

Modelado de Cambios de Nivel en Series de Tiempo No Lineales Usando Redes Neuronales Artificiales

Paola Andrea Sánchez y Juan David Velásquez

UNIVERSIDAD NACIONAL DE COLOMBIA.

Facultad de Minas. Escuela de Sistemas

{pasanche ; jdvelasq}@unalmed.edu.co

Recibido para revisión Sep 2004, aceptado Nov 2004, versión final recibida Nov 2004

Resumen: Es comúnmente aceptado que muchas variables físicas y económicas presentan comportamientos dinámicos no lineales, cuya complejidad hacen que sea imposible formular un modelo econométrico basado en leyes físicas o económicas que representen su evolución adecuadamente. El problema de la formulación del modelo se ve agravado por la presencia de observaciones atípicas y cambios estructurales, para las cuales no existen modelos matemáticos que permitan su representación en el caso no lineal. En este estudio, se presentan tres nuevos modelos que permiten la representación de cambios de nivel en series de tiempo no lineales usando perceptrones multicapa, los cuales son ejemplificados para una serie ficticia y una serie real ampliamente conocida y estudiada. Los resultados indican las bondades de los modelos propuestos, permitiendo que ellos sean recomendados como parte integral de las metodologías para la modelación de series de tiempo usando redes neuronales artificiales.

Palabras Clave: Series de Tiempo, Cambios Estructurales, Cambios en el Nivel, Series de Tiempo No-Lineales, Redes Neuronales Artificiales, Perceptrones Multicapa.

Abstract: It is usually accepted that many physical and economical variables shown nonlinear dynamical behaviors which complexity makes it impossible to formulate an econometric model based only on physical or economic laws that properly represents its evolution. The formulation of model is aggravated with the presence of outliers and structural breaks, for which there are not any mathematical models that allow its representation in the nonlinear case. In this paper, three new models are presented which allow the representation of level shifts in nonlinear time series using multilayer perceptrons, which are illustrated for one artificial and one real series, and it is proposed a methodological strategy for the construction and specification of such models, based on a previous linear modeling. The results indicate the goodness of the proposed models, allowing them to be recommended as an integral part of the methodologies for modeling time series using artificial neural networks.

Keywords: Time Series, Structural Breaks, Level Shifts, Nonlinear Time Series, Artificial Neural Networks, Multilayer Perceptrons.

1 INTRODUCCIÓN

El análisis estadístico de series de tiempo ha sufrido un importante desarrollo en la última centuria, especialmente en el caso lineal, cuyos modelos y métodos han sido ampliamente estudiados y claramente entendidos [véase por ejemplo Box y Jenkins (1970)]; como evidencias de ello, las metodologías desarrolladas han sido aplicadas con éxito en la solución de problemas reales en áreas como la economía y la ingeniería. No obstante, se ha reconocido ampliamente que muchas variables económicas y financieras poseen un compor-

tamiento dinámico no lineal, y que en muchos de los casos, no existen suficientes leyes físicas o económicas que permitan especificar completamente un modelo estadístico para su representación. En consecuencia, diferentes modelos paramétricos no lineales han sido propuestos en las últimas décadas [una amplia recopilación es presentada en Tong (1990)], intentando caracterizar la forma de no linealidad presente en los datos, los cuales son especificados de manera empírica a partir de la información disponible [véase Granger y Teräsvirta (1993)]; su éxito en casos prácticos depende del grado de ajuste entre las características de la serie temporal estudiada y

las propiedades del modelo utilizado para representar su dinámica. No obstante, la formulación de cada familia de modelos requiere un tipo determinado de no linealidad, por lo que no son universalmente aplicables; es así como las técnicas no paramétricas de regresión no lineal surgen como alternativas atractivas, ya que permiten extraer las relaciones determinísticas no lineales presentes entre los regresores y la variable dependiente, sin realizar ningún tipo de supuestos. De dichas técnicas, los perceptrones multicapa son uno de los modelos más ampliamente usados en la solución de problemas de regresión y series de tiempo [véase Masters (1995)].

La especificación del modelo de la serie temporal se ve comúnmente afectada por la presencia de observaciones atípicas y cambios estructurales, los cuales modifican las propiedades estadísticas de la serie temporal estudiada, puesto que el modelo final no está en capacidad de representar adecuadamente la dinámica de la serie, y en consecuencia, durante su interpretación económica se pueda llegar a conclusiones incorrectas o inexactas [véase Tsay (1988) y Junttila (2001)].

Dicho problema ha sido ampliamente estudiado en el caso lineal, y se han formulado varias metodologías encaminadas a la detección de cambios en la estructura de la serie temporal y la presencia de datos atípicos [véase por ejemplo Tsay (1988), Fox (1972), Box y Tiao (1975), y Chen y Liu (1993)]. Aunque se han realizado algunos trabajos en el contexto de modelos no lineales [véase Van Dijk, Franses y Lucas (1996) y Tolvi (1998)], existe realmente muy poca experiencia sobre este tópico.

Este artículo se concentra en la problemática del modelamiento de cambios de nivel en series de tiempo no lineales, cuando ellas son modeladas usando perceptrones multicapa, siendo nuestro principal aporte, la formulación de varios modelos híbridos que permiten dar cuenta de este tipo de fenómeno.

Para alcanzar el objetivo propuesto, se ha organizado el resto de este artículo en la siguiente forma: en la Sección 2, se discute acerca del modelamiento tradicional de series de tiempo con cambios estructurales, para introducir posteriormente en la Sección 3, el modelamiento de series de tiempo utilizando perceptrones multicapa; seguidamente en la Sección 4, se presenta una propuesta metodológica para el modelamiento de una serie temporal con cambios de nivel mediante la utilización de perceptrones multicapa; en la Sección 5, se ilustran la metodología propuesta con un caso de ejemplo simulado y un caso real; y finalmente en la Sección 6, se discuten las principales lecciones aprendidas durante la realización de este trabajo.

2 MODELAMIENTO TRADICIONAL DE CAMBIOS ESTRUCTURALES EN SERIES DE TIEMPO

En el modelamiento estructural de series de tiempo [véase Harvey (1989) y Pole, West y Harrison (1994)] se asume que una serie temporal puede ser descompuesta en varios componentes aditivos que reflejan sus diferentes propiedades. Dichas componentes son: el nivel o tendencia, los patrones estacionales, los ciclos de largo plazo, la influencia de variables exógenas, y finalmente, la componente irregular.

Los cambios estructurales se presentan cuando hay modificaciones instantáneas o graduales en uno o más de las componentes previamente enunciadas, debido a eventos específicos [véase Rodríguez (2002)]. Así mismo, pueden ocurrir observaciones puntuales con muy baja probabilidad de ocurrencia, conocidas usualmente como outliers [véase Fox (1972), Tolvi (2001) y Tsay (1988)]. Además, ante el conocimiento *a priori* de la presencia de un cambio estructural o un outlier en la serie temporal, el modelador puede introducir modificaciones exógenas al pronóstico del modelo, las cuales son conocidas con el nombre de intervenciones [véase Box y Tiao (1975)].

Un cambio estructural se caracteriza por afectar de forma permanente la componente determinística de la serie temporal. En el caso lineal, se han desarrollado procedimientos generales basados en modelos de tipo ARMA y ARIMA, orientados especialmente a su detección, los cuales han sido ampliamente utilizados para explicar el efecto en la economía de eventos tales como las reformas políticas, fiscales y económicas. No obstante, el proceso generador de la serie temporal puede ser no lineal.

Los cambios estructurales han sido usualmente modelados a través de una componente aditiva que representa la perturbación exógena, la cual es usualmente modelada como una función paramétrica $f(t)$ Tsay (1988), de tal forma que la serie temporal y_t es descrita por el modelo:

$$y_t = f(t) + Z_t \quad (1)$$

donde Z_t representa la componente regular y determinística de la serie que a menudo es modelada con representaciones de tipo ARMA, de la forma:

$$Z_t = \frac{\theta_q(B)}{\phi_p(B)} \nu_t \quad (2)$$

donde $\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ y $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$, son respectivamente los polinomios autorregresivos y de medias móviles, y ν_t representa los errores de la serie los cuales siguen un proceso de ruido blanco [véase Box y Jenkins (1970)]. Este artículo se centra en la representación autorregresiva de la serie y se asume que la parte de medias móviles puede

ser contenida dentro de los errores, de tal forma que $\theta_d(B) = 1$.

Según el planteamiento tradicional, el efecto de irregularidades en una serie temporal representado por $f(t)$ puede ser capturado por una variable ficticia explicativa $I_d(t)$, que lo modela. La función $f(t)$ puede ser descrita de forma general como:

$$f(t) = \omega_0 \cdot \phi_t \cdot I_d(t) \quad (3)$$

donde, $I_d(t)$ representa la variable explicativa que indica la ocurrencia de una perturbación en el instante de tiempo d ; ω_0 equivale al impacto inicial de la perturbación; y ϕ_t es el parámetro que acompaña la variable explicativa y determina la forma dinámica de dicha perturbación [véase Box y Tiao (1975)].

Es así como un evento transitorio puede modelarse mediante una variable explicativa de tipo pulso la cual es usada para representar un cambio en un instante de tiempo que desaparece posteriormente; o como una variable transitoria que representa un cambio en un instante de tiempo cuya influencia va disminuyendo en el tiempo.

Mientras que un evento permanente, como los presentados en este trabajo, es modelado mediante una variable explicativa de tipo escalón o salto, que es usada para representar cambios en un instante de tiempo cuyo efecto ejerce una influencia sostenida sobre la serie, modificando así su comportamiento, y toma el valor de cero antes del evento y uno a partir de él, del tal forma que $I_d(t)$ se representa de la forma:

$$I_d(t) = \begin{cases} 0 & \text{Si } t < d \\ 1 & \text{Si } t \geq d \end{cases} \quad (4)$$

Ahora bien, ha sido comúnmente asumido que la dinámica de la serie temporal y_t tiene un comportamiento lineal, y que si existen evidencias de no linealidad, ésta es causada por la presencia de los cambios estructurales de carácter permanente [véase Tolvi (2001)]. No obstante, la dinámica de y_t puede ser no lineal, y si este hecho no es tenido en cuenta, se puede llegar a una especificación inadecuada de su modelo, manifestado mediante una representación errónea de las propiedades de ésta, y en consecuencia, puede conducir a conclusiones incorrectas en su interpretación económica [véase Tsay (1988) y Junttila (2001)].

Si bien en las últimas décadas se han hecho acercamientos a esta problemática vía modelos paramétricos no lineales, la principal falencia que se tiene en estos modelos es que aplican un patrón de no linealidad específico para la serie analizada, no siendo adecuados para modelar otros tipos de no linealidad.

Otra forma alternativa de abordar el problema es la utilización de métodos no paramétricos de regresión no lineal, como son las redes neuronales, los cuales han

demostrado tener la capacidad de reconocer y modelar comportamientos atípicos [véase Franses y Van Dijk (2000)], sin una forma de no linealidad predefinida. No obstante, se desconoce si existen señales que demuestren que la red neuronal está en capacidad de detectar dichas anomalías como cambios estructurales en la serie, por lo que la meta de éste trabajo es la investigación de ese comportamiento.

3 MODELADO DE SERIES DE TIEMPO USANDO PERCEPTRONES MULTICAPA

Los perceptrones multicapa son un tipo de red neuronal artificial (RNA) 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 de tiempo, el modelo del perceptrón 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. Un perceptrón multicapa con una capa oculta usando la función de transferencia lineal para realizar la activación de su capa de salida, puede ser representado matemáticamente como una función, que relaciona el grupo de regresores x_t de la serie temporal Z_t , tal que:

$$Z_t = \alpha_0 + \sum_{j=1}^q \alpha_j g[\beta_{0j} + (\sum_{i=1}^p \beta_{ij} x_t^{(i)})] + \epsilon_t \quad (5)$$

donde α_j representa los pesos de la capa oculta a la capa de salida; los coeficientes β_{ij} son los pesos de la capa de entrada a la capa oculta; 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, que corresponde en este caso a la función sigmoidea bipolar:

$$g(u) = \frac{2}{1 + \exp(-u)} - 1 \quad (6)$$

Según el modelo expresado en (5), la entrada a cada neurona de la capa oculta es la parte determinística de un modelo autoregresivo, y en consecuencia, puede considerarse que la red neuronal es la suma ponderada de la transformación no lineal de tantos modelos autoregresivos como neuronas posea en la capa oculta. Igualmente, usando una función de activación logística, 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

de tiempo es equivalente a una combinación no lineal de modelos autoregresivos, y puede expresarse de la forma:

$$Z_t = \alpha_0 + \sum_{j=1}^q \alpha_j g[w_j(B) \cdot x_t] + \epsilon_t \quad (7)$$

donde $w_j(B) = \beta_{1j}B + \beta_{2j}B^2 + \dots$ representa los parámetros β_{ij} del modelo asociados al operador de rezagos $B = 0 + B + B^2 + \dots$, correspondientes al modelo autorregresivo.

El atractivo que poseen las RNA en el modelamiento de series de tiempo es su habilidad para identificar dependencias ocultas con base en una muestra finita, especialmente de orden no lineal [véase Kaastra y Boyd (1996) y Masters (1995)]. La principal ventaja de éste enfoque sobre otros modelos es que no parten de supuestos *a priori* sobre la forma de la no linealidad presente en la serie temporal, realizando su aprendizaje a partir de la muestra de observaciones.

Adicionalmente, como se presenta en Franses y Van Dijk (2000), las redes neurales han demostrado tener la capacidad de reconocer y modelar comportamientos atípicos, tales como "outliers" o cambios estructurales en las series de tiempo. No obstante, no existe claridad acerca del comportamiento de los perceptrones multicapa en el modelamiento de series de tiempo con cambios en alguna de sus componentes estructurales, y más aún hay incertidumbre acerca de las señales que conducen a la detección de dichos cambios.

1 PROPUESTA METODOLÓGICA

Como se discutió anteriormente, la presencia de no linealidades en la serie temporal combinada con la existencia de cambios estructurales de carácter permanente puede conducir a una especificación errónea del modelo. Ante tal problemática, la utilización de los modelos lineales tradicionales queda desestimada, siendo más atractiva la utilización de métodos no paramétricos que permitan la aproximación de la serie en presencia de no linealidades, como es el caso de los perceptrones multicapa.

1.1 Modelo Propuesto de Series de Tiempo No Lineales con Cambio en el Nivel

Tomando como base el planteamiento tradicional de series de tiempo con cambios estructurales presentado en (1) y (3), y teniendo en cuenta que una red neuronal, analizada en el contexto del modelamiento de series de tiempo, puede ser interpretada como una combinación no lineal de modelos autorregresivos, tal como se expresa en (7), es posible extrapolar dicho planteamiento para obtener un modelo general de la parte autoregresiva de una serie temporal con cambios estructurales utilizando redes neuronales, mediante tres enfoques, donde

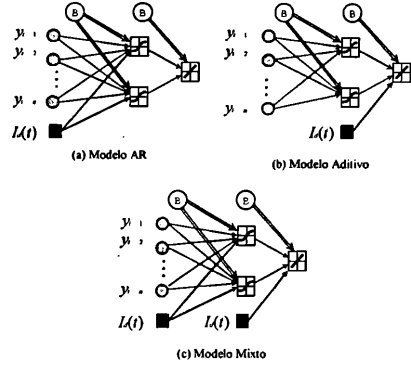


Figura 1: Esquema Representativo de un Perceptrón Multicapa estándar, autorregresivo, aditivo y mixto

$f(t)$ está únicamente representada por su variable explicativa $I_d(t)$, como se expresa en (4):

$$y_t = \alpha_0 + \sum_{j=1}^q \alpha_j g[\beta_{0j} + (\sum_{i=1}^p \beta_{ij} y_{t-i}) + \phi_1 \cdot I_d(t)] + \epsilon_t \quad (8)$$

$$y_t = \alpha_0 + \sum_{j=1}^q \alpha_j g[\beta_{0j} + (\sum_{i=1}^p \beta_{ij} y_{t-i})] + \phi_1 \cdot I_d(t) + \epsilon_t \quad (9)$$

$$y_t = \alpha_0 + \sum_{j=1}^q \alpha_j g[\beta_{0j} + (\sum_{i=1}^p \beta_{ij} y_{t-i}) + \phi_1 \cdot I_d(t)] + \phi_2 \cdot I_d(t) + \epsilon_t \quad (10)$$

donde, el primero [véase la expresión (8) y la Figura 1(a)] es equivalente a la representación de un modelo autorregresivo, para el cual la perturbación $I_d(t)$ en la serie temporal es tomada como una entrada adicional a la red neuronal, y ϕ_1 es el parámetro que acompaña la perturbación; el segundo [véase la expresión (9) y la Figura 1(b)] equivale a un modelo aditivo, en el cual la perturbación es tomada como una entrada que afecta la salida de la capa oculta de la red; y el tercero [véase la expresión (10) y la Figura 1(c)] es una combinación de los dos enfoques anteriores.

La Figura 1 esquematiza una perceptrón multicapa estándar con n regresores en la capa de entrada, dos neuronas en la capa oculta, una en su capa de salida y los tres esquemas propuestos.

Ahora bien, como caso especial de cambio estructural se considera el cambio en el nivel de la serie. Dicho cambio representa un evento cuyo efecto sobre la serie

afecta de forma permanente el nivel de la serie y por ende, modifica su estructura. Estos han sido investigados en Box y Tiao (1975), con un tratamiento extendido presentado en Tsay (1988), el cual hace uso de un procedimiento iterativo para su detección basado en un modelamiento inicial de tipo ARIMA [véase además Chen y Liu (1993), Fox (1972) y Vaage (2000)]. Siguiendo el planteamiento (8) [se procede de igual manera para representar la red neuronal según los planteamientos (9) y (10)], la forma simplificada del modelo de la red neuronal para una serie temporal con cambios de nivel es equivalente a:

$$y_t = \begin{cases} \alpha_0 + \sum_{j=1}^q \alpha_j g[\beta_{0j} + (\sum_{i=1}^p \beta_{ij} y_{t-i})] + \epsilon_t & \text{Si } t < d \\ \alpha_0 + \sum_{j=1}^q \alpha_j g[\beta_{0j} + (\sum_{i=1}^p \beta_{ij} y_{t-i} + \phi_1)] + \epsilon_t & \text{Si } t \geq d \end{cases} \quad (11)$$

o, expresado de otra forma

$$y_t = \alpha_0 + \sum_{j=1}^q \alpha_j g[\beta_{0j} + (\sum_{i=1}^p \beta_{ij} y_{t-i} + \phi_1 \cdot I_d(t))] + \epsilon_t \quad (12)$$

donde el modelo mostrado en (12) expresa que un cambio de nivel de magnitud ϕ_1 ocurre en un tiempo $t = d$ y que dicho cambio tiene un efecto permanente en la serie temporal y_t .

4.2 Estrategia para la Especificación y Construcción del Modelo

Para la especificación del modelo de la serie temporal no lineal con cambio en el nivel, diferentes estrategias han sido propuestas en la literatura para el caso lineal [véase por ejemplo Tsay (1988)] y estas pueden ser ajustadas para el caso no-lineal; es así como en la Tabla 1, se propone un procedimiento tentativo para la especificación y construcción del modelo de la serie temporal, el cual se compone básicamente de tres pasos:

Tabla 1: Estrategia de Especificación y Construcción del Modelo de una Serie Temporal No Lineal con Cambio en el Nivel

Paso 1:	Construcción de un Modelo Lineal
	<i>Análisis Exploratorio Preliminar</i>
	<i>Identificación - Estimación - Diagnóstico</i>
	<i>Prueba de Linealidad</i>
	<i>Especificación del Modelo</i>
Paso 2:	Construcción de un Modelo No Lineal
Paso 3:	Diagnóstico del Modelo No Lineal

4.2.1 Paso 1: Construcción de un Modelo Lineal

En el paso 1, se estima un modelo lineal para la serie temporal antes del cambio de nivel [dicho cambio se asume como conocido], de acuerdo a las directrices generales para la construcción de un modelo ARIMA [véase procedimiento propuesto por Box y Jenkins (1970)]; este Paso se realiza con la finalidad de identificar el conjunto de regresores y_{t-i} que afectan la serie. Adicionalmente se evalúa la presencia de no linealidades en la serie, con el fin de justificar la aplicación de tal forma que se justifique la utilización del modelo no lineal propuesto. En este artículo la evaluación de la linealidad se hace mediante la utilización de la prueba RESET para contrastar linealidad propuesta por Ramsey (1969) [véase Granger y Teräsvirta (1993)].

4.2.2 Paso 2: Construcción de un Modelo No Lineal

Con base en el paso anterior, en el Paso 2, se estima un modelo no lineal según (8) [se procede de igual forma para el cálculo de los modelos (9) y (10)], donde los regresores del modelo lineal calculado en el Paso 1, son tomados como entradas a la red.

El cálculo del número de neuronas en la capa oculta de la red neuronal se hace mediante un algoritmo constructivo, en el cual la red es inicializada con una neurona y de acuerdo a su necesidad nuevas unidades son agregadas. En cada ciclo una nueva neurona es adicionada a la capa oculta y el perceptrón multicapa completo es reentrenado totalmente hasta alcanzar el mínimo error para la arquitectura actual; esta nueva neurona es aceptada si el error decrece. Igualmente se utiliza una estrategia combinada para el entrenamiento de la red neuronal donde los parámetros del modelo son estimados conjuntamente con el número de neuronas ocultas. Las soluciones del sistema de ecuaciones así generado, son obtenidas usando mínimos cuadrados.

4.2.3 Paso 3: Diagnóstico del Modelo No Lineal

Finalmente, con el fin de validar la adecuación del mejor modelo encontrado a la serie se examinan los residuales de éste. Aquí se verifica que dichos residuales cumplen con los supuestos establecidos sobre ellos, es decir, que son incorrelacionados y normalmente distribuidos con media cero y varianza constante. En este artículo el análisis es realizado mediante la evaluación de los gráficos de los residuales con respecto al tiempo y los residuales cuadráticos, para validar los supuestos acerca de la media y la varianza; el autocorrelograma de los residuales, para la validación del supuesto de incorrelación, y el histograma de los residuales estandarizados, para validar la hipótesis de normalidad. El no cumplimiento de los supuestos sobre los residuales conduce a una nueva etapa de identificación.

Una vez estimado el modelo, es posible comparar su desempeño respecto a otros modelos, con base en un criterio común. En éste artículo, los modelos encontrados son comparados con los de un preceptrón multicapa simple calculado según (5), en donde no se tienen en cuenta las perturbaciones en el cálculo del modelo de la red, con base en el criterio del MSE [Mean Square Error].

El ajuste del modelo encontrado a la serie, se realiza mediante la inspección visual del gráfico de dispersión entre la serie y su modelo calculado. En éste un buen ajuste se ve reflejado en una distribución uniforme de los datos alrededor de una línea de tendencia de 45°.

5 RESULTADOS Y ANÁLISIS

Para demostrar las cualidades y la efectividad de los modelos propuestos, se realizaron dos pruebas, la primera con una serie temporal no lineal simulada, y la siguiente con una serie real.

La serie simulada fue generada de acuerdo a un proceso de orden no-lineal LSTAR (*Logistic Smooth Transition Autorregresivo*) y consiste de 250 observaciones para la cual se aplicó un cambio estructural permanente en su nivel a partir de la observación 200.

La serie de Muertes por accidentes automovilísticos en Gran Bretaña, es usada como caso real para evaluar el método propuesto. Esta serie fue estudiada en Harvey (1989) mediante la utilización de modelos estructurales y posee la característica de tener un comportamiento no lineal y presentar además un cambio estructural en su nivel.

5.1 Caso 1: Serie Simulada

5.1.1 Descripción de la Serie

La serie no lineal simulada Z_t fue construida según el modelo:

$$Z_t = (\phi_{1,0} + \phi_{1,1} \cdot Z_{t-1}) \cdot (1 - G(Z_{t-1}; \gamma, c)) + (\phi_{2,0} + \phi_{2,1} \cdot Z_{t-1}) \cdot G(Z_{t-1}; \gamma, c) + \epsilon_t \quad (13)$$

con:

$$G(s_t; \gamma, c) = \frac{1}{1 + \exp\{-\gamma \cdot (s_t - c)\}}$$

donde, el modelo expresado es no lineal con componente AR(1), y los errores son *iid* y se distribuyen como una normal (0,1). En la Tabla 2 aparecen los valores seleccionados para cada parámetro de la ecuación (13).

Ahora bien, la serie de tiempo y_t fue construida adicionando al modelo no lineal calculado Z_t , una perturbación permanente $I_d(t)$ a partir del instante de tiempo $d = 200$. El valor para el parámetro de la función de

Tabla 2: Parámetros Seleccionados para el Modelo No Lineal

$\phi_{1,0}$	$\phi_{1,1}$	$\phi_{2,0}$	$\phi_{2,1}$	γ	c
1.5	-0.5	-1.5	0.5	1.5	0.5

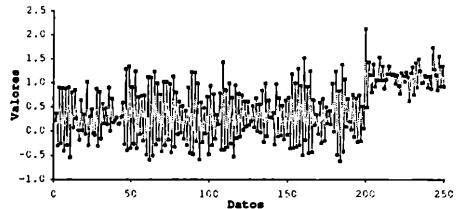


Figura 2: Serie de Tiempo No Lineal Simulada con Cambio en el Nivel

perturbación es unitario. En la Figura 2, aparece la serie de tiempo no lineal simulada con cambio estructural permanente en el nivel.

5.1.2 Identificación de la Serie

El análisis preliminar de la serie presente en la Figura 5.1.1, muestra que aunque antes del cambio de nivel [observación 200] la serie parece oscilar alrededor de un valor medio, la presencia de este cambio modifica su comportamiento, además ésta presenta altos niveles de volatilidad asociados, por lo tanto, su media y su varianza no son constantes.

El modelamiento lineal preliminar calculado para la serie indica la existencia de una relación causal entre los valores con rezagos 1, 2, 3 y 4, y el valor actual. Cabe aclarar que no es el objetivo de éste trabajo el modelamiento lineal, por lo que se omiten los resultados de éste proceso.

Para ratificar la no linealidad de la serie se aplicó la prueba *RESET*, en cuyo contraste se obtuvo que $R = 35.3597 > VC = 2.0477$, lo cual implica que con una significancia de 95% la hipótesis nula de linealidad es rechazada, por lo que es correcto afirmar que la serie estudiada es no lineal.

Hasta este punto se ha probado que la serie simulada tiene un comportamiento no lineal, que posee un cambio en su nivel en la observación 200 y que además presenta una relación causal entre los valores con rezagos 1, 2, 3 y 4, y el valor actual.

5.1.3 Resultados de la Estimación y Diagnóstico del Modelo No Lineal

Según el procedimiento planteado, el segundo paso es desarrollar un modelo de red neuronal para la serie estudiada. La representación de la red neuronal utilizada consiste de un perceptrón multicapa con una capa oculta activada con función logística bipolar, y una capa de salida activada con transferencia lineal. Los pesos de las conexiones son calculados usando un algoritmo híbrido que combina Estrategias de Evolución y Mínimos Cuadrados. Las entradas de la red equivalen a los respectivos valores rezagados de la serie simulada para un modelo lineal, rezagos 1, 2, 3 y 4, y los valores correspondientes a la variable de perturbación $I_d(t)$ con $d = 200$.

Siguiendo las condiciones mencionadas, fueron estimados los respectivos modelos propuestos en (8), (9) y (10), de la serie simulada, así como también, los modelos de perceptrón multicapa simple para cada caso (este modelo no tiene en cuenta las perturbaciones para su cálculo). En la Tabla 3, aparecen los valores calculados del MSE de cada uno de los modelos. Según dichos resultados, el modelo que mejor se ajusta a la serie simulada equivale a uno mixto con perturbación y con una neurona en la capa oculta [véase Figura 3], puesto que presenta el menor MSE.

Tabla 3: Valores del MSE de los Modelos Propuestos para la Serie No Lineal Simulada con Cambio en el Nivel

	Sin Cambio	Con Cambio
Modelo NN - AR	0.07821	0.05230
Modelo NN - Aditivo	0.07821	0.04871
Modelo NN - Mixto	0.07821	0.04851

En comparación con el perceptrón multicapa simple, tanto el modelo mixto, como los demás propuestos presentan valores menores para el MSE, obteniéndose mejoras del 33.1%, 37.7% y 38.0%, respectivamente; lo que permite afirmar que la adición de la variable de perturbación al modelo mejora la adecuación de éste a la serie. La Figura 3, muestra la serie simulada real [círculos], así como el mejor modelo calculado [línea continua], en ésta puede observarse que aunque existen puntos donde el modelo se desvía de los valores reales, en general, éste se ajusta tanto a los valores como al comportamiento de la serie.

La Figura 4, presenta el gráfico de dispersión entre la serie simulada y el modelo Mixto calculado para ésta. La nube de puntos oscila alrededor de una línea de tendencia aproximada de 45°, para este tipo de gráficos ese

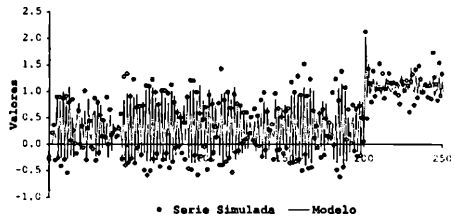


Figura 3: Modelo NN - Mixto calculado para la Serie No Lineal Simulada con Cambio en el Nivel

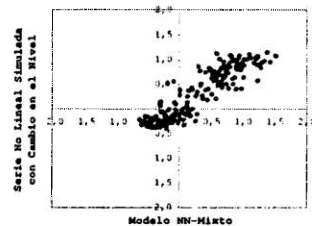


Figura 4: Gráfico de Dispersión entre la Serie No Lineal Simulada con Cambio en el Nivel y el Modelo NN-Mixto Ajustado

resultado demuestra que el modelo calculado se ajusta bastante bien a la serie, lo cual permite afirmar que dicho modelo es considerado un buen estimador de ésta.

Finalmente, para evaluar el comportamiento de los residuales del modelo ajustado se construyó una ilustración resumen de algunos gráficos estadísticos calculados para estos, dichos gráficos son presentados en la Figura 5. Los gráficos del lado izquierdo que representan los residuales del modelo y su autocorrelación, indican que dichos valores son incorrelacionados y se distribuyen uniformemente alrededor de cero, por lo que puede afirmarse que se comportan como ruido blanco; el gráfico superior derecho que corresponde a los residuales cuadráticos revela que aunque sí bien existen regiones con altos valores en los errores, lo que refleja lo expresado acerca de la Figura 4, estos tienden a presentar valores bajos y un comportamiento constante; por último, el gráfico inferior derecho que corresponde al histograma de frecuencias indica que los residuales tienen una distribución aproximadamente normal.

De lo anterior, los residuales del modelo ajustado se distribuyen independientemente como una normal ($0, \sigma_e^2$), lo cual permite reafirmar que el modelo calculado es un buen estimador de la serie.

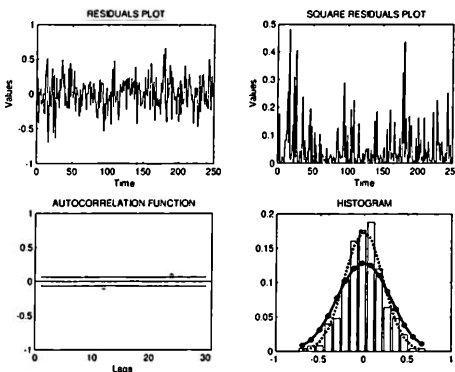


Figura 5: Gráficos de Residuales del Modelo NN - Mixto calculado para la Serie No Lineal Simulada con Cambio en el Nivel

5.2 Caso 2: Muertes por Accidentes Automovilísticos en Gran Bretaña

La Serie mensual de muertes por accidentes automovilísticos en Gran Bretaña para el periodo entre enero de 1981 y diciembre de 1984, fue estudiada en Harvey (1989), y contiene un total de 48 observaciones. La gráfica de la serie mostrada en la Figura 6, sugiere la presencia de un cambio estructural en el nivel alrededor de enero de 1983 [observación 25], debido a la reglamentación del uso obligatorio del cinturón de seguridad a partir de dicha fecha.

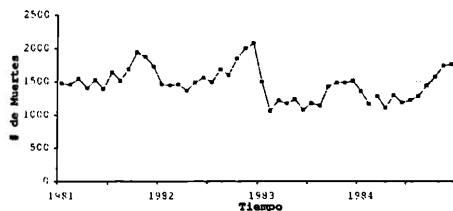


Figura 6: Muertes por Accidentes Automovilísticos en Gran Bretaña ene/81-dic/84

5.2.1 Identificación de la Serie

En la Figura 6, se observa la presencia de pequeñas volatilidades, la permanencia de un nivel aproximadamente constante antes y después del cambio estructural, y ciclos estacionales de periodo 12, cuyos mayores valores se presentan en los meses de diciembre y los menores

entre abril y junio.

El modelo lineal ajustado a la serie refleja la existencia de una relación causal entre los valores con rezagos 1 y 13, y el número de muertes actuales.

La prueba *RESET* aplicada a la serie, produjo un $R = 22.9215 > VC = 2.3358$, lo cual implica que con una significancia de 95% es correcto afirmar que la serie de Muertes en Gran Bretaña es no lineal.

De acuerdo a lo dicho hasta este punto, es posible afirmar que la serie de Muertes por accidentes automovilísticos en Gran Bretaña presenta un comportamiento no lineal, posee un cambio estructural en su nivel en enero de 1983 y que refleja una relación causal entre los valores con rezagos 1 y 13, y el número de muertes actuales.

5.2.2 Resultados de la Estimación y Diagnóstico del Modelo No Lineal

Siguiendo la misma representación para la red neuronal utilizada en la serie simulada y tomando como entradas de la red los valores rezagados 1 y 13, y los valores correspondientes a la variable de perturbación $I_d(t)$ con $d = 25$, fueron estimados los respectivos modelos propuestos en (8), (9) y (10) para la serie de Muertes por accidentes automovilísticos en Gran Bretaña y sus correspondientes representaciones de perceptrón multicapa simple [sin perturbación].

La Tabla 4 presenta los resultados calculados de $\sigma_a = \sqrt{MSE}$ para cada uno de los modelos. Los resultados sugieren que para el caso simple los tres modelos se ajustan de forma equivalente a la serie, mientras que al incluir la perturbación, tanto el modelo aditivo como el mixto se comportan de idéntica forma, presentando iguales valores para el parámetro σ_a , pero mejorando notablemente sus valores al incluir la variable de perturbación [los modelos NN-Aditivo y NN-Mixto logran una mejoraría del 48.8% respecto al modelo de red neuronal simple].

Tabla 4: Valores de σ_a de los modelos propuestos para la Serie de Muertes por Accidentes Automovilísticos en Gran Bretaña

	Sin Cambio	Con Cambio
Modelo NN - AR	162.9138	162.9138
Modelo NN - Aditivo	162.9138	83.4512
Modelo NN - Mixto	162.9138	83.4520

En la Figura 7, se presenta la serie real [círculos] y el modelo aditivo calculado [línea continua], donde en general puede observarse que el modelo se ajusta tanto a los valores como al comportamiento de la serie, dicho

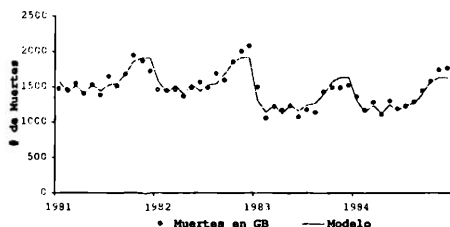


Figura 7: Modelo NN- Aditivo calculado para la Serie de Muertes por Accidentes Automovilísticos en Gran Bretaña

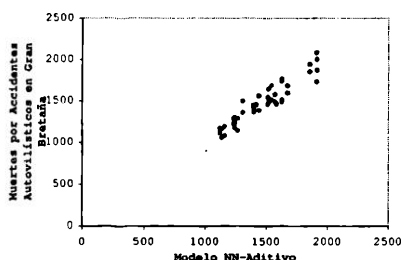


Figura 8: Gráfico de Dispersión entre la Serie de Muertes por Accidentes Automovilísticos en Gran Bretaña y el Modelo NN-Aditivo Ajustado

resultado demuestra que el modelo calculado captura el comportamiento de la serie.

La Figura 8, presenta el gráfico de dispersión entre la serie y el modelo aditivo calculado. Como puede observarse, la nube de puntos oscila alrededor de una línea de tendencia de 45° y no presenta valores extremos; dicho resultado demuestra que el modelo calculado se ajusta adecuadamente al comportamiento de la serie, lo que permite afirmar que dicho modelo es un buen estimador de ésta.

6 CONCLUSIONES

En éste artículo se aborda la problemática de modelamiento de series de tiempo con cambios en el nivel. Con la aplicación de un modelo general de perceptrón multicapa que aprovecha las bondades de la representación tradicional de dichos cambios, se demostró la viabilidad de modelar la parte autorregresiva de una serie temporal no lineal con cambios en su nivel.

Se realizó además una propuesta metodológica para la especificación y estimación del modelo de una serie temporal no lineal con cambio en el nivel, la cual fue evaluada posteriormente, mediante la aplicación a dos casos, uno simulado y uno real. Los resultados de la apli-

cación sugieren una clara ganancia en la representación de series con las características citadas, toda vez que reducen los niveles de error y permiten un análisis del efecto individual del cambio de nivel en la serie.

El modelo general de perceptrón multicapa considera tres diferentes enfoques, autorregresivo, aditivo y mixto, dependiendo de la especificación que se haga de las perturbaciones como entradas a la red. En la aplicación los tres modelos presentaron resultados aproximadamente similares, no obstante, no se ha evaluado cual de ellos es más adecuado para problemas de éste tipo.

Los resultados preliminares aquí expuestos, indican la viabilidad de utilizar modelos de perceptrón multicapa en la representación de series con las características citadas, y resaltan la importancia de realizar investigaciones posteriores que permitan formalizar de una manera más transparente una estrategia de especificación de modelos de series de tiempo basados en redes neuronales artificiales.

REFERENCIAS

- Box, G. y Jenkins, G. (1970), *Time Series Analysis: Forecasting and Control*, Holden-Day Inc, San Francisco.
- Box, G. y Tiao, G. (1975), 'Intervention analysis with applications to economic and environmental problems', *Journal of the American Statistical Association* **70**, 70-79.
- Chen, C. y Liu, L. (1993), 'Joint estimation and outlier effects in time series', *Journal of the American Statistical Association* **88**, 284-297.
- Fox, A. (1972), 'Outliers in time series', *Journal of the Royal Statistical Society Series B* **34**, 350-363.
- Franses, P. y Van Dijk, D. (2000), *Non-linear time series models in empirical finance*, Cambridge University Press.
- Granger, C. y Teräsvirta, T. (1993), *Modeling Nonlinear Economic Relationships*, Oxford University Press.
- Harvey, A. (1989), *Forecasting, Structural Time Series and the Kalman Filter*, Cambridge University Press.
- Junttila, J. (2001), 'Structural breaks, ARIMA model and Finnish inflation forecasts', *International Journal of Forecasting* **17**, 203-230.
- Kaasra, I. y Boyd, M. (1996), 'Designing a neural network for forecasting financial and economic series', *Neurocomputing* **10**, 215-236.
- Masters, T. (1995), *Neural, Novel and Hybrid Algorithms for Time Series Prediction*, John Wiley and Sons Inc.
- Pole, A., West, M. y Harrison, J. (1994), *Applied Bayesian Forecasting and Time Series Analysis*, Chapman & Hall, NY.

- Rodríguez, G. (2002), Modelos estructurales en el contexto de las series temporales económicas, Documento de trabajo, Universidad de La Laguna y Universidad de Las Palmas de Gran Canaria.
- Tolvi, J. (1998), Outliers en time series: A review, Research reports 76, University of Turku, Department of Economic.
- Tolvi, J. (2001), Nonlinear model selection in the presence of outliers, Research reports 90, University of Turku, Department of Economic.
- Tong, H. (1990), *Non-linear Time Series: A Dynamical System Approach*, Oxford Statistical Science Series, Clarendon Press Oxford.
- Tsay, R. (1988), 'Outliers, level shifts, and variance changes in time series', *Journal of Forecasting* 7, 1–20.
- Vaage, K. (2000), 'Detection of outliers and level shifts in time series: An evaluation of two alternative procedures', *Journal of Forecasting* 19, 23–37.
- Van Dijk, D., Franses, P. y Lucas, A. (1996), Testing for smooth transition nonlinearity in the presence of outliers, Research reports Report 9622/A, Erasmus University, Rotterdam. Econometric Institute.

Paola A. Sánchez es Ingeniera Industrial y Magister en Ingeniería de Sistemas de la Universidad Nacional de Colombia, y actualmente adelanta sus estudios de Doctorado en Ingeniería en la misma institución. Sus preferencias investigativas están relacionadas con la aplicación de técnicas de Inteligencia Artificial en el mejoramiento de modelos de series de tiempo y pronóstico.

Juan David Velásquez es profesor asistente en la Facultad de Minas, y candidato a doctor en el programa académico del Doctorado en Ingeniería - Sistemas Energéticos de la Universidad Nacional. Sus intereses de investigación incluyen la aplicación de técnicas de Inteligencia Artificial en el campo de la Econometría Financiera, y el modelado y predicción de series temporales no lineales en Mercados Energéticos.

