

# “Aplicación de Redes Neuronales Artificiales para el Análisis de la Inflación en Bolivia”

---

4to. Encuentro de Economistas de Bolivia

Julio 2011

Daniela Urquizo Rojas

Fabián Mendivil Gomez

## RESUMEN (ABSTRACT)

El principal objetivo de la política monetaria es estabilizar el nivel de precios, lo cual, en el marco de una economía de mercado resulta altamente complejo por la no-linealidad de las relaciones entre las variables que interactúan con la inflación de forma directa o indirecta.

El presente documento de investigación analiza la inflación en Bolivia a partir de variables ampliamente sustentadas por la teoría económica, empero, la herramienta utilizada para tal fin es poco ortodoxa al tratarse de Redes Neuronales Artificiales (RNA). Las RNA se presentan actualmente como una interesante alternativa para el análisis de fenómenos que involucran el contexto social y de expectativas que métodos convencionales no capturan con precisión. Dicho esto, se examina la inflación y su relacionamiento con variables como ser: emisión monetaria, IGAE, tipo de cambio y la misma inflación rezagada, recurriendo a información estadística correspondiente al periodo post-inflacionario.

**Palabras Clave:** *Red neuronal artificial, no linealidad, nivel de precios, política monetaria, inflación.*

**Clasificación JEL:** *E3, C53*

Información de contacto: [fabianmg157@gmail.com](mailto:fabianmg157@gmail.com), [danielaurquizo@gmail.com](mailto:danielaurquizo@gmail.com)

## **Introducción**

Bolivia, como economía pequeña, abierta y dolarizada, presenta características que la distinguen de sus principales socios comerciales respecto al comportamiento de sus variables macroeconómicas y por su puesto al manejo de su política económica, lo cual comprende a las políticas monetaria y cambiaria que aplica el Banco Central de Bolivia.

Siendo la inflación, la principal variable en la que se enfoca el Banco Central para desarrollar sus herramientas de control, y las medidas que procuren estabilizarla, es menester profundizar no sólo el análisis convencional que se puede hacer sobre dicha variable sino, expandir el alcance de otros instrumentos que permitan conocer mejor su comportamiento y mejor aún lograr predecir su valor a corto plazo.

El presente documento de investigación tiene como objetivo estudiar empíricamente a la inflación mediante el uso de Redes Neuronales Artificiales (RNA), para explorar la relación que tiene con sus posibles determinantes, entre ellos la emisión monetaria, variación del tipo de cambio nominal de compra, el índice global de actividad económica (IGAE) y dos variables de rezago para la inflación.

Las RNAs han demostrado ser de gran utilidad para efectuar análisis de variables económicas y financieras debido a que con frecuencia se presentan relaciones no lineales que son de complejo estudio para modelos econométricos convencionales. Durante los últimos años han sido tomadas como una excelente alternativa porque con ellas se elaboran modelos computacionales que de manera simplificada, intentan replicar el funcionamiento complejo del cerebro humano. Asimismo, poseen una gran capacidad de aprendizaje a través de ensayos repetidos, lo cual ha aumentado su popularidad por el aporte que trae aplicación en las ciencias económicas y sociales, entre otras.

Este artículo, incorpora algunas innovaciones en la modelación de la inflación y sus determinantes que permiten generar pronósticos más confiables, gracias a que el modelo se aproxima con mayor exactitud a la realidad. El estudio se divide en cuatro secciones principales: en la primera, se exponen elementos teóricos y conceptuales acerca de la inflación, desde un enfoque teórico y empírico; en la segunda sección, se despliega una introducción a las RNAs, su estructura básica, funcionamiento y lo que compone un análisis de sensibilidad. La tercera sección corresponde a la aplicación de redes neuronales en el caso de la inflación en Bolivia. Finalmente, la cuarta sección presenta las conclusiones derivadas de la investigación.

## I. Elementos conceptuales sobre la teoría de la inflación

Por definición se conoce a la inflación como el cambio porcentual del nivel general de precios en un periodo determinado, de manera persistente y sostenida, se han planteado teorías desde distintos enfoques para explicar su comportamiento.

### I.1 Teoría Cuantitativa del dinero

Esta teoría, apoyada por la escuela monetaria es considerada la más tradicional ya que propone la existencia de una relación directa entre la cantidad de dinero y la inflación. La ecuación de transacciones de Irving Fischer establece una relación entre la oferta monetaria  $M$ , la velocidad del dinero  $V$ , el nivel de precios  $P$  y el producto  $Y$ , además del volumen de transacciones  $T$ .

$$MV \equiv PY$$

Suponiendo que la velocidad del dinero está determinada por factores institucionales en el sector monetario y que, a corto plazo, permanece constante, y si además se considera que el sector real determina el volumen de transacciones, entonces el nivel de precios será proporcional a la oferta monetaria  $M$ .

$$P = \frac{V}{Y}M$$

### I.2 La Curva de Phillips

Planteada por A.W. Phillips, expone la relación entre la inflación y el desempleo, dando por hecho que el propósito de mantener la estabilidad de precios es conservar el salario real de los individuos de una economía.

La curva de Phillips dio lugar a la incorporación de la NAIRU (non-accelerating inflation rate of unemployment), en adición a la Ley de Okun, para establecer una nueva curva de Phillips aumentada por las expectativas<sup>1</sup>. De esta forma, cuando la inflación resulta igual a la anticipada, el desempleo se ubica en su tasa natural. Asimismo, si la inflación supera lo anticipado, el desempleo resulta ser mayor que su tasa natural.

---

<sup>1</sup> LARRAÍN, SACHS "Macroeconomía en la economía global" Segunda Edición, extraído del Capítulo "Inflación desempleo y estabilización", pág. 357.

### I.3 Modelo Keynesiano

Planteado por John Maynard Keynes, el modelo de la “brecha inflacionaria” propone un numerador que se halla en relación directa con el incremento de la inflación, partiendo de la siguiente ecuación:

$$Y_t = C_t + I_t = Y_0$$

Donde  $Y_0$  es la producción real en condiciones de pleno empleo. La función de consumo se plantea de la siguiente manera,

$$C_t = a + b \frac{Y P_{t-1}}{P_t}$$

Donde  $a$  es el consumo autónomo real,  $b$  es la propensión marginal a consumir,  $Y P_{t-1}$  es el ingreso monetario del periodo anterior, y  $P_t$  es el nivel de precios corriente. Suponiendo que, además el nivel de inversión real es exógeno:

$$I_t = I_0$$

$$Y P_t = Y_0 P_t$$

Sustituyendo las anteriores ecuaciones:

$$\frac{P_t}{P_{t-1}} = \frac{b Y_0}{Y_0 - (a + I_0)}$$

$$P_t = \frac{(a + I_0) + b Y_0 - Y_0}{Y_0 - (a + I_0)}$$

Se puede notar claramente que el numerador del lado izquierdo es la conocida “brecha inflacionaria”, la cual, mientras más grande es, representa una mayor tasa de inflación. Esta teoría explica la inflación como surgida de un exceso en la demanda agregada por encima de la capacidad de pleno empleo de la economía.

## I.4 Enfoque Neoclásico

Conocido también como el de las “Expectativas Racionales”, tiene dos supuestos fundamentales: el equilibrio automático de los mercados o sea que estos tienden permanentemente al equilibrio y que los agentes forman sus expectativas de forma eficiente y racional. A través de este enfoque, desaparecería la brecha entre la inflación esperada y la efectiva.

El siguiente es un modelo de expectativas racionales, suponiendo que la tasa de crecimiento del dinero es exógena, a través de una función de demanda de dinero de Cagan:

$$\frac{M_t}{P_t} = Y^{\beta_0} e^{\beta_1 - \beta_2 \pi_{t+1}}$$

Donde  $\frac{M^d}{P}$  es la demanda de saldos reales,  $P$  es el nivel de precios,  $Y$  es el ingreso real y  $\pi_{t+1}$  es la tasa de inflación esperada.

## I.5 Revisión de estudios empíricos sobre la inflación en Bolivia

### I.5.1 Nivel de inflación basado en el tipo de cambio

Un modelo teórico de inflación para Bolivia, de Comboni y De la Viña<sup>2</sup> expresado para el período de postestabilización económica en Bolivia: 1986 – 1991, presenta un rasgo notable como la tasa de depreciación del tipo de cambio la cual está determinada en forma simultánea con la tasa de inflación.

Además de considerar la tasa de crecimiento del tipo de cambio, el modelo también considera incrementos en el precio de los hidrocarburos y variaciones en el nivel internacional de precios (inflación externa), como determinantes (exógenos) de la tasa de inflación doméstica. A partir de este pequeño modelo estructural descrito anteriormente, se deriva una especificación dinámica para la forma reducida del modelo, la misma que es aproximada mediante un vector autorregresivo (VAR). Esta estructura puede utilizarse para realizar tests de cointegración entre la tasa de crecimiento del tipo de cambio y las tasas de inflación doméstica y extranjera, de

---

<sup>2</sup> COMBONI y DE LA VIÑA (2002), “Precios y Tipo de Cambio en Bolivia: Evidencia Empírica del Período de Post-Estabilización” Pág. 4

acuerdo a la relación prevista por la paridad del poder de compra (PPC). El modelo básico viene definido por las siguientes ecuaciones:

$$\begin{aligned}\pi_t &= \varphi_0 + \varphi_1 \hat{e}_t + \varphi_2 p_t + \mu_t \\ \Delta \hat{e}_t &= \beta(\hat{e}_t^* - \hat{e}_{t-1}) \\ \hat{e}_t^* &= \tau_0 + \tau_1(\pi_{t-1} - \pi_{t-1}^*) + \varepsilon_t\end{aligned}$$

La primera ecuación del modelo expresa la tasa de inflación en función de la tasa de crecimiento del tipo de cambio nominal y del incremento en el precio de los hidrocarburos; la segunda ecuación simplemente muestra un mecanismo de ajuste parcial por el que (la tasa de crecimiento de) el tipo de cambio nominal converge hacia su valor de equilibrio  $\hat{e}_t^*$ , con una velocidad dada por el parámetro  $\beta$ .

Cuando el valor de este coeficiente es igual a la unidad, entonces el tipo de cambio se ajusta a su nivel de equilibrio en forma inmediata; por el contrario, cuando  $\beta$  se aproxima a cero, el tipo de cambio se asemeja a un proceso autorregresivo de primer orden.

### 1.5.2 Determinantes de la Inflación

Según Orellana y Requena<sup>3</sup> (1999), es posible determinar el nivel de inflación basándose en un modelo de vectores autorregresivos, formado a partir de dos ecuaciones: la primera, expresa la tasa de inflación ( $\pi_t$ ) en función de la inflación desfasada ( $\pi_{t-1}$ ), de la depreciación del tipo de cambio nominal ( $e_{t-1}$ ), del crecimiento de la emisión desfasada ( $E_{t-1}$ ), y de la inflación internacional medida en dólares ( $\pi_{t-1}^*$ ); además se emplearon un conjunto de variables *dummies* para capturar el efecto del incremento en el precio de los hidrocarburos y parte de la estacionalidad en la inflación (D). Las ecuaciones del modelo se expresan de la siguiente manera:

$$\pi_t = \delta_1 + \alpha_{11}\pi_{t-1} + \alpha_{12}e_{t-1} + \beta_{11}E_{t-1} + \beta_{12}D + \beta_{13}\pi_{t-1}^* + u_1 \quad (1)$$

$$e_t = \delta_2 + \alpha_{21}\pi_{t-1} + \alpha_{22}e_{t-1} + \beta_{21}E_{t-1} + \beta_{22}D + \beta_{23}\pi_{t-1}^* + u_2 \quad (2)$$

Combinando las ecuaciones (1) y (2) obtiene la representación dinámica de este modelo en términos de un vector autorregresivo de primer orden.

$$X_t = A + BY_t + CX_{t-1} + U_t$$

<sup>3</sup> ORELLANA, REQUENA (1999) "Determinantes de la Inflación en Bolivia", pág. 18

Los resultados del modelo VAR indican que el efecto transmisión de la depreciación a la inflación se ha reducido en los últimos años, es decir, que en la medida que la depreciación del boliviano se ha reducido el efecto transmisión del tipo de cambio a la inflación también cayó. De esta manera, el estudio sugiere que la relación entre la tasa de depreciación de la moneda local la inflación es no lineal, demostrando también que el ancla nominal es efectiva en condiciones de elevada inflación pero reduciendo su efecto en condiciones de baja inflación.

### **I.5.3 Metas de Inflación en Economías Dolarizadas**

Según Liederman, Maino y Parrado (2005), la adaptación de un esquema de metas explícitas de inflación es un adecuado régimen de política monetaria para economías con altos o bajos niveles de dolarización. El análisis concluye que un esquema de Metas de Competitividad con Miedo a Flotar (MCMF) – además de los ajustes rezagados del tipo de cambio – es que la fijación estricta de metas de tipo de cambio tiende a promover la dolarización. En cambio, el esquema de Metas de Inflación Completamente Desarrollado (MICD) debe tender a reducirla.

El estudio emplea un modelo VAR que incluye siete variables: precio mundial del petróleo, tasa de interés internacional, el PBI desestacionalizado, la tasa de inflación, la tasa de interés de referencia, la masa monetaria y el tipo de cambio nominal. En el caso de Bolivia, la variable dependiente es la tasa de Crawl y no la tasa de interés como en el caso de los demás países que se incluyen en el estudio, esto quiere decir, que no está permitida la flotación del tipo de cambio<sup>4</sup>.

### **I.5.4 Choques Exógenos, Expectativas y Credibilidad de la Política Monetaria**

Los autores Mendieta, Cerezo y Cossío (2008), realizaron la estimación de un modelo para las expectativas inflacionarias, hallando sus determinantes para varios países de Sudamérica, entre ellos, Bolivia. Mediante el uso de modelos de panel para la inflación esperada, tomaron como variables a la inflación rezagada con un periodo, la brecha del tipo de cambio efectivo y real, la brecha del producto, la credibilidad (variable que mide la reputación en el cumplimiento de las metas en el último año móvil), la tasa de política monetaria, la variable meta y la variable dummy de metas de inflación. Asimismo, el documento analiza empírica y teóricamente los determinantes de las expectativas, para el caso de Bolivia, mediante el siguiente modelo<sup>5</sup>:

---

<sup>4</sup> LIEDERMAN, MAINO, PARRADO (2005) “Metas de Inflación en Economías Dolarizadas”, pág. 66

<sup>5</sup> MENDIETA, CERESO, COSSIO (2008) “La inflación está de vuelta en Sudamérica? Choques exógenos, expectativas y credibilidad de la política monetaria” pág. 134

$$E_t \pi_{t+12} = 0,723 - 0,073\bar{\pi} + 0,816E_{t-1}\pi_{t+11} + 0,468(\pi_{t-1} - \bar{\pi}) + 0,433\tilde{y} + 0,043i_{t-1} - 1,766\tilde{R}$$

Para concluir la sección de reseña bibliográfica de la inflación, se presenta el siguiente cuadro resumen de las variables consideradas en los estudios empíricos citados:

**Tabla 1: RESUMEN DE VARIABLES CONSIDERADAS  
EN ESTUDIOS EMPÍRICOS DE LA INFLACIÓN**

Autor	Título	Variable Endógena	Variables Exógenas	Signo de la relación	Metodología
Mendieta, Cerezo, Cossío (2008)	¿La inflación está de vuelta en Sudamérica? Choques exógenos, expectativas y credibilidad de la política monetaria	Inflación Esperada	Inflación Rezagada	Positivo (+)	Modelo de Datos de Panel  Regresión sobre la media Efectos Aleatorios Efectos Fijos
			Dummy Metas de Inflación	Negativo (-)	
			Meta	Positivo (+)	
			Tasa de Política Monetaria	Negativo (-)	
			Credibilidad	Positivo (+)	
			Brecha del Producto	Negativo (-)	
			Brecha del TCR	Positivo (+)	
Cossío, Laguna, Martín, Mendieta, Mendoza, Palmero, Rodríguez (2007)	La Inflación y Políticas del Banco Central de Bolivia	Núcleo Inflacionario	Índice Global de Actividad Económica	Positivo (+)	Vectores Autorregresivos (VAR)
			Tipo de Cambio nominal	Positivo (+)	
			Emisión Monetaria	Positivo (+)	
			Remonetización de la economía	Débilmente negativo (-)	
			Gasto del Gobierno General	Débilmente positivo (+)	
Leiderman, Maino, Parrado (2005)	Metas de Inflación en Economías Dolarizadas	Tasa de Crawl (tasa de depreciación de la moneda local)	Tasa de inflación	Negativo (-)	Vectores Autorregresivos (VAR)
			Tasa de variación del TCRE	Negativo (-)	
			Tasa de variación de las RIN	Negativo (-)	
			PIB desestacionalizado	Negativo (-)	
			Tasa de interés de referencia	Negativo (-)	
Orellana y Requena (1999)	Determinantes de la inflación en Bolivia	Inflación	Inflación Desfasada	Positivo (+)	Vectores Autorregresivos (VAR)
			Depreciación del tipo de cambio nominal (elevado al cuadrado)	Positivo (+)	
			Crecimiento de la emisión desfasada	Positivo (+)	
			Inflación Internacional	Débilmente positivo (+)	
			Precio de los hidrocarburos (dummy)	Positivo (+)	
			Fenómeno del Niño y parte de la estacionalidad de la inflación (dummy)	Positivo (+)	

**Elaboración:** Propia en base a los estudios empíricos descritos

## I.6 El Rol del Banco Central en el Control de la Inflación

La principal prescripción de la política monetaria es la estabilidad de precios del consumidor. Para alcanzar su objetivo, el Banco Central interviene mediante Operaciones de Mercado Abierto, estableciendo encajes legales de cumplimiento obligatorio y también facilitando liquidez

a las entidades del sistema financiero a través de créditos con garantía del Fondo RAL, como con operaciones de reporto<sup>6</sup>.

El BCB emplea un esquema de metas intermedias fijando límites a la expansión de su Crédito Interno Neto (CIN) y un piso a la variación de las Reservas Internacionales Netas (RIN), ya que Bolivia es una economía dolarizada.

A través de la política cambiaria, el BCB emplea al tipo de cambio como ancla nominal de los precios, por el efecto que tiene este sobre el precio de los bienes importados, manteniendo a su vez la competitividad cambiaria y promoviendo la estabilidad del sistema financiero.

El tipo de cambio, como ancla nominal, mantiene el nivel de precios lejos de crecer o caer demasiado rápido y de esa manera preserva el valor del dinero en una economía. De esta manera, ayuda a promover la estabilidad de precios atando las expectativas de inflación a bajos niveles directamente a través de coaccionar la variable con el valor de la moneda. La razón más ingeniosa por la que se considera importante a un ancla nominal es porque puede limitar el problema de inconsistencia dinámica, en el que la política monetaria conducida con discrecionalidad resulta pobre a largo plazo.

## **II. Redes Neuronales Artificiales (RNAs)**

La búsqueda de nuevas técnicas y métodos que permitan mejorar la eficiencia en la ejecución de diferentes tipos de tareas, siempre ha sido la característica fundamental del hombre, puesto que en su afán por optimizar el tiempo, incrementar la eficiencia y automatizar las operaciones mecánicas, ha logrado una gran variedad de avances científicos a través del tiempo. Precisamente un desarrollo relativamente reciente está dado por las Redes Neuronales Artificiales (RNA).

Las RNA representan un nuevo paradigma en el manejo y análisis de la información, inspirada en el funcionamiento de las miles de millones de neuronas biológicas que el ser humano posee en su sistema nervioso. Las neuronas son células que se encuentran alojadas principalmente en el cerebro que es el lugar en el que se genera el “pensamiento”. Prácticamente es en el comportamiento individual y especialmente colectivo de cada una de estas neuronas, donde se encuentra el secreto de la inteligencia humana.

---

<sup>6</sup> COSSÍO, LAGUNA, MARTIN, MENDIETA, MENDOZA, PALMERO y RODRÍGUEZ (2007) “La Inflación y Políticas del Banco Central de Bolivia”, Jornadas Monetarias del BCB

Las redes de neuronas del cerebro constituyen sistemas paralelos, distribuidos y adaptativos. Paralelos porque las neuronas trabajan todas al mismo tiempo en la resolución de un mismo problema, es decir, trabajan de forma paralela. Constituyen sistemas distribuidos ya que a diferencia de un computador cuya memoria se ubica en un lugar específico, en este caso, ésta se encuentra diseminada a través de toda la red y más específicamente en sus diferentes sinapsis (puntos de comunicación o conexión entre neuronas). Finalmente son sistemas adaptativos ya que mediante la dinámica de la sinapsis o conexiones entre neuronas logran imitar una de las capacidades más importantes del cerebro humano: el aprendizaje.

Las RNA son sistemas que aproximan a nivel básico todo este complejo hardware biológico, resultando es un esquema artificial semi-paramétrico (McNelis, pag. 20, 2005) que permite el análisis de variables financieras y económicas en las que el mecanismo de relacionamiento subyacente es especialmente no lineal.

En este sentido, la unidad básica de una RNA es la neurona, denominada también procesador elemental, representa un dispositivo simple de cálculo que recibe señales exteriores de otras neuronas y proporciona una respuesta a dichos impulsos. La sinapsis a nivel biológico constituye un intercambio de sustancias químicas denominadas neurotransmisores.

A nivel artificial representa la dinámica existente entre las variables inputs que se relacionan con la variable output por medio de neuronas intermedias y un conjunto de pesos o ponderaciones que representan la importancia que cada una de éstas posee. En un sentido más analítico, el modelo general de neurona artificial contendrá al menos los siguientes elementos:

- Un conjunto de entradas  $x_{j,t}$
- Un conjunto de pesos o ponderaciones  $w_{ij}$  que constituyen la intensidad de la interacción entre una neurona presináptica  $j$  (que envía información) y una neurona postináptica  $i$  (que recibe la información)
- Una regla de propagación  $f_i[x_{j,t}, w_{ij}]$  que proporcionará el valor del potencial postináptico  $p_{i,t} = f_i[x_{j,t}, w_{ij}]$
- Una función de activación  $L_i[a_{i,t-1}, p_{i,t}]$  que determina el estado de activación actual de la neurona  $i$   $a_{i,t} = L_i[a_{i,t-1}, p_{i,t}]$ , que depende de su estado de activación anterior  $a_{i,t-1}$  y de su potencial postsináptico actual
- Una función de salida  $S_i[a_{i,t}]$ , que determinará la salida actual  $y_i(t) = S_i[a_{i,t}]$

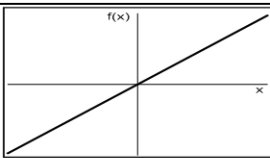

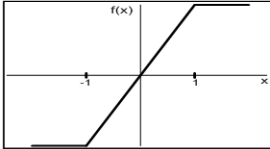
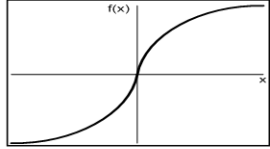
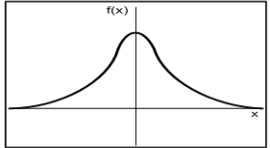
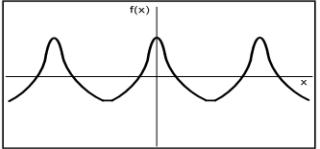
De esta manera, la dinámica de comportamiento de la neurona  $i$  estaría dada por la siguiente ecuación:

$$N_{i,t} = S_i\{L_i[a_{i,t-1}, f_i(x_{j,t}, w_{ij})]\} \quad (1)$$

Por lo general la regla de propagación  $f(\bullet)$  se reduce a una sumatoria de las diferentes entradas  $x_j$  ponderadas por sus respectivos pesos  $w_{ij}$ , de la forma  $\sum (x_{j,t} \cdot w_{ij})^7$ , que una vez superado cierto umbral, es trasladado a una función de activación o transferencia  $L(\bullet)$ .

Las funciones de activación cumplen además la función de achatar el valor del potencial postsináptico y mantenerlo dentro una escala específica de valores, tal como se expone en la Tabla 2.

**Tabla 2: TIPOS DE FUNCIONES DE ACTIVACIÓN**

NOMBRE	FUNCIÓN	RANGO	GRÁFICA
Identidad	$y = x$	$[-\infty, +\infty]$	
Escalón	$y = \text{signo}(x)$ $y = H(x)$	$[-1, +1]$ $[0, +1]$	
Lineal a tramos	$y = \begin{cases} -1, & \text{si } x < -1 \\ x, & \text{si } -1 \leq x \leq 1 \\ +1, & \text{si } x > 1 \end{cases}$	$[-1, +1]$	
Sigmoidea o Logística	$y = \frac{1}{1 + e^{-x}}$ $y = \text{tanh}(x)$	$[0, +1]$ $[-1, +1]$	
Gaussiana	$y = a \cdot e^{-bx^2}$	$[0, +1]$	
Sinusoidal	$y = a \cdot \sin(\omega x + \phi)$	$[-1, +1]$	

Fuente: Martín del Brío, Sanz M., 1997

<sup>7</sup> Ciertos tipos de RNA como los mapas autoorganizados de Kohonen y las redes en base radial, hacen uso de reglas de propagación de diferente, como la distancia euclídea:  $f_i^2(t) = \sum (x_j - w_{ij})^2$ .

La capacidad de las RNA de capturar no linealidades entre los inputs y el output se origina precisamente en el uso de estas funciones que suelen ser de tipo logístico, tangente hiperbólico, trigonométrico, entre otros (Zhang G., Patuwo E., Hu M., 1997). Esta función determina, tal como lo indica su nombre, el nivel o estado de activación de la neurona y que además representa la señal de salida que será enviada hacia otra neurona.

Desde el punto de vista de la economía y la teoría financiera, se suele resaltar el beneficio del uso de funciones de activación no lineales a partir de la dinámica presente en los movimientos en variables económicas fundamentales, como el caso de la tasa de interés que cuando se encuentra en niveles extremos, sus variaciones no tienen efectos importantes sobre el consumo y la inversión. Incluso, en el caso de la función de activación logística, se hace analogía de su forma con el aprendizaje y la formación de expectativas de los agentes económicos (McNelis, pag. 24, 2005).

Las RNA están constituidas por grupos de neuronas que se agrupan a su vez en unidades estructurales denominadas capas o superficies. Es posible distinguir tres tipos principales de capas: la capa de entrada compuesta por variables inputs cada una de las cuales constituyen neuronas sin sinapsis entrantes, la capa oculta o intermedia que es aquella que no posee una conexión directa con el entorno y constituye más bien un escenario de reprocesamiento de las señales provenientes de la capa de entrada y que posteriormente determinarán en gran medida los resultados. Finalmente la capa de salida que proporcionan la respuesta de la red y está constituida por la variable output que es una neurona sin sinapsis saliente.

A pesar de que la inclusión de una superficie oculta aflora una serie de dudas sobre el procesamiento interno que las RNA realizan, dando lugar a la crítica de la “caja negra”<sup>8</sup>, su incorporación está plenamente justificada desde una óptica de comportamiento de las variables económicas. Aquí el concepto de variables latentes cobra crucial importancia puesto que es un hecho comprobado que el efecto de la variación de variables macroeconómicas o microeconómicas sobre otras de similar naturaleza no siempre es directo. Incrementos en las tasas de interés pueden afectar negativamente al crecimiento económico disminuyendo primero el consumo y la inversión, un incremento en la tasa de rendimiento de bonos del banco central puede incrementar las tasas de interés para préstamos afectando en primera instancia los

---

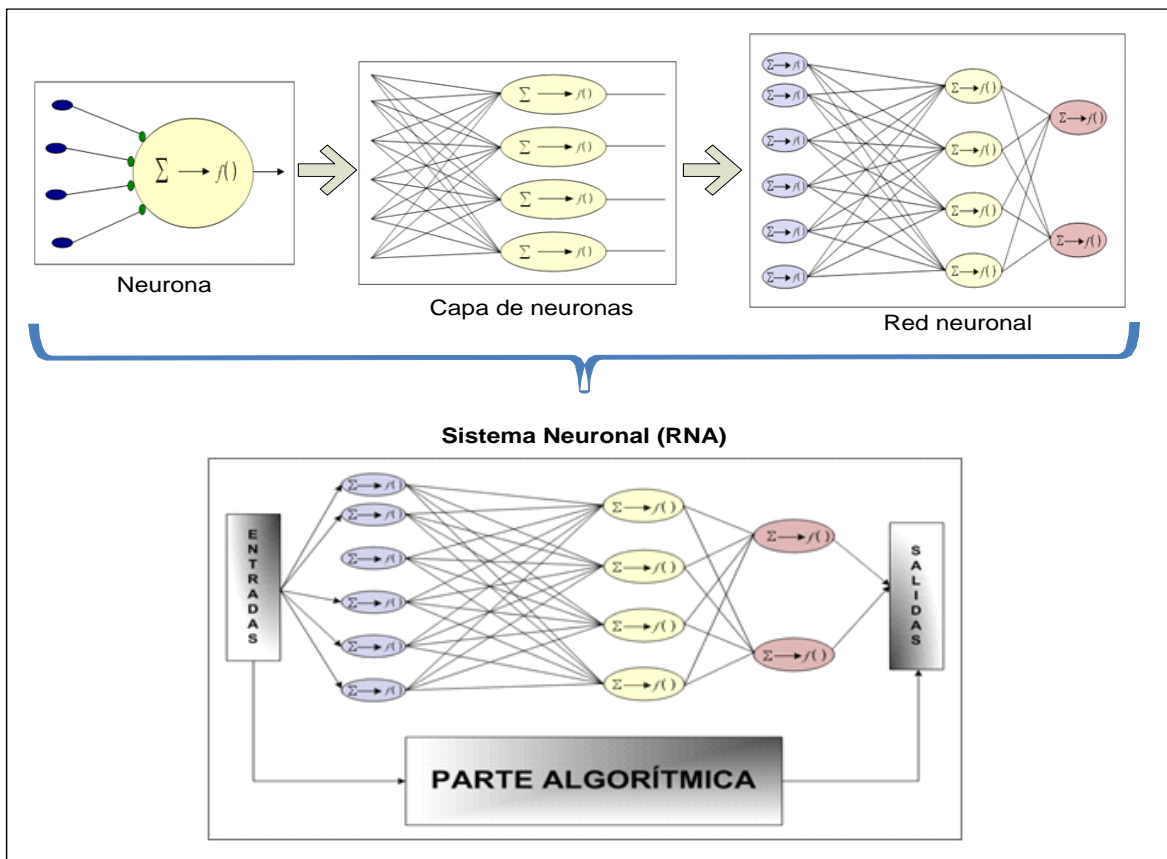
<sup>8</sup> Esta crítica tiene que ver con el hecho de que las RNA no proveen mucha información acerca de cómo la solución es encontrada, ya que la misma se genera en un proceso interno dado en las neuronas de la capa oculta de la red. Sin embargo, como se verá más adelante, actualmente, con los nuevos avances y estudios en el área, ésta crítica está perdiendo fuerza y hoy en día es posible entender cada vez más la forma cómo operan internamente las RNA, principalmente a partir del análisis de sensibilidad.

depósitos del público en el sistema bancario y las tasas de interés pasivas. La mora en el sistema bancario para colocaciones en dólares podría incrementarse como resultado de la apreciación de la moneda extranjera, que erosiona primero los ingresos en moneda doméstica de los prestatarios e incrementa su probabilidad de incumplimiento.

Entre otras justificaciones, la incorporación de ésta capa oculta además provee a la RNA la capacidad de generalización (Kaastra, Boyd, 1996) es decir, de responder eficientemente a nueva información o nuevos datos, debido a que captura adecuadamente el patrón de conexión existente entre los inputs y el output.

El Gráfico 1 ilustra la forma como deben concebirse las RNA empezando de una unidad básica denominada neurona, cuya agrupación permite generar capas, la agrupación de éstas a su vez generan una red. Finalmente, el sistema, para fines prácticos y de utilización, deberá contener variables de entrada, salida y un conjunto de relaciones matemáticas que permitirán la interacción de todos estos elementos.

**Gráfico 1: ESTRUCTURA JERÁRQUICA DE LAS RNA**



**Fuente:** Elaboración propia con Martin del Brío, Sanz M., 1997

Matemáticamente, un sistema neuronal se representa mediante el siguiente sistema de ecuaciones:<sup>9</sup>

$$n_{k,t} = w_{k,0} + \sum_{i=1}^{i^*} w_{k,i} x_{i,t} \quad (2)$$

$$N_{k,t} = L(n_{k,t}) = \frac{1}{1+e^{-n_{k,t}}} \quad (3)$$

$$\hat{y}_t = \gamma_0 + \sum_{k=1}^{k^*} \gamma_k N_{k,t} \quad (4)$$

El sistema presenta  $i^*$  variables input y  $k^*$  neuronas. El potencial postsináptico en está determinado por la función de propagación de tipo sumatoria que combina cada uno de los inputs ( $x_{i,t}$ ) observados en el momento  $t$  con sus respectivos pesos ( $w_{k,i}$ ). Ambos elementos junto con el término constante  $w_{k,0}$  (conjunto de pesos iniciales) forman la variable  $n_{k,t}$  que posteriormente es achatada a través de una función de activación de tipo logística  $L(n_{k,t})$ , y representando la neurona  $N_{k,t}$  en el momento de observación  $t$ . Finalmente, el conjunto de  $k^*$  neuronas observadas en el tiempo  $t$ , se combinan linealmente con el vector de coeficientes  $\gamma_k$  y una constante  $\gamma_0$  que representan los pesos de las neuronas intermedias con el output. Como se aprecia, este conjunto de ecuaciones intenta aproximar el Proceso Generador de Datos (PGD) que subyace a la variable output con las variables de entrada o inputs.

Otro tipo de función de activación muy utilizada es la tangente hiperbólica cuyo rango varía entre  $[-1, 1]$  a diferencia de  $[0, 1]$  que era el rango de acción de la función logística. En este caso la ecuación 3 se vería de la siguiente manera:

$$N_{k,t} = T(n_{k,t}) = \frac{e^{n_{k,t}} - e^{-n_{k,t}}}{e^{n_{k,t}} + e^{-n_{k,t}}} \quad (5)$$

No existe una regla definida para la elección del tipo de función de activación. Algunos autores señalan la existencia de ciertas reglas heurísticas (Zhang G., Patuwo E., Hu M., 1997) como el uso de la función logística en estudios que involucran el aprendizaje del comportamiento normal entre las variables, como son los problemas de clasificación y la función tangente hiperbólica para aprendizaje de comportamientos que se desvían del promedio o implican variabilidad como es el pronóstico.

Las superficies ocultas juegan un rol importantísimo en la identificación de patrones ocultos en la información y para capturar las no linealidades existentes. Sin éstas, el sistema de ecuaciones descrito anteriormente equivaldría a un modelo de regresión lineal normal. No

---

<sup>9</sup> Este análisis se basa en el capítulo 2 de "Neural Networks in Finance, Gaining Predictive Edge in the Market" de Paul Mc Nelis.

obstante, al igual que en el caso de las funciones de activación no existe una regla para la determinación de la cantidad de superficies ocultas a ser utilizadas en el análisis de un problema o la cantidad de neuronas que éstas deben contener. No obstante, problemas de mayor complejidad requerirán muy probablemente mayor cantidad de capas ocultas y viceversa.

Una vez definida la estructura de la RNA, el problema fundamental se reduce al siguiente grupo de ecuaciones:

$$MIN \Psi(\Omega) = \sum_{t=1}^T (y_t - \hat{y}_t)^2 \quad (6)$$

Donde:

$$\hat{y}_t = f(x_t, \Omega) = \gamma_0 + \sum_{k=1}^{k^*} \gamma_k N_{k,t}$$

$$\Omega = \{w_{k,i}, \gamma_k\} \quad (7)$$

El objetivo fundamental de todo análisis con RNA consiste en minimizar el cuadrado de los errores entre el valor real de la variable output, que posee T observaciones y una función  $f(x_t, \Omega)$  que refleja a la red misma y que permitirá aproximar el PGD de la variable output dado por la ecuación 4. En el tenor de lo señalado algunos párrafos atrás, dicha función depende fundamentalmente de dos parámetros: el conjunto de variables input ( $x_{i,t}$ ) y el conjunto de pesos sinápticos ( $w_{k,i}, \gamma_k$ )

Existe un número importante de métodos o algoritmos que permiten estimar los parámetros del modelo de RNA minimizando 6 pero sobresalen principalmente: El método de retropropagación o basado en gradientes<sup>10</sup> locales, que consiste en computar primeras y segundas derivadas de la función  $\Psi$  con respecto a los parámetros del conjunto  $\Omega$ , partiendo de valores iniciales  $\Omega_0$  y actualizándose constantemente en función de criterios de parada o de convergencia previamente definidos. El método de búsqueda estocástica o también denominado *simulated annealing* que realiza actualizaciones aleatorias del conjunto inicial de pesos hasta que se alcance el criterio de parada definido. Finalmente el método de búsqueda estocástica evolucionaria o más conocido como algoritmos genéticos, que implica la actualización de los pesos iniciales utilizando la selección genética, *breeding* y mutación a través de una serie de generaciones hasta que el mejor conjunto de coeficientes sea encontrado.<sup>11</sup>

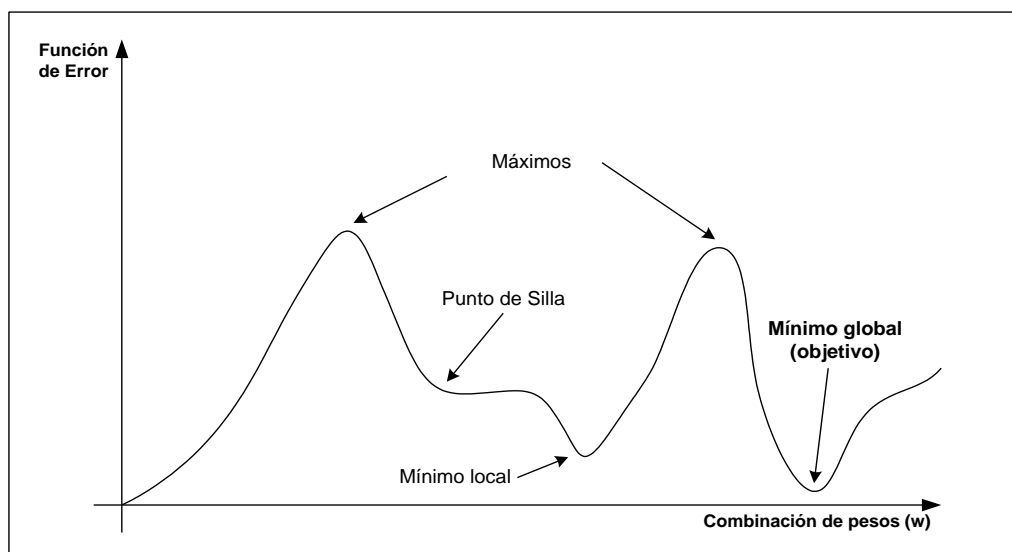
<sup>10</sup> Un gradiente se define como el vector de derivadas parciales de una función respecto a cada una de las variables. (Aristizábal M., 2006)

<sup>11</sup> Para mayor detalle sobre los tres métodos se sugiere al lector se sugiere al lector revisar el capítulo N° 3 de Neural Networks in Finance, Gaining Predictive Edge in the Market, sección N° 3.2.

En general, al proceso sistemático e iterativo de variación de los pesos sinápticos en un RNA originado en una estimulación del entorno (inputs), se conoce como aprendizaje y se distinguen principalmente tres tipos: aprendizaje supervisado, que implica fijar un objetivo o salida deseada e iterar sistemáticamente los diferentes parámetros de tal forma que la RNA alcance el valor deseado o minimice la diferencia entre el error generado entre el valor de output que la red genera y el real. Aprendizaje no supervisado o autoorganizado, que no requiere de un valor output específico deseado y se lo puede describir como: “la estimación de la función de densidad de probabilidad  $p(x)$  que describe la distribución de patrones  $x$  pertenecientes a un espacio de entrada a partir de muestras”. (Martin del Brío, Sanz M., 1997). Finalmente el aprendizaje reforzado en el que si bien no se provee un valor de salida deseado, sí se suministran a la red señales de error que le dan una idea general de su rendimiento diciéndole únicamente si el output generado va yendo en la dirección correcta, sin mayores detalles.

En general todos los métodos de entrenamiento/aprendizaje de las RNA buscan encontrar un error mínimo global, pero en vista de que esto se realiza en una superficie altamente no lineal, tal como se muestra en el Gráfico 2, la tarea no resulta fácil.

**Gráfico 2: COMPLEJIDAD EN LA BÚSQUDA DEL ERROR MÍNIMO GLOBAL**



**Fuente y Elaboración:** Propia

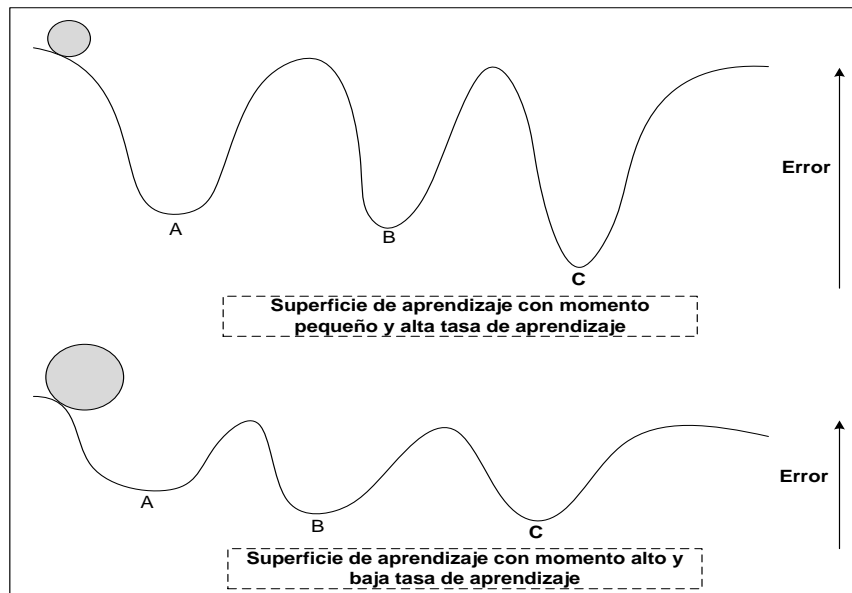
A medida que los pesos son ajustados durante el entrenamiento, existe la posibilidad de quedar atascado en cualquiera de los muchos puntos en los que la derivada equivale a cero o la curva presenta un área plana como un punto silla. En general, no existe un algoritmo que garantice encontrar el punto mínimo global en un problema no lineal dado, existen únicamente estrategias

que implican reestimaciones o búsquedas estocásticas comunes o evolucionarias como las mencionadas anteriormente.

Una forma de mejorar los métodos de descensos de gradientes constituye la incorporación dentro del esquema de un nuevo parámetro denominado momento (Zhang G., Patuwo E., Hu M., 1997), permitiendo que más grandes tasas de aprendizaje resulten en una convergencia más rápida minimizando la tendencia a oscilar. El término momento tiene como principal objetivo hacer que el próximo cambio del peso sináptico sea en alrededor de la misma dirección del cambio previo reduciendo así la oscilación.

Como explican Nisbet R., Elder J. y Miner G. (2009), entrenar una RNA es similar a un balón rodando a través de una serie de colinas y caídas. El tamaño del balón y especialmente la masa asociada a su peso representa el parámetro momento y la tasa de aprendizaje la pendiente de camino por donde el balón debe rodar. Esto implica que la utilización de un momento pequeño podría derivar en que la solución (el balón) quede fácilmente atascado en un mínimo local como el punto A o B del Gráfico 3.

**Gráfico 3: DINÁMICA DE BÚSQUEDA DEL MÍNIMO GLOBAL**



**Fuente y Elaboración:** Propia

El algoritmo no presenta suficiente tendencia a seguir buscando el mínimo global, subiendo a través de las colinas ya que éstas son demasiado pronunciadas por la tasa de aprendizaje alta (parte superior de la gráfica). Una modificación a este escenario utilizando momentos mayores y tasas de aprendizaje menores (parte inferior de la gráfica), incrementaría de forma importante la

probabilidad de encontrar el punto óptimo global, permitiendo que la rutina de minimización del error se mueva con mayor facilidad en la superficie de aprendizaje hasta encontrar el punto C que constituye el óptimo global.

Según Nisbet R., Elder J. y Miner G. (2009), en RNA configuradas manualmente, una tasa de aprendizaje del orden del 0.9 obtiene los mejores resultados, junto a momentos configurados entre 0.1 y 0.3. Sin embargo, como señala en Zhang G., Patuwo E., Hu M., 1997, el hecho de que se pueda definir un momento con valores entre 0 y 1 hace prácticamente imposible realizar una búsqueda exhaustiva para encontrar la mejor combinación, junto con la tasa de aprendizaje, implicando el diseño de RNA supervisadas una tarea que implica necesariamente cierto ensayo y error.

Otro parámetro que es necesario configurar adecuadamente es la tasa de disminución de aprendizaje, puesto que comúnmente la tasa de aprendizaje va decayendo a medida que se realiza el entrenamiento. Esto tiene el efecto importante de aplanar la superficie de búsqueda y de esta manera permitir encontrar los mínimos globales de una forma más fácil.

## **II.1 Análisis de sensibilidad y significancia con RNA**

Una de las principales desventajas que se suele señalar en la utilización de RNA es el hecho de que constituyen una suerte de “cajas negras” debido a que el entendimiento completo de las interrelaciones internas generadas entre inputs, el output y las neuronas ocultas no es posible. Shachmurove Y. (2002) señala este hecho como una indeterminación observacional que surge a partir de la naturaleza misma autopoietica o de autoorganización que presentan las RNA, que hace extremadamente complejo asimilar cómo las relaciones en las capas ocultas son estimadas. Sin embargo, la tendencia a considerar estos sistemas como cajas negras ha disminuido considerablemente los últimos años. Como señala Nisbet R., Elder J. y Miner G. (2009), estudios recientes están abriendo la caja negra en gran medida por el estudio del efecto o contribución que cada variable input presenta sobre el output a la luz principalmente del análisis de sensibilidad. Esta serie de avances está dando lugar a que en la actualidad se hable más bien de “cajas grises” ya que las RNA modernas, a partir de ciertas técnicas, ya pueden proveer una medida de importancia o significación de las diferentes variables inputs utilizadas para construir la red, permitiendo así, apreciar ciertos detalles sobre la operativa interna de cálculo.

El análisis de sensibilidad se refiere a cómo el output es afectado por perturbaciones en los inputs y/o en los pesos sinápticos (Yeung D., Cloete Ian, Shi D. Wing W.Y., 2010). Dichas

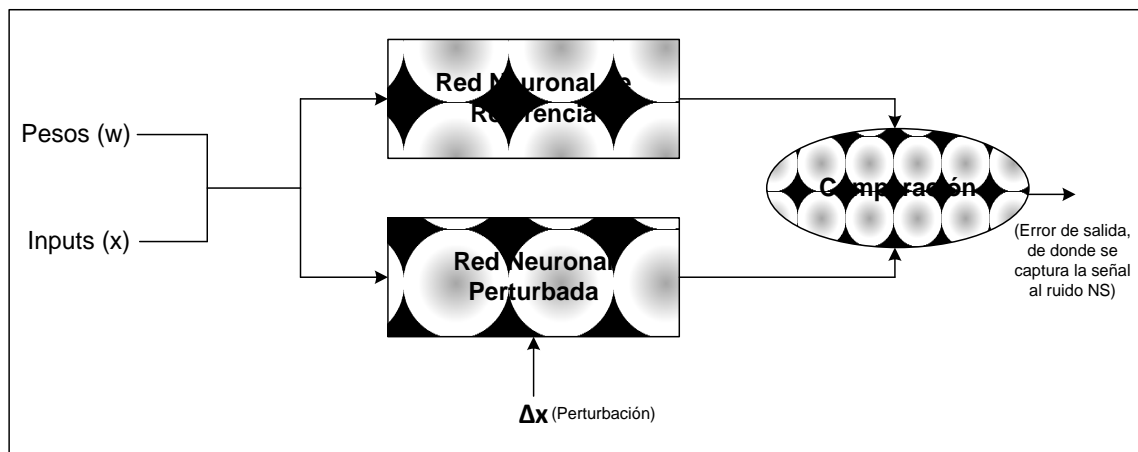
perturbaciones son ocasionadas comúnmente por imprecisiones o ruido en la información que para el análisis de sensibilidad son generadas a través de simulación y aplicadas a los inputs.

En forma analítica, si expresamos el efecto de una pequeña perturbación  $\Delta\Theta$  sobre una función de costo  $g$  a partir de la expansión de Taylor tendríamos lo siguiente:

$$g(\Theta + \Delta\Theta) = g(\Theta) + \frac{\Delta\Theta}{1!} g'(\Theta) + \frac{\Delta\Theta^2}{2!} g''(\Theta) + \dots \quad (8)$$

Obviamente el parámetro  $\Theta$  incluye al conjunto de inputs ( $x_i$ ) y los pesos ( $\Omega$ ) y la variación de la función se captura a través del concepto de señal al ruido NS (noise-to-signal). Como se aprecia, en este contexto las derivadas (variaciones marginales) juegan un rol muy importante para determinar el efecto de las perturbaciones y el análisis de sensibilidad permitirá definir cómo deben computarse para calcular la respuesta del sistema a las perturbaciones en los parámetros. La idea que subyace a este análisis se puede entender de mejor forma visualizando el siguiente gráfico.

**Gráfico 4: ANÁLISIS DE SENSIBILIDAD CON RNA**



**Fuente y Elaboración:** Propia con Martin del Brío, Sanz M., 1997

El análisis requiere de una RNA de referencia cuyos pesos hayan sido estimados y se mantengan en estado de recuerdo o constantes. A los inputs de ésta se le aplican una serie de shocks generando una nueva RNA perturbada cuyo nuevo output se analiza y contrasta con la red original. De esta manera se realiza un análisis comparativo entre las dos redes tomando en cuenta las variaciones que se han podido observar en el error del output. El presente trabajo muestra el resultado de aplicar pequeñas perturbaciones a los inputs de las mejores RNA construidas.

Si se toma en cuenta que una RNA opera en un mapping no lineal y diferenciable  $\Gamma: \mathbb{R}^I \rightarrow \mathbb{R}^K$  desde un vector de inputs  $x = (x_1, x_2, \dots, x_I)$  hasta un vector de outputs  $o = (o_1, o_2, \dots, o_K)$ . Supóngase además que  $x^{(n)} \in \Phi$ , donde  $\Phi$  es un conjunto abierto. Ya que  $o$  es diferenciable en  $x^{(n)}$  se tiene que:

$$o(x + \Delta x) = o(x^{(n)}) + J(x^{(n)})\Delta x + g(\Delta x)$$

Donde  $J(x^{(n)})$  representa la matriz Jacobiana definida como:

$$\begin{pmatrix} \frac{\partial o_1}{\partial x_1} & \frac{\partial o_1}{\partial x_2} & \dots & \frac{\partial o_1}{\partial x_I} \\ \frac{\partial o_2}{\partial x_1} & \frac{\partial o_2}{\partial x_2} & \dots & \frac{\partial o_2}{\partial x_I} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial o_K}{\partial x_1} & \frac{\partial o_K}{\partial x_2} & \dots & \frac{\partial o_K}{\partial x_I} \end{pmatrix}$$

Donde además se debe cumplir que:

$$\lim_{\Delta x \rightarrow 0} \frac{g(\Delta x)}{(\Delta x)} = 0$$

En este trabajo, la perturbación aplicada a cada input ( $\Delta x$ ) corresponde al valor de una desviación estándar de la variable input respectiva. Tal como se cita en Engelbrecht, 1999, el análisis de sensibilidad puede ser utilizado para una gran variedad de tareas como ser: optimización, robustez y análisis de estabilidad, generalización, medición de la no linealidad en el relacionamiento inputs/outputs, inferencia y causalidad, aprendizaje selectivo o reducción del conjunto de inputs, entre otros. Aquí se los utiliza para llevar a cabo una tarea de análisis inferencial y de causalidad que ciertas variables económico-financieras podrían tener sobre los niveles de inflación durante el periodo de tiempo estudiado.

### III. Aplicación de redes neuronales al caso de la inflación en Bolivia

#### III. 1 Observaciones generales sobre la base de datos

Para cada variable, se aplicaron las pruebas de Raíz Unitaria Augmented Dickey-Fuller (ADF), y la prueba de Phillips – Perron (P-P) porque esta última utiliza métodos estadísticos no paramétricos que evitan la correlación serial en los términos de error, sin que añadan términos de diferencia rezagados. Los resultados de las pruebas se resumen en la Tabla 3 de la siguiente manera:

**Tabla 3: PRUEBAS DE RAÍZ UNITARIA**

Variables	Estadísticos	Niveles	Valores Críticos MaCKinnon		
			1%	5%	10%
Tasa de Inflación (PI)	ADF	-4.9145	-3.4584	-2.8733	-2.5730
	P-P	-10.9320	-3.4580	-2.8731	-2.5729
Tipo de Cambio (TC)	ADF	-2.0376	-3.4582	-2.8732	-2.5729
	P-P	1.7331	-3.9978	-3.4290	-3.1377
Variación del TC (TTC)	ADF	-2.5532	-3.4583	-2.8733	-2.5730
	P-P	-5.4307	-3.9980	-3.4290	-3.1377
Índice Global de Actividad Económica (IGAE)	ADF	-3.6215	-3.9984	-3.4292	-3.1378
	P-P	-7.8452	-3.9978	-3.4290	-3.1377
Variación del IGAE (TIGAE)	ADF	-13.8073	-3.4585	-2.8734	-2.5730
	P-P	-17.5088	-3.4581	-2.8732	-2.5729
Emisión Monetaria (EMI)	ADF	4.4903	-3.4584	-2.8733	-2.5730
	P-P	5.2093	-3.4580	-2.8731	-2.5729
Variación de la EMI (TEMI)	ADF	-8.7344	-3.4585	-2.8734	-2.5730
	P-P	-27.1662	-3.4581	-2.8732	-2.5729
Agregado Monetario M'3 (M3)	ADF	3.6719	-3.4610	-2.8745	-2.5736
	P-P	4.9814	-3.4606	-2.8743	-2.5735
Variación del M3 (TM3)	ADF	-7.0095	-3.4610	-2.8745	-2.5736
	P-P	-13.9130	-3.4607	-2.8744	-2.5736

**Fuente y Elaboración:** Propia con la base de datos

Las pruebas de raíz unitaria señalan que para todas las variables, (excepto para el tipo de cambio nominal), se rechaza la hipótesis nula, es decir, se rechaza la presencia de raíz unitaria, por tanto estas variables son estacionarias<sup>12</sup>. El tipo de cambio nominal, que presenta raíz unitaria, es una variable no estacionaria debido al comportamiento que tuvo durante los últimos años, es decir, luego de presentar un prolongado periodo de depreciaciones, la variable

<sup>12</sup> HARRIS, SOLLIS (2003), "Applied time series modelling and forecasting", Capítulo 2 : Short – and long run models, pág. 29: "Si la variable tiene raíz unitaria, será no-estacionaria y explosiva"

comenzó a presentar apreciaciones y para después presentar otro semi-prolongado periodo de no variación seguido de más apreciaciones.

Adicionalmente, se calculó la Matriz de Correlación que se muestra en la Tabla 4, donde se puede observar que la diagonal principal, constituida por la correlación de una variable consigo misma da como resultado la unidad. Asimismo, la matriz muestra la correlación que tiene la variable dependiente con las independientes y las independientes entre sí. En este contexto, gran parte de las variables de la matriz se correlacionan a un nivel mayor a 0.5, con lo que se afirma la existencia de un problema de multicolinealidad.

**Tabla 4: MATRIZ DE CORRELACIÓN**

ARIABLES	PI	TC	TTC	IGAE	TIGAE	EMI	TEMI	M3	TM3
PI	1.0000								
TC	-0.1344	1.0000							
TTC	-0.1749	-0.5218	1.0000						
IGAE	-0.0967	0.7596	-0.6803	1.0000					
TIGAE	-0.2120	0.0024	0.0026	0.1618	1.0000				
EMI	0.0453	0.4740	-0.6742	0.8655	0.0150	1.0000			
TEMI	0.0984	0.0601	-0.1538	0.1857	0.3444	0.0855	1.0000		
M3	-0.0580	0.6331	-0.6349	0.9266	0.0111	0.9584	0.0538	1.0000	
TM3	0.1373	-0.1757	-0.0770	-0.0160	0.1088	0.0510	0.3368	-0.0336	1.0000

Asimismo, se utilizó el programa E – Views para describir estadísticamente las variables del estudio, como lo muestra la Tabla 4.

**Tabla 5: DESCRIPCIÓN ESTADÍSTICA DE LA BASE DE DATOS**

	PI	TC	TTC	IGAE	TIGAE	EMI	TEMI	M3	TM3
Media	0.4971	6.3077	0.0027	152.8838	0.0059	5,038,998	0.0159	30,324,405	0.0122
Mediana	0.4374	6.6129	0.0036	147.5470	0.0006	2,153,185	0.0121	27,249,969	0.0126
Máximo	3.0479	8.0800	0.0114	229.8400	0.1939	24,585,622	0.2587	84,382,319	0.1349
Mínimo	-2.5473	3.7676	-0.0148	96.3066	-0.1693	666,564	-0.2299	5,490,858	-0.1222
Desvío Estándar	0.6580	1.3282	0.0043	31.4626	0.0636	5,664,263	0.0759	19,021,572	0.0223
Variabilidad	1.3237	0.2106	1.5862	0.2058	10.8517	1.1241	4.7725	0.6273	1.8180
Asimetría	0.4945	-0.2617	-1.0706	0.4672	0.1232	1.5272	0.1434	1.0709	-0.2590
Curtosis	6.4391	1.7160	5.1039	2.3696	3.8975	3.9412	5.4317	3.4599	11.8840
Jarque-Bera	121.12	18.18	85.23	12.02	8.19	96.62	56.71	45.39	749.05
Probabilidad	0.0000	0.0001	0.0000	0.0025	0.0166	0.0000	0.0000	0.0000	0.0000

**Fuente y Elaboración:** Propia con la base de datos

De acuerdo con la Tabla 5, las variables presentan características que hacen poco recomendable un análisis econométrico tradicional; entre algunas características están:

- Asimetría: Las variables presentan asimetría para todos los casos, aún tratándose de la tasa de variación de algunas de ellas, ya sea mediante valores positivos o negativos, existe una asimetría que las aleja de tener una distribución normal.
- Curtosis: La curtosis es elevada para la mayor parte de las variables, lo cual también refleja y confirma que las variables no tienen distribución normal.
- Jarque-Bera y Probabilidad: El estadístico Jarque-Bera, para todas las variables presenta valores discordantes, que acompañados del valor de las probabilidades, las cuales deberían tender a uno, corroboran que las variables no son candidatas para efectuar un análisis econométrico.

Por todo lo anteriormente descrito, las redes neuronales se convierten en la alternativa oportuna para realizar el análisis de la inflación, dado que estadísticamente, no se cumplen los supuestos suficientes y necesarios de consistencia técnica para efectuar otro tipo de análisis convencional de naturaleza econométrica.

### **III. 2 Aplicación de las RNA al estudio de inflación.**

La presente investigación ha utilizado como input de las RNA construidas información mensual de la variación mensual del Índice de Precios al Consumidor (IPC), conocida como Tasa de Inflación (PI). Como variables output se ha hecho uso de la variación mensual del tipo de cambio de compra TTC, el Índice Global Actividad Económica (IGAE) y la Emisión Monetaria (EMI), mismas que, como se pudo advertir en la sección precedente, se encuentran muy ligadas a la inflación. Asimismo se realizó un análisis de significancia estadística ex-ante generando algunos modelos estadísticos e identificando a través de ellos aquellos rezagos de la variable inflación que podrían ser significativos,<sup>13</sup> de esta manera se eligieron los rezagos 1 (PI1) y 6 (PI6) como variables input adicionales.

En general conforme el análisis del grado de integración de las variables input presentado anteriormente, únicamente el TC presentó raíces unitarias, por tal motivo se trabajó con su diferencia logarítmica (TTC). El resto de las variables constituyen procesos estocásticos estacionarios y se los incorporó en la RNA en niveles. Individualmente todas las variables en la prueba econométrica fueron estadísticamente significativas al 1%, exceptuando la TTC. El

---

<sup>13</sup> Los modelos se pueden apreciar en el Anexo A

realizar un análisis econométrico de las relaciones entre las variables, constituye únicamente un ejercicio destinado a conocer de forma adelantada la capacidad de los inputs para explicar el output, es decir, una guía inicial. Sin embargo, la metodología de RNA utiliza una modelización diferente y sus resultados pueden fácilmente discrepar de los encontrados a la luz de la econometría.<sup>14</sup>

Antes de comenzar el entrenamiento y diseño de las RNA, las variables elegidas fueron normalizadas restando cada observación por su media y dividiendo el resultado por la desviación estándar. Posteriormente fueron escalonadas a una dimensión que variaba entre -1 y 1, debido fundamentalmente a que se utilizaron funciones de activación de tipo tangente hiperbólico.

Se ha utilizado la topología denominada Perceptrón Multicapa (MLP) que es de más amplio uso en las diferentes aplicaciones financieras de RNA. Se utilizaron los cinco inputs explicados anteriormente, una capa oculta con cuatro neuronas y la tasa de inflación como output. De esta forma, la estructura que presenta es de tipo (5,4,1). Siguiendo el mismo camino que Aristizábal M., 2006, el total de unidades escondidas (4) se determinó realizando un balance entre la capacidad de captura del comportamiento no lineal de las variables y evitar problemas de sobreentrenamiento. De las 252 observaciones para cada variable, en todos los casos se destinó 80% de la información para entrenamiento, 10% para validación cruzada (VC) y 10% para realizar un test a la capacidad de generalización de la RNA. Si se toma en cuenta que una RNA con una estructura suficientemente grande puede proporcionar un error de aprendizaje tan pequeño como se desee puesto que tiene la capacidad de aprender de la información, la eficacia real o capacidad de generalización, sin embargo, puede decaer a medida que el entrenamiento se extiende. Este problema se lo conoce como sobreaprendizaje y una forma de evitarlo consiste precisamente en entrenar y validar simultáneamente la RNA hasta encontrar y detenerse en un punto óptimo en el que el error de generalización sea el mínimo. A continuación se presentan los resultados de las dos RNA que se construyeron.

En la primera fila de la tabla se detallan los números de repetición exactos en que se generaron esos sinápticos que arrojaron errores MSE mínimos, mismos que son detallados en las filas debajo. Como se explicó, también se utilizó un conjunto de datos para realizar un test que permita determinar la capacidad de generalización de las RNA. Necesario resaltar el hecho antes de la determinación de los grupos entrenamiento, validación cruzada y test, los datos

---

<sup>14</sup> Como se aprecia en el Anexo A, donde se detallan los modelos estadísticos estimados, a pesar del alto nivel de significatividad individual de las variables, el modelo en su conjunto presenta un coeficiente de determinación muy bajo lo que refleja un poder explicativo global bajo.

fueron ordenados de forma aleatoria, de tal manera que los tres grupos dispongan de información correspondiente a todos los periodos analizados que en este estudio representan 20 años de historia y de esta manera durante el análisis de sensibilidad se pueda capturar eficientemente el relacionamiento de las variables.<sup>15</sup>

**Tabla 6: DESCRIPCIÓN DE LAS RNA CONSTRUIDAS**

<i>Regla de Aprendizaje</i>	<i>LEVENBERG-MARQUARDT (L-M)</i>		<i>QUICKPROPAGATION (QP)</i>	
<i>ENTRENAMIENTO</i>	<i>SET DE ENTRENAMIENTO</i>	<i>SET DE VALIDACIÓN CRUZADA</i>	<i>SET DE ENTRENAMIENTO</i>	<i>SET DE VALIDACIÓN CRUZADA</i>
Repetición	222	19	10000	9997
Error Cuadrático Medio Mínimo (MSEM) <sup>16</sup>	0.0057	0.0051	0.0137	0.0069
Error Cuadrático Medio FINAL (MSE)	0.0057	0.0170	0.0137	0.0089

#### **Análisis de sensibilidad de la variable Inflación a las variables inputs.**

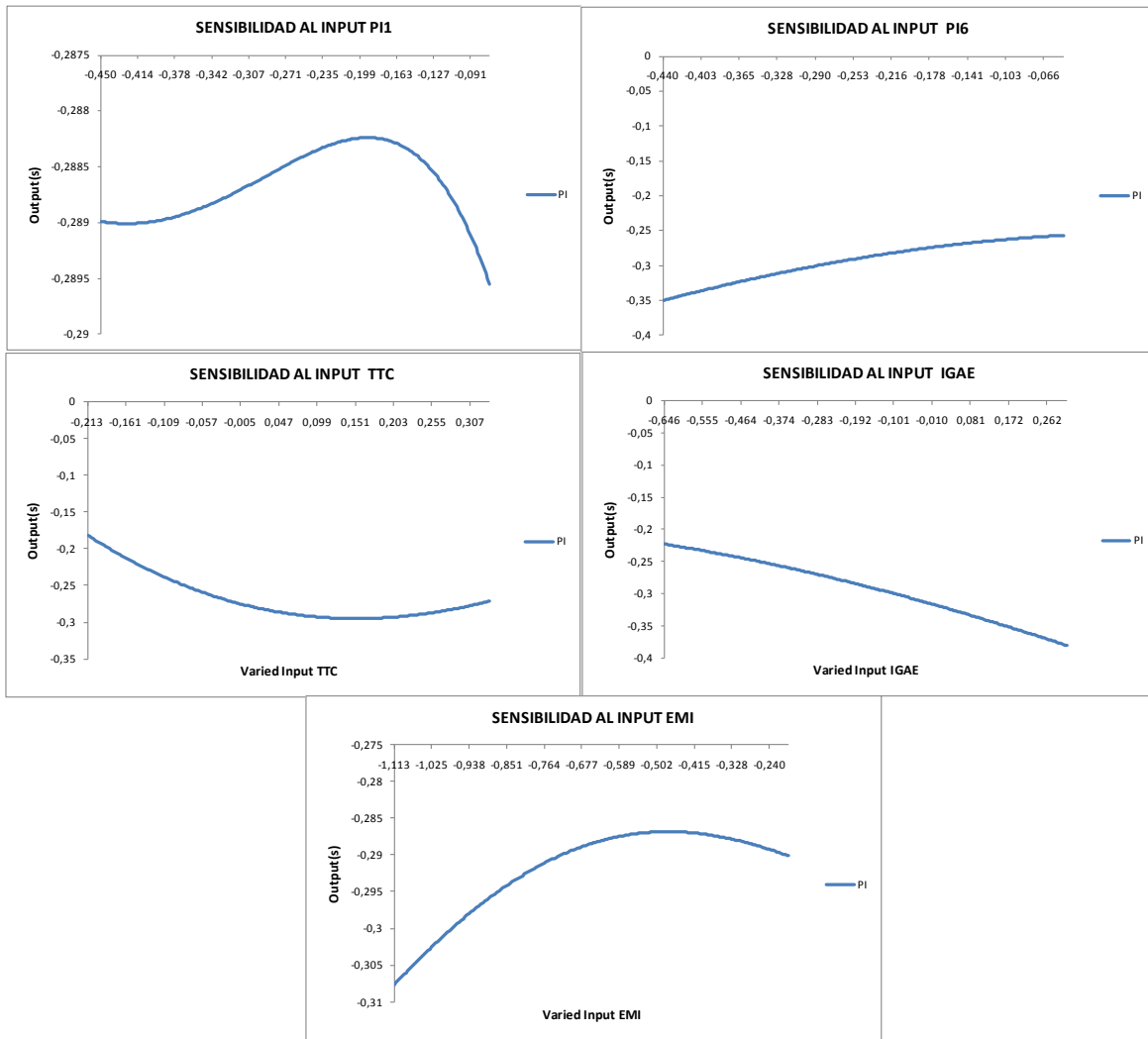
Como se mencionó desde un inicio, el objetivo fundamental de este trabajo de investigación fue el de analizar la inflación a partir de una serie de variables admitidas de forma general como sus principales determinantes. En esta sección se explican los resultados encontrados a partir de aplicar un choque equivalente a un desvío estándar en cada uno de los inputs sobre la inflación, midiendo la sensibilidad que ésta presenta a sus variables determinantes, además del tipo de variación que causan.

El eje horizontal de las gráficas mide el valor de la variable output (inflación) cuando se aplica un choque (equivalente a un desvío estándar) a un input, manteniendo los demás inputs constantes y utilizando los mejores pesos sinápticos encontrados durante la etapa de entrenamiento. Como se aprecia este efecto tiende a mostrar una senda que refleja la relación que subyace entre las dos variables.

<sup>15</sup> Las estadísticas correspondientes a la realización del test se presentan en el Anexo B

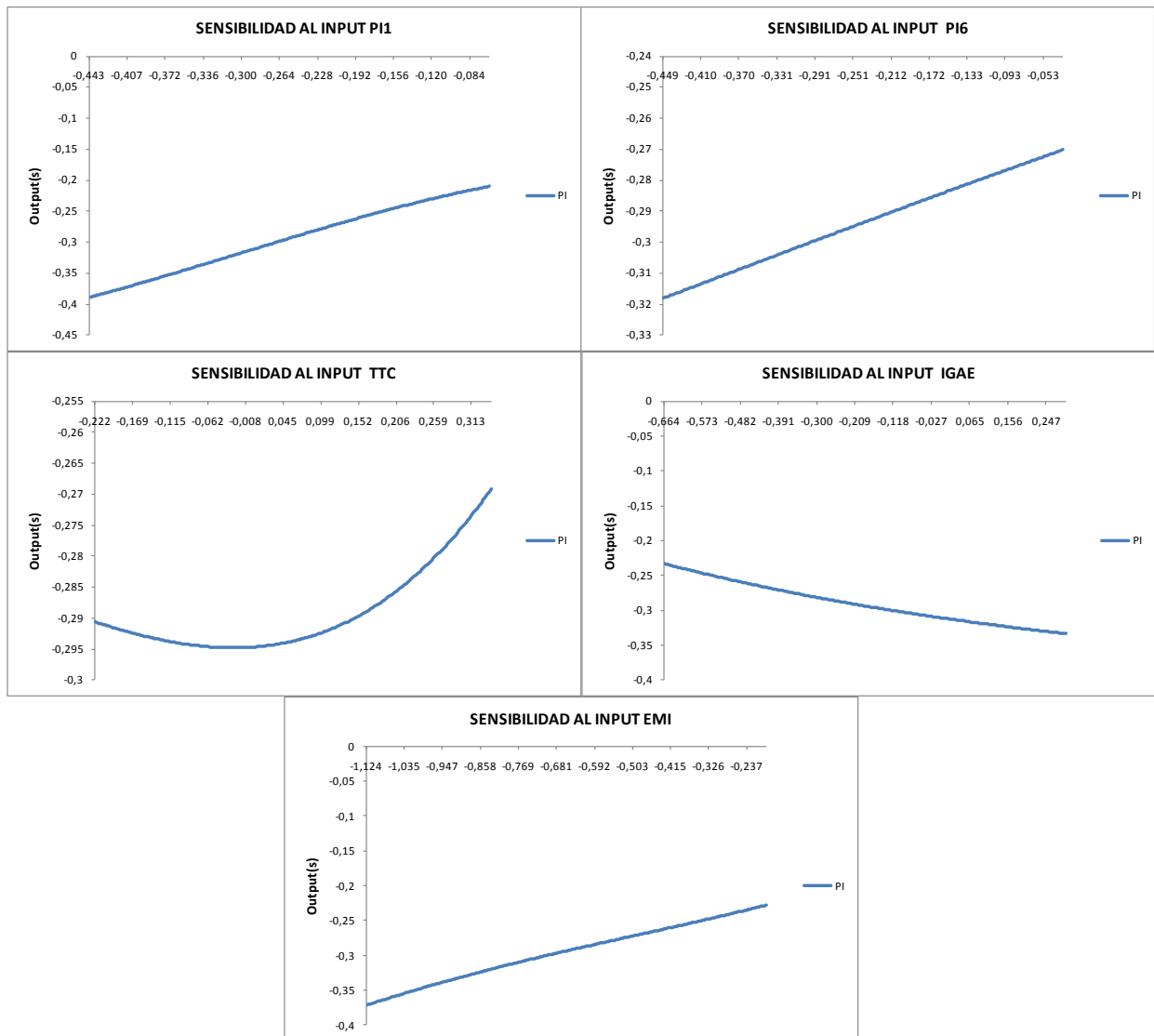
<sup>16</sup> El Error Cuadrático Medio se calcula mediante la siguiente fórmula:  $\frac{\sum_{j=0}^p \sum_{i=0}^n (d_{ij} - y_{ij})^2}{n * p}$  donde p número de neuronas output, n total de observaciones y<sub>ij</sub> output del ejemplar i al procesar el elemento j y d<sub>ij</sub> output objetivo para el ejemplar i al procesar el elemento j.

**Gráfico 5: ANÁLISIS DE SENSIBILIDAD DE LA INFLACIÓN**  
(RNA entrenada con LEVENBERG-MARQUARDT)



En general, el signo de la variación de la Inflación (PI) ante choques en las variables coincide con los resultados encontrados en estudios relacionados que fueron revisados en una sección anterior pero que utilizan la metodología econométrica. El impacto que tiene la inflación en periodos anteriores (PI1 y PI6) sobre la inflación actual es positivo, aunque tiende a disminuir en el tiempo. La senda que sigue la inflación ante choques en la variable IGAE refleja la relación intrínsecamente negativa existente entre ambas. De igual manera, la emisión muestra que tiene un impacto en general positivo que va disminuyendo paulatinamente.

**Gráfico 6: ANÁLISIS DE SENSIBILIDAD DE LA INFLACIÓN**  
**(RNA entrenada con QUICKPROPAGATION)**



Analizando la sensibilidad de la inflación con la RNA entrenada con el método Quickpropagation, los resultados son similares, aunque levemente diferentes en lo que respecta a la senda que sigue esta variable ante choques en los inputs. Se observa también una respuesta positiva con respecto a los valores pasados de la inflación, una relación negativa con respecto al indicador de actividad económica como es el IGAE y una respuesta positiva a incrementos en la emisión monetaria. Es muy probable que las sendas de comportamiento a los diferentes choques no sean iguales en forma (pero si en signo) debido a que los métodos de aprendizaje difieren y por lo tanto las RNA están capturando estructuras de relación entre las variables que ponderan más algunos aspectos en relación a otros.

El método L-M es conocido como un método de aprendizaje adaptativo de cuasi segundo orden en vista de que, a diferencia de los métodos de primer orden como el QP, hace uso de la matriz Hessiana o de segundas derivadas parciales para determinar la actualización de los pesos utilizando la curvatura en vez de solamente la pendiente de la superficie de rendimiento. Las reglas de primer orden utilizan solamente una aproximación local de la pendiente de esta superficie para determinar la mejor dirección de movimiento de los pesos sinápticos. Este hecho puede que explique la sensibilidad encontrada para el caso de la variación del tipo de cambio (TTC). En ambas RNA se encontraron sensibilidades relativamente negativas en un inicio pero que paulatinamente tienden a hacerse positivas. Esto concuerda con los estudios revisados en una sección anterior en el que en ciertos casos variables como la inflación o la inflación subyacente mostraron una relación positiva con la variación del tipo de cambio, pero como se muestra en Leiderman, Maino y Parrado (2005) el signo de la relación también era negativo.

#### **IV. Conclusiones**

Dado que las series económicas han presentado más relaciones no lineales que lineales en las variables de estudio, el empleo de RNAs representa la alternativa para obtener resultados más confiables, en un contexto que no cumple con las exigencias de los modelos econométricos tradicionales.

Se exploró la relación entre la inflación y sus determinantes: el tipo de cambio nominal, la emisión monetaria, el IGAE y la inflación rezagada en uno y seis periodos, a través de dos RNA, encontrando resultados que coinciden con los hallados con herramientas estadísticas de uso más generalizado. Dicho análisis permitió capturar relaciones de tipo no lineal entre las variables, mismos que se aprecian en las diferentes sendas de sensibilidad que presenta la inflación a choques en sus variables determinantes, encontrados en la tercera sección.

Es necesario profundizar en la investigación de las relaciones capturadas al interior de las RNA, cuyo operativa y características internas se conocen comúnmente como “caja negra”, pero que a la luz del análisis de sensibilidad se hace más viable su interpretación y entendimiento de las estructuras que subyacen a la interrelación entre inputs y outputs.

## Referencias Bibliográficas

Orellana W., Requena J. (1999). *“Determinantes de la Inflación en Bolivia”*

Larraín, Sachs *“Macroeconomía en la economía global”* Segunda Edición

Comboni, De La Viña (2002), *“Precios y Tipo de Cambio en Bolivia: Evidencia Empírica del Período de Post-Estabilización”*

Liederman, Maino, Parrado (2005), *“Metas de Inflación en Economías Dolarizadas”*

Mendieta, Cerezo, Cossío (2008), *“¿La inflación está de vuelta en Sudamérica? Choques exógenos, expectativas y credibilidad de la política monetaria”*

Cossío, Laguna, Martín, Mendieta, Mendoza, Palmero y Rodríguez (2007) *“La Inflación y Políticas del Banco Central de Bolivia”*, Jornadas Monetarias del BCB

McNeils P. (2005), *“Neural Networks in Finance, Gaining Predictive Edge in the Market”*

Aristizábal M.C. (2006) *“Evaluación asimétrica de una red neuronal artificial: aplicación al caso de la inflación en Colombia”*

Harris R., Sollis R. (2003), *“Applied time series modeling and forecasting”*

**ANEXO A**  
**MODELOS ECONOMÉTRICOS INICIALES ESTIMADOS**

Variable Dependiente: PI  
Sample(adjusted): 1990:07 2010:12  
Included observations: 246 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.399312	0.504347	2.774503	0.0060
PI(-1)	0.264294	0.063629	4.153700	0.0000
PI(-6)	0.178713	0.061683	2.897293	0.0041
IGAE	-0.008890	0.003390	-2.622410	0.0093
EMI	4.82E-08	1.71E-08	2.826146	0.0051
TTC	6.824037	14.59421	0.467585	0.6405
R-squared	0.221091	Mean dependent var		0.576057
Adjusted R-squared	0.204863	S.D. dependent var		0.781969
S.E. of regression	0.697285	Akaike info criterion		2.140843
Sum squared resid	116.6895	Schwarz criterion		2.226339
Log likelihood	-257.3237	F-statistic		13.62463
Durbin-Watson stat	1.978623	Prob(F-statistic)		0.000000

VARIABLE DEPENDIENTE: PI  
Sample(adjusted): 1990:07 2010:12  
Included observations: 246 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.480958	0.414911	3.569338	0.0004
PI(-1)	0.270235	0.063101	4.282545	0.0000
PI(-6)	0.176068	0.060151	2.927101	0.0038
IGAE	-0.009230	0.003046	-3.030503	0.0027
EMI	4.61E-08	1.71E-08	2.704334	0.0073
D(TTC)	25.63456	20.30558	1.262439	0.2080
R-squared	0.225524	Mean dependent var		0.576057
Adjusted R-squared	0.209389	S.D. dependent var		0.781969
S.E. of regression	0.695298	Akaike info criterion		2.135135
Sum squared resid	116.0253	Schwarz criterion		2.220631
Log likelihood	-256.6216	F-statistic		13.97739
Durbin-Watson stat	1.980206	Prob(F-statistic)		0.000000

## ANEXO B ANÁLISIS DE LA CAPACIDAD DE GENERALIZACIÓN DE LAS RNA

<b>LEVENBERG-MARQUARDT</b>		<b>QUICKPROPAGATION</b>	
<b>Performance</b>	<b>PI</b>	<b>Performance</b>	<b>PI</b>
MSE	0.0296	MSE	0.0334
NMSE	0.8374	NMSE	0.8324
MAE	0.1384	MAE	0.1188
Min Abs Error	0.0089	Min Abs Error	0.0061
Max Abs Error	0.3582	Max Abs Error	0.7110
r	0.4610	r	0.4258

