



UNIVERSIDAD DEL BÍO-BÍO



“Evaluación predictiva del Modelo ARIMA optimizado con fuerza bruta operacional aplicado en el precio del cobre y en el índice bursátil Dow Jones. 2011-2019”

Programa Articulado

Autor : Catalina Viveros Navarrete

Profesor Guía : Dr. Luis Améstica Rivas

Chillán, Marzo de 2020

Evaluación predictiva del Modelo ARIMA optimizado con fuerza bruta operacional aplicado en el precio del cobre y en el índice bursátil Dow Jones. 2011-2019

Catalina Viveros Navarrete¹

RESUMEN

Eliminar la incertidumbre que se genera en los mercados bursátiles ha sido objeto de estudio desde sus inicios, razón por la cual esta investigación busca estimar la capacidad predictiva del modelo ARIMA optimizado con fuerza bruta para determinar el signo de la variación de los precios del cobre y del índice bursátil Dow Jones (DJI) para el periodo comprendido entre los años 2011 y 2019. La investigación tiene un enfoque cuantitativo y de carácter exploratorio, en donde se utilizaron los precios de cierre semanales de ambas variables en el periodo anteriormente mencionado. Este estudio busca comprobar si el modelo utilizado logra una capacidad predictiva mayor o igual al 60%, arrojando como resultado arrojaron una capacidad predictiva de un 49% para el cobre y un 51% para el DJI, no estadísticamente significativa según test de Pesaran y Timmermann.

Palabras claves: modelos predictivos, ARIMA, fuerza bruta, Dow Jones, cobre.

ABSTRACT

Solving doubts and uncertainly generated by the stock markets has been studied since its beginnings, which is why this research seeks to estimate the capacity of the ARIMA model optimized with brute force to determine the sign of the variation in the prices of copper and Dow Jones stock index for the period between 2011 and 2019.

The research will have a quantitative and exploratory approaches, in which the weekly closing prices of both variables will be used in the previous mentioned period. The purpose of this study is to verify if the model used achieves a predictive capacity greater or equal than 60%.

Key words: predictive models, ARIMA, brute force, Dow Jones, copper.

¹ Magíster en Gestión de Empresas, Facultad de Ciencias Empresariales de la Universidad del Bío-Bío, Chillán, Chile. Email: catalina.viveros1401@alumnos.ubiobio.cl

INTRODUCCIÓN

El comportamiento del mercado de activos es incierto, la incertidumbre que se genera en este mercado influye directamente en la toma de decisiones de inversionistas, es por esta razón que el poder predecir el comportamiento de los precios de las acciones y/o commodities que se tranzan en estos mercados ha concentrado el interés de diferentes personas que se desempeñan en esta área, tales como: inversionistas, académicos, analistas entre otros. Este escenario hizo necesario la creación de modelos con capacidad predictiva que permiten disminuir la incertidumbre y por ende, el riesgo generado en este mercado. Asimismo predecir los cambios y tendencias en los precios de los activos, permitiría crear proyecciones para obtener mayores beneficios económicos.

La literatura muestra que se han creado distintos modelos predictivos a través de los años, los cuales han ido evolucionando en su complejidad y su relación con otras disciplinas que han aportado a las finanzas. Estudios anteriores han sido aplicados en los commodities que se transan a nivel mundial, como lo es la investigación de Foix, Weber (2007) donde se evaluó el desempeño predictivo de diferentes redes neuronales de tipo multilayer perceptron, construidas sobre la base de precio rezagado del cobre y variables derivadas del mismo, los resultados obtenidos se contrastaron con los generados mediante la aplicación de los modelos más tradicionales y exitosos de series de tiempo (ARIMA. Caminata aleatoria y promedio móvil). Los resultados arrojaron que los modelos no lineales de pronóstico, basados en redes neuronales multilayer perceptron, superaron la exactitud de los modelos lineales de series de tiempo más comúnmente usados. De esta manera, las redes neuronales demostraron ser capaces de aprovechar mejor la información contenida en los precios históricos considerados en el estudio.

Por otra parte Parisi, Améstica y Chileno, (2019) en su investigación logra desarrollar un modelo ARIMA optimizado con fuerza bruta con una capacidad predictiva superior al 60% en el porcentaje de variación de precio del petróleo, para llegar a este porcentaje se realizaron cien mil iteraciones. En esta misma línea Parisi, Améstica y Lobos (2016) desarrollan un modelo similar al mencionado anteriormente, pero esta vez se realizan un millón de iteraciones y es aplicado en las opciones binarias en el índice bursátil NASDAQ, logrando también, un porcentaje de predicción del signo del precio superior al 60%. En este contexto, es que se optó por utilizar un modelo de regresión lineal (ARIMA) optimizado con fuerza

bruta para la predicción del signo de la variación del precio del cobre y del índice bursátil Dow Jones Average. Estas variables fueron elegidas debido a la importancia e influencia que tienen en la economía; en el caso del cobre, este cuenta con una alta participación en el producto interno bruto del país, un alza o una baja de este commodity afecta fuertemente a la economía chilena, en el caso de DJI este es un fiel reflejo del estado de la economía mundial, por lo tanto poder predecir su comportamiento disminuye la incertidumbre en los inversionistas y les permite anteponerse a los posibles escenarios negativos.

MARCO TEORICO

Buroto y Fabris (2016), señalan que en términos macroeconómicos, el nivel y la volatilidad en el precio de las materias primas constituyen también un tema importante para los países desarrollados y la economía mundial, por dos razones. La primera, debido al potencial efecto del aumento del nivel de precios para alimentar presiones inflacionarias no deseables. Y lo segundo, por el creciente vínculo entre activos financieros y productos primarios que ha tenido lugar en múltiples carteras de inversión. Este último fenómeno se conoce en la literatura como la “financiarización del mercado de commodities” Cheng y Xiong (2013). Desde la aparición de este fenómeno, una de las principales hipótesis para explicar la volatilidad en los precios remite al mayor peso que los nuevos instrumentos financieros representa en la determinación del valor de las materias primas. Es decir, la explicación tradicional que relaciona el precio de los commodities con la oferta y demanda mundial ha dado paso a una segunda explicación que hace referencia a factores globales de carácter monetario y financiero (Bastourre, Carrera, Ibarlucia y Sardi, 2012).

Para reducir esta incertidumbre existente en el mercado se han creado diferentes modelos de predicción. Ejemplo de ello, aquellos basados en la llamada inteligencia artificial, que buscan obtener un resultado confiable, dentro de estos se encuentran modelos, tales como: los Autómatas celulares, modelo matemático propuesto por Neumann y Ulam (1940), definido como un sistema dinámico formado por un conjunto de elementos sencillos idénticos entre sí, pero que en conjunto son capaces de demostrar comportamientos complejos globales. Estos son usados para modelar sistemas complejos de cualquier índole, por lo que no sorprende que en las finanzas y la economía los expertos hayan hecho lo mismo. Varios especialistas han hecho investigaciones y han utilizado a los autómatas celulares para

predecir los cambios en los signos de los precios de las acciones y los resultados han sido positivos.

Así mismo en 1943 McCulloch y Pitts desarrollan el modelo de Redes neuronales, como concepto, puede ser explicada como un poderoso sistema computacional paralelo, capaz de resolver problemas altamente complejos. En el mundo de las finanzas este modelo ha sido utilizado en diferentes estudios, uno de ellos por Parisi et al (2006) en el cual se obtiene como resultado que las redes neuronales tienen una gran capacidad de predicción del signo de las variaciones de los índices, pudiéndose demostrar que la mayor capacidad predictiva permite mayores retornos ajustados por riesgo, incluso en los índices más volátiles.

Por otra parte en 1970, Holland propone los algoritmos genéticos, métodos adaptativos que pueden usarse para resolver problemas de búsqueda y optimización. Esto se ve reflejado en el estudio de Pereira (1996) donde los utilizó para encontrar los valores óptimos de los parámetros usados por tres reglas de transacción distintas para el tipo de cambio dólar estadounidense/dólar australiano: los parámetros obtenidos mostraron resultados intramuestrales positivos, los cuales disminuyeron al aplicar las reglas fuera de la muestra, aun cuando continuaron siendo rentables.

Todavía cabe señalar el modelo ARIMA, el cual es un modelo estadístico implementado por Box y Jenkins (1976) los cuales realizan las previsiones basándose en lo ocurrido en el pasado. Para ello tienen en cuenta tanto los valores observados en períodos anteriores (componente auto regresivo) como los factores aleatorios o de innovación en cada uno de ellos (componente de media móvil). Este modelo ha sido utilizado en el área de las finanzas para poder obtener una predicción del precio de diferentes commodities.

PREDICCIÓN EN COMMODITIES EN EL CONTEXTO INTERNACIONAL.

En el contexto de series de tiempo, pronosticar significa predecir valores futuros de una variable aleatoria basándose en el estudio de la estructura definida por las observaciones pasadas de variables que explican su variación, suponiendo que la estructura del pasado se conserva en el futuro (Pajouh y Danech, 2002). En consecuencia, se supone que las regularidades que puedan contener las series históricas de variaciones de precios e índices bursátiles son percibidas por los modelos de series de tiempo, de redes neuronales y de algoritmos genéticos. No obstante, el grado de predictibilidad de los rendimientos es por lo

general considerado económicamente no significativo, por lo que Hodgson y Nicholls (1991) sugieren evaluar la significancia económica de predecir la dirección de los cambios en los precios de los activos.

Igualmente, estudios como el de Wu y Zhang (1997), han sugerido que las estrategias de transacción basadas en estimar la dirección del cambio en el nivel de precios son más efectivas y generan mayores beneficios que las basadas una predicción puntual del nivel de precios, también determinaron que las estrategias de transacción basadas en la estimación del cambio en la dirección de los niveles de precios son más efectivas y eficientes, y permiten obtener beneficios mayores que aquellas estrategias basadas en la predicción específica de los niveles de precios de los instrumentos financieros.

Con respecto a los modelos de predicción lineal más utilizados en la literatura, se encuentra el Modelo Autorregresivo Integrado de Media Móvil (ARIMA) fue desarrollado por Box y Jenkins (1970), y tienen como objetivo proyectar una variable en función del comportamiento pasado de series de tiempo o, mejor dicho, variables temporales. Este modelo utiliza información más reciente como valor de partida y luego analiza los errores de proyección más reciente para así ajustarlo apropiadamente.

Dicho lo anterior, diferentes estudios han utilizado el modelo ARIMA en sus predicciones como lo es el de Broz y Viego (2014), evaluaron la capacidad predictiva de los modelos ARIMA aplicados a los precios de multiproductos forestales industrializables para la zona norte de Misiones (Argentina), siguiendo la metodología de BoxJenkins. Esto nace de la necesidad de aplicar herramientas de gestión forestal para tomar mejores decisiones de inversión y manejo de los montes debido a que en el noroeste argentino se encuentra el polo forestal de mayor importancia del país, centrándose en las provincias de Misiones y Corrientes más del 60% de la producción, las cuales abastecen a 640 actividades industriales de un total de 731 que se desarrollan en la provincia. Se propuso un modelo autorregresivo de media móvil (ARIMA) basado en la metodología de Box-Jenkins para predecir los precios de cuatro productos de *Pinus* spp. Para ello se utilizan series temporales de precios correspondientes al periodo julio 2002-septiembre 2013. Los modelos propuestos predicen precios futuros con errores de predicción entre 0,9% y 1,8%.

En la misma línea, Ceballos y Pire (2015), en el cual plantearon una manera confiable para pronosticar el precio de la tonelada de arroz con el fin de planificar racionalmente la toma de

decisiones basada en estos pronósticos, dado el estudio el precio de la tonelada del arroz se ubicó en US \$665 bajo condiciones normales del mercado, en caso de ocurrir cambios en las políticas agrarias de los países productores, se presenten eventos o fenómenos naturales que afecten a la producción el precio de la tonelada se ubicara en US \$2.309. En caso extremo de que la producción sea superior a las expectativas de los productores, empresarios, agrícolas y comercializadores, el precio bajaría a US \$191,7.

MODELO ARIMA

George E. P. Box, y Gwilym M. Jenkins, en 1976 implementan un modelo basado en series de tiempo llamado ARIMA, Por sus siglas en inglés AutoRegressive Integrated Moving Average. Este modelo econométrico consta con tres componentes:

Proceso Autorregresivo (AR): Se define como modelo autorregresivo si la variable endógena de un periodo t es explicada por sí misma en las observaciones o datos pasados, multiplicados por un coeficiente que le da un peso específico a la información pasada.

Proceso Integrado (I): Se refiere al estado de la variable, es decir si se va a trabajar sobre el valor sin modificación sobre su primera o segunda diferencia. Entendiendo la primera diferencia simplemente como la primera variación de la serie en estudio. Por ejemplo, una serie de precios se entiende como integración en cero. Es decir, se trabajará con la variable pura, lo cual en términos generales no es recomendable dado que tienen tendencia y no se pueden modelar en esas condiciones. Un grado de integración 2 significa que el modelo se construirá sobre la variación de la serie en estudio, es decir, no se modela el precio, sino la variación del precio, pero sí se obtiene un modelo para la variación de precios, se le suma el precio anterior y se obtiene la proyección de precios, es decir en niveles ARIMA y cero.

Proceso de Media Móvil (MA): Es aquel que explica el valor de una determinada variable en un periodo t en función de un término independiente y una sucesión de errores correspondientes a periodos precedentes, ponderados convenientemente. A continuación, se muestra la simbología y componentes de un modelo ARIMA:

ARIMA (p, d, q)

P = AR () autorregresivo como variable explicativa.

D= Integrado

Q= Error como variable explicativa (media móvil de los errores).

Por otra parte, la metodología de Box y Jenkins supone que no se conocen o que no están disponibles aquellas variables que determinan el valor proyectado, siendo en esta situación donde los Modelos ARIMA tienen una gran ventaja sobre los modelos de regresión típicos, y a principios de los 80 fueron ampliamente utilizados para modelar la media de un activo (Alonso & García, 2009) y según (Pai & Lin, 2004) el modelo ARIMA, ha sido uno de los enfoques más utilizados en predicción.

Para comprender mejor, el caso más simple de modelos no estructurales para datos longitudinales es el proceso de media móvil. Sea u_t (con $t = 1, 2, \dots, T$) un proceso ruido blanco, con $E(u_t) = 0$ y $\text{var}(u_t) = \sigma^2$, luego:

$$y_t = \mu + u_t + \theta_1 * u_{t-1} + \theta_2 * u_{t-2} + \dots + \theta_q * u_{t-q} \quad (1)$$

$$y_t = \mu + \sum_{i=1}^q \theta_i * u_{t-i} + u_t \quad (2)$$

En este caso, t es el tiempo; $E(u_t)$ es la esperanza matemática de u_t ; $\text{var}(u_t)$ es la varianza constante; σ^2 es una varianza; y_t es la media móvil; μ es un parámetro perteneciente a la media móvil.

También, un proceso de media móvil es una combinación lineal de procesos de ruido blanco, de modo que y_t depende de valores presentes y pasados de un término de perturbación caracterizado como ruido blanco. De este modo, $E(y_t) = \mu$.

En un modelo autorregresivo (AR), el valor presente de y depende de valores pasados de y más un término de error. Un proceso autorregresivo de orden p se expresa en la ecuación (4):

$$y_t = \mu + \tau_1 * y_{t-1} + \tau_2 * y_{t-2} + \dots + \tau_p * y_{t-p} \quad (4)$$

$$y_t = \mu + \sum_{i=1}^p \varphi_i * y_{t-i} + u_t \quad (5)$$

En este caso φ y τ son parámetros autorregresivos.

En particular, la propiedad que se desea en los modelos autorregresivos es la estacionalidad de los coeficientes, de no cumplirse esta condición, el modelo causaría que los valores

anteriores del término de error tengan un efecto decreciente sobre el valor de y_t a medida que pasa el tiempo. La estacionalidad de un modelo AR de orden p se visualiza en raíces de la ecuación característica dentro del círculo unitario (Viego & Broz, 2012).

TECNICA DE LA FUERZA BRUTA

La técnica fuerza bruta (Parisi, 2015; Parisi, Améstica-Rivas y Chileno, 2016), utiliza la capacidad de las computadoras para poder encontrar la mejor solución a un problema de optimización. Esta técnica aplicada a los modelos multivariados dinámicos simula la inteligencia humana, puesto que genera diferentes escenarios en los cuales cada uno de ellos brinda una solución única al problema. La función de este modelo multivariado dinámico con fuerza bruta es comparar los nuevos escenarios generados con los anteriores y elegir el mejor. Dicho de otra manera, la técnica de la fuerza bruta recuerda al igual que un humano, el propio comportamiento para ofrecer una mejor solución a determinado problema; si ese comportamiento solucionó el problema, cada vez que suceda un escenario parecido se utilizará el mismo comportamiento. De la misma manera, los modelos multivariados dinámicos con fuerza bruta utilizan el mejor modelo.

Como afirma Durán (2006), la fuerza bruta consiste en enlistar todos los casos y para cada uno calcular la solución, identificando de este modo el caso que ofrezca la mejor solución. Asimismo, Riveros (2015), en un estudio para encontrar la solución óptima al problema del camino más corto para una empresa de logística, comenta que la solución más directa es con fuerza bruta, es decir evaluar todas las posibles combinaciones (de recorridos) y quedarse con el trayecto que utiliza una menor distancia.

En el caso de los modelos ARIMA, que son modelos de regresión, la aplicación de la técnica fuerza bruta permite generar infinitos coeficientes de un universo establecido, para darle un peso a cada variable establecida y evaluada en el modelo. Incluso se puede afirmar que usando fuerza bruta se puede contemplar todos los escenarios del universo establecido, siendo así, una mejora a los algoritmos genéticos, los cuales sólo buscan alrededor de un punto en el universo que ofrece una solución de primera instancia óptima (Parisi, Améstica-Rivas y Chileno, 2019).

METODOLOGIA

El método de investigación utilizado es cuantitativo de carácter exploratorio y corte longitudinal, donde se utilizaron precios de cierre semanal desde agosto de 2011 hasta agosto del 2019, tanto del cobre como del Dow Jones. Estos precios fueron obtenidos desde las paginas web de cochilco (http://boletin.cochilco.cl/estadisticas/grafico.asp?tipo_metal=1) y yahoo finanzas (<https://es.finance.yahoo.com/quote/%5EDJI?p=^DJI&.tsrc=fin-srch>) respectivamente. Para la construcción del modelo ARIMA optimizado con fuerza bruta fue necesario la elaboración de una macro en Microsoft Visual Basic, la cual permitió obtener la predicción de los precios, y así calcular el porcentaje de predicción de signo en el precio de las variables a estudiar. Se utilizaron 1.500.000 de iteraciones.

ANALISIS ESTADISTICO

Se construyó un modelo ARIMA optimizado con fuerza bruta computacional para poder evaluar su capacidad predictiva para frecuencias diarias. De esta manera se buscó determinar el comportamiento de este modelo en su función predictiva para conocer la variación del signo del precio del cobre para el valor en estudio. El planteamiento de la técnica fuerza bruta de Parisi (2015) y utilizado en Parisi, Améstica y Chileno (2016), tiene como objetivo encontrar los coeficientes óptimos que maximicen el porcentaje de predicción de signo de las variaciones semanales de las cotizaciones para el valor en estudio. El modelo autorregresivo integrado de promedio móvil de orden p, d, q - ARIMA (p, d, q): Al igual que un modelo ARMA, corresponde a la combinación de los modelos autorregresivo y al de medias móviles, con la característica de incluir un proceso de restablecimiento (el cual se designa como integración) de inestabilidad original presente en una serie de tiempo.

La forma general de un modelo ARIMA es semejante al de un modelo ARMA:

$$Y'_t = \varphi_1 Y'_{t-1} + \varphi_2 Y'_{t-2} + \varphi_3 Y'_{t-3} + \dots + \varphi_p Y'_{t-p} + \alpha_t - \theta_1 \alpha_{t-1} - \theta_2 \alpha_{t-2} - \dots - \theta_q \alpha_{t-q} \quad (6)$$

Donde: Y'_t : Es la serie inducida a la estabilidad

En ocasiones es apropiado incluir un término constante μ a los modelos AR, MA y ARIMA. En un modelo que contiene sólo parámetros de media móvil, el valor de la constante es la

media de los valores de la serie ($\square = \mu$). Por otro lado, en un modelo autorregresivo $\square = \mu (1 - \phi_1 - \phi_2 - \dots - \phi_p)$ (Arнау 2001). La inserción del término depende de la serie en estudio. Concretamente, depende del valor de la media. Si la media de todas las observaciones es cero o extremadamente cercana, no se incluye término constante. Si la media es significativamente distinta de cero, se incluye el término.

EVALUACIÓN DE LA PREDICCIÓN

Esta etapa consiste en evaluar la calidad de cada modelo en función a su porcentaje de predicción de signo alcanzado (PPS). La evaluación se realiza sobre la base de un conjunto extramuestral de 418 datos semanales para el modelo. Se utiliza la muestra total para estimar los coeficientes α , β y θ de cada modelo respectivamente, por medio de la minimización de la suma del cuadrado de los residuos del modelo, como para evaluar la capacidad predictiva de los modelos.

Además, para realizar esto se comparó el signo de la proyección con el signo de la variación observada en cada i -ésimo período, en el que $i = 1, 2, \dots, m$. Si los signos entre la proyección y el observado coinciden, entonces aumentan la efectividad del modelo analizado y, en caso contrario, disminuye su capacidad predictiva. Una vez proyectado el signo de la variación del precio para el período $n+1$, la variación observada correspondiente se incluye en la muestra de tamaño de n con objeto de reestimar los coeficientes del modelo, contando ahora con una observación más. Así, el mismo modelo, pero con sus coeficientes nuevamente calculados, este es utilizado para realizar la proyección correspondiente al período $n+2$. Este procedimiento recursivo se efectuó una y otra vez hasta concluir con las observaciones del conjunto extramuestral. Finalmente, el PPS de cada modelo se calculó de la siguiente forma:

$$\text{Porcentaje} = \frac{\sum_{t=n+1}^{n+m} p_i}{m} \quad (7)$$

$$p_i = \left\{ \begin{array}{l} 1, \\ \text{si} \\ 0, \end{array} \quad \left. \begin{array}{l} \Delta \text{Precio}_i * \hat{\Delta} \text{Precio}_i > 0 \\ \\ \text{en otro caso} \end{array} \right\}, i = [n + 1, \dots, n + m] \quad (8)$$

En la que Δ en la ecuación (8), representa la variación observada, Δ^{\wedge} la variación estimada, $n = 0$ y $m = 418$, se refiere al precio observado en el período i que es representado por t que también se refiere al período o tiempo en el que se observa el precio. De esta manera, los modelos multivariados dinámicos construidos optimizados con técnica fuerza bruta fueron evaluados en función de su capacidad para predecir el signo de los movimientos del precio del valor en estudio.

EVALUACIÓN ESTADÍSTICA

En esta etapa se aplicó la prueba de acierto direccional de Pesaran y Timmermann (1992), que proponen Parisi y Parisi (2010), para medir la capacidad de predicción de los modelos de ARIMA, el cual permite medir la significancia estadística de los modelos y su capacidad de predicción. El test de Pesaran y Timmermann testea las siguientes hipótesis: H_0 : la proyección basada en el modelo y el valor observado son independientes H_1 : la proyección basada en el modelo y el valor observado no son independientes

A continuación las ecuaciones 9,10..15. Presentan fórmula para el cálculo del test DA de Pesaran y Timmermann

$$DA = \frac{(SR - SRI)}{\sqrt{Var(SR) - Var(SRI)}} \quad (9)$$

Donde:

$$SR = PPS = \frac{\text{Cantidad de signos acertados}}{\text{Número de predicciones}} = \frac{\sum_{i=1}^n I_i(Y * \hat{Y} > 0)}{m} \quad (10)$$

Luego:

$$SRI = P * \hat{P} + (1 - P) * (1 - \hat{P}) \quad (11)$$

Donde:

$$P = \frac{\text{Signos } > 0 \text{ Observados}}{\text{Datos totales}} = \frac{\sum_{i=1}^n I_i(Y > 0)}{m} \quad (12)$$

$$\hat{P} = \frac{\text{Signos } > 0 \text{ proyectados}}{\text{Datos totales}} = \frac{\sum_{i=1}^n I_i(\hat{Y} > 0)}{m} \quad (13)$$

Finalmente:

$$Var(SRI) = \frac{m(2\hat{P}-1)^2 * P(1-P) + m(2P-1)^2 * \hat{P}(1-\hat{P}) + 4P\hat{P}(1-P)*(1-\hat{P})}{m^2} \quad (14)$$

$$Var(SR) = \frac{SRI*(1-SRI)}{m} \quad (15)$$

El Test DA tiene como supuesto que los valores obtenidos de la aplicación del test se distribuyen según una función normal, por lo tanto, los puntos críticos con un 95% de confianza son [-1,96; +1,96]. Este intervalo representa la zona de no rechazo de la hipótesis nula “H0”; si se rechaza la hipótesis nula, entonces se probará que los resultados obtenidos no son producto del azar.

RESULTADOS

En la tabla 1, se presenta el análisis de ambas series de valores de precios de cierre del cobre y del DJI. Ambas variables arrojaron como un resultado un coeficiente de asimetría negativo, esto significa que la distribución presenta valores que tienden a agruparse hacia la izquierda de la media. Es importante mencionar el valor entregado en estadístico de la curtosis el cual resultado ser negativo, lo cual significa que la distribución de los valores observados en ambas variables adoptaron una forma platicúrtica, es decir que la distribución tiene una forma menos apuntada y con colas menos gruesas que la normal.

Tabla 1: estadísticos descriptivos

Cobre	Estadístico	DJI
294,1827	Media	18524,5892
2,35264	Error típico	217,46204
298,8750	Mediana	17701.2598
263,95	Moda	10771.48
48,09982	Desviación estándar	4446,02193
2313,592	Varianza de la muestra	19767111,04
-0,643	Curtosis	-0,965
-0,068	Coefficiente de asimetría	-0,397
220,09	Rango	16560,55
193,95	Mínimo	10771,48
414,04	Máximo	27332,03
414,04	Suma	7743278,30
418	Cuenta	418

Fuente: elaboración propia

A continuación, se presenta el mejor modelo multivariado dinámico para cada caso de estudio de acuerdo con el porcentaje de predicción de signo:

$$\Delta Cobre_t = \alpha_1 \cdot \Delta Cobre_{t-1} + \alpha_2 \cdot \Delta Cobre_{t-2} + \alpha_3 \cdot \Delta Cobre_{t-3} + \alpha_4 \cdot \Delta Cobre + \beta_1 \cdot \varepsilon_{t-1} + \beta_2 \cdot \varepsilon_{t-2} + \beta_3 \cdot \varepsilon_{t-3} + \beta_4 \cdot \varepsilon_{t-4} + \varepsilon_t \quad (16)$$

$$\Delta DJI_t = \alpha_1 \cdot \Delta DJI_{t-1} + \alpha_2 \cdot \Delta DJI_{t-2} + \alpha_3 \cdot \Delta DJI_{t-3} + \alpha_4 \cdot \Delta DJI + \beta_1 \cdot \varepsilon_{t-1} + \beta_2 \cdot \varepsilon_{t-2} + \beta_3 \cdot \varepsilon_{t-3} + \beta_4 \cdot \varepsilon_{t-4} + \varepsilon_t \quad (17)$$

En la tabla 2 se muestran los coeficientes α_i y β_i del mejor modelo para el caso del cobre y del DJI estos en términos del porcentaje de predicción de signo que se obtiene por el modelo,

los cuales maximizan el PPS, estos datos se obtuvieron luego de desarrollar le modelo utilizando fuerza bruta operacional, alcanzando 1.5 millones de iteraciones.

Tabla 2: resumen de parámetros

Variables	Coefficiente	Variables	Coefficiente
Δcobre_{t-1}	-0,028628476	ΔDJI_{t-1}	-0,73817021
Δcobre_{t-2}	0,063393098	ΔDJI_{t-2}	0,326926446
Δcobre_{t-3}	-0,080233417	ΔDJI_{t-3}	-0,212887407
Δcobre_{t-4}	-0,070023907	ΔDJI_{t-4}	0,890839907
Δcobre_{t-5}	0,683424619	ΔDJI_{t-5}	-0,275215317
Δcobre_{t-6}	-0,314891447	ΔDJI_{t-6}	-0,651966186
Δe_{t-1}	0,317763727	Δe_{t-1}	-0,330062598
Δe_{t-2}	-0,553964805	Δe_{t-2}	-0,730937059
Δe_{t-3}	-0,407420843	Δe_{t-3}	0,379256375
Δe_{t-4}	0,567819561	Δe_{t-4}	0,18724776
Δe_{t-5}	-0,712817062	Δe_{t-5}	0,786218525
Δe_{t-6}	0,506540222	Δe_{t-6}	-0,459558718

En cuanto a la Tabla3, se exponen los porcentajes de predicción de signo arrojados para ambas variables, estos resultados se obtuvieron desde los mejores modelos producidos por ARIMA optimizado con fuerza bruta operacional. La capacidad predictiva fue estimada en un conjunto extramuestral de 418 datos semanales para ambos casos, la cual no resultó estadísticamente significativa en cada uno de los valores, de acuerdo con la prueba de acierto direccional (DA), rechazándose así la hipótesis de que si existe capacidad predictiva en los modelos ARIMA con fuerza bruta para el caso del precio del cobre y del DJI.

Tabla 3: Resumen prueba DA

Modelo	PPS	Prueba DA^a
Cobre	49%	-0,80846226*
DJI	51%	-0,49675693*

^aEl valor de z crítico es de 1.96 y 1.64, para una significación de 5% y 10%, respectivamente.

*Significativo al 5% y al 10%.

Finalmente, el valor del test DA se compara con el valor crítico de una tabla de distribución normal (z^* , el cual es del 1,96, para un nivel de significancia de 5%. Debido a que el valor absoluto del test DA es menos que Z^* , no se rechaza la hipótesis nula que plantea que la proyección basada en el modelo ARIMA optimizado con fuerza bruta aplicado tanto el cobre como en el DJI y los valores observados -0,808 y -0,496 respectivamente son independientes y, en consecuencia, se concluye que el modelo no tiene capacidad para predecir el signo de las variaciones del precio del cobre observado.

CONCLUSIÓN

No es factible construir un modelo predictivo con una capacidad de predicción superior al 60% para el caso del cobre y DJI. Los modelos se construyeron con un millón y medio de iteraciones con fuerza bruta, dado que la optimización por simplex y/o solver no alcanzaron el resultado esperado.

El modelo ARIMA optimizado con fuerza bruta operacional aplicado al precio del cobre y del DJI, bajo las mismas condiciones (418 datos semanales, tomados desde agosto del año 2011 hasta agosto del 2019, 1,5 millones de iteraciones para cada variable), entrego un mejor PPS para el caso del DJI con un 51%, mientras que para el cobre entrego un PPS del 49%, si bien el modelo tiene una capacidad predictiva, esta no es mayor al 60 % y no presenta significancia estadística según el resultado obtenido en el test DA para ambas variables.

De acuerdo a los resultados obtenidos queda abierta la investigación para que analistas, investigadores y/o interesados en el área, puedan utilizar otros métodos que permitan obtener una predicción de la variación del precio del signo de cobre y del DJI más precisa.

REFERENCIAS BIBLIOGRÁFICAS

Alonso, J., & García, J. (2009). ¿Qué tan buenos son los patrones del IGBC para predecir su comportamiento? Una aplicación con datos de alta frecuencia. *Estudios Gerenciales*, 25(112), 13-36.

Box, G., & Jenkins, G. (1976). *Time series analysis: Forecasting and Control*. San Francisco: Holden Day.

Bastourre, D., Carrera, J., Ibarlucia, J. & Sardi, M. (2012), "Dos síntomas y una causa: Flujos de capitales, precios de los commodities y determinantes globales", *Documentos de Trabajo 2012* (BCRA), nro. 57.

Buroto, N. & Fabris, J. (2016). Transmisión de la volatilidad entre los mercados de materias primas y los mercados de activos financieros. *Revista de Investigación en Modelos Financieros, Vol 2*, 65-83.

Broz, D. & Viego, V. (2014). Predicción de precios de productos de Pinus spp. Con modelos ARIMA. *Madera y bosques, Vol 20*, 37-46

Ceballos, S. & Pire, R. (2015). Estimación del precio internacional del arroz (*Oryza sativa* L.) bajo el modelo ARIMA. *Revista Mexicana de Ciencias Agrícolas*, (11), 2083-2089.

Cheng, I. & Xiong, W. (2013), "The Financialization of Commodity Markets", Working Paper 19642, National Bureau of Economic Research.

Durán, G. (2006). Investigación de operaciones, modelos matemáticos y optimización. Universidad de Chile. Obtenido de: http://old.dii.uchile.cl/~gduran/docs/charlas/junaeb_willy_8.pdf.

Hodgson, A., & Nicholls, D. (1991). The impact of index futures markets on australian sharemarket volatility. *Journal of business finance & accounting*, Vol 18 (2), 267-280.

Pajouh, & Danech. Methodology for traffic forecasting. The French National Institute for Transport and Safety Research (INRETS). Arcuel. 2002.

Pai, P., & Lin, C. (2004). A hybrid arima and support vector machines model in stock price forecasting. *Omega*, 497-505.

Parisi, A., Améstica, L., & Chileno, E. (2016). Modelo predictivo para variaciones de precio del petróleo. Optimización de ARIMA utilizando fuerza bruta operacional. *TEC Empresarial*, 13 (1), 1-24,

Parisi, A., Améstica-Rivas, L., & Lobos, F. (2016). Modelos ARIMA para opciones binarias en el índice bursátil Nasdaq. Optimización utilizando fuerza bruta operacional. *XVI international finance conference* Santiago Chile, 0-7.

Parisi, F. & Norambuena, M. (2007). *Proyección del Signo de las Variaciones del Precio de las Acciones: Modelos Multivariados Dinámicos*. Santiago: Universidad de Chile.

Pereira, R. (1996). "Selecting Parameters for Technical Trading Rules Using Genetic Algorithms", *Journal of Applied Finance and Investment*, vol. 1, núm. 3, julio-agosto, pp. 27-34.

Pesaran, M. & Timmermann, A. (1992). A Simple Nonparametric Test of Predictive Performance. *Journal of Business & Economic Statistics*, 1992, vol. 10, issue 4, 561-65.

Riveros, D. (2015). *Aplicación de la investigación de operaciones al problema de la distribución a una empresa logística*. Universidad Nacional Mayor de San Marcos. Obtenido de: http://cybertesis.unmsm.edu.pe/bitstream/cybertesis/4365/1/Riveros_vd.pdf

Viego, V., & Broz, D. (2012). *Un modelo de predicción de precios forestales basados en la metodología Box-Jenkins*. Argentina: Técnicas Forestales de Entre Ríos.

Wu, Y. & Zhang, H. (1997). Forward premiums as unbiased predictors of future currency depreciation: A non-parametric analysis. *Journal of International Money and Finance*, Vol 16 (4), 609-623.