

ÓRDENES DE FLUJO, TASA DE INTERÉS Y TASA DE CAMBIO NOMINAL: UN EJEMPLO DE REDES NEURONALES PARA COLOMBIA¹ 2005

Octavio José Salcedo Parra*

Marco Aguilera Prado**

Resumen

El presente documento muestra un modelo de Redes Neuronales Artificiales, RNA, para el pronóstico de la tasa de cambio nominal en Colombia, que incluye órdenes de flujo y el diferencial de las tasas de interés como variables de entrada al modelo, hallando relaciones no lineales entre las mismas. Adicionalmente se plantean algunas conclusiones metodológicas a partir de los tradicionales esquemas econométricos para el manejo de las series de tiempo.

Palabras clave

Tipo de cambio nominal, pronóstico, modelo no-lineal, redes neuronales artificiales.

Clasificación JEL: C53 - E37

Abstract

The current document shows a model of Artificial Neuronal Networks ANN for the forecast of the rate of nominal exchange rate in Colombia; what includes flow orders and the differential of the interest rates like variables of entrance to the model. We did find nonlinear relations between the same ones. Additionally some method conclusions from the traditional econometrics ways for the handling of the time series were considered.

Keywords: Nominal exchange rate, forecast, non-linear models, artificial neuronal networks.

1. Este artículo se recibió el 01/06/05 y fue aprobado el 08/08/06.

* Ingeniero de Sistemas, Magíster en Economía, octavionetworking@gmail.com

** Economista UAO maquilera75@yahoo.com

Introducción

El análisis para la modelación y el pronóstico del tipo de cambio de una divisa con respecto a otras ha tenido varias etapas relevantes. En este sentido los primeros modelos desarrollados tenían como punto de partida los equilibrios de flujos entre los países; estos planteaban que el tipo de cambio podía determinarse mediante funciones de oferta y demanda de divisas correspondiente a cada uno de los países. Entre los modelos que hacen parte de esta línea de desarrollo están los de Meade (1951) y Mundell - Fleming (1963), mencionados por Manrique (2001).

Posteriormente, con el dinamismo que adquirieron los mercados financieros y la caída del Bretton Woods aparecieron modelos en los que el comercio internacional, los flujos internacionales de comercio, el precio de las exportaciones, los bienes domésticos y los portafolios internacionales de activos se tomaron como variables relevantes para la determinación de la tasa de cambio. Este enfoque se denominó equilibrio de stocks o mercado de activos, el cual se encuentra dividido en modelos monetarios (precios flexibles, precios rígidos) y modelos de equilibrio de cartera, que pretenden explicar las fluctuaciones del tipo de cambio mediante un proceso análogo al que se ven sometidos los precios de otros activos financieros.

En esta misma corriente se originaron los modelos seminales y obligados en

la literatura del tipo de cambio: Rogoff y Meese (1983, 1988), los cuales examinan los modelos monetarios y de balance de portafolios para pronósticos en horizontes cortos (1 a 12 meses).

No obstante, los trabajos de Meese y Rogoff (1983a, 1983b, 1988) reciben fuertes críticas en Evans y Lyons (1999), quienes consideran que los modelos de tasa de cambio nominal han estado en crisis desde la aparición de dicho modelo, al hacer todos ellos una aproximación que empíricamente falla. Adicionalmente Frankel y Rose (1995) escriben: "El análisis de Meese y Rogoff en horizontes cortos nunca ha sido convincentemente explicado. Se continuó trayendo un efecto pesimista en el campo de la modelación de la tasa de cambio empíricamente en particular y en las finanzas internacionales en general".²

En ese orden de ideas aparecen modelos que tratan de explicar las causas de los pobres resultados de los modelos tradicionalmente macro de TCN. Flood y Rose (1995), por ejemplo, se encaminan a que los determinantes más críticos de la volatilidad de la tasa de cambio no son macroeconómicos. Blanchard (1979), Dornbusch (1982), Meese (1986) y Evans (1986), entre otros, se mueven en la corriente que propone se deben tomar variables exógenas para explicar la volatilidad de la TCN a partir de burbujas financieras. En contravía, Flood y Hodrick (1990)

2. Evans y Lyons (1999)

concluyen que las alternativas de burbuja son poco convincentes. De otra parte, Domínguez (1986), Frankel y Froot (1987) y Hau (1998), entre otros, explican los componentes exógenos a partir de irracionalidades como errores en las variables expectativas.³

Con esta motivación surge el trabajo de Evans y Lyons (1999), el cual hace un aporte significativo: la inclusión de una variable microeconómica para explicar y pronosticar las variaciones en la TCN. La idea central de este modelo es considerar elementos propios de la microestructura de mercados financieros sobre la valoración de activos, para incluir variables explicativas (microeconómicas) no consideradas en los trabajos hasta aquí realizados. Fundamentalmente se hace un desarrollo sobre las órdenes de flujo, entendidas como la medida de la presión a la venta o a la compra del activo financiero, tomándolas como variable explicativa del modelo de TCN. Este modelo explica las variaciones de la tasa de cambio en una función lineal del diferencial de la tasa de interés y el orden de flujo del yen vs. US dólar y marco alemán vs. US dólar para el periodo comprendido de mayo 1 a agosto 31 de 1996.

Las conclusiones de este documento apuntan en cuatro direcciones: la robustez, la causalidad, cambios en el portafolio y la capacidad de pronóstico fuera de la muestra.

Otros trabajos se movieron en una dirección alterna, no de especificación, sino de la forma funcional de la tasa de cambio, en el sentido que hicieron apreciaciones sobre la no linealidad de la tasa de cambio. Se destacan Baillie y McMahon (1989), quienes precisaron que la tasa de cambio no es linealmente previsible; Hsieh (1989) observó que los cambios de la tasa de cambio pueden ser no linealmente dependientes. A pesar de ello Meese y Rose (1991) encontraron que el pobre poder explicativo de los modelos no podía ser atribuido a las no linealidades, pese a que previamente Meese y Rose (1990) utilizaron un estimador no-paramétrico para manejar las no linealidades, sin que proporcionara ninguna mejora significativa en la explicación mensual de la tasa de cambio.⁴

En este contexto, recientemente Evans y Lyons (2004) persisten en la microestructura como determinante de la tasa de cambio, mediante un modelo de equilibrio dinámico. Este trabajo concluye que aún existen vacíos en la explicación de la TCN a partir de fundamentos macro. Los problemas en la explicación surgen por la volatilidad de las variables macro, porque hay variaciones en la TCN cuando no hay movimientos en las variables macro, o viceversa.

Adicionalmente y en contraposición a los tradicionales modelos ARIMA, de caminata aleatoria, martingale, cade-

3. *Ibid*

4. Gradojevic y Yang, 2000.

nas de Markov, utilizados, entre otros, por Baillie y Bollerslev (1989), Engel y Hamilton (1990), Verkooijen (1996) y Plasmans, Verkooijen y Daniels (1998) utilizaron modelos macroeconómicos y las redes neuronales artificiales (RNA), para probar si la relación subyacente es no lineal; a la vez que Hu (1999) demostró (con datos diarios y semanales) que las redes neuronales artificiales –RNA- son un método más robusto de pronóstico que el modelo de caminata aleatoria. Previamente Kuan y Liu (1995) habían usado una técnica hacia el futuro y redes neuronales artificiales recurrentes para producir pronósticos de media condicional de la tasa de cambio.⁵

Contrariamente, en Colombia la literatura de TCN no es muy amplia, en contraste con la de tasa real. Cárdenas (1997) menciona los trabajos de Wiesner (1978), Urrutia (1981), López (1987) y Steiner (1987) sobre los procesos de minidevaluaciones en el sistema de crawling peg colombiano. El trabajo de Cárdenas (1997) analiza los determinantes de la tasa de cambio nominal en el periodo 1985-1986, en los dos regímenes, crawling peg y el sistema de bandas,⁶ mediante el modelo monetario simple, el modelo monetario con precios fijos y el modelo de balance de portafolio. La conclusión sobre

los determinantes apunta a que el modelo monetario con precios flexibles se ajusta en buena manera al comportamiento de la TCN colombiana. Según el modelo la TCN responde a cambios en la oferta monetaria y las tasas de interés. Más recientemente Oliveros y Huertas (2002) presentan un documento que considera “la estimación de los desequilibrios nominales y reales del tipo de cambio en Colombia” incluyendo dos componentes: uno de tendencia estocástica y uno asociado al ciclo. Sintéticamente la descomposición apunta a que la tendencia se asocia al equilibrio y el componente estocástico con los desequilibrios. Para el TCN en la serie presentada (información trimestral de marzo 1980 a marzo 2002) la paridad en el poder de compra no se valida en el sentido estricto. Adicionalmente se construye un modelo de *Behavioral Equilibrium Exchange Rate* –BEER–, el cual arroja conclusiones similares.

En el horizonte metodológico Jalil y Melo (1999) presentan una manera de testear la no-linealidad de la inflación y el crecimiento de los medios de pago, particularmente comparando un AR(p) con un modelo autorregresivo de transición suave STAR.⁷ Son relevantes las conclusiones que van de la mano con la identificación de relaciones no-

5. *Ibid*

6. Colombia entre 1967 y octubre de 1991 mantuvo un tipo de cambio fijo ajustado con minidevaluaciones (*crawling peg*); a partir de noviembre de 1991 adoptó un sistema de bandas hasta septiembre de 1999, cuando se adopta el tipo de cambio de flotación libre.

7. Particularmente este test es aplicable a modelos no-lineales con formas funcionales definidas y no es factible su aplicación a RNA, ya que estas no tienen formas funcionales definidas.

lineales entre inflación y el crecimiento de los medios de pago y con la alta capacidad de pronóstico del modelo no lineal en el largo plazo.

En este contexto, fundamentado en la microestructura y las RNA, y teniendo en cuenta las críticas hacia los modelos enteramente macro de Evans y Lyons (1999) y lo expresado por Barkoulas, Baum, Caglayan y Chakraborty (2001) sobre los procesos tipo martingale y la memoria de largo plazo, el presente documento examina el comportamiento de un modelo de pronóstico de la TCN del peso respecto al US dólar, al introducir una variable de la microestructura del mercado (orden de flujo) en un sistema de observaciones diarias con variables macroeconómicas (tasas de interés), con una modelación no lineal de RNA para series de tiempo diarias de un (1) año, buscando medir el poder predictivo y el comportamiento de la modelación mediante la raíz del error cuadrático medio (RMSE) y el error medio porcentual absoluto (MAPE).⁸

Para ello el artículo se dividió en cuatro secciones: la primera sección cubre la introducción. La sección dos revisa lo relacionado con las RNA y la microestructura de los mercados financieros. La sección tres describe los datos, analiza y presenta los resultados. La sección cuatro corresponde a conclusiones y recomendaciones.

Redes Neuronales Artificiales – RNA- y la micro-estructura de los mercados financieros

• Redes Neuronales Artificiales – RNA-

Una Red Neuronal Artificial (RNA) es un intento de poder realizar una simulación computacional del comportamiento de partes del cerebro humano mediante la réplica en pequeña escala de los patrones que éste desempeña en la formación de resultados a partir de los sucesos percibidos. Más formalmente las RNA son modelos estadísticos no lineales utilizados principalmente para la clasificación y predicción de datos y variables inspirados por el sistema nervioso biológico, que tratan de simular el proceso de aprendizaje humano con el convencimiento de que habiendo sido creados por el proceso de selección natural el mecanismo debe ser eficiente. (Montenegro, 2001).

Se denominan neuronales porque están basadas en el funcionamiento de una neurona biológica cuando procesa información. Ésta tiene tres partes bien diferenciadas: la soma, las dendritas y el axón. La neurona biológica toma los datos por las dendritas, que están conectadas a otras neuronas o centros de información nerviosa; éstas a su vez son el árbol receptor de la red; son como fibras nerviosas que cargan de señales eléctricas el cuerpo

8. Una amplia revisión de estos modelos econométricos puede encontrarse en Mills (1999) y Brooks (2002). De hecho, Mills (1999) presenta los modelos aquí mencionados para datos de la libra esterlina en diferentes períodos.

de la célula. La conexión, denominada sinapsis, es un proceso complicado que se realiza mediante los llamados neurotransmisores, siendo el punto de contacto entre un axón de una célula y una dendrita de otra célula. En la sinapsis los estímulos procedentes de las conexiones sinápticas son lo suficientemente grandes para que la neurona se active y envíe a través del axón una corriente eléctrica destinada a liberar neurotransmisores hacia otras neuronas. El soma, el cuerpo de la célula, realiza la suma de esas señales de entrada. En la Figura 1, se pueden detallar las partes.

La estructura de una neurona artificial es una emulación de una neurona biológica, de modo que se podría hacer el siguiente paralelo con la neurona biológica:

- Las entradas X_j representan los pulsos discretos que provienen de otras neuronas y que son absorbidas por las dendritas (UTP, 2000).

- Los pesos W_j son la intensidad de la sinapsis que conecta dos neuronas.
- θ , es el umbral, valor que la neurona debe sobrepasar para que se produzca el proceso biológico dentro de la célula al activarse (Analogía que puede apreciarse en la Figura 2).

De la cual se deduce que:

$$H_i(t) = \sum_{j=1}^n X_j W_j,$$

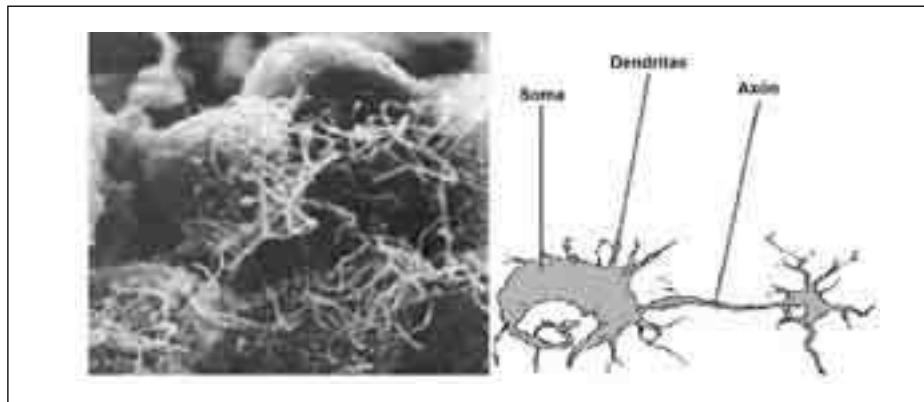
donde: $H_i(t)$: Potencial sináptico de la neurona i en el momento t ,

X_j : Entrada de datos procedentes de la fuente de información j ,

W_j : El peso sináptico asociado a la entrada X_j .

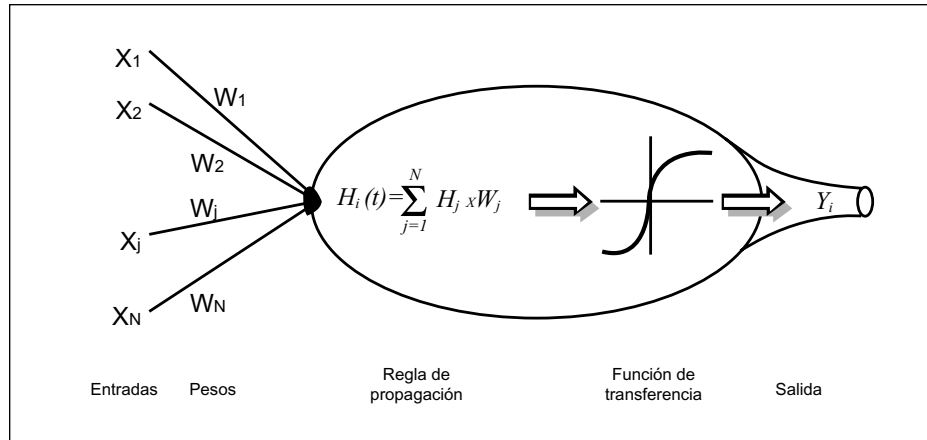
Como una de las características más relevantes detrás de las RNA es adquirir conocimiento mediante datos

Figura 1. Estructura de una neurona biológica



Fuente: Brian Resech Institut. UCLA

Figura 2. Estructura de una neurona artificial



Fuente: Brian Resch Institut. UCLA

pasados, se hace necesario el proceso de aprendizaje en el cual la RNA modifica sus pesos (W_j) en respuesta a una información de entrada (X_j). Los cambios que se producen durante el proceso de aprendizaje se reducen a la destrucción, modificación y creación de conexiones entre neuronas. Por tanto, el proceso de aprendizaje (entrenamiento) de la red artificial depende del tipo de red neural modelada y se catalogan como algoritmos de aprendizaje supervisados y no supervisados. Para las redes neurales utilizadas en este trabajo es necesario emplear algoritmos de aprendizaje supervisados, que consisten en modificación sucesiva de los parámetros de la red (pesos y umbrales) hasta que la salida de la red sea lo más próxima posible a la salida deseada o esperada para cada patrón de entrenamiento.

En este sentido se utiliza la arquitectura RNA *Multilayer Perceptron* -MLP- con la técnica de aprendizaje *back-propagation*, en la que su topología puede tener una capa de entrada con n neuronas. Para este caso de estudio se utilizó una capa con dos neuronas para el primer modelo (la orden de flujo -ODF- y la diferencia entre la DTF y la libor -DDL-, con los siguientes rezagos: $ODF_{t-1}, ODF_{t-2}, ODF_{t-9}, ODF_{t-13}, ODF_{t-36}, DDL_{t-1}, DDL_{t-5}, DDL_{t-6}, DDL_{t-8}, DDL_{t-9}, DDL_{t-17}$), y cinco neuronas ($TRM_{t-1}, TRM_{t-2}, TRM_{t-3}, TRM_{t-14}$ y TRM_{t-16}) para el segundo modelo,⁹ por lo menos una capa oculta (con cuatro y ocho neuronas para el primer y segundo modelo RNA, respectivamente), también con n neuronas, y una capa de salida con m neuronas. Para estos modelos se utilizó solamente una neurona de salida -TRM-. En síntesis, la RNA conside-

9. Se utilizó el δ - test, ver Anexo A.

rada posee una arquitectura con una capa de entrada, una capa oculta y una capa de salida y por tanto puede ser expresada como RNA(I, H, O).

Es de resaltar que el tamaño de la capa de entrada (ODF, DDL) corresponde a la caracterización propia del problema. De acuerdo con esta consideración, la generación de la capa de entrada es proporcional a las variables de mayor relevancia sugeridas en la solución del problema; aplicándole el δ -test se obtuvo el porcentaje de significancia de cada una de ellas para el ajuste y pronóstico de la tasa de cambio nominal (TRM). El tamaño de la capa oculta puede ser un factor multiplicativo de la capa de entrada. La capa de salida depende del patrón de resultados que deseamos obtener para la TRM.

La técnica de aprendizaje se basa en la regla delta generalizada, la cual se basa a su vez en el descenso usando la técnica del gradiente para minimizar una función objetivo $W(w_i)$ que es diferenciable con respecto a cada uno de los pesos en la red. En este sentido es realmente una técnica de optimización distribuida basada en el descenso, usando la dirección del gradiente, de tal forma que las expresiones de descenso que usan el gradiente aparecen como reglas discretas de actualización de los pesos que utilizan únicamente información localmente disponible a cada neurona (nodo) y una cantidad que es retropropagada desde las neuronas (nodos) en la penúltima capa. A continuación se deduce la regla que

se utilizó en el modelo de pronóstico de RNA:

La arquitectura funcional de la red es:

$$Y_k = g\left(\sum_{h=1}^H c_h g\left(\sum_{i=1}^I x_i w_{hi} + \theta_h\right) + d\right),$$

donde los c_h son los pesos que unen la neurona h de la capa oculta con la neurona de la capa de salida y d el umbral de la neurona en la capa de salida. Los pesos (w , c) y umbrales (θ , d) son ajustados durante el entrenamiento de la red. La fórmula para el ajuste de los pesos de la red depende de la posición de las capas que conectan los pesos, particularmente si los pesos están en la capa oculta o en la capa de salida.

Actualización de los pesos de la capa de salida: El error de una unidad de la red de salida MLP puede escribirse como:

$$\delta_k = (Y_k - \hat{Y}_k), \quad (1)$$

t se refiere al t -ésimo vector de entrenamiento y k a la k -ésima unidad de salida (Como en el modelo MLP sólo se utilizó una neurona en la capa de salida $k = 1$) (Méndez 2002).

La regla delta generalizada del algoritmo BP minimiza la SCE de todas las unidades de salida.

$$E_t = \frac{1}{2} \sum_{k=1}^K \delta_k^2. \quad (2a)$$

El término $1/2$ aparece sólo por conveniencia.

Para determinar el sentido en que se deben cambiar los pesos, se calcula el valor negativo del gradiente de Et, ∇E_t respecto a los pesos W_{kj} .

$$E_t = \frac{1}{2} \sum_{i=1}^n (Y_{it} - \hat{Y}_{it})^2 \quad (2b)$$

$$\frac{\partial E_t}{\partial W_{kj}^*} = -(Y_{it} - \hat{Y}_{it}) \frac{\partial F_k^*}{\partial (netu_{it}^*)} \frac{\partial (netu_{it}^*)}{\partial W_{kj}^*} \quad (3)$$

$$\frac{\partial (netu_{it}^*)}{\partial W_{kj}^*} = \left(\frac{\partial}{\partial W_{kj}^*} \sum_{l=1}^n W_{lj}^* S_{it} + d_{it}^* \right) = S_{it} \quad (4)$$

Con nodos de salida, salida de cada nodo de la capa oculta y salidas de los nodos de la capa de entrada como sigue:

$$netu_{it}^* = \sum_{j=1}^n W_{ij}^* S_{jt} + d_{it}^* \quad (5)$$

$$S_{it} = F_i^*(netu_{it}^*), \quad (6)$$

$$netu_{jt}^* = \sum_{k=1}^l W_{jk}^* X_{kt} + \theta_j^* \quad (7)$$

La magnitud del cambio de los pesos será proporcional al gradiente negativo de:

$$\frac{\partial E_t}{\partial W_{kj}^*} = -(Y_{it} - \hat{Y}_{it}) \frac{\partial F_k^*}{\partial (netu_{it}^*)} S_{it} \quad (8)$$

$$\Rightarrow \Delta_j W_{kj}^* = \eta (Y_{it} - \hat{Y}_{it}) \frac{\partial F_k^*}{\partial (netu_{it}^*)} S_{it} \quad (9)$$

Los pesos de la capa de salida se actualizan según

$$W_{ij}^*(iteración) = W_{ij}^*(iteración - 1) + \Delta_j W_{ij}^*(iteración - 1) \quad (10a)$$

Si la función de activación es la función sigmoide (logística):

$$F(x) = \frac{1}{1 + e^{-x}},$$

su derivada es: $F'(x) = F(x)(1 - F(x))$. (11)

De (9), (10a) y (11) la dinámica del vector de pesos de la capa de salida se calcula como:

$$W_{ij}^*(iteración) = W_{ij}^*(iteración - 1) + \eta (Y_{it} - \hat{Y}_{it}) \hat{Y}_{it} (1 - \hat{Y}_{it}) S_{it}, \quad (10b)$$

donde en (10b) η es la tasa de aprendizaje. Un valor alto de esta tasa de aprendizaje puede ayudar a la rápida convergencia de la red, pero se corre el peligro de

saltarse el mínimo global de la suma de los cuadrados de los errores (SCE).

De esta manera queda manifiesto que las modelaciones mediante RNA se refieren típicamente a modelos estadísticos no lineales del tipo:

$$x_t = f(x_{t-1}, x_{t-2}, \dots, x_{t-d}) + g(r_t, r_{t-1}, \dots, r_{t-k}),$$

donde f puede ser cualquier tipo de función relacionado con la salida x_t , dado un conjunto de variables de entrada. Adicionalmente cabe mencionar que según el proceso iterativo y la forma funcional que precede cada respuesta, los valores de los parámetros no interesan en el sentido cualitativo, tal como interesan en modelos lineales o regresivos.

Actualización de los pesos de la capa oculta: Para determinar la dinámica del vector de pesos de la capa oculta se calcula el gradiente negativo de la suma de los cuadrados de los errores con respecto a los pesos de la capa oculta, y el cambio en el vector de pesos se hace en la dirección del negativo de este gradiente (Méndez, 2002).

Partiendo de: (2b)

$$E_t = \frac{1}{2} \sum_{k=1}^n (Y_k - F_k^*(net_{a_k}^*))^2 \quad (2c)$$

$$= \frac{1}{2} \sum_{k=1}^n (Y_k - F_k^*(\sum W_{kj}^* S_j + d_k^*))^2 \quad (2d)$$

Como S_j depende de los pesos de las capas ocultas,

$$\frac{\partial E_t}{\partial W_{jk}^*} = -\frac{1}{2} \sum \frac{\partial}{\partial W_{jk}^*} (Y_k - \hat{Y}_k)^2 \quad (12a)$$

$$\frac{\partial E_t}{\partial W_{jk}^*} = -\sum (Y_k - \hat{Y}_k) \frac{\partial \hat{Y}_k}{\partial (net_{a_k}^*)} \frac{\partial (net_{a_k}^*)}{\partial S_j} \frac{\partial S_j}{\partial (net_{a_j}^*)} \frac{\partial (net_{a_j}^*)}{\partial W_{jk}^*} \quad (12b)$$

$$\frac{\partial E_t}{\partial W_{jk}^*} = -\sum (Y_k - \hat{Y}_k) F_k^{**}(net_{a_k}^*) W_{kj}^* F_j^{**}(net_{a_j}^*) X_j. \quad (12c)$$

Los pesos de la capa oculta se ajustan proporcionalmente al negativo de la ecuación (12c)

$$\Delta_j W_{jk}^* = \eta F_j^{**}(net_{a_j}^*) X_j \sum (Y_k - \hat{Y}_k) F_k^{**}(net_{a_k}^*) W_{kj}^* \quad (13)$$

De (13) el cambio en el vector de pesos de la capa oculta se calcula como:

$$W_{jk}^*(iter) = W_{jk}^*(iter-1) + \eta F_j^{**}(net_{a_j}^*) X_j \sum (Y_k - \hat{Y}_k) F_k^{**}(net_{a_k}^*) W_{kj}^* \quad (14)$$

La idea básica detrás de este cálculo de señales de error para los nodos ocultos es propagar los errores hacia atrás, basados en la discrepancia observada entre los valores de los nodos de salida y los esperados para un patrón de entrenamiento. Es de anotar que los pesos cambian cuando se presenta cada patrón de entrenamiento.

En este contexto y considerando la convergencia del algoritmo de acuerdo con las técnicas de gradiente descendiente, es conveniente avanzar por la superficie de error con incrementos pequeños de los pesos, ya que se tiene información local de superficie y no se sabe lo lejos o lo cerca que se está del punto mínimo, y si se toman valores grandes de incrementos es posible no evaluar valores pequeños que factiblemente sean una buena medida del valor pronóstico. En muchos entrenamientos la tasa de aprendizaje (alfa factor de decaimiento exponencial entre 0 y 1 que determina la contribución relativa del gradiente actual y los anteriores al cambio en los pesos), de peso debe ser pequeña y se le introduce un término de aceleración para asegurar que la red converja; por

ello, casi siempre la red debe hacer un número significativo de iteraciones.

El modelo base de la red neuronal artificial trabajada es:

$$Y_t = g\left(\sum_{k=1}^H c_k g\left(\sum_{i=1}^I x_{ki} w_{ki} + \theta_k\right) + d\right).$$

Para una salida de la capa oculta de la siguiente forma:

$$O_{th} = g\left(\sum_{i=1}^I w_{hi} x_{ti} + \theta_h\right),$$

donde:

- O_{th} Es la de salida de la neurona h de la capa oculta dado el vector de entrada en t ,
- X_{ti} Es la señal de entrada¹⁰ i que pertenece al vector de entrada (vector de entrada en el tiempo t),
- W_{hi} Son los pesos que conectan las neuronas entre capas,
- θ_h Es el umbral de la neurona h ubicada en la capa oculta,
- I Es la dimensión del vector de entrada,
- g Es la función de activación, función logística:

10. La selección del vector de entrada no es inmediata y, dada su importancia, debe utilizarse un método adecuado. Hong Pi y Carsten Peterson (1994) diseñaron el método de la prueba δ (δ -test), el cual establece las dependencias funcionales dadas en una secuencia de medidas. Este método se basa en el cálculo de las probabilidades condicionales y posterior derivación de las dependencias entre variables. La virtud del método para preprocesar datos en el contexto de redes neuronales de tipo *feedforward* está demostrado tal como el perceptrón multicapa -MPL-, utilizado para diseñar la RNA (Buitrago y Alcalá, 1998). Ver anexo A, para una explicación detallada.

$$g(x) = \frac{1}{1 + \exp^{-x}}$$

Esta función limita la señal de salida al intervalo $[0,1]$ porque $g(-\infty) = 0$ y $g(\infty) = 1$. Aplicando esta conceptualización y haciendo un paralelo con el modelo clásico de regresión lineal, el interés no radica en los parámetros del modelo o en el signo de éstos; lo que interesa realmente es que iterando con unos pesos y con una función de activación (se aplica la función sigmoide), seguramente se podrá ajustar el modelo y encontrar un buen pronóstico.

Debido a que las RNA no definen un modelo matemático determinado o una forma funcional establecida, no es posible establecer una distribución para los errores.

• *Microestructura de mercados financieros y tasa de cambio*

La microestructura de los mercados financieros es el estudio de los procesos y resultados que se producen en el intercambio de activos con reglas de negociación explícitas (Marín y Rubio, 2001).

Dicha microestructura se centra en la interacción entre los mecanismos del proceso de negociación y sus resultados en términos de precios y cantidades negociadas. Se reconoce que las reglas específicas según las cuales se produce el proceso de negociación afectan directamente los resultados de tales procesos, es decir, el comportamiento de los agentes en el juego de demanda y oferta determina el precio y los volúmenes de transacción.

Cuestiones como la eficiencia de un determinado mecanismo de contratación, la naturaleza del proceso de ajuste de los precios ante la llegada de nueva información, y el papel de la información asimétrica entre los diversos agentes que intervienen en las negociaciones son piezas claves de la microestructura de mercados.

La introducción de microestructuras al estudio de la tasa de cambio permite considerar tres elementos que complementan la visión del mercado de activos: información que no es del dominio público, la existencia de diversos actores en el mercado que intervienen de diferente manera, y la presencia de varias formas de intercambio que afectan los precios. La idea central en la cual se sustenta este nuevo *approach* es considerar que verdaderamente quienes definen el precio son los que intervienen en la compra o venta del activo, y en ese sentido son relevantes sus preferencias dentro del mercado y la manera en que toman las decisiones acerca de comprar o vender. Lo anterior permite una visión complementaria a algunos *puzzles* de la tasa de cambio: explicar la alta volatilidad de la tasa de cambio, el problema de especificación a partir de fundamentos macro, y los inexplicables pero predecibles retornos excesivos a la especulación, a partir de la forma en que las decisiones son tomadas con expectativas racionales y no racionales, y con información pública y privada por parte de los actores del mercado.

A su vez el sustento de las microestructuras para el estudio de la TCN

se resume en *order flow* y *spread*. El primer concepto se refiere al volumen de transacciones y el sentido de las mismas, es decir, el volumen a transar y si se compra o se vende, lo que se asimila al exceso de demanda, toda vez que se realiza una operación; no necesariamente implica un equilibrio de suma cero. El segundo, a su vez, es condicionante del precio en la medida que se adentra en las asimetrías de la información en el mercado financiero. A modo de resumen, la idea fundamental es que los mecanismos específicos de negociación, contratación y liquidación, así como la organización concreta sobre la que trabajan los intermediarios afectan el comportamiento de los precios. Esto implica que los determinantes de los precios no pueden medirse exclusivamente en términos de rendimiento esperado y múltiples fuentes de riesgo. La propia estructura del mercado es en sí misma un determinante de los precios.

Es relevante igualmente mencionar que el enfoque de microestructuras en TCN¹¹ modela en la siguiente dirección:

$$\Delta P_t = f(X, I, Z) + \varepsilon_t,$$

donde el cambio en la tasa de cambio nominal entre dos transacciones (ΔP_t) (en contravía de variaciones promedio mensuales de los enfoques macro) es explicado a partir del orden de flujo (X), una medida de la posición del *dealer* o inventario (I) y un conjunto de variables

micro adicionales (Z). Sin embargo, el presente trabajo se mueve en la dirección de los híbridos, donde el pronóstico de la tasa de cambio (P_t) es explicado por variables micro y macro de la siguiente manera:

$$P_t = f(X, I, Z) + g(i) + \varepsilon_t,$$

donde las órdenes de flujo se agregan y se introduce una variable macro: tasas de interés nacionales y extranjeras (i). Esta modelación trata de solucionar los problemas de ajuste de los modelos macro de mercado de activos, en los que éste es asociado con las órdenes de flujo no consideradas, que en ambos casos se recoge por el término de error de la función (ε_t), pero se explica a partir de consideraciones micro. En ese sentido, el presente está interesado en la “demanda” entendida como el ofrecimiento de compra de una determinada divisa que contiene la información necesaria para identificarla o divulgarla y valorizarla; y la “oferta”, ofrecimiento de venta de una determinada divisa o contrato que contiene la información necesaria para identificarla o divulgarla y valorizarla. El sistema permite la transacción de divisas y/o contratos a través del sistema de mercado spot y sistema de mercado *Next day* (BVC, 2003), como variable micro y el diferencial de las tasas de cambio (DTF y LIBOR) como variable macro, como determinantes de las variaciones diarias de la TCN.

11. Como modelos originales provenientes de Lyons y Evans (1999).

Sobre las órdenes de flujo en Colombia cabe mencionar que las operaciones legales de compra y venta de dólares se desarrollan a través de la Bolsa de Valores de Colombia, específicamente en el Sistema Electrónico de Transacciones e Información del mercado de divisas SET FX, que puede definirse como un mecanismo electrónico a través del cual las entidades afiliadas pueden mediante estaciones de trabajo conectadas a una red computacional, en sesiones de negociación, ingresar ofertas y demandas, cotizar y/o celebrar entre ellas las operaciones, contratos y transacciones propios de su régimen legal en el mercado cambiario, de forma que quedan registradas las operaciones a nivel intradiarias.

Adicionalmente es de relevancia mencionar que el mercado de intercambio nacional está organizado como un mercado de negociantes descentralizados con N negociantes, y clientes de los negociantes, los cuales ejecutan sus transacciones concurrentemente con todos los participantes conectados al sistema SET FX. La negociación se inicia todos los días hábiles de lunes a viernes de 8:00 a.m. a 1:00 p.m.; generalmente antes del inicio de cada sesión los negociantes se reúnen con sus respectivos asesores en riesgo y analistas económicos para evaluar la tendencia del mercado de divisas y especular sobre las necesidades de los

clientes, lo cual va a marcar el ajuste del precio y de la forma hasta cuánto los clientes están dispuestos a apostarles a la compra y venta de divisas. Durante cada sesión los negociantes ejecutan las órdenes de flujo, ya sea ejerciendo presión a la venta o a la compra, dependiendo del comportamiento del mercado y de acuerdo con las necesidades de liquidez o comercialización que tengan sus clientes. Es de resaltar que cada una de las realizaciones que ejecutan los negociantes y los clientes de los negociantes son independientes, distribuidas normalmente, no públicamente observadas entre clientes y negociantes pero sí observables entre negociantes.

Análisis empírico

• Datos utilizados

La disponibilidad de bases de datos detalladas sobre la actividad intradiaria y diaria en los mercados financieros ha abierto la posibilidad a investigaciones econométricas sobre el funcionamiento de este tipo de mercados; por tanto, se tomó orden de flujo ODF, TRM y el diferencial de la Tasa de Interés Libor DDL, desde el 16 de abril del 2003 al 16 de abril de 2004.¹² El *set* de entrenamiento de la red neuronal artificial se definió de acuerdo con una medida del error entre los datos generados y el conjunto de entrenamiento (datos reales). Por lo general este valor varía, considerando valores pequeños.

12. El mercado de divisas en el país es relativamente nuevo (1991). Anteriormente las operaciones eran registradas, entre otros operadores, por el Citibank, y la disponibilidad de datos anteriores al proceso administrado por la BVC es compleja.

En el proceso de entrenamiento es probable que si la RNA sufre un proceso de sobreentrenamiento, lo que provoca que la RNA entre en un *loop*, o un subentrenamiento, entonces para ambos procesos la RNA perderá su capacidad de ajuste, pronóstico y generalización (Buitrago y Alcalá, 1998). En lo relacionado con la variable diferencial de tasa de interés (DDL) se tomó la *libor* a noventa días como componente macroeconómico de la economía estadounidense. La fuente de los datos es el sistema Reuter y por Colombia se tomó la DTF a noventa días. Estos datos tienen su origen en la página Web de Corfinsura.

Finalmente y teniendo en cuenta la definición antes dada de orden de flujo, éste se calcula como la diferencia entre todas las operaciones iniciadas a la compra (T) y todas las operaciones iniciadas a la venta (P) en el mismo día, de tal forma que si $OD > 0$ hay más intención de compra y por lo tanto una presión al alza en la tasa de cambio – la cual no necesariamente es explicada por variables macro del

mercado –; y cuando $OD < 0$ se produce el proceso inverso.

• *Análisis de resultados*

La Figura 3 presenta las variables en niveles. Es de esperar una alta volatilidad, característica de las series de tiempo financieras. Mediante los tests ADF y KPSS se determinó la estacionariedad de las series.

Los resultados muestran que a un nivel de significancia de 1%, las series TRM y Diferencial de tasas de interés son no estacionarias. La serie ODF es estacionaria. Adicionalmente se presenta el test Ljung-Box para el residuo de la regresión auxiliar en cada caso y entre paréntesis el p-value asociado a la prueba.

En el caso de modelos paramétricos se debe considerar esta propiedad de estacionariedad (más adelante se muestran los resultados del test de co-integración de Johansen). Para el caso de modelos univariados de pronóstico de TRM se trabajó con su primera diferencia en logaritmo natural (que corresponde al rendimiento compuesto continuamente).

Cuadro 1. Pruebas de raíz unitaria

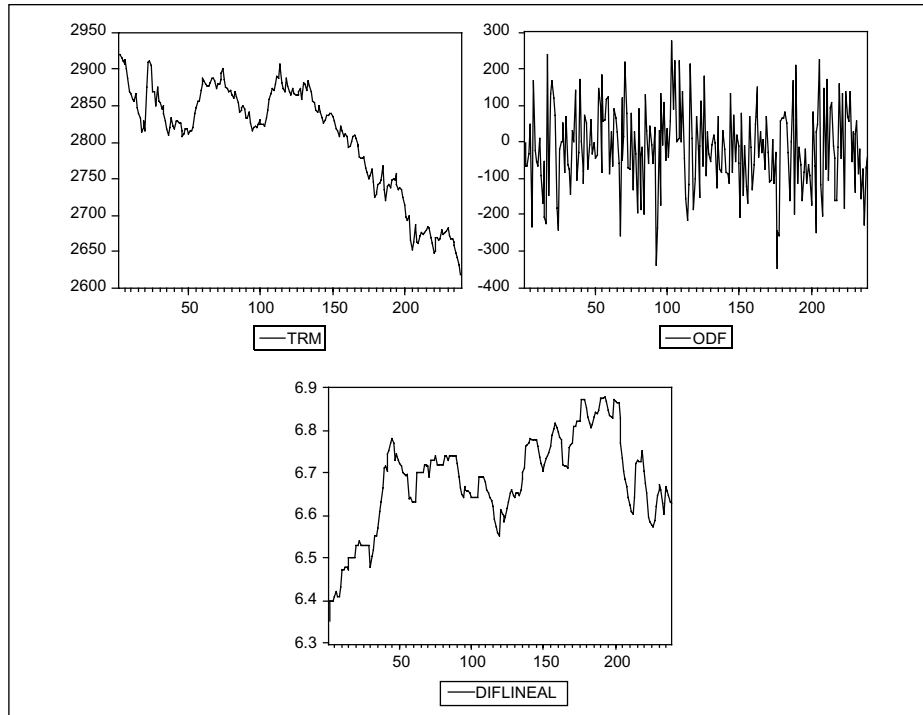
	ADF	Valor crítico 1%	KPSS	Valor crítico 1%	Decisión
TRM (1)	-1.829393	-2.574714	1.550905	0.739	I(1)
ODF (2)	-15.03944	-3.457747	0.173559	0.739	I(0)
Diferencial (3)	-2.66635	-3.457865	0.851827	0.739	I(1)

(1) $Q(36) = 44.25(0.163)$

(2) $Q(36) = 43.19(0.191)$

(3) $Q(36) = 22.937 (0.955)$

Fuente: Cálculos de los autores. Salida de programa E-views

Figura 3. Series en nivel

Fuente: Cálculos de los autores. Salida de programa E-views

• Modelos de Redes Neuronales Artificiales – RNA-

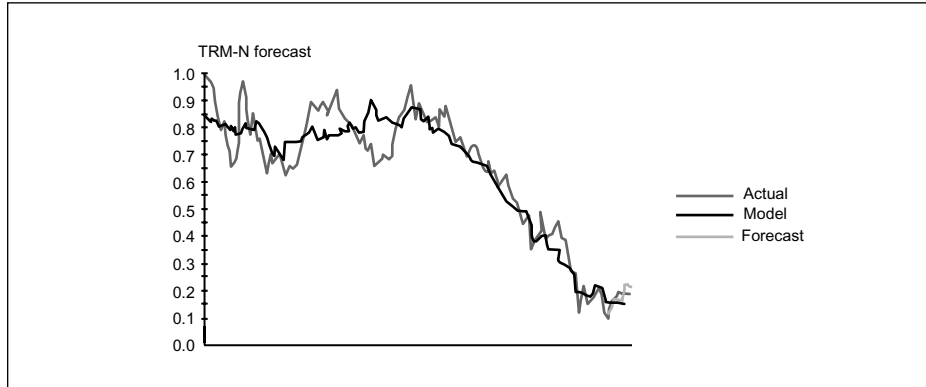
Como se explicó previamente, las RNA no tienen un modelo paramétrico específico; lo que se hizo fue variar el vector de entrada, variables independientes, orden de flujo, diferencial de tasa de interés (DTF-Libor) en un tiempo t y rezagadas en un τ .

Para el primer modelo, al aplicar el δ - test se obtuvieron los siguientes rezagos más significativos para el modelo: ODF_{t-1} , ODF_{t-2} , ODF_{t-9} , ODF_{t-13} , ODF_{t-36} , DDL_{t-1} , DDL_{t-5} , DDL_{t-6} , DDL_{t-8} , DDL_{t-9} , DDL_{t-17} .

Para el segundo modelo se considera la TRM rezagada en un $\tau = 1$, como si fuera un $AR(p)$; para este caso el δ - test arroja significancia para los rezagos TRM_{t-1} , TRM_{t-2} , TRM_{t-3} , TRM_{t-14} , TRM_{t-16} .

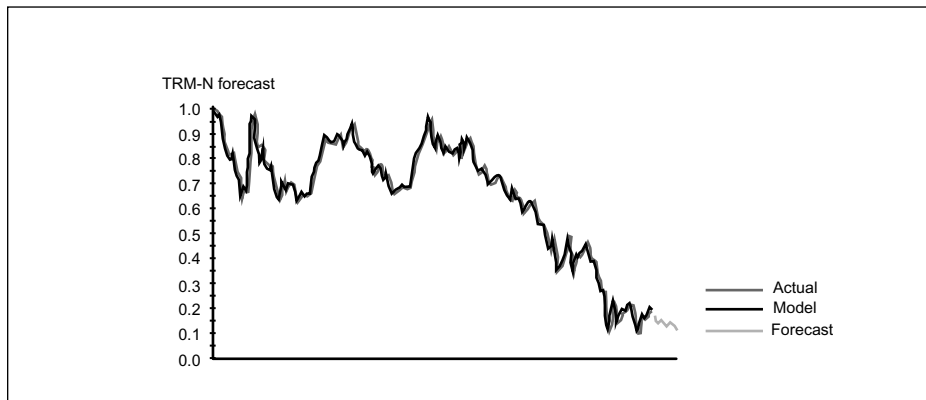
Según los pronósticos se observa que a medida que pasa el tiempo (días) el modelo 2 de RNA pierde capacidad de predicción; sin embargo, sus estadísticas de prueba (MAPE, RMSE y el aproximado de R^2) se comportan adecuadamente y con valores aceptables, lo cual hace de las técnicas de RNA para pronóstico de la TRM una herramienta viable, de la manera presentada.

Figura 4. Ajuste del modelo 1. RNA pronóstico h=10



Fuente: Cálculos de los autores. Salida de programa 4ANN

Figura 5. Ajuste del modelo 2 RNA. Pronóstico h=10



Fuente: Cálculos de los autores. Salida de programa 4ANN

Cuadro 2. Estadísticos de los modelos AR(p) y RNA – Un año

	RMSE	MAPE	R ²
	Pronóstico 10 datos		
RNA Modelo 1	0.0775	10.54870	93.54
RNA Modelo 2	0,036	5.671	97.3662

Fuente: Cálculos de los autores. Salida de programa 4ANN

• *Modelos de Series de Tiempo*

Análisis multivariado: En el Cuadro 3 se presenta el test de cointegración de Johansen. En todos los casos se presenta el test de la traza. Los resultados fueron obtenidos utilizando dos rezagos.

En todos los casos se encuentra que existe un vector de cointegración; sin embargo, como se vio anteriormente

una de las variables es $I(0)$, lo cual puede afectar esta conclusión; en particular, esta variable puede ser la que conforma el vector de cointegración. Utilizando el criterio de Schwartz, el modelo que se consideró adecuado fue el primero. Al estimar el VEC se obtuvieron los resultados expuestos en el Cuadro 4.

Cuadro 3. Test de Cointegración de Johansen

Vectores	Modelo 1	Valor crítico 1%	Modelo 2	Valor crítico 1%	Modelo 3	Valor crítico 1%	Modelo 4	Valor crítico 1%
Ninguno	99.38	29.75	104.03	41.07	100.82	35.65	106.29	48.45
Al menos 1	7.90(*)	16.31	12.36(*)	24.6	9.59(*)	20.04	14.94(*)	30.45
Al menos 2	2.42	6.51	2.54	12.97	0.00	6.65	3.52	16.26

Fuente: salida de programa E-views

Cuadro 4. Vector de Cointegración

	TRM	ODF	Diferencial
Coefficiente	1	32.33773	446.6811
Estadístico t		[10.4216]	[16.9962]

Fuente: salida de programa E-views

Cuadro 5. Matriz Alpha

Coefficiente	-0.00268	0.024488	-4.18E-07
Estadístico t	[-9.72161]	[5.37546]	[-0.61896]

Fuente: salida de programa E-views

En el anexo B se presenta el componente de corto plazo y la descomposición de varianzas para este modelo. Se observa que en el corto plazo únicamente la variable ODF causa la tasa de cambio; sin embargo, en el componente de largo plazo las variables ODF

y diferencial son significativas. Sobre este modelo se realizó el pronóstico 1 paso adelante para 10 periodos.

Análisis univariado: Se consideró un modelo alternativo sobre la primera diferencia de la tasa de cambio: Un modelo GARCH(1,1).

Cuadro 6. Causalidad de Granger y Test Portmanteau

	Causada: TRM	
Exclude	Tes	P-value
D(ODF)	96.86	0.00
D(DIFLINEAL)	2.43	0.30
	Causada: ODF	
D(TRM)	2.32	0.31
D(DIFLINEAL)1.04	0.59	
	Causada: DIFERENCIAL	
D(TRM)	1.120	0.571
D(ODF)	1.163	0.559
Portamentau(40)	6.99439	0.6377

Fuente: salida de programa E-views

Cuadro 7. Modelo Garch(1,1)

	Coefficiente	Z-estadístico	P-Value
C	-0.05	-2,05	0,04
D21	1,52	9,95	0,00
D25	-1,48	-3,24	0,00
D203	0,87	6,13	0,00
AR(1)	-0,34	-1,23	0,22
MA(1)	0,55	2,31	0,02
Variance Equation			
C	0,034	1,689	0,091
ARCH(1)	0,272	2,971	0,003
GARCH(1)	0,457	2,193	0,028
	Test	P-Value	
Ljung-Box	34,747	0,177	
Jarque-Bera	0,31	0,85	
Arch(4) Modelo de- finitivo	1,62	0,805	

Fuente: salida de programa E-views

Cuadro 8. Evaluación del pronóstico con modelos

	GARCH(1,1)	VEC	MODELO 1 RNA	MODELO 2 RNA
RMSE	22.90	20.66	31.98	17.24
MAPE	17.94	15.72	26.93	12.78

Fuente: Cálculos de los autores. Salida de programa 4ANN

Los resultados de evaluación del pronóstico muestran que el modelo 2 de redes neuronales presenta mejor pronóstico que los modelos paramétricos presentados.

• *Verificando la no linealidad*

Este documento propone un modelo no lineal para verificar la existencia de no linealidad. Para cada modelo se han desarrollado diversas pruebas como el Test Arch (referenciado anteriormente) o el test LSTR (mencionado en la introducción) aplicado a modelos con formas funcionales específicas.

Mills (1999) y Brooks (2002) presentan una discusión al respecto y

proponen cómo en el caso que no se tenga una forma funcional previamente definida, una forma general de verificar la no linealidad es el test de Ramsey sobre los residuales de un modelo lineal, generalmente un modelo autorregresivo. Igualmente, los autores proponen una variación denominada test de Tsay, el cual en lugar de incluir términos pronosticados a diferentes potencias adiciona el producto de términos rezagados. Las ventajas de estos tests es que no requieren una forma funcional no lineal específica. La forma funcional para verificar la no linealidad en el test de Tsay es:

$$\tilde{u}_t = \Psi' w_t + \sum_{i=1}^P \sum_{j=i}^P \varphi_{ij} y_{t-i} y_{t-j} + u_t^*$$

donde u son los términos residuales del modelo lineal. Con la hipótesis nula el modelo es lineal y el estadístico tiene distribución *chi* cuadrado con tantos grados de libertad como términos adicionales se incluyan en la anterior regresión.

En el caso particular de verificar la posibilidad de no linealidad de un tipo de red neuronal, Terasvirta *et al.* (1994:2931) proponen la siguiente forma funcional:

$$\tilde{u}_t = \Psi' w_t + \sum_{i=1}^P \sum_{j=i}^P \varphi_{ij} y_{t-i} y_{t-j} + \sum_{i=1}^P \sum_{j=i}^P \sum_{k=j}^P \varphi_{ijk} y_{t-i} y_{t-j} y_{t-k} + u_t^*$$

Igualmente, con la hipótesis nula se tendría linealidad. Tanto el estadístico de Tsay como este particular para redes neuronales se distribuyen *chi* cuadrado. La enorme ventaja de esta modificación es que permite discernir la no linealidad que se analiza en este trabajo.

A continuación se presentan los resultados para verificar la existencia de no linealidad en la tasa de cambio.

Para todos los casos se supuso como número de rezagos, $p=2$.

Los resultados muestran que se rechaza la hipótesis nula para el test de no linealidad tipo red neuronal artificial, según lo explicado por Mills (1999). Este resultado refuerza la conclusión parcial presentada previamente, según la cual el modelo alternativo más adecuado debería ser aquel que incluyese este tipo de linealidad.

Cuadro 9. Pruebas de no linealidad (Ho: Modelo es lineal)

	Test	P-value
Test de Tsay	4.154917	0.2452
Test no lineal RNA	15.79502	0.0033

Fuente: Cálculos de los autores. Programa utilizado STATA

Conclusiones y recomendaciones

Este trabajo se mueve en la misma dirección de los recientes trabajos de Evans y Lyons de las microestructuras y tasa de cambio, haciendo una primera aproximación al caso colombiano unido a una nueva metodología no-lineal para modelar –RNA– planteado para la inflación por Misas, López y Querubín (2002). Este enfoque permite analizar cómo las señales son percibidas de diferente forma por los agentes y de alguna manera se reflejan en la microestructura de la información, siendo, de todas maneras, asimetrías propias del mercado de divisas.

El uso de RNA, que son típicamente modelos estadísticos no lineales que pueden ser expresadas como un modelo genérico de aproximación de funciones, permitió pronósticos aceptablemente buenos para un modelo de TRM, según el RMSE, con la inclusión de la variable microeconómica orden de flujo.

Adicionalmente se verificó la existencia de no linealidad del tipo RNA. Quedaría por verificar el comportamiento de otros modelos del tipo heterocedasticidad condicionada, y por vectores autorregresivos de análisis univariado y multivariado para robustecer la formalidad econométrica y aventurar de esa manera conclusiones mayormente fuertes.

El poder explicativo de las variables macroeconómicas, tal como lo men-

cionan Evans y Lyons (1999), en el comportamiento del tipo de cambio en el corto plazo -diario- es poco significativo, ya que para los modelos desarrollados se comportan como no significativas. En este sentido, el mercado financiero con su comportamiento especulativo, explicado en gran parte por las órdenes de flujo, es un determinante del precio del dólar en el corto plazo, ya que el volumen diario de negociación supera en gran medida el volumen de la oferta y demanda reales de la economía, lo que explica en gran parte el poder del mercado financiero en la determinación del precio de las mismas.

La propia estructura del mercado es en sí misma un determinante de los precios y de la tasa de cambio nominal; específicamente el orden de flujo como indicador de los mecanismos de negociación, contratación y liquidación, afecta el comportamiento de los precios.

Como recomendación metodológica sería adecuado iniciar el estudio de líneas de investigación que fusionen diferentes técnicas de la computación bio-inspirada, como los algoritmos genéticos y la lógica difusa, para así obtener métodos predictivos y de optimización más eficaces y potentes, ya que ellas tratan de simular comportamientos humanos de maneras diferentes a las tradicionales modelaciones lineales y computacionales.

Bibliografía

- BAILLIE, R. AND T. BOLLERSLEV (1989). "COMMON STOCHASTIC TRENDS IN A SYSTEM OF EXCHANGE RATES", *Journal of Finance*, 44, 167-181.
- Bar koul as, T., Baum C., Cagl ayan, M., Chakr aborty, A (2000). "Persistent dependence in foreign exchange rates? A reexamination". Department of Economics and Accounting. University of Liverpool.
- Buitr ago, A., Alcal á, J., (1998). "Análisis, diseño e implementación de un prototipo de sistema neuronal para pronóstico de series de tiempo económicas". Departamento de Ingeniería de Sistemas. Universidad Nacional
- Bl anchar d, O., (1979). "Speculative bubbles, crashes, and rational expectations", *Economics Letters*, 14: 387-389.
- Cár denas, M. (1997). "La tasa de cambio en Colombia", Editores Tercer Mundo S.A. Primera Edición. Bogotá.
- Diebold, F. and J. Nason (1990). "Nonparametric exchange rate prediction?", *Journal of International Economics*, 28, 315-332.
- Dominguez, K., (1986). "Are foreign exchange forecasts rational? New evidence from survey data", *Economic Letters*, 21: 277-281.
- Dornbusch, R., (1982). "Equilibrium and disequilibrium exchange rates, *Zeitschrift fur Wirtschafts und Sozialwissenschaften*", 102: 573-799; reprinted in R. Dornbusch, *Dollars, Debts and Deficits*, MIT Press: Cambridge, MA.
- Engel, C. and J. D. Hamilton (1990). "Long swings in the dollar: Are they in the data and do markets know it?", *American Economic Review*, 80, 689-713.
- Evans, M., Lyons, R. (1999). "Order Flow and Exchange Rate dynamics". NBER Working Paper No. 7317.
- Evans, M., Lyons, R. (2004). "A New Micro Model of Exchange Rate Dynamics".
- Frankel, J., and A. Rose, (1995). Empirical research on nominal exchange rates, in G.
- Grossman and K. Rogoff (eds.), *Handbook of International Economics*, Elsevier Science: Amsterdam, 1689-1729.
- Flood, R., and A. Rose, (1995). "Fixing exchange rates: A virtual quest for fundamentals", *Journal of Monetary Economics* 36, 3-37.
- Gradojevic, N., Yang, J. (2000). "The Application Of Artificial Networks To Exchange Rate Forecasting: The Role Of Market Microstructure Variables" Bank of Canada December 2000, working paper 2000-23.
- Hau, H., (1998). "Competitive entry and endogenous risk in the foreign exchange market", *Review of Financial Studies*, 11: 757-788.
- Hu, M.Y., G. Zhang, C.X. Jiang, and B.E. Patuwo. (1999). "A Cross-Validation Analysis of Neural Network Out-of-Sample Performance in Exchange

- Rate Forecasting." *Decision Sciences* 30 (1): 197–216.
- Kuan, C-M. and T. LIU (1995). "Forecasting exchange rates using feedforward and recurrent neural networks", *Journal of Applied Econometrics*, 10, 347-364.
- Jalil, M. y Melo, L.F. (1999). "Una relación no lineal entre la inflación y los medios de pago". Borradores de Economía, Banco de la República.
- Lyons, R (2001). "The microstructure approach to exchange rates". MIT press.
- López, A.(1987). "Las minidevaluaciones en Colombia: un largo camino de búsqueda de una tasa de cambio libre, pero intervenida", *En: 20 años del régimen de cambios y de comercio exterior*. Banco de la República, Bogotá.
- Marín, C., Rubio, G. (2001). *Economía Financiera*. Barcelona: Antoni Bosch.
- Manrique, C. (2001). "La modelización del tipo de cambio de la peseta y el marco alemán durante el período 1987-1996". *Fundamentos de Análisis Económico*.
- Meese, R., and K. Rogoff, (1983a). "Empirical exchange rate models of the seventies, *Journal of International Economics*", 14: 3-24.
- Meese, R., and K. Rogoff, (1983b). "The out-of-sample failure of empirical exchange rate models, in J. Frenkel (ed.), *Exchange Rate and International Macroeconomics*, University of Chicago Press: Chicago.
- MEESE, R., AND K. ROGOFF, (1988). "WAS IT REAL? THE EXCHANGE RATE-INTEREST DIFFERENTIAL RELATION OF THE MODERN FLOATING-RATE PERIOD," *Journal of Finance*, 43: 933-948.
- Meese, R. A. and A. K. Rose (1991). "An empirical assessment of nonlinearities in models of exchange rate determination", *Review of Economic Studies*, 80, 192-196.
- Meese, R.A. and A.K. Rose. (1990). "Nonlinear, Nonparametric, Nonesential Exchange Rate Estimation." *The American Economic Review* 80 (2): 678–91.
- Meese, R.A. and A.K. Rose. (1991). "An Empirical Assessment of Non-Linearities in Models of Exchange Rate Determination." *Review of Economic Studies* 58 (3): 603–19.
- Montenegro, A.(2001). "*Redes Neuronales Artificiales*". Documentos de investigación, Departamento de Economía, Pontificia Universidad Javeriana.
- Oliveros C, H. y Huertas C, C. (2002). "Desequilibrios nominales del tipo de cambio en Colombia". Borradores de economía. Banco de la República.
- Plasmans, J., W. Verkoijen, and Daniels, H. (1998). "Estimating Structural Exchange Rate Models by Artificial Neural Networks." *Applied Financial Economics* 8: 541–51.
- Steiner, R.(1987). "El manejo de la tasa de cambio y el estatuto cambiario: los objetivos de largo plazo y los desarrollos coyunturales", *En: 20 años del régimen de cambios y de comercio exterior*. Banco de la República, Bogotá.

- TORRES, L., HERNÁNDEZ, G., NIÑO, L. (1994). REDES NEURONALES. BOGOTÁ. DEPARTAMENTO DE MATEMÁTICA. UNIVERSIDAD NACIONAL DE COLOMBIA.
- URRUTIA, M. (1981). "EXPERIENCE WITH THE CRAWLING PEG IN COLOMBIA" IN: J. WILSON, ED., EXCHANGE RATE RUIES, ST. MARTINUS PRESS. NY
- UNIVERSIDAD TECNOLÓGICA DE PEREIRA. (2000). "REDES NEURONALES ARTIFICIALES". FACULTAD DE INGENIERÍA ELÉCTRICA. DISPONIBLE EN <http://ohm.utp.edu.co/neuronales>
- Wiesner, E. (1978). "Devaluación y mecanismos de ajuste en Colombia" En: Política económica externa en Colombia. Asociación Bancaria de Colombia, Bogotá, pp. 123-224.
- VERKOOIJEN, W. (1996.). "A NEURAL NETWORK APPROACH TO LONG-RUN EXCHANGE RATE PREDICTION." *Computational Economics* 9: 51-65.
- Zhang, G. and M.Y. Hu. 1998. "Neural Network Forecasting of the British Pound/US Dollar Exchange Rate." *International Journal of Management Science* 26 (4): 495-506.
- MILLS, T (1999) The Econometric of Financial Modelling. Second Edition. Cambridge University Press.
- Brooks, J (2002) Introductory Econometrics For Finance. Cambridge University Press.
- Terasvirta, T *et al.* (1994). Aspects of Modelling Nonlinear Time Series. Handbook of Econometrics. Chapter 48.

Anexo A

Formulación General del δ - _test

Sea x_t una serie de tiempo, $t = 1, 2, 3, \dots, N$

$$x_t = f(x_{t-1}, x_{t-2}, \dots, x_{t-d}) + g(r_t, r_{t-1}, \dots, r_{t-d}). \quad (1)$$

La serie de tiempo es representada como una sucesión de N puntos $z(i)$ en un espacio de dimensión $(d+1)$, $(d=0,1,2,\dots)$

$$z(i) = (z_0(i), z_1(i), \dots, z_k(i), \dots, z_d(i)), \quad (2)$$

donde $z_k(i) = x_{t-k}$ definiendo :

$$\begin{aligned} l_k(i, j) &= |z_k(i) - z_k(j)|, & k &= 0, 1, \dots, d \\ n(\epsilon \leq l \leq \delta) &= n(\epsilon \leq l_0 \leq \delta, l_1 \leq \delta, \dots, l_d \leq \delta) \\ n(l \leq \delta) &= n(l_0 \leq \delta, l_1 \leq \delta, \dots, l_d \leq \delta). \end{aligned} \quad (3)$$

Se obtienen los valores de probabilidades para la dimensión:

$$\begin{aligned} P(l_0 \leq \epsilon, l \leq \delta) &= \frac{1}{N_{par}} n(l_0 \leq \epsilon, l \leq \delta), \\ P(l \leq \delta) &= \frac{1}{N_{par}} n(l \leq \delta). \end{aligned} \quad (4)$$

donde N_{par} es el número total de parejas de vectores. A partir de las anteriores definiciones se calculan las probabilidades condicionales de la ecuación 5.

$$\begin{aligned} P_d(\epsilon; \delta) &= P(l_0 \leq \epsilon, l \leq \delta) = \frac{P(l_0 \leq \epsilon, l \leq \delta)}{P(l \leq \delta)}, \\ P_d(\epsilon) &= \max_{\delta} P_d(\epsilon; \delta) = P_d(\epsilon; \delta)_{\delta \rightarrow \infty}. \end{aligned} \quad (5)$$

$P_d(\epsilon)$ mide qué tan bien la dinámica puede ser modelada en términos de la variable d . Para cuantificar la dependencia de cada una de las variables, es conveniente definir el *índice de dependencia*:

$$\lambda_d(\varepsilon) = \frac{P_d(\varepsilon) - P_{d-1}(\varepsilon)}{1 - P_d(\varepsilon)}, \quad d = 1, 2, \dots \quad (6)$$

y su promedio sobre ε

$$\bar{\lambda}_d(\varepsilon) = \frac{\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} d\varepsilon \lambda_d(\varepsilon) (1 - P_d(\varepsilon))}{\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} d\varepsilon (1 - P_d(\varepsilon))} = \frac{\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} d\varepsilon (P_d(\varepsilon) - P_{d-1}(\varepsilon))}{\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} d\varepsilon (1 - P_d(\varepsilon))}. \quad (7)$$

Para un sistema determinístico libre de ruido, $P_d(\varepsilon)$ se satura a 1 para $\varepsilon \geq d_0$ y se tiene

$$\sum_{d=1}^{d_0} \bar{\lambda}_d = \sum_{d=1}^{d_0} \lambda_d(\varepsilon) = 1. \quad (8)$$

Anexo B

Componente de corto plazo – Vec

Error Correction:	D(TRM)	D(ODF)	D(DIFLINEAL)
CointEq1	-0.00268 [-9.72161]	0.024488 [5.37546]	-4.18E-07 [-0.61896]
D(TRM(-1))	-0.16032 [-2.50996]	-1.605227 [-1.51920]	5.39E-05 [0.34402]
D(TRM(-2))	-0.07862 [-1.55135]	0.008388 [0.01000]	-0.000121 [-0.97518]
D(ODF(-1))	-0.05091 [-6.15994]	-0.214042 [-1.56548]	-9.01E-06 [-0.44446]
D(ODF(-2))	-0.00274 [-0.48561]	-0.101893 [-1.09048]	-1.30E-05 [-0.93642]
D(DIFLINEAL(-1))	-4.59097 [-0.18186]	-50.08267 [-0.11993]	0.23435 [3.78351]
D(DIFLINEAL(-2))	-35.1767 [-1.42513]	-380.8994 [-0.93286]	0.137912 [2.27720]
D9	-29.2282 [-4.02189]	-73.35969 [-0.61023]	0.02447 [1.37233]
D21	44.99588 [6.14487]	105.7153 [0.87274]	-0.005439 [-0.30276]
D22	31.92607 [4.04969]	146.2174 [1.12120]	-0.000185 [-0.00957]
D25	-21.5619 [-2.96531]	-23.33663 [-0.19401]	0.001495 [0.08381]
D120	2.336862 [0.32214]	67.2767 [0.56064]	0.06987 [3.92559]
D62	-0.49423 [-0.06887]	-75.50492 [-0.63609]	0.06922 [3.93156]
D214	2.653032 [0.36252]	100.8995 [0.83346]	0.06946 [3.86834]
D57	-7.45664 [-1.03198]	66.56916 [0.55694]	-0.05681 [-3.20445]
R-squared	0.598634	0.506211	0.304177

Fuente: Cálculos de los autores. Salida de programa E-views

Anexo C. Descomposición de varianza – Modelo VEC

.....				
.....
.
.
.
.
.
.
.
.
.
..
.....				
.....
.
.
.
.
.
.
.
.
.
..
.....				
.....
.
.
.
.
.
.
.
.
..

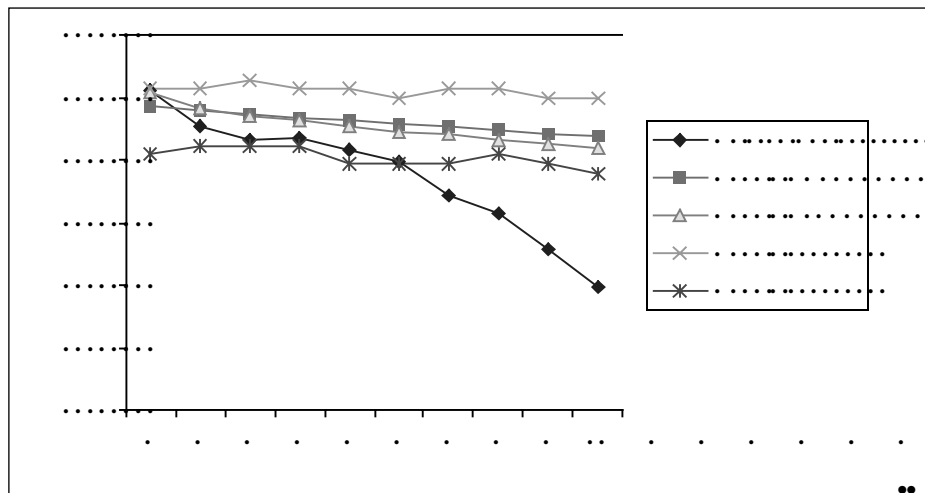
Fuente: Cálculos de los autores. Salida de programa E-views

Anexo D.

Resultados de los pronósticos por modelo

Valores Reales	Modelos Lineales		Modelos de Redes neuronales	
	Modelo GARCH	Modelo VEC	Modelo 1	Modelo 2
2682,09	2677,13	2681,51	2682,68	2661,65
2671,01	2675,72	2676,35	2682,68	2664,66
2666,55	2674,49	2674,28	2685,69	2664,66
2667,06	2673,51	2672,41	2682,68	2664,66
2663,05	2672,53	2670,79	2682,68	2658,65
2659,05	2671,53	2669,21	2679,68	2658,65
2648,8	2670,51	2667,94	2682,68	2658,65
2642,55	2669,51	2666,34	2682,68	2661,65
2631,72	2668,50	2664,89	2679,68	2658,65
2619,59	2667,50	2663,81	2679,68	2655,64

Fuente: Cálculos de los autores. Programa Excel



Fuente: Cálculos de los autores. Programa Excel