

**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

**DIVISIÓN DE ESTUDIOS DE POSGRADO
FACULTAD DE INGENIERÍA**

ANÁLISIS GEOESTADÍSTICO

**DE LA SISMICIDAD OBSERVADA EN LA ZONA DEL EMBALSE
DE LA PRESA AGUAMILPA, ESTADO DE NAYARIT**

T E S I S

**QUE PARA OBTENER EL TÍTULO DE
MAESTRO EN INGENIERÍA**

(EXPLORACIÓN DE RECURSOS ENERGÉTICOS DEL SUBSUELO)

P R E S E N T A :

MIGUEL MACÍAS CARRASCO

**DIRECTOR DE TESIS:
DR. VINICIO SURO PÉREZ**

CIUDAD UNIVERSITARIA

DICIEMBRE DE 2000



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

Para

mi esposa

Ana Gabriela

y mis hijos

Sami y Gabrielito

Agradecimientos.

Al Dr. Vinicio Suro Pérez por sus enseñanzas, su paciencia y tiempo, pero sobre todo por la motivación constante para hacer este trabajo.

Al Dr. Antonio Uribe Carvajal por su total apoyo para la realización de mis estudios de maestría y por compartir su valioso conocimiento sobre el tema. También, a todos los compañeros de trabajo por sus comentarios y ayuda así como al Ing. José Antonio Maza Álvarez, Gerente de Estudios de Ingeniería Civil de la CFE, por las facilidades otorgadas.

A mis padres, mis hermanos y sus familias, por tener su confianza a lo largo de mi vida.

Resumen

Se analiza la sismicidad registrada dentro del embalse de la presa Aguamilpa, a partir de su llenado, para estudiar la predicción en espacio, sin consideraciones temporales, de la magnitud sísmica. Primeramente se exploraron, con histogramas, datos de magnitud para diferentes ventanas de tiempo que al compararlos con la forma del histograma general, se observó similitud hasta en el caso de información trimestral, independientemente se realizó el análisis del valor de b para poblaciones trimestrales, que resultó consistente con lo anterior. Se hizo estadística bivariada para el cálculo de variogramas para conocer la continuidad espacial de la variable bajo estudio, para lo cual se propuso la hipótesis de que la magnitud podría comportarse como variable regionalizada. Los variogramas experimentales obtenidos permitieron establecer modelos de variabilidad espacial de la magnitud y reflejaron el patrón espacial de esta, mayor continuidad a lo largo de fallas y mayor cubrimiento con el tiempo. Estos modelos fueron la base para la estimación espacial de la magnitud y la incertidumbre asociada, utilizando las técnicas geoestadísticas de kriging ordinario y simulación estocástica secuencial, cuyos resultados se presentaron en mapas de magnitud y mapas de probabilidad de excedencia de valores extremos de magnitud, siendo estos últimos los más adecuados para representar el fenómeno bajo estudio. Por lo tanto, el considerar la magnitud sísmica como variable regionalizada y definir la distribución espacial de esta con modelos de incertidumbre a partir de simulación estocástica, es significativo en la delimitación de fuentes sísmicas y el análisis de peligro sísmico asociado.

CONTENIDO

Resumen.	v
1 Introducción.	1
1.1 Planteamiento del problema.	4
1.2 Antecedentes.	4
1.2.1 Sismicidad.	5
1.2.2 Marco geológico-estructural.	6
1.2.3 Estudios.	8
1.3 Objetivos.	9
2 Hipótesis.	11
2.1 Marco teórico.	11
2.2 Hipótesis.	12
3 Metodología.	13
4 Análisis geoestadístico.	16
4.1 Análisis estructural.	16
4.1.1 Exploración de datos.	16
4.1.1.1 Análisis sin consideraciones temporales.	20
4.1.1.2 Análisis con ventanas de tiempo.	22
4.1.1.3 Análisis comparativo entre ventanas de tiempo.	25
4.1.2 Variogramas.	29
4.1.2.1 Estacionariedad de $Z(x)$.	33
4.1.2.1 Variogramas, isotropía y modelos para $Z(x)$.	35
4.1.2.2 Variogramas indicadores, isotropía y modelos.	43
4.2 Estimación.	47
4.2.1 Técnica de kriging ordinario.	47
4.2.2 Técnica de kriging indicador.	57
4.3 Simulación.	64
4.3.1 Simulación secuencial.	65
4.3.2 Modelo de incertidumbre.	70
5 Investigaciones futuras.	75
6 Discusión y conclusiones.	77
7 Bibliografía.	80

Apéndice 1. 85
Apéndice 2. 88
Apéndice 3. 89

1 Introducción

La sismicidad asociada a grandes acumulaciones de agua, como son las presas o lagos artificiales, es un fenómeno cuyas características son complejas ya que la sismicidad natural propia de la región circundante a estas obras puede incrementarse, en cuanto a número de sismos, magnitud o área sismogénica, o bien, que la sismicidad natural disminuya debido a los cambios en el estado de esfuerzos preexistente (Gupta et al, 1972).

En general, no existe un patrón regular de la sismicidad inducida por presas (Gough, 1978), ya que puede haber desde microsismos con magnitud $M \leq 2$ (magnitud Richter) hasta sismos de magnitud importante $M = 6$. Tampoco hay una relación directa entre mayor volumen y profundidad de una presa con una sismicidad más intensa y numerosa. Se ha observado que al llenar el embalse de una presa puede haber existido sismicidad previa a este llenado, y posteriormente no haber sismicidad o viceversa. Además, puede presentarse sismicidad durante el llenado o después de varios años de concluido éste (Simpson, 1976; Rothé, 1973). Sin embargo, de acuerdo a lo observado por los investigadores citados y otros, las conclusiones sobre sismicidad inducida por presas suponen que (Packer et al, 1977):

- 1) El volumen de una determinada presa no es un indicador confiable del peligro de sismicidad inducida.
- 2) Se puede esperar migración de eventos sísmicos con el tiempo.
- 3) Los mecanismos focales observados en diferentes presas son consistentes con el tipo de esfuerzos preexistentes en la vecindad de las presas.
- 4) Un cambio rápido en el nivel del agua (hacia abajo o hacia arriba) del embalse puede ser un mecanismo importante de sismicidad inducida.
- 5) No es posible dar un límite superior en cuanto a magnitud de sismos inducidos, ya que el llenado del embalse actúa sólo como disparador de los esfuerzos preexistentes, es decir, el esfuerzo regional acumulado.

El caso de la Central Hidroeléctrica Aguamilpa, Nayarit (C.H. Aguamilpa, Nay.), cuya localización es mostrada en la figura 1.1, es un ejemplo notable de sismicidad inducida (Uribe, 1994). La zona definida en un área de 100 x 100 km aproximadamente, con centro en la C.H. Aguamilpa, se consideró prácticamente asísmica previo al llenado del embalse (25 de junio de 1993), habiéndose instalado sismógrafos por parte del personal de la Comisión Federal de Electricidad (CFE) en diferentes puntos a partir de 1987 y hasta la fecha del llenado. Estos sismógrafos permitieron registrar la sismicidad natural de la región, nueve sismos de magnitud menor o igual a 3.1 grados (Delgado, 1992), que agregada a la información sísmica obtenida de catálogos nacionales y mundiales (Council of the national seismic system, 1996) desde 1900, permitió obtener los resultados mostrados en la figura 1.2a. En la figura 1.2b se grafican los sismos localizados (1403) correspondientes a 1994

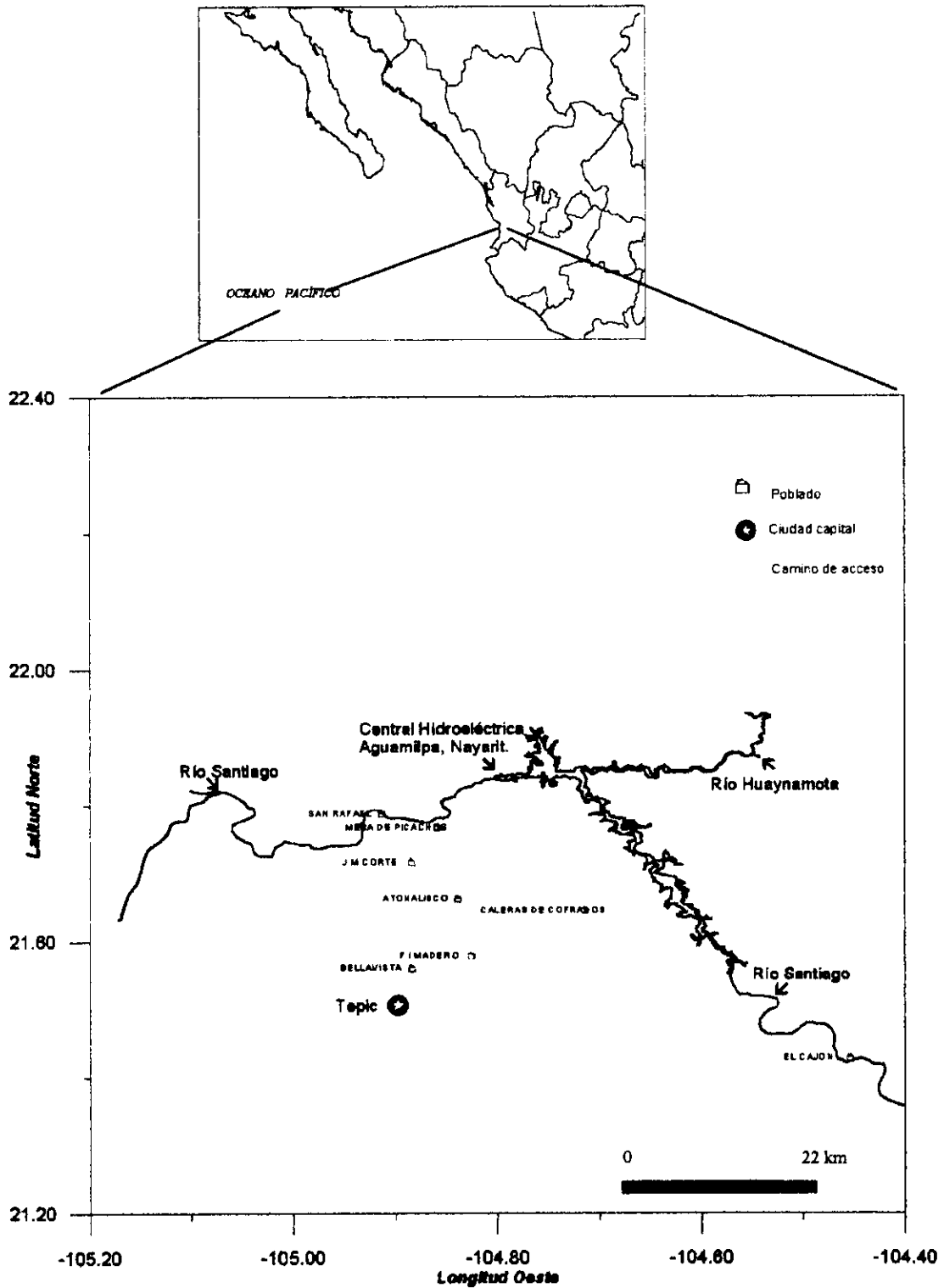


Figura 1.1. Localización de la Central Hidroeléctrica Aguamilpa en el Estado de Nayarit. Se observa la zona del embalse, los ríos que alimentan a la presa y los poblados más cercanos a ésta.

como ejemplo de la sismicidad generada a partir del llenado del embalse. Es importante aclarar que la sismicidad observada en la zona sureste del embalse, dada la poca profundidad de éste en etapas tempranas de sismicidad (1993), no se le asocia a un proceso de cambio significativo en el estado de esfuerzos preexistente. Bajo estas condiciones, es de esperarse que la ocurrencia de sismos en ésta área se relacione con las condiciones geotérmicas en dicha región (Quijano y Chacón, 1985).

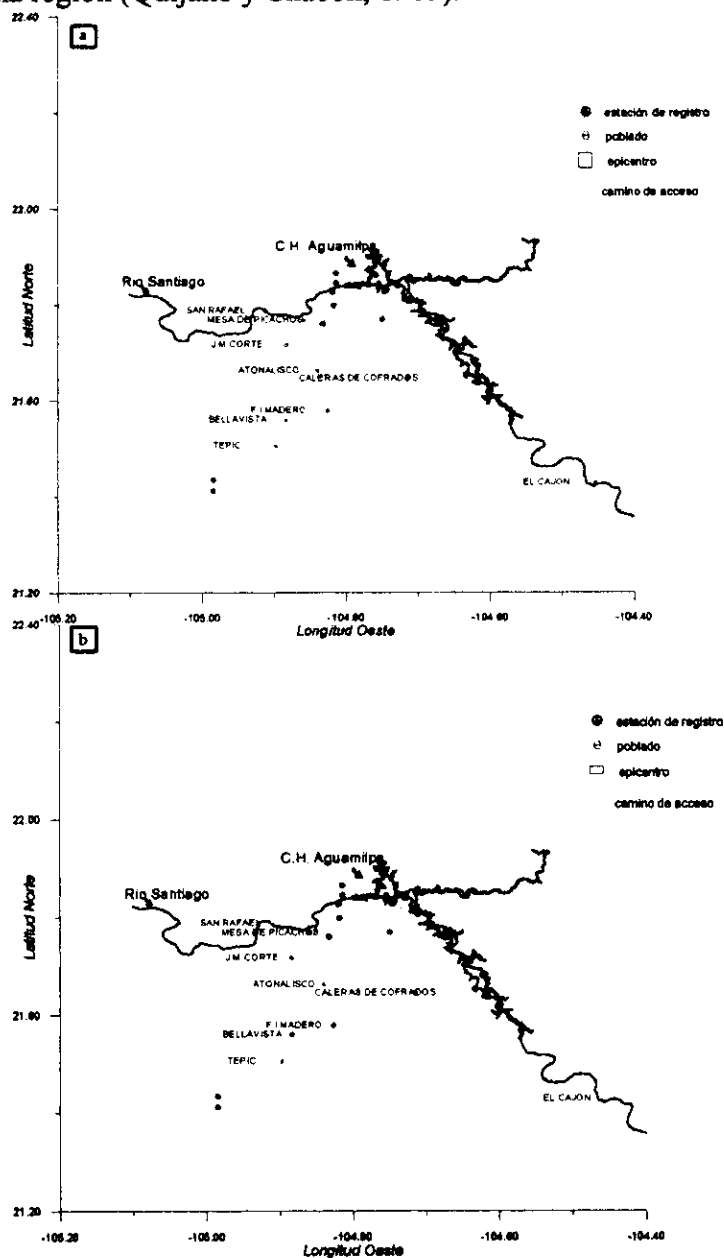


Figura 1.2. Sismicidad en la Central Hidroeléctrica Aguamilpa, Nayarit. (a) Sismicidad natural registrada de 1900 a junio de 1993. El número de sismos es de nueve, con magnitud menor o igual a 3.1 y ninguno cae en el área del embalse. (b) Ejemplo de la sismicidad ocurrida durante 1994, período en el que se registraron 1403 sismos de magnitudes entre 0.5 y 4.63 grados.

1.1 Planteamiento del problema

La sismicidad observada, a partir de julio de 1993 en la zona del embalse de la C.H. Aguamilpa, Nay., puede ser un factor de peligro para la seguridad de las estructuras de la presa y los poblados aledaños a ésta. Durante la evolución de esta sismicidad se ha observado migración de eventos sísmicos (en superficie y profundidad), así como un incremento en la magnitud de los sismos más fuertes. Se ha podido delimitar zonas sismogénicas donde ocurren los sismos de mayor magnitud, las cuales están estrechamente relacionadas con la geometría de las fallas más importantes presentes en el área de estudio (Delgado, 1998).

Bajo estas condiciones la pregunta que surge es, ¿cómo será la distribución espacial en el tiempo de la magnitud sísmica en la zona analizada?. Lo anterior es un cuestionamiento no fácil de responder, sin embargo, una parte de éste es examinado y constituye el problema a resolver en esta investigación: la estimación, sobre el área de interés y sin consideraciones temporales, de valores extremos de magnitud sísmica con sus probabilidades asociadas. Hay que destacar que de antemano se renuncia a la predicción en tiempo de la distribución espacial de la magnitud sísmica. No obstante, se plantea teóricamente el problema de inferencia en el tiempo.

1.2 Antecedentes

El modelado de la distribución espacial de la sismicidad observada en una determinada región, para un intervalo de tiempo específico, tiene implicaciones importantes en el análisis y evaluación del peligro sísmico asociado (Cornell, 1968). Uno de los puntos sensibles en el estudio del peligro sísmico de un determinado sitio o región, es la de definir las fuentes sísmicas que puedan afectar a la zona evaluada. Dentro de las técnicas actuales, mayormente reconocidas en la evaluación de éste peligro, está la del enfoque probabilístico, que requiere de la delimitación de estas fuentes sísmicas. Investigaciones recientes han buscado versiones suavizadas de la sismicidad original de la zona analizada, en lugar de la delimitación subjetiva de las áreas consideradas fuentes sísmicas (Musson y Winter, 1993). En este contexto, los resultados esperados en esta investigación, pueden ser considerados como una alternativa en la delimitación de estas zonas sismogénicas.

Es importante enfatizar que el modelado de los mecanismos que rigen el fenómeno de sismicidad inducida o la evaluación del peligro sísmico asociado, no constituyen el problema a resolver en este trabajo. Sin embargo, los resultados esperados pueden ser de utilidad en ambas cuestiones. Por otro lado, con base en la información sísmica disponible, el área analizada está comprendida entre las coordenadas 104.65° a 104.85° de longitud O y 21.75° a 21.90° de latitud N

1.2.1 Sismicidad

La sismicidad registrada alrededor del embalse de la C.H. Aguamilpa, abarca desde 1987 hasta la actualidad, teniendo un período prácticamente asísmico desde ese año hasta finales de julio de 1993, esto es 30 días después de iniciado el llenado del embalse (25 de junio de 1993). Posteriormente empezó la actividad sísmica, cuya evolución en cuanto a la cantidad, magnitud y migración de sismos se resume a continuación:

- 1) El número de sismos ha variado: de un incremento importante durante el llenado del embalse y los dos siguientes años (1994 y 1995), a un decremento en los tres últimos años (1996 - 1998).
- 2) La magnitud ha evolucionado de una microsismicidad ($M \leq 2.5$) abundante con pocos sismos de mayor intensidad ($2.5 \leq M \leq 4.5$) a una microsismicidad menos frecuente, con menor número de sismos de magnitud mayor a los de la primera etapa ($M = 5.2$, máximo valor registrado).
- 3) Se han relacionado zonas sismogénicas con los sismos de mayor magnitud.

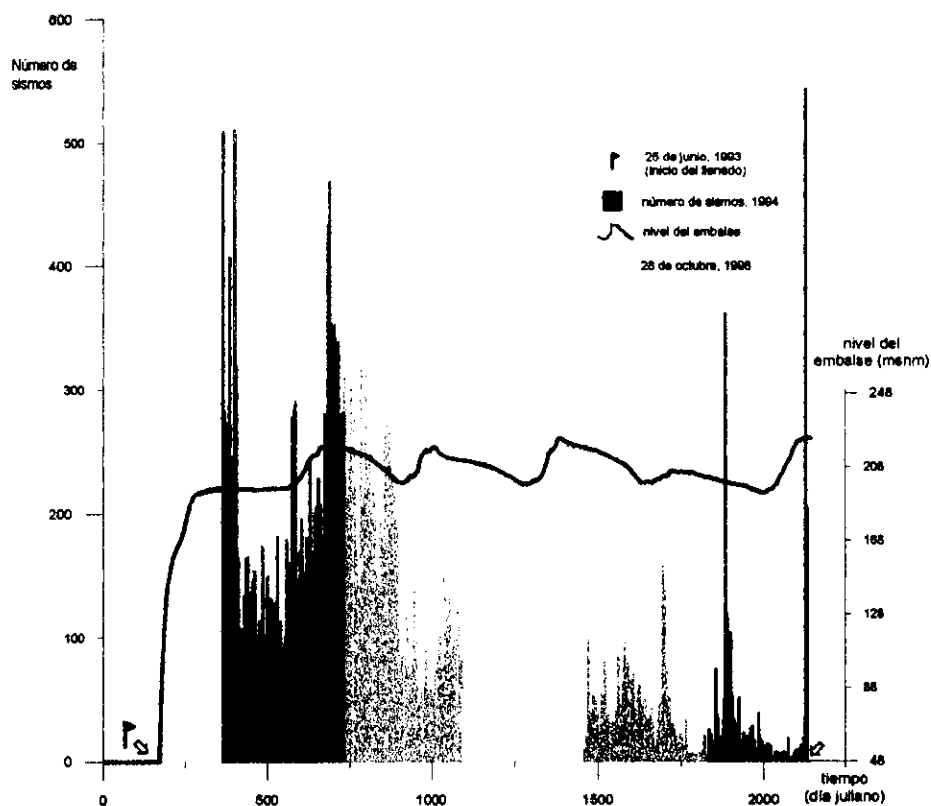


Figura 1.3. Número de sismos y nivel del embalse (metros sobre el nivel del mar) contra el tiempo (día juliano), a partir del llenado de la presa Aguamilpa, Nayarit. Observar la forma cíclica del nivel del embalse, consecuencia de la temporada de lluvias y sequía en la región, y la forma también cíclica del número de sismos en el tiempo.

1.2.2 Marco geológico-estructural

Con el objeto de conocer las fallas y fracturas más importantes que afectan la zona estudiada, es necesario comprender el marco tectónico que rige la dinámica de la corteza terrestre en esta zona. Sin embargo, es importante recordar que la sismicidad regional que se ubica fuera del área analizada y que si bien está determinada por este marco tectónico, no es considerada para efectos del análisis propuesto, ya que se ubica fuera de los objetivos planteados en este trabajo.

La C.H. Aguamilpa se localiza en un área de transición entre dos provincias geológicas: Faja Volcánica Transmexicana (FVT), de edad cenozoica, origen volcánico y ambiente geotectónico de arco continental, y la Faja Ignimbrítica Mexicana (FIM) con las mismas características de la anterior (Ortega, et al., 1992). En la figura 1.4a se muestran estas provincias, mientras que en la figura 1.4b se aprecia la tectónica a nivel de placas que rige en la región, la cual involucra las del Pacífico, Cocos, Rivera y Norteamérica. Están delimitadas por rasgos estructurales de características diversas: La Dorsal del Pacífico con carácter divergente, las fracturas de Rivera y Tamayo como límites conservativos y dinámica de falla de transformación. Además, la zona de subducción denominada Trinchera Mesoamericana que actúa como límite convergente. También, se aprecia una zona que corresponde al denominado Graben Tepic-Zocoalco. La presa está construida sobre una parte del cauce del Río Santiago ubicada en la zona del graben mencionado. Asimismo, se han interpretado los rasgos estructurales de carácter lineal a partir de imágenes de satélite, reflejando dos sistemas de fallas para esta zona y su entorno. Un sistema con orientación NO-SE y el otro sistema (conjugado del primero) con orientación NE-SO. El primer sistema es paralelo a el Graben Tepic-Zocoalco, que es la estructura más importante de esta región, la cual se encuentra prácticamente cubierta por rocas volcánicas pliocuaternarias y cuyos centros de emisión, grandes y pequeños, también presentan una orientación NO-SE, reflejando zonas de debilidad cortical asociadas a la tectónica que afecta la zona estudiada (Delgado et al, 1999).

En la región ha prevalecido un régimen tectónico de carácter extensional a partir del Mioceno Medio a Superior, como probable resultado del cese de la subducción de la Paleoplaca Farallón, e inicio de la deriva de la Península de Baja California hacia el NO. Este régimen se mantuvo durante el Plioceno y Cuaternario, cuando se generaliza la actividad extrusiva de los grandes estratovolcanes presentes en la región, cuya tendencia NO-SE, se ve reflejada en el mismo Río Santiago en su tramo comprendido entre la localidad de las Adjuntas (confluencia de los Ríos Santiago y Huaynamota) y el Río Joraviejo, lineamiento que persiste hacia el NO con las Fallas Rosario Viejo, El Sauz y El Sauz-Guachera. El fracturamiento NE-SO, conjugado al anterior, se refleja principalmente en los lineamientos señalados por los cauces de algunas corrientes superficiales, como el caso de la zonas de fracturas Adjuntas-Palmita y Guachera (figura 1.5). La columna geológica del área de la boquilla en esta central hidroeléctrica está conformada por las unidades de roca denominadas Aguamilpa, Colorín y Picachos, de naturaleza volcánica y composiciones andesítica, riódacítica y traquiandesítica, respectivamente. Estas unidades

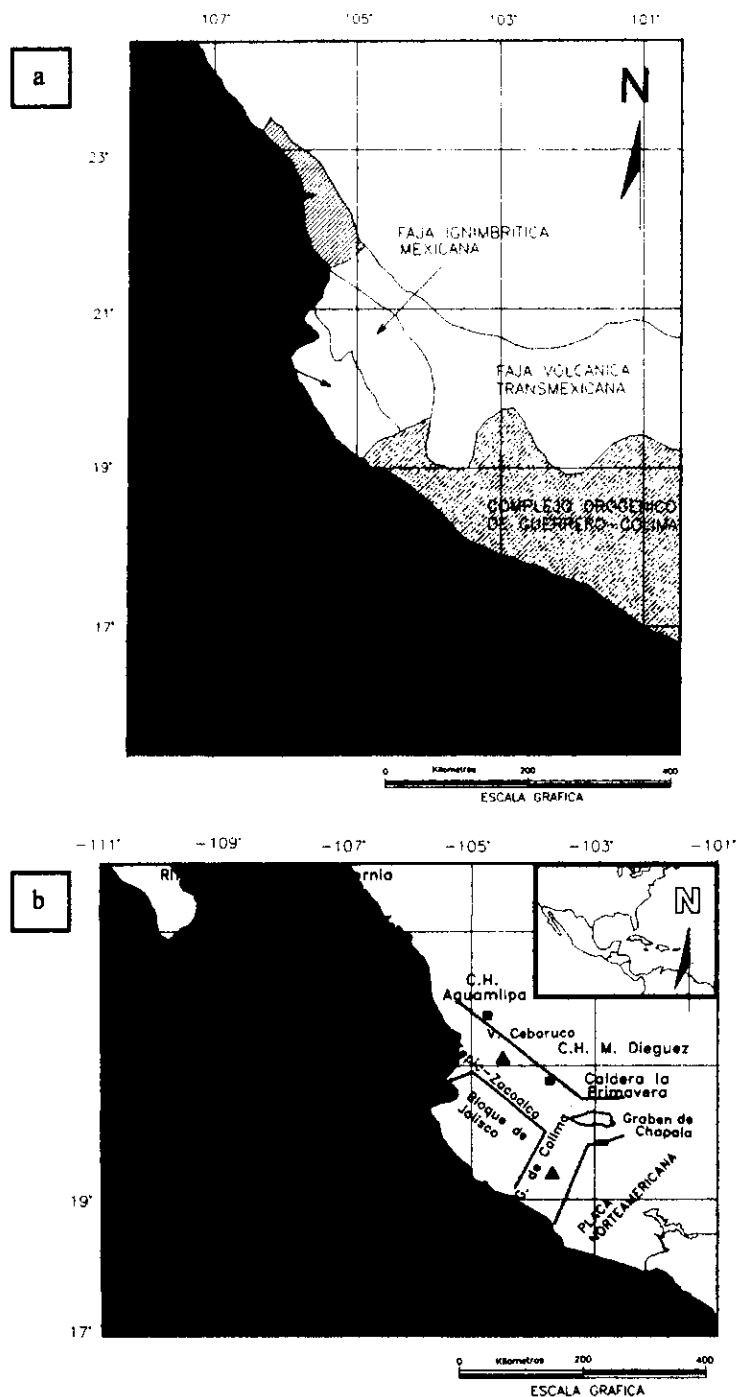


Figura 1.4. Contexto geológico de la Central Hidroeléctrica Aguamilpa, Nayarit. (a) La Faja Ignimbrítica Mexicana y la Faja Volcánica Transmexicana son las provincias geológicas que delimitan la región en el cual se ubica la C.H. Aguamilpa, Nay. También se observan las provincias geológicas adyacentes a las primeras. (b) El Graben Tepic-Zacoalco es la megaestructura representativa del ambiente tectónico en la región estudiada, en el cual se ubica la C.H. Aguamilpa. En la figura se pueden apreciar las principales placas y zonas de fractura involucradas.

fueron afectadas por una serie de intrusivos en forma de troncos y diques, estos últimos emplazados en zonas de debilidad con direcciones muy semejantes a los patrones estructurales actuales (Moreno y Soto, 1994). Cabe añadir que las fallas y fracturas mencionadas son las estructuras más importantes en la zona del embalse de la presa, por lo que su ubicación es relevante para el análisis e interpretación de resultados.

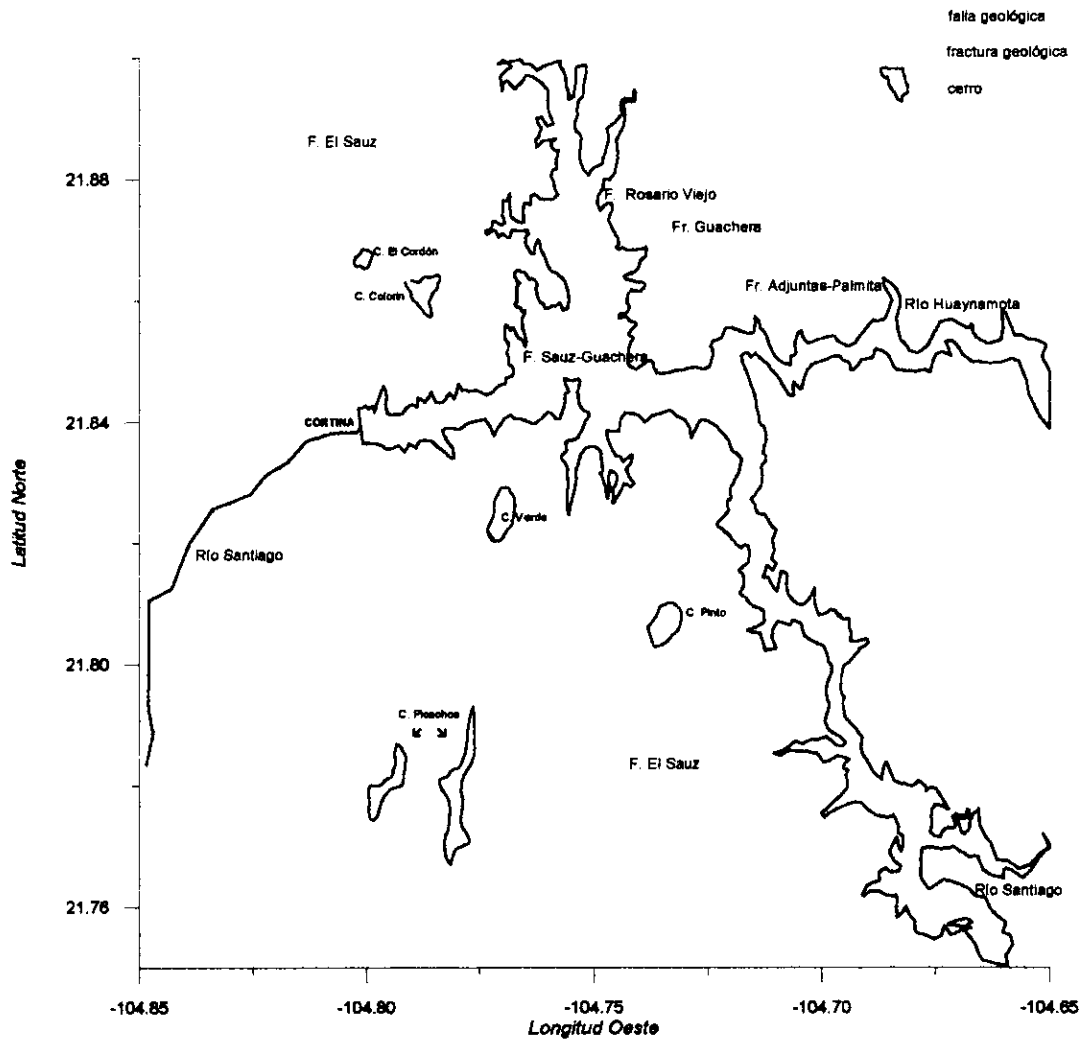


Figura 1.5. Principales fallas y fracturas geológicas presentes en la zona del embalse de la C.H. Aguamilpa, Nay.

1.2.3 Estudios

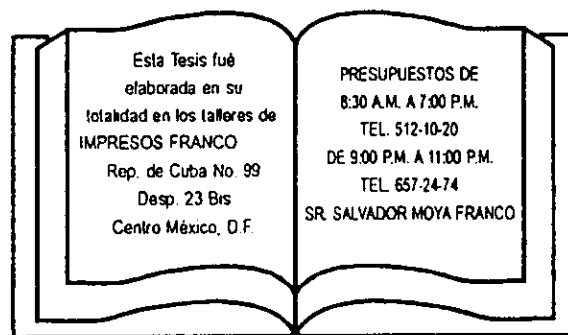
Las primeras investigaciones sobre sismicidad inducida se remontan a la década de los 40 en donde Carder (1945) asoció las variaciones de la sismicidad con el llenado del embalse del Lago Mead, Colorado, E.U. Basados en más de cien casos de sismicidad inducida por presas, Guha y Patil (1990) clasificaron los sismos inducidos en intensos ($M \geq 6$), moderados ($3.1 \leq M \leq 5.9$) y microsismos ($M \leq 3.0$). Gupta (1992) compiló más de 70 casos de sismicidad inducida y sólo en pocos casos observó decremento en la sismicidad con el llenado del embalse. Baecher y Kenney (1987) encontraron que la profundidad de un embalse y el volumen de este, son el primer y segundo atributo respectivamente, de mayor correlación con la sismicidad inducida. Por otro lado, el problema de la estimación de la magnitud máxima esperada de sismos inducidos por presas ha sido estudiado con enfoques diversos. Srivastava y Dube (1997) utilizaron análisis de componentes principales para expresar la magnitud máxima en términos de parámetros como la profundidad, volumen y defasamiento en tiempo de la ocurrencia de sismos con respecto a la historia del llenado del embalse. Otras referencias de casos en donde se han obtenido modelos de estimación de magnitud de sismicidad inducida por presas, en función de los parámetros ya mencionados, son: Habibagahi (1998); Chen y Yin (1996); Chang (1989); Fu et al. (1995). En ninguno de los trabajos citados se ha utilizado el enfoque geoestadístico. Sin embargo, para el caso de sismicidad natural, se han hecho análisis de peligro sísmico utilizando técnicas geoestadísticas: Glass (1978); Carr (1983); Carr y Glass (1984); Carr y Bailey (1985); Carr (1996). También, apoyados en técnicas geoestadísticas, De Rubeis et al. (1993) realizaron estudios sismotectónicos de la parte central de Italia. Con respecto a los estudios relacionados con el caso de la C.H. Aguamilpa, Nay., éstos se han orientado a el monitoreo, registro y posible correlación, de manera cualitativa, de la sismicidad observada con la historia del llenado del embalse y con las condiciones geológico-estructurales locales y regionales (Delgado et al, 1998).

1.3 Objetivos

Definir un modelo geoestadístico que permita:

- a) Estimar valores de magnitud de sismos para localidades no conocidas en función de la historia sísmica del área (el número de sismos contra el tiempo), de los valores conocidos de magnitud sísmica y de la variabilidad espacial en distintas direcciones de estos mismos datos.
- b) Estimar zonas de probabilidad de excedencia de valores extremos de magnitud de sismos para intervalos de tiempo preestablecidos.
- c) En función de los resultados de los incisos anteriores, proponer una solución teórica al problema de la distribución espacial en el tiempo de la magnitud sísmica.

Los objetivos propuestos obedecen a dos razones. Una es que las técnicas geoestadísticas son utilizadas, con buenos resultados, en problemas de toma de decisiones basados en un entorno de riesgo (Cromer, 1996). Sin embargo, en el caso del riesgo sísmico este enfoque no es una práctica común no obstante que ha dado resultados importantes (Carr, 1996). La segunda razón se refiere a que existe una gran cantidad de información geotécnica, generada durante el desarrollo de proyectos de construcción de obras civiles, que puede estar relacionada con el fenómeno sísmico y de la cual, con técnicas como las geoestadísticas, se pudieran obtener resultados con valor agregado de problemas ya analizados, y consecuentemente utilizar esta nueva información en problemas por resolver.



2 Hipótesis

Para realizar un análisis geoestadístico de un fenómeno, la(s) variable(s) representativa(s) del mismo deberá(n) comportarse como variable(s) regionalizada(s). Los conceptos básicos de geoestadística sobre variables regionalizadas son expuestos en éste capítulo y utilizados para establecer la hipótesis de ésta investigación.

2.1 Marco teórico

La ocurrencia de temblores, descrita por parámetros que los ubican en espacio y tiempo así como su magnitud asociada, es un fenómeno con características fuertemente aleatorias, por lo que en la estimación de dichos parámetros se pueden esperar incertidumbres o probabilidades asociadas.

Desde hace dos décadas, el enfoque probabilístico en el análisis del fenómeno sísmico y los riesgos asociados ha ganado terreno (Semih y Gulkan, 1994), sin embargo, sólo hasta últimas fechas se ha considerado, en algunos estudios, a los parámetros de la sismicidad como variables regionalizadas y por lo tanto, factibles para el enfoque geoestadístico. Lo anterior implica que además de considerar la propiedad aleatoria se toma en cuenta el aspecto espacial del fenómeno. Las características principales de las variables regionalizadas son (Matheron, 1963):

- a) Una variable regionalizada es una variable aleatoria que está espacialmente referida y sus variaciones ocurren en un espacio denominado campo geométrico de la variable.
- b) La variable puede mostrar mayor o menor continuidad espacial. Se pueden tener casos en donde muestras adyacentes de la variable tengan valores muy similares, que correspondería a una continuidad absoluta, o bien, casos donde haya una fuerte variación en los valores de muestras adyacentes y variaciones más suaves entre muestras no cercanas entre sí, lo que correspondería a una falta de continuidad (efecto pepita) y a una continuidad promedio respectivamente.
- c) La variable puede ser anisotrópica, es decir, pueden existir direcciones hacia las cuales la variable muestre mayor continuidad, y algunas en las cuales el fenómeno varíe más rápidamente con la distancia. Mas aún, puede tener comportamiento isotrópico donde la variable cambia de manera similar en todas direcciones.

Cabe añadir que con base en la información disponible y de acuerdo a los conceptos existentes de sismicidad inducida, se considera que:

- 1) La historia sísmica del área (número de sismos detectados en el tiempo) y
- 2) Las zonas de fallas y fracturas preexistentes,

constituyen información relevante en el entendimiento de la magnitud sísmica y su distribución espacial.

2.2 Hipótesis

Se plantea que la magnitud sísmica puede ser modelada como una variable regionalizada y por lo tanto, usar métodos geoestadísticos para inferir la distribución espacial de la magnitud, y su probabilidad de ocurrencia. Además, se espera que esta variable y las variables conocidas (datos de la C.H. Aguamilpa, Nay.) que se deriven de la consideración hecha en el inciso 2.1 (información relevante para el análisis geoestadístico), estén correlacionadas entre sí, de tal forma que puedan ser sujetas a este análisis y permita establecer un modelo de la distribución espacial, para periodos de tiempo preestablecidos, de la sismicidad observada y de la incertidumbre asociada a esta estimación.

Antes de iniciar la descripción de la metodología y el análisis de información disponible, es conveniente mencionar que además de la información mencionada en el apartado 2.1, se contó con el catálogo de sismos localizados en la región de interés elaborado a partir de julio de 1993. El paquete de programas utilizado para el análisis geoestadístico fue el GSLIB (Deutsch y Journel, 1992) complementado con programas de cómputo personales.

3 Metodología

Debido a la gran versatilidad de las técnicas geoestadísticas para caracterizar fenómenos en espacio, tiempo o espacio-tiempo, su uso se ha extendido a diversas ramas de ciencias de la Tierra. En problemas de Geotecnia e Ingeniería Sísmica aun no es una práctica extendida, no obstante su gran potencial para analizar situaciones de riesgo y apoyar en el proceso de toma de decisiones (Cromer, 1996).

Uno de los objetivos principales en el análisis e interpretación geoestadística, en el contexto de riesgo, es el de proveer una predicción de la posible o probable distribución espacial de la propiedad bajo estudio. Hay dos formas básicas de predicción geoestadística: la estimación y la simulación. En la estimación se pretende tener un estimado, estadísticamente eficiente, de la ocurrencia espacial de la propiedad estudiada, con base en los datos y su correlación espacial, capturada a través de un modelo de covariancia o variograma. Lo anterior define las técnicas comúnmente conocidas como *kriging*. En la simulación se pueden generar muchas imágenes, igualmente probables, de la distribución espacial de la propiedad estudiada con el mismo modelo de correlación espacial usado en *kriging*. Sin embargo, la diferencia entre las diferentes imágenes permite tener una medida de la incertidumbre de esta distribución espacial, y así responder preguntas relacionadas con el análisis de riesgo (Journel, 1989; Isaaks y Srivastava, 1989).

La medida de semejanza entre parejas de datos espacialmente distribuidos (covariancia) o de continuidad/discontinuidad (variograma) es capturada a través de modelos. En general, estos modelos pueden funcionar en tres situaciones:

- 1) Dos diferentes atributos P y Q (variables) medidos en la misma localidad x : $[P(x), Q(x)]$
- 2) El mismo atributo Z medido en diferentes localidades, (x) y $(x+h)$: $[Z(x), Z(x+h)]$
- 3) Dos diferentes atributos medidos en diferentes localidades: $[P(x), Q(x+h)]$

De lo anterior se derivan diferentes tipos de estimadores dependiendo si el caso es univariado o multivariado, y si la media de la(s) variable(s) considerada(s) es o no conocida y si esta última es constante o variable. La forma de los estimadores es (Isaaks y Srivastava, 1989):

$$Z^*(x_0) = m + \sum_{\alpha=1}^n \lambda_{\alpha} (Z(x_{\alpha}) - m), \quad \text{kriging simple}$$

$$Z_k^*(x_0) = m_k + \sum_{k=1}^K \sum_{\alpha=1}^{n_k} \lambda_{k\alpha} (Z_k(x_{\alpha}) - m_k), \quad \text{cokriging simple}$$

$$Z^*(x_0) = \sum_{\alpha=1}^n \lambda_{\alpha} Z(x_{\alpha}), \quad \text{kriging ordinario}$$

$$Z_{k^*}(x_0) = \sum_{k=1}^K \sum_{\alpha=1}^{n_k} \lambda_{k\alpha} Z_k(x_{\alpha}), \quad \text{cokriging ordinario}$$

donde $Z(x_{\alpha})$ son las variables conocidas (datos) en sus respectivas posiciones (x_{α}) , (x_0) es la localidad a estimar, λ son los pesos del estimador respectivo, n y K son el número de datos y variables consideradas respectivamente, m es el valor esperado de $Z(x)$ y m_k es el valor esperado de cada una de las K variables. Todos estos estimadores funcionan bajo criterios de insesgamiento y variancia mínima, es decir, la distribución del error $[Z(x)-Z^*(x)]$ tiene media cero y variancia mínima. Asimismo, se considera estacionaridad del fenómeno. Ahora bien, cuando se quiere una medida de la incertidumbre de la estimación en una localidad no muestreada, se busca conocer la probabilidad de ocurrencia de esa estimación, es decir, su función de densidad (Journel, 1989). Para lograr lo anterior, se hace uso de la función indicador:

$$I(x; z_c) = \begin{cases} 1, & \text{si } Z(x) \leq z_c \\ 0, & \text{si } Z(x) > z_c \end{cases},$$

Esta función indicador, al ser evaluada, toma el valor de 1 en las localidades de un mapa donde se cumple la condición respecto del umbral z_c . Entonces, el valor esperado de esta función representará una proporción equivalente a la probabilidad acumulada, es decir:

$$E[I(x; z_c)] = \text{Prob} \{Z(x) \leq z_c\} = p.$$

El problema de estimar la probabilidad de ocurrencia de sismos en cierta área geográfica está asociado a la capacidad de inferir estadística de los sismos tanto en el tiempo como en el espacio. De esta manera, una componente clave en la estimación de tal probabilidad es el acceso a información histórica, así como la capacidad para modelar tal información desde el punto de vista estadístico.

Por lo tanto, con base en los conceptos de Geoestadística recién mencionados donde se considera tanto el enfoque de la estimación como el de la simulación, la estrategia a seguir en el análisis propuesto consiste en tres etapas:

- 1) **Análisis estructural**, donde se prueba la hipótesis planteada a partir de la exploración de la variabilidad espacial de la magnitud sísmica. También en esta etapa se analiza la información disponible complementaria, que es considerada relevante para la investigación.
- 2) **Estimación**, donde se utilizan los resultados de la etapa 1 para estimar valores de magnitud sísmica con el objeto de observar tendencias en la distribución espacial de esta variable. Además se obtiene un modelo de incertidumbre de esta estimación para dar una primera respuesta al problema planteado.
- 3) **Simulación**, donde se utiliza también los resultados de la primera etapa para generar muchas imágenes o realizaciones de la distribución espacial de la magnitud sísmica. La diferencia entre estas imágenes proporciona un modelo de incertidumbre en la estimación de la variable estudiada. El resultado de esta etapa, además de dar una respuesta alterna al problema planteado, es la base para atacar el problema pendiente de la distribución espacial de la magnitud sísmica con el tiempo.

4 Análisis geoestadístico

Como fue mencionado en el capítulo anterior, el análisis geoestadístico propuesto se realiza en tres fases que responden a distintos cuestionamientos. Estas fases son mostradas a continuación.

4.1 Análisis estructural

El término análisis estructural se refiere al establecimiento de modelos de variogramas a partir de los variogramas experimentales de un determinado grupo de datos. Previo a esta evaluación de la variabilidad espacial de la información, se requiere el análisis univariado para reconocer grupos de datos representativos. También, se busca el aseguramiento de la calidad de los datos.

4.1.1 Exploración de datos

La información utilizada ha sido obtenida de registros sísmicos de la zona del embalse de la presa Aguamilpa, ubicada en el Estado de Nayarit. Estos sismogramas son generados por la red de microsismicidad que se mantiene en operación, desde junio de 1993, en la zona de interés. Los datos se refieren a parámetros hipocentrales de sismos, consistentes en la localización epicentral (longitud O y latitud N), la magnitud, la profundidad, la fecha de ocurrencia y otros datos asociados a la calidad de la localización epicentral.

Esta información ha sido usada para crear una base de datos constituida principalmente (82.8 %) por sismos de 1993 y 1994, y en menor porcentaje (17.2%) por eventos ocurridos en el período de 1995 a 1998. Esta diferencia en porcentaje es debida a que el trabajo de lectura de sismogramas y localización de eventos sísmicos, se realizó principalmente en los años de 1993 y 1994, mientras que de 1995 a 1998 se trabajó solo en localización de sismos fuertes. Desde 1987, se mantuvo una estación sísmica permanente (*El Casco*), que sirvió como índice de sismicidad.

La figura 4.1 muestra la localización epicentral y magnitud (agrupada en diferentes rangos) de sismos, mientras que la tabla 4.1 contiene un ejemplo de la información recopilada, en este caso, información de mayo de 1994. También en la figura 4.1 se muestran las principales fallas y estructuras geológicas presentes alrededor del embalse de la presa, así como la ubicación de las estaciones sismológicas de la red de microsismicidad arriba mencionada (Delgado et al, 1998).

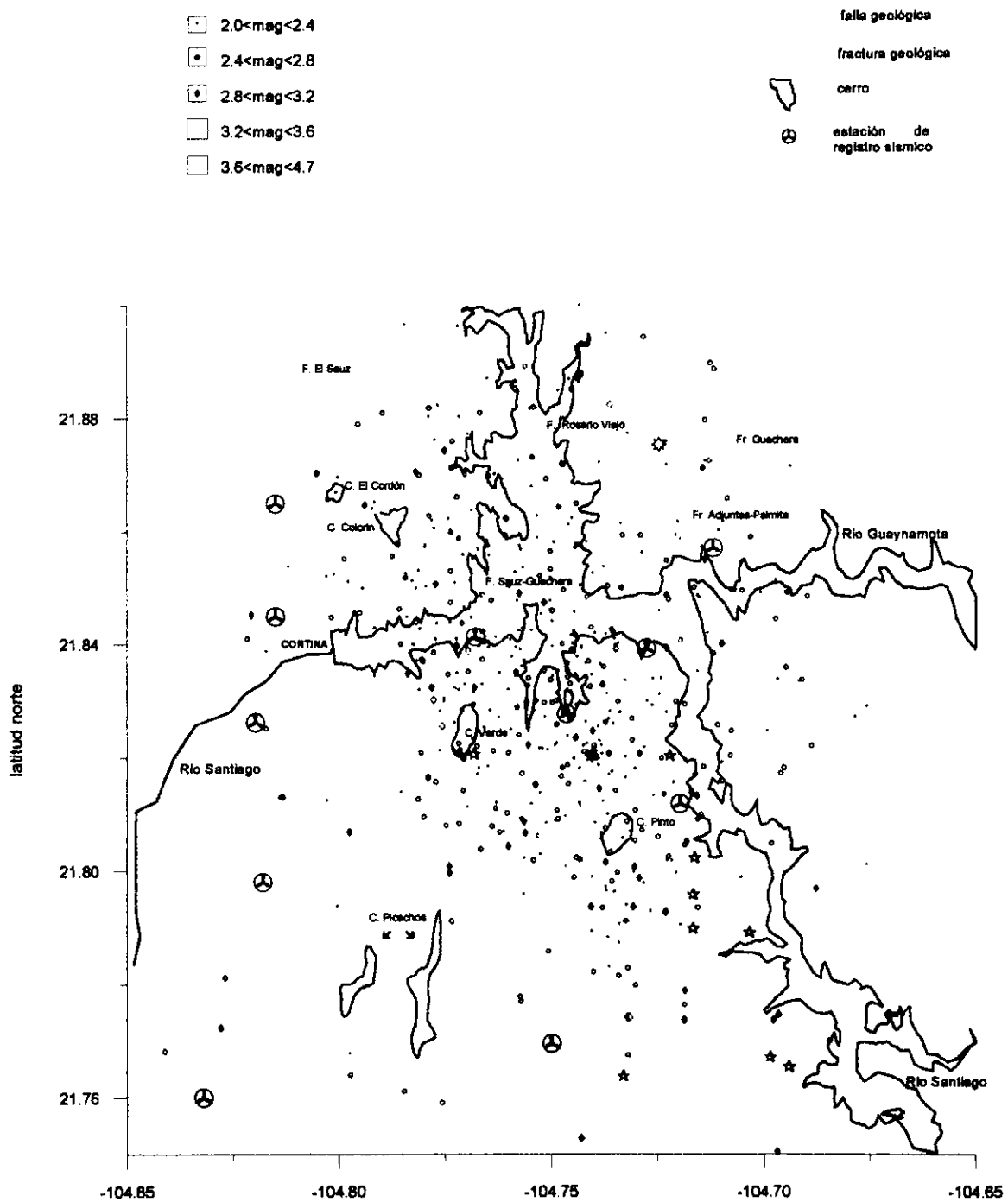


Figura 4.1. Sismicidad observada durante 1994. Se muestran las fallas y las zonas de fractura principales en la zona del embalse de la presa Aguamilpa, Nayarit, así como las estaciones de la red de microsismicidad de la CFE.

Tabla 4.1. Ejemplo de los datos disponibles de parámetros hipocentrales.

Fecha	T. Origen	Lat. N	Long. O	Prof.	Mag.	N	GAP	D.Min	RMS
940501	253 51.66	21-51.18	104-47.22	3.97	1.91	6	139	3.8	0.11
940501	1233 2.17	21-51.56	104-41.91	7.48	2.27	5	185	2.4	0.16
940501	16 6 49.27	21-51.52	104-41.31	7.47	2.42	5	188	3.4	0.12
940503	1143 47.44	21-49.40	104-42.89	9.00	2.68	7	111	3.9	0.16
940503	17 8 56.28	21-52.51	104-44.87	6.25	2.50	6	166	3.4	0.14
940503	22 5 25.28	21-49.49	104-43.90	9.46	2.49	8	106	3.8	0.12
940505	1819 8.51	21-48.80	104-44.99	4.61	2.57	7	125	5.7	0.63
940515	621 46.49	21-48.52	104-43.73	9.37	2.22	9	110	5.5	0.16
940515	728 18.53	21-48.90	104-45.29	8.75	2.69	8	102	5.9	0.15
940515	729 19.86	21-49.27	104-44.69	8.08	2.89	9	104	4.7	0.13
940516	847 22.51	21-55.32	104-46.14	6.34	2.90	9	220	8.6	1.00
940518	037 8.44	21-51.74	104-44.17	6.01	2.64	7	191	1.6	0.51
940528	922 38.83	21-51.04	104-43.78	11.18	2.64	9	131	1.2	0.42

Los parámetros Fecha y T. Origen ubican en tiempo a un determinado sismo, por lo que un valor de 940501 y de 123302.17 respectivamente, se refiere a un evento que se originó el primero de mayo de 1994 a las 12 horas, 33 minutos y 02.17 segundos. Los parámetros Lat. N, Long. O y Prof. ubican geográficamente al temblor considerado, siendo los dos primeros las coordenadas geográficas (Latitud y Longitud) en grados, minutos y fracciones de minutos del epicentro respectivo, mientras que el parámetro Prof. es el valor en kilómetros de la profundidad del sismo, medida a partir de cero metros sobre el nivel del mar bajo el epicentro. El parámetro Mag. (en grados) se refiere a la magnitud del sismo, que es una medida de la energía liberada por éste, calculada de acuerdo a la duración de la señal sísmica, es decir, el tiempo transcurrido entre el arribo de la fase P y hasta el punto en que la señal se confunde con el nivel de ruido del sitio de la estación de registro. El parámetro N es el número de fases leídas para la localización del temblor, en donde las fases de onda sísmica P y S, para la misma estación, se consideran como dos lecturas. El parámetro GAP es el ángulo mayor formado entre dos estaciones de la red sísmica

respectiva y un tercer punto, el epicentro, que viene a ser el vértice de dicho ángulo. El parámetro D_{min} representa la distancia, medida en kilómetros, del epicentro a la estación de registro más cercana a este último. El parámetro RMS es el error cuadrático medio de los residuales de tiempo en segundos:

$$RMS = \left[\frac{1}{n} \sum_{i=1}^n R_i^2 \right]^{1/2},$$

donde R_i es el residual de tiempo para la i ésima estación y n es el número de estaciones.

Para evaluar la calidad de la localización epicentral, se consideraron como prioritarios los parámetros RMS y N. El valor de RMS es estimado utilizando residuales entre tiempo observado y tiempo calculado para cada una de las fases leídas en cada estación, por lo que este valor refleja de manera cualitativa la calidad de las lecturas de los tiempos de las fases P y S, y lo acertado del modelo de espesores y velocidades de corteza utilizado en la estimación del tiempo de viaje de una determinada fase. Por otro lado, para obtener una localización epicentral correcta, se requiere de lecturas en tres o más estaciones, por lo que un valor de seis fases leídas sería un mínimo aceptable para el parámetro N. Cabe mencionar que un valor de RMS de 0.3 o menor, implica localizaciones epicentrales precisas, ya que fuentes habituales de error como la precisión de la regla para leer registros sísmicos analógicos, o la precisión en las marcas de tiempo del aparato de registro, son del orden de 0.1 s. Consecuentemente, la información original fue filtrada con base en estos dos parámetros y los valores antes descritos.

Esta información ha sido usada para efectuar un análisis geoestadístico que requiere de la exploración y modelado de las posibles relaciones en espacio y tiempo de la magnitud sísmica. Por ello, en primera instancia se efectúa un análisis estadístico univariado, con y sin consideración de diferentes ventanas de tiempo para esta variable.

La base de datos antes descrita, se complementa con un registro del número total de sismos que han ocurrido a partir del inicio del llenado del embalse, 25 de junio de 1993. La gráfica de este registro contra el número consecutivo de días (día juliano) es mostrada en la figura 4.2. Estos sismos han sido registrados en la estación *El Casco* (estación testigo), que tiene una sensibilidad para detectar temblores de magnitudes muy pequeñas, como es observado en el histograma de la figura 4.3, registrándose temblores menores a 0.5 grados de magnitud. Esta gráfica es importante ya que permite observar que no obstante el fuerte cambio en cuanto al número de sismos, éstos no desaparecen en su totalidad y tienden a estabilizarse en un número reducido. Asimismo, es evidente que no hay un valor promedio constante de la sismicidad en el tiempo, sin embargo, se puede observar periodicidad en cuanto a los cambios fuertes en el número de sismos. La observación anterior apoya, como se verá más adelante, las conclusiones en el análisis con ventanas de tiempo de la información sísmica localizada.

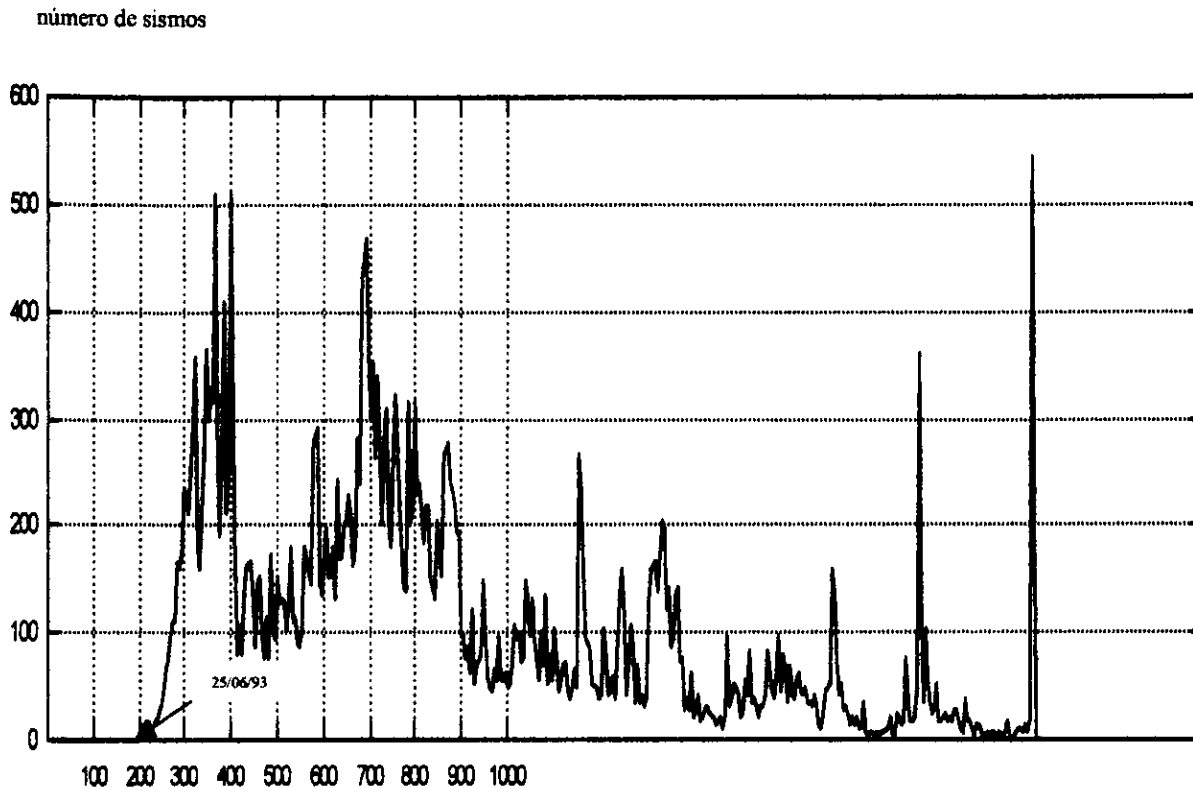


Figura 4.2. Evolución histórica del número de sismos en la estación *El Casco* (permanente) ubicada en la zona del embalse de la presa Aguamilpa, Nayarit. El día 200 corresponde a la fecha del 25 de junio de 1993. Este registro incluye eventos con magnitudes hasta de 0.3 grados.

4.1.1.1 Análisis sin consideraciones temporales

El propósito es observar las características principales del histograma y la estadística básica de la magnitud de sismos, abarcando el periodo de 1993 a 1998, para posteriormente considerar ventanas de tiempo por años, semestres, trimestres y meses, y así establecer comparaciones entre las diferentes poblaciones y encontrar el tamaño de ventana mínimo representativo del histograma general. Esta estrategia permite identificar grupos de datos con los cuales es posible explorar los cambios, en espacio y tiempo, observados en la sismicidad de la zona de estudio. La figura 4.3 muestra el histograma y la estadística básica de la totalidad de la información de magnitud sísmica contenida en la base de datos.

Número de sismos

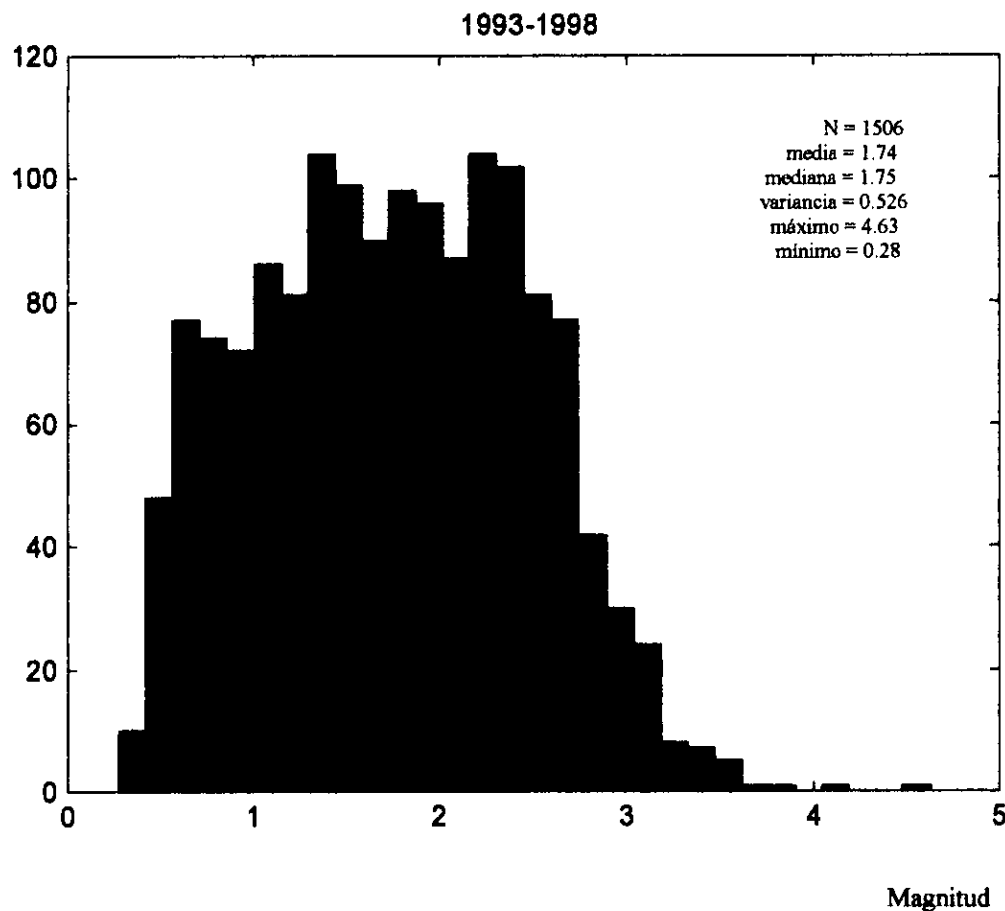


Figura 4.3. Histograma y estadística básica de los datos de magnitud sísmica durante el período de 1993 a 1998 en la zona del embalse de la presa Aguamilpa, Nayarit.

Cabe recordar que esta información corresponde a sismos localizados, y que solo representa una muestra de la sismicidad presente en la zona. Es notable que el valor de mayor frecuencia está en 1.35 (± 0.05), ligeramente menor que la media y la mediana, acusando una distribución con mayor número de sismos hacia los valores bajos, lo cual es esperado en un proceso de sismicidad establecido, en donde existe un mayor número de sismos pequeños con respecto a sismos de magnitud mayor. Sin embargo, el corte observado alrededor del valor de 2.2, corresponde al valor mínimo de detección de la red sísmica respectiva, ya que la continuidad en cuanto al incremento del número de sismos para magnitudes menores, se pierde en este valor. También, es notable la carencia de información para los valores de 1.6 y 2.0, lo cual acusa que la muestra está incompleta para valores menores de 2.0.

4.1.1.2 Análisis con ventanas de tiempo

El análisis con ventanas de tiempo se realizó con información de sismos considerando el rango completo de magnitud de eventos localizados. La figura 4.4 (a-c) muestra los histogramas para ventanas de tiempo de un año, seis meses, tres meses y un mes respectivamente. Sin embargo, es importante recordar que para ventanas de tiempo dentro de los años de 1995 y 1998 es inapropiado elaborar histogramas debido a la carencia de información.

En la figura 4.4a, es notable como el histograma de 1994 tiene la forma del histograma general y acusan la carencia de información alrededor del valor 2.0. El histograma de 1993 también carece de información cercana a 2.0, y no obstante que es menor el número de datos que el de 1994, este histograma tiende a reproducir la forma del histograma general. Los histogramas de 1996 y 1997 presentan una distribución no esperada y es notable la ausencia de información.

Número de sismos

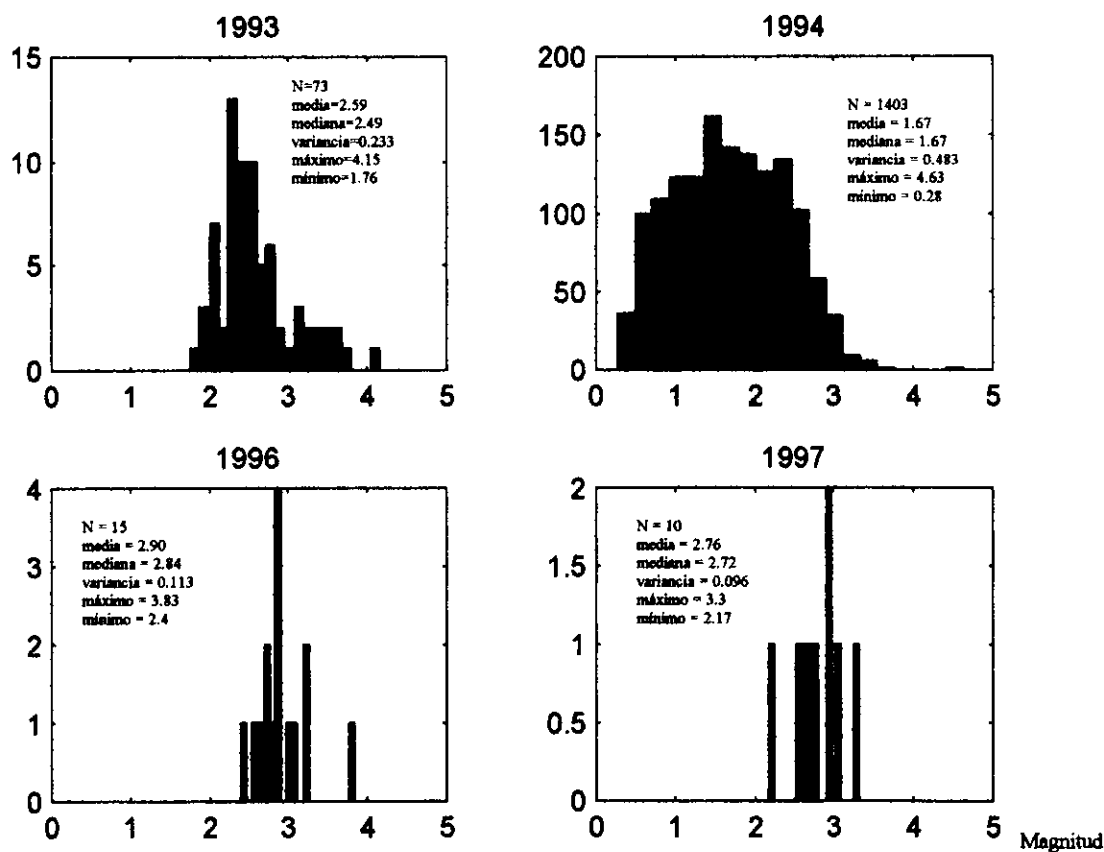


Figura 4.4a. Histogramas y estadística básica de los años 1993, 1994, 1996 y 1997. Para 1995 y 1998 no fue posible elaborar histogramas.

En los histogramas semestrales y trimestrales de 1994, figura 4.4b, se observa un comportamiento similar al del histograma general, pero también es notable un cambio en el valor del nivel mínimo de detección entre los distintos histogramas. Sin embargo, en el caso del trimestre octubre-diciembre de 1994, la forma del histograma no está definida para valores menores a 2.0 y el número de datos es menor al de los otros trimestres, debido a la disminución de la sismicidad en este período (figura 4.2, período entre los días 700 y 800), sin embargo, al comparar con los otros trimestres, es de esperarse que la muestra de este trimestre sea incompleta.

Número de
sismos

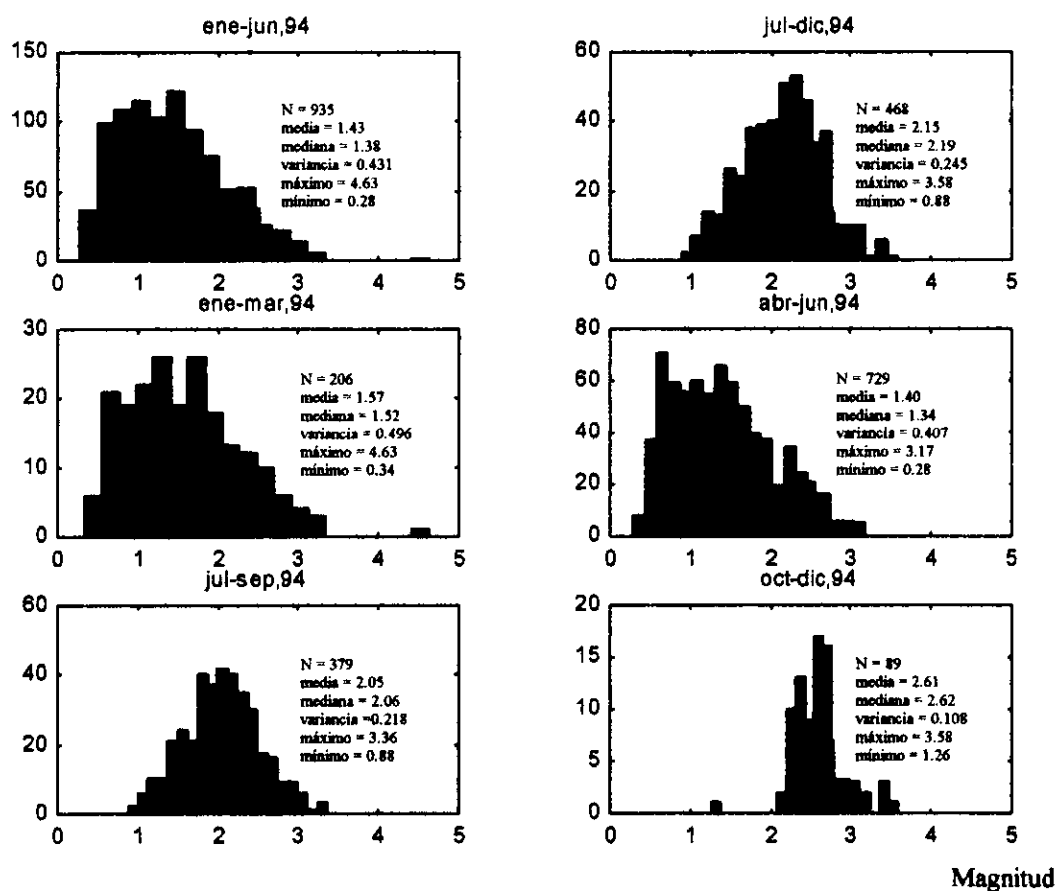
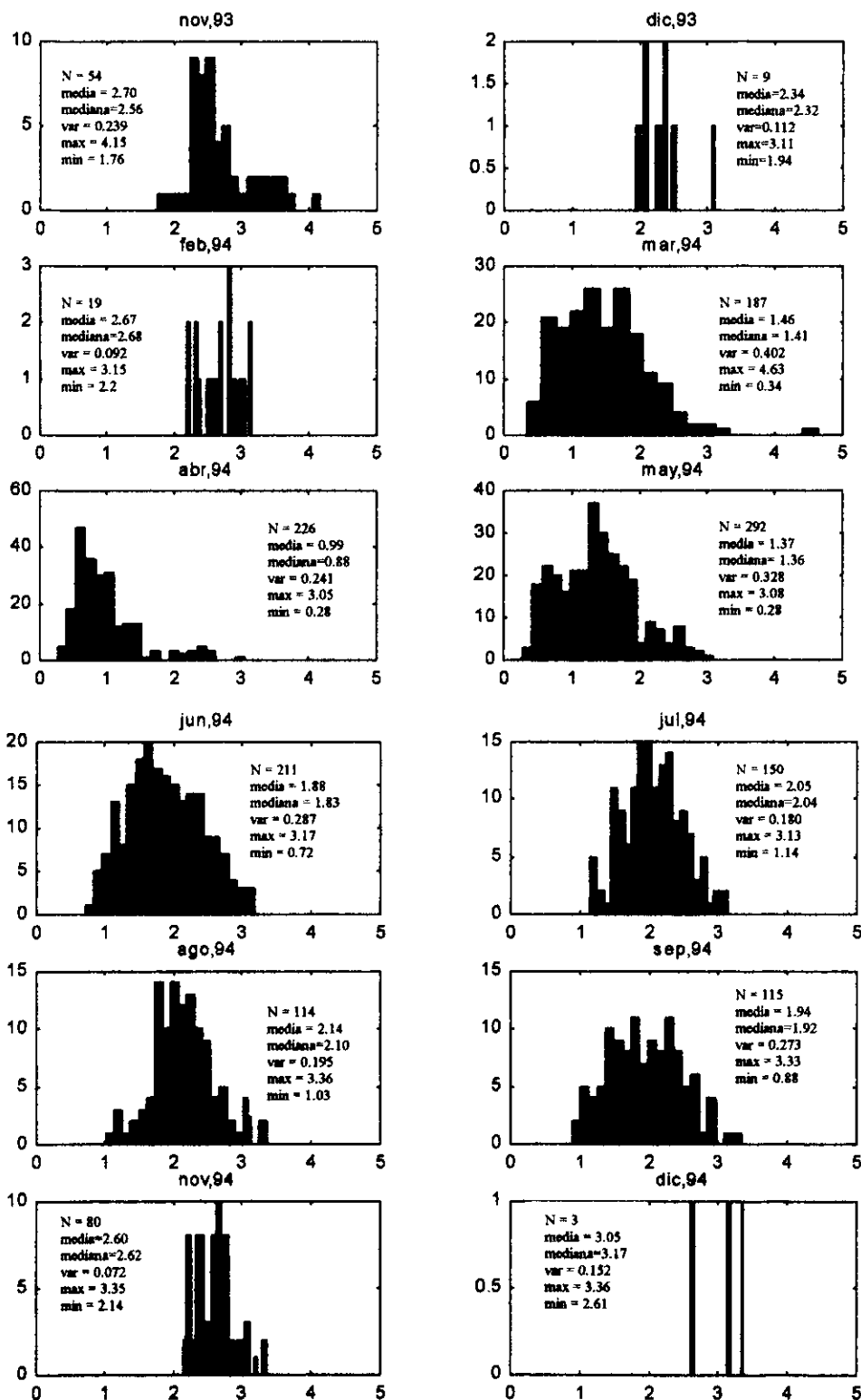


Figura 4.4b. Histogramas y estadística básica de los semestres y trimestres de 1994.

En el caso de las ventanas de tiempo mensuales, figura 4.4c, seis de catorce histogramas tienen escasa o nula información y no tienen una forma similar a los histogramas trimestrales. Es decir, diciembre de 1993, febrero y diciembre de 1994, carecen de información. Situación también observada para los periodos octubre de 1993, enero y octubre de 1994. Por lo anterior, se puede decir que las ventanas mensuales de información disponible no son representativas del fenómeno analizado, mientras que las ventanas trimestrales o mayores reproducen la forma del histograma general.



Magnitud

Figura 4.4c. Histogramas y estadística básica para los dos últimos meses de 1993 y para cada mes de 1994, excepto enero y octubre por no tener información disponible.

4.1.1.3 Análisis comparativo entre ventanas de tiempo

Del histograma de la totalidad de la base de datos (1993-1998), se ha identificado que el valor del nivel mínimo de detección se ubica en 2.2 ± 0.05 grados de magnitud, pero también se observa, considerando ventanas de tiempo trimestrales, una variación de 1.4 a 2.6 en este valor. Esto es explicado por un criterio de selección de sismos, para su localización, no constante durante el período de análisis (Lomas, 1995).

Por la razón antes expuesta, se decidió acotar la información con un límite inferior de 2.0, ya que este valor y la forma del histograma general se mantienen en las ventanas de tiempo de los años de 1993, 1994, los semestres de 1994 y aún en los tres primeros trimestres de 1994. Para cada uno de los años de 1995 a 1998, los histogramas correspondientes no son, en buena medida, similares a la forma establecida por el histograma general, debido a la ausencia de información en estos periodos.

Por otro lado, con base en el registro del número total de sismos detectados (no necesariamente localizados) en el tiempo (estación índice), se obtiene la gráfica del número acumulado de sismos contra el tiempo. A partir de esta gráfica, mostrada en la figura 4.5, se puede establecer que la continuidad de la sismicidad, así como los cambios en la pendiente de esta curva, definen periodos con tasa de sismicidad dominante. Estos periodos o pendientes comienzan a partir del día juliano 230, 25 de julio de 1993.

Las diferentes tasas de sismicidad analizadas estuvieron comprendidas entre los días julianos 200 y 700. Los periodos establecidos en días julianos, con base en los cambios de pendiente de la gráfica de la figura 4.5, y su correlación con respecto a la información trimestral, se incluyen en la tabla 4.2.

Tabla 4.2. Periodos, en días julianos, con base en la pendiente de la curva de sismicidad acumulada y su relación con información trimestral localizada.

Período (día juliano)	Trimestre de información epicentral
200-300	-
300-400	octubre-diciembre/1993
400-575	enero-marzo/1994 y abril-junio/1994
575-675	julio-septiembre/1994

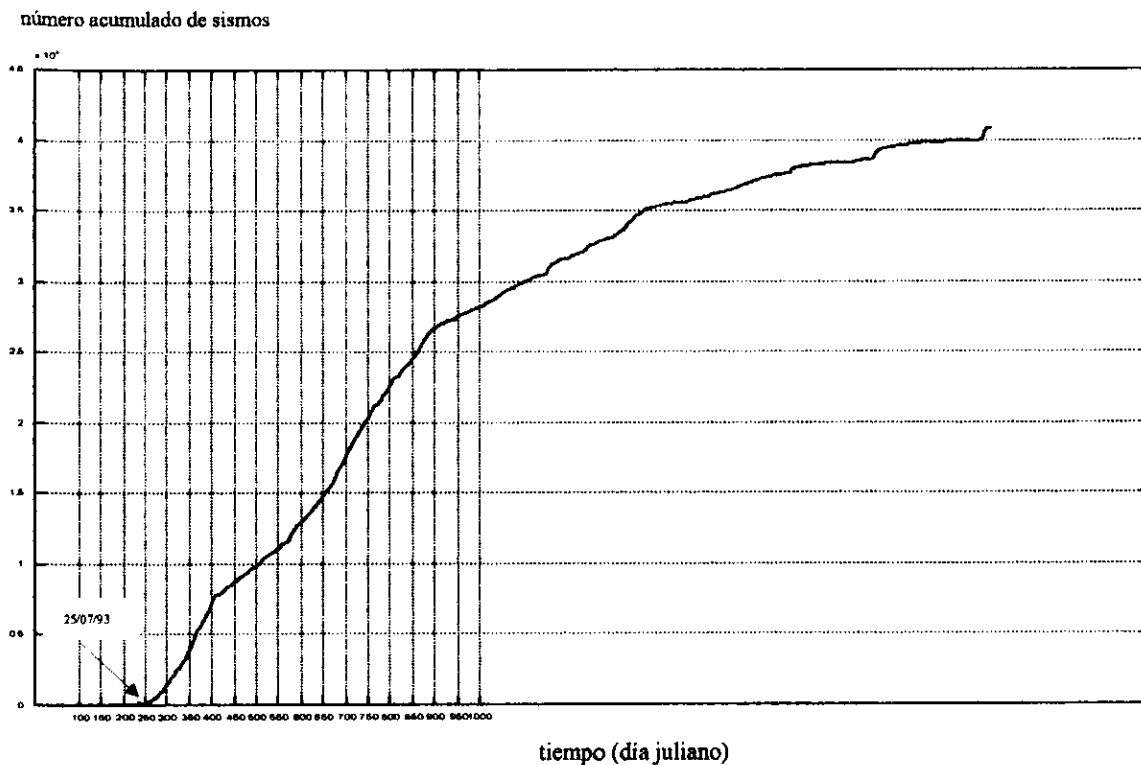


Figura 4.5. Evolución histórica del número acumulado de sismos en la zona del embalse de la presa Aguamilpa, Nayarit. Obsérvese que en el día 230, 25 de julio de 1993, se presenta el primer evento sísmico en la zona, es decir, 30 días después del comienzo del llenado del embalse.

Valor de "b"

Para determinar el tamaño mínimo de ventana de tiempo, representativa del histograma general de sismos localizados, se ha considerado el análisis del valor de b (Gutenberg y Richter, 1954). Esto es, la pendiente de la recta ajustada en una gráfica del logaritmo base 10 del número acumulado de sismos, por arriba de un determinado umbral de magnitud, contra este mismo valor de magnitud, para un conjunto de datos representativo de la sismicidad en una zona específica y en un período de tiempo predeterminado. La figura 4.6 muestra los puntos, recta ajustada y valores de la pendiente del ajuste respectivo, para los diferentes trimestres de información localizada (epicentros) antes mencionados. Se consideraron diferentes valores de nivel mínimo de detección en el cálculo del valor de b para los diferentes trimestres. La forma de la relación a la cual se ajustan las observaciones (Gutenberg y Richter, 1954) es:

$$\log N = a - b(M),$$

donde N es el número acumulado de sismos, de magnitud mayor o igual al umbral considerado (valor de M), dentro de un periodo de tiempo fijo, mientras que a y b son los coeficientes de la recta.

Log (número de sismos)

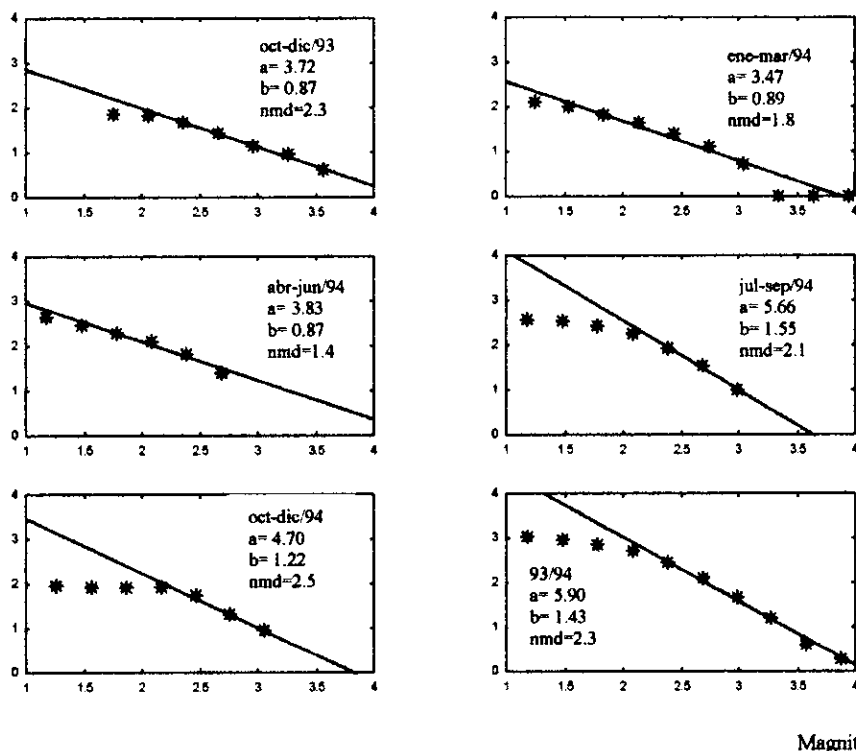


Figura 4.6. Rectas ajustadas de acuerdo a la relación Gutenberg y Richter (1954) para la información trimestral de 1993 y 1994. Estos ajustes se hicieron considerando el número acumulado de sismos para valores de magnitud definidos, entre el valor del nivel mínimo de detección y el valor máximo observado, con incrementos cada 0.3 grados y para cada histograma de las ventanas de tiempo.

El valor del coeficiente b (pendiente de la recta) representa la frecuencia de la magnitud de la sismicidad. De manera cualitativa, al comparar dos valores de este parámetro, calculados en diferentes periodos de un catálogo de sismicidad de una determinada región, indica que para el valor más alto, se requerirá un número mayor de sismos de magnitudes pequeñas para tener un sismo de magnitud mayor, y para el valor de b relativamente más bajo, se requerirá menor número de sismos de magnitudes pequeñas para tener uno de magnitud mayor. También de manera cualitativa, se puede relacionar con el estado de esfuerzos o grado de fracturamiento que puede estar presente en una zona determinada.

En la figura 4.6 se puede observar que el cambio en el valor de b del trimestre octubre-diciembre/1993 al de enero-marzo/1994 es mínimo, sin embargo, se tiene un cambio importante en la pendiente de la curva del número acumulado de sismos entre estos

periodos, detectado en el punto del día 400 de la figura 4.5. El cambio en el valor de b de abril-junio/1994 a julio-septiembre/1994 está asociado a la inflexión de la curva mencionada para el día 575. Asimismo, la poca variación de este valor entre enero-marzo/1994 y abril-junio/1994 es congruente con la continuidad de la pendiente entre los días 400 y 575. Lo anterior implica que el valor de b en el trimestre octubre-diciembre/1993 puede estar subvaluado, ya que el número de sismos (localizados) correspondiente a este intervalo (figura 4.4a, período 1993) es mínimo, lo que resulta evidente al observar la gran cantidad de sismos detectados para el intervalo en cuestión (del día 300 al 400, figura 4.2). Por lo tanto, si se incluye la información faltante de magnitudes pequeñas (cercanas al nivel mínimo de detección), el valor de b para este período sería mayor al calculado.

Por consiguiente, es importante enfatizar que:

- ◆ La información de los años de 1995 a 1998 no es representativa del fenómeno a analizar. La forma de los histogramas y los valores de la tasa de magnitud de sismicidad (valor de b) no son los esperados de acuerdo a la figura del número acumulado de sismos contra el tiempo para la zona de interés. Lo anterior es debido a la carencia de información en estos periodos.
- ◆ De manera natural, se espera que el número acumulado de sismos de magnitudes pequeñas sea mayor al número acumulado para magnitudes mayores. Sin embargo, es notable que para valores menores o iguales a 2.2 (60.05), el número acumulado de sismos en el histograma general empieza a decrecer. Lo anterior acusa el nivel mínimo de detección de la red sísmica correspondiente. Asimismo, dadas las características de la información disponible, se acotó la base de datos con un límite inferior de 2.0 grados de magnitud. Por lo tanto, valores abajo de este nivel mínimo no fueron considerados en el análisis geoestadístico.
- ◆ Es plausible considerar ventanas de tiempo de seis meses, e inclusive hasta de tres meses, con la información de los años de 1993 y 1994, para analizar posibles cambios en la ocurrencia espacial de la sismicidad en la zona del embalse de la presa Aguamilpa, Nayarit. Por otro lado, la escasez de datos en los años de 1995 a 1998 es una restricción importante para el análisis espacio-temporal en este período.
- ◆ La información de 1993 y 1994 constituye la parte principal de la base de datos definitiva.

4.1.2 Variogramas

Para modelar las relaciones espaciales de la magnitud de sismos, es decir, establecer las direcciones de mayor o menor continuidad y/o variabilidad, incluyendo la distancia máxima hasta la cual los datos son correlacionables, se calcularon variogramas experimentales para seis diferentes direcciones azimutales (0° , 30° , 60° , 90° , 120° , 150°). La información considerada correspondió a los epicentros con sus magnitudes asociadas del año de 1993 y 1994, tomando en cuenta ventanas de tiempo trimestral y sin ventanas de tiempo.

La expresión usada para estimar el variograma experimental (Journel y Huijbregts, 1978) es:

$$g_{xy} = 1/2N \sum_{i=1}^N [x_i - y_i]^2, \dots\dots (1)$$

donde x_i e y_i representan realizaciones de variables aleatorias X e Y. Como se aprecia, el variograma es un promedio de las diferencias al cuadrado de parejas de datos.

Asimismo, al incluir la estadística bivariada, las dos variables aleatorias de la pareja, pueden representar:

- a) Dos diferentes atributos medidos en la misma localidad.
- b) El mismo atributo medido en dos diferentes localidades.
- c) Dos diferentes atributos medidos en dos diferentes localidades.

En el caso b, se puede definir un vector \mathbf{h} que capture la dirección y distancia entre las dos localidades, de tal forma que las variables aleatorias pueden ser expresadas como:

$$X = Z(x) \ ; \ Y = Z(x + \mathbf{h}).$$

Si se agrupan todas las parejas de datos $n(\mathbf{h})$ separadas a distancias y direcciones similares (vector \mathbf{h} similar), la evaluación del variograma experimental proporcionará las características promedio de variabilidad espacial de la variable Z. La expresión del variograma experimental, que ahora será una función vectorial, tendrá la forma:

$$g_A(\mathbf{h}) = 1/n(\mathbf{h}) \sum_{a=1}^{n(\mathbf{h})} [Z(x_a) - Z(x_a + \mathbf{h})]^2, \dots\dots (2)$$

donde A es el dominio espacial de la variable Z(x).

Si se analiza el límite del variograma experimental cuando h tiende a cero, se tiene que:

$$\lim_{h \rightarrow 0} g_A(h) = \lim_{h \rightarrow 0} \frac{1}{n(h)} \sum_{a=1}^{n(h)} [Z(x_a) - Z(x_a + h)]^2 = 0, \dots \quad (3)$$

lo cual implica que se está evaluando con parejas de datos integradas por un mismo dato repetido. Sin embargo, lo interesante es observar que el variograma tendrá valores pequeños o que tienden a cero para un h pequeño, cuando el fenómeno analizado sea muy continuo o las muestras analizadas sean semejantes. Caso contrario es cuando el variograma tiene valores altos cuando h tiende a cero, lo que implica un fenómeno con alta variabilidad espacial o poca continuidad en el origen, es decir, hay una gran variación en las muestras a distancias pequeñas, no obstante que la variable analizada pueda tener mayor estabilidad o continuidad a distancias mayores. Una situación como la antes descrita se le conoce como efecto pepita. El reconocer este efecto en el modelo de la variabilidad espacial de la variable $Z(x)$ es relevante para la estimación de valores de $Z(x)$, a distancias pequeñas de los datos considerados.

Para poder analizar el límite cuando h tiende a infinito es necesario introducir el concepto de covariancia. En una distribución bivariada, como es el caso analizado, se consideran dos variables aleatorias X e Y . Utilizando la propiedad del operador valor esperado y la forma de la variancia se tiene que:

$$E[X + Y] = E[X] + E[Y] = m_X + m_Y$$

$$\begin{aligned} \text{var}[X + Y] &= E[(X + Y - m_X - m_Y)^2] = E[(X - m_X + Y - m_Y)^2] \\ &= E[(X - m_X)^2 + 2(X - m_X)(Y - m_Y) + (Y - m_Y)^2] \\ &= E[(X - m_X)^2] + 2E[(X - m_X)(Y - m_Y)] + E[(Y - m_Y)^2] \end{aligned}$$

\therefore

$$\text{var}[X + Y] = \text{var}[X] + 2E[(X - m_X)(Y - m_Y)] + \text{var}[Y],$$

donde el término $E[(X - m_X)(Y - m_Y)]$ se define como la covariancia de las variables aleatorias X y Y , por lo que,

$$\text{var}[X + Y] = \text{var}[X] + 2\text{cov}[X, Y] + \text{var}[Y],$$

entonces,

$$\text{cov}[X, Y] = E[(X - m_X)(Y - m_Y)] = E[XY - X m_Y - Y m_X + m_X m_Y] = E[XY] - m_X m_Y - m_Y m_X + m_X m_Y$$

∴

$$\text{cov}[X, Y] = E[XY] - m_X m_Y = E[(X - m_X)(Y - m_Y)]. \quad \dots \quad (4)$$

La covariancia representa una medida de la dependencia lineal de dos variables (Journel, 1989). Ahora bien, la expresión (1) puede ser expresada en términos de h y utilizando el operador de valor esperado como.

$$g(h) = \frac{1}{2}E[(Z(x) - Z(x+h))^2], \quad \dots \quad (5)$$

y la expresión (4) como

$$\text{cov}[Z(x), Z(x+h)] = E[Z(x)Z(x+h)] - m_X m_{X+h}, \quad \dots \quad (6)$$

al desarrollar la expresión (5) se tiene

$$\begin{aligned} 2g(h) &= E[Z^2(x) - 2Z(x)Z(x+h) + Z^2(x+h)] \\ &= E[Z^2(x)] - 2E[Z(x)Z(x+h)] + E[Z^2(x+h)], \end{aligned}$$

que al sumar y restar el producto $m_X m_{X+h}$ en el término central de la expresión anterior, se tiene,

$$\begin{aligned} 2g(h) &= E[Z^2(x)] - 2\{E[Z(x)Z(x+h)] - m_X m_{X+h}\} - 2m_X m_{X+h} + E[Z^2(x+h)] \\ &= E[Z^2(x)] - m_X^2 + m_X^2 - 2\text{cov}[Z(x), Z(x+h)] - 2m_X m_{X+h} + E[Z^2(x+h)] - m_{X+h}^2 + m_{X+h}^2 \\ &= \text{var}[Z(x)] + m_X^2 - 2\text{cov}[Z(x), Z(x+h)] - 2m_X m_{X+h} + \text{var}[Z(x+h)] + m_{X+h}^2 \end{aligned}$$

∴

$$g(h) = \frac{1}{2}\{\text{var}[Z(x)] + \text{var}[Z(x+h)] - \text{cov}[Z(x), Z(x+h)] - m_X m_{X+h} + \frac{1}{2}(m_X^2 + m_{X+h}^2)\}, \dots \quad (7)$$

la expresión anterior establece la relación entre variograma y covariancia. Ahora, si se considera que dos variables son independientes:

$$E[XY] = E[X] E[Y],$$

entonces el límite de $g(h)$ cuando h tiende a infinito es,

$$\lim_{h \rightarrow 0} g(h) = \frac{1}{2} \{ \text{var}[Z(x)] + \text{var}[Z(x+h)] - m_x m_{x+h} + \frac{1}{2}(m_x^2 + m_{x+h}^2) \}$$

y si además se asumen condiciones de estacionariedad de primer orden:

$$E[Z(x)] = E[Z(x+h)] \quad \text{y} \quad \text{var}[Z(x)] = \text{var}[Z(x+h)]$$

entonces,

$$\lim_{h \rightarrow 0} g(h) = \text{var}[Z(x)] - m_x^2 + m_x^2 = \text{var}[Z(x)], \quad \dots \dots \quad (8)$$

por lo tanto, el variograma está acotado con los límites:

$$\lim_{h \rightarrow 0} g(h) = 0 \quad \text{y} \quad \lim_{h \rightarrow \infty} g(h) = \text{var}[Z(x)].$$

Al considerar las condiciones de estacionariedad de primer orden en la ecuación (7) se tiene,

$$g(h) = \text{var}[Z(x)] - \text{cov}[Z(x), Z(x+h)],$$

que bajo hipótesis de estacionariedad de segundo orden, donde la covariancia no depende de la posición sino del vector h ,

$$\text{cov}[Z(x), Z(x+h)] = \text{cov}(h) \quad \text{y por lo tanto} \quad g(h) = \text{var}[Z(x)] - \text{cov}(h),$$

pero como la variancia es la autocovariancia de la misma variable, es decir, la covariancia cuando $h = 0$, se tiene que:

$$g(h) = \text{cov}(0) - \text{cov}(h), \quad \dots \dots \quad (9)$$

lo que indica que el variograma y la covariancia es la misma medida espacial de una variable con diferente representación.

4.1.2.1 Estacionariedad de $Z(x)$.

Como ya fue expuesto anteriormente, el variograma de una función aleatoria, concebida ésta como un conjunto de variables aleatorias dependientes, caracteriza el grado de variabilidad espacial entre dos cualesquiera de estas variables $Z(x)$, $Z(x+h)$, cuya distancia y dirección entre ambas está definida por el vector h . Este variograma es modelado después de que el promedio espacial definido por la expresión (2) es evaluado, y dicho modelo representa la estadística promedio sobre el dominio A . La decisión de este modelado es necesariamente subjetiva e implica que el variograma ya no depende de la posición x dentro de A , es decir: $\gamma(h) = \frac{1}{2}E[(Z(x) - Z(x+h))^2]$ es independiente de $x \in A$, y se dice que $\gamma(h)$ es estacionario. Por lo tanto, la estacionariedad es una característica del modelo y no una propiedad intrínseca de la distribución real de los valores de $Z(x)$ sobre A . La decisión de estacionariedad permite considerar un conjunto de datos, sobre los cuales, se calculan promedios experimentales y proporciones que se asumen representativos de la población como un todo (dominio A) y no de una localidad x en particular.

En lo particular, está claro que la evolución temporal de la magnitud sísmica en el caso estudiado es un aspecto relevante en el análisis de esta información. La hipótesis geoestadística supone la conveniencia de verificar, con los mismos datos, la aplicabilidad de un modelo estacionario o de uno no estacionario. Así, la construcción de variogramas permite identificar no solamente la variabilidad espacial, sino también la probable existencia de patrones temporales. De esta manera, al considerar la magnitud sísmica como una variable espacial y temporal, $Z(x,t)$, el semivariograma respectivo puede ser expresado como:

$$\gamma(h,t') = \frac{1}{2}E[(Z(x,t) - Z(x+h,t'))^2],$$

siendo h y t' los desplazamientos espacial y temporal que permiten distinguir y modelar la variabilidad en el espacio y en el tiempo. La modelación del variograma anterior supone la existencia de información, en espacio y tiempo, y la necesidad de promediarla con la finalidad de obtener valores del semivariograma para diferentes valores de h y t . Este proceso es un recurso para considerar la irregularidad espacial y temporal de las muestras disponibles, y así obtener, lo que se denomina un semivariograma experimental. Este análisis puede modelar a $Z(x,t)$ como una variable estacionaria, en tiempo o espacio, o como una variable no estacionaria. Esto es, el variograma experimental y el análisis estadístico univariado permitirán modelar a la variable en cuestión como:

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

o

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

o bien como,

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

además de modelar, también, el variograma como una función del desplazamiento espacial h y temporal t' , es decir, se podrá determinar si es aceptable que $Z(x,t)$ pueda ser modelada como una variable aleatoria de segundo orden.

Como se puede apreciar, elementos como el análisis estadístico de la información disponible, la construcción de semivariogramas, y su análisis, son tareas importantes para definir como modelar $Z(x,t)$. En el caso de la magnitud sísmica observada en la zona del embalse de la presa Aguamilpa, el análisis estadístico univariado y el uso de ventanas de tiempo ha permitido modelar a $Z(x,t)$ como una variable aleatoria de segundo orden en cada una de las k ventanas de tiempo, es decir:

$$E[Z(x,t_k)] = m_k,$$

con t_k siendo un intervalo de tiempo. Esto permite suponer que $E[Z(x,t)]$ no depende de la posición, y de que en cada ventana de tiempo este valor es constante para $\forall x$. Observar que en esta concepción no hay hipótesis acerca de la variación en el tiempo de este valor esperado. El modelo propuesto considera estacionariedad de segundo orden en cada ventana de tiempo para x , pero no extrapola tal supuesto para el conjunto de ventanas de tiempo, es decir, para t . Esta observación es relevante, ya que la aplicación de las técnicas de kriging y de simulación estocástica secuencial, en este contexto, son aplicadas en cada ventana de tiempo k donde las condiciones mencionadas han sido aceptablemente reproducidas por la información disponible. La implicación de esta aseveración es que un semivariograma experimental es obtenido para cada k , como se verá en el apartado siguiente, y es usado para propósitos de estimación y simulación. Esto es, en general se satisface que:

$$g_k(h) \neq g_{k'}(h) \quad \forall k \neq k',$$

indicando k el período de tiempo definido por t_k , y k un índice referido a tal período. Así, esta aproximación al problema modela $Z(x,t)$ como una variable aleatoria estacionaria, de segundo orden, para cada k y no desconoce la posible no estacionariedad en el tiempo.

Desde luego que el aspecto temporal es importante en el contexto de una predicción temporal y no sólo en el de la predicción espacial, considerado aquí. Este aspecto será explorado en el siguiente capítulo, desde el punto de vista teórico, presentando una alternativa al problema de inferir un modelo no estacionario a una variable aleatoria temporal, cuyas realizaciones aún no han ocurrido, esto es, el problema es modelar la no estacionariedad primero, y después seleccionar las técnicas para efectuar la predicción temporal.

4.1.2.2 Variogramas, isotropía y modelos para $Z(x)$.

La posición de la variable explorada está definida por las coordenadas geográficas (longitud O y latitud N) de los epicentros. Estas coordenadas están expresadas en grados y fracción de grado, cuyo máximo rango de variación es de 0.2 de grado. El área considerada va de 104.65° a 104.85° de longitud O y de 21.75° a 21.90° de latitud N.

Es importante anotar que la confiabilidad del variograma experimental, depende de la cantidad de parejas de datos para los diferentes vectores h . Algunos de los parámetros que intervienen en la búsqueda de parejas de datos son: tolerancia angular, espaciamiento del vector h , distancia máxima de h , tolerancia del espaciamiento y ancho de banda para limitar la tolerancia del ángulo de análisis (Deutsch y Journel, 1992). La tabla 4.3 indica los valores usados de estos parámetros en la estimación del variograma. La figura 4.7 muestra los diferentes variogramas experimentales obtenidos de la información sísmica de los años 1993 y 1994 en su totalidad, y para ventanas de tiempo trimestrales a partir de octubre de 1993 hasta septiembre de 1994. También, puede observarse en uno de los variogramas de cada gráfica, el número de parejas de datos encontrado en la evaluación de cada h en particular. Para los otros variogramas estos números son similares.

Tabla 4.3. Valores de los parámetros para la estimación de variogramas experimentales con la información sísmica de 1993 y 1994.

Parámetro	Valor en grados	Valor en km (aproximado)
Tolerancia angular	15°	-
Espaciamiento de h	0.005	0.5
Distancia máxima de h	0.15	15
Tolerancia del espaciamiento	0.0025	0.25
Ancho de banda	0.005	0.5

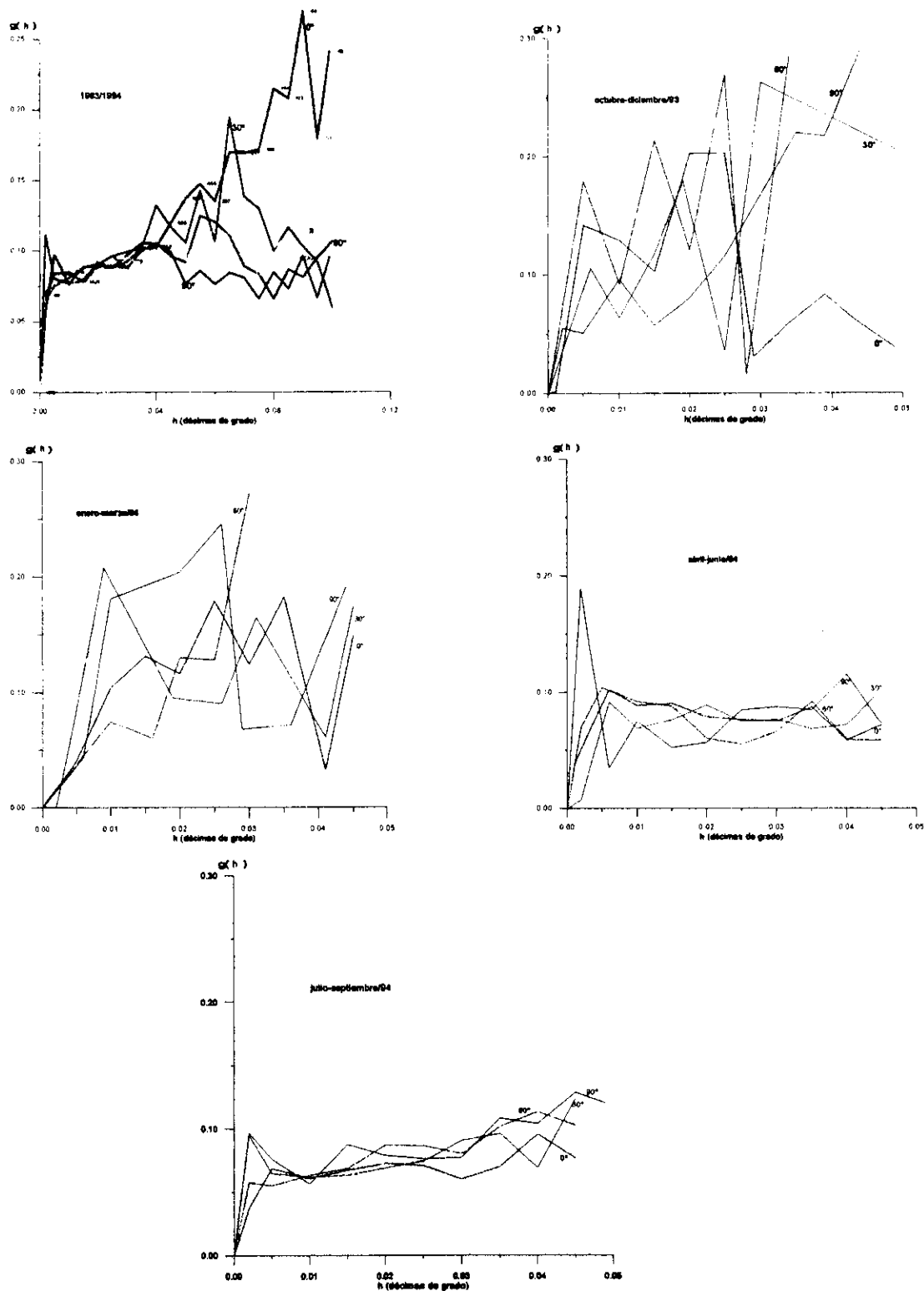


Figura 4.7. Variogramas experimentales de la magnitud de sismos para seis direcciones azimutales (0° , 30° , 60° , 90° , 120° y 150°). h representa el vector que separa cada pareja de datos para cada una de las direcciones analizadas.

Independientemente de la interpretación de los variogramas experimentales, el modelado de éstos es esencial en el proceso de estimación de la variable $Z(x)$, o bien, de la estimación de probabilidades asociadas a dicho proceso. Los modelos más comunes de variogramas experimentales son (Isaaks y Srivastava, 1989):

a) Modelo Gaussiano: $g(h) = 1 - \exp(-3h^2/a^2)$,

indicado por $g(h) = c \text{ Gauss}(a)$, donde c se refiere al factor de escala o contribución del modelo y a es el rango utilizado en la evaluación de $g(h)$, por ejemplo: $0.5\text{Gauss}(0.1)$, es un modelo gaussiano multiplicado 0.5 veces y evaluado con un rango de 0.1. Las expresiones utilizadas para los otros modelos son similares.

b) Modelo Exponencial: $g(h) = 1 - \exp(-3h/a)$,

indicado por $g(h) = c \text{ Exp}(a)$.

c) Modelo Esférico: $\gamma(h) \begin{cases} = 1.5h/a - 0.5(h/a)^3 & \text{si } h \leq a \\ \text{de otra forma} & \end{cases}$,

indicado por $g(h) = c \text{ Esf}(a)$.

d) Modelo Lineal: $g(h) = ? h ?$,

indicado por $g(h) = c |h|$.

e) Modelo Efecto pepita: $\gamma_0(h) \begin{cases} = 0 & \text{si } h=0 \\ = 1 & \text{de otra forma} \end{cases}$,

indicado por $g_0(h) = c$,

donde el parámetro a es el rango de la función, es decir, el valor de h por arriba del cual el variograma indica que no hay correlación estadística entre las parejas $[Z(x), Z(x+h)]$.

La idea de modelar los variogramas experimentales obedece a la necesidad de caracterizar la variabilidad espacial, en todas direcciones, del fenómeno en estudio, además de que los resultados obtenidos con los variogramas experimentales, provienen en general de datos limitados e imprecisos. Cabe mencionar que, una vez capturada la variabilidad

espacial en un modelo de variograma, éste será utilizado durante el proceso de estimación de la variable $Z(x)$ a través de técnicas de kriging, más adelante expuestas.

Para modelar el variograma experimental se puede ajustar un solo modelo, o combinar dos o más de los modelos mencionados. La figura 4.8 muestra las gráficas de estos modelos.

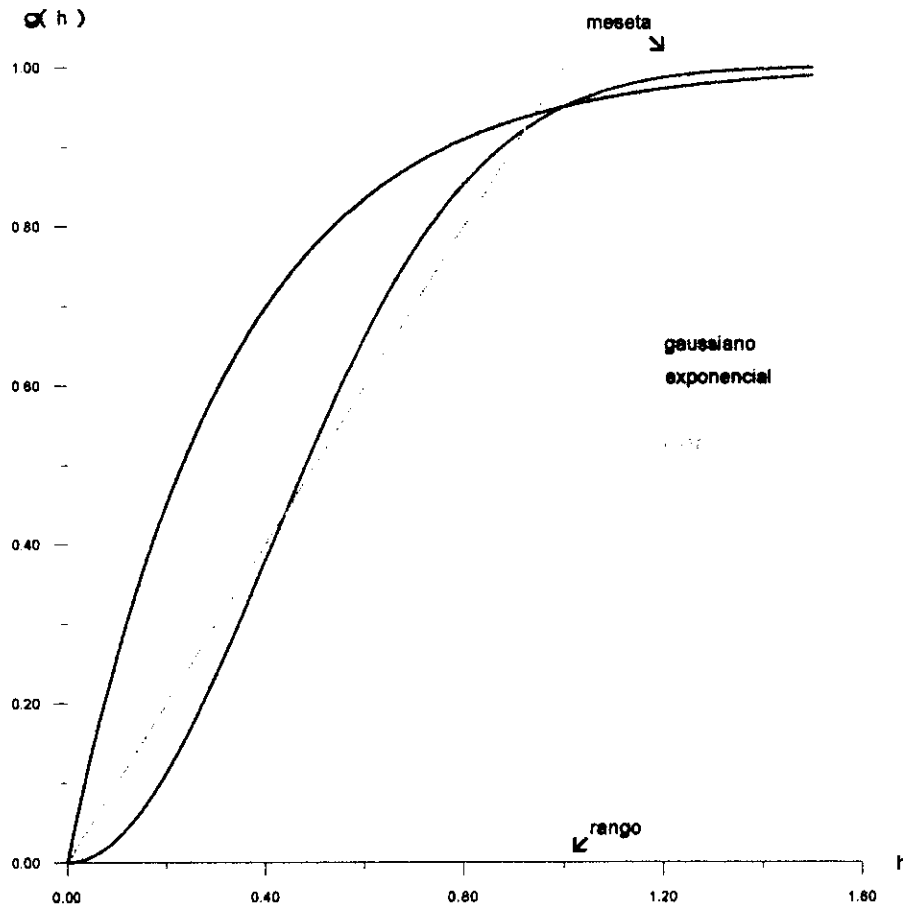


Figura 4.8. Modelos más comunes de variogramas. Todos los modelos representados tienen un rango = 1 y meseta = 1, excepto el modelo lineal que no tiene meseta y rango.

Cuando el fenómeno analizado presenta una variabilidad espacial semejante en cualquier dirección, todos los variogramas experimentales pueden ser resumidos en un solo modelo denominado isotrópico. Por el contrario, si los variogramas experimentales denotan alguna dirección de mayor o más rápida variación con respecto a otra(s), se está ante un caso de anisotropía. Esta puede ser geométrica, cuando se tienen variancias semejantes en todas direcciones pero diferentes rangos, o anisotropía zonal, cuando hay diferentes variancias (mesetas) para las diversas direcciones (Isaaks y Srivastava, 1989). La figura 4.9

ejemplifica los casos antes citados, mientras que en la figura 4.10 se muestra un esquema bidimensional del caso de anisotropía geométrica.

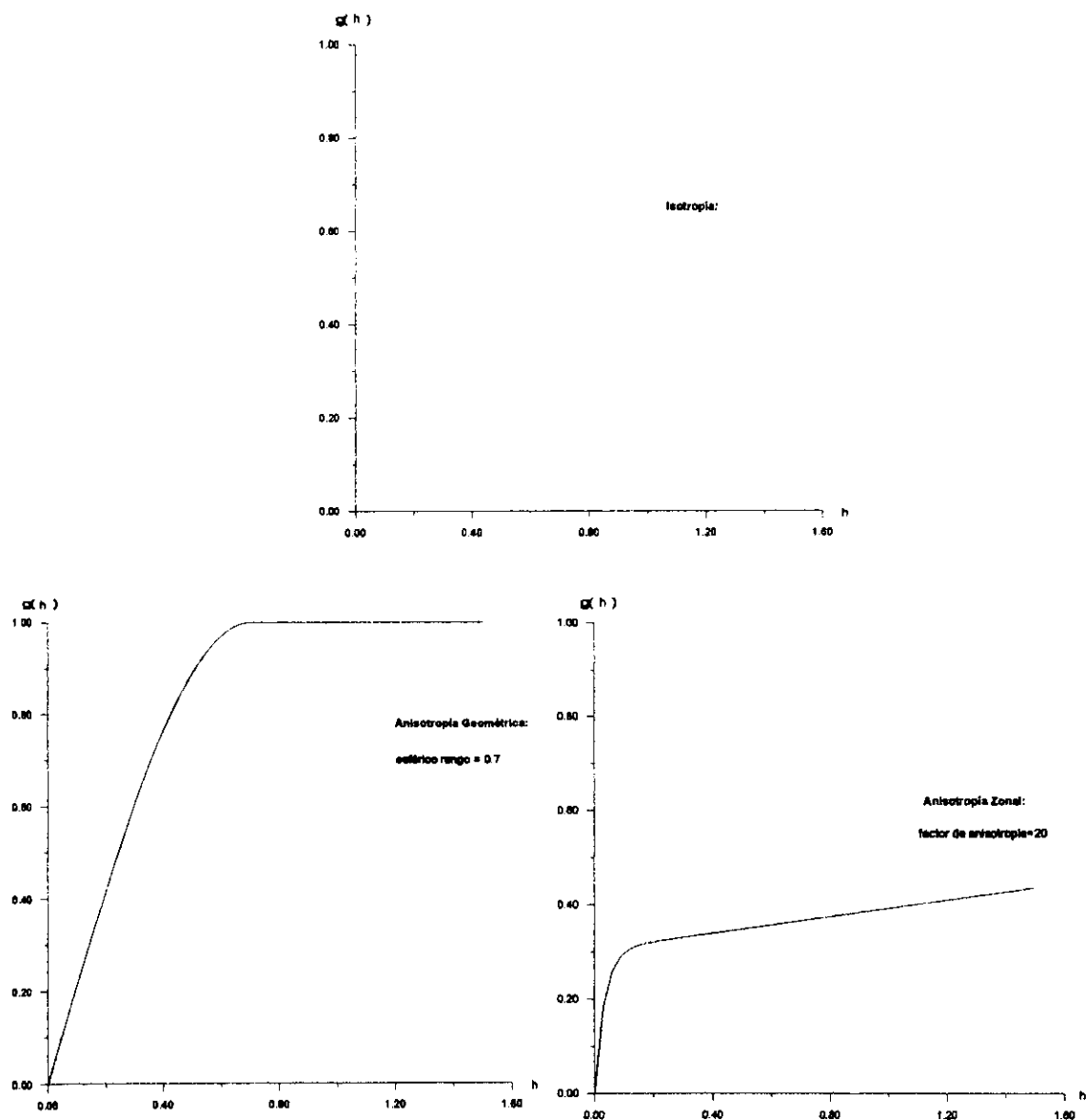


Figura 4.9. Modelos representativos de isotropía, anisotropía geométrica y anisotropía zonal.

En el esquema de la figura 4.10 están representadas las dos direcciones de mayor y menor continuidad, eje rojo y azul respectivamente, del caso de anisotropía geométrica de la figura 4.9. La elipse representa una superficie de igual valor de la función variograma. En la dirección del eje rojo el fenómeno varía más lentamente que en la dirección del eje azul, de tal forma que llega al valor de meseta, $C = 1$, cuando h toma el valor del rango mayor,

mientras que en la dirección del eje azul el variograma llega al valor de su meseta cuando h toma el valor de rango menor

El poder capturar esta anisotropía es muy importante en la estimación de $Z(x)$ ya que, como será expuesto más adelante, para estimar valores de $Z(x)$ en localidades no muestreadas, se considera la información de una vecindad alrededor del punto a estimar, pero esta información puede no contribuir de igual manera a dicha estimación. Lo anterior puede ser explicado utilizando los puntos 1, 2, 3 y 4 de la figura 4.10, donde aparentemente los puntos se encuentran a la misma distancia del punto de intersección de los ejes, punto 0, sin embargo, si se considera la distancia en términos de la covariancia y no distancia euclidiana, resulta que los puntos 1 y 3 se ubican en la dirección de mayor continuidad, y los puntos 2 y 4 en la dirección de menor rango o mayor variación, por lo que no tienen el mismo peso en la estimación de $Z(x)$ en el punto 0, ya que sus valores de covariancia y/o variograma, dada la ecuación (9), no son iguales para un mismo valor de h .

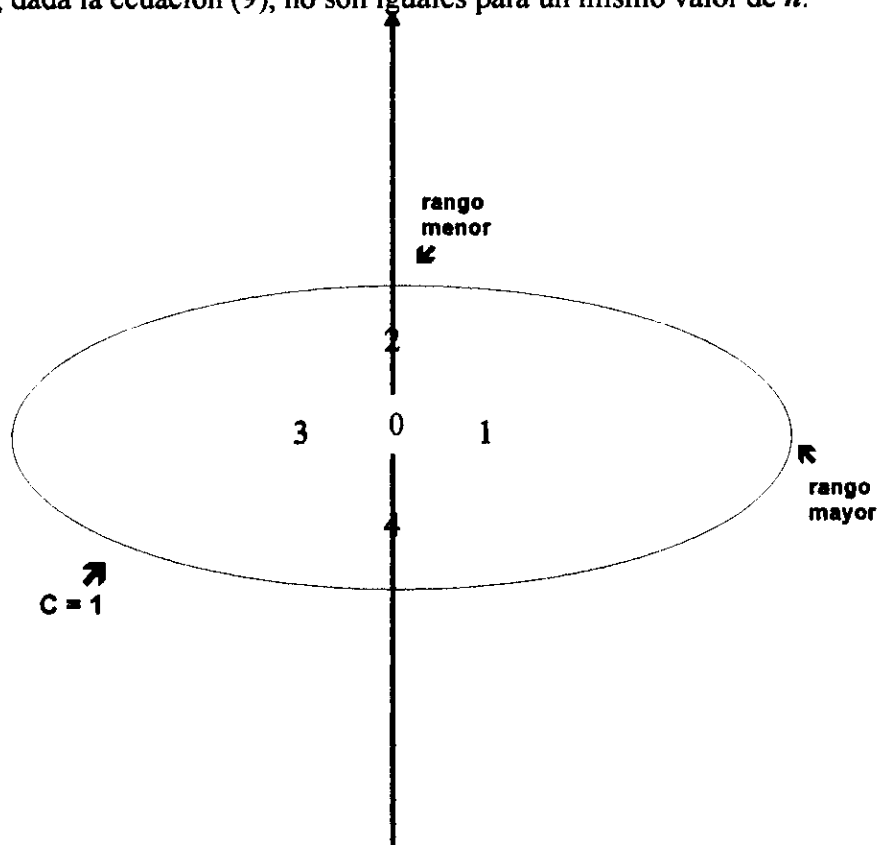


Figura 4.10. Esquema representativo del caso de anisotropía geométrica. La elipse representa la superficie de igual valor de la función variograma, en este caso $C = 1$. Los puntos 1, 2, 3 y 4 representan una vecindad alrededor del punto a estimar, punto 0.

Con base en lo anterior, se definieron modelos con los que se efectuó el proceso de estimación de la magnitud sísmica. Para el caso de los variogramas de 1993/1994 se obtuvieron dos modelos: uno isotrópico y el otro considera anisotropía zonal. Para el caso

de las ventanas de tiempo trimestrales, se obtuvo un modelo isotrópico para cada una de ellas. La figura 4.11 y 4.12 muestran, en seis gráficas, estos modelos con sus respectivos variogramas experimentales, mientras que en la tabla 4.4 se indican las expresiones para todos los casos considerados.

Tabla 4.4. Modelos de variabilidad espacial de la magnitud sísmica registrada de 1993-1994 en la zona del embalse de la presa Aguamilpa, Nayarit. Se consideraron diferentes ventanas de tiempo.

Intervalo de tiempo	Modelo	$\gamma(h)$
Oct/93-sep/94	Isotrópico	$0.05+0.025\text{Exp}(0.01)+0.045\text{Esf}(0.1)$
Oct/93-sep/94	Anisotropía zonal	$0.05+0.02\text{Exp}(0.003)+0.035\text{Esf}(0.13)$
Oct-dic/93	Isotrópico	$0.05+0.165\text{Esf}(0.04)$
Ene-mar/94	Isotrópico	$0.0+0.2\text{Exp}(0.04)$
Abr-jun/94	Isotrópico	$0.04+0.03\text{Exp}(0.01)+0.01\text{Esf}(0.08)$
jul-sep/94	Isotrópico	$0.04+0.02\text{Exp}(0.005)+0.04\text{Esf}(0.07)$

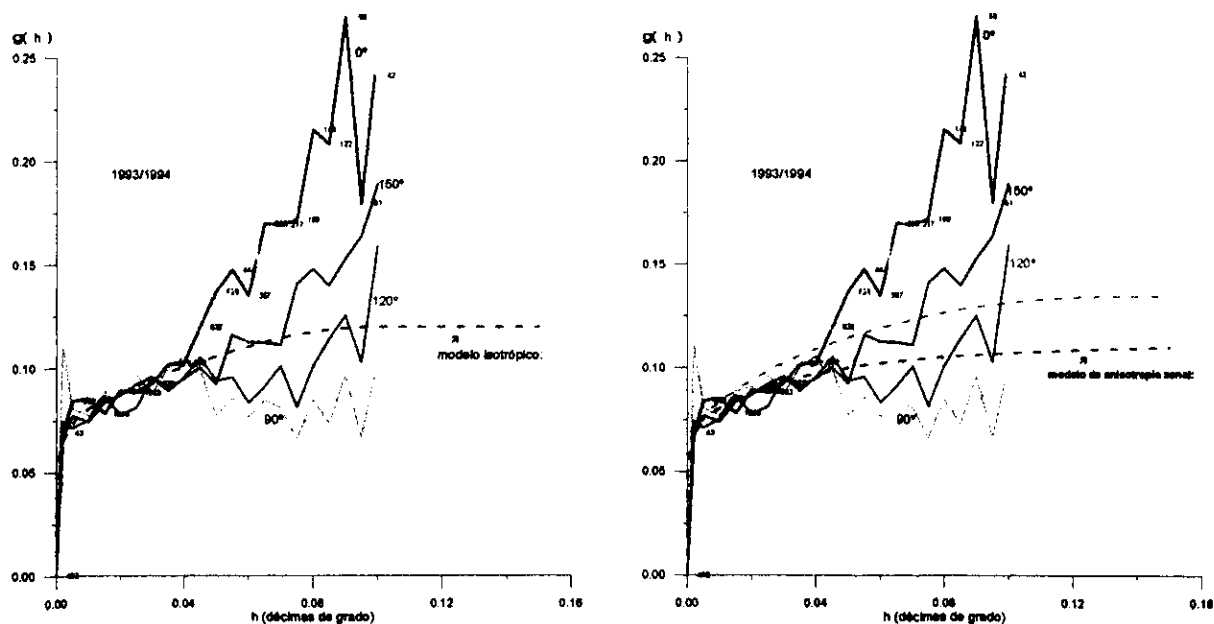


Figura 4.11. Variogramas experimentales y modelos de variabilidad espacial para la magnitud sísmica observada, de 1993 a 1994, en la zona del embalse de la presa Aguamilpa, Nayarit.

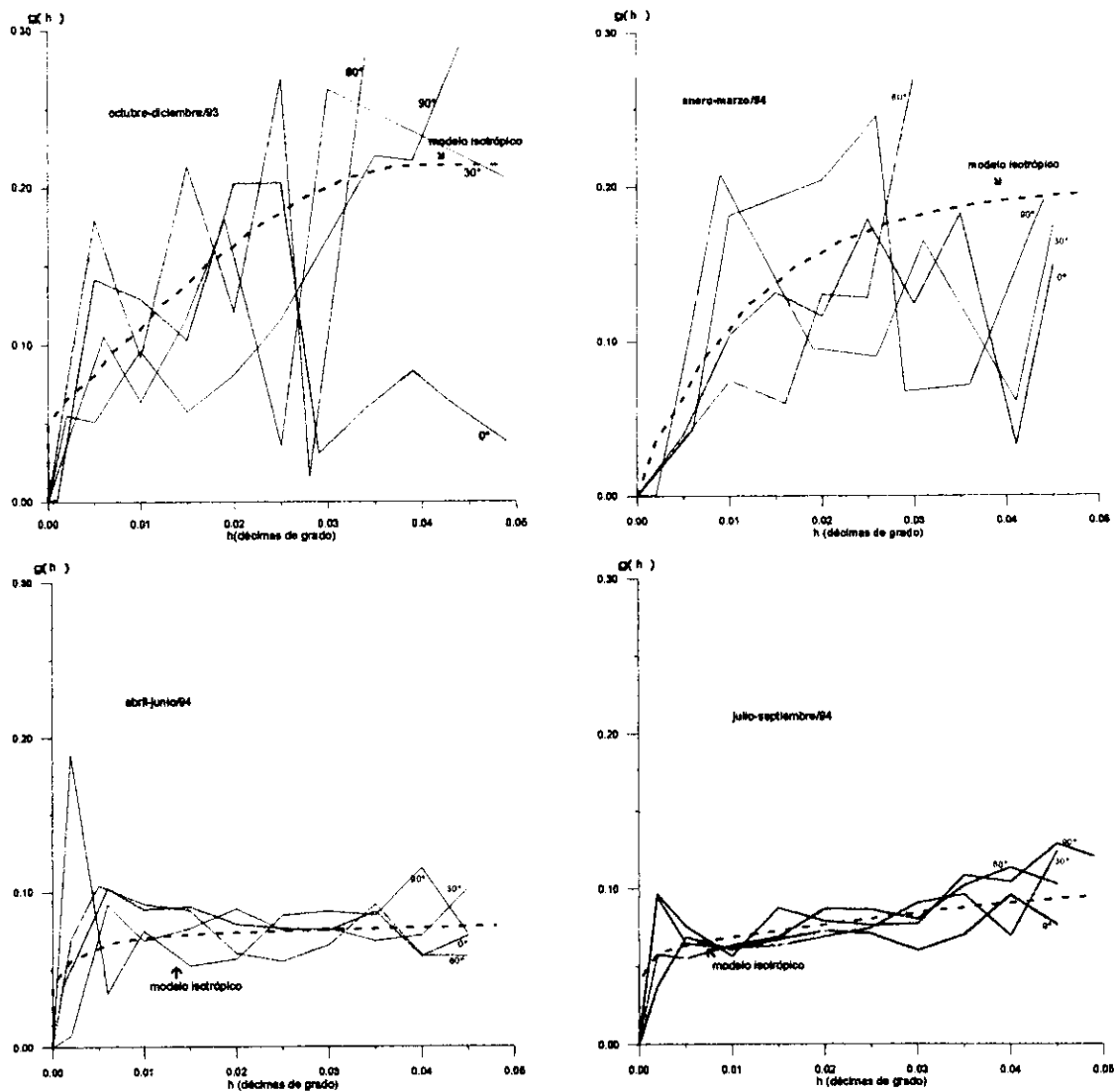


Figura 4.12. Variogramas experimentales y modelos de variabilidad espacial para la magnitud sísmica en ventanas de tiempo trimestrales. Zona del embalse de la presa Aguamilpa, Nayarit.

Para el caso de la información sísmica que cubre de octubre de 1993 a septiembre de 1994, es notable cómo los variogramas para todas las direcciones se comportan de manera similar hasta un valor de $h \approx 0.05$. Después de este valor, los variogramas para las direcciones de 0° , 30° y 150° , es decir, direcciones con tendencia N-S, presentan valores mayores de meseta que las direcciones 60° , 90° y 120° . También, es notable que el rango para las direcciones orientadas E-O es menor que para las orientadas N-S. En el caso de los variogramas de la figura 4.12 (información trimestral), el comportamiento de estos es similar para todas las direcciones, por lo que los modelos isotrópicos se ajustan

adecuadamente no obstante la dispersión que se tiene en los casos de los trimestres octubre-diciembre/1993 y enero-marzo/1994. Lo anterior es debido a la escasez de datos ya detectada en la fase exploratoria (histogramas y valor del parámetro b) de los mismos.

Cabe recordar que el rango representa la distancia, módulo del vector h , por arriba de la cual hay ausencia de correlación entre los datos. En consecuencia, al observar los valores de rango de los modelos de la tabla 4.4, puede decirse que los datos de magnitud sísmica, de la zona del embalse de la presa Aguamilpa, Nayarit, están correlacionados entre sí a distancias mayores a medida que la sismicidad se establece, ya que los valores de rango para periodos de sismicidad temprana son del orden de 0.04 grados de longitud, 4 km aproximadamente, mientras que para periodos de sismicidad posteriores, de abril a septiembre de 1994, los valores de rango están entre 0.07 y 0.08 (7-8 km). De tal forma que al abarcar todo el período de análisis, 1993 y 1994, el rango observado oscila entre 0.1 y 0.13 grados de longitud (10-13 km).

Las observaciones anteriores reflejan que la sismicidad se presenta preferentemente en direcciones con tendencia N-S o bien que tiene mayor continuidad en estas direcciones, además que la ocurrencia de sismos a partir del llenado del embalse, abarca áreas mayores con el tiempo. También, es notable que la mayoría de los modelos tienen un efecto pepita importante, lo cual implica que para distancias pequeñas existe una alta variabilidad en la magnitud sísmica, es decir, ocurren temblores de magnitudes pequeñas y grandes a distancias relativamente cercanas o bien, existen agrupaciones de sismos en zonas específicas del área estudiada.

4.1.2.3 Variogramas indicadores, isotropía y modelos.

El modelado de variogramas experimentales es necesario e importante en el proceso de estimación. Sin embargo, si se requiere un modelo de incertidumbre alrededor del valor estimado se deberá obtener la distribución de probabilidad de $Z(x)$. Esta distribución puede ser modelada a partir de datos ubicados en la vecindad de $Z(x)$ (Journel, 1989).

Si los datos con los que se pretende estimar $Z(x)$ fueran espacialmente independientes, es decir, datos tanto cercanos como lejanos al punto de estimación contribuyen con el mismo peso en dicha estimación, un posible modelo de incertidumbre alrededor del estimado de $Z(x)$ estaría dado por la distribución de valores de los n datos de la vecindad considerada. Además, se puede decir que la probabilidad de que el valor estimado de $Z(x)$ sea menor o igual a un determinado umbral z_c , corresponde a la proporción de datos de la vecindad considerada que sean iguales o menores a dicho umbral, es decir,

$$\text{Prob}\{Z(x) \leq z_c | (n)\} \approx p = \{\text{proporción de datos } z(x_a) \leq z_c\}, a = 1, \dots, n, \dots \quad (10)$$

donde (n) corresponde a la vecindad de datos, z_c es el umbral considerado.

Una manera de estimar dicha proporción hace uso de la variable indicador (Journel, 1989):

$$i(x_\alpha; z_c) = \begin{cases} 0, & \text{si } z_c < z(x_\alpha) \\ 1, & \text{si } z_c \geq z(x_\alpha) \end{cases} \quad \alpha = 1, \dots, n, \dots \quad (11)$$

entonces, como la probabilidad que se requiere conocer corresponde a la función de distribución acumulada, se tiene que:

$$F(x; z_c | (n)) = \text{Prob}\{Z(x) \leq z_c | (n)\} \cong 1/n \sum_{a=1}^n i(x_a; z_c) \in [0,1], \dots \quad (12)$$

donde la sumatoria significa el número acumulado de datos que cumplen la condición del umbral, que al dividirse entre n se logra la estimación de la proporción buscada. Nótese que la relación (12) establece una aproximación a la probabilidad buscada, ya que al asumir independencia espacial entre los datos se tiene que:

$$\text{Prob}\{Z(x) \leq z_c | (n)\} = \text{Prob}\{Z(x) \leq z_c\} \cong E[I(x_a; z_c)] = 1/n \sum_{a=1}^n i(x_a; z_c).$$

Ahora bien, para el caso general de dependencia espacial entre los datos,

$$F(x; z_c | (n)) = \text{Prob}\{Z(x) \leq z_c | (n)\} \cong \sum_{a=1}^n a_a(x) i(x_a; z_c) \in [0,1], \dots \quad (13)$$

entonces, la probabilidad buscada es un promedio pesado de la variable indicador. La ecuación (13), como será expuesto en la fase de estimación, puede ser vista como una estimación de la variable indicador.

En consecuencia, como fue mencionado en el apartado anterior, para estimar una variable espacialmente distribuida, se requiere conocer un modelo de variabilidad espacial de dicha variable. Además, para estimar la función de distribución acumulada, se requiere hacer el análisis para diferentes umbrales, por lo que se requerirá estimar variogramas indicadores para diferentes umbrales o valores de corte.

De acuerdo con lo anterior, se calcularon para las diferentes ventanas de tiempo y para el período mayor de la información considerada (octubre/1993 a diciembre/1994), variogramas indicadores para cinco diferentes umbrales: $z_c = 2.23, 2.41, 2.63, 2.80$ y 3.15 ,

que corresponden a los cuantiles del 0.25, 0.50, 0.75, 0.85 y 0.95 de la información sin ventanas de tiempo. Las gráficas de los variogramas experimentales de $i(x_a, z_c)$ con sus respectivos modelos, para la información de 1993 a 1994 considerando las mismas direcciones y parámetros de los variogramas de la variable $Z(x)$, son mostradas en la figura 4.13, mientras que en la tabla 4.5 se incluyen las formulas de los modelos para todos los casos. Todos estos modelos han sido considerados en la estimación de la incertidumbre asociada a la estimación de $Z(x)$, ya sea con el enfoque de la técnica de kriging indicador (KI) o en la fase de simulación estocástica con variables indicadoras. Las técnicas mencionadas serán expuestas en los capítulos siguientes.

Tabla 4.5. Modelos de variabilidad espacial de la variable indicador de la magnitud sísmica. Se incluyen los modelos del análisis con y sin ventanas de tiempo para los umbrales $z_c = 2.23, 2.41, 2.63, 2.80$ y 3.15 .

Intervalo de tiempo	Umbral	$g(h)$
Oct/93-sep/94	2.23	$0.15+0.05\text{Exp}(0.01)+0.0175\text{Esf}(0.1)$
	2.41	$0.15+0.08\text{Exp}(0.01)+0.02\text{Esf}(0.1)$
	2.63	$0.08+0.06\text{Exp}(0.01)+0.042\text{Esf}(0.1)$
	2.80	$0.05+0.04\text{Exp}(0.01)+0.032\text{Esf}(0.1)$
	3.15	$0.0+0.02\text{Exp}(0.01)+0.0186\text{Esf}(0.1)$
Oct/93-dic/93	2.23	$0.05+0.07\text{Exp}(0.02)$
	2.41	$0.1+0.14\text{Exp}(0.02)$
	2.63	$0.05+0.18\text{Exp}(0.03)$
	2.80	$0.05+0.15\text{Esf}(0.05)$
	3.15	$0.0+0.13\text{Esf}(0.05)$
Ene/94-mar/94	2.23	$0.1+0.11\text{Exp}(0.01)$
	2.41	$0.05+0.2\text{Exp}(0.015)$
	2.63	$0.05+0.17\text{Exp}(0.02)$
	2.80	$0.0+0.19\text{Exp}(0.03)$
	3.15	$0.0+0.185\text{Esf}(0.07)$
Abr/94-jun/94	2.23	$0.1+0.11\text{Exp}(0.018)$
	2.41	$0.1+0.15\text{Exp}(0.024)$
	2.63	$0.05+0.1\text{Exp}(0.018)$
	2.80	$0.0+0.06\text{Exp}(0.0045)+0.032\text{Esf}(0.05)$
	3.15	$0.0+0.009\text{Exp}(0.01)+0.0055\text{Esf}(0.04)$
Jul/94-sep/94	2.23	$0.1+0.135\text{Exp}(0.012)$
	2.41	$0.1+0.1\text{Exp}(0.015)+0.05\text{Esf}(0.06)$
	2.63	$0.05+0.03\text{Exp}(0.008)+0.07\text{Esf}(0.07)$
	2.80	$0.0+0.05\text{Exp}(0.0075)+0.05\text{Esf}(0.07)$
	3.15	$0.0+0.005\text{Exp}(0.01)+0.01\text{Esf}(0.07)$

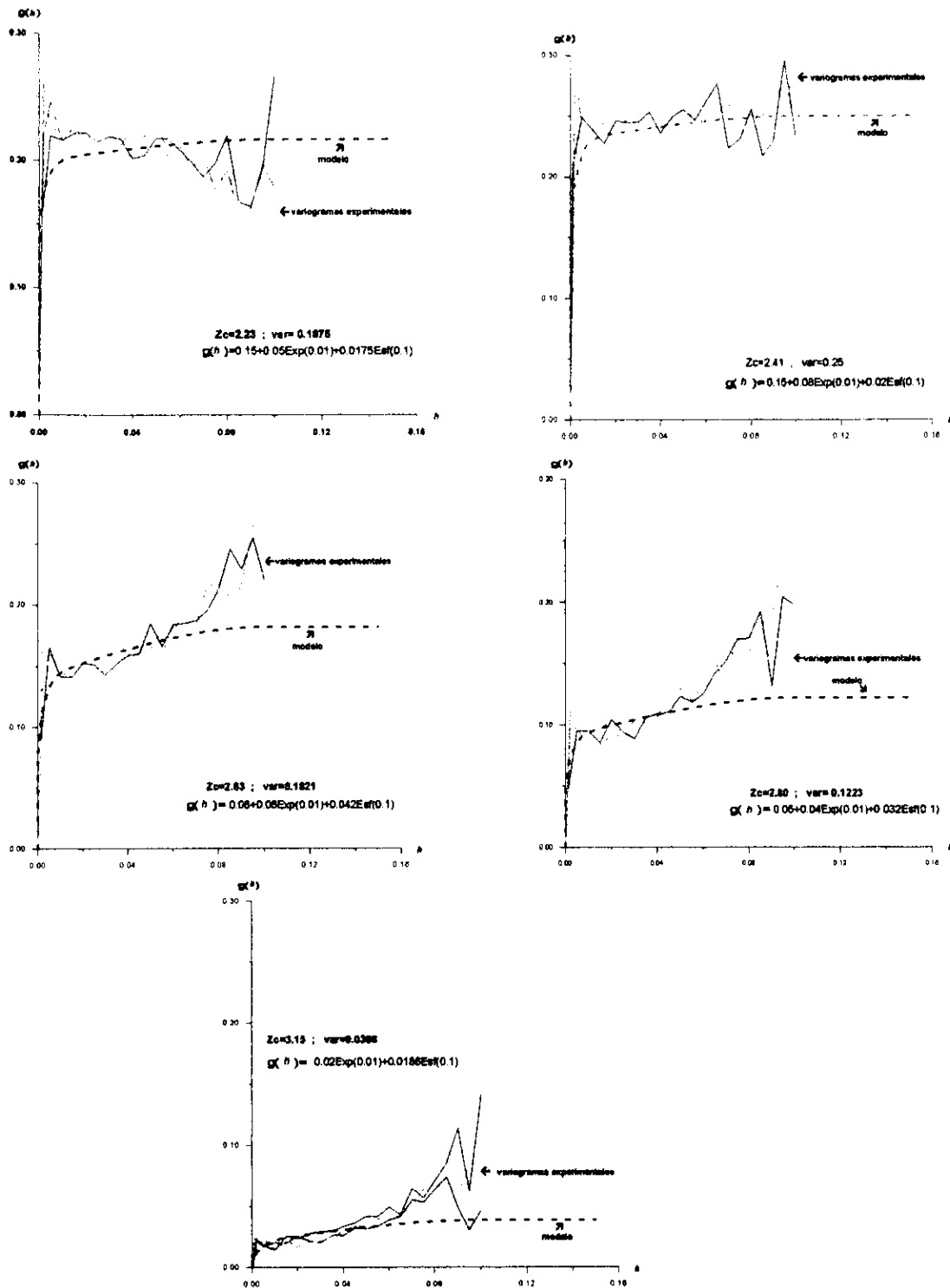


Figura 4.13. Variogramas experimentales para la variable indicador de la magnitud sísmica, y sus respectivos modelos. La información considerada abarca de octubre de 1993 a septiembre de 1994. El análisis se hizo para $z_c = 2.23, 2.41, 2.63, 2.80$ y 3.15 .

4.2 Estimación

El enfoque a utilizar en esta fase de estimación es el de kriging y específicamente la técnica de kriging ordinario (KO). Kriging es un término adoptado para una familia de algoritmos de estimación basados en la teoría de la regresión lineal. La idea básica es la de estimar un valor desconocido en una localidad no muestreada x_o , a partir de una combinación lineal de n valores conocidos en localidades específicas x_a , $a=1, \dots, n$. (Journel, 1998). La forma general del estimador, denominado kriging simple (KS), es:

$$Z^*(x_o) = 1_o + \sum_{a=1}^n 1_a Z(x_a), \quad \dots \quad (14)$$

en donde al evaluarse la expresión anterior se tendrá el valor estimado $z^*(x_o)$. La diferencia entre el valor real $z(x_o)$ y el valor estimado $z^*(x_o)$ define el error de estimación $z(x_o) - z^*(x_o)$. Cabe aclarar, con respecto a la notación utilizada, que las minúsculas se refieren a los datos o valores estimados, mientras que las mayúsculas denotan variables aleatorias asociadas a estos valores. Dos de las características del estimador de KS, y que marcan la diferencia con otras técnicas de estimación espacial son:

1) La propiedad de insesgamiento del error, es decir,

$$E[Z(x_o) - Z^*(x_o)] = 0.$$

2) La variancia del error es mínima,

$$s^2_E(x_o) = \text{var}[Z(x_o) - Z^*(x_o)],$$

por lo anterior, las técnicas geoestadísticas de kriging resultan apropiadas para interpolar o extrapolar información en localidades no muestreadas, ya que el estimador es insesgado y emplea el modelo de variabilidad espacial de $Z(x)$ en la solución de los pesos del estimador, a través de las covariancias de los datos con el punto a estimar y las covariancias entre los mismos datos, lo cual será expuesto en el siguiente apartado.

4.2.1 Técnica de kriging ordinario.

Debido a las condiciones de variabilidad espacial de la magnitud sísmica, es difícil esperar que esta variable tenga un valor constante distribuido sobre el área analizada. Por el contrario, se espera que el valor medio de $Z(x)$, considerando una vecindad alrededor del punto a estimar, sea diferente y represente la variabilidad sísmica observada en la zona del embalse de la presa Aguamilpa, Nayarit. Con base en lo anterior, el empleo de KO resulta

conveniente para la estimación de la magnitud sísmica en localidades no muestreadas, ya que:

$$E[Z(x_0) - Z^*(x_0)] = E[Z(x_0)] - l_0 - \sum_{a=1}^n l_a E[Z(x_a)] = m_0 - l_0 - \sum_{a=1}^n l_a m_a = 0$$

∴

$$l_0 = m_0 - \sum_{a=1}^n l_a m_a,$$

por lo que

$$Z^*(x_0) = m_0 + \sum_{a=1}^n l_a [Z(x_a) - m_a],$$

que para el caso de interpolación espacial,

$$Z^*(x_0) - m = \sum_{a=1}^n l_a [Z(x_a) - m],$$

donde m es el valor esperado o media, que se asume conocido de las n variables aleatorias $Z(x_a)$, $a = 1, \dots, n$. Sin embargo, para el caso que nos ocupa, no se puede asumir una media conocida, común o constante, sobre el dominio de análisis, por lo que independientemente del valor de m y al considerar nuevamente la propiedad de insesgamiento sobre el estimador de KS:

$$E[Z(x_0) - Z^*(x_0)] = m - l_0 - \sum_{a=1}^n l_a m = -l_0 + m(1 - \sum_{a=1}^n l_a) = 0$$

ecuación que se cumple para:

$$l_0 = 0 \text{ y}$$

$$\sum_{a=1}^n l_a = 1,$$

por lo que el estimador, llamado ahora como KO, toma la forma

$$Z^*(x_o) = \sum_{a=1}^n \lambda_a Z(x_a), \text{ con la restricción: } \sum_{a=1}^n \lambda_a = 1. \quad \dots\dots (15)$$

Una vez establecido el modelo de variación espacial de $Z(x)$ sobre el área de interés, se está en posibilidad de usar el algoritmo de KO, requiriendo del conocimiento de la matriz de covariancias entre los mismos datos y la matriz de covariancias entre los datos y la posición a estimar. Es decir, para obtener los pesos λ_a del estimador de KO con su respectiva restricción (15), se deberá resolver el sistema de ecuaciones (ver apéndice 1):

$$\sum_{b=1}^n \lambda_b \text{Cov}(x_a, x_b) + m = \text{Cov}(x_a, x_o), \quad a=1, \dots, n, \quad \dots\dots (16)$$

el cual requiere de conocer los términos $\text{Cov}(x_a, x_b)$ y $\text{Cov}(x_a, x_o)$, que representan respectivamente, las matrices de covariancias entre los mismos datos y covariancias entre los datos y el punto a estimar. En la expresión (16), m es el multiplicador de Lagrange dada la restricción de que los pesos λ sumen 1.

Las matrices de covariancias antes mencionadas son definidas a partir del modelo de variograma previamente establecido, ya que de acuerdo a la expresión (9), la covariancia puede ser descrita en términos del variograma por:

$$C(\mathbf{h}) = C(0) - \gamma(\mathbf{h}),$$

donde $C(0)$ representa la variancia de los datos (Isaaks y Srivastava, 1989).

Es importante resaltar que cualquier elemento de las matrices de covariancias arriba mencionadas, corresponde a la covariancia entre dos posiciones, ya sea entre datos o entre un dato y la posición a estimar. Estas dos posiciones, que están en una dirección y distancia específicas, definen al vector \mathbf{h} con el cual se evalúa la covariancia entre estas dos posiciones. Suponiendo un arreglo de tres datos para estimar la variable asociada a estos datos en un determinado punto, la forma matricial de la ecuación (16) es:

$$\begin{bmatrix} C(x_1, x_1) & C(x_1, x_2) & C(x_1, x_3) & m \\ C(x_2, x_1) & C(x_2, x_2) & C(x_2, x_3) & m \\ C(x_3, x_1) & C(x_3, x_2) & C(x_3, x_3) & m \\ 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} l_1 \\ l_2 \\ l_3 \\ 1 \end{bmatrix} = \begin{bmatrix} C(x_1, x_0) \\ C(x_2, x_0) \\ C(x_3, x_0) \\ 1 \end{bmatrix}$$

Asimismo, se observa que en la ecuación (15) interviene el número de datos definido por el valor de n ($a=1, \dots, n$). Lo anterior implica que en el algoritmo de kriging ordinario se requiere elegir este valor n (número de muestras a considerar en la estimación) a través de una vecindad. Podría esperarse que entre mayor sea el número de datos mayor será el área de la vecindad, sin embargo, si la densidad de datos alrededor del punto a estimar es alta, el valor de n puede ser elevado con vecindades de áreas pequeñas y viceversa. Si la densidad de datos es pobre, se requerirá una vecindad amplia para alcanzar una n representativa. Con base en lo anterior se pueden esperar dos situaciones de estimación:

- En lugares donde la densidad de datos es pobre, la estimación de $Z(x)$ tenderá a utilizar un valor de media local de una vecindad mayor, mientras que en zonas de alta densidad de datos, se estará estimando $Z(x)$ con medias representativas de una vecindad menor. Lo anterior es debido a que en zonas de baja densidad de información, para satisfacer el número de datos de la vecindad considerada, se toma datos a distancias mayores que el rango de correlación del modelo de variabilidad y en estos casos, el estimador de KO reconoce que la mejor estimación es la media de los datos abarcados (ver apéndice 2).
- Entre mayor sea el número de datos a considerar para la estimación, se tenderá más al suavizamiento de los datos en la estimación, lo cual es consecuencia de la propiedad de mínima variancia.

En las figuras 4.14, 4.15, 4.16 y 4.17 se ejemplifica la situación de estimación, explicada en el inciso (a), para los datos de las diferentes ventanas de tiempo. Mientras que en la figura 4.18 se ejemplifica el caso de estimación del inciso (b) para una ventana de tiempo.

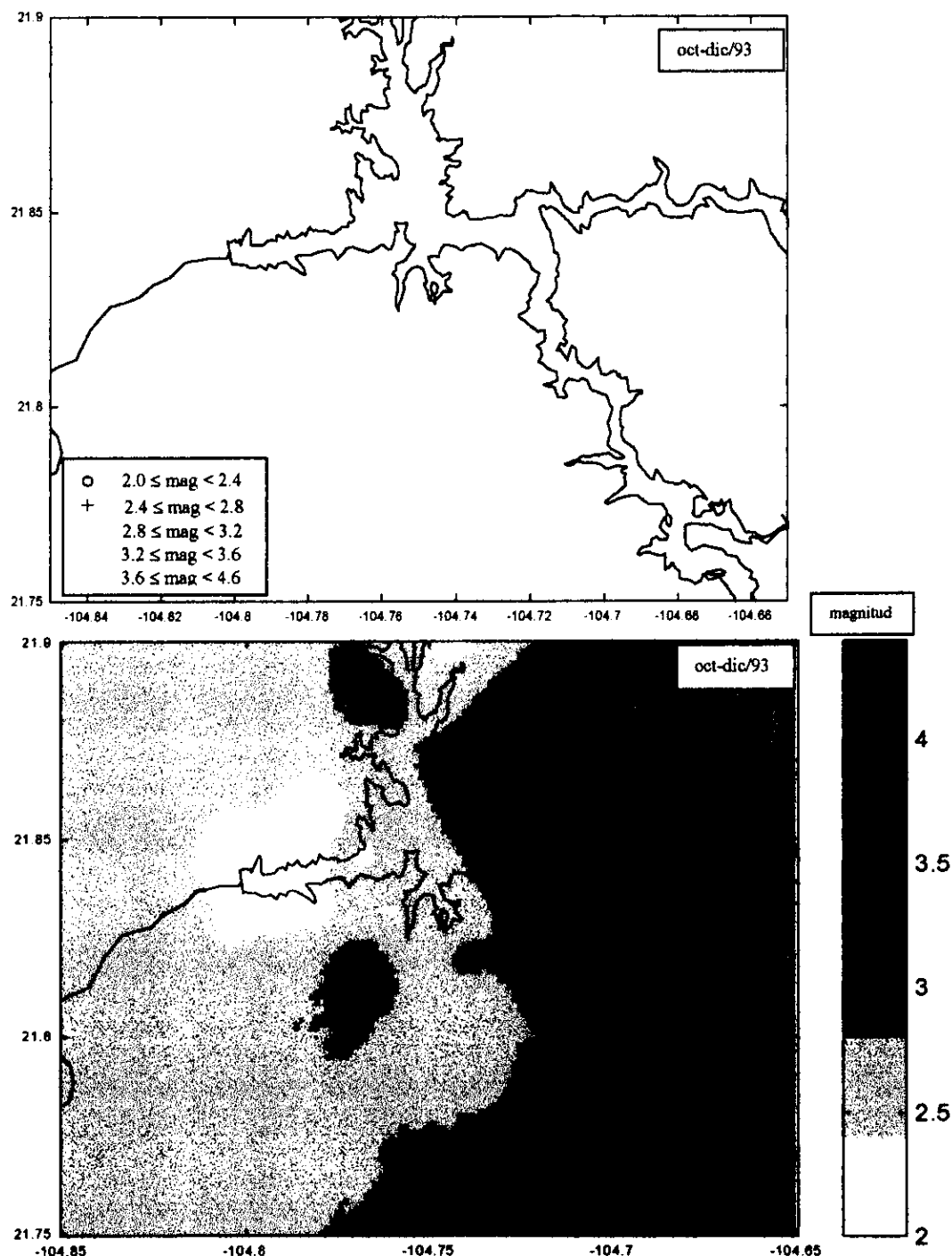


Figura 4.14. En la gráfica superior se observa la distribución espacial de los datos de magnitud sísmica del período octubre-diciembre de 1993. El mapa inferior es el resultado del kriging ordinario para estos datos, donde se aprecia la influencia del rango del modelo de variograma respectivo (0.04) en la dimensión de las áreas cerradas alrededor de los datos y en los valores estimados. Mientras que en zonas de pocos datos, los valores estimados representan la media aritmética de los datos abarcados en la vecindad.

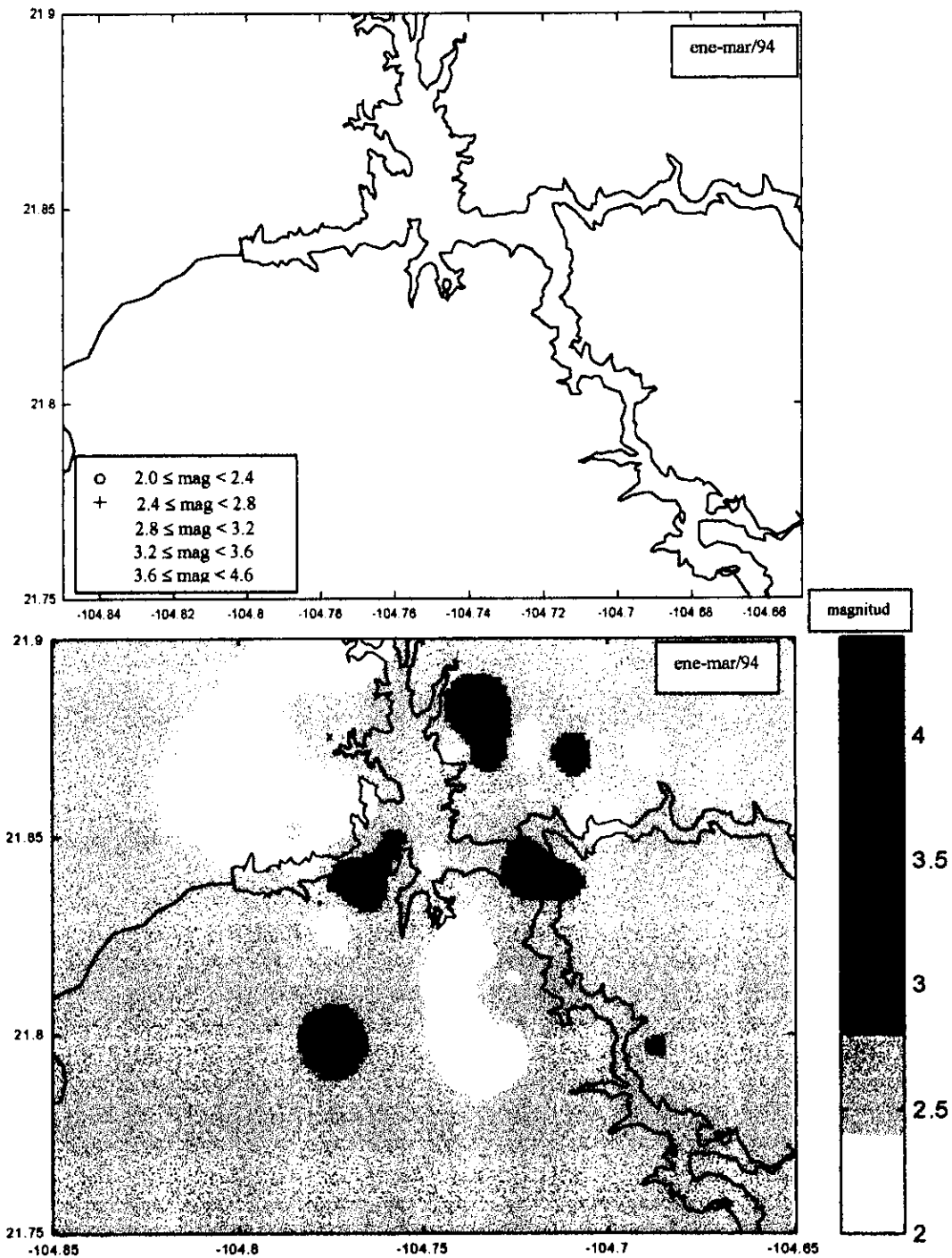


Figura 4.15. En la gráfica superior se observa la distribución de datos del período enero-marzo de 1994. El mapa inferior es el resultado del kriging ordinario para estos datos. Es notable la influencia del rango del variograma respectivo y los valores conocidos en la dimensión de las áreas cerradas y los valores estimados en éstas. Las zonas con poca o nula información son estimadas con la media aritmética de la vecindad considerada.

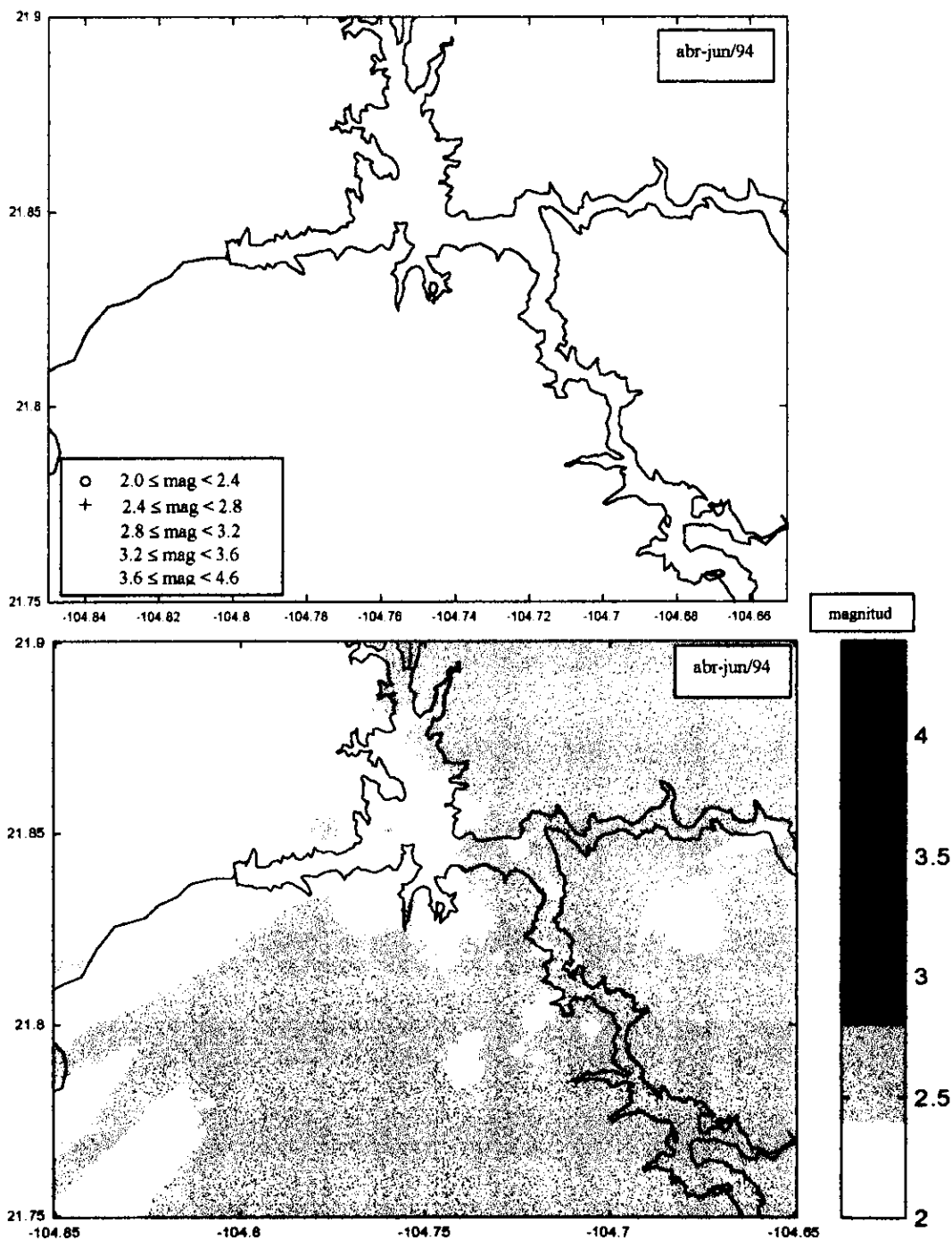


Figura 4.16. El mapa superior es la distribución de datos de magnitud sísmica del período abril-junio de 1994, mientras que el mapa inferior es el resultado del kriging ordinario para estos datos. Debido a que la mayoría de los datos varían en un intervalo entre 2.0 y 2.8 y dada la alta densidad de éstos y el rango del variograma respectivo (0.8), los valores estimados se agrupan solo en dos zonas de magnitud.

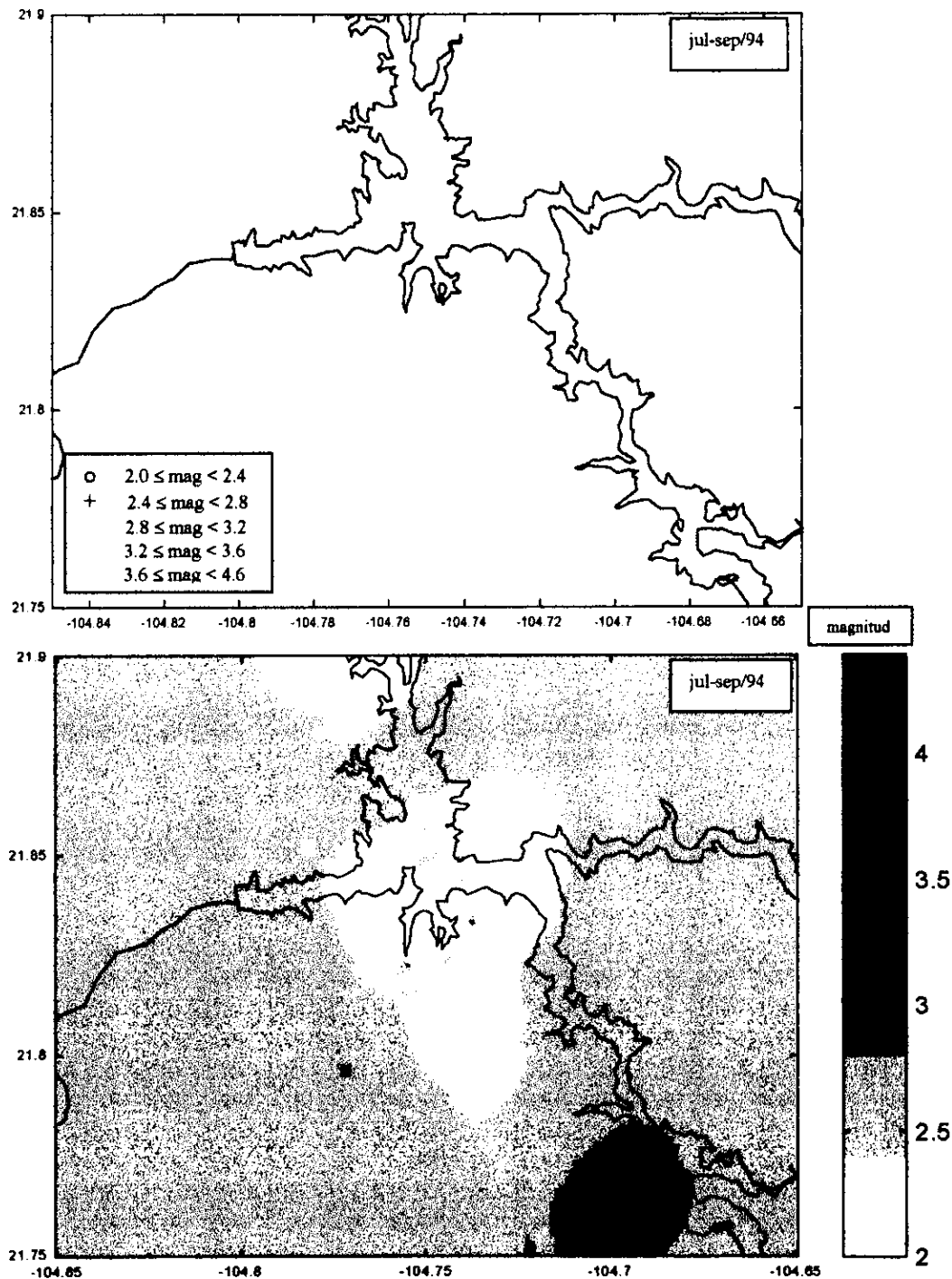


Figura 4.17. El mapa superior representa la distribución espacial de los datos del período julio-septiembre de 1994, mientras que el mapa inferior es el resultado de la estimación con kriging ordinario, donde es notable la influencia de la alta densidad de datos en la zona central y el rango de correlación respectivo. En la zona sureste del mapa los resultados reflejan la influencia de un dato aislado en la vecindad considerada.

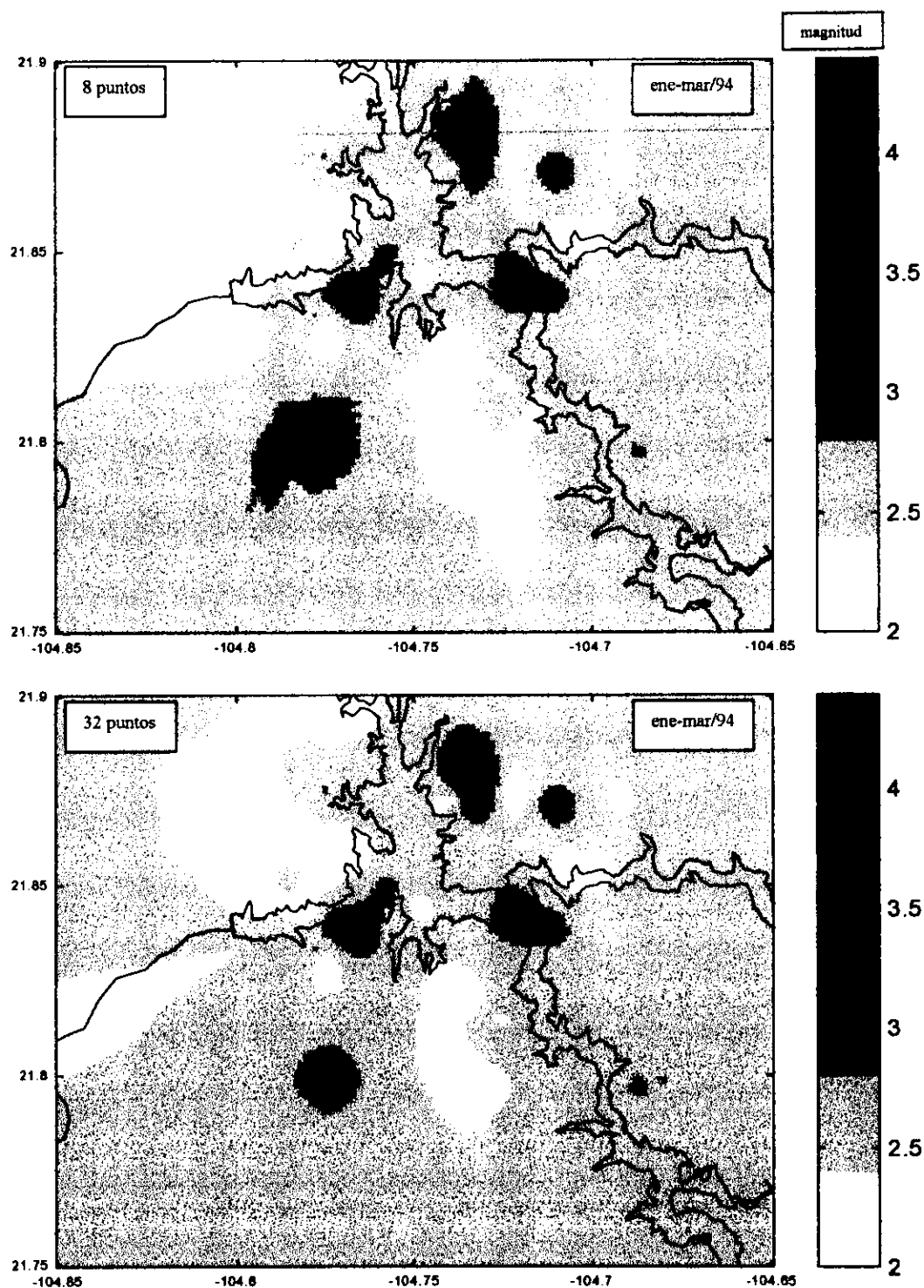


Figura 4.18. En el mapa superior se aprecia el resultado de estimación para el periodo enero-marzo de 1994 considerando una vecindad de 8 puntos, mientras que el mapa inferior es la estimación considerando 32 puntos en la vecindad. Es notable el efecto de mayor suavizamiento al considerar mayor número de datos en la vecindad.

Con base en los resultados observados en las figuras de la 4.14 a la 4.17, es importante enfatizar desde el punto de vista sismológico que:

- ◆ La utilidad de los mapas de estimación de magnitud sísmica, utilizando KO, es principalmente el observar tendencias de este fenómeno y la influencia del rango de correlación, de los modelos de variograma respectivos, en los valores estimados.
- ◆ Los valores estimados en localidades ubicadas a distancias, medidas con respecto a la ubicación de los datos, mayores que el rango del variograma respectivo, no necesariamente reflejan la posibilidad de ocurrencia de temblores en estas zonas, ya que en la estimación de estos valores no se han considerado variables de tipo geológico como la presencia de fallas o fracturas o las distancia a estas, que en última instancia controlan el fenómeno sísmico, además de la alta incertidumbre de los valores estimados en dichas zonas, lo cual será expuesto en el apartado y capítulo siguiente.

4.2.2 Técnica de kriging indicador

Como fue mencionado en la sección de variogramas indicadores, si se requiere modelar la incertidumbre asociada a la interpolación espacial de la magnitud sísmica, se deberá considerar una distribución de posibles valores de $Z(x)$ (Journel, 1989). Esta distribución podría ser la distribución acumulada, la cual necesita ser evaluada para diferentes umbrales. Un modelo de esta función es:

$$F(x; z_c | (n)) = \sum_{a=1}^n a_a(x; z_c) i(x_a; z_c) \in [0,1], \quad \dots \quad (17)$$

donde los pesos $a_a(x; z_c)$ suman uno.

Un punto a resaltar en la ecuación (17) es que el modelo de incertidumbre asociado a la estimación de $Z(x)$ puede ser visto como una estimación de la variable indicador para cada umbral, ya que de acuerdo a la forma del estimador de KO (15), los pesos 1_a corresponden a los pesos $a_a(x; z_c)$ y $Z(x_a)$ corresponde a $i(x_a; z_c)$ de la ecuación (17) respectivamente. Además, al considerar que los datos tiene pesos no iguales, se está reconociendo la existencia de patrones de dependencia espacial, capturados en los variogramas indicadores de los diferentes umbrales, entre los datos $z(x_a)$ y la incógnita $Z(x)$, consecuentemente, dependencia entre los datos indicadores $i(x_a; z_c)$ y la función indicadora desconocida $I(x; z_c)$ (Journel, 1989). La selección de los umbrales, como fue mencionado en el apartado de variogramas indicadores, se hizo en función de los cuantiles de 0.25, 0.50, 0.75, 0.85 y 0.95 del total de la información analizada, de tal forma que la función de distribución acumulada para cada punto a estimar, fue evaluada para cada uno de estos umbrales.

De la función de distribución acumulada se pueden derivar intervalos de probabilidad, es decir,

$$\text{Prob}\{Z(x) \in [a, b] | (n)\} = F(b; x | (n)) - F(a; x | (n)),$$

y por consiguiente, probabilidades de excedencia de $Z(x)$ con respecto a un determinado umbral,

$$\text{Prob}\{Z(x) > z_c | (n)\} = 1 - F(x; z_c | (n)), \quad \dots \quad (18)$$

a partir de la expresión (18) se pueden configurar mapas de igual valor de probabilidad de exceder un determinado umbral.

En las figuras 4.19, 4.20, 4.21 y 4.22 se muestran, en mapas de probabilidad de excedencia de magnitud sísmica, los resultados del kriging indicador (KI) para dos umbrales y las diferentes ventanas de tiempo. La malla que se utilizó en la obtención de dichos resultados fue de 100 x 100 puntos, comprendidos en un rango de 0.2 grados de longitud (20 km aproximadamente) y 0.15 grados de latitud (15 km aproximadamente).

Es notable en estas imágenes, cómo los valores de probabilidad varían considerablemente en zonas donde existe una alta densidad de datos (poca incertidumbre) con valores muy diferentes, como por ejemplo, la zona central del mapa inferior de la figura 4.19 o la parte central de los mapas de la figura 4.20, mientras que en zonas de baja densidad de datos, los valores de probabilidad varían poco, o bien, presentan lineamientos que no corresponden a tendencias sino a artefactos, como por ejemplo, las zonas de las esquinas de las figuras mencionadas y de las figuras 4.21 y 4.22. El hecho de que los valores de probabilidad en zonas de baja densidad de datos estén muy suavizados, obedece a que la estimación de la variable indicador es realizada a distancias mayores del rango de correlación del variograma indicador respectivo, por lo que el valor estimado corresponde a la media de los datos en la vecindad (ver apéndice 2). En el caso donde una localidad estimada coincide con una localidad de valor conocido (dato indicador), la función de distribución de probabilidad está representada por una función impulso unitario, que tiene valor de uno para el valor del dato $i(x_a; z_c)$, o bien, por un escalón unitario en el caso de la función de distribución acumulada.

Con respecto a los resultados mostrados en los mapas de probabilidad de excedencia de magnitud sísmica y su relación con las condiciones geológico-estructurales (principales fallas y fracturas) y argumentos sobre sismicidad inducida establecidos en el capítulo 1 de este trabajo, se observó lo siguiente:

- a) Cambio en el patrón de ocurrencia de la sismicidad, pasando de una sismicidad que cubre una amplia área ubicada principalmente sobre la zona Este del embalse, a una sismicidad más restringida en área (focos) y ubicada a mayor distancia del embalse con el tiempo. Sin embargo, en la zona SE del embalse, la sismicidad mayor a 2.8 persiste y los cambios son menores.
- b) Zonas dentro de las cuales se tienen los sismos más fuertes, observables en los mapas para $z_c = 2.8$, las cuales pueden ser asociadas con zonas sismogénicas.
- c) Estas zonas sismogénicas están presentes a lo largo de las principales fallas o fracturas y en el cruce de estas.
- d) Las zonas sismogénicas se han manifestado en diferentes lugares durante el período de análisis, sin embargo, en la zona sudeste del embalse los cambios son mínimos.

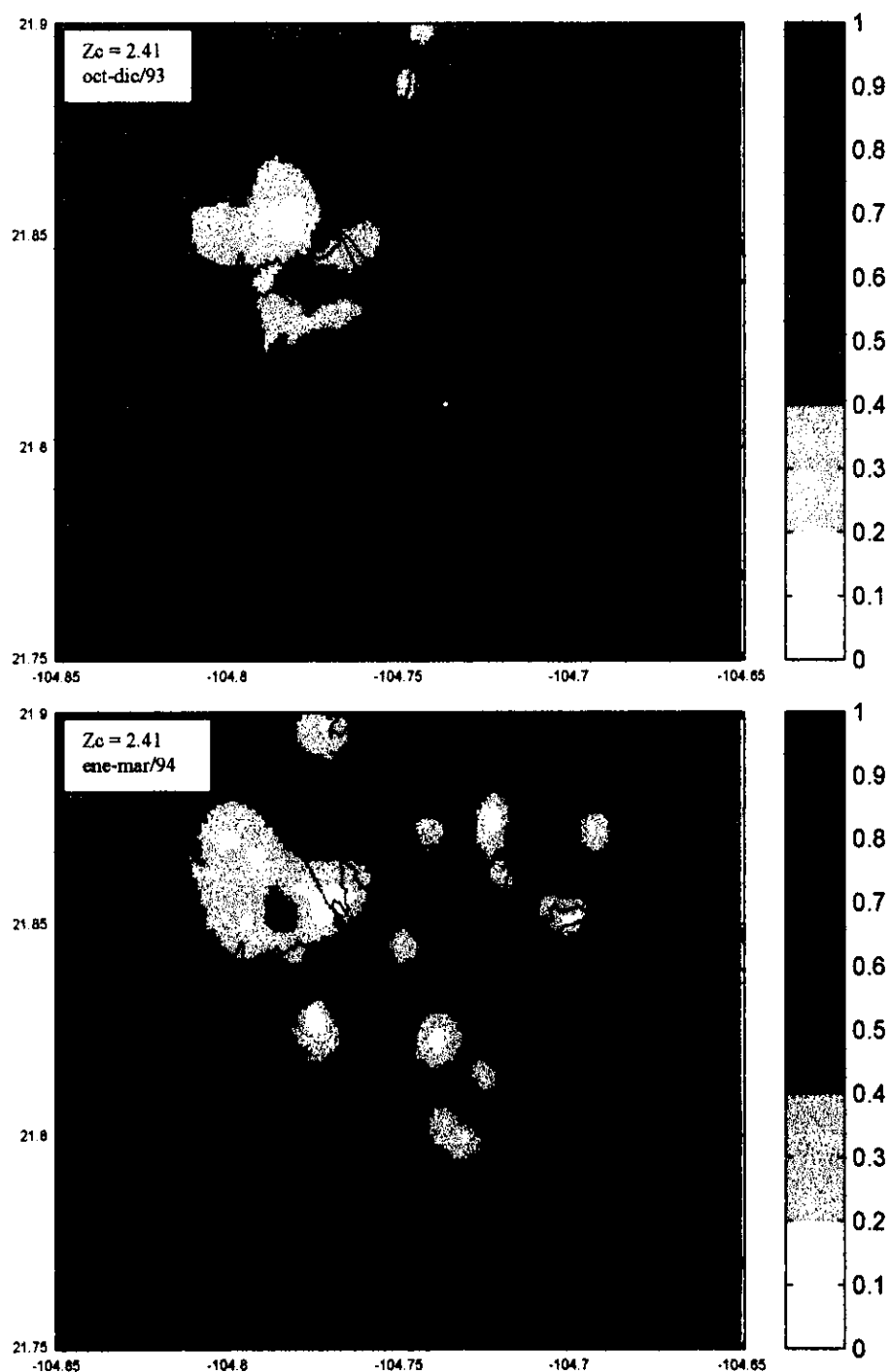


Figura 4.19. Probabilidad de exceder $z_c = 2.41$ obtenida al utilizar kriging indicador con información de octubre-diciembre de 1993 y enero-marzo de 1994. Es notable, al observar por ejemplo los valores de 0.6 y mayores, el cambio en el patrón de ocurrencia de la sismicidad mayor o igual a 2.41, de una sismicidad extendida (mapa superior) a una sismicidad reflejada en áreas más pequeñas (mapa inferior).

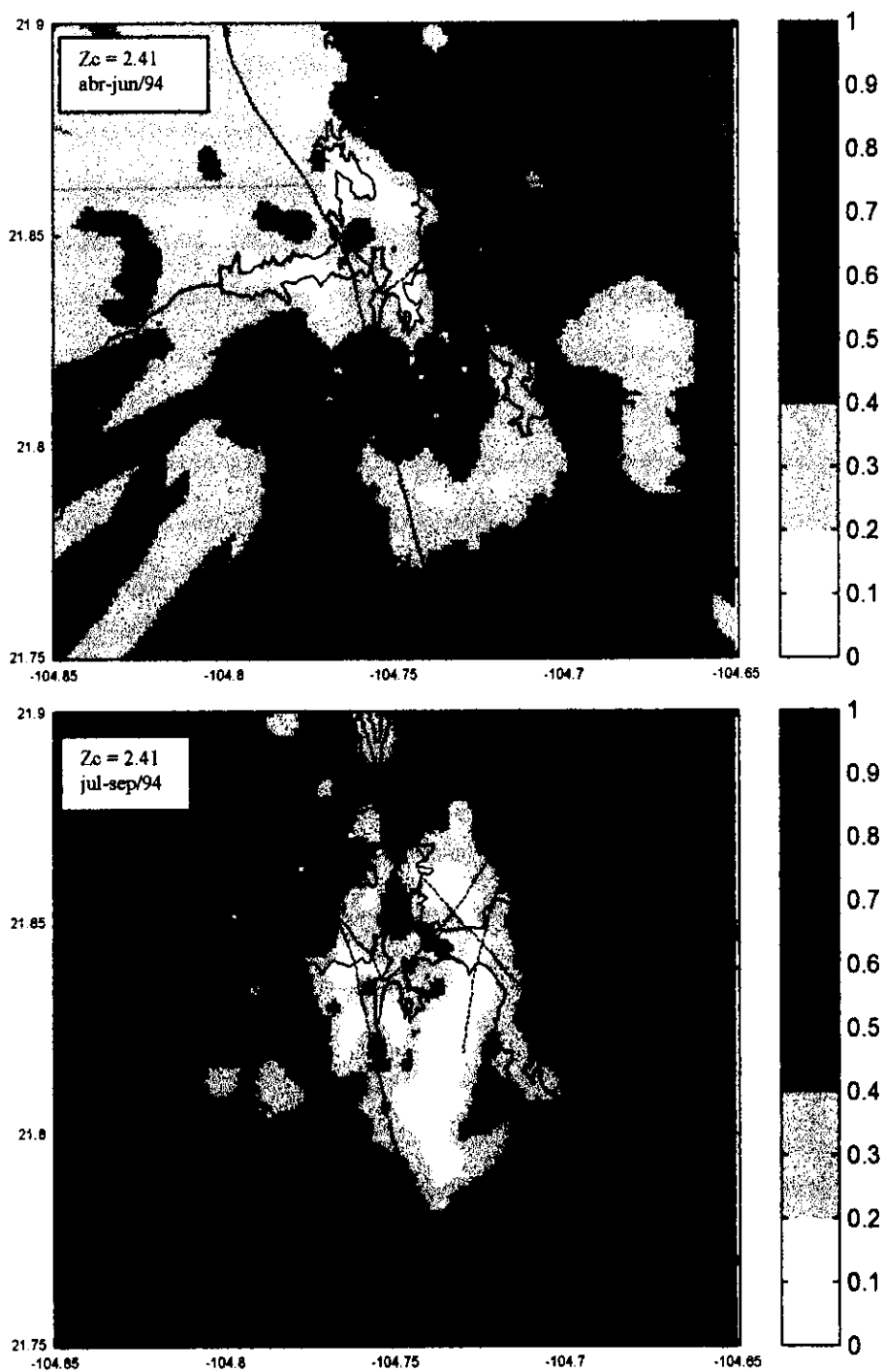


Figura 4.20. Resultados de kriging indicador, para $z_c = 2.41$, con información de abril-junio y julio-septiembre de 1994. Los valores altos de probabilidad (0.6 y mayores) reflejan que la sismicidad mayor a 2.41, si bien se manifiesta cerca del embalse (mapa superior), es notable también su presencia a mayor distancia de este último (mapa inferior).

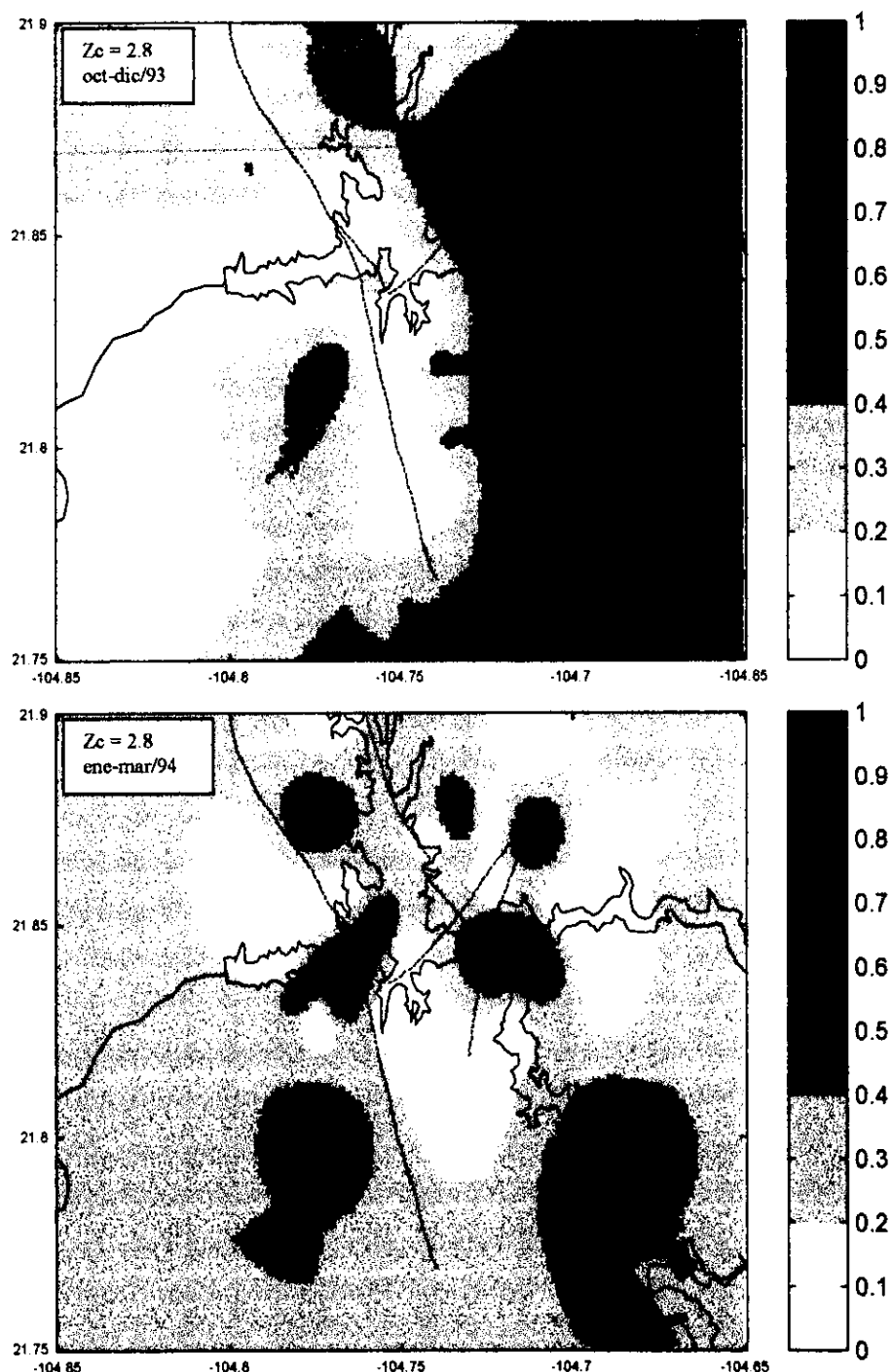


Figura 4.21. Probabilidad de exceder $z_c = 2.8$ obtenida al utilizar kriging indicador con información de octubre-diciembre de 1993 y enero-marzo de 1994. Observar el cambio (valores altos de probabilidad) en la distribución espacial de la sismicidad fuerte (mayor o igual a 2.8) así como la mejor definición en las áreas de igual probabilidad y la menor ocurrencia de artefactos.

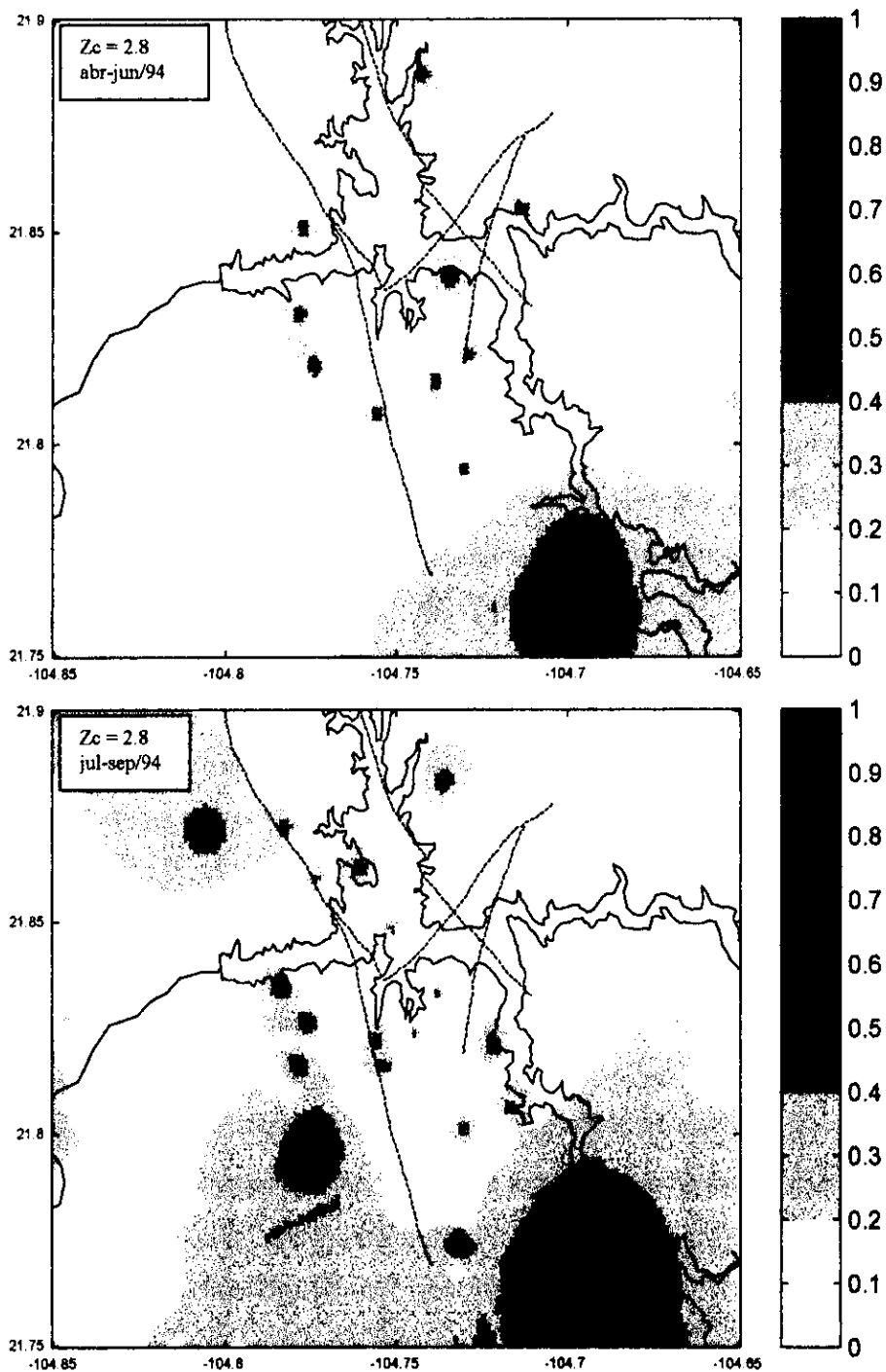


Figura 4.22. Probabilidad de exceder $z_c = 2.8$ en los periodos abril-junio y julio-septiembre de 1994. La sismicidad mayor o igual a 2.8 ocurre en zonas más pequeñas (valores altos de probabilidad) y relativamente más cercanas al embalse o a las fallas consideradas (mapa superior), para luego establecerse en áreas más grandes ubicadas a distancias relativamente mayores del embalse y de las fallas presentes en la zona analizada.

Al observar los mapas superiores de las figuras 4.19 y 4.21, en el período de octubre a diciembre de 1993, se refleja el mecanismo de disparo relacionado con el fuerte cambio en el nivel del agua a partir del llenado del embalse, ya que los valores altos de probabilidad de excedencia, para los dos umbrales, están cubriendo una amplia zona preferentemente hacia el Este del mapa y con tendencia Norte-Sur. Cabe recordar que para este período, la profundidad del embalse en la zona SE del mapa es mínimo, por lo que la sismicidad observada se asocia con la cercanía de focos termales en esta zona (Quijano y Chacón, 1985). A medida que pasa el tiempo, estos valores altos de probabilidad se restringen en área, acusando zonas sismogénicas que posteriormente, en los últimos trimestres, migran a distancias mayores del embalse, con excepción del caso de la zona SE. Las zonas sismogénicas están mayormente reflejadas en el mapa del umbral de 2.8 que en el de 2.41, ya que en este último las zonas de alta probabilidad son más amplias, mientras que en el mapa de 2.8 estas zonas son de menor área, lo cual sostiene que la ocurrencia de sismos mayores está asociada con estas zonas sismogénicas.

Los resultados obtenidos, al utilizar KI en el análisis de la información sísmica de cada ventana de tiempo y para cada uno de los umbrales considerados, son muy sensibles a la cantidad de datos abarcados en la estimación de la variable indicador, ya que como puede verse en las figuras de la 4.19 a la 4.22, se definen mejor las áreas de igual probabilidad en zonas de alta densidad de datos, mientras que en zonas de baja densidad de datos los resultados son muy suavizados y se tiene la tendencia a tener artefactos. Lo anterior puede ser visto al comparar los resultados del umbral $z_c = 2.8$ contra los del umbral $z_c = 2.41$, ya que al considerar $z_c = 2.8$, el número de datos abarcados en el análisis es mayor y por consiguiente los resultados están mejor definidos y los artefactos tienden a desaparecer.

Los mapas antes descritos dan una primera respuesta al problema de la distribución espacial de la magnitud sísmica sin consideraciones temporales, ya que en cada localidad estimada se tiene una función de densidad de probabilidad que permite conocer, entre otras cosas, cual sería el valor de magnitud más esperado en dicha localidad para el intervalo de tiempo analizado.

4.3 Simulación

En la estimación del modelo de incertidumbre alrededor de la interpolación de $Z(x)$, la técnica de KI genera para cada localidad a estimar un solo modelo, es decir, una función de distribución acumulada para cada localidad. En general, los mapas obtenidos con KI probabilístico tienen mayor variación en zonas donde hay mayor densidad de datos, y prácticamente no tienen variación en zonas con pocos datos. Las estimaciones representadas en estos mapas son obtenidas bajo el criterio de minimización del error de estimación local y pueden no reproducir patrones de continuidad espacial presentes en el área analizada (Journel, 1989). Es decir, el modelo de incertidumbre dado por la distribución condicional:

$$F(x; z_c | (n)) = \text{Prob}\{Z(x) \leq z_c | (n)\},$$

no provee información sobre la incertidumbre espacial conjunta, por ejemplo, la probabilidad de que una traza de puntos $x_j, j = 1, \dots, N$ sea valuada por encima de un cierto umbral z_c , lo cual puede caracterizar una fractura, falla o una trayectoria preferencial de sismicidad. Si este es el caso, las técnicas de simulación tienen propiedades que favorecen la solución de este problema, ya que al comparar con los algoritmos de interpolación es relevante resaltar lo siguiente:

- ◆ En la totalidad de los algoritmos de interpolación, el objetivo es obtener la mejor estimación local $z^*(x)$ en localidades no muestreadas sin considerar la estadística espacial resultante de las estimaciones $z^*(x), x \in A$. En simulación, las características globales como el histograma y el variograma son tan importantes como la exactitud local.
- ◆ Al utilizar kriging como algoritmo de interpolación, se obtiene un modelo numérico único de estimación considerado como óptimo en el sentido de variancia mínima, mientras que la simulación provee muchos modelos numéricos alternativos, que son representaciones del fenómeno analizado, a partir de la reproducción de su variabilidad espacial. Las diferencias entre estos modelos alternativos o realizaciones proveen una medida de la incertidumbre espacial conjunta.

El proceso de generar modelos alternativos igualmente probables de la distribución espacial de $Z(x)$, denotados como $\{Z^l(x), x \in A\}, l = 1, \dots, L$ donde L es el número total de realizaciones, se le conoce como simulación estocástica (Dutsch y Journel, 1992; Journel, 1989). Las características principales de estos conjuntos de valores simulados se resumen en:

- 1) Reproducción de la función de distribución acumulada univariada, comúnmente representada con el histograma de la información original (ver apéndice 3).

2) Reproducción de los valores de los datos en sus localidades respectivas:

$$z^l(x_a) = z(x_a), \text{ para todas las realizaciones } l = 1, \dots, L.$$

3) Reproducción de cualquier número K de covariancias indicadoras $F(z_k, z_k; h)$ del tipo (ver apéndice 3):

$$F(z, z'; h) = E[I(x; z) \cdot I(x+h; z')] = \text{Prob}\{Z(x) \leq z, Z(x+h) \leq z'\}.$$

4) Reproducción no sólo de los datos $z(x_a)$, sino también de datos locales tales como rangos de variación de $z(x_a) \in [a_a, b_a]$ y/o distribuciones de probabilidad a priori de los datos $z(x_a)$.

4.3.1 Simulación secuencial

Con base en las propiedades antes expuestas y dadas las características de la variabilidad espacial de la magnitud sísmica en la zona del embalse de la presa Aguamilpa, las técnicas de simulación resultan adecuadas para modelar de manera más realista esta sismicidad. En lo particular, debido a que hay suficiente cantidad de datos, se utilizó el algoritmo de simulación secuencial de la variable indicador. El algoritmo está basado en el principio de simulación secuencial y consiste en lo siguiente (Deutsch y Journel, 1992):

a) En cada nodo $x_j, j = 1, \dots, N$ a ser simulado a través de una trayectoria aleatoria dentro de una malla considerada, donde N es el número total de nodos, se estima, aplicando KI para K umbrales, una función de distribución acumulada condicional.

$$F^*(x_j; z_k | (n)) = \text{Prob}^*\{Z(x_j) \leq z_k | (n)\}, \quad k = 1, \dots, K.$$

b) Se interpola, para mayor definición y continuidad, la función de distribución acumulada condicional considerando todos los umbrales z_k y todo el rango de variación de los datos $z_a \in [z_{\min}, z_{\max}]$.

c) Se obtiene la simulación tipo Monte-Carlo de una realización $z^l(x_j)$, al seleccionar un número aleatorio uniforme $p^l \in [0, 1]$ y extraer el cuantil respectivo a p^l de la función de distribución acumulada condicional, es decir,

$$z^l(x_j) = F^{*-1}(x_j; p^l | (n)), \quad \text{tal que:} \quad F^*(x_j; z^l(x_j) | (n)) = p^l.$$

- d) El conjunto de datos indicadores, considerando cada uno de los umbrales z_k , es actualizado con el valor simulado $z^l(x_j)$, para luego aplicar KI en el siguiente nodo x_{j+1} a lo largo de la trayectoria aleatoria.

De esta manera, una vez que todas las localidades $x_j, j = 1, \dots, N$ han sido simuladas, se obtiene una imagen estocástica $\{ z^l(x_j), x_j \in A \}$ de la distribución espacial de $Z(x)$. Todo el proceso de simulación secuencial puede ser repetido a través de una nueva trayectoria aleatoria, y así obtener otra realización independiente tal que:

$$\{ z^{l'}(x_j), x_j \in A \}, l' \neq l.$$

En la figura 4.23, a manera de resaltar los puntos 1 y 3 de las características principales de los valores simulados ya mencionadas, se presentan las gráficas, para la información trimestral de octubre-diciembre de 1993 y enero-marzo de 1994, de la función de distribución acumulada de los datos y diez funciones de distribución acumulada de las respectivas simulaciones. Asociadas a estas gráficas y dentro de la misma figura están los modelos de variogramas (indicadores) de los datos y los variogramas experimentales de la variable indicador, para el umbral $z_c = 2.8$, de algunas de las simulaciones consideradas. La diferencia observada entre la función de distribución acumulada de los datos y las funciones de distribución acumulada de las simulaciones no es mayor del 5 %, lo que implica una adecuada reproducción de la estadística univariada en las simulaciones obtenidas. Las diferencias observadas entre los modelos de variograma indicador y los variogramas experimentales respectivos, para las dos ventanas de tiempo consideradas, se incrementan para valores de h mayores que el rango respectivo y son menores para valores de h dentro del rango de correlación, sin embargo, las diferencias observadas se presentan en proporciones similares tanto por arriba como por abajo del modelo de variograma respectivo. Lo anterior significa entre otras cosas, una reproducción adecuada de la distribución espacial de $Z(x)$ tanto en el sentido local como en el sentido global.

En las figuras 4.24 y 4.25 se muestran simulaciones de la distribución espacial de $Z(x)$, para las dos ventanas de tiempo mencionadas, con sus respectivas gráficas de datos asociados. Es notable en estas gráficas la reproducción de los valores de los datos (punto 2 de las propiedades de los valores simulados), y las características globales resultantes (textura) reflejando la tendencia espacial de los mismos datos. Al comparar estos mapas con los obtenidos con la técnica de KO (figuras 4.14 y 4.15), se puede observar que la simulación no genera resultados suavizados o que reflejen las tendencias locales, sin embargo, sí reflejan los cambios fuertes de la magnitud sísmica en áreas más restringidas y alineaciones o tendencias globales de la sismicidad, no obstante que también se generan valores de magnitud sísmica no esperados en áreas diversas, como puede apreciarse en las figuras mencionadas.

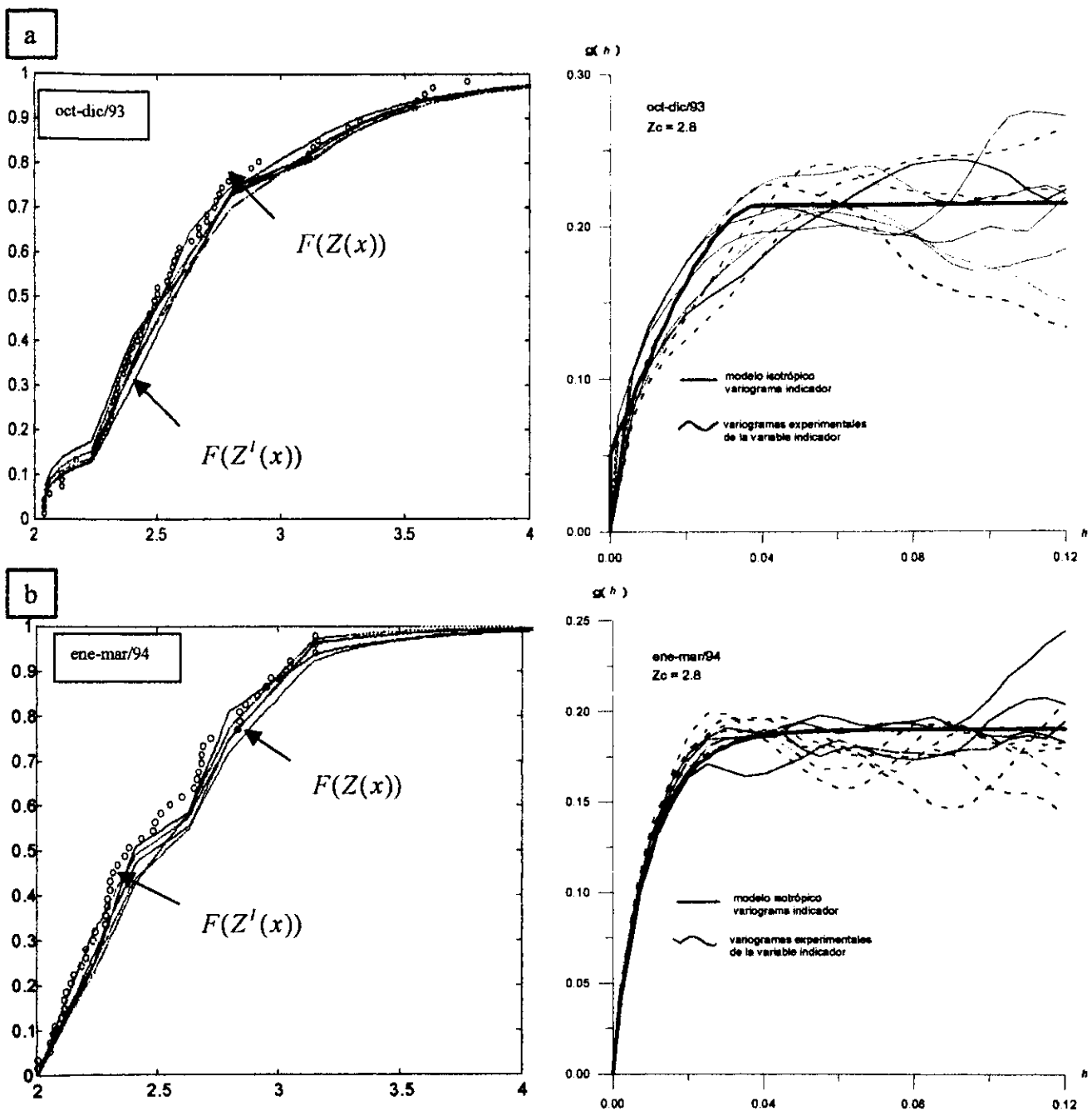


Figura 4.23. a) Funciones de distribución acumulada de los datos del periodo octubre-diciembre de 1993 y de los valores simulados $Z^l(x)$. La diferencia máxima, entre los valores de frecuencia acumulada, de los datos y las simulaciones es de 10% (para $Z_c = 2.8$), sin embargo, para todo el rango de variación de la magnitud, esta diferencia es de 5%. La gráfica superior derecha indica el modelo de variograma indicador, para el umbral y período mencionados, y los variogramas experimentales a partir de valores simulados. Las diferencias entre variogramas experimentales y el modelo son proporcionales, lo que implica una reproducción adecuada de la variabilidad espacial promedio de los datos. b) Las mismas observaciones hechas en las gráficas superiores son válidas para las del periodo enero-marzo de 1994.

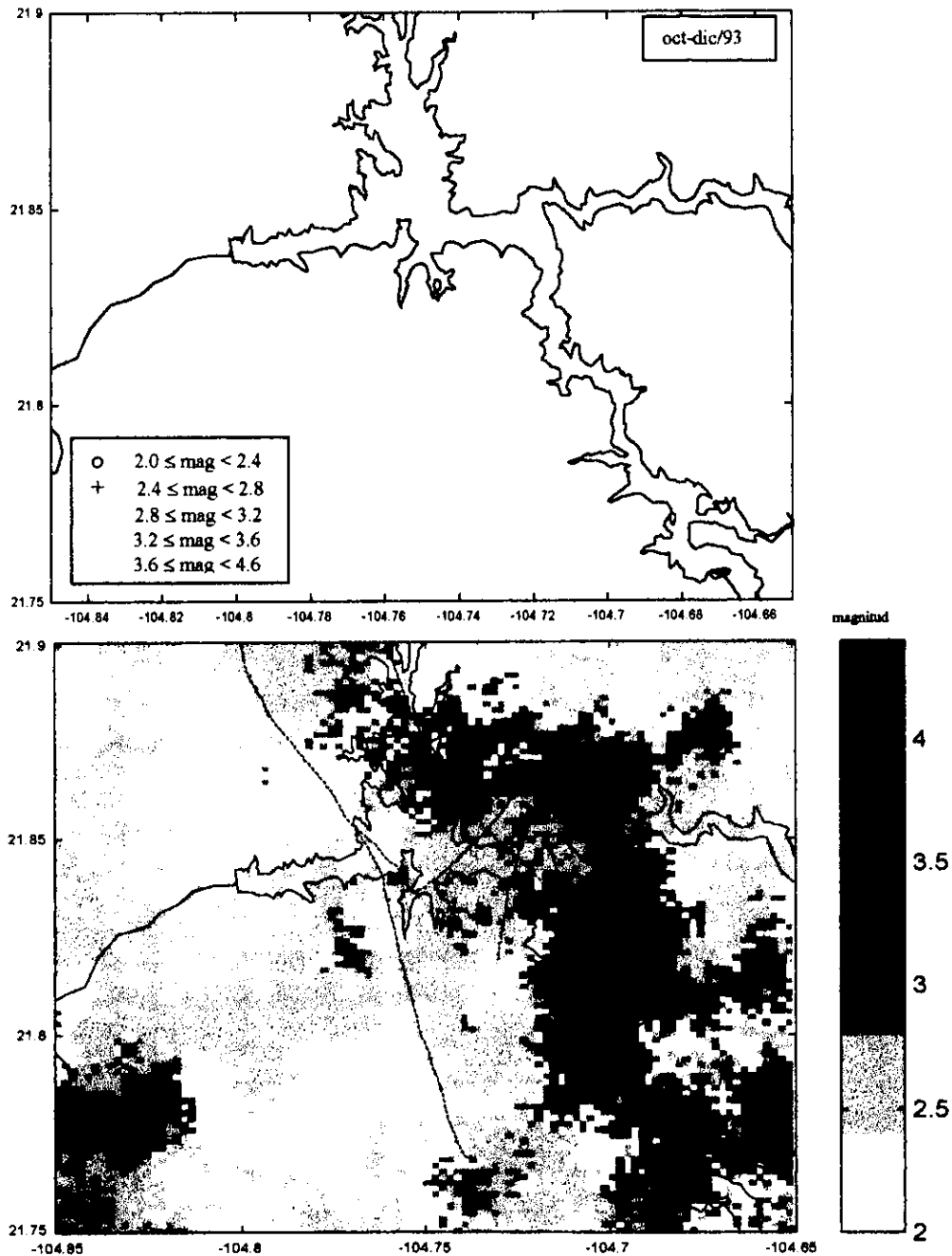


Figura 4.24. El mapa superior muestra la distribución espacial de la magnitud sísmica para el período octubre-diciembre de 1993, mientras que el mapa inferior es el resultado de la simulación estocástica de esta información. Se observa la reproducción de los valores de los datos y la influencia de éstos en los valores simulados en zonas cercanas a dichos datos. Asimismo, el resultado refleja: el patrón noroeste de la sismicidad mayor de 3.2, la sismicidad entre 2.4 y 3.2 en la porción sur-central del mapa y la sismicidad menor a 2.4 (zonas blancas). Los valores altos de la esquina sudoeste no son coherentes con los datos, generados por la imposición del histograma en el modelo, y corresponden a zonas de alta incertidumbre.

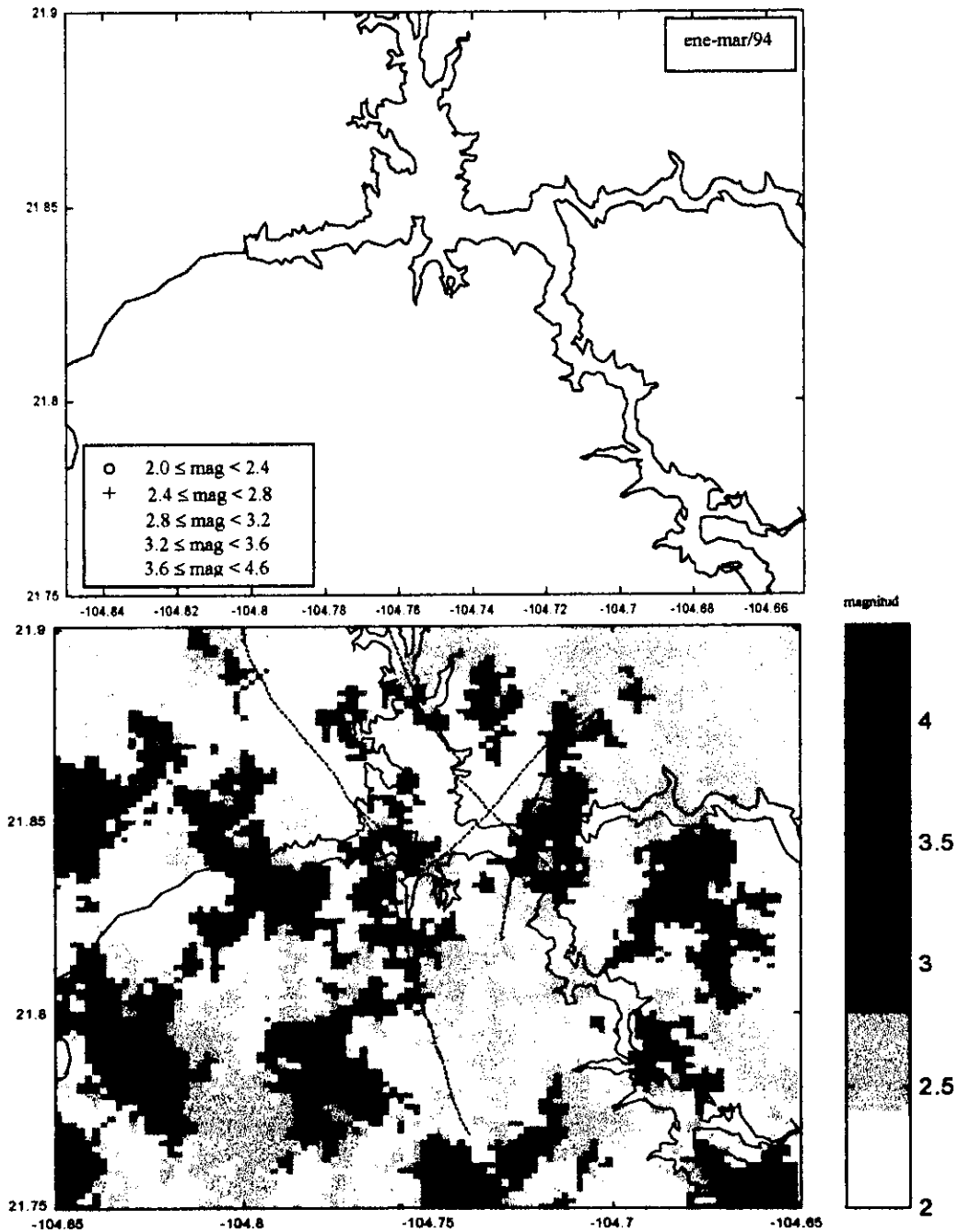


Figura 4.25. El mapa superior es la distribución espacial de la magnitud sísmica del período enero-marzo de 1994. El mapa inferior es la simulación estocástica de dicha información. El resultado refleja los cambios fuertes en la sismicidad en áreas pequeñas, como es la zona superior central en donde se tienen puntos negros (altos) junto con áreas blancas, reproduciendo el evento de magnitud 4.15 adyacente a sismicidad de $M \leq 2.4$. También se reconoce el patrón de sismicidad menor que 2.8 con dirección noroeste en la zona central del mapa (tonos claros). Se observan, también, valores no coherentes con los datos y que corresponden a zonas de alta incertidumbre.

4.3.2 Modelo de incertidumbre

Las simulaciones de la distribución espacial de $Z(x)$ por si mismas no tienen gran utilidad, sin embargo, es importante ratificar en estas imágenes las propiedades de reproducción de textura, estadística univariada global, variabilidad espacial y reproducción de datos. La parte relevante, para efectos de obtener modelos de incertidumbre espacial de $Z(x)$, es observar las diferencias entre todas las imágenes para cada grupo de datos.

Se simularon, con mallas de 100x100 puntos (medida arbitraria), 100 imágenes para cada ventana de tiempo. Bajo estas condiciones, considerando los cien valores asociados a cada localidad para cada grupo de datos, se obtuvieron funciones de distribución acumulada condicional para cada punto y sobre toda la malla, es decir,

$$F^*(x_j; z_k|(n)) = \text{Prob}^*\{Z(x_j) \leq z_k | (n)\}, \text{ para cada } k = 1, \dots, K \text{ y para cada } j = 1, \dots, N,$$

donde (n) representa el total de datos muestrados de la ventana de tiempo analizada, es decir, son los datos condicionantes. K es el número de umbrales considerados y N es el total de nodos simulados. A partir de esta información, se pueden generar mapas de probabilidad de excedencia de un determinado umbral de magnitud sísmica. Estos mapas, para cada período de análisis, dan respuesta a uno de los objetivos planteados en este trabajo.

En las figuras 4.26, 4.27, 4.28 y 4.29 se muestran, para dos umbrales, los mapas de probabilidad de excedencia de magnitud sísmica para cada una de las ventanas de tiempo. En estos mapas, los cuales son comparables con los de las figuras de la 4.19 a la 4.22 obtenidos con la técnica de KI, también son válidas las observaciones y comentarios hechos con respecto a la sismicidad y su relación con las fallas, fracturas y argumentos sobre sismicidad inducida. Sin embargo, cabe resaltar que los resultados de los modelos de incertidumbre obtenidos con simulación secuencial son menos suavizados que los obtenidos con la técnica de KI, más restringidos en área, representativos y coherentes con las tendencias generales de los datos (textura). En síntesis, para un fenómeno como la ocurrencia de temblores, donde se tiene una fuerte variabilidad en la magnitud sísmica a distancias cortas y a su vez, se pueden presentar tendencias regionales de sismicidad, el enfoque geoestadístico de la simulación estocástica resultó ser más adecuado y representativo, para la estimación de la distribución espacial de la magnitud sísmica y la incertidumbre asociada, que el enfoque de KI.

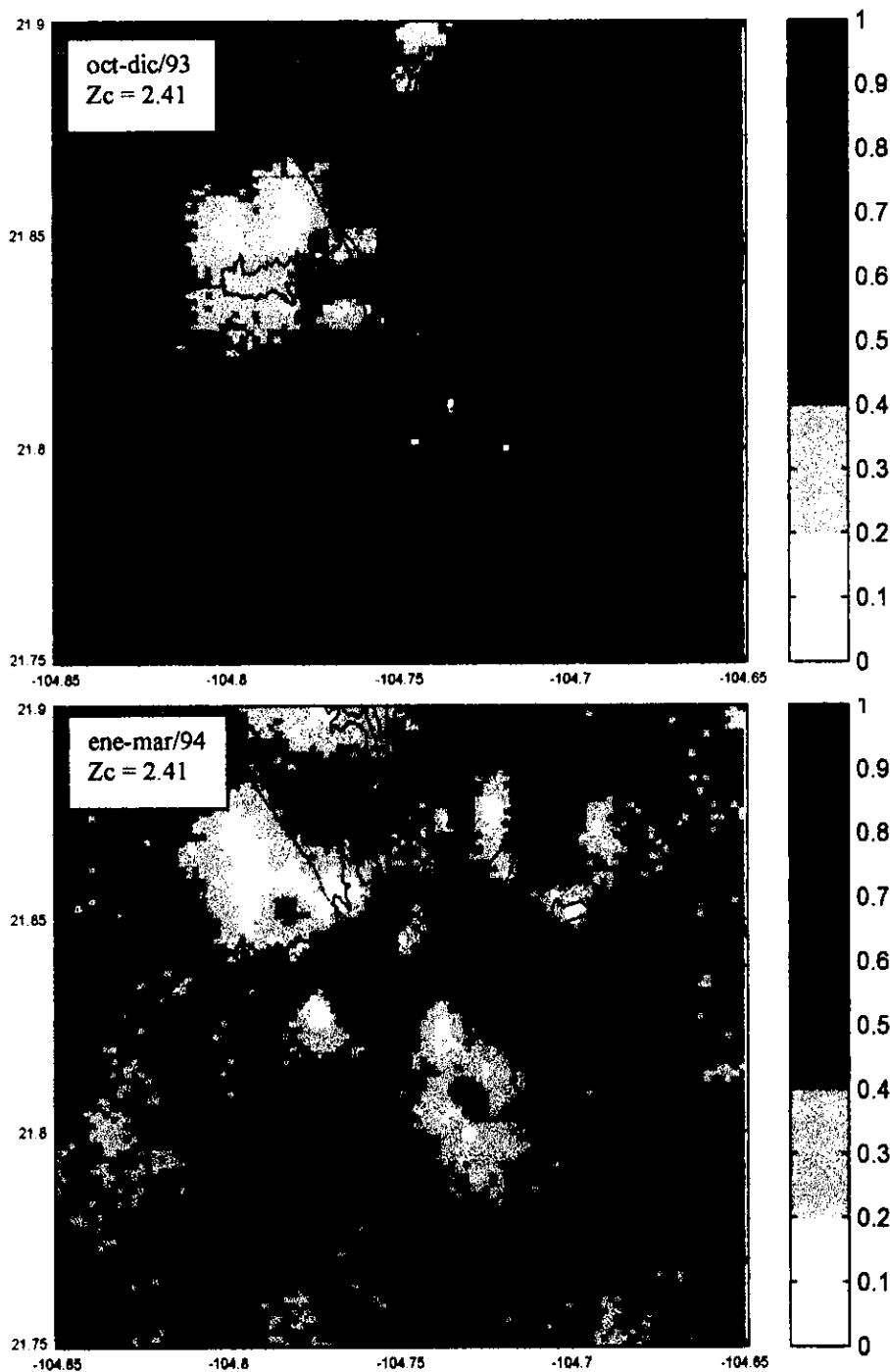


Figura 4.26. Probabilidad de exceder la magnitud 2.41 para dos ventanas de tiempo. En el período octubre-diciembre de 1993, la sismicidad por arriba de 2,41 se distribuyó en una amplia zona al Este del área considerada, limitada al Oeste por la falla El Sauz. En el período siguiente, esta sismicidad se manifestó en áreas más pequeñas, cercanas a la falla y al embalse. La sismicidad menor a 2,41 (tonos claros) es constante en la zona NO, sin embargo, en el segundo período se manifestó en la parte sur-central del área analizada.

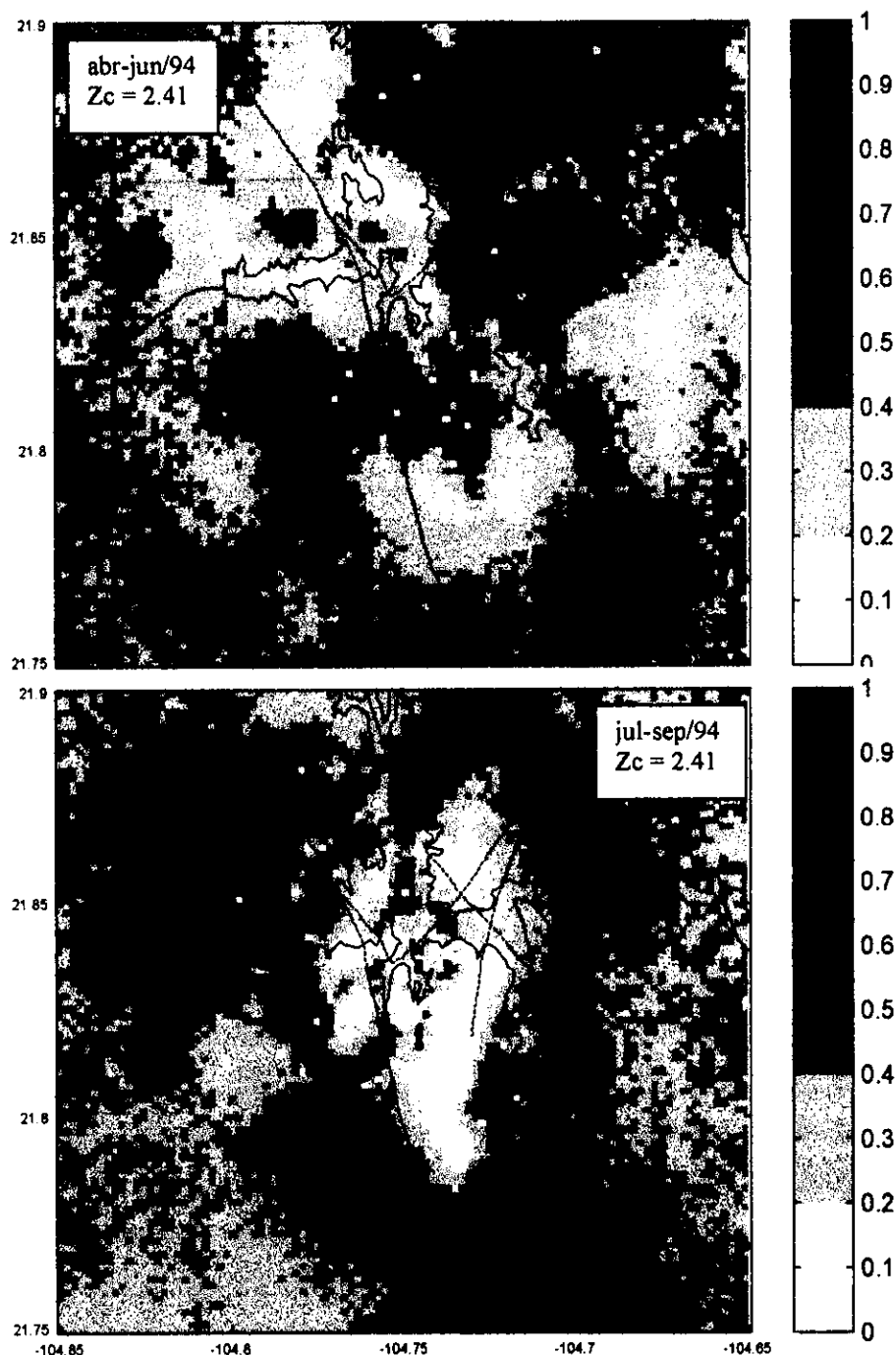


Figura 4.27. Probabilidad de exceder la magnitud 2,41 para dos ventanas de tiempo. En el mapa superior, abril-junio de 1994, se observa que la sismicidad mayor a 2,41 difiere de la del período anterior en cuanto a que hay presencia de ésta en la zona norte del embalse, desapareciendo en la zona central del mismo, y dando lugar a sismicidad menor a 2,41 al Norte y Sur del graben limitado por las fallas El Sauz y Rosario Viejo. Esta tendencia se ve acentuada en el siguiente período (mapa inferior) con nueva presencia en la parte NO y menor presencia en la zona NE:

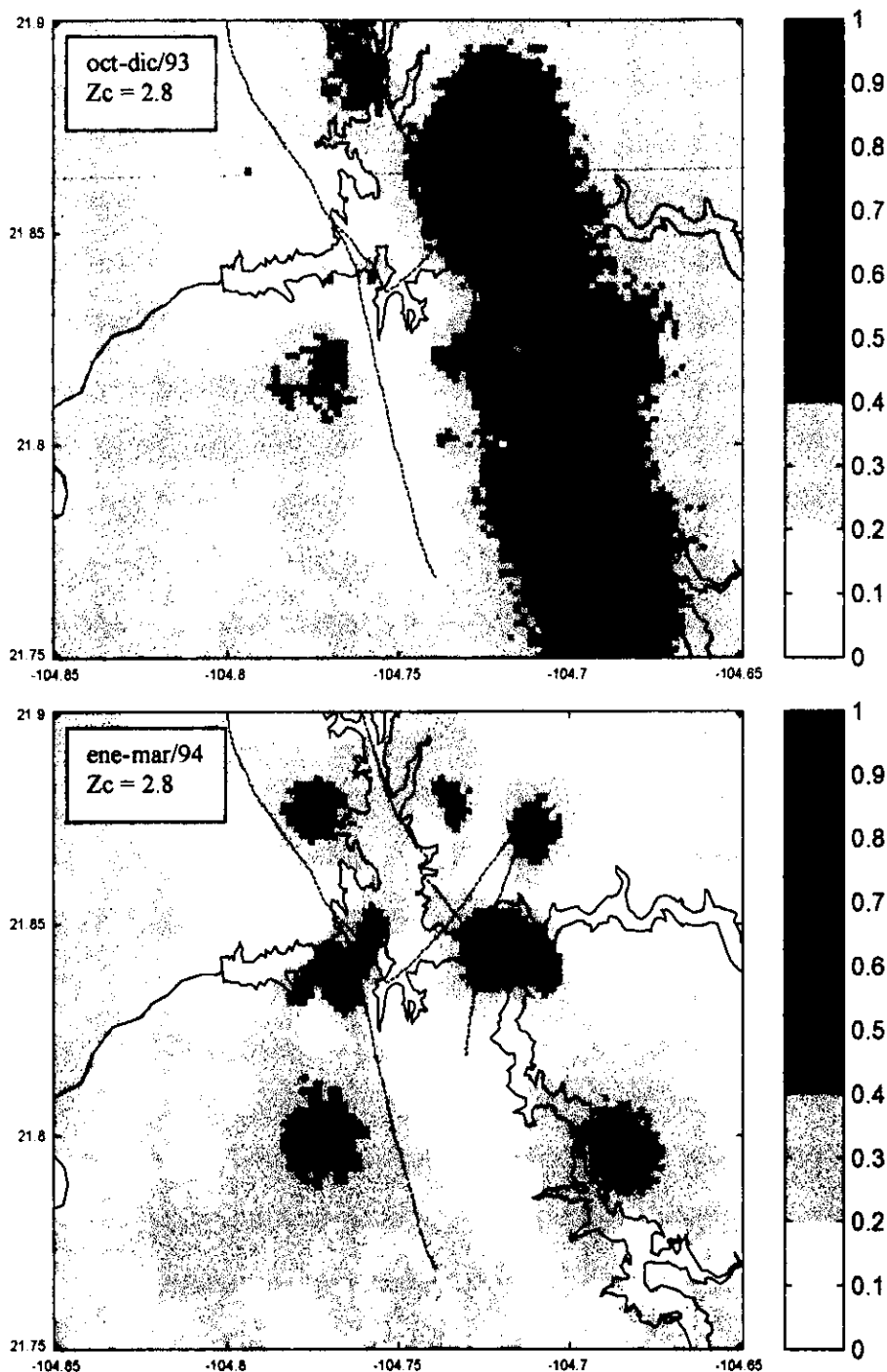


Figura 4.28. Probabilidad de exceder la magnitud 2,8 para los periodos octubre-diciembre de 1993 y enero-marzo de 1994. En el mapa superior se observa la tendencia regional NO de esta sismicidad con concentraciones sobre una parte del embalse y en la zona de fracturas. La sismicidad menor a 2,8 se manifiesta del centro del graben hacia el Oeste. En el siguiente período, se observa también la tendencia NO pero las concentraciones son en áreas menores y cercanas al embalse, fallas y fracturas consideradas.

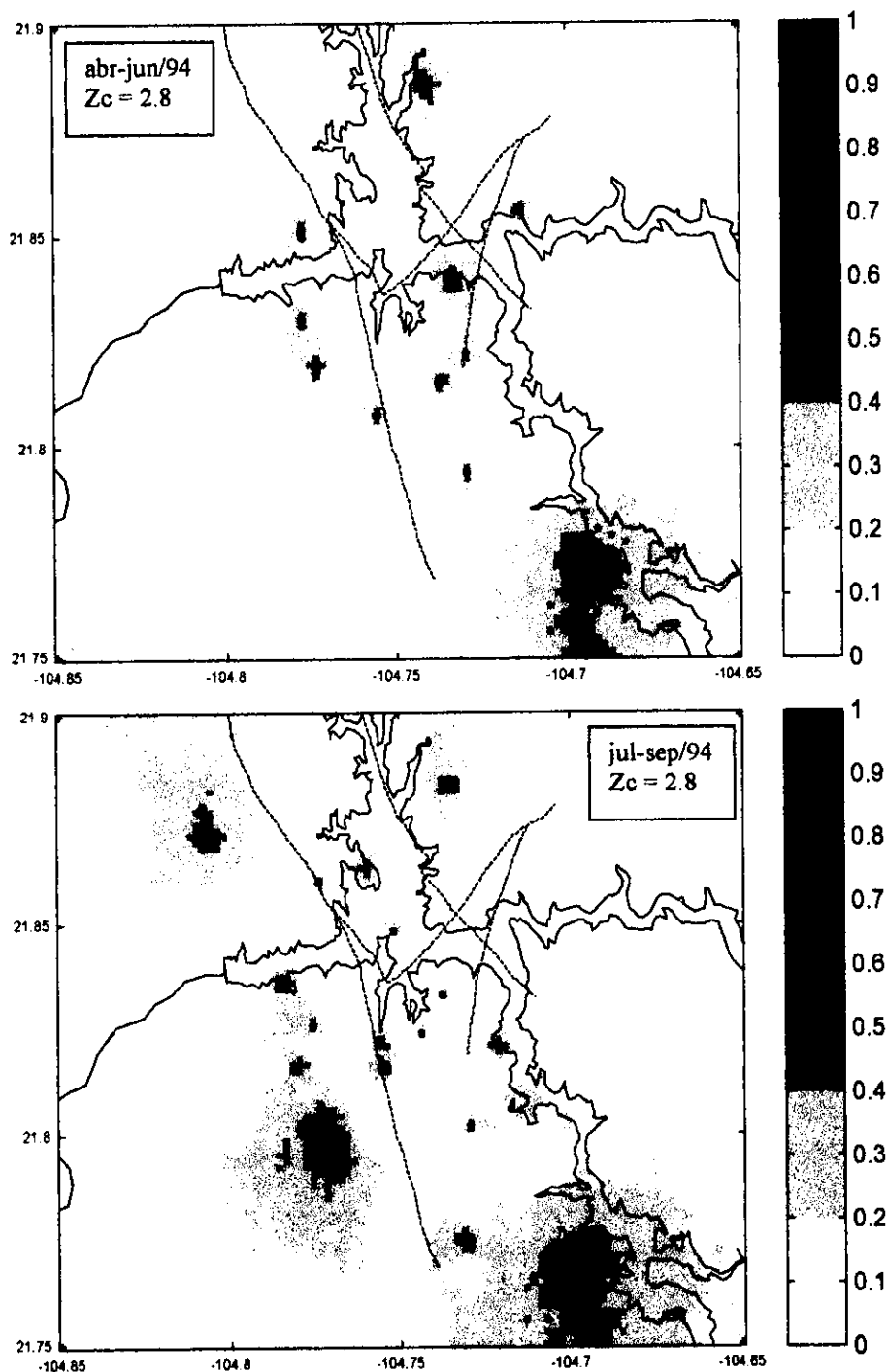


Figura 4.29. Probabilidad de exceder la magnitud 2.8 durante abril-junio y julio-septiembre de 1994. En el primer período esta sismicidad se manifestó en áreas menores a las del período anterior, aunque persiste la cercanía a fallas y fracturas y la tendencia NO. Para el siguiente período (mapa inferior), se mantiene la tendencia NO, con áreas de influencia mayores y hacia el Oeste de la falla El Sauz. Esta sismicidad, en la zona SE del embalse, tiene mínima variación durante los cuatro periodos analizados.

5 Investigaciones futuras

Este trabajo ha estado dirigido a predecir espacialmente la variabilidad de la magnitud sísmica, y también ha establecido un procedimiento para determinar la probabilidad, de que en cierta locación, la magnitud sísmica exceda cierto umbral. Además, este modelo se ha fundamentado considerando estacionariedad hasta de segundo orden en el espacio, en cada una de las ventanas seleccionadas. La estacionariedad en el tiempo, aunque importante, es un tópico relacionado con la predicción temporal que no ha estado dentro del alcance de este trabajo. Sin embargo, considerando su importancia y dentro de las actividades de investigación a continuar en el futuro, un esquema de solución es presentado a continuación.

Supongamos que cada ventana de tiempo es identificada con el atributo t_k . Así, si hay K ventanas de tiempo, estas son denominadas t_1, t_2, \dots, t_K . La magnitud sísmica entonces puede ser modelada como $Z(x, t_k)$ para cada k ventana de tiempo. Ahora, se ha establecido que es razonable que:

$$E[Z(x, t_k)] = m_k,$$

es decir, el valor esperado es constante para toda x en cada ventana de tiempo, pero entre ventanas de tiempo, el valor esperado puede ser diferente. Así, el problema a resolver es la siguiente predicción temporal:

$$Z(x, t_{k+1}) = m_{k+1} + \sum_{k=1}^K \sum_{\alpha=1}^{n_k} \lambda_{k\alpha} [Z(x_\alpha, t_k) - m_k]. \quad \dots\dots\dots (19)$$

En esta expresión, claramente los ponderadores λ son desconocidos y también el valor esperado m_{k+1} . Este último elemento, más adelante será discutido en cuanto a como establecerlo, aunque cabe mencionar que en la expresión referida también destaca que la doble sumatoria no hace mención de información en el futuro, esto es, de t_{k+1} . La información consiste de la existente y que ha ocurrido hasta ahora, desde t_1 hasta t_k . Asimismo, también puede apreciarse que la expresión (19) no es otra que la correspondiente a cokriging simple, cuya solución es obtenida a través de los criterios de insesgamiento y mínima variancia, resultando en:

$$\sum_{l=1}^K \sum_{\beta=1}^{n_l} \lambda_{l\beta} \text{Cov}(Z(x_\alpha, t_k), Z(x_\beta, t_l)) = \text{Cov}(Z(x_\alpha, t_k), Z(x, t_{k+1})) \quad \begin{matrix} k = 1, \dots, K \\ \alpha = 1, \dots, n_k \end{matrix} \quad \dots\dots\dots (20)$$

El lado izquierdo de la ecuación presenta la información estadística, en la forma de autocovariancias y covariancias cruzadas, de la magnitud sísmica entre las K ventanas de

tiempo. El lado derecho de la ecuación es la covariancia cruzada entre las K ventanas de tiempo y la ventana donde se efectuará la predicción espacial en x y la temporal en $K+1$.

Como se puede apreciar, el sistema de ecuaciones resultante contiene en la matriz de covariancias, que corresponde a aquella que se obtiene de resolver el problema de cokriging simple, autocovariancias definidas como:

$$Cov(h; t_k, t_k) = E[(Z(x+h, t_k) - m_k)(Z(x, t_k) - m_k)],$$

las cuales son calculadas de los variogramas utilizados en la predicción espacial, o pueden ser derivadas si así es deseado, de las simulaciones estocásticas obtenidas con anterioridad. Cualesquiera que sea la decisión, estas covariancias son obtenidas de los datos medidos y la inferencia de ellas es razonablemente obtenida. Sin embargo, otro de los términos en el sistema de ecuaciones corresponde a covariancias cruzadas de la forma siguiente:

$$Cov(h; t_k, t_{k'}) = E[(Z(x+h, t_k) - m_k)(Z(x, t_{k'}) - m_{k'})]. \quad \dots\dots\dots (21)$$

Estas expresiones, considerando la escasez de información en las ventanas de tiempo, se propone sean obtenidas a través de las simulaciones estocásticas realizadas en cada ventana de tiempo. El procedimiento es calcular las covariancias cruzadas para cada h y para cada pareja de ventanas de tiempo k y k' . Así, se pueden obtener diferentes valores del producto $(Z'_s(x+h, t_k) - m_k)(Z'_s(x, t_{k'}) - m_{k'})$, donde la notación s indica que son valores simulados e i, j que corresponden a realizaciones diferentes en las ventanas de tiempo k, k' . El número de estos productos es el producto del número de realizaciones de la ventana k por el número de realizaciones de la ventana de tiempo k' , de tal suerte que el proceso para obtener (21) asegura su inferencia cuando k o k' corresponden a ventanas de tiempo en donde se hayan efectuado simulaciones estocásticas. Si el número de realizaciones para k es 50 y para k' también, el número de productos para calcular (21) es 2,500. Esta cantidad permite una robusta estimación de la covariancia cruzada.

Conviene aclarar con respecto a las covariancias cruzadas que éstas pueden ser no estacionarias considerando que para cada h están siendo calculadas. Esto relaja la condición de estacionariedad de segundo orden, al no depender las covariancias de h y si depender de la posición x . Sin embargo, esto también implica que no existe garantía de que el sistema de ecuaciones tenga solución, ya que las covariancias cruzadas no han sido modeladas con funciones positivo definidas (Journel, 1989) debido a que son obtenidas en un proceso automático. Si se quiere modelarlas, se tendrá que aceptar estacionariedad de segundo orden en espacio y tiempo.

6 **Discusión y conclusiones**

Dadas las condiciones de la información y el tipo de ésta, sobresale lo siguiente:

- ◆ La base de datos utilizada en el análisis geoestadístico fue formada a partir de catálogos sísmicos que corresponden a dos diferentes campañas de medición (redes sísmicas). Por ello, el muestreo del fenómeno no se realizó bajo las mismas condiciones, sugiriendo un posible sesgo en la información. Sin embargo, el área de una de estas redes está contenida en la otra y ambas capturaron información durante el período analizado. Los datos fueron filtrados a una magnitud mínima de 2.0, con objeto de homogeneizar la base de datos, además de que las técnicas geoestadísticas reconocen, al utilizar modelos de variabilidad espacial, que los datos geográficamente están dispersos.
- ◆ La sismicidad en regiones alrededor de presas o lagos artificiales es un fenómeno controlado por las condiciones geológico-estructurales de la zona afectada. El grado de fracturamiento, presencia de fallas y tipos de roca son condiciones que determinan la difusión de la presión de poro generada por el fluido infiltrado. Los efectos mecánicos de la presión de poro determinan el patrón espacial y temporal de lo que se denomina sismicidad inducida por presas. Sin embargo, la sismicidad real observada en estas regiones puede estar influenciada por el efecto químico del agua, corrosión, sobre el material de relleno en fracturas o fallas preexistentes (Talwani y Acree, 1985). Por lo tanto, el modelado de la distribución espacial de la sismicidad inducida por presas o lagos artificiales, deberá incluir variables geológicas como presencia de fallas o fracturas, tipos de roca y/o variables geohidrológicas relacionadas con la presión de poro, como la compresibilidad del fluido, porosidad de la roca, difusividad hidráulica (Talwani y Acree, 1985).

Una vez establecido lo anterior, es evidente la característica multivariada del fenómeno. Sin embargo, hay que recordar que las técnicas geoestadísticas permiten, como ha sido mencionado en este trabajo, el análisis y modelado de la distribución espacial de fenómenos multivariados como el analizado, que requiere de información suficiente y confiable para su aplicación. Este argumento da el marco de esta investigación, con base en la información disponible y el análisis geoestadístico. Este enfoque ha permitido determinar que:

- 1) El análisis exploratorio univariado aplicado en la información sísmica disponible correspondiente al período de octubre de 1993 a octubre de 1998, de la zona alrededor del embalse de la presa Aguamilpa, Nay., validó que los datos correspondientes a los años de 1993 y 1994 son representativos de toda la información considerada, mientras que los datos de 1995 a 1998 no son representativos de la información general. Por ello, la información de 1993 a 1994 constituyó la base de datos para el análisis geoestadístico.

- 2) Los histogramas de datos agrupados en ventanas de tiempo semestrales, y aún trimestrales, a partir de la base de datos definitiva, tienen la forma del histograma general. Asimismo, utilizando el análisis del valor de b (Gutenberg y Richter, 1954) se reconocieron los cambios en la tasa de sismicidad, observados en la gráfica de la evolución histórica del número de sismos acumulado. Lo anterior definió el tamaño mínimo de la ventana de tiempo representativa de la información general, permitiendo explorar los cambios espacio-temporales de la sismicidad observada en la región de interés.
- 3) El análisis de la variabilidad espacial de la magnitud sísmica, considerando ventanas de tiempo trimestrales, probó que la magnitud sísmica, para el caso analizado, puede ser modelada como una variable regionalizada estacionaria. Además, la magnitud sísmica presentó mayor continuidad en direcciones con tendencia N-S que en las de tendencia E-O y, los rangos de correlación de los modelos de variogramas obtenidos, para información trimestral y para todo el período de la base de datos, fueron mayores a medida que avanzó el tiempo, es decir, la sismicidad cubrió áreas mayores con el tiempo sin dejar de comportarse como variable regionalizada.
- 4) Los variogramas de $Z(x)$, magnitud sísmica, mostraron tendencias de acuerdo a lo observado en la sismicidad de la zona. Sin embargo, para los variogramas de $I(x; z_c)$ considerando el umbral $z_c = 2.23$ y dado que la información fue acotada en $z_{\min} = 2.0$, se redujo el número de datos $z(x_a) \leq 2.23$, por lo que los variogramas para el umbral referido tuvieron información escasa. Sin embargo, esta situación no afectó el análisis de la incertidumbre para umbrales mayores.
- 5) La estimación de la distribución espacial de la magnitud sísmica a partir de la técnica de KO, para los diferentes periodos considerados, resaltó las tendencias locales de los datos y la compatibilidad de los modelos de variograma respectivos. La estimación de $Z(x)$ a partir de la simulación estocástica, reflejó tanto las tendencias locales como las regionales. Por otro lado, en la estimación de $Z(x)$ con KO, en zonas de baja densidad de datos, se generaron artefactos, los cuales no corresponden a tendencia alguna de la información. La utilidad de los mapas de estimación, con cualquiera de las dos técnicas, sólo es en el sentido de observar tendencias locales y regionales del fenómeno.
- 6) Los mapas de probabilidad de excedencia de un determinado umbral, obtenidos a partir de los modelos de incertidumbre, son más suavizados en el caso de KO que los resultados de la simulación estocástica. Estos últimos reflejaron tanto tendencias locales (cambios fuertes de la magnitud sísmica a distancias pequeñas) como regionales (mayor continuidad de la magnitud sísmica en direcciones con tendencia N-S). Por lo tanto, la técnica de simulación estocástica de la variable indicador, resultó ser adecuada para modelar la distribución espacial de la magnitud sísmica en la zona del embalse de la presa Aguamilpa, Nayarit, y permitió hacer conjeturas sobre la evolución en tiempo de los esfuerzos actuantes en el área analizada, lo cual, por supuesto, no fue materia de discusión en este trabajo.

- 7) Los resultados derivados del modelo de incertidumbre, obtenido a partir de la técnica de simulación estocástica de la variable indicador, son importantes en la definición y análisis del peligro sísmico y el riesgo asociado. Estos mapas de probabilidad de excedencia de un determinado umbral de magnitud sísmica, permitirían establecer diferentes escenarios de peligro sísmico, para cada ventana de tiempo y considerando diferentes umbrales, y por lo tanto evaluar las posibles consecuencias asociadas a cada escenario, es decir, intensidades de movimiento (aceleración máxima del terreno) para toda el área considerada. Lo anterior, sería información relevante para la toma de decisiones en la planeación y diseño de nuevas obras civiles en la región analizada.
- 8) Variables como son la profundidad del evento sísmico, tipo de roca, presencia de falla o fractura y la distancia de éstas a los puntos de interés, son importantes en la estimación de la distribución espacial de la magnitud sísmica y la incertidumbre asociada a este proceso. Por lo tanto, estas variables deberán considerarse en investigaciones futuras. Técnicas geoestadísticas como cokriging o cosimulación, permitirían abordar el problema considerando las diferentes variables involucradas.

7 Bibliografía

- Baecher, G. B. y R. L. Kenney, 1992. *Statistical examination of reservoir induced seismicity*, Bull. Seism. Soc. Am., V. 72, pp. 553-569.
- Carder, D. S., 1945. *Seismicity investigations in the Boulder Dam area, 1940-1944 and influence of reservoir loading on earthquake activity*, Bull. Seism. Soc. Am., V. 35, pp. 175-192.
- Carr, J. R., 1983. *Application of the theory of regionalized variables to earthquake parametric estimation and simulation*, disertación doctoral no publicada, Universidad de Arizona, 259 p.
- Carr, J. R. y C. E. Glass, 1984. *A regionalized variables model for seismic hazard assessment*, eighth world conference on earthquake engineering, Prentice Hall, Englewood Cliffs, New Jersey, V.1, pp. 207-213.
- Carr, J. R. y R. E. Bailey, 1985. *An indicator kriging model for the investigation of seismic hazard*, Mathematical Geology, V.18, N° 4, pp. 409-428.
- Carr, J. R., 1996. *Revisiting the characterization of seismic hazard using geostatistics: a perspective after the 1994 Northridge, California Earthquake*, Geostatistics for environmental and geotechnical applications, American Society for Testing and Materials, 1996.
- Chang, B., 1989. *Preliminary study on the prediction of reservoir earthquakes*, Earthquake Research in China, V.3, N° 2, pp. 201-216.
- Chen, X. y X. Yin, 1996. *Variation of the load/unload response ratio before the main shocks of the reservoir induced earthquakes*, Earthquake Research in China, V.10, N° 2, pp. 141-149.
- Cornell, C.A., 1968. *Engineering seismic risk analysis*, Bull. Seis. Soc. Am., V.58, pp. 1583-1606.

Council of the National Seismic System, 1996. *Global seismic catalog*, www.quake.geo.berkeley.edu

Cromer, V. M., 1996. *Geostatistics for environmental and geotechnical applications: a technology transfered*, Geostatistics for environmental and geotechnical applications, American Society for Testing and Materials, 1996.

De Rubeis, V., Gasparini, C. y P. Tosi, 1993. *Statistical comparison of geophysical parameters and seismicity in central Italy*, Natural Hazards, V.7, pp. 25-39.

Delgado, V. M.A., 1992. *Sexto informe técnico sobre monitoreo sísmico del Río Santiago*, Departamento de Sismotectónica de la Gerencia de Estudios de Ingeniería Civil de la Comisión Federal de Electricidad, reporte interno, 69 p.

Delgado, V. M.A., Vázquez, R. R., Sánchez, A. E. y Malagón, M. A., 1998. *Informe del enjambre sísmico del 28 de octubre de 1998 asociado a la C.H. Aguamilpa*, Departamento de Sismotectónica de la Gerencia de Estudios de Ingeniería Civil de la Comisión Federal de Electricidad, reporte interno, 38 p.

Delgado, V. M.A., Vázquez, R. R., Sánchez, A. E. y Malagón, M. A., 1999. *Reconsideración sismotectónica en la región de la C.H. Aguamilpa y Santa Rosa, Estados de Nayarit y Jalisco*, Departamento de Sismotectónica de la Gerencia de Estudios de Ingeniería Civil de la Comisión Federal de Electricidad, reporte interno inédito, 60 p.

Deutsch, C. V. y A. G. Journel, 1992. *GSLIB: Geostatistical software library and user's guide*, Oxford University Press, New York, 339 p.

Fu, Z., Jiang, L. y G. Li, 1995. *Analysis of the time-space-intensity distribution of earthquakes-induced life losses*, Journal of Earthquake Prediction Research, V.4, N° 4, pp. 61-470.

Glass, C. E., 1978. *Application of regionalized variables to microzonation*, Proc. 2nd international conference on microzonation for safer construction - research and application, V.1, pp. 509-521.

- Gough, D. I., 1978. *Induced seismicity*, The assessment of mitigation of earthquake risk, UNESCO, 341.
- Guha, S. K. y D. N. Patil, 1990. *Large water reservoir related induced seismicity*, Geophysik, V.99, pp. 3-5.
- Gupta, H. K., 1992. *Reservoir induced earthquakes*, Current Science, V.62, pp. 183-192.
- Gupta, H. K., Rastogi, B. K. y M. Narain, 1972. *Common features of the reservoir associated seismic activities*, Bull. Seismol. Soc. Amer., V.62, pp. 481-492.
- Gutenberg, B. y C. F. Richter, 1954. *Seismicity of the Earth and associated phenomena*, Princeton, N. J., Princeton University Press.
- Habibagahi, G., 1998. *Reservoir induced earthquakes analyzed via radial basis function networks*, Soil Dynamics and Earthquake Engineering, V.17, N°1, pp. 53-56.
- Isaaks, E. H. y R. M. Srivastava, 1989. *An introduction to applied geostatistics*, Oxford University Press, New York, 561 p.
- Journel, A. G., 1989. *Fundamentals of geostatistics in five lessons*, V.8, Short course in geology, American Geophysical Union, Washington, D.C., 40 p.
- Journel, A. y C. Huijbregts, 1978. *Mining Geostatistics*, Academic Press, 600 p.
- Lomas, D. E., 1995. *Sismicidad inducida en la C.H. Aguamilpa, Estado de Nayarit, de marzo a septiembre de 1994*, Oficina de Sismología de la Gerencia de Ingeniería Experimental y Control de la Comisión Federal de Electricidad, reporte interno, 10 p.
- Matheron, G., 1963. *Principles of geostatistics*, Economic Geology, V.58, pp. 1246-1266.
- Moreno, G. S. y G. M.A. Soto, 1994. *Monografía Geológica, Proyecto Hidroeléctrico Aguamilpa, Estado de Nayarit*, Gerencia de Estudios de Ingeniería Civil de la

Comisión Federal de Electricidad, reporte interno G-28, 289 p.

Musson, R.M. W. y P. W. Winter, 1993. *Seismic hazard methodology for a hazard map of the U.K.*, AEA technology report N° GNSR(DTI), 149 p.

Ortega, F., Mitre, L. M., Roldán, J., Aranda, J., Morán, D., Alaniz, S. y A. Nieto, 1992. *Texto explicativo de la quinta edición de la Carta Geológica de la República Mexicana, escala 1: 2 000 000*, Instituto de Geología de la Universidad Nacional Autónoma de México, Secretaría de Energía, Minas e Industria Paraestatal, Consejo de Recursos Minerales.

Packer, D. R., Lovegreen, J. R. y J. L. Born, 1977. *Reservoir Induced Seismicity*, V.6, Woodward-Clyde Consultants, San Francisco, 124p.

Quijano, L. J.L. y F. M. Chacón, 1985. *Evolución geoquímica de la zona geotérmica del volcán Ceboruco, Nay.*, Departamento de Exploración de la Gerencia de Proyectos Geotermoeléctricos de la Comisión Federal de Electricidad, reporte interno 4/85.

Rothe, J. P., 1973. *Summary: geophysical report, in man made lakes: their problems and environmental effects*, Washington, Amer. Geophys. Union (Geophysical monograph 17), pp. 441-454.

Semih, M. y P. Gulkan, 1994. *Seismic hazard analysis with randomly located sources*, Natural Hazards, V.9, pp. 215-233.

Simpson, D. W., 1976. *Seismicity changes associated with reservoir loading*, Engineering Geology, V.10, pp. 123-150.

Srivastava, H. N. y R. K. Dube, 1997. *Reservoir induced seismic hazard using principal component analysis*, Geofísica Internacional, V.36, N° 1, pp. 3-7.

Talwani, P. y S. Acree, 1985. *Pore pressure diffusion and the mechanism of reservoir induced seismicity*, Pageoph, V.122, 19p.

Uribe, C. A., 1994. *Aguamilpa, uno de los mejores casos de sismicidad inducida documentada en el Mundo*, Annual meeting of the Mexican Geophysical Union, Pto. Vallarta, México.

Apéndice 1

El estimador de kriging ordinario es

$$Z^*(x_0) = \sum_{\alpha=1}^n \lambda_{\alpha} Z(x_{\alpha}),$$

que al aplicar la propiedad de insesgamiento sobre el error de estimación

$$E[Z(x) - Z^*(x_0)] = 0$$

se tiene,

$$E[Z(x) - \sum_{\alpha=1}^n \lambda_{\alpha} (Z(x_{\alpha}))] = 0$$

$$E[Z(x)] - \sum_{\alpha=1}^n \lambda_{\alpha} E[Z(x_{\alpha})] = 0,$$

considerando estacionariedad de primer orden,

$$m - \sum_{\alpha=1}^n \lambda_{\alpha} m = 0 \Rightarrow m(1 - \sum_{\alpha=1}^n \lambda_{\alpha}) = 0, \text{ que se cumple para}$$

$$\sum_{\alpha=1}^n \lambda_{\alpha} = 1 \quad \dots (1).$$

Ahora bien, para poder aplicar la propiedad de mínima variancia al error de estimación bajo la restricción (1), se hace uso del enfoque de los multiplicadores de Lagrange, por lo que,

$$\sigma_e^2 = E[(Z(x_0) - Z^*(x_0))^2], \quad \dots (2), \text{ ya que } E[Z(x) - Z^*(x_0)] = 0.$$

La ecuación (2), llamada variancia de estimación, es la expresión a minimizar bajo la restricción (1), por lo que al desarrollar la expresión (2) se tiene,

$$E[(Z(x_0) - \sum_{\alpha=1}^n \lambda_{\alpha} Z(x_{\alpha}))^2] = E[Z^2(x_0) - 2Z(x_0) \sum_{\alpha=1}^n \lambda_{\alpha} Z(x_{\alpha}) + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} Z(x_{\alpha}) Z(x_{\beta})]$$

$$= E[Z^2(x_0)] - 2 \sum_{\alpha=1}^n \lambda_{\alpha} E[Z(x_0)Z(x_{\alpha})] + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} E[Z(x_{\alpha})Z(x_{\beta})] \quad \dots (3)$$

al sumar y restar en el segundo término de la expresión (3):

$$2 \sum_{\alpha=1}^n \lambda_{\alpha} m^2, \text{ y en el tercer término la expresión:}$$

$$\sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} m^2, \text{ la expresión (3) es equivalente a,}$$

$$\begin{aligned} \sigma_e^2 = & E[Z^2(x_0)] - 2 \sum_{\alpha=1}^n \lambda_{\alpha} \{E[Z(x_0)Z(x_{\alpha})] - m^2\} - 2 \sum_{\alpha=1}^n \lambda_{\alpha} m^2 + \\ & \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} \{E[Z(x_{\alpha})Z(x_{\beta})] - m^2\} + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} m^2, \end{aligned}$$

utilizando la relación $Cov(Z(x), Z(x+h)) = E[Z(x)Z(x+h)] - m^2$ y denotando

$$Cov(Z(x_{\alpha}), Z(x_{\beta})) = Cov(x_{\alpha}, x_{\beta}), \text{ se tiene,}$$

$$\sigma_e^2 = E[Z^2(x_0)] - 2 \sum_{\alpha=1}^n \lambda_{\alpha} Cov(x_0, x_{\alpha}) - 2 \sum_{\alpha=1}^n \lambda_{\alpha} m^2 + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} Cov(x_{\alpha}, x_{\beta}) + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} m^2,$$

considerando la restricción (1) en la expresión anterior,

$$\sigma_e^2 = E[Z^2(x_0)] - 2 \sum_{\alpha=1}^n \lambda_{\alpha} Cov(x_0, x_{\alpha}) - 2m^2 + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} Cov(x_{\alpha}, x_{\beta}) + m^2$$

y si también se considera que $Var(Z(x_0)) = E[Z^2(x_0)] - m^2$ y que esta variancia es equivalente a la autocovariancia de $Z(x_0)$, denotada por $Cov(0)$, se tiene que

$$\sigma_e^2 = Cov(0) - 2 \sum_{\alpha=1}^n \lambda_{\alpha} Cov(x_0, x_{\alpha}) + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} Cov(x_{\alpha}, x_{\beta}) \quad \dots (4),$$

expresión sujeta a la restricción (1). Por lo que la función objetivo a minimizar, de acuerdo a la técnica de los multiplicadores de Lagrange es:

$$V = \sigma_e^2 + 2\mu \left(\sum_{\alpha=1}^n \lambda_{\alpha} - 1 \right), \quad \dots (5)$$

derivando (5) con respecto a las $\lambda_\alpha, \alpha=1, \dots, n$ y con respecto a μ , se generan $n + 1$ ecuaciones, por lo que

$$\frac{\partial V}{\partial \lambda_\alpha} = -2Cov(x_\alpha, x_0) + 2 \sum_{\beta=1}^n \lambda_\beta Cov(x_\alpha, x_\beta) + 2\mu = 0$$

$$\frac{\partial V}{\partial \mu} = 2 \sum_{\alpha=1}^n \lambda_\alpha - 1 = 0$$

\therefore

$$\sum_{\beta=1}^n \lambda_\beta Cov(x_\alpha, x_\beta) + \mu = Cov(x_\alpha, x_0), \dots, \alpha=1, \dots, n, \dots \quad (6)$$

la expresión (6) más la restricción (1) constituyen el sistema de ecuaciones de orden $(n+1) \times (n+1)$ cuya solución es igual a los pesos λ del estimador de kriging ordinario.

Apéndice 2

El límite de la función variograma $\gamma(h)$ cuando $h \rightarrow \infty$ corresponde a la variancia de los datos, $Cov(0)$, situación que se da cuando hay independencia o no hay correlación espacial entre estos datos. También, corresponde cuando h es mayor que el rango de correlación del modelo de variograma respectivo. Si este es el caso, la matriz de covariancias entre los datos, $Cov(x_\alpha, x_\beta)$, se reduce a una matriz diagonal cuyos elementos son iguales a la variancia común $Cov(0) = Var(Z(x))$ para todas las x , y la matriz de covariancias entre los datos y la posición a estimar, $Cov(x_\alpha, x_0)$, es igual a cero. Bajo esta circunstancia, el sistema de kriging ordinario se reduce a:

$$\lambda_\alpha Cov(x_\alpha, x_\alpha) + \mu = 0, \quad \alpha = 1, \dots, n,$$

bajo la restricción

$$\sum_{\alpha=1}^n \lambda_\alpha = 1$$

por lo que,

$$\lambda_\alpha = -\frac{\mu}{Cov(x_\alpha, x_\alpha)}, \quad \alpha = 1, \dots, n,$$

que al considerar la restricción sobre los pesos λ_α ,

$$\sum_{\alpha=1}^n \lambda_\alpha = \sum_{\alpha=1}^n -\frac{\mu}{Cov(x_\alpha, x_\alpha)} = -\frac{n\mu}{Cov(x_\alpha, x_\alpha)} = 1 \Rightarrow -\mu = \frac{Cov(x_\alpha, x_\alpha)}{n}$$

de donde,

$$\lambda_\alpha = \frac{Cov(x_\alpha, x_\alpha)}{nCov(x_\alpha, x_\alpha)} = \frac{1}{n}.$$

Por lo tanto, cada peso λ de los n datos es igual a $\frac{1}{n}$ y el resultado del estimador será:

$$Z^*(x_0) = \sum_{\alpha=1}^n \lambda_\alpha Z(x_\alpha) = \frac{1}{n} \sum_{\alpha=1}^n Z(x_\alpha)$$

que corresponde a la media aritmética de los datos de la vecindad.

Apéndice 3

La técnica utilizada en la simulación estocástica de $Z(x)$ fue la de simulación secuencial de la variable indicador. La idea fundamental es la de aplicar el algoritmo de KI para completar la información de dicha variable, con valores de probabilidad entre 0 y 1 para cada localidad simulada, donde:

- a) Se completa, en primera instancia, la información de la variable indicador para los K umbrales considerados, en una determinada localidad x por medio de KI.
- b) Se extrae de esta información de la variable indicador (función de distribución acumulada a posteriori) un valor $z'(x)$.
- c) Se considera este valor simulado $z'(x)$ dentro de la información condicionante, y se repite el proceso en otra localidad x' hasta cubrir todas las localidades.

Por lo tanto, la estimación de la variable indicador, a partir de los datos originales y los valores previamente simulados, determinan las características de la simulación de la distribución espacial de $Z(x)$.

Aplicando el estimador de KS,

$$Z^*(x_0) - m = \sum_{\alpha=1}^n \lambda_{\alpha} [Z(x_{\alpha}) - m]$$

a la variable indicador, se tiene:

$$i^*(x_0; z_k) - p_k = \sum_{\alpha=1}^n \lambda_{\alpha} [i(x_{\alpha}; z_k) - p_k] \quad \dots\dots\dots(1)$$

ya que $p_k = F(z_k) = E[I(x; z_k)]$, que corresponde a la media de la variable indicador. El sistema de ecuaciones cuya solución son los pesos λ del estimador de KS, aplicado a la variable indicador y considerando los diferentes umbrales es:

$$\sum_{\beta=1}^n \lambda_{\beta} Cov_k(x_{\beta}, x_{\alpha}) = Cov_k(x_0, x_{\alpha}) \text{ para } \alpha = 1, \dots, n \quad \dots\dots\dots(2)$$

Al considerar que la covariancia de la función aleatoria original, constituida por las variables aleatorias $Z(x)$, no es mas que un momento estadístico de la función de distribución acumulada bivariada,

$F(h; z, z') = \text{Prob}\{Z(x) \leq z, Z(x+h) \leq z'\}$, es decir,

$\text{Cov}(h) = E[Z(x)Z(x+h)] - (E[Z(x)])^2$, se tiene que para la variable indicador,

$$\text{Cov}_k(h) = F(h, z_k, z_k) - p_k^2 \quad \dots\dots\dots(3)$$

que representa un modelo de la covariancia centrada de la variable indicador para el umbral z_k . Ahora bien, dado que el estimador de KS tiene la propiedad de insesgamiento:

$E[I(x; z_k) - I^*(x; z_k)] = 0$, se tiene que,

$$E[I^*(x; z_k)] = p_k = F(z_k),$$

lo que implica que las estimaciones de la variable indicador, de donde se extraen los valores simulados $z'(x)$, reproducen la función de distribución acumulada univariada $F(z_k)$.

Considerando ahora la covariancia del estimador de la variable indicador $I^*(x; z_k)$ con cualquiera de los n datos indicadores:

$$E[(I^*(x; z_k) - p_k)(I(x_\alpha; z_k) - p_k(x))] =$$

$$\sum_{\beta=1}^n \lambda_\beta E[(I(x_\beta; z_k) - p_k)(I(x_\alpha; z_k) - p_k)] =$$

$$\sum_{\beta=1}^n \lambda_\beta \text{Cov}_k(x_\beta, x_\alpha) = \text{Cov}_k(x_0, x_\alpha)$$

de acuerdo con la relación (3) para $\alpha = 1, \dots, n$.

Por lo que el estimador $I^*(x; z_k)$, y por consiguiente $z'(x)$, es tal que reproduce todos los valores del modelo de la covariancia indicador $\text{Cov}_k(x_0, x_\alpha)$.