

## Predicción del IPC mexicano combinando modelos econométricos e inteligencia artificial

Luis Manuel León Anaya  
Víctor Manuel Landassuri Moreno  
Héctor Rafael Orozco Aguirre<sup>1</sup>  
Maricela Quintana López  
Universidad Autónoma del Estado de México

(Primera recepción: 10/junio/2017, última recepción: 14/marzo/2018  
aceptado: 25/mayo/2018)

### Resumen

El objetivo de este trabajo es descomponer los factores de comportamiento del Índice de Precios y Cotizaciones (IPC) mexicano para ser pronosticado mediante modelos econométricos y redes neuronales artificiales evolutivas. La metodología empleada consiste en reducir la complejidad de análisis y eliminar el ruido en los datos del IPC mediante la descomposición empírica en modos (DEM), combinando las funciones de modo intrínseco (FMIs) resultantes con las variantes de los modelos autorregresivo integrado de promedio móvil (ARIMA) y autorregresivo con heterocedasticidad condicional (ARCH), y el algoritmo de selección de características de programación evolutiva de redes (FS-EPNet) para pronosticar su comportamiento. La configuración experimental y resultados se presentan y analizan mediante tres fases de predicción del IPC. Las limitaciones son que el IPC mexicano no es estacionario, implicando que algunas FMIs tampoco lo sean. La originalidad consiste en la combinación de la DEM con el algoritmo FS-EPNet para analizar la evolución del mercado bursátil mexicano a través de su IPC, con lo cual se demuestra y concluye que genera una mejor predicción que la obtenida a partir de los datos originales.

*Clasificación JEL: C22, C45, C53*

*Palabras clave: Pronóstico, Índice Bursátil, Series de Tiempo, Descomposición Empírica en Modos, Redes Neuronales Artificiales Evolutivas*

## Mexican IPC Prediction Combining Econometric Models and Artificial Intelligence

### Abstract

The purpose of this paper is to decompose the behavioral factors of the Mexican Price and Quotation Index (IPC for its acronym in Spanish) to be forecast using econometric models and evolutionary artificial neural networks. The methodology used consists on reducing the analysis complexity and eliminating the noise in the IPC data through empirical mode decomposition (EMD), combining the intrinsic mode functions (IMFs) resulting with the variants of the autoregressive integrated mobile average (ARIMA) and autoregressive conditional heteroskedasticity (ARCH) models, as well as the algorithm for selection of characteristics of evolutionary network programming (FS-EPNet) to forecast its behavior. The experimental configuration and results are shown and are analyzed using three prediction phases of the IPC. The limitations are that the Mexican IPC is

<sup>1</sup>Universidad Autónoma del Estado de México  
Blvd. Universitario s/n Predio San Javier (Km. 11.5 Carretera Atizapán-Nicolás Romero) Atizapán de Zaragoza, Estado de México. Tel 01 (55) 5827-0703, Ext. 1157. Correo Electrónico: hrorozcoa@uaemex.mx

**Abstract**

not stationary, which implies that some IMFs are also not stationary. The originality of this consists on the combination of DEM with the FS-EPNet algorithm to analyze the evolution of the Mexican Stock Exchange through its IPC, which is used to show and conclude that it generates a better prediction than that obtained from the original data.

*JEL Classification: C22, C45, C53*

*Keywords: Forecast, Stock Index, Time Series, Empirical Mode Decomposition, Evolutionary Artificial Neural Networks*

**1. Introducción**

En el ámbito económico, el lograr una predicción certera del comportamiento futuro del mercado de valores resulta de especial interés, puesto que es una pieza fundamental para el desarrollo de la industria y del comercio, ya que incide en la economía de una nación. El mercado bursátil puede ser atractivo para los inversionistas puesto que existe una gran liquidez ya que se pueden obtener altas tasas de rendimiento a diferencia de otros instrumentos de inversión, por el hecho de que hay un gran potencial de ganancia. Por esas razones tanto el gobierno, la industria, el banco central y los inversionistas mantienen una estrecha vigilancia en el comportamiento del mercado de valores, y sobre los diferentes acontecimientos que pudieran repercutir sobre él.

Típicamente, el análisis y la posterior predicción del mercado de valores se realizan a través de los diferentes índices. La construcción de un índice es una mezcla de diversas organizaciones que se encuentran en sectores distintos de la economía y por ende dicho indicador es confeccionado de manera artificial. Aunado a ello, una gran cantidad de circunstancias inciden en el mercado de valores como pueden ser las tasas de interés, sucesos políticos, políticas monetarias, factores tecnológicos, situaciones económicas, y expectativas de los inversionistas. Es un hecho bien conocido que las series de tiempo (ST) financieras poseen características inherentes de no estacionalidad, no linealidad, volatilidad y ruido. Tradicionalmente, se considera que el ruido afecta la estimación óptima de los parámetros de los modelos de predicción. De esta forma, se puede considerar que la obtención de una predicción acertada del mercado bursátil es una tarea compleja e incluso, si se toma como válida la hipótesis del mercado eficiente (EMH, por sus siglas en inglés Efficient-Market Hypothesis) (Fama, 1970), la elaboración de una predicción precisa es una tarea imposible. Se considera que el análisis del comportamiento del mercado bursátil mediante métodos matemáticos comienza con la tesis doctoral de Louis Bachelier (1900), intitulada "La Teoría de la Especulación". En dicha investigación, Bachelier mediante un proceso estocástico, introduce la aplicación de la hipótesis del paseo aleatorio a la cotización de opciones de la bolsa de París (Hol, 2003). Posteriormente, Eugene Fama (1965; 1970) se encargó de popularizar y desarrollar la EMH en su trabajo doctoral, donde sentó las bases, argumentando que en cualquier mercado existe una gran cantidad de inversores bien informados, y que el valor presente de los activos refleja toda la información disponible.

En las últimas décadas del siglo anterior, los métodos tradicionales estadísticos y econométricos como el modelo de valoración de activos (CAPM, por sus siglas en inglés Capital Asset Pricing Model) (Sharpe, 1964), el modelo de Black-Scholes (1973), la regresión lineal, los modelos autorregresivo integrado de media móvil (ARIMA, por sus siglas en inglés Autoregressive Integrated Moving Average), de heterocedasticidad condicional

auto-regresiva (ARCH, por sus siglas en inglés Autoregressive Conditional Heteroskedasticity) y generalizado ARCH (GARCH, por sus siglas en inglés Generalized ARCH), así como los vectores autorregresivos (VAR) han sido ampliamente utilizados en la predicción de los mercados financieros. Los que tuvieron un mayor auge para el análisis y predicción de los mercados financieros en la década de los ochentas fueron los modelos ARIMA. Posteriormente, en la década de los noventa la volatilidad de los activos toma un papel preponderante y se deja de lado la media de los datos, siendo los más populares para el pronóstico del mercado bursátil los modelos ARCH y sus derivaciones (Alonso García, 2009). Sin embargo, los modelos tradicionales trabajan de manera adecuada cuando los datos están normalizados y sujetos a cumplir con media estacionaria, coeficientes no negativos, etc.

Las ST financieras se comportan de una manera muy similar a un proceso de caminata aleatoria, considerando la tasa de éxito generado, pero suelen ser muy ruidosas, debido a la aleatoriedad en las variaciones del día a día (Magdon-Ismail, Nicholson, Abu-Mostafa, 1998). Además, son no lineales, no estacionarias, y altamente volátiles. Tomando en consideración los factores que inciden en el análisis y la predicción del mercado de valores, dado que los métodos tradicionales se encuentran limitados, surge la motivación de emplear métodos o modelos alternos que puedan capturar, de un modo más fiable, el comportamiento del mercado bursátil (McNelis, 2005; Brooks, 2014; Abbasi, Aghaei, fard, 2015). Una alternativa son las redes neuronales artificiales (RNAs) provenientes de la inteligencia artificial (IA), denominadas también como modelos no paramétricos. Las RNAs tienen la habilidad de manejar información tanto lineal como no lineal con un mayor grado de adaptabilidad, modificando sus pesos de entrada, ofreciendo tolerancia a fallos debido a su capacidad de almacenar información de forma redundante. Una RNA es una herramienta flexible, no sujeta a supuestos que restrinjan su modelado.

Como panorama de esta investigación, se presenta la descomposición empírica en modos (DEM) del índice de precios y cotizaciones (IPC) mexicano en sus factores de comportamiento representados por funciones de modo intrínseco (FMIs); con el propósito de generar y analizar su pronóstico al combinar variantes de los modelos ARIMA-ARCH, y el algoritmo de selección de características de programación evolutiva de redes (FS-EPNet, por sus siglas en inglés Feature Selection of Evolutionary Programming of Networks). Para este fin, la configuración experimental se realiza mediante tres fases de predicción del IPC. Analizando los resultados obtenidos, es posible concluir que la predicción del IPC en combinación con la DEM genera mejores resultados que aquella obtenida a partir de datos sin descomposición en FMIs.

El resto del documento tiene la siguiente estructura, el segundo apartado versa sobre la revisión literaria. En el tercer apartado se introduce el método de descomposición empírica en modos (DEM) apreciando como se obtienen funciones de modo intrínseco (FMIs). Posteriormente, el cuarto apartado esboza un panorama general del modelo ARIMA-ARCH. El quinto apartado describe al modelo no paramétrico, que abarca a las RNAs, así como también las particularidades del algoritmo de selección de características de programación evolutiva de redes (FS-EPNet, por sus siglas en inglés Feature Selection of Evolutionary Programming of Networks). La configuración experimental es expuesta en el quinto apartado, mientras que la fase experimental y los resultados son exhibidos en el sexto apartado. Finalmente, se presentan las conclusiones derivadas del estudio.

## 2. Revisión literaria

El empleo de las RNAs en el análisis de los mercados financieros no es un tema reciente. En la década de los 80's, White (1988) comenzó la aplicación de las RNAs al mercado bursátil, utilizando un modelo de tipo perceptrón multicapa (MLP, por sus siglas en inglés MultiLayer Perceptron) y como método de aprendizaje el algoritmo de retropropagación (BP, por sus siglas en inglés BackPropagation). En dicho estudio, el MLP se utilizó como un modelo de regresión lineal simple sobre los datos históricos de la cotización de IBM, donde se determinó que el mercado no es predecible. Posteriormente, se entrenó el MLP con un conjunto de datos derivados de los retornos accionarios IBM por 1,000 días. En la primera fase del pronóstico, las RNAs arrojaron un pronóstico notablemente mejor que el modelo lineal. Sin embargo, con 2 grupos de datos distintos del conjunto de entrenamiento, el resultado fue desfavorable para ambos modelos, siendo el porcentaje de aciertos cercano al 50 %.

A partir del trabajo mencionado, el uso de las RNAs en el análisis y predicción de los mercados financieros ha prosperado y se ha popularizado de manera notable, sobre todo en los últimos años. A continuación, se mencionan dos trabajos destacados.

Lee et al. (2008) comparan el rendimiento de un modelo estacional autorregresivo integrado de promedio móvil (SARIMA, por sus siglas en inglés Seasonal Autoregressive Integrated Moving Average) y de una RNA de retropropagación (BPNN, por sus siglas en inglés BackPropagation Neural Network) en la predicción del índice bursátil de Corea del Sur (KOSPI, por sus siglas en inglés KOREA composite Stock Price Index). Los resultados de la investigación demuestran que el modelo SARIMA fue superior en la predicción del indicador KOSPI. Sin embargo, el modelo neuronal demostró superioridad en la predicción de los retornos del mismo.

Por medio del índice Riga Stock Exchange (RSE) hoy en día "Nasdaq Riga", el mercado de valores de Letonia es analizado por Zorin y Borisov (2002). En dicha investigación, se evalúa la predicción de un modelo de RNA y ARIMA. Los resultados indican que el modelo RNA supera al ARIMA. Además, la investigación concluyó que la predicción del indicador RSE es aceptable por un período de dos semanas.

A pesar de que los modelos de RNAs son herramientas útiles para el análisis de los mercados financieros, aún persisten varios factores por resolver al momento de crearlos, debido a que el diseño topológico de una RNA tiene una gran influencia en el desempeño mostrado y en su capacidad de generalización. En este trabajo de investigación, se propone utilizar un algoritmo evolutivo (AE) que facilite la configuración arquitectónica lo más cercana posible a un nivel óptimo, es decir el menor error posible. Para ello, se emplea el algoritmo de selección de características de programación evolutiva de redes (FS-EPNet, por sus siglas en inglés Feature Selection of Evolutionary Programming Networks). Este algoritmo es descrito en (Landassuri-Moreno Bullinaria, 2009), el cual se encarga de la configuración arquitectónica y de los pesos de una RNA de manera evolutiva.

Los datos provenientes del mercado bursátil reflejan otro factor importante que hay que considerar, ya que juegan un papel preponderante en la búsqueda de una predicción precisa del comportamiento del mercado accionario. Dado que la información que se recaba del mercado bursátil es muy ruidosa y compleja, siendo esta la materia prima con la que trabajan los modelos de pronósticos, es muy factible que la tarea predictiva, se convierta en la mayoría de los casos en un esfuerzo infructuoso e imposible de realizar de manera conveniente.

En este trabajo, se aplica el principio de divide y vencerás, usando la técnica de descomposición empírica en modos (DEM) con el fin de aminorar la complejidad de una

ST. De esta forma, la metodología del presente trabajo propone simplificar la tarea predictiva de los modelos ARIMA-ARCH y RNA mediante la división de una tarea predictiva compleja, en subtarefas relativamente más sencillas. En primera instancia, mediante la aplicación de la técnica DEM a la serie original del promedio industrial Dow Jones (DJIA, por sus siglas en inglés Dow Jones Industrial Average), obteniendo así, un número finito y frecuentemente pequeño de las funciones de modo intrínseco (FMIs). Posteriormente, cada componente FMI y el residuo es modelado y predicho de manera individual, por un lado, con un modelo independiente de RNA el cual es diseñado y configurado mediante el algoritmo FS-EPNet. Por otro lado, con un modelo ARIMA-ARCH, que también se le denomina modelo paramétrico de manera indistinta en esta investigación, se pretende que los patrones subyacentes de cada descomposición puedan ser capturados con un grado mayor de exactitud.

La predicción final de cada modelo se obtiene mediante la adición de cada una de las descomposiciones pronosticadas. Finalmente, se toma en cuenta lo que indica la literatura, “que una combinación de pronósticos conlleva a una predicción más precisa” (Adams, 1978; Makridakis et al., 1982; Bessler Chamberlain, 1987; Gonzales Luna, 2011). En lugar de sólo elegir la predicción producida por el modelo que se considere más apto. Por tal razón, en esta investigación se realiza una combinación de pronósticos, ya que se asume que cada modelo, dadas sus características, es capaz de capturar diferentes aspectos del fenómeno de una ST. Como consecuencia de lo anterior, se toma como premisa que el modelo ARIMA-ARCH es capaz de capturar de manera más adecuada la dinámica lineal. Por otro lado, se presupone que un modelo de RNA se puede emplear para modelar, comprender y asimilar el comportamiento no lineal del mercado bursátil, más aún al usar la DEM.

### 3. Descomposición empírica en modos

Existen diversas técnicas de procesamiento de señales que parten del principio de descomponer una señal original, con el objetivo de aminorar la complejidad del fenómeno, es decir, separarla en componentes más sencillos para trabajar de manera aislada cada componente. El método más popular es el análisis de Fourier. Sin embargo, la transformada de Fourier sufre de restricciones, ya que el sistema debe ser lineal y los datos deben ser estrictamente periódicos o estacionarios, de lo contrario, el espectro de Fourier resultante tendría poco sentido físico (Huang, Shen, Long, 1999). Por el contrario, el método de DEM propuesto en un inicio por Huang et al. (1998), no se encuentra sujeto a tales restricciones, siendo capaz de tratar señales no lineales. La DEM es un método que descompone señales de forma adaptiva, en un conjunto finito de componentes oscilatorios denominados FMIs, que incorporan solamente componentes de frecuencias semejantes. Así, este método consiste en un proceso de eliminación de ruido de una señal de modo iterativo, para adquirir los componentes oscilantes, dicho procedimiento es denominado algoritmo de cernido. Su aplicación se ha expandido a diferentes ámbitos como las ciencias marítimas (Huang, Shen, Long, 1999), diagnóstico de averías (Liu, Riemenschneider, Xu, 2006), procesamiento de imágenes (Bi, Sun, Huang, Yang, Huang, 2007), reconocimiento de patrones (Yanga, Yang, Qic, Suene, 2006) y predicción del consumo eléctrico (León Anaya et al., 2013).

Una FMI es una función que debe cumplir dos condiciones: primera, en todo el conjunto de datos, el número de extremos y el número de cruces por cero debe ser igual o como máximo diferenciarse por uno; y segunda, el valor medio de la envolvente definida por los máximos y mínimos locales debe ser cero o un valor muy cercano a cero en todos los puntos. Una FMI, al cumplir estas condiciones asegura que se encuentra implicada en

solamente un modo oscilatorio, esto da como resultado que no se admitan ondas complejas incrustadas en la señal. El proceso de cernido se le aplica a la señal con el fin de adquirir las FMIs de forma iterativa a través de los siguientes pasos:

1. Identificar el total de extremos máximos y mínimos locales de la señal de entrada discreta.
2. Formar la envolvente superior  $e_{sup}(t)$  y la envolvente inferior  $e_{inf}(t)$  definidos por los puntos máximos y mínimos locales respectivamente, mediante una interpolación polinomial de orden 4.
3. Calcular punto por punto el valor promedio  $m(t)$  a partir de los envolventes superior e inferior.

$$m_1(t) = \frac{e_{sup}(t) + e_{inf}(t)}{2} \quad (1)$$

4. Realizar la resta entre la señal de entrada  $x(t)$  y el valor promedio  $m(t)$  para obtener el primer componente  $h_1$

$$h_1(t) = x(t) - m(t) \quad (2)$$

5. El componente  $h_1(t)$  es examinado para verificar si cumple con los requisitos para ser una FMI. Si dicho componente no logra ser una FMI, entonces el procedimiento de cernido se repite  $k$ -veces de los pasos del 1 al 4, y el componente toma el lugar de  $x(t)$  como la señal de entrada.

$$h_{11}(t) = h_1(t) - m_{11}(t) \quad (3)$$

6. El proceso de cernido se itera  $k$ -veces hasta que  $h_{1k}$  alcance los requisitos para ser una FMI.

$$h_{1k}(t) = h_{1(k-1)} - m_{1k} \quad (4)$$

El primer componente que alcanza la cualidad de ser una FMI es llamado  $c_1(t)$ . A partir de los datos, se logra obtener la primer componente FMI, la cual contiene la amplitud más fina, es decir, el componente con el periodo más corto de la señal, denotado como sigue:

$$c_1(t) = h_{jk} \quad (5)$$

Los criterios de paro en el proceso de cernido determinan el momento en el cual el proceso ha sido completado y se ha adquirido una nueva FMI, es decir, estos criterios marcan la cantidad de iteraciones necesarias para obtener una FMI. El criterio de paro se rige bajo las siguientes tres condiciones:

1. En cada punto: amplitud media  $<(\text{umbral}2 \text{ amplitud del envolvente})$ .
2. Media de la matriz booleana:  $((\text{amplitud media} / \text{amplitud de la envolvente}) > \text{umbral}) < \text{tolerancia}$ .
3. El número de cruces por cero y el número de extremos es menor o igual a uno.

Los parámetros umbral, umbral2 y tolerancia deben ser predeterminados antes de comenzar el proceso de cernido y los criterios más populares son los propuestos por Huang et al. (1988). Es mediante la manipulación de dichos criterios de paro en el proceso de cernido, que es posible adquirir una cantidad mayor o menor de FMIs provenientes de la señal de entrada original. Dicho criterio calcula la similitud entre dos componentes, prototipos de FMI, resultantes de iteraciones sucesivas, analizando el parecido o semejanza entre los dos componentes, mediante la medición de la desviación que hay entre ambas por medio de la desviación estándar SDj entre dos prototipos de FMI sucesivos  $h_{j-1}(t)$  y  $h_j(t)$ :

$$SD_j = \sum_{t=0}^T \left\{ \frac{(h_{1(j-1)}(t) - h_{1j}(t))^2}{h_{1(j-1)}^2(t)} \right\} < U \quad (6)$$

En esta investigación, se descompone una ST mediante el software matemático Matlab y la función propuesta por Rilling, Flandrin, Gonçalves (2003). Los parámetros de paro por defecto son: [0.05, 0.5, 0.05]-[umbral, umbral2, tolerancia]. Se descompusieron un total de 1554 datos y se obtuvieron 7 FMIs y un residuo. Sin embargo, cuando se manipularon los parámetros de paro para obtener un mayor número de FMIs, se disminuyó el error de las predicciones estimadas.

Después de una exhaustiva búsqueda, se determinaron los nuevos parámetros de paro en: [0.0031, 0.013, 0.0031] y se obtuvieron un total de 11 FMIs y un residuo. Para determinar cuáles eran los criterios de paro idóneos, se partió como primer premisa que los parámetros tenían que ser menores a los establecidos por defecto. De esta forma, se adquiriría una cantidad mayor de FMIs. Los datos se organizaron en 2 subconjuntos. El primero contenía los 35 datos originales, en el segundo se insertaron 35 datos a pronosticar. Se comenzó con unos parámetros muy pequeños, que se aumentaron paulatinamente. La cantidad de FMIs idóneas se obtuvo hasta obtener la misma cantidad de FMIs, en ambos subconjuntos. De esta forma, al mantener la misma cantidad de FMIs en las dos submuestras, la expansión de los datos mediante la descomposición no crea una notable disrupción en la estructura de las FMIs adquiridas. Una vez que se ha obtenido el componente  $c_1$  se separa del resto de los datos:

$$r_1 = x(t) - c_1 \quad (7)$$

Es muy posible que el componente  $r_1$  aun contenga información que se puede utilizar, por ello, el proceso de cernido debe de volver a realizarse, y el componente  $r_1$  pasa a ser el dato de entrada con el fin de adquirir un nuevo componente residual  $r_n$ . El método es iterado n-veces,  $r_2 = r_1 - c_2, \dots, r_n = r_{n-1} - c_n$ , hasta que  $r_n(t)$  sea más pequeño que un valor fijado previamente, o la señal se transforme en una señal monótona decreciente de la cual ya no sea posible extraer información para obtener una FMI. La suma que se obtiene al finalizar la experimentación con la DEM y mediante la cual se puede volver a recuperar la señal de entrada es como sigue:

$$x(t) = \sum_{j=1}^j IMF_j(t) + r_n(t) \quad (8)$$

Donde  $j$  es la suma total de FMIs y  $r_n(t)$  es la función monótona que marca el fin del procedimiento. Diversas causas afectan el número de FMIs que se pueden obtener al concluir el proceso de cernido: la condición de interrupción para obtener una FMI, el tamaño de la muestra de datos que se descompone, el grado de complejidad de la señal, entre otros.

#### 4. Variantes ARIMA-ARCH

En la década de los 70's del siglo pasado, Box Jenkins plantearon y desarrollaron una metodología que tenía como finalidad identificar, estimar y diagnosticar modelos dinámicos de ST, donde el factor tiempo asume el rol de una variable preponderante. Dicha metodología, tiene como premisa permitir que sean los propios datos de la variable los que indiquen las características del modelo probabilístico adecuado al fenómeno analizado. De esta forma, el modelador queda exento en muchos de los casos de la exhaustiva y compleja tarea de especificar el modelo de forma tradicional, y que en ocasiones, resultaba en una infructuosa especificación del modelo, ya que las variables explicativas elegidas en base a la teoría subyacente del fenómeno no resultaban ser las idóneas.

Los modelos ARIMA basan la explicación de una variable mediante su pasado. Esta manera de abordar la explicación de un fenómeno en cuestión ofrece la desventaja de sólo realizar el análisis del fenómeno en cuestión con una sola variable, cuando en algunos casos es posible realizar un análisis más enriquecedor con una colección de variables que pudieran explicar la evolución de manera más fehaciente de la variable analizada. Sin embargo, el análisis univariante ofrece algunas ventajas que son: evitar la forma tradicional de identificación y especificación del modelo, puesto que en muchas ocasiones la identificación de las variables explicativas es poco clara y las relaciones entre las variables no son constantes en el tiempo; y sólo son necesarios los datos sobre la variable en estudio. En esta investigación, se utilizaron cuatro variantes del modelo, las cuales son las siguientes.

1. La primera de ellas, es la más general y con un uso más amplio, el modelo ARIMA puro, que basa su predicción en una combinación lineal de sus valores pasados y una conjunción de errores o innovaciones (Box Jenkins, 1990).
2. La segunda, es el modelo SARIMA que se utiliza cuando el proceso analizado posee un comportamiento estacional (Makridakis, Wheelwright, Hyndman, 1988).
3. A principios de la década de los ochenta surge una modelización que trata de resolver la volatilidad de los mercados financieros, de esta forma surgen los modelos de volatilidad condicional heterocedástica que conceptualmente son idóneos para tratar la inestabilidad de la varianza en el transcurso del tiempo. En el año de 1982 Robert Engle (1982) introduce los modelos de volatilidad condicional (ARCH). En esencia, el modelado del proceso toma en cuenta los cambios significativos en la varianza. El conjunto o familia de los modelos ARCH tratan de captar los rasgos más representativos de las STs financieras que se encuentran con mayor reincidencia, entre dichos rasgos sobresalen el agrupamiento de la volatilidad, distribución leptocúrtica y el efecto de apalancamiento (Brooks, 2014). Lo anterior da la tercera variante en la combinación ARIMA-ARCH.
4. Finalmente, la cuarta variante es una combinación con la versión generalizada del modelo ARCH (GARCH), para emplear el llamado ARIMA-GARCH.

#### 5. Algoritmo FS-EPNet

La aplicación de las RNAs en diferentes ámbitos se ha incrementado de manera notable en las últimas décadas, y el ámbito financiero no es la excepción. Como se había mencionado, las RNAs han demostrado ampliamente su efectividad como técnicas en la resolución de problemas, no obstante, la edificación efectiva de una RNA no es una tarea trivial, por ejemplo en el MLP, es indispensable determinar a priori el número de capas y de neuronas

por capa, la tasa de aprendizaje, la inicialización de los pesos y del umbral, entre otros parámetros como las funciones de activación.

En los últimos años, se ha incrementado significativamente la investigación en esta área, logrando destacados avances tanto teóricos como empíricos. El problema de diseñar la arquitectura de una RNA cercana al óptimo para una aplicación sigue sin resolverse (Yao Liu, 1997), puesto que su diseño suele realizarse a partir de la metodología de prueba y error, donde es el experto quien define con base en su experiencia y criterio los parámetros que componen una RNA, es decir, el diseño a mano de RNAs (HDNN, por sus siglas en inglés Hand Design Neural Networks). Se puede concluir, que los valores óptimos de los parámetros de la RNA que permitan lograr su mejor calidad no se encuentran garantizados (Ortiz-Rodriguez et al., 2013) al diseñarlas usando el HDNN. Por ello, se propone utilizar un algoritmo evolutivo (AE) que se encargue de diseñar los parámetros de la RNA por medio de la evolución, con el fin de lograr un diseño topológico más eficiente de la red.

Los AEs retoman las ideas de la teoría Darwiniana de la evolución natural, basando sus fundamentos en los principios que determinan la evolución biológica de las especies. Estos algoritmos se consideran métodos de resolución de problemas que son capaces de alcanzar cualquier zona del espacio de búsqueda, ya que su exploración es una mezcla entre una búsqueda dirigida y una aleatoria. La evolución de RNAs o también denominada como Neuroevolución (NE) es un método por el cual se adaptan las RNAs, donde el AE es el instrumento que tiene como función minimizar o maximizar la función objetivo, mediante la búsqueda de los parámetros óptimos de la red, esta metodología ha demostrado su efectividad en diversas investigaciones (Yao, 1999; Bullinaria, 2007; Landassuri-Moreno Bullinaria, 2009; León-Anaya et al., 2013).

El algoritmo FS-EPNet (Landassuri-Moreno Bullinaria, 2009) es una ampliación del algoritmo EPNet introducido inicialmente por Yao Liu (1997), la extensión realizada al algoritmo EPNet se enfoca en adaptar la dimensión del espacio de entrada durante la evolución al mismo tiempo que se adaptan los demás parámetros de la red. De esta forma, la inclusión del proceso de selección de características al algoritmo EPNet, pretende eliminar información innecesaria o redundante, para conservar solamente las características o la información más trascendental. Así, los beneficios se pueden ver en términos de: ayudar a mejorar el aprendizaje automático en términos de precisión predictiva, comprensibilidad, eficiencia en el aprendizaje, modelos compactos, y una eficiente recopilación de datos (Sammur, 2011).

Por otro lado, si se desestima la inclusión del proceso de selección de características al modelo neuronal, es factible que se tenga un exceso de entradas que dé como resultado una RNA de un tamaño mayor al necesario, esto se traduce en un mayor consumo de tiempo en la fase de aprendizaje, aunado a ello se inserta mayor ruido al sistema, que tiene como consecuencia una disminución en la exactitud de la predicción (Landassuri-Moreno Bullinaria, 2009), y se puede perder la generalización de la red, la cual es la capacidad de responder correctamente a patrones no vistos durante la etapa de aprendizaje. Cabe destacar que el modelo utilizado por el algoritmo FS-EPNet es el MLP generalizado (Werbos, 1994), este modelo se configura mediante una capa de entrada, una capa de salida y un conjunto de neuronas ocultas que se pueden interconectar entre ellas, así como también se admiten conexiones entre cualquier neurona anterior a una delantera. De esta forma, se pueden modelar RNAs con neuronas de entrada conectadas a neuronas de salida, además de que es factible conectar neuronas de salida entre sí. Dada la flexibilidad que se le proporciona a una RNA del tipo MLP generalizada (GMLP, por sus siglas en Inglés Generalized MLP), se pueden crear redes muy complejas pero a la vez muy compactas

que contengan un número menor de neuronas, a diferencia de las basadas en un típico MLP con una o varias capas ocultas.

El algoritmo FS-EPNet podría parecer un procedimiento simple, sin embargo, cuenta con diversas características como se mencionan a continuación:

- a. El algoritmo FS-EPNet se considera un procedimiento híbrido que es empleado para evolucionar RNAs de propagación hacia delante (feedforward), ya que mezcla la búsqueda estructural evolutiva con un método de descenso por gradiente del tipo BP.
- b. Se cimienta en el algoritmo para evolucionar arquitecturas de RNAs, donde solamente son empleados los operadores de selección y mutación, siendo este último el único operador por medio del cual se recorre el espacio de búsqueda de las topologías neuronales.
- c. Al descartar la utilización del operador de cruzamiento en el proceso evolutivo, se solventa el inconveniente conocido como el problema de las convenciones competitivas (Hancock, 1992; Schaffer, Whitley, Eshelman, 1992) o el problema de la permutación (Hancock, 1992; Radcliffe, 1993).
- d. El algoritmo alterna entre los cambios que se producen en la confección estructural de la RNA y la transformación de los pesos mediante el entrenamiento, de esta manera, se pretende no romper los vínculos entre los padres y sus descendientes. Así mismo, se logra una búsqueda eficiente dado que el algoritmo FS-EPNet se construyó bajo las mismas bases en las cuales fue diseñado e implementado EPNet, y con el fin de no agregarle mayor complejidad al modelo, la evolución de las entradas opera de igual forma a como se realiza en su algoritmo predecesor EPNet, es decir, el modelo trabaja de manera análoga sobre las entradas como sobre cualquier nodo de la RNA.
- e. El algoritmo FS-EPNet fomenta la edificación de una red neuronal evolutiva parsimoniosa, es decir, una RNA con la menor cantidad de parámetros, o bien, con el menor grado de complejidad de entre varias, pero a la vez con el mayor poder predictivo posible. Esto es realizado mediante el empleo de diversas mutaciones de forma sucesiva en cada una de las generaciones y solamente una de estas mutaciones será efectuada en cada generación.

El algoritmo FS-EPNet emplea 9 operadores de mutación distintos para transformar la conectividad y los pesos de la red, los cuales se dan a continuación:

1. El entrenamiento híbrido, donde se mezclan 2 métodos de entrenamiento; un algoritmo de retropropagación modificado (MBP, por sus siglas en inglés Modified Backpropagation) y el algoritmo de Recocido Simulado (SA, por sus siglas en inglés Simulated Annealing).
2. Eliminación de nodo oculto
3. Eliminación de conexión
4. Eliminación de entrada
5. Eliminación de retardo, se determina si las entradas de la red son consecutivas o espaciadas.
6. Adición de nodo.
7. Adición de conexión.

---

8. Adición de retardo

9. Por último, adición de entrada.

El orden secuencial de las mutaciones es fundamental ya que es uno de los objetivos al crear RNAs compactas. Consecuentemente, se pueden diseñar redes parsimonias y solamente se incrementan sus dimensiones si estas no son capaces de alcanzar un desempeño aceptable con la tarea asignada. Mediante esta forma progresiva de aplicar mutaciones, el algoritmo FS-EPNet introduce una novedosa y simple alternativa para manejar el compromiso existente entre la varianza y el sesgo (bias-variance trade-off), sin embargo, puesto que las RNAs son evolucionadas con una sección de los datos disponibles (incluyendo los de entrenamiento), es muy probable que la mayoría de los individuos (RNAs) de la población tengan un sesgo pequeño (Liu Yao, 2001). Por otro lado, la calidad de la adaptación de un individuo (RNA) en la población en cada generación es medido únicamente por una métrica de desempeño de la varianza, en este caso, se emplea la técnica de normalización de la raíz del error cuadrático medio (NRMSE, por sus siglas en inglés Normalized Root Mean Square Error) (Landassuri-Moreno Bullinaria, 2009). De esta manera, no se introduce mayor complejidad al modelo, y se da un balance entre el sesgo y la varianza. Los principales pasos del algoritmo FS-EPNet se detallan a continuación:

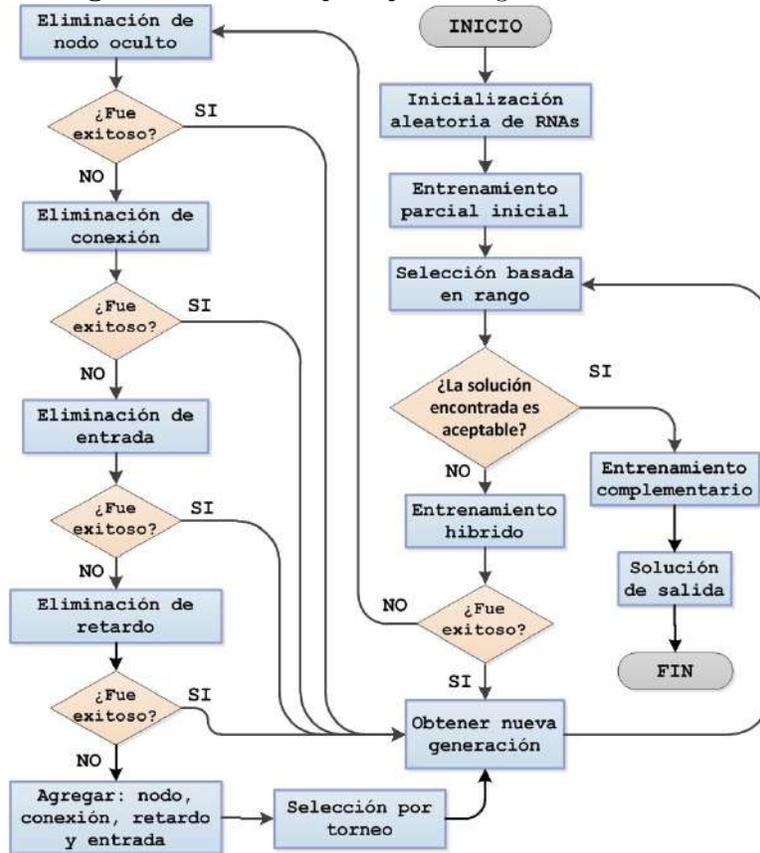
1. Generar una población inicial de redes aleatoriamente. El número de nodos y la densidad de las conexiones iniciales para cada red se generan al azar dentro de ciertos rangos.
2. Entrenar parcialmente cada red de la población por un cierto número de épocas usando el algoritmo de MBP con tasas de aprendizaje adaptativo.
3. Se clasifican las redes en la población de acuerdo a la evaluación del error cometido (valor de adaptabilidad), del mejor al peor.
4. Si se ha alcanzado el número máximo de generaciones determinado previamente o si la mejor red encontrada es aceptable (valores establecidos por el usuario), detener el proceso evolutivo. En caso contrario, continuar y elegir a una red de la población (usada como padre para generar descendientes a partir de ella).
5. Entrenar la red seleccionada mediante el entrenamiento híbrido para obtener un nuevo descendiente. Si el entrenamiento híbrido reduce notablemente el error, calificar al descendiente como exitoso y reemplazar al padre con su descendiente en la población actual. De lo contrario, descartar al descendiente y continuar.
6. Determinar la cantidad de nodos a eliminar de manera aleatoria entre 1 y 3, posteriormente, se eliminan los nodos ocultos de la red seleccionada de manera aleatoria. A continuación, se entrena parcialmente la red podada con el algoritmo MBP para obtener el nuevo descendiente. Si el descendiente es mejor que la peor red en la población actual, se sustituye a la peor red con el descendiente y se va al paso 3. De lo contrario, se descarta al descendiente y se continúa.
7. Calcular la importancia aproximada de cada conexión en la red padre utilizando un método no convergente (Finnoff, Hergert, Zimmermann, 1993). Determinar el número de conexiones que se desea eliminar de la misma manera que se ha descrito en el paso 6. Tomando como base el cálculo de importancia de las conexiones se procede a eliminar conexiones de manera aleatoria y entrenar parcialmente la red.

7. En caso de que el descendiente sea mejor que el peor individuo de la población actual, se sustituye al peor individuo con el nuevo descendiente y se va al paso 3. De lo contrario, se descarta a este descendiente y se continúa.
8. Determinar la cantidad de entradas a eliminar de manera aleatoria entre 1 y 2, posteriormente, se elige de forma aleatoria que entradas se eliminarán de la red padre. Se entrena parcialmente la red podada con el algoritmo de MBP para obtener el nuevo descendiente. Si el descendiente es mejor que la peor red en la población actual, se sustituye a la peor red con el descendiente y se va al paso 3. De lo contrario, se descarta al descendiente y se continúa.
9. El número de retardos a eliminar se determina de forma aleatoria, posterior a la modificación se ajustan los patrones de entrenamiento y se entrena parcialmente la red podada. En caso de que el descendiente sea mejor que el peor individuo de la población actual, se sustituye al peor individuo con el nuevo descendiente y se va al paso 3. De lo contrario, se descarta al descendiente y se continúa.
10. Determinar el número de conexiones, nodos ocultos, entradas y retardos, que se agregarán a la red y generar 4 descendientes a partir de ello. Posteriormente, entrenar parcialmente a los descendientes con el algoritmo MBP y someterlos a una selección por torneo, donde se elige al individuo mejor adaptado para pasar a la siguiente generación. Se reemplaza al peor individuo de la población actual con el nuevo descendiente elegido y se va al paso 3.
11. Una vez concluido el proceso evolutivo, se realiza un entrenamiento complementario al mejor individuo.

La figura 1 muestra el proceso evolutivo del algoritmo FS-EPNet. Uno de los rasgos distintivos del FS-EPNet al igual que su predecesor el EPNet es el empleo de la técnica de selección por estado estacionario (Steady-State), en el algoritmo de estado estacionario sólo se reemplazan unos cuantos individuos en cada generación, a diferencia de los algoritmos generacionales donde se sustituye a toda la población en cada iteración.

En la mayoría de los modelos neuroevolutivos, cuando se aplica una operación de mutación topológica aditiva a una capa, se agrega un nodo que se encuentra completamente conectado, y los pesos relacionados a estas conexiones se inicializan aleatoriamente. Esta forma de agregar nodos tiende a eliminar el comportamiento adquirido por el padre, dando como resultado una ruptura entre el vínculo funcional entre el padre y su descendiente. Por el contrario, en el algoritmo FS-EPNet se emplea una serie de procedimientos como son: un entrenamiento parcial posterior a una mutación arquitectónica y la división de un nodo existente. De esta forma, no se quebranta dicho vínculo entre ancestro-descendiente, simulando así, una evolución Lamarckiana (Yao Liu, 1997; Landassuri-Moreno Bullinaria, 2009), siendo diferente a evolucionar redes con una evolución Darwiniana, donde sólo se hereda la arquitectura y el valor de los pesos aprendidos es descartado.

**Figura 1.** Estructura principal del algoritmo FS-EPNet.



Fuente: Elaboración propia.

## 6. Configuración experimental

La parte experimental se encuentra dividida en 3 fases o etapas (A, B y C), donde es conveniente resaltar que en todas las etapas, los modelos predictivos emplearon la misma técnica de pronóstico, llamada predicción a múltiples pasos o iterada, es decir, en todos se utilizan valores predichos para extender el horizonte de predicción, esto a pesar de recibir diferentes nombres a lo largo de las publicaciones, como se puede notar a continuación: en (Landassuri-Moreno, Bustillo-Hernández, Carbajal-Hernández, Sánchez Fernández, 2013) se le llama predicción multipaso (MSP, por sus siglas en inglés Multi-Step Prediction) mientras que Brooks (2014) lo denominó como pronóstico de varios pasos adelante (MSF, por sus siglas en inglés Multi-Step-Ahead Forecasts). El horizonte de predicción en las 3 etapas es de 35 días.

Los modelos ARIMA-ARCH se construyeron por medio de la estrategia de Box-Jenkins, mediante el software econométrico Eviews. Se emplearon 3 variantes del modelo: ARIMA, SARIMA y la familia de los modelos ARCH. Para utilizar el modelo ARIMA-ARCH es necesario que los datos introducidos en el modelo sean estacionarios. La serie del IPC presenta un proceso no estacionario y como tal, algunas FMIs así como el residuo muestran signos de no estacionariedad. El orden de diferenciación de las STs para convertirlas en un proceso estacionario fue determinado por la prueba de raíz unitaria

(Brooks, 2014). Así, se inicia con la normalización de la ST en el intervalo  $[-1, 1]$ , rango usado por las funciones de transferencia de las RNAs. Posteriormente, se descompone la ST mediante la técnica de DEM, utilizando el software matemático Matlab y la función propuesta por Rilling, Flandrin, Gonçalves (2003). A través del procedimiento propuesto por Box-Jenkins (1990), se estimaron los parámetros ARIMA-ARCH donde los criterios utilizados para identificar el mejor modelo son: los criterios de información Akaike (AIC, por sus siglas en inglés Akaike Information Criterion) y el Bayesiano (BIC, por sus siglas en inglés Bayesian Information Criterion). La transformación de los datos originales a un proceso lineal en el modelo neuronal no es necesario que se realice, teniendo en cuenta esto, los datos se introdujeron sin filtro ni alteración y sólo aplicando la normalización de los datos anteriormente señalada. Para los experimentos del modelo no paramétrico se llevaron a cabo 12 ejecuciones independientes con fines estadísticos. Los datos provenientes del IPC se dividieron en 4 subconjuntos:

1. Conjunto de entrenamiento: se utiliza para llevar a cabo la tarea de aprendizaje con el algoritmo MBP o la técnica de SA.
2. Conjunto de validación: se emplea para evitar el sobre-ajuste en el proceso de aprendizaje;
3. Conjunto interior de prueba: se utiliza para simular una predicción MSP real (predicción fuera de la muestra) con el fin de medir la aptitud de las RNAs.
4. Conjunto final de prueba: utilizado una vez concluido el proceso evolutivo para evaluar el rendimiento de los diferentes individuos pertenecientes a la población final, para luego ser clasificados.

Para producir las predicciones del modelo DEM/FS-EPNet/RNA, se promedió a los cinco individuos mejor clasificados tomando como base el criterio de desempeño NRMSE en cada descomposición para ser parte de la reconstrucción. Es decir, de entre 12 ejecuciones independientes se eligieron los resultados provenientes de las cinco redes neuronales que obtuvieron el error menor en la predicción del conjunto final de prueba para cada FMI y el residuo. El resultado de la predicción de estas cinco RNAs elegidas a su vez se promedió para obtener una sola predicción por cada descomposición.

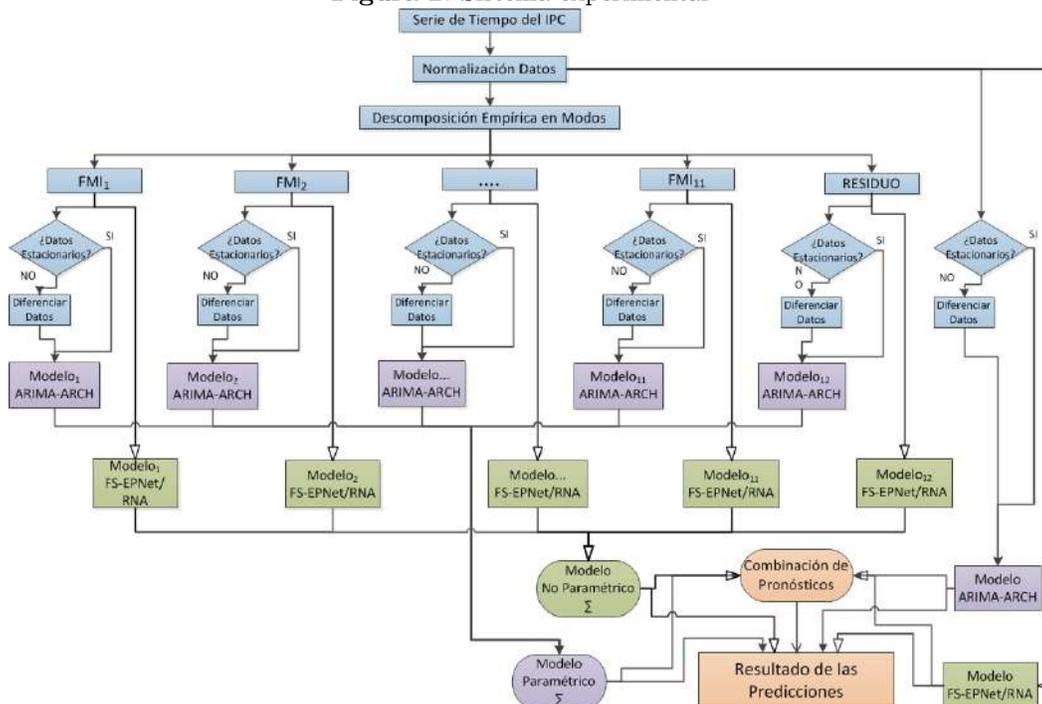
Para reconstruir una predicción final del modelo DEM/FS-EPNet/RNA, el promedio resultante de cada FMI y el residuo se sumó para adquirir una sola predicción. Por otro lado, los pronósticos generados mediante la fórmula FS-EPNet/RNA son el resultado de promediar los pronósticos generados mediante 12 ejecuciones independientes tomando como entrada la serie original del indicador analizado en cuestión. La configuración de los parámetros del algoritmo FS-EPNet se basó en la estrategia utilizada por Landassuri-Moreno (2012). Algunos parámetros se determinaron previamente antes de que comenzara el proceso evolutivo.

Los parámetros establecidos para esta investigación fueron los mismos para todos los modelos: tasa de aprendizaje 0.153, las mutaciones de los nodos ocultos y las conexiones entre 1 y 3, las mutaciones de entrada y de los retardos entre 1 y 2, el tamaño poblacional de 30 individuos, con 1,400 generaciones esperadas en el proceso evolutivo. No se emplearon criterios de interrupción temprana (early stopping) con el fin de evitar un sobre-entrenamiento de las RNAs, puesto que con base en experimentos previos, se infirió que no era necesario. Dada la complejidad de la tarea a resolver y también derivado de las pruebas preliminares, se estableció el número de épocas de entrenamiento parcial durante la evolución en 1,200 y de 3,600 al finalizar el proceso evolutivo.

La figura 2 ilustra la parte experimental de esta investigación, se utilizaron 4 sistemas, se diseñaron 2 modelos que trabajaron con la serie original del IPC, un modelo ARIMA-ARCH, así como también, un modelo no paramétrico de RNA diseñado mediante el algoritmo FS-EPNet. Posteriormente, los otros 2 sistemas tanto el paramétrico como el no paramétrico, que operaron con las FMIs y el residuo, edificaron para cada una de las descomposiciones un modelo ARIMA-ARCH. De igual forma, para cada FMI y su residuo se diseñó una RNA, donde cada una de las 11 FMIs y el residuo fueron pronosticados de manera individual en cada modelo (paramétrico y no paramétrico), finalmente se reconstruyó una predicción para cada modelo, mediante la suma de cada una de las descomposiciones pronosticadas, de manera que cada modelo produjo su propia predicción. De esta forma, se obtuvieron 4 pronósticos uno por cada sistema, los cuales fueron: 1) ARIMA-ARCH; 2) FS-EPNet/RNA; 3) DEM/ARIMA-ARCH; y 4) DEM/FS-EPNet/RNA.

Por último, se combinaron las predicciones de los modelos DEM/ARIMA-ARCH y DEM/FS-EPNet/RNA para crear un único pronóstico. El experimento se dividió en 3 etapas, la fase A es la parte que se puede considerar como el pronóstico dentro de la muestra, y las fases B y C representan al pronóstico fuera de la muestra. Esta etapa de investigación se encuentra implicada específicamente en la predicción de la bolsa mexicana de valores (BMV), en un periodo de crisis. La muestra de datos analizada es la correspondiente al precio de cierre del IPC para el período comprendido entre el 08/07/2002 y el 16/12/2008 tomada de ADVFN (2018); se contabilizó un total de 1624 datos del IPC.

Figura 2. Sistema experimental

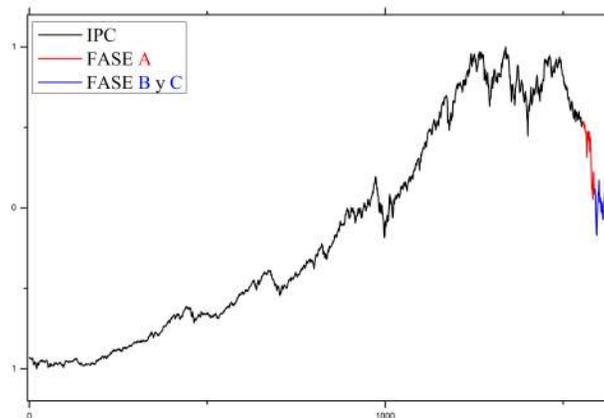


Fuente: Elaboración propia.

Para efectos de análisis estadístico, los criterios de confiabilidad en la predicción que son considerados son los siguientes: el error cuadrático medio (MSE, por sus siglas en inglés Mean Squared Error), la raíz del error cuadrático medio (RMSE, por sus siglas en inglés Root Mean Square Error) y la raíz del error cuadrático medio normalizado (NRMSE, por sus siglas en inglés Normalized Root Mean Square Error). Con lo cual, al comparar los valores de errores obtenidos en los pronósticos, el valor menor indica el mejor resultado en la predicción.

La figura 3 representa las diferentes fases de experimentación realizadas para el IPC.

**Figura 3.** Fases de experimentación del IPC.



Fuente: Elaboración propia.

Las tres fases de experimentación del IPC se describen a continuación:

### 6.1 Fase A del pronóstico del IPC

El objetivo es analizar el desempeño de los diferentes modelos propuestos en este estudio, con el fin de obtener los pesos de los modelos en la etapa de combinación de pronósticos (fase C). En esta primera fase, el modelo no paramétrico fue entrenado con un total de 1,554 datos, del 08/07/2002 al 28/08/2008, este conjunto de datos se encuentra representado en la figura 6 en color negro, dicha línea aparece en el subcuadro con la leyenda IPC. Mientras que para el modelo paramétrico se estimaron sus parámetros con un total de 350 datos, del 04/12/2007 al 28/08/2008. Se utilizó una cantidad menor de datos para el modelo paramétrico puesto que se obtuvo un mejor rendimiento en los resultados, además de que se facilitó la adquisición de los parámetros del modelo ARIMA-ARCH. El horizonte de predicción abarcó del 29/08/2008 al 17/10/2008, los cuales se encuentran ilustrados en la figura 6 en color rojo, en el subcuadro se indica bajo la leyenda FASE A. Por último en esta primera fase o etapa, se presenta una combinación de pronósticos con una igualdad en la ponderación.

Esta fase sirve para adquirir los pesos de los modelos de predicción para la fase C, donde se combinan los modelos paramétricos y no paramétricos. Puesto que el periodo de tiempo de mayor interés en esta investigación es cuando tiene ocurrencia la caída bursátil del indicador, esta fase vendría siendo una etapa donde se prueba el funcionamiento y se comprueban las diferentes técnicas de predicción, así como la técnica DEM. Es así que esta etapa se podría considerar como el periodo dentro de la muestra con la salvedad de que para realizar la predicción de esta etapa, a los modelos de predicción se les oculta el periodo de datos a pronosticar, ya que en esta investigación se estima que no resulta de utilidad la predicción dentro de la muestra.

## 6.2 Fase B del pronóstico del IPC

En esta segunda fase de la experimentación, el periodo de predicción transcurrió del 20/10/2008 al 08/08/2008; estos datos se representan en color azul en la figura 6. El día de la crisis bursátil fue el 22/10/2008, con una caída del 7.54%. En este periodo de experimentación, los datos fueron desplazados hacia adelante 35 días (puntos de pronóstico) a partir de los utilizados en la fase anterior, para estimar los parámetros de ambos modelos de predicción. Los parámetros del modelo de ARIMA-ARCH fueron estimados con los datos que abarcan el período comprendido entre el 01/06/2007 al 17/10/2008, un total de 350 datos. El periodo de la muestra empleado para el entrenamiento de las RNAs fue del 08/07/2002 hasta 17/10/2008, resultando en un total de 1,589 observaciones, en esta fase, al igual que en la tercera, los datos utilizados para las estimaciones están representados en la figura 6, a través de la adición de los datos de la serie representados en color negro y rojo.

## 6.3 Fase C del pronóstico del IPC

En la fase final, las predicciones resultantes de los modelos paramétricos y no paramétricos se utilizaron para realizar una predicción final combinada, con el objetivo de obtener una mejor predicción en los resultados. En esta fase se analizó el mismo período de tiempo que en la experimentación hecha en la fase B.

## 7. Discusión de resultados

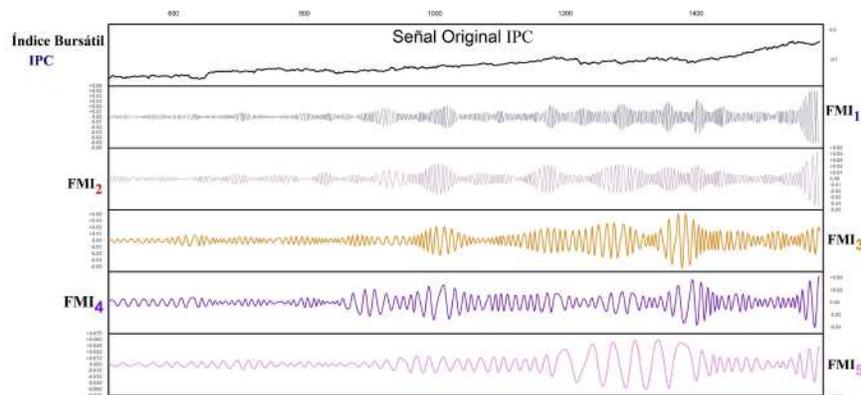
La figura 4, muestra la serie original del IPC, así como las 12 descomposiciones resultantes al aplicarle la técnica DEM, de la descomposición se obtuvieron 11 FMIs y un elemento residual. Se ilustra cómo la primera FMI se distingue por contener oscilaciones rápidas y conforme avanza el proceso de cernido, las FMIs que se adquieren incluyen oscilaciones gradualmente más lentas hasta que se alcanza el residuo que tiene como característica ser una función monótona, a esta serie ya no se le puede extraer más información.

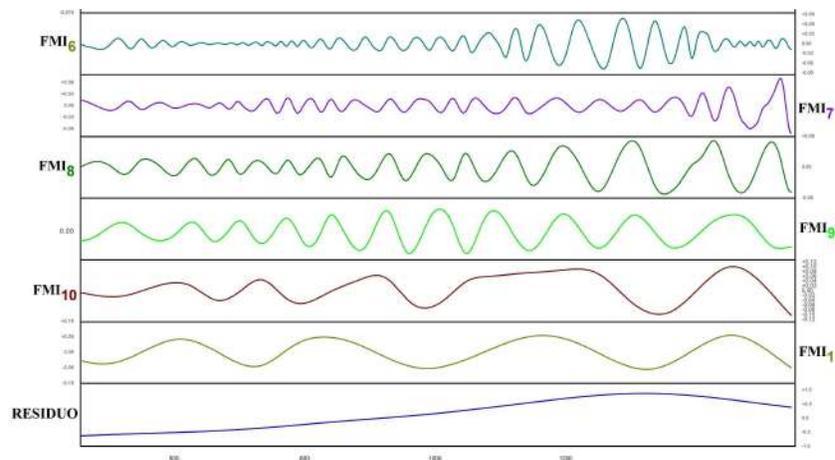
Los resultados de las fases de experimentación son mostrados como sigue:

### 7.1 Resultados de la fase A del pronóstico del IPC

En esta fase se comparan los 4 modelos anteriormente descritos. En la figura 5 se muestra el desempeño de los diferentes modelos propuestos para esta etapa. En tal figura, se ilustra como el modelo paramétrico ARIMA-GARCH(1,1) es el modelo con el peor desempeño, además cabe destacar como el modelo DEM/ARIMA-ARCH después de aproximadamente 22 días de predicción tiene un desempeño muy errático. Por otra parte, es visible que el modelo FS-EPNet/RNA es el de mejor ajuste a los 35 días de pronóstico.

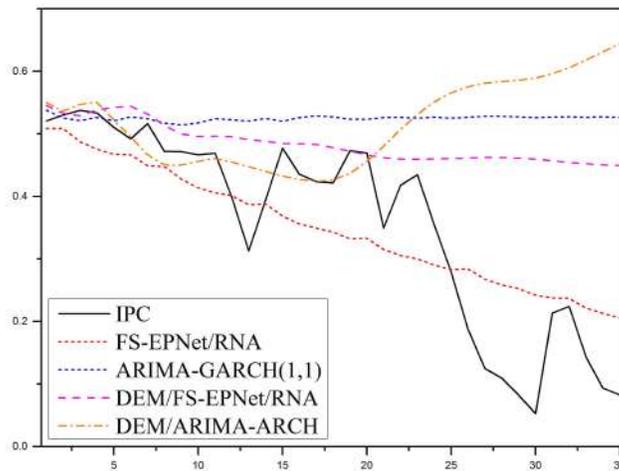
**Figura 4.** IPC al precio de cierre del 08/07/2002 al 13/03/2009, y sus correspondientes 11 FMIs y su residuo.





Fuente: elaboración propia.

Figura 5. fase A del IPC, del 29/08/2008 al 17/10/2008



Fuente: elaboración propia.

La tabla 1, indica que para los 10, 15 y 20 periodos, el modelo que mejor se ajusta a los datos es el DEM/ARIMA-ARCH y que para los 35 periodos, el modelo FS-EPNet/RNA tiene un error menor.

Tabla 1. Medición de errores en la fase A del pronóstico del IPC.

10 días de predicción					15 días de predicción				
Criterio	FS-EPNet/ RNA	ARIMA- GARCH(1,1)	DEM/FS- EPNet/RNA	DEM/ARIMA- ARCH	Criterio	FS-EPNet/ RNA	ARIMA- GARCH(1,1)	DEM/FS- EPNet/RNA	DEM/ARIMA- ARCH
MSE	0.00189	0.00085	0.00081	<b>0.0005</b>	MSE	0.00291	0.00605	0.00351	<b>0.00191</b>
RMSE	0.04344	0.02912	0.02838	<b>0.02245</b>	RMSE	0.05393	0.07778	0.06171	<b>0.04373</b>
NRMSE	0.28785	0.19294	0.18807	<b>0.14877</b>	NRMSE	0.39323	0.58376	0.47457	<b>0.3425</b>

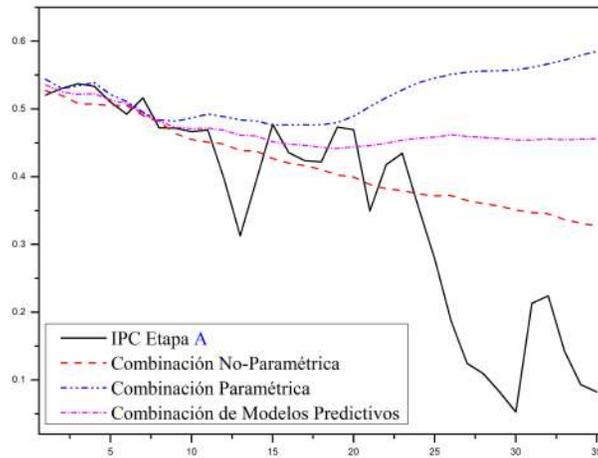
  

20 días de predicción					35 días de predicción				
Criterio	FS-EPNet/ RNA	ARIMA- GARCH(1,1)	DEM/FS- EPNet/RNA	DEM/ARIMA- ARCH	Criterio	FS-EPNet/ RNA	ARIMA- GARCH(1,1)	DEM/FS- EPNet/RNA	DEM/ARIMA- ARCH
MSE	0.00483	0.00622	0.00338	<b>0.00161</b>	MSE	0.00817	0.05321	0.03482	<b>0.06958</b>
RMSE	0.06952	0.07884	0.05816	<b>0.04007</b>	RMSE	0.09037	0.23068	0.1866	<b>0.26378</b>
NRMSE	0.5661	0.64203	0.47359	<b>0.32633</b>	NRMSE	0.57783	1.475	1.19319	<b>1.68665</b>

Fuente: elaboración propia

En esta etapa se propuso una combinación igualitaria de pesos de los modelos (promedio simple). La figura 6 muestra la combinación de 3 modelos predictivos. La primera mezcla es la combinación no-paramétrica que incluye los modelos FS-EPNet/RNA y DEM/FS-EPNet/RNA.

**Figura 6.** Combinación de pronósticos en la fase A del IPC, del 29/08/2008 al 17/10/2008.



Fuente: elaboración propia.

El segundo modelo combina al modelo paramétrico ARIMA-GARCH(1,1) y al modelo DEM/ARIMA-ARCH, el resultado de esta combinación produce una combinación paramétrica. La tercera y última combinación, llamada combinación de modelos predictivos, incluye la combinación no-paramétrica y la combinación paramétrica. La tabla 2, muestra el comportamiento de las predicciones en un esquema de combinación con pesos igualitarios. La misma tabla indica que para los 11 periodos, la predicción con un mejor desempeño es la combinación de modelos predictivos. Sin embargo, para los 35 periodos la combinación no-paramétrica es la predicción que posee el error más bajo en esta fase del IPC.

**Tabla 2.** Combinación de pronósticos y medición de errores en la fase A del IPC, del 29/08/2008 al 17/10/2008.

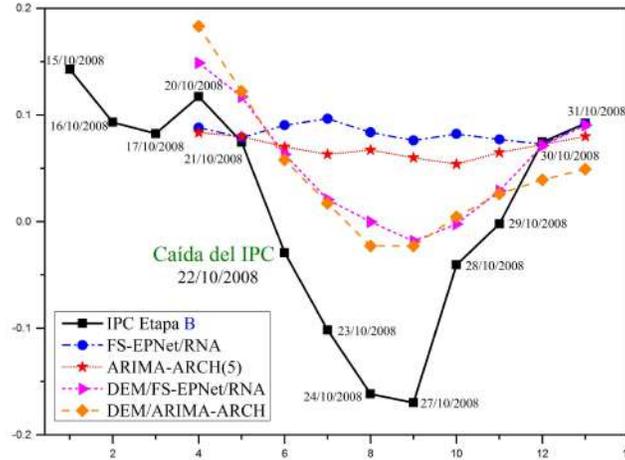
11 Periodos de Predicción				35 Periodos de Predicción			
Criterio	Combinación no-paramétrica	Combinación paramétrica	Combinación de modelos predictivos	Criterio	Combinación no-paramétrica	Combinación paramétrica	Combinación de modelos predictivos
MSE	0.00029	0.00022	<b>0.00016</b>	MSE	<b>0.01592</b>	0.06029	0.03371
RMSE	0.01688	0.01475	<b>0.0125</b>	RMSE	<b>0.12618</b>	0.24554	0.18359
NRMSE	0.11489	0.1065	<b>0.08083</b>	NRMSE	<b>0.80669</b>	1.57005	1.17391

Fuente: elaboración propia.

## 7.2 Resultados de la fase B del pronóstico del IPC

Aquí, la predicción comenzó del 20/10/2008, dos días antes de la caída bursátil acontecida el miércoles 22/10/2008, hasta el 08/12/2008, dando un total de 35 días de pronóstico. La figura 7 muestra en detalle el periodo de tiempo del 15/10/2008 al 31/10/2008, un intervalo de 13 días, donde se muestra el comportamiento del IPC, así como también se muestra el comportamiento de las diferentes predicciones de los modelos paramétricos y no paramétricos del 20/10/2008 al 31/10/2008, las predicciones abarcaron un período de 10 días.

**Figura 7.** Fase B del IPC datos del 15/10/2008 al 31/10/2008 con predicción del 20/10/2008 al 31/10/2008.



Fuente: elaboración propia.

La tabla 3 muestra el rendimiento de los modelos de predicción en diferentes puntos de predicción. A los 5, 10 y 15 puntos de pronóstico los modelos que utilizan la técnica DEM demuestran una amplia superioridad sobre los modelos que produjeron su predicción solamente con los datos del IPC originales. Sin embargo, a los 35 días de predicción el modelo ARIMA-ARCH(5) es el menos deficiente de entre todos los modelos propuestos en esta etapa. La tabla 4, muestra los parámetros utilizados para cada descomposición por los modelos predictivos. Esta tabla muestra que las descomposiciones FMI1, FMI2, FMI3, FMI4 y FMI7 se encuentran definidas y regidas por un proceso ARCH. Por otro lado, el modelo SARIMA se emplea en la predicción de la FMI8 y la FMI11, esto debido a que sus procesos tienen características de estacionalidad. La FMI6 es conducida por un modelo ARIMA debido a que no se le aplicó diferenciación, y en última instancia, las descomposiciones FMI9, FMI10 y el residuo se modelaron mediante el modelo ARIMA.

Por otra parte, la experimentación del modelo no paramétrico demuestra la alta flexibilidad que poseen las RNAs para manejar las cambiantes facetas y capturar los patrones subyacentes de las diferentes descomposiciones derivados de la ST del IPC. Además, el algoritmo FS-EPNet demuestra que puede producir redes muy compactas. El diseño para predecir la FMI3, con tan sólo 2 entradas, 6 neuronas en la capa oculta y 18 conexiones, como edificación topológica se muestra en la figura 7b.

**Tabla 3.** Medición de errores en la fase B del pronóstico del IPC.

5 Días de Predicción					10 Días de Predicción				
Criterio	FS-EPNet/ RNA	ARIMA- ARCH(5)	DEM/FS- EPNet/RNA	DEM/ARIMA- ARCH	Criterio	FS-EPNet/ RNA	ARIMA- ARCH(5)	DEM/FS- EPNet/RNA	DEM/ARIMA- ARCH
MSE	0.02293	0.01812	0.01048	<b>0.00952</b>	MSE	0.01965	0.01569	0.00778	<b>0.00751</b>
RMSE	0.15141	0.1346	0.10239	<b>0.0976</b>	RMSE	0.14018	0.12525	0.0882	<b>0.08668</b>
NRMSE	1.32332	1.17636	0.89485	<b>0.85296</b>	NRMSE	1.30183	1.16325	0.81915	<b>0.80497</b>

15 Días de Predicción					35 Días de Predicción				
Criterio	FS-EPNet/ RNA	ARIMA- ARCH(5)	DEM/FS- EPNet/RNA	DEM/ARIMA- ARCH	Criterio	FS-EPNet/ RNA	ARIMA- ARCH(5)	DEM/FS- EPNet/RNA	DEM/ARIMA- ARCH
MSE	0.01386	0.01125	0.00759	<b>0.0064</b>	MSE	0.00897	<b>0.00888</b>	0.01537	0.02273
RMSE	0.11775	0.10604	0.08714	<b>0.07997</b>	RMSE	0.09469	<b>0.09423</b>	0.12398	0.15076
NRMSE	1.18473	1.06696	0.87677	<b>0.80465</b>	NRMSE	1.23459	<b>1.2286</b>	1.61647	1.96566

Fuente: elaboración propia.

Otras investigaciones, como la mencionada anteriormente en la sección de revisión literaria en el trabajo realizado por Zorin Borisov (2002), se empleó con una arquitectura de RNA con: 59 neuronas de entrada y 35 neuronas ocultas. Dicha diferencia entre ambos diseños arquitectónicos, a pesar de que no resuelven la misma tarea estas dos RNAs, puede otorgar un atisbo de las dimensiones tan compactas que puede lograr diseñar el algoritmo FS-EPNet. Además, la figura 7 en cada recuadro ilustra por cuestiones de apreciación un fragmento de la serie a resolver, la topología de la RNA encargada de solucionar la tarea, así como también el correlograma de dicha serie y el modelo paramétrico que mejor se ajustó a los datos. Cabe mencionar, que en la figura 7a se muestra la serie original del IPC, así como su correspondiente correlograma, donde es posible apreciar la dificultad de encontrar patrones a esta serie mediante este gráfico de autocorrelaciones. Por el contrario, en la figura 7b, 7c y 7d, los diferentes correlogramas de las FMIs muestran una mayor facilidad en la estimación de los parámetros del modelo ARIMA.

**Tabla 4.** Parámetros de los modelos paramétrico y no-paramétrico en la fase B del pronóstico del IPC.

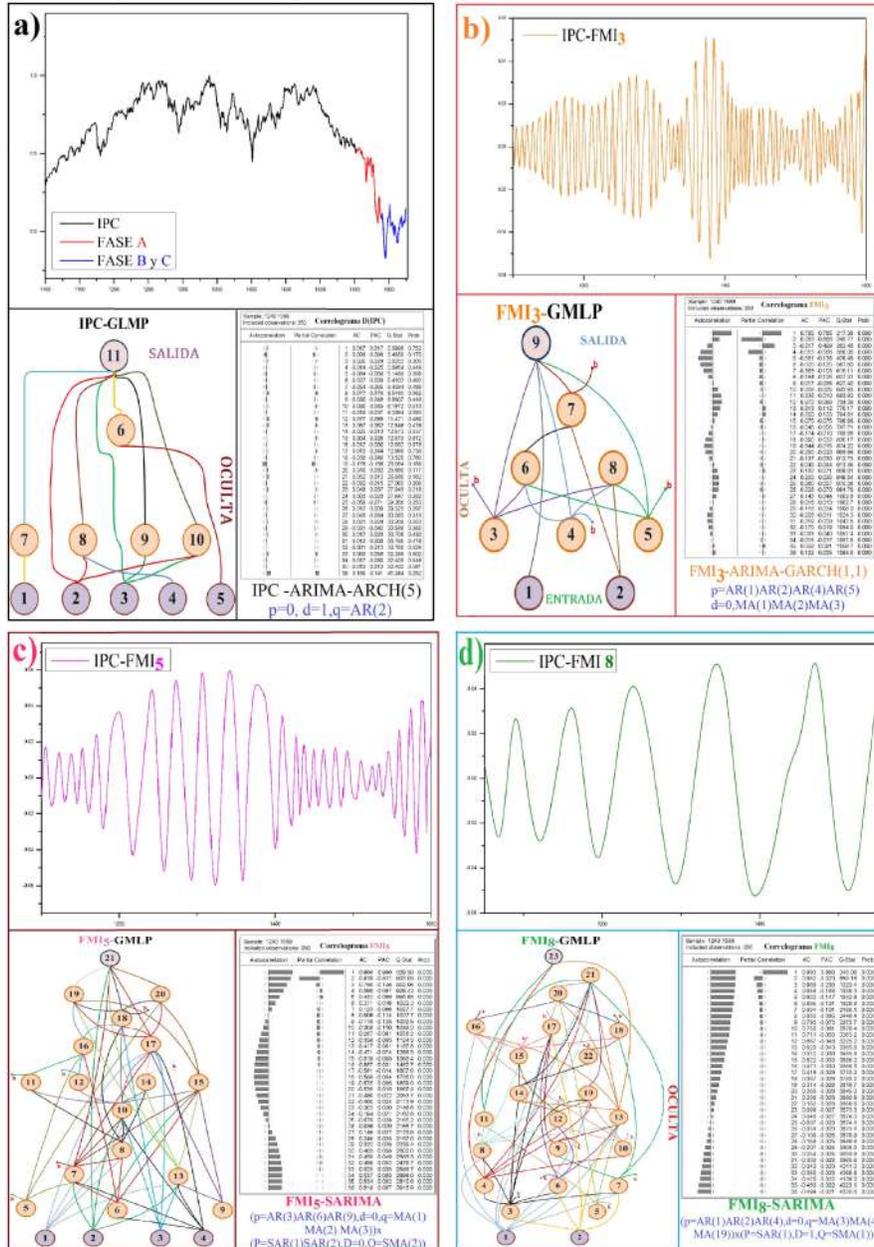
Modelo	Parámetros del modelo	Modelo paramétrico	Parámetros del modelo
FMI1			
FS-EPNet/RNA	Entradas:6; Ocultas:5; Conexiones:14; Retardos:6;	ARIMA-GARCH(1,1)	$p=AR(1)AR(2), d=0, q=0$
FMI2			
FS-EPNet/RNA	Entradas:3; Ocultas:8; Conexiones:24; Retardos:2;	ARIMA-ARCH(2)	$p=AR(6), d=0,$ $q=MA(1)MA(2)MA(3) MA(4)MA(5)MA(6)$
FMI3			
FS-EPNet/RNA	Entradas:2; Ocultas:6; Conexiones:18; Retardos:1;	ARIMA-GARCH(1,1)	$p=AR(1)AR(2)AR(4)AR(5),$ $d=0, q=MA(1)MA(2)MA(3)$
FMI4			
FS-EPNet/RNA	Entradas:3; Ocultas:5; Conexiones:16; Retardos:1;	ARIMA-GARCH(1,1)	$p=AR(1)AR(3)AR(4)AR(5)AR(6),$ $d=0, q=MA(1)MA(3)MA(4)$
FMI5			
FS-EPNet/RNA	Entradas:4; Ocultas:16; Conexiones:70; Retardos:7;	SARIMA	$(p=AR(3)AR(6)AR(9), d=0, q=MA(1)MA(2)$ $MA(3))x(P=SAR(1)SAR(2), D=0, Q=SMA(2))$
FMI6			
FS-EPNet/RNA	Entradas:2; Ocultas:12; Conexiones:48; Retardos:2;	ARIMA	$p=AR(1)AR(2)AR(3)AR(5)AR(7),$ $d=0, q=MA(1)MA(2)MA(7)MA(8)$
FMI7			
FS-EPNet/RNA	Entradas:2; Ocultas:15; Conexiones:59; Retardos:4;	ARIMA-ARCH(3)	$p=AR(1)AR(7),$ $d=0, q=MA(1)MA(2)MA(3)$
FMI8			
FS-EPNet/RNA	Entradas:2; Ocultas:20; Conexiones:78; Retardos:2;	SARIMA	$(p=AR(1)AR(2)AR(4), d=0, q=MA(3)MA(4)$ $MA(19))x(P=SAR(1), D=1, Q=SMA(1))$
FMI9			
FS-EPNet/RNA	Entradas:2; Ocultas:9; Conexiones:33; Retardos:4;	ARIMA	$p=AR(1)AR(2) AR(3), d=1,$ $q=MA(1)MA(2)$
FMI10			
FS-EPNet/RNA	Entradas:3; Ocultas:8; Conexiones:24; Retardos:4;	ARIMA	$p=AR(1)AR(2) AR(3)AR(5)$ $AR(7)AR(8), d=1, q=MA(1)MA(4)$
FMI11			
FS-EPNet/RNA	Entradas:4; Ocultas:17; Conexiones:80; Retardos:5;	SARIMA	$(p=AR(1)AR(3)AR(7), d=1, q=MA(1) MA(2)$ $MA(3))x(P=SAR(1), D=0, Q=0)$
Residuo			
FS-EPNet/RNA	Entradas:6; Ocultas:14; Conexiones:60; Retardos:5;	ARIMA	$p=AR(1)AR(4), d=2,$ $q=MA(1)MA(2)MA(3)MA(4)$

Fuente: elaboración propia.

Por último, la figura 8 muestra la evolución del IPC y el comportamiento de las predicciones de los modelos paramétricos y no paramétricos que produjeron su predicción solamente con la serie del IPC, así como también los modelos que trabajaron de manera conjunta con la técnica de DEM, para un periodo de 35 puntos o días de pronóstico. Dado

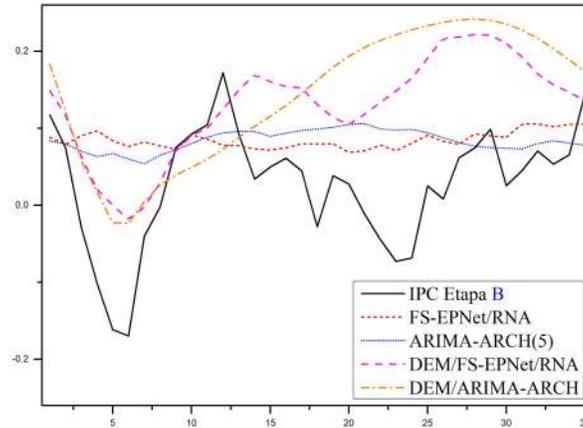
que la caída del IPC se produjo el 22/10/2008, la predicción comienza 2 días y continúa hasta 32 días posteriores, dando así los 35 días pronosticados.

**Figura 8.** a) diseño topológico del GMLP mediante FS-EPNet para predecir el IPC, b) diseño topológico del GMLP mediante FS-EPNet para pronosticar la FMI3 del IPC, c) topología del GMLP diseñado mediante FS-EPNet para pronosticar la FMI5 del IPC, y d) diseño topológico del GMLP mediante FS-EPNet para predecir la FMI8 del IPC.



Fuente: elaboración propia.

**Figura 9.** fase B del pronóstico del IPC, predicción del 20/10/2008 al 08/12/2008.



Fuente: elaboración propia.

### 7.3 Resultados de la fase C del pronóstico del IPC

Los métodos de combinación propuestos y utilizados en esta fase o etapa son: la combinación ponderada (CP) y el promedio simple (PS). La CP se basa en la estrategia propuesta por Granger Ramanatha (1984), donde los modelos utilizan un sistema de ponderación de combinaciones con pesos muy cercanos a la igualdad y se encuentra restringida la suma de pesos de los modelos a combinar a la unidad. Por lo tanto, la medida que se eligió fue el periodo de 11 pasos hacia adelante del criterio NRMSE, ya que en dicho periodo el error de ambos modelos es muy cercano uno del otro.

Las predicciones que se incluyen en la CP son los derivados de los modelos paramétricos y no paramétricos que trabajaron en conjunto con la técnica de DEM. Para llevar a cabo la combinación, los pesos se determinaron con base en el rendimiento de los modelos de la etapa A. Para realizar la combinación ponderada, cada una de las predicciones de la fórmula DEM/ARIMA-ARCH se multiplicó por 0.56766 y para la fórmula DEM/FS-EPNet/RNA cada pronóstico se multiplicó por 0.43234. Posteriormente, ambas predicciones ponderadas son adicionadas y se obtiene como resultado la CP. Cabe recalcar que la suma de los coeficientes  $PMP + PMNP = 1$  dan como resultado la unidad.

Adicionalmente, se propone una combinación de predicciones con pesos igualitarios (PS). Por un lado, se realizó la combinación no-paramétrica que incluye a los modelos FS-EPNet/RNA y DEM/FS-EPNet/RNA. Por otro lado, la combinación paramétrica es una mezcla de los modelos ARIMA-ARCH(5) y DEM/ARIMA-ARCH. La fórmula que se utilizó para determinar la ponderación de pesos se muestra en la tabla 5.

La figura 9 ilustra el comportamiento de los modelos: combinación no-paramétrica, combinación paramétrica y combinación ponderada, así como también la ST del IPC, en el mismo periodo experimental llevado a cabo anteriormente en la fase B del pronóstico del IPC. A partir de una inspección visual de la figura 9, se puede apreciar que la predicción producida por la CP por aproximadamente 12 días es el modelo que se ajusta mejor al periodo de la caída bursátil. La tabla 6 confirma lo anteriormente ilustrado en la figura 9, ya que muestra el rendimiento de los modelos predictivos para 5, 15 y 35 periodos de predicción. También, la misma tabla expone cómo el modelo que mejor se ajusta a los datos en los primeros periodos es el modelo de CP. Sin embargo, para los 35 puntos de predicción, el error acumulado por la CP es el mayor de entre los modelos expuestos en la fase C del pronóstico del IPC.

## 8. Conclusiones

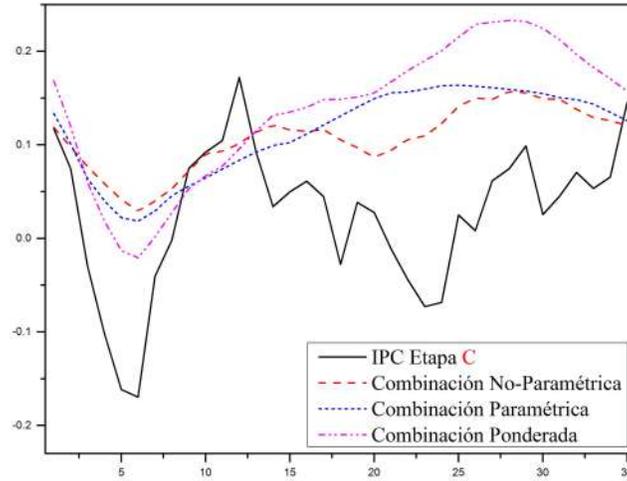
En décadas recientes, se han propuesto una gran cantidad de herramientas y técnicas que tratan de predecir el comportamiento del mercado bursátil. Entre los modelos con mayor popularidad se encuentran los modelos ARIMA, sin embargo, estos sólo son capaces de captar relaciones y características lineales en los datos bursátiles. Dadas las carencias de los modelos lineales, es que surgió la necesidad de abordar el análisis del mercado bursátil con herramientas no paramétricas y no lineales. Los resultados de esta investigación así lo confirman. Sin embargo, si los modelos paramétricos combinan su predicción con otras técnicas, en el caso de esta investigación la DEM, los resultados pueden ser mejorados notablemente. De igual forma, el resultado en las predicciones de los modelos no paramétricos puede ser enriquecido con la colaboración de una herramienta que descompone y elimina el ruido en las series. Tomando en consideración lo anteriormente expuesto y los resultados obtenidos, es posible inferir que la aplicación de la técnica DEM al pronóstico de la ST del IPC puede resultar beneficiosa, ya que aquellos patrones que se encontraban ocultos debido al ruido intrínseco de la ST, DEM los hace emerger y le facilita la predicción tanto al modelo paramétrico como al no paramétrico.

**Tabla 5.** Formulación de la combinación ponderada.

EGNRMSE = error global NRMSE MPNRMSE = modelo paramétrico–NRMSE MNPNRMSE = modelo no-paramétrico–NRMSE EGNRMSE = MPNRMSE + MNPNRMSE
<b>Para determinar el peso se tomó como tasa de valoración el criterio NRMSE a 11 días de la fase A.</b>
MPNRMSE = DEM/ARIMA-ARCH = 0.14576 MNPNRMSE = DEM/FS-EPNet/RNA = 0.19139 Dando como resultado el error global: EGNRMSE = 0.14576 + 0.19139 = 0.33716
<b>PMP = peso del modelo paramétrico</b> El peso del modelo paramétrico se determinó: $PMP = 1 - \frac{0.14576}{0.33716}$ PMP = 0.56766
<b>PMNP = peso del modelo no-paramétrico</b> El peso del modelo no-paramétrico se determinó: $PMNP = 1 - \frac{0.19139}{0.33716}$ PMNP = 0.43234
Yt = predicción del modelo paramétrico Xt = predicción del modelo no paramétrico CP = Combinación ponderada CP = PMP(Yt ) + PMNP(Xt ) CP = PMP(Predicciones-DEM/ARIMA-ARCH) + PMNP(Predicciones-DEM/FS-EPNet/RNA)

Fuente: elaboración propia.

**Figura 10.** fase C del pronóstico del IPC para una combinación de sus predicciones.



Fuente: elaboración propia.

En este estudio, se contrastaron 2 herramientas predictivas como modelos, el primero, el modelo paramétrico ARIMA-ARCH de carácter lineal. El segundo de estos fue diseñado mediante el AE FS-EPNet, el modelo no paramétrico de RNAs con cualidades no lineales. Ambos modelos se compararon en 2 intervalos de tiempo, en la primera etapa de predicción los modelos DEM/FS-EPNet/RNA, DEM/ARIMA-ARCH y la combinación de estos 2 últimos fueron los que mejor se ajustaron a la evolución del indicador en los primeros días de predicción.

**Tabla 6.** combinación de pronósticos en la fase C del pronóstico del IPC. Medición de errores: predicción del IPC para 5, 15 y 35 días.

Criterio	5 Periodos de Predicción			15 Periodos de Predicción				35 Periodos de Predicción			
	Combinación no paramétrica	Combinación paramétrica	Combinación ponderada	Criterio	Combinación no paramétrica	Combinación paramétrica	Combinación ponderada	Criterio	Combinación no paramétrica	Combinación paramétrica	Combinación ponderada
MSE	0.01577	0.01272	<b>0.00985</b>	MSE	0.0097	0.0082	<b>0.0066</b>	MSE	<b>0.01067</b>	0.01338	0.01906
RMSE	0.12359	0.11276	<b>0.09926</b>	RMSE	0.09848	0.09057	<b>0.08124</b>	RMSE	<b>0.10328</b>	0.11568	0.13806
NRMSE	1.09764	0.98551	<b>0.86751</b>	NRMSE	0.99087	0.91132	<b>0.81741</b>	NRMSE	<b>1.34661</b>	1.5083	1.80013

Fuente: elaboración propia.

Cabe señalar que todos los modelos de predicción pierden su poder predictivo en un periodo por más de 20 días aproximadamente en esta primera etapa de experimentación. En el segundo periodo de predicción, se intentó predecir una caída del bursátil acaecida el 22/10/2008, donde el IPC tuvo una baja del 7.01%. En este periodo, los modelos que basaron su predicción en las FMIs y el residuo resultantes de la descomposición del IPC fueron las herramientas capaces de anticipar la caída bursátil. Estos resultados pueden llevar a concluir que los modelos que trabajaron con DEM son capaces de capturar de manera un poco más precisa el movimiento cíclico del mercado bursátil.

Para obtener una predicción más acertada de una crisis bursátil, parece no ser suficiente el empleo de sólo una variable endógena, en este caso el pasado de la variable. Por ello, se propone incluir otras variables, que pudieran enriquecer y fortalecer el estudio de los mercados financieros. Es factible concluir, que el estudio de la crisis financiera, así como la configuración de los modelos predictivos deben tomar en cuenta las repercusiones que ocasionan las malas noticias en los inversionistas, así como las órdenes informáticas que son empleadas por ellos mismos con el objetivo de aminorar las pérdidas.

Finalmente, los acontecimientos históricos han demostrado, que en un momento de crisis los inversionistas son propensos a entrar en pánico y reaccionar de manera visceral vendiendo todos sus activos. Por ello, el investigador al igual que el inversionista debe considerar abordar el estudio del mercado bursátil desde otra perspectiva. Empleando diferentes herramientas predictivas, así como afinar aquellas que han demostrado ser valiosas. De igual forma, debe tomar en cuenta el uso de técnicas de descomposición de señales, ya que estas han demostrado ser herramientas que apoyan de manera sobresaliente en el análisis y posterior predicción del mercado bursátil.

## Referencias

- Abbasi, N. Mousazadeh; Aghaei, M.; fard, M. Moradzadeh, M. (2015). Forecasting Stock Market Using Wavelet Transforms and Neural Networks and ARIMA (Case study of price index of Tehran Stock Exchange). *International Journal of Applied Operational Research*, 31-40
- Adams, A. J. (1978). Modeling and forecasting seasonal time series of product sales across territories : a comparative study. Tesis maestría: University of Iowa.
- ADVFN. (2017). Datos Históricos IPC México – ME. <https://mx.advfn.com/bolsa-de-valores/BMV/ipc-mexico-ME/historico/more-datos-historicos>.
- Alonso, J. C., García, J. C. (2009). ¿Qué tan buenos son los patrones del IGBC para predecir su comportamiento? *Estudios Gerenciales*, 25(112), 13-36.
- Bessler, D. A., Chamberlain, P. J. (1987). On Bayesian composite forecasting. *Omega*, 43-48.
- Bi, N., Sun, Q., Huang, D., Yang, Z., Huang, J. (2007). Robust Image Watermarking Based on Multiband Wavelets and Empirical Mode Decomposition. *IEEE Transactions on Image Processing*, 1956-1966.
- Black, F., Scholes, M. S. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 637-654.
- Box, G. E., Jenkins, G. (1990). *Time Series Analysis: Forecasting and Control*. Holden-Day, Incorporated.
- Brooks, C. (2014). *Introductory Econometrics for Finance*. Cambridge: Cambridge University Press.
- Bullinaria, J. A. (2007). Understanding the advantages of modularity in neural systems. In *Proceedings of the Twenty-eighth Annual Conference of the Cognitive Science Society*, 119–124, Mahwah, NJ: Lawrence Erlbaum Associates.
- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987–1007.
- Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 383-417.
- Fama, E. F. (1965). The Behaviour of Stock Market Prices. *Journal of Business*, 34-105.
- Finnoff, W., Hergert, F. Zimmermann, H. G. (1993). Improving model selection by nonconvergent methods. *Neural Networks*, 6, 771-783.
- Gonzales Martínez, R., Luna Acevedo, B. (2011). Predicciones del WTI para el año 2011: para el año 2011: Combinaciones de los Pronósticos de un Modelo Autorregresivo de Redes Neuronales Artificiales y de un Modelo ARIMA-GARCH. 4to Encuentro de economistas de Bolivia.
- Granger, C. W., Ramanathan, R. (1984). Improved Methods of Combining Forecasts. *Journal of Forecasting*, 3(2), 197-204.
- Hancock, P. (1992). Genetic algorithms and permutation problems: a comparison of recombination operators for neural net structure specification. *Combinations of Genetic Algorithms and Neural Networks, 1992, COGANN-92. International Workshop on (108-122)*. IEEE.
- Hellström, T., Holmström, K. (1998). Predicting the Stock Market. (S. Mälardalen University, Ed.) *Opuscula*.
- Hol, E. M. (2003). *Empirical Studies on Volatility in International Stock Markets*. Amsterdam: Springer Science Business Media.
- Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. Cambridge, MA, USA: MIT Press.
- Huang, N. E., Shen, Z., Long, S. R. (1999). A new view of nonlinear water waves: the Hilbert Spectrum. *Annual review of fluid mechanics*, 31(1), 417-457.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., y otros. (1988). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London (págs. 903-995)*. The Royal Society.
- Landassuri-Moreno, V. M. (2012). Evolution of modular neural networks. Tesis Doctoral. School of Computer Science. University of Birmingham.
- Landassuri-Moreno, V., Bullinaria, J. A. (2009). Feature Selection in Evolved Artificial Neural Networks using the Evolutionary Algorithm EPNet. *Proceedings of the 2009 UK Workshop on Computational Intelligence*. Nottingham, UK: University of Nottingham: UKCI '2009.

- Landassuri-Moreno, V., Bustillo-Hernández, C., Carbajal-Hernández, J., Sánchez Fernández, L. P. (2013). Single-step-ahead and Multi-step-ahead Prediction with Evolutionary Artificial Neural Networks. En J. Ruiz-Shulcloper, G. Sanniti di Baja, *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications* (págs. 65-72). Havana: 18th.
- Lee, K. J., Chi, A. Y., Yoo, S., Jin, J. J. (2008). Forecasting Korean Stock Price Index (KOSPI) Using Back Propagation Neural Network Model, Bayesian Chiao's Model, and Sarima Model. *Journal of Management Information and Decision Sciences*, 11(2), 372-378.
- León Anaya, L., Landassuri-Moreno, V., Bustillo-Hernández, C., Orozco Aguirre, H.R. (2013). Predicción de la Demanda de Energía Eléctrica Comparando Métodos Clásicos contra Redes Neuronales Artificiales Evolucionadas. RVP-AI/2013, Memoria Técnica (págs. 1-6). Acapulco Gro.: IEEE Sección México.
- Liu, B., Riemenschneider, S., Xu, Y. (2006). Gearbox fault diagnosis using empirical mode decomposition and Hilbert spectrum. *Mechanical Systems and Signal Processing*, 718-734.
- Liu, Y., Yao, X. (2001). Evolving neural networks for Hang Seng stock index forecast. 256-260: *Evolutionary Computation, 2001. Proceedings of the 2001 Congress*.
- Magdon-Ismail, M., Nicholson, A., Abu-Mostafa, Y. S. (1998). Financial markets: very noisy information processing. *Proceedings of the IEEE*, 86, págs. 2184-2195.
- Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., y otros. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting*, 111-153.
- McNelis, P. D. (2005). *Neural networks in finance: gaining predictive edge in the market*. Burlington, M: Elsevier AP.
- Ortiz-Rodríguez, J. M., Martínez-Blanco, M. d. R., Cervantes Viramontes, J. M., y Vega-Carrillo, H. R. (2013). Robust design of artificial neural networks methodology in neutron spectrometry. In Suzuki, P. K., editor, *Artificial Neural Networks-Architectures and Applications*. InTech.
- Radcliffe, N. J. (1993). Genetic set recombination and its application to neural network topology optimization. *Neural Computing and Applications*, 1(1):67-90.
- Rilling, G., Flandrin, P., Gonçalves, P. (2003). On empirical mode decomposition and its algorithms. *IEEE-EURASIP workshop on nonlinear signal and image processing*, 8-11.
- Roberts, H. (1967). *Statistical versus Clinical Prediction in the Stock Market*. Unpublished paper presented to the Seminar on the Analysis of Security Prices, University Chicago.
- Sammut, C. (2011). *Encyclopedia of Machine Learning*. Springer.
- Samuelson, P. A. (1965). Proof that Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review*, 41-49.
- Saravanan, N., Fogel, D. B. (1995). Evolving Neural Control Systems. *IEEE Expert: Intelligent Systems and Their Applications*, 23-27.
- Schaffer, J., Whitley, D., Eshelman, L. (1992). Combinations of genetic algorithms and neural networks: a survey of the state of the art. In Whitley, D. and Schaer, J., editors, *International Workshop on Combinations of Genetic Algorithms and Neural Networks (COGANN-92)*, 1-37, Piscataway, New Jersey. IEEE Press.
- Schwefel, H.-P. (1981). *Numerical Optimization of Computer Models*. New York, NY, USA: John Wiley Sons, Inc.
- Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 425-442.
- Werbos, P. J. (1994). *The roots of backpropagation: From ordered derivatives to neural networks and political forecasting*. New York: Wiley.
- White, H. (1988). Economic prediction using neural networks: the case of IBM daily stock returns. *International Symposium on Neural Networks*.
- Yanga, Z., Yang, L., Qic, D., Suene, C. Y. (2006). An EMD-based recognition method for Chinese fonts and styles. *Pattern Recognition Letters*, 1692-1701.
- Yao, X. (1999). Evolving artificial neural networks. *Proceedings of the IEEE*, 87(9):1423-1447.
- Yao, X. y Liu, Y. (1997a). EPNNet for chaotic time-series prediction. In *SEAL'96: Selected papers from the First Asia-Pacific Conference on Simulated Evolution and Learning*, 146-156, London, UK. Springer-Verlag.
- Zhang, G. (2000). Neural networks for classification: a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 451-462.
- Zhang, G. P., Berardi, V. L. (2001). Time Series Forecasting with Neural Network Ensembles: An Application for Exchange Rate Prediction. *The Journal of the Operational Research Society*, 652-664.
- Zorin, A. y Borisov, A. (2002). Modelling Riga stock exchange index using neural networks. In *Proceedings of the International Conference "Traditions and Innovations in Sustainable Development of Society. Information Technologies"*, 312-320.

