



BANCO DE LA REPÚBLICA
SUBGERENCIA DE ESTUDIOS ECONÓMICOS

**La inflación en Colombia:
una aproximación desde las redes neuronales**

Martha Misas Arango
Enrique López Enciso
Pablo Querubín Borrero*

Bogotá, febrero de 2002

* Se agradecen de manera especial los comentarios y sugerencias de Miguel Urrutia M. Como también, la colaboración de Norberto Rodríguez N. y Rocío Betancourt G. en la descripción y evaluación de los modelos lineales. Los resultados y opiniones son responsabilidad exclusiva de los autores y no comprometen al Banco de la República ni a su Junta Directiva.

Introducción

Las redes neuronales (ANN) ¹ son modelos computacionales diseñados para simular el funcionamiento del cerebro y, en particular, la forma como éste procesa información. En el contexto de análisis de series de tiempo, se clasifican como modelos no lineales entrenados para (i) realizar conexiones entre los valores pasados y presentes de una serie de tiempo y (ii) extraer estructuras y relaciones escondidas que gobiernan el sistema de información. El atractivo de este enfoque, inspirado en la neurología, es su habilidad para aprender, es decir, para identificar dependencias con base en una muestra finita, de manera que el conocimiento adquirido pueda ser generalizado a muestras no observadas (Herbrich et.al, 1999). Si bien, como señalan Kuan y White (1994), las redes neuronales y sus algoritmos de aprendizaje asociados están todavía lejos de ofrecer una descripción acertada de cómo funciona el cerebro, éstas se han constituido en un marco de modelación muy poderoso e interesante cuyo potencial ha sido comprobado en diversas aplicaciones en todas las ciencias². Para Moshiri y Cameron (1998), los investigadores son atraídos hacia ese enfoque porque las redes neuronales no están sujetas a supuestos restrictivos como la linealidad, que suele ser necesaria para la aplicación de los modelos matemáticos tradicionales.

La reputación de las redes neuronales para manejar datos e información compleja puede haber contribuido considerablemente a la difusión e implementación de modelos de redes neuronales en la economía y la econometría. Herbrich, et.al (1999) señalan tres campos principales en los cuales se ha concentrado la aplicación de redes neuronales en economía: i) clasificación de agentes económicos, ii) pronóstico de series de tiempo y iii) modelaje de agentes con racionalidad limitada. El segundo campo ha sido de especial importancia. Franses y van Dijk (2000) señalan que anualmente se publican alrededor de 20 o 30 artículos relacionados con el pronóstico y modelación, a través de redes neuronales, de los precios de las acciones, la tasa de cambio, las tasas de interés, el producto, y la inflación entre otros³. Franses y van Dijk (2000), como Tkacz y Hu (1999), atribuyen la creciente popularidad de estos modelos ANN y su aplicación a series de tiempo, a la capacidad que poseen para permitir relaciones no lineales muy generales entre las variables. En efecto, al ser especificados de manera apropiada, pueden aproximar cualquier función no lineal. Adicionalmente, las redes neuronales son una herramienta importante en la modelación de variables en las cuales la existencia de un modelo estructural no

¹ Conocidas en la literatura como redes neuronales artificiales o ANN (Artificial Neural Network).

² Durante los años sesenta las redes neuronales fueron desarrolladas como una herramienta para la clasificación y reconocimiento de patrones. Otras aplicaciones mencionadas por Kuan y White (1994) son: conversión de texto escrito a voz, reconocimiento de caracteres escritos a mano, realización de complejas tareas de coordinación, juegos de *backgammon*, diagnóstico de dolor de pecho y decodificación de caos determinístico.

³ Se destacan los artículos de Gradojevic y Yang (2000), Medeiros, Veiga y Pedreira (2001) y Evans (1997) que modelan y pronostican tasa de cambio, los artículos de Moshiri y Cameron (1998), Tkacz (2000) y Gazely y Binner (2000) para inflación, el trabajo de Moody, Levin y Rehfuss (1993) para el índice de producción industrial, Aiken (1996) y Draisma et.al (1995) para la tasa de desempleo, Tkacz y Hu (1999) para el producto interno bruto, y el trabajo de Swanson y White (1997) para el pronóstico de nueve series macroeconómicas como desempleo, producto, exportaciones, entre otras.

es clara, pues no parten de supuestos *a-priori* sobre los datos para el pronóstico y todo lo que de ellas puede decirse es inherente a las observaciones (Evans, 1997).

El objetivo de este trabajo es modelar la relación entre dinero e inflación utilizando un modelo de redes neuronales. Los agregados monetarios han sido utilizados tradicionalmente como determinantes o variables explicativas de la inflación. Sin embargo, la existencia de asimetrías entre la política monetaria y la inflación al igual que la evidencia de fijación asimétrica de precios en los agentes económicos, pueden justificar la existencia de no-linealidades entre dinero e inflación. Por lo tanto, la aplicación de esta técnica, capaz de capturar estas no linealidades, puede generar pronósticos más precisos de la inflación constituyéndose en una herramienta de pronóstico y de modelación muy valiosa.

Este trabajo está ordenado de la siguiente forma: en la segunda sección se discuten algunos elementos teóricos sobre la relación entre dinero e inflación y las posibles fuentes de no-linealidades entre estas dos variables. La tercera sección presenta la metodología e introduce al lector a las redes neuronales. La cuarta sección discute posibles interpretaciones económicas del *hidden layer* o superficie escondida, un elemento central de las redes neuronales artificiales. La quinta sección reporta los resultados del trabajo y finalmente en la última sección se mencionan algunas conclusiones.

2. Dinero e inflación

2.1 Importancia del análisis de los agregados monetarios.

Hoy en día la mayor parte de los economistas están de acuerdo en que en el largo plazo la inflación es un fenómeno monetario. De hecho, esa aseveración constituye uno de los pilares básicos de la teoría económica moderna. De otro lado, la relación de largo plazo entre el dinero y los precios ha sido confirmada por un gran número de estudios empíricos, ya sea por ejercicios de corte transversal entre países o en series de tiempo para el caso de economías particulares.

La relación fundamental entre dinero y precios y el monopolio que sobre la creación de la base monetaria tienen los bancos centrales lleva a que éstos últimos tengan un interés natural en la evolución de los agregados monetarios. En los años ochenta muchos bancos centrales buscaron cumplir con objetivos monetarios como forma de hacer política monetaria. Esa elección coincidió con que en ese período las innovaciones financieras y el cambio estructural que se dieron en la economía mundial modificaron las estructuras financieras. La consecuencia final fue la inestabilidad de la demanda de dinero con lo cual se hacía difícil interpretar la evolución de los agregados monetarios, y el abandono de las metas monetarias.

Hoy se arguye, como fruto de esa experiencia, que los bancos centrales no deben utilizar los agregados monetarios como indicadores del impacto de sus políticas en la economía⁴. Esa percepción también responde a que en el análisis moderno de la política monetaria, o por lo menos en una de sus vertientes más importantes, los agregados monetarios no tienen un papel explícito, como se puede constatar en los modelos teóricos y en los empíricos⁵. En esos modelos la característica típica es que el banco central utiliza la tasa de interés como su instrumento de política. Se supone que la política monetaria afecta directamente a variables importantes como la inflación y el producto, sin que la cantidad de dinero tenga algún papel intermedio en su determinación. Si bien el acervo de dinero que debe ser ofrecido para llegar a un nivel dado de tasa de interés puede ser determinado por medio de una función de demanda de dinero, esto no es realmente necesario para caracterizar la economía. El dinero, en consecuencia, es en esencia superfluo en el modelo.

Sin embargo, la mayoría de los bancos centrales, sea cual sea la estrategia de política monetaria que formalmente lleven a cabo, continúan monitoreando los agregados monetarios. Existen al menos dos argumentos para sustentar ese comportamiento. En primer lugar, el dinero se encuentra estrechamente relacionado con el crédito, y debe ser una parte muy importante del canal de crédito de la transmisión de la política monetaria. En segundo lugar, se reconoce que las variables monetarias contienen información muy valiosa.

Con relación al primer punto, Altissimo et al. (2001) hacen un ejercicio en el que investigan si los residuales de las ecuaciones de demanda de dinero y de crédito del modelo central de la Banca de Italia están correlacionados con los errores de pronóstico de las variables económicas esenciales. Sus resultados sugieren que la evolución de los agregados monetarios y del crédito revela información importante para entender los mecanismos de transmisión de la política monetaria asociados al comportamiento de los *spreads* de tasa de interés o al racionamiento cuantitativo de crédito.

El reconocimiento que se hace acerca del contenido de información de las variables monetarias es de una amplia gama. De una parte, se encuentran las explicaciones estructurales de la inflación donde el dinero tiene un papel activo y dominante. De otro lado, se da un tratamiento al dinero en el que como máximo se le reconocen sus virtudes como indicador sintético, sin ninguna explicación causal en la evolución de la variable precios.

Con relación a ese último punto, Friedman (1984) reconoce que los agregados monetarios pueden ser excelentes indicadores del proceso inflacionario futuro o de los

⁴ De hecho muchos bancos centrales abandonaron las metas monetarias mientras que en otros la importancia dada a los indicadores monetarios disminuyó. En muchos bancos centrales europeos, sin embargo, los agregados monetarios y de crédito continuaron jugando un papel esencial en las décadas de los ochenta y de los noventa (ECB, 2001).

⁵ Véase, por ejemplo, los modelos teóricos de Svensson (1997) y Svensson y Woodford (1999). Dentro de los modelos empíricos un ejemplo patente es el modelo FRB/US.

mecanismos de transmisión de la política monetaria, a pesar que no jueguen un papel causal ni estructural. Desde ese punto de vista, aún si la inflación se entiende como el resultado de un exceso de demanda o de la presión de costos, la evolución de los agregados monetarios provee información para las decisiones de política monetaria, dado que permite a las autoridades identificar mejor la naturaleza de los choques que afectan a la economía y/o predecir las tendencias en la evolución futura de los precios.

Una explicación de las propiedades de indicador líder que tiene el dinero para predecir la inflación futura, se encuentra en que los agregados monetarios pueden estar relacionados con variables macroeconómicas importantes en la transmisión de la política monetaria. Por ejemplo, los estudios de demanda de dinero sugieren que el crecimiento del dinero está relacionado positivamente con el crecimiento del producto y negativamente con la tasa de interés. En ese contexto, el crecimiento del producto más allá de un cierto límite provocará un incremento en el crecimiento monetario y el surgimiento de presiones inflacionarias. Desde ese punto de vista, el crecimiento del dinero resume la información acerca de los determinantes de la demanda de dinero, con influencia, también, en la evolución futura de los precios.

2.2 Análisis monetario y no linealidad.

Es muy amplia la variedad de instrumentos analíticos y técnicas disponibles para hacer análisis monetario y extraer información de los agregados monetarios. Al respecto, un punto de partida natural es la estimación de la demanda de dinero⁶. Los estudios recientes que presentan este tipo de estimaciones utilizan en su mayoría especificaciones de corrección de errores y técnicas de modelado basadas en la cointegración (Pill, 2001). En los bancos centrales de Canadá y Reino Unido se utilizan una gran variedad de técnicas econométricas que, sorprendentemente, son muy similares en los dos casos: modelos VAR y VEC (Hauser, 2001; Masuch, 2001).

Una variante de los modelos VEC son los modelos P^* de la inflación (Hallman et al, 1989, 1991)⁷. Este tipo de modelos utiliza la desviación de la cantidad de dinero corriente de un nivel consistente con un equilibrio de largo plazo como un indicador adelantado para la inflación. Dos desarrollos recientes de este tipo de modelos son los de Gerlach y Svensson (2000) y Orphanides y Porter (2001). Los primeros autores, en el contexto de la discusión sobre la política monetaria en Europa, encuentran que el modelo P^* tiene un gran poder predictivo sobre la inflación futura, lo cual no se encuentra para el caso del crecimiento de $M3$ ⁸. Para los segundos, si se quiere obtener buenos pronósticos de la inflación en los modelos P^* , se debe modelar como variable en el tiempo a la velocidad de circulación del dinero y no asumirla como una constante

⁶ Laidler (1993) revisa los aportes al respecto.

⁷ Existe una aplicación para Colombia realizada por Misas, López y Melo (1999). Ver anexo No.4

⁸ Nicoletti Altamari (2001) encuentra un resultado similar para una brecha monetaria extraída de un modelo P^* , pero uno diferente para el caso de $M3$.

como frecuentemente se hace. Resulta conveniente también poder prever los cambios en la velocidad por medio de un análisis institucional detallado, para así poder tener un conocimiento adecuado de la situación económica.

Recientemente, los analistas han avanzado hacia modelos que consideran la existencia de una estructura no lineal en la relación del dinero con otras variables. Como anota Tkacz (2000), el supuesto de linealidad no puede ser justificado si se piensa que los efectos de la política monetaria en la economía son asimétricos. Ya Friedman en 1968 señala, que una política contraccionista tendrá un impacto relativamente mayor sobre el crecimiento del PIB que una política expansionista equivalente⁹.

Un impulso muy importante al tratamiento de las no linealidades proviene de los modelos recientes de demanda de dinero. Lütkepohl et al. (1999) encuentran que a partir de la unificación de Alemania su estimación de la demanda de dinero presenta inestabilidad. Por tal razón los autores sugieren una especificación no lineal de la demanda de dinero (LSTR)¹⁰. Según esos autores la fuente de la no linealidad se encuentra en la variable de tasa de inflación. Esta última, además de recoger una porción sustancial de la estacionalidad del modelo, es la variable que permite controlar que el ajuste del modelo de demanda se de en términos reales o nominales, como predice la teoría económica.

Soto (1997), quien construye un modelo de redes neuronales, ha profundizado en las posibles fuentes de no linealidad en la demanda de dinero. Para este autor existen tres elementos que pueden llevar a una estructura no lineal en la relación entre el dinero, el ingreso y la tasa de interés: la especificación microeconómica de la función, la agregación y la intermediación financiera.

Con base en un modelo genérico de costo de transacción, Soto demuestra la existencia de fuentes potenciales de no linealidad que surgen del comportamiento microeconómico de los agentes. Con el fin de obtener un tipo de especificación de la demanda de dinero que pueda ser estimada con métodos econométricos, se suponen funciones explícitas de utilidad y de costos de transacción. Estas funciones muy restrictivas suponen, por ejemplo, que la tasa de descuento es constante e independiente de la riqueza de los hogares o que la función de utilidad es separable y es una función CES. Este tipo de funciones no cumplen con los requerimientos matemáticos de la teoría o al basarse en el agente representativo no permiten capturar el comportamiento del consumo agregado. Este tipo de enfoque se aleja de los modelos cerrados simples y genera estructuras no lineales caracterizadas, como ejemplo, por *thresholds* o comportamiento dependiente del tiempo por parte de los agentes.

⁹ Esta asimetría en el impacto de la política monetaria sobre el PIB podría extenderse a una asimetría de la política monetaria sobre la inflación si se considera una Curva de Phillips de corto plazo.

¹⁰ Logistic Smooth Transition Regression Model. Ver sección 3.

Las no linealidades que surgen de la agregación de funciones individuales se producen porque se asume la existencia de n individuos idénticos. Se puede dar el caso, sin embargo, en que el comportamiento combinado de los agentes afectados por *shocks* idiosincráticos o colectivos lleve a desviaciones radicales del comportamiento individual. En ese caso se tendría una respuesta no lineal a los *shocks*.

También es posible obtener no linealidades en la respuesta de la demanda de dinero a cambios en los fundamentales, como consecuencia del comportamiento de los intermediarios financieros. Como ha mostrado la literatura del canal del crédito, los intermediarios surgen como consecuencia de las asimetrías de información típicas de los mercados financieros. Si se consideran las innovaciones financieras, su presencia puede llevar a una respuesta no lineal de los agregados monetarios a cambios en los fundamentales.

Rhee y Rich (1995), retoman la discusión de precios rígidos de la literatura nekeynesiana para mostrar que en estos modelos los *shocks* monetarios tienen un efecto asimétrico sobre la inflación. La inflación responde en mayor grado a los *shocks* monetarios positivos que a los *shocks* negativos. La asimetría estaría positivamente relacionada con los movimientos en la inflación promedio. El ejercicio empírico realizado para Corea provee evidencia acerca de la existencia de esa asimetría y de que esta es más marcada si hay movimientos positivos en la inflación promedio.

Para Gazely y Binner (2000), para el caso de la relación entre dinero e inflación cualquier relación funcional particular especificada es arbitraria y cualquier conclusión derivada está sujeta a una especificación arbitraria del modelo, motivo por el cual la flexibilidad de las redes neuronales y su capacidad para aproximar cualquier función no lineal representan elementos esenciales para el pronóstico y modelación de estas variables. De otro lado, estas autoras sugieren que buena parte del fracaso en los *targets* monetarios se debe a un problema de mala medición de los agregados. Su propuesta es construir índices monetarios tipo Divisia y utilizarlos como variable insumo dentro de una red neuronal construida para pronosticar la inflación¹¹.

3. Metodología

En esta sección se explican, de manera general, algunos tópicos relacionados con la arquitectura de la red neuronal biológica vs. la red neuronal artificial, como también, los procesos de estimación, selección y validación de una red neuronal artificial.

¹¹ De acuerdo con este enfoque el problema radica en que los agregados monetarios son construidos por medio de una suma simple. Lo ideal sería efectuar una suma ponderada, donde el factor de ponderación provenga del rendimiento del activo respectivo.

Como lo señalan Granger y Teräsvirta (1993), la construcción de un modelo que relacione a una variable y_t con su propia historia y/o con la historia de otras variables, X_t , puede llevarse a cabo a través de una variedad de alternativas. Estas dependen de la forma funcional mediante la cual se aproxima la relación, como también, de la relación existente entre dichas variables, es decir, de si ésta es de carácter lineal o no lineal. Las diferentes alternativas pueden ser clasificadas de la siguiente forma:

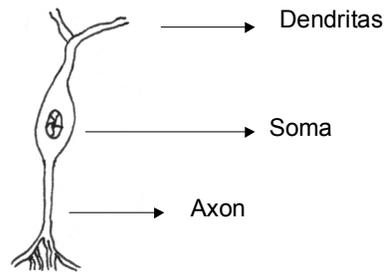
- No paramétrico: $y_t = f(X_t) + e_t$ donde f no está restringida a pertenecer a una clase específica de funciones.
- Paramétricos: supone una forma funcional específica para $f(\)$ usualmente con parámetros que deben ser estimados. Por ejemplo:
 - Lineales: $y_t = \beta' X_t + e_t$
 - No lineales:
 - Transición suave¹²: $y_t = \beta_1' X_t + F(X_t) \beta_2' X_t + e_t$ donde la función $F(\)$ captura la transición del modelo
 - Redes Neuronales: $y_t = \Phi_0 + X_t' \Phi + \sum_{j=1}^q \beta_j G(Z_t' \gamma_j) + e_t$
- Semiparamétricos: $y_t = \beta' \bar{X}_t + f(Z_t) + e_t$ donde las variables entran en el modelo de forma paramétrica y no paramétrica.

3.1 ¿Qué es una red neuronal artificial?, ¿cuál es su relación con la red neuronal biológica?

La pregunta que surge inicialmente al abordar el tema de redes neuronales artificiales es ¿qué es una red neuronal artificial?. Una respuesta a esta pregunta está dada por Kuan y White (1994) y Swanson y White (1995, 1997 a, b), quienes señalan que las redes neuronales artificiales, propuestas por científicos cognitivos, son una clase de modelos no lineales flexibles, capaces de aprender a través de ensayo y error, los cuales están inspirados en ciertas características asociadas a la forma como el cerebro humano procesa información.

De acuerdo a Klerfors (1998) y Plummer (2000), la relación entre una red neuronal artificial y el esquema de procesamiento del cerebro humano puede ser entendida a través de la siguiente explicación: el elemento básico del cerebro humano es un tipo específico de célula, llamada neurona, la cual nos provee, en nuestro quehacer continuo, de capacidad de pensar, recordar y conectar con experiencias previas. Cada una de estas neuronas puede conectarse con otras células similares. Así, la potencia del cerebro proviene del número de dichos elementos básicos y de las múltiples conexiones generadas entre ellos.

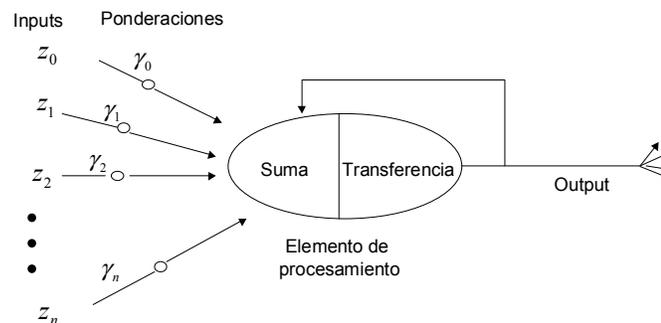
FIGURA 1



Las neuronas naturales o biológicas tienen tres componentes morfológicos básicos: dendritas, soma y axon y un espacio de comunicación o sinapse. Básicamente, una neurona recibe *inputs* de diversas fuentes, los cuales son combinados de formas distintas, generalmente a través de esquemas no lineales, para producir finalmente un *output*. La figura 1 presenta una neurona biológica simplificada y la relación de sus componentes. Existen varios aspectos fundamentales para resaltar la forma como la red neuronal biológica procesa la información: (i) paralelismo masivo¹³, (ii) respuesta no lineal ante un input¹⁴, (iii) procesamiento a través de múltiples superficies conformadas por unidades neuronales¹⁵ y (iv) retroalimentación dinámica entre unidades.

Las redes neuronales artificiales son modelos computacionales construidos siguiendo la estructura de la neurona biológica. Así, la unidad básica es la neurona artificial, la cual imita las cuatro funciones básicas de la neurona biológica. En la Figura 2 se presenta una neurona artificial.

FIGURA 2



¹²Dos trabajos para Colombia que utilizan este tipo de modelo son los de Arango y González (1998) y Jalil y Melo (2000).

¹³ En los sistemas neuronales biológicos el número de unidades de procesamiento pueden ser del orden de billones o más (masivo). Tal número nunca es encontrado en los modelos económicos. Como lo señalan Kuan y White (1994), el procesamiento en paralelo es lo común.

¹⁴ Esta es la tendencia de ciertos tipos de neuronas que solo reaccionan ante cierto nivel de estímulo o actividad input, es decir, se vuelven activas cuando éste supera cierto límite particular. Este punto introduce el hecho fundamental de la respuesta no lineal al paradigma de las redes neuronales.

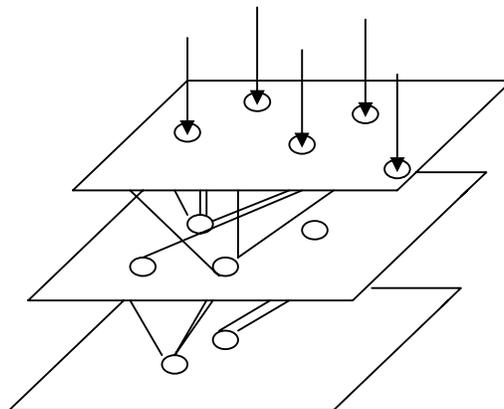
¹⁵ Es común observar procesos que se realizan a través de múltiples superficies. Por ejemplo, en la corteza cerebral humana se encuentran seis superficies de procesamiento diferentes.

Como se observa en la Figura 2, la neurona artificial recibe distintos *inputs* (z_n). Estos, antes de llegar al área de procesamiento, son multiplicados por una ponderación específica o conector ponderante, (γ_i). En el caso más sencillo, en esta área dichos productos son agregados y filtrados para producir un *output*.

Tanto la red neuronal biológica como la red neuronal artificial presentan una estructura topológica similar. En tal estructura la conexión con el mundo real se lleva a cabo a través de neuronas que reciben de éste sus distintos *inputs*, otras que lo proveen con los *outputs* y las restantes que se encuentran “escondidas” de éste. En la Figura 3 se observa cómo las neuronas se agrupan en diferentes superficies, dependiendo de su conexión con el mundo exterior¹⁶. Así, cuando la superficie-*input* recibe una señal o *input* sus neuronas producen un *output*, el cual se convierte en *input* de la siguiente superficie. El proceso continúa hasta que ciertas condiciones son alcanzadas y la superficie-*output* produce el resultado final.

La comunicación entre neuronas se da a través de un conjunto de conectores. Así, una neurona recibe *inputs* de diferentes neuronas pero produce un único *output*, el cual es comunicado a otras. La conexión entre superficies se realiza de diferentes formas, una de ellas, conocida como alimentación hacia adelante (*feed forward*), se tiene cuando las neuronas de la primera superficie envían sus *outputs* a las neuronas de la segunda superficie, sin recibir *inputs* de retorno¹⁷. Adicionalmente, si cada neurona de la primera superficie está conectada con cada neurona de la segunda superficie, se dice que la red está completamente conectada (*fully connected*).

FIGURA 3



¹⁶ Es de señalar que entre la superficie *input* y la superficie *output* pueden existir varias superficies escondidas

¹⁷ Generalizable a cualquier par de superficies continuas, donde, debido al orden de la red, una jugará el papel de la primera y la siguiente de la segunda.

El proceso de aprendizaje del cerebro humano se basa esencialmente en la experiencia. La red neuronal artificial es definida como un algoritmo mecánico de aprendizaje debido a que cambios en las ponderaciones de sus conectores la conducen a aprender la solución para un problema. El peso de las conexiones entre neuronas es almacenado como valor de ponderación de los conectores. Así, el sistema adquiere nuevos conocimientos ajustando el conjunto de ponderaciones asociados a dichos conectores.

La habilidad del aprendizaje de una red neuronal artificial está determinada por su arquitectura o construcción y por el algoritmo seleccionado para su entrenamiento o aprendizaje. El método de entrenamiento puede seguir diferentes esquemas. Uno de éstos es el conocido como de propagación hacia atrás (*back propagation*), donde no sólo se tiene en cuenta la manera como la red realiza su tarea, sino que, la información sobre los errores es filtrada y retornada al sistema para ajustar las conexiones entre las diferentes superficies y de esta forma incrementar el desempeño de la red. Es decir, para conformar un proceso de aprendizaje supervisado.

3.2 Aplicaciones de redes neuronales en el análisis de series de tiempo

Las redes neuronales artificiales (ANN) pueden ser entendidas, en el contexto de análisis de series de tiempo, como modelos no lineales entrenados para (i) realizar conexiones entre los valores pasados y presentes de una serie de tiempo y (ii) extraer estructuras y relaciones escondidas que gobiernan el sistema de información, (véase, Azoff 1996). El aumento en la investigación sobre redes neuronales se debe, primordialmente, a la gran capacidad que éstas muestran en el reconocimiento de patrones de comportamiento¹⁸. Estas han demostrado aproximar casi cualquier función no lineal de una manera acertada¹⁹, motivo por el cual, al ser aplicadas a series de tiempo, pueden detectar las dinámicas no lineales y lograr un mejor ajuste en comparación con modelos lineales (Franses y van Dijk, 2000). Adicionalmente, como lo presenta Tckaz y Hu (1999) las redes neuronales han demostrado tener la capacidad de reconocer y 'modelar' comportamientos atípicos, tales como '*outliers*' o cambios de nivel en las series de tiempo, lo cual puede denotar precisamente, un alejamiento con respecto al mundo lineal.

La red neuronal aquí considerada presenta la forma funcional dada en las ecuaciones (1) y (2).

$$y_t = F(X_t; \Theta) \quad (1)$$

donde

¹⁸ Por ejemplo, al considerar una serie de tiempo que presenta una relación dinámica no lineal, las ANN podrían detectar dicho comportamiento y ajustar la serie de una manera más adecuada que lo que la ajustaría un procedimiento lineal.

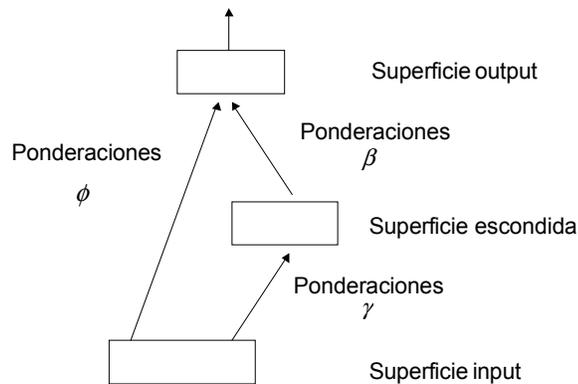
$$F(X_t; \Theta) = \Phi_0 + X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t \quad (2)$$

siendo $Z_t \subseteq X_t$, $X_t = \{y_{t-1}, \dots, y_{t-p}, w_1, \dots, w_m\}$ con y_{t-j} , $j = 1, \dots, p$ rezagos de la variable dependiente, w_j , $j = 1, \dots, m$ variables exógenas y G la función logística, es decir:

$$G(w) = \frac{1}{1 + \exp(-w)} \quad (3)$$

La red neuronal presentada en (2) se nota como ANN(K, K', Q) o red neuronal ampliada²⁰, de tal forma que, $K = p + m$, número de *inputs* en la parte lineal, $K' \leq K$, número de *inputs* en la parte no-lineal y Q número de componentes logísticas. Es de señalar que, el término $X_t' \Phi$ puede ser visto como la representación de la parte lineal de la relación entre y_t y X_t , en tanto que, la componente logística caracteriza la parte no lineal. En la figura 4 se presenta el esquema correspondiente a este modelo.

FIGURA 4

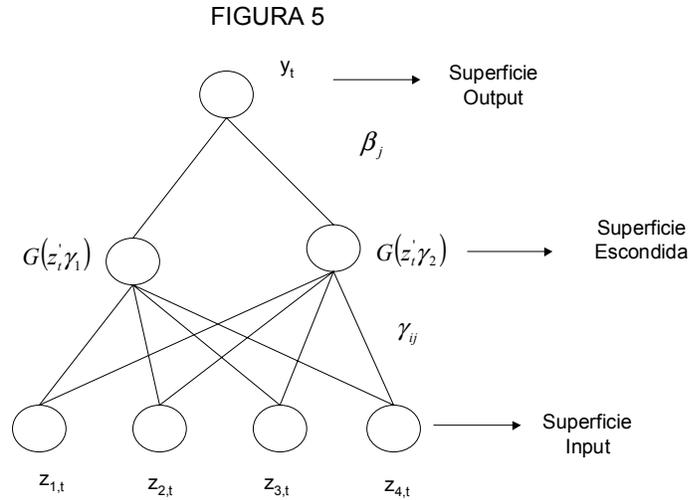


En la figura 5 se muestra un ejemplo gráfico de la componente no-lineal de la ecuación (2), suponiendo $K = K'$ y $Q = 2$, Franses y van Dijk (2000). Este ejemplo permite relacionar la ecuación (2) con los conceptos de redes neuronales artificiales y biológicas dadas anteriormente. En este caso, la red está constituida por tres superficies diferentes. La superficie base o superficie-*input* está conformada por las variables explicatorias, las cuales determinan a Z_t y se conocen como *inputs*. Dichos *inputs* son multiplicados por los valores de ponderación de los conectores, γ_{ij} , los cuales atenúan o amplifican las señales originales para posteriormente entrar a la superficie escondida²¹, donde, en este caso, existen dos unidades escondidas asociadas a funciones logísticas. Es importante resaltar que dichas unidades pueden ser entendidas como variables latentes que enriquecen al modelo lineal. En esta superficie se

¹⁹ Esto se conoce como la propiedad de aproximación universal de las redes neuronales, lo cual no es posible si no se introduce una superficie oculta o *hidden layer*. (Tckaz y Hu, 1999).

²⁰ Esto por considerar tanto parte lineal como parte no lineal.

forman las combinaciones $Z_t' \gamma_j$ ($j=1, \dots, K$) que son transformadas a valores entre cero y uno mediante las funciones de activación²² $G(\cdot)$. Finalmente, éstas son multiplicadas por ponderaciones β_j para dar como resultado el *output* y_t ²³. Dado que esta red presenta una única superficie escondida y que la información va en una única dirección *input-output* se conoce como *single hidden layer feedforward network*.



Como lo señalan Kuan y White (1994), el proceso de aprendizaje de la red neuronal artificial es de carácter secuencial. Así, el aprendizaje es un proceso por el cual la red adquiere conocimiento paso a paso, siendo éste definido como la acumulación de experiencias ocurridas. En las ANN, el conocimiento es adquirido a través de los conectores o parámetros $\Theta = \{\gamma_{ij}, \beta_j, \phi_k\}$. En un esquema general, dado el conocimiento $\hat{\Theta}_t$ en el tiempo (t), el conocimiento $\hat{\Theta}_{t+1}$ en el tiempo (t+1) está conformado por :

$$\hat{\Theta}_{t+1} = \hat{\Theta}_t + \Delta_t \quad (4)$$

donde Δ_t representa un incremento en el conocimiento (aprendizaje). Un procedimiento de aprendizaje satisfactorio debe, por consiguiente, especificar alguna forma apropiada que permita la actualización Δ_t , con base en (i) el conocimiento previo y (ii) los nuevos valores observados

$W_t = (y_t, X_t)$, de tal forma que, se encuentre una función apropiada Ψ_t ($\Delta_t = \Psi_t(W_t, \hat{\Theta}_t)$).

²¹ No relacionada de manera directa con el mundo real.

²² Dichas funciones son asociadas a la forma como la neurona biológica responde ante un estímulo, (véase, Nota de pie de la página 8).

²³ Swanson y White (1995) se refieren a los parámetros γ_{ij} como ponderaciones *input to hidden unit*, mientras que los parámetros β_j son ponderaciones *hidden unit to output*. A su vez, los parámetros Φ de la parte lineal en (2) se conocen como ponderaciones *input to output*.

Formalmente, los conectores o parámetros del modelo presentado en (2) son estimados minimizando la suma de residuales al cuadrado de la siguiente función:

$$S(\Theta) = \sum_{t=1}^n [y_t - F(X_t; \Theta)]^2 \quad (5)$$

donde el vector Θ contiene $(K+1)+Q(K+1)$ parámetros y n corresponde al número de observaciones consideradas. En general, cualquier algoritmo de mínimos cuadrados no-lineales puede ser utilizado para obtener $\hat{\Theta}$. El proceso de estimación se lleva a cabo a través de un esquema recursivo. Así, dados, en la r -ésima iteración, $\hat{\Theta}^{(r)}$ vector de parámetros estimados, $S(\hat{\Theta}^r)$ suma de residuales al cuadrado y $\nabla S(\hat{\Theta}^r)$ gradiente asociado, la estimación $r+1$ -ésima se obtiene a través de la siguiente formulación²⁴:

$$\hat{\Theta}^{r+1} = \hat{\Theta}^r - \lambda A(\hat{\Theta}^r)^{-1} \nabla S(\hat{\Theta}^r) \quad (6)$$

donde, de acuerdo al algoritmo de propagación hacia atrás o *backpropagation*, la matriz $A(\hat{\Theta}^r)$ es igual a una matriz idéntica. Es importante anotar que, bajo este algoritmo particular, diferentes elementos del vector gradiente puedan ser estimados recursivamente²⁵ y, de esta forma, el error cometido en (r) es de vital importancia para la estimación en ($r+1$), hecho que, como ya se había señalado, coincide con el proceso de aprendizaje del cerebro humano en lo referente a la consideración de los errores pasados. Así, como afirman Moshiri y Cameron (1998), en el contexto de las series de tiempo modeladas por medio de las redes neuronales, el aprendizaje o *training* se define como un proceso de iteración y ajuste de parámetros que permite llevar el *output* producido por la red neuronal, en línea con el *output* observado de tal forma que se minimice la función objetivo.

Como lo sugiere Franses y van Dijk (2000), las propiedades numéricas del método de estimación de mínimos cuadrados no-lineales pueden ser mejoradas si: (i) las variables son re-escaladas de tal forma que tengan media cero y desviación estándar unitaria y (ii) si en la función objetivo del proceso de estimación, ecuación (5), se involucra un término de penalización o *weight decay*, de tal forma que se evite la presencia de parámetros muy grandes. Así, la nueva función objetivo es:

²⁴ Quedando especificada la forma funcional Ψ_t . Es de señalar que, tal forma funcional está asociada al algoritmo de optimización utilizado.

²⁵ Por ejemplo, la derivada parcial de $S(\Theta)$ respecto a β_j es igual a: $-2 \sum_{t=1}^n \varepsilon_t G(x_t, \gamma_j)$

$$S(\Theta) = \sum_{t=1}^n [y_t - F(X_t; \Theta)]^2 + r_\phi \sum_{i=0}^{k'} \phi_i^2 + r_\beta \sum_{j=1}^Q \beta_j^2 + r_\gamma \sum_{j=1}^Q \sum_{i=0}^k \gamma_{ij}^2 \quad (7)$$

La implementación de una red neuronal artificial requiere de la selección de cuatro elementos básicos: (i) la función de activación, (ii) el número de unidades escondidas o *hidden units*, (Q), (iii) el número de variables input, (K) y (iv) la determinación de dos conjuntos de información: el primero definido como información dentro de muestra o *training sample*, el cual es utilizado para 'entrenar' a la red y estimar los parámetros; y el segundo, definido como información fuera de muestra o *forecasting sample*, utilizado para evaluar los pronósticos. En general, la función logística es utilizada como función de activación y por consiguiente, dicho punto se pierde como elemento de selección.

Existen diferentes estrategias que pueden ser utilizadas para determinar el número de unidades escondidas y el conjunto de variables input como lo presentan Moody y Utans (1991), Moody (1994, 1995), Levin et al. (1994), Franses y van Dijk (2000) y Swanson y White (Op. Cit. 1995, 1997a). Estos últimos autores proponen, como esquema para definir el conjunto de variables input, la adopción de un estrategia *step-wise* en la componente lineal, de tal forma que, en cada paso, la variable adicionada sea considerada en el conjunto input si ésta mejora los criterios de selección²⁶. Una vez definido el conjunto de variables input de la componente lineal, X_t ecuación (2), se conforma el conjunto Z_t , $Z_t \subseteq X_t$. En primera instancia, el conjunto Z_t incluye tan sólo la primera variable del conjunto input, luego se adiciona a dicho conjunto la segunda variable input de tal forma que en el último paso $Z_t = X_t$. La estimación de la red se lleva a cabo considerando cada uno de estos conjuntos Z_t y diferente número de unidades escondidas, (Q) ($Q = 1, \dots, Q^*$). De acuerdo con Gradojevic y Yang (2000), el número de unidades escondidas Q a elegir presenta un *trade-off*. Demasiadas unidades conllevan a un sobre entrenamiento o sobre ajuste lo cual evita que la red aprenda una solución general; muy pocas unidades ocultas por otro lado, inhiben el aprendizaje del patrón entre el *input* y el *output*. La selección del mejor modelo se lleva a cabo a través de medidas tradicionales de evaluación de pronóstico dentro y fuera de muestra²⁷.

La relevancia de las variables *input* en el modelo seleccionado se lleva a cabo a través de un análisis de sensibilidad, el cual puede realizarse mediante dos enfoques: Franses (2000) y

²⁶ Criterios tales como AIC y BIC.

²⁷ Swanson y White (1995, Op.Cit.) señalan que al existir modelos que resultan mejores de acuerdo a alguna medida de evaluación pero no de acuerdo a otras, se deben mirar todas las medidas y se podría incluso especificar una función de pérdida que asigne ponderaciones a las diferentes medidas. Adicionalmente, un modelo que resulte mejor según las medidas dentro de muestra, no necesariamente es adecuado para pronóstico fuera de muestra por lo cual, al escoger el mejor modelo, deben tenerse en cuenta todas las medidas.

Draisma et al. (1995). En tal análisis se compara la variable y_t observada con la variable output estimada, \hat{y}_t , a través de la red. El esquema de Draisma et al. (1995) evalúa la contribución de una variable input (z_j) sobre el output rompiendo las conexiones de la red con dicha variable²⁸.

Como se mencionó anteriormente, las redes neuronales²⁹ se clasifican dentro de los modelos paramétricos no lineales. Sin embargo, no es común pensar en ellas como procesos generadores de datos. Así, carece de sentido realizar una prueba sobre existencia de no linealidades tipo red neuronal. Sin embargo, a través de las redes neuronales puede llevarse a cabo un test sobre no linealidad general como el desarrollado por Lee et. al. (1993).

Finalmente, es importante señalar algunos problemas asociados a la implementación de las redes neuronales. Por un lado, los parámetros son casi imposibles de interpretar y no brindan información sobre qué tipo de modelo paramétrico no lineal sería adecuado para describir los patrones o dinámicas no lineales identificados. Por esto suele hacerse referencia a las redes neuronales como 'cajas negras' en el sentido de que sirven sólo para identificar patrones en la serie y para pronóstico, pero no permiten interpretar ni dar sentido económico a los parámetros estimados. Adicionalmente, un buen ajuste de la red dentro de muestra (que puede lograrse introduciendo un número suficientemente grande de componentes logísticos no lineales) no garantiza un buen desempeño de la red para pronóstico fuera de muestra; es decir, existe el riesgo de sobre ajuste debido a la alta flexibilidad de estos modelos tal y como se señaló anteriormente.

4. Interpretación económica del *hidden layer* o superficie escondida

En la modelación de variables económicas, la existencia de asimetrías entre el *input* y el *output* pueden dar origen a no linealidades en su relación. Tal y como se vio anteriormente, las componentes logísticas o unidades escondidas caracterizan la parte no lineal de la relación entre el input Z_t y el output y_t .

Esta relación entre Z_t y y_t puede no ser directa, en cuyo caso pueden existir variables intermedias o latentes que se vean afectadas por las variables *input* antes que el *output*³⁰. Tkacz

²⁸ Tal rompimiento se logra imponiendo ceros a los conectores o coeficientes vinculados con la variable en cuestión, tanto en la componente lineal como en la no lineal. Como se sugiere en Draisma et al. (1995), este procedimiento es comparable con el análisis de contribución parcial de términos lineales.

²⁹ Es de señalar que su estructura puede gozar de algún tipo de interpretación

³⁰ Tkacz y Hu (1999), mencionan que para el caso de la relación entre tasas de interés y producto, es de esperarse que variables intermedias afectadas por la tasa de interés tales como la inversión, se vean afectadas antes que el producto. Por ejemplo, una caída en la inversión presente causada por un aumento en las tasas de interés, conlleva a una caída en el producto presente. Sin embargo, esta caída en la inversión actual conduce también a una caída en el producto futuro debido a la pérdida de flujos futuros de

y Hu (1999) señalan que estas variables intermedias, que son precisamente los *hidden layers* o componentes logísticos en la red neuronal, son etapas intermedias de procesamiento donde los *inputs* Z_j , y sus ponderaciones γ_j , son sujetos a otra re-ponderación, antes de afectar a y_t . Al modelar y pronosticar series económicas con redes neuronales, estas variables intermedias no tienen que ser identificadas y pueden tratarse como desconocidas u ocultas³¹.

En el caso de modelos de redes neuronales para inflación, en la sección 2 presentamos evidencia de diferentes trabajos que confirman asimetrías entre la política monetaria y el nivel de precios, lo que justifica la consideración de no linealidades entre variables monetarias e inflación.³² Si bien las variables intermedias o latentes que operan en el *hidden layer* de tales modelos no tienen que ser conocidas, algunas variables que se ven afectadas por la política monetaria antes que la inflación pueden mencionarse especulativamente, constituyéndose en mecanismos de transmisión importantes que pueden estar siendo capturados por el *hidden layer*. Tal y como señala Tkacz (2000), la política monetaria (ya sea a través de instrumentos de tasa de interés o de agregados monetarios) puede afectar el nivel de inversión, lo cual a su vez tiene un efecto sobre el nivel de actividad económica y el producto, que vía una curva de Phillips de corto plazo, afecta el nivel de precios. De esta manera, a pesar de no incluir explícitamente variables como la inversión o la brecha del producto como *inputs* del modelo, estos mecanismos de transmisión pueden estar siendo modelados en el *hidden layer* de manera que su relación y sus dinámicas no lineales con la inflación estén siendo consideradas en la superficie escondida de la red.

5. Resultados

Este trabajo se llevó a cabo con información trimestral correspondiente a la primera diferencia del logaritmo de las series IPC, M1 y M3. El estudio abarca el período comprendido entre el primer trimestre de 1969 y el segundo trimestre del 2001. Con el propósito de mejorar las propiedades de estimación, como se mencionó anteriormente, las variables son re-escaladas en el intervalo (0,1)³³.

ingresos resultante de la inversión en capital que no fue llevada a cabo. El *hidden layer* está capturando estas dinámicas intermedias u ocultas entre la tasa de interés y la inversión.

³¹ Moody, Levin y Refhuss (1993) y Moody (1995) señalan que esta dificultad para identificar y conocer todas las variables que afectan las relaciones económicas, ha conducido en muchos casos a la ausencia de modelos científicos confiables o convincentes, y ha conducido a la utilización de modelos 'caja negra' para pronosticar series económicas que no requieren del conocimiento o de la identificación de todas las variables que operan entre el *input* y el *output*.

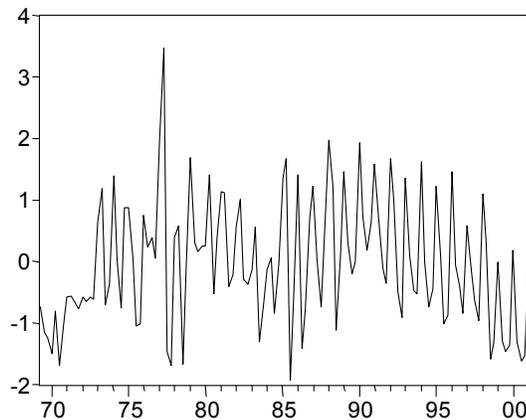
³² Algunos de estos trabajos son Friedman (1968), Cover (1992), Morgan (1993) y Rhee y Rich (1995).

³³ $x_t^* = \frac{x_t - \bar{x}_t}{\sigma(x_t)}$; $x_t = \ln(x_t) - \ln(x_{t-1})$. Es de señalar que la notación "*" se utiliza para referirse a la

transformación de la diferencia del logaritmo de cada serie.

Como *training sample* o período de entrenamiento se consideran dos conjuntos de información diferentes, el primero, desde 1969:2 hasta 1997:4 y el segundo, desde 1969:2 hasta 1999:4. Así, para el primer conjunto, la información para pronóstico fuera de muestra está conformada por 14 observaciones 1998:01-2001:02, en tanto que, para el segundo, se reduce a seis, 2000:01-2001:02. Como se observa en las Gráficas 1 y 2, a partir de 1998 el proceso desinflacionario experimentado por el país explica el comportamiento decreciente de la serie hasta el año 2001. Este hecho justifica la selección de los dos conjuntos de información como *training samples*, ya que esto permite que la red ‘aprenda’ e ‘identifique’ las dinámicas y patrones asociados a dicha caída y que por consiguiente, pueda mejorar su capacidad de pronóstico. Es importante resaltar que la consideración de los dos *training samples* hace parte de una estrategia que trata de tomar en cuenta, por un lado, la dinámica de la caída y, por otro lado, el número de observaciones fuera de muestra.

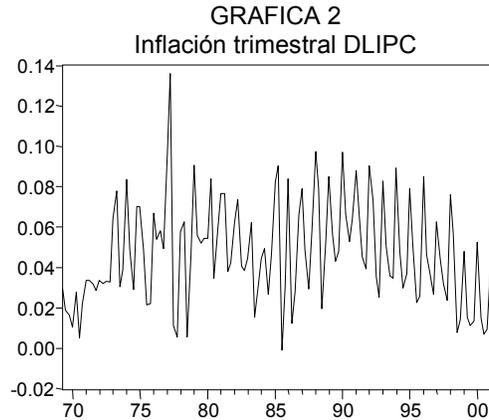
GRAFICA 1
Inflación trimestral DLIPC*



*Transformación de re-escalamiento

La estrategia *stepwise*, propuesta por Swanson y White (1995,1997 a) y descrita anteriormente, se lleva a cabo con el propósito de determinar el conjunto de variables inputs de la componente lineal. Así, se parte de una regresión lineal cuya variable dependiente es $\Delta LIPC_t^*$ y cuyas variables explicativas varían de acuerdo al modelo considerado. En el primer modelo, el conjunto de variables explicativas, o inputs de la componente lineal, considera los primeros ocho rezagos de la inflación, es decir, $\Delta LIPC_{t-1}^*$ hasta $\Delta LIPC_{t-8}^*$. En el segundo y tercer modelo se consideran tan solo variables monetarias como explicativas. El segundo

incluyen los seis primeros rezagos de M1, $\Delta M1_{t-1}^*$ hasta $\Delta M1_{t-6}^*$, en tanto que, en el tercero la variable monetaria considerada es M3. El cuarto y quinto modelo incluyen rezagos de inflación y dinero en su conjunto de información y tienen en cuenta el mismo número de rezagos que se consideró para los tres primeros modelos, sólo que en el cuarto modelo la variable monetaria considerada es M1 mientras que en el quinto es M3³⁴.



El Cuadro 1 resume los resultados de la estrategia *stepwise* para cada modelo y presenta las variables que conforman el conjunto de información X_t , señalado en (2), del cual se selecciona el conjunto de información $Z_t \subseteq X_t$ que conforma la parte no lineal.

CUADRO 1
RESULTADOS ESTRATEGIA *STEPWISE*

$y_t = \Delta LIPC_t^*$ $m_{1t} = \Delta LM1_t^*$ $m_{3t} = \Delta LM3_t^*$		
	<i>Training Sample 1969:2-1997:4</i>	<i>Training Sample 1969:2-1999:4</i>
Modelo 1	$X_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$	$X_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$
Modelo 2	$X_t = \{m_{1t-1}, m_{1t-2}\}$	$X_t = \{m_{1t-1}, m_{1t-2}, m_{1t-5}\}$
Modelo 3	$X_t = \{m_{3t-1}, m_{3t-2}, m_{3t-3}, m_{3t-5}\}$	$X_t = \{m_{3t-1}, m_{3t-2}, m_{3t-3}, m_{3t-5}\}$
Modelo 4	$X_t = \{y_{t-2}, y_{t-8}, m_{1t-1}, m_{1t-2}\}$	$X_t = \{y_{t-1}, y_{t-8}, m_{1t-1}, m_{1t-2}\}$
Modelo 5	$X_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{3t-1}\}$	$X_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{3t-1}, m_{3t-2}\}$

³⁴ El motivo por el cual se consideran los primeros ocho rezagos de la inflación y los primeros seis rezagos del dinero se debe a que diferentes trabajos como Misas et al (1999) y Jalil et al (2000) muestran que el rezago con el cual el dinero afecta la inflación es de año y medio, mientras que el componente autorregresivo de la inflación suele considerar los últimos 2 años.

Una vez seleccionado el conjunto X_t , se lleva a cabo el proceso de estimación de la red neuronal, ecuación (2), para distintas configuraciones del conjunto de información Z_t , $Z_t \subseteq X_t$, y para los dos *training samples* seleccionados. Para cada modelo, se estima la red incluyendo en el conjunto Z_t la primera variable de X_t , y para Q *hidden units* desde 1 hasta 4. Luego se incluye en Z_t , las primeras dos variables de X_t y se estima la red para Q desde 1 hasta 4. El proceso se repite hasta que se estima la red con $Z_t = X_t$ ³⁵.

El método de optimización numérica, utilizado en el proceso de estimación, es el Quasi-Newton de Broyden, Fletcher, Goldfarb y Shano, (BFGS). Como señala Franses y van Dijk (2000), la convergencia en el proceso de estimación no garantiza la obtención del mínimo global³⁶. Por consiguiente, se llevan a cabo múltiples estimaciones de cada una de las diferentes redes utilizando distintos valores iniciales del vector de parámetros $\hat{\Theta}$. En particular, cada red se estima para 30 valores iniciales diferentes del vector de parámetros. Tales valores iniciales son obtenidos aleatoriamente a partir de una distribución uniforme entre [-2,2] y los parámetros del término de *weight decay* en la función objetivo $S(\Theta)$ fueron fijados de la siguiente manera: $r_\phi = 0.01$, $r_\beta = r_\gamma = 0.0001$. La selección de la estimación óptima, en cada modelo, se realiza considerando dos criterios: (i) menor valor de la función objetivo y (ii) vector de gradientes, asociado a los parámetros de la estimación, sin elementos superiores a 1×10^{-3} . Finalmente, para cada red estimada, se calculan medidas de evaluación dentro y fuera de muestra, las cuales permiten elegir la mejor red para cada modelo.³⁷

En los Cuadros 2 y 3 se presentan, para cada conjunto de variables inputs, los modelos o redes neuronales seleccionados, de acuerdo a las medidas de evaluación dentro y fuera de muestra para los períodos comprendidos entre 1969:02 – 1997:04 y 1969:02-1999:04, respectivamente³⁸.

De acuerdo a los resultados³⁹ presentados en el Cuadro 2, el Modelo 4 se constituye como el mejor modelo de redes neuronales⁴⁰ para el período comprendido entre 1969:02 y 1997:04. Este modelo considera: (i) un conjunto de variables *input* en la parte lineal, X_t , igual al de la parte no lineal, Z_t , conformado por el segundo y el octavo rezago de la inflación mensual y por los dos primeros del crecimiento de M1 y (ii) cuatro unidades escondidas Q . Si bien en

³⁵ En las tablas de resultados se incluye el caso $Q=0$, que hace referencia al modelo que solo tiene en cuenta el componente lineal de la red.

³⁶ De hecho, según Franses y van Dijk (2000), la función $S(\Theta)$ posee muchos mínimos locales.

³⁷ Las medidas de evaluación consideradas son las siguientes: AIC, BIC, RMSE, RMSPE, MAE, MAPE y SR dentro de muestra y RMSE, RMSPE, MAE, MAPE, SR y U-THEIL para fuera de muestra. Dichas medidas se describen en el Anexo 2.

³⁸ En el anexo 1 se presentan los resultados de las medidas de evaluación de la red neuronal dentro y fuera de muestra para cada uno de los períodos y modelos bajo estudio.

³⁹ Los diferentes resultados, presentados en este artículo, son obtenidos a través de programación bajo el módulo IML de SAS, versión 8.2.

⁴⁰ Dentro del conjunto de redes neuronales estudiadas en este trabajo.

algunas medidas, el modelo 4 es superado por los otros modelos, una mirada global a todas las medidas demuestra que el modelo 4 resulta ser el mejor según AIC, RMSE, MAE, MAPE (dentro de muestra) y resulta ser el segundo mejor en medidas como BIC, RMSE y MAE (fuera de muestra).

CUADRO 2
 MODELOS SELECCIONADOS
 1969:02 – 1997:04

Modelo General:												
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon \quad , \quad G(z) = (1 + \exp(-z))^{-1}$ $y_t = \Delta LIPC_t^* \quad , \quad m_{1t} = \Delta LM1_t^* \quad , \quad m_{3t} = \Delta LM3_t^*$												
Dentro de Muestra							Fuera de Muestra					
AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
Modelo 1 : Inputs: $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$, $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$, $Q=4$												
-0.949	-0.449	0.516	6.306	0.386	2.076	83.01	1.259	1.390	1.039	1.131	53.84	0.177
Modelo 2 : Inputs: $X_t = \{m_{1t-1}, m_{1t-2}\}$, $Z_t = \{m_{1t-1}, m_{1t-2}\}$, $Q=3$												
-0.884	-0.594	0.578	6.141	0.454	2.117	79.46	1.106	5.671	0.938	3.010	100.0	0.919
Modelo 3 : Inputs: $X_t = \{m_{3t-1}, m_{3t-2}, m_{3t-3}, m_{3t-5}\}$, $Z_t = \{m_{3t-1}, m_{3t-2}, m_{3t-3}\}$, $Q=3$												
-0.649	-0.281	0.631	5.226	0.464	1.952	78.90	0.461	3.202	0.379	1.527	100.0	0.383
Modelo 4 : Inputs: $X_t = \{y_{t-2}, y_{t-8}, m_{1t-1}, m_{1t-2}\}$, $Z_t = \{y_{t-2}, y_{t-8}, m_{1t-1}, m_{1t-2}\}$, $Q=4$												
-1.090	-0.490	0.463	7.416	0.360	1.838	81.13	0.904	3.319	0.812	1.745	100.0	0.424
Modelo 5 : $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{3t-1}\}$, $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{3t-1}\}$, $Q=4$												
-0.286	0.314	0.693	7.400	0.556	2.281	66.98	1.022	1.257	0.904	1.085	53.85	0.161

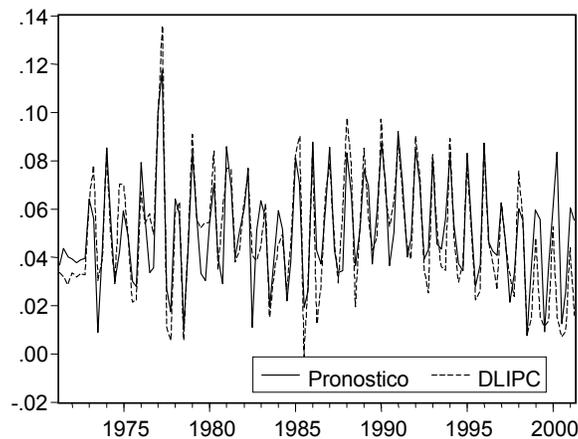
El análisis de los resultados presentados en el Cuadro 3 permiten concluir que el modelo 5 es el mejor modelo de redes neuronales para el período 1969:02-1999:04. Este modelo, al igual que el anterior, presenta el mismo conjunto de variables *input* en la parte lineal como en la no lineal. Dicho conjunto está conformado por los rezagos dos, cuatro y ocho de la inflación mensual y los dos primeros rezagos del crecimiento de M3. En este modelo el número de unidades escondidas es igual a tres.

CUADRO 3
 MODELOS SELECCIONADOS
 1969:02 – 1999:04

Modelo General:												
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon \quad , \quad G(z) = (1 + \exp(-z))^{-1}$ $y_t = \Delta LIPC_t^* , \quad m_{1t} = \Delta LM1_t^* , \quad m_{3t} = \Delta LM3_t^*$												
Dentro de Muestra							Fuera de Muestra					
AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
Modelo 1 : Inputs: $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$, $Z_t = \{y_{t-2}, y_{t-4}\}$, $Q=4$												
-0.815	-0.415	0.572	8.281	0.433	2.190	83.96	0.979	3.007	0.858	1.547	84.61	0.384
Modelo 2 : Inputs: $X_t = \{m_{1t-1}, m_{1t-2}, m_{1t-5}\}$, $Z_t = \{m_{1t-1}, m_{1t-2}\}$, $Q=4$												
-0.813	-0.438	0.582	17.780	0.448	3.444	80.34	1.026	3.353	0.891	2.080	80.00	0.904
Modelo 3 : $X_t = \{m_{3t-1}, m_{3t-2}, m_{3t-3}, m_{3t-5}\}$, $Z_t = \{m_{3t-1}, m_{3t-2}, m_{3t-3}\}$, $Q=4$												
-0.685	-0.216	0.599	13.722	0.446	3.221	78.63	0.405	1.436	0.297	0.807	100.0	0.357
Modelo 4 : Inputs: $X_t = \{y_{t-1}, y_{t-8}, m_{1t-1}, m_{1t-2}\}$, $Z_t = \{y_{t-1}, y_{t-8}, m_{1t-1}, m_{1t-2}\}$, $Q=4$												
-1.141	-0.568	0.458	12.333	0.351	2.965	85.96	1.077	4.493	0.933	2.674	80.00	0.907
Modelo 5 : $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{3t-1}, m_{3t-2}\}$, $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{3t-1}, m_{3t-2}\}$, $Q=3$												
-1.059	-0.558	0.490	11.492	0.375	2.358	85.08	0.492	0.694	0.295	0.511	80.00	0.140

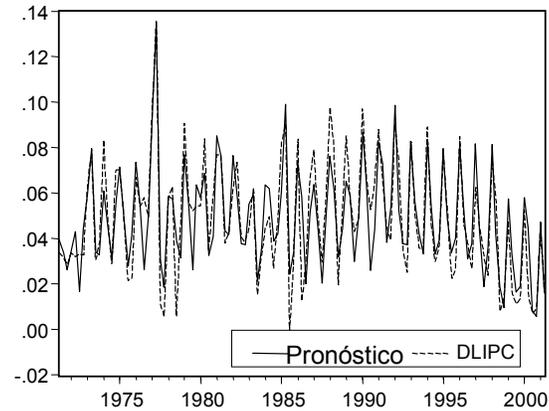
Las Gráficas 3 y 4 presentan la inflación mensual observada y pronosticada a través de la red.

GRÁFICA 3
 Inflación observada y pronóstico bajo el Modelo 4
 (Training sample: 1969:02 – 1997:04)



En la Gráfica 3 la inflación pronosticada proviene del modelo 4 correspondiente al período de entrenamiento o *training sample* 1969:02 – 1997:04. En esta Gráfica se consideran pronósticos fuera de muestra a aquellos reportados para el período 1998:01-2001:02

GRÁFICA 4
Inflación observada y pronóstico bajo el Modelo 5
(*Training sample*: 1969:02 – 1999:04)



En la Gráfica 4 la inflación pronosticada proviene del modelo 5 correspondiente al período de entrenamiento o *training sample* 1969:02 – 1999:04. Los pronósticos fuera de muestra en este caso son los correspondientes al período 2000:01-2001:02

Al comparar los pronósticos fuera de muestra correspondientes al año 2000 y 2001⁴¹, se observa un mejor ajuste de aquellos provenientes del modelo 5, Gráfica 4. Este mejor ajuste se debe a que este modelo considera mayor información, la cual permite incrementar el proceso de aprendizaje de la red neuronal. Esto confirma la importancia de haber considerado un segundo *training sample* que le permita a la red incorporar la información relacionada con la caída en la inflación desde 1998. De hecho, puede verse en la Gráfica 3, que la red seleccionada para el primer *training sample*, la cual no tiene en cuenta la caída de la serie entre 1998 y 2001 y por consiguiente no ‘aprende’ la nueva relación entre dinero e inflación experimentada durante este período desinflacionario, tiende a sobre-pronosticar la inflación.

Los Cuadros 4 y 5 presentan los resultados del análisis de sensibilidad, sugerido por Draisma et. al (1995), correspondientes a los modelos 4 (1969:02 – 1997:04) y 5 (1969:02 – 1999:04), respectivamente. Como se observa, al comparar las medidas de evaluación de pronóstico dentro y fuera de muestra del modelo 4, Cuadro1, con las obtenidas en el análisis de sensibilidad, el modelo presenta su mayor pérdida de capacidad predictiva, dentro y fuera de muestra, al desconectar el primer rezago del crecimiento monetario⁴². Sin embargo, a partir de la

⁴¹ Comunes para los dos períodos de entrenamiento.

⁴² En el Anexo 3 se presenta un resultado similar siguiendo la metodología de análisis de sensibilidad propuesta por Franses y van Dijk (2000).

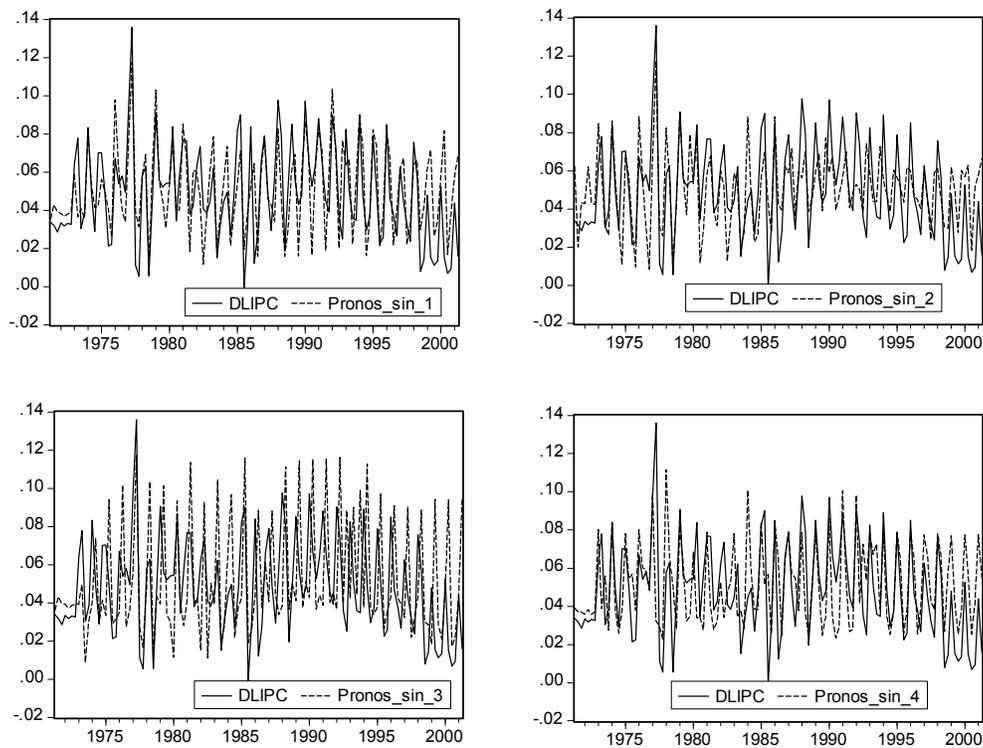
Gráfica 5⁴³ se concluye que en este modelo todas las variables inputs son relevantes para el pronóstico tanto dentro como fuera de muestra.

CUADRO 4

ANÁLISIS DE SENSIBILIDAD SOBRE LA RED NEURONAL SELECCIONADA: MODELO 4
1969:02 – 1997:04

$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon \quad , \quad G(z) = (1 + \exp(-z))^{-1}$														
$y_t = \Delta LIPC_t^*$														
$m_{1t} = \Delta LM1_t^*$														
$\text{Inputs: } X_t = \{y_{t-2}, y_{t-8}, m_{1t-1}, m_{1t-2}\}, Z_t = \{y_{t-2}, y_{t-8}, m_{1t-1}, m_{1t-2}\}$														
$Q=4$														
	Dentro de Muestra							Fuera de Muestra						
Desconexión	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL	
(1) y_{t-2}	-0.274	0.325	0.697	14.472	0.548	3.729	76.4	1.261	3.185	1.013	1.807	69.2	0.407	
(2) y_{t-8}	0.042	0.642	0.816	21.209	0.629	3.993	66.0	1.170	3.076	0.991	1.761	69.2	0.393	
(3) m_{1t-1}	0.919	1.519	1.265	29.477	1.011	8.001	56.6	1.625	6.264	1.212	2.851	46.2	0.800	
(4) m_{1t-2}	0.330	0.930	0.942	6.358	0.674	2.709	68.9	1.130	6.481	1.029	3.189	100	0.828	

GRÁFICA 5

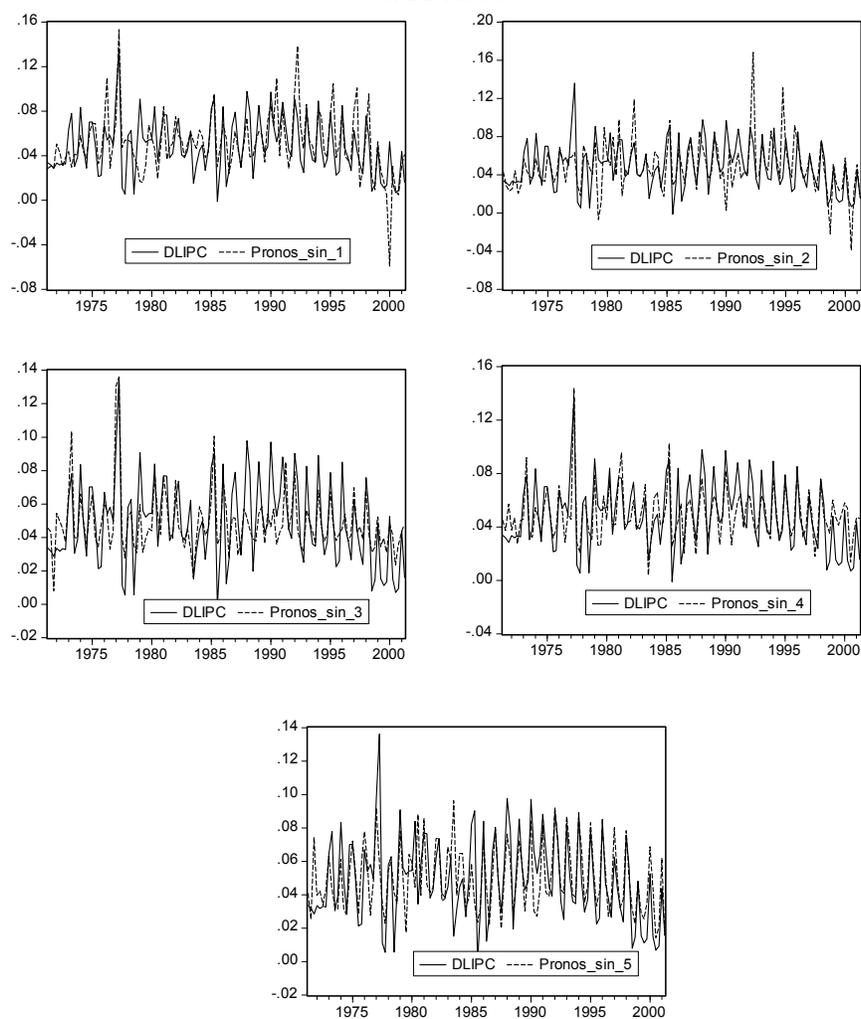


⁴³ La Gráfica presenta la inflación trimestral re-escalada observada y la pronosticada, desconectando cada una de las variables inputs de la red.

CUADRO 5
ANÁLISIS DE SENSIBILIDAD SOBRE LA RED NEURONAL SELECCIONADA: MODELO 5
 1969:02 – 1999:04

$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t, \quad G(z) = (1 + \exp(-z))^{-1}$													
$y_t = \Delta LIPC_t^*$ $m_{3t} = \Delta LM3_t^*$													
Inputs: $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{3t-1}, m_{3t-2}\}$, $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{3t-1}, m_{3t-2}\}$ $Q=3$													
	Dentro de Muestra							Fuera de Muestra					
Desconexión	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
(1) y_{t-2}	0.144	0.645	0.895	16.570	0.653	4.271	69.3	1.872	10.401	1.136	4.986	40.0	2.101
(2) y_{t-4}	0.445	0.946	1.041	6.479	0.692	2.436	70.2	0.819	0.968	0.574	0.792	100	0.195
(3) y_{t-8}	-0.317	0.184	0.711	6.854	0.554	2.112	76.3	0.839	0.741	0.700	0.673	80.0	0.150
(4) m_{3t-1}	-0.479	0.022	0.656	14.458	0.512	2.716	80.7	0.898	0.901	0.717	0.822	100	0.182
(5) m_{3t-2}	-0.133	0.369	0.780	8.154	0.541	2.339	77.2	0.645	2.325	0.537	1.544	100	0.470

GRÁFICA 6



Al comparar las medidas de evaluación de pronóstico dentro y fuera de muestra correspondientes al modelo 5, Cuadro 2, con aquellas obtenidas al llevar a cabo el análisis de sensibilidad, Cuadro 5, se observa que la mayor pérdida de capacidad predictiva se alcanza al desconectar el rezago de orden dos de la inflación. Al igual que en el caso anterior, la Gráfica 6 permite concluir que todas las variables *input* son relevantes tanto para el pronóstico dentro de muestra como para el fuera de muestra.

Dado que la estimación de un modelo de redes neuronales exige la existencia de una relación no lineal entre la variable dependiente y las variables *input*, en nuestro caso inflación y crecimiento monetario, es necesario verificar estadísticamente tal existencia. El Cuadro 6 presenta los resultados del test de Lee et. al. (1993)⁴⁴, basado en redes neuronales⁴⁵, llevado a cabo sobre los dos *training samples*. Para la muestra comprendida entre 1969:02-1997:04 y a un nivel de significancia $\alpha = 10\%$ existe evidencia para rechazar la hipótesis nula de relación lineal entre las variables en cuestión. Al considerar el segundo *training sample* 1969:02-1999:04, se encuentra a un nivel de significancia $\alpha = 5\%$ que existe evidencia para rechazar la hipótesis nula de linealidad. Así, hay evidencia sobre la existencia de una relación no lineal entre la inflación y los crecimientos de M1 y M3 para los períodos de entrenamiento respectivo.

CUADRO 6
DETECCIÓN DE NO-LINEALIDAD
TEST DE LEE, WHITE Y GRANGER

REPLICACIONES	MUESTRA	MODELO SELECCIONADO	P-VALUE TEST LEE, WHITE Y GRANGER
5000	1969:02 – 1997:04	Modelo 4	0.067321
5000	1969:02 – 1999:04	Modelo 5	0.044263

Los Cuadros 7 y 8 presentan las medidas de evaluación de pronóstico dentro y fuera de muestra de la red neuronal seleccionada vs algunos modelos de carácter lineal, Anexo 4, para los períodos 1969:02-1997:04 y 1969:02-1999:04, respectivamente. Es de señalar que, las medidas de evaluación de pronóstico se llevan a cabo sobre la inflación anual en unidades originales (no transformadas).

⁴⁴ Este test, tipo pormanteau, plantea linealidad bajo la hipótesis nula en tanto que, bajo la alterna plantea no linealidad.

CUADRO 7
DESEMPEÑO DE LA RED VERSUS OTROS MODELOS
1969:02 – 1997:04

Modelos	Dentro de Muestra					Fuera de Muestra					
	RMSE	RMSPE	MAE	MAPE	SR	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
Red Neuronal Artificial											
Modelo 4	0.0110	2.516	0.009	0.535	84.38	0.0225	1.715	0.020	1.140	100.0	0.871
Modelos Lineales											
C. Phillips	0.0103	3.658	0.008	0.627	85.94	0.0176	1.208	0.015	0.911	92.31	0.613
P* tradicional	0.0105	1.787	0.009	0.424	87.50	0.0102	0.645	0.008	0.473	92.31	0.327
Escandinavo	0.0113	2.864	0.009	0.546	89.06	0.0161	1.219	0.013	0.884	92.31	0.619
Precios relativos	0.0080	1.789	0.006	0.359	93.75	0.0162	0.814	0.014	0.694	76.92	0.413
ARIMA	0.0148	4.780	0.011	0.878	87.50	0.0249	1.848	0.023	1.526	76.92	0.938

CUADRO 8
DESEMPEÑO DE LA RED VERSUS OTROS MODELOS
1969:02 – 1999:04

Modelos	Dentro de Muestra					Fuera de Muestra					
	RMSE	RMSPE	MAE	MAPE	SR	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
Red Neuronal Artificial											
Modelo 5	0.0125	2.983	0.010	0.625	90.28	0.0125	0.827	0.007	0.470	80.00	0.589
Modelos Lineales											
C. Phillips	0.0111	3.076	0.008	0.596	87.50	0.0058	0.534	0.005	0.332	100.0	0.380
P* tradicional	0.0104	1.532	0.009	0.411	87.50	0.0103	0.521	0.008	0.396	100.0	0.371
Escandinavo	0.0118	2.528	0.009	0.543	88.89	0.0087	0.434	0.007	0.363	100.0	0.309
Precios relativos	0.0085	1.880	0.007	0.402	90.28	0.0102	0.726	0.008	0.509	100.0	0.517
ARIMA	0.0149	4.653	0.011	0.878	87.50	0.0167	1.148	0.015	1.006	100.0	0.817

Los resultados de los Cuadros 7 y 8 permiten concluir que los modelos seleccionados de redes neuronales pueden ser considerados competitivos dentro del conjunto de modelos trimestrales utilizados para predecir la inflación en Colombia. El modelo 4 supera a los modelos Escandinavo y ARIMA en las medidas dentro de muestra y compite con los demás modelos en RMSE, MAE y SR. Para las medidas fuera de muestra, el modelo 4 supera al ARIMA en todas las medidas y supera a todos los modelos lineales en el SR. El modelo 5, elegido para el segundo *training sample*, supera al ARIMA en todas las medidas dentro de muestra, supera también a todos los modelos en el SR dentro de muestra, como también, al modelo C. Phillips en el RMSPE y compite con todos los modelos en el MAE. En las medidas fuera de muestra, el modelo supera al ARIMA en todas las medidas excepto SR, supera a todos los modelos (con excepción del C. Phillips) en MAE y supera al modelo de precios relativos en MAPE.

⁴⁵ En cada training sample se utilizó la red neuronal seleccionada.

6. Conclusiones

Los resultados de este trabajo demuestran el potencial que tienen los modelos de Redes Neuronales en el pronóstico de la inflación, al competir y en algunos casos superar a los modelos lineales tradicionales. Con estos resultados se amplía y mejora la familia de modelos de que hoy se dispone para pronóstico de la inflación y, en particular, se fortalece el trabajo de los modelos que contienen como insumo variables monetarias.

La prueba de no-linealidad, al igual que diferentes argumentos teóricos mencionados en el trabajo, señalan la importancia de modelar la inflación con técnicas no lineales como las redes neuronales. La principal ventaja de esta técnica es que explota la riqueza de su estructura no lineal y la habilidad para aprender en una forma adaptativa del proceso generador de datos subyacente.

Con el fin de mejorar la capacidad de pronóstico de este modelo, algunos elementos adicionales deben considerarse en trabajos futuros como son la ampliación del conjunto de información tanto en observaciones como en variables. Tkacz y Hu (1999) señalan que el número de observaciones utilizadas en la red se vuelve un factor determinante dado: (i) el número de parámetros que deben ser estimados bajo las diferentes especificaciones y (ii) la optimización del proceso de aprendizaje, la cual se logra en la medida en que se cuenta con un mayor número de observaciones en el conjunto de información que posibilita una mejor identificación de los patrones y dinámicas inherentes a la relación estudiada⁴⁶. En este sentido, un ejercicio con datos mensuales y otras variables explicativas adicionales⁴⁷ se llevará a cabo con el fin de mejorar la capacidad predictiva de la red⁴⁸.

⁴⁶ Estos autores recomiendan tener por lo menos 300 observaciones como *training sample* en cuyo caso 30 años de información trimestral pueden resultar insuficientes.

⁴⁷ Aunque, como se señaló en el trabajo, los *hidden layers* pueden estar modelando y capturando el efecto de algunas variables intermedias o mecanismos de transmisión que no se consideraron directamente como *inputs* del modelo

⁴⁸ Tkacz y Hu (1999) señalan que al utilizar series de alta frecuencia, el pronóstico de la red neuronal debe superar ampliamente el pronóstico de los modelos lineales.

BIBLIOGRAFIA

- Altissimo F., E. Gaiotti and A. Locarno (2001), "Monetary Analysis in the Bank of Italy prior to EMU: The Role of Real and Monetary Variables in the Models of the Italian Economy" in *Monetary Analysis: Tools and Applications* Edited by Klöckers H. y Willeke C., European Central Bank.
- Aiken, M. (1996), "A neural network to predict civilian unemployment rates", *Journal of International Information Management*, Vol 5, No.1.
- Arango L.E. y A. González (1998), "Some evidence of smooth transition nonlinearity in Colombian inflation", *Borradores de Economía*, No. 105, Banco de la República.
- Azoff, E. M. (1996), *Neural Network. Time Series Forecasting of Financial Markets*, Wiley, A Wiley Finance Edition.
- Ball, L. And G. Mankiw (1994), "Asymmetric Price Adjustment and Economic Fluctuations", *The Economic Journal*, Vol. 104, No. 423.
- Cover, J. (1992), "Asymmetric Effects of Positive and Negative Money-Supply Shocks", *Quarterly Journal of Economics*, No. 107.
- Draisma, G., J. Kaashoek and H. K. van Dijk (1995), "A neural network applied to economic time series", *Econometric Institute, Erasmus University, Rotterdam and Tinbergen Institute Rotterdam*. S.F.
- Evans, O.V.D.(1997), "Short-Term Currency Forecasting Using Neural Networks" *ICL Systems Journal* 11(2).
- European Central Bank (2001), *The Monetary Policy of the ECB*, European Central Bank.
- Fillion, J.F. and A. Leonard (1997), "La courbe de Phillips au Canada: un examen de quelques hypotheses", *Departement des Recherches, Banque du Canada*.
- Franses P.H. and D. van Dijk (2000), *Non-linear time series models in empirical finance*, Cambridge University Press.
- Friedman, M. (1968), "The Role of Monetary Policy." *American Economic Review* 58: 1-17.
- Friedman B. M. (1984), "The Value of Intermediate Targets in Implementing Monetary Policy" in FRB Kansas City (Ed.), *Price Stability and Public Policy*.
- Gazely A. and J. Binner (2000), "The Application of Neural Networks to the Divisia Index Debate: Evidence from three Countries", *Applied Economics*, Vol. 10 No.12.
- Gerlach, S. and Svensson L. (2000), "Money and Inflation in the Euro Area: A case for Monetary Indicators ?" Working Paper No.8025, National Bureau of Economic Research.
- Gradojevic, N. and J. Yang (2000), "The Application of Artificial Neural Networks to Exchange Rate Forecasting: The Role of Market Microstructure Variables", Working paper 2000-23, Bank of Canada.
- Granger, C.W.J. and T. Terasvirta (1993), *Modelling Nonlinear Economic Relationships*, Advanced Texts in Econometrics, Oxford University Press.
- Hallman, J., Porter R. and Small D. (1989), "M per Unit of Potential GNP as an anchor for the Price Level" Staff Study 157, Board of Governors of the Federal Reserve System.

----- (1991), "Is the Price Level Tied to the M2 Monetary Aggregate in the Long Run" *American Economic Review* Vol.81 No.4.

Hauser, A.(2001), "Money and Credit in an Inflation-Targeting Regime: The Bank of England's Quarterly Monetary Assessment" in *Monetary Analysis:Tools and Applications* Edited by Klöckers H. y Willeke C., European Central Bank.

Herbrich,R., M. Keilbach, Graepel, T., Bollmann, P. and K. Obermayer (1999), "Neural Networks in Economics: Background, Applications and New Developments", Technische Universität Berlin.

Jalil, M. y L.F. Melo (2000), "Una relación no lineal entre la inflación y los medios de pago", *Borradores de Economía*, No. 145, Banco de la República

Klerfors, D. (1998), "Artificial Neural Networks: What are they?, How do they work?, In what areas are they used?", Saint Louis University, School of Business & Administration.

Kuan C. M. and H. White (1994), "Artificial Neural Networks: An Econometric Perspective", *Econometric Reviews* 13.

Laidler, D.(1993), *The Demand for Money: Theories, Evidence and Problems*, Harper Collins, New York, NY.

Lee, T.H., H. White and C. W.J. Granger (1993), "Testing for neglected nonlinearity in time series models", *Journal of Econometrics*, No. 56.

Levin, A., T. Leen and J. Moody (1994), "Fast Pruning Using Principal Components", *Advances in Neural Information Processing* 6, J. Cowan, G. Tesauro and J. Alspector, eds., Morgan Kaufmann, San Mateo, C.A.

López, E. y M. Misas (1998), "Un examen empírico de la curva de Phillips en Colombia" *Ensayos de Política Económica* No. 34, Banco de la República.

Lütkepohl,H., T. Teräsvirta and J. Wolters (1999), "Investigating Stability and Linearity of a German M1 Money Demand Function", *Journal of Applied Econometrics* Vol14 No.5.

Makridakis, S. and S. Wheelwright (1989), *Manual de Técnicas de Pronóstico*, Editorial Limusa.

Masuch, K. , H.Pill and C. Willeke (2001), "Framework and Tools of Monetary Analysis" in *Monetary Analysis:Tools and Applications*, Edited by Klöckers H. y Willeke C., European Central Bank.

Medeiros, M., A. Veiga, and C. Pedreira (2001), "Modelling Exchange Rates: Smooth Transitions, Neural Networks, and Linear Models", Dept. of Economics, Pontifical Catholic University of Rio de Janeiro.

Misas M., E. López y L.F. Melo (1999), "La Inflación desde una Perspectiva Monetaria: un modelo P* para Colombia" *Revista Ensayos de Política Económica* No. 35, Banco de la República.

Moody, J. (1994), "Prediction Risk and Architecture Selection for Neural Networks", *Statistics to Neural Networks: Theory and Pattern Recognition Applications*, NATO ASI Series F, Springer-Verlag.

----- (1995), "Economic Forecasting: Challenges and Neural Network Solutions", *International Symposium on Artificial Neural Networks*, Taiwan.

-----, U. Levin and S. Rehfuss (1993), "Predicting the U.S. Index of Industrial Production" Appears in *Proceedings of Parallel Applications in Statistic and Economics*, Special Issue of neural Network World, Vol. 3 No.6.

----- and J. Utans (1991), "Principled Architecture Selection for Neural Networks: Application to Corporate Bond Rating Prediction", *Advances in Neural Information Processing Systems* 4, Morgan Kaufmann Publishers, San Mateo, CA.

Morgan, D.P. (1993), "Asymmetric Effects of Monetary Policy", *Federal Reserve Bank of Kansas City, Economic Review*, No. 78.

Moshiri, S. and N. Cameron (1998), "Neural Networks vs. Econometric Models in Forecasting Inflation", Department of Economics, University of Manitoba.

Nicoletti Altimari S. (2001), "Does Money Lead Inflation in the Euro Area" WP No.63, Working Papers Series . European Central Bank.

Orphanides A. and R. Porter (2001), "Money and Inflation the Role of Information Regarding the Determination of M2 Behaviour" in *Monetary Analysis: Tools and Applications*, Edited by Klöckers H. y Willeke C., European Central Bank.

Pill H. (2001) "Monetary Analysis: Tools and Applications" in *Monetary Analysis: Tools and Applications*, Edited by Klöckers H. y Willeke C., European Central Bank.

Plummer, E. (2000), "Time Series Forecasting with Feed-Forward Neural Networks: Guidelines and Limitations", Department of Computer Science, Graduate School of The University of Wyoming.

Rhee, W. and R.W. Rich (1995), "Term Structure Forecasts of Inflation", *Economic Journal*, No. 102.

Ricketts and Rose (1995), "Inflation, Learning and Monetary Policy in the G7 economies" Bank of Canada, Working Paper 95-7.

Pindyck, R. and D. Rubinfeld (1997), *Econometric Models and Economic Forecasts*, McGraw-Hill International Editions, Fourth Edition.

Soto R. (1997), "Non-linearities in the Demand for Money: A Neural Network Approach." Documento de Investigación No.107, Ilades/Georgetown University.

Swanson, N. R. and H. White (1995), "A Model-Selection Approach to Assessing the Information in the Term Structure Using Linear Models and Artificial Neural Networks", *Journal of Business & Economic Statistics*, Vol. 13, No.3.

----- (1997a), "A model Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks", *The Review of Economics and Statistics*, No. 79.

----- (1997 b), "Forecasting economic time series using flexible versus fixed specification and linear versus nonlinear econometric models", *International Journal of Forecasting*, No. 13.

Svensson, Lars (1999), "Monetary Policy Issues for the Eurosystem" *Carnegie-Rochester Conferences Series on Public Policy* 71, 79-136, marzo.

Tkacz. G. (2000), "Non-Parametric and Neural Networks Models of Inflation Changes", Working Paper 2000-7, Bank of Canada.

Tkacz, G. and S. Hu (1999), "Forecasting GDP Growth Using Artificial Neural Networks", Working Paper 99-3, Bank of Canada.

ANEXO 1
CUADRO 1
EVALUACIÓN DE LA RED NEURONAL DENTRO DE MUESTRA: MODELO 1
1969:02 – 1997:04

Modelo General:								
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$								
$y_t = \Delta LIPC_t^*$ $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$								
Inputs	Q	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR
X_t	0	-0.5207	-0.4208	0.7425	8.8670	0.5547	2.3881	80.1887
X_t $Z_t = \{y_{t-2}\}$	1	-0.5601	-0.4852	0.7348	8.0033	0.5489	2.4358	80.1887
	2	-0.5289	-0.3790	0.7258	7.8699	0.5401	2.3363	81.1321
	3	-0.4896	-0.2647	0.7197	9.2940	0.5303	2.4657	83.9623
	4	-0.4424	-0.1426	0.7165	10.8654	0.5253	2.7302	83.9623
X_t $Z_t = \{y_{t-2}, y_{t-4}\}$	1	-0.6142	-0.5143	0.7086	10.9761	0.5370	3.1547	79.2453
	2	-0.5891	-0.3892	0.6912	9.2798	0.5284	2.8439	82.0755
	3	-0.6018	-0.3021	0.6616	11.8206	0.5029	3.3014	80.1887
	4	-0.8152	-0.4155	0.5729	8.2815	0.4337	2.1901	83.9623
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$	1	-0.6052	-0.4803	0.7051	11.7873	0.5293	2.8649	81.1321
	2	-0.7547	-0.5049	0.6245	7.5984	0.4845	2.0646	81.1321
	3	-0.8857	-0.5110	0.5582	9.7283	0.4324	2.8461	81.1321
	4	-0.9491	-0.4495	0.5161	6.3064	0.3863	2.0763	83.0189

CUADRO 1A
EVALUACIÓN DE LA RED NEURONAL FUERA DE MUESTRA: MODELO 1
1969:02 – 1997:04

Modelo General:							
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$							
$y_t = \Delta LIPC_t^*$ $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$							
Inputs	Q	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
X_t	0	1.1193	3.4192	0.9986	1.9075	84.6154	0.4368
X_t $Z_t = \{y_{t-2}\}$	1	1.2038	3.3321	1.0883	1.9663	84.6154	0.4256
	2	1.1918	3.4899	1.0841	2.0428	100.0000	0.4458
	3	1.1729	3.5078	1.0475	1.9601	100.0000	0.4481
	4	1.0376	4.2706	0.9565	2.2729	100.0000	0.5455
X_t $Z_t = \{y_{t-2}, y_{t-4}\}$	1	1.1057	4.1373	0.9473	2.2236	76.9231	0.5285
	2	1.0914	2.6964	0.9599	1.6207	76.9231	0.3444
	3	1.6855	3.1905	1.3024	2.0149	38.4615	0.4075
	4	0.9790	3.0070	0.8586	1.5474	84.6154	0.3841
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$	1	1.1372	4.5496	0.9571	2.4278	76.9231	0.5811
	2	1.0808	2.7868	0.9560	1.6487	92.3077	0.3560
	3	1.1544	3.6217	1.0362	2.0098	76.9231	0.4626
	4	1.2595	1.3902	1.0392	1.1312	53.8462	0.1776

CUADRO 2
EVALUACIÓN DE LA RED NEURONAL DENTRO DE MUESTRA: MODELO 1
1969:02 – 1999:04

Modelo General:								
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$								
$y_t = \Delta LIPC_t^*$ $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$								
Inputs	Q	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR
X_t	0	-0.5079	-0.4124	0.7492	25.0201	0.5561	4.1579	81.5789
X_t $Z_t = \{y_{t-2}\}$	1	-0.5530	-0.4814	0.7389	18.0961	0.5503	3.6112	79.8246
	2	-0.5341	-0.3909	0.7267	15.8042	0.5378	3.4644	81.5789
	3	-0.5168	-0.3020	0.7141	17.0295	0.5261	3.6196	84.2105
	4	-0.4789	-0.1925	0.7091	15.6497	0.5163	3.4573	84.2105
X_t $Z_t = \{y_{t-2}, y_{t-4}\}$	1	-0.6319	-0.5364	0.7042	21.1881	0.5359	4.5703	78.9474
	2	-0.5837	-0.3928	0.6967	18.7814	0.5237	4.3199	80.7018
	3	-0.7488	-0.4623	0.6196	24.6474	0.4736	3.8286	84.2105
	4	-0.8149	-0.4330	0.5789	22.6112	0.4507	3.7354	81.5789
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$	1	-0.6205	-0.5011	0.7021	20.4006	0.5274	4.2705	83.3333
	2	-0.8141	-0.5754	0.6102	19.4882	0.4614	3.3204	78.0702
	3	-0.6896	-0.3315	0.6217	15.2846	0.4783	3.8174	83.3333
	4	-0.8834	-0.4060	0.5403	6.5983	0.4182	1.9970	85.9649

CUADRO 2A
EVALUACIÓN DE LA RED NEURONAL FUERA DE MUESTRA: MODELO 1
1969:02 – 1999:04

Modelo General:							
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$							
$y_t = \Delta LIPC_t^*$ $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$							
Inputs	Q	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
X_t	0	0.8252	1.8791	0.7751	1.4616	100.0000	0.3795
X_t $Z_t = \{y_{t-2}\}$	1	0.7567	1.1981	0.6947	1.0053	100.0000	0.2420
	2	0.7226	1.2408	0.6543	0.9207	80.0000	0.2506
	3	0.7515	1.4800	0.6851	1.0291	80.0000	0.2989
	4	0.7286	1.4044	0.6638	0.9891	100.0000	0.2836
X_t $Z_t = \{y_{t-2}, y_{t-4}\}$	1	0.7291	1.9929	0.6870	1.4754	100.0000	0.4025
	2	0.6698	1.6579	0.6173	1.2408	80.0000	0.3348
	3	0.7541	1.1050	0.6813	0.9756	100.0000	0.2232
	4	0.6439	1.0220	0.5900	0.8854	100.0000	0.2064
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$	1	0.6746	2.0899	0.6226	1.4874	100.0000	0.4221
	2	0.9689	1.9668	0.8362	1.5462	100.0000	0.3972
	3	0.9308	3.7951	0.7758	1.9457	80.0000	0.7665
	4	1.0971	2.1748	0.9661	1.6824	80.0000	0.4392

CUADRO 3
EVALUACIÓN DE LA RED NEURONAL DENTRO DE MUESTRA: MODELO 2
1969:02 – 1997:04

Modelo General:								
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$								
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM1_t^*$ $X_t = \{m_{t-1}, m_{t-2}\}$								
Inputs	Q	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR
X_t	0	-0.6035	-0.5318	0.7137	4.5009	0.5415	1.8800	77.0642
X_t $Z_t = \{m_{t-1}\}$	1	-0.5852	-0.5127	0.7268	6.5148	0.5399	2.1174	78.5714
	2	-0.5734	-0.4286	0.7119	7.5537	0.5361	2.2302	77.6786
	3	-0.5511	-0.3338	0.7010	8.5148	0.5160	2.3097	78.5714
	4	-0.5228	-0.2332	0.6924	7.8720	0.5092	2.2169	77.6786
X_t $Z_t = \{m_{t-1}, m_{t-2}\}$	1	-0.7349	-0.6383	0.6684	6.7338	0.5206	2.1327	76.7857
	2	-0.7990	-0.6059	0.6248	7.4446	0.4893	2.2671	78.5714
	3	-0.8840	-0.5944	0.5780	6.1407	0.4544	2.1174	79.4643
	4	-0.9034	-0.5172	0.5525	7.3827	0.4427	2.2256	79.4643

CUADRO 3A
EVALUACIÓN DE LA RED NEURONAL FUERA DE MUESTRA: MODELO 2
1969:02 – 1997:04

Modelo General:							
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$							
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM1_t^*$ $X_t = \{m_{t-1}, m_{t-2}\}$							
Inputs	Q	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
X_t	0	1.1159	4.9296	0.9289	2.8071	100.0000	0.6297
X_t $Z_t = \{m_{t-1}\}$	1	1.1698	5.7603	0.9573	2.9766	100.0000	0.9718
	2	1.1625	5.8963	0.9419	3.0069	100.0000	0.9657
	3	1.1788	6.2891	0.9615	3.4235	100.0000	0.9792
	4	1.1854	6.3972	0.9955	3.1691	100.0000	0.9847
X_t $Z_t = \{m_{t-1}, m_{t-2}\}$	1	1.1482	5.7500	0.9718	2.8953	100.0000	0.9538
	2	1.7271	5.2579	1.2820	2.9947	92.3077	1.4347
	3	1.1064	5.6706	0.9375	3.0097	100.0000	0.9191
	4	1.4015	11.5745	1.1693	4.7926	92.3077	1.1643

CUADRO 4
EVALUACIÓN DE LA RED NEURONAL DENTRO DE MUESTRA: MODELO 2
1969:02 – 1999:04

Modelo General:								
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$								
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM1_t^*$ $X_t = \{m_{t-1}, m_{t-2}, m_{t-5}\}$								
Inputs	Q	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR
X_t	0	-0.6462	-0.5757	0.7057	4.3544	0.5389	1.8465	78.6325
X_t $Z_t = \{m_{t-1}\}$	1	-0.6918	-0.6213	0.6898	28.5850	0.5182	4.2404	79.4872
	2	-0.6691	-0.5282	0.6802	27.1352	0.5137	4.0869	77.7778
	3	-0.6481	-0.4368	0.6701	28.5232	0.5014	4.4355	77.7778
	4	-0.6135	-0.3318	0.6647	28.0877	0.4859	4.2442	78.6325
X_t $Z_t = \{m_{t-1}, m_{t-2}\}$	1	-0.7774	-0.6834	0.6554	16.1416	0.4872	3.6844	79.4872
	2	-0.8407	-0.6529	0.6138	27.1769	0.4771	4.0232	80.3419
	3	-0.8715	-0.5898	0.5842	15.9586	0.4486	3.5025	80.3419
	4	-0.8132	-0.4375	0.5815	17.7798	0.4477	3.4436	80.3419
X_t $Z_t = \{m_{t-1}, m_{t-2}, m_{t-5}\}$	1	-0.7682	-0.6508	0.6528	28.3387	0.4926	4.1935	80.3419
	2	-0.8511	-0.6163	0.6003	22.5813	0.4611	3.6086	79.4872
	3	-0.9343	-0.5821	0.5520	27.2421	0.4387	3.9898	80.3419
	4	-0.9750	-0.5054	0.5184	6.8500	0.3881	1.8350	78.6325

CUADRO 4A
EVALUACIÓN DE LA RED NEURONAL FUERA DE MUESTRA: MODELO 2
1969:02 – 1999:04

Modelo General:							
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$							
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM1_t^*$ $X_t = \{m_{t-1}, m_{t-2}, m_{t-5}\}$							
Inputs	Q	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
X_t	0	1.4427	5.6104	1.3969	3.7558	100.0000	1.1331
X_t $Z_t = \{m_{t-1}\}$	1	1.4814	5.5623	1.4311	3.7559	100.0000	1.3050
	2	1.4385	5.4389	1.3919	3.6699	100.0000	1.2672
	3	1.5021	5.5483	1.4401	3.7618	100.0000	1.3232
	4	1.5128	5.7463	1.4174	3.8567	100.0000	1.3326
X_t $Z_t = \{m_{t-1}, m_{t-2}\}$	1	1.4584	5.6363	1.4063	3.7589	100.0000	1.2847
	2	1.4300	5.2416	1.3936	3.5893	100.0000	1.2597
	3	1.4505	5.5288	1.4125	3.7376	100.0000	1.2777
	4	1.0260	3.3534	0.8914	2.0796	80.0000	0.9038
X_t $Z_t = \{m_{t-1}, m_{t-2}, m_{t-5}\}$	1	1.3975	4.7008	1.3492	3.2807	100.0000	1.2310
	2	1.2274	4.6343	1.0823	3.0558	80.0000	1.0812
	3	1.3699	5.5559	1.3187	3.6384	80.0000	1.2067
	4	1.2237	5.6798	1.1682	3.6546	80	1.0780

CUADRO 5
EVALUACIÓN DE LA RED NEURONAL DENTRO DE MUESTRA: MODELO 3
1969:02 – 1997:04

Modelo General:								
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$								
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM3_t^*$ $X_t = \{m_{t-1}, m_{t-2}, m_{t-3}, m_{t-5}\}$								
Inputs	Q	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR
X_t	0	-0.2935	-0.1953	0.8327	7.6428	0.6258	2.1332	65.1376
X_t $Z_t = \{m_{t-1}\}$	1	-0.3729	-0.2992	0.8076	5.7700	0.6025	1.9319	74.3119
	2	-0.3839	-0.2366	0.7815	5.8149	0.5781	1.8812	71.5596
	3	-0.3320	-0.1110	0.7805	6.5083	0.5770	1.9582	69.7248
	4	-0.3482	-0.0536	0.7534	5.9897	0.5547	1.9391	73.3945
X_t $Z_t = \{m_{t-1}, m_{t-2}\}$	1	-0.3552	-0.2570	0.8074	6.2892	0.6069	2.1164	71.5596
	2	-0.4267	-0.2303	0.7512	5.9954	0.5675	1.8423	73.3945
	3	-0.5189	-0.2243	0.6917	5.2377	0.5158	1.7438	77.0642
	4	-0.5061	-0.1133	0.6713	6.2590	0.4983	1.9671	77.0642
X_t $Z_t = \{m_{t-1}, m_{t-2}, m_{t-3}\}$	1	-0.4589	-0.3361	0.7597	6.5113	0.5866	2.2672	74.3119
	2	-0.5113	-0.2658	0.7071	3.2931	0.5123	1.4464	76.1468
	3	-0.6490	-0.2807	0.6307	5.2259	0.4636	1.9516	78.8991
	4	-0.5882	-0.0972	0.6213	7.3041	0.4562	2.3449	76.1468
X_t $Z_t = \{m_{t-1}, m_{t-2}, m_{t-3}, m_{t-5}\}$	1	-0.4978	-0.3505	0.7383	7.6244	0.5764	2.5946	74.3119
	2	-0.4988	-0.2042	0.6987	4.5141	0.5553	1.9126	74.3119
	3	-0.5176	-0.0757	0.6555	2.4067	0.4873	1.2163	77.9817
	4	-0.5986	-0.0094	0.5960	11.5622	0.4529	2.4795	73.3945

CUADRO 5A
EVALUACIÓN DE LA RED NEURONAL FUERA DE MUESTRA: MODELO 3
1969:02 – 1997:04

Modelo General:							
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$							
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM3_t^*$ $X_t = \{m_{t-1}, m_{t-2}, m_{t-3}, m_{t-5}\}$							
Inputs	Q	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
X_t	0	0.5558	1.8394	0.5071	1.0988	92.3077	0.2350
X_t $Z_t = \{m_{t-1}\}$	1	0.7160	2.2485	0.6427	1.3158	84.6154	0.5947
	2	0.6671	2.3326	0.6079	1.3263	84.6154	0.5542
	3	0.6620	2.1902	0.5989	1.2648	84.6154	0.5499
	4	0.7573	2.0390	0.6681	1.2309	84.6154	0.6291
X_t $Z_t = \{m_{t-1}, m_{t-2}\}$	1	0.6215	2.5240	0.5535	1.2438	84.6154	0.5163
	2	0.5524	3.3601	0.4950	1.6386	84.6154	0.4589
	3	0.7341	2.3146	0.6571	1.3441	76.9231	0.6098
	4	1.7368	3.7251	0.9209	1.9896	76.9231	1.4428
X_t $Z_t = \{m_{t-1}, m_{t-2}, m_{t-3}\}$	1	0.5893	5.1514	0.4442	2.2856	84.6154	0.4896
	2	0.6206	1.1931	0.5259	0.8590	84.6154	0.5155
	3	0.4607	3.2021	0.3794	1.5265	100.0000	0.3827
	4	0.5417	3.4494	0.4241	1.7065	76.9231	0.4500
X_t $Z_t = \{m_{t-1}, m_{t-2}, m_{t-3}, m_{t-5}\}$	1	0.6167	4.9076	0.4787	2.2558	76.9231	0.5123
	2	0.6240	4.6766	0.5074	2.1795	100.0000	0.5183
	3	0.8668	8.4975	0.7987	3.2762	61.5385	0.7201
	4	0.5588	2.4812	0.4699	1.1735	76.9231	0.4642

CUADRO 6
EVALUACIÓN DE LA RED NEURONAL DENTRO DE MUESTRA: MODELO 3
1969:02 – 1999:04

Modelo General:								
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$								
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM3_t^*$ $X_t = \{m_{t-1}, m_{t-2}, m_{t-3}, m_{t-5}\}$								
Inputs	Q	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR
X_t	0	-0.371	-0.277	0.803	12.968	0.607	2.989	66.67
X_t $Z_t = \{m_{t-1}\}$	1	-0.435	-0.364	0.784	14.596	0.590	3.021	72.65
	2	-0.453	-0.312	0.758	15.545	0.566	3.115	71.79
	3	-0.413	-0.201	0.754	13.129	0.562	2.789	70.94
	4	-0.362	-0.080	0.754	13.734	0.560	2.869	72.65
X_t $Z_t = \{m_{t-1}, m_{t-2}\}$	1	-0.441	-0.347	0.775	6.799	0.583	2.083	71.79
	2	-0.511	-0.323	0.724	19.887	0.545	3.213	72.65
	3	-0.517	-0.235	0.698	21.300	0.528	3.309	77.78
	4	-0.467	-0.091	0.691	6.031	0.519	1.879	72.65
X_t $Z_t = \{m_{t-1}, m_{t-2}, m_{t-3}\}$	1	-0.548	-0.430	0.729	26.279	0.555	4.286	74.36
	2	-0.526	-0.291	0.706	26.579	0.532	4.170	76.07
	3	-0.677	-0.325	0.628	19.341	0.470	3.015	75.21
	4	-0.685	-0.216	0.599	13.722	0.446	3.221	78.63
X_t $Z_t = \{m_{t-1}, m_{t-2}, m_{t-3}, m_{t-5}\}$	1	-0.535	-0.394	0.727	22.039	0.565	3.837	70.09
	2	-0.582	-0.300	0.675	4.688	0.500	1.638	72.65
	3	-0.616	-0.193	0.631	4.098	0.472	1.773	74.36
	4	-0.748	-0.185	0.561	11.980	0.430	3.050	80.34

CUADRO 6A
EVALUACIÓN DE LA RED NEURONAL FUERA DE MUESTRA: MODELO 3
1969:02 – 1999:04

Modelo General:							
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$							
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM3_t^*$ $X_t = \{m_{t-1}, m_{t-2}, m_{t-3}, m_{t-5}\}$							
Inputs	Q	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
X_t	0	0.400	1.820	0.321	0.966	100.00	0.368
X_t $Z_t = \{m_{t-1}\}$	1	0.554	2.257	0.484	1.355	80.00	0.488
	2	0.513	2.161	0.439	1.274	100.00	0.452
	3	0.370	1.900	0.254	0.989	100.00	0.326
	4	3.209	3.132	1.602	2.055	80.00	2.827
X_t $Z_t = \{m_{t-1}, m_{t-2}\}$	1	0.416	1.160	0.396	0.844	100.00	0.366
	2	0.578	2.923	0.484	1.698	80.00	0.509
	3	0.588	2.904	0.510	1.713	80.00	0.518
	4	1.434	8.091	0.838	3.819	80.00	1.263
X_t $Z_t = \{m_{t-1}, m_{t-2}, m_{t-3}\}$	1	0.733	3.615	0.615	2.175	100.00	0.646
	2	0.690	3.721	0.523	2.178	80.00	0.608
	3	0.555	2.122	0.527	1.362	100.00	0.489
	4	0.405	1.436	0.297	0.807	100.00	0.357
X_t $Z_t = \{m_{t-1}, m_{t-2}, m_{t-3}, m_{t-5}\}$	1	0.869	4.062	0.772	2.498	100.00	0.765
	2	0.786	4.049	0.647	2.498	60.00	0.693
	3	0.655	2.713	0.604	1.746	60.00	0.577
	4	0.825	4.421	0.658	2.657	60.00	0.727

CUADRO 7
EVALUACIÓN DE LA RED NEURONAL DENTRO DE MUESTRA: MODELO 4
1969:02 – 1997:04

Modelo General:								
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$								
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM1_t^*$ $X_t = \{y_{t-2}, y_{t-8}, m_{t-1}, m_{t-2}\}$								
Inputs	Q	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR
X_t	0	-0.6776	-0.5777	0.6865	5.6872	0.5020	2.1233	80.1887
X_t $Z_t = \{y_{t-2}\}$	1	-0.7337	-0.6588	0.6738	5.9158	0.4886	2.1419	79.2453
	2	-0.7205	-0.5706	0.6595	4.5781	0.4866	1.9368	83.0189
	3	-0.6834	-0.4586	0.6532	5.4526	0.4869	2.1059	82.0755
	4	-0.6596	-0.3598	0.6428	4.6836	0.4616	1.8531	84.9057
X_t $Z_t = \{y_{t-2}, y_{t-8}\}$	1	-0.7462	-0.6463	0.6633	5.4814	0.4898	2.0322	79.2453
	2	-0.7070	-0.5071	0.6517	5.0829	0.4706	1.9039	79.2453
	3	-0.8037	-0.5039	0.5981	7.2091	0.4381	2.3897	83.9623
	4	-0.6886	-0.3888	0.6335	5.8688	0.4737	1.9358	83.0189
X_t $Z_t = \{y_{t-2}, y_{t-8}, m_{t-1}\}$	1	-0.7505	-0.6256	0.6557	5.5065	0.4703	2.1416	81.1321
	2	-0.7645	-0.5147	0.6215	6.3598	0.4483	2.1867	84.9057
	3	-0.8439	-0.4693	0.5700	11.3043	0.4288	2.9172	82.0755
	4	-1.0027	-0.5031	0.5024	2.2839	0.3795	1.1833	84.9057
X_t $Z_t = \{y_{t-2}, y_{t-8}, m_{t-1}, m_{t-2}\}$	1	-0.7986	-0.6488	0.6342	4.1588	0.4589	1.7491	80.1887
	2	-0.9049	-0.6051	0.5686	7.1340	0.4359	2.0057	80.1887
	3	-0.9487	-0.4991	0.5259	9.8449	0.3964	2.6259	81.1321
	4	-1.0897	-0.4902	0.4634	7.4156	0.3596	1.8377	81.1321

CUADRO 7A
EVALUACIÓN DE LA RED NEURONAL FUERA DE MUESTRA: MODELO 4
1969:02 – 1997:04

Modelo General:							
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$							
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM1_t^*$ $X_t = \{y_{t-2}, y_{t-8}, m_{t-1}, m_{t-2}\}$							
Inputs	Q	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
X_t	0	1.0685	5.8738	0.9132	3.1183	100.0000	0.7503
X_t $Z_t = \{y_{t-2}\}$	1	1.0525	5.5035	0.8963	2.9823	100.0000	0.7030
	2	1.0813	7.2215	0.9253	3.5553	100.0000	0.9225
	3	1.0174	6.3747	0.8639	3.2673	100.0000	0.8143
	4	1.2379	6.8298	0.9984	3.4866	100.0000	0.8724
X_t $Z_t = \{y_{t-2}, y_{t-8}\}$	1	1.0002	6.0665	0.8556	3.1317	100.0000	0.7749
	2	0.9963	5.6594	0.8469	2.9839	100.0000	0.7229
	3	1.1122	5.5692	0.9232	3.0145	100.0000	0.7114
	4	1.1997	6.6676	0.9857	3.4128	69.2308	0.8517
X_t $Z_t = \{y_{t-2}, y_{t-8}, m_{t-1}\}$	1	1.0549	4.8970	0.9015	2.7622	100.0000	0.6255
	2	1.1341	5.7475	0.9251	3.1449	100.0000	0.7342
	3	3.5696	4.2586	1.7762	2.7133	76.9231	0.5440
	4	0.8108	3.9655	0.7030	2.0295	100.0000	0.5065
X_t $Z_t = \{y_{t-2}, y_{t-8}, m_{t-1}, m_{t-2}\}$	1	1.1798	6.6500	1.0446	3.2796	100.0000	0.8494
	2	1.2743	8.2147	1.0433	3.6678	100.0000	1.0493
	3	1.1425	5.7102	0.9380	3.0987	100.0000	0.7294
	4	0.9035	3.3191	0.8116	1.7446	100.0000	0.4240

CUADRO 8
EVALUACIÓN DE LA RED NEURONAL DENTRO DE MUESTRA: MODELO 4
1969:02 – 1999:04

Modelo General:								
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$								
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM1_t^*$ $X_t = \{y_{t-1}, y_{t-8}, m_{t-1}, m_{t-2}\}$								
Inputs	Q	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR
X_t	0	-0.7176	-0.5983	0.6688	17.4226	0.4851	3.7016	79.8246
X_t $Z_t = \{y_{t-1}\}$	1	-0.8322	-0.7606	0.6426	16.5296	0.4685	3.7111	78.0702
	2	-0.9252	-0.7820	0.5976	8.5668	0.4463	2.3106	78.0702
	3	-0.8887	-0.6739	0.5930	4.6121	0.4470	1.7833	78.9474
	4	-0.8881	-0.6017	0.5779	7.4661	0.4359	2.1861	80.7018
X_t $Z_t = \{y_{t-1}, y_{t-8}\}$	1	-0.8185	-0.7230	0.6415	14.1471	0.4707	3.6363	82.4561
	2	-0.8958	-0.7048	0.5960	7.9442	0.4438	2.1975	78.0702
	3	-1.0931	-0.8067	0.5216	9.7241	0.4073	2.6986	80.7018
	4	-1.1300	-0.7481	0.4945	11.5594	0.3788	2.5935	78.9474
X_t $Z_t = \{y_{t-1}, y_{t-8}, m_{t-1}\}$	1	-0.8505	-0.7311	0.6258	16.4857	0.4504	3.7547	78.9474
	2	-0.9761	-0.7374	0.5627	9.9525	0.4133	2.5861	78.9474
	3	-0.9571	-0.5991	0.5439	14.7626	0.4079	3.6528	76.3158
	4	-1.0908	-0.6134	0.4871	4.6424	0.3509	1.4633	79.8246
X_t $Z_t = \{y_{t-1}, y_{t-8}, m_{t-1}, m_{t-2}\}$	1	-0.8826	-0.7394	0.6105	19.3160	0.4649	4.0556	80.7018
	2	-0.9559	-0.6695	0.5586	27.9471	0.4326	4.7221	82.4561
	3	-1.1502	-0.7205	0.4811	23.4688	0.3792	4.4244	83.3333
	4	-1.1408	-0.5680	0.4588	12.3336	0.3517	2.9655	85.9649

CUADRO 8A
EVALUACIÓN DE LA RED NEURONAL FUERA DE MUESTRA: MODELO 4
1969:02 – 1999:04

Modelo General:							
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$							
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM1_t^*$ $X_t = \{y_{t-1}, y_{t-8}, m_{t-1}, m_{t-2}\}$							
Inputs	Q	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
X_t	0	1.1395	5.6767	0.9405	3.4430	80.0000	1.1465
X_t $Z_t = \{y_{t-1}\}$	1	1.3956	4.7374	1.3149	3.3061	100.0000	0.9568
	2	1.2335	4.1654	1.1684	2.8985	100.0000	0.8413
	3	1.2635	3.8758	1.1735	2.7026	100.0000	0.7828
	4	1.2294	3.9986	1.1633	2.8111	100.0000	0.8076
X_t $Z_t = \{y_{t-1}, y_{t-8}\}$	1	1.3130	3.7239	1.2241	2.7076	100.0000	0.7521
	2	1.2399	4.0908	1.1822	2.8717	100.0000	0.8262
	3	1.4851	3.9920	1.3314	2.9117	80.0000	0.8063
	4	1.4196	5.3194	1.3184	3.6092	100.0000	1.0744
X_t $Z_t = \{y_{t-1}, y_{t-8}, m_{t-1}\}$	1	1.3791	4.5960	1.2986	3.2227	100.0000	0.9283
	2	1.2703	4.2585	1.1943	2.9682	100.0000	0.8601
	3	1.6135	4.9487	1.5178	3.5078	80.0000	0.9995
	4	1.4400	5.2996	1.3470	3.5471	80.0000	1.0704
X_t $Z_t = \{y_{t-1}, y_{t-8}, m_{t-1}, m_{t-2}\}$	1	1.0096	4.7869	0.9274	3.0704	100.0000	0.9668
	2	1.2491	4.3763	1.1218	2.9672	100.0000	0.8839
	3	1.0468	4.6304	0.8695	2.5427	80.0000	0.9352
	4	1.0773	4.4930	0.9330	2.6742	80.0000	0.9074

CUADRO 9
EVALUACIÓN DE LA RED NEURONAL DENTRO DE MUESTRA: MODELO 5
1969:02 – 1997:04

Modelo General:								
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$								
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM3_t^*$ $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{t-1}\}$								
Inputs	Q	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR
X_t	0	-0.5798	-0.4799	0.7209	5.6904	0.5238	1.9558	81.1321
X_t $Z_t = \{y_{t-2}\}$	1	0.7119	0.7868	1.3880	12.4324	1.1628	3.7598	29.2453
	2	0.7341	0.8839	1.3647	12.4119	1.1383	3.6724	30.1887
	3	0.7394	0.9642	1.3305	11.3499	1.1083	3.4700	32.0755
	4	0.7415	1.0413	1.2951	9.9993	1.0659	3.2252	33.9623
X_t $Z_t = \{y_{t-2}, y_{t-4}\}$	1	-0.2781	-0.1782	0.8382	6.7910	0.6401	2.1638	66.9811
	2	0.1570	0.3569	1.0038	6.5523	0.8303	2.5777	55.6604
	3	0.3238	0.6235	1.0510	9.2730	0.8477	2.9594	48.1132
	4	0.5428	0.9425	1.1296	13.5578	0.9131	3.8295	41.5094
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$	1	-0.2765	-0.1516	0.8311	7.1640	0.6351	2.2572	68.8679
	2	0.0615	0.3113	0.9392	5.4591	0.7431	2.0090	56.6038
	3	-0.3891	-0.0144	0.7155	6.7607	0.5806	1.9689	78.3019
	4	0.0423	0.5419	0.8472	9.2330	0.6856	2.8585	61.3208
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{t-1}\}$	1	-0.2656	-0.1158	0.8279	5.7338	0.6524	2.1107	71.6981
	2	0.0002	0.3000	0.8940	5.5612	0.7089	2.0477	56.6038
	3	-0.3912	0.0584	0.6950	8.5849	0.5253	2.6131	78.3019
	4	-0.2858	0.3137	0.6927	7.3999	0.5558	2.2812	66.9811

CUADRO 9A
EVALUACIÓN DE LA RED NEURONAL FUERA DE MUESTRA: MODELO 5
1969:02 – 1997:04

Modelo General:							
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$							
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM3_t^*$ $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{t-1}\}$							
Inputs	Q	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
X_t	0	0.8622	2.9610	0.7875	1.6881	84.6154	0.3782
X_t $Z_t = \{y_{t-2}\}$	1	0.9557	4.1363	0.8811	2.0752	38.4615	0.5284
	2	0.9345	3.8224	0.8546	1.9166	53.8462	0.4883
	3	0.9523	4.5127	0.8905	2.2039	46.1538	0.5764
	4	0.9377	2.9497	0.8573	1.6475	61.5385	0.3768
X_t $Z_t = \{y_{t-2}, y_{t-4}\}$	1	0.8043	1.3527	0.7214	1.0296	76.9231	0.1728
	2	1.0437	5.4534	0.9087	2.1520	53.8462	0.6966
	3	1.3704	3.0135	1.2610	2.0773	61.5385	0.3849
	4	1.0126	4.3880	0.9804	2.4161	46.1538	0.5605
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$	1	0.8376	1.4521	0.7521	1.0857	76.9231	0.1855
	2	0.8933	1.0983	0.7830	0.9617	61.5385	0.1403
	3	1.0228	1.4296	0.8617	1.0706	76.9231	0.1826
	4	1.1907	9.5047	1.1319	4.1412	38.4615	1.2141
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{t-1}\}$	1	1.1707	1.6526	1.0478	1.3962	69.2308	0.2111
	2	2.0704	1.7678	1.2369	1.3211	61.5385	0.2258
	3	1.1869	2.6666	1.0267	1.6159	69.2308	0.3406
	4	1.0224	1.2566	0.9044	1.0854	53.8462	0.1605

CUADRO 10
EVALUACIÓN DE LA RED NEURONAL DENTRO DE MUESTRA: MODELO 5
1969:02 – 1999:04

Modelo General:								
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$								
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM3_t^*$ $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{t-1}, m_{t-2}\}$								
Inputs	Q	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR
X_t	0	-0.6284	-0.4852	0.6933	10.5464	0.5099	2.6813	83.3333
X_t $Z_t = \{y_{t-2}\}$	1	-0.6986	-0.6270	0.6870	9.9896	0.5055	2.6830	83.3333
	2	-0.6660	-0.5227	0.6803	12.8556	0.5028	3.0454	84.2105
	3	-0.6281	-0.4133	0.6755	10.8271	0.4990	2.8204	85.0877
	4	-0.5937	-0.3073	0.6695	9.0988	0.4879	2.5028	85.0877
X_t $Z_t = \{y_{t-2}, y_{t-4}\}$	1	-0.7871	-0.6916	0.6516	11.7115	0.4828	3.0785	80.7018
	2	-0.7967	-0.6057	0.6263	15.4645	0.4719	3.6573	81.5789
	3	-0.9087	-0.6222	0.5720	14.8434	0.4282	3.1173	84.2105
	4	-0.8110	-0.4291	0.5800	7.7694	0.4208	1.8281	84.2105
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$	1	-0.8044	-0.6851	0.6404	12.5409	0.4911	3.2908	83.3333
	2	-0.8231	-0.5844	0.6074	6.1237	0.4514	1.8657	82.4561
	3	-0.8916	-0.5336	0.5620	7.1244	0.4339	2.3678	85.0877
	4	-0.8940	-0.4166	0.5374	6.3384	0.4001	2.0966	84.2105
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{t-1}\}$	1	-0.8058	-0.6626	0.6344	12.8441	0.4797	3.2731	83.3333
	2	-0.8660	-0.5796	0.5843	12.6014	0.4533	3.1363	85.9649
	3	-0.9282	-0.4986	0.5376	20.1254	0.4102	3.9495	83.3333
	4	-1.1852	-0.6123	0.4488	16.1427	0.3279	3.3875	78.0702
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{t-1}, m_{t-2}\}$	1	-0.7786	-0.6115	0.6375	16.2655	0.4754	3.5016	83.3333
	2	-1.0437	-0.7096	0.5254	12.8725	0.3921	2.7820	85.0877
	3	-1.0597	-0.5585	0.4904	11.4929	0.3755	2.3584	85.0877
	4	-1.3126	-0.6442	0.4067	13.5146	0.3107	2.4637	79.8246

CUADRO 10A
EVALUACIÓN DE LA RED NEURONAL FUERA DE MUESTRA: MODELO 5
1969:02 – 1999:04

Modelo General:							
$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon_t$ $G(z) = (1 + \exp(-z))^{-1}$							
$y_t = \Delta LIPC_t^*$ $m_t = \Delta LM3_t^*$ $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{t-1}, m_{t-2}\}$							
Inputs	Q	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
X_t	0	0.4429	0.8343	0.3678	0.6208	100.0000	0.1685
X_t $Z_t = \{y_{t-2}\}$	1	0.3820	0.5718	0.3208	0.4376	100.0000	0.1155
	2	0.4875	1.2123	0.4112	0.8690	100.0000	0.2449
	3	0.4211	0.6646	0.3642	0.5431	100.0000	0.1342
	4	0.3911	0.4275	0.3383	0.4012	100.0000	0.0863
X_t $Z_t = \{y_{t-2}, y_{t-4}\}$	1	0.3626	1.1786	0.3076	0.8042	100.0000	0.2381
	2	0.4310	0.5464	0.3225	0.4584	100.0000	0.1104
	3	0.5827	1.3857	0.5441	1.0667	80.0000	0.2799
	4	1.4371	1.5355	0.8628	1.1846	100.0000	0.3101
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}\}$	1	0.3504	0.6583	0.2951	0.5234	100.0000	0.1330
	2	0.7217	1.7046	0.5587	1.0885	60.0000	0.3443
	3	0.4849	1.0953	0.4056	0.7123	100.0000	0.2212
	4	4.4314	3.5721	2.4632	2.3088	40.0000	0.7214
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{t-1}\}$	1	0.3135	0.6642	0.2682	0.5144	100.0000	0.1341
	2	0.6204	1.4967	0.5054	0.8893	100.0000	0.3023
	3	0.3758	1.0255	0.3560	0.7401	100.0000	0.2071
	4	0.6912	1.2948	0.5713	0.9933	80.0000	0.2615
X_t $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{t-1}, m_{t-2}\}$	1	0.4190	0.6221	0.3533	0.5285	100.0000	0.1256
	2	1.0065	1.7483	0.8483	1.4227	80.0000	0.3531
	3	0.4922	0.6946	0.2953	0.5118	80.0000	0.1403
	4	0.9964	2.4962	0.8037	1.7648	60.0000	0.5042

ANEXO 2⁴⁹

MEDIDAS DE EVALUACIÓN DE PRONÓSTICO

$$AIC(k) = \ln(\hat{\sigma}^2) + \frac{2k}{n}$$

$$BIC(k) = \ln(\hat{\sigma}^2) + \frac{k \ln(n)}{n}$$

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^n \hat{\varepsilon}_t^2}{n}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{y}_t - y_t)^2}$$

$$RMSPE = \sqrt{\frac{1}{n} \sum_{t=1}^n \left(\frac{\hat{y}_t - y_t}{y_t} \right)^2}$$

$$MAE = \frac{1}{n} \sum_{t=1}^n |\hat{y}_t - y_t|$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{\hat{y}_t - y_t}{y_t} \right|$$

$$U - THEIL = \sqrt{\frac{\sum_{t=1}^n \hat{\varepsilon}_t^2}{(y_t - y_{t-1})}}$$

$$SRP = \frac{1}{n} \sum_{t=1}^n I_t [\{(y_t - y_{t-1}) \cdot (\hat{y}_t - \hat{y}_{t-1})\} > 0]$$

$$SRN = \frac{1}{n} \sum_{t=1}^n I_t [\{(y_t - y_{t-1}) \cdot (\hat{y}_t - \hat{y}_{t-1})\} < 0]$$

$$SR = SRP + SRN$$

donde:

n es el número de observaciones consideradas

$\hat{\varepsilon}_t$ residuales obtenidos a partir de la red neuronal

k número de parámetros de la red neuronal

\hat{y}_t valor estimado por la red neuronal

⁴⁹ Véanse, Franses y van Dick (2000), Pindyck y Rubinfeld, (1997), y Makridakis y Wheelwright (1989)

ANEXO 3

En este anexo se reportan los resultados de la prueba de sensibilidad de Franses y van Dijk (2000) en la cual, a diferencia de la prueba desarrollada por Draisma et al. (1995), se mantienen únicamente las conexiones de la variable de interés y se rompen las conexiones con las demás variables.

CUADRO 1
ANALISIS DE SENSIBILIDAD SOBRE LA RED NEURONAL SELECCIONADA: MODELO 4
1969:02 – 1997:04

$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon \quad , \quad G(z) = (1 + \exp(-z))^{-1}$													
$y_t = \Delta LIPC_t^*$													
$m_{1t} = \Delta LM1_t^*$													
Inputs: $X_t = \{y_{t-2}, y_{t-8}, m_{1t-1}, m_{1t-2}\}$, $Z_t = \{y_{t-2}, y_{t-8}, m_{1t-1}, m_{1t-2}\}$ $Q=4$													
Única Conexión	Dentro de Muestra							Fuera de Muestra					
	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
y_{t-2}	0.970	1.570	1.298	23.600	0.978	5.657	61.3	1.021	2.155	0.880	1.409	46.2	0.275
y_{t-8}	0.874	1.474	1.237	14.102	0.981	4.768	50.9	0.964	3.860	0.925	2.120	30.8	0.493
m_{1t-1}	0.512	1.112	1.032	18.218	0.835	4.897	68.9	1.125	1.141	0.877	0.988	69.2	0.146
m_{1t-2}	1.694	2.293	1.864	44.832	1.534	12.74	51.9	2.241	7.907	1.893	4.678	46.2	1.010

CUADRO 2
ANALISIS DE SENSIBILIDAD SOBRE LA RED NEURONAL SELECCIONADA: MODELO 5
1969:02 – 1999:04

$y_t = X_t' \Phi + \sum_{j=1}^Q \beta_j G(Z_t' \gamma_j) + \varepsilon \quad , \quad G(z) = (1 + \exp(-z))^{-1}$													
$y_t = \Delta LIPC_t^*$													
$m_{3t} = \Delta LM3_t^*$													
Inputs: $X_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{3t-1}, m_{3t-2}\}$, $Z_t = \{y_{t-2}, y_{t-4}, y_{t-8}, m_{3t-1}, m_{3t-2}\}$ $Q=3$													
Única Conexión	Dentro de Muestra							Fuera de Muestra					
	AIC	BIC	RMSE	RMSPE	MAE	MAPE	SR	RMSE	RMSPE	MAE	MAPE	SR	U-THEIL
y_{t-2}	0.227	0.728	0.933	13.946	0.754	3.088	65.8	1.194	1.050	1.015	0.988	100	0.212
y_{t-4}	0.328	0.830	0.982	13.718	0.782	3.030	55.3	0.965	1.969	0.921	1.593	100	0.398
y_{t-8}	0.510	1.011	1.075	16.890	0.901	4.434	42.1	1.394	2.323	1.270	1.815	80	0.469
m_{3t-1}	0.690	1.191	1.176	18.500	0.933	4.996	57.0	0.531	1.450	0.496	1.055	80	0.293
m_{3t-2}	1.407	1.908	1.683	67.848	1.392	15.51	50.0	1.516	7.061	1.313	4.521	40	1.426

ANEXO 4

1. Modelo Curva de Phillips⁵⁰:

Este modelo está basado en el modelo triangular convencional de Phillips, el cual tiene en cuenta los cambios en las expectativas influenciados por el desarrollo de la política monetaria. La estimación de este modelo incorpora además de los choques de demanda y de oferta, los regímenes de inflación basados en los resultados de un modelo de Markov-switching, a través de una variable que mide la probabilidad de estado de mayor media de inflación. A nivel internacional, este procedimiento fue propuesto por Ricketts y Rose (1995) y Fillion y Leonard⁵¹. En la actualidad, este modelo utiliza como variables explicativas la inflación rezagada, la inflación de los bienes importados, una estimación de la brecha del producto y una variable switching que recoge la probabilidad.

$$\pi_t = \beta_0 + \beta_1 \pi_{t-4} + \beta_2 \pi_{t-6} + \beta_3 \pi_{t-2}^m + \beta_4 y_{t-1} + \beta_5 P_{1t} + \delta' s_{it} + \beta_6 \delta_{1t} + \varepsilon_t$$

Definiciones:

- π_t es la primera diferencia del logaritmo del IPC fin de período
- y_t es la brecha del producto calculado por el modelo de Mecanismos de Transmisión a partir de una función de producción para el PIB potencial.
- π_t^M es la primera diferencia del logaritmo de los precios de los bienes importados.
- P_{1t} es la probabilidad de estado de mayor media de inflación según el modelo de Switching de Hamilton.
- δ_{1t} es una variable dicótoma que toma el valor de uno en el segundo trimestre de 1986 y cero en el resto.
- s_{it} son las variables dummies estacionales

2. Modelo P estrella

El modelo P estrella original para la inflación fue propuesto por Hallman et al. (1989, 1991)⁵², el cual está basado en la teoría cuantitativa del dinero. En él se asume un ajuste lento de los precios ante cambios no anticipados de la oferta de dinero, de tal forma que las desviaciones de los precios de su nivel de largo plazo generan presiones inflacionarias.

El P* tradicional asume que la velocidad de circulación del dinero es estacionaria, por lo que el nivel de precios depende de la brecha de la velocidad del agregado monetario, en este caso M1, y de la brecha del producto (calculado por el filtro de Holdrick-Prescott del PIB).

⁵⁰ Véanse, López y Misas (1998) '.

⁵¹ Véanse, Ricketts and Rose (1995) y Fillion, J.F. and A. Leonard (1997).

⁵² Véanse, Hallman, J., Porter R. and Small D. (1989, 1991)

P estrella tradicional:

$$\pi_t = \beta_0 + \beta_1 \pi_{t-1} + \beta_2 \pi_{t-4} + \beta_3 \pi_{t-6} + \beta_4 y_{t-1} + \beta_5 \eta_{t-1} + \delta' s_{it} + \beta_6 \delta_{1t} + \beta_7 \delta_{2t}$$

Definiciones:

- y_t es la brecha del producto obtenida a través del filtro de Hodrick-Prescott.
- η_t es la brecha de la velocidad del agregado monetario M1 obtenido a través del filtro de Hodrick-Prescott.
- δ_{2t} es una variable dicótoma que toma el valor de uno en el primer y segundo trimestres de 1985 y cero en el resto.

3. Modelo de Precio Relativo de Alimentos:

Este modelo busca capturar los efectos del choque de oferta más común como es el de los precios de los alimentos, así como controlar los choques de demanda utilizando una estimación de la brecha del producto y de la oferta monetaria. La proyección del precio de los alimentos se considera exógena y es construida con otro modelo que proyecta la inflación de alimentos en función de su dinámica pasada y de la probable evolución de la cantidad de lluvias. El modelo está complementado con la evolución de la cantidad de dinero definido como M1, con la brecha del producto y con rezagos de la inflación.

$$\pi_t = \beta_0 + \beta_1 \pi_{t-1} + \beta_2 \pi_{t-4} + \beta_3 \pi_{t-6} + \beta_4 \pi^R_t + \beta_5 \pi^R_{t-5} + \beta_6 y_t + \beta_7 M_{1t-3} + \delta' s_{it} + \beta_8 \delta_{1t} + \beta_9 \delta_{2t} + \beta_{10} \delta_{3t}$$

Definiciones:

- π^R_t es la primera diferencia del logaritmo del precio relativo de alimentos (IPC alimentos/ IPC total)
- M_{1t} es la primera diferencia del logaritmo de M1.
- δ_{3t} es una variable dicótoma que toma el valor de uno en el primer y segundo trimestres de 1999 y cero en el resto

4. Modelo Escandinavo:

Este modelo sigue la tradición "escandinava", en donde la inflación de una economía abierta y pequeña está determinada principalmente por factores de costo como el crecimiento de los salarios reales y por la devaluación real.

$$\pi_t = \beta_0 + \beta_1 \pi_{t-1} + \beta_2 \pi_{t-4} + \beta_3 \pi_{t-6} + \beta_4 \chi_{t-3} + \beta_5 \omega_t + \beta_6 \omega_{t-1} + \beta_7 \omega_{t-3} + \beta_8 \delta_{1t} + \beta_9 \delta_{2t} + \beta_{10} \delta_{3t}$$

Definiciones:

- χ_t es el crecimiento trimestral de la tasa de cambio real.
- ω_t es el crecimiento trimestral del salario real de la industria.