaciones hasta de 8 °C. En cuanto a la salinidad, fue de 34.95 ‰ a 37.57 ‰ con una variabilidad hasta de 10‰.
- Periodo de Nortes: Comprenden los meses de septiembre a febrero. La corriente predominante cambió de ser marcadamente unidireccional con componente al Suroeste, a ser predominante al Sureste. La temperatura promedio varió entre 23.48 a 26.82 °C decreciendo a través del tiempo. La salinidad se mantuvo en 34 ‰ y con poca variación.

3.3 Análisis Exploratorio

Este análisis se realizó por separado para los 3 conjuntos de datos, dentro de cada uno estos, se hizo otro análisis por cada nivel de profundidad y para cada una de las variables. Aquí sólo se reportarán los resultados para Fe y Mn, el resto se encuentra en el documento oficial y de manera parcial (SEDEMAR, 1992).

Las medidas de tendencia central y de dispersión de la información se presentan en la Tabla 3.1. Los histogramas y las gráficas de papel normal asociadas a los datos se presentan en las Figuras 3.1.1-1 y 2 a Figuras 3.1.6-1 y 2.

Variables	Nivel de Profundidad (m)	No. de Datos Observados por temporada		Media	Desviación Estándar	Mínimo	1er cuartil	Mediana	3er cuartil	Máximo	
		s	n								
Fe	0	s	22	3.636	2.816	0.900	1.375	2.800	4.450	9.600	
		ll	25	4.500	2.936	1.000	4.500	7.200	9.100	14.300	
		n	25	7.092	9.713	1.700	3.200	4.100	5.000	49.800	
	4	s	24	4.754	5.881	1.000	1.825	3.050	4.525	29.800	
		ll	24	16.333	19.050	3.800	6.200	8.000	11.325	72.200	
		n	20	6.780	10.537	1.700	2.525	3.350	4.450	46.900	
	8	s	20	4.035	3.337	0.400	1.200	3.150	5.500	13.200	
		ll	20	15.615	22.229	6.000	7.425	8.900	12.100	105.800	
		n	20	10.960	11.865	2.400	3.850	5.050	13.600	38.000	
	12	s	21	6.08	8.948	0.400	3.100	4.200	7.400	43.400	
		ll	19	20.610	35.648	4.700	7.200	8.700	13.850	160.000	
		n	17	18.429	31.624	2.400	4.100	5.500	19.500	127.500	
	16	s	7	9.028	11.394	1.600	3.950	5.300	6.950	34.500	
		ll	9	7.022	2.164	2.300	7.000	7.200	8.000	9.900	
		n	10	5.150	2.442	2.200	3.825	4.500	5.650	10.000	
	Global	s	94	5.100	6.400	0.400	1.800	3.400	6.000	43.400	
		ll	97	13.800	21.300	1.000	6.700	8.100	9.900	160.0	
		n	92	9.749	16.568	1.700	3.075	4.200	6.550	127.500	
	Mn	0	s	20	1.050	0.787	0.100	0.575	1.000	1.350	3.500
			ll	25	1.720	0.980	0.000	1.000	2.000	2.000	3.000
			n	25	3.612	1.144	2.200	2.800	3.500	4.200	6.100
4		s	21	0.986	0.511	0.200	0.700	0.900	1.400	1.900	
		ll	24	2.875	2.092	1.000	1.750	2.000	3.250	10.000	
		n	20	3.420	0.854	1.700	2.925	3.650	4.025	4.500	
8		s	18	1.317	1.556	0.100	0.500	0.900	1.400	7.000	
		ll	20	2.400	1.353	1.000	2.000	2.000	3.000	7.000	
		n	20	3.730	1.378	2.000	2.975	3.450	4.200	8.400	
12		s	16	1.975	3.410	0.300	0.575	1.050	1.600	14.300	
		ll	19	2.263	0.933	0.000	2.000	3.000	3.000	3.000	
		n	17	4.559	3.476	1.700	3.300	3.600	4.100	15.400	
16		s	5	1.060	0.754	0.400	0.600	0.800	1.200	2.300	
		ll	9	2.556	1.236	1.000	2.000	3.000	3.000	5.000	
		n	10	3.540	0.947	2.500	2.900	3.250	3.950	5.300	
Global		s	80	1.300	1.800	0.100	0.600	0.900	1.400	14.300	
		ll	97	2.300	1.500	0.000	1.000	2.000	3.000	10.000	
		n	92	3.763	1.806	1.700	2.800	3.500	4.125	15.400	

Tabla 3.1.1 Medidas de Tendencia Central y de Dispersión para las variables Fe y Mn de los tres muestreos, por renglón se representa la información de cada temporada (s: secas, ll: lluvias, n: nortes).

a) Primer Muestreo

Para la variable Fe, el comportamiento de los datos observados no presenta una distribución aproximada a la normal, esto se debe a los valores extremos que tiene, ver Figura 3.1.1-1(a), (c), (e) y Figura 3.1.1-2 (a), (c), (e). La mayoría de los datos se agrupan en el intervalo de 0.0 a $10.0 \times 10^{-3} \mu\text{g/g}$ en todos los niveles de profundidad, excepto el de 12 m, el cual tiene valores de hasta $43.400 \times 10^{-3} \mu\text{g/g}$, ver Tabla 3.1. Los niveles 12 y 16 m son los que tienen mayor concentración, lo cual era de esperarse por ser esta variable un metal.

Al observar las gráficas en papel normal (Figura 3.1.1-1(b), (d), (f) y Figura 3.1.1-2 (b), (d), (f)) se puede ver que para los niveles de 4, 8, 12 y 16 m hay un valor extremo que influye por nivel de profundidad, que va de 13.200 a $43.400 \times 10^{-3} \mu\text{g/g}$. Esto puede comprobarse si se analiza los valores del tercer cuartil que van de 4.525 a $7.400 \times 10^{-3} \mu\text{g/g}$, y que son de 3 a 6 veces menores en proporción a su respectivo valor máximo, ver Tabla 3.1.

En cuanto a las medidas de tendencia central (media y mediana), ambas tienden a crecer en los primeros 4 niveles de profundidad, con lo cual se piensa que esta variable se está sedimentando.

Con lo que respecta a la variabilidad de los datos, al igual que la media, la desviación estándar crece en función de la profundidad para los niveles de 0, 4, 8 y 12 m. A los 16 metros este valor disminuye y se debe en gran parte a que el número de datos con los que se calcula son pocos.

De manera global, los datos presentan un comportamiento muy similar al nivel de profundidad 12 m, tanto en el histograma como en la gráfica en papel normal (Figura 3.1.1-2 (a), (b)), mucho tiene que ver en este nivel se encuentra los valores extremos tanto el mínimo como el máximo. Si se compara las otras medidas descriptivas como la media, mediana y cuartiles, los valores globales resultan ser menor en aproximadamente una unidad.

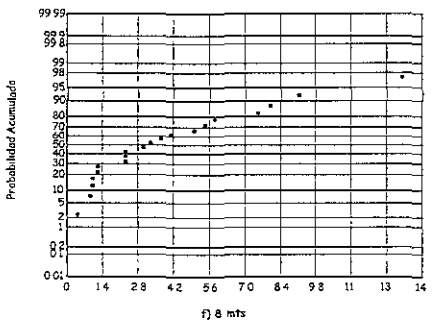
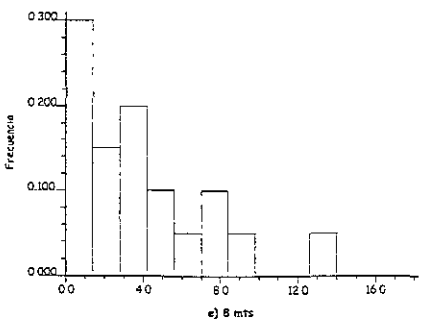
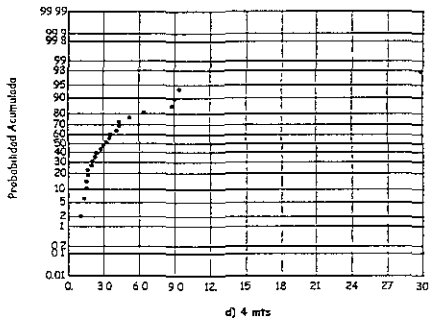
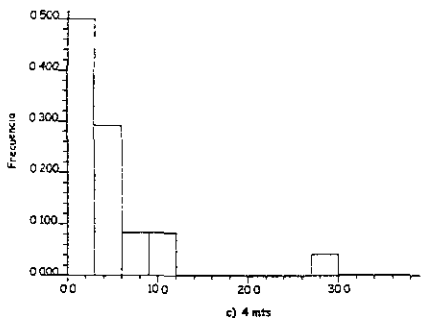
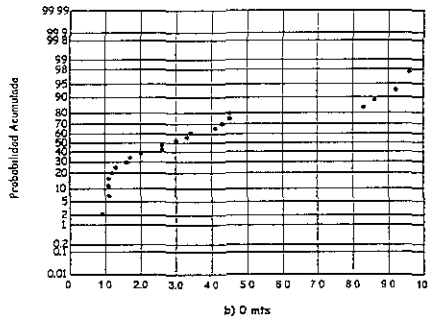
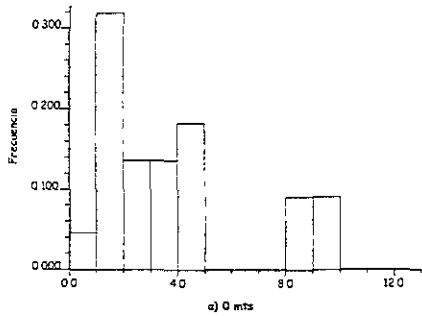


Figura 3.1.1-1 Histogramas y gráficas en papel normal para Fe por nivel de profundidad, Primer Muestreo

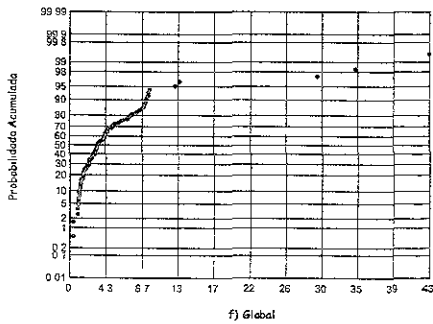
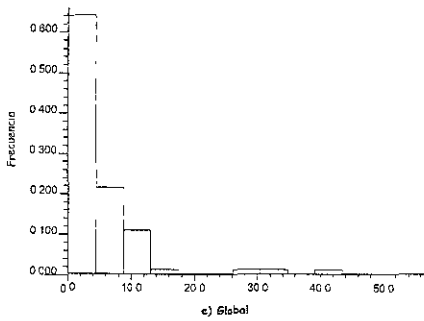
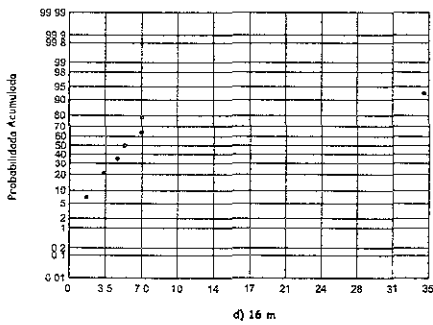
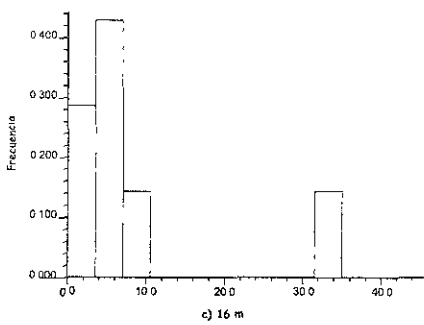
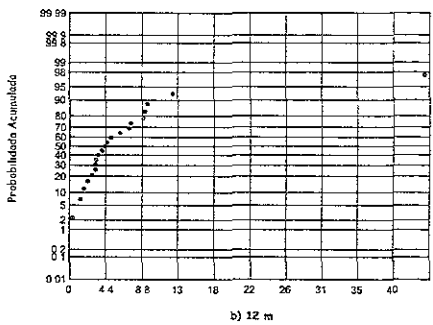
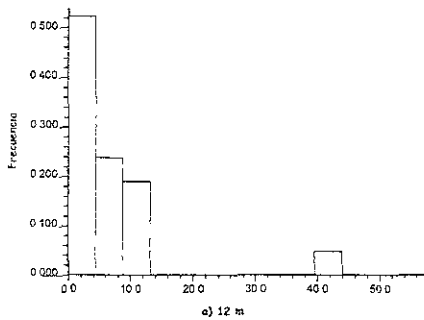


Figura 3.1 1-2 Histogramas y gráficas en papel normal para Fe por nivel de profundidad y global, Primer Muestreo

Para la variable Mn, tampoco se presenta una distribución aproximada a la normal (Figura 3.1.2-1 y Figura 3.1.2-2), y también se debe a la magnitud de sus valores extremos, principalmente en los niveles 8, 12 y 16 m y en especial en los dos primeros. Los terceros cuartiles van de 6 a 9 veces menores en proporción con sus respectivos máximos. La media no presenta mucha diferencia de un nivel a otro, solo en el nivel de 12 m, que es donde se encuentra el valor máximo de toda la muestra, ver Tabla 3.1.

Las gráficas en papel normal muestran los datos alineados a excepción del valor extremo. Si este valore no estuviera, los datos tendrían un comportamiento aproximada a la normal.

En cuanto a la desviación estándar, no cambia mucho a excepción del nivel de 12 m de profundidad. De nuevo esto refleja la presencia del valor extremo.

De manera global, se puede decir que quitando el valor extremo de $14,300 \times 10^{-3} \mu\text{g/g}$ que hay en el nivel de 12 m, la variable Mn parecería un comportamiento normal pero con varianzas diferentes.

b) Segundo Muestreo

Este segundo conjunto de datos de la variable Fe presenta valores mayores que en el primer muestreo en general, es decir, los valores máximos son de 4 a 6 veces mayores en los mismos niveles. Esto habla de una menor dispersión del metal en este muestreo.

Aproximadamente el 75% de los datos se encuentra en el rango de 0.0 a $15.0 \times 10^{-3} \mu\text{g/g}$ (ver tercer cuartil, Tabla 3.1). Se puede hablar de una distribución normal en forma aproximada para los niveles de 0 y 16 m (ver Figura 3.1.3-1 (a), (b) y Figura 3.1.3-2 (c), (d)), mientras que en los otros niveles, hay un valor que distorsiona y altera el comportamiento de la nube de datos. Con respecto al valor de la media, este crece en función de los niveles de profundidad, excepto en el último. La mediana no presenta mayor cambio, pues se encuentra entre 7.200 a

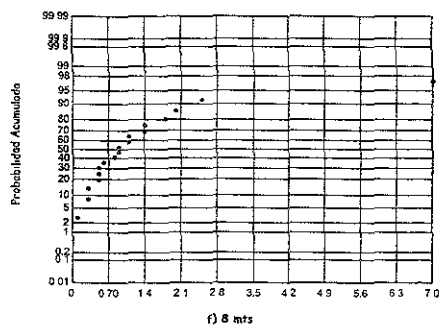
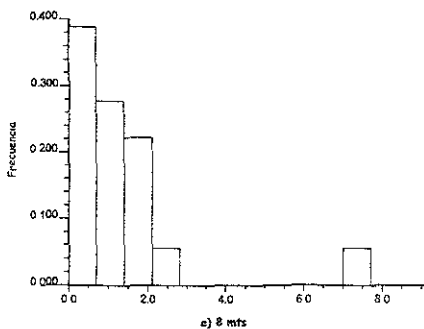
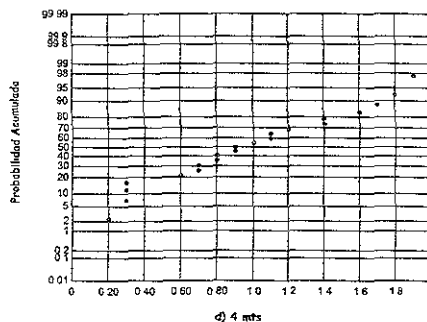
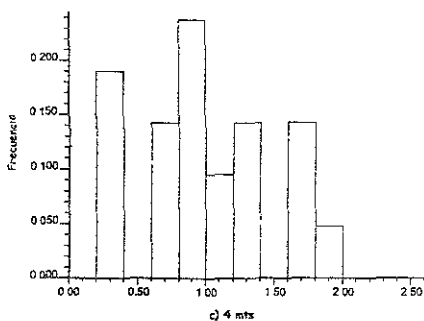
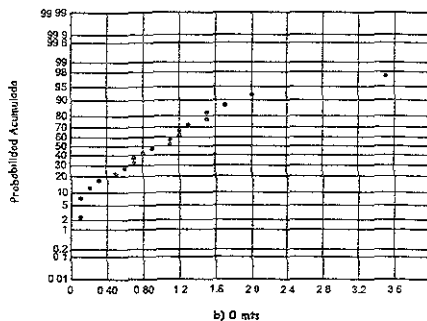
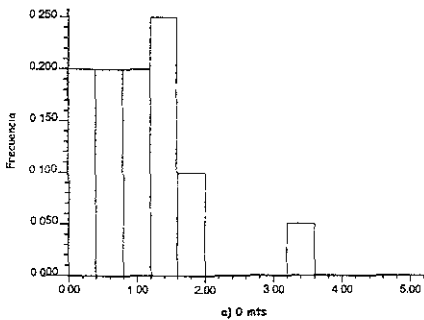


Figura 3.1.2-1 Histogramas y gráfica en papel normal para Mn por nivel de profundidad, Primer Muestreo

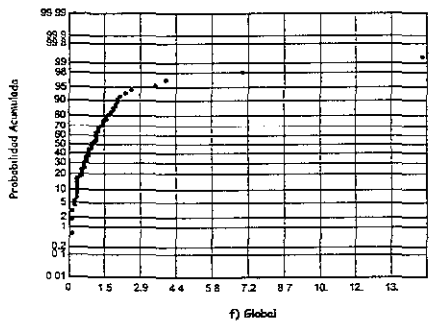
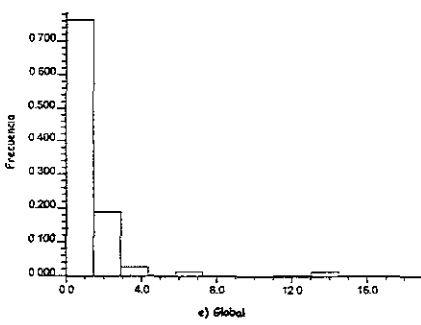
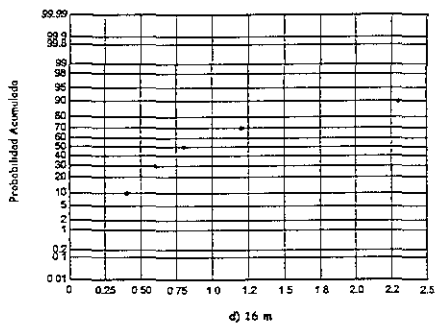
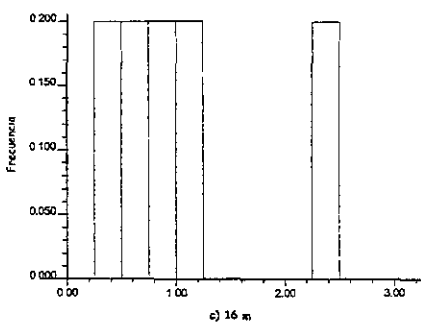
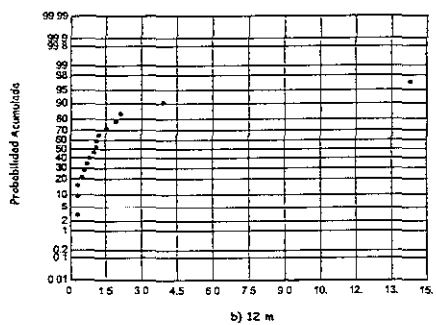
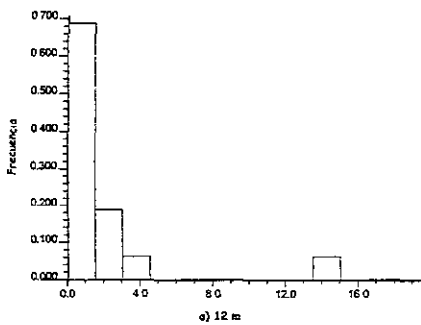


Figura 3.1.2-2 Histogramas y gráficas en papel normal para Mn por nivel de profundidad y global, Primer Muestreo

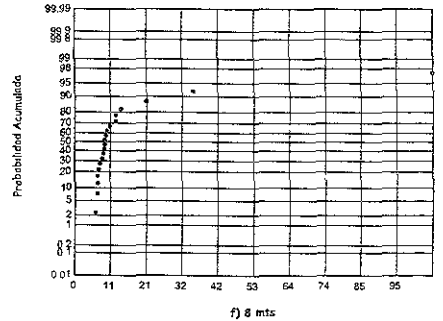
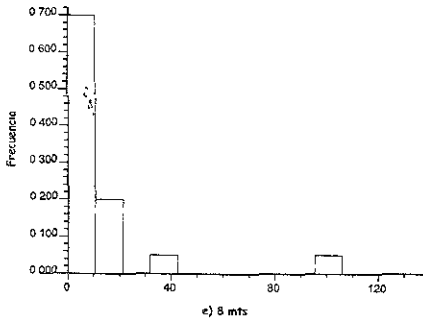
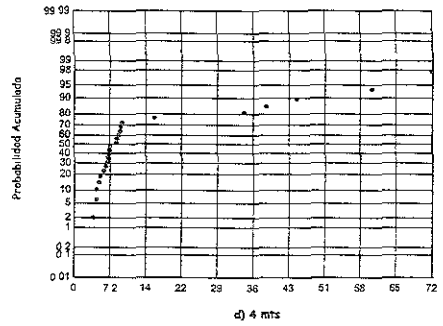
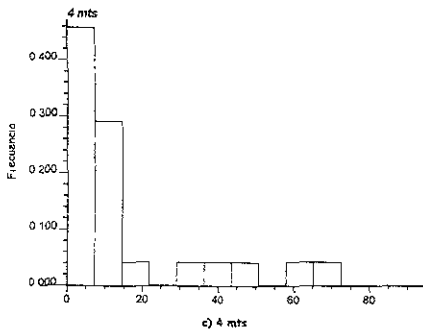
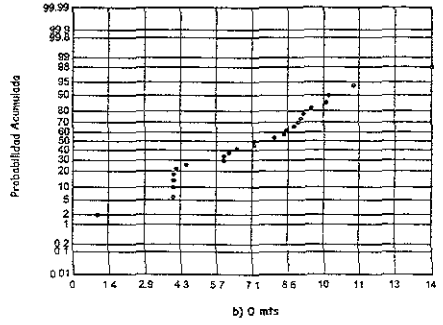
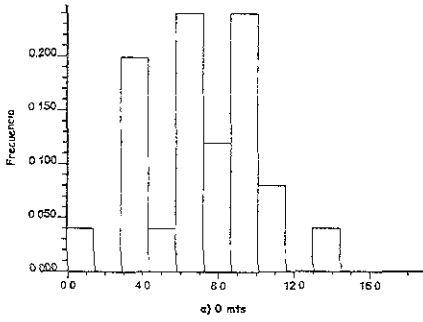


Figura 3.1.3-1 Histogramas y gráficas en papel normal para Fe por nivel de profundidad, Segundo Muestreo

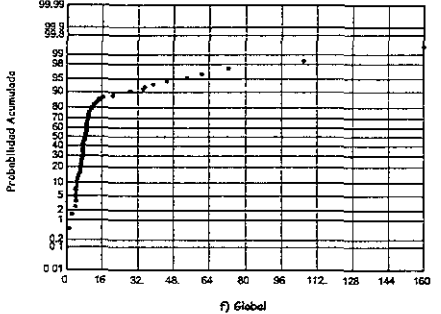
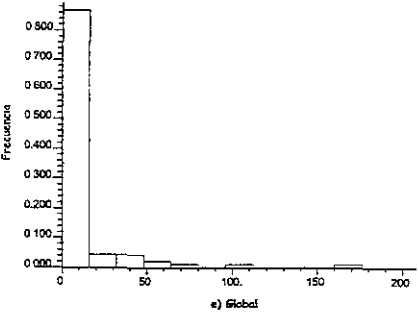
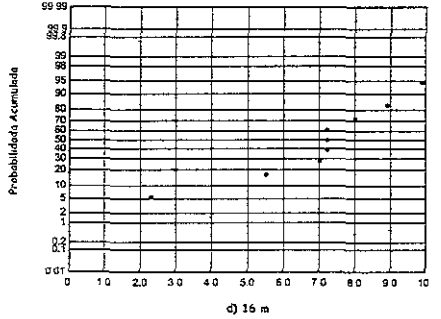
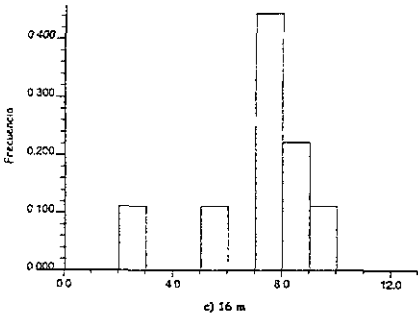
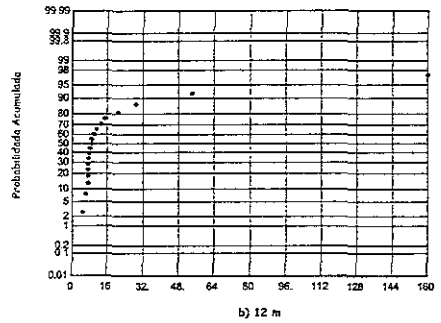
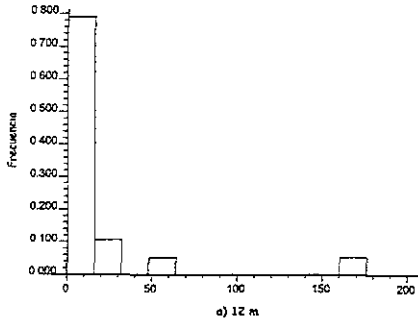


Figura 3.1.3-2 Histogramas y gráficas en papel normal para Fe por nivel de profundidad y global, Segundo Muestreo

$8.900 \times 10^{-3} \mu\text{g/g}$. Para los niveles 0 y 16 m, los valores de desviación estándar son muy parecidos y para el resto es creciente en función de la profundidad.

En los 3 niveles que presentan mayor variabilidad, el valor extremo no es el único que afecta el comportamiento global de los datos, en algunos casos hay hasta tres valores que pueden considerarse aberrantes. Así, no se puede pensar que al quitar uno solo de ellos habría cambio en la distribución y quitar los tres no reflejaría el comportamiento real de la muestra.

En la variable Mn, el nivel de profundidad que presenta mayor variabilidad es el de 4 m y es porque contiene el valor más grande de la muestra en general. El rango del resto de los niveles, que es aproximadamente un 75% de los datos de la muestra, va de 0.0 a $3.0 \times 10^{-3} \mu\text{g/g}$, ver Tabla 3.1. La media y mediana no muestran un incremento importante a través de los niveles de profundidad y las gráficas en papel normal, a pesar de tener un comportamiento escalonado, podrían considerarse aproximadamente normales, ver Figura 3.1.4-1 y Figura 3.1.4-2.

c) Tercer Muestreo

Al igual que los otros 2 muestreos, la variable Fe presenta valores extremos en cada nivel de profundidad a excepción de los 16 m (Tabla 3.1 y Figura 3.1.5-2 (c) y (d)). El nivel de los 8 m, tiene valores a lo largo de todo su rango que va de 0.0 a $40.0 \times 10^{-3} \mu\text{g/g}$ (Figura 3.1.5-1 (c) y (d)), mientras que los niveles 0, 4, y 12 solo tienen un valor extremo en cada uno de ellos, y el cual influye en sus medidas descriptivas (Tabla 3.1). Los datos no presentan una distribución aproximada a la normal, de hecho si se pensara en quitar el valor máximo, continuaría siendo no normal.

Los valores de media y mediana tienen un comportamiento senoidal tal que en los 2 primeros niveles decrecen, pero para los siguientes 2 crece, y vuelven a disminuir en el último. La misma situación se presenta con el resto de los valores descriptivos.

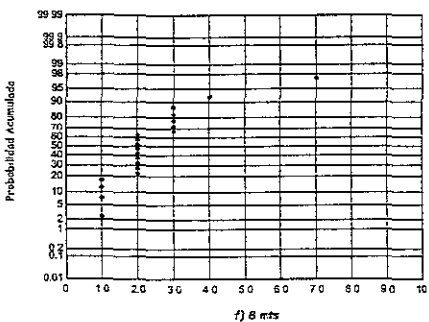
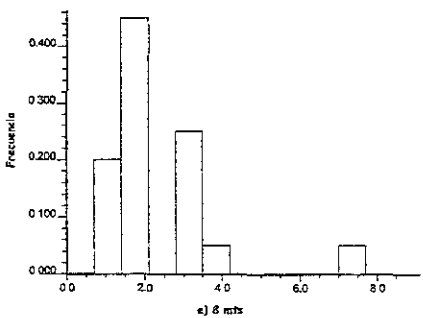
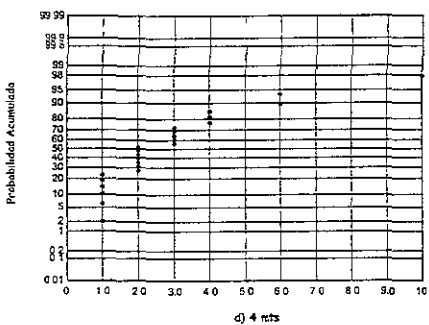
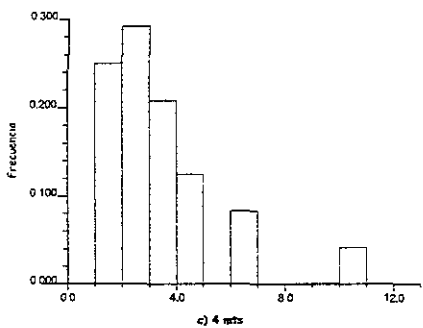
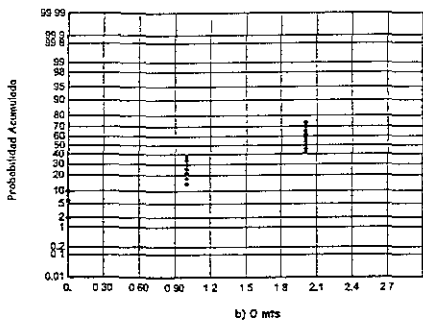
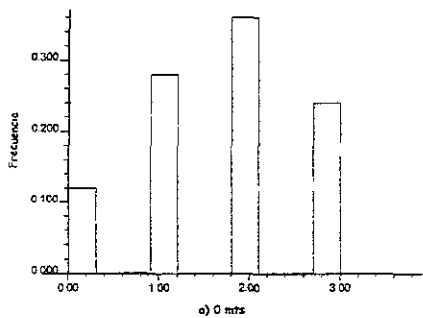


Figura 3.1.4-1 Histogramas y gráficas en papel normal para Mn por nivel de profundidad, Segundo Muestreo

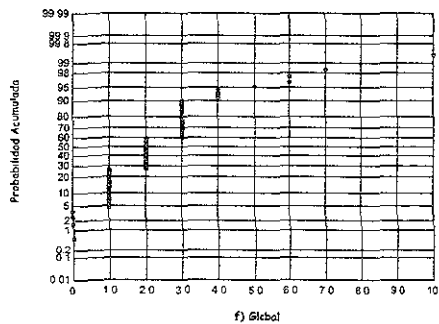
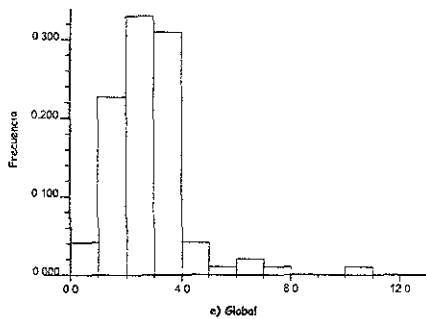
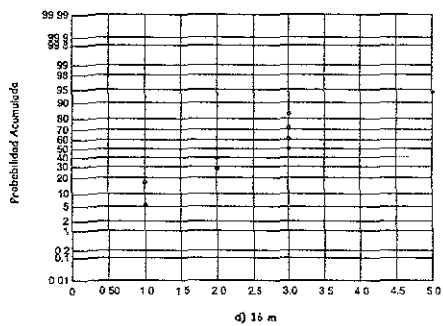
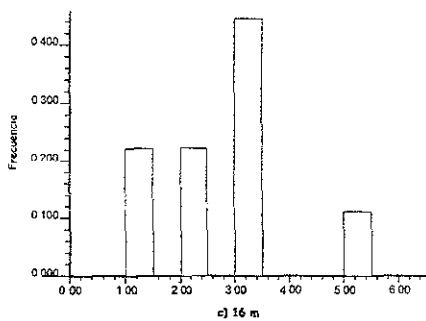
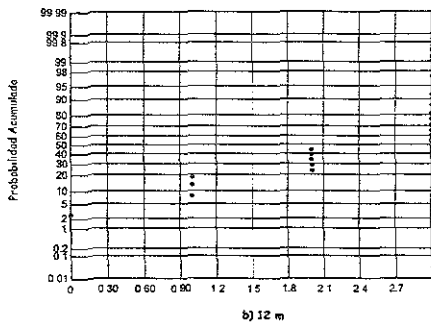
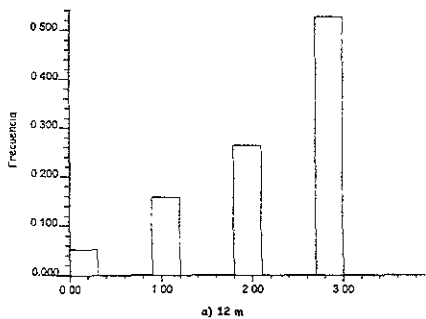


Figura 3.1.4-2 Histogramas y gráfica en papel normal para Mn por nivel de profundidad y global, Segundo Muestreo

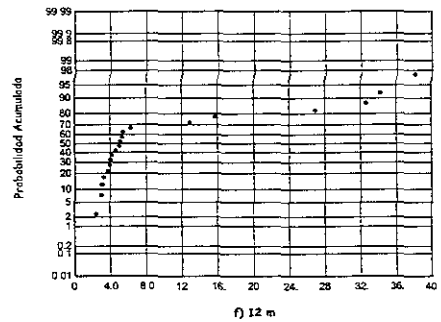
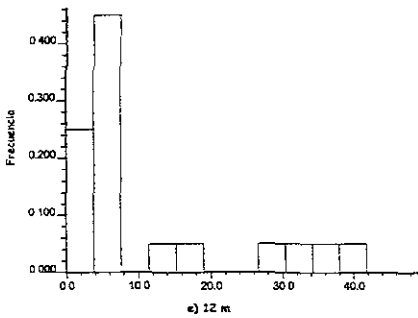
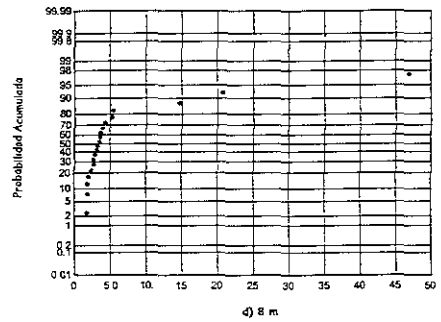
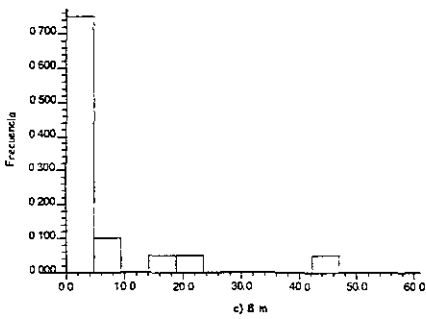
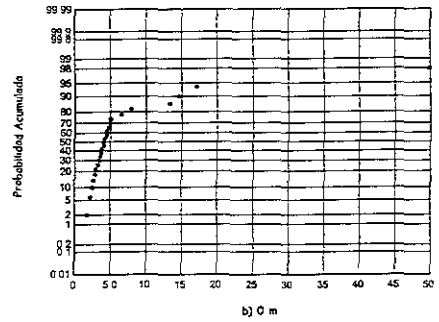
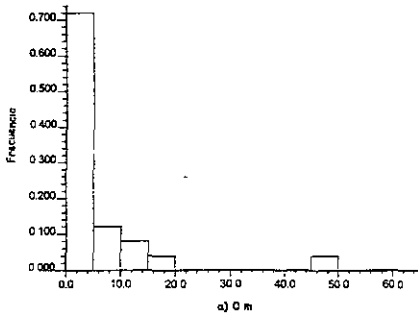


Figura 3.1.5-1 Histogramas y gráficas en papel normal para Fe por nivel de profundidad, Tercer Muestreo

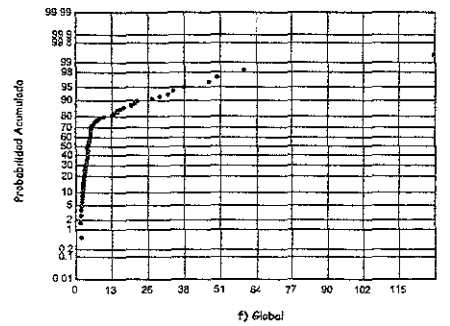
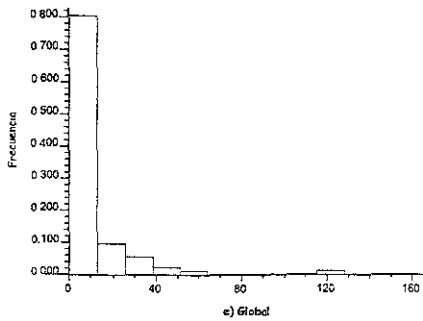
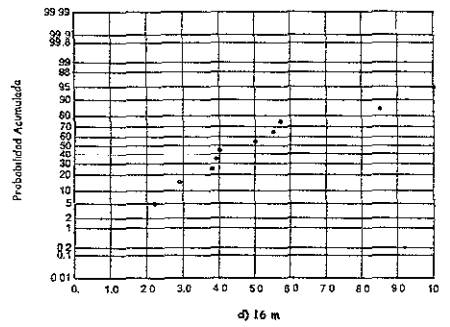
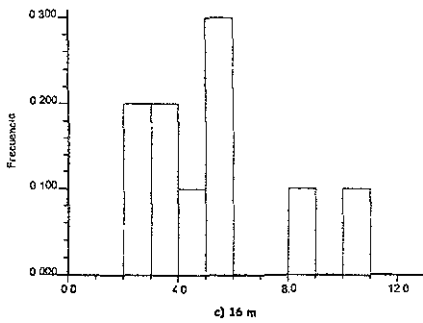
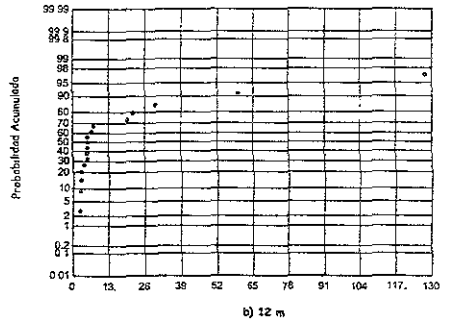
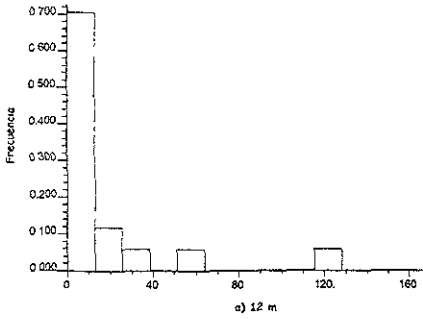


Figura 3.1.5-2 Histogramas y gráficas en papel normal para Fe por nivel de profundidad y global, Tercer Muestreo

Al observar las gráficas y medidas descriptivas de los datos, de manera global, se puede decir que la mayoría de los valores se encuentran en el rango de 0.0 a $40.0 \times 10^{-3} \mu\text{g/g}$ y que un solo valor es mayor a $120.0 \times 10^{-3} \mu\text{g/g}$ (Figura 3.1.5-2 (e) y (f) y Tabla 3.1).

Para la variable Mn, en todos los niveles de profundidad, la mayoría de los datos se concentran en el rango de 0.0 a $6.0 \times 10^{-3} \mu\text{g/g}$ y solamente los niveles de 8 y 12 m presentan los de mayor concentración. Las medidas de tendencia central presentan un patrón similar al del Fe, al igual que la desviación estándar. Los histogramas y gráficas en papel normal podrían considerarse como semejantes a los de una distribución normal (Figura 3.1.6-2 (e) y (f)).

3.3.1 Transformaciones

Para los tres muestreos, el Fe es la variable que muestra un comportamiento no normal y que daría pie para trabajarlos bajo una transformación que quizás ayudara a tener esta distribución. Se optó por buscar una familia de transformaciones que fuera adecuada y para ello se utilizó la familia de Box y Cox (Atkinson, 1987) cuyo propósito es encontrar el valor de λ que indicara la transformación idónea para que los datos tuvieran una distribución normal, además de que, como valor agregado, la varianza fuera homogénea. Esto se aplicó en cada uno de los muestreos pero solo en el segundo y tercer muestreo se obtuvo que los valores adecuados para λ fuera 0, es decir, se tendría que calcular el logaritmo de los datos en ambos conjuntos para los niveles de profundidad 4, 8, 12 y 0, 4, 8 y 12 m.

Los histogramas y gráficas en papel normal de los datos transformados se presentan en las Figura 3.1.7 y Figura 3.1.8 y las medidas de tendencia central y de dispersión en la Tabla 3.2. Con lo que respecta al primer muestreo y al nivel cero del segundo muestreo, los valores de λ fueron muy cercanos a 1 lo cual indica que no era necesario hacer ninguna transformación.

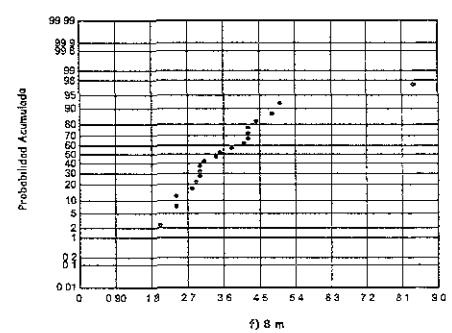
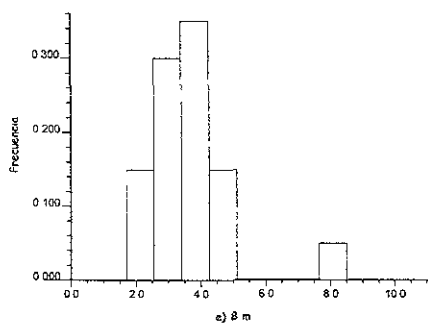
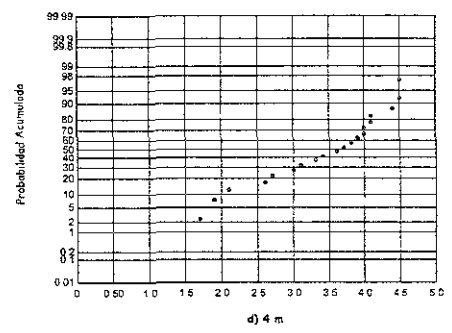
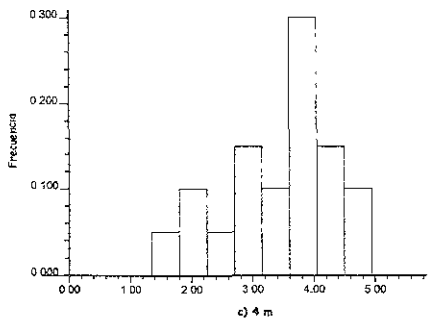
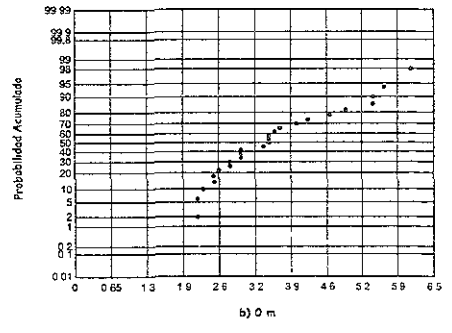
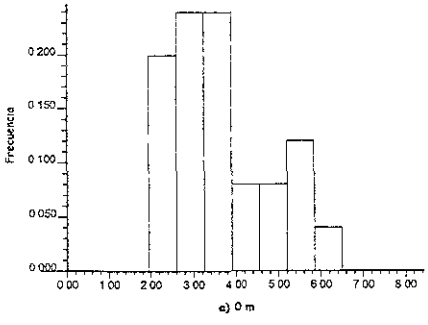


Figura 3.1.6-1 Histogramas y gráficas en papel normal para Mn por nivel de profundidad, Tercer Muestreo

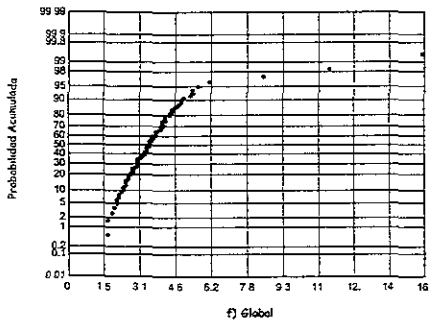
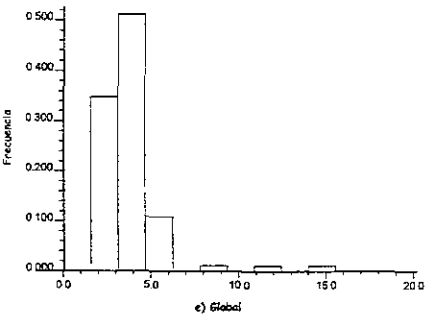
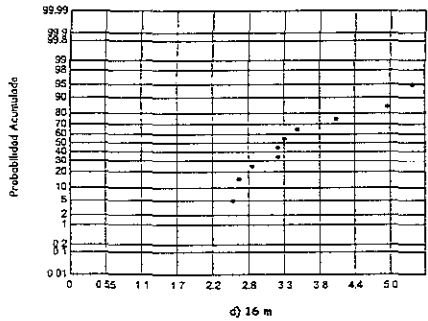
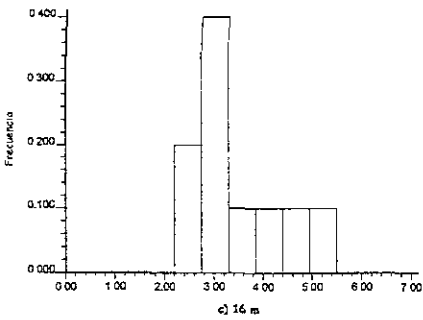
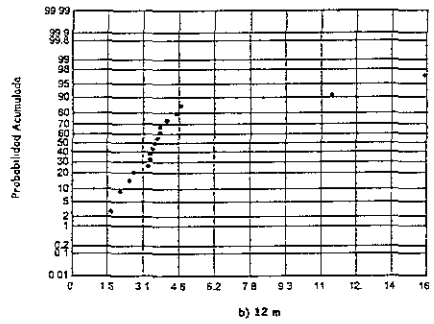
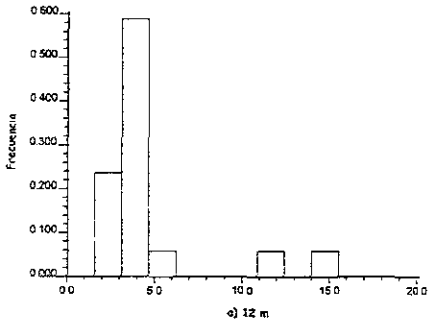


Figura 3.1.6-2 Histogramas y gráficas en papel normal para Mn por nivel de profundidad y global, Tercer Muestreo

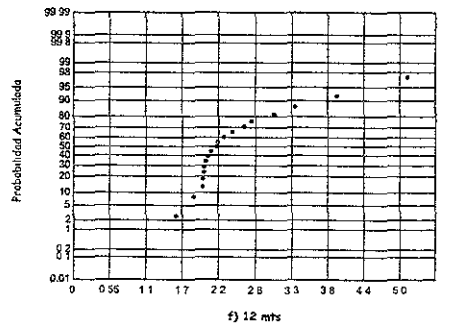
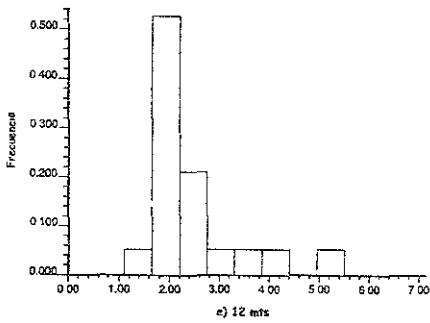
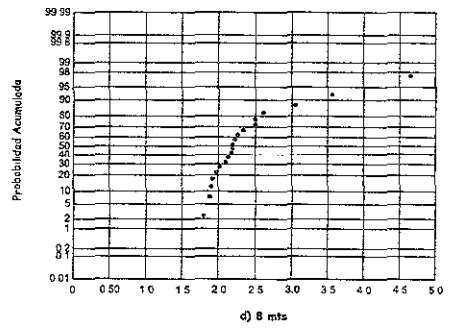
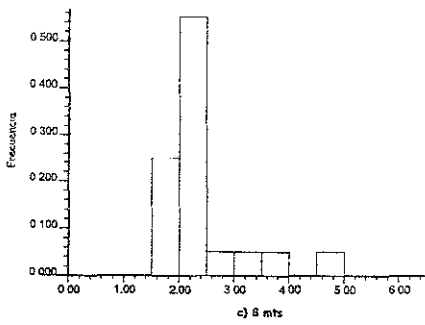
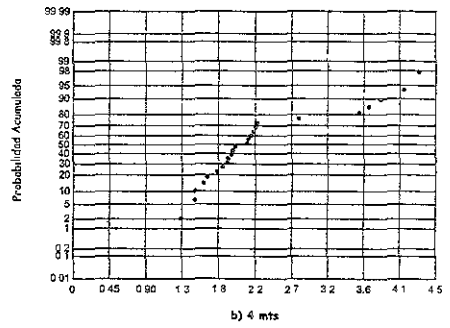
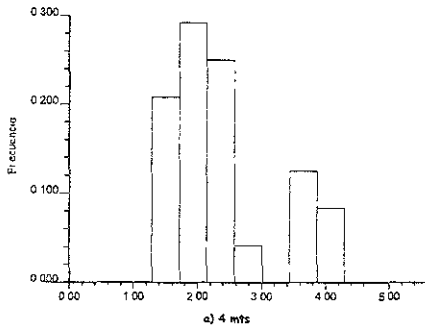


Figura 3.1.7 Histogramas y gráfica en papel normal para $\ln(\text{Fe})$ por nivel de profundidad, Segundo Muestreo

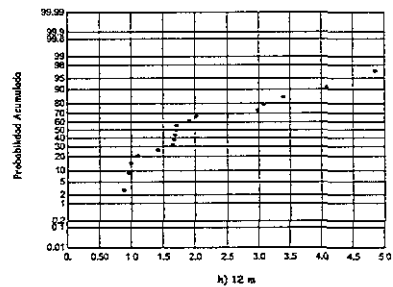
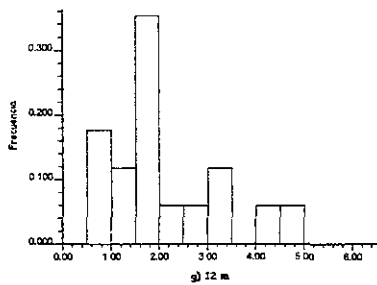
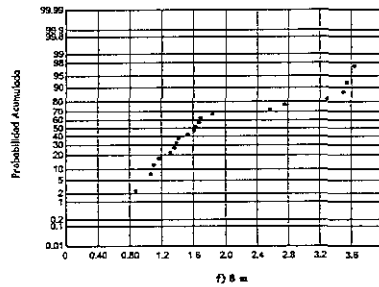
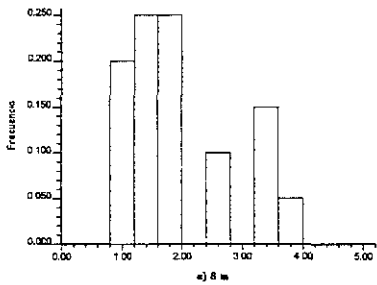
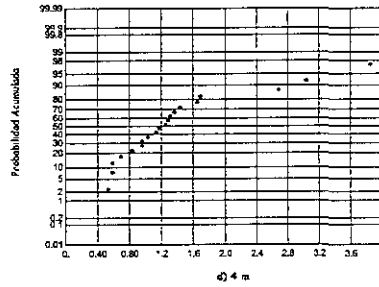
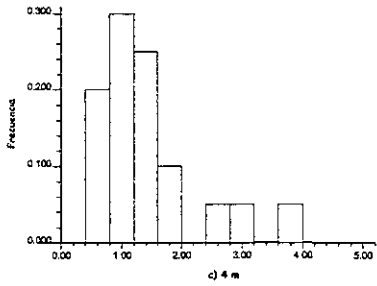
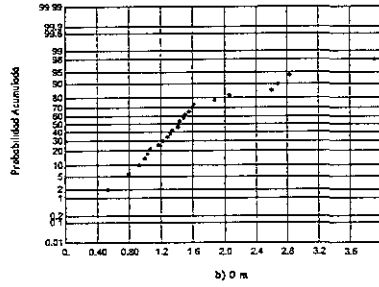
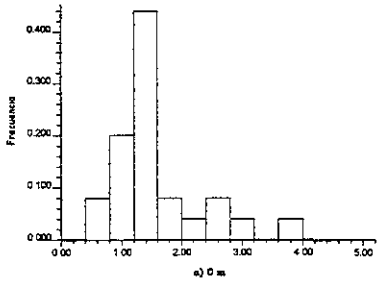


Figura 3.1.8 Histogramas y gráficas en papel normal para $\ln(\text{Fe})$ por nivel de profundidad, Tercer Muestreo

Número de Muestreo	Nivel de Profundidad (m)	No de Datos Observados	Media	Desviación Estándar	Mínimo	1er cuartil	Mediana	3er cuartil	Máximo
2	4	24	2.350	0.866	1.335	1.824	2.077	2.400	4.179
	8	20	2.395	0.679	1.792	2.005	2.186	2.493	4.662
	12	19	2.477	0.860	1.548	1.974	2.163	2.627	5.075
3	0	25	1.582	0.742	0.531	1.163	1.411	1.609	3.908
	4	20	1.400	0.860	0.531	0.925	1.208	1.488	3.848
	8	20	1.944	0.910	0.875	1.348	1.619	2.606	3.638
	12	17	2.119	1.153	0.875	1.411	1.705	2.970	4.848

Tabla 3.1.2 Medidas de Tendencia Central y de Dispersión para la variable $\ln(\text{Fe})$, 2o y Tercer Muestreo

Al analizar los datos transformados, con sus gráficas y medidas descriptivas, se puede observar que los histogramas son semejantes a una distribución normal. Las gráficas en papel normal muestran un comportamiento lineal, en comparación con el conjunto de datos originales, y la pendiente es pronunciada. Las medidas descriptivas presentan las mismas tendencias explicadas anteriormente pero con menor variabilidad.

Aunque en Geoestadística no necesariamente supone que las variables tengan una distribución particular, si se recomienda considerar alguna distribución para este tipo de análisis (Journel y Huijbregts, 1978). Otros autores comentan que se puede trabajar a las variables tal cual se presentan (Diggle y Tawn, 1998). Sin embargo, para esta trabajo se decidió manejar ambos casos, es decir, analizar los datos originales, bajo una transformación para el Fe y posteriormente comparar los resultados.

3.4 Análisis Geoestadístico en \mathbb{R}^2

3.4.1 Descripción de datos en 2 dimensiones

Como se mencionó anteriormente, se marcaron los datos de acuerdo al rango de valores definidos por los valores extremos y los intercuantiles, es decir se definieron 4 grupos de la siguiente manera:

Grupos	Rango
Menores	[Min, 1er cuartil)
Menores Medios	[1er cuartil, Mediana)
Mayores Medios	[Mediana, 3er cuartil)
Mayores	[3er cuartil, Max]

Y se ubicaron espacialmente en las Figuras 3.2.1 a 3.2.6. De acuerdo a esta ubicación se presentaron los patrones que a continuación se describen y pensando que la malla muestreada es una matriz de datos de 5 x 5.

a) Primer Muestreo

Durante esta época, los valores menores de Fe en sedimentos se registraron en las estaciones de la playa. Se establecieron 2 núcleos: al Sureste y al Noreste de la red; asociado con la presencia de limos y arenas finas. Fue evidente una coloración marrón-amarillenta en los sedimentos del área cercana al difusor, (SEDEMAR, 1992).

Al observar cada uno de los niveles de profundidad, se puede ver que las mayores concentraciones se encuentran en los extremos norte y sur de la red, siendo los de mayor concentración a los 4, 8 y 12 m (estación 6), al noroeste de la red. Con lo que respecta a los otros valores, se observa la presencia del difusor pues son valores pequeños en el sentido de estar en los rangos del mínimo al 3er cuartil. Sin embargo, al nivel de 12 m los valores de mayor concentración se encuentran en el segundo y tercer renglón, lo que hablaría de puntos de concentración o de sedimentación del metal, ver Figura 3.2.1. Es importante mencionar que la corriente en esta época fue con dirección noroeste (sección 3.2).

La distribución del Mn mostró núcleos de altas concentraciones al Noroeste, Sureste y Noreste de la red, asociados a sedimentos limo-arenosos en los 2 primeros y limo-arcillosos en el último, mismos que poseen mayor capacidad de adhesión con los metales pesados, (SEDEMAR, 1992).

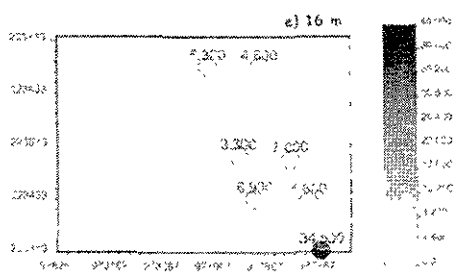
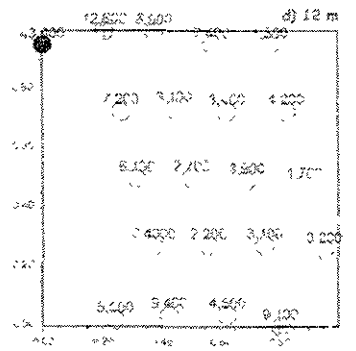
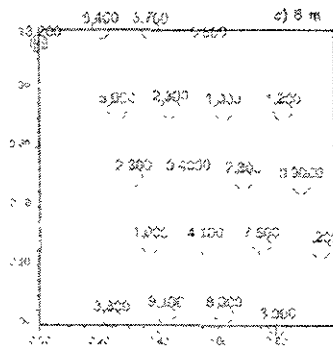
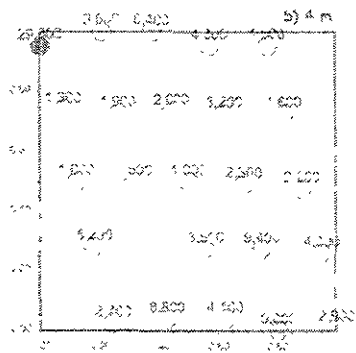
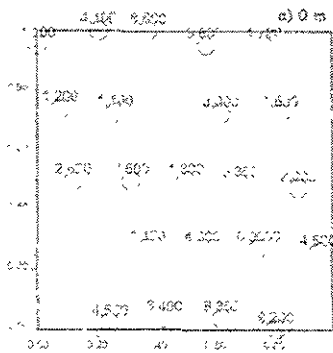


Figura 3.2.1 Mapa de datos para Fe par nivel de profundidad, Primer Muestreo

De manera similar al Fe, la máxima concentración de Mn se da al noroeste, Norte y Sur de la red y a los 4, 8 y 12 m de profundidad (estación 6). En los renglones 2 y 3 se presentan los valores que van por debajo de la mediana y que habla de la presencia del difusor. En los niveles de 4 y 8 m, hay mayor concentración en los 2 últimos renglones, podría pensarse que el Mn tiende a sedimentarse en estas estaciones, pero hay que tomar en cuenta que en la parte superior de la malla hay datos faltantes, ver Figura 3.2.2.

b) Segundo Muestreo

De los 3 muestreos, este es el que presenta mayor variabilidad, lo cual es de esperarse porque hay un factor que es determinante que es la presencia de lluvia.

En la época de lluvias fue posible apreciar la formación de un núcleo de altas concentraciones de Fe, a partir del extremo final del difusor, el cual presentó un decremento concéntrico, (SEDEMAR, 1992).

Al observar los valores de mayor concentración de Fe se puede ver que estos se encuentran entre los renglones de 2 a 4 y en todas las profundidades y a pesar de que al Norte y Sur (filos 1 y 5) tienen valores altos, no son tan grandes como los de en medio. Esto puede dar a pensar que la fuerza con que el difusor expulsa el agua con residuos no es suficiente como tampoco lo es la corriente para dispersarlos; si uno observa los mapas de 8 y 12 m de profundidad, los puntos con valores más altos se encuentran en dirección hacia el Sur. Con los que respecta a valores pequeños tienden a presentarse hacia el Noreste y Sureste, ver Figura 3.2.3.

En el periodo de lluvias se registró un núcleo de altas concentraciones de Mn en el área localizada a lo largo del difusor, con un decremento paulatino y registros elevados al Noroeste y Noreste de la red, (SEDEMAR, 1992).

Debido a como fue reportada la información, no se puede decir mucho del comportamiento de los datos, ya que los intervalos que se definieron anteriormente están traslapados y no son de

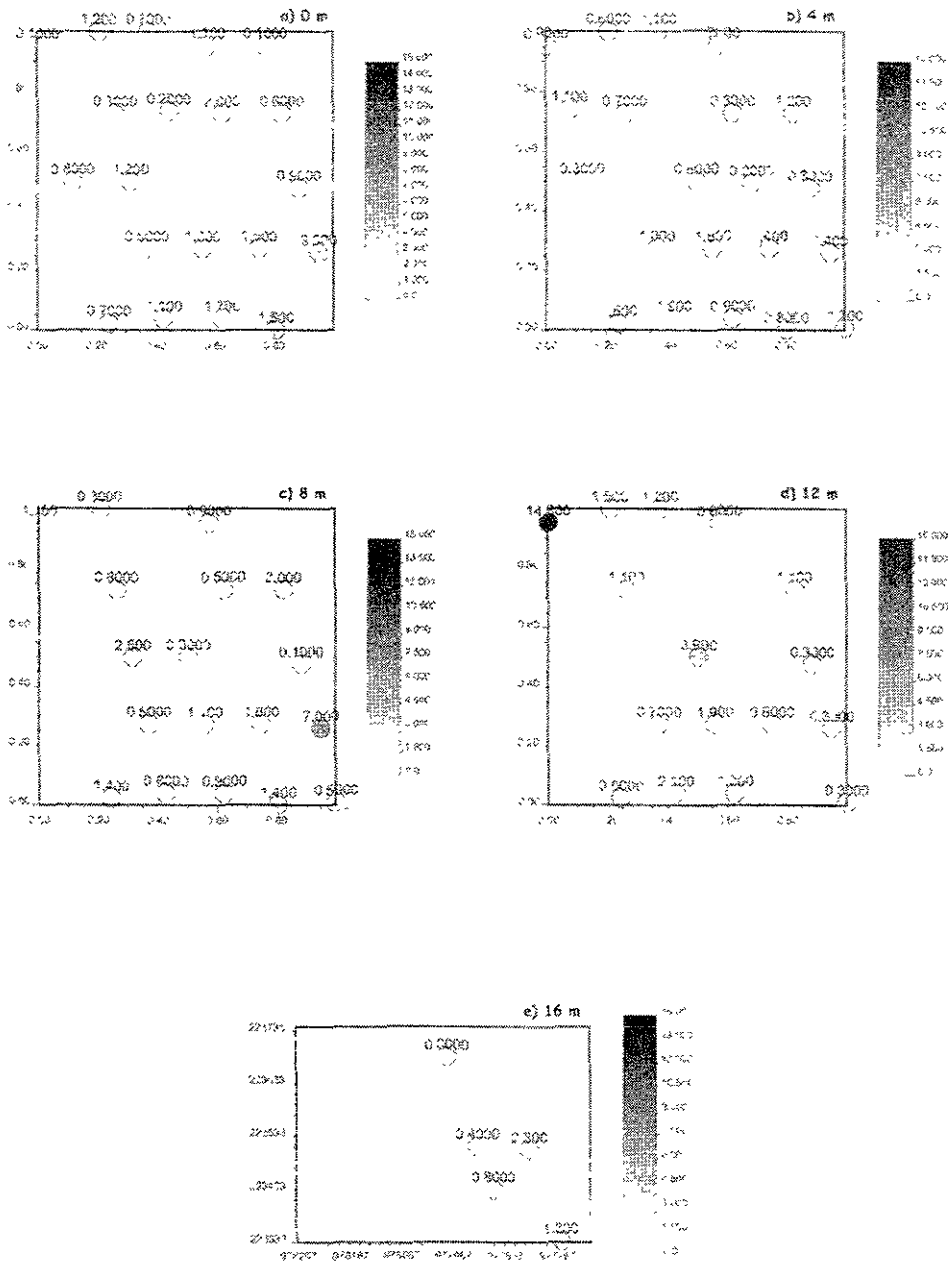


Figura 3.2.2 Mapa de datos para Mn por nivel de profundidad, Primer Muestreo

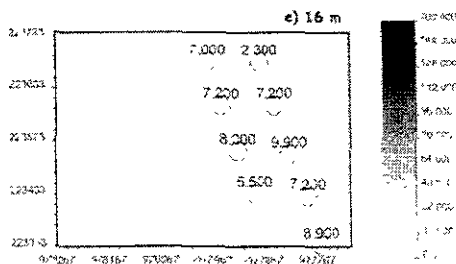
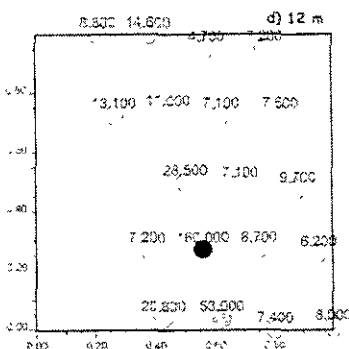
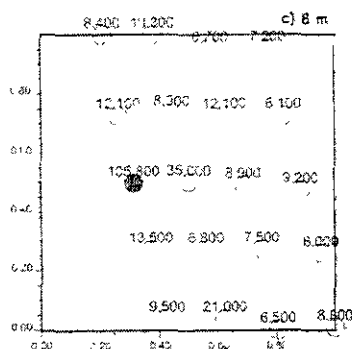
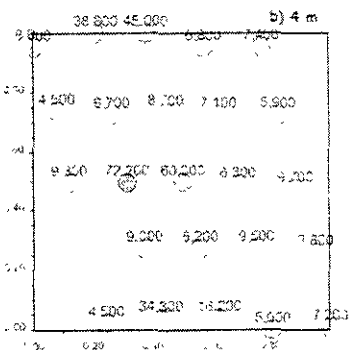
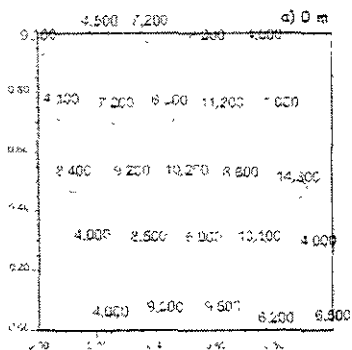


Figura 3.2.3 Mapa de datos para Fe por nivel de profundidad, Segundo Muestreo

mucha ayuda. Solo se puede decir que valores pequeños de Mn tienden a irse hacia el Sureste de la malla de muestreo y que los valores de mayor magnitud, se encuentran en medio de la malla (tercer renglón) en los niveles de 4 y 8 m de profundidad, que es donde aproximadamente se encuentra el difusor, ver Figura 3.2.4.

c) Tercer Muestreo

Durante la temporada de nortes, la tendencia de la distribución de Fe en los sedimentos fue a aumentar desde la costa, hacia altamar. Fue notorio el establecimiento al Norte, Noroeste, Sureste y en el área localizada a lo largo del difusor, (SEDEMAR, 1992).

En este muestreo se nota la presencia del difusor y la corriente del agua pues si uno observa los valores de mayor concentración de Fe se encuentran hacia el norte de la red, mientras que al sur están los de menor concentración. Con lo que respecta al nivel de 12 m, la estación de mayor concentración es la que se ubica exactamente en medio de la malla y es donde termina el difusor. Si se observa la columna completa, en este nivel, todos los valores que se presentan en ella están por arriba de su media y con valores altos al Norte y bajos al Sur, Figura 3.2.5

Si se observa el perímetro de la malla de Mn al nivel de 0 m de profundidad (ver Figura 3.2.6 (a)), 10 de las 16 estaciones presentan valores mayores a su respectiva mediana. Por otro lado, al estudiar la cuarta columna de la malla, casi todos son valores extremos y eso se repite para los niveles de 4 y 8 m. Si se pensara en la malla como un plano cartesiano, el cuarto cuadrante tendría los mayores valores en todos los niveles de profundidad. Si se considera el comportamiento del Fe, se podría pensar que el difusor separa los elementos, es decir, el Fe hacia el Norte y el Mn al Sur.

3.4.2 Descripción General del Análisis Estructural

Para este análisis, se identificó la ubicación geográfica de los valores extremos de cada nivel de profundidad y se observó que éstos variaban en función de la situación climatológica de la

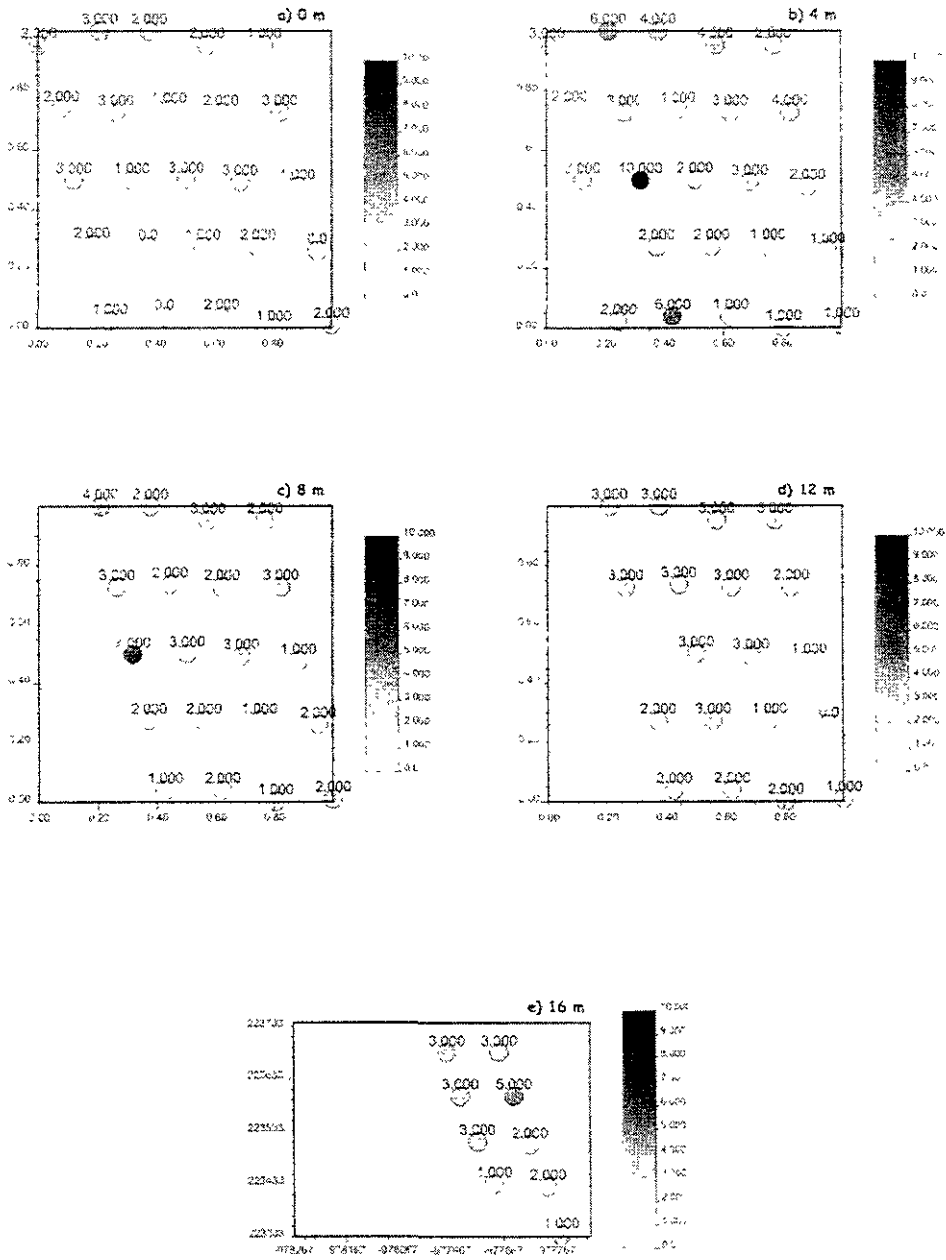


Figura 3.2.4 Mapa de datos para Mn por nivel de profundidad, Segundo Muestreo

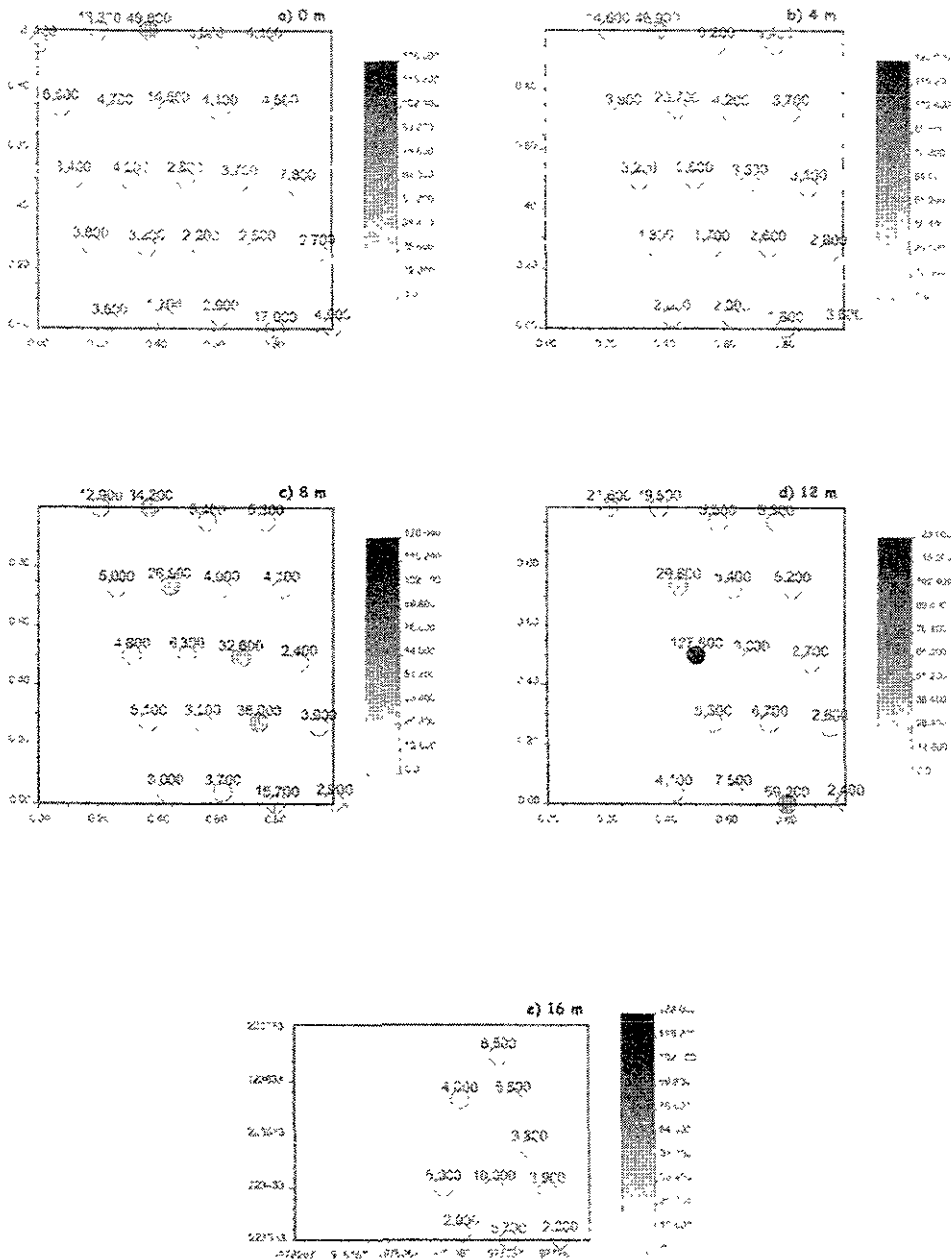


Figura 3.2.5 Mapa de datos para Fe por nivel de profundidad, Tercer Muestreo

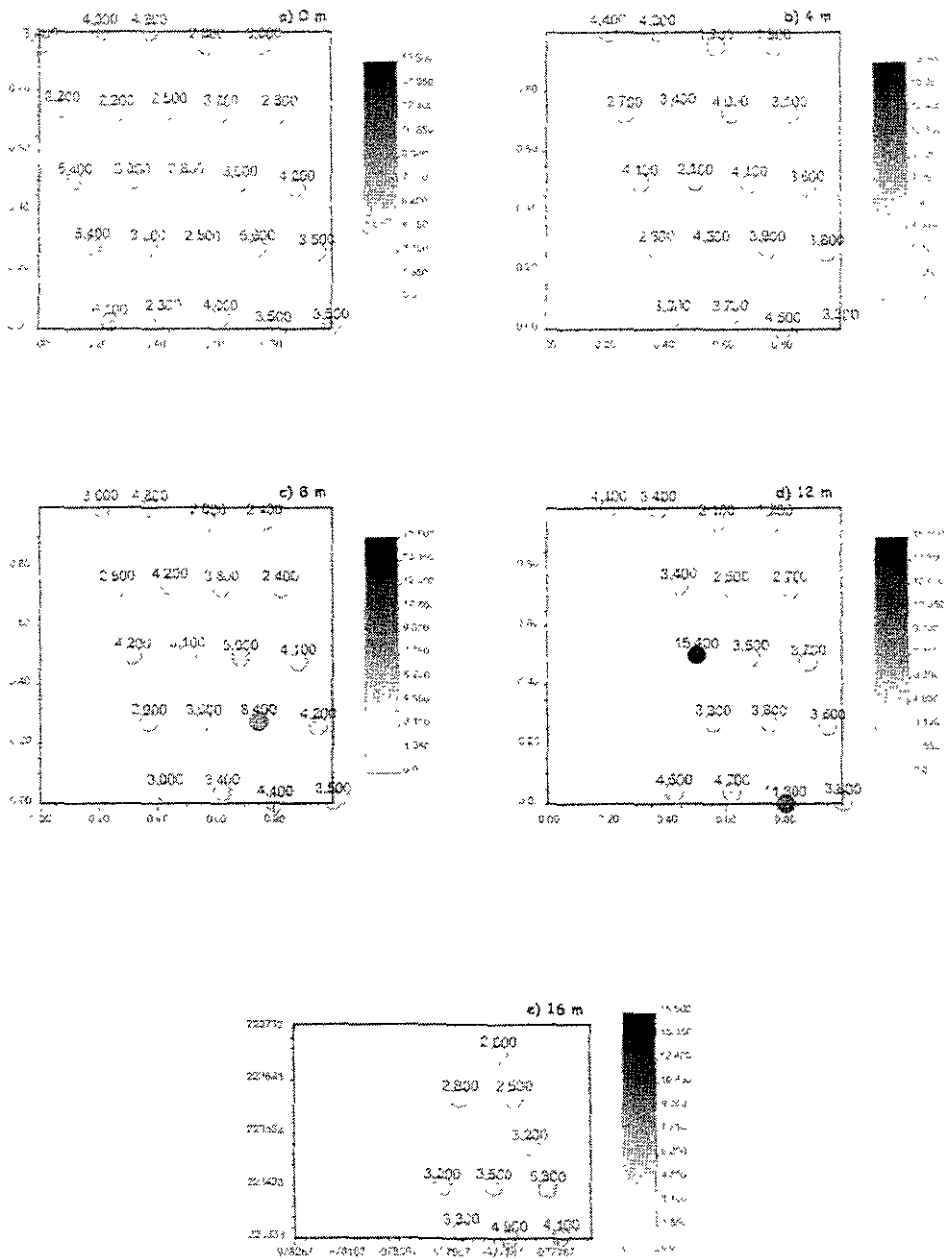


Figura 3.2.6 Mapa de datos para Mn por nivel de profundidad, Tercer Muestreo

región (sección 3.2). Esta identificación, junto con el análisis del resto de los datos, permitió observar que todos los niveles presentaban comportamiento no uniforme en la malla del muestreo. Se optó por buscar direcciones que indicarán la continuidad espacial de los datos y en esa dirección se calculó el variograma al cual se le agregó un ángulo de tolerancia de 20 ° a 30 ° según el caso. El rango a través del cual se eligieron las direcciones fue de 0° a 180°.

Los variogramas experimentales que finalmente se utilizaron para proponer algún variograma teórico fueron aquellos que tuvieran un comportamiento creciente y mayor número de pares o parejas de valores para su construcción. Una vez elegidos estos variogramas, se observó el comportamiento de los datos en el origen y se consideró que el mejor modelo era uno de tipo gaussiano para explicar el comportamiento de la información, ver Tablas 3.3 a 3.5. De lo anterior se desprende una característica general que presentan en este tipo de variables y es que el mejor modelo para los metales es de tipo gaussiano (Isaaks y Srivastava, 1989).

Posteriormente, con los modelos de los variogramas seleccionados, se utilizaron para interpolar la zona mediante la técnica de kriging y se construyeron los mapas de distribución espacial de los contaminantes en cada nivel. Esto se hizo para cada uno de los 3 muestreos efectuados en la zona del estudio. Los mapas resultantes de este proceso y de su desviación estándar se encuentran en las Figuras 3.3.1 a 3.3.8 y 3.4.1 a 3.4.8 respectivamente. De esto se hablará en la sección 3.4.3.

Para evaluar la confiabilidad de los resultados anteriores, se realizó el análisis de validación cruzada (Isaaks y Srivastava, 1989, Deutsch y Journel, 1998) y se presentan resultados en la sección 3.4.5.

Toda esta parte del análisis se realizó con el programa GEOEAS, versión 1.2.1.

El nivel de 16 m de profundidad no fue considerado en este análisis debido a que cuenta con poca información y a su participación espacial muy dispersa.

3.4.3 Variogramas

Como se mencionó en el Capítulo 2, el comportamiento del variograma permite conocer la relación que existe entre la estructura espacial y la precisión de la estimación. Debido a que no es posible utilizar un variograma experimental de manera directa, lo que se hace en general es proponer algún modelo teórico que refleje la información de los datos. Mucho tiene que ver el comportamiento que tiene el variograma en el origen.

Las Tablas 3.3 a 3.5 presentan los modelos teóricos y parámetros asociados que se utilizaron para cada conjunto de datos. El modelo que se aplicó con mayor frecuencia fue el gaussiano, de hecho en el Primer y Tercer Muestreo todos son de este tipo, mientras que para el Segundo, también se utilizó el exponencial.

A continuación se describe estos modelos en función de sus parámetros, es decir, meseta, rango y ángulo.

a) Primer Muestreo

Para la variable Fe, el modelo que se utilizó fue el de tipo gaussiano. El valor de meseta variaba en función de la profundidad, el de mayor valor es a los 12 m que también es el de mayor variabilidad, mientras que el de menor es a los 4 m en ambos casos, ver Tabla 3.3.

Con lo que respecta a los rangos, la situación es muy similar a la meseta, es decir, cambios en función de la profundidad, siendo mayor a los 12 m y la menor a los 4 m.

Variable	Profundidad (m)	Angulo (°)	Meseta	Rango
Fe	0	60	17.000	350
		130	14.000	400
	4	150	1.500	600
		160	1.300	600
	8	130	25.000	450
		150	25.000	450
	12	120	700.000	1000
		150	800.000	1000
Mn	0	80	1.000	300
		100	5.000	850
		110	5.000	900
	4	0	0.350	400
	8	0	8.200	400
		160	4.000	400
	12	130	90.000	1000
		150	100.000	1100

Tabla 3.3 Variogramas de tipo gaussiano por nivel de profundidad, Primer Muestreo

En cuanto a los ángulos, para los niveles 4, 8 y 12 m varían entre 120° a 160°, que no es significativo si el ángulo de tolerancia va de 20° a 30°. También a los 0 m, hay un modelo apropiado con 130° y hay otro de 60° lo cual habla de menor uniformidad, o mayor anisotropía para este nivel.

Al igual que en Fe, el Mn al nivel de 12 m su valor de meseta es el mayor y el menor a los 4 m y la variabilidad también se da de la misma manera, ver Tabla 3.3.

Para los rangos, se puede ver 2 grupos, el primero que va de 300 a 400 en los niveles de 0, 4 y 8 m, y el segundo de 900 a 1100 en los niveles de 0 y 12 m. Para los ángulos se podría hacer lo mismo, definir 2 grupos: los que tiene 0° (niveles 4 y 8 m) y los que van de 80 a 160° (0, 8 y 12 m).

b) Segundo Muestreo

Cuando uno propone algún modelo, el objetivo principal que se debe tener en mente es el de tratar de captar la mayor característica espacial de la variable y una parte muy importante en la selección del modelo tiene que ver con el comportamiento que tiene en el origen. Modelos con un desarrollo parabólico en el inicio deberían ser usados solamente para fenómenos que sean considerados continuos, en otros casos el modelo lineal funciona adecuadamente (Goovaerts, 1997). En el caso de metales o en general de concentraciones el modelo gaussiano funciona adecuadamente (Isaaks & Srivastava, 1989).

En este muestro se ajustaron 2 tipos de modelos gaussiano y exponencial, sin que alguno sea específico para una de las variables o para algún nivel. Mucho tuvo que ver el comportamiento que se observó en el variograma experimental, para algunos niveles de profundidad el comportamiento fue de tipo lineal y en particular el modelo exponencial se consideró adecuado.

Para la variable Fe, el modelo que se utilizó con mayor frecuencia fue el de tipo gaussiano, excepto para el nivel de 4 m en donde se ajustó el de tipo exponencial, ver Tabla 3.4.

Variable	Profundidad (m)	Modelo	Angulo (°)	Meseta	Rango
Fe	0	Gaussiano	0	26.000	500
		Gaussiano	40	26.000	500
		Gaussiano	170	11.000	150
	4	Exponencial	90	350.000	300
		Exponencial	120	650.000	400
		Exponencial	130	370.000	400
	8	Gaussiano	0	1500.000	400
		Gaussiano	130	1000.000	400
		Gaussiano	160	850.000	350
		Gaussiano	170	1700.000	450
	12	Gaussiano	80	1800.000	100
		Gaussiano	90	1800.000	230
		Gaussiano	140	1700.000	120
Gaussiano		150	1650.000	120	
Mn	0	Exponencial	60	1.600	550
		Exponencial	120	1.600	400
	4	Gaussiano	0	8.000	400
		Gaussiano	100	4.000	200
	8	Gaussiano	0	4.100	400
		Gaussiano	130	3.000	300
		Gaussiano	160	3.500	300
		Gaussiano	170	6.000	400
	12	Gaussiano	70	0.500	300
		Gaussiano	100	1.600	400
Gaussiano		130	2.400	400	
Gaussiano		170	2.300	450	
Ln(Fe)	4	Gaussiano	110	1.200	600
		Gaussiano	150	1.100	200
	8	Gaussiano	0	1.100	350
		Gaussiano	160	1.100	350
		Gaussiano	170	1.100	350
	12	Gaussiano	80	1.200	200
		Gaussiano	120	1.000	290
Gaussiano		130	0.800	200	

Tabla 3.4 Variogramas por nivel de profundidad, Segundo Muestreo

Con lo que respecta a la meseta, el nivel de 0 m presenta valores menores a 30; a 4 m de profundidad, este parámetro está entre 300 a 650 m, y en el resto varió de los 850 a 1800, siendo mayor a los 12 m. De lo anterior parecería que el valor de la meseta crece en función de la profundidad.

Con respecto al rango, los 3 primeros niveles varían de 300 a 500 m, mientras que a los 12 m va de 100 a 230.

Para el ángulo, en casi todos los niveles varían entre 80 a 170° a excepción del nivel 0 m pues contiene ángulos de 0° y 40°.

Al igual que la variable Fe, la mayoría de los modelos que se utilizaron para Mn son de tipo gaussiano, solo el nivel de 0 m fue modelado con los de tipo exponencial.

Si se observa el valor de las mesetas, estas están dentro de un rango de 0.500 a 8.000, siendo las más pequeñas a los 0 m, las mayores y más variables a los 4 m.

Los rangos se mantienen semejantes en las diferentes profundidades pues varían de 200 a 550 siendo el de mayor variabilidad al nivel de 4 m.

Para los ángulos, podría pensarse de nuevo en 2 grupos en cada nivel, definidos como aquellos que son menores o iguales a 70°, predominando el valor de 0°, y los que son mayores a 100° y definir con ello las zonas anisotrópicas de cada nivel.

Para la transformación $\ln(\text{Fe})$, se analizó los niveles de 4 a 12 m de profundidad y todos los modelos propuestos son de tipo gaussiano.

El valor de las mesetas no varió mucho pues está dentro del rango de 1 a 1.200.

En el parámetro de rango, el que tuvo mayor variabilidad es a los 4 m, de 200 a 600, siendo estos valores el mínimo y el máximo en general. En cuanto a los ángulos, el nivel de 8 m,

presenta uno de 0° y el resto mayor o igual a 160° , mientras que para los otros niveles los rangos de variación fueron similares.

c) Tercer Muestreo

Como se mencionó anteriormente, en este muestreo todos los modelos son gaussianos así que ha de suponerse ésto en cada uno de los siguientes apartados y tener en cuenta la información de la Tabla 3.5.

El valor de la meseta para Fe varía de 180 a 2600, siendo este último valor el único en este muestreo que se utilizó a los 12 m de profundidad. Si se ordenara el valor de las mesetas en función del rango en cada nivel, se obtendría el orden de 8, 0, 4 y 12 m.

Los rangos se podrían considerar similares, excepto a los 12 m que es de 150, mientras que el resto de los valores varió de 300 a 550.

Los ángulos varían de 60° a 160° , los niveles de 0, 4, y 8 m presentan diferencias similares entre ellos.

El rango de variación de los valores de meseta para la variable Mn, va de 0.600 a 3.200 siendo los más homogéneos a los 4 y 8 m y el más alejado se presenta a los 12 m.

Los rangos varían de nivel pero no con un patrón definido. El nivel que presenta mayor variabilidad es el de 4 m.

Para los niveles de 0, 4 y 8 m, se puede de nuevo definir 2 grupos en cuanto al valor de los ángulos. Estos grupos son: menores o iguales a 90° y los mayores o iguales a 120° , de igual manera que en el caso anterior esto podría definir los ángulos de anisotropía.

Variable	Profundidad (m)	Angulo (°)	Meseta	Rango
Fe	0	60	180.000	300
		130	200.000	450
	4	80	520.000	420
		130	380.000	550
		160	220.000	400
	8	80	240.000	370
		110	245.000	360
	12	60	2600.000	150
	Mn	0	40	2.200
80			1.600	200
160			1.450	280
4		0	0.800	300
		120	0.800	250
		150	0.600	400
8		90	3.200	230
		150	3.000	230
12		170	2.400	450
		100	1.600	400
		120	2.000	400
		130	2.400	450
Ln(Fe)		0	60	0.800
	80		1.600	380
	120		1.400	420
	150		0.800	400
	4	80	1.200	320
		110	3.000	600
		130	2.700	650
		170	1.400	475
	8	110	1.550	340
		120	1.300	500
		130	1.300	400
	12	80	1.200	200
		120	0.950	240

Tabla 3.5 Variogramas de tipo gaussiano por nivel de profundidad, Tercer Muestreo

A diferencia del segundo muestreo, aquí se llevó a cabo esta transformación para los datos en los 4 niveles.

Para el parámetro de meseta de $\ln(\text{Fe})$ se obtuvo que, el rango de variación está entre 0.800 a 3.000 y si se observan las magnitudes en función de la profundidad, éstas van creciendo desde el nivel 0 m hasta el nivel 8 m y decrecen en los 12 m. La mayor variabilidad se presentó a 8 m.

En cuanto al rango, este varió de 200 a 650. El nivel de 4 m es el que presenta mayor variabilidad y se observaron los valores más altos. Para los niveles de 8 y 12 m dicho valor decrece, lo mismo que a los 0 m. Estos resultados parecen detectar la presencia del difusor.

Los ángulos presentan un patrón similar al rango en todos los casos.

3.4.4 Kriging

La idea de describir cada una de las variables por muestreo, es la de observar si los variogramas que se eligieron son adecuados para construir el Kriging de cada nivel de profundidad y para cada temporada. Si esto fuera el caso, entonces se tendría un análisis completo de cómo el difusor está dispersando los residuos por elemento, lo cual podría ayudar para ajustar el modelo del difusor bajo estudio.

El número de valores predcidos para cada variable fue de 100, 10 valores para cada dirección y considerando el origen en (-978267, 223333), con incrementos de 56.556 y 40.778 en Longitud y Latitud respectivamente. El tipo de Kriging que se utilizó fue ordinario y en bloques de 2x2 y el programa con el que se calculó fue GEOEAS (Versión 1.2.1).

Para explicar los resultados de este análisis será necesario tomar en cuenta 3 gráficas, la primera es donde se presentan los datos muestreados (Figuras 3.2.1 a 3.2.6), la segunda será la de Kriging (Figuras 3.3.1 a 3.3.8) y por último la gráfica la Desviación Estándar del Kriging (Kriging(DE), Figuras 3.4.1 a 3.4.8). Se debe tomar en cuenta que la zona a la que se hará

referencia será aquella que fue muestreada exclusivamente ya que en las de Kriging y Kriging(DE) hay partes que no deben estar y que por la interpolación y por el tipo de Kriging utilizados fueron calculadas.

Posteriormente se llevó a cabo un proceso de validación cruzada y se estimaron los errores asociados a Kriging. Con dichos valores se construyeron histogramas y gráficas en papel normal para analizar su comportamiento (Figuras 3.5.1 a 3.5.8) y evaluar la confianza de los resultados.

a) Primer Muestreo

Para la variable Fe, si se observan las Figuras 3.2.1 y 3.3.1 se puede ver que las concentraciones altas se encuentran al Norte y al Sur de la zona muestreada y valores bajos en el centro, recuperando la forma y dirección del difusor. También la acumulación del metal a los 12 m queda representada en dichas figuras.

Las imágenes de la Figuras 3.4.1 muestran valores cercanos a 0 en los niveles 4, 8 y 12m, lo cual habla de la poca diferencia que hay entre los valores observados y los predichos. Así que se puede pensar que se obtuvo un buen ajuste y que los variogramas y el kriging utilizados fueron bien seleccionados.

Para la variable Mn, los valores estimados por kriging muestran un comportamiento similar al que se describió anteriormente, es decir, valores altos al Norte, Sur y Noroeste de la malla principalmente en los niveles 4, 8 y 12 m, ver Figura 3.3.2.

Las imágenes que se muestran en la Figura 3.4.2 también presentan valores bajos en la zona muestreada.

b) Segundo Muestreo

Si se observan las imágenes de la Figura 3.3.3 (a) a (c), para la variable Fe, se puede ver que las concentraciones de mayor valor se encuentran a lo largo de la imagen, partiendo en dos cada una de éstas. Al nivel de 0 m de profundidad (Figura 3.3.3(a)) se puede ver que hay mayor variabilidad, pero a los 12 m (Figura 3.3.3(d)) se observa que hay una alta concentración de Fe hacia el Sur. Dicho comportamiento refleja la presencia de lluvia en la superficie y la poca corriente que tiene el mar al menos en esta zona.

Las imágenes de la desviación estándar del kriging de Fe ($KSD(Fe)$) (Figura 3.4.3), al igual que las anteriores, muestran poca variabilidad en la zona muestreada.

Para la variable Mn se comentó que las concentraciones más altas se encuentran alrededor del difusor, esto queda bien representado en las imágenes de kriging como se ve en la Figura 3.3.4. A los 12 m de profundidad, en la parte norte se ve que el Mn queda ahí, con concentraciones altas lo cual también queda representado en la imagen, ver Figura 3.3.4(d).

En cuanto a las imágenes de la desviación estándar de kriging de Mn ($KSD(Mn)$), reflejan un buen ajuste pues en la zona muestreada los valores son cercanos a 0, ver Figura 3.4.4.

Es importante recordar que los valores que se observaron en esta variable fueron muy planos en el sentido de que se contaba con valores enteros y quizás el kriging que se utilizó no haya sido el más adecuado.

Para el $\ln(Fe)$, las imágenes resultaron ser muy parecidas a las de Fe (ver Figura 3.3.3 y 3.3.5) en cuanto a los puntos de mayor concentración pues quedan bien representados. En cuanto al resto de la información, para $\ln(Fe)$ la superficie es más homogénea con el cambio de escala porque los valores no son tan extremos.

Las imágenes de la desviación estándar del kriging de $\ln(\text{Fe})$ ($\text{KSD}(\ln(\text{Fe}))$), presenta menor variabilidad que las de Fe , lo cual indica que se tuvo un mejor ajuste y que este se debió al cambio de escala que tuvieron los datos (ver Figura 3.4.5).

c) Tercer Muestreo

Para este muestreo, las concentraciones altas de Fe se encuentran en la parte Norte de las imágenes de kriging en los niveles de 0 y 4 m; a los 8 m, al Norte y Sureste y a los 12 m en la parte media y un poco al Sureste (Figura3.3.6).

Las imágenes de la $\text{KSD}(\text{Fe})$, muestran valores cercanos a 0 en la zona muestreada a excepción del último nivel que presenta mayor variabilidad (Figura 3.4.6).

Para la variable Mn , los valores de mayor concentración se ubican en la dirección Sureste para los niveles 0, 4 y 8 m. A los 12 m, el comportamiento es muy similar al Fe , es decir, valores altos en la parte media y un poco hacia el Sureste (Figura3.3.7).

En cuanto a las imágenes de $\text{KSD}(\text{Mn})$, en la zona muestreada los valores son cercanos a 0, así que las estimaciones hechas para este muestreo fueron buenas en cuanto a que la variabilidad fue baja (Figura 3.4.7).

Al igual que en el muestreo anterior, la variable $\ln(\text{Fe})$ tiene un comportamiento muy similar al Fe (Figura 3.3.8), y en general se tiene una interpolación más homogénea al compararla con Fe .

Las imágenes de $\text{KSD}(\ln(\text{Fe}))$ son similares en cuanto a comparación gráfica y en cuanto a variabilidad, esta es menor por el cambio de escala (Figura 3.4.8).

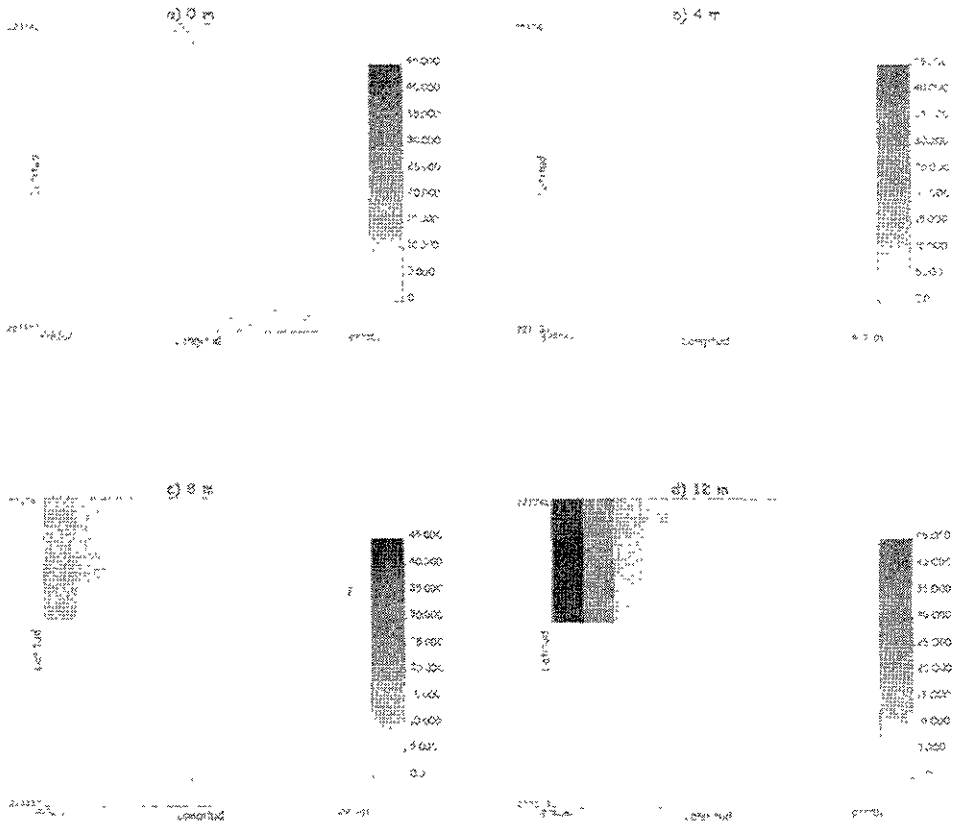


Figura 3.3.1 Kriging de Fe por nivel de profundidad Primer Muestreo

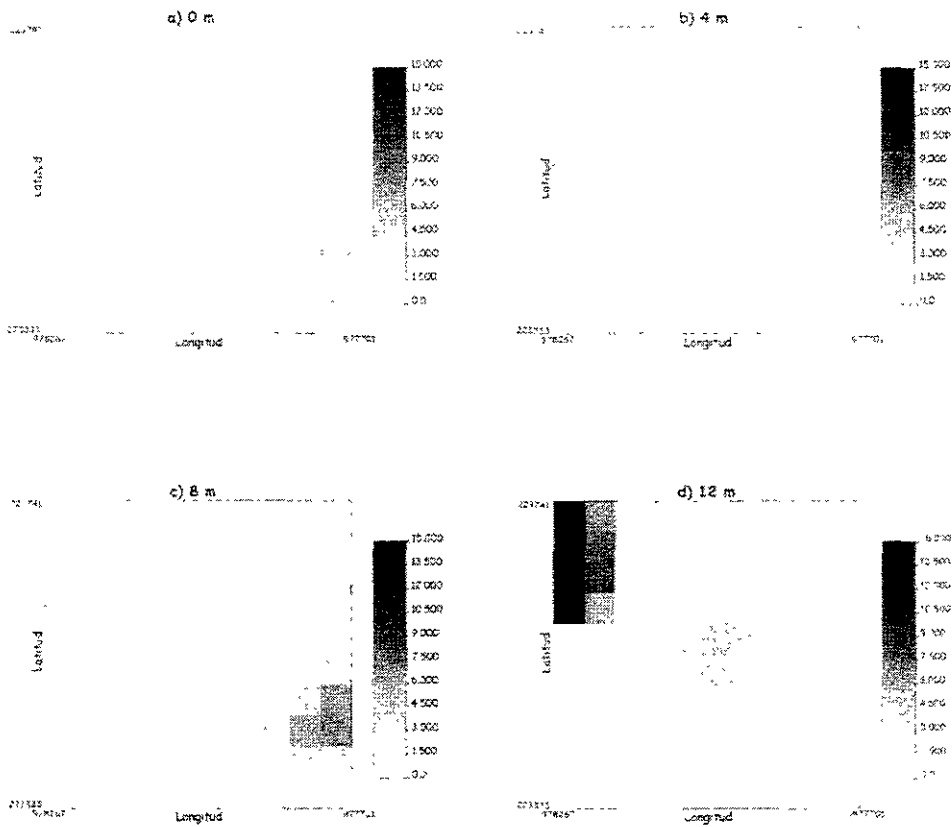


Figura 3.3.2 Kriging de Mn por nivel de profundidad, Primer Muestreo

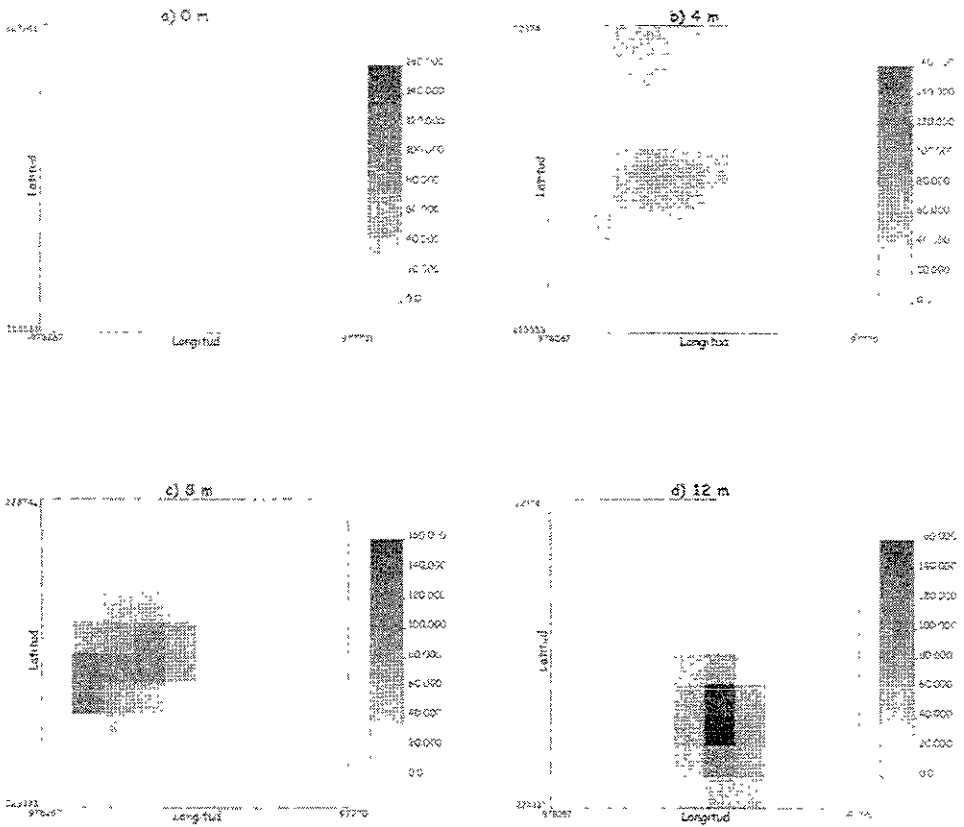


Figura 3.3.3 Kriging de Fe por nivel de profundidad, Segundo Muestras

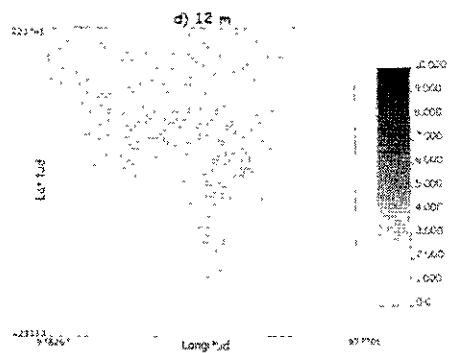
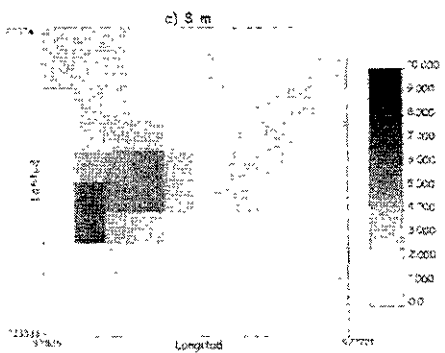
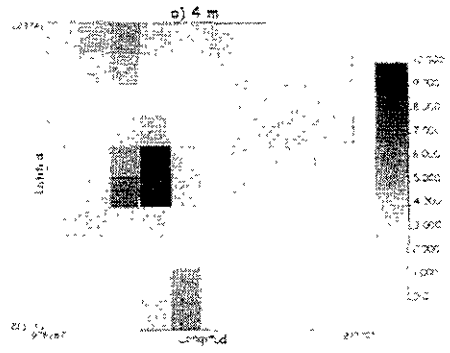
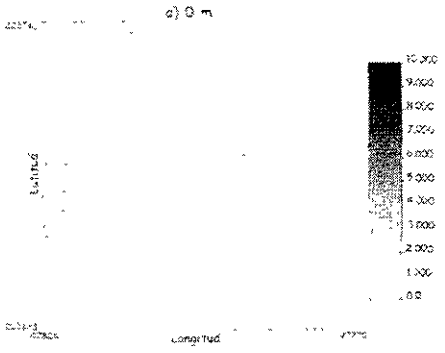


Figura 3.3.4 Kriging de Mn por nivel de profundidad, Segundo Muestreo

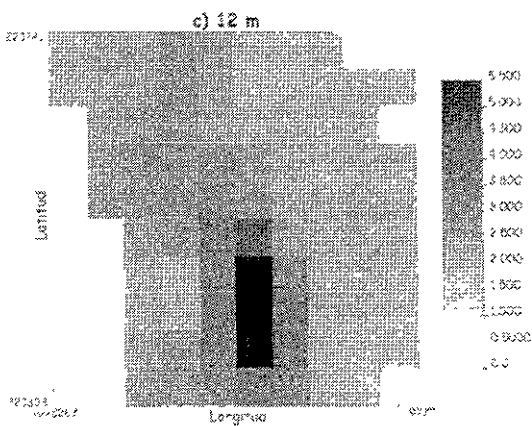
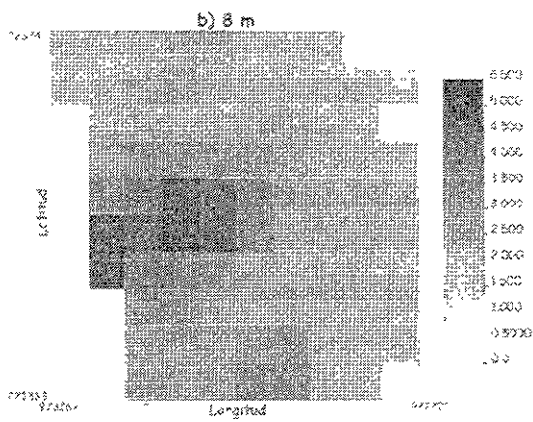
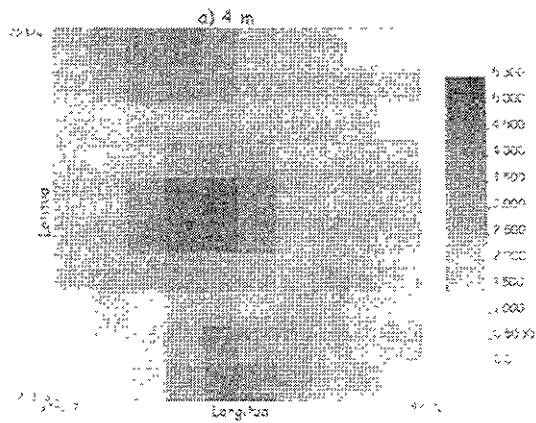


Figura 3.3.5 Kriging de In(Fe) por nivel de profundidad, Segunda Muestra

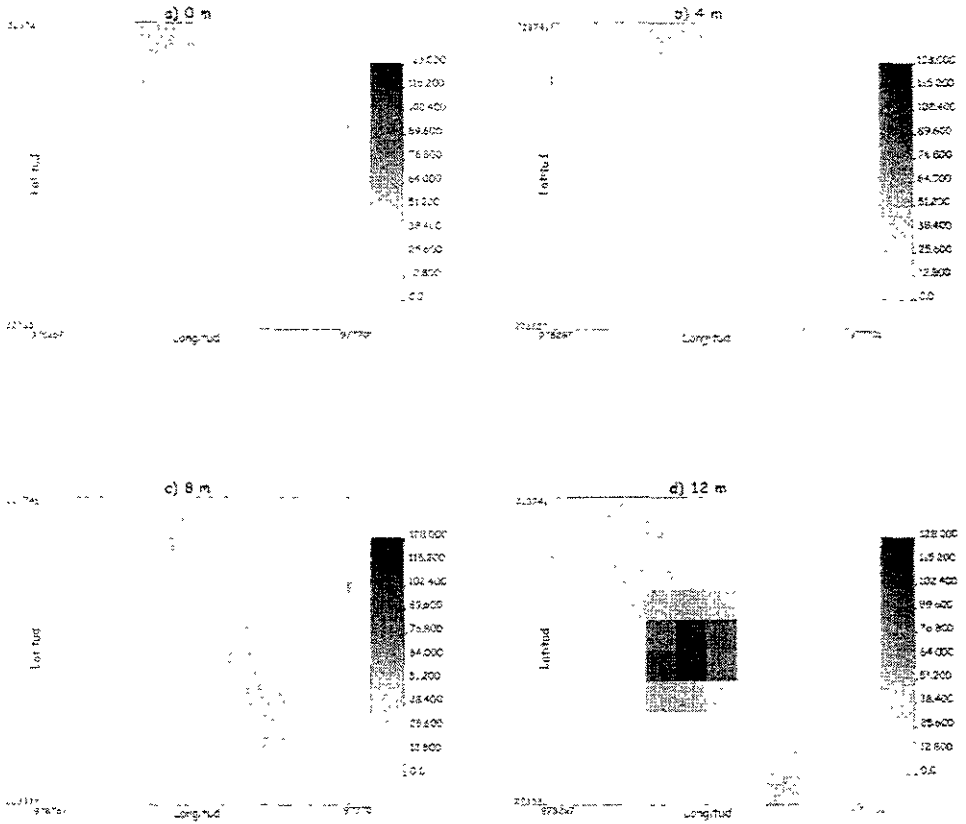


Figura 3.3.6 Kriging de Fe por nivel de profundidad, Tercer Muestreo

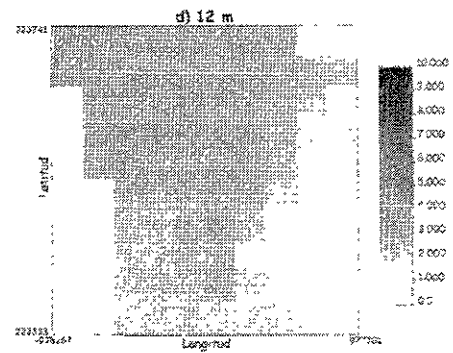
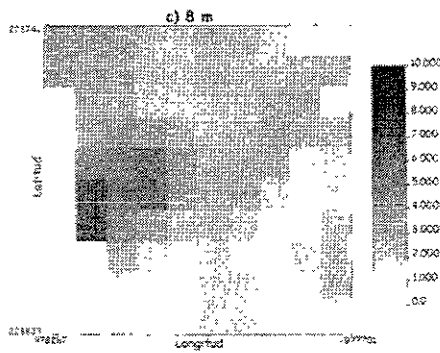
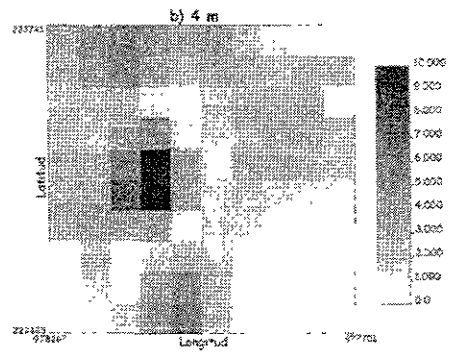
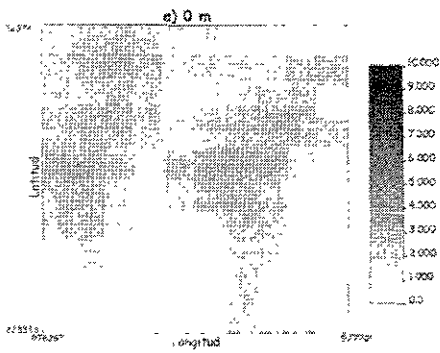


Figura 3.3.7 Kriging de Mn por nivel de profundidad, Tercer Muestreo

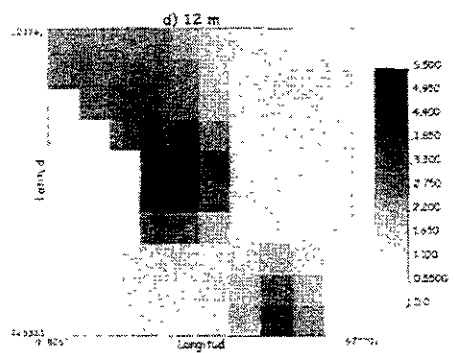
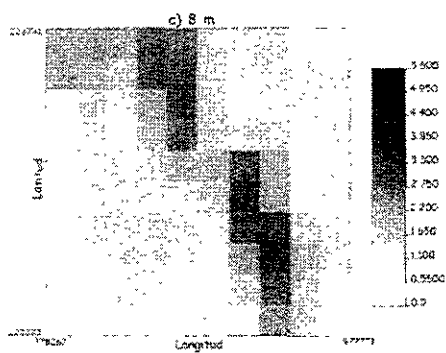
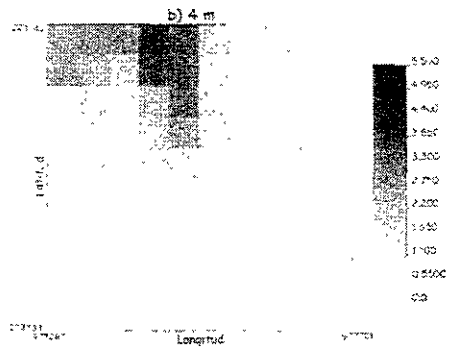
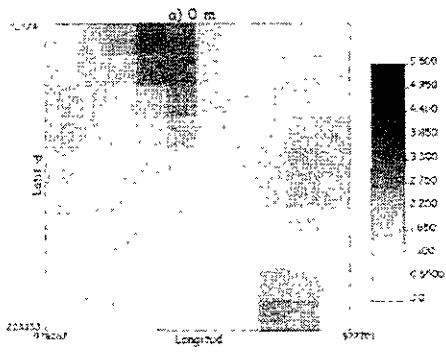


Figura 3.3.8 Kriging de ln(Fe) por nivel de profundidad, Tercer Muestreo

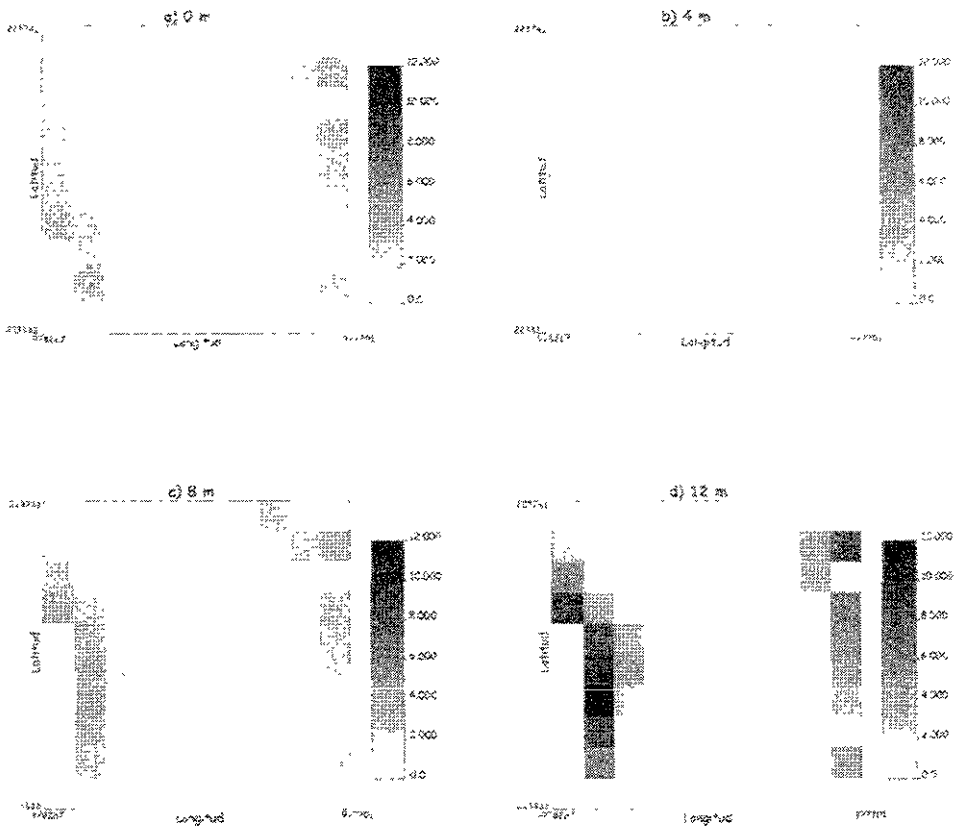


Figura 3.4.1 Kriging (DE) de Fe por nivel de profundidad, Primer Muestreo

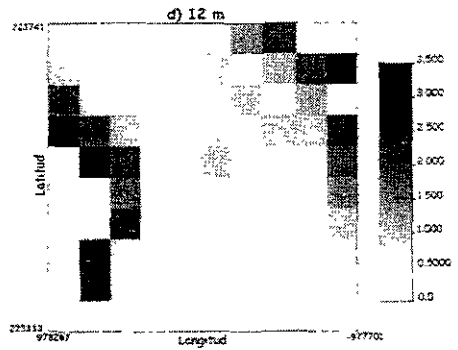
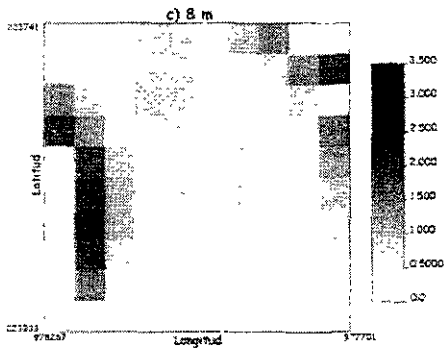
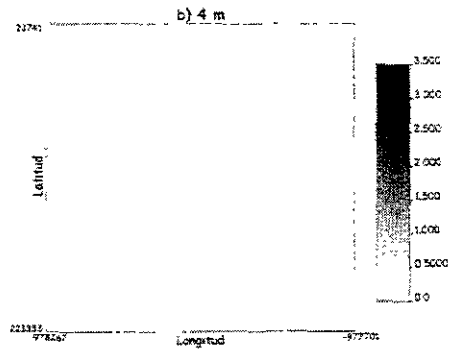
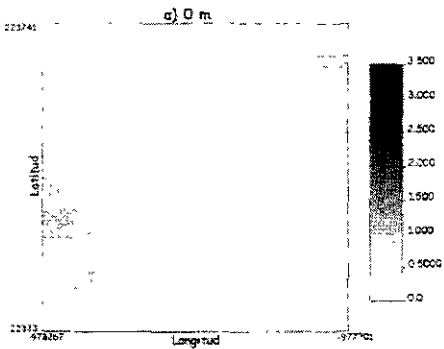


Figura 3.4.2 Kriging (DE) de Mn por nivel de profundidad, Primer Muestreo

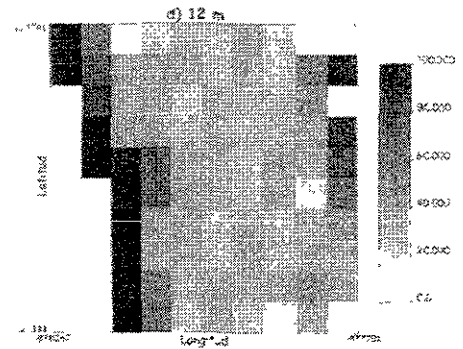
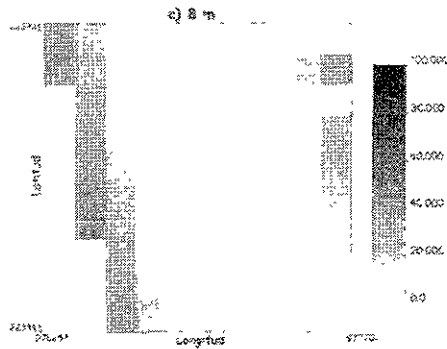
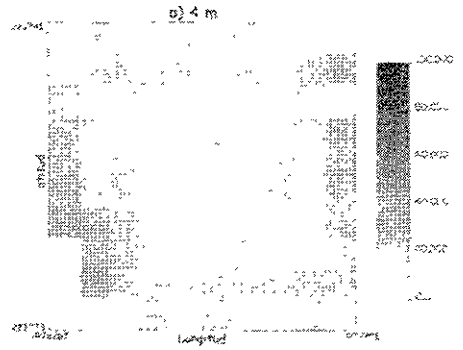
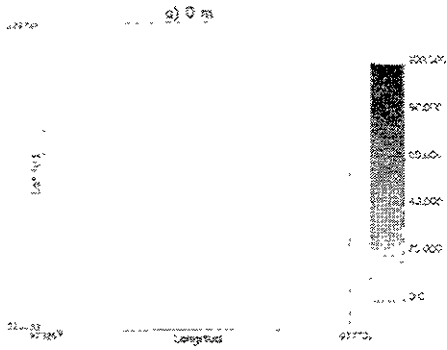


Figura 3.4.3 Kriging (DE) de Fe por nivel de profundidad, Segundo Muestreo

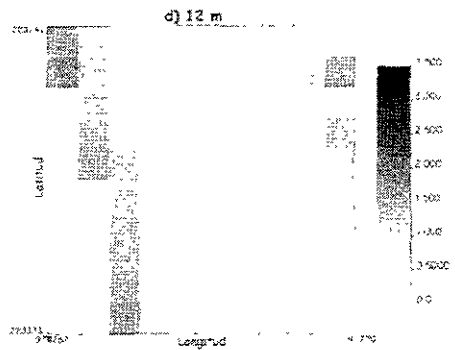
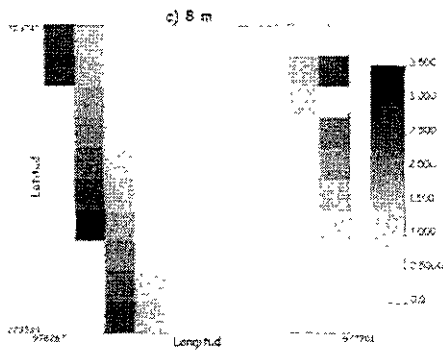
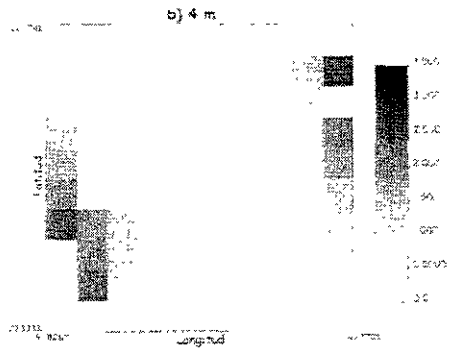
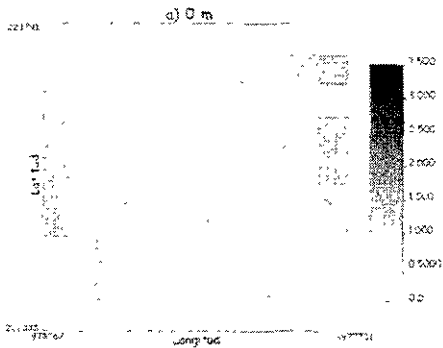


Figura 3.4.4 Kriging (DE) de Mn por nivel de profundidad, Segunda Muestreo

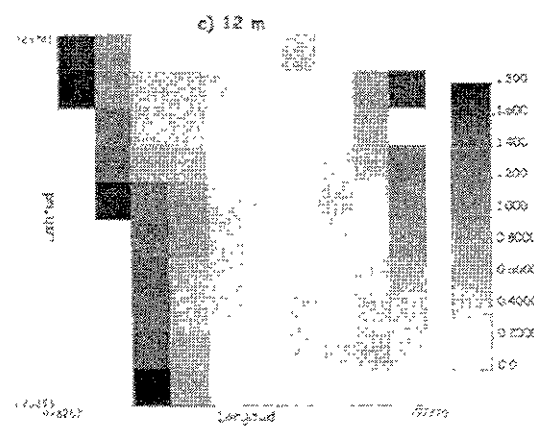
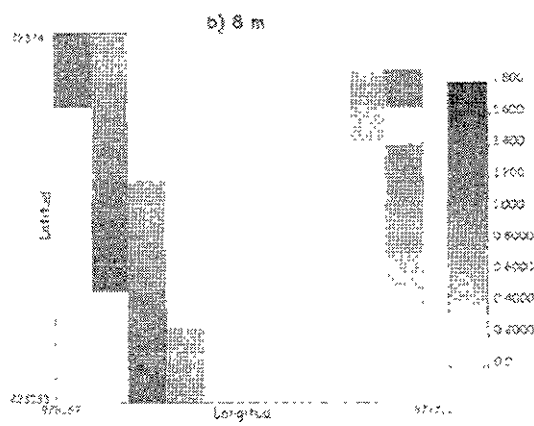
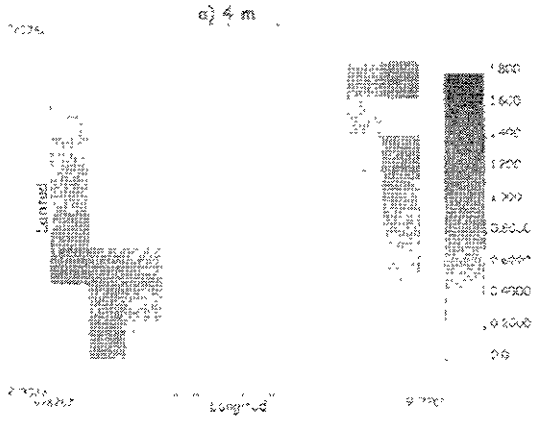


Figura 3 4.5 Kriging (DE) de ln(Fe) por nivel de profundidad, Segundo Muestreo

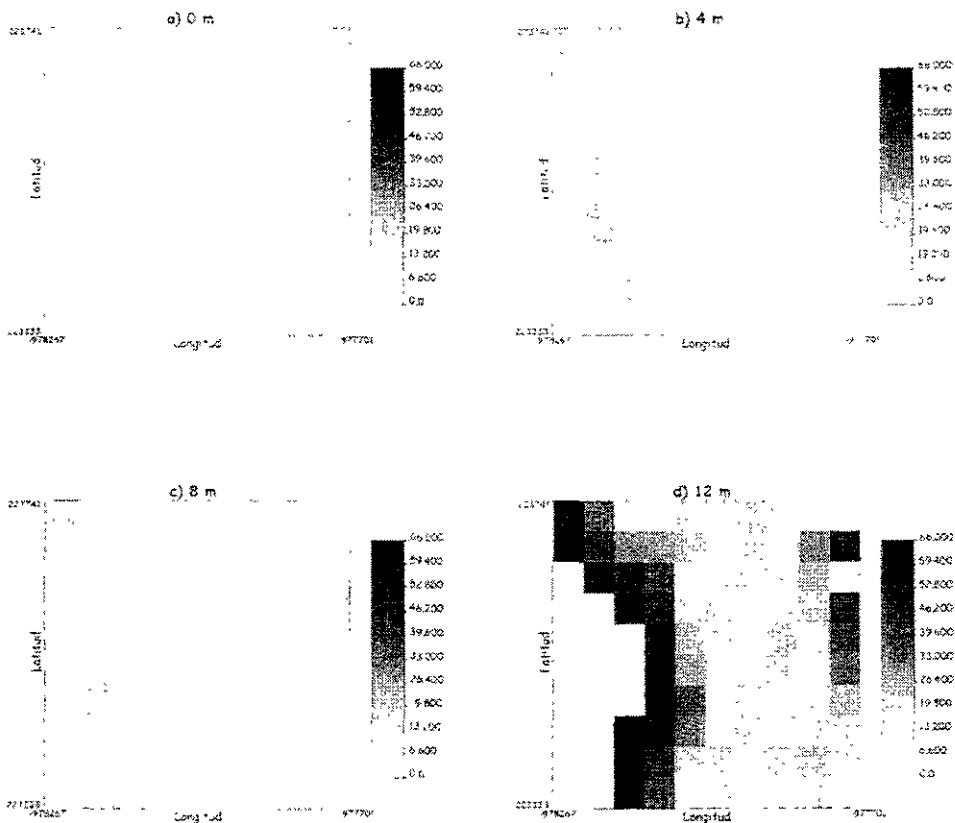


Figura 3.4.6 Kriging (DE) de Fe por nivel de profundidad, Tercer Muestreo

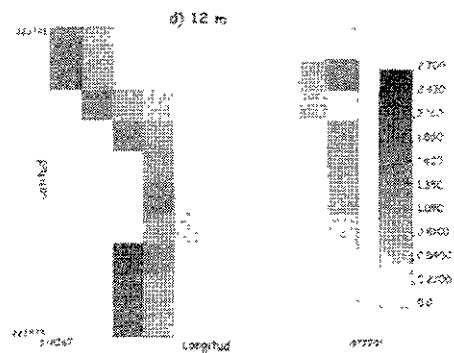
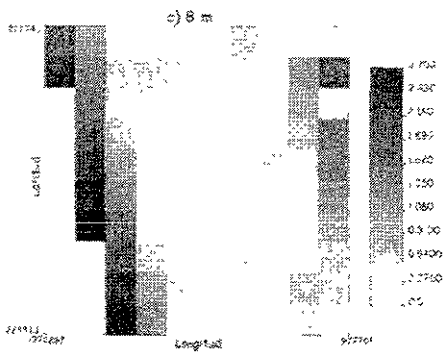
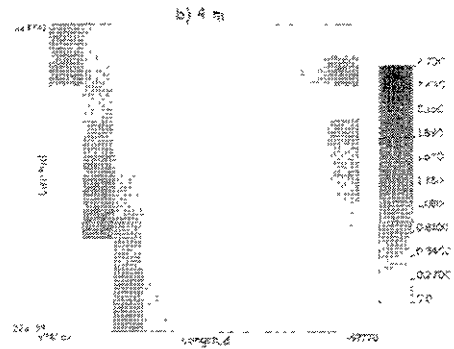
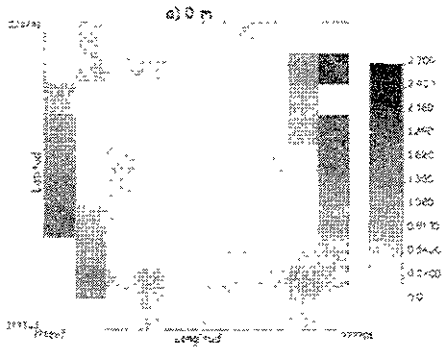


Figura 3.4.7 Kriging (DE) de Min por nivel de profundidad, Tercer Muestreo

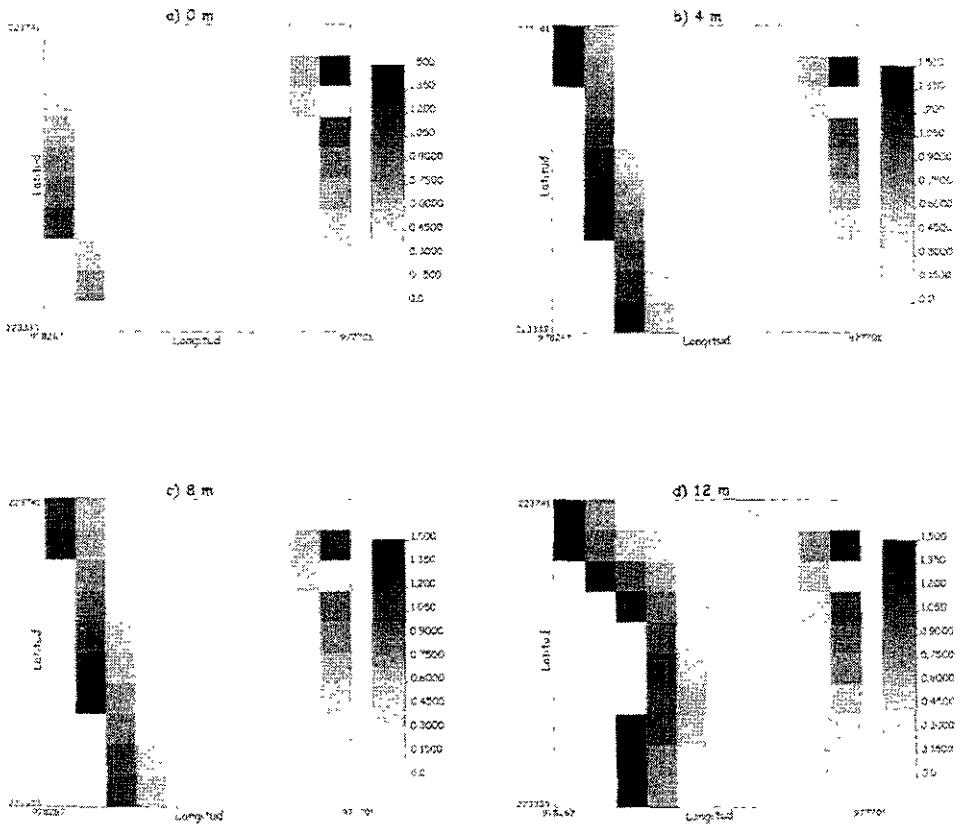


Figura 3.4.8 Kriging (DE) de $\ln(\text{Fe})$ por nivel de profundidad, Tercer Muestreo

3.4.5 Validación Cruzada

En el capítulo 2, se mencionó que la validación cruzada es un método que permite comparar los valores predichos con los datos originales con la idea de probar si los variogramas y los parámetros utilizados en ellos pueden reproducir la variabilidad espacial de las muestras. Este parte del trabajo muestra que teniendo una buena elección de variogramas a interpolar por medio de kriging, proporciona buenas estimaciones de los datos en general.

Para cada uno de los muestreos se realizó esta validación con las mismas consideraciones que se tuvieron para realizar la interpolación por medio de kriging. Se obtuvieron valores estimados para los puntos muestreados, la desviación estándar del kriging, el error estimado y el error estándar (error estimado dividido por la desviación estándar del kriging). Se construyeron histogramas y gráficas en papel normal para los errores estimados por ser estos el reflejo de la buena estimación que se tuvo en general de los datos. A continuación se describen estas gráficas.

a) Primer Muestreo

Los errores estimados para la variable Fe muestran que a los 4 m se tuvo el error de estimación más grande pues el rango va de -30 a 20 y son 3 puntos como se puede ver en la gráfica en papel normal (Figura 3.5.1 (d)). Los otros histogramas y gráficas en papel normal muestran un buen comportamiento en el sentido de que podrían considerarse aproximadas a una distribución normal, Figura 3.5.1.

Para Mn, los valores de los errores muestran un mejor comportamiento que los de Fe. El rango de mayor magnitud se da a los 8 m, de -8 a 6. A los 12 m, la mayoría de los datos se encuentran en el rango de -2 a 2 y sólo uno valor de 6 que habla de una sobre estimación de algún punto, Figura 3.5.2.

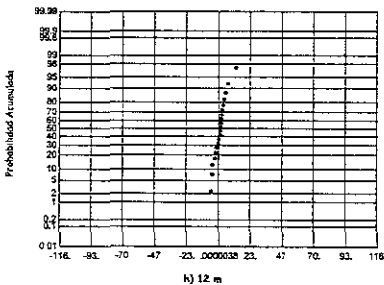
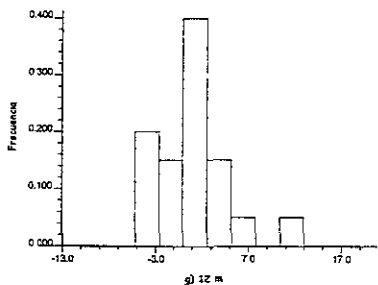
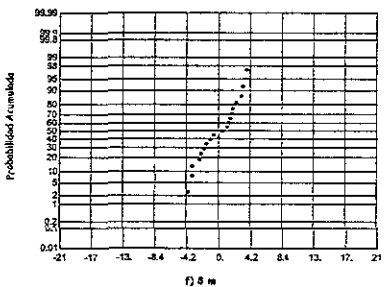
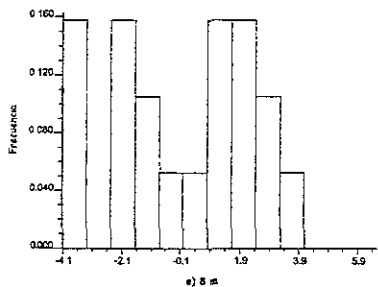
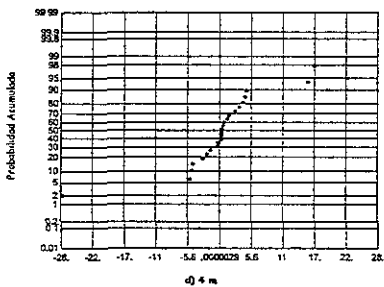
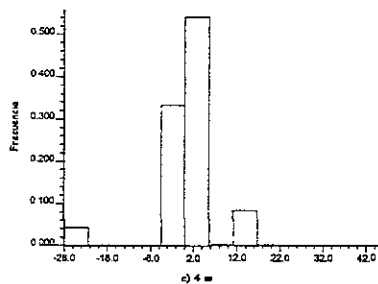
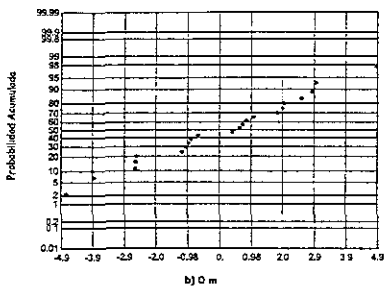
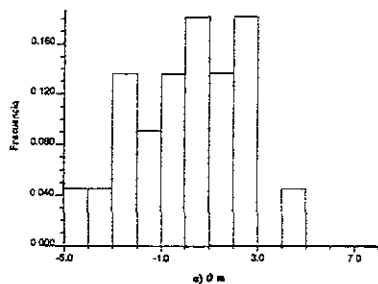


Figura 3.5.1 Histogramas y gráfica en papel normal para $e(\text{Fe})$ por nivel de profundidad, Primer Muestreo

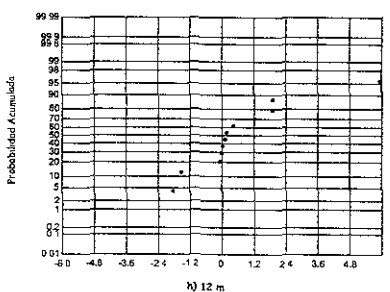
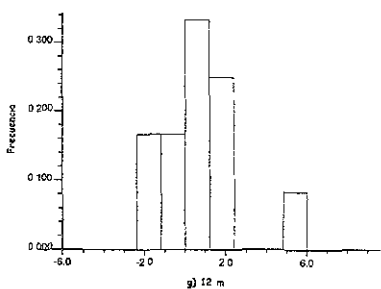
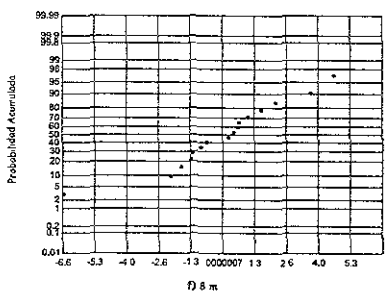
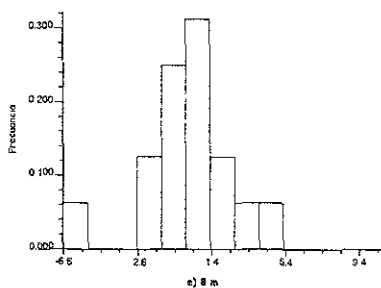
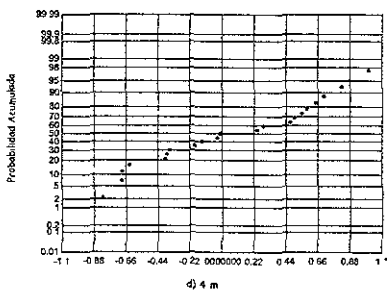
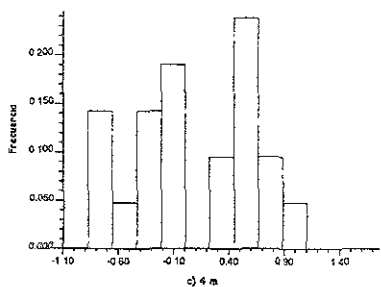
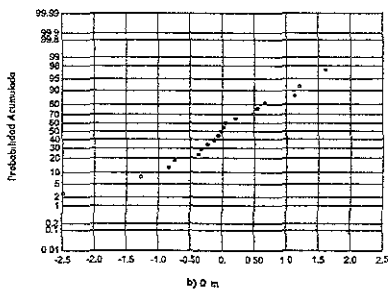
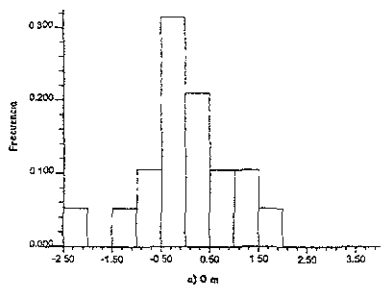


Figura 3.5.2 Histogramas y gráficas en papel normal para $e(Mn)$ por nivel de profundidad, Primer Muestreo

b) Segundo Muestreo

Para la variable Fe, los errores estimados tuvieron un rango de variabilidad alto, lo cual era de esperarse ya que es el muestreo que se caracteriza por ser variable. Los histogramas y gráficas en papel normal no muestran un mal comportamiento a excepción del nivel de 12 m que presenta un valor subestimado menor a -100, Figura 3.5.3.

Para los errores estimados de Mn, al nivel de 0 m, el histograma presenta frecuencias y valores muy estáticos, sin embargo la gráfica en papel normal muestra continuidad en los valores y no un comportamiento escalonado que es la característica principal de esta variable en este muestreo. El resto de los niveles de profundidad tiene un buen comportamiento, Figura 3.5.4.

Los errores estimados del $\ln(\text{Fe})$ muestran varias modas en los valores -1 y 1, a los 4 m (Figura 3.5.7 (a) y (b)). A los 8 m, el valor de la moda se recorre hacia la derecha si se compara con el histograma de Fe (Figura 3.5.5 (c) y Figura 3.5.3 (e) respectivamente). El rango de los datos es menor que en Fe, lo cual es lógico por el cambio de escala que hubo al transformar los datos. Las gráficas en el papel normal podrían considerarse bien portadas, Figura 3.5.5.

c) Tercer Muestreo

Al igual que en el muestreo anterior, los errores de estimación del Fe presentan rangos de -40 a 20 a excepción del nivel de 12 m que va de -120 a 40 aproximadamente. Las gráficas en papel normal tienen un comportamiento similar a excepción del nivel de 8 m, que presenta un cambio en el rango de -1 a 0 y en los otros niveles se presenta de -2 a -1, Figura 3.5.6.

Los errores estimados del Mn presentan un buen comportamiento en general, sólo a los 12 m el rango es mayor pues va de -10 a 5 aproximadamente, mientras que en el resto va de -3 a 2, Figura 3.5.7.

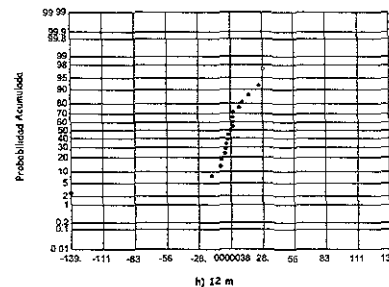
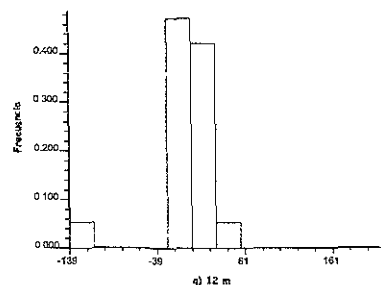
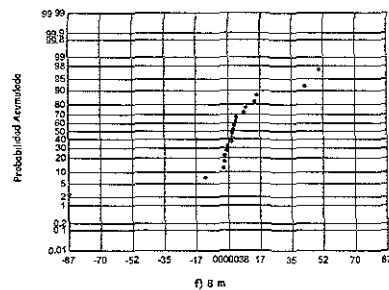
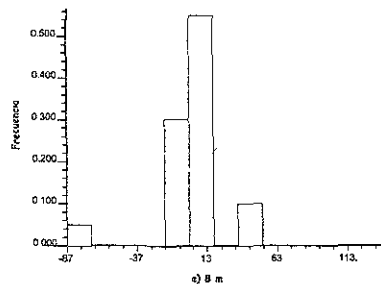
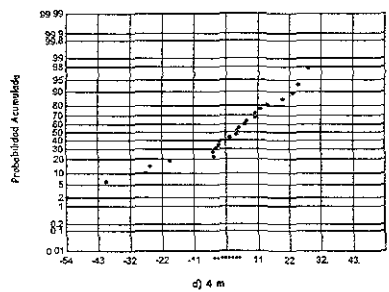
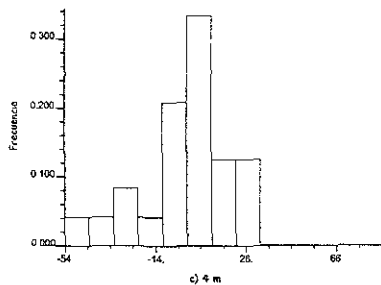
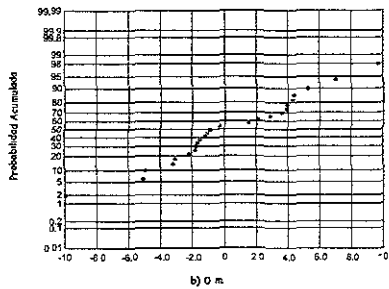
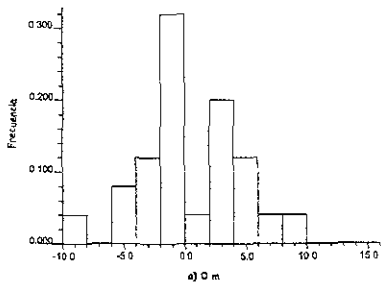


Figura 3.5.3 Histogramas y gráficas en papel normal para $e(\text{Fe})$ por nivel de profundidad, Segundo Muestreo

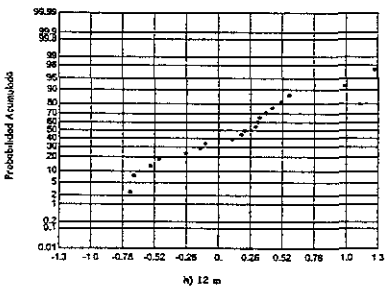
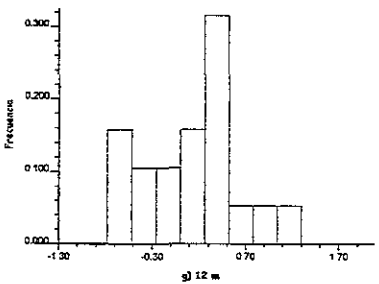
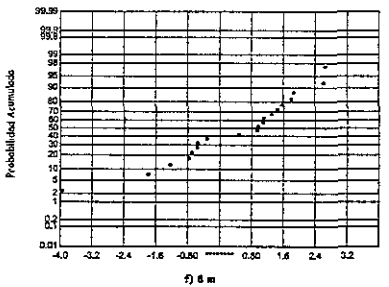
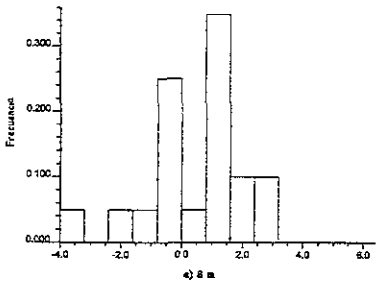
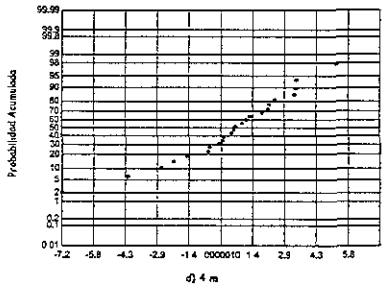
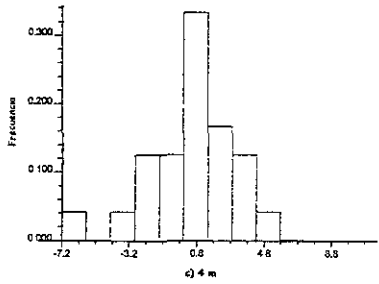
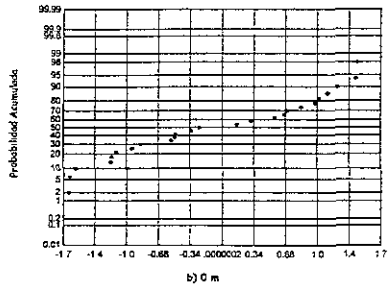
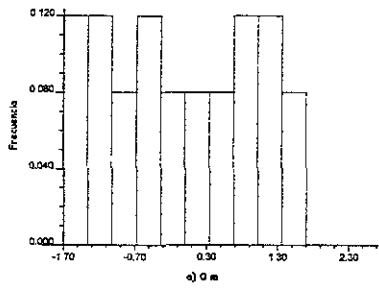


Figura 3.5.4 Histogramas y gráfica en papel normal para $e(Mn)$ por nivel de profundidad, Segundo Muestreo

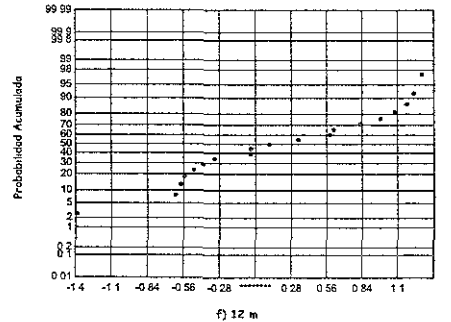
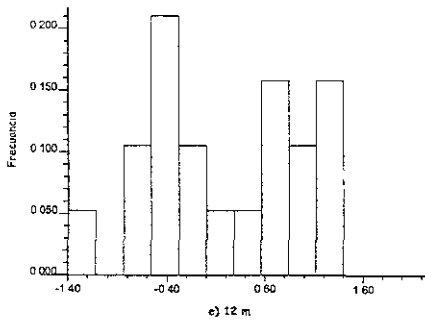
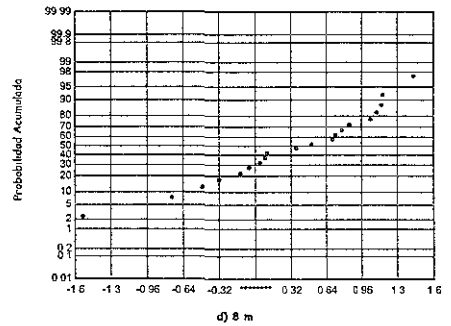
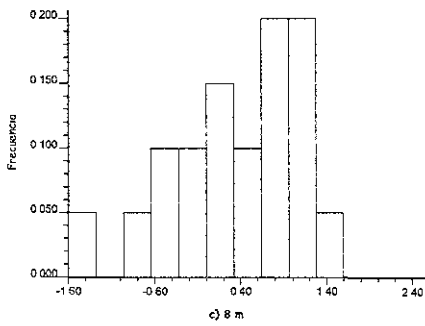
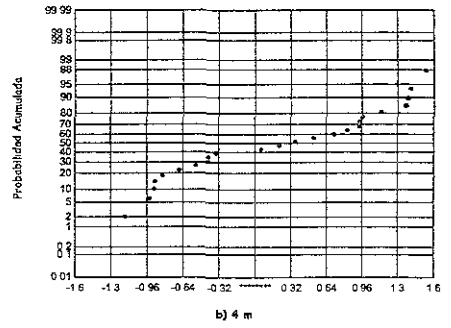
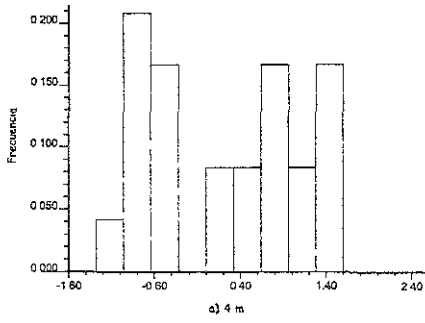


Figura 3.5.5 Histogramas y gráfica en papel normal para $z(\ln(\text{Fe}))$ por nivel de profundidad, Segundo Muestreo

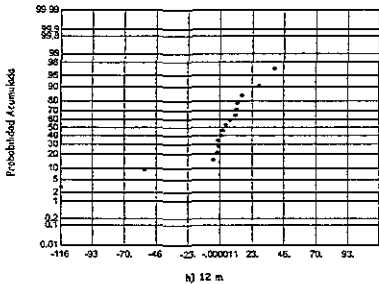
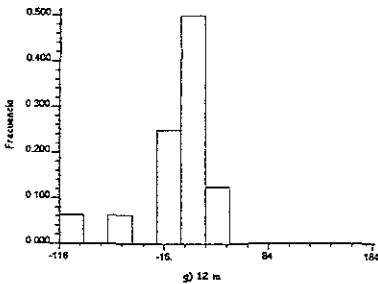
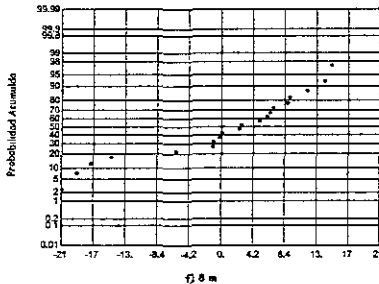
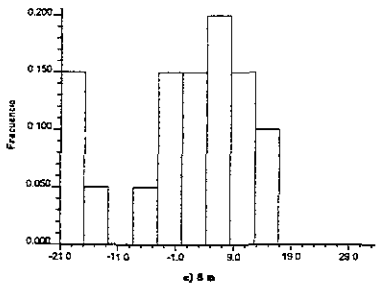
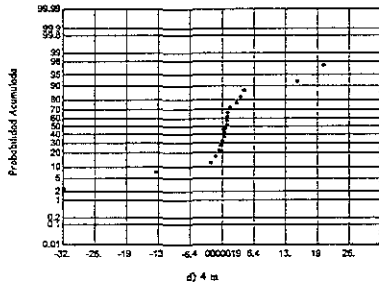
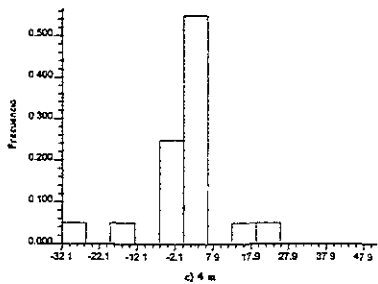
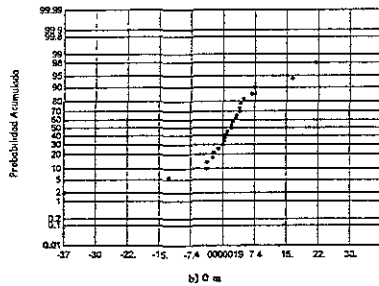
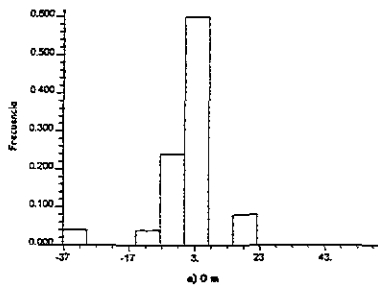


Figura 3.5.6 Histogramas y gráficas en papel normal para $e(\text{Fe})$ por nivel de profundidad, Tercer Muestreo

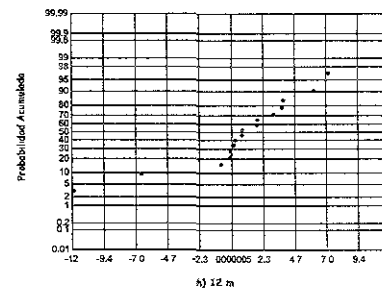
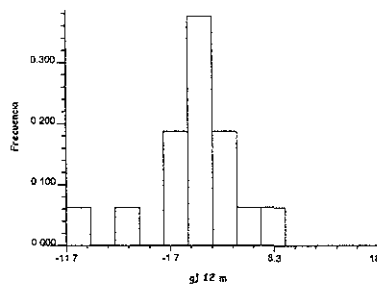
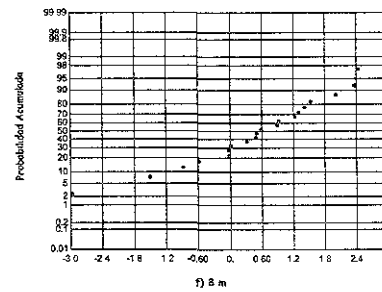
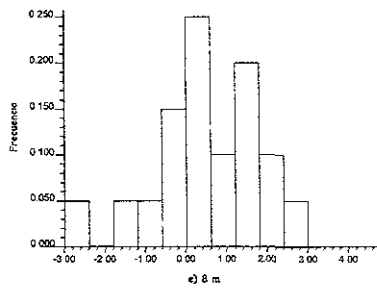
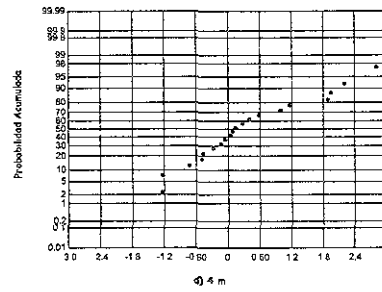
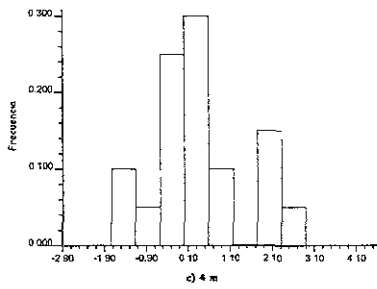
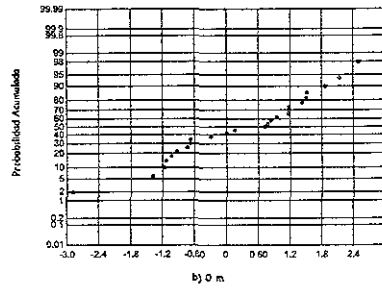
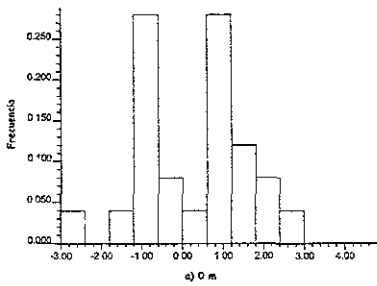


Figura 3.5.7 Histogramas y gráficas en papel normal para $e(Mn)$ por nivel de profundidad, Tercer Muestreo

Para la variable $\ln(\text{Fe})$, los errores estimados tienen un mejor comportamiento que en el muestreo anterior, pues el cambio de escala, al menos en este muestreo, ayudó a tener una mejor estimación que se refleja en las gráficas de estos errores, Figura 3.5.8.

3.5 Análisis Geoestadístico en \mathbb{R}^3

3.5.1 Descripción General del Análisis Estructural

Para esta parte del análisis, se tomó en cuenta los resultados que se obtuvieron en \mathbb{R}^2 como fueron las direcciones de los variogramas y la localización espacial de los valores extremos, entre otras cosas.

Antes de comenzar el análisis hubo necesidad de transformar las coordenadas de los datos puesto que se requería que tuvieran rangos similares, en esta caso las variables de longitud y latitud estaban en km y la profundidad en mts. La transformación que se utilizó fue la siguiente:

$$z' = \frac{z - z_{\min}}{z_{\max} - z_{\min}}$$

en donde

z es el valor original (longitud, latitud o profundidad)

z_{\min} es el valor mínimo de z

z_{\max} es el valor máximo de z

z' es el valor transformado de z

Una vez transformadas las coordenadas, lo que se hizo fue identificar los ángulos que fueran iguales o similares en cada uno de los niveles para cada una de las variable. Después se procedió a hacer el cambio de valor de los ángulos pues en el programa GEOEAS los ángulos se miden como en geometría plana, es decir, el ángulo se mide partiendo del eje horizontal (X)

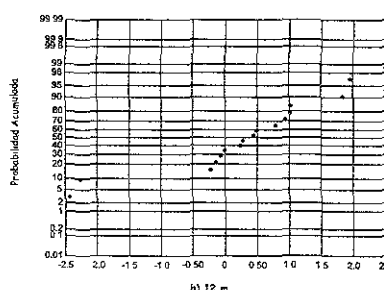
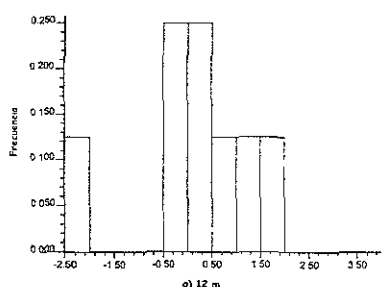
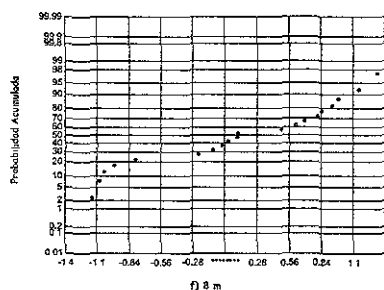
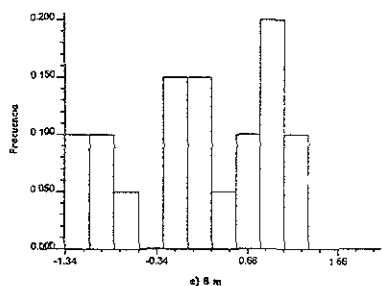
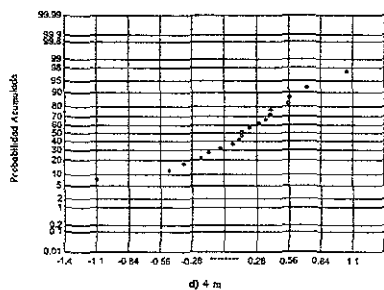
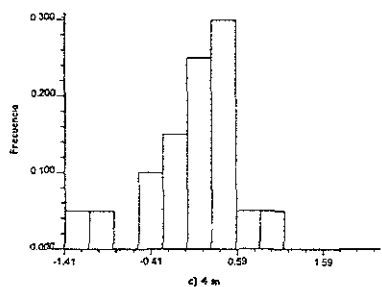
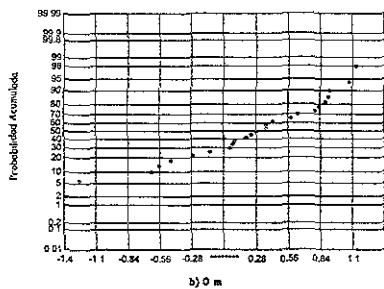
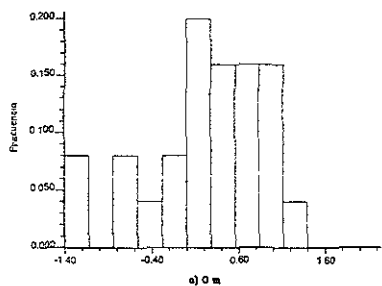


Figura 3.5.8 Histogramas y gráficas en papel normal para $e(\ln(Fe_2))$ por nivel de profundidad, Tercer Muestreo

en sentido contrario a las manecillas del reloj de 0 a 180°, mientras que las bibliotecas de GSLIB, miden los ángulos partiendo de la dirección Norte hacia el sentido de las manecillas del reloj 0 a 360° (azimut), como se hacen en las mediciones terrestres en general. Una vez que se tuvo identificados los ángulos, se trazó un círculo y sobre él las direcciones de cada nivel, es decir, para el primer muestreo en el nivel 0 fueron 60 y 130° y se trazaron en el círculo como se puede ver en la Figura 3.6 (a) con una línea continua, posteriormente se marcaron las del nivel de 4 mts cuyos ángulos son 0, 150 y 160° con una línea punteada y así sucesivamente, ver Tablas 3.3 a 3.5 para las direcciones.

Estos círculos a lo que ayudaron fue a observar el rango sobre el cual se debería buscar alguna dirección sobre el plano (ángulo azimut). Una vez identificados los rangos se calcularon los variogramas para diferentes direcciones ahora tomando en cuenta la profundidad (ángulo dip) (ver apéndice 1), se modelaron solo aquellos que tuvieran un comportamiento creciente y con un número grande de pares o parejas en cada rezo. Las Figuras 3.7.1 a 3.7.8 muestran los variogramas experimentales estandarizados seleccionados y los modelos propuestos.

El modelo que mejor se ajustó para los variogramas experimentales fue el gaussiano, lo cual era de esperarse ya que en el análisis de \mathcal{R}^2 fue el que mejor resultó y también por el tipo de variables que se están trabajando como se había mencionado anteriormente.

La Tabla 3.6 muestra los parámetros de los modelos propuestos los cuales son menos en cuanto a cantidad de modelos seleccionados en \mathcal{R}^2 . Esto suena lógico ya que al considerar otra variable de ubicación, en este caso la profundidad, la continuidad espacial es más restringida.

Con los variogramas seleccionados en cada uno de los muestreos, se interpoló el volumen de la zona por medio de kriging. Los mapas de kriging y los de su desviación estándar se muestran en las Figuras 3.8.1 a 3.8.8 y 3.9.1 a 3.9.8 plano XY, es decir, vistos como en \mathcal{R}^2 . De esto se hablará más adelante y con mayor detalle en la sección 3.5.3.

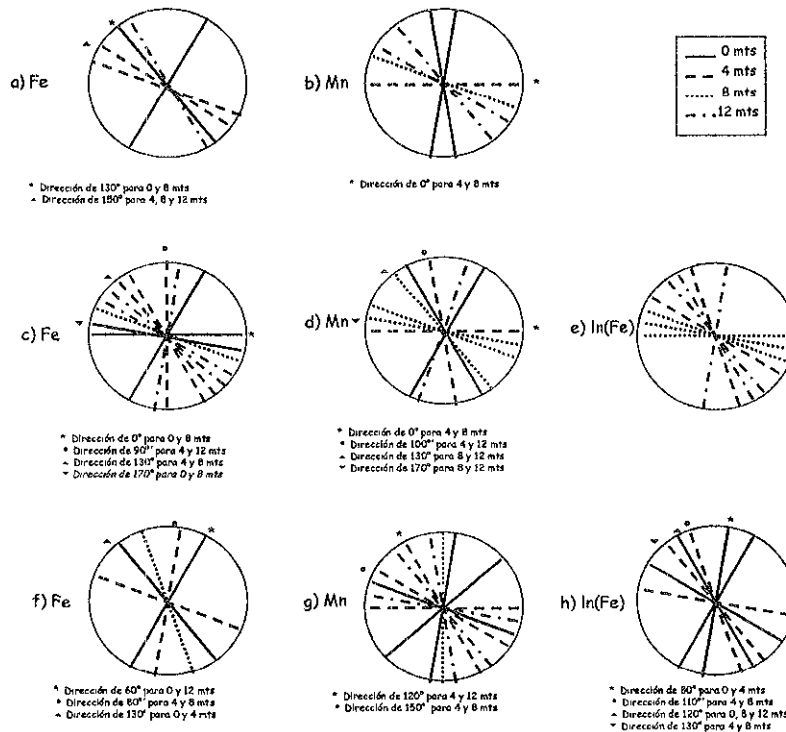


Figura 3.6 Círculos en donde se identifican las direcciones para el análisis de 2 dimensiones, cada renglón es un muestreo y los direcciones marcadas (*, °, ^ y >) son comunes en distintos niveles

Posteriormente la validación de la interpolación fue hecha por medio de la validación cruzada y los resultados se presentan en la sección 3.5.4.

3.5.2 Variogramas

Al analizar los parámetros de los variogramas en cada uno de los muestreos, se observó que hay direcciones que son claves en cuanto a los valores que tomaron tanto el rango como la meseta. Esto hablaría de que son direcciones sobre las cuales se están dispersando o estancando los metales. La Tabla 3.6 es muy importante para las observaciones marcadas y los ángulos o direcciones, estos últimos serán referidos como (ángulo Azimut, ángulo Dip) durante esta parte del análisis.

a) Primer muestreo

En este muestreo las direcciones que son comunes para ambas variables son dos: (120, 0) y (140, 0), ambas están marcadas en la Tabla 3.6 con diferentes tonos de gris. Para la primera dirección los valores de meseta coinciden y los de rango se podrían decir que son parecidos. En la segunda dirección, ambos parámetros son diferentes, ver Figura 3.7.1 y Figura 3.7.2.

Hay otra cosa a tomar en cuenta y es que el valor del ángulo Dip es 0, lo que querría decir que la continuidad espacial sólo se da a nivel del mar.

Epoca de Muestreo	Variable	Azimut	Dip	Rango	Meseta
Secas	Fe	120	0	1.25	1.6
		140	0	1.3	2.5
		160	0	1.2	1.0
	Mn	120	0	1.0	1.6
		140	0	0.9	2.0
Lluvias	Fe	110	0	0.7	1.8
		170	0	0.80	1.10
		110	-30	0.5	2.5
		200	-120	0.8	2.3
	Mn	110	0	1.0	1.5
		110	-30	0.5	1.5
		110	-150	0.4	1.3
		140	-150	1.2	2.0
	Ln(Fe)	110	0	1.0	2.0
		110	-30	0.6	1.9
110		-150	0.6	1.4	
Nortes	Fe	140	0	1.3	2.2
		190	0	0.25	2.5
		110	-120	0.6	1.7
	Mn	110	0	0.45	1.2
		140	0	1.2	1.7
	Ln(Fe)	140	0	1.6	4.0
		190	0	0.5	2.5

Tabla 3.6 Parámetros para modelos gaussianos de los 3 muestreos para análisis en \mathcal{R}^3

b) Segundo Muestreo

Para este muestreo son 2 las direcciones que coinciden, en este caso para todas las variables son las siguientes (110, 0), (110, -30), y otra dirección más para dos de ellas (Mn, Ln(Fe)) que es (110, -150). Aquí el valor del ángulo Dip varía en cuanto a que se está tomando en cuenta los

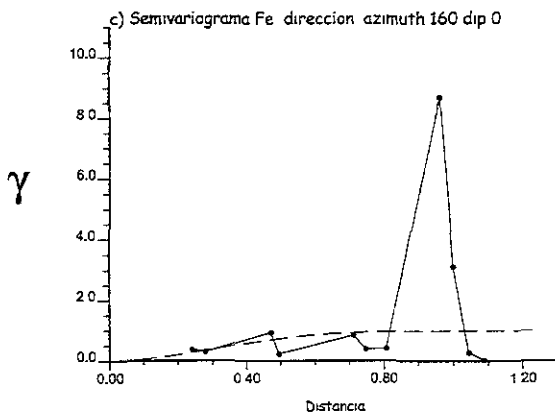
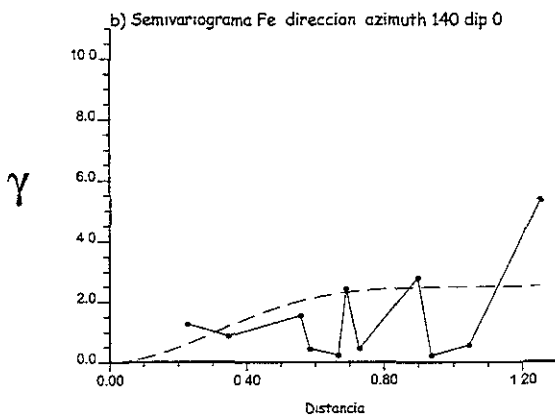
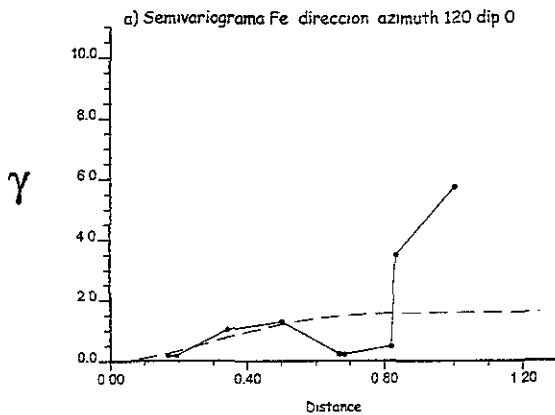


Figura 3.7.1 Variogramas experimentales y modelo gaussiano para la variable Fe en el primer muestreo

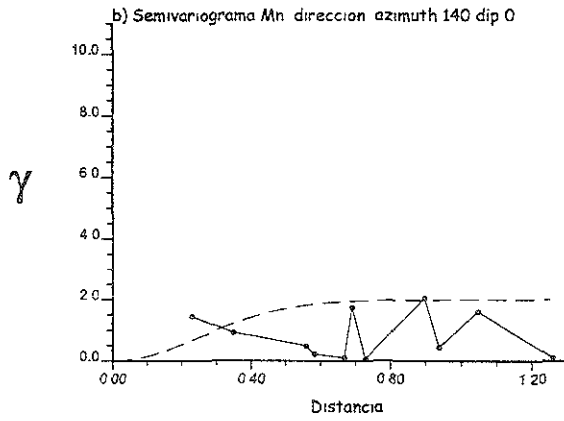
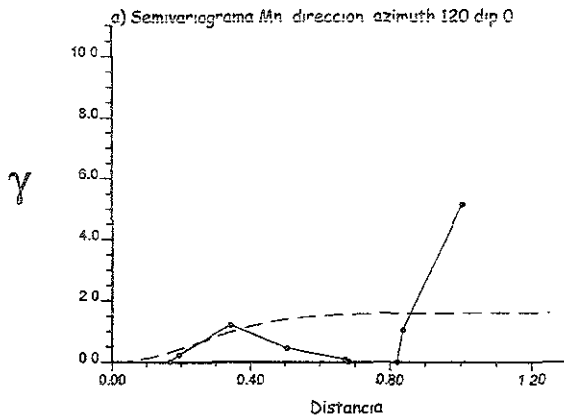


Figura 3.7 2 Variogramas experimentales y modelo gaussiano para la variable Mn en el primer muestreo

cambios que hay a nivel de profundidad. Cabe recordar que el segundo muestreo se caracteriza por tener un factor adicional que es la lluvia y que, de alguna manera, siempre se le ha considerado importante para la interpretación de los datos, ver Tabla 3.6.

Los valores de meseta con respecto a la dirección de (110, 0) varían de 1.5 a 2, mientras que el valor del rango va de 0.7 a 1.0, coincidiendo este último valor para Mn y ln(Fe). La segunda dirección, la meseta va de 1.5 a 2.5 y el rango de 0.5 y 0.6, estos valores de rango se podrían considerar parecidos. La última dirección para valores de meseta son similares de 1.3 y 1.4 y de rango de 0.4 y 0.6, ver Figura 3.7.3 a Figura 3.7.5.

b) Tercer muestreo

En este muestreo sólo hay una dirección que se da de manera conjunta para las tres variables, esta es (140, 0). Los valores de meseta en esta dirección varían mucho de 1.7 a 4, y el rango de 1.2 a 1.6. Los valores altos en estos parámetros son para la variables ln(Fe). Ver Tabla 3.6 y Figura 3.7.6 a Figura 3.7.8.

Hay otra dirección que coincide para las variables Fe y ln(Fe) y es (190, 0). En este caso el valor de la meseta coincide, 2.5, y el rango es de 0.25 y 0.5 respectivamente.

3.5.3 Kriging

En el análisis se tomaron no sólo los variogramas direccionales propuestos, sino también hubo necesidad de definir el tamaño de un elipsoide cuyo objetivo es el de definir la vecindad de datos a trabajar. Los valores que definieron el elipsoide en cada muestreo fueron establecidos de acuerdo a los rangos y mesetas de los variogramas seleccionados y se presentan en la Tabla 3.7 así como también los ángulos sobre los que se define, ver especificaciones del modelo para el variograma en apéndice 1. Estos parámetros son requerimientos de las bibliotecas utilizadas que son explicadas en los métodos de búsqueda del apéndice 1.

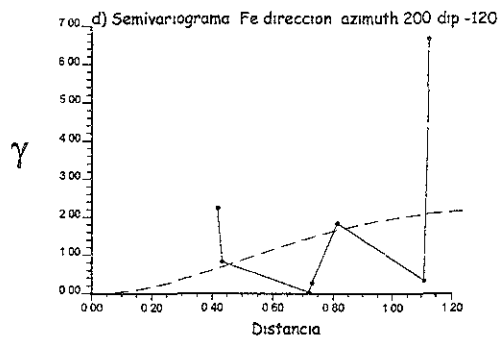
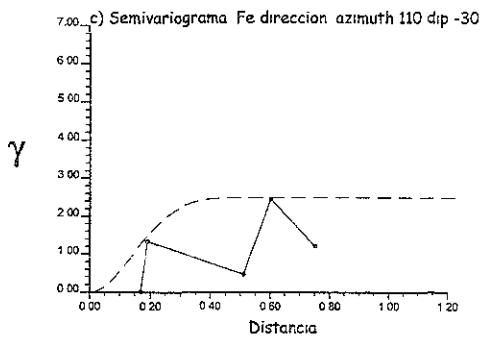
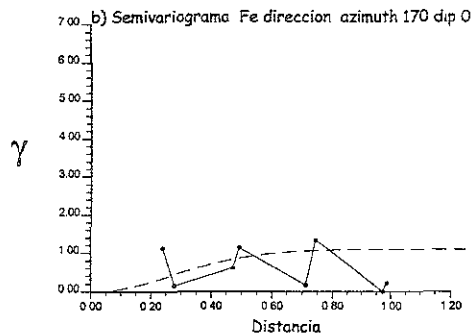
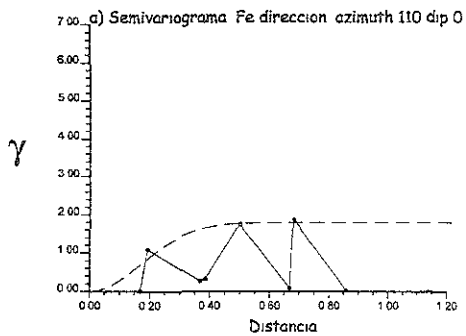


Figura 3.7.3 Variogramas experimentales y modelo gaussiano para la variable Fe en el segundo muestreo

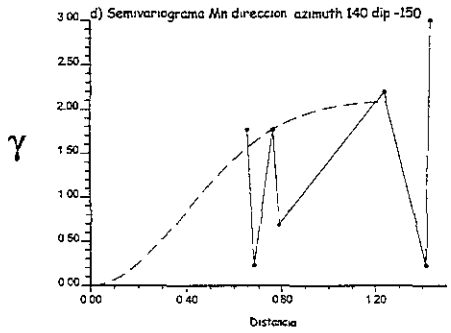
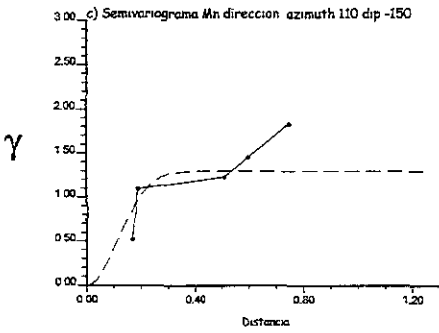
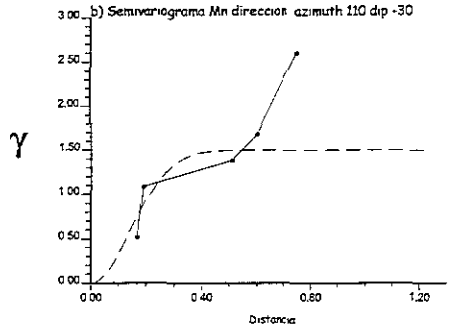
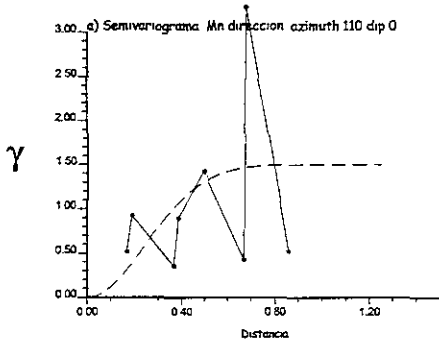


Figura 3.7.4 Variogramas experimentales y modelo gaussiano para la variable Mn en el segundo muestreo

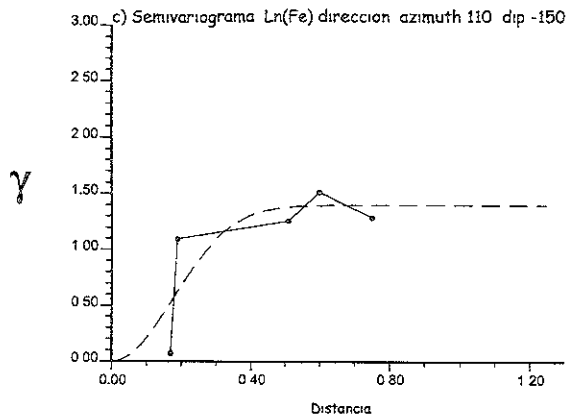
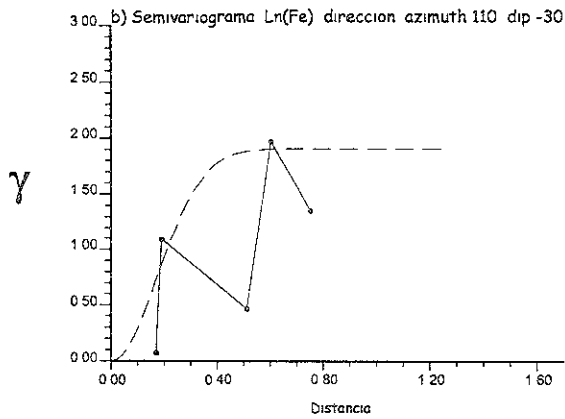
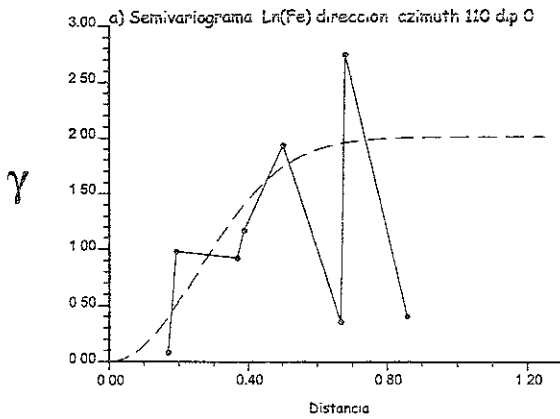


Figura 3.7.5 Variogramas experimentales y modelo gaussiano para la variable $\ln(\text{Fe})$ en el segundo muestreo

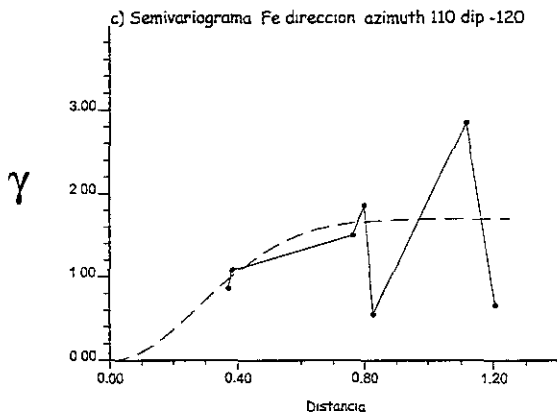
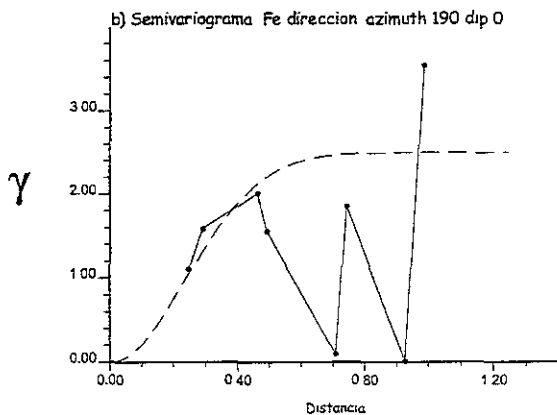
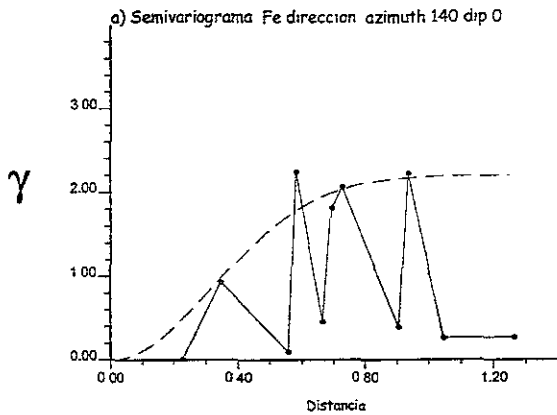


Figura 3.7.6 Variogramas experimentales y modelo gaussiano para la variable Fe en el tercer muestreo

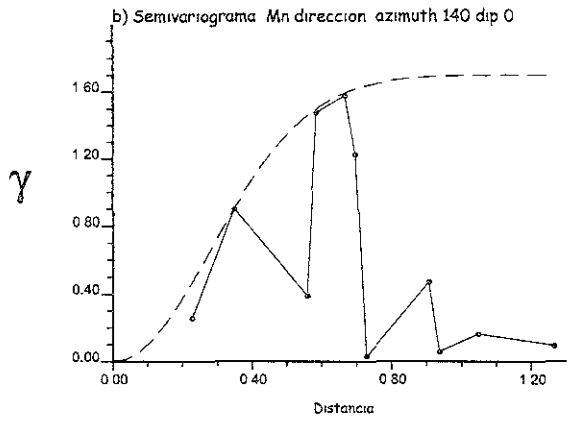
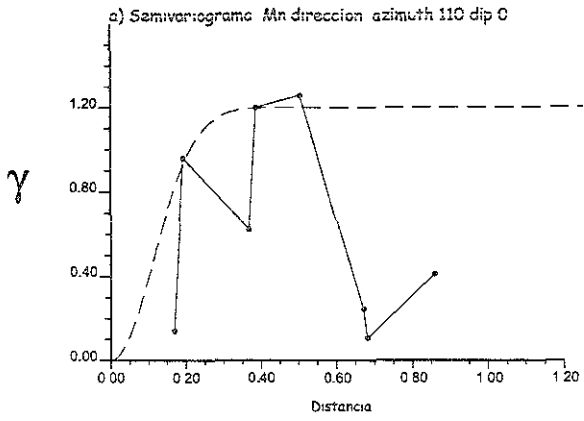


Figura 3.7.7 Variogramas experimentales y modelo gaussiano para la variable Mn en el tercer muestreo

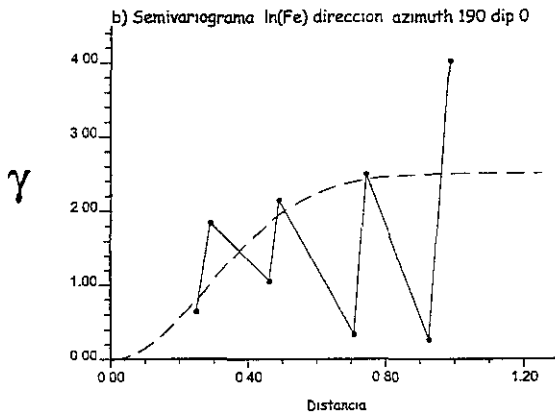
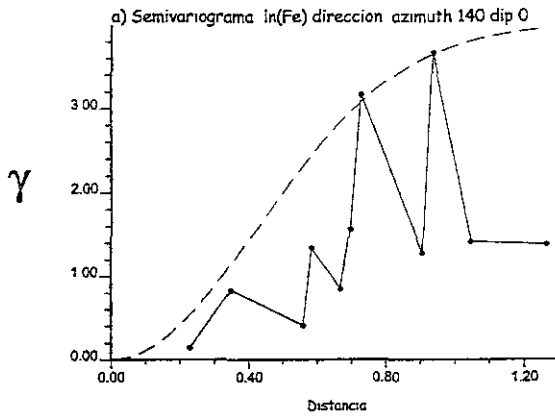


Figura 3.7.8 Variogramas experimentales y modelo gaussiano para la variable $\ln(\text{Fe})$ en el tercer muestreo

Variable	No. de Muestreo	Radio de Búsqueda	Angulos de Búsqueda
Fe	1	(1.3, 1.2, 1.5)	(0, 120, -110)
	2	(1.2, 0.6, 7.0)	(200, 110, -120)
	3	(1.0, 2.0, 0.75)	(110, -120, 140)
Mn	1	(1.9, 0.9, 0.8)	(0, 120, -110)
	2	(1.5, 0.4, 3.0)	(110, -30, -150)
	3	(1.0, 2.0, 0.75)	(110, -120, 140)
Ln(Fe)	2	(1.7, 0.6, 3.0)	(110, 0, -150)
	3	(1.0, 2.0, 0.75)	(110, -120, 140)

Tabla 3.7 Parámetros para definir el elipsoide de búsqueda

El número de valores predichos para cada variable fue de 0 a 1 en todas las dimensiones, con incrementos de 0.1 o 0.2 dependiendo del número de valores a predecir en cada eje. El tipo de kriging que se utilizó fue el de bloques de 2x2x2 y los programas que se utilizaron fueron de las bibliotecas de GSLIB, kt3d principalmente.

Al igual que en el análisis en dos dimensiones será necesario tomar en cuenta tres gráficas, la de los datos muestreados (Figuras 3.2.1 a 3.2.6), la de kriging en 3 dimensiones (Figuras 3.8.1 a 3.8.8) y la de desviación estándar de kriging (Figuras 3.9.1 a 3.9.8).

En general el comportamiento que se describirá en el siguiente punto es parecido al análisis anterior pero con menos modelos de variogramas y tomando en cuenta a la profundidad como un parámetro más de ubicación.

De igual manera que en dos dimensiones, la validación cruzada fue aplicada y se construyeron histogramas y gráficas en papel normal de los errores de predicción (Figuras 3.10.1 a 3.10.8) con la idea de observar el comportamiento de los mismos.

a) Primer Muestreo

Para la variable Fe, los valores de mayor concentración se observan en las partes Norte y Sur de la zona muestreada, las bajas concentraciones se encuentran en el centro como se puede ver en la Figura 3.7.1. El rango que se está manejando en este conjunto de valores predichos es de 0 a $25 \times 10^{-3} \mu\text{g/g}$, que sería un poco más del tercer cuartil global de la muestra observada en este muestreo, ver Tabla 3.1. Las gráficas de Kriging (DE), Figura 3.9.1 muestran valores, entre 0 y 2 lo cual habla de que la variabilidad en este conjunto de datos es pequeña y homogénea a través de la malla.

El comportamiento del Mn es similar al de Fe, valores de alta concentración son en el Norte y Sur de la malla muestreada y bajos en el centro, con la particularidad que el rango de predicción de estos valores es muy parecido a los datos reales, es decir 0 a $15 \times 10^{-3} \mu\text{g/g}$, ver Figura 3.8.2 y Tabla 3.1 para esta variable. Las gráficas de Kriging (DE), Figura 3.9.2 muestra valores bajos en casi todas ellas. Las Figuras 3.8.2 (c) y (d) presentan algunas manchas en la parte Sur de la malla pero la magnitud de estas es menor a $0.5 \times 10^{-3} \mu\text{g/g}$ según la escala de grises puesta en ellas, así que se podría considerar que la variabilidad es homogénea.

b) Segundo Muestreo

Para la variable Fe en este muestreo la presencia de la profundidad se hace notar puesto que al comparar las Figuras 3.3.3(a) y 3.8.3(a) son diferentes porque en esta última es notoria la concentración del metal al centro de la imagen y que al relacionarla con la Figura 3.8.3(b) resulta ser parecida en cuanto a la ubicación de la concentración del metal. El comportamiento de que el metal no se ha dispersado adecuadamente es evidente a través de las imágenes pues como se mencionó en el análisis de dos dimensiones esto sólo sucede en los dos primeros niveles y en los dos últimos se muestra en donde está quedando el metal. La Figura 3.9.3 de la desviación estándar del kriging de Fe muestra que la variabilidad es baja en la zona muestreada y al compararla con la Figura 3.4.3 es muy baja.

Al igual que el Fe, la variable Mn muestra concentraciones altas en la parte central. En este caso, la profundidad en la generación del volumen del metal se observa en las Figura 3.8.4 (b) a (d). Al comparar las Figura 3.3.4 y Figura 3.8.4 se puede decir que el Mn tarda más en sedimentarse que el Fe ya que las mayores concentraciones se observan en los niveles de 4 y 8 m. Nuevamente la Figura 3.9.4 de la desviación estándar del kriging de Mn muestra poca variabilidad en la zona muestreada y es más homogénea que la que se presenta en el análisis en dos dimensiones (ver Figura 3.4.4).

Para la variable del $\ln(\text{Fe})$, las superficies que se obtienen son más homogéneas que las que se generaron para Fe pues la transformación hace que los cambios de determinado punto a otro no sean tan extremos (ver Figura 3.8.5). La variabilidad del kriging generado es también mucho más homogénea que la del Fe (ver 3.9.5).

c) Tercer Muestreo

El comportamiento del Fe en este muestreo presenta una clara dispersión del metal hacia la zona Norte en todos los niveles de profundidad, a los 8 y 12 m también existen dos manchas en la parte sureste de la malla y en la parte media a los 12 m (ver Figura 3.8.6). Nuevamente se puede observar que al involucrar a la profundidad en el proceso del krigado resulta beneficioso para el análisis no solo en la predicción de los datos sino también en que la variabilidad es menor y mas homogénea, esto se puede ver en la Figura 3.9.6 en la zona muestreada.

Las imágenes que se observan para el Mn en este muestreo presentan acumulaciones del metal en las partes Sureste y Suroeste de la malla, la parte central muestra zonas de baja concentración que marcan la presencia del difusor y la sedimentación del metal a los 12 m de profundidad (ver Figura 3.8.7). En la Figura 3.9.7 se puede ver que la variabilidad es baja y homogénea en los procesos de predicción del kriging.

Para la variable $\ln(\text{Fe})$ se puede observar que las imágenes son muy parecidas a las del Fe pero más homogéneas en el sentido de que en zonas de baja concentración del $\ln(\text{Fe})$ no son tan

claras como se ve en las del Fe, esto habla que la transformación ayuda a que la superficie sea menos sensible a los cambios extremos entre los puntos muestreados, ver Figuras 3.8.6 y 3.8.8. Las gráficas de desviación estándar del Kriging de $\ln(\text{Fe})$ muestran valores bajos y homogéneos, ver Figura 3.9.8.

3.5.4 Validación Cruzada

El objetivo de la validación cruzada es el de ver si los modelos de variograma elegidos son capaces de predecir los valores que ya han sido observados, es decir, los valores muestreados. Los errores de predicción serán los indicativos de qué tan buena ha sido. El hecho de poder pronosticar estos valores hablaría de que con esta metodología se podría reproducir la variabilidad espacial de las observaciones. Al igual que en el análisis de dos dimensiones se construyeron histogramas y gráficas en papel normal de los errores y de igual maneras serán descritos, también se incluyeron gráficas de la ubicación de los errores en la malla de muestreo. Es importante mencionar que la información que se presenta es para todo el volumen y no por profundidad.

a) Primer Muestreo

La mayoría de los errores de predicción para la variable Fe se encuentran agrupados entre los valores de -10 a 10 y equivaldría casi al 90% de los valores, solamente 5 de ellos pueden considerarse valores extremos como se puede ver en el histograma y gráfica de papel normal 3.10.1 (a) y (b). Estos 5 valores se encuentran en los extremos de la malla en el extremo superior izquierdo de las gráficas de la Figura 3.11.1.

En el histograma y gráfica de papel normal para los errores de Mn muestran un comportamiento normal, la mayoría de ellos se ubican alrededor del 0 y solamente existe un valor grande negativo que se ubica en el extremo superior izquierdo de la malla, ver Figura 3.10.1 (c) y (d) y Figura 3.11.2 (d).

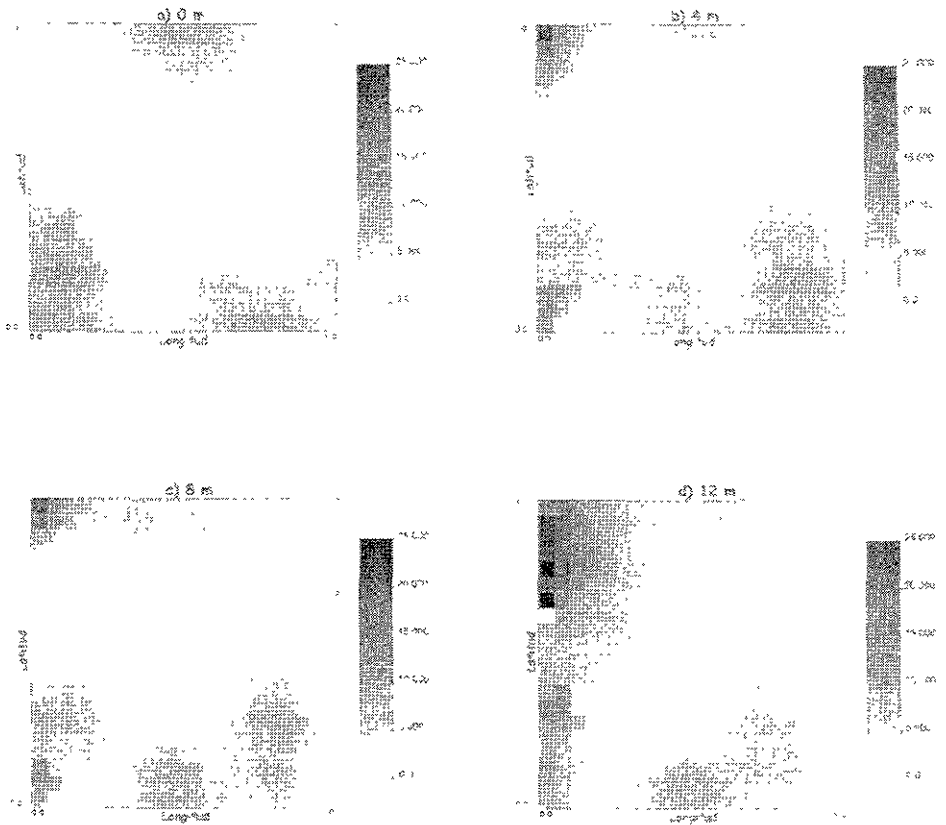


Figura 3.6 1 Kriging en 3D para Fe por niveles de profundidad, Primer Muestreo

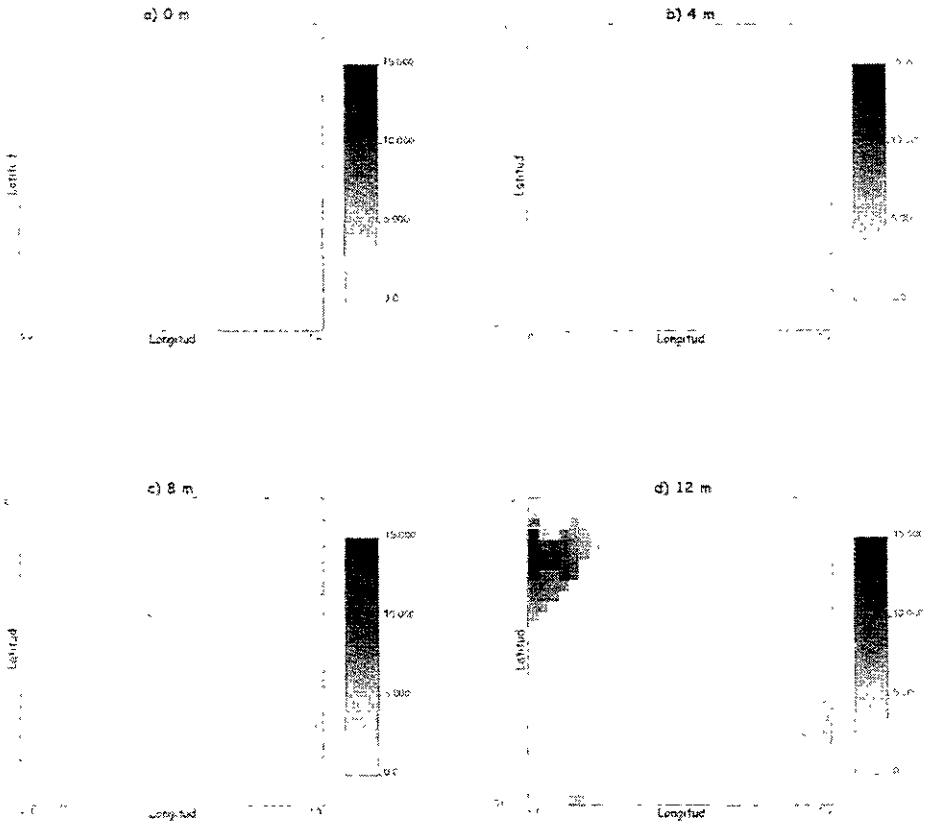


Figura 3.8.2 Kriging en 3D para Mn por niveles de profundidad, Primer Muestreo

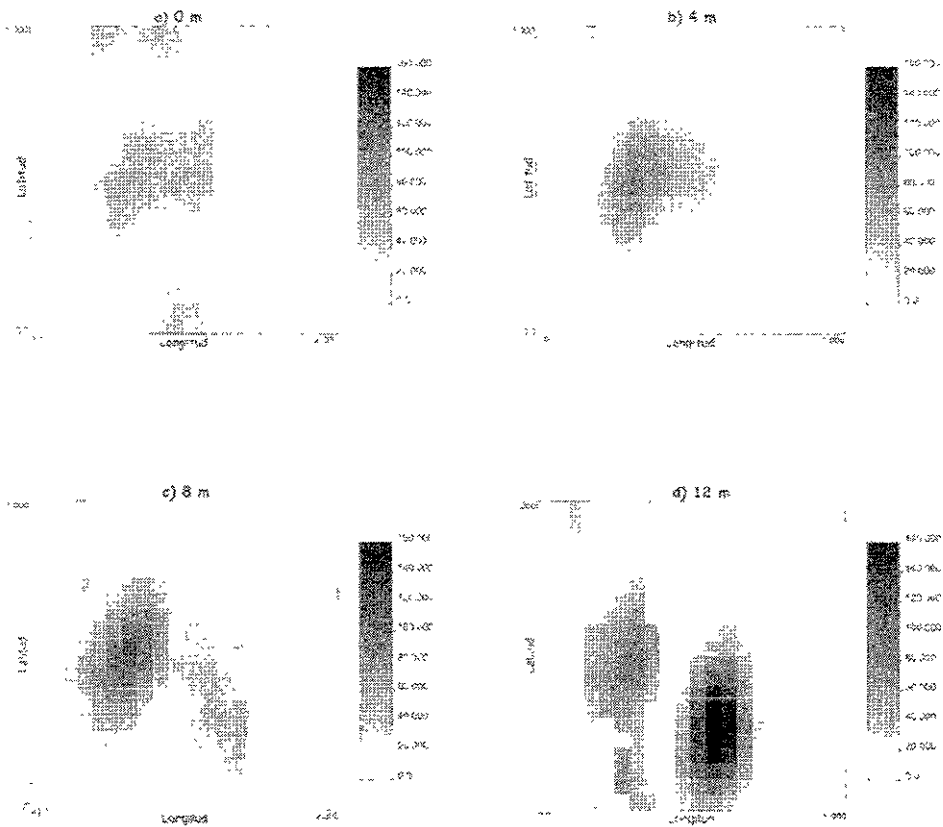


Figura 3.8.3 Kriging en 3D para Fe por niveles de profundidad, Segundo Muestreo

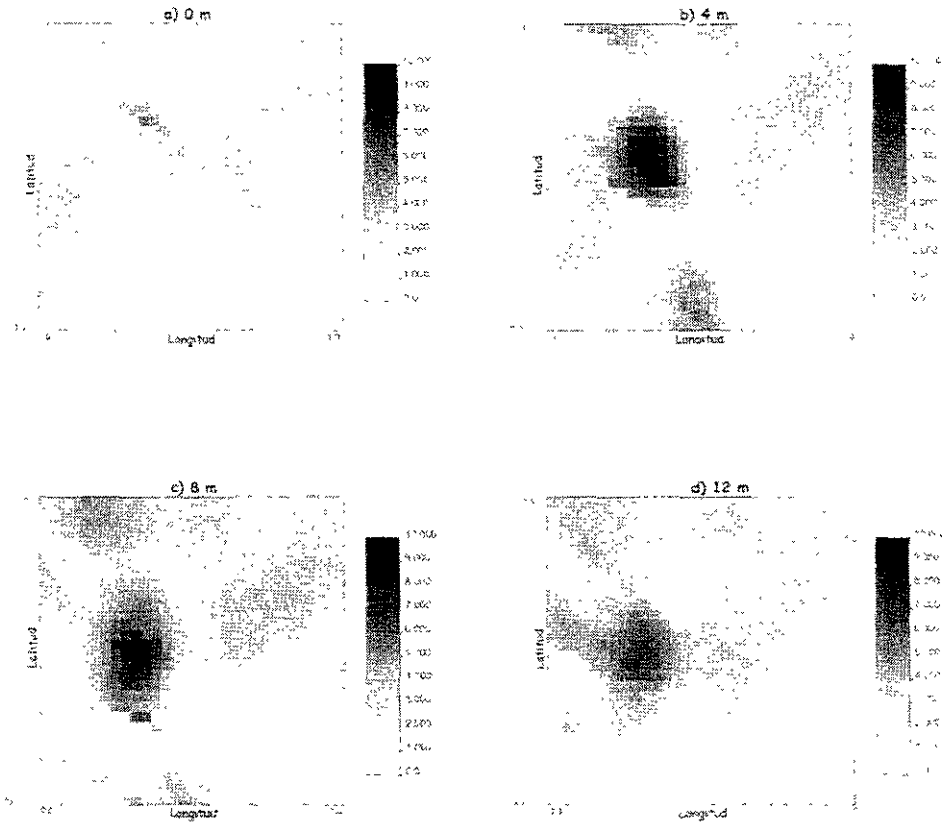


Figura 3.8.4 Kriging en 3D para Mn por niveles de profundidad, Segundo Muestreo

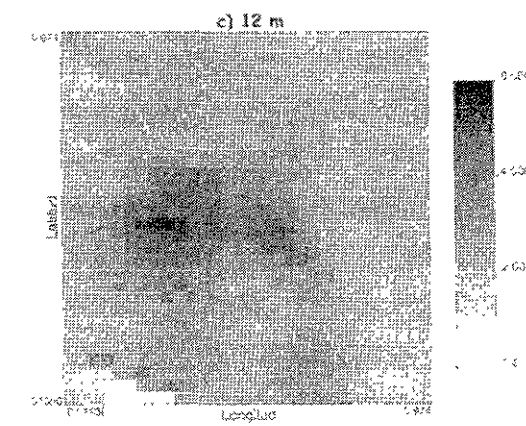
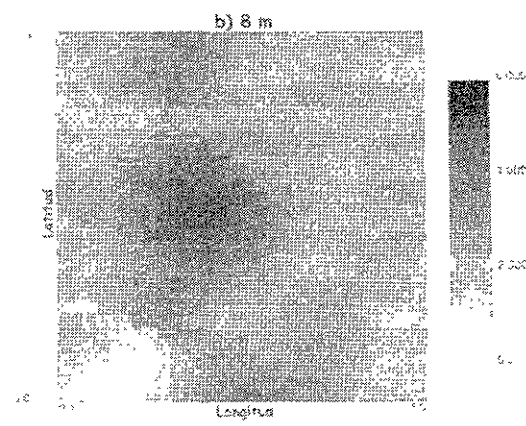
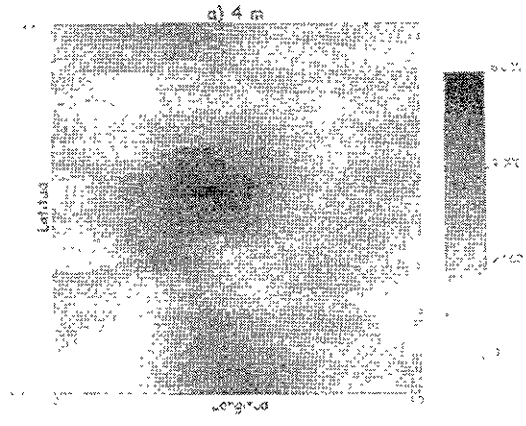


Figura 3.8.5 Kriging en 3D para $\ln(\text{Fe})$ por niveles de profundidad. Segundo Muestreo

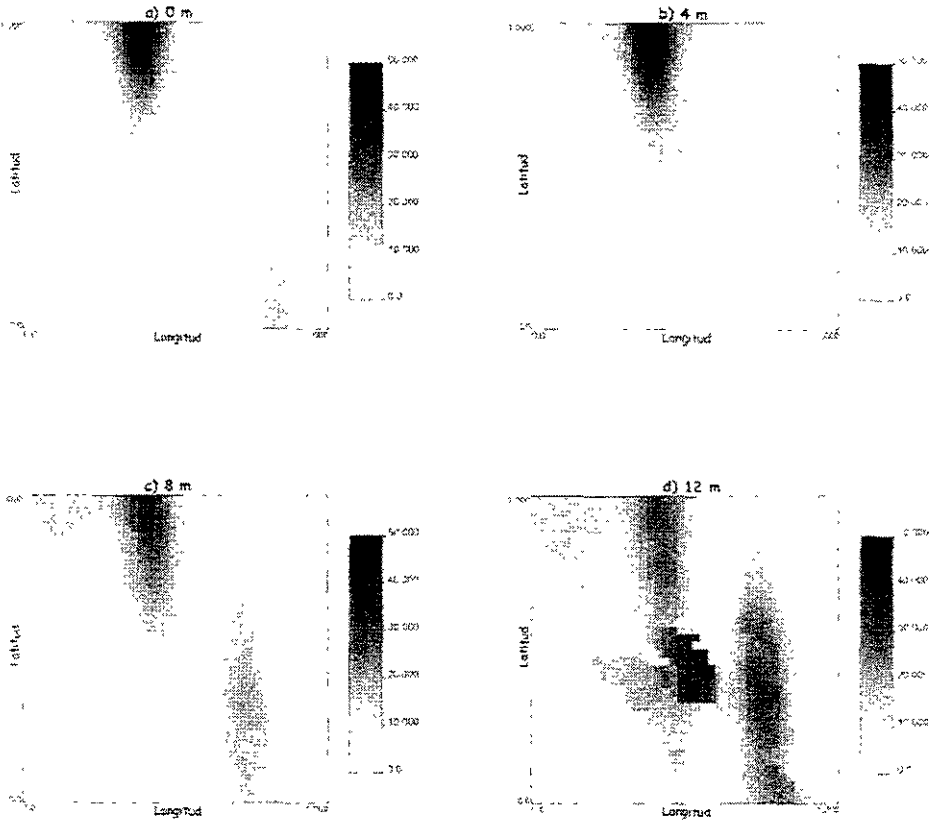


Figura 3.8.6 Kriging en 3D para Fe por niveles de profundidad. Tercer Muestreo

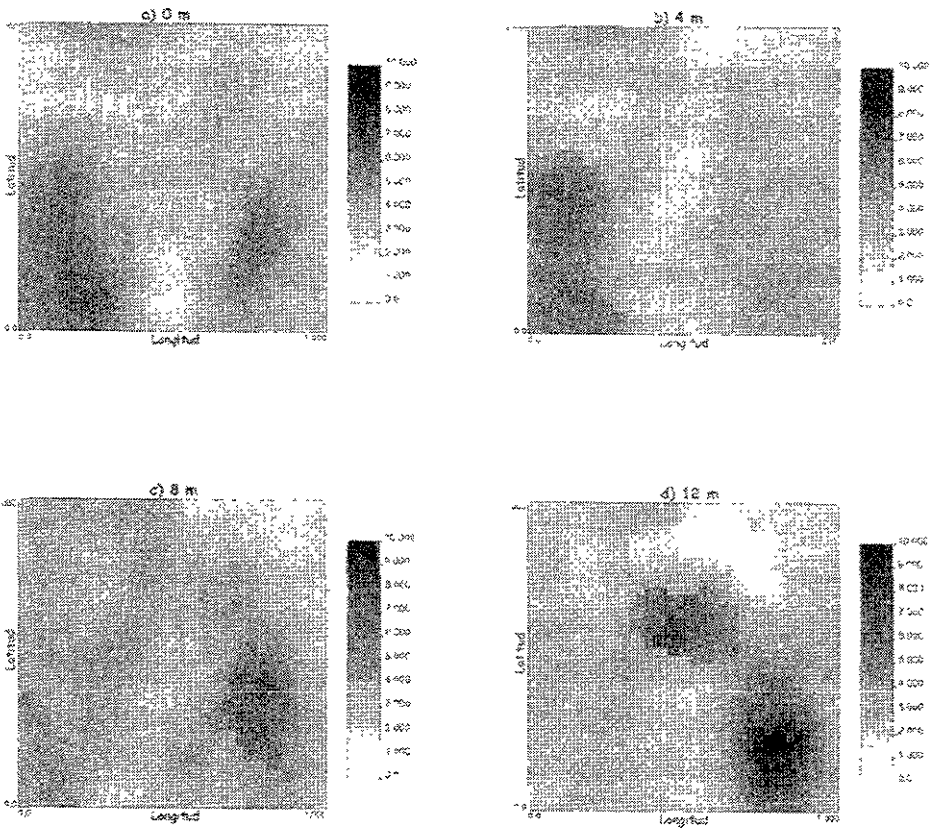


Figura 3.8.7 Kriging en 3D para Mn por niveles de profundidad, Tercer Muestreo

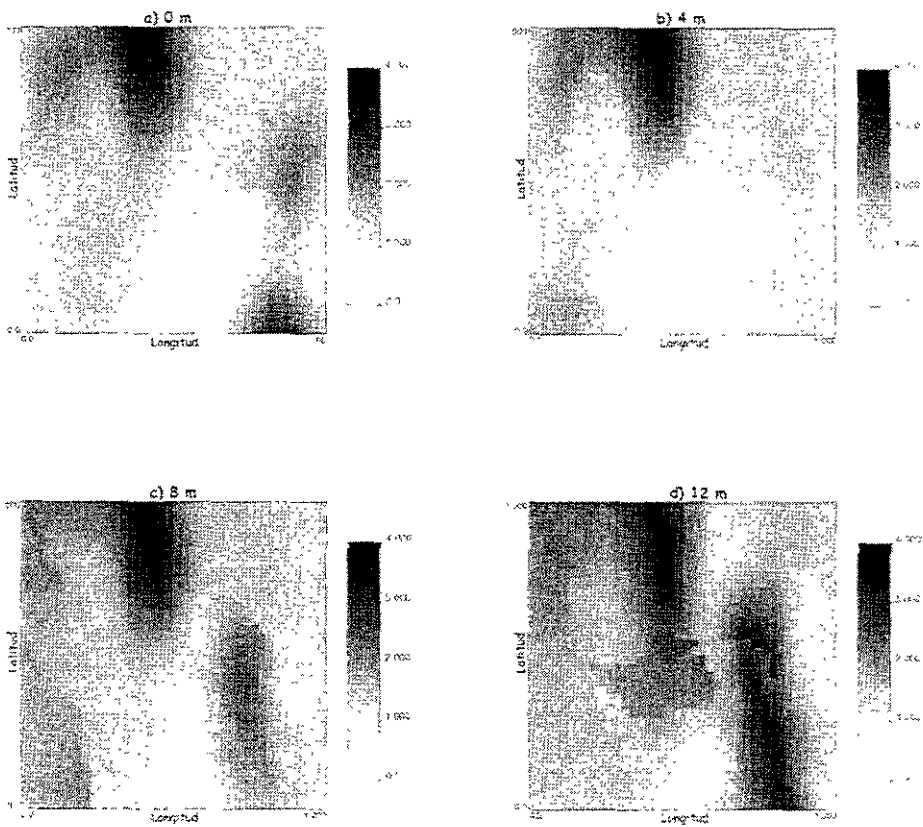


Figura 3.8.8 Kriging en 3D para $\ln(\text{Fe})$ por niveles de profundidad, Tercer Muestreo

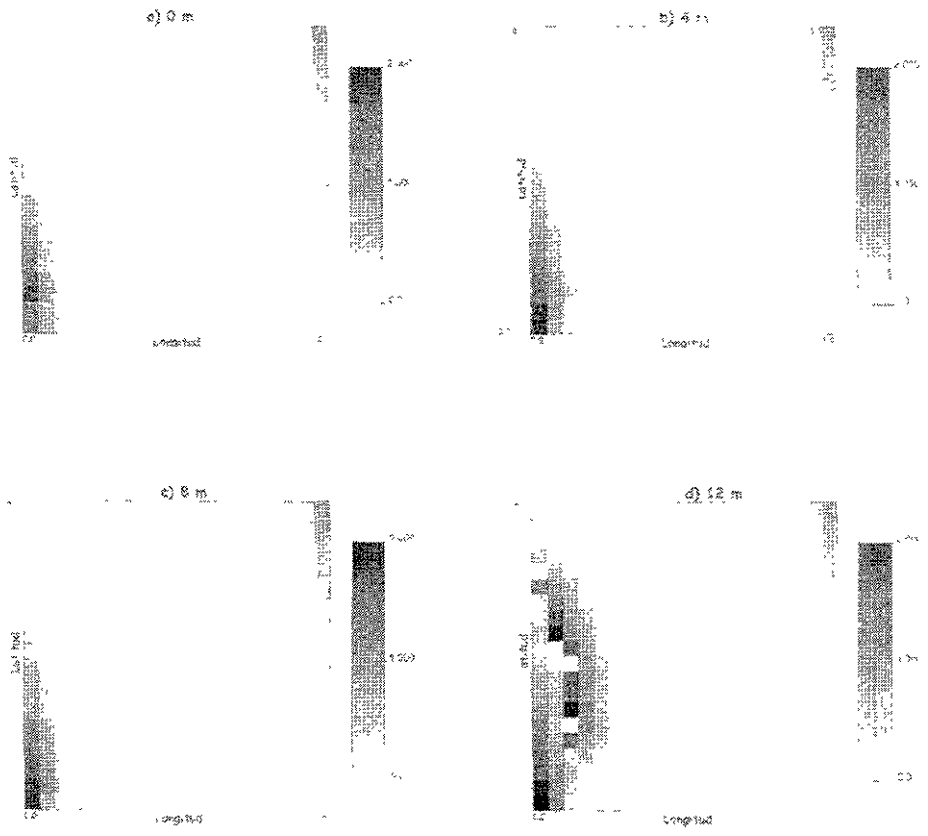


Figura 3.9.1 Kriging(DE) en 3D para Fe por niveles de profundidad, Primer Muestreo

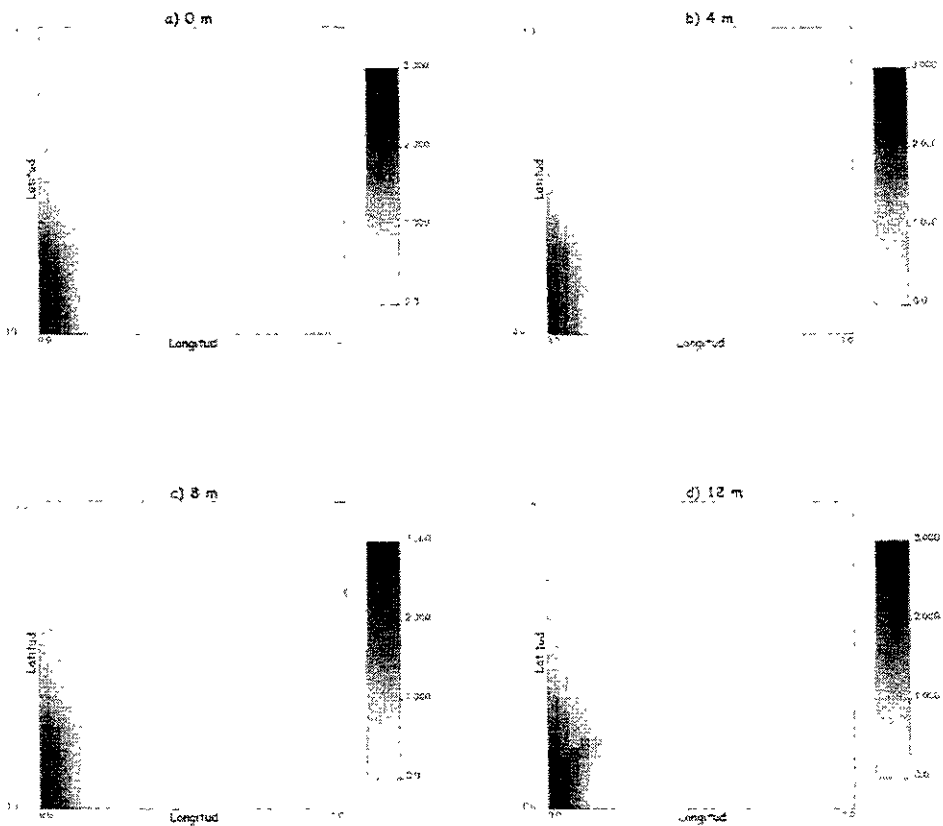


Figura 3 9 2 Kriging(DE) en 3D para Mn por niveles de profundidad, Primer Muestreo

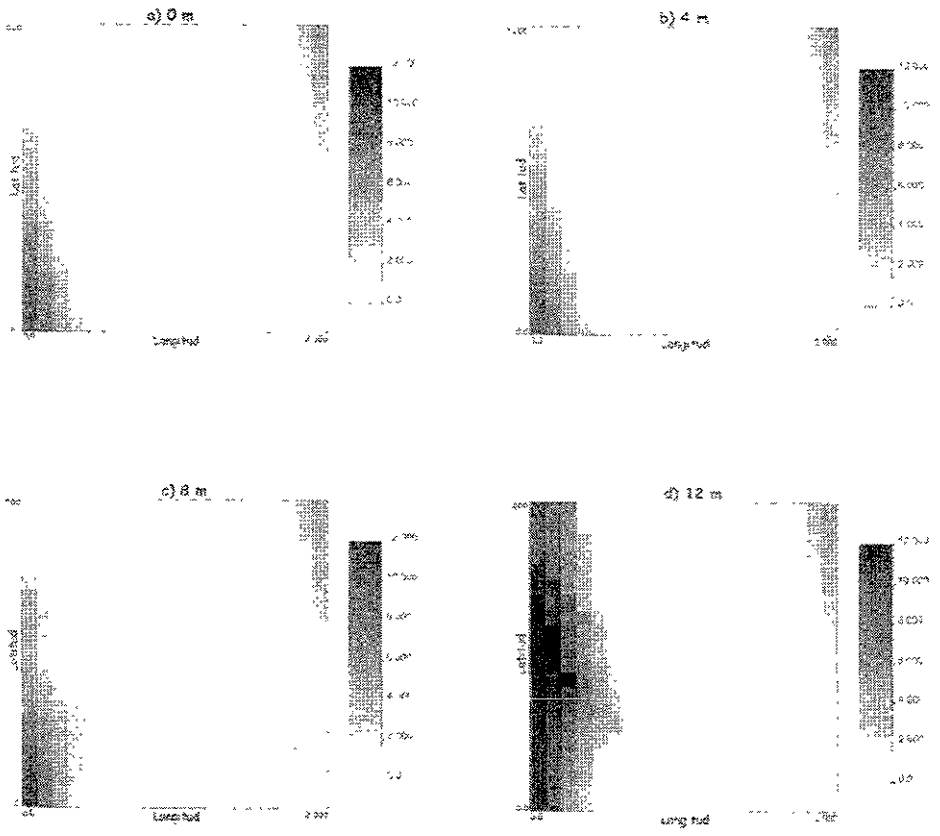


Figura 3.9.3 Kriging(DE) en 3D para Fe por niveles de profundidad, Segundo Muestreo

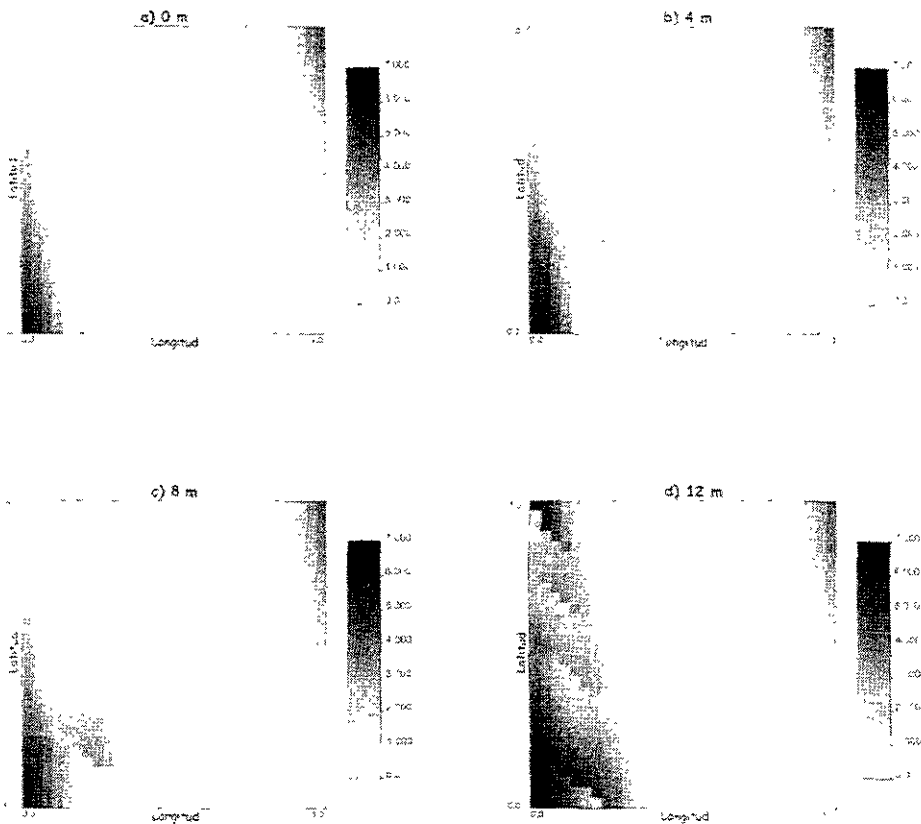


Figura 3.9.4 Kriging(DE) en 3D para Mn por niveles de profundidad, Segundo Muestreo

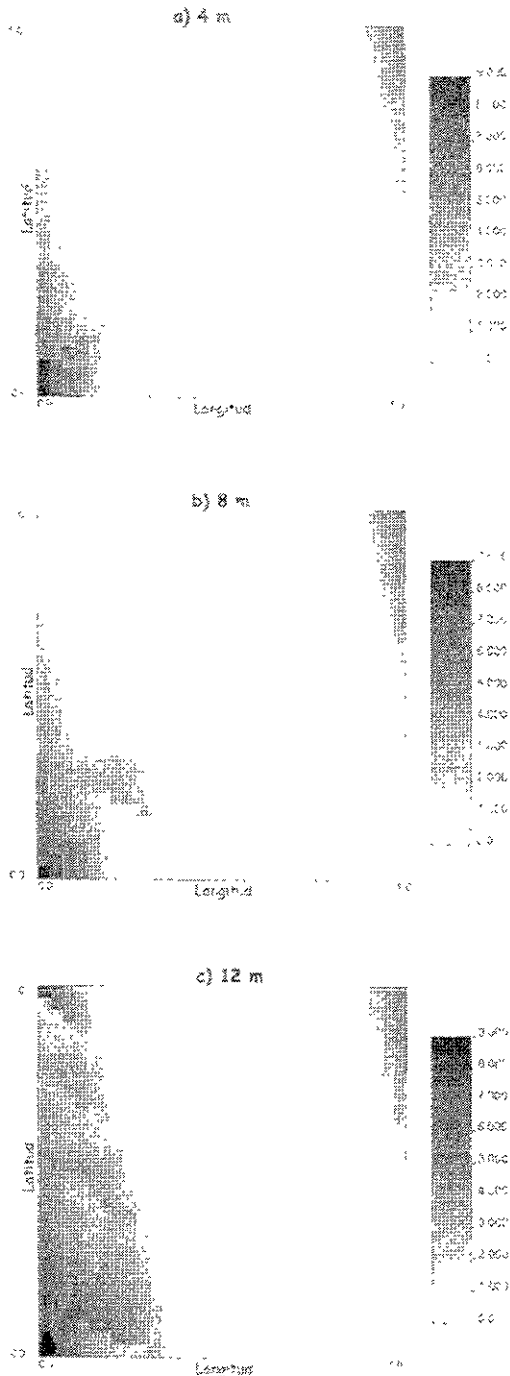


Figura 3.9.5 Kriging(DE) en 3D para In(Fe) por niveles de profundidad, Segunda Muestra

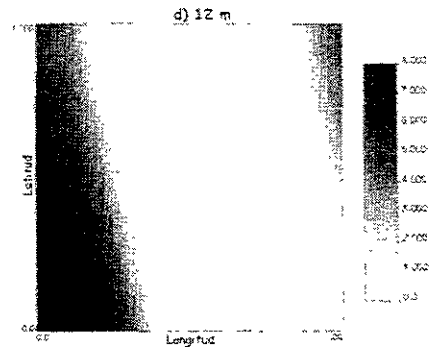
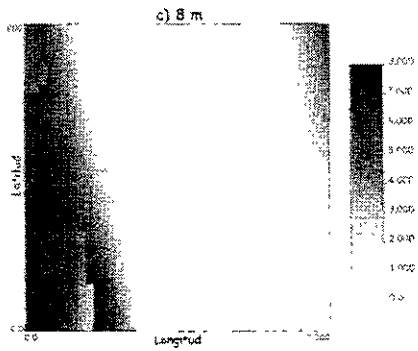
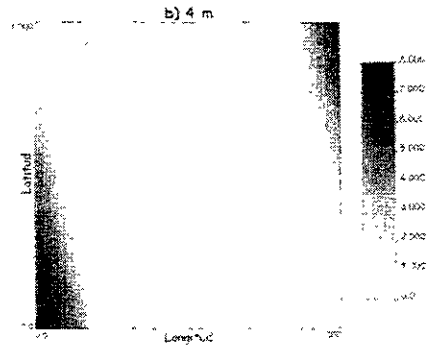
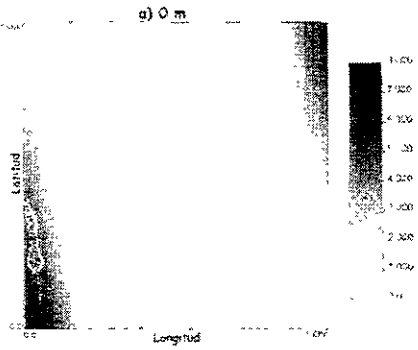


Figura 3.9.6 Kriging(DE) en 3D para Fe por niveles de profundidad, Tercer Muestreo

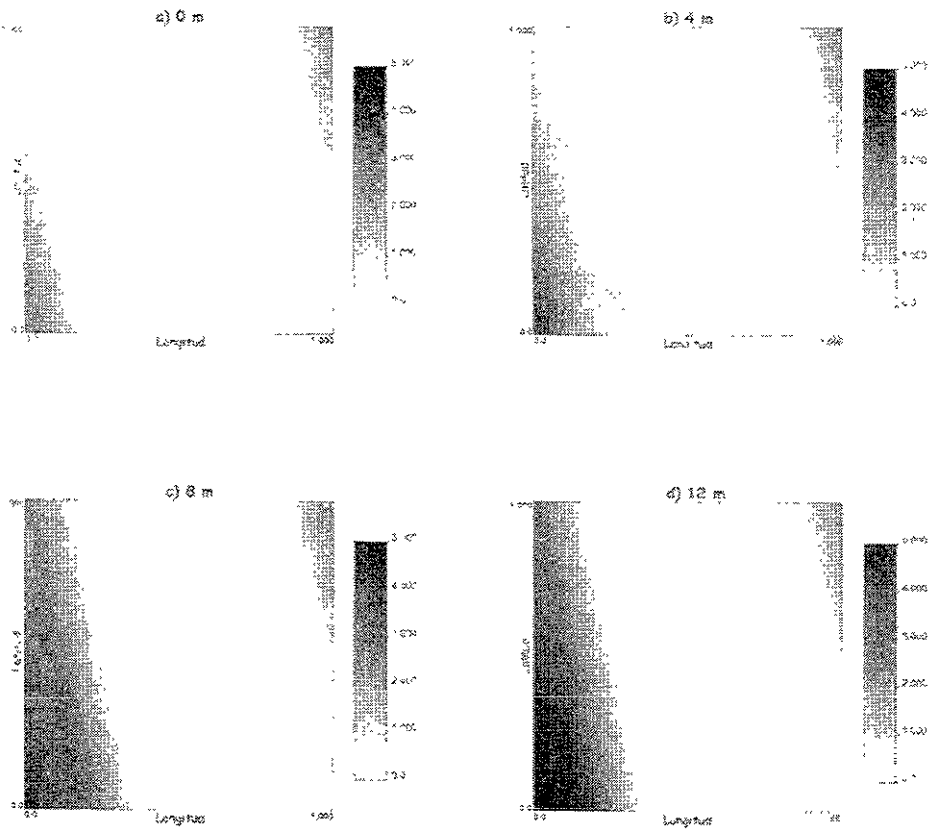


Figura 3.9.7 Kriging(DE) en 3D para Mn por niveles de profundidad, Tercer Muestreo

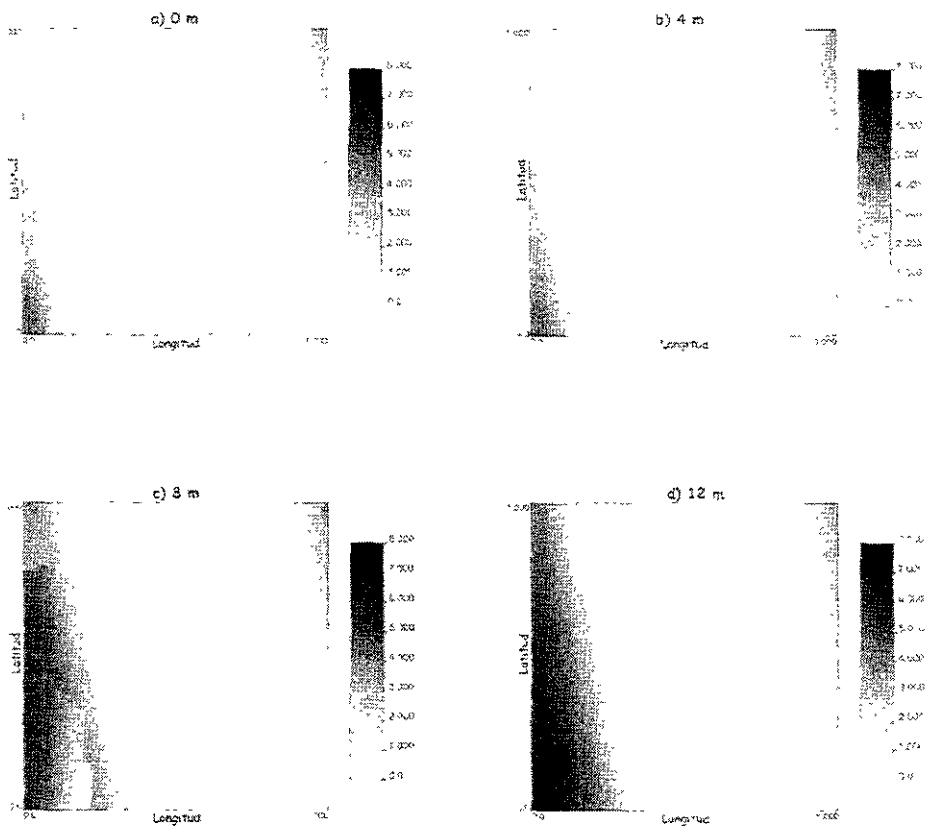


Figura 3.9.8 Kriging(DE) en 3D para $\ln(\text{Fe})$ por niveles de profundidad, Tercer Muestreo

b) Segundo Muestreo

En este muestreo los valores de los errores de predicción de Fe de -147 a 30 aproximadamente, es el rango más grande que hay para esta variable en los tres eventos analizados pues es el que tiene mayor variabilidad. A pesar de esto, la mayoría de ellos se concentran cercanos al valor 0 como se muestra en las Figura 3.10.2 (a) y (b). Hay 3 valores subestimados que podrían considerarse grandes y estos se ubican en la parte central de la malla en los niveles de 4, 8 y 12 m ver Figura 3.11.3 (b) a (d).

El histograma y la gráfica de papel normal para los errores de Mn son la que tiene una distribución normal bien definida como se puede observar en la Figura 3.10.2 (c) y (d). Los valores extremos en esta distribución se encuentran ubicados en la parte central de la malla a los 4 m de profundidad, ver Figura 3.11.4 (b).

La transformación $\ln(\text{Fe})$ hace que la predicción de los valores sea mejor, esto ya se había mencionado anteriormente pero la distribución de los errores es reflejo de que la transformación fue adecuada ya que el comportamiento que muestran la Figura 3.10.2 (e) y (f) es bueno. Los valores extremos que se observan en estas gráficas se encuentran en la parte central de la malla o muy cerca de esta parte en los niveles de 8 y 12 m, ver Figura 3.11.5 (b) y (c).

c) Tercer Muestreo

Los histogramas de Fe y Mn en este muestreo presentan un comportamiento muy similar pues los errores se acumulan en la parte central y algo parecido sucede también con las gráficas en papel normal, ver Figura 3.10.3 (a), (c) y (b), (d), respectivamente. Los valores extremos de los errores de Fe y Mn se encuentran ubicados en la parte central y Sureste de la malla a los 12 m, ver Figura 3.11.6 (d) y Figura 3.11.7 (d).

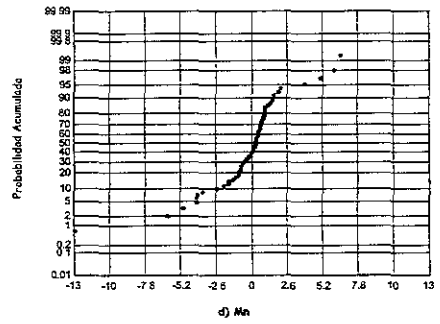
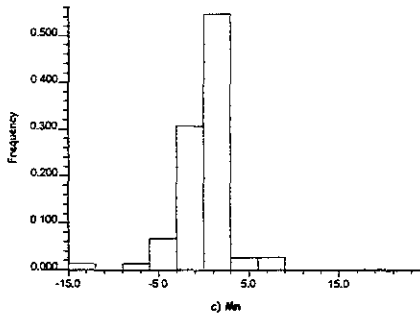
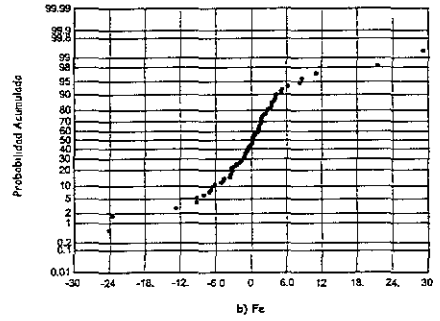
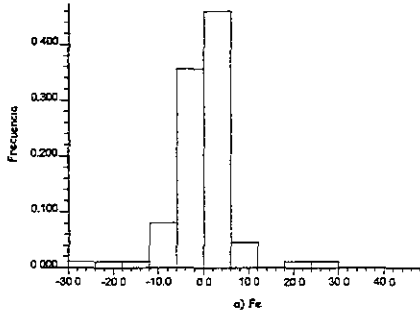


Figura 3.10.1 Histogramas y gráficas en papel normal para $e(\text{Fe})$, $e(\text{Mn})$ y $e(\ln(\text{Fe}))$ por niveles de profundidad, Primer Muestreo

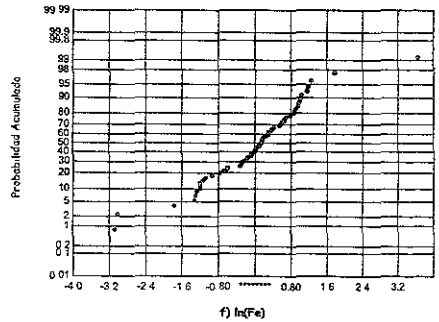
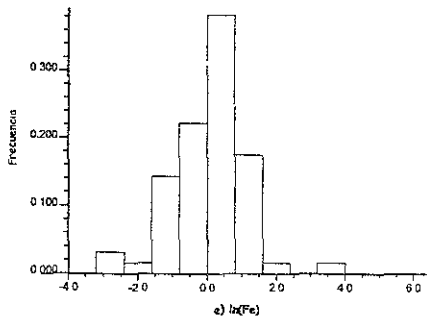
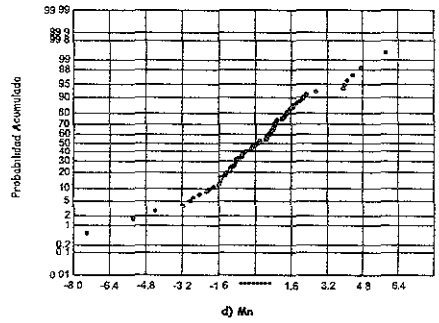
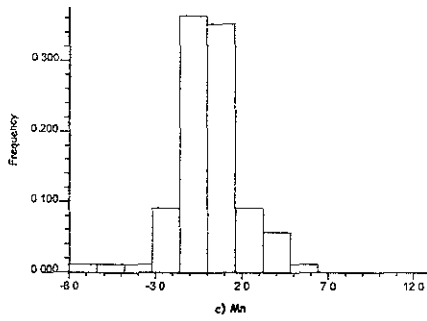
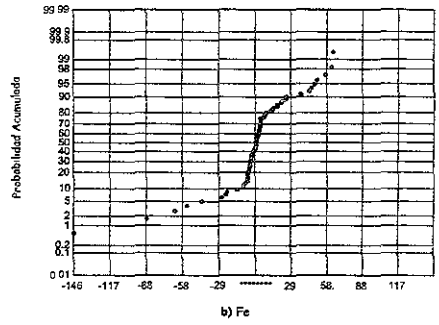
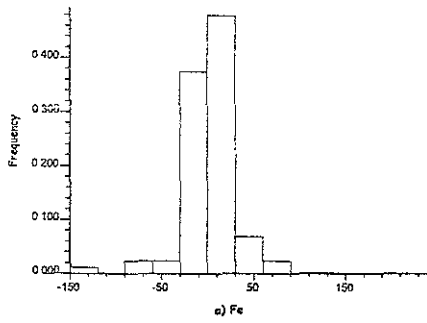


Figura 3.10.2 Histogramas y gráficas en papel normal para $e(\text{Fe})$, $e(\text{Mn})$ y $e(\ln(\text{Fe}))$ por niveles de profundidad, Segundo Muestra

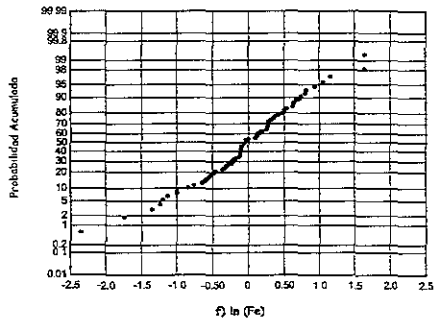
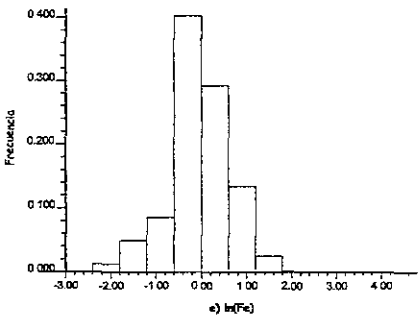
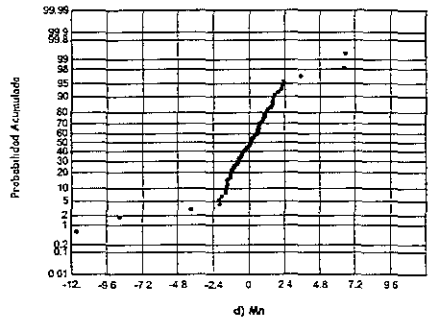
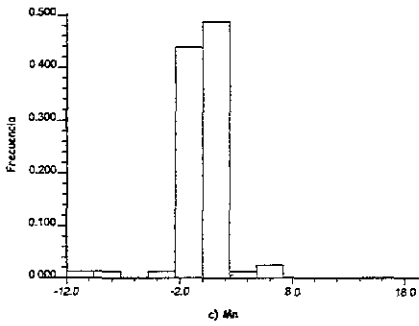
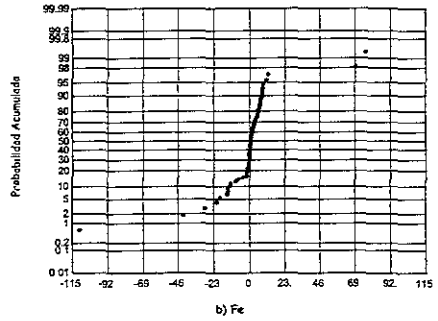
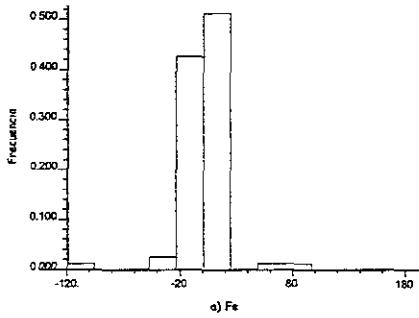


Figura 3.10.3 Histogramas y gráficas en papel normal para $e(\text{Fe})$, $e(\text{Mn})$ y $e(\ln(\text{Fe}))$ por niveles de profundidad, Tercer Muestreo

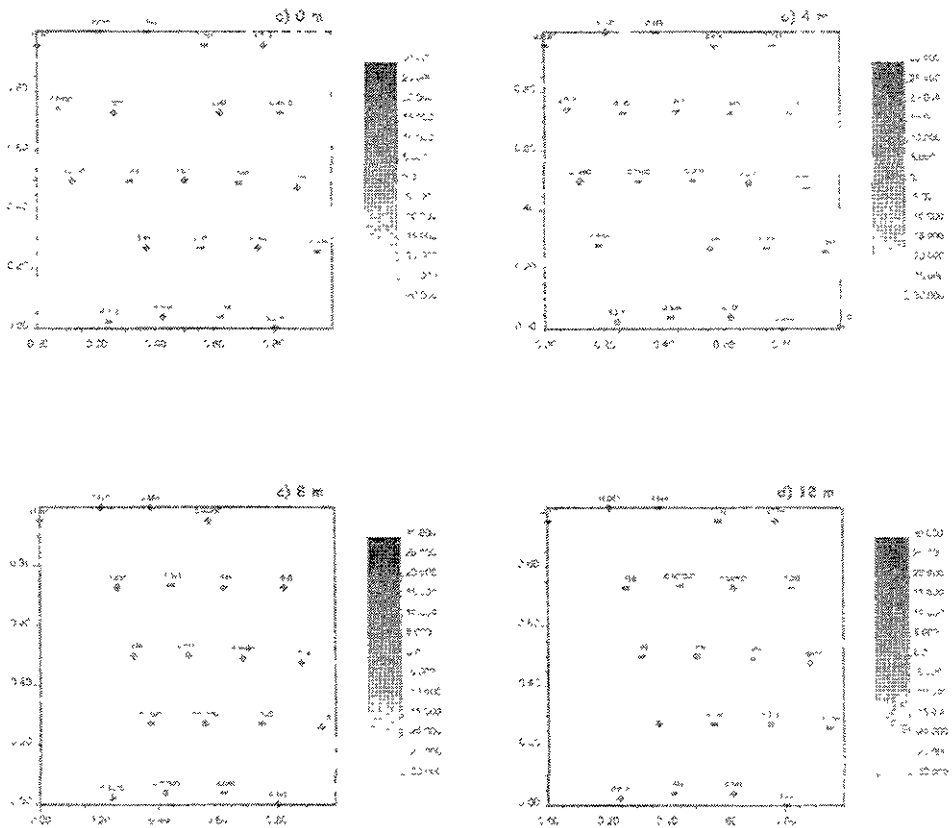


Figure 3.11.1 Mapa de Datos para e(Fe) por nivel de profundidad, Primer Muestreo

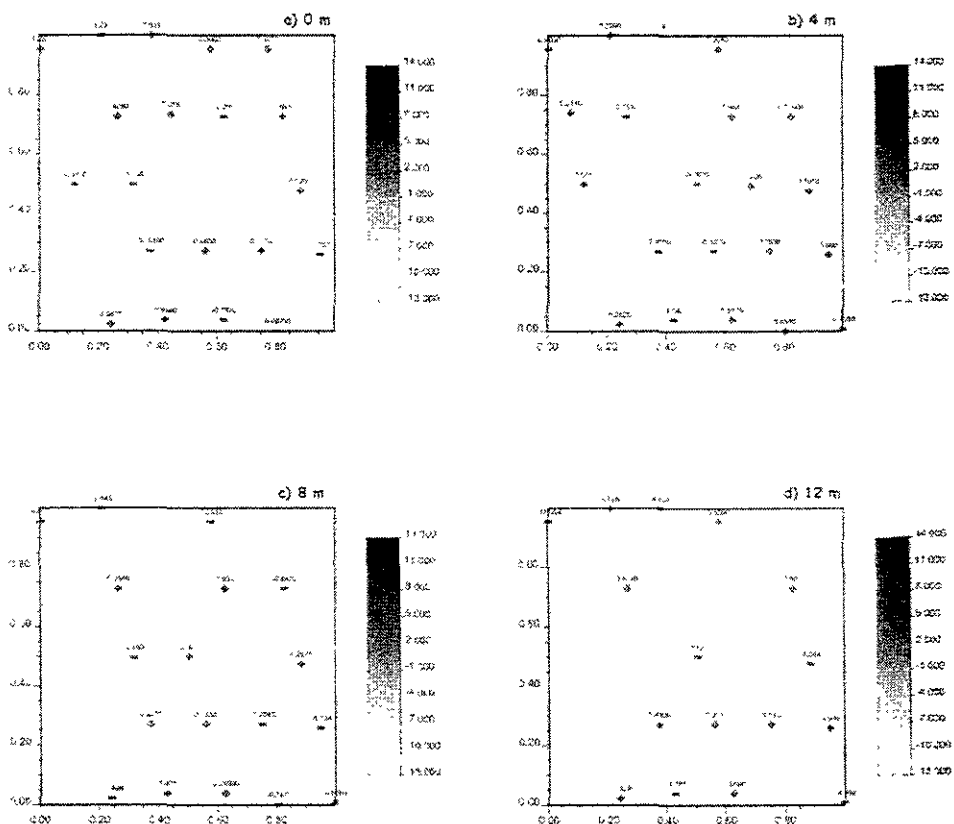


Figura 3 11.2 Mapa de Datos para e(Mn) por nivel de profundidad, Primer Muestreo

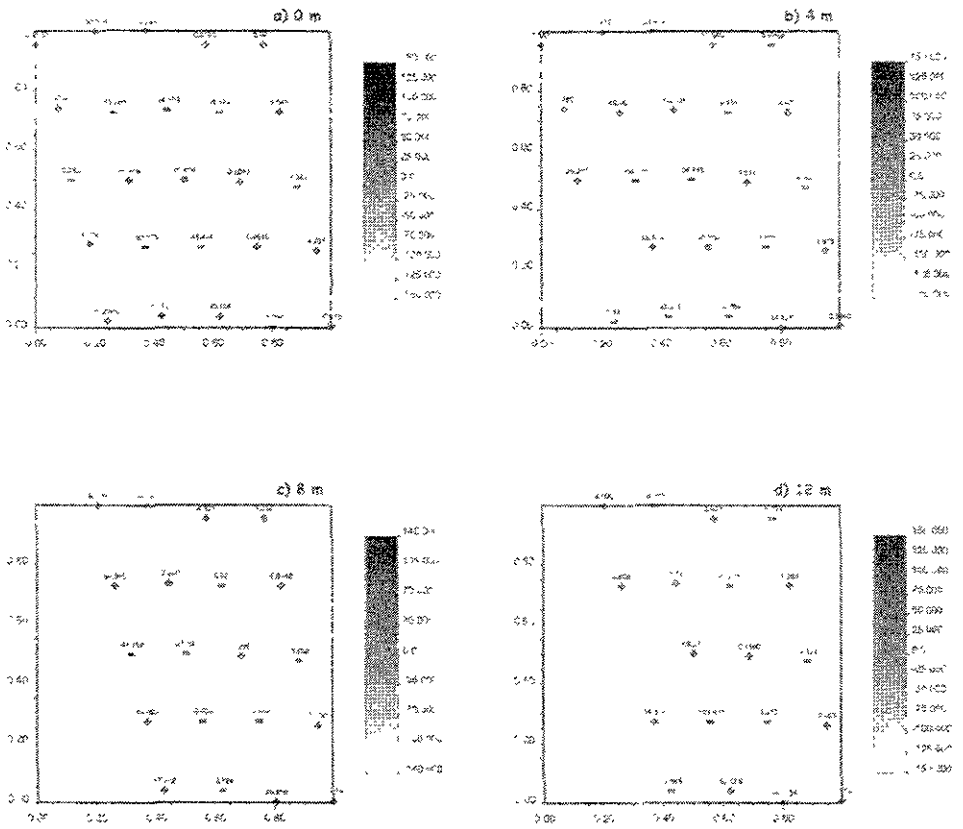


Figura 3.11.3 Mapa de Datos para a(Fe) por niveles de profundidad, Segundo Muestreo

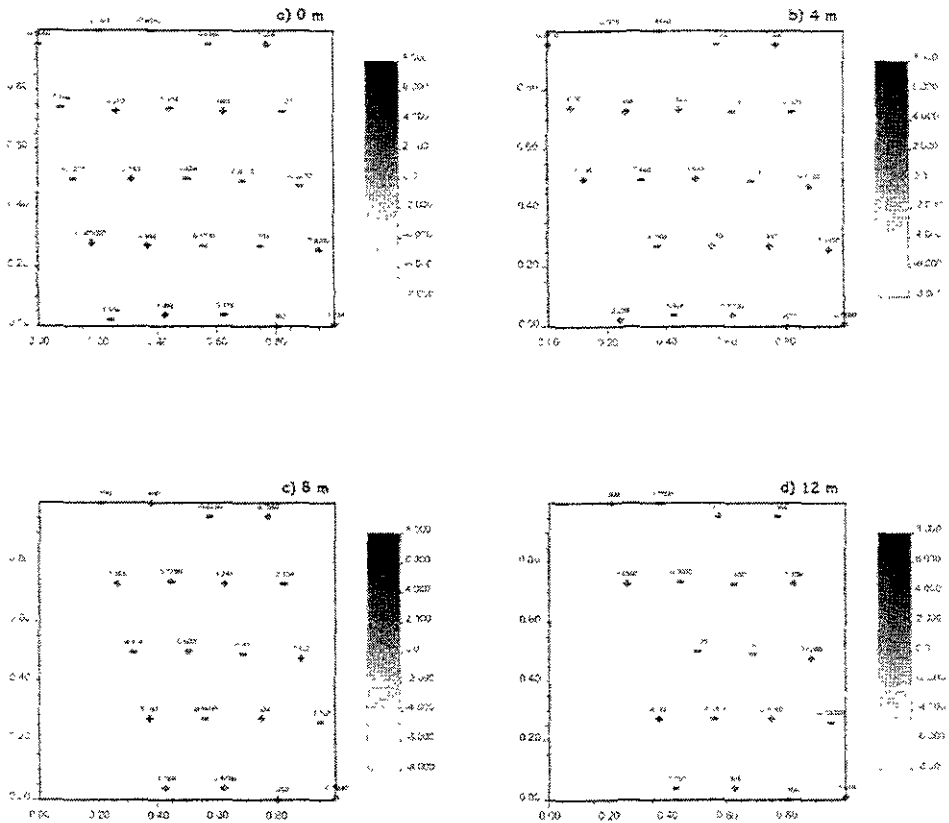


Figura 3.11 4 Mapa de Datos para e(Mn) por nivel de profundidad, Segundo Muestreo

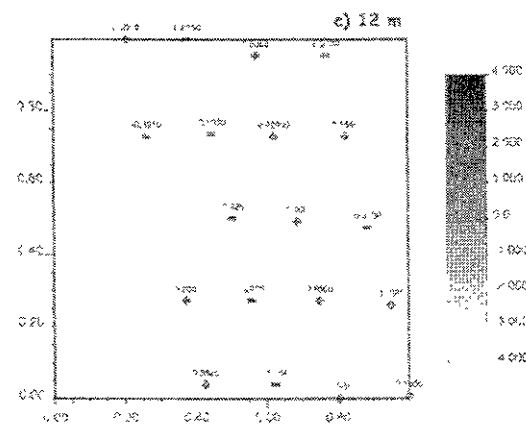
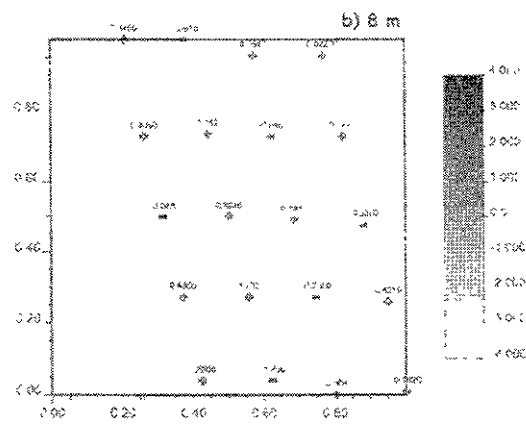
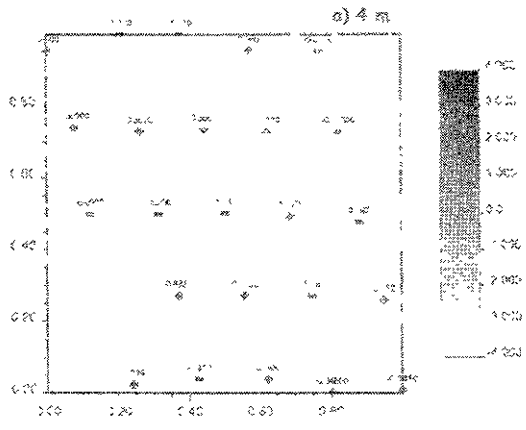


Figura 3.11.5 Mapa de Datos para $e(\ln(\text{Fe}))$ por nivel de profundidad. Segundo Muestreo

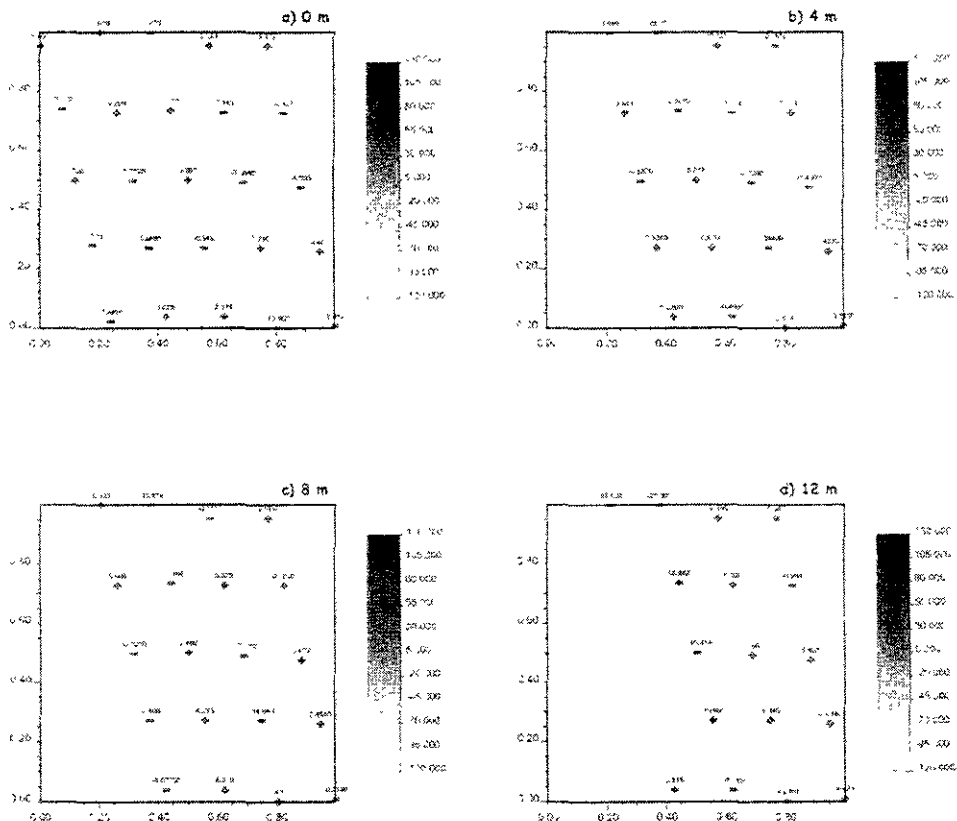


Figura 3.11.6 Mapa de Datos para e(Fe) por nivel de profundidad, Tercer Muestreo

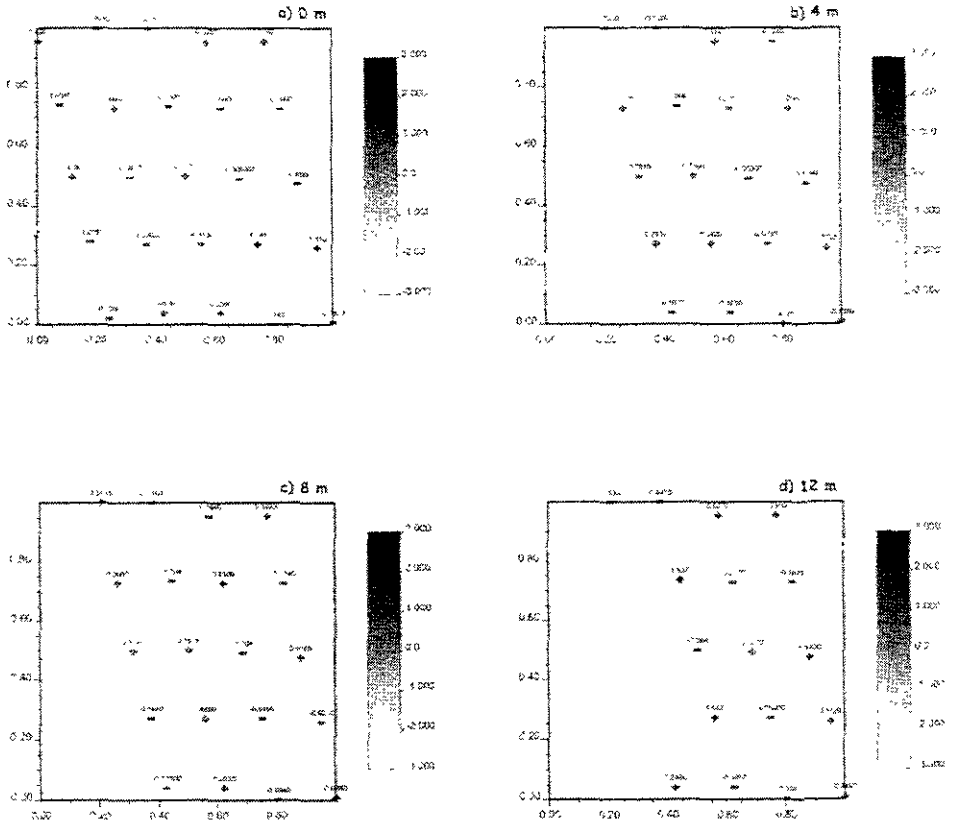


Figura 3 11.8 Mapa de Datos para $e(\ln(\text{Fe}))$ por nivel de profundidad, Tercer Muestreo

Al igual que en el muestreo anterior la variable $\ln(\text{Fe})$ muestra una buena distribución de los errores, ver Figura 3.10.3 (e) y (f). La ubicación de los valores extremos está en parte central de la malla a los 12 m, ver Figura 3.11.8.

3.6 Discusión

3.6.1 Antecedentes

Los sólidos, entre otros efectos, incrementan la turbidez del agua, en elevadas concentraciones inhibe la penetración de la luz solar y, por consecuencia, la fotosíntesis de la flora acuática; puede ocasionar la muerte de los peces por el taponamiento de las vías respiratorias o las de los organismos bentónicos, a causa de la obstrucción debida a la acumulación de sólidos en su habitat, lo que puede producir asfixia (IMP, 1987).

En México se recomienda que para el desarrollo de la vida acuática, la concentración de sólidos en la región costera no debe reducir la profundidad de compensación de la fotosíntesis en más del 10% (Diario Oficial, 13-XII-89). Al respecto, en el área no se cuenta con registros de esta naturaleza, pero debido a los valores de sólidos totales tan elevados, es muy factible que se exceda dicha condición (SEDEMAR, 1992).

Dentro del orden de toxicidad de los metales pesados para los organismos marinos, el Fe y el Mn ocupan el 11° y 12° lugar entre los 15 más importantes respecto a las fases más sensibles de sus ciclos de vida (Waldichuk, 1977).

El Fe elemental usualmente no es un contaminante significativo en el mar, sin embargo a valores de pH comunes en agua marina, forma principalmente óxidos casi insolubles u óxidos de hidruro (Clark, 1986, Ahrland, 1985).

Por otra parte, el Fe es un factor importante en el transporte de mercurio y de otros metales pesados durante la formación de minerales ferromanganosos. Las concentraciones de

Fe en aguas oceánicas, en general, varían entre 0.003 y 0.0034 mg/L (Alvarez y González, 1986).

En los criterios ecológicos de calidad de agua (CE-CCA-001/89) (Diario Oficial, 13-XII-89), se establece que un nivel máximo permisible de Fe para la protección de la vida acuática en aguas marinas costeras de 0.05 mg/L.

El Mn se puede encontrar en agua con valencias II, III y IV en estado soluble, así como en suspensión y en formas de complejo. Su solubilidad depende de los valores de pH, oxígeno disuelto y agentes quelantes. Los límites establecidos en Estados Unidos, Canadá y en algunos otros países europeos para aguas destinadas a diferentes usos son de 0.05 mg/L. Las sales de Mn sólo se consideran tóxicas para la vida acuática en concentraciones muy elevadas (Rodier, 1981). La Agencia de Protección Ambiental de los Estados Unidos (EPA) dentro de sus criterios de calidad del agua para la vida acuática, establecen que los niveles menores a 0.02 mg/L presentan riesgos mínimos y de 0.1 mg/L en adelante ponen en peligro la vida marina (Van der Leeden et al., 1990).

Al igual que el Fe, el Mn a valores de pH comunes en agua de mar, forman óxidos prácticamente insolubles u óxidos de hidruro (Ahrland, 1985). Su efecto como contaminante antropogénico se considera mínimo, pero es un factor de importancia en el transporte del mercurio y otros metales pesados durante la formación de minerales ferromanganosos (Alvarez y González, 1986).

El máximo valor permisible establecido en los criterios ecológicos de la calidad del agua (CE-CCA-001/89) (Diario Oficial, 13-XII-89) para el Mn en agua marina, no está definido, por lo cual las concentraciones en promedio no pueden ser comparadas.

3.6.2 Resultados

Si bien las notas anteriores suenan alarmantes y lo son cuando se comparan los valores reportados con los de la Tabla 3.1 y se observa que son mayores a los permitidos, también lo es cuando no se toman medidas pertinentes para remediar este tipo de problema.

Al relacionar los resultados observados durante todo el trabajo y tomando en cuenta otro tipo de factores que ayuden a comprender el porqué del problema será mucho más fácil entender la situación y tal vez contemplar una posible solución.

a) Fe

Durante la temporada de secas, la distribución del Fe en la columna de agua, registró los valores más altos en las parte Norte y Sur de la malla. La concentración del metal se fue incrementando conforme la profundidad aumentaba. Fue evidente la presencia del difusor y la influencia de las corrientes marinas locales. En la parte Noroeste de la malla, que fue donde se dieron las concentraciones más altas del metal, fue influenciada por una descarga ubicada al Norte de la planta de Du Pont.

En la época de lluvias, las mayores concentraciones se dieron en la parte central de la malla distribuyéndose hacia la zona Sur. La corriente marina no ayudo en mucho a la dispersión del metal pues este se sedimenta en el fondo según se puede apreciar en los mapas.

El comportamiento del Fe en la columna de agua en la temporada de Nortes, tiene un comportamiento encontrado ya que la distribución se da contraria en los dos últimos niveles, pues las direcciones de las concentraciones van en direcciones Sur-Norte y viceversa, posiblemente ésto se deba a la corriente marina que, según los informes, es de Noreste-Sureste.

b) Mn

El comportamiento del Mn en la época de secas fue similar al del Fe ya que las concentraciones altas se dieron en las partes Norte y Sur de la malla pero el valor mas alto fue en la parte Noroeste al nivel de 12m. Como se mencionó anteriormente, posiblemente ésto se debe a la presencia de otra descarga que se ubica al Norte de compañía Du. Pont. La variación vertical del metal en la columna de agua fue menor que la del Fe pero se dio de manera similar, es decir, conforme aumentaba la profundidad había mayor concentración de Mn.

En la época de lluvias, el Mn se comportó también parecido al Fe, con la apariencia de que fue menos denso que el Fe pues las concentraciones altas se dan en los niveles 4 y 8 m y la dispersión del metal se da con mayor homogeneidad que en el Fe.

La distribución del Mn en la columna de agua durante la temporada de nortes fue relativamente homogénea en los tres primeros niveles. A los 12 m se nota una alta concentración en la parte Noreste de la malla. También fue posible apreciar la presencia de cantidades importantes de Mn en superficie y probablemente se deba a la influencia de alguna otra fuente.

Conclusiones

El análisis de los datos y la metodología aplicada a través de todo este trabajo lleva a las siguientes conclusiones:

- El análisis descriptivo de la información y la ubicación de la misma referenciada de manera espacial, ayuda a comprender mejor la problemática del área muestreada.
- El trabajar la información original resulta interesante por que los resultados que se generarán será más veraces y la alternativa de transformar alguna variables ayuda a tener una mejor predicción de los datos cuando ésta es adecuada.
- El considerar una variable más de ubicación, en este caso la profundidad, fue buena para el análisis en general, pues presentó una mejor predicción de los datos más apegada a la realidad y ésto se confirma con el decremento de la variabilidad en cada caso.
- Las estaciones de tiempo, en este caso, no mostraron correlación alguna pues cada temporada podría considerarse independiente o afectada por otras variables no controlables como lo fueron la lluvia, la velocidad del viento, etc.
- El análisis geoestadístico ofrece una respuesta rápida y eficiente para la evaluación de modelos de difusores.
- Aunque el análisis que aquí se presenta es del año de 1992 para una caso muy específico, la metodología puede ser usada para la evaluación de otros difusores, para la evaluación de otras zonas contaminadas por diferentes factores o simplemente para la valoración de zonas ambientales, minerales, etc.

- La geoestadística, en general, muestra una mejor alternativa para el análisis de variables georeferidas desde el punto de vista estadístico, aun siendo sólo de manera descriptiva.
- Las bibliotecas de GSLIB son una buena alternativa para el análisis de tres dimensiones.

Comentarios Adicionales

El análisis particular de estos conjuntos de datos aún no ha terminado, pues aún quedan cosas por hacer como proponer otros métodos de krigado, ya sea por kriging como un modelo de tendencia o el de median polish. También analizar estas dos variables de manera conjunta con Cokriging y el evaluar los paquetes utilizados para el análisis. Estos son trabajos que se contemplan hacer en el futuro.

Apéndice 1 Especificaciones en GSLIB

A.1 Especificaciones del Modelo para Variograma

Cada tipo de estructura requiere de 2 o 6 parámetros que definan su propia anisotropía geométrica en 2 o 3 dimensiones. La Figura A.1 muestra los parámetros de ángulo y rango para el caso de 2 dimensiones.

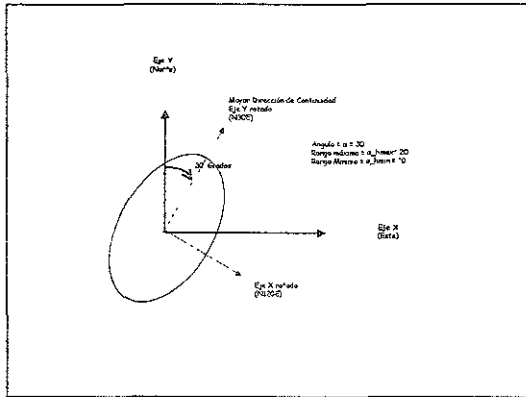


Figura A.1 Ejemplo de los dos parámetros necesarios para definir la anisotropía geométrica de la estructura del variograma en dos dimensiones

La dirección de máxima continuidad es especificada por la rotación del ángulo correspondiente al ángulo azimutal que es medido en grados y con dirección en el sentido de las manecillas del reloj; el ángulo 0 está ubicado en la dirección positiva de Y o dirección Norte (N). El rango en esta dirección horizontal máxima es especificado como a_{hmax} (meseta). El rango en la dirección perpendicular o la mínima dirección horizontal es especificada por a_{hmin} .

Un rango grande implica que la contribución de la varianza no será agregada a la dirección, una característica conocida como anisotropía zonal.

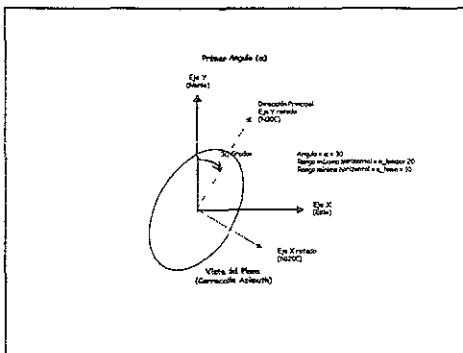
Dentro de las bibliotecas de GSLIB, la distancia es corregida tal que ésta se cuenta para una anisotropía específica. Esto es, la componente de la distancia a lo largo del eje rotado X es corregida por el radio a_{h_max}/a_{h_min} . Esta convención permite que el modelo de potencia anisotrópico sea manejado de una manera intuitivamente correcta; la distancia anisotrópica es calculada y la potencia w permanece sin cambio.

La Figura A.2 muestra sólo los ángulos y rangos requeridos en 3 dimensiones. Muchos programas toman un camino más corto y solamente utilizan 2 ángulos y 2 rangos. La complejidad agregada de los tres ángulos no es la programación sino la documentación. Esto está completamente estructurado para visualizar y documentar un fenómeno que es sumergido con respecto a la horizontal, a la profundidad azimuthal que no está alineada con un eje coordenado. El tercer ángulo es requerido para el conteo concepto de hundimiento o de rastreo. Un ejemplo que requiere el tercer ángulo es modelar la anisotropía geométrica dentro de la extremidad de un sinclinal.

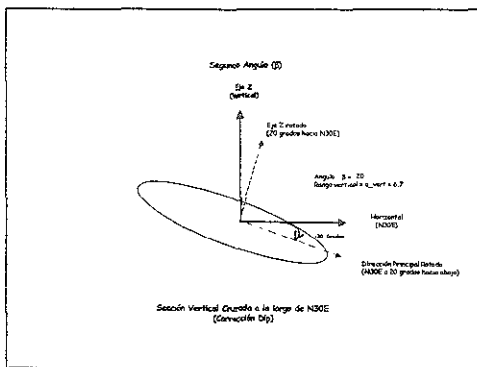
La manera más sencilla de describir los 3 ángulos y los tres rangos es la de imaginar las rotaciones y el aplastamiento que se requiere para transformar una esfera en un elipsoide. Se referirá al eje original Y como la dirección principal y se consideraran las rotaciones tal que su terminación será la dirección estructural principal, la máxima continuidad:

- La rotación del primer ángulo (α) rota del eje original Y (dirección principal) en el plano horizontal, este ángulo es medido en el sentido de las manecillas del reloj.
- La rotación del segundo ángulo (β) la dirección principal desde la horizontal, este ángulo es medido en grados negativos hacia abajo desde la horizontal.
- La rotación del tercer ángulo deja la dirección principal definida por α y β sin cambio. Las dos direcciones principales ortogonales a la dirección principal son rotadas en el sentido de las manecillas del reloj, a la dirección principal cuando se vea hacia el origen.

Primer Paso



Segundo paso



Tercer paso

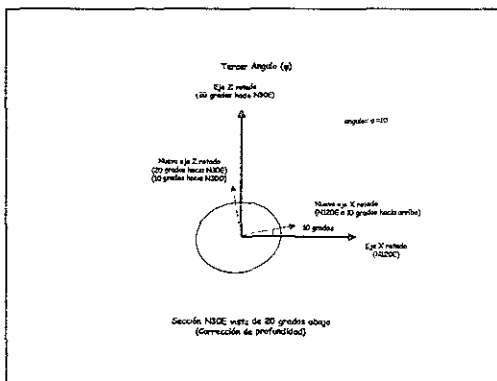


Figura A.2 Ejemplo de los seis parámetros necesarios para definir la anisotropía geométrica de la estructura del variograma en tres dimensiones

- La rotación del tercer paso en la Figura A.2 aparece como *segundero* de un reloj y *girando como tal*.

A.2. Estrategias de Búsqueda

La mayoría de los algoritmos de kriging y simulación consideran un número limitado de datos al considerar una vecindad. La primera razón por la cual se debe limitar es por el tiempo de procesamiento de la computadora y a los requerimientos de memoria. El tiempo de proceso requerido para resolver un sistema de ecuaciones de kriging se incrementa en el número de datos al cubo, es decir, si el número de datos se duplicara implicaría que el tiempo de proceso sería de 8 veces. Con lo que respecta a los requerimientos de almacenaje para la matriz principal de kriging se vería incrementada en el número de datos al cuadrado, por ejemplo, si el número de datos se duplicara, la memoria ocupada se incrementaría en 4 veces más.

Además, si se adopta la búsqueda global de la vecindad, se requeriría del conocimiento de la covarianza para una mayor separación de la distancia entre los datos. Usualmente la covarianza es poco utilizada en la segunda mitad o en la última tercera parte de las distancias. La búsqueda local en la vecindad no toma los valores de la covarianza fuera del elipsoide de búsqueda.

La tercera razón para limitar la vecindad de búsqueda es la de permitir reescalar la media local cuando se aplica el kriging ordinario. Todos los datos deberán ser estandarizados de manera conjunta para crear un histograma y variograma confiables. Sin embargo es mejor relajar la decisión de estacionariedad de manera local y usar solamente los datos cercanos.

Las condiciones que son utilizadas para establecer cuáles son los datos cercanos a ser incluidos en el análisis son las siguientes:

1. Solamente aquellos datos que caen dentro del elipsoide centrado en la zona a estimar serán tomados en cuenta. Esta búsqueda del elipsoide anisotrópico se determina de la

misma manera que en el variograma, es decir, la búsqueda del radio en dos o tres direcciones.

2. Los datos permisibles deberán ser restringidos por un número máximo previamente especificado. Los datos captados por este máximo son los que serán utilizados para el análisis. En todos los programas y en la simulación donde los datos originales son buscados independientemente de los nodos simulados, los más cercanos serán medidos por la distancia euclidiana (posible anisotropía). En el algoritmo secuencial, los nodos de la red simulada previamente son buscados por el variograma de la distancia.
3. La búsqueda del octavo está disponible como una opción para asegurar que los datos son tomados por todos los lados del punto que está siendo estimado o predicho. Esto es particularmente importante cuando se trabaja en tres dimensiones con los datos alineados a lo largo de la malla muestreada, pues garantiza que en la búsqueda serán tomados en cuenta más de un punto muestreado. La búsqueda del octavo deberá ser especificada por la elección del número de datos que serán considerados en cada uno de ellos.

Si además son pocos los datos encontrados, entonces la localización del nodo considerado será menos uniforme. Esto restringe la predicción o simulación a áreas donde los datos son suficientes.

Son 3 los tipos de algoritmos de búsqueda han sido implementados en diferentes programas:

Búsqueda Exhaustiva: La aproximación más simple es la de corroborar sistemáticamente todos los datos y tomar el número máximo de los más cercanos (ver segunda condición) que consideran las tres condiciones antes mencionadas. Esta estrategia es ineficiente cuando los datos son muchos y se estuviera adoptando solamente la estructura del programa en dos dimensiones (Deutsch & Journel, 1998).

Búsqueda de Súper Bloque: Esta estrategia de búsqueda particiona los datos dentro de una red sobrepuesta en el área considerada. Cuando se está prediciendo un punto entonces es posible limitar la búsqueda a aquellos datos que caen dentro de los súper bloques. Esta búsqueda ha sido adaptada para aquellos datos que no tienen una estructura bien definida en la mayoría de los programas de kriging y simulación (Deutsch & Journel, 1998).

Búsqueda Espiral: Esta estrategia de búsqueda es para aquellos datos que tienen una estructura bien definida. La idea es la de visitar el área más cercana a los nodos de la red y después en forma espiral, hasta que cada uno tenga una cantidad suficiente de datos o que se hayan rebasado los límites de búsqueda. Esta estrategia ha sido adaptada en todos los programas de simulación secuencial, (Deutsch & Journel, 1998).

Bibliografía

Ahrland, S. 1985. Inorganic Chemistry of the Ocean. In: Irgolic, K. J. And A.E. Martell (eds.). Environmental Inorganic Chemistry. V. C. H. Publisher, Inc. USA. P. 65-88.

APHA. 1985. Standard Methods for the Examination of Water and Waste Water. American Public Health Association . American Water Works, Pollution Control Federation, 16th de. Washington, D. C. USA.

Atkinson, A. C. 1985. Plots, Transformations, and Regression. An introduction to graphical methods of diagnostic regression analysis. Oxford University Press.

Alvarez, R. y C. González. 1986. Niveles actuales de hidrocarburos y Metales Pesados en aguas costeras del Océano Pacífico Mexicano. Revista del Instituto Mexicano del Petróleo. Vol. XVIII, No. 3. Julio de 1986. P 59-68.

Capurro, L.A. 1972. La Circulación Oceánica en el Golfo de México. Mem. IV Congr. Nac. Ocean. (México): 3-12.

Chica Olmo, M. 1988. Análisis Geoestadístico en el Estudio de la Explotación de los Recursos Minerales. Tesis Doctoral. Universidad de Granada. Granada, España.

Clark, R. B. 1986. Marine Pollution. Claredon Press Oxford.

Cox, R. A., R.J. Mc Cartney and F. Culkin. 1970. The specific gravity/ salinity/ temperature relationship in natural sea water. Deep Sea Res. 17: 679-689.

Cressie, N. 1989. *Geostatistics*. The American Statistician, Vol. 43 No. 4. Pp. 197-202.

Cressie, N. 1991. *Statistics for Spatial Data*. John Wiley & Sons.

Deutsch, C. V. and A. G. Journel. 1998. *GSLIB: Geostatistical Software Library and User's Guide*. Second Edition. Oxford University Press.

Diario Oficial. 1989. Acuerdo por lo que se establecen los criterios ecológicos de calidad del agua CE-CCA-001/89. Secretaría de Desarrollo Urbano y Ecología. Diario Oficial de la Federación.

Diaz Avalos, C. 1992. *Construcción de una Red de Monitoreo Oceanográfica utilizando la Teoría de Variables Regionalizadas*. Tesis de Maestría. UACPyP, UNAM. México.

Diggie, P. J., Tawn, J. A. and Moyeed, R. A. 1998. *Model-Based Geostatistics*. Applied Statistics, Vol. 47 pp. 299-350.

Englund, E. and Sparks, A. 1991. *GEO-EAS 1.2.1 User's Guide*. United States Environmental Protection Agency.

Gendzwill D. J and Stauffer, M. 1981. Analysis of triaxial ellipsoids: Their shapes, plane sections, and plane projections. *Math. Geology*. 13(2): 135-152.

Goovaerts, P. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press.

IMP. 1987. *Memorias del Curso Internacional "Impacto Ambiental"*. Subdirección de Desarrollo Profesional. Instituto Mexicano del Petróleo. 178 pp.

Isaaks, E. and Srivastava. 1989. *An Introduction to Applied Geostatistics*. Oxford University Press.

- Journel, A. G and Huijbregts CH. J. 1978. *Mining Geostatistics*. Academic Press.
- Rangel Montes, V. R. 1988. Bioensayos con aguas residuales de Du Pont, S.A. de C.V. División Pigmentos y Productos Químicos, Planta Altamira, Tamps. México, para determinar el límite de tolerancia media de la lebrancha (*Mugil curema*) Cuvier y Valenciennes. Universidad del Noreste. Tampico, Tams. Escuela de Ciencias Biológicas. Tesis profesional.
- Rodier, J. 1981. *Análisis de las Aguas , Aguas Naturales, Aguas Residuales, Aguas de Mar*. Ediciones Omega. Barcelona. 1059 pp.
- Royle A, Clark, I, Brooker, P. I, Parker, H, Journel, A, Rendu, J. M, Sandefur, R, Grant D. C, Moussett-Jones, P. 1980. *Geostatistics*. Mc Graw-Hill.
- SEDEMAR. 1992. *Caracterización del Ambiente Acuático en la Zona de Descarga del Difusor Submarino de Du Pont, S. A. de C. V., Planta Altamira*. Reporte de Investigación.
- Van Der Leeden, F., F. L. Troised and D. Keith Todd. 1990. *The Water Encyclopedia 2nd de Geraghty & Miller Ground Water Series*. Lewis Publishers. 808 pp.
- Velleman, P. F. and Hoaglin D. C. 1981. *Applications, Basics and Computing of Exploratory Data Analysis*. Chapter 8 : Median Polish. Duxbury Press.
- Waldichuk, M. 1977. *La Contaminación Mundial del Mar: Una Recapitulación*. Comisión Oceanográfica Intergubernamental/ UNESCO.