

**MODELADO DEL PRECIO DEL ORO: APLICACIÓN DE MODELOS CLÁSICOS
DE SERIES DE TIEMPO VS. REDES NEURONALES ARTIFICIALES**

LEIDY SAMARA GUERRERO GONZÁLEZ

UNIVERSIDAD DEL VALLE
FACULTAD DE CIENCIAS SOCIALES Y ECONÓMICAS
PROGRAMA ACADÉMICO DE ECONOMÍA
SANTIAGO DE CALI

2010

**MODELADO DEL PRECIO DEL ORO: APLICACIÓN DE MODELOS CLÁSICOS
DE SERIES DE TIEMPO VS. REDES NEURONNALES ARTIFICIALES**

LEIDY SAMARA GUERRERO GONZÁLEZ

Trabajo de Grado para optar por el Título de
Economista

Director
Jhon Alexander Méndez Sayago
Magister en Economía

UNIVERSIDAD DEL VALLE
FACULTAD DE CIENCIAS SOCIALES Y ECONÓMICAS
PROGRAMA ACADÉMICO DE ECONOMÍA
SANTIAGO DE CALI

2010

CONTENIDO

	Pág.
INTRODUCCIÓN.....	1
1. REVISIÓN DE LITERATURA.....	3
1.1 REVISIÓN DE LITERATURA EN EL ÁMBITO INTERNACIONAL.....	3
1.2 REVISIÓN DE LITERATURA EN EL ÁMBITO NACIONAL.....	5
2. MARCO TEÓRICO.....	8
2.1 EL ORO EN LA ECONOMÍA MUNDIAL.....	8
2.2 MODELOS DE PREDICCIÓN.....	9
2.2.1 Caminata aleatoria.....	10
2.2.2 Técnica Box-Jenkins.....	11
2.2.2.1 Proceso autorregresivo $AR(p)$	11
2.2.2.2 Proceso de media móvil $MA(q)$	12
2.2.2.3 Proceso autorregresivo y de media móvil $ARMA(p,q)$	12
2.2.3 Modelo de redes neuronales artificiales de tipo Perceptrón Multicapa (MLP).....	12
2.2.3.1 Estructura de la red.....	13

2.2.3.2 Funciones de activación.....	15
2.2.3.3 Características de las redes neuronales artificiales.....	15
2.2.3.4 Proceso de enseñanza y pesos.....	15
2.2.3.5 Aprendizaje supervisado.....	16
2.2.3.6 Algoritmo de aprendizaje.....	17
2.2.3.7 Tamaño de la red.....	17
2.2.3.8 Duración del entrenamiento.....	17
 3. METODOLOGÍA.....	 19
 3.1 ANÁLISIS DESCRIPTIVO.....	 19
3.1.1 Descripción de la muestra, variables del modelo, selección de los datos y softwar's.....	19
 3.2 ANÁLISIS INFERENCIAL.....	 20
3.2.1 Estimación e interpretación de resultados de los modelos de series de tiempo: Caminata Aleatoria con deriva y ARMA(p,q).....	20
 3.2.2 Estimación e interpretación de resultados de los modelos Perceptrón Multicapa de Redes Neuronales Artificiales	24
 4. CONCLUSIONES.....	 32
 REFERENCIAS BIBLIOGRÁFICAS.....	 34
 ANEXOS.....	 38

LISTA DE TABLAS

	Pág.
Tabla1. Estimación de modelos de series de tiempo.....	21
Tabla 2. Criterios de comparación modelos lineales	22
Tabla 3. Medidas de evaluación de predicción.....	23
Tabla 4. Serie original y Predicción ARIMA(4,1,4).....	23
Tabla 5. Criterios de comparación modelos Perceptrón Multicapa.....	26
Tabla 6. Medidas de evaluación de predicción modelos Perceptrón Multicapa.....	28
Tabla 7. Serie original y Predicción Modelo 1 de RNA.....	28
Tabla 8. Comparación de tendencias.....	30
Tabla A1. Correlograma serie lprecio.....	40
Tabla A2. Correlograma serie dlprecio.....	41
Tabla A3. Test Dickey Fuller Aumentado, intercepto.....	42
Tabla A4. Test Dickey Fuller Aumentado, ninguno.....	42
Tabla A5. Test Dickey Fuller Aumentado, intercepto y tendencia.....	43
Tabla A6. Test Dickey Fuller Aumentado, intercepto.....	43
Tabla A7. Test Dickey Fuller Aumentado, ninguno.....	43
Tabla A8. Test Dickey Fuller Aumentado, intercepto y tendencia.....	44
Tabla A9. Test de comprobación de modelos.....	44
Tabla A10. Test Ljung Box de autocorrelación ARIMA(2,1,2).....	45
Tabla A11. Test Ljung Box de autocorrelación ARIMA(4,1,4).....	46
Tabla A12. Test Ljung Box de autocorrelación Caminata Aleatoria.....	47

LISTA DE FIGURAS

	Pág.
Figura 1. Arquitectura de una red MLP con una capa oculta	14
Figura 2. Serie original vs. Predicción ARIMA(4,1,4).....	24
Figura 3. Serie original vs. Predicción Modelo 1 de RNA.....	29
Figura A1. Rango-Media.....	38
Figura A2. Comportamiento serie lprecio.....	39
Figura A3. Comportamiento serie dlprecio.....	41
Figura A4. Conjunto de datos de entrenamiento y validación.....	48
Figura A5. Rendimiento del entrenamiento, validación y test.....	48
Figura A6. Ajuste de entrenamiento, validación y test en conjunto.....	49

LISTA DE ANEXOS

	Pág.
Anexo A. Prueba rango-media y test Dickey Fuller	38
Anexo B. Comprobación de supuestos de los modelos.....	44
Anexo C. Indicadores del mejor modelo seleccionado de RNA: Rendimientos y ajuste...	48

RESUMEN

El objetivo de este trabajo de investigación es analizar el ajuste y nivel de predicción que presentan diferentes modelos lineales y no lineales sobre el precio del oro cotizado en la bolsa de Nueva York. Serán empleados los modelos lineales de Caminata Aleatoria y diversas especificaciones ARIMA de la metodología Box-Jenkins (1970) y modelos no lineales Perceptrón Multicapa de Redes Neuronales Artificiales. Los datos empleados son tomados de la página web www.kitco.com en enero de 2010, para el periodo que transcurre del 2 de enero del 2002 al 29 de diciembre del 2005. Como principal resultado se encontró que el modelo Perceptrón Multicapa presenta mejores medidas de evaluación Error Absoluto Medio (MAE) y Raíz del Error Cuadrático Medio (RMSE), pero la especificación ARIMA(4,1,4) presenta una mayor cantidad de aciertos al modelar la tendencia.

Palabras claves: Series de tiempo, caminata aleatoria, metodología Box-Jenkins, perceptrón multicapa, redes neuronales artificiales.

Clasificación JEL: C22, C45,

INTRODUCCIÓN

El oro es considerado como un importante refugio en tiempos de crisis para muchos inversionistas alrededor del mundo por ser un activo sólido y confiable que se convierte en una vía alterna de inversión cuando cae una moneda fuerte, cuando caen los precios de otros *commodities*, por factores geopolíticos y por ser una fuente de protección en tiempos inflacionarios (Bertolotto, 2007). Sin embargo, el precio del oro está influenciado por diversos fenómenos y variables que actúan de forma sorpresiva y conjunta a nivel mundial (problema de multicolinealidad), ya que existe dependencia del precio del oro con otras variables y éstas a su vez dependen del precio del oro (problema de endogeneidad), al igual que las acciones de los agentes resultan ser impredecibles y la estructura del mercado es competitiva.

Lo anterior lleva a pensar en la gran importancia que tiene para las diversas economías refugiarse en un *commodity* como el oro, pero a su vez, es notorio el riesgo al que se ven enfrentados a la hora de realizar la inversión, por lo cual surgen las siguientes preguntas: ¿Es posible controlar las variables que afectan el precio del oro? ¿Puede conocerse cuál será el precio del oro que regirá el día de mañana? Si puede predecirse dicho precio ¿Cuál es el mejor método para hacerlo? En primer lugar, no es posible controlar las variables que afectan el precio del oro por estar en un mercado competitivo donde las acciones de un agente no determinan el precio de un bien, además, su precio se ve afectado por un gran conjunto de variables, que van desde el punto de vista de la oferta y la demanda del oro hasta las decisiones que realiza un presidente de una economía influyente.

Activos tan interesantes por su seguridad, pero tan fluctuantes en su precio como el oro son atractivos para inversionistas de todo el mundo, los cuales están muy inmersos en el tema de las finanzas internacionales y sin embargo es para ellos un desafío lograr predecir el precio de este activo. Emplean un conjunto de técnicas a través del análisis técnico y fundamental y, aunque las dos metodologías son muy importantes y han logrado resultados muy satisfactorios, no puede establecerse una regla general o un modelo determinado que logre con exactitud y veracidad arrojar un resultado exacto. En este orden de ideas, es posible plantear modelos lineales de series de tiempo y no lineales de Redes Neuronales

Artificiales con el fin de observar el desempeño que presentan al realizar predicción de corto plazo de un paso adelante y de esta manera, contribuir a minimizar el riesgo de los inversionistas y los individuos interesados en el tema. Aunque los modelos ARMA no incorporan información previa (teoría económica por ejemplo), en la práctica tiene un amplio uso ya que permite la obtención de proyecciones razonables en el corto plazo (Aranibar y Humérez, 1996) y las RNA por su parte, son capaces de aprender de la experiencia, de generalizar de casos anteriores a nuevos casos, de abstraer características esenciales a partir de entradas que presentan información irrelevante, entre otros, lo cual hace de estos modelos una herramienta muy útil en la predicción de corto plazo (Richter y Toro, 2006).

Los modelos clásicos de series de tiempo que serán empleados son la Caminata Aleatoria y diferentes especificaciones ARIMA de la metodología Box-Jenkins (1970) y modelos no lineales de tipo Perceptrón Multicapa de Redes Neuronales Artificiales que tienen el enfoque insumo-producto de Mc. Culloch y Pitts (1943). Con la aplicación de dichas metodologías estadísticas, el objetivo del trabajo de investigación se dirige a analizar el ajuste y nivel de predicción que presentan diferentes modelos lineales y no lineales sobre el precio del oro cotizado en la bolsa de Nueva York, donde las medidas de evaluación empleadas serán el Error Absoluto Medio (MAE) y Raíz del Error Cuadrático medio (RMSE)

Las técnicas lineales han sido utilizadas y desarrolladas ampliamente en el transcurso del tiempo dentro del ámbito financiero, lo cual hace que la bibliografía y software sean abundantes para este tipo de técnicas por su fácil aplicación. Sin embargo, las técnicas no lineales y específicamente las Redes Neuronales Artificiales han ido ganando terreno por ser sistemas novedosos que permiten un alto grado de predictibilidad al eliminar los problemas típicos de no linealidad en la predicción de activos financieros y lograr un buen ajuste con la propuesta de especificación de modelos lineales (Fríz, 2004).

El presente trabajo se organiza de la siguiente manera: En la sección 1 se realiza una revisión de literatura a nivel nacional e internacional, en la sección 2 se presentan los modelos y el sustento teórico al problema de investigación. Posteriormente, en la sección 3 se presenta la descripción de los datos, variables, estimación de modelos y análisis de resultados y finalmente en la sección 4 se presentan las conclusiones y bibliografía.

1. REVISIÓN DE LITERATURA

A continuación se exponen los trabajos más importantes sobre la predicción de precios y volatilidades de *commodities* como el oro a partir de las técnicas Box-Jenkins y Redes Neuronales Artificiales.

1.1 REVISIÓN DE LITERATURA EN EL ÁMBITO INTERNACIONAL

La hipótesis del mercado eficiente establece que *en cualquier momento, el precio de una acción refleja plenamente toda la información que se encuentra contenida en dicho activo* mientras que, por otro lado, la teoría del Caos intenta demostrar que el mercado de valores es esencialmente aleatorio; temas abordados por Aranibar y Humérez (1996) quienes intentan predecir el comportamiento de 9 minerales por medio de la técnica Box-Jenkins y el trabajo de Lawrence (1997); el cual realiza un paralelo de diferentes técnicas de predicción. La primera investigación encuentra un comportamiento de caminata aleatoria para todos los *commodities* y define a los mercados como eficientes, mientras que Lawrence (1997) desvirtúa la hipótesis del mercado eficiente argumentando que el comportamiento de las acciones, por ejemplo, es posible predecirlo por medio de las RNA ya que son sistemas muy completos capaces de modelar sistemas caóticos.

El artículo de Foster (2002) dirige el objetivo principal de su investigación a encontrar un método de inteligencia artificial que prediga los valores de activos futuros a través de la modelación del precio de la soja por medio de 3 enfoques de inteligencia artificial: las redes neuronales artificiales, los algoritmos genéticos y el híbrido de redes neuronales con algoritmos genéticos, teniendo en cuenta los mismos datos para la prueba y validación. Los resultados arrojados por esta investigación revelan que las redes neuronales puras presentan una predicción mucho mejor que los modelos híbrido, aunque la relación horas-hombre es mucho mayor en las redes; otro factor interesante es la cantidad de recursos empleados por

la computadora, el cual es un aspecto contribuye a catalogar al modelo de algoritmos genéticos como el mejor.

Friz (2004) bajo la dirección de Antonino Parisi, realiza un estudio de redes neuronales para modelar el precio del *commodity*¹ oro empleando la red de tipo Perceptrón Multicapa y las técnicas de funcionamiento *rolling*² y recursivo, donde las variables explicativas son el precio del oro y el índice Dow-Jones rezagados y es empleada una técnica llamada *bootstrap*³ para comprobar la robustez del modelo. Los resultados revelan que las redes neuronales presentan un nivel de predicción mayor con la técnica *rolling* que con la técnica recursiva, esto es un 73,05% de predictibilidad contra un 64,39% respectivamente. Gutiérrez (2004) también bajo la dirección de Antonino Parisi, evalúa la rentabilidad que habría obtenido un inversionista por la conformación semanal de sus carteras durante casi 4 años por medio de la utilización de RNA y un modelo basado en la conformación de betas; donde este último tiene en cuenta dos factores, el riesgo y el retorno, mientras que las RNA le dan una mayor importancia a la maximización de los retornos sujeto a un nivel de riesgo máximo. Los resultados muestran que la conformación de carteras por RNA logran una mayor rentabilidad que las carteras conformadas por betas, convirtiéndolas en una potente alternativa para la generación de retornos proyectados y conformación de carteras.

Bajo la línea de investigación anterior, Richter y Toro (2006) y Foix y Weber (2007) utilizan las RNA en la modelación del precio del cobre y el precio del oro respectivamente, en los que estos últimos implementan modelos clásicos de series de tiempo y modelos híbridos. Los dos trabajos de investigación llegan a la conclusión de definir a las RNA como una excelente herramienta de predicción, aunque Foix y Weber (2007) señalan que es posible conseguir pronósticos más precisos con la aplicación de modelos híbridos en

¹ Commodity: Es todo bien que es producido en masa por el hombre, o del cual existen enormes cantidades disponibles en la naturaleza, que tiene valor o utilidad y un muy bajo nivel de diferenciación o especialización.

² Rolling: Tipo de funcionamiento que trata de capturar el hecho de que los agentes al formar sus expectativas no utilizan toda la información histórica, sino que desechan la más antigua a medida que ésta se va haciendo obsoleta e incorporan la información más reciente a su set de información relevante para tomar decisiones.

³ Bootstrap: Técnica estadística que consiste en generar nuevas series a partir de la combinación de series ya existentes, con el fin de poder realizar inferencia, obtener la distribución de los resultados, establecer intervalos de confianza, etc.

trabajo conjunto con modelos de series de tiempo y redes neuronales. La predicción del precio del oro también es investigada por Ismail, Yahya y Shabri (2009) por medio de la utilización de modelos de regresión lineal múltiple; estos investigadores identifican las siguientes variables como las más influyentes sobre el comportamiento del precio del oro: El índice CRB, el tipo de cambio EURO/USD, la tasa de inflación, la oferta de dinero (M1), el índice compuesto de NYSE, Standard and Poor's 500, letras del tesoro y el índice del dólar de los EE.UU. (USDX), pero también resaltan que los problemas de multicolinealidad son enormes al emplear todas estas variables como explicativas dentro del modelo, utilizando solo la inflación, la oferta monetaria (M1), el índice CRB y el tipo de cambio EURO/USD. Los resultados arrojados demuestran que la técnica y las variables empleadas logran predecir el precio del oro en un 85,2%.

1.2 REVISIÓN DE LITERATURA EN EL ÁMBITO NACIONAL

Dentro del contexto nacional, el trabajo de Misas, López y Querubín (2002) proponen un modelo de pronóstico para la inflación colombiana de 1969 a 1997 a partir de la técnica RNA, donde recurren a una red con una capa oculta y propagación hacia adelante y la inclusión de los rezagos de variables explicativas tales como la inflación, M1 y M3; con el fin de encontrar un modelo que lograra un mayor ajuste, a partir de la especificación de diferentes modelos a través de prueba y error y combinando las variables mencionadas anteriormente, además de la incorporación de modelos lineales con el fin de realizar comparaciones entre ellos. Finalmente, se concluye que las redes pueden modelar el comportamiento de la inflación colombiana y que estos modelos son tan buenos o mejores que los modelos lineales clásicos. El trabajo de Santana (2006) también se enfoca en la modelación y predicción de la inflación colombiana en el periodo comprendido de julio a diciembre del 2005, para ello aplican RNA, la metodología SARIMA de Box-Jenkins y suavizamiento exponencial. El trabajo concluye que las RNA son más precisas que las metodologías tradicionales.

Otros factores económicos como el IPC y la TRM son analizados por los trabajos de Velásquez y Montoya (2005) y Ayala y Castillo (2006) respectivamente. La primera investigación propone un modelo híbrido basado en una descomposición estructural de la serie con el objetivo de remover cualquier patrón fácilmente detectable en los datos, el cual es contrastado con modelos Box-Jenkins y modelos convencionales de RNA. El segundo artículo emplea dos modelos RNA, uno está basado en el entrenamiento con 20 días precedentes y el otro con 5 días precedentes al valor de la predicción. Los investigadores señalan que las RNA son una excelente técnica de predicción, sin embargo, Velásquez y Montoya (2005) resaltan un mejor desempeño por parte del modelo híbrido y Ayala y Castillo (2006) obtuvieron mejores resultados con el modelo de 5 entradas, esto demuestra que la información más reciente es mucho más importante que la inclusión de información más antigua para este tipo de investigaciones. Jalil y Misas (2007) también realizan una investigación sobre el tipo de cambio empleando RNA con técnica *rolling* para el pronóstico y un modelo ARIMA con el fin de evaluarlos con medidas simétricas y asimétricas. Los resultados muestran que los modelos de RNA tienen una mayor capacidad de pronóstico para funciones de pérdida que no son asimétricas.

Velásquez y Aldana (2007) intentan modelar el precio del café colombiano cotizado en la bolsa de Nueva York por medio de la utilización de diversos modelos como son el AR, ARCH, AR-MLP y AR-MLP-ARCH y, aunque este trabajo demuestra la aplicación de diversos modelos lineales y no lineales en la predicción del comportamiento del precio del café y permite a la vez comprender la dinámica de este *commodity*, los resultados obtenidos no son los más satisfactorios, pese a que el modelo que presentó un mayor acercamiento es el AR-MLP-ARCH.

El pronóstico del precio de la energía cotizado en la bolsa es un tema que llama la atención de muchos investigadores, como por ejemplo las realizadas por Toro, Garcés y Galvis (2007), Cadavid y Molina (2008) y Botero y Cano (2008). El primer trabajo implementa una red de tipo Perceptrón Multicapa con la técnica de *backpropagation* para el entrenamiento y utiliza como datos de entrada el precio en bolsa del mes anterior, el mes considerado, los datos del embalse equivalente y la demanda esperada. Cadavid y Molina (2008) emplean dos estructuras de RNA, en la primera emplean la serie de precios diarios

como datos de entrada y en la segunda la serie de precio diarios más el nivel medio de embalses los cuales son contrastados con un modelo GARCH(1,1) y, por último, el artículo de Botero y Cano (2008) hace uso de modelos Autorregresivos donde es empleada la serie del precio de la energía en bolsa para explicar el comportamiento del precio de la energía.

Los modelos que emplearon RNA obtuvieron resultados satisfactorios tanto dentro como fuera de la muestra, pero señalan que un análisis y procesamiento más riguroso de los datos de entrada así como una búsqueda de mejores topologías permiten mejorar la precisión del pronóstico. Por otro lado, la investigación de Botero y Cano (2008) empleó modelos Autorregresivos para indicar que existió un periodo de intervención por parte de un regulador y que sólo cuando fueron incorporadas estas intervenciones a las técnicas de estimación se logró obtener un modelo que cumpliera todos los estadísticos de prueba. Sin embargo, se resalta que la estacionalidad y la parte autorregresiva del modelo son determinantes en el momento de modelar series de este tipo, por lo tanto, el modelo se debe retroalimentar de forma permanente o se requiere de modelos más elaborados.

2. MARCO TEÓRICO

2.1 EL ORO EN LA ECONOMÍA MUNDIAL

Desde la antigüedad, el oro ha tenido una gran importancia para el hombre, ya que ha sido empleado para fines ornamentales y de culto, para adorno personal, para tratamientos médicos, para la fabricación de aparatos tecnológicos, como símbolo de poder y como reserva monetaria, todo esto, gracias a sus características de maleabilidad, belleza, difícil falsificación y resistencia. Además, este precioso metal ha jugado un importante papel en la política monetaria; los grandes movimientos en la producción del oro, las fluctuaciones en precios en patrón monetario y el sistema general le confieren una importancia fundamental para las transacciones internacionales (Heredia, 2009).

En el siglo XIX el desarrollo del sistema monetario internacional estaba basado en el respaldo de billetes por una cantidad de oro, fenómeno conocido como el “Patrón oro”; lo cual significó un verdadero internacionalismo donde Europa y América eran los continentes con mayor poder económico y Gran Bretaña era el administrador del patrón oro. Sin embargo, al estallar la Primera Guerra Mundial (1914), la financiación de los gastos y reconstrucción de la posguerra hizo incrementar la cantidad de billetes en circulación sin poder soportarlo con reservas de oro lo que excluyó la circulación de monedas con el fin de concentrarlo en los bancos (Witker y Valenzuela, 2002).

Durante la posguerra, Gran Bretaña hizo muchos intentos por recuperar el poderío por medio del patrón oro, pero sus intentos fueron fallidos y sus políticas no arrojaron los mejores resultados, en cambio una nueva potencia, Estados Unidos, y un nuevo mercado, Nueva York, comenzaban a concentrar los capitales financieros. En 1944 se realizó la conferencia de Bretton Woods donde se establecieron las reglas para las relaciones comerciales y financieras entre los países más industrializados del mundo, se decidió la creación del Banco Mundial y el Fondo Monetario Internacional, el uso del dólar como

moneda internacional, y el restablecimiento del sistema “Patrón cambio oro”, lo cual generó un ambiente de estabilidad y desarrollo a la mayoría de las economías mundiales durante el periodo 1944-1971, conocido en la historia como “Edad de oro” (Witker y Valenzuela, 2002).

Todas las características propias del oro lo han convertido en uno de los depósitos de valor más importantes frente a la incertidumbre y la volatilidad de los mercados en la actualidad, además de ser fuente de refugio contra el dólar y tiempos inflacionarios, ya que son aspectos que generan temor sobre la pérdida de poder adquisitivo del dinero induciendo a los inversionistas a proteger sus riquezas a través de la adquisición de dicho metal precioso (Heredia, 2009).

2.2 MODELOS DE PREDICCIÓN

Los modelos de series temporales se construyen sobre la premisa de que las series de tiempo tienen una historia estadística recurrente particular que puede ser modelada y explotada para fines de pronóstico. Detrás de esta metodología, está la idea ecléctica que no podemos conocer lo suficiente acerca de la estructura de una economía que permita la construcción de un modelo estructural detallado que permita la obtención de buenos pronósticos (Araníbar y Humérez, 1996).

Rodríguez y Fermín (2006) señalan que la caminata aleatoria describe muy bien el comportamiento de las series financieras y es una clara muestra de encontrarse en una situación concerniente a un mercado eficiente, donde los precios de los títulos negociados reflejan toda el contenido informativo disponible y se ajustan total o rápidamente a la nueva información, con lo cual, los cambios en los precios de los valores se tornan imprevisibles.

Por otro lado, la técnica Box-Jenkins analiza las propiedades probabilísticas o estocásticas de las series de tiempo económicas bajo la filosofía de “permitir que la información hable por sí misma”, esta técnica ha sido ampliamente utilizada en el pronóstico de series de

tiempo arrojando muy buenos resultados cuando se cuenta con una serie estacionaria o que puede convertirse en estacionaria a través de transformaciones (Cruz, 2008).

Por último, la especificación de Redes Neuronales Artificiales RNA son definidos por Cruz (2008) como modelos matemáticos con base al funcionamiento de los sistemas nerviosos biológicos que intenta contar con un sistema que imite y simule el funcionamiento natural de las neuronas, es decir, que puedan “pensar” y “aprender” como el cerebro humano, el cual se puede ver concebirse como un conjunto interconectado de neuronas.

Hipotéticamente, se tiene que un modelo basado en Redes Neuronales Artificiales (RNA) es capaz de modelar el comportamiento de una serie de tiempo financiera e incluso de mejorar los resultados arrojados por modelos Box-Jenkins y Caminata Aleatoria, ya que las RNA pueden capturar las relaciones no lineales presentes en los mercados financieros, como se ha demostrado en diversos estudios.

2.2.1 Caminata aleatoria

La hipótesis de la caminata aleatoria sugiere que el movimiento de los precios de los activos financieros es imprevisible, lo cual indica que no hay tendencias perceptibles con las cuales se pueda realizar inversión, si existiera, los inversores podrían obtener ganancias extraordinarias; este modelo supone que los rendimientos sucesivos son independientes y que los rendimientos están idénticamente distribuidos en el tiempo. (Martínez, 2007).

En síntesis, la caminata aleatoria define un mercado eficiente en forma débil si el precio más reciente contiene toda la información disponible, esto indica que el precio más reciente es quien predice de una mejor forma los precios futuros. En esta medida, se plantea el siguiente modelo de caminata aleatoria con deriva como alternativa en la determinación de los precios del oro en el futuro:

$$Y_t = \beta + Y_{t-1} + E_t \quad (1)$$

$$dY_t = \beta + E_t \quad (2)$$

Donde Y_t es la variable dependiente que se intenta predecir en el periodo t , β es un término *drift* o deriva arbitrario, dY_t es el cambio en el índice y E_t es un error aleatorio que cumple con las condiciones $E(\varepsilon_t)=0$, $E(\varepsilon_t^2) = \sigma_\varepsilon^2$ y $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0$ para todo $s \neq t$.

2.2.2 Técnica Box-Jenkins

La metodología Box-Jenkins consiste en encontrar un modelo matemático que represente el comportamiento de una serie temporal de datos y permita hacer previsiones únicamente introduciendo el periodo de tiempo correspondiente (Chatfield, 1989); los modelos Box-Jenkins son los llamados Autorregresivos [AR(p)], Medias Móviles [MA(q)], Mixtos [ARMA (p, q)] y Autorregresivos Integrados de Medias Móviles [ARIMA (p, d, q)], es decir, toda una familia de modelos para explicar la evolución de una variable a lo largo del tiempo a partir de las observaciones pasadas de la serie y errores pasados de previsión.

Gómez, Madariaga y Ugarte (1988) citando a Pankratz (1983) señalan tres ventajas que justifican y aconsejan la utilización de los modelos ARIMA: En primer lugar, los métodos tradicionales son, en su mayor parte, modelos “ad hoc” o intuitivos, sin un fundamento sólido de estadística matemática y teoría de la probabilidad. En segundo lugar, los modelos ARIMA, no son un único modelo sino una familia completa de posibles modelos. Por último, se puede demostrar que un modelo ARIMA adecuado produce predicciones muy buenas, es decir, ningún otro modelo univariante consigue predicciones con menor error medio cuadrático, aunque cabe señalar que este tipo de modelos presentan mejor desempeño en predicciones de corto plazo que en predicciones de largo plazo.

2.2.2.1 Proceso autorregresivo AR(p)

En el proceso autorregresivo de orden p la observación actual Y_t es generada por un promedio ponderado de observaciones pasadas que se remontan p periodos, junto con una perturbación aleatoria en el periodo actual, la ecuación que describe este proceso es:

$$Y_t = \mu + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + E_t \quad (3)$$

2.2.2.2 Proceso de media móvil MA(q)

En el proceso de media móvil de orden q cada observación Y_t es generada por un promedio ponderado de perturbaciones aleatorias que se remontan q periodos, la ecuación que describe este proceso es:

$$Y_t = \mu + E_t + \alpha_1 E_{t-1} + \alpha_2 E_{t-2} + \dots + \alpha_q E_{t-q} \quad (4)$$

El modelo de media móvil supone que las perturbaciones aleatorias están distribuidas en forma independiente a lo largo del tiempo, es decir, son generadas por un proceso de ruido blanco.

2.2.2.3 Proceso autorregresivo y de media móvil ARMA(p,q)

Este tipo de modelos mixtos explican el comportamiento de muchas series que no pueden modelarse sólo como Autorregresivos o Medias Móviles, la ecuación que describe este proceso es:

$$Y_t = \mu + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + E_t + \alpha_1 E_{t-1} + \alpha_2 E_{t-2} + \dots + (5)$$

Donde Y_t es la variable que se intenta modelar, los coeficientes de los parámetros $\beta_1, \beta_2, \beta_p, \alpha_1, \alpha_2, \alpha_q$ son determinados a partir de los datos a través de cualquier método estadístico consistente y los términos de error E_t son variables aleatorias normales, independientes e idénticamente distribuidas con $E(\varepsilon_t)=0$, $E(\varepsilon_t^2) = \sigma_\varepsilon^2$ y $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0$ para todo $s \neq t$.

2.2.3 Modelo de redes neuronales artificiales de tipo Perceptrón Multicapa (MLP)

“Una red neuronal puede ser descrita como un modelo de regresión no lineal cuya estructura se inspira en el funcionamiento del sistema nervioso. En términos generales, una red consiste en un gran número de unidades simples de proceso, denominas neuronas, que actúan en paralelo y están conectadas mediante vínculos ponderados.

Cada neurona recibe entradas desde otras neuronas y genera un resultado que depende sólo de la información localmente disponible, ya sea almacenada internamente o plasmada en los ponderadores de las conexiones. El resultado generado por la neurona servirá de entrada para otras neuronas.

Mediante la adecuada modificación de los ponderadores de la red, en un proceso denominado aprendizaje, la red mejorará su desempeño en el desarrollo de la tarea para la cual fue construida.” (Foix y Weber, 2007)

2.2.3.1 Estructura de la red

Este estudio empleará una red Perceptrón Multicapa (MLP de sus siglas en inglés *Multi Layer Perceptron*) con información recurrente hacia adelante, las cuales poseen al menos tres niveles de neuronas, la capa de entrada se encarga de almacenar la información suministrada a la red, la capa oculta se encarga de recibir, procesar y memorizar la información y la capa de salida contiene la respuesta de la red dando un resultado final. Aunque es posible emplear más de una capa oculta, este estudio no lo hará ya que según Caicedo y López (2009) proponer más de una capa oculta aumenta fuertemente la complejidad computacional del algoritmo de aprendizaje.

Este tipo de red es muy utilizada en la predicción o modelación del comportamiento de los precios de *commodities*, las cuales según Friz (2004) actúan como funciones mapeadoras universales, desempeñándose muy bien con las series de tiempo; estas redes se caracterizan porque pueden aplicarse diferentes funciones de activación a los *slabs* de la capa oculta, para detectar diferentes características en los patrones procesados a través de la red.

La red MLP tiene la siguiente forma funcional:

$$\hat{y}_t = \lambda \left[\sum_{h=1}^Q c_h G \left(\sum_{i=1}^p w_{hi} x_{ti} + \theta_h \right) + d \right] \quad (6)$$

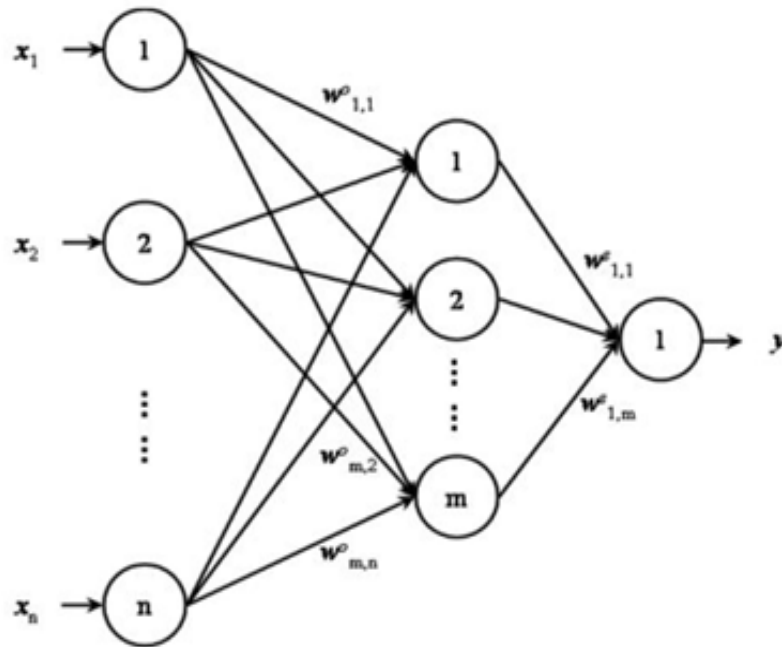
Donde $\hat{y}_t$ es el valor de salida de la red en el periodo t , $G(.)$ y $\lambda(.)$ representan la función de activación, w_{hi} representa el peso de la conexión de la neurona h a la neurona

i , x_{ti} representa las variables explicativas o inputs $\{x_{t1}, x_{t2}, x_{t3}, \dots, x_{tp}\}$, θ_h representa los umbrales de las neuronas.

Si $\lambda(\cdot)$ es la función lineal:

$$\hat{y}_t = \sum_{h=1}^q c_h G(\sum_{i=1}^p w_{hi} x_{ti} + \theta_h) + d \quad (7)$$

Figura 1. Arquitectura de una red MLP con una capa oculta



Fuente: Hilera y Martínez (1994)

La figura anterior es una red de tipo Perceptrón Multicapa (MLP por sus siglas en inglés) con n número de variables o nodos en la capa de entrada, m número de nodos en su única capa oculta y una única salida y que representa el resultado o predicción del modelo. Las diferentes combinaciones $w^o_{m,n}$ representan los pesos sinápticos que proporciona el sistema entre la capa de entrada y oculta y las combinaciones $w^e_{1,m}$ representa los pesos sinápticos que proporciona el sistema entre la capa oculta y la de salida.

2.2.3.2 Funciones de activación

Según Friz (2004) las funciones de activación son los elementos de la red que determinan en qué nivel de datos se requiere el mayor énfasis. Por ejemplo, si se utiliza una función lineal todos los datos son tratados de la misma manera; por el contrario, si se usara una función tangente hiperbólica, ésta consideraría más relevante la información alrededor de cero que aquella que se encuentra en los extremos de uno y menos uno. En este caso específico será empleada una función de activación de tipo sigmoideal en la capa oculta, según Caicedo y López (2009) esta función es la que ha resuelto mayor cantidad de problemas con redes MLP, pues le garantiza la capacidad de procesamiento no-lineal y en la capa de salida sugieren la utilización de una función de activación lineal ya que es muy útil en el modelo e identificación de sistemas, además que esta función no satura la salida de la red.

2.2.3.3 Características de las redes neuronales artificiales

Las RNA se caracterizan por actuar de 3 formas: la primera es la forma paralela que consiste en que cada uno de los procesadores elementales trabaja con una parte de un problema mayor, la segunda es la forma distribuida donde los cálculos asignados se distribuyen entre todas las neuronas, cada una de las cuales trata una fracción de un problema mayor y la tercera es la versión adaptativa donde las RNA se adaptan al entorno por medio del cambio en los pesos y sinapsis, de tal manera que aprenden de la experiencia encontrando soluciones a los problemas.

2.2.3.4 Proceso de enseñanza y pesos

El proceso de enseñanza se debe realizar previamente a la generación del modelo con el fin de obtener una RNA “pensante” y capaz de discernir la información de forma adecuada. Los datos empleados en los estudios de redes neuronales artificiales pueden ser empleados en dos grupos como son aprendizaje y validación ó en tres grupos aprendizaje, validación y verificación o test. El entrenamiento se emplea en la primera fase de ejecución del

algoritmo de aprendizaje, la validación mide la capacidad de la red para responder correctamente a los patrones que no fueron ingresados durante el entrenamiento y el test sirve para verificar el grado de aprendizaje. Este trabajo empleará entrenamiento, validación y test siguiendo las sugerencias de Caicedo y López y los porcentajes para cada grupo estarán entre 50% a 70% para el entrenamiento, de 30% a 50% para validación y de 10% a 20% en la verificación, de tal forma que la suma de la distribución de los tres grupos de un 100%.

La formación de la red debe incluir unos pesos sinápticos o ponderadores (w_{hi}) que cada neurona le asigna a la información que recibe de las neuronas que la preceden; los diferentes valores que reciben las neuronas modificados por los pesos sinápticos se suman para producir la entrada neta, donde esta última determina si la neurona se activa o no y esta activación o no de la neurona depende a su vez de la función de activación. El proceso de aprendizaje está dado por:

$$W(t+1) = w(t) + \Delta w(t) \quad (8)$$

Donde, $w(t+1)$ es el valor actualizado del peso sináptico, $w(t)$ es el valor actual del peso sináptico y $\Delta w(t)$ es la variación del peso sináptico.

2.2.3.5 Aprendizaje supervisado

La red aquí propuesta estará bajo aprendizaje supervisado, el cual según Caicedo y López (2009) se caracteriza porque el proceso de entrenamiento es controlado por un agente externo llamado *supervisor o maestro*. El mundo real corresponde al problema a modelar que se representa mediante un conjunto de datos de entrada y salida. El supervisor opera como un “maestro” que guía el aprendizaje en la red, y conoce las salidas deseadas correspondientes a las respectivas entradas.

Los datos de entrada se le presentan al supervisor y a la red neuronal simultáneamente, el supervisor propone la salida deseada para ser comparada con la salida de la red neuronal artificial. Para que la labor del supervisor sea exitosa se define en él, un error de

entrenamiento como la diferencia entre la salida deseada o esperada y la salida que produce la RNA.

$$error = d - y \quad (9)$$

Donde, y es la salida de la red neuronal artificial y d es la salida deseada.

2.2.3.6 Algoritmo de aprendizaje

El algoritmo más conocido para el entrenamiento de la red es el de Retropropagación del Error o *backpropagation* con gradiente descendente, el cual según Caicedo y López (2009) encuentra un valor mínimo de error (local o global) navegando con el gradiente descendente por la superficie, donde la velocidad de convergencia es controlada por el parámetro α . Sin embargo, el algoritmo *backpropagation* resulta ser un método de entrenamiento lento para ciertas aplicaciones donde requeriremos una alta velocidad de convergencia, por lo cual este trabajo empleara el algoritmo Levenberg Marquardt, quien presenta como una de sus grandes fortalezas la combinación de dos estrategias de minimización, el método de gradiente descendente y el método de Newton, además de ajustar el parámetro α a cada modelo.

2.2.3.7 Tamaño de la red

El tamaño de la red no tiene una regla determinada, su tamaño óptimo se logra por medio de la experimentación, pero es sabido que la capa oculta está relacionada con la capacidad de mapeo de la red, por lo que un número grande de neuronas en esta capa contribuye a la capacidad para memorizar el conjunto de entrenamiento, pero aumentar el tamaño de la red exageradamente tiene consecuencias muy negativas porque la red puede llegar a un punto en el que no pueda generalizar la información obteniendo resultados impredecibles.

2.2.3.8 Duración del entrenamiento

Otro problema es la detención del proceso de entrenamiento. El proceso debe detenerse cuando entre dos entrenamientos consecutivos no disminuye el error cuadrático medio, para ello se tienen dos alternativas: la primera es fijar un ECM en un punto determinado para cuando el entrenamiento llegue a este punto se detenga, la segunda alternativa es fijar un número máximo de eventos desde que se alcanzó el último mínimo. Adicionalmente, es de vital importancia para cualquier red saber distinguir entre memorización y generalización, esta última es definida por (Friz, 2004) como la capacidad de proporcionar una respuesta correcta ante patrones que no han sido empleados en su entrenamiento.

3. METODOLOGÍA

3.1 ANÁLISIS DESCRIPTIVO

3.1.1 Descripción de la muestra, variables del modelo, selección de los datos y softwares

Desde 1918 la bolsa de Nueva York se ha convertido en la principal casa de bolsa del mundo, el centro de inversión mundial y la más importante del mundo, los organismos de control la regulan con el fin de lograr transacciones limpias y que no perjudiquen el bienestar de las economías. Todas las características propias de esta bolsa como su puntualidad, arbitraje, tecnología, entre muchas otras, hacen de ella una fuente confiable de inversión. Es por ello, que este trabajo de investigación emplea la información de la bolsa de Nueva York y los datos son obtenidos de la página web www.kitco.com reconocida y utilizada ampliamente a nivel mundial en todo lo relacionado con la cotización de metales preciosos.

Los datos originales están dados en términos de USD/onza, el periodo de tiempo que se analiza en este trabajo de investigación fue considerado de acuerdo a la tesis de Fríz(2004), quien emplea datos recientes y un rango de estudio de pocos años obteniendo muy buenos resultados de predicción y análisis, además, se tuvo en cuenta que en análisis de series de tiempo es necesaria una buena cantidad de datos para la muestra, este periodo está conformado por 992 datos desde el 2-enero-2002 al 29- diciembre-2005 y el periodo que se intenta modelar o predecir es del 29-noviembre-2005 a 29-diciembre-2005, el cual está conformado por 21 datos.

En la estimación de los modelos se emplearon diferentes rezagos del precio del oro como variables explicativas y se escoge el mejor modelo de acuerdo a las pruebas y resultados obtenidos, según Foix y Weber (2007) la generación de proyecciones del precio empleando otras variables explicativas tiene el potencial problema de requerir el pronóstico de dichas

variables, lo que complica su utilización práctica. Sin embargo, si las redes neuronales demuestran tener un mejor desempeño que los modelos de series de tiempo, se habrá establecido que las redes neuronales pueden sacar mayor provecho de los precios históricos y se tendrá una evaluación conservadora de su eficiencia predictiva que podrá alentar a la realización de investigaciones adicionales. El *software* empleado en la realización de los modelos lineales es Eviews 6.0 y para modelar la red neuronal MLP fue empleado el *software* Matlab 7.7.0

3.2 ANÁLISIS INFERENCIAL

3.2.1 Estimación e interpretación de resultados de los modelos de series de tiempo: Caminata Aleatoria con deriva y ARIMA(p,d,q)

En primera medida, se realiza un análisis exploratorio de la serie que consiste en la observación de la tendencia, estacionalidad y ciclos por medio de la visualización gráfica de la serie con el fin de conocer su estacionariedad, este proceso se le conoce como identificación de la serie. El análisis gráfico lleva a pensar que la serie presenta problemas de estacionariedad, pero para corroborar la aplicación de los logaritmos se realiza la prueba de rango-media y se realizan los test de raíz unitaria para comprobar cuantas veces es oportuno realizar diferenciación⁴.

Menguados los problemas de no estacionariedad de la serie se procede a modelar; el correlograma sugiere que la serie asume un comportamiento de caminata aleatoria, sin embargo, algunos rezagos p , q sobresalen por lo cual se estiman diferentes modelos $AR(p)$, $MA(q)$ y $ARIMA(p,d,q)$ con el fin de escoger el mejor de acuerdo a los criterios de comparación y mayor nivel de pronóstico medido a partir del mínimo RMSE y MAE, pero antes de realizar predicción se debe verificar el modelo a través de unos supuestos que deben cumplir los coeficientes y el término de error ϵ_t .

⁴ Para un análisis detallado de los gráficos de las series, correlogramas, prueba de rango-media y test Dickey-Fuller véase Anexo A

De acuerdo a la información obtenida a partir del correlograma se sugiere la implementación de un modelo de caminata aleatoria, sin embargo, se emplean los modelos ARIMA(2,1,2) y ARIMA(4,1,4) ya que son significativos:

$$Precio_t = \beta + E_t \quad (10)$$

$$Precio_t = E_t + \sum_{i=1}^p \phi_i Precio_{t-i} + \sum_{i=1}^q \theta_i E_{t-i} \quad (11)$$

Donde, $Precio_t$ representa la diferencia del logaritmo del precio del oro en el periodo t , E_t es un error aleatorio que cumple con las condiciones $E(s_t)=0$, $E(s_t^2) = \sigma_s^2$ y $Cov(s_t, s_s) = 0$ para todo $s \neq t$.

Los resultados obtenidos a partir de las estimaciones de los modelos Caminata Aleatoria y los diferentes modelos ARIMA(p,d,q) son presentados en la tabla1.

Tabla1. Estimación de modelos de series de tiempo

	Variables	Coefficient	Std, Error	t-Statistic	Prob,
ARIMA(2,1,2)	C	0,0006	0,0003	2,1348	0,0330
	AR(1)	-1,4999	0,0092	-162,6951	0,0000
	AR(2)	-0,9784	0,0087	-112,6685	0,0000
	MA(1)	1,5067	0,0055	275,0069	0,0000
	MA(2)	0,9916	0,0049	202,0609	0,0000
ARIMA(4,1,4)	C	0,0006	0,0003	2,0990	0,0361
	AR(1)	-0,4834	0,0111	-43,3930	0,0000
	AR(2)	-0,4875	0,0110	-44,2810	0,0000
	AR(3)	-0,4858	0,0107	-45,4756	0,0000
	AR(4)	-0,9471	0,0109	-87,1366	0,0000
	MA(1)	0,4924	0,0049	99,5918	0,0000
	MA(2)	0,4989	0,0062	80,6057	0,0000
	MA(3)	0,4879	0,0055	88,2649	0,0000
	MA(4)	0,9860	0,0041	238,3338	0,0000
CAMINATA ALEATORIA	C	0,0006	0,0003	2,1368	0,0329

Fuente: Cálculos propios

En la tabla anterior se puede observar que todos los coeficientes son estadísticamente significativos al 5%. Al realizar el proceso de estimación para los diferentes modelos, se encuentra que existen problemas de leptokurtosis producto de la gran volatilidad que presentan los datos (*outlayers*), además, con el fin de menguar posibles problemas de heterocedasticidad Eviews 6.0 presenta una herramienta muy útil que ayuda a menguar este tipo de problemas a través del coeficiente de consistencia heterocedastica de White. Ahora, se procede a comparar los diferentes modelos por medio de los criterios de comparación Shwarz, Akaike, Log-Likelihood y Hanna Quinn:

Tabla 2. Criterios de comparación modelos lineales

MODELO	SHWARZ	AKAIKE	LOG- LIKELIHOOD	HANNA QUINN
ARIMA(2,1,2)	-6,5360	-6,5608	3249,2960	-6,5513
ARIMA(4,1,4)	-6,5197	-6,5644	3248,5080	-6,5474
CAMINATA ALEATORIA	-6,5575	-6,5624	3252,6900	-6,5606

Fuente: Cálculos propios

El análisis por criterios de comparación sugiere que el mejor modelo es la Caminata Aleatoria, pero antes de realizar cualquier tipo de predicción se procede a realizar la verificación de los modelos donde $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = \sigma_\varepsilon^2$ y $\text{Cov}(\varepsilon_s, \varepsilon_t) = 0$ para todo $s \neq t$; la etapa de verificación tiene su origen en la idea de que todo modelo es erróneo, puesto que son representaciones simples de la realidad (Rodríguez, 2001).⁵ Cumplidos los supuestos de verificación de los modelos se realiza la predicción y se escoge el mejor modelo de acuerdo a las medidas de evaluación de predicciones Error Absoluto Medio (EAM) y Raíz del Error Cuadrático Medio (RECM)⁶

⁵ Para un análisis detallado de la verificación de los supuestos de los modelos véase Anexo B.

⁶ El Error Absoluto Porcentual Medio (EAPM) también es una medida de evaluación muy buena, sin embargo, tiene el inconveniente de que para valores de la serie y_t muy pequeños, la contribución relativa en el periodo t puede ser enorme.

Tabla 3. Medidas de evaluación de predicción modelos lineales

MODELOS	MAE	RMSE
ARIMA(2,1,2)	10,4816473	14,08988674
ARIMA(4,1,4)	10,0149783	13,58186367
CAMINATA ALEATORIA	10,5569377	14,13867179

Fuente: Cálculos propios

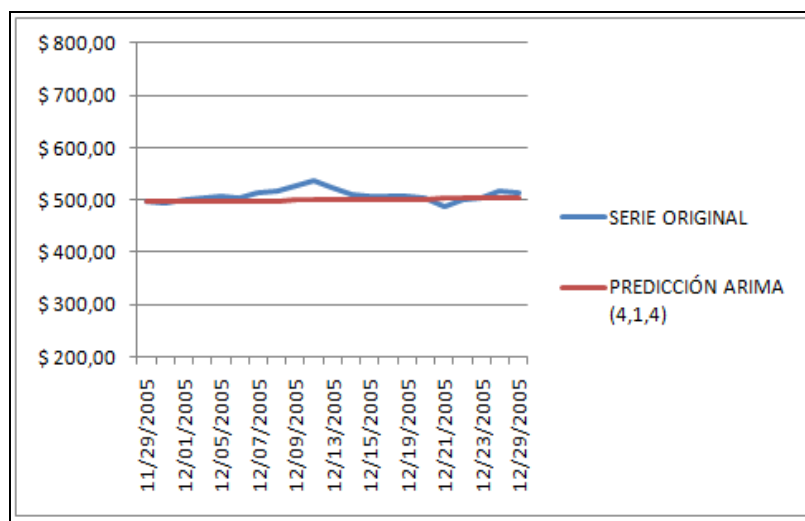
Los resultados de la tabla 3 muestran que las medidas de evaluación de predicción no son tan cercanas a cero como es lo ideal y contrario a lo que mostraron los criterios de comparación, el modelo que logra acercarse en una mayor medida a los datos originales o que presenta un mayor nivel de predicción es el ARIMA(4,1,4) lo que le concede el título del “mejor modelo de predicción” de series de tiempo de este estudio de acuerdo a las medidas de evaluación antes mencionadas. Ahora, será realizado un análisis al gráfico de la serie original versus la predicción del modelo ARIMA(4,1,4) con el fin de observar que tanto se asemeja el comportamiento de las series.

Tabla 4. Serie original y Predicción ARIMA(4,1,4)

FECHA (m/d/a)	SERIE ORIGINAL	PREDICCIÓN ARIMA (4,1,4)
11/29/2005	496,00	496,78
11/30/2005	495,65	497,42
12/01/2005	499,75	498,57
12/02/2005	502,50	498,82
12/05/2005	505,65	498,47
12/06/2005	504,25	498,43
12/07/2005	515,40	498,17
12/08/2005	515,70	498,89
12/09/2005	525,50	499,83
12/12/2005	536,50	500,32
12/13/2005	522,50	500,87
12/14/2005	509,50	500,53
12/15/2005	506,25	500,34
12/16/2005	507,00	500,64
12/19/2005	508,75	500,95
12/20/2005	502,50	501,95
12/21/2005	489,00	502,50
12/22/2005	500,00	502,61
12/23/2005	503,60	502,75
12/28/2005	518,00	502,42
12/29/2005	513,00	502,75

Fuente: Elaboración propia

Figura 2. Serie original vs. Predicción ARIMA(4,1,4)



Fuente: Elaboración propia

En la figura 2 se puede apreciar que la serie original presenta un comportamiento fluctuante para el último mes del año 2005 muy marcada por ascensos y descensos del precio del metal, en contraste, la predicción obtenida con el mejor modelo lineal de series de tiempo ARIMA(4,1,4) de este estudio también muestra algunas fluctuaciones pero éstas no se acercan de manera estricta a la serie objetivo, este hecho pudo ser apreciado con anterioridad en las medidas de evaluación que informaron de modelos no tan buenos en el proceso de predicción del precio exacto del *commodity* oro.

Los resultados encontrados para los modelos de series de tiempo sugieren que no es posible encontrar una predicción exacta sobre el precio del oro empleando como variables explicativas la información pasada de la misma variable, esto se debe en gran medida a que los agentes económicos que operan en los mercados no pueden obtener utilidades más allá de lo normal con la utilización de información públicamente disponible.

3.2.2 Estimación e interpretación de resultados de los modelos Perceptrón Multicapa de Redes Neuronales Artificiales

El uso de las series originales o normalizadas varía de acuerdo a los autores que realicen la investigación, este trabajo empleará datos normalizados ya que según Foster (2002) este

proceso optimiza más fácilmente el rendimiento de la red neuronal porque no la satura con el manejo de datos muy elevados.

De acuerdo a la expresión (6), se tiene que la formulación del modelo definitivo Perceptrón Multicapa de Redes Neuronales Artificiales a estimar es:

$$\widehat{Precio}_t = \lambda \left[\sum_{h=1}^Q c_h G \left(\sum_{i=1}^p w_{hi} x_{ti} + \theta_h \right) + d \right] \quad (12)$$

Donde $\widehat{Precio}_t$ es el valor de salida de la red del precio del oro en el periodo t , $G(.)$ es la función de activación sigmoideal de la capa oculta y $\lambda(.)$ representan la función de activación lineal de capa de salida, w_{hi} representa el peso de la conexión de la neurona h a la neurona i , x_{ti} representa las variables explicativas o inputs $\{x_{t1}, x_{t2}, x_{t3}, \dots, x_{tp}\}$ que en este caso son los precios del oro rezagados hasta en 10 periodos, θ_h representa los umbrales de las neuronas.

El tamaño de la red fue ajustado gradualmente a través de la experimentación hasta obtener los seis mejores modelos de acuerdo a las funciones de evaluación de predicción Error Absoluto Medio (EAM) y Raíz del Error Cuadrático Medio (RECM), aunque cabe anotar que es importante observar otras variables como son el rendimiento de la red, gradiente, alfa o mu, ajuste de los datos, entre otros.

Este trabajo de investigación obtuvo que los mejores resultados se obtuvieron empleando un 70% de los datos en el entrenamiento, 20% en validación y 10% en el test. El proceso de entrenamiento y validación, en el cual la red neuronal aprende la dinámica de una función $y=f(t)$, fue muy similar para los seis mejores modelos y se puede decir que la red logra un buen nivel de aprendizaje, se define además, que el proceso de detención del entrenamiento está dado por el número de iteraciones que son regulados bajo supervisión con el fin de obtener las mejores medidas de evaluación de predicción e impedir el sobre-entrenamiento de la red.

En la capa de entrada fueron empleados 7, 8, 9 y 10 rezagos del precio del oro o neuronas y en la capa oculta, siguiendo la sugerencia de Caicedo y López (2009), se empleo el proceso de experimentación el cual consiste en empezar por un muy alto número de neuronas y evaluar su funcionamiento y luego, paulatinamente, ir decreciendo el número de neuronas hasta el instante en el cual, el error de aprendizaje y verificación se mantengan en el mínimo requerido; toda esta dinámica arroja que los mejores modelos se desempeñaron con el uso de 40, 30, 25 y 20 neuronas.

Tabla 5. Criterios de comparación modelo Perceptrón Multicapa

	Neuronas capa entrada	Neuronas capa oculta	Número de iteraciones	Al cierre de entrenamiento				Mejor rendimiento en proceso de validación		Ajuste de los datos
				Iteraciones	Rendimiento	Gradiente	Alfa o Mu	Iteraciones	Rendimiento	
Modelo1	8	40	20	26	0,000579	0,009540	0,001	6	0,00072576	0,99802
Modelo2	10	30	15	22	0,000629	0,002440	0,010	7	0,00073063	0,99800
Modelo3	9	30	20	29	0,000526	0,000645	0,010	9	0,00077882	0,99802
Modelo4	7	30	15	21	0,000581	0,001530	0,010	6	0,00079953	0,99792
Modelo5	10	25	15	20	0,000601	0,010600	0,001	5	0,00080635	0,99799
Modelo6	10	20	15	21	0,000558	0,000609	0,010	6	0,00093234	0,99787

Fuente: Cálculos propios

En la tabla 5. se puede observar que la red presenta dos tipos de rendimiento, uno es el rendimiento al cierre de entrenamiento o lo que es lo mismo, el rendimiento del entrenamiento que se define como la diferencia entre la salida de la red y el valor deseado, este es el error que se utiliza para modificar los pesos de la red, y el otro rendimiento es el del proceso de validación que se define como la diferencia entre la salida de la red y el valor deseado, pero evaluado en un dato de entrada que pertenece al nuevo conjunto de datos de validación; se aclara que este error no es considerado para modificar los pesos sinápticos de la red, es decir, que este error permite observar el comportamiento de la red frente a nuevos datos que pertenecen al conjunto universo del problema, pero diferentes a los utilizados en el entrenamiento.

Se puede ver entonces, que el modelo 1 presento el mejor rendimiento en el proceso de validación en la iteración número 6, pero no obtuvo el mejor rendimiento en el periodo de entrenamiento, que fue obtenido por el modelo 3, esto se debe a que el entrenamiento continúa su proceso hasta completar las 20 iteraciones fijadas inicialmente de tal forma que en este transcurso la red puede cambiar de forma trascendental su entrenamiento.

El gradiente por su parte, ayuda en el proceso de minimización de los errores al encontrar un mínimo local o global por medio del algoritmo de entrenamiento, se nota entonces que el menor gradiente lo presenta el modelo 6 y que la red se ha ubicado en un mínimo local. La velocidad de convergencia se controla con el valor del parámetro α , entre más pequeño sea este parámetro tiene como consecuencia un aprendizaje seguro, pero si por el contrario este parámetro es muy alto, puede generar oscilaciones en el aprendizaje, se puede decir entonces que el aprendizaje más seguro lo presentan los modelos 1 y 5. Finalmente, el ajuste de los datos está conformado por el ajuste del entrenamiento, validación y test a los datos originales o datos objetivo, por lo cual, se puede concluir que el mejor ajuste lo presenta el modelo 1⁷.

En general, todos los modelos presentan resultados bastante similares, pero se podría pensar que el mejor modelo es el número 1, ya que presenta la mayor parte de los mejores criterios de comparación, no obstante, las medidas de evaluación de predicción establecidas para la selección del mejor modelo son el Error Absoluto Medio (EAM) y Raíz del Error Cuadrático medio (RECM) que se exponen a continuación.

Tabla 6. Medidas de evaluación de predicción modelos Perceptrón Multicapa

⁷ Para un análisis detallado del rendimiento del entrenamiento y validación y ajuste del mejor modelo de redes neuronales artificiales véase Anexo C.

MODELO	MAE	RMSE
Modelo1	7,59626237	10,9093782
Modelo2	8,58963015	12,0739199
Modelo3	9,17515408	12,2276672
Modelo4	8,21297792	11,6842024
Modelo5	7,99847617	10,9858272
Modelo6	8,27519449	11,0486509

Fuente: Cálculos propios

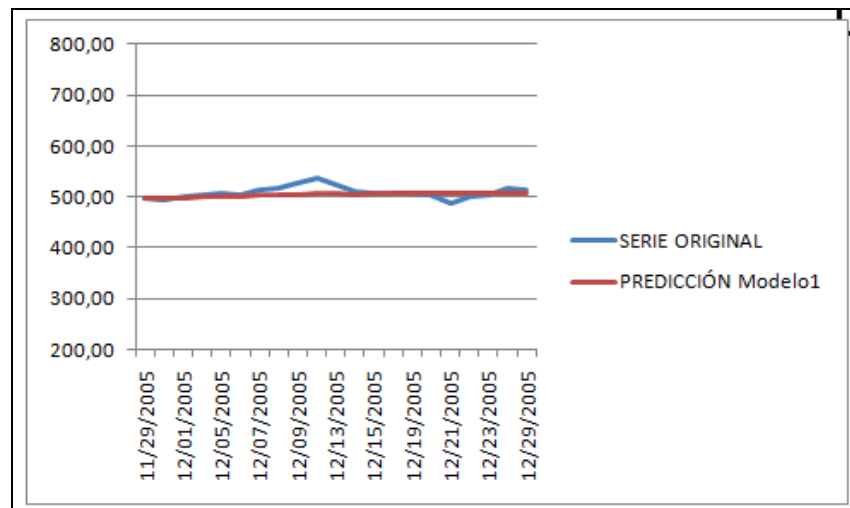
Los resultados obtenidos en la tabla 6 demuestran que las mejores medidas de evaluación de predicción para el modelo Perceptrón Multicapa de Redes Neuronales Artificiales de este estudio de investigación las presenta el modelo 1, tal y como lo habían indicado los criterios de comparación.

Tabla 7. Serie original y Predicción Modelo 1 de RNA

FECHA (m/d/a)	SERIE ORIGINAL	PREDICCIÓN Modelo1
11/29/2005	496,00	498,22
11/30/2005	495,65	496,89
12/01/2005	499,75	498,65
12/02/2005	502,50	500,21
12/05/2005	505,65	499,20
12/06/2005	504,25	502,21
12/07/2005	515,40	504,82
12/08/2005	515,70	504,70
12/09/2005	525,50	504,89
12/12/2005	536,50	505,77
12/13/2005	522,50	506,06
12/14/2005	509,50	505,11
12/15/2005	506,25	505,96
12/16/2005	507,00	507,28
12/19/2005	508,75	506,73
12/20/2005	502,50	506,49
12/21/2005	489,00	507,18
12/22/2005	500,00	507,11
12/23/2005	503,60	506,59
12/28/2005	518,00	507,25
12/29/2005	513,00	508,18

Fuente: Elaboración propia

Figura 3. Serie original vs. Predicción Modelo 1 de RNA



Fuente: Elaboración propia

De la tabla 7 y figura 3 se puede decir que la serie original asume un comportamiento fluctuante mayor que el de la predicción del modelo no lineal de RNA y la predicción no sigue con exactitud el comportamiento de la serie objetivo, pero se logra capturar algunos efectos aleatorios.

Muchos trabajos de investigación le confieren a los modelos de redes neuronales propiedades muy superiores a los modelos lineales, los resultados obtenidos en este trabajo de investigación demuestran que para este caso específico el modelo Perceptrón Multicapa presentó un mejor desempeño que el modelo lineal de acuerdo a las medidas de evaluación de predicción aunque las diferencias entre las medidas no fueron significativamente grandes.

Tabla 8. Comparación de tendencias

FECHA (m/d/a)	TENDENCIA ORIGINAL	TENDENCIA MODELO ARIMA(4,1,4)	TENDENCIA MODELO1
11/29/2005			
11/30/2005	Baja	Sube	Baja
12/01/2005	Sube	Sube	Sube
12/02/2005	Sube	Sube	Sube
12/05/2005	Sube	Baja	Baja
12/06/2005	Baja	Baja	Sube
12/07/2005	Sube	Baja	Sube
12/08/2005	Sube	Sube	Baja
12/09/2005	Sube	Sube	Sube
12/12/2005	Sube	Sube	Sube
12/13/2005	Baja	Sube	Sube
12/14/2005	Baja	Baja	Baja
12/15/2005	Baja	Baja	Sube
12/16/2005	Sube	Sube	Sube
12/19/2005	Sube	Sube	Baja
12/20/2005	Baja	Sube	Baja
12/21/2005	Baja	Sube	Sube
12/22/2005	Sube	Sube	Baja
12/23/2005	Sube	Sube	Baja
12/28/2005	Sube	Baja	Sube
12/29/2005	Baja	Sube	Sube

Fuente: Elaboración propia

Predecir el precio exacto del precio del oro es una tarea excesivamente compleja, es por ello que, conocer la tendencia se convierte en el foco y objetivo principal de los inversionistas; en esta medida, se muestra el comportamiento tendencial diario para la serie original y la predicha por el modelo ARIMA(4,1,4) y Perceptrón Multicapa en la tabla 8.

Se puede notar que los dos modelos presentan variaciones en su tendencia, el modelo lineal por su parte presenta 12 aciertos de 20 posibles mientras que el de Redes Neuronales Artificiales logró capturar el efecto tendencial para 10 días de 20 posibles, por lo cual, el modelo lineal resulta ser mejor en la predicción de la tendencia que el de RNA para este estudio.

Se concluye finalmente, que el modelo1 de Redes Neuronales Artificiales es el mejor modelo de este trabajo de investigación para la predicción del precio del oro, definido a partir de las medidas de evaluación de predicción MAE y RMSE, pero en la predicción de

la tendencia el modelo lineal ARIMA(4,1,4) es quien demuestra una mayor cantidad de aciertos.

Otro aspecto importante a resaltar, es que como variables explicativas se están empleando solamente rezagos del precio oro sin tener en cuenta otras variables como por ejemplo el índice Dow-Jones, el precio del dólar, las tasas de interés, la masa de dinero, entre muchas otras variables que pueden mejorar la calidad de las predicciones pero que implicarían la realización de pruebas y técnicas más elaboradas en los modelos RNA.

4. CONCLUSIONES

- El mejor modelo de series de tiempo de este estudio de acuerdo a las medidas de evaluación MAE y RMSE es la especificación ARIMA(4,1,4), el cual presenta aleatoriedad en su predicción logrando acercarse en diversos puntos a la serie financiera original, pero el modelo Perceptrón Multicapa es quien logra un mayor acercamiento a la serie objetivo durante todo el estudio.
- El modelo Perceptrón Multicapa de RNA que presentó un mejor desempeño tanto en entrenamiento como en validación está compuesto por 8 neuronas en la capa de entrada, 40 neuronas en la capa oculta y 20 iteraciones; su predicción es ascendente y aleatoria, es decir, que este modelo no lineal logra capturar en algunos puntos la tendencia descendente de la serie y los efectos no incluidos en el entrenamiento, dicho de otra manera, tiene capacidad de generalización.
- La especificación ARIMA(4,1,4) demostró tener una mayor cantidad de aciertos en la modelación de la tendencia que el modelo de redes neuronales artificiales, sin embargo, otras especificaciones Perceptrón Multicapa lograron capturar la tendencia de una mejor manera pero no sobre el mismo nivel, por lo cual las medidas de evaluación de predicción eran más grandes.
- Es importante resaltar que las redes neuronales no son una herramienta mágica, sino que trabajan en el reconocimiento de patrones y su principal ventaja es su alta capacidad para el procesamiento de datos. Aunque este estudio encontró que el modelo no lineal sólo presentó un mejor desempeño en la predicción del precio y no de la tendencia, también es claro que sólo fueron empleados rezagos del precio del oro como variables explicativas, lo cual puede ser insuficiente para la modelación de este tipo de *commodities*, pero implementar otro tipo de variables trae consigo la implementación de técnicas más elaboradas y rigurosas.
- Es posible encontrar modelos que logren un ajuste muy bueno con las series financieras, pero es importante tener en cuenta que el factor humano seguirá siendo

fundamental, esto se debe a que detrás de los precios de los activos financieros como *commodities*, acciones, bonos, *swaps*, entre otros, tienen tras de ellos a seres humanos, lo que implica que siempre se verán afectados por muchos otros factores aparte de la evolución del mercado, eventos inesperados como un ataque terrorista, decisiones de un presidente de una economía influyente, el anuncio del pago o no pago de dividendos de una empresa, la renuncia o despido de una persona clave en una organización y otros sucesos que no se pueden anticipar por medios estadísticos, en estos casos sólo un ser humano será capaz de incorporarlos para encontrar el impacto que generará sobre el precio del activo.

REFERENCIAS BIBLIOGRÁFICAS

ARANÍBAR, Jaime y HUMÉREZ, Julio (1996): “Modelos de series de tiempo para el pronóstico de precios de minerales”, *Revista de análisis económico*. Unidad de Análisis de Políticas Sociales y Económicas.

AYALA, Mauricio y CASTILLO, Ruben (2006): *Un Modelo de Predicción para el Valor TRM: Un Acercamiento desde las Redes Neuronales Artificiales*, Publicación de administración y finanzas, Universidad del Valle.

BERTOLOTTO, Manuel (2007): *Los fundamentales del precio del oro*, Publicación de economía, Universidad de San Andrés.

BOTERO, Sergio y CANO, Jovan (2008): “Análisis de series de tiempo para la predicción de los precios de la energía en la bolsa de Colombia”, *Cuadernos de Economía*, Vol. XXVII, No. 48, pp. 173-208.

CAICEDO, Eduardo y LÓPEZ, Jesús (2009): *Una aproximación práctica a las redes neuronales artificiales*, Programa editorial Universidad del Valle.

CHATFIELD, C. (1989): *The Analysis of Time Series: An Introduction*, Cuarta Edición, Chapman y Hall.

CRUZ, Iván (2008): *Pronósticos en el Mercado de Derivados Utilizando Redes Neuronales y Modelos ARIMA: Una Aplicación al Cete de 91 Días en el MexDer*, Tesis de maestría, Universidad Nacional Autónoma de México.

FOIX, Cristian y WEBER, Richard (2007): “Pronóstico del precio del cobre mediante redes neuronales”, *Revista Ingeniería de Sistemas*, Vol. 21, pp. 63-86.

FOSTER, Ernest (2002): *Commodity Futures Price Prediction, an Artificial Intelligence Approach*, Thesis submitted to the graduate faculty, University of Georgia.

FRIZ, Rodolfo (2004): *Redes Neuronales Aplicadas a la Predicción del Precio del Oro y Medición de la Robustez de los Resultados Utilizando Bootstrap*, Tesis de pregrado, Universidad de Chile.

GÓMEZ, Fernando; MADARIAGA, José y UGARTE José (1988): “La eficiencia en el mercado bursátil español”, *Actualidad Financiera*, No. 42, pp. 2238-2250.

GUTIÉRREZ, Marcelo (2004): *Administración de Carteras con Redes Neuronales Mediante Metodología Rolling*, Tesis de pregrado, Universidad de Chile.

HEREDIA, Rodrigo (2009): *El Oro: Pros y Contrás*, Publicación Grupo Financiero IXe

HILERA, J. R. y MARTÍNEZ, V. J. (1994): *Redes Neuronales Artificiales. Fundamentos, Modelos y Aplicaciones*, Rama.

JALIL, Munir y MISAS, Martha (2007): “Evaluación de pronóstico del tipo de cambio utilizando redes neuronales y funciones de pérdida asimétricas”, *Revista Colombiana de Estadística*, Vol. 30, No. 1, pp. 143-161.

LAWRENCE, Ramón (1997): *Using Neural Networks to Forecast Stock Market Prices*, Thesis submitted for degree, University of Manitoba.

MARTÍNEZ, Josefina (2007): *La Hipótesis de los Mercados Eficientes, el Modelo del Juego Justo y el Recorrido Aleatorio*, Publicación de la Universidad de A Coruña.

MISAS, Martha; LÓPEZ, Enrique y QUERUBÍN, Pablo (2002): “La inflación en Colombia: Una aproximación desde las redes neuronales”, *Subgerencia de Estudios Económicos*, Banco de la República.

RICHTER, Katharine y TORO, Daniela (2006): *Redes Neuronales Artificiales, Una Herramienta para las Finanzas*, Tesis de pregrado, Universidad de Concepción.

RODRÍGUEZ, Luis y FERMÍN, José (2006): “Mercado eficiente y caminata aleatoria en la bolsa de valores de Caracas”, *Red de Revistas Científicas de América Latina y el Caribe, España y Portugal*, Vol. 31, No. 012, pp. 888-893.

SANTANA, Juan (2006); “Predicción de series temporales con redes neuronales: Una aplicación a la inflación colombiana”, *Revista Colombiana de Estadística*, Vol. 29, No. 1, pp. 77-92.

TORO, Eliana; GARCÉS, Alejandro y GALVIS, Juan (2007): “Redes neuronales aplicadas a la predicción de precios en bolsa de energía”, *Revista Colombiana de Tecnologías de Avanzada*, Vol. 2, No. 10, pp. 10-14.

VELÁSQUEZ, Juan y MONTOYA, Santiago (2005): “Modelado del índice de precios al consumidor usando un modelo híbrido basado en redes neuronales artificiales”, *Revista DYNA*, Vol. 72, No. 147, pp. 85-93.

VELÁSQUEZ, Juan y ALDANA, Mario (2007): “Modelado del precio del café colombiano en la bolsa de Nueva York usando redes neuronales artificiales”, *Revista Facultad Nacional de Agronomía, Medellín*, Vol. 60, No. 2, pp. 4129-4144.

VILLADA, Fernando; CADAVID, Diego y MOLINA, Juan (2008): “Pronóstico del precio de la energía eléctrica usando redes neuronales artificiales”, *Revista Facultad de Ingeniería de la Universidad de Antioquia*, No. 44, pp. 111-118.

WITKER, Jorge y VALENZUELA, Elvira (2002): *El Sistema Monetario Internacional Contemporáneo*, Publicación de la Universidad Nacional Autónoma de México.

Z. Ismail, A. Yahya y A. Shabri (2009): “Forecasting gold price using multiple linear regression method”, *American Journal of Applied Sciences*, Vol. 6, pp. 1509-1514.

ANEXOS

ANEXO A. PRUEBA RANGO – MEDIA Y TEST DICKEY FULLER

Cuando una serie temporal no es estacionaria, puede ser convertida a estacionaria realizando transformaciones como son:

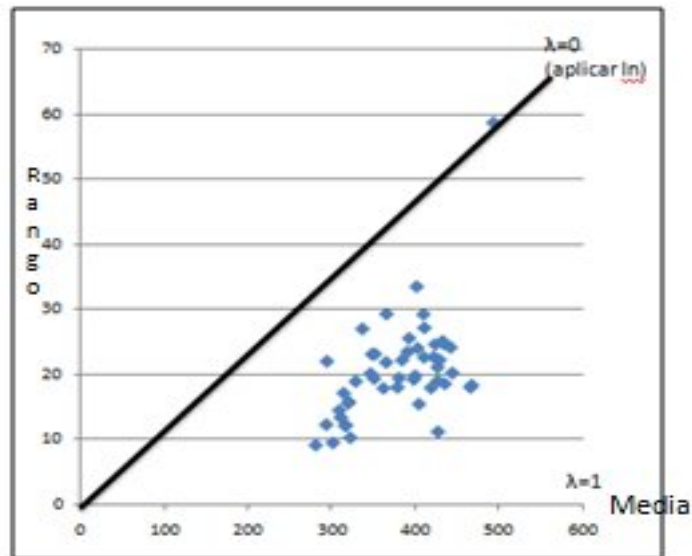
1. Transformación Box-Cox: Determinación de λ
2. Diferencias sucesivas: Determinación de d

Cuando la serie no es estacionaria en varianza se realiza la transformación Box-Cox:

$$Y_t^{(\lambda)} = \begin{cases} \frac{Y_t^{(\lambda)} - 1}{\lambda} & \text{si } \lambda \neq 0 \\ \ln(Y_t) & \text{si } \lambda = 0 \end{cases} \quad (1)$$

Para el valor $\lambda = 0$ se aplica una transformación logarítmica y para el valor $\lambda = 1$ no se modifican los datos. Con el fin de detectar si la serie es estacionaria en varianza se realiza un diagrama Rango/Media y se encuentra el valor de λ según la pendiente de la recta. Si la recta presenta pendiente > 0 se transforma la serie, pero si la recta no presenta pendiente no se transforma la serie.

Figura A1. Rango-Media



Fuente: Elaboración propia

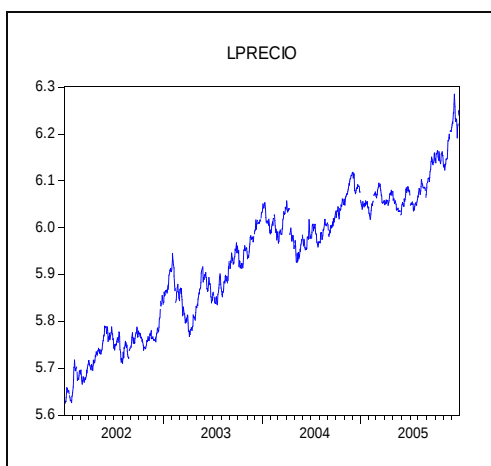
De acuerdo a la información que brinda el gráfico, se concluye que se debe realizar transformación logarítmica a la serie, además que realizar esta transformación a las series financieras es muy común porque presentan algunas ventajas.

Ahora, cuando la serie no es estacionaria en media se transforma calculando diferencias sucesivas:

$$W_t = \Delta^d Y_t^{(A)} \quad (3)$$

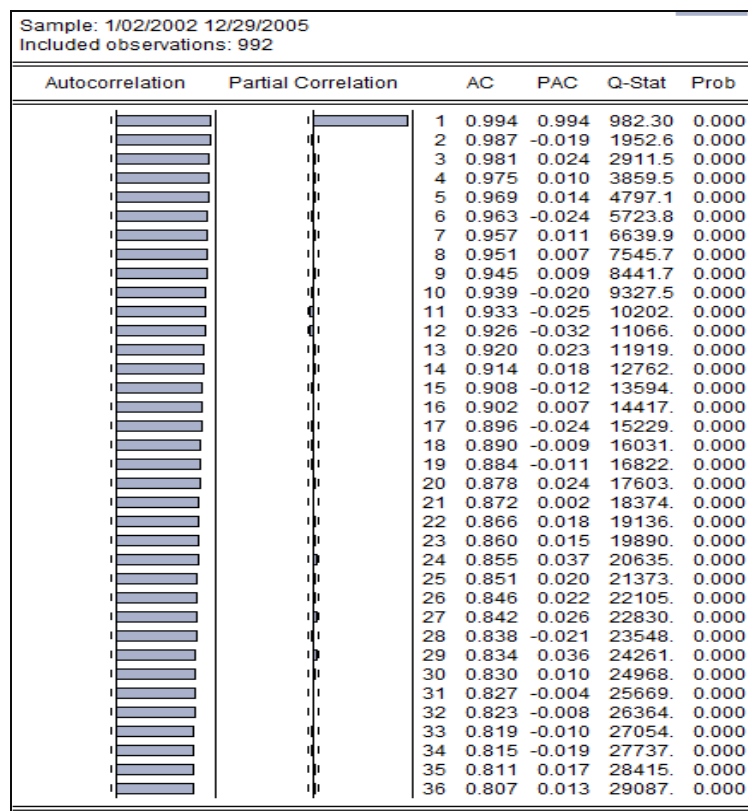
Para conocer el valor de d se debe analizar el gráfico de la serie original y su correlograma.

Figura A2. Comportamiento serie lprecio



Fuente: Elaboración propia

Tabla A1. Correlograma serie lprecio

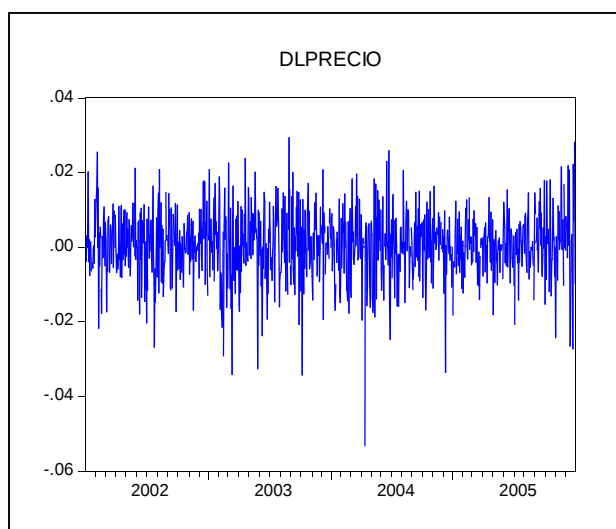


Fuente: Elaboración propia

Al observar el gráfico y correlograma de la serie lprecio, se nota que la serie no oscila en torno a un valor promedio, la FACM comienza en un valor muy próximo a 1 y desciente muy lentamente y la FACPM tiene el primer valor cercano a 1 y el resto son casi cero; este comportamiento es muy característico de una serie no estacionaria por lo cual, se sugiere diferenciar.

Se procede a diferenciar la serie y se realiza el mismo procedimiento, analizar el gráfico del comportamiento de la serie y el correlograma de la serie dlprecio.

Figura A3. Comportamiento serie dlprecio



Fuente: Elaboración propia

Tabla A2. Correlograma serie dlprecio

Sample: 1/02/2002 12/29/2005 Included observations: 991						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.017	0.017	0.2995	0.584
		2	-0.000	-0.001	0.2995	0.861
		3	-0.012	-0.012	0.4465	0.930
		4	0.023	0.023	0.9732	0.914
		5	-0.003	-0.004	0.9834	0.964
		6	-0.016	-0.016	1.2369	0.975
		7	-0.043	-0.042	3.0666	0.879
		8	-0.031	-0.030	4.0302	0.854
		9	-0.001	-0.001	4.0323	0.909
		10	-0.013	-0.014	4.2094	0.937
		11	-0.005	-0.003	4.2329	0.963
		12	-0.040	-0.039	5.8362	0.924
		13	-0.032	-0.033	6.8804	0.908
		14	0.010	0.009	6.9899	0.935
		15	0.031	0.027	7.9374	0.926
		16	0.053	0.052	10.753	0.824
		17	0.037	0.036	12.125	0.792
		18	0.018	0.014	12.438	0.824
		19	-0.009	-0.014	12.514	0.863
		20	-0.012	-0.019	12.666	0.891
		21	-0.042	-0.044	14.499	0.847
		22	-0.023	-0.019	15.018	0.861
		23	-0.014	-0.008	15.224	0.886
		24	-0.029	-0.024	16.061	0.886
		25	-0.039	-0.036	17.637	0.857
		26	-0.004	-0.002	17.653	0.888
		27	-0.013	-0.012	17.814	0.909
		28	-0.026	-0.024	18.521	0.912
		29	-0.009	-0.006	18.609	0.931
		30	0.032	0.031	19.669	0.925
		31	0.034	0.026	20.871	0.915
		32	-0.012	-0.024	21.019	0.931
		33	0.008	-0.003	21.079	0.946
		34	-0.053	-0.064	23.958	0.900
		35	-0.013	-0.020	24.139	0.916
		36	-0.004	-0.003	24.157	0.934

Fuente: Elaboración propia

Se puede notar que el gráfico de la serie oscila en torno a un valor promedio, pero la FACM no decrece rápidamente, en este caso todos sus valores son muy cercanos a cero pero decrecen y aumentan continuamente y algo parecido sucede con la FACPM. Este comportamiento sugiere un modelo de caminata aleatoria.

Con el fin de contrastar los resultados obtenidos sobre la estacionariedad en media de la serie, se procede a aplicar el test Dickey-Fuller Aumentado si existe raíz unitaria.

Ho: La serie tiene raíz unitaria

Ha: La serie no tiene raíz unitaria

Tabla A3. Test Dickey Fuller Aumentado, intercepto

Variable dependiente: D(LPRECIO)			
Variables exógenas: Intercepto			
		t-estadístico	Prob.
Test D-F		-0,9772	0,763
Test valores críticos			
	1%	-3,4367	
	5%	-2,8642	
	10%	-2,5682	

Fuente: Cálculos propios

Tabla A4. Test Dickey Fuller Aumentado, ninguno

Variable dependiente: D(LPRECIO)			
Variables exógenas: Ninguno			
		t-estadístico	Prob.
Test D-F		2,1121	0,9921
Test valores críticos			
	1%	-2,5672	
	5%	-1,9411	
	10%	-1,6164	

Fuente: Cálculos propios

Tabla A5. Test Dickey Fuller Aumentado, intercepto y tendencia

Variable dependiente: D(LPRECIO)			
Variables exógenas: Tendencia e intercepto			
		t-estadístico	Prob.
Test D-F		-3,5126	0,0385
Test valores críticos	1%	-3,9673	
	5%	-3,4143	
	10%	-3,1293	

Fuente: Cálculos propios

De acuerdo a los resultados obtenidos del Test de Dickey Fuller aumentado, se puede concluir que no se puede rechazar la hipótesis nula de raíz unitaria, por lo cual se hace necesario el proceso de primeras diferencias para la serie LPRECIO. Ahora se realiza nuevamente el test de D-F para la serie DLPRECIO.

Tabla A6. Test Dickey Fuller Aumentado, intercepto

Variable dependiente: D(DLPRECIO)			
Variables exógenas: Intercepto			
		t-estadístico	Prob.
Test D-F		-3,087	0
Test valores críticos	1%	-3,4367	
	5%	-2,8642	
	10%	-2,5682	

Fuente: Cálculos propios

Tabla A7. Test Dickey Fuller Aumentado, ninguno

Variable dependiente: D(DLPRECIO)			
Variables exógenas: Ninguna			
		t-estadístico	Prob.
Test D-F		-3,0748	0

Test valores críticos	1%	-2,5673
	5%	-1,9411
	10%	-1,6164

Fuente: Cálculos propios

Tabla A8. Test Dickey Fuller Aumentado, intercepto y tendencia

Variable dependiente: D(DLPRECIO)			
Variables exógenas: Intercepto, tendencia			
		t-estadístico	Prob.
Test D-F		-3,0855	0
Test valores críticos	1%	-3,9673	
	5%	-3,4143	
	10%	-3,1293	

Fuente: Cálculos propios

De acuerdo a los resultados obtenidos al aplicar nuevamente el test Dickey Fuller Aumentado a la serie, se concluye que la serie DLPRECIO no necesita ser diferenciada, es decir, que la serie ya es estacionaria con una diferenciación.

ANEXO B. COMPROBACIÓN DE SUPUESTOS DE LOS MODELOS

Tabla A9. Test de comprobación de modelos

Modelo	Media	Test ARCH de heterocedasticidad		Test de Normalidad Jarque-Bera	
		Obs*R-squared	Prob. Chi-Square(1)	Jarque-Bera	Prob.
ARIMA(2,1,2)	-0,000000785	0,0293	0,8641	209,5373	0
ARIMA(4,1,4)	-0,00000178	0,0002	0,9892	188,1688	0
CAMINATA ALEATORIA	5,85E-19	0,0531	0,8178	194,796	0

Fuente: Cálculos propios

Tabla A10. Test Ljung Box de autocorrelación ARIMA(2,1,2)

Date: 01/14/11 Time: 00:27					
Sample: 1/07/2002 12/29/2005					
Included observations: 989					
Q-statistic probabilities adjusted for 4 ARMA term(s)					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1	0.017	0.017	0.2772
		2	-0.004	-0.004	0.2929
		3	-0.006	-0.006	0.3252
		4	0.018	0.018	0.6448
		5	-0.001	-0.002	0.6468
		6	-0.013	-0.013	0.8209
		7	-0.049	-0.048	3.1965
		8	-0.025	-0.024	3.8017
		9	-0.005	-0.005	3.8284
		10	-0.014	-0.014	4.0252
		11	-0.001	0.001	4.0256
		12	-0.046	-0.045	6.1069
		13	-0.028	-0.028	6.9022
		14	0.010	0.008	6.9954
		15	0.028	0.024	7.7771
		16	0.057	0.057	11.066
		17	0.033	0.031	12.172
		18	0.019	0.017	12.528
		19	-0.006	-0.012	12.568
		20	-0.017	-0.023	12.846
		21	-0.037	-0.038	14.256
		22	-0.024	-0.022	14.855
		23	-0.015	-0.008	15.093
		24	-0.024	-0.020	15.686
		25	-0.045	-0.043	17.762
		26	0.001	0.003	17.762
		27	-0.014	-0.014	17.971
		28	-0.027	-0.025	18.734
		29	-0.006	-0.002	18.769
		30	0.028	0.027	19.576
		31	0.036	0.029	20.900
		32	-0.012	-0.024	21.039
		33	0.005	-0.005	21.064
		34	-0.049	-0.060	23.507
		35	-0.016	-0.023	23.786
		36	-0.003	-0.002	23.794

Fuente: Elaboración propia

Tabla A11. Test Ljung Box de autocorrelación ARIMA(4,1,4)

Date: 01/14/11 Time: 00:36					
Sample: 1/09/2002 12/29/2005					
Included observations: 987					
Q-statistic probabilities adjusted for 8 ARMA term(s)					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.010	0.010	0.0904	
		2 -0.002	-0.002	0.0958	
		3 -0.002	-0.002	0.1018	
		4 0.007	0.007	0.1459	
		5 0.006	0.006	0.1827	
		6 -0.014	-0.014	0.3716	
		7 -0.050	-0.049	2.8303	
		8 -0.016	-0.015	3.0924	
		9 -0.014	-0.014	3.2792	0.070
		10 -0.012	-0.012	3.4192	0.181
		11 0.001	0.002	3.4204	0.331
		12 -0.049	-0.049	5.8431	0.211
		13 -0.019	-0.020	6.2174	0.286
		14 0.006	0.003	6.2519	0.396
		15 0.027	0.024	6.9653	0.433
		16 0.058	0.056	10.321	0.243
		17 0.026	0.025	11.021	0.274
		18 0.023	0.022	11.545	0.317
		19 -0.007	-0.013	11.594	0.395
		20 -0.019	-0.023	11.958	0.449
		21 -0.035	-0.037	13.205	0.432
		22 -0.030	-0.028	14.096	0.443
		23 -0.014	-0.008	14.296	0.503
		24 -0.020	-0.017	14.708	0.546
		25 -0.045	-0.042	16.778	0.469
		26 0.004	0.006	16.798	0.537
		27 -0.018	-0.017	17.109	0.582
		28 -0.030	-0.028	18.043	0.585
		29 -0.003	-0.002	18.052	0.646
		30 0.028	0.027	18.856	0.654
		31 0.040	0.033	20.476	0.613
		32 -0.011	-0.022	20.590	0.663
		33 0.000	-0.009	20.590	0.715
		34 -0.050	-0.062	23.129	0.626
		35 -0.018	-0.025	23.480	0.659
		36 -0.001	-0.001	23.482	0.709

Fuente: Elaboración propia

Tabla A12. Test Ljung Box de autocorrelación Caminata Aleatoria

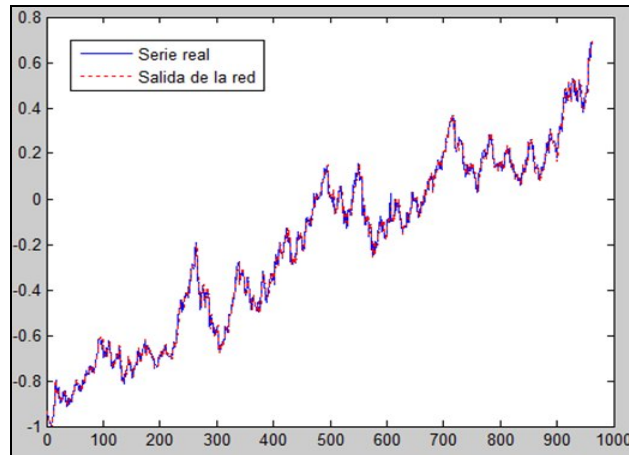
Date: 01/14/11 Time: 00:39						
Sample: 1/03/2002 12/29/2005						
Included observations: 991						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.017	0.017	0.2995	0.584
		2	-0.000	-0.001	0.2995	0.861
		3	-0.012	-0.012	0.4465	0.930
		4	0.023	0.023	0.9732	0.914
		5	-0.003	-0.004	0.9834	0.964
		6	-0.016	-0.016	1.2369	0.975
		7	-0.043	-0.042	3.0666	0.879
		8	-0.031	-0.030	4.0302	0.854
		9	-0.001	-0.001	4.0323	0.909
		10	-0.013	-0.014	4.2094	0.937
		11	-0.005	-0.003	4.2329	0.963
		12	-0.040	-0.039	5.8362	0.924
		13	-0.032	-0.033	6.8804	0.908
		14	0.010	0.009	6.9899	0.935
		15	0.031	0.027	7.9374	0.926
		16	0.053	0.052	10.753	0.824
		17	0.037	0.036	12.125	0.792
		18	0.018	0.014	12.438	0.824
		19	-0.009	-0.014	12.514	0.863
		20	-0.012	-0.019	12.666	0.891
		21	-0.042	-0.044	14.499	0.847
		22	-0.023	-0.019	15.018	0.861
		23	-0.014	-0.008	15.224	0.886
		24	-0.029	-0.024	16.061	0.886
		25	-0.039	-0.036	17.637	0.857
		26	-0.004	-0.002	17.653	0.888
		27	-0.013	-0.012	17.814	0.909
		28	-0.026	-0.024	18.521	0.912
		29	-0.009	-0.006	18.609	0.931
		30	0.032	0.031	19.669	0.925
		31	0.034	0.026	20.871	0.915
		32	-0.012	-0.024	21.019	0.931
		33	0.008	-0.003	21.079	0.946
		34	-0.053	-0.064	23.958	0.900
		35	-0.013	-0.020	24.139	0.916
		36	-0.004	-0.003	24.157	0.934

Fuente: Elaboración propia

De las tablas A9, A10, A11 y A12 se puede concluir que los modelos cumplen con el proceso de verificación, en la medida que presentan media muy cercana a 0, varianza constante ya que la prueba de heterocedasticidad demuestra que no hay efectos ARCH y de acuerdo a las pruebas Ljung Box de autocorrelación serial ninguno de los modelos presenta problemas de autocorrelación, aunque los problemas de no normalidad persisten.

ANEXO C. INDICADORES DEL MEJOR MODELO SELECCIONADO DE RNA: RENDIMIENTOS Y AJUSTE

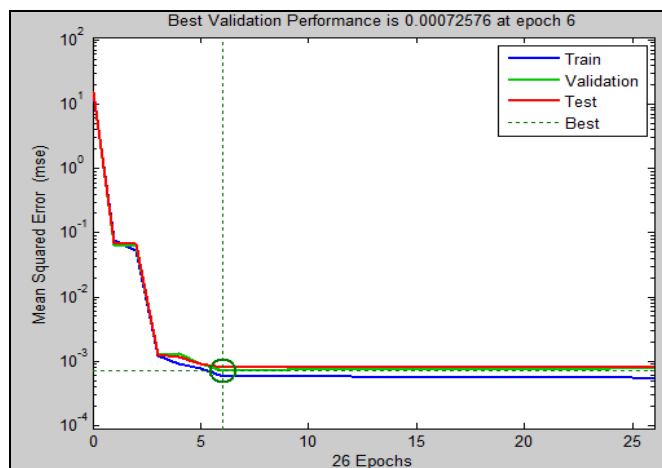
Figura A4. Conjunto de datos de entrenamiento y validación



Fuente: Elaboración propia

En la figura A4. son presentados los datos de entrenamiento y validación para el mejor modelo seleccionado, donde la red neuronal aprenderá la dinámica de la función $y=f(t)$. Para el entrenamiento llamado *serie real* es tomada la línea azul y para el proceso de validación llamado *salida de la red* es tomada la línea punteada roja, se puede notar que los datos de validación y entrenamiento logran acoplarse muy bien.

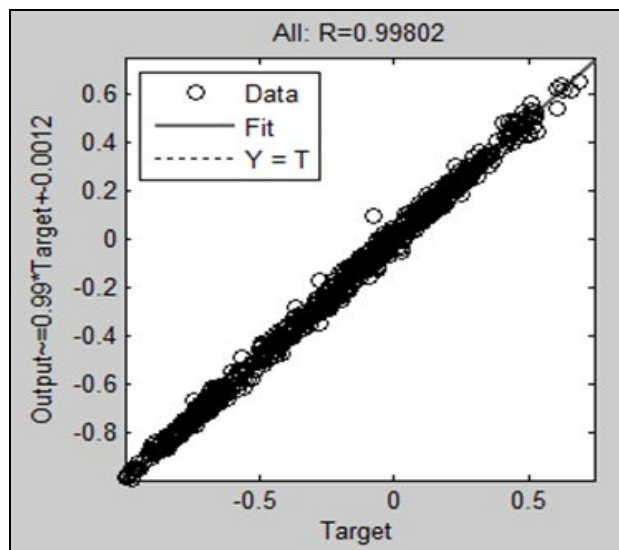
Figura A5. Rendimiento del entrenamiento, validación y test



Fuente: Elaboración propia

La figura A5. Muestra el comportamiento asumido por la red durante el entrenamiento, validación y prueba, cuando la línea azul cae y la línea verde asciende, quiere decir que la red se encuentra sobre-entrenada puesto que el error seguirá bajando pero para el conjunto de los datos de entrenamiento, pero para la validación el error tendrá una tendencia ascendente. El comportamiento obtenido para esta red muestra que el error mínimo fue encontrado en la sexta iteración y que a pesar de haber fijado 20 iteraciones para el proceso de entrenamiento, el error no se hizo más pequeño, es decir que la red no se encuentra sobre-entrenada.

Figura A6. Ajuste de entrenamiento, validación y test en conjunto



Fuente: Elaboración propia

Se sabe que los datos de entrenamiento, validación y prueba presentan un buen ajuste cuando el indicador R es cercano a 1, de la figura A6 se puede concluir que los datos presentaron un muy buen ajuste en predicción.