
Modelo de Pronostico PIB Financiero

Presentado por:
**Jhon Wilfer
Virgüez Sierra**



LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Fundación Universitaria Los Libertadores
Facultad de Ingeniería y Ciencias Básicas
Especialización en Estadística Aplicada
Bogotá D.C. Colombia
2018

Modelo de Pronostico PIB Financiero

Presentado por
Jhon Wilfer Virgüez Sierra
en cumplimiento parcial de los requerimientos para optar al
título de **Especialista en Estadística Aplicada**

Dirigida por
Sebastien Lozano
Profesor



LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Fundación Universitaria Los Libertadores
Facultad de Ingeniería y Ciencias Básicas
Especialización en Estadística Aplicada
Bogotá D.C. Colombia

2018

Notas de aceptación



LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Firma del presidente del
jurado

Firma del jurado

Firma del jurado

Bogotá DC, febrero de 2019.

Las directivas de la Fundación Universitaria Los Libertadores, los jurados calificadores y el cuerpo docente no son responsables por los criterios e ideas expuestas en el presente documento. Estos corresponden únicamente a los autores y a los resultados de su trabajo.



LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Dedicatoria

Dedico este logro a mi familia, que sin su apoyo esto no hubiera sido posible, fue fundamental contar con su tiempo, paciencia y voz de ánimo.



LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Agradecimientos

Agradezco la paciencia mantenida por mi familia, especialmente el sacrificio de tiempo que demando el desarrollo de este producto. Agradecimientos infinitos a Olga, su paciencia para comprender mis continuas frustraciones fue fundamental para lograr alcanzar este objetivo. Igualmente, agradezco la oportuna orientación de los docentes y su compromiso para que este producto llegue a buen término



LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Índice general

Modelo de Pronostico del PIB Financiero	1
Capítulo 1 Introducción	3
Capítulo 2. Planteamiento del Problema.....	5
2.1. Objetivos.....	6
2.2. Objetivo General.....	6
2.3. Objetivos Específicos.....	6
2.4. Justificación	7
Capítulo 3 Marco Teórico / Conceptual.....	9
3.1. Crecimiento del PIB	9
3.2. Intermediación Financiera.....	12
3.3. Series de tiempo	15
Capítulo 4 Marco Metodológico.....	11
4.1. Modelos para series estacionarias Autorregresivos (AR) y Media Móvil (MA)	12
4.2. Modelos para series No Estacionarias (ARIMA y SARIMA).....	14
4.3. Metodología Box-Jenkins	17
Capítulo 5 Análisis y Resultados.....	21
5.1. Exploración de la serie	21
5.2. Postulación de los Modelos.....	30
5.3. Elección del Modelo y verificación de supuestos.....	32
5.4. Pronóstico.....	34
5.5. Análisis	38
Capítulo 6 Conclusiones y Recomendaciones	15
Bibliografía	21
Apéndice A Primer Apéndice.....	23

Índice de figuras

Gráfica 1. PIB valor agregado 2017.....	11
Gráfica 2 Entidades que componen el sector financiero	13
Gráfica 3. Metodología de Box-Jenkins	18
Gráfica 4 Serie Original	21
Gráfica 5. Descomposición de la serie y dummy	22
Gráfica 6 Descomposición de la serie.....	22
Gráfica 7 Autocorrelograma Simple.....	23
Gráfica 8 Autocorrelograma Parcial.....	23
Gráfica 9 Quantile-Quantile	24
Gráfica 10 Residuos de muestra sin tendencia	26
Gráfica 11 Serie Diferenciada.....	28
Gráfica 12.Descomposición Serie diferenciada.....	28
Gráfica 13. ACF y PACF Serie diferenciada.....	29
Gráfica 14. Modelo 1 Vs Datos Originales	30
Gráfica 15. Modelo 2 Vs Datos Originales.....	31
Gráfica 16. Modelo 3 Vs Datos Originales.....	31
Gráfica 17. Modelo 4 Vs Datos Originales.....	32
Gráfica 18. Modelos Vs Datos Originales	32
Gráfica 19 Comprobación de los errores del modelo.....	34
Gráfica 20. Evaluación de la predicción.....	34
Gráfica 21 Comportamiento del Suavizamiento	35
Gráfica 22 Combinación ARIMA con el Suavizamiento	36
Gráfica 23 Combinación del ARIMA con el Suavizamiento para toda la serie.....	37
Gráfica 24. Verificación de residuos para la combinación	37
Gráfica 25. Predicción de dos años.....	38

Índice de cuadros

Tabla 1 Pruebas de raíz unitaria DF y PP para toda la serie.....	25
Tabla 2 Pruebas de raíz unitaria DF y DF Aumentada para el conjunto de prueba.....	25
Tabla 3 Prueba de Raíz Unitaria KPSS con tendencia.....	26
Tabla 4 Prueba Raíz unitaria DF y DF aumentada para muestra sin tendencia	27
Tabla 5 Prueba de Raíz Unitaria KPSS sin tendencia.....	27
Tabla 6 Pruebas de raíz unitaria DF y PP para la serie diferenciada.	29
Tabla 7 Criterios de Información	33
Tabla 8 Prueba de Wald para los coeficientes	33
Tabla 9 Precisión del modelo en el conjunto de prueba.....	35
Tabla 10 Precisión del suavizamiento en el conjunto de prueba	36
Tabla 11 Precisión de la Combinación del Modelo con el Suavizamiento conjunto de prueba	36
Tabla 12 Precisión de la Combinación del Modelo con el Suavizamiento para toda la serie ...	37
Tabla 13 Prueba de Wald para los coeficientes del modelo definitivo	39

Modelo de Pronostico del PIB Financiero

Resumen

El presente trabajo se orienta a estimar un modelo de pronóstico para el PIB Financiero, a partir de la serie de intermediación financiera que publica trimestralmente el DANE en el reporte del PIB por ramas de la actividad económica. A partir de esta serie, que va del primer trimestre de 2005 al tercer trimestre de 2018, se estimó un modelo ARIMA (1,0,0) X (1,0,0) [4] con parámetros, $ARIMA (1,0,0) X (1,0,0) [4] (1 - 0.6968B)(1 - 0.7075B^4)(y_t - 3429.2321 - 117.4861 - 201.8461) = \varepsilon_t$ este modelo incluye intercepto, tendencia y una variable dummy para el cuarto trimestre de 2008. Dada la limitación en la potencia de la predicción, el modelo se mejoró con una combinación del algoritmo de suavizamiento exponencial, dando como resultado un nuevo ARIMA (1,0,0) X (1,0,0) [4] con parámetros $(1 - 0.7192B)(1 - 0.7004B^4)(y_t - 3402.981 - 120.541 - 200.565) = \varepsilon_t$ con todos los parámetros significativos al 99%, y un error promedio de -6.740 y un error promedio absoluto de 108.449, útil y preciso para hacer predicciones en un horizonte de 8 periodos o 2 años.

Palabras claves: PIB, Intermediación Financiera, ARIMA. Escriba aquí las palabras clave.

Capítulo 1 Introducción

El Producto Interno Bruto es un indicador del flujo de la economía para un país en un periodo de tiempo determinado (Mankiw, 2014), es también el mejor indicador de los resultados de la economía. Por su importancia, es calculado de manera trimestral por las instituciones de estadística de los países, pues permite identificar cual ha sido el comportamiento de la economía, permitiendo conocer la renta de los miembros de la economía o el gasto en la producción de bienes y servicios, lo que da una idea de la renta y la producción para satisfacer la demanda de los hogares, las empresas, el gobierno y el mercado extranjero. Y en términos generales, permite anticipar cual es el comportamiento de la economía en su conjunto o de sectores particulares al permitir pronosticar el comportamiento en el mediano y largo plazo con la ayuda de las herramientas de análisis de series de tiempo.

Para el caso concreto del análisis del presente documento, se considera solo el componente del sector establecimientos financieros y de esta rama de la actividad económica se analiza la serie de la intermediación financiera. Para lo anterior, se toma el valor agregado del sector establecimientos financiero y de ahí se extrae la serie de la intermediación financiera. El sector establecimientos financieros tiene un peso promedio del 20% para el periodo de análisis, sobre el valor total del PIB, y cerca de un cuarto es aportado por el componente de la intermediación financiera (el restante, de la rama de establecimientos financieros, es el resultado del comportamiento de las actividades inmobiliarias y alquiler de vivienda, y las actividades empresariales y de alquiler). De esta manera, se toma la serie desestacionalizada y a precios constante de 2005. Su importancia, radica tanto en el aporte directo al PIB total, como en que en que dichas actividades son un indicador del comportamiento de la economía como un todo.

La importancia de las actividades de intermediación tiene su justificación en que permite conocer cuál es el comportamiento del sistema financiero, a través del nivel de ahorro e inversión que tiene el país en un determinado momento, permite proyectar en el largo plazo el crecimiento que tenga la economía y permite anticipar el costo de la inversión.

El objetivo de este documento es, por tanto, determinar y estimar un modelo de pronóstico sobre el PIB financiero, entendido este como el comportamiento de la actividad de intermediación financiera, a partir del uso de las herramientas de análisis de series de tiempo, particularmente el uso de la metodología Box Jenkins.

Capítulo 2. Planteamiento del Problema

Las actividades de intermediación han empezado a tener un peso preponderante dentro del sector establecimientos financieros, seguros, actividades inmobiliarias y servicios a las empresas, y dentro del mismo PIB. Para el primer caso, paso de representar, en promedio para el año 2000, el 21%, a cerca del 32%, en promedio para el 2017, mostrando un crecimiento del 48%, convirtiendo dicho componente en el segundo dentro del sector. Así mismo, el peso de la intermediación dentro del PIB total paso de representar, en promedio para el 2000, el 4% a un 7%, en promedio para el 2017, mostrando un crecimiento del 64%, este crecimiento es más significativo si se tiene en cuenta que la participación del sector financiero se mantuvo casi constante en los años del análisis. Sumado a lo anterior, el componente de la intermediación financiera permite tener una idea de cuál es el comportamiento del ahorro y la inversión de un país, y la necesidad de conocer su comportamiento futuro justifica el plantear y estimar un modelo estadístico que permita anticipar y conocer cuál será el comportamiento de la serie a partir del análisis de su trayectoria.

Para tratar el problema de conocer el comportamiento futuro de la intermediación financiera se han implementado diversos modelos, sin embargos casi todos son de tipo econométricos que buscan explicar la relación que existe entre esta y otras variables económicas, como el crecimiento (Levine, Loayza, & Beck, 2000), (Fernando G & Garcia M, 1995) o como variable explicativa del crecimiento (Fernando G & Garcia M, 1995) y (Kuosmanen & Vataja, 2014), o en agregado como los modelos univariados y multivariados que presenta (Hernandez R, 2004) que se orientan a estimar un modelo sobre todo el componente de los establecimientos financieros, seguros, inmuebles y servicios a las empresas. La novedad del presente trabajo es que, consiente de la limitación de la información disponible, se plantea un modelo simple, que bajo el principio de parsimonia (Navaja de Ockham) donde siempre será mejor la explicación simple sobre la compleja, se estime un modelo univariado con solo la información de intermediación financiera y que tenga capacidad para hacer pronósticos en el corto y mediano plazo.

2.1. Objetivos

2.2. Objetivo General

- Estimar un modelo de pronóstico de la intermediación financiera, tomando como base el periodo comprendido entre el primer trimestre de 2005 al tercer trimestre de 2018

2.3. Objetivos Específicos

- Identificar los componentes del PIB colombiano y cada una de sus variables
- Plantear por lo menos 3 modelos que cumpla con los supuestos, tengan en cuenta las particularidades de los datos.
- Elegir el mejor modelo y elaborar el pronóstico.

2.4. Justificación

Se estima importante modelar el componente de intermediación financiera, y con ello hacer previsiones sobre su valor futuro, en la medida que da cuenta sobre el comportamiento del ahorro y la inversión de un país. En otras palabras, la actividad de intermediación financiera consiste en captar recursos (bajo la forma de depósitos, certificados a término, bonos, títulos, etc.), transformarlos (calcular un tiempo estimado del depósito) y ponerlos en el mercado (prestamos, inversión, etc.). En este sentido, su importancia es capital en la medida que permitirá conocer cuál será el nivel de financiación que tendrán los inversionistas, la capacidad de ahorro de los ciudadanos y el comportamiento del mercado financiero como un todo.

Por otro lado, ayudara conocer cuál será el valor que pueda tener la inversión, en la medida que el patrón de ahorro e inversión de un país da una idea de lo competitivo que puede volverse el mercado financiero. De igual manera, conocer el desempeño de la variable de intermediación, permitirá calcular el nivel de crecimiento en el mediano plazo y el nivel de desarrollo que tendrá el país, en la medida que sirve como variable explicativa en el cálculo de la acumulación de capital vía el componente de ahorro que en el largo plazo se convierte en inversión, dando una idea del movimiento de la curva de crecimiento de la economía en el largo plazo.

En otras palabras, el presente trabajo pretende modelar el componente de intermediación financiera para allanar el camino de futuras investigaciones que busquen mostrar la magnitud de estas relaciones descritas. Es decir, su justificación son dichas relaciones que no se comprobaran en el presente trabajo, pero que se consideran determinantes para prever el futuro económico del país y que en todo caso requieren del modelamiento de la variable de intermediación financiera.

Capítulo 3 Marco Teórico / Conceptual

Dado nuestro objetivo principal es estimar un modelo de pronóstico de la intermediación financiera, y teniendo en cuenta su influencia sobre el crecimiento del PIB, a través del análisis de series de tiempo, a continuación, se describirán las siguientes categorías para comprender como se presentación sus relaciones y dan respuesta a nuestro objetivo.

3.1. Crecimiento del PIB

Para establecer cómo se da el crecimiento del PIB, en primer lugar, se debe explicar que es y cómo está construido, así como señalar como crece o que actividades generales determinan su crecimiento. Cuando se habla de la riqueza de un país esta se mide a través de lo que se conoce como Producto Interno Bruto el cual:

“Mide el valor monetario de los bienes y servicios finales es decir, los que adquiere el consumidor final producidos por un país en un período determinado y cuenta todo el producto generado dentro de las fronteras Abarca los bienes y servicios producidos para la venta en el mercado, pero incluye también otros, como los servicios de defensa y educación suministrados por el gobierno.” (Callen, 2008)

Esta medición solo tiene en cuenta las actividades que generan un tipo de remuneración o valor que pueda contabilizarse mediante alguno de los mecanismos de información legales, bancos, reporte de impuestos, etc.; por ello, lo que se hace en lo que se conoce como mercado negro no se contabiliza.

Calle menciona que el Producto Interno Bruto se concibe desde tres enfoques uno de producción, de gasto y de ingreso:

- Enfoque de producción: suma el “valor agregado” en cada etapa de producción lo que corresponde al total de ventas menos el valor de los insumos intermedios utilizados en la producción.
- Suma el valor de las adquisiciones realizadas por los usuarios finales; por ejemplo, el consumo de alimentos, televisores y servicios médicos por parte de los hogares; la inversión en maquinarias por parte de las empresas, y las adquisiciones de bienes y servicios por parte del gobierno y de extranjeros
- Suma los ingresos generados por la producción; por ejemplo, la remuneración que perciben los empleados y el superávit operativo de las empresas (que equivale aproximadamente a las ventas menos los costos)

De otro lado, cuando se habla de los componentes del PIB se habla de los principales agregados, que, en los libros de texto, explican el comportamiento de este, el Consumo, la Inversión, el Gasto Público y el Mercado externo (Exportaciones menos Importaciones)

Sin embargo y de acuerdo con la forma como se contabilice el PIB los componentes pueden aludir a los sectores o cuentas en las que se agrupan las actividades económicas. Que, entre otras cosas, permite medir el desempeño de los sectores o actividades con mayor relevancia en la actividad económica. Para Colombia (DANE, 2010) se hablan de nueve sectores, estos son; agricultura, ganadería, caza, silvicultura y pesca; explotación de minas y canteras; industrias manufactureras; Suministro de electricidad, gas y agua; construcción; comercio, reparación, restaurantes y hoteles; transporte, almacenamiento y comunicaciones; establecimientos financieros, seguros, actividades inmobiliarias y servicios a las empresas; actividades de servicios sociales, comunales y personales

De esta manera el PIB busca hacer una medición de la economía de un país para poder compararlo con otras y así determinar su nivel de crecimiento, inversión y riesgo, ya que los elementos sobre los que se miden incluyen el gasto público, los ingresos generados por las diferentes actividades económicas, la capacidad de ahorro, etc. Este PIB al dividirse por el número de habitantes del país permite obtener el PIB per cápita.

“El PIB, junto con otras variables, se utiliza como medida del crecimiento económico de un país. Se suele asumir que los habitantes de un país con un PIB per cápita mayor se encuentran en una situación mejor que los de otro cuyo PIB per cápita sea menor. Así, los individuos de países con un nivel de bienestar material mayor, medido éste a través del PIB, estarían mejor que aquellos residiendo en un país en el que se disfrute un nivel de bienestar material inferior. Una mayor renta per cápita, es igual al PIB per cápita, significa asimismo un menú de opciones más amplio en cuanto a viajes, servicios de salud, etc.” (Armentia, 2014, p. 11)

Como se señaló anteriormente el PIB se puede medir de tres formas, una de ellas es la de oferta y valor agregado, en un Documento de la Universidad Católica de Oriente se señala que esta corresponde a la cantidad ofrecida de ciertos productos en forma generalizada y que en un país se mide a través de sus sectores económicos:

- Sector Primario: Abarca las actividades relacionadas con la explotación de los recursos naturales, tales como la agricultura, la caza, la pesca y la silvicultura. Las actividades de este sector generalmente no llevan procesos de transformación; es decir, los productos se venden tal como se extraen de la tierra.
- Sector Secundario: En este se incluyen las actividades en las cuales se transforman productos generalmente del sector primario o del mismo sector secundario, tales como la agroindustria, la producción de alimentos procesados, el plástico, los textiles, entre otros.
- Sector Terciario: Se agrupan en este sector las actividades en las cuales se producen bienes intangibles y generalmente dichas actividades sirven de soporte a los dos

sectores anteriores. Se encuentran el comercio, el transporte, la salud, la educación, el sector financiero, etc.

Esta medición permite establecer el peso de cada sector en el crecimiento de un país, de allí que sea la medición que se tendrá en cuenta en este documento, en tanto que se pretenden determinar la incidencia de una de las actividades del sector financiero en el crecimiento económico.

En la gráfica a continuación se presenta la composición actual del PIB en Colombia por sector para el año 2017:

Gráfica 1. PIB valor agregado 2017



Fuente: Dane

Como se observa el valor agregado del PIB para el 2017 fue de 7 puntos porcentuales, donde el sector con más crecimiento económico fue el sector financiero y de seguros el cual represento en este periodo el 5% del PIB; por otro lado, el sector de Comercio al por mayor y al por menor, reparación de vehículos automotores y motocicletas, transporte y almacenamiento, alojamiento y servicios de comida representó el 18% del PIB pero con un crecimiento del 6%; adicionalmente es importante mencionar que el sector con menor crecimiento fue el de la construcción con un crecimiento de -3% siendo el 7% del PIB.

En cifras del DANE del periodo 2005 a 2017, se observa que sectores como un crecimiento promedio de más del 10% fueron los sectores de construcción, actividades financieras y de seguros, así como el sector de actividades profesionales, científicas, y técnicas, actividades de servicios administrativos y de apoyo; por otro lado, los sectores con menor crecimiento fueron el de industrias manufactureras y el sector de la información y comunicaciones.

Como se observa la composición del PIB en Colombia, es principalmente del sector secundario y terciario, siendo este último quien mayor peso tiene, es decir, que el crecimiento económico como valor agregado depende principalmente de la transformación de productos y la prestación de servicios.

En la siguiente sección y teniendo en cuenta el comportamiento del sector financiero en el PIB se describirá la composición de este sector, es decir cada una de las actividades que aportan a la generación de valor a partir de los servicios financieros y de seguros, para finalizar en la tercera sección en la importancia de analizar esta información apoyados en el uso de series de tiempo.

3.2. Intermediación Financiera

Como vimos en el anterior capítulo el sector financiero y de seguros es uno de los sectores que ha mantenido un crecimiento alto en el PIB Colombiano, estos datos ya se registraban en 2014 cuando señalaba que el sector “que más incidencia tiene en la formación del PIB es, el de los establecimientos financieros, seguros, actividades inmobiliarias y servicios a las empresas, que representó en el año 2014 el 18,8% del PIB” (Escuela Nacional Sindical, 2015, p. 11)

Este sector está constituido por “las instituciones financieras y sus fondos administrados, bajo la vigilancia de la Superintendencia Financiera donde se encuentran las siguientes clases de instituciones: Los establecimientos de Crédito, las sociedades de servicios financieros, las sociedades de capitalización, las entidades aseguradoras y los intermediarios de seguros y reaseguros

Dentro del primer grupo se encuentran los establecimientos bancarios, las corporaciones financieras, las compañías de financiamiento comercial y las cooperativas financieras; en el segundo grupo se encuentran las sociedades fiduciarias, los almacenes generales de depósito, las sociedades administradoras de pensiones y cesantías y las sociedades de intermediación cambiaria y de servicios financieros especiales, SICA y SFE; En el tercer grupo el de las sociedades de capitalización no se encuentran divididas en grupos; el cuarto grupo el de las entidades aseguradoras está constituido por las compañías de seguros, las compañías de Reaseguros y las cooperativas de seguros; el último grupo se encuentra conformado por los corredores de seguros, las agencias de seguros, los agentes de seguros y los corredores de reaseguros.

La Superintendencia financiera (s.f) define cada uno de estos grupos:

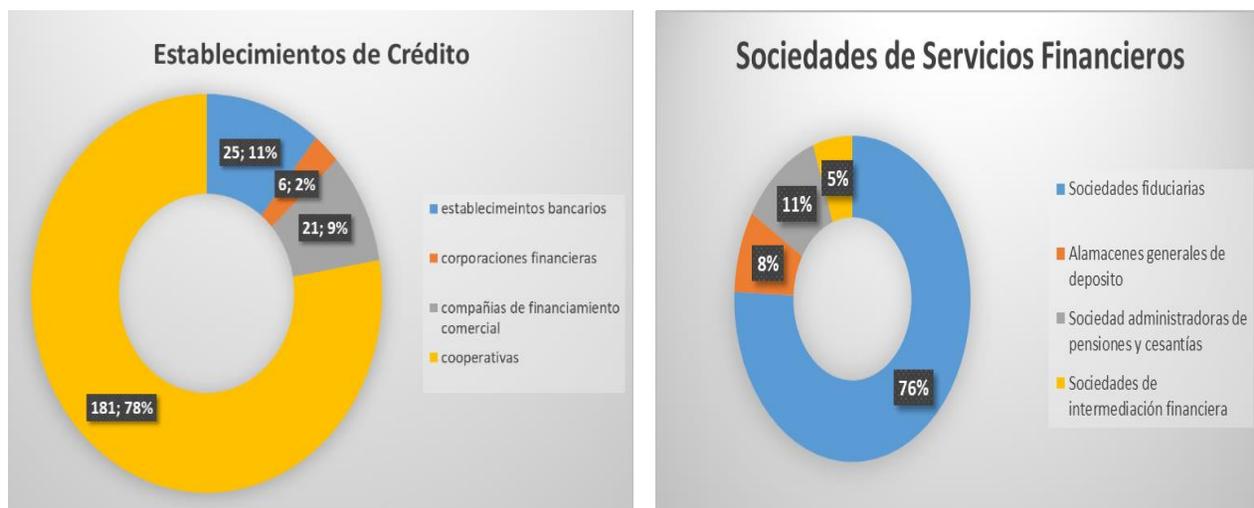
- Establecimientos de crédito: son instituciones financieras cuya función principal consiste en captar en moneda legal recursos del público en depósitos a la vista (cuentas de ahorro, corriente) o a término (CDT y CDAT´S), para colocarlos nuevamente a través

de préstamos, descuentos, anticipos u otras operaciones activas de crédito. Son establecimientos de crédito” (Superintendencia Financiera de Colombia, s.f.)

- Las sociedades de servicios financieros son sociedades que tienen por función la realización de las operaciones previstas en el régimen legal que regula su actividad, si bien captan recursos del ahorro público, por la naturaleza de su actividad se consideran como instituciones que prestan servicios complementarios y conexos con la actividad financiera.
- Las sociedades de capitalización: Son instituciones financieras cuyo objeto consiste en estimular el ahorro mediante la constitución, en cualquier forma, de capitales determinados, a cambio de desembolsos únicos o periódicos, con posibilidad o sin ella de reembolsos anticipados por medio de sorteos.
- Entidades Seguradoras: Su objeto es la realización de operaciones de seguro, bajo las modalidades y los ramos facultados expresamente
- Intermediarios de seguros y reaseguros: Es quien hace o sirve de enlace o mediador entre dos o más personas (naturales o jurídicas); y más específicamente entre productores y consumidores”; es decir, entre la empresa aseguradora y el propuesto tomador-asegurado-beneficiario de una póliza determinada de seguros.

De estas entidades, las que son de relevancia para este documento son los establecimientos de crédito y las sociedades de servicios financieros, de las primeras existen actualmente 233 de las cuales las cooperativas financieras representan el 78% de este grupo; en cuanto a las sociedades de servicios financieros existen 37 de las cuales el 76% lo representan las sociedades fiduciarias.

Gráfica 2 Entidades que componen el sector financiero



Fuente: Superfinanciera

En cuanto a la participación de este sector de la economía el informe anual de la superintendencia financiera señala:

“El sistema financiero cerró el 2017 con un desempeño positivo: Los activos del sistema financiero registraron un crecimiento real anual de 7.6% alcanzando un valor de \$1,566 billones. En términos absolutos, los activos presentaron un incremento anual de \$167.3, comportamiento que permitió alcanzar un índice de profundización frente al PIB de 173% esto se ve reflejado principalmente por el aumento en las inversiones y la cartera neta en \$115.2b y \$25.7b, respectivamente. El crecimiento mensual de los activos fue de \$10.2b, de las inversiones \$7.7b y de la cartera de \$1.6b.” (Superintendencia Financiera de Colombia, 2018)

Es decir, el crecimiento en el sector ha sido fuerte y estable y la tendencia se mantiene, esta situación la evidencia además el crecimiento de los activos de los establecimientos de crédito:

“Los activos de los establecimientos de crédito aumentaron 34.5b en 2017. Los activos de los Establecimientos de Crédito (EC) se ubicaron en \$608.5b tras presentar un incremento anual de 34.5b (1.8% real anual), principalmente por incrementos en la cartera de créditos y operaciones de leasing financiero y en las inversiones y operaciones con derivados de \$20.3b y \$10.5b, respectivamente. Frente a noviembre, el saldo en inversiones y operaciones con derivados pasó de \$113.5b a \$113.7b, donde los instrumentos de patrimonio de emisores nacionales aportaron la mayor contribución, con \$780.1mm” (Superintendencia Financiera de Colombia, 2018)

Dentro de este sector las actividades que se desarrollan y que permiten que estos crecimientos se den son la intermediación financiera que en palabras de la superintendencia consiste en captar dinero del público y prestarlo pero que solo puede ser ejercida por entidades sometidas a la inspección, vigilancia y control de la Superintendencia.

Esta actividad de captación de recursos del público se hace mediante la realización de operaciones pasivas o de recepción de fondos con el fin de colocarlos, también en forma masiva, a través de la ejecución de operaciones activas o de otorgamiento de créditos (prestamos), gestión que por su naturaleza requiere previa autorización administrativa señala la superintendencia financiera (2017).

Teniendo en cuenta el crecimiento del sector y las actividades de intermediación financiera, es coherente señalar que existe una relación directa en entre estas actividades y más cuando:

“Un mayor desarrollo de los servicios que el sistema bancario presta al sector privado, incluyendo los servicios tradicionales de intermediación y los relacionados con el manejo del riesgo y la información, propicia mayores tasas de crecimiento económico. Algo similar puede esperarse de una mayor capacidad del sistema para generar crédito y favorecer una mayor redistribución de éste hacia el sector privado” (Fernando G & Garcia M, 1995),

Este elemento de intermediación financiera se mide a través de la tasa de margen de intermediación financiera la cual se hace cuatrimestral; estos datos miden los ingresos percibidos por la actividad de intermediación en el país, por ello y como nuestro objetivo es determinar el nivel de influencia que esta ha tenido en el crecimiento económico de Colombia en el periodo de 2012 a 2015 , el proceso requerirá un análisis de series de tiempo, que permita hacer un proceso objetivo y transparente sobre la información existente en esta materia.

3.3. Series de tiempo

Como señalamos anteriormente nuestro objetivo es estimar un modelo de pronóstico de la actividad de intermediación financiera, para ello se usará el análisis de series de tiempo que en su concepto más sencillo se refiere a “los datos estadísticos que se recopilan, observan o registran en intervalos de tiempo regulares (diario, semanal, semestral, anual, entre otros)” (Universidad de Sonora Mexico)

Las series de tiempo tienen varios componentes que unidos generan cambios en la serie analizada: Tendencia Secular, variación estacional, variación cíclica y variación irregular, a continuación, se describirá cada una de ellas de acuerdo con el curso de estadística de la universidad de Sonora México:

- **Tendencia secular:** La tendencia secular o tendencia a largo plazo de una serie es por lo común el resultado de factores a largo plazo. En términos intuitivos, la tendencia de una serie de tiempo caracteriza el patrón gradual y consistente de las variaciones de la propia serie, que se consideran consecuencias de fuerzas persistentes que afectan el crecimiento o la reducción de la misma, tales como: cambios en la población, en las características demográficas de la misma, cambios en los ingresos, en la salud, en el nivel de educación y tecnología. Las tendencias a largo plazo se ajustan a diversos esquemas. Algunas se mueven continuamente hacia arriba, otras declinan, y otras más permanecen igual en un cierto período o intervalo de tiempo.
- **Variación estacional:** El componente de la serie de tiempo que representa la variabilidad en los datos debida a influencias de las estaciones, se llama componente estacional. Esta variación corresponde a los movimientos de la serie que recurren año tras año en los mismos meses (o en los mismos trimestres) del año poco más o menos con la misma intensidad. Por ejemplo: Un fabricante de albercas inflables espera poca actividad de ventas durante los meses de otoño e invierno y tiene ventas máximas en los de primavera y verano, mientras que los fabricantes de equipo para la nieve y ropa de abrigo esperan un comportamiento anual opuesto al del fabricante de albercas
- **Variación cíclica:** Con frecuencia las series de tiempo presentan secuencias alternas de puntos abajo y arriba de la línea de tendencia que duran más de un año, esta variación se mantiene después de que se han eliminado las variaciones o tendencias estacional e irregular. Un ejemplo de este tipo de variación son los ciclos comerciales cuyos períodos recurrentes dependen de la prosperidad, recesión, depresión y recuperación, las cuales no dependen de factores como el clima o las costumbres sociales

- Variación Irregular: Esta se debe a factores a corto plazo, imprevisibles y no recurrentes que afectan a la serie de tiempo. Como este componente explica la variabilidad aleatoria de la serie, es impredecible, es decir, no se puede esperar predecir su impacto sobre la serie de tiempo. Existen dos tipos de variación irregular: a) Las variaciones que son provocadas por acontecimientos especiales, fácilmente identificables, como las elecciones, inundaciones, huelgas, terremotos. b) Variaciones aleatorias o por casualidad, cuyas causas no se pueden señalar en forma exacta, pero que tienden a equilibrarse a la larga.

Las series de tiempo adicionalmente se dividen en dos grupos como señala (Villavicencio, 2012)

- Estacionarias: cuando la serie es estable a lo largo del tiempo, es decir, cuando la media y varianza son constantes en el tiempo. Esto se refleja gráficamente en que los valores de la serie tienden a oscilar alrededor de una media constante y la variabilidad con respecto a esa media también permanece constante en el tiempo.
- No estacionarias. Son series en las cuales la tendencia y/o variabilidad cambian en el tiempo. Los cambios en la media determinan una tendencia a crecer o decrecer a largo plazo, por lo que la serie no oscila alrededor de un valor constante

Adicionalmente a lo señalado, lo que se espera al hacer un análisis de una serie de tiempo es ver su tendencia, esta puede ser lineal, o no lineal; la primera muestra que algo aumenta o disminuye a un ritmo constante y el método que se utiliza para obtener la línea recta de mejor ajuste es el Método de Mínimos Cuadrados; la segunda tendencia es cuando el comportamiento de la serie es curvilíneo, la cual puede ser polinomial, logarítmica, exponencial y potencial, entre otras.

Una serie de tiempo puede estar influenciada por datos atípicos que no permiten ver la tendencia de manera clara, para ello se usan los métodos de suavización de serie, entre estos se encuentra el método de promedio móvil se construye sustituyendo cada valor de una serie por la media obtenida con esa observación y algunos de los valores inmediatamente anteriores y posteriores, el otro método son los promedios móviles ponderados que consiste en asignar un factor de ponderación distinto para cada dato. Generalmente, a la observación o dato más reciente a partir del que se quiere hacer el pronóstico, se le asigna el mayor peso, y este peso disminuye en los valores de datos más antiguos.

Un tercer método de suavización señala el mismo curso es suavizamiento exponencial el cual emplea un promedio ponderado de la serie de tiempo pasada como pronóstico; es un caso especial del método de promedios móviles ponderados en el cual sólo se selecciona un peso o factor de ponderación.

Este tipo de metodología de análisis estadístico es la mejor herramienta para analizar la evolución de información es decir sobre el crecimiento, variación, etc; por eso es

pertinente para el análisis de información económica como lo es el PIB y cada uno de sus componentes esto tal como lo señala (Wicherm & Hanke, 2006)

“Para que las organizaciones grandes, medianas o pequeñas, privadas o públicas, puedan reducir el grado de incertidumbre causado por el cambio constante del entorno, deben respaldar sus decisiones en algo más que la experiencia o la intuición, deben respaldarlo en la elaboración de pronósticos precisos y confiables que sean suficientes para satisfacer las necesidades en la planeación de la organización” (Contreras Juárez, Zuniga, Martínez Flores, & Sánchez Partida, 2016)

Partiendo de esto y teniendo en cuenta que, a través de las series de tiempo, analizando la cadena de información se pueden determinar las principales situaciones que en nuestro caso particular hacen variaciones en el proceso de intermediación financiera haciendo que el crecimiento del país igualmente varié.

Varios estudios señalan que el uso de series de tiempo para el análisis del crecimiento permite adicionalmente pronosticar los comportamientos de este en un determinado periodo, estos pronósticos se elaboran a través de modelos, en los cuales se busca “manipular” cada uno de los elementos que componen las series de tiempo, para que sean objetivas, sin afectación de datos atípicos y acercadas a la realidad.

De allí que Hernández et al, señale que:

“El manejo de las Series de Tiempo es de vital importancia en planeación y en áreas del conocimiento donde evaluar el efecto de una política basada sobre una variable, y/o conocer predicciones de sus valores futuros, aportan criterios que disminuyen el riesgo en la toma de decisiones o en la implementación de políticas futuras” (HERNANDEZ S. , PEDRAZA M., & ESCOBAR DIAZ , 2008)

Capítulo 4 Marco Metodológico

El presente marco metodológico está basado en los módulos preparados como guías de clase, puntualmente el trabajo de (Lozano F, 2017). Para temas más precisos, sobre gráficos o pruebas de hipótesis, se usó el libro de Hyndman (Hyndman & Athanasopoulos, 2018) y Cowperwait y Metcalfe (Cowperwait & Metcalfe, 2009). Para otro particular se informará la fuente de la información.

Como se explicó con anterioridad se trabajará con las herramientas de análisis de series, es decir considerado que los datos son un conjunto de observaciones indexadas por el tiempo, los cuales cumplen con los supuestos de que cada observación es consecutiva y esta espaciada en igual periodo de tiempo y que además es discreto. Y se busca, con la historia de las propias observaciones, hacer un pronóstico. Para esto, se debe asegurar en un primer momento, que la serie cumpla con el supuesto de la estacionariedad, de tal manera que se pueda inferir o generalizar los resultados. También, es necesario asegurar la dependencia que tenga la serie de su propia historia, es decir estableciendo su autocorrelación y el nivel de varianza que presenta la serie a lo largo de su recorrido en un análisis de autocovarianza.

Para establecer el nivel de relación de la serie consigo misma se utilizan los autocorrelogramas, que busca evaluar que rezagos u observaciones anteriores son las que más afectan el presente y futuro de la serie, tanto con una relación directa como por el efecto de sus errores. Se busca de este modo estimar parámetros que den información sobre el comportamiento de la serie y que, dado el cumplimiento de los supuestos de estacionariedad, varianza y media constante, se pueda estimar futuros posibles comportamientos de la serie. La aplicación y verificación de los expuesto se usará la función ACF Y PACF de R.

Para los pronósticos de proyección de la intermediación financiera, se utilizará un modelo univariado tipo ARIMA (p, d, q) , este se estima siguiendo la metodología Box-Jenkins (1978), para este tipo de modelos se parte de considerar que el comportamiento del proceso estocástico puede ser explicado por el efecto de los valores pasados y por la suma ponderada de los errores de dichos rezagos.

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - L^d)Y_t = \delta + (1 + \theta_1 B + \dots + \theta_1 B^p)Z_t$$

Donde ϕ_i es el coeficiente autorregresivo de orden i . El parámetro θ_j Corresponde al coeficiente de promedio móvil de orden j . El valor de d , hace referencia al numero de diferencias que requiere la variable para cumplir el supuesto de estacionariedad. δ Es un valor constante. B^k Es un operador que cumple con $B^k Y_t = Y_{t-k}$ y Z_t es un proceso de ruido blanco con media cero y varianza σ^2 constante.

Para entender mejor se explicará cada uno de los componentes del Modelo ARIMA. En un primer momento se explicará el Modelo AR, luego el MA, el ARMA, el operador de diferencia para entender los componentes del modelo ARIMA y finalmente el componente estacional.

4.1. Modelos para series estacionarias Autorregresivos (AR) y Media Móvil (MA)

En el Modelo de Autorregresivo (AR), se busca establecer el nivel de dependencia directa de la observación presente con su pasado, pues es una regresión que toma como variable dependiente el presente y_t y lo regresa con sus rezagos. El orden (p) indica el número de rezagos que más le afectan. De esta manera $y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$, Donde, ϕ_1, \dots, ϕ_p son la pendiente de los parámetros autorregresivos, el error es un ruido blanco (con media cero y varianza constante). Esta ecuación se puede reescribir utilizando un operador de desfase, quedando como $y_t = c + \phi_1 B y_t + \phi_2 B^2 y_t + \dots + \phi_p B^p y_t + \varepsilon_t$, esta ecuación se puede despejar y factorizar para hallar el polinomio auto-regresivo $\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$, dando como resultado, $\phi(B)y_t = c + \varepsilon_t$ que es la nomenclatura o fórmula del modelo auto-regresivo de orden (p).

El modelo puede entenderse de manera práctica, como un proceso en el que lo que pasa hoy esta explicado por el efecto de una constante, más la interacción de la pendiente por el efecto de ayer, más un error. Debe tenerse presente que el error es un ruido blanco (es decir presenta una media cero y una varianza constante). O si se considera, se puede quitar el efecto de la media restando a ambos lados de la ecuación, que de manera formal quedaría; $y_t - \mu = \phi(y_{t-1} - \mu) + \varepsilon_t$ para un modelo AR (1).

El modelo AR (1) tiene un gran poder explicativo; permite entre otras cosas dimensionar las propiedades del parámetro de la pendiente ϕ indicando:

- Si $\phi = 0$, se tiene un paseo aleatorio.
- Si $\phi \neq 0$, y_t depende o del rezago o del error y se habla de un proceso autocorrelacionado.
- Si ϕ presenta valores positivos altos, hay presencia de una alta autocorrelación, por el contrario, si se trata de valores negativos, se trata de series de tiempo oscilatorias.

En todo caso la ϕ debe ser un valor entre -1 y 1, y en casos de un mayor orden, la suma no debe ser superior a 1, ni inferior a -1. La identificación de la autocorrelación se puede hacer por medio de las gráficas de autocorrelogramas; uno simple nos muestra la persistencia o no de dicha autocorrelación según su configuración (una caída rápida indica una persistencia baja, es decir que los efectos del pasado tienen poco impacto en el presente de la serie, al contrario, una caída lenta indica una alta persistencia del pasado en el presente de las observaciones), el parcial indica

el orden de autocorrelación.

Una vez explicado el proceso de auto regresión, se procederá a explicar el componente de Medias Móviles o modelo MA, al igual que el proceso de autoregresión, este tiene en cuenta la influencia del pasado de la serie para explicar el presente y futuro de la serie, sin embargo, este proceso considera la influencia de los errores, es decir la diferencia entre el valor estimado versus el valor real de la serie. En este caso el MA tiene un proceso de orden q . $y_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \dots + \theta_q\varepsilon_{t-q}$. Utilizando el operador de desfase, como se hizo para el modelo AR, se obtiene que el MA(q)= $y_t = \mu + \theta_1B + \theta_2B^2 + \dots + \theta_qB^q + \varepsilon_t$. De igual manera, el polinomio de media móvil $\theta(B) = (1 - \theta_1B - \theta_2B^2 - \dots - \theta_qB^q)$, donde la media de la variable es la esperanza de la misma, que es igual a su media, y su varianza esta es $Var(y_t) = \sigma^2(1 + \theta_1^2 + \dots + \theta_q^2)$, y las autocovarianzas son
$$\begin{cases} (\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \dots + \theta_{j+q}\theta_q)\sigma^2, & \text{Para todo } j = 1, 2, \dots, q, \\ 0, & \text{Para todo } j > q \end{cases}$$

De manera práctica, el modelo MA (1) puede ser entendido como la observación presente igual a la media, más el ruido o error, más la pendiente multiplicada por el error o ruido de ayer, o formalmente $y_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1}$. Donde, el ruido o error se asume como un ruido blanco, es decir con media cero y varianza constante, y los parámetros de media móvil tienen las siguientes propiedades;

- Si $\theta \neq 0$, y_t depende o del rezago o del error y se habla de un proceso autocorrelacionado.
- Si θ presenta valores positivos altos, hay presencia de una alta autocorrelación, por el contrario, si se trata de valores negativos, se trata de series de tiempo oscilatorias.

En todo caso la θ debe ser un valor entre -1 y 1, y en casos de un mayor orden, la suma no debe ser superior a 1, ni inferior a -1. La identificación de la media móvil se puede hacer por medio de las gráficas de autocorrelogramas; uno parcial nos muestra la persistencia o no de dicha autocorrelación según su configuración (una caída rápida indica una persistencia baja, es decir que los efectos del pasado tienen poco impacto en el presente de la serie, al contrario, una caída lenta indica una alta persistencia del pasado en el presente de las observaciones), el simple indica el orden de la media móvil.

Continuando con la explicación de los componentes del Modelo, se procederá con el modelo ARMA (p, q). Este busca incorporar las propiedades de los modelos autorregresivos y de media móvil, es decir $y_t = \phi_1B y_t + \phi_2B^2 y_t + \dots + \phi_pB^p y_t + \theta_1B + \theta_2B^2 + \dots + \theta_qB^q + \varepsilon_t$ donde el error es un ruido blanco (con media cero y varianza constante) y ϕ y θ son los parametros autoregresivos y de media movil. Usando el operador de desfase se puede escribir como; $\phi(B)y_t = \theta(B)\varepsilon_t$ siendo $\phi(B)$ y $\theta(B)$ los polinomios AR y MA. De manera formal el modelo queda expresado como $y_t - \mu = \phi(y_{t-1} - \mu) + \varepsilon_t + \theta_1\varepsilon_{t-1}$.

4.2. Modelos para series No Estacionarias (ARIMA y SARIMA)

En términos prácticos, el Modelo ARMA, y por tanto sus componentes AR y MA, se orientan hacia series que cumplen con el supuesto de estacionariedad (media y varianza constante en el tiempo), sin embargo en la realidad garantizar el cumplimiento de este supuesto de estabilidad es difícil, las series económicas, por ejemplo, rara vez tienen un comportamiento determinista y por tanto se debe acudir a herramientas que permitan transformar la serie y permitir el cumplimiento de los supuestos para poder extrapolar los resultados hacia el comportamiento futuro de la serie.

Una primera forma de lograr la transformación de la serie es el uso del operador de diferencia. Este busca, a través de la eliminación de la magnitud de los valores, estabilizar los datos alrededor de la media. El procedimiento consiste en la resta, al valor presente el valor anterior, esta resta (en términos absolutos) permite conocer cuál es el crecimiento o cambio que tiene la variable de un periodo a otro dejando de lado valores de magnitud que distorsionan el valor de las variables. Es decir, pone la atención únicamente en el cambio de un periodo a otro sufrido por la variable que es donde reviste importancia cualquier análisis o estimación que se desee hacer. Se puede decir que se quita el ruido a los datos. La diferenciación tiene el inconveniente de acortar la serie, en el entendido que, al restar el valor anterior a todos los valores de la serie, se pierde el primer valor. Sin embargo, el poder de la diferenciación, de coaccionar los datos hacia la estacionariedad, es altamente útil, pero a pesar de esto no debe hacerse más de tres veces.

En términos formales el operador diferencia se define como; $\Delta^n y_t = \Delta^{n-1} y_t - \Delta^{n-1} y_{t-1} = \sum_{r=0}^n (-1)^r \binom{n}{r} y_{t-r}$

El uso del operador de diferencia implica el cálculo de un nuevo parámetro que transforma los modelos explicados y da paso a uno nuevo modelo, el ARIMA (p, d, q), donde la d indica el orden de diferenciación que se requiere para transformar la serie y permitir cumplir con el supuesto de estacionariedad. Es decir si la serie no es estacionaria una integración de orden d puede hacer que sea estacionaria; $\Delta^n y_t = (1 - B)^d y_t$ De manera General; $\phi(B)[(1 - B)^d y_t - \mu] = \theta(B)\varepsilon_t$, donde $\phi(B)$ y $\theta(B)$ son los polinomios AR y MA, y $(1 - B)^d$ es el orden general de integración necesario para hacer que la serie sea estacionaria.

Como se ha repetido constantemente, las series económicas presentan problemas de tendencia, que implican el incumplimiento del supuesto de estacionariedad, por lo que se debe buscar identificar cuál es su tipo para poder hacer un análisis acorde a los datos. Una forma para determinar la existencia de esta, más allá de un diagnostico visual o gráfico, son las pruebas de raíz unitaria, que además de determinar la existencia de la raíz unitaria ayuda a determinar el tipo de procedimiento a seguir para poder asegurar un buen comportamiento. A continuación, se presentan dos de las pruebas más usadas la Dickey-Fuller y la Phillips-Perron. En la parte práctica del presente trabajo también se usará la Kwiatkowski.

La Prueba de raíz unitaria Dickey-Fuller (DF) sirve para comprobar que la serie es integrada de orden uno, es decir que no presenta problemas de estacionariedad, de manera directa con el paquete ts series de R. De otro lado, sirve para determinar cuál tipo de prueba es mejor para descartar la presencia de la raíz unitaria, es decir ayuda a determinar cuál es la que mejor aproxima a la realidad de la serie original, se basa en asumir que la serie se puede aproximar por un proceso AR (1) con tres variantes: media cero, media diferente de cero (con una constante) y tendencia lineal. En el paquete URCA de R se presenta por medio de tres argumentos none, drift, y trend. La prueba tiene como hipótesis nula que la serie es no estacionaria, por tanto, se busca rechazar esta hipótesis, para aceptar la nula (que indica que la serie es estacionaria) y se cumple cuando tau sea menor que el valor crítico. A continuación, se explica el contenido de la prueba dado el alto valor interpretativo que da. La explicación está basada en Enders (Enders, 2014).

- En el primer caso se supone que $y_t \sim AR(1)$ con media es cero, es decir la prueba con el argumento “none” $\Delta y_t = \gamma y_{t-1} + \varepsilon_t$; La hipótesis nula es **H₀**: $\gamma = 0$ versus la alterna **H_a**: $\gamma < 0$ El estadístico de la prueba se denota por τ (tau) y su distribución bajo H₀ permite calcular los valores críticos, de tal forma que el criterio de rechazo es $\tau < \hat{\tau}_{0,05}$, con $\hat{\tau}$ es valor calculado del estadístico. Es decir, cuando el estadístico de la prueba está dentro de las tres regiones (1%, 5%, 10%) y no se puede rechazar la hipótesis nula, se debe suponer que los datos son una caminata aleatoria, es decir, que hay una raíz unitaria presente.
- En el segundo caso se supone que $y_t \sim AR(1)$ con media diferente de cero, es decir con constante, la prueba de raíz unitaria Dickey-Fuller con el argumento “drift” $\Delta y_t = \alpha + \gamma y_{t-1} + \varepsilon_t$. Donde α (alfa) es refiere a la constante, o término de deriva, "tau2" sigue siendo la hipótesis gamma es cero ($\gamma=0$) nula. Se presume una raíz unitaria, gamma=0. El término phi1 se refiere a la segunda hipótesis, que es una hipótesis nula combinada de $\alpha = \gamma = 0$ (alfa y gamma son igual a cero). Esto significa que AMBOS valores son probados para ser cero al mismo tiempo. Si $p < 0.05$, se rechaza la hipótesis nula, y se supone que POR LO MENOS uno de los dos es falso - es decir, uno o ambos de los términos α o γ (alfa o gamma) no son cero. No rechazar la hipótesis nula implica que AMBOS α o γ (alfa o gamma) son igual a cero, implicando 1) que $\gamma = 0$ (gamma igual a cero) por lo tanto una raíz de unidad está presente, Y 2) $\alpha = 0$ (alfa igual a cero), así que no hay término de deriva.
- Finalmente, el tercer caso para $y_t \sim AR(1)$ con tendencia lineal, la prueba de raíz unitaria Dickey-Fuller con el argumento “trend” $\Delta y_t = \alpha + \beta_t + \gamma y_{t-1} + \varepsilon_t$. donde β_t es un término de tendencia temporal. Las hipótesis son las siguientes: tau: $\gamma=0$ (gamma es igual a cero), phi3: $\beta_t = \gamma = 0$ phi2: $\alpha = \beta_t = \gamma = 0$. En este caso la hipótesis nula es que la serie es no estacionaria. Lo que tau3 implica, como se explicó antes, es que no se puede rechazar la hipótesis nula de que existe raíz unitaria, implicando que en efecto hay presencia de una raíz unitaria. No rechazar phi3 implica dos cosas: 1) $\gamma = 0$. (raíz de la unidad) Y 2) no hay un término de tendencia temporal, es decir, $\beta_t = 0$.. Si rechazamos esta hipótesis nula, implicaría que uno o ambos de estos términos no era cero No rechazar phi2 implica tres cosas: 1) $\gamma = 0$. Y 2) ningún término de tendencia temporal Y 3) ningún término de deriva,

es decir, que $\alpha = \beta_t = \gamma = 0$. Rechazar la hipótesis nula implica que uno, dos, o los tres términos NO eran cero. En este sentido estamos en presencia de una raíz unitaria en la tendencia.

En términos prácticos lo que busca la prueba es determinar la existencia de la raíz unitaria y orientar si está asociada a una constante o a la tendencia lineal.

La prueba Phillips-Perrón (PP), al igual que la DF, se orienta a determinar la existencia de una raíz unitaria en la serie a analizar, toma como hipótesis nula la no estacionariedad de la serie y la regla de rechazo será si el p-valor es menor al valor crítico. La prueba se hará con los paquetes tseries y URCA del software estadístico r, particularmente con las funciones pp.test y ur.pp.

De otro lado, la presencia de la no estacionaridad se puede explicar por el efecto cíclico de las observaciones, es decir un fenómeno puede estar explicado por la periodicidad con la que son tomadas las series, la presencia de eventos que están atados a un ciclo temporal, etc. Es decir, el evento puede tener ocurrencia por factores que solo se dan cada periodo, ejemplo el volumen de ventas debe tener en cuenta el ciclo laboral, educativo, productivo, etc., o la serie misma muestra estos ciclos, el PIB por ejemplo, puede estar influido por lo que sucede anualmente; por el tema de las compras de fin de año, o por el consumo de energía en ciertos periodos, y depende de la periodicidad con la que los datos son tomados, lo que configura gráficamente un ciclo, es decir una forma oscilatoria que coincidirá con la forma con la que se toman los datos.

La situación anterior demanda la utilización de herramientas de tipo estacional que tengan en cuenta la periodicidad descrita, para este tipo de series se puede hacer uso de modelos SARIMA, que no es otra cosa que Modelos de tipo ARIMA con un componente que tiene en cuenta el problema periódico de la serie. Teniendo un modelo General ARIMA (p,d,q), explicado con anterioridad, $\phi(B)[(1-B)^d y_t - \mu] = \theta(B)\varepsilon_t$ que presenta ciclos periodicos (gráficamente se observan después de diferenciar), se debe considerar el uso del componente estacional (seasonal, s) u operador estacional; $\Delta_s^D = (1-B^s)^D$ (donde D indica el número de diferencias estacionales necesarias para que la serie logre capturar el efecto periódico) que debe corresponder con un operador Autorregresivo (AR) de orden p y estacional ; $\Phi(B)^s = 1 - \Phi_1 B^s - \dots - \Phi_p B^{sP}$, o de Media Movil (MA) de orden q y estacional; $\Theta(B)^s = 1 - \Theta_1 B^s - \dots - \Theta_q B^{sQ}$, de esta manera el Modelo ARIMA (p,d,q) se convierte a un SARIMA (p,d,q)X(P,D,Q)s, donde la parte con letras mayúscula corresponde al componente estacional y la parte en letras minúsculas corresponde al componente no estacional que siempre se ha usado, y estará determinado por;

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - \Phi_1 B^s - \dots - \Phi_p B^{sP})[(1-B)^d(1-B^s)^D - \mu]y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)(1 - \Theta_1 B^s - \dots - \Theta_q B^{sQ})\varepsilon_t.$$

El error se comporta como un ruido blanco, con media cero y varianza σ^2 constante. Usando el operador de desfase y factorizando se puede reescribir la anterior formula como;

$$\phi(B)\Phi(B)^s[(1-B)^d(1-B^s)^D - \mu]y_t = \theta(B)\Theta(B)^s\varepsilon_t.$$

En términos prácticos; un ARIMA (1,1,1) X (1,1,1) o SARIMA (1,1,1) X (1,1,1) para una serie trimestral sería;

$$\frac{(1 - \phi_1 B)}{\text{AR}(1)} \frac{(1 - \Phi_1 B^4)}{\text{AR}(1)} \frac{(1 - B)}{\text{Diferenciacion}} \frac{(1 - B^4)}{\text{Diferencia}} y_t = \frac{(1 + \theta_1 B)}{\text{MA}(1)} \frac{(1 + \Theta_1 B^4)}{\text{MA}(1)} \frac{\varepsilon_t}{RB}$$

NO estacional estacional NO estacional estacional NO estacional Estacional (0, σ^2)

Teniendo en cuenta lo anterior y que el proceso de ajustar un modelo es un proceso de calibración, no automático, se hará uso de la metodología de Box-Jenkins, se trata de una metodología iterativa para hacer estimaciones (predicciones) a partir de una serie de tiempo que cumpla con criterios de estabilidad en cada momento del tiempo (es decir ser estacionaria; tener una media y una varianza constante), lo que garantiza que la probabilidad de ocurrencia de los eventos sea estable y que las conclusiones de un momento sean generalizables a todo el proceso estocástico.

4.3. Metodología Box-Jenkins

La metodología se aplica a dos tipos puntuales de Modelos; los Autorregresivos de Media Móvil (ARMA), y los Modelos Autorregresivos Integrados de Media Móvil (ARIMA), se basa en el uso de los mismos datos para determinar cuál es el mejor (se entiende como mejor aquel que, por el principio de parsimonia, integre el menor número de parámetros). La metodología consta de 4 pasos, que como se dijo anteriormente deben repetirse hasta lograr la estimación del mejor modelo;

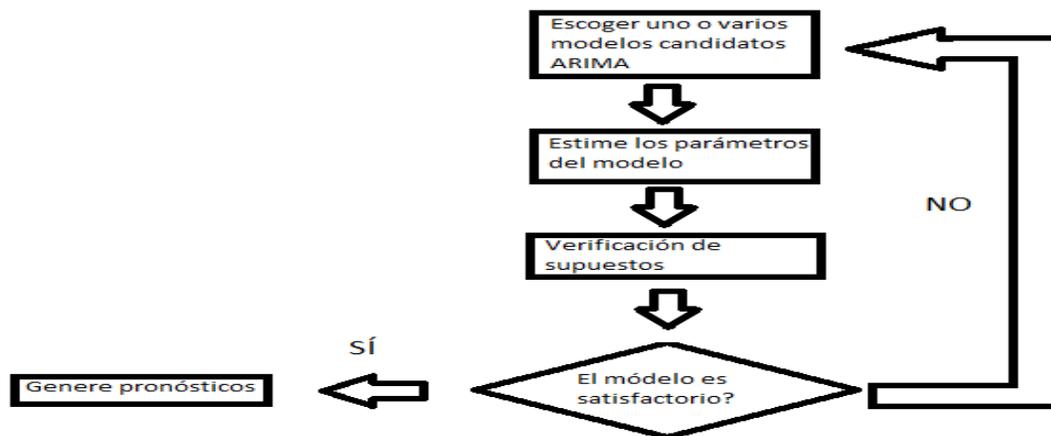
- a. Se postula una clase de modelo; si es auto regresivo (a partir de estimar su autocorrelación parcial se logra identificar cuáles son los rezagos que mayor impactan en la ocurrencia del evento en el presente. Esto es, se hace una regresión con los sus rezagos y se escogen los estadísticamente más significativos), si es Media Móvil (se identifican a partir de los rezagos cuales son los que presentan mayor impacto o afectación por eventualidades externas. A partir de la estimación de la función de autocorrelación se logra identificar el comportamiento de los rezagos y como ha variado la media en cada momento del tiempo, con lo que se logra estimar el comportamiento en el futuro de esta), o mixtos.
- b. De acuerdo con los datos, se verifica o identifica el modelo al cual se ajustan los datos, teniendo en cuenta lo descrito en el punto anterior, es decir en cuál es mejor la configuración los datos de la serie, en este punto se recomienda elaborar los autocorrelogramas a partir de la estimación de los coeficientes de autocorrelación; simple para identificar los modelos de media móvil y parcial para identificar los modelos autoregresivos.
- c. Estime los parámetros del modelo seleccionado; con la función de autocorrelación identifique los rezagos, es decir el orden de cada uno de los modelos y la estacionalidad

para las medias móviles. Se busca estimar cuales son los coeficientes de la función a estimar, se pueden utilizar los métodos de máxima verosimilitud o mínimos cuadrados.

- d. Evalúe el modelo, a partir de los criterios de información (AIKE, AICC, BIC) y estime cual es el mejor modelo.

El paso a paso del proceso iterativo se describe mejor en el flujograma siguiente, ver grafica 3

Gráfica 3. Metodología de Box-Jenkins



Fuente: (Lozano F, 2017)

De otro lado, para hacer estimaciones usando este tipo de metodologías se deben cumplir con una serie de supuestos o ajustes hacia donde deben converger los datos, puntualmente los residuos o errores. De tal manera, que la predicción que se haga se ajuste a la realidad. Particularmente, se tiene que cumplir con el supuesto de que los residuos no presenten autocorrelación, que los errores presenten una distribución normal y que los errores sean independientes. Para cumplir con estos supuestos se tienen las pruebas de Box-Pierce, Portmantau, Ljung-Box, Breusch-Godfrey y Durbin-Watson.

La idea general detrás del ajuste de la verificación del modelo es corroborar que los datos ajustados se comporten lo más parecido posible a la serie original y que los residuos, es decir la diferencia entre la serie original y la ajustada sean lo más pequeño posibles, no se relacionen en ningún momento con las variables, y se comporten de manera aleatoria o como un ruido blanco (media cero y varianza σ^2 constante). Esto último es lo que buscan corroborar las pruebas que se mencionaron anteriormente, básicamente se orientan a determinar la no existencia de correlación serial mediante el rechazo de la hipótesis nula con un nivel de significancia de $100(1 - \alpha)\%$ cuando el estadístico de prueba (Q) sea mayor al valor crítico de chi-cuadrado con uno menos alfa rezagos como grados de libertad;

$$\left\{ \begin{array}{l} \text{Ho Rechazo si } Q = n(n + 2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n - k} > \chi_{1-\alpha}^2(h) \\ \text{H1, Acepto si } Q = n(n + 2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n - k} < \chi_{1-\alpha}^2(h) \end{array} \right.$$

Donde, $\hat{\rho}_k^2$ es la auto correlación muestral en el rezago k , y k es el numero de rezagos utilizados. De otro lado, la comprobación del cumplimiento del supuesto se puede hacer a través de la ayuda grafica que ofrece el software estadístico R con la función de `tsdiag` o la función `checkresiduals` que se encuentra en el paquete `forecast`. Básicamente, se observan los auto-correlogramas simples de los residuos, donde interesa que que ningún rezago sea significativo, es decir que ninguno sobrepase la línea punteada azul. El grafico QQplot que muestre una distribución normal de los errores, es decir que ningún punto se encuentre alejado de la línea transversal o por fuera de la región sombreada. La grafica de los errores, debe asemejar a un ruido blanco, es decir que se presente una varianza constante, los datos se muevan alrededor del cero y no se evidencie un patrón. Finalmente, los p-valores del Ljung-Box, que deben ser significativos, es decir todos deben estar por encima de 0,05, en el grafico todos los puntos deben mantenerse por encima de la línea punteada.

Capítulo 5 Análisis y Resultados

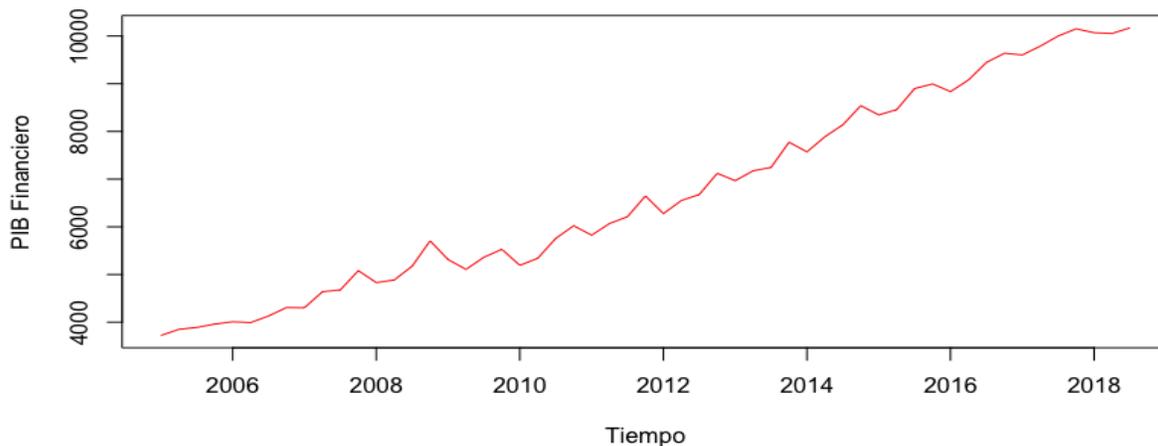
Una vez explicado el componente metodológico, se procederá a presentar los datos y las principales características descriptivas de los mismo.

Los datos que se utilizaran es la serie del Producto Interno Bruto PIB del Departamento Administrativo Nacional de Estadística DANE por sectores económicos, se tomara el componente de valor agregado del sector establecimientos financieros y de esta se extrae la serie de intermediación financiera, esta serie se encuentra a precios constantes de 2005, las cifras están en miles de millones de pesos, las observaciones se dan por una separación temporal de tres meses es decir por trimestres y cubre el periodo del primer semestre del año 2005 hasta el tercer trimestre del 2018.

5.1. Exploración de la serie

Para estimar el modelo, comprobar su rendimiento, y por conveniencia, se dividirán los datos en dos subconjuntos; una muestra para estimar y probar el modelo, y otra para evaluar su precisión. El primer conjunto de observaciones estará conformado por los datos que van del primer trimestre del 2005 al primer trimestre del 2015. El segundo grupo, que será para hacer la evaluación, tomará las observaciones desde el segundo trimestre de 2015 hasta el tercer trimestre de 2018. Como se explicó anteriormente la estimación del PIB financiero se hará con la serie de la Intermediación financiera.

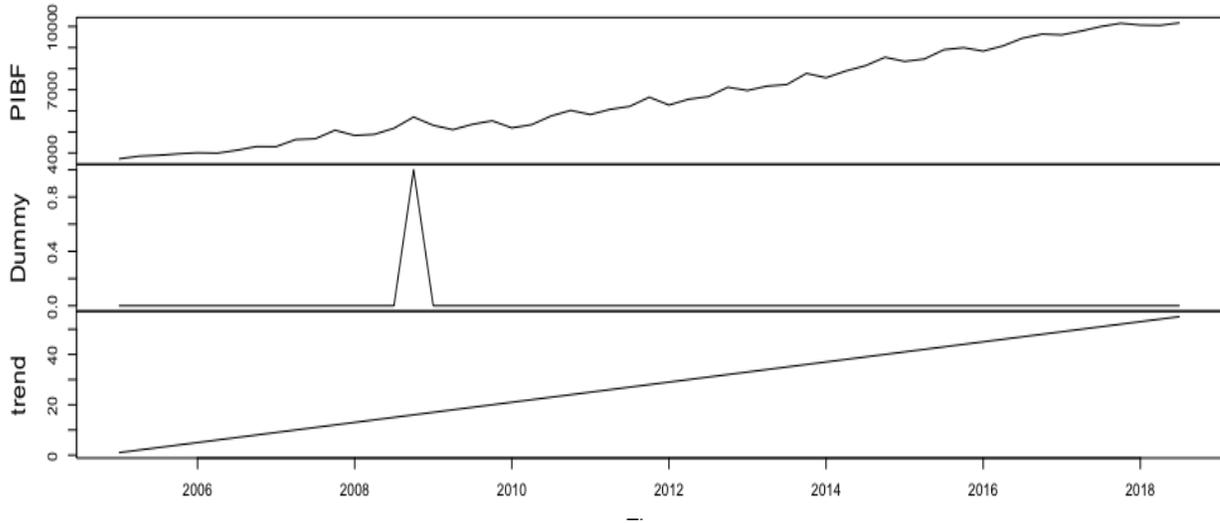
Gráfica 4 Serie Original



Como se puede observar la serie tienen una tendencia. Sin embargo, se procederá a comprobar por medio de las diferentes pruebas estadísticas, si hay presencia de raíz unitaria. Aunque, normalmente se aconseja que cuando se encuentra una tendencia se proceda a quitarla, es importante también tratar de entender por qué se da. Se aconseja no recurrir inmediatamente a la diferenciación ya que puede involuntariamente condicionar los datos para que den una forma

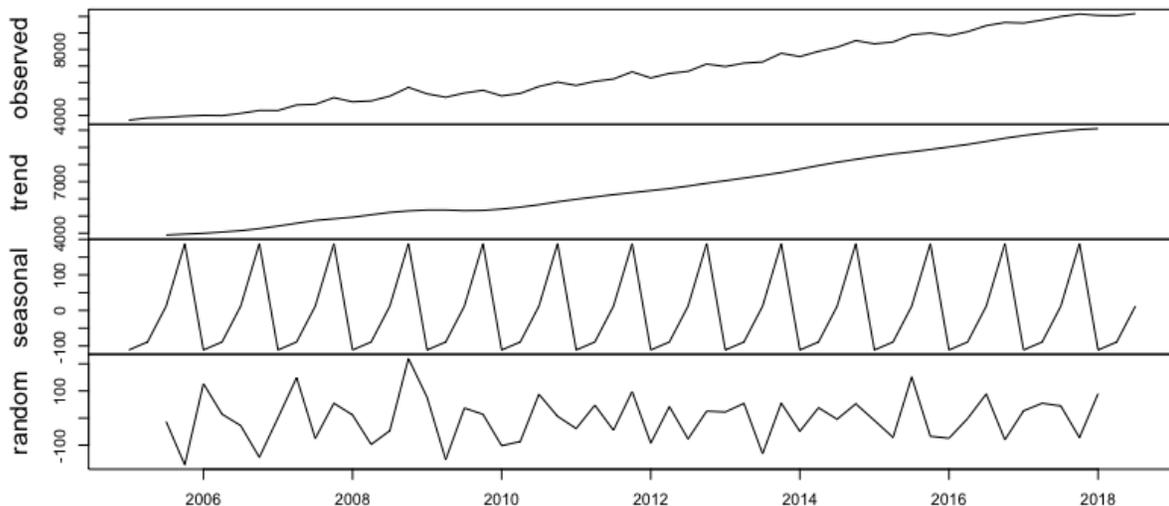
esperada que pueden llevar a resultados contrarios o sesgados. Es muy útil, por tanto, hacer una descomposición de la serie para observar la tendencia, el componente estacional, y los ciclos de esta.

Gráfica 5. Descomposición de la serie y dummy



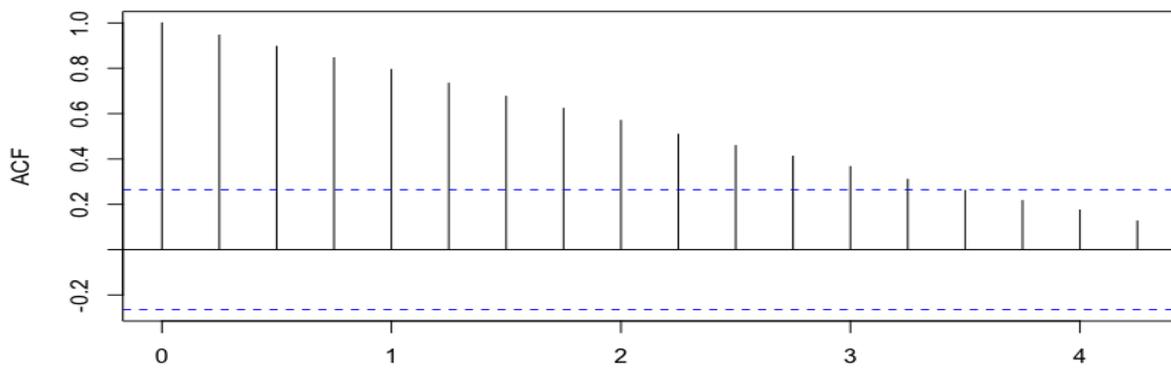
Es evidente la tendencia, sin embargo es llamativo que la serie presenta un comportamiento anormal en el cuarto trimestre del 2008, quizá la serie toma alguno de los efectos que tuvo la crisis financiera que ese año estallo en los Estados Unidos, ahí el diferencial entre lo que paga el banco y lo que gana por el préstamo (que es la intermediación) presenta un comportamiento que sale de la media de los datos, es por esto que se identifica en la serie, y se estima necesario el uso de una variable dummy que resalte este valor, que como único no debe afectar el comportamiento de la serie. Además, como se observará a continuación claramente se puede evidenciar un componente de tendencia y un componente estacional (periódico)

Gráfica 6 Descomposición de la serie



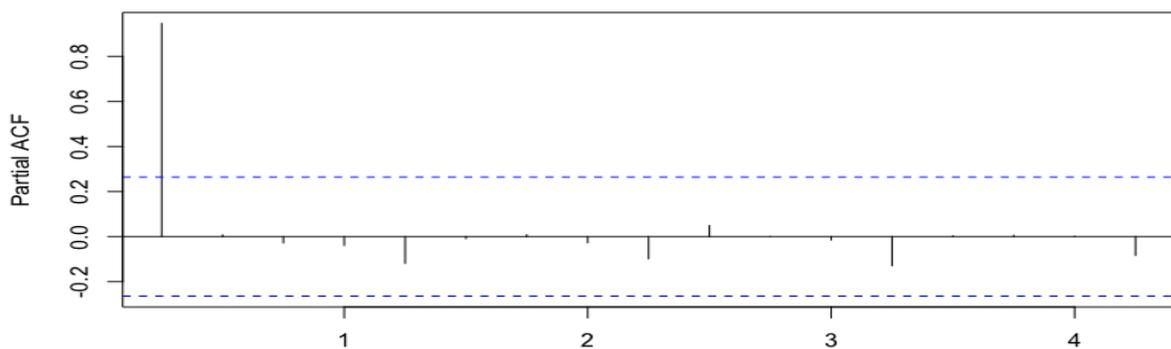
Una vez observadas las particularidades de la serie, se procederá a hacer diversas pruebas para estar seguro de cuáles son los comportamientos fuertes o señales que deben tenerse en cuenta en la formulación del modelo. Un primer acercamiento, hacia conocer el comportamiento de la serie, es revisar el patrón de configuración de las funciones de autocorrelación de la serie, buscando con ello identificar el orden y parámetros de regresión y de media móvil. Esto determinara el tipo de modelo a proponer. Queda claro que la serie no se comporta de manera “adecuada”, es decir no es estacionaria y por tanto un primer modelo a considerar es uno de tipo ARIMA, por la presencia de la tendencia y estacionalidad de la serie, para confirmar lo anterior posteriormente se hará una prueba de raíz unitaria que validará la no estacionariedad de la serie y la necesidad de diferenciar para poder continuar.

Gráfica 7 Autocorrelograma Simple



Previamente se explicó que era necesario identificar la autocorrelación que presentan los datos para tener una idea sobre el modelo a estimar. Como se puede apreciar, en la función de autocorrelación simple la configuración del autocorrelograma muestra una caída lenta, que desde el punto de vista del componente auto regresivo indica una alta dependencia del pasado. Desde el modelo de media móvil, la función de autocorrelograma simple indica el orden de rezagos que son necesarios estimar para tener un buen ajuste del modelo, en el componente de media móvil es necesario tener en cuenta por lo menos hasta el tercer rezago. Sin embargo, de acuerdo a los resultados que se obtengan de las pruebas de raíz unitaria y la necesidad de configurar el modelo este número de rezagos a usar es meramente informativo.

Gráfica 8 Autocorrelograma Parcial

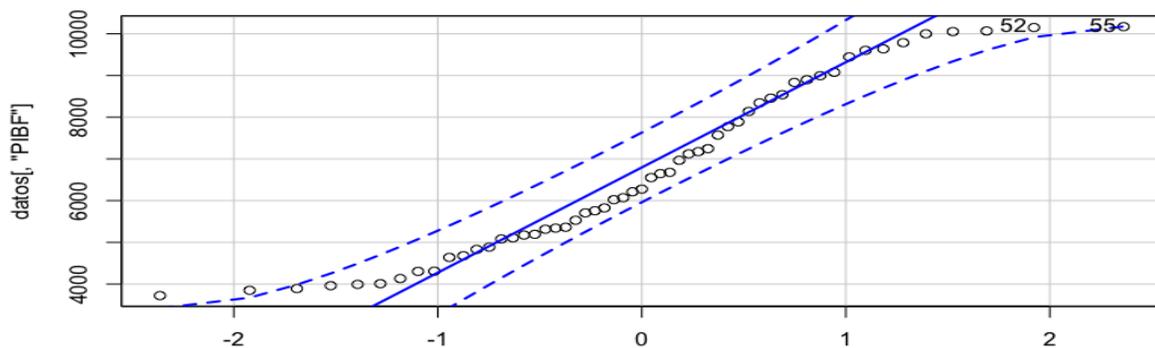


Como se pudo apreciar en el autocorrelograma parcial se presenta un rezago importante y con efecto significativo en el resultado de la serie para el componente autorregresivo. Por el lado del componente de media móvil se puede ver una baja persistencia de los errores del pasado en la situación actual de la serie, por esta razón no es claro el orden, como se vio en el autocorrelograma parcial, donde indica que por lo menos es de orden tres., esto confirma lo que se dijo anteriormente que el número de rezagos para el proceso de media móvil no es evidente ni tan preciso solo con la opción visual o gráfica.

Sin embargo hay que tener precaución con los resultados anteriores, pues “en un proceso y_t con raíz unitaria, por ejemplo, ARIMA(p, 1, q) las FAC muestral converge a cero con mucha lentitud y la FACP muestral presenta un valor cercano a 1 en el rezago $k = 1$ y el resto son aproximadamente cero” (Giraldo Gomez, 2006), ejemplo lo que se muestra en las imágenes anteriores (ver gráfico 7 y 8). El autor también aconseja, que antes de diferenciar es bueno hacer una prueba de hipótesis y descartar la presencia de raíces unitarias

Continuando con la revisión de los supuestos de estacionariedad de la serie, se procederá a ver su distribución, en el sentido de mantener o no su estabilidad. Se procederá a graficar la distribución de la serie en un gráfico qqplot, para poder observar la variabilidad o varianza de los datos a través del tiempo, la estabilidad de la media, algunos supuestos de normalidad de los datos y con ello confirmar las condiciones de estacionariedad, al menos desde el punto de vista grafico o visual.

Gráfica 9 Quantile-Quantile



Como se puede apreciar los datos no se distribuyen uniformemente, por tanto, se procederá a descartar la existencia de alguna raíz unitaria y determinar si la serie está bien compartimentada para poder hacer cualquier estimación, es decir se descartará que ésta presenta un patrón sistemático que es impredecible. Para lo anterior se procederá a hacer la prueba de **Dickey-Fuller Aumentado** y la prueba de **Phillips-Perron** del paquete series. En estas pruebas, la hipótesis nula es que la serie tiene raíces unitarias, por tanto, no es estacionaria. Por el contrario, la hipótesis alternativa es que la serie es estacionaria (alternativa= “stationary”) y k , corresponde a la orden de lags o retrasos. Acá el criterio de decisión es que, con cualquier valor, del p-value, por debajo de 0,05 se rechaza la hipótesis nula de no estacionariedad, es decir se acepta que la serie es estacionaria.

Tabla 1 Pruebas de raíz unitaria DF y PP para toda la serie.

Test	P-valor $H_0: y_t \sim I(1)$	Valor crítico $\alpha = 0,05$	Regla de rechazo Interpretación
Dickey-Fuller	0,7304	0,05	P-valor < α Rechazo H_0 p-valor > α Acepto H_0 La prueba falla en rechazar la hipótesis nula, por tanto, se asume que la serie no es estacionaria.
Phillips-Perron	0.08073	0,05	p-valor < α Rechazo H_0 p-valor > α Acepto H_0 La prueba falla en rechazar la hipótesis nula, por tanto, se asume que la serie no es estacionaria

En las dos pruebas no se puede rechazar la hipótesis nula, de esta forma se confirma la existencia de una raíz unitaria en la serie y hay que determinar en donde se encuentra.

Para continuar, y comprobar la presencia de raíz unitaria se procederá a hacer la prueba DF y DF Aumentada. Como se comentó con anterioridad se hizo una división de los datos en dos conjuntos. El primer conjunto, se usará para el aprendizaje, comprende los datos desde el primer semestre del 2005 al primer semestre del 2015 y corresponde al 80% de los datos. El segundo conjunto son los demás datos, es decir los que van del segundo semestre del 2015 al tercer trimestre del 2018 y son los datos que se usaran para hacer las pruebas de ajuste y pronóstico del modelo. Teniendo en cuenta la división las pruebas de raíz unitaria DF y DF aumentada se harán al primer grupo de datos.

Tabla 2 Pruebas de raíz unitaria DF y DF Aumentada para el conjunto de prueba

Modelo $y_t \sim AR(1)$	H_0	Valor Estadístico de prueba (τ)	Valores críticos $\hat{\tau}_{0,05}$			Regla de rechazo $\tau < \hat{\tau}_{0,05}$	
			1 pct	5 pct	10 pct		
“none” $\Delta y_t = \gamma y_{t-1} + \varepsilon_t$	$\gamma = 0$	2.4278	tau	-2.62	-1.95	-1.61	No se puede rechazar H_0 , se debe suponer la presencia de una caminata aleatoria, que confirma la presencia de raíz unitaria
“drift” $\Delta y_t = \alpha + \gamma y_{t-1} + \varepsilon_t$	$\gamma = \alpha = 0$	1.7834 3.4905	Tau Phil1	-3.58 7.06	-2.93 4.86	-2.60 3.94	No se puede rechazar H_0 , se debe suponer la presencia de una raíz unitaria. Sin embargo, al no poder rechazar la segunda prueba se reconoce que gamma o alfa es diferente de cero, por lo que se puede modelar con un drift (constante)
“trend” $\Delta y_t = \alpha + \beta_t + \gamma y_{t-1} + \varepsilon_t$	$\gamma = \alpha = \beta_t = 0$	-1.1206 2.6404 2.5089	Tau Phil1 Phil3	-4.15 7.02 9.31	-3.50 5.13 6.73	-3.18 4.31 5.61	No se puede rechazar H_0 , se debe suponer la presencia de una raíz unitaria. Sin embargo, al poder rechazar la segunda y tercera prueba (phil 1 y3) se reconoce que gamma, alfa o beta es diferente de cero, por lo que se puede modelar con una tendencia y un drift (constante)

De acuerdo a la tabla 2 se confirma la presencia de la raíz unitaria y se plantea necesario agregar en el modelo un parámetro de una constante y un parámetro de tendencia tal como lo explica Hyndman en el capítulo 8 que plantea los componentes de los modelos ARIMA (Hyndman & Athanasopoulos, 2018)

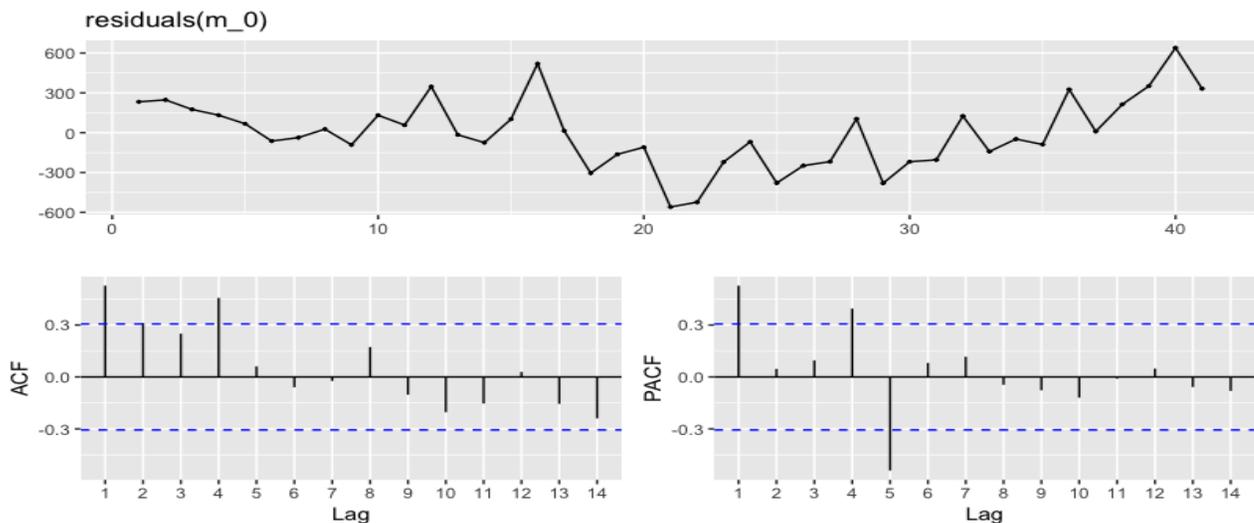
Como se tiene la sospecha de que la serie no es estacionaria por la tendencia se trata de confirmar a través de una prueba de Kwiatkowski et al, se prueba con una presencia de una tendencia lineal (τ). La mecánica de la prueba es que se parte de una **hipótesis nula de que la serie es estacionaria**, y se rechaza la hipótesis nula si el valor p asociado con el estadístico de prueba es inferior a un nivel de significación elegido, por ejemplo, el valor $p < 0,05$. De lo contrario, no se rechaza la hipótesis nula. En este caso **no se rechaza que la serie es estacionaria** y se confirma que hay estacionariedad en tendencia

Tabla 3 Prueba de Raíz Unitaria KPSS con tendencia

Test	Valor estadístico $H_0: y_t \sim I(0)$	Valor crítico $\alpha = 0,05$	Regla de rechazo Interpretación
Kwiatkowski et al	0,1430	0,146	P-valor $< \alpha$ Rechazo H_0 p-valor $> \alpha$ Acepto H_0 . Se rechaza H_0 por tanto se asume que la serie no es estacionaria en presencia de una tendencia lineal

Con los resultados de las pruebas no se rechaza la hipostasis nula, de manera tal que se acepta que no hay estacionariedad en la serie, por tanto, se debe buscar eliminar dicho efecto y buscar ajustar para que la serie se comporte de tal manera que permita generalizar los resultados. Para continuar se debe remover la tendencia y probar que con esto se soluciona el problema de estacionariedad para esto se prueba con una regresión de la serie contra su tendencia y se evalúan los residuos:

Gráfica 10 Residuos de muestra sin tendencia



Como puede apreciarse los residuos no se comportan como un ruido blanco, en los residuos se puede apreciar una persistencia con respecto al tiempo. Desde el punto visual se aprecia que el procedimiento no dio fruto, ahora con las pruebas se busca confirmar que las series sin la tendencia, es estacionaria, es decir que eliminando la tendencia se descarta la presencia de la raíz unitaria. Ya en párrafos anteriores se explicó como es el funcionamiento de las pruebas por lo que solo se tomaran los valores del estadístico y se confrontaran con el de los valores críticos: Primero se mostraran los datos de la prueba de raíz unitaria Dickey-Fuller

Tabla 4 Prueba Raíz unitaria DF y DF aumentada para muestra sin tendencia

Modelo $y_t \sim AR(1)$	H_0	Valor Estadístico de prueba (τ)	Valores críticos $\hat{\tau}_{0,05}$			Regla de rechazo $\tau < \hat{\tau}_{0,05}$	
			1 pct	5 pct	10 pct		
“none” $\Delta y_t = \gamma y_{t-1} + \varepsilon_t$	$\gamma = 0$	-1.4191	tau	-2.62	-1.95	-1.61	No se puede rechazar H_0 , se debe suponer la presencia de una caminata aleatoria, hay presencia de raíz unitaria
“drift” $\Delta y_t = \alpha + \gamma y_{t-1} + \varepsilon_t$	$\gamma = \alpha = 0$	-1.3159 0.9705	Tau Phil1	-3.58 7.06	-2.93 4.86	-2.60 3.94	No se puede rechazar H_0 , se debe suponer la presencia de una raíz unitaria. Sin embargo, al no poder rechazar la segunda prueba se reconoce que gamma o alfa es diferente de cero, por lo que se puede modelar con un drift (constante)
“trend” $\Delta y_t = \alpha + \beta_t + \gamma y_{t-1} + \varepsilon_t$	$\gamma = \alpha = \beta_t = 0$	-1.1206 1.749 2.5089	Tau Phil1 Phil3	-4.15 7.02 9.31	-3.50 5.13 6.73	-3.18 4.31 5.61	No se puede rechazar H_0 , se debe suponer la presencia de una raíz unitaria. Sin embargo, al poder rechazar la segunda y tercera prueba (phil 1 y3) se reconoce que gamma, alfa o beta es diferente de cero, por lo que se puede modelar con una tendencia y un drift (constante)

Como puede apreciarse en la tabla anterior NO se rechaza la hipótesis nula y se acepta que la serie es NO estacionaria, la eliminación de la tendencia no soluciona el problema de la serie. Para confirmar que eliminando la tendencia NO se logra estacionariedad se ensaya con la prueba de Kwiatkowski et al. En este caso **también, se rechaza que la serie es estacionaria en la tendencia** y se confirma que No hay estacionariedad, por tanto, se debe buscar ensayar con otras pruebas que permitan solucionar la presencia de la raíz unitaria.

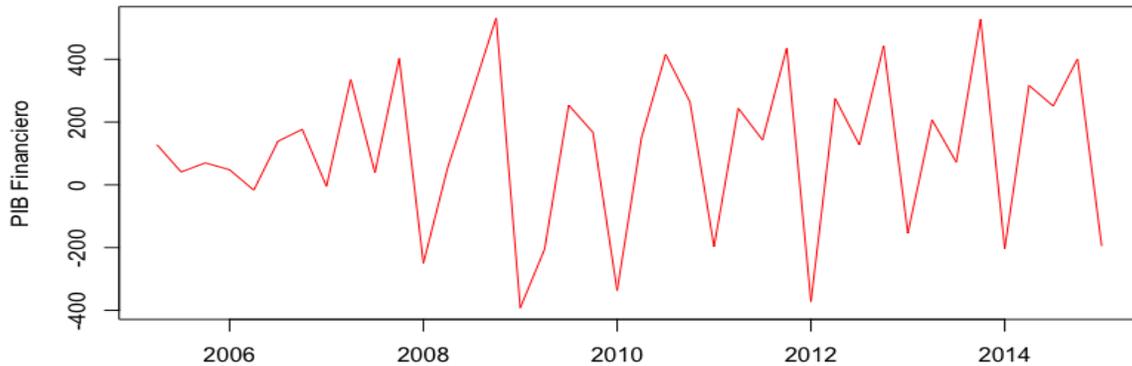
Tabla 5 Prueba de Raíz Unitaria KPSS sin tendencia

Test	Valor estadístico $H_0: y_t \sim I(0)$	Valor crítico $\alpha = 0,05$	Regla de rechazo Interpretación
Kwiatkowski et al	0,1403	0,463	P-valor < α Rechazo H_0 p-valor > α Acepto H_0 Se rechaza H_0 por tanto se asume que la serie no es estacionaria en presencia de una tendencia lineal

Conociendo el comportamiento de la serie frente a las pruebas de raíz unitaria realizadas, puntualmente la persistencia de la no estacionariedad, se pretende a través de la diferenciación

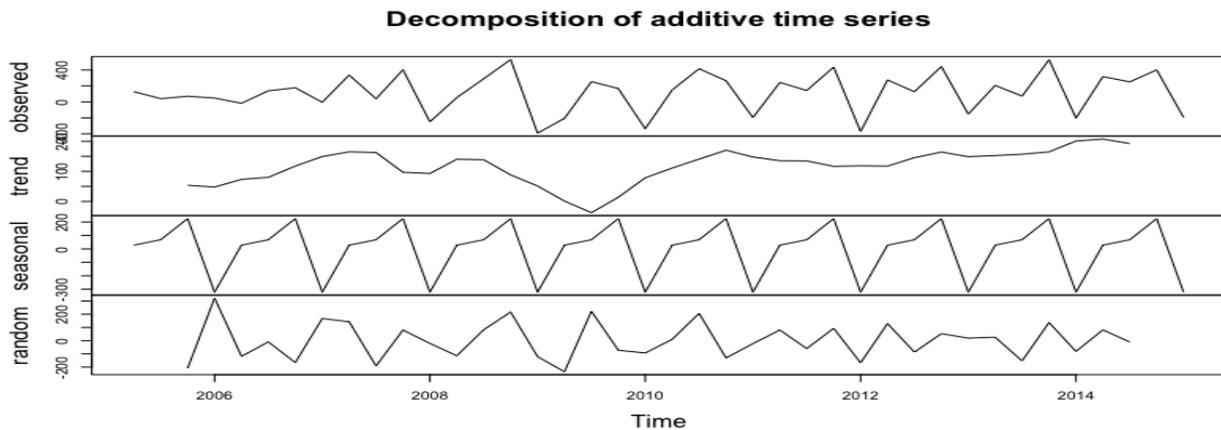
transformar la serie y ensayar un mejor ajuste que permita estimar un modelo más acorde con los datos originales.

Gráfica 11 Serie Diferenciada



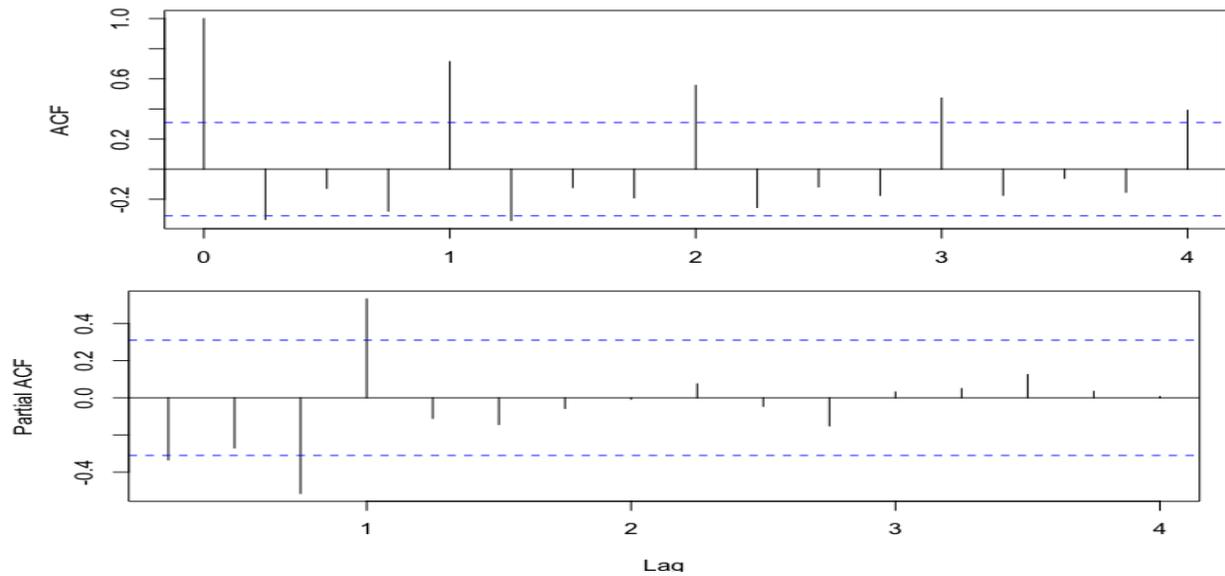
Ahora se observa cómo está compuesta esta “nueva serie”, para esto se presenta una descomposición de la serie y se construyen los diferentes correlogramas (parcial y simple) para determinar los efectos de los rezagos y de los errores. También, es prudente revisar el comportamiento de los residuos y determinar si cumplen los supuestos de normalidad, ya con esto se puede empezar a construir el modelo que mejor ajuste a la serie.

Gráfica 12. Descomposición Serie diferenciada



Como se puede apreciar, la diferenciación sirvió para quitar la tendencia, estabilizando la serie. Sin embargo, este procedimiento parece aumentar el efecto periódico de la serie, será necesario hacer una nueva prueba para determinar si la serie está integrada estacionalmente; para esto es útil la prueba de Hylleberg mejor conocido como la prueba de Hegy que además ayuda a verificar que frecuencia estacional es significativa para desarrollar el modelo. Así mismo, se procederá a hacer las pruebas de raíz unitaria para determinar la efectividad de la diferenciación. De otro lado, se procederá a revisar los autocorrelogramas, parcial y simple, para determinar los diferentes efectos y poder integrarlos en el modelo.

Gráfica 13. ACF y PACF Serie diferenciada



Al revisar las pruebas de hipótesis se presenta que para la prueba Dickey-Fuller aumentada existe una raíz unitaria (no se puede rechazar la hipótesis nula), pero para la prueba Phillips-Perrón la serie es estacionaria, esto puede sugerir la presencia del efecto de la periodicidad temporal, que es evidente en los gráficos anteriores donde la correlación parcial muestra el efecto de un rezago, esta información se tendrá en cuenta para estructurar el modelo.

Tabla 6 Pruebas de raíz unitaria DF y PP para la serie diferenciada.

Test	P-valor $H_0: y_t \sim I(1)$	Valor crítico $\alpha = 0,05$	Regla de rechazo Interpretación
Dickey-Fuller	0.7625	0,05	P-valor < α Rechazo H_0 p-valor > α Acepto H_0 La prueba falla en rechazar la hipótesis nula, por tanto, se asume que la serie no es estacionaria.
Phillips-Perron	0.01	0,05	p-valor < α Rechazo H_0 p-valor > α Acepto H_0 Se rechaza la hipótesis nula, por tanto, se asume que la serie es estacionaria

Como se evidencio en la descomposición de la serie, el hecho de haber diferenciado acentuó el efecto de la periodicidad en la serie, este hecho pone de manifiesto que es necesario considerar esta situación y proponer un modelo tipo SARIMA, es decir que tenga en cuenta la diferencia estacional.

Como se ha mencionado en párrafos anteriores el modelo que se busca estimar debe considerar las particularidades de la serie y tratar de emular sus características, de tal manera que se ajuste, y sirva para generalizar en un pronóstico dichos comportamientos. De esta manera el modelo que se pretende construir debe considerar la estacionariedad en tendencia, los efectos de la periodicidad temporal (estacionalidad), el efecto drift, los respectivos rezagos que se identificaron

y el efecto que se identificó en el 2008.

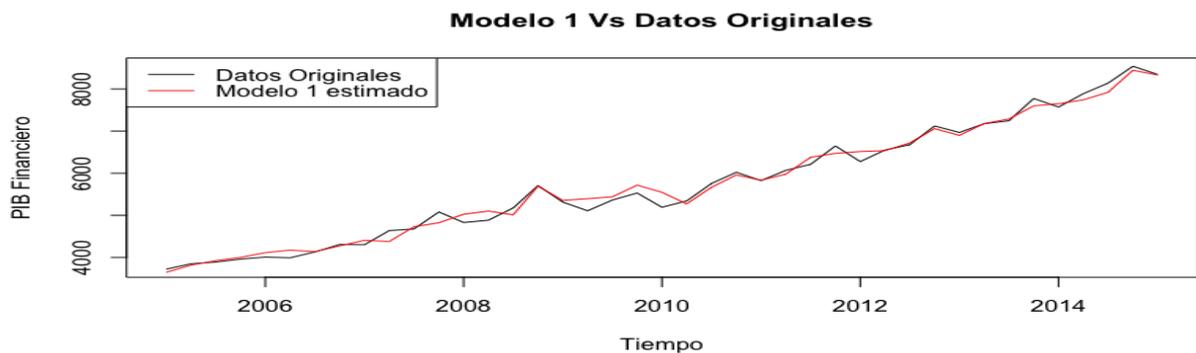
Al aplicar las pruebas de estacionariedad, a la serie transformada, es decir doblemente diferenciada, se logra descartar la no estacionariedad y se prepara para poder hacer predicciones, bueno primero ajustando el modelo que debe tener en cuenta los hallazgos mencionados.

5.2. Postulación de los Modelos

Teniendo en cuenta lo hallado en la exploración de los datos, es decir que se debe tener en cuenta una dummy que identifica el valor anormal o atípico del cuarto trimestre del 2008, que hay un componente estacional de tipo trimestral, que además es necesario trabajar con una constante y una tendencia, es imperativo que el modelo que se estime considere estas señales o particularidades. A continuación, se presentarán cada uno de los modelos construidos, en cada paso se hace una verificación muy general del cumplimiento de los supuestos de normalidad e independencia de los residuos, no se mostrarán en el documento, pero se tuvieron en cuenta en la construcción y formulación de cada uno de los modelos que se presentaran. Estos modelos son candidatos a mejor modelo, y a partir de las pruebas de criterios de información se seleccionará aquellos, que, bajo el principio de parsimonia, presenten la mejor descripción de los datos, con el menor número de parámetros.

Se estima un **Primer modelo** que considera ARIMA (1,0,0) X (1,0,0) [4], esta notación está basada en lo explicado por Hyndman (Hyndman R. , 2019) en su sitio web sobre las constantes en los Modelos ARIMA en R. De esta manera, se parte de un Modelo General: $ARIMA(1,1,1)X(1,1,1)[4] \rightarrow (1 - \phi_1 B)(1 - \Phi_1 B^4)(1 - B)(1 - B^4)y_t = (1 + \theta_1 B)(1 + \Theta_1 B^4)\varepsilon_t(1)$, que cancelando los términos no necesarios y agregando los componentes o parámetros que se identificaron por medio de las diferentes pruebas; tendencia, constante y la dummy (que identifica el dato atípico que se mostró en la descomposición de la serie,) y los diferentes procesos AR y de MA que se evidenciaron en los correlogramas, se postula un modelo ARIMA (1,0,0)X(1,0,0)[4], que en terminos formales sera; $(1 - \phi_1 B)(1 - \Phi_1 B^4)(y_t - \alpha - \beta_t) = \varepsilon_t$ (2). Ahora, reemplazando los parámetros en (2) tenemos; $(1 - 0.6968B)(1 - 0.7075B^4)(y_t - 3429.2321 - 117.4861 - 201.8461) = \varepsilon_t$ (3)

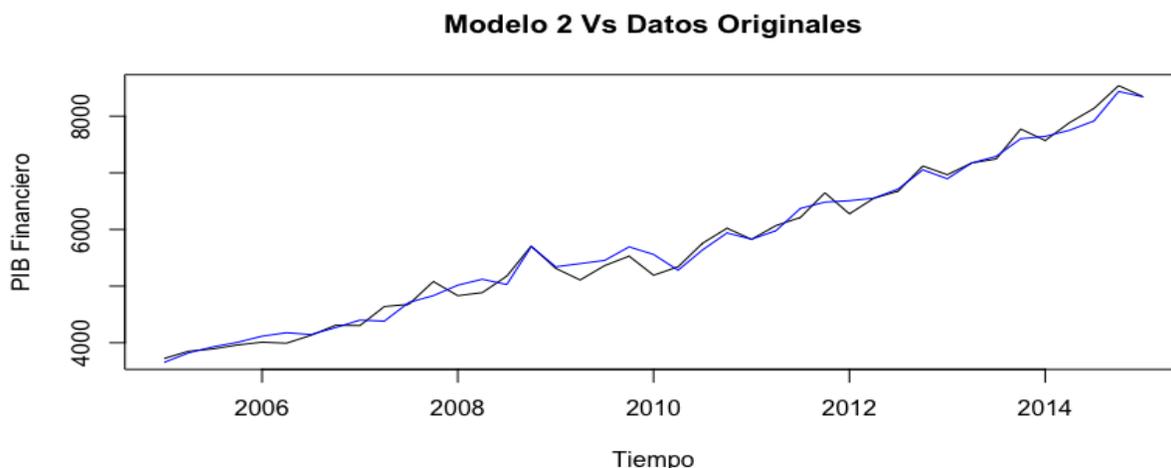
Gráfica 14. Modelo 1 Vs Datos Originales



Un Segundo modelo ARIMA (1,0,1) X (1,0,0) [4]

Con el mismo procedimiento descrito anteriormente y partiendo de (1), se postula un ARIMA (1,0,1)X(1,0,0)[4] $\rightarrow (1 - \phi_1 B)(1 - \Phi_1 B^4)(y_t - \alpha - \beta_t) = (1 + \theta_1 B)\varepsilon_t$ (4). Reemplazando los parámetros en (4) tenemos; $(1 - 0.7519B)(1 - 0.6974B^4)(y_t - 3438.3361 - 117.6089 - 225.2852) = (1 - 0.1088B)\varepsilon_t$ (5)

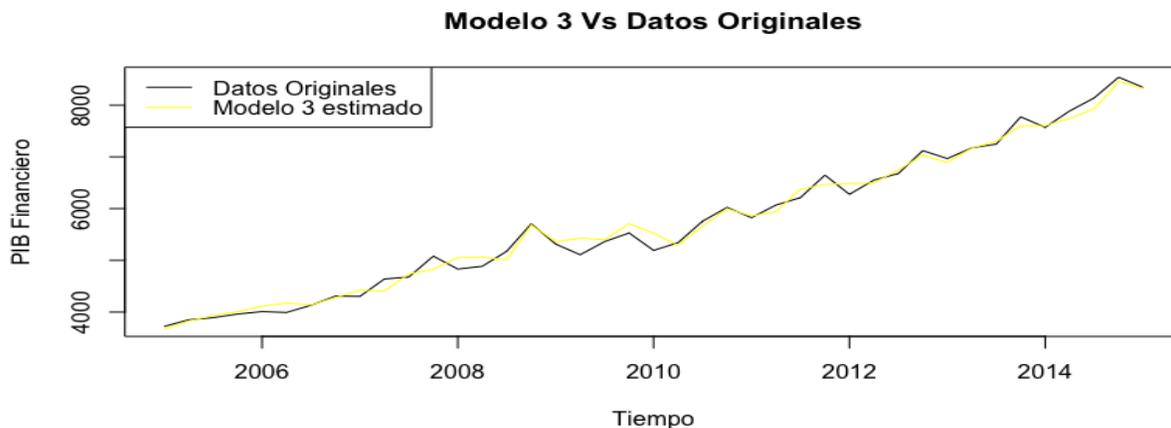
Gráfica 15. Modelo 2 Vs Datos Originales



Un Tercer modelo ARIMA (1,0,0) X (1,0,1) [4]

Con el mismo procedimiento descrito anteriormente y partiendo de (1), se postula un ARIMA (1,0,0)X(1,0,1)[4] $\rightarrow (1 - \phi_1 B)(1 - \Phi_1 B^4)(y_t - \alpha - \beta_t) = (1 + \Theta_1 B^4)\varepsilon_t$ (6). Reemplazando los parámetros en (6) tenemos; $(1 - 0.7381B)(1 - 0.8249B^4)(y_t - 3456.5637 - 117.5997 - 222.7976) = (1 - 0.2564B^4)\varepsilon_t$ (7)

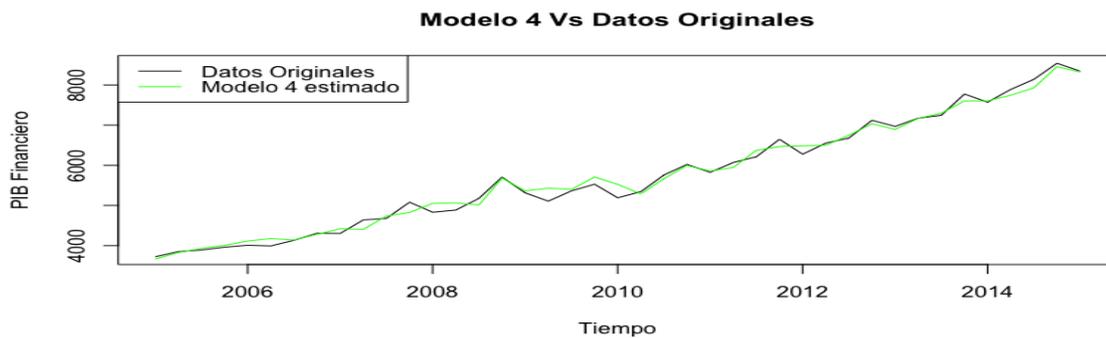
Gráfica 16. Modelo 3 Vs Datos Originales



Un Cuarto modelo ARIMA (1,0,1) X (1,0,1) [4]

Con el mismo procedimiento descrito anteriormente y partiendo de (1), se postula un ARIMA (1,0,0)X(1,0,1)[4] $\rightarrow (1 - \phi_1 B)(1 - \Phi_1 B^4)(y_t - \alpha - \beta_t) = (1 + \theta_1 B)(1 + \Theta_1 B^4)\varepsilon_t$ (8). Reemplazando los parámetros en (8) tenemos; $(1 - 0.8291B)(1 - 0.8423B^4)(y_t - 3491.5140 - 118.1007 - 268.0411) = (1 - 0.1800B)(1 - 0.3252B^4)\varepsilon_t$ (9)

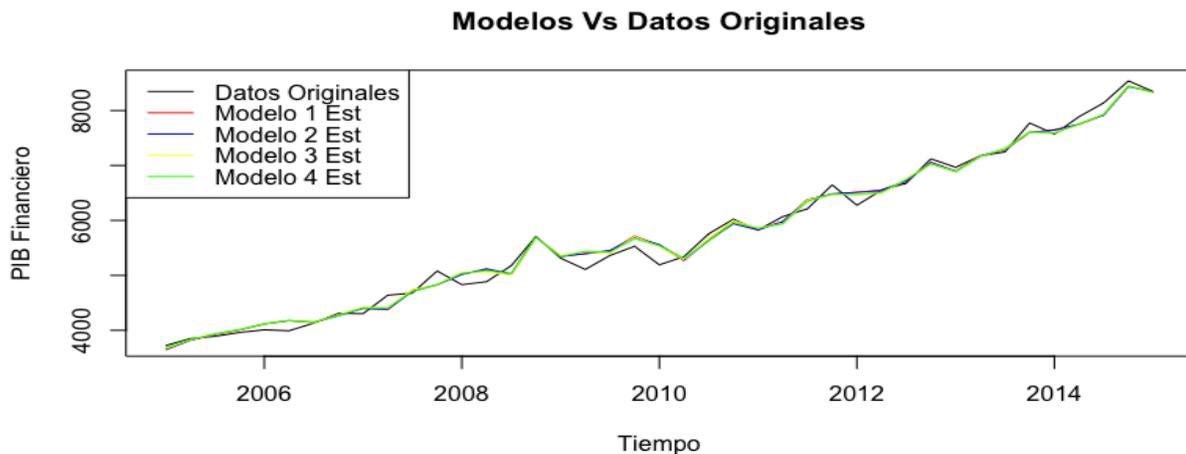
Gráfica 17. Modelo 4 Vs Datos Originales



5.3. Elección del Modelo y verificación de supuestos

Al super poner un modelo sobre el otro, podemos observar si gráficamente se ajustan a los datos originales. En la gráfica es evidente que quien mejor ajusta es el modelo 4, de una manera muy sutil. Sin embargo, no parece encontrarse diferencias significativas entre los modelos 1, 2 y 3, por esta razón y para una conclusión más contundente se decidirá cuál es el mejor a partir del uso de criterios de información, pues determinara a partir del principio de parsimonia cual es el que mejor describe la realidad de la manera más sencilla.

Gráfica 18. Modelos Vs Datos Originales



Al aplicar criterios de información se obtienen los resultados con los cuales pretende seleccionar el mejor modelo. En este caso, después de verificar los resultados, se procederá de manera más estricta a la revisión de los supuestos de los errores y así poder empezar a construir los pronósticos.

Tabla 7 Criterios de Información

Modelo	Especificación ARIMA	de	Prueba	
			AIC	BIC
M1	(1,0,0) X (1,0,0) [4]	6	538.8504	549.1318
M2	(1,0,1) X (1,0,0) [4]	7	540.5809	552.5759
M3	(1,0,0) X (1,0,1) [4]	7	539.8542	551.8492
M4	(1,0,1) X (1,0,1) [4]	8	541.0797	554.7883

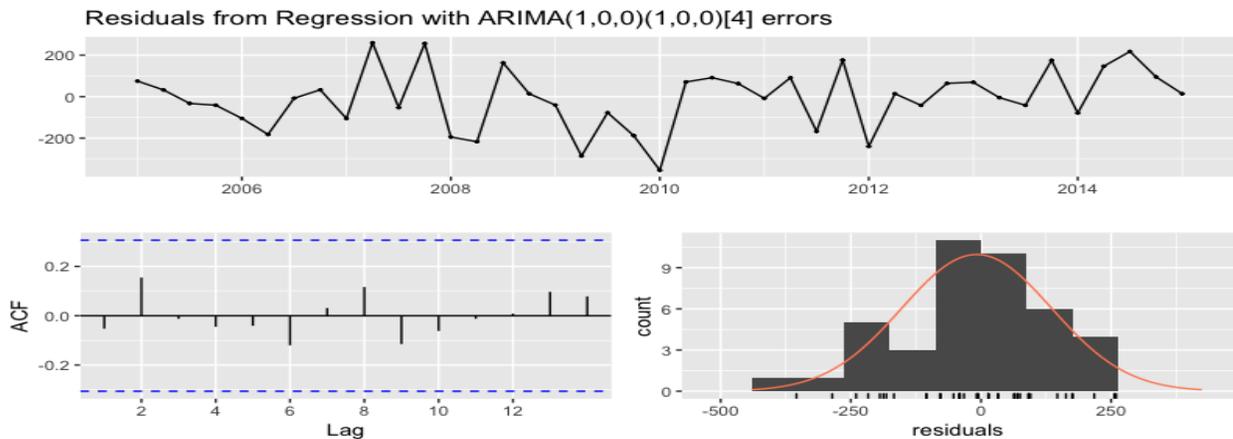
Como se puede apreciar, las diferencias entre un modelo y otro son sutiles. Sin embargo, el mejor modelo, según el menor número de criterios y parámetros, es el primero, es decir el ARIMA ((1, 0,0) X (1, 0,0) [4], bajo el principio de parsimonia es el que mejor explica la realidad de manera más sencilla. Ahora se procederá a verificar los supuestos de los residuos y la significancia estadística de los parámetros que arroja el modelo

Tabla 8 Prueba de Wald para los coeficientes

Parámetro	Valor estimado	Error Estándar	Z-Valor	Pr(> z)	Interpretación
ar1	0.69679	0.10770	6.4700	9.802e-11 ***	El valor del parámetro es significativo
sar1	0.70746	0.10259	6.8958	5.355e-12 ***	El valor del parámetro es significativo
intercept	3429.23213	305.060	11.2411	< 2.2e-16 ***	El valor del parámetro es significativo
drift	117.48605	10.932	10.7469	< 2.2e-16 ***	El valor del parámetro es significativo
d1	201.84606	96.990	2.0811	0.03742 *	El valor del parámetro es significativo
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Como se aprecia todos los parámetros son significativos. Ahora se desarrollará la prueba de Jarque Bera para determinar la normalidad de los residuos. X-squared = 0.63853, p-value = 0.7267, como se puede apreciar se acepta que los residuos se distribuyen de manera normal, por lo que se procederá a hacer la verificación gráficamente y luego hacer los pronósticos.

Gráfica 19 Comprobación de los errores del modelo

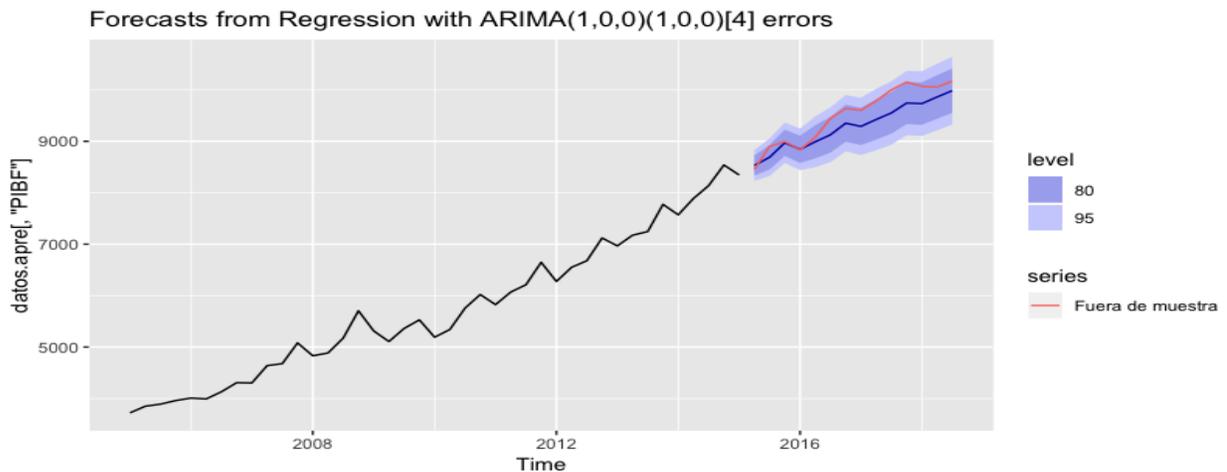


Como se puede apreciar los residuos cumplen los supuestos. Los residuos se comportan como un ruido blanco. En la gráfica de autocorrelogramas ninguno de los rezagos presenta significancia, la distribución de los residuos tiene una forma de campana que puede considerarse normal, por lo que se puede concluir que no presenta ninguna dependencia respecto a las variables analizadas. Ahora se procederá con los pronósticos. Como se explicó en un inicio se hicieron dos conjuntos de datos, unos de entrenamiento y los otros para probar los pronósticos, de esta manera con el primer grupo se hará un pronóstico que se comparara con el segundo conjunto y se evaluará el poder de predicción que tenga el modelo.

5.4. Pronóstico

Siguiendo con la metodología Box- Jenkins, una vez elegido el mejor modelo y comprobando el supuesto de comportamiento de los residuos, se procede a hacer los pronósticos, en este caso con el modelo ARIMA Ganador $ARIMA(1,0,0) \times (1,0,0)[4]$, $(1 - 0.6968B)(1 - 0.7075B^4)(y_t - 3429.2321 - 117.4861 - 201.8461) = \varepsilon_t$

Gráfica 20. Evaluación de la predicción



Como se puede apreciar el modelo captura de manera adecuada los impulsos de la serie original, y se encuentra dentro del margen de error de la predicción, por lo que puede decirse que representa de manera correcta lo que sucede al interior de la serie. Sin embargo, a medida que aumenta el tiempo del pronóstico este tiende a alejarse de los datos de la muestra de evaluación, por esta razón es necesario implementar una ayuda que permita mantener el pronóstico dentro de los márgenes para hacerlo más preciso. Un candidato fuerte para esta tarea es el suavizamiento exponencial, que busca minimizar la autocorrelación de los errores. A continuación, se presentan las medidas de precisión del modelo frente a los datos que se reservaron para entrenar el mismo

Tabla 9 Precisión del modelo en el conjunto de prueba.

	Error Promedio	Error Promedio cuadrático	Error Promedio absoluto	Error porcentual Promedio	Promedio Absoluto Porcentaje de error	Error medio absoluto escalado	Autocorrelación de errores en el rezago 1
Pronostico Modelo	-8.295	142.544	111.883	-0.298 6	2.053481	0.2288	-0.0522
Datos de evaluación	220.397	270.002	232.139	2.2472	2.385382	0.47490	0.55301

Como puede apreciarse los errores del pronóstico están dentro de lo normal y lo esperado. Para comprobar, gráficamente, el poder y precisión del pronóstico anterior se procederá a compararlo con el suavizamiento exponencial, luego aprovechando el poder de predicción de este se procederá a unirlos y ganar en la precisión del modelo, para ajustarlo con la información completa de la serie. Así mismo, se hará la verificación grafica del cumplimiento de los supuestos de los residuos y finalmente se hará el pronóstico por fuera de la muestra.

Gráfica 21 Comportamiento del Suavizamiento

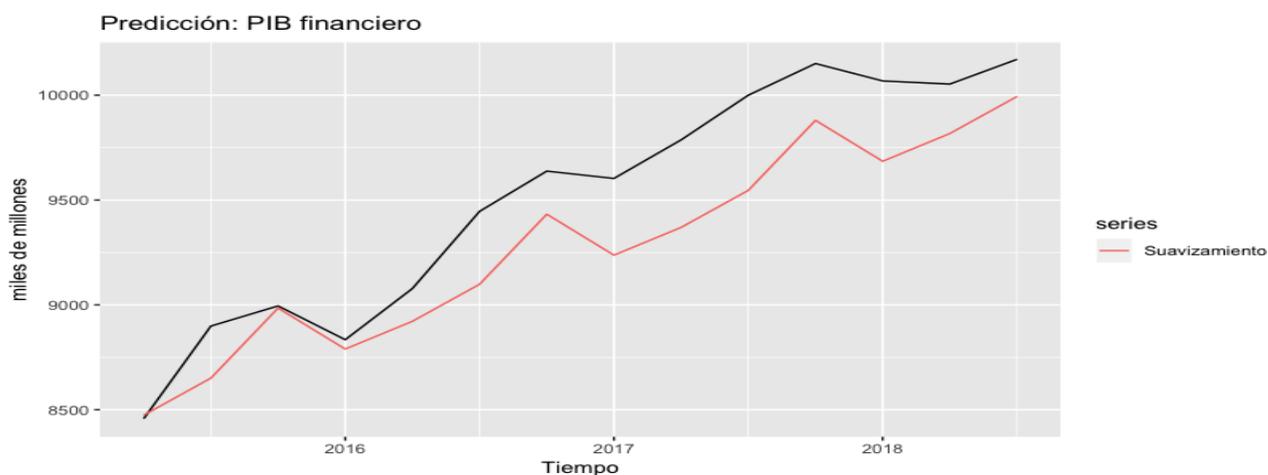
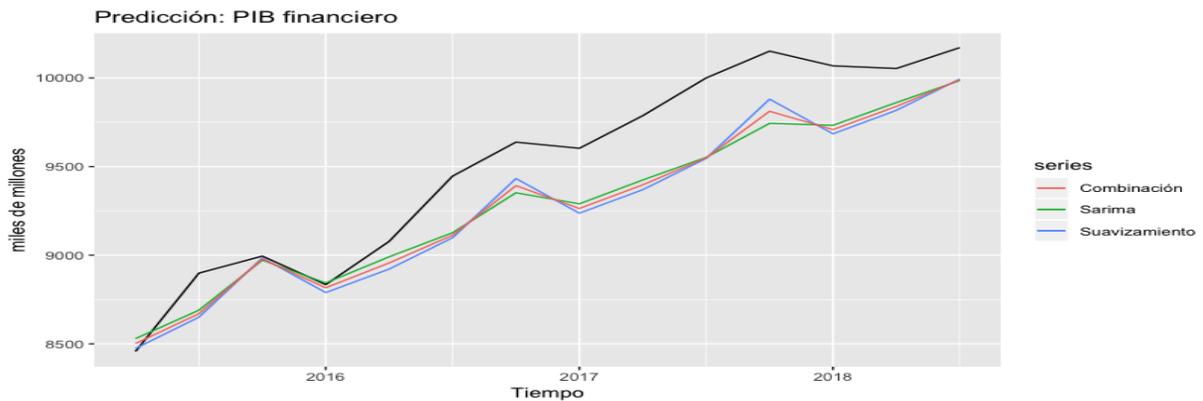


Tabla 10 Precisión del suavizamiento en el conjunto de prueba

	Error Promedio	Error Promedio cuadrático	Error Promedio absoluto	Error porcentual al Promedio	Promedio Absoluto Porcentaje de error	Error medio absoluto escalado	Autocorrelación de errores en el rezago 1
Pronostico suavizamiento	6.232	139.232	112.774	-0.0704	2.1368	0.2307	-0.00752
Datos de evaluación	235.656	2760.794	238.099	2.4222	2.34511	0.4871	0.3802

Como se observa gráficamente y con los resultados de los errores, el suavizamiento tiende a subestimar el pronóstico, ubicándose por debajo del valor esperado, y el “problema” del modelo es que sobre estima el valor estimado (ubicándolo por encima del valor real aumentando conforme aumenta el tiempo del pronóstico). Por lo tanto, su combinación permitirá acercar más los valores estimados al valor real, es decir disminuyendo el error se aumenta la precisión del modelo. De esta manera, se combinará el poder del algoritmo de suavizamiento con el modelo ARIMA estimado.

Gráfica 22 Combinación ARIMA con el Suavizamiento



Como puede apreciarse el poder del suavizamiento hace que el pronóstico sea más confiable, al estar más cerca y constante del valor real, permitiendo captar los cambios más importantes o ciclos de la serie. A continuación, se presentan los errores que arroja la combinación de las dos herramientas, en este caso ya no aparecen errores para la combinación, sino respecto a los datos de evaluación.

Tabla 11 Precisión de la Combinación del Modelo con el Suavizamiento conjunto de prueba

	Error Promedio	Error Promedio cuadrático	Error Promedio absoluto	Error porcentual Promedio	Promedio Absoluto Porcentaje de error	Autocorrelación de errores en el rezago 1
Datos de evaluación	228.027	271.833	234.437	2.3347	2.410	0.4942

A continuación, se presenta la serie y la combinación de las dos herramientas para asegurar que el comportamiento de esta combinación sea adecuado y capture en su totalidad los impulsos y señales de la serie, y que no solo corresponda al conjunto de prueba de los datos, que era lo que se mostraba con antelación

Gráfica 23 Combinación del ARIMA con el Suavizamiento para toda la serie

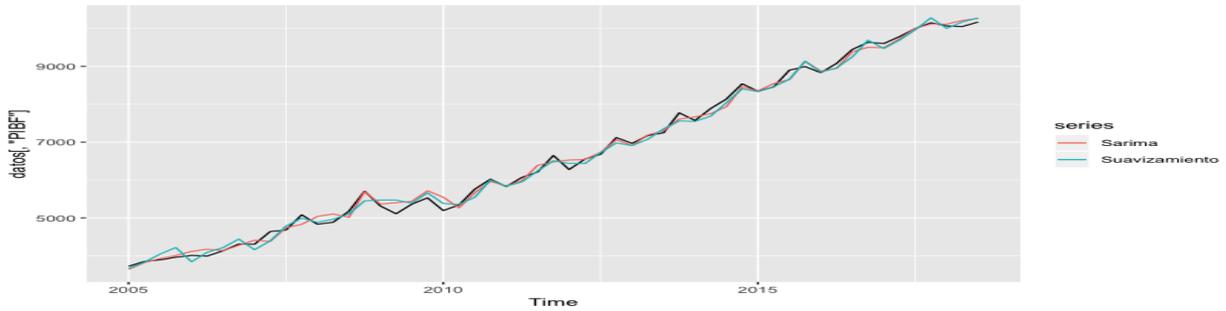
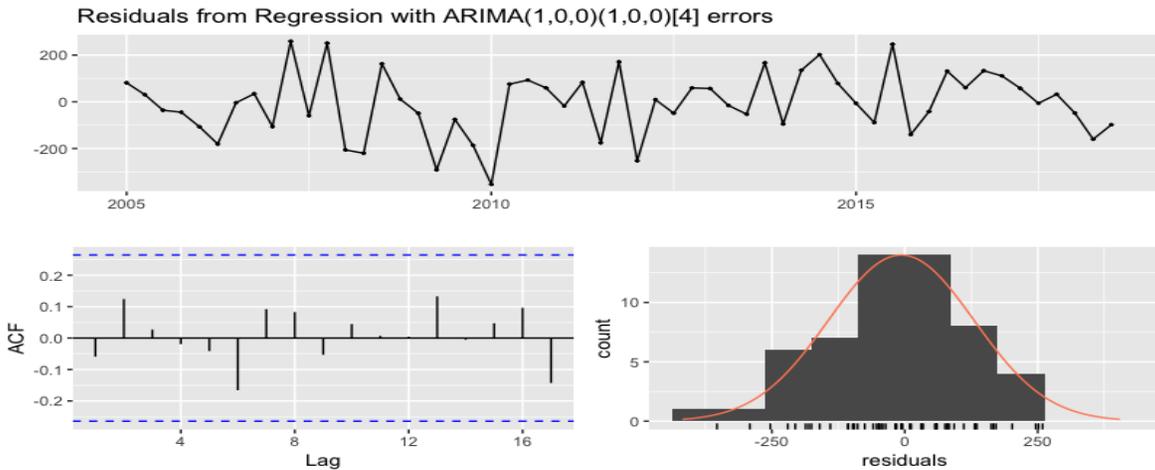


Tabla 12 Precisión de la Combinación del Modelo con el Suavizamiento para toda la serie

	Error Promedio	Error Promedio cuadrático	Error Promedio absoluto	Error porcentual Promedio	Promedio Absoluto Porcentaje de error	Error medio absoluto escalado	Autocorrelación de errores en el rezago 1
Datos de evaluación	-6.740	136.067	108.449	-0.2434	1.8079	0.2159	-0.0590

Finalmente, y una vez se ha ajustado el modelo con el poder de ambas herramientas, se procederá a hacer la verificación de los residuos para estar seguros de los pronósticos.

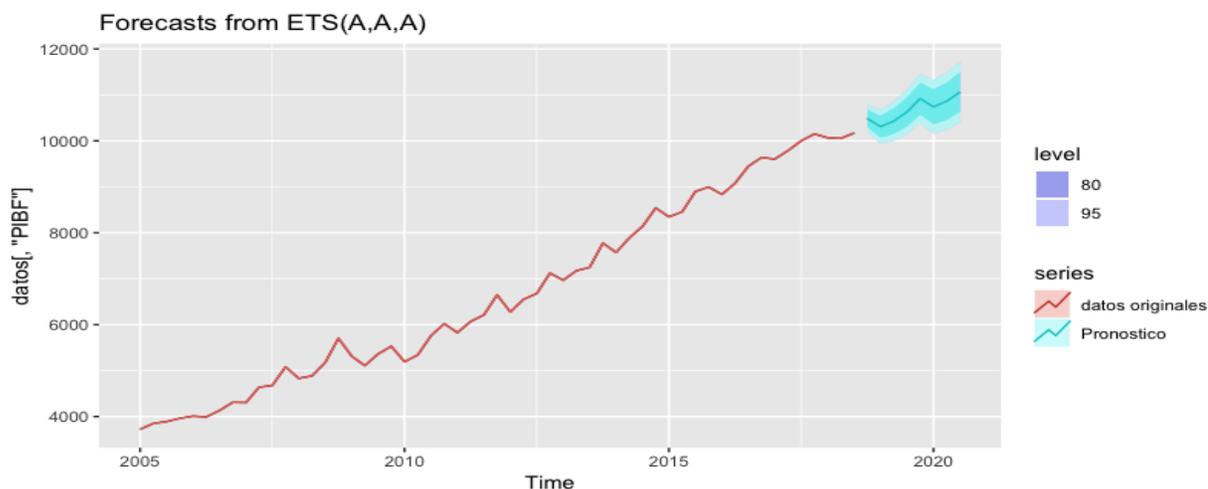
Gráfica 24. Verificación de residuos para la combinación



Como se puede apreciar los residuos cumplen los supuestos. Los residuos se comportan como

un ruido blanco. En la gráfica de autocorrelogramas ninguno de los rezagos presenta significancia, la distribución de los residuos tiene una forma de campana que puede considerarse normal, por lo que se puede concluir que no presenta ninguna dependencia respecto a las variables analizadas. Ahora se procederá con los pronósticos. Predicción de dos años por fuera de la muestra

Gráfica 25. Predicción de dos años



5.5. Análisis

Como puede observarse en la imagen, el modelo con la ayuda del suavizamiento exponencial captura las oscilaciones de la serie. De esta manera, se presenta un crecimiento para el cuarto trimestre del 2018 (primer periodo del pronóstico), comportamiento usual para el último periodo de cada año a través de la serie. Luego hay una caída para el primer trimestre del año 2019 (segundo periodo del pronóstico), también acorde con lo visto a lo largo de la serie. Y finalmente continua en crecimiento hasta el cuarto trimestre, volviendo a caer para el primer semestre del 2020.

Sin embargo, al combinar las herramientas los parámetros del modelo cambian, por tanto el modelo resultante será un $ARIMA(1,0,0) \times (1,0,0)$ [4], con intercepto o constante, tendencia y una variable dummy $(1 - 0.7192B)(1 - 0.7004B^4)(y_t - 3402.981 - 120.541 - 200.565) = \varepsilon_t$ con todos los parámetros significativos al 99%.

Tabla 13 Prueba de Wald para los coeficientes del modelo definitivo

Parámetro	Valor estimado	Error Estándar	Z-Valor	Pr(> z)	Interpretación
ari	0.7192	0.0906	7.9410	2.006e-15 ***	El valor del parámetro es significativo
sar1	0.7004	0.0865	8.0958	5.691e-16 ***	El valor del parámetro es significativo
intercept	3402.981	297.477	11.4395	< 2.2e-16 ***	El valor del parámetro es significativo
drift	120.541	8.302	14.5195	< 2.2e-16 ***	El valor del parámetro es significativo
d1	200.565	91.427	2.1937	0.028262 *	El valor del parámetro es significativo
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Cabe destacar que el propósito del trabajo era la estimación de un modelo que solo tuviera en cuenta la memoria de la serie, objetivo logrado. Sin embargo, se resalta que se deja a disposición de futuras investigaciones el establecer la magnitud de las relaciones potenciales de esta variable y demás indicadores de la actividad real de la economía; es decir cómo se traduce el comportamiento futuro de la variable intermediación con los crecimientos en inversión y que tanto impacta el ahorro en la inversión de capital, o de qué manera se hace competitivo el mercado financiero con el comportamiento de la variable estimada. Estos son futuras investigaciones donde el conocer con precisión el comportamiento de la variable intermediación financiera será altamente valorado, y donde se espera que el aporte del presente modelo sea considerado.

Capítulo 6 Conclusiones y Recomendaciones

- Se estimo un modelo de pronóstico ARIMA (1,0,0) X (1,0,0) [4] $(1 - 0.6968B)(1 - 0.7075B^4)(y_t - 3429.2321 - 117.4861 - 201.8461) = \varepsilon_t$, con intercepto o constante, tendencia y una variable dummy para el cuarto trimestre de 2008, este modelo se mejoró con una combinación del algoritmo de suavizamiento exponencial, dando como resultado un nuevo ARIMA con parámetros $(1 - 0.7192B)(1 - 0.7004B^4)(y_t - 3402.981 - 120.541 - 200.565) = \varepsilon_t$ con todos los parámetros significativos al 99%, que sirve para elaborar el pronóstico de la intermediación financiera en un horizonte de ocho periodos, o dos años. El modelo presenta un error promedio de -6.740, un error promedio cuadrático de 136.067, un error promedio absoluto de 108.449, un error porcentual promedio de 0.2434, un porcentaje promedio absoluto de error de 1.8079 y un error medio absoluto escalado de -0.2159.
- Para estimar el modelo, fue necesario conocer de donde provenían los datos, lo cual hizo necesario revisar los componentes del PIB, de donde se extrajo la serie de intermediación financiera. A partir de esta serie se estimaron cuatro modelos que tenían en cuenta la memoria de corto plazo de la serie a través de los procesos Autorregresivos y de Media Móvil. Sin embargo, como era evidente desde el punto de vista gráfico, era necesario establecer una estrategia para tratar el valor atípico que se presentó en el cuatro trimestre del 2008 y la tendencia que presentaba la serie, para esto se elaboraron una serie de pruebas de hipótesis que dieron como resultado la inclusión de la tendencia y constante en el modelamiento de los datos. De esta estrategia, se obtuvo un modelo que presentaba una sobreestimación de los errores, por tanto, se recurrió a los algoritmos de suavizamiento con lo que se logra mitigar el efecto de los errores. Una vez comprobado el cumplimiento de los supuestos de los residuos se procedió a hacer los pronósticos y verificar su precisión, tanto de manera gráfica como con la minimización de los errores.
- Finalmente, como recomendación, en el análisis de la serie es necesario considerar otras variables, para determinar la relación con las mismas y evaluar así la predicción en la explicación de las variables reales de la economía.
- ...

Bibliografía

- Callen, T. (2008). *Finanzas & Desarrollo*. Obtenido de <https://www.imf.org/external/pubs/ft/fandd/spa/2008/12/pdf/basics.pdf>
- Contreras Juárez, A., Zuniga, C., Martínez Flores, J., & Sánchez Partida, D. (2016). *Estudios Gerenciales*. Obtenido de Análisis de series de tiempo en el pronóstico de la demanda de almacenamiento de productos perecederos: https://ac.els-cdn.com/S0123592316300754/1-s2.0-S0123592316300754-main.pdf?_tid=7fe508d1-fce0-4ac9-8308-5a2d40a41fob&acdnat=1536950949_1e83e0213b74eaf1d010fecf73acabed
- Cowpertwait, P. S., & Metcalfe, A. V. (2009). *Introductory Time Series With R*. New York: Springer.
- DANE. (2010). *Metodología nueva base de las cuentas nacionales año 2005*. Bogotá: Dirección de Síntesis y Cuentas Nacionales.
- Enders, W. (2014). *Applied econometric time series*. University of Alabama: Wiley.
- Fernando G, T., & Garcia M, G. (1995). Intermediación financiera y crecimiento económico. *Cuaderno de Economía*, 14(23), 178-196.
- Giraldo Gomez, N. D. (2006). *Series de Tiempo con R. Notas de clase*. Medellín: Universidad Nacional de Colombia.
- Hernandez R, V. G. (2004). Modelos de Pronostico para el PIB de los Establecimientos Financieros, seguros, Inmuebles y servicios a las empresas. *Archivos de Economía*(274), 1-30.
- HERNANDEZ S. , C., PEDRAZA M., L., & ESCOBAR DIAZ , A. (2008). *Scientia et Technica. Universidad Tecnología de Pereira*. Obtenido de Aplicaciones de las series de tiempo en modelos de tráfico para una red de datos: <file:///C:/Users/olbernal/Downloads/Dialnet-AplicacionesDeLasSeriesDeTiempoEnModelosDeTraficoP-4748386.pdf>
- Hyndman, R. (6 de enero de 2019). *Constants and ARIMA models in R*. Obtenido de <https://robjhyndman.com/hyndsight/arimaconstants/>.
- Hyndman, R., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice*. Clayton, Australia: OTexts.
- Kuosmanen, P., & Vataja, J. (2014). Forecasting GDP growth with financial market data in Finland: Revisiting stylized facts in a small open economy during the financial crisis. *Elservier*, 23(2), 90-97.
- Levine, R., Loayza, N., & Beck, T. (2000). Financial intermediation and growth: Causality and causes. *Journal of Monetary Economics*(46), 31-77.
- Lozano F, S. (2017). *Series de Tiempo I y II*. Bogotá: Fundación Universitaria Los libertadores.

Mankiw, G. (2014). *Macroeconomía*. Barcelona: Antonio Bosh.

Superintendencia Financiera de Colombia. (2018). Obtenido de Actualidad del Sistema Financiero. Diciembre 2017:
<https://m.superfinanciera.gov.co/descargas/.../comsectorfinanciero122017.pdf>

Superintendencia Financiera de Colombia. (s.f.). *Aprenda con la SUPER*. Obtenido de Conformación del Sistema Financiero Colombiano:
<https://www.superfinanciera.gov.co/jsp/loader.jsf?lServicio=Publicaciones&lTipo=publicaciones&lFuncion=loadContenidoPublicacion&id=11268&dPrint=1>

Universidad de Sonora Mexico. (s.f.). *Proyecto de seguimiento a los cursos de estadística*. Obtenido de Estadística:
<http://www.estadistica.mat.uson.mx/Material/seriesdetiempo.pdf>

Villavicencio, J. (2012). *Introducción a Series de Tiempo*. San Juan: Estadísticas de Gobierno.

Wicherm, D., & Hanke, J. (2006). *Pronósticos en los negocios*. Juarez: Pearson.

Apéndice A Primer Apéndice