

0163  
7



UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO

---

---

DIVISIÓN DE ESTUDIOS DE POSGRADO  
DE LA FACULTAD DE INGENIERÍA

EVALUACIÓN DE LA RESPUESTA SÍSMICA  
EN EL VALLE DE MÉXICO APLICANDO  
REDES NEURONALES ARTIFICIALES

NEFTALÍ SARMIENTO SOLANO

T E S I S

PRESENTADA A LA DIVISIÓN DE  
ESTUDIOS DE POSGRADO DE LA  
FACULTAD DE INGENIERÍA  
DE LA

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
COMO REQUISITO PARA OBTENER EL GRADO DE  
**MAESTRO EN INGENIERÍA**  
**(MECÁNICA DE SUELOS)**

DIRECTOR DE TESIS DR MIGUEL P ROMO ORGANISTA

0163



MÉXICO D.F.,

ABRIL 2001



Universidad Nacional  
Autónoma de México



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

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

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

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

---

## AGRADECIMIENTOS

---

Agradezco al Dr. Miguel P. Romo por su atinada dirección durante el desarrollo de este trabajo de tesis, por su paciencia en la revisión del texto y por compartir sus conocimientos académicos.

Le doy las gracias a cada uno de los sinodales: Dr. Efraín Ovando S , Dr Francisco J. Chávez G. Dr Víctor Taboada U., y M. I. Silvia R. Benítez G ; por los atinados consejos en las correcciones de esta tesis.

Le agradezco al Instituto de Ingeniería de la Universidad Nacional Autónoma de México, por la beca recibida durante el curso de materias y al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el patrocinio (bajo el proyecto 33032-U) para la realización de este trabajo

También agradezco a los profesores de la División de Estudios de Posgrado de la Facultad de Ingeniería, compañeros de clases, amigos y personal del área de geotecnia del Instituto de Ingeniería, que de alguna manera contribuyeron a la realización de esta tesis.

---

# ÍNDICE

---

	<i>págs</i>
<b>CAPÍTULO 1. INTRODUCCIÓN</b>	1
1.1 GENERALIDADES	1
1.2 ANTECEDENTES	2
1.3 OBJETIVO	2
1.4 ALCANCES	3
1.5 RESUMEN	3
<b>CAPÍTULO 2. PRINCIPIOS FUNDAMENTALES DE LAS REDES NEURONALES ARTIFICIALES</b>	4
2.1 GENERALIDADES	4
2.1.1 Red Neuronal Biológica	4
2.1.2 Red Neuronal Artificial	5
2.1.3 Tipos de Redes Neuronales Artificiales	5
2.2 MODELO DE REDES NEURONALES	6
2.2.1 Modelo de una Neurona	7
2.2.2 Funciones de Activación	8
2.2.3 Arquitectura de Redes	9
2.3 SUPERFICIES DE EJECUCIÓN Y OPTIMIZACIÓN	11
2.3.1 Características de las Superficies de Error	12
2.3.2 Métodos Clásicos de Optimización	14
2.3.2.1 Algoritmo del descenso más pronunciado	14
2.3.2.2 Método de Newton	16
2.3.2.3 Gradiente conjugado	17
2.4 ENTRENAMIENTO DEL PERCEPTRÓN MULTI-CAPA (PMC)	17
2.4.1 Arquitectura del Perceptrón Multi-Capa	18
2.4.2 Regla de Aprendizaje del PMC	18
2.4.3 Superficies de Error del PMC	19
2.4.4 Retropropagación	20
2.4.4.1 Algoritmo básico	20

2.5	VARIACIONES DEL ALGORITMO DE RETROPROPAGACIÓN	22
2.5.1	Modificaciones Heurísticas	23
2.5.1.1	Aplicación de momentum	23
2.5.1.2	Coefficiente de aprendizaje variable	24
2.5.2	Técnicas de Optimización Numéricas	25
2.5.2.1	Gradiente conjugado	25
2.5.2.2	Algoritmo Levenberg-Marquardt	27
<b>CAPÍTULO 3. RESPUESTA SÍSMICA DEL SITIO</b>		<b>30</b>
3.1	INTRODUCCIÓN	30
3.1.1	Efectos de la Topografía de Superficie y la Geometría de la Cuenca	31
3.2	EFFECTO DE LAS CONDICIONES LOCALES DEL SUELO EN LA RESPUESTA DEL TERRENO	33
3.2.1	Análisis Teóricos de Respuesta Sísmica	33
3.2.1.1	Análisis lineales de respuesta	34
3.2.1.2	Variaciones de los análisis de respuesta sísmica	37
3.2.2	Funciones de Amplificación Medidas	41
3.2.3	Movimientos de Superficie Medidos	41
3.2.4	Resumen del Efecto de las Condiciones Locales del Suelo en los Movimientos Sísmicos del Terreno	44
3.3	EFFECTO DE LA CONSOLIDACIÓN REGIONAL EN LA RESPUESTA SÍSMICA DE LOS DEPÓSITOS ARCILLOSOS DE LA CIUDAD DE MÉXICO	48
<b>CAPÍTULO 4. MODELO DE RESPUESTA SÍSMICA EN EL VALLE DE MÉXICO MEDIANTE REDES NEURONALES ARTIFICIALES</b>		<b>52</b>
4.1	DESCRIPCIÓN DEL MODELO	52
4.2	BASE DE DATOS DEL MODELO DE RESPUESTA SÍSMICA CON RNA	53
4.2.1	Zonificación Geotécnica y Sísmica del Valle de México	53
4.2.2	Movimientos Sísmicos del Terreno y Características de los Sitios empleados en el Modelo	56
4.2.2.1	Excitación sísmica	56
4.2.2.2	Caracterización de los depósitos de suelo	57
4.2.2.3	Movimientos de respuesta medidos en superficie	59
4.2.2.4	Variación de los movimientos sísmicos de un depósito de suelo	61
4.3	DATOS PARA LA ETAPA DE ENTRENAMIENTO Y PRUEBA DE LA RNA	62

<b>CAPÍTULO 5. RESULTADOS</b>	64
5.1 RESPUESTA SÍSMICA CON RNA. ENTRENAMIENTO CON DATOS MEDIDOS	64
5.1.1 Módulo LAGO	65
5.1.2 Módulo TRANS	70
5.2 ENTRENAMIENTO DEL MODELO DE RNA CON DATOS GENERADOS	71
5.3 COMPARACIÓN CON MODELOS TEÓRICOS DE RESPUESTA SÍSMICA	76
<b>CAPÍTULO 6. CONCLUSIONES</b>	79
<b>BIBLIOGRAFÍA Y REFERENCIAS</b>	82

---

# CAPÍTULO 1

## INTRODUCCIÓN

---

### 1.1 GENERALIDADES

Desde los comienzos de la ingeniería sísmica geotécnica se ha reconocido la influencia que tienen las condiciones locales del sitio en las características de los movimientos sísmicos del terreno, principalmente en aquellas regiones de alta sismicidad donde existen depósitos estratificados de suelos blandos con espesores considerables. Tal es el caso de la cuenca del valle de México, donde se ha demostrado que los suelos de esta localidad amplifican significativamente los movimientos sísmicos de los depósitos profundos (terreno firme) subyacentes. A partir de los eventos sísmicos de septiembre de 1985, estos efectos se apreciaron notablemente presentándose amplificaciones hasta cinco veces las aceleraciones máximas del terreno y ocho veces las amplitudes espectrales, con cambios significativos en los contenidos de frecuencias. La existencia de estaciones acelerométricas cuando ocurrieron tales eventos, permitió confirmar lo anterior, de ahí la importancia que tiene la captura de la información relacionada con los fenómenos físicos.

Las Redes Neuronales Artificiales (RNA) son técnicas computacionales inspiradas en la forma como operan las neuronas del cerebro humano. Estos sistemas están integrados por una gran cantidad de unidades de procesamiento, llamadas neuronas, que están altamente interconectadas unas con otras, por lo general en paralelo, y que procesan la información hasta aprender de los datos presentados. Estos modelos de procesamiento y aprendizaje neuronal constituyen una fuerte herramienta para resolver problemas que son difíciles para los procesos de computación convencionales; por lo cual, su aplicación ha cubierto varios campos de la ingeniería (control, sistematización, procesamiento de datos, clasificación, identificación, etc.). El punto de partida para aplicar esta técnica es disponer de una base de datos representativa del problema que se trata modelar.

## 1.2 ANTECEDENTES

Desde los inicios de la sismología, la respuesta sísmica del terreno fue uno de los temas principales de investigación (Reid, 1910; Gutenberg, 1957; Seed, Ugas y Lysmer, 1976), y desde entonces ya se conocía que las condiciones locales del sitio modifican los movimientos sísmicos incidentes.

En México son bien conocidos estos efectos desde mediados del siglo pasado (Rosenblueth, 1952). Los suelos blandos que se encuentran en el valle de la ciudad amplifican notablemente los movimientos sísmicos registrados en terreno firme. En estudios subsiguientes (Herrera y Rosenblueth, 1965; Romo y Jaime, 1986, 1987; Romo y Seed, 1986) se concluyó nuevamente que los depósitos de arcilla afectan la intensidad y el contenido de frecuencias de las ondas incidentes en la cuenca. También, se ha demostrado que las características de los movimientos del terreno en la zona geotécnica de Lago son grandemente controladas por las características de los depósitos de arcilla, por lo cual, muchos de los aspectos de la respuesta sísmica de estos depósitos pueden reproducirse con procedimientos de propagación de ondas cortantes unidimensionales simples (Romo y Jaime, 1986; Romo y Seed, 1986), y desde el sismo de septiembre de 1985 las características de los movimientos del terreno pueden correlacionarse con las propiedades dinámicas y espesores del suelo arcilloso, mostrando una clara correlación entre intensidad de daño sísmico y espesores de los depósitos de arcilla.

Después de los eventos sísmicos de 1985, varios proyectos de investigación se iniciaron para comprender la naturaleza de los movimientos sísmicos del terreno dentro de la cuenca de México. Esto incluyó la instalación de instrumentos de medición de movimientos fuertes (estaciones acelerométricas) a través de la ciudad. Muchos de estos aparatos se colocaron en la superficie del terreno, dentro de los depósitos de arcilla en arreglos verticales (pozos) y algunos otros en edificios. Todo este sistema de monitoreo ha integrado lo que se conoce como la Red Acelerométrica de la Ciudad de México (RACM) lo que ha permitido medir y almacenar información muy valiosa. Estos datos han ayudado al desarrollo de varias investigaciones sísmicas geotécnicas (comportamiento de suelo, propagación de ondas, comportamiento sísmico de edificios, interacción dinámica suelo-estructura, etc.)

De acuerdo con la disponibilidad de esta gran cantidad de información, los procedimientos basados en el aprendizaje, como es el caso de las RNA, resultan ser técnicas alternativas para evaluar con detalle el efecto que tienen las condiciones locales del suelo en los movimientos sísmicos de terreno firme dentro del valle de México. Esta herramienta es capaz de capturar y registrar el conocimiento del fenómeno relacionado directamente de la información medida en campo. Es decir, cuando las RNA se ajustan a la información disponible, puesto que son estimadores universales, capturan todas las leyes de la mecánica que el fenómeno de respuesta sísmica del terreno obedece. Por lo cual, esta modelación aproximada es una alternativa de interés sobre los modelos basados totalmente en técnicas matemáticas.

## 1.3 OBJETIVO

El objetivo principal de este trabajo consiste en aplicar la técnica de Redes Neuronales Artificiales para reproducir la respuesta sísmica de los depósitos de suelos localizados en las zonas geotécnicas de Lago y Transición dentro del valle de México. Este modelo neuronal considera como datos de alimentación, los espectros de respuesta de aceleración medidos en terreno firme y las características o propiedades del sitio en particular; para obtener como salida de la red, los espectros de respuesta de aceleración en la superficie del terreno.

También, parte de este trabajo se enfoca a determinar el efecto que tiene el fenómeno de concordación regional en la respuesta sísmica de los depósitos arcillosos de la ciudad de México.

## 1.4 ALCANCES

Para lograr el propósito principal de este trabajo es necesario abarcar los siguientes aspectos

- Integración de la base de datos para el modelo con las aceleraciones espectrales calculadas a partir de los movimientos medidos en la superficie de los sitios seleccionados. Los datos que caracterizan los depósitos de suelo se determinan de estudios realizados en campo y laboratorio.
- Análisis de los parámetros que integran la base de datos (variables representativas del fenómeno de respuesta sísmica) para la representación más adecuada de ellos en el modelo neuronal.
- Definición de la topología de red para modelar el fenómeno de respuesta sísmica. Utilizar durante el entrenamiento y prueba del modelo neuronal únicamente Redes multi-capas de propagación hacia adelante y propagación del error hacia atrás
- Comparación de los resultados del modelo neuronal con los valores medidos en campo.

## 1.5 RESUMEN

En el segundo capítulo de este trabajo se describe básicamente en que consiste la técnica de Redes Neuronales Artificiales (RNA). Se explica la estructura del modelo a partir de la forma como operan las redes neuronales biológicas del cerebro humano. También se mencionan los aspectos matemáticos que integran los algoritmos de ejecución de esta técnica (procesamiento y aprendizaje)

En el siguiente capítulo se mencionan los factores principales que afectan la respuesta sísmica del terreno. Se describen algunas de las técnicas más usadas para determinar la respuesta sísmica de los depósitos de suelo y el efecto que tienen las condiciones locales del sitio en los movimientos de campo libre dentro del valle de México

En el cuarto capítulo se presentan las bases que integran el modelo de respuesta sísmica con RNA. Se describe el modelo neuronal y las variables representativas del fenómeno de respuesta sísmica. Se indica la información que integrará la base de datos para el entrenamiento y generalización del modelo

En el quinto capítulo se muestran los resultados obtenidos con el modelo neuronal de respuesta sísmica, describiendo primeramente la topología de la RNA utilizada. Los resultados presentados se obtienen usando Redes Neuronales Multi-Capa con propagación hacia adelante con un algoritmo de entrenamiento con retropropagación. Se presentan los resultados del modelo neuronal para sitios ubicados en la zona de Lago y Transición durante las etapas de entrenamiento y predicción. También, se compara la predicción de la técnica de RNA con otros modelos de respuesta sísmica.

En el último capítulo se presentan las conclusiones del trabajo, donde se indica la capacidad de las RNA para reproducir los movimientos sísmicos dentro del valle de México (zona de Lago y Transición). Así también, se mencionan las limitaciones, ventajas y nuevas opciones de esta herramienta que podrían mejorar los resultados obtenidos con el modelo neuronal de respuesta sísmica presentado en este trabajo.

---

## CAPITULO 2

---

### PRINCIPIOS FUNDAMENTALES DE LAS REDES NEURONALES ARTIFICIALES

---

En este capítulo se describe básicamente en que consiste la técnica de Redes Neuronales Artificiales (RNA). Se comienza por explicar la estructura del modelo a partir de su similitud con las redes neuronales biológicas. Después se mencionan cuales son los aspectos matemáticos que integran sus algoritmos de ejecución (procesamiento y aprendizaje).

#### 2.1 GENERALIDADES

Las RNA son sistemas inspirados en la forma como trabaja el cerebro humano. El desarrollo de las teorías modernas acerca del aprendizaje y procesamiento neuronal artificial se iniciaron en los años cuarenta con el trabajo elaborado por Warren McCulloch y Walter Pitts (1943), cuando ya se comprendía en rasgos generales la forma de operar de las redes neuronales biológicas. Aunque por una década este tipo de investigación languideció, con la publicación de los artículos de Marvin Minsky y Seymour Papert (1969), estas teorías tuvieron un nuevo auge durante los años ochenta cuando aparecieron nuevas computadoras de gran capacidad. A partir de ese momento, las computadoras digitales han sido utilizadas como herramientas para modelar tanto neuronas individuales, así como agrupaciones de neuronas en forma de redes. Durante los años siguientes, se introdujeron nuevos conceptos importantes, tales como el uso de la mecánica estadística en el funcionamiento de una cierta clase de red (Hopfield, 1982), y el desarrollo del algoritmo de retropropagación para el entrenamiento de una red perceptrón multi-capas (Rumelhart y McClelland, 1986).

En la actualidad, se aprecia un notable crecimiento en la investigación y desarrollo de las redes neuronales artificiales. Muchos de estos avances se han relacionado con nuevos conceptos tales como arquitecturas sofisticadas y algoritmos de aprendizaje o técnicas de entrenamiento más eficientes. Sin olvidar que el avance computacional ha sido un factor importante para aplicar estos conceptos.

##### 2.1.1 Red Neuronal Biológica

Las neuronas biológicas que conforman el cerebro están formadas básicamente por las siguientes partes: las dendritas, el soma, el axón y la sinapsis (ver fig. 2.1a). Las dendritas son ramificaciones de fibras nerviosas muy delgadas que reciben la información de otras neuronas; el soma es el cuerpo de la célula que recibe y procesa la información entrante, el axón es una sola fibra larga que transporta la información del soma a otras neuronas; y la sinapsis es el punto de contacto entre el axón de una célula y la dendrita de otra célula, en donde se desarrollan procesos químicos complejos que establecen la función de la red neuronal.

### 2.1.2 Red Neuronal Artificial

En la fig. 2.1b se presenta el esquema de una RNA, la cual se define como un conjunto de unidades de procesamiento simples llamadas neuronas, nodos o células y que operan en paralelo. Estas neuronas se organizan en capas y se conectan entre sí, ya sea en una misma capa o en capas adyacentes. La resistencia de cada conexión se expresa con un valor numérico llamado peso ( $w$ ), el cual usualmente se modifica en la etapa de entrenamiento o aprendizaje, que es el proceso iterativo de retroalimentación de la red neuronal.

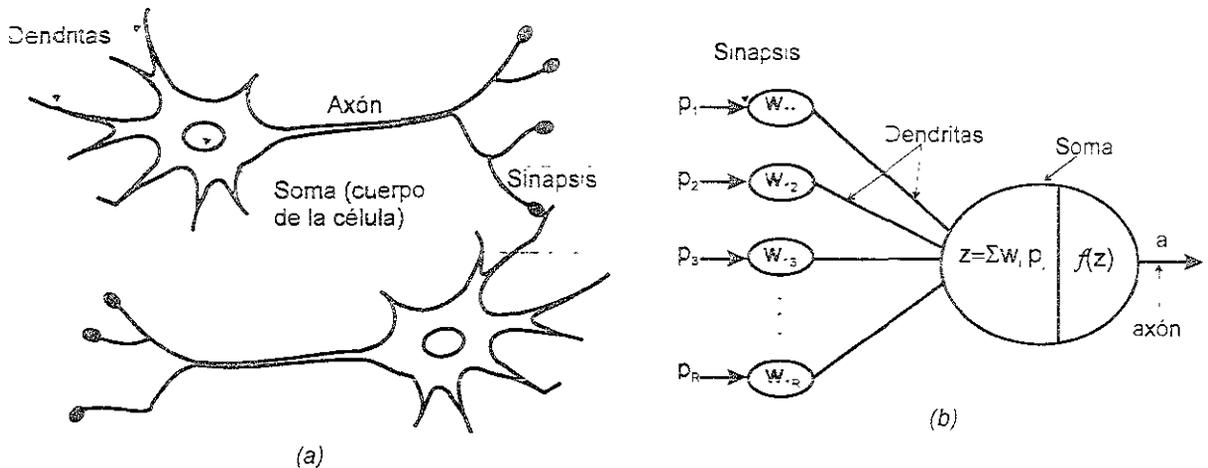


Fig 2.1 (a) Red Neuronal Biológica y (b) Red Neuronal Artificial

De cualquier forma, las RNA son modelos burdos de las biológicas, muy simples y que de alguna manera no se compararían siquiera con la complejidad del cerebro humano, pero que pueden en cierta medida realizar funciones muy útiles. Aún así, existen dos similitudes importantes entre las redes neuronales biológicas y artificiales (ver fig. 2.1): la primera consiste en que la unidad básica de ambas redes son dispositivos computacionales simples, que están altamente interconectados. La segunda es que las conexiones entre neuronas determinan el funcionamiento de la red.

En la actualidad, esta técnica de RNA es de gran utilidad para resolver problemas que son difíciles para los procesos de computación convencionales. Sus aplicaciones han aumentado grandemente no sólo en ingeniería, ciencia y matemática, sino también en medicina, finanzas y literatura. Las redes neuronales se entrenan para calcular funciones complejas en varios campos de aplicación incluyendo *reconocimiento de patrones, identificación, clasificación, reconocimiento de voz, sistemas de visión y control*.

### 2.1.3 Tipos de Redes Neuronales Artificiales

Las RNA se clasifican básicamente por su arquitectura (topología de sus conexiones sinápticas) y del tipo de regla de aprendizaje que emplean. Se tienen en general dos tipos de aprendizaje: supervisado y no supervisado. En el *aprendizaje supervisado*, se indica a la red cómo debe ejecutarse (aprendizaje de reforzamiento) o cuál es el correcto comportamiento (aprendizaje totalmente supervisado). En el *aprendizaje no supervisado*, la red es autónoma, aprende a partir de la regla de aprendizaje seleccionada, junto con los datos individuales de entrenamiento. Estos procesos de entrenamiento de la red se definen más ampliamente en las secciones siguientes.

Tabla 2.1 Tipos de Redes Neuronales Artificiales más comunes

Modo de Aprendizaje	Redes de retroalimentación	Redes de alimentación hacia adelante
I. No supervisado	<ul style="list-style-type: none"> <li>* Grossberg Adaptativo (AG)</li> <li>* Grossberg Desviado (SG)</li> <li>* Teoría de Resonancia Adaptable binaria (ART1)</li> <li>* Teoría de Resonancia Adaptable analógica (ART2, ART2a)</li> <li>* Hopfield discreto (DH)</li> <li>* Hopfield continuo (CH)</li> <li>* Memoria asociativa bidireccional discreta (BAM)</li> <li>* Memoria asociativa temporal (TAM)</li> <li>* Memoria asociativa bidireccional adaptable (ABAM)</li> <li>* Kohonen Mapa organizando por sí mismo / Mapa de topología conservativa (SOM/TPM)</li> <li>* Aprendizaje competitivo</li> </ul>	<ul style="list-style-type: none"> <li>* Matriz de Aprendizaje (LM)</li> <li>* Conducción - Aprendizaje de reforzamiento (DR)</li> <li>* Memoria Asociativa lineal (LAM)</li> <li>* Memoria Asociativa Lineal Óptima (OLAM)</li> <li>* Memoria Asociativa Distribuida Esparsificada (SDM)</li> <li>* Memoria Asociativa Borrosa (FAM)</li> <li>* Contrapropagación (CPN)</li> </ul>
II. Supervisado	<ul style="list-style-type: none"> <li>* Mapa Congrositivo borroso (FCM)</li> <li>* Máquina de Boltzmann (BM)</li> <li>* Correlación de Cascada Recurrente (RCC)</li> <li>* Retropropagación a través del tiempo (BPTT)</li> <li>* Aprendizaje recurrente de tiempo real (RTRL)</li> <li>* Filtro de Kalman Extendido Recurrente (EKF)</li> </ul>	<ul style="list-style-type: none"> <li>* Perceptrón</li> <li>* Adaline, Madaline</li> <li>* Retropropagación (BP)</li> <li>* Máquina de Cauchy (CM)</li> <li>* Crítico Heurístico Adaptable (AHC)</li> <li>* Red Neuronal con Retraso de Tiempo (TDNN)</li> <li>* Multa Premiada Asociativa (ARP)</li> <li>* Filtro Competitivo de Avalancha (AMF)</li> <li>* Retropercepción (Perc)</li> <li>* Artmap</li> <li>* Red de Lógica Adaptable (ALN)</li> <li>* Correlación de Cascada (CasCor)</li> <li>* Filtro Kalman Extendido (EKF)</li> <li>* Vector de Aprendizaje de Cuantificación (LVQ)</li> <li>* Red Neuronal Probabilística (PNN)</li> <li>* Red Neuronal de Regresión General (GRNN)</li> </ul>

Todos estos métodos de aprendizaje se relacionan estrechamente con un cierta clase de topología de red. Una red es de alimentación hacia adelante si sus conexiones forman ciclos sinápticos no cerrados, y una red es de retroalimentación si tienen conexiones que se alimentan hacia atrás (ciclos cerrados) En la Tabla 2.1 se muestra una lista en la que se incluyen los tipos de RNA más comunes.

## 2.2 MODELO DE REDES NEURONALES

Un modelo de red neuronal se define a partir de sus partes físicas, junto con sus algoritmos o reglas de ejecución. Se ha mencionado, que una RNA consiste físicamente de unidades de procesamiento llamadas neuronas y de conexiones entre ellas, las cuales definen su topología o arquitectura. Sin embargo, en cualquier modelo donde se procesa información, es necesario también definir un algoritmo de procesamiento, que en este caso es ejecutado entre las neuronas y es conocido como regla de aprendizaje.

### 2.2.1 Modelo de una Neurona

Una neurona es la unidad de procesamiento primaria de una red neuronal, en la fig. 2.2a se muestra una neurona de una sola entrada y en la fig. 2.2b una neurona con múltiples entradas. La neurona recibe una o varias entradas mediante conexiones denominadas sinapsis, cuya fuerza está representada por el peso sináptico  $w$ . Estas entradas  $p_j$  son las activaciones de los nodos de entrada (datos de alimentación) que se multiplican por los pesos sinápticos  $w_{1j}$ . Dentro de la neurona se tiene una suma real ponderada:

$$z = \sum (w_{1j} p_j) \quad (2.1)$$

Esta sumatoria se transforma (lineal o lineal equivalente) por medio de la expresión:

$$a = f(z) = f\left[\sum (w_{1j} p_j)\right] \quad (2.2)$$

donde  $f$  es un funcional, denominado función de activación, con el cual se obtiene la activación o respuesta de la neurona. Además, se puede presentar, si se desea o no en el modelo neuronal otra entrada, la cual en la mayoría de los casos es 1 y se multiplica por un factor de compensación  $b$  denominado sesgo, el cual es similar a un peso pero que permanece constante, y posteriormente pasa al sumando, entonces la expresión (2.1) se generaliza como:

$$z = \sum (w_{1j} p_j + b) \quad (2.3)$$

Haciendo la similitud con una neurona biológica, los pesos  $w_{1j}$  representan la resistencia sináptica, el cuerpo de la célula es representado por la sumatoria  $\Sigma$  y la función de activación, y la salida de la neurona ( $a$ ) sería la señal que pasa al axón.

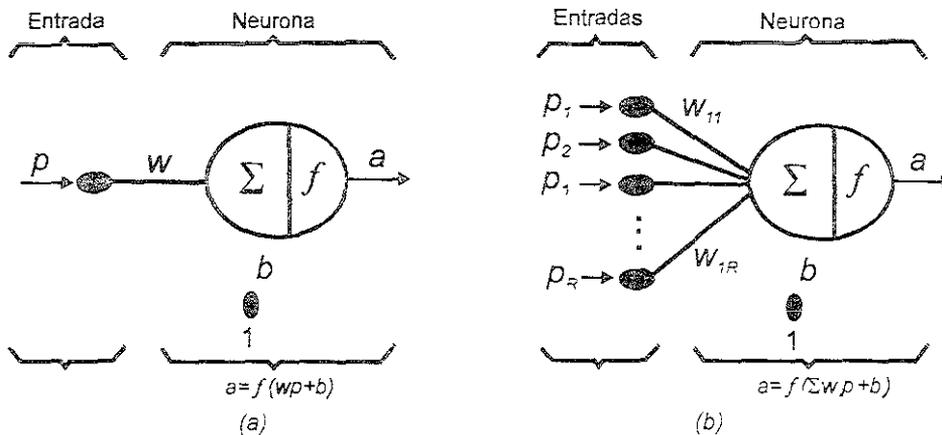


Fig. 2.2 (a) Neurona de entrada simple y (b) Neurona de múltiples entradas

La matriz  $w_{1j}$  y el escalar  $b$  son ajustados por alguna regla de aprendizaje, mientras que la salida real de la neurona dependerá de la función de activación  $f$  que seleccione el usuario. Para el caso de varias neuronas, como se mostrará más adelante, la matriz de pesos se define como  $w_{ij}$ , el sesgo es ahora el vector  $b_i$  y las activaciones o salidas de las neuronas se representan con el vector  $a_i$ .

## 2.2.2 Funciones de Activación

Para obtener la *activación* o respuesta de la neurona, la sumatoria representada por la expresión (2.3) se transforma a un valor neto usando una función de umbral (ecuación (2.2)). Esta función de umbral llamada función de activación o de transferencia, puede ser alguna forma de función lineal o no lineal de  $z$ . Sin embargo, en la mayoría de los problemas tratados con redes se usan funciones no lineales (común en las redes de varias capas), debido a que si la neurona no transforma su entrada a un valor neto, se dice que tiene una función de activación *identidad o lineal*.

Esta función de activación funciona como un límite en el rango de valores que se desea que la red reproduzca, es decir, se selecciona para satisfacer alguna característica del problema que la neurona trata de resolver. Existe una gran variedad de funciones de activación y un ejemplo de éstas, se muestra a continuación.

### Función de activación logarítmica sigmoideal

La función de activación log-sigmoideal (fig. 2.3a), se utiliza comúnmente en redes multi-capas que son entrenadas usando el algoritmo de retropropagación (sección 2.4.4), principalmente porque esta función es diferenciable. Esta función transforma el argumento,  $z$ , a una salida limitada en el rango de 0 a 1, según la expresión:

$$a = \frac{1}{1 + e^{-z}}$$

Se observa que el argumento,  $z$ , puede tener cualquier valor en el rango  $-\infty$  a  $+\infty$ . La misma función se muestra en la fig. 2.3b para una neurona de una sola entrada, donde ahora se ilustra el efecto del *peso* y *sesgo*. El esquema representativo de esta función se presenta con un *ícono* que reemplaza a la  $f$  general para mayor simplificación de los diagramas de RNA.

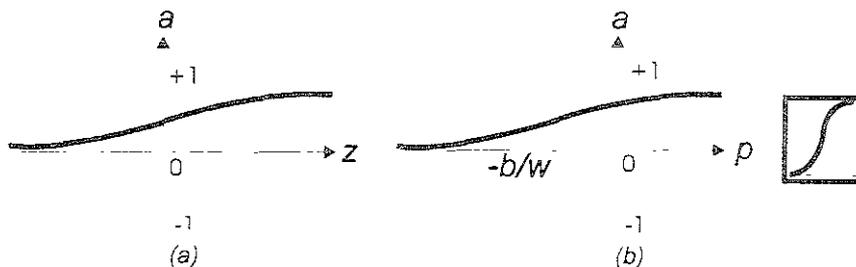


Fig.2.3 Función de activación log-sigmoideal

En la Tabla 2.2 se presentan las funciones de activación más comunes, y por supuesto es posible definir otro tipo de funciones además de las mostradas. La función de activación en redes multi-capas (sección 2.4) debe ser no lineal. Una red multi-capas, integrada por  $S$  capas de neuronas ocultas y con funciones de activación lineales, no sería mejor que el *perceptrón lineal simple*, el cual no tiene neuronas ocultas sólo unidades de entrada y de salida (Rosenblatt, 1958). Casi cualquier función no lineal es útil, aunque para el aprendizaje de retropropagación es necesario que la función sea diferenciable lo cual limita el uso de todas las funciones no lineales. Las funciones sigmoideas como la logística (logsig), la tangente hiperbólica (tanh), así como la función gaussiana (gauss) son las opciones más comunes. Las funciones como la tanh que produce valores positivos y negativos, ayuda a que la ejecución del entrenamiento sea de manera más rápida que las funciones que producen solo valores positivos como la logsig, debido que se tiene un mejor condicionamiento numérico.

Tabla 2.2 Funciones de activación o de transferencia más comunes

Nombre	Relación entrada (n)/salida (a)	Gráfica
Límites definidos	$a=0 \quad n < 0$ $a=1 \quad n \geq 0$	
Límites definidos simétricos	$a=-1 \quad n < 0$ $a=+1 \quad n \geq 0$	
Lineal	$a = n$	
Lineal de saturación	$a=0 \quad n < 0$ $a=n \quad 0 \leq n \leq 1$ $a=1 \quad n > 1$	
Lineal de saturación simétrica	$a=-1 \quad n < -1$ $a=n \quad -1 \leq n \leq 1$ $a=1 \quad n > 1$	
Log-sigmoidal	$a = \frac{1}{1 + e^{-n}}$	
Sigmoidal tangente hiperbólico	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$	
Lineal positivo	$a=0 \quad n < 0$ $a=n \quad 0 \leq n$	

Para las unidades de salida (output), se debe seleccionar una función de activación que satisfaga a la distribución de los valores deseados (objetivos). Por ejemplo, para salidas en un rango entre 0 y 1, la función logsig es una opción excelente. Para valores deseados continuos dentro de un rango limitado, las funciones logsig y tanh son de nuevo útiles, siempre y cuando las salidas se distribuyan en el rango de los valores deseados u objetivos, o distribuir los objetivos al rango de la función de activación de salida de la red (normalización al rango de las funciones de activación). Pero si los valores deseados no tienen ningún rango limitado conocido, es mejor usar una función de activación ilimitada, donde a menudo la función de identidad o lineal es apropiada (solo útil en las neuronas de salida). Si los valores deseados son positivos pero no tienen ningún límite superior conocido, puede usarse una función de activación de salida exponencial

### 2.2.3 Arquitectura de Redes

La arquitectura o topología de una RNA se define por las conexiones entre las unidades de procesamiento, la cual es de gran importancia para la ejecución o capacidad de la red. La arquitectura de red más general es cuando todos sus nodos interactúan totalmente (ver fig 2.4). Los casos que se derivan del anterior, son las *redes de retroalimentación* cuando se tienen conexiones que alimentan hacia atrás, y las *redes de alimentación hacia adelante* que, como su nombre lo indica, sus conexiones no forman ciclos cerrados

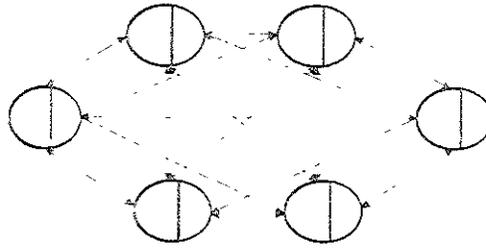


Fig 2.4 Arquitectura de red caso general

En una red de alimentación hacia adelante siempre se pueden identificar las unidades de entrada (no existen conexiones que entran a ellas) y las unidades de salida (no hay conexiones que salen de ellas). Entre estas unidades extremas, hablando de redes multi-capa, están las capas ocultas arregladas en unidades de procesamiento que ejecutan todas las operaciones de la red.

Un conjunto de varias unidades de procesamiento operando en paralelo constituyen una capa de neuronas. En la fig. 2.5, se muestra una red de una sola capa con S neuronas y de entradas múltiples, donde cada elemento del vector de entradas  $p$ , se conecta a cada neurona a través de la matriz de peso  $w_{ij}$ , cada neurona tiene un sesgo  $b_i$ , un sumando, una función de activación  $f$  y una salida  $a_i$ . De tal forma, que la capa incluye la matriz de pesos, los sumandos, el vector sesgo, la función de activación y el vector de salida.

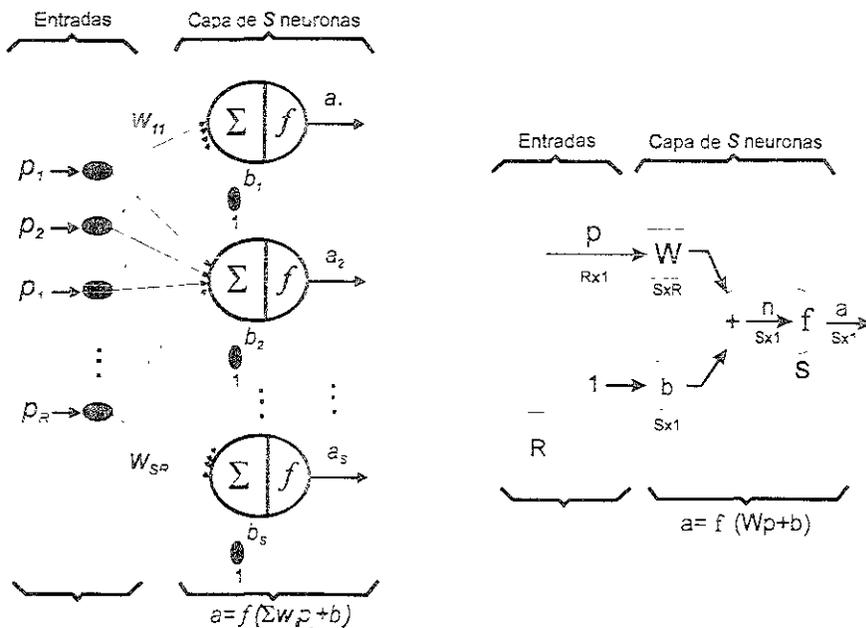


Fig 2.5 Red de una sola capa de S neuronas y en notación abreviada

Los elementos del vector de entrada se introducen en la red a través de la matriz de pesos  $W$

$$W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1R} \\ w_{21} & w_{22} & \dots & w_{2R} \\ \vdots & \vdots & & \vdots \\ w_{S1} & w_{S2} & \dots & w_{SR} \end{bmatrix}$$

En la matriz de pesos  $W$ , los índices de renglón indican la neurona de destino asociada con ese peso, mientras que los índices de columna indican la fuente de entrada para ese peso. Así, los índices en  $w_{2,1}$  dicen que este peso representa la conexión a la segunda neurona de la primera fuente. En notación abreviada, de acuerdo con la fig. 2.5,  $p$  es un vector de dimensión  $R$ ,  $W$  es una matriz de  $S \times R$ , y  $a$  y  $b$  son vectores de dimensión  $S$ .

Si ahora se considera una red con varias capas (*red multi-capas*), cada capa tiene su propia matriz de peso  $W$ , su propio vector del sesgo  $b$ , un vector de entrada neta  $z$  y un vector de salida  $a$ . Específicamente, añadimos el número de la capa con otro subíndice en cada una de estas variables. Esta notación se usa en la red de tres capas mostrada en la fig. 2.6

Se observa que existen  $R$  entradas y que la primera, segunda y tercera capas, tienen  $S_1$ ,  $S_2$  y  $S_3$  neuronas respectivamente. Las salidas de la primera capa son las entradas para la segunda capa; y asimismo, las salidas de la segunda capa son entradas para la tercera. Una capa que constituye la salida de la red se llama *capa de salida*; las otras capas se llaman *capas ocultas*. La red mostrada en la fig. 2.6 tiene una capa de salida y dos capas ocultas. Las redes multi-capas pueden aproximar bastante bien la mayoría de las funciones, lo que las redes de una sola capa no pueden hacer.

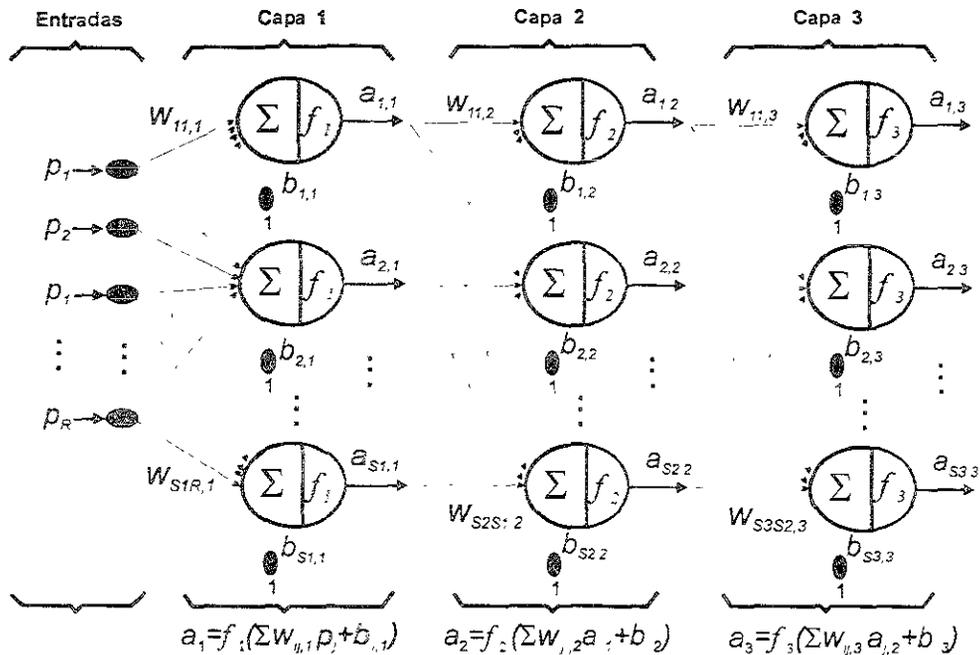


Fig 2.6 Red de tres capas (dos capas ocultas)

### 2.3 SUPERFICIES DE EJECUCIÓN Y OPTIMIZACIÓN

Durante el entrenamiento de redes multi-capas, las funciones o reglas de aprendizaje ajustan los parámetros de la red, básicamente los pesos y los sesgos, para optimizar su ejecución. Para valorar una regla de aprendizaje se define el *índice de funcionamiento o ejecución* como una medida cuantitativa de funcionamiento del entrenamiento de la red, que tiene un valor pequeño cuando el entrenamiento es aceptable y grande cuando el entrenamiento es pobre. De esta forma, un proceso de optimización equivale a determinar el espacio de parámetros para reducir el índice de ejecución. Este índice de ejecución dentro del entrenamiento de redes multi-capas es en términos de una función de error  $E$

Antes de describir los algoritmos de aprendizaje de redes neuronales de varias capas, es esencial conocer sus características principales. Una forma muy útil para describir las tareas de ejecución de los algoritmos de entrenamiento es en términos de *optimización de funciones*, es decir, la minimización de la función de error  $E$ . Desde este punto de vista, el entrenamiento de la red se convierte en un proceso de minimización del error o de optimización, y por lo tanto, se define en cada tarea de entrenamiento, una *superficie de error* multidimensional no negativa (o superficie de energía) que se forma al graficar el valor de  $E$  para todos los pesos  $w$  de la red. Esta aproximación es de gran utilidad, debido a que se han desarrollado fuera del campo de redes neuronales muchas estrategias eficientes para la optimización de funciones, las cuales en su mayoría pueden adaptarse sin ningún problema al entrenamiento de RNA.

### 2.3.1 Características de las Superficies de Error

De acuerdo a sus propiedades generales existen diferentes clases de superficies de error; sin embargo nos enfocamos en las propiedades de una sola clase: *superficies de error suaves*; las cuales, son muy comunes en la práctica, y la mayoría de los algoritmos de entrenamiento de redes multi-capas están diseñados para este tipo de superficies de error.

La suavidad de una superficie de error depende de los grados para los cuales la función de error  $E$  es diferenciable continua. La función  $E$  es continua si para pequeños cambios en los pesos de la red se producen pequeños cambios en  $E$ ; y es diferenciable continua si la primera derivada de la función  $E$  ( $\nabla E$ ) existe y es continua. De la misma forma, la función  $E$  es doblemente diferenciable continua si su segunda derivada ( $\nabla^2 E$ ) existe y es continua. Entonces, se define al gradiente  $g(w)$  y al Hessiano  $G(w)$  como

$$g_i = \frac{\partial E(w)}{\partial w_i}, \quad 1 \leq i \leq W \quad G_{ij} = \frac{\partial^2 E(w)}{\partial w_i \partial w_j}, \quad 1 \leq i, j \leq W$$

De tal forma que si  $E$  es dos veces diferenciable continua, el Hessiano será simétrico ( $G=G^T$ )

Un conjunto de características más comunes y útiles de las superficies de error son los *puntos estacionarios*, en donde el gradiente  $g(w)$  es cero. Existen diferentes tipos de puntos estacionarios conocidos: *máximo*, *mínimo*, *punto de inflexión* (caso unidimensional), y *punto silla* (caso multidimensional). Estas categorías pueden ser fuertemente divididas en puntos *fuertes* y *débiles*. Las diferencias entre estos puntos estacionarios son más fáciles de entender por medio de un caso unidimensional (ver figs. 2.7a y b). En un caso multidimensional, un *mínimo* puede ser el punto más bajo de una cuenca en la superficie de error, y un *máximo* el punto más alto de una cuenca invertida. Un *punto silla*, como su nombre lo indica, es el equivalente multidimensional de un punto de inflexión. Sin embargo, un punto estacionario *fuerte* consiste de un solo punto, mientras que los puntos estacionarios *débiles* son caracterizados por una línea en el caso unidimensional y puede ser también una línea o una superficie en el caso multidimensional.

Considerando que la función de error  $E$  es dos veces diferenciable y continua, la diferencia entre máximo, mínimo y punto silla se refleja en los eigenvalores de la matriz Hessiana  $G$  (un escalar  $\lambda$  es un valor característico de  $G$  si hay un vector  $x$  no cero tal que  $Gx = \lambda x$ ). El punto estacionario  $w^*$  es definitivamente un máximo fuerte si  $G(w^*)$  es negativo definido (todos los eigenvalores de  $G$  son estrictamente negativos), y definitivamente un mínimo fuerte si  $G(w^*)$  es positivo definido (todos los eigenvalores de  $G$  son estrictamente positivos). La situación es menos clara para máximos y mínimos débiles, es decir, el punto estacionario  $w^*$  puede ser un máximo débil si  $G(w^*)$  es negativo semi-definido (todos los eigenvalores de  $G$  son no positivos), y puede ser un mínimo débil si  $G(w^*)$  es positivo semi-definido (todos los eigenvalores de  $G$  son no negativos).

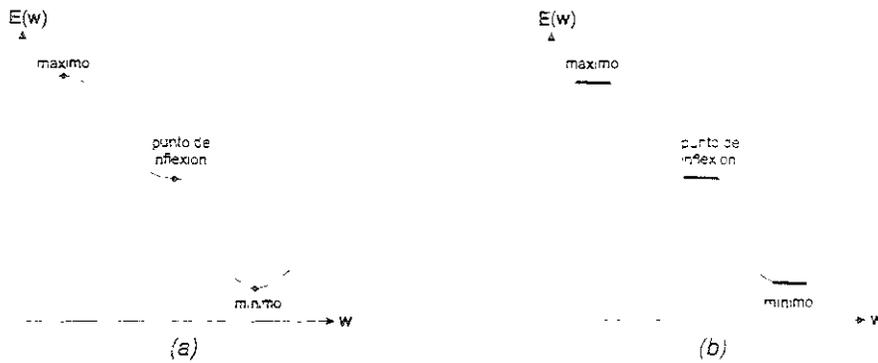


Fig. 2.7 Superficies de error con (a) puntos estacionarios fuertes y (b) puntos estacionarios débiles

El objetivo principal es minimizar la función  $E$ , por lo cual, el *mínimo* es en particular la característica más importante de la superficie de error. Cada mínimo tiene una cuenca de atracción asociada (una región alrededor del mínimo del cual es solamente posible escapar por pasos sobre regiones grandes (o por deformación de la superficie de error en algunos casos). Cerca del mínimo, la cuenca será convexa o estrictamente convexa. La definición de una función *convexa* depende de las definiciones de un conjunto convexo, es decir, un conjunto de puntos que si contiene a los puntos  $x$  y  $y$ , también contienen el segmento de línea entre los puntos  $x$  y  $y$ . La diferencia entre funciones convexas, estrictamente convexas y no convexas se explica fácilmente en los casos unidimensionales (ver fig. 2.8a, b y c). Sin embargo, dentro de las funciones convexas existen rarezas, donde la *convexidad local* (convexidad en la vecindad de un mínimo) es un concepto importante en los métodos de optimización. En términos de la matriz del Hessiano  $G(w^*)$  para un mínimo  $w^*$ , la *convexidad local* corresponde a un Hessiano semi-definido positivo y *convexidad local estricta* a un Hessiano definido positivo.

Si el mínimo  $w^*$  es el punto más bajo de toda la superficie de error, se conoce como un *mínimo global* (o mínimo absoluto). Si existe un mínimo  $w^*$  con un nivel de error más alto que el mínimo global, se trata de un *mínimo local*. En general, no existe un camino fácil y efectivo, ya sea un mínimo  $w^*$  local o global es necesario conocer la información de toda la superficie de error y no únicamente en la vecindad de  $w^*$ , lo cual dificulta el problema.

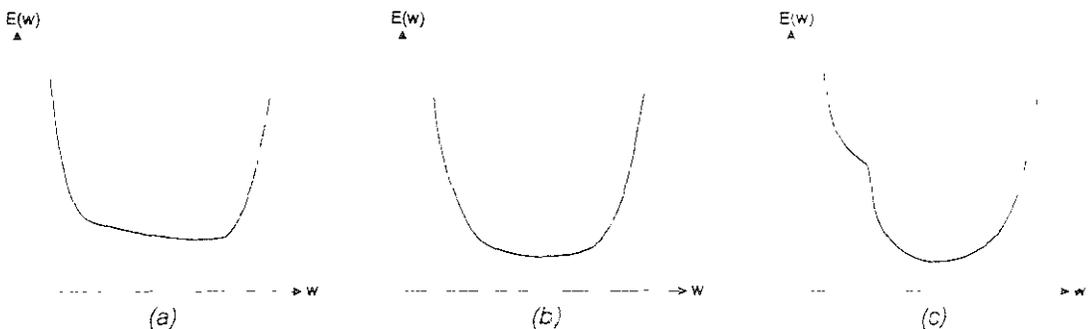


Fig. 2.8 (a) Función convexa, (b) Función estrictamente convexa y (c) Función no convexa

Otras características importantes en las superficies de error son los *valles* y *mesetas multidimensionales*. Los *valles estrechos* corresponden a convoluciones y elipsoides de alta excentricidad en la superficie de error; en casos especiales donde  $E$  es cuadrático con un Hessiano positivo definido constante, es posible cuantificar con precisión el grado de excentricidad usando los eigenvalores de la matriz Hessiana  $G$  (Luenberger, 1984). Las *mesetas* son regiones donde la pendiente es plana en dimensiones múltiples simultáneamente y cuando es suficientemente plana es seguramente un punto estacionario débil.

### 2.3.2 Métodos Clásicos de Optimización

El objetivo de los métodos clásicos de optimización no lineal es determinar el punto óptimo  $\mathbf{x}^*$  (mínimo de una función suave no lineal arbitraria  $F(\mathbf{x})$ ) donde el vector  $\mathbf{x}$  contiene  $N$  variables independientes parámetros libres,  $x_i$  ( $i=1, \dots, N$ ). Cuando  $N$  es más grande que 1 y no hay restricciones en los valores aceptables de  $\mathbf{x}$ , se tiene un problema de *minimización no restringida multivariada* (Shepherd, 1997). En términos de  $F$  y  $\mathbf{x}$ , los elementos de vector gradiente  $\mathbf{g}$  y de la matriz Hessiana  $\mathbf{G}$  están dados por

$$g_i = \frac{\partial F(\mathbf{x})}{\partial x_i}, \quad 1 \leq i \leq N \quad \mathbf{G}_{ij} = \frac{\partial^2 F(\mathbf{x})}{\partial x_i \partial x_j}, \quad 1 \leq i, j \leq N \quad (2.4)$$

Estos métodos pueden usarse en el entrenamiento del perceptron multi-capas (sección 2.5.4) a partir de hecho de que cualquier combinación de arquitectura de red, función de error y conjunto de entrenamiento se define una función no lineal  $F(\mathbf{x})$ , donde los pesos  $\mathbf{w}$  de la red son equivalentes a los parámetros libres. Todos los métodos de optimización clásica que se describen en esta sección presentan las siguientes características:

a) Se derivan, algebraicamente, de la expansión de la serie de Taylor de una función multivariada  $F$  en la vecindad de un punto arbitrario  $\mathbf{x}$ :

$$F(\mathbf{x} + \Delta\mathbf{x}) = F(\mathbf{x}) + \mathbf{g}(\mathbf{x})^T \Delta\mathbf{x} + \frac{1}{2} \Delta\mathbf{x}^T \mathbf{G}(\mathbf{x}) \Delta\mathbf{x} + \dots \quad (2.5)$$

b) Son algoritmos de descenso iterativos, es decir, el mínimo  $\mathbf{x}^*$  se determina en una serie de pasos, con  $F(\mathbf{x}_{k+1}) \leq F(\mathbf{x}_k)$  para cada intervalo de paso  $k$ . El paso tomado para cada iteración  $k$  (equivalente a  $\mathbf{x}_{k+1} - \mathbf{x}_k$ ) será denotado como  $\Delta\mathbf{x}_k$ . Es decir, estos algoritmos tienen como condición inicial  $\mathbf{x}_0$  que se actualiza con la relación:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta\mathbf{x}_k \quad (2.6)$$

c) Requieren que la dirección de búsqueda  $\mathbf{p}_k$  sea una dirección descendente ( $\mathbf{g}_k^T \mathbf{p}_k < 0$ ). En términos de dirección de búsqueda  $\mathbf{p}_k$ , el paso  $\Delta\mathbf{x}_k$  es definido por:

$$\Delta\mathbf{x}_k = \alpha_k \mathbf{p}_k \quad (2.7)$$

donde  $\alpha_k$  es un escalar no negativo llamado *coeficiente de aprendizaje* y que determina la longitud de paso.

#### 2.3.2.1 Algoritmo del descenso más pronunciado

Es un método de optimización de primer orden, que se basa en el modelo lineal. Considerando la serie de Taylor  $F(\mathbf{x})$  de primer orden deducida de la expresión (2.5) se tiene:

$$F(\mathbf{x}_k + \Delta\mathbf{x}_k) \approx F(\mathbf{x}_k) + \mathbf{g}(\mathbf{x}_k)^T \Delta\mathbf{x}_k \quad (2.8)$$

donde  $\mathbf{g}_k$  es el gradiente evaluado en la condición anterior  $\mathbf{x}_k$ . La dirección donde la función decrece más rápidamente, ocurre cuando  $\mathbf{g}_k^T \mathbf{p}_k$  es más negativo (siendo  $\mathbf{p}_k$  constante, solamente cambia la dirección) es decir, cuando el vector dirección sea el gradiente negativo ( $\mathbf{p}_k = -\mathbf{g}_k$ ). Sustituyendo lo anterior en la ecuación (2.6) se tiene el método del *descenso más pronunciado*:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \mathbf{g}_k \quad (2.9a)$$

En este algoritmo el coeficiente de aprendizaje  $\alpha_k$  se determina a partir de dos métodos generales

- a) Minimizando el índice de ejecución  $F(\mathbf{x}_k - \alpha_k \mathbf{g}_k)$  usando una *estrategia de línea de búsqueda* (Hagan y Demuth, 1995), donde el valor óptimo de  $\alpha_k$ , para un índice de ejecución  $F$  cuadrático y un Hessiano positivo definido constante  $\mathbf{G}$ , está dado por:

$$\alpha_k = \frac{\mathbf{g}_k^T \mathbf{g}_k}{\mathbf{g}_k^T \mathbf{G} \mathbf{g}_k} \quad (2.9b)$$

La convergencia del algoritmo utilizando una minimización a lo largo de una línea de búsqueda se presenta en la fig. 2.9

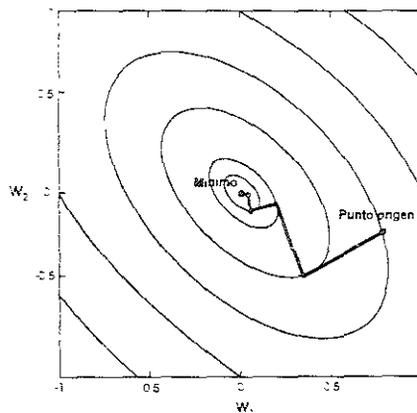


Fig. 2.9 Descenso más pronunciado con minimización a lo largo de una línea

- b) Utilizando un valor fijo de  $\alpha_k$  (p. e.  $\alpha_k = 0.01$ ) o variándolo con respecto a cada iteración (p. e.  $\alpha_k = 1/k$ ). La selección de un valor pequeño de  $\alpha_k$  puede asegurar la convergencia del método a un punto estacionario  $\mathbf{x}^*$ ; sin embargo, la velocidad de convergencia del método es lento. Utilizar un valor de  $\alpha_k$  grande para llegar a la convergencia más rápidamente puede ocasionar inestabilidad en el algoritmo, como se ilustra en la fig. 2.10. Un valor de  $\alpha_k$  máximo estable, considerando un índice de ejecución  $F$  cuadrático, está dado por  $\alpha < (2/\lambda_{\max})$ ; donde el escalar  $\lambda_{\max}$  es el eigenvalor más grande de  $\mathbf{G}$ .

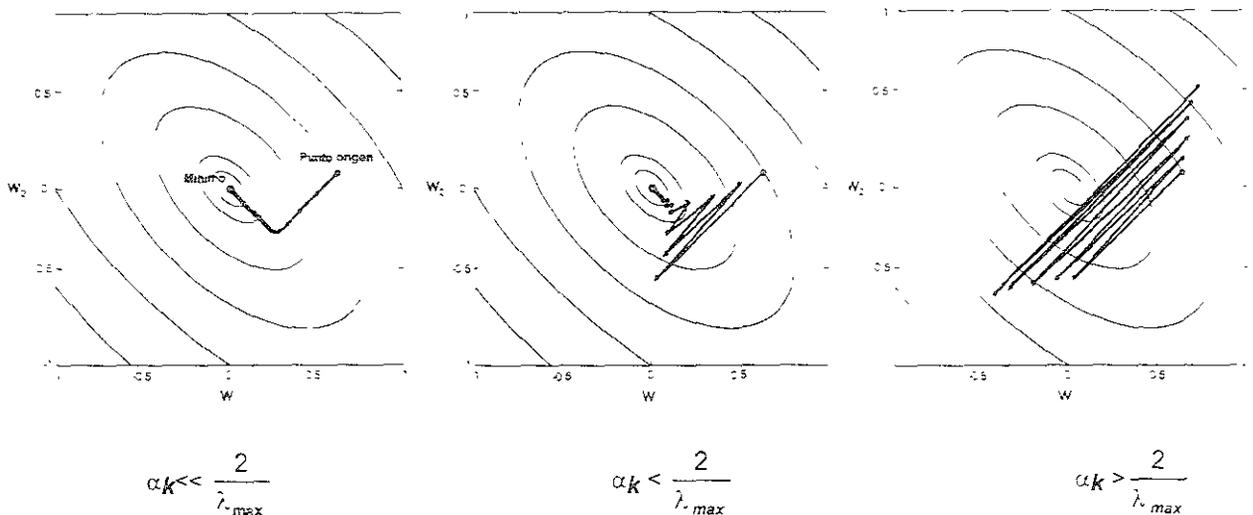


Fig. 2.10 Trayectoria del descenso más pronunciado para diferentes  $\alpha_k$ .

### 2.3.2.2 Método de Newton

El método de Newton es un algoritmo de optimización de segundo orden basado en el modelo cuadrático

$$F(\mathbf{x}_k - \Delta\mathbf{x}_k) \approx F(\mathbf{x}_k) + \mathbf{g}(\mathbf{x}_k)^T \Delta\mathbf{x}_k + \frac{1}{2} \Delta\mathbf{x}_k^T \mathbf{G}(\mathbf{x}_k) \Delta\mathbf{x}_k \quad (2.10)$$

La velocidad de convergencia teórica y la ejecución práctica de los métodos de segundo orden son generalmente superiores a aquellos de primer orden (descenso más pronunciado), si el índice de ejecución  $F$  es suficientemente suave. Si las sucesivas direcciones de búsqueda son *mutuamente conjugadas* con respecto a la matriz del Hessiano  $\mathbf{G}$ , es decir  $\mathbf{p}_i^T \mathbf{G} \mathbf{p}_j = 0$  para toda  $i \neq j$ , entonces el método se aproximará con una velocidad de convergencia rápida. De lo anterior, la regla de actualización para este método se define como:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{G}_k^{-1} \mathbf{g}_k \quad (2.11)$$

Para funciones  $F(\mathbf{x})$  cuadráticas, con un mínimo fuerte y con un Hessiano positivo definido, el método de Newton converge en una sola iteración (ver fig. 2.11). Para funciones no cuadráticas, la velocidad de convergencia local será cuadrática, es decir, en varios pasos. De hecho, no se puede garantizar que el método convergerá para todas las funciones, lo cual depende de la función y de las condiciones iniciales.

Lo anterior indica que el método de Newton proporciona una mayor exactitud y convergencia rápida, siempre que las funciones analíticas que representan el problema se aproximen con funciones cuadráticas en una zona cercana a un mínimo fuerte. Sin embargo, el método no puede distinguir entre un mínimo local y global, debido a que la aproximación es una función cuadrática donde sólo se tiene un mínimo, y por lo tanto, no se pueden conocer las características globales de la función. Por otra parte, la disponibilidad de trabajar con la matriz  $\mathbf{G}$ , ayuda a distinguir entre punto silla y un mínimo; es decir, si el punto estacionario  $\mathbf{x}^*$  es un punto silla ( $\mathbf{G}$  es indefinida), es posible escapar de la posición de  $\mathbf{x}^*$  siguiendo una *dirección de curvatura negativa* (Shepherd, 1997). Se requiere entonces calcular y almacenar la matriz Hessiana y su inversa, lo que implica un mayor costo computacional y de almacenamiento, y por tal razón este método no es recomendado en la práctica. La dirección de búsqueda del método de Newton es la misma que la del descenso más pronunciado cuando  $\mathbf{A}_k = \mathbf{A}_k^{-1} = \mathbf{I}$ , por lo cual, se han propuesto algoritmos que caen de alguna forma entre estos dos métodos.

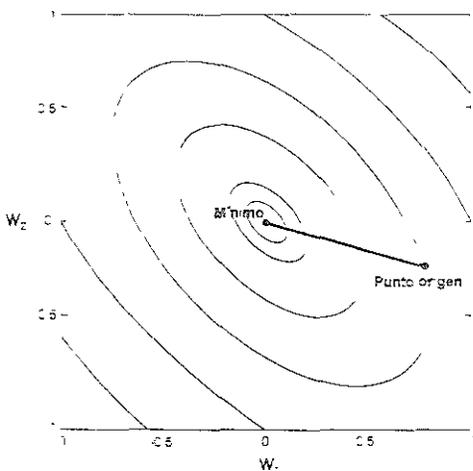


Fig. 2.11 Trayectoria de convergencia del método de Newton

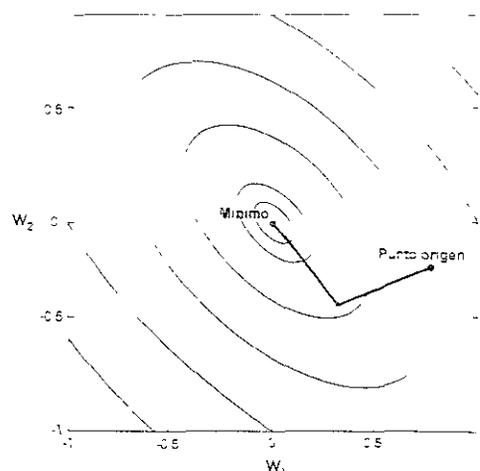


Fig. 2.12 Trayectoria de convergencia de algoritmo de gradiente conjugado

### 2.3.2.3 Gradiente conjugado

Debido a que el método de Newton implica un alto costo de cálculo y almacenamiento cuando el número de parámetros  $N$  es grande, lo cual es común en muchas aplicaciones prácticas de RNA, se han propuesto otros algoritmos como el de *gradiente conjugado* donde básicamente se requiere sólo calcular las primeras derivadas pero conservando la terminación cuadrática (Hestenes y Stiefel, 1952). Los métodos de gradiente conjugado se basan en que una sucesión de direcciones de búsqueda mutuamente conjugadas ( $\mathbf{p}_i^T \mathbf{G} \mathbf{p}_i = 0$ ) es una serie  $\{\mathbf{p}_i\}$ ,  $i=0, \dots, k$ , que puede generarse iterativamente sin necesidad de calcular la matriz del Hessiano, por medio de la siguiente expresión

$$\mathbf{p}_k = -\mathbf{g}_k - \beta_k \mathbf{p}_{k-1} \quad (2.12)$$

usando como dirección de búsqueda inicial la seleccionada para el algoritmo del descenso más pronunciado ( $\mathbf{p}_0 = -\mathbf{g}_0$ ). La diferencia entre los métodos de gradiente conjugados se encuentra en la forma de calcular el escalar  $\beta_k$  de la ecuación (2.12). Las tres más comunes fórmulas de gradiente conjugado para  $\beta_k$  son las ecuaciones de Hestenes-Seifel, de Fletcher-Reeves y de Polak-Ribière.

$$\begin{array}{lll} \text{Hestenes-Seifel} & \text{Fletcher-Reeves} & \text{Polak-Ribière} \\ \beta_k = \frac{(\mathbf{g}_k - \mathbf{g}_{k-1})^T \mathbf{g}_k}{(\mathbf{g}_k - \mathbf{g}_{k-1})^T \mathbf{p}_{k-1}} & \beta_k = -\frac{\mathbf{g}_k^T \mathbf{g}_k}{\mathbf{p}_{k-1}^T \mathbf{g}_{k-1}} & \beta_k = \frac{(\mathbf{g}_k - \mathbf{g}_{k-1})^T \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{g}_{k-1}} \end{array} \quad (2.13)$$

Una trayectoria de convergencia del método de gradiente conjugado es como la que se presenta en la fig. 2.12.

## 2.4 ENTRENAMIENTO DEL PERCEPTRÓN MULTI-CAPA (PMC)

En el modelo de las primeras neuronas artificiales, los parámetros de la red fueron propuestos y no se disponía de algún método de entrenamiento (McCulloch y Pitts, 1943). Más adelante se desarrolló la red neuronal llamada perceptrón (Rosenblatt, 1958), en la que se presentaba una regla de aprendizaje para entrenar a la red. Sin embargo, todavía la red perceptrón tenía limitaciones (Minsky y Papert, 1969), las cuales se superaron con la aparición del perceptrón multi-capas (PMC) y las reglas de aprendizaje asociadas. A partir de entonces, la red perceptrón multi-capas es una de las clases de redes neuronales más comúnmente usadas.

Las redes PMC se emplean en la *clasificación de patrones* y *aproximación de funciones*. Uno de los casos con los que se demostró la capacidad de este tipo de redes, y al mismo tiempo significó el renacimiento de la técnica de redes en general, fue la solución del problema clásico excluyente XOR (Rumelhart y McClelland, 1986), que es un claro ejemplo de clasificación de dos categorías linealmente separables. Los casos de aplicaciones del PMC en aproximación de funciones son en *sistemas de control* (encontrar una función de retroalimentación apropiada que mapee las salidas medidas a las entradas de control) y en *filtros adaptables* (encontrar una función que mapee los valores retrasados de una señal de entrada a una señal de salida apropiada). Según referencias (Hagan, Demuth y Beale, 1996), una red de dos capas, con funciones de activación sigmoideas en la capa oculta y lineales en la capa de salida, pueden aproximar virtualmente cualquier función de interés a cualquier grado de exactitud.

El entrenamiento del PMC envuelve el ajuste de la red, es decir, determinar la configuración de pesos tal que ésta sea capaz de producir una salida específica para un dado conjunto de patrones de entrada.

### 2.4.1 Arquitectura del Perceptrón Multi-Capa

Como se ha mencionado, la arquitectura del perceptrón multi-capa consiste de unidades o nodos arreglados en dos o más capas, donde la capa de entrada no se cuenta debido a que solamente sirve para distribuir las entradas de cada patrón del conjunto de entrenamiento. Los nodos están conectados con pesos de valor real, pero no hay conexiones entre nodos en una misma capa. La notación para el PMC es la misma que se presenta en la sección 2.2.3, como por ejemplo el perceptrón de tres capas mostrado en la fig 2.6, donde la salida de la primera capa es la entrada de la segunda capa y la salida de la segunda es la alimentación de la tercera capa. Cada capa puede tener diferentes números de neuronas, y cada una de ellas una función de activación diferente. De esta forma, para identificar la estructura de una red multi-capa, se usará la notación simbólica abreviada mostrada en la fig. 2.13 (red de 3 capas), donde el número de entradas es seguido por el número de neuronas en cada capa

$$R - S_1 - S_2 - S_3$$

El número de nodos en las capas de entrada y salida de la red PMC se determina, respectivamente, por el tamaño de patrones de entrada y el tamaño de salidas deseadas de la tarea de entrenamiento seleccionada.

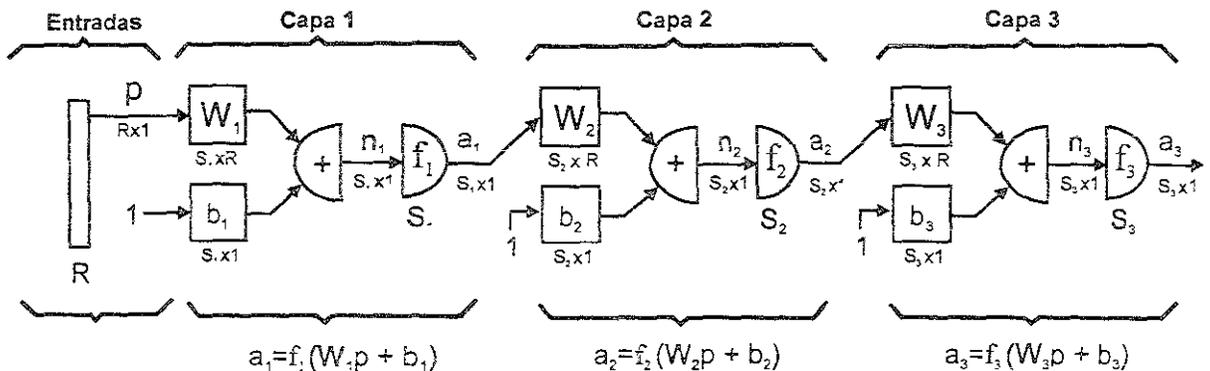


Fig. 2.13 Red de tres capas en notación abreviada (dos capas ocultas)

### 2.4.2 Regla de Aprendizaje del PMC

El procedimiento para modificar los pesos y sesgos de una red PMC se conoce como *regla de aprendizaje* o algoritmo de entrenamiento. Esta regla tiene el propósito de entrenar o educar a la red para realizar alguna tarea específica. Existen muchos tipos de algoritmos de entrenamiento, sin embargo, todos ellos caen dentro de las siguientes categorías de entrenamiento:

- Entrenamiento supervisado.** En este tipo de entrenamiento se proporciona con anticipación un conjunto de datos de entrenamiento que proporcionará la conducta apropiada a la red. Siendo por notación,  $p_q$  una entrada a la red y  $t_q$  la correspondiente salida correcta (*salidas objetivo o deseadas*). El algoritmo ajusta los pesos y los sesgos de la red para que sus salidas se acerquen a las salidas objetivo.
- Entrenamiento reforzado.** Es idéntico al supervisado, pero ahora la salida de la red es una calificación en vez de proporcionar la salida correcta a cada entrada de la red. Esta calificación es una medida de la ejecución del trabajo de la red sobre algún conjunto de datos de entrada, y este tipo de entrenamiento tiene mucha aplicación en sistemas de control.

c) **Entrenamiento no supervisado.** En este tipo de entrenamiento no existe ninguna salida objetivo disponible, los pesos y los sesgos se modifican únicamente en respuesta al conjunto de datos de entrada. La mayoría de estos algoritmos realizan algún tipo de funcionamiento agrupado.

El entrenamiento del perceptrón multi-capas es un proceso iterativo que involucra en cada iteración o época, el cálculo de las salidas de la red para cada patrón del conjunto de entrenamiento ajustando los pesos y sesgos de la red en función de la diferencia entre la salida de la red actual  $a_p$  para un patrón dado y la salida deseada  $t_p$  para ese patrón; siendo de esta forma, un entrenamiento de tipo *supervisado*. Dados una arquitectura de red y un algoritmo de entrenamiento convenientes, los pesos de la red se ajustarán progresivamente hasta que la salida de la red sea cercanamente aceptable a la salida deseada para cada patrón en el conjunto de entrenamiento.

El índice de ejecución dentro de entrenamiento del PMC en términos de una función de error o de energía  $E$  es de la siguiente forma:

$$E = \sum_{p=1}^P E_p$$

donde  $E_p$  es la contribución parcial del patrón  $p$  para el error total  $E$  de la red. La selección de la función de error puede tener un impacto significativo en la ejecución de una red PMC como la selección del algoritmo de entrenamiento. Las dos funciones más populares de error son la *suma-de-cuadráticos* (o *cuadráticos mínimos*):

$$E = \frac{1}{2} \sum_{p=1}^P \sum_{i=1}^{S_M} (t_i - a_i)^2$$

y una versión normalizada de la suma-de-cuadráticos llamada error *cuadráticos mínimo* (MSE) también conocida como error *medio cuadrático* (LMS):

$$E = \frac{1}{2pS_M} \sum_{p=1}^P \sum_{i=1}^{S_M} (t_i - a_i)^2$$

La ventaja de la función *MSE* sobre la función de error *suma-de-cuadráticos* es que, siendo insensitivas ambas al número de patrones en el conjunto de entrenamiento y el número de nodos de salida de la red, ésta puede usarse para hacer comparaciones más significativas entre diferentes tareas de entrenamiento. Una característica importante de la *MSE* y la función de error *suma-de-cuadráticos* es que la reducción en  $E$  puede estar asociada a un incremento del número de patrones de entrenamiento clasificados incorrectamente; en efecto, reduciendo  $E$  a su mínimo nivel logrado con esas funciones de error no significa necesariamente que el número mínimo de patrones clasificados erróneamente han sido logrados. Esta propiedad puede tener consecuencias no deseables, como por ejemplo incrementar la probabilidad que la red converja a regiones en el espacio de pesos donde la vasta mayoría de patrones son correctamente clasificados, pero un pequeño número sean severamente clasificados erróneamente.

### 2.4.3 Superficies de Error del PMC

Para cualquier combinación de arquitectura del PMC, problema de prueba y función de error, se tiene una correspondiente superficie de error con  $w+1$  dimensiones para un PMC con  $w$  pesos. Debido a esta característica, resulta un problema determinar la forma precisa de la superficie de error del PMC para un problema dado. Producir un mapa detallado de una superficie de error resultaría impráctico y extenso aún para pequeños problemas de entrenamiento y arquitecturas.

Por lo anterior, en el entrenamiento del PMC se considera como parámetro de control las curvas de error global de entrenamiento para sucesivas iteraciones o épocas. Es común decir que un PMC ha quedado atrapado en un mínimo local cuando la curva de entrenamiento se allana para un nivel comparativamente alto de error. Sin embargo, existen un número de causas igualmente posibles para este comportamiento que no están relacionadas a un mínimo local, por ejemplo: la red parece estar convergiendo a un punto silla; el problema tiene residuos largos en la solución; la red está lentamente cruzando una meseta; la red está zigzagueando bajo una superficie fondeada (valle angosto): uno o más pesos de la red están comenzando a saturarse; o la red está tomando pasos muy pequeños para cada época. Debido a esta multitud de causas diferentes es improbable que exista una sola estrategia efectiva que pueda mejorar una ejecución de entrenamiento pobre o tomar en cuenta todos los estos detalles

Las mesetas, valle estrechos y mínimos locales en la superficie de error son obstáculos potenciales para un entrenamiento favorable. Los mínimos locales tienen correlaciones físicas en términos de entrenamiento del PMC, como son: *nodos ocultos redundantes*, *nodos ocultos saturados* y *regiones muertas del espacio de pesos* (todos los nodos ocultos están inactivos). Sin embargo, la existencia de mínimos locales en la superficie de error no significa que la red es común a sufrir dificultades en el entrenamiento, debido a que el mínimo local puede ser pequeño y esparcirse. Las características de las superficies de error del PMC dan una buena indicación de las clases de estrategias que son comunes para aplicar entrenamientos eficientes y confiables (Shepherd, 1997), por ejemplo

- La suavidad de la superficie de error sugiere que la optimización clásica con derivadas será efectiva.
- Algunos métodos clásicos no están muy lejos de provocar un progreso lento en *valles estrechos*.
- Los efectos de la precisión de punto flotante son comúnmente importantes en regiones planas cercanas.
- Si existen mínimos locales, las estrategias de *minimización global* pueden ser necesarias para garantizar una aceptable probabilidad de eficiencia de entrenamiento

La *minimización global* es un concepto extenso que abarca una amplia gama de estrategias para adaptar los cambios de convergencia a un mínimo global, más que a uno local. Se han propuesto muchas aproximaciones para reducir la probabilidad de que un PMC quede atrapado en una cuenca de atracción de un mínimo local, como por ejemplo: métodos estocásticos, estrategias determinísticas, cambiando la función de error, esquemas de inicialización de pesos, esquemas de cambios dinámicos del número de nodos ocultos, y métodos de tuneleo (Shepherd, 1997; Wessels y Barnard, 1992; Hirose et al., 1991)

#### 2.4.4 Retropropagación

El algoritmo de retropropagación es de los métodos más usados de entrenamiento del PMC, el cual es una aproximación al algoritmo del descenso más pronunciado (sección 2.3 2.1) en donde el índice de ejecución es el *error medio cuadrático*. Para una red lineal de una sola capa (ver fig. 2.2) el error es una función explícita de los pesos de la red, y fácilmente pueden calcularse sus derivadas con respecto a los pesos. En redes multi-capa, con funciones de transferencia no lineales, la relación entre los pesos de la red y el error es más compleja.

##### 2.4.4.1 Algoritmo básico

En el algoritmo se proporciona un conjunto de ejemplos de conducta apropiada de la red (aprendizaje supervisado):

$$\{p_1, t_1\}, \{p_2, t_2\}, \dots, \{p_q, t_q\}$$

donde  $\mathbf{p}_q$  es la entrada de la red, y  $\hat{t}_q$  es la salida deseada para el patrón  $q$  correspondiente. Considerando como índice de ejecución el error medio cuadrático y como regla de aprendizaje la del descenso más pronunciado, el algoritmo de retropropagación básico se puede resumir en los siguientes pasos.

**a) Propagación hacia adelante de las entradas a través de la red**

Para facilitar la explicación del algoritmo se empleará la notación abreviada (matricial) que se muestra en la red multi-capas (tres capas) de la fig. 2.13. Las ecuaciones que describen la regla de propagación de la información hacia adelante son:

$$\mathbf{a}_{m-1} = \mathbf{f}_{m-1}(\mathbf{W}_{m-1}\mathbf{a}_m + \mathbf{b}_{m-1}) \quad m = 0, 1, \dots, M-1 \quad (2.14)$$

donde  $M$  es el número de capas en la red. Las neuronas de la primera capa reciben las entradas externas ( $\mathbf{a}_0 = \mathbf{p}$ ) que son las condiciones iniciales para la expresión (2.14). Las salidas de las neuronas en la última capa son consideradas las salidas de la red ( $\mathbf{a} = \mathbf{a}_M$ ).

**b) Propagar las sensibilidades hacia atrás a través de la red:**

Las sensibilidades del índice de ejecución  $\hat{F}$  (error medio cuadrático) se definen como los cambios en el  $i$ -ésimo elemento de la entrada neta a la capa  $m$ , con la siguiente expresión:

$$\mathbf{s}_m \equiv \frac{\partial \hat{F}}{\partial \mathbf{n}_{i,m}} \quad (2.15)$$

donde:

$$\mathbf{n}_{i,m} = \sum_{j=1}^{s_{m-1}} \mathbf{w}_{i,j,m} \mathbf{a}_{j,m-1} + \mathbf{b}_{i,m} \quad (2.16)$$

El proceso para calcular la sensibilidad  $\mathbf{s}_m$  consiste en una relación recurrente donde la sensibilidad en la capa  $m$  se calcula a partir de la sensibilidad de la capa  $m+1$ , es decir, la sensibilidad se propaga hacia atrás, de ahí el concepto de *retropropagación*. El algoritmo de *retropropagación* emplea la misma técnica de aproximación del algoritmo del *descenso más pronunciado*, siendo la única diferencia que para calcular el gradiente se necesitan primero retropropagar las sensibilidades. Las ecuaciones que describen lo anterior, partiendo de la capa final, son:

$$\mathbf{s}_M = -2\dot{\mathbf{F}}_M(\mathbf{n}_M)(\hat{t} - \mathbf{a}) \quad (2.17)$$

y en las capas ocultas:

$$\mathbf{s}_m = \dot{\mathbf{F}}_m(\mathbf{n}_m) (\mathbf{W}_{m-1})^T \mathbf{s}_{m-1}, \quad m = M-1, \dots, 2, 1 \quad (2.18a)$$

siendo:

$$\dot{\mathbf{F}}_m(\mathbf{n}_m) = \begin{bmatrix} \dot{f}_m(n_{1,m}) & 0 & \dots & 0 \\ 0 & \dot{f}_m(n_{2,m}) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \dot{f}_m(n_{s_m,m}) \end{bmatrix} \quad (2.18b)$$

en donde  $f'_m$  es la primera derivada de la función de activación y  $n_{i,m}$  se define con la expresión (2.16). Como se muestra en las expresiones (2.17) y (2.18), las sensibilidades se propagan hacia atrás a través de la red de la última capa a la primera.

$$s_M \rightarrow s_{M-1} \rightarrow \dots \rightarrow s_2 \rightarrow s_1$$

**c) Actualización de los pesos y sesgos**

Finalmente los pesos y sesgos se actualizan usando la aproximación de la regla del descenso más pronunciado.

$$\begin{aligned} W_m(k+1) &= W_m(k) - \alpha s_m (a_{m-1})^T \\ b_m(k+1) &= b_m(k) - \alpha s_m \end{aligned} \tag{2.19}$$

donde  $\alpha$  es el coeficiente de aprendizaje.

**2.5 VARIACIONES DEL ALGORITMO DE RETROPROPAGACIÓN**

La diferencia principal entre una red de una sola capa lineal y redes multi-capas no lineales es la superficie del error medio cuadrático. Mientras la superficie para una red de una sola capa lineal tiene un solo punto mínimo y curvatura constante, la superficie para una red multi-capas puede tener muchos puntos mínimos locales y la curvatura puede variar ampliamente en regiones diferentes del espacio de parámetros o pesos. Por esta razón, será difícil de escoger un *coeficiente de aprendizaje* apropiado para el algoritmo de retropropagación básico (descenso más pronunciado) descrito en la sección anterior. En la superficie de error, algunas regiones pueden ser planas necesitando un coeficiente de aprendizaje grande, mientras que en otras regiones la curvatura es alta, lo que requeriría un coeficiente de aprendizaje pequeño. En la fig. 2.14 se puede observar el efecto de escoger un coeficiente de aprendizaje muy grande para este algoritmo en la solución de un problema determinado (Hagan, Demuth y Beale, 1996), donde se grafica una sección de la superficie de error al variar únicamente dos parámetros y dejar fijas las demás condiciones (óptimas de un problema conocido). se muestra que en zonas planas el algoritmo es inestable

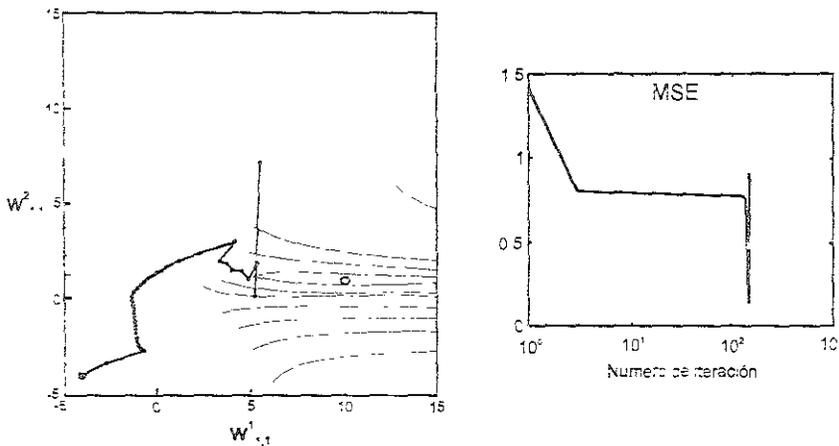


Fig 2.14 Trayectoria de convergencia con coeficiente de aprendizaje muy grande

También, de las características de las superficies de parámetros de redes multi-capa se tienen algunos inconvenientes al proporcionar las *condiciones iniciales de pesos* para el algoritmo de retropropagación básico (sección 2.3.1) Primero, no se deben fijar los pesos iniciales a cero, debido a que el origen del espacio de parámetros tiende a ser un punto silla en la superficie de error. Segundo, no se deben dar a los pesos iniciales valores grandes, porque existen regiones planas en la superficie de error.

Además, el algoritmo de retropropagación básico resulta ser demasiado lento en la ejecución del aprendizaje para la mayoría de las aplicaciones prácticas en RNA, por lo cual se han derivado variaciones de este algoritmo para proporcionar una mayor rapidez de ejecución, pero que básicamente conservan el proceso de propagar el error hacia atrás (las derivadas se procesan de la última capa de la red a la primera) por lo que siguen llamándose algoritmos de retropropagación. Los métodos que pueden acelerar la convergencia de este algoritmo se pueden dividir principalmente en dos categorías.

- a) El desarrollo de técnicas heurísticas o empíricas.
- b) Adaptación de técnicas de optimización numéricas más eficientes.

Las diferencias entre los diferentes algoritmos que existen se presentan en la forma como se usan las derivadas para actualizar los pesos.

## 2.5.1 Modificaciones heurísticas

A continuación se presentan algunos procedimientos por mejorar el algoritmo de retropropagación básico aplicando métodos heurísticos.

### 2.5.1.1 Aplicación de Momentum

El momentum es una modificación que se basa en mejorar la convergencia si se suavizan las oscilaciones de la trayectoria de búsqueda (ver fig. 2.10), lo cual es posible con un filtro pasa-bajas. Un ejemplo del efecto de un filtro de primer orden (Hagan, Demuth y Beale, 1996) se muestra en la fig. 2.15, donde  $w(k)$  es la entrada al filtro (onda seno),  $y(k)$  es la salida del filtro y  $\gamma$  es el *coeficiente de momentum* que debe satisfacer:  $0 < \gamma < 1$ . En la gráfica se puede observar que el filtro tiende a reducir la cantidad de oscilación, mientras aún se rastrea el valor medio

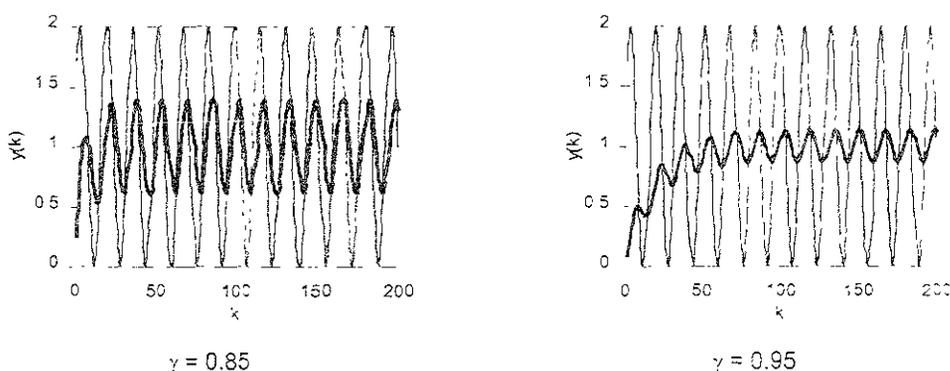


Fig 2.15 Efecto del Suavizado con Momentum

En una red neuronal, recordando primero que la regla de actualización de pesos y sesgo del algoritmo de retropropagación básico está dada por las ecuaciones (2.19), cuando el filtro de momentum se agrega a la regla de aprendizaje, se obtienen las ecuaciones:

$$\begin{aligned}\Delta W^m(k) &= \gamma \Delta b^m(k-1) - (1-\gamma) \alpha s^m (a^{m-1})^T \\ \Delta b^m(k) &= \gamma \Delta b^m(k-1) - (1-\gamma) \alpha s^m\end{aligned}\quad (2.20)$$

Aplicando estas expresiones modificadas se tiene una trayectoria de ejecución como la mostrada en la fig. 2.16, la cual corresponde a las mismas condiciones iniciales y el mismo coeficiente de aprendizaje del ejemplo mostrado en la fig. 2.14, pero ahora considerando un coeficiente de momentum,  $\gamma = 0.8$ . El algoritmo se estabiliza aún empleando un coeficiente de aprendizaje más grande, y además se tiende a acelerar la convergencia cuando la trayectoria está entrando en una dirección consistente.

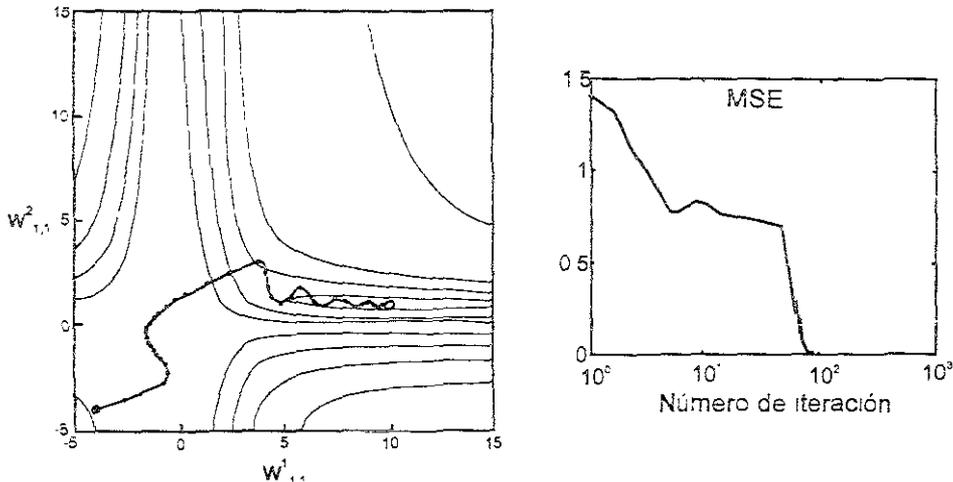


Fig 2.16 Trayectoria de convergencia con Momentum

### 2.5.1.2 Coeficiente de aprendizaje variable

Es posible acelerar la convergencia del método al aumentar el coeficiente de aprendizaje en superficies planas y disminuirlo en regiones de curvatura alta. Sin embargo, la superficie de error para una red multicapa es muy diferente en varias regiones del espacio parámetros como se mencionó al inicio de esta sección. A continuación se describe un procedimiento grupal muy directo donde el coeficiente de aprendizaje se ajusta durante la ejecución del algoritmo. Las reglas del algoritmo de retropropagación con coeficiente de aprendizaje variable son:

1. Si el error cuadrático (sobre el conjunto total de entrenamiento) aumenta más que algún porcentaje fijo  $\zeta$  (entre 1 y 5%) después de una actualización de pesos, entonces tal actualización de pesos se desecha, y el coeficiente de aprendizaje se multiplica por un factor  $\rho$  (entre 0 y 1) y el coeficiente de momentum (si se utiliza) se fija en cero.
2. Si el error cuadrático decrece después de una actualización de pesos, entonces tal actualización de pesos se acepta y el coeficiente de aprendizaje se multiplica por algún factor  $\eta > 1$ . Si  $\gamma$  se ha fijado en cero previamente se restablece a su valor original.
3. Si el error cuadrático aumenta menos que  $\zeta$ , entonces la actualización de pesos se acepta pero no se modifica el valor del coeficiente de aprendizaje. Si previamente se fijó en cero se restablece a su valor de original.

La trayectoria de ejecución del algoritmo de retropropagación usando un coeficiente de aprendizaje variable es como la mostrada en la fig. 2.17a, con un coeficiente de aprendizaje inicial y momentum similares a los del ejemplo presentado en la fig. 2.16: con valores de  $\eta = 1.05$ ,  $\rho = 0.7$  y  $\zeta = 4\%$ . Esto muestra cómo el tamaño del paso (coeficiente de aprendizaje) tienden a aumentar cuando la trayectoria está viajando en una línea recta con error decreciente constante. Este mismo efecto se puede observar en la fig. 2.17b donde se grafica el error cuadrático y el coeficiente de aprendizaje contra el número iteraciones o épocas. Cuando la trayectoria de ejecución llega a un valle estrecho, el coeficiente de aprendizaje disminuye rápidamente si el error aumenta más de 4%. Luego el coeficiente de aprendizaje aumenta y se acelera la convergencia; pero al rebasar el punto mínimo (cuando el algoritmo casi ha convergido) el coeficiente se reduce de nuevo.

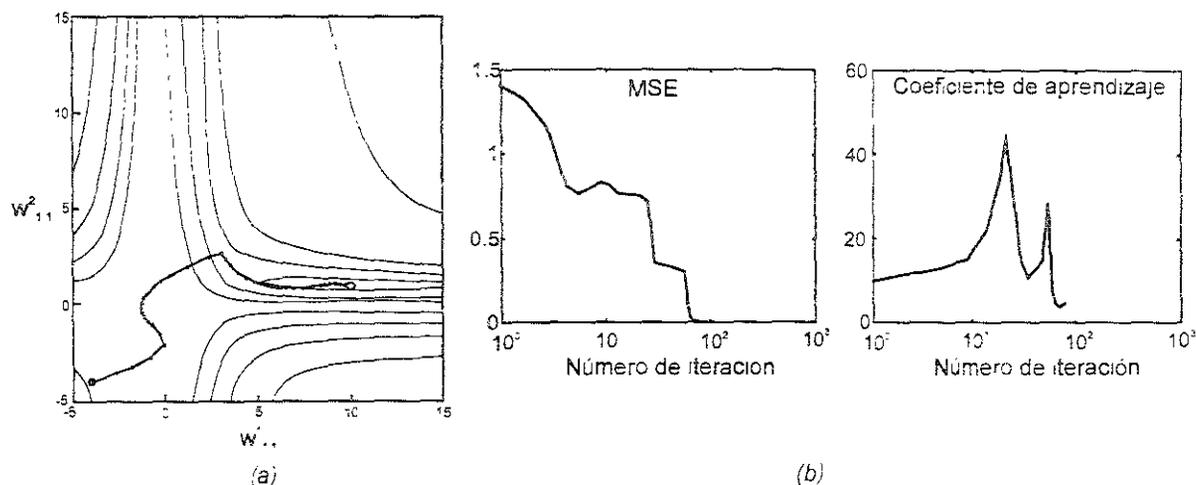


Fig 2.17 (a) Trayectoria y (b) características de convergencia con coeficiente de aprendizaje variable

Existen muchas variaciones a este algoritmo para ajustar el coeficiente de aprendizaje en los trabajos presentados por Jacobs (1988), Tollenaere (1990) y Fahlman (1988). Las modificaciones heurísticas al algoritmo de retropropagación básico pueden resolver más rápido algunos problemas; sin embargo existen dos desventajas principales a estos métodos empíricos:

- Las modificaciones requieren fijar varios parámetros (p.e.  $\eta$ ,  $\rho$  y  $\zeta$ ), lo cual afecta de alguna forma la ejecución del algoritmo.
- A veces estas modificaciones no pueden converger en problemas para los que el algoritmo básico tiene una solución.

## 2.5.2 Técnicas de optimización numéricas

A continuación se presentan algunas adaptaciones de los *métodos de optimización de segundo orden* (presentados en la sección 2.3.2), con el objetivo de mejorar la ejecución del algoritmo de retropropagación básico (del descenso más pronunciado) para el entrenamiento de redes PMC.

### 2.5.2.1 Gradiente conjugado

El algoritmo de gradiente conjugado, como se mencionó en la sección 2.3.2.3, no requiere el cálculo de las segundas derivadas, manteniendo la propiedad de convergencia para una función cuadrática. La adaptación de esta técnica de optimización al entrenamiento de redes PMC se conoce como algoritmo retropropagación de gradiente conjugado (CGBP).

Como se ha mencionado, el algoritmo de gradiente conjugado (GC) consiste en los siguientes pasos

1. Seleccionar como dirección de búsqueda inicial al gradiente negativo.  $\mathbf{p}_0 = -\mathbf{g}_0$ . Recordando que el gradiente  $\mathbf{g}$  se determina con la expresión (2.4)
2. Tomar un paso  $\Delta \mathbf{x}_k = \alpha_k \mathbf{p}_k$ , seleccionando el coeficiente de aprendizaje  $\alpha_k$  para minimizar la función a lo largo de la dirección de búsqueda, de la forma  $\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta \mathbf{x}_k$
3. Seleccionar una dirección de búsqueda:

$$\mathbf{p}_k = -\mathbf{g}_k - \beta_k \mathbf{p}_{k-1} \tag{2.21}$$

escogiendo un  $\beta_k$  con cualquiera de las fórmulas mostradas en la expresión (2.13).

4. Si no ha convergido el algoritmo, regresar al paso 2.

Este algoritmo de GC no puede aplicarse directamente a la tarea de ejecución de RNA, debido a que el índice de ejecución no es cuadrático, lo cual afecta al algoritmo de dos formas: no se puede usar la ecuación (2.9b) para minimizar la función a lo largo de una línea (paso 2); y el mínimo exacto normalmente no se alcanza en un número finito de pasos y por consiguiente el algoritmo necesitará ser restablecido después de algún número de iteraciones fijo. De esta forma, primero se necesita un procedimiento general para localizar el mínimo de una función en una dirección específica, lo cual involucra dos pasos: *localizar el intervalo* y *reducirlo*.

Para *localizar el intervalo* donde se encuentra un mínimo se utiliza un método de comparación de funciones, el cual se ilustra en la fig. 2.18a. Primero se evalúa el índice de ejecución en el punto inicial  $\mathbf{a}_1$  (que corresponde a los valores actuales de pesos y sesgos de la red), es decir se evalúa para  $F(\mathbf{x}_0)$ . El siguiente paso es evaluar la función en un segundo punto  $\mathbf{b}_1$ , que está a una distancia  $\epsilon$  del punto inicial a lo largo de la primera dirección de búsqueda  $\mathbf{p}_0$ , es decir se evalúa para  $F(\mathbf{x}_0 + \epsilon \mathbf{p}_0)$ . Este procedimiento se realiza sucesivamente para nuevos puntos  $\mathbf{b}_i$  doblando la distancia entre los puntos, el cual se detiene en el momento que la función aumenta entre dos evaluaciones consecutivas, entre  $(\mathbf{a}_3, \mathbf{b}_3)$  y  $(\mathbf{a}_4, \mathbf{b}_4)$ ; lo que indica que el mínimo puede encontrarse en el intervalo  $(\mathbf{a}_5, \mathbf{b}_5)$ . Una vez localizado el intervalo, el próximo paso es su *reducción*, lo cual se consigue al evaluar la función dentro del intervalo seleccionado. Es necesario evaluar la función por lo menos en dos puntos interiores, como se muestra en la fig. 2.18b, donde  $c$  y  $d$  son los puntos a evaluar. Si  $F(\mathbf{c}) > F(\mathbf{d})$ , entonces el mínimo debe ocurrir en el intervalo  $(\mathbf{c}, \mathbf{b})$ ; y viceversa si  $F(\mathbf{c}) < F(\mathbf{d})$ , entonces el mínimo debe ocurrir en el intervalo  $(\mathbf{a}, \mathbf{d})$ .

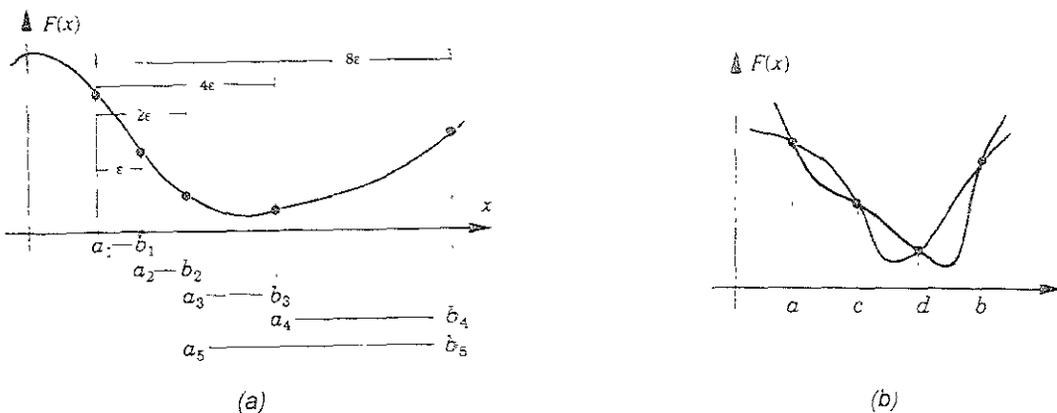


Fig 2.18 (a) Localización del intervalo y (b) reducción del intervalo

El procedimiento anterior requiere de un método para seleccionar los puntos **c** y **d**. Existen varias maneras; sin embargo, la forma más común es aplicar el método de búsqueda de la *Sección Dorada* (Hagan, Demuth y Beale, 1995) diseñado para reducir el número de evaluaciones de una función dada, donde en cada iteración se requiere una nueva evaluación de la función. Es decir, en el caso ilustrado en la fig. 2.18b el punto **a** se descartaría y **c** llegaría a ser un punto externo, entonces un nuevo punto **c** se pondría entre los puntos originales **c** y **d**.

Antes de aplicar el método de GC al algoritmo de entrenamiento de RNA, se requiere de otra modificación más. Si el índice de ejecución del error medio cuadrático para redes multi-capa no es cuadrático, el algoritmo no convergerá al mínimo en  $n$  iteraciones (donde  $n$  es el número de parámetros para optimizar). El método más simple para solucionar esta situación es restablecer la dirección de búsqueda empleando la dirección de descenso más pronunciado (pendiente negativa) después de las  $n$  iteraciones.

Finalmente, se tiene el algoritmo en modo grupal, es decir, se utiliza el algoritmo de retropropagación para calcular el gradiente (ecuación (2.15)) y el algoritmo de gradiente conjugado para actualizar los pesos y sesgos. En la fig. 2.19 se presenta la trayectoria de convergencia aplicando el algoritmo de retropropagación de gradiente conjugado para las mismas condiciones iniciales del ejemplo mostrado en la fig. 2.16. Se observa cómo este algoritmo grupal converge en pocas iteraciones en comparación con los otros algoritmos que se han presentado hasta el momento; sin embargo, en cada iteración se requieren más procesos de cómputo que los métodos anteriores.

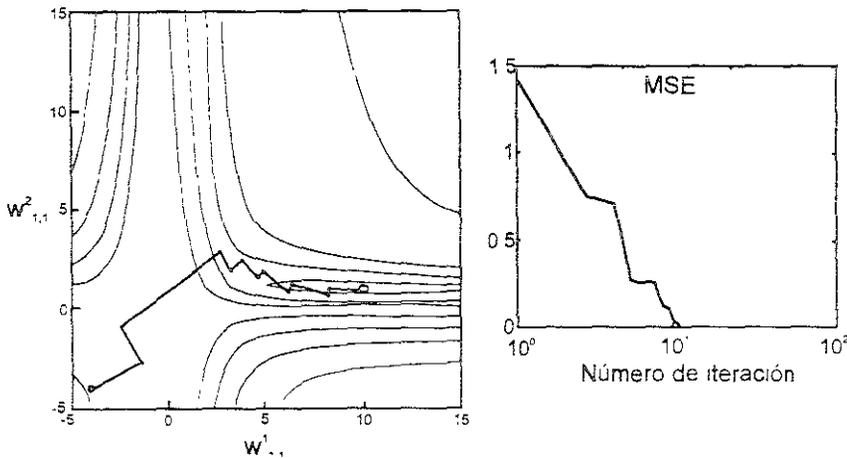


Fig 2.19 Trayectoria de Gradiente Conjugado

### 2.5.2.2 Algoritmo Levenberg-Marquardt

El algoritmo de Levenberg-Marquardt es una variación del método de Newton para entrenar una red PMC donde el índice de ejecución es el error medio cuadrático. El proceso iterativo del algoritmo de retropropagación con Levenberg-Marquardt (LMBP) se muestra a continuación.

1. Se presenta a la red la información  $p_q$  y  $t_q$  (entradas/salidas objetivo). Se calculan las salidas correspondientes de la red con la ecuación (2.14) y los errores por medio de la expresión:

$$e_q = t_q - a_{q,M}$$

Se determina la suma de los errores cuadráticos sobre todas las entradas usando la siguiente ecuación.

$$F(\mathbf{x}) = \sum_{q=1}^Q (\mathbf{t}_q - \mathbf{a}_q)^T (\mathbf{t}_q - \mathbf{a}_q) = \sum_{q=1}^Q \mathbf{e}_q^T \mathbf{e}_q = \sum_{q=1}^Q \sum_{j=1}^{S_M} (\mathbf{e}_{j,q})^2 = \sum_{i=1}^I (v_i)^2 \quad (2.22)$$

donde  $\mathbf{e}_{j,q}$  es el  $j$ -ésimo elemento del error para el  $q$ -ésimo patrón entrada-salida deseada, y  $Q$  el número total de patrones.

2. Después, se calcula la matriz Jacobiana con la expresión:

$$J(\mathbf{x}) = \begin{bmatrix} \frac{\partial \mathbf{e}_{1,1}}{\partial \mathbf{w}_{1,1}} & \frac{\partial \mathbf{e}_{1,1}}{\partial \mathbf{w}_{1,2}} & \dots & \frac{\partial \mathbf{e}_{1,1}}{\partial \mathbf{w}_{S_1,R,1}} & \frac{\partial \mathbf{e}_{1,1}}{\partial \mathbf{b}_{1,1}} & \dots \\ \frac{\partial \mathbf{e}_{2,1}}{\partial \mathbf{w}_{1,1}} & \frac{\partial \mathbf{e}_{2,1}}{\partial \mathbf{w}_{1,2}} & \dots & \frac{\partial \mathbf{e}_{2,1}}{\partial \mathbf{w}_{S_1,R,1}} & \frac{\partial \mathbf{e}_{2,1}}{\partial \mathbf{b}_{1,1}} & \dots \\ \vdots & \vdots & & \vdots & \vdots & \\ \frac{\partial \mathbf{e}_{S_M,1}}{\partial \mathbf{w}_{1,1}} & \frac{\partial \mathbf{e}_{S_M,1}}{\partial \mathbf{w}_{1,2}} & \dots & \frac{\partial \mathbf{e}_{S_M,1}}{\partial \mathbf{w}_{S_1,R,1}} & \frac{\partial \mathbf{e}_{S_M,1}}{\partial \mathbf{b}_{1,1}} & \dots \\ \frac{\partial \mathbf{e}_{1,2}}{\partial \mathbf{w}_{1,1}} & \frac{\partial \mathbf{e}_{1,2}}{\partial \mathbf{w}_{1,2}} & \dots & \frac{\partial \mathbf{e}_{1,2}}{\partial \mathbf{w}_{S_1,R,1}} & \frac{\partial \mathbf{e}_{1,2}}{\partial \mathbf{b}_{1,1}} & \dots \\ \frac{\partial \mathbf{e}_{1,2}}{\partial \mathbf{w}_{1,1}} & \frac{\partial \mathbf{e}_{1,2}}{\partial \mathbf{w}_{1,2}} & \dots & \frac{\partial \mathbf{e}_{1,2}}{\partial \mathbf{w}_{S_1,R,1}} & \frac{\partial \mathbf{e}_{1,2}}{\partial \mathbf{b}_{1,1}} & \dots \\ \vdots & \vdots & & \vdots & \vdots & \end{bmatrix} \quad (2.23)$$

donde el tercer subíndice corresponde al número de capa de la red. Se determinan las sensibilidades, partiendo de la última capa, con las relaciones de recurrencia:

$$\tilde{\mathbf{s}}_{q,M} = -\dot{F}_M(\mathbf{n}_{q,M}) \quad (2.24a)$$

y para las capas ocultas

$$\tilde{\mathbf{s}}_{q,m} = \dot{F}_m(\mathbf{n}_{q,m}) (\mathbf{w}_{m-1})^T \tilde{\mathbf{s}}_{q,m-1} \quad (2.24b)$$

donde  $F_m(\mathbf{n}_m)$  se obtiene de la expresión (2.18b). Luego, se aumentan las matrices individuales de sensibilidad de Marquardt usando la expresión:

$$\tilde{\mathbf{S}}_m = [\tilde{\mathbf{S}}_{1,m} | \tilde{\mathbf{S}}_{2,m} | \dots | \tilde{\mathbf{S}}_{q,m}] \quad (2.25)$$

Se calculan los elementos de la matriz Jacobiana con las ecuaciones, para los pesos:

$$[J]_{h,i} = \frac{\partial v_h}{\partial x_i} = \frac{\partial \mathbf{e}_{k,q}}{\partial \mathbf{n}_{i,q,m}} \left( \frac{\partial \mathbf{n}_{i,q,m}}{\partial \mathbf{w}_{i,j,m}} \right) = \tilde{\mathbf{s}}_{i,h,m} \left( \frac{\partial \mathbf{n}_{i,q,m}}{\partial \mathbf{w}_{i,j,m}} \right) = \tilde{\mathbf{s}}_{i,h,m} (\mathbf{a}_{j,q,m-1}) \quad (2.26a)$$

y para los sesgos:

$$[J]_{h,i} = \frac{\partial v_h}{\partial x_i} = \tilde{\mathbf{s}}_{i,h,m} \left( \frac{\partial \mathbf{n}_{i,q,m}}{\partial \mathbf{b}_{i,m}} \right) = \tilde{\mathbf{s}}_{i,h,m} \quad (2.26b)$$

siendo:

$$\tilde{\mathbf{s}}_{i,h,m} = \frac{\partial v_h}{\partial \mathbf{n}_{i,q,m}} = \frac{\partial \mathbf{e}_{k,q}}{\partial \mathbf{n}_{i,q,m}}$$

donde  $h = (q-1) S_M + k$

3. Posteriormente, para obtener  $\Delta \mathbf{x}_k$  se resuelve la siguiente ecuación

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [\mathbf{J}^T(\mathbf{x}_k)\mathbf{J}(\mathbf{x}_k) + \mu_k \mathbf{I}]^{-1} \mathbf{J}^T(\mathbf{x}_k) \mathbf{v}(\mathbf{x}_k) \quad (2.27)$$

donde  $\mathbf{v}$  es el vector de error:

$$\mathbf{v}^T = [v_1, v_2, \dots, v_N] = [e_{1,1}, e_{2,1}, \dots, e_{S_M,1}, e_{1,2}, \dots, e_{S_M,2}]$$

y  $\mathbf{x}$  el vector de parámetros:

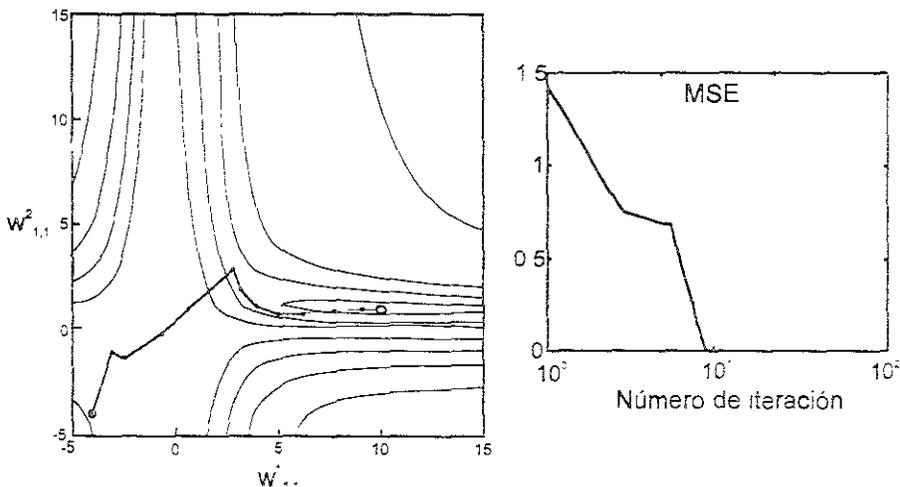
$$\mathbf{x}^T = [x_1, x_2, \dots, x_n] = [w_{1,1}, w_{1,2}, \dots, w_{S,R}, b_{1,1}, \dots, b_{S,1}, w_{1,2}, \dots, b_{S,M}]$$

siendo  $N = Q \times S_M$  y  $n = S_1(R+1) + S_2(S_1+1) + \dots + S_M(S_{M-1}+1)$

4. Por último, se recalcula la suma de los errores cuadráticos usando  $\mathbf{x}_k + \Delta \mathbf{x}_k$ . Si esta nueva suma de cuadráticos es más pequeña que la calculada en el paso 1, entonces se divide  $\mu$  por  $\nu$ , dejando  $\mathbf{x}_k + \Delta \mathbf{x}_k$  y regresamos al paso 1. Si la suma de cuadráticos no se reduce, entonces se multiplica  $\mu$  por  $\nu$ , y regresamos al paso 3.

Se considera que el algoritmo converge cuando la norma del gradiente,  $\mathbf{g} = 2 \mathbf{J}^T(\mathbf{x}) \mathbf{v}(\mathbf{x})$ , es menor que algún valor predeterminado, o cuando la suma de cuadráticos ha llegado a algún error fijado.

La fig. 2.20 muestra la trayectoria de convergencia del algoritmo LMBP, con  $\mu = 0.01$  y  $\nu = 5$ , para las mismas condiciones iniciales del ejemplo presentado en la fig. 2.14. Se observa que el método converge en menos iteraciones que cualquiera de las técnicas presentadas. Sin embargo, este algoritmo también requiere más proceso de cómputo por iteración que cualquiera de los otros métodos, debido a que se requiere calcular la inversa de una matriz (ecuación (2.27)).



La Fig 2.20 Trayectoria de convergencia para el algoritmo de LMBP

El inconveniente más importante del algoritmo LMBP es el almacenamiento requerido para guardar la matriz del Hessiano aproximada  $\mathbf{J}^T \mathbf{J}$ . Por tal razón, resulta poco práctico el uso del algoritmo de Levenberg-Marquardt cuando el número de parámetros es muy grande.

---

## CAPÍTULO 3

### RESPUESTA SÍSMICA DEL SITIO

---

En este capítulo se mencionan los factores principales que afectan la respuesta sísmica del terreno. Se describen algunas de las técnicas más usadas para determinar la respuesta sísmica de los depósitos de suelos y el efecto que tienen las condiciones locales del sitio en los movimientos de campo libre dentro del valle de México.

#### 3.1 INTRODUCCIÓN

Uno de los temas principales dentro de la ingeniería sísmica geotécnica es la determinación de la respuesta dinámica del terreno; con el objetivo de desarrollar, por ejemplo, espectros de respuesta para diseño, determinar esfuerzos y deformaciones dinámicas para la evaluación de problemas de licuación de suelos y para obtener las fuerzas sísmicas que pueden ocasionar la inestabilidad de estructuras de tierra y de retención. La evaluación de la respuesta sísmica en un sitio específico involucra aspectos como la identificación de los tipos de fallas potencialmente activas donde se originan los sismos, la evaluación de la sismicidad asociada con fuentes individuales, la forma como se propagan las ondas sísmicas de la fuente de origen a un sitio de roca particular, la determinación de la respuesta sísmica de los depósitos de suelo y de sistemas suelo-estructura, y la evaluación de la estabilidad sísmica de tales estructuras ante un ambiente sísmico de diseño. Los primeros tres aspectos se relacionan principalmente con procesos geológicos y geofísicos que determinan el medio ambiente sísmico en el terreno firme de la localidad. Sin embargo, a pesar de que las ondas sísmicas viajan a través de varios kilómetros de roca y sólo unas decenas de metros en el suelo, en la mayoría de los casos son los depósitos de suelo el factor más importante que afecta las características de los movimientos del terreno. Por lo cual, el problema de respuesta dinámica en este capítulo se reduce básicamente a evaluar el efecto que tienen las condiciones locales del sitio sobre las características de las ondas sísmicas incidentes.

Las condiciones del suelo y las características geológicas locales que afectan la respuesta del sitio son numerosas, algunas de las más importantes se presentan a continuación de acuerdo con la fig. 3.1.

1. Mientras mayor sea la extensión horizontal ( $L_1$  o  $L_2$ ) de los suelos más blandos, menores serán los efectos de frontera (limitrofes) de la roca subyacente en la respuesta del sitio.
2. La profundidad del suelo ( $H_1$  o  $H_2$ ) que sobreyace al lecho de roca afecta la respuesta dinámica debido a que el periodo natural de vibración del suelo se incrementa conforme aumenta la profundidad, lo cual incide en las frecuencias de ondas que son filtradas por el depósito de suelo.

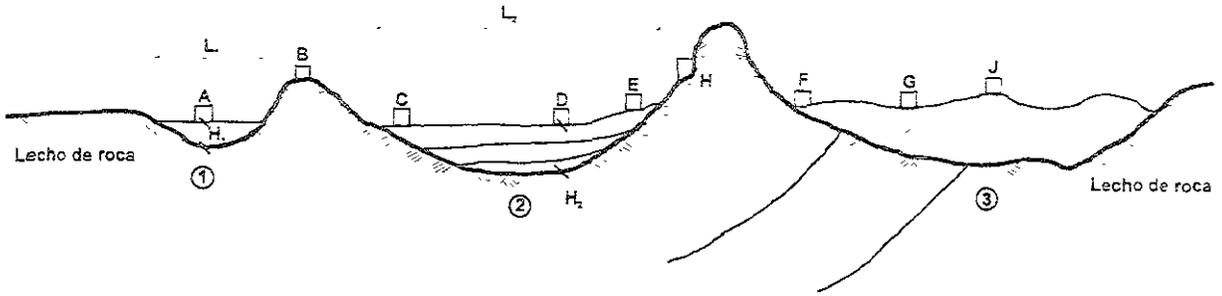


Fig 3.1 Diagrama esquemático mostrando la geología local y las características del suelo

- 3 La pendiente de los planos de estratificación (valles 2 y 3 en la fig 3.1) afecta la respuesta; sin embargo, el análisis del problema se complica.
4. Los cambios horizontales en el tipo de suelo a lo largo de un sitio (F y G en la fig 3.1), afectan en forma local la respuesta dentro del sitio, así como la seguridad de un edificio desplantado en dos tipos de suelos.
5. La topografía del lecho de roca y los depósitos de suelo, ocasionan varios efectos sobre las ondas sísmicas que llegan, tales como reflexión, refracción, enfoque y difracción. Muchos de estos efectos se consideran como hipótesis; por ejemplo, es posible predecir confiablemente los efectos debido a las características presentadas en los valles 1 y 2 para los sitios A y D, lo cual no es posible para los sitios G y J, por las condiciones del valle 3 (fig. 3.1). Bien puede ser que las características geológicas, tales como irregularidades ocultas en la topografía del lecho de roca, expliquen las diferencias entre las respuestas de dos sitios con igual estratigrafía.
6. Las cordilleras (sitio B en la fig. 3.1) son otra característica topográfica que afecta la respuesta, donde puede presentarse una amplificación del movimiento del lecho de roca.
7. Las pendientes de los depósitos sedimentarios pueden ocasionar una falla durante un sismo. En terrenos escarpados (como en el sitio H), la falla puede ser en forma catastrófica. Sin embargo, en suelos estratificados con pendientes suaves también existe la posibilidad de falla si se tienen estratos propensos a licuación (sitio E).
8. El contenido de agua del suelo es un factor importante en la respuesta del sitio, lo cual se aplica no solamente a suelos estratificados con pendientes suaves, sino que el efecto de licuación puede ocurrir en un terreno plano compuesto de suelos no cohesivos saturados.

En resumen, el efecto que tienen las condiciones locales del sitio en los movimientos sísmicos de roca basal, dependen de aspectos sismológicos, condiciones geológicas, características geotécnicas y peculiaridades geométricas del sitio, tales factores se presentan en la Tabla 3.1

### 3.1.1 Efectos de la Topografía de Superficie y la Geometría de la Cuenca

Aunque resulta difícil de evaluar el efecto de las irregularidades topográficas y la geometría de una cuenca aluvial en la respuesta sísmica, existen algunos casos que se basan en análisis teóricos y mediciones reales que pueden de alguna manera proporcionar un panorama más real de cómo son afectados los movimientos del terreno por estos factores.

Tabla 3.1 Principales factores que influyen en los efectos locales del sitio (Romo, Mendoza y García 2000)

Sismológicos	<ul style="list-style-type: none"> <li>* Intensidad y contenido de frecuencias de los movimientos sísmicos de roca basal</li> <li>* Duración de los movimientos en roca basal</li> </ul>
Geológicos	<ul style="list-style-type: none"> <li>* Estructuras geológicas locales</li> <li>* Tipo de roca subyacente</li> <li>* Espesor de los depósitos de suelo</li> <li>* Características estratigráficas</li> <li>* Tipos de suelos</li> </ul>
Geotécnicos	<ul style="list-style-type: none"> <li>* Características de vibración elástica de los depósitos de suelo</li> <li>* Comportamiento no lineal del suelo</li> <li>* Impedancia relativa entre la roca basal y los depósitos de suelo sobreyacentes</li> </ul>
Geométricos	<ul style="list-style-type: none"> <li>* Depósitos de suelo estratificados no horizontales</li> <li>* Topografía de la interfaz suelo-roca basal</li> <li>* Configuración de la cuenca</li> </ul>

### Topografía

Evaluar el efecto de las irregularidades topográficas es difícil debido a la interacción de ondas que se propagan en el medio, lo cual puede producir modelos complejos de amplificación y atenuación dependiendo, tanto de la geometría de la irregularidad, como de los ángulos de incidencia, frecuencias y tipos de ondas (Sánchez-Sesma y Campillo, 1993). A pesar de lo anterior, los movimientos sísmicos en la cima de lomas o cordilleras rocosas estrechas por lo general se amplifican, como lo muestra el estudio realizado en la presa Pacoima en California del sur durante el sismo de San Fernando en 1971, donde se registraron valores de aceleraciones horizontales mayores a los esperados (Trifunac y Hudson, 1971). Estos mismos efectos se presentaron durante el sismo de Northridge en 1994, donde las amplificaciones en la cresta de la colina de Tarzana fueron mayores debido al efecto topográfico (Bouchon y Barker, 1996). También, un caso característico en términos de daño sísmico se presentó en una villa de Italia durante el sismo de Irpinia en 1980, donde los mayores daños se tuvieron cerca de la cresta de una loma, siendo insignificantes lejos de ella (Castellani, et al., 1982). Otro ejemplo, es el estudio realizado en el Cañón de Feitsui (Taiwan), donde a partir de datos observados y simulaciones numéricas, se concluye que los movimientos del terreno son modificados significativamente por la topografía a lo largo del cañón (Huang y Chiu, 1999).

Los efectos causados por irregularidades topográficas sencillas (lomas de geometría regular) pueden estimarse a partir de soluciones exactas de problemas idealizados (Aki, 1988), donde se demuestra también, que los movimientos en la cima son amplificados. Así también, existen soluciones analíticas para cañones semicirculares, semi-épticos y cilíndricos semiparabólicos (Trifunac, 1973; Wong y Trifunac, 1974; Todorovska y Lee, 1991). Estudios más detallados (Geli et al., 1988; Sánchez-Sesma, 1990; Faccioli, 1991) han considerado varios tipos de geometrías de las irregularidades topográficas y campos de ondas diferentes.

Por lo tanto, la evaluación de los efectos de las irregularidades topográficas del terreno requiere de análisis de respuesta bidimensionales (2-D), y en algunos casos, tridimensionales (3-D), lo que implica mayor tiempo y detalle para la caracterización del sitio. Para este tipo de análisis, se han desarrollado diferentes herramientas de cómputo, como es el caso del programa de respuesta sísmica 2-D, FLUSHPLUSH (Deng, 1993), aplicado a evaluar el efecto de las características topográficas de superficie en la intensidad de los movimientos de base (Athanasopoulos y Leonidou, 1999).

La predicción aproximada de movimientos del terreno en presencia de irregulares topográficas (montaña, cañón, coma, etc.) requiere de simulaciones 3-D, debido a que la respuesta para estos casos depende mucho del azimut, ángulo y tipo de ondas incidentes. Los casos de superficies irregulares de tipo cordillera (o valle) se han tratado en una forma 2-D; sin embargo, solamente un número limitado de resultados están disponibles para configuraciones tridimensionales del problema (Sánchez-Sesma et al., 1994). Estudios realizados mediante el método de elementos de frontera para determinar la dispersión 3-D de las ondas sísmicas de topografías irregulares y valles profundos debido a la propagación de ondas P, S y Rayleigh, han mostrado que las mayores amplificaciones se presentan en montañas muy escarpadas, mientras que en montañas con pendiente suave exhiben poca amplificación de los movimientos sísmicos del terreno (Reinoso, et al., 1997).

### **Cuencas aluviales**

La curvatura de una cuenca aluvial rellena de suelos blandos, puede atrapar las ondas de cuerpo incidentes y causar que algunas ondas se propaguen a través de los depósitos de suelo como ondas de superficie (Graves, 1993; Chávez-García, et al., 1995), causando una prolongación en la duración y amplificación de los movimientos del terreno. Las ondas que quedan atrapadas en una cuenca de depósitos sedimentarios han causado daños considerables como en Caracas (1977), en la ciudad de México (Michoacán, 1985), en Santa Mónica y al oeste de Los Ángeles (1994), y en Kobe (1995).

El efecto de reflexión múltiple de ondas en los límites de la cuenca, junto con el efecto de enfocamiento (concentración de energía en determinados lugares) de ondas de cuerpo causado por la curvatura tridimensional de la interfaz aluvión-basamento (Somerville, 1998), indican la necesidad de utilizar simulaciones 3-D para estos casos. Sin embargo, la respuesta sísmica del terreno se puede predecir por análisis de respuesta unidimensionales (1-D) que consideran la propagación vertical de ondas de corte, si los depósitos de suelo aluvial son de gran extensión y el sitio de análisis no se encuentra cerca de los límites de la cuenca.

Las diferencias observadas entre las funciones de amplificación al centro y en los límites de una cuenca aluvial (Caracas, 1967; San Fernando, 1971, Leninakan, Armenia, 1988) explican porque los movimientos en estos puntos son considerablemente diferentes. También, se ha mostrado la similitud de resultados entre análisis unidimensionales y bidimensionales en la respuesta sísmica al centro de un valle plano poco profundo (Bard y Gariel, 1986). Para los casos de valles aluviales de forma irregular tales como combinaciones de regiones cóncavas-convexas, los estudios teóricos (Rial et al., 1992) indican que los movimientos pueden resultar muy complejos.

## **3.2 EFECTO DE LAS CONDICIONES LOCALES DEL SUELO EN LA RESPUESTA DEL TERRENO**

El efecto de las condiciones locales del suelo pueden ilustrarse principalmente por análisis teóricos de respuesta del terreno, por medición de los movimientos de superficie reales y del subsuelo para un mismo sitio, y por medición de los movimientos del terreno reales en sitios superficiales con condiciones del subsuelo diferentes.

### **3.2.1 Análisis Teóricos de Respuesta Sísmica**

Se han desarrollado métodos cuantitativos para predecir la influencia de las condiciones locales del sitio en los movimientos fuertes del terreno. En general, estos métodos se agrupan de acuerdo a la dimensión del problema, aunque muchas de las técnicas 2-D y 3-D son extensiones relativamente directas de sus correspondientes técnicas unidimensionales.

La mayoría de los métodos teóricos 1-D simples se basan en la suposición que la respuesta en campo libre de un depósito de suelo se debe a la propagación ascendente de las ondas de corte. Los procedimientos analíticos basados en este concepto, como el modelo presentado en la fig 3.2, han mostrado cómo los depósitos de suelos aluviales amplifican los movimientos del lecho de roca

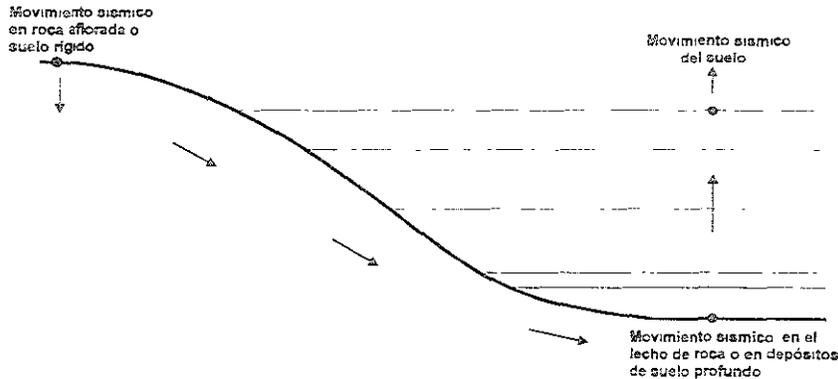


Fig 3.2 Modelo de análisis de respuesta unidimensional para suelos con estratificación horizontal

De esta forma, los métodos de análisis unidimensionales de respuesta del terreno más empleados involucran tres hipótesis básicas:

- La respuesta del depósito de suelo se debe predominantemente a la propagación vertical de ondas de corte desde la formación inferior de roca a la superficie (Kanai, 1952).
- Todas las fronteras entre los diferentes materiales que forman el depósito de suelo, tanto en la superficie como en la roca basal, son esencialmente horizontales y de extensión infinita
- El material de cada estrato de suelo es homogéneo y linealmente elástico o viscoelástico.

Varias investigaciones (Idriss y Seed, 1968; Romo y Seed, 1986, Romo, 1991; Karantoni y Bouckovalas, 1997) han comprobado que las dos primeras hipótesis proporcionan resultados razonables para muchos sitios donde se localizan depósitos de suelo aluvial con estratificación horizontal. Con respecto a la tercera suposición, la linealidad de los estratos de suelo resulta ser inapropiada cuando los movimientos sísmicos son muy intensos, sin embargo, en varios casos se ha observado que el comportamiento no lineal del suelo puede ser tratado de manera práctica aceptable utilizando un método lineal equivalente (Martin, 1975; Seed e Idriss, 1969).

### 3.2.1.1 Análisis lineales de respuesta

Una clase importante de técnicas para evaluar la respuesta sísmica se basa en el uso de *Funciones de Transferencia*, con las cuales se expresan varios parámetros de respuesta tales como: desplazamientos, velocidades, aceleraciones, esfuerzos y deformaciones cortantes. Una función de transferencia relaciona los movimientos de campo libre en la superficie del terreno (respuesta) con los movimientos en la roca basal (excitación) para un sitio específico en particular (depósito de suelo). con lo cual, es posible determinar cómo cada frecuencia en el movimiento del lecho de roca se amplifica, o atenúa, por el depósito del suelo. Como se sabe, esta aproximación se limita al análisis de sistemas lineales únicamente; sin embargo, el comportamiento no lineal puede aproximarse usando un procedimiento iterativo con propiedades del suelo linealmente equivalentes

La aproximación lineal de respuesta sísmica consiste básicamente en evaluar las funciones de transferencia, las cuales se derivan para ciertas condiciones geotécnicas simples y complejas. Estos casos han podido ilustrar algunos de los efectos importantes que los depósitos de suelo tienen en las características del movimiento del terreno.

### Depósitos de suelo homogéneo sobre roca rígida

Considerando un estrato de suelo homogéneo, isotrópico, elástico lineal sobreyaciendo la roca basal rígida como se muestra en la fig. 3.3. Los desplazamientos en cualquier punto dentro del depósito de suelo pueden expresarse como:

$$u(z,t) = A_i e^{i(\omega t - kz)} + A_r e^{i(\omega t + kz)} \quad (3.1)$$

donde  $\omega$  es la frecuencia circular del movimiento del terreno,  $k$  el número de onda ( $\omega/v_s$ ), y  $A_i$  y  $A_r$  son las amplitudes de ondas que viajan en las direcciones  $-z$  (ascendente) y  $+z$  (descendente) respectivamente

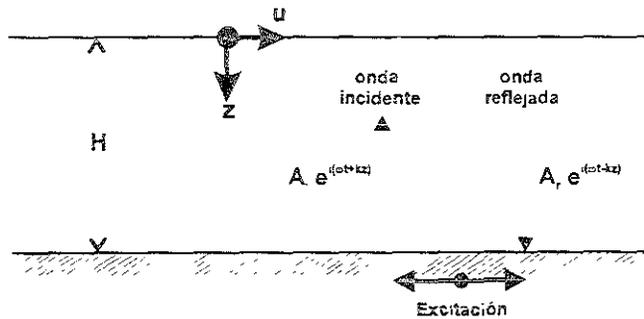


Fig 3.3 Depósito de suelo elástico lineal de espesor  $H$  sobreyaciendo al lecho de roca rígido

Aplicando las condiciones de frontera: en la superficie libre ( $z=0$ ) los esfuerzos y deformaciones cortantes son nulos, y reacomodando la expresión (3.1) la cual se satisface cuando  $A_i = A_r$ , los desplazamientos pueden expresarse entonces:

$$u(z,t) = 2A_i e^{i\omega t} \cos(kz) \quad (3.2)$$

La ecuación (3.2) puede usarse para definir una función de transferencia que describe la relación de amplitudes de desplazamiento en dos puntos cualquiera en la capa del suelo. Seleccionando estos dos puntos en la superficie y en el fondo de la capa de suelo tenemos la función de transferencia:

$$F_t(\omega) = \frac{u_{\max}(0,t)}{u_{\max}(H,t)} = \frac{2A_i e^{i\omega t}}{2A_i \cos(kH) e^{i\omega t}} = \frac{1}{\cos(kH)} = \frac{1}{\cos(\omega H / v_s)} \quad (3.3)$$

El módulo de la función de transferencia es la *función de amplificación*. La expresión (3.3) indica que la amplificación tiende a infinito (*resonancia*) cuando  $\cos(kH)=0$ , lo cual ocurre para

$$k_n H = \frac{\pi}{2} + n\pi$$

puesto que:

$$k_n = \frac{\omega_n}{v_s} = \frac{2\pi f_n}{v_s}$$

donde  $f$  es la frecuencia en Hz y  $v_s$  es la velocidad de ondas de corte entonces se definen las frecuencias naturales del depósito del suelo como:

$$f_n = \frac{v_s}{4H}(2n+1) \quad n = 0, 1, 2, \dots \quad (3.4)$$

Este modelo muy simple nos muestra que la respuesta de un depósito de suelo depende mucho de la frecuencia del movimiento en la base, y que las frecuencias en las que ocurren amplificaciones fuertes dependen de la geometría ( $H$ ) y las propiedades del material ( $v_s$ ) del depósito de suelo

En el caso de propagación de ondas de corte en materiales viscosos se presenta una disipación de energía debido al amortiguamiento del suelo, por lo tanto, no puede existir una amplificación infinita como la que presenta el análisis anterior, proporcionando resultados más acordes con la realidad. Asumiendo que el suelo es un material viscoso tipo Kelvin-Voigt y utilizando el principio de correspondencia (el cual establece que la solución amortiguada se obtiene de la solución elástica reemplazando las constantes elásticas por las constantes complejas correspondientes), los desplazamientos resultantes pueden expresarse con la misma ecuación (3.1) pero sustituyendo  $k$  por  $k^*$ , donde  $k^*$  es un número de onda complejo con parte real  $k_1$  y parte imaginaria  $k_2$ . Considerando las manipulaciones algebraicas anteriores, la función de transferencia para el caso de un suelo amortiguado sobre la roca rígida puede expresarse como:

$$F_2(\omega) = \frac{1}{\cos(k^*H)} \quad (3.5)$$

Si el módulo rigidez y la velocidad de onda complejos se define respectivamente como:

$$G^* = G(1+i2\beta) \quad v_s^* = \sqrt{G^*/\rho} = v_s(1+i\beta)$$

entonces el número de onda complejo, para pequeños valores de  $\beta$ , puede escribirse de la forma

$$k^* = \frac{\omega}{v_s^*} = k(1-i\beta)$$

Finalmente, la función de amplificación se expresa:

$$|F_2(\omega)| = \frac{1}{\sqrt{\cos^2(kH) + (\beta kH)^2}} = \frac{1}{\sqrt{\cos^2(\omega H / v_s) + [\beta(\omega H / v_s)]^2}} \quad (3.6)$$

De acuerdo con la expresión (3.6), la amplificación en una capa de suelo amortiguado también varía con la frecuencia, para relaciones de amortiguamiento pequeñas. La amplificación alcanzará un máximo local siempre que  $k_n H = \pi/2 + n\pi$  pero nunca alcanzará un valor infinito (para  $\beta > 0$ ). Como se ha mencionado, las frecuencias que corresponden a los máximos locales se conocen como frecuencias naturales del depósito de suelo. De acuerdo con la expresión (3.4), el factor de amplificación pico disminuye con el incremento de la frecuencia natural, el más grande factor de amplificación ocurrirá aproximadamente para la frecuencia natural más baja, también conocida como la *frecuencia fundamental*. El periodo de vibración que corresponde a la frecuencia fundamental se conoce como el *periodo característico o fundamental del sitio*:

$$T_s = \frac{4H}{v_s} \tag{3.7}$$

La fig. 3.4 muestra la variación del factor de amplificación con la frecuencia para diferentes niveles de amortiguamiento. Si estas curvas se comparan con el caso no amortiguado de acuerdo con la expresión 3.3 (donde las amplitudes tienden a infinito en las frecuencias naturales del depósito), se aprecia que el amortiguamiento afecta la respuesta más para frecuencias altas que para frecuencias bajas.

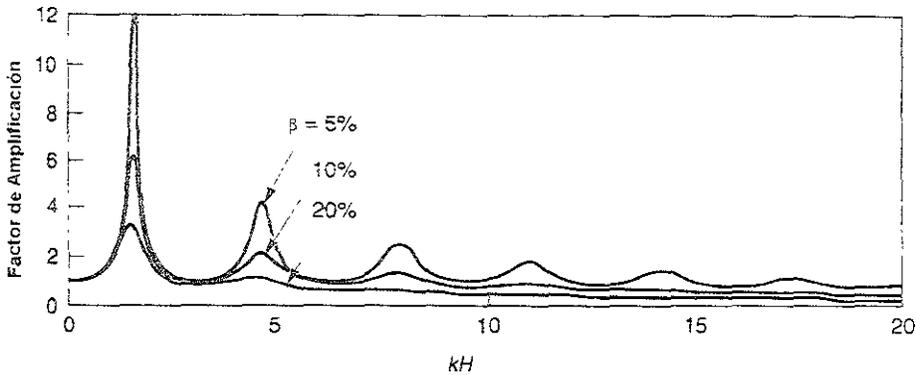


Fig 3.4 Influencia del amortiguamiento en la respuesta sísmica de estratos de suelo homogéneo

Como se puede apreciar, los análisis teóricos pueden mostrar cómo se afectan los movimientos superficiales del terreno por las condiciones locales del suelo. Por lo general, los movimientos en terreno firme son amplificados por los depósitos de suelo, porque en la mayoría de los sitios, la densidad ( $\rho$ ) y la velocidad de propagación de las ondas S ( $v_s$ ) de los materiales cerca de la superficie son menores que a profundidades mayores. De esta forma, si se desprecian los efectos de variabilidad y amortiguamiento del material, la conservación de la energía de las ondas elásticas requiere que el flujo de energía ( $\rho v_s u^2$ ) sea constante desde la profundidad a la superficie del terreno. Puesto que  $\rho$  y  $v_s$  disminuyen conforme se aproximan las ondas a la superficie de terreno, entonces la velocidad de la partícula,  $u$ , debe aumentar.

Las características de los depósitos locales del suelo también influyen en la amplificación del movimiento del terreno cuando la impedancia específica ( $\rho v_s$ ) es constante. Esto se ilustra considerando por ejemplo, dos depósitos de suelo con geometrías idénticas (ver fig. 3.5a), pero uno de ellos más rígido ( $v_s=320$  m/s) que el otro ( $v_s=80$  m/s), se asume además, que los dos depósitos de suelo son linealmente elásticos y la roca basal es rígida. Las funciones de amplificación para estos casos se determinan con la expresión (3.6), donde se muestra claramente (ver fig. 3.5b) que en el sitio más blando se amplifican los movimientos de roca en las frecuencias bajas (periodos largos) más que en el sitio menos blando, lo contrario se observaría para los movimientos de roca en las frecuencias altas (periodos cortos).

### 3.2.1.2 Variaciones de los análisis de respuesta sísmica

Los casos simples de análisis teóricos presentados anteriormente ilustran aspectos importantes del efecto que tienen los depósitos de suelo en la respuesta sísmica, que desde el punto de vista práctico son bastante aceptables; sin embargo, existen análisis más complejos que se acercan más a la realidad

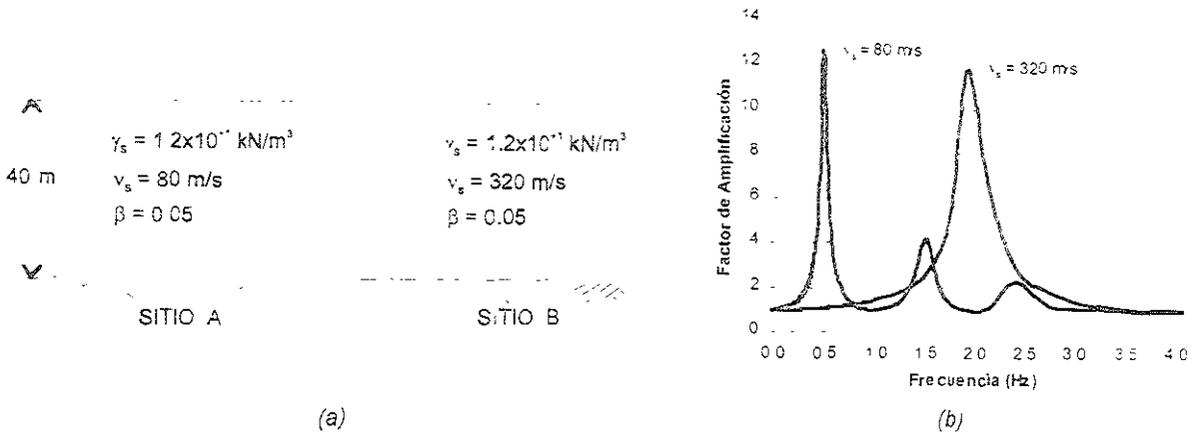


Fig 3.5 (a) Dos depósitos de suelo hipotéticos sobre lecho de roca rígido y (b) Funciones de amplificación para los sitios A y B

### Depósitos de suelo homogéneo sobre roca basal elástica

Es posible determinar los movimientos sísmicos en depósitos de suelo sobreyaciendo un lecho de roca elástica, en donde a diferencia del caso anterior, roca basal rígida, los movimientos que se presentan son afectados por la presencia de la roca basal, y sólo se reflejarán parcialmente las ondas de esfuerzo que viajan descendientemente al llegar a la frontera suelo-roca y parte de su energía se transmitirá a través de la frontera para continuar viajando descendientemente. Si la roca se extiende a gran profundidad, parte de la energía elástica de estas ondas se disipará del depósito de suelo (amortiguamiento por radiación) ocasionando que las amplitudes del movimiento en la superficie libre sean menores que aquéllas para el caso con lecho de roca rígido. En la fig. 3.6 se presenta el efecto de la rigidez de la roca basal en la magnitud de la amplificación (caso no amortiguado) y para ciertas relaciones de impedancia,  $\rho_1 v_1 / \rho_2 v_2$  (donde el estrato 1 sobreyace al 2). Se observa la semejanza entre los efectos por amortiguamiento y por la elasticidad del lecho de roca.

Debido a que un sismo produce movimientos en el lecho de roca sobre un rango de frecuencias, se amplificarán más algunas componentes del movimiento de lecho de roca real que otras. Considerando una roca basal elástica, la amplificación del movimiento del sitio local será influenciada por la impedancia específica del lecho de roca. Es decir, se esperaría que en una roca basal más rígida se produzcan factores de amplificación aproximadamente del 50% más altos que aquéllos asociados con roca basal más suave, para condiciones de suelo equivalentes (Jacob, 1991). Por consiguiente, cualquier descripción de las condiciones locales del sitio debería incluir la densidad y rigidez del lecho de roca.

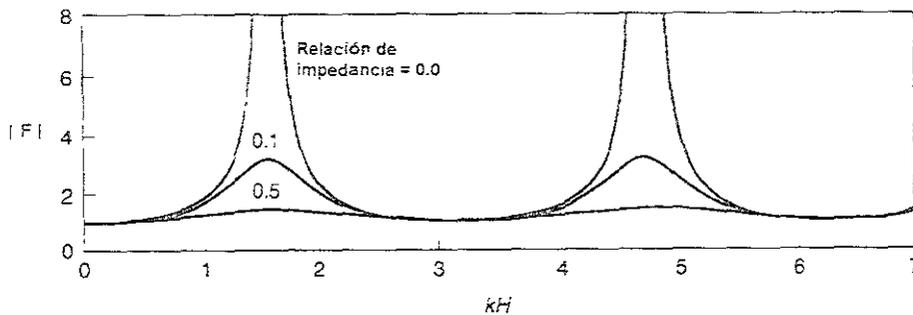


Fig. 3.6 Efecto de la relación de impedancia en el factor de amplificación para el caso de suelo no amortiguado

### Depósitos de suelo estratificados

Por otra parte, los problemas de respuesta en el terreno normalmente involucran depósitos de suelo estratificados con características de rigidez y amortiguamiento diferentes, con fronteras en las cuales la energía de onda elástica se reflejará y/o se transmitirá. Para tales condiciones se han desarrollado métodos de análisis que emplean funciones de transferencia donde se considera un depósito de suelo constituido de  $N$  capas horizontales donde la  $N$ -ésima capa es el lecho de roca (fig 3.7). Tales funciones de transferencia indican que el movimiento en cualquier capa puede determinarse del movimiento en cualquier otra capa, de lo cual, si el movimiento en cualquier punto en el perfil de suelo es conocido, el movimiento en cualquier otro punto puede definirse (Schnabel, et al., 1972)

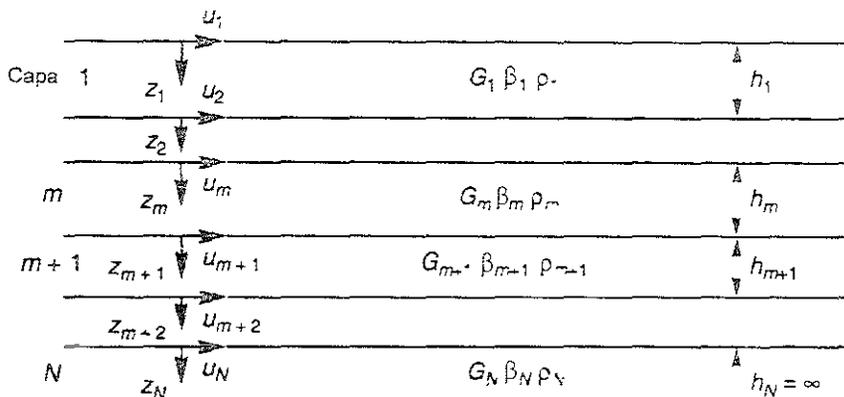


Fig 3.7 Modelo de respuesta sísmica para depósitos de suelo estratificados en roca basal elástica

### Aproximación lineal equivalente de respuesta no lineal del suelo

El efecto del comportamiento esfuerzo-deformación histerético no lineal del suelo bajo carga cíclica en la respuesta sísmica de depósitos de suelos blandos estratificados, se puede aproximar utilizando el método lineal equivalente (Seed e Idriss, 1969), que consiste básicamente en actualizar el módulo de rigidez al cortante,  $G$ , y la relación de amortiguamiento,  $\beta$ , del suelo, con base en el nivel de deformaciones (o esfuerzos) de cortante.

Puesto que el nivel de deformación calculada depende de los valores de las propiedades lineales equivalentes y de la intensidad del sismo, se requiere un procedimiento iterativo que asegure que las propiedades usadas en el análisis sean compatibles con el nivel de deformación calculado en todas las capas que constituyen el depósito. El proceso de iteración para obtener propiedades de suelo compatibles con la deformación permite aproximar un comportamiento de suelo no lineal, es importante recordar que esta aproximación sigue siendo un método de análisis lineal. Las propiedades de suelo compatibles con la deformación son constantes en toda la duración del sismo, a pesar de que las deformaciones para un tiempo particular sean pequeñas o mas grandes. Por ello, el método es incapaz de representar los cambios en la rigidez del suelo que realmente ocurren durante el sismo; sin embargo, esta aproximación ha sido aceptable en muchos casos (Romc y Seed, 1986, Dickenson, 1994, Chang, 1996).

La aproximación lineal equivalente para análisis unidimensional de respuesta del terreno de sitios estratificados ha sido programado en un código muy conocido llamado SHAKE (Schnabel et al., 1972). Una versión de este programa que utiliza la teoría de vibraciones aleatorias y la del valor extremo, fue desarrollada por Bárcena y Romo (1993). La diferencia básica entre estos dos códigos es que mientras en SHAKE la excitación se define en términos de acelerogramas, en el programa RADSH la sollicitación sísmica se da en términos de un espectro de respuesta (o un espectro de potencia), el cual puede ser de aceleraciones, velocidades o desplazamientos.

Por otra parte, se han desarrollado varios modelos no lineales para describir la respuesta de depósitos de suelo ante cargas sísmicas, y que se construyen a partir de una ecuación constitutiva para deformación cortante unidimensional en la base de un modelo histerético suponiendo que los coeficientes elásticos y de amortiguamiento son función de los valores medios de deformación (p.e. Muravskii y Frydman, 1998). Los resultados obtenidos con este tipo de modelos se han comparado con aproximaciones lineales equivalentes (Martin, 1975) mostrando algunas de las limitaciones que este último método tiene.

### **Propagación de diferentes tipos de ondas**

El análisis de registros medidos en campo ha permitido comprobar la existencia de diferentes tipos de ondas sísmicas que se propagan en el medio, como son las ondas de cuerpo (de compresión y de corte) y superficiales. En las ondas de compresión (P) no se presentan deformaciones en las direcciones paralelas al frente de onda, e involucra cambios de volumen sin distorsión cortante. En las ondas de corte (S) el movimiento es normal a la dirección de propagación de la onda y existen esfuerzos cortantes en los planos normales, es decir, no se tienen cambios de volumen y sólo hay fuerzas de corte. Las ondas de superficie resultan de la interacción de las ondas de cuerpo con la superficie y estratos superficiales del terreno, donde las más conocidas son las ondas Rayleigh y Love. También, a partir de mediciones realizadas en campo, se han detectado otro tipo de ondas de corteza de periodo-corto (Lg y Rg) con características similares a las superficiales y que viajan solamente a lo largo de trayectorias continentales.

Las contribuciones relativas de los diferentes tipos de ondas al movimiento total del suelo y los tiempos de llegada dependen de varios factores tales como: la distancia epicentral, la profundidad focal, reflexiones y refracciones múltiples, y la dispersión a lo largo de las trayectorias de propagación. En cierta forma, todos los registros sísmicos contienen algunas componentes de todas las ondas mencionadas. Sin embargo, la proporción exacta de cada una de ellas resulta muy difícil de conocer, debido a que las propiedades de la fuente del sismo, y los detalles físicos y geométricos de la geología no se pueden definir exactamente.

Sin embargo, se han desarrollado varias teorías para el análisis de la respuesta de sitios estratificados horizontalmente sujetos a ondas de cuerpo armónicas planas con ángulos de incidencia definidos (Hannon, 1964; Silva, 1978). De la misma forma, existen varias investigaciones (Lysmer y Drake, 1972; Bocheva, 1977; Swanger, 1978) sobre depósitos estratificados sujetos a ondas de superficie.

### **Métodos numéricos**

La necesidad de simular la propagación de ondas sísmicas en circunstancias más reales, ha conducido al uso de métodos numéricos como diferencias finitas, elementos finitos o elementos espectrales. Existe una gran variedad de lecturas relacionadas con los temas (p.e. Ohmiato y Chouet, 1997; Moczo et al., 1997; Faccioli et al., 1997); sin embargo, es interesante mencionar brevemente algunos casos que se han presentado para evaluar el efecto de los depósitos de suelos en la respuesta sísmica.

El método de elementos finitos se emplea para estudiar por ejemplo, el problema de variaciones locales en los movimientos de campo libre (Chen, Lysmer y Seed, 1981), donde el medio ambiente sísmico se considera compuesto por una proporción conocida de ondas de cuerpo con un cierto ángulo de incidencia y ondas de Rayleigh y de Love. Una de las conclusiones importantes del análisis anterior es que para depósitos de suelo en los que la rigidez se incrementa con la profundidad, las respuestas horizontales determinadas con un modelo unidimensional (ondas S propagándose verticalmente) son similares a las calculadas horizontalmente suponiendo ondas Rayleigh. Además, se observó que el componente vertical calculado para ondas Rayleigh se reproducía bastante bien si la respuesta del depósito se calculaba suponiendo un modelo 1-D y ondas P propagándose verticalmente. Entonces, el análisis desacoplado (SH y P) podría usarse en lugar del acoplado (R). Estos resultados explican, al menos en parte, porqué la hipótesis de propagación vertical de ondas se cumple satisfactoriamente para sitios estratificados.

Por otra parte, el método de elementos de frontera indirecto puede utilizarse para simular la propagación de ondas bidimensionales en un medio elástico estratificado irregular (Vai, et al., 1999). Los resultados obtenidos mediante este procedimiento han permitido evaluar los efectos de la geología de superficie en la respuesta sísmica del terreno. Análisis similares también permiten estudiar el efecto de irregularidades aleatorias en la superficie de la roca de apoyo, considerando un medio estratificado arbitrario (Fuks y Voronovich, 2000; Jin y Li, 2000).

### **3.2.2 Funciones de Amplificación Medidas**

Las suposiciones idealizadas de los análisis unidimensionales de respuesta del terreno (materiales uniformes, capas horizontales, ondas de corte propagándose verticalmente, etc.) producen funciones de amplificación suaves (ver fig. 3.5b), lo cual no ocurre con las funciones de amplificación reales, debido a que tales condiciones raramente existen en el campo.

La función de amplificación real se determina a partir de los datos de registro obtenidos en sitios donde se instalaron instrumentos de medición de movimientos fuertes en arreglos verticales o pozos, con lo cual es posible evaluar la variación vertical de los movimientos sísmicos desde la superficie de los depósitos de suelo hasta la roca basal o terreno firme (Johnson y Silva, 1981; Chang et al., 1986; Romo, 1995; Taboada, et al., 1998). Por ejemplo, en un lugar con depósitos de suelo estratificados pero con velocidades de ondas de corte relativamente uniformes, las funciones de amplificación reales son cualitativamente similares a lo predicho por los análisis de respuesta del terreno. Para los sitios con condiciones del subsuelo más complicadas, o para sismos más fuertes donde el comportamiento no lineal del suelo es más significativo, disminuye la efectividad de los análisis de respuesta teóricos para predecir la forma general de la función de amplificación real, principalmente en los picos irregulares. El comportamiento no lineal del suelo también ocasiona que las funciones de amplificación de movimientos fuertes sean diferentes a aquéllas de movimientos débiles (Aki, 1993).

### **3.2.3 Movimientos de Superficie Medidos**

La comparación de los movimientos superficiales del terreno medidos en diferentes sitios es de gran utilidad para mostrar el efecto que tienen las condiciones locales del suelo en la respuesta sísmica del terreno. Ejemplos de esta evidencia significativa se mencionan a continuación dando un mayor énfasis al caso de la ciudad de México en 1985.

### **San Francisco, 1957**

Los movimientos registrados en varias zonas de San Francisco durante el sismo de 1957 de magnitud cercana a 5.3, indican que los movimientos superficiales del terreno en los afloramientos de roca son bastante similares. pero los movimientos en términos de amplitud y contenido de frecuencias para los sitios donde se presentan depósitos de suelo de espesor considerable son diferentes (Idriss y Seed 1968).

### **Caracas, 1967**

Los movimientos sísmicos registrados en Caracas, Venezuela en el año de 1967 también indican una amplificación en el centro de la ciudad debido a la presencia de depósitos aluviales (Duval, et al., 1998 Semblat, et al., 1999). La frecuencia de resonancia de estos depósitos de suelo es alrededor de 0.6 Hz ocasionando los mayores daños a varias estructuras con frecuencia de resonancia muy cercanas a las del depósito

### **Loma Prieta, California, 1989**

El 19 de octubre de 1989, ocurrió un sismo de magnitud de 7.1 con epicentro en Loma Prieta (a 100 km al sur de San Francisco y Oakland, California). Las intensidades más altas de este evento se presentaron en el área de la Bahía de San Francisco, donde se tienen depósitos aluviales de arcillas y limos, con algunas capas de arena y grava.

Las aceleraciones horizontales máximas se atenuaron más rápidamente para sitios de roca-suelo residual superficial que en zonas con depósitos aluviales. Las aceleraciones máximas en los afloramientos de roca fueron de 0.06 g en la direcciones de E-W y 0.03 g en la direcciones de N-S, los valores correspondientes en la superficie de los depósitos de la arcilla de la Bahía fueron de 0.16 g y 0.11 g. Estos resultados muestran como la presencia de los depósitos de suelos blandos causaron una amplificación significativa del movimiento del lecho de roca subyacente (Seed et al., 1990)

### **Dinar, 1995**

En octubre de 1995 ocurrió un sismo ( $M_w = 6$ ) provocando daños considerables en la ciudad de Dinar Turquía. La distribución del daño sísmico y las variaciones en las características de frecuencia y amplitud de las aceleraciones registradas durante el evento, permiten identificar tres tipos de zonas: montañosas, transición y planas. Los datos registrados indican que la relación de respuesta espectral entre sitios de suelo blando y suelo rígido es del orden de 2-3 dentro del rango de frecuencias de 0.5-3 Hz. A pesar de que tales amplificaciones no se atribuyen por completo a los efectos de sitio (debido a que las distancias relativas entre estaciones en comparación con la distancia epicentral del sismo, no son suficientemente grandes) este ejemplo nos muestra en general que los depósitos de suelos aluviales tienden a amplificar los movimientos sísmicos (Durukal et al., 1998)

### **Pujilí, 1996**

El sismo del 28 de marzo de 1996 ( $M_w=5.7$ ) produjo un extenso daño en Pujilí, una pequeña ciudad localizada en la parte central del valle Inter-Andino de Ecuador. La distribución de daños en el lugar permiten suponer una amplificación en los movimientos del terreno relacionados a efectos locales del sitio, debido a que el área de mayor daño se localizó sobre una zona de depósitos aluviales (Guéguen et al., 1998).

**Ciudad de México, 1985**

Los movimientos superficiales del terreno registrados para diferentes sitios (Mena et al. 1986) en la ciudad de México durante el sismo del 19 de septiembre de 1985 ( $M_s=8.1$ ) con epicentro en las costas del Pacífico (Guerrero-Michoacán), ilustraron la importancia de las condiciones locales del suelo en los movimientos del terreno superficial (Romo y Jaime, 1986; Romo y Seed, 1986). En ese entonces, existían instrumentos en Ciudad Universitaria (CU), Secretaría de Comunicaciones y Transportes (SCT), Central de Abastos Oficinas (CAO) y en el Frigorífico (CAF), Tiáhuac Deportivo (TD), Tiáhuac Bombas (TB), Viveros Coyoacán (VC) y Tacubaya (TYC). La ubicación de estos instrumentos de medición para movimientos fuertes en la ciudad de México se muestra en la fig. 3.8, donde los sitios CU y TYC se encuentran en la zona de Lomas (terreno firme), los sitios SCT, CAO, CAF, TD y TB se localizan en la zona Lago (depósitos de arcilla de alta compresibilidad de espesores mayores a 20 m) y el sitio VC en la zona de Transición (depósitos de arena aluvial y capas limosas intercaladas con capas de arcilla).

A pesar de la gran distancia epicentral (400 km), se produjeron aceleraciones en el sitio CU de sólo 0.04 g, en la orientación E-W. En la zona de Transición (sitio VC), las aceleraciones máximas del terreno fueron ligeramente mayores que aquéllas que se presentaron en el sitio CU. Sin embargo, en los sitios CAO y SCT (zona de Lago), las aceleraciones máximas del terreno fueron del orden de dos y cinco veces más grandes que las registradas en CU, respectivamente (ver Tabla 3.2). Los niveles fuertes de sacudimiento persistieron sobre una duración muy larga en estos sitios ubicados en la zona de Lago.

Los contenidos de frecuencias de los movimientos en los sitios SCT, CAO y CAF fueron muy diferentes a los presentados en el sitio CU. Los espectros de respuesta mostrados en la fig. 3.9 ilustran los efectos pronunciados de los depósitos de suelos blandos, donde para periodos aproximados de 2 s (valor del periodo fundamental,  $T_s$ , del sitio SCT), las aceleraciones espectrales fueron del orden de 10 veces más altas que aquéllas registradas en terreno firme.

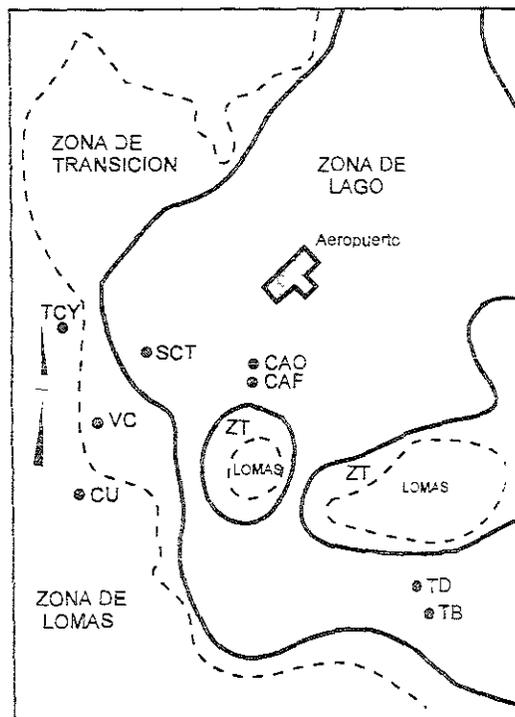


Fig. 3.8 Zonificación Geotécnica y localización de algunas estaciones acelerométricas en la ciudad de México

El daño estructural fue despreciable en la zona de Lomas y mínimo en la zona de Transición. En la zona de Lago ocurrieron los mayores daños, específicamente en áreas donde los espesores de arcilla variaban de 25 a 46 m (Romo, 1987) con periodos característicos de los depósitos de suelo blando estimados entre 1.9 a 2.8 s. Incluso dentro de estas zonas, el daño fue mayor para aquellas estructuras donde su periodo dominante resultó ser aproximado al periodo del sitio característico. Esta condición de *doble resonancia* (amplificación de los movimientos en lecho de roca por el depósito de suelo y la amplificación de los movimiento de campo libre por la estructura) combinada con las deficiencias estructurales de diseño y construcción, causaron un daño local severo

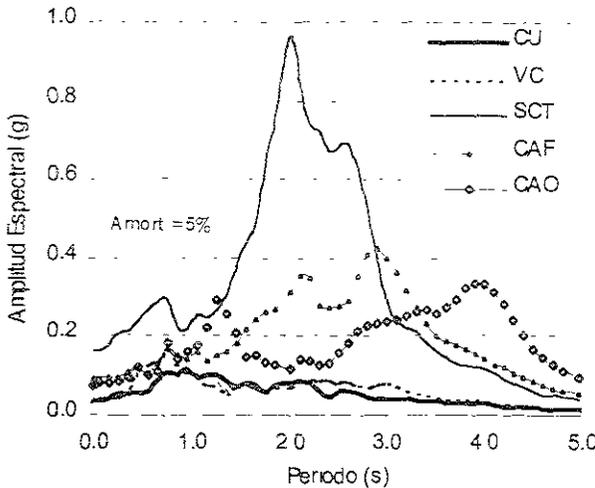


Tabla 3.2 Características del sismo 89/09/19 orientación E-W, en los sitios de estudio

Síto	$A_{max}$ (g)	$T_s$ (s)
CU	0.035	-
TCY	0.032	-
VC	0.041	0.8
SCT	0.163	2.0
CAF	0.092	2.9
CAO	0.078	3.5
TB	0.103	3.6
TD	0.114	3.8

Fig. 3.9 Espectros de respuesta de aceleraciones de los movimientos medidos durante el sismo 89/09/19, orientación E-W

### 3.2.4 Resumen del Efecto de las Condiciones Locales del Suelo en los Movimientos Sísmicos del Terreno

Las secciones anteriores indican claramente que las condiciones locales del suelo influyen grandemente en la magnitud de las aceleraciones máximas. Las relaciones entre las aceleraciones máximas registrada en afloramientos de roca y en sitios con diferentes perfiles de suelo (roca, suelos rígidos y suelos cohesivos profundos) muestran distintas tendencias en el comportamiento de amplificación (Seed et al., 1976). Como se aprecia en la fig. 3.10 las aceleraciones máximas en la superficie de los depósitos de suelo son ligeramente más grandes cuando las aceleraciones máximas en la roca son pequeñas (menos de 0.1 g) y algo menores para niveles de aceleración altos (más de 0.1 g)

Basándose en datos registrados (ciudad de México y Bahía de San Francisco) y análisis adicionales de respuesta sísmica, se han obtenido relaciones entre las aceleraciones máximas del terreno en los sitios de depósitos de suelo blando y en los sitios de roca (driss, 1990) que se muestran en la fig. 3.11. Se puede observar que para niveles bajos a moderados de aceleración (menos de 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 las superficies de los depósitos de suelo blando. Sin embargo, para niveles de aceleración más altos (mayor de 0.4 g) no se desarrollan aceleraciones máximas tan grandes como aquéllas registradas en roca, debido básicamente a la baja rigidez y el alto amortiguamiento generados por el marcado comportamiento no lineal de los suelos blandos.

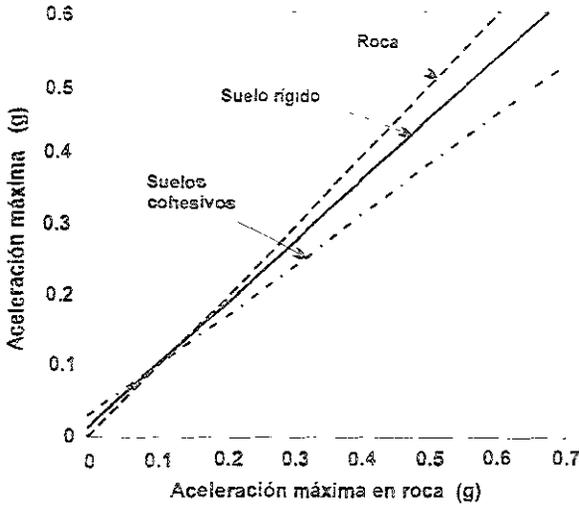


Fig 3.10 Relaciones aproximadas entre aceleraciones máximas en roca y condiciones de sitio local (tomado de Seed et al., 1976)

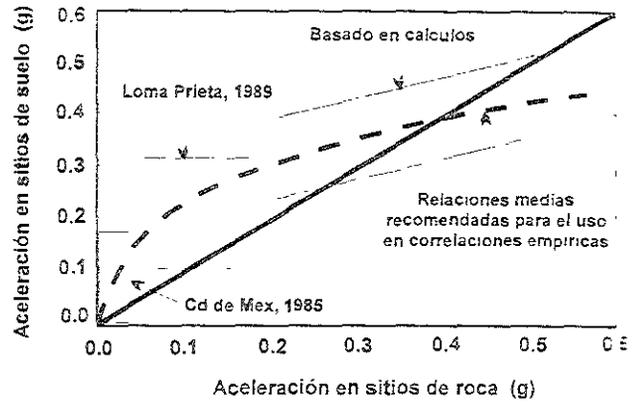


Fig 3.11 Relaciones aproximadas entre aceleraciones máximas en roca y sitios de suelo blando (tomado de Idriss, 1990)

Las condiciones locales del suelo también influyen en el contenido de frecuencias de los movimientos sísmicos superficiales y por lo tanto, en la amplitud y forma de sus espectros de respuesta. Se han calculado espectros de respuesta de los movimientos del terreno registrados para diferentes sitios: roca, suelos rígidos (menos de 61 m de profundidad), suelos cohesivos profundos (mayor de 76 m de profundidad) y depósitos de arcilla blanda a media rígida (Seed et al., 1976). Si se normalizan los espectros calculados con respecto a la aceleración máxima del terreno, es posible observar los efectos de las condiciones locales del suelo en las formas espectrales como se muestra en la fig. 3.12. Se observa que para periodos alrededor de 0.5 s, las amplificaciones espectrales son mucho más altas para los sitios de suelo que para los sitios de roca. Para periodos más largos, la amplificación espectral aumenta con la disminución de la rigidez en los depósitos de suelo. Se muestra claramente (ver fig. 3.12) que los depósitos de suelos blandos y profundos amplifican los movimientos en los periodos largos (bajas frecuencias). Estos resultados indican también, que no es apropiado el uso de una sola forma de espectro de respuesta para todas las condiciones de sitio que se presentan.

Un análisis similar al caso anterior se realizó ahora para los depósitos de suelo lacustre de la ciudad de México. En la fig. 3.13 se presentan las formas espectrales (FE) normalizadas para diferentes condiciones del subsuelo y considerando valores espectrales medios ( $\mu$ ), +1, +1.5 y +2 desviaciones estándar ( $\sigma$ ), donde este último caso representa prácticamente la envoltura de los valores espectrales para cada caso considerado. Estos resultados se obtuvieron a partir de datos medidos en campo para varios eventos sísmicos disponibles en cada región (zonas II y III) de acuerdo a la zonificación geotécnica de la ciudad de México, originalmente propuesta por Marsal y Mazari (1957), y recientemente afinada por Auvinet, Méndez y Juárez (2000), como se indica en la fig. 4.2. Se conservan las tres zonas I (de Lomas), II (de Transición) y III (de Lago).

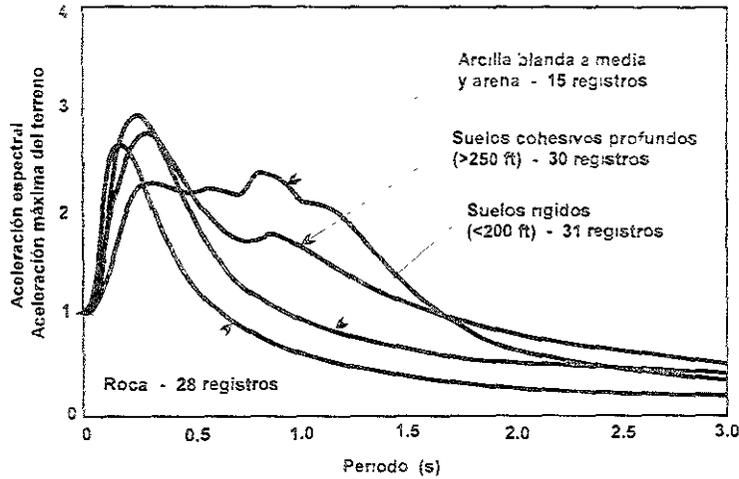


Fig. 3 12 Espectros de respuesta normalizados medios (5% de amortiguamiento) para diferentes condiciones de sitio local (tomado de Seed et al., 1976)

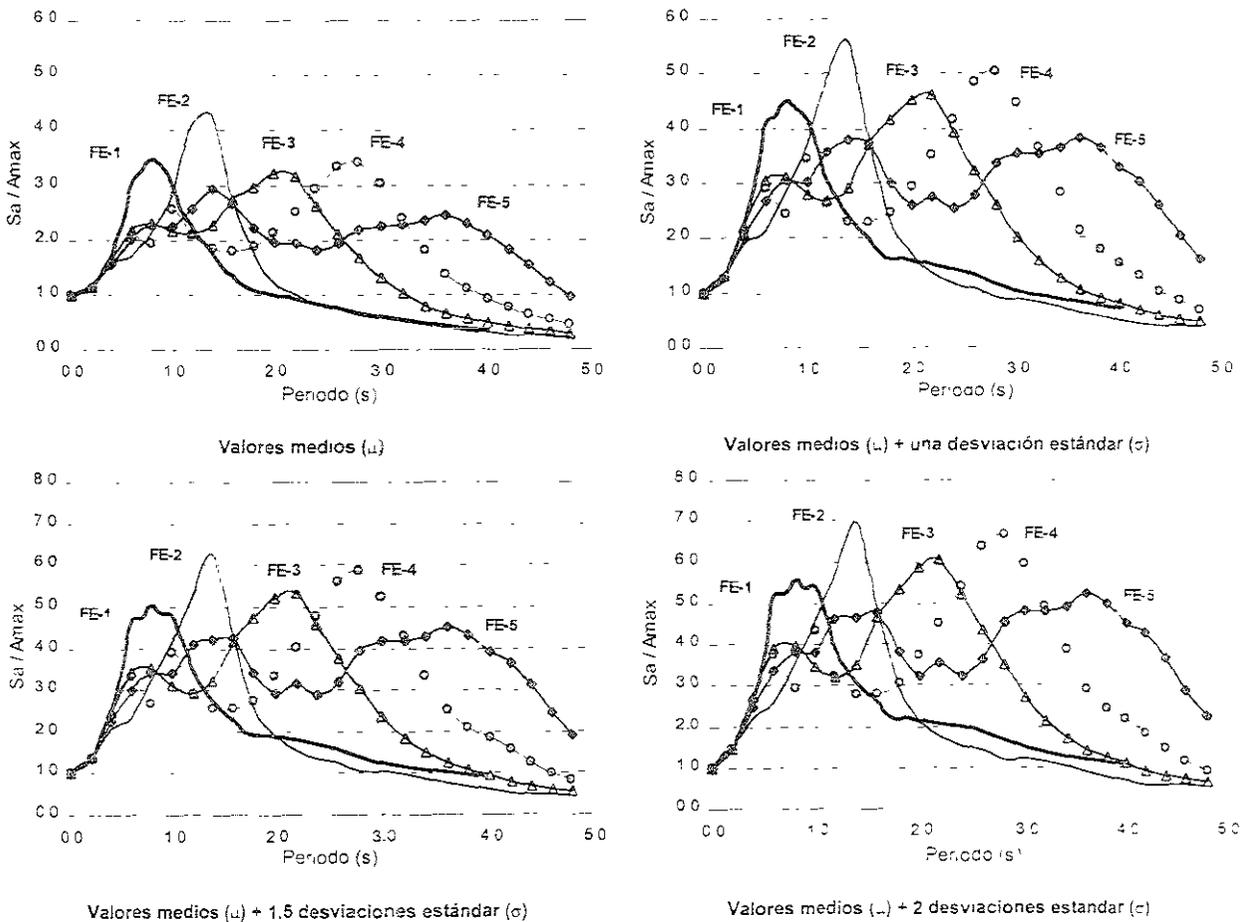


Fig. 3 13 Espectros de respuesta normalizados (5% de amortiguamiento) para diferentes condiciones locales del suelo de acuerdo a la zonificación geotécnica del valle de México (fig. 4 2)

La base de datos con que se realizó el estudio estadístico anterior se presenta en la Tabla 3.3 donde se muestran los sitios (estaciones acelerométricas) para cada zona geotécnica y los eventos sísmicos disponibles para cada uno de ellos. La forma espectral FE-1 corresponde a la zona de Transición, que como se ha mencionado, se tienen depósitos de arena aluvial y capas limosas intercaladas con capas de arcilla, mientras que para la zona de Lago (depósitos de arcilla blanda) se determinaron cuatro formas espectrales (FE-2 a la FE-5) debido a que el rango de periodos donde se presentan las amplitudes máximas es bastante amplio. Los sitios y el rango de periodos fundamentales (Ts) considerados para cada forma espectral de la zona de Lago, se presentan en la misma Tabla 3.3. Estos resultados indican la importancia de evaluar los movimientos sísmicos en los depósitos arcillosos, ya que se pueden tener variaciones importantes en las características de los movimientos sísmicos que se pueden traducir en graves daños a las estructuras, como ocurrió en 1985.

Tabla 3.3 Base de datos del análisis estadístico de respuesta sísmica del valle de México

Forma Espec	Zona Geo	Sitios	Sismos (año/mes/día)												
FE-1	Transición	CO47	93/05/15	94/12/10	93/10/24	95/10/09	95/09/14								
		DR16	93/07/29	94/05/23	94/12/10	95/10/21	95/10/30	93/10/24	95/10/09						
		COYS	93/05/15	93/07/29	94/05/23	94/12/10	95/10/21	95/10/30	93/10/24	95/10/09	85/09/19				
		EO30	94/05/23	94/12/10	93/10/24	95/10/09	95/09/14								
		GR27	93/05/15	94/05/23	94/07/04	94/12/10	95/10/21	95/10/30	93/10/24	94/03/14	95/10/09	95/09/14			
		ES57	93/05/15	94/05/23	94/07/04	94/12/10	95/10/21	95/10/30	93/10/24	94/03/14	95/10/09	95/09/14			
		AU46	94/05/23	94/12/10	95/10/21	95/10/30	94/03/14	95/10/09	95/09/14						
		ME52	93/05/15	94/05/23	94/12/10	93/10/24	95/10/09	95/09/14							
		DX37	93/05/15	94/07/04	94/12/10	95/10/21	95/10/30	93/10/24	94/03/14	95/10/09	95/09/14				
		AO24	93/05/15	94/12/10	95/10/21	95/10/09	95/09/14								
FE-2	Lago Ts(s) (1.2-1.4)	DFRO	90/05/11	93/05/15	94/05/23	94/07/04	94/12/10	95/10/21	95/10/30	89/04/25	93/10/24	95/10/09	95/09/14		
		SI53	93/05/15	94/05/23	94/07/04	94/12/10	95/10/21	95/10/30	93/10/24	95/10/09	95/09/14				
		JC54	94/05/23	94/07/04	94/12/10	93/10/24	94/03/14	95/10/09	95/09/14						
		MI15	93/07/29	94/05/23	94/07/04	94/12/10	93/10/24	94/03/14	95/10/09	95/09/14					
		CH84	94/05/23	94/07/04	94/12/10	93/10/24	94/03/14	95/10/09	95/09/14						
		IB22	94/05/23	94/12/10	93/10/24	95/10/09	95/09/14								
FE-3	Lago Ts(s) (2.0-2.5)	SCT2	90/05/11	94/05/23	94/12/10	97/01/11	99/06/15	99/09/30	89/04/25	93/10/24	95/10/09	95/09/14	85/09/19		
		CJ03	94/12/10	95/10/30	94/03/14	95/10/09	95/09/14								
		TL08	93/07/29	94/05/23	94/07/04	94/12/10	95/10/30	93/10/24	94/03/14	95/10/09	95/09/14				
		CI05	93/07/29	95/10/30	94/05/23	94/07/04	94/12/10	94/03/14	93/10/24	95/09/14	95/10/09				
		TL55	93/07/29	94/05/23	94/07/04	94/12/10	93/10/24	94/03/14	95/09/14						
		GA62	93/07/29	94/05/23	94/07/04	94/12/10	93/10/24	94/03/14	95/10/09	95/09/14					
		RMAS	90/05/11	93/07/29	94/05/23	94/07/04	97/01/11	93/10/24	95/09/14						
		LI58	93/05/15	95/10/30	93/10/24	94/03/14	95/10/09								
		BL45	93/07/29	94/05/23	94/07/04	94/12/10	95/10/30	93/10/24	94/03/14	95/10/09	95/09/14				
		VG09	93/07/29	94/05/23	94/07/04	94/12/10	93/10/24	94/03/14	95/10/09	95/09/14					
RM48	93/07/29	94/05/23	94/07/04	94/12/10	93/10/24	94/03/14	95/10/09	95/09/14							
FE-4	Lago Ts(s) (2.6-3.1)	HJ72	93/07/29	94/05/23	94/07/04	94/12/10	93/10/24	94/03/14	95/10/09	95/09/14					
		XPC6	93/05/15	94/05/23	94/12/10	93/10/24	94/03/14	95/10/09							
		BO39	94/05/23	94/12/10	93/10/24	94/03/14	95/10/09	95/09/14							
		BA49	93/05/15	94/05/23	94/12/10	95/10/30	93/10/24	94/03/14	95/10/09	95/09/14					
		VM29	93/07/29	94/05/23	94/07/04	94/12/10	93/10/24	95/10/09	95/09/14						
		CA59	94/05/23	94/12/10	95/10/30	93/10/24	94/03/14	95/10/09	95/09/14						
		JA43	93/05/15	94/05/23	94/12/10	93/10/24	94/03/14	95/10/09	95/09/14						
AP68	94/05/23	94/12/10	94/03/14	95/10/09	95/09/14										
FE-5	Lago Ts(s) (3.4-4.2)	AR14	93/05/15	94/05/23	94/12/10	94/03/14	95/10/09	95/09/14							
		CDAO	88/02/08	90/05/11	97/01/11	99/09/30	89/04/25	95/09/14	85/09/19						
		UNKS	94/05/23	94/12/10	93/10/24	95/10/09	95/09/14								
		TLAS	93/05/15	94/05/23	94/12/10	95/10/30	93/10/24	95/10/09	95/09/14						
		DM12	93/07/29	94/05/23	94/07/04	94/12/10	93/10/24	94/03/14	95/10/09	95/09/14					
		ZARS	93/07/29	94/07/04	94/12/10	95/10/30	89/04/25	93/10/24	95/10/09	95/09/14					
		HA41	94/05/23	94/07/04	94/12/10	93/10/24	94/03/14	95/10/09	95/09/14						
AU11	93/05/15	94/05/23	94/07/04	94/12/10	93/10/24	94/03/14	95/10/09	95/09/14							

Los valores medios espectrales normalizados para los depósitos de arcilla blanda en la ciudad de México (fig. 3.13), en comparación con los resultados presentados por Seed et al. (1976) en la fig. 3.12 para el mismo tipo de suelo, muestran ordenadas espectrales mayores y rangos de periodos fundamentales más largos. Se observa también, que las formas espectrales para los depósitos de la ciudad de México definen mejor el rango de periodos característicos, lo cual es muy importante en una región (zona de Lago) donde las características de los movimientos sísmicos (amplitud y frecuencias) varían considerablemente.

En las demás gráficas de la misma fig. 3.13, donde se presentan los espectros normalizados medios  $+1$ ,  $+1.5$  y  $+2$  desviaciones estándar, se muestra que tanto varían las ordenadas espectrales de los movimientos sísmicos superficiales debido a los depósitos arcillosos de la ciudad de México, siendo entre 6 y 7 veces la aceleración máxima del terreno para las zonas de mayor amplificación sísmica comprendida entre periodos de 1.4 y 2.6 s ( $\mu+2\sigma$ ). Con base en el análisis de estas formas espectrales se corrobora que una zonificación más detallada para la zona de Lago permitiría utilizar espectros de respuesta más apropiados para el diseño sísmico de estructuras cimentadas en dicha zona. Además, utilizando conjuntamente la información geotécnica disponible con la sísmica, es posible definir una zonificación geosísmica para la zona Metropolitana. Esta clasificación tendría ventajas evidentes sobre el uso de una zonificación geotécnica y espectros específicos, como se recomienda actualmente en el Reglamento de construcciones del distrito Federal.

Observando los espectros normalizados para periodos característicos mayores a 2.0 s (**FE-3**, **FE-4** y **FE-5** para  $\mu+2\sigma$ ), se aprecia que las ordenadas espectrales en los periodos del segundo modo de vibrar son considerables (entre 4 a 5 veces la aceleración máxima del terreno para periodos de 0.8 a 1.6 s) Lo anterior, por ejemplo, podría afectar a edificios o estructuras esbeltas que al diseñarse considerando el primer periodo característico (p.e. entre 3 y 4 s para el caso de la forma espectral **FE-5** donde se aprecia más notablemente este efecto), el periodo dominante de la estructura podría coincidir con el segundo periodo característico del suelo, ocasionando una doble resonancia y aumentando los movimientos sísmicos en tal estructura.

### 3.3 EFECTO DE LA CONSOLIDACIÓN REGIONAL EN LA RESPUESTA SÍSMICA DE LOS DEPÓSITOS ARCILLOSOS DE LA CIUDAD DE MÉXICO

Es importante considerar el efecto del fenómeno de consolidación regional debido a la extracción de agua del subsuelo en la respuesta sísmica de los depósitos arcillosos, como es el caso en la ciudad de México. Para mostrar este efecto, se presenta un caso en que se consideran características y propiedades típicas de la estratigrafía que se encuentra en la zona de Lago.

En la modelación del problema se supone un estrato de arcilla con espesor promedio de 40 m y debajo del cual se encuentran los depósitos profundos (estrato de suelo firme). Para este análisis se consideran como propiedades iniciales representativas de las arcillas de la ciudad de México las mostradas en la Tabla 3.4, así como también se tiene una distribución de presión de poro,  $u$ , hidrostática como condición inicial. Aquí se hace la hipótesis de que existe una fuente de suministro de agua que mantiene el nivel de aguas freáticas al nivel de la superficie del terreno. Se aplica la teoría de consolidación de Terzaghi para modelar el proceso de disipación de presión de poro, considerando que la extracción de agua se realiza en los depósitos profundos. Para estas condiciones, en la fig. 3.14 se presenta la distribución de disipación de  $u$  en función de la profundidad y del factor tiempo,  $T_v$ .

Tabla 3.4 Valores de propiedades representativas de las arcillas de la ciudad de México

$\gamma_{so}$ t/m <sup>3</sup>	$\omega_n$ %	$\omega_L$ %	$\omega_P$ %	$a_v$ cm <sup>2</sup> /kg	$e_o$	$S_s$	$G_{so}$ t/m <sup>2</sup>
1.20	328.60	345.60	108.10	1.37	7.95	2.79	782.90

Con el incremento del esfuerzo efectivo ( $\Delta\sigma'$ ) para cada factor tiempo (fig. 3.14) se determina el incremento del módulo de rigidez al cortante ( $\Delta G$ ) con la siguiente expresión (Romo y Ovando, 1994).

$$\Delta G = 122 p_a \left( \frac{1}{I_p - I_r} \right)^{I_p - I_r} \left( \frac{\Delta\sigma'_c}{p_a} \right)^{0.82}$$

donde:

$I_p$  = Índice plástico.

$I_r$  = Consistencia relativa.

$\Delta\sigma'_c$  = Incremento de esfuerzo de consolidación efectivo.

$p_a$  = Constante para definir las unidades de  $\Delta G$  (p. e. 1kg/cm<sup>2</sup>; 10t/m<sup>2</sup>)

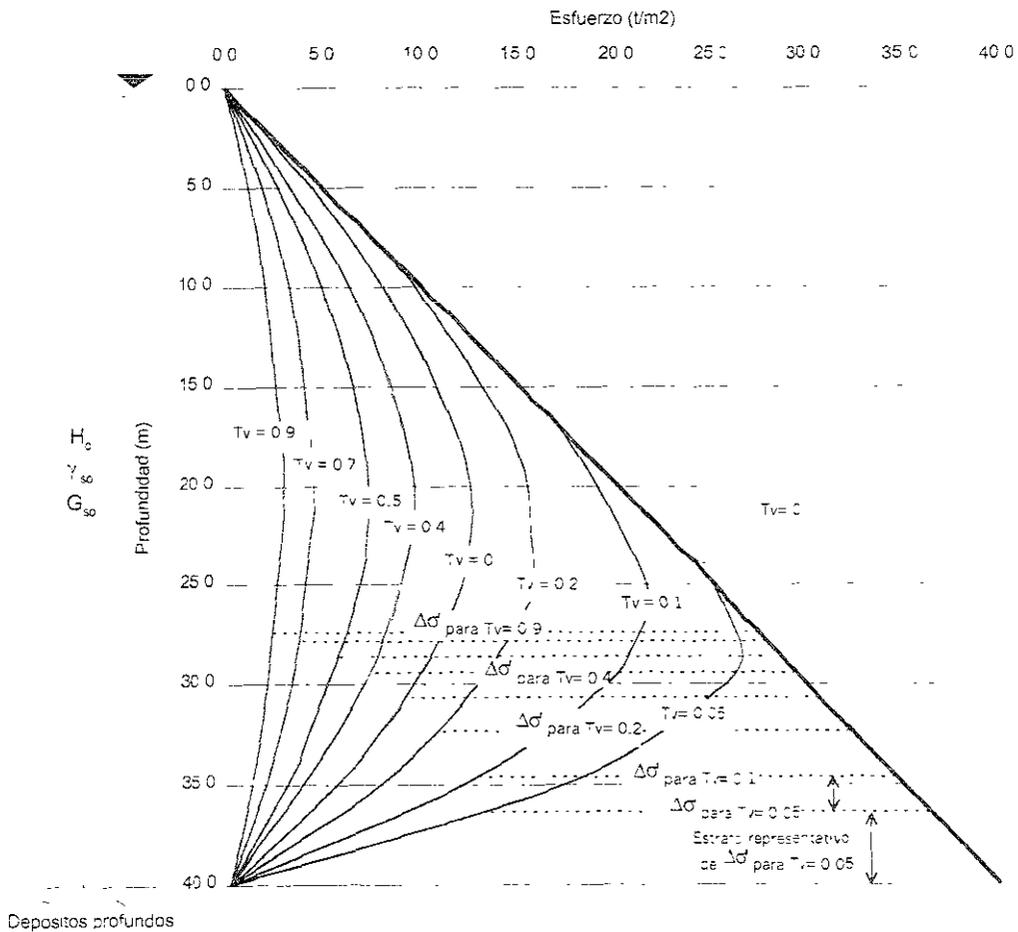


Fig. 3.14 Curvas isocronas para una distribución de exceso de presión de poro inicial triangular

Así mismo, se calcula el incremento del peso volumétrico del suelo ( $\Delta\gamma_s$ ) con la ecuación

$$\gamma_{s1} = \left( \frac{S_s - e_1}{1 + e_1} \right) \gamma_o \quad e_1 = e_o - \Delta e,$$

siendo  $\Delta e_1 = a_v (\Delta\sigma'_v)$  y  $\gamma_o = 1 \text{ t/m}^3$ . Estos incrementos de  $G$  y  $\gamma_s$  afectan únicamente al estrato representativo del área de disipación de  $u$  para cada  $T_v$  (ver fig. 3.14). La respuesta de los depósitos arcillosos se supuso lineal y se calculó con un programa de respuesta que considera la propagación bidimensional de ondas de corte SH en suelos estratificados horizontalmente, RADSH (Bárcena y Romo, 1993). Se utilizó como excitación en la base de los depósitos, el espectro de respuesta de aceleraciones del sismo del 19 de septiembre de 1985 registrado en terreno firme (estación CU). En este ejercicio, el amortiguamiento del suelo se mantuvo constante (2.0%) durante el proceso de consolidación. Desde luego que se está consciente que al aumentar  $G$  disminuye el amortiguamiento, por lo que las ordenadas espectrales calculadas subestiman a las que se hubieran obtenido de haberse modificado el amortiguamiento en función de  $G$ .

La respuesta en términos de espectros de aceleraciones con 5% de amortiguamiento se presenta en la fig. 3.15 para cada factor tiempo, donde se observa como las amplitudes espectrales máximas y el periodo fundamental de los depósitos de suelo blando disminuyen con el tiempo. Esto se debe al incremento global en  $G_{so}$  y a la disminución en  $H_o$ . En la fig. 3.15, también se muestra la variación del periodo elástico natural del suelo con el factor tiempo,  $T_v$ . Para  $T_v$  pequeños la tasa de variación de  $T_s$  es alta. Esta disminuye sensiblemente con el paso del tiempo, de acuerdo con el fenómeno de consolidación.

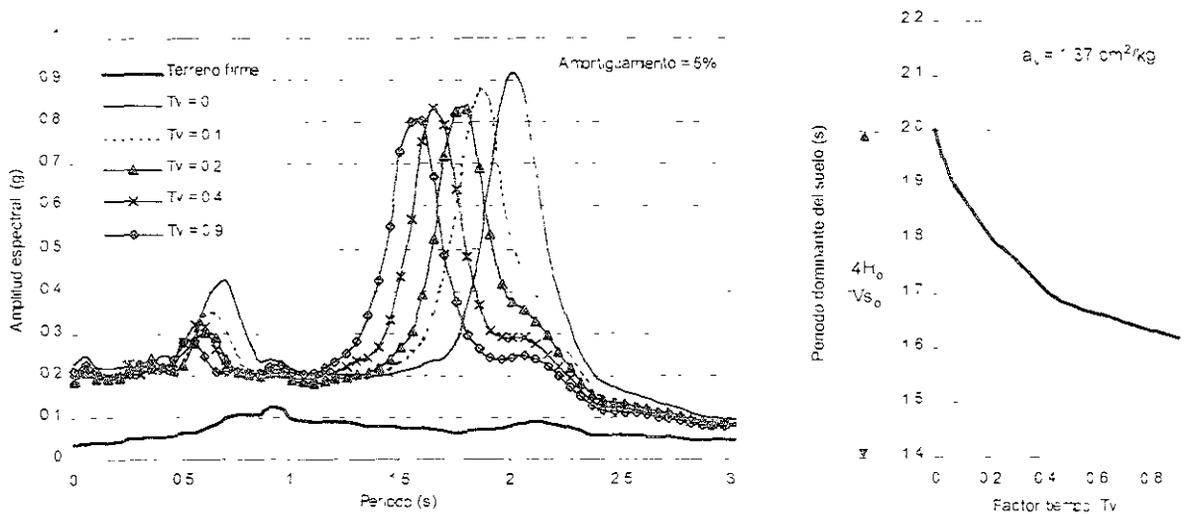


Fig 3.15 Efecto de la consolidación regional en la respuesta sísmica

Cabe mencionar, que cuando ocurrió el sismo de 1985 el proceso de consolidación en la ciudad de México ya estaba presente, por lo que la hipótesis de partida son condiciones extremas, representativas del problema en general. Sin embargo, es posible analizar otros aspectos, como por ejemplo, en la fig. 3.16 se muestra cómo varía la respuesta del terreno debido a la consolidación regional para el mismo problema planteado, pero ahora considerando una profundidad promedio del nivel freático de 2.5 m, y también que el sismo de 1985 ocurrió para un  $T_v=0.2$ . De nueva cuenta, se presenta el mismo efecto observado anteriormente.

Es notorio que el efecto de consolidación regional, en intervalos de  $T_v$  más acordes con la realidad es de poca importancia en cuanto al cambio en el periodo natural del terreno. Desde el punto de vista de diseño de estructuras, los resultados muestran que es saludable tener un edificio con un periodo natural ligeramente superior al del suelo (para las condiciones de éste en el momento de la construcción), ya que con el tiempo la severidad de los movimientos sísmicos del terreno decrecerá. Por consiguiente el tiempo mitiga el peligro sísmico en la estructura.

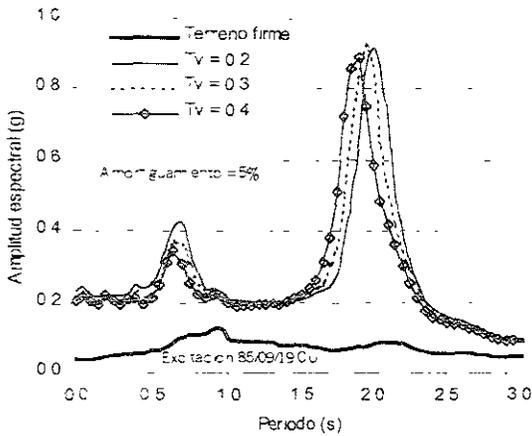


Fig 3.16 Variación de la respuesta sísmica de los depósitos arcillosos

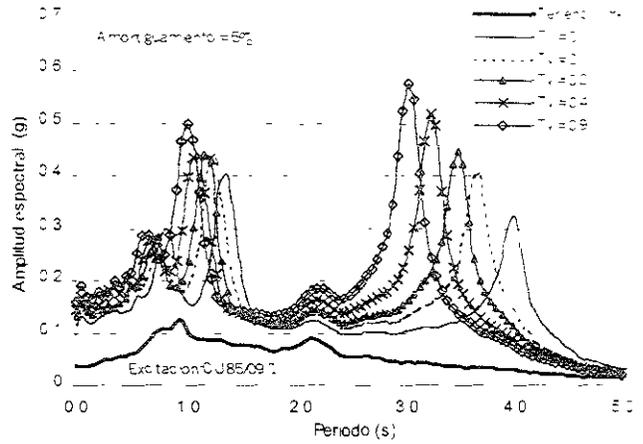


Fig 3.17 Variación de la respuesta sísmica de los depósitos arcillosos para  $H = 80\text{ m}$

En la fig. 3.17 se presenta la variación de la respuesta sísmica para las mismas condiciones del problema planteado inicialmente (ver fig. 3.14), pero ahora considerando un espesor del depósito arcilloso de 80 m. Se observa de nuevo, que el periodo característico del sitio disminuye con el tiempo, sin embargo, las aceleraciones espectrales aumentan, lo cual ocasionaría un mayor peligro sísmico para estructuras cimentadas en tales espesores arcillosos y con periodos naturales menores al del suelo para condiciones iniciales de construcción. De la misma forma, se pueden aplicar otras hipótesis al problema como sería considerar que el estrato de suelo donde se extrae el agua se encuentra a una cierta distancia más abajo de los depósitos de arcilla, lo cual ocasionaría que la presión hidrostática en la base del suelo arcilloso no se dispararía completamente para un instante después de  $T_v=0$ , entre otras.

---

## CAPÍTULO 4

# MODELO DE RESPUESTA SÍSMICA EN EL VALLE DE MÉXICO MEDIANTE REDES NEURONALES ARTIFICIALES

---

Se presenta en este capítulo, el modelo de respuesta sísmica aplicando la técnica de RNA. Primero se describe el modelo neuronal, después se detallan las variables representativas del fenómeno de respuesta sísmica, y al final se indica que información integra la base de datos para el entrenamiento y generalización del modelo con RNA.

### 4.1 DESCRIPCIÓN DEL MODELO

En la ciudad de México se han reconocido los efectos que tienen las condiciones locales del suelo en la respuesta sísmica desde mediados del siglo XX (Rosenblueth, 1952; Herrera y Rosenblueth, 1965); sin embargo, fue hasta los eventos sísmicos de septiembre de 1985 cuando tales efectos se apreciaron notablemente, siendo una fuerte evidencia de que los depósitos arcillosos en esta localidad modifican significativamente los movimientos del terreno superficial aún ante pequeñas variaciones en sus propiedades. Desde entonces, se han desarrollado varias investigaciones en el área de ingeniería sísmica y geotécnica para tratar de explicar el efecto que tienen las condiciones locales del sitio en los movimientos sísmicos incidentes dentro del valle de México (p.e. Romo y Jaime, 1986; Romo y Seed, 1986; Sánchez-Sesma, 1989; Chávez-García y Bard, 1994; Romo, 1995). El impulso de estas investigaciones permitió establecer una serie de procedimientos para calcular los movimientos superficiales y a profundidad del terreno de campo libre (algunos de ellos mencionados en el capítulo anterior) y establecer los criterios sísmicos para el diseño de edificios cimentados en los depósitos de suelos blandos.

También desde entonces la colocación de varios instrumentos de medición han proporcionado información muy valiosa para el desarrollo de nuevas herramientas de análisis (Ordaz et al., 1998; Romo, 1999). Asimismo, estos eventos sísmicos medidos y almacenados desde 1985 han constituido una base de datos importante, lo que permite aplicar procedimientos basados en el aprendizaje (RNA) para modelar la respuesta sísmica de los depósitos de suelo en el valle de México.

En la fig. 4.1 se presenta esquemáticamente el modelo neuronal de respuesta sísmica que se propone en este trabajo, donde las RNA operan en forma equivalente a una función de transferencia empírica, es decir, a partir de los movimientos sísmicos en terreno firme y de las características dinámicas del sitio se determinan los movimientos superficiales de campo libre en términos de valores espectrales

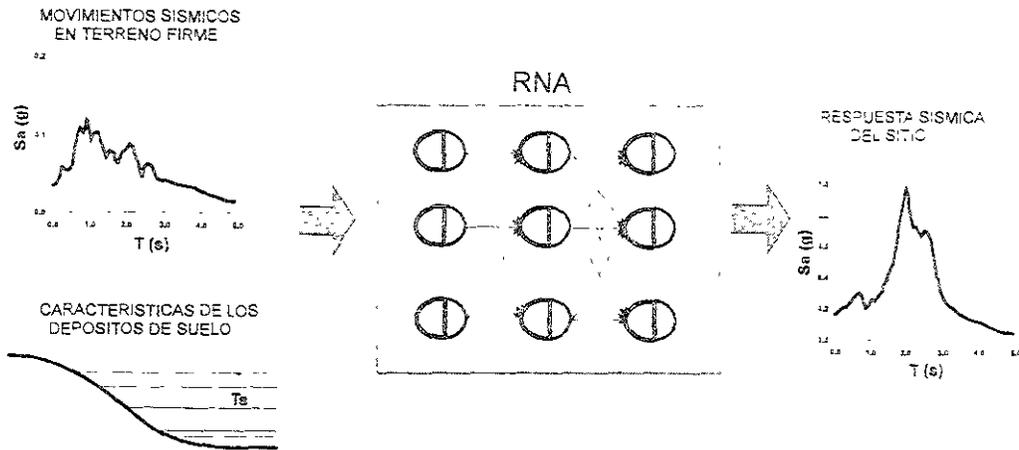


Fig 4.1 Modelo de respuesta sísmica mediante Redes Neuronales Artificiales

De acuerdo con la fig. 4.1, las variables principales que alimentan este modelo son:

- El periodo ( $T$ ) que es la abscisa del espectro de respuesta en terreno firme.
- La aceleración espectral (ordenada) en terreno firme ( $Sa_{CU}$ ) para el correspondiente periodo  $T$ .
- El periodo característico del sitio ( $T_s$ ), donde se desea conocer la respuesta

La respuesta del modelo neuronal (variable de salida o del comportamiento deseado) es la aceleración espectral en el sitio ( $Sa_{SIT}$ ) para el correspondiente periodo  $T$ . Dentro del grupo de aplicaciones de las RNA, el modelo presentado corresponde al de Aproximación de Funciones, donde los datos en la unidad de salida (respuesta) deben ser los correspondientes a los datos que se dan de entrada. De esta forma, las variables anteriores integran la base de datos para el diseño (etapas de entrenamiento y predicción) de las RNA.

#### 4.2 BASE DE DATOS DEL MODELO DE RESPUESTA SÍSMICA CON RNA

Como se ha mencionado, uno de los aspectos importantes en la aplicación de la herramienta de RNA es la base de datos disponible, que se integra con los parámetros que servirán como datos de entrada y salida para el modelo considerado en la fig. 4.1. A continuación se presenta una descripción de la información disponible y del modelo de respuesta sísmica.

##### 4.2.1 Zonificación Geotécnica y Sísmica del Valle de México

La ciudad de México se divide geotécnica y sísmicamente en tres regiones con diferentes condiciones del subsuelo (ver fig. 4.2):

- **Zona de Lomas** (zona I). Se encuentran formaciones de rocas por derrames de lava o suelos generalmente firmes (tobas volcánicas), aunque pueden existir algunos estratos arenosos en estado suelto o cohesivos relativamente blandos ya sea intercalados o en la superficie.
- **Zona de Transición** (zona II). Integrada predominantemente por estratos arenosos y limo-arenosos intercalados con capas de arcilla lacustre. Su espesor es variable entre decenas de centímetros y no más de 20 m aproximadamente, donde más abajo se encuentran los depósitos profundos.

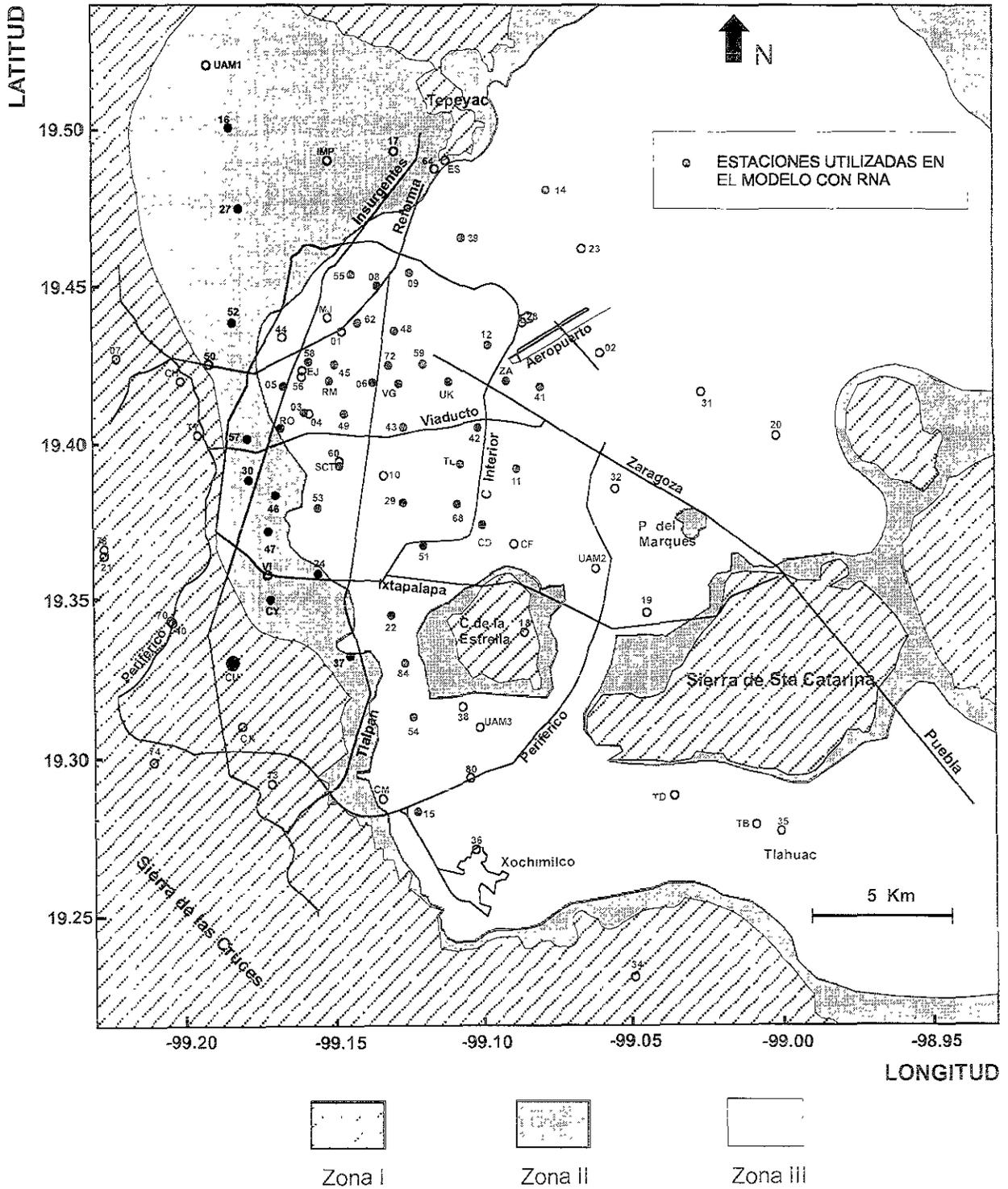


Fig 4.2 Zonificación Geotécnica del valle de México (propuesta por Auvinet, Méndez y Juárez, 2000) y ubicación de las estaciones acelerométricas de la RACM

- **Zona de Lago** (zona III). Está constituida por potentes depósitos de arcilla de alta compresibilidad y alto contenido de agua (espesores entre 20 y 100 m), donde se encuentran intercaladas capas arenolimosas o arcillosas de consistencia firme a muy dura con espesores variables de centímetros a varios metros. Estos depósitos arcillosos suelen estar cubiertos superficialmente por suelos aluviales y rellenos artificiales, y debajo de ellos se tienen los llamados depósitos profundos (capas rígidas de arena limosa cementadas)

En la fig 4.3 se presenta una estratigrafía típica de la ciudad de México donde la toba volcánica de la zona de Lomas se localiza al oeste del centro de la ciudad, mientras que los suelos blandos de la zona de Lago generalmente consisten de dos formaciones blandas (la superior y la inferior) separadas por una capa de arena compacta de espesor variable entre 0.0 y 6.0 m llamada capa dura. El nivel de aguas freáticas generalmente se encuentra a 2.5 m de profundidad promedio en la mayor parte de la zona de Lago. En el mismo perfil estratigráfico se muestra también, cómo en la zona de Transición se tienen estratificaciones erráticas de depósitos aluviales

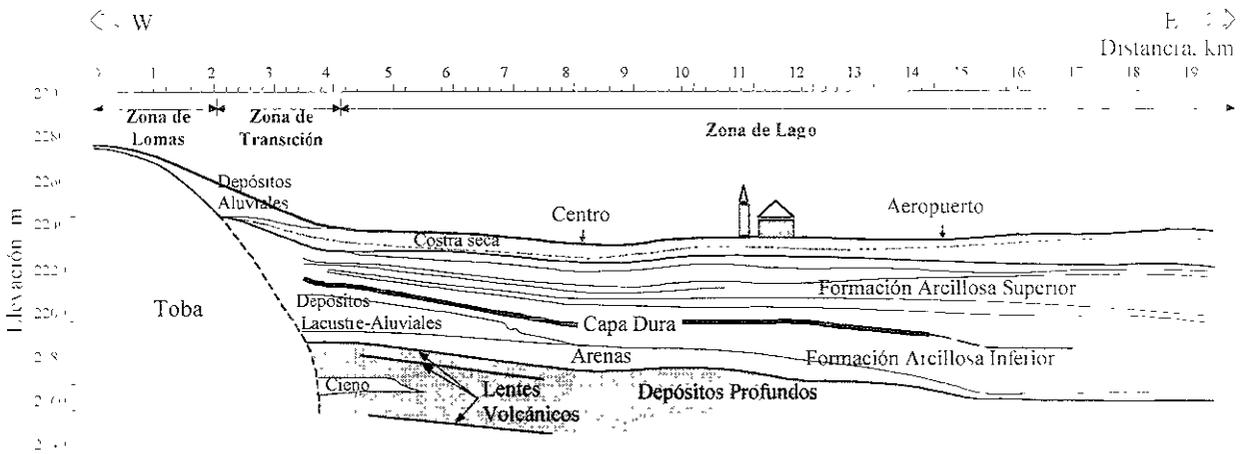


Fig 4.3 Estratigrafía general de la ciudad de México (propuesta por Mooser, 2000)

Por otra parte, la República Mexicana está ubicada en una zona de alta sismicidad puesto que en las costas del Pacífico se generan movimientos bruscos en la corteza terrestre a lo largo de las fallas que separan las diferentes placas que la forman en esta región (placa de Cocos y placa Norteamericana). Los sismos que afectan a la ciudad de México se clasifican en cuatro grupos (Rosenblueth y Ovando, 1990) a) los locales, que se originan en el interior del valle de México o en sus inmediaciones, b) los que se originan en el resto de la placa Norteamericana, c) los de subducción de la placa de Cocos bajo la Norteamericana y d) los de la placa de Cocos subducida originados a profundidades entre 50 y 100 Km. Los temblores originados por la subducción de la placa de Cocos son los que más estragos han causado en la ciudad de México, de hecho a este grupo pertenece el 85/09/19

Debido a la alta sismicidad y ante las amplificaciones de los movimientos superficiales del terreno por las condiciones locales del suelo en el valle de México, se han desarrollado diferentes estudios e investigaciones dentro de la ingeniería sísmica geotécnica. Gran parte de estos avances se han logrado gracias al incremento del número de instrumentos de medición, lo que ha permitido tener una mejor observación e interpretación de los registros de movimientos fuertes ocurridos en los últimos años (desde 1986).

En 1985 operaban en la ciudad de México del orden de diez estaciones acelerométricas; en la actualidad, existen más de doscientas, las cuales integran la conocida Red Acelerométrica de la Ciudad de México (RACM). En la fig. 4.2 se presentan algunas de las estaciones de la RACM junto con la zonificación geotécnica del valle de México antes descrita y las principales vías de la ciudad.

#### 4.2.2 Movimientos Sísmicos del Terreno y Características de los Sitios empleados en el Modelo

Los registros de movimientos sísmicos utilizados para formar la base de datos, con la cual se entrenaron las RNA, se obtuvieron a partir de la Base Mexicana de Datos de Sismos Fuertes (Aicántara, et al., 1997). Este sistema integra toda la información que disponen las diferentes instituciones que conforman la RACM (Instituto de Ingeniería, Fundación JBS, Fundación ICA, CENAPRED, etc.). En la fig. 4.2 se indican con punto negro los sitios seleccionados de la RACM para integrar la base de datos usada para el diseño de la RNA con más de 40 estaciones en la superficie libre del terreno, donde en cada una de ellas se han registrado entre 5 y 11 eventos sísmicos. Se observa también, que la mayoría de los sitios se ubican en la zona de Lago donde se han presentado las mayores amplificaciones de los movimientos del terreno.

##### 4.2.2.1 Excitación sísmica

La excitación sísmica del modelo neuronal mostrado en la fig. 4.1, se caracteriza por los movimientos superficiales registrados en la estación acelerométrica CU (Ciudad Universitaria), la cual se ubica al noreste del centro de la ciudad en la zona geotécnica I (ver fig. 4.2). Los sismos registrados en este sitio y utilizados para alimentar el modelo neuronal se presentan en la Tabla 4.1. Estos eventos se seleccionaron considerando sus características principales, las cuales se muestran en la misma Tabla 4.1.

Tabla 4.1 Características generales de los movimientos sísmicos de entrada (sitio CU), orientación E-W

Fecha (año/mes/día)	Epicentro		Prof (km)	Distancia epicentral (km)	Ms	Amax $\times 10^{-3}$ (g)
	Lat N	Long W				
85/09/19	18.081	102.942	15	419	8.1	35.00
88/02/08	17.494	101.157	19.2	291	5.5	2.20
89/04/25	16.603	99.400	19	304	6.7	10.30
90/05/11	17.046	100.840	11.7	308	5.1	1.60
93/05/15	16.430	98.740	<5	326	5.8	1.44
93/07/29	17.380	100.650	<5	266	5.0	0.70
93/10/24	16.540	99.980	<5	321	6.5	4.10
94/03/14	15.670	93.010	95	770	6.5	0.60
94/05/23	18.030	100.570	23	205	5.6	4.50
94/07/04	14.830	97.290	31	539	5.9	0.50
94/12/10	18.020	101.560	20	290	6.3	5.90
95/09/14	16.310	98.880	22	337	7.3	3.20
95/10/09	18.740	104.670	5	580	7.5	2.60
95/10/30	16.380	98.500	21	336	5.3	0.40
97/01/11	18.090	102.860	17	411	7.3	5.80
99/06/15	18.200	97.470	92	220	6.7	11.85
99/09/30	15.890	97.070	10	443	7.4	7.80

Estos movimientos superficiales en CU se proporcionan a la RNA en términos de espectros de respuesta de aceleración absoluta correspondientes a un sistema de un grado de libertad con comportamiento elástico lineal y un amortiguamiento del 5% ( $Sa_{CU}$ ). En la fig. 4.4 se presentan los espectros de respuesta de estos movimientos (Tabla 4.1) normalizados con respecto a la aceleración máxima de terreno ( $A_{max}$ ), junto con las curvas de valores medios ( $\mu$ ), más una y más dos desviaciones estándar ( $\sigma$ ), lo cual nos indica la variación en amplitud y contenido de frecuencias de las señales seleccionadas. Se observa que el rango de periodos (abscisa del espectro) utilizado en el modelo para definir los espectros de respuesta de las señales de excitación (sitio CU) es de 0.0 a 5.0 s, mientras que las ordenadas espectrales varían de 0.1 a 4.3 veces la aceleración máxima del terreno ( $A_{max}$ ).

La normalización de los espectros en la fig. 4.4, elimina el factor de intensidad de los eventos registrados en roca y muestra únicamente el efecto que tienen las fuentes donde se libera la energía sísmica y las diferentes trayectorias que siguen las ondas hasta el sitio, lo cual se observa en el intervalo de periodos donde se tiene la amplitud espectral máxima ( $Sa_{max}$ ) que varía entre 0.30 y 1.60 s

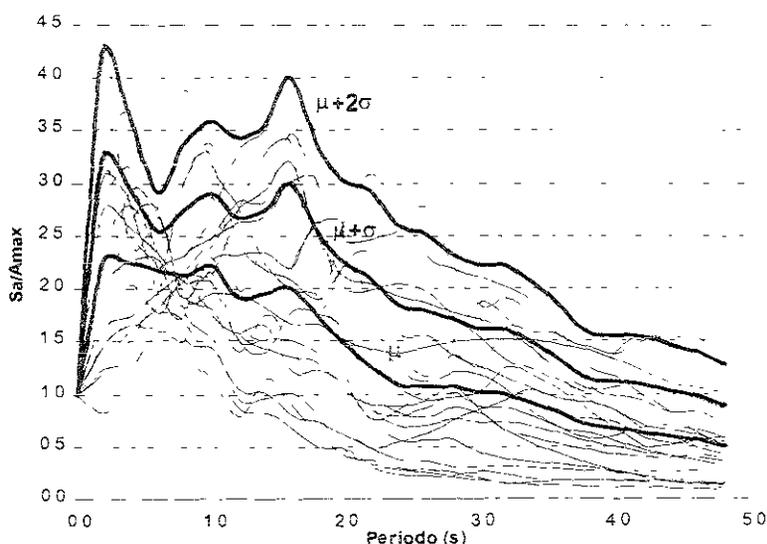


Fig. 4.4 Formas espectrales de los movimientos sísmicos de excitación (estación CU) utilizados en el modelo

La estación CU se seleccionó como sitio de referencia para alimentar el modelo neuronal debido principalmente a que en este sitio se han registrado la mayor cantidad y los principales eventos sísmicos incidentes en el valle de México, y además se ha utilizado como base en estudios de respuesta sísmica (Romo, 1987). De cualquier forma, la entrada del modelo neuronal es relativa, es decir, se puede utilizar otro sitio de referencia, como por ejemplo CDAO, y de ahí determinar la respuesta en los demás sitios centro de la ciudad de México. La RNA debe ser capaz de encontrar la relación que existe entre los datos de entrada y de salida, siempre y cuando la fuente de entrada sea la misma y se disponga de la información suficiente. Por lo tanto, el manejo de la base de datos disponible es un aspecto importante en el diseño adecuado de una RNA. Mientras mayor sea la información utilizada, los resultados con la técnica de RNA y con cualquier modelo deberán mejorar.

#### 4.2.2.2 Caracterización de los depósitos de suelo

Debido a que se tienen diferentes condiciones del subsuelo en el valle de México, el modelo neuronal se separa en dos módulos generales para el diseño (entrenamiento) de las RNA. Uno modela la respuesta sísmica en sitios ubicados en la zona de Lago (LAGO), y el otro en la zona de Transición (TRANS)

Los depósitos de suelo en ambas zonas se caracterizan dinámicamente por sus periodos naturales elásticos ( $T_s$ ) que se definieron de previos estudios de campo y laboratorio para algunos sitios donde se tenía como objetivo evaluar la rigidez dinámica de los estratos de suelo en la ciudad de México. Para los demás sitios donde no se dispone de tal información, sus periodos característicos se aproximaron por medio de funciones de transferencia empíricas a partir de sus espectros de pseudovelocidades relativas al 2% de amortiguamiento. Estos periodos ( $T_s$ ) coinciden aproximadamente con el mapa de contornos de isoperiodos dominantes propuestos por Lermo et al. (1988). En la Tabla 4.2 se presentan las características generales de los sitios considerados para el módulo LAGO (zona Lacustre), el cual se integra por 33 sitios (ver fig. 4.2) con periodos característicos que varían entre 1.2 y 4.2 s. Estos sitios se seleccionaron de acuerdo a la cantidad de sismos registrados en cada uno de ellos

Tabla 4.2 Características generales de los sitios de suelo ubicados en la zona de Lago

No	Clave (mapa)	Nombre de la estación	Ubicación		Ts (s)	No	Clave (mapa)	Nombre de la estación	Ubicación		Ts (s)
			Lat N	Long W					Lat N	Long W	
1	JC54 (54)	Jard de Coyacacan	19 313	99 127	1.20	18	HJ72 (72)	Hospital Juarez	19 425	99 130	2.60
2	MI15 (15)	Mramontes	19 283	99 125	1.30	19	XP06 (06)	Xochpilli	19 420	99 135	2.60
3	DFRO (RO)	Roma	19 405	99 166	1.30	20	BO39 (39)	Bondoyrto	19 465	99 105	2.70
4	IB22 (22)	Ibero	19 345	99 130	1.40	21	BA49 (49)	Buenos Aires	19 410	99 145	2.80
5	CH84 (84)	Culhuacan	19 330	99 125	1.40	22	VM29 (29)	Villa del Mar	19 381	99 125	2.90
6	SI53 (53)	San Simon	19 375	99 148	1.40	23	CA59 (59)	Candelana	19 426	99 118	3.00
7	SCT (SCT)	SCT	19 393	99 147	2.00	24	JA43 (43)	Jamaica	19 405	99 125	3.00
8	CJ03 (03)	C U Juarez	19 410	99 157	2.05	25	AP68 (68)	Apilaco	19 381	99 107	3.10
9	TL08 (08)	Tlatelolco	19 450	99 134	2.10	26	AR14 (14)	Aragon	19 481	99 076	3.40
10	CI05 (05)	Cibeles	19 419	99 165	2.10	27	CDAO (CD)	Central Apastos Ofic	19 372	99 096	3.50
11	TL55 (55)	Tlatelolco	19 436	99 143	2.15	28	UNKS (UK)	Unid Kennedy Sup	19 419	99 111	3.50
12	RMAS (RM)	Roma A Sup	19 420	99 155	2.20	29	TLAS (TL)	Tlacotal Sup	19 397	99 105	3.60
13	GA62 (62)	Ganbaici	19 439	99 140	2.20	30	DM12 (12)	Deport Moctezuma	19 431	99 096	3.60
14	LI58 (58)	Liverpool	19 426	99 157	2.20	31	ZARS (ZA)	Zaragoza Sup	19 419	99 088	4.00
15	VG09 (09)	Valle Gomez	19 454	99 123	2.40	32	HA41 (41)	Hangares	19 418	99 079	4.00
16	BL45 (45)	Balderas	19 425	99 148	2.40	33	AU11 (11)	Autódromo	19 392	99 087	4.20
17	RM48 (48)	Rodolfo Menendez	19 436	99 128	2.50						

Tabla 4.3 Características generales de los sitios de suelo ubicados en la zona de Transición

No	Clave (mapa)	Nombre de la estación	Ubicación		Ts (s)
			Lat N	Long W	
1	CO47 (47)	Coyoacán	19 371	99 170	0.50
2	DR16 (16)	Deport. Reynosa	19 501	99 183	0.55
3	COYS (CY)	Est. No. 6, Coy	19 348	99 169	0.60
4	EO30 (30)	Esparza Oteo	19 389	99 177	0.65
5	GR27 (27)	Granjas	19 475	99 180	0.80
6	ES57 (57)	Escandon	19 402	99 177	0.85
7	AU46 (46)	Angel Urraza	19.383	99 168	0.85
8	ME52 (52)	Manano Escobedo	19 438	99 182	0.90
9	DX37 (37)	Xotepingo	19 332	99 144	0.95
13	AO24 (24)	Alberca Olímpica	19 358	99 154	1.00

En la Tabla 4.3 se presentan también, las características de los sitios seleccionados para el módulo **TRANS** y que se ubican en la fig. 4.2 con un punto negro. Se muestra que la base de datos para este módulo se integra con 10 sitios, donde sus periodos dominantes varían entre 0.5 y 1.0 s. El rango de periodos para esta zona es bastante estrecho, en comparación los sitios en la zona III, lo cual debe facilitar el entrenamiento del modelo neuronal. Mas adelante (sección 4.3) se indica que eventos sísmicos se disponen en cada sitio para ambas zonas geotécnicas, y cuales se consideran como datos de entrenamiento y predicción para alimentar el modelo con RNA.

#### 4.2.2.3 Movimientos de respuesta medidos en superficie

El patrón objetivo del modelo con RNA son los movimientos sísmicos superficiales medidos en los sitios ubicados en las zonas de Lago y Transición (ver Tabla 4.2 y 4.3); los cuales se presentan a la red en términos de aceleraciones espectrales (**SaSIT**) correspondientes a un sistema de un grado de libertad con comportamiento elástico lineal y un amortiguamiento del 5%. En la fig. 4.5 se muestran los espectros de respuesta normalizados con respecto a  $A_{max}$  de los eventos sísmicos considerados para cada uno de los sitios de la zona de Lago.

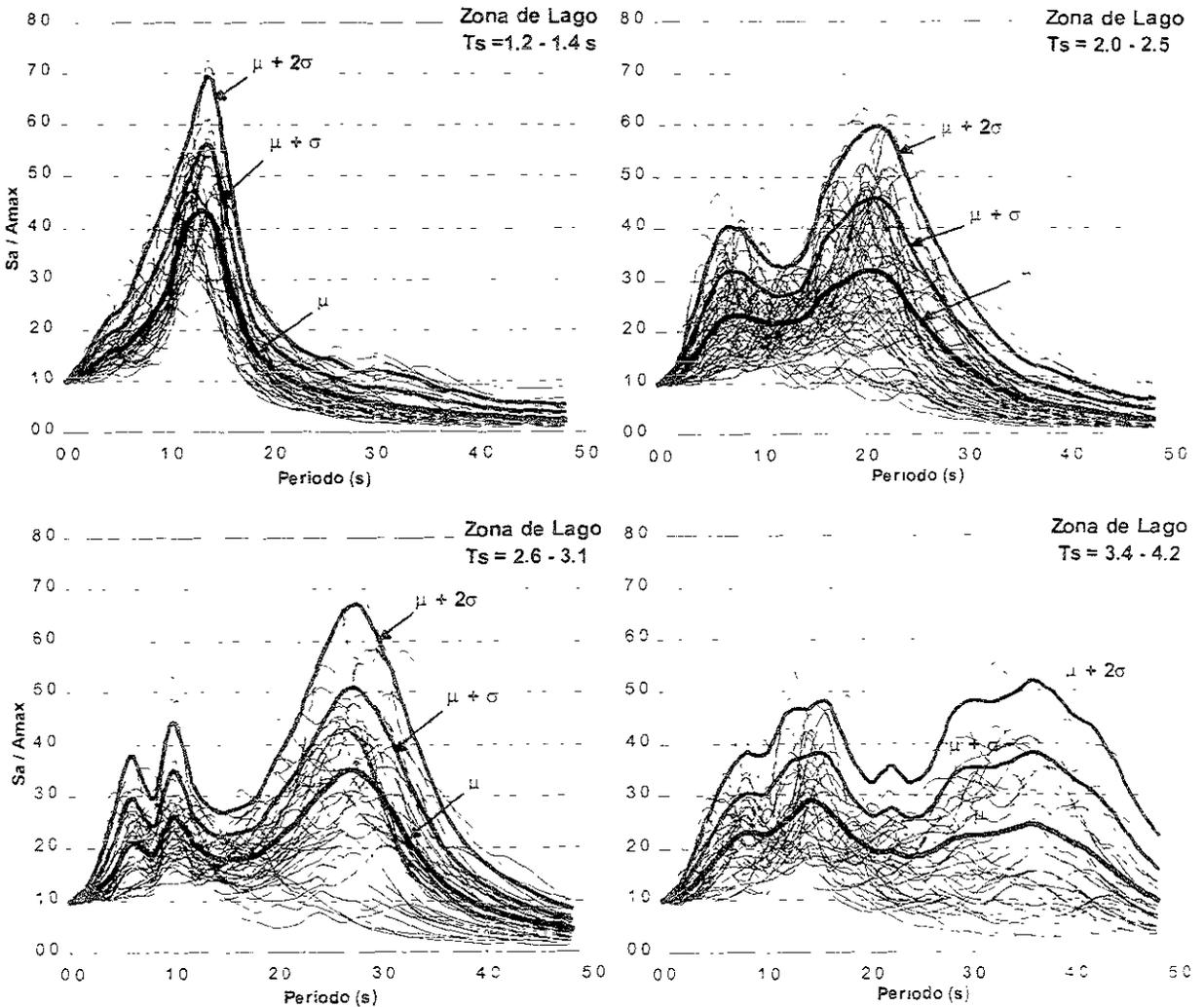


Fig. 4.5 Formas espectrales de los movimientos sísmicos de respuesta medidos (zona de Lago)

Debido a que los periodos naturales de los sitios ( $T_s$ ) en esta zona III varían en un amplio rango de valores (1.2 - 4.2 s), se ha propuesto subdividir la base de datos del módulo LAGO en cuatro subcasos. Los intervalos de periodos característicos propuestos para cada caso se presentan en las mismas gráficas de la fig. 4.5, y posteriormente en la Tabla 4.4 se indica qué eventos sísmicos se utilizan. Esta subdivisión mejora la presentación de los datos a las RNA, y por consiguiente se pueden obtener entrenamientos en menor tiempo.

En las gráficas anteriores (fig. 4.5) se aprecia la variación que existe entre las ordenadas espectrales y el contenido de frecuencias de los movimientos de respuesta medidos en superficie. El rango de estas ordenadas espectrales varía de 0.1 a 7.0 veces la aceleración máxima del sitio ( $A_{max}$ ), mientras que el periodo (abscisa del espectro) varía entre 0.0 y 5.0 s

De la misma forma, en la fig. 4.6 se presentan los espectros normalizados de los sismos registrados en los sitios que se ubican en la zona II (Transición) de acuerdo con la Tabla 4.3. Estas formas espectrales definen mejor el rango de periodos característicos de sitio en esta zona geotécnica ( $T_s$  entre 0.5 y 1.0 s) en comparación con los de la zona de Lago. La magnitud de las ordenadas espectrales varía entre 0.1 y 6.0 veces la aceleración máxima del sitio ( $A_{max}$ ), mientras que los valores de las abscisas (periodo,  $T$ ) varían de 0.0 a 4.0 s.

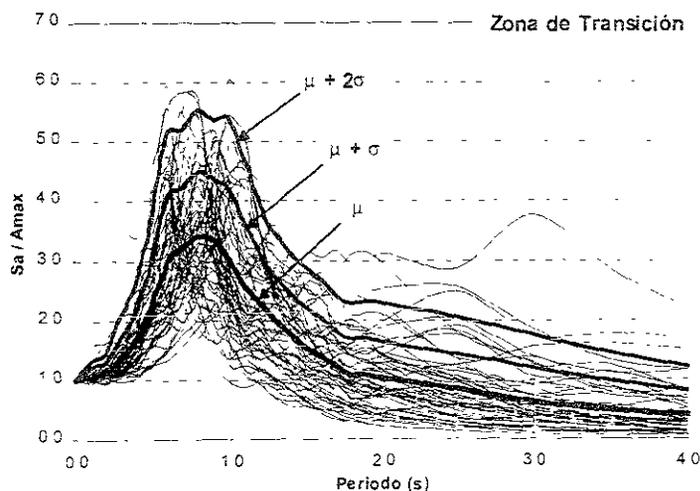


Fig 4.6 Formas espectrales de los movimientos sísmicos de respuesta (zona de Transición)

En las figuras anteriores (4.5 y 4.6) también se muestran las curvas para valores medios ( $\mu$ ), más una y más dos desviaciones estándar ( $\sigma$ ), donde se puede apreciar el comportamiento general de los movimientos sísmicos de los depósitos arcillosos del valle de México, estos mismos resultados se presentaron y discutieron con anterioridad en la sección 3.2.4 (fig. 3.13). Sin embargo, es importante mencionar que las formas espectrales presentadas son muestra de la variabilidad espacial de los datos disponibles, principalmente dentro del rango de periodos dominantes que es una de las características más importantes de la respuesta sísmica y que el modelo con RNA debe reproducir en periodo y magnitud espectral. Por ello, es necesario seleccionar primero la información que servirá para formar la base de datos de entrenamiento del modelo y también la base de datos de prueba (predicción). En estas etapas, es necesario encontrar la mejor forma de proporcionar los datos a la RNA, para que la ejecución de la etapa de entrenamiento sea eficiente.

4.2.2.4 Variación de los movimientos sísmicos de un depósito de suelo

Hasta el momento se han descrito las variables que se consideran importantes para modelar el fenómeno de respuesta sísmica con RNA (ver fig. 4.1), sin embargo, existen otros factores que afectan los movimientos sísmicos de los depósitos de suelos como los presentados en la Tabla 3.1. En este trabajo se muestra un modelo neuronal sencillo como una primera aplicación de método, pero también es importante tomar en cuenta todas las variables, que aunque no están directamente relacionadas con el problema, para las redes si son representativas.

En las figs. 4.7a y b se presentan las funciones de amplificación empíricas para dos sitios en el valle de México con referencia a los movimientos registrados en el sitio CU. El primer sitio (GR27) está ubicado en la zona de Transición y el segundo (SCT) en la zona del Lago. Como se mencionó en la fig. 4.1, idealmente el modelo con RNA debe reproducir estas funciones de amplificación empíricas. Sin embargo, se puede apreciar que aún en un mismo sitio existen variaciones considerables en las amplitudes espectrales principalmente en sus periodos característicos. Se observa cómo el factor de amplificación en el periodo dominante del sitio varía conforme aumenta la aceleración máxima en terreno firme ( $A_{maxCU}$ ) que es una medida del grado de intensidad del sismo, lo que indica que esta variable puede ser representativa en el modelo neuronal para calcular la respuesta sísmica

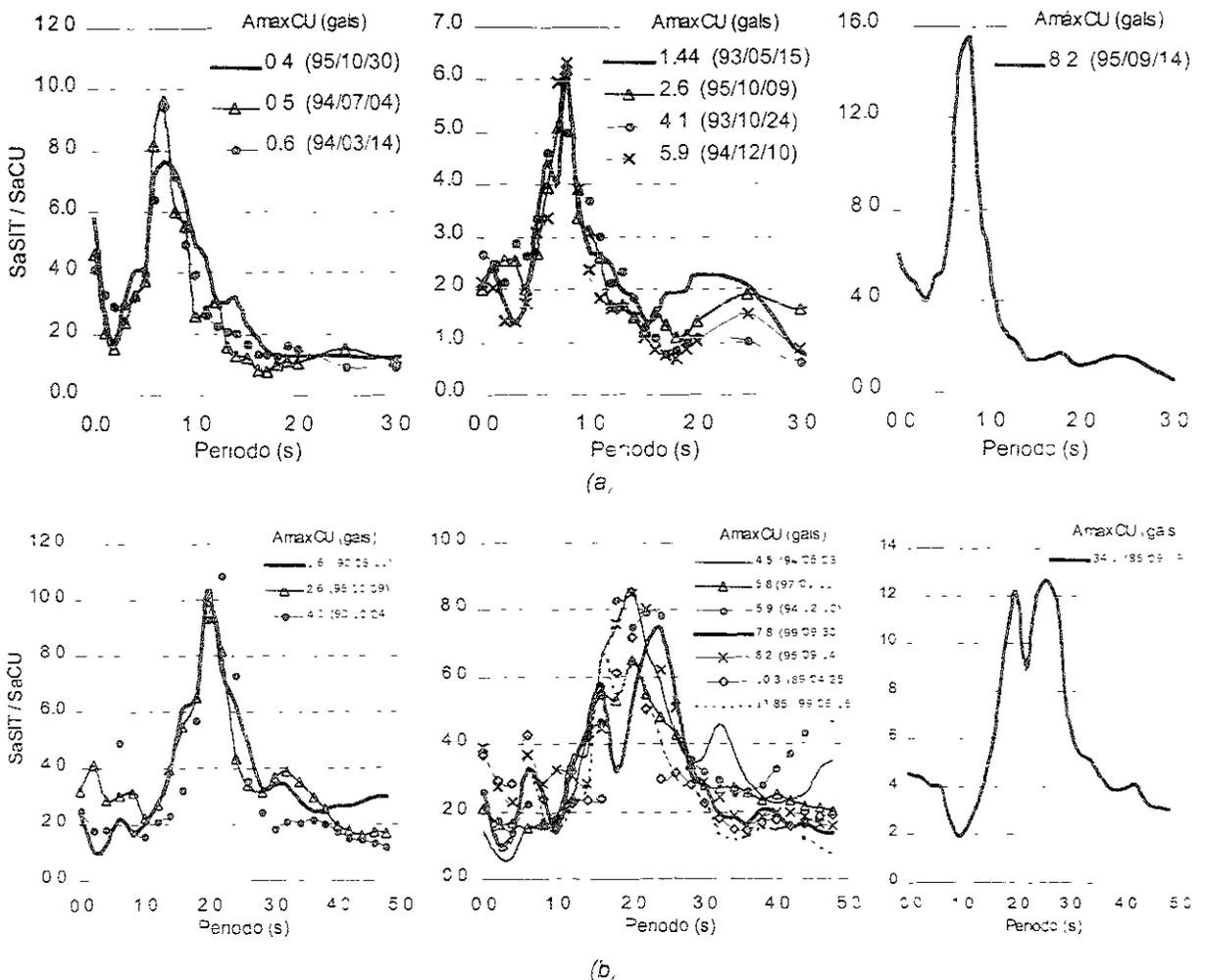


Fig 4.7 Funciones de amplificación empíricas, (a) sitio GR27 ( $T_s=0.8$  s) y (b) sitio SCT ( $T_s=2$ s)

Las distintas funciones de amplificación medidas en un mismo sitio indican que el tipo de ondas sísmicas incidentes, el contenido de frecuencias, la energía sísmica, etc., afectan la respuesta sísmica del depósito. Todos estos factores, incluso aquellos que se desconocen, influyen en la forma espectral de la señal registrada en terreno firme, una de las variables de entrada del modelo (**SaCU**). Sin embargo, se requiere de otra variable en el modelo neuronal, además de las presentadas, que indique a la red que toda la forma espectral influye en la respuesta sísmica (**SaSIT**) para un determinado periodo ( $T$ ). De acuerdo con la fig. 4.7 y el razonamiento anterior, se considera la aceleración máxima en el sitio CU ( $A_{máx, CU}$ ) como otra variable de entrada del modelo neuronal. Aunque esta variable ya se considera en la amplitud espectral **SaCU** (periodo  $T=0$ ) para la RNA es un indicador del nivel de intensidad de los movimientos en terreno firme.

Uno de los factores que influye en las características de los movimientos del terreno es el comportamiento no lineal del suelo. Resultados obtenidos a partir de estudios de campo y laboratorio realizados por Romo y Jaime (1986) mostraron que las arcillas de la ciudad de México presentan un comportamiento casi elástico lineal. Sin embargo, estudios espectrales de registros sísmicos en la zona de Lago (Barcena y Romo, 1993) señalaron que en ciertos sitios se han desarrollado efectos no lineales. Estos resultados concuerdan con otras investigaciones teóricas (Romo, 1991) y experimentales (Romo, Jaime y Taboada, 1989; Romo y Ovando, 1995; Romo, 1995) en las cuales se ha señalado que el comportamiento de la arcilla de la ciudad de México depende fundamentalmente del índice de plasticidad ( $I_p$ ), donde los efectos no lineales disminuyen al aumentar este parámetro. El comportamiento no lineal es más importante cuando  $I_p \leq 150\%$ . En las zonas de Lago y Transición existen depósitos de arcillas con índices de plasticidad menor de 150%, como en los sitios presentados en las figs. 4.7a y b. Sin embargo, la existencia de efectos no lineales todavía es un tema de discusión debido a que no se ha comprobado completamente este fenómeno a partir de registros sísmicos considerando variaciones por incertidumbre.

### 4.3 DATOS PARA LA ETAPA DE ENTRENAMIENTO Y PRUEBA DE LA RNA

El grupo de datos para la etapa de entrenamiento de una RNA, es un subconjunto de la muestra total de datos disponibles. Después de entrenar una RNA (o al mismo tiempo) se evalúa o prueba la eficiencia de la red con otro subconjunto de la muestra total de datos y que no se consideran en el grupo de entrenamiento. El proceso de evaluación se conoce comúnmente como etapa de generalización o predicción de la RNA y en la cual se determina la arquitectura de red óptima.

En la Tabla 4.4 se presentan los eventos seleccionados para los sitios considerados en el módulo **LAGO**, donde se indica también cuales son los casos para entrenamiento y prueba. Esta misma información pero ahora para los sitios del módulo **TRANS** se muestra en la Tabla 4.5. Las características generales de los eventos sísmicos y de los sitios se presentaron en las Tablas 4.1, 4.2 y 4.3

Estos sitios se seleccionaron a partir de la cantidad de eventos sísmicos registrados en cada uno de ellos y de sus características dinámicas (**Ts**). En el caso de la zona de Transición, los sitios seleccionados son en su mayoría los disponibles (ver fig. 4.2). Para la zona de Lago se trató de abarcar primordialmente la región donde se han presentado las mayores amplificaciones de los movimientos incidentes.

Los eventos sísmicos para entrenamiento y prueba se seleccionaron arbitrariamente. Se eligieron como datos de prueba al menos un sitio con todos sus eventos sísmicos en la zona de Transición y en cada subcaso del módulo LAGO. Además, se seleccionaron arbitrariamente algunos eventos para los demás sitios.

Tabla 4.4 Eventos disponibles en los sitio de suelo utilizados en el módulo LAGO

Subcaso	Clave de la Estación (mapa)	EVENTOS (año/mes/día)																
		88/02/08	90/05/11	93/05/15	93/07/29	94/05/23	94/07/04	94/12/10	95/10/30	97/01/11	99/06/15	99/09/30	99/04/25	93/10/24	94/03/14	95/10/09	95/09/14	85/09/19
I Ts (s) 1.2 - 1.4	JC54 (54)				*	*	*	*						*	*	*	*	
	MI15 (15)			p	*	*	*	*						*	*	*	*	
	CH84 (84)					*	*	*	*					*	*	*	*	
	DFRO (RD)		c	p		p	p	p	p				p	p	*	*	p	p
	SI53 (53)			p		p	p	p	p					p	*	*	*	*
	IB22 (22)					*	*	*	*					*	*	*	*	*
II Ts (s) 2.0 - 2.5	SCT (SCT)		*			*	*	*	*	*	*	p	*	*	*	p	p	*
	CJ03 (03)								p	p					p	p	p	
	TL08 (08)				*	*	*	*	*	*				*	p	*	*	*
	CI05 (05)				*	p	*	*	*	*				*	*	*	*	*
	TL55 (55)				p	p	p	p						p	p		p	
	GA62 (62)				*	*	*	*	*	*				p	*	*	*	*
	RMAS (RM)		p		*	*	*	*	*	*	*			*	*	*	*	*
	LI58 (58)			*		*	*	*	*	*	*			*	*	*	*	*
	BL45 (45)				*	*	*	*	*	*	*			*	p	*	*	*
	VG09 (09)				*	*	*	*	*	*	*			*	*	*	*	*
RM48 (48)				*	*	*	*	*	*	*			*	*	*	*	*	
III Ts (s) 2.6 - 3.1	HJ72 (72)				*	*	*	*	*	*				*	p	*	*	*
	XP06 (06)			*		*	*	*	*	*				*	*	*	*	*
	BO39 (39)					*	*	*	*	*				p	*	*	*	*
	BA49 (49)			*		*	*	*	*	*				*	*	*	*	*
	VM29 (29)				*	p	*	*	*	*				*	*	*	*	*
	CA59 (59)				*	*	*	*	*	*				*	*	*	*	*
	JA43 (43)			p		p	p	p						p	p	p	p	
AP68 (68)					*	*	*	*	*	*			*	*	*	*	*	
IV Ts (s) 3.4 - 4.2	AR14 (14)			*		*	*	*	*	*				*	*	*	*	*
	CDAO (CD)	*	*			*	*	*	*	*	p	*	*	*	*	*	*	*
	UNKS (UK)			*		*	*	*	*	*				*	*	*	*	*
	TLAS (TL)			*		*	*	*	*	*				*	*	*	*	*
	DM12 (12)					a	p	p	p					p	p	p	p	
	ZARS (ZA)				*	*	*	*	*	*				*	*	*	*	*
	HA41 (41)					p	p	p						p	p	p	p	
	AU11 (11)			*		*	*	*	*	*				*	*	*	*	*

\* Entrenamiento  
p Predicción

Tabla 4.5 Eventos disponibles en los sitio de suelo utilizados en el módulo TRANS

Clave de la Estación (mapa)	EVENTOS (año/mes/día)									
	93/05/15	93/07/29	94/05/23	94/07/04	94/12/10	95/10/30	93/10/24	94/03/14	95/09/14	95/10/09
CO47 (47)	*			*	*	*	*	*	*	*
DR16 (16)	*	*	*	*	*	*	*	*	*	*
COYS (CY)	*	*	p	*	*	*	*	*	*	*
EO30 (30)			*	*	*	*	*	*	*	*
GR27 (27)	*		p	*	*	*	*	*	*	p
ES57 (57)	*		p	*	*	*	*	*	*	*
AU46 (46)			p	p	p			p	p	p
ME52 (52)	*		*	*	*	*	*	*	*	*
DX37 (37)	*		*	*	*	*	*	*	p	*
AO24 (24)	*		*	*	*	*	*	*	*	*

\* Entrenamiento  
p Predicción

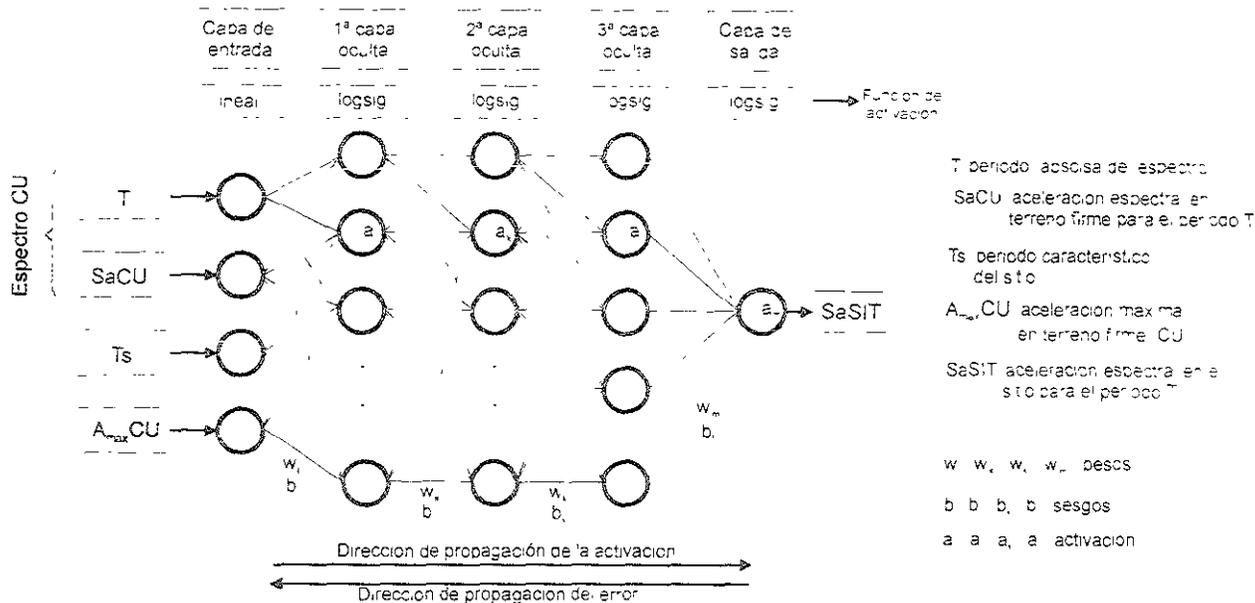


Fig 5.1 RNA para análisis de respuesta sísmica en el valle de México

### 5.1.1 Módulo LAGO

En el capítulo anterior se presentó el desarrollo del modelo neuronal de respuesta sísmica para los sitios localizados en la zona III (módulo LAGO), donde se describe básicamente la base de datos medidos para las etapas de entrenamiento y predicción (prueba). Como se mencionó, este módulo se dividió en cuatro subcasos debido a que el rango de periodos característicos de sitio ( $T_s$ ) es bastante amplio en esta zona (ver fig. 4.5 y Tabla 4.4). Para determinar la topología de red apropiada se utilizó la base de datos del subcaso I ( $T_s$  entre 1.2 y 1.4 s) y en la fig. 5.2a se presenta el análisis para determinar la arquitectura de red óptima, considerando solamente tres variables de entrada ( $T$ ,  $SaCU$  y  $T_s$ ) y empleando funciones de activación sigmoideas (logsig) en todas las capas de la red.

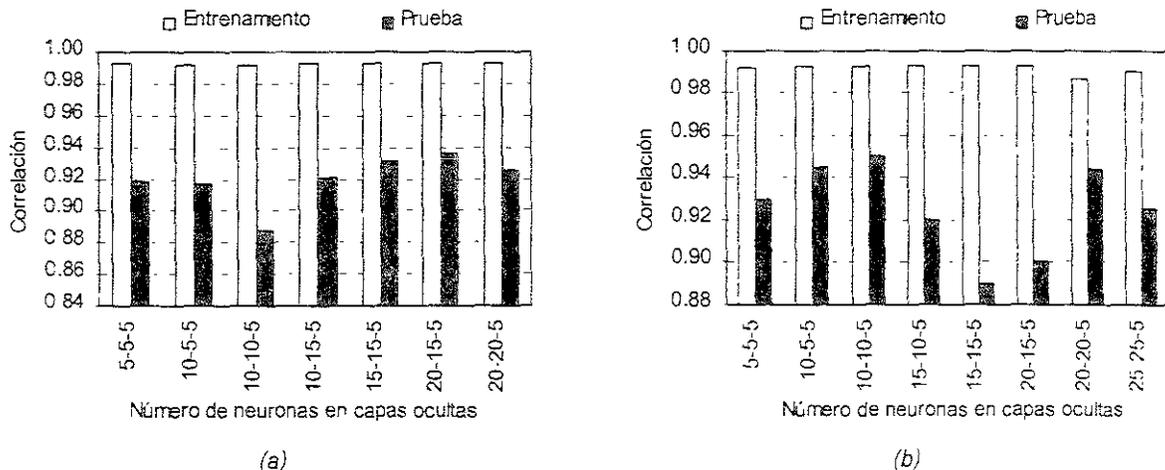


Fig 5.2 Etapas de entrenamiento y prueba del modelo neuronal para el subcaso I del módulo LAGO, (a) considerando tres variables de entrada ( $T$ ,  $SaCU$ ,  $T_s$ ) y (b) considerando cuatro variables de entrada ( $T$ ,  $SaCU$ ,  $T_s$ ,  $A_{maxCU}$ )

En el ejemplo anterior (fig. 5.2a) se obtiene como red óptima una arquitectura de 20-15-5 neuronas en las capas ocultas con una correlación de 93.5% en la etapa de predicción. Sin embargo, e mismo análisis de estabilidad pero ahora considerando la aceleración máxima en terreno firme ( $A_{max}CU$ ) como otra variable de entrada, se obtiene una estructura óptima de 10-10-5 neuronas en las capas ocultas con una correlación de 95.0% en la etapa de prueba, como se muestra en la fig. 5.2b. Lo anterior indica que esta variable  $A_{max}CU$  mejora los resultados en la prueba del modelo neuronal.

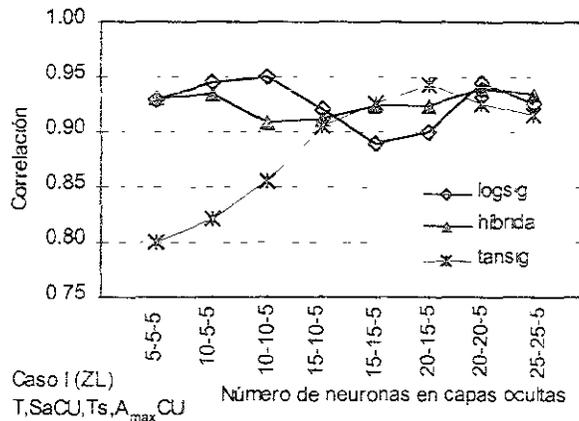


Fig 5.3 Evaluación de la estabilidad de arquitecturas de RNA durante la etapa de prueba

En la fig. 5.3, se presenta el análisis de sensibilidad para el subcaso I, considerando diferentes funciones de activación en el modelo. En el primer análisis se emplean funciones log-sigmoidales en todas las capas de la red, en el segundo se utiliza una combinación de funciones log-sigmoidales y lineales (red híbrda), y en el último se aplican funciones tan-sigmoidales en todas las capas de la red. Se observa que en cualquiera de los análisis se obtienen correlaciones muy similares en la etapa de predicción (generalización de la red) para las estructuras óptimas. Sin embargo, utilizar solamente funciones de activación log-sigmoidales en todas las capas de la red asegura tener valores positivos en la salida del modelo, lo que concuerda con las características de la base de datos disponible (valores no negativos y normalizados entre 0 y 1). Este es la razón de utilizar funciones de activación *logsig* en todas las arquitecturas de RNA para modelar e fenómeno de respuesta sísmica.

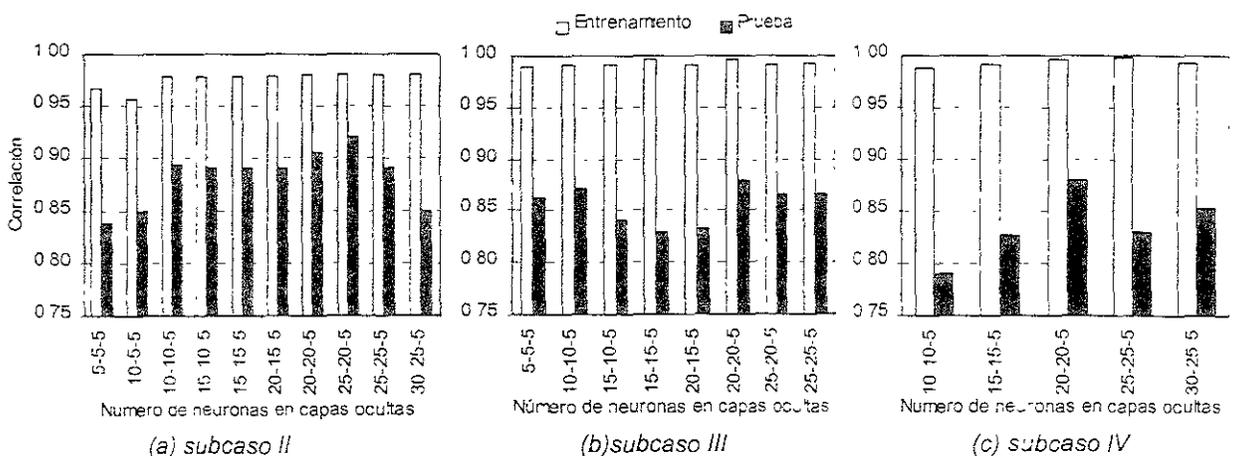


Fig 5.4 Determinación de las estructuras de RNA óptimas para el módulo LAGO durante las etapas de entrenamiento y predicción

Los ejemplos anteriores (figs. 5.2a, b y 5.3) se realizaron para demostrar que la topología de RNA mostrada en la fig. 5.1 es la que proporciona mejores resultados durante la etapa de generalización del modelo. En la fig. 5.4 se muestran los análisis para determinar las arquitecturas de red óptimas para los demás casos que integran el módulo LAGO. Se observa que todas las correlaciones para el grupo de entrenamiento resultaron mayores al 97.0%, mientras que los resultados en la etapa de predicción (prueba) muestran correlaciones bajas para las arquitecturas óptimas (donde se presenta el mínimo error entre la respuesta de la red y el valor medido), principalmente en los subcasos III y IV. Sin embargo, se tienen algunos eventos donde el modelo neuronal predice con bastante aproximación a los datos medidos; y que se muestran en las figs. 5.5 a la 5.8.

La fig. 5.5 presenta algunos de los eventos de predicción para la zona de Lago (subcaso I), en los cuales se observa la aproximación del modelo neuronal en la respuesta sísmica de los depósitos arcillosos con periodos dominantes entre 1.2 y 1.4 s. La correlación del conjunto total de prueba para este caso fue de 95.0% para una arquitectura de red óptima de 10-10-5 neuronas en las capas ocultas (ver fig. 5.2a), lo cual es bastante aceptable para estos casos que el modelo neuronal no había visto durante la etapa de entrenamiento.

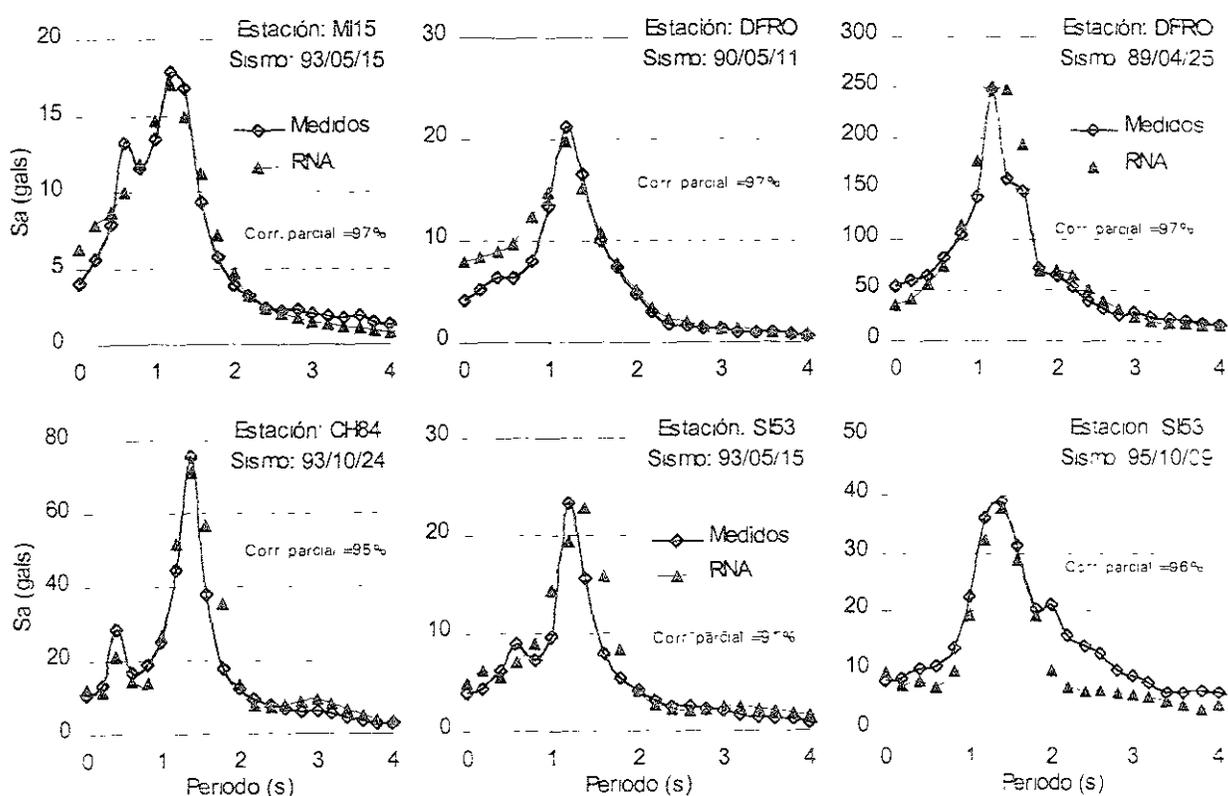


Fig 5.5 Predicción del modelo neuronal para sitios ubicados en la zona de Lago subcaso I

Resultados similares se obtuvieron con la RNA para el subcaso II ( $T_s$  entre 2.0 y 2.6 s), con una arquitectura de red óptima de 25-10-5 neuronas en las capas ocultas y una correlación en prueba de 92.0% (ver fig. 5.4a). Algunos de los eventos de prueba pertenecientes a este subcaso se muestran en la fig. 5.6, donde los valores obtenidos por la técnica de RNA son muy cercanos a los valores reales.

En la misma fig. 5.6, se presentan tres eventos de prueba que corresponden al sitio SCT ( $T_s=2.0$  s), aunque en el evento 99/09/30 no se reproduce exactamente la forma espectral en sus picos, el periodo de la ordenada espectral máxima obtenido con RNA es cercano al periodo dominante del depósito, como lo muestran los otros dos eventos para el mismo sitio (95/10/09 y 95/09/14). Sin embargo, existen eventos sísmicos con formas espectrales complejas (no suaves) que se pueden reproducir con el modelo neuronal como es el caso del evento 95/09/14 para el sitio CJ03 ( $T_s=2.1$  s).

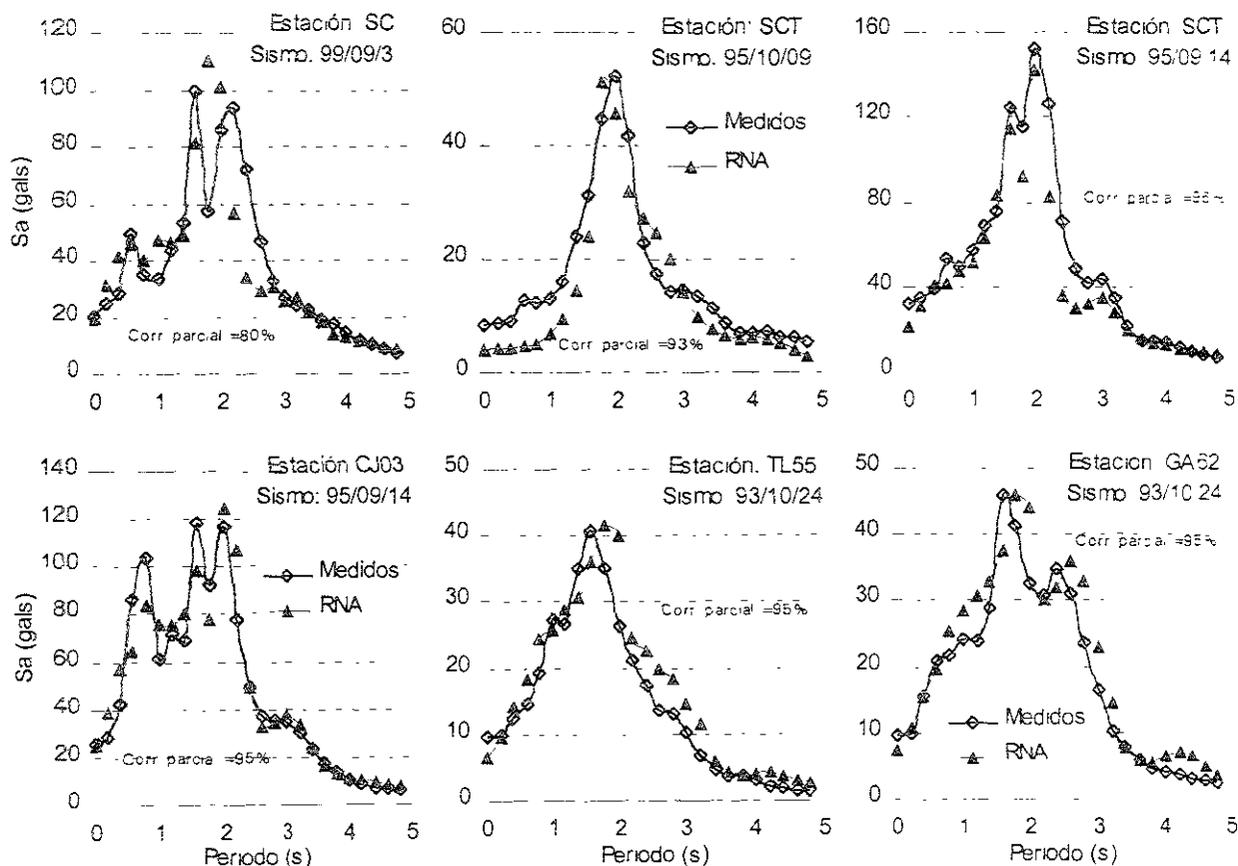


Fig 5.6 Predicción del modelo neuronal para sitios ubicados en la zona de Lago, subcaso II

En las figs 5.7 y 5.8 se presentan algunos de los resultados obtenidos con las RNA óptimas para los subcasos III y IV (20-20-5 neuronas en las capas ocultas en ambos casos). Estos resultados al compararse con los valores medidos muestran que la aproximación de esta técnica es aceptable y es posible reproducir las formas espectrales de la respuesta sísmica para los eventos presentados. Sin embargo, las correlaciones que se obtuvieron en los conjuntos de prueba total fueron bajas (menos de 90.0%) muy similar en ambos casos como lo muestran los análisis de estabilidad (figs. 5.4b y c), lo que indica que la generalización de estas RNA es deficiente. Esto se debe posiblemente a que las características de los eventos sísmicos que integran las bases de datos de estos subcasos (ver fig. 4.5) son bastante complejas, con un contenido amplio de frecuencias y variaciones considerables en sus amplitudes espectrales.

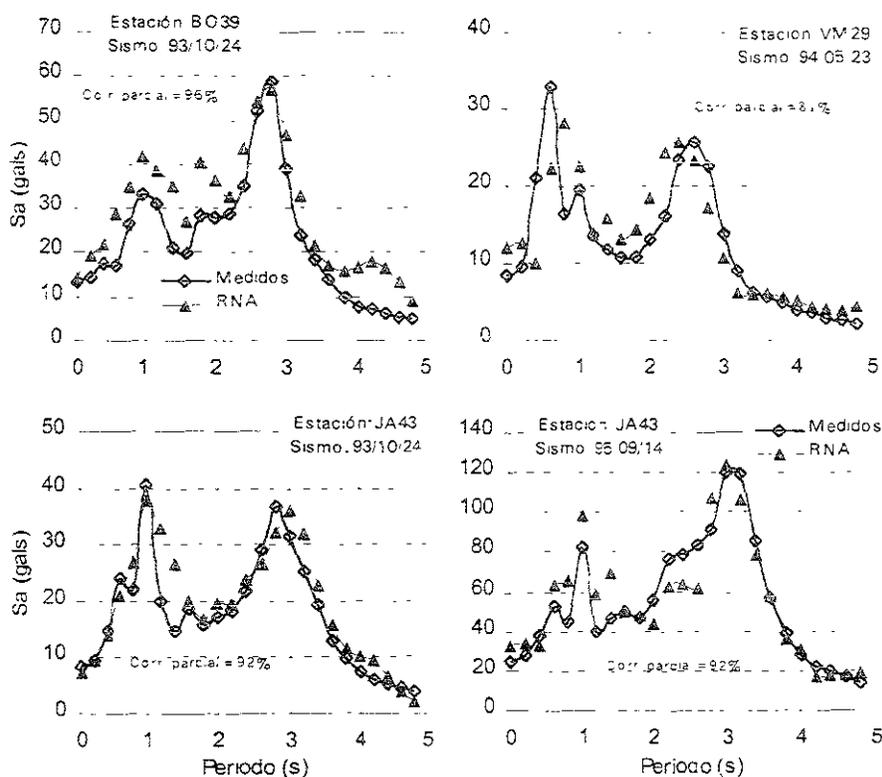


Fig 5.7 Predicción del modelo neuronal para sitios ubicados en la zona de Lago, subcaso III

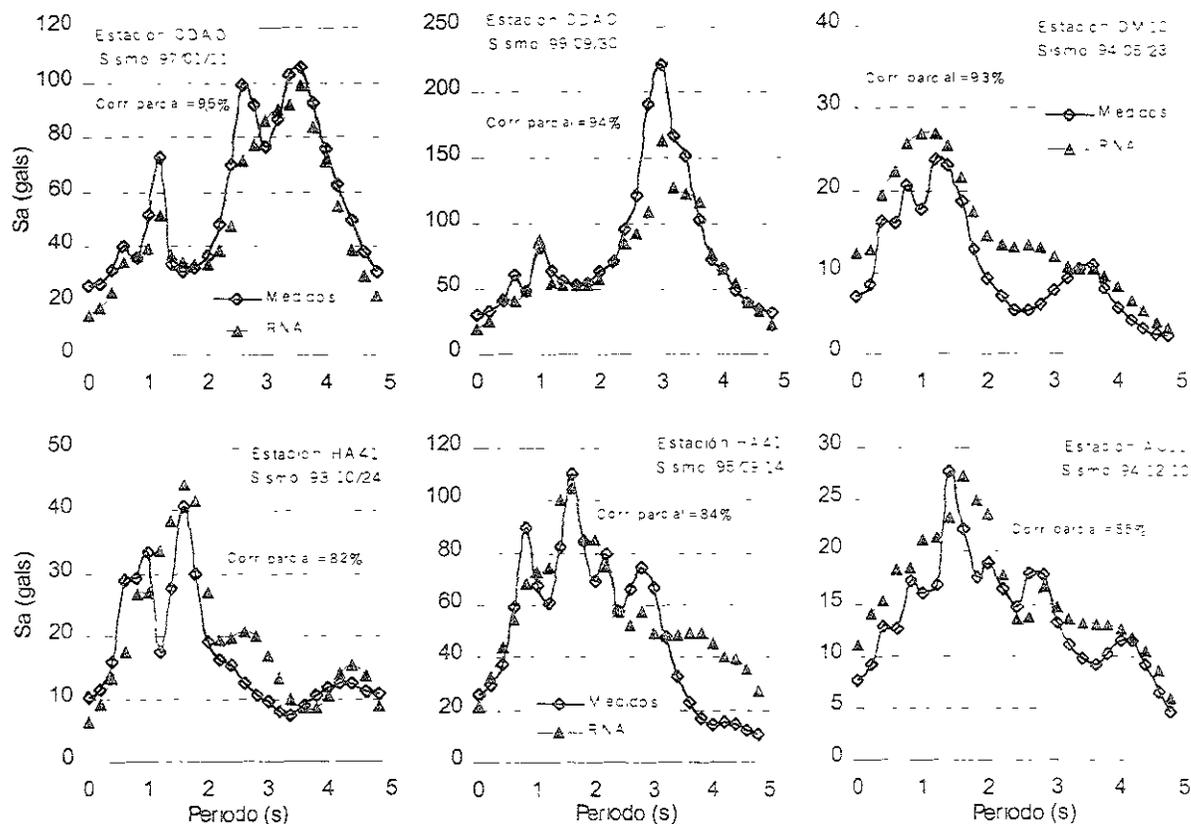


Fig 5.8 Predicción del modelo neuronal para sitios ubicados en la zona de Lago, subcaso IV

### 5.1.2 Módulo TRANS

El desarrollo del modelo de respuesta sísmica con RNA para los depósitos de suelos ubicados en la zona de Transición (módulo **TRANS**) se presentó en el capítulo anterior, donde básicamente se describe la información disponible para los conjuntos de entrenamiento y prueba (sección 4.3.1). De acuerdo con las características de la respuesta sísmica en estos sitios (fig. 4.6 y Tabla 4.5) se utilizó la misma topología de RNA mostrada en la fig. 5.1, considerando las mismas variables de entrada ( $T$ ,  $Sa_{CU}$ ,  $T_s$ ,  $A_{maxCU}$ ) y salida ( $Sa_{SIT}$ ) del modelo.

Las etapas de entrenamiento y prueba realizadas para este caso se presentan en la fig. 5.9. De estos resultados se observa que una combinación de 10-10-5 neuronas en las capas ocultas nos proporciona la mayor correlación en el conjunto de predicción, siendo la arquitectura de red más adecuada para este módulo (mínimo error entre salida de la red y valores medidos). Se muestra también, que para esta red óptima se obtuvo una correlación en el grupo de entrenamiento de 99.0%, mientras que en el grupo de prueba se llegó a una correlación máxima de 95.2%. Estos resultados son bastante aceptables, considerando que el problema de respuesta sísmica de los sitios del valle de México es complejo y las condiciones del suelo que se tienen en esta zona geotécnica (sección 4.2.1) son muy erráticas.

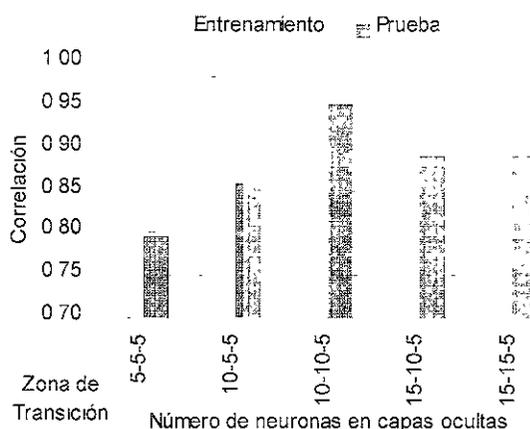


Fig. 5.9 Determinación de la estructura de red óptima para el módulo TRANS (etapas de entrenamiento y predicción)

En la fig. 5.10 se muestran algunos de los casos de predicción con RNA, donde la respuesta del modelo es bastante aproximado a los valores medidos. Los resultados muestran que esta técnica puede reproducir las amplitudes espectrales máximas y sus periodos correspondientes, en estos sitios donde se presentan periodos característicos entre 0.5 y 1.0 s.

La arquitectura de red óptima para este módulo es la misma que la obtenida para el subcaso I del módulo LAGO, esto se debe posiblemente a que los periodos característicos de los sitios considerados para ambos casos son muy cercanos. Además, las formas espectrales de los eventos sísmicos disponibles en la zona de Transición (fig. 4.6), muestran poca dispersión de datos en términos de amplitudes espectrales y contenidos de frecuencias. Esto es otra de las razones de que el modelo con RNA para estos casos presentó mejor comportamiento en sus resultados (mayor correlación en el conjunto de prueba).

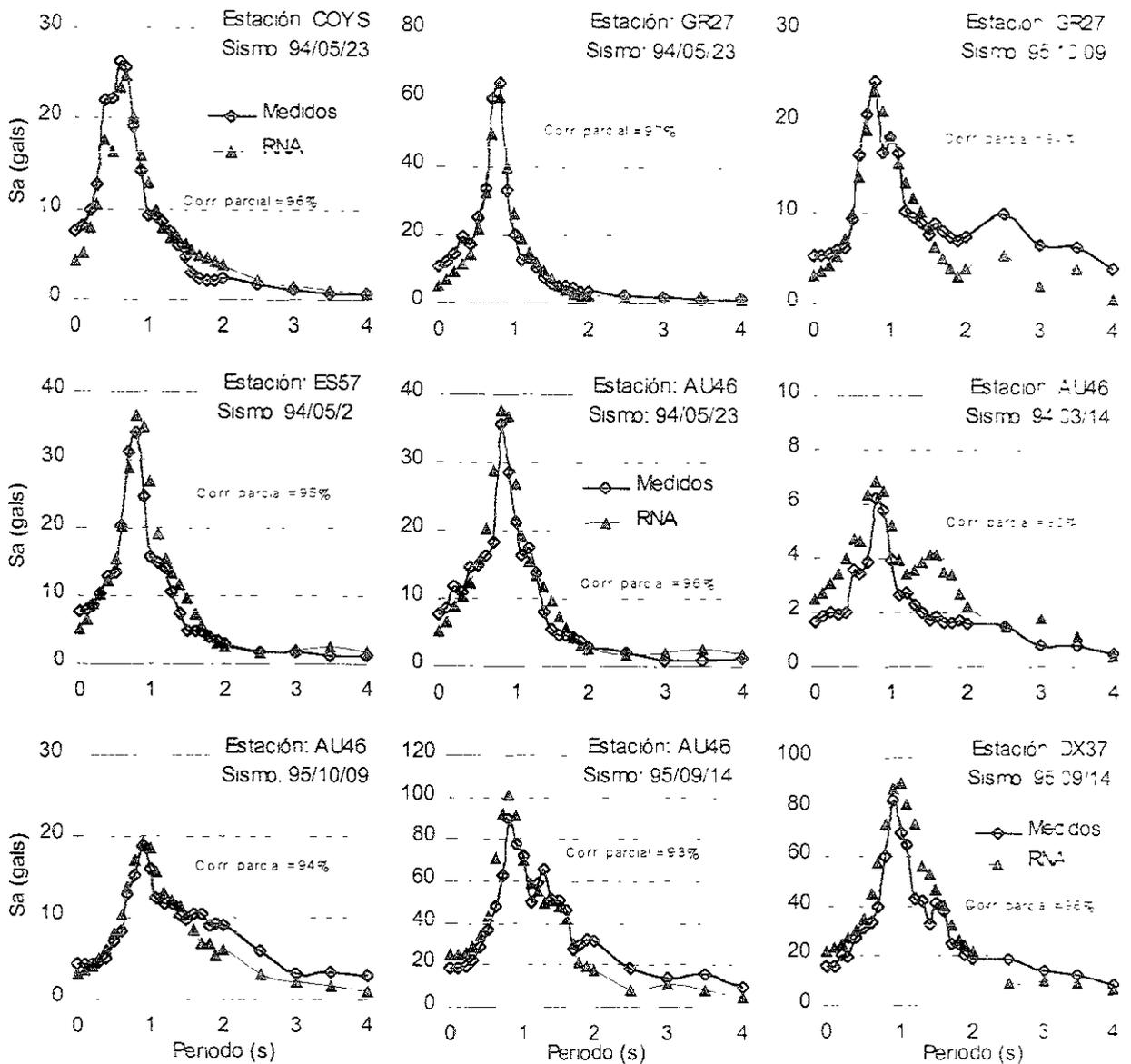


Fig. 5.10 Predicción del modelo neuronal para sitios ubicados en la zona de Transición

## 5.2 ENTRENAMIENTO DEL MODELO DE RNA CON DATOS GENERADOS

Los resultados obtenidos con la herramienta de RNA que se presentaron en la sección anterior, nos muestran la capacidad predictiva del modelo. Sin embargo, en varios casos las predicciones son un tanto deficientes como lo indican las bajas correlaciones de los subcasos III y IV del módulo LAGO en la fig 5.4b y c. Lo anterior se debe posiblemente a que las redes presentadas se basan únicamente en la información seleccionada; es decir, en los sitios y eventos sísmicos que aparecen en las Tablas 4.4 y 4.5. Puesto que las RNA son técnicas de aprendizaje, las hace dependientes de los datos con que se alimentan y la forma como le son presentados.

Los registros sísmicos utilizados para integrar la base de datos de entrenamiento y prueba del modelo neuronal tienen una alta relación señal-ruido, lo que indica que estas señales no incluyen efectos de factores externos que no están relacionados directamente con el problema. De cualquier forma, el fenómeno de respuesta sísmica de los depósitos dentro del valle de México es un problema complejo como se indicó en el capítulo anterior al observar la variación de las formas espectrales de las figs. 4.4, 4.5 y 4.6. A pesar de lo anterior, el modelo neuronal presentado puede reproducir la respuesta sísmica en varios sitios dentro del valle de México de acuerdo con las figs. 5.5 a la 5.8 y 5.10.

Por otra parte, después de obtener las redes óptimas de los casos anteriores, se evaluaron sus capacidades predictivas para depósitos de suelos no considerados durante el entrenamiento de la red. Algunos de estos resultados se presentan en los ejemplos de la fig. 5.11a, donde para casos de sitios con periodos dominantes de 3.05 y 3.15 s (casos interpolados) las respuestas de la red (RNA del subcaso III, zona de Lago) resultaron erróneas

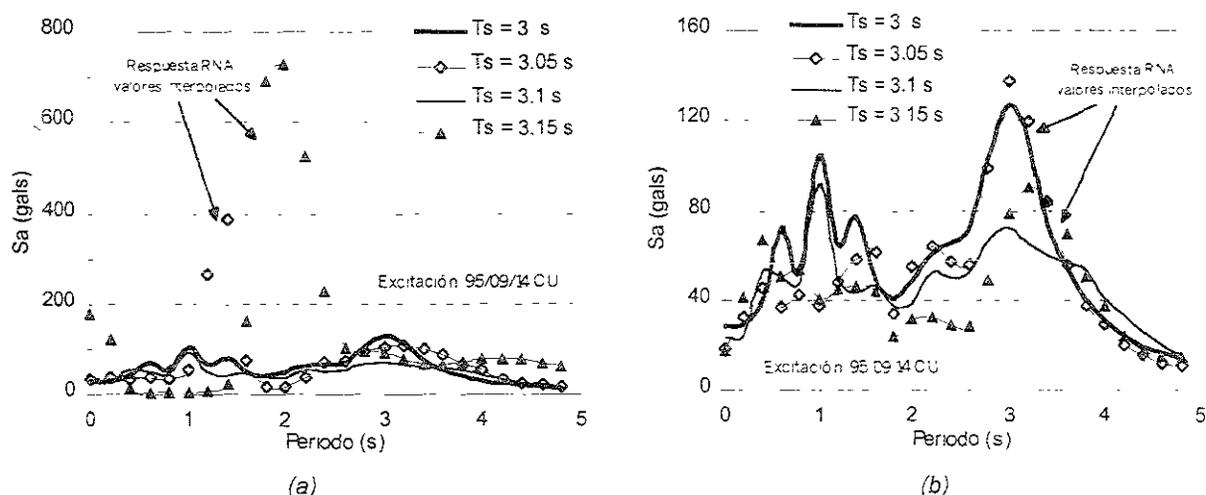


Fig. 5.11 Resultados del modelo sísmico con RNA para sitios no conocidos por la red durante su entrenamiento, (a) datos medidos y (b) datos generados

Lo anterior indica que la información utilizada para formar la base de datos de entrenamiento de las redes presentadas en la sección anterior, es insuficiente. Por lo cual, se realizó otro modelo de respuesta sísmica con RNA en la zona de Lago, idéntico al anterior (ver fig. 5.1), pero ahora utilizando una mayor cantidad de datos en el conjunto de entrenamiento. En este caso se generaron más datos partir de la información seleccionada (ver Tabla 4.5).

Para generar datos representativos de la información registrada en las estaciones acelerométricas ubicadas en la zona de Lago (fig. 4.2), se propone lo siguiente. A partir de los espectros de respuesta en terreno firme correspondientes a los espectros disponibles en un sitio específico ( $S_aSIT$ ), se obtienen los valores máximos y mínimos para cada periodo. Estos límites (envolventes) nos definen una banda de la información medida en terreno firme, como se muestra en la fig. 5.12a. Lo mismo se realiza con las funciones de amplificación empíricas del mismo sitio (fig. 5.12b). A partir de la banda de aceleración espectral en terreno firme ( $S_aCU$ ), se generan más espectros ( $S_i$ ) en forma aleatoria o con un  $\Delta S_a$  constante, de tal forma que dichos espectros  $S_i$  queden dentro de las franjas máximas y mínimas de la información real disponible. A estos espectros  $S_i$  se calculan sus correspondientes espectros de potencia ( $P_i$ ) mediante la teoría del valor extremo. En este caso se utilizó el programa RADSH (Bárcena y Romo, 1993).

De la misma forma, se generan funciones de amplificación ( $fa_i$ ) a partir de la banda de factores de amplificación medidos (SaSIT/SaCU) del sitio considerado. Los espectros de potencia del sitio ( $P_s$ ) se calculan mediante la siguiente expresión:

$$P_s = |fa_i|^2 P_i$$

y finalmente del espectro de potencia  $P_s$  calculado, se obtiene su respectivo espectro de respuesta ( $S_s$ ) con el mismo programa RADSH (fig. 5.12c).

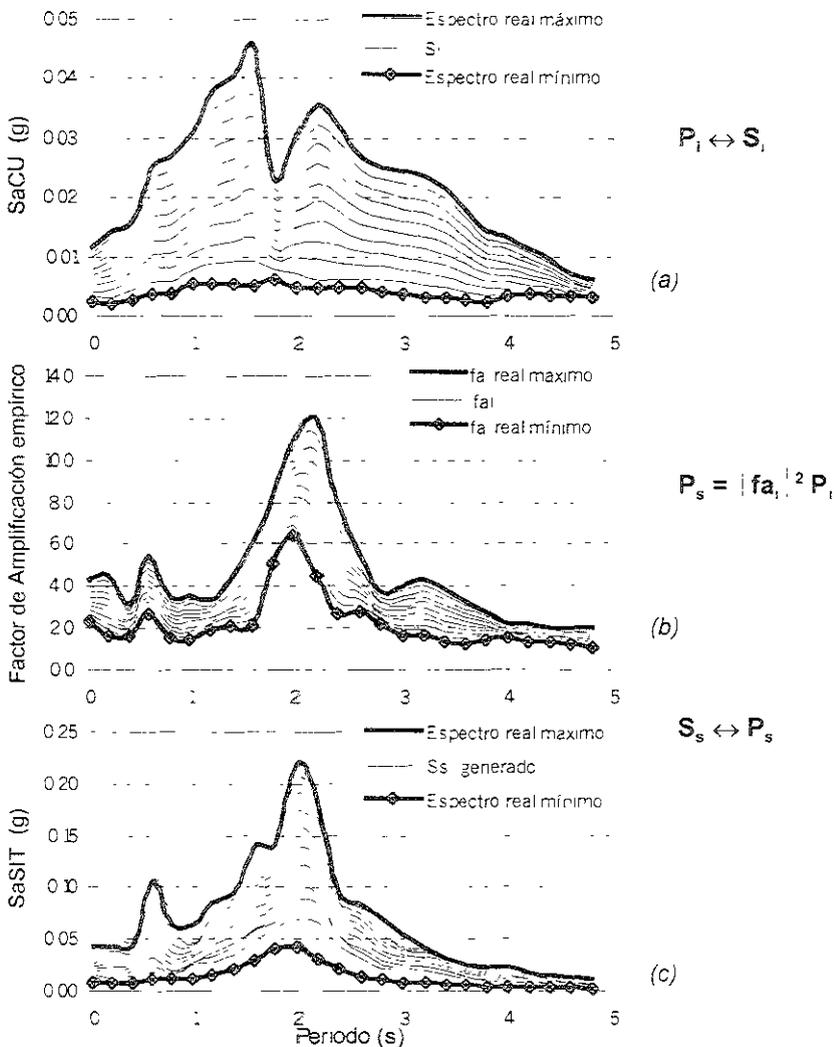


Fig 5.12 Procedimiento de generación de datos a partir de los datos medidos

Con el procedimiento anterior, se generaron alrededor de 15 espectros para cada uno de los sitios en la zona de Lago presentados en la Tabla 4.2, los cuales se utilizaron como base de datos de entrenamiento para la misma topología de RNA mostrada en la fig 5.1, mientras que el grupo de predicción se integró con todos los eventos reales mostrados en la Tabla 4.4. Estas bases de datos de entrenamiento y prueba con espectros generados se dividieron en dos subconjuntos, debido principalmente al amplio rango de periodos característicos en la zona, a la cantidad de datos manejados y, por consiguiente, al tiempo de ejecución durante la etapa de entrenamiento del modelo.

De esta forma, el primer subconjunto del modelo con datos generados se integró con sitios con periodos naturales entre 1.2 s y 2.6 s (caso 1), mientras que el segundo subconjunto se incluyen sitios con periodos característicos entre 2.7 s y 4.2 s (caso 2).

Los análisis de estabilidad de redes (entrenamiento y prueba) de los dos subconjuntos que integran el modelo se presentan en las figs. 5.13a y b. En estos análisis se observa que las arquitecturas de redes más adecuadas, donde se presenta la máxima correlación en el conjunto de prueba resultaron una combinación de 25-15-5 y 25-20-5 neuronas en las capas ocultas, para los casos 1 y 2 respectivamente. En las etapas de entrenamiento de los casos presentados se obtuvieron correlaciones de 100.0% para las arquitecturas de redes óptimas. Sin embargo, las correlaciones en el conjunto de predicción fueron de 93.0% y 91.0%, para los casos 1 y 2 respectivamente.

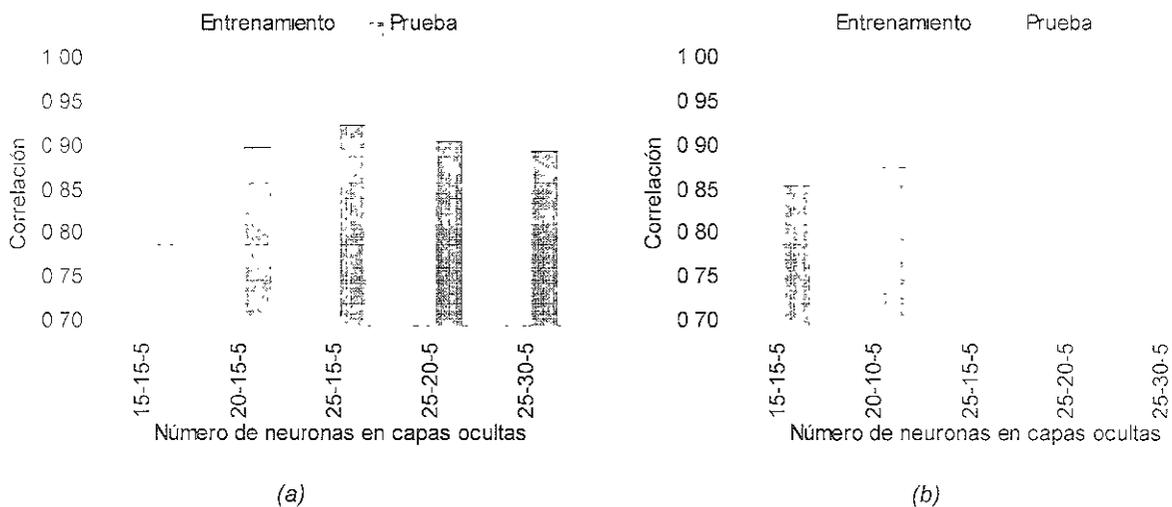


Fig 5.13 Análisis de entrenamiento y prueba de las RNA con datos generados, (a) caso 1 y (b) caso 2

Aunque las correlaciones en el grupo de prueba parecen ser bajas, algunos de los resultados presentados en la fig. 5.14, indican que estas redes (entrenadas con datos generados) pueden calcular con bastante aproximación la respuesta sísmica medida en los sitios ubicados en la zona de Lago. De acuerdo con los resultados mostrados, la técnica con RNA puede predecir los movimientos sísmicos en varios sitios con periodos característicos entre 1.2 y 4.2 s, y para eventos sísmicos de magnitudes entre 6.3 y 7.4.

El anterior ejemplo (datos generados en el grupo de entrenamiento) demostró que la predicción con RNA mejora si se aumenta la cantidad de datos en el grupo de entrenamiento. Aunque las correlaciones para los conjuntos de prueba de las redes entrenadas con datos medidos (figs. 5.2a y 5.4) no se pueden comparar con las presentadas en la fig. 5.13, debido a que los conjuntos de predicción en ambos análisis son diferentes, sí se puede evaluar la capacidad predictiva de la técnica empleada para sitios no vistos por la red en su entrenamiento (casos interpolados). Este análisis se presenta en la fig. 5.11b, donde para periodos dominantes de sitio interpolados ( $T_s$  de 3.05 y 3.15 s), la respuesta obtenida es congruente con los valores medidos cercanos, mientras que la respuesta calculada con el modelo neuronal entrenado con datos medidos es incorrecto.

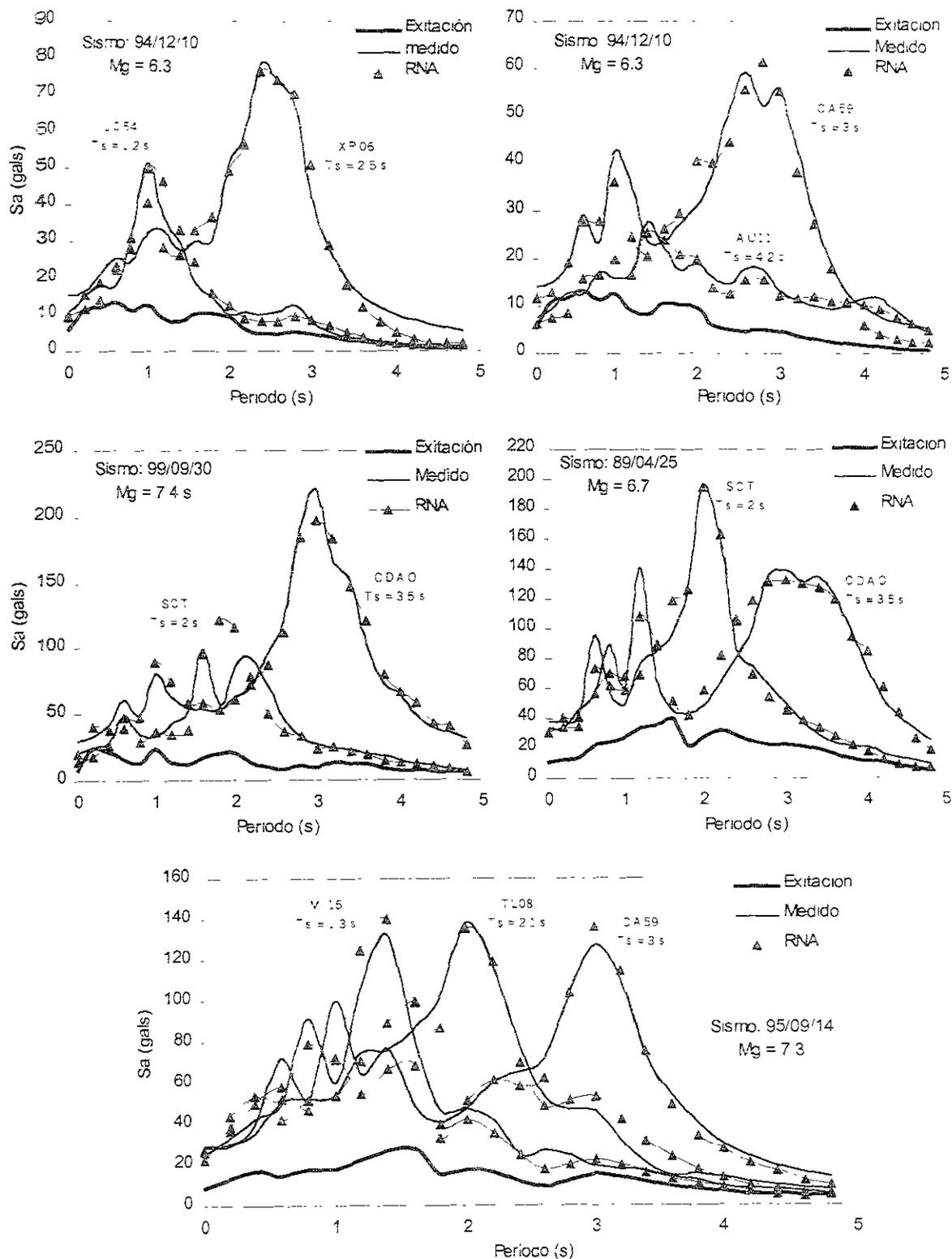


Fig 5.14 Respuesta sísmica medida y predicción con RNA para varios sitios en el valle de México, utilizando datos generados en la etapa de entrenamiento

### 5.3 COMPARACIÓN CON OTROS MODELOS DE RESPUESTA SÍSMICA

Existen una gran variedad de procedimientos para evaluar la respuesta sísmica del terreno que, como se ha mencionado en el capítulo 3, la mayoría de ellos se basan en métodos analíticos, probabilísticos y empíricos. Muchas de estas técnicas han sido verificadas ampliamente a través de comparaciones con mediciones reales, proporcionando en general resultados muy aproximados.

Para las características específicas de los depósitos de suelos que se encuentran en el valle de México, Bárcena y Romo (1994) desarrollaron una metodología simple para determinar la respuesta sísmica del sitio si se conocen los movimientos en terreno firme (excitación) y las características del suelo. Este procedimiento analítico se basa en la propagación unidimensional (1-D) de las ondas de corte en depósitos estratificados horizontalmente, en teorías de vibración aleatoria, en teoría del valor extremo y en matrices de propagación propuestas por Haskell; el cual se resume en un programa de computadora llamado RADSH.

Sin embargo, existen otros estudios (Kawase y Aki, 1989; Chávez-García y Bard, 1994) que sostienen que los análisis 1-D no reproducen satisfactoriamente los movimientos sísmicos en relación a la larga duración observada en el valle de México, principalmente durante el sismo de septiembre de 1985. Cabe mencionar, por otra parte, que una de las posibles razones por la que la duración de la respuesta no se reprodujo con modelos 1-D puede deberse a que los registros obtenidos en CU no contenían la duración total de los movimientos en este sitio, debido a la sensibilidad de los instrumentos existentes. De hecho, investigaciones realizadas por Singh y Ordaz (1993) indicaron que los sismos registrados en terreno firme utilizando sismógrafos de banda ancha tienen una mayor duración que los obtenidos con los acelerógrafos existentes en el valle de México. Considerando esto en los análisis 1-D es posible reproducir la larga duración observada en el sismo de 1985 y eventos posteriores. Pero también, debido a la variación de los movimientos sísmicos que se han observado en varios sitios ubicados en terreno firme no se puede asegurar que los movimientos en el sitio CU sean representativos a los que se presentan en los depósitos profundos.

Una RNA puede sustituir a los análisis teóricos para obtener la respuesta sísmica del sitio en las zona de Lago y de Transición. En la fig. 5.15 se presentan los resultados proporcionados por la RNA, junto con los valores medidos y los resultados con el programa RADSH (análisis lineal). Se observa cómo en la mayoría de los casos el modelo neuronal tiene mejor aproximación a los valores medidos, que el método 1-D de respuesta sísmica. Estos resultados son los obtenidos con las redes entrenadas utilizando la base de datos generados y el criterio presentado en la sección 5.2, donde únicamente el evento sísmico de septiembre de 1985 (85/09/19) se utilizó dentro del grupo de entrenamiento, mientras que los demás casos son de predicción. En los análisis con el programa RADSH se utilizaron la estratigrafía y las propiedades de los suelos del sitio conocidas en estudios anteriores (Bárcena y Romo, 1994).

Estos resultados (fig. 5.15) muestran que la RNA puede estimar la respuesta sísmica sin utilizar completamente las características estratigráficas de los depósitos de suelo y sin plantear tantas hipótesis en los análisis. Este es uno de los aspectos importantes en la aplicación de la herramienta de RNA, donde no se establece ningún tipo de hipótesis, puesto que el modelo solamente se basa en los datos medidos representativos del problema. Sin embargo, lo anterior no quiere decir que el especialista que aplica esta técnica no tenga un conocimiento suficiente del fenómeno físico. Como se ha mencionado en secciones anteriores, la selección de los parámetros representativos del problema a partir de la información disponible y la presentación de ellos a la red, es uno de los aspectos importantes de la aplicación de la técnica de RNA, lo cual requiere de una comprensión amplia del fenómeno a modelar.

Además, estos resultados (fig. 5.15) indican que la técnica con RNA puede predecir con aproximación aceptable la respuesta superficial del terreno ante eventos que provienen de distintas fuentes sísmicas (ver Tabla 4.1). El ejemplo más claro es el sismo 99/06/15 (Tehuacán), que en el sitio SCT ( $T_s=2.0$  s) los resultados con RNA presentan mejor aproximación a lo medido, que los análisis 1-D. En este caso se observa que el modelo neuronal define mejor el periodo fundamental del sitio. Resultados similares se tienen para el evento sísmico 99/09/30 donde el periodo natural con RNA se acerca más al valor medido que con el programa RADSH.

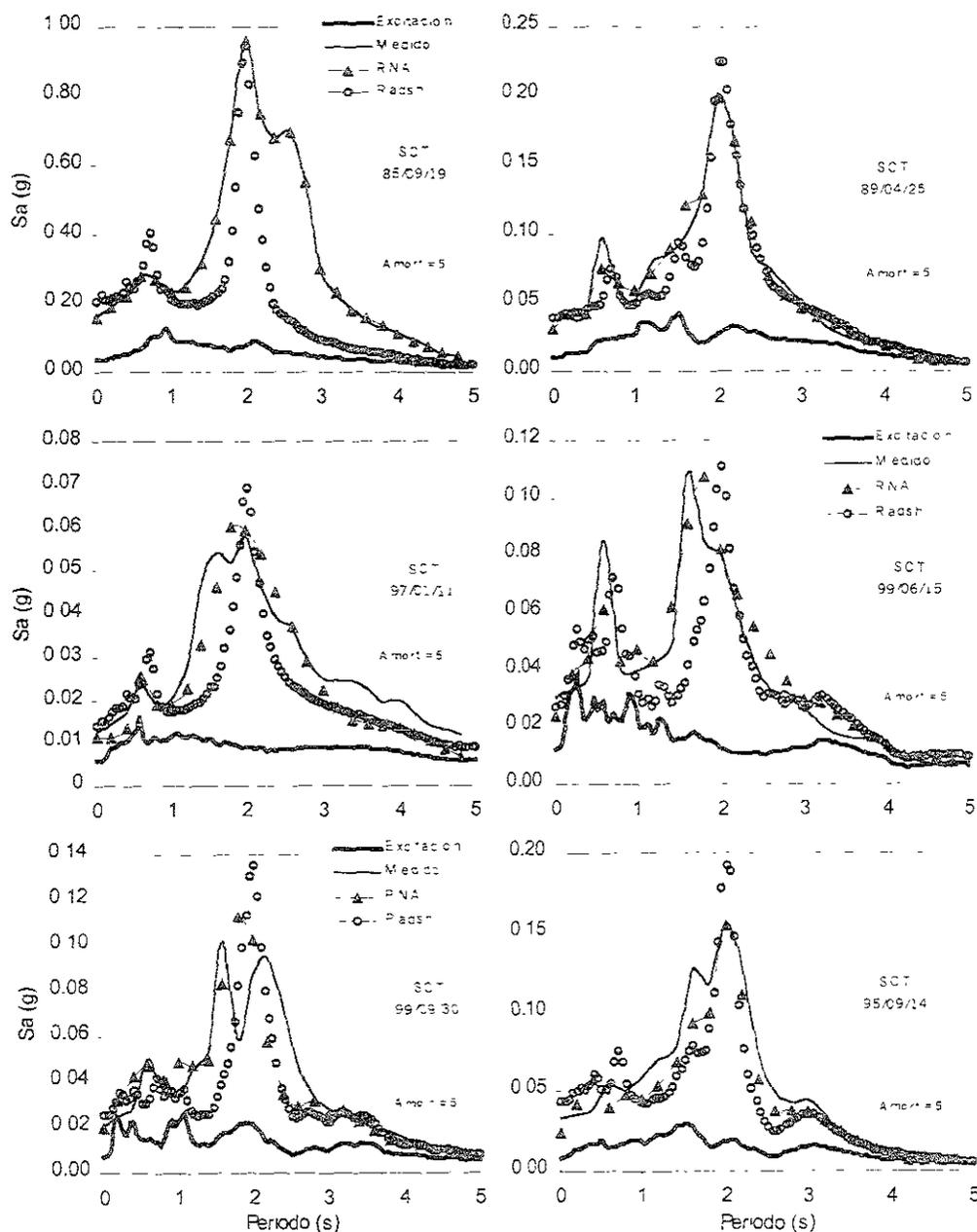


Fig. 5.15 Comparación de resultados del modelo neuronal con análisis teóricos de respuesta sísmica

Por otra parte, para evaluar el modelo neuronal con respecto al de cocientes espectrales se realizaron algunas comparaciones de esta técnica con el programa Z (Ordaz et al., 1998). Estos resultados se presentan en la fig. 5.16, donde se observa una aproximación adecuada de los dos métodos para estos

casos. Sin embargo, con RNA los detalles de las formas espectrales se acercan más a lo medido, que con el programa Z.

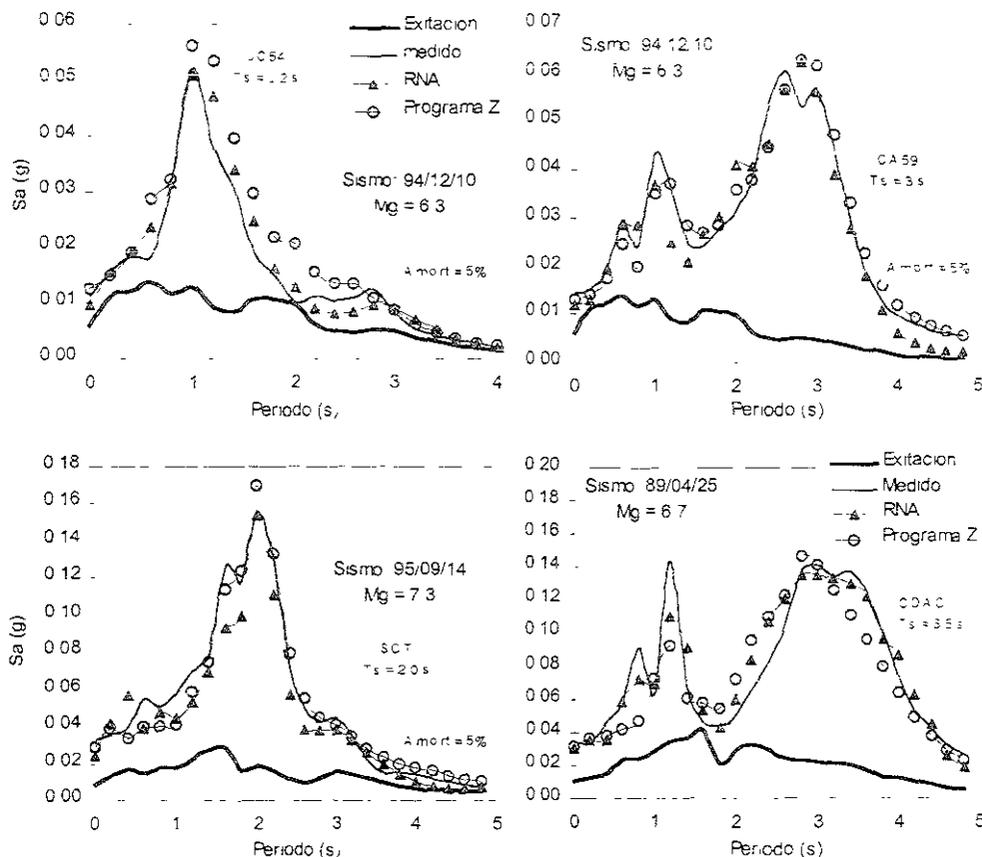


Fig 5.16 Comparación de resultados de RNA y del programa Z (Ordaz et al. 1998) con lo medido

También, en la fig. 5.17 se presentan dos casos (sitios TLAS y RMCS) donde se comparan los resultados de RNA y del programa Z con los medidos, para el evento sísmico de Tehuacán (99/06/15). Se observa que con el programa Z se sobreestiman los valores medidos para ambos casos, mientras que con RNA se obtiene mejor aproximación. Estos resultados indican que la RNA aprende de los datos presentados y por lo tanto, puede estimar mejor la respuesta sísmica que los modelos semiempíricos cuando el sismo proviene de otra fuente sísmica.

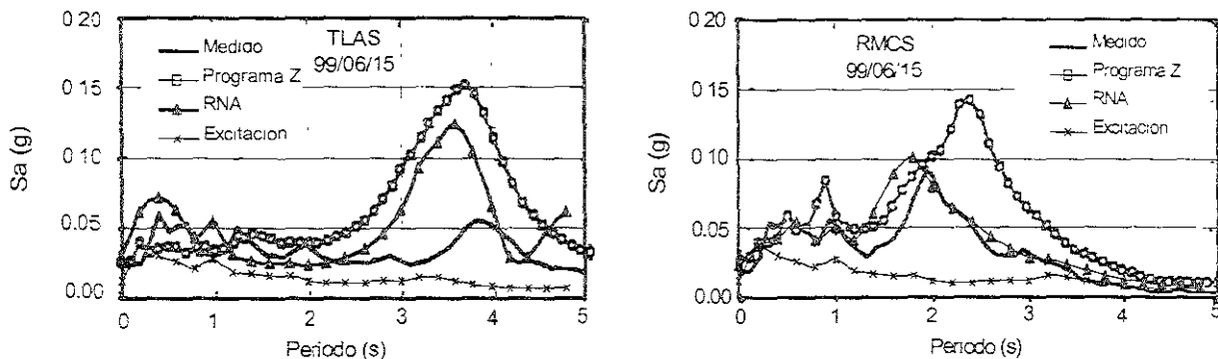


Fig 5.17 Resultados de RNA y del programa Z, para el evento sísmico de Tehuacán (99/06/15 EW)

---

## CAPÍTULO 6

### CONCLUSIONES

---

A partir de los resultados obtenidos con el modelo de respuesta sísmica aplicando la herramienta de Redes Neuronales Artificiales (RNA), se extraen las siguientes conclusiones:

- La técnica de RNA puede estimar la respuesta sísmica en términos de aceleraciones espectrales de los depósitos de suelos dentro del valle de México, si se conocen el espectro de respuesta en terreno firme (Ciudad Universitaria) y el periodo natural del sitio (en las zonas de Lago y Transición). La respuesta de las RNA se ha comparado con los valores medidos, donde la aproximación obtenida es aceptable para los casos presentados.
- También resulta interesante comparar esta técnica con otros modelos. En este trabajo se compararon los resultados de RNA con un método 1-D, debido a que desde el punto de vista de la práctica en Ingeniería, estos procedimientos son útiles para estimar de manera aproximada los movimientos sísmicos del suelo. De la misma forma, se realizaron algunas comparaciones de RNA con el programa Z (Ordaz, et al., 1998), donde se obtienen con los dos métodos aproximaciones aceptables para los casos presentados. Sin embargo, con RNA los detalles de las formas espectrales se acercan más a lo medido, que con el programa Z.
- Con RNA no se pierde el sentido físico del problema, debido a que se incluyen las variables que están directamente relacionadas con el fenómeno de respuesta sísmica en el valle de México. Las comparaciones del modelo neuronal con los valores medidos confirman lo anterior.
- Se puede predecir aproximadamente la respuesta superficial del terreno ante eventos que provienen de distintas fuentes sísmicas, como los utilizados para el entrenamiento del modelo neuronal. El caso más claro es el evento sísmico de Tehuacán (99/06,15), donde los resultados con redes se aproximan mejor a los medidos, que los obtenidos con un modelo 1-D y con el programa Z. Esto indica que el modelo neuronal es más útil en la práctica para estimar los movimientos del terreno en términos de aceleraciones espectrales.

- Las RNA tienen el atractivo de que no requieren suponer hipótesis particulares (tipo de ondas incidentes, características geométricas de la cuenca, características estratigráficas del sitio, comportamiento no lineal, condiciones de frontera, etc.), en comparación con los análisis de respuesta teóricos conocidos. Los movimientos sísmicos utilizados para el desarrollo de modelo neuronal (información medida) necesariamente incluyen los efectos causados por los factores que influyen en la respuesta sísmica del sitio.
- Para aplicar esta técnica es necesario utilizar las variables que están relacionadas directamente con el fenómeno de respuesta sísmica. En este trabajo se utilizaron como variables de alimentación del modelo neuronal las presentadas en la fig. 5.1. Desde luego se puede explorar el uso de otros arreglos de parámetros de entrada al modelo, pero es importante realizar una preselección de variables representativas, en base al conocimiento que se tenga del fenómeno natural o aplicando otras alternativas de selección como, por ejemplo, algoritmos genéticos.
- Se seleccionó el sitio Ciudad Universitaria como punto de referencia para alimentar el modelo neuronal, debido a que en tal sitio se han registrado los principales eventos sísmicos incidentes en el valle de México y se ha utilizado como base en varios estudios previos. De cualquier forma, la entrada del modelo neuronal es relativa, es decir, así como se utiliza CU como punto de referencia se puede utilizar otro sitio, como por ejemplo CDAO y de ahí calcular la respuesta en los demás sitios de la zona de Lago, en la de Transición e incluso en terreno firme.
- Aumentar la base de datos para el entrenamiento de las redes con información generada a partir de la observada, puede mejorar la predicción del modelo para sitios no considerados (casos interpolados) durante su entrenamiento. Lo cual demuestra que una mayor cantidad de información medida puede mejorar los resultados con RNA.
- La división del problema en varios subcasos, ofrece un menor costo computacional durante el entrenamiento de las RNA, lo que conduce al desarrollo de redes en módulos como una opción para la generalización del método.
- La confiabilidad de estas técnicas basadas en el aprendizaje, reside principalmente en la calidad y cantidad de los datos medidos (o experimentales) con los que se alimenta el modelo y de la selección apropiada de la topología de RNA.

Las limitaciones del modelo de respuesta sísmica presentado en este trabajo y los posibles avances para mejorar el método se describen a continuación:

- Los modelos mostrados aquí son válidos dentro del rango de valores de los parámetros (características de los movimientos sísmicos y de los sitios) que se consideran en la base de datos de entrenamiento de las redes. Aunque, esta limitación puede resolverse si en la base de datos de entrenamiento se incluye información más completa y reciente. Esto nos indica que la actualización de las RNA ayudará a la predicción de la respuesta sísmica.
- Aunque estos modelos neuronales proporcionan resultados muy aproximados a los valores medidos, existen algunos aspectos que todavía es necesario mejorar como es la generalización del método (etapa de predicción). La existencia de otras topologías de RNA (en cascada, de regresión generalizada, recurrentes, etc.) y de nuevos algoritmos de entrenamiento pueden mejorar los resultados presentados aquí (Romo, 1999, García, et al., 2000).

- Como se ha mencionado, la eficiencia de la técnica de RNA depende de la calidad y cantidad de los datos utilizados para el aprendizaje; sin embargo, en la mayoría de los fenómenos naturales existe una gran cantidad de parámetros que están directamente relacionados con el fenómeno, pero también se presentan variables que no lo están (perturbaciones internas). Por lo cual, la aplicación de técnicas de preprocesamiento de datos para analizar la información disponible antes de aplicar la herramienta de RNA puede ayudar al entrenamiento y predicción del modelo. Un ejemplo de lo anterior, son los Análisis de Cuantificación Recurrente (AQR) para series de tiempo, los cuales pueden gráficamente detectar patrones ocultos, cambios estructurales en datos o ver similitudes a través de la información en estudio. Esto ayudaría al modelo con RNA a identificar nuevos patrones de comportamiento en sitios con periodos característicos iguales, pero diferente respuesta sísmica ante un mismo movimiento de excitación.

De la experiencia obtenida durante este trabajo se concluye que para la aplicación de la técnica de RNA a cualquier problema de ingeniería sísmica geotécnica, se requiere principalmente del entendimiento del fenómeno físico a modelar, de la experiencia para la selección de la topología de redes más adecuadas para resolver el problema específico, y de estrategias para el manejo, selección de variables y presentación de la información seleccionada a las redes.

Del análisis de la base de datos medidos de respuesta sísmica de los depósitos de suelo del valle de México, utilizada en el aprendizaje de las RNA se puede decir lo siguiente:

- Las formas espectrales de los movimientos sísmicos en los depósitos de la zona de Lago en la ciudad de México (figs. 3.13 y 4.5) presentan ordenadas espectrales máximas entre 6 y 7 veces la aceleración máxima del terreno, y rangos de periodos fundamentales entre 1.2 y 4.2 s. Las zonas de aceleraciones espectrales máximas comprenden periodos entre 1.4 y 2.6 s. En la zona de Transición, los depósitos de suelo presentan aceleraciones espectrales no mayores a 6.0 veces la aceleración máxima del terreno, y periodos característicos entre 0.5 y 1.0 s (fig. 4.6).
- Los espectros normalizados para periodos característicos entre 3.0 y 4.0 s, muestran que las ordenadas espectrales en los segundos periodos naturales del sitio son considerables (entre 4 a 5 veces la aceleración máxima del terreno)
- La información presentada comprueba que una zonificación más detallada para la zona de Lago, permitiría utilizar espectros de respuesta más apropiados para el diseño sísmico de estructuras cimentadas en dicha zona.

Por otra parte, uno de los objetivos particulares de este trabajo es mostrar la existencia de otros factores que influyen en la respuesta sísmica y que su consideración puede mejorar la modelación del problema. Los resultados presentados sobre el efecto de la consolidación regional en la respuesta sísmica en los depósitos de suelo arcilloso en la ciudad de México, permiten sugerir que los periodos característicos ( $T_s$ ) de los depósitos de suelos blandos disminuyen con el tiempo, mientras que las amplitudes espectrales máximas pueden disminuirse o incrementarse, dependiendo del periodo natural del depósito de suelo y las características frecuenciales de la excitación. De esta forma, la variación temporal de las propiedades dinámicas debido al fenómeno de consolidación se puede tomar en cuenta de manera explícita en una RNA como la desarrollada en este trabajo, modificando el periodo fundamental del sitio

---

## BIBLIOGRAFÍA Y REFERENCIAS

---

- Aki, K., (1988), Local Site Effects on Strong Ground Motion, Earthquake Engineering and Soil Dynamics II. Recent Advances in Ground-Motion Evaluation. Geotechnical Special. Publication No 20 June, pp. 103-155.
- Aki, K., (1993), Local site effects on weak and strong ground motion, Tectonophysics, Vol. 218, No. 1-3 pp. 93-112.
- Alcántara, et al., (1997), La base mexicana de datos de sismos fuertes en disco compacto. Revista Ingeniería Civil, CICM, No. 338, pp. 13-22
- Anderson, J. A., (1972), A Simple Neural Network Generating an Interactive Memory Mathematical Biosciences, Vol. 14
- Athanasopoulos G.A.; Pelekis, P.C.; Leonidou, E.A. (1999), Effects of surface topography on seismic ground response in the Egion (Greece) 15 June 1995 earthquake, Soil Dynamics and Earthquake Engineering. Vol 18, No. 1, ISSN 0267-7261, pp. 135-149.
- Auvinet, G, Méndez, E. y Juárez, M., (2000), Archivos personales.
- Avilés, J. y Perez-Rocha, L. E., (1998), Site effect and soil-structure interaction in the Valley of Mexico, Soil Dynamics and Earthquake Engineering, Vol. 17, No. 1, ISSN 0267-7261, January. pp. 29-39
- Bárcena, A. y Romo, M. P., (1993), Efectos no lineales en las respuestas medidas en los depósitos arcillosos de la ciudad de México, Memoria del X Congreso Nacional de Ingeniería Sísmica, Puerto Vallarta, Jal. Méx pp.263-269.
- Bárcena, A. y Romo, M. P., (1994), RADSH: Programa de computadora para analizar depósitos de suelo estratificados horizontalmente sujetos a excitaciones dinámicas aleatorias. Reporte interno Instituto de Ingeniería, UNAM.
- Barnard, E., (1992). Optimization for Training Neural Nets. IEEE, Transactions on Neural Network. Vol. 3, No 2
- Bard, P.Y. y Ganel, J.C., (1986), The seismic response of two-dimensional sedimentary deposits with large vertical velocity gradients, Bulletin of the Seismological Society of America. Vol 76, pp. 343-356
- Battiti, R., (1992). First and Second Order Methods for Learning: Between Steepest Descent and Newton's Method. Neural Computation, Vol. 4, No 2, pp. 141 – 166
- Bocheva, H. (1977), Soil Amplification Factor of Surface Waves. Proc. 6th World Conf. Earthquake Engrg., Vol 6, pp. 176-189. New Delhi.
- Bouchon, M. y J.S. Barker (1986), Seismic response of a hill: the example of Tarzana, California, Bull Seism Soc Am., 86, 66-72.
- Castellani A., Peano A, Sardella L, (1982), Seismic response of topographic irregularities, In: Proceedings of the Third international earthquake microzonation conference, Vol. II, pp. 533-40
- Chang, S. W. (1996), Seismic response of deep stiff soil deposits, Dissertation submitted in partial satisfaction of requirements for the degree of doctor of Philosophy in civil Engineering of the University of California at Berkeley.
- Chang, C.-Y., Power, M. S., Idriss, I. M., Somerville, P. G., Silva, W. y Chen P. C. (1986), Engineering characterization of ground motion, Task II: observational data on spatial variations of earthquake ground motions, Report NUREG/CR-3805, Vol. 3. U. S. Nuclear Regulatory Commission, Washington, D C
- Charalambous, C., (1992). Conjugate Gradient Algorithm for Efficient Training of Artificial Neural Networks. IEEE, Proceedings, Vol. 139, No 3.
- Chávez-García, F.J. y Bard, P.Y., (1994), Site effects in Mexico City eight years after the September Michoacan earthquake, Soil Dyn Earthq Engrg; Vol. 13, pp. 229-47.
- Chávez-García, F.J., Ramos-Martínez, J., y Romero-Jiménez, E., (1995). Surface Wave Dispersion Analysis in Mexico City, Bulletin of the Seismological Society of America Vol. 85, pp. 1116-1126

- Chávez-García, F.J., y Romero-Jiménez, E., (1996), Análisis de los acelerogramas registrados en la ciudad de México durante el temblor del 25-04-89. Evidencias de una posible explicación para la gran duración del movimiento fuerte, Revista de Ingeniería Sísmica, No. 54, pp 1-15.
- Chávez-García, F.J., Rodríguez, M., Field, E. y Hatzfeld, D., (1997), Topographic site effects. A comparison of two nonreference methods, Bulletin of the Seismological Society of America, Vol. 87, pp. 1667-1673.
- Chen, J. Ch., Lysmer, J. y Seed, H. B. (1981), Analysis of local variations in free field seismic ground motion EERC Report No. UCB/EERC-81/03, January
- Colindres, R., (1983), Dinámica de Suelos y Estructuras : Aplicadas a la Ingeniería Sísmica, edit. Limusa México.
- Deng, N. (1991), Two-Dimensional site Response Analyses. Thesis presented to University of California at Berkeley, Berkeley, California in partial satisfaction of the requirements for the degree of Doctor of Philosophy
- Deng N. (1993), FLUSHPLUS: A computer program for approximate 3-D analysis of soil-structure interaction problems. An enhanced PC version of the original FLUSH program published in 1975 by Lysmer, Udaya Tasa, Seed., (program conversion and modification) University of California, Berkeley
- Dickenson, S. E. (1994), Dynamic response of soft and deep cohesive soils during the Loma Prieta earthquake of October 17, 1989, Dissertation submitted in partial satisfaction of requirements for the degree of doctor of Philosophy in civil Engineering of the University of California at Berkeley
- Durukal, E., Erdik, M., Avci, J., Yuzugullu, O., et al, (1998), Analysis of the strong motion data of the 1995 Dinar, Turkey earthquake, Soil Dynamics and Earthquake Engineering, Vol. 17, No. 7-8, ISSN 0267-7261, October/December, pp 557-558.
- Duval, A.M., Meneround, J.P., Vidal, S., Bard, P.Y., (1998), Relation between curves obtained from microtremor and site effects observed after Caracas 1967 earthquake. 11<sup>th</sup> European Conference on Earthquake Engineering, Paris, France.
- 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, pp. 1779-1797.
- Faccioli, E., Maggio, F., Paolucci, R., Quarteroni, A., (1997), 2D y 3D elastic wave propagation by a pseudo-spectral domain decomposition method, J. Seis., 1:pp. 237-251.
- Fahlman, S. E., (1988), 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. Ed. San Mateo CA: Morgan Kaufmann, pp. 38 - 51.
- Freeman, J. y Skapura, D., (1991), Neural Networks: Algorithms, Applications, and Programming Techniques. Reading, MA: Addison-Wesley. (Chapter 14).
- Fucks, I.M. y Voronovich, A.G., (2000), Wave diffraction by rough interfaces in an arbitrary plane-layered medium, Waves in Random Media, A Journal Recognized by the European Physical Society, Vol. 10, No. 2, April, pp. 253-272.
- García, S. R., Romo, M. P. y Sarmiento, N., (2000), Analysis of ground motions in Mexico City with Artificial Neural Networks, Simposio Internacional de Riesgos Geológicos y ambientales en la Ciudad de México. Octubre.
- Geil, L., Bard, P. Y., y Jullien, B., (1988), The effect of topography on earthquake ground motion: A review and new results, Bulletin of the Seismological Society of America, Vol. 78, No. 1, pp. 42-63
- Graves, R. W., (1993), Modeling three-dimensional site response effects in Marina District Basin, San Francisco California, Bulletin of the Seismological Society of America. Vol. 83, 1042-1063
- Grossberg, S., (1976), Adaptive Pattern Classification and Universal Recording: I. Parallel Development and coding of Neural Feature detectors. Biological Cybernetics, Vol. 23.
- Guéguen P., Jean-Luc Chatelain, Bertrand Guillier, Hugo Yepes y Jose Egred, (1998), Site effect and damage distribution en Pujili (Ecuador) after the 28 March 1996 earthquake, Soil Dynamics and Earthquake Engineering, Vol. 17, No. 5, ISSN 0267-7261. Julio, pp. 329-334.
- Gutenberg, B., (1957), Effects of ground on earthquake motion Bulletin of the Seismological Society of America Vol. 217, pp. 221-250.
- Hagan, M., Demuth, H., and Beale, M., (1996). Neural Network Design. Ed. PWS Compañía, Boston
- Hagan, M., Menhaj, M., (1994). Training Feedforward Networks with the Marquardt Algorithm. IEEE, Transactions on Neural Network, Vol. 5, No 6.
- Hannon, W. J., (1964), An applications of the Haskell-Thomson matrix method to the synthesis of the surface motion due to dilatational waves, Bull. Seis. soc. Am., vol. 54, No. 6, pp. 2067-2079, December.

- Herrera, Y. y Rosenblueth, E., (1965), Response Spectra on Stratified Soil, Proceedings, III World Conference on Earthquake Engineering, New Zealand, vol. I, pp. 44-60.
- Hestenes, M. R. y Stiefel, E. (1952) Methods of conjugate gradients for solving linear systems. Journal of Research of the National Bureau of Standards, 49, pp. 409-436.
- Hirose, Y., Yamashita, K. y Hijiya, S. (1991) Back-propagation algorithm which varies the number of hidden units, Neural Networks; 4, pp. 61-66.
- Hopfield, J. J., (1982). Neural Networks and Physical Systems with Emergent Collective Computational Properties. Proceedings of the National Academy of Sciences, Vol. 79, pp. 2554-2558.
- Hornik, K. M., Stinchcombe, M., and White, H., (1989). Multilayer Feedforward Networks are Universal Approximators. Neural Networks, Vol. 2, No 5.
- Huang, H. C. y Chiu, H. C. (1999), Canyon topography effects on ground motion at Feitsui damsite, Soil Dynamics and Earthquake Engineering, Vol. 18, No. 2, ISSN 0267-7261, August, pp. 87-89.
- Idriss, I. M. y Seed, H. B. (1968), Seismic response of horizontal soil layers, JSMFD, ASCE, Vol. 94, No. SM4, pp 100-1031, July.
- 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, pp. 273-289
- Jacob, K. H., (1991), Seismic zonation and site response: Are building-code soil-factors adequate to account for variability of site conditions across the US?. Proceeding, 4th International Conference on Microzonation, EERI, Stanford University, Palo Alto, California, Vol. 1 pp. 695-702
- Jacobs, R. A., (1988). Increased Rates of Convergence Through Learning Rate Adaptation. Neural Networks Vol. 1, No 4.
- Jaime, A. (1988), Características dinámicas de las arcillas del valle de México, Ph.D. Thesis, DEPEI, Facultad de Ingeniería, UNAM, México.
- Jin, Y.Q. Y Li, G., (2000), Detection of a scatter target over randomly rough surface by using the angular correlation function in a finite-element approach, Waves in Random Media, A Journal Recognized by the European Physical Society. Vol. 10, No. 2, April, pp. 273-280.
- Johnson, L. R. y Silva, W., (1981), The effects unconsolidated sediment upon the ground motion during local earthquakes. Bull. Seis. Soc. Am., Vol. 71, No. 1, pp. 127-142.
- Karantoni, F. V. y Bouckovalas, G., (1997), Description and analysis of building damage due to Pyros, Greece earthquake, Soil Dynamics and Earthquake Engineering, Vol. 16, No. 2, ISSN 0267-7261, pp. 141-150.
- Kanai, K., (1952), Relation between the nature of surface layer and amplitudes of earthquake motions, Bull. Tokyo Earth. Res. Ins., Vol. 30, pp. 31-37.
- Kawase, H. y Aki, K., (1989), A study on the response of a soft basin for incident S, P, and Rayleigh waves with special reference to the long duration observed in Mexico City, Bull. Seis. Soc. Ame. Vol. 79, No. 5, pp 1361-1382.
- Kohonen, T., (1972). Correlation Matrix Memories. IEEE, Transactions on Computers. Vol. 21.
- Kramer, S.L., (1996). Geotechnical Earthquake Engineering, Prentice-Hall International Series in Civil Engineering and Engineering Mechanics. University of Washington
- Lermo, J., Rodriguez, M., Singh, S. K. (1988), The Mexico earthquake of September 19, 1985 -- Natural period of sites in the Valley of Mexico from microtremor measurements and strong motion data, Earthquake Spectra, Vol. 4, No. 4, Nov., pp. 805-814.
- Luenberger, D. G (1984) Linear and Nonlinear Programming, 2nd edn, Addison-Wesley, Reading MA.
- Lysmer, J. y Drake, L. A., (1972), A finite element method for seismology, Methods in Computational Physics, Vol. 11, Ch. 6, academic Press.
- Marsal, R.J y Mazari, M., (1957), El subsuelo de la Ciudad de México, Facultad de Ingeniería, Universidad Nacional Autónoma de México.
- Martin, P. P., (1975), Non-linear methods for dynamic analysis of ground response, Ph.D. Dissertation, University of California, Berkeley, June.
- Matlab, The Language of Technical Computing, Versión 5.2.0, January 17, 1998. Paquete de programación.
- McCulloch, W., and Pitts, W., (1943). A logical calculus of the ideas immanent in Nervous Activity. Bulletin of Mathematical Biophysics, Vol. 5.
- Mena, E., et al. (1986). Catálogo de acelerogramas procesados del sismo del 19 de septiembre de 1985, Parte I: Ciudad de México, Instituto de Ingeniería, No. 497, Julio, México.
- Minsky, M. and Papert, S., (1969). Perceptrons. Cambridge, MA: MIT Press (Chapters 1, 4).

- Moczo, P., Bystricky E., Kristek, J., Carcione, J.M., Bouchon, M. (1997), Hybrid modeling of P-SV seismic motion at inhomogeneous viscoelastic topographic structure, Bull. Seis. Soc. Ame 87:1305-1323.
- Mooser, F., (2000), Archivos personales.
- Muravskii, G.T. y Frydman, S., (1998) Site response analysis using a non-linear hysteretic model, So. Dynamics and Earthquake Engineering, Vol. 17, No. 4. ISSN 0267-7261, June, pp 227-238.
- Ordaz, M., Singh, S.K., Reinoso, E., Lermo, J., Espinosa, J. M. y Domínguez, T., (1988), Estimation of response spectra in the lakebed zone of the valley of México during the Michoacán earthquake, Earthquake Spectra 4, pp. 815-834.
- Ordaz, M., Pérez-Rocha, L.E., Reinoso, E., Montoya C., y Arboleda, J. (1998). Programa Z. Espectros de sitio en el D.F., Instituto de Ingeniería, UNAM, CIS-FJBS, CENAPRED
- Ohminato T., y Chouet B. A., (1997), A free-surface boundary conditions for including 3D topography in the finite difference method, Bull. Seis. Soc. Ame., 87:494-515.
- Reid, H.F., (1910), The California earthquake of April 18, 1906, The Mechanics of the Earthquake, Vol. 2, Report of the State Earthquake Investigation Commission, Carnegie Institute of Washington, Publication 87, Washington D. C.
- Reinoso, E., Wrobel, L.C., y Power, H., (1997). Three-dimensional scattering of seismic waves from topographical structures, Soil Dynamics and Earthquake Engineering, Vol. 16, 41-61
- Rial, J.A., Saltzman, N.G. y Ling, H. (1992). Earthquake-induced resonance in sedimentary basin American Scientist, Vol. 80, No. 6, pp. 566-578.
- Romo, M.P. y Jaime, A. (1986). Características dinámicas de las arcillas del valle de México y análisis de respuesta sísmica del suelo, Reporte Técnico, Instituto de Ingeniería, UNAM.
- Romo, M.P. y Seed, H. B., (1986). 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, pp. 148-162, Mexico.
- Romo, M.P., (1987). Foundation Engineering in Mexico City: Seismic Aspects, Proceedings International Symposium on Geotechnical engineering of soft soil, 2, 213-224, Ciudad de México.
- Romo, M.P., Jaime, A. y Reséndiz, D. (1988). General Soil Conditions and Clay Properties in the Valley of Mexico, Earthquake Spectra, Vol 4, Num 4.
- Romo, M. P., Jaime, A., y Taboada, V. M., (1989). Cyclic behavior of México City clay, Internal Report, Institute of Engineering, UNAM.
- Romo, M.P., Ovando, E., Jaime, A. y Hernández, G. (1989). Local site effects on Mexico City Ground Motions, XII International Conference on soil Mechanics and Foundation Engineering, Rio de Janeiro, August.
- Romo, M. P., Magaña, R. y Bárcena, A., (1990). Respuesta no lineal de depósitos de suelo, Informe Interno Instituto de Ingeniería, UNAM. México.
- Romo, M. P., (1991). Comportamiento dinámico de la arcilla de la ciudad de México y sus repercusiones en la Ingeniería de cimentaciones, Sismodinámica 2, pp. 125-143.
- Romo, M.P., y Ovando, E., (1994). Comportamiento dinámico y estático de los suelos del sitio Proyecto Alameda, Informe del Instituto de Ingeniería elaborado para Reichmann International, junio.
- Romo, M.P., y Ovando, E., (1995). Comportamiento dinámico de las arcillas de la sonda de Campeche, Informe del Instituto de Ingeniería, UNAM.
- Romo, M.P., (1995). Clay behavior, ground response and soil-structure interaction studies in Mexico City, state of the art paper, III International Conference on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics, St. Louis Missouri, USA, April.
- Romo, M.P., Rangel, J.L., Flores, O. y García, S.R., (1998), Aplicación de redes neuronales artificiales a la geotecnia, XIX Reunión Nacional de Mecánica de Suelos, SMMS, Puebla, Mex pp. 418-427.
- Romo, M.P., (1999). Earthquake Geotechnical Engineering and Artificial Neural Networks, IV Arthur Casagrande Lecture, XI Panamerican congress on soil Mechanics and Geoenvironmental Engineering, Foz do Iguassu, Brasil, Vol. IV.
- Romo, M.P., Mendoza, M. J. y García, S. (2000). Geotechnical factors in seismic design of foundations state-of-the-art report, Cuaderno Técnico No. 1, S'MMS.
- Rosenblueth, E. (1952), Teoría del diseño sísmico sobre mantos blandos, Ediciones ICA, Serie B. 14 pp. 3-12, México.
- Rosenblueth, E. y Ovando, E., (1990), Riesgo sísmico en el Valle de México: una perspectiva geotécnica. El subsuelo en la cuenca del Valle de México y su relación con la ingeniería de cimentaciones a cinco años del sismo, SMMS, pp. 83-94.

- Rosenblatt, F., (1958). The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. Psychological Review, Vol. 65.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J., (1986). Learning representation by Back propagation Errors. Nature, Vol. 323.
- Rumelhart, D. E., and McClelland, J. L., Eds., (1986). Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1. Cambridge, MA:MIT Press.
- Sánchez-Sesma, F. J., Campillo, M., Bard, P.Y., Gariel, J.C. y Aki, K. (1989), The Great 1985 Michoacan Earthquake: a Unified Approach Considering Source, path and site effects, in Engineering Seismology and Site Response, Eds. Cakmac A. S. y Herrera Y, Computational Mechanics Publications, south Hampton, pp. 53-75.
- Sánchez-Sesma, F., (1990). Elementary solutions for response of a wedge-shaped medium to incident SH- and SV-waves, Bulletin of the Seismological Society of America. Vol. 80, pp. 737-742.
- Sánchez-Sesma, F. y Campillo, M. (1993). Topografic effects for incident P, SV, and Rayleigh waves. Tectonophysics, Vol. 218, No. 1-3, pp. 113-125.
- Sánchez-Sesma, F. J., Chavez-Perez, S., Aviles, J., (1994). Scattering of elastic waves by three-dimensional topographies, Proceedings of the 8th world conference on earthquake engineering, san Francisco, Vol. II, pp. 639-46.
- Seed, H.B. y Idriss, I., (1969). The Influence of Soil Conditions on Ground Motions during Earthquakes. Journal of Soils Mechanics and Foundations Division, ASCE, Vol. 94, No. SM1., pp. 120-137.
- Seed, H.B. Ugas, c. y Lysmer, J., (1976). Site-dependent spectra for Earthquake-Resistant Desing, Report. No EERC 74-12, November.
- Seed, H.B., Romo, M.P., Sun, Jaime, A. y Lysmer, J., (1987). Relationships Between Soil Conditions and Earthquake Ground Motions in Mexico City in the Earthquake of september 19, 1985, Report No. UCB/EERC-87/15, October
- Seed, H.B., Romo, M.P., Sun, Jaime, A. y Lysmer, J., (1987). Relationships Between Soil Conditions and Earthquake Ground Motions in Mexico City in the Earthquake of september 19, 1985. Report No UCB/EERC-87/15, October.
- Seed, H.B., Dickenson, S., Reimer, M. Bray, J., Sitar, N., Mitchell, J. Idriss, Y. Kayen, R. Kropp, A., Harder, L., y Power, M. (1990). Rpreliminary report on the principal geotechnical aspects of the October 17, 1989 Loma Prieta earthquake, Report No. UCB/EERC-90/05, EERC, University California, Berkeley, 137 pp.
- Semblat, J.F., Dangla, P. y Duval, A.M., (1999), BEM analysis of seismic wave amplification in Caracas, 7th International symposium on Numerical Models in Geomechanics (NUMOG), Graz, Australia, Balkema, pp 275-80.
- Scales, L. E., (1985). Introduction to Non-Linear Optimization. New York: Springer-Verlag.
- Schnabel, Lysmer y Seed, (1972), SHAKE: A computer Program for Earthquake Response Analysis of Horizontally Layered Sites, Report N. EERC/72-12, University of California, Berkeley, December
- Seed, Bray, Chang y Dickenson (1997), Site-dependent seismic response including recent strong motion data. Seismic Behavior of ground and Geotechnical Structures, Edit. Seco Pinto, Balkema, Rotterdam.
- Shanno, D. F., (1990). Recent Advance in Numerical Techniques for Large-Scale Optimization. In Neural Networks for Control, Miller, Sutton and Werbos, Eds, Cambridge, MA:MIT Press.
- Shepherd, A. J., (1997). Second-Order Methods for Neural Networks. Fast and Reliable Training Methods for Multi-Layer Perceptrons (Perspectives in Neural Computing), Ed. Springer, London, UK.
- Silva W. (1978), Wave propagation in anelastic media with application to seismology, PhD Dissertation University of California, Berkeley.
- Singh, S. K. y Ordaz, M. (1993), On the origin of long coda observed in the lake-bed strong-motion records of Mexico City, Bulletin of the Seismological Society of America, vol. 83, pp. 1298-1306.
- Singh, S. K., Ordaz, M., Pacheco, J. F., Quass, r., Alcantara, L., Alcocer, S., Gutierrez, c., Meli, R., y Ovando, E. (1999), A preliminary report on the Tehuacán, México Earthquake of June 15, 1999 (Mw=7.0). Seismological Research Letters, Vol. 70 No. 5, Sep/Oct. pp. 489-504
- Somerville, P. (1998), Emerging art: earthquake ground motion, Geotechnical Earthquake Engineering and Soil Mechanics III, Vol. Y. ASCE, Proc. of speciality conference.
- Specht, D. F. (1991). A General Regression Neural Network. IEEE, Transactions on Neural Networks, Vol. 2, No. 6.
- Swanger, H. J. y Boore, D. M. (1978), Simulation of strong displacements using surface-wave model superposition, Bull. Seis. Soc. am., Vol. 68, No. 4, pp. 247-263, March.

- Taboada, V.M., Martínez, H.E. y Romo, M.P. (1998), **Análisis de algunos registros de aceleración en el sitio de Central de Abastos Oficinas**, XIX Reunión Nacional de Mecánica de Suelos, SMMS, Puebla, Mex. pp. 313-318.
- Trifunac M.D. (1973), **Scattering of plane SH waves by a semi-cylindrical canyon**, Earthquake Engng. Struct. Dynam., 1:267-281
- Trifunac M.D. y Hudson, E.D. (1971), **Sanalysis of the Pacoima Dam accelerograms – San Fernando earthquake of 1971**, Bulletin of the Seismological Society of America. Vol. 61, No. 5. pp. 1393-1411.
- Todorovska, M.I. y Lee, V.W. (1991), **A note on scattering of Rayleigh waves by shallow circular canyons: analytical approach**, Bull. Indian Soc. Earthquake Tech. 20:1-16.
- Tollenaere, T., (1990). **SuperSAB: Fast Adaptive Back Propagation with Good Scaling Properties**. Neural Networks, Vol. 3, No 5.
- Vai, Rossana; Castillo-Covarrubias, José Manuel; Sánchez-Sesma, Francisco J.; Komatitsch, Dimitri; Vilotte, Jean-Pierre. (1999), **Elastic wave propagation in an irregularly layered medium**, pp. 11-18 , January, vol 18, No 1
- Vidale, J. E. y Helmberger, D.V. (1988), **Elastic finite difference of the 1971 San Fernando earthquake**, Bull Seis. Soc. Am., Vol.78, No. 1, pp. 122-141.
- Wessels, L. F. A. y Barnard, E. (1992) **Avoiding false local minima by proper initialization of connections**, IEEE Transactions on Neural Networks, 3(6), pp. 899-905.
- Wong H.L. y Trifunac M.D. (1974), **Scattering of plane SH waves by a semi-elliptical canyon**, Earthquake Engng. Struct. Dynam., 3:157-169.