

Predicción de índices bursátiles mediante un sistema híbrido basado en modelos ocultos de Markov y redes neuronales artificiales

A Hybrid HMM/ANN method for stock index forecasting

Jabbour, Georges¹ y Maldonado, José^{2*}

¹Escuela de Ingeniería de Sistemas Facultad de Ingeniería ULA

²IEAC, Núcleo La Liria, Edif. G. ULA

Mérida-Venezuela

*maldonaj@ula.ve

Recibido: 01-10-2008

Revisado: 25-06-2009

Resumen

El enfoque de la mezcla de expertos comprende una gran variedad de modelos cuya filosofía consiste en descomponer una serie de tiempo en varios estados, de manera que cada estado se modela a través de un experto, con lo cual se logran capturar los patrones de la serie de tiempo de una manera eficiente. Los Expertos Temporales Ocultos de Markov (ETOM) representan uno de los métodos más avanzados pertenecientes a esta corriente, que no son más que modelos híbridos basados en Redes Neuronales Artificiales (RNA) y Modelos Ocultos de Markov (MOM). Una característica distintiva de los ETOM es que la transición de estados de la serie de tiempo se modela a través de un Modelo Oculito de Markov, en el cual la matriz de transición de estados es variante en el tiempo. El objetivo de esta investigación consistió en evaluar el desempeño de los ETOM en la predicción de índices bursátiles, y realizar una comparación con resultados obtenidos mediante RNA puras. Los experimentos fueron realizados con 15 índices bursátiles, y se obtuvo que los modelos ETOM superan sustancialmente a las RNA tanto en precisión como en su capacidad para capturar patrones.

Palabras clave: Series de tiempo, mezcla de expertos, redes neuronales artificiales, modelos ocultos de Markov.

Abstract

The mixture of expert approaches includes a great variety of models, whose philosophy consists in breaking down a time series in various states, so that each state is modeled through an expert, which results in capturing the patterns of the time series in an efficient way. The Time-Line Hidden Markov Experts (THMEs) represents one of the most advanced methods based on this approach, which are just hybrid models based on Artificial Neural Networks (ANNs) and Hidden Markov Models (HMMs). A distinctive characteristic of the THMEs is that the state transitions of the time series is modeled through a HMM, whose state transition matrix is time-variant. The aim of this research consisted in assessing the performance of the THME in the prediction of Stock Market Indexes, and to compare it with the results obtained through pure ANNs. The experiments were carried out using 15 stock market indexes, and the results obtained were that the THMEs substantially overcomes the ANNs, both in accuracy as in its capacity to detect patterns.

Key words: Time series, mixture of experts, artificial neural networks, hidden Markov models.

1 Introducción

Desde el momento en el que se originaron las bolsas de valores, se ha tratado de descifrar el comportamiento de los precios de las acciones que en ellas se comercializan. Diversos enfoques y aportes multidisciplinarios han dado

origen a ciertas técnicas y metodologías para el análisis financiero, las cuales, indiferentemente de su eficiencia, siguen empleándose en la actualidad (Lynch y Rothchild, 1996). Esto se debe a que las series de tiempo financieras poseen patrones que son difíciles de detectar, lo cual generalmente conlleva a concluir que estas series de tiempo son

caminos aleatorios, y que las predicciones obtenidas son prácticamente inútiles, tanto a corto como a largo plazo (Burton, 1996). Sin embargo, tales aseveraciones han sido objetadas por los analistas de valores y los investigadores pertenecientes a enfoques innovadores (Wei, Yoshiteru y Shou-Yang, 2005), de hecho, autores ampliamente reconocidos por sus contribuciones en el área de las finanzas, afirman que es posible cierto grado de predicción de los eventos futuros a partir del conocimiento del pasado (Chorafas, 1994). Desde los años 90, este punto de vista ha sido la génesis de una gran cantidad de investigaciones sobre la aplicación de enfoques alternativos y multidisciplinarios en el análisis de los sistemas financieros, tales como: la Teoría de Caos y Fractales, Wavelets, Máquinas de Vectores de Soporte, Inteligencia Artificial, Modelos Ocultos de Markov e incluso, combinaciones de estos métodos (Lijuan, 2002).

Uno de los métodos más recientes que pertenecen a esta corriente, es el basado en Expertos Temporales Ocultos de Markov (ETOM), que consiste en descomponer la serie de tiempo en estados, que internamente presentan un comportamiento homogéneo y sin dinámicas complejas. En este sentido, la idea principal de los modelos ETOM es entrenar un "sub-modelo" para cada estado, con el objetivo de convertirlo en "experto" en ese ambiente. Posteriormente, las respuestas de todos los expertos se combinan por medio de un algoritmo para generar una salida global que representa la predicción arrojada por el modelo ETOM.

Bajo este enfoque, el comportamiento de una serie de tiempo obedece a un proceso estocástico oculto, es decir, para cada instante de tiempo, la serie de tiempo se encuentra bajo un estado en particular que ejerce influencia sobre el comportamiento presente y futuro de la misma. Por lo tanto, es necesario conocer la manera en la que se comporta la serie de tiempo cuando se encuentra en cada uno de los posibles estados, y precisamente este es el objetivo de cada uno de los expertos, sin embargo, esto no es suficiente para que el modelo ETOM pueda ofrecer buenas predicciones, ya que es necesario entender la forma en la que evolucionan los referidos estados de la serie de tiempo, de hecho, antes de predecir el valor de la serie de tiempo es necesario predecir el estado de la misma. En este sentido, los Modelos Ocultos de Markov (MOM) son apropiados para modelar el proceso estocástico que representa al estado de la serie de tiempo, y por lo tanto, sirven para controlar, o administrar, las salidas de expertos. No obstante, un MOM convencional no es capaz de describir, para series de tiempo complejas, las transiciones de estado para todo instante de tiempo, ya que sus probabilidades de transición de estado son constantes en el tiempo, lo cual impide que el MOM pueda cumplir su objetivo de la forma adecuada, trayendo como consecuencia predicciones inexactas. Para resolver tal problema, la matriz de transición de estados del MOM debe ser variante con respecto al tiempo. Matemáticamente, esto significa que en lugar de utilizar una matriz de transición A constante, se genera una matriz de transición $A(t)$ que varía

de un instante a otro, de acuerdo a la dinámica de la serie de tiempo bajo consideración.

En este trabajo, el objetivo central consistió en utilizar los ETOM para hacer predicción de series de tiempo financieras. Las actividades generales desarrolladas comprendieron una revisión teórica sobre el enfoque de los ETOM, su arquitectura, entrenamiento y operación. En cuanto al entrenamiento, se tomó en cuenta que un ETOM involucra un MOM no clásico, donde los estados y las probabilidades de las transiciones de estado variantes en el tiempo, fueron modelados con RNA entrenadas a partir de patrones obtenidos a través de técnicas de clustering.

Se realizaron pruebas de predicción, por un lado, con los ETOM, y por otro lado, con un tipo de RNA pura, para finalmente comparar las dos técnicas en base a los resultados obtenidos en cada caso.

2 Los expertos temporales ocultos de Harkov

La arquitectura de un ETOM se muestra en la Fig. 1 (Wang et al, 2003). Los expertos pueden ser cualquier tipo de modelo de regresión, dependiendo del problema bajo estudio (Lijuan, 2002). En este caso los expertos son RNA que modelan los estados de la serie de tiempo, los cuales se obtienen a través del método Clustering Difuso. Para este último, el criterio de clasificación de un patrón de la serie (en algún estado), es la situación dinámica de la serie de tiempo (ΔX_t) para el instante al que pertenece tal patrón, de manera que las muestras de datos que tienen una estructura similar se incluyen en un mismo grupo. Por otro lado, la red de transición de estados es también un modelo de regresión, utilizado para predecir las transiciones de estado en función de ΔX_t , con el objetivo de identificar a los expertos asociados a ciertos patrones de cambio de la serie de tiempo. Cabe destacar que para predecir las transiciones de estado se utiliza como entrada ΔX y no X . Cuando se utiliza X , significa que las regiones de los expertos o las transiciones de estado están determinadas por la posición de la serie de tiempo y no por su situación dinámica al momento de generar la predicción.

El MOM combina los resultados de los expertos por medio de las probabilidades a priori de estar en cada estado, las cuales son obtenidas a partir de las probabilidades de transición de estado y de la cadena de estados previos, tal como se puede observar en la Fig. 1.

En el MOM se asume que para cada estado, la densidad de probabilidad de las observaciones $b_i(t)$ es Gaussiana. La estimación de la media condicional de esta función se toma directamente de la salida del respectivo experto, mientras que la varianza se determina durante el proceso de entrenamiento, en el cual se ajusta el nivel de ruido de cada estado. Esto no solo ofrece una explicación sobre la distribución de la salida de los expertos para cada instante de tiempo, sino que permite el uso de la ley de Bayes para calcular las probabilidades a posteriori que se utilizan en la predicción de una observación un paso (o una unidad de

tiempo) adelante.

3 Entrenamiento de los modelos ETOM

3.1 Descomposición de la serie en M estados

Se generaron M subconjuntos de patrones, con los respectivos grados de pertenencia difusos, de manera que cada

patrón puede pertenecer a uno o más grupos (o estados). Para ello, se aplicó el Clustering Difuso de C-Medias (CDCM), utilizando como criterio de agrupación, la situación dinámica (D_t) asociada a cada patrón:

$$D_t = \Delta X_t = X_t - X_{t-1} = [(y_{t-1} - y_{t-2}), (y_{t-2} - y_{t-3}), \dots, (y_{t-L} - y_{t-(L+1)})] \quad (1)$$

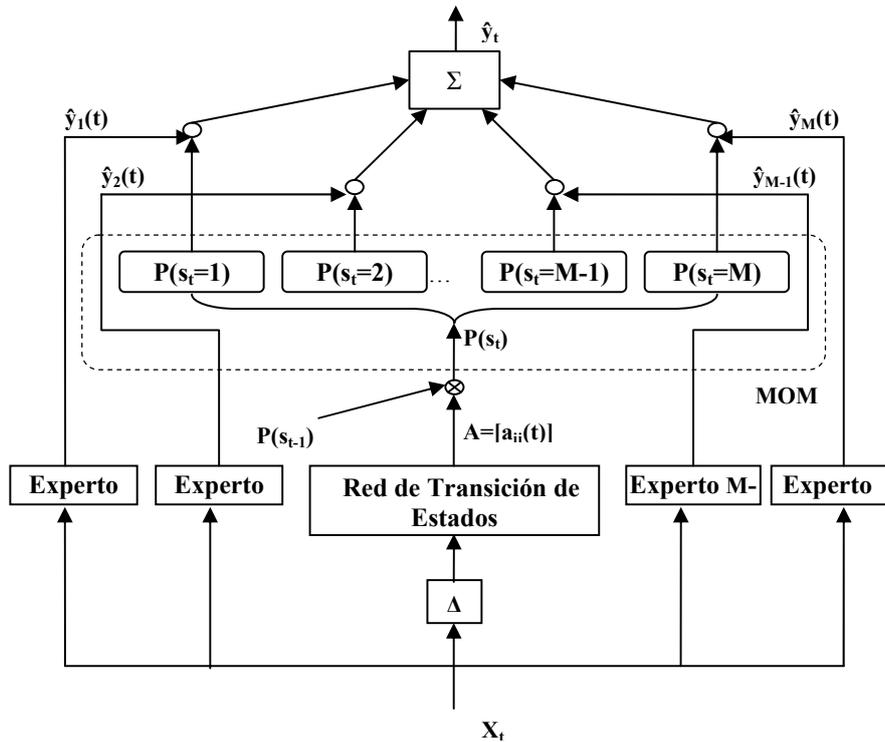


Fig. 1. Diagrama del ETOM. $P(s_t=i)$ es la probabilidad de estar en el estado i en el tiempo t . “ Σ ” denota una operación de suma. El símbolo “O” denota una multiplicación convencional. “ \otimes ” denota un producto matricial

El algoritmo CDCM obtiene los clusters minimizando gradualmente la siguiente función objetivo:

$$O = \sum_{t=1}^T \sum_{j=1}^M [\mu_j(t)]^2 \|D_t - C_j\| \quad (2)$$

donde D es un conjunto de T datos, y C , un número predeterminado de clusters.

El proceso culmina cuando la función (2) alcanza su valor mínimo (Jang et al, 1997). El resultado es la matriz $UN \times C = \{\mu_{ij}\}$, en donde μ_{ij} (representada a veces como $\mu_j(i)$) es el grado de pertenencia del i -ésimo patrón al j -ésimo estado; $j=1, \dots, M$. Cada una de las filas de la matriz U debe sumar exactamente 1, ya que la probabilidad de que una observación no pertenezca a ninguno de los estados es cero.

3.2 Entrenamiento de los Expertos

Para cada uno de los M grupos de patrones obtenidos

en el paso anterior, se construyó una red neuronal multicapa de conexión hacia adelante para desempeñar el papel de experto en el respectivo estado. Cada experto local se entrenó utilizando el algoritmo de retropropagación del error (Gupta et al, 2003).

3.3 Entrenamiento del MOM

El MOM con la propiedad de transición de estado variable en el tiempo se entrenó con un algoritmo de Baum-Welch modificado, basado en el principio EM (Rabiner, 1989). Suponiendo que la distribución de probabilidad del MOM es Gaussiana, para el estado j se tiene que:

$$b_j(y_t) = P(y_t | s_t = j, X_t, \lambda') = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left\{-\frac{[y_t - \hat{y}_j(X_t)]^2}{2\sigma_j^2}\right\} \quad (3)$$

en (3), $\hat{y}_j(X_t)$ corresponde a la salida del j -ésimo experto; $j=1, \dots, M$; $t=1, \dots, T$ y $s_t \in \zeta$ (ζ es el espacio de estados de la serie de tiempo).

En el algoritmo de Baum-Welch modificado, el entrenamiento del MOM se realizó mediante las siguientes expresiones (las cuales garantizan estimaciones máximo verosímiles), (Jang et al, 1997):

$$\tilde{\pi}_i = \frac{P(Y, s_0 = i | \lambda')}{P(Y | \lambda')}; \quad i=1, \dots, M \quad (4)$$

$$\tilde{\alpha}_{ij}(t) = \frac{P(Y, s_{t-1} = i, s_t = j | \lambda')}{P(Y, s_{t-1} = i | \lambda')}; \quad i=1, \dots, M; j=1, \dots, M; t=1, \dots, T \quad (5)$$

$$\tilde{\sigma}_i^2 = \frac{\sum_{t=1}^T [\gamma_i(t)(y_t - \hat{y}(X_t))^2]}{\sum_{t=1}^T \gamma_i(t)}; \quad i=1, \dots, M; \quad (6)$$

3.4 Entrenamiento de la Red de transición de estados

Se utilizó una RNA de base radial (Witten y Frank, 2005) para predecir las probabilidades de transición de estado, las cuales al ser combinadas con las salidas de los expertos por medio del MOM, producen la salida global del modelo. A diferencia de los expertos locales, esta red recibe como entrada ΔX_t en lugar de X_t , donde X_t es un vector que contiene los valores pasados de la serie que se utilizan para predecir el valor en el instante t , y por lo tanto, las transiciones de estado están determinadas por la situación dinámica de la serie.

4 Predicción de series de tiempo mediante Etom

Para generar la predicción de un paso, \hat{y}_t , se utilizan tanto las probabilidades a priori de los estados como las probabilidades a posteriori. En el primer caso, en la predicción a priori, se toman en cuenta las observaciones anteriores al instante t . En el segundo caso, en la predicción a posteriori, se obtiene a partir de la información anterior a t y del resultado de la predicción a priori. Estas predicciones fueron calculadas a través de los siguientes pasos (Jang et al, 1997):

- Cálculo de las probabilidades a priori de que la serie se encuentre en cada uno de los estados en el instante t :

$$P(s_t = j | Y_{t-1}, \lambda') = \sum_{i=1}^M a_{ij}(t) P(s_{t-1} = i | Y_{t-1}, \lambda') \quad (7)$$

- Cálculo de la predicción a priori:

$$\hat{y}_{pr}(t) = \sum_{i=1}^M P(s_t = i | Y_{t-1}, \lambda') \hat{y}_i(X_t) \quad (8)$$

- Cálculo de las probabilidades a posteriori de que la serie se encuentre en cada uno de los estados en el instante t :

$$P(s_t = i | Y_t, \lambda') = \frac{P(Y_t | s_t = i, \lambda')}{\sum_{j=1}^M P(Y_t | s_t = j, \lambda')} = \frac{p(y_t | s_t = i, \lambda') P(s_t = i | Y_{t-1}, \lambda')}{\sum_{j=1}^M p(y_t | s_t = j, \lambda') P(s_t = j | Y_{t-1}, \lambda')} \quad (9)$$

- Cálculo de la predicción a posteriori:

$$\hat{y}_{po}(t) = \sum_{i=1}^M P(s_t = i | Y_t, \lambda') \hat{y}_i(X_t) \quad (10)$$

Las ecuaciones (7-10) generan dos predicciones distintas, una a priori y otra a posteriori, y a pesar de que esta última es teóricamente más adecuada, no existen garantías de que siempre va a ser mejor, por lo tanto, al evaluar un modelo es necesario observar ambos tipos de predicciones para finalmente seleccionar la mejor. Matemáticamente, esto significa igualar a \hat{y}_t con $\hat{y}_{pr}(t)$ o con $\hat{y}_{po}(t)$. Por otro lado, estas expresiones ofrecen el mecanismo para obtener predicciones de un paso. Cuando se desea generar una predicción de múltiples pasos, es decir, la predicción para un instante ubicado h unidades de tiempo en el futuro, se deben generar h predicciones de un paso, comenzando por el instante t y terminando en el instante $t+h-1$, de manera que cada predicción obtenida se utiliza como entrada para generar la predicción del siguiente instante, y así sucesivamente hasta llegar a la h -ésima unidad de tiempo. Este esquema de predicciones tiene un sentido más realista que las predicciones de un solo paso, razón por la cual fue utilizado en este trabajo para realizar la validación de los modelos.

5 Experimentos y resultados

Se llevaron a cabo predicciones de 15 series de tiempo financieras, 3 de origen venezolano y 12 extranjeras. En la tabla 1, se describen estas series.

Estas series fueron analizadas en base a sus promedios semanales. Esto se hizo con el objetivo de reducir la duración del entrenamiento, dada la cantidad de series de tiempo y dada la cantidad de modelos a evaluar por cada serie. Las series de promedios semanales tenían entre 260 y 280 datos, con lo cual se logró reducir drásticamente la duración de cada entrenamiento, sin afectar la calidad de los resultados; al contrario, se lograba eliminar parte del ruido presente en cada serie. Con esto tampoco se estaría limitando el alcance de las predicciones, ya que generalmente los analistas financieros no están interesados en una predicción de un día para otro o una cantidad específica de días, sino en predicciones para un tiempo futuro múltiplo de una semana.

Por otro lado, se utilizó, de izquierda a derecha, el 85% de cada serie para entrenamiento y el 15% restante para la validación.

5.1 Comparación en cuanto a exactitud

En la tabla 2, se muestran los resultados de las mejores RNA y los mejores modelos ETOM encontrados, luego de evaluar en promedio 10,67 RNA y 11,72 modelos ETOM para cada serie. Los criterios de parada para el entrenamiento de cada RNA fueron: un error de entrenamiento igual a 10^{-5} o un número de ciclos igual a 10000, lo que ocurriera

primero, sin provocar el sobreentrenamiento; mientras que en el caso de los modelos ETOM, se utilizaron los siguientes parámetros: 0.01 como error de entrenamiento y un máximo de 1000 ciclos para el MOM, un error de entrenamiento de 0.01 para la Red de transición de estados, y un error de entrenamiento de 10^{-5} y un máximo de 2000 ciclos para las RNA expertas.

Tabla 1. Series de tiempo bajo estudio

Nombre	Descripción	Período de Estudio	No. de Datos
IBC	Índice Bursátil Caracas	02/01/1997 al 28/06/2002	1316
IFC	Índice Financiero Caracas	02/01/1997 al 28/06/2002	1315
IIC	Índice Industrial Caracas	02/01/1997 al 28/06/2002	1315
BOVESPA	Índice Bursátil Brasileño	02/01/1998 al 27/12/2002	1232
CAC40	Índice Bursátil Francés	02/01/1998 al 31/12/2002	1264
DAX	Índice Bursátil Alemán	04/01/1999 al 30/12/2003	1265
DJ	Índice Bursátil Dow Jones	03/01/2000 al 31/12/2004	1256
FTSE100	Índice Bursátil Inglés	02/01/1997 al 30/12/2001	1260
NASDAQ	Índice Bursátil Nasdaq	02/01/1997 al 30/12/2001	1256
NIKKEI225	Índice Bursátil Japonés	04/01/2000 al 30/12/2004	1229
NYSE	Índice Bursátil de la Bolsa de N. York	03/01/2000 al 31/12/2004	1256
SEOUL	Índice Bursátil Coreano	05/01/1998 al 30/12/2002	1222
SHANGHAI	Índice Bursátil de Hong Kong	02/01/2001 al 30/12/2005	1200
SP500	Índice Bursátil S&P 500	03/01/2000 al 31/12/2004	1256
SM	Índice Bursátil Suizo (Swiss Market)	01/04/1997 al 28/12/2001	1195

Tabla 2. Errores cuadráticos medios de las predicciones de las RNA y los ETOM

Serie de Tiempo	Total RNA	Total ETOM	Corto Plazo		Mediano Plazo		Largo Plazo	
			RNA	ETOM	RNA	ETOM	RNA	ETOM
IBC	126670	75080	83180	67760	128560	78340	168270	79760
IFC	203190	193870	201490	170750	184760	154060	225310	264610
IIC	446640	153960	200090	68900	806870	158750	345040	249270
BOVESPA	2158400	598470	560200	430430	2547200	799770	3587600	559150
CAC40	1228200	217350	268200	190910	1749200	358140	1747000	95010
DAX	189940	181060	23480	19050	134390	128720	452310	434390
DJ	19598	18729	13096	9719	32902	4749	13978	43356
FTSE100	120800	119170	61210	48500	204930	211720	107100	110130
NASDAQ	21848	34052	10094	20344	40765	63852	16822	17743
NIKKEI	76350	71380	156350	129810	24640	27400	42390	54050
NYSE	70310	63001	21720	35438	33250	84941	168150	71640
SEOUL	29217	17962	8201	8878	29402	11491	53854	36344
SHANGHAI	12666	2344	17360	3316	8320	2860	11859	633
SP500	1278	1049	323	257	920	405	2798	2688
SM	463090	174600	113910	133140	698030	263410	640810	134800

Note que en la tabla 2, las columnas numéricas están agrupadas de dos en dos, en donde cada grupo de 2 representa la comparación de las RNA con los ETOM en algún sentido particular. En las 2 primeras columnas, se comparan los Errores Cuadráticos Medios (ECM) de las predicciones basadas en el total de los datos de validación, el 15% del total de cada serie de tiempo. En las siguientes 6 columnas se realiza un análisis comparativo de los ECM de las predicciones en distintos horizontes de tiempo: corto, mediano y largo plazo; los cuales vienen dados por la primera, segunda y tercera parte de las predicciones, respectivamente; siendo estos 3 grupos del mismo tamaño. Esta descomposición tiene por objetivo analizar la calidad de las predicciones de cada método con respecto al tiempo, es decir, se evalúa la efectividad de cada método en distintos horizontes de tiempo.

Para ilustrar lo antes expuesto, considérese el caso de la serie de tiempo "IBC". Originalmente, en la frecuencia diaria, esta serie contaba con 1316 observaciones, pero al ser transformada en una serie de frecuencia semanal, su tamaño se redujo a 273. De estos 273 datos, 232 (el 85%) fueron utilizados para el entrenamiento de modelos; y los 41 datos restantes (el 15%) fueron empleados en la validación. Por lo tanto, en la validación de cada modelo, fueron generadas 41 predicciones, llámense las predicciones globales, de las cuales, las 14 primeras corresponden a las predicciones de corto plazo; las 14 siguientes, las de mediano plazo; y las 13 últimas, las de largo plazo.

En la tabla 2, las celdas con valores resaltados en ne-

grita y cursiva, indican que la RNA fue mejor en ese aspecto (dependiendo de la columna) que el modelo ETOM, mientras que para los casos restantes, ocurrió lo contrario, por ejemplo: el ECM total obtenido con la RNA para la serie NASDAQ, igual a 21848, está resaltado, y por lo tanto, para esta serie la RNA tuvo un ECM menor al del ETOM; ahora bien, para la serie SEOUL, la RNA tuvo un mejor desempeño sólo en las predicciones a corto plazo, es decir, para la predicción global y para las predicciones de mediano y largo plazo, el modelo ETOM fue superior.

Por lo tanto, los resultados que contiene la tabla 2 muestran que los modelos ETOM superan ampliamente a las RNA en cuanto a exactitud, ya que en el horizonte de tiempo total (o global), tuvieron un menor ECM para 14 de las 15 series de tiempo. Además, en las predicciones a corto, mediano y largo plazo, tuvieron un menor error en 12 de 15, 11 de 15 y 10 de 15 series, respectivamente. Sin embargo, no tiene sentido comparar las predicciones de 2 o más series mediante el ECM, ya que esta medida está estrictamente asociada al rango de la serie, y las series de tiempo utilizadas presentan rangos muy distintos entre sí. En este sentido, se realizó una transformación del ECM para poder comparar resultados no solo de distintos modelos, sino también de distintas series. Para un modelo y una serie en particular, esta nueva medida no es más que la relación porcentual entre el ECM de las predicciones de cada uno de los plazos y el ECM de las predicciones globales, tal como se muestra en la tabla 3.

Tabla 3. Relación porcentual entre las predicciones de cada plazo y la predicción global

Serie de Tiempo	Corto Plazo		Mediano Plazo		Largo Plazo	
	RNA	ETOM	RNA	ETOM	RNA	ETOM
IBC	65,67%	90,25%	101,49%	104,34%	132,84%	106,23%
IFC	99,16%	88,07%	90,93%	79,47%	110,89%	136,49%
IIC	44,80%	44,75%	180,65%	103,11%	77,25%	161,91%
BOVESPA	25,95%	71,92%	118,01%	133,64%	166,22%	93,43%
CAC40	21,84%	87,84%	142,42%	164,78%	142,24%	43,71%
DAX	12,36%	10,52%	70,75%	71,09%	238,13%	239,91%
DJ	66,82%	51,89%	167,88%	25,36%	71,32%	231,49%
FTSE100	50,67%	40,70%	169,64%	177,66%	88,66%	92,41%
NASDAQ	46,20%	59,74%	186,58%	187,51%	77,00%	52,11%
NIKKEI	204,78%	181,86%	32,27%	38,39%	55,52%	75,72%
NYSE	30,89%	56,25%	47,29%	134,82%	239,16%	113,71%
SEOUL	28,07%	49,43%	100,63%	63,97%	184,32%	202,34%
SHANGHAI	137,06%	141,47%	65,69%	122,01%	93,63%	27,01%
SP500	25,27%	24,50%	71,99%	38,61%	218,94%	256,24%
SM	24,60%	76,25%	150,73%	150,86%	138,38%	77,21%
Promedio	58,94%	71,70%	113,13%	106,38%	135,63%	127,33%

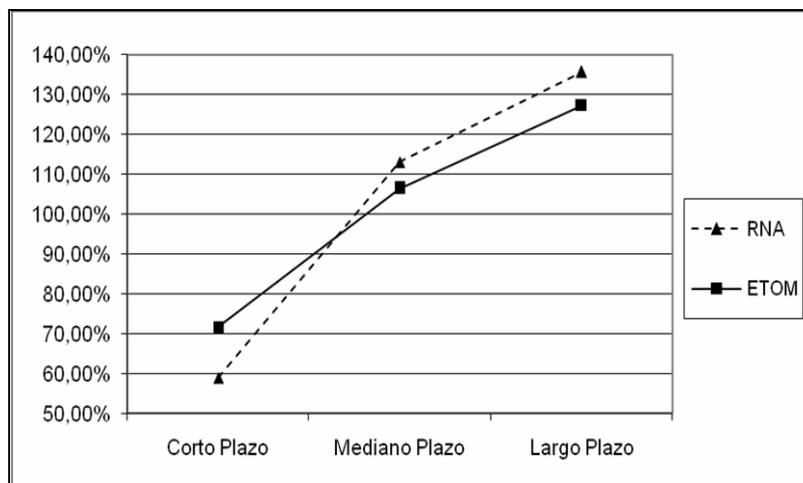


Fig. 2. Estabilidad de las predicciones de cada método

Para comprender la procedencia de los valores presentados en la tabla 3, considérese el siguiente ejemplo: el valor de la primera fila y primera columna (esquina superior izquierda), igual a 65.67%, proviene de dividir 83180 (el ECM de la predicción a corto plazo de la serie IBC utilizando RNA) entre 126670 (el ECM de la predicción global de la serie IBC utilizando RNA), y luego multiplicando el resultado por 100%. De esta tabla, la información más interesante se encuentra en la última fila, en donde se promedian los porcentajes asociados a cada método, para cada plazo.

Cabe destacar que estos promedios no representan la exactitud de cada método en cada plazo, por ejemplo, el que se haya obtenido un promedio del 58,94% para las RNA en las predicciones a corto plazo, y un promedio del 71,70% para los modelos ETOM en las mismas predicciones, no significa que las RNA hayan sido más exactas en las predicciones a corto plazo, de hecho, en este horizonte, los ETOM superaron a las RNA en 12 de 15 series. Este 58,94% sólo significa que en promedio, el ECM de las RNA a corto plazo fue igual al 58,94% del ECM de las pre-

dicciones globales de las RNA. La interpretación de estos porcentajes se puede comprender más fácilmente mediante el gráfico de la Fig. 2.

Del gráfico de la Fig. 2 se obtiene un resultado que se traduce en otra ventaja de los modelos ETOM, además de su exactitud. En este caso, la conclusión es que, a medida

que se alejan las predicciones, la pérdida de exactitud de los modelos ETOM es menor que la de las RNA. Eso puede ser deducido de las pendientes de los gráficos de cada método, ya que mientras más grande sea una pendiente, mayor será la pérdida de exactitud de un método.

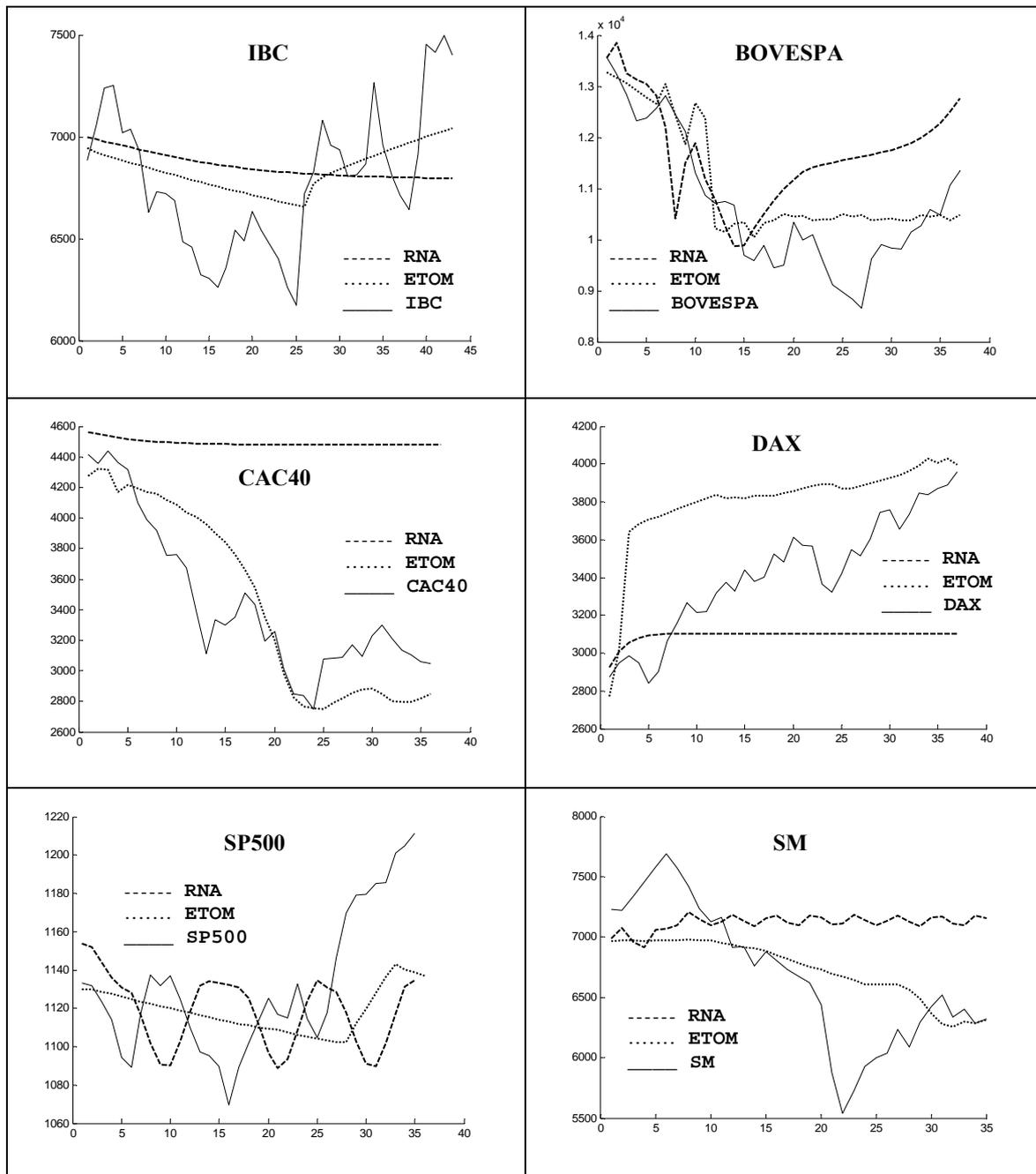


Fig. 3. Gráficos de las predicciones de los mejores modelos

5.2 Comparación gráfica entre las predicciones de los ETOM y las RNA

Además de las comparaciones antes mostradas, se realizó una comparación entre las RNA y los modelos ETOM por medio de sus gráficas. En las gráficas se puede analizar la dinámica de las predicciones más que su exactitud, para comprobar si las tendencias de las predicciones ocurren en el momento adecuado, y en general, la dirección que toman estas predicciones a lo largo del tiempo. En la Fig. 3 se muestran los gráficos de algunas de las predicciones obtenidas, seleccionadas aleatoriamente.

En los gráficos de la Fig. 3 se puede apreciar que, en general, los modelos ETOM generaron predicciones cuyos patrones y tendencias se asemejan a las verdaderas series en mayor medida que las predicciones arrojadas por las RNA, por lo tanto, se puede afirmar que los ETOM poseen una mayor capacidad para capturar las dinámicas de series de

tiempo financieras. Vale la pena mencionar, que los modelos ETOM tienen una capacidad para predecir las tendencias y los patrones de una manera aún más acertada que la que se obtuvo en los experimentos asociados a la Fig. 3, pero esta capacidad no pudo ser aprovechada al máximo debido a que el objetivo era minimizar el error de generalización durante el proceso de búsqueda del mejor modelo, es decir, el grado de similitud entre las predicciones y las verdaderas series fue ignorado durante el proceso de búsqueda de los mejores modelos. En muchas ocasiones se observó que las predicciones de los ETOM presentaban una gran capacidad de predicción de los patrones, pero en una posición que afectaba significativamente al error de generalización, es decir, se predecían muy bien las tendencias a las alzas y a las bajas (los momentos en los que ocurrían y el tiempo que durarían) pero a una distancia considerable de la verdadera curva. Esto ocurrió en una gran cantidad de corridas, entre las cuales se encuentran las mostradas en los gráficos de la Fig. 4.

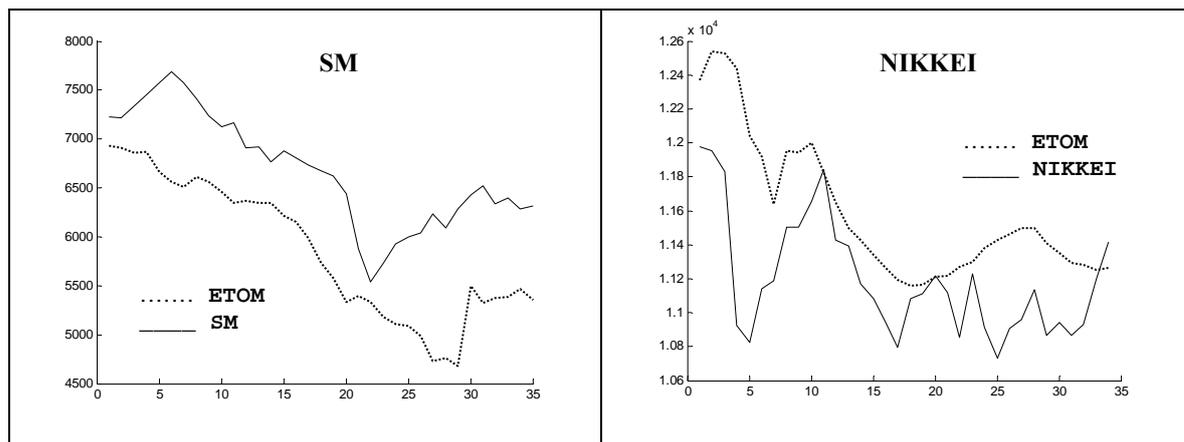


Fig. 4. Gráficos de predicciones con modelos ETOM alternativos

Los gráficos de la Fig. 4 corresponden a predicciones generadas por modelos ETOM distintos a los mejores modelos encontrados para las series SM y NIKKEI. En ambos casos, se puede observar que existe una distancia (entre la verdadera serie y la predicción) que perjudica la calidad de los resultados, y es por ello que los modelos asociados a estos gráficos no llegaron a ser los mejores modelos para estas series. Sin embargo, resulta interesante la similitud entre las curvas predichas y las verdaderas curvas, mostrándose la capacidad de los modelos ETOM en la captura de los patrones y la dinámica de series de tiempo financieras.

6 Conclusiones

Con respecto a la exactitud de las predicciones, se puede concluir que, en general, los modelos ETOM producen mejores predicciones que las RNA.

Además de presentar un menor error de generalización, una de las virtudes más importantes de los modelos ETOM es que hacen predicciones cuyos patrones se asemejan significativamente a los de la serie real, al menos hacen esto mejor que las RNA, incluso en los casos en los cuales estos dos métodos presentan un ECM del mismo orden. De hecho, esta característica estuvo presente, incluso, en los modelos ETOM que poseían un ECM alto, ofreciendo predicciones eficientes de las tendencias de las series, tanto en los momentos en los que ocurrirían como la duración que tendrían, y muchas veces hacían esto mejor que los modelos que produjeron el menor error de generalización. Esto demuestra el gran potencial de los modelos ETOM, es decir, lo que lleva a pensar que los resultados de este enfoque pueden ser aún mejores que los actuales, si se introducen nuevos elementos en la estructura del modelo que logren aprovechar esta capacidad al máximo.

A pesar de lo antes expuesto, no se puede afirmar que los modelos ETOM son absolutamente mejores que las RNA, menos en un caso particular como el tratado, por lo que no se puede descartar el uso de las RNA puras en la predicción de series de tiempo financieras. Las RNA presentan ventajas en lo que respecta a la implementación, permitiendo experimentar más fácilmente con los datos. Además, cuando se está experimentando con RNA no es necesario evaluar tantos modelos para llegar al mejor, como ocurre con los ETOM, aparte de que el entrenamiento de una RNA dura menos tiempo que el entrenamiento de un modelo ETOM. En este sentido, las RNA consumen menos recursos computacionales y un menor esfuerzo por parte del investigador.

Agradecimiento

Los autores desean agradecer al CDCHT de la Universidad de Los Andes (Venezuela), por su apoyo a través de los proyectos E-252-06-02-B e I-1047-07-02-C.

Referencias

Burton M, 1996, A, Random Walk Down Wall Street. Editorial Norton. New York.

Chorafas D, 1994, Chaos Theory in the Financial Markets. Editorial Irwin.

Gupta M, Jin, L y Homma N, 2003, Static And Dynamic Neural Networks. Editorial John Wiley And Sons. New York.

Jang J, Sun C y Mizutani E, 1997, Neuro-Fuzzy and Soft Computing. Editorial Prentice Hall. Englewood Cliffs, NJ.

Lijuan C, 2003, Support vector machines experts for time series forecasting. Neurocomputing, Vol.51 pp. 321-339.

Lynch P y Rothschild J, 1996, Learn to Earn. Editorial John Wiley & Sons. New York.

Rabiner L, 1989, A Tutorial on Hidden Markov Models. Proceedings of the IEEE, Vol. 77, No.2. pp. 257-286.

Wang X, Whirgham P y Deng D, 2003, Time-Line Hidden Markov Experts and its Application in Time Series Prediction. Universidad de Otago.

Wei H, Yoshiteru N y Shou-Yang W, 2005, Forecasting stock market movement direction with support vector machine. Computers & Operations Research, Vol. 32 pp. 2513-2522.

Witten I y Frank E, 2005, Data Mining. Practical Machine Learning Tools and Techniques. Editorial Elsevier. E.E.U.U.