

# SISTEMAS ADAPTATIVOS DE INFERENCIA NEURODIFUSA CON ERRORES HETEROCEDÁSTICOS PARA EL MODELADO DE SERIES FINANCIERAS\*

*Elizabeth Catalina Zapata Gómez\*\**

*Juan David Velásquez Henao\*\*\**

*Ricardo Agustín Smith Quintero\*\*\*\**

---

\* Artículo de investigación científica y tecnológica. Es producto de la tesis de Maestría de E. C. Zapata *Modelado de series temporales usando sistemas adaptativos de inferencia neuro difusa (ANFIS) con heterocedasticidad condicional autorregresiva*. Patrocinado por la Facultad de Minas, Universidad Nacional de Colombia, Medellín, Colombia. El artículo se recibió el 07-03-2007 y se aprobó el 07-11-2008.

\*\* Magíster en Ingeniería de Sistemas, Universidad Nacional de Colombia, sede Medellín, Colombia, 2007; Ingeniera de Sistemas, Universidad Nacional de Colombia, 2004. Miembro del Grupo de Computación Aplicada, Facultad de Minas, Universidad Nacional de Colombia. Correo electrónico: eczapata@unal.edu.co.

\*\*\* Candidato a doctor en Ingeniería-Área Sistemas Energéticos, Universidad Nacional de Colombia, sede Medellín, Colombia; Magíster en Ingeniería de Sistemas, Universidad Nacional de Colombia, sede Medellín, Colombia, 1997; Ingeniero Civil, Universidad Nacional de Colombia, 1994. Profesor asociado de la Escuela de Sistemas, Facultad de Minas, Universidad Nacional de Colombia, sede Medellín. Director del Grupo de Computación Aplicada, Facultad de Minas, Universidad Nacional de Colombia. Correo electrónico: jdvelasq@unal.edu.co.

\*\*\*\* PhD in Civil Engineering, Colorado State University, Estados Unidos, 1981; Master of Science in Civil Engineering, Colorado State University, 1977. Profesor titular de la Escuela de Geociencias y Medio Ambiente, Facultad de Minas, Universidad Nacional de Colombia, sede Medellín, Colombia. Correo electrónico: rasmith@unal.edu.co.

## Sistemas adaptativos de inferencia neurodifusa con errores heterocedásticos para el modelado de series financieras

### RESUMEN

En este trabajo se propone una nueva clase de modelos híbridos no lineales. En el modelo propuesto, la no linealidad en la media se representa usando un sistema adaptativo neurodifuso de inferencia (ANFIS, por su sigla en inglés), mientras la varianza se representa usando una componente autorregresiva heterocedástica condicional. Se presenta la formulación matemática de este tipo de modelos y se propone un método para su estimación; adicionalmente, se desarrolla para el modelo propuesto una estrategia de especificación basada en una batería de pruebas estadísticas que incluyen pruebas para la especificación de los modelos de regresión con transición suave (STR, por su sigla en inglés) y la prueba del radio de verosimilitud. Como un caso de estudio, se modela la dinámica de la serie de los cambios en los precios de cierre de las acciones de IBM, la cual se usa comúnmente como referente en la literatura de series de tiempo. Los resultados indican que el modelo desarrollado representa mejor que otros modelos de características similares la dinámica de la serie estudiada.

**Palabras clave:** ANFIS, ARCH, heterocedasticidad, series temporales, modelos no lineales.

## Adaptive Neuro-Fuzzy Inference Systems with Heteroscedastic Errors for Financial Series Modeling

### ABSTRACT

This paper proposes a new kind of non-linear hybrid model. In the proposed model, mean non-linearity is represented by using an adaptive neuro-fuzzy inference system (ANFIS) whereas variance is represented using a conditional self-regressive heteroscedastic component. The mathematical formula for this type of model is shown and a method to estimate it is proposed. In addition, a specification strategy is developed for the proposed model, based on a battery of statistical soft transaction regression (STR) tests and on verosimilarity radius testing. As a case study, the IBM stock closing price series dynamics were modeled, which is commonly used as a benchmark in the literature on time series. Results indicate that the model developed represents the dynamics of the studied series better than other models with similar characteristics.

**Key words:** ANFIS, ARCH, heteroscedasticity, time series, non-linear models.

## Sistemas adaptativos de inferência em neuro difusão com erros heterocedásticos para o modelado de séries financeiras

### RESUMO

Neste trabalho propõem-se uma nova classe de modelos híbridos não lineais. No modelo proposto, a não linearidade em média representa-se utilizando um sistema adaptativo de neuro difusão de inferência (ANFIS, por sua sigla em inglês), enquanto a variação se representa usando um componente auto-regressivo heterocedástico condicional. Apresenta-se a formulação matemática deste tipo de modelos e propõem-se um método para sua estimação; adicionalmente, desenvolve-se para o modelo proposto uma estratégia de especificação baseada em uma bateria de provas estatísticas que incluem provas para a especificação dos modelos de regressão com transição suave (STR, por sua sigla em inglês), e a prova de rádio de verosimilitude. Como um caso de estudo, modela-se a dinâmica da série dos câmbios nos preços de feixe nas ações de IBM, a qual utiliza-se comumente como referente na literatura de séries de tempo. Os resultados indicam que o modelo desenvolvido representa melhor que outros modelos de características similares a dinâmica da série estudada.

**Palavras chave:** ANFIS, ARCH, heterocedasticidade, séries temporais, modelos não lineais.

## Introducción

El análisis y la predicción de series temporales se usan como un medio para entender su comportamiento histórico y realizar predicciones de muchas variables físicas, sociales, financieras y económicas (Tong, 1990). Ambas tareas se basan en la identificación, especificación, estimación y diagnóstico de un modelo matemático que represente la dinámica de la serie estudiada; para ello se tienen en cuenta las propiedades particulares de la serie que están relacionadas con sus características visuales, como la presencia de tendencias primarias o secundarias, los patrones repetitivos, los ciclos de largo plazo (Harvey, 1989), las linealidades, la varianza cambiante en el tiempo y los *clusters* de volatilidad. El gran número de dichas características, la complejidad de sus interacciones y la amplia gama de modelos propuestos en la literatura hacen que la especificación de un modelo para una serie particular sea tanto una ciencia como un arte.

Tradicionalmente, el análisis y la predicción de series temporales han estado ligados al uso de modelos lineales, particularmente a la aproximación propuesta por Box y Jenkins (1970), que se usa como un punto de partida para otras clases de modelos. Sin embargo, al reconocer que en muchos casos los modelos lineales no capturan de forma adecuada ciertos comportamientos de las series reales se abren las puertas al uso de modelos no lineales.

Consecuentemente, se han desarrollado muchas técnicas no lineales, tanto paramétricas como no paramétricas, fundamentadas en la

estadística, las matemáticas y las ciencias de la computación. De la amplia gama de modelos no lineales existentes, las técnicas de regresión no lineal emergidas del campo de la inteligencia artificial han tenido aplicaciones empíricas en el modelado y la predicción de series temporales.

Por ejemplo, Velásquez, Dyner y Souza (2004) demuestran que los sistemas adaptativos de inferencia neurodifusa (ANFIS, por su sigla en inglés) son una generalización de los modelos de regresión con transición suave (STR, por su sigla en inglés), y proponen una aproximación metodológica novedosa basada en criterios estadísticos para su especificación. De este modo, la aproximación desarrollada se usa para el modelado del precio de la electricidad en el mercado brasileño.

Velásquez y González (2006) desarrollan un nuevo modelo híbrido que combina redes neuronales artificiales y modelos heterocedásticos y lo aplican al modelado del índice de tipo de cambio real colombiano, usando redes neuronales artificiales. Zapata, Velásquez y Smith (2004) evalúan las bondades de los ANFIS para predecir caudales mensuales en el sector eléctrico colombiano y las comparan con otros modelos tradicionales, que encuentran que los ANFIS son más precisos en la mayoría de los casos.

No obstante, la mayor parte de los trabajos citados en la literatura más relevante están relacionados con el uso de modelos de redes neuronales artificiales —véase, por ejemplo, la revisión general presentada por Zhang, Patuwo y Hu (1998), o los ejemplos de aplicaciones específicas que son presentados

por Heravi, Osborn y Birchenhall (2004); Swanson y White (1997a y 1997b); Chatfield (1998); Darbellay y Slama (2000); Kuan y Liu (1995); entre muchos otros—.

Más aún, se han desarrollado esfuerzos importantes por establecer marcos metodológicos para especificar modelos de redes neuronales basados en criterios estadísticos —véase, por ejemplo, a Anders y Korn (1999) y Fukumizu (2003)—. De esta forma, existe muy poca experiencia en el modelado y la predicción de series temporales usando sistemas de inferencia difusa, particularmente para los ANFIS (Jang, 1993); sólo recientemente se ha planteado la relación de los ANFIS con otras técnicas econométricas (Velásquez et al., 2004), que abren la posibilidad de establecer un proceso sistemático de especificación fundamentado en pruebas estadísticas, para desarrollar así un marco de trabajo formal y riguroso.

Por otro lado, se ha observado que muchas series temporales presentan una volatilidad que puede ser cambiante en el tiempo y que, además, puede agruparse en *clusters*. Este comportamiento ha sido investigado principalmente en el campo financiero y sugiere la existencia de una estructura determinística en la volatilidad. Para explicar estos comportamientos, Engle (1982) desarrolló los modelos de heterocedasticidad condicional autorregresiva (ARCH, por su sigla en inglés), los cuales permiten representar la varianza de las series como una función de las perturbaciones aleatorias pasadas. Otros modelos más complejos han surgido como generalizaciones de la propuesta de Engle (1982), como los modelos GARCH (Bollers-

lev, 1986), IGARCH (Nelson, 1991), entre muchos otros.

Tradicionalmente, los modelos ARCH, y sus generalizaciones, se han usado para modelar la volatilidad de las series temporales, mientras la media esperada se representa con la aproximación de Box y Jenkins (1970). Así, los modelos obtenidos son lineales en media y no lineales en varianza. No obstante, es natural pensar que los modelos de Box y Jenkins (1970) utilizados en esa aproximación podrían ser reemplazados por otras técnicas no lineales para obtener una representación no lineal en la media y en la varianza.

Aunque esta idea fue esbozada de forma general por Tong (1990), en la literatura más relevante se han presentado pocas aproximaciones: Chan y McAleer (2003) presentan un modelo de transición suave combinado con un modelo GARCH para los errores en presencia de observaciones extremas; entre tanto, Chen y Mike (2006) proponen la combinación de un modelo de transición brusca (tipo *threshold*), que modela la no linealidad en la media, con un modelo GARCH, que representa la estructura de la volatilidad, y reseñan varios casos de aplicación en que el modelo híbrido postulado brinda una mejor representación, para los casos analizados, en relación con otros modelos tradicionales.

De acuerdo con lo anterior, resulta natural extender la propuesta de Velásquez et al. (2004), donde se considera un ANFIS homocedástico (varianza constante en el tiempo), que permite que la varianza de los errores siga un proceso ARCH. Igualmente, este mismo enfoque puede verse como una

extensión del trabajo de Chan, Marinova y McAleer (2005), al sustituir el modelo STR por un modelo adaptativo de inferencia neurodifusa.

Por ende, el objetivo de este trabajo es formular un marco metodológico para seleccionar modelos ANFIS-ARCH basados en criterios estadísticos. Como caso de aplicación se caracteriza la serie de los cambios relativos en los precios de cierre de las acciones de IBM, la cual ha sido comúnmente utilizada por diferentes investigadores como un *benchmark* para evaluar las bondades de nuevos modelos propuestos.

El resto de este artículo está organizado como sigue: en la sección 1 se desarrolla el marco metodológico para la especificación de los modelos ANFIS-ARCH, en la sección 2 se describen los datos usados para el caso de aplicación, en la sección 3 se realiza el modelado de la serie estudiada y, finalmente, se concluye.

## 1. El modelo ANFIS-ARCH: su estimación y especificación

En esta sección se describe la estructura matemática de los modelos ANFIS (sección 1.1), ARCH (sección 1.2) y ANFIS-ARCH, así como la estrategia de especificación para el modelo propuesto en este artículo (sección 1.3).

### 1.1 Sistemas adaptativos de inferencia neurodifusa

Un ANFIS (Jang, 1993) es un tipo de red neuronal artificial que emula el proceso de

inferencia realizado en los sistemas difusos. Desde un punto de vista estadístico, es una técnica no paramétrica de regresión no lineal con la capacidad de aproximar cualquier función definida en un dominio compacto, de tal forma que se le considera un aproximador universal de funciones.

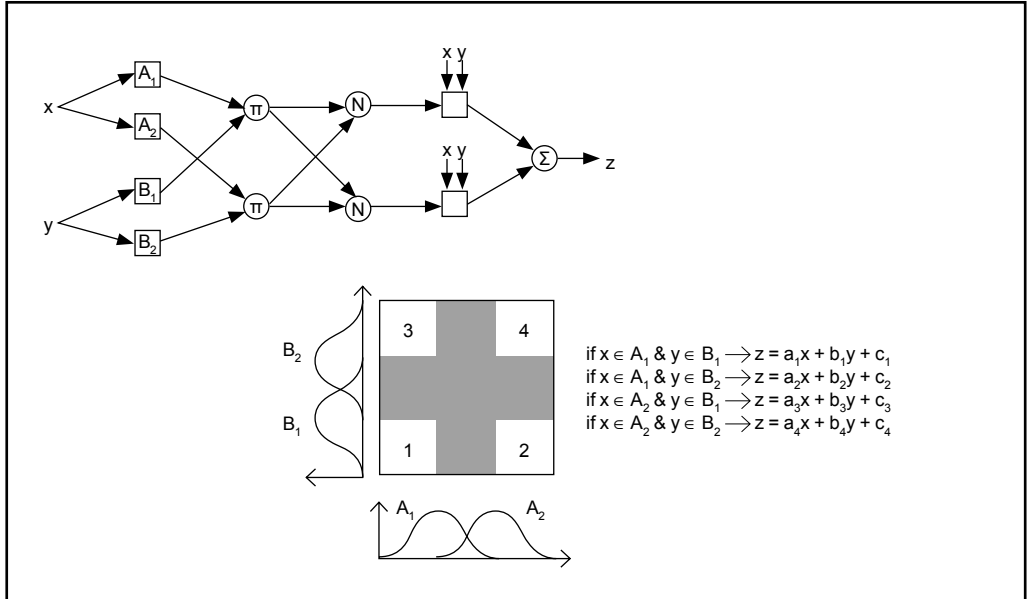
La idea fundamental de los ANFIS consiste en dividir en dos o más regiones cada una de las variables de entrada o regresores. De esta forma, el dominio del problema resulta dividido en un conjunto de regiones que surge de la intersección de las regiones en que ha sido dividido cada regresor. La arquitectura típica de los ANFIS se presenta en el Gráfico 1. Donde  $x$  y  $y$  son las variables independientes del problema y  $z$  es la salida del sistema o variable dependiente. El dominio de  $x$  se ha dividido en las regiones  $A_1$  y  $A_2$ , y el dominio de  $y$ , en las regiones  $B_1$  y  $B_2$ . De esta forma, el dominio del problema, esto es, el plano  $X \times Y$ , queda dividido en las regiones 1, 2, 3 y 4 (véase la parte inferior del Gráfico 1), que resultan de la intersección de  $A_1$ ,  $A_2$ ,  $B_1$  y  $B_2$ .

A cada una de las cuatro regiones en que se divide el dominio del problema se le asigna un modelo lineal de la forma  $z = ax + by + c$ , el cual representa la relación de dependencia entre  $z$  y  $(x, y)$ . El modelo final es representado como un conjunto de reglas, donde el antecedente permite determinar a qué región pertenece el punto que se desea evaluar  $(x, y)$ ; mientras el consecuente corresponde al modelo lineal (Gráfico 1).

La principal diferencia con los modelos de partición o basados en regiones es que los

Gráfico 1

**Arquitectura típica de los ANFIS**



Fuente: Jang (1993).

ANFIS permiten que un punto pertenezca simultáneamente a dos regiones (de ahí, que admite la ambigüedad) y, consecuentemente, el valor de  $z$  se calcula teniendo en cuenta los dos modelos lineales. Esta zona de ambigüedad es notada en color gris en el Gráfico 1.

En este caso, la pertenencia de un punto a una región se modela a través de un conjunto difuso, el cual se caracteriza por su función de pertenencia,  $\mu_A(x)$ , que indica el grado en que  $x$  está asociado al conjunto difuso  $A$ , donde un grado de 1 indica la certeza absoluta de que  $x$  pertenece al conjunto  $A$ ; mientras que 0 indica la certeza absoluta de que  $x$  no pertenece al conjunto  $A$ . Los valores entre 0 y 1 indican ambigüedad respecto a si el punto

pertenece o no. El proceso de inferencia realizado para calcular el valor de  $z$  dada una entrada  $(x, y)$  es el siguiente (Velásquez et al., 2004):

- Calcular las funciones de pertenencia:  $\mu_{A1}(x)$ ,  $\mu_{A2}(x)$ ,  $\mu_{B1}(y)$  y  $\mu_{B2}(y)$ .
- Estimar para cada regla el producto  $w_j = \mu_{A_j}(x) \cdot \mu_{B_j}(y)$ .
- Establecer el porcentaje en que cada regla aporta a la solución final:  $w_j = w_j / \sum_{i=1}^N w_i$ .
- Finalmente, calcular el resultado del sistema como:  $\sum_i w_i \cdot f_i$ . Donde  $f_i$  es usualmente una combinación lineal de las variables del consecuente.

Jang (1993) propone un algoritmo de aprendizaje híbrido para la estimación de los parámetros de los ANFIS, el cual combina el uso de técnicas de optimización basadas en gradientes con la estimación basada en mínimos cuadrados. El proceso de especificación tradicionalmente usado (Jang, 1994) se basa en árboles de clasificación y regresión para encontrar la especificación inicial del modelo, que consiste en definir las variables de entrada, la variable de salida y la cantidad de conjuntos difusos (regiones) en que se dividirá el dominio de cada variable independiente.

Velásquez et al. (2004) demuestran que el modelo anterior puede generalizarse para el problema específico del modelado y la predicción de series temporales. Consideran el caso en que el antecedente tiene una sola variable de decisión  $z$ , cuyo dominio está cubierto por dos conjuntos borrosos y los consecuentes pueden considerar un conjunto de regresores  $x$ . De esta forma, el sistema de inferencia descrito está conformado por dos reglas:

$$\begin{aligned}
 \text{Si } z \in A_1 &\rightarrow y_t = \tilde{a}_0^{(1)} + \sum_{k=1}^U \tilde{a}_k^{(1)} \cdot x_t^{(k)} \\
 \text{Si } z \in A_2 &\rightarrow y_t = \tilde{a}_0^{(2)} + \sum_{k=1}^U \tilde{a}_k^{(2)} \cdot x_t^{(k)}
 \end{aligned} \tag{1}$$

Donde los regresores  $x_t^{(k)}$  (para  $k=1, \dots, U$ ) y la variable de transición  $z$  pueden corresponder a rezagos de la serie analizada o a variables exógenas que puedan explicar su comportamiento. Un ejemplo del modelo obtenido es presentado en la ecuación (9). Los dos conjuntos borrosos que cubren el dominio de  $z$  pueden especificarse a través de las funciones  $S(\cdot)$  y  $Z(\cdot)$ :

$$\begin{aligned}
 S &= \frac{1}{1 + \exp(-\gamma_1(x - c_1))} \\
 Z &= 1 - \frac{1}{1 + \exp(-\gamma_2(x - c_2))}
 \end{aligned} \tag{2}$$

Donde  $c_i$  es un parámetro que define el punto de inflexión de la función de transición, y  $\gamma_i$  define su pendiente. Tal como es mostrado por Velásquez et al. (2004), la formulación propuesta en (2) es equivalente a un modelo LSTR o *logistic smooth transition regression* cuando se hace  $c_1=c_2$  y  $\gamma_1=\gamma_2$ , ya que para la ponderación asignada a cada regla se obtiene que  $w_1=1-w_2$ , de acuerdo con el algoritmo de cálculo previamente descrito.

La función de transición puede especificarse de igual modo por medio de una función exponencial dada por:

$$\begin{aligned}
 S &= \exp(-\gamma_1(x - c_1)^2) \\
 Z &= 1 - \exp(-\gamma_2(x - c_2)^2)
 \end{aligned} \tag{3}$$

Un problema de esta función es que converge a un valor constante cuando  $\gamma_1$  y  $\gamma_2$  tienden a cero. Igualmente, cuando se imponen las mismas restricciones anteriores, el modelo es equivalente a un ESTR o *exponential smooth transition regression*.

La ventaja que presenta esta formulación para los ANFIS en el contexto del análisis y la predicción de series de tiempo es que permite especificar una función de transición equivalente más compleja que los modelos STR tradicionales. Así, para los modelos LSTR, la función de transición entre regímenes está definida como la función  $S(\cdot)$  presentada en (2). Dicha función posee un

punto de inflexión en  $(c_1, 0,5)$  y es simétrica alrededor de  $c_1$ .

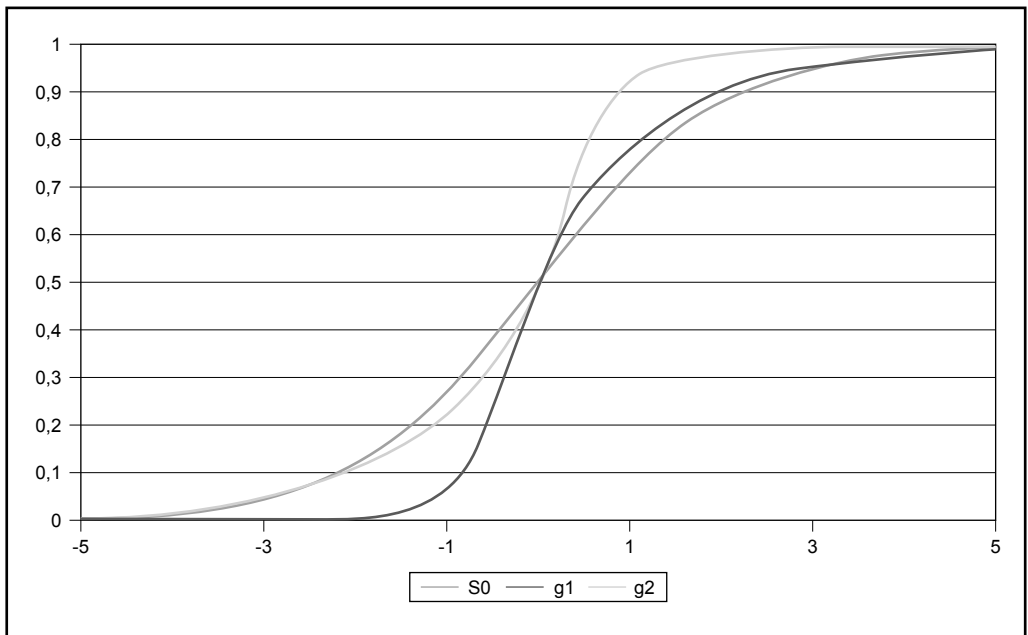
En el Gráfico 2 se presenta la función de transferencia  $S(.)$  con  $\gamma_1=1$  y  $c_1=0$ , y la función equivalente de transición para los ANFIS para  $c_1=c_2=0$ ; la curva g1 representa el caso  $\gamma_1=3$  y  $\gamma_2=1$ , mientras que la curva g2 representa  $\gamma_1=1$  y  $\gamma_2=3$ . Se observa que la función de transición equivalente para los ANFIS no es simétrica, lo que podría, en términos generales, ser una ventaja sobre los modelos LSTR tradicionales, puesto que permite una mejor adaptación del modelo.

### 1.2 Modelo de heterocedasticidad condicional autorregresiva

Los modelos ARCH (Engle, 1982) y sus generalizaciones, como GARCH (Bollerslev, 1986) e IGARCH (Nelson, 1991), han sido una forma efectiva de modelar la no linealidad en varianza, o heterocedasticidad, o varianza cambiante en el tiempo, que poseen algunas series temporales. Este comportamiento se visualiza como una varianza cambiante en el tiempo y como grupos de volatilidad similar.

Gráfico 2

#### Funciones de transferencia logística y equivalentes para los sistemas adaptativos de inferencia neurodifusa



Fuente: elaboración propia.



Su uso tradicional ha estado relacionado con los modelos ARIMA, pero recientemente se han propuesto diversos modelos híbridos, donde se usan técnicas no lineales para aproximar la media de la serie, y los modelos ARCH para modelar la varianza (Chang, 2006; Chan et al., 2005). Específicamente, Engle (1982) propone representar una serie temporal  $y_t$  como:

$$y_t = f(I_{t-1}) + h(I_{t-1}) \cdot \varepsilon_t \quad (4)$$

Donde  $I_t$  es la información disponible hasta el instante  $t$ . Así, el valor esperado de  $y_t$  es  $f(\cdot)$  mientras  $h(\cdot)$  representa la desviación estándar de  $y_t$ .  $\varepsilon_t$  es una variable aleatoria normal estándar. Si  $f(\cdot)$  es lineal, se dice que la serie es lineal en la media. Si  $h(\cdot)$  es constante, se dice que la serie es homocedástica, o de varianza constante; de lo contrario, se dice que es heterocedástica.

De esta forma, la heterocedasticidad puede definirse como la presencia de una estructura causal en la varianza que la hace dependiente del tiempo. Una forma de modelar la heterocedasticidad es representar la varianza de los residuales como una combinación lineal de los residuales anteriores, así:

$$\sigma_t^2 = \omega_0 + \sum_{i=1}^I \omega_i \sigma_{t-i}^2 \quad (5)$$

La cual define un modelo ARCH de parámetros  $w_p, i=0:I$ . En este caso,  $e_t = h(I_{t-1}) \cdot \varepsilon_t$ .

### 1.3 Modelo ANFIS-ARCH

Un modelo ANFIS-ARCH consiste en una extensión del modelo desarrollado por Ve-

lázquez et al. (2004), presentado en la ecuación (1), que se obtiene al considerar que los residuales  $e_t$  siguen un proceso ARCH. De esta forma, un modelo ANFIS-ARCH se define como:

$$\hat{y}_t = f(\Psi, x) + e_t, \quad e_t = \sigma_t^2 \varepsilon_t \quad (6)$$

Donde  $f(\cdot)$  es definida en (4),  $\Psi$  es el vector de parámetros del modelo; y  $\sigma_t^2$  es la varianza de los residuales, la cual sigue un proceso ARCH definido en (5). Los parámetros del modelo  $\Psi$  son obtenidos maximizando el logaritmo de la función de verosimilitud de los residuales:

$$L = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log \sigma_t^2 - \frac{1}{2} \sum_{t=1}^T \frac{e_t^2}{\sigma_t^2} \quad (7)$$

El uso de (7) es la base del procedimiento estadístico de inferencia para especificar el modelo, el cual se describe más adelante. La estimación de parámetros del modelo ANFIS-ARCH propuesto es un problema difícil de abordar debido a la presencia de múltiples puntos de mínima local causados por su no linealidad; además, su componente ARCH incorpora una dificultad adicional, ya que la función de costo, en este caso la función de verosimilitud de los residuales, se hace muy sensible a pequeñas variaciones en los parámetros de la componente ARCH.

Pensando en ello, se propone una técnica de gradiente descendente adaptativo denominada *iRprop+* (Igel y Hüsken, 2000), que al tener tamaños de paso individuales para cada componente del vector de parámetros, pro-

porciona un método robusto y preciso para su estimación.

El procedimiento de especificación del modelo es descrito en el siguiente algoritmo, el cual se basa principalmente en la propuesta de Terasvirta, Tjostheim y Granger (1994) para los modelos STR:

- Paso 1: se selecciona un modelo autorregresivo de orden  $P$  que aproxime la serie, usando un criterio de información, tal como el de Akaike, Hannan-Quinn o Schwarz.
- Paso 2: se realiza una prueba de contraste para determinar si un modelo STR (Terasvirta et al., 1994) puede representar mejor la dinámica de la serie frente a un modelo lineal. Si se acepta la hipótesis nula, que favorece la alternativa lineal, se detiene el proceso; de lo contrario, se sigue con el paso 3.
- Paso 3: se determina el tipo de función de transición más adecuada (logística o exponencial), así como la variable de transición, siguiendo el proceso de especificación propio de los modelos STR; la prueba aparece descrita en Terasvirta et al. (1994).
- Paso 4: se considera el modelo ANFIS con dos reglas presentado en (1). Cada regla es un modelo autorregresivo cuyo orden  $P$  fue determinado en el paso 1. Igualmente se considera que la varianza de los errores es constante.
- Paso 5: Se realiza la estimación del vector de parámetros óptimos del modelo  $\Psi$  maximizando (7).
- Paso 6: se realizan pruebas de diagnóstico para demostrar si los residuales  $e_t$ , para  $t=1, \dots, T$ , son normales e incorrelacionados. Cuando se cumplen estas dos condiciones, se continúa el proceso de especificación; de lo contrario, se reinicia. La homocedasticidad de los residuales es determinada a partir del contraste de Engle (1982). En caso de que los residuales resulten homocedásticos, se termina el proceso de especificación.
- Paso 7: a partir de los resultados del contraste de Engle, se postula un modelo ARCH para representar la varianza.
- Paso 8: el proceso de eliminación de regresores irrelevantes está basado en la prueba del radio de verosimilitud (Vuong, 1989), es sistemático y puede ser descrito de la siguiente forma, primero para las variables de los consecuentes y luego para las componentes ARCH:
  - Sea  $\log L$  la función especificada en (7); si se tienen dos modelos  $\Theta$  y  $\Theta^*$  tales que el segundo es un modelo restringido del primero, el radio de verosimilitud  $LR$  se define como:
 
$$LR = 2(\log L\Theta - \log L\Theta^*) \quad (8)$$
  - Para cada variable del modelo, tanto de los consecuentes como del modelo

ARCH, se estiman modelos restringidos donde dicha variable es eliminada y se calcula LR para cada uno.

- Se selecciona la entrada correspondiente al modelo con el menor valor crítico (LR).
- Si el valor crítico es mayor del 5%, se elimina la entrada y se vuelve al primer paso, considerando el modelo restringido que eliminaba la entrada seleccionada como el nuevo modelo. De lo contrario se termina el proceso.
- Paso 9: se realizan las pruebas de diagnóstico descritas en el paso 6 a los residuales normalizados. Si se rechaza el contraste de Engle, se retorna al paso 7; en caso contrario, el modelo ya está especificado.

### 1.4 Contribuciones

Aunque el modelo propuesto surge de la combinación de dos modelos ya estudiados en la literatura, se debe resaltar que:

- Los ANFIS constituyen una técnica propia de la inteligencia computacional desarrollada como un modelo general de regresión que desconoce aspectos fundamentales del modelado de series temporales. En este trabajo se amplía la propuesta presentada por Velásquez et al. (2004), al contextualizar a los ANFIS como un modelo de series de tiempo, cuyo proceso de especificación se basa en criterios formales de corte estadístico. En la propuesta original, la varianza de los residuales se

supone constante en el tiempo, por lo que el modelo es no lineal en la media y lineal en la varianza.

- Al considerar que los residuales siguen un proceso ARCH, se extiende el modelo original para que considere que la varianza es cambiante en el tiempo. Esto amplía las posibilidades de modelado, ya que permitiría detectar otras características de los datos que quedarían ocultas al considerar otras aproximaciones. Combinar los ANFIS con un modelo ARCH en una sola ecuación es algo trivial; sin embargo, la dificultad emerge cuando se intenta aplicar esta aproximación a una serie real. En primer lugar, dada una especificación cualquiera, es necesario estimar sus parámetros a partir de los datos, lo que implica el desarrollo de algoritmos de optimización. En segundo lugar, es necesario desarrollar y probar experimentalmente una estrategia que permita especificar adecuadamente el modelo propuesto. Una de las principales contribuciones de este trabajo es que se formula, precisamente, un marco metodológico para especificar la nueva clase de modelos propuestos.

### 2. Descripción de los datos del caso de aplicación

En la sección anterior se desarrolló el modelo propuesto en este artículo y se formuló una estrategia para su especificación. En esta sección se valida el modelo propuesto contra otros especificados en la literatura para un caso real. La serie escogida corresponde a los precios diarios de IBM, que fue utiliza-

da originalmente por Box y Jenkins (1970) para demostrar el proceso de especificación de sus modelos.

Como ya se indicó, nuestro objetivo es establecer las bondades del modelo propuesto respecto a otros establecidos en la literatura para un caso de aplicación bien conocido y validado, más que ahondar en la serie misma y los desarrollos y avances que se han dado en la última década.

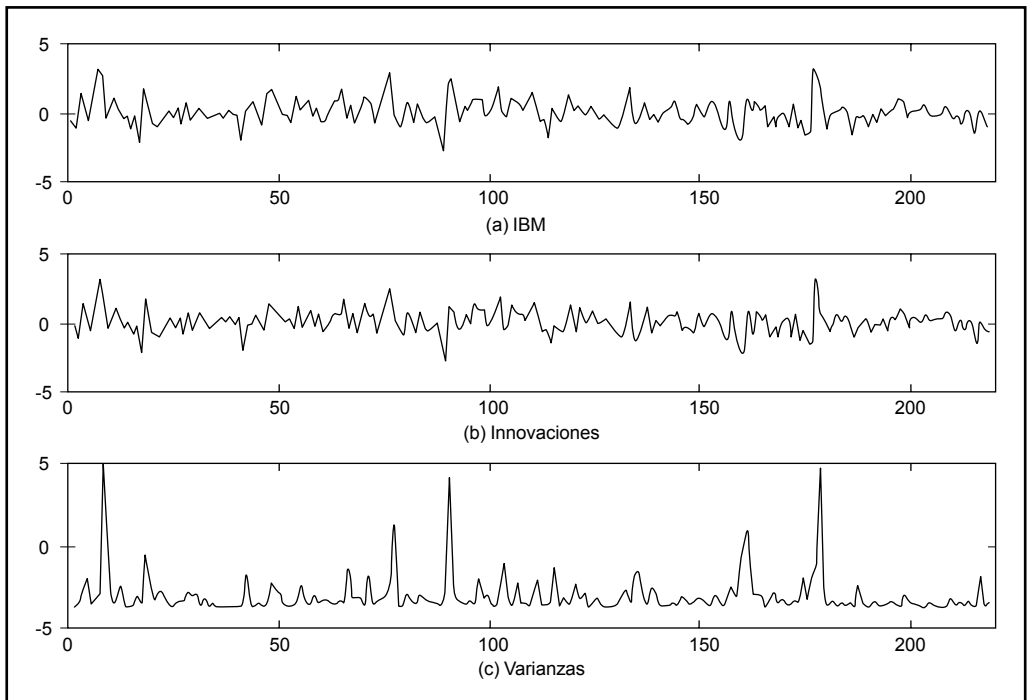
Desde entonces, esta serie se ha analizado en la literatura como un *benchmark* y se ha demostrado, entre otros aspectos, sus características de no linealidad y heterocedasticidad

(Tong, 1990; Tsay, 2002). Muchos enfoques se han usado para representar la dinámica de esta serie; sin embargo, el modelo de inteligencia computacional más citado en la literatura para modelarla corresponde a las redes neuronales artificiales, el cual fue propuesto por White (1993) y sirvió como precedente para el uso de las redes neuronales en el análisis de otras series financieras.

La serie de precios de IBM utilizada en este trabajo corresponde a la misma utilizada por Tong (1990), quien demuestra que su comportamiento es no lineal; sus datos van desde el 18 de mayo de 1961 hasta 30 de marzo de 1962, con un total de 118 observaciones. El Gráfico 3

Gráfico 3

**(a) IBM, (b) innovaciones del modelo ANFIS-ARCH y (c) varianzas del modelo ANFIS-ARCH**



Fuente: elaboración propia.

presenta la serie temporal. La media de la serie es 0,07 y su varianza es de 0,9239.

### 3. Caracterización de la serie

El Gráfico 4 resume las propiedades estadísticas de la serie. La función de autocorrelación simple y la función de autocorrelación parcial (FACP) se presentan en los gráficos 4a y 4b, respectivamente, e indican que la serie se podría aproximar con un modelo  $AR(P)$ , de orden 2. En el espectro de potencias (Grá-

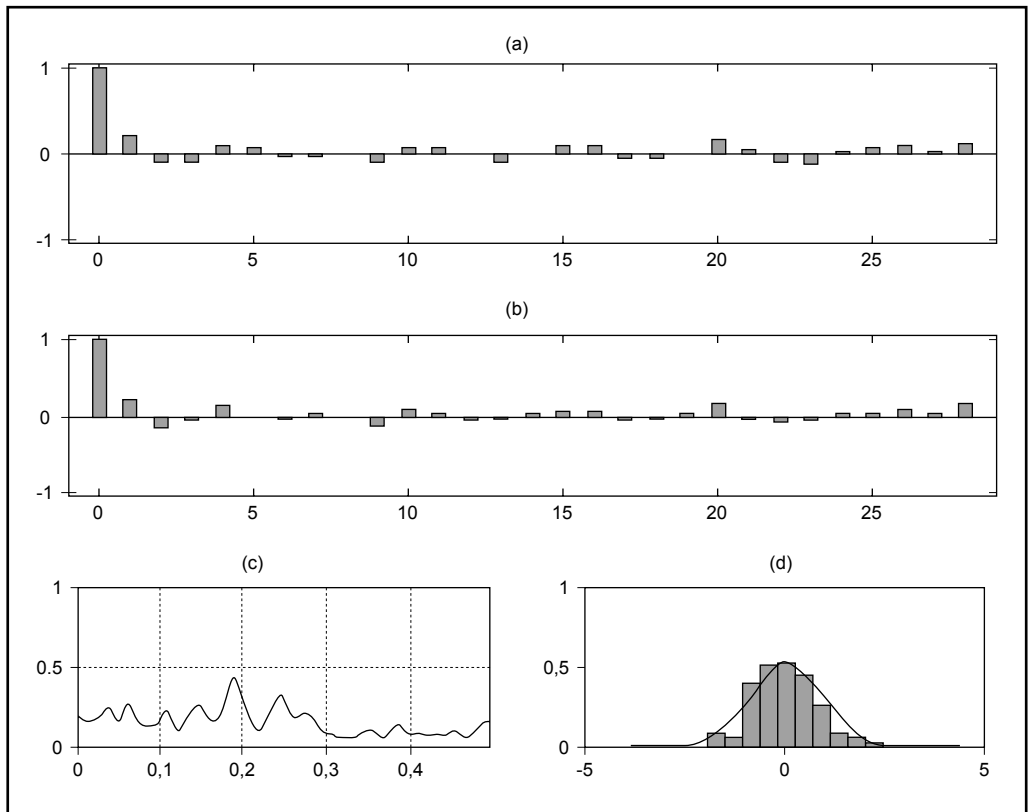
fico 4c) no se observa una concentración de energía de frecuencias muy altas.

El histograma (Gráfico 4d), además, muestra que los datos de la serie siguen una distribución aproximadamente normal. A continuación se presentan los pasos de la metodología propuesta para el modelo desarrollado.

*Paso 1.* En el Cuadro 1 se reportan los valores de los diferentes criterios de información estimados para modelos autorregresivos de orden

Gráfico 4

**Propiedades estadísticas de la serie: (a) autocorrelograma simple, (b) autocorrelograma parcial, (c) espectro y (d) histograma**



Fuente: elaboración propia.

$P=1, \dots, 12$ . CAK es el criterio de información de Akaike; CHQ es el criterio de información de Hannan-Quinn; CSC es el criterio de información de Schwarz. Los criterios de información utilizados permiten seleccionar el orden tentativo del modelo autorregresivo inicial; en este caso, el error está compuesto por una componente de ajuste a la muestra de datos usada para la calibración, la cual disminuye a medida que aumenta el orden del modelo, más una componente de variabilidad del error, que aumenta a medida que se incrementan los parámetros. Cada uno de los criterios de información utilizados es una aproximación a dicho error total, de tal forma que se prefiere aquel modelo que minimice el valor del criterio de información utilizado. Así el orden  $P$  del modelo  $AR(p)$  óptimo para las reglas de los consecuentes es 4.

Cuadro 1

**Valor de los criterios de información para el modelo  $AR(p)$**

p	CAK	CHQ	CSC
1	2,737	2,756	2,783
2	2,722	2,747	2,784
3	2,729	2,760	2,806
4	2,721	2,759	2,814
5	2,730	2,774	2,839
6	2,738	2,789	2,863
7	2,747	2,803	2,887
8	2,756	2,818	2,911
9	2,747	2,816	2,918
10	2,749	2,825	2,936
11	2,759	2,840	2,960
12	2,764	2,852	2,981

CAK: criterio de información de Akaike; CHQ: criterio de información de Hannan-Quinn; CSC: criterio de información de Schwarz.

Fuente: elaboración propia.

*Paso 2.* En el Cuadro 2 se presentan los estadísticos obtenidos al aplicar las pruebas clásicas de no linealidad de Kennan, Ramsey (prueba RESET), Tsay y White, las cuales fueron descritas de forma detallada por Tong (1990). Como hipótesis nula se supone que un modelo  $AR(p)$  captura adecuadamente la dinámica de los datos, y que los residuales obtenidos son ruido blanco. Como hipótesis alternativa se plantea que existe una componente no lineal remanente en los residuales que el modelo  $AR(p)$  no fue capaz de capturar.

Las cantidades entre paréntesis corresponden al valor crítico o  $p$ -valor para el estadístico calculado. Si el  $p$ -valor es mayor o igual a 0,06, se considera que no existen indicios de que el modelo  $AR(p)$  sea inadecuado, por lo que la dinámica de la serie estudiada es lineal y, por consiguiente, no existe mérito para usar un modelo no lineal, por ejemplo, ANFIS. En el caso contrario ( $p$ -valor  $< 0,06$ ), hay suficientes indicios para rechazar la hipótesis nula y, por ende, se justifica el uso de un modelo no lineal.

En el Cuadro 3 se reportan los resultados del contraste de no linealidad desarrollado para los modelos STR. En primer lugar, se parte de la hipótesis nula de que un modelo  $AR(p)$  es adecuado para modelar la serie frente a la hipótesis alternativa de que existe una componente no lineal remanente en los datos, la cual podría capturarse de manera adecuada por un modelo STR.

Cuadro 2

**Pruebas de no linealidad**

Rez	Kennan	RESET	Tsay	White
1	15,12 (0,000)	10,06 (0,000)	15,26 (0,000)	15,67 (0,039)
2	10,44 (0,001)	6,19 (0,002)	6,22 (0,000)	13,53 (0,115)
3	10,47 (0,001)	5,15 (0,007)	2,97 (0,008)	14,17 (0,084)
4	14,58 (0,000)	9,53 (0,000)	2,44 (0,009)	15,16 (0,051)
5	14,38 (0,000)	9,45 (0,000)	1,85 (0,031)	15,25 (0,049)
6	11,30 (0,001)	7,90 (0,000)	1,52 (0,075)	12,41 (0,202)
7	8,93 (0,003)	8,18 (0,000)	1,51 (0,058)	15,05 (0,054)
8	4,80 (0,030)	8,96 (0,000)	1,22 (0,203)	12,61 (0,182)
9	6,27 (0,013)	5,99 (0,003)	1,34 (0,097)	9,62 (0,813)
10	3,58 (0,060)	5,10 (0,007)	1,22 (0,177)	-0,01 (1,000)
11	4,03 (0,046)	5,18 (0,006)	1,16 (0,238)	9,70 (0,784)
12	3,07 (0,081)	4,13 (0,017)	1,27 (0,120)	11,14 (0,381)

Fuente: elaboración propia.

Cuadro 3

**Resultados de la prueba STR**

Var	H0		H1		H2		H3		Tipo
Lag 1	2,84	(-0,001)	0,92	(-0,453)	2,01	(-0,095)	5,51	(0,00)	L
Lag 2	1,19	(-0,29)							
Lag 3	0,54	(-0,889)							
Lag 4	1,00	(-0,45)							

Fuente: elaboración propia.

Para el caso analizado, sólo se presentan indicios de no linealidad para  $z_t = y_{t-1}$ . Cuando hay evidencias de no linealidad, se realizan los contrastes notados como H1, H2 y H3, los cuales permiten escoger el tipo de función de transición (logística o exponencial). El análisis indica que la función logística (*L*) equivalente para la transición entre regiones es la más adecuada.

*Pasos 3, 4 y 5.* En el Cuadro 4 se resumen las propiedades estadísticas de los residuales obtenidos al estimar un modelo ANFIS

homocedástico, esto es, con varianza constante;  $\log L$  es el logaritmo de la función de verosimilitud de los residuales, que es igual a -278. El contraste de normalidad de Jarque-Bera permite rechazar la hipótesis nula en que se asume que los residuales siguen una distribución normal. Esto es consistente con el resultado de la prueba ARCH, que indica que los residuales siguen un proceso ARCH(1). La prueba de Ljung-Box señala que las autocorrelaciones de los residuales no son significativamente diferentes de cero. Este resultado indica que es necesario incor-

Cuadro 4

**Pruebas estadísticas para el modelo ANFIS homocedástico**

Tamaño de la muestra	218			
Sumatoria	-0,6978			
Media	-3,20E-03			
Mediana	-0,03827			
Desviación estándar	1,011			
Asimetría	0,0837			
Curtosis	4			
Exceso de curtosis	0,9001			
Mínimo	-3,381			
Máximo	3,652			
Test de normalidad de Jarque-Bera	7.038 (-0,03)			
Test de normalidad de Lin-Mudholkar	-0.37(-0,645)			
LLF	-2,78E+02			
Rez	Box-Pierce	Ljung-Box	McLeod-Lee	ARCH
1	0,01 (0,906)	0,01 (0,905)	3,45 (0,063)	3,45 (0,063)
2	0,08 (0,962)	0,08 (0,962)	3,52 (0,172)	1,72 (0,424)
3	0,13 (0,988)	0,14 (0,987)	3,58 (0,310)	1,16 (0,764)

Fuente: elaboración propia.

porar una componente ARCH de orden uno al modelo ANFIS.

*Paso 6.* El proceso de eliminación de regresores irrelevantes se llevó a cabo según el algoritmo descrito, y se obtuvieron como relevantes en el modelo los rezagos del 1 al 3 de la serie. El modelo final elegido corresponde a un modelo ANFIS-ARCH de la forma:

$$y_t = \frac{S(y_{t-1})}{S(y_{t-1}) + Z(y_{t-1})} \{ \phi_0^{(1)} + \phi_1^{(1)}y_{t-1} + \phi_2^{(1)}y_{t-2} + \phi_3^{(1)}y_{t-3} \} + \frac{S(y_{t-1})}{S(y_{t-1}) + Z(y_{t-1})} \{ \phi_0^{(2)} + \phi_1^{(2)}y_{t-1} + \phi_2^{(2)}y_{t-2} + \phi_3^{(2)}y_{t-3} \} + a_t^2 \tag{9}$$

Donde  $a_t = h_t \varepsilon_t$  con  $h_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2$ . De esta forma,  $h_t^2$  sigue un proceso ARCH(1). Los parámetros de este modelo son presentados en el Cuadro 5.

Cuadro 5

**Parámetros del modelo ANFIS-ARCH**

Parámetro	Valor	STD	Est. T
c(1)	0,96	3,46	0,278
γ(1)	2,73	1,74	1,567
c(2)	0,15	0,01	28,905
γ(2)	285,43	1,00	284,619
Φ0(1)	-0,37	3,11	-0,121
Φ1(1)	-0,43	4,50	-0,095
Φ2(1)	0,03	1,74	0,019
Φ3(1)	-0,02	1,42	-0,015
Φ0(2)	0,70	4,21	0,167
Φ1(2)	0,15	4,15	0,036
Φ2(2)	-0,28	2,15	-0,129
Φ3(2)	-0,17	1,73	-0,100
α0	0,52	1,43	0,364
α1	0,35	2,49	0,139

Fuente: elaboración propia.



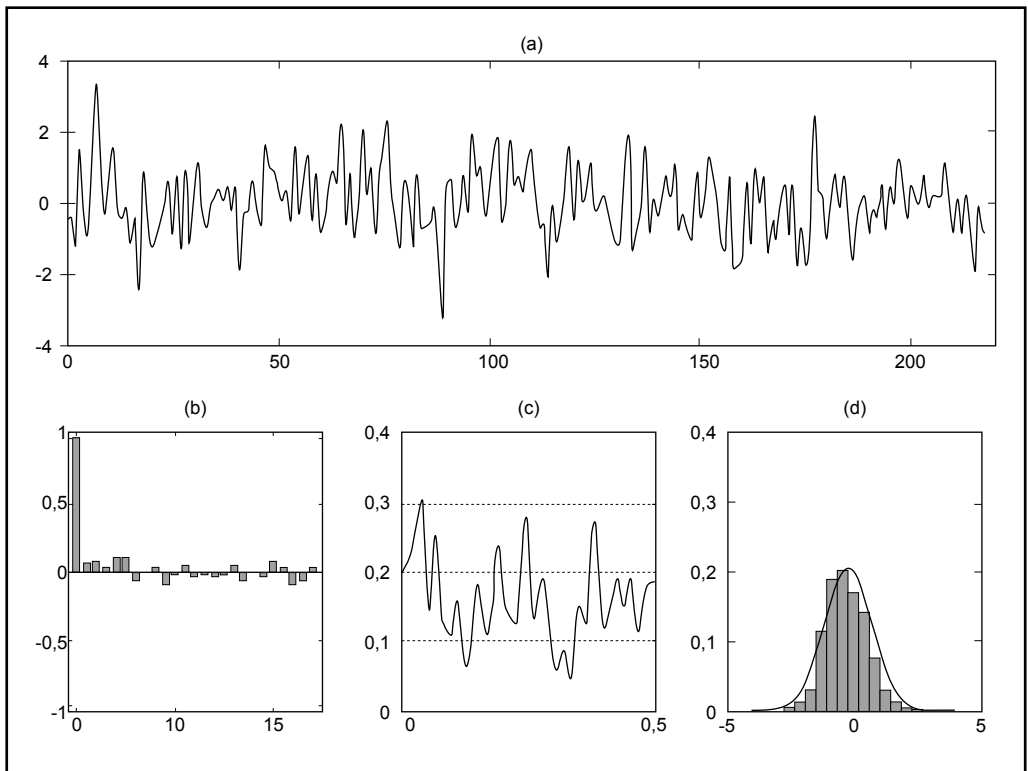
El Gráfico 5 resume las principales propiedades estadísticas para los residuales normalizados obtenidos al estimar (9). La función de autocorrelación simple (FAC) y el espectro indican que los residuales parecen estar incorrelacionados; además, el histograma indica que son aproximadamente normales. Las pruebas estadísticas para este modelo, reportadas en el Cuadro 6, confirman que los residuales siguen una distribución normal. Los contrastes aplicados indican que son homocedásticos y que no existen correlaciones remanentes.

En el Gráfico 3a se presentó la serie analizada. En los gráficos 3b y 3c se presentan las innovaciones estimadas y la varianza en función del tiempo. Se observa que el modelo captura la relación entre las varianzas altas y las innovaciones de mayor magnitud.

El Gráfico 6 muestra la función de transición entre regímenes como función del tiempo y la variable de transición (Gráfico 5b). El Gráfico 5a muestra fluctuaciones rápidas de un régimen a otro, lo que evidencia la no linealidad de la serie.

Gráfico 5

**Propiedades estadísticas de los residuales del modelo ANFIS-ARCH:  
(a) residuales, (b) FAC, (c) espectro y (d) histograma**



Fuente: elaboración propia.

Cuadro 6

**Pruebas estadísticas para el modelo ANFIS-ARCH**

Tamaño de la muestra	218			
Sumatoria	-0,5431			
Media	-2,49E-03			
Mediana	-0,02791			
Desviación estándar	1,0030			
Asimetría	0,1145			
Curtosis	3,40600			
Exceso de curtosis	0,406			
Mínimo	-3,231			
Máximo	3,426			
Test de normalidad de Jarque-Bera	1,747 (0,418)			
Test de normalidad de Lin-Mudholkar	-0,5799 (0,719)			
LLF	-2,73E+02			
Rez	Box-Pierce	Ljung-Box	McLeod-Lee	ARCH
1	0,65 (0,418)	0,66 (0,415)	0,18 (0,672)	0,18 (0,675)
2	1,77 (0,413)	1,80 (0,407)	0,35 (0,840)	0,17 (0,920)
3	1,88 (0,597)	1,92 (0,590)	0,70 (0,873)	0,23 (0,972)

Fuente: elaboración propia.

El Gráfico 7 presenta el pronóstico determinístico para la serie;  $y_{T+1}$  se calculó utilizando los datos históricos. Para calcular  $y_{T+2}$  se utilizaron los datos reales hasta  $y_T$  y el pronóstico para  $y_{T+1}$ . Los restantes valores se calcularon de forma similar. En todos los casos se asume que los residuales son cero. Se observa cómo la serie pronosticada converge rápidamente a un punto de equilibrio donde las entradas y la salida del modelo son iguales numéricamente.

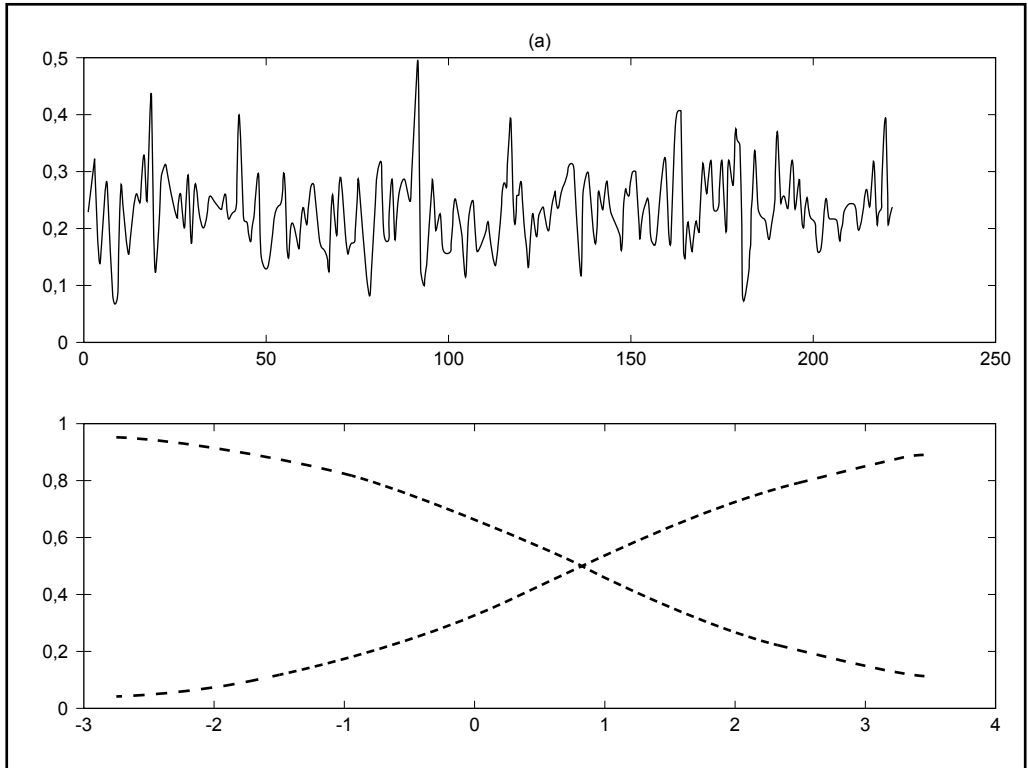
El Gráfico 8 muestra la convergencia a largo plazo del modelo; ella se obtiene al realizar el pronóstico determinístico utilizando diferentes puntos de arranque obtenidos de los datos históricos. En general, los mode-

los no lineales pueden presentar procesos de convergencia a uno o más puntos estables, procesos cíclicos y procesos divergentes. Es común que, dependiendo del punto de inicio, un mismo modelo presente distintos comportamientos. Para el modelo obtenido se presenta un proceso de convergencia a un único punto de equilibrio, y la velocidad con que el modelo converge a este es aproximadamente constante e independiente del tipo de fenómeno extremo.

Las propiedades estadísticas de los datos son reportadas en el Cuadro 7 y contrastadas con las propiedades para una serie sintética simulada de 10.000 observaciones. Este cuadro permite analizar qué tan bien son reproduci-

Gráfico 6

**Función de transición para el modelo ANFIS-ARCH:  
(a) en el tiempo y (b) como función de la variable de transición**



Fuente: elaboración propia.

das las propiedades estadísticas de los datos. En general, puede concluirse que el modelo reproduce adecuadamente las propiedades de la serie original.

Cuadro 7

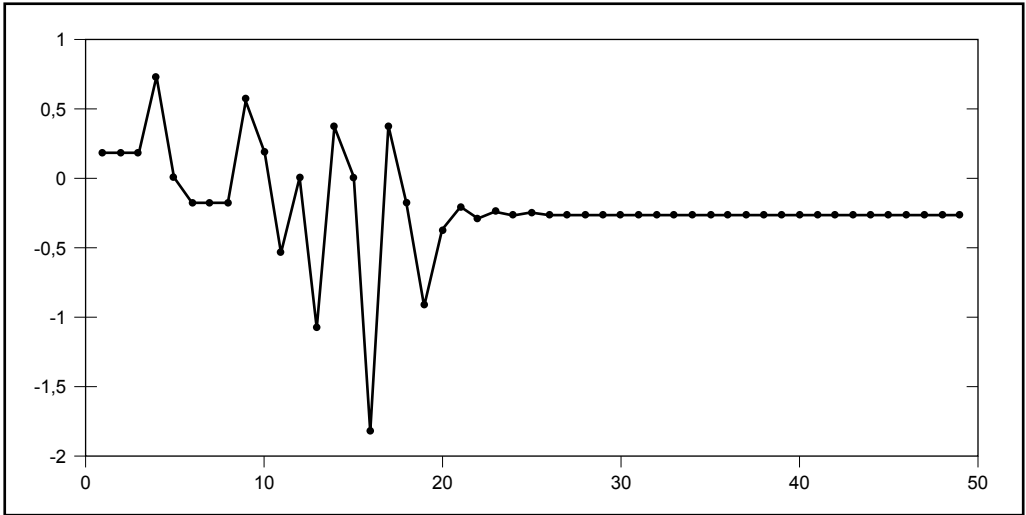
**Momentos de la serie real y el modelo**

Modelo	Media	Varianza	Curto-sis	Asimetría
Real	0,0722	0,9239	4,2643	0,4667
ANFIS-ARCH	0,071	0,9274	4,6762	0,4355

Fuente: elaboración propia.

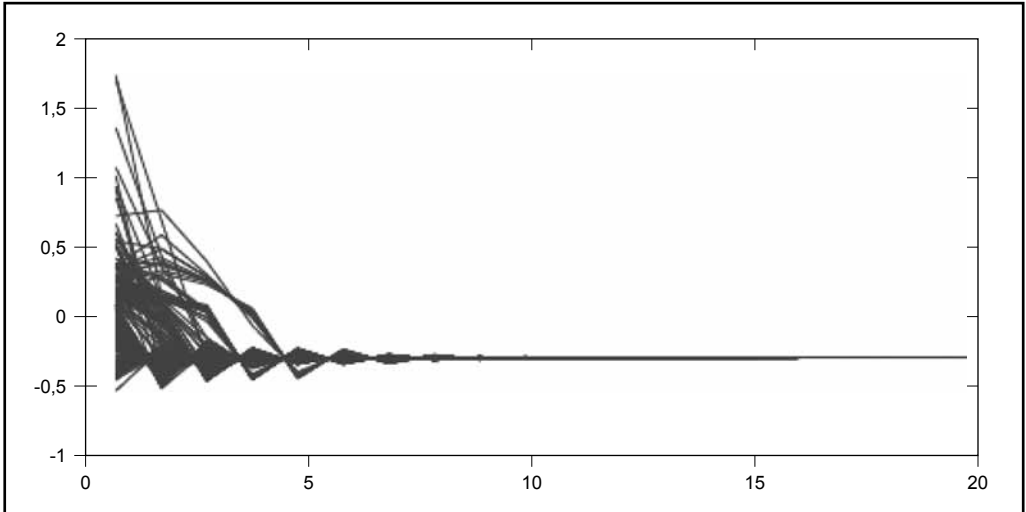
El Cuadro 8 presenta los estadísticos de diferentes modelos con los que se desea comparar el modelo ANFIS-ARCH propuesto. De acuerdo con la relación que establecen Velásquez et al. (2004) entre los modelos ANFIS y STR, los modelos STR-ARCH y ARX-ARCH, se puede afirmar que dichos modelos poseen la misma estructura de los modelos ANFIS-ARCH, a efectos de comparar el ajuste de la serie.

Gráfico 7  
**Pronóstico determinístico**



Fuente: elaboración propia.

Gráfico 8  
**Convergencia de largo plazo**



Fuente: elaboración propia.

Cuadro 8

**Estadísticos de diferentes modelos**

	ANFIS-ARCH	STR-ARCH	ARX-ARCH
T	218	218	218
R <sup>2</sup>	0,1808	0,1414	0,06361
AIC	2,656	2,672	2,674
Asimetría	0,6997	0,04787	-0,07948
Exceso de curtosis	0,2639	0,2672	0,5189
LJB	0,6731 (0,714)	0,5937 (0,743)	2,377 (0,305)
LM <sub>ARCH</sub>	1,08 (0,359)	1,16 (0,327)	0,98 (0,405)
LLF	-273,5103081	-279,2202421	-285,4672007

Fuente: elaboración propia.

El contraste entre cada uno de los modelos presentados se describe a continuación:

- El primer modelo presentado en el cuadro de comparaciones es el STR-ARCH. El estadístico R<sup>2</sup> es mayor para el modelo ANFIS-ARCH, mientras el criterio de información de Akaike es mayor para el modelo STR-ARCH, y el log *L* también es mayor para los modelos ANFIS-ARCH. Los tres estadísticos confirman que el modelo ANFIS-ARCH reproduce mejor la dinámica de la serie. Finalmente, los estadísticos de Ljung-Box y LM<sub>ARCH</sub> indican que los residuales son incorrelacionados y homocedásticos para ambos modelos. Se realizó, además, la prueba del radio de verosimilitud para determinar cuál modelo representaba mejor la dinámica de la serie. Se consideró el modelo ANFIS-ARCH como el modelo completo, y el modelo STR-ARCH como el modelo restringido. El estadístico de la prueba fue 11,420, y su valor crítico 0,022. Se puede concluir así que el

modelo ANFIS-ARCH captura mejor la dinámica de la serie.

- El otro modelo usado para comparar es el ARX-ARCH contra el cual los estadísticos también indicaron que el modelo ANFIS-ARCH capturaba mejor la dinámica de la serie. El criterio de información de Akaike fue mayor para el modelo ARX-ARCH, y tanto el estadístico R<sup>2</sup> como el log *L* fueron mayores para el modelo ANFIS-ARCH. Además, la prueba del radio de verosimilitud, cuyo estadístico fue 23,914 y su valor crítico 0,008, indicó que efectivamente los modelos ANFIS-ARCH se comportaban mejor que los modelos ARX-ARCH.

## Conclusiones

Tradicionalmente, en la literatura de series temporales, el problema de la no linealidad se ha tratado considerando ya sea la no linealidad en la media (redes neuronales y neurodifusas, modelos STR, etc.) o la no linealidad

en la varianza (ARCH y sus derivados). No obstante, la dinámica de una serie, usualmente financiera o económica, puede presentar ambos tipos de no linealidad al tiempo simultánea, de tal forma que dichas características no se capturan y se representan de manera adecuada cuando se aplican aproximaciones tradicionales. Esto puede llevar a conclusiones erróneas sobre las propiedades de los datos analizados y a pronósticos más imprecisos en términos del valor futuro esperado y de su volatilidad.

En este artículo se ha propuesto un nuevo modelo para el modelado y la predicción de series temporales que presentan simultáneamente no linealidades en la media y la varianza. Igualmente, se ha desarrollado un proceso para su especificación basado en criterios estadísticos. Así, se han alcanzado aportes tanto conceptuales como metodológicos a las áreas de la inteligencia computacional y a la econometría no lineal, ya que el nuevo modelo propuesto integra conceptos y metodologías de ambas áreas.

El caso analizado ejemplifica tanto el modelo ANFIS-GARCH propuesto como la metodología de especificación desarrollada. Al comparar este modelo con modelos alternativos más parsimoniosos (con menos parámetros) y más tradicionales, tal como el modelo AR-GARCH, se demuestra que la aproximación aquí presentada es, en términos estadísticos, significativamente diferente y que permite modelar mejor las propiedades de la serie analizada.

## Lista de referencias

- Anders, U. y Korn, O. (1999). Model selection in neural networks. *Neural Networks*, 12, 309-323.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31, 307-327.
- Box, G. y Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. San Francisco: Holden-Day, Inc.
- Chan, F., Marinova, D. y McAleer, M. (2005). Modelling thresholds and volatility in US ecological patents. *Environmental Modelling & Software*, 20, 1369-1378.
- Chan, F. y McAleer, M. (2003). Estimating smooth transition autoregressive models with garch errors in the presence of extreme observations and outliers. *Applied Financial Economics*, 13 (8), 581-592.
- Chang, B. (2006). Applying nonlinear generalized autoregressive conditional heteroscedasticity to compensate ANFIS outputs tuned by adaptive support vector regression. *Fuzzy Sets and Systems*, 17, 1832-1850.
- Chatfield, J. F. C. (1998). Time series forecasting with neural networks: A comparative study using the airline data. *Journal of Applied Statistics*, 47, 231-250.
- Chen, C. y Mike, S. (2006). On a threshold heteroscedastic model. *International Journal of Forecasting*, 22(1), 73-89.

- Darbellay, G. y Slama, M. (2000). Forecasting the short-term demand for electricity: Do neural networks stand a better chance? *International Journal of Forecasting*, 16, 71- 83.
- Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations. *Econometrica*, 50, 987-1007.
- Fukumizu, K. (2003). Likelihood ratio of unidentified models and multilayer neural networks. *Annals of Statistics*, 31 (3), 833-851.
- Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge: Cambridge University Press.
- Heravi, S., Osborn, D. y Birchenhall, C. (2004). Linear versus neural network forecasts for european industrial production series. *International Journal of Forecasting*, 20, 435- 446.
- Igel, C. y Hüsken, M. (2000). Improving the Rprop learning algorithm. En H. Bothe y R. Rojas (Eds.), *Proceedings of the Second International ICSC Symposium on Neural Computation (NC 2000)* (pp. 115-121). Genova: ICSC Academic Press.
- Jang, J. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 23, 665-684.
- Jang, J. (1994). *Structure determination in fuzzy modeling: a fuzzy cart approach*. Documento presentado en el IEEE International Conference on Fuzzy Systems.
- Kuan, C. y Liu, T. (1995). Forecasting exchange rates using feedforward and recurrent neural networks. *Journal of Applied Econometrics*, 10, 347-364.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59 (2), 347-370.
- Swanson, N. y White, H. (1997a). Forecasting economic time series using adaptive versus non-adaptive and linear versus non-linear econometric models. *International Journal of Forecasting*, 13, 439-461.
- (1997b). A model selection approach to real time macroeconomic forecasting using linear models and artificial neural networks. *Review of Economics and Statistics*, 39, 540-550.
- Terasvirta, T., Tjostheim, D. y Granger, C. (1994). Aspects of modelling nonlinear time series. *Handbook of Econometrics*, 4 (48), 2917-2957.
- Tong, H. (1990). *Non-linear time series a dynamical system approach*. Oxford: Clarendon Press.
- Tsay, R. (2002). *Analysis of financial time series: Financial econometrics*. Philadelphia: John Wiley & Sons.
- Velásquez, J. D., Dyner, I. y Souza, R. C. (2004). Modelación de serie temporales usando ANFIS. *Inteligencia Artificial*, 8 (23), 47-64.
- Velásquez, J. D. y González, L. M. (2006). Modelado del índice de tipo de cambio real colombiano usando redes neuronales artificiales. *Cuadernos de Administración*, 19 (32), 319-336.

Vuong, Q. H. (1989), Likelihood ratio test for model selection and non-nested hypotheses. *Econometrica*, 57, 307-333.

White, H. (1993). Economic prediction using neural networks: The case of IBM daily stock returns. *IEEE International Conference on Neural Networks in Finance and Investing*, 2, 315-328.

Zapata, E., Velásquez, J. y Smith, R. (2004). Modelamiento de series de caudal usando ANFIS. *Revista Avances en Recursos Hidráulicos*, 11, 79-90.

Zhang, G., Patuwo, B. y Hu, M. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14, 35-62.