

Modelado Estructural No Lineal de Series Temporales

Paola Andrea Sánchez y Juan David Velásquez

UNIVERSIDAD NACIONAL DE COLOMBIA.

Escuela de Sistemas. Postgrado en Ingeniería de Sistemas
{pasanche; jdvelasq}@unalmed.edu.co

Recibido para revisión Octubre 2005, aceptado Noviembre 2005, versión final recibida Noviembre 2005

Resumen: Los modelos estructurales son herramientas conceptualmente útiles en el modelado de series temporales, toda vez que permiten la representación individual de cada componente estructural de la serie estudiada; sin embargo, la dificultad que estos presentan para representar relaciones no lineales ha desestimado su utilización en series reales. Alternativamente, los modelos basados en redes neuronales artificiales (RNA) son una alternativa promisoría para el modelado de series con características no lineales; no obstante, su incapacidad para dar una explicación económica de los parámetros calculados, hace que estas sean vistas como cajas negras. En éste estudio, se propone una metodología híbrida que combina las bondades del modelado estructural en la representación explícita de las características de la serie y de las redes neuronales en la captura de relaciones no lineales. Los resultados experimentales con series reales indican que la combinación de modelos de éste tipo puede ser una vía más efectiva para el modelado de series temporales no lineales que los modelos usados individualmente.

Palabras Clave: Modelos de Series Temporales; Modelos Estructurales Estáticos; Redes Neuronales Artificiales; Modelos Híbridos.

Abstract: Structural models are tools conceptually useful in time series modeling, because they allow the individual interpretation behavior of each structural component in the series; however, the difficult in the representation of nonlinear relationships of these models, have despised their use in real series. The Artificial Neural Networks (RNA) models are a promising alternative by the nonlinear time series modeling; nevertheless, the inability for to explain the parameters calculate of these models do that RNA have seen as black box. In this paper, we propose a hybrid approach that uses an explicit modeling of the structural characteristics of time series by structural models, joint with modeling of nonlinear relationships provided by neural networks. Experimental results with real series suggest the combination of this models can be an effective way to nonlinear time series modeling that the individual models.

Keywords: Time Series Models; Structural Static Models; Artificial Neural Networks; Hybrid Models.

1 INTRODUCCIÓN

El objetivo del modelado de series temporales es proveer al modelador de una representación matemática de la serie, que permita capturar, total o parcialmente, las características más relevantes del fenómeno real, a partir de la información contenida en los datos. Si bien, en la literatura se han propuesto diversos modelos orientados a la representación de series temporales, su utilidad depende del grado de similitud entre la dinámica del proceso generador de la serie y la formulación matemática del modelo con que ella se represente, por lo que muchos esfuerzos se han dedicado para el desarrollo de nuevos métodos y técnicas que superen las falencias encontradas.

En las aproximaciones tradicionales, la metodología de Box y Jenkins (1970) es aparentemente la más conocida, y más utilizada, debido a su sólida fundamentación matemática, y la existencia de un gran número de herramientas computacionales que facilitan su aplicación a casos reales. Sin embargo, sus mayores críticas se basan en el supuesto de una estructura de correlación lineal entre la variable dependiente y los regresores, los cuales están conformados por los valores pasados de la serie misma y los valores actuales y pasados de las variables explicativas que pueden dar cuenta del fenómeno; además, la aplicación de la metodología parte del supuesto de que el proceso generador de la serie sea estacionario en sentido homogéneo para el caso de los modelos univariados.

de tal forma, que componentes como la tendencia o los patrones cíclicos deben ser eliminados previo a la especificación del modelo, eliminado de la serie original aquella información potencialmente sobre la dinámica del proceso.

Una clase general de modelos lineales la constituyen los modelos estructurales, los cuales se basan en la representación de la serie mediante las componentes estructurales que la constituyen, como son el nivel y la tendencia, los patrones estacionales, los ciclos de largo plazo y la componente irregular, y donde otros modelos como los ARIMA, son sólo casos particulares [véase Harvey (1989) y Pole, West y Harrison (1994)]. Dichos modelos permiten la interpretación individual de la influencia de cada componente dentro de la serie, evitando su eliminación en el proceso de construcción y especificación del modelo, por lo que pueden considerarse conceptualmente más transparentes; sin embargo, su principal crítica es que asumen relaciones de orden lineal.

Por otra parte, ha sido ampliamente reconocido que muchas series temporales reales poseen comportamientos no lineales, y que en muchos de los casos, no existen suficientes leyes físicas o económicas que permitan especificar completamente un modelo estadístico para su representación [véase a Granger y Teräsvirta (1993) entre otros]. En consecuencia, diferentes modelos paramétricos no lineales han sido propuestos en la literatura [una amplia recopilación de estos es presentada por Tong (1990)], el cual parte de una forma supuesta de no linealidad, y realiza la especificación del modelo a partir de la información disponible; su éxito en casos prácticos depende del grado en que el modelo utilizado logre representar las características reales de la serie temporal estudiada. Su principal crítica está relacionada con que la formulación de cada familia de modelos requiere la especificación de un tipo determinado de no linealidad, por lo que no son universalmente aplicables ni válidos para todos los casos. Es así como las técnicas no paramétricas de regresión no lineal han surgido como alternativas atractivas para abordar dicho problema, puesto que permiten extraer las relaciones determinísticas no lineales desconocidas presentes entre los regresores y la serie, a partir de la información contenida en las observaciones, sin necesidad de realizar ningún tipo de supuestos [véase Zhang, Patuwo y M.Y. Hu (1998)]. De esta familia de técnicas, las RNA, y en especial los modelos tipo perceptrón multicapa (Multi-layer Perceptron - MLP), entendidos como modelos de regresión estadística no lineal, han recibido una gran atención entre los investigadores de la comunidad de Inteligencia Artificial y la de Estadística; es así, como los MLP han sido exitosos en la representación de numerosas aplicaciones en campos como la ingeniería, la medicina, las finanzas, la economía, entre otros [véase Masters (1995)].

En éste artículo, proponemos un acercamiento

híbrido al modelado de series temporales usando modelos estructurales estáticos y redes neuronales. La motivación del modelo híbrido viene desde las siguientes perspectivas: primero, en la práctica es a menudo dificultoso determinar si la serie en estudio es generada desde un proceso lineal o no lineal, o si un método particular es más efectivo que otro en el modelado de dicha serie, constituyéndose la elección de la técnica adecuada para la representación en una tarea compleja; segundo, las series temporales reales son en contadas ocasiones puramente lineales o no lineales [véase Zhang (2003)]: a menudo éstas series contienen una combinación de ambos patrones. Ante tal situación, la utilización de un modelo puramente lineal o no lineal da como resultado representaciones inadecuadas. Cada uno de los modelos citados, ha demostrado ser poderoso en su dominio de aplicación, pero no son universalmente aplicables en todas las circunstancias, es así como la aproximación de los modelos lineales a problemas no lineales puede no ser adecuada, mientras que el uso de modelos no lineales en series con características lineales, igualmente no presenta un buen desempeño. En consecuencia modelos híbridos, que combinan modelado lineal y no lineal, han surgido como alternativa para modelar estructuras de autocorrelación complejas y de orden no lineal en los datos con mayor precisión, puesto que permiten la captura de diferentes patrones presentes en estos: tercero, ha sido ampliamente aceptado en la literatura de modelado de series temporales que un método simple no es el mejor para todas las situaciones [véase Zhang (2003), Chatfield (1988), Makridakis, Anderson, Carbone, Fildes, Hibdon, Lewandowski, Newton, Parzen y Winkler (1982)]; esto es debido en gran parte a que los problemas reales son de naturaleza compleja y un modelo de este tipo puede no ser adecuado para capturar diferentes patrones. Estudios empíricos sugieren que combinando diferentes modelos, la precisión de la representación puede ser mejor que para el caso individual [véase Clemen (1989), Makridakis et al. (1982), Makridakis, Chatfield, Hibon, Lawrence, Millers, Ord y Simmons (1993), Newbold y Granger (1974)]. Por lo tanto, la unión de modelos con características diferentes incrementa la posibilidad de capturar diferentes patrones en los datos y proporciona una representación más acorde de la serie temporal. El modelado híbrido surge entonces, de forma natural como la unión de técnicas similares o diferentes con características complementarias.

En la literatura de series temporales varias combinaciones de métodos han sido propuestas. Sin embargo, muchas de ellas combinan modelos similares; es así, como diferentes estudios acerca de combinaciones de técnicas de modelado lineal son encontrados en la literatura tradicional. Si bien, éste tipo de combinaciones ha demostrado su capacidad de mejorar la precisión de las representaciones hechas, se considera que una vía más

efectiva de combinación podría estar basada en modelos con características diferentes.

Tanto la evidencia teórica como la empírica sugieren que la combinación de modelos disímiles, o que discrepen fuertemente de otros, conduce a una disminución en los errores del modelo [véase Granger (1989), Krogh y Vedelsby (1995), Perrone y Cooper (1993), Makridakis (1989), Palm y Zellner (1992) y Winkler (1989)] y permite, adicionalmente, reducir su incertidumbre [Chatfield (1996)]. El modelo combinado es así más robusto para estimar los posibles cambios en la estructura de los datos. De dichos modelos, las combinaciones de modelos lineales con técnicas basadas en Inteligencia Computacional han tenido gran aceptación [véase trabajos de Luxhoj, Riis y Stensballe (1996), Ginzburg y Ilorn (1994), Pelikan, De Groot y Wurtz (1992), Wedding II y Cios (1996) y Zhang (2003)]; no obstante, las principal crítica a estos trabajos es que exigen un preprocesamiento de la serie, el cual va encaminado a eliminar los componentes estructurales de ésta.

Este artículo se concentra en la problemática del modelado de series temporales no lineales, siendo su principal aporte, la formulación de un modelo híbrido estructural no lineal que permite la representación de las componentes estructurales de la serie utilizando un modelo estructural estático, así como las relaciones no lineales mediante RNA del tipo Perceptrón Multicapa. Para alcanzar el objetivo propuesto, se ha organizado el resto de este artículo en la siguiente forma: en la Sección 2, se discuten los acercamientos hechos al modelado de series temporales mediante modelos estructurales y modelos de RNA; la metodología híbrida propuesta para la representación de series temporales no lineales es presentada en la Sección 3, mientras que en la Sección 4 se aplica a dos casos reales; y finalmente en la Sección 5, se discuten las principales lecciones aprendidas durante la realización de este trabajo.

2 MODELADO DE SERIES TEMPORALES

Existen diferentes acercamientos al modelado de series temporales: los modelos estadísticos tradicionales que incluyen medias móviles, modelos autorregresivos, suavizamiento exponencial, modelos estructurales y modelos ARIMA, se basan en la hipótesis de relaciones lineales, donde se asume que el valor actual de la serie se puede representar mediante una combinación lineal de sus valores pasados y los de sus variables explicativas. Por su relativa simplicidad en el entendimiento e implementación, los modelos lineales han sido el foco de investigación y aplicación durante varias décadas. Modelos como medias móviles, modelos autorregresivos, suavizamiento exponencial y ARIMA exigen un preprocesamiento de la serie orientado a la eliminación de las componentes estructurales de esta, no siendo así

para el caso de modelos estructurales, los cuales permiten una representación de la serie sin tratamientos a priori. No obstante, la imposibilidad de estos modelos para representar relaciones no lineales, ha llevado a los investigadores a volcar su interés hacia áreas más prometedoras. Es así como en la literatura se han propuesto modelos que intentan representar de forma paramétrica las relaciones no lineales presentes en la serie temporal, tales como los modelos bilineales [Granger y Andersen (1978)], los modelos de regímenes autorregresivos (*threshold autoregressive model - TAR*) [Tong (1990)], y los modelos autorregresivos de heterocedasticidad condicional (*autoregressive conditional heteroscedastic model - ARCH*) [Engle (1982)]. Aunque se han conocido aplicaciones exitosas de estos modelos no lineales, la ganancia en su aplicación está limitada al hecho que el modelo logre representar las características de la serie [De Gooijer y Kumar (1992)], dado que dichas metodologías se han desarrollado asumiendo patrones específicos de no linealidad, por lo que no son universalmente aplicables. Estudios recientes sugieren los modelos basados en RNA como una alternativa efectiva en el modelado de series temporales con características no lineales; la gran fortaleza que presentan estos modelos es precisamente su flexibilidad para modelar relaciones no lineales. En esta Sección, se discutirá acerca de la problemática del modelado estructural y de redes neuronales.

2.1 El Modelo Estructural

En el modelado estructural de series temporales [véase Pole et al. (1994) y Harvey (1989)] se asume que la serie puede ser descompuesta en varios componentes aditivos que reflejan sus diferentes propiedades. Dichos componentes son: la componente tendencia μ_t , los patrones estacionales $\gamma_{p/t}$, los ciclos de largo plazo ψ_t , los i rezagos de la serie temporal y_{t-i} , y la componente irregular ϵ_t , de tal forma que la serie puede ser descrita vía representación estructural como:

$$y_t = \mu_t + \gamma_{p/t} + \psi_t + \sum_{i=1}^n \delta_i y_{t-i} + \epsilon_t \quad (1)$$

Se asume que la componente irregular ϵ_t , se comporta como un proceso de ruido blanco; esto es como una secuencia de variables aleatorias no correlacionadas serialmente, independientes e idénticamente distribuidas con media constante e igual a cero y varianza constante σ_ϵ^2 .

La peculiaridad de los modelos estructurales radica en su flexibilidad para representar adecuadamente movimientos de la serie temporal, los cuales provienen de diferentes propiedades, obteniéndose así una mejor descripción de la serie y sus componentes. La importancia de una representación de la serie mediante estos modelos reside precisamente, en que ellos proveen una

descripción de la serie en términos de sus variables de interés, sin requerir un procedimiento previo de eliminación de componentes, como lo exigen otros modelos.

La principal ventaja del modelado estructural es que permite modelar el comportamiento de las componentes de una serie separadamente y luego expresar juntos los submodelos para formar un sólo modelo que abarca a toda la serie; sin embargo, su utilidad se ve opacada por la restricción en el modelado sólo de relaciones lineales.

En el modelado estructural de series temporales, diferentes estrategias han sido propuestas para la especificación del modelo de la serie [véase Harvey (1989)]. Estos procedimientos pueden ser ajustados, de tal forma que permitan representar relaciones de orden no lineal. La estrategia propuesta por (Harvey, 1989) presenta tres pasos iterativos: la identificación del modelo, la estimación de los parámetros y la validación o diagnóstico.

La identificación del modelo consiste en un análisis exploratorio de la serie temporal y sus características, con el fin de determinar las componentes que deben ser incluidas en el modelo y presentar así, una especificación de éste para su evaluación posterior. En éste punto no es necesario desestacionarizar la serie [transformar y diferenciar para obtener una media y varianza constantes], como lo requieren los modelos ARIMA, puesto que la estimación posterior contempla la representación de cada componente.

Una vez el modelo tentativo es identificado, se calculan los parámetros del modelo, mediante algún método de minimización del error. Diferentes algoritmos lineales y no lineales han sido propuestos para dicha minimización; sin embargo, aquellos basados en la minimización de los residuales cuadráticos del modelo ofrecen ventajas representativas.

Finalmente y con el fin de validar la adecuación del modelo encontrado a la serie, se examinan sus residuales verificándose que sean incorrelacionados y normalmente distribuidos con media cero y varianza constante. Las desviaciones de los supuestos hechos sobre los residuales son una fuerte evidencia ya sea por una falla en el proceso de especificación, la ausencia de componentes significativos o la inclusión de componentes no relevantes. El incumplimiento de los supuestos sobre los residuales conduce a una nueva etapa de identificación.

2.2 Modelo de Series Temporales Usando RNA

Cuando se conoce la existencia de no linealidades en la serie, un gran número de estructuras no lineales pueden ser usadas para describir la evolución y el comportamiento de ésta. Un buen modelo no lineal debe ser lo "suficientemente general para capturar las relaciones no lineales contenidas en los datos" [De Gooijer y Kumar (1992)]. Las RNA son modelos con una alta capacidad para aproximar series temporales no lineales sólo a par-

tir de la información contenida en las observaciones, y sin asumir patrones de comportamientos a priori [véase Kaastra y Boyd (1996) y Masters (1995)].

Una ventaja significativa que poseen los modelos de RNA sobre otros modelos no lineales es su capacidad para aproximar con un alto grado de precisión cualquier relación no lineal. Su poder viene del procesamiento paralelo que hace de la información a partir de los datos.

Los MLP son un tipo de red neuronal donde las neuronas se encuentran agrupadas en capas, y cada una de ellas está conectada con las neuronas de la capa siguiente. En el contexto de series temporales, el modelo MLP consta de una capa de entrada en donde existe una neurona por cada regresor de la serie temporal, de una o más capas ocultas donde se realiza el procesamiento de la información, y finalmente una capa de salida donde la red entrega los valores pronosticados de la variable dependiente. La salida de la capa oculta se produce al aplicar a la entrada una función de transformación, conocida como función de activación.

Un MLP con una capa oculta y transferencia lineal en la activación de su capa de salida, puede ser definido matemáticamente como un modelo y_t , que representa las relaciones existentes entre el grupo de regresores de la serie temporal y_{t-i} , visibles en el instante de tiempo t , tal que:

$$y_t = \alpha_0 + \sum_{j=1}^q \alpha_j g[\beta_{0j} + (\sum_{i=1}^p \beta_{ij} y_{t-i})] + \epsilon_t \quad (2)$$

donde α_j y β_{ij} son los parámetros del modelo a menudo llamados pesos; p es el número de neuronas en la capa de entrada y q el número de neuronas en la capa oculta; ϵ_t son los errores aleatorios del modelo, los cuales se asume que son independientes e idénticamente distribuidos (iid) con media cero y varianza constante σ_ϵ^2 ; y $g(u)$, representa la función de activación de la capa oculta, la cual a menudo equivale a la función logística:

$$s(u) = \frac{1}{1 + \exp(-u)} \quad (3)$$

Según el modelo (2), la entrada a cada neurona de la capa oculta es la parte determinística de un modelo autorregresivo, y en consecuencia, puede considerarse que la RNA es la suma ponderada de tantos modelos autorregresivos como neuronas posea en la capa oculta. y_t resulta definida en función de sus valores rezagados y de los pesos de la red, en una forma funcional de orden no lineal, por lo que éste modelo en el contexto de series temporales es equivalente a una combinación no lineal de modelos autorregresivos.

Con el diseño de una red neuronal artificial se pretende conseguir que para ciertas entradas de la red, ésta sea capaz de generar una salida deseada. Para ello, además de una topología de red adecuada (arquitectura),

se requiere un proceso de aprendizaje o entrenamiento, que permita modificar los pesos de las neuronas de la capa oculta y de la salida hasta encontrar una configuración acorde con la relación medida por algún criterio y de esta forma estimar los parámetros de la red.

Si bien, las aplicaciones de redes neuronales presentes en la literatura han demostrado los beneficios de éste enfoque, han sido también criticadas por la dificultad que presentan para replicar una solución obtenida previamente, el peligro de sobreentrenamiento y el alto número de parámetros que deben ser experimentalmente seleccionados para generar la salida deseada, como son: la selección de variables de entrada a la red neuronal a partir de un conjunto usualmente grande de posibles entradas; la selección de la arquitectura interna de la red; y la estimación de los valores asociados a los pesos de las conexiones [véase Kaastra y Boyd (1996)].

Para cada uno de los problemas mencionados anteriormente se han propuesto en la literatura diferentes acercamientos [véase DasGupta, Siegelmann y Sontag (1995), Peretto (1992), Mehra y Benjamin (1997)].

La selección de las variables de entrada depende en gran medida del conocimiento que posea el modelador acerca de la serie temporal, y es tarea de éste elegir según algún criterio previamente fijado. La necesidad de cada variable dentro del modelo. Como medida alternativa para apoyar la decisión del modelador, estudios recientes han planteado la utilización de procedimientos basados en la utilización de análisis decisional [véase Velásquez (2003)], o en los métodos estadísticos tradicionales, tales como las funciones de autocorrelación [véase Box y Jenkins (1970)].

En cuanto al problema de la selección de la arquitectura interna de la red, es decir, el número de capas ocultas y de neuronas en cada capa, teóricamente una red neuronal con una capa oculta y un número suficiente de neuronas puede aproximar con cierta precisión cualquier función continua definida en un dominio compacto [véase Kaastra y Boyd (1996), Hornik, Stinchcombe y White (1990)]. Sin embargo, la determinación en la práctica del número adecuado de capas, así como del número de neuronas en cada capa es una tarea compleja. Algunos autores sugieren el uso de una capa oculta cuando la serie temporal es continua, y de dos sólo si existe algún tipo de discontinuidad. El incremento del número de capas ocultas sólo aumenta el tiempo computacional y el peligro de sobreentrenamiento. Ahora bien, con un número pequeño de neuronas en la capa oculta la red no puede aprender adecuadamente las relaciones existentes en los datos, mientras que un número superior al requerido hace que la red memorice los datos con una generalización pobre, por lo que es de poca utilidad para la predicción. Es así como para la selección del número óptimo de neuronas ocultas diversos métodos han sido propuestos; sin embargo, los métodos basados en algo-

ritmos constructivos tienen ciertas ventajas sobre otros, puesto que permiten evaluar la conveniencia de adicionar o no una nueva neurona a la red, durante la etapa de entrenamiento, según ella disminuya el término del error, lo que los convierte en métodos computacionalmente más eficientes.

La estimación de los parámetros de la red neuronal es a menudo hecha según el criterio de minimización de la suma de errores cuadráticos (MSE); sin embargo, este resultado requiere de un proceso iterativo de entrenamiento. El objetivo del entrenamiento es encontrar un conjunto óptimo de pesos de las neuronas que determine el mínimo global de la función del error. Algoritmos potencialmente útiles para la estimación de los parámetros basados en la utilización de gradiente decente (GRG2) han sido desarrollados [véase Hung y Denton (1993), Subramanian y Hung (1993)], los cuales ajustan los pesos reduciendo la pendiente de la superficie del error; sin embargo, es posible formular una estrategia combinada para el entrenamiento de la red neuronal donde los pesos son estimados conjuntamente con el número de neuronas ocultas.

Otra limitación asociada a la aplicación de RNA en el modelado de series temporales es que los parámetros arrojados por la red son casi imposibles de interpretar y no brindan información sobre qué tipo de modelo paramétrico no lineal sería adecuado para describir los patrones o dinámicas no lineales identificados, razón por la cual suele hacerse referencia a estas como "cajas negras" en el sentido de que sirven sólo para identificar patrones en la serie, pero no permiten interpretar ni dar sentido económico a los parámetros estimados.

3 LA METODOLOGÍA HÍBRIDA

La representación estructural de una serie temporal, dada por la ecuación (1), permite explicar de manera independiente el comportamiento de cada componente estructural de la serie. La finalidad del modelado estructural es encontrar una función para cada componente estructural de la serie, que aproxime con algún grado de precisión dichas componentes a su forma real, valiéndose de la información proporcionada por la serie de tiempo, tal que ella es representada por:

$$y_t = \hat{\mu}_t + \hat{\gamma}_p/t + \hat{\psi}_t + \sum \hat{\delta}_i y_{t-i} + \hat{\epsilon}_t \quad (4)$$

donde, $\hat{\mu}_t$, $\hat{\gamma}_p/t$, $\hat{\psi}_t$, $\hat{\delta}_i$ y $\hat{\epsilon}_t$, representan las funciones estimadas para cada componente de la serie.

Si bien, el modelo mostrado en 4, en la práctica ha demostrado ser adecuado en la representación de múltiples casos [véase ejemplos de aplicaciones en Harvey (1989)], la crítica más acentuada que se hace, es que dicho modelo asume relaciones de orden lineal entre sus

componentes, representando de forma inadecuada una serie con características no lineales.

En el caso ideal, posterior a la estimación de las contribuciones de cada componente estructural, los residuales restantes, atribuibles al componente irregular, serían ruido blanco. No obstante, si la serie presenta relaciones de orden no lineal, y bajo el supuesto de que las componentes estructurales observables de la serie son identificables y calculables, dichas relaciones quedarían alojadas dentro del término del error, afectando así, de forma sustancial la representación que se haga de éste y por lo tanto de la serie. Lo anterior, justifica la necesidad de modelar independientemente las características no lineales de dicho componente. La componente irregular quedaría definida como:

$$\epsilon_t = y_t - (\hat{\mu}_t + \hat{\gamma}_{\eta/t} + \hat{\psi}_t + \sum \hat{\delta}_i y_{t-i}) \quad (5)$$

Ahora bien, teniendo en cuenta el poder que tienen las RNA en la aproximación de funciones no lineales, es posible representar las relaciones ocultas dentro del componente de error, mediante la aplicación de una RNA, con n nodos de entrada, de la forma:

$$\epsilon_t = f(\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-n}) + \eta \quad (6)$$

donde f es una función no lineal determinada por la red neuronal y η es el error. Denominando NN al término del error estimado en (6) mediante redes neuronales, el modelo general de la serie temporal quedaría definido de la forma:

$$y_t = \hat{\mu}_t + \hat{\gamma}_{\eta/t} + \hat{\psi}_t + \sum \hat{\delta}_i y_{t-i} + NN \quad (7)$$

La representación híbrida hecha de una serie temporal no lineal planteada en (7) permite solventar problemas independientes de cada técnica [por ejemplo: para el caso de los modelos estructurales, que estos sólo permiten la representación de relaciones lineales, o en la redes neuronales, que no admiten una interpretación económica de los parámetros resultantes], e incrementa la probabilidad de capturar diferentes patrones presentes en los datos, combinando así, convenientemente, los beneficios que poseen éstas [modelos estructurales y RNA] en la representación de series temporales.

La metodología para la construcción del modelo híbrido consiste de dos pasos; en el primero, un modelo estructural es utilizado para representar las componentes lineales del problema; en el segundo, la red neuronal es desarrollada para modelar los residuales del modelo estructural. Como el modelo estructural no está en capacidad de capturar la estructura no lineal de los datos, una vez se determinen las componentes estructurales de la serie, los residuales del modelo contendrán la información acerca de la no linealidad remanente. El proceso de estimación de los parámetros del modelo estructural

y de la red se hace de manera conjunta mediante un algoritmo iterativo basado en la reducción del término del error.

4 RESULTADOS EMPÍRICOS

La efectividad y el desempeño del modelo propuesto fueron evaluados mediante su aplicación a dos casos reales: la serie de Muertes por Accidentes Automovilísticos en Gran Bretaña y la de Anomalías en la Temperatura de la Superficie del Mar en la Zona Niño 3.5. Estas series provienen de áreas diferentes y poseen diferentes características estadísticas.

La serie de Accidentes Automovilísticos en GB contiene el número de Muertes por Accidentes Automovilísticos en Gran Bretaña entre enero de 1969 y diciembre de 1984, y posee un comportamiento no lineal, tal como puede observarse en la Figura 1; ésta ha sido estudiada por varios autores desde diferentes ópticas [véase Harvey (1989) para un modelado estructural y con efecto de intervenciones, Durbin y Harvey (1985)].

La serie Niño 3.5 contiene el número de Anomalías en la Temperatura de la Superficie del Mar en la Zona Niño 3.5. La serie, mostrada en la Figura 2, consiste de 645 observaciones, y al igual que la serie anterior posee un comportamiento no lineal.

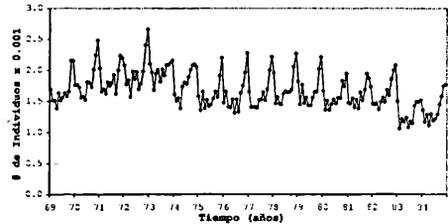


Figura 1: Muertes por Accidentes Automovilísticos en Gran Bretaña ene/69-dic/84 [véase Harvey (1989)]

• Valores escalados por 10^{-3}

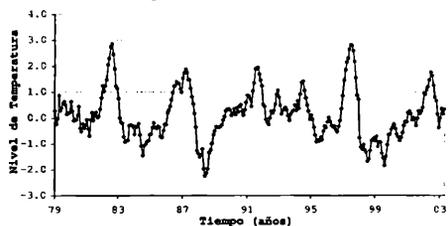


Figura 2: Anomalías en la Temperatura de la Superficie del Mar en la Zona Niño 3.5 jun/79-sep/03

4.1 Caso 1: Serie de Muertes por Accidentes Automovilísticos en Gran Bretaña

En primer lugar se evalúa la serie mensual de Muertes por accidentes automovilísticos en Gran Bretaña para el periodo enero de 1969 y diciembre de 1984, para un total de 192 observaciones. Como se mencionó anteriormente, dicha serie ha sido estudiada desde diferentes ópticas obteniéndose resultados similares [véase Harvey (1989), Durbin y Harvey (1985)].

A) *Identificación de la Serie:* la evaluación visual de la serie en la Figura 1, sugiere la presencia de un ciclo estacional constante de periodo anual, cuyos mayores valores se presentan en el mes de diciembre [picos] y los menores entre abril y junio [valles]. Presenta también cambios bruscos en su media y varianza, y se descartan la presencia de componentes asociados a tendencia y ciclos de largo plazo.

El test de linealidad *RESET* aplicado a la serie indica con un nivel de confianza del 95%, que dicha serie es no lineal, toda vez que $R = 8.1953 > VC = 1.9942$.

Según la evaluación preliminar, se puede afirmar que la serie de accidentes automovilísticos en GB presenta un comportamiento no lineal, lo cual justifica la utilización del modelo propuesto y además un componente estacional de periodo 12, una media que se asume, para efectos de la representación planteada, constante y diferente de cero; así como, la presencia de volatilidades que pueden ser recogidas mediante la inclusión en el modelo de una variable de rezagos. Se descarta la presencia de tendencia y ciclos de largo plazo.

B) *Resultados de la Estimación y Diagnóstico del Modelo:* Se procede a desarrollar un modelo híbrido estructural no lineal (7) para la serie estudiada. El modelo considerado contiene una componente estacional de periodo 12, un parámetro para el nivel y las variables con rezago uno y dos de la serie, con la finalidad de recoger volatilidades y comportamientos pasados. Mientras que la red neuronal, esta determinada por una variable de entrada equivalente a los valores con rezago uno de los residuales, una neurona oculta y una variable de salida.

Para evaluar el comportamiento de los residuales del modelo ajustado se construyó una ilustración resumida de algunos gráficos estadísticos calculados para estos, los cuales son presentados en la Figura 3. En la parte superior se presentan los gráficos de los residuales [izquierda] y los residuales cuadráticos [derecha]; en el primero los datos se distribuyen uniformemente alrededor de cero y en el segundo los datos presentan valores constantes; también se

revela en ambos gráficos la presencia de volatilidades aunque con amplitudes bajas, en especial al final del periodo [año 1983] donde los residuales aparecen persistentemente por debajo de cero; de la evaluación de estos gráficos puede decirse que los residuales se distribuyen uniformemente alrededor de una media constante e igual a cero; el autocorrelograma de los errores es mostrado en el gráfico inferior izquierdo, en él, a pesar de que existen datos que se salen de las franjas, son pocos y de baja magnitud y no parecen seguir algún tipo de patrón, por lo que puede considerarse los residuales del modelo como incorrelacionados; finalmente, el gráfico inferior derecho que corresponde al histograma de frecuencias exhibe una forma [línea fina] generalizada de una distribución normal estándar [línea punteada], por lo que no se puede concluir que los residuales se distribuyan normalmente. La desviación del supuesto de normalidad que exhiben los residuales de la serie, es reflejo de los cambios presentes en su media, por lo que se obviará una nueva identificación del modelo y se asumirá que éste representa adecuadamente las características de la serie.

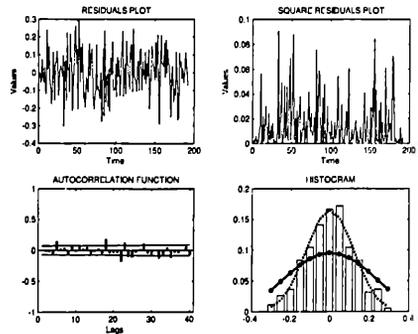


Figura 3: Gráficos de Residuales del Modelo Estructural No Lineal calculado para la Serie de Muertes por Accidentes Automovilísticos en Gran Bretaña

El desempeño del modelo híbrido calculado fue comparado con una RNA y un modelo Estructural mediante el cálculo de su MSE. En la Tabla 1, aparecen los valores del MSE calculados para cada uno de los modelos.

Los resultados de la Tabla 1 indican, que el modelo estructural no lineal se ajusta mejor que los demás presentados a las características de la serie, toda vez que su valor calculado para el MSE es menor que en los otros casos.

Tabla 1: Valores del MSE de los modelos propuestos para la serie de Muertes por Accidentes Automovilísticos en Gran Bretaña

	MSE
Modelo RNA	0.0259
Modelo Estructural	0.0176
Modelo Estructural No Lineal	0.0144

El modelo estructural considera la influencia de la componente estacional, la tendencia, y los rezagos 1 y 12; mientras que para la red neuronal se asumen como significativos los rezagos 1, 2, 12 y 13 de la serie, y contiene además dos neuronas en su capa oculta. Ambos modelos permiten la representación que características disímiles de la serie, tales como la presencia de componentes estructurales y su condición de no linealidad; sin embargo, la diferencia entre sus valores sugiere que ninguno de los dos modelos es apto para representar todos los patrones en los datos.

Los resultados del modelo híbrido indican que la combinación de ambos modelos puede reducir significativamente los errores. En términos de MSE, los porcentajes de mejora del modelo híbrido propuesto son de 44.6% y 18.3% con respecto a la red neuronal y al modelo estructural, respectivamente.

En la Figura 4, se presentan los gráficos de los modelos calculados [línea punteada], en contraste con la serie real [línea continua]. La evaluación visual de estos resultados muestra que para el modelo de RNA [véase Figura 4(a)], los valores aunque intentan seguir el comportamiento de la serie, exhibe altos niveles de volatilidad y se queda corto al reflejar los movimientos estacionales presentes en la serie; en contraste, los dos modelos restantes [véase Figuras 4(b) y 4(c)] representan claramente dentro de ellos la influencia de la componente estacional, pero el modelo estructural no lineal refleja detalles singulares de la dinámica de la serie que el modelo estructural tiende a generalizar.

El gráfico de dispersión entre la serie de Muertes por accidentes automovilísticos en Gran Bretaña y el modelo estructural no lineal calculado es mostrado en la Figura 5. Como puede observarse, los valores se agrupan alrededor de una línea de tendencia de 45°, lo cual indica que el ajuste del modelo a la serie es apropiado, lo que permite afirmar que el modelo puede considerarse un buen estimador de la serie.

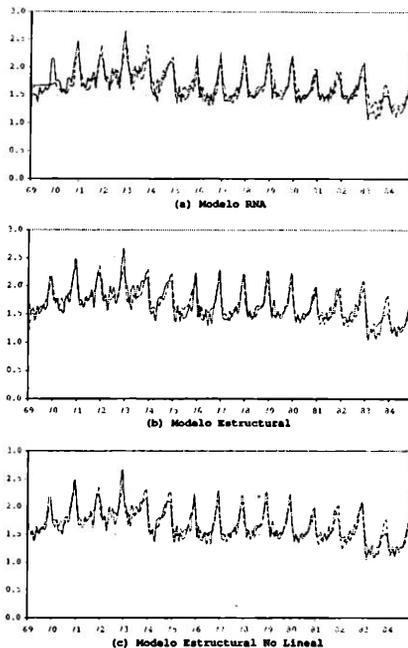


Figura 4: Modelos calculados para la Serie de Muertes por Accidentes Automovilísticos en Gran Bretaña. (a) Modelo RNA, (b) Modelo Estructural y (c) Modelo Estructural No Lineal

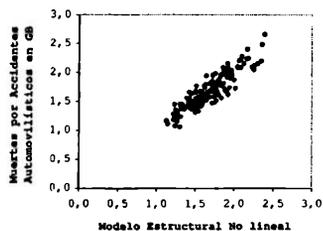


Figura 5: Gráfico de Dispersión entre la Serie de Muertes por Accidentes Automovilísticos en Gran Bretaña y el Modelo Estructural No Lineal Ajustado

De éste caso de aplicación se puede concluir que aunque si bien, el modelo estructural no lineal ajustado a la serie de Muertes por Accidentes Automovilísticos en Gran Bretaña manifiesta algunas características de la serie que son difíciles de ex-

plicar, y que son ignoradas por otros modelos, como es la no linealidad, el incumplimiento del supuesto de normalidad conduce a pensar que existen aún comportamientos dentro de la serie que el modelo no logra reflejar y que posiblemente requieran un tratamiento adicional, como puede ser el modelado explícito dentro de éste de los cambios en su media.

4.2 Caso 2: Serie de Anomalías en la Temperatura de la Superficie del Mar en la Zona Niño 3.5

De igual forma que en el apartado anterior se estudió la serie de Anomalías en la Temperatura de la Superficie del Mar en la Zona Niño 3.5. La serie mensual para el periodo comprendido entre junio de 1979 y septiembre de 2003, está compuesta de 292 observaciones, y fue obtenida del servidor NOAA. La serie presenta ciclos de largo plazo debido a la influencia del fenómeno del Niño, los cuales se ven reflejados por aumentos en las temperaturas para los años donde se manifiesta el evento.

A) *Identificación de la Serie:* La inspección visual de la serie en la Figura 2, revela la presencia de una media aproximadamente constante, ciclos de largo plazo y volatilidades. No manifiesta la presencia de ciclos estacionales de corto plazo [periodos menores a un año], ni tendencia. Los ciclos presentes en la serie son de amplitud y periodo variable, de aproximadamente cada 4 años, no obstante para efectos de este trabajo se considerarán dichos ciclos como determinísticos, y por tanto, calculables.

La prueba *RESET* aplicada a la serie, produjo un $R = 2.7667 > VC = 2.1129$, lo cual indica con una significancia de 95%, que es correcto afirmar que la serie estudiada es no lineal.

De acuerdo al análisis preliminar de la serie Niño 3.5, podemos afirmar que ésta exhibe ciclos de largo plazo de amplitud y periodo variables, una media constante e igual a cero, volatilidades que pueden ser recogidas mediante la inclusión en el modelo de una variable de rezagos y se descarta la presencia de tendencia y componentes estacionales. La serie presenta además un comportamiento no lineal, lo que justifica la utilización del modelo propuesto.

B) *Resultados de la Estimación y Diagnóstico del Modelo:* El modelo considerado para la representación de la serie, contiene una componente cíclica de periodo 4, lo que equivalente a una frecuencia de 6 ciclos presentes en el periodo de medición, y las variables con rezago uno y tres de la serie. La red neuronal, contiene una variable de entrada, equivalente a los valores con rezago uno de los residuales, dos neuronas en la capa oculta y una neurona en la capa de salida.

El comportamiento de los residuales del modelo ajustado es evaluado mediante el análisis de los gráficos presentados en la Figura 6. Los gráficos de los residuales [superior izquierda] y los residuales cuadráticos [superior derecha] indican, que los errores del modelo se distribuyen uniformemente alrededor de cero, en el primer gráfico, y conserva valores constantes y de poca magnitud, en el segundo: El gráfico inferior izquierdo presenta el autocorrelograma de los errores, en él a pesar de que existen datos que se salen de las franjas, son pocos y de baja magnitud, por lo que puede considerarse incorelacionados; por último, en el gráfico inferior derecho se presenta el histograma de frecuencias de los residuales, en el cual, la forma de campana adoptada por la frecuencia de los residuales [línea gruesa], similar a la distribución normal ideal [línea punteada], indica que estos pueden considerarse normalmente distribuidos.

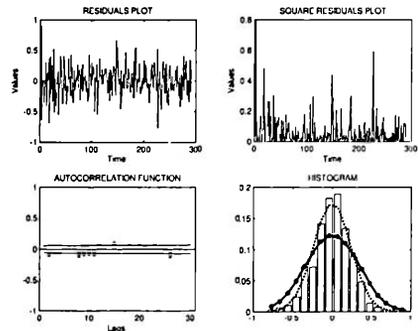


Figura 6: Gráficos de Residuales del Modelo Estructural No Lineal calculado para la Serie de Anomalías en la Temperatura de la Superficie del Mar en la Estación Niño 3.5

El modelo híbrido calculado fue comparado con una RNA y un modelo Estructural mediante el cálculo del MSE, con el fin de evaluar su desempeño. Los valores del MSE calculados para cada modelo aparecen en la Tabla 2.

Tabla 2: Valores de MSE de los Modelos Propuestos para la Serie de Anomalías en la Temperatura de la Superficie del Mar en la Estación Niño 3.5

	MSE
Modelo RNA	0.0716
Modelo Estructural	0.0631
Modelo Estructural No Lineal	0.0562

Los resultados de la Tabla 2 sugieren que el modelo estructural no lineal representa mejor las características de la serie considerada que los demás modelos, ya que presenta valores menores para el término del error. El modelo Estructural está influenciado por la presencia de ciclos de largo plazo y por los aportes hechos por los rezagos 1, 3 y 5. Igualmente, la RNA considerada contiene tres variables de entrada correspondientes a los rezagos 1, 3 y 5 de la serie, dos neuronas en la capa oculta y una neurona en la capa de salida. Los resultados sugieren una mejoría del 21.4% y 10.9%, del modelo estructural no lineal propuesto con respecto a los modelos RNA y Estructural, esto debido a que el modelo propuesto además de reconocer las características no lineales de la serie, incluye dentro de sí el aporte de las componentes estructurales.

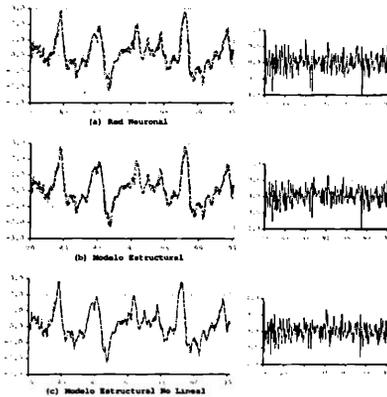


Figura 7: Modelos Calculados para la Serie de Anomalías en la Temperatura de la Superficie del Mar en la Zona Niño 3.5. (a) Modelo RNA, (b) Modelo Estructural y (c) Modelo Estructural No Lineal

La Figura 7 presenta los gráficos de cada modelo calculado [línea punteada], comparados con la serie real [línea continua]. Como se observa en los gráficos, el modelo RNA refleja de forma adecuada el comportamiento general de la serie, sin embargo, presenta altos grados de volatilidad [véase la Figura 7(a)]; para el caso estructural [véase la Figura 7(b)], el modelo tiene en cuenta tanto valores rezagados de la serie como los ciclos de largo plazo presentes en ella, por lo que dicho comportamiento es bien recogido por el modelo, no obstante en puntos extremos, el modelo se queda corto al intentar replicar el comportamiento de la serie; en contraste, el modelo es-

tructural no lineal captura de forma más apropiada dichos comportamientos de la dinámica de la serie [véase la Figura 7(c)], lo cual se refleja de forma más clara en los gráficos de los errores adjuntos, en los cuales se observa que para el mismo nivel, el modelo estructural no lineal presenta menos volatilidad.

En la Figura 8 se presenta el gráfico de dispersión entre la serie de anomalías en la temperatura de la superficie del mar en la zona Niño 3.5 y el modelo estructural no lineal calculado. En éste se observa como los datos se concentran siguiendo una línea de tendencia de 45° , lo que es reflejo de un buen ajuste del modelo a la serie y permite afirmar que el modelo calculado puede considerarse un buen estimador de ésta.

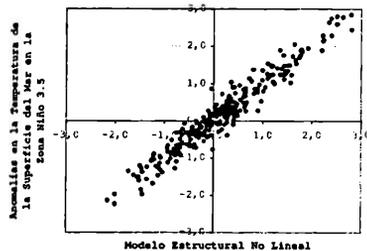


Figura 8: Gráfico de Dispersión entre la Serie de Anomalías en la Temperatura de la Superficie del Mar en la Zona Niño 3.5 y el Modelo Estructural No Lineal Ajustado

Del modelado de la serie Niño 3.5, se puede concluir que la representación hecha de las características de la serie mediante el modelo estructural no lineal logra capturar mejor comportamientos de la serie que otros modelos, por lo que según los resultados presentados, dicho modelo, puede considerarse un buen estimador de la serie.

5 CONCLUSIONES

En éste trabajo se demostró como la representación explícita de los componentes estructurales de una serie temporal hecha por los modelos estructurales, unida al modelado de las relaciones no lineales mediante redes neuronales, permite la formulación de un modelo híbrido estructural no lineal, que combina las ventajas de ambas herramientas.

El modelo planteado fue evaluado mediante la aplicación a dos series reales. El modelado de los dos casos demostró que el modelo planteado en comparación con representaciones de orden lineal y no lineal presenta

mejores resultados, toda vez que reduce los valores del error.

El modelo híbrido estructural no lineal demostró su capacidad de capturar diferentes comportamientos lineales y no lineales que pueden estar presentes dentro de una serie temporal, lo cual constituye una ganancia tanto a nivel conceptual, puesto que permite interpretar individualmente la influencia de cada componente de la serie, incluyendo la no linealidad en los residuales, como a nivel matemático, pues comparativamente con modelos individuales, el híbrido propuesto es numéricamente mejor.

REFERENCIAS

- Box, G. y Jenkins, G. (1970), *Time Series Analysis: Forecasting and Control*, San Francisco.
- Chatfield, C. (1988), 'What is the 'best' method of forecasting?', *J. Appl. Statist.* **15**, 19-39.
- Chatfield, C. (1996), 'Model uncertainty and forecast accuracy', *J. Forecasting* **15**, 495-508.
- Clemen, R. (1989), 'Combining forecasts: a review and annotated bibliography with discussion', *Int. J. Forecasting* **5**, 559-608.
- DasGupta, B., Siegelmann, H. T. y Sontag, E. (1995), 'On the complexity of training neural networks with continuous activation functions', *IEEE Transactions on Neural Networks* **6**, 6.
- De Gooijer, J. y Kumar, K. (1992), 'Some recent developments in non-linear time series modelling, testing, and forecasting', *Int. J. Forecasting* **8**, 135-156.
- Durbin, J. y Harvey, A. (1985), The effects of seat belt legislation on road casualties in Great Britain: Report on assessment of statistical evidence, Technical report, HMSO, London.
- Engle, R. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations', *Econometrica* **50**, 987-1007.
- Ginzburg, I. y Horn, D. (1994), 'Combined neural networks for time series analysis', *Adv. Neural Inf. Process. Systems* **6**, 224-231.
- Grainger, C. (1989), 'Combining forecasts-twenty years later', *J. Forecasting* **8**, 167-173.
- Granger, C. y Andersen, A. (1978), *An Introduction to Bilinear Time Series Models*.
- Granger, C. y Teräsvirta, T. (1993), *Modeling Nonlinear Economic Relationships*.
- Harvey, A. (1989), *Forecasting, Structural Time Series and the Kalman Filter*.
- Hornik, K., Stinchcombe, M. y White, H. (1990), 'Using multi-layer feedforward networks for universal approximation', *Neural Networks* **3**, 551-560.
- Hung, M. y Denton, J. (1993), 'Training neural networks with the GRG2 nonlinear optimizer', *Eur. J. Oper. Res.* **69**, 83-91.
- Kaastera, I. y Boyd, M. (1996), 'Designing a neural network for forecasting financial and economic series', *Neurocomputing* **10**, 215-236.
- Krogh, A. y Vedelsby, J. (1995), 'Neural network ensembles, cross validation, and active learning', *Adv. Neural Inf. Process.* **7**, 231-238.
- Luxhoj, J., Riis, J. y Stensballe, B. (1996), 'A hybrid econometric-neural network modeling approach for sales forecasting', *Int. J. Prod. Econ.* **43**, 175-192.
- Makridakis, S. (1989), 'Why combining works?', *Int. J. Forecasting* **5**, 601-603.
- Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibdon, M., Lewandowski, R., Newton, J., Parzen, E. y Winkler, R. (1982), 'The accuracy of extrapolation (time series) methods: results of a forecasting competition', *J. Forecasting* **1**, 111-153.
- Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Millers, T., Ord, K. y Simmons, L. (1993), 'The M-2 competition: a real-life judgmentally based forecasting study', *Int. J. Forecasting* **9**, 5-29.
- Masters, T. (1995), *Neural, Novel and Hybrid Algorithms for Time Series Prediction*.
- Mehra, P. y Benjamin, W. (1997), *Artificial Neural Networks: Concepts and Theory*.
- Newbold, P. y Granger, C. (1974), 'Experience with forecasting univariate time series and the combination of forecasts (with discussion)', *J. R. Statist. Soc.* **137**, 131-164.
- Palm, F. y Zellner, A. (1992), 'To combine or not to combine? issues of combining forecasts', *J. Forecasting* **11**, 687-701.
- Pelikan, E., De Groot, C. y Wurtz, D. (1992), 'Power consumption in west-bohemia: improved forecasts with decorrelating connectionist networks', *Neural Network World* **2**, 701-712.
- Peretto, P. (1992), *An Introduction to the Modeling of Neural Networks*.
- Perrone, M. y Cooper, L. (1993), *Neural Networks for Speech and Image Processing*, Chapman & Hall, London, chapter When networks disagree: ensemble method for hybrid neural networks, pp. 126-142.
- Pole, A., West, M. y Harrison, J. (1994), *Applied Bayesian Forecasting and Time Series Analysis*, NY.
- Subramanian, V. y Hung, M. (1993), 'A GRG2-based system for training neural networks: design and computational experience', *ORSA J. Comput.* **5**, 386-394.

- Tong, H. (1990), *Non-linear Time Series: A Dynamical System Approach*, Oxford Statistical Science Series.
- Velásquez, J. (2003), Construcción de Escenarios de Pronóstico del Precio de Electricidad en Mercados de Corto Plazo, Propuesta de tesis doctoral, Facultad de Minas, Universidad Nacional de Colombia.
- Wedding II, D. y Cios, K. (1996), 'Time series forecasting by combining rbf networks, certainty factors, and the box-jenkins model', *Neurocomputing* 10, 149-168.
- Winkler, R. (1989), 'Combining forecasts: a philosophical basis and some current issues', *Int. J. Forecasting* 5, 605-609.
- Zhang, G. (2003), 'Time series forecasting using a hybrid ARIMA and neural network model', *Neurocomputing* 50, 159-175.
- Zhang, G., Patuwo, E. y M.Y. Hu, M. (1998), 'Forecasting with artificial neural networks: the state of the art', *Int. J. Forecasting* 14, 35-62.



ADPOSTAL

Llegamos a todo el mundo!