

Centro de Investigación Científica y de Educación Superior de Ensenada



Investigación Sobre Variantes del Método de Cadenas de Markov para la Estimación de Peligro Sísmico

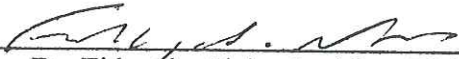
TESIS

DOCTORADO EN CIENCIAS

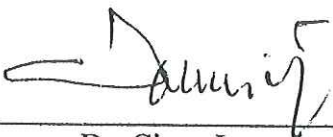
CLAUDIA SOLEDAD HERRERA OLIVA

ENSENADA BAJA CFA, MEXICO OCTUBRE DE 2005

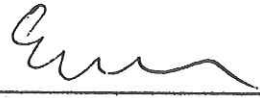
TESIS DEFENDIDA POR
Claudia Soledad Herrera Oliva
Y APROBADA POR EL SIGUIENTE COMITÉ


Dr. Fidencio Alejandro Nava Pichardo

Director del Comité


Dr. Cinna Lomnitz

Miembro del Comité


Dr. Ewa Glowacka

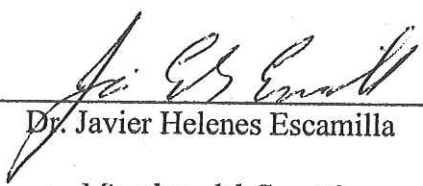
Miembro del Comité


Dr. Luis Munguía Orozco

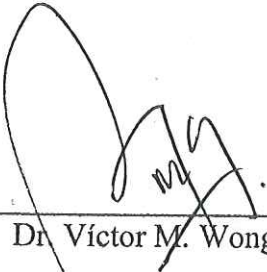
Miembro del Comité


Dr. Cecilio Javier Rebollar Bustamante

Miembro del Comité


Dr. Javier Helenes Escamilla

Miembro del Comité


Dr. Víctor M. Wong Ortega
Coordinador de Posgrado en Ciencias de la Tierra


Dr. Raúl Ramón Castro Escamilla
Director de Estudios de Posgrado

07 de Octubre del 2005

**CENTRO DE INVESTIGACIÓN CIENTÍFICA Y DE EDUCACIÓN SUPERIOR
DE ENSENADA**



**PROGRAMA DE POSGRADO EN CIENCIAS
EN CIENCIAS DE LA TIERRA CON ORIENTACIÓN EN SISMOLOGÍA**

**INVESTIGACIÓN SOBRE VARIANTES DEL MÉTODO DE CADENAS DE MARKOV
PARA LA ESTIMACIÓN DE PELIGRO SÍSMICO**

TESIS

que para cubrir parcialmente los requisitos necesarios para obtener el grado de
DOCTOR EN CIENCIAS

Presenta:

CLAUDIA SOLEDAD HERRERA OLIVA

Ensenada, Baja California, México, Octubre del 2005.

RESUMEN de la tesis de **Claudia Soledad Herrera Oliva**, presentada como requisito parcial para la obtención del grado de DOCTOR EN CIENCIAS en CIENCIAS DE LA TIERRA con orientación en SISMOLOGÍA. Ensenada, Baja California, México. Octubre de 2005.

INVESTIGACIÓN SOBRE VARIANTES DEL MÉTODO DE CADENAS DE MARKOV PARA LA ESTIMACIÓN DE PELIGRO SÍSMICO.

Resumen aprobado por:



Dr. Fidencio Alejandro Nava Pichardo
Director de Tesis

Un estudio previo presentó un nuevo método para obtener estimaciones de peligro sísmico para un *sistema*, que es un área geográfica que consta de varias regiones sísmogénicas, expresando la probabilidad de ocurrencia de patrones de presencia o ausencia de sismicidad, con magnitud mayor o igual a una magnitud umbral M_r , en las regiones como una cadena de Markov. La aplicación de este método, basado en el conteo de transiciones entre estados, al que llamaremos aquí *directo*, al área de Japón dió resultados muy satisfactorios.

Esta tesis presenta nuevos métodos de estimación de peligro sugeridos por los resultados del método directo. Dado que la limitante más importante del método directo es la relativa escasez de eventos de alta magnitud, algunos de estos métodos exploran la posibilidad de que sismicidad con magnitud umbral M_r^m contenga información pertinente a la ocurrencia de terremotos con $M_r^M > M_r^m$. Estos métodos, llamados aquí *cruzados*, permiten estimar peligro sísmico para mayores magnitudes que el método directo. El método de conteo cruzado mejora en 6.6% los resultados del método directo.

Otros métodos aprovechan la capacidad de aprendizaje de una red neuronal elemental (perceptrón) para modelar el comportamiento markoviano del sistema; métodos neuronales tanto directos como cruzados, fueron implementados con resultados que superan en 67% los de los métodos de conteo.

Como parte del estudio, se enfrentó el problema de evaluación del desempeño de los métodos de estimación de peligro. Se optó por emplear funciones calificadoras, *i.e.* funciones matemáticas que balancean características positivas contra características negativas del método, a manera de asignar una calificación normalizada que permite comparar diferentes modelos y realizaciones.

Palabras clave: Peligro sísmico, cadenas de Markov, redes neuronales, catálogos sísmicos.

ABSTRACT of the thesis of **Claudia Soledad Herrera Oliva**, presented as partial requirement to obtain the DOCTOR IN SCIENCES degree in EARTH SCIENCES with orientation in SEISMOLOGY. Ensenada, Baja California, Mexico. October 2005.

INVESTIGATION ABOUT VARIANTS OF THE METHOD OF MARKOV CHAINS FOR SEISMIC HAZARD EVALUATION.

ABSTRACT

A previous work introduced a new method for seismic hazard evaluation in a *system*, which is a geographic area with distinct, but related seismogenic regions, based on modeling the transition probabilities of *states* *i.e.* patterns of presence or absence of seismicity, with magnitude greater or equal to a threshold magnitude M_r , in the regions, as a Markov chain. Application of this method, based on a straightforward count of observed transitions between states, called here *direct*, to the Japan area gave very satisfactory results.

This thesis presents new methods suggested by the results of the direct method. Given that the most important limitation of the direct method is the relative scarcity of large magnitude events, we decided to explore the possibility that seismicity with magnitude $M \geq M_r^m$ contains information about the future occurrence of earthquakes with $M \geq M_r^M > M_r^m$. These methods, called here *mixed methods*, allow to estimate seismic hazard for larger magnitudes than the direct method. The mixed counting method improves the results of the direct method by 6.6%.

Other methods use of the learning capacity of an elementary neuronal network (perceptron) to model the markovian behavior of the system; neuronal methods (direct and mixed), were implemented and gave results 67% better than the counting methods.

As part of this study, the problem of performance evaluation of hazard estimation methods is addressed, leading to the use of grading functions, *i.e.* mathematical functions that evaluate a tradeoff between the positive and negative characteristics of the method, so as to assign a normalized grade that allows comparison of different models and realizations.

Key words: Seismic hazard, Markov chains, neural networks, seismic catalogues.

DEDICATORIA

A mis padres
Sra. María Teresa Oliva y Sr. Nicanor Herrera.

A mis hermanas
Gaby y Faby

A mis abuelitos
Soledad Ponce Miguel Herrera
María Teresa Contreras Rubén Oliva

A mi amore
José Rubén Campos Gaytán

A DIOS
Por todo.

AGRADECIMIENTOS

Al CONACyT por la beca crédito otorgada. Al finalizar ésta el apoyo económico dimanó el fondo de becas de la División de Ciencias de la Tierra del CICESE.

A mi asesor y amigo Alejandro Nava Pichardo por su apoyo, confianza y valiosa dirección.

A los miembros del comité de tesis Dr. Cinna Lomnitz A., Dr. Luis Munguía Orozco, Dr. Cecilio J. Rebollar Bustamante, Dr. Javier Helenes Escamilla y Dra. Ewa Glowacka, por sus valiosos comentarios y sugerencias, que enriquecieron ampliamente el contenido de la versión final de la tesis.

A Francisco Farfán e Ignacio Méndez por su asesoría en la adquisición de catálogos sísmicos.

A Miguel A. Garcia Arthur, Humberto Benítez y Daniel Peralta por su invaluable asistencia computacional.

A Martha Elba Barrera, Citlalli Romero, Silvia Cruz, Maria del Carmen Pérez y Concepción González por su ayuda en las labores administrativas.

Al personal de la biblioteca del CICESE, Lupita Morales, Cecilia González y Guadalupe Zamora por sus atenciones.

A mis buenos amigos del CICESE: Rubén, Tere, Carlos, Adriana, Yanina, Silvia, Miguel Ángel y especialmente a Rosalía por su calidez y apoyo incondicional.

A las Familias Campos Gaytán, Moreno Campos, Campos Ferret, León Campos, León Félix, Campos Villavicencio, Aguilar Campos y León Serrano por hacerme sentir parte de ellos.

CONTENIDO

Página

I. INTRODUCCIÓN.....	1
I.1. Planteamiento del problema.....	1
I.2. Objetivo y metas.....	3
I.2.1. Objetivo.....	3
I.2.2. Metas.....	3
II. ANTECEDENTES.....	4
II.1. Conceptos básicos.....	4
II.1.1. Rebote elástico.....	5
II.1.2. Tiempos de Recurrencia.....	5
II.1.3. Gap Sísmico.....	8
II.1.4. Migración.....	8
II.1.5. Relación Gutenberg-Richter.....	8
II.2. Métodos clásicos para la determinación de peligro sísmico.....	9
II.2.1. Modelo de Poisson.....	9
II.2.2. Modelo de Hagiwara.....	10
II.2.3. Modelos basados en la teoría del gap sísmico.....	11
II.2.4. Métodos de reconocimiento de patrones.....	11
III. METODOLOGÍA.....	13
III.1 Método Markoviano para la estimación de peligro sísmico.....	13
III.1.1 Sistema.....	13
III.1.2. Premisas.....	14
III.1.3. Parámetros umbral.....	15
III.1.4. Método Directo.....	15
III.2. Nuevos métodos.....	16
III.2.1 Métodos markovianos.....	16
III.2.1.1 Método Markoviano cruzado.....	17
III.2.1.2. Método de caminos preferenciales.....	18
III.2.2. Métodos Neuronales.....	19
III.2.2.1 Perceptrón simple.....	20
III.2.2.1.1. Entrada y salida del perceptrón.....	21
III.2.2.1.2 Entrenamiento del perceptrón.....	21
III.2.2.1.3. Perceptrón simple con bias.....	22
III.2.2.2. Perceptrón con información de magnitudes.....	22
III.2.2.3. Perceptrón cruzado.....	23
III.2.2.4. Perceptrón con historia.....	23
III.2.2.4.1 Perceptrón con entrada simultánea de 2 pasos.....	23
III.2.2.4.2. Perceptrón con entrada de 2 pasos codificados.....	24
III.2.2.5. Redes neuronales de 2 o más capas.....	25

CONTENIDO (continuación)

	Página
III.2.2.6. Perceptrón por regiones.....	25
IV. EVALUACIÓN DEL DESEMPEÑO DE LOS MÉTODOS.....	26
IV.1. Probabilidad de transición observada promedio.....	26
IV.2. Pronósticos.....	27
IV.2.1. Pronósticos exitosos, falsas alarmas y eventos perdidos.....	27
IV.2.2. Multiplicidad.....	28
IV.2.3. Error regional.....	29
IV.2.4. Utilidad de los pronósticos.....	30
IV.2.5. Factor de éxito.....	31
IV.3. Funciones calificadoras.....	31
IV.4 Pronósticos y postnósticos.....	33
V. APLICACIÓN Y RESULTADOS.....	34
V.1. Localización de la zona de estudio.....	34
V.2. Tectónica regional.....	34
V.2.1. Procesos de subducción y arcos de isla.....	34
V.2.2. Fallas activas.....	36
V.3. Sismicidad.....	38
V.4. Descripción de la base de datos.....	39
V.5. Aplicación de los métodos a Japón.....	39
V.5.1 Elección de los parámetros umbral.....	39
V.5.1.1 Parámetro umbral de espacio.....	39
V.5.1.2. Umbrales de magnitud.....	41
V.5.1.3. Umbral de tiempo.....	41
V.5.2. Estados del Sistema.....	42
V.6. Resultados de la aplicación de los métodos al área de Japón.....	43
V.6.1 Resultados I.....	43
V.6.1.1. Evaluación del desempeño del método directo.....	44
V.6.1.2. Evaluación del desempeño del método cruzado.....	46
V.6.2. Resultados II.....	50
V.6.2.1. Evaluación del desempeño del perceptrón simple	50
V.6.2.1. Evaluación del desempeño del perceptrón cruzado	54
V.6.2.3. Resultados del método de caminos preferenciales, y de los perceptrones con historia.....	57
V.6.2.4 Resultados de la aplicación del perceptrón con información de magnitudes.....	58

CONTENIDO (continuación)

	Página
VI. DISCUSIÓN Y CONCLUSIONES.....	60
VI.1. Discusión.....	60
VI.2. Conclusiones.....	62
REFERENCIAS.....	65
APÉNDICE A.....	69
APÉNDICE B.....	71
APÉNDICE C.....	75

LISTA DE FIGURAS

Figura	Página
1 Esquematización de los modelos de recurrencia: (a) modelo periodico. (b) modelo predecible. (c) modelo de corrimiento predecible, τ_1 y τ_2 son los esfuerzos que actúan sobre la falla.....	7
2 Ejemplo del perceptrón con entrada simultánea de dos pasos para un sistema con $S = 4$ posibles estados.....	24
3 Mapa tectónico de la parte occidental del Océano Pacífico; el recuadro marca la zona de estudio.....	35
4 Distribución de fallas activas en y al derredor de las islas japonesas. Tomada de Matsuda y Kinugusa, 1991.....	37
5 Sismicidad del Japón reportada por el ISC (1964-2002).....	38
6 El sistema y sus regiones identificadas por los índices correspondientes.....	40
7 Valores de las funciones calificadoras del método directo en función de M_r . Izquierda <i>postnósticos</i> . Derecha <i>pronósticos</i>	44
8 Valores de las funciones calificadoras del método <i>cruzado</i> en función de M_r^m . Izquierda <i>postnósticos</i> . Derecha <i>pronósticos</i>	47
9 Representación de los valores de las matrices de probabilidad de transición codificados según la banda a la derecha. Izquierda: Método directo, para $M_r = 6.2$. Derecha: Método cruzado para $M_r^m = 5.7$ y $M_r^M = 6.2$	48
10 Valores de las funciones calificadoras del método <i>cruzado</i> para $M_r^M = 6.5$ en función de M_r^m . Izquierda <i>postnósticos</i> . Derecha <i>pronósticos</i>	49
11 Valores de los elementos de las matrices de probabilidades de transición codificados según las bandas mostradas a la derecha para $M_r = 6.1$	51

LISTA DE FIGURAS (continuación)

Figura	Página
12 Valores de las funciones calificadoras del perceptrón <i>simple</i> en función de la magnitud umbral M_r . Izquierda <i>postnósticos</i> . Derecha <i>pronósticos</i> . Las líneas discontinuas representan valores de los discriminantes observados por el método markoviano directo.....	52
13 Valores de las funciones calificadoras del perceptrón <i>cruzado</i> en función de la magnitud umbral. Izquierda <i>postnósticos</i> . Derecha <i>pronósticos</i> . Las líneas discontinuas representan valores de los discriminantes observados por el método directo para $M_r = 6.2$	55
14 Representación de los valores de las matrices de probabilidad de transición codificados según la banda a la derecha. Izquierda: Peceptrón directo, para $M_r = 6.2$. Derecha: Perceptrón cruzado para $M_r^m = 5.7$ y $M_r^M = 6.2$	56
15 Valores de las funciones calificadoras del perceptrón <i>cruzado</i> para $M_r^M = 6.5$ en función de M_r^m . Izquierda <i>postnósticos</i> . Derecha <i>pronósticos</i>	57

LISTA DE TABLAS

Tabla	Página
I Parámetro umbral de magnitud y número de sismos con $M \geq M_r = 6.1$ para cada región, durante el periodo 1964-2002.....	41
II Estados del sistema.....	42
III Resultados de la evaluación de postnósticos. n_t es el número de transiciones, f_x es el factor de éxito, n_x es el número de éxitos, p^b es la probabilidad binomial de Bernoulli de observar n_x éxitos en n_t transiciones con una probabilidad u	45
IV Resultados de la evaluación de pronósticos. Cantidades como en la tabla III.....	46
V Resultados de la evaluación de postnósticos de ambos perceptrones para altas magnitudes. n_t es el número de transiciones, f_x es el factor de éxito, n_x es el número de éxitos, p^b es la probabilidad binomial de Bernoulli de observar n_x éxitos en n_t transiciones, cada una con probabilidad $p = u$	53
VI Resultados de la evaluación de pronósticos de los perceptrones para altas magnitudes. Cantidades como en la tabla V.....	54

INVESTIGACIÓN SOBRE VARIANTES DEL MÉTODO DE CADENAS DE MARKOV PARA LA ESTIMACIÓN DE PELIGRO SÍSMICO

I. INTRODUCCIÓN

Actualmente no es posible *predecir* adecuadamente la ocurrencia de terremotos, esto es: especificar, anticipada y determinísticamente, la ocurrencia de un sismo específico, con localización espaciotemporal y magnitud determinadas, y con márgenes pequeños en dichas determinaciones. Por tanto, una de las metas más importantes de los sismólogos es la determinación de peligro sísmico (entendiendo por peligro sísmico la probabilidad de ocurrencia de uno o más sismos en un área geográfica específica durante intervalos de tiempo y magnitud determinados), porque las estimaciones de peligro son la base de la determinación de riesgo sísmico, el cual juega un papel muy importante en la planeación de muchas actividades humanas; particularmente económicas y de defensa civil.

I.1. *Planteamiento del problema*

Las estimaciones de peligro sísmico, están frecuentemente basadas en análisis estadísticos de la historia sísmica del área de interés, y suelen determinarse mediante el ajuste de alguna distribución de probabilidad a modelos de ocurrencia de terremotos; modelos como el asociado con la distribución Gutenberg-Richter (Gutenberg y Richter, 1944) y como variantes probabilísticas de modelos basados en la teoría del rebote elástico (Reid, 1910). Ejemplo de estas variantes son los ampliamente utilizados modelos de tiempo predecible, deslizamiento predecible o modelos mixtos (Shimazaki y Nakata, 1980) y de vacancia sísmica (Fedotov, 1965). Sin embargo, las estimaciones no han resultado

suficientemente confiables dada la gran cantidad de suposiciones implícitas y de factores desconocidos, sobre los que aún no se tiene control, que contienen los modelos. Otras estimaciones probabilísticas de peligro sísmico se basan en modelos de sismicidad Poissoniana (e.g. Lomnitz, 1974; Brillinger, 1982; Lomnitz y Nava, 1983); pero, si bien este tipo de sismicidad puede ser apropiado para la sismicidad mundial o para regiones afectadas por terremotos originados en muchas regiones, definitivamente no es apropiada para el estudio de una región sismogénica en particular, donde debe ser aplicable el modelo sísmico causal del rebote elástico (Lomnitz y Nava, 1983).

Para contribuir a la evaluación de peligro sísmico, Herrera (2001) y Nava *et al.* (2005) presentaron un modelo estadístico que explora la posibilidad de modelar la probabilidad de ocurrencia de patrones de sismicidad en un área geográfica con varias regiones sismogénicas (un *sistema*) como una cadena de Markov. Este método, al que llamaremos *directo* y que estima el peligro sísmico a partir del conteo directo de la incidencia de transiciones observadas, dio resultados altamente satisfactorios (aciertos en la evaluación de probabilidades muy por encima de los resultantes de dos modelos sin memoria: el de distribución uniforme de probabilidades y el altamente utilizado modelo de Poisson). Las matrices de probabilidad de transición obtenidas mediante la aplicación del método directo son distintas a las matrices poissonianas, y la aplicación del método directo a catálogos sintéticos poissonianos dió resultados muy inferiores a los obtenidos para el catálogo real.

Esta tesis presenta nuevos métodos de estimación de peligro sísmico sugeridos por los resultados del método directo, así como una manera de evaluar el desempeño de los métodos. Algunos de estos métodos son variaciones del método directo, que permiten

conocer mejor el sistema o extender su aplicabilidad a mayores magnitudes. Otros métodos aprovechan la capacidad de aprendizaje de una red neuronal elemental para modelar el comportamiento markoviano del sistema.

I.2. *Objetivo y metas*

I.2.1. *Objetivo*

Implementar y probar métodos estadísticos, sugeridos por el método markoviano directo de Herrera (2001) y Nava *et al.* (2005), para obtener estimaciones confiables de peligro sísmico.

I.2.2. *Metas*

Diseño e implementación de los métodos.

Diseño e implementación de medidas para evaluar el desempeño de los métodos.

Aplicación de los métodos a un área específica para:

Evaluar su aplicabilidad y desempeño

Obtener estimaciones de peligro para esa área.

II. ANTECEDENTES

Como se mencionó en el capítulo anterior la predicción sísmica *i.e.* la estimación determinista acerca del tiempo, lugar y magnitud de un terremoto es prácticamente imposible porque el proceso de generación de un sismo es muy complejo, y además dicho proceso tienen su origen kilómetros bajo la superficie de la tierra, lo que hace imposible la medición u observación directa de los procesos en cuestión. Por tanto, para manejar en forma promediada los efectos de factores imposibles de medir in situ, como macro y microfallas, distribuciones de asperidades, concentraciones de esfuerzo, etc., es necesario formular modelos para la ocurrencia de terremotos en terminos probabilísticos. Estos modelos, frecuentemente se basan en conceptos como los descritos someramente a continuación. Algunos modelos son puramente estadísticos, mientras que otros son basados en consideraciones físicas o fenomenológicas, pero también requieren análisis estadístico.

Casi todos los modelos con componente física se basan, directa o indirectamente, en el modelo de rebote elástico de Reid discutido más abajo; entre ellos podemos mencionar: *tiempos de recurrencia* determinados a partir de razonamientos sismotectónicos (e.g. Shimazaki y Nakata, 1980); *gaps sísmicos (o vacancias)* (e.g. Fedotov, 1965) y *migración de sismicidad* (e.g. Richter, 1958; Mogi, 1968b).

II.1. Conceptos básicos

Con la finalidad de que los modelos empleados en la evaluación de peligro sísmico queden claros, en esta sección se explicará brevemente los conceptos básicos, mencionados anteriormente, en que se apoyan estos modelos.

II.1.1. *Rebote elástico*

El movimiento relativo de placas tectónicas adyacentes produce acumulación de esfuerzos en el terreno. Cuando los esfuerzos alcanzan un valor crítico, que puede ser el esfuerzo de rompimiento o *resistencia* de la roca, o la fricción estática de una falla preexistente, el material se rompe súbitamente (*falla*) y los esfuerzos son relajados; esta ruptura súbita con la correspondiente radiación de ondas elásticas es lo que se conoce como sismo.

La teoría de que los sismos son producidos por fallamiento cuando son liberados los esfuerzos acumulados es conocida con el nombre de *rebote elástico* y fue propuesta por H.F. Reid en 1910.

II.1.2. *Tiempos de Recurrencia*

Dado que la velocidad de los movimientos relativos entre las placas puede considerarse constante para tiempos de interés sismológico, la acumulación de esfuerzos (suponiendo elasticidad lineal) es también a velocidad constante. Así, si suponemos que en un segmento de la falla un terremoto ocurre cuando el esfuerzo alcanza el nivel de resistencia y libera éste total o parcialmente, y que es necesario que el esfuerzo vuelva a alcanzar el nivel de la resistencia antes de que ocurra otro terremoto, entonces debe haber intervalos entre los terremotos. Hay tres modelos basados en esta premisa y otras suposiciones que han sido usados para el cálculo de riesgo y que se describe a continuación.

En el modelo más simple, los esfuerzos se acumulan hasta alcanzar la resistencia del medio y entonces ocurre un sismo que relaja los esfuerzos hasta un valor mínimo (asociado generalmente con el valor de la fricción dinámica en la falla), tras de lo cual los esfuerzos comienzan a acumularse de nuevo (figura 1a). En este caso, los terremotos, todos del mismo tamaño, se repiten a intervalos iguales de tiempo conocidos como *intervalos*, o *tiempos de recurrencia*.

Este modelo es poco realista ya que, si bien se puede considerar constante la tasa de acumulación de esfuerzos suponiendo constante la velocidad de deslizamiento de las placas, la heterogeneidad del material causa que ni el esfuerzo de ruptura ni la relajación de esfuerzos sean siempre los mismos, por lo que difícilmente los terremotos serán siempre del mismo tamaño. También, la energía acumulada puede ser liberada por varios sismos en vez de por un solo sismo mayor y, por otro lado, existe la posibilidad de que parte de los esfuerzos acumulados se relajen por deslizamiento asísmico. Por estas causas, que son parte integral del ciclo sísmico, es difícil pensar que exista un tiempo de recurrencia determinado.

Un modelo más realista es el llamado *de tiempo predecible*, propuesto por Shimazaki y Nakata (1980). En este modelo los esfuerzos sobre una falla se relajan cuando llegan a un cierto valor crítico, como en el caso anterior, pero no siempre con la misma caída de esfuerzos. Los esfuerzos son relajados por terremotos de distinto tamaño y los intervalos entre éstos no son iguales, sino que dependen del tamaño del terremoto precedente. El modelo permite por lo tanto, conocidos la resistencia y el tamaño del terremoto precedente, predecir el tiempo del siguiente terremoto, pero no su tamaño

(figura 1b). De nuevo este modelo es demasiado simplista, por lo que sus predicciones no son confiables.

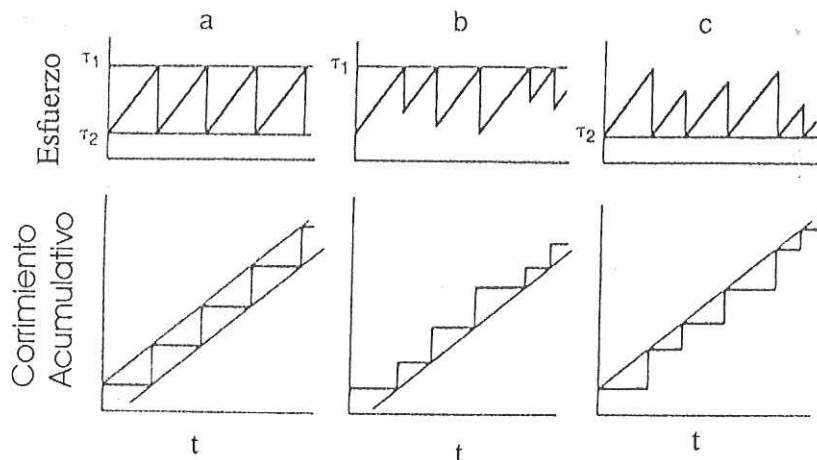


Figura 1. Esquematzación de los modelos de recurrencia: (a) modelo periódico. (b) modelo de tiempo predecible. (c) modelo de corrimiento predecible, τ_1 y τ_2 son los esfuerzos que actúan sobre la falla.

Otro modelo, llamado *de corrimiento predecible* (figura 1c), dice que cada vez que ocurre un sismo el nivel de esfuerzos en la falla baja a un valor determinado (el nivel de la fricción dinámica, supuesto cero para modelos que predicen fusión o separación temporal en la falla), para lo cual el corrimiento de la falla en un sismo dado debe ser tal que reponga la deficiencia de corrimiento que causa el esfuerzo. De esta manera, se puede predecir el tamaño de un terremoto que ocurra a un tiempo dado, pero no se puede decir cuándo será éste.

II.1.3. *Gap Sísmico*

Fedotov observó en 1965 que los grandes sismos *recurren*, rompiendo áreas muy similares a las de terremotos previos comparables a ellos; y que, las áreas de ruptura de los grandes sismos casi no se traslapan con las adyacentes. Esto llevó al concepto de “*gap*” (también llamado en español *hueco, brecha o vacancia*) *sísmico*, que es una sección de frontera entre placas donde se sabe que han ocurrido sismos grandes en el pasado, pero no recientemente durante un tiempo suficiente para almacenar la energía necesaria para un nuevo gran sismo.

II.1.4. *Migración*

Los primeros estudios que sugirieron la idea de que los epicentros de los terremotos *migran*, es decir, definen una trayectoria que puede indicar la dirección y el tiempo aproximado en que ocurrirá el siguiente sismo, fueron realizados por Richter (1958) y Mogi (1968b). La idea de la migración es razonable si se considera que los sismos producen concentraciones de esfuerzo en las áreas vecinas a la ruptura que, a su vez, pueden disparar nuevos terremotos (Nava, 1987).

II.1.5. *Relación Gutenberg-Richter*

Los primeros estudios estadísticos de sismicidad fueron hechos por Gutenberg y Richter (1944), quienes encontraron que el número N de sismos con magnitudes mayores o iguales a M , que ocurre en un intervalo de tiempo y una región determinados, es función lineal de dicha magnitud:

$$\log N(M) = a - bM, \quad (1)$$

donde b es una constante que tiene valores característicos para distintas regiones de la Tierra y a es una constante que depende de la región y del tiempo de muestreo. La relación Gutenberg-Richter dice cuántos sismos de determinada magnitud ocurrirán por unidad de tiempo, pero no cómo están distribuidos en el tiempo.

Muchas estimaciones de peligro sísmico están basadas en el número de sismos arriba de una magnitud umbral que predice la relación Gutenberg-Richter.

II.2. Métodos clásicos para la determinación de peligro sísmico

A continuación se hace una breve descripción de algunos de los modelos clásicos empleados en la evaluación de peligro sísmico.

II.2.1. Modelo de Poisson

Muchas evaluaciones de peligro sísmico se basan en el análisis de registros históricos de sismicidad, apoyándose en modelos en que la dependencia temporal de la sismicidad es considerada como sistemas de tipo poissoniano. Es decir, el número n de eventos de determinada magnitud que ocurre durante un tiempo t se distribuye con probabilidad:

$$p(n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad (2)$$

donde λ es el promedio del número de ocurrencias por unidad de tiempo. Para esta distribución el intervalo temporal T entre eventos se distribuye exponencialmente:

$$f(T) = \lambda e^{-\lambda T}. \quad (3)$$

Un comportamiento de acuerdo con una distribución de Poisson implica un proceso sin memoria (Ross, 1970; Nava y Espíndola, 1993); la ocurrencia del terremoto, no importa cuán grande sea, no cambia las probabilidades de eventos futuros. Aunque λ puede ser función del tiempo, su valor no depende, en principio, de la sismicidad pasada. Un modelo sin memoria es incompatible con el modelo del rebote elástico (Reid, 1910) y con la teoría del gap sísmico (Fedotov, 1965) explicados anteriormente.

El que el modelo poissoniano sea tan socorrido, puede explicarse por el hecho de que la sismicidad mundial o regional de sismos principales se adecúa a una distribución de Poisson debido a que es una superposición de varios sistemas, cada uno de los cuales es causal y con memoria y, en general, con ciclos y fases diferentes, de manera que el sistema compuesto resulta sin memoria.

De este modo, una estimación de peligro sísmico basada en una distribución de Poisson puede ser apropiada para estudios a nivel mundial, o bien para regiones afectadas por eventos originados en muchas subregiones pero, definitivamente, no es apropiado para el estudio de una zona sismogénica en particular y no tiene valor substancial para predicciones a corto plazo (Lomnitz y Nava, 1993).

II.2.2. Modelo de Hagiwara

Hagiwara (1975) propone un modelo estocástico para la ocurrencia de terremotos y la deformación horizontal que los acompaña. Este trabajo considera que el modelo propuesto por Utsu (1972) en el cuál la probabilidad de ocurrencia de un terremoto puede ser representada por un proceso de Markov en el cual los elementos de la matriz de

transición son funciones de distribución binomial con parámetro μ relacionado con la escala de deformación, es una buena simulación de la ocurrencia de los terremotos.

II.2.3. Modelos basados en la teoría del gap sísmico

Con base en la hipótesis del gap sísmico, si en cierta región donde se sabe que han ocurrido sismos grandes en el pasado, y que no han ocurrido durante un tiempo suficiente para almacenar de nuevo la energía necesaria para un nuevo gran sismo, entonces la probabilidad de que ocurra un gran sismo es mayor, lo anterior ha permitido la elaboración de mapas de peligro sísmico como los presentados por (McCann *et al.*, 1979; Nishenko 1991).

II.2.4. Métodos de reconocimiento de patrones

Existen otras técnicas estadísticas basadas en la detección de variaciones espacio-temporales de la sismicidad (técnicas de reconocimiento de patrones). Una de ellas ha recibido considerable atención, Keilis Borok *et al.*, (1990) presentaron una técnica para el reconocimiento de patrones, donde se dan pesos subjetivos a la actividad sísmica pasada con objeto de obtener una predicción. Sus algoritmos M8 y CN se utilizan para determinar el tiempo cuando la probabilidad de ocurrencia de grandes terremotos es mayor (TIP) de sus siglas en inglés.

La evidencia de que información pertinente a la localización y tiempo de ocurrencia de futuros sismos grandes puede ser encontrada en patrones de ocurrencia de sismos menores que los preceden va en aumento. Por ejemplo (Agnew y Jones, 1991) indican que en una escala de tiempo pequeña, algunos grandes terremotos son precedidos por sismos menores que ocurren en zonas muy cercanas a la fuente del gran evento. Por otra parte Rhoades y Evison (2005) presentan un método para el pronóstico el terremotos basado en la noción de que el incremento en la escala de precursores ocurre en todas las escalas en los procesos simogénicos.

III. METODOLOGÍA

En este capítulo se describe varios nuevos métodos para la determinación de peligro sísmico con base en el análisis estadístico de sismicidad reportada en catálogos, propuestos y explorados como parte de este trabajo de tesis. Cuatro de estos métodos resultaron exitosos y serán objeto de aplicación y futuro desarrollo; otros métodos no resultaron tan satisfactorios y se mencionan aquí porque, aun siendo resultados negativos, iluminan algunos aspectos del problema de cálculo de peligro y porque documentarlos puede, posiblemente, ahorrar trabajo a otros investigadores trabajando sobre el tema.

III.1 *Método Markoviano para la estimación de peligro sísmico*

Dado que el trabajo de esta tesis es elaboración de la metodología presentada por Herrera (2001) y Nava *et al.* (2005) para la determinación de peligro sísmico en un sistema, y, por tanto, utiliza los mismos conceptos y nomenclatura, en esta sección se describe brevemente dicha metodología.

III.1.1 *Sistema*

El método trabaja sobre un *sistema*, definido como un área geográfica que incluye R regiones sismogénicas. Dados un catalogo sísmico y un tiempo inicial, durante un intervalo de tiempo específico Δt , el estado de cada región, s_r , tiene uno de dos valores: 0 ó 1 que corresponden, respectivamente, a la ausencia o presencia de sismos con magnitud

mayor o igual a un valor umbral M_r . El estado total del *sistema* es una de las $S = 2^R$ combinaciones posibles de los estados regionales:

$$s = \sum_{r=0}^{R-1} 2^r s_r . \quad (4)$$

Se puede representar el estado del sistema en forma binaria por una palabra de R bits formada por la concatenación de los estados regionales, que va de $s = 0 = 0\dots 000$ hasta $s = S - 1 = 1\dots 111$, representación que permite visualizar fácilmente cuáles regiones son activas y cuáles no.

III.1.2. Premisas

El método original y, por tanto, las elaboraciones sobre éste fueron motivados por las siguientes premisas; pero cabe mencionar que los métodos **no** dependen de éstas y deben ser juzgados por su desempeño únicamente.

1. El estado de esfuerzos que determina la posible ocurrencia de sismos de magnitudes grandes en cada región del sistema, durante un intervalo de tiempo dado, está relacionado con la historia sísmica de ésta y las demás regiones. Sin embargo, la actividad sísmica de cada región no está directamente relacionada con alguna o algunas de las demás de manera que permita hacer predicciones deterministas de la actividad en dicha región con base en la actividad de las otras.
 2. Cada estado del sistema es resultado de toda la historia sísmica de las regiones.
- Por lo tanto, el próximo estado, que depende de la misma historia sísmica a

través del estado actual, está determinado probabilísticamente por éste. Esta premisa permite aplicar la metodología de Cadenas de Markov véase (apéndice A).

III.1.3. *Parámetros umbral*

En el apéndice C se presenta los criterios empleados para la determinación de los parámetros libres del sistema, (el área de estudio y sus regiones, la magnitud umbral y el tamaño del intervalo al que se aplicará cada evaluación de peligro). Aquí solamente mencionaremos que una elección preliminar de los parámetros se hace a manera de obtener definición y cobertura óptimas dentro de los límites de la longitud de catálogo y, en cierto grado, de su precisión; la elección final de los parámetros se hace en forma iterativa y es pragmáticamente determinada por la combinación de éstos que resulta en el mejor desempeño (cap IV).

III.1.4. *Método Directo*

Sustentado en las definiciones y premisas expuestas anteriormente (III.1.1 y III.1.2) el método, al que de aquí en adelante nos referiremos como *método directo*, se describe a continuación.

En este método, el estado del sistema es evaluado para cada intervalo Δt sucesivo, y de la sucesión de estados construimos la matriz de transición Θ , cuyos elementos θ_{ij} corresponden al número de veces que el estado $S_{n+1} = j$ ocurrió después del estado $S_n = i$.

También construimos el vector Ξ , donde cada componente $\xi_i = \sum_j \theta_{ij}$ corresponde al número total de transiciones originadas desde el estado $S_n = i$.

Tras procesar todos los datos del catálogo, construimos la matriz de probabilidades de transición P , con elementos

$$p_{ij} = \frac{\theta_{ij}}{\xi_i} \quad (5)$$

Con base en el estado del sistema observado durante un intervalo Δt dado, esta matriz da estimaciones de peligro para los posibles estados del sistema para el siguiente Δt .

III.2. Nuevos métodos

En esta sección son descritos nuevos métodos para la estimación de peligro sísmico, sugeridos por los resultados de la aplicación del *método directo* al área de Japón (véase cap. V) y desarrollados en el presente trabajo.

III.2.1 Métodos markovianos

Estos métodos están basados en la estimación de probabilidades de transición a partir del conteo directo de transiciones.

III.2.1.1 Método Markoviano cruzado

Dado que estamos interesados en estimaciones de peligro sísmico para sismos grandes, potencialmente peligrosos, y que la limitante principal del método directo es la relativa escasez de eventos de gran magnitud, decidimos explorar la posibilidad de que sismicidad con magnitud mayor o igual a cierta magnitud umbral, $M \geq M_r^m$, contenga información pertinente a la ocurrencia de terremotos con magnitud $M \geq M_r^M > M_r^m$. Llamamos a este método *cadena de Markov cruzadas*, ya que estima la probabilidad de ocurrencia de sismos con $M \geq M_r^M$ con base en la ocurrencia observada en un paso anterior de sismos con $M \geq M_r^m$.

En este método, para cada intervalo Δt se determina dos estados del sistema S_n^m y S_n^M (donde n indica el número del intervalo), con base en la ocurrencia o no ocurrencia de sismos con magnitud $M \geq M_r^m$ y $M \geq M_r^M$, respectivamente. De este modo, tras procesar todos los datos del catalogo tenemos dos historias paralelas de estados del sistema: $S^m = \{S_n^m; n=1, \dots, N\}$ y $S^M = \{S_n^M; n=1, \dots, N\}$, donde N es el número total de intervalos.

De estas dos historias de estados del sistema, construimos una matriz de transiciones Θ , con elementos θ_{ij} que corresponden al numero de veces que el estado $S_{n+1}^M = j$ ocurre después del estado $S_n^m = i$. Tras procesar todos los datos del catálogo, construimos la matriz de probabilidades de transición P , con elementos,

$$p_{ij} = \frac{\theta_{ij}}{\xi_i}, \quad (6)$$

donde $\xi_i = \sum_j \theta_{i,j}$ corresponde al número total de transiciones originadas desde el estado

$S_n^m = i$. El método directo es el caso particular $M_r^m = M_r^M$.

Aplicaremos este método a la misma base de datos utilizada por el método directo y por los modelos neuronales y compararemos los resultados.

III.2.1.2. Método de caminos preferenciales

Para probar si un sistema es verdaderamente markoviano, además de determinar si tiene memoria, es necesario determinar si la probabilidad de un estado depende únicamente del estado anterior. El método presentado aquí explora la hipótesis nula que consiste en la existencia de *camino preferenciales*, i.e. secuencias características de 2 o más transiciones que lleven a un estado dado.

De la historia de estados, se cuenta el número $n_{i,j,...k,m}$ de veces que una secuencia dada, que pasa por los estados $i, j, ..., k$, termina en cada estado m del sistema; por ejemplo $n_{9,3,1}$ es el número de veces que el sistema pasó del estado 9 al estado 3 y, de allí, al estado 1; $n_{7,9,3,1}$ es cuántas veces la misma trayectoria previa comenzó, un paso antes, en el estado 7, etc. La probabilidad de que la secuencia termine en el estado v es:

$$\Pr(v | i, j, ..., k) = \frac{n_{i,j,...k,v}}{N_{i,j,...k}}, \text{ donde}$$

$$N_{i,j,...k} = \sum_{m=0}^S n_{i,j,...k,m}.$$

Si $\Pr(v | i, j, ..., k)$ concuerda con $p_{i,j,...k,v} = p_{i,j} \cdot \dots \cdot p_{k,v}$ de la matriz de probabilidades de transición de Markov, entonces no hay caminos preferenciales que tengan mayor valor predictivo que la estimación markoviana de peligro.

III.2.2. Métodos Neuronales

La construcción de la matriz Markoviana de probabilidades de transición, (construida con base en la información contenida en el catálogo) puede considerarse como un proceso de aprendizaje de acuerdo con el siguiente esquema:

- a) Comenzamos con una matriz P nula, lo que significa que se considera que no hay sismicidad (una matriz que supusiera sismicidad uniformemente aleatoria tendría $p_{ij} = \frac{1}{s} \quad \forall i, j$); los renglones cuyos elementos suman cero no pueden ser interpretados en términos de probabilidad (sin observaciones es imposible la estimación de peligro).
- b) Si, tras observar la primera transición, digamos $k \rightarrow m$, evaluamos la matriz de probabilidad tendremos un renglón, el k , cuyo elemento $m = 1$ y cuya suma de elementos es 1 (ya que los demás elementos son cero) y puede ser considerado ya como una estimación de probabilidad (aunque se trate de una sola opción segura).
- c) Para la siguiente transición, digamos $m \rightarrow r$, tendremos una estimación probabilística para el renglón m , luego para el r , y así en adelante.
- d) Cuando comienzan a ocurrir transiciones desde un estado con otra u otras transiciones previamente observadas, se puede ver cómo cada nueva transición modifica o ratifica las probabilidades observadas antes y esto es un proceso de aprendizaje gradual. La observación de cada una de las transiciones, contribuye con peso de $(1/\text{número total de transiciones})$ a dicho aprendizaje. Queda sólo

mencionar el posible caso en que nunca se presente un estado dado; entonces la matriz tendrá un renglón y una columna iguales a cero y éstas pueden eliminarse, obteniéndose así una matriz de un orden menor, pero estocástica y perfectamente útil para nuestros propósitos).

Consideramos la posibilidad de aplicar otros métodos de aprendizaje al problema de evaluación de peligro sísmico; y las redes neuronales artificiales son candidatos naturales para este objetivo.

III.2.2.1 *Perceptrón simple*

El tipo más simple de red neuronal, conocido como perceptrón, trabaja con patrones de entrada-salida binarios, y su funcionamiento, se realiza en dos fases: una primera en la que se presenta a la red un conjunto de entradas y las correspondientes salidas que deseamos proporcione la red. En esta fase la red aprende la salida que debe dar para cada entrada, adaptando los valores de los pesos w_{ij} con base en la diferencia entre salida calculada y salida deseada. En la fase siguiente, la red es capaz de responder más acertadamente cuando se le vuelve a presentar los patrones de entrada (véase apéndice B); en caso de no haber inconsistencias o contradicciones en los datos y de contar con una serie suficientemente larga de pares entrada-salida, el perceptrón puede llegar a responder sin errores.

III.2.2.1.1. *Entrada y salida del perceptrón*

Las entradas y salidas deseadas del perceptrón considerado aquí corresponden a la historia de estados del sistema, observada a lo largo de todo el catálogo. Para cada intervalo n , cada par de estados sucesivos (S_n, S_{n+1}) corresponde, respectivamente, a una entrada y a la salida correspondiente que deseamos proporcione la red.

Cada estado está representado por palabras binarias cuyos S elementos son todos ceros excepto el elemento correspondiente al número del estado, el cual vale uno; por ejemplo, el estado 10 (actividad en las regiones 1 y 3) $s = 1010$ en esta representación es $z = 0000000000100000$. Esta representación, que llamaremos z para distinguirla de la s de cuatro bits, ortogonaliza los estados en S dimensiones lo que nos permite usar perceptrones. La función de activación empleada es la función identidad (*purelin*) $f(x) = x$; es decir, solamente transmite la entrada que recibe cada neurona, que es un promedio pesado de las entradas.

III.2.2.1.2 *Entrenamiento del perceptrón*

Con base en la historia de transiciones observada, se entrena al perceptrón según la ecuación B.2 (apéndice B), considerando diversos valores iniciales de los pesos w_{ij} y diferentes valores de amortiguamiento η . Para conservar la interpretación probabilística de la matriz W , usamos un tope *poslin* (hacemos cero los valores $w_{ij} < 0$) al aplicar las correcciones Δw_{ij} y, después, se normaliza la matriz (por renglones). La matriz W nos proporciona estimaciones de peligro para el siguiente intervalo temporal.

III.2.2.1.3. *Perceptrón simple con bias*

Como se mencionó anteriormente, utilizar la representación de palabras de S bits con un solo bit distinto de cero en el elemento correspondiente al número del estado, ortogonaliza, de hecho, ortonormaliza los estados en S dimensiones y, posiblemente, por ser las palabras explícitamente ortogonales, los bias tienden a ser nulos y su presencia no cambia mayormente los resultados; *i.e.* no necesitamos trabajar con bias.

III.2.2.2. *Perceptrón con información de magnitudes*

Utilizar solamente ausencia ó presencia de sismos descarta cualquier información que puedan contener las magnitudes en sí (aparte de ser mayores o iguales a la magnitud umbral). Para explorar la posibilidad de que la cantidad de energía liberada sea un factor que escoja entre las opciones dadas por las varias probabilidades, se procedió a codificar la *magnitud equivalente* M_{eq} (magnitud que corresponde al total de la energía sísmica liberada en todas las regiones) en una escala, digamos, del 1 al 3: donde 1, 2, y 3 representan baja, media, y alta, M_{eq} respectivamente. Por ejemplo, según este código, la palabra utilizada para representar el estado 3 con M_{eq} media es: 0 0 0 2 0 0 ... 0. Este código desnormalizaría los estados (que continuarían siendo ortogonales) y para evitar comportamiento lineal del sistema sería necesario utilizar bias (véase apéndice B). Con base en la historia de transiciones observada, se entrena al perceptrón según las ecuaciones B.2 y B.3, considerando varios valores iniciales de los pesos w_{ij} y, b_i y diferentes valores de amortiguamiento η . Para conservar la interpretación probabilística de la matriz W , usamos un tope *poslin*, al aplicar las correcciones Δw_{ij} y, después, se normaliza la matriz

(por renglones). La matriz W nos proporciona estimaciones de peligro para el siguiente intervalo temporal.

III.2.2.3. *Perceptrón cruzado*

En este método cada par de estados sucesivos (i, J) corresponde respectivamente, a la entrada y a la salida que deseamos proporcione la red; esto es, corresponden a los estados del sistema $z_n^m = i, z_{n+1}^M = J$ determinados con base en la ocurrencia o no ocurrencia de sismos con magnitud $M \geq M_r^m$ y $M \geq M_r^M$, respectivamente. El entrenamiento de este perceptrón se hace según lo explicado en la sección III.2.2.1.2.

III.2.2.4. *Perceptrón con historia*

III.2.2.4.1. *Perceptrón con entrada simultánea de 2 pasos*

Perceptrón consistente en S neuronas, con o sin bias, que es alimentado a cada paso con dos palabras: z_n , que representa el estado al paso n , y z_{n-1} que corresponde al estado previo, la salida deseada corresponde al estado subsecuente z_{n+1} . En este caso, la matriz W es una matriz de $2S \times S$ elementos y el peligro para el paso $n+1$ está dado por la normalización del vector resultante de la suma de los renglones correspondientes a los estados z_n y z_{n-1} (véase Figura 2).

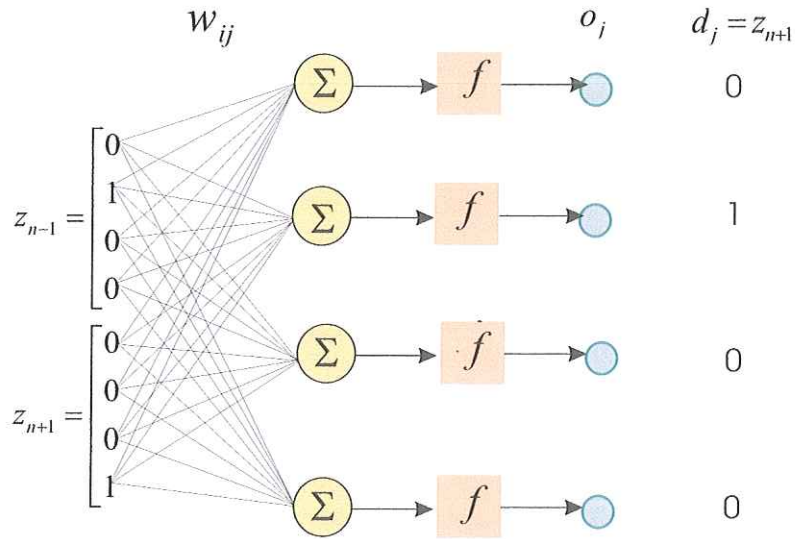


Figura 2. Ejemplo del perceptrón con entrada simultánea de dos pasos para un sistema con $S = 4$ posibles estados.

III.2.2.4.2. Perceptrón con entrada de 2 pasos codificados

Perceptrón consistente en S neuronas, con o sin bias, que es alimentado al paso n por la palabra $\tilde{z}_n = a z_n + b z_{n-1}$, donde z_n es el estado al paso n , z_{n-1} el estado al paso previo y a y b son constantes, tales que $a + b = 1$, que determinan qué tanta importancia se da al pasado con respecto al presente. En este caso, dado que la entrada no consiste de unos y ceros, el peligro para el paso $n + 1$ no corresponde a algún renglón de W , sino que es la salida $\tilde{z}_n W$ (para el perceptrón simple esta salida, corresponde a un renglón).

III.2.2.5. *Redes neuronales de 2 o más capas*

En una red neuronal de varias capas, la primera capa trabaja directamente combinando las entradas y las siguientes capas haciendo combinaciones de la salida de la capa previa. Para una entrada de tipo z (p.ej. las usadas para los perceptrones directo y cruzado) la salida de la primera capa es simplemente un renglón de W ; las siguientes capas no tienen información adicional que añadir al hacer combinaciones de elementos de dicho renglón, por lo que utilizar varias capas no produce mejoras significativas en el desempeño.

Falta explorar el desempeño de redes neuronales multicapa para entradas que incluyan más información, como pueden ser entradas con historia o incorporación de otras observables con potencial premonitor.

III.2.2.6. *Perceptrón por regiones*

Perceptrón consistente en una neurona por región, con o sin bias, que (a diferencia de lo descrito en III.2.2.1.1.) es alimentado con la palabra de cuatro bits correspondiente a un estado y cuya salida debe ser la palabra correspondiente al estado subsecuente. Este perceptrón es claramente inoperante dada la no-ortogonalidad de las entradas y salidas. Una red neuronal con más capas podría identificar el estado y, después, elaborar la salida; claramente no contaría con más información que el perceptrón simple y la mayor complejidad cancela la elegancia de tener entrada y salida correspondiente a la actividad por región.

IV. EVALUACIÓN DEL DESEMPEÑO DE LOS MÉTODOS

Evaluar el desempeño de diferentes métodos o modelos (diferentes combinaciones de parámetros para un método dado) de estimación de peligro, no es inmediato ya que, como veremos más adelante, hay que evaluar varios factores, a veces contradictorios, como confiabilidad, y utilidad.

IV.1. *Probabilidad de transición observada promedio*

Dado que el principal resultado de los métodos consiste en las probabilidades de ocurrencia de las posibles transiciones, una medida natural de su desempeño es la *probabilidad observada* promedio $\hat{p}$. Si antes de la n -ésima transición el sistema está en el estado i , las probabilidades p_{ij} constituyen la estimación de peligro para el subsecuente estado; si entonces ocurre el estado k , la *probabilidad observada* es $\hat{p}_n = p_{ik}$. Podemos decir entonces que la estimación de peligro fue buena si $\hat{p}_n$ fue grande; en cambio, un valor bajo de $\hat{p}_n$ implica que la estimación fue mala. De este modo,

$$\hat{p} \equiv \frac{\sum_{n=1}^N \hat{p}_n}{N}, \quad (7)$$

donde N es el número de transiciones observadas, es un buen estimador de la “calidad” de los pronósticos. Nótese que el valor de p_{ik} puede ir cambiando conforme aumenta el número de observaciones. Existe una probabilidad de referencia “natural” contra la cual se puede comparar los resultados, que es la de la distribución uniforme

$p_{ij}^u \equiv u = (S)^{-1} = \langle p_{ij} \rangle$ correspondiente al caso donde todos los estados tienen la misma probabilidad de ocurrir y pronosticar un estado es mera adivinación aleatoria.

Sin embargo, aunque es una buena estimación del desempeño, $\hat{p}$ no toma en cuenta que el objetivo principal de las estimaciones de peligro sísmico es pronosticar la ocurrencia de terremotos para que la sociedad pueda prepararse para ellos. Una estimación **útil**, debe tomar en cuenta factores como; el nivel de probabilidad de pronóstico (discutido más adelante), pronósticos múltiples (necesariamente contradictorios), falsas alarmas, pronósticos erróneos, etc. Una estimación dada puede dar muy buenos resultados para un factor, pero puede verse superada en otros factores por otras estimaciones, que sin embargo no resultan tan bien en el primer factor, por ello es difícil elegir el parámetro que dé la mejor estimación de peligro. A continuación discutimos algunas medidas cuantitativas de los factores mencionados.

IV.2 Pronósticos

IV.2.1. Pronósticos exitosos, falsas alarmas y eventos perdidos

Definimos un pronóstico como la declaración que un determinado estado dado tiene probabilidad alta de ocurrir, donde se define *probabilidad alta* como aquella probabilidad mayor que una probabilidad umbral:

$$p_x = f_x u, \quad (8)$$

donde el factor f_x , al que llamamos *factor de éxito*, expresa el umbral en términos de la probabilidad uniforme u . Un pronóstico es considerado *exitoso* cuando se realiza la transición correspondiente, y denotamos al número de pronósticos exitosos observados

como n_x . Cuando un pronóstico no es exitoso, entonces es una falsa alarma, y el número de falsas alarmas es contado como n_f . El número de transiciones para las cuales hubo pronóstico es n_{pred} y n_t es el número total de transiciones.

IV.2.2. Multiplicidad

Llamamos *multiplicidad*, m_i , al número de elementos mayores que p_x en el renglón i de la matriz de probabilidades de transición, i.e. al número de pronósticos para el estado inicial i . Cuando $m_i = 1$, entonces no hay problema en contar el resultado como un pronóstico exitoso o como una falsa alarma. Cuando $m_i > 1$, no es inmediato cómo contar los éxitos; dos posibles métodos son:

- a) Contando cada éxito como 1 (llamamos a la suma n_{x1}).
- b) Contando cada éxito como $1/m_i$ (llamamos a la suma n_{xm}). De esta manera, un valor pequeño de p_x no resulta en un absurdamente alto número de éxitos, necesariamente contradictorios.

Cuando $m_i = 0$, entonces no hay pronóstico (no hay éxito o falsa alarma) y decimos que el estado resultante de la transición es un *evento perdido*; el número de eventos perdidos es n_s .

Es conveniente tener $m_i \cong 1$ y, por ende $\langle m \rangle \approx 1$, porque esto significa tener un sólo pronóstico por transición, evitar tanto pronósticos múltiples (con sus implícitas falsas alarmas) como eventos perdidos. Tres estimaciones de la multiplicidad promedio son:

- a) Como la media $\langle m \rangle = \langle m_i \rangle$,
- b) Como la razón $\langle m \rangle = \frac{n_{x1}}{n_{xm}}$,
- c) Como un promedio ponderado de (i) y (ii).

IV.2.3. *Error regional*

El *error regional* para una transición dada es el número de regiones cuya actividad (ocurrencia o no ocurrencia de terremotos) fue pronosticada erróneamente. Para $m_i = 1$ el error puede tomar valores desde 0 (si el estado pronosticado fue correcto) hasta el número de regiones (si ocurre un estado completamente distinto). Para $m_i > 1$, dos formas de evaluar el error son:

- a) Sumando los errores de los diferentes pronósticos. Este error toma valores de 1 (es imposible que acierte a 2 estados distintos a la vez) a valores muy altos. (Suponiendo 8 predicciones fallidas con un error promedio de 2 por predicción, podría llegar en el peor de los casos a 16).
- b) Asignando un error de 1 a cada contradicción de los estados pronosticados, y contando las diferencias entre la actividad pronosticada y la actividad observada. Este error puede tomar valores desde 0 hasta el número de regiones.

Nosotros usaremos la forma (a) por ser más estricta.

IV.2.4. Utilidad de los pronósticos

Queda un ultimo factor por considerar, y es el de la *utilidad* de los pronósticos. Existen dos casos triviales: uno es cuando la magnitud umbral es mayor que la maxima magnitud observada; en este caso sólo hay transiciones $0 \rightarrow 0$, $p_{00} = 1$ y todas las demás probabilidades son nulas. El segundo caso trivial es cuando la magnitud umbral es muy pequeña y en consecuencia $p_{S-1, S-1} = 1$; este caso no es importante para nosotros, dado que estamos interesados únicamente en sismos grandes. Para ambos casos triviales las medidas descritas anteriormente resultarían en valores óptimos, pero los pronósticos serían completamente inútiles porque no proporcionan información alguna (los resultados son siempre conocidos de antemano).

Para evitar los casos triviales, o cercanos a lo trivial, es necesario penalizar la superabundancia de las transiciones que terminan en los estados 0 o $S-1$; haremos esto considerando la perdida de la información contenida de las probabilidades de transición correspondientes.

La información (en bits) contenida en un pronóstico con probabilidad p comúnmente se define como $I(p) = -\log_2(p)$ (c.f. Goldman, 1953); aquí, usaremos unidades de información *ubit* que asignan valor unitario a la información en la probabilidad uniforme, i.e. $I(u) = 1$ ubit. Si un total de N_t transiciones ha sido utilizado para evaluar las probabilidades de transición, entonces la probabilidad de cualquier transición al estado 0 es:

$$p_0 = \frac{\sum_i p_{i0}}{N_t}, \quad (9)$$

y, para $p_0 > u$. Si $I(p_0)$ (en ubits) es menor que uno, todas las probabilidades $\hat{p}_{i0}$ y pronósticos exitosos son multiplicados la información antes de ser contados.

IV.2.5. Factor de éxito

Todas las medidas relacionadas con pronósticos dependen de la elección del factor de éxito; un valor grande de f_x es deseable, porque queremos hacer pronósticos para altas probabilidades, ya que esto minimiza el número de falsas alarmas, pero un valor muy grande de f_x puede reducir demasiado el número de pronósticos (incrementando el número de eventos perdidos) y por lo tanto de éxitos, haciendo el modelo casi inútil. Un valor bajo de f_x tendrá como resultado alta multiplicidad y un gran número de falsas alarmas y error regional (ambos no deseables). Por lo que, debe encontrarse el valor óptimo de f_x para obtener el mejor desempeño de un determinado modelo (una determinada combinación de los parámetros del sistema, Δt , M_r^m , y M_r^M ; para los modelos directos $M_r^m = M_r^M = M_r$).

IV.3. Funciones calificadoras

Elegir el mejor método, modelo o f_x no es sencillo, porque usualmente hay una serie de intercambios entre factores deseados y no deseados. La inspección directa de los valores resultantes de las diferentes medidas de desempeño no es práctica para elegir el valor óptimo para un parámetro de un determinado modelo, y puede ser altamente confusa cuando se trata de comparar distintos métodos. Por lo tanto, para evaluar el desempeño

decidimos utilizar funciones calificadoras, i.e. funciones matemáticas cuyo valor incrementa con los resultados deseados y disminuye con los no deseados. En general estas funciones dan diferentes pesos a cada factor, acordes con sus relativos tamaño e importancia. Ajustar por tamaño es necesario, ya que menudo se compara cantidades con diferente orden de magnitud. Ponderar por importancia es muy subjetivo, pero refleja un consenso de deseabilidad para los diferentes factores.

Los parámetros que controlan el tamaño y peso de los distintos factores son ajustados empíricamente, a manera de obtener valores altos para predictores considerados buenos y bajas para los malos. Una vez ajustados, las funciones calificadoras son utilizadas en forma global para poder comparar los distintos métodos y modelos.

No hay una forma predeterminada para las funciones calificadoras; nosotros probamos varias funciones calificadoras distintas; lineal, no lineal, producto y mixtas. Consideramos que es mejor usar más de una función calificadora para juzgar los resultados. Los funciones que finalmente escogimos son:

$$d_0 = 2 + 5\hat{p} + \frac{10 n_x - n_f - e - n_s}{n_t} + (0.00001)f_x \quad (10)$$

$$d_1 = \frac{(400 + 0.00002f_x) \cdot \hat{p}^2 n_x^2}{(n_f + e + n_s)n_t} + 2. \quad (11)$$

Por supuesto que los valores absolutos de una función calificadora son bastante arbitrarios, y pueden ser cambiados escalando los parámetros (los valores escogidos aquí para ambas funciones calificadoras se escalaron con objeto de que los resultados observados pudieran ser graficados utilizando las mismas escalas); pero estamos más

interesados en los valores relativos, o en los valores óptimos relativos (utilizando siempre los mismos parámetros en la función, por supuesto).

Sin embargo, una idea de qué significan los valores absolutos (i.e., no relativos) puede ser obtenida mediante comparación con valores resultantes para los modelos sin memoria Poisson y uniforme (mera adivinación).

IV.4 Pronósticos y postnósticos

Para hacer un buen pronóstico, es necesario contar con suficientes observaciones **previas** que permitan hacer una estimación *robusta* de las probabilidades de transición; i.e. que éstas no cambien significativamente a la luz de una nueva observación. Desafortunadamente, la longitud de nuestro catálogo es apenas suficiente para hacer pronósticos, basados en estimaciones robustas, para unos cuantos (~20) de los últimos intervalos de tiempo; cantidad demasiado pequeña para ser estadísticamente significativa en la evaluación del desempeño de los métodos.

Por tanto, recurrimos como solución a este problema a hacer *postnósticos*, i.e. pronósticos basados en probabilidades de transición estimadas usando toda la información disponible del catálogo, incluyendo el intervalo para el cual se evalúa el peligro y, en general, intervalos posteriores a éste. Aunque estos postnósticos no son una medida de la capacidad del método para pronosticar, sí son una estimación grosera de qué podemos esperar del método, suponiendo estacionaridad en los procesos sismogénicos, y son una base para evaluar los resultados de los verdaderos pronósticos (hechos sólo con sismicidad previa).

V. APLICACIÓN Y RESULTADOS

A continuación se presenta los resultados de la aplicación de los métodos descritos en el capítulo III, al área de Japón, empleando los datos reportados por el *Internacional Seismological Centre*.

V.1. Localización de la zona de estudio

El archipiélago de Japón se localiza cerca de los márgenes de las placas Eurasia, Pacífico, Norte América y Filipinas, entre 25° y 55° N, y 125° y 155° E (Figura 3). A pesar de que son más de 3,000 islas e islotes los que conforman el archipiélago, las islas Honshu, Hokkaido, Kyushu y Shikoku, densamente pobladas, representan el mayor porcentaje de su superficie total.

V.2. Tectónica regional

A continuación se hace una breve descripción de la tectónica predominante en la región.

V.2.1. Procesos de subducción y arcos de isla

Las zonas de subducción del Pacífico son clasificadas, *grosso modo*, en dos tipos; el tipo Mariana, donde el postarco presenta centros de dispersión activos, y el tipo Chileno, donde el postarco es sometido a un régimen compresional de esfuerzos (Uyeda y Kanamori, 1979; Uyeda, 1982).

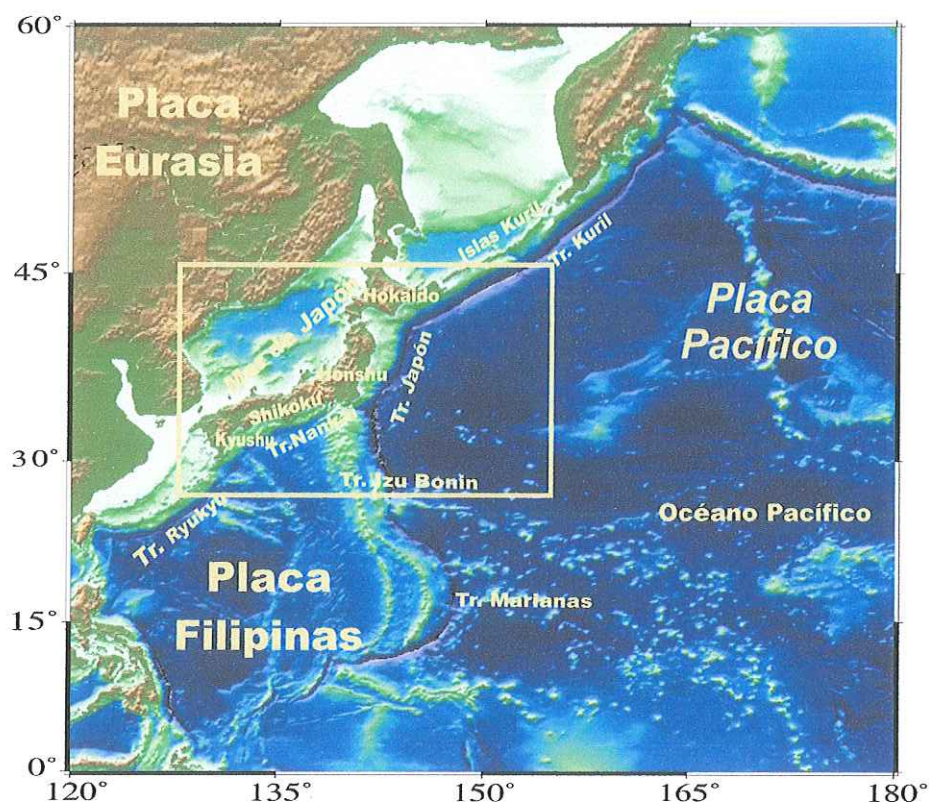


Figura 3. Mapa tectónico de la parte occidental del Océano Pacífico; el recuadro marca la zona de estudio.

El noroeste del Océano Pacífico presenta dos sistemas arco-trinchera: el primero está constituido por las trincheras Kamchatka-Kuril, Japón-Izu-Bonin y Mariana, el segundo lo integran las trincheras Nankai y Ryukyu (figura 1). Estos sistemas determinan tres zonas de subducción; en la primera, la placa Pacífico subduce en dirección NW-NNW bajo la placa euroasiática a lo largo de la trinchera Japón-Kuril-Kamchatka con velocidad $\sim 9\text{cm/a}$; en la segunda, la misma placa Pacífico subduce en dirección aproximadamente W-WSW bajo la placa de Filipinas en la trinchera Izu-Bonin-Marianas con velocidad $\sim 3.8\text{cm/a}$; finalmente, en la tercera de estas zonas de subducción la placa de Filipinas

subduce bajo la placa Euroasiática a lo largo de la trinchera Ryukyu-Nankai con velocidad de 3-5cm/a (Seno *et. al.*, 1993; Nava y Espíndola, 1993).

Estas tres zonas de subducción convergen en un punto triple localizado al Sureste de Tokio, ésta es la única unión de este tipo en el planeta (McKenzie y Morgan, 1969).

V.2.2. Fallas activas

En Japón y sus alrededores son conocida más de 1,500 fallas activas con razones de deslizamiento muy variadas (Matsuda y Kinugasa, 1991). La figura 2 muestra la distribución de estas fallas que en su mayoría son de longitud menor de 100 km. Sólo la línea tectónica media (MTL) en el occidente de Japón y la línea tectónica Itoigawa-Shizuoka (IS) en Japón central son de longitud mayor de 100 km. La MTL en la isla de Shikoku es la falla más activa de la región y tiene velocidad de deslizamiento de 5 a 10 mm/a (Okada, 1980).

En la parte noreste de Japón predominan fallas inversas, con dirección N-S ó NNE-SSW, paralelas a la trinchera del Japón. Las partes central y occidental de Japón se caracterizan por tener una densa distribución de fallas inversas y fallas laterales. La densidad de fallas activas en esta zona no sólo es la más alta en Japón, sino probablemente la más alta en la región del Cinturón Circumpacífico (Matsuda y Kinugasa, 1991).

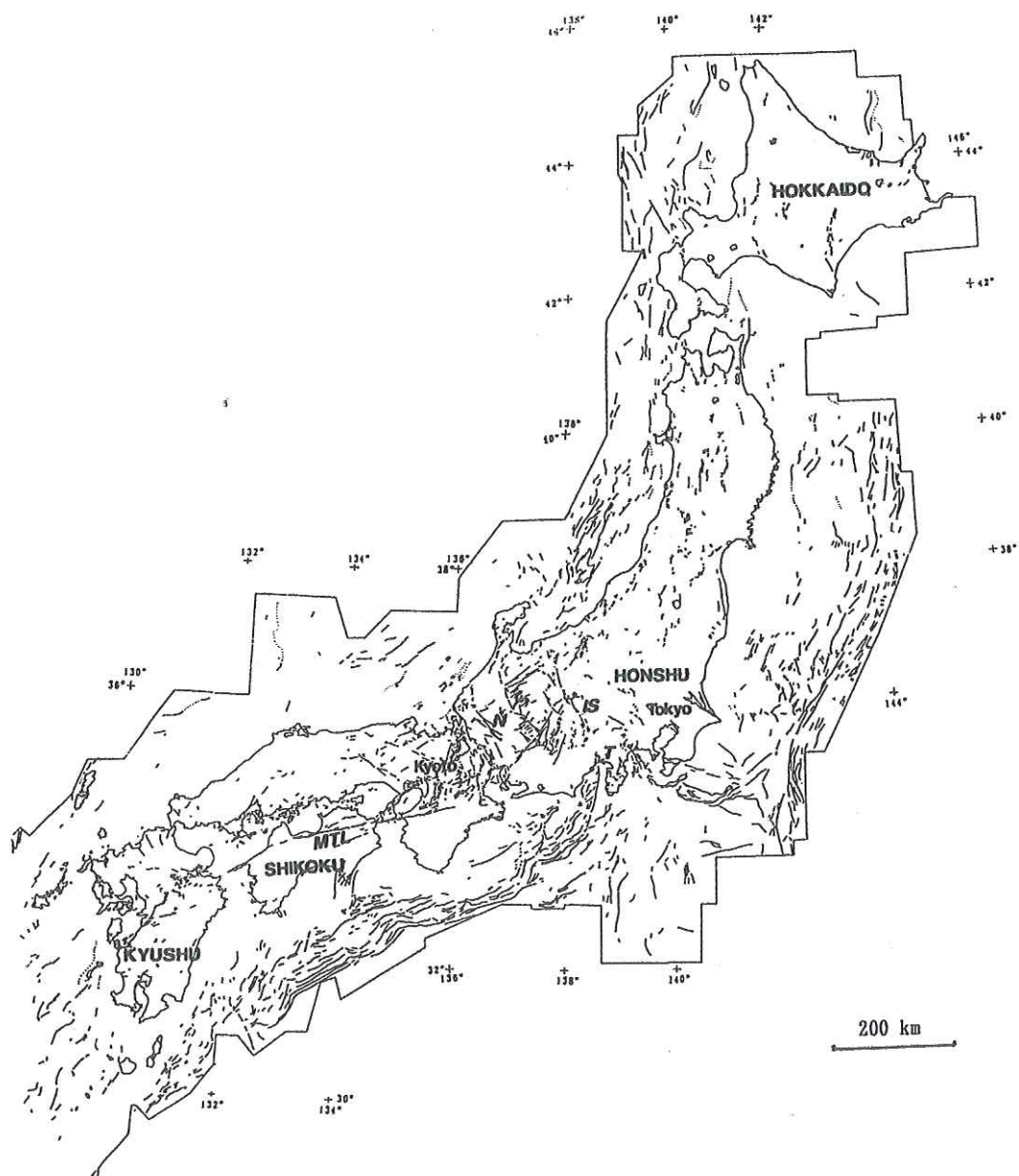


Figura 4. Distribución de fallas activas en y al alrededor de las islas japonesas. Tomada de Matsuda y Kinugasa, 1991.

V.3. Sismicidad

El noroeste del Pacífico es una de las zonas sísmicas más activas del mundo. Dicha actividad se debe principalmente a la interacción entre las placas Pacífico, Eurasia y Filipinas. La mayoría de los sismos, particularmente los de magnitud $M > 8$, ocurren a lo largo de las trincheras Japón, Sagami y Nankai (Wenosky, *et. al.*, 1984). Todos los arcos de la región: Kuril, Japón, Izu Bonnin, Ryukyu, Mariana, y Nankai definen claramente zonas de Benioff (Uyeda, 1991; Kanamori, 1977).

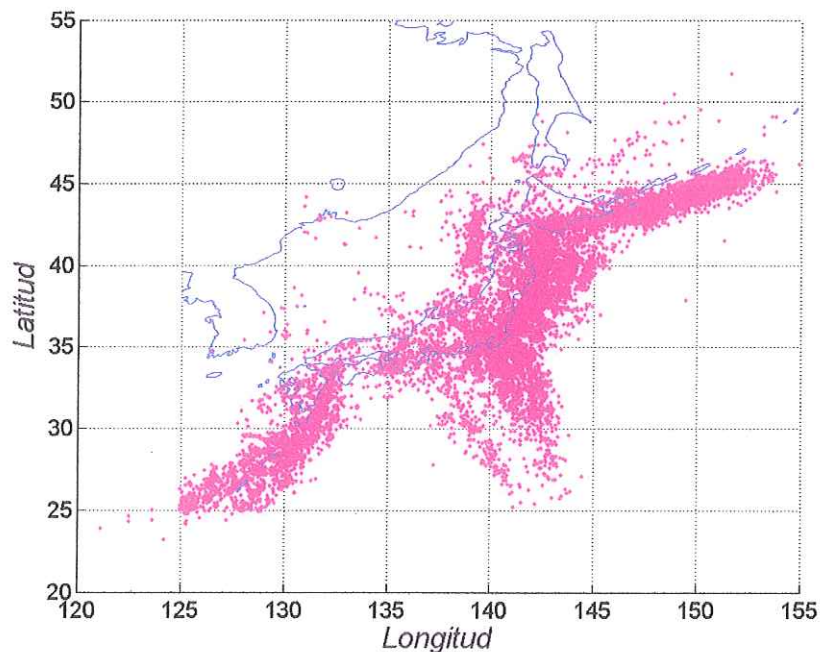


Figura 5. Sismicidad del Japón reportada por el ISC (1964-2002).

V.4. Descripción de la base de datos

El catálogo sísmico empleado para aplicación de los métodos propuestos en el presente trabajo al área de Japón, es el del *International Seismological Centre* (ISC). Los datos empleados por el ISC provienen de 3,000 estaciones que cubren las zonas sismogénicas del planeta. Esta base de datos incluye sismos localizados entre 25° y 55° N, y 125° y 155° E, con magnitudes que van de 3.0 a 7.9, ocurridos entre el 1 de Enero de 1964 y el 31 de Mayo de 2002.

V.5. Aplicación de los métodos a Japón

A continuación se presenta los resultados de la aplicación de los métodos descritos en el capítulo III al área de Japón.

V.5.1 Elección de los parámetros umbral

V.5.1.1 Parámetro umbral de espacio

El sistema fue elegido por tener varias regiones sismogénicas, alta sismicidad y por contar con un buen catálogo. La figura 4 muestra el sistema formado por cuatro regiones escogidas según se explicó en (C.1). La región 0, *Islas Kuriles*, contiene los sismos generados por la subducción de la placa Pacífico bajo la placa Euroasiática desde el extremo SW de la trinchera Kuril hasta los 155°E de longitud. La región 1, *Japón Central*, contiene sismos asociados con la subducción en la trinchera de Japón. La región 2, *Japón SE*, agrupa la actividad debida a la subducción en las trincheras Nankai e Izu-Bonin. La región 3, *Islas Ryukyu*, contiene la sismicidad asociada con la subducción en la trinchera del mismo nombre.

Dado que es muy importante identificar correctamente la región en la que ocurre cada sismo, los límites precisos de las regiones están separados por bandas de baja concentración de epicentros, de manera que ignorar la sismicidad en ellas no causara un error importante en el método.

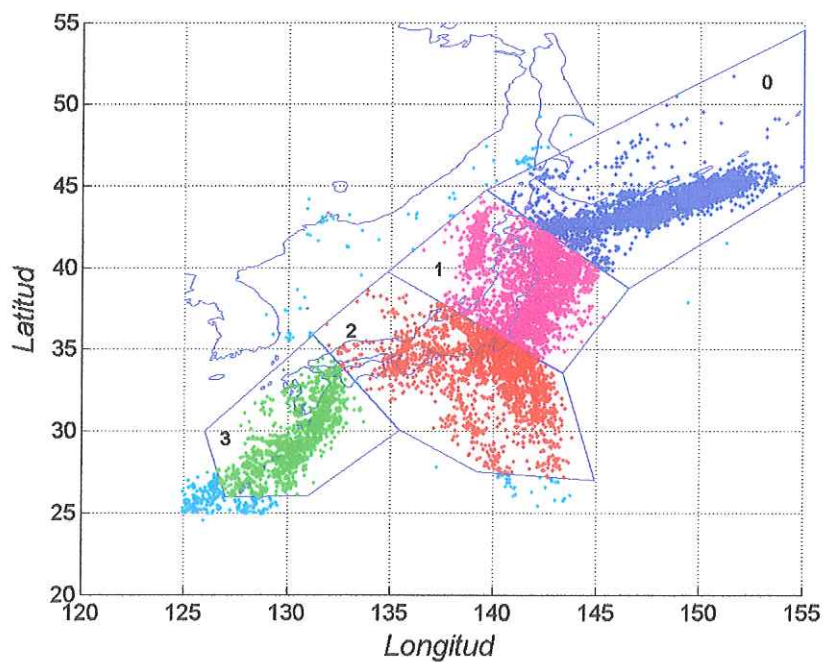


Figura 6. El sistema y sus regiones identificadas por los índices correspondientes.

V.5.1.2. *Umbrales de magnitud*

Dado que las estimaciones de peligro sísmico están relacionadas con la ocurrencia de sismos grandes, posiblemente destructivos, en este trabajo consideramos como magnitud mínima de interés para cada región $M_r \geq 6.1$ (elegida con base en los resultados obtenidos de la evaluación del desempeño de los métodos, presentados más adelante). La elección preliminar de este valor se hizo según lo explicado en (C.2). La Tabla I muestra el número total de sismos con $M \geq M_r$.

Tabla I. Parámetro umbral de magnitud y número de sismos con $M \geq M_r = 6.1$ para cada región, durante el periodo 1964-2002.

Región	No. Sismos
0. Islas Kuriles	110
1. Japón Central	107
2. Japón Sureste	88
3. Islas Ryukyu	31

V.5.1.3. *Umbral de tiempo*

El catálogo reportado por el ISC de 1964 a 2002 fue analizado usando varios Δt ; finalmente, $\Delta t = 0.10$ años fue escogido porque para la magnitud umbral seleccionada proporcionó los mejores resultados de acuerdo con lo explicado en la sección C.3.

V.5.2. Estados del Sistema

Con $R = 4$ regiones, resultan $S = 2^4 = 16$ estados posibles, mostrados en la tabla II.

Tabla II. Estados del sistema

Estados	Regiones			
	3	2	1	0
0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
5	0	1	0	1
6	0	1	1	0
7	0	1	1	1
8	1	0	0	0
9	1	0	0	1
10	1	0	1	0
11	1	0	1	1
12	1	1	0	0
13	1	1	0	1
14	1	1	1	0
15	1	1	1	1

V.6. Resultados de la aplicación de los métodos al área de Japón

Dado que, los métodos para la estimación de peligro sísmico propuestos aquí, tienen filosofías distintas, dividimos esta sección en dos partes.

V.6.1 Resultados I

En esta sección se presenta los resultados de la primera parte de esta investigación, que corresponde a los resultados de la evaluación del desempeño de los métodos Markovianos (directo y cruzado) de conteo. La evaluación del desempeño del método directo obedece a que se definió nuevos criterios de evaluación, y dado que el método directo constituye la base sobre la cuál se fundamenta el presente trabajo, los resultados del desempeño del método directo son un buen punto de referencia para evaluar el desempeño de los nuevos métodos.

Los resultados mostrados fueron obtenidos para las regiones especificadas en V.5.1.1, con base en todos los datos del catálogo ISC de 1964 a 2002 (incluye la actualización del catálogo más reciente) y $\Delta t = 0.10$ años, que resultan en un total de 384 transiciones. Para evaluar el desempeño de *postdicciones* fueron utilizadas las 384 transiciones; para el de *predicciones* fueron utilizadas las 20 transiciones más recientes.

Para cada modelo, *i.e.* cada magnitud umbral del método directo o combinación de magnitudes umbral del método cruzado, se determinó el valor óptimo del factor de éxito f_x , que es el que resulta en los máximos valores de las funciones calificadoras para el modelo. Estos valores máximos de las funciones calificadoras son los usados para comparar el desempeño de los distintos modelos y métodos y son los que se muestran a continuación.

V.6.1.1. Evaluación del desempeño del método directo

La figura 7 muestra el desempeño del método directo como función de la magnitud umbral M_r . Para postnósticos (izquierda) los mejores resultados son para $M_r = 6.1$; los correspondientes valores del factor de éxito y de las funciones calificadoras son mostrados en la Tabla III. La última columna de la Tabla III muestra la probabilidad de obtener n_x éxitos en n_t transiciones al azar como proceso de Bernoulli con probabilidad $p = u$; como puede verse, los resultados del método directo tienen una probabilidad muy pequeña de ser obtenidos al azar.

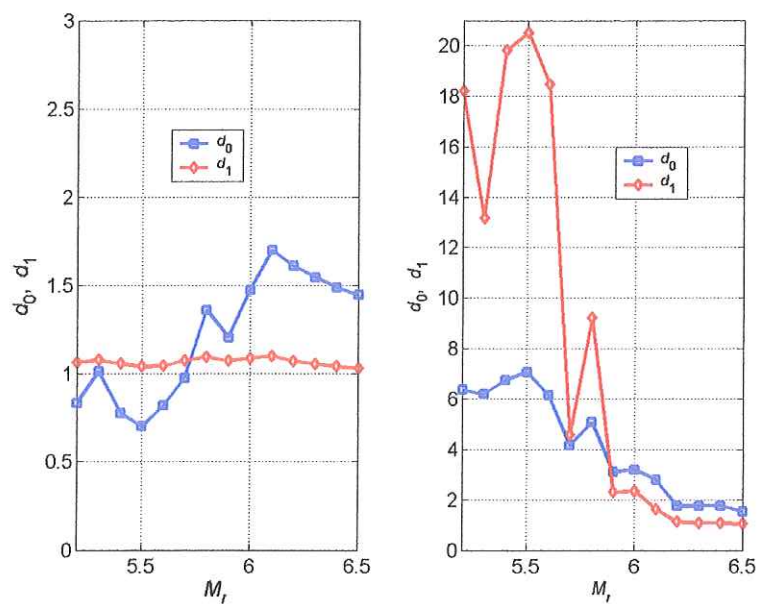


Figura 7. Valores de las funciones calificadoras del método directo en función de M_r . Izquierda *postnósticos*. Derecha *pronósticos*.

Tabla III. Resultados de la evaluación de postnósticos. n_t es el número de transiciones, f_x es el factor de éxito, n_x es el número de éxitos, p^b es la probabilidad binomial de Bernoulli de observar n_x éxitos en n_t transiciones con una probabilidad u .

Método	M_r	n_t	f_x	n_x	d_0	d_1	p^b
Directo	$M_r = 6.1$	384	6.2	49	1.699	1.100	$1.2 \cdot 10^{-6}$
Directo	$M_r = 6.2$	384	7.5	42	1.610	1.071	$1.7 \cdot 10^{-4}$
Directo	$M_r = 6.3$	384	5.5	39	1.547	1.059	$9.6 \cdot 10^{-4}$
Directo	$M_r = 6.4$	384	6.2	38	1.489	1.041	$4.2 \cdot 10^{-3}$
Directo	$M_r = 6.5$	384	5.0	32	1.445	1.030	$2.0 \cdot 10^{-2}$
Cruzado	$M_r^M = 6.1$ & $M_r^m = 5.7$	384	5.0	51	1.661	1.095	$2.3 \cdot 10^{-7}$
Cruzado	$M_r^M = 6.2$ & $M_r^m = 5.7$	384	5.0	51	1.717	1.100	$2.3 \cdot 10^{-7}$
Cruzado	$M_r^M = 6.3$ & $M_r^m = 5.7$	384	6.2	43	1.680	1.077	$8.9 \cdot 10^{-5}$
Cruzado	$M_r^M = 6.4$ & $M_r^m = 5.8$	384	7	37	1.607	1.058	$2.7 \cdot 10^{-3}$
Cruzado	$M_r^M = 6.5$ & $M_r^m = 6.1$	384	6.5	33	1.525	1.039	$1.4 \cdot 10^{-2}$

Para pronósticos (Figura 7, derecha) las funciones calificadoras tienen un máximo local para $M_r = 5.8$ ($f_x = 7.5$) pero presentan valores muy grandes para magnitudes pequeñas. Dado que 20 muestras no constituyen un muestreo estadísticamente confiable y que en la práctica son de interés sólo los pronósticos de sismos con magnitudes mayores, futuros pronósticos se harían en este momento a partir del modelo con $M_r = 6.1$ (Tabla IV), que corresponde al modelo para el que el método directo tuvo su mejor desempeño para postnósticos.

Tabla IV. Resultados de la evaluación de pronósticos. Cantidades como en la tabla III.

Método	M_r	n_t	f_x	n_x	d_0	d_1	p^b
Directo	$M_r = 6.1$	20	7.5	4	2.794	1.632	$2.6 \cdot 10^{-2}$
Directo	$M_r = 6.2$	20	8.0	2	1.758	1.125	$2.3 \cdot 10^{-1}$
Directo	$M_r = 6.3$	20	8.0	2	1.752	1.073	$2.3 \cdot 10^{-1}$
Directo	$M_r = 6.4$	20	8.0	2	1.752	1.073	$2.3 \cdot 10^{-1}$
Directo	$M_r = 6.5$	20	3.0	2	1.5018	1.040	$2.3 \cdot 10^{-1}$
Cruzado	$M_r^M = 6.1$ & $M_r^m = 5.5$	20	7.5	6	4.680	6.686	$9.4 \cdot 10^{-4}$
Cruzado	$M_r^M = 6.2$ & $M_r^m = 5.5$	20	5	6	4.420	5.804	$9.4 \cdot 10^{-4}$
Cruzado	$M_r^M = 6.3$ & $M_r^m = 5.8$	20	7.5	4	3.327	2.349	$2.6 \cdot 10^{-2}$
Cruzado	$M_r^M = 6.4$ & $M_r^m = 5.8$	20	7.5	4	3.327	2.349	$2.6 \cdot 10^{-2}$
Cruzado	$M_r^M = 6.5$ & $M_r^m = 5.8$	20	7.5	4	3.457	2.995	$2.6 \cdot 10^{-2}$

V.6.1.2. Evaluación del desempeño del método cruzado

Dado que el método directo tuvo su mejor desempeño para $M_r = 6.1$, se decidió explorar el desempeño del método cruzado para $M_r^M = 6.1$. Este método tuvo su mejor desempeño para $M_r^m = 5.7$ y, aunque no mejoró los valores de las funciones calificadoras observados para el método directo, sí mejora un poco la cantidad de éxitos (véase Tabla III). Como se verá más adelante, el método cruzado obtiene mejoras importantes para magnitudes umbral mayores.

La figura 8 muestra los valores de las funciones calificadoras del método cruzado para $M_r^M = 6.2$, como función de M_r^m ; las líneas discontinuas horizontales indican los valores óptimos observados para el método directo ($M_r = 6.2$). Puede apreciarse claramente que el desempeño del método cruzado para postnósticos es mejor que el del método directo; ambas funciones calificadoras muestran mejoras para $M_r^m = 5.6, 5.7, 5.8, 6.0$ y 6.1 , con el máximo valor en 5.7 . El desempeño del método cruzado para la combinación $M_r^M = 6.2$ y $M_r^m = 5.7$ es aún mejor que el mejor del método directo (Tabla III).

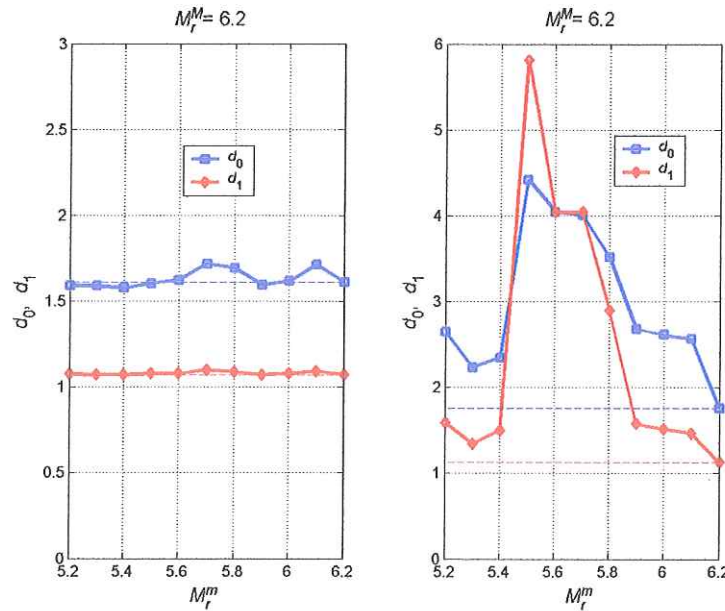


Figura 8. Valores de las funciones calificadoras del método *cruzado* en función de M_r^m . Izquierda *postnósticos*. Derecha *pronósticos*.

La figura 9 muestra los valores codificados en escala de color de las matrices de probabilidad de transición P obtenidas por el método directo para $M_r = 6.2$ (izquierda), y para el método cruzado con $M_r^m = 5.7$ y $M_r^M = 6.2$ (derecha). Puede verse que, aunque ambas matrices tienen regiones parecidas donde predominan valores bajos, sus altos ocurren para transiciones distintas; las diferencias entre ellas son debidas a la información aportada por los sismos con $M_r^m \leq M < M_r^M$.

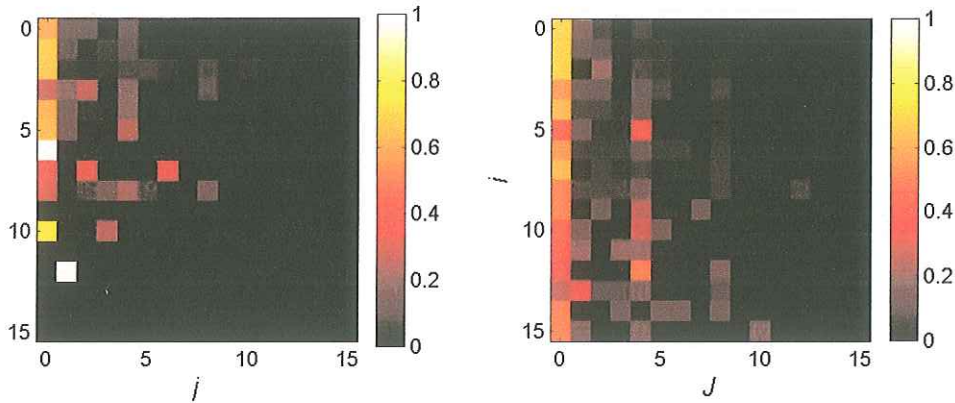


Figura 9. Representación de los valores de las matrices de probabilidad de transición codificados según la banda a la derecha. Izquierda: Método directo, para $M_r = 6.2$. Derecha: Método cruzado para $M_r^m = 5.7$ y $M_r^M = 6.2$

De la figura 8 (derecha) puede verse que el desempeño del método cruzado para pronósticos con $M_r^M = 6.2$ resulta, para todos los valores de M_r^m considerados, mucho mejor que el del método directo con $M_r = 6.2$; aunque es necesario insistir en que el muestreo de pronósticos es escaso.

Fue posible aplicar ambos métodos, directo y cruzado, hasta $M_r = M_r^M = 6.5$; la figura 10 muestra el comportamiento de las funciones calificadoras como función de M_r^m para este caso límite.

Aunque los valores máximos de las funciones calificadoras para el método cruzado ocurren para $M_r^M = 6.2$, para magnitudes grandes el desempeño del método cruzado es consistentemente mejor que el del método directo, tanto para postnósticos como para pronósticos, como puede apreciarse en las tablas III y IV.

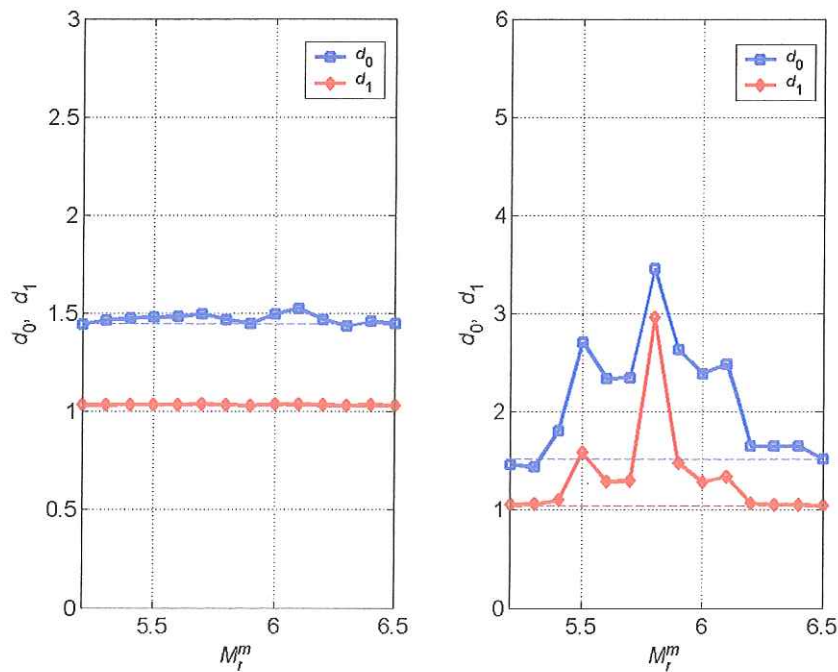


Figura 10. Valores de las funciones calificadoras del método *cruzado* para $M_r^M = 6.5$ en función de M_r^m . Izquierda *postnósticos*. Derecha *pronósticos*.

V.6.2. Resultados II

A continuación se presenta los resultados de la aplicación de los métodos neuronales al área de Japón.

Con objeto de poder comparar directamente los resultados obtenidos por los distintos métodos propuestos aquí, para el entrenamiento de los perceptrones fueron utilizados los mismos parámetros umbral usados por los métodos Markovianos, *i.e* mismas regiones y magnitudes umbral, y $\Delta t = 0.10$ años. Los perceptrones (*simple* y *cruzado*) fueron entrenados según la ecuación B.2 (véase apéndice B), usando diferentes valores iniciales para los pesos w_{ij} , así como valores varios de amortiguamiento η . Los renglones de la matriz W fueron normalizados a manera de obtener una matriz estocástica susceptible de ser interpretada de manera probabilística, según se explicó en la sección III.2.2.1.2. Para aprendizaje en redes es posible y válido “reciclar” la información del catálogo. En la práctica, reciclar más de 5 veces no cambia los resultados.

Al igual que para los modelos Markovianos, para cada modelo se determinó el valor óptimo del factor de éxito f_x , que es el que da los máximos valores de las funciones calificadoras para el modelo. Estos valores óptimos de las funciones calificadoras son los usados para comparar el desempeño de los distintos modelos y métodos que se muestran a continuación.

V.6.2.1. Evaluación del desempeño del perceptrón simple

Para el caso particular, de función de activación lineal, pesos iniciales muy pequeños (casi nulos) y fuerte amortiguamiento, la matriz de pesos resultante del

aprendizaje del perceptrón es, casi igual a la de probabilidades de transición del método directo (Figura 10). La máxima similitud se obtiene para valores iniciales de $w_{ij} = 0.00001 \quad \forall i, j$ y $\eta = 0.002604 (=384^{-1})$, con máxima diferencia absoluta entre valores correspondientes $\delta_{\max} = \text{Max}(|w_{ij} - p_{ij}|; i, j = 0, \dots, 15) = 0.0541$ y diferencia absoluta promedio $\langle \delta \rangle = \frac{1}{16^2} \sum_{i,j=0}^{15} |w_{ij} - p_{ij}| = 0.0011$.

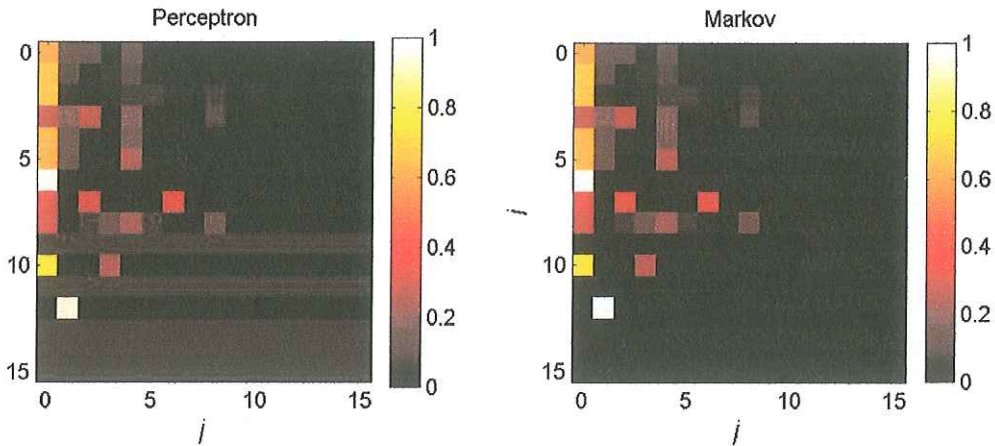


Figura 11. Valores de los elementos de las matrices de probabilidades de transición codificados según las bandas mostradas a la derecha para $M_r = 6.1$.

La similitud entre los resultados indica que la misma estimación probabilística puede obtenerse por dos métodos con filosofías diferentes; dado que el tratamiento hecho con el perceptrón de activación lineal equivale a realizar una estimación de mínimos cuadrados (Bishop, 1995), resulta particularmente interesante que el perceptrón dé valores similares a los de un modelo que simplemente contabiliza las ocurrencias.

Naturalmente, el desempeño de este perceptrón es idéntico al del método directo para pronosticar el estado de sismicidad del sistema, que con base en los resultados de las funciones calificadoras presentados en la sección (V.6.1.1) esta combinación de valores iniciales y constante de amortiguamiento **no** es la que resulta en resultados óptimos para el perceptrón. Estos resultados fueron reportados en Herrera y Nava (2004).

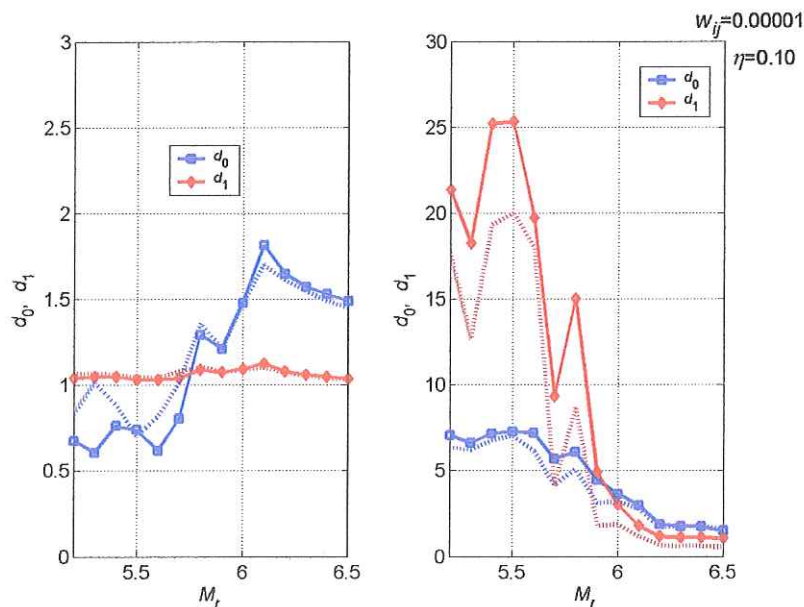


Figura 12. Valores de las funciones calificadoras del perceptrón *simple* en función de la magnitud umbral M_r . Izquierda *postnósticos*. Derecha *pronósticos*. Las líneas discontinuas representan valores de las funciones calificadoras observados por el método markoviano directo.

La figura 11 muestra el desempeño del perceptrón *simple* como función de la magnitud umbral M_r , para el caso particular de función de activación lineal $f(x) = x$, pesos iniciales $w_{ij} = 0.00001 \quad \forall i, j$ y amortiguamiento $\eta = 0.10$. Para postnósticos (izquierda) las funciones calificadoras indican que el desempeño del perceptrón simple es

mejor que el del método markoviano directo para distintos valores de M_r ; particularmente en los valores de d_0 , que muestran mejoras para $M_r = 6.1, 6.2, 6.3, 6.4$ y 6.5 , con el máximo valor para $M_r = 6.1$; los valores del factor de éxito y de las funciones calificadoras para este máximo son mostrados en la tabla V.

Tabla V. Resultados de la evaluación de postnósticos de ambos perceptrones para altas magnitudes. n_t es el número de transiciones, f_x es el factor de éxito, n_x es el número de éxitos, p^b es la probabilidad binomial de Bernoulli de observar n_x éxitos en n_t transiciones, cada una con probabilidad $p = u$.

Perceptrón	M_r	n_t	f_x	n_x	d_0	d_1	p^b
Directo	$M_r = 6.1$	384	5.0	54	1.814	1.124	$1.7 \cdot 10^{-8}$
Directo	$M_r = 6.2$	384	6.2	43	1.647	1.077	$8.9 \cdot 10^{-5}$
Directo	$M_r = 6.3$	384	6.0	39	1.569	1.059	$9.6 \cdot 10^{-4}$
Directo	$M_r = 6.4$	384	6.5	37	1.527	1.047	$2.7 \cdot 10^{-3}$
Directo	$M_r = 6.5$	384	7.5	32	1.487	1.035	$2.0 \cdot 10^{-3}$
Cruzado	$M_r^M = 6.1$ & $M_r^m = 5.7$	384	4.5	51	1.670	1.098	$2.3 \cdot 10^{-7}$
Cruzado	$M_r^M = 6.2$ & $M_r^m = 6.1$	384	5.0	49	1.816	1.112	$1.2 \cdot 10^{-6}$
Cruzado	$M_r^M = 6.3$ & $M_r^m = 5.7$	384	6.5	44	1.733	1.084	$4.6 \cdot 10^{-5}$
Cruzado	$M_r^M = 6.4$ & $M_r^m = 5.8$	384	6.2	39	1.654	1.063	$9.6 \cdot 10^{-4}$
Cruzado	$M_r^M = 6.5$ & $M_r^m = 6.1$	384	6.5	33	1.554	1.043	$1.4 \cdot 10^{-2}$

Para pronósticos (Figura 11, derecha) el desempeño del perceptrón resultó mejor que el del método markoviano directo para todas las magnitudes M_r consideradas en este análisis; los valores de ambas funciones calificadoras son mayores que los observados por el método directo (Figura 11, líneas discontinuas). Si bien el desempeño del perceptrón *simple* es mejor que el del método directo en todos los aspectos, son particularmente buenas

las mejoras observadas para grandes magnitudes, dado que en la práctica sólo son de interés las estimaciones de peligro para sismos grandes.

Tabla VI. Resultados de la evaluación de pronósticos de los perceptrones para altas magnitudes. Cantidades como en la tabla V.

Método	M_r	n_t	f_x	n_x	d_0	d_1	p^b
Directo	$M_r = 6.1$	20	6.0	4	2.980	1.769	$2.6 \cdot 10^{-2}$
Directo	$M_r = 6.2$	20	6.0	3	1.864	1.130	$9.3 \cdot 10^{-2}$
Directo	$M_r = 6.3$	20	9.0	2	1.737	1.069	$2.3 \cdot 10^{-1}$
Directo	$M_r = 6.4$	20	9.0	2	1.737	1.069	$2.3 \cdot 10^{-1}$
Directo	$M_r = 6.5$	20	4.0	2	1.492	1.035	$2.3 \cdot 10^{-1}$
Cruzado	$M_r^M = 6.1 \text{ \& } M_r^m = 5.5$	20	6.0	6	4.469	5.194	$9.4 \cdot 10^{-4}$
Cruzado	$M_r^M = 6.2 \text{ \& } M_r^m = 5.5$	20	6.0	7	4.407	5.715	$9.4 \cdot 10^{-4}$
Cruzado	$M_r^M = 6.3 \text{ \& } M_r^m = 5.8$	20	9	4	3.520	3.092	$2.6 \cdot 10^{-2}$
Cruzado	$M_r^M = 6.4 \text{ \& } M_r^m = 5.8$	20	9	4	3.520	3.092	$2.6 \cdot 10^{-2}$
Cruzado	$M_r^M = 6.5 \text{ \& } M_r^m = 5.8$	20	9	4	3.598	4.185	$2.6 \cdot 10^{-2}$

V.6.2.1. Evaluación del desempeño del perceptrón cruzado

Dado que el perceptrón directo tuvo su mejor desempeño para $M_r = 6.1$, se decidió explorar el desempeño del perceptrón *cruzado* para magnitud $M_r^M = 6.1$. Análogamente a lo ocurrido para el mismo caso en los métodos markovianos de conteo, el método cruzado no mejoró los valores de las funciones calificadoras observados para el perceptrón simple; podemos especular que cuando la información a una magnitud umbral dada es suficiente para un buen desempeño de los métodos directos, las magnitudes menores no contribuyen una cantidad significativa de información. Sin embargo, como se verá más adelante, el

perceptrón cruzado sí mejora significativamente los resultados para magnitudes umbral mayores, cuando el método directo ya no es tan eficiente.

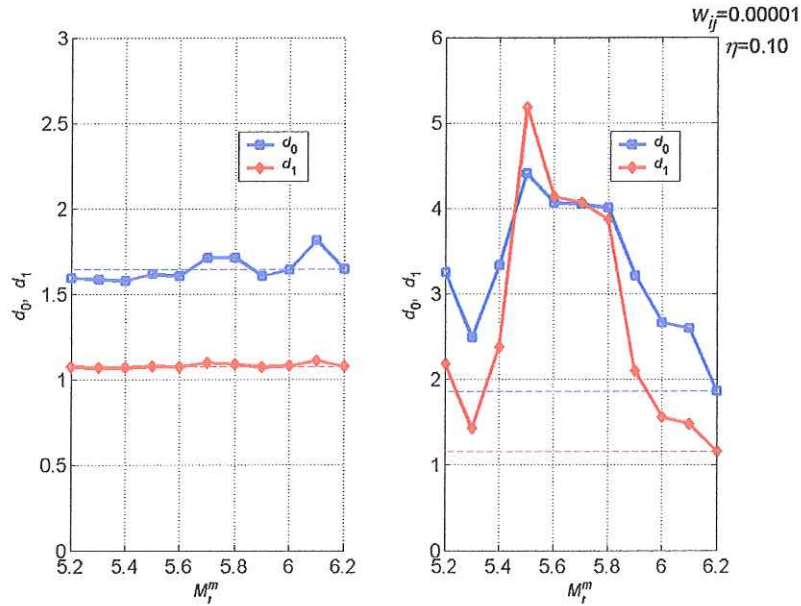


Figura 13. Valores de las funciones calificadoras del perceptrón cruzado en función de la magnitud umbral. Izquierda postnósticos. Derecha pronósticos. Las líneas discontinuas representan valores de las funciones calificadoras observados por el método directo para $M_r = 6.2$.

La figura 12 muestra los valores de las funciones calificadoras del perceptrón cruzado para $M_r^M = 6.2$, como función de la magnitud M_r^m ; las líneas discontinuas horizontales indican los valores óptimos observados para el perceptrón simple ($M_r = 6.2$). Puede apreciarse claramente que para postnósticos el desempeño del perceptrón cruzado es mejor que el del perceptrón directo para varios valores de M_r^m ; ambas funciones calificadoras muestran mejoras para $M_r^m = 5.7, 5.8, 6.0$ y 6.1 , con el máximo valor para 6.1 .

Para pronósticos el método cruzado funciona mejor, en todos los aspectos, que el perceptrón directo para todas las M_r^M consideradas en este análisis (Figura 12).

La figura 13 muestra los valores codificados en color de las matrices de probabilidad de transición W obtenidas por el perceptrón simple para $M_r = 6.2$ (izquierda), y para el perceptrón cruzado con $M_r^m = 6.1$ y $M_r^M = 6.2$ (derecha). Puede verse que, aunque ambas matrices tienen regiones parecidas donde predominan valores bajos, sus altos ocurren para transiciones distintas; las diferencias entre ellas son debidas a la información aportada por los sismos con $M_r^m \leq M < M_r^M$.

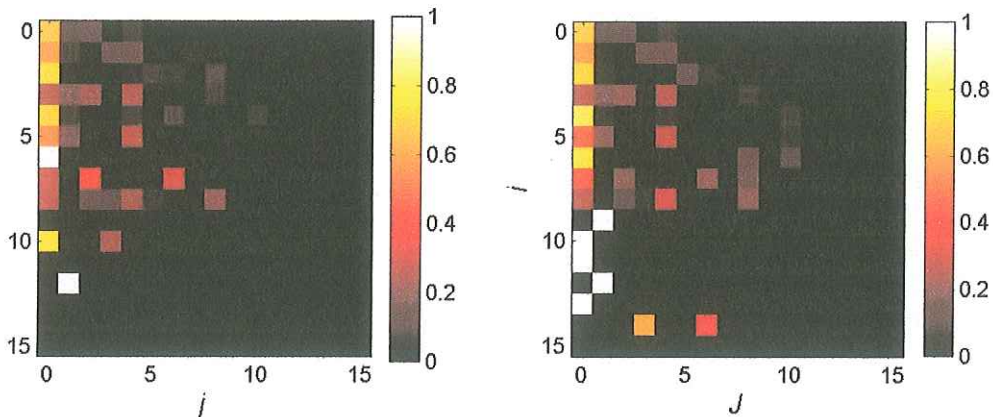


Figura 14. Representación de los valores de las matrices de probabilidad de transición codificados según la banda a la derecha. Izquierda: Peceptrón directo, para $M_r = 6.2$. Derecha: Perceptrón cruzado para $M_r^m = 5.7$ y $M_r^M = 6.2$.

Con base en el desempeño observado por los perceptrones fue posible aplicar estos métodos hasta $M_r = M_r^M = 6.5$; la figura 14 muestra el comportamiento de las funciones calificadoras como función de M_r^m para este caso límite.

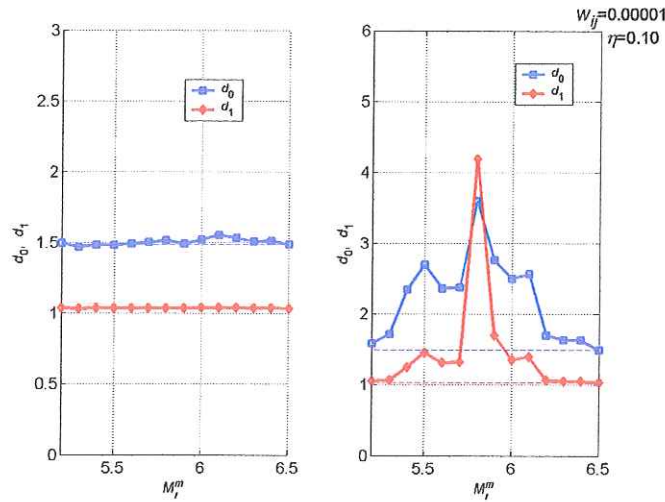


Figura 15. Valores de las funciones calificadoras del perceptrón *cruzado* para $M_r^M = 6.5$ en función de M_r^m . Izquierda *postnósticos*. Derecha *pronósticos*.

Se observa que aunque los valores máximos de las funciones calificadoras para el perceptrón cruzado ocurren para $M_r^M = 6.2$, para magnitudes grandes el perceptrón cruzado es consistentemente mejor que el perceptrón directo, tanto para postnósticos como para pronósticos, como puede apreciarse en las tablas V y VI.

V.6.2.3. Resultados del método de caminos preferenciales, y de los perceptrones con historia

El método descrito en la sección (III.2.1.2.) fue aplicado al sistema estudiado para historias de transiciones (directas) correspondientes a un amplio rango de magnitudes umbral y trayectorias de 2 y 3 pasos (antes del estado final).

No hay un número significativo de trayectorias de 2 pasos para $M_r < 5.8$ o de 3 pasos para $M_r < 6.1$; a partir de dichas magnitudes, todas las trayectorias que aparecen

varias veces incluyen 2 o más veces al estado cero y reflejan la tendencia del sistema al caso trivial. En cualquier caso, para un número de incidencias significativo, la probabilidad observada coincide bastante bien con la markoviana, o equivalentemente el número de sismos esperado de las probabilidades markovianas coincide ± 1 (excepcionalmente, 2) evento(s) con el número observado.

Esta similitud indica que el sistema en efecto tiene memoria de un solo paso, es decir, que efectivamente el sistema es Markoviano.

Este resultado es confirmado por los resultados de la aplicación de los perceptrones con historia, que también parecen indicar que estados previos al estado inmediato anterior no contienen información pertinente a la ocurrencia de estados futuros, que resulte en un desempeño mejor que el obtenido de perceptrones alimentados solamente por el estado anterior.

V.6.2.4 Resultados de la aplicación del perceptrón con información de magnitudes

Dado que el perceptrón directo tuvo su mejor desempeño para $M_r = 6.1$, se decidió explorar el desempeño del perceptrón *directo con información de magnitudes* para $M_r = 6.1$, resultando (para postnósticos) en un desempeño óptimo superior en 5.5% al observado por el perceptrón directo, e incluso superior en 4.4% al desempeño óptimo observado por el perceptrón cruzado para $M_r^m = 6.1$ y $M_r^M = 6.2$. Por su parte, el perceptrón *cruzado con información de magnitudes* mejora en un 1.5% el desempeño del perceptrón cruzado $M_r^m = 6.1$ y $M_r^M = 6.2$. Para pronósticos el perceptrón directo con

información de magnitudes para $M_r = 6.1$, resultó mejor que el perceptrón directo en 3.0%. Para pronósticos y magnitud umbral $M_r^M = 6.2$, el perceptrón cruzado con información de magnitudes mejoró en 20% el desempeño del perceptrón cruzado, Si bien los perceptrones con información de magnitudes (directo y cruzado) mejoran las estimaciones de peligro, estas mejoras son tan pequeñas que gráficamente son difíciles de apreciar.

VI. DISCUSIÓN Y CONCLUSIONES

VI.1. *Discusión*

Los nuevos métodos propuestos aquí, mejoran los resultados obtenidos por Herrera (2001) y Nava *et al.* (2005) para el método markoviano directo. Dichos nuevos métodos permiten hacer estimaciones útiles y confiables de peligro sísmico a partir del análisis estadístico de la sismicidad reportada en catálogos, y son adaptables a diferentes áreas sismogénicas.

Se implementó un criterio de evaluación del desempeño mediante el uso de funciones calificadoras; estas funciones, que asignan una calificación numérica al desempeño de cada modelo, permiten comparar distintos métodos y modelos y, por tanto, elegir el método y modelo óptimos para la estimación de peligro sísmico.

La razón $r = (d_0^{mx} - d_0^{md}) / d_0^{md}$ para postnósticos del catálogo completo (384 transiciones) y pronósticos de las últimas 20 transiciones, indica que el método cruzado da estimaciones de peligro 6.6% mejores que las estimadas por el método directo (para postnósticos y magnitudes umbral $M_r = 6.2$ y $M_r^M = 6.2$). Para pronósticos y magnitudes umbral $M_r = 6.1$ y $M_r^M = 6.1$, r indica una mejoría de 67% en las estimaciones de peligro del método cruzado. Por su parte, el desempeño del perceptrón cruzado para postnóstico y magnitud umbral $M_r^M = 6.2$ mostró un desempeño 10% mejor al observado por el método directo para $M_r = 6.2$. Para pronósticos y magnitudes umbral $M_r = 6.1$ y $M_r^M = 6.1$, r indica una mejoría de 50% en las estimaciones de peligro del perceptrón cruzado.

De la comparación del desempeño de los perceptrones y los métodos markovianos de conteo, resultó que el perceptrón simple da estimaciones de peligro 6.6% veces mejores que las estimadas por el método directo para ambos casos (postnósticos y pronósticos para $M_r = 6.1$). Para postnósticos y magnitudes umbral $M_r^m = 6.1$ y $M_r^M = 6.2$, el perceptrón cruzado mostró un desempeño 5.6% mejor al obtenido con el método markoviano cruzado. Para pronósticos y magnitudes umbral $M_r^m = 5.8$ y $M_r^M = 6.3$, r indica una mejoría de 5.6% en las estimaciones de peligro del perceptrón cruzado sobre las del método markoviano.

Es importante subrayar que, aunque para altas magnitudes los valores de las funciones calificadoras observados por los métodos cruzados no son tan grandes como en los casos óptimos ($M_r^M = 6.2$ para póstnosticos y $M_r^M = 6.3$ para pronósticos), sí son mucho mejores que los observados por los métodos directos; por ejemplo, el desempeño del método markoviano cruzado para póstnosticos con $M_r^m = 6.1$ y $M_r^M = 6.5$ resulto 5.5% mejor que el del método directo con $M_r = 6.5$. Por su parte, el perceptrón cruzado con $M_r^m = 6.1$ y $M_r^M = 6.5$ tuvo un desempeño superior en 4.5% al del perceptrón simple para y $M_r = 6.5$ para postnósticos. Es muy importante que el método cruzado permita extender el rango de estimaciones con buen desempeño a mayores magnitudes, ya que son éstas las más importantes de pronosticar.

Los resultados negativos para todos los modelos que incluyen historia de más de un paso, son una indicación de que el sistema estudiado no sólo tiene memoria, como fue reportado por Nava *et al.* (2005), sino que es verdaderamente markoviano.

VI.2. Conclusiones

- ❖ La aplicación de los nuevos métodos markovianos propuestos aquí al área de Japón, para *postnósticos* del catálogo completo y *pronósticos* de las últimas 20 transiciones, da resultados altamente satisfactorios que tienen probabilidades despreciables de ser obtenidos al azar. El buen desempeño de estos métodos confirma que este sistema sismogénico es un sistema con memoria, como fue reportado por Nava *et al.* (2005), y extiende la metodología propuesta en dicho trabajo para obtener estimaciones confiables de peligro sísmico a mayores magnitudes.
- ❖ Para la aplicación del método directo, las funciones calificadoras sugieren que, para intervalos de tiempo de $\Delta t = 0.10$ años, se obtiene los mejores resultados cuando se *pronostica* la ocurrencia o no ocurrencia de sismos con magnitudes $M \geq M_r = 6.1$; estos resultados son mejores que los reportados por Nava *et al.* (2005) para $M_r = 5.5$, y consideramos esto un buen resultado porque las magnitudes grandes son de mayor importancia para estimaciones de peligro.

- ❖ Para magnitudes umbral mayores que la magnitud óptima del método directo, el método markoviano cruzado da resultados consistentemente mejores que el método directo, particularmente para pronósticos (aunque, 20 pronósticos no constituyen un muestreo estadísticamente confiable). Este resultado indica que los eventos con magnitudes $M \geq M_r^m < M_r^M$ contienen información pertinente a la ocurrencia de sismos con $M \geq M_r^M$; lo que permite extender la magnitud umbral límite de un catálogo dado, para hacer estimaciones de peligro sísmico para eventos mayores.
- ❖ La aplicación de métodos neuronales al mismo sistema simogénico, para 384 postdicciones y 20 predicciones, da resultados altamente satisfactorios que tienen probabilidades despreciables de ser obtenidos al azar.
- ❖ Para el caso particular de $f(x) = x$, $w_{ij} = 0.00001$ y $\eta = 384^{-1}$ las probabilidades de transición obtenidos por el perceptrón son casi idénticas a las obtenidas mediante el método markoviano directo; esto indica que el método directo puede considerarse como un tipo particular de proceso de aprendizaje.
- ❖ La comparación de los valores de las funciones calificadoras para el método markoviano directo y el perceptrón directo indica que, si bien el método directo da resultados bastante buenos, no es el método óptimo para aprovechar la información contenida en los catálogos para la evaluación de peligro sísmico en sistemas sismogénicos.
- ❖ De todos los métodos probados, el perceptrón cruzado da los mejores resultados, particularmente en la estimación de peligro para sismos con $M_r^M = 6.5$.

- ❖ Resta por explorar si los métodos propuestos aquí funcionan igualmente bien en otras áreas sismogénicas; posiblemente, los métodos cruzados permitan la aplicación a regiones cuyos catálogos no sean tan completos como el utilizado. Determinar si sistemas en otras regiones son o no Markovianos y a qué escalas de tiempo, puede ser un dato valioso para estudios, no sólo de peligro y riesgo sísmicos, sino también de tectónica.
- ❖ Los métodos propuestos aquí sugieren, a su vez, nuevas variantes y métodos para la estimación de peligro sísmico en un sistema que restan por explorar. De particular interés es la integración de información adicional extrasísmica, como información geodésica, geoquímica y otras, mediante la consideración de procesos puntuales marcados y el uso de redes neuronales más complejas.

LITERATURA CITADA

- Agnew, D.C. y L.M. Jones 1991. Prediction probabilities from foreshocks. *J. Geophys. Res.* **96**: 11,959-11,971p.
- Anderson, J.A. y E. Rosenfeld. 1988. *Neurocomputing: Foundations of research*, M.I.T. Press, Cambridge, MA. 729pp.
- Bishop, C.M. 1995. *Neural networks for pattern recognition*, Calendon Press Oxford, UK. 482pp.
- Brillinger, D. 1982. Seismic risk assessment: some statistical aspects, *Earthquake Prediction Research*. **1**: 183-195p.
- Fedotov, S. 1965. Regularities of the distribution of strong earthquakes in Kamchatka, the Kuril Islands, and northeast Japan, *Trudy Isnt. Fiz. Zemli. Acad. Nauk. SSSR*. **36**: 66-94p.
- Goldman, S. 1953. *Information theory*. Dover Publs. Inc., USA. 385 pp.
- Gutenberg, B. y C. Richter. 1944. Frequency of earthquakes in California. *Bull. Seism. Soc. Am.* **34**: 185-188p.
- Hagiwara, Y. 1975. An stochastic model of earthquake occurrence and the accompanying horizontal land deformations. *Tectonophysics*. **26**: 91-101p.
- Heerman, P.D. y N. Khazenie. 1992. Clasificación de multiespectral remote sensing data using a back propagation neural network. *IEEE Transacciones on Geoscience and Remote sensing*. **30**: 81-88p.
- Herrera, C. 2001. Determinación de peligro sísmico mediante cadenas de Markov. Tesis de Maestría, CICESE. 86 pp.
- Herrera, C. Y F.A. Nava. 2004. La estimación Markoviana Directa de Peligro Sísmico en sistemas sismogénicos como un caso particular de aprendizaje neuronal tipo perceptrón. *GEOS*. **24**: 446-450p.
- Hertz, J., K. Anders y R.G. Palmer. 1991. *Introduction to the theory of neural computation*, Addison Wesley, Boston, MA, USA. 327pp.

- Hines, W. y C. Montgomery. 1993. Probabilidad y estadística para ingeniería. Segunda edición, Continental, México. 834 pp.
- International Seismological Centre. *On-line Bulletin*. <http://www.isc.ac.uk/Bull>. Fecha de consulta 05/Enero/2002.
- Kanamori, H. 1977. Seismic and aseismic slip along subduction zones and their tectonic implications, in Island Arcs, Deep Sea Trenches and Back-arc basins, Maurice Ewing Series I, editado por M. Talwani. y W. C. Pitman III, pp 163-174. Am. Geophys. Union. Washington D.C. USA.
- Keilis-Borok, V.I. y V.G. Kossobokov. 1990a. Premonitory activation of earthquake flow-algorithm M8. *Physics of the Earth and Planetary Interiors*. **61**: 73-83p.
- Keilis-Borok, V.I. y V.G. Kossobokov. 1990b. Times of increased probability of strong earthquakes ($M \geq 7.5$) diagnosed by algorithm M8 in Japan and adjacent territories. *Journal of Geophysical Research*. **95**: 12413-12442p.
- Lomnitz, C. 1974. *Global Tectonics and Earthquake Risk*. Elsevier Scientific Publishing Co, Amsterdam – London – New York. 320pp.
- Lomnitz, C. y F. Nava. 1983. The predictive power of seismic gaps. *Bull. Seism. Soc. Am.* **73**: 1815-1824p.
- Matsuda, T. y Y. Kinugasa. 1991. Active faults in Japan. *Episodes*. **14**: 199-204p.
- McCann, W.R., S. P. Nishenko, S.P. Sykes y J. Krause. 1979. Seismic gap and plate tectonics: seismic potential for major boundaries. *Pure Appl. Geophys.* **117**:1082-1147p.
- McClelland, J.L. y D. F. Rumelhart. 1986. "Paralell distributed processing". M.I.T. Press, Cambridge, MA. 632pp.
- McKenzie, D.P. y W. J. Morgan. 1969. Evolution of the triple junctions. *Nature*. **224**: 125-133p.
- Minsky, M.L. y S. Papert. 1969. *Perceptrons: An introduction to computational geometry*, M.I.T. Press, Cambridge, MA. 275pp.

- Mogi, K. 1968. Migration of Seismic Activity. Bull. of Earthquake Research Institute. **46**: 53-74.
- Nava, F.A. 1987. Terremotos. Fondo de Cultura Económica, México. 187pp.
- Nava, F.A. y V. Espíndola. 1993. Seismic hazard assessment through semi-estochastic simulation. Bull. Seism. Soc. Am. **83**: 450-468p.
- Nava, F.A, C. Herrera, J. Frez, y E. Glowacka. 2005. Seismic hazard evaluation using Markov chains; Application to the Japan area. Pure Appl. Geophys. **162**: 1347-1366p.
- Nishenko, S.P. 1991. Circum-Pacific seismic potential 1989-1999. Pure Appl. Geophys. **135**: 169-259p.
- Nilsson, N. 1965. Learning machines, McGraw-Hill, New York, 138pp.
- Okada, A. 1980. Quaternary faulting along the Median Tectonic Line of southwest Japan: Memoirs of the Geological Society of Japan. **18**: 79-108p.
- Reid, H. F. 1910. The mechanism of the Earthquake. The California Earthquake of April 18, 1906. En: Report of the State Earthquake Investigation Commission (Editors) Carnegie Institution. Washington D.C. 2:16-18p.
- Richter, C. 1958. Elementary Seismology. W. H. Freeman, San Francisco. 521pp.
- Roadhes, D.A. y F.F. Evison. 2005. Test of the EEPAS forecasting model on the Japan earthquake catalogue. Pure Appl. Geophys. **162**: 1271-1290p.
- Rosenblatt, F. 1958. The perceptron: a probabilistic model for information storage and organization in the brain. Psychological Review, **65**:386-408p. Reprinted in: J.A. Anderson and E. Rosenfeld 1988. Neurocomputing: Foundations of Research. MIT Press Cambridge, MA, USA. 729pp.
- Rosenblatt, F. 1962. Principles of neurodynamics. Spartan, New York. 215pp.
- Ross, S. 1970. Applied probability models with optimization applications .Dover, USA. 198pp.
- Seno, T., S. Stein, y A. Griff. 1993. A model for the motion of Philippine Sea plate consistent with NUVEL-1 and geological data. Journal Geophysic Research. **98**: 941-948p.

- Shimazaki, K. y T. Nakata. 1980. Time predictable recurrence model for large earthquakes, Geophys. Res. Lett. 7: 279-282p.
- Utsu, T. 1972. Aftershocks and earthquake statistics (IV). Journal of the Faculty of Science. Hokkaido University Series VII. 4: 1-42p.
- Uyeda, S. y H. Kanamori. 1979. Backarc opening and the mode for subduction. Journal Geophysical Research. 84: 1049-1061p.
- Uyeda, S. 1982. Subduction zones an introduction to comparative subductology Tectonophysics. 81: 133-159p.

APÉNDICE A

En este apéndice se presenta una breve recapitulación de los conceptos elementales de las cadenas de Markov.

A.1 Cadenas de Markov

Una Cadena de Markov es una sucesión de estados S_i (cada estado puede ser considerado resultado de una prueba) que satisface las dos propiedades siguientes:

1. Cada estado S_i pertenece a un conjunto finito de estados $\{S_0, S_1, S_2, \dots, S_N\}$ llamado espacio de estados del sistema.

2. La probabilidad de ocurrencia del estado para un tiempo dado $S([n+1] \Delta t) \equiv S_{(n+1)}$ depende solamente del estado inmediatamente precedente $S(n \Delta t) \equiv S_{(n)}$ y no de cualquier otro estado previo:

$$P\{S_{(n+1)} = j | S_{(n)} = i\} = \Pr\{S_{(n+1)} = j | S_{(n)} = i, S_{(n-1)} = k, \dots, S_{(0)} = z\} \quad (12)$$

para toda $n = 0, 1, 2, \dots$ y cualquier $j, i, k, \dots, z$; propiedad denominada propiedad Markoviana.

Para cada par de estados (i, j) se establece la probabilidad condicional, o probabilidad de transición de un paso p_{ij} que es la probabilidad de que j suceda inmediatamente después de i :

$$p_{ij} \equiv \Pr\{S_{(n+1)} = j | S_{(n)} = i\}, \quad (13)$$

y una cadena de Markov se puede describir por medio de la matriz de probabilidades de transición $P = [p_{ij}]$.

Se dice que p_{ij} es estacionaria si

$$P\{S_{(n+1)} = j \mid S_{(n)} = i\} = \Pr\{S_{(1)} = j \mid S_0 = i\}; \quad n = 0, 1, 2, \dots, \quad (14)$$

es decir, las probabilidades de transición permanecen invariables en el tiempo.

La existencia de probabilidades de transición estacionarias de un paso implica que:

$$p_{ij}^{(n)} = \Pr\{S_{t+n} = j \mid S_t = i\} = \Pr\{S_n = j \mid S_0 = i\}; \quad t = 0, 1, 2, \dots \quad (15)$$

Si las probabilidades de transición de un paso son estacionarias, es posible calcular las probabilidades de transición de varios pasos mediante las ecuaciones de Chapman-Kolmogorov (Hines, 1993; Ross, 1970):

$$p_{ij}^{(n)} = \sum_{l=0}^m p_{il}^{(\nu)} p_{lj}^{(n-\nu)}; \quad i = 0, 1, 2, \dots, m \quad j = 0, 1, 2, \dots, m \quad 0 \leq \nu \leq n, \quad (16)$$

que, para los casos $\nu = 1$ y $\nu = n-1$ resultan en

$$p_{ij}^{(n)} = \sum_{l=0}^m p_{il} p_{lj}^{(n-1)} = \sum_{l=0}^m p_{il}^{(n-1)} \cdot p_{lj}; \quad i = 0, 1, 2, \dots, m; \quad j = 0, 1, 2, \dots, m; \quad \forall n, \quad (17)$$

de donde la matriz de probabilidades de transición a n pasos $P^{(n)}$, puede obtenerse de la de las probabilidades de un paso como

$$P^{(n)} = P^n \quad (18)$$

lo que significa que, para evaluar la probabilidad de una transición a n pasos es necesario considerar todas las posibles combinaciones de transiciones intermedias que permitan llegar del estado inicial al final (combinaciones correspondientes a los productos renglón por columna de la matriz de probabilidades).

APÉNDICE B

En este apéndice, se presenta la teoría fundamental de las redes neuronales artificiales, principalmente, la arquitectura (estructura o patrón de conexiones) de una red neuronal tipo perceptrón, y el algoritmo de aprendizaje, conocido como regla de aprendizaje de Hebb.

Las redes neuronales artificiales (RNA) son máquinas computacionales de aprendizaje, que funcionan emulando el comportamiento de las redes neuronales biológicas (e.g. Hertz *et al.*, 1991; Anderson, 1988). Constan de un conjunto de unidades de entrada, una red de unidades básicas de procesamiento, llamadas neuronas, generalmente interconectadas entre sí, y una salida producida por éstas con base en la entrada. Cada neurona produce una salida determinada por la información que recibe, ya sea directamente de la entrada o a través de otras neuronas.

B.1 *Perceptrón*

La RNA más sencilla es la tipo *perceptrón* (Rosenblatt, 1958, 1962; Nilsson, 1965; Minsky & Papert, 1969) que consta de una serie de entradas $[z] = \{z_i; i = 1, 2, \dots, n\}$, una sola capa de neuronas, y un conjunto de salidas $[o] = \{o_j; j = 1, 2, \dots, n\}$ (Fig.B.1). El perceptrón se caracteriza por ser capaz de resolver problemas linealmente separables.

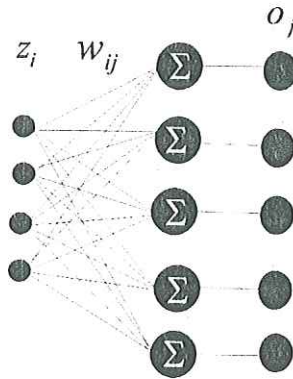


Figura 16. Perceptrón. Los círculos con sumatoria representan las neuronas de la red.

A partir de la entrada $[z]$, la neurona j produce la salida

$$o_j = f\left(\sum_{i=1}^n w_{ij} z_i + b_i\right) \quad (19)$$

donde las w_{ij} son pesos asignados a las líneas de comunicación de cada entrada a cada neurona (representan el conocimiento que tiene la RNA acerca de un determinado problema), y f es una *función de activación* (usualmente igual para todas las neuronas) que determina cómo es la salida. La función de activación es escogida de acuerdo con las especificaciones del problema que la neurona tenga que resolver y, en las redes clásicas que emulan los procesos sinápticos de las neuronas, es a menudo una función binaria que determina si la neurona es activada (y propaga la información) o permanece inhibida (salida nula). Generalmente, se añade al conjunto de pesos de las neuronas un parámetro constante adicional b , denominado “bias”, para dar mayor estabilidad al sistema.

B.2 Algoritmo de Aprendizaje

La característica más importante de las RNA es que pueden *aprender*, esto es, para cada entrada $[z]$ dada, ir ajustando los valores iniciales de los pesos, con base en el error cometido, de manera que la salida de la red tienda a la salida deseada $[d]$. Rosenblatt (1958, 1962) demostró que, presentando a la red un conjunto de patrones de entrada-salida y empleando una simple regla de corrección de los pesos con base en los residuales de las respuestas obtenidas, se puede obtener los parámetros que configuran la respuesta óptima de la red, w_{ij} , en un número finito de pasos. Este procedimiento determina la regla de aprendizaje de la red, conocida como regla de Hebb:

$$\begin{aligned} w_{ij}^{k+1} &= w_{ij}^k + \Delta w_{ij}^k \\ \Delta w_{ij}^k &= \eta (d_j^k - o_j^k) z_i^k \end{aligned} \quad (20)$$

$$b_i^{k+1} = b_i^k + \eta (d_j^k - o_j^k) \quad (21)$$

donde el parámetro de amortiguación η , llamado *razón de aprendizaje* o *factor de relajación*, controla la rapidez con que son modificados los pesos durante el entrenamiento y toma usualmente valores entre 0 y 1. En algunos casos η puede ser constante; en otros es reducida conforme progresa el entrenamiento hasta alcanzar la estabilidad. McClelland &

Rumelhart (1986) proponen como valor óptimo del parámetro de amortiguación, $\eta = 1/n$, donde n es el número de neuronas de la red. Por su parte, Heermann & Khazenie (1992) añaden dos factores a dicho valor:

$$\eta = C_0 \times \frac{1}{p} \times \frac{1}{n} \quad (22)$$

donde C_0 es una constante empírica y p es el número de patrones de entrenamiento. Sin embargo, en el presente trabajo los valores del parámetro de amortiguación obtenidos con las expresiones anteriores no resultaron en el desempeño óptimo de la red; el valor final de este parámetro fue determinado empíricamente como el que dio el desempeño óptimo de la red, juzgado según los discriminantes definidos en la sección III.3.

APÉNDICE C

Este apéndice presenta la definición de los parámetros umbral, así como los criterios empleados para su determinación. Como se verá, la elección de parámetros umbral no es única, y resulta generalmente de una serie de compromisos entre características positivas y negativas en la elección de cada parámetro así como de la influencia de éste sobre los demás parámetros.

C.1. *Parametro umbral de espacio*

Está dado por la delimitación espacial del sistema (área de estudio) y de cada una de las regiones que lo conforman. El sistema está constituido por:

1. Regiones sismogénicas, escogidas con base en la tectónica y la sismicidad del área de interés, de acuerdo con la premisa 1.
2. “Zonas de Nadie” que son bandas entre las fronteras o límites de las regiones, cuyos sismos son ignorados para evitar asignarlos a la región equivocada debido a errores de localización. El ancho de estas bandas corresponde aproximadamente al error epicentral esperado.

C.2. *Parámetro umbral de magnitud*

Definimos como magnitud umbral M_r a la magnitud mínima de interés. La elección de la magnitud umbral M_r , es gobernada por los siguientes factores:

- a) M_r debe ser suficientemente grande para que la evaluación de peligro sea importante.
- b) M_r debe ser suficientemente grande para evitar la ocurrencia esencialmente aleatoria de los sismos pequeños.
- c) M_r debe ser suficientemente pequeña para que los sismos no estén ausentes de todos o de la mayoría de los intervalos considerados, ya que entonces el muestreo no sería suficiente para su predicción.

C.3. *Parámetro umbral de tiempo*

El umbral de tiempo es la longitud o tamaño del intervalo Δt durante el cual la ausencia o presencia de sismos ocurridos en cada región define el estado del sistema, y, por lo tanto, es también el intervalo al que se aplicará cada evaluación de peligro.

El parámetro Δt se determina con base en las siguientes consideraciones:

- a) Idealmente, Δt debe ser pequeño, para que las estimaciones de peligro sean útiles.
- b) Un Δt demasiado pequeño, ocasiona que el estado de ausencia de sismos en todas las regiones, el estado 0, sea el más frecuente, de manera que las probabilidades distintas de p_{00} sean tan pequeñas que no tengan valor predictivo.

- c) Un Δt demasiado grande, ocasiona que el estado con presencia de sismos en todas las regiones, el estado $S-1$, sea mucho más frecuente que cualquier otro, de manera que la transición de $S-1$ a $S-1$ sea dominante sobre todas las demás y las probabilidades distintas de $p_{S-1 S-1}$ sean tan pequeñas que no tengan valor predictivo.
- d) Para una cobertura de catálogo determinada, Δt grande disminuye el número de transiciones muestreadas y hace menos robustas las estimaciones de p_{ij} .
- e) Finalmente, si las premisas son correctas, Δt debe ser lo suficientemente grande para que se lleve acabo la interacción entre regiones.

De este modo, una vez que se ha elegido las regiones y la magnitud umbral, se hace la estimación preliminar de Δt a partir de los ceros de $\theta_{00} - \theta_{S-1 S-1}$ (el número de transiciones del estado 0 al estado 0 menos el número de transiciones del estado $S-1$ al estado $S-1$) y de $\xi_0 - \xi_{S-1}$ (el número total de ocurrencias del estado 0 menos el número total de ocurrencias del estado $S-1$).

C.4. Determinación del tiempo inicial

El tiempo inicial se fija en el intervalo $t_{\min} - \Delta t < t < t_{\min} + \Delta t$, donde $t_{\min}$ es el tiempo del primer sismo con $M \geq M_r^0$ y $t_{\min} - \Delta t$ es cubierto por el catálogo. El uso de tiempo iniciales ligeramente distintos (que pueden ser escogidos aleatoriamente) permite

estudiar la estabilidad del método, al medir la variabilidad de los resultados de una realización a otra.

La elección final de los parámetros umbral se hace en forma iterativa y está pragmáticamente determinada por la combinación de éstos que resulta en el mejor desempeño juzgado mediante discriminantes apropiados (IV.3).