

Modelos de redes neuronales aplicados a la predicción del tipo de cambio del dólar observado en Chile

Antonino Parisi F.

Franco Parisi F.

Universidad de Chile

José Luis Guerrero

Georgetown University

Extracto

Este estudio analiza la capacidad de las redes neuronales para predecir el signo de las variaciones diarias del dólar observado, entendiendo que la predicción de la dirección del movimiento es relevante para desarrollar estrategias de transacción efectivas (Leung, Daouk y Chen, 2000). Se usaron modelos de redes neuronales de algoritmo de aprendizaje supervisado de propagación hacia atrás: el Perceptrón Multicapa, la Red Recurrente Jordan-Elman y la Red Ward, con funcionamiento estándar y recursivo. La evaluación se hizo sobre la base de un conjunto extramuestral de 398 datos y sobre 337 subconjuntos extramuestrales, de 60 observaciones diarias cada uno. El desempeño relativo de los

Los autores desean agradecer a Edinson Cornejo, por su labor como asistente de investigación, a los comentarios realizados por Jorge Gregoire, Fernando Bravo, Jorge Niño, Alex Silemberger, y también agradecen a un árbitro anónimo por sus valiosos comentarios y a los participantes en el seminario realizado en el Departamento de Administración de la Universidad de Chile

modelos fue medido por el número de predicciones correctas (*hits*) del signo de la variación del tipo de cambio, aplicando para ello el *test* de acierto direccional de Pesaran y Timmermann (1992). A su vez, los resultados de la mejor red neuronal fueron contrastados con los de un modelos ingenuo y con los de una estrategia *buy and hold*. La red ward de tres capas obtuvo el mejor rendimiento y su capacidad para predecir el signo de las variaciones diarias del tipo de cambio del dólar observado resultó significativa para el período analizado. Sin embargo, al utilizar un proceso recursivo para re-estimar período a período los pesos de la red no se generaron mejoras significativas. Además, al contrastar el porcentaje de predicción de signo (PPS) de la red ward estándar y recursiva con el de un modelo ingenuo, la de este último fue significativamente mayor. No obstante, la capacidad predictiva de los modelos no se tradujo en retornos significativamente más altos respecto a una estrategia pasiva, a pesar de que los retornos anormales promedios fueron positivos. Los resultados se adecuan a lo esperado para el caso chileno, dado el bajo número de agentes que participan en este mercado, la alta volatilidad originada en sendas crisis financieras internacionales y los significativos cambios en el mercado a partir de la participación del Banco Central.

Abstract

This study analyzes the forecast ability of the sign variation of the daily exchange in neural networks. The forecast of price's direction is important to effective trade strategies (Leung, Daouk & Chen, 2000). We used backpropagation neural network models: the Multilayer Perceptron, the Jordan-Elman Recurrent network and the Ward network, all with standard and recursive method. To test the forecast ability we use an out-of-sample set of 398 observations and over 337 out-of-sample sub-set of 60 observations. The relative performance of the models was measured by the number of hits of the exchange rate variation sign, using for that the Pesaran & Timmermann's (1992) *directional accuracy test*. Then, the results of the best neural network model was compared with the ones of AR(1) model and with a *buy and hold* strategy. The neural network of three layers had the best performance and it forecast ability was significant in statistical and economical terms. When we used a recursive process for re-estimate period to period the weights of the network, we do not find a better performance. In addition, we compare the percentage of sign prediction between ward standard network, ward recursive network and the AR(1) model, and the last

was significantly higher. However, the model's forecast capacity did not generate higher significant returns, in relation to a passive investment strategy, even when the abnormal average returns was positive. The results were the expected for the Chilean case due to: the few agents that have this market, the high volatility of rate exchange as consequences of international financial crisis, and the significant changes in the market produced by the Central Bank's interventions.

1. Introducción

Este estudio analiza la capacidad de los modelos de redes neuronales para predecir el signo de las variaciones diarias del dólar observado en Chile, y la compara con la capacidad predictiva de un modelo AR(1) no optimizado, donde ambos modelos fueron aplicados de manera estándar y recursiva. Esta investigación continúa la línea de otros estudios como el de Min Qi (2001), donde se utilizan modelos de redes neuronales con funcionamiento recursivo para la predicción de recesiones en Estados Unidos, y cuyos resultados muestran que el *spread* de la tasa de interés, los indicadores principales del departamento de comercio y el índice S&P500 son útiles para ese fin. Wilson y Sharda (1994) compararon el rendimiento de las redes neuronales con el análisis discriminante, y encontraron que las redes neuronales fueron superiores en la predicción de firmas en bancarrota. Por su parte, Hill et al. (1994) sostiene que el poder de predicción de las redes neuronales, en comparación con los modelos estadísticos estándares, puede de hecho estar en un contexto específico y que ninguna técnica domina claramente a la otra. Sin embargo, dicho estudio sugiere que las redes neuronales artificiales tienden a predecir mejor que los métodos estadísticos alternativos cuando los datos son de tipo financieros, estacionales y no lineales.

Continuando con la línea del estudio anterior, Callen et al. (1996) compararon la habilidad relativa de los modelos de redes neuronales y de los modelos de series de tiempo lineales para predecir las ganancias contables trimestrales para una gran cantidad de firmas que transan en la New York Stock Exchange, concluyendo que los primeros no son necesariamente superiores a los segundos, aún cuando los datos sean financieros, estacionales y no lineales, lo cual contrasta con los resultados

obtenidos en el estudio mencionado anteriormente. Por su parte, Dasgupta et als. (1994) compararon las redes neuronales con otras técnicas como el análisis discriminante y el modelo de regresión *logistic*, con el objeto de medir su capacidad para identificar y clasificar consumidores basados en su disposición para tomar riesgos financieros y comprar productos de inversión no tradicionales. En dicho estudio, las redes neuronales superaron a las otras técnicas en términos de porcentaje de predicción.

La mayoría de las prácticas de transacción adoptadas por los analistas financieros confían en la predicción precisa de los niveles de precios de los instrumentos financieros. No obstante, estudios recientes han sugerido que las estrategias de transacción basadas en la proyección de la dirección del cambio en el nivel de precios son más efectivas, y pueden generar mayores beneficios, que aquellas basadas en una predicción puntual. En esta línea, Leung, Daouk y Chen (2000) compararon la capacidad predictiva de los modelos de clasificación¹ con los de estimación de nivel², y concluyeron que los primeros (los cuales fueron propuestos para predecir el signo o la dirección de las variaciones de los índices bursátiles S&P500, FTSE100 y Nikkei 225) se desempeñan mejor que los segundos, en términos de su tasa de acierto, y son capaces de generar beneficios más altos. Maberly (1986) analizó la relación entre la dirección de los cambios de precio interdía e intradía, y O'Connor, Remus y Griggs (1997) apoyan la utilidad de proyectar la dirección del cambio en el nivel de precios más que el nivel del precio en sí.

Tal como lo indica Parisi, Parisi y Cañas (2002), el mercado chileno del dólar se compone de varios operadores que transan dólares entre ellos, los que componen el "Mercado Cambiario Formal" (MCF). Los operadores más importantes son los bancos, y es por ello que es común referirse al tipo de cambio en este mercado como precio del dólar *interbancario*³. En Chile, existen cerca de 20 instituciones bancarias que transan dólares y activos derivados de la divisa, destacando el Citibank, el Chase Manhattan Bank y el Banco Santander, entre los operadores

¹Linear discriminant analysis, logit model, probit model y probabilistic neural network

²Adaptive exponential smoothing, vector autoregression with Kalman filter, multivariate transfer function y multilayered feedforward neural network.

³Precio asignado al intercambio de dólares entre mesas de dinero bancarias.

internacionales. En tanto, por parte de los operadores nacionales se encuentran el Banco de Chile, Banco Edwards y el Banco Santiago.

En un día promedio, se transan cerca de US\$ 500 millones en el mercado formal que opera de 10:00 a 13:00 horas, de lunes a viernes. Cabe señalar que los bancos extranjeros continúan operando, pero fuera de horario. Los operadores bancarios transan mediante dos formas, por teléfono y por medio de computadores, existiendo varios software para ello. El software usado en Chile para el intercambio de dólares se llama Datatec, y sólo los bancos tienen acceso a éste. Aunque para los bancos el precio que rige en todo momento es el que aparece en el Datatec, para los demás actores del mercado de divisa el precio de referencia es el que publica Reuters, además de otras agencias de información tales como El Mercurio Online, Bridge, Dow Jones, o Bloomberg⁴. Un problema que surge en este mercado es que las agencias informativas logran los precios de la divisa por medio telefónico al llamar a los bancos, información que luego es entregada a los demás agentes del mercado, creándose un desfase entre el precio que aparece en el Datatec y el proporcionado al resto del mercado. Además, muchas veces los operadores tratan de manejar el precio al mentirles a las agencias de "a cuánto" están transando. Por lo tanto, existe un diferencial entre el precio al que transan los bancos y el precio al que pueden llegar a transar las empresas y otros miembros que se nutran de las agencias informativas.

Es conveniente precisar, al igual que Parisi, Parisi y Cañas (2002), que el Banco Central de Chile dejó de intervenir formalmente el mercado de la divisa en Septiembre de 1999. Luego, indirectamente intervino por medio de la venta de Bonos Reajustables en Dólares Estadounidenses (PRDs) justo cuando eventos internacionales e iliquidez interna generaron un alza en el tipo de cambio. Antes de la fecha indicada, existían bandas de precios para la divisa evitando que el tipo de cambio subiese o bajase de ciertos niveles en períodos cortos de tiempo, mecanismo conocido como "flotación sucia". En ocasiones en que la divisa tocaba un límite de la banda, el Banco Central debía intervenir al vender o comprar dólares alterando sus reservas internacionales, presionando el tipo de cambio al interior de la banda. El 16 de agosto de 2001, el Banco Central reimpone la

⁴Para mayor detalle ver Parisi, Kohn y Castillo (2001).

intervención en el mercado cambiario, con el propósito de reducir los efectos en el tipo de cambio producto de la crisis de Argentina.

El presente estudio sobre el mercado cambiario chileno se concentró en el empleo de las redes neuronales multicapas con aprendizaje supervisado, específicamente en el *perceptrón multicapa* y algunas de sus variantes. Además, se estimaron los resultados generados a partir de un modelo AR(1). Estos modelos fueron aplicados en forma estándar y recursiva, evaluándose la capacidad predictiva de cada uno de ellos a través del *test de precisión direccional*. La importancia de este trabajo radica en la posibilidad de encontrar un modelo que permita predecir el comportamiento diario del tipo de cambio \$/US\$. Esta variable juega un papel sumamente importante en la economía chilena, afectando no sólo a las empresas que participan en el sector exportador e importador, sino que también a la liquidez del sistema financiero nacional, como consecuencia de las operaciones especulativas que se llevan a cabo con la divisa.

La red ward de tres capas obtuvo el mejor rendimiento y su capacidad para predecir el signo de las variaciones diarias del tipo de cambio del dólar observado resultó significativa para el período analizado. Sin embargo, al utilizar un proceso recursivo para re-estimar período a período los pesos de la red no se generaron mejoras significativas. Además, al contrastar el PPS de la red ward estándar y recursiva con el de un modelo ingenuo, la de este último fue significativamente mayor. No obstante, la capacidad predictiva de los modelos no se tradujo en retornos significativamente más altos respecto a una estrategia pasiva, a pesar de que los retornos anormales promedios fueron positivos. El documento se divide en cuatro secciones: la sección dos presenta la metodología empleada en el estudio; la sección tres aborda el análisis de los resultados y, finalmente, la sección cuatro presenta las conclusiones del estudio.

2. Metodología y datos

Se usaron 1.201 observaciones de precios de cierre diarios de las variables incluidas en el modelo, correspondientes al período comprendido entre el 08 de Octubre de 1996 y el 23 de Mayo de 2001. En el modelo formulado, la variable de salida está dada por la variación diaria del dólar observado

correspondiente al período t , mientras que las variables de entrada se refieren a las variaciones de un conjunto de variables que se detallan más adelante. El modelo incluye rezagos de las variables de entrada, para así considerar el desfase lógico que existe entre el momento en que se realiza la predicción, el período actual o $t-1$, y el momento futuro para el cual la proyección es válida, al que nos referimos como período t . Así, debido a que la información de hoy está desfasada con respecto a la proyección, se utiliza un modelo dinámico con rezago en las variables independientes.

Las redes neuronales pueden entenderse como modelos multiecuacionales o multi-etapas, en los que el *output* de unos constituye el *input* de otros. En el caso de las redes multicapas, existen etapas en las cuales las ecuaciones operan en forma paralela. Las redes neuronales, al igual que, por ejemplo, los modelos de suavizamiento exponencial y de análisis de regresión, utilizan *inputs* para generar un *output* en la forma de una proyección. La diferencia radica en que las redes neuronales incorporan inteligencia artificial en el proceso que conecta los *inputs* con los *outputs*, Kuo y Reitsch, (1995-96).

Según Martín del Brío y Sanz (1997), las redes neuronales artificiales "son sistemas de procesamiento que copian esquemáticamente la estructura neuronal del cerebro para tratar de reproducir sus capacidades". En consecuencia, son una clase de modelos no lineales flexibles que se caracterizan por ser sistemas *paralelos*⁵, *distribuidos*⁶ y *adaptativos*⁷, donde estos tres conceptos se traducen en un mejor rendimiento y en una mayor velocidad de procesamiento. Por su parte, Herbrich, Keilbach, Graepel, Bollmann-Sdorra y Obermayer (2000) señalan que la característica más importante de las redes neuronales es su capacidad para aprender dependencias basadas en un número finito de observaciones, donde el término *aprendizaje* significa que el conocimiento adquirido a partir de las muestras puede ser empleado para proporcionar una respuesta correcta ante datos no utilizados en el entrenamiento de la red. La literatura sugiere que las redes neuronales poseen varias ventajas potenciales sobre los métodos

⁵Cuentan con una gran cantidad de neuronas o procesadores elementales (PE), cada uno de los cuales trabaja paralelamente con una pequeña parte de un problema mayor.

⁶Cuentan con muchos PEs a través de los cuales distribuyen su memoria.

⁷Tienen la capacidad de adaptarse al entorno modificando sus pesos y sinapsis de manera de encontrar una solución aceptable al problema.

estadísticos tradicionales, destacándose el que éstas pueden ser aproximaciones de funciones universales aún para funciones no lineales, Hornik, Stinchcombe y White, (1989), lo que significa que ellas pueden aproximar automáticamente cualquier forma funcional (lineal o no lineal) que mejor caracterice los datos, permitiéndole a la red extraer más señales a partir de formas funcionales subyacentes complejas, Hill, Márquez, O'Connor y Remus, (1994).

En este estudio se emplearon diferentes modelos de redes neuronales que usan el algoritmo de aprendizaje supervisado de propagación hacia atrás⁸, con el objeto de predecir el signo de las variaciones diarias del tipo de cambio "\$/US\$". Los modelos analizados fueron: el Perceptrón Multicapa, la Red Recurrente Jordan-Elman y la Red Ward.

Se utilizó el software de redes neuronales *Neuroshell 2*, el cual requiere para su funcionamiento que la base de datos se divida en tres conjuntos diferentes: (a) conjunto de entrenamiento (conjunto de datos empleados para que la red aprenda el problema), (b) conjunto de prueba (conjunto de datos utilizados para evitar el sobreaprendizaje⁹ de la red), y (c) conjunto de producción o extramuestral (datos no incorporados en los conjuntos anteriormente mencionados, los cuales fueron usados para probar la capacidad de predicción de la red ante datos nunca vistos)¹⁰. Los datos del conjunto de entrenamiento y de prueba se eligen aleatoriamente, cubren el periodo comprendido entre el 08 de Octubre de 1996 y el 10 de Noviembre de 1999, y cuentan en un principio con 400 datos cada uno. Por su parte, el conjunto extramuestral posee inicialmente 398 datos y cubre el

⁸Estos modelos se caracterizan porque en ellos el resultado es conocido y la red se entrena a sí misma hasta que es capaz de predecir aquel resultado asociado con los datos de entrada (Dasgupta, Dispensa y Ghose, 1994).

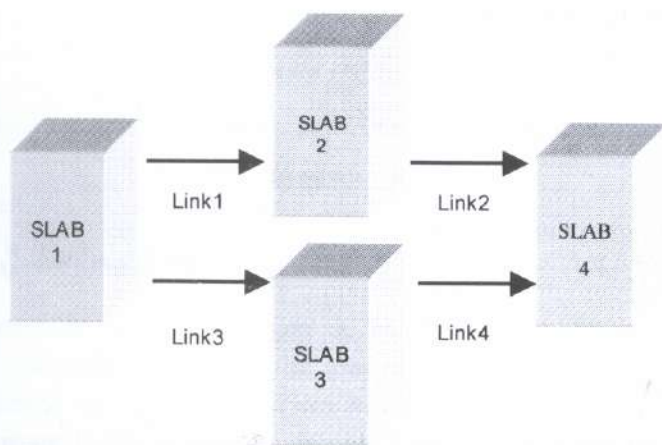
⁹El sobreajuste a los datos de la muestra o sobreaprendizaje de la red se produce cuando el sistema se ajusta demasiado a los datos de entrenamiento, aprendiendo incluso el ruido presente en ellos, por lo que crece el error ante patrones diferentes a los empleados en el entrenamiento y disminuye la precisión de la proyección).

¹⁰Cabe mencionar que otros autores dividen el conjunto total de datos en dos segmentos, donde el primero es utilizado para entrenar la red mientras que el segundo es usado para determinar y validar la arquitectura y la especificación del modelo (Chen y Leung, 1998; Leung, Daouk y Chen, 2000). En este sentido, McCann y Kalman (1994), al predecir la tendencia del precio del oro, dividieron el conjunto total de datos en dos grupos: uno para entrenar la red y otro para probarla.

periodo comprendido entre el 11 de Noviembre de 1999 y el 23 de Mayo de 2001¹¹. No obstante, con el objeto de analizar trimestralmente la estabilidad de la capacidad predictiva del mejor modelo, a partir de los 398 datos del conjunto extramuestral se constituyeron 337 subconjuntos de producción de 60 observaciones diarias cada uno. Con respecto a las limitaciones que presenta el software, estas se refieren principalmente a (a) la gran cantidad de datos que se requiere para conformar los conjuntos de entrenamiento, de prueba y extramuestral, y (b) la falta de una función que permita realizar el proceso de recursividad en forma automática.

Luego de probar con numerosas arquitecturas y parámetros, los mejores resultados en términos de capacidad predictiva se obtuvieron con la Red Ward de tres capas: una capa de entrada, una oculta y otra de salida (ver figura 1).

Figura 1
Red Ward



En términos generales, la capa de entrada (Slab 1) posee una neurona por cada *i*-ésima variable de entrada, la capa oculta posee “*j*” neuronas (las cuales se dividen en partes iguales entre los Slabs 2 y 3), y la capa de salida

¹¹Los datos del conjunto de entrenamiento, de prueba y extramuestral suman 1.198, debido a la pérdida de 3 observaciones como consecuencia de haber usado primeras diferencias y un máximo de dos rezagos.

(Slab 4) posee una neurona. De esta forma, la salida neta de la capa oculta viene dada por la ecuación (1):

$$i_{pj}^h = f_j^h (Neta_{pj}^h), \quad \text{con } j=1,2,\dots,n. \quad (1)$$

donde la expresión (2) se emplea para los Slabs 2 y 3, respectivamente:

$$Neta_{pj}^h = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1i} \\ w_{21} & w_{22} & \dots & w_{2i} \\ \vdots & \vdots & \dots & \vdots \\ w_{\lfloor \frac{j}{2} \rfloor 1} & w_{\lfloor \frac{j}{2} \rfloor 2} & \dots & w_{\lfloor \frac{j}{2} \rfloor i} \end{bmatrix}^h \times \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \end{bmatrix}_{(i \times 1)} + \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{\lfloor \frac{j}{2} \rfloor} \end{bmatrix}^h = \begin{bmatrix} Neta_{p1} \\ Neta_{p2} \\ \vdots \\ Neta_{p\lfloor \frac{j}{2} \rfloor} \end{bmatrix}^h \quad (2).$$

La salida neta de la capa de salida queda expresada de acuerdo a la ecuación (3), para los Slabs 2 y 3:

$$o_{pk} = f_k^o (Neta_{pk}^o) \quad (3)$$

donde:

$$Neta_{pk}^o = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{\lfloor \frac{j}{2} \rfloor} \end{bmatrix}^o \times \begin{bmatrix} i_1 \\ i_2 \\ \vdots \\ i_{\lfloor \frac{j}{2} \rfloor} \end{bmatrix}^h + [\theta_1]_{(1 \times 1)}^o = [Neta_1]_{(1 \times 1)}^o \quad (4).$$

Para una explicación más detallada del algoritmo, ver el anexo 1. Las variables de entrada o independientes utilizadas para explicar la variación diaria del tipo de cambio son: la variación diaria del tipo de cambio "\$/US\$" rezagada (Tipo), la variación diaria de los índices bursátiles Bovespa, Merval, Nikkei 225 (Nikkei), Dow Jones Industry (Dow) e IGPA, y la variación de la tasa de interés interbancaria promedio diaria (TI). De esta manera se busca capturar el efecto de las variaciones pasadas del tipo de cambio, de los cambios experimentados por los principales centros

bursátiles internacionales (Brasil, Argentina, Japón y Estados Unidos) y por el mercado de valores local, y de las intervenciones en el mercado de divisas realizadas por el Banco Central a través de la tasa de política monetaria (TPM), la cual impacta en la tasa de interés interbancaria. La forma funcional del modelo se expresa en la ecuación (5):

$$\delta Tipo_t = g(\delta Tipo_{t-1}, \delta Tipo_{t-2}, \delta Bovespa_t, \delta Merval_t, \delta Nikkei_t, \delta DowJ_t, \delta GPA_t, \delta TL_t) \quad (5)$$

donde:

$$\delta Tipo_t = \frac{Tipo_t - Tipo_{t-1}}{Tipo_{t-1}} \quad (6).$$

Las características de la red ward usada para proyectar el signo de la variación diaria del dólar observado se presentan en el cuadro 1.

Cuadro 1
Arquitectura de la red ward.

Funciones de activación y N° de neuronas por Slab							
Slab 1		Slab 2		Slab 3		Slab 4	
Lineal[-1,1]		Gaussiana		Gauss. Complemento		Logistic	
8		8		8		1	
Tasas de aprendizaje y de momento.							
Link 1		Link 2		Link 3		Link 4	
Aprendizaje	Momento	Aprendizaje	Momento	Aprendizaje	Momento	Aprendizaje	Momento
0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

Al momento de predecir el signo de las variaciones que componen el conjunto extramuestral, además de usar el método estándar se aplicó un procedimiento recursivo. Esta última metodología ha sido empleada para medir el desempeño de modelos de redes neuronales que buscan predecir periodos de recesión en los Estados Unidos (Qi, 2001; Estrella & Mishkin, 1998). De esta forma, si consideramos que el conjunto extramuestral tiene *m* datos, al evaluar el funcionamiento de la red se consideró sólo la predicción del primer valor. Posteriormente, el dato analizado sale del conjunto extramuestral (el cual queda con *m-1* datos) y pasa a formar parte del conjunto de entrenamiento, siguiendo un proceso aleatorio, por lo que la muestra de *n* datos que contiene los valores de entrada se incrementa a

$n+1$. Luego se realizaron nuevas iteraciones y, en consecuencia, los pesos del modelo se re-estimaron en cada una de las $m-1$ proyecciones. Este proceso se repitió hasta que en el conjunto extramuestral quedó sólo un dato, lo que le permitió a la red aprender del error cometido en la predicción y ajustar los pesos estimados para los valores de entrada de las capas de salida y oculta (w_{jk} y w_{ij} , respectivamente), recalculándolos hasta el momento en que la última observación (correspondiente al período $t-1$) fue empleada para proyectar el valor que la variable de salida podría alcanzar en el momento t , el cual representa el futuro inmediato.

El proceso de evaluación empírica de la red neuronal y del modelo AR(1) se hizo sobre la base de los datos del conjunto extramuestral (compuesto de 398 observaciones diarias) y en función del desempeño de los modelos a lo largo de los 337 subperíodos extramuestrales (compuestos de 60 datos diarios cada uno). En esta etapa, el desempeño relativo de los modelos fue medido por el número de predicciones correctas (*hits*) del signo de la variación del tipo de cambio. Así, al igual que Kanas (2001), se aplicó el test de *certeza direccional* (DA) de Pesaran & Timmermann (1992) a fin de medir la precisión direccional de los modelos de proyección. El test compara el signo de la proyección, \hat{y}_{n+i} , con el del valor observado, y_{n+i} , para cada i -ésima observación del conjunto extramuestral ($i=1,2, \dots, m$), donde el signo indica la dirección en que se moverá el tipo de cambio: al alza si es positivo, o a la baja si es negativo. Si los signos coinciden, aumentan la efectividad de la red y, en caso de no existir coincidencia, aumenta el error de predicción de la red.

El ratio de éxito (SR) se define como:

$$SR = m^{-1} \sum_{i=1}^m I_i [y_{n+i}, \hat{y}_{n+i} > 0] \quad (7)$$

donde $I_i[\cdot]$ es una función indicador que toma el valor de 1 cuando su argumento es cierto y 0 en otro caso. Además:

$$P = m^{-1} \sum_{i=1}^m I_i [y_{n+i} > 0] \quad (8)$$

y

$$\hat{P} = m^{-1} \sum_{i=1}^m I_i [\hat{y}_{n+i} > 0] \quad (9).$$

El ratio de éxito en el caso de independencia de \hat{y}_{n+i} y y_{n+i} , SRI, está dada por:

$$SRI = P \cdot \hat{P} + (1 - P)(1 - \hat{P}) \quad (10)$$

cuya varianza es:

$$VAR[SRI] = m^{-2} [m(2 \cdot \hat{P} - 1)^2 P(1 - P) + m(2 \cdot P - 1)^2 \hat{P}(1 - \hat{P}) + 4 \cdot P \cdot \hat{P}(1 - P)(1 - \hat{P})] \quad (11).$$

Por su parte, la varianza de SR se define como:

$$VAR[SR] = m^{-1} \cdot SRI(1 - SRI) \quad (12).$$

Finalmente, el *test* DA de Pesaran y Timmermann (1992) está dado por:

$$DA = (VAR[SR] - VAR[SRI])^{-1/2} (SR - SRI) \quad (13).$$

Pesaran y Timmermann (1992) mostraron que, bajo la hipótesis nula de que \hat{y}_{n+i} y y_{n+i} están independientemente distribuidos, dicho *test* sigue una distribución normal estándar. También se utilizó el *test* de Friedman y el *test* de Tukey, los cuales permitieron determinar la existencia de diferencias significativas entre las frecuencias y las rentabilidades de los modelos de predicción.

La capacidad de los modelos de redes neuronales para predecir el signo de la variación del tipo de cambio, junto con la rentabilidad que un inversionista habría obtenido de haber seguido las recomendaciones de compra y venta arrojadas por el modelo, fue contrastada con la capacidad predictiva y con el retorno de un modelo AR(1) no optimizado. A su vez,

los resultados de ambos fueron comparados con la rentabilidad generada por una estrategia de inversión pasiva o *buy and hold*¹².

El modelo AR(1), estimado en forma estándar y recursiva, fue el siguiente:

$$\delta Tipo_t = \alpha_1 \cdot \delta Tipo_{t-1} + \varepsilon_t \quad (14)$$

donde α_1 corresponde al coeficiente que acompaña a la variable independiente y es estimado por el modelo, mientras que ε_t representa el término de error de la regresión.

3. Análisis de los resultados

El cuadro 2 presenta un resumen de los estadísticos de las series consideradas en el estudio. Para las variables que presentan un coeficiente de asimetría positivo, la distribución está desplazada o presenta un sesgo hacia la derecha como consecuencia de la existencia de valores extremos y poco usuales. Las series de variaciones de precios de cierre de los índices DJI y Merval, que evidencian un coeficiente de asimetría negativo, tienen un sesgo hacia la izquierda.

Por otra parte, el valor de la kurtosis fue superior a 3 para todas las variables. En consecuencia, la distribución de las observaciones adoptó una forma leptokúrtica, con las observaciones concentradas en un estrecho rango de valores y con una masa relativamente baja en las colas de las funciones de densidad. De acuerdo a los resultados del *test de Jarque-Bera*, en todos los casos se rechazó la hipótesis nula de que las variaciones de los valores de cierre diarios siguen una distribución normal, con un nivel de significancia del 5%. Además, al analizar la serie del tipo de cambio, los coeficientes de autocorrelación para AR(1) y AR(2) resultaron ser estadísticamente significativos y no significativos, respectivamente, al nivel de significancia del 5%.

¹²La estrategia "*buy and hold*" es una estrategia de inversión pasiva que no aplica gestión alguna. Consiste en comprar el activo en cuestión y mantenerlo durante todo el período de inversión, para venderlo al final de este. La rentabilidad de esta estrategia está dada por la diferencia entre el valor de venta y el valor de compra del activo analizado.

Cuadro 2

Estadísticos descriptivos de la serie de variaciones de valores de cierre diarios.
Entre paréntesis se entrega el test $-t$. Periodo: 08 de Octubre 1996 al 23 de Mayo de 2001.
1200 observaciones

ESTADÍSTICOS	TIPO CAMBIO	BOVESPA	DOW	NIKKEI	IGPA	MERVAL	TASA INTERBANCARIA
Media	0.000325	0.001043	0.000587	-0.000219	-0.0000260	-0.00000235	0.022496
Mediana	0.0000187	0.000343	0.000368	0.000000	-0.0000686	0.000000	0.000000
Máximo	0.024982	0.334190	0.049802	0.079560	0.045659	0.122604	19.75000
Mínimo	-0.023544	-0.158090	-0.071830	-0.069785	-0.036845	-0.137266	-0.935484
Desv. Est.	0.003602	0.027923	0.011745	0.015272	0.007618	0.021894	0.586691
Skewness	0.748358	1.362621	-0.312635	0.203306	0.179843	-0.283317	31.79619
Kurtosis	10.80979	23.87905	6.222948	5.468960	6.688468	8.863449	1066.554
Jarque-Bera	3161.651	22168.07	538.9179	313.0549	686.7086	1735.056	56759546
AR(1)	0.072293 (2.507553)*	-	-	-	-	-	-
AR(2)	0.013172 (0.455522)	-	-	-	-	-	-

Los valores críticos de los estadísticos Jarque-Bera y t de Student son 4.43 y 1.96, respectivamente, para un $\alpha=5\%$.

(*): Significativo al 5%.

El cuadro 3 muestra la matriz de correlación entre las variables. Se observa una correlación levemente positiva entre la variación del tipo de cambio en el período t y la variación del precio de la divisa en el período $t-1$ y $t-2$. A su vez, la variación del tipo de cambio evidencia una correlación negativa con las variaciones de los índices bursátiles Bovespa, DJI, Nikkei, IGPA y Merval, señalando que una caída en estos mercados está asociada a un aumento en el tipo de cambio, y viceversa. Entre las variaciones del tipo de cambio y las fluctuaciones de la tasa de interés promedio interbancaria también existe una correlación negativa.

Al evaluar la capacidad predictiva de los modelos, medida a través del porcentaje de predicción del signo (PPS) de las fluctuaciones diarias del tipo de cambio observado y calculada sobre un conjunto extramuestral de 398 datos, el mejor resultado se obtuvo con el modelo ingenuo o AR(1). El modelo ingenuo estimado en forma estándar arrojó un PPS (60,05%) y una rentabilidad acumulada (14,58%) idéntica a la del modelo ingenuo estimado en forma recursiva. No obstante, a pesar de que el modelo ingenuo evidenció la mayor capacidad predictiva durante el período de tiempo analizado, la red ward estándar generó la mayor rentabilidad acumulada (16,66%). Se observó que todos los modelos analizados registraron un PPS estadísticamente significativo y que, además, generaron una rentabilidad superior a la lograda por una estrategia *buy and hold*. Cabe señalar que la red ward con funcionamiento estándar logró un PPS de 57,54% y que, al considerar un funcionamiento recursivo, este se redujo levemente a un 57,04%. Estos resultados difieren de los obtenidos por Parisi (2002). En dicho estudio, de las diversas arquitecturas de redes neuronales utilizadas para predecir el signo de las variaciones semanales del IPSA, la red ward con funcionamiento recursivo obtuvo el mejor rendimiento, alcanzando una capacidad predictiva extramuestral de 72% y una rentabilidad acumulada extramuestral (para la cartera conformada por el IPSA) de un 24,42%.

Cuadro 3
Matriz de coeficientes de correlación

TC	1	0.07229	0.01317	TC(-2)	BOVESPA(-1)	DOW(-1)	IGPA(-1)	MERVAL(-1)	NIKKEI(-1)	TI(-1)
TC(-1)	0.07229	1	0.07223	1	0.03077	0.04281	-0.05005	-0.00851	-0.04327	-0.02113
TC(-2)	0.01317	0.07223	1	0.30777	1	0.48698	0.44534	0.02884	-0.00056	-0.00394
Bovespa(-1)	-0.09480	-0.02747	0.03077	1	0.48698	1	0.36538	0.67990	0.09949	-0.00020
Dow(-1)	-0.12801	-0.00619	0.04281	0.48698	1	0.48698	1	0.51476	0.07240	0.00606
IGPA(-1)	-0.11578	-0.08050	-0.05005	0.44534	0.36538	0.36538	1	0.44138	0.16642	-0.01670
Merval(-1)	-0.13733	-0.00851	0.02884	0.67990	0.51476	0.51476	0.44138	1	0.10700	0.01842
Nikkei(-1)	-0.15430	-0.04327	-0.00056	0.09949	0.07240	0.07240	0.16642	0.10700	1	-0.06811
TI(-1)	-0.03263	-0.02113	-0.00394	-0.00020	0.00606	0.00606	-0.01670	0.01842	-0.06811	1

Lo anterior llama la atención, debido a que se esperaba un mejor desempeño del modelo recursivo, ya que este permite que los pesos de la red se re-estimen día a día, recogiendo así la nueva información y las nuevas expectativas del mercado. Los resultados se presentan en el cuadro 4.

Cuadro 4
Resumen de medidas estadísticas

PERÍODO EXTRAMUESTRAL COMPLETO DE 398 DATOS				
MODELO	PPS	TEST DA	RENTABILIDAD	σ [TIPO CAMBIO]
R. Ward Estándar	57.54%	2.81	16.66%	0.395%
R. Ward Recursiva	57.04%	2.55	14.99%	0.395%
AR(1) Estándar	60.05%	3.96	14.58%	0.395%
AR(1) Recursivo	60.05%	3.96	14.58%	0.395%
Buy & Hold	-	-	11.75%	-

Medidas estadísticas promedio de los 337 sub-períodos extramuestrales.
Entre paréntesis se entrega la desviación estándar de los estadísticos, calculada considerando los 337 sub-períodos extramuestrales.

MODELO	PPS PROMEDIO	Nº PERÍODOS SIGNIFICATIVOS (*)	RETORNO PROMEDIO	ESTRATEGIA PASIVA
R. Ward Estándar	56.89% (8.07%)	114	2.49% (3.23%)	2.08% (3.98%)
R. Ward Recursiva	57.27% (7.05%)	108	2.45% (3.35%)	2.08% (3.98%)
AR(1) Estándar	59.84% (8.09%)	167	2.43% (2.97%)	2.08% (3.98%)

El test DA se contrastó con un valor crítico de $z=1.645$, para un nivel de significancia del 10%.

(*): Número de períodos extramuestrales (de un total de 31) en los cuales la capacidad predictiva de los modelos resultó ser estadísticamente significativa, de acuerdo al test DA (valor crítico de $z=1.96$; $\alpha=5\%$).

Con el objeto de evitar el problema de *data snooping*¹³, y de despejar las dudas respecto de si la capacidad predictiva se debe a la bondad del modelo, a las características de la muestra de observaciones sobre la que ha

¹³El término *data snooping* (también conocido como *data mining*), "ocurre cuando un determinado conjunto de datos es usado más de una vez para propósitos de inferencia o selección de modelos. Cuando esta re-utilización de datos ocurre, siempre existe la posibilidad de que cualquier resultado satisfactorio que se haya obtenido pueda deberse simplemente a la suerte, en vez de a algún mérito inherente al modelo que generó aquellos resultados." (White, 2000).

sido aplicado o simplemente al factor *suerte*, se evaluó cada modelo sobre 337 conjuntos extramuestrales de 60 datos diarios cada uno¹⁴.

El cuadro 4 muestra que el PPS promedio¹⁵ del modelo ingenuo estándar¹⁶ (59,84%), para un período de 60 días, superó al PPS de la red ward estándar (56,89%) y de la red ward recursiva (57,27%). Luego, al analizar la estabilidad de la capacidad predictiva de los modelos, se encontró que el PPS del modelo ingenuo también registró la mayor variabilidad (un 8,09%, versus un 8,07% y un 7,05% correspondientes a la red ward estándar y recursiva, respectivamente). El modelo predictivo más estable, en términos relativos, fue la red ward con funcionamiento recursivo.

Al aplicar el test DA sobre los 337 períodos extramuestrales, encontramos que la significancia de la capacidad predictiva de los modelos no es clara, ya que ésta pasa de ser significativa en un período a no serlo en el siguiente, variando a través del tiempo. Esto plantea la necesidad de reconstruir los modelos período a período y de re-estimar sus pesos y coeficientes, y sugiere que no existe un único modelo explicativo de la evolución del tipo de cambio \$/US\$. Esta conclusión es similar a la obtenida por Parisi, Parisi y Guerrero (2003) en su estudio sobre modelos de redes neuronales aplicados a la predicción de índices bursátiles internacionales.

¹⁴Estos 337 períodos extramuestrales, de 60 datos cada uno, se obtuvieron a partir del conjunto extramuestral compuesto por 398 observaciones diarias. Debemos señalar que otra forma de probar la bondad de los modelos y la validez de sus resultados, independientemente de la muestra de datos sobre las que han sido aplicados, es aplicar un proceso de *bootstrap*. El *bootstrap* es un proceso de generación de observaciones ficticias a partir de datos históricos, a fin de dar solución al problema de escasez de datos y, de esta forma, poder contar con suficiente información para elaborar diferentes conjuntos extramuestrales sobre los cuales probar la validez de los modelos. No obstante, en el presente estudio se optó por dividir la serie histórica de observaciones en datos intramuestrales (usados para estimar los pesos y coeficientes de los modelos) y extramuestrales (usados para evaluar su capacidad predictiva), construyendo 337 subconjuntos extramuestrales a partir de este último.

¹⁵El PPS promedio, la desviación estándar del PPS y la rentabilidad acumulada promedio de los modelos y de la estrategia pasiva, fueron calculados a partir de los resultados que se obtuvieron de los 337 períodos extramuestrales, donde cada uno de ellos estuvo compuesto de 60 observaciones diarias.

¹⁶Sólo se presentan los resultados del modelo ingenuo o AR(1) con funcionamiento estándar, ya que el modelo ingenuo recursivo presentó resultados idénticos.

A pesar de que la capacidad predictiva de los modelos no fue significativa en más de la mitad de los períodos extramuestrales analizados, los modelos lograron predecir correctamente el signo de las variaciones del tipo de cambio con una frecuencia promedio superior al 50%. Se observó que el modelo ingenuo evidenció una capacidad predictiva significativa durante un mayor número de períodos extramuestrales (167, versus 114 de la red estándar y 108 de la red recursiva), registrando a su vez la capacidad predictiva promedio más alta. Pese a ello, el modelo AR(1) generó la menor rentabilidad promedio.

Al analizar el retorno promedio de los modelos durante los 337 períodos extramuestrales, encontramos que la red ward con funcionamiento estándar superó en rentabilidad a la red ward recursiva y al modelo ingenuo (un 2,49%, 2,45% y 2,43%, respectivamente, donde la rentabilidad corresponde a la acumulada durante un período de 60 días). Se observa además que, independientemente de la significancia estadística de la capacidad predictiva de los modelos, estos superaron en rentabilidad a la estrategia de inversión pasiva, la que generó un retorno acumulado promedio de 2,08%, a la vez que presentaron una menor volatilidad. El cuadro 5 muestra el retorno anormal promedio, calculado al comparar las rentabilidades de los diferentes modelos entre sí.

Cuadro 5

*Retorno anormal promedio entre los modelos analizados.
Entre paréntesis se entrega la desviación estándar de los retornos anormales, calculada considerando los 337 sub-períodos extramuestrales.*

MODELOS COMPARADOS	RETORNO ANORMAL PROMEDIO
Red Ward Estándar v/s Estrategia Pasiva	0,41% (1,65%)
Red Ward Recursiva v/s Estrategia Pasiva	0,37% (1,45%)
AR(1) Estándar v/s Estrategia Pasiva	0,35% (1,52%)
Red Ward Recursiva v/s Red Ward Estándar	-0,04% (0,55%)
Red Ward Estándar v/s AR(1) Estándar	0,06% (1,16%)
Red Ward Recursiva v/s AR(1) Estándar	0,02% (1,05%)

El cuadro 6 muestra el resultado de los *test* de Friedman y de Tukey. De acuerdo al *test* de Friedman, existen diferencias estadísticamente significativas entre el PPS de los modelos y también entre sus rentabilidades. Para individualizar estas diferencias se aplicó el *test* de Tuckey, el cual arrojó evidencia de que el modelo ingenuo y la red ward recursiva presentan diferencias significativas en cuanto a su capacidad predictiva, siendo mayor la del primer modelo. Un resultado similar se obtuvo al comparar el PPS del modelo ingenuo y de la red ward estándar, mientras que se concluyó que la red estándar y la red recursiva presentaron una capacidad de predicción equivalente en términos estadísticos.

Cuadro 6
Test de Friedman y de Tukey.

Test de Friedman: Diferencia entre medias	Estadístico S
PPS	27.83
Retorno Modelos	50.47
Retorno Anormal	8.06
Test de Tukey: PPS	P-Value
Red Ward Estándar – Red Ward Recursiva	0.8084
AR(1) - Red Ward Recursiva	0.0000
AR(1) – Red Ward Estándar	0.0000
Test de Tukey: Retorno Modelos	P-Value
Red Ward Estándar – Red Ward Recursiva	0.9987
AR(1) - Red Ward Recursiva	0.9997
Estrategia Pasiva – Red Ward Recursiva	0.4841
AR(1) – Red Ward Estándar	0.9950
Estrategia Pasiva – Red Ward Estándar	0.3919
Estrategia Pasiva – AR(1)	0.5406
Test de Tukey: Retorno Anormal	P-Value
Red Ward Estándar – Red Ward Recursiva	0.6594
AR(1) - Red Ward Recursiva	0.3326
AR(1) – Red Ward Estándar	0.8485

Para el caso del PPS y del retorno anormal, el valor crítico de χ^2 fue de 19.50, para 336 grados de libertad en el numerador, 2 grados de libertad en el denominador y un nivel de significancia de 0,05. Para el caso de la rentabilidad de los modelos, el valor crítico de χ^2 fue de 8,53, para 336 grados de libertad en el numerador, 3 grados de libertad en el denominador y un nivel de significancia de 0,05. El nivel de significancia, con el cual es comparado el *p-value*, es de 0,05.

No obstante, al comparar la rentabilidad de los modelos, el *test* de Tukey encontró que no existen diferencias estadísticamente significativas, incluso al contrastar la rentabilidad de los modelos predictivos con el retorno de la estrategia pasiva. Similares resultados se obtuvieron al analizar los retornos anormales. De esta manera, el análisis del tipo de cambio chileno no apoya las conclusiones de Tang et al. (1991), cuyo estudio señaló que, para series de tiempo de corto plazo, las redes neuronales muestran mejores resultados que las técnicas de Box-Jenkins.

4. Conclusiones

De las diversas arquitecturas de redes neuronales utilizadas en este estudio, la red ward de tres capas fue la que obtuvo el mejor rendimiento en la predicción del signo de las variaciones diarias del tipo de cambio observado. No obstante, la capacidad predictiva de la red ward estándar (evaluada sobre un conjunto extramuestral de 398 datos diarios) fue superada por la del modelo ingenuo. Además, el utilizar un proceso recursivo para re-estimar período a período los pesos de la red no generó mejoras significativas.

Al evaluar los resultados de los modelos sobre 337 períodos extramuestrales de 60 datos cada uno, se encontró que el PPS promedio del modelo ingenuo superó tanto a la red ward estándar como recursiva. El *test* de Tuckey arrojó evidencia de que el modelo ingenuo y la red ward recursiva presentan diferencias significativas en cuanto a su capacidad predictiva, siendo mayor la del primer modelo. Un resultado similar se obtuvo al comparar el PPS del modelo ingenuo y de la red ward estándar, mientras que se concluyó que la red estándar y la red recursiva presentan una capacidad de predicción equivalente en términos estadísticos. Luego, al comparar la rentabilidad de los modelos, el *test* de Tukey no encontró diferencias significativas, aun al comparar con la estrategia de inversión pasiva. Con relación al exceso de retorno de las técnicas respecto de la estrategia pasiva, si bien en promedio estos fueron positivos, no fueron significativos en términos estadísticos. En consecuencia, la capacidad predictiva de los modelos analizados, aun cuando resultó significativa, no se tradujo en mayores beneficios económicos. Estos resultados se adecuan a

lo esperado para el caso chileno, dado el bajo número de agentes que participan en este mercado, la alta volatilidad originada en sendas crisis financieras internacionales y los significativos cambios en el mercado a partir de la participación del Banco Central.

REFERENCIAS

- CALLEN, J., C. KWAN, P. YIP, y Y. YUAN, (1996). "Neural network forecasting of quarterly accounting earnings". *International Journal of Forecasting*, 12, p.p 475-482.
- CHEN, A. y M. LEUNG, (1998). "Dynamic foreign currency trading guided by adaptive forecasting". *Review of Pacific Basin Financial Markets and Policies*, 1, p.p. 383-418.
- DASGUPTA, C. G., G. S. DISPENSA, y S. GHOSE, S. (1994). "Comparing the predictive performance of a neural network model with some traditional market response models". *International Journal of Forecasting*, (10)2, pp. 235-244.
- ESTRELLA, A., y F. S. Mishkin, (1998), "Predicting US recessions: Financial variables as leading indicators", *The Review of Economics and Statistics*, 80(1), p.p. 45-61.
- HERBRICH, R., M. KEILBACH, T. GRAEPEL, P. BOLLMANN-SDORRA, y K. OBERMAYER, (2000), "Neural networks in economics: Background, applications and new developments", In T. Brenner, editor, *Advances in Computational Economics: Computational Techniques for Modelling Learning in Economics* 11, 169-196. Kluwer Academics.
- HILL, T., L. MARQUEZ, M. O'CONNOR, y W. REMUS, (1994), "Artificial neural network models for forecasting and decision making", *International Journal of Forecasting* (10)1, pp. 5-15.
- HORNIK, K., M. STINCHCOMBE y H. WHITE, (1989), "Multilayer feedforward networks are universal approximators", *Neural Networks* 5, 359-366.
- KANAS, ANGELOS (2001), "Neural networks linear forecasts for stocks returns", *International Journal of Finance and Economics*, 6, p.p. 245-254.
- KUO, C. y A. REITSCH, (winter 1995-96), "Neural networks vs. conventional methods of forecasting", *Journal of Business Forecasting*, 14, 17-22.

- LEUNG MARK T., DAOUK HAZEM y CHEN AN-SING (2000), "Forecasting stock indices: a comparison of classification and level estimation models", *International Journal of Forecasting*, (16)2, pp.173-190.
- MABERLY, E. D. (1986), "The informational content of the interday price change with respect to stock index futures", *Journal of Futures Markets*, 6, p.p. 385-395.
- MARTÍN DEL BRÍO, B. y A. SANZ, (1997), "Redes neuronales y sistemas borrosos: Introducción, teórica y práctica", Primera Edición. Ra-ma. 387 pp.
- O'CONNOR, M., W. REMUS, y K. GRIGGS, (1997), "Going up-going down: How good are people at forecasting trends and changes in trends?", *Journal of Forecasting*, 16, p.p. 165-176.
- PARISI, A. (2002), "Evaluación de modelos de redes neuronales de predicción del signo de las variaciones del IPSA", *Estudios de Administración*, Vol. 9, N°1.
- PARISI, A., F. PARISI, y J. L. GUERRERO, (2003), "Modelos predictivos de redes neuronales en índices bursátiles", *El Trimestre Económico*, México, (Forthcoming).
- PARISI, F., A. PARISI, y E. CAÑAS, (2002), "Reglas simples de análisis técnico y modelos autoregresivos en el mercado cambiario chileno entre 1995 al 2001" *Estudios de Administración*, Vol9, N° 2.
- PESARAN, M.H. y A. TIMMERMANN, (1992), "A simple nonparametric test of predictive performance", *Journal of Business and Economic Statistics*, 10, pp. 461-465.
- QI, M. (2001), "Predicting US recessions with leading indicators via neural network models", *International Journal of Forecasting*, 17, p.p. 383-401.
- TANG, Z., C. DE ALMIEDA y P. FISHWICK, (1991), "Time-series forecasting using neural networks vs. Box-Jenkins methodology", *Simulation* 57, 303-310.
- WHITE, H. (2000), "A reality check for data snooping", *Econométrica*, Vol. 68, No. 5, 1097-1126.
- WILSON, R. L. y R. SHARDA, (1994), "Bankruptcy prediction using neural networks", *Decision Support Systems*, 11(5), June, p.p. 545-557.