



**VALIDACIÓN DE MEDIDAS DE EVALUACIÓN PARA EL PRONÓSTICO DE LA
TASA DE CAMBIO EN COLOMBIA**

**Julián Vélez Correa
Pedro Nieto Figueroa**

**Colegio de Estudios Superiores de Administración –CESA–
Maestría en Finanzas Corporativas
Bogotá
2016**

**VALIDACIÓN DE MEDIDAS DE EVALUACIÓN PARA EL PRONÓSTICO DE LA
TASA DE CAMBIO EN COLOMBIA**

**Julián Vélez Correa
Pedro Nieto Figueroa**

Directores:

**Javier Cadena Lozano
Miller Ariza Garzón**

**Colegio de Estudios Superiores de Administración –CESA–
Maestría en Finanzas Corporativas
Bogotá
2016**

Tabla de contenido

1.	INTRODUCCIÓN	6
1.1.	Discusión	6
1.2.	Alcance	8
1.3.	Organización del documento	9
2.	CONTEXTO ACTUAL.....	11
2.1.	Pronóstico y evaluación.....	11
2.2.	El pronóstico en Colombia.....	14
3.	LOS MODELOS DE PRONÓSTICO	18
3.1.	Los modelos cuantitativos de pronóstico.....	18
3.2.	Modelos ARIMA	22
3.3.	Movimiento browniano.....	24
4.	MEDIDAS Y <i>TEST</i> DE EVALUACIÓN DE PRONÓSTICOS	26
4.1.	Principales criterios de aceptabilidad estadística.....	26
4.2.	Los criterios de información y selección	28
4.2.1.	Criterio de R^2 y R^2 ajustado.....	29
4.2.2.	Akaike.....	31
4.2.3.	Schwarz.....	32
4.2.4.	Hannan y Quinn	33
4.3.	Funciones de pérdida	34
4.4.	Las medidas de precisión o de evaluación de pronósticos.....	36
4.4.1.	Clasificación	36
4.4.2.	Medidas dependientes de la escala	38
4.4.3.	Medidas basadas en porcentajes	42
4.4.4.	Medidas basadas en errores relativos.....	48
4.4.5.	Medidas con errores escalados.....	51
4.4.6.	Estadístico U de Theil.....	54
4.4.7.	Resumen y comparación de las medidas de precisión	56
4.5.	<i>Test</i> de verificación de exactitud.....	58

4.5.1.	<i>Test</i> MGN	59
4.5.2.	<i>Test</i> DM	61
4.5.3.	<i>Test</i> HLN.....	64
4.5.4.	<i>Test</i> Giacomini y White	66
5.	LA VARIABLE: TASA DE CAMBIO	69
5.1.	La relevancia de la tasa de cambio	70
5.2.	Los estudios en Colombia.....	71
6.	PLANTEAMIENTO METODOLÓGICO.....	75
6.1.	Fuentes de información.....	75
6.2.	Validación teórica	75
6.3.	Modelación econométrica.....	76
7.	RESULTADOS	78
7.1.	Los modelos	78
7.1.1.	ARIMA	78
7.1.2.	Movimiento browniano.....	79
7.1.3.	Los pronósticos	80
7.2.	Las medidas de evaluación de pronósticos	83
7.3.	Los <i>test</i> de comparación de pronósticos	88
7.4.	Conclusiones.....	90
	Bibliografía.....	95

Anexos

Anexo 1.	Pruebas de ruido blanco y de estacionariedad e invertibilidad	104
Anexo 2.	Movimiento browniano a partir de una caminata aleatoria	106
Anexo 3.	Pronóstico modelo ARIMA	108
Anexo 4.	Pronóstico modelo movimiento browniano.....	110
Anexo 5.	Pronóstico de ajuste	112

Índice de gráficas

Gráfica 1. Devaluación de monedas frente al dólar americano	71
Gráfica 2. Expectativas tasa de cambio - Percentiles	72
Gráfica 3. Modelo ARIMA - Tasa real vs. Pronósticos	82
Gráfica 4. Movimiento browniano - Tasa real vs. Pronósticos	82
Gráfica 5. Pronóstico de ajuste (Tasa real vs Pronósticos ARIMA y Movimiento Browniano).....	83

Índice de tablas

Tabla 1. Variables pronosticadas y medidas de evaluación en Colombia.....	17
Tabla 2. Funciones de pérdida y definiciones matemáticas	35
Tabla 3. Clasificación de las medidas de evaluación	36
Tabla 4. Criterios de evaluación y su importancia determinada por académicos y profesionales	37
Tabla 5. Características de las medidas de evaluación.....	56
Tabla 6. Valoración de las medidas de evaluación a partir de los criterios estadísticos y de usuario.....	58
Tabla 7. Encuesta de expectativas tasa de cambio.	73
Tabla 8. Pronóstico metodología <i>rolling</i>	81
Tabla 9. Resultados medidas de evaluación de pronóstico para el modelo ARIMA	84
Tabla 10. Resultados medidas de evaluación de pronóstico para el modelo de movimiento browniano....	85
Tabla 11. Resultados de los <i>test</i> y validación de hipótesis nula.....	89
Tabla 12. Características de las medidas de precisión	90
Tabla 13. Desempeño de las medidas de precisión	91

1. INTRODUCCIÓN

1.1. Discusión

El futuro es tal vez uno de los mayores enigmas que tiene la vida; todos sabemos que pasó ayer, todos tenemos la capacidad de construir o cambiar el hoy, pero ninguno puede garantizar 100% el mañana. Son diversos los factores externos, además de los internos, los que pueden interferir, alterar o cambiar en cualquier momento nuestros próximos minutos y días.

Si este es el panorama de nuestra vida personal, *¿cómo podrá ser el de la vida empresarial?* Para nadie es un secreto que conocer el futuro resulta el mayor deseo de cualquier administrador además, por supuesto, de generar grandes números; lo cual puede resultar una rentable consecuencia si se logra el primer objetivo. La joya de la corona sería entonces saber con anticipación que necesitarán, desearán y soñarán las compañías.

Producto de la naturaleza inquieta del ser humano, que lo lleva a romper aquellos límites que no le permiten obtener sus resultados, nacen disciplinas como la prospectiva, “ciencia que ayuda a anticipar las decisiones para tomar acciones hoy, para beneficiar el consumidor mañana; prospectar no es proyectar, es definir los posibles escenarios de los cursos de acción del mercado, del comprador y del consumidor en diferentes entornos buscando la mejor solución en todos los contextos” (EAFIT, 2014). Si bien la prospectiva mira hacia adelante, es el pronóstico quien nos permite allanar ese camino preparando el presente y mirando el pasado.

Con esta herramienta, la predicción mediante el pronóstico, se limita la incertidumbre, ya no se habla de un futuro a ciegas, sino de un futuro con “posibles escenarios”, una “sutil” diferencia semántica para algunos, que parte en dos la forma de planear, crear y administrar compañías: Planeación financiera, proyección de ventas, comportamientos de variables endógenas y exógenas del mercado y análisis riesgos son algunos de los campos donde el pronóstico juega un papel importante en la toma de decisiones y en el norte que se observe al ser parte de un sistema económico, una compañía y de un mercado que cambia a diario.

Con “ciencias” como la pronosticación, es claro que el margen de incertidumbre no desaparece pero si disminuye para el administrador; quienes tengan el conocimiento y la capacidad de emplearla adecuadamente pueden reorientar sus ejercicios estratégicos de manera un poco más certera.

El pronóstico entonces se puede definir como aquella metodología que intenta conocer el comportamiento en el futuro de cierta variable con algún grado de certeza. Los pronósticos permiten tomar decisiones en el presente sobre sucesos aún por ocurrir tal como lo menciona Capistrán (2010).

El problema en este escenario se plantea de la siguiente manera ¿Qué pasa cuando no son cientos sino miles las compañías que tienen una misma proyección del futuro, una misma información, en un mercado tan competitivo? ¿Quiénes podrán aprovechar mejor esta ventana al futuro, las más grandes, las más innovadoras, las de mayor capacidad de inversión llegando a una monopolización? ¿Se corre el riesgo de limitar las ofertas de valor y las propuestas de las compañías bajo este escenario? ¿Se puede caer en la tentación de creer que no existe un futuro diferente al que muestran los estudios?

A pesar de que el mercado mismo se encarga de equilibrar las ecuaciones planteadas en las preguntas anteriores, a través de “la incertidumbre”, nace el planteamiento de quién puede tomar ventaja y la respuesta es sencilla: Quien logre acertar en sus pronósticos. De todo lo anterior surgen las inquietudes con respecto a que tan adecuados son los resultados de los pronósticos y cómo podrían evaluarse los mismos.

Aquí es precisamente donde toman relevancia las medidas de precisión, que son aquellas que permitirán validar los pronósticos desde diferentes criterios estadísticos. Podría pensarse entonces en una clasificación inicial dependiendo de si estas medidas evalúan de manera individual o con respecto a otros modelos, para este último caso los llamados *test* de evaluación.

El presente trabajo de investigación nace entonces de la inquietud de identificar el estado del arte de estas medidas y establecer qué tan efectivas son, así como las ventajas y desventajas de cada

una de ellas a partir de la comparación de las mismas tanto de manera teórica como práctica a través de dos modelos de pronóstico de la tasa de cambio dólar – peso (TCRM)¹.

Si bien es cierto que varias de las medidas de evaluación de pronóstico son utilizadas en la actualidad como herramientas que generan un alto grado de confianza, no existe literatura precisa que permita evidenciar cual es la más adecuada, para qué o en qué casos se utiliza cada una, ni tampoco se han identificado textos que permitan comparar sus resultados. Esto se acentúa en mayor grado si se limita en el caso Colombiano.

Parte del propósito de esta investigación es validar y evaluar varias de estas medidas. Se espera que en el desarrollo del proceso investigativo y de acuerdo con la metodología de la investigación propuesta en este documento, los resultados aporten de manera significativa al propósito definido y sean fuente de consulta académica para siguientes ejercicios de investigación. De igual forma, dada la importancia estadística de los instrumentos de evaluación, su aplicación puede ser considerada como referente para la toma de decisiones.

1.2. Alcance

Con todo lo anterior, la pregunta que se espera responder al final de la presente investigación es: ¿Cuáles son los elementos y bajo qué condiciones una medida de evaluación de un pronóstico presenta un mejor desempeño con respecto a otra? Esto último abarca los criterios estadísticos así como los del usuario. Así mismo y de la mano con esta pregunta, se pretende validar la hipótesis que el desempeño de las medidas de evaluación de un pronóstico difiere debido a la formulación misma de las medidas y a las características del pronóstico en si.

Consecuentemente con la hipótesis anterior el objetivo general es caracterizar las estadísticas de evaluación de pronóstico de series de tiempos (medidas y *test*) sobre una serie económica-financiera.

Lo anterior implica propósitos intermedios como lo son el identificar las principales medidas y *test* de evaluación de pronóstico teniendo en cuenta su propósito, estructura (unidades, función de

¹ Para efectos del presente documento cuando se mencione tasa de cambio se hará referencia a la relación dólar-

perdida y formulación matemática), ventajas y desventajas; clasificar las mismas acorde con sus características y usos; distinguir aquellas más utilizadas para la tasa de cambio en Colombia; comparar los *test* y las medidas de precisión; definir los modelos de pronóstico para ejemplificar los cálculos de las medidas y; analizar sus resultados para un pronóstico financiero que sumado al horizonte de tiempo, características y propósito del mismo, permita la selección del modelo más apropiado.

Al final, el resultado del presente documento será el análisis teórico y práctico de las medidas y *test* de evaluación de pronósticos, mediante el cual se pretende que sea un punto de mejora en la estimación de variables financieras y económicas.

1.3. Organización del documento

El presente documento se divide en dos grandes partes: La primera que presenta el estado del arte y marco teórico y, la segunda el ejercicio práctico. Para la primera sección se han desarrollado los capítulos 2, 3, 4 y 5 con el cual se abarca no solo el marco conceptual sobre el cual se basa la presente investigación sino también el análisis de información previa en Colombia y los estudios recientes. Los capítulos 6 y 7 presentan de manera general la metodología desarrollada y la modelación correspondiente para los pronósticos de la tasa de cambio.

Si se revisa cada una de la secciones, el capítulo 2 contextualiza el estado del arte y clasifica la literatura al respecto. El capítulo 3 inicia con una presentación general de lo que es un modelo de pronóstico, la clasificación de las regresiones, los pronósticos de series de tiempo dentro de los cuales se detallan los modelos ARIMA y de movimiento browniano.

Continuando, el capítulo 4 desglosa una a una un conjunto de medidas y *test* de evaluación de pronósticos así como incluye los conceptos de criterios de aceptabilidad estadística, criterios de información y funciones de pérdida. Por su parte, el capítulo 5 presenta la relevancia de la tasa de cambio en el día a día de la economía y de las finanzas incluyendo una revisión de algunos estudios realizados en Colombia con esta variable.

La segunda sección inicia con el planteamiento metodológico que permite argumentar la investigación en el capítulo 6 y termina con los resultados obtenidos en el capítulo 7. Al final de este último, se presentan las conclusiones correspondientes junto con algunas ideas para adelantar exploraciones futuras o dar continuidad al tema objeto de este documento.

2. CONTEXTO ACTUAL

El presente capítulo tiene como objetivo la contextualización al lector del estado del arte en cuanto a los ejercicios metodológicos que han realizado diferentes autores dentro y fuera de Colombia. Se ha dado especial énfasis en información relacionada con la actividad de pronóstico así como la manera de evaluar los modelos correspondientes a través de medidas de precisión. A lo largo de este capítulo se mencionarán algunas de ellas como son MAPE, GMRAE, MdRAE, MdAPE, MASE, RMSSE, MdASE, MAE, RMSE, RMSPE, RAE, MSE y el U de Theil. Estas medidas serán explicadas en el Capítulo 4.

2.1. Pronóstico y evaluación

El pronóstico involucra el uso de toda la información disponible en cuanto a series de datos, modelos formales e informales, supuestos subjetivos e incluso la pericia misma del investigador con el objetivo de poder realizar afirmaciones acerca de la probabilidad de ocurrencia de sucesos futuros (Arrow & Intriligator, 2006). De estos elementos que mencionan los autores, precisamente se da comienzo a la calidad del pronóstico que se obtendrá de un modelo.

Idealmente, los métodos de pronóstico deben ser evaluados en las situaciones para las cuales ellos serán empleados. La evaluación consiste en cuatro pasos: “Probar las asunciones, evaluar los datos y métodos de cálculos, replicar las salidas y, valorar y validar nuevamente las salidas” (Armstrong J. , Evaluating Forecasting Methods, 2001, pág. 443). La mayoría de los métodos de pronóstico están basados en los procedimientos metodológicos como los criterios pre especificados o la obtención de una muestra de errores. Por lo anterior es posible efectuar un chequeo o *test* de ciertos principios de evaluación, tales como: Usar alternativas razonables, hacer pruebas a datos y métodos, replicar salidas e identificar errores y evaluar los resultados.

Para el análisis aplicado de los modelos de pronóstico se encuentran múltiples autores. Para iniciar, Moskowitz y Wright (1982) mencionan que el pronóstico predice lo que pasará si las tendencias históricas no cambian, si esto no sucede, cada pronóstico deberá ajustarse, de otra manera resultará engañoso. Por su parte, Granger y Newbold (1977) buscaron determinar si el pronóstico de consenso incorpora eficientemente información disponible de acuerdo con las propiedades que provienen del marco conceptual de pronósticos óptimos.

De igual manera, Munir y Misas (2007) hacen referencia a la caracterización del pronóstico óptimo y hacen una breve definición al respecto. Los autores mencionan que la calidad de “óptimo” se da por la minimización de la pérdida esperada, tomando como referencia, entre otros, el artículo de Granger (1999). En este mismo sentido, Teräsvirta, van Dijk y Medeiros (2004), en relación a la construcción de pronósticos óptimos, exponen algunos planteamientos para su estructura y estimación para diferentes horizontes de tiempo. Clements, Joutz y Stekler (2007) señalan que los pronósticos entre horizontes facilitan la interpretación y aumentan el poder de las pruebas de eficiencia.

Otros autores en sus investigaciones incorporan la validación de pronósticos a partir de medidas de precisión o *test* diferenciales. Por ejemplo, Hanke y Reitsch (1995) y Bowerman, O’Connell y Koehler (2004) recomiendan la utilización de la medida MAPE (*Mean Absolute Percentage Error*, en español media de los errores absolutos en términos porcentuales). Makridakis, Wheelwright y Hyndman (1998) están en desacuerdo con el uso de esta medida en algunas circunstancias que advierten en sus casos de investigación: Cuando la serie no tiene una escala con un origen que tenga un significado real y cuando las series tienen ceros como valores.

Por otro lado, Armstrong y Collopy (1992) aconsejan la utilización de medidas tales como GMRAE (*Geometric Mean Relative Absolute Error*), MdRAE (*Median Relative Absolute Error*) y MdAPE (*Median Absolute Percentage Error*) que adicionan facilidades de medición de factores importantes en las series de tiempo como la tendencia, estacionalidad y ciclicidad. Fildes (1992) igualmente recomienda MdAPE y GMRAE.

Por su parte, Hyndman y Koehler (2005), en sus estudios escalan² los errores mediante el uso de un pronóstico base (el método ingenuo³ – *naïve method*). Estos autores proponen que las medidas basadas en errores escalados deben convertirse en la aproximación estándar cuando se compara la precisión de pronósticos en escalas diferentes. Así mismo, se inclinan por el uso de la medida MASE (*Mean Absolute Scaled Error*) porque es menos sensible a valores atípicos cuando se comparan los resultados con los de la medida RMSSE (*Root Mean Square Scaled Error*) y menos variable en muestras pequeñas que MdASE (*Median Absolute Scaled Error*).

En cuanto a las medidas relativas, las cuales remueven la escala de los datos comparando los pronósticos generalmente con una caminata aleatoria⁴ o con un modelo ingenuo (*naïve method*), algunos autores encuentran que tienen el inconveniente que los errores relativos tienen una distribución estadística con media y varianza indefinida. Una situación común donde no es posible usar estas medidas es cuando se está validando la precisión fuera de la muestra en un único horizonte de proyección a través de múltiples series (Hyndman & Koehler, 2005).

Ahora bien, las medidas de evaluación discutidas por varios autores son el método clásico de validar un pronóstico. Sin embargo, algunos investigadores han propuesto mecanismos de pruebas inferenciales para validar la precisión cuando se comparan varios modelos. Aquí tienen cabida todos los *test* que buscan comprobar la hipótesis nula que corresponde a que dos modelos tienen la misma precisión.

Los *test* más conocidos son⁵:

² Escalar en este contexto significa que los errores son divididos por una escala que normalmente es un resultado que arroja un pronóstico de comparación. Adicionalmente hace referencia a las medidas escaladas que se explican en el Capítulo 4 – Medidas y *test* de evaluación de pronósticos.

³ El método ingenuo es uno de los más simples: Basado en el principio de inexistencia de cambios es aquel en el que la predicción es igual a la última observación de la variable.

⁴ En general, una caminata aleatoria es una serie de tiempo no estacionaria en la cual el valor siguiente depende del valor actual y un choque puramente aleatorio. Una serie de tiempo y_t sigue una caminata aleatoria si el cambio en y_t es independiente e idénticamente distribuido, esto es si, $y_t = y_{t-1} + \epsilon_t$ (con ϵ_t independiente e idénticamente distribuido). Adicionalmente se cumple que $E(\epsilon_t | y_{t-1}, y_{t-2}, \dots) = 0$ y $E(y_t | y_{t-1}, y_{t-2}, \dots) = y_{t-1}$. Esto último significa que el mejor pronóstico del valor de mañana es el valor de hoy.

⁵ Los *test* aquí mencionados serán explicados en el Capítulo 4 – Medidas y *test* de evaluación de pronósticos.

- El *test* MGN de Granger y Newbold (1977) que toma las bondades del trabajo realizado por Morgan (1939-40). Este *test* calcula el estadístico de prueba y busca comprobar la hipótesis nula de que las varianzas de los errores de dos modelos son iguales.
- El *test* DM de Diebold y Mariano (1995) quienes introducen los errores que no se distribuyen como una normal y correlacionados temporalmente. El mismo Diebold (2013) veinte años después discute el desempeño de su estadístico.
- Los *test* HLN de Harvey, Leybourne y Newbold (1997) que trabajan sobre los estadísticos de los dos anteriores corrigiendo entonces la posible heterocedasticidad⁶ del *test* MGN (*test* HLN-MGN) y modificando el DM para muestras pequeñas (*test* HLN-DM).
- El *test* de Giacomini y White (2006) quien en relación con DM adicionaron dos innovaciones: La muestra de los estimadores son asintóticas e incluyen objetivos de evaluación condicional.

2.2. El pronóstico en Colombia

En Colombia por su parte, la literatura relacionada con la validación de pronósticos está asociada normalmente a las estimaciones que se realizan sobre las variables de interés económico y financiero. Las investigaciones típicamente buscan pronosticar una o varias variables y usan las medidas de precisión y/o *test* para contrastar sus modelos. Sin embargo, los estudios están más orientados al pronóstico en sí que a la evaluación de los mecanismos que permiten validar y comparar la calidad de los mismos.

La mayoría de las investigaciones se encuentran publicadas por el Banco Central de Colombia (Banco de la República) e incluyen algunos fundamentales de relevancia macroeconómica tales como la inflación, la tasa de cambio nominal y el PIB (Producto Interno Bruto), entre otros.

Para el caso de la inflación, Castaño y Melo (1998) la estimaron mediante el uso de 15 modelos diferentes: 10 de regresiones simples y cinco de series de tiempo (VAR⁷ y ARIMA). Para la

⁶ El término heterocedasticidad surge de la unión de dos palabras: “hetero” (diferente) y “cedasticidad” (dispersión). Se dice entonces que un modelo lineal presenta heterocedasticidad cuando la varianza de las perturbaciones no es constante a lo largo de las observaciones.

evaluación utilizaron las medidas MAE (*Mean Absolute Error*), MAPE, RMSE (*Root Mean Square Error*), RMSPE (*Root Mean Square Percentage Error*) y el estadístico U-Theil así como un conjunto de métodos de combinación de pronósticos⁸.

En una investigación similar, Melo y Nuñez (2004) utilizaron siete modelos de pronóstico diferentes (sobresalen los modelos de series de tiempo ARIMA y STR⁹ de transición suave) para la estimación de la inflación. La evaluación de los resultados se realizó con los cálculos de las medidas RMSFE, RMSPFE, MAFE, MAPFE, U-Theil y RAFE¹⁰. Igualmente los autores utilizaron los *test* DM y HLN-DM para la comparación de todas las combinaciones posibles de sus siete modelos.

Aristizábal (2006) por su parte, exploró la relación no lineal entre el dinero y la inflación en Colombia a través de una red neuronal artificial¹¹ utilizando información mensual de la variación del IPC (Índice de Precios al Consumidor) y del agregado monetario M3 con datos mensuales del año 1982 al año 2003. El autor concluye basado en los resultados de las medidas RMSE, MAE, RMSPE y MAPE que el desempeño es mejor en un modelo de redes neuronales comparado con un modelo ARIMA.

⁷ La esencia de los modelos VAR es un sistema de ecuaciones autorregresivas o vector autorregresivo, con tantas ecuaciones como series a analizar o predecir sin distinguir entre variables endógenas y exógenas. Cada variable es explicada por los retardos de sí misma y por los retardos de las demás variables.

⁸ Es la combinación de dos modelos de pronóstico diferentes con el objetivo de incorporar de una manera óptima la información de los modelos y generar uno mejor. Para lo anterior existen técnicas lineales que se desarrollaron a partir de los estudios de Bates y Granger (1969) y no lineales basadas por ejemplo en modelos STR o en modelos redes neurales desarrolladas por Deutsch, Granger y Teräsvirta (1994) y Donalson y Kamstra (1996) respectivamente.

⁹ Los modelos regresivos de transición suave STR (Smooth Transition Regressive) fueron introducidos por Chan y Tong (1986); Luukkonen, Saikkonen y Teräsvirta (1988), Teräsvirta (1994) y extendidos a un contexto multivariable por Granger y Teräsvirta (1993). Los modelos STR permiten el cambio endógeno de un régimen a otro y las ecuaciones para cada régimen son estimadas simultáneamente.

¹⁰ La letra F incluida en las medidas dentro del documento de los autores viene de la palabra *Forecast*. De esta manera para utilizar la misma nomenclatura del presente documento, RMSFE = RMSE, RMSPFE = RMSPE, MAFE = MAE, MAPFE = MAPE y RAFE = RAE .

¹¹ Un modelo de red neuronal es no lineal que para uno de los casos involucra procedimientos estadísticos de estimación recursiva con una estructura de funciones multicapa con capacidad de aprendizaje adaptativo y auto organización.

Otras investigaciones que involucran el pronóstico de la inflación son las de Rodríguez y Siado (2003) y Núñez (2005). Los primeros contrastan un modelo no paramétrico¹² con uno ARIMA y otro tipo STAR¹³ utilizando MAE, MAPE, RMSE, RMSPE, el U-Theil así como el *test* de DM. El segundo continúa la investigación realizada por él mismo y Núñez (2004) comparando varios modelos con medidas de precisión y *test* diferenciales (MGN, DM, HLN-DM).

Si se analizan las investigaciones de la tasa de cambio nominal, Patiño y Alonso (2005) evaluaron cuatro pronósticos utilizando la medida RMS y el coeficiente U-Theil. Munir & Misas (2007) pronosticaron la variable a través de un modelo lineal ARIMA y de uno no lineal que considera una red neuronal artificial autorregresiva. Estos autores realizaron la evaluación con medidas simétricas (MAPE, MAE, RMSE y RMSPE) y asimétricas (lin-lin)¹⁴. Por último, con la tasa de cambio, está el modelo de redes neuronales desarrollado por Ayala y Castillo (2008) y validado con MSE (*Mean Square Error*).

Con respecto al PIB, Salazar (2009) lo pronosticó a través de un modelo de red neuronal con datos trimestrales de principios del año 1980 al segundo trimestre de 2008. Igualmente Cristiano, Hernández y Pulido (2012) realizaron una evaluación de pronóstico para el período desde el primer trimestre de 2005 al tercer trimestre de 2011 y calcularon los errores de pronóstico respecto a los crecimientos anuales del PIB. Las medidas usadas en ambas investigaciones fueron MAE, MAPE, RMSPE y RMSE.

Otros ejercicios de estimación incluyen los pronósticos de los precios de las acciones entre los que sobresale el realizado por Villada, Muñoz y García (2012) para los instrumentos Ecopetrol y Preferencial Bancolombia (medidas MAPE y RMSE); pronósticos hidrológicos de nivel de agua, del caudal líquido y del caudal sólido realizados directamente por el IDEAM (Rivera, y otros,

¹² Un modelo no paramétrico “(...) es un método de inferencia estadística válido cuando no se hacen hipótesis paramétricas sobre la distribución de los datos” (Delicado, 2008, pág. vi). Existen dos interpretaciones: La primera de ellas se basa en estadísticos libres de la distribución de los datos, mientras que la segunda se fundamenta en la estimación de funciones relacionadas con la distribución de probabilidad de los datos.

¹³ Los modelos autorregresivos de transición suave (Smooth Transition Regressive) fueron introducidos por Teräsvirta (1994) e involucra al igual que el modelo STR, el cambio de regímenes o niveles.

¹⁴ Ver definición de función de pérdida en numeral 4.3 del presente documento.

2008)¹⁵ y pronósticos de variables macroeconómicas en general. En todos es común el uso de medidas de precisión para con base en ellas tomar conclusiones acerca de los modelos planteados.

Para el último caso de variables macroeconómicas sobresale la investigación de Cajiao, Melo y Parra (2014). Los autores pronosticaron el IPC (Índice de Precios al Consumidor), el agregado monetario M1, el PIB, el IPI (Índice de Producción Manufacturera) y la inversión pero incluyeron el *test* de Giacomini y White (2006) como aporte fundamental en su estudio.

La Tabla 1 resume los autores que han realizado pronósticos de diversas variables así como las medidas de evaluación utilizadas.

Tabla 1. Variables pronosticadas y medidas de evaluación en Colombia

Variab les	Autores	Medidas Utilizadas
Inflación	Castaño y Melo (1998)	MAE, MAPE, RMSE, RMSPE y U de Theil
	Melo y Núñez (2004)	MAE, MAPE, RMSE, RMSPE, U de Theil y RAE <i>Test</i> DM y HLN-DM
	Aristizábal (2006)	MAE, MAPE, RMSE, RMSPE
	Rodríguez y Siado (2003)	MAE, MAPE, RMSE, RMSPE y U de Theil <i>Test</i> DM
	Núñez (2005)	MAE, MAPE, RMSE, RMSPE, U de Theil y RAE <i>Test</i> MGN, DM y HLN-DM
Tasa de Cambio	Patiño y Alonso (2005)	RMSE y U de Theil
	Múnir y Misas (2007)	MAE, MAPE, RMSE y RMSPE
	Ayala y Castillo (2008)	MSE
	Fayad, Fortich, Vélez (2009)	MAE, MAPE, RMSE y U de Theil
PIB	Salazar (2009)	MAE, MAPE, RMSE y RMSPE
	Cristiano, Hernández y Pulido (2012)	MAE, MAPE, RMSE y RMSPE
Precio de acciones	Villada, Muñoz y García (2012)	MAPE y RMSE
Nivel de agua y caudal	Rivera y otros (2008)	MSE, MAE y MAPE
IPC, M1, IPI	Cajiao, Melo y Parra (2014)	MSE, MAE y RMSE <i>Test</i> GW

Fuente: Elaboración propia

¹⁵ Si bien el libro fue trabajado por varios autores, el propietario del mismo es el IDEAM (Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia).

3. LOS MODELOS DE PRONÓSTICO

El pronosticar siempre ha sido una búsqueda a lo largo de la historia de la humanidad como iniciativa de sobrevivencia, anticipación y decisión. Los antiguos Egipcios pronosticaban las cosechas a partir del nivel que alcanzaba el río Nilo en la temporada de inundaciones. En el imperio romano se hablaba de predecir el futuro y fue el Emperador Constantino quien emitió un edicto donde prohibió las consultas directas a eruditos en predicción (Armstrong J. , 2001).

En el campo económico y financiero, ya en el siglo XVII Sir William Petty había lanzado su sistema cíclico de 7 años que analizaba el desempeño de los negocios en ese periodo de tiempo. En el siglo XIX se empiezan a desarrollar por su parte los modelos causales iniciales de regresión simple y múltiple junto con los modelos deterministas de series de tiempo de una sola variable.

Aquí se destacan los trabajos de von Thunen, Walrás, Pareto y Wicksell entre otros. Antes de la Gran Depresión en Estados Unidos, ya se había desarrollado la “industria del pronóstico” entre los años 1910 y 1930 que particularmente falló en la predicción de la crisis financiera de la época (Hawkins, 1996).

Todo lo anterior lleva a pensar que la actividad de pronóstico ha sido una inquietud constante del ser humano. Por esta razón y como parte del marco conceptual, en este capítulo se realizará una breve descripción de los modelos de pronóstico haciendo énfasis en el modelo ARIMA así como en el de movimiento browniano, los cuales serán utilizados para el pronóstico de la TCRM.

3.1. Los modelos cuantitativos de pronóstico

Según la teoría econométrica, hay dos tipos de metodologías para la realización de pronósticos cuantitativos: Los de regresión y los de series de tiempo. En los primeros una variable llamada dependiente o endógena es explicada por una o varias variables denominadas independientes,

exógenas o explicativas mediante un modelo econométrico. En los segundos, la variable dependiente es explicada por ella misma pero teniendo en cuenta su pasado.

El objetivo de la regresión es determinar la dependencia de la variable explicada de aquellas independientes o explicativas. Para esto se define una forma funcional de un modelo, que podría ser una función lineal, una cuadrática, una logarítmica o cualquier otra que el investigador y la teoría económica indique o presuponga.

Esta definición será la base para la estimación de los coeficientes y parámetros que le dan peso a las variables explicativas y que posteriormente serán validados para garantizar su mejor asertividad (Rosales & Bonilla, 2006). La regresión, como su nombre lo indica, implica conocer el pasado para que una vez se hayan completado los tres pasos mencionados, se pueda reproducir los comportamientos hacia el futuro, es decir, pronosticar.

Es importante mencionar que la estimación de los parámetros del modelo definido, normalmente se hace a través de dos métodos de estimación frecuentes: 1) El de mínimos cuadrados ordinarios (MCO) desarrollado por Carl Friedrich y, 2) el de máxima verosimilitud (MV o MLE) y los conceptos de probabilidad, suficiencia, eficiencia e información propuestos por R.A. Fisher entre 1912 y 1922 (Aldrich, 1997) aunque con algunos bosquejos iniciales de Bernoulli (1700-1782) y Euler en el siglo XIX.

Los modelos de regresión tienen diferentes clasificaciones dependiendo de cómo se analicen y comparen. Algunas clasificaciones de los modelos de regresión se presentan a continuación:

1. Según el número de variables. Pueden ser simples, cuando una variable dependiente es explicada por solo una independiente, y múltiples cuando es explicadas por más de una regresora.
2. Según sus relaciones. Pueden ser modelos uniecuacionales, donde únicamente existe una variable endógena, o multiecuacionales donde existen varias variables que se desean explicar. Dentro de estos últimos sobresalen los de ecuaciones simultáneas donde las variables están determinadas simultáneamente por un cierto número de variables predeterminadas.

3. Según su relación. Pueden ser modelos de regresión lineales o no lineales. Es importante tener en cuenta que la relación aquí especificada hace referencia a las variables y no a los parámetros ya que para estos últimos, existe el supuesto de linealidad según el Teorema de Gauss-Markov.
4. Según su retardo. Pueden ser modelos estáticos los cuales no tienen en consideración los retardos de sus variables explicativas o modelos dinámicos si por el contrario, lo hacen. Algunos autores como Barbancho (1976) y Chakravarty (1959) consideran la inclusión del retardo como el factor que introduce el carácter dinámico a un modelo mientras que otros autores como Tinbergen (1956) definen el dinamismo con solo el hecho que una variable se refiera a periodos de tiempo.
5. Según los datos que se utilizan. Pueden ser modelos de corte transversal que implican observaciones de una variable en un momento dado, de series de tiempo u observaciones a lo largo del tiempo o una combinación de los dos anteriores denominada tipo panel (Alvarez, y otros, 2011).
6. Según el sector exterior. Pueden ser abiertos si se refieren a economías abiertas o cerrados si por el contrario se analizan contextos específicos.
7. Según su ámbito de cobertura. Pueden ser macroeconómicos o microeconómicos.

Por otro lado y analizando los modelos que observan solo la misma variable pero en su pasado, se tiene que, en términos generales, hay cinco enfoques de los pronósticos económicos basados en series de tiempo: 1) métodos de suavizamiento exponencial, 2) modelos de regresión uniecuacionales, 3) modelos de regresión de ecuaciones simultáneas, 4) modelos autorregresivos integrados de promedios móviles (ARIMA) y, 5) modelos de vectores autorregresivos” (Gujarati & Porter, 2010, pág. 773).

Los modelos de suavizamiento exponencial buscan básicamente ajustar una curva a los datos que se disponen de la serie de tiempo. Un ejemplo es el suavizamiento exponencial lineal que busca la tendencia de línea recta de los datos dando mayor prioridad a los más recientes. Aquí sobresalen según Makridakis, Wheelwright y Hyndman (1998) los modelos de Brown para un parámetro que incluye únicamente el suavizador; el modelo de Holt para dos parámetros, que adicional al suavizador incluye los componentes de tendencia; y el modelo de Holt Winter para tres parámetros, el cual agrega a los dos anteriores el factor de estacionalidad que suelen incorporar los ciclos económicos y los términos de patrón estacional constante para casos aditivos y de patrón estacional creciente para casos multiplicativos.

Los modelos de una sola ecuación, es decir, modelos en los cuales había una sola variable dependiente Y y una o más variables regresoras, se explicaron en los métodos de regresión. Sin embargo, a partir de las series de tiempo se busca igualmente estimar un modelo con información de las variables en su pasado.

Un ejemplo de estos son los modelos ARIMA, los cuales corresponden a una combinación de los modelos de auto regresión (AR) y los de media móvil (MA). Mientras los primeros expresan la variable en función del pasado de la misma variable y de un impacto aleatorio, los segundos lo hacen en función de un conjunto de términos de error de periodos precedentes así como de un término independientes. La integración (I) de ambos mediante la diferenciación de la serie para hacer la serie de tiempo estacionario, componen los modelos ARIMA.

En contraste, están los modelos de ecuaciones simultáneas: “Esto sucede cuando Y está determinada por las X y algunas X, a su vez, están determinadas por Y. En otras palabras, hay una relación en dos sentidos, o simultánea, entre Y y (algunas) X, que hace dudar del valor de la distinción entre variables dependientes y explicativas” (Gujarati & Porter, 2010, pág. 671).

Los modelos de vectores autorregresivos (VAR) desarrollados por Christopher Sims (1980) toman como base precisamente las ecuaciones simultáneas incluyendo los términos autorregresivos o valores rezagados de la variable dependiente así como los vectores de dos o más variables. Una de las críticas que realiza Sims y que lo llevó al desarrollo de su modelo, es que si hay una verdadera simultaneidad entre un conjunto de variables, todas deben tratarse en

igualdad de condiciones, es decir, no debe haber ninguna distinción a priori entre variables endógenas y exógenas.

Si bien se mencionaron algunos ejemplos, es importante aclarar que existen múltiples modelos para la realización de pronóstico. Los modelos de la familia de suavización, los polinomiales, las redes neuronales, son algunos de ellos. Para el presente caso, es de interés el detallar los modelos ARIMA y el movimiento browniano.

3.2. Modelos ARIMA

Box y Jenkins (1970) con sus estudios en los 60s y 70s desarrollaron el cuerpo metodológico de estos modelos que rompen con el paradigma de especificar un modelo (revisión de marco teórico, identificación de variables y definición de forma funcional) de tal manera que son los propios datos de las variables los que indican la estructura probabilística subyacente y el mecanismo para pronosticar (de Arce & Mahia).

La gran ruptura de los estudios de Box y Jenkins fue reconocer que es posible definir y aproximar una amplia cantidad de patrones de comportamiento con solo unos pocos parámetros. Tal como lo indican Shewart y Wilks (2011), es la idea de parsimoniosidad¹⁶ lo que los guiaron a desarrollar los modelos ARIMA.

Estos modelos ya incluyen la descomposición de las series en su tendencia (componente de largo plazo que representa el crecimiento o disminución en la serie sobre un periodo amplio), ciclo (fluctuación alrededor de la tendencia), su factor estacional (patrón que se repite con frecuencia de corto plazo) y el componente aleatorio (la variabilidad de la serie).

Los ARIMA son modelos del tipo estocástico y requieren como insumo series estacionarias, las cuales en caso de no serlo, se buscan a través de la diferenciación de la variable misma.

¹⁶ El término de parsimoniosidad se deriva del latín parsimonia y surgió en el siglo XIV por el inglés Guillermo Ockham quien manifestaba que la explicación más sencilla es, probablemente, "más correcta que la más difícil y compleja". Es decir, esta implícito el concepto de que la naturaleza 'prefiere' lo simple antes que lo complejo (Universidad de Talca, sin fecha). En estadística es la "distancia" mínima entre puntos, por ejemplo para el cálculo de una regresión lineal se considera la suma mínima de cuadrados (De Luna, 1996).

Adicionalmente, tienen como supuesto fundamental que la serie es producida por un proceso estocástico cuya naturaleza puede ser identificada y caracterizada mediante un modelo que será al final, útil para el pronóstico.

Como se mencionó anteriormente, los modelos ARIMA son la mezcla de expresiones autorregresivas y de medias móviles. Un modelo autorregresivo por su parte, explica la variable endógena Y_t por las observaciones de ella misma en periodos anteriores incorporando un término de error. En el caso de procesos estacionarios con distribución normal, Y_t puede expresarse como una combinación lineal de sus valores pasados (la parte sistemática) más el error (parte no sistemática) según González (2009). La expresión genérica de un modelo autorregresivo de orden p $AR(p)$ es de la siguiente manera:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + a_t \quad (1)$$

Así por ejemplo, un modelo autorregresivo de orden 1 $AR(1)$ está definido por:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + a_t \quad (2)$$

Por otro lado, un modelo de medias móviles es aquel que explica la variable en función de un término independiente o deriva y un conjunto de errores correspondientes a periodos anteriores. La expresión genérica de un modelo de medias móviles de orden q $MA(q)$ es de la siguiente manera:

$$Y_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (3)$$

Así por ejemplo, un modelo de orden 1 $MA(1)$ está definido por:

$$Y_t = \mu + \theta_1 a_{t-1} \quad (4)$$

El modelo $ARMA(p,q)$ es su forma más general podría expresarse como la unión de las expresiones (1) y (3):

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (5)$$

Los modelos de series de tiempo se basan en el supuesto de que las series de tiempo consideradas son débilmente estacionarias¹⁷ (Gujarati & Porter, 2010). Una serie de tiempo puede ser diferenciada para volverla estacionaria: En general, si una serie de tiempo es de orden $I(d)$, después de diferenciarla d veces se obtiene una serie $I(0)$ que es estacionaria. Adicionado el orden para lograr estacionariedad, se obtiene el modelo genérico ARIMA(p,d,q).

3.3. Movimiento browniano

El primer registro de lo que es el Movimiento Browniano está relacionado a las observaciones del botánico Robert Brown en 1828 cuando reportó en una revista científica que los granos de polen suspendidos en una cierta substancia realizaban movimientos irregulares relacionados con las múltiples colisiones aleatorias de las moléculas del líquido con los granos del polen. Si bien esto definió el fenómeno físico, el modelo matemático que demostró la existencia de tal proceso¹⁸ fue desarrollado por Norbert Wiener en 1923. Anterior a esto, ya Louis Bachelier había trabajado en su tesis Teoría de la Especulación en 1900 asociando los precios de las acciones con movimientos brownianos.

A pesar de lo anterior, el movimiento browniano solo obtuvo una gran relevancia con los trabajos de Osborne (1959), Osborne (1962) y del premio Nobel Paul Samuelson (1965). El primero consideró los precios de las acciones y más específicamente los logaritmos de los mismos como un ensamble de decisiones en equilibrio estadístico con propiedades análogas a un ensamble de partículas que siguen un movimiento browniano.

El segundo introdujo el término de geométrico al movimiento browniano con la propiedad que cada dólar del mercado de valores está sujeto a las fluctuaciones multiplicativas o porcentuales por unidad de tiempo sin importar el valor absoluto del precio de la acción. Lo anterior conlleva a un proceso log-normal para el cual el valor de la garantía tiene dos propiedades deseadas: Para

¹⁷ La media y la varianza de una serie de tiempo son débilmente estacionarias cuando estas son constantes y su covarianza es invariante en el tiempo.

¹⁸ Un proceso estocástico $\{B_t = W_t: t \geq 0\}$ con $B_0 = 0$, con trayectoria continuas, con incrementos independientes y con cada variable B_t distribuyéndose como una $N(0, t)$.

tiempos cortos, una buena aproximación es \sqrt{t} y para $t \rightarrow \infty$, el valor se acerca al valor de la acción.

El movimiento browniano es uno de los modelos más utilizados para describir el precio de una acción con cuatro supuestos: Los cambios en el precio son estacionarios, son independientes (no correlación con cambios anteriores), son continuos y se distribuyen con una $N(0, t)$. A pesar de lo anterior, los datos empíricos no se ajustan a los supuestos anteriores y en el caso del último, la distribución difiere de la distribución normal (Muñoz San Miguel, 2005).

En general, el movimiento browniano geométrico se expresa como:

$$S_t = S_0 e^{\left(\left(\mu - \frac{\sigma^2}{2}\right)t + W_t \sigma\right)} \quad (6)$$

Donde S_0 es el valor inicial, μ el desplazamiento medio, σ la volatilidad, t el tiempo de la observación que se desea calcular y W_t es un proceso de movimiento browniano. Con este modelo suelen realizarse las simulaciones de n trayectorias o simulaciones de Montecarlo¹⁹.

¹⁹ El método de Montecarlo es un método no determinista usado para aproximar expresiones matemáticas complejas mediante el uso la simulación de variables aleatorias que pueden tomar múltiples valores. El método aprovecha la teoría del límite central al tener un error absoluto de la estimación que decrece como $\frac{1}{\sqrt{N}}$.

4. MEDIDAS Y *TEST* DE EVALUACIÓN DE PRONÓSTICOS

En este capítulo se revisarán desde la teoría un conjunto de medidas de evaluación de pronósticos y *test* de evaluación con el objetivo de conducir al lector en el mundo de la comprobación de la exactitud de los pronósticos. Al inicio del mismo se presenta una visión general de los principales criterios estadísticos; se continua con una breve explicación de las funciones de pérdida; posteriormente la clasificación, definición teórica, características, ventajas y desventajas de las medidas de precisión y por último, la explicación de los *test* diferenciales de comparación de pronósticos.

4.1. Principales criterios de aceptabilidad estadística

Es claro que la estadística económica se relaciona, entre otros, con la recopilación, procesamiento y presentación de cifras económicas por lo general en forma de gráficos y tablas. Éste es el trabajo del estadístico económico, cuya actividad principal consiste en recopilar cifras relacionadas con el comportamiento de algunos fundamentales como el producto nacional bruto (PNB), empleo, desempleo, precios, etc (Gujarati & Porter, 2010).

Los datos así reunidos constituyen la materia prima del trabajo econométrico. Pero el estadístico económico no va más allá de la recolección de información, pues no le conciernen las cifras recopiladas para probar las teorías económicas. Sin duda, es el econometrista quien se ocupa de realizar esta labor (Gujarati & Porter, 2010).

De igual manera, todo dato o grupo de datos obtenido, antes de ser totalizado y utilizado requiere un examen crítico, sobre aspectos de exactitud, precisión y representatividad (Vidal Tovar, 2013).

Por lo anterior, vale la pena señalar a continuación algunos de los principales criterios de aceptabilidad estadística que esta investigación tuvo en cuenta como requisitos para la selección de la serie (TCRM) que luego fue pronosticada.

Makridakis e Hibn (1995) refiere dos criterios de vital importancia:

- a) **Confiabilidad:** Es la habilidad de cada medida para generar resultados similares cuando se trabaja con muestras diferentes de la misma serie. Entre menor sea la variación mejor se puede catalogar la medida al no estar influenciada por cada muestra, por valores extremos o atípicos u otras características propias de cada serie.
- b) **Discriminación:** Las medidas deben ser capaces de discriminar entre los diferentes modelos para un mismo conjunto de datos de tal manera que indiquen cuales son los más y menos apropiados.

No menos importantes, se observan los siguientes criterios, de los cuales Chavarría Izos (2011) advierte exactitud y precisión:

- c) **Exactitud:** Falta de error sistemático, grado en que los resultados medios de varias determinaciones se aproximan al valor verdadero, en particular cuando se aproximan a la concentración en la que podría adoptarse una medida o, en el caso de una encuesta, por encima del intervalo de las concentraciones previstas. La exactitud está relacionada con el sesgo (desviación, inclinación) de una estimación, cuanto menor es el sesgo más exacta es una estimación.
- d) **Precisión:** Se suele confundir con la anterior, sin embargo la precisión se refiere a la dispersión del conjunto de valores obtenidos de mediciones repetidas de una magnitud. Cuanto menor es la dispersión mayor la precisión. Una medida común de la variabilidad es la desviación estándar de las mediciones y la precisión se puede estimar como una función de ella. Aunque exactitud y precisión son bastante parecidas, sus definiciones difieren en el hecho de que una tiene que ver con la cercanía al valor real y la otra se refiere a dar el mismo

resultado en distintas mediciones; todo esto nos lleva a deducir que se puede ser exacto mas no preciso y viceversa.

- e) Validación: Se refiere al hecho de que una prueba sea de tal manera concebida, elaborada y aplicada y que mida lo que se propone medir. En pocas palabras, se refiere al significado de la medida como cierta y precisa. Baechle y Earle (2007) citados por Ruiz (2011), señalan que la validez es el grado en que una prueba o ítem de la prueba mide lo que pretende medir, es la característica más importante de una prueba. Al referirse a la validez relativa a un criterio definen a éste como la medida en que los resultados de la prueba se asocian con alguna otra medida de la misma aptitud.

Consideran los autores igualmente, que en muchas ocasiones la validez relativa a un criterio se estima en forma estadística utilizando el coeficiente de correlación de Pearson (también denominado tabulación cruzada, a este tipo de validez se le denomina validez concurrente).

- f) Ajuste: Describe lo cercano o no que un modelo estadístico calculado se alinea con una muestra de datos. Las medidas de ajuste normalmente calculan la diferencia entre los valores observados o reales y aquellos que resultan después de la formulación de un modelo.

4.2. Los criterios de información y selección

Como se ha mencionado, el proceso de pronóstico en el análisis econométrico involucra aspectos relacionados con la estadística misma así como otros que dependen de la variable, de los objetivos de análisis, del entorno y del contexto en general que se salen de la realidad matemática y estadística como tal.

Sin embargo, cuando se inicia este proceso y se escogen diferentes mecanismos para pronosticar una misma variable endógena, el resultado final deriva en una serie de modelos que hacen un bosquejo matemático de la variable correspondiente. La pregunta que surge aquí es: ¿Cómo elegir o bajo qué criterios seleccionar un modelo sobre otro?

Varios autores han tratado de responder la pregunta anterior desarrollando diferentes estadísticos que pretenden medir la calidad relativa de un modelo para un conjunto determinado de datos: El punto importante está por lo tanto, en evaluar la calidad del modelo de pronóstico cuando se quiere hacer una predicción de una observación en el futuro.

Es aquí donde se involucran los criterios de información, que son precisamente aquellos estadísticos “(...) que fueron construidos como estimadores aproximadamente insesgados de la log-verosimilitud esperada, o, equivalentemente, de la discrepancia de la Información de Kullback-Leibler²⁰ entre la verdadera función $g(z)$ y un modelo estadístico desde un punto de vista predictivo” (Caballero, 2011, pág. 49). Dentro de los criterios de información que sobresalen están:

- Criterio R^2 y \bar{R}^2 ajustado
- Criterio de la función de verosimilitud y del logaritmo de la función de verosimilitud
- Suma de Cuadrados Residuales
- Criterio de Akaike (AIC)
- Criterio de Schwartz (BIC)
- Criterio de Hannan-Quinn
- Criterio de Bozdogan

Para efectos del presente documento, se detallarán R^2 , \bar{R}^2 ajustado, AIC, BIC y Hannan-Quinn en las secciones siguientes.

4.2.1. Criterio de R^2 y \bar{R}^2 ajustado

Se denominan también coeficientes de determinación, bien sea el r^2 (caso de dos variables) o R^2 (regresión múltiple). El criterio está definido por:

²⁰ La información de Kullback-Leibler es básicamente una medida de la distancia entre la densidad de probabilidad generada por el modelo que está siendo desarrollado y la realidad. Fue introducida por Kullback y Leibler (1951) y establece que la información entre el proceso real $g(z)$ y el modelo aproximado $f(z|\theta_k)$ con respecto a $G(z)$ será:

$$l(\theta_k) = E \left\{ \ln \frac{g(z)}{f(z|\theta_k)} \right\}, \text{ donde } E \text{ corresponde al valor esperado.}$$

$$R^2 = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT} \quad (7)$$

Donde SCE es la Suma de Cuadrados Explicados (calculado con la regresión), SCR es la Suma de Cuadrados de los Residuos y SCT es la Suma de los Cuadrados Total.

La cantidad R^2 se conoce como coeficiente de determinación muestral múltiple y es la medida más común de la bondad del ajuste de una línea de regresión. Igualmente es una estadística descriptiva que calcula la proporción de la “varianza” de la variable dependiente explicada por las variables regresoras, que bajo las asunciones del modelo clásico lineal Gaussiano²¹, podría ser un *test* de significancia (Dufour, 2011). Pueden observarse dos propiedades iniciales de R^2 :

- Es una cantidad no negativa.
- Sus límites son $0 \leq R^2 \leq 1$.

Los inconvenientes que presenta el criterio son:

- Mide bien la bondad dentro de la muestra pero no hay garantías de hacerlo fuera de ella.
- Al comparar dos o más valores, la variable dependiente debe ser la misma.
- R^2 no disminuye cuando se agregan más variables.

Precisamente por el último punto, Henry Theil (1961) desarrolló la R^2 ajustada la cual ya incluye una penalización cuando se agregan más variables dependientes. En este sentido $\bar{R}^2 \leq R^2$ y es una mejor medida a pesar que la variable a explicar debe ser la misma. El criterio está definido por:

$$R^2_{ajustada} = \bar{R}^2 = 1 - (1 - R^2) \frac{n-1}{n-k} \quad (8)$$

Donde k es el número de parámetros y n el tamaño muestral.

²¹ El modelo clásico de Gauss hace referencia a un modelo correctamente especificado, lineal en los parámetros, media condicional cero, con homocedasticidad, sin correlación en las perturbaciones, sin covarianza entre los errores y las variables regresoras y donde el número de observaciones es mayor al número de parámetros.

4.2.2. Akaike

El criterio de Akaike es considerado como el primer criterio de selección de modelos y hoy es ampliamente usado en la mayoría de herramientas estadísticas. Fue desarrollado por Hirotugu Akaike (1974), quien extendió el principio de la máxima probabilidad de estimación de parámetros desconocidos y dimensión y estructura definidas, a un mecanismo metodológico donde ambos, los parámetros y la dimensión del modelo son no conocidos y deben por lo tanto ser determinados del conjunto de datos. Al final la estimación y selección pueden ser desarrolladas simultáneamente.

El marco conceptual parte de un conjunto de datos z que son generados por un proceso no conocido $g(z)$ y sobre el cual se busca un modelo basado en parámetros que provea una aproximación cercana. Se define entonces $\mathcal{F}(k) = \{f(z|\theta_k) \mid \theta_k \in \Theta(k)\}$ como una función de densidad paramétrica en k dimensiones (k parámetros) en la cual el espacio de parámetros $\Theta(k)$ consiste en vectores de k dimensiones cuyos componentes son funcionalmente independientes. Así mismo, $\hat{\theta}_k$ denota un vector estimado obtenido mediante la función de máxima probabilidad o función de verosimilitud de las observaciones $f(z|\theta_k)$ sobre $\Theta(k)$. El modelo ajustado será entonces $f(z|\hat{\theta}_k)$.

El objetivo es entonces, de un conjunto diferente de funciones $\mathcal{F} = \{\mathcal{F}(k_1), \mathcal{F}(k_2), \dots, \mathcal{F}(k_L)\}$ para el modelo ajustado, definir cuál es la mejor aproximación de $g(z)$. El desarrollo matemático realizado por el autor lleva a que el criterio de información de Akaike (AIC) sea un estimador muestral de $E[\ln f(z|\theta_k)]$, es decir, la esperanza de la log-verosimilitud y que finalmente viene dado por la siguiente expresión:

$$AIC(k) = -2 \ln f(z|\hat{\theta}_k) + 2k = -2 \ln(\text{máxima probabilidad}) + 2k \quad (9)$$

Cuando hay diferentes especificaciones de $f(z|\hat{\theta}_k)$ correspondientes a diferentes modelos, se tomará aquella $f(z|\hat{\theta}_k)$ que arroje el valor mínimo de $AIC(k)$. Así mismo, cuando la probabilidad máxima es la misma para dos modelos se toma aquel con el menor número de parámetros involucrados.

Consecuentemente, la idea “(...) subyacente al uso del AIC para la selección de modelos es la maximización de la log-verosimilitud esperada de un modelo determinado” (Caballero, 2011, pág. 52) o lo que es similar, minimizar la pérdida entre los datos reales y el modelo a ser evaluado. Por lo anterior, el primer término de la ecuación podría ser considerado como la bondad del ajuste y el segundo la penalización conforme incrementa el número de parámetros (conforme al Principio de Parsimonia).

En este sentido, el criterio AIC impone una penalización por añadir regresoras o variables independientes. Sin embargo, es importante tener en cuenta que este criterio no busca el modelo correcto, ya que parte de la premisa que el modelo verdadero puede no estar dentro del conjunto de modelos a evaluar y por ende, su objetivo es seleccionar el modelo que proporcione las mejores predicciones (Peña, 2002). Por último, AIC puede utilizarse para el desempeño de predicciones dentro de la muestra, para predicciones fuera de la muestra, para modelos anidados²² y no anidados (Gujarati & Porter, 2010).

4.2.3. Schwarz

Este criterio fue introducido por Gideon Schwarz (1978) quien plantea una metodología opcional al de Akaike para la selección de los modelos ya que este último podría no ser asintóticamente justificable (Stone, 1979). Su propuesta se basa en el cálculo de los estimadores de máxima probabilidad a través de los límites de los estimadores de Bayes²³. En este sentido, el autor busca la apropiada modificación de la máxima probabilidad a través del comportamiento asintótico de los estimadores de Bayes (de allí precisamente el nombre de BIC – *Bayesian Information Criterion*).

²² Se dice que un modelo A está anidado en el modelo B porque es un caso especial del modelo B, es decir, A se reduce a B.

²³ Conceptualmente los problemas de estimación en el enfoque bayesiano se resuelven mediante los principios de cálculo de probabilidades incorporando información a priori. Esta fortaleza se convierte en su debilidad porque exige representar la información inicial mediante una distribución a priori que según los contradictores puede incluir los prejuicios del investigador (Peña, 2002).

Si se parte entonces del modelo de Akaike, se puede observar que el segundo término, correspondiente a $2k$ y que como se mencionó es la penalización por el número de parámetros, no depende del tamaño muestral. Lo anterior conlleva a que para casos con muestras pequeñas y grandes con el mismo número de parámetros no hay diferenciación mediante este criterio.

Precisamente por esto se dice que el AIC no es un estimador consistente del número adecuado de factores comunes (Caballero, 2011). El BIC entonces penaliza el número de parámetros con $\ln n$ en lugar de 2 como lo hace Akaike. El criterio será entonces definido por:

$$BIC(k) = -2 \ln f(z|\hat{\theta}_k) + (\ln n)k = -2\ln(\text{máxima probabilidad}) + (\ln n)k \quad (10)$$

Donde $f(z|\hat{\theta}_k)$ es la función de verosimilitud de las observaciones, $\hat{\theta}_k$ la estimación máximo verosímil, k el número de parámetros y n el tamaño muestral.

Bajo la definición de BIC anterior, entre un conjunto de modelos se prefiere aquel que resulte con una menor cantidad en (10). Con lo anterior, el criterio BIC igualmente escogerá modelos más parsimoniosos, es decir, aquellos que dispongan una menor cantidad de parámetros. “El criterio BIC trata de seleccionar el modelo correcto, con máxima probabilidad a posteriori, y puede demostrarse que es un criterio consistente, de manera que la probabilidad de seleccionar el modelo correcto tiende a uno cuando crece el tamaño muestral” (Peña, 2002, pág. 350).

4.2.4. Hannan y Quinn

El criterio HQIC fue desarrollado por E. J. Hannan y B. G. Quinn (1979) quienes desarrollaron un procedimiento para la estimación del orden de una autoregresión a partir de las iteración logarítmica de las autocorrelaciones parciales. Según Caballero (2011), el desarrollo realizado por los autores podría ser una variación del BIC en la medida que introduce una penalización adicional de la magnitud del tamaño muestral. Por su parte Cavanaugh (2012) menciona que es un criterio intermedio que busca la consistencia de un criterio como el BIC y la eficiencia asintótica de uno como el AIC.

El criterio es definido por:

$$HQIC(k) = -2 \ln f(z|\hat{\theta}_k) + 2 \ln((\ln n))k$$

$$HQIC(k) = -2 \ln(\text{máxima probabilidad}) + 2 \ln((\ln n))k \quad (11)$$

Donde $f(z|\hat{\theta}_k)$ es la función de verosimilitud de las observaciones, $\hat{\theta}_k$ la estimación máximo verosímil, k el número de parámetros y n el tamaño muestral.

En este caso el término de penalización es precedido por una constante $c=2$ que incluso podría ser mayor que el valor asignado (Burnham & Anderson, 2002). Estos autores igualmente anotaron que el criterio HQIC al igual que el BIC, pero a diferencia del AIC, no realiza una estimación de la divergencia de Kullback-Leibler.

Si se analiza entonces el término de penalización, el cual contiene la iteración logarítmica $(2 \ln((\ln n))k)$, se encuentra que este provee una función que disminuye tan rápido como es posible para una estimación consistente en la medida que el tamaño de la muestra aumenta (Khim Sen & Shitan, 2000). Como igual se adiciona penalización, este término corrige el desempeño del criterio BIC cuando las muestras son pequeñas.

4.3. Funciones de pérdida

Las funciones de pérdida se constituyen en un elemento fundamental en la teoría de optimización de problemas en varias ciencias: Teoría de decisión (donde juega un gran papel la Teoría Bayesiana propuesta por el Ingles Thomas Bayes en el siglo XVIII), estimación, pronóstico, inversión financiera y en la econometría entre otros (Lee, 2007). La definición más sencilla de una función de pérdida podría darse de la siguiente manera: Dada una predicción \hat{y}_t y un nivel real y_t , la función de pérdida $L_{t+1}(y_{t+1}, \hat{y}_t)$ mide la diferencia o discrepancia entre el algoritmo o modelo de predicción y la salida deseada.

Diferentes criterios de estimación podrían reflejarse de la utilización de diferentes funciones de pérdida tal como lo menciona Tussel (2007) ya que de éstas depende la estimación misma de los parámetros del modelo. Se encuentra en la teoría diferentes funciones de pérdida para la evaluación del conjunto de pares de datos. Algunos ejemplos de las funciones clásicas son las cuadráticas, las de valor absoluto, las de cuantiles y la 0-1. Cada una de estas funciones se soportan en un estadístico (la media, la media al cuadrado o la mediana misma), el cual se convierte en un criterio importante para la decisión del modelo de estimación.

Por otro lado, dependiendo de cómo tratan las magnitudes de los errores y el signo de los mismos, las funciones de pérdida podrían ser clasificadas como simétricas y asimétricas. Las más populares según Elliot y Timmermann (2007) son las simétricas pero con su contraparte, aunque requieren una construcción matemática más compleja, se puede lograr una medición más real desde el punto de vista económico y no matemático: Por ejemplo al considerar que errores positivos y negativos de la misma magnitud no tengan la misma pérdida.

De las funciones de pérdida surgen todas las medidas de evaluación de pronósticos que se explicarán en este mismo capítulo puesto que, cada medida utiliza como tal una función. Son precisamente las medidas las que condicionan el cálculo matemático de la función misma para la evaluación de modelo de pronóstico. En la Tabla 2 se presentan algunas funciones de pérdida con su formulación matemática correspondiente.

Tabla 2. Funciones de pérdida y definiciones matemáticas

Función de pérdida	Definición matemática
Error al cuadrado	$L_{t+1}(y_{t+1}, \hat{y}_t) = (y_{t+1} - \hat{y}_t)^2$
Error absoluto	$L_{t+1}(y_{t+1}, \hat{y}_t) = y_{t+1} - \hat{y}_t $
Orden α asimétrica (lin-lin)	$L_{t+1}(y_{t+1}, \hat{y}_t) = (\alpha - 1)(y_{t+1} - \hat{y}_t < 0)(y_{t+1} - \hat{y}_t)$ para $\alpha \in (0,1)$
Linex	$L_{t+1}(y_{t+1}, \hat{y}_t) = \exp(a(y_{t+1} - \hat{y}_t)) - a(y_{t+1} - \hat{y}_t) - 1, a \in \mathbb{R}$
Cambio de dirección	$L_{t+1}(y_{t+1}, \hat{y}_t) = \begin{cases} 1, & \text{signo}(y_{t+1} - \hat{y}_t) \neq \text{signo}(\hat{y}_t - y_t) \\ 0, & \text{signo}(y_{t+1} - \hat{y}_t) = \text{signo}(\hat{y}_t - y_t) \end{cases}$
Predictiva log-likelihood	$L_{t+1}(y_{t+1}, \hat{y}_t) = \log \hat{y}_t(y_{t+1})$ donde \hat{y}_t es el pronóstico de densidad de y_{t+1} .

4.4. Las medidas de precisión o de evaluación de pronósticos

4.4.1. Clasificación

En la Tabla 3 se presenta las medidas de evaluación de pronósticos clasificados por el tipo de medida y su evaluación. Las medidas pueden ser absolutas si calculan el error sin importar el signo, relativas a una base si utilizan un valor para compararse y relativas al tamaño de los errores si arrojan un valor porcentual en sus cálculos. Así mismo, pueden clasificarse como de método simple si el valor obtenido de la medida ya indica la asertividad del pronóstico, de más de un método si se requieren dos o más pronósticos para su cálculo o de comparación si utiliza dentro de su cálculo un modelo de referencia.

Tabla 3. Clasificación de las medidas de evaluación

		Tipo de Evaluación		
		Método Simple	Más de un Método	Comparación con un modelo de referencia
Tipo de Medida	Absoluta	MSE MAE RMSE RMdSE		
	Relativa a una base u método			U de Theil MRAE MdRAE GMRAE MSSE RMSSE MASE MdASE
	Relativa al tamaño de los errores	MAPE MdAPE RMSPE RMdSPE	MAPE MdAPE RMSPE RMdSPE	sMAPE sMdAPE

Fuente: Traducido y adaptado de Makridakis y Hibon (1995)

En este punto es importante mencionar un estudio que realizaron Carbone y Armstrong (1982) en el Primer Simposio International de Pronóstico (*First International Symposium on Forecasting*) realizado en Quebec, Canadá, en el cual midieron cuales eran los criterios preferidos para escoger una medida de evaluación de pronóstico. Este ejercicio se realizó con un universo de 70 académicos y 75 profesionales. La Tabla 4 tabula el número de veces que un criterio de los

presentados fue reportado por los participantes. En adición, para el caso específico del criterio de exactitud, igualmente se presenta la medida de evaluación que fue indicada.

Tabla 4. Criterios de evaluación y su importancia determinada por académicos y profesionales

Criterio	Académicos	Profesionales
	70	75
Exactitud		
R ²		2
MSE	30	20
GMSE ²⁴	1	
Varianza Mínima	2	4
U Theil	3	1
MPE ²⁵	5	5
MAE	12	14
MAPE	17	7
MMAE (Minimax AE) ²⁶	2	
Errores de pronóstico aleatorio	1	2
Sin medida específica	8	14
Facilidad en interpretación	26	29
Costo/Tiempo	24	25
Facilidad en uso e implementación	26	18
Adaptación a nuevas condiciones	10	13
Universalidad	3	10
Captura de puntos de inflexión	5	6
Robustez	10	3
Incorporación de entradas de juicio	4	2

Fuente: Traducido de Carbone y Armstrong (1982)

De la tabla se puede concluir que los criterios de exactitud, facilidad en interpretación, costo/tiempo y facilidad uso e implementación prevalecen como puntos de referencia a la hora de evaluar una medida de evaluación. Un detalle adicional que los autores encontraron es la similitud en proporciones en el grupo de académicos con el de profesionales.

²⁴ GMSE o *Geometric Mean Squared Error* hace referencia a la media geométrica de los errores al cuadrado.

²⁵ MPE o *Median Percentage Error* es similar a MAPE dada en (16) pero tiene en cuenta los signos (no incluye el valor absoluto)

²⁶ MMAE o *MiniMax Absolute Error* es una medida que busca minimizar la máxima pérdida o error.

4.4.2. Medidas dependientes de la escala

4.4.2.1. MSE (Mean Square Error)

Se define MSE (*Mean Square Error*) como la media de e_t^2 , es decir el promedio de los errores entre el estimador y lo que se estima al cuadrado:

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad (12)$$

Donde n es el número de muestras y \hat{y}_t es la estimación de y_t . De la fórmula anterior se deduce que la función de pérdida de la medida es la cuadrática o error medio al cuadrado.

Algunas características

- EL MSE es el segundo momento del error y por lo tanto relaciona la varianza del estimador así como su sesgo: Entre más pequeño es el MSE más se ajusta el estimador a los datos reales.
- El elevar las diferencias al cuadrado hace que los datos negativos y positivos no se cancelen entre sí. Adicionalmente, al ser de naturaleza cuadrática, otorga mayor peso a aquellos errores de mayor tamaño siendo útil en los casos en los que el pronóstico implica una valoración de los errores de esta manera pero no en aquellos donde no se dé relevancia a lo anterior.
- Las unidades del MSE no son las mismas de los datos ni del estimador.
- Toma valores entre 0 y ∞ .

Desventajas

Hay dos desventajas básicas en el uso de esta medida para la evaluación de pronósticos:

- Por su fórmula en sí, es una medida sensible a valores estimados atípicos quizás con poca frecuencia de ocurrencia. Esto arrojará un valor superior en la diferencia con el dato real que se incrementará con la potencia cuadrada y por ende, tendrá un peso en el cálculo del promedio haciéndola poco confiable.

- Por la diferencia de unidades en relación con los datos originales es difícil de interpretar.

Adicionalmente, por ser una medida que estadísticamente trabaja en el segundo momento, se hace necesario la comparación con la varianza: Ashley (1983) desarrolló un teorema mediante el cual demuestra que si el MSE es mayor que la varianza de la variable explicativa, la inclusión de ésta en un modelo de pronóstico arrojará peores resultados que si la omite. Así mismo encontró que es muy usual encontrar pronósticos de variables macroeconómicas con MSE que exceden su varianza (Ashley R. , 1988).

4.4.2.2. RMSE (Root Mean Square Error)

Se define RMSE (*Root Mean Square Error*) como la raíz cuadrada de la media de los errores al cuadrado:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (13)$$

Donde n es el número de muestras, \hat{y}_t corresponde a los valores observados de y_t y y_t se modela en valores en tiempo o lugar t . La función de pérdida de la medida es la cuadrática o error medio al cuadrado.

Algunas características

El error cuadrático medio (RMSE) también llamado desviación cuadrática media es una medida de uso frecuente de la diferencia entre los valores pronosticados por un modelo y los valores realmente observados. Estas diferencias individuales son también llamadas residuos y el RMSE sirve para agregar en una sola medida la capacidad de predicción.

El RMSE es más apropiado que el MAE para representar el rendimiento del modelo cuando se espera que la distribución de error sea gaussiano acorde con Chai y Draxler (2014). La diferencia con la medida MSE es que el resultado está en las unidades originales de la información histórica.

La medida es buena cuando se quiere impresionar a la gente (algunos analistas prefieren un medición compleja a una simple). Otra razón para el uso de la RMSE surge cuando hay una función de pérdida cuadrática; en otras palabras, cuando el costo asociado con un error aumenta como el cuadrado del error.

Desventajas

- Se debería evitar el uso de del RMSE cuando la evaluación de error es tan crudo que hay valores atípicos (grandes errores de medición). Estos valores tendrán un fuerte efecto en la medida en que se elevan al cuadrado. Al igual que la MSE, RMSE es una herramienta muy popular por su relevancia teórica en la modelación estadística, sin embargo, varios autores las han rechazado precisamente por su sensibilidad a valores atípicos, entre ellos Armstrong (2001).
- Willmott y Matsuura (2005) han sugerido que el RMSE no es un buen indicador de promedio del rendimiento del modelo y podría ser un indicador engañoso de error promedio.
- Un problema asociado con el uso de la medida RMSE o similares es el hecho de que la previsión de la varianza de error varía a través del tiempo. Puede variar a causa de no linealidades en el modelo y debido a la variación en las variables exógenas si se incluyen en el modelo (Woschnagg & Cipan, 2004).

4.4.2.3. MAE (Mean Absolute Error)

Se define MAE (*Mean Absolute Error*) como la magnitud promedio de los errores de un ejercicio de pronóstico sin tener en cuenta su signo, es decir, el promedio de los valores absolutos de los errores calculados:

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (14)$$

Donde n es el número de muestras y \hat{y}_t es la estimación de y_t . La función de pérdida de la medida es la del error absoluto.

Algunas características

- MAE arroja un número que puede ser directamente interpretado puesto que la pérdida se encuentra en las mismas unidades de la variable de salida.
- El peso asignado a cada uno de los errores calculados es el mismo.
- El uso de esta medida es aconsejable para describir errores uniformemente distribuidos.
- Por ser una medida natural del error como llaman a esta métrica Willmott y Matsuura (2005), hay otra métricas, por ejemplo la RMSE, que se convierten funciones de MAE.
- El hecho mismo de ser una medida absoluta podría convertirse en una desventaja, sin embargo, por no ser de naturaleza cuadrática, recibe menos influencia de valores atípicos.
- Toma valores entre 0 y ∞ .

Desventajas

- Puesto que la mayoría de modelos tienen una asunción en el sentido que los errores se encuentran distribuidos como una normal y no como una función uniforme, es posible encontrar métricas que se ajustan de una mejor manera a los datos: Este es el caso por ejemplo de RMSE según Chai y Draxler (2014).
- Un importante aspecto en las métricas de error es su capacidad de discriminar entre los resultados del modelo. Cuando se utiliza MAE, los resultados podrían estar afectados por una gran cantidad de valores de error promedio sin necesariamente reflejas errores de gran tamaño.
- Por la inclusión de valores absolutos en su cálculo, el uso de MAE podría significar dificultades en los cálculos de la sensibilidad del valor de MAE con respecto a los parámetros del modelo en cuestión.

4.4.2.4. MDAE (Median Absolute Error)

Se define MDAE (*Median Absolute Error*) como la mediana de los errores de un ejercicio de pronóstico sin tener en cuenta su signo, es decir, la mediana de los valores absolutos de los errores calculados:

$$MdAE = \text{Mediana} (|y_t - \hat{y}_t|) \text{ para } t = 1 \dots n \quad (15)$$

Donde n es el número de muestras y \hat{y}_t es la estimación de y_t . La función de pérdida de la medida es la del error absoluto.

Es una medida de precisión dependiente a la escala de los datos. Es útil para comparar métodos del mismo conjunto de datos, pero no cuando éstos tienen diferente escala. MdAE no está influenciada por los valores extremos sin embargo, esta ventaja es también una debilidad pues no maximiza el uso de la información disponible sobre los errores, un rasgo que debe compartir toda medida "robusta" según Swanson, Tayman y Bryan (2011).

4.4.3. Medidas basadas en porcentajes

4.4.3.1. MAPE (Mean Absolute Percentage Error)

MAPE (*Mean Absolute Percentage Error*) mide el tamaño del error en términos porcentuales lo cual lo hace no dependiente de la escala. Se calcula como el promedio de los errores porcentuales sin tener en cuenta el signo:

$$MAPE^{27} = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (16)$$

Donde n es el número de muestras, y_t es el valor actual y \hat{y}_t es la estimación. La función de pérdida de la medida es la del error absoluto.

Algunas características

- El uso de esta métrica está asociado a series de tiempo homogéneas e igualmente espaciadas.

²⁷ Normalmente las medidas basadas en porcentajes se expresan en términos porcentuales. Es por esta razón que en algunos casos las fórmulas incluyen un número 100 como multiplicando. Para efectos del presente trabajo esta constante no se ha incluido.

- La métrica es sensible a la escala y no debe ser utilizada en casos donde hay un bajo nivel de datos.
- Puesto que está dado en términos porcentuales, es más fácil de entender que los valores arrojados por otras métricas y por ello, es común que MAPE sea usada para comparar diferentes modelos de pronóstico con conjuntos de datos diferentes.
- La métrica tiene propiedades estadísticas valiosas por cuanto usa todas las observaciones y tiene la más pequeña variabilidad de muestra a muestra (Levy & Lemeshow, 1991).

Desventajas

- Por la fórmula misma, si hay valores en cero, esto significará una división entre cero. A pesar que lo anterior se soluciona desestimando aquellos datos en cero, varios problemas podrían surgir cuando se tienen valores de y_t muy pequeños y \hat{y}_t grandes ya que esto redundaría en valores de MAPE extremadamente grandes que podrían hacer las comparaciones inútiles. De esta manera, casos en los cuales donde el valor cero es importante, por ejemplo pronósticos de temperaturas en grados Fahrenheit o Centígrados, este tipo de medidas podría no ser de las preferidas para la comparación de modelos.
- Un mismo valor de error puede significar un porcentaje diferente: Casos en que Si y_t es mayor que \hat{y}_t , esto arrojará un porcentaje más pequeño que cuando y_t es menor que \hat{y}_t : Tómese el caso en que y_t sea 150 y \hat{y}_t 100. El porcentaje de error absoluto es 33.33%. Si los valores se invierten el porcentaje será 50% (Makridakis & Hibon, 1995).
- Esta métrica no es robusta en el sentido que pocos valores atípicos pueden alejar su cálculo así como no se ajusta a varias distribuciones (Hoaglin, Mosteller, & Tukey, 1983).
- La validez de MAPE es cuestionable puesto que la distribución de porcentajes de valores absolutos es frecuentemente asimétrica y con sesgo a la derecha según el *National Research Council* (1980).
- Toma valores entre 0 y ∞ .

4.4.3.2. MdAPE (Median Absolute Percentage Error)

MdAPE (*Median Absolute Percentage Error*) mide el tamaño del error en términos porcentuales y se calcula como la mediana de los errores porcentuales sin tener en cuenta el signo:

$$MdAPE = Mediana \left(\left| \frac{y_t - \hat{y}_t}{y_t} \right| \right) \text{ para } t = 1 \dots n \quad (17)$$

Donde n es el número de muestras, y_t es el valor actual y \hat{y}_t es la estimación. La función de pérdida de la medida es la del error absoluto.

Algunas características

El porcentaje de error absoluto de la mediana es similar al MAPE (ya sea regular o simétrica) pero en vez de resumir los errores en un porcentaje absoluto (APE), calcula su promedio. Es decir, toda la APE están ordenadas desde la más pequeña a la más grande y el APE en el medio (para el caso, hay un número par de APE entonces el promedio de los dos medio se calcula) se utiliza para denotar la mediana.

Algunas de sus ventajas son:

- No está influenciada por los valores atípicos.
- No tienen el inconveniente de la escala, razón por la cual puede ser utilizada para diferentes conjuntos de datos.

Desventajas

- Las medidas basadas en porcentajes de error, por lo general, tienen la desventaja de ser infinitas o indefinidas si $y_t = 0$ para cualquier t en el período de interés, y que tiene una distribución extremadamente sesgada cuando cualquier y_t está cerca de cero. Esto significa, por ejemplo, que la MAPE es a menudo sustancialmente mayor que la MdAPE. (Hyndman, 2006).

- Su mayor desventaja es que su significado es menos intuitivo. El uso de la simétrica APE reduce las posibilidades de los valores atípicos y reduce la necesidad utilizar MdAPE. Además, es difícil combinar MdAPE a través de horizontes y / o series según y cuando la nueva base de datos de que se disponga (Makridakis S. , 1993).

4.4.3.3. RMSPE (Root Mean Square Percentage Error)

Se define RMSPE (*Root Mean Square Percentage Error*) como la raíz cuadrada del promedio de los errores en términos porcentuales al cuadrado:

$$RMSPE = \sqrt{\frac{1}{n} \sum_{t=1}^n \left(\frac{y_t - \hat{y}_t}{y_t} \right)^2} \quad (18)$$

Donde n es el número de muestras, y_t es el valor actual y \hat{y}_t es la estimación. La función de pérdida de la medida es la cuadrática o error al cuadrado.

Algunas características

- RMSPE tiene las mismas propiedades que RMSE con la diferencia que los resultados son expresados en porcentajes (Swanson, Tayman, & Bryan, 2011).

Desventajas

- Como las medidas porcentuales, RMSPE tiene la desventaja de tomar un valor infinito o indefinido cuando y_i es igual a cero o podría tener una distribución sesgada cuando y_i es muy cercano a cero. Por esto y al igual que otras medidas que arrojan porcentajes, no es común el uso de esta métrica en aquellos modelos donde los datos se presentan como pequeños conteos.
- Toma valores entre 0 y ∞ .

4.4.3.4. RMdSPE (Root Median Square Percentaje Error)

Se define RMdSPE (*Root Median Square Percentage Error*) como la raíz cuadrada de la mediana de los errores en términos porcentuales al cuadrado:

$$RMdSPE = \sqrt[2]{\text{Mediana} \left(\left(\frac{y_t - \hat{y}_t}{y_t} \right)^2 \right)} \text{ para } t = 1 \dots n \quad (19)$$

Donde n es el número de muestras, y_t es el valor actual y \hat{y}_t es la estimación. La función de pérdida de la medida es la cuadrática o error al cuadrado.

Esta medida tiene la ventaja de ser independiente de la escala de datos, por lo que se utilizan con frecuencia para comparar el rendimiento de previsión a través de diferentes conjuntos de datos.

4.4.3.5. sMAPE (Symmetric Mean Absolute Percentage Error)

sMAPE (*Symmetric Mean Absolute Percentage Error*) mide igualmente el tamaño del error en términos porcentuales (no dependiente de la escala). Se calcula como el promedio de los valores absolutos de los errores dividido entre los promedios de los valores reales y pronosticados:

$$sMAPE = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{(|y_t| + |\hat{y}_t|)/2} \quad (20)$$

Donde n es el número de muestras, y_t es el valor actual y \hat{y}_t es la estimación. La función de pérdida de la medida es la del error absoluto. Hay algunas variaciones a (20) entre las cuales se mencionan dos:

$$sMAPE = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{(y_t + \hat{y}_t)/2} \quad (21)$$

$$sMAPE = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{(y_t + \hat{y}_t)} \quad (22)$$

Mientras (21) permite la toma de valores negativos y positivos al igual que (22), este último adicionalmente limita a valores de 0 a 100%.

El MAPE simétrico fue diseñado para solucionar algunas de las limitaciones del MAPE. Las revisiones iniciales involucraron a Armstrong (1985), Flores (1986) y finalmente a Makridakis (1993).

Algunas características

- Como correcciones a la métrica MAPE; sMAPE busca que los valores extremos tengan una menor influencia así como contrarresta la asimetría que se genera con el término $|y_t - \hat{y}_t|$ cuando el pronóstico \hat{y}_t es mayor o menor que el dato real y_t (ver ejemplo presentado para la métrica MAPE): El ejemplo en mención arroja un resultado de 40% para ambos casos haciéndolos simétricos.
- Toma valores entre 0% y 200%.

Desventajas

- A pesar del tipo de simetría que corrige sMAPE y que se explicó en este mismo numeral, cuando la medición del error $|y_t - \hat{y}_t|$ es igual pero los valores de y_t y \hat{y}_t no se intercambian vuelve y se genera asimetría: Esto es lo que se llama la no corrección cuando para un mismo valor de y_t hay un pronóstico por encima (*over forecasting*) y uno por debajo (*under forecasting*). Para ejemplificar lo anterior, si y_t toma un valor de 100, sMAPE arrojará resultados de 4,76% y 5,26% para valores de \hat{y}_t de 110 y 90 respectivamente (nótese que $|y_t - \hat{y}_t|$ es igual para ambos casos): Hay una mayor penalización para pronósticos que arrojen un valor más alto (Hyndman & Koehler, 2005).
- Por los valores que toma, sMAPE es un poco más difícil de entender. Por lo anterior, si los problemas que genera MAPE no son tan graves en el caso que se estudia, se recomienda el uso de MAPE sobre sMAPE (Armstrong J. S., 1985).

4.4.3.6. SMdAPE (Symmetric Median Absolute Percentage Error)

SMdAPE (*Symmetric Median Absolute Percentage Error*) se define como la mediana de los valores absolutos de los errores dividido entre los promedios de los valores reales y pronosticados:

$$sMdAPE = \text{Mediana} \frac{|y_t - \hat{y}_t|}{(|y_t| + |\hat{y}_t|)/2} \text{ para } t = 1 \dots n \quad (23)$$

Donde n es el número de muestras, y_t es el valor actual y \hat{y}_t es la estimación. La función de pérdida de la medida es la del error absoluto.

Algunas características

- Al igual que sMAPE toma valores entre 0% y 200%
- Los problemas derivados de valores pequeños de y_t pueden ser menos severos para sMdAPE.
- Hyndman y Koehler (2005) hacen una breve explicación de esta medida y citan en su literatura a Makridakis (1993) quien señala que, tanto MAPE como MdAPE tienen la desventaja de que se castigan más los errores positivos que los negativos. Esta observación condujo a la utilización de las denominadas medidas “simétricas” para evitar lo anterior.

4.4.4. Medidas basadas en errores relativos

4.4.4.1. MRAE (Mean Relative Absolute Error)

Se define MRAE (*Mean Relative Absolute Error*) como la media aritmética de los errores absolutos de un modelo de pronóstico dividido entre los errores absolutos de un modelo de referencia, caminata aleatoria por lo general:

$$MRAE = \frac{1}{n} \sum_{t=1}^n RAE_t \quad (24)$$

$$\text{Donde } RAE_t = \frac{\frac{|y_t - \hat{y}_t|}{y_t}}{\frac{|y_t - FN_t|}{y_t}} \quad (25)$$

Donde n es el número de muestras, y_t es el valor actual, \hat{y}_t es la estimación del modelo y FN_t es el pronóstico seleccionado de comparación o referencia. La función de pérdida es la de error absoluto.

Algunas características

Como todas las medidas que incluyen el RAE, esta métrica es independiente de la escala que arroja un resultado sin unidades debido a la división como tal de los errores del modelo de pronóstico y de los errores del modelo de referencia.

Desventajas

Por su misma expresión matemática, el tener en el denominador el valor real de la serie que se está analizando, hace que para series en la cuales el cero es un número válido o en las que los valores muy bajos son parte de la serie se genere un cálculo erróneo o atípico que no permita la interpretación debida. Si bien hay autores que plantean correcciones como se verá en el numeral de la métrica GMRAE, esto mismo puede reflejar la arbitrariedad en el manejo de los datos (Prestwich, Rossi, Tarim, & Hnich, 2014).

4.4.4.2. MdRAE (Median Relative Absolute Error)

Se define MdRAE (*Median Relative Absolute Error*) como una medida de precisión que calcula la mediana de los errores absolutos de un modelo de pronóstico dividido entre los errores absolutos de un modelo de referencia. Se expresa como:

$$MdRAE = \text{Mediana } RAE_t \text{ para } t = 1 \dots n \quad (26)$$

Donde n es el número de muestras, y_t es el valor actual, \hat{y}_t es la estimación del modelo, FN_t es el pronóstico seleccionado de comparación y RAE_t es igual a (25). La función de pérdida es la de error absoluto.

El error absoluto relativo medio se encuentra al ordenar el error relativo absoluto (RAE) desde el valor más pequeño hasta el más grande y el uso de su valor medio (la media de los dos valores medios si m es un número par).

Al respecto, el MdRAE es similar a la MdAPE excepto que la expresión se utiliza para calcular el error utilizado en la búsqueda de la mediana en lugar de la APE. La ventaja de la MdRAE es que no está influenciada por los valores extremos, al tiempo que permite comparaciones con un método de referencia.

Desventajas

Aunque la medida MdRAE es altamente fiable, no discrimina lo suficiente para decir con confianza cual método de pronóstico es mejor que otro. Su desventaja, como la de la MdAPE, es que su significado no está claro, más que la de MdAPE.

4.4.4.3. GMRAE (Geometric Mean Relative Absolute Error)

Se define GMRAE (*Geometric Mean Relative Absolute Error*) como la media geométrica de los errores absolutos de un modelo de pronóstico dividido entre los errores absolutos de un modelo de referencia::

$$GMRAE = (\prod_t RAE_t)^{\frac{1}{n}} \quad (27)$$

Donde n es el número de muestras, y_t es el valor actual, \hat{y}_t es la estimación del modelo, FN_t es el pronóstico seleccionado de comparación y RAE_t es igual a (25). La función de pérdida es la de error absoluto.

Algunas características

- La ventaja de la media geométrica relativa es que no se ve tan afectada por valores atípicos y que es mucho más fácil de entender y comunicar que el estadístico U de Theil (Armstrong & Collopy, 1992). Sin embargo estos mismos autores recomiendan para evitar los valores atípicos: a) el descarte de los errores grandes y pequeños o b) el suavizar los extremos mediante la definición de límites superior e inferior, 10 y 0.01 respectivamente.
- La métrica GMRAE es utilizada para la calibración de los parámetros de un modelo y es útil para pronósticos con un limitado conjunto de series de datos.

- Los valores que calcula la métrica son independientes igualmente de la escala por cuanto el error del modelo seleccionado se divide por el error de un modelo de referencia.

Desventajas

- A pesar de las correcciones que plantearon Armstrong y Collopy (1992) del recorte de valores atípicos, la sola posibilidad de arrojar una varianza infinita cuando el término del denominador de la fórmula de RAE sea arbitrariamente pequeña, se convierte en una desventaja tal como lo sostienen Hyndman y Koehler (2005) y más aún en casos donde se trabajan series intermitentes que llevan a esta situación (Hyndman, 2006).
- Makridakis y Hibon (1995) clasifican a GMRAE como una métrica de poca o nula intuición en la información que resulta de sus cálculos.

4.4.5. Medidas con errores escalados

4.4.5.1. MSSE (Mean Square Scaled Error)

Se define MSSE (*Mean Square Scaled Error*) como la media de los errores al cuadrado escalados con un modelo de referencia:

$$MSSE = \frac{1}{n} \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\left(\frac{1}{n-1} \sum_{i=2}^n |y_i - y_{i-1}|\right)^2} = Media(q_t^2) = \frac{1}{n} \sum_{t=1}^n q_t^2 \quad (28)$$

Donde n es el número de muestras, y_t es el valor actual y \hat{y}_t es la estimación del modelo. El denominador de la fórmula anterior $\frac{1}{n-1} \sum_{i=2}^n |y_i - y_{i-1}|$, corresponde al modelo de referencia que para este caso es el modelo ingenuo. La función de pérdida para MSSE es la cuadrática o error al cuadrado. Esta métrica junto con las demás que escalan el error fue introducida por Hyndman (2006) y fue definida como una alternativa de uso general.

4.4.5.2. RMSSE (Root Mean Square Scaled Error)

Se define RMSSE (*Root Mean Square Scaled Error*) como la raíz de la media de los errores al cuadrado escalados con un modelo de referencia:

$$RMSSE = \sqrt{\frac{1}{n} \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\left(\frac{1}{n-1} \sum_{i=2}^n |y_i - y_{i-1}|\right)^2}} = \sqrt{\frac{1}{n} \sum_{t=1}^n q_t^2} \quad (29)$$

Donde n es el número de muestras, y_t es el valor actual y \hat{y}_t es la estimación del modelo. El denominador de la fórmula anterior $\frac{1}{n-1} \sum_{i=2}^n |y_i - y_{i-1}|$, corresponde al modelo de referencia que para este caso es el modelo ingenuo. La función de pérdida para RMSSE es la cuadrática.

Algunas características

- La métrica RMSSE es independiente de la escala por cuanto se toman los errores del pronóstico y se dividen por un modelo de referencia. Esta métrica es usada para comparar modelos de pronóstico para una misma serie de datos o modelos para diferentes series.
- Esta métrica se ajusta para series de demanda intermitente porque nunca arroja un valor infinito o indefinido excepto en los casos irrelevantes cuando los datos históricos son iguales.
- Hyndman y Koehler (2005) proponen esta métrica como un estándar para casos en los cuales se requiere la comparación de modelos de series de datos de diferentes escalas.

Desventajas

Por la complejidad de la fórmula, es más compleja la interpretación de su resultado en comparación con la propuesta que realizan los autores de la métrica MASE (*Mean Absolute Scaled Error*). Así mismo, ellos mencionan que RMSSE podría ser sensible a valores atípicos.

4.4.5.3. MASE (Mean Absolute Scaled Error)

Se define MASE (*Mean Absolute Scaled Error*) como la media absoluta de los errores escalados con un modelo de referencia. Hyndman y Koehler (2005), sostienen que es una medida de error

para pronósticos de diferente escala y que es la mejor medida disponible. Se expresa de la siguiente forma:

$$MASE = \frac{1}{n} \frac{\sum_{t=1}^n |y_t - \hat{y}_t|}{\left(\frac{1}{n-1} \sum_{i=2}^n |y_i - y_{i-1}|\right)} = \text{Media}|q_t| \quad (30)$$

Donde n es el número de muestras, y_t es el valor actual y \hat{y}_t es la estimación del modelo. El denominador de la fórmula anterior $\frac{1}{n-1} \sum_{i=2}^n |y_i - y_{i-1}|$, corresponde al modelo de referencia que para este caso es el modelo ingenuo. La función de pérdida para MASE es la del error absoluto.

Medidas como MdASE se puede definir de forma análoga, pero Billah (2005) usó una medida de error similar cuando se calcula el valor absoluto del error de pronóstico como porcentaje de la muestra en desviación estándar. Sin embargo, este enfoque tiene la desventaja de que el denominador crece con el tamaño de la muestra para la serie no estacionaria que contiene una raíz unitaria. Al escalar dentro de un modelo ingenuo, MAE sólo asume que la serie no tiene raíz por más de una unidad, que casi siempre es cierto para los datos reales.

Sin embargo cuando MASE es <1 el método propuesto da, sobre los errores de la media, más pequeño que el de un solo paso errores del método tradicional. Si se calculan previsiones de varios pasos, es posible escalar por la muestra en el MAE y calcula a partir de las previsiones ingenuas de varios pasos.

Algunas características

MASE se puede utilizar para series de datos debido a la ocurrencia de valores infinitos e indefinidos.

Esta medida también es fácilmente interpretable: Valores de MASE mayor que uno indican que las estimaciones son peores, en promedio, que dentro de la muestra de un solo paso por estimaciones del método o modelo ingenuo.

MASE es menos sensible a los valores atípicos y más fácil de interpretar que RMSSE.

4.4.5.4. MdASE (Median Absolute Scaled Error)

Se define MASE (*Median Absolute Scaled Error*) como la mediana de los errores absolutos escalados con un modelo de referencia:

$$MdASE = Mediana \frac{|y_t - \hat{y}_t|}{\left(\frac{1}{n-1} \sum_{i=2}^n |y_i - y_{i-1}|\right)} = Mediana |q_t| \text{ para } t = 1 \dots n \quad (31)$$

Donde n es el número de muestras, y_t es el valor actual y \hat{y}_t es la estimación del modelo. El denominador de la fórmula anterior $\frac{1}{n-1} \sum_{i=2}^n |y_i - y_{i-1}|$, corresponde al modelo de referencia que para este caso es el modelo ingenuo. La función de pérdida para MdASE es la del error absoluto.

Algunas características

Al igual que RMSSE, la métrica MdASE no depende de la escala por la división que se realiza entre los errores del pronóstico y los errores del modelo de referencia. De la misma manera, es usada para comparar modelos de pronóstico para una misma serie de datos o modelos para diferentes series.

Desventajas

Hyndman y Koehler (2005) mencionan que MdASE es una métrica variable para casos donde el número de muestras es reducido en comparación con otras métricas escaladas como MASE.

4.4.6. Estadístico U de Theil

Henri Theil propuso dos medidas de error en diferentes momentos. El estadístico U1 hace referencia a una medida de exactitud (Theil H. , 1958) mientras que el U2 hace referencia a una medida de calidad del pronóstico (Theil H. , Applied Economic Forecasting, 1966):

$$\text{Estadístico U1} = \frac{\sqrt{\sum_{t=1}^n (y_t - \hat{y}_t)^2}}{\sqrt{\sum_{t=1}^n y_t^2 + \sum_{t=1}^n \hat{y}_t^2}} \quad (32)$$

$$\text{Estadístico } U2 = \sqrt{\frac{\sum_{t=1}^n \frac{(y_t - \hat{y}_t)^2}{y_t} / n}{\sum_{t=1}^n \frac{(y_t - FN_t)^2}{y_t} / n}} \quad (33)$$

Donde n es el número de muestras, y_t es el valor actual, \hat{y}_t es la estimación del modelo y FN_t es el pronóstico seleccionado de comparación que suele ser el de caminata aleatoria o el modelo ingenuo o *naive*.

El estadístico U1 tiene algunos defectos con la calidad de información que presenta: Theil no especificó si los valores de y_t y \hat{y}_t deben ser usados como absolutos o como deltas de cambio. Bliemel (1973) interpretó el U1 para ambos y finalmente concluye que sin importar cuál sea el caso, el U1 tiene poco valor como un índice de la exactitud de un pronóstico: Sin importar si se trabaja con datos absolutos o deltas de cambio, los valores serán delimitados entre 0 (el mejor caso) y 1 (el peor caso).

El valor de 1 puede ser obtenido cuando el valor pronosticado es cero. Para todos los demás casos el estadístico arrojará un valor menor a 1 sin precisar si el modelo es mejor o no que uno de referencia. El estadístico U2 suple lo anterior y tiene defectos menores que el U1. Por lo anterior las características y desventajas que se presentan a continuación hacen referencia al estadístico U2.

Algunas características

El estadístico U2 toma valores entre 0 y ∞ . Un valor de 1 significa que la exactitud del modelo de pronóstico seleccionado es la misma que el modelo de referencia. Un valor por debajo de 1 significa que el modelo es mejor que el de referencia y un valor mayor a 1 significa lo opuesto: Se ha interpretado como la división del RMSE del modelo trabajado entre el RMSE del modelo de referencia.

Desventajas

- El estadístico U de Theil es altamente influenciado por los valores atípicos. En el límite inferior cuando el numerador de la expresión es cercano a cero, su cuadrado será aún más cercano a cero. Por el contrario en el límite superior, cuando y_t y FN_t tienden a cero, el estadístico tendrá una tendencia hacia infinito.
- No es obvio el significado de un valor de 0.85 en el estadístico y mucho menos cuanto es esto mejor que un valor de 0.82 (Makridakis & Hibon, 1995). La confusión final entre el U1 y el U2 al final son el resultado que estas métricas no son bien entendidas (Armstrong J. S., 1985).

4.4.7. Resumen y comparación de las medidas de precisión

4.4.7.1. Un resumen de las características

La Tabla 5 presenta un resumen con las características de las medidas descritas en los numerales anteriores. Si bien la clasificación puede resultar subjetiva para las dos últimas columnas es relevante incluir la posición del autor que inicialmente la desarrolló así como la nuestra en las medidas que se adicionaron. Es importante aclarar para la columna de ¿Fácil de entender? que la respuesta “moderado” significa que en algunos ámbitos podría serlo y en otros no.

Tabla 5. Características de las medidas de evaluación

Medida	Escala	¿Libre de unidades?	¿Función de pérdida simétrica?	¿Valida nivel y cambios?	¿Información relevante para tomar decisiones?	¿Fácil de entender?
MSE	Intervalo	No	Si	Si	Si	Moderado
RMSE	Intervalo	No	Si	Si	Si	Moderado
MAE	Intervalo	No	Si	Si	Si	Si
MdAE	Intervalo	No	Si	Si	No	Si
MAPE	Razón (%)	Si	No	Si	Si	Si
MdAPE	Razón (%)	Si	Si	Si	Si	Si
RMSPE	Razón (%)	Si	Si	Si	Moderado	Si
RMdSPE	Razón (%)	Si	Si	Si	Moderado	Si
Smape	Razón (%)	Si	Si	Si	Moderado	Si
sMdAPE	Razón (%)	Si	Si	Si	Moderado	Si
MRAE	Razón (%)	Si	No	Si	Moderado	No
MdRAE	Razón (%)	Si	No	Si	Moderado	No
GMRAE	Razón (%)	Si	No	Si	Moderado	No

Medida	Escala	¿Libre de unidades?	¿Función de pérdida simétrica?	¿Valida nivel y cambios?	¿Información relevante para tomar decisiones?	¿Fácil de entender?
MSSE	Razón (%)	Si	No	Si	Moderado	No
RMSSE	Razón (%)	Si	No	Si	Moderado	No
MASE	Razón (%)	Si	No	Si	Moderado	No
MdASE	Razón (%)	Si	No	Si	Moderado	No
U Theil	Intervalo	Si	Si	No	No	No

Fuente: Traducido y adaptado de Armstrong (1985)

4.4.7.2. Una comparación

Desde un punto de vista académico hay dos criterios básicos para la validación de las diferentes medidas de evaluación de pronósticos y con los cuales se han clasificado las mismas en la Tabla 6: El estadístico y el uso de los usuarios (Makridakis & Hibon, 1995). En el primer caso se evalúan la confiabilidad y la discriminación de modelos mientras que en el segundo, la información que entrega cada una así como la facilidad de entendimiento.

Si se empieza con la arista estadística, la confiabilidad es la habilidad de cada medida para generar resultados similares cuando se trabaja con muestras diferentes de la misma serie. Entre menor sea la variación mejor se puede catalogar la medida al no estar influenciada por cada muestra, por valores extremos o atípicos u otras características de cada propias de cada sub serie. Así mismo, las medidas deben ser capaces de discriminar entre los diferentes modelos para un mismo conjunto de datos de tal manera que indiquen cuales son los más y menos apropiados. Entre más altas sean las variaciones mejor será la medida puesto que indicará lo anterior.

En el extremo opuesto, es decir desde la arista de los usuarios, la información que entrega cada medida es única y dependiendo del contexto, servirán o no para tomar decisiones o definir políticas corporativas: Algunas entregan información para reportes, otras permiten hacer comparaciones entre métodos de pronósticos y un último grupo permiten ambos. Igualmente, los usuarios valoran la facilidad de entender la información que entregan las medidas para a partir de la misma, definir su uso.

Tabla 6. Valoración de las medidas de evaluación a partir de los criterios estadísticos y de usuario

	Criterio estadístico						
	Confiabilidad			Discriminación			
	Alto	Medio	Bajo	Alto	Medio	Bajo	
Información	Reportes		sMAPE	MSE			
				RMSE			
				MAPE			
				MAE			
				MdAE	MSE		
				MdAPE	MAPE	MdAE	MSSE
				RMSPE	MAE	RMSE	RMSSE
				RMdSPE	MdAPE	sMAPE	MASE
				MSSE			MdASE
				RMSSE			
				MASE			
				MdASE			
				Criterio de usuario	Hacer comparaciones	U Theil	MRAE
GMRAE	MdAPE	GMRAE	U Theil				
MdRAE	RMSPE	MAPE	sMAPE				
sMAPE	RMdSPE	MdAPE					
Sentido común		sMAPE	MAPE		MAPE	sMAPE	
Algún significado			MAE				
			MdAE				
			MdAPE		MAE	MdAE	
			RMSPE		MdAPE	MdRAE	
			RMdSPE				
Entendimiento	No intuitivo	U Theil	MRAE				
			GMRAE				
			MdRAE				
				MSE			
				RMSE		MSSE	
				MSSE		RMSSE	
				RMSSE	MSE	MASE	
				MASE		MdASE	
				MdASE			

Fuente: Traducido y adaptado de Makridakis y Hibon (1995)

4.5. Test de verificación de exactitud

Una de las herramientas básicas para la toma de decisiones en instituciones de gobierno, en compañías y en general en el mundo económico y financiero es la estimación de pronósticos que puedan darnos una idea del futuro con base en la información del pasado. Numerosos autores han desarrollado modelos precisamente para la generación de pronósticos. El mayor compendio de investigación de modelos podría ser el que empezó Makridakis y Hibon (1979) y que se convirtieron en las competiciones M1 (Makridakis, y otros, 1982), M2 (Makridakis, y otros, 1993) y M3 (Makridaki & Hibon, 2000).

El problema que surge entonces es el cómo seleccionar el mejor método de un conjunto de dos o más modelos de pronóstico. La respuesta a esta inquietud se fue dando con el desarrollo de *test* que comparan la habilidad de predecir de dos modelos dado una función de pérdida fija o que pueda ser modificada. Existe una amplia literatura al respecto entre los cuales sobresalen Granger y Newbold (1977); Diebold y Mariano (1995); Harvey, Leybourne y Newbold (1997) y; Giacomini y White (2006); . Al respecto, esta investigación repasará los planteamientos realizados por los autores anteriores. Adicionalmente se invita al lector a consultar otros complementarios tales como Meese y Rogoff (1988); West (1996) y; Clark y McCracken (2012).

4.5.1. *Test* MGN

Su nombre obedece a sus autores Granger y Newbold (1977), quienes en su trabajo toman los aportes realizados por Morgan (1939-40). Este *test* busca comprobar la hipótesis nula de que las varianzas de los errores de dos modelos son iguales y por ende su exactitud. Para su aplicación se deben considerar los supuestos que las pérdidas son una función cuadrática y los errores son de media cero, gaussianos y serialmente correlacionados (Mariano, 2000). La hipótesis nula H_0 se ilustra a continuación:

$$H_0: V(e_{1t}) = V(e_{2t}) \forall t \text{ (Los modelos tienen la misma exactitud)}$$

$$H_1: V(e_{1t}) \neq V(e_{2t}) \text{ (Los modelos tienen diferente exactitud)}$$

Manteniendo estas hipótesis desarrollan una prueba de exactitud de pronóstico basado en la ortogonalización. Al respecto, Diebold y Mariano (1995) señalan que la idea del *test* MGN es evitar asumir la independencia contemporánea a través de una transformación denominada ortogonalización para la cual los errores primero se transforman según:

$$x_t = (e_{1t} + e_{2t}) \tag{34}$$

$$z_t = (e_{1t} - e_{2t}) \tag{35}$$

En este caso, la hipótesis nula de media cero de pérdida diferencial es equivalente a la igualdad de las variaciones de los dos errores de pronóstico, o equivalente a cero covarianza entre x_t y z_t , ya que sigue directamente de la definición de x_t y z_t , y la covarianza de estas dos variables es:

$$\text{cov}(x_t, z_t) = E(e_{1t}^2 - e_{2t}^2) \quad (36)$$

Por lo tanto el estadístico de prueba es:

$$MGN = r / [(1 - r^2) / (T - 1)]^{1/2} \quad (37)$$

Donde:

$$r = x'z / [(x'x) / (z'z)]^{1/2} \quad (38)$$

Y x y z son los vectores $T \times 1$ con t elementos x_t y z_t respectivamente y x' y z' son los vectores transpuestos de los mismos. MGN tiene una distribución *t-student*²⁸ con $T - 1$ grados de libertad.

Por lo tanto, la hipótesis de igualdad de error es que no existe correlación entre x y z . Y si los errores se distribuyen normalmente, la hipótesis del error cuadrado medio igual puede ser probado mediante un *test* de correlación estándar entre x y z . Esta prueba se basa en la suposición que los errores de pronóstico son de tipo ruido blanco²⁹, por lo tanto la prueba sólo es

²⁸ Si z_1 es una variable normal estandarizada [es decir, $z_1 \sim N(0, 1)$] y otra variable z_2 sigue la distribución ji cuadrada con k grados de libertad y está distribuida independientemente de z_1 , entonces la variable definida como: $t = \frac{z_1}{\sqrt{(z_2)/k}} = \frac{z_1\sqrt{k}}{\sqrt{z_2}}$, sigue la distribución *t-Student* con k grados de libertad. Una variable distribuida t se

designa con frecuencia como t_k , donde el subíndice k denota los grados de libertad. Las propiedades de la distribución *t-student* son: La distribución t , lo mismo que la distribución normal, es simétrica, pero es más plana que la normal. Sin embargo, a medida que aumentan los grados de libertad, la distribución t se aproxima a la distribución normal. La media de la distribución t es cero y su varianza es $\frac{k}{(k-2)}$.

²⁹ El ruido blanco es una señal aleatoria en la cual no existe ninguna correlación estadística entre sus valores en dos tiempos diferentes. Por lo anterior, su densidad espectral es constante. Se dice que el ruido blanco es gaussiano cuando su función de densidad corresponde a una distribución normal.

aplicable a las predicciones de un solo paso. Además, la prueba es válida como una prueba de la igualdad de la exactitud del pronóstico sólo en la pérdida de error al cuadrado (Mariano, 2000).

4.5.2. *Test DM*

El *test DM* fue propuesto por Francis Diebold y Roberto Mariano en el año de 1995 con el objetivo de comparar la exactitud de dos pronósticos diferentes para una misma variable en un ambiente de independencia del modelo. El *test* fue propuesto para funciones de pérdida no cuadráticas, pronósticos de múltiples periodos y errores no normales, no de media cero, serialmente correlacionados y contemporáneamente correlacionados (Mariano, 2000).

Como ellos mismos exponen (1995), la intención fue proponer un *test* de amplia aplicabilidad considerando una hipótesis nula de no diferencia entre los datos originales y los modelos de pronóstico que se quieren comparar.

Si se tiene entonces un conjunto de datos y_t que corresponden a los valores reales y \hat{y}_{1t} y \hat{y}_{2t} dos modelos de pronóstico, la pregunta que surge es entonces: ¿Son esos modelos igualmente buenos para definir y_t ? Para esto, los autores proponen un modelo que se explicará a continuación:

Sean los errores de cada modelo de pronóstico:

$$e_{1t} = \hat{y}_{1t} - y_t \quad (39)$$

$$e_{2t} = \hat{y}_{2t} - y_t \quad (40)$$

Existe una función $g(e_{1t})$ y otra $g(e_{2t})$ que corresponden a las pérdidas asociadas en cada uno de los momentos t de los datos originales. Normalmente esta es una función que toma el valor de cero (0) cuando no hay error, no es nunca negativa e incrementa en tamaño en la medida que los errores aumenten. Teniendo las dos funciones se define entonces el diferencial de los dos modelos que se comparan como:

$$d_t = g(e_{1t}) - g(e_{2t}) \quad (41)$$

Si los dos pronósticos tienen la misma exactitud, d_t o función diferencial será cero para todo t . Es aquí entonces donde Diebold y Mariano (1995) introducen la hipótesis nula y la alternativa:

$$H_0: E(d_t) = 0 \quad \forall t \text{ (Los modelos tienen la misma exactitud)}$$

$$H_1: E(d_t) \neq 0 \quad \forall t \text{ (Los modelos tienen diferente exactitud)}$$

Ahora, tomando como referencia el análisis de Umberto Triacca (2006), se tienen:

$$\bar{d} = \sum_{t=1}^T d_t \quad (42)$$

$$f_d(0) = \frac{1}{2\pi} \{ \sum_{k=-\infty}^{\infty} \gamma_d(k) \} \quad (43)$$

\bar{d} en (42) es la media de la muestra de la pérdida diferencial y ; $f_d(0)$ y $\gamma_d(k)$ en (43) son la densidad espectral de la pérdida diferencial a una frecuencia de cero y la autocovarianza de la pérdida diferencial con un retraso k respectivamente.

$$\text{Bajo la } H_0, \frac{\bar{d}}{\sqrt{\frac{2\pi f_d(0)}{T}}} \rightarrow N(0,1)$$

De esta manera, con el propósito de validar la hipótesis nula, Diebold-Mario desarrollaron un estadístico de prueba que asume una covarianza estacionaria y una memoria corta. Este se define como:

$$DM = \frac{\bar{d}}{\sqrt{\frac{2\pi \hat{f}_d(0)}{T}}} \quad (44)$$

$\hat{f}_d(0)$ en (44) es un estimado de $f_d(0)$ equivalente a:

$$\hat{f}_d(0) = \frac{1}{2\pi} \sum_{k=-(T-1)}^{T-1} l\left(\frac{k}{h-1}\right) \hat{\gamma}_d(k) \quad (45)$$

$$\widehat{\gamma}_d(k) = \frac{1}{T} \sum_{t=|k|+1}^T (d_t + \bar{d}) (d_{t-|k|} - \bar{d}) \quad (46)$$

$$l\left(\frac{k}{h-1}\right) = \begin{cases} 1 & \text{para } \left|\frac{k}{h-1}\right| \leq 1 \\ 0 & \text{en otros casos} \end{cases} \quad (47)$$

Sin embargo,

$$\widehat{\gamma}_d(k) = \widehat{\gamma}_d(-k) \text{ y } l\left(\frac{k}{h-1}\right) = 0 \text{ cuando } |k| > h - 1 \quad (48)$$

Por lo tanto,

$$\widehat{f}_d(0) = \frac{1}{2\pi} \left(\widehat{\gamma}_d(0) + 2 \sum_{k=1}^{h-1} \widehat{\gamma}_d(k) \right) \quad (49)$$

Bajo la hipótesis nula, el estadístico DM está distribuido normalmente asintóticamente $N(0,1)$. La hipótesis nula de no diferencia será rechazada cuando $|DM| > z_{\alpha/2}$, siendo z el valor de la tabla estándar normal correspondiente al 50% del nivel α deseado para el *test*.

Desventajas

- Como los experimentos realizados por Diebold y Mariano (1995) mostraron, la distribución normal puede ser una aproximación pobre del estadístico DM para muestras finitas. Sus resultados mostraron que el *test* DM puede tener el tamaño incorrecto rechazando la hipótesis nula frecuentemente cuando se depende del grado de correlación serial entre los errores de los modelos de pronóstico y el tamaño de la muestra T (Triacca, 2006).
- Diebold (2013), casi veinte años después del desarrollo de su estadístico, plantea que el estadístico inicial DM fue entendido para comparar pronósticos y no modelos como muchos estudios lo han hecho: La esencia de DM es tomar pronósticos que son calificados como primitivos y realizar asunciones sobre los mismos y más exactamente sobre la pérdida diferencial. El estadístico no fue planteado para la comparación de los modelos en si teniendo en cuenta que estos últimos requieren la estimación de los parámetros y este proceso podría inducir una covarianza no estacionaria.

4.5.3. Test HLN

Los autores Harvey, Leybourne y Newbold (1997) trabajaron con base en los estadísticos de los *test* MGN y DM generando para cada uno de ellos una versión modificada. De esta manera buscan corregir la posible heterocedasticidad del MGN y las limitaciones del DM para muestras pequeñas.

4.5.3.1. MGN modificado

Los autores llevaron el *test* MGN a un modelo de regresión de la forma $x_t = \beta z_t + \varepsilon_t$. El estadístico en (37) es equivalente entonces al estadístico para probar la hipótesis nula que $\beta = 0$ en la regresión en mención. Teniendo en cuenta lo anterior, reformulan el estadístico inicial expresándolo de la siguiente forma:

$$MGN = \beta / (s^2 / z' z)^{1/2} \quad (50)$$

Y dónde:

$$\beta = x' z / z' z \quad (51)$$

$$s^2 = (x - \beta z)' (x - \beta z) / (T - 1) \quad (52)$$

Esta prueba sería válida y de hecho más potente e imparcial que la anterior dada en (37) en la situación en que se consideren supuestos ideales (Mariano, 2000).

Cuando los errores de pronóstico provienen de una distribución de cola pesada, los autores argumentan que la estimación de la varianza de β , que aparece en el denominador de MGN es sesgada y por lo tanto recomiendan la siguiente modificación de MGN:

$$HLN - MGN \text{ modificado} = \beta / [(\sum z_t^2 \hat{\varepsilon}_t^2) / (\sum z_t^2)^2]^{1/2} \quad (53)$$

Donde $\hat{\varepsilon}_t$ corresponde al residuo de mínimos cuadrados ordinarios en el tiempo t . En esta modificación se aprecia entonces una corrección para la heterocedasticidad en la estimación de la

varianza de β . Harvey, Leybourne y Newbold (1997) sugieren comparar el estadístico HLN-MGN con valores críticos de la distribución t con $T - 1$ grados de libertad.

El estudio de simulación concluye que la prueba original MGN tiene tamaños empíricos que son iguales a los tamaños nominales cuando los errores de pronóstico se han extraído de una distribución de Gauss. Sin embargo, cuando el error de pronóstico en el proceso de generación es t_6 (distribución t con seis grados de la libertad), el *test* MGN se empieza a sobredimensionar, obteniendo deficiencias peores a medida que aumenta el tamaño de la muestra.

Para la prueba de MGN modificado, los resultados de la simulación son variables. Como teoría sugiere que el *test* registra el tamaño correcto cuando la muestra es grande. Sin embargo, hay un sobredimensionamiento cuando el tamaño de las muestras es pequeño, tanto para el caso en que la distribución de los errores es una normal o t_6 . De hecho, en el último caso (con muestras pequeñas), el *test* modificado resulta peor que el *test* original MGN.

Estos resultados de la simulación llevaron a los autores a considerar todavía otra variación, una con enfoque no paramétrico utilizando *test* de rangos de Spearman³⁰ para la correlación entre x y z . Sin embargo, la gran desventaja de las modificaciones realizadas en general, es la limitación de su aplicabilidad para predicciones con un horizonte a un paso ($h = 1$) y para funciones de pérdida cuadráticas.

4.5.3.2. DM modificado

De la misma manera como trabajaron con el *test* MGN, Harvey, Leybourne y Newbold (1997) propusieron una modificación al *test* de Diebold y Mariano cuyo objetivo principal fue poder trabajar con muestras pequeñas. La primera de las modificaciones se basa en el uso de un estimador insesgado de la varianza de d_t asumiendo entonces que para horizontes de pronóstico, la pérdida diferencial d_t tiene autocovarianza. La varianza de d_t está dada por:

³⁰ Hace referencia a la relación que existe entre dos variables usando una función monotónica, es decir, aquella que conserva el mismo orden. La relación está definida entonces por el coeficiente de correlación de Spearman ρ (debido a Charles Spearman).

$$V(\bar{d}) = \frac{1}{T} \left[\gamma_0 + \frac{2}{T} \sum_{k=1}^{h-1} (T-k) \gamma_k \right] \quad (54)$$

Los autores logran una aproximación del valor esperado de $V(\bar{d})$ así:

$$E\left(\widehat{V(\bar{d})}\right) \sim V(\bar{d}) \left[(T+1-2h+h(h-1)/T)/T \right] \quad (55)$$

Y por lo tanto sugirieron modificar el estadístico inicial así:

$$HLN - DM \text{ modificado} = DM \sqrt{\frac{T+1-2h+h(h-1)/T}{T}} \quad (56)$$

Una modificación adicional intuitivamente razonable del *test* DM es comparar el estadístico con los valores críticos de la distribución *t-Student* con $n - 1$ grados de libertad en lugar de utilizar la distribución normal estándar. La tendencia del estadístico original de sobredimensionamiento en la medida que el horizonte de pronóstico crece sobretodo en muestras pequeñas, se corrigió de manera significativa con la modificación del mismo según los resultados de las simulaciones de MonteCarlo realizadas por los autores.

4.5.4. *Test* Giacomini y White

Giacomini y White (2006) propusieron un mecanismo de comparación entre pronósticos basado en la inferencia sobre las expectativas condicionales de los mismos así como de sus errores en lugar de hacerlo sobre las expectativas incondicionales que son el foco de los demás *test*. Lo anterior significa que capturan información no solo del modelo como las demás herramientas de comparación sino que también prestan atención a determinantes importantes como el procedimiento mismo de estimación de los parámetros y la ventana de datos a usar: Esto lo hacen porque a pesar que un modelo esté ajustado a los datos no necesariamente puede generar buenos pronósticos por razones como por ejemplo parámetros no estimados precisamente.

Los *test* de GW pueden ser aplicados a pronósticos de múltiples pasos, intervalos y múltiples funciones de probabilidad y densidad. Los autores toman como base el trabajo realizado por West (1996) el cual consiste en probar la hipótesis nula de igual exactitud de dos pronósticos:

$$H_0: E \left[(Y_{t+1} - f_t(\beta_1^*))^2 - (Y_{t+1} - g_t(\beta_2^*))^2 \right] = 0 \quad (57)$$

Donde Y_{t+1} es la variable de interés; $f_t(\beta_1)$ y $g_t(\beta_2)$ son dos modelos de pronóstico y β_1^* y β_2^* son estimados de los parámetros. Nota: Los pronósticos son paramétricos por la dependencia de los parámetros β_1 y β_2 . De esta manera, la hipótesis nula se cumple cuando los dos modelos son en promedio igual de exactos. Por su parte cuando se rechaza, se escogería el modelo con la menor cantidad de pérdidas (errores). Lo que Giacomini y White aducen con respecto a lo anterior es que la hipótesis nula permitirá seleccionar un pronóstico basado en un modelo correctamente especificado, esto es, el modelo que mejor se aproxima al proceso de generación de datos Y_{t+1} .

Sin embargo, no se puede garantizar que el modelo correctamente especificado pronostique de la mejor manera por lo cual, si lo que se pretende es encontrar el modelo que mejor pronostique de cara al futuro, la hipótesis nula no es lo suficiente clara para inferirlo. Es aquí donde Giacomini y White proponen su idea central, la habilidad predictiva condicional: Una expectativa condicionada al conjunto de datos F_t disponible en un tiempo t con unas pérdidas que dependen de la estimación de los parámetros $\widehat{\beta}_{1t}$ y $\widehat{\beta}_{2t}$ en t y no en sus límites de probabilidad:

$$H_0: E \left[(Y_{t+1} - f_t(\widehat{\beta}_{1t}))^2 - (Y_{t+1} - g_t(\widehat{\beta}_{2t}))^2 \mid F_t \right] = 0 \text{ con } t = 1, 2, \dots \quad (58)$$

Con la modificación realizada por los autores sobre la propuesta de West y la condicionalidad en el tiempo, se obtienen las siguientes ventajas:

- El *test* GM es aplicable para aquellos casos en los cuales los datos son heterogéneos, es decir, con cambios subyacentes que varían en el tiempo. Como mencionaron Clements y Hendry (1999) esto puede ser más realista a la hora de analizar variables económicas.
- La heterogeneidad igualmente afecta el espacio de tiempo en el cual se toman los datos de la variable de interés. Con el *test* propuesto los autores no consideran una ventana de datos que

se alarga en el tiempo sino que por el contrario parten de la premisa que el espacio de tiempo puede ir cambiando y que esto hace parte de la estimación del modelo.

- La propuesta de GM permite la comparación de modelos anidados y no anidados.

Desventajas

- La debilidad principal de los resultados del *test* de Giacomini y White es no puede ser usado en esquemas recursivos (Giacomini & Rossi, 2013).
- Si bien el considerar determinantes como la estimación de los parámetros y el tamaño de las ventanas a usar hace el *test* de Giacomini y White robusto en sus consideraciones teóricas, también lo convierte en un método más complejo sobre todo en ambientes donde el esquema de *rolling* que plantean para diferentes conjuntos de datos F_t no es ampliamente conocido. Clark y McCracken (2010) mencionan que si bien el método de *rolling* es común en las personas que trabajan en el mundo del pronóstico, éste no es universal.
- Clark y McCracken (2010) igualmente mencionan que en el estudio de Giacomini y White (2006) no hay evidencia en las simulaciones de Monte Carlo de las magnitudes de P y R (*out-of-sample*³¹ y *in-sample*³² respectivamente). Estos mismos autores, Clark y McCracken (2012), mencionan que su *test* solo aplica para modelos estimados con una ventana fija y de tamaño finito.

³¹ *Out-of-sample* hace referencia a periodos diferentes de tiempo a los de los datos con los cuales se trabajó. Así por ejemplo si se habla de pronósticos *out-of-sample* estos corresponden a pronósticos diferentes a los de la muestra con la cual se construyeron los modelos.

³² El término *in-sample* corresponde a aquella validación que se realiza con los datos de la muestra. Así por ejemplo si se habla de un pronóstico *in-sample* o de ajuste, este corresponde a aquel modelo que valida el pronóstico para los mismos datos con los cuales se generó.

5. LA VARIABLE: TASA DE CAMBIO

Sin lugar a dudas el éxito empresarial exige que se observen de manera permanente los factores o variables externas que tiene impacto en los negocios. Esto obliga a las compañías a monitorear su entorno social, político, económico y de mercado. Si esto es así, estos fenómenos implican un impacto significativo en sus arcas, razón por la cual se debe observar con recelo el factor financiero.

En ese entendido, es de suma importancia que las compañías incluyan dentro de sus componentes de gestión unas adecuadas herramientas y técnicas que les permitan observar esos entornos. Esto sugiere incorporar o desarrollar metodologías de planeación que apoyen finalmente a la toma de decisiones. De esta manera se puede evaluar el presente y tratar de representar el futuro.

Es acá donde toma vital importancia el pronóstico en las finanzas, donde surgen cierto tipo de preguntas que el empresario quisiera resolver, por ejemplo: ¿Cómo visualizar el futuro? ¿Qué variables impactan la operación? ¿Cómo estimar el comportamiento de esas variables? ¿Cómo afectarán los resultados de la empresa?

Torres Barrón (Sin fecha) resalta las bondades de la aplicación de los pronósticos en la planeación administrativa, lo anterior como un intento por reducir la incertidumbre y respaldar la toma de decisiones en algo más que suponga la sola intuición de los empresarios. Así mismo, Hanke y Wichern (2006) sustentan que las empresas puedan reducir el grado de incertidumbre y deben respaldar sus decisiones en algo más que la intuición, esto a través de la elaboración de pronósticos correctos y precisos que sean suficientes para satisfacer las necesidades de planeación de la organización.

Se puede afirmar entonces que los pronósticos financieros son importantes en el sentido que el empresario no solo debe considerar intuiciones o suposiciones subjetivas, las cuales pueden ser válidas por la experiencia o el conocimiento de un sector específico, sino que adicionalmente es necesario acudir a mecanismos que le permitan conocer el entorno en cifras para suponer el futuro. Esto requiere acudir a supuestos que se fundamentan en bases conocidas, ejemplo: datos

y comportamientos históricos, índices o variaciones, márgenes, etc. Finalmente estos resultados apoyaran de manera significativa la toma de decisiones.

La introducción anterior permite involucrar los objetivos del presente capítulo:

- a) Explicar la relevancia en el mercado financiero de la tasa de cambio, variable que se modelará en los siguientes capítulos para comparar de manera práctica las medidas e instrumentos de evaluación de pronósticos.
- b) Presentar los estudios específicamente en Colombia que hayan tratado de modelar la tasa de cambio presentando las medidas usadas (si así lo hicieron) para validar sus pronósticos.

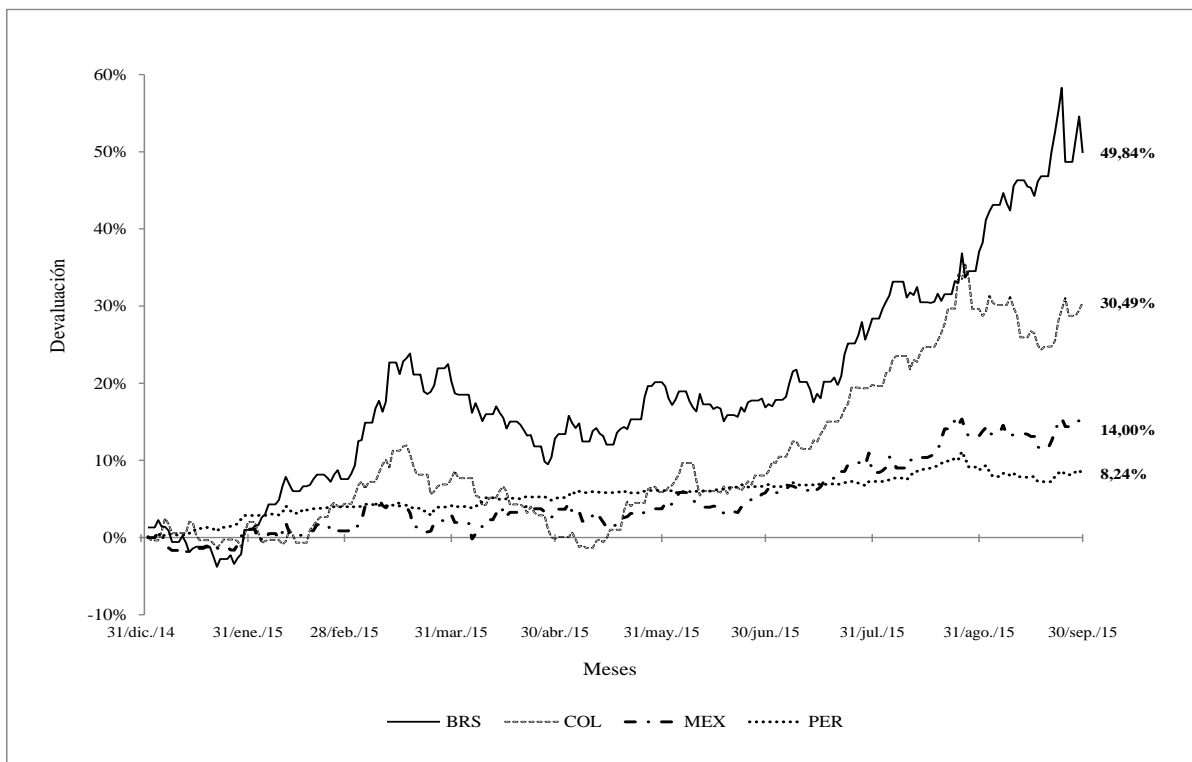
5.1. La relevancia de la tasa de cambio

El 2015 sin duda fue un año difícil para el comercio internacional, principalmente en las operaciones de bienes y servicios requeridos por Sudamérica. América Latina presentó una depreciación considerable de sus monedas frente al dólar americano, cifras como: Brasil al 49,84%, Colombia al 30,49%, México al 14,00% y Perú al 8,24% tal como lo muestra la Gráfica 1, castigaron fuertemente al sector importador.

Lo anterior sin mencionar casos como Argentina y Venezuela donde las cifras son igualmente salidas de lo común. Esto directamente afecta desde las compañías mismas hasta el ciudadano que tenía como propósito, por ejemplo viajar al exterior y ve que su presupuesto ya no es suficiente.

La devaluación de la moneda afecta directamente la balanza comercial y por ende la balanza de pagos de un país, que se suma a otros fundamentales económicos, tales como la inflación, el desempleo y el PIB (Producto Interno Bruto), entre otros. Ahora en el ámbito local, los presupuestos de las compañías que efectúan operaciones en moneda extranjera, se ven impactados negativa o positivamente ante las fluctuaciones que pueda tener la tasa de cambio. Seguramente muchas empresas optaron por frenar proyectos de inversión o crecimiento por las implicaciones de la contratación de servicios o compra de bienes con tarifas en dólares.

Gráfica 1. Devaluación de monedas frente al dólar americano



Fuente: Elaboración propia con información extraída de www.banrep.gov.co y www.banxico.org.mx.

Es indudable entonces la relevancia de esta variable en Colombia, donde el intercambio comercial con Estados Unidos, como socio y aliado estratégico en operaciones de exportación e importación de bienes y servicios, o como fuente directa en la inversión extranjera, impacta de manera significativa el crecimiento económico del país.

Quizás la pregunta más consecuente que podemos hacernos ante este panorama es: ¿Alguien en sus pronósticos logró prever la depreciación del peso frente al dólar en los porcentajes observados?

5.2. Los estudios en Colombia

En Colombia varios estudios se han realizado sobre la tasa de cambio e investigadores de diferentes ciencias han intentado pronosticarla. Entidades como el Banco de la República, quien

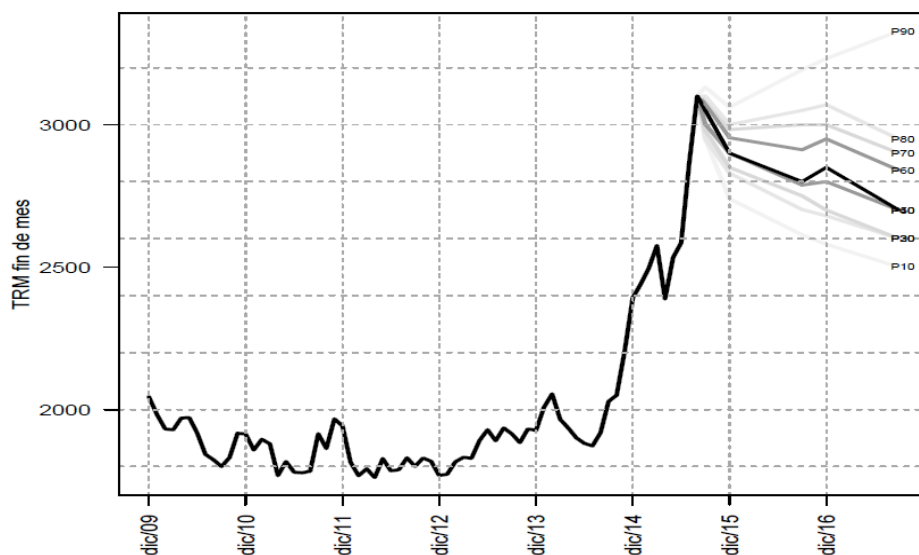
cumple un rol importante como Banco Central en el país; entidades privadas como Fedesarrollo; así como otras entidades del sector financiero y comisionistas, tales como Grupo Bancolombia, Helm Bank, Grupo Aval, Citibank, Corficolombiana, Alianza Valores, entre otros, tienen grupos de profesionales expertos que monitorean y analizan constantemente esta variable, así como muchas otras.

Sobresalen entonces los portales web de estas entidades donde se consignan las expectativas financieras y económicas del país con relación a la tasa de cambio, tasas de interés, inflación, índice de desempleo, etc.

Los datos estimados más cercanos con relación a la tasa de cambio se pudieron observar en la encuesta mensual de expectativas de analistas económicos que realiza el Banco de la República, la cual está disponible en su página web y contiene expectativas desde el año 2003.

En la Gráfica 2 se puede observar el resultado de la encuesta de expectativas del mes de septiembre de 2015 en la cual se informan las perspectivas de tasa de cambio de septiembre 2015 a septiembre de 2017, fin de mes.

Gráfica 2. Expectativas tasa de cambio - Percentiles



Fuente: Banco de la República, gráficas y series históricas de la encuesta de expectativas de analistas económicos www.banrep.gov.co

Con relación a la evaluación de pronósticos, el Banco Central de Colombia también publica en su página web una evaluación de pronósticos mensual y acumulado de la tasa de cambio, donde simplemente se informan las diferencias entre los pronósticos de algunas las entidades financieras y los valores observados. En la Tabla 7 se puede observar la encuesta correspondiente al mes de agosto de 2015.

Tabla 7. Encuesta de expectativas tasa de cambio.

Fecha de realización: Del 10 al 12 de agosto de 2015		
TRM 31 de agosto de 2015 \$ 3.101,10		
Puesto ocupado	Entidad	Diferencia en pesos
1	Ultrabursátiles	\$ 85,10
2	Alianza Valores S.A	\$ 101,10
2	Corredores Davivienda S.A.	\$ 101,10
3	Fondo Latinoamericano de Reservas	\$ 116,10
4	Finagro S.A	\$ 121,10
5	Old Mutual Pensiones y Cesantías	\$ 131,10
5	Old Mutual comisionista de bolsa	\$ 131,10

Nota: La encuesta se realizó a 40 entidades privadas y del sector financiero, se muestran las cinco primeras posiciones. La distancia entre el pronóstico y el valor observado se refiere al valor absoluto, es decir, el pronóstico puede estar por encima o por debajo del valor observado.

Fuente: Elaboración propia con información extraída de la encuesta expectativas de inflación, tasa de cambio y tasa de intervención Banco de la República www.banrep.gov.co.

Desafortunadamente los modelos utilizados por estas entidades así como los del Banco de la República no son públicos. De igual manera, la información técnica es confidencial y se desconoce el detalle de los métodos y medidas de precisión que son empleados para el contraste y evaluación de los pronósticos.

Como resultado de lo anterior, se analizaron los estudios realizados por autores en Colombia con artículos de índole público con el objetivo de validar sus métodos de pronóstico y los resultados obtenidos. Gran parte de ellos están vinculados al sector académico o relacionados con el Banco de la República.

Un primer estudio corresponde al de Patiño y Alonso (2005) quienes evaluaron cuatro pronósticos de la tasa de cambio nominal: El modelo de enfoque monetario de precios rígidos³³,

³³ Modelos propuestos por Dornbusch (1976) y Frankel (1979) que suponen entre otros movilidad perfecta de capitales, cumplimiento del PPA y que la variación del tipo de interés nominal provoca entradas o salidas de capital que afectan al tipo de cambio nominal.

el modelo monetario de diferenciales de productividad³⁴, el modelo de paridad de poder adquisitivo (PPA)³⁵ y un modelo de caminata aleatoria. El análisis se realizó para el periodo de 1984 a 2004. Los pronósticos fueron evaluados utilizando la medida RMS y el coeficiente U-Theil, los cuales comprobaron que el modelo más acertado era el de caminata aleatoria. No se encontró en el documento la explicación del porqué del uso de tales medidas.

Un segundo estudio fue realizado por Munir y Misas (2007) quienes evaluaron con medidas simétricas (MAPE, MAE, RMSE y RMSPE) y asimétricas (lin-lin)³⁶ el pronóstico de la tasa de cambio nominal a partir de un modelo lineal ARIMA y de uno no lineal que consideró una red neuronal artificial autorregresiva.

El periodo analizado abarcó desde el 8 de febrero de 2000 y el 1 de marzo de 2005 dejando las últimas 60 observaciones para la evaluación *Rolling* fuera de muestra. Uno de los resultados obtenidos fue que “(...) a la luz de las funciones de pérdida asimétricas, los modelos no lineales tienen una mejora considerable en capacidad de pronóstico, con respecto a los modelos lineales” (Munir & Misas, 2007, pág. 158). Cabe resaltar que al igual que el primer estudio, no se encontró información de las razones de selección de las medidas empleadas para la evaluación.

Un modelo igualmente soportado en redes neuronales fue el desarrollado por Ayala y Castillo (2008). Los autores modelaron una red neuronal del tipo perceptrón³⁷ para la base de datos que abarcó enero de 1995 a junio de 2004. Las conclusiones de su investigación estuvieron orientadas al modelo de red como tal con resultados favorables a corto plazo evaluada con MSE.

³⁴ Modelo que incluye una medida de productividad según la hipótesis Balassa-Samuelson que predice una relación positiva entre el diferencial de productividades sectoriales y la evolución del tipo de cambio real.

³⁵ La teoría de la paridad del poder adquisitivo afirma que los tipos de cambio entre las divisas deben ser tales que permiten que una moneda tenga el mismo poder adquisitivo en cualquier parte del mundo.

³⁶ Ver formulación de la función de pérdida lin-lin en numeral 4.3 del presente documento

³⁷ Son un tipo de redes neuronales ampliamente estudiadas que se caracterizan por su estabilidad, eficiencia para minimizar el error y porque sus neuronas están dispuestas en capas o niveles comunicándose una a una.

6. PLANTEAMIENTO METODOLÓGICO

El tipo de investigación que se realizó fue de carácter exploratorio ejemplificando un caso específico, la tasa de cambio en Colombia. En ese sentido se pretende someter a prueba algunos instrumentos o metodologías para encontrar limitaciones o fallas al respecto (Arias Galicia, 2012).

6.1. Fuentes de información

Para el presente ejercicio la variable tasa de cambio de Colombia hace referencia a la relación de pesos colombianos por dólar americano. La fuente primaria para la recolección de la información de esta variable fue la página web de la Superintendencia Financiera de Colombia³⁸. La tasa de cambio corresponde al promedio aritmético de las tasas promedio ponderadas de compra y venta de divisas de las operaciones interbancarias y de transferencias.

Tal como se menciona en la página web de la Entidad, la Superintendencia Financiera de Colombia tiene la obligación de calcular la Tasa de Cambio Representativa del Mercado o TCRM, divulgarla al mercado y al público en general, conforme a lo dispuesto en el artículo 80 de la Resolución Externa No. 08 de 2000, expedida por la Junta Directiva del Banco de la República.

6.2. Validación teórica

El desarrollo de la investigación objeto del presente documento, inició con la búsqueda a fondo de bibliografía pertinente así como del estado del arte relacionado con las medidas de precisión y *test* de validación de pronósticos.

³⁸ <https://www.superfinanciera.gov.co/>

A partir de la documentación obtenida, se realizó una comparación teórica de las diferentes metodologías de evaluación de pronóstico, identificando ventajas y desventajas de cada una de ellas. Así mismo, se revisó la literatura correspondiente y se incluyó al ejercicio anterior, los *test* que han desarrollado varios autores y que tienen por objeto la validación de la precisión de dos pronósticos de una misma variable.

6.3. Modelación econométrica

Como ya se mencionó, se tomó como variable la TCRM, serie de impacto dada la importancia estructural en la economía y su comportamiento a lo largo del año 2015, la cual generó expectativas de incertidumbre no solo para el gobierno sino también para las compañías que impulsan el mercado industrial, comercial y financiero en Colombia.

Para la modelación y con el objeto de evaluar la variable en términos de inversión a mediano plazo, se procesaron los datos de cierre de cada mes. Los datos de entrada comprendieron un periodo de 180 meses, octubre de 2000 a septiembre de 2015, de los cuales 120 se utilizaron para la obtención de los modelos de pronóstico y 60 para validación y evaluación de los mismos.

Con la información obtenida de la TCRM de octubre 2000 a septiembre de 2010 se realizó entonces la modelación para explicar la dependencia de la variable consigo misma. Se calcularon dos modelos de pronóstico: Un modelo ARIMA (modelo autorregresivo integrado de promedios móviles) y un modelo de movimiento browniano. Es importante mencionar que el propósito del presente trabajo de investigación no es escoger el mejor modelo de pronóstico sino validar cómo funcionan las medidas de evaluación de pronósticos.

Una vez terminado el ejercicio anterior y a partir de las funciones de cada modelo, se pronosticó la TCRM por periodos sucesivos de seis meses: El primer pronóstico comprendió el período octubre de 2010 a marzo de 2011. Consecuentemente, se calcularon los errores para cada uno de los pronósticos tomando los datos reales de la TCRM de cierre de mes.

Este ejercicio de pronóstico se realizó iterativamente en 60 oportunidades utilizando la metodología de *Rolling*³⁹. Esto significa que se realizaron pronósticos a seis meses integrando el dato real del mes inmediatamente anterior a los modelos de pronóstico. Como ejemplo, los tres primeros grupos de meses en los que se pronosticó la tasa de cambio fueron octubre de 2010 a marzo de 2011, noviembre de 2010 a abril de 2011 y diciembre de 2010 a mayo de 2011. El primer grupo no incluyó ningún dato real puesto que los modelos tuvieron en cuenta la información de los primeros 120 meses mencionados. Sin embargo, el segundo y tercer grupo incluyó la información real de los meses de octubre y noviembre respectivamente.

A medida que se realizó el pronóstico a seis meses, se obtuvieron grupos de datos de la tasa de cambio pronosticada así como los errores respectivos. A esto se adicionó la información del pronóstico del periodo *in-sample* así como los errores del mismo (periodo inicial de 120 meses).

Por último y ya con las tablas de datos obtenidos, se calcularon uno a uno los valores de las medidas de precisión de los pronósticos así como los *test* de comparación diferenciales. La exploración práctica y análisis de los datos será una consecuencia de lo anterior.

Cabe resaltar que para todo el proceso descrito, se tomaron únicamente los modelos obtenidos con la información de los 120 meses iniciales. Las iteraciones del *Rolling* incluyeron uno a uno los 60 siguientes datos reales incorporándolos a las funciones de los modelos iniciales. Esto significa que el *Rolling* no generará iterativamente nuevas funciones de modelos de pronóstico. Lo anterior por cuanto se considera que los modelos ARIMA y de movimiento browniano obtenidos son estables.

³⁹ La metodología de pronóstico de *Rolling* consiste en un proceso FIFO (First In First Out) de entrada y salida de datos de tal manera que se mantiene una ventana de tiempo fija. Puesto que tal ventana requiere supervisión, se denomina igualmente como una metodología iterativa en la que se comparan, incluyen y eliminan datos pronosticados y datos reales.

7. RESULTADOS

7.1. Los modelos

Para la estimación de los pronósticos se utilizaron dos modelos, el ARIMA y el de movimiento browniano. Estas metodologías fueron explicadas en el capítulo 3.

7.1.1. ARIMA

El modelo se desarrolló a partir de 120 datos históricos de la serie TCRM para el período 2000M10 a 2010M09⁴⁰. Con estas observaciones se realizó el pronóstico de 60 datos para estimar el periodo 2010M10 a 2015M09. Con el fin de validar la información histórica de la serie, se evaluó como primera medida que los datos viniesen de una serie estacionaria⁴¹.

El resultado obtenido con la prueba de raíz unitaria de Dickey-Fuller⁴² llevó a la transformación en diferencias de la serie TCRM. Posteriormente, se analizó el correlograma para identificar el proceso estocástico que determinó la serie. Esto dio como resultado los ordenes de los términos del modelo una vez revisados los índices de autocorrelación y la correlación parcial.

Los términos autorregresivos (AR) y de medias móviles (MA) se presentan en (59). En el proceso de inclusión y descarte de los términos se utilizó a manera de selección el criterio de Akaike. El modelo⁴³ es el siguiente:

⁴⁰ 2000M10 a 2010M09 significa el periodo comprendido entre octubre de 2.000 y septiembre de 2.010. En este caso la M intermedia hace referencia a Mes.

⁴¹ Una serie es estacionaria cuando su media y varianza son constantes en el tiempo.

⁴² Este *test* debe su nombre a David Dickey y Wayne Fuller. En nivel formal, la estacionariedad se verifica averiguando si la serie de tiempo contiene una raíz unitaria, esta prueba sirve para ese propósito (Gujarati & Porter, 2010).

⁴³ Las pruebas de ruido blanco y de estacionariedad e invertibilidad del modelo o proceso generador se presentan en el Anexo 1.

$$(1 - 0,138 L^5 - 0,264 L^{14} - 0,238 L^{33}) (1 - L) TRM = (1 + 0,228 L + 0,164 L^4 - 0,208 L^6 - 0,152 L^{11} - 0,168 L^{13}) \varepsilon_t \quad (59)$$

Con respecto a la calidad de precisión de este modelo y para el desarrollo del presente ejercicio, no se involucró el efecto ARCH ⁴⁴, pues lo que se busca es que la varianza incondicional sea constante y ARCH parte de la premisa que existe una varianza condicional. Seguramente en modelos de riesgos es importante incluirlo porque interesa la volatilidad. Adicionalmente, las pruebas ARCH no validan si el modelo ARIMA esta bien especificado ni mejora la calidad de la precisión.

7.1.2. Movimiento browniano

El modelo de Movimiento Browniano parte de la justificación que este es una caminata aleatoria, demostración que se presenta en el Anexo 2. Si se toma como base la definición del movimiento geométrico browniano obtenido, esto es $S_t = S_0 e^{t\mu + Z(\sigma\sqrt{t})}$, para el modelo de pronóstico de la TCRM, se tendría entonces:

- μ va a ser estimado a través de la media de los retornos de la muestra de datos 2000M10 a 2010M09.
- σ va a ser estimado a través de la desviación estándar de los retornos de la muestra de datos 2000M10 a 2010M09.
- Z es una variable aleatoria que se distribuye como una $N(0,1)$.

⁴⁴ Con mucha frecuencia, un valor d significativo es una indicación de que existen errores de especificación del modelo. Uno de estos errores es el efecto ARCH, conocido como efecto de heteroscedasticidad condicional autorregresiva. Por lo general, los precios de los activos financieros se caracterizan por un fenómeno conocido como acumulación de volatilidad, lo que significa que existen lapsos en los que muestran amplias variaciones durante prolongados periodos seguidos por un intervalo de tranquilidad relativa. Los modelos con heteroscedasticidad condicional autorregresiva (ARCH) o modelos con heteroscedasticidad condicional autorregresiva generalizada (GARCH) representan la mencionada acumulación de volatilidad (Gujarati & Porter, 2010).

- S_0 es el valor u observación real inmediatamente del periodo inicial, que en algunos casos representa el precio anterior.
- t es el horizonte del pronóstico que equivaldría en el presente caso a $h = 1, \dots, 6$.

Puesto que Z es una variable aleatoria que se distribuye como una $N(0,1)$, para efectos de generación de los pronósticos se realizó una simulación de Montecarlo con N 10.000 escenarios posibles para esta variable.

7.1.3. Los pronósticos

Se procedió entonces a realizar el pronóstico para el periodo 2010M10 a 2015M09, el cual es el periodo de la muestra de evaluación de los pronósticos (pronóstico *out-of-sample*), tanto para el modelo ARIMA como para el de movimiento browniano. Este procedimiento se llevo a cabo mediante la metodología de *Rolling* para horizontes de seis meses ($h = 1, \dots, 6$).

Como se aprecia en la Tabla 8, con los datos de muestra base que van hasta 2010M09, se realizó un primer pronóstico de los siguientes seis meses (periodo 2010M10 a 2011M03) tabulando los resultados en columnas independientes y correspondientes a cada horizonte de $h = 1$ a $h = 6$. En la tabla en mención estos valores de pronóstico se pueden identificar con el color gris claro.

Este proceso se repitió para pronosticar el siguiente periodo de seis meses, periodo 2010M11 a 2011M04, tomando los mismos modelos estimados e incorporando la siguiente observación real, es decir, aquella del mes 10 del año 2010 (2010M10). El resultado obtenido para cada horizonte $h = 1, \dots, 6$ son los pronósticos con color azul de la Tabla 8. Este ejercicio se realizó T veces hasta llegar al periodo 2015M09.

En la tabla se ilustran igualmente tres ejemplos adicionales cuando se incorporan observaciones reales en 2010M11, 2010M12 y 2011M01 (ver pronósticos con color verde, amarillo y gris oscuro respectivamente).

Tabla 8. Pronóstico metodología *rolling*

Periodo	Tasa Real Yt	Pronóstico para h=1	Pronóstico para h=2	Pronóstico para h=3	Pronóstico para h=4	Pronóstico para h=5	Pronóstico para h=6
I_t 201009	\$ 1.799,89						
I_{t+1} 201010	\$ 1.831,64	\$ 1.820,82	1DP				
I_{t+2} 201011	\$ 1.916,96	\$ 1.984,33	\$ 1.852,98	2DP			
201012	\$ 1.913,98	\$ 1.932,62	\$ 1.987,45	\$ 1.843,42	3DP		
201101	\$ 1.857,98	\$ 1.863,01	\$ 1.858,25	\$ 2.039,48	\$ 1.823,94	4DP	
201102		\$ 1.879,84	\$ 1.778,71	\$ 1.809,24	\$ 1.983,23	\$ 1.902,67	
201103			\$ 1.926,85	\$ 1.808,14	\$ 1.803,11	\$ 1.998,29	\$ 1.797,39
201104				\$ 1.897,56	\$ 1.745,83	\$ 1.806,45	\$ 2.059,48
201105					\$ 1.960,76	\$ 1.808,45	\$ 1.812,52
201106						\$ 1.951,02	\$ 1.806,85
201107							\$ 2.087,31
⋮							
I_{t+T} 201509							

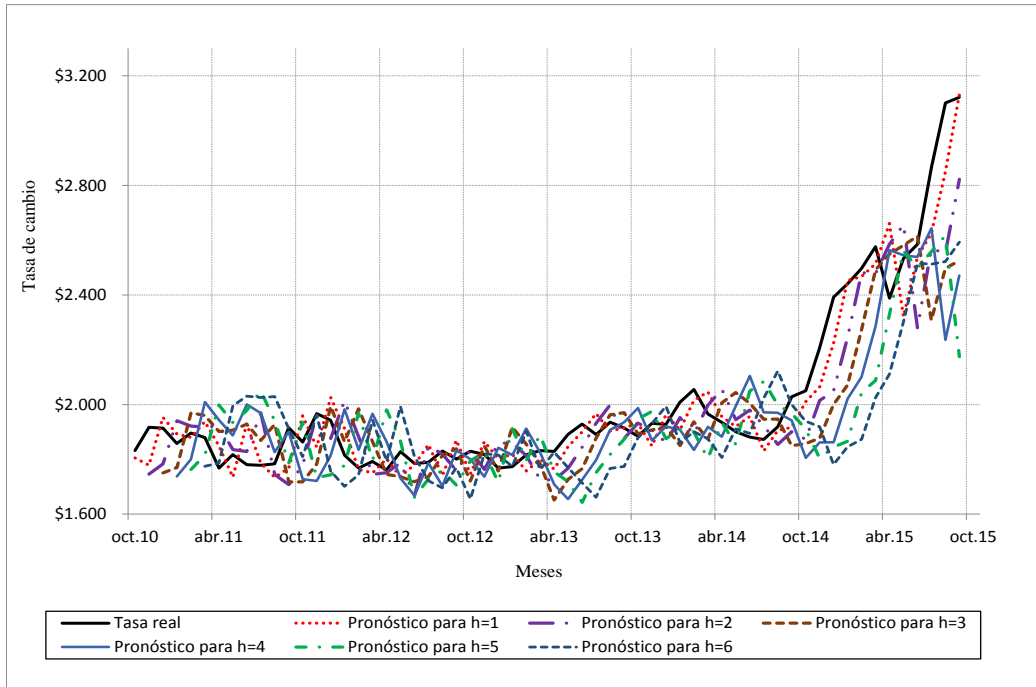
Fuente: Elaboración propia

Dos puntos importantes por resaltar es que en la medida que se acerca al periodo 2015M09 ya no se pueden pronosticar los seis horizontes: Se van reduciendo en uno hasta poder hallar uno solo con la observación real de I_{t+T-1} . Así mismo, al inicio de los pronósticos tal como lo presenta la Tabla 8, se va perdiendo un dato en la medida que el horizonte avanza: Para $h = 2$ se pierde un dato⁴⁵, para $h = 3$ se pierden dos datos y así sucesivamente.

El resultado del pronóstico hasta el periodo 2015M09 para el modelo ARIMA se presenta en el Anexo 3 donde se puede visualizar la tasa real por cada mes así como el pronóstico correspondiente para cada uno de los diferentes horizontes $h = 1, \dots, 6$. Esta misma información se esboza en la Gráfica 3. Esto mismo puede apreciarse para el modelo browniano en el Anexo 4 y Gráfica 4. Cabe anotar que para efectos de análisis también se realizó el pronóstico de ajuste o *in-sample* con cada uno de los modelos, los cuales se presentan en el Anexo 5 y Gráfica 5.

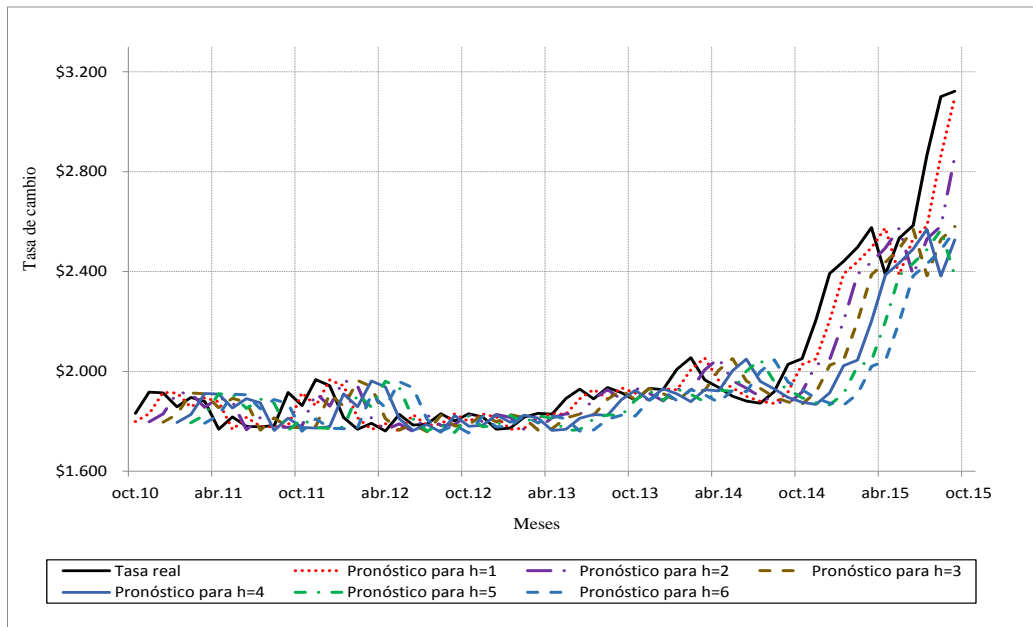
⁴⁵ En la Tabla 8 esto se visualiza como 1DP, 2DP, 3DP, 4DP o 5DP que significa 1, 2, 3, 4 o 5 Datos Perdidos.

Gráfica 3. Modelo ARIMA - Tasa real vs. Pronósticos



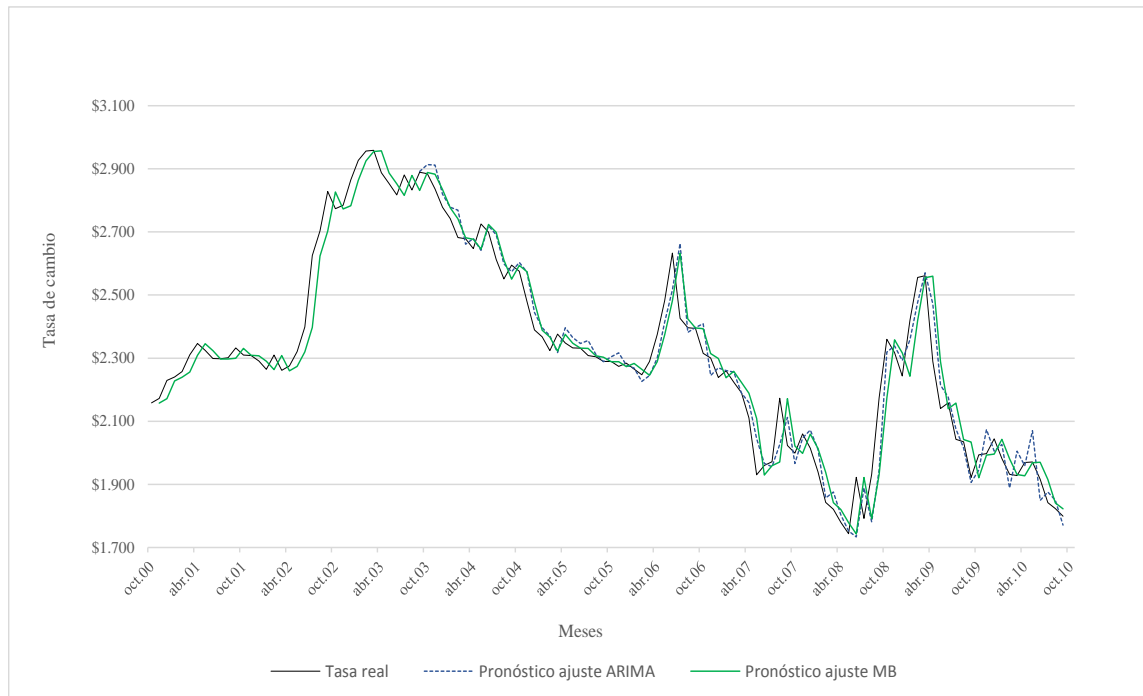
Fuente: Elaboración propia

Gráfica 4. Movimiento browniano - Tasa real vs. Pronósticos



Fuente: Elaboración propia

Gráfica 5. Pronóstico de ajuste (Tasa real vs Pronósticos ARIMA y Movimiento Browniano)



Fuente: Elaboración propia

7.2. Las medidas de evaluación de pronósticos

Una vez realizada la modelación de la tasa de cambio y el pronóstico de la misma a 60 meses durante el periodo de 2010M10 a 2015M09 con el *Rolling* a uno y hasta seis meses (pronósticos con $h = 1$ hasta $h = 6$) se realizaron los cálculos de las medidas. Se incluyó igualmente el pronóstico de ajuste dentro de la muestra (*in-sample*).

Para el cálculo de las medidas se tuvieron en cuenta las siguientes consideraciones:

- Para aquellos casos en los que la medida introduce implícitamente una comparación con un pronóstico de referencia, se utilizó el segundo modelo: La referencia es el mismo modelo ARIMA cuando se realizaron los cálculos para el movimiento browniano y viceversa, es decir, se comparan entre ellos.
- El cálculo de las medidas para el horizonte $h = 1$ tiene en cuenta 60 datos (periodo 2010M10 a 2015M09). El horizonte $h = 2$ pierde un dato y se calcula con base en 59 datos y así

sucesivamente para $h = 3, \dots, 6$. Esto se da precisamente por el pronóstico tipo *Rolling* en el cual cada paso hacia adelante a partir del primer mes implica la pérdida de un dato de pronóstico.

- Así mismo las medidas escaladas que contienen en el denominador el término base ($\sum_{i=2}^n |y_i - y_{i-1}|$), pierden desde el horizonte $h = 1$ un dato puesto que el primer cálculo puede darse desde $i=2$. Cada vez que se avanza a $h = 2, \dots, 6$, la medida pierde un dato adicional por cada periodo de pronóstico.

Los resultados para el modelo ARIMA se presentan en la Tabla 9. En la misma se listan cada una de las 17 medidas y los dos estadísticos U Theil de con su respectivo resultado para cada uno de los horizontes $h = 1, \dots, 6$.

Tabla 9. Resultados medidas de evaluación de pronóstico para el modelo ARIMA

Medida	Pronóstico para h=1	Pronóstico para h=2	Pronóstico para h=3	Pronóstico para h=4	Pronóstico para h=5	Pronóstico para h=6	Pronóstico de ajuste
MSE	8.908,21	21.503,08	36.214,85	49.266,76	58.843,60	56.228,51	5.052,13
RMSE	94,38	146,64	190,30	221,96	242,58	237,13	71,08
MAE	72,36	108,29	134,65	155,06	168,97	165,83	52,38
MdAE	54,13	84,08	91,14	103,28	120,59	128,82	41,30
MAPE	3,51%	5,07%	6,23%	7,17%	7,82%	7,58%	2,38%
MdAPE	2,69%	4,36%	4,96%	5,48%	6,60%	6,66%	1,85%
RMSPE	4,31%	6,31%	7,91%	9,19%	10,15%	10,06%	3,25%
RMdSPE	2,69%	4,36%	4,96%	5,48%	6,60%	6,66%	1,85%
sMAPE	3,53%	5,17%	6,44%	7,46%	8,16%	7,94%	2,38%
sMdAPE	2,70%	4,34%	4,91%	5,47%	6,62%	6,89%	1,83%
MRAE	270,15%	239,78%	517,59%	325,60%	281,47%	169,03%	369,52%
MdRAE	128,18%	143,69%	141,03%	129,01%	123,20%	117,86%	91,47%
GMRAE	133,52%	133,58%	165,23%	137,81%	127,72%	91,95%	91,73%
MSSE	251,09%	615,24%	1002,49%	1360,48%	1602,97%	1490,90%	123,47%
RMSSE	158,46%	248,04%	316,62%	368,85%	400,37%	386,12%	111,12%
MASE	121,49%	183,17%	224,03%	257,67%	278,88%	270,02%	81,88%
MdASE	90,87%	142,22%	151,64%	171,62%	199,04%	209,77%	64,56%
U Theil 1	0,023	0,036	0,048	0,056	0,061	0,060	0,016
U Theil 2	1,191	1,170	1,154	1,145	1,150	1,092	0,813

Fuente: Elaboración propia

De la misma manera, los resultados para el movimiento browniano se presentan en la Tabla 10.

Tabla 10. Resultados medidas de evaluación de pronóstico para el modelo de movimiento browniano

Medida	Pronóstico para h=1	Pronóstico para h=2	Pronóstico para h=3	Pronóstico para h=4	Pronóstico para h=5	Pronóstico para h=6	Pronóstico de ajuste
MSE	6.823,37	16.993,67	29.183,34	39.668,36	47.730,28	52.254,49	6.537,80
RMSE	82,60	130,36	170,83	199,17	218,47	228,59	80,86
MAE	59,25	89,42	109,20	130,67	146,15	153,45	57,95
MdAE	43,19	56,74	61,20	90,26	96,08	93,99	42,88
MAPE	2,79%	4,09%	4,91%	5,89%	6,54%	6,79%	2,55%
MdAPE	2,16%	3,05%	3,33%	4,68%	4,98%	5,34%	1,85%
RMSPE	3,62%	5,39%	6,86%	8,03%	8,83%	9,21%	3,61%
RMdSPE	2,17%	3,05%	3,33%	4,68%	4,98%	5,34%	1,85%
sMAPE	2,82%	4,18%	5,09%	6,14%	6,86%	7,17%	2,55%
sMdAPE	2,18%	3,01%	3,39%	4,57%	4,99%	5,20%	1,83%
MRAE	175,83%	145,28%	87,99%	111,42%	136,30%	342,50%	303,20%
MdRAE	78,18%	69,60%	70,92%	77,51%	81,18%	84,85%	109,33%
GMRAE	74,89%	74,86%	60,52%	72,56%	78,30%	108,75%	109,01%
MSSE	192,33%	486,22%	807,85%	1095,42%	1300,23%	1385,53%	159,78%
RMSSE	138,68%	220,50%	284,23%	330,97%	360,59%	372,23%	126,41%
MASE	99,48%	151,26%	181,69%	217,14%	241,22%	249,87%	90,61%
MdASE	72,52%	95,97%	101,82%	149,98%	158,58%	153,05%	67,04%
U Theil 1	0,020	0,032	0,043	0,050	0,055	0,058	0,017
U Theil 2	0,839	0,854	0,867	0,874	0,870	0,916	1,231

Fuente: Elaboración propia

En general, de acuerdo a lo observado en las tablas 9 y 10, se evidencia que el desempeño del pronóstico para la serie TCRM favorece la precisión del modelo de movimiento browniano comparado con el modelo ARIMA, salvo para el pronóstico de ajuste. Se infiere entonces que el modelo ARIMA se ajusta a la muestra *in-sample* de mejor manera debido al proceso mismo de selección de los términos autorregresivos y de medias móviles. De igual manera se observa, para los dos modelos, que cada vez que se pierde un dato en los diferentes horizontes de tiempo, $h = 1, \dots, 6$, el valor pronosticado se aleja del valor real.

Las medidas dependientes de escala cuyas cifras se presentan en valores numéricos o intervalos, suelen tener un criterio de medición más simple, al utilizarse entre otros como se puede apreciar

en (12) a (15), el promedio de los errores, la raíz cuadrada de la media de los errores al cuadrado, el promedio de los valores absolutos de los errores calculados o la mediana de los valores absolutos de los errores calculados respectivamente. Su interpretación indica que entre más pequeño sea el resultado de la medida más se ajusta el estimador a los datos reales.

Sin embargo, en el desarrollo del ejercicio se observa que los valores aumentan en la medida que el horizonte de pronóstico avanza, lo cual puede indicar que estas medidas son sensibles a valores atípicos y a la incertidumbre de pronosticar a periodos mas largos. Por ejemplo, para la medida MSE en el modelo de movimiento browniano los valores inician con 6823,37 en $h = 1$ a 52 254,49 en $h = 6$ y en ARIMA inicia con 8908,21 en $h = 1$ y su resultado en $h = 6$ es 56 228,51, es decir la cifra se incrementa 7,7 y 6,3 veces respectivamente.

Los resultados de las medidas basadas en porcentajes que buscan evaluar el tamaño del error en términos porcentuales y que se expresan en (16) a (20) y (23), indican la bondad en su formulación al no depender de la escala de datos ni ser tan influenciadas por valores atípicos precisamente por su denominador. Por ejemplo la medida MdAPE, que hace referencia a la mediana de los errores porcentuales, arroja valores en $h = 1$ de 2,16% para el movimiento browniano y 2,69% para ARIMA.

En efecto, medidas como MdAPE y RMdSPE que muestran unos menores porcentajes de error, validan las apreciaciones de Armstrong y Collopy (1992), quienes aconsejan la utilización de MdAPE, por ejemplo, porque adiciona facilidades de medición de factores importantes en las series de tiempo.

Para las medidas basadas en errores relativos en (24), (26) y (27), los cálculos incorporan, la media aritmética, la mediana o la media geométrica de los errores absolutos de un modelo, dividido entre los errores absolutos del modelo de referencia de tal manera que cualquier cambio sustancial en el denominador afecta significativamente el resultado. Prueba de ello se evidencia en la medida MRAE, para ARIMA, la cual presenta disminuciones de $h = 1$ a $h = 2$, un aumento considerable en $h = 3$ y disminución en los tres últimos horizontes de tiempo $h = 4$ a $h = 6$.

Algo similar se puede evidenciar con las medidas escaladas definidas en (28) a (31) puesto que se encuentran menores porcentajes. Es importante recordar, que valores de MASE mayor que uno indican que las estimaciones son menos asertivas, en promedio, que dentro de la muestra de un solo paso por estimaciones del método o modelo ingenuo. En ese sentido el resultado para el modelo de movimiento browniano en $h = 1$ cuyo valor es de 0,99, estuvo por debajo del referente, lo cual valida este argumento.

Adicionalmente, los resultados del ejercicio son consecuentes con los estudios de Hyndman y Koehler (2005) quienes sostienen que escalando el error absoluto calculado de la medida MAE, por ejemplo, mediante el uso de un pronóstico base redundará en un menor valor que aquel obtenido del pronóstico base. Siguiendo con esta teoría, los resultados también evidencian que el uso de la medida MASE es menos sensible a valores atípicos cuando se comparan los resultados con los de la medida RMSSE y menos variable en muestras pequeñas que MdASE.

Por último, con respecto a los estadísticos U de Theil, los resultados muestran igualmente mejores cifras para el modelo de movimiento browniano que para el ARIMA. En efecto el estadístico U1 en (32) y que hace referencia al criterio de exactitud del pronóstico, presenta márgenes de error muy bajos, aun cuando se pierde un dato en los diferentes horizontes de tiempo, $h = 1, \dots, 6$. Los datos oscilan entre 0,020 y 0,058 para el modelo de movimiento browniano y 0,023 y 0,060 para el ARIMA, valores todos por debajo de uno que se traduce en que el modelo de movimiento browniano es más preciso que el modelo ARIMA.

El estadístico U2 que hace referencia a la calidad, favorece nuevamente al modelo de movimiento browniano sobre el modelo ARIMA. Se observa que los pronósticos en el modelo Browniano, en los diferentes horizontes de tiempo estan por debajo de uno. Por ejemplo $h = 1$ para movimiento browniano es de 0,839 contra un 1,191 en ARIMA.

7.3. Los test de comparación de pronósticos

Utilizando las fórmulas presentadas en (37), (44), (53) y (56) se realizaron los cálculos matemáticos para obtener los estadísticos de cada uno de los *test*. Para el caso específico del *test* DM así como su correspondiente modificación HLN-DM se trabajaron las funciones de pérdida cuadrática y absoluta. Así mismo, importante tener en cuenta que cada uno de los estadísticos se calcularon para los horizontes $h = 1, \dots, 6$ así como la el pronóstico de ajuste.

La Tabla 11 presenta los resultados de los seis estadísticos trabajados así como los valores de referencia con los cuales se comparán los mismos, estos son, el t^* y el valor-p (*p-value*) para las distribuciones *t-student* y normal respectivamente. La validación de la hipótesis nula H_0 , en la que los modelos tienen la misma exactitud, se incluye en esta misma tabla para cada uno de los grados de significancia $\alpha = 1\%, 5\% \text{ o } 10\%$ (NRHo corresponde a no rechazar la hipótesis nula H_0 mientras que RHo si lo hace).

Se puede observar que para todos los casos a excepción del pronóstico de ajuste, los estadísticos de MGN y HLB-MGN son positivos. Los vectores ortogonales x_t y z_t presentados en (34) y (35) respectivamente, incluyen los términos de error en donde e_{1t} corresponde al modelo ARIMA y e_{2t} al movimiento browniano. Ahora, si se revisa (37), el posible cambio de signo del estadístico depende de su numerador. Si se analiza a su vez la expresión de éste en (38), el cambio de signo estará dado por el término $x'z^{46}$, lo cual permite concluir que al ser positivo en todos los horizontes $h = 1, \dots, 6$, el modelo ARIMA presenta mayores errores que el movimiento browniano.

La conclusión anterior en parte se corrobora haciendo un análisis similar con el estadístico DM y el HLN-DM en donde los resultados son números positivos con la excepción del horizonte $h = 3$. En este caso el diferencial expresado en (41) tiene como referencia igualmente que e_{1t} corresponde al modelo ARIMA y e_{2t} al movimiento browniano y es precisamente el promedio de este diferencial el que define el signo del estadístico en (44). Al ser positivo, nuevamente lleva a la conclusión que el movimiento browniano se acomoda mejor a la serie de la TCRM.

⁴⁶ Si se observa (51) y (53) para el estadístico HLN-MGN, el término $x'z$ es igualmente quien define el signo.

La excepción a la regla mencionada está relacionada con los resultados obtenidos en el pronóstico de ajuste. Obsérvese que todos los estadísticos tienen signo negativo de tal forma que para los datos dentro de la muestra (*in-sample*), el modelo ARIMA tiene mejor precisión desde el punto de vista de la magnitud de los errores, que el movimiento browniano. A pesar que esto es coherente con las bondades que tienen este tipo de modelos, no resulta útil construir un modelo que se ajuste con los datos dentro de la muestra y no tan bien fuera de ella.

Por otro lado, en cuanto a la validación de la hipótesis nula H_0 , nótese que para los pronósticos de ajuste, los modelos resultan con idéntica exactitud y precisión bajo todos los estadísticos así como para todos los grados de significancia $\alpha = 1\%$, 5% o 10% . Para los pronósticos fuera de la muestra, se observan algunas similitudes con los resultados de los estadísticos con $\alpha = 5\%$ y $\alpha = 10\%$ para los horizontes $h = 1, 2, 4, 5$. En los demás casos, los *test* arrojan diferentes conclusiones con respecto a la aceptación o no de la hipótesis nula H_0 .

Tabla 11. Resultados de los *test* y validación de hipótesis nula

<i>Test</i>	Pronóstico para h=1	Pronóstico para h=2	Pronóstico para h=3	Pronóstico para h=4	Pronóstico para h=5	Pronóstico para h=6	Pronóstico de ajuste
MGN	2,510**	2,652**	2,796***	2,857***	2,643**	0,928****	(3,618)****
DM cuadrática	2,108**	7,991***	(1,748)****	2,542**	3,624***	1,826*	(3,019)****
DM absoluta	2,960***	3,549***	(0,232)****	5,278***	5,139***	3,081***	(2,661)****
HLN-MGN	2,535**	3,576***	3,403***	2,666**	2,360**	1,147****	(0,398)****
HLN-DM cuadrática	2,091**	7,788***	(1,673)****	2,286**	3,332***	1,643****	(3,001)****
HLN-DM absoluta	2,935***	3,458***	(0,222)****	4,954***	4,726***	2,773***	(2,646)****

Nota: * es significativo al 10%, ** es significativo al 5%, *** es significativo al 1% y **** no hay significancia o igualdad de precisión en el pronóstico.

Fuente: Elaboración propia

7.4. Conclusiones

La Tabla 12 presenta las características de las medidas de precisión y complementa la Tabla 5 del Capítulo 4. Se adicionaron las columnas de sensibilidad a valores atípicos (medido en una escala de alta, media y baja) y la de uso conjunto. Precisamente esta última tiene como objetivo aconsejar al lector a no usar una única medida de precisión en el momento de las valoraciones sino a usarlas en conjunto.

Por ejemplo el uso de MAE, que al criterio de los autores es una medida poco útil por si sola, podría complementarse con MAPE o MRAE. Lo anterior, con el propósito de tener la información de la diferencia en unidades que arroja MAE con la dispersión porcentual que resulta con MAPE o MRAE. La información final puede apreciarse en la Tabla 12.

Tabla 12. Características de las medidas de precisión

Medida	Escala	Unidades	Función de pérdida	Simetría en función de pérdida	Sensibilidad a valores atípicos	Usar en conjunto con
MSE	Intervalo	Datos al cuadrado	Cuadrática	Si	Alto	-
RMSE	Intervalo	Datos	Cuadrática	Si	Alto	MAPE / MRAE
MAE	Intervalo	Datos	Absoluta	Si	Medio	MAPE / MRAE
MdAE	Intervalo	Datos	Absoluta	Si	Medio	-
MAPE	Porcentaje	No tiene	Absoluta	No	Bajo	MSE / MRAE
MdAPE	Porcentaje	No tiene	Absoluta	Si	Bajo	-
RMSPE	Porcentaje	No tiene	Cuadrática	Si	Alto	-
RMdSPE	Porcentaje	No tiene	Cuadrática	Si	Alto	-
sMAPE	Porcentaje	No tiene	Absoluta	Si	Bajo	-
sMdAPE	Porcentaje	No tiene	Absoluta	Si	Bajo	-
MRAE	Porcentaje	No tiene	Absoluta	No	Bajo	MSE / MAPE
MdRAE	Porcentaje	No tiene	Absoluta	No	Bajo	-
GMRAE	Porcentaje	No tiene	Absoluta	No	Bajo	-
MSSE	Porcentaje	No tiene	Cuadrática	No	Alto	-
RMSSE	Porcentaje	No tiene	Cuadrática	No	Alto	-
MASE	Porcentaje	No tiene	Absoluta	No	Bajo	-
MdASE	Porcentaje	No tiene	Absoluta	No	Bajo	-
U Theil 1	Porcentaje	No tiene	Cuadrática	No	Alto	-
U Theil 2	Intervalo	No tiene	Cuadrática	Si	Alto	MSE / MAPE

Fuente: Elaboración propia

Aunado a lo anterior, la Tabla 13 complementa la Tabla 6 del Capítulo 4. La información allí consignada se considera a partir de los resultados obtenidos y analizada con criterios más amplios. Si bien se utilizan los estadísticos y los de usuario, en el primero se han incluido los seis criterios de aceptación descritos en el aparte 4.1. En cuanto a los usuario, se incluyeron dos, la calidad de información que entrega cada medida así como la facilidad de entender el resultado⁴⁷.

Tabla 13. Desempeño de las medidas de precisión

Medida	Criterio estadístico						Criterio de usuario	
	Confiabilidad	Discriminación	Validación	Ajuste	Exactitud	Precisión	Calidad de información	Facilidad de entendimiento
MSE	Bajo	Alto	Bajo	Bajo	Bajo	Bajo	Medio	Bajo
RMSE	Bajo	Medio	Medio	Bajo	Bajo	Bajo	Alto	Bajo
MAE	Bajo	Alto	Alto	Bajo	Bajo	Bajo	Alto	Alto
MdAE	Bajo	Alto	Bajo	Bajo	Bajo	Bajo	Medio	Medio
MAPE	Bajo	Alto	Alto	Medio	Medio	Bajo	Alto	Alto
MdAPE	Bajo	Medio	Bajo	Medio	Medio	Bajo	Medio	Medio
RMSPE	Bajo	Bajo	Bajo	Medio	Medio	Bajo	Alto	Bajo
RMdSPE	Bajo	Bajo	Bajo	Medio	Medio	Bajo	Medio	Bajo
sMAPE	Medio	Alto	Alto	Medio	Medio	Medio	Alto	Medio
sMdAPE	Medio	Medio	Bajo	Medio	Medio	Medio	Medio	Medio
MRAE	Medio	Alto	Alto	Medio	Medio	Medio	Alto	Medio
MdRAE	Medio	Alto	Bajo	Medio	Medio	Medio	Alto	Medio
GMRAE	Medio	Medio	Medio	Medio	Medio	Medio	Medio	Medio
MSSE	Bajo	Bajo	Bajo	Alto	Alto	Bajo	Bajo	Bajo
RMSSE	Bajo	Bajo	Bajo	Alto	Alto	Bajo	Bajo	Bajo
MASE	Bajo	Bajo	Medio	Alto	Alto	Bajo	Bajo	Bajo
MdASE	Bajo	Bajo	Bajo	Alto	Alto	Bajo	Bajo	Bajo
U Theil 1	Bajo	Medio	Bajo	Medio	Medio	Bajo	Bajo	Bajo
U Theil 2	Alto	Alto	Alto	Medio	Medio	Alto	Alto	Medio

Fuente: Elaboración propia

La confiabilidad y la precisión parecen estar más relacionados con los modelos mismos que con las medidas. Se podría pensar que muchas medidas son confiables y precisas⁴⁸ al ser muy similares, sin embargo, los resultados mayormente incrementales en función del horizonte h , contrastan con lo anterior. Aquí el estadístico U de Theil fue el más confiable y preciso en los dos modelos porque “contrarresta” la función de pérdida cuadrática con la raíz cuadrada en (33) y por

⁴⁷ Cada uno de los criterios fue calificado en una escala de alto, medio y bajo, donde este último significa la peor calificación.

⁴⁸ Por ejemplo cuando se revisan los resultados del pronóstico $h = 1$ y del pronóstico de ajuste en el caso del movimiento browniano.

incluir la comparación con un modelo de referencia. Con respecto a esto último, se concluye que los errores tanto del modelo inicial como del de referencia van incrementando de manera proporcional.

El uso de un modelo de referencia podría entenderse como un factor que mejora el desempeño de las medidas puesto que al tener en el denominador el valor observado (medidas PE⁴⁹) o a un modelo de referencia (medidas RAE⁵⁰ y U de Theil) conlleva consigo una reducción del sesgo que puede generar los valores atípicos.

Por su parte, varias medidas sobresalen por su discriminación y esto es independiente de la presencia de escala: MAE, MAPE, MRAE y el mismo U de Theil son ejemplos. El desempeño de las medidas para este criterio no es influenciado por su formulación matemática, el tamaño de la muestra ni los valores atípicos. Sin embargo, si lo es, para el caso de las medidas escaladas. En la medida que la escala⁵¹ aumenta, la posibilidad de diferenciar disminuye, así, series con valores atípicos muy altos podrían no ser propicias para ser evaluadas con estas medidas.

El análisis precedente difiere en el caso del criterio de validación. Si la perspectiva que se adopta para revisar este criterio es similar al de la facilidad con que se entiende la información de cada medida, MAE, MAPE, sMAPE, MRAE y el U de Theil tendrían la más alta calificación. El desempeño de la validez es visto aquí como la intuitividad de cada medida.

Para el ajuste y la exactitud, las calificaciones más altas las tienen aquellas medidas que realizan una comparación, esto es, las porcentuales, las relativas y las escaladas. Estas últimas precisamente por tener una constante en su denominador son las de valor mínimo más aun cuando se presentan valores atípicos.

Las conclusiones derivadas de los criterios estadísticos son relevantes cuando los usuarios tienen un conocimiento básico de su significado. No obstante, la afirmación anterior podría verse avocadas a un escenario gerencial donde quizás prime la practicidad. En consecuencia, la calidad

⁴⁹ Las medidas PE son descritas en el numeral 4.4.3 y son MAPE, MdAPE, RMSPE, RMdSPE, sMAPE, sMdAPE.

⁵⁰ Las medidas RAE son descritas en el numeral 4.4.4 y son MRAE, MdRAE y GMRAE.

⁵¹ La escala es el denominador de las medidas y que corresponde a $\frac{1}{n-1} \sum_{i=2}^n |y_i - y_{i-1}|$.

de información y la facilidad de entendimiento de los números y porcentajes que arrojan las medidas adquieren una alta importancia. Las evidencias de la investigación permiten inferir que las medidas absolutas son más fáciles de entender que aquellas cuadráticas por su formulación. Por su parte, las medidas porcentuales y relativas por si solas entregan ya una referencia que avalan o no la comparación misma y por ende, el valor de la información que resulta de sus cálculos.

En contraposición a las medidas, si bien el estadístico calculado de los *test* de evaluación permiten inducir cual modelo se ajusta mejor, su fin inicial es determinar si dos modelos tienen o no la misma precisión. Fundamentos en cada *test* como la ortogonalización en el MGN, la pérdida diferencial en el DM, las correcciones a los sesgos por HLN y la selección de una u otra distribución de probabilidad, son los mecanismos que finalmente corroboran la conclusión a la que se llegó con los resultados de las medidas de precisión, el modelo browniano se ajusta mejor que el ARIMA en los datos fuera de la muestra pero no dentro de la misma.

Al final lo mencionado evidencia que una simulación de Montecarlo para una variable aleatoria con 10.000 repeticiones bajo un movimiento geométrico browniano, explica de manera más precisa la serie, que un modelo que contempla el histórico en términos autorregresivos y de medias móviles de diferentes rezagos. Sin embargo, no es intuitiva la deducción anterior con los estadísticos calculados en cada uno de los *test*.

No deja de llamar la atención el hecho que no hay ningún patrón en la aceptación o rechazo de H_0 en la medida que los horizontes avanzan para los diferentes grados de significancia (ver por ejemplo el patrón que se rompe en $h = 3$ y $h = 6$ para $\alpha = 5\%$ o 10%). Para los *test* MGN y HLN-MGN es el término $x'z$ el que condiciona el estadístico. Si bien los errores van en aumento a medida que el horizonte crece, es probable que cuando se generan los vectores, los errores en promedio se cancelen entre ellos mismos lo cual conlleva a no rechazar la H_0 .

Ahora bien, cabe la pena anotar que las correcciones introducidas por Harvey, Leybourne y Newbold (1997) para los *test* MGN y DM pueden validarse en la Tabla 11 en aquellos casos que el estadístico modificado no rechaza H_0 habiéndolo hecho el estadístico inicial. Las correcciones

del *test* HLN-DM sobre el DM se validan fácilmente con la disminución de los valores del estadístico. En el caso de MGN esto solo es remarcable para ciertos horizontes. En conclusión, las modificaciones de los *test* HLN, los hace más adecuados para la evaluación de la serie estudiada.

Por otro lado, por lo que se refiere a la funcionalidad de los *test*, la cual es determinar si dos modelos son igualmente precisos, es importante tener en cuenta que adicionar los resultados de un grupo de tres a cuatro medidas junto a la valoración de los estadísticos calculados, permitirá inferir suficiente información para seleccionar el modelo que mejor se ajuste a la variable en estudio. Siempre será indicado trabajar con el modelo más sencillo para efectos de pronóstico.

Para terminar, la presente investigación puede ser el punto inicial a trabajos complementarios que a juicio de los autores podrían estar enfocados en dos frentes. El primero con el análisis de las medidas de precisión para un número mayor de modelos de pronóstico. Sería interesante ver los resultados cuando se incluyen modelos no lineales como los de la familia de transición sobre los parámetros; los modelos polinomiales y los modelos de redes neuronales.

El segundo con la inclusión de nuevos *test* de evaluación de pronósticos como podrían ser los de West (1996); Giacomini y White (2006) y; Clark y McCracken (2012). Incluso si se piensa en la unión de los dos planteamientos, podría resultar un gran compendio de clasificación de medidas de precisión y de *test* de evaluación de pronóstico.

Bibliografía

- Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions On Automatic Control*, 19, 716-723.
- Aldrich, J. (1997). R. A. Fisher and the Making of Maximum Likelihood 1912-1922. *Statistical Science*, 12(3), 162-176.
- Alvarez, J., Corrales, H., González, Y., Lorenzo, C., Prieto, M., & Zarzosa, P. (2011). *Material Docente de Econometría*. Obtenido de Universidad de Valladolid: http://www2.eco.uva.es/econometria-ADE/material_2011-12/material_teoría1_2011-12.pdf
- Arango, L., & Flórez, L. (2004). Expectativas de Actividad Económica en Colombia y Estructura a Plazo: Un Poco Más de Evidencia. *Borradores de Economía Banco de la República*.
- Aristizábal, M. C. (2006). Evaluación asimétrica de una red neuronal artificial: Aplicación al caso de la inflación en Colombia. *Borradores de economía Banco de la República*.
- Armstrong, J. (2001). Evaluating Forecasting Methods. En J. Armstrong, *Principles of Forecasting: A Handbook for Researchers and Practitioners* (pág. Capítulo 14). Norwell: Kluwer Academic Publishers.
- Armstrong, J. (2001). *Principles or Forecasting*. Dordrecht: KLUWER ACADEMIC PUBLISHERS.
- Armstrong, J. S. (1985). *Long-range Forecasting: From Crystal Ball to Computer*. New York: John Wiley.
- Armstrong, J., & Collopy, F. (1992). Error Measures for Generalizing about Forecasting Methods: Empirical Comparisons. *International Journal of Forecasting*(8), 69-80.
- Arrow, K., & Intriligator, M. (2006). *Handbook of Economic Forecasting*. San Diego, California, Estados Unidos: North Holland.
- Ashley, R. (1983). On the Usefulness of Macroeconomic Forecasts As Inputs to Forecasting Models. *Journal of Forecasting*, 2, 211-223.
- Ashley, R. (1988). On the Relative Worth of Recent Macroeconomic Forecasts. *International Journal of Forecasting*, 4, 363-376.
- Ayala, M., & Castillo, R. (2008). Un Modelo de Predicción para el Valor TRM: Un Acercamiento desde las Redes Neuronales Artificiales. Obtenido de http://www.usa.edu.co/observatorio_economico/articulos_profesores/trm.pdf

- Baechle, T. R., & Earle, R. W. (2007). *Principios del entrenamiento de la fuerza y el acondicionamiento físico*. Madrid: Panamericana.
- Barbancho, A. (1976). *Fundamentos y Posibilidades de la Econometría*. Barcelona: Ariel.
- Bates, J., & Granger, C. (1969). The Combination of Forecasts. *20*, 451-468.
- Billah, B., King, M., & Snyder, R. (2005). Exponential smoothing model selection for forecasting. En *Working Paper 6/05*. Monash University, Australia.
- Bliemel, F. (1973). Theil's Forecast Accuracy Coefficient: A Clarification. *Journal of Marketing Research*, *10*, 444-446.
- Bowerman, B., O'Connell, R., & Koehler, A. (2004). *Forecasting, Time Series and Regression: An Applied Approach*. Belmont, California, Estados Unidos: Thomson Brooks/Cole.
- Box, G., & Jenkins, G. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
- Burnham, K., & Anderson, D. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York: Springer-Verlag.
- Caballero, F. (2011). *Selección de Modelos mediante Criterios de Información en Análisis Factorial. Aspectos Teóricos y Computacionales*. Granada, España: Selección de Modelos mediante Criterios de Información en Análisis Factorial. Aspectos Teóricos y Computacionales.
- Cajiao, S., Melo, L., & Parra, D. (2014). Pronósticos para una Economía menos Volátil: El Caso Colombiano. *Coyuntura Económica: Investigación Económica y Social*, 35-59.
- Capistrán, C. (2010). Las expectativas macroeconómicas, una evaluación de pronósticos de corto plazo. *El trimestre económico*, 275-312.
- Carbone, R., & Armstrong, J. (1982). Evaluation of Extrapolative Forecasting Methods: Results of a Survey of Academicians and Practitioners. *Journal of Forecasting*, *1*, 215-217.
- Castaño, E., & Melo, L. (1998). Métodos de Combinación de Pronósticos: Una Aplicación a la Inflación Colombiana. *Borradores de Economía Banco de la República*.
- Cavanaugh, J. (23 de October de 2012). *The University of Iowa*. Obtenido de Model Selection: Criteria for Time Series: http://myweb.uiowa.edu/cavaugh/ms_lec_9_ho.pdf
- Chai, T., & Draxler, R. (2014). *Root mean square error (RMSE) or mean absolute error (MAE)*. Copernicus Publications on behalf of the European Geosciences Union.

- Chai, T., & Draxler, R. (2014). Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)? – Arguments Against Avoiding RMSE in the Literature. *Geoscientific Model Development*, 7, 1247-1250.
- Chakravarty, S. (1959). *The Logic of Investment Planning*. Amsterdam: North Holland Publishing.
- Chan, K., & Tong, H. (1986). On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, 7, 179-194.
- Chase, C. (2009). *Demand-Driven Forecasting: A Structured Approach to Forecasting*.
- Chavarría Izos, A. R. (23 de 06 de 2011). Obtenido de Blogspot: Recuperado el 08 de abril de 2016 de <http://exactitudyprecision.blogspot.com.co/2011/06/exactitud-y-precision-quimica.html>
- Clark, T., & McCracken, M. (2010). Testing for Unconditional Predictive Ability. *Research Division Federal Reserve Bank of St. Louis - Working Papers*.
- Clark, T., & McCracken, M. (2012). Nested Forecast Model Comparisons: A New Approach to Testing Equal Accuracy. *Research Division Federal Reserve Bank of St. Louis - Working Papers*.
- Clements, M., & Hendry, D. (1999). *Forecasting Non-stationary Economic Time Series*. MIT Press.
- Clements, M., Joutz, F., & Stekler, H. (2007). An Evaluation of the Forecast of the Federal Reserve: A Pooled Approach. *Journal of Applied Econometrics*, 22, 121-136.
- Council, N. R. (1980). *Estimating Population and Income for Small Places*. Washington, D.C.: National Academy Press.
- Cristiano, D., Hernández, M., & Pulido, J. D. (2012). Pronósticos de corto plazo en tiempo real para la actividad económica colombiana. *Borradores de Economía*.
- de Arce, R., & Mahia, R. (s.f.). *Modelos Arima: Definiciones Básicas*. Obtenido de Universidad Autónoma de Madrid: https://www.uam.es/personal_pdi/economicas/anadelsur/pdf/Box-Jenkins.PDF
- De Luna, E. (1996). Epistemología de la investigación taxonómica: inferencias filogenéticas y su evaluación. *Boletín de la sociedad Botánica de México*.
- Delicado, P. (2008). *Curso de Modelos no Paramétrico*. Barcelona, España: Universitat Politècnica de Catalunya.
- Deutsch, M., C.W., G., & Teräsvirta, T. (1994). The combination of Forecasts Using Changing Weights. *International Journal of Forecasting*, 10, 47-57.

- Diebold, F. (2013). Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold-Mariano Test. *Journal of Business and Economic Statistics Invited Lecture*.
- Diebold, F., & Mariano, R. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, 253-265.
- Donaldson, R., & Kamstra, M. (1996). Forecast Combining with Neural Networks. *Journal of Forecasting*, 15, 49-61.
- Dornbusch, R. (1976). Expectations and Exchange Rate Dynamics. *Journal of Political Economy*, 84, 1161-1176.
- Dufour, J.-M. (2011). Coefficients of Determination. *McGill University*.
- EAFIT. (7 de Julio de 2014). *Prospectiva en Marketing: el marketing del mañana se puede anticipar*. Obtenido de http://www.eafit.edu.co/cec/programas/Paginas/prospectiva-marketing.aspx#.VHJpx_mG-So
- Elliott, G., & Timmermann, A. (2007). Economic Forecasting.
- Fayad, C., Fortich, R., & Vélez, I. (2009). Proyección de la Tasa de Cambio de Colombia bajo Condiciones de PPA: Evidencia Empírica usando VAR. *Estudios Gerenciales*, 25(113), 211-226.
- Fildes, R. (1992). The Evaluation of Extrapolative Forecasting Methods. *International Journal of Forecasting*(8), 81-98.
- Flores, B. (1986). A Pragmatic View of Accuracy Measurement in Forecasting. En *Omega The International Journal of Management Science* (Vol. 14, págs. 93-98). Omega.
- Frankel, J. A. (1979). On the Mark: A Theory of Floating Exchange Rates Based on Real Interest Differentials. *American Economic Review*, 69, 610-622.
- Giacomini, R., & Rossi, B. (2013). Forecasting in Macroeconomics . En N. Hashimzade, & M. Thornton, *Handbook of Research Methods and Applications on Empirical Macroeconomics* (págs. 381-480). Cheltenham, UK; Northampton, MA, USA: Edward Elgar Publishing.
- Giacomini, R., & White, H. (Noviembre de 2006). Test of Conditional Predictive Ability. *Econometrica*, 74(6), 1545-1578.
- González, M. d. (2009). *Análisis de Series Temporales: Modelos ARIMA*. Bilbao: Facultad de Ciencias Económicas y Empresariales Universidad del País Vasco.
- Granger, C. (1999). Outline of Forecast Theory using Generalized Cost Functions. *Spanish Economic Review*, 161-173.

- Granger, C. (1999). Outline of Forecast Theory using Generalized Cost Functions. *Spanish Economic Review*, 161-173.
- Granger, C., & Newbold, P. (1977). Forecasting Economic Time Series. *Academic Press*.
- Granger, C., & Teräsvirta, T. (1993). *Modelling nonlinear economic relationship*. New York: Oxford University Press.
- Gujarati, D., & Porter, D. (2010). *Econometría*. (M. Hill, Ed.) Ciudad de México D.F, México.
- Hanke, J. E., & Wichern, D. W. (2006). *Pronósticos en los negocios*. Pearson Prentice Hall.
- Hanke, J., & Reitsch, A. (1995). *Business Forecasting* (Quinta ed.). Englewood Cliffs, New Jersey: Prentice-Hall.
- Hannan, E. J., & Quinn, B. G. (1979). The Determination of the Order of an Autoregression. *Journal of the Royal Statistical Society*, 41(2), 190-195.
- Harvey, D., Leybourne, S., & Newbold, P. (Junio de 1997). Testing the Equality of Predicting Mean Squared Errors. *International Journal of Forecasting*, 13(2), 281-291.
- Hawkins, J. (1996). *Economic forecasting: History and Procedures*. Obtenido de The Australian Government - The Treasury: http://cprs.treasury.gov.au/documents/987/PDF/02_eco_forecasting.pdf
- Hoaglin, D., Mosteller, F., & Tukey, J. (1983). Introduction to More Refined Estimators. En D. Hoaglin, F. Mosteller, & J. Tukey, *Understanding Robust and Exploratory Data Analysis* (págs. 283-296). New York: John Wiley.
- Hyndman, R. (2006). Another Look at Forecast-Accuracy Metrics for Intermittent Demand. *Foresight*, 4, 43-46.
- Hyndman, R., & Koehler, A. (20 de Mayo de 2005). Another Look at Measures of Forecast Accuracy.
- Khim Sen, L., & Shitan, M. (2000). The Performance of AICC as an Orden Sepection Criterion in ARMA Time Series Models. *Pertanika Journal of Science*, 10(1), 25-33.
- Kullback, S., & Leibler, R. (1951). On Information and Sufficiency. *Annals of Mathematical Statistics*, 22, 79-86.
- La República. (1 de enero de 2015). En 2015, Bancóldex llega con recursos de \$60.000 millones para emprendedores.
- Lee, T.-H. (Marzo de 2007). *Loss Functions in Time Series Forecasting*. Obtenido de Universidad de California Riverside : <http://www.faculty.ucr.edu/~taelee/paper/lossfunctions.pdf>

- Levy, P., & Lemeshow, S. (1991). *Sampling of Populations: Methods and Applications*.
- Luukkonen, R., Saikkonen, P., & Teräsvirta, T. (1988). Testing linearity against smooth transition autoregressive models. *Biometrika*, 75(3), 491-499.
- Makridaki, S., & Hibon, M. (2000). The M3-Competition: Results, conclusions and implications. *International Journal of Forecasting*, 16, 451-476.
- Makridakis, S. (1993). Accuracy Measures: Theoretical and Practical Concerns. *International Journal of Forecasting*, 9, 527-529.
- Makridakis, S., & Hibon, M. (1979). Accuracy of Forecasting: An Empirical Investigation. *Journal of the Royal Statistical Society*, 142, 97-145.
- Makridakis, S., & Hibon, M. (1995). Evaluating Accuracy Measures. *INSEAD Working Paper*, 31.
- Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., . . . Winkler, R. (1982). The Accuracy of Extrapolation (time series) Methods: Results of a Forecasting Competition. *Journal of Forecasting*, 1, 111-153.
- Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K., & Simmons, L. (1993). The M2-Competition: A real-time judgmentally based forecasting study. *International Journal of Forecasting*, 9, 5-22.
- Makridakis, S., Wheelwright, S., & Hyndman, R. (1998). *Forecasting Methods and Applications* (Tercera ed.). New York: John Wiley & Sons.
- Mariano, R. (2000). Testing Forecast Accuracy.
- Meese, R., & Rogoff, K. (1988). Was it real? The exchange rate-interest differential. *The Journal of Finance*, 43, 933-948.
- Melo, L., & Núñez, H. (2004). Combinación de Pronósticos de la Inflación en Presencia de Cambios Estructurales. *Borradores de Economía Banco de la República*.
- Morgan, W. A. (1939-40). A test for significance of the difference between the two variances in a sample from a normal bivariate population. *Biometrika* 31, 13-19.
- Moskowitz, H., & Wright, G. (1982). *Investigación de Operaciones*. Ciudad de México D.F., México: Prentice Hall.
- Munir, A., & Misas, M. (2007). Evaluación de Pronósticos del Tipo de Cambio Utilizando Redes Neuronales y Funciones de Pérdida Asimétricas. *Revistas Colombiana de Estadística*, 30(1), 143-161.

- Muñoz San Miguel, J. (15 y 16 de Septiembre de 2005). *Universidad de Valencia*. Obtenido de Jornadas de ASEPUMA Asociación Española de Profesores Universitarios de Matemáticas para la Economía y la Empresa: http://www.uv.es/asepuma/XIII/comunica/comunica_37.pdf
- Núñez, H. (2005). Una Evaluación de los Pronósticos de Inflación en Colombia bajo el Esquema de "Inflación Objetivo"* . *Revista de Economía del Rosario*, 8, 151-185.
- Osborne, M. (1959). Brownian Motion in the Stock Market. *Operations Research*, 25, 145-173.
- Osborne, M. (1962). Periodic Structure in the Brownian Motion of Stock Prices. *Operations Research*, 10(3), 345-379.
- Patiño, C., & Alonso, J. (Julio de 2005). *Evaluación de Pronósticos para la Tasa de Cambio en Colombia*. Recuperado el 11 de Abril de 2015, de Universidad Icesi: http://www.icesi.edu.co/revistas/index.php/estudios_gerenciales/article/view/171/html
- Peña, D. (2002). *Análisis de Datos Multivariantes*. Madrid: McGraw-Hill.
- Prestwich, S., Rossi, R., Tarim, S., & Hnich, B. (2014). Mean-Based Error Measures for Intermittent Demand Forecasting. *International Journal of Production Research*, 52, 6782-6791.
- Rivera, H., González, H., Martínez, O., Domínguez, E., Romero, H., Fajardo, M., . . . Carvajal, M. (2008). *Protocolo para la Emisión de los Pronósticos Hidrológicos*. Bogotá, Colombia: Imprenta Nacional de Colombia - IDEAM.
- Rodríguez, N., & Siado, P. (2003). Un Pronóstico no Paramétrico de la Inflación Colombiana. *Revista Colombiana de Estadística*, 26, 89-128.
- Rosales, R., & Bonilla, J. (2006). *Introducción a la Econometría*. Bogotá DC, Colombia: Ediciones Uniandes.
- Ruiz Medina, M. I. (2011). *Políticas Públicas en salud y su impacto en el sdeguro popular en Culiacán, Sinaloa, México*.
- S., M. R., & Mariano, R. S. (2000). *Testing Forecast Accuracy*.
- Salazar. (2009). Evaluación de pronóstico de una red neuronal sobre el PIB en Colombia. *Borradores de Economía*.
- Samuelson, P. (1965). Rational Theory of Warrant Pricing. *Industrial Management Review*, 6(2), 13-39.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2), 461-464.
- Shewhart, W., & Wilks, S. (2011). *Time Series Analysis and Forecasting by Example*. Hoboken, New Jersey, Estados Unidos: John Wiley & Sons.
- Sims, C. (1980). Macroeconomics and Reality. *Econometrica*, 1-48.

- Stone, M. (1979). Comments on Model Selection Criteria of Akaike and Schwarz. *Journal of the Royal Statistical Society*, 41(2), 276-278.
- Swanson, D., Tayman, J., & Bryan, T. (2011). MAPE-R: A Rescaled Measure OF Accuracy for Cross-Sectional Forecasts. *Journal of Population Research*, 28, 225-248.
- Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89(425), 208-218.
- Teräsvirta, T., van Dijk, D., & Medeiros, M. (2004). Linear Models, Smooth Transition Autoregressions, and Neural Networks for Forecasting Macroeconomic Time Series: A reexamination. *PUC Rio Departamento de Economía*(485).
- Theil, H. (1958). *Economic Forecasts and Policy*. Amsterdam: North Holland Publishing Co.
- Theil, H. (1961). *Economic Forecasts and Policy* (Segunda ed.). Amsterdam: North-Holland.
- Theil, H. (1966). *Applied Economic Forecasting*. Chicago, Illinois, Estados Unidos: Rand McNally.
- Theil, H. (1966). *Applied Economic Forecasting*. Amsterdam: North Holland Publishing Co.
- Tinbergen, J. (1956). *Economic Policy: Principles and Design (Contributions to Economic Analysis)*. Amsterdam: North Holland.
- Torres Barrón, M. d. (Sin fecha). Pronósticos, una herramienta clave para la planeación de las empresas. *Instituto Tecnológico de Sonora*.
- Triacca, U. (2006). *International Doctoral Program in Economics Scuola Superiore Sant'Anna*. Obtenido de Lesson19: Comparing Predictive Accuracy of two Forecasts: The Diebol-Mariano Test: <http://www.phdeconomics.sssup.it/documents/Lesson19.pdf>
- Tussel, F. (19 de Septiembre de 2007). *Estadística Matemática*. Obtenido de Universidad del País Vasco: <http://www.et.bs.ehu.es/~etptupaf/nuevo/ficheros/estad5/deci.pdf>
- Universidad de Talca, C. (sin fecha). *El principio de parsimonia ¿ De que se trata?* Obtenido de Portal Educativo: Recuperado el 08 de abril de 2016 de <http://www.educativo.otalca.cl/link.cgi/profesores/noticias/1964>
- Vidal Tovar, C. R. (2013). *UNAD Universidad Nacional Abierta y a Distancia*. Obtenido de Recuperado el 08 de abril de 2016 http://datateca.unad.edu.co/contenidos/211621/PROY-GRADO_EN_LINEA/leccin_41_importancia_del_analisis_estadstico_de_los_datos.html
- Villada, F., Muñoz, N., & García, E. (2012). Aplicación de las Redes Neuronales al Pronóstico de Precios en el Mercado de Valores. *Información Tecnológica*, 23, 11-20.

West, K. (1996). Asymptotic Inference About Predictive Ability. *Econometrica*, 64(5), 1067-1084.









































































Willmott, C., & Matsuura, K. (2005). Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing Average Model Performance. *Climate Research*, 30, 79-82.

Woschnagg, E., & Cipan, J. (2004). Evaluating Forecast Accuracy. *University of Vienna*.

Anexos

Anexo 1. Pruebas de ruido blanco y de estacionariedad e invertibilidad

Prueba de Ruido Blanco

	AC	CP	#	AC	CP	P
			1	0.026	0.026	
			2	0.073	0.072	
			3	-0.028	-0.032	
			4	0.005	0.001	
			5	0.015	0.019	
			6	-0.009	-0.011	
			7	-0.039	-0.041	
			8	-0.056	-0.052	
			9	-0.064	-0.056	0.157
			10	0.038	0.047	0.333
			11	-0.034	-0.031	0.502
			12	0.070	0.065	0.555
			13	-0.050	-0.045	0.647
			14	0.045	0.036	0.728
			15	0.045	0.048	0.790
			16	-0.097	-0.117	0.734
			17	-0.020	-0.024	0.810
			18	0.021	0.042	0.868
			19	-0.057	-0.061	0.886
			20	-0.012	-0.016	0.925
			21	-0.082	-0.061	0.911
			22	0.019	0.018	0.940
			23	-0.189	-0.179	0.661
			24	0.183	0.187	0.365
			25	-0.064	-0.072	0.391
			26	-0.051	-0.085	0.433
			27	0.040	0.068	0.482
			28	-0.006	-0.004	0.547
			29	0.010	-0.031	0.609
			30	0.039	0.035	0.653
			31	-0.054	-0.037	0.680
			32	0.089	0.065	0.658
			33	0.030	0.061	0.703
			34	-0.007	-0.057	0.751
			35	-0.023	0.005	0.790
			36	-0.135	-0.181	0.677

Nota: AC es Autocorrelación, CP es Correlación Parcial y P Probabilidad.

Fuente: Adaptado de herramienta EViews

Prueba de estacionariedad e invertibilidad

Raíces AR	Valor
0.978630	0.978630
0.230075 ± 0.950406i	0.977858
-0.867802 ± 0.445938i	0.975675
-0.634825 ± 0.736610i	0.972418
0.565395 ± 0.786339i	0.968504
0.892772 ± 0.364523i	0.964322
-0.145420 ± 0.949114i	0.960189
-0.953005 ± 0.079955i	0.956353
-0.299588 ± 0.904692i	0.953006
0.806964 ± 0.501864i	0.950293
0.676048 ± 0.665019i	0.948309
-0.746688 ± 0.584472i	0.948235
0.388088 ± 0.864739i	0.947832
0.931886 ± 0.169777i	0.947225
-0.486264 ± 0.812728i	0.947090
0.057279 ± 0.944982i	0.946716
-0.904230 ± 0.280033i	0.946599

Raíces MA	Valor
-0.645665 ± 0.659154i	0.922696
-0.892516 ± 0.217004i	0.918518
0.902055	0.902055
0.461711 ± 0.772928i	0.900330
0.743277 ± 0.441498i	0.864512
-0.324151 ± 0.778868i	0.843629
0.092226 ± 0.770036i	0.775539

Fuente: Elaboración propia

Anexo 2. Movimiento browniano a partir de una caminata aleatoria

Si se parte de una caminata aleatoria con deriva, se tiene entonces:

$$X_t = X_{t-1} + \epsilon_t + \mu \quad (60)$$

Donde ϵ_t es un choque aleatorio definido como ruido blanco gaussiano que se distribuye $N(0, \sigma^2)$. Si se definen los términos X_{t-1} y X_{t-2} se tiene:

$$X_{t-1} = X_{t-2} + \epsilon_{t-1} + \mu \quad (61)$$

$$X_{t-2} = X_{t-3} + \epsilon_{t-2} + \mu \quad (62)$$

Si se reemplaza (61) en (60) y (62) en (61), se obtiene,

$$X_t = X_{t-2} + \epsilon_{t-1} + \mu + \epsilon_t + \mu \quad (63)$$

$$X_t = X_{t-3} + \epsilon_{t-2} + \mu + \epsilon_{t-1} + \mu + \epsilon_t + \mu \quad (64)$$

Con lo anterior se deduce entonces,

$$X_t = X_0 + \sum_{i=1}^t \epsilon_i + t\mu \quad (65)$$

Si se tiene que $E(\epsilon_t) = 0$ y $V(\epsilon_t) = \sigma^2$ entonces:

$$E(X_t) = X_0 + t\mu, \quad (66)$$

$$V(X_t) = t\sigma^2 \quad (67)$$

Con lo cual X_t se distribuye $N(X_0 + t\mu, t\sigma^2)$. X_t se puede estandarizar en una variable Z_t así:

$$Z = \frac{X_t - (X_0 + t\mu)}{\sigma\sqrt{t}} \rightarrow N(0,1). \text{ Despejando } X_t,$$

$$X_t = (X_0 + t\mu) + Z (\sigma\sqrt{t}) \quad (68)$$

Ahora si $X_t = \ln(S_t)$ y $X_0 = \ln(S_0)$ se tendría entonces:

$$\ln(S_t) = (\ln(S_0) + t\mu) + Z (\sigma\sqrt{t}), \quad (69)$$

con lo cual,

$$\ln\left(\frac{S_t}{S_0}\right) = t\mu + Z (\sigma\sqrt{t}) \quad (70)$$

y por lo tanto,

$$S_t = S_0 e^{t\mu + Z (\sigma\sqrt{t})} \quad (71)$$

Siendo S_t la definición de un movimiento geométrico browniano. Nótese que $\ln\left(\frac{S_t}{S_0}\right)$ son los retornos logarítmicos. En este sentido el movimiento geométrico browniano corresponde a una caminata aleatoria cuando S_t y S_0 son precios sobre los cuales pueden calcularse retornos logarítmicos.

Anexo 3. Pronóstico modelo ARIMA

No.	Mes	Tasa real	Pronóstico para h=1	Pronóstico para h=2	Pronóstico para h=3	Pronóstico para h=4	Pronóstico para h=5	Pronóstico para h=6
1	201010	\$ 1.831,64	\$ 1.805,15					
2	201011	\$ 1.916,96	\$ 1.778,25	\$ 1.745,72				
3	201012	\$ 1.913,98	\$ 1.953,40	\$ 1.783,04	\$ 1.750,51			
4	201101	\$ 1.857,98	\$ 1.891,99	\$ 1.940,41	\$ 1.770,05	\$ 1.737,52		
5	201102	\$ 1.895,56	\$ 1.879,80	\$ 1.921,58	\$ 1.969,99	\$ 1.799,64	\$ 1.762,74	
6	201103	\$ 1.879,47	\$ 1.937,82	\$ 1.918,46	\$ 1.960,24	\$ 2.008,65	\$ 1.815,44	\$ 1.774,86
7	201104	\$ 1.768,19	\$ 1.849,58	\$ 1.921,24	\$ 1.901,89	\$ 1.943,66	\$ 1.998,57	\$ 1.786,08
8	201105	\$ 1.817,34	\$ 1.735,13	\$ 1.835,09	\$ 1.906,75	\$ 1.887,40	\$ 1.934,78	\$ 1.995,17
9	201106	\$ 1.780,16	\$ 1.929,52	\$ 1.828,55	\$ 1.928,51	\$ 2.000,17	\$ 1.978,22	\$ 2.030,33
10	201107	\$ 1.777,82	\$ 1.786,12	\$ 1.969,56	\$ 1.868,59	\$ 1.968,55	\$ 2.049,83	\$ 2.025,68
11	201108	\$ 1.783,66	\$ 1.733,76	\$ 1.743,95	\$ 1.927,39	\$ 1.826,42	\$ 1.939,79	\$ 2.029,17
12	201109	\$ 1.915,10	\$ 1.768,42	\$ 1.707,13	\$ 1.717,32	\$ 1.900,76	\$ 1.786,25	\$ 1.910,93
13	201110	\$ 1.863,06	\$ 1.958,88	\$ 1.778,73	\$ 1.717,44	\$ 1.727,63	\$ 1.935,67	\$ 1.809,74
14	201111	\$ 1.967,18	\$ 1.845,15	\$ 1.962,84	\$ 1.782,68	\$ 1.721,39	\$ 1.732,95	\$ 1.961,76
15	201112	\$ 1.942,70	\$ 2.025,56	\$ 1.875,69	\$ 1.993,37	\$ 1.813,22	\$ 1.743,71	\$ 1.756,42
16	201201	\$ 1.815,08	\$ 1.912,55	\$ 2.014,32	\$ 1.864,45	\$ 1.982,13	\$ 1.777,82	\$ 1.701,37
17	201202	\$ 1.767,83	\$ 1.763,19	\$ 1.882,90	\$ 1.984,67	\$ 1.834,80	\$ 1.968,27	\$ 1.743,57
18	201203	\$ 1.792,07	\$ 1.750,46	\$ 1.744,76	\$ 1.864,47	\$ 1.966,24	\$ 1.796,26	\$ 1.943,05
19	201204	\$ 1.761,20	\$ 1.802,04	\$ 1.750,93	\$ 1.745,24	\$ 1.864,95	\$ 1.980,37	\$ 1.793,43
20	201205	\$ 1.827,83	\$ 1.737,21	\$ 1.787,38	\$ 1.736,27	\$ 1.730,57	\$ 1.866,34	\$ 1.993,28
21	201206	\$ 1.784,60	\$ 1.779,79	\$ 1.668,49	\$ 1.718,66	\$ 1.667,55	\$ 1.661,08	\$ 1.810,40
22	201207	\$ 1.789,02	\$ 1.853,91	\$ 1.848,00	\$ 1.736,71	\$ 1.786,87	\$ 1.728,90	\$ 1.721,80
23	201208	\$ 1.830,50	\$ 1.739,95	\$ 1.819,64	\$ 1.813,73	\$ 1.702,44	\$ 1.759,33	\$ 1.695,58
24	201209	\$ 1.800,52	\$ 1.869,90	\$ 1.758,68	\$ 1.838,38	\$ 1.832,47	\$ 1.706,25	\$ 1.768,82
25	201210	\$ 1.829,89	\$ 1.746,76	\$ 1.831,97	\$ 1.720,76	\$ 1.800,45	\$ 1.793,75	\$ 1.654,93
26	201211	\$ 1.817,93	\$ 1.864,94	\$ 1.762,84	\$ 1.848,05	\$ 1.736,83	\$ 1.827,22	\$ 1.819,85
27	201212	\$ 1.768,23	\$ 1.801,41	\$ 1.859,15	\$ 1.757,05	\$ 1.842,26	\$ 1.716,12	\$ 1.815,53
28	201301	\$ 1.773,24	\$ 1.819,12	\$ 1.859,88	\$ 1.917,61	\$ 1.815,51	\$ 1.912,15	\$ 1.773,43
29	201302	\$ 1.816,42	\$ 1.757,43	\$ 1.813,78	\$ 1.854,54	\$ 1.912,27	\$ 1.796,47	\$ 1.902,76
30	201303	\$ 1.832,20	\$ 1.799,15	\$ 1.726,70	\$ 1.783,05	\$ 1.823,81	\$ 1.889,28	\$ 1.761,93
31	201304	\$ 1.828,79	\$ 1.764,41	\$ 1.723,82	\$ 1.651,37	\$ 1.707,72	\$ 1.753,94	\$ 1.825,95
32	201305	\$ 1.891,48	\$ 1.846,80	\$ 1.767,73	\$ 1.727,14	\$ 1.654,69	\$ 1.718,60	\$ 1.769,43
33	201306	\$ 1.929,00	\$ 1.899,79	\$ 1.844,92	\$ 1.765,85	\$ 1.725,26	\$ 1.643,09	\$ 1.713,37
34	201307	\$ 1.890,33	\$ 1.967,27	\$ 1.931,40	\$ 1.876,53	\$ 1.797,46	\$ 1.751,42	\$ 1.661,06
35	201308	\$ 1.935,43	\$ 1.903,68	\$ 1.998,18	\$ 1.962,31	\$ 1.907,44	\$ 1.817,76	\$ 1.767,13
36	201309	\$ 1.914,65	\$ 1.915,25	\$ 1.876,26	\$ 1.970,76	\$ 1.934,89	\$ 1.872,66	\$ 1.774,03
37	201310	\$ 1.884,06	\$ 1.931,09	\$ 1.931,83	\$ 1.892,84	\$ 1.987,34	\$ 1.946,66	\$ 1.878,21

No.	Mes	Tasa real	Pronóstico para h=1	Pronóstico para h=2	Pronóstico para h=3	Pronóstico para h=4	Pronóstico para h=5	Pronóstico para h=6
38	201311	\$ 1.931,88	\$ 1.848,68	\$ 1.906,45	\$ 1.907,19	\$ 1.868,20	\$ 1.975,37	\$ 1.930,63
39	201312	\$ 1.926,83	\$ 1.963,48	\$ 1.861,30	\$ 1.919,06	\$ 1.919,80	\$ 1.875,58	\$ 1.993,45
40	201401	\$ 2.008,26	\$ 1.907,54	\$ 1.952,52	\$ 1.850,34	\$ 1.908,11	\$ 1.908,95	\$ 1.860,31
41	201402	\$ 2.054,90	\$ 2.015,24	\$ 1.891,50	\$ 1.936,51	\$ 1.834,33	\$ 1.899,84	\$ 1.900,77
42	201403	\$ 1.965,32	\$ 2.046,26	\$ 1.997,55	\$ 1.873,81	\$ 1.918,82	\$ 1.802,93	\$ 1.874,98
43	201404	\$ 1.935,14	\$ 1.955,76	\$ 2.055,18	\$ 2.006,46	\$ 1.882,72	\$ 1.933,77	\$ 1.806,32
44	201405	\$ 1.900,64	\$ 1.919,32	\$ 1.944,65	\$ 2.044,06	\$ 1.995,34	\$ 1.855,01	\$ 1.911,15
45	201406	\$ 1.881,19	\$ 1.956,54	\$ 1.979,48	\$ 2.004,81	\$ 2.104,22	\$ 2.048,97	\$ 1.894,63
46	201407	\$ 1.872,43	\$ 1.830,62	\$ 1.923,16	\$ 1.946,10	\$ 1.971,43	\$ 2.084,18	\$ 2.023,42
47	201408	\$ 1.918,62	\$ 1.905,67	\$ 1.854,32	\$ 1.946,86	\$ 1.969,80	\$ 1.998,53	\$ 2.122,53
48	201409	\$ 2.028,48	\$ 1.916,90	\$ 1.900,99	\$ 1.849,65	\$ 1.942,18	\$ 1.968,20	\$ 1.999,80
49	201410	\$ 2.050,52	\$ 2.009,94	\$ 1.872,90	\$ 1.856,99	\$ 1.805,65	\$ 1.910,60	\$ 1.939,21
50	201411	\$ 2.206,19	\$ 2.064,56	\$ 2.014,73	\$ 1.877,68	\$ 1.861,78	\$ 1.803,54	\$ 1.918,97
51	201412	\$ 2.392,46	\$ 2.223,44	\$ 2.049,50	\$ 1.999,66	\$ 1.862,62	\$ 1.844,58	\$ 1.780,53
52	201501	\$ 2.441,10	\$ 2.452,54	\$ 2.244,96	\$ 2.071,02	\$ 2.021,18	\$ 1.865,75	\$ 1.845,91
53	201502	\$ 2.496,99	\$ 2.467,72	\$ 2.481,77	\$ 2.274,19	\$ 2.100,25	\$ 2.043,72	\$ 1.872,79
54	201503	\$ 2.576,05	\$ 2.514,58	\$ 2.478,63	\$ 2.492,69	\$ 2.285,10	\$ 2.087,83	\$ 2.025,67
55	201504	\$ 2.388,06	\$ 2.660,97	\$ 2.585,48	\$ 2.549,53	\$ 2.563,58	\$ 2.328,16	\$ 2.111,19
56	201505	\$ 2.533,79	\$ 2.321,43	\$ 2.656,62	\$ 2.581,13	\$ 2.545,18	\$ 2.561,12	\$ 2.302,20
57	201506	\$ 2.585,11	\$ 2.540,27	\$ 2.279,46	\$ 2.614,65	\$ 2.539,15	\$ 2.498,38	\$ 2.515,91
58	201507	\$ 2.866,04	\$ 2.624,17	\$ 2.569,10	\$ 2.308,29	\$ 2.643,47	\$ 2.557,86	\$ 2.513,01
59	201508	\$ 3.101,10	\$ 2.849,12	\$ 2.552,06	\$ 2.497,00	\$ 2.236,18	\$ 2.616,33	\$ 2.522,17
60	201509	\$ 3.121,94	\$ 3.132,12	\$ 2.822,65	\$ 2.525,59	\$ 2.470,52	\$ 2.174,72	\$ 2.592,80

Fuente: Elaboración propia

Anexo 4. Pronóstico modelo movimiento browniano

No.	Mes	Tasa real	Pronóstico para h=1	Pronóstico para h=2	Pronóstico para h=3	Pronóstico para h=4	Pronóstico para h=5	Pronóstico para h=6
1	201010	\$ 1.831,64	\$ 1.799,05					
2	201011	\$ 1.916,96	\$ 1.830,78	\$ 1.797,78				
3	201012	\$ 1.913,98	\$ 1.916,06	\$ 1.829,50	\$ 1.796,45			
4	201101	\$ 1.857,98	\$ 1.913,08	\$ 1.914,72	\$ 1.828,14	\$ 1.795,08		
5	201102	\$ 1.895,56	\$ 1.857,11	\$ 1.911,74	\$ 1.913,30	\$ 1.826,75	\$ 1.793,69	
6	201103	\$ 1.879,47	\$ 1.894,67	\$ 1.855,81	\$ 1.910,32	\$ 1.911,84	\$ 1.825,34	\$ 1.792,29
7	201104	\$ 1.768,19	\$ 1.878,59	\$ 1.893,34	\$ 1.854,43	\$ 1.908,87	\$ 1.910,36	\$ 1.823,91
8	201105	\$ 1.817,34	\$ 1.767,36	\$ 1.877,27	\$ 1.891,94	\$ 1.853,02	\$ 1.907,39	\$ 1.908,87
9	201106	\$ 1.780,16	\$ 1.816,49	\$ 1.766,12	\$ 1.875,88	\$ 1.890,50	\$ 1.851,58	\$ 1.905,90
10	201107	\$ 1.777,82	\$ 1.779,33	\$ 1.815,21	\$ 1.764,81	\$ 1.874,45	\$ 1.889,04	\$ 1.850,13
11	201108	\$ 1.783,66	\$ 1.776,99	\$ 1.778,08	\$ 1.813,87	\$ 1.763,47	\$ 1.873,00	\$ 1.887,56
12	201109	\$ 1.915,10	\$ 1.782,83	\$ 1.775,74	\$ 1.776,76	\$ 1.812,49	\$ 1.762,10	\$ 1.871,53
13	201110	\$ 1.863,06	\$ 1.914,20	\$ 1.781,57	\$ 1.774,42	\$ 1.775,41	\$ 1.811,08	\$ 1.760,72
14	201111	\$ 1.967,18	\$ 1.862,19	\$ 1.912,86	\$ 1.780,25	\$ 1.773,07	\$ 1.774,03	\$ 1.809,67
15	201112	\$ 1.942,70	\$ 1.966,26	\$ 1.860,88	\$ 1.911,44	\$ 1.778,90	\$ 1.771,70	\$ 1.772,64
16	201201	\$ 1.815,08	\$ 1.941,79	\$ 1.964,88	\$ 1.859,50	\$ 1.909,99	\$ 1.777,52	\$ 1.770,31
17	201202	\$ 1.767,83	\$ 1.814,23	\$ 1.940,43	\$ 1.963,42	\$ 1.858,09	\$ 1.908,51	\$ 1.776,13
18	201203	\$ 1.792,07	\$ 1.767,00	\$ 1.812,96	\$ 1.938,99	\$ 1.961,93	\$ 1.856,65	\$ 1.907,01
19	201204	\$ 1.761,20	\$ 1.791,23	\$ 1.765,76	\$ 1.811,61	\$ 1.937,51	\$ 1.960,41	\$ 1.855,19
20	201205	\$ 1.827,83	\$ 1.760,38	\$ 1.789,97	\$ 1.764,45	\$ 1.810,23	\$ 1.936,01	\$ 1.958,87
21	201206	\$ 1.784,60	\$ 1.826,97	\$ 1.759,14	\$ 1.788,65	\$ 1.763,11	\$ 1.808,83	\$ 1.934,50
22	201207	\$ 1.789,02	\$ 1.783,77	\$ 1.825,69	\$ 1.757,84	\$ 1.787,29	\$ 1.761,75	\$ 1.807,42
23	201208	\$ 1.830,50	\$ 1.788,18	\$ 1.782,51	\$ 1.824,34	\$ 1.756,50	\$ 1.785,90	\$ 1.760,37
24	201209	\$ 1.800,52	\$ 1.829,64	\$ 1.786,93	\$ 1.781,19	\$ 1.822,95	\$ 1.755,14	\$ 1.784,50
25	201210	\$ 1.829,89	\$ 1.799,68	\$ 1.828,36	\$ 1.785,60	\$ 1.779,84	\$ 1.821,54	\$ 1.753,76
26	201211	\$ 1.817,93	\$ 1.829,03	\$ 1.798,41	\$ 1.827,00	\$ 1.784,24	\$ 1.778,46	\$ 1.820,11
27	201212	\$ 1.768,23	\$ 1.817,08	\$ 1.827,75	\$ 1.797,08	\$ 1.825,61	\$ 1.782,86	\$ 1.777,06
28	201301	\$ 1.773,24	\$ 1.767,40	\$ 1.815,80	\$ 1.826,39	\$ 1.795,71	\$ 1.824,20	\$ 1.781,47
29	201302	\$ 1.816,42	\$ 1.772,41	\$ 1.766,16	\$ 1.814,46	\$ 1.825,00	\$ 1.794,32	\$ 1.822,77
30	201303	\$ 1.832,20	\$ 1.815,57	\$ 1.771,17	\$ 1.764,85	\$ 1.813,08	\$ 1.823,59	\$ 1.792,92
31	201304	\$ 1.828,79	\$ 1.831,34	\$ 1.814,29	\$ 1.769,85	\$ 1.763,51	\$ 1.811,67	\$ 1.822,16
32	201305	\$ 1.891,48	\$ 1.827,93	\$ 1.830,06	\$ 1.812,95	\$ 1.768,51	\$ 1.762,14	\$ 1.810,25
33	201306	\$ 1.929,00	\$ 1.890,60	\$ 1.826,65	\$ 1.828,70	\$ 1.811,57	\$ 1.767,14	\$ 1.760,76
34	201307	\$ 1.890,33	\$ 1.928,10	\$ 1.889,27	\$ 1.825,30	\$ 1.827,31	\$ 1.810,17	\$ 1.765,75
35	201308	\$ 1.935,43	\$ 1.889,45	\$ 1.926,74	\$ 1.887,87	\$ 1.823,91	\$ 1.825,89	\$ 1.808,75
36	201309	\$ 1.914,65	\$ 1.934,52	\$ 1.888,12	\$ 1.925,32	\$ 1.886,43	\$ 1.822,50	\$ 1.824,46
37	201310	\$ 1.884,06	\$ 1.913,75	\$ 1.933,17	\$ 1.886,72	\$ 1.923,85	\$ 1.884,97	\$ 1.821,07

No.	Mes	Tasa real	Pronóstico para h=1	Pronóstico para h=2	Pronóstico para h=3	Pronóstico para h=4	Pronóstico para h=5	Pronóstico para h=6
38	201311	\$ 1.931,88	\$ 1.883,18	\$ 1.912,41	\$ 1.931,73	\$ 1.885,28	\$ 1.922,36	\$ 1.883,49
39	201312	\$ 1.926,83	\$ 1.930,98	\$ 1.881,86	\$ 1.910,99	\$ 1.930,26	\$ 1.883,82	\$ 1.920,85
40	201401	\$ 2.008,26	\$ 1.925,93	\$ 1.929,62	\$ 1.880,46	\$ 1.909,54	\$ 1.928,77	\$ 1.882,35
41	201402	\$ 2.054,90	\$ 2.007,32	\$ 1.924,58	\$ 1.928,19	\$ 1.879,03	\$ 1.908,06	\$ 1.927,26
42	201403	\$ 1.965,32	\$ 2.053,94	\$ 2.005,91	\$ 1.923,15	\$ 1.926,72	\$ 1.877,58	\$ 1.906,57
43	201404	\$ 1.935,14	\$ 1.964,40	\$ 2.052,50	\$ 2.004,42	\$ 1.921,69	\$ 1.925,23	\$ 1.876,10
44	201405	\$ 1.900,64	\$ 1.934,23	\$ 1.963,02	\$ 2.050,97	\$ 2.002,90	\$ 1.920,20	\$ 1.923,72
45	201406	\$ 1.881,19	\$ 1.899,75	\$ 1.932,88	\$ 1.961,57	\$ 2.049,41	\$ 2.001,35	\$ 1.918,69
46	201407	\$ 1.872,43	\$ 1.880,31	\$ 1.898,42	\$ 1.931,44	\$ 1.960,07	\$ 2.047,83	\$ 1.999,78
47	201408	\$ 1.918,62	\$ 1.871,55	\$ 1.878,99	\$ 1.897,01	\$ 1.929,97	\$ 1.958,56	\$ 2.046,22
48	201409	\$ 2.028,48	\$ 1.917,72	\$ 1.870,24	\$ 1.877,60	\$ 1.895,57	\$ 1.928,48	\$ 1.957,02
49	201410	\$ 2.050,52	\$ 2.027,53	\$ 1.916,38	\$ 1.868,85	\$ 1.876,17	\$ 1.894,10	\$ 1.926,97
50	201411	\$ 2.206,19	\$ 2.049,56	\$ 2.026,11	\$ 1.914,95	\$ 1.867,43	\$ 1.874,71	\$ 1.892,61
51	201412	\$ 2.392,46	\$ 2.205,16	\$ 2.048,12	\$ 2.024,61	\$ 1.913,50	\$ 1.865,99	\$ 1.873,25
52	201501	\$ 2.441,10	\$ 2.391,34	\$ 2.203,61	\$ 2.046,60	\$ 2.023,06	\$ 1.912,02	\$ 1.864,52
53	201502	\$ 2.496,99	\$ 2.439,96	\$ 2.389,66	\$ 2.201,98	\$ 2.045,05	\$ 2.021,50	\$ 1.910,52
54	201503	\$ 2.576,05	\$ 2.495,82	\$ 2.438,24	\$ 2.387,89	\$ 2.200,30	\$ 2.043,46	\$ 2.019,91
55	201504	\$ 2.388,06	\$ 2.574,84	\$ 2.494,07	\$ 2.436,44	\$ 2.386,07	\$ 2.198,60	\$ 2.041,86
56	201505	\$ 2.533,79	\$ 2.386,94	\$ 2.573,04	\$ 2.492,22	\$ 2.434,58	\$ 2.384,23	\$ 2.196,87
57	201506	\$ 2.585,11	\$ 2.532,60	\$ 2.385,27	\$ 2.571,13	\$ 2.490,32	\$ 2.432,70	\$ 2.382,36
58	201507	\$ 2.866,04	\$ 2.583,90	\$ 2.530,83	\$ 2.383,50	\$ 2.569,17	\$ 2.488,40	\$ 2.430,79
59	201508	\$ 3.101,10	\$ 2.864,70	\$ 2.582,09	\$ 2.528,95	\$ 2.381,68	\$ 2.567,18	\$ 2.486,45
60	201509	\$ 3.121,94	\$ 3.099,65	\$ 2.862,69	\$ 2.580,17	\$ 2.527,02	\$ 2.379,84	\$ 2.565,17

Fuente: Elaboración propia

Anexo 5. Pronóstico de ajuste

No.	Mes	Tasa real	ARIMA	MB *	No.	Mes	Tasa real	ARIMA	MB *
1	200010	\$ 2.158,36			61	200510	\$ 2.289,57	\$ 2.304,15	\$ 2.288,54
2	200011	\$ 2.172,84		\$ 2.157,35	62	200511	\$ 2.274,04	\$ 2.317,09	\$ 2.288,50
3	200012	\$ 2.229,18		\$ 2.171,82	63	200512	\$ 2.284,22	\$ 2.277,76	\$ 2.272,98
4	200101	\$ 2.240,80		\$ 2.228,14	64	200601	\$ 2.265,65	\$ 2.268,81	\$ 2.283,15
5	200102	\$ 2.257,45		\$ 2.239,75	65	200602	\$ 2.247,32	\$ 2.226,10	\$ 2.264,59
6	200103	\$ 2.310,57		\$ 2.256,39	66	200603	\$ 2.289,98	\$ 2.245,53	\$ 2.246,27
7	200104	\$ 2.346,73		\$ 2.309,49	67	200604	\$ 2.375,03	\$ 2.300,46	\$ 2.288,91
8	200105	\$ 2.324,98		\$ 2.345,63	68	200605	\$ 2.482,41	\$ 2.414,75	\$ 2.373,92
9	200106	\$ 2.298,85		\$ 2.323,89	69	200606	\$ 2.633,12	\$ 2.515,92	\$ 2.481,25
10	200107	\$ 2.298,27		\$ 2.297,77	70	200607	\$ 2.426,00	\$ 2.663,50	\$ 2.631,89
11	200108	\$ 2.301,23		\$ 2.297,19	71	200608	\$ 2.396,63	\$ 2.381,77	\$ 2.424,86
12	200109	\$ 2.332,19		\$ 2.300,15	72	200609	\$ 2.394,31	\$ 2.397,33	\$ 2.395,51
13	200110	\$ 2.310,02		\$ 2.331,10	73	200610	\$ 2.315,38	\$ 2.409,14	\$ 2.393,19
14	200111	\$ 2.308,59		\$ 2.308,94	74	200611	\$ 2.300,42	\$ 2.244,98	\$ 2.314,30
15	200112	\$ 2.291,18		\$ 2.307,51	75	200612	\$ 2.238,79	\$ 2.269,06	\$ 2.299,34
16	200201	\$ 2.264,82		\$ 2.290,11	76	200701	\$ 2.259,72	\$ 2.260,83	\$ 2.237,74
17	200202	\$ 2.309,82		\$ 2.263,76	77	200702	\$ 2.224,12	\$ 2.255,72	\$ 2.258,66
18	200203	\$ 2.261,23		\$ 2.308,74	78	200703	\$ 2.190,30	\$ 2.189,78	\$ 2.223,08
19	200204	\$ 2.275,35		\$ 2.260,17	79	200704	\$ 2.110,67	\$ 2.159,45	\$ 2.189,28
20	200205	\$ 2.321,16		\$ 2.274,29	80	200705	\$ 1.930,64	\$ 2.045,63	\$ 2.109,68
21	200206	\$ 2.398,82		\$ 2.320,07	81	200706	\$ 1.960,61	\$ 1.966,20	\$ 1.929,74
22	200207	\$ 2.625,06		\$ 2.397,70	82	200707	\$ 1.971,80	\$ 1.956,46	\$ 1.959,69
23	200208	\$ 2.703,55		\$ 2.623,83	83	200708	\$ 2.173,17	\$ 2.026,49	\$ 1.970,88
24	200209	\$ 2.828,08		\$ 2.702,29	84	200709	\$ 2.023,19	\$ 2.112,12	\$ 2.172,15
25	200210	\$ 2.773,73		\$ 2.826,76	85	200710	\$ 1.999,44	\$ 1.966,14	\$ 2.022,24
26	200211	\$ 2.784,21		\$ 2.772,43	86	200711	\$ 2.060,42	\$ 2.046,87	\$ 1.998,50
27	200212	\$ 2.864,79		\$ 2.782,91	87	200712	\$ 2.014,76	\$ 2.072,92	\$ 2.059,46
28	200301	\$ 2.926,46		\$ 2.863,45	88	200801	\$ 1.939,60	\$ 2.010,85	\$ 2.013,82
29	200302	\$ 2.956,31		\$ 2.925,09	89	200802	\$ 1.843,59	\$ 1.857,53	\$ 1.938,69
30	200303	\$ 2.958,25		\$ 2.954,93	90	200803	\$ 1.821,60	\$ 1.875,90	\$ 1.842,73
31	200304	\$ 2.887,82		\$ 2.956,87	91	200804	\$ 1.780,21	\$ 1.803,64	\$ 1.820,75
32	200305	\$ 2.853,33		\$ 2.886,47	92	200805	\$ 1.744,01	\$ 1.753,05	\$ 1.779,38
33	200306	\$ 2.817,32		\$ 2.852,00	93	200806	\$ 1.923,02	\$ 1.733,69	\$ 1.743,19
34	200307	\$ 2.880,40		\$ 2.816,00	94	200807	\$ 1.792,24	\$ 1.889,63	\$ 1.922,12
35	200308	\$ 2.832,94		\$ 2.879,05	95	200808	\$ 1.932,20	\$ 1.781,22	\$ 1.791,40
36	200309	\$ 2.889,39	\$ 2.891,53	\$ 2.831,61	96	200809	\$ 2.174,62	\$ 1.946,39	\$ 1.931,30
37	200310	\$ 2.884,17	\$ 2.913,13	\$ 2.888,04	97	200810	\$ 2.359,52	\$ 2.319,53	\$ 2.173,60

No.	Mes	Tasa real	ARIMA	MB *	No.	Mes	Tasa real	ARIMA	MB *
38	200311	\$ 2.836,05	\$ 2.911,80	\$ 2.882,82	98	200811	\$ 2.318,00	\$ 2.338,53	\$ 2.358,42
39	200312	\$ 2.778,21	\$ 2.819,51	\$ 2.834,72	99	200812	\$ 2.243,59	\$ 2.293,07	\$ 2.316,92
40	200401	\$ 2.742,47	\$ 2.778,31	\$ 2.776,91	100	200901	\$ 2.420,26	\$ 2.357,98	\$ 2.242,54
41	200402	\$ 2.682,34	\$ 2.769,11	\$ 2.741,19	101	200902	\$ 2.555,89	\$ 2.477,11	\$ 2.419,13
42	200403	\$ 2.678,16	\$ 2.661,09	\$ 2.681,09	102	200903	\$ 2.561,21	\$ 2.570,65	\$ 2.554,69
43	200404	\$ 2.646,99	\$ 2.679,23	\$ 2.676,91	103	200904	\$ 2.289,73	\$ 2.472,46	\$ 2.560,01
44	200405	\$ 2.724,92	\$ 2.641,54	\$ 2.645,75	104	200905	\$ 2.140,66	\$ 2.214,47	\$ 2.288,66
45	200406	\$ 2.699,58	\$ 2.719,88	\$ 2.723,65	105	200906	\$ 2.158,67	\$ 2.176,50	\$ 2.139,66
46	200407	\$ 2.612,44	\$ 2.690,40	\$ 2.698,32	106	200907	\$ 2.043,37	\$ 2.075,68	\$ 2.157,66
47	200408	\$ 2.551,43	\$ 2.598,51	\$ 2.611,22	107	200908	\$ 2.035,00	\$ 2.015,66	\$ 2.042,41
48	200409	\$ 2.595,17	\$ 2.573,35	\$ 2.550,24	108	200909	\$ 1.922,00	\$ 1.906,67	\$ 2.034,05
49	200410	\$ 2.575,19	\$ 2.603,29	\$ 2.593,96	109	200910	\$ 1.993,80	\$ 1.945,81	\$ 1.921,10
50	200411	\$ 2.479,10	\$ 2.573,01	\$ 2.573,99	110	200911	\$ 1.997,47	\$ 2.073,84	\$ 1.992,87
51	200412	\$ 2.389,75	\$ 2.445,10	\$ 2.477,94	111	200912	\$ 2.044,23	\$ 2.005,79	\$ 1.996,54
52	200501	\$ 2.367,76	\$ 2.396,76	\$ 2.388,63	112	201001	\$ 1.982,29	\$ 2.026,74	\$ 2.043,27
53	200502	\$ 2.323,77	\$ 2.370,50	\$ 2.366,65	113	201002	\$ 1.932,32	\$ 1.888,63	\$ 1.981,36
54	200503	\$ 2.376,48	\$ 2.317,39	\$ 2.322,68	114	201003	\$ 1.928,59	\$ 2.005,93	\$ 1.931,42
55	200504	\$ 2.348,32	\$ 2.396,89	\$ 2.375,37	115	201004	\$ 1.969,75	\$ 1.959,25	\$ 1.927,69
56	200505	\$ 2.332,79	\$ 2.364,95	\$ 2.347,22	116	201005	\$ 1.971,55	\$ 2.070,09	\$ 1.968,83
57	200506	\$ 2.331,81	\$ 2.346,28	\$ 2.331,70	117	201006	\$ 1.916,46	\$ 1.849,27	\$ 1.970,63
58	200507	\$ 2.308,49	\$ 2.355,68	\$ 2.330,72	118	201007	\$ 1.842,79	\$ 1.876,24	\$ 1.915,56
59	200508	\$ 2.304,30	\$ 2.312,14	\$ 2.307,41	119	201008	\$ 1.823,74	\$ 1.850,31	\$ 1.841,93
60	200509	\$ 2.289,61	\$ 2.289,12	\$ 2.303,22	120	201009	\$ 1.799,89	\$ 1.771,25	\$ 1.822,89

Nota: MB es Movimiento Browniano

Fuente: Elaboración propia