



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**PROGRAMA DE MAESTRÍA Y DOCTORADO EN
INGENIERÍA**

**ATENUACIÓN ESPECTRAL SÍSMICA
EN EL CAMPO LEJANO:
UNA APROXIMACIÓN CON REDES NEURONALES
ARTIFICIALES**

T E S I S

COMO REQUISITO PARA OBTENER EL GRADO DE

**DOCTORA EN INGENIERÍA
(MECÁNICA DE SUELOS)**

QUE PRESENTA:

GLORIA ELENA ECHEVERRI RAMÍREZ

DIRECTOR DE TESIS
DR. MIGUEL P. ROMO ORGANISTA

MÉXICO, D.F., 2006



Agradecimientos

Doy gracias a Dios.

Quiero expresar mis agradecimientos:

Al Dr. Miguel Pedro Romo Organista por la dirección de este trabajo y por el valioso aporte a mi formación profesional.

Al jurado, conformado por los doctores Daniel Reséndiz Núñez, Jesús Alberro Aramburu, Francisco Cervantes Pérez, Mario Gustavo Ordaz Schroeder, Shri Krishna Singh y Juan Diego Jaramillo Fernández por sus indicaciones constructivas y la revisión del manuscrito.

A la Universidad EAFIT y al Instituto de Ingeniería de la UNAM que con su apoyo hicieron posible mi especialización en México.

A mi familia por su apoyo incondicional.

A todas las personas que de alguna manera contribuyeron al logro de este objetivo.

Contenido	Página
Resumen	i
Contenido	ii
1 Introducción	1
2 Consideraciones sobre el movimiento sísmico en superficie y criterios de estimación	4
2.1 Parámetros que caracterizan movimientos fuertes	4
2.2 Variables asociadas a los efectos sísmicos	7
2.2.1 Dislocación sísmica y factores asociados	7
2.2.2 Trayectoria de las ondas y factores asociados	11
2.2.3 Efectos de sitio y factores asociados	12
2.3 Criterios de evaluación	15
2.3.1 Predicciones empíricas	15
2.3.2 Predicciones teóricas y semiempíricas	17
2.3.3 Duración del movimiento fuerte	22
2.4 Modelo sismológico de atenuación sísmica en el campo lejano	25
2.4.1 Hipótesis	25
2.4.2 Cálculo del desplazamiento	27
2.4.3 Simulación teórica	27
3 Redes neuronales artificiales	29
3.1 Introducción	29
3.2 Estructura de un sistema neuronal artificial	31
3.2.1 Modelo neuronal artificial	31
3.2.1.1 Modelo general de una neurona artificial	31
3.2.1.2 Modelo estándar de neurona artificial	35
3.3 Arquitectura de redes neuronales	36
3.4 Operación	37
3.4.1 Fase de aprendizaje	37
3.4.2 Fase de ejecución	38
3.5 Clasificación de los modelos neuronales	40
3.5.1 El asociador lineal	40
3.5.2 Regla de aprendizaje de Hebb	40
3.5.3 Regla de la pseudoinversa	41
3.5.4 Redes neuronales de aprendizaje supervisado y arquitectura no realimentada	42
3.5.4.1 Perceptrón simple	43
3.5.4.2 Adalina	43
3.5.4.3 Perceptrón multicapa	45
3.6 Redes multicapa vs. métodos estadísticos	49
3.7 Nota de observación sobre la estructura de los datos	50
3.8 Limitaciones de las redes neuronales artificiales	50
4 Ambiente sismogeotécnico y base de datos	51
4.1 Caso México	51
4.1.1 Ambiente sismotectónico	51
4.1.2 Base de datos	52
4.2 Caso Colombia	54
4.2.1 Ambiente sismotectónico	54
4.2.2 Distribución de la actividad sísmica en el país	57
4.2.3 Base de datos	62
4.3 Ciudad de Medellín	64
4.3.1 Ambiente sismotectónico	64
4.3.2 Base de datos	66

5	Cuantificación neuronal de parámetros sísmicos	68
5.1	Variables sensoriales	68
5.2	Modelo neuronal	70
5.3	Resultados	74
5.3.1	Caso México	74
5.3.2	Caso Ciudad de Medellín	93
5.3.3	Caso Colombia	100
5.4	Validación del método	106
5.4.1	Comparación con los valores observados	106
5.4.2	Comparación estadística con otros trabajos	106
5.5	Modelo paramétrico vs. redes neuronales para la estación sísmica CU	106
5.5.1	Modelo lineal paramétrico	106
5.5.2	Modelo neuronal para la estación de registro sísmico CU	115
5.6	Simulación hipotética sobre la aceleración máxima	116
5.7	Incertidumbre	119
6	Análisis de sensibilidad sobre las variables sensoriales	122
7	Conclusiones y recomendaciones	131
	Agradecimientos	133
	Referencias	134

Resumen

Con base en los procedimientos de Inteligencia Artificial, específicamente las Redes Neuronales Artificiales e incorporando la información contenida en la Base Mexicana de Datos de Sismos Fuertes, Red Acelerográfica de Medellín y Red Nacional de Acelerógrafos de Colombia, se identificaron los modelos cognoscitivos que son capaces de calcular las relaciones funcionales para predecir la duración de la parte intensa del movimiento sísmico, el componente horizontal de la aceleración máxima y espectros de respuesta elásticos. Estos modelos neuronales se obtuvieron mediante transformaciones algorítmicas de los parámetros que caracterizan la fuente sísmica, la trayectoria de las ondas y los efectos de sitio del lugar de interés. Se consideran las zonas sismogénicas de subducción de la Costa del Océano Pacífico, desde el estado de Michoacán hasta el estado de Oaxaca, y los sismos interplaca e intraplaca generados dentro del territorio colombiano. Para el entrenamiento de las Redes Neuronales y la verificación de su capacidad de generalización, se utilizó un conjunto de 516 acelerogramas en el caso mexicano y 560 acelerogramas en el caso colombiano. Los resultados obtenidos con las modelaciones neuronales se presentan como: 1) mapas de contornos de duración de la fase intensa y de isoaceleraciones para varios sismos; 2) curvas de atenuación para las diferentes zonas geotécnicas; 3) comparación gráfica entre los espectros de respuesta elásticos calculados a partir del registro original y los estimados con los Sistemas Neuronales Artificiales.

Abstract

Based on the Artificial intelligence techniques, specifically Artificial Neural Networks and incorporating the information contained in the Mexican Strong Motion Database, the Local Accelerographs Network of Medellín and the National Seismological Network of Colombia, cognitive models were identified which are able to calculate functional relations to predict the duration of the intense part of seismic shaking, the horizontal peak ground acceleration, and the acceleration response spectra with 5% damping. These neural models were obtained by means of algorithmic transformations of parameters that characterize seismic source, wave trajectory and site effects at a particular location. The functional approximators were developed for the subduction zone along the Pacific Ocean coast, from Michoacan to Oaxaca, and for the interplate, intraplate, and subduction earthquakes generated within the Colombian territory. In training the Neuronal Networks and verification of their generalization capabilities, a set of 516 accelerograms was used in the Mexican case and 560 accelerograms in the Colombian case. The results obtained with the neural models are presented in terms of: 1) contour maps of equal duration of the intense motion phase and of equal peak ground acceleration for several earthquakes; 2) attenuation curves for different geotechnical zones; and 3) graphic comparison among acceleration response spectra calculated from the original records and those predicted with Artificial Neural Systems.

Capítulo 1

1 Introducción

La calidad de los estudios de amenaza sísmica requeridos en los códigos antisísmicos depende fundamentalmente de los haberes para estimar la naturaleza y la severidad de la ocurrencia de movimientos del terreno provocados por sismos fuertes. Con el fin de contribuir en esta dirección, se planteó como objetivo identificar y aplicar procedimientos cognoscitivos que permitan estimar en forma cuantitativa y en términos ingenieriles las características fundamentales del movimiento del suelo debido a acciones sísmicas.

La respuesta superficial del suelo ante excitación sísmica es un fenómeno complejo gobernado por un gran número de variables que dificultan la modelación analítica del problema. Los movimientos que pueden esperarse en un determinado sitio cambian de evento a evento y dependen de factores que varían en el tiempo y en el espacio. La variabilidad está relacionada con la fuente sísmica y sus procesos sismogénicos, la trayectoria de propagación de las ondas sísmicas desde la fuente al sitio, y la interacción de las ondas que llegan al sitio con las condiciones fuertemente heterogéneas que caracterizan la geología de la mayoría de los perfiles del subsuelo.

Desde décadas atrás, el problema de atenuación y amplificación sísmica ha sido abordado con técnicas de regresión múltiple y modelos teóricos. Ambos, soportados en la evidencia empírica recabada durante años y en los significativos avances físicos, sismológicos y matemáticos que representan la generación, la transmisión, la disipación y difracción de la energía sísmica contenida en las ondas incidentes. Un gran número de estas contribuciones han sido documentadas en la literatura técnica [Brune 70, Ambraseys 73, Donovan 73, Esteva 74, Trifunac 76, McGuire 80, Boatwright 82, Boore 83, Boore 84, Luco 85, Singh 87, Boatwright 88, Joyner 88, Ordaz 92, Campbell 93, McVerry 93, Abrahamson 97, Anderson 97, Somerville 98]. No obstante, ninguno de estos criterios ha conducido a resultados suficientemente satisfactorios.

Hay cierta tendencia a establecer un nuevo campo en la ciencia de la computación, que integra métodos de solución que no pueden ser descritos mediante un enfoque algorítmico tradicional. Se trata de métodos derivados de inteligencia artificial.

La Inteligencia Artificial AI, tienen su origen en la emulación burda del comportamiento de los sistemas biológicos. Se trata de un enfoque capaz de manejar las imprecisiones e incertidumbres que se presentan al resolver algún problema (reconocimiento de patrones, toma de decisiones, etc.), ofreciendo soluciones robustas y de fácil implantación, e incluye: la Teoría de Conjuntos Difusos o Lógica Borrosa, el Razonamiento Aproximado, los Sistemas Expertos, los Algoritmos Genéticos, la Programación Genética, los Sistemas Clasificadores, Aprendizaje de Árboles de Decisión, la Teoría del Caos, y las Redes Neuronales, entre otros. Estas técnicas son aplicables en numerosos campos del conocimiento entre ellos la ingeniería sísmica [Dowla 90, Emami 96, Essenreite 96, Romo 00, Lin 00, Paolucci 00, Romo 01, Hurtado 01a, Hurtado 01b, Sarmiento 01]. Aunque, en principio, se trata de enfoques diferentes, suelen buscarse combinaciones de ellas.

La Lógica Difusa es una técnica innovadora que permite trasladar aseveraciones del lenguaje natural a la formalidad matemática. Trabaja con conjuntos que no tienen límites perfectamente definidos a través de funciones de pertenencia y por lo mismo dan flexibilidad a la modelación. Se puede decir que la Lógica Difusa es una generalización de la lógica tradicional, en la que las variables y sus relaciones no están definidas con nitidez sino con cierta gradación. Es una forma de razonamiento que permite incorporar esquemas de pensamiento típicamente humanos. Los sistemas basados en la lógica difusa han demostrado su utilidad en una gran variedad de operaciones y tareas de reconocimiento de patrones. Una de las principales potencialidades de la lógica difusa cuando se le compara con otras metodologías que tratan con datos imprecisos es que las bases de conocimiento de un sistema de este tipo son fáciles de examinar, entender, mantener y actualizar.

En casi todas las aplicaciones de la inteligencia artificial es necesario procesar información incierta o imprecisa y evaluarla mediante modelos no deterministas. Dentro del campo del razonamiento incierto, se denomina razonamiento aproximado al que utiliza métodos numéricos, para diferenciarlo de los métodos simbólicos de tratamiento de la incertidumbre. Los principales métodos de razonamiento aproximado, son el modelo bayesiano subjetivo, los factores de certeza, la teoría de la evidencia de Dempster y Shafer, y la teoría de la posibilidad.

Los Sistemas Expertos son sistemas informáticos capaces de emular el comportamiento de un experto humano en un área limitada del conocimiento. Las reglas son las características principales de los sistemas expertos. Esto significa que contienen un juego predefinido de conocimientos que se utiliza para todas las decisiones. El sistema usa las reglas y las inferencias que se le indican para producir sus resultados. Dependiendo del tipo de entrada y las reglas utilizadas, los sistemas expertos pueden emplearse como herramientas cuantitativas o cualitativas. Un sistema experto genérico, consiste de dos módulos principales: la base de conocimientos y el motor de inferencia. La base, contiene el conocimiento disponible sobre un tema específico para el que se diseña el sistema. Este conocimiento se codifica con cierta notación que incluye reglas, predicados, redes semánticas y objetos. El motor de inferencia combina los hechos y las preguntas particulares utilizando el pilar de conocimiento, y seleccionando los datos y pasos apropiados para presentar los resultados.

Los Algoritmos Genéticos consisten en métodos generalizados de búsqueda, diseño y optimización que simulan el proceso y su evolución. Su objetivo principal es evolucionar a partir de una población de soluciones para un problema, intentando producir nuevas generaciones de soluciones que sean mejores que las anteriores. Son algoritmos matemáticos aplicables a problemas de optimización, con fundamento en la teoría de la evolución de Darwin; operan en un ciclo simple de selección y reproducción, implicando una recombinación y mutación del "material genético" de las soluciones. Una población de posibles soluciones se genera al azar y éstas se evalúan con respecto a un objetivo; las más aptas se combinan entre sí para producir nuevas soluciones. El ciclo se repite hasta llegar a una solución aceptable o a la óptima. Al igual que en los procesos de selección natural, debe existir equilibrio entre la explotación de soluciones buenas (cruzamiento) y la búsqueda de nuevas posibilidades (mutación) para mantener una razonable diversidad en la población que evoluciona generación tras generación hasta alcanzar un punto de equilibrio que es o está muy cerca del óptimo absoluto del problema. Algunas características de los Algoritmos Genéticos están representadas en el uso de parámetros codificados como una cadena de longitud finita sobre un alfabeto finito, en algoritmos de funcionamiento paralelo, en uso de operadores probabilistas, en la menor sensibilidad a los máximos locales y que están menos restringidos por continuidad y unimodalidad.

La Programación Genética en IA, es una forma de progreso para generar programas automáticamente mediante un juego de instrucciones primitivas. Extiende los principios de los algoritmos genéticos para lograr que un programa informático evolucione. Así, conduce al predominio de los individuos más aptos para un ambiente dado. Cada nueva generación se obtiene por la reproducción de los programas.

Los Sistemas Clasificadores ejecutan un tipo de aprendizaje que genera reglas a partir de ejemplos. Utilizan, así mismo, algoritmos genéticos para obtener reglas de producción, que luego alimentarán a los sistemas.

El Aprendizaje de Árboles de Decisión sirve para clasificar instancias caracterizadas como conjuntos de atributos. Cada nodo terminal del árbol representa un atributo y tendrá tantas ramas como valores distintivos puede tomar dicho atributo. Para clasificar una instancia, se comienza por la raíz y se va siguiendo el camino que indica los valores de los atributos hasta llegar al nodo terminal, que indica la clase a la que pertenece. El algoritmo más representativo de este tipo es denominado ID3. Este algoritmo emplea como heurística para evaluar el mejor atributo una medida estadística conocida como ganancia de información, que refleja el grado en que un atributo sirve para discriminar entre las distintas clases. Si un atributo es muy discriminante, forma grupos de ejemplos homogéneos (de la misma clase) para cada uno de sus valores. Este comportamiento tiende a producir árboles pequeños con pocos niveles, ya que se seleccionan aquellos atributos que tienen mayor probabilidad de originar hojas o conjuntos totalmente análogos en pocos pasos.

La Teoría del Caos es un conjunto de técnicas para examinar y determinar relaciones altamente complejas entre datos inicialmente clasificados como aleatorios. Una característica de este tipo de datos es que una diferencia en las condiciones iniciales puede repercutir en cambios drásticos de los resultados en un período muy breve. Fundamental en la teoría del caos es restar importancia a la unidad individual y otorgarle trascendencia a las situaciones entre los diferentes niveles del sistema. Esta teoría describe el comportamiento de sistemas no lineales partiendo de la hipótesis que la dinámica y el comportamiento regular de un conjunto complejo de datos tiene un orden subyacente. Igualmente, demuestra la tendencia de los resultados a oscilar alrededor de un área en oposición a estabilizarse en un nivel de equilibrio.

Las Redes Neuronales RN son modelos computacionales que asemejan el funcionamiento de porciones del cerebro humano. La emulación se realiza en forma paralela, distribuida y adaptiva mediante muchos procesadores conectados entre sí. Son un conjunto de técnicas matemáticas para modelar las conexiones entre una familia de datos. Por medio de un proceso de entrenamiento, las redes neuronales aprenden de los datos que se les suministran. El conocimiento adquirido se almacena como pesos sinápticos adjudicados a los distintos enlaces entre las neuronas de la red.

Las redes neuronales artificiales RNAs se caracterizan por su eficiencia en el reconocimiento y aprendizaje de patrones en un conjunto de datos no lineales, incompletos, con ruidos y hasta compuestos por ejemplos contradictorios. La habilidad de esta herramienta para manipular datos imprecisos, incrementa la actividad de las redes en el procesamiento de información sin reglas claras o que no puedan ser formuladas fácilmente. Sin reglas convencionales, una red neuronal obtiene experiencia analizando automática y sistemáticamente una base de datos, para determinar pautas de comportamiento. Con base en estas normas inferidas, las redes neuronales, realizan predicciones sobre nuevos casos, siempre que estos pertenezcan a la misma familia de las incluidas en la base de datos.

Dada la complejidad del problema del movimiento sísmico, en el que interviene un gran número de variables, y considerando el potencial que ofrecen las RNAs con respecto a las otras metodologías derivadas de AI para identificar conexiones al utilizar relaciones no lineales entre los datos, en esta tesis se optó por esta técnica cognoscitiva para estimar la respuesta superficial del suelo frente a acciones sísmicas.

Haciendo uso de la información contenida en la Base Mexicana de Datos de Sismos Fuertes, la Red Acelerográfica de Medellín (RAM) y la Red Nacional de Acelerógrafos de Colombia, se identificaron los modelos de RNAs que permiten predecir la duración de la parte intensa del movimiento sísmico, la aceleración máxima en el componente horizontal y su espectro de respuesta elástica de campo libre, para México y Colombia. Esto se logró a través de la transformación algorítmica de los parámetros que caracterizan la fuente sísmica, la trayectoria de las ondas y los efectos de sitio de la zona de interés, considerando las zonas sismogénicas de subducción de la Costa del océano Pacífico, desde el estado de Michoacán hasta el estado de Oaxaca, y los sismos interplaca e intraplaca generados dentro del territorio colombiano.

Lo peculiar de las enormes amplificaciones registradas en la ciudad México y la necesidad de contar con procedimientos que permitan evaluar la atenuación del movimiento sísmico para Colombia con base en la información local, fueron las razones que llevaron a elegir a México y Colombia como casos de aplicación de las herramientas basadas en las técnicas de RNAs para la predicción de parámetros sísmicos que caracterizan los movimientos del terreno en superficie.

En el cuerpo del presente escrito se desarrolla la modelación cognoscitiva, de la siguiente manera:

El capítulo 2 incluye las consideraciones teóricas del movimiento fuerte en superficie, en términos de la respuesta del subsuelo en diferentes sitios ante acciones sísmicas, y describe algunos de los criterios usuales en la literatura técnica para estimar y evaluar los parámetros que caracterizan la respuesta del terreno con fines ingenieriles.

En el capítulo 3 se describen los sistemas neuronales artificiales, su evolución, sus componentes y modos de operación, así como la representación matemática de reglas y algoritmos clásicos. Se

presta especial atención a las redes neuronales empleadas como estimadores de funciones, dado el interés de esta aplicación.

El capítulo 4 contiene una exposición resumida de los principales rasgos tectónicos regionales de los casos estudiados, así como información geotécnica y de la base de datos utilizada.

En el capítulo 5 se presentan los modelos neuronales y la cuantificación de parámetros y resultados en forma de mapas de contornos de duración de la fase intensa y de isoaceleración para varias parejas fuentes-sitio, curvas de atenuación para las zonas constituidas por material rocoso, material de transición y suelo blando y la comparación entre espectros de respuesta elásticos calculados a partir del registro original y los estimados con RNAs, para diferentes estaciones dentro de la ciudad de México, así como los obtenidos para la ciudad de Medellín y otros lugares en Colombia.

Un análisis de sensibilidad en términos cualitativos, sobre las variables sensoriales que conforman el vector de entrada de las redes neuronales artificiales implementadas en este estudio, es el objetivo del capítulo 6.

Finalmente, en el capítulo 7 se enuncian las conclusiones y recomendaciones para utilizar modelos neuronales en el estudio de la atenuación espectral sísmica.

Capítulo 2

Consideraciones sobre el movimiento sísmico en superficie y criterios de estimación

Con el tiempo, la medición instrumental del fenómeno sísmico en superficie que contiene la configuración fuente–estación, ha proporcionado información valiosa en el desarrollo de procedimientos analíticos para la valoración cuantitativa del movimiento, lo que ha representado un importante avance en la comprensión y predicción de los efectos y los alcances de su influencia. Los registros sísmicos y los espectros de respuesta asociados, muestran la influencia de tres aspectos de origen diferente no disjuntos en sus efectos: la fuente sísmica, la trayectoria de las ondas y los efectos de las condiciones locales en el sitio de interés. Una sinopsis de tales aspectos y algunos criterios de evaluación constituye la materia de este apartado.

2.1 Parámetros que caracterizan movimientos fuertes

Varios parámetros dinámicos han sido propuestos para caracterizar la amplitud, el contenido frecuencial y la duración de la fase intensa de temblores, ver tabla 2.1. De todos los parámetros ampliamente reconocidos que reflejan los cambios del movimiento sísmico y sus efectos en las estructuras, tres son los más relevantes y corresponden a las variables propiamente dichas del movimiento -desplazamiento, velocidad y aceleración- en términos de valores máximos o de ordenadas espectrales en función de la frecuencia; y algún criterio que defina cierta porción del acelerograma durante el cual el movimiento fuerte ha tenido lugar.

Debido a la complejidad del proceso de desplazamiento del tren ondulatorio y la disipación de energía asociada, los estudios de atenuación durante décadas se orientaron, en su mayoría, a evaluar los parámetros máximos como indicadores de la intensidad relativa del movimiento. Específicamente, el valor del componente horizontal de la aceleración máxima desarrollada en algún sitio en particular durante un sismo, ha sido el índice más comúnmente usado por su significado en el diseño ingenieril. Su importancia radica, en el escalamiento de espectros de respuesta para distintos grados de amortiguamiento y en la estimación groso modo de las fuerzas de inercia que proporciona coeficientes para el diseño sísmico en función del tipo de estructura. De este modo, el riesgo sísmico queda expresado en términos de la esperanza del valor presente de las pérdidas por sismo, en función del coeficiente basal de diseño, que es el parámetro usado para definir la capacidad de las estructuras ante carga lateral [Ordaz, 92]. Por tanto, la aceleración máxima sigue siendo extensamente utilizada para caracterizar el movimiento en las altas frecuencias.

No obstante lo útil de los valores máximos estimados para un sitio específico, estos no describen el contenido de frecuencias del patrón ondulatorio, ni el efecto en la forma del espectro, sólo de una manera indirecta en su correspondiente zona de influencia. En consecuencia, resulta necesario predecir espectros de respuesta debido a la ventaja que representan al mostrar la intensidad de la señal en diversas frecuencias y la concentración de la energía en algunas de ellas.

Así, los movimientos en terreno firme generados por las acciones sísmicas, se especifican en términos de espectros de aceleración para diferentes porcentajes de amortiguamiento. La obtención de un espectro de aceleraciones en terreno firme es el resultado de filtrar el registro acelerográfico de un determinado sitio a través de osciladores amortiguados. Este procedimiento conduce a los denominados espectros de respuesta de aceleración, velocidad y desplazamiento, que son el conjunto de los valores absolutos máximos respectivos del movimiento de la masa del oscilador. Se puede aseverar que el espectro de respuesta es la base para el diseño antisísmico.

Tabla 2.1 Parámetros dinámicos que reflejan la variación del movimiento fuerte del suelo

Parámetro	Concepto	Características del movimiento sísmico		
		Amplitud	Contenido de frecuencias	Duración
Aceleración máxima, a_{max}	Obtenida directamente del acelerograma	•		
Velocidad máxima, V_{max}	Calculada del registro de velocidades	•		
Período predominante, T_p	Definido como el período de vibración correspondiente al máximo valor del espectro de amplitudes de Fourier		•	
Frecuencia central, Ω , es una medida de la frecuencia donde el espectro de potencia está concentrado. Puede ser usada para calcular la aceleración máxima teórica	$\Omega = \sqrt{\frac{\lambda_2}{\lambda_0}} ; \lambda_n = \int_0^{\omega_n} \omega^n G(\omega) d\omega$ $a_{max \text{ Teórica}} = \sqrt{2\lambda_0 \ln(2.8 \frac{\Omega T_d}{2\pi})}$		•	
Factor de forma, δ , indica la dispersión de la función de densidad del espectro de potencia	$\delta = \sqrt{1 - \frac{\lambda_1^2}{\lambda_0 \lambda_2}}$		•	
V_{max} / a_{max} : cantidad asociada con el movimiento de frecuencias diferentes	La cantidad $2\pi (V_{max} / a_{max})$, puede ser interpretada como el período de vibración de una onda armónica equivalente y provee una indicación de cuáles períodos de movimiento son más significativos		•	
Duración, T_d	T_d es igual al inverso de la frecuencia de esquina, f_0 , ($T_d = 1/2\pi f_0$)			•
Intensidad de Arias, IA	$IA = \frac{\pi}{2g} \int_0^{\infty} [a(t)]^2 dt$	•	•	•
Aceleración cuadrática media, a_{rms}^2	$a_{rms}^2 = \frac{1}{T_d} \int_0^{T_d} [a(t)]^2 dt$	•	•	
Intensidad característica, I_c	$I_c = a_{rms}^{1.5} T_d^{0.5}$ <p>Es un parámetro relacionado directamente con el índice de daño de la estructura.</p>	•	•	•

La forma espectral de diseño requerida en los estudios de riesgo sísmico, generalmente se obtiene del estudio de acelerogramas reales captados durante sismos fuertes. Sin embargo, en el momento en que el espectro hace tránsito de uno de respuesta a uno de diseño, surge la duda de cómo manejar, en el espectro diseño, variables intrínsecas en el acelerograma que inhiben su estudio como una variable independiente. Dentro de estas variables se encuentra, la duración de la parte intensa del movimiento.

El interés por integrar de alguna forma la duración del movimiento fuerte del terreno al análisis estructural se debe fundamentalmente al deterioro que sufren los edificios por el aumento del número y amplitud de ciclos de carga durante sismos importantes. El daño estructural depende de la historia de deformación que sufre la estructura antes y después de las máximas demandas del movimiento.

Se ha observado que espectros similares con duraciones diferentes provocan varios niveles de daño a las estructuras, causando obviamente, menor daño los movimientos sísmicos de menor duración. También, se han registrado casos en los que el movimiento sísmico es de corta duración pero con un alto valor de aceleración pico, PGA, y ha generado pocos daños. Por otro lado, se sabe que entre mayor sea el período de vibración de la estructura, se requerirá una mayor duración para llegar a la máxima respuesta. Esto se manifiesta en el hecho de que para eventos de larga permanencia, crece

la probabilidad de que se dañen estructuras esbeltas porque la señal contiene ondas de períodos intermedios y largos.

Durante sismos recientes, en México se ha podido comprobar que ni la aceleración máxima del terreno ni los desplazamientos del terreno ni los espectros de respuesta están estrechamente correlacionados con los daños observados en obras civiles. Sin embargo, en los sitios donde el movimiento fue intenso, la duración es un parámetro bien correlacionado con los efectos de los sismos en las estructura [Guerrero 97]

En la actualidad la manera como se introduce la variable duración dentro del espectro de diseño es un proceso cualitativo que eleva el grado de mesura en la zona de períodos largos del espectro.

2.2 Variables asociadas a los efectos sísmicos

Se sabe que la interacción y el movimiento relativo entre las diferentes placas tectónicas producen la mayoría de los sismos, debido a la liberación súbita de los esfuerzos que se generan en las zonas de contacto de placas y al interior de ellas. Desde el ámbito de la sismología, los eventos sísmicos son tratados como un problema de radiación de ondas elásticas que se propagan en el interior de la tierra y en superficie. Se modelan dos grupos principales, las ondas de cuerpo y las de superficie, en ambos casos, considerando que en su recorrido las ondas tienden a un decremento en su amplitud debido principalmente a las refracciones y reflexiones, a la atenuación geométrica, la disipación de la energía debido a procesos viscosos y al comportamiento no lineal del material térreo. Por lo tanto, la amplitud y contenido de frecuencias del movimiento del suelo dependen de la atenuación o amplificación de las ondas que lo producen. Algunos factores asociados a la generación, la transmisión y la disipación o amplificación de la energía sísmica y sus efectos, se listan en la tabla 2.2.

Tabla 2.2 Aspectos y factores asociados a los efectos sísmicos

Aspectos	Factores	Parámetro físico indicador	Referencia
Fuente sísmica	Fuente interplaca o intraplaca	Profundidad, h	[Aki 80], [Papageorgiou 83], [Bullen 85], [Nishenko 87], [Lay 95]
	Energía liberada	Caida de esfuerzos, $\Delta\sigma$ Momento sísmico, M_0 Magnitud Momento, M_w	[Brune 70], [Kanamori 77] [Hanks 79, 82], [Aki 80], [Singh 90]
	Geometría de la falla, mecanismo de ruptura, dirección de la distribución de la energía radiada	Azimut, ϕ Buzamiento, δ Dirección del movimiento, λ Patrón de radiación	[Aki 80], [Boatwright 82, 88], [Singh 88], [Boore 89], [Somerville 98]
Propagación	Características de la trayectoria de las ondas	Modelo de corteza -velocidad Distancia a la fuente, R Ángulo de directividad, i Coeficientes de atenuación, p.e., factor de calidad Q, y κ	[Knopoff 64], [Aki 80], [Boore 84], [Singh 88] [Boore 89], [Somerville 98]
Condiciones locales	Excitación sísmica	Tipo de onda y azimut de la onda incidente	[Aki 88], [Somerville 98]
	Estratigrafía local	Litología en términos de densidad ρ y velocidad de propagación de ondas de cortante, β . Contraste de impedancia. Espesor o profundidad de los estratos, H. Período dominante de vibración, T_p . Anomalías estratigráficas.	[Seed 69], [Seed 74], [Dobry 76], [Romo 86a], [Romo 86b], [Seed 87], [Silva 88], [Singh 88], [Idriss 90], [Rosenblueth 92], [Graves 93], [Somerville 98], [Romo 95], [Ovando 03].
	Topografía	Forma y dimensión de la característica topográfica. Geometría de la cuenca o de la depresión del terreno. Contraste de impedancia.	[Boore 73], [Jibson 87], [Aki 88], [Geli 88], [Finn 91], [Faccioli 91], [Sánchez-Sesma 93], [Huang 99]

2.2.1 Dislocación sísmica y factores asociados

Tipos de fuentes

Con base en la teoría sobre la tectónica de placas y la forma de ocurrencia de los sismos, las fuentes sísmicas se agrupan en general, en dos categorías: sismos interplaca, ocurridos en bordes

convergentes o en áreas próximas a la periferia de las placas litosféricas (se sitúan a lo largo del plano inclinado de la placa subducida o zona de Benioff) y sismos intraplaca, generados en las zonas alejadas de los bordes o límites conocidos entre placas litosféricas o en fallas geológicas producto del tectónismo local.

Este, es el caso colombiano en donde se presenta gran diversidad de sismofuentes siendo más recurrentes los temblores generados en fallas geológicas. Según Somerville [98], se ha observado una mayor variabilidad en el movimiento del suelo en sismos producidos por fallas geológicas que para los sismos de categoría tectónica. En el caso de México, las características de la fuente no son tan "disímiles" en referencia a la génesis de los sismos que se generan en el territorio colombiano, y corresponden en su mayoría a las zonas sismogénicas de subducción de la Costa del Océano Pacífico, pero hay condiciones geológicas regionales y locales que hacen que los movimientos generados por un sismo varíen perceptiblemente en sitios diferentes que se encuentran aproximadamente a la misma distancia de la fuente.

Profundidad focal

Indudablemente por su mecanismo, hay una relación estrecha entre la génesis de la fuente sísmica y la profundidad del temblor. De manera arbitraria se han otorgado diferentes calificativos (p.e., se denominan sismos superficiales los generados a menos de 20 km por debajo de la superficie; sismos someros los generados entre 20 y 70 km por debajo de la superficie; sismos intermedios los generados entre 70 y 300 km por debajo de la superficie; sismos profundos aquellos cuyo hipocentro se localiza a distancias verticales mayores de 300 km por debajo de la superficie). En fallas geológicas locales, la profundidad generalmente es superficial, mientras que en las zonas de subducción la localización de los epicentros aumenta gradualmente a medida que la placa oceánica se inclina y se sumerge bajo la placa continental, así los focos sísmicos son poco profundos frente a la costa pero tierra adentro se localizan a profundidades mayores.

La magnitud no expresa el grado de daño que puede causar un sismo, pero se considera que la profundidad si es un parámetro significativo en la intensidad sísmica. Por ejemplo, un evento de magnitud grande y foco profundo puede que no cause daños significativos; al contrario, uno de magnitud moderada en un foco localizado a poca profundidad puede ser muy destructivo (p.e., los terremotos de Managua en 1972, 10 km de profundidad; San Salvador en 1986, 8 km de profundidad; Armenia (Colombia) del 25 de enero de 1999, menos de 15 km de profundidad). Este último, de magnitud moderada ($M_L = 6.1$), causó graves daños en la zona cafetera de Colombia, afectando las ciudades de Armenia, Pereira y un gran número de pequeñas poblaciones. La severidad de los daños se explica aparentemente por lo superficial del foco y, para las ciudades de Armenia y Pereira, por su cercanía al lugar del evento y la amplificación de las ondas sísmicas por suelos blandos (secuencia de ceniza volcánica con espesores del orden de 25 m) procedente de la cadena de volcanes de la cordillera central.

Los sismos de foco profundo producen en los sismogramas trazos muy diferentes a los registrados por sismos superficiales, aun cuando la energía liberada sea similar para ambos eventos. Cerca de la dislocación sísmica, a distancias menores de 150 km, prevalecen en los registros ondas de cuerpo, sea porque llegan directamente o porque arriban después de ser reflejadas o refractadas. A distancias mayores comienzan a predominar las ondas de superficie (de Love y de Rayleigh) de menor amplitud. A su vez, la amplitud de las ondas de superficie disminuye exponencialmente con la profundidad, de modo que en eventos de foco profundo imperan las ondas de cuerpo que tienen amplitud y contenido de energía mayores; por ejemplo, para reflejar lo observado en cuanto a que los eventos profundos presentan caída de esfuerzos ($\Delta\sigma$) mayor que los superficiales, se asumen para los superficiales $\Delta\sigma=100$ bares y para los profundos $\Delta\sigma=200$ bares, [Ordaz 89].

Magnitud o energía liberada

Si bien el movimiento fuerte del suelo que produce un sismo es el resultado de un proceso complejo, el uso de la magnitud como un índice de la energía liberada en el foco es una forma simple de describir la fuente. La magnitud es una medida cuantitativa e instrumental del tamaño del evento, relacionada con la energía sísmica liberada durante el proceso de ruptura en la falla, por lo que es

una constante única que se asigna a un determinado sismo y es independiente del sitio de observación. Se calcula a partir de varias escalas de magnitudes. En este estudio se utilizó la magnitud momento, M_w [Kanamori 77, Hanks 79].

La escala de magnitud momento o magnitud de energía, se obtiene de la estimación de la energía cinética radiada por un temblor y depende del momento sísmico (M_0). El parámetro M_0 es el producto del módulo de rigidez del medio donde se produce el sismo (G), el área de dislocación de la ruptura (A) y el desplazamiento del medio (μ)

$$M_0 = G A \mu \quad (2.1)$$

En la mayoría de los casos es difícil estimar las dimensiones de la falla, por lo que el momento sísmico, se calcula con regularidad a partir de las ondas de período largo debido a que son menos afectadas por la atenuación inelástica y la complejidad estructural, que las ondas de período corto. La relación entre la magnitud momento y el momento sísmico está dada por la siguiente expresión

$$M_w = \frac{2}{3} \log M_0 - 10.7 \quad (2.2)$$

Desde el punto de vista cualitativo, la magnitud afecta directamente las amplitudes de las ondas sísmicas, por lo tanto, a mayor magnitud, en general, las aceleraciones tienden a incrementarse.

En la medida en que las ondas sísmicas viajan, los registros que se obtienen muestran en sus espectros un decremento en la región de períodos cortos, lo cual se manifiesta en un corrimiento de la cima del espectro hacia la derecha, esto es, hacia la zona de períodos largos. Un ejemplo de este comportamiento, se ilustra en la figura 2.1.

Por otro lado, se ha observado que cuanto mayor es el tamaño del temblor, mayor es la eficiencia en la generación de ondas de período largo. Como resultado, el período dominante se incrementa con el incremento en la magnitud y la distancia.

Mecanismo de ruptura

El influjo del mecanismo de ruptura en las características de los temblores se revela en los contenidos de frecuencias y en las particularidades del movimiento del terreno. Es importante no perder de vista que a la fecha, los mecanismos de ruptura y su influencia en los temblores que pueden surgir en cierta fuente son imposibles de predecirlos en detalle. Lo que se sabe es que los diferentes ambientes algún rol desempeñan. Los ambientes de subducción son menos conocidos que los correspondientes a fallamientos del tipo horizontal.

La representación de los procesos que tienen lugar en su foco, está definida por los parámetros del mecanismo, que permiten establecer la orientación del plano de fractura. Estos están definidos por el azimut de su traza, el buzamiento y la dirección del movimiento sobre él.

Se ha observado que la intensidad de movimiento sísmico varía con la posición azimutal del sitio respecto a la fuente sísmica y con la dirección en que se propaga la ruptura a lo largo de la falla, especialmente en lo que se denomina el campo cercano. La propagación de la dislocación en determinada dirección genera ondas sísmicas que son registradas en forma diferente a ambos lados de la fuente. En sitios localizados en la dirección de propagación de la ruptura se registran mayores intensidades y frecuencias más altas, en cambio, en sitios localizados en el lado opuesto a la dirección de propagación se experimentan menores intensidades y frecuencias más bajas. Este fenómeno conocido como efecto de direccionalidad, se cree que se debe a que las ondas tiendan a concentrarse o polarizarse en la dirección en que se propaga la ruptura, en forma similar al efecto Doppler en las ondas acústicas.

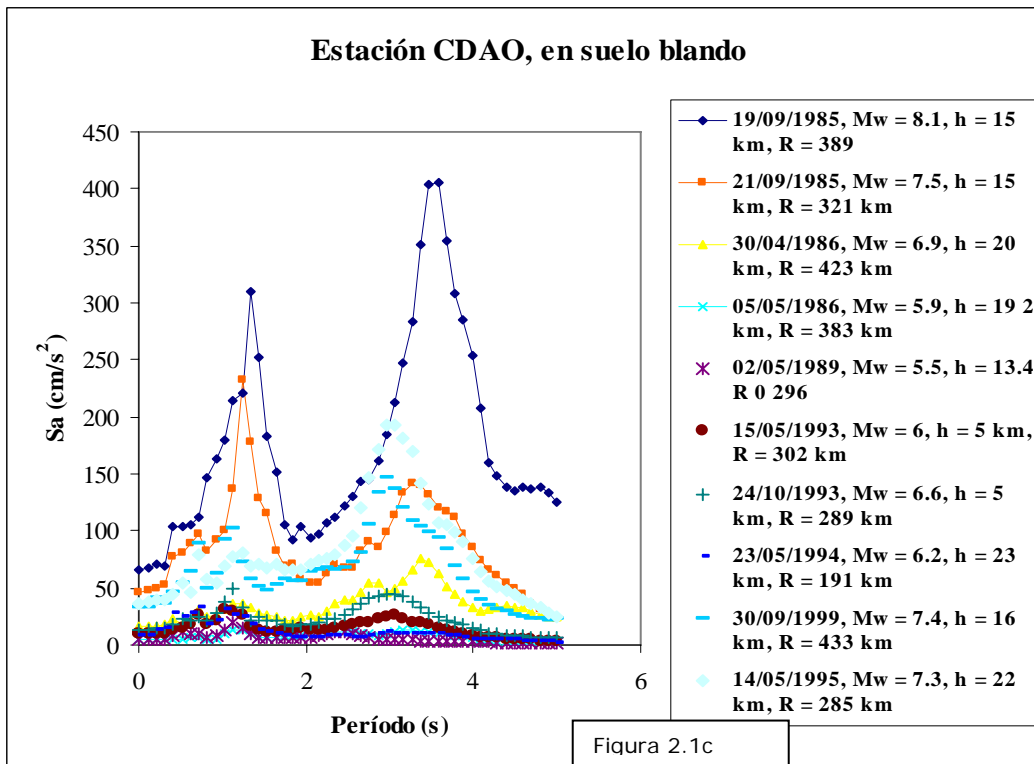
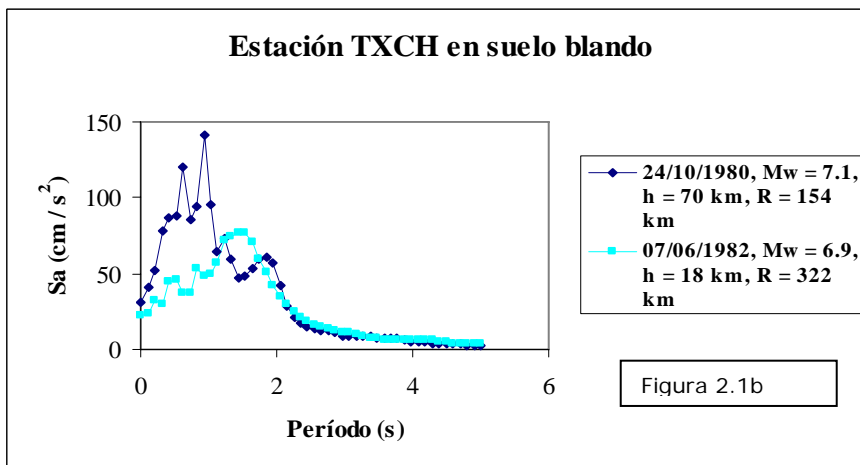
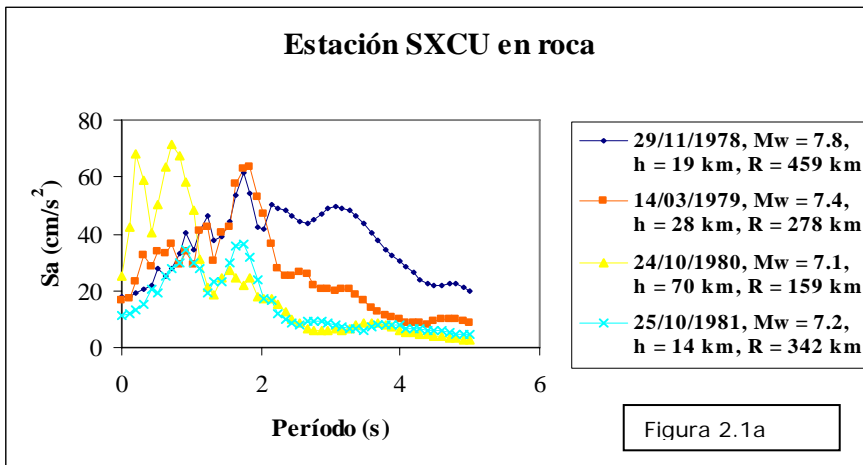


Figura 2.1 Contiene los espectros de respuesta obtenidos a partir de la señal original registrada en la ciudad de México para algunos temblores generados dentro de la República mexicana. a) Estación SXCUC localizada en roca. A juzgar por lo observable simplemente por inspección visual, en primera instancia se puede apreciar los efectos del tipo de fuente, el sismo de 24/10/1980 con foco localizado en el centro de México y originado en falla tipo normal con respecto a los demás eventos cuyos focos fueron localizados en la zona de subducción del oeste mexicano y con falla inversa, presenta el mayor valor de aceleración máxima para la menor magnitud y concentración en altas frecuencias. Comparando entre los tres eventos de fuentes similares el período predominante es alrededor del mismo valor. b) Estación TXCH localizada en suelo blando. Se aprecia los efectos de tipo de fuente, desplazamiento de amplitudes espectrales máximas hacia la zona de bajas frecuencias. c) Estación CDAO en suelo blando, resalta la similitud en la forma de los espectros y la dependencia de la amplitud con la magnitud es concluyente.

La evidencia de la influencia de tales factores, se ha identificado en la práctica a partir de registros que muestran características diferentes en estaciones cercanas, ubicadas en sitios con condiciones locales similares, o en una misma estación, a partir de eventos distintos originados en la misma fuente. Por ejemplo, cuando el ángulo que se forma entre un vector que va del sitio a la fuente es pequeño, el movimiento del terreno sufrirá un incremento en amplitud, que deberá observarse con mayor intensidad cerca de la fuente y con tendencia a decrecer lejos de ésta [Boatwright 82]. Singh [88] analizó diversos sitios en la ciudad de México localizados en terreno blando y encontró que las variaciones de temblor a temblor, para la mayoría de los sitios, son en general, moderados e independientes del temblor, prevaleciendo como característica relevante la frecuencia de los primeros modos de vibración del suelo en cada sitio.

La falta de información sobre el efecto de la variación azimutal en la intensidad respecto a la orientación de la falla y la dirección de propagación de la ruptura ha sido suplida parcialmente con modelos matemáticos que consideran la propagación de las ondas, el mecanismo de la fuente y la dirección en que se propaga la ruptura. De esta manera, conocido el mecanismo focal de un temblor, es posible calcular el patrón de radiación de una dislocación. En este caso y con fines demostrativos se ha simulado el modelo sísmológico propuesto en [Aki 80] para una fuente sísmica en el campo lejano, como un ejemplo que ilustran los efectos de directividad al considerar variaciones en los parámetros que definen el mecanismo y la geometría de la fuente. Esta simulación y sus resultados se presentan al final del capítulo.

Sin embargo, de acuerdo con lo reportado por Ordaz [92]: "las observaciones muestran que, por efectos de difracción, la energía de alta frecuencia es aproximadamente isotrópica. Lo que significa que, de corregirse por un patrón de radiación que no tome en cuenta este efecto, se podrían cometer errores graves en los puntos nodales para ciertas frecuencias". Según este mismo autor, por tal razón resulta más robusto y congruente con las observaciones usar un valor promedio del patrón de radiación independiente de la frecuencia.

Boore [89] considera que el efecto asociado con la directividad podría estar contenido en forma implícita en las variables derivadas directamente del registro sísmico (aceleración máxima, duración de la fase intensa, intensidad de Arias, espectros de respuesta coligados, etc.) y no requerir por tanto un parámetro adicional, por lo que señala que los modelos teóricos tienen la desventaja de no incluir los efectos azimutales.

2.2.2 Trayectoria de las ondas y factores asociados

La composición de frecuencias no es sólo función del mecanismo en la fuente sino también de la propagación de las ondas, específicamente de la estructura cortical entre la fuente y sitio y la distancia entre ambos. Las características de las ondas, durante su trayectoria, son afectadas por la geología y las propiedades físicas (densidad ρ , factor de calidad Q , velocidad de las ondas transversales β , etc.) de las formaciones existentes entre el foco y la región de interés. Durante su recorrido, las ondas tienen además un decremento en su amplitud debido principalmente a las refracciones y reflexiones que las afectan.

La distancia fuente-sitio como una variable que contribuye a identificar los efectos de la trayectoria de propagación de las ondas sísmicas, es un parámetro muy ambiguo. Varias medidas han sido propuestas para estimar el movimiento del terreno en superficie. No hay una base clara para escoger una función que represente la dependencia de la amplitud del movimiento con la distancia. Sin embargo, con el tiempo, la tendencia ha sido a usar la distancia epicentral por su disponibilidad [Abrahamson 97]. También, desde la perspectiva sísmológica se han acuñado los conceptos de campo lejano y campo cercano como términos cualitativos para relacionar la longitud de la trayectoria foco-sitio. Así, a partir de esta representación, un emplazamiento puede considerarse situado en el campo lejano cuando su distancia al origen es suficiente como para atenuar las ondas de cuerpo, y el campo cercano como la región situada a una distancia más pequeña que la longitud de la fractura.

La atenuación geométrica, como consecuencia de los efectos de amplificación del frente de onda, contribuye en diferentes proporciones a dicho decremento tanto en las ondas de cuerpo como de superficie. En el primer caso la reducción de la amplitud se supone proporcional al recíproco de la

distancia fuente-sitio, $1/R$ y en el segundo como el recíproco de la raíz cuadrada de la distancia foco-sitio $(1/R)^{1/2}$.

Como se mencionó, la distancia fuente-sitio es un parámetro que se prefiere sustituir por la distancia epicentral, que es más asequible. El período dominante de las ondas está asociado a la magnitud del evento y a la distancia epicentral. Según los referentes sismológicos, las ondas de cuerpo causan cerca del epicentro vibraciones de mayor amplitud y de alta frecuencia, del orden de 5 a 20 hertz (períodos dominantes muy cortos, 0.20 a 0.005 s), y la amplitud de su componente vertical es significativa. Para distancias mayores de 100 km o campo lejano, se considera que el movimiento fuerte deja de ser controlado por las ondas de período corto, en virtud de la existencia de múltiples reflexiones y refracciones a través de las capas que forman la corteza de la Tierra.

Así, la sismología cuantitativa indica que las altas frecuencias de las ondas sísmicas son filtradas cuando se propagan por el medio sólido de la tierra y se atenúan más rápidamente que las restantes, debido precisamente a que tienen un mayor número de ciclos. Es decir, con la distancia se reduce su amplitud más rápidamente que en las ondas de menor frecuencia. De este modo, a medida que éstas se alejan de la fuente, predominan las ondas de períodos largos. A distancias epicentrales grandes, las ondas de cuerpo pueden tener períodos dominantes de 2 a 3 segundos, mientras que para las ondas de superficie suelen tener períodos hasta de 20 segundos.

2.2.3 Efectos de sitio y factores asociados

Las condiciones locales de un lugar suelen ser un factor significativo en la respuesta superficial del terreno. La frecuencia dominante de las ondas en la última etapa de su trayectoria son determinadas por las condiciones propias del sitio y no por el mecanismo de falla ni la trayectoria previa de ellas.

En algunos casos, los efectos de sitio modifican los movimientos sísmicos incidentes como ocurrió con las enormes amplificaciones registradas en el temblor del 19 de septiembre de 1985 en la zona de lago de la Ciudad de México. También son los efectos de sitio los que explican la severidad de los movimientos generados por el terremoto de Armenia-Colombia, el 25 de enero de 1999, como consecuencia de un foco somero más la amplificación sísmica causada por suelos blandos y formas topográficas favorables para tal efecto. Estos ejemplos y muchos otros (p.e., Caracas- Venezuela, 1979; Loma Prieta-California, 1988; San Mónica, 1994; Kobe-Japón, 1995; Kocaeli-Turquía, 1999; etc.) acentúan la trascendencia de las condiciones locales e indican el potencial destructivo de estos efectos. Ordaz [92] reporta para el Valle de México amplificaciones de hasta 10 veces en el espectro de Fourier en zonas firmes con respecto a otros sitios fuera de valle que se encuentran a distancias equidistantes y, para la zona de lago con respecto a la zona firme en dicho espectro, amplificaciones de 50 veces, lo que implica un factor de amplificación total de 500 [Ordaz 92].

Es admitida la influencia de las condiciones del sitio en la intensidad del movimiento del terreno y la severidad del daño ocasionado en las diferentes estructuras. De acuerdo con la ley de conservación de la energía, el desplazamiento de las partículas asociadas a ondas sísmicas que se propagan en un cuerpo sólido se incrementa al pasar de un medio rígido a uno blando. Registros obtenidos en arreglos verticales, como resultados teóricos indican que la aceleración asociada con las ondas sísmicas aumentan en la superficie al transitar de una base rocosa a un estrato de sedimentos blandos; la variabilidad depende de la profundidad de la columna de suelo y de la relación de impedancia entre los dos medios [Seed 69, Seed 74, Romo 86a, Romo 86b, Seed 87, Silva 88, Romo 95].

Un factor, importante en este caso, es la posición relativa del foco con respecto al sitio de interés, pues el tipo y amplitud de las ondas sísmicas que causan la alteración está supeditado a éste. Cuando las ondas de cuerpo alcanzan la superficie, gran parte de su energía se refleja nuevamente hacia el interior de la tierra, generando en la superficie una amplificación de las ondas respecto al tren de ondas incidentes.

El hecho de que las ondas de cuerpo sufran múltiples refracciones al viajar a través de la corteza terrestre y su velocidad crezca con la profundidad, permiten suponer una propagación vertical cuando atraviesan estratos blandos horizontales. Sin embargo, no siempre es posible partir de dicho

indicio como se considera en los modelos unidimensionales basados en el uso de la ecuación de onda y la técnica del Elemento Finito en una dimensión, pues la aportación de otras componentes al movimiento depende de la relación entre la longitud de la onda superficial y el ancho de la depresión o cuenca sedimentaria de que se trate. Por ejemplo, si dicha relación es menor que la unidad el movimiento en la frontera podría ser no uniforme.

La amplitud y forma de los espectros de respuesta en función de los efectos locales está bien documentada [Seed 69, Seed 74], tal como se observa en la figura 2.2, tomada de Seed [74]. Las curvas mostradas resultan de promediar las ordenadas espectrales para cada período. La forma de los espectros de respuesta normalizados con respecto a la aceleración máxima del terreno para diferentes condiciones del sitio es un reflejo de ello. Se aprecia en la figura que para períodos cercanos a 0.5 s, las amplificaciones espectrales son superiores para suelo que para roca y para períodos más largos, la amplificación espectral aumenta con la disminución de la rigidez en los depósitos de suelo. Es decir, los depósitos de suelos blandos y profundos amplifican los movimientos en las bajas frecuencias. Los suelos extremadamente blandos como los sedimentos saturados de la Ciudad de México exhiben períodos fundamentales de vibración susceptibles de entrar en resonancia con ondas sísmicas de período largo y su amplificación es función de la magnitud del evento. Ovando y Romo [03] reportan el efecto del tiempo en las propiedades estáticas y dinámicas de los suelos compresibles de la zona de lago de la ciudad de México. Estos cambios influyen directamente en las características vibratorias de los depósitos de suelo. Consecuentemente en el peligro sísmico.

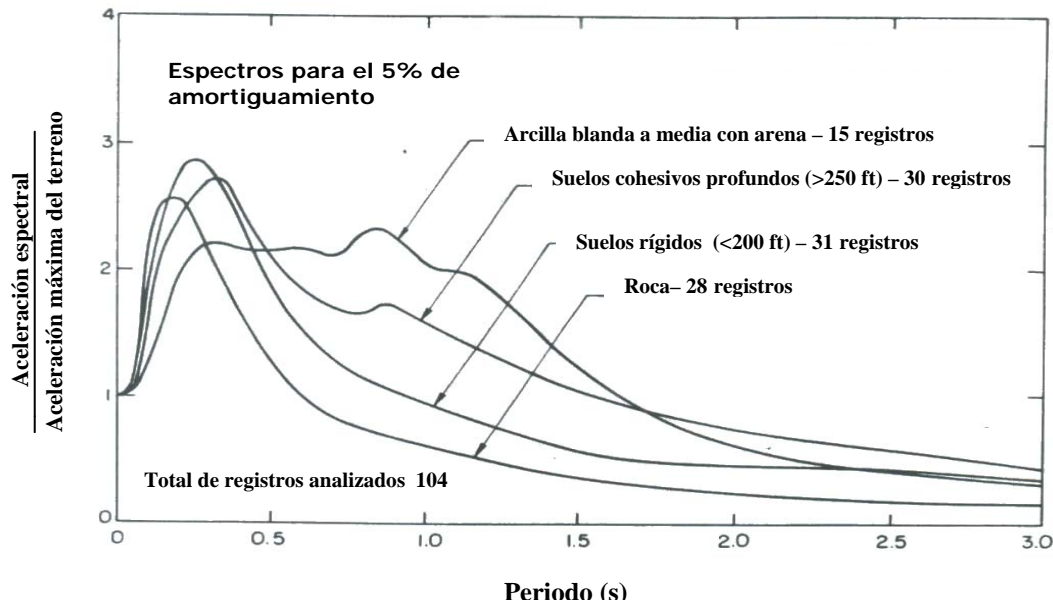


Figura 2.2 Promedio de espectros de aceleración para diferentes condiciones de sitio. Figura tomada de [Seed 74].

Relación de impedancia entre el manto superficial y el estrato subyacente

La aplicabilidad de modelos unidimensionales a suelos con pequeñas diferencias de impedancia entre el manto superficial y el lecho que lo soporta, es más limitada que para los casos en que dicha diferencia es elevada, puesto que en aquellos casos existe mayor incertidumbre respecto a la influencia de los tipos de ondas que dan lugar a un predominio de ondas de cortante con propagación vertical [Ruiz 77].

Para sitios donde el subsuelo exhibe un período dominante alto, hay diversos estudios que comprueban la validez de un modelo unidimensional. Pese a, existen inquietudes sobre las

condiciones en que las ondas superficiales dejan de ser significativas y pasa a ser dominante la propagación vertical de ondas de cortante. Todavía más, es claro que la razón por la que la teoría de propagación vertical de ondas unidimensionales de cortante produce buenos resultados, como en el caso del Valle de México, es la gran diferencia de impedancia entre los suelos superficiales y el terreno subyacente. En el caso mencionado, se sabe que, al menos para los sismos generados en la zona de subducción, las ondas que llegan son principalmente del tipo superficial, pues las de cuerpo deben haber sufrido una atenuación significativa con la distancia [Ruiz 77].

Geometría de la cuenca

El efecto de la geomorfología en la respuesta sísmica resulta difícil de evaluar. Sin embargo, se han observado concentraciones de ondas en la superficie de depósitos blandos, apoyados sobre roca cuando su frontera inferior tiene geometría irregular y las dimensiones de las irregularidades son comparables al espesor del manto superficial. Las irregularidades geométricas y la heterogeneidad de las formaciones obligan a estudiar el problema bidimensionalmente [Sánchez-Sesma 93].

La solución general del problema en un espacio bidimensional con procedimientos analíticos tiene grandes complicaciones matemáticas, por lo que la mayoría de los estudios incluyen simplificaciones como linealidad, homogeneidad e isotropía de los materiales, tipo de excitación estacionaria, frente de onda plano, aproximación de sistemas semi-infinitos mediante fronteras equivalentes, continuidad del medio, etc. No obstante, permiten la evaluación cualitativa de ciertos fenómenos.

Comportamiento del suelo ante cargas dinámicas

Idriss [90] relaciona la aceleración máxima del terreno en sitios de depósitos blandos y en sitios de roca. Sus resultados muestran que para niveles de aceleración menores a 0.4 g es probable que las aceleraciones máximas en los sitios de suelo blando sean mayores que en los sitios de roca o terreno firme. Las aceleraciones en roca relativamente bajas pueden causar aceleraciones altas en la superficie de los depósitos blandos, mientras que para niveles de aceleración mayores a 0.4 g casi no hay amplificaciones, debido básicamente al alto amortiguamiento de los suelos blandos en el intervalo de comportamiento no lineal. Para el Valle de México Singh [88] y Ordaz [92] presentan estudios que llevan a considerar que una gran porción de los suelos blandos permanece en el intervalo lineal, incluso bajo la fuerte excitación sísmica del temblor del 19 de septiembre de 1985, consideración que otros autores no comparten.

Topografía de la superficie

El fenómeno de amplificación observado en resaltos topográficos, conformados generalmente por roca, advierte que no sólo deben considerarse los espesores de suelo blando como principal contribuyente del fenómeno de amplificación, sino también las condiciones locales de frontera.

Las irregularidades de la superficie del suelo pueden alterar el movimiento que éste tendría si fuese horizontal, dando lugar a amplificaciones o atenuaciones locales y generales. La extensión de los efectos depende principalmente de la relación entre la longitud de onda sísmica y la dimensión de las irregularidades topográficas. Las irregularidades cuyas dimensiones son comparables a la longitud dominante de las ondas, pueden producir amplificación notable, mientras que cuando la longitud de onda es mucho mayor que la profundidad o altura de las anomalías, el efecto es despreciable [Sánchez-Sesma 93]. Para las ondas de alta frecuencia la topografía puede ser un factor muy importante en el movimiento del terreno.

Se han hecho modelaciones teóricas, con procedimientos analíticos en un espacio de más de una dimensión [Geli 88, Graves 93, Sánchez-Sesma 94, Huang 99]. Los resultados apuntan a que para irregularidades en forma de colinas y valles se presentan grados de amplificación diversos dependiendo del contenido de frecuencias, de la dirección y del ángulo de incidencia de las ondas sísmicas, con aumento de la amplitud del movimiento en los bordes de los valles, en la cresta y en las laderas de las colinas. Jibson [87] comprobó que en general, la amplificación en las crestas es mayor para frecuencias 2 a 4 hertz y que la variación en la intensidad del movimiento sísmico debido a asperezas topográficas puede ser muy grande.

Pese a que se reconoce la influencia de las formas topográficas sobre la intensidad del movimiento sísmico, todavía no se tienen los medios apropiados para obtener determinaciones cuantitativa detalladas, por lo que en los códigos de diseño antisísmico aun no se considera este aspecto en la evaluación de las fuerzas sísmicas a que estarían sujetos los edificios.

2.3 Criterios de evaluación

La mayoría de los modelos y funciones de predicción denominados leyes de atenuación sísmica se han desarrollado para los casos en que los efectos del mecanismo de falla se imponen sobre los efectos locales. Cuando ocurre esto, la amplitud del movimiento decae con la distancia, dando lugar al fenómeno de atenuación sísmica. Para estimar movimientos fuertes en terrenos donde el efecto de las condiciones locales impera, se hace uso de funciones de transferencia empíricas o cocientes espectrales.

Con base en registros se calculan las funciones de transferencia del suelo para cada estación. Tales funciones permiten evaluar aproximadamente las modificaciones que sufre el movimiento sísmico entre la base rocosa y la superficie libre del terreno, y se utilizan para calibrar los modelos teóricos que permiten estimar los movimientos en la superficie para movimientos en la base rígida más intensos que los registrados. Estas funciones han sido calculadas como cocientes entre espectro de amplitudes de Fourier, y/o cocientes entre espectros de respuesta en un mismo sitio de los registros en superficie y en roca, o mediante procedimientos de aproximación lineal equivalente y métodos numéricos. Los análisis de relación de respuesta espectral se suelen separar por niveles de aceleración en roca, con el fin de inferir efectos de no-linealidad del suelo.

Utilizando la técnica de los cocientes espectrales o funciones de transferencia empíricas, los datos de la red acelerográfica de la ciudad de México, han servido para medir las amplificaciones relativas de los sitios en zonas de lago y transición con respecto a los de terreno firme. El cociente representa una medida de la amplificación empírica en el dominio de la frecuencia y se ha considerado que refleja las características dinámicas del sitio como el período o frecuencia dominante.

2.3.1 Predicción empírica

Para estimar el movimiento en un determinado sitio en función de las variables que afectan la respuesta superficial del terreno, se han desarrollado relaciones empíricas obtenidas por regresión múltiple, a partir de un conjunto de datos, hasta lograr un curva ajuste.

Estas relaciones incorporan las observaciones y evidencias recopiladas durante años para establecer relaciones entre los parámetros del temblor y los de la intensidad en el sitio. La estructura del funcional es seleccionada de manera que refleje lo observado; así una forma general es

$$\ln Y = C_1 + C_2 M + C_3 M^{C_4} + C_5 \ln[R + C_6 \exp(C_7 M)] + C_8 R$$

donde: Y , parámetro ligado a la intensidad; M , magnitud; R , distancia; $C_1, C_2, C_3, C_4, C_5, C_6, C_7$ y C_8 coeficientes empíricos obtenidos a partir del análisis de regresión.

La incertidumbre del valor obtenido con la relación de predicción es evaluada a través del parámetro, $\sigma_{\ln Y}$, que representa la desviación estándar del $\ln Y$, con la magnitud y la distancia. Históricamente los valores $\sigma_{\ln Y}$ se han supuesto constantes, pero estudios recientes indican que varía con la magnitud [Abrahamson 97].

Para una magnitud dada, la probabilidad de que Y exceda un valor Y^* es definida como $1 - F_Z(z^*)$. Donde, $F_Z(z^*)$ es el valor de la función de distribución estándar acumulada y $z^* = (\ln Y^* - \ln \bar{Y}) / \sigma_{\ln Y}$

Durante décadas pasadas se llevaron a cabo varios estudios para estimar la aceleración pico con base en registros sísmicos. Por ejemplo, Campbell [81] usó datos de diferentes partes del mundo para desarrollar una ley de atenuación general de la aceleración máxima (a_{max}) en sitios distantes no más de 50 km de la falla, para temblores con magnitudes entre 5 y 7.7. Obtuvo la siguiente expresión

$$\ln a_{max}(g) = -4.141 + 0.868M - 1.09 \ln(R + 0.0606 \exp^{(0.7M)}); \quad \sigma_{\ln a_{max}} = 0.37 \quad (2.3)$$

Esta relación de atenuación para valores máximos de aceleración en función de la magnitud y la distancia y valores de $\sigma_{\ln Y}$ constantes representó el estado del arte en esa época.

Posteriormente el propio Campbell [94], amplió la base de datos de registros sísmicos y propuso una ley de atenuación afinada en la que incorpora términos adicionales para reflejar las características de la fuente y del sitio, lo que condujo a la ecuación (2.4).

$$\begin{aligned} \ln a_{max}(gales) = & -3.512 + 0.904M_W - 1.328 \ln \sqrt{R^2 + [0.149 \exp(0.647M_W)]^2} \\ & + (1.125 - 0.112 \ln R - 0.0957M_W)F + (0.440 - 0.171 \ln R)S_{SR} \\ & + (0.405 - 0.222 \ln R)S_{HR} \end{aligned} \quad (2.4)$$

$$\sigma_{\ln a_{max}} = \begin{cases} 0.889 - 0.0691M & M \leq 7.4 \\ 0.38 & M > 7.4 \end{cases}$$

donde, R es la mínima distancia en km a la falla causativa (≤ 60 km). El término F denota el tipo de fuente y toma un valor de 0 para falla normal y de 1 para falla inversa; los términos S_{SR} y S_{HR} describen las características estratigráficas del sitio

$S_{SR} = S_{HR} = 0$ para depósitos aluviales,
 $S_{SR} = 1$ para sitios suelo-roca (depósitos sedimentarios del terciario),
 $S_{HR} = 1$ para sitios en roca competente (sedimentaria, metamórfica y cristalina).

Las leyes de atenuación espectral para las variables del movimiento son una extensión a varios periodos del funcional preestablecido para valores máximos. El procedimiento consiste en calcular para cada periodo de la estructura los coeficientes correspondientes por medio de una regresión. A la fecha, se han derivado un sinnúmero de aproximaciones estadísticas como las anteriores. Ver por ejemplo Abrahamson [97].

Para México han sido de gran utilidad las relaciones empíricas propuestas por Esteva [74], Bufaliza [84], Singh [87], Castro [88], Ordaz [89] y Anderson [97]. Para Colombia, Ojeda [97] presenta tres funciones empíricas para el cálculo del componente horizontal de la aceleración máxima, considerando dos ambientes tectónicos diferentes. El primero incluye los datos de sismos que provienen de regiones tectónicas continentales y el segundo corresponde a los datos de la zona de subducción. No obstante la consideración simotectónica, los resultados obtenidos muestran una gran dispersión, lo que hace poco confiable su aplicación, como ellos mismos concluyen.

2.3.2 Predicciones teóricas y semiempíricas

Los modelos teóricos que representan la generación, la transmisión y la disipación de la energía sísmica, tales como lo de Haskell, el modelo de grieta, el de falla circular, el de barrera, el de aspereza, el de fuente de Brune, etc., han proporcionado las bases para establecer relaciones de atenuación. Sin embargo, la generación de temblores es un proceso complejo cuya incertidumbre obliga a considerar las variables en juego como aleatorias, por lo que estos procedimientos deben ajustarse a la teoría de probabilidades [Abrahamson 92, Ordaz 92].

Con los modelos de fuente estocástica las predicciones teóricas se apegan más a los registros sísmicos. Dos modelos de fuente estocástica que han sido ampliamente implementados con éxito en la predicción de movimientos sísmicos son el modelo de omega cuadrada [Hanks 81] y el modelo de barrera [Papageorgiou 83].

Para la predicción del movimiento del suelo con modelos de fuente estocástica, se utilizan la simulación de Monte Carlo en el dominio del tiempo y la teoría de vibraciones aleatorias en el de la frecuencia. Los métodos se complementan; los detalles se pueden consultar en Boore [83, 84, 86, 87].

En ambos modelos la amplitud del espectro del movimiento del suelo está expresada como una función de la frecuencia del tipo

$$R(f) = C * S(f) * A(f) * D(f) * I(f) \quad (2.5)$$

donde los factores C , S , A , D e I representan, respectivamente, un factor de escala, el espectro de fuente, un factor de amplificación, un factor de atenuación y un factor de la respuesta del instrumento.

El factor de escala es expresado en términos de las siguientes variables

$$C = \frac{R_{\Theta\Phi} F V}{4\pi\rho_0\beta_0^3 R} \quad (2.6)$$

En la anterior ecuación, $R_{\Theta\Phi}$ es el patrón de radiación promedio sobre un rango apropiado del azimut y el ángulo de salida [Boore 84]; F , factor de corrección que toma en cuenta la amplificación de la onda por efecto de superficie libre (usualmente se aplica un factor de 2); V representa la partición de la energía que se genera en la fuente en dos componentes horizontales perpendiculares supuestas de igual amplitud $(1/\sqrt{2})$; ρ_0 es la densidad de masa del material en que se propaga la onda; β_0 la velocidad de las ondas de corte en la región de la fuente; y R es el factor de atenuación geométrico. Para las ondas de cuerpo dentro del campo cercano (100 km) $R = r$, donde r es la distancia hipocentral. Para distancias más allá de la fuente cercana, R se considera proporcional a \sqrt{r} .

El factor de fuente $S(f)$ de acuerdo con [Hanks 81], es definido como

$$S(f) = \frac{M_0}{\left[1 + \left(\frac{f}{f_0}\right)^2\right]} \quad (2.7)$$

M_0 denota el momento sísmico, f_0 la frecuencia de esquina y f es la frecuencia. Estos dos parámetros se relacionan mediante la siguiente expresión [Brune 70]

$$f_0 = 4.9 \times 10^6 \beta_0 (\Delta\sigma / M_0)^{1/3} = 2.34\beta_0 / 2\pi r_0 \quad (2.8)$$

En la ecuación anterior $\Delta\sigma$ es la caída de esfuerzos medida en bares, r_0 es una dimensión característica del área de ruptura y β_0 está dada en km/s y M_0 en dina-cm

Una versión modificada de $S(f)$ fue proporcionada por Joyner [84], para ajustar la dislocación de forma similar a la que debe ocurrir cuando el momento excede al crítico, M_{0c} , correspondiente a la zona de ruptura de todo el ancho del sismograma

$$S(f) = \frac{M_0}{\left(1 + \frac{if}{f_B}\right)^{1/2}} \quad f \leq f_A, \text{ donde } i = \sqrt{-1} \quad (2.9)$$

$$S(f) = \frac{M_0 \left(\frac{f_A}{f}\right)^{3/2}}{\left(1 + \frac{if}{f_B}\right)^{1/2}} \quad f \geq f_A \quad (2.10)$$

Aquí

$$f_A = 4.9 \times 10^6 \beta_0 \lambda^{-1/4} \left(\frac{\Delta\sigma}{M_0}\right)^{1/3} \quad M_0 \leq M_{0c} \quad (2.11)$$

$$f_B = 4.9 \times 10^6 \beta_0 \lambda^{3/4} \left(\frac{\Delta\sigma}{M_0}\right)^{1/3} \quad M_0 \leq M_{0c} \quad (2.12)$$

$$f_A = 4.9 \times 10^6 \beta_0 \lambda^{-1/4} \Delta\sigma^{1/3} M_{0c}^{1/6} M_0^{-1/2} \quad M_0 \geq M_{0c} \quad (2.13)$$

$$f_B = 4.9 \times 10^6 \beta_0 \lambda^{3/4} \left(\frac{\Delta\sigma}{M_0}\right)^{1/3} \quad M_0 \geq M_{0c} \quad (2.14)$$

λ es la razón entre el largo y el ancho de la falla, cuyo valor típico es de 4 [Joyner 84].

El factor de amplificación puede ser estimado por diferentes caminos. El más usual es la función de transferencia dependiente de la frecuencia que resulta de la propagación de la onda en una secuencia de capas [Boore 84]. La amplificación puede también ser representada en términos de la impedancia del sitio. La conservación de la energía requiere que la amplitud se incremente como consecuencia del decremento de la impedancia entre la región de la fuente y el sitio de registro. El factor de amplificación puede aproximarse a la siguiente relación $\sqrt{\rho_0 \beta_0 / \rho_r \beta_r}$, donde el subíndice r se refiere al material próximo al sitio del registro y el subíndice 0 al material próximo a la fuente.

El factor de atenuación es descrito como

$$D(f) = \exp\left[\frac{-\pi f r}{Q(f)\beta}\right] P(f) \quad (2.15)$$

donde Q es una función de atenuación que depende de la frecuencia, β la velocidad de propagación promedio sobre el trayecto, y P es un filtro pasa alta.

De acuerdo con Aki [en Joyner 88], esta función de atenuación puede definirse como

$$Q = 29.4 \frac{1 + \left(\frac{f}{0.3}\right)^{2.9}}{\left(\frac{f}{0.3}\right)^2} \quad (2.16)$$

El filtro, P , es utilizado en el factor de atenuación, para explicar el decaimiento en altas frecuencias observado en el espectro de aceleración. En general, los espectros muestran una brusca disminución de la energía en altas frecuencia, arriba de la frecuencia de corte, f_m , [Hanks 82]. Este filtro puede ser representado por el filtro Butterworth [Boore 83], o también mediante la siguiente expresión

$$P(f) = \exp(-\pi\kappa_0 f) \quad (2.17)$$

En la expresión anterior la forma y la notación es de Anderson [84]. Para distancias pequeñas y grandes valores de Q , el factor $P(f)$ contribuye a la mayor parte de la atenuación, el filtro $D(f)$ es aproximadamente equivalente a un filtro de corte pasa-alta con frecuencia de corte f_m $(1/\pi) \kappa_0$ y κ_0 es una constante que depende del sitio de observación.

Para dar forma a los espectros, se recurre al factor de filtro $I(f)$. Específicamente, para el espectro de aceleración este factor se representa como

$$I(f) = (2\pi f)^2, \text{ donde } i = \sqrt{-1} \quad (2.18)$$

Una vez calculado el espectro $R(f)$ haciendo uso de la teoría del valor extremo pueden obtenerse los valores pico del parámetro seleccionado. El método se basa en el trabajo de Cartwright y Longuet-Higgins [1956].

Se calculan los momentos estadísticos m_0 , m_2 y m_4 del espectro de densidad de energía por medio de la siguiente ecuación

$$m_k = \frac{1}{\pi} \int_0^{\infty} \omega^k |r(f)|^2 d\omega \quad \text{donde } \omega = 2\pi f \quad (2.19)$$

La raíz cuadrada del valor medio cuadrático rms está dada por

$$y_{rms} = (m_0 / T_r)^{1/2} \quad (2.20)$$

T_r denota la duración de la respuesta, y es igual a T_w para el cálculo de la aceleración pico, donde

$$T_w = \frac{1}{f_0} + 0.05r \text{ para el modelo de Hanks-McGuire} \quad (2.21)$$

$$T_w = \frac{1}{f_A} + 0.05r \text{ para el modelo modificado de Joyner} \quad (2.22)$$

La determinación de T_r para la predicción de los valores del espectro de respuesta es descrita más adelante.

El valor esperado del parámetro pico $[E(y_{max})]$ puede calcularse mediante una formulación exacta o asintótica dependiendo del valor del parámetro banda-ancha ξ y del valor extremo, N

$$\xi = \frac{m_2}{(m_0 m_4)^{1/2}} \quad (2.23)$$

N es el más grande entero contenido en el número $2\tilde{f}T_\omega$, donde

$$\tilde{f} = \frac{1}{2\pi} (m_4 / m_2)^{1/2} \quad (2.24)$$

Si N es menor que $14.7/\xi - 8$, entonces el valor esperado del parámetro pico es dado por la siguiente expresión

$$E(Y_{\max}) = y_{rms} \left(\sqrt{\frac{\pi}{2}} \sum_{l=1}^N (-1)^l \frac{C_l^N}{\sqrt{l}} \xi^l \right) \quad (2.25)$$

donde C_l^N son los coeficientes binomiales ($= N! / (l!(N-l)!$). Para valores grandes de N podrían presentarse problemas numéricos con la fórmula exacta.

Para la solución asintótica

$$E(Y_{\max}) = y_{rms} \left\{ \left[2 \ln(N) \right]^{1/2} + \gamma / \left[2 \ln(N) \right]^{1/2} \right\} \quad (2.26)$$

γ es la constante de Euler (0.577216). En la aproximación asintótica las cantidades \tilde{f} y N son recalculadas mediante las ecuaciones

$$\tilde{f} = \frac{1}{2\pi} (m_2 / m_0)^{1/2} \quad (2.27)$$

$$N = 2\tilde{f}T_r \quad (2.28)$$

En sismos pequeños a moderados, o a bajas frecuencias de oscilación o alto amortiguamiento, la duración del movimiento podría no ser suficiente para generar una respuesta pseudo-estacionaria. Boore y Joyner [84b] han desarrollado una corrección empírica para la duración de la respuesta T_r

$$T_r = T_\omega + \frac{T_0}{2\pi\zeta} \left(\frac{\gamma^3}{\gamma^3 + 1/3} \right) \quad (2.29)$$

donde T_0 es el período de un oscilador simple en segundos, ζ el amortiguamiento del oscilador expresado como una fracción del crítico, y $\gamma = T_\omega / T_0$

Una limitante del modelo descrito consiste en que sólo es aplicable a sitios cuya distancia epicentral sea grande comparada con las dimensiones de la fuente.

Joyner [88] desarrolló un método de simulación estocástica que es aplicable a distancias pequeñas. Se empieza por generar funciones que representan la superficie de ruptura en cada punto para todo el deslizamiento, usando el método de Monte Carlo para obtener el espectro de la fase aleatoria. El espectro tiene amplitud constante definida por el momento sísmico M_0 y los valores críticos son controlados por el parámetro $\Delta\sigma$. La caída de esfuerzos se supone igual a f_A/v , donde f_A es definida por las ecuaciones (2.11 a 2.13) y v es la velocidad de ruptura. La función que representa el total de movimiento es calculada mediante el inverso de la transformada de Fourier multiplicada por el filtro $1/(1+if/f_B)^{1/2}$ para obtener el espectro $S(f)$. El espectro se multiplica por los factores C , A , D e I definidos anteriormente y transformados al dominio del tiempo para la estimación del parámetro pico.

En esta dirección, para el caso de México, Ordaz [92] propuso tres métodos para estimar el movimiento fuerte ante temblores de subducción con diferentes distancias epicentrales. Las expresiones son de aplicación directa para la estimación probabilista de las amplitudes del espectro de Fourier (aceleraciones). Los modelos son semiempíricos, derivados de consideraciones teóricas del modelo de dislocación de Haskell y los parámetros fueron ajustados para reproducir las características de los acelerogramas. Para la zona epicentral se usó un modelo que toma en cuenta la naturaleza finita del área de ruptura de grandes temblores desarrollado por Singh [89]. Gallego [00] utilizó este mismo procedimiento de espectros fuente para construir leyes de atenuación en términos de aceleración máxima para Colombia. A fin de tomar en cuenta las condiciones del valle de México, Ordaz [92] acude a funciones de transferencia empíricas.

Ordaz [94] deriva relaciones de atenuación en términos del espectro de amplitudes de Fourier (aceleración) para la estación CU de la ciudad de México aplicando la técnica de regresión lineal bayesiana, la cual consiste en tomar como variables aleatorias los parámetros de la ley de atenuación. Se estiman a priori los valores esperados de los parámetros, según el conocimiento previo que se tenga de estos, es decir, sin considerar los datos registrados, y luego se actualizan con el teorema de Bayes incorporando los datos registrados. Este procedimiento condujo a una solución numérica que es más estable que la lograda con procedimientos de regresión convencional.

Utilizando la forma espectral normalizada y una ley de atenuación espectral Reyes [99] estimó el espectro de respuesta para la estación CU de la ciudad de México. El funcional incorporado es el propuesto por Joyner [88]; la regresión empleada es como la usada por Ordaz [94]. Además, para estimar el espectro de respuesta en la zona de terreno blando del Valle de México, utilizó la técnica de coeficientes de espectro de respuesta CER, tomando como punto de control la estación CU.

Sarmiento [01] reprodujo la respuesta sísmica de los depósitos de suelos localizados en las zonas geotécnicas de Lago y Transición del valle de México, aplicando redes neuronales artificiales al procedimiento empírico de función de transferencia, tomando como base para el entrenamiento de la red la media de los espectros de respuesta registrados en la estación CU.

2.3.3 Duración del movimiento fuerte

A partir del registro sísmico se ha inferido la relación directa que la duración de los eventos sísmicos tiene con la magnitud del temblor y la distancia epicentral. Cerca del foco, el movimiento del terreno puede durar pocos segundos y en sismos de magnitud baja el registro está representado por unas pocas oscilaciones. A distancias más allá del campo cercano y a medida que las ondas sísmicas se dispersan al alejarse de la fuente, el movimiento puede durar varios minutos si es generado por un sismo de gran magnitud, mientras en la fase en que dominan las ondas de corte la duración se mantiene constante [Joyner 88].

La duración en un acelerograma depende del tipo de instrumento con el que se registre. La confiabilidad de los registros depende de las amplitudes y duraciones mínimas de las señales impuestas por los umbrales de disparo de los aparatos, y por los problemas de digitalización consecuencia de la baja ganancia de los instrumentos.

Desde la óptica ingenieril, el interés radica en la duración de la fase intensa. No es posible definir la duración del movimiento fuerte como una simple ventana de tiempo entre el inicio y el final del acelerograma. Numerosas propuestas existen en la literatura para aislar la porción del acelerograma representativa del movimiento fuerte. Bommer [96] ha clasificado algunas de estas definiciones en cuatro categorías, ver tabla 2.4 de [Bommer 00]:

- Categoría 1. Duración puntual: es simplemente el intervalo entre la primera y la última excursión de un umbral de amplitud.
- Categoría 2. Duración uniforme: es la suma de todos los intervalos de tiempo durante el cual la amplitud del registro está por encima de un cierto umbral.
- Categoría 3. Duración significativa: se basa en el intervalo de duración en el cual una porción del total de la intensidad de Arias es acumulada.

- Categoría 4. Duración de la respuesta estructural: determinada por la aplicación de una de las tres definiciones anteriores a la respuesta de un oscilador libre de un grado de libertad.

El umbral usado para definir la duración de las categorías puede ser un valor absoluto de la aceleración, o un valor relativo establecido como una porción del máximo. Es importante resaltar que para un cierto acelerograma, la duración establecida de acuerdo con las diferentes definiciones puede variar en un factor mayor de 2.

Usualmente, se acepta como definición de duración el lapso entre la primera y última excursión de aceleración que sobrepase un cierto nivel de aceleración. En este estudio, se consideró la definición de duración propuesta por Trifunac [75], que corresponde a la porción de la señal que se desarrolla entre el 5% y 95% de la intensidad de Arias en un acelerograma. No obstante, la falta de unificación de criterios para definir la fase intensa del movimiento sísmico no permite una comparación realista entre los diversos enfoques.

Ordaz [87] analizó registros obtenidos en el Valle de México y su conclusión es que la duración de Arias es una medida adecuada. Por otro lado, Singh [90] encontró que la duración de la ruptura no difiere mucho del inverso de la frecuencia esquina.

La intensidad de Arias es un parámetro característico del movimiento fuerte del terreno por ser una medida de la intensidad basada en la capacidad de daño de un sismo, independientemente de la presencia o no de estructuras en la zona.

Arias encontró que la intensidad en una cierta dirección y evaluada en un punto, es proporcional a la integral en el tiempo de la aceleración al cuadrado, según la siguiente expresión

$$I_{xx}(x) = f(x) \int_{0.1}^{t_0} a_x^2(t) dt \quad (2.30)$$

donde $I_{xx}(x)$ = intensidad en un punto a lo largo del eje x ; $f(x) = \cos^{-1}[1/(1-\xi^2)^{1/2}]$
 ξ = amortiguamiento con respecto al crítico; t_0 = duración total del acelerograma
 $a_x(t)$ = aceleración del suelo según el eje x en el instante t; g = aceleración de la gravedad

En general, los valores de amortiguamiento requeridos en los problemas prácticos varían entre 0.02 y 0.20, por lo que $f(x)$ para esta consideración se encuentra en el intervalo 1.4 a 1.57; valores que pueden aproximarse a $\pi/2$. En consecuencia, la intensidad a lo largo del eje es prácticamente una constante. Arias, para evaluar la intensidad, consideró un amortiguamiento con respecto al crítico igual a cero.

$$I_{xx} = I(0) = \frac{\pi}{2g} \int_0^{t_0} a_x^2(t) dt \quad (2.31)$$

Considerando aceleraciones según tres componentes ortogonales $a_x(t)$, $a_y(t)$, $a_z(t)$; se puede obtener un tensor simétrico \mathbf{J} , el cual Arias denominó tensor de intensidades en el punto considerado.

La traza de la matriz \mathbf{J} es un invariante para una rotación de los ejes coordenados, lo cual lo cataloga como un escalar bien definido en cada punto donde se logre medir. Arias deduce la intensidad escalar en un punto a partir de la siguiente expresión

$$I = I_{xx} + I_{yy} + I_{zz} = \frac{\pi}{2g} \int_0^{t_0} (a_x^2 + a_y^2 + a_z^2) dt \quad (2.32)$$

Relaciones empíricas para predecir la duración del movimiento fuerte han sido propuestas por Esteva [64], Bolt [73], Trifunac [75, 82], Novikova [93], Singh [93], Trifunac [95] y Reinoso [01].

Esteva [64] encuentra que la esperanza de duración D de un sismo en segundos, se puede estimar en función de la magnitud del sismo (M) y la distancia focal del sitio (R), mediante de la siguiente ecuación

$$D = 0.02e^{0.74M} + 0.3R \quad (2.33)$$

Recientemente, Reinoso [01] proporciona la expresión (2.34) para la duración de la fase intensa, en términos de la magnitud, la distancia al área de ruptura y el periodo dominante, usando registros acelerográficos de sismos en la costa mexicana del Pacífico.

$$D = 0.01e^M + (0.036M - 0.07)R + (4.8M - 16)(T_s - 0.5) \quad (2.34)$$

En la ecuación anterior, los términos del primero al tercero representan respectivamente, los efectos de la fuente, la distancia y el sitio en la duración del movimiento. Para sitios en roca cuyo periodo dominante resulta menor que 0.5 s en la ecuación T_s , debe considerarse igual a 0.5.

Para el caso colombiano no ha sido documentado ningún trabajo que permita predecir esta cantidad.

Tabla 2.4 Clasificación de duración del movimiento fuerte

Definición de duración	Puntual		Uniforme		Significativa		Estructural	
	Relativa	Absoluta	Relativa	Absoluta	Relativa	Absoluta	Relativa	Absoluta
Rosenblueth & Bustamante (1962)							o	
Housner (1965)		•						
Ambraseys & Sarma (1967)		•						
Husid (1969)					•			
Page <i>et al.</i> , (1972)		•						
Donavan (1972)					•			
Bolt (1973)		o		o				
Housner (1975)					•			
Trifunac & Brady (1975)					•			
Apitkaev (1977)	•							
Trifunac & Westermo (1977)					o			
Saragoni (1977)					•			
McGuire & Barnhard (1979)	•	•			•			
McCann & Shah (1979)					•			
Pérez (1980)								o
Takizawa & Jennings (1980)					•			
Vanmarcke & Lai (1980)					•			
Zahrah & Hall (1984)								o
Zhou & Xie (1984)					•			
Theofanopoulos & Drakopoulos (1986)					•			
Elghadamsi <i>et al.</i> , (1988)					•			
Xie & Zhang (1988)								o
Mohraz & Peng (1989)					o			
Theofanopoulos & Watabe (1989)					•			
Kawashima & Aizawa (1989)	•	•						
Sarm & Casey (1990)			•					
Papazachos <i>et al.</i> , (1992)		•						
Bommer & Martinez- Pereira (1996)					•	•		
Somerville <i>et al.</i> , (1997)					•			
Wembo & Kezhong (1997)	•							
Safak (1998)							o	
Bommer & Martinez- Pereira (1999)						•		

(•) duración independiente de la frecuencia (o) duración dependiente de la frecuencia

2.4 Modelo sismológico de atenuación sísmica en el campo lejano

Como un ejemplo para ilustrar los efectos de directividad en función de los parámetros focales que definen el mecanismo y la geometría de la fuente, se presentan los resultados de la simulación teórica realizada con base en el modelo sismológico para una fuente sísmica en el campo lejano propuesto en [Aki 80].

2.4.1 Hipótesis

Fuente sísmica. Se considera la dislocación como una fuente puntual en la que actúan sistemas de fuerzas que representan la fractura. La geometría de la falla se establece a partir de los parámetros de orientación del plano de fractura: el azimut de su traza (strike), el buzamiento del plano (dip), la dirección en que ha tenido lugar el movimiento (slip), según se ilustra en la figura 2.3 por medio de un modelo hipotético.

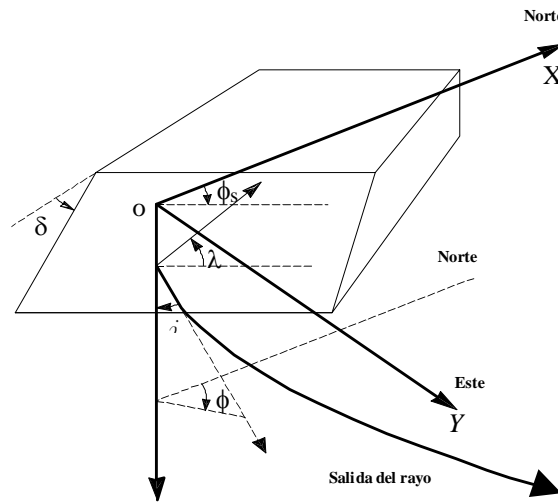


Figura 2.3 Configuración esquemática de un plano de falla que ilustra los parámetros focales: azimut (strike) ángulo ϕ_s , buzamiento (dip) ángulo δ , dirección en que ha tenido lugar el movimiento (slip) ángulo λ , ángulo de salida del rayo i y el azimut entre la fuente y la estación de registro ϕ .

Dentro del marco de esta premisa, un modelo elemental para representar el mecanismo de un sismo es definido por una fractura de cizalla de superficie plana, cuyas dimensiones son pequeñas comparadas con la distancia al punto de observación. Así, una fractura de cizalla se equipara a un sistema de dos pares de fuerzas ortogonales sin momento resultante, sistema conocido como doble par de fuerzas.

Patrón de radiación. Se asume el diagrama de radiación de las amplitudes de las ondas P y S para un doble par de fuerzas en una fuente puntual. Establecido por una serie de lóbulos, donde la forma del patrón de radiación está en función del azimut entre la fuente y la estación de registro ϕ , el ángulo i que definen la dirección de salida del rayo, y los ángulos ϕ_s , δ y λ para una dislocación con orientación arbitraria.

La distribución azimutal de amplitudes queda así relacionada con la geometría de la fuente por medio de las siguientes expresiones

$$F^P = \text{sen}^2 i (\cos^2 \phi M_{xx} + \text{sen} 2\phi M_{yy} + \text{sen}^2 \phi M_{yy} - M_{zz}) + 2\text{sen } i \cos i (\cos \phi M_{xz} + \text{sen} \phi M_{yz}) \quad (2.35)$$

$$F^{SV} = \text{sen } i \cos i (\cos^2 \phi M_{xx} + \text{sen} 2\phi M_{xy} + \text{sen}^2 \phi M_{yy} - M_{zz}) + \cos 2i (\cos \phi M_{xz} + \text{sen} \phi M_{yz}) \quad (2.36)$$

$$F^{SH} = \text{sen } i [\text{sen} \phi \cos \phi (M_{yy} - M_{xx}) + \cos 2\phi M_{xy}] + \cos i [\cos \phi M_{yz} - \text{sen} \phi M_{xz}] \quad (2.37)$$

donde

$$\begin{aligned} M_{xx} &= -(\text{sen} \delta \cos \lambda \text{sen} 2\phi_s + \text{sen} 2\delta \text{sen} \lambda \text{sen}^2 \phi_s) \\ M_{xy} &= M_{yx} = (\text{sen} \delta \cos \lambda \cos 2\phi_s + \text{sen} 2\delta \text{sen} \lambda \text{sen} \phi_s \cos \phi_s) \\ M_{yy} &= (\text{sen} \delta \cos \lambda \text{sen} 2\phi_s - \text{sen} 2\delta \text{sen} \lambda \cos^2 \phi_s) \\ M_{xz} &= M_{zx} = -(\cos \delta \cos \lambda \cos \phi_s + \cos 2\delta \text{sen} \lambda \text{sen} \phi_s) \\ M_{yz} &= M_{zy} = -(\cos \delta \cos \lambda \text{sen} \phi_s - \cos 2\delta \text{sen} \lambda \cos \phi_s) \\ M_{zz} &= \text{sen} 2\delta \text{sen} \lambda \end{aligned} \quad (2.38)$$

Propagación. La propagación se considera a través de la teoría clásica de rayos, en donde el rayo sísmico está constituido por los elementos infinitamente pequeños de las normales al frente de onda. La ecuación diferencial para el rayo en coordenadas espaciales viene dada por la expresión

$$\frac{d}{ds} \left(\frac{1}{c} \frac{d\mathbf{x}}{ds} \right) = \nabla \left(\frac{1}{c} \right) \quad (2.39)$$

aquí c es la velocidad de transmisión de las ondas del medio y \mathbf{x} el vector de posición que define la trayectoria del rayo.

Atenuación geométrica. La atenuación geométrica es definida como la disminución de la energía que transportan las ondas sísmicas en el interior de la tierra, la cual se esparce sobre un espacio cada vez mayor, es decir, la densidad de energía va disminuyendo a medida que la onda se aleja de la fuente. Ésta es evaluada como $1/R$, donde R es función del área transversal de un tubo definido por las trayectorias de rayos que parten con ángulos muy cercanos en la fuente. Se incrementa con la distancia entre la sismofuente y el punto de observación.

Disipación de la energía. La disipación de la energía se introduce a partir de un coeficiente de atenuación física que depende de la naturaleza de los materiales que componen el medio. Esta disminución se asume como se plantea en la literatura, con un decremento temporal exponencial, y se expresa en términos de la frecuencia, ω ; el tiempo, t ; y el factor de calidad, Q ; de la siguiente forma

$$e^{-\frac{\omega t}{2Q}} \quad (2.40)$$

2.4.2 Cálculo del desplazamiento

Usando la formulación presentada en [Aki 80] y teniendo en cuenta la teoría rayos, se puede calcular el campo de desplazamiento, considerando las hipótesis enunciadas mediante las siguientes expresiones

$$U_k^P(x, t) = \frac{1}{4\pi\rho_s\alpha_s^3} \frac{1}{R} F^P \dot{M}_{pq} \left(t - \frac{R}{\alpha_s} \right) \left(e^{-\frac{\omega t}{2Q}} \right) \quad (2.41)$$

$$U_k^{SV}(x, t) = \frac{1}{4\pi\rho_s\beta_s^3} \frac{1}{R} F^{SV} \dot{M}_{pq} \left(t - \frac{R}{\beta_s} \right) \left(e^{-\frac{\omega t}{2Q}} \right) \quad (2.42)$$

$$U_k^{SH}(x, t) = \frac{1}{4\pi\rho_s\beta_s^3} \frac{1}{R} F^{SH} \dot{M}_{pq} \left(t - \frac{R}{\beta_s} \right) e^{-\frac{\omega t}{2Q}} \quad (2.43)$$

donde ρ_s densidad del medio; α_s , β_s velocidad de las ondas longitudinales y transversales respectivamente.

2.4.3 Simulación teórica

El resultado de la simulación teórica se muestra en las figuras 2.4, en la que ilustra la variación espacial de la intensidad sísmica con los parámetros focales. En estas imágenes se puede observar un patrón de radiación que disminuye con la distancia, pero la forma de la atenuación varía en cada caso en función de los valores de las cantidades geométricas que definen la fuente, reflejando la anisotropía en la atenuación.

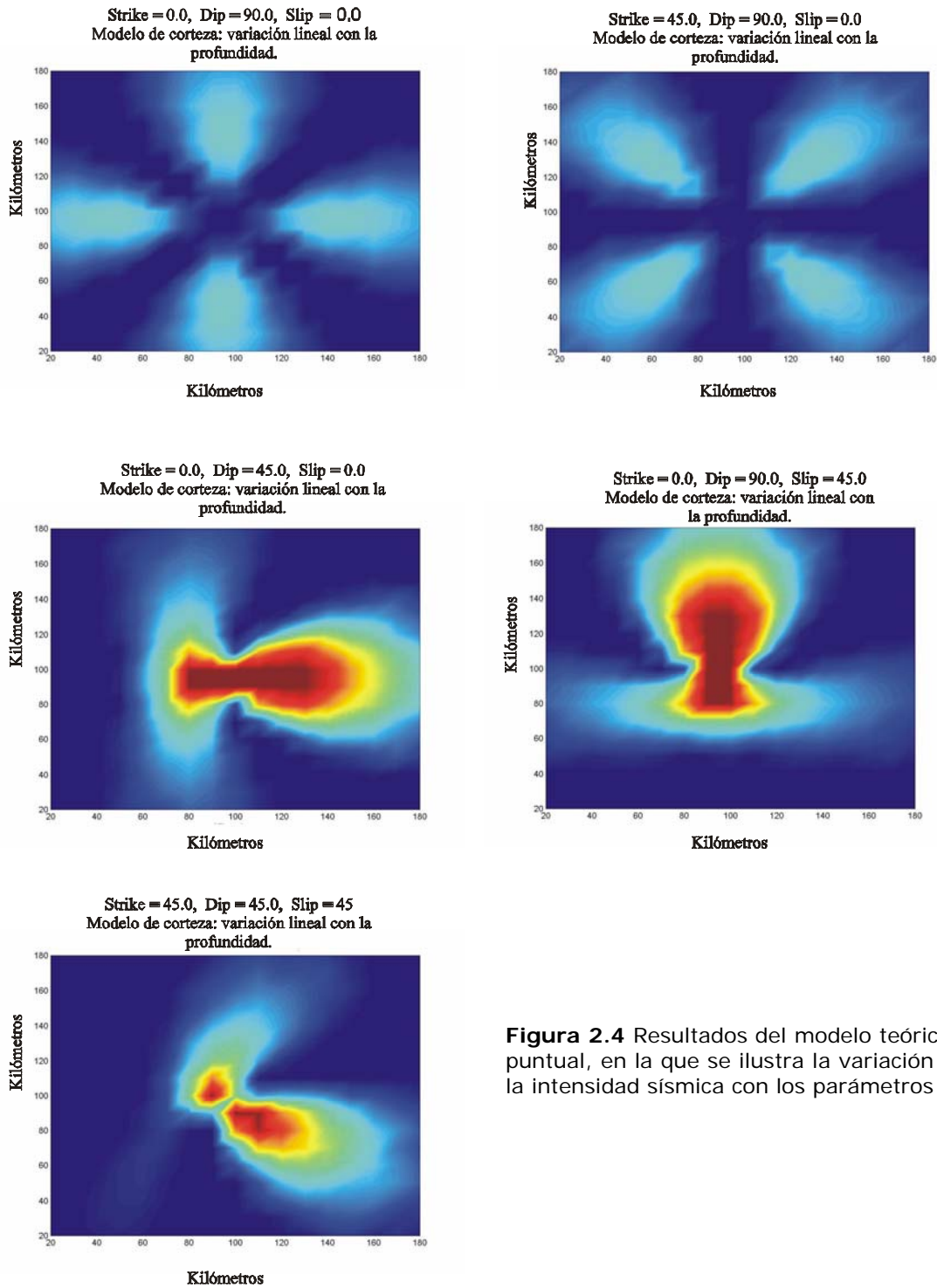


Figura 2.4 Resultados del modelo teórico de fuente puntual, en la que se ilustra la variación espacial de la intensidad sísmica con los parámetros focales.

Capítulo 3

Redes Neuronales Artificiales

Este capítulo comienza con una breve descripción histórica del concepto y técnica de las redes neuronales artificiales. Luego se describe la estructura del elemento básico de procesamiento, la forma de operar de algunos modelos clásicos y finalmente los algoritmos de ejecución.

3.1 Introducción

Las redes neuronales artificiales RNAs son sistemas dinámicos de aprendizaje conformados por un conjunto de técnicas matemáticas para modelar las conexiones entre una colección de datos. La era moderna, empieza con el trabajo pionero de McCulloch y Pitts en 1943. En su clásico artículo McCulloch [43] describe el primer modelo abstracto de una neurona en términos neurocientíficos y matemáticos. Ambos autores demostraron que mediante redes basadas en este tipo de modelo de neurona se podría realizar cualquier función lógica. En términos generales, se puede demostrar que un nodo de tipo umbral puede implementar funciones separables linealmente [Lippmann 87].

El primer modelo de operación neuronal, fue mejorado en sus aspectos biológicos por Hebb [49], quien propuso un mecanismo de aprendizaje para la neurona biológica. Hebb postuló que cuando un axón presináptico causa la activación de cierta neurona postsináptica, la eficiencia de las sinapsis que las relaciona se refuerza. Su postulado de gran influencia contribuyó al desarrollo de modelos computacionales de aprendizaje y sistemas adaptativos.

Rosenblatt [57], introdujo el Perceptrón como herramienta computacional. La estructura del Perceptrón se fundamentó en las primeras etapas de procesamiento de los sistemas sensoriales de los animales, en los que la información atraviesa sucesivas capas de neuronas, realizando un procesamiento progresivo de mayor nivel. Este modelo fue la base de la arquitectura de las redes neuronales artificiales y se usa en la actualidad como reconocedor de patrones. El perceptrón tiene la facultad de generalizar, es decir, después de aprender una serie de patrones puede reconocer otros patrones similares que no se le han presentado anteriormente. Sin embargo, carece de capacidad para resolver el problema de la función OR-exclusiva y es inhábil para catalogar clases no separables linealmente.

Una variante del perceptrón es la ADALINE (adaptive linear element), desarrollada por Widrow [60], en este modelo Widrow introduce el algoritmo LMS (least mean-square). A diferencia del nodo de asociador lineal, la ADALINE incorpora un parámetro adicional denominado sesgo que le proporciona un grado de libertad adicional permitiéndole realizar además de las combinaciones lineales de las entradas (rotaciones y dilataciones) de un asociador lineal, transformaciones traslacionales.

A finales de los años 60 surgieron numerosas críticas que frenaron el crecimiento de las redes neuronales artificiales. El trabajo más influyente fue el análisis matemático detallado del perceptrón que presentó Minsky [69], en el que puso claramente de manifiesto las limitaciones de éste y considero que la extensión a perceptrones multinivel era completamente estéril. No obstante, Rumelhart [86a] demostró que algunos problemas imposibles para los perceptrones simples, pueden ser resueltos por redes de arquitectura multinivel con funciones de activación no lineal, utilizando procesos sencillos de entrenamiento.

Posteriormente, Hopfield [82] aplicó la idea de una función de energía con modelos de vidrio espín de la mecánica estadística, para formular una nueva ruta en la neurocomputación por medio de redes recurrentes con conexiones sinápticas simétricas. Esta analogía abrió el camino para que un diluvio de teorías físicas fueran incorporadas en la modelación neuronal. De este modo, Hopfield transformó el campo de las redes neuronales. Particularmente las redes neuronales con alimentación hacia atrás, ha ganado un gran espacio en el curso del tiempo.

Otro importante desarrollo fue la investigación publicada por Kohonen [82], denominada mapas autoorganizados de Kohonen. Este modelo trata de reproducir la evidencia de que la información en la corteza cerebral aparece con frecuencia organizada en forma espacial. El objetivo de Kohonen era demostrar que un estímulo externo por sí solo, suponiendo una estructura propia y una descripción funcional del comportamiento de la red, era suficiente para forzar la formación de mapas. El modelo tiene dos variantes, denominadas LVQ (learning vector quantization) y SOFM (self-organizing feature map). Ambos se basan en el principio de formación de mapas topológicos para establecer características comunes entre la información de entrada a la red, aunque difieren en las dimensiones, siendo de una sola dimensión en el caso de LVQ y bidimensional, e incluso tridimensional en la red SOFM. Los Mapas autoorganizados de propiedades requiere de una red de una capa en la que se establece un orden de vecindad entre las neuronas, a través de un retículo (uni, bi, multi) dimensional de neuronas que calculan un discriminante, un mecanismo determina la neurona con mayor valor de salida; se activa la neurona seleccionada y sus vecinas; el proceso de aprendizaje hace que se refuerce los discriminantes de las neuronas activas con respecto a las señales de entrada.

En 1984, Kirkpatrick [84], describe un nuevo procedimiento llamado Enfriamiento o Templado Simulado (simulated annealing), un procedimiento también arraigado en la mecánica estadística. Partiendo de un estado inicial de alta temperatura y gran desorden interno (entropía), simula un enfriamiento progresivo que ayuda a encontrar un estado estable en el mínimo global de la función de energía. La idea del templado simulado fue usada posteriormente por Hinton [84], en el desarrollo de la denominada máquina de Boltzmann, la cual fue la primera realización exitosa de una red neuronal multinivel, indicando que la especulación de Minsky [69] fue incorrectamente fundamentada.

También durante los años de 1980, Barto [84], amplió el algoritmo de aprendizaje reforzado desarrollando el modelo asociativo con recompensa y penalización, que se aplica en redes con conexión hacia delante de dos capas cuyas neuronas de salida presentan una función de activación estocástica. En este caso existe una única señal de error, que representa un índice global del rendimiento de la red. Penalizando errores vs. descubriendo similitudes, el aprendizaje basado en la similitud encuentra que elementos del conjunto de entrenamiento están codificados de manera insuficiente por la red para adicionarlo y modificar el conocimiento almacenado para adaptarlo a las características del ambiente. Esto hace que la información acumulada previamente por la red no sea desechada a medida que la red interactúa con el ambiente, realizando así un aprendizaje estable. El proceso de búsqueda regula de forma interna la modificación la topología y de los pesos de la red.

Rumelhart [86b], reportó el desarrollo del algoritmo de retropropagación de errores (backpropagation), que consistió en la solución al problema de entrenamiento de los nodos dispuestos en las capas ocultas de redes multicapas.

Grossberg [87] establece un nuevo principio de auto-organización conocido como la ART (adaptive resonance theory). La teoría de resonancia adaptativa fue propuesta como una teoría de la cognición humana, bajo esta teoría se han formulado modelos de aprendizaje destinados a la generación no supervisada de categorías y al reconocimiento de patrones, se basa en la idea de hacer resonar la información de entrada con los prototipos de las categorías que reconoce la red. Si la red entra en resonancia con alguno de ellos, hace la consideración que pertenece a dicha categoría y realiza la adaptación del prototipo. Cuando no resuena con ningún prototipo, la red se encarga de crear una nueva categoría con el dato de entrada como prototipo de la misma.

Broomhead [88] describe un procedimiento para el diseño de una red con capa de alimentación hacia delante usando funciones de base radial lo cual provee una alternativa más al perceptrón multicapa.

La aparición de ordenadores rápidos en los que pueden simularse redes complejas de gran tamaño y el descubrimiento de potentes algoritmos de aprendizaje, han posibilitado el progreso vertiginoso de las redes neuronales y sus numerosas aplicaciones en diversos campos del conocimiento. Esta diversidad de aplicaciones, ha afianzado el grado de madurez de los sistemas neuronales artificiales, dejando atrás la fase experimental y objeto de moda e instituyéndose en una técnica más a añadir a la amplia panoplia de métodos existentes.

En la actualidad con el campo más asentado han surgido cuantiosos algoritmos de aprendizaje más eficientes, pero basados en el hecho de que los modelos nacientes tienen a construirse sobre los existentes, lo que ha representado una contribución de importancia notable en numerosos campos del conocimiento, que también ha sido extendido en diversas aplicaciones a los dominios de la ingeniería sísmica [Dowla 90, Emami 96, Essenreite 96, Romo 00, Lin 00, Paolucci 00, Romo 01, Hurtado 01a, Hurtado 01b, Sarmiento 01].

3.2 Estructura de un Sistema Neuronal Artificial

La idea que emerge en los sistemas neuronales artificiales ANS, para abordar la clase de problema que el cerebro resuelve con eficiencia, está concebida en construir sistemas que reproduzcan en parte la estructura de las redes neuronales biológicas. Significa, imitar la estructura hardware del sistema nervioso y construir sistemas de procesamiento de información paralelos, distribuidos y adaptativos que presenten un cierto comportamiento inteligente, con el fin de alcanzar una funcionalidad similar.

Las neuronas, las cuales son los elementos básicos de un sistema biológico se agrupan en conjuntos compuestos por millones de unidades organizadas en capas, conformando un sistema con funcionalidad propia. En la implementación de ANS la emulación se lleva a cabo con una estructura similar donde el elemento primario de partida es por tanto la neurona artificial dispuesta igualmente en capas. Como tal, varias capas constituyen una red neuronal, las que junto con las interfaces de entrada y salida más los módulos convencionales adicionales necesarios, conforman el sistema global del proceso como se ilustra en figura 3.1.

De acuerdo con Rumelhart [86a], cinco elementos componen un sistema neuronal o conexionista: un conjunto de procesadores primarios o neuronas artificiales, un patrón de conectividad o arquitectura, una dinámica de activaciones, una regla de aprendizaje y el entorno donde opera.

3.2.1 Modelo de Neurona Artificial

En primer lugar, se describe la estructura de una neurona artificial genérica y luego la versión simplificada, que posee una estructura próxima a la neurona McCulloch-Pitts clásica, de amplio uso en los modelos orientados a aplicaciones prácticas.

3.2.1.1 Modelo general de neurona artificial

Así, en el marco establecido por Rumelhart [86b], se denomina procesador elemental a un dispositivo simple de cálculo que a partir de un vector de entrada de procedencia externa o de otra neurona, proporciona una única respuesta, como de ilustra en la figura 3.2.

Con base en la figura 3.2, los elementos que constituyen la neurona de etiqueta i son los siguientes

- Conjunto de entradas, $x_j(t)$.
- Peso sináptico de la neurona i , w_{ij} que representan la intensidad de interacción entre cada neurona presináptica j y la neurona postsináptica i .
- Regla de propagación $\sigma(w_{ij}, x_j(t))$, que provee el valor del potencial postsináptico $h_i(t) = \sigma(w_{ij}, x_j(t))$ de la neurona i en función de sus pesos y entradas.
- Función de activación $f_i(a_i(t-1), h_i(t))$ de la neurona i que suministra el estado de activación actual $a_i(t) = f_i(a_i(t-1), h_i(t))$ de la neurona i en función de su estado anterior $a_i(t-1)$ y de su potencial postsináptico actual.
- Función de salida $F_i(a_i(t))$, que proporciona la salida actual $y_i(t) = F_i(a_i(t))$ de la neurona i en función de sus estados de activación.

De este modo, la operación de la neurona i que integra una serie de entradas y entrega una cierta respuesta, puede expresarse como

$$y_i(t) = F_i(f_i[a_i(t-1), \sigma(w_{ij}, x_j(t))]) \quad (3.1)$$

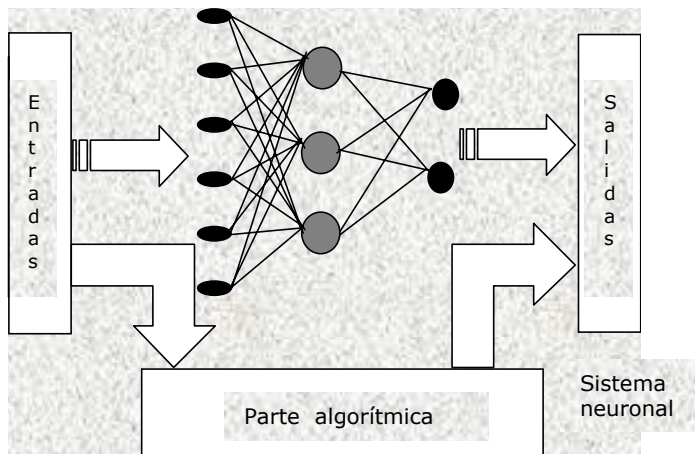
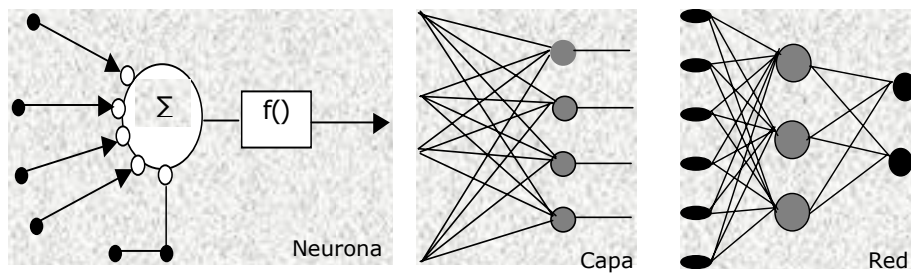


Figura 3.1 Estructura jerárquica de un sistema basado en ANS

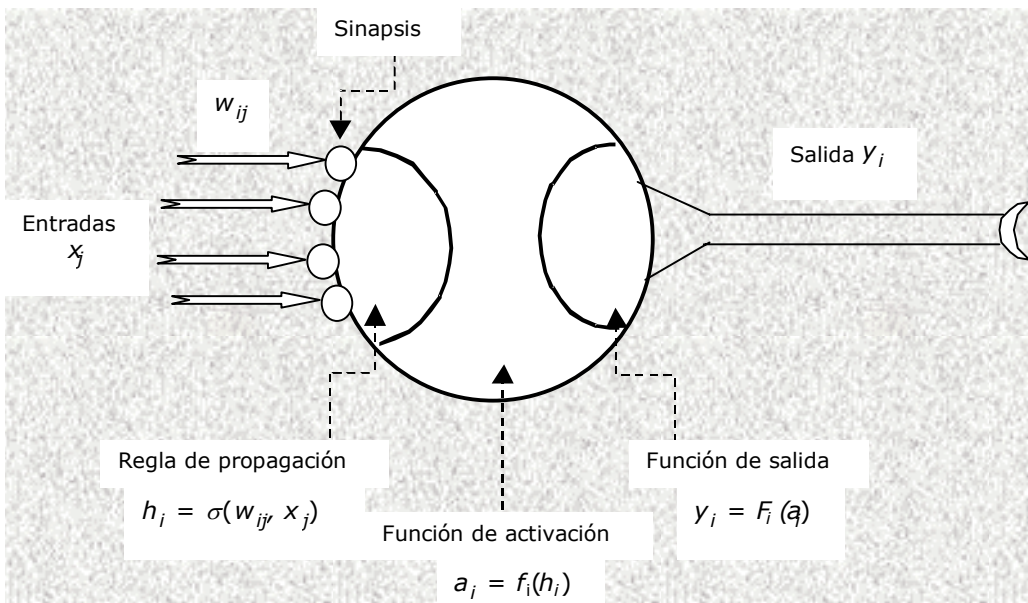


Figura 3.2 Modelo genérico de una neurona artificial

Entradas y salidas

Las variables de entrada y salida pueden ser binarias (digitales) o continuas (analógicas), dependiendo del modelo y aplicación. Por ejemplo, un perceptrón multicapa o MLP admite varios tipos de señales. Para tareas de clasificación opera con salidas digitales $\{0, +1\}$, mientras que para un problema de ajuste funcional de una aplicación multivariable continua, se utilizan salidas continuas pertenecientes a un cierto intervalo.

Dependiendo del tipo de salida, las neuronas suelen recibir nombres específicos. Las neuronas estándar cuya salida sólo puede tomar valores 0 ó 1 se denominan genéricamente neuronas de tipo McCulloch-Pitts, mientras que aquellas neuronas que únicamente pueden tener por salidas -1 o $+1$ se les designa neuronas tipo Ising. Si pueden adoptar diversos valores discretos en la salida (p.e., $-2, 0, +1, +2$), se dice que se trata de una neurona de tipo Potts [Kosko 92]. En ocasiones, el rango de los valores que una neurona de salida continua puede proporcionar se suele limitar a un intervalo definido.

Regla de propagación

La regla de propagación permite obtener, a partir de las entradas y los pesos, el valor del potencial postsináptico h_i de la neurona

$$h_i(t) = \sigma(w_{ij}, x_j(t)) \quad (3.2)$$

La función más habitual es la de tipo lineal, y se basa en la suma ponderada de las entradas con los pesos sinápticos

$$h_i(t) = \sum_j w_{ij} x_j \quad (3.3)$$

que formalmente también puede ser interpretada como el producto escalar de los vectores de entrada y pesos

$$h_i(t) = \sum_j w_{ij} x_j = \mathbf{w}_i^T \cdot \mathbf{x} \quad (3.4)$$

El peso sináptico w_{ij} define en este caso la intensidad de interacción entre la neurona presináptica j y la postsináptica i . Dada una entrada positiva procedente de un sensor o simplemente la salida de otra neurona, si el peso es positivo tenderá a excitar a la neurona postsináptica, si es negativo a inhibirla. Se habla entonces, de sinapsis excitadoras e inhibitoras.

Una regla de tipo no lineal, de uso más limitado, es la siguiente

$$h_i(t) = \sum_{j_1 j_2 \dots j_p} w_{ij_1 j_2 \dots j_p} x_{j_1} x_{j_2} \dots x_{j_p} \quad (3.5)$$

que implica una interacción de tipo multiplicativo entre las entradas de la neurona. El uso de esta última regla de propagación determina que una neurona se denomine de orden superior o neurona sigma-pi al emplear sumas y productos [Kosko 92], e implica una mayor complejidad en el estudio de la dinámica de la red neuronal

La distancia euclidiana, es otra regla de propagación de amplio uso, especialmente en los modelos basados en el cálculo de distancias entre vectores (e. p., los mapas autoorganizados de Kohonen). Representa la distancia al cuadrado existente entre el vector de entradas y el vector de pesos. Cuando ambos vectores son muy similares, la distancia es muy pequeña; y si son muy diferentes, la distancia crece. Por lo tanto opera de manera diferente a las ya citadas reglas

$$h_i^2(t) = \sum_j (x_j - w_{ij})^2 \quad (3.6)$$

Función de transferencia o de activación

La función de transferencia proporciona el estado de activación actual $a_i(t)$ a partir del potencial postsináptico $h_i(t)$ y del propio estado de activación anterior $a_i(t-1)$

$$a_i(t) = f_i(a_i(t-1), h_i(t)) \quad (3.7)$$

Sin embargo, en varias RNAs se considera que el estado actual de la neurona no depende de su estado anterior, sino únicamente del actual

$$a_i(t) = f_i(h_i(t)) \quad (3.8)$$

En general, en las aplicaciones prácticas, la función de activación $f(.)$ se considera determinista, monótona creciente y continua. La forma $y = f(x)$ de las funciones de activación más empleadas en los ANS se enuncian en la tabla 3.1 en la que se designa con x al potencial postsináptico, y y el estado de activación. La más elemental de todas las funciones, es la de identidad, empleada por ejemplo en la ADALINE. Otro caso también sencillo es la función escalón, utilizada en la neurona clásica de McCulloch-Pitts, en el Perceptrón Simple, en la red de Hopfield discreta, y otras.

Normalmente los algoritmos de aprendizaje requieren de una función de activación que cumpla la condición de ser derivable. En este sentido, la más plausible es la función de tipo sigmoide. Otra función clásica es la gaussiana, que se utiliza junto con reglas de propagación que involucran el cálculo de cuadrados de distancias entre los vectores de entrada y pesos (p.e., euclidiana). En ocasiones se emplean funciones senoidales, como en el caso en que se requiere expresar explícitamente una periodicidad temporal. Cualquier función definida en un intervalo de posibles valores de entrada, con un incremento monótonico y que tenga ambos límites superior e inferior, podrá realizar la función de activación o de transferencia de forma satisfactoria.

Tabla 3.1 Algunas funciones de transferencia de uso frecuente

Denominación	Función	Rango
Identidad	$y = x$	$[-\infty, +\infty]$
Escalón	$y = \text{signo}(x)$ $y = H(x)$	$\{-1, +1\}$ $\{0, +1\}$
Lineal a tramos	$y = -1$, si $x < -1$ $y = x$, si $-1 \leq x \leq +1$ $y = +1$, si $x > +1$	$[-1, +1]$
Sigmoide	$y = 1/(1+e^{-x})$ $y = \text{tgh}(x)$	$[0, +1]$ $[-1, +1]$
Gaussiana	$y = A \cdot e^{-Bx^2}$	$[0, +1]$
Sinusoidal	$y = A \cdot \text{sen}(\omega x + \varphi)$	$[-1, +1]$

Función de Salida

Esta función proporciona la salida global de la neurona $y_i(t)$ en términos de su estado de activación actual $a_i(t)$. Con frecuencia la función de salida es simplemente la de identidad $F(x) = x$. De modo que el estado de activación de la neurona se considera como la propia salida

$$y_i(t) = F_i(a_i(t)) = a_i(t) \quad (3.9)$$

Esto ocurre en los modelos clásicos, como el MLP o la ADALINA. La función de salida puede ser también de tipo escalón, lo que supone que la neurona no se dispare hasta que la activación supere

un cierto umbral. En otros modelos, como en la máquina de Boltzmann [Hinton 84], la función de activación es estocástica, apropiando a la neurona un comportamiento probabilista.

3.1.1.2 Modelo estándar de neurona artificial

En la práctica se utiliza un modelo de neurona más simple llamado neurona estándar, que es un caso particular del modelo de Rumelhart considerando la regla de propagación como la suma ponderada y la función de identidad como la función de salida, su representación gráfica se muestra en la figura 3.3. De esta forma, la neurona estándar consiste en

- Un conjunto de entradas x_j y pesos sinápticos w_{ij} .
- Una regla de propagación $h_i(t) = \sigma(w_{ij}, x_j(t))$; $h_i(t) = \sum w_{ij} x_j$ es la más común.
- Una función de activación $y_i(t) = f_i(h_i(t))$, que representa simultáneamente la salida de la neurona y su estado de activación.

A menudo se agrega al conjunto de pesos de la neurona un parámetro adicional θ_i , llamado umbral, el cual se resta del potencial postsináptico, resultando el argumento de la función de activación como

$$\sum w_{ij} x_j - \theta_i \tag{3.10}$$

Esta adición, implica añadir un grado de libertad adicional a la neurona. En el caso de nodos de respuesta todo-nada, el umbral representa el indicio de disparo de la neurona. Es decir, el nivel mínimo que debe alcanzar el potencial de membrana para que la neurona se dispare o se active. Así, el modelo de neurona estándar se expresa como

$$y_i(t) = f_i \left(\sum_j w_{ij} x_{ij} - \theta_i \right) \tag{3.11}$$

Si en la ecuación anterior se hace que los índices i y j comiencen en 0, se puede definir $w_{i0} \equiv \theta_i$ y $x_{0j} \equiv -1$, con lo que el potencial local se obtiene realizando la suma desde $j=0$

$$y_i(t) = f_i \left(\sum_{j=0}^n w_{ij} x_{ij} \right) \tag{3.12}$$

Precisada de este modo la neurona estándar, basta con establecer la forma de la función de activación para una completa determinación.

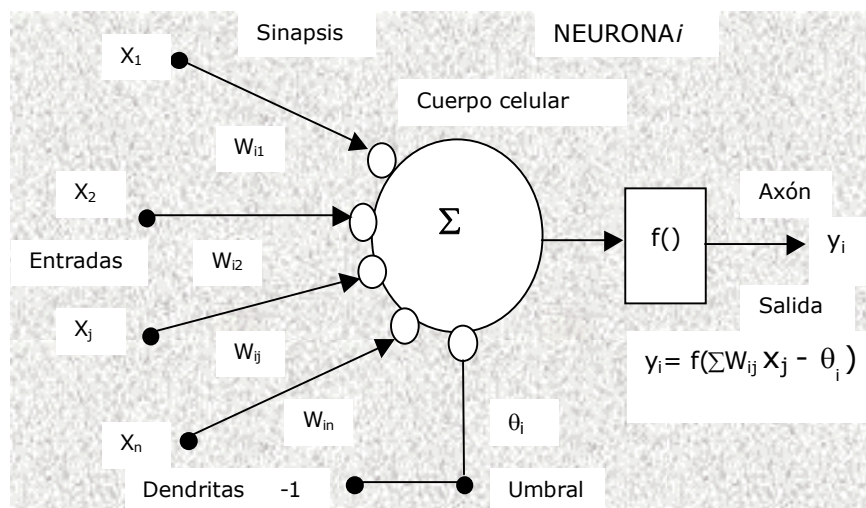


Figura 3.3
Representación de la neurona estándar.

Dispositivo de umbral

Cuando en el modelo de neurona estándar se considera que las entradas son digitales, por ejemplo $x_i = \{0, 1\}$, y la función de activación escalón o Heaviside $H(\cdot)$, definida entre 0 y 1, se obtiene

$$y_i(t) = H\left(\sum_j w_{ij}x_j - \theta_i\right) \quad (3.13)$$

Como $H(x) = 1$ cuando $x \geq 0$, y $H(x) = 0$ cuando $x < 0$, se tiene

$$y_i = \begin{cases} 1, & \text{si } \sum w_{ij} \geq \theta_i \\ 0, & \text{si } \sum w_{ij} < \theta_i \end{cases} \quad (3.14)$$

Es decir, si el potencial de membrana supera el valor umbral de disparo θ_i , la neurona se activa, si no, sencillamente la neurona no se activa. Este es el modelo de neurona del perceptrón original denominado en algunos casos dispositivo de tipo umbral.

Ahora, si en el dispositivo de tipo umbral, se considera que la acción de las entradas inhibitoras es absoluta (ante la presencia de una sola señal inhibitora, las neuronas ya no se disparan) y se introduce retardos en la propagación de las señales, se obtiene el modelo original de neurona de McCulloch-Pitts.

Neurona continua tipo sigmoide

Se implementa como función de activación una sigmoide cuando en el esquema de neurona estándar se considera que las entradas pueden ser tanto digitales como continuas y las salidas exclusivamente continuas. Un sigmoide es una función continua y diferenciable en cierto intervalo, por ejemplo $[-1, +1]$ o en el $[-0, +1]$, dependiendo de la función concreta que se elija. Las dos funciones más usuales de este tipo son las siguientes

$$y = f(x) = \frac{1}{1 + e^{-x}}, \text{ con } y \in [0, 1] \quad (3.15)$$

$$y = f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \text{tgh}(x), \quad \text{con } y \in [-1, 1] \quad (3.16)$$

Este modelo de neurona es el utilizado en el perceptrón multicapa. El requisito de trabajar con funciones diferenciables como ya se mencionó es imponer la regla de aprendizaje.

3.3 Arquitecturas de redes neuronales

La arquitectura o patrón de conexión de una red neuronal proporciona los canales por los que discurre su dinámica y determina el comportamiento de la red a través de su propia conexión sináptica.

Una red neuronal está conformada por tres capas con funciones diferentes. La capa sensorial, compuesta por neuronas que reciben los datos y las señales procedentes del entorno; la capa de salida que proporciona la respuesta de la red neuronal; y una capa oculta que no tiene conexión directa con el dominio externo, no se enlaza directamente a órganos sensores ni a efectores. Esta

capa escondida suministra a la red neuronal grados de libertad mediante los cuales puede encontrar representaciones internas correspondientes a determinados rasgos del entorno.

Las conexiones entre neuronas pueden ser instigadoras cuando el peso sináptico es positivo o inhibitoras si el peso sináptico es negativo. Usualmente, no se suele definir una conexión como de un tipo o de otra, sino que por medio del aprendizaje se obtiene un valor del peso que incluye signo y magnitud.

Por otro lado, se pueden distinguir conexiones intracapa que tienen lugar entre las neuronas pertenecientes a un mismo nivel, denominadas enlaces laterales; y conexiones intercapa que se producen entre las neuronas de diferentes estratos. Existen además conexiones realimentadas con un sentido contrario al de entrada-salida, en algunos casos puede existir realimentación incluso de una neurona con signo misma.

Con base en los distintos conceptos, se establecen diferentes tipos de arquitecturas neuronales. De acuerdo con su organización en capas, se habla de redes monocapa y de redes multicapa. Asimismo, considerando el flujo de datos en la red, éstas se clasifican en unidireccionales (*feedforward*) y recurrentes (*feedback*). Una red es unidireccional cuando no presenta bucles cerrados de conexiones, la información se propaga en un sentido desde la neurona de entrada hasta la neurona de salida. En las redes recurrentes la información puede circular entre las capas en cualquier sentido, incluida la de salida-entrada, el flujo de información puede encontrar un bucle realimentador de atrás hacia delante.

También se habla de redes autoasociativas y heteroasociativas. Con frecuencia se interpreta la operación de una red neuronal como la de una memoria asociativa, que ante un determinado patrón de entradas responde con un cierto patrón de salida. Si una red se entrena para que con la presentación de un patrón **A** responda con otro diferente **B**, se dice que la red es heteroasociativa; y cuando una red se educa para que asocie un patrón **A** consigo misma, se clasifica como autoasociativa. El interés de este tipo de redes, como en el caso de la de Hopfield, reside en que ante la presentación del patrón $\mathbf{A}' = \mathbf{A} + \text{ruido}$, de como respuesta el patrón original **A**, eliminando el ruido presente en la señal de entrada.

3.4 Operación

Se distinguen dos modos de operación en los sistemas neuronales artificiales, el modo de convergencia o fase aprendizaje y el modo de estabilidad o fase de ejecución.

3.4.1 Fase de aprendizaje

En el contexto de las redes neuronales puede definirse el aprendizaje como la fase en la que se produce el ajuste de los parámetros libres de la red con base en un proceso de estimulación producido por el entorno que la rodea. El tipo de aprendizaje está determinado por la forma en que dichos parámetros son adaptados. En la mayoría de los casos el aprendizaje reside sin más, en determinar un conjunto de pesos sinápticos que permita a la red realizar correctamente el tipo de procesamiento necesario.

Cuando se construye un sistema neuronal, se parte de un cierto modelo de neurona y de una determinada arquitectura, estableciéndose los pesos sinápticos iniciales como nulos o aleatorios. Para que la red resulte operativa es necesario entrenarla, lo que constituye el modo aprendizaje. El entrenamiento se puede llevar a cabo en dos niveles. El convencional es el modelado de las sinapsis, que consiste en modificar los pesos sinápticos siguiendo una cierta regla de aprendizaje, construida normalmente a partir de la optimización de una función de error, que mide la eficacia actual de la operación de la red. Si se denomina $w_{ij}(t)$ al peso que conecta la neurona presináptica j con la postsináptica i en la iteración t , el algoritmo de aprendizaje en función de las señales que en el instante t llegan procedentes del entorno, proporcionará el valor $\Delta w_{ij}(t)$ que corresponde a la modificación que debe incorporarse en dicho peso y el cual queda actualizado de la forma

$$\Delta w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t).$$

El proceso de aprendizaje es iterativo, renovándose de la manera anterior una y otra vez, hasta que la red alcanza el rendimiento deseado.

Algunos modelos neuronales incluyen otro nivel en el aprendizaje como la creación o destrucción de neuronas, en el que se modifica la propia arquitectura de la red. En cualquier caso, en el aprendizaje la información contenida en los datos de entrada queda incorporada en la propia estructura de la red neuronal, la cual almacena la representación de una cierta imagen de su entorno.

Los tipos básicos de aprendizaje son el supervisado y el no supervisado, cuya distinción tiene su origen en el campo del reconocimiento de patrones. Ambas modalidades pretenden estimar funciones entrada/salida multivariable, pero mientras que en el aprendizaje supervisado se proporciona una cierta información sobre estas funciones (como la distribución de las clases, etiquetas de los patrones de entrada o salidas asociadas a cada patrón, etc.), en el autoorganizado no se suministra información alguna. Las reglas de aprendizaje supervisados suelen ser computacionalmente más complejas, pero proporcionan mejores resultados. Además de las formas básicas de aprendizaje pueden distinguirse otros tipos como el híbrido y el reforzado.

En el aprendizaje supervisado se presenta a la red una colección de patrones, junto con la salida deseada, e iterativamente la red ajusta sus pesos hasta que su salida tiende a ser la deseada, utilizando para ello información detallada del error en cada paso. De este modo, la red es capaz de estimar relaciones entrada/salida sin necesidad de una función prescrita.

Sea $E[W]$ un funcional que representa el error esperado de la operación de la red, expresado en términos de sus pesos sinápticos W . En el aprendizaje supervisado se pretende estimar una cierta función multivariable desconocida $f: \mathcal{X}^n \rightarrow \mathcal{Y}^m$ (la que representa la red neuronal) a partir de las muestras (\mathbf{x}, \mathbf{y}) ($\mathbf{x} \in \mathcal{X}^n$, $\mathbf{y} \in \mathcal{Y}^m$) tomadas aleatoriamente, por medio de la minimización iterativa de $E[W]$ mediante aproximación estocástica.

Las técnicas de aproximación estocástica estiman valores esperados a partir de cantidades aleatorias observadas. Usualmente se implementan en forma de algoritmo discreto del tipo de descenso por el gradiente.

Varios algoritmos de aprendizaje, se basan en métodos numéricos iterativos que tratan de minimizar una función costo. Lo que puede dar lugar en ocasiones a problemas en la convergencia del algoritmo. Este aspecto no puede abordarse de un modo general, debe ser estudiado para cada algoritmo. En un sentido riguroso, la convergencia es una manera de comprobar si una determinada arquitectura, junto con su regla de aprendizaje, es capaz de resolver un problema, pues el grado de error que se mide durante el proceso de aprendizaje describe la precisión del ajuste.

En el proceso de entrenamiento es importante distinguir entre el nivel de error alcanzado al final de la fase de aprendizaje para el conjunto de datos de entrenamiento y el error que la red ya entrenada incurre ante patrones no utilizados en el aprendizaje, lo cual mide la capacidad de generalización de la red. Interesa más una buena generalización que un error muy pequeño en el entrenamiento, pues ello indica que la red ha capturado correctamente el mapeo subyacente en los datos.

3.4.2 Fase de ejecución

Una vez que el sistema ha sido entrenado, el aprendizaje se desconecta en la mayoría de los casos no en todos los modelos, por lo que los pesos y la estructura quedan fijos y la red dispuesta para procesar datos.

En las redes unidireccionales, ante un patrón de entrada, las neuronas responden proporcionando directamente la salida del sistema. Al no existir bucles de realimentación no existe ningún problema en relación con su estabilidad. Por el contrario, las redes con realimentación son sistemas dinámicos no lineales, que requieren ciertas condiciones para que su respuesta converja a un estado estable o punto fijo. Una serie de teoremas generales indican las condiciones que aseguran la estabilidad de la respuesta en una amplia gama de redes neuronales, bajo determinadas condiciones [Kosko, 1992].

Los teoremas que permiten demostrar la estabilidad del sistema (p.e., teorema de Cohen-Grossberg, teorema de Cohen-Grossberg-Kosko [Kosko 92]) se basan en el método de Lyapunov, como alternativa al arduo método directo.

Básicamente, el método de Lyapunov establece que si en un sistema dinámico -como es una red neuronal- de varias entradas (x_1, x_2, \dots, x_n) y descrito por el siguiente sistema de ecuaciones diferenciales

$$\dot{x}_i \equiv \frac{dx_i}{dt} = F(t, x_1, x_2, \dots, x_n) \quad (3.17)$$

se cumplen las condiciones

- el sistema está en reposo solamente en el origen
- existen las derivadas de las ecuaciones que lo describen en todo el dominio
- las variables están acotadas

y se puede encontrar una función de Lyapunov V de las variables x_i , $V: \mathfrak{R}^n \rightarrow \mathfrak{R}$, tal que

$$\dot{V} = \sum_{i=1}^n \frac{\partial V}{\partial x_i} \leq 0, \forall x_i \quad (3.18)$$

entonces el sistema converge para todas las posibles entradas (x_1, x_2, \dots, x_n), y es globalmente estable.

La función de Lyapunov denominada frecuentemente función de energía de Lyapunov constituye una generalización del concepto físico de energía. Es un instrumento matemático que permite probar la estabilidad global de un sistema dinámico, erige una manera asequible de estudiar la consistencia de sistemas dinámicos. Con esta formulación matemática simplemente se está expresando que si se encuentra una cierta función energía del sistema, que disminuya siempre en su operación, entonces el sistema es estable. Una técnica similar empleó Hopfiel para demostrar que su modelo de red completamente interconectado era estable en el caso de que la matriz de pesos sináptica fuese simétrica y de diagonal nula.

La técnica de Lyapunov es también la que Cohen, Grossberg y Kosko han aplicado en sus teoremas para demostrar la estabilidad de una amplia clase de redes neuronales realimentadas, autoasociativas y heteroasociativas. Así, el teorema de Cohen-Grossberg, determina las condiciones de estabilidad para redes autoasociativas no adaptativas y el de Cohen- Grossberg – Kosko, establece los requisitos de solidez para redes autoasociativas adaptativas. Como ilustración se enuncia el teorema de Cohen-Grossberg [Haykin, 99].

Teorema de Cohen-Grossberg

Para cualquier sistema dinámico no lineal que pueda describirse de la forma siguiente

$$\dot{x}_i = \alpha_i(x_i) \left[\beta_i(x_i) - \sum_{j=1}^n m_{ji} S_j(x_j) \right] \quad (3.19)$$

tal que

- la matriz $\|m_{ij}\|$ es simétrica y $m_{ij} \geq 0, \forall i, j$
- la función $\alpha_i(\xi)$ es continua $\forall \xi \geq 0$
- $\alpha_i(\xi) \geq 0, \forall \xi \geq 0; S_i(\xi) \geq 0, \forall \xi \geq 0$
- la función $S_i(\xi)$ es diferenciable y no decreciente $\forall \xi \geq 0$

la función (3.20), es una función de energía de Lyapunov para el sistema y el sistema es estable.

$$V = (1/2) \sum_{i=1}^n \sum_{j=1}^n m_{ij} S_i(x_i) S_j(x_j) - \sum_{i=1}^n \int_0^{x_i} S_i'(\theta_i) \beta(\theta_i) d\theta \quad (3.20)$$

Si se interpreta la ecuación anterior como la descripción de la activación en el tiempo de una red neuronal autoasociativa de una o más capas, no adaptativa, con matriz de peso simétrica y que cumple las condiciones enunciadas, entonces dicha red es estable.

3.5 Clasificación de los modelos neuronales

Existen diversos tipos de RNAs en las que cada una tiene sus propias características. Ellas difieren en su arquitectura (número de capas, número de neuronas por capa, grado de conectividad y tipo de conexión entre neuronas); mecanismo de aprendizaje (supervisado por corrección de error, supervisado reforzado, supervisado estocástico, no supervisado e híbrido); tipos de asociación entre la información de entrada y salida (heteroasociativa o autoasociativa) y en la forma como se representa la información (analógica o binaria).

Una clasificación que procede considerando el tipo de aprendizaje y la arquitectura se encuentra bien documentada en [Haykin 99]. Sobresalen los mapas autoorganizados de Kohonen como modelo neuronal no supervisado; el modelo de Hopfield empleado como memoria asociativa; la máquina de Boltzmann que resulta de emplear el modelo de neurona de Hopfield desde un punto probabilista; las funciones de base radial de tipo híbrido al incorporar aprendizaje supervisado y no supervisado; y la red neuronal de aprendizaje supervisado y arquitectura recurrente denominada cascada de correlación. En este caso, dado el interés de la aplicación de las RNAs como estimadores de funciones, se describen las redes neuronales más representativas, de aprendizaje supervisado y arquitectura no realimentada.

3.5.1 El asociador lineal

El asociador lineal es un sencillo ejemplo de red unidireccional. Este modelo, mediante una transformación lineal, asocia un conjunto de patrones de entrada a otros de salida. Consta de una capa de neuronas lineales, denotadas por \mathbf{x} las entradas, por \mathbf{y} las salidas que a su vez constituye el vector de respuesta de la red neuronal y por $\mathbf{W} = \{w_{ij}\}$ la matriz de pesos sinápticos. Cada fila de \mathbf{W} contiene los pesos de una neurona \mathbf{w}_i

$$\mathbf{W} = (\mathbf{w}_1 \ \mathbf{w}_2 \ \dots \ \mathbf{w}_m)^T \quad (3.21)$$

La operación del asociador lineal es simplemente

$$\mathbf{y} = \mathbf{W}\mathbf{x} \quad (3.22)$$

Por lo tanto, cada neurona i del asociador lineal lleva a cabo la suma ponderada de las entradas con sus propios pesos sinápticos.

El asociador lineal debe aprender a asociar p pares entradas-salidas, $\{(\mathbf{x}^\mu, \mathbf{t}^\mu) / 1 \leq \mu \leq p\}$, ajustando sus pesos \mathbf{W} de modo que ante un cierto patrón de entrada \mathbf{x}^μ responda con \mathbf{t}^μ , y que ante entradas similares, $(\mathbf{x}^\mu + \varepsilon)$, responda con salidas también próximas $(\mathbf{t}^\mu + \delta)$, con ε y δ cantidades pequeñas. El problema se centra en encontrar la matriz de pesos \mathbf{W} haciendo uso de una regla de aprendizaje que proporcione el conjunto óptimo de pesos.

3.5.2 Regla de aprendizaje de Hebb

Se denomina aprendizaje hebbiano a la forma de aprendizaje que involucra una modificación de los pesos Δw_{ij} proporcional al producto de una entrada j por la salida i de la neurona. Así, el aprendizaje descrito por Hebb puede calcularse como

$$\Delta w_{ij} = \varepsilon y_i x_j \quad (3.23)$$

donde ε es el ritmo de aprendizaje con valores entre 0 y 1.

Para el asociador lineal, la regla de Hebb se formula mediante la expresión

$$\Delta W_{ij}^{\mu} = t_i^{\mu} x_j^{\mu} \quad (3.24)$$

por lo tanto

$$W_{ij}^{new} = W_{ij}^{old} + \Delta W_{ij}^{\mu} \quad (3.25)$$

Si los pesos de partida son nulos, el valor final de \mathbf{W} para las p asociaciones será

$$\mathbf{W} = \mathbf{t}^1 \mathbf{x}^{1T} + \mathbf{t}^2 \mathbf{x}^{2T} + \dots + \mathbf{t}^p \mathbf{x}^{pT} \quad (3.26)$$

Empleando la regla de Hebb para el entrenamiento del asociador lineal, si los vectores de entrada $\{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^p\}$ son ortonormales se cumple

$$\mathbf{W} \mathbf{x}^{\mu} = (\mathbf{t}^1 \mathbf{x}^{1T} + \mathbf{t}^2 \mathbf{x}^{2T} + \dots + \mathbf{t}^p \mathbf{x}^{pT}) \cdot \mathbf{x}^{\mu} = \mathbf{t}^1 (\mathbf{x}^{1T} \cdot \mathbf{x}^{\mu}) + \dots + \mathbf{t}^p (\mathbf{x}^{pT} \cdot \mathbf{x}^{\mu}) = \mathbf{t}^{\mu} \quad (3.27)$$

Así, ante la entrada \mathbf{x}^{μ} se reproduce la respuesta aprendida \mathbf{t}^{μ} . El problema reside en que las condiciones son restrictivas. Para que las asociaciones sean correctas los patrones de entrada deben ser ortogonales y de longitud unitaria, por lo que si la dimensión del espacio de entradas es n , solamente podrá aprender hasta n asociaciones. Para almacenar más pares entrada-salida se requiere utilizar otras estrategias.

Manteniendo el requisito de vector de longitud unitaria y eliminando la condición de ortogonalidad, se tiene

$$\mathbf{W} \mathbf{x}^{\mu} = \mathbf{t}^1 (\mathbf{x}^{1T} \mathbf{x}^{\mu}) + \mathbf{t}^2 (\mathbf{x}^{2T} \mathbf{x}^{\mu}) + \dots + \mathbf{t}^p (\mathbf{x}^{pT} \mathbf{x}^{\mu}) = \mathbf{t}^{\mu} + \sum_{v \neq \mu} \mathbf{t}^v (\mathbf{x}^{vT} \mathbf{x}^{\mu}) = \mathbf{t}^{\mu} + \delta \quad (3.28)$$

Expresión denominada expansión señal-ruido, porque proporciona la salida deseada más un término adicional, que se interpreta como el ruido superpuesto en la señal. Haciendo uso de reglas como la pseudoinversa o la de Widrow-Hoff, se obtiene una matriz de pesos que logra además que el ruido δ sea pequeño comparado con la señal.

3.5.3 Regla de la pseudoinversa

En general lo que se pretende en las redes neuronales es conseguir un algoritmo de aprendizaje a partir de un cierto criterio a optimizar. Por lo que se hace necesario definir el significado de óptimo en cada caso. Es decir, hay que establecer un criterio que mida el rendimiento de la red para encontrar una regla de actualización de pesos que lo optimice. Un procedimiento comúnmente utilizado es el error cuadrático medio de las salidas actuales de la red respecto a las medidas. Para el asociador lineal se tiene

$$E\{W_{ij}\} = (1/p) \sum_{\mu=1}^p |\mathbf{t}^{\mu} - \mathbf{W} \mathbf{x}^{\mu}|^2 = (1/p) \sum_{\mu=1}^p \sum_{i=1}^n (t_i^{\mu} - W x_i^{\mu})^2 \quad (3.29)$$

De este modo, un algoritmo de aprendizaje para el asociador lineal deberá obtener un conjunto de pesos que minimicen esta expresión de error. Si los vectores \mathbf{x}^{μ} son ortonormales, el error que proporciona la regla de Hebb (3.24) según (3.29) es cero, lo que indica que esta regla es óptima respecto a la medida del error.

Considerando la norma $\|M\|^2 = \|(m_{ij})\|^2 = \sum_{i=1}^p \sum_{j=1}^q m_{ij}^2$ y denominando a una matriz $n \times p$ como \mathbf{X} que tiene por columnas los vectores de entrada \mathbf{x}^μ , $\mathbf{X} = (\mathbf{x}^1 \ \mathbf{x}^2 \ \dots \ \mathbf{x}^p)$, y una matriz $m \times p$ como \mathbf{Y} cuyas columnas son los vectores de salida \mathbf{y}^μ , $\mathbf{Y} = (\mathbf{y}^1 \ \mathbf{y}^2 \ \dots \ \mathbf{y}^p)$, la ecuación (3.29) se transforma en

$$E\{\mathbf{w}_{ij}\} = (1/p)\|\mathbf{Y} - \mathbf{W}\mathbf{X}\|^2 \quad (3.30)$$

La regla de Hebb con esta nomenclatura, se expresa de la forma

$$\mathbf{W} = \mathbf{Y}\mathbf{X}^T \quad (3.31)$$

De esta forma, una regla de aprendizaje que se sirva de la matriz pseudoinversa puede escribirse como

$$\mathbf{W} = \mathbf{Y}\mathbf{X}^+ \quad (3.32)$$

Donde \mathbf{X}^+ denota la pseudoinversa de \mathbf{X} [Hetch-Nielsen 90]. Se puede demostrar que esta elección para \mathbf{W} minimiza el error cuadrático medio, ecuaciones (3.29 y 3.30). Es decir, que es óptima respecto de este error. En [Ritter 91] se deduce la ecuación (3.32) a partir de la minimización algebraica de la ecuación (3.30).

Así como con la regla de Hebb se pueden almacenar hasta n vectores ortonormales, con la pseudoinversa se pueden almacenar hasta n vectores linealmente independientes. Si se pretende almacenar más pares entrada-salida, se presentará sus errores pero el mapeo lineal implementado seguirá siendo óptimo en el sentido del menor error cuadrático medio posible.

Al ser ambas reglas óptimas con el mismo criterio, indica que deben estar relacionadas, lo cual es apreciable al considerar un conjunto de vectores de entrada ortonormales donde la regla de la pseudoinversa se convierte en la Hebb. Si se realiza la expansión en serie de la ecuación (3.32) de la pseudoinversa [Haykin 99], el primer término de la serie es precisamente la ecuación (3.31) de la regla de Hebb. Así, la regla de Hebb representa un caso particular de la pseudoinversa.

Para el cálculo de la pseudoinversa se utiliza el teorema de Greville [Hetch-Nielsen 90, Kosko 92], pese a que presenta el inconveniente de que para aprender un nuevo patrón se debe recalcular toda la matriz de pesos, lo que no encaja dentro de los conceptos de redes neuronales que buscan que los modelos sean locales y operen incrementalmente. Se ha demostrado en la práctica [Kosko 92] que una forma aproximada del teorema de Greville que proporciona resultados correctos es como

$$w_i^{new} = w_i^{old} + \varepsilon \cdot (t_i^\mu - (w_i^{old})^T \mathbf{x}^\mu) \mathbf{x}^\mu \quad (3.33)$$

siendo ε el ritmo de aprendizaje, parámetro que indica la rapidez en la actualización ($0 < \varepsilon < 1$). En este esquema iterativo, los patrones deben ser presentados a la red repetidamente, obteniéndose una aproximación a la matriz pseudoinversa mediante cálculos simples y locales.

Esta expresión coincide con la regla de Widrow-Hoff de la adalina, en [Haykin 99] se relacionan los dos algoritmos indicando que ambas reglas son equivalentes cuando se trata de realizar un aprendizaje estadístico.

3.5.4 Redes neuronales de aprendizaje supervisado y arquitectura no realimentadas

El conjunto de redes neuronales de este tipo son las más numerosas, por lo que se describen los modelos de relevancia histórica y generalidad con el fin de indicar los aspectos que con frecuencia aparecen en el campo de las redes neuronales, como es la estimación funcional, la clasificación de patrones, la memoria asociativa, etc.

3.5.4.1 Perceptrón simple

El perceptrón simple fue introducido por Rosenblatt [57], es un modelo unidireccional, compuesto por dos capas de neuronas, una sensorial y otra de salida. La operación de una red de este tipo, con n neuronas de entrada y m de salida, se puede expresar como

$$y_i(t) = f\left(\sum_{j=1}^n w_{ij}x_j - \theta_i\right), \forall i, 1 \leq i \leq m \quad (3.34)$$

las neuronas de entradas no realizan ningún cómputo, únicamente envían información a las neuronas de salida. La función de activación de las neuronas de la capa de salida es de tipo escalón. Por lo que la operación del perceptrón simple puede escribirse como

$$y_i = H\left(\sum_{j=1}^n w_{ij}x_j - \theta_i\right), \forall i, 1 \leq i \leq m \quad (3.35)$$

El perceptrón es útil como clasificador y para representar funciones booleanas ya que la neurona es de tipo MacCulloch-Pitts y de salida binaria.

Insertar el algoritmo de aprendizaje al perceptrón, consiste en modificar la respuesta que proporciona el perceptrón según la regla de Hebb. Así, para un conjunto de p patrones \mathbf{x}^μ , $\mu=1, \dots, p$, con salidas deseadas t^μ en el intervalo -1 ó 1 , y arquitectura con pesos inicialmente aleatorios, la actualización de los pesos es

$$\Delta w_{ij}^\mu(t) = \begin{cases} 2\varepsilon(t_i^\mu x_j^\mu), & \text{si } y_i^\mu \neq t_i^\mu \\ 0, & \text{si } y_i^\mu = t_i^\mu \end{cases} \quad (3.36)$$

Actualización que también se puede escribir como

$$\Delta w_{ij}^\mu(t) = \varepsilon(t_i^\mu - y_i^\mu)x_j^\mu \quad (3.37)$$

La ecuación (3.37) es la forma usual de expresar la regla del perceptrón. En su utilización práctica, se debe llegar a un compromiso para el valor del ritmo de aprendizaje ε , puesto que un valor pequeño implica un aprendizaje lento y uno excesivamente grande puede conducir a oscilaciones en el entrenamiento e introducir variaciones en los pesos extremadamente amplias. Al ser las entradas y las salidas discretas $\{-1, +1\}$, también lo será la actualización de sus pesos, que únicamente podrán tomar los valores 0 ó $\pm 2\varepsilon$. El ajuste de los pesos en la iteración t debido a todo el conjunto de aprendizaje será

$$w_{ij}(t+1) = w_{ij}(t) + \sum_{\mu=1}^p \Delta w_{ij}^\mu(t) \quad (3.38)$$

3.5.4.2 ADALINE

El modelo Adaptive Linear Element, introducido por Widrow [60], es un modelo clásico que utiliza una neurona similar a la del perceptrón pero de respuesta lineal y cuyas entradas pueden ser continuas. Incorpora un parámetro denominado sesgo que proporciona un grado de libertad adicional y no un umbral de disparo como el del perceptrón. La operación de la ADALINE se expresa como

$$y_i(t) = \sum_{j=1}^n w_{ij}x_j - \theta_i, \forall i, 1 \leq i \leq m \quad (3.39)$$

La diferencia con el asociador lineal y el perceptrón reside en la regla de aprendizaje, que para este caso es la regla de Widrow-Hoff o LMS (Least Mean Squares), también conocida como la regla delta para un caso particular. Esta regla conduce a actualizaciones de tipo continuo proporcionales al error en el que incurre la neurona. La ADALINE se limita a patrones linealmente independientes; sin embargo, ante patrones no separables linealmente los resultados que proporciona son en promedio mejor que los del perceptrón al operar reduciendo el error cuadrático medio al mínimo posible.

La regla LMS conduce a mejores asociaciones cuando son linealmente independientes, proporcionando cuando no lo son una matriz de pesos óptima desde el punto de vista de los mínimos cuadrados. Así la regla delta se considera como una versión iterativa aproximada de la basada en la pseudoinversa para el caso de vectores estocásticos [Haykin 99].

La forma de derivar la regla LMS a partir de la optimización de cierta función energía, es la manera como en la computación neuronal se obtiene el algoritmo de aprendizaje.

El procedimiento que permite derivar de forma sistemática reglas de aprendizaje para arquitecturas concretas, consiste en proponer una función error que mida el rendimiento actual de la red, función que dependerá de los pesos sinápticos. Dada esta función error, se introduce un procedimiento general de optimización que sea capaz de proporcionar una configuración de pesos que corresponda a un mínimo de la función propuesta. El método de optimización aplicado a la función error proporciona una regla de actualización de pesos, que en función de los patrones de aprendizaje modifican iterativamente los pesos hasta alcanzar el punto óptimo de la red neuronal.

Un método de minimización usualmente empleado es el denominado descenso por el gradiente. Para esto, se comienza definiendo una función $E(\cdot)$ que proporcione el error actual E que comete la red neuronal y que será una función del conjunto de pesos W , $E=E(W)$, $E:\mathcal{R}^n \rightarrow \mathcal{R}$. Así se puede construir la representación de esta función, como una hipersuperficie con colinas y valles. El objetivo del aprendizaje es encontrar la configuración de pesos que corresponde al mínimo global de la función error.

Para encontrar la configuración de pesos óptima mediante descenso por el gradiente se opera del siguiente modo. Se parte en $t=0$ de una cierta configuración $W(0)$, y se calcula el sentido de la máxima variación de la función $E(W)$ en $W(0)$, que vendrá dado por su gradiente en $W(0)$. El sentido de la máxima variación apunta hacia una colina del paisaje de la hipersuperficie de $E(\cdot)$. Luego se modifican los parámetros W siguiendo el sentido contrario al indicado por el gradiente de la función de error. De este modo se lleva a cabo un descenso por la hipersuperficie del error, aproximándose en una cierta cantidad al valle. El proceso se itera hasta alcanzar el mínimo y matemáticamente se expresa como

$$W(t+1) = W(t) - \varepsilon \nabla E(W) \quad (3.40)$$

En la ecuación anterior, ε indica el tamaño del paso tomado en cada iteración, que idealmente debe ser infinitesimal y puede ser diferente para cada peso.

Está comprobado matemáticamente [Haykin 99], que una actualización de este tipo conduce a un mínimo de la función de error $E(W)$. Para una matriz de pesos $W = \{w_{ij}\}$, el cálculo de la variación que en $E(W)$ se produce en la iteración t , es

$$\delta(E(w_{ij})) = \sum_{ij} \left(\frac{\partial E(w_{ij})}{\partial w_{ij}} \right) \delta w_{ij} \quad (3.41)$$

la variación de los pesos es $-\varepsilon$ (para el cálculo de δw_{ij} debe ser infinitesimal) multiplicado por el gradiente, por lo tanto

$$\delta(f(w_{ij})) = \sum_{ij} \frac{\partial E(w_{ij})}{\partial w_{ij}} \left(-\varepsilon \frac{\partial E(w_{ij})}{\partial w_{ij}} \right) = -\varepsilon \sum_{ij} \left(\frac{\partial E(w_{ij})}{\partial w_{ij}} \right)^2 \leq 0 \quad (3.42)$$

luego la variación en la función es siempre menor que cero, por lo que siempre disminuye. Mediante este procedimiento se asegura alcanzar un mínimo local de la función, aunque puede no coincidir con el mínimo global.

Para evitar el estancamiento en un mínimo relativo, usualmente se recurre a la regla de Widrow-Hoff, la cual se utiliza en el modelo ADALINE.

Para una muestra finita, se plantea la siguiente función error

$$E[w_{ij}] = \frac{1}{2} \sum_{\mu=1}^p \sum_{i=1}^n (t_i^\mu - y_i^\mu)^2 \quad (3.43)$$

Mediante esta función se obtiene el error cuadrático medio correspondiente a las salidas de la red respecto de los objetivos. El proceso de optimización es el de descenso por el gradiente requiriéndose calcular

$$\frac{\partial E[w_{ij}]}{\partial w_{ij}} = -\sum_{\mu=1}^p (t_i^\mu - y_i^\mu) \frac{dy_i^\mu}{dw_{ij}} = -\sum_{\mu=1}^p (t_i^\mu - y_i^\mu) x_j^\mu \quad (3.44)$$

por tanto, el incremento de pesos

$$\Delta w_{ij} = -\varepsilon \frac{\partial E[w_{ij}]}{\partial w_{ij}} = \varepsilon \sum_{\mu=1}^p (t_i^\mu - y_i^\mu) x_j^\mu \quad (3.45)$$

Esta expresión es la regla LMS. Así, la regla LMS se acerca asintóticamente a la solución, pues el tamaño de los incrementos es cada vez menor.

Debido a la linealidad de la neurona de la ADALINE, la función error $E[w_{ij}]$ es cuadrática en los pesos definiendo una superficie en forma de paraboloides. Un paraboloides, como la parábola en el plano, posee un único mínimo, aunque en ocasiones presenta una forma degenerada con uno o más canales, pero todos de la misma profundidad. En cualquiera de los casos, la función $E[w_{ij}]$ es mínima, y la regla (3.45) lleva directamente al mínimo puesto que siempre desciende por la superficie de error. Por esto, la regla LMS invariablemente alcanza el mínimo global sin importar la configuración de los pesos de partida, asentando uno de los pocos casos en redes neuronales en el que se pueda realizar una afirmación de este tipo.

3.5.4.3 Perceptrón multicapa

Partiendo de un perceptrón monocapa y considerando sus limitaciones computacionales, se llegó a la arquitectura MLP (Multi-Layer Perceptron). Aplicaciones en numerosos problemas han permitido comprobar experimentalmente que con esta topología se pueden representar complejos mapeos y abordar problemas de clasificación de gran envergadura, de una manera eficaz y relativamente simple.

Si se añaden capas intermedias a un perceptrón simple, se obtiene un perceptrón multinivel. Configuración que se entrena mediante el algoritmo denominado retropropagación de error BP, o haciendo uso de algunas de sus variantes.

Si para una red de arquitectura multicapa se denomina x_i a las entradas, y_i a las salidas de la capa oculta, z_k a las de la capa final y globales de la red, t_k la salida objetivo, w_{ij} los pesos de la capa

oculta, θ_j sus umbrales, w'_{kj} los pesos de la capa de salida y θ'_k sus umbrales. La operación de un MPL con una capa oculta y neuronas de salida lineal se expresa como

$$z_k = \sum_j w'_{kj} y_j - \theta'_k = \sum_j w'_{kj} f \left(\sum_i w_{ji} x_i - \theta_j \right) - \theta'_k \quad (3.46)$$

Esta es la arquitectura más común de MPL, estructura que constituye un aproximador de funciones. Los aproximadores universales tales como polinomios de Lagrange, polinomios ortogonales de Chebyshev, aproximantes de Padé y series de Fourier truncadas pueden modelar procesos continuos de manera adecuada [Hornik 89, Chen 93, Haykin 99].

McCulloch [43] demostró que la modelación de una neurona como dispositivo umbral podía representar cualquier función booleana. En 1987, Denker [87] comprobó que toda función booleana encaja en una red unidireccional multicapa de un sólo nivel oculto. Posteriormente Lippmann [87], mostró que un perceptrón con dos capas ocultas bastaba para representar regiones de decisión arbitrariamente complejas. Lapedes [87] demostró que un perceptrón de dos capas ocultas es suficiente para representar cualquier función arbitraria no necesariamente booleana. Funahashi [89] demostró matemáticamente que un MLP convencional de una única capa oculta (ecuación 3.46), constituía en efecto, un aproximador universal de funciones [Chen 93, Haykin 99].

Teorema Funahashi

Sea $f(x)$ una función no constante, acotada y monótona creciente. K un subconjunto compacto acotado y cerrado de \mathfrak{R}^n . Sea un número real $\varepsilon \in \mathfrak{R}$, y sea un entero $k \in \mathbb{Z}$, tal que $k \geq 3$. En estas condiciones se tiene que

Cualquier mapeo $\mathbf{g}: \mathbf{x} \in K \rightarrow (g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_m(\mathbf{x})) \in \mathfrak{R}^m$, con $g_i(\mathbf{x})$ sumables en K , puede ser aproximado en el sentido de la topología L_2 en K por el mapeo entrada-salida representado por una red neuronal unidireccional (MPL) de k capas ($k-2$ ocultas), con $f(x)$ como función de transferencia de las neuronas ocultas y funciones lineales para las de las capas de entrada y de salida. Es decir,

$\forall \varepsilon > 0, \exists$ un MPL de las características anteriores, que implementa el mapeo

$$\mathbf{g}': \mathbf{x} \in K \rightarrow (g'_1(\mathbf{x}), g'_2(\mathbf{x}), \dots, g'_m(\mathbf{x})) \in \mathfrak{R}^m \quad (3.47)$$

de manera que

$$d_{L_2(K)}(\mathbf{g}, \mathbf{g}') = \left(\sum_{i=1}^m \int_K |g_i(x_1, \dots, x_n) - g'_i(x_1, \dots, x_n)| dx \right)^{1/2} < \varepsilon$$

(3.49)

Un MPL de una sola capa oculta puede aproximar hasta el nivel deseado cualquier función continua en un intervalo, por lo tanto las redes neuronales multicapas unidireccionales son aproximadores universales de funciones. A partir de la expresión que define la operación de este tipo de red

$$g'_k(\mathbf{x}) = \sum_j w'_{kj} y_j - \theta'_k = \sum_j w'_{kj} f \left(\sum_i w_{ji} x_i - \theta_j \right) - \theta'_k \quad (3.50)$$

donde $\mathbf{g}'(\mathbf{x})$ dada por el MPL representa una cierta función $\mathbf{g}(\mathbf{x})$ como un desarrollo en función sigmoideas $f(x)$, lo cual posee una analogía con la representación convencional de una función periódica como un desarrollo en series de Fourier de senoides.

Retropropagación del error

Una solución al problema de entrenar los nodos de las capas ocultas de redes con arquitectura multicapa, fue el algoritmo BP (Backpropagation). La deducción del BP es una consecuencia natural de expandir el algoritmo LMS a redes multicapa, por esto se plantea con un funcional de error similar al de la ecuación (3.43). Este, se deriva no sólo en función de los pesos de la capa de salida, sino también en función de los pesos de las neuronas ocultas haciendo uso de la regla de la cadena. En consecuencia, las funciones de transferencia de las neuronas deben ser derivables.

Sea un MPL de tres capas, cuya arquitectura se presenta en la figura 3.4, con entradas, salidas, pesos y umbrales de las neuronas definidas anteriormente. Entonces dado un patrón de entradas \mathbf{x}^μ , ($\mu = 1, \dots, p$) y considerando que la operación global de esta arquitectura se expresa como

$$z_k^\mu = \sum_j w'_{kj} y_j^\mu - \theta'_k = \sum_j w'_{kj} f\left(\sum_i w_{ji} x_i^\mu - \theta_j\right) - \theta'_k \quad (3.51)$$

Las funciones de activación de las neuronas ocultas $f(h)$ son de tipo sigmoide, con potencial postsináptico h . Si se parte de una función de error cuadrático medio, como en el caso de la ADALINE, se obtiene

$$E(w_{ji}, \theta_j, w'_{kj}, \theta'_k) = (1/2) \sum_\mu \sum_k \left[t_k^\mu - f\left(\sum_j w'_{kj} y_j^\mu - \theta'_k\right) \right]^2 \quad (3.52)$$

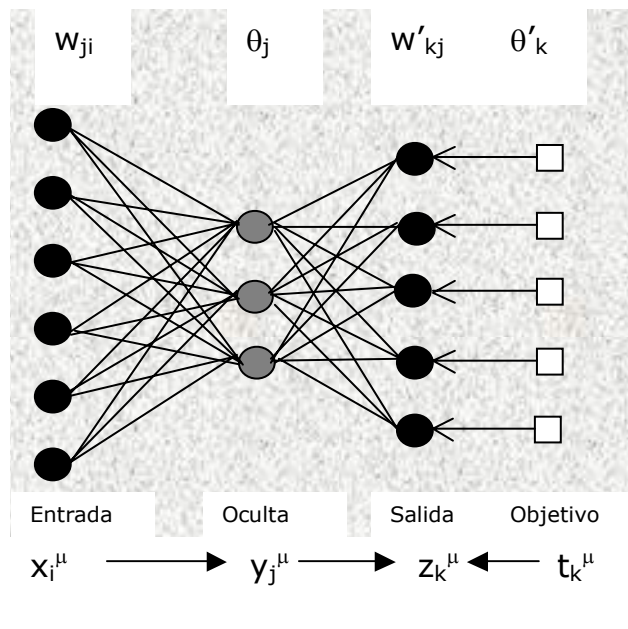


Figura 3.4 Arquitectura del perceptrón multicapa

La minimización se lleva a cabo mediante descenso por el gradiente, considerando en este caso un gradiente respecto a los pesos de la capa de salida y otro respecto a los de la capa oculta

$$\delta w'_{kj} = -\varepsilon \frac{\partial E}{\partial w'_{kj}} \quad \delta w_{ji} = -\varepsilon \frac{\partial E}{\partial w_{ji}} \quad (3.53)$$

Las expresiones de actualización de los pesos se obtienen considerando las dependencias funcionales y aplicando la regla de la cadena, así

$$\delta W'_{kj} = \varepsilon \sum_{\mu} \Delta_k^{\mu} y_j^{\mu}, \text{ con } \Delta_k^{\mu} = [t_k^{\mu} - f(v_k^{\mu})] \frac{\partial f(v_k^{\mu})}{\partial v_k^{\mu}} \quad (3.54)$$

$$\delta W_{ji} = \varepsilon \sum_{\mu} \Delta_j^{\mu} y_i^{\mu}, \text{ con } \Delta_j^{\mu} = \left(\sum_k \Delta_k^{\mu} w'_{kj} \right) \frac{\partial f(v_j^{\mu})}{\partial v_j^{\mu}} \quad (3.55)$$

La actualización de los umbrales, se realiza haciendo uso de estas mismas expresiones, asumiendo que el umbral es un caso particular de peso sináptico, cuya entrada es una constante igual a -1 .

En primer lugar se calcula la expresión Δ_k^{μ} , -denominada señal de error, por ser proporcional al error de la salida actual de la red- con la que se actualiza $\delta W'_{kj}$ de los pesos de la capa de salida. Luego se propagan hacia atrás los errores Δ_k^{μ} a través de las sinapsis, proporcionando de este modo las señales de error Δ_j^{μ} , correspondiente a las sinapsis de la capa oculta. Con éstas se actualiza δW_{ji} de las sinapsis ocultas. El algoritmo puede extenderse a topologías con varias capas ocultas siguiendo el mismo procedimiento.

Conviene señalar que el método BP, intrínsecamente tiene varias desventajas entre las que se suman, su lentitud de convergencia; degradación progresiva del aprendizaje o de su capacidad de generalización, denominado sobreentrenamiento; procesos de aprendizaje estancados en mínimos locales, al no garantizar el mínimo global de la función de error, etc. Para resolver estos inconvenientes, han surgido varias modificaciones. La primera variante la propusieron los propios creadores del modelo, al introducir un término denominado momento el cual consiste en añadir al cálculo de la variación de los pesos ecuación (3.53), un término adicional proporcional al incremento de la iteración anterior. Es decir, introducir un término inercial que amortigua los cambios bruscos de dirección en el gradiente y refuerza la modificación de los pesos si los gradientes están alineados

$$\delta W'_{kj}(t+1) = -\varepsilon \left. \frac{\partial E}{\partial W'_{kj}} \right|_t + \alpha \delta W'_{kj}(t-1) \quad \delta W_{ji}(t+1) = -\varepsilon \left. \frac{\partial E}{\partial W_{ji}} \right|_t + \alpha \delta W_{ji}(t-1) \quad (3.56)$$

con α entre 0 y 1, en general próximo a 1. De esta manera, si los incrementos en un determinado peso tienen siempre el mismo signo, las actualizaciones en cada iteración serán mayores. Sin embargo, si los incrementos en cierto peso oscilan a veces positivos otros negativos, el incremento efectivo acumulado se reduce al cancelarse. De esta forma, en zonas estrechas y profundas de la hipersuperficie de error, los pesos correspondientes a la dimensión estrecha sufren incrementos pequeños, mientras que los de las direcciones que desciende directamente al fondo se ven potencializados. Ésta es una manera de aumentar el ritmo de aprendizaje efectivo en determinada dirección.

Existen otras técnicas que permiten acelerar el aprendizaje, como los métodos de segundo orden que se basan en realizar el descenso utilizando también la información proporcionada por el ritmo de cambio de la pendiente, $H = \partial^2 E(W) / \partial W_{ij} \partial W_{kl}$. Los algoritmos de Gradiente Conjugado, Levenberg-Marquardt, Quickprop son algunos procedimientos más confiables en el sentido de que disminuyen la posibilidad de quedar atrapado en un mínimo relativo. Cada uno emplea distintas aproximaciones que soslayan el esfuerzo computacional que representa el cálculo directo de la matriz Hessiana [Shepherd 97]. En el método de gradiente conjugado, básicamente se calculan las primeras derivadas pero conservando la terminación cuadrática; la adaptación de este procedimiento a las redes multicapa es lo que le da el nombre al algoritmo, el cual propiamente es repropagación de gradiente conjugado. El algoritmo Levenberg-Marquardt es una modificación del método de Gauss-Newton también incorporado en redes multicapa, el inconveniente de esta variante es el almacenamiento requerido para guardar la matriz del Hessiano aproximada del Jacobiano, resultando un procedimiento poco práctico cuando el número de parámetros es muy grande [Shepherd 97]. El algoritmo Quickprop implementado por [Fahlman 89], consigue acelerar el aprendizaje procesando las entradas, asigna un ritmo de aprendizaje diferente a cada peso denominado ritmo adaptativo, en el

cual incorpora una heurística que da información sobre la forma de las segundas derivadas sin tener que derivar.

En este estudio se exploraron los algoritmos de aprendizaje de segundo orden, específicamente los tres mencionados. De la experimentación el que mejor índice de ejecución proporcionó fue la regla de propagación rápida QP de Fahlman.

3.6 Redes multicapa vs. métodos estadísticos

En términos estadísticos, las redes neuronales son estimadores no paramétricos que realizan aproximaciones denominadas de modelo libre [Haykin 99]. El método convencional de ajuste a una línea recta mediante mínimos cuadrados es un estimador paramétrico que impone al problema un determinado modelo de partida, cuyos parámetros se ajustan según la muestra disponible. A diferencia de los paramétricos, el MPL y muchas otras redes son estimadores de modelo libre pues no se impone ninguna forma funcional para el ajuste. Específicamente, el MLP realiza un tipo de regresión multidimensional no lineal, sin suposición inicial de una determinada forma funcional. En este sentido el MPL con frecuencia supera la regresión u otras técnicas tradicionales para espacios de dimensión alta.

Las limitaciones de las redes neuronales y del MPL en particular, pueden entenderse mediante el problema que también surge en la estimación paramétrica, el dilema de la varianza y el sesgo [Haykin 99]. En términos neuronales para realizar un ajuste óptimo el número de patrones de entrenamiento debe tender a infinito, pues para un conjunto finito, los estimadores no paramétricos suelen ser muy sensibles a los casos particulares de pares entra-salida seleccionados para el aprendizaje. Esto es debido a que la red neuronal, estimador libre, posee inherentemente una gran varianza. La red neuronal puede implementar numerosos diferentes mapeos, pero solo implementará el correcto si se entrena con conjunto de tamaño idealmente infinito. La única forma de controlar la elevada varianza que la red neuronal posee inicialmente, es introduciendo en su arquitectura algún tipo de sesgo o información conocida sobre el problema a resolver [Haykin 99]. Es decir, se trata de conseguir que el estimador neuronal, en principio no paramétrico, tienda a ser paramétrico en un cierto grado, de manera que de partida se encuentra en cierta medida viciado hacia el tipo de solución que interesa.

La disyuntiva que se plantea es precisamente en qué medida se debe viciar el modelo neuronal. Si se emplea una red de bajo sesgo y alta varianza, es decir, con muchos parámetros o pesos y sin introducir información conocida, es necesario para entrenar la red correctamente un número de ejemplos muy alto. Si el problema a resolver carece de cantidades suficientes de patrones, se deberá introducir cierto sesgo en la arquitectura, el cual debe corresponder fielmente con la realidad. Por lo tanto, una red de alto sesgo tendrá poca flexibilidad, pero una de alta varianza ajustará también el ruido presente en los datos. La solución ideal del problema varianza sesgo es encontrar en la evolución del error de generalización el punto que dará la arquitectura óptima.

Incorporar en la red neuronal restricciones que de antemano se sabe debe cumplir el problema, contribuye a reducir los grados de libertad. Estas restricciones suponen introducir información disponible sobre el problema en la propia arquitectura de la red. Las denominadas técnicas de regularización abordan el problema insertando restricciones que implica que el mapeo que implementa la red neuronal sea suave. Éstas pueden ser incluidas en la función error en forma de términos adicionales $\phi(w_\alpha)$, que miden la desviación de los resultados actuales de la restricción planteada

$$E[w_\alpha] = \frac{1}{2} \sum_{\mu} \sum_k [t_k^\mu - F_k^\mu(x)]^2 + \lambda \sum_{\alpha} \phi(w_\alpha) \quad (3.57)$$

siendo λ un parámetro de regularización, que controla el compromiso entre el grado de suavidad de la solución frente al nivel de los datos de entrenamiento que alcanza.

Otro tipo de restricciones que puede introducirse es limitar los pesos en el tamaño. En este caso se elige una función de regularización de la forma $\varphi(w_\alpha) = \varphi(w_\alpha)^2$, lo que conduce a una regla que implementa el decaimiento de pesos.

La aplicación de técnicas de regularización ha inspirado modelos neuronales como las funciones de base radial RBF (Radial Basis Functions) que son aproximadores locales.

Si bien no existe ningún estudio concluyente que permita asegurar la superioridad de un método de predicción no lineal en concreto sobre los otros, el análisis de los trabajos realizados por diversos investigadores proporciona elementos de juicio suficientes para escoger las redes neuronales.

3.7 Nota de observación sobre la estructura de los datos

Tanto la inferencia estadística como la bayesiana imponen condiciones a los datos muestrales. La estadística inductiva, por ejemplo, impone un muestreo aleatorio, lo que implica ciertas técnicas de extracción de las muestras y que no exista dependencia entre sus elementos ni disparidad entre sus valores. Estas prescripciones implican que los datos sean independientes, homogéneos y que la población tenga distribución normal tanto para la estimación puntual como por intervalos de confianza u otros. En consecuencia, se requiere verificar que tales condiciones se cumplan mediante diversos procedimientos desarrollados para tales fines (p.e., contraste por autocorrelación, tratamiento de valores atípicos, contraste de Pearson o de Kolmagorov-Smirnov, etc.). En los sistemas neuronales artificiales estas condiciones no son necesarias para la aplicación del modelo, lo que es una ventaja de las RNAs, pues como ya se ha reseñado, los registros sísmicos están afectados por factores que no son claramente separables.

3.8 Limitaciones de las redes neuronales artificiales

Para identificar la solución que podría proporcionar de forma eficiente una RNA hay que recurrir al procedimiento de prueba y error. Quizá el mayor inconveniente es que las RNAs no proporcionan información sobre la naturaleza del fenómeno. Una vez entrenada una RNA, su funcionamiento es el de una caja negra. Esto ha motivado una línea de intensa investigación en la que se trata de extraer las reglas que ha aprendido la red neuronal en el proceso de aprendizaje, para ello se emplea la equivalencia existente entre algunos modelos de redes unidireccionales y los sistemas basados en lógica borrosa.

Capítulo 4

Ambiente sismogeotécnico y base de datos

4.1 Caso México

4.1.1 Ambiente sismotectónico

Uno de los fenómenos en ambiente de convergencia que se presenta en el territorio mexicano responde a la subducción como resultado de la interacción entre las placas de Cocos, Rivera y Norteamericana, la cual se extiende a lo largo del occidente de México entre 92° y 108°W. En la figura 4.1, se ilustra la estructura regional de convergencia de este cinturón tectónico. Los estudios de mecanismo focal indican que las placas de Cocos y Rivera empujan debajo de la de Norteamérica en una dirección N30°E con velocidad de 2 a 8 cm/año [Nishenko 87]. En general, la mayor parte de la tensión acumulada a lo largo de las nueve zonas sismogénicas de subducción de la Costa del océano Pacífico es liberada por la ocurrencia de grandes sismos en fallas inversas.

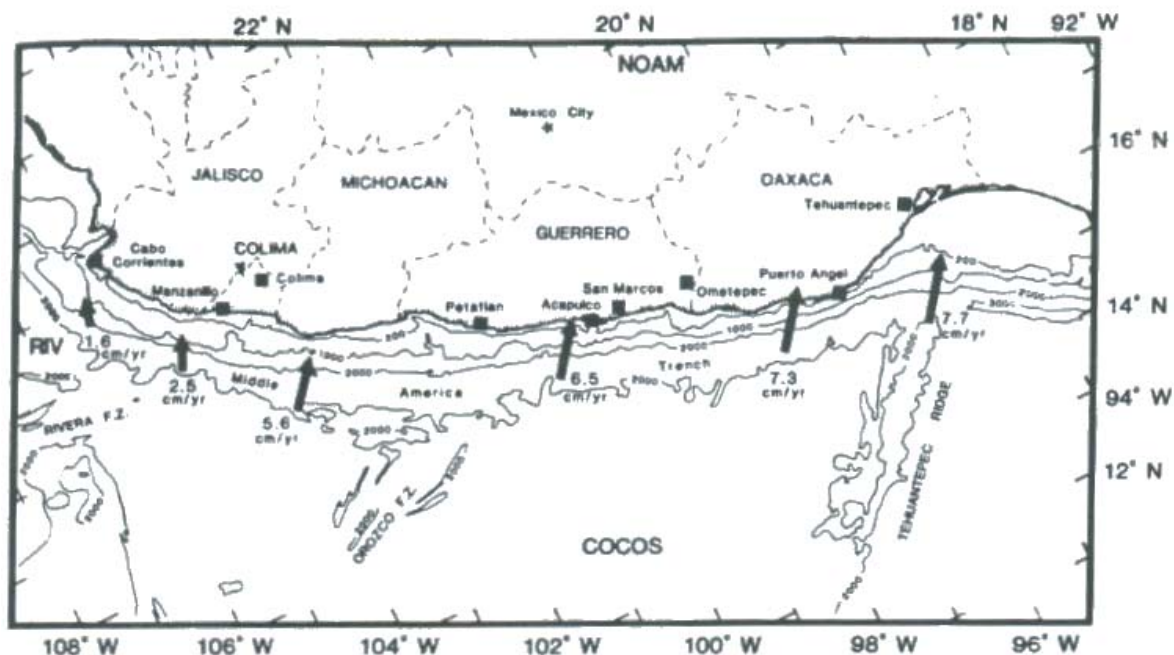


Figura 4.1 Localización de la zona de subducción de la Costa del océano Pacífico. Figura tomada de [Nishenko 87].

La Ciudad de México se ve afectada por sismos de diferentes orígenes como: sismos locales que se originan en el interior de la placa de Norteamérica, temblores de profundidad intermedia de falla normal causados por rompimientos de la placa de Cocos subducida, y sismos de subducción de la costa del océano Pacífico; la mayoría de los grandes sismos registrados son de subducción.

Las características geotécnicas del subsuelo de la ciudad, más la alta sismicidad a la que está expuesta, la han convertido en un referente internacional.

La Ciudad de México se encuentra emplazada en el llamado Valle de México que tiene aproximadamente un área de 2000 km² y está circundada por un relieve montañoso. La ciudad ha sido dividida en tres áreas geotécnicas denominadas: zona I o de Lomas, conformada principalmente por tobas volcánicas; zona II o de Transición constituida por secuencias de espesor variable hasta 20 m de estratos de arena y limo arenoso intercalados con capas de arcilla lacustre; zona III o de Lago, área con potentes depósitos de arcilla de 20 a 100 m de espesor, de muy baja densidad y propiedades mecánicas pobres, intercalada en ciertas porciones por una capa de arena limosa de alta consistencia. En esta última zona es donde se han evidenciado los mayores efectos locales. Los efectos de sitio tan marcados se deben a las propiedades casi-elásticas de las arcillas y al cambio de rigidez (impedancia) al pasar de, la roca basal (200 m de profundidad) a los depósitos profundos (aproximadamente 60 a 100 m) a los suelos lacustres.

El primer macrosismo para el cual se tuvieron instrumentos para registrar movimientos del terreno en algunos puntos de la ciudad de México fue el sismo de Michoacán del 19 de septiembre de 1985. A partir de esta fecha, se ha ampliado significativamente la red sísmica lo que ha permitido medir y almacenar información valiosa la cual se encuentra en la Base Mexicana de Datos de Sismos Fuertes (CD-ROM Vol. 2, 2000). En la Figura 4.2, se muestra la localización de algunas estaciones acelerográficas que registraron el evento de Michoacán y la zonificación geotécnica de la ciudad.

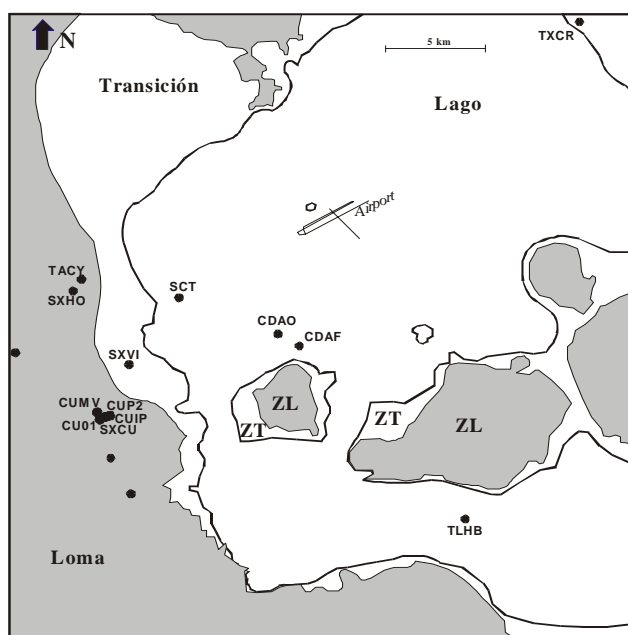


Figura 4.2 Zonificación geotécnica de la Ciudad de México y localización de algunas estaciones acelerográficas que registraron el sismo del 19 de septiembre de 1985

4.1.2 Base de datos

La información de movimientos fuertes que conforman la base de datos utilizada en esta tesis, fue extraída de la Base Mexicana de Datos de Sismos Fuertes. Los datos empleados corresponden a 516 acelerogramas, registrados en 107 estaciones para 21 eventos con magnitud entre 5.3 y 8.1, y distancias epicentrales de 133 a 498 km. La información general de los sitios de registro y los parámetros focales de los eventos sísmicos se presentan en las tablas 4.1 y 4.2 respectivamente.

Tabla 4.1 Información general de las estaciones sísmicas de la red de acelerógrafos de la Ciudad de México utilizadas en este estudio

Estación De registro	Latitud	Longitud	Localización geotécnica	Variable ficticia asignada a la localización geotécnica de la Ciudad México para el modelo neuronal
AE02	19.429	99.058	Lago	100
AL01	19.436	99.145	Lago	100
AO24	19.358	99.154	Transición	10
AP68	19.381	99.107	Lago	100
AR14	19.481	99.076	Lago	100
AU11	19.392	99.087	Lago	100
AU46	19.383	99.168	Transición	10
BA46	19.410	99.145	Lago	100
BL45	19.425	99.148	Lago	100
BO39	19.465	99.105	Lago	100
CA59	19.426	99.118	Lago	100
CADO	19.372	99.096	Lago	100
CDAF	19.366	99.086	Lago	100
CE18	19.340	99.085	Transición	10
CE32	19.386	99.059	Lago	100
CENA	19.314	99.170	Lomas	1
CI05	19.419	99.165	Lago	100
CJ03	19.410	99.157	Lago	100
CJ04	19.410	99.157	Lago	100
CJ04	19.410	99.157	Lago	100
CO47	19.371	99.170	Transición	10
CO56	19.422	99.159	Lago	100
COYS	19.347	99.168	Transición	10
CP28	19.439	99.084	Transición	10
CS78	19.366	99.226	Transición	10
CT64	19.488	99.140	Transición	10
CU01	19.330	99.183	Lomas	1
CU02	19.330	99.183	Lomas	1
CU03	19.330	99.183	Lomas	1
CU80	19.294	99.104	Lago	100
CU80	19.294	99.104	Lago	100
CUIP	19.330	99.183	Lomas	1
CUMV	19.330	99.183	Lomas	1
CUP1	19.330	99.183	Lomas	1
CUP2	19.330	99.183	Lomas	1
CUP3	19.330	99.183	Lomas	1
CUP4	19.330	99.183	Lomas	1
CUP5	19.330	99.183	Lomas	1
CH84	19.330	99.125	Lago	100
DFRO	19.405	99.166	Transición	10
DFVG	19.419	99.126	Lago	100
DM12	19.431	99.096	Lago	100
DR16	19.501	99.183	Transición	10
DX37	19.332	99.144	Lago	100
EO30	19.389	99.177	Transición	10
ES57	19.402	99.177	Transición	10
FJ74	19.299	99.210	Transición	10
GA62	19.439	99.140	Lago	100
GC38	19.316	99.160	Lago	100
GR27	19.475	99.160	Transición	10
HA41	19.418	99.180	Transición	10
HJ72	19.425	99.130	Lago	100
IB22	19.345	99.130	Lago	100
IM40	19.343	99.203	Transición	10
JA43	19.405	99.125	Lago	100
JC54	19.313	99.127	Lago	100
LI33	19.306	98.963	Lago	100
LI58	19.426	99.157	Lago	100
LV17	19.493	99.128	Lago	100
M15	19.283	99.125	Lago	100
ME52	19.438	99.182	Transición	10
MI15	19.283	99.125	Transición	10
MT50	19.425	99.190	Transición	10
MY19	19.346	99.043	Lago	100
NZ20	19.403	99.000	Lago	100
NZ31	19.417	99.025	Lago	100
PA34	19.202	99.049	Lago	100
PD42	19.406	99.100	Lago	100
PE10	19.390	99.132	Lago	100
PENR	19.332	99.041	Transición	10

Continúa tabla 4.1

RIDA	19.520	99.190	Transición	10
RIDI	19.360	99.060	Transición	10
RIDX	19.310	99.100	Transición	10
RM48	19.436	99.128	Lago	100
SCT B1	19.393	99.147	Lago	100
SCT B2	19.393	99.147	Lago	100
SCT1	19.393	99.147	Lago	100
SCT2	19.393	99.147	Lago	100
SI53	19.375	99.148	Lago	100
SP51	19.366	99.119	Lago	100
SXCU	19.326	99.182	Lomas	1
SXHO	19.000	99.205	Lomas	1
SXVI	19.358	99.171	Lago	100
TACY	19.403	99.194	Lomas	1
TE07	19.427	99.222	Lomas	1
TH35	19.279	99.000	Lago	100
TL08	19.450	99.134	Lomas	1
TL55	19.436	99.143	Lago	100
TLHB	19.279	99.008	Lago	100
TLHD	19.293	99.035	Lago	100
TP1	19.292	99.171	Lomas	1
TP13	19.292	99.171	Lomas	1
TXCL	19.480	98.990	Lago	100
TXCO	19.557	98.941	Lago	100
TXCR	19.518	99.805	Lomas	1
TXCH	19.430	98.950	Lago	100
TXS1	19.493	99.973	Lago	100
TXS2	19.492	98.770	Lago	100
TXSO	19.580	99.019	Lago	100
U121	19.365	99.226	Lomas	1
UC44	19.434	99.165	Lago	100
UI21	19.365	99.226	Lomas	1
UM29	19.381	99.125	Lago	100
VG09	19.454	99.123	Lago	100
VM29	19.381	99.125	Lago	100
XO36	19.271	99.102	Lago	100
XP06	19.420	99.135	Lago	100

Tabla 4.2 Sismos utilizados en el diseño de la RNA para la Ciudad de México

Fuente sísmica	Fecha	Latitud	Longitud	Profundidad (km)	Magnitud (Mw)	Nº de registros
Oaxaca	29/11/1978	16.010	99.590	19.0	7.8	2
Guerrero	14/03/1979	17.810	101.280	28.0	7.4	4
Centro México	24/10/1980	18.210	98.240	70.0	7.1	4
Michoacán	25/10/1981	18.010	102.110	14.0	7.2	5
Guerrero	07/06/1982	16.620	98.150	18.0	6.9	4
Centro México	04/06/1984	18.320	98.290	85.0	5.3	2
Guerrero	02/07/1984	16.760	98.510	33.0	6.1	2
Michoacán	19/09/1985	18.180	102.570	15.0	8.1	11
Guerrero	21/09/1985	17.820	101.670	15.0	7.5	9
Michoacán	30/04/1986	18.410	102.980	20.0	6.9	5
Michoacán	05/05/1986	18.240	102.530	19.2	5.9	2
Guerrero	07/06/1987	16.800	98.740	23.0	5.3	1
Guerrero	08/02/1988	17.660	101.040	19.2	5.8	33
Guerrero	10/03/1989	17.610	101.020	17.6	5.4	26
Guerrero	25/04/1989	16.810	99.380	19.0	6.9	57
Guerrero	02/05/1989	16.820	99.350	13.4	5.5	40
Guerrero	15/05/1993	16.740	98.360	5.0	6.0	48
Guerrero	24/10/1993	16.780	98.730	5.0	6.6	61
Guerrero	23/05/1994	18.310	100.520	23.0	6.2	65
Guerrero	14/09/1995	16.840	98.600	22.0	7.3	69
Oaxaca	30/09/1999	16.060	96.930	16.0	7.4	71

4.2 Caso Colombia

4.2.1 Ambiente sismotectónico

Toda la costa oeste de Sudamérica ha sido considerada como una zona de subducción activa de placas oceánicas debajo de la litósfera continental. Colombia es parte de esta zona por ubicarse en la

esquina noroccidental del continente suramericano entre los $4^{\circ} 13' 30''$ de latitud sur y $12^{\circ} 27' 46''$ de latitud norte y entre los $66^{\circ} 50' 54''$ y $79^{\circ} 01' 23''$ de longitud al oeste de Greenwich. En ella convergen la placa oceánica de Nazca, la placa continental de Sudamérica y la placa oceánica del Caribe; sometiendo al territorio colombiano a una tectónica compleja interplaca e intraplaca. Este proceso ha sido la fuerza dominante en el desarrollo de la estructura geológica del país, que según Etayo [86], está constituida por 34 fragmentos diversos de placas yuxtapuestas como se ilustra en la figura 4.3. Cada fragmento, denominado terreno, se caracteriza por poseer litología, estratigrafía, estilo estructural e historia propios. Estos terrenos están separados entre sí por fallas con direcciones predominantes N-S a NE-SW.

El carácter de esta tectónica compresiva ha sido confirmado por investigaciones neotectónicas, sismológicas y en particular por estudios de mecanismos focales de algunos sismos importantes. La dirección de compresión máxima parece ser relativamente uniforme con orientación NNW-SSE entre las placas Sudamérica, de corteza sílica, y la placa Nazca, de corteza simática.

Un mapa de terrenos estratigráficos-tectónico representa, por decirlo así, una "hipótesis de cartografía geológica" que trata de identificar e individualizar segmentos de corteza terrestre oceánica o continental en un territorio, con base en características estratigráficas propias de cada segmento y que a la vez sean distintas (pero no por cambios faciales) de las de los segmentos contiguos; los terrenos quedan demarcados por las trayectorias de megafracturas o suturas, visibles o inferidas. Se deduce que los terrenos tienen historias geológicas diversas y que, en función de sus posiciones tectónicas actuales, pueden ser considerados como autóctonos o alóctonos.

Comúnmente los mapas estructurales geológicos constituyen un dato muy importante, pues los aspectos tectónicos determinan qué tipos de estructuras están presentes y cómo se relacionan entre sí. En la carta estructural simplificada (figura 4.4), se observa que la distribución de las rupturas de superficie presenta una organización ligada a la interacción entre placas y a los esfuerzos tectónicos.

En términos globales el territorio colombiano puede considerarse dividido en dos regiones de distinta naturaleza. Hacia el oriente hay una región plana que es parte de la plataforma sudamericana y hacia el occidente una región montañosa formada por la trifurcación orogénica de la cordillera de los Andes conformada por las cordilleras Oriental, Central y Occidental; arreglos tectónicos de diferentes edades separados entre sí por cuencas intermontanas y con algunos sistemas orográficos aparentemente independientes del andino.

Entre las principales fallas tectónicas del territorio colombiano se destacan la falla Romeral, que se despliega a lo largo del borde occidental de la Cordillera Central, caracterizada por una elevada actividad sísmica en sus regiones central y austral; la falla Cauca-Patía, que se extiende paralela a la falla Romeral en el borde oriental de la Cordillera Occidental y cuya parte más activa es la central, en donde se han registrado sismos de magnitud $M_s=7.5$; la falla de Guaicaramo, considerada como el límite oriental del bloque Andino, exhibe un movimiento hacia el noreste de 1 cm/año. Estudios de mecanismos focales indican que es una falla dextral con esfuerzos principales máximos en dirección E-W. Otra falla de importancia es la de Bucaramanga, rasgo geológico de segundo orden al que se han adjudicado numerosos eventos sísmicos relacionados con su traza principal; en el sector comprendido entre las coordenadas 72.5° - 73.5° W y 6.3° - 7.3° N es donde presenta mayor actividad sísmica.

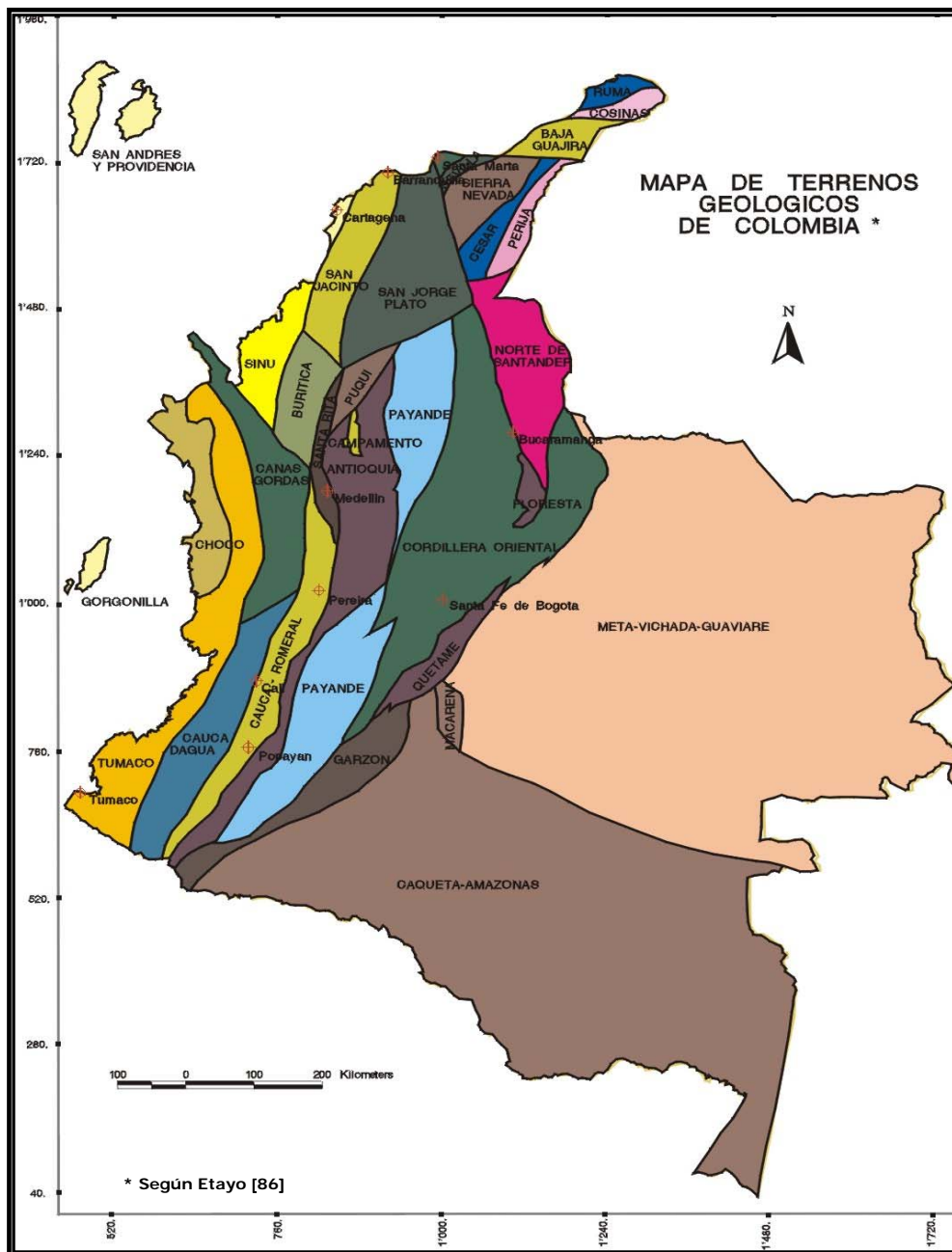


Figura 4.3 Mapa de terrenos estratigráfico-tectónico de Colombia



Figura 4.4 Mapa tectónico de Colombia contiene las principales fallas geológicas activas.

4.2.2 Distribución de la actividad sísmica en el país

De acuerdo con la información reportada en el catálogo de sismos INGEOMINAS y actualizada con los datos obtenidos de las publicaciones seriadas U.S. Department of the Interior Geological Survey, se ha localizado el registro sísmico histórico e instrumental sobre el mapa de terrenos geológicos de Colombia en función de la profundidad del foco y la magnitud, con el fin de ilustrar la actividad sísmica del país durante el período comprendido entre 1566 y 1998. Las cartas sísmicas elaboradas se presentan en las figuras 4.5, 4.6 y 4.7.

En los mapas de localización de epicentros, se aprecia para eventos de foco profundo que la sismicidad se hace más difusa del occidente hacia el este. También se observa en dicha carta que

existen vacíos de simicidad por ejemplo en la parte norte del terreno geológico de San Jacinto y San Jorge-Plato, así como patrones claramente definidos en la cordillera Occidental al oeste del eje cafetero, el Nido de Bucaramanga y la zona paralela al eje de la cordillera Oriental al suroeste del Nido de Bucaramanga.

La sismicidad cortical (menos de 40 km), muestra enjambres importantes; es notable la actividad en la zona de Choco, región noroccidental del país limítrofe entre Colombia y Panamá, donde hay una concentración epicentral significativa en número y magnitud. Además, una actividad alineada en dirección N-S paralela al valle del río Magdalena aproximadamente a 5° N de latitud y el frente del borde llanero que es otro rasgo visible a partir de la sismicidad también perceptiblemente apreciables.

El borde occidental del escudo de las Guayanas que constituye la zona plana del territorio nacional correspondiente a los Llanos Orientales y la Amazonía, conforman la región asísmica del país. Los pocos sismos que se han observado tienen hipocentro bastante profundo del orden de 300 km o más.

Según recuentos históricos, se considera que los sismos ocurridos a lo largo de la Cordillera Oriental son los que en Colombia han ocasionado la mayor destrucción y pérdidas de vidas humanas, esto antes del sismo del 25 de enero de 1999, que destruyó la Ciudad de Armenia y otras áreas urbanas de la zona cafetera de Colombia, asociado a un ramal de la falla Romeral. En los macizos montañosos de las Cordilleras Central y Occidental, también se observan los efectos de movimientos tectónicos internos en forma de sismos fuertes. En los cañones de los ríos Magdalena y Cauca continuamente se registran sismos originados en las múltiples fallas de la zona. Los departamentos de Nariño y Cauca se caracteriza por una larga historia sísmica con temblores de tipo tanto tectónicos como volcánicos. Cabe mencionar el sismo de Popayán ocurrido el 31 de marzo de 1983 con magnitud 5.5 y foco superficial, indujo intensidades locales del orden de IX en la escala Mercalli modificada y muchas replicas, y en consecuencia daños materiales superiores a los trescientos millones de dólares; la tragedia generada por la erupción del volcán del Nevado del Ruiz ocurrida el 13 de noviembre de 1985 (fenómeno asociado a las mismas fuerzas internas de la Tierra) ocasionó daños materiales directos por valor superior a doscientos millones de dólares y más de 22.000 pérdidas humanas; el sismo de Páez del 6 de junio de 1994, con epicentro en las estribaciones del Nevado del Huila y magnitud de 6.4 con foco a 10 km de profundidad, dio origen a enormes deslizamientos de laderas que a su vez conformaron grandes avalanchas cósmicas que arrasaron poblaciones con saldo de miles de desaparecidos. Además, en el Océano Pacífico, a lo largo de la costa, se observa un cinturón de alta sismicidad. De esta franja se destacan el terremoto de Tumaco de enero de 1906 con magnitud $M_s=8.9$ (uno de los más fuertes de la historia) y el terremoto de Tumaco de 1979 de magnitud $M_s=7.9$.

En las profundidades del macizo de Santander, en la Cordillera Oriental, se encuentra una zona activa bien definida, en la cual se localiza un "Nido" de sismos a profundidades intermedias conocido como "Nido de Bucaramanga". Se trata de una zona limitada de un radio no mayor de 15 km, dentro del cual se concentran hipocentros de sismos con profundidades entre 150-170 km. El Nido de Bucaramanga está ubicado cerca de las coordenadas 6.8° N y 73.1° W y aunque se encuentra en una zona de relativa baja actividad sísmica, el Nido propiamente dicho, es catalogado como una de las zonas de mayor actividad sísmica en el mundo, junto con las bien conocidas regiones sísmicas de Brancha en los Cárpatos Orientales e Indo-Kush en el Pamir. El Nido de Bucaramanga es relacionado con un posible punto de unión triple de las placas Caribe, Sudamericana y Nazca.

Con base en el marco geotectónico regional, la historia sísmica de la región contenida en los catálogos sísmicos histórico e instrumental y la distribución focal en planta, se han establecido siete sismofuentes en el territorio colombiano. Estas se demarcan en la figura 4.8. Así, las zonas sísmicas del país están definidas por las sismofuentes de Subducción, del Viejo Caldas, Piedemonte llanero, Nido de Bucaramanga, Darien, Muriondó y Frontino.

De acuerdo con la información sísmica tanto histórica como instrumental, se ha identificado para la mayoría de los eventos sísmicos un marcado control de la distribución de las intensidades por las estructuras geológicas regionales, indicando una anisotropía de la atenuación. Para Colombia la dirección del eje mayor de las isosistas coincide, en general, con la dirección de los sistemas de fallas

y los cuerpos geológicos (N-S a NE-SW), dicha tendencia pone de manifiesto el efecto de las fallas y los cuerpos geológicos en la atenuación del movimiento.

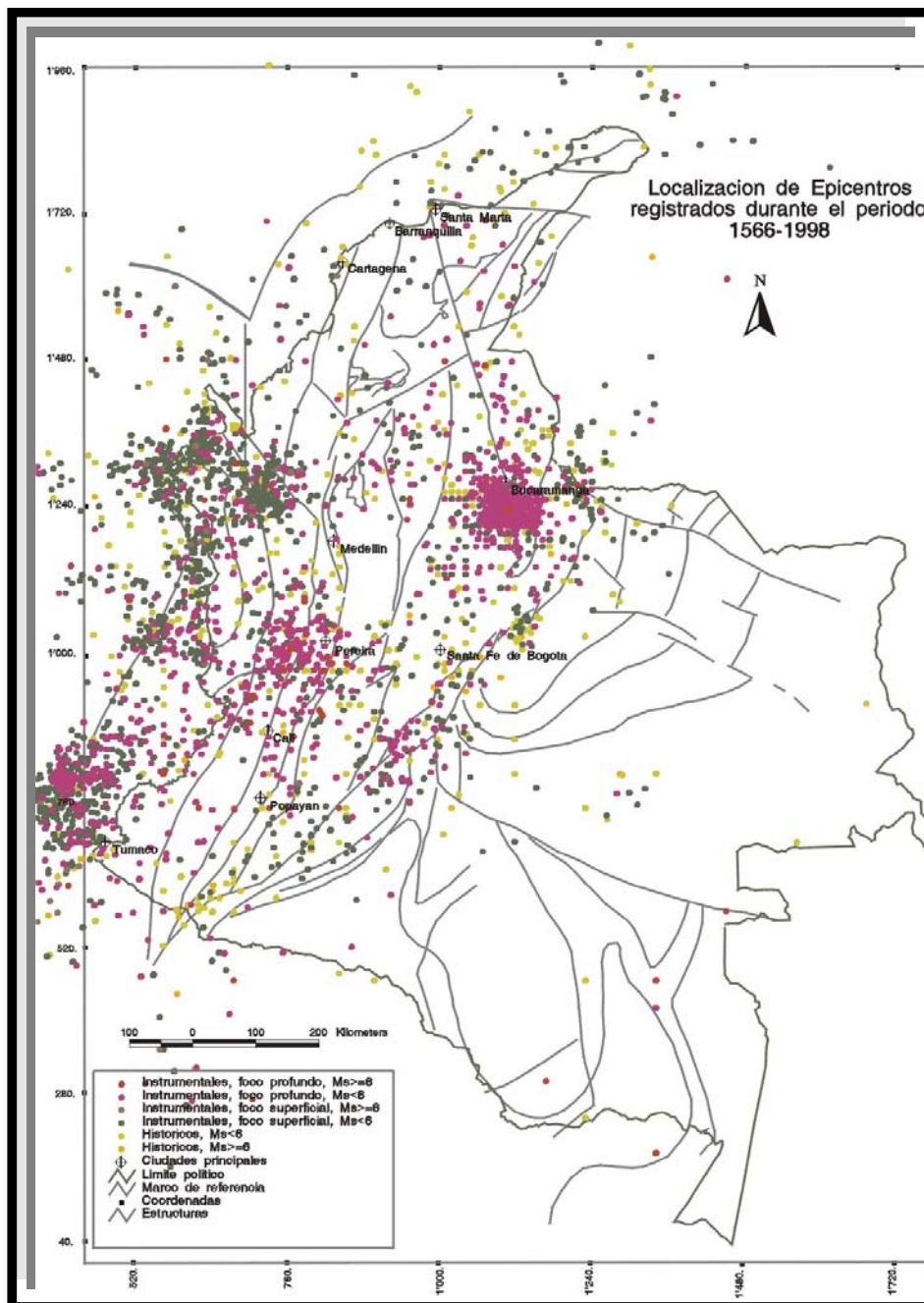


Figura 4.5 Carta sísmica del registro histórico e instrumental sobre el mapa de terrenos geológicos de Colombia en función de la profundidad del foco y la magnitud para el período comprendido entre 1566 y 1998.

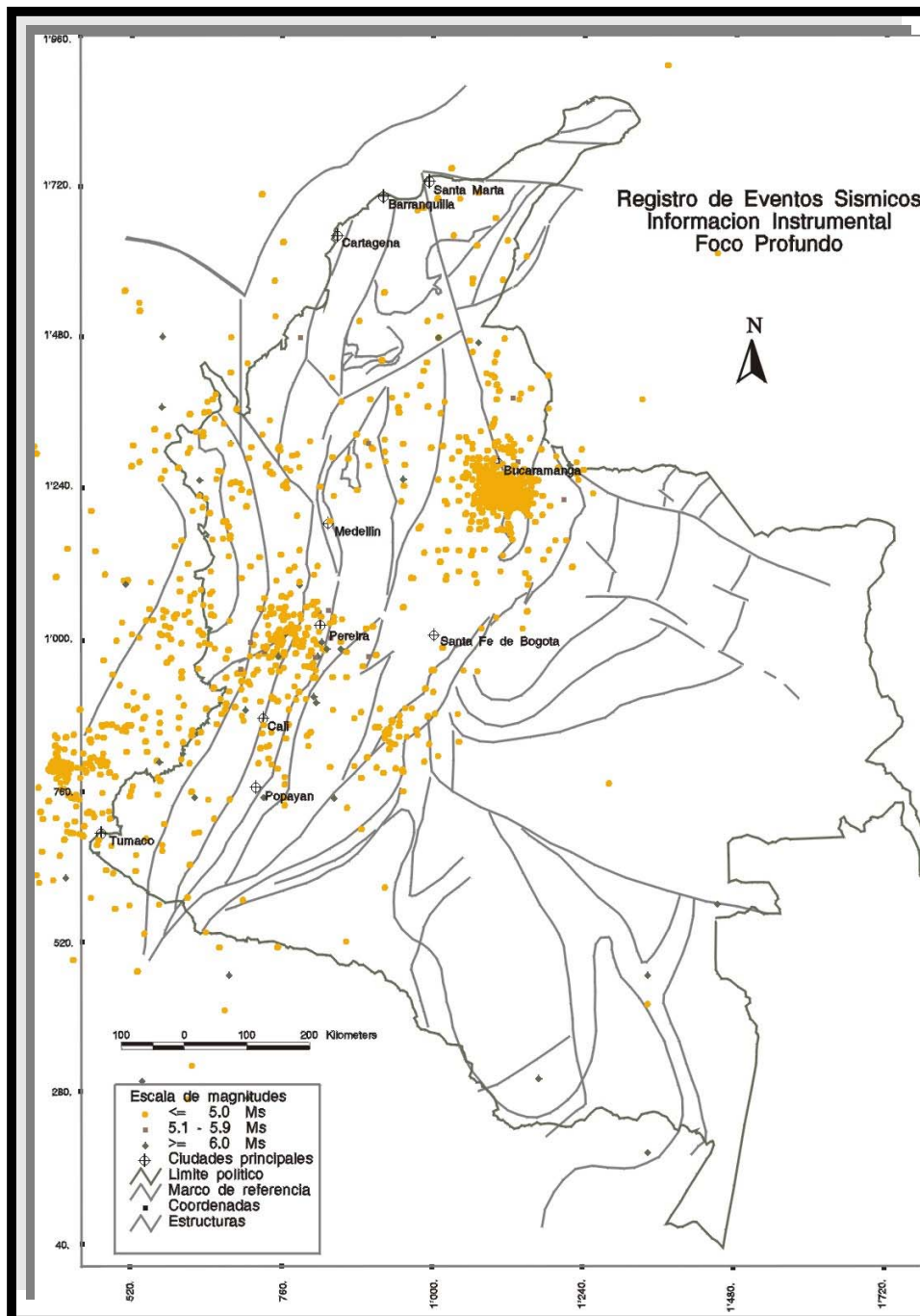


Figura 4.6 Distribución focal en planta de los eventos sísmicos con foco profundo sobre el mapa de terrenos geológicos de Colombia en función de la magnitud.

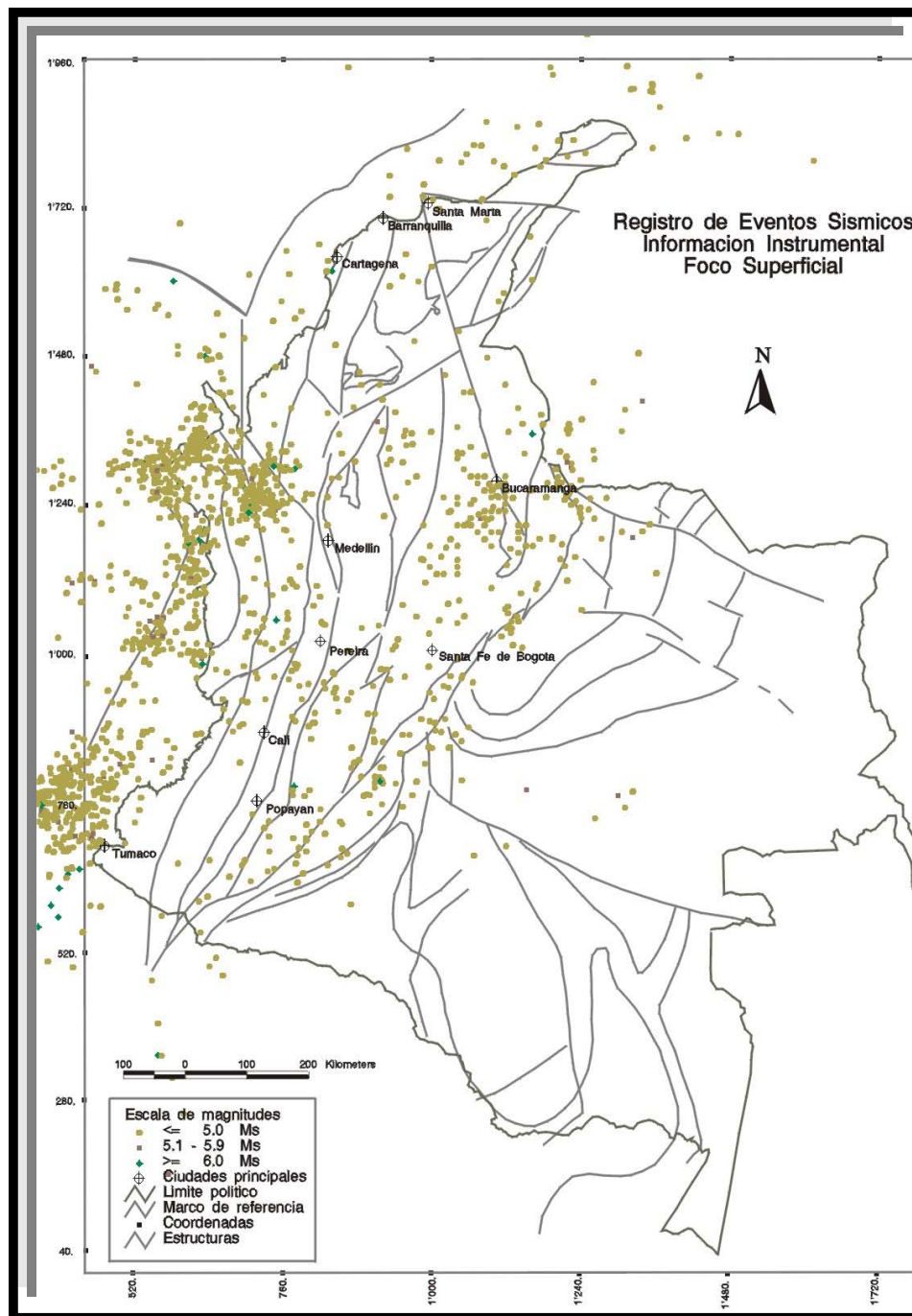


Figura 4.7 Distribución focal en planta de los eventos sísmicos con foco superficial sobre el mapa de terrenos geológicos de Colombia en función de la magnitud.

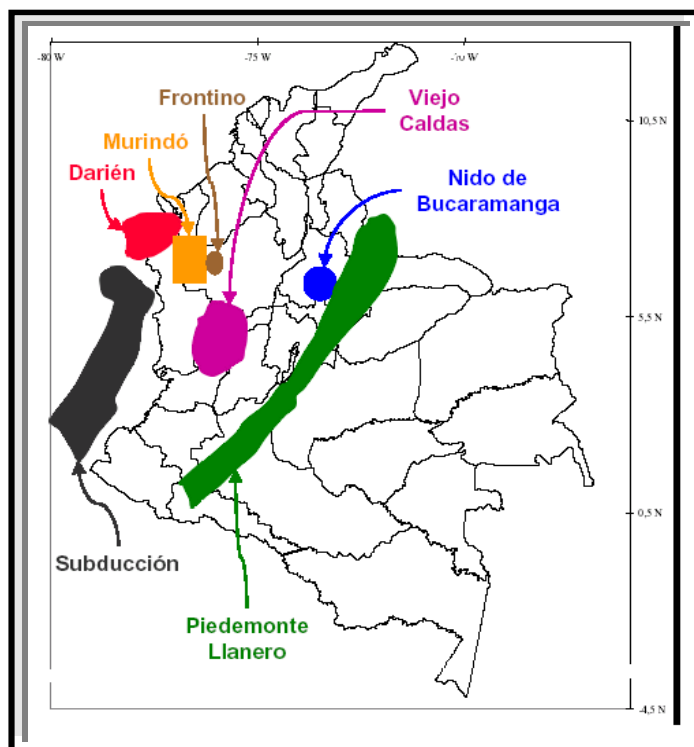


Figura 4.8 Localización de las principales fuentes sísmicas en el territorio colombiano

4.2.3 Base de datos

Para el caso colombiano la base de datos empleada está conformada por 280 acelerogramas registrados en 197 estaciones procedentes de 59 eventos sísmicos con magnitudes entre 4.1 y 7.2, la información general de esta muestra se presenta en las tablas 4.3 y 4.4. El registro instrumental fue proporcionado por la Red Nacional de Acelerógrafos de Colombia - INGEOMINAS que opera desde 1993. La mayoría de los eventos considerados están asociados a la sismofuente del Nido de Bucaramanga y a la del Viejo Caldas, mientras que los generados a profundidades menores de 35 km, aparecen relacionados directamente con fallas activas de la región.

Tabla 4.3 Información general de las estaciones de la Red Nacional de acelerógrafos de Colombia que registraron los eventos utilizados en este estudio

Estación de registro	Latitud	Longitud	Localización Geotécnica	Velocidad de onda de cortante estimadas (m/s ²)	Estación de registro	Latitud	Longitud	Localización Geotécnica	Velocidad de onda de cortante estimadas (m/s ²)
ANAP005	4.586	74.518	Roca	760	ESTERE	4.930	73.080	Roca	760
ANDE	5.660	75.910	Roca	760	FILA	5.290	75.560	Roca	760
ANDES011	5.659	75.907	Roca	760	FILA010	5.300	75.570	Roca	760
ANSE6006	5.250	75.820	Roca	760	FILAD001	5.290	75.560	Roca	760
ANSER	5.250	75.820	Roca	760	FILAD002	5.290	75.560	Roca	760
BETA004	2.720	75.440	Roca	760	FILAD13	5.297	75.568	Roca	760
BETA005	2.720	75.440	Roca	760	FUQUE	5.472	73.739	Roca	760
BETAN001	2.720	75.440	Roca	760	FUQUE018	5.472	73.739	Roca	760
BOG1007	4.640	74.080	Suelo	180	FUQUE03	5.472	73.739	Roca	760
BOG1008	4.642	74.080	Suelo	180	FUQUE07	5.472	73.739	Roca	760
BOG1009	4.642	74.080	Suelo	180	GALER	4.710	73.990	Roca	760
BUC1	7.100	73.130	Roca	760	GO2008	4.710	73.990	Roca	760
BUC1001	7.070	73.072	Roca	760	GUAD	5.060	74.590	Roca	760
BUC1008	7.100	73.060	Roca	760	GUAD004	5.065	74.590	Roca	760
BUC1014	7.100	73.060	Roca	760	GUAD005	5.065	74.590	Roca	760

Continúa tabla 4.3

BUC1020	7.070	73.072	Roca	760	GUAD006	5.065	74.590	Roca	760
BUC1021	7.070	73.072	Roca	760	GUAD008	5.065	74.590	Roca	760
BUC1026	7.070	73.072	Roca	760	GUADU	5.060	74.590	Roca	760
BUC2	7.140	73.110	Suelo	180	GUADU003	5.060	74.590	Roca	760
BUC402	7.062	73.166	Suelo	180	GUADU006	5.060	74.590	Roca	760
BUC6001	7.141	73.110	Suelo	180	GUADU006	5.060	74.590	Roca	760
BUC6002	7.070	73.104	Suelo	180	GUADU008	5.060	74.590	Roca	760
BUIS	7.141	73.119	Suelo	180	IBA1	4.450	75.230	Suelo	180
BUIS002	7.141	73.119	Suelo	180	IBAR003	4.475	75.254	Roca	760
BUIS006	7.141	73.119	Suelo	180	IBASU003	4.444	75.232	Suelo	180
CA003	8.200	73.400	Roca	760	IPIA	0.810	75.580	Roca	760
CALAR	4.460	75.620	Roca	760	MAZP	4.810	75.690	Suelo	180
CALIM	3.880	76.580	Roca	760	MEL1001	5.560	74.890	Roca	760
CANAP	4.549	74.514	Roca	760	NOBS	5.770	72.930	Roca	760
CANAP	4.55	74.514	Roca	760	NORC	5.560	74.890	Roca	760
CANDE	5.660	75.910	Roca	760	NORCA	5.560	74.890	Roca	760
CANDES021	5.660	75.910	Roca	760	NORCA001	5.560	74.890	Roca	760
CANSE	5.260	75.810	Roca	760	OCA009	8.200	73.400	Roca	760
CARBE	4.280	74.430	Roca	760	OCA009	8.200	73.400	Roca	760
CARBE002	4.280	74.430	Roca	760	OCA015	8.200	73.400	Roca	760
CBETA002	2.710	75.430	Roca	760	OCA024	8.200	73.400	Roca	760
CBUC1	7.100	73.060	Roca	760	OCANA	8.200	73.400	Roca	760
CBUC1044	7.100	73.060	Roca	760	PAMP	7.380	72.640	Roca	760
CBUC1046	7.100	73.060	Roca	760	PAMP002	7.379	72.637	Roca	760
CBUC2002	7.140	73.110	Suelo	180	PAMP07	7.379	72.637	Roca	760
CBUIS	7.140	73.110	Suelo	180	PAS2	1.210	77.260	Roca	760
CBUIS002	7.140	73.110	Suelo	180	PENSI	5.390	75.160	Roca	760
CBUIS024	7.14	73.110	Roca	760	PENSI015	5.390	75.160	Roca	760
CCALA	4.510	75.830	Roca	760	PENSI018	5.390	75.160	Roca	760
CCC1012	7.878	72.508	Roca	760	PENSY023	5.388	75.160	Roca	760
CFILA	5.297	75.567	Roca	760	PENSY024	5.388	75.160	Roca	760
CFLAN	4.690	75.620	Suelo	180	PLAY005	6.293	74.942	Suelo	180
CFONE	5.400	73.790	Roca	760	PLAY006	6.293	74.942	Suelo	180
CGUAD	5.070	74.590	Roca	760	POISS	2.470	76.590	Suelo	180
CGUAD001	5.065	74.590	Roca	760	POP2	2.470	76.590	Suelo	180
CGUAD003	5.070	74.590	Roca	760	POPAR	2.440	76.630	Roca	760
CIBA1	4.440	75.230	Suelo	180	PRAD	3.760	74.890	Roca	760
CIBA2	4.460	75.240	Roca	760	PUTP	4.790	75.690	Suelo	180
CLROS012	4.848	74.360	Roca	760	QUETA	4.330	73.860	Roca	760
CMAZP	4.81	75.690	Suelo	180	QUETA002	4.330	73.860	Roca	760
CNORC	5.560	74.890	Roca	760	QUETA003	4.330	73.860	Roca	760
CNORC002	5.560	74.884	Roca	760	RIOSU	5.420	75.720	Roca	760
CPAMP009	7.370	72.640	Roca	760	RIOSU001	5.420	75.720	Roca	760
CPAMP018	7.379	72.637	Roca	760	RIOSU002	5.420	75.720	Roca	760
CPENS	5.390	75.160	Roca	760	ROSA	4.848	74.326	Roca	760
CPLAT002	2.380	75.910	Roca	760	ROSA003	4.848	74.326	Roca	760
CPOP3	2.424	76.625	Roca	760	ROSA004	4.848	74.326	Roca	760
CPRAD	3.763	74.893	Roca	760	ROSA004	4.848	74.326	Roca	760
CPRAD002	3.760	74.890	Roca	760	ROSA005	4.848	74.326	Roca	760
CPTEJ	3.230	76.440	Suelo	180	ROSA006	4.848	74.326	Roca	760
CPTJ	3.23	76.440	Suelo	180	ROSA007	4.848	74.326	Roca	760
CQUET	4.328	73.861	Roca	760	ROSA008	4.848	74.326	Roca	760
CROSA002	4.850	74.330	Roca	760	ROSAL	4.850	74.330	Roca	760
CSANBAR2	4.640	74.080	Roca	760	ROSO	4.840	75.680	Suelo	180
CSEVI	4.250	75.930	Roca	760	SANLUI8	6.002	74.994	Roca	760
CSLUI	6.050	74.990	Roca	760	SARD002	8.080	72.790	Roca	760
CSLUI016	6.052	74.995	Roca	760	SARD004	8.057	72.974	Roca	760
CSTRC	4.880	75.630	Suelo	180	SARD005	8.057	72.794	Roca	760
CSUES	5.717	72.957	Roca	760	SARDIN	8.080	72.790	Roca	760
CSUES001	5.720	72.960	Roca	760	SEVI	4.250	75.930	Roca	760
CSUES002	5.720	72.960	Roca	760	SEVIL	4.260	75.940	Roca	760
CSUES002	5.720	72.960	Roca	760	SLUI001	6.002	74.994	Roca	760
CSUES006	5.72	72.96	Roca	760	SLUIS	4.760	73.140	Roca	760
CSUES007	5.720	72.960	Roca	760	STHE001	6.230	75.550	Roca	760
CSUES007	5.720	72.960	Roca	760	SUES018	5.717	72.957	Roca	760
CSVIC005	6.340	75.270	Roca	760	SUES060	5.720	72.960	Roca	760
CSVIC014	6.342	75.271	Roca	760	SUESC	5.717	72.957	Roca	760
CTOCH	3.650	76.090	Roca	760	SUESC61	5.717	72.950	Roca	760
CTORI	2.98	76.290	Suelo	180	TADO	5.260	76.560	Roca	760
CUC1021	7.878	72.509	Roca	760	TRUJIL	4.230	76.330	Roca	760
CUC1022	7.878	72.509	Roca	760	TUTU	2.980	76.290	Roca	760
CUC1023	7.878	72.509	Roca	760	TUTUN	5.740	76.550	Roca	760
CUCT003	7.860	72.520	Roca	760	UIS001	7.140	73.110	Roca	760
CUCU1009	7.878	72.509	Roca	760	VHER	5.040	75.120	Roca	760
CUCU1010	7.878	72.509	Roca	760	VHER003	5.032	75.120	Roca	760
CUCUR	7.878	72.509	Roca	760	VHER004	5.032	75.120	Roca	760
CVHER002	5.040	75.120	Roca	760	VHER006	5.032	75.120	Roca	760
CHIN001	4.550	73.630	Roca	760	VIL	4.180	73.640	Roca	760
CHIN002	4.554	73.631	Roca	760	VILHER	5.036	75.122	Roca	760
CHING	4.550	73.630	Roca	760	VILHER7	5.036	75.122	Roca	760
CHING002	4.550	73.630	Roca	760	VILR006	4.177	73.640	Roca	760
CHING004	4.550	73.630	Roca	760	VILL009	4.177	73.640	Roca	760
CHING07	4.550	73.630	Roca	760	VILLRC7	4.150	73.640	Roca	760
DAGUA	3.660	76.710	Roca	760					

Tabla 4.4 Eventos que conforman la base de datos para el estudio de la atenuación sísmica con procedimientos neuronales para el caso colombiano

Fuente sísmica	Fecha	Latitud	Longitud	Profundidad (km)	Magnitud (Mw)	Nº de registros
Yali - Antioquia	18/11/1993	6.73	74.77	17	4.7	1
Sismo de Páez - Cauca	06/06/1994	2.85	76.078	1	6.5	2
Sismo de Bolívar - Antioquia	07/02/1994	5.96	75.99	1	4.7	4
Nido de Bucaramanga	10/12/1994	6.77	73.25	150	6	3
Riosucio - Chocó	13/09/1994	7.24	76.7	28	6.3	5
Nido de Bucaramanga	18/02/1994	6.8	73.42	160	5.9	1
Santa Isabel - Antioquia	19/09/1994	7	74.6	1	5.2	1
Nido de Bucaramanga	26/02/1994	6.81	73.1	160	5.6	1
Pueblo Rico - Risaralda	28/02/1994	5.09	76.11	120	4.9	2
Nido de Bucaramanga	28/03/1994	6.74	73.05	160	5.5	1
S.J. Rioseco - Cundinamarca	04/06/1995	4.87	74.61	8	4.8	5
Tauramena - Casanare	06/04/1995	5.14	72.77	1	5.4	2
Calima - Valle	08/02/1995	4.13	76.74	102	6.8	17
Nido de Bucaramanga	15/03/1995	6.92	73.08	155	5.1	1
Tauramena - Casanare	19/01/1995	5.01	72.95	25	6.7	14
Risaralda - Caldas	19/08/1995	5.11	75.71	110	6.8	6
Tauramena - Casanare	20/01/1995	5.13	72.89	1	5.5	4
Otanche - Boyacá	20/05/1995	5.64	73.89	150	5.9	5
Garagoa - Boyacá	21/01/1995	5.02	73.13	1	5.6	3
La Uruía - Boyacá	22/01/1995	5.09	72.9	1	5.7	6
Nido de Bucaramanga	22/04/1995	6.83	73.18	150	5.5	2
Tauramena - Casanare	23/01/1995	5.04	72.77	1	5	2
El Provenir	23/01/1995	4.94	72.9	1	4.9	2
Nido de Bucaramanga	23/02/1995	6.78	73.17	155	5.7	2
Yopal - Casanare	23/04/1995	5.17	72.36	1	5.7	1
Nido de Bucaramanga	23/07/1995	6.83	73.13	165	5.6	1
Juradó - Chocó	04/11/1996	7.43	77.53	1	6.5	4
Sipí - Chocó	10/09/1996	4.5	76.7	100	4.7	5
Nido de Bucaramanga	12/10/1996	6.83	73.13	150	5.3	5
El Dovio - Valle del Cauca	17/11/1996	4.5	76.29	110	5.1	6
Nido de Bucaramanga	01/01/1997	6.81	73.1	160	6.2	10
Génova Quindío	02/09/1997	3.96	75.87	230	7.2	7
Nido de Bucaramanga	07/03/1997	6.78	73.29	150	5.9	7
Nido de Bucaramanga	11/05/1997	6.8	73.09	160	5.9	8
Nido de Bucaramanga	01/06/1997	6.83	72.97	160	6.1	6
Nido de Bucaramanga	01/06/1997b	6.85	73.05	170	6.2	4
Génova - Quindío	11/12/1997	4	75.95	220	6.8	11
Cubarral - Meta	17/07/1997	3.83	74.09	1	5.4	4
Sipí - Chocó	19/02/1997	4.54	76.52	120	6.1	5
Jamundí - Valle	19/11/1997	3.28	76.49	100	4.9	3
Nido de Bucaramanga	24/06/1997	6.81	73.03	160	5.5	2
Nido de Bucaramanga	01/18/1998	6.82	72.97	160	5	1
Landázuri - Santander	06/03/1998	6.38	73.9	1	5.4	13
Cimitarra	08/03/1998	6.27	73.89	1	5.5	10
Nido de Bucaramanga	30/03/1998	6.75	73.18	150	5.7	7
Nido de Bucaramanga	26/10/1998	6.81	73.13	160	5.6	8
Bosconia - Cesar	18/03/1999	10.07	73.6	100	6.1	4
Nido de Bucaramanga	14/04/1999	6.82	73.14	170	6.1	15
Pulí - Cundinamarca	15/05/1999	4.67	74.75	18.7	4.8	7
Guayabetal - Cundinamarca	01/06/1999	4.29	73.76	1	5.2	4
Sativasur - Boyacá	17/07/1999	6.1	72.74	1	5.6	1
Nido de Bucaramanga	11/08/1999	6.9	73.15	160	5.4	2
Sismo del Quindío	25/01/1999	4.41	75.71	1	6.2	11
Nido de Bucaramanga	15/01/2000	6.87	73.08	160	4.9	1
Fortul - Arauca	17/10/2000	6.67	72.04	1	5.9	1
Nido de Bucaramanga	04/02/2000	6.81	73.15	150	4.1	2
Nido de Bucaramanga	05/02/2000	6.77	73.21	160	6	1
Nido de Bucaramanga	24/03/2000	6.79	73.03	161	5.4	1
El Bagre - Antioquia	28/03/2000	7.68	74.53	1	5.3	1

4.3 Ciudad de Medellín

4.3.1 Información sismogeotécnica

El Valle de Aburrá se encuentra localizado sobre la Cordillera Central en la parte septentrional de los Andes Colombianos a unos 6° 15' de Latitud norte y 75° 30' de Longitud oeste. La forma del valle es la de una depresión con dirección predominante N-S. La Ciudad Medellín se encuentra situada dentro del Valle de Aburrá y se ha desarrollado en la planicie aluvial del río que lleva su mismo nombre y en

las laderas que delimitan el valle; tiene un área aproximadamente de 110 km². Las vertientes empinadas que rodean la ciudad están conformadas por varias unidades litológicas de rocas ígneas y metamórficas. En el área urbana afloran cuerpos petrogénicos de cuarzodiorita, gabro, anfibolita, dunita y neis; todos con diferentes grados de intemperismo generando diferentes tipos de suelos residuales y gran diversidad de condiciones geotécnicas.

La actividad sísmica a la que está sujeta la Ciudad de Medellín proviene del complejo marco geodinámico de interacción de las placas tectónicas que se encuentran en el territorio nacional. En general, la ciudad ha sido perturbada durante su historia por sismos de intensidad media que provienen de diferentes fuentes lejanas como la zona de Benioff del viejo Caldas, el nido de Bucaramanga, los ramales sur de la falla Romeral y la falla Murindó. Aunque los eventos sísmicos como los de Murindó ocurridos el 17 y 18 de octubre de 1992, han producido aceleraciones máximas en la roca del orden de 15 gales, los daños han sido cuantiosos, en términos de pérdidas materiales que se estimaron en 14 millones de dólares.

La diversidad de suelos que pueden encontrarse y el amplio intervalo de variación de sus propiedades geotécnicas, identificadas en estudios geológicos y geotécnicos, han llevado a establecer catorce zonas homogéneas con base en la respuesta sísmica superficial de los perfiles de suelos característicos de la ciudad, como se ilustra en la figura 4.9 [Universidad Eafit, 99].

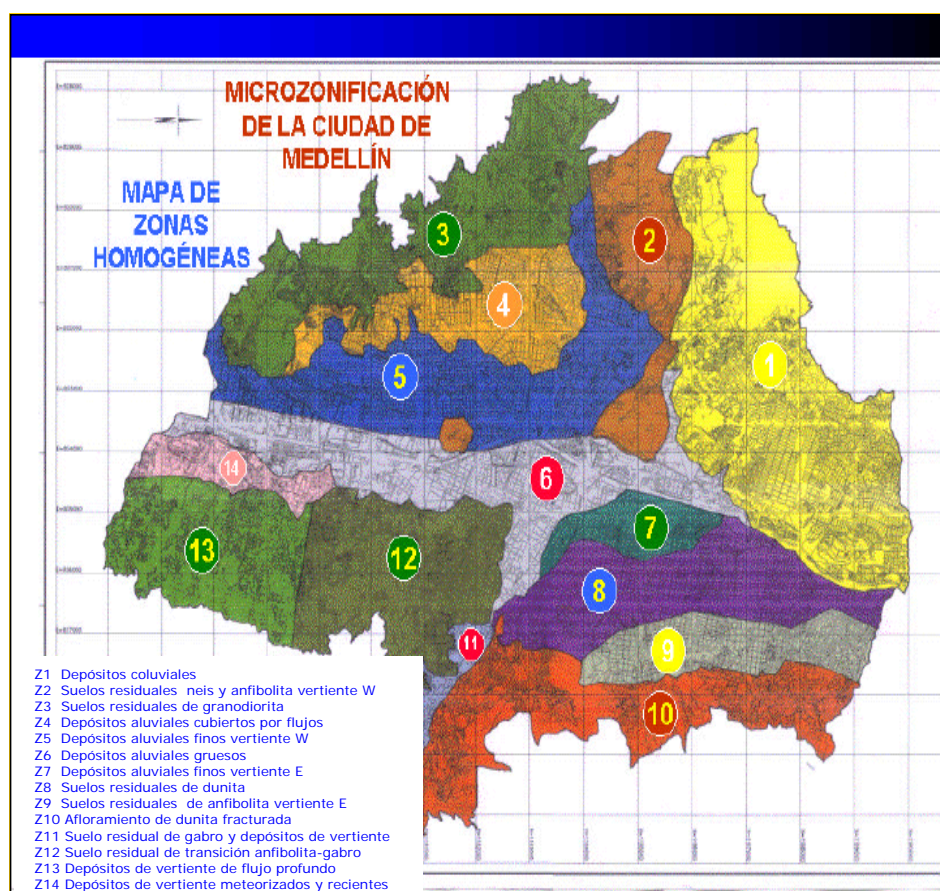


Figura 4.9 Unidades litológicas que afloran en el área urbana de la ciudad. Figura tomada de [Universidad Eafit, 99].

La compleja distribución espacial de suelos y rocas más las diferentes características de las fuentes sísmicas que pueden afectar la ciudad, dificultan el desarrollo de expresiones que describan la

atenuación de la energía sísmica, implicando para este caso una evaluación para cada pareja fuente-sitio.

4.3.2 Base de datos

La base de datos utilizada, fue proporcionada por la Red acelerográfica de Medellín (RAM). La red sísmica del municipio de Medellín está conformada por 22 estaciones en superficie, dos de las cuales tienen emplazados equipos profundos en pozo hasta el basamento rocoso. Es considerada como la segunda red local más importante de Latinoamérica después de la instalada en Ciudad de México. Comenzó a operar en noviembre de 1996 bajo la responsabilidad de la Universidad EAFIT. La información general de las estaciones de registro sísmico de la RAM, se presenta en la tabla 4.5, y en la figura 4.10 la distribución de estas estaciones en el área urbana de la ciudad cuyas localizaciones cubren un gran número de las formaciones geológicas presentes.

El conjunto de datos está conformado por 278 registros acelerográficos derivados de 26 eventos sísmicos con magnitud entre 3.8 y 6.8; la mayoría proceden de la sismofuente denominada Nido de Bucaramanga y corresponden a los registrados por la red acelerográfica hasta el año 2000. La localización y los parámetros focales de los 26 sismos se presentan en la tabla 4.6.



Figura 4.10 Localización de las estaciones sísmicas de la RAM. Figura tomada de Universidad Eafit. 991.

Tabla 4.5 Información general de la red local de acelerógrafos RAM

Estación de registro	Latitud	Longitud	Altitud (m)	Categoría del tipo de perfil según código NSR 98	Velocidad de onda de cortante medidos in situ (m/s ²)
MAN	6.22	75.61	1605	C	360
SPE	6.35	75.50	1499	B	760
SEM	6.23	75.55	1877	C	360
EVH	6.26	75.55	1730	B	760
ECC	6.27	75.61	1450	C	360
COM	6.28	75.57	1563	C	360
EAU	6.30	75.60	2026	C	360
CSJ	6.25	75.55	1620	B	760
ESE	6.19	75.53	2820	A	1500
FMI	6.28	75.60	1537	B	760
SOL	6.24	75.55	1660	C	360
EVT	6.24	75.54	1822	B	760
ISA	6.26	75.55	1730	C	360
UEA	6.20	75.58	1518	C	360
POL	6.25	75.62	1633	C	360
EMO	6.28	75.54	2170	C	360
LIC	6.24	75.55	1688	B	70
ITM	6.18	75.66	2048	C	360
ISJ	6.37	75.59	1427	B	760
UDM	6.23	75.61	1595	D	180
UNA	6.26	75.58	1488	C	360
EET	6.20	75.56	1650	C	360

Tabla 4.6 Sismos usados en el diseño de la RNAs para la Ciudad de Medellín

Fuente sísmica	Fecha	Latitud	Longitud	Profundidad (km)	Magnitud (Mw)	Nº de registros
Camposeco - Santander	22/11/1996	6.16	74.28	1	4.7	5
Villanueva - Santander-	01/01/1997	6.73	73.18	160	6.2	5
Sipí - Choco	19/02/1997	4.62	76.58	100	6.0	12
Zapatoca - Santander	07/03/1997	6.76	73.23	150	5.8	15
Villanueva - Santander	11/05/1997	6.74	73.18	150	5.8	14
Cepita - Santander	11/06/1997	6.81	72.96	160	6.0	12
Santos - Santander	11/06/1197	6.84	73.09	150	6.1	14
Zapatoca - Santander	24/06/1997	6.87	73.23	150	5.5	5
Roncesvalles - Quindío	02/09/1997	3.93	75.68	150	6.5	16
Génova - Quindío	11/12/1997	4.00	75.95	220	6.8	19
Cimitarra - Santander	06/03/1998	6.28	73.89	30	5.4	16
Cimitarra - Santander	08/03/1998	6.30	73.88	30	5.4	15
Cimitarra - Santander	27/03/1998	6.37	73.95	20	4.5	8
Bucaramanga - Santander	30/03/1998	6.70	72.93	175	4.5	5
Abriaquí - Antioquia	13/07/1998	6.56	76.05	20	3.8	15
San Andrés - Santander	26/10/1998	6.85	72.90	160	5.7	4
Los Santos - Santander	04/12/1998	6.8	73.18	150	5.6	10
Los Santos - Santander	10/12/1998	6.81	73.18	160	5.4	3
Córdoba - Quindío	25/01/1999	4.40	75.71	40	6.0	14
Zaragoza - Antioquia	10/02/1999	7.52	74.88	30	4.6	13
Los Santos - Santander	14/04/1999	6.78	73.08	160	6.1	11
Sativasur - Boyacá	17/07/1999	6.07	72.75	30	5.5	7
Urrao - Antioquia	18/09/1999	6.50	76.16	30	4.7	8
Piedecuesta - Santander	08/11/1999	6.93	73.07	160	6.7	15
Betania - Antioquia	16/01/2000	5.74	76.06	30	6.2	11
Nido de Bucaramanga	05/02/2000	6.76	72.93	159	4.8	6

Capítulo 5

Cuantificación neuronal de parámetros sísmicos

Considerando el potencial que ofrecen las RNAs con respecto a los otros procedimientos establecidos en IA, en el sentido ya indicado de identificar conexiones al utilizar relaciones no lineales entre los datos, modelar cualquier tipo de distribución, manejar datos con redundancia e inconsistencia en la información, tolerar las imprecisiones e incertidumbres del entorno, reconocer patrones entre los ejemplos ilustrativos, y la significativa ventaja de los sistemas neuronales con relación a los métodos convencionales, al no requerir compromisos entre los datos y alguna función prescrita ni la condición de independencia de las variables; se eligió como metodología para estimar algunos parámetros dinámicos que permiten cuantificar la respuesta superficial del suelo.

5.1 Variables sensoriales

El primer paso en la modelación neuronal es adquirir conocimiento de los aspectos que pueden influir en el resultado del proceso de adopción de medidas, y seleccionar los parámetros externos que conforman el vector de entrada.

Con base en el análisis preliminar de los datos y considerando que para México no existe correlación entre la forma espectral y la región de la zona de subducción donde ocurre el temblor, se separó la base de datos en función de la zonificación geotécnica de la Ciudad. No obstante, al observar formas espectrales poco uniformes asociadas a largos períodos dominantes y varios modos de vibración en registros sísmicos de estaciones localizadas en roca, para ser más congruente con lo observado, se dividió la base de datos en tres grupos con fronteras establecidas por el valor del periodo dominante, así

- Grupo 1, para los casos en que la característica fundamental de vibración de cada uno de los perfiles de suelo o período dominante de este, sea menor igual a 1.5 segundos ($T_p \leq 1.5$ s)
- Grupo 2, para los casos en que la característica fundamental de vibración de cada uno de los perfiles de suelo o periodo dominante de este, se ubique en el intervalo 1.5 a 3 segundos ($1.5 < T_p < 3$ s)
- Grupo 3, para los casos en que la característica fundamental de vibración de cada uno de los perfiles de suelo o el periodo dominante de éste, sea igual o mayor a 3 segundos ($T_p \geq 3$ s)

Para el caso colombiano dada la diversidad de las fuentes sísmicas y la variabilidad del movimiento en superficie en función de la profundidad del foco, se disgregó la base de datos con este criterio, pues se sabe que los sismos superficiales se atenúan mucho más rápido que los sismos profundos. Las agrupaciones se denominaron de foco superficial y de foco lejano.

El disgregar la información e implementarla en la red como grupos con determinadas formas espectrales decrece la varianza en la red neuronal, pues se está introduciendo información a priori a la red sin afectar su flexibilidad y reducir los grados de libertad.

Las variables significativas del fenómeno para la representación del modelo neuronal denominadas variables sensoriales, se seleccionaron con base en las características sustantivas de los aspectos sabidos -fuente, trayectoria, sitio- que interviene en la respuesta superficial del suelo (capítulo 2). Estas se listan en la tabla 5.1.

Tabla 5.1 Parámetros externos

Variabes sensoriales de entrada	Aspecto a caracterizar
1. Magnitud momento, M_w 2. Profundidad, h 3. Azimut del plano de falla 1, ϕ_1 (strike 1) 4. Buzamiento del plano de falla 1, δ_1 (dip 1) 5. Dirección del movimiento del plano de falla 1, λ_1 (slip1) 6. Azimut del plano de falla 2, ϕ_2 (strike 2) 7. Buzamiento del plano de falla 2, δ_2 (dip2) 8. Dirección del movimiento del plano de falla 2, λ_2 (slip2)	Fuente sísmica
9. Distancia epicentral, R 10. Coseno del ángulo horizontal que se forma entre la línea que une el epicentro y estación de registro con la línea de coordenada geográfica longitud, α_i	Trayectoria
11. Frecuencia dominante de vibración, f_p 12. Frecuencia de esquina, f_e 13. Frecuencia de corte, f_c 14. Relación entre la velocidad máxima y la aceleración máxima, V_{max}/A_{max} 15. Factor de atenuación en función de la frecuencia dominante $Q(f_p)$ 16. Factor de atenuación en función del período dominante, $Q(T_p)$ 17. Duración de la fase intensa del movimiento, D_i 18. Intensidad de Arias, I_A 19. Aceleración máxima, A_{max} 20. Clasificación del tipo de suelo, C_s , variable ficticia asignada a la localización geotécnica (tabla 4.1) 21. Velocidad de onda de cortante del subsuelo donde está localizada la estación de registro, β (tabla 4.5)	Condiciones locales

Los parámetros que conforman el vector de entrada en su mayoría fueron obtenidos directamente de registros reales de diferentes temblores y seleccionados con base en el soporte proporcionado por las investigaciones sobre el fenómeno sísmico reportado en la literatura técnica. De esta manera la modelación neuronal se fundamentó en información sísmológica.

Para modelar el concepto de atenuación se utilizó el factor de forma definido por Aki en Joyner [88]

$$Q(f_p) = (29.4) \left(1 + (f_p / 0.3)^{2.9} \right) \div (f_p / 0.3)^2$$

La variable sensorial denominada clasificación del tipo de suelo C_s , es una variable ficticia que se introdujo para clasificar los datos de acuerdo con las condiciones locales de Ciudad de México. Así, para la zona I, $C_s=1$; para la zona II, $C_s=10$; y para la zona III, $C_s=100$.

Con el fin de representar la anisotropía en la atenuación de las ondas sísmicas se utilizó el coseno del ángulo horizontal que se forma entre la línea que une el epicentro y estación de registro con la línea de coordenada geográfica longitud. Pues para los Andes colombianos, en general, la energía liberada por los sismos se atenúa más rápidamente en sentido E-W que en sentido N-S.

La frecuencia de esquina, f_e y frecuencia de corte, f_c obtenidas del espectro de amplitudes de Fourier de cada acelerograma son parámetros que acotan la rama ascendente, la de movimiento fuerte y la de desvanecimiento del espectro y por tanto contribuyen a indicar los puntos de quiebre en los valores espectrales.

Cabe mencionar que las frecuencias de esquina y de corte dejan de ser significativas como parámetros sensoriales para el caso de la Ciudad de Medellín, ya que las mayores amplificaciones están concentradas en bajos periodos; y que para el caso colombiano por ausencia de información sobre los mecanismos focales de los eventos sísmicos considerados, los parámetros que definen la

geometría de la fuente sísmica –azimut, buzamiento y dirección del movimiento- no fueron incluidos en la modelación neuronal.

El cociente entre la velocidad máxima y la aceleración máxima (V_{max}/A_{max}), es una cantidad asociada con el movimiento de frecuencias diferentes que puede ser interpretada como el periodo de vibración de una onda armónica equivalente y provee una indicación de cuáles periodos de movimiento son más significativos.

Es importante resaltar que los parámetros externos utilizados en la modelación son propiedades directas de cada sitio en particular y de cada epifoco, lo que conlleva a una evaluación bastante conservadora y a una predicción del movimiento fuerte para cada pareja fuente-sitio.

5.2 Modelo neuronal

El diseño de una red neuronal artificial consiste en: 1) definir la arquitectura, la función de procesamiento, la función de entrada, la función de transferencia, el algoritmo de aprendizaje y el criterio de error; 2) monitorear la respuesta; y 3) analizar la sensibilidad para identificar las contribuciones de las variables de entrada.

No hay un procedimiento claro para diseñar la topología de una RNA. La arquitectura de las redes que se consideran óptimas se consigue mediante experimentación (prueba y error). Se puede escoger entre diversos modelos de ajuste y comprobar, a través de los coeficientes de error de predicción y correlación, el mejor de ellos, o bien realizar una nueva selección, hasta encontrar el que mejor responda, esto es el que minimice el error de predicción. Sin embargo, teniendo en cuenta que los RNA han demostrado conseguir buenos acercamientos en el campo de la predicción, se experimentó con los algoritmos de aprendizaje de segundo orden para el perceptrón multicapa como aproximador universal de funciones.

En este estudio, la eficacia del sistema se midió con el error promedio en la etapa de prueba antes que la red se degradara y entrara en el proceso de sobreaprendizaje, es decir, se aplicó como criterio de generalización, el concepto de validación cruzada ampliamente utilizado en la fase de desarrollo de una red neuronal supervisada. El proceso de validación cruzada consiste en entrenar y validar a la vez para detenerse en el punto óptimo correspondiente al menor error promedio logrado por la red aunque sea un mínimo local. No obstante, para algunas topologías no fue necesario aplicar la técnica de parada temprana y se dejó que el aprendizaje discurriera hasta un error sensiblemente constante.

El fenómeno de sobreaprendizaje se presenta en la fase de entrenamiento después de que la red se ha adaptado progresivamente al conjunto de aprendizaje. En un momento dado el sistema se ajusta demasiado a las particularidades de los patrones empleados en el entrenamiento, aprendiendo incluso el ruido presente en los datos, en consecuencia crece el error de generalización o de los patrones empleados en la predicción que obviamente son diferentes a los utilizados en el entrenamiento. Bajo estas circunstancias la red no ajusta correctamente el mapeo, sino que simplemente está memorizando los patrones del conjunto de aprendizaje.

La explicación del sobreajuste radica en que las RNAs, son estimadores no lineales poderosos, capaces de modelar situaciones muy complejas. En las herramientas lineales, por ejemplo en ajuste mediante polinomios, la complejidad del modelo viene dada por el número de parámetros libres a ajustar (coeficientes), mientras que en los sistemas no lineales como son las RNAs, el modelo depende tanto del número de parámetros como de su valor actual. Al ser las RNAs modelos de elevada complejidad pueden memorizar cualquier cosa, motivo por el cual incurren algunas veces en sobreaprendizaje.

Como resultado de las combinaciones entre las alternativas de decisión relacionadas con las diferentes opciones disponibles de las RNAs y los factores que intervienen en el movimiento fuerte del suelo, se identificaron los modelos neuronales que permiten estimar las relaciones funcionales para predecir parámetros que caracterizan el movimiento fuerte en los casos mencionados. Las características de las redes se presentan en las tablas 5.2, 5.3 y 5.4 para los casos modelados;

Ciudad de México, Ciudad de Medellín y Colombia respectivamente. En la figura 5.1, a manera de ejemplo, se ilustra mediante una representación esquemática las variables involucradas en el cálculo de la salida de una neurona y la topología de la red neuronal para el caso de la aceleración máxima de la Ciudad de México.

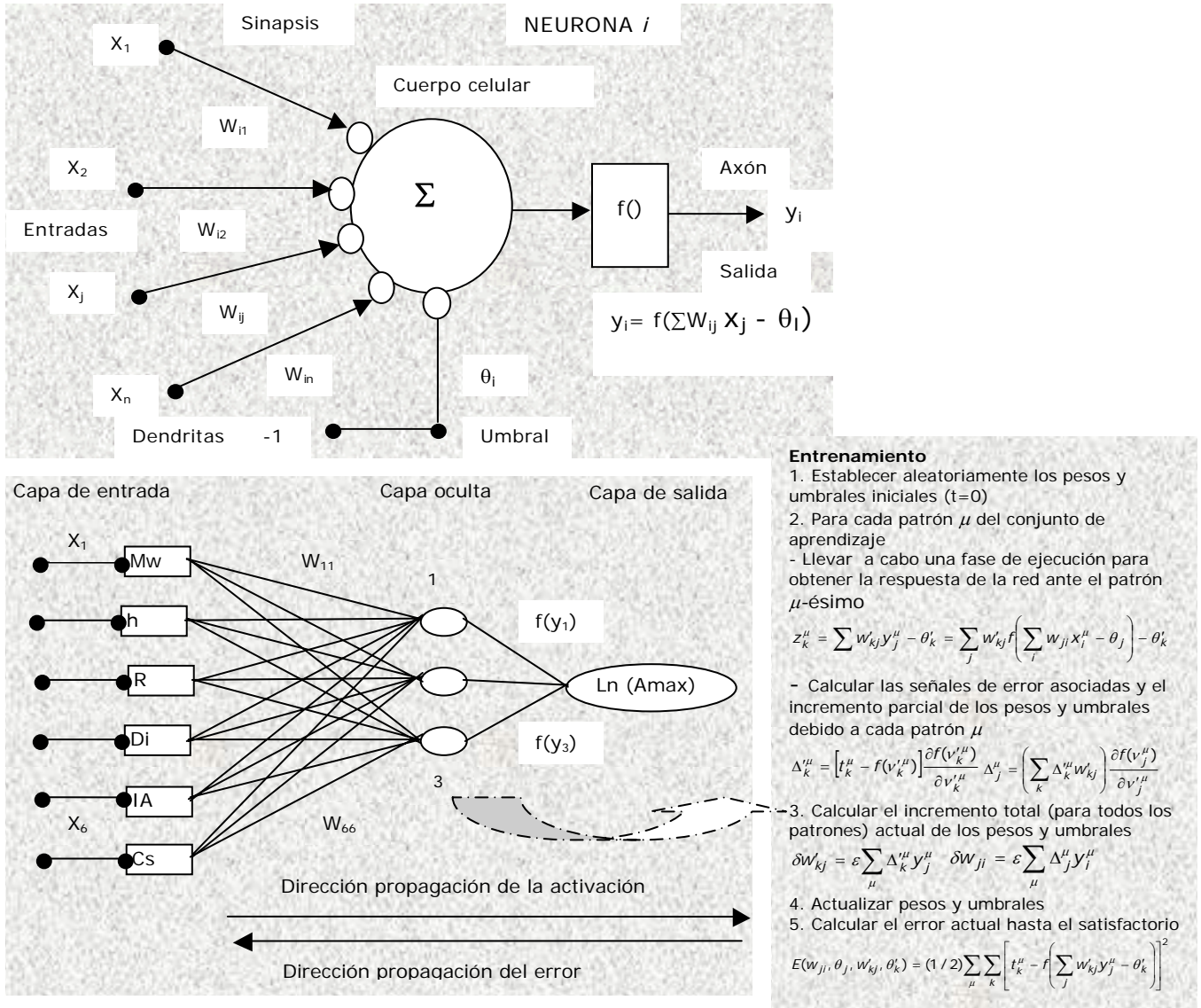


Figura 5.1 Representación esquemática de las variables involucradas en el cálculo de la salida de una neurona y topología de la red neuronal para el caso de la aceleración máxima de la Ciudad de México con un recorrido de la información dentro de la red hacia delante.

Los datos para la transformaciones algorítmicas se dan en la tabla 5.5; y los parámetros estadísticos empleados para medir el entrenamiento y la capacidad de generalización de las redes con patrones diferentes a los utilizados en el aprendizaje se presentan en la tabla 5.6. El error de estimación muestra la desviación estándar de los residuales.

Una idea más intuitiva del proceso de aprendizaje del perceptrón multicapa, puede resumirse de la siguiente forma:

Partiendo de un punto aleatorio $W(0)$ del espacio R^{n_w} donde n_w es el número de parámetros de la red, pesos más umbrales; el proceso de aprendizaje desplaza el vector de parámetros $W(n-1)$ en el espacio R^{n_w} siguiendo la dirección negativa del gradiente del error en dicho punto, alcanzando así un nuevo punto en dicho espacio $W(n)$, que estará más próximo al mínimo de la función error que el anterior. El proceso continúa hasta que se encuentre un mínimo de la función error E , lo cual sucede cuando $\frac{\partial E}{\partial w} \approx 0$.

En este momento, los parámetros dejan de sufrir cambios significativos de una iteración a otra y el proceso de aprendizaje finaliza.

Durante el proceso de aprendizaje la red extrae las características de la muestra para poder así responder correctamente a patrones diferentes. Esto se conoce como la capacidad de la red para generalizar las características presentadas en el conjunto de datos. De este modo, el perceptrón multicapa representado en la figura 5.1, con alimentación hacia delante que consta de seis celdas sensoriales (magnitud momento, profundidad del sismo, distancia epicentral, duración de la parte intensa del movimiento, intensidad de Arias y variable ficticia de clasificación del tipo de suelo); función de procesamiento normalizada; una capa oculta con tres elementos de procesamiento artificial; algoritmo de aprendizaje QP supervisado; función de entrada producto punto; función de transferencia sigmoide; 494 patrones de entrada para el modo de convergencia y 22 patrones en fase de ejecución, produjo un ajuste de parámetros libres que condujo a una generalización del 98.5% con error en términos de la desviación estándar del logaritmo natural de la aceleración máxima establecida en la señal de salida de 0.21.

Este resultado es un ejemplo más, que sitúa al perceptrón multicapa como una clase de funciones para interpolar relaciones no lineales entre datos de entrada y salida, y considerarlo como un aproximador universal, en el sentido de que cualquier función continua sobre un compacto de R^n puede aproximarse con un perceptrón multicapa con al menos una capa oculta de neuronas.

Tabla 5.2 Topología básica de las RNAs para el caso mexicano

Característica	Duración parte intensa Ciudad de México	Aceleración máxima Ciudad de México y para sitios por fuera del Valle de México	Espectros de respuesta elásticos Ciudad de México
Arquitectura	Cascada	Multicapa con alimentación hacia adelante	Cascada
VARIABLES EXTERNAS	Seis	seis	Diecinueve
Función de procesamiento	Máximos y mínimos	Normalizada	Normalizada
Número de capas ocultas	Una	Una	Una
Neuronas en capa oculta	Dieciséis	tres	Diecinueve
Algoritmo de aprendizaje	Propagación rápida QP	Propagación rápida QP	Propagación rápida QP
Función de entrada	Producto punto	Producto punto	Producto punto
Función de Transferencia	Sigmoide	Sigmoide	Sigmoide
Criterio de error	Medio cuadrático	Medio cuadrático	Medio cuadrático
Capa de salida	Ln (Di)	Ln (Amax)	Ln (Sa)

Tabla 5.3 Topología básica de las RNAs para la Ciudad de Medellín

Característica	Duración parte intensa	Aceleración máxima	Período predominante	Espectros de respuesta elásticos
Arquitectura	Cascada	Multicapa con alimentación hacia adelante	Cascada	Cascada
Variables externas	Seis	Seis	Cinco	Diez
Función de procesamiento	Máximos y mínimos	Normalizada	Máximos y mínimos	Normalizada
Número de capas ocultas	Una	Una	Una	Una
Neuronas en capa oculta	Diez	Tres	Cinco	Diez
Algoritmo de aprendizaje	Propagación rápida OP	Propagación rápida QP	Propagación rápida QP	Propagación rápida OP
Función de entrada	Suma de regresión	Producto punto	Sigma Pi	Producto punto
Función de Transferencia	Sigmoide	Sigmoide	Sigmoide	Sigmoide
Criterio de error	Medio cuadrático	Medio cuadrático	Entropía cruzada	Medio cuadrático
Capa de salida	Ln (Di)	Ln (Amax)	Tp	Ln (Sa)

Tabla 5.4 Topología básica de las RNAs para el caso colombiano

Característica	Aceleración máxima	Espectros de respuesta elásticos
Arquitectura	Cascada	Cascada
Variables externas	Seis	Diez
Función de procesamiento	Normalizada	Normalizada
Número de capas ocultas	Una	Una
Neuronas en capa oculta	Seis	Diez
Algoritmo de aprendizaje	Propagación rápida OP	Propagación rápida OP
Función de entrada	Producto punto	Producto punto
Función de Transferencia	Sigmoide	Sigmoide
Criterio de error	Medio cuadrático	Medio cuadrático
Capa de salida	Ln(Amax)	Ln (Sa)

Tabla 5.5 Variables externas para cada modelo neuronal

Modelo	Vector de entrada
Duración fase intensa Ciudad de México	$M_w, R, f_p, Q(f_p), Q(T_p), C_s$
Aceleración máxima Ciudad de México	M_w, h, R, D_i, I_A, C_s
Aceleración máxima, sitios fuera del Valle de México	M_w, h, R, D_i, I_A, C_s
Espectro de respuesta elástico Ciudad de México grupo 1, $T_p < 1.5$	$M_w, h, \phi_1, \delta_1, \lambda_1, \phi_2, \delta_2, \lambda_2, R, \alpha_i, f_p, f_c, V_{max}/A_{max}, Q(f_p), D_i, I_A, A_{max}, C_s$
Espectro de respuesta elástico Ciudad de México grupo 2 $T_p 1.5-3$	$M_w, h, \phi_1, \delta_1, \lambda_1, \phi_2, \delta_2, \lambda_2, R, \alpha_i, f_p, f_c, V_{max}/A_{max}, Q(f_p), D_i, I_A, A_{max}, C_s$
Espectro de respuesta elástico Ciudad de México grupo 3, $T_p > 3$	$M_w, h, \phi_1, \delta_1, \lambda_1, \phi_2, \delta_2, \lambda_2, R, \alpha_i, f_p, f_c, V_{max}/A_{max}, Q(f_p), D_i, I_A, A_{max}, C_s$
Duración fase intensa Ciudad de Medellín	$M_w, R, f_p, Q(f_p), Q(T_p), \beta$
Aceleración máxima Ciudad de Medellín	$M_w, h, R, D_i, I_A, \beta$
Período dominante Ciudad de Medellín	$M_w, h, R, \alpha_i, \beta$
Espectro de respuesta elástico Ciudad de Medellín foco superficial	$M_w, h, R, \alpha_i, f_p, Q(f_p), Q(T_p), D_i, I_A, A_{max}$
Espectro de respuesta elástico Ciudad de Medellín foco lejano	$M_w, h, R, \alpha_i, f_p, Q(f_p), Q(T_p), D_i, I_A, A_{max}$
Aceleración máxima República de Colombia	$M_w, h, R, D_i, I_A, \beta$
Espectro de respuesta elástico Colombia foco superficial	$M_w, h, R, \alpha_i, f_p, Q(f_p), Q(T_p), D_i, I_A, A_{max}$
Espectro de respuesta elástico Colombia foco lejano	$M_w, h, R, \alpha_i, f_p, Q(f_p), Q(T_p), D_i, I_A, A_{max}$

Tabla 5.6 Parámetros estadísticos

Modelo	Error en términos de la desviación estándar		Coeficiente de correlación	
	Aprendizaje	Prueba	Aprendizaje	Prueba
Duración fase intensa Ciudad de México	0.385	0.311	0.934	0.937
Aceleración máxima Ciudad de México	0.220	0.210	0.964	0.985
Aceleración máxima, sitios fuera del Valle de México	0.210	0.280	0.974	0.943
Espectro de respuesta elástico Ciudad de México grupo 1, $T_p < 1.5$	0.305	0.318	0.927	0.966
Espectro de respuesta elástico Ciudad de México grupo 2 $T_p 1.5-3$	0.277	0.349	0.956	0.943
Espectro de respuesta elástico Ciudad de México grupo 3, $T_p > 3$	0.223	0.337	0.984	0.883
Duración fase intensa Ciudad de Medellín	0.210	0.460	0.95.9	0.827
Aceleración máxima Ciudad de Medellín	0.270	0.340	0.972	0.906
Período dominante Ciudad de Medellín	0.220	0.270	0.902	0.845
Espectro de respuesta elástico Ciudad de Medellín foco superficial	0.292	0.376	0.873	0.909
Espectro de respuesta elástico Ciudad de Medellín foco lejano	0.328	0.505	0.958	0.829
Aceleración máxima República de Colombia	0.279	0.254	0.949	0.978
Espectro de respuesta elástico Colombia foco superficial	0.431	0.487	0.828	0.872
Espectro de respuesta elástico Colombia foco lejano	0.376	0.561	0.802	0.820

5.3 Resultados

5.3.1 Caso México

Los resultados se presentan en forma gráfica e incluyen:

- Comparación entre los valores de duración de la fase intensa calculados con base en la definición de duración propuesta Trifunac [75], que corresponde a la porción de la señal que se desarrolla entre el 5 % y 95% de la intensidad de Arias en un acelerograma; y los estimados con RNAs para la ciudad de México, ver figura 5.2. El ajuste logrado en este caso en la etapa de validación es bastante aceptable, se alcanzó un coeficiente de correlación de 0.937 para la base de datos seleccionada que incluye sitios localizados en las tres zonas geotécnica de la ciudad. El error promedio de 0.311 oscila alrededor de los valores reportados en la literatura. En [Bommer 00] la desviación estándar para la regresión $\log(D_E) = 0.69M_w - 3.7$ es de 0.28 usando 32 registrados instalados en roca, en esta ecuación D_E es la duración efectiva [Bommer 99]. En [Reinoso 01], ver ecuación (2.34) capítulo 2, el error promedio para todos los registros utilizados fue de 0.24 considerando la duración de la parte intensa del movimiento como el tiempo entre 2.5 y 97.5 % de la intensidad de Arias.
- Valores de aceleración máxima observados y los estimados con RNAs tanto para la cCudad de México como para sitios fuera del Valle de México, figuras 5.3 y 5.4 respectivamente. En ambos casos es apreciable una disminución de la dispersión (ver también tabla 5.6), salvo para niveles bajos de aceleración máxima en los que RNA sobreestima los valores observados, pero las diferencias son admisibles.
- Valores estimados de aceleración horizontal máxima. Con base en la modelación neuronal especificada en la tabla 5.2 y considerando las variables externas indicadas en la tabla 5.5,

se divisó que el parámetro externo denominado clasificación del tipo de suelo C_s , resulta estadísticamente significativo (ver análisis de sensibilidad capítulo 6). Este indicio permite construir curvas de atenuación para las diferentes zonas geotécnicas de la ciudad. En las figuras 5.5, 5.6, 5.7 se muestra respectivamente, la variación de los valores estimados de aceleración horizontal máxima componente NS, en función de la distancia para diferentes magnitudes M_w , en cada zona; y se contrastan con algunos sismos no incluidos en la modelación durante la fase de aprendizaje. Las gráficas de la figura 5.8 representan las ecuaciones de atenuación propuestas por Sing [87] y Ordaz [89] para su comparación con los resultados obtenidos en este estudio. Con base en la inferencia anotada, se realizó la representación espacial de la duración de la fase intensa y aceleración máxima considerando en la simulación la magnitud, la localización y la profundidad del sismo del 19 de septiembre de 1985. Los resultados se presentan en los mapas de contornos de la figura 5.9. Para estos casos la intensidad de Arias se estimó con las correlaciones que se presentan en la tabla 5.7, y la duración de la parte intensa del movimiento requerida como dato de entrada para estimar la aceleración máxima con RNAs, se calculó con el sistema neuronal entrenado para la duración de la fase intensa.

Tabla 5.7 Correlación entre las variables M_w e IA para la Ciudad de México

Zona	Ajuste	Error estándar	Coefficiente de correlación
I	$IA = 3.0527141 \cdot 10^{-10} \exp(2.9010966 \cdot M_w)$	0.3111	0.950
II	$IA = 17.69594 - 6.5229314 \cdot M_w + 0.60075464 \cdot (M_w)^2$	0.4274	0.927
III	$IA = 7.9581147 \cdot 10^{-11} \exp(3.4046626 \cdot M_w)$	2.2485	0.984

- Espectros de respuesta elásticos asociados al registro original y los estimados con la RNA a partir del vector externo conformado por las siguientes variables M_w , h , ϕ_1 , δ_1 , λ_1 , ϕ_2 , δ_2 , λ_2 , R , α_i , f_p , f_c , V_{max}/A_{max} , $Q(f_p)$, D_i , IA , A_{max} , C_s ; y con topologías estipuladas en la tabla 5.2 para la etapa de predicción, en cada una de las categorías consideradas de acuerdo con la escisión establecida en función del período dominante. Dada la cantidad de gráficas, se presentan por separado los resultados del conjunto de patrones utilizados en la validación de la modelación de cada grupo en las figuras 5.10, 5.11 y 5.12.

En este caso es de resaltar que las características de la RNA considerada como la óptima son las mismas para los tres grupos, lo que puede interpretarse como un procedimiento estable. En general, los resultados obtenidos en la fase de ejecución, para 30 espectros del grupo 1, 35 del grupo 2 y 16 del grupo 3 con parámetros estadísticos reportados en la tabla 5.6, permiten concluir que el modelo neuronal reproduce los patrones de validación con un ajuste aceptable dentro de las aplicaciones prácticas del campo profesional, aunque en varios eventos la red neuronal no plasma exactamente la forma espectral ni sus máximos como era de esperarse dada la complejidad del fenómeno y la naturaleza de la modelación basada en aprendizaje, lo que conlleva a resultados fuertemente dependientes de los datos que se utilicen en el entrenamiento y la forma como le sean presentados a la red. Estos aspectos se pueden mejorar considerando un modelo híbrido que acople las RNAs con la lógica borrosa y los algoritmos genéticos.

Reconociendo que existe una acumulación del error de predicción sobre las variables aceleración máxima e intensidad de Arias en la estimación de espectros de respuesta, se diseñó un sistema neuronal en serie. El diagrama de flujo se presenta en la figura 5.13 y en las figuras 5.14 y 5.15 los resultados de los bloques aceleración máxima e intensidad de Arias respectivamente.

El sistema en serie se aplicó en la estimación neuronal de espectros de respuesta del grupo 2 de Ciudad de México, dichos resultados se presentan la figura 5.11 denotados con azul. En esta misma figura se incluyeron los resultados con redes neuronales sin considerar los siguientes parámetros f_c , V_{max}/A_{max} , D_i , IA y A_{max} en el vector de entrada; los resultados corresponden al trazo en amarillo.

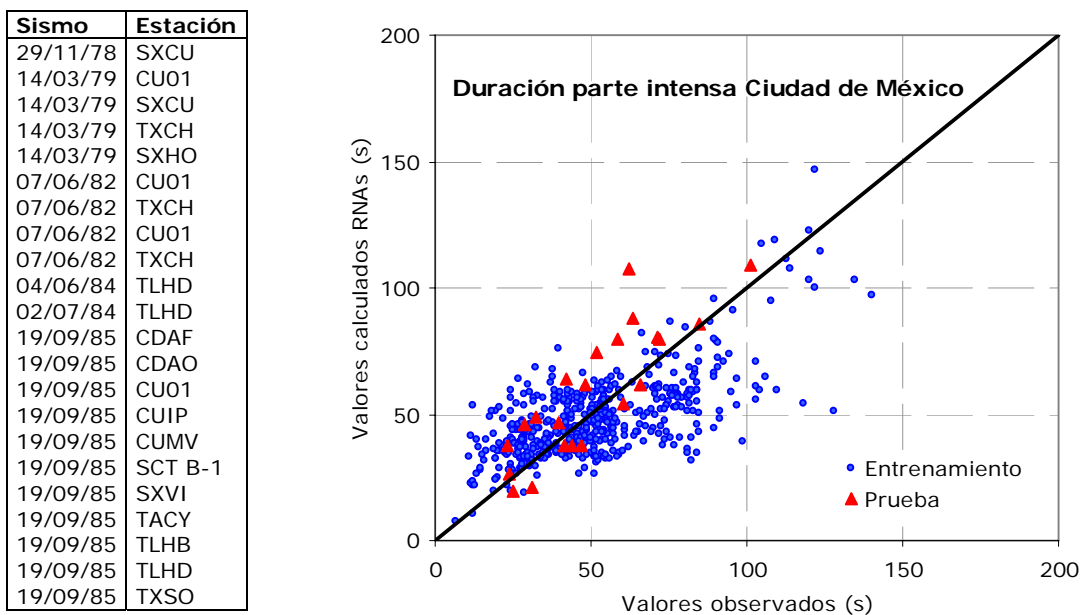


Figura 5.2 Comparación entre los valores de duración de fase intensa observados y los estimados con RNAs. A la izquierda datos utilizados en etapa de prueba.

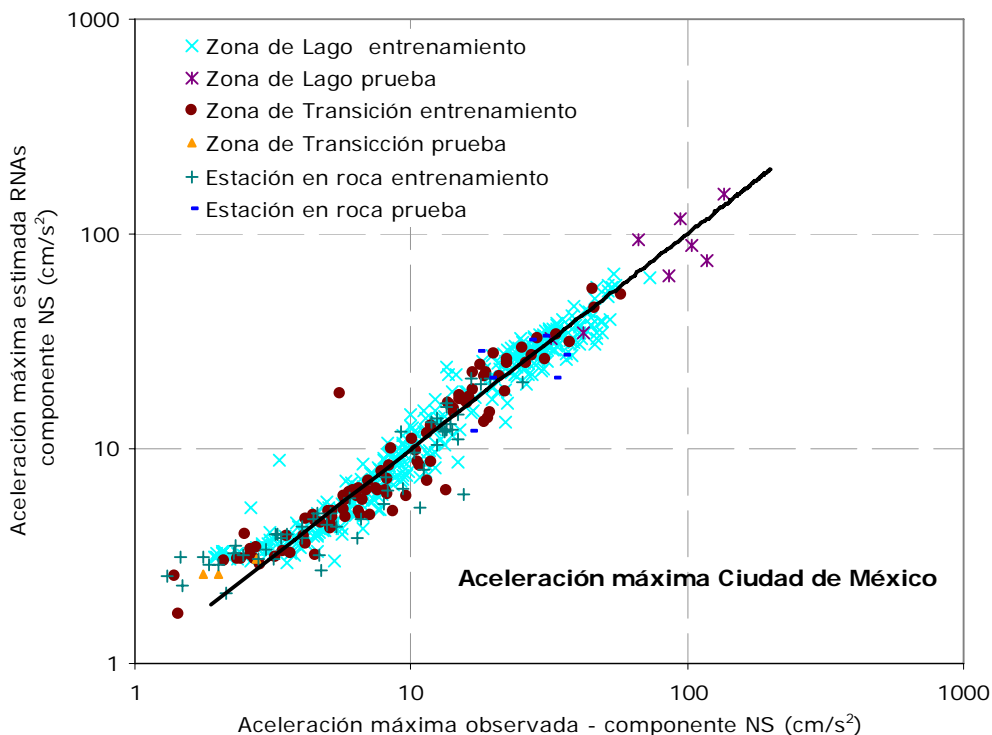


Figura 5.3 Comparación entre los valores de aceleración máxima observados y los estimados con RNAs. Datos utilizados en la etapa de prueba igual a los de duración parte intensa (tabla superior izquierda).

Tabla 5.8 Eventos sísmico y estaciones de registro *fuera del Valle de México* utilizado en la neuro computación en la etapa de entrenamiento para la aceleración máxima componente NS.

Sismo	Estación	Sismo	Estación	Sismo	Estación
14/03/79	ACAP	25/04/89	ACAZ	24/10/93	IGUA
14/03/79	ACAS	25/04/89	ATYC	24/10/93	MEZC
24/10/80	MINA	25/04/89	CAYA	24/10/93	MSAS
24/10/80	OAXM	25/04/89	COMD	24/10/93	OCLL
24/10/80	PAJA	25/04/89	COPL	24/10/93	OCTT
25/10/81	CHL1	25/04/89	COYC	24/10/93	POZU
25/10/81	SXPV	25/04/89	CSER	24/10/93	PTQL
07/06/82	CHL1	25/04/89	FIC2	24/10/93	SMR2
07/06/82	OAXM	25/04/89	LLAV	24/10/93	VIGA
07/06/82	SMAR	25/04/89	MAGY	24/10/93	VNTA
21/09/85	ACAP	25/04/89	MSAS	23/05/94	ACAD
21/09/85	ACAS	25/04/89	OCLL	23/05/94	ACAJ
21/09/85	ATYC	25/04/89	OCTT	23/05/94	ACAZ
21/09/85	AZIH	25/04/89	PARS	23/05/94	AZIH
21/09/85	CAYA	25/04/89	SMR2	23/05/94	BALC
21/09/85	CHI1	25/04/89	TEAC	23/05/94	CAYA
21/09/85	COYC	25/04/89	VIGA	23/05/94	CHIL
21/09/85	FICA	25/04/89	VNTA	23/05/94	COMD
21/09/85	PAPN	25/04/89	XALT	23/05/94	COYC
21/09/85	PARS	02/05/89	ACAC	23/05/94	CSER
21/09/85	SUCH	02/05/89	ACAD	23/05/94	CUER
21/09/85	SXPU	02/05/89	ACAR	23/05/94	FIC2
21/09/85	UNIO	02/05/89	ACAZ	23/05/94	IGUA
21/09/85	VNTA	02/05/89	COPL	23/05/94	MSAS
21/09/85	XALT	02/05/89	COYC	23/05/94	NUXC
21/09/85	ZACA	02/05/89	CPDR	23/05/94	OCLL
30/04/86	ARTG	02/05/89	FIC2	23/05/94	OCTT
30/04/86	CALE	02/05/89	MSAS	23/05/94	PAPN
30/04/86	COLI	02/05/89	OCLL	23/05/94	PETA
30/04/86	GUAC	02/05/89	OCTT	23/05/94	SUCH
30/04/86	ZACA	02/05/89	PARS	23/05/94	ZIIG
05/05/86	CALE	02/05/89	SMR2	14/09/95	ACAC
05/05/86	GUAC	02/05/89	TEAC	14/09/95	ACAD
05/05/86	ZACA	02/05/89	VIGA	14/09/95	ACAJ
07/06/87	ACAD	02/05/89	VNTA	14/09/95	ACAN
07/06/87	ACAP	02/05/89	XALT	14/09/95	ACAR
07/06/87	COYC	15/05/93	ACAC	14/09/95	ACAZ
07/06/87	FIC2	15/05/93	ACAD	14/09/95	ATYC
07/06/87	MSAS	15/05/93	ACAJ	14/09/95	CAIG
07/06/87	OCLL	15/05/93	ACAR	14/09/95	CHIL
07/06/87	SMR2	15/05/93	ACAZ	14/09/95	COPL
07/06/87	TEAC	15/05/93	CHIL	14/09/95	COYC
07/06/87	TNLP	15/05/93	COYC	14/09/95	CUER
07/06/87	VIGA	15/05/93	CSER	14/09/95	IGUA
07/06/87	XALT	15/05/93	CUER	14/09/95	OCLL
08/02/88	ATYC	15/05/93	ESTA	14/09/95	POZU
08/02/88	AZIH	15/05/93	MEZC	14/09/95	RICA
08/02/88	BALC	15/05/93	MSAS	14/09/95	TNLP
08/02/88	CAYA	15/05/93	OCTT	14/09/95	VIGA
08/02/88	COMD	15/05/93	SMR2	30/09/99	ACAC
08/02/88	COYC	15/05/93	TNLP	30/09/99	ACAD
08/02/88	LLAV	15/05/93	VIGA	30/09/99	ACAJ
08/02/88	MAGY	24/10/93	CACD	30/09/99	ACAR
08/02/88	NARJ	24/10/93	ACAJ	30/09/99	CHIL
08/02/88	PAPN	24/10/93	ACAN	30/09/99	COPL
08/02/88	PARS	24/10/93	ACAR	30/09/99	COYC
08/02/88	SUCH	24/10/93	ACAZ	30/09/99	COYQ
08/02/88	TEAC	24/10/93	CAYA	30/09/99	CSER
08/02/88	TNLP	24/10/93	CHIL	30/09/99	HUIG
08/02/88	UNIO	24/10/93	COPL	30/09/99	IGUA
25/04/89	ACAC	24/10/93	CPDR	30/09/99	PLIG
25/04/89	ACAD	24/10/93	CSER	30/09/99	VIGA
25/04/89	ACAP	24/10/93	CUER	30/09/99	ZIIG
25/04/89	ACAR	24/10/93	ESTA		
25/04/89	ACAS	24/10/93	FIC2		

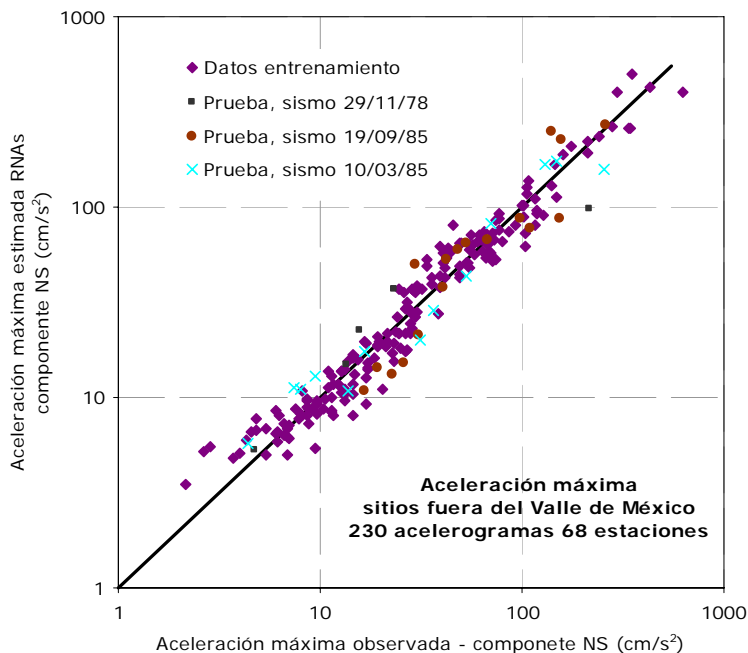


Figura 5.4 Comparación entre los valores de aceleración máxima observados y los estimados con RNAs. Datos utilizados en la modelación fase entrenamiento en tabla 5.8.

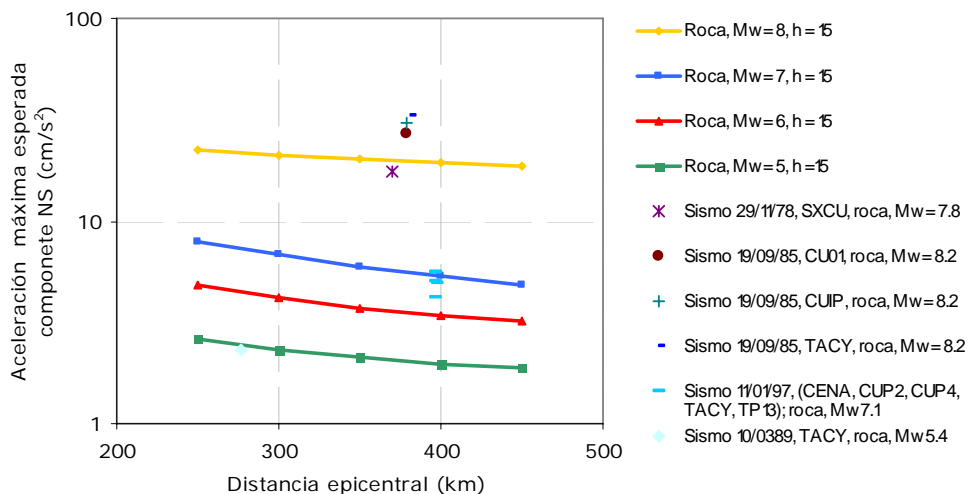


Figura 5.5 Valores estimados de aceleración del suelo con RNAs en función de R y Mw. Zonas I de la Ciudad de México.

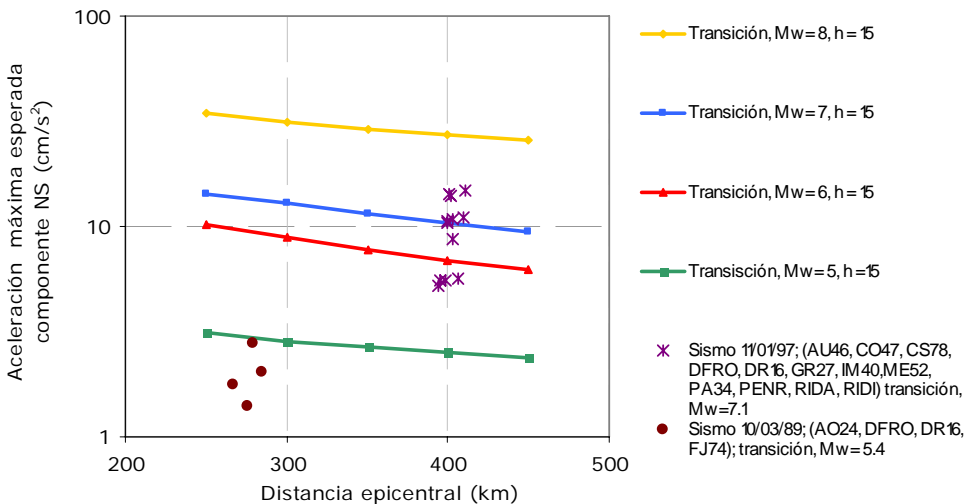


Figura 5.6 Valores estimados de aceleración del suelo con RNAs en función de R y Mw. Zonas II de la Ciudad de México.

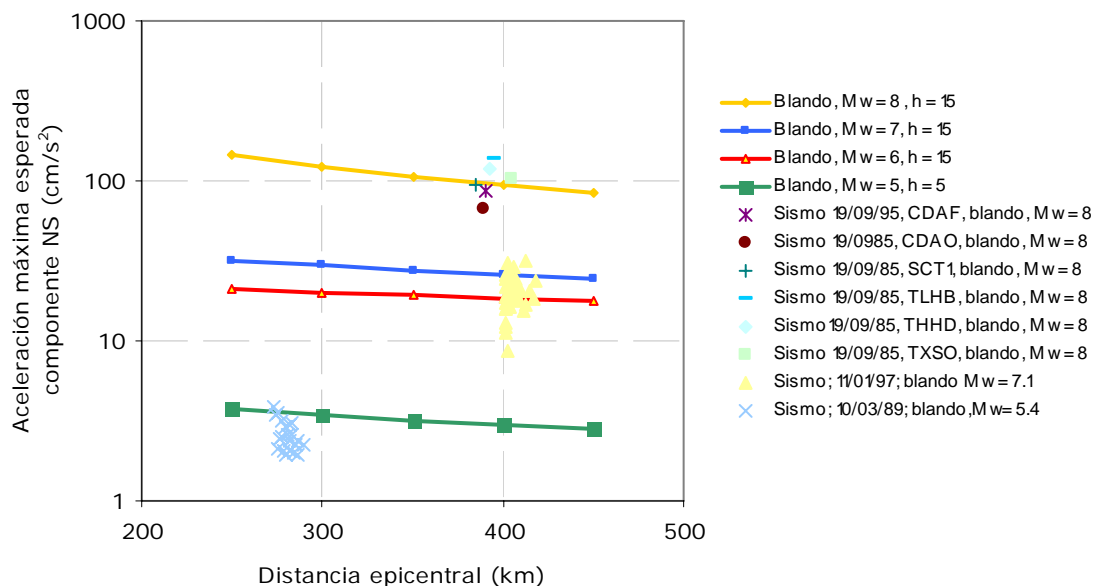


Figura 5.7 Valores estimados de aceleración del suelo con RNAs en función de R y Mw. Zona III de la Ciudad de México.

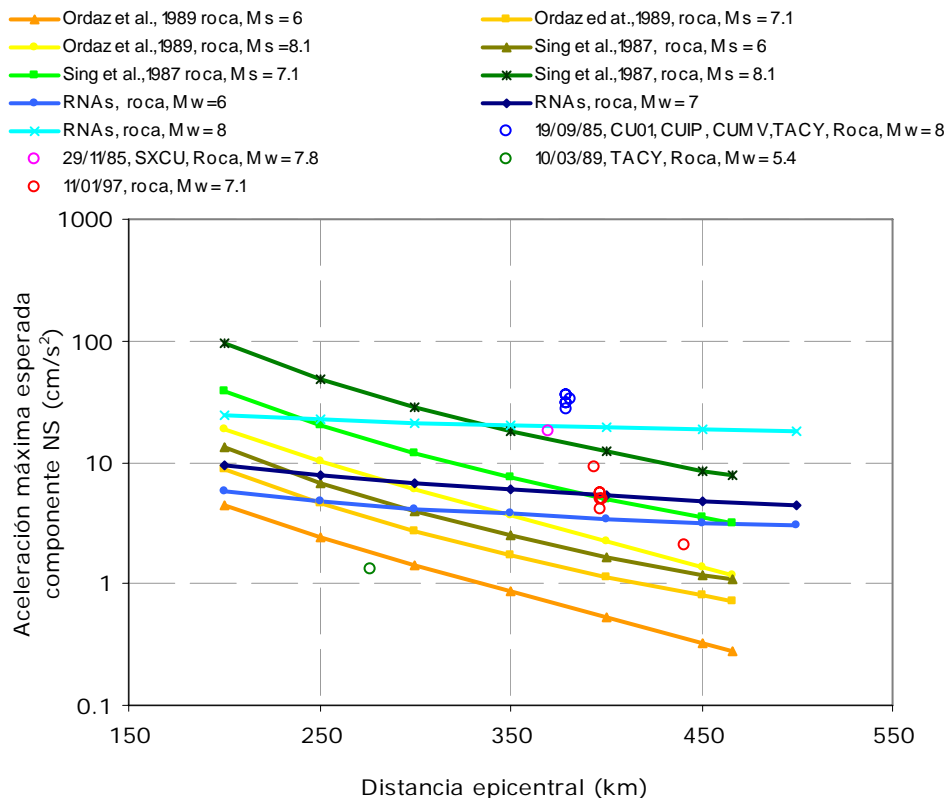


Figura 5.8 Valores estimados de aceleración horizontal máxima con RNAs en función de la distancia para diferentes magnitudes Mw, Zonas I de la Ciudad de México. Comparados para diferentes sismos y con las curvas obtenidas en otros estudios.

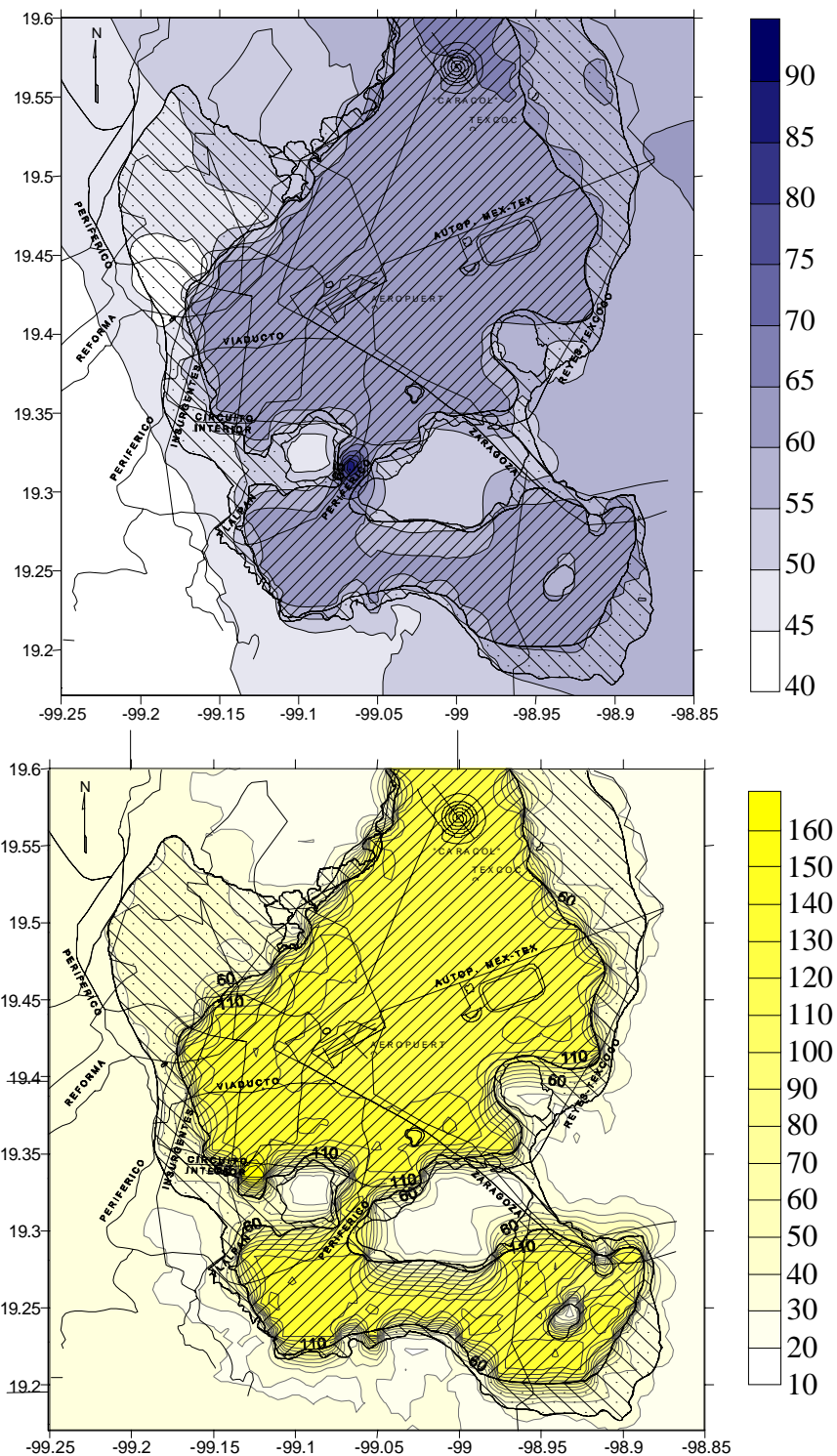


Figura 5.9 Parte superior de la figura, contornos de igual duración de la fase intensa del movimiento, en segundos. Parte inferior contornos de isoaceleración, en cm/s^2 . Simulación para el sismo del 19 de septiembre de 1985.

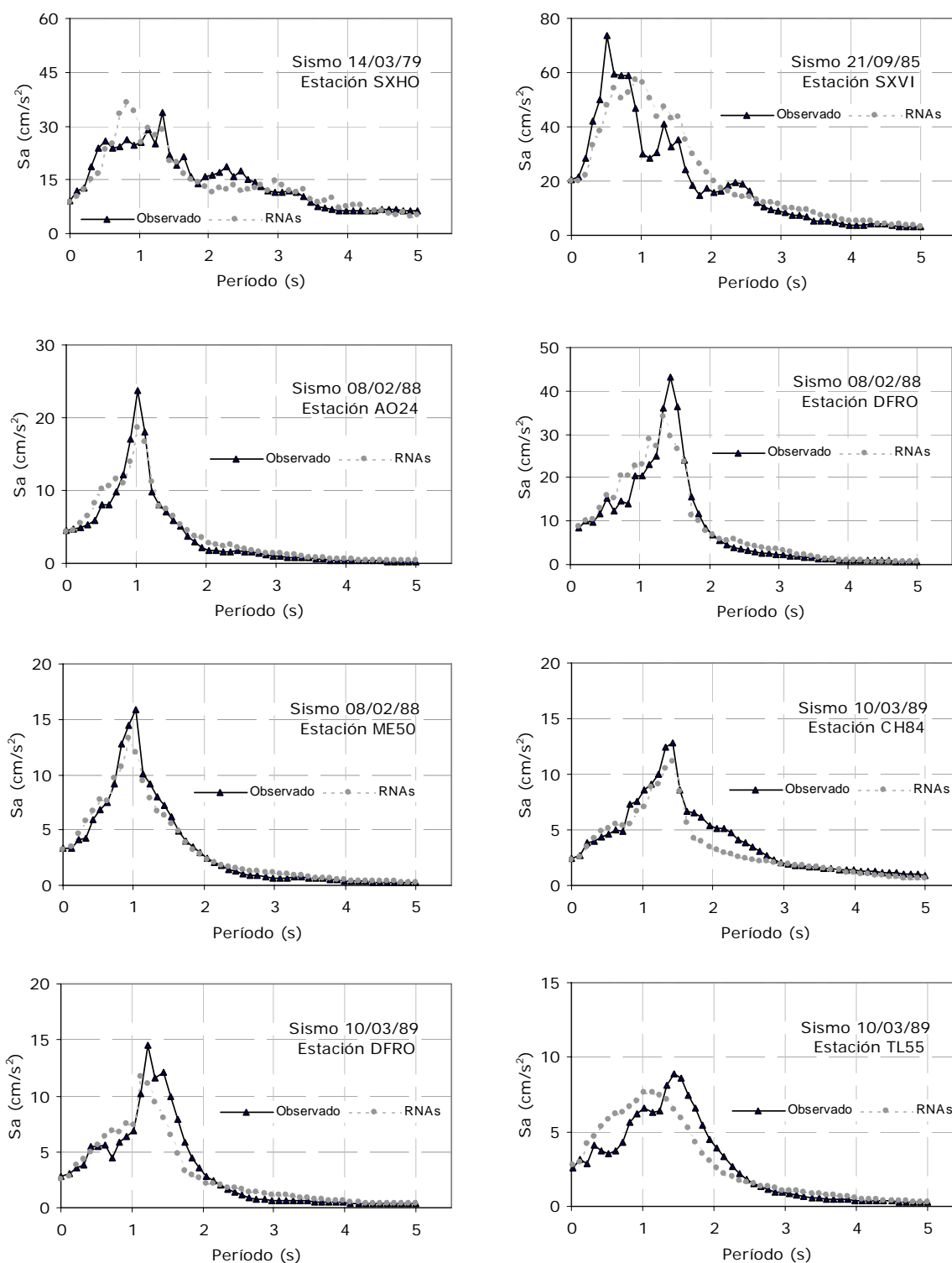
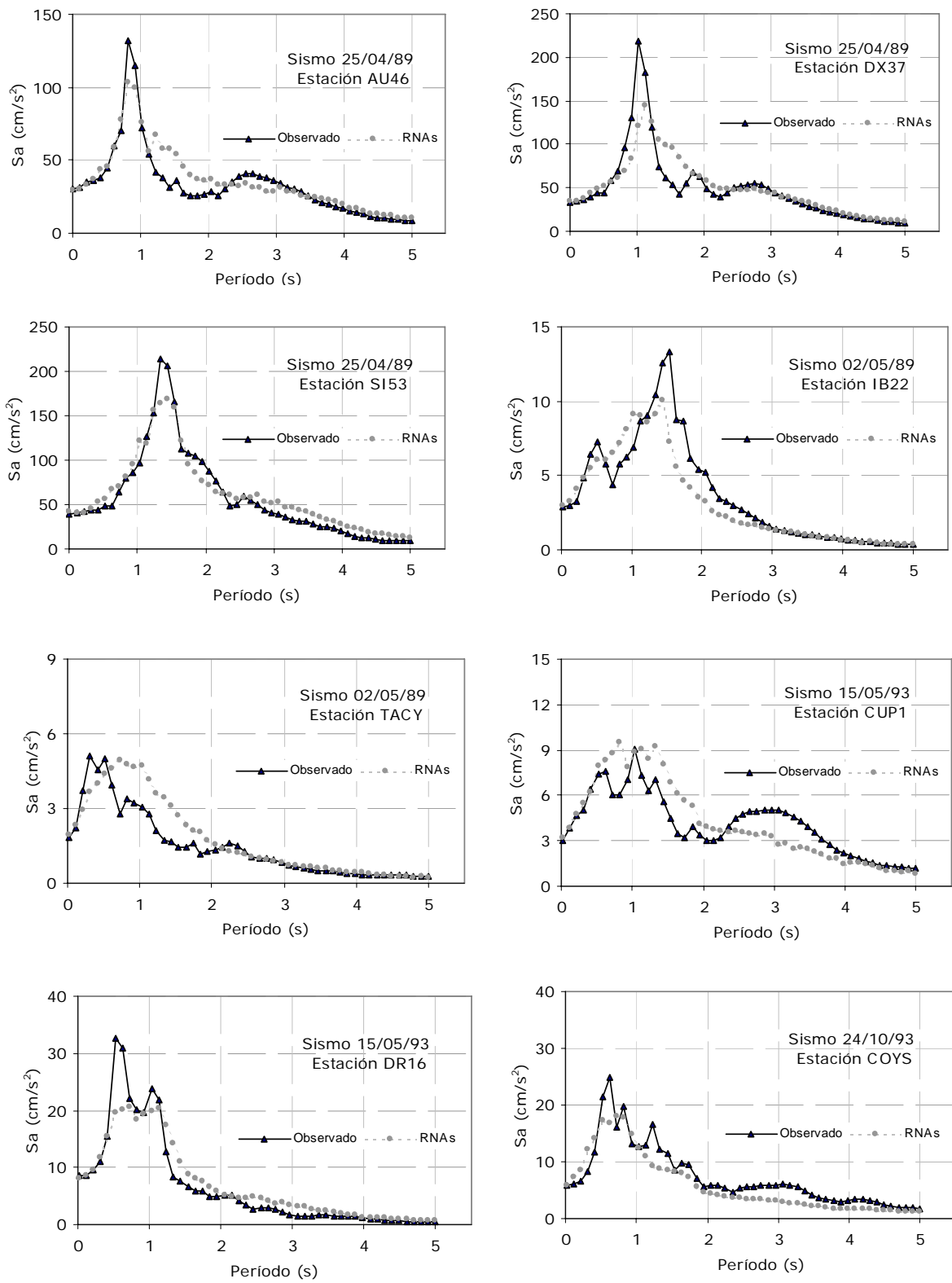
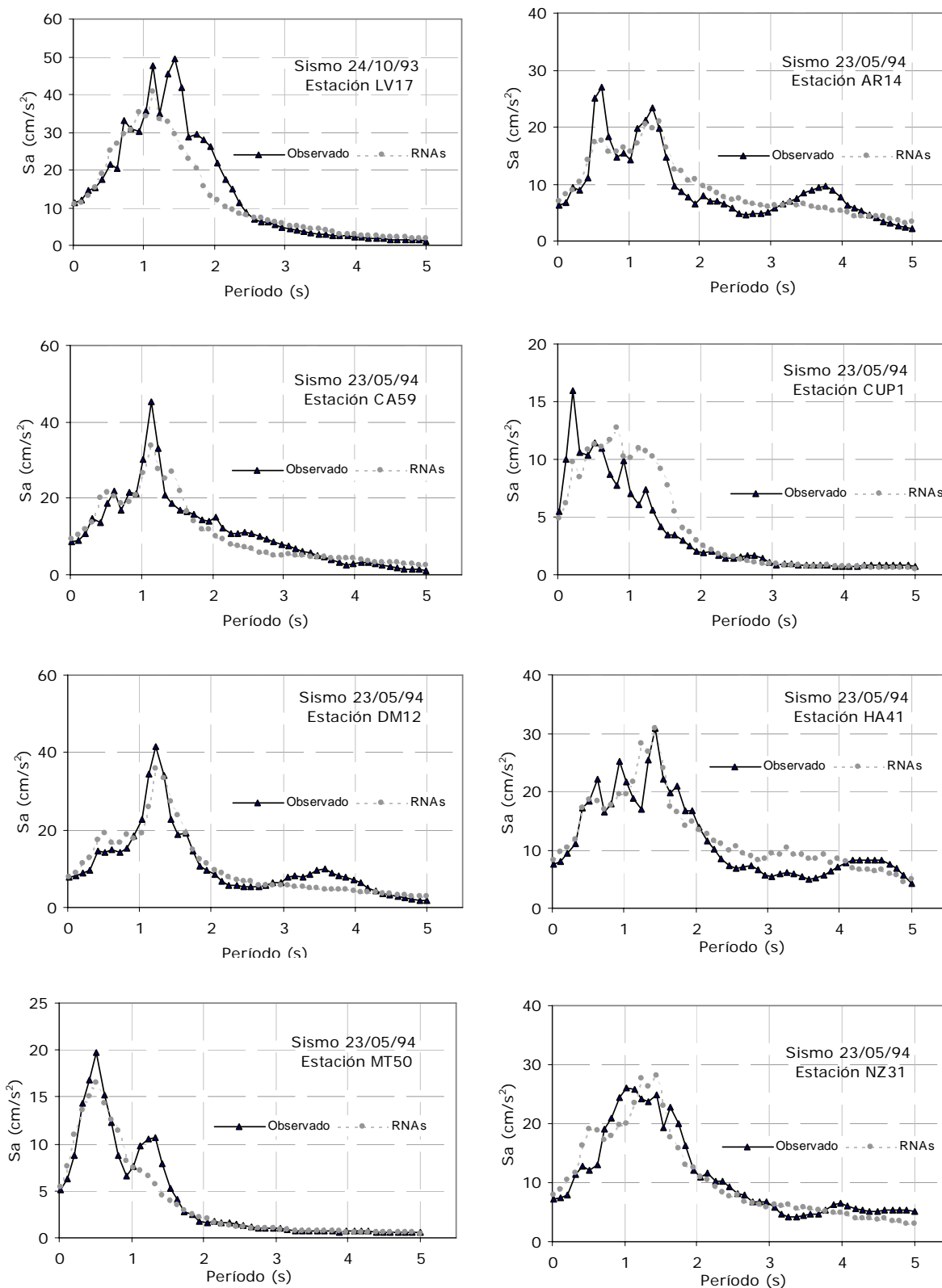


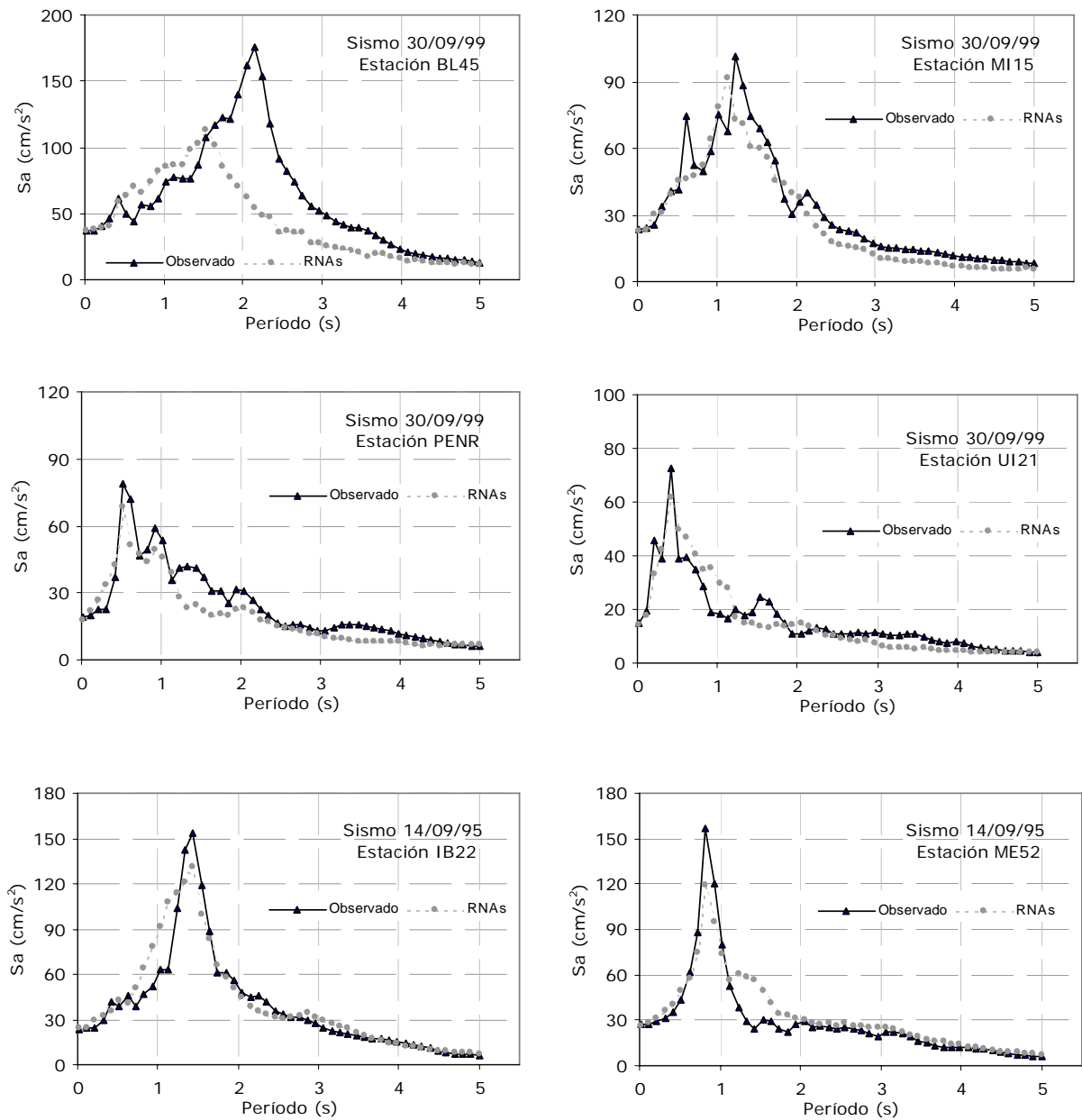
Figura 5.10 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para el grupo 1 establecido para la Ciudad de México. Estos corresponden a los casos en que la característica fundamental de vibración de cada uno de los perfiles de suelo o período predominante de este, sea menor igual a 1.5 segundos ($T_p \leq 1.5$ s).



Continúa figura 5.10 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para el grupo 1 establecido para la Ciudad de México.



Continúa figura 5.10 Espectros de respuesta elásticos(5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para el grupo 1 establecido para la Ciudad de México.



Continúa figura 5.10 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para el grupo 1 establecido para la Ciudad de México.

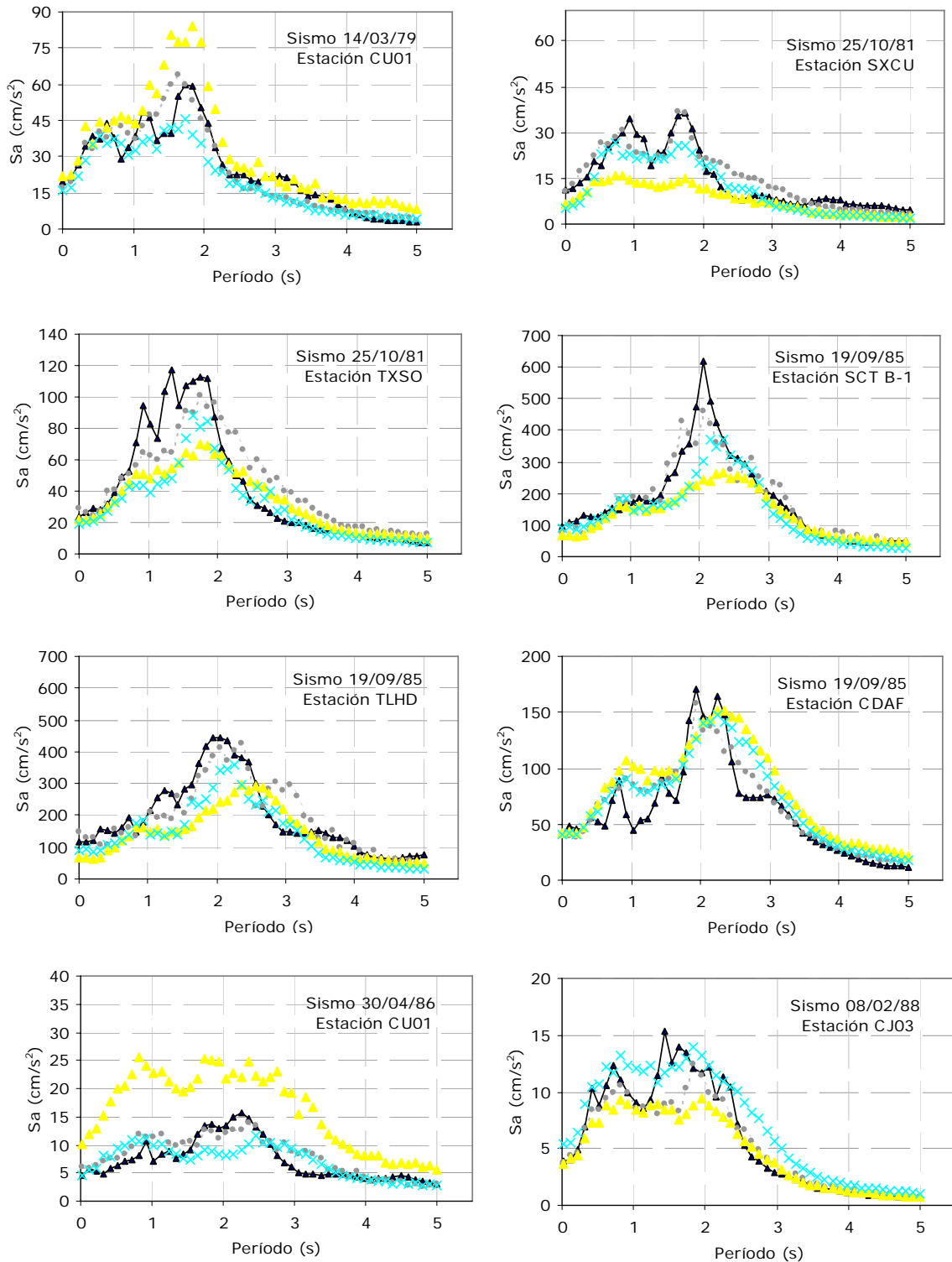
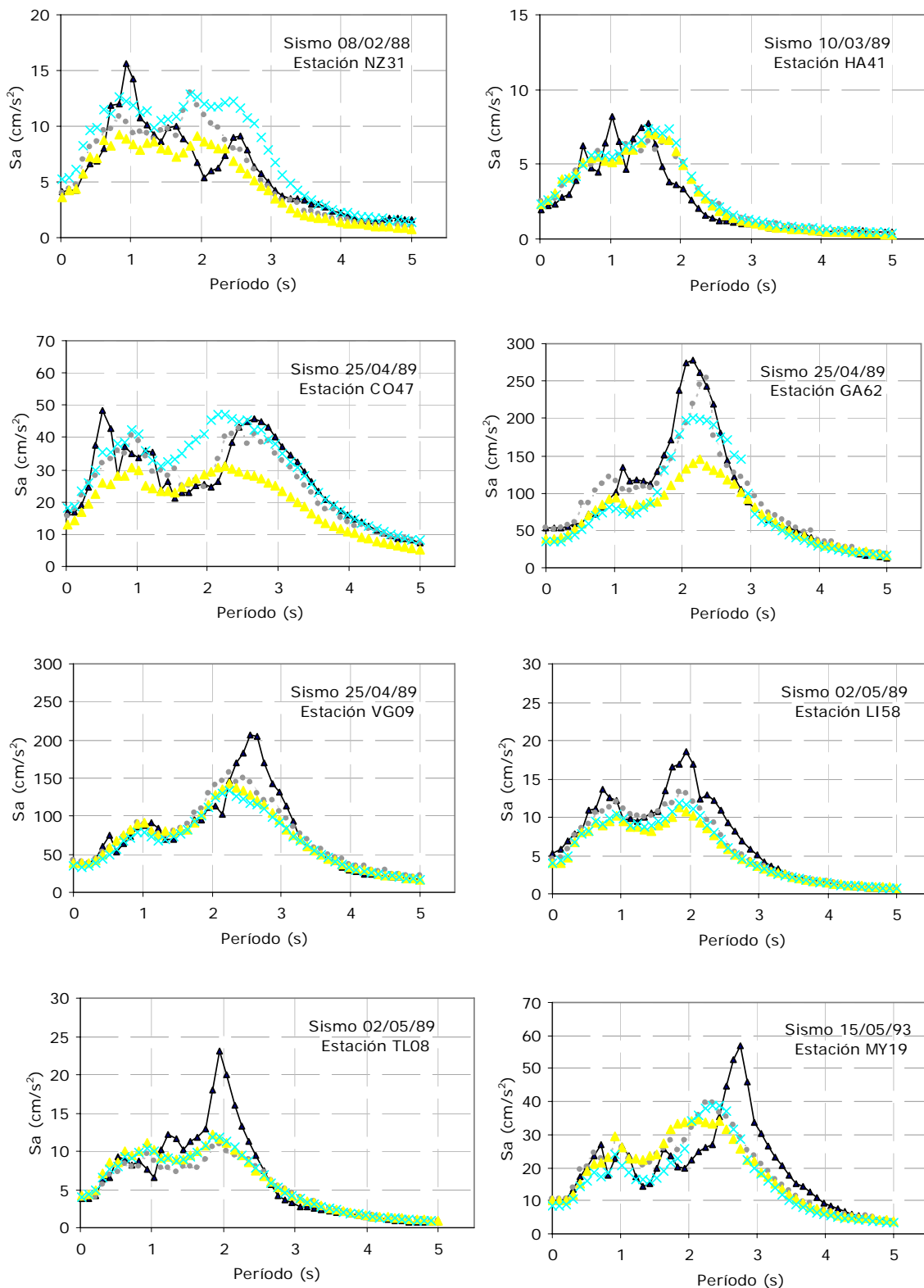
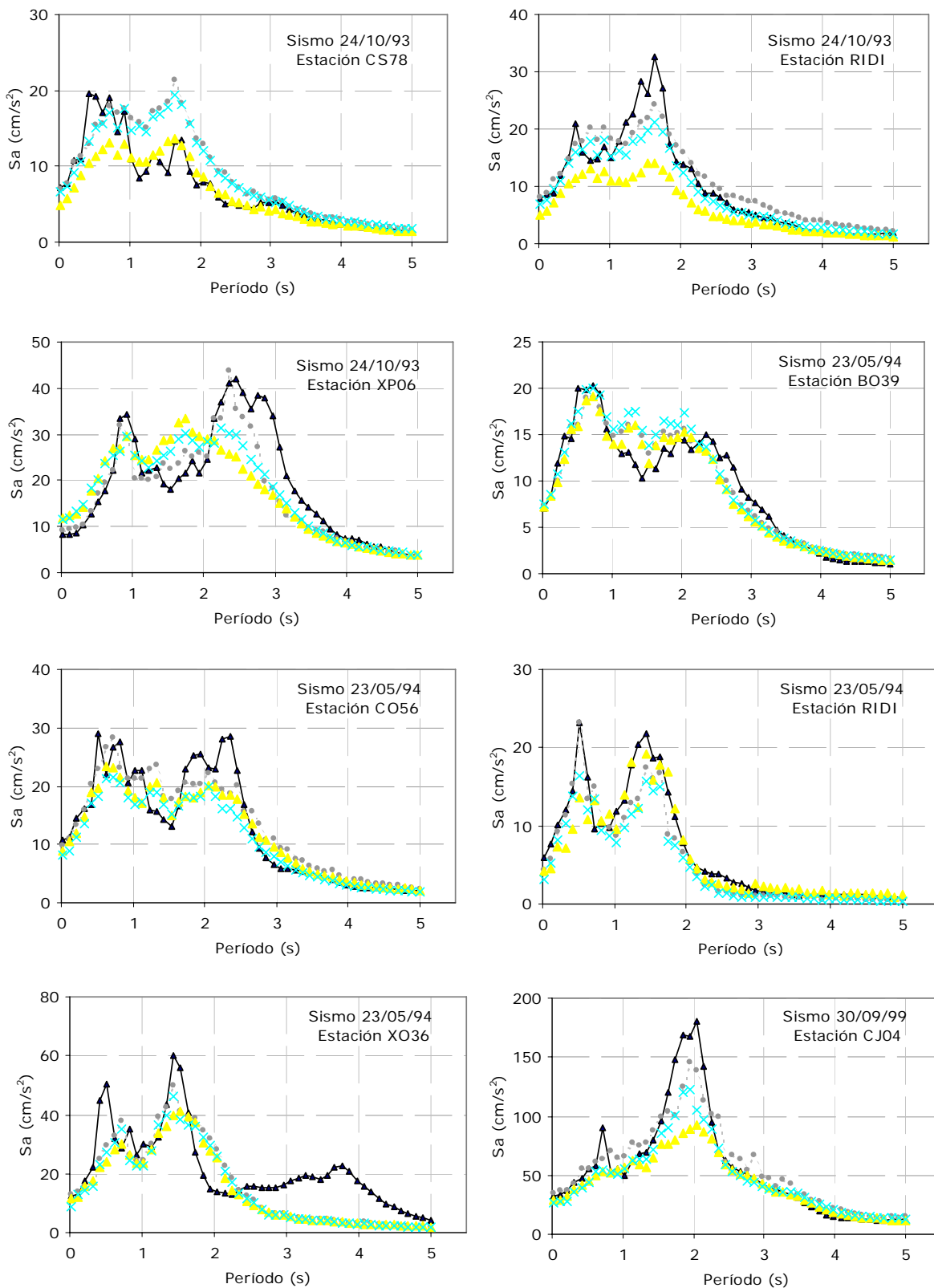


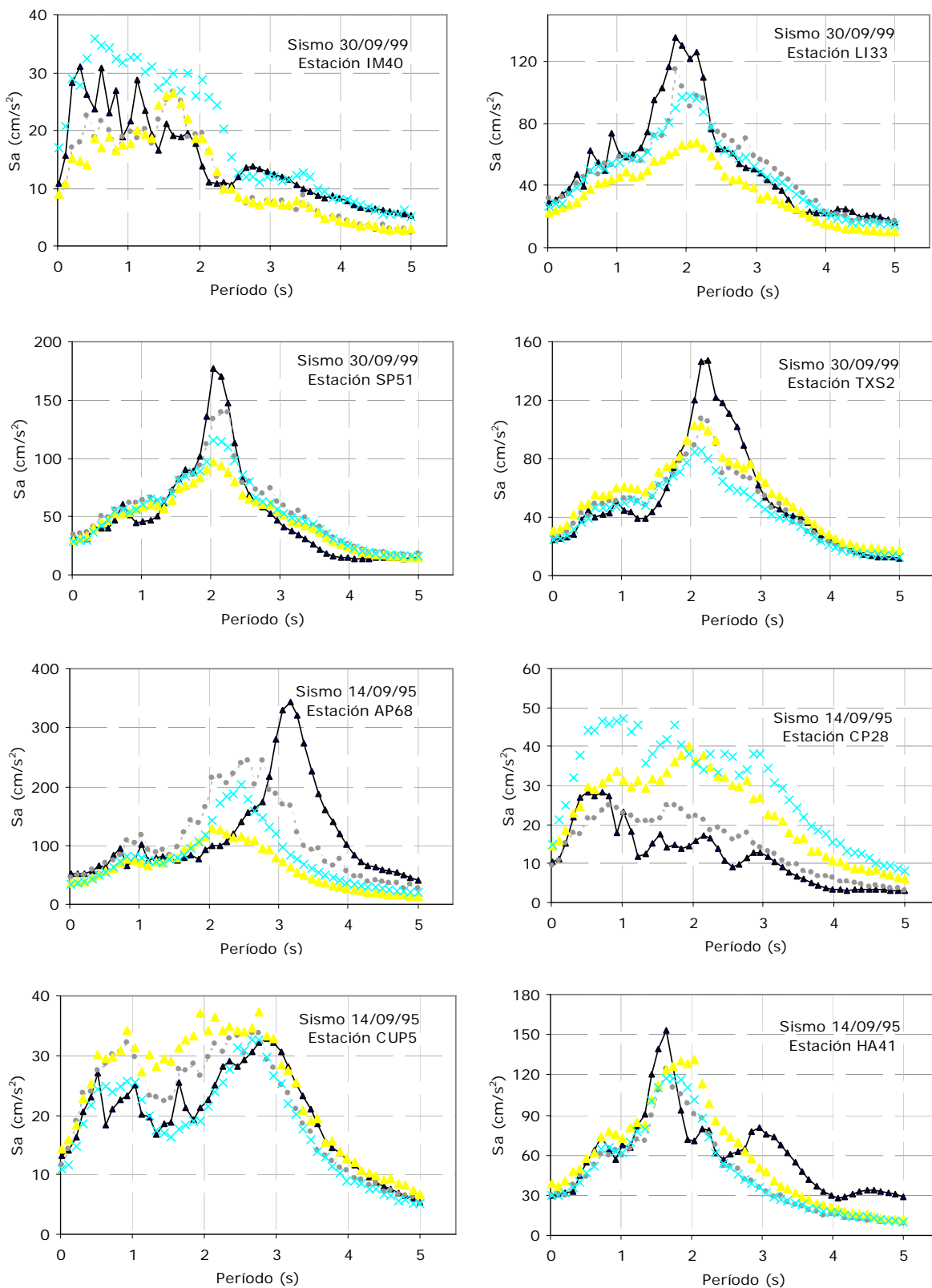
Figura 5.11 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para el grupo 2 del caso Ciudad de México. Corresponden a los casos en que la característica fundamental de vibración de cada uno de los perfiles de suelo o periodo dominante de este, se ubiquen en el intervalo 1.5 a 3 segundos ($1.5 \text{ s} < T_p < 3 \text{ s}$). Espectros observados (negro) y estimados (en gris RNA con 19 variables, en azul sistema neuronal en serie y en amarillo RNA con 14 variables).



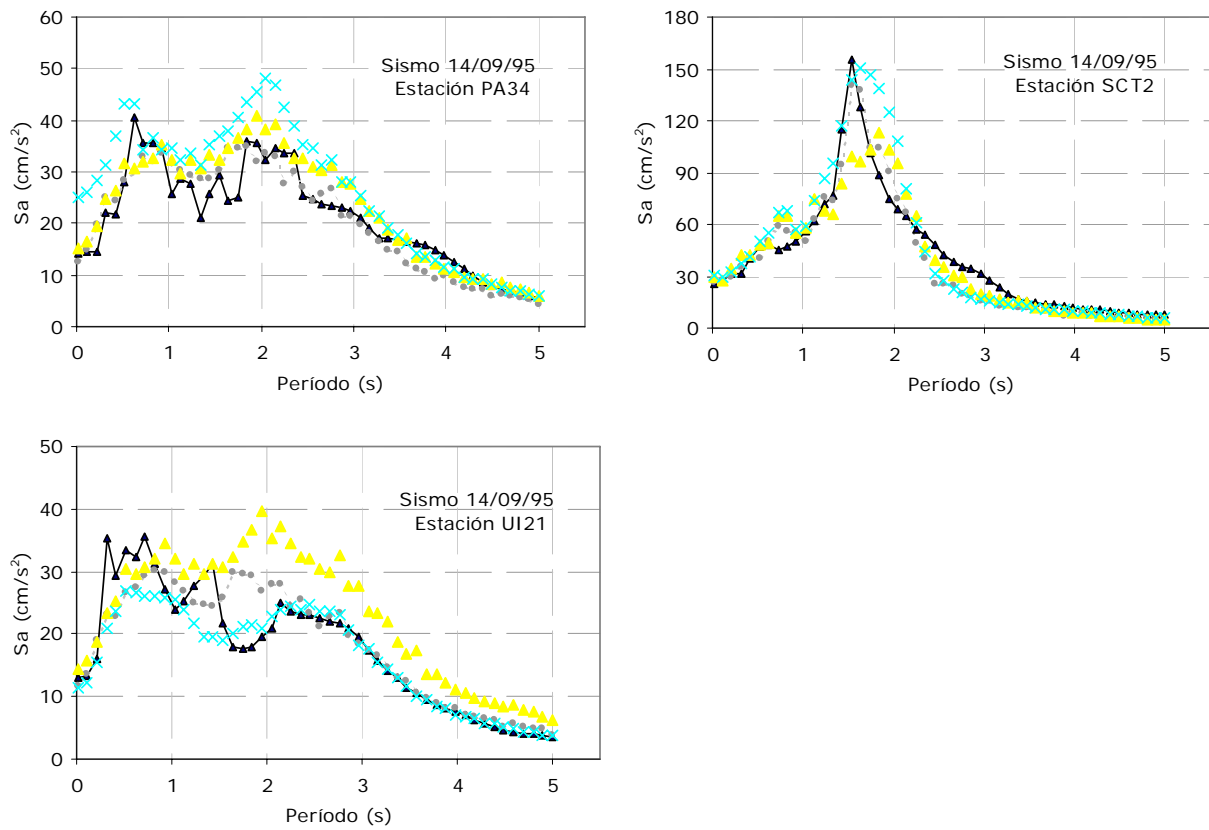
Continúa figura 5.11 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para el grupo 2 establecido para la Ciudad de México. Espectros observados (negro) y estimados (en gris RNA con 19 variables, en azul sistema neuronal en serie y en amarillo RNA con 14 variables).



Continúa figura 5.11 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para el grupo 2 establecido para la Ciudad de México. Espectros observados (negro) y estimados (en gris RNA con 19 variables, en azul sistema neuronal en serie y en amarillo RNA con 14 variables).



Continúa figura 5.11 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para el grupo 2 establecido para la Ciudad de México. Espectros observados (negro) y estimados (en gris RNA con 19 variables, en azul sistema neuronal en serie y en amarillo RNA con 14 variables).



Continúa figura 5.11 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para el grupo 2 establecido para la Ciudad de México. Espectros observados (negro) y estimados (en gris RNA con 19 variables, en azul sistema neuronal en serie y en amarillo RNA con 14 variables).

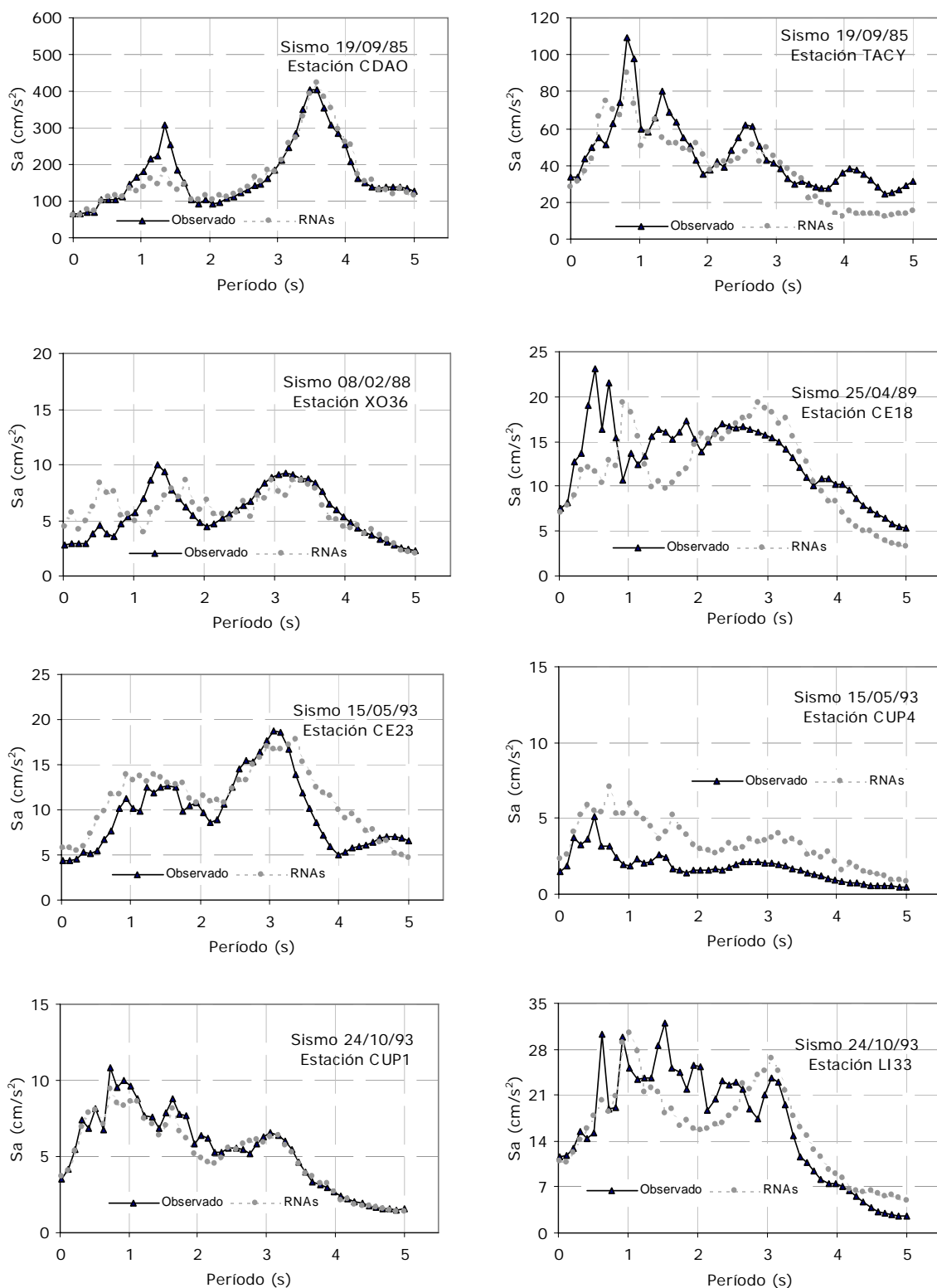
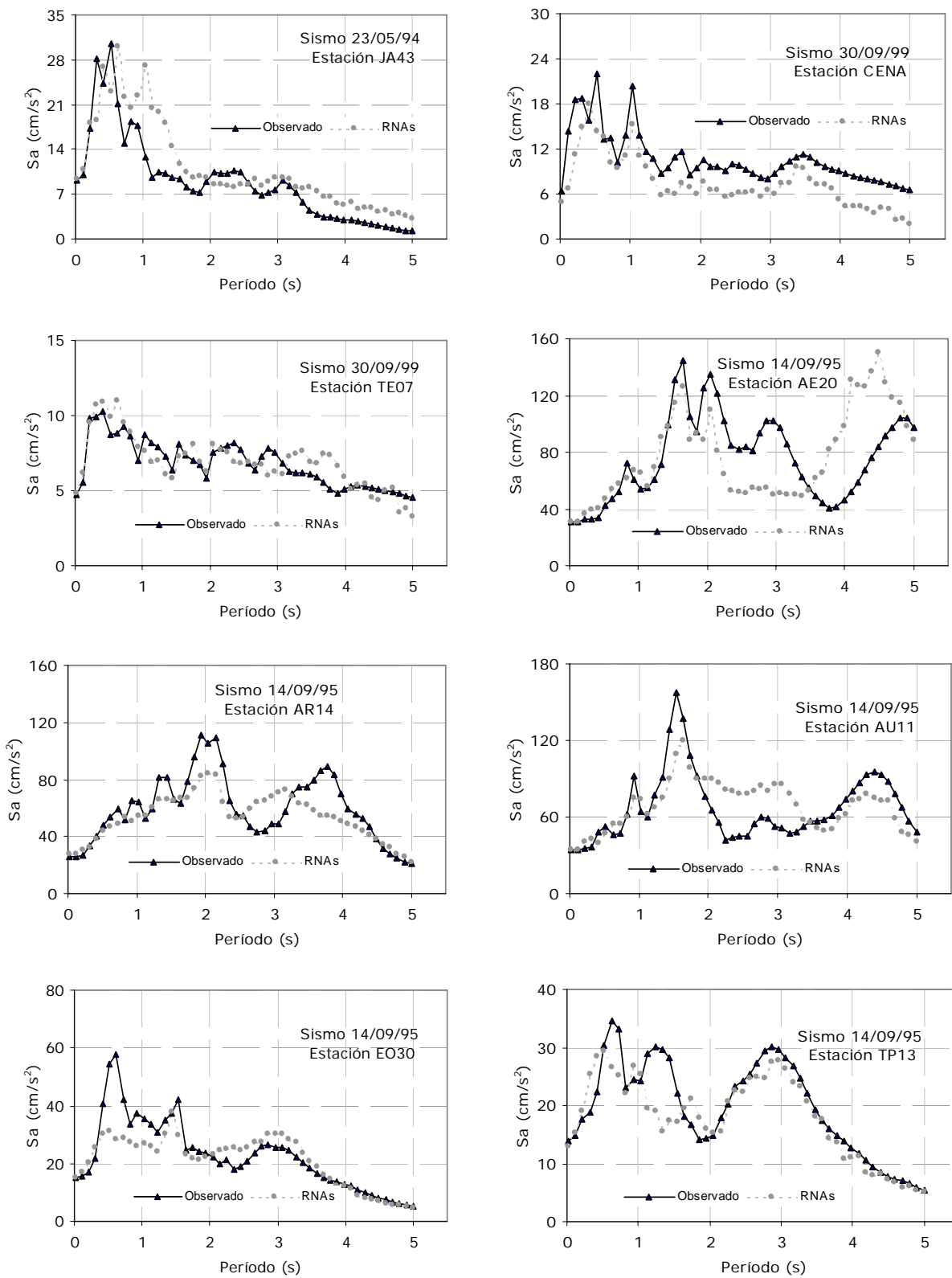


Figura 5.12 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para el grupo 3 del caso Ciudad de México. Corresponden a los casos en que la característica fundamental de vibración de cada uno de los perfiles de suelo o el periodo predominante de éste, sea igual o mayor a 3 segundos ($T_p \geq 3$ s).



Continúa figura 5.12 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de predicción para el grupo 3 establecido para la Ciudad de México.

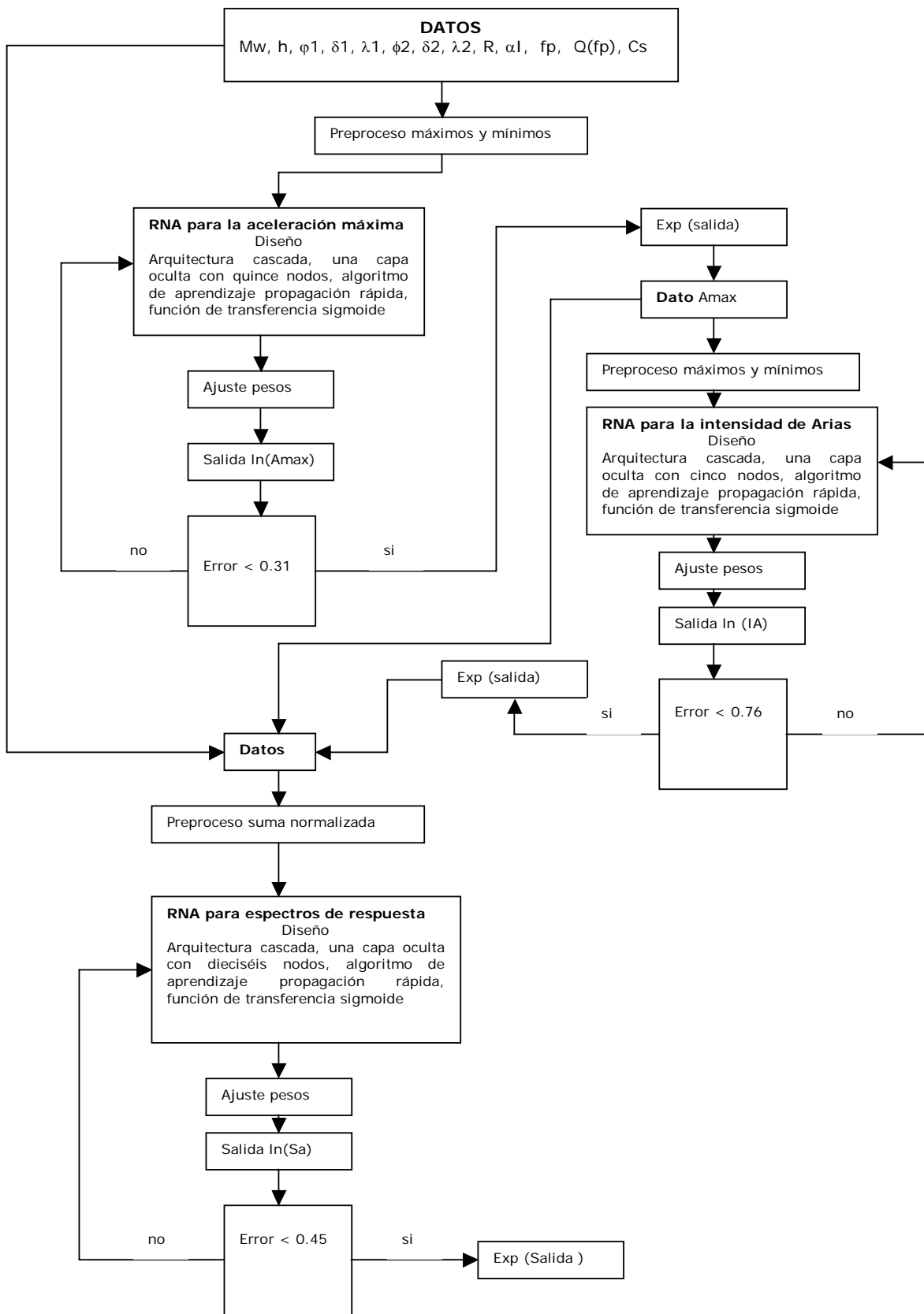


Figura 5.13 Diagrama de flujo del sistema neuronal desarrollado en serie.

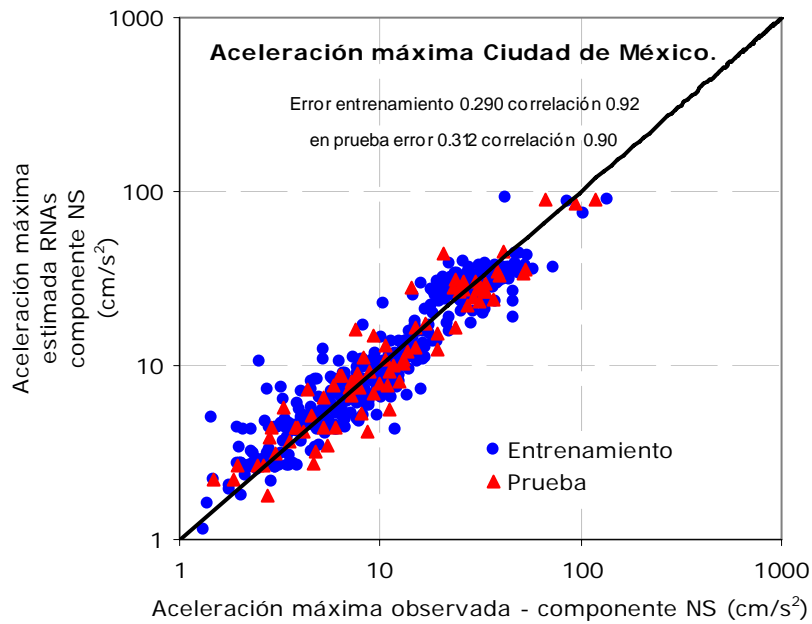


Figura 5.14 Bloque neuronal para la aceleración máxima del sistema en serie.

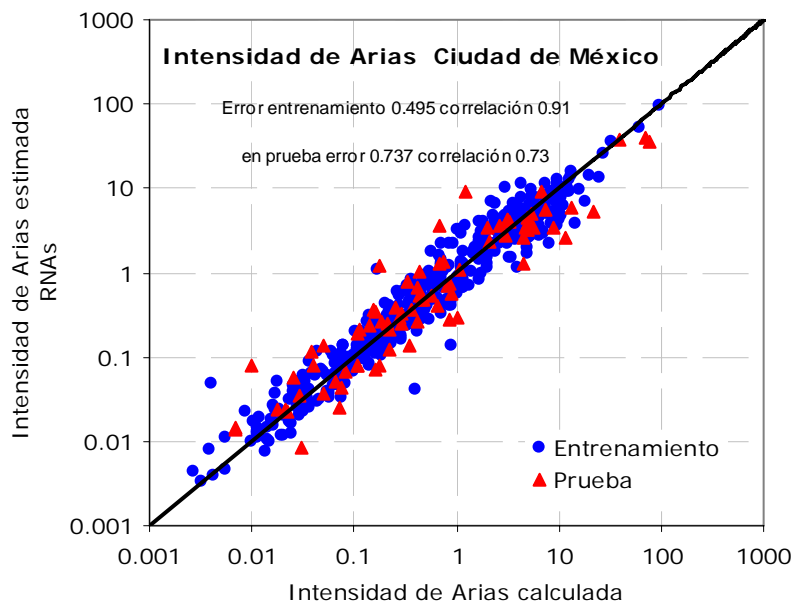


Figura 5.15 Bloque neuronal para la intensidad de Arias del sistema en serie.

5.3.2 Caso Ciudad de Medellín

Se ilustra gráficamente la comparación entre los valores observados y los estimados con las RNAs para la duración de la parte intensa de movimiento, aceleración máxima horizontal componente NS y el período dominante en las figuras 5.16, 5.17 y 5.18; así como los espectros de respuesta elásticos calculados a partir de los acelerogramas registrados y los espectros estimados con RNAs en la etapa de predicción para las categorías establecidas, que en este caso fueron foco superficial y foco lejano en las figuras 5.19 y 5.20 respectivamente.

De acuerdo con los parámetros estadísticos indicados en la tabla 5.6; la estimación conseguida en la modelación de la duración, la aceleración y el período dominante fue un poco más pobre pero aceptable. Esto explicable considerando la compleja distribución espacial de los suelos de la ciudad y el amplio rango de variación de las propiedades geotécnicas, cuya representación estuvo sólo soportada en el parámetro denominada velocidad de onda de cortante del subsuelo, β , donde está localizada la estación de registro.

Los resultados obtenidos a partir del modelo neuronal puntualizado en la tabla 5.3 para la atenuación espectral de la ciudad de Medellín, muestran que esta técnica puede reproducir las amplitudes espectrales máximas y sus períodos correspondientes con bastante aproximación respecto a los valores observados. No obstante, las correlaciones logradas en el conjunto de prueba 90.6% para foco superficial y 84.5% para foco lejano, son inferiores a las alcanzadas en el caso de la Ciudad de México, pese a que las formas espectrales de los eventos sísmicos disponibles muestra poca dispersión en términos de amplitudes espectrales y contenidos de frecuencia. Cabe recordar aquí, que por ausencia de información sobre los mecanismos focales de los eventos sísmicos considerados, los parámetros que definen la geometría de la fuente sísmica no fueron incluidos como variables en el vector de entrada de la red neuronal, quizás esto sea un reflejo de dicha omisión. También es importante subrayar que para la ciudad de Medellín dada su geomorfología y estratigrafía, es cuestionable la aplicabilidad de modelos unidimensionales en el análisis de la amplificación local, debido a que las condiciones reales están lejos de las hipótesis establecidas para estos modelos en los que se debe cumplir simultáneamente que, la influencia de los diferentes tipos de ondas den lugar a un predominio de ondas de corte con propagación vertical, los estratos que conforman el subsuelo debe ser aproximadamente horizontal con dimensiones grandes comparadas con la profundidad y relativamente homogéneos.

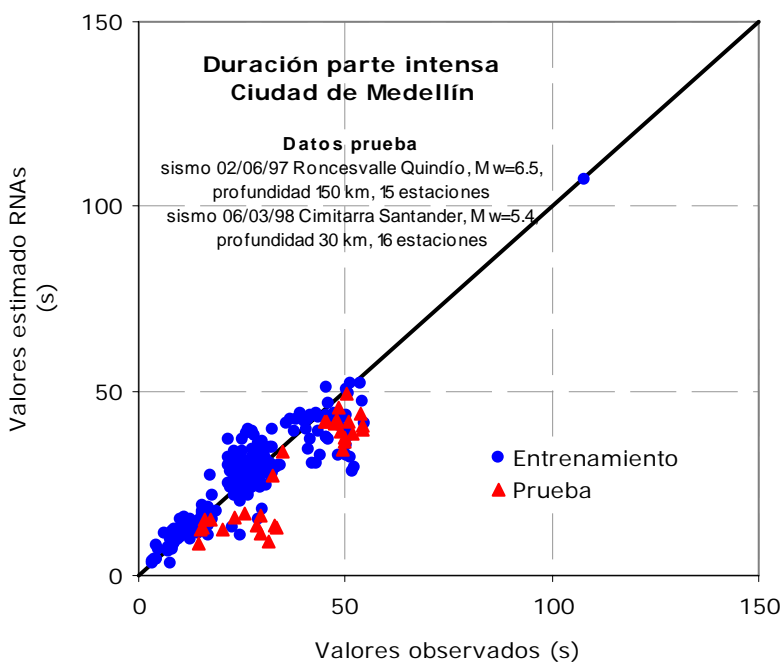


Figura 5.16 Comparación entre los valores de duración de fase intensa observados y los estimados con RNAs.

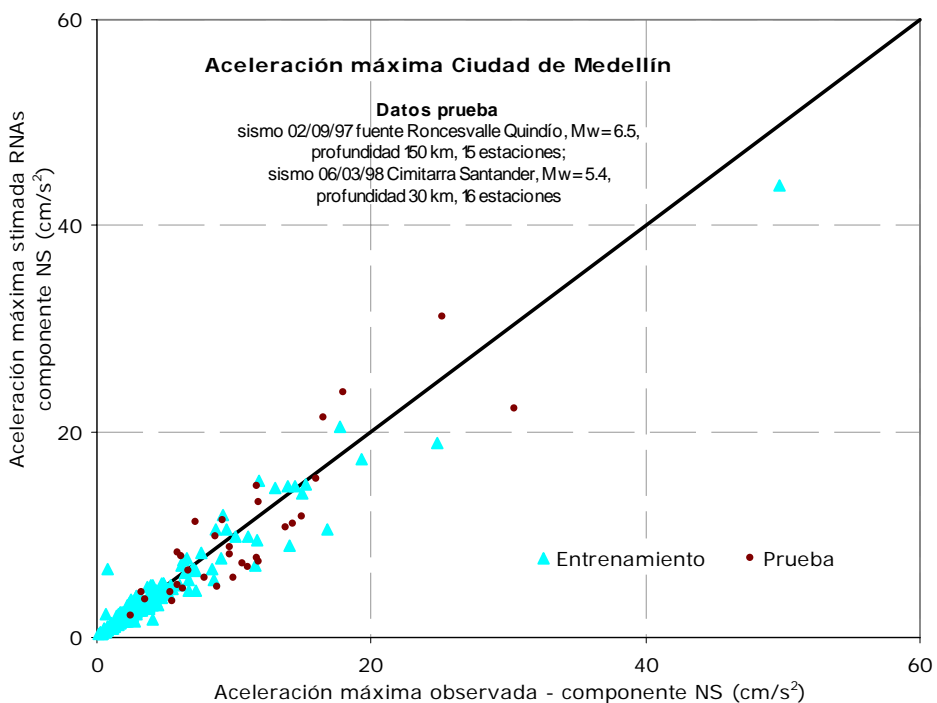


Figura 5.17 Comparación entre los valores de aceleración máxima observados y los estimados con RNAs.

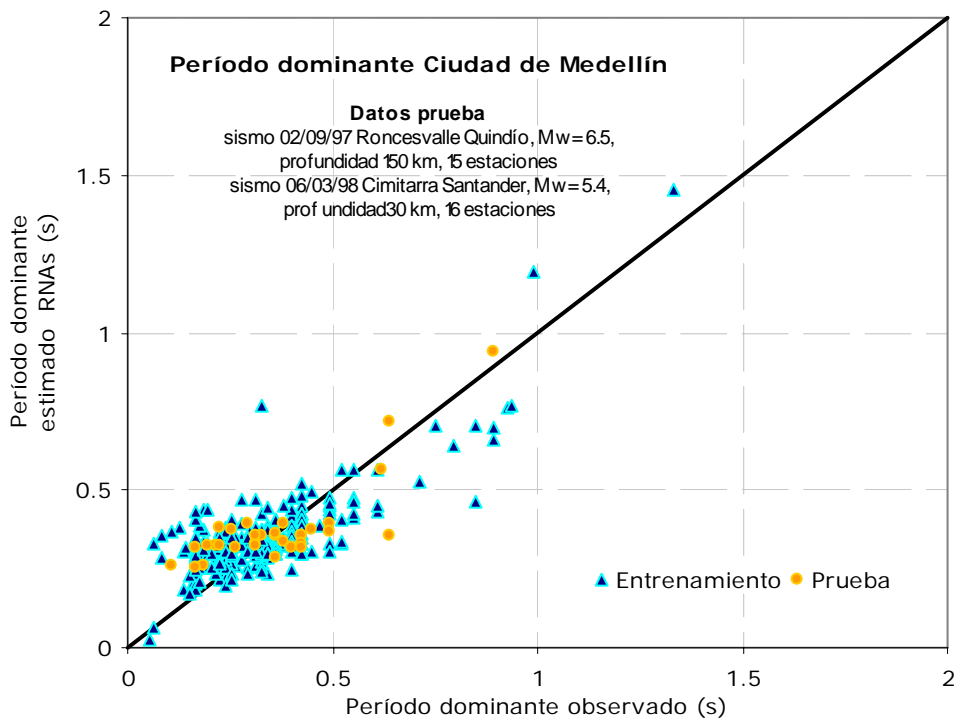


Figura 5.18 Comparación entre los valores del período predominante observados y los estimados con RNAs.

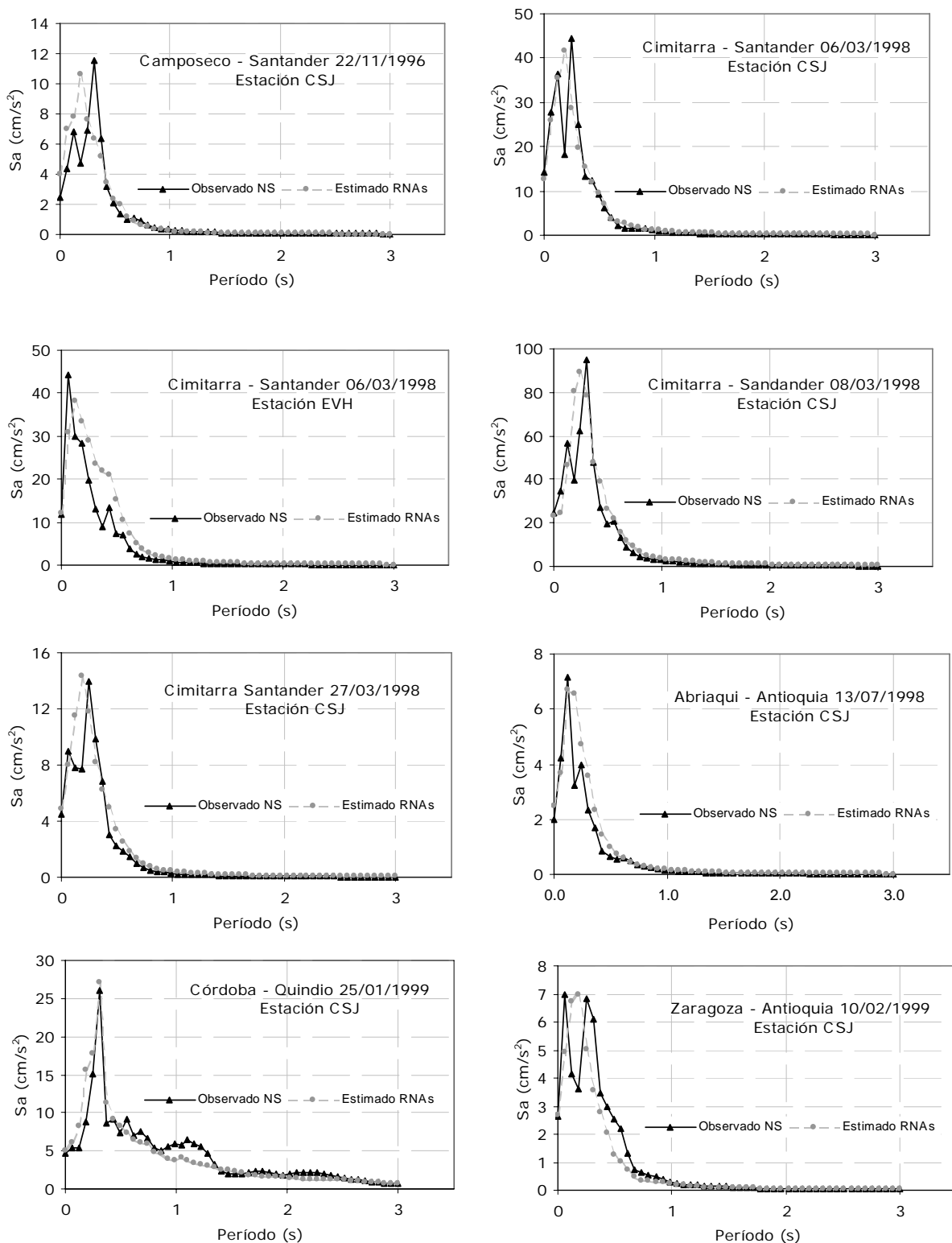
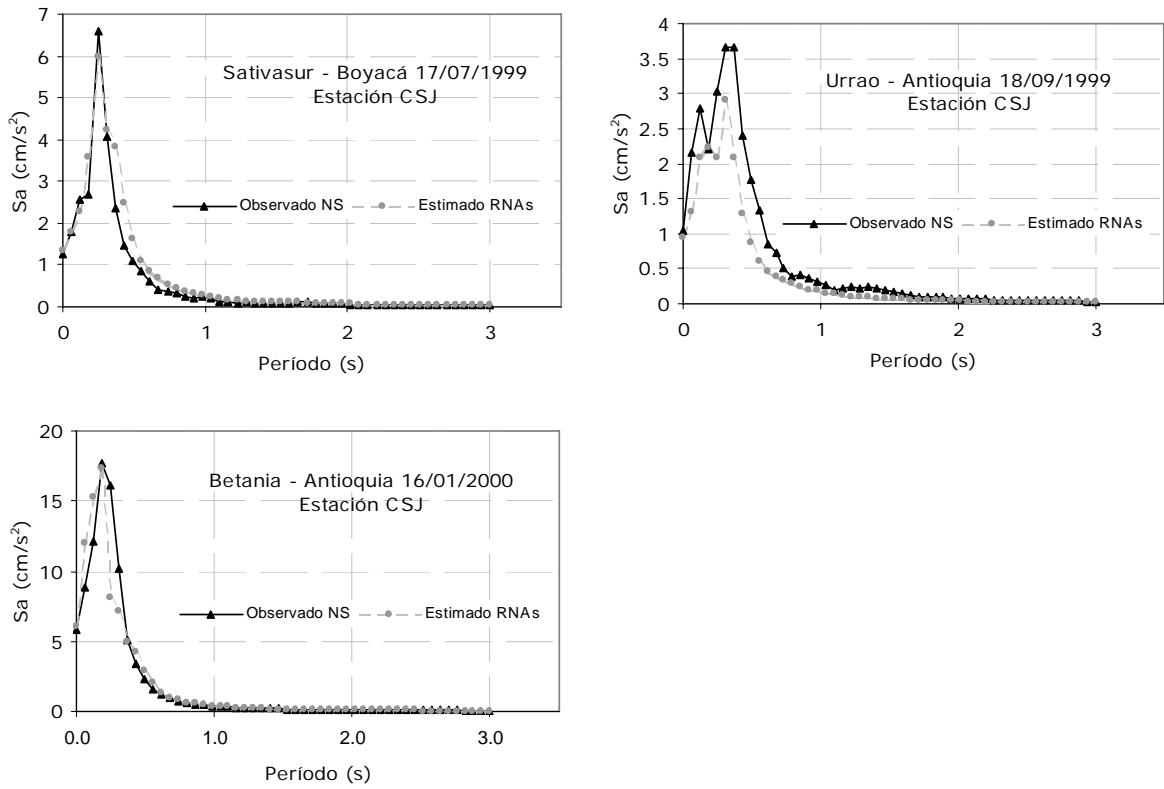


Figura 5.19 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para foco superficial del caso Ciudad de Medellín



Continúa figura 5.19 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para foco superficial del caso Ciudad de Medellín

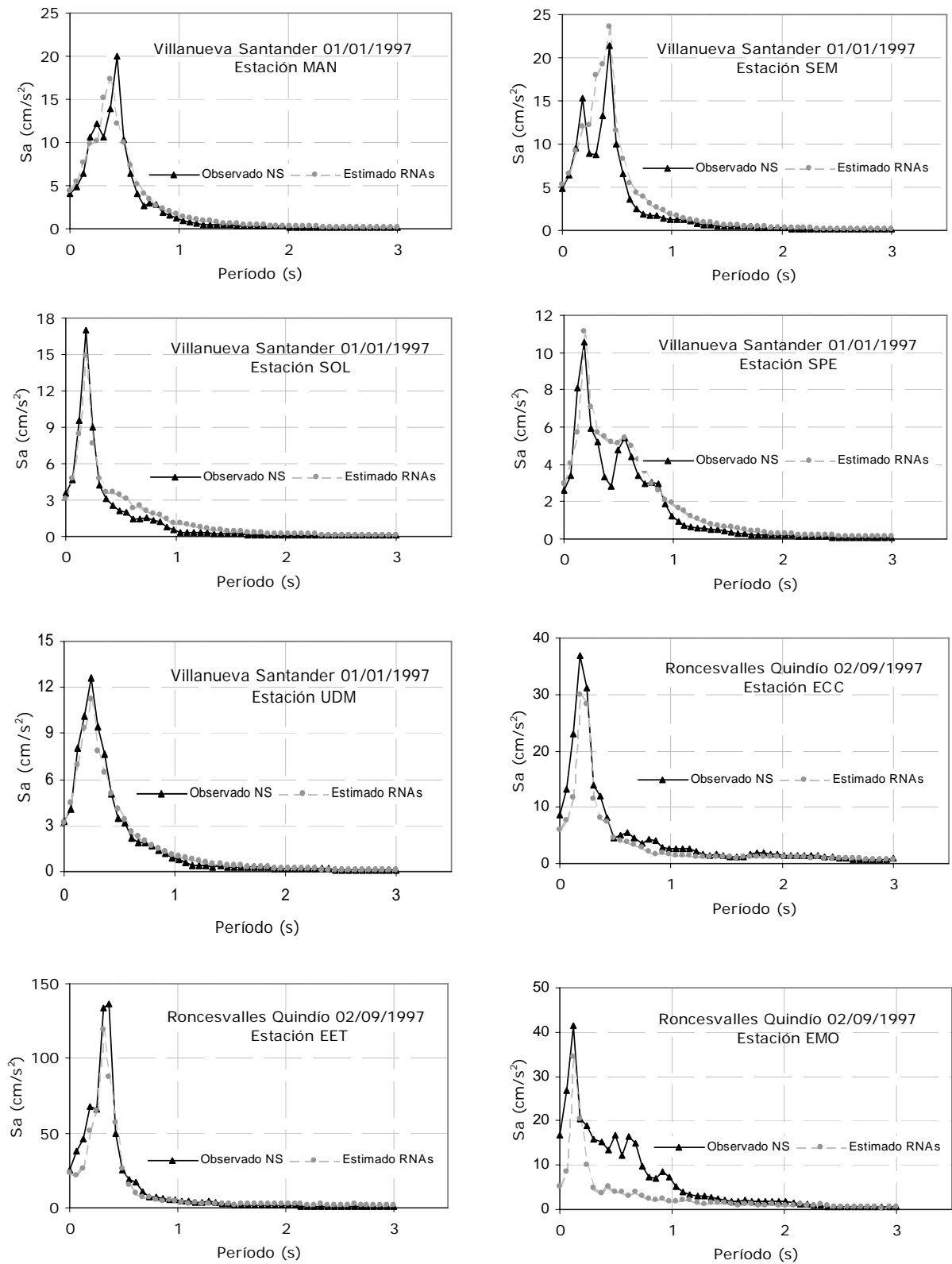
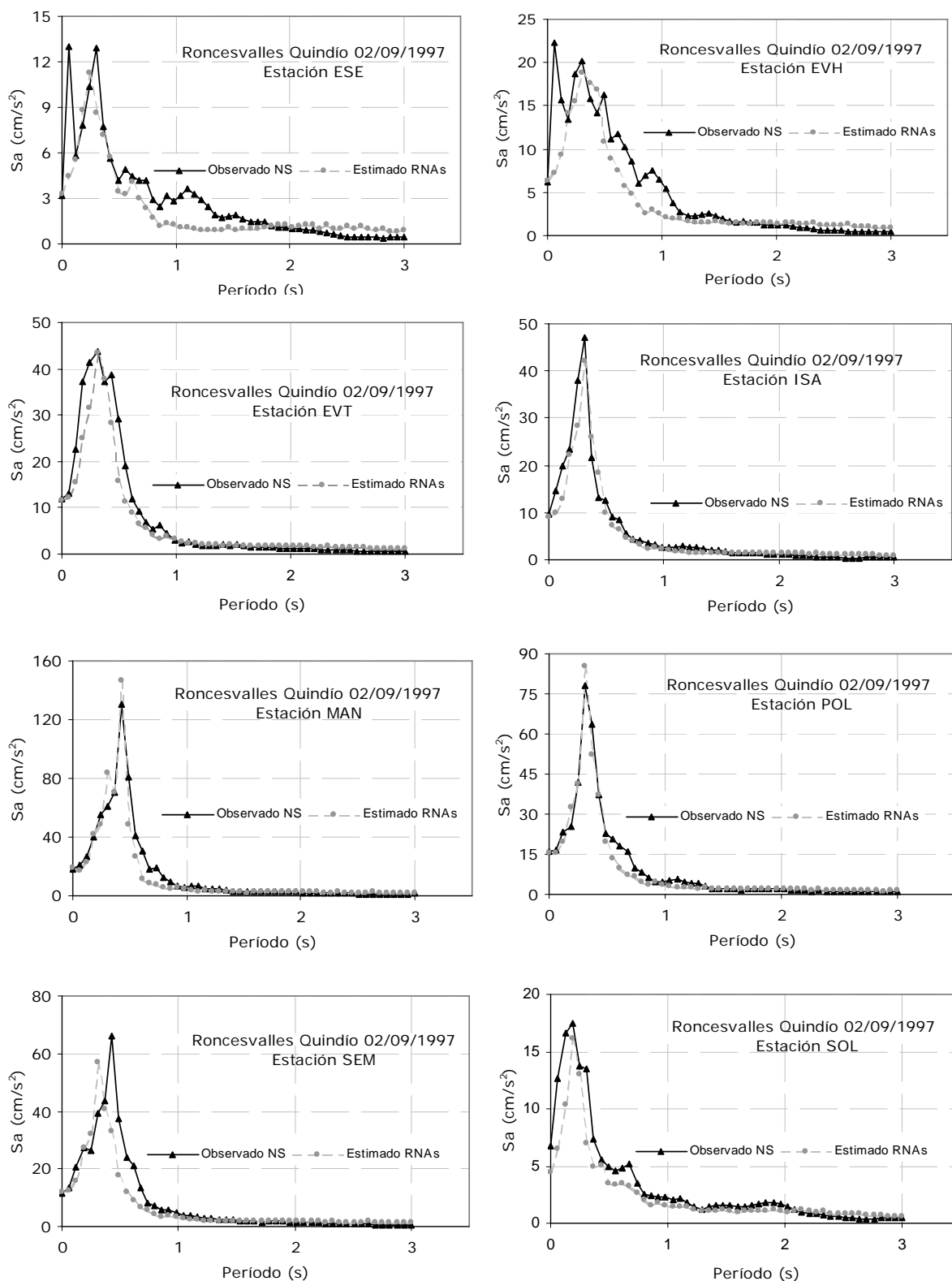
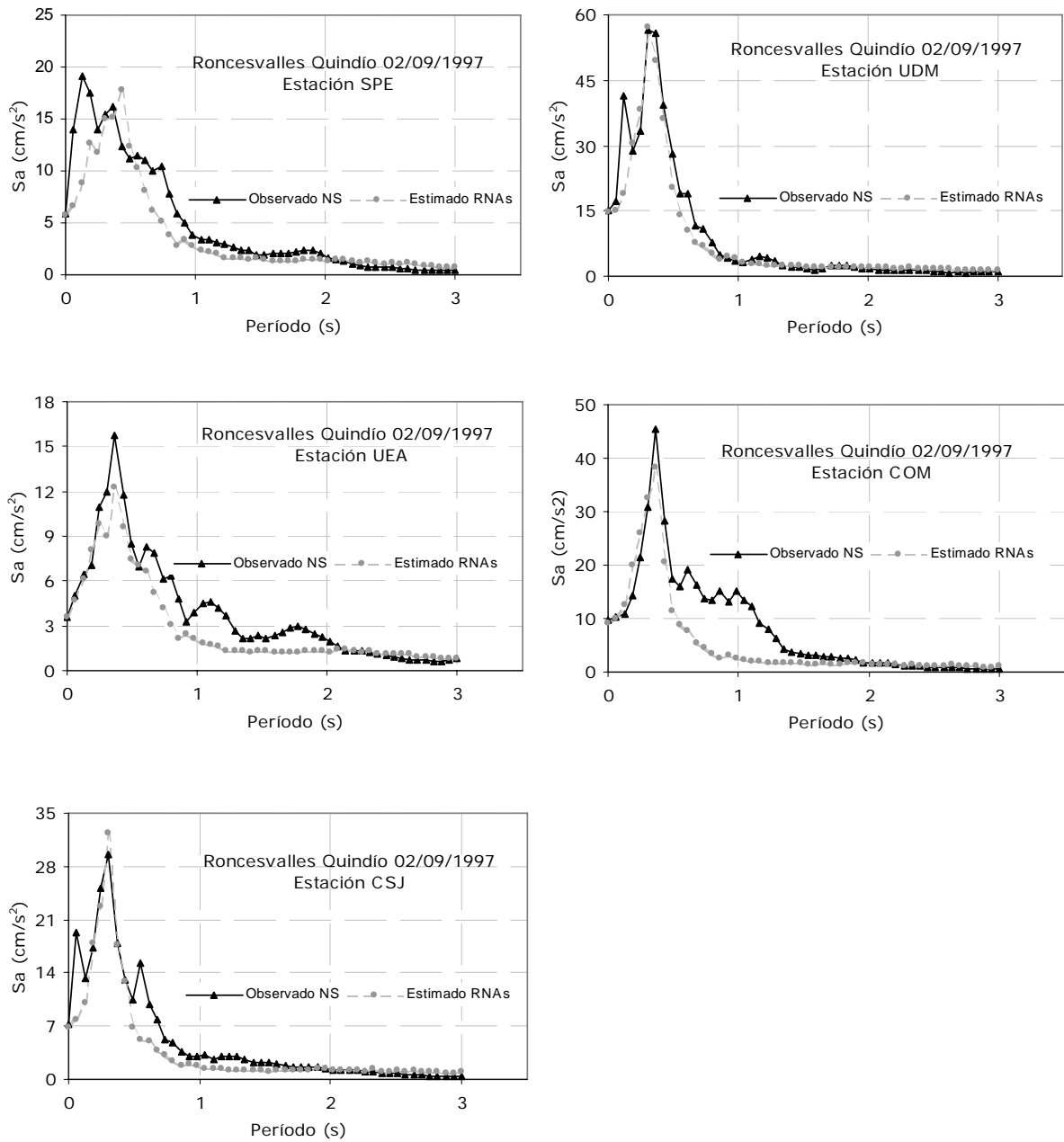


Figura 5.20 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para foco lejano del caso Ciudad de Medellín.



Continúa figura 5.20 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para foco lejano del caso Ciudad de Medellín.



Continúa figura 5.20 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para foco lejano del caso Ciudad de Medellín.

5.3.3 Caso Colombia

Igual que para los caso anteriores, se muestra la comparación entre los valores de aceleración máxima horizontal componente NS medidos y los estimados con la RNA, cuya representación gráfica corresponde a la figura 5.21; y los resultados de los espectros de respuesta elásticos calculados a partir del registro original y los predichos con el procedimiento neuronal en la etapa de ejecución para las agrupaciones fijadas en el caso colombiano, foco superficial figura 5.22 y foco lejano figura 5.23.

De estas representaciones se infiere que los resultados obtenidos con el modelo neuronal indicado en la tabla 5.4 para la aceleración máxima, son bastantes congruentes con los valores medidos; se alcanzó un porcentaje de correlación en la fase de validación de 97.8 (tabla 5.6). La estimación de espectros de respuesta elásticos con la RNA cuyas características se detallan también en la tabla 5.4, igualmente muestran la capacidad predictiva del modelo, aunque las correlaciones en grupo de prueba parecen ser bajas (87% foco superficial y 82% foco lejano) y en varios casos las predicciones son un tanto deficientes como se aprecia para los eventos de foco superficial Tauramena–Casanare estación PRAD, Juradó–Chocó estación PAMP, Sismo del Quindío estación CSANBAR2, y los sismos de foco lejano Calima-Valle estación PRAD, Calima-Valle estación TADO, Génova-Quindío estación CSTRC, los demás resultados 10 de 14 en el caso de foco superficial y 21 de 24 en el caso de foco lejano indican que la RNA puede calcular con suficiente aproximación la respuesta sísmica observada en sitios ubicados a todo lo largo y ancho del país. Estos resultados se consiguieron a partir de una muestra de 280 acelerogramas, registrados la mayoría en estaciones sísmicas localizadas en roca y procedentes de 59 eventos con magnitudes entre 4.1 y 7.2. En la medida en que se pueda ampliar la base de datos e incorporarla en la RNA como sistema dinámico autoadaptativo que es, se disminuirá la dispersión.

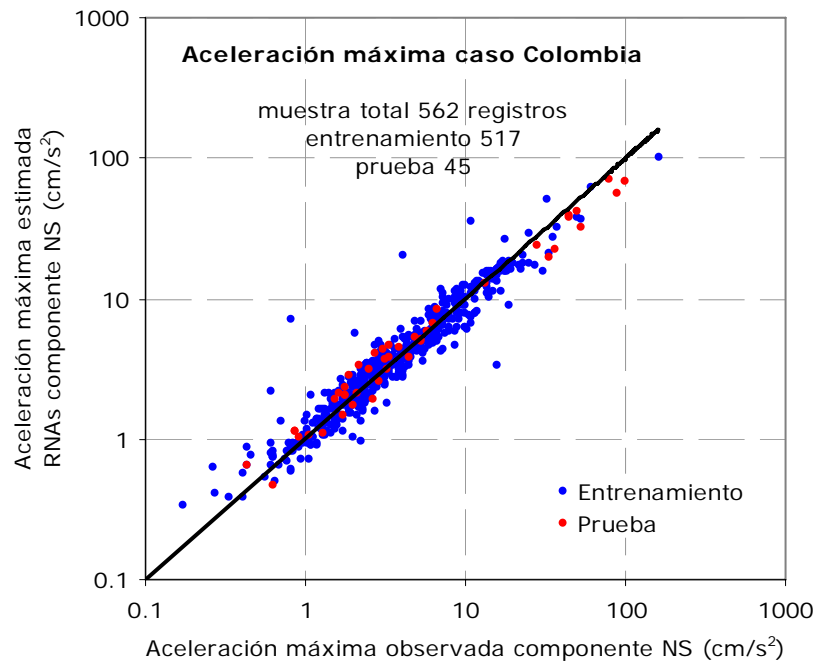


Figura 5.21 Comparación entre los valores de aceleración máxima componente NS observados y los estimados con RNAs.

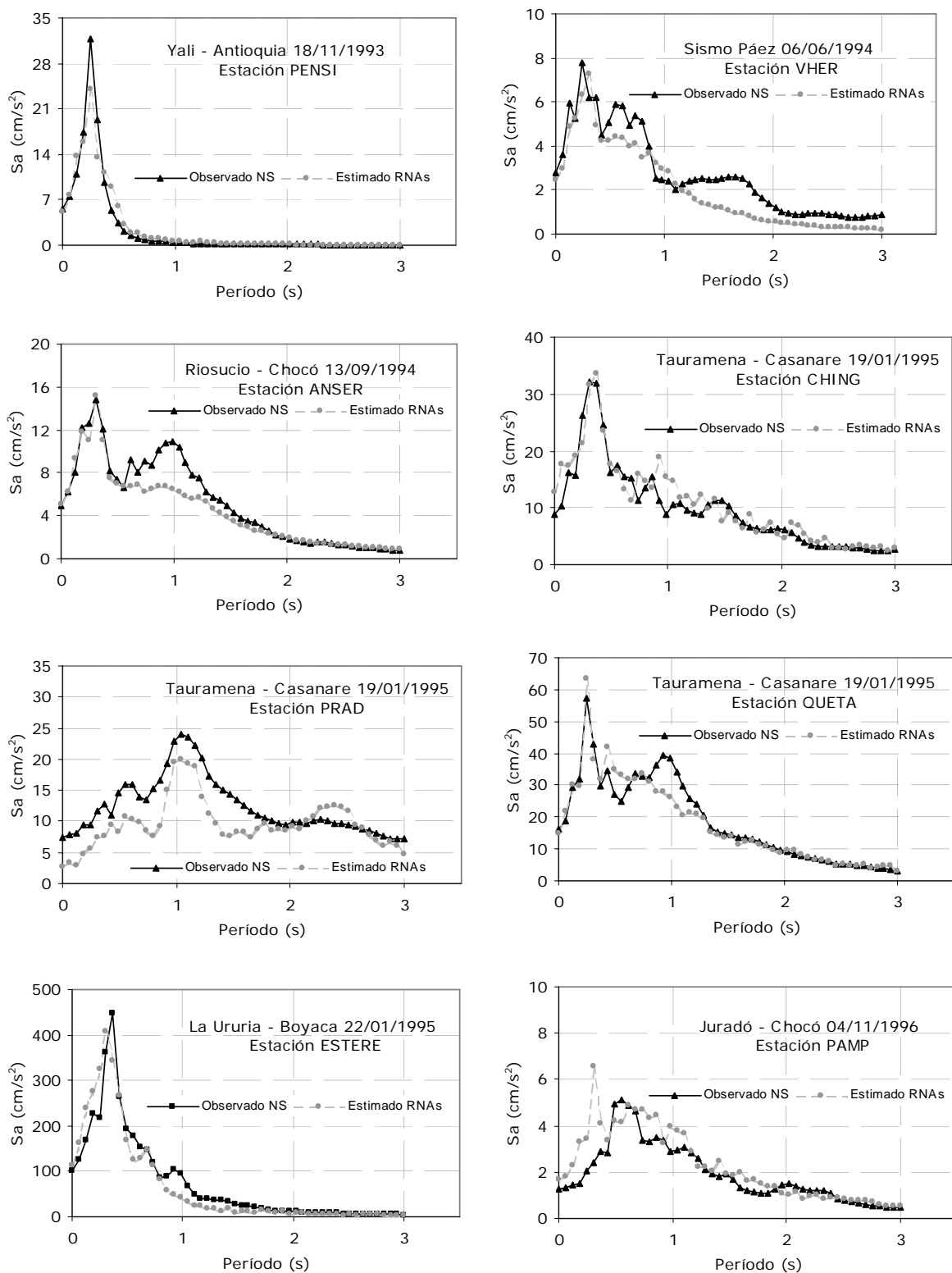
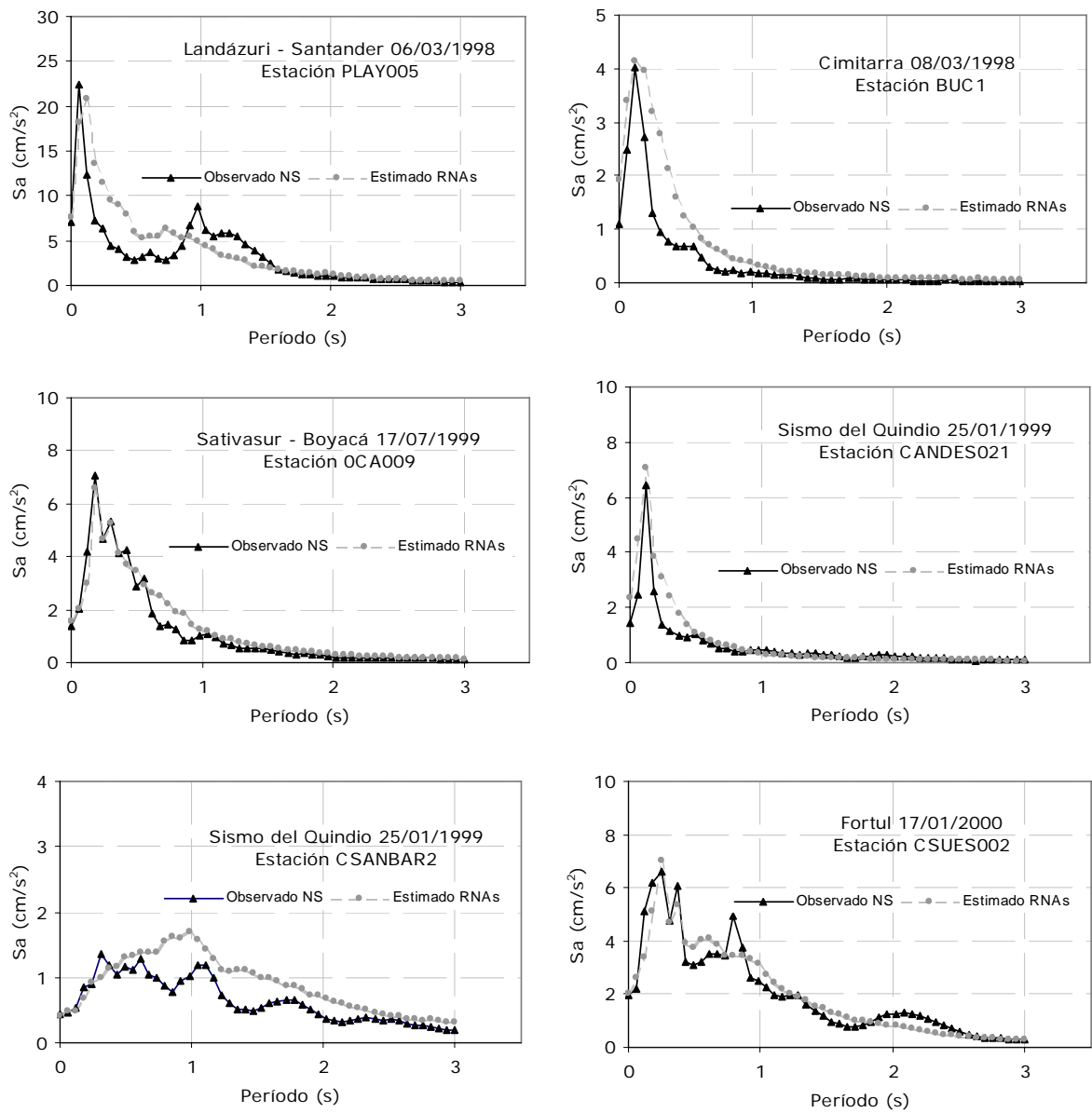


Figura 5.22 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para foco superficial caso Colombia.



Continúa figura 5.22 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para foco superficial caso Colombia.

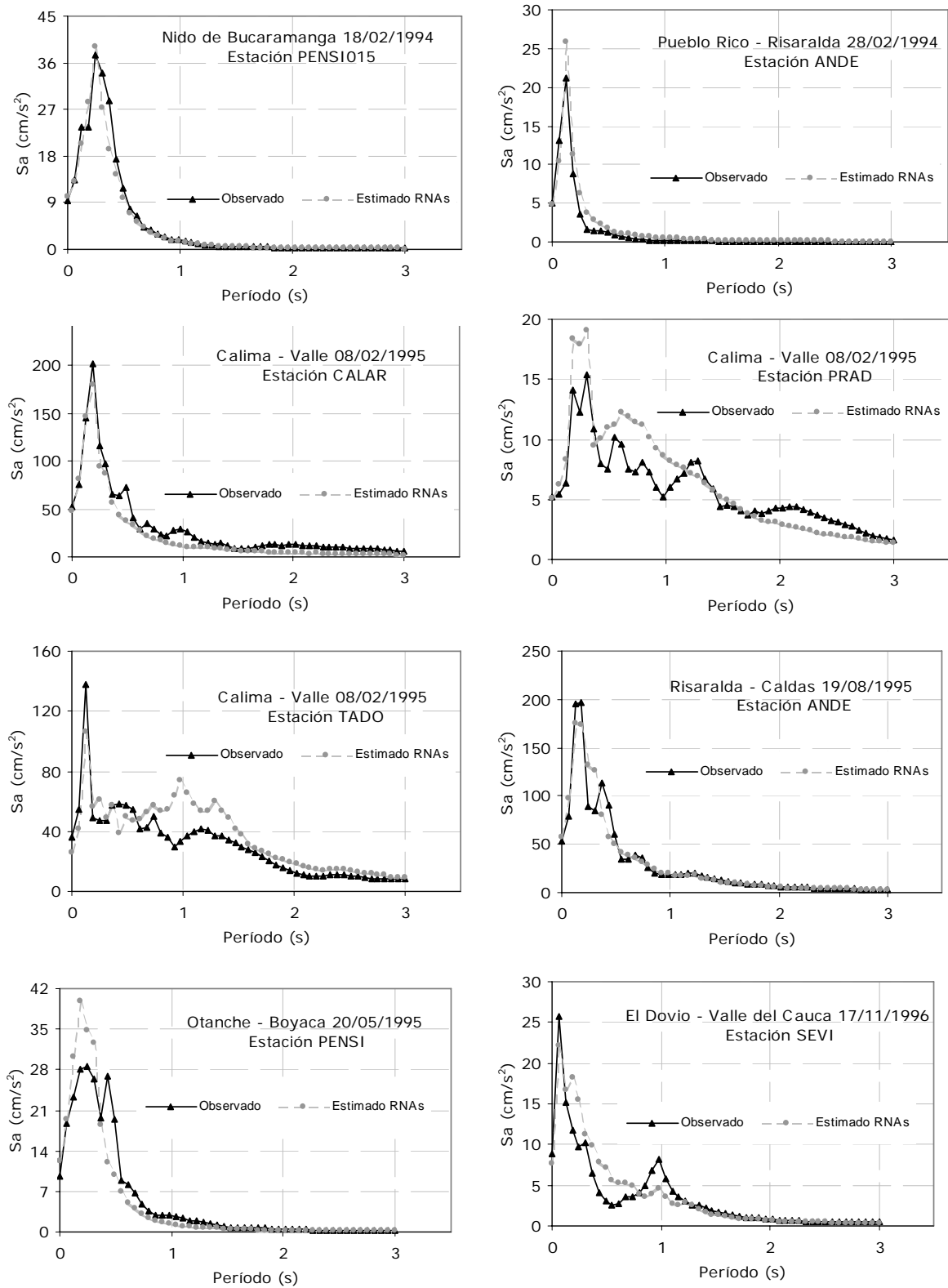
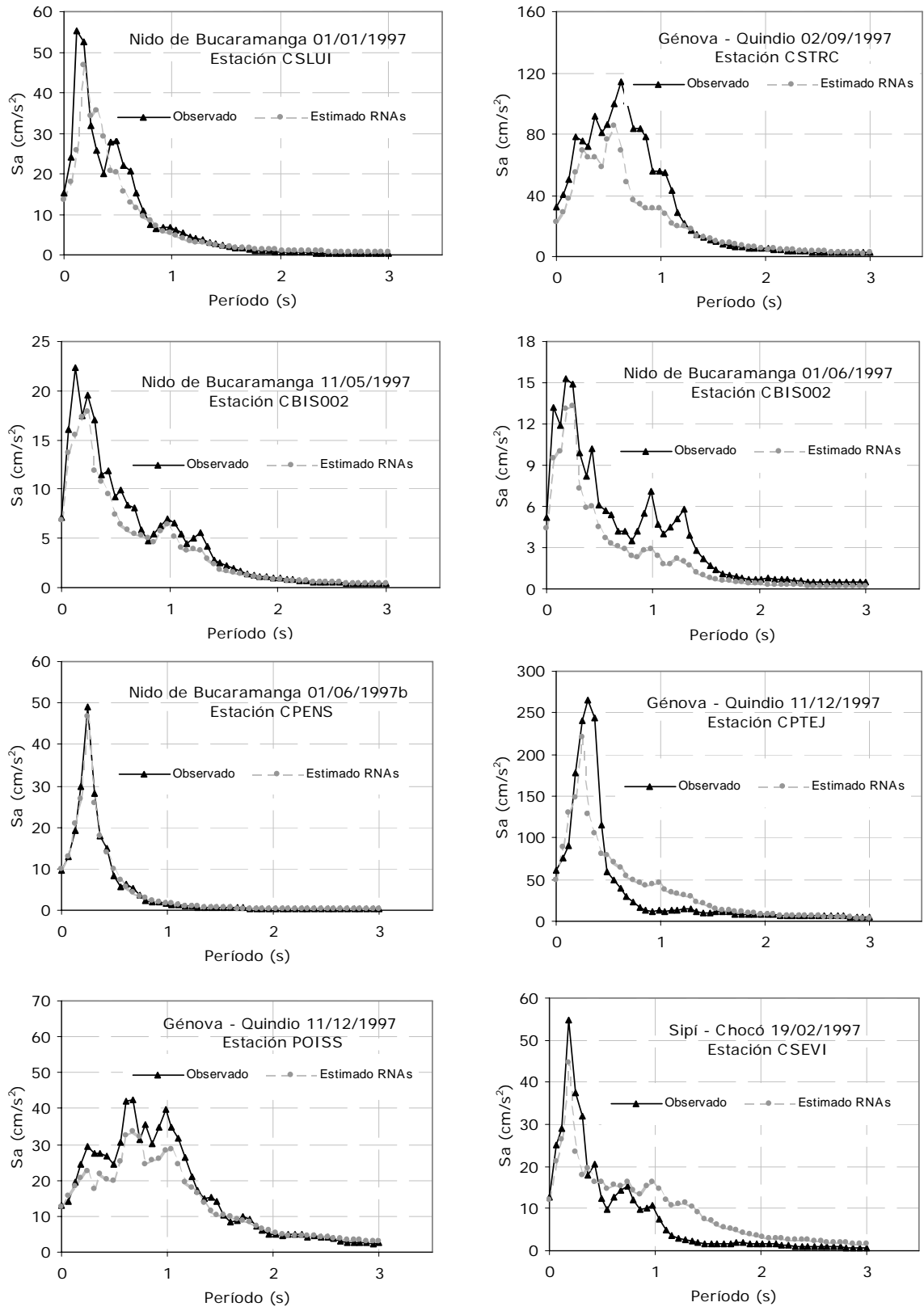
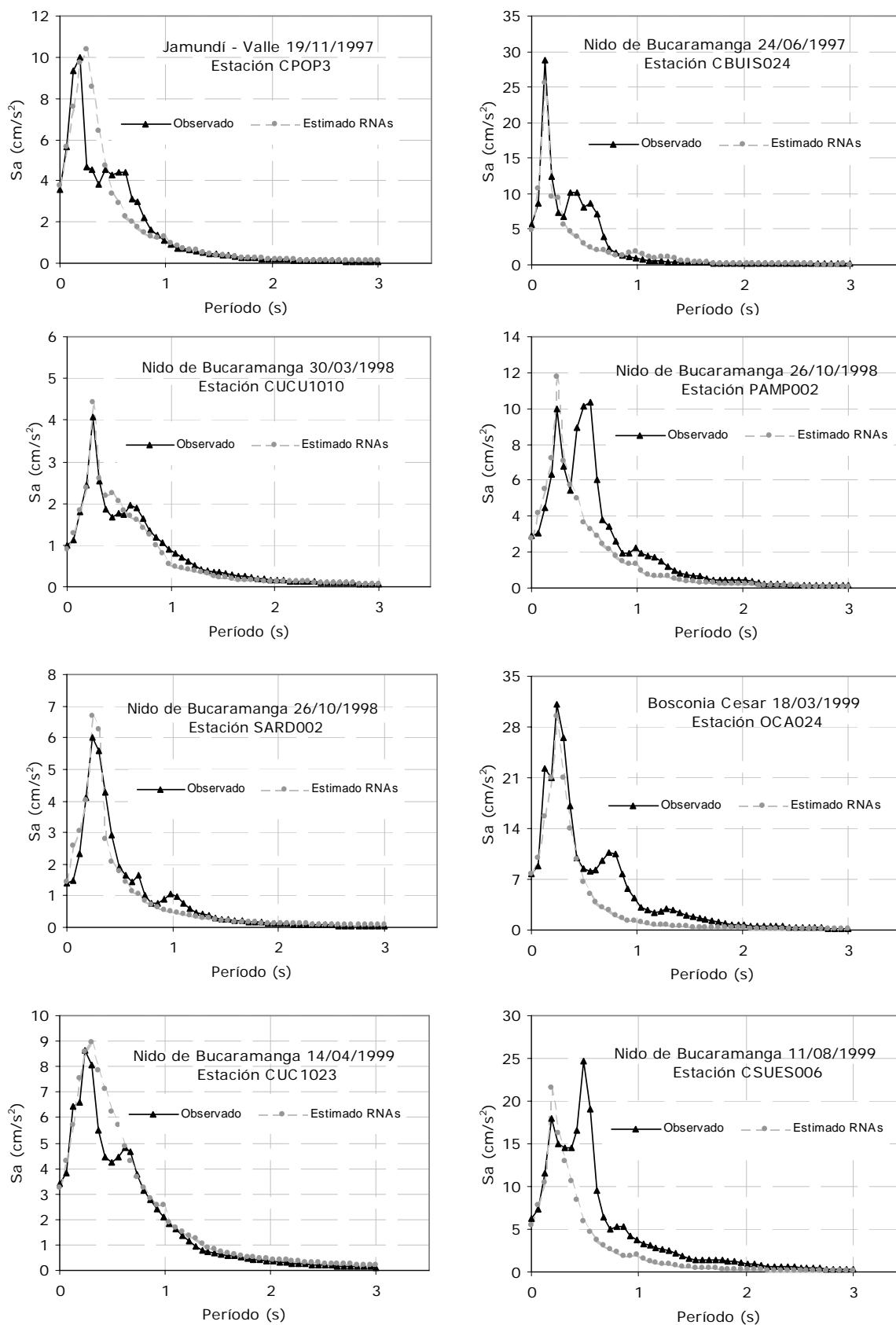


Figura 5.23 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para foco lejano caso Colombia.



Continúa figura 5.23 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para foco lejano caso Colombia.



Continúa figura 5.23 Espectros de respuesta elásticos (5% de amortiguamiento) calculados a partir del registro original y los estimados con RNAs en la etapa de prueba para foco lejano caso Colombia.

5.4 Validación del método

5.4.1 Comparación con los valores observados

La comparación gráfica entre lo observado y lo estimado ilustrado en las figuras anteriores, es por sí misma una forma (quizá la más rigurosa) de validar los resultados. Si bien, no se logró una generalización óptima, los índices de ejecución son satisfactorios en términos prácticos. Un mayor acercamiento al ajuste óptimo se puede conseguir incrementando el número de patrones de entrenamiento, pues los estimadores no paramétricos suelen ser muy sensibles a los casos particulares de pares entrada-salida seleccionados para realizar el aprendizaje. Por otro lado, La eficiencia del sistema neuronal puede mejorar si se introduce explícitamente el conocimiento a priori en los modelos neuronales en la medida que se comprenda el fenómeno.

De acuerdo con los resultados obtenidos, se puede concluir que los sistemas neuronales artificiales proveen una buena solución a problemas con grandes bases de datos que almacenan ingentes cantidades de información y en los que existen muchos casos particulares. Aunque no resuelve el problema con rigor matemático, las soluciones aproximadas que se consiguen es una característica valiosa al no requerir ninguna función prescrita ni la independencia de las variables, y al permitir obviar el problema de la no linealidad en sistemas complejos proporcionando soluciones de fácil implementación.

5.4.2 Comparación estadística con otros trabajos

Se presenta a continuación el comportamiento estadístico a partir de los valores de desviación estándar, los cuales reflejan en forma simple el comportamiento global de las modelaciones en la predicción, tanto para los procedimientos convencionales como para los sistemas de inteligencia artificial. Se consideró este parámetro como base de comparación por ser un indicador estadístico de dispersión ampliamente aceptado en la literatura técnica [Abrahamson 97], [Somerville 98].

El error cometido al estimar el logaritmo natural del parámetro de interés (D_i , A_{max} , S_a) en términos de la desviación estándar en los modelos neuronales se presentó en la tabla 5.6, estos se comparan a continuación con el valor de la desviación estándar para leyes de atenuación propuesta por otros autores cuyos valores se indican en la tabla 5.9. Según los datos especificados, los errores cometidos con los modelos neuronales son significativamente menores a los obtenidos por otros procedimientos. A partir de este criterio se infiere que el procedimiento con RNAs se pueden emplear fructíferamente en la ingeniería sísmica.

Tabla 5.9 Desviación estándar del logaritmo de la aceleración máxima del terreno $\sigma_{\log A_{max}}$

Ley de atenuación	$\sigma_{\log A_{max}}$ (cm/s ²)	Ley de atenuación	$\sigma_{\log A_{max}}$ (cm/s ²)
Esteva y Villaverde (1974)	0.28	Reyes (1999)	0.19
Joyner y Boore (1981)	0.26	Gallego (2000)	0.29
Boore y Joyner (1982) en [Reyes 99]	0.30	En este estudio (Ciudad de México)	0.10
Bufuliza (1984)	0.27	En este estudio (fuera del Valle de México)	0.12
Singh et al (1987)	0.16	En este estudio (Ciudad de Medellín)	0.19
Anderson (1997)	0.39	En este estudio (República Colombia)	0.11

5.5 Modelo paramétrico vs. redes neuronales para la estación sísmica CU

Con base en la información seleccionada en este estudio para la estación CU (base de datos capítulo 4) y considerando las variables que se indican en la tabla 5.10, se compara la solución alcanzada con la técnica de regresión lineal y el procedimiento con RNAs. Para ambos casos, la validación se llevó a cabo separando una muestra como se indica en la misma tabla.

5.5.1 Modelo lineal paramétrico

Los datos de movimiento fuerte se ajustaron por regresión lineal múltiple utilizando el método de estimación Marquardt, con el funcional de la ecuación (5.1). Esta ecuación se extendió para diferentes períodos con el fin de estimar la atenuación espectral para varios períodos

$$\ln Sa(T) = \alpha_1(T) M_w + \alpha_2(T) R + \alpha_3(T) h + \alpha_4(T) \alpha_i + \alpha_5(T) f_p + \alpha_6(T) \tag{5.1}$$

donde $Sa(T)$ aceleración espectral en cm/s^2 ; T período del sistema de un grado de libertad en s ; y M_w, R, h, α_i y f_p corresponden a las cantidades previamente definidas en la tabla 5.1.

Haciendo uso de la información reunida en la tabla 5.10 se calcularon los coeficientes empíricos del funcional ($\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ y α_6) para cada período estructural. El error estándar de la estimación en términos de la desviación estándar de los residuos y el coeficiente de correlación se reportan en la tabla 5.11. Dado lo reducido del tamaño de la muestra se especifican los intervalos de confianza en el mismo cuadro. En la tabla 5.12, se registra la desviación estándar de los residuos del conjunto de datos utilizados en la validación. Para una mejor y rápida visualización, estos últimos valores se graficaron en función del período como se ilustra en la figura 5.24. Admite resaltar, que de acuerdo con Abrahamson [97] la mayoría de los modelos de atenuación sísmica muestran un incremento en la desviación estándar con el período, no siendo así en este caso, lo que representa un mejor ajuste.

El contraste entre los espectros observados de los eventos sísmicos utilizados para probar la regresión y los estimados a través de la ley de regresión y con la RNA, se presentan en la figura 5.25. Dichos resultados permiten apreciar que el funcional de la ecuación (5.1) proporciona un ajuste bastante aceptable (excepto para el sismo 14/09/1995 estación CUP5) y que la estimación alcanzada con la RNA reproduce con mejor similitud los espectros observados. El detalle importante del funcional propuesto es la incorporación de los últimos tres términos profundidad del sismo, parámetro azimutal y frecuencia dominante. Variables que hasta ahora, ninguna de ellas han sido incluidas explícitamente en los modelos de regresión para la estimación de movimientos fuertes.

Tabla 5.10 Información sísmica de la estación de registro sísmica CU de la Ciudad de México seleccionada para generar la ley de atenuación espectral y para la modelación de la RNA

Datos empleados en la regresión para obtener los coeficientes empíricos y para el entrenamiento de la RNA						
Sismo	Estación	Magnitud momento	Distancia epicentral km	Profundidad km	Coseno del ángulo entre epifoco-sitio	Frecuencia predominante
29/11/1978	SXCU	7.80	459.0	19.0	0.99251548	0.550
14/03/1979	SXCU	7.40	278.0	28.0	0.58568802	0.626
24/10/1980	SXCU	7.10	159.0	70.0	0.76416491	1.140
25/10/1981	SXCU	7.20	342.0	14.0	0.40995015	0.550
07/06/1982	CU01	6.90	319.0	18.0	0.93441671	0.610
07/06/1982	CU01	6.90	317.0	18.0	0.95380862	0.579
19/09/1985	CU01	8.00	379.0	15.0	0.32150675	0.397
21/09/1985	CUIP	7.50	311.0	15.0	0.51898710	0.643
21/09/1985	CUMV	7.50	311.0	15.0	0.51898710	0.424
30/04/1986	CU01	6.90	413.0	20.0	0.23548281	0.446
08/02/1988	CU01	5.80	269.0	19.2	0.66867710	1.030
25/04/1989	CU01	6.90	280.0	19.0	0.99695830	0.424
25/04/1989	CU02	6.90	280.0	19.0	0.99695830	0.424
25/04/1989	CU02	6.90	280.0	19.0	0.99695830	0.424
15/05/1993	CUP1	6.00	300.0	5.0	0.95304169	0.925
15/05/1993	CUP2	6.00	300.0	5.0	0.95304169	0.925
15/05/1993	CUP3	6.00	300.0	5.0	0.95304169	0.925
15/05/1993	CUP4	6.00	300.0	5.0	0.95304169	0.327
24/10/1993	CUP1	6.60	286.0	5.0	0.98458468	0.327
24/10/1993	CUP2	6.60	286.0	5.0	0.98458468	0.643
24/10/1993	CUP3	6.60	286.0	5.0	0.98458468	0.327
24/10/1993	CUP4	6.60	286.0	5.0	0.98458468	0.327
23/05/1994	CUP1	6.20	181.0	23.0	0.60654475	0.812
23/05/1994	CUP4	6.20	181.0	23.0	0.60654475	0.812
14/09/1995	CUP1	7.30	282.0	22.0	0.97366787	0.327
14/09/1995	CUP4	7.30	282.0	22.0	0.97366787	0.327
Datos de comprobación para ambos procedimientos regresión y RNA						
Sismo	Estación	Magnitud momento	Distancia epicentral km	Profundidad km	Coseno del ángulo entre epifoco-sitio	Frecuencia predominante
14/03/1979	CU01	7.40	278.0	28.0	0.58688532	0.643
19/09/1985	CUIP	8.00	379.0	15.0	0.32150675	0.402
19/09/1985	CUMV	8.00	379.0	15.0	0.32150675	0.495
14/09/1995	CUP5	7.30	282.0	22.0	0.97366787	0.335

Tabla 5.11 Coeficientes obtenidos de la regresión lineal aplicada a eventos registrados en la estación sísmica CU de la Ciudad de México.

T = 0.01 Coeficiente de correlación = 0.927035 Error estándar de la estimación = 0.195932			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.3637$	0.169748	1.00961	1.71779
$\alpha_2 = -0.00254344$	0.00125523	-0.00516182	0.0000749334
$\alpha_3 = 0.00677478$	0.00655541	-0.0068996	0.0204491
$\alpha_4 = 0.326377$	0.240983	-0.176306	0.82906
$\alpha_5 = 0.479286$	0.330079	-0.209249	1.16782
$\alpha_6 = -7.20712$	1.12015	-9.54372	-4.87052
T = 0.11184 Coeficiente de correlación = 0.916612 Error estándar de la estimación = 0.205442			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.23154$	0.177986	0.860264	1.60281
$\alpha_2 = -0.00360424$	0.00131615	-0.0063497	-0.000858783
$\alpha_3 = 0.00981762$	0.00687357	-0.00452043	0.0241557
$\alpha_4 = 0.0959716$	0.252679	-0.431109	0.623052
$\alpha_5 = 0.567526$	0.346099	-0.154425	1.28948
$\alpha_6 = -5.7253$	1.17452	-8.1753	-3.27529
T = 0.21367 Coeficiente de correlación = 0.929744 Error estándar de la estimación = 0.192076			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.21217$	0.166407	0.86505	1.55929
$\alpha_2 = -0.00530458$	0.00123053	-0.00787142	0.0225726
$\alpha_3 = 0.0091674$	0.00642638	-0.00423783	0.0225726
$\alpha_4 = -0.043787$	0.23624	-0.536576	0.449002
$\alpha_5 = 0.586321$	0.323582	-0.0886608	1.2613
$\alpha_6 = -4.72124$	1.0981	-7.01185	-2.43063
T = 0.3151 Coeficiente de correlación = 0.962268 Error estándar de la estimación = 0.141832			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.4845$	0.122964	1.22801	1.741
$\alpha_2 = -0.00507197$	0.000909282	-0.0069687	-0.00317524
$\alpha_3 = 0.00213298$	0.00474869	-0.00777265	0.0120386
$\alpha_4 = 0.285299$	0.174567	-0.0788416	0.649439
$\alpha_5 = 0.729054$	0.239107	0.230285	1.22782
$\alpha_6 = -6.79685$	0.81143	-8.48947	-5.10423
T = 0.41735 Coeficiente de correlación = 0.889385 Error estándar de la estimación = 0.291569			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.49866$	0.1785	1.12632	1.87101
$\alpha_2 = -0.00494368$	0.00139124	-0.00784577	-0.00204158
$\alpha_3 = -0.00338025$	0.00769371	-0.0194291	0.0126686
$\alpha_4 = 0.0991684$	0.265187	-0.454002	0.652339
$\alpha_5 = 0.714835$	0.350046	-0.0153498	1.44502
$\alpha_6 = -6.62969$	1.19177	-9.11568	-4.14371
T = 0.51918 Coeficiente de correlación = 0.915746 Error estándar de la estimación = 0.194243			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.35058$	0.168284	0.999545	1.70162
$\alpha_2 = -0.00391829$	-0.00391829	-0.00651409	-0.00132249
$\alpha_3 = 0.00111832$	0.00649889	-0.0124381	0.0146748
$\alpha_4 = 0.290553$	0.238905	-0.207796	0.788902
$\alpha_5 = 0.5778$	0.327233	-0.104798	1.2604
$\alpha_6 = -5.91818$	1.11049	-8.23463	-3.60172
T = 0.62102 Coeficiente de correlación = 0.880026 Error estándar de la estimación = 0.247692			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.36408$	0.21459	0.916457	1.81171
$\alpha_2 = -0.00328472$	0.00158683	-0.0065948	0.0000253506
$\alpha_3 = 0.00406309$	0.00828716	-0.0132237	-0.0132237
$\alpha_4 = 0.150091$	0.304644	-0.485386	0.785569
$\alpha_5 = 0.84436$	0.417276	-0.026065	1.71478
$\alpha_6 = -6.34154$	1.41606	-9.29541	-3.38768
T = 0.72286 Coeficiente de correlación = 0.921914 Error estándar de la estimación = 0.21114			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.36293$	0.182923	0.981355	1.7445
$\alpha_2 = -0.00268582$	0.00135266	-0.00550742	0.000135785
$\alpha_3 = 0.00989051$	0.00706421	-0.00484521	0.0246262
$\alpha_4 = 0.394097$	0.259687	-0.147603	0.935796
$\alpha_5 = 0.457405$	0.355698	-0.284571	1.19938
$\alpha_6 = -6.48279$	1.20709	-9.00075	-3.96483
T = 0.82469 Coeficiente de correlación = 0.912905 Error estándar de la estimación = 0.23713			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.45451$	0.20544	1.02597	1.88306
$\alpha_2 = -0.00187835$	0.00151916	-0.00504728	0.00129058
$\alpha_3 = 0.0102352$	0.00793378	-0.00631439	0.0267848
$\alpha_4 = 0.43791$	0.291654	-0.17047	1.04629
$\alpha_5 = 0.534686$	0.399483	-0.298623	1.36799
$\alpha_6 = -7.48221$	1.35568	-10.3101	-4.65431

Continúa tabla 5.11

T = 0.92653 Coeficiente de correlación = 0.905382 Error estándar de la estimación = 0.236403			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.46567$	0.20481	1.03844	1.89289
$\alpha_2 = -0.00181406$	0.00151451	-0.00497327	-0.00497327
$\alpha_3 = -0.00497327$	0.00790947	-0.011127	0.0218708
$\alpha_4 = 0.235296$	0.29076	-0.37122	-0.37122
$\alpha_5 = 0.70595$	0.398259	-0.124805	1.5367
$\alpha_6 = -7.32997$	1.35153	-10.1492	-4.51073
T = 1.0284 Coeficiente de correlación = 0.888373 Error estándar de la estimación = 0.23963			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.61733$	0.207473	1.18587	2.0488
$\alpha_2 = -0.00179165$	0.00153183	-0.00497727	0.00139397
$\alpha_3 = -0.00397953$	0.00801334	-0.0206442	0.0126852
$\alpha_4 = 0.685761$	0.292035	0.0784399	1.29308
$\alpha_5 = 1.22286$	0.412812	0.361747	2.08397
$\alpha_6 = -8.95978$	1.36993	-11.8087	-6.11084
T = 1.1302 Coeficiente de correlación = 0.910229 Error estándar de la estimación = 0.221563			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.59229$	0.191953	1.19189	1.9927
$\alpha_2 = -0.000850063$	0.00141943	-0.00381095	0.00211083
$\alpha_3 = -0.00434014$	0.00741294	-0.0198033	0.011123
$\alpha_4 = 0.528137$	0.272507	-0.0403037	1.09658
$\alpha_5 = 0.933914$	0.373257	0.155312	1.71252
$\alpha_6 = -8.82856$	1.26668	-11.4708	-6.1863
T = 1.232 Coeficiente de correlación = 0.902003 Error estándar de la estimación = 0.225273			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.56092$	0.195167	1.15381	1.96804
$\alpha_2 = -0.00125436$	0.0014432	-0.00426484	0.00175611
$\alpha_3 = -0.00778596$	0.00753707	-0.0235081	0.00793613
$\alpha_4 = 0.265469$	0.27707	-0.312491	0.843428
$\alpha_5 = 0.800185$	0.379508	0.00854368	1.59183
$\alpha_6 = -8.20254$	1.28789	-10.889	-5.51604
T = 1.3339 Coeficiente de correlación = 0.882519 Error estándar de la estimación = 0.245544			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.594$	0.21273	1.15025	2.03775
$\alpha_2 = -0.00100973$	0.00157308	-0.00429112	0.00227166
$\alpha_3 = -0.0111907$	0.00821534	-0.0283276	0.00594627
$\alpha_4 = 0.397844$	0.302004	-0.232127	1.02781
$\alpha_5 = 0.907908$	0.41366	0.0450266	1.77079
$\alpha_6 = -8.66818$	1.40379	-11.5964	-5.73991
T = 1.4357 Coeficiente de correlación = 0.890262 Error estándar de la estimación = 0.27696			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.67745$	0.239946	1.17693	2.17797
$\alpha_2 = -0.000176164$	0.00177433	-0.00387736	0.00352503
$\alpha_3 = -0.00285349$	0.00926638	-0.0221829	0.0164759
$\alpha_4 = 0.458114$	0.340641	-0.252453	1.16868
$\alpha_5 = 0.728173$	0.466582	-0.245103	1.70145
$\alpha_6 = -9.56967$	1.58339	-12.8726	-6.26677
T = 1.5376 Coeficiente de correlación = 0.889841 Error estándar de la estimación = 0.316896			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.77465$	0.274519	1.20201	2.34728
$\alpha_2 = 0.000364732$	0.00202999	-0.00386976	0.00459922
$\alpha_3 = 0.00185993$	0.0106015	-0.0202545	0.0239744
$\alpha_4 = 0.581859$	0.389723	-0.231091	1.39481
$\alpha_5 = 0.416199$	0.53381	-0.697312	1.52971
$\alpha_6 = -10.3494$	1.81153	-14.1282	-6.57064
T = 1.6394 Coeficiente de correlación = 0.889054 Error estándar de la estimación = 0.356863			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.80282$	0.309171	1.15789	2.44774
$\alpha_2 = 0.00102167$	0.00228623	-0.00374733	0.00579067
$\alpha_3 = 0.00745575$	0.0119397	-0.0174502	0.0323617
$\alpha_4 = 0.63246$	0.438917	-0.283106	1.54803
$\alpha_5 = -0.102219$	0.601192	-1.35629	1.15185
$\alpha_6 = -10.5064$	2.0402	-14.7622	-6.25062
T = 1.7412 Coeficiente de correlación = 0.878932 Error estándar de la estimación = 0.379063			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.80476$	0.328195	1.12223	2.48728
$\alpha_2 = 0.00229761$	0.00242315	-0.00274162	0.00733685
$\alpha_3 = 0.0087092$	0.0126761	-0.0176522	0.0350706
$\alpha_4 = 0.570246$	0.461961	-0.390455	1.53095
$\alpha_5 = -0.0777791$	0.642324	-1.41765	1.26209
$\alpha_6 = -10.9722$	2.16705	-15.4788	-6.46551

Continúa tabla 5.11

T = 1.8431 Coeficiente de correlación = 0.727995 Error estándar de la estimación = 0.373889			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.83778$	0.323922	1.16209	2.51347
$\alpha_2 = 0.00219004$	0.00239531	-0.00280649	0.00718657
$\alpha_3 = 0.00729115$	0.0125094	-0.018803	0.0333853
$\alpha_4 = 0.830895$	0.459858	-0.128353	1.79014
$\alpha_5 = 0.234177$	0.629875	-1.07972	1.54808
$\alpha_6 = -11.5289$	2.13754	-15.9878	-7.0701
T = 1.9449 Coeficiente de correlación = 0.879531 Error estándar de la estimación = 0.858409			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.46161$	0.369103	0.69167	2.23154
$\alpha_2 = 0.000829982$	0.00263792	-0.00467263	0.00633259
$\alpha_3 = 0.0139839$	0.0147384	-0.0167599	0.0447276
$\alpha_4 = 0.374824$	0.489527	-0.646314	1.39596
$\alpha_5 = -0.525479$	0.764635	-2.12048	1.06953
$\alpha_6 = -7.8924$	2.54065	-13.1921	-2.5927
T = 2.0467 Coeficiente de correlación = 0.888122 Error estándar de la estimación = 0.374395			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.74705$	0.324361	1.07044	2.42365
$\alpha_2 = 0.00214013$	0.00239855	-0.00286317	0.00714342
$\alpha_3 = 0.0118075$	0.0125263	-0.014322	0.0379371
$\alpha_4 = 0.993584$	0.46048	0.0330372	1.95413
$\alpha_5 = -0.435434$	0.630728	-1.75111	0.880243
$\alpha_6 = -10.9194$	2.14043	-15.3843	-6.45455
T = 2.1486 Coeficiente de correlación = 0.881526 Error estándar de la estimación = 0.384882			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.66795$	0.333446	0.972395	2.36351
$\alpha_2 = 0.00255057$	0.00246573	-0.00259287	0.00769402
$\alpha_3 = 0.0146668$	0.0128772	-0.0121946	0.0415282
$\alpha_4 = 1.15328$	0.473379	0.165825	2.14073
$\alpha_5 = -0.574707$	0.648395	-1.92724	0.777824
$\alpha_6 = -10.5771$	2.20038	-15.167	-5.98716
T = 2.2504 Coeficiente de correlación = 0.862448 Error estándar de la estimación = 0.413907			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.65411$	0.358592	0.906101	2.40212
$\alpha_2 = 0.00333307$	0.00265168	-0.00219825	0.00886439
$\alpha_3 = 0.013531$	0.0138483	-0.0153562	0.0424181
$\alpha_4 = 1.32914$	0.509077	0.267218	2.39106
$\alpha_5 = -0.425519$	0.697292	-1.88005	1.02901
$\alpha_6 = -10.9537$	2.36632	-15.8897	-6.0176
T = 2.3522 Coeficiente de correlación = 0.850702 Error estándar de la estimación = 0.424379			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.65677$	0.367664	0.88984	2.42371
$\alpha_2 = 0.00383981$	0.00271876	-0.00183144	0.00951106
$\alpha_3 = 0.00982318$	0.0141986	-0.0197947	0.0394411
$\alpha_4 = 1.49396$	0.521956	0.405181	2.58275
$\alpha_5 = -0.22551$	0.714932	-1.71684	1.26582
$\alpha_6 = -11.3206$	2.42619	-16.3816	-6.2597
T = 2.4541 Coeficiente de correlación = 0.843288 Error estándar de la estimación = 0.429468			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.67653$	0.372073	0.900397	2.45266
$\alpha_2 = 0.00346004$	0.00275137	-0.00227923	0.00919931
$\alpha_3 = 0.00453677$	0.0143689	-0.0254363	0.0345099
$\alpha_4 = 1.65281$	0.528216	0.550971	2.75465
$\alpha_5 = -0.204054$	0.723506	-1.71326	1.30516
$\alpha_6 = -11.4126$	2.45528	-16.5342	-6.29094
T = 2.5559 Coeficiente de correlación = 0.837618 Error estándar de la estimación = 0.435204			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.71199$	0.377043	0.925493	2.49849
$\alpha_2 = 0.00289344$	0.00278812	-0.00292249	0.00870937
$\alpha_3 = -0.000477486$	0.0145609	-0.030851	0.029896
$\alpha_4 = 1.69438$	0.535271	0.577818	2.81093
$\alpha_5 = -0.208382$	0.73317	-1.73775	1.32099
$\alpha_6 = -11.4485$	2.48808	-16.6386	-6.25849
T = 2.6578 Coeficiente de correlación = 0.837209 Error estándar de la estimación = 0.428588			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.73281$	0.371311	0.95827	2.50736
$\alpha_2 = 0.00219201$	0.00274573	-0.00353551	0.00791952
$\alpha_3 = -0.00577128$	0.0143395	-0.035683	0.0241404
$\alpha_4 = 1.65731$	0.527134	0.557723	2.75689
$\alpha_5 = -0.223362$	0.722025	-1.72948	1.28276
$\alpha_6 = -11.2825$	2.45025	-16.3936	-6.17131

Continúa tabla 5.11

T = 2.7596 Coeficiente de correlación = 0.832129 Error estándar de la estimación = 0.430376			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.69021$	0.37286	0.91244	2.46799
$\alpha_2 = 0.00229679$	0.00275719	-0.00345462	0.0080482
$\alpha_3 = -0.00513702$	0.0143993	-0.0351735	0.0248995
$\alpha_4 = 1.74914$	0.529333	0.644966	2.85331
$\alpha_5 = -0.232638$	0.725037	-1.74504	1.27977
$\alpha_6 = -11.1605$	2.46048	-16.2929	-6.02799
T = 2.8614 Coeficiente de correlación = 0.839599 Error estándar de la estimación = 0.404892			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.65999$	0.367953	0.892454	2.42753
$\alpha_2 = 0.00278869$	0.0027209	-0.00288702	0.00846441
$\alpha_3 = -0.00279636$	0.0142098	-0.0324376	0.0268449
$\alpha_4 = 1.97478$	0.522367	0.885144	3.06442
$\alpha_5 = -0.265874$	0.715495	-1.75837	1.22662
$\alpha_6 = -11.3469$	2.42809	-16.4118	-6.28193
T = 2.9633 Coeficiente de correlación = 0.850481 Error estándar de la estimación = 0.416317			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.649$	0.36068	0.896637	2.40137
$\alpha_2 = 0.0035575$	0.00266712	-0.00200603	0.00912102
$\alpha_3 = -0.000867548$	0.0139289	-0.0299229	0.0281878
$\alpha_4 = 2.19615$	0.512041	1.12805	3.26425
$\alpha_5 = -0.218115$	0.701352	-1.68111	1.24488
$\alpha_6 = -11.7881$	2.3801	-16.753	-6.82334
T = 3.051 Coeficiente de correlación = 0.888546 Error estándar de la estimación = 0.388581			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.70908$	0.336651	1.00684	2.41132
$\alpha_2 = 0.00349337$	0.00248943	-0.0016995	0.00868624
$\alpha_3 =$	0.013001	-0.029894	0.0243452
$\alpha_4 = 2.33451$	0.477928	1.33756	3.33145
$\alpha_5 = -0.0455335$	0.654626	-1.41106	1.3200
$\alpha_6 = -12.3967$	2.22153	-17.0307	-7.76262
T = 3.1669 Coeficiente de correlación = 0.880727 Error estándar de la estimación = 0.365402			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.74129$	0.316569	1.08093	2.40164
$\alpha_2 = 0.00343167$	0.00234094	-0.00145144	0.00831479
$\alpha_3 = -0.00392766$	0.0122255	-0.0294296	0.0215742
$\alpha_4 = 2.29046$	0.449419	1.35298	3.22793
$\alpha_5 = 0.0742167$	0.615578	-1.20986	1.35829
$\alpha_6 = -12.6578$	2.08902	-17.0154	-8.30018
T = 3.2688 Coeficiente de correlación = 0.878872 Error estándar de la estimación = 0.368206			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.73766$	0.318998	1.07224	2.40308
$\alpha_2 = 0.00369589$	0.0023589	-0.00122469	0.00861647
$\alpha_3 = -0.00115382$	0.0123192	-0.0268514	0.0245437
$\alpha_4 = 2.26806$	0.452867	1.32339	3.21272
$\alpha_5 = 0.147521$	0.620301	-1.14641	1.44145
$\alpha_6 = 1.44145$	2.10504	-17.2418	-8.45973
T = 3.3706 Coeficiente de correlación = 0.842349 Error estándar de la estimación = 0.372362			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.69541$	0.322599	1.02248	2.36834
$\alpha_2 = 0.00404007$	0.00238552	-0.000936049	-0.000936049
$\alpha_3 = 0.00311119$	0.0124583	-0.0228764	0.0290988
$\alpha_4 = 2.2532$	0.457979	1.29787	3.20853
$\alpha_5 = 0.216317$	0.627302	-1.09221	-1.09221
$\alpha_6 = -12.8318$	2.1288	-17.2724	-8.39119
T = 3.4724 Coeficiente de correlación = 0.869251 Error estándar de la estimación = 0.37706			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.65538$	0.326669	0.973958	2.3368
$\alpha_2 = 0.00427082$	0.00241562	-0.000768088	0.00930973
$\alpha_3 = 0.00601876$	0.0126155	-0.0202968	0.0323343
$\alpha_4 = 2.23209$	0.463758	1.26471	3.19948
$\alpha_5 = 0.228046$	0.635217	-1.097	1.55309
$\alpha_6 = -12.7517$	2.15567	-17.2483	-8.25501
T = 3.5743 Coeficiente de correlación = 0.861084 Error estándar de la estimación = 0.384476			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.62026$	0.333092	0.925443	2.31508
$\alpha_2 = 0.00427205$	0.00246312	-0.000865932	0.00941003
$\alpha_3 = 0.00708091$	0.0128635	-0.019752	0.0339138
$\alpha_4 = 2.10981$	0.472876	1.12341	3.09622
$\alpha_5 = 0.202942$	0.647707	-1.14815	1.55404
$\alpha_6 = -12.5057$	2.19805	-17.0908	-7.92066

Continúa tabla 5.11

T = 3.6761 Coeficiente de correlación = 0.85163 Error estándar de la estimación = 0.392784			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.61763$	0.341419	0.905439	2.32982
$\alpha_2 = 0.00426324$	0.00252469	-0.00100319	0.00952967
$\alpha_3 = 0.00797971$	0.0131851	-0.019524	0.0354834
$\alpha_4 = 2.01588$	0.484698	1.00481	3.02694
$\alpha_5 = 0.228402$	0.663899	-1.15647	1.61327
$\alpha_6 = -12.5293$	2.253	-17.229	-7.82963
T = 3.778 Coeficiente de correlación = 0.852715 Error estándar de la estimación = 0.39183			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.59698$	0.339465	0.888867	2.30509
$\alpha_2 = 0.0040428$	0.00251024	-0.00119348	0.00927908
$\alpha_3 = 0.00855973$	0.0131096	-0.0187866	-0.0187866
$\alpha_4 = 1.92301$	0.481923	0.917735	2.92829
$\alpha_5 = 0.168759$	0.660099	-1.20819	1.5457
$\alpha_6 = -12.2996$	2.2401	-16.9724	-7.62681
T = 3.8798 Coeficiente de correlación = 0.853612 Error estándar de la estimación = 0.39955			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.595$	0.340292	0.885164	2.30484
$\alpha_2 = 0.00393199$	0.00251636	-0.00131705	0.00918104
$\alpha_3 = 0.00837886$	0.0131416	-0.0190341	0.0357918
$\alpha_4 = 0.0357918$	0.483098	0.837635	2.85309
$\alpha_5 = 0.130002$	0.661707	-1.2503	1.5103
$\alpha_6 = -12.2476$	2.24556	-16.9318	-7.56342
T = 3.9816 Coeficiente de correlación = 0.853792 Error estándar de la estimación = 0.388454			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.63273$	0.33654	0.930717	2.33474
$\alpha_2 = 0.00364488$	0.00248861	-0.00154629	0.00883605
$\alpha_3 = 0.00603764$	0.0129967	-0.0210731	0.0331483
$\alpha_4 = 1.70536$	0.477771	0.708744	2.70198
$\alpha_5 = 0.160127$	0.654412	-1.20495	-1.20495
$\alpha_6 = -12.3614$	2.2208	-16.9939	-7.72886
T = 4.0835 Coeficiente de correlación = 0.853362 Error estándar de la estimación = 0.386882			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.66156$	0.336738	0.959137	2.36399
$\alpha_2 = 0.0034757$	0.00249008	-0.00171852	0.00866992
$\alpha_3 = 0.00431625$	0.0130043	-0.0228104	0.0314429
$\alpha_4 = 1.6606$	0.478052	0.663403	2.65781
$\alpha_5 = 0.20274$	0.654796	-1.16314	1.56862
$\alpha_6 = -12.5488$	2.22211	-17.184	-7.91354
T = 4.1853 Coeficiente de correlación = 0.834409 Error estándar de la estimación = 0.519568			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.58521$	0.318082	0.921706	2.24872
$\alpha_2 = 0.00370867$	0.00247916	-0.00146279	0.00888012
$\alpha_3 = 0.0101277$	0.01371	-0.0184709	0.0387263
$\alpha_4 = 1.3574$	0.472556	0.371662	2.34313
$\alpha_5 = 0.0335296$	0.623773	-1.26764	1.3347
$\alpha_6 = -11.9083$	2.1237	-16.3383	-7.47836
T = 4.2871 Coeficiente de correlación = 0.841951 Error estándar de la estimación = 0.394964			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.63739$	0.342165	0.923641	2.35113
$\alpha_2 = 0.00303292$	0.00253021	-0.00224501	0.00831086
$\alpha_3 = 0.0035921$	0.0132139	-0.0239717	0.0311559
$\alpha_4 = 1.36362$	0.485757	0.350348	2.37689
$\alpha_5 = 0.178245$	0.665349	-1.20965	1.56614
$\alpha_6 = -12.0882$	2.25792	-16.7981	-7.37823
T = 4.389 Coeficiente de correlación = 0.841038 Error estándar de la estimación = 0.39104			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.6057$	0.338781	0.899017	2.31239
$\alpha_2 = 0.00305125$	0.00250518	-0.00217448	0.00827698
$\alpha_3 = 0.00368275$	0.0130832	-0.0236084	0.0309739
$\alpha_4 = 1.19055$	0.480952	0.187303	2.1938
$\alpha_5 = 0.151382$	0.658769	-1.22279	1.52555
$\alpha_6 = -11.7698$	2.23559	-16.4332	-7.10644
T = 4.4908 Coeficiente de correlación = 0.837294 Error estándar de la estimación = 0.394742			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.60758$	0.341988	0.8942	2.32095
$\alpha_2 = 0.00289024$	0.0025289	-0.00238496	0.00816543
$\alpha_3 = 0.0029507$	0.0132071	-0.0245988	0.0305002
$\alpha_4 = 1.05207$	0.485505	0.0393263	2.06482
$\alpha_5 = 0.14357$	0.665004	-1.24361	1.53075
$\alpha_6 = -11.6452$	2.25675	-16.3527	-6.93766

Continúa tabla 5.11

T = 4.5927 Coeficiente de correlación = 0.836417 Error estándar de la estimación = 0.396441			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.63504$	0.34346	0.918594	2.35149
$\alpha_2 = 0.00272822$	0.00253979	-0.00256969	0.00802613
$\alpha_3 = 0.00802613$	0.0132639	-0.0260435	0.0292927
$\alpha_4 = 1.00206$	0.487595	-0.0150517	2.01916
$\alpha_5 = 0.18592$	0.667868	-1.20723	1.57907
$\alpha_6 = -11.7896$	2.26647	-16.5173	-7.06179
T = 4.6945 Coeficiente de correlación = 0.828628 Error estándar de la estimación = 0.521256			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.61576$	0.319115	0.950093	2.28142
$\alpha_2 = 0.00319116$	0.00248722	-0.00199709	0.00837941
$\alpha_3 = 0.00645777$	0.0137545	-0.0222338	0.0351493
$\alpha_4 = 0.879457$	0.474091	-0.109482	1.8684
$\alpha_5 = 0.170581$	0.6258	-1.13482	1.47598
$\alpha_6 = -11.8211$	2.1306	-16.2655	-7.37678
T = 4.7963 Coeficiente de correlación = 0.845859 Error estándar de la estimación = 0.38478			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.69248$	0.333358	0.997107	2.38785
$\alpha_2 = 0.00244699$	0.00246508	-0.00269509	0.00758907
$\alpha_3 = -0.00140924$	0.0128738	-0.0282636	0.0254451
$\alpha_4 = 0.937691$	0.473253	-0.0495002	1.92488
$\alpha_5 = 0.257725$	0.648223	-1.09445	-1.09445
$\alpha_6 = -12.1244$	2.1998	-16.7131	-7.5357
T = 4.8982 Coeficiente de correlación = 0.849167 Error estándar de la estimación = 0.381282			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.72001$	0.330327	1.03096	2.40906
$\alpha_2 = 0.0022677$	0.00244267	-0.00282763	0.00736303
$\alpha_3 = -0.00294515$	0.0127568	-0.0295553	0.023665
$\alpha_4 = 0.915717$	0.468951	-0.0624992	1.89393
$\alpha_5 = 0.273166$	0.64233	-1.06671	1.61305
$\alpha_6 = -12.2697$	2.1798	-16.8167	-7.72271
T = 5.00 Coeficiente de correlación = 0.852053 Error estándar de la estimación = 0.378671			
Parámetro	Error estándar	Intervalo de confianza 95 % inferior	Intervalo de confianza 95 % superior
$\alpha_1 = 1.74747$	0.328065	1.06313	2.4318
$\alpha_2 = 0.00203167$	0.00242594	-0.0030287	0.0070921
$\alpha_3 = -0.00472744$	0.0126694	-0.0311554	0.0217005
$\alpha_4 = 0.875719$	0.465739	-0.095798	1.84724
$\alpha_5 = 0.267748$	0.637931	-1.06296	1.59845
$\alpha_6 = -12.3684$	2.16487	-16.8842	-7.85252

Tabla 5.12 Desviación estándar de los residuos del conjunto de datos utilizados en la validación de la regresión lineal aplicada a eventos registrados en la estación sísmica CU de la Ciudad de México

T (s)	0.0100	0.11184	0.21367	0.31551	0.41735	0.51918	0.62102	0.72286	0.82469	0.92653
σ_{InSa} (cm/s ²)	0.248	0.258	0.233	0.180	0.362	0.309	0.291	0.340	0.443	0.337
T (s)	1.0284	1.1302	1.232	1.3339	1.4357	1.5376	1.6394	1.7412	1.8431	1.9449
σ_{InSa} (cm/s ²)	0.288	0.330	0.244	0.268	0.203	0.164	0.317	0.377	0.415	0.421
T (s)	2.0467	2.1486	2.2504	2.3522	2.4541	2.5559	2.6578	2.7596	2.8614	2.9633
σ_{InSa} (cm/s ²)	0.425	0.428	0.359	0.317	0.305	0.333	0.326	0.364	0.443	0.510
T (s)	3.0651	3.1669	3.2688	3.3706	3.4724	3.5743	3.6761	3.778	3.8798	3.9816
σ_{InSa} (cm/s ²)	0.488	0.496	0.482	0.449	0.418	0.381	0.344	0.313	0.279	0.215
T (s)	4.0835	4.1853	4.2871	4.389	4.4908	4.5927	4.6945	4.7963	4.8982	5.0000
σ_{InSa} (cm/s ²)	0.226	0.210	0.287	0.329	0.312	0.314	0.332	0.351	0.383	0.405

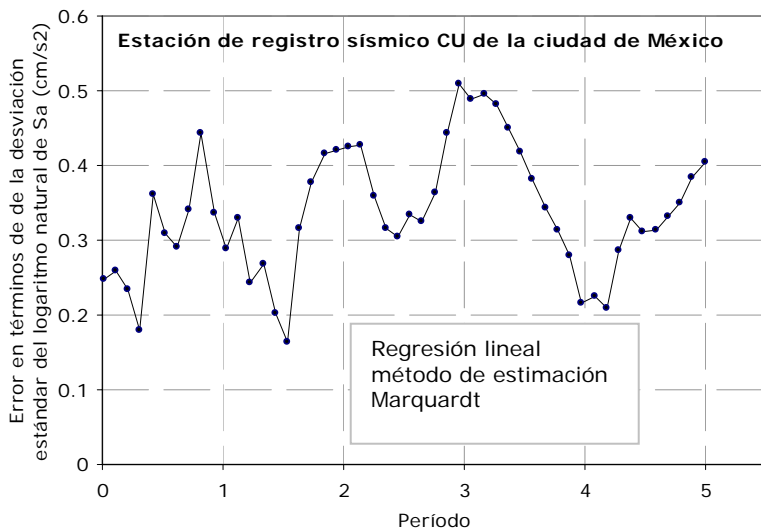


Figura 5.24 Desviación estándar de los residuos con el período para el junto de datos utilizados en la validación de la regresión lineal aplicada a eventos registrados en la estación sísmica CU de la Ciudad de México.

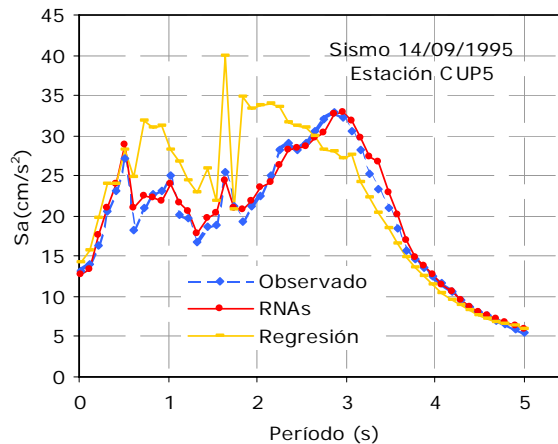
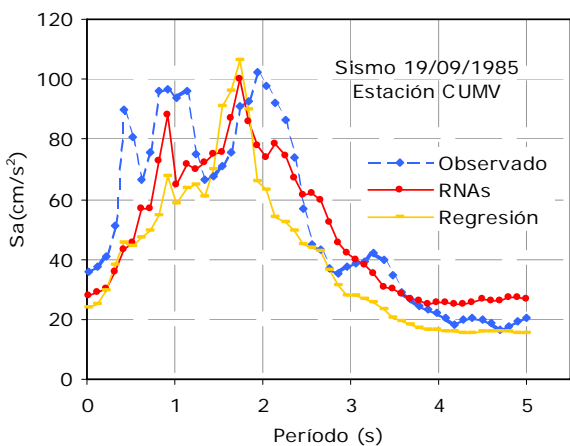
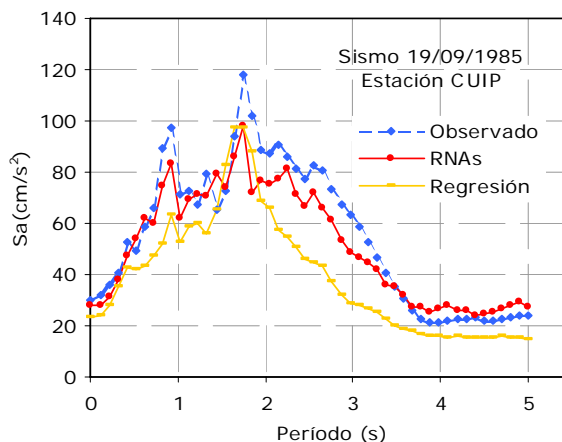
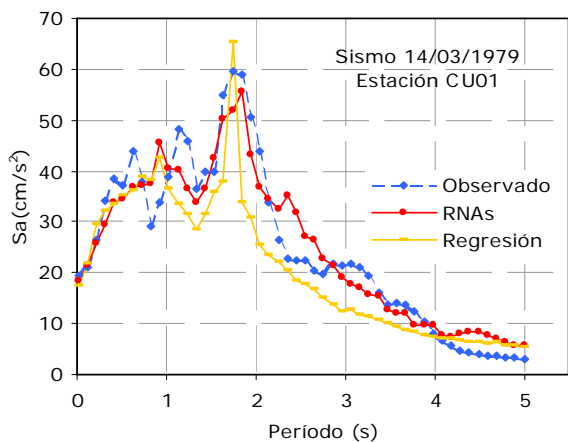


Figura 5.25 Comparación entre los espectros observados y los estimados (5% de amortiguamiento) con la ley de regresión de la ecuación 5.1, y con la RNA para la estación de registro sísmico CU.

5.5.2 Modelo neuronal para la estación de registro sísmico CU

Empleando los mismo los datos utilizados en el modelo paramétrico (tabla 5.10), se desarrolló el modelo neuronal con las características indicadas en la tabla 5.13, el cual permitió lograr el mejor índice de ejecución para los patrones de aprendizaje presentados. Los resultados en términos estadísticos se presentan anexos en la tabla 5.13, el peso individual de las variables de entrada asociadas al proceso en porcentaje se reporta en la figura 5.26, y los resultados en la figura 5.25. Estos resultados muestran las ventajas de las RNAs sobre los procedimientos convencionales utilizados en la ingeniería sísmica, pues es evidente que la mejor predicción la proporcionó la red neuronal artificial, superando la regresión para espacios de alta dimensión y sin imponer ninguna forma funcional de partida.

Tabla 5.13 Topología básica de la RNA y parámetros estadísticos para la estación CU

Característica	Espectros de respuesta elásticos			
Arquitectura	Cascada			
Celdas sensoriales	Cinco (M_w , h , R , α_l , f_p)			
Función de procesamiento	Máximo y mínimos			
Número de capas ocultas	Una			
Neuronas en capa oculta	Seis			
Algoritmo de aprendizaje	Propagación rápida QP			
Función de entrada	Producto punto			
Función de Transferencia	Sigmoide			
Criterio de error	Medio cuadrático			
Capa de salida	Ln (S_a)			
Modelo	Error en términos de la desviación estándar		Porcentaje de correlación	
	Aprendizaje	Prueba	Aprendizaje	Prueba
Espectros de respuesta elásticos estación de registro sísmico CU	0.153	0.217	0.974	0.967

Estación de registro sísmico CU - etapa de entrenamiento

Estación de registro sísmico CU - etapa de prueba

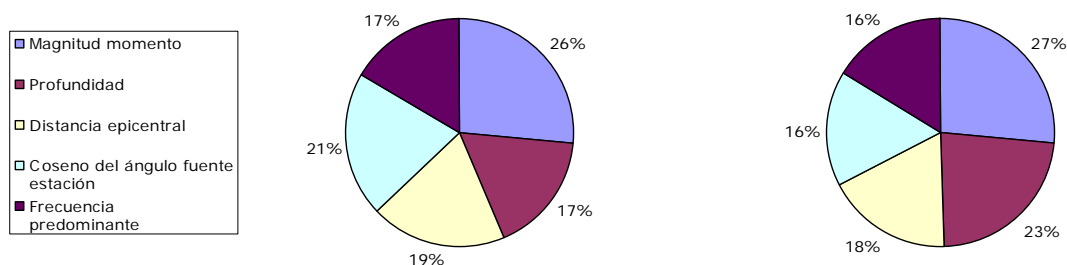


Figura 5.26 Contribución de las variables externas proporcionadas a la RNA con topología indicada en la Tabla 5.12 para la estación de registro sísmico CU de la Ciudad de México.

5.6 Simulación hipotética sobre la aceleración máxima

Considerando la siguiente ley de atenuación sobre la aceleración máxima (Ordaz, comunicación personal)

$$\text{Log } a_{max} = 1.76 + 0.3M_s - \log R - 0.0031R + \varepsilon \quad (5.2)$$

donde

M_s magnitud de ondas superficiales

R distancia mínima al área de ruptura en km

a_{max} aceleración máxima expresada en cm/s^2

ε variable aleatoria con distribución normal media cero y desviación estándar 0.3

Se simuló $a_i(M_s, R, E(\log a), \varepsilon)$. Los resultados de la simulación se reportan en la tabla 5.14. De los cuales los primeros 32 datos (resaltados en cursiva), se utilizaron para el cálculo de los coeficientes empíricos del funcional (ecuación (5.3)) y para entrenar la red neuronal, los datos restantes (496) se emplearon en ambos casos para evaluar la capacidad de predicción.

$$E(\log(a_i)) = 2.94465 + 0.189769M - \log R - 0.00429908R \quad (5.3)$$

Los resultados conseguidos en la simulación hipotética sobre la aceleración máxima a través de la ley de regresión y con la RNA para la población de validación muestran una solución semejante y una alta dispersión, como se ilustra en figura 5.27. Las características de la RNA con la que se alcanzó la mejor generalización para los patrones de aprendizaje presentados, el error estándar de la estimación en términos de la desviación estándar de los residuos y el coeficiente de correlación se indican en la tabla 5.15.

Este es el caso en que no se puede concluir que los modelos neuronales sobrepasan claramente en eficiencia a la técnica estadística. Sin embargo, más que enfrentar las técnicas comparadas, la intención es identificar el paralelismo entre estos procedimientos y su solución, pues de hecho a los sistemas neuronales también se les denomina técnicas de ajuste estadístico inspiradas en la biología.

Se han realizado estudios para identificar el símil entre los métodos estadísticos y los neuronales, por ejemplo, se equiparan: el perceptrón simple con nodo tipo umbral y el análisis discriminante; el perceptrón simple con nodo tipo sigmoide y la regresión logística; la ADALINE y la regresión lineal; el perceptrón multicapa y la regresión no lineal simple y multivariada; el aprendizaje hebbiano no supervisado y el análisis de componentes principales; la red simple de Kohonen y análisis cluster mínimos cuadrados; cuantificación de vectores LVQ y análisis discriminante; etc. No obstante, hay algunos modelos de RNAs para los que no se ha encontrado una técnica estadística a asemejar, como es el caso del modelo de Hopfield.

Tabla 5.15 Topología básica de la RNA y parámetros estadísticos

Característica		a_i	
Arquitectura		Multicapa con alimentación hacia adelante	
Celdas sensoriales		Dos (M_s , R_i)	
Función de procesamiento		Normalizada	
Número de capas ocultas		Una	
Neuronas en capa oculta		Tres	
Algoritmo de aprendizaje		Propagación rápida QP	
Función de entrada		Producto punto	
Función de Transferencia		Sigmoide	
Criterio de error		Medio cuatrático	
Capa de salida		$\text{Log}(a_i)$	
Error en términos de la desviación estándar		Porcentaje de correlación	
Aprendizaje	Prueba	Aprendizaje	Prueba
0.356	0.297	0.533	0.695

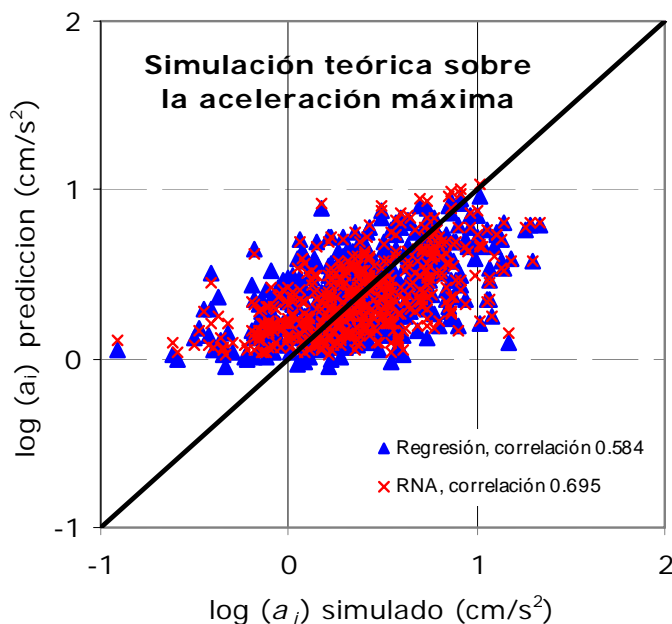


Figura 5.27 Resultados de la simulación hipotética sobre la aceleración máxima valores de predicción para ambos procedimientos

5.7 Incertidumbre

Los primeros pasos a verter en la especificación de los espectros de diseño para estructuras civiles, son requerimientos que se establecen a partir de la estimación del riesgo sísmico, el que a su vez depende en gran medida del conocimiento que se tenga sobre el potencial de las fuentes que generan los temblores y del estudio de la atenuación de las ondas sísmicas. La sismicidad local, entendida como el grado de actividad sísmica en un volumen de la corteza terrestre y que se cuantifica por la magnitud máxima que puede alcanzar un temblor, el número de temblores que ocurren en un intervalo de tiempo y la energía liberada; así como por las características propias del temblor (magnitud, tipo de falla, posición de su foco, etc.) y las intensidades en los sitios de interés, son

fenómenos que por su naturaleza (y mientras no se cuente con suficientes registros que permitan estimar y medir cada acción y su efecto) son inherentemente inciertos.

Según Somerville [98], se distinguen dos tipos de incertidumbres: la incertidumbre epistémica debida a la falta de conocimiento del proceso que liga la causa con el efecto o incertidumbre científica sobre el estado verdadero de la naturaleza, la cual puede reducirse al incrementar en cantidad la base de datos; y la incertidumbre aleatoria o paramétrica debida a la naturaleza propia del fenómeno en la que intervienen múltiples causas desconocidas. Para disminuir la incertidumbre aleatoria se requiere de vasto conocimiento de las variables, por esta razón interesa distinguir entre resultados más probables que otros.

En la modelación neuronal, se sabe que pocas variables de entradas restringen el espacio de búsqueda de parámetros y llevan a una pobre generalización, y que muchas variables implican una alta dimensión del espacio de búsqueda y conducen a un exceso de varianza en la arquitectura. En cualquiera de los dos casos se está en el dilema del sesgo frente a la varianza ya comentado en el capítulo 3. Pero queda la inquietud de cómo establecer los resultados más probables de las variables que integran el vector de entrada para sismos postulados. Este asunto podría abordarse con redes Bayesianas que son un modelo de dependencia probabilística representado mediante grafos construido a partir de sistemas expertos basados en probabilidad. Los grafos son herramientas muy útiles para definir sistemas expertos y otros modelos utilizados en el área de la inteligencia artificial. Una red Bayesiana es un par (D, P) , donde D es un grafo dirigido acíclico y P es un conjunto de n funciones de probabilidad condicionada, una para cada variable.

Capítulo 6

Análisis de sensibilidad sobre las variables sensoriales

Para compensar la ausencia de medidas estadísticas concerniente al significado de las variables que conforman el vector entrada (extraídas de los datos observados y de la información consignada en la literatura técnica), se realiza un análisis de sensibilidad sobre dichas variables.

En las figuras subsecuentes se muestra la representación gráfica de los pesos individuales en porcentaje de las características asociadas al proceso para los diferentes modelos neuronales que permitieron estimar con base en el criterio de convergencia elegido, los parámetros dinámicos incluidos en este estudio. Esta información es la cuantificación de las características importantes contenidas en el vector de entrada que extraen las neuronas artificiales escondidas de la red neuronal.

El grado de representatividad de cada variable dentro del modelo tiene una importancia relativa, que depende de los valores numéricos del colectivo y del número de variables involucradas, las cuales cambian con la complejidad del problema y cuyo número crece en mira de lograr una mejor representación de la física del fenómeno y ampliar el espacio de búsqueda de parámetros de la red neuronal.

En la figura 6.1 se indica la contribución de las variables de entrada en la neuro computación de la duración de la fase intensa y la aceleración máxima para el caso mexicano. En la margen izquierda etapa de aprendizaje y margen derecha fase de predicción. Respecto a la duración de fase intensa, en esta figura los pesos de los parámetros sensoriales guardan una alta coherencia con lo reportado en la literatura [Reinoso 01] en cuanto a la dependencia de dicho parámetro con respecto a la magnitud del sismo, la distancia, el perfil de suelo y el periodo dominante de éste. Contrariamente, en la estimación neuronal de la aceleración máxima, la magnitud no es una característica relevante para la red neuronal, quizás una explicación puede estar en la estrecha relación entre la magnitud del evento y la intensidad de Arias. Es de resaltar que las anotaciones expuestas también son observables para los casos ciudad de Medellín y Colombia ver figuras 6.5 y 6.7.

En las figuras 6.2, 6.3 y 6.4, se presenta la aportación de las variables de entrada en el cálculo neuronal de los espectros de respuesta elásticos para caso de la Ciudad de México, tanto en etapa de aprendizaje como en la fase predicción. En la modelación de la atenuación espectral para los tres grupos establecidos en este estudio, las variables relevantes corresponden a la aceleración máxima y a la intensidad de Arias, el resto de las variables tienen una distribución casi equitativa. No obstante, resalta el acumulado de los parámetros azimutales que caracterizan la geometría de la falla, el cual es de aproximadamente un 40%.

La contribución de las variables de entrada en la neuro computación para la duración fase intensa del sismo, aceleración máxima y período dominante para la Ciudad de Medellín, se representan en figura 6.5; y en la figura 6.6 los espectros de respuesta elásticos tanto para foco superficial como para foco lejano. Llama la atención la participación de la variable coseno del ángulo fuente-estación incluso para ambos casos (foco superficial, foco lejano), lo que induce a conjeturar sobre la alta incidencia de la geomorfología del Valle de Aburrá donde esta emplazada la ciudad en el registro sísmico.

En figura 6.7 se ilustra el peso de las variables externas para los modelos neuronales del caso colombiano en los que se estimaron la aceleración máxima y los espectros de respuesta elásticos. Se observa que los porcentajes de los pesos de las variables que conforman el vector de entrada, también son muy equitativos en la modelación de foco superficial y foco lejano. En ambos casos, las mayores contribuciones, reposan en la frecuencia dominante y los factores de atenuación en función

de la frecuencia dominante. Resultado que podría adjudicarse a los efectos que induce la variabilidad del subsuelo, ya que las estaciones sísmicas están distribuidas a lo largo y ancho del país sin que se conozca con exactitud las características geotécnicas de cada perfil.

Es de comentar que los efectos individuales de las variables de poca importancia con respecto al conjunto, tiene un rol significativo en la capacidad de generalización de las RNAs. Esto se compró corriendo la red neuronal sin considerar las variables externas de baja contribución.

También es importante anotar que ninguna de las variables muestra una tendencia definida a crecer o disminuir entre la etapa de aprendizaje y la etapa de ejecución para los casos estudiados. Esta estabilidad, se puede interpretar como una consecuencia del mapeo implementado por la red y su capacidad de generalización para los diversos rangos de valores de todas las entradas (patrones de aprendizaje y validación). Si bien en el presente, es difícil llevar a cabo un análisis de sensibilidad de las variables de entrada en términos cuantitativos, tal vez en un futuro, este examen cuantitativo podría emprenderse, explorado el espacio de la sinapsis y planteando el modelo inverso a través del concepto de memoria asociativa bidireccional y así obtener una información más inteligible sobre el comportamiento de las variables.

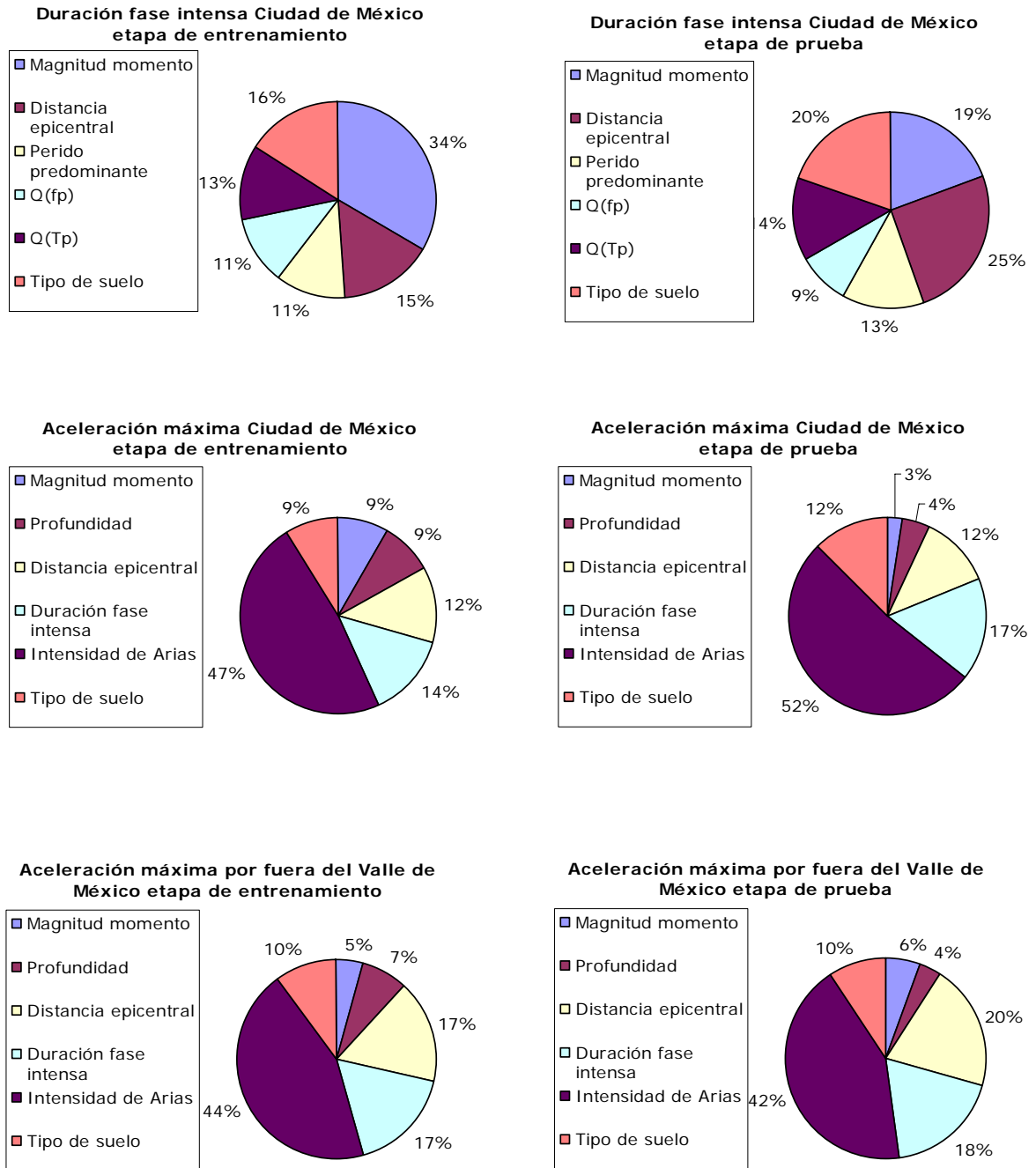
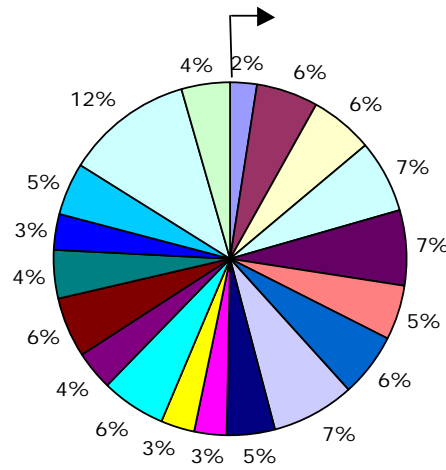


Figura 6.1 Contribución de las variables de entrada en la neuro computación de la duración de la fase intensa y la aceleración máxima para el caso mexicano. Margen izquierda etapa de aprendizaje y margen derecha fase de prueba.

Caso Ciudad de México grupo 1- etapa de entrenamiento

- Magnitud Momento
- Profundidad
- Azimut 1
- Buzamiento 1
- Dirección del movimiento 1
- Azimut 2
- Buzamiento 2
- Dirección del movimiento 2
- Distancia epicentral R
- Coseno del ángulo fuente y estación
- Frecuencia dominante
- Frecuencia de esquina
- Frecuencia de corte
- Velocidad máx/aceleración máx
- Factor de atenuación
- Duración fase intensa
- Intensidad de Arias
- Aceleración máxima
- Tipo de suelo



Caso Ciudad de México grupo 1- etapa de prueba

- Magnitud Momento
- Profundidad
- Azimut 1
- Buzamiento 1
- Dirección del movimiento 1
- Azimut 2
- Buzamiento 2
- Dirección del movimiento 2
- Distancia epicentral R
- Coseno del ángulo fuente y estación
- Frecuencia dominante
- Frecuencia de esquina
- Frecuencia de corte
- Velocidad máx/aceleración máx
- Factor de atenuación
- Duración fase intensa
- Intensidad de Arias
- Aceleración máxima
- Tipo de suelo

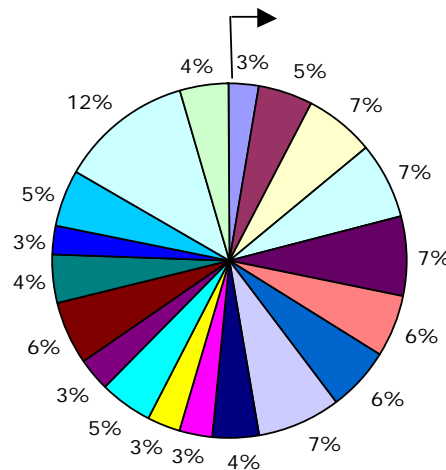
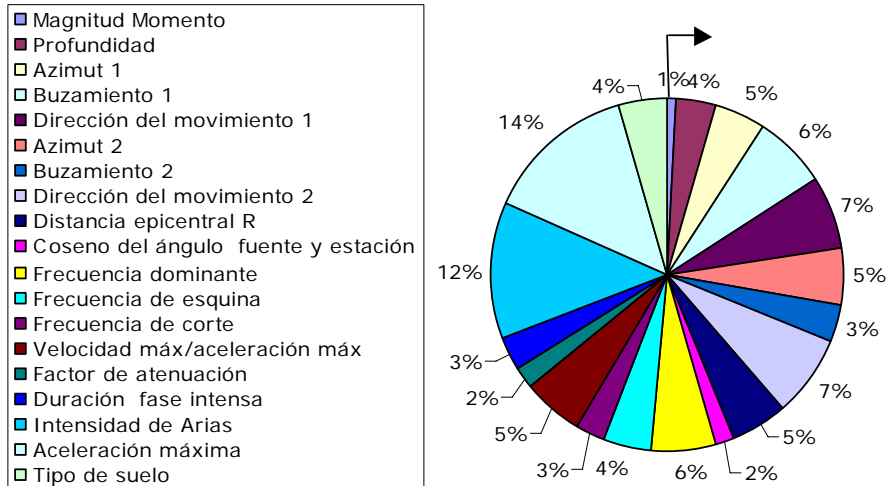


Figura 6.2 Contribución de las variables de entrada en la neuro computación de los espectros de respuesta elásticos. Parte superior etapa de aprendizaje, parte inferior en predicción. Caso Ciudad de México grupo 1.

Caso Ciudad de México grupo 2 - etapa de entrenamiento



Caso Ciudad de México grupo 2 - etapa de prueba

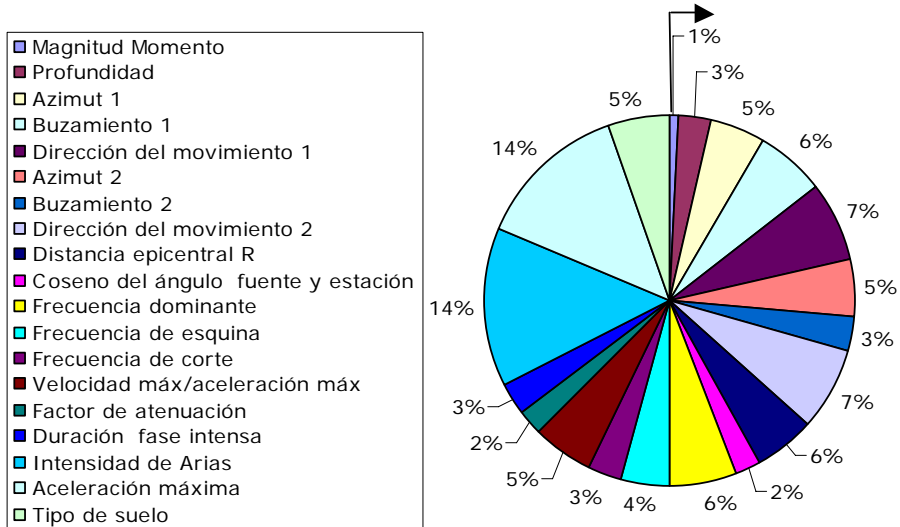
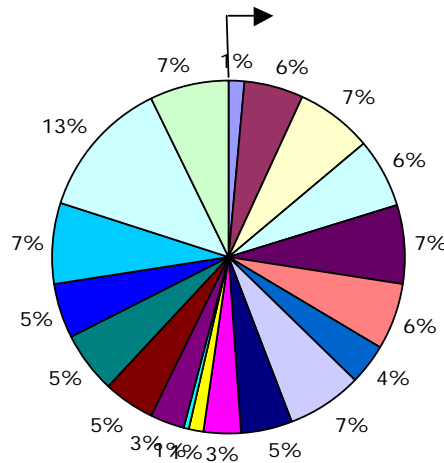
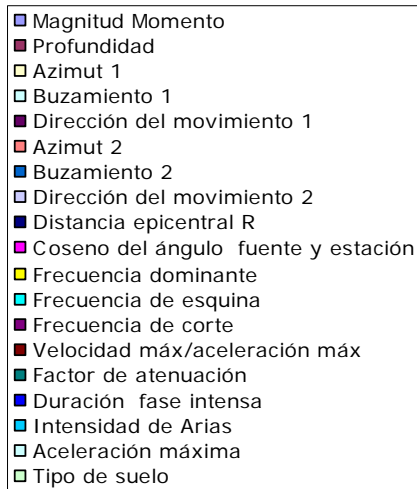


Figura 6.3 Contribución de las variables de entrada en la neuro computación de los espectros de respuesta elásticos. Parte superior etapa de aprendizaje, parte inferior en predicción. Caso Ciudad de México grupo 2.

Caso Ciudad de México grupo 3 - etapa de entrenamiento



Caso Ciudad de México grupo 3 - etapa de prueba

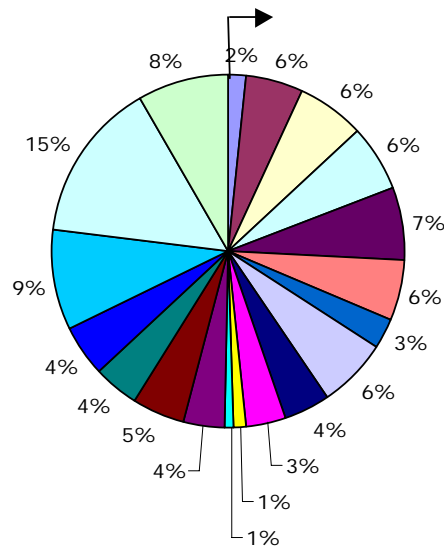
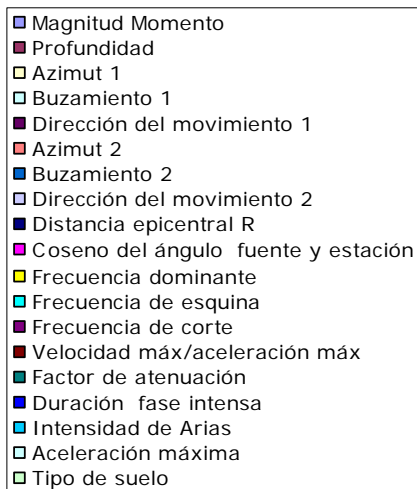


Figura 6.4 Contribución de las variables de entrada en la neuro computación de los espectros de respuesta elásticos. Parte superior etapa de aprendizaje, parte inferior en predicción. Caso Ciudad de México grupo 3.

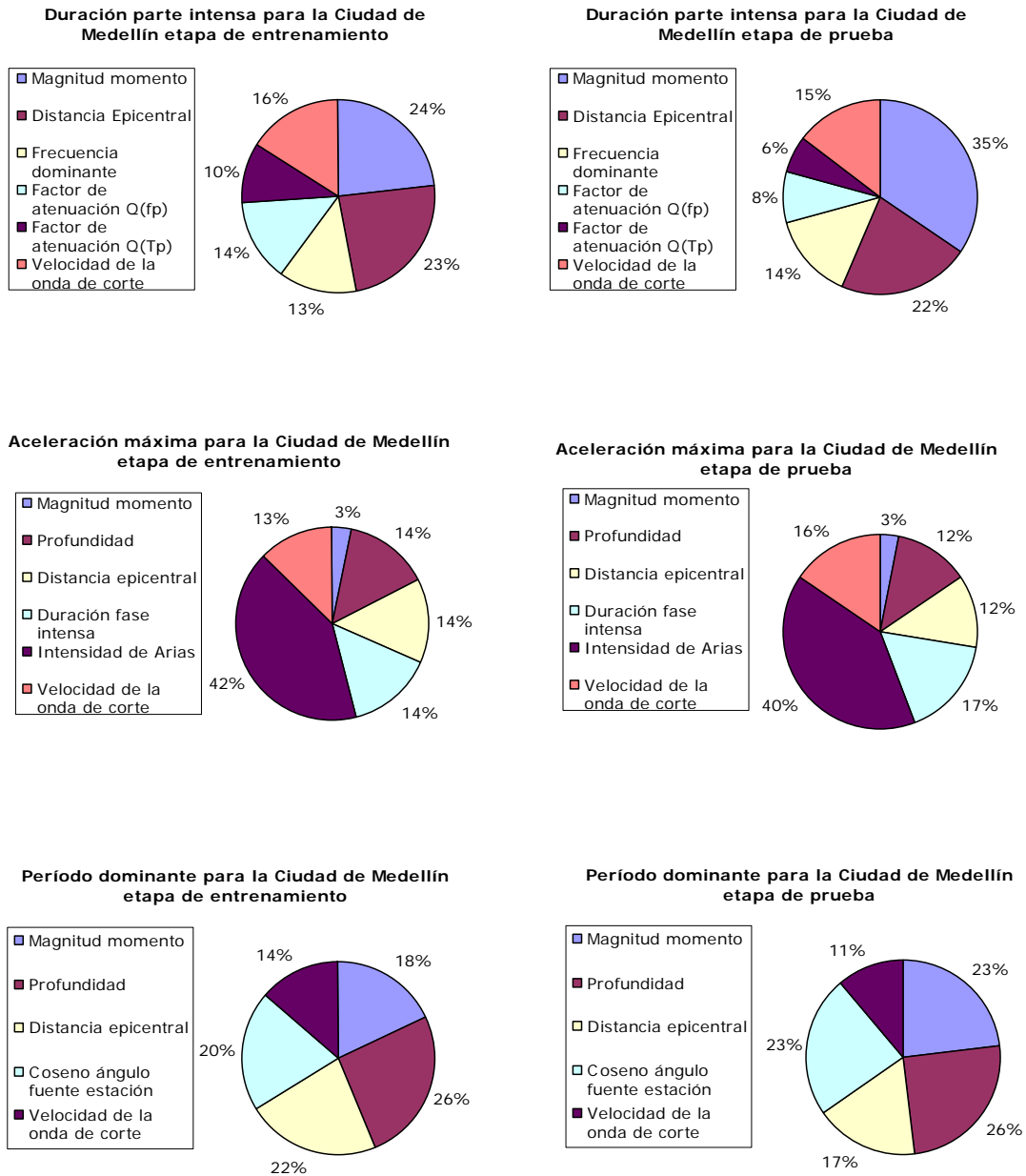


Figura 6.5 Contribución de las variables de entrada en la neuro computación para la duración parte intensa, aceleración máxima y período dominante para la Ciudad de Medellín. Margen izquierda etapa de aprendizaje, margen derecha fase de predicción.

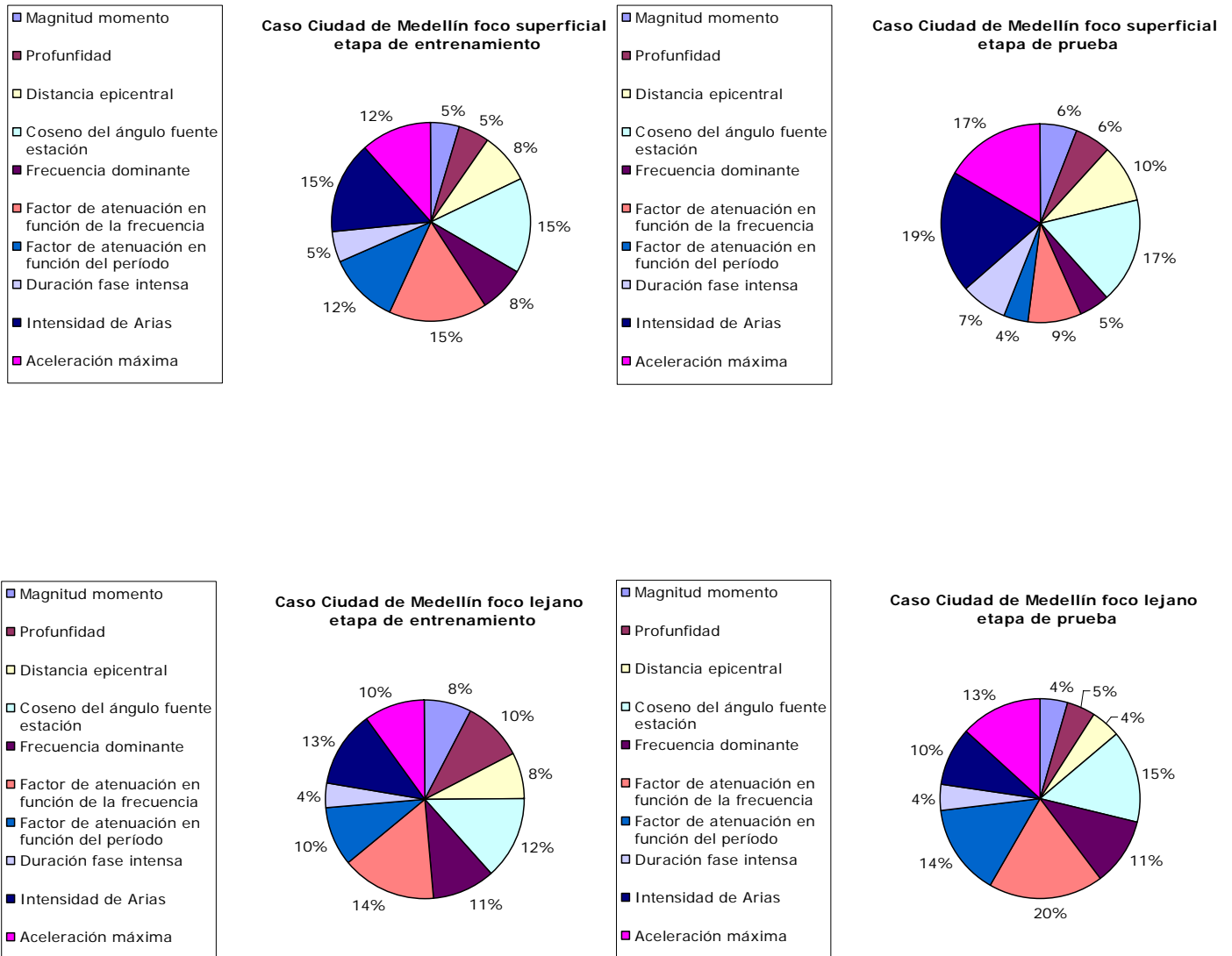


Figura 6.6 Contribución de las variables de entrada en la neuro computación de los espectros de respuesta elásticos para la Ciudad de Medellín. Margen izquierda etapa de aprendizaje, margen derecha fase de predicción.

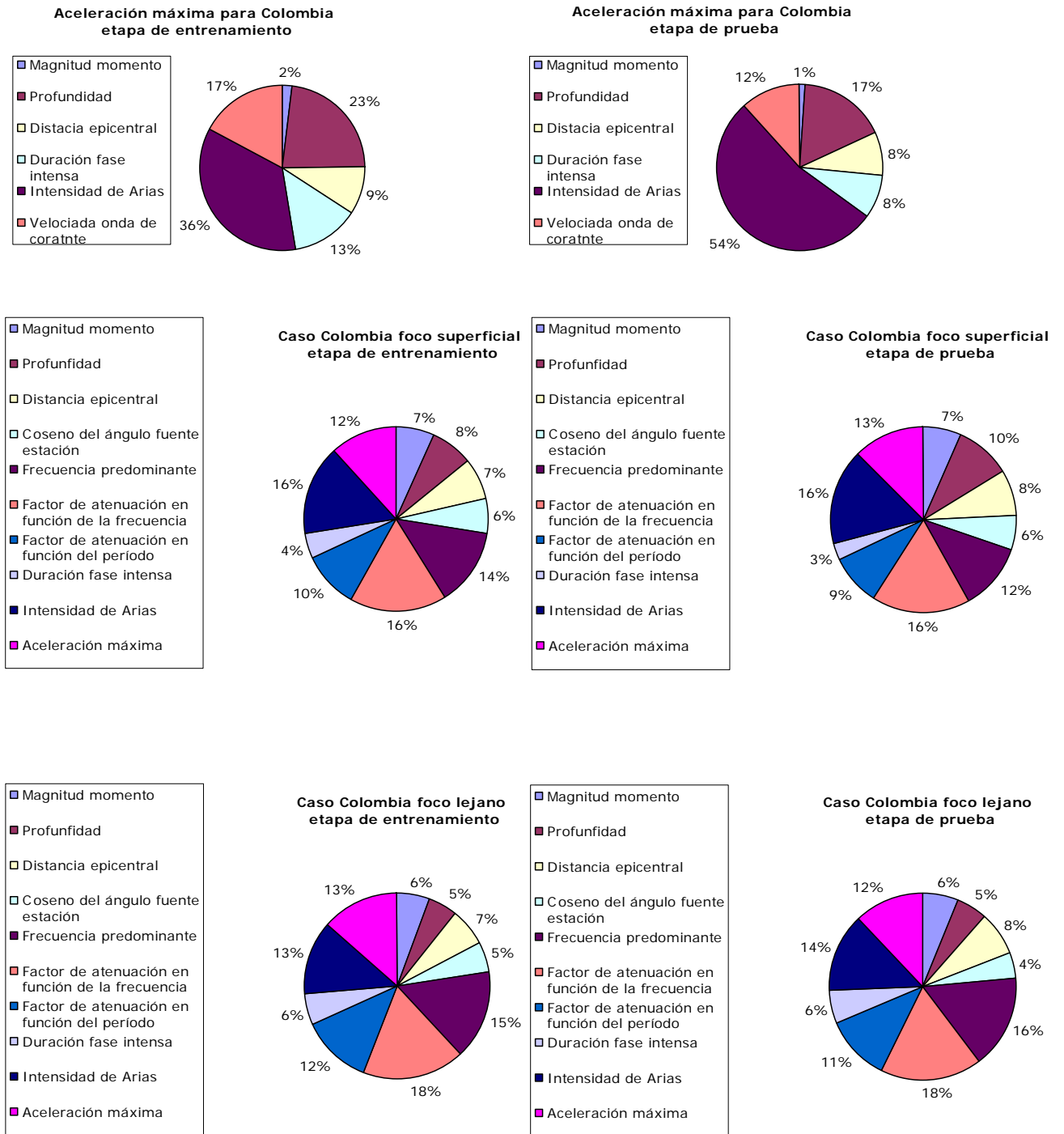


Figura 6.7 Contribución de las variables de entrada en la neuro computación para la aceleración máxima y los espectros de respuesta elásticos para el caso Colombiano. Margen izquierda etapa de aprendizaje, margen derecha fase de predicción.

Capítulo 7

Conclusiones y recomendaciones

- La respuesta superficial del suelo ante excitación sísmica es un fenómeno complejo gobernado por un gran número de variables que dificultan la modelación analítica del problema. Con el tiempo, la medición instrumental del fenómeno sísmico en superficie que contiene la configuración fuente–estación, ha proporcionado información valiosa en el desarrollo de procedimientos analíticos para la valoración cuantitativa del movimiento, lo que ha representado un importante avance en la comprensión y predicción de los efectos y su trascendencia. No obstante, se requiere contar con estudios más detallados pues es poco común que se cumpla con rigor las hipótesis en que se sustentan.
- En este estudio se presentan dos ejemplos de la aplicabilidad de las redes neuronales artificiales en la atenuación espectral sísmica para casos con ambientes sismo geotécnicos diferentes. En los casos estudiados (México y Colombia), el perceptrón multicapa como aproximador universal de funciones, entrenado con un algoritmo de aprendizaje de segundo orden y con un vector de entradas conformado por parámetros elementales que intervienen en la amplitud y forma de los espectros de respuesta elásticos y en la dimensión de las otras cantidades de interés, provee resultados dentro del margen de error admisible en la práctica ingenieril, pero con la limitación de que no proporciona información sobre la física del fenómeno y su ajuste, sujeto a la muestra disponible, es muy sensible a los casos particulares de pares entrada-salida seleccionados para el aprendizaje.
- Las RNAs constituyen un método de solución en el que hay que recurrir al procedimiento de prueba y error, con un gran esfuerzo computacional, ya que se desconoce la topología del sistema neuronal que podría proporcionar la solución satisfactoria. Sin embargo, una vez entrenada una red, la solución es de fácil implementación.
- El MLP realiza un tipo de regresión multidimensional no lineal sin necesidad de suponer una determinada forma funcional, contrario a los procedimientos convencionales de la ingeniería sísmica, como la regresión u otras técnicas para espacios de alta dimensión. Esto elimina la adopción de hipótesis difíciles de validar. Con base en los resultados obtenidos las mejores predicciones se alcanzan con los modelos neuronales artificiales usados, a partir de grandes bases de datos (más de 1000 acelerogramas en este caso).
- Aunque los sistemas neuronales artificiales no resuelven el problema con rigor matemático, las soluciones aproximadas que se consiguen son valiosas al no requerir ninguna función prescrita ni la independencia de las variables, y por obviar el problema de la no-linealidad en sistemas complejos. Esto representa una ventaja, los registros sísmicos confieren efectos no claramente separables.
- El conjunto de variables externas seleccionadas como representativas (tabla 5.1), el arreglo de estos parámetros de entrada a los modelos neuronales y la transformación algorítmica usada permiten estimar el movimiento fuerte del suelo en términos del espectro de respuesta con una similitud aceptable entre lo observado y lo estimado, para cada pareja fuente-sitio y sin tener que recurrir en el caso de terrenos blandos a funciones de transferencia ni al uso de cocientes espectrales ligados a algún punto de referencia. En consecuencia, se pueden emplear fructíferamente los modelos neuronales, pese a que resulta poco hacedero interpretar su funcionamiento.
- Debido a las incertidumbres inherentes al fenómeno, para mejorar la estimación obtenida con RNAs para México y Colombia, se recomienda el uso de un modelo híbrido que acople los

sistemas neuronales y los basados en lógica borrosa. Esto se puede lograr estableciendo funciones de membresía para cada variable externada de la RNA.

- Con el fin de poder realizar análisis cuantitativos sobre la sensibilidad de los parámetros que conforman el vector de entrada de las redes neuronales artificiales, se recomienda explorar el espacio de la sinapsis planteando el modelo inverso a través del concepto de memoria asociativa bidireccional.
- Dado que los espectros convencionales sólo recogen cierta información de la historia de respuesta, correspondiente a los picos de la misma e ignorando lo que sucede tras ella, se sugiere investigar la capacidad de los procedimientos neuronales para estimar espectros de energía, los cuales acopian toda la información de la historia de respuesta de un sistema de un grado de libertad. El modelo neuronal podría desarrollarse considerando las variables que influyen en dicho espectro con base en la relación que guarda el mismo con otras medidas de la acción sísmica, pues se ha observado que todos los valores que afectan el espectro de Fourier afectan igualmente al de energía [Sawada 92], por ejemplo las estrechas correlaciones que median entre los parámetros del espectro de potencia con la sencilla relación a_{\max}/V_{\max} , donde se ha identificado que la frecuencia correspondiente al pico del espectro de potencia está en relación inversa con el factor a_{\max}/V_{\max} .

Referencias

- [Abrahamson 92] Abrahamson, N. A., Youngs, R. R. (1992), A stable algorithm for regression analysis using the random effects model, *BSSA*, **82**, 505-510.
- [Abrahamson 97] Abrahamson, N. A., Shedlock K. (1997), Overview, *Seismological Research Letters*, **68**:1, 9-23.
- [Anderson 84] Anderson, J. G., Hough, S. E. (1984), A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies, *BSSA*, **74**, 1969-1993.
- [Anderson 97] Anderson, J. G. (1997), Nonparametric description of peak acceleration above a subduction thrust, *Seismological Research*, **68**: 1, 86-93.
- [Ambraseys 73] Ambraseys, N. N. (1973), Dynamics and response of foundation materials in epicentral regions of strong earthquakes, *Proc. World Conf. Earthquake Eng. 5th*.
- [Aki 80] Aki, K., Richards, P. G. (1980), *Quantitative Seismology*, W. H. Freeman
- [Aki 88] Aki, K. (1988), Local site effects on strong motion, *Earthquake Engineering and Soil Dynamics II, Recent Advances in Ground-Motion Evaluation*, Geotechnical Special, Publication N° 20, 103-155.
- [Ang 90] Ang, A. H. S. (1990), Reliability bases for seismic safety assessment and design, *Proceedings Fourth U.S. National Conference on Earthquake Engineering*, Earthquake Engineering Research Institute, Palm Spring, **1**, 29-45.
- [Barto 83] Barto, A. G., Sutton, R. S., Anderson, C. W. (1983), Neuron-like adaptive elements that can solve difficult learning control problems, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-13, 834-846.
- [Boatwright 82] Boatwright, J. (1982), A dynamic model for far-field acceleration, *BSSA*, **72**, 1049-1068.
- [Boatwright 88] Boatwright, J. (1988), The Seismic radiation from composite models of faulting, *BSSA*, **78**, 489-508.
- [Bolt 73] Bolt, B. A. (1973), Duration strong motion, *Proceeding of the 5th World Conference on Earthquake Engineering*, 292.
- [Bommer 96] Bommer, J. J., Martinez-Pereira, A. (1996), The Prediction of strong-motion duration for engineering design, *Proceedings 11WCEE*, Acapulco, Paper No. 84.
- [Bommer 99] Bommer, J. J., Martinez-Pereira, A. (1999), The effective duration of earthquake strong motion, *Journal of Earthquake Engineering*, **3**, 127-172.
- [Bommer 00] Bommer, J. J., Martinez-Pereira, A. (2000), Strong-motion parameters : definition, usefulness and predictability, *Proceedings 12WCEE*.
- [Boore 73] Boore, D. M. (1973), The effect of simple topography on seismic waves: implications for the accelerations recorded at Pacoima Dam, San Fernando Valley, California, *BSSA*, **63**, 1575-1598.
- [Boore 83] Boore, D. M. (1983), Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra, *BSSA*, **73**, 1865-1894.
- [Boore 84a] Boore, D. M., Boatwright, J. (1984a), Average body-wave radiation coefficient, *BSSA*, **74**, 1615-1621.
- [Boore 84b] Boore, D. M., Joyner, W. B. (1984b), A note on the use of random vibration theory to predict peak amplitudes of transient signals, *BSSA*, **74**, 2035-20-39.
- [Boore 84c] Boore, D. M., Joyner, W. B. (1984c), Ground motions and response spectra at soil sites from seismological models of radiated spectra, *Proceedings 8WCEE*, **2**, 457-464.
- [Boore 86] Boore, D. M., (1986), Short-period *P* and *S* wave radiation from large earthquake: implications for spectral scaling relations, *BSSA*, **76**, 43-64.
- [Boore 87] Boore, W. B., Atkinson, G. M. (1987), Stochastic prediction of ground motion and spectral response parameters at hard-rock sites in eastern North America, *BSSA*, **77**, 440-467.

- [Boore 89] Boore, W. B., Joyner W. B. (1989), The effect of directivity on the stress parameter determined from ground motion observations, *BSSA*, **79**, 1984-1988.
- [Broomhead 88] Broomhead, D., Lowe, D. (1988), Multivariable function interpolation and adaptive networks. *Complex Systems*, 2: 321-355.
- [Brune 70] Brune, J. N. (1970), Tectonic stresses and spectra of seismic waves from earthquake, *Journal Geophys Res*, **75**, 4995-5009.
- [Bullen 85] Bullen, K. E., Bolt, B. A. (1985), *An Introduction to the Theory of Seismology*, Cambridge University Press.
- [Bufaliza 84] Bufaliza, M. (1984), *Atenuación de intensidades sísmicas con la distancia en sismos mexicano*, tesis de maestría, Facultad de Ingeniería, UNAM.
- [Campbell 81] Campbell, K. W. (1981), Near source attenuation of peak horizontal acceleration, *BSSA*, **71**, 2039-2070.
- [Campbell 93] Campbell, K. W. (1993), Comparison of contemporary strong-motion attenuation relationships, *Proceedings of the International Workshop on Strong Motion Data*, 2, 49-82.
- [Campbell 94] Campbell, K. W., Bozorgnia, Y. (1994), Near-source attenuation of peak horizontal acceleration from worldwide accelerograms recorded from 1957 to 1993, *Proceedings, Fifth U.S. National Conference on Earthquake Engineering*, Earthquake Engineering Research Institute, Berkeley, California, Vol. 1, 283-292.
- [Castro 88] Castro, R., Singh, S. K., Mena, E. (1988), An empirical model to predict Fourier amplitude spectra of horizontal ground motion, *Earthquake Spectra*, **4**, 675-686.
- [Chen 93] Chen, D. S., Jain, R. C. (1993), *A robust back-propagation learning algorithm for function approximation*, en Artificial intelligence frontiers in statistics, Chapman & Hall, London. pp 217-240.
- [Clough 93] Clough, R. W., Penzien, J. (1993), *Dynamics of structures*, McGraw-Hill.
- [Chávez-García 95] Chávez-García, F. J., Ramos-Mrtínez, J., Romero-Jimenez, E. (1995), Surface wave dispersion analysis in Mexico City, *BSSA*, **85**, 1116-1126.
- [Denker 87] Denker, J. S., Schwartz, B., Solla, S., Howard, R., Jackel, L., Hopfield, J. (1987), Large automatic learning, rule extraction and generalization, *Complex Systems*, 1: 877-922.
- [Dobry 76] Dobry, R., Oweis, I., Urzua, A. (1976), Simplified procedures for estimating the fundamental period of soil profile, *BSSA*, **66**, 1293-1321
- [Donovan 73] Donovan, N. C. (1973). A Statistical evaluation of strong motion data Including the February 9 1971, San Fernando Earthquake, *Proceedings World Conference Earthquake Engineering 5th*.
- [Dowla 90] Dowla, F. U., Taylor, S. R., Anderson, R. W. (1990), Seismic discrimination with artificial network: preliminary results with regional spectra data, *BSSA*, **80**, 1346-1373.
- [Emami 96] Emami, S. M. R., Iwao, Y., Harada, T. (1996), A Method for Prediction of Peak Horizontal Acceleration by Artificial Neural Networks. *Proceeding of the Eleventh World Conference on Earthquake Engineering*, paper 1238. Elsevier Science Ltd., Rotterdam.
- [Esteva 64] Esteva, L., Rosenblueth, E. (1964), Espectros de temblores a distancias moderadas y grandes. *BSSA*, **2**, 1-18.
- [Esteva 74] Esteva, L. (1974), *Seismicity in seismic risk and engineering decisions*, C. Lomnitz and E. Rosenblueth eds. Elsevier Scientific Publishing Company, Amsterdam, 174-224.
- [Etayo 86] Etayo F. et al., (1986), Mapa de terrenos geológicos de Colombia, *publ. Esp. Ingeominas*, Bogotá, 14: 235 p.
- [Essenreite 96] Essenreite, R. (1996), *Geophysical Deconvolution and Inversion with Neural Network*, PhD Thesis, Geophysikalisches Institut Universität Karlsruhe, Karlsruhe.

- [Faccioli 91] Faccioli, E. (1991), Seismic amplification in the presence of geological and topographic irregularities, *Proceedings, 2nd international Conference on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics*, St. Louis Missouri, Vol. 2, 1779-1797.
- [Fahlman 89] Fahlman, S. E. (1989), Fast learning variations on back – propagation: An empirical study, *In Proceedings of the 1988 Connectionist Models Summer School*, D. Touretzky, G. Hinton and T. Sejnowski, (eds), San Mateo, CA: Morgan Kaufmann, 38-51.
- [Fahlman 91] Fahlman, S. E., Christian, L. (1991), *The cascade-correlation learning architecture*, CMU-CS-90-100, School of Computer Science, Carnegie Mellon University, Pittsburgh.
- [Finn 91] Finn, W. D. L., Ledbetter, R. H., Fleming, R. L., Templeton, A. E., Forrest, T. W., Stacy, S. T. (1991), Dam on liquefiable foundation: Safety assessment and remediation, *Proceeding, 17th International Conference on Large Dams*, Vienna, 531-553.
- [Freeman 92] Freeman, J., Skapura, D. M. (1992), *Neural networks: algorithms, applications, and programming techniques*, Addison-Wesley.
- [Funahashi 89] Funahashi, K. (1989), On the approximate realization of continuous mappings by neural networks, *Neural Networks*, Vol. 2, 183-192.
- [Gallego 00] Gallego, M. (2000), *Estimación de riesgo sísmico en la República de Colombia*, Tesis de Maestría, División de Estudios de Posgrado, Facultad de Ingeniería, Universidad Nacional Autónoma de México.
- [Geli 88] Geli, L., Bard, P. Y., Jullien, B. (1988), The effect of topography on earthquake ground motion: A review and new results, *BSSA*, **78**, 42-63.
- [Gish 93] Gish, H. (1993), *Maximum likelihood training of neural*, en *Artificial intelligence frontiers in statistics*, Chapman & Hall, London. pp 241-255
- [Graves 93] Graves, R. W. (1993), Modeling three-dimensional site response effects in Marina District, San Francisco California, *BSSA*, **83**, 1042-1063.
- [Grossberg 87] Grossberg, S. (1987), Competitive learning: From interactive activation to adaptive resonance theory, *Cognitive science*, 11: 23-63.
- [Guerrero 97] Guerrero, R. (1997), *Duración del movimiento durante sismos: implicaciones en la degradación estructural*, Tesis de Maestría, División de Estudios de Posgrado, Facultad de Ingeniería, Universidad Nacional Autónoma de México.
- [Hand 93] Hand, D. J., (1993), *Artificial intelligence frontiers in statistics*, Chapman & Hall, London. p 410.
- [Hanks 79] Hanks, T. C., Kanamori, H., (1979), A moment magnitude scale, *Journal of Geophysical Research*, **84**, 2348-2350.
- [Hanks 81] Hanks, T. C., McGuire, R. K. (1981), The character of high-frequency strong ground motion, *BSSA*, **71**, 2071-2095.
- [Hanks 82] Hanks, T. C. (1982), f_{max} , *BSSA*, **72**, 1867-1879.
- [Harvey 94] Harvey, R. (1994), *Neural network principles*, Prentice-Hall, New Jersey. p 196.
- [Haykin 99] Haykin, S. (1999), *Neural networks: a comprehensive foundation*, Prentice Hall, New Jersey. p 842.
- [Hebb 49] Hebb, D. (1949), *Organization of Behavior*, John Wiley & Sons, New York. p 842.
- [Hecht-Nielsen 90] Hecht-Nielsen, R. (1990), *Neurocomputing*, Addison-Wesley
- [Hinton 84] Hinton, G., Ackley, D., Sejnowski, T. (1984), *Boltzmann machines: Constraint satisfaction networks that learn*, Technical Report CMU-CS-84-119, Carnegie-Mellon University, Department of Computer Science.
- [Hopfield 82] Hopfield, J. (1982), Neural networks and physical systems with emergent collective computational abilities, *In Proceedings of the National Academy of Science*, vol. 81, 3088-3092.
- [Hornik 89] Hornik, K., Stinchcombe, M., White, H. (1989), Multiple layer feed-forward networks are universal approximators, *Neural Networks*, **2**, 359-366.

- [Huang 99] Huang, H. C., Chiu, H. C. (1999), Canyon topography effects on ground motion at Feitsui damsite, *Soil Dynamics and Earthquake Engineering*, **18**, 87-89.
- [Hurtado 96] Hurtado, J. E., Barbat, A. H., Canas, J. A. (1996), Simulación de acelerogramas compatibles con espectros de energía sísmica inducida, *Revista Internacional de métodos Numéricos para Cálculo y Diseño en Ingeniería*, **12-1**, 73-91.
- [Hurtado 01a] Hurtado, J. E., Londoño, J. M., Meza, M. A. (2001), On the applicability of neural networks for soil dynamics amplification analysis, *Soil Dynamics and Earthquake Engineering*, **21**, 919-631.
- [Hurtado 01b] Hurtado, J. E. (2001), Neural networks in Stochastic mechanics, *Archives of Computational Methods in Engineering*, State of the reviews, **8**, 303-342.
- [Idriss 90] Idriss, I. M. (1990), *Response of soil soft sites during earthquake*, in J. M. Duncan, de., Proceedings, H. Bolton Seed Memorial Symposium, BiTech Publishers, Vancouver, British Columbia, Vol. 2, 273-289.
- [Jibson 87] Jibson, R. (1987), Summary of research on the effects of topographic amplification of earthquake shaking on slope stability, Open-File Report 87-268, *U.S. Geological Survey*, California.
- [Joyner 81] Joyner, W. B., (1981), Peak horizontal acceleration and velocity from strong motion records including records from the 1979 Imperial Valley California, Earthquake, *BSSA*, **71**, 2011-2038.
- [Joyner 84] Joyner, W. B., (1984), A scaling law for the spectra of large earthquake, *BSSA*, **74**, 1167-1188.
- [Joyner 88] Joyner, W. B., Boore, D. M. (1988), Measurement, characterization, and prediction of strong ground motion, *State of the Art Report, Earthquake Engineering and Soil Dynamics II – Recent Advances in Ground-Motion Evaluation*, Geotechnical Special Publication No. 20, ASCE, 43-102.
- [Kanamori 77] Kanamori, H., (1977), The energy release in great earthquake, *Journal of Geophysical Research*, **82**, 2981-2987.
- [Kirkpatrick 84] Kirkpatrick, S., (1984), Optimization by simulated annealing: Quantitative Studies, *Journal of Statistical Physics*, **34**, 975-986.
- [Knopoff 64] Knopoff, L. (1964), "Q", *Review of Geophysics*, **2**, 625-660.
- [Kohonen 82] Kohonen, T. (1982), Self-organized formation of topologically correct feature maps, *Biological Cybernetics*, 43:59-69.
- [Kosko 92] Kosko, B. (1992), *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach To Machine Intelligence*, Prentice – Hall, New Jersey. p 449.
- [Kosko 93] Kosko, B., Isaka, S. (1993), Lógica borrosa, *Investigación y Ciencia*, 60-65.
- [Lapedes 87] Lapedes, A. S., Faber, R. M. (1987), *Nonlinear signal processing using neural networks: Prediction and system modeling*, Los Álamos, technical report LA-UR-87-2662.
- [Lay 95] Lay, T., Wallace, T. C. (1995), *Modern Global Seismology*, Academic Press.
- [Lin 96] Lin, C. T., Lee, C.S. G. (1996), *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*, Prentice Hall, New Jersey. p 797.
- [Lin 00] Lin, C. J., Ghaboussi, J. (2000), Recent progress on neural network based methodology for generating artificial earthquake accelerograms, *Proceedings 12WCEE*.
- [Lippmann 87] Lippmann, R. P. (1987), An introduction to computing with neural nets, artificial neural networks: concepts and theory, *IEEE Computer Society Press*, 13-31.
- [Lisboa 92] Lisboa, P. G. L. (1992), *Neural Networks: Current Applications*, Chapman and Hall, New York.
- [Luco 85] Luco, J. E. (1985), On strong ground motion estimates based on models of the radiated spectrum, *BSSA*, **75**, 641-650.
- [Martín 02] Martín, B. K., Sanz, A. (2002), *Redes Neuronales y Sistemas Difusos*, Alfaomega, Madrid.

- [McGuire 80] McGuire, R. K., Hanks. (1980), RMS Accelerations and spectral amplitudes of strong-ground motion during the San Fernando, California Earthquake, *BSSA*, **70**, 1907-1919.
- [McCulloch 43] McCulloch, J., Piitts, W. (1943), A logical calculus of the ideal immanent in nervous activity, *Bulletin of Mathematical Biophysics*, **7**: 115-133.
- [McVerry 93] McVerry, G. H., Dowrick, D. J., Sritharam, S., Cousins, W. J., Porritt, T. E. (1993), Attenuation of peak ground accelerations in New Zealand, *Proceeding of the International Workshop on Strong Motion Data*, **2**, 23-38.
- [Minsky 69] Mainisky, M., Papert, S. (1969), *Perceptrons: An introduction to computational geometry*, The MIP Press.
- [Newmark 71] Newmark, N. M., Rosenblueth, E. (1971), *Fundamentals of Earthquake Engineering*, Prentice hall.
- [Nishenko 87] Nishenko, S. P., Singh, S. K. (1987), Conditional Probability for the recurrence of large and great interplate earthquakes along the Mexican subduction zone, *BSSA*, **77**, 2095-3114.
- [Novikova 93] Novikova, E. I., Trifunac, M. D. (1993), Duration of strong motion: physical basis and empirical equation, *Report N° 93-02*, Department of Civil Engineering, University of Southern California.
- [Ojeda 97] Ojeda, C., Martinez, A. (1997), Modelo para la atenuación de la energía sísmica en Colombia a partir de sismos registrados por la Red Nacional de Acelerógrafos, *Asociación colombiana de Ingeniería Sísmica*, Boletín técnico N- 53
- [Ordaz 87] Ordaz, M., Reinoso, E. (1987), Uso de teoría de vibraciones aleatorias en la determinación de los espectros de diseño del Reglamento para las Construcciones del Distrito Federal, *Memorias VII Congreso Nacional de Ingeniería Sísmica*, Querétaro, México, A155-A167.
- [Ordaz 89] Ordaz, M., Jara, J. M., Singh, S. K. (1989), *Riesgo Sísmico y Espectros de Diseño en el Estado de Guerrero*, Reporte del Instituto de Ingeniería, UNAM y el Centro de Investigación Sísmica A.C. al Gobierno del Estado de Guerrero.
- [Ordaz 92] Ordaz, M. (1992), *Estimación de Movimientos Fuertes ante Temblores Mexicanos de Subducción*, Tesis Doctoral, División de Estudios de Posgrado, Facultad de Ingeniería, Universidad Nacional Autónoma de México.
- [Ordaz 93] Ordaz, M., Sing, S. K. (1993), Source spectral and spectral attenuation of seismic waves from Mexican earthquakes, and evidence of amplification in the hill zone of Mexico City, *BSSA*, **82**, 24-43.
- [Ordaz 94] Ordaz, M., Sing, S.K., Arciniega, A. (1994), Bayesian attenuation regressions: an application to Mexico City, *Geophys. J. Int.*, **117**, 335-344
- [Ovando 03] Ovando, E., Romo, M. P., Contreras, N., Giralt, A. (2003), Effects on soil properties of future settlements in downtown Mexico City due to ground water extraction, *Geofísica Internacional*, Vol. 42, Num. 2, pp. 185-204.
- [Papageorgiou 83] Papageorgiou, A. S., Aki, K. (1983), A specific barrier model for the quantitative description of inhomogeneous faulting and the prediction of strong ground motion, *BSSA*, **73**, 693-722 y 953-978.
- [Reinoso 97] Reinoso, E., Wrobel, L. C., Power, H. (1997), Three-dimensional scattering of seismic waves from topographical structures, *Soil Dynamics and Earthquake Engineering*, **16**, 41-61.
- [Reinoso 01] Reinoso, E., Ordaz, M. (2001), Duration of strong ground motion during Mexican earthquakes in terms of magnitude, distance to the rupture area and dominant site period, *Earthquake Engng Struct. Dyn.* **30**:653-673.
- [Reyes 99] Reyes, J. C., (1999), *El estado límite de servicio en el diseño sísmico de edificios*, Tesis Doctoral, División de Estudios de Posgrado, Facultad de Ingeniería, Universidad Nacional Autónoma de México.
- [Ritter 91] Ritter, H., Martinez, T., Schulten, K. (1991), *Neural Computation and Self-Organizing Maps*, Addison-Wesley.
- [Romo 86a] Romo, M. P., Jaime, A. (1986a), *Características dinámicas de las arcillas del valle de México y análisis de respuesta sísmica del suelo*, Reporte Técnico, Instituto de Ingeniería, UNAM.

- [Romo 86b] Romo, M. P., Seed, H. B. (1986b), Analytical Modelling of Dynamics Soil Response in the Mexico Earthquake of September 19, 1985, *Proc. International Conference of the Mexico Earthquake-1985: Factors Involved and Lessons Learned*, Mexico, pp 148-162.
- [Romo 95] Romo, M. P. (1995), Clay Behavior, Ground Response and Soil-Structure Interaction Studies in Mexico City, state of art paper, *III International Conference on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics*, St. Louis Missouri, USA, April.
- [Romo 00] Romo, M. P. (2000), *Earthquake Geotechnical Engineering and artificial neural networks*, 4th Arthur Casagrande Lecture, SMMS, Cuaderno Técnico 1.
- [Romo 01] Romo, M. P., García, S. R., Mendoza, M. J., Taboada, V. (2001), Recurrent and constructive-algorithm networks for sand behavior modeling, *The International Journal of Geomechanics*, **1**, 371-387.
- [Rosenblatt 57] Rosenblatt, F. (1957), *The perceptron: A perceiving and recognizing automation*, Technical Report 85-460-1, Cornell Aeronautical Laboratory.
- [Rosenblueth 92] Rosenblueth, E., Arciniega, A. (1992), Response spectral ratios, *Earthquake Engineering and Structural Dynamics*, **21**, 483-492.
- [Ruiz 77] Ruiz, S. (1977), Influencia de las condiciones locales en las características de los sismos. Instituto de Ingeniería 387, Universidad Nacional Autónoma de México.
- [Rumelhart 86a] Rumelhart, D. E., McClelland, J. L. (1986a), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1, Cambridge, MA: MIT Press.
- [Rumelhart 86b] Rumelhart, D. E., Hinton, G., Williams, R. (1986b), Learning representations by back-propagating errors, *Nature*, 322:533-536.
- [Safak 98] Safak, E. (1998), 3D response spectral: a method to include in response spectra, *Proceedings 11ECEE*, Paris.
- [Sánchez- Sesma 93] Sánchez-Sesma, F., Campillo, M. (1993), Topografic effects for incident P, SV, and Rayleigh waves, *Tectonophysics*, **218**, 1-3, 113-125.
- [Sarmiento 01] Sarmiento, N. (2001), *Evaluación de la respuesta sísmica en el valle de México aplicando redes neuronales artificiales*, Tesis de Maestría, División de Estudios de Posgrado, Facultad de Ingeniería, Universidad Nacional Autónoma de México.
- [Sawada 92] Sawada, T., Hirao, K., Tsujihara, O., Yamamoto, H. (1992), Relationship between Maximum amplitude Ratio (a/v , ad/v^2) and Spectral Parameters of Earthquake Ground Motion, *Proceeding of the Tenth World Conference on Earthquake Engineerig*, Vol 2, 617-622, Madrid.
- [Seed 69] Seed, H. B., Idriss, I. (1969), The influence of soil conditions on ground motion during earthquake, *Journal of Soil Mechanics and Foundations Division*, ASCE, 120-137.
- [Seed 74] Seed, H. B., Ugas, C., Lysmer, J., (1974), *Site-dependent spectra for Earthquake-Resistant Desing*, Report. N° EERC 74-12, November.
- [Seed 87] Seed, H. B., Romo, M. P., Sun, Jaime, A., Lysmer, J. (1987), *Relationships Between Soil Conditions and Earthquake Ground Motions in Mexico City in the Earthquake of September 19, 1985*, Report No. UCB/EER-87/15, October.
- [Shepherd 97] Shepherd, A. J. (1997), *Second-Order Methods for Neural Networks: fast and reliable training methods for multi-layer perceptrons*, Springer, 144 p.
- [Silva 88] Silva, W. J. (1988), *Soil response to earthquake ground motion*, EPRI Report NP- 5747, Electric Power Research Institute, California.
- [Singh 82] Singh, S. K., Apsel, R. J., Fried, J., Brune, J. N. (1982), Spectral attenuation of SH waves along the Imperial Fault, *BSSA*, **72**, 2003-2016.
- [Singh 87] Singh, S. K., Mena, E., Castro, C. (1987), Empirical prediction of ground motion in Mexico City from coastal earthquake, *BSSA*, **77**, 1862-1867.
- [Singh 88] Singh, S. K., Lermo, J., Domínguez, T., Ordaz, M., Espinosa, J. M., Mena, E., Quaas, R. (1988), A study of amplification of seismic waves in the Valley of Mexico with respect to a hill zone site (CU), *Earthquake Spectra*, **4**, 653-673.

- [Singh 89] Singh, S. K., Ordaz, M., Anderson, J. G., Rodríguez, M., Quaas, R., Mena, E., Ottaviani, M., Almora, D., (1989), Analysis of near-source strong motion recordings along the Mexican subduction zone, *BSSA*, **79**, 1697-1717.
- [Singh 90] Singh, S. K., Mena, E., Anderson, J. G., Quaas, R., Lermo, J. (1990), Source spectral and RMS acceleration of Mexican Subduction zone earthquakes, *PAGEOPH*, **133**, 447-474.
- [Singh 93] Singh, S. K., Ordaz, M. (1993), On the origin of long coda observed in the lake-bed strong-motion records of Mexico City, *BSSA*, **83**, 1298-1306.
- [Somerville 98] Somerville, P. (1998), Emerging art: Earthquake ground motion, *Geotechnical Earthquake Engineering and Soil Dynamics III*, Proceedings of a Specialty Conference, Geotechnical Special Publication N° 75, ASCE, 1, 43-102.
- [Trifunac 75] Trifunac, M. D., Brady A. G. (1975), A study of the duration of strong earthquake ground motion, *BSSA*, **65**, 581-626.
- [Trifunac 76] Trifunac, M. D. (1976), Preliminary analysis of the peaks of strong earthquake ground motion-dependence of peaks, on earthquake magnitude, epicentral distance, and recording site condition, *BSSA*, **66**, 89-219.
- [Trifunac 82] Trifunac, M. D., Westermo, B.D. (1982), Duration of strong earthquake shaking. *Soil Dynamics and Earthquake Engineering*, **1**, 117-121.
- [Trifunac 95] Trifunac, M. D., Novikova, E. I. (1995), Duration of earthquake fault motion in California. *Earthquake Engineering and Structural Dynamics*, **24**, 781-799.
- [Universidad EAFIT 99] Universidad EAFIT., Integral S. A., INGEOMINAS., Facultad Nacional de Minas. (1999), Instrumentación y Microzonificación Sísmica del Área Urbana de Medellín, Publicación del Municipio de Medellín, Secretaria Privada, 135p.
- [Vemuri 88] Vemuri, V. (1988), Artificial neural networks: An introduction, en Artificial neural networks: Theoretical Concepts, *IEEE Computer Society Press*, 1-12.
- [Welstead 94] Welstead, S. T. (1994), *Neural Network and Fuzzy Logic Application in C/C++*, John Wiley & Sons Ink., p 494.
- [Widrow 60] Widrow, B. (1960), *An adaptative "adaline" neuron using chemical "memistors"*, Technical Report 1553-2, Stanford Electronics Laboratory.