

---

PRONÓSTICO DE MORA PAR 30 DE CONTACTAR  
BAJO LA METODOLOGIA BOX JENKINS  
UTILIZANDO UNA BASE DE DATOS CON PERIODO  
2010 - 2019

---

Presentado por:

INGRID VANESSA MENESES BUSTOS  
Economista



LOS LIBERTADORES  
FUNDACIÓN UNIVERSITARIA

Fundación Universitaria Los Libertadores  
Facultad de Ingeniería y Ciencias Básicas  
Especialización en Estadística Aplicada

Bogotá D.C, Colombia

2019

---

**PRONÓSTICO DE MORA PAR 30 DE  
CONTACTAR BAJO LA METODOLOGIA BOX  
JENKINS UTILIZANDO UNA BASE DE DATOS  
CON PERIODO 2010 - 2019**

---

Presentado por:

**INGRID VANESSA MENESES BUSTOS**  
Economista

En cumplimiento parcial de los requerimientos para optar al  
título de

**Especialista en Estadística  
Aplicada**

*Asesores:*

Sebastian Lozano

Adriana Gallego

Fundación Universitaria Los Libertadores

Facultad de Ingeniería y Ciencias Básicas

Especialización en Estadística Aplicada

Bogotá D.C, Colombia

2019



---

## Agradecimientos

---

En primer lugar, agradezco a Dios por permitirme cumplir el anhelo de cursar esta especialización y por darme fuerza y valentía para salir siempre adelante, agradezco infinitamente a mi esposo Gabriel por el amor con que me apoyo en el transcurso de este proceso, por su comprensión, motivación y compañía.

También agradezco a mis profesores por compartir su conocimiento con tanto esmero y Finalmente a la universidad que bajo esta modalidad permite que mas personas podamos acceder a educación superior de calidad.

# Índice general

---

CAPÍTULO 1. Introducción.....	9
CAPÍTULO 2 Planteamiento del Problema .....	11
2.1 Descripción del problema. ....	11
2.1.1 Formulación del problema.....	13
2.1.2 Sistematización del problema. ....	13
2.2 Objetivos .....	13
2.2.1 Objetivo general.....	13
2.2.2 Objetivos específicos .....	13
2.3 Justificación.....	14
CAPÍTULO 3 Marco Teorico - Conceptual .....	16
3.1 Series de Tiempo .....	16
3.1.1 Metodología Box-Jenkins.....	16
3.1.2 Modelos AR .....	18
3.1.3 Modelos MA.....	18
3.1.6 Criterios ACF y PACF .....	19
3.1.7 Criterio de información Bayesiana (BIC) .....	21
3.1.8 Criterio de información Akaike (AIC) .....	21
3.2 Gestión de Riesgos .....	21
3.2.1 Política de administración de riesgos. ....	21
3.2.2 Objetivos de la gestión administración de riesgos.....	22
CAPÍTULO 4. Marco Metodológico .....	23
CAPÍTULO 5. Resultados .....	29
5.1 Metodología Box Jenkins.....	30
5.1.1 Identificación del modelo más adecuado .....	30
5.1.2 Estimación de los parámetros del modelo.....	32
5.1.3 Calculo de ACF y PACF.....	33
5.1.4 Evaluación y diagnostico del modelo .....	34
5.1.5 Predicción.....	34
CAPÍTULO 6. Conclusion y Recomendaciones.....	37
Referencias y Netgrafia .....	39

---

## Índice de figuras

---

	Pág.
Metodología Box Jenkins .....	17
Mora par 20.....	30
Prueba raíz unitaria.....	32
Diagnóstico del modelo SARIMA .....	33
Predicción R studio hasta agosto 2019.....	35
Predicción R studio hasta febrero 2020 .....	36

---

## Índice de cuadros

---

	Pág.
Modelo sarima (1,1,1) x (1,0,0) <sub>12</sub> .....	30
Modelo sarima (1,1,0) x (1,0,0) <sub>12</sub> .....	31
Modelo sarima (1,1,0) x (1,0,1) <sub>12</sub> .....	31
Modelo sarima (1,1,1) x(2,1,0) <sub>12</sub> .....	31
Predicción con niveles de confianza. ....	34
Comparativo predicción áreas contactar vs realizada vs real .....	35
Predicción con niveles de confianza .....	36

---

## Resumen

---

Las instituciones financieras en general están expuestas a diferentes tipos de riesgos, entre ellos el riesgo de crédito, de liquidez, el riesgo operativo, el riesgo reputacional, entre otros, sin embargo, en las entidades microfinancieras se puede incrementar el porcentaje de riesgo de estos, debido a la facilidad con la cual se otorgan créditos, por lo cual la incertidumbre por el incumplimiento de pago de los agentes económicos es más alta. De este modo la presente investigación busca aportar a través de la metodología Box Jenkins un modelo que permita pronosticar la mora par 30 de Contactar que servirá como alerta temprana y será de gran utilidad a la hora de tomar decisiones, como cambios de política, estrategias comerciales y financieras.

# CAPÍTULO 1

---

## Introducción

---

En las entidades microfinancieras se puede incrementar el porcentaje de riesgo debido a la facilidad con la cual se otorgan créditos, algunos de los principales riesgos son la informalidad del negocio, garantía del negocio, competencia del mercado, capacidad de pago, destino real del crédito, falta de experiencia en la actividad económica, voluntad de pago, por lo cual la incertidumbre por el incumplimiento de pago de los agentes económicos es más alta.

Debido a lo anterior las microfinancieras desarrollan mecanismos y estrategias que les permita hacer una buena colocación de los créditos, por lo cual hacen un análisis más exhaustivo sobre las características de los clientes.

De este modo la presente investigación busca aportar a través de la metodología Box Jenkins un modelo que permita pronosticar la mora par 30 de Contactar con lo cual se pueden generar estrategias que sirvan como alerta temprana frente al comportamiento de la mora, el cálculo de deterioro de valor, la activación de campañas que equilibren la mora, entre otras que impactará directamente sobre la

cartera de contactar, por otro lado, es importante tener en cuenta que el modelo desarrollado no espera ser el más sofisticado sino el más adecuado a la institución, por lo cual se busca verificar que no siempre los modelos más complejos son los mejores predictores en la realidad.

En el segundo capítulo se plantea el problema del trabajo y la justificación, en el tercer capítulo se establece el marco teórico y conceptual, se explican las principales definiciones, con el fin de que sean más claras las metodologías a utilizar, así como los tipos de modelos para aplicar y los criterios estadísticos que permitirán seleccionar el mejor modelo. En el cuarto capítulo se encuentra el marco metodológico, es relevante contar con investigaciones previas en el tema y que las recomendaciones permitan mejorar la calidad del presente trabajo.

En el capítulo 5 se mostrarán los modelos más significativos, así como también los resultados obtenidos con ayuda del software R studio, gracias a la disponibilidad de información, se tomará un gran número de observaciones, que serán factor clave para la robustez de los resultados, permitiendo generar un análisis a través de conclusiones y recomendaciones consignadas en el capítulo 6. Finalmente se presentan los referentes bibliográficos.

## CAPÍTULO 2

---

### Planteamiento del Problema

---

#### **2.1 Descripción del problema.**

A nivel internacional existen acuerdos de supervisión y regulación bancaria, emitidos por el Comité de Basilea, que son de gran relevancia para que las entidades financieras puedan prevenir distintos tipos de riesgos económicos a los que se enfrentan a diario.

El comité de Basilea en su primer acuerdo, conocido como “Basilea I”, establece los requerimientos mínimos para otorgar créditos con un capital regulatorio que no comprometa el nivel de liquidez de la entidad, ni represente riesgos de ningún tipo, entre ellos de crédito, mercado y tipo de interés. Este acuerdo ha sido de gran ayuda en el fortalecimiento de los sistemas bancarios, sin embargo el acuerdo tenía algunas limitaciones en sus definiciones por lo cual fue sustituido por el acuerdo “Basilea II”, en el cual se complementa y se hace recomendaciones sobre la legislación y regulación bancaria, su principal objetivo es la creación de un estándar internacional que pueda servir de modelo a los reguladores bancarios para que se dé mayor protección a las entidades frente a los riesgos operativos y financieros.

Como consecuencia de la crisis de las hipotecas Subprime, se genera un tercer acuerdo, conocido como “Basilea III”, en el cual se obliga a las entidades financieras para aumentar sus reservas de capital e incluso crear gradualmente colchones financieros que puedan ser útiles en tiempos de nuevas crisis, lo cual fortalecerá a las entidades frente a posibles riesgos sistémicos o resultados inesperados.

A nivel nacional la entidad encargada de regular y supervisar las entidades bancarias es la Superintendencia Financiera, mediante la circular básica contable y financiera en el capítulo II, establece los criterios generales para evaluar el riesgo de crédito, con lo cual las entidades se puedan fortalecer para generar practicas adecuadas que administren y aminoren el riesgo, para esto deberán desarrollar un sistema de administración de riesgo crediticio o SARC.

En el caso de la Corporación Nariño, Empresa y Futuro Contactar en calidad de ONG, no es vigilada por la Superintendencia Financiera, sin embargo, en pro de satisfacer uno de sus objetivos estratégicos, como es el de implementar un modelo de gestión por procesos y proyectos, garantizando el óptimo crecimiento y expansión de la organización con el retorno esperado, implementa el sistema de administración de riesgo crediticio SARC.

Como complemento a esta iniciativa, es de vital importancia que la entidad implemente nuevas estrategias que le permitan disminuir la incertidumbre frente a los clientes y mejorar la toma de decisiones sobre el otorgamiento de los créditos, así pues, con ayuda de estadísticas y creación de modelos econométricos se podrá predecir diferentes indicadores, entre ellos la mora que sirvan para orientar la toma de decisiones y aminoren el riesgo.

2.1.1 Formulación del problema. ¿A través de la metodología Box Jenkins se puede formular modelos adecuados para pronosticar la mora par 30 de Contactar y a la vez verificar que no siempre los modelos más complejos son los mejores predictores de la realidad?

2.1.2 Sistematización del problema. ¿Por qué a través de la metodología Box Jenkins se puede formular el mejor modelo para pronosticar la mora par 30 de Contactar?

- ¿Cuáles supuestos estadísticos permitirán determinar el mejor modelo de pronóstico de serie de tiempo para la mora par 30 de Contactar?
- ¿Por qué no siempre los modelos más complejos son los mejores predictores en la realidad?
- ¿Cómo se espera contribuir a la estrategia de colocación de Contactar con la generación del pronóstico de mora par 30?

## **2.2 Objetivos**

### 2.2.1 Objetivo general

Formular distintos modelos utilizando la metodología Box Jenkins para generar el mejor pronóstico de mora par 30 de Contactar y verificar que no siempre los modelos más complejos son los mejores predictores y se ajustan a la realidad.

### 2.2.2 Objetivos específicos

- Formular el mejor modelo de pronóstico de serie de tiempo para la mora par 30 de Contactar.
- Verificar que no siempre los modelos más complejos son los mejores predictores en la realidad.

- 
- Generar una propuesta que sirva de estrategia en la colocación de Contactar a través del pronóstico generado de mora par 30.

### 2.3 Justificación

Teniendo en cuenta el papel que cumplen las entidades financieras en el ámbito económico, se empieza a observar un fenómeno de gran impacto y es el protagonismo que han ido adquiriendo las entidades de intermediación financiera (EIF)<sup>1</sup> y particularmente las microfinancieras en el mercado, quienes les conceden a pequeños empresarios créditos con menos requisitos y con mayores facilidades, lo cual implica mayores riesgos para la entidad, de ahí que se debe hacer un mayor seguimiento, análisis y una evaluación permanente con el fin de identificar correctamente los diferentes factores que influyen en el cliente para que el pago del crédito sea oportuno y el solicitante cumpla a cabalidad con sus obligaciones financieras, para esto, las microfinancieras diseñan diferentes políticas que les permite aminorar el riesgo, sin embargo existe la necesidad de innovarse y estar generando continuamente herramientas que permitan mejorar la toma de decisiones por parte de la organización.

---

<sup>1</sup> EIF. Entidad de intermediación financiera. En inglés finance intermediation. Es la intervención que realizan las instituciones nacionales de crédito, organismos auxiliares, instituciones nacionales de seguros y fianzas y demás entidades autorizadas legalmente para constituirse como medio de enlace entre el acreditante de un financiamiento y el acreditado, obteniendo una comisión por su labor al concertar los créditos en los mercados de Dinero nacional e internacional. **Diccionario Economía**. Disponible en <[http://www.eco-finanzas.com/diccionario/I/INTERMEDIACION\\_FINANCIERA.htm](http://www.eco-finanzas.com/diccionario/I/INTERMEDIACION_FINANCIERA.htm)>. Con acceso el 07 de Septiembre de 2019.

En Colombia, la superintendencia Financiera de Colombia (SFC) mediante la Circular básica contable y financiera (Circular externa 100 de 1995) en su capítulo II, exige que todas las entidades financieras deben implementar un Sistema de Administración del Riesgo Crediticio (SARC)<sup>2</sup>, en este sentido y teniendo en cuenta la volatilidad de las distintas variables financieras, la Corporación Nariño, Empresa y Futuro Contactar se anticipa ha dicho requerimiento, con la intención de estar atento y tener un sistema financiero sólido.

Es por eso por lo que el estudio propuesto, pretende generar valor agregado a la institución, para lo cual se desarrollara distintos modelos predictivos de la mora par 30 de Contactar con el fin de orientar las decisiones en muchas áreas y a partir de ellas generar estrategias que permitan aminorar el riesgo de liquidez de la Institución, además de verificar que no siempre los modelos más complejos sirven para predecir mejor los indicadores de análisis financiero como la mora.

---

<sup>2</sup> SUPERINTENDENCIA FINANCIERA DE COLOMBIA. Circular básica contable y financiera (Circular externa 100 de 1995, Cap. II. Disponible en <<https://www.superfinanciera.gov.co/jsp/loader.jsf?IServicio=Publicaciones&ITipo=publicaciones&IFuncion=loadContenidoPublicacion&id=15466>>. Con acceso el 06 de Septiembre de 2019.

## CAPÍTULO 3

---

### Marco Teórico / Conceptual

---

#### **3.1 Series de Tiempo**

En esta capítulo, se presentan de forma muy general los modelos estadísticos tradicionales en series de tiempo univariados. En general, se presenta la teoría general.

##### **3.1.1 Metodología Box-Jenkins**

En el análisis de Series de Tiempo, la metodología de Box-Jenkins, nombrada así en honor a los estadísticos George E. P. Box y Gwilym Jenkins, se aplica a los modelos autorregresivos de media móvil (ARMA) o a los modelos autorregresivos integrados de media móvil (ARIMA) para encontrar el mejor ajuste de una serie temporal de valores, a fin de que los pronósticos sean más acertados.

En general la mejor forma de explicar dicha metodología es con el siguiente resumen:

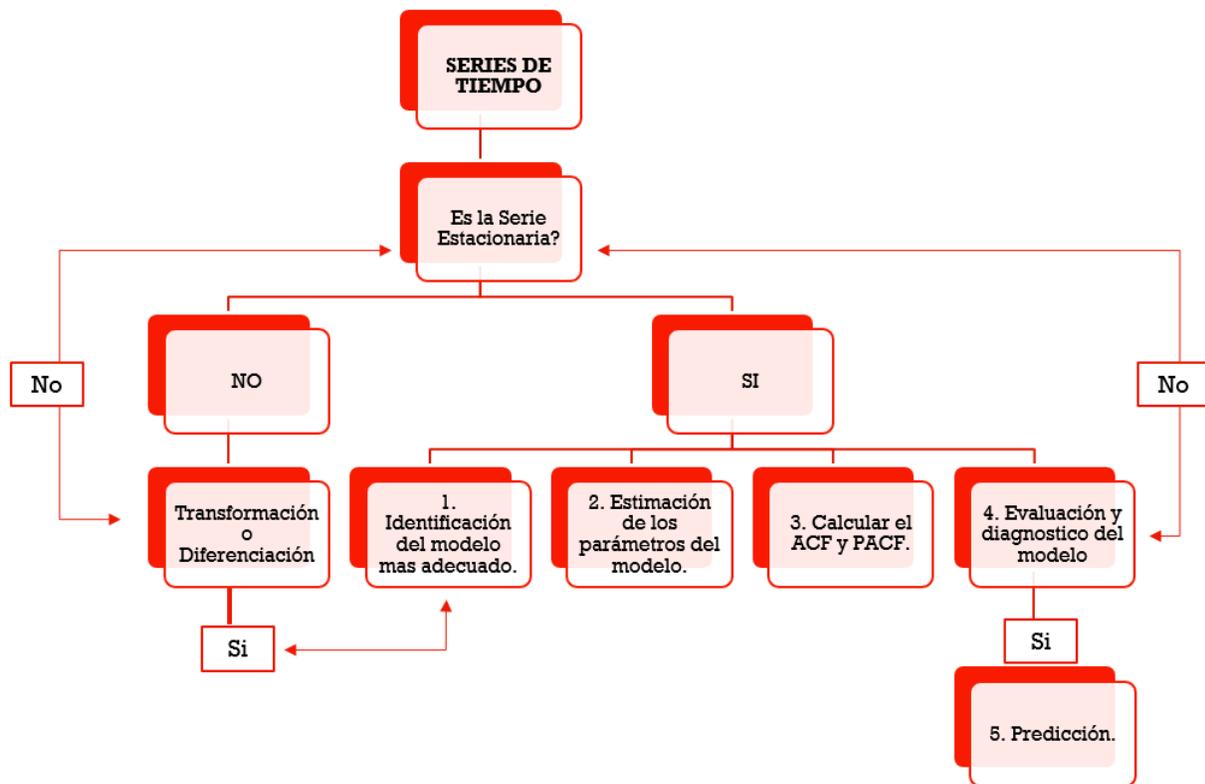


Figura 1: Metodología de Box-Jenkins

Según el cuadro anterior, si la serie es estacionaria o en otras palabras no tiene tantas fluctuaciones en el espacio temporal que se esté analizando se puede empezar a tratar, para eso se observa los gráficos de la serie original y sus diferencias, a continuación se debe evaluar cual es el mejor modelo para predecir, para eso se analizan las pruebas BIC O AIC con ayuda del software R, con las cuales se puede observar que tipo de modelo puede ser el mejor o el óptimo para arrancar hacer los respectivos análisis (auto.arima).

Después se estima los parámetros del modelo o se hace un análisis de que tan buenos son los coeficientes que arroja y se pasa a evaluar el modelo con otro tipo de pruebas estadísticas, finalmente se realiza la predicción.

En el día a día las series de tiempo tienen comportamientos muy inestables por lo cual se hace necesario realizar algunas transformaciones de los datos, entre las principales esta rezagar uno o más periodos, colocarlas en logaritmos, medias móviles, entre otras más complejas, en este punto hay que colocar especial atención, ya que al transformar los datos no debe perderse el comportamiento inicial porque podrían generarse predicciones no confiables.

### 3.1.2 Modelos AR

En un modelo AR, el valor de la serie en el momento  $t$  es una combinación lineal de las ultimas  $p$  observaciones de la variable. En los casos más sencillos se analiza la variable con un periodo anterior. Las observaciones fluctúan alrededor de  $\mu$ , que es la media de la serie.

Ejemplo:

$$AR(1) = y(t) = 5 + 0.2 * y(t - 1) + a(t)$$

$$\text{La media} = \frac{5}{1 - 0.2} = 6.25$$

Los procesos AR no pueden representar series de memoria muy corta, donde el valor actual de la serie solo esta correlacionado con un número pequeño de valores anteriores de la serie.

El parámetro  $\theta_1$  está relacionado con la memoria de la serie. Cuanto más cerca este de cero, la memoria es más corta. A medida que se incrementa, la memoria es mayor y la dependencia con el pasado es más fuerte. Un proceso  $AR(p)$  tendrá una función de autocorrelación parcial distinta de cero para los valores que estén dentro del grado de libertad del modelo, y para los siguientes será de valor cero.

### 3.1.3 Modelos MA

Los modelos MA siempre son estacionarios, por lo tanto, no hay que analizar ninguna condición de estacionariedad.

Un modelo MA será invertible en un AR ( $\infty$ ) estacionario si se cumple lo siguiente, raíces de la ecuación característica  $< 1$ , raíces del polinomio característico  $> 1$ . Los procesos MA son función de un número finito de innovaciones anteriores, no de todas.

En los casos de modelo MA(q) la función de autocorrelación parcial se comporta de la misma manera, con la acotación que es de forma exponencial sinusoidal amortiguada, como para la correlación completa de los modelos AR(p).

#### 3.1.4 Modelos ARMA

Los modelos ARMA son modelos mixtos que tienen componentes autorregresivos (p) y media móviles (q). Un modelo arma es una herramienta que como combina ambas formas de predecir una serie, en algunos casos ajusta mejor dicha predicción.

Un modelo ARMA(p,q) se puede considerar estacionario si su parte autorregresiva es estacionaria y su componente de medias móviles es invertible. Los modelos ARMA solo se pueden aplicar a series que no muestren tendencia, por lo que es muy relevante considerar la estacionariedad e invertibilidad.

Este tipo de modelos se utilizan en mayor proporción en modelos donde las variables en un instante t arrojan resultados no determinísticos, es decir, modelos estocásticos.

Las autocorrelaciones son muy parecidas a las del modelo AR(p), sin embargo, el decaimiento no inicia desde el principio. funciona solo dentro de los grados de libertad, con autocorrelaciones completas y fuera de ellas no tiene sentido la determinación.

#### 3.1.6 Criterios ACF y PACF

Se entiende que ACF es una función de autocorrelación (completa) que nos proporciona valores de autocorrelación de cualquier serie con sus valores rezagados.

En términos simples, describe qué tan bien se relaciona el valor presente de la serie con sus valores pasados. Una serie temporal puede tener componentes como tendencia, estacionalidad, cíclico y residual. ACF considera todos estos componentes mientras encuentra correlaciones, por lo tanto, es un 'gráfico de autocorrelación completo'

Por su parte se entiende que PACF es una función de autocorrelación parcial. Básicamente, en lugar de encontrar correlaciones del presente con retrasos como ACF, encuentra la correlación de los residuos (que permanece después de eliminar los efectos que ya se explicaron por los retrasos anteriores) con el siguiente valor de retraso, por lo tanto, 'parcial' y no 'completo' a medida que eliminamos las variaciones ya encontradas antes de encontrar la siguiente correlación.

Entonces, si hay alguna información oculta en el residuo que pueda modelarse en el próximo retraso, se podría obtener una buena correlación y mantener ese próximo retraso como una característica durante el modelado.

Con ayuda de los gráficos ACF y PACF se puede determinar qué tipo de modelo es, si Autorregresivo AR o de Media móvil MA, en los Modelos AR: ACF decae en el tiempo y PACF marca los rezagos en el gráfico, en el caso de los modelos MA es lo contrario. Cuando son modelos ARMA se puede observar en el gráfico que, aunque se presenten oscilaciones, los gráficos muestran decaimiento en el tiempo.

Lo anterior debido a que se encuentran características óptimas u orden del proceso AR utilizando el gráfico PACF, ya que elimina las variaciones explicadas por los retrasos anteriores, por lo que se obtienen solo las características relevantes. Se encuentran características óptimas u orden del proceso de MA utilizando el gráfico ACF, ya que al ser un proceso MA no tiene componentes estacionales y de tendencia, por lo que solo se obtiene la relación residual con los retrasos de las series de tiempo en el gráfico ACF. ACF actuando como una trama parcial.

### 3.1.7 Criterio de información Bayesiana (BIC)

Fue propuesto por Schwarz en (1978), ha sido uno de los métodos más populares usado para la selección de modelos. Este es un criterio de evaluación de modelos en términos de sus probabilidades posteriores, es una medida alternativa para el AIC para obtener información sobre el mejor modelo que se puede utilizar. El BIC se define como:  $BIC = -2 \log(L) + k \log(T)$ , donde  $k$  es el número de parámetros que varían ( $k = p + q$  en los modelos ARMA) y  $T$  es la cantidad de datos disponibles.

### 3.1.8 Criterio de información Akaike (AIC)

El AIC fue propuesto por Akaike (1974) como un estimador insesgado asintótico de la información de Kullback-Leibler esperada, entre un modelo candidato ajustado y el verdadero modelo. El AIC se define como:  $AIC = -2\ln(\theta_{bn}) + 2p^3$ .

## 3.2 Gestión de Riesgos

Constituye una práctica inherente a la actividad empresarial, razón por la cual no obedece a prescripciones jurídicas o normativas, sino a la intencionalidad estratégica de la Organización para preservar la integridad de los recursos empresariales, incrementar la ventaja competitiva y garantizar la continuidad del negocio frente a los diferentes riesgos a los que se encuentra expuesta.

### 3.2.1 Política de administración de riesgos.

Contactar gestiona los riesgos a los que se ve expuesta a través del fomento de una cultura para la gestión de riesgos y el desarrollo de modelos para cada tipo de riesgo con el fin de mitigarlos y contribuir al logro de los objetivos organizacionales

---

<sup>3</sup> Montesinos López, Abelardo. Estudio del AIC y BIC en la selección de modelos de vida con datos censurados. Disponible en <https://probayestadistica.cimat.mx/sites/default/files/PDFs/TE414MontesinosLopez.pdf>.

Con acceso el 21 de oct. de 19

### 3.2.2 Objetivos de la gestión administración de riesgos

Los objetivos de la gestión integral de riesgos en CONTACTAR son los siguientes:

- Fortalecer el Sistema de Administración de riesgos como un proceso estructurado riguroso y transparente, que contribuya efectivamente al logro de los objetivos y metas de Contactar.
- Unificar y fortalecer la cultura de Administración del riesgo y de control a lo largo de la organización, con base en estándares y modelos de “mejores prácticas” mundiales.
- Garantizar que las exposiciones de riesgo estén dentro de los parámetros aceptados por Contactar.
- Implementar herramientas para la identificación, gestión y control de riesgos
- Proteger el patrimonio de la institución.

## CAPÍTULO 4

---

### Marco Metodológico

---

A continuación, se citan algunos estudios anteriores a esta investigación con el fin de identificar aspectos relevantes que sirvan de guía en el pronóstico de mora por 30 de contactar bajo la metodología Box Jenkins utilizando una base de datos con periodo 2015 – 2019.

**Rodian Jasef, Saby Sanabria (2016).** En su trabajo “Pronostico de series de tiempo geodésicas mediante el uso del lenguaje de programación R y la implementación de los métodos estadísticos box-jenkins y holt-winters”, encontró las siguientes conclusiones importantes.

La evaluación de la autocorrelación sirve de soporte a la hora de seleccionar el modelo más adecuado para pronosticar una serie de tiempo, el análisis de los gráficos de autocorrelación simple y parcial ayudan al investigador a escoger, de una mejor manera, que tipo de modelo (autorregresivo (AR), media móvil (MA), ARMA (autorregresivo con media móvil), ARIMA (procesos integrados), entre otros) es el que mejor podrá representar el comportamiento de los datos a futuro.

El empleo de la función auto. arima facilita la selección del mejor modelo, dado que internamente el algoritmo escoge el que mejor resultados de pronóstico presentaría. Según la comparación de los parámetros AIC, BIC y máxima verosimilitud, pero se sigue realizando la evaluación de la autocorrelación dado que por medio de sus gráficos (simples y parciales) se puede asegurar que los modelos que selecciona la función auto. arima estén correctamente interrelacionados con la naturaleza de los datos.

Aunque los procesos de evaluación de autocorrelación entre los datos de una serie de tiempo se ejecutan indistintamente al método de pronóstico a utilizar, si sirven como punto de partida y comparación cuando se escoge Box-Jenkins, dado que, con base en sus gráficos, se pueden encontrar los parámetros que construirán los diferentes modelos según sea el caso.

Se hace necesario realizar una coherente evaluación previa de los datos insumo, que permita eliminar datos erróneos o atípicos de la serie de tiempo y con esto disminuir el riesgo de realizar análisis sobre datos que durante la ejecución de los procedimientos generen desviaciones y reduzcan las precisiones en los resultados finales.

No existe un único camino para pronosticar una serie de tiempo geodésica, la selección de un método u otro depende de las características propias de los datos, la precisión requerida y el tiempo de pronóstico que se quiere calcular, dado que, métodos como Box-Jenkins son reconocidos por su capacidad para proyectar datos a corto plazo, mientras que Holt-Winters tiene una mejor capacidad para pronosticar datos cuando las series de tiempo tienden a ser lineales y a presentar factores estacionales.

La evaluación y análisis de las autocorrelaciones simples y parciales ayudan a seleccionar el modelo que mejor describa la serie de tiempo a futuro, pero deben

estar acompañados de otra serie de análisis y pruebas estadísticas que garanticen que el(los) método(s) escogidos respondan real y precisamente al comportamiento pasado, presente y futuro de los datos.

Los modelos identificados, desarrollados, comprobados y ejecutados se basaron en métodos univariados, es decir, buscaron establecer una característica en particular durante un periodo de tiempo determinado, basándose única y exclusivamente en la historia de las propias series de tiempo por cada una de sus componentes.

A pesar de su grado de complejidad, formulaciones y análisis previos requeridos, el método Box-Jenkins permite establecer modelos de pronóstico mucho más consistentes a corto y mediano plazo, por eso, en el medio científico es uno de los métodos apetecidos por los investigadores a la hora de realizar estudios sobre series de tiempo. Para la escogencia de los mejores modelos predictivos se usaron las funciones `auto.arima` y `HoltWinters`, que de forma automática seleccionan los parámetros que mejor pronosticaran los comportamientos futuros de cada una de las series de tiempo por componente basados en el cumplimiento de una serie de requisitos estadísticos, como lo son, menores valores en los estadísticos AIC, BIC y máxima verosimilitud (para BoxJenkins) y menor error medio cuadrático (para HoltWinters).

Garantizar que los residuales generados tras la modelización de una serie de tiempo por medio del método Box-Jenkins cumplen con los supuestos de normalidad, estacionalidad, homocedasticidad y no autocorrelación es clave a la hora de entender de forma precisa y certera, el comportamiento pasado, presente y futuro de los datos vinculados a la serie de tiempo.

Siempre que se ejecute un modelo predictivo, por cualquiera de los métodos disponibles en la rama estadística, se hace necesario realizar sobre ellos

---

validaciones de testeo mediante el establecimiento de datos de entrenamiento y prueba, con el fin de evaluar la precisión y veracidad que tendrá el modelo y sus parámetros cuando sean pronosticados los valores requeridos en el tiempo.

El lenguaje estadístico R se convierte en un aliado fundamental para todo aquel interesado en desarrollar técnicas estadísticas y geodésicas a la hora de pronosticar series de tiempo, dado que en la actualidad cuenta con una infinidad de herramientas, paquetes y funciones que facilitan la identificación, cálculo, evaluación y determinación de modelos cada vez más precisos y acertados.<sup>4</sup>

**Según Diaz, Liens y Oliva de Colon Fidel (2016).** En su trabajo “Aproximación a la metodología Box-Jenkins para la predicción de la tasa de cambio EUR/USD” han evolucionado diversas teorías y modelos empíricos que intentan explicar los determinantes de los tipos de cambio, los cuales incluyen factores del ámbito financiero, componentes de expectativas y variables de la esfera real de la economía. Cada teoría contrastada selecciona un método de estimación y muchas veces se obtienen resultados contradictorios.

A la par del desarrollo teórico en países más desarrollados, en la región latinoamericana se han identificado numerosos trabajos en torno al tema, que van desde un análisis de los determinantes del tipo de cambio real, hasta el objetivo explícito de explicar las fluctuaciones del tipo de cambio spot para su predicción. En general, las técnicas más utilizadas son las de cointegración que traen asociado un mecanismo de corrección del error. No obstante, se utilizan los enfoques ARIMA y otros métodos menos tradicionales.

---

<sup>4</sup> Rodian Jasef, Saby Sanabria, Pronostico de series de tiempo geodésicas mediante el uso del lenguaje de programación R y la implementación de los métodos estadísticos box-jenkins y holt-winters (2016. Disponible e <https://repository.unimilitar.edu.co/bitstream/handle/10654/15481/SabySanabriaRodianJasef2016.pdf?sequence=1&isAllowed=y>). Con acceso el 17 de oct de 2019

En tal sentido, existe mucha controversia y debate sobre las potencialidades de los diversos métodos econométricos para explicar y predecir el comportamiento del tipo de cambio. Ya sean métodos deterministas, de alisamiento, modelos uniecuacionales, de ecuaciones simultáneas, enfoques ARIMA o modelos VAR para a cointegración, todos enfrentan defensores y detractores. Por tanto, no se puede afirmar que un método sea apropiado por sí mismo para explicar cualquier situación cotidiana, sino que deben contrastarse diferentes variantes y seleccionar aquella que reporte un mejor ajuste de la serie real o menor error de predicción.

En el trabajo, se apeló a las potencialidades que ofrece la metodología Box-Jenkins o método ARIMA para el tratamiento de series temporales, teniendo en cuenta que se centra en el análisis de las propiedades probabilísticas o estocásticas de las series de tiempo por sí mismas. De ahí que se aplicara esta perspectiva para el estudio de la serie de tipo de cambio EUR/USD, se comprobó que la serie utilizada no es estacionaria en niveles, pero se alcanzó la condición con una primera diferencia. Este fue un resultado esperado, pues es una característica que comparte la mayoría de las variables económicas. Además, se contrastaron 18 regresiones con componentes autorregresivos y de medias móviles.

Una vez evaluada la significación de los parámetros estimados, el cumplimiento de las condiciones de estacionariedad y la comparación de los criterios Akaike, se seleccionaron cuatro especificaciones con componentes AR y MA. No obstante, los resultados de la raíz del error cuadrático medio y coeficiente de Theil calculados en las proyecciones, de conjunto con las nociones teóricas sobre el tipo de cambio, condujeron a seleccionar las especificaciones MA(1) y MA(7) como mejores variantes de estimación.<sup>5</sup>

---

<sup>5</sup> Díaz, Liens y Oliva de Colon Fidel (2016). "Aproximación a la metodología Box-Jenkins para la predicción de la tasa de cambio EUR/USD" Disponible en <[http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S2073-60612016000100005](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2073-60612016000100005)>. Con acceso el 18 de octubre de 2019.

**Según García, Portilla, Nunez, Gonzales y Jaramillo (2007)**, En el artículo “Análisis de series de tiempo univariante aplicando metodología de Box-Jenkins para la predicción de ozono en la ciudad de Cali, Colombia”. Revista Facultad de Ingeniería de la Universidad de Antioquia, se concluyó que una vez estabilizada la serie se debe proceder a estudiar la presencia de regularidades para identificar posibles modelos matemáticos. Para ello se calculó la función de autocorrelación tanto simple como parcial y se compararon con los diferentes patrones que suministra la metodología Box-Jenkins, que son típicos de las diferentes funciones generadoras de datos, seleccionando los modelos que mejor se ajusten o adecuen a la forma de las funciones de autocorrelación obtenidas.

Después de la selección de los mejores modelos, se estimaron los coeficientes de los mismos y se procedió a efectuar un análisis de los residuos (diferencia entre el valor realmente observado y el valor previsto por el modelo), verificando la presencia de un comportamiento estacionario o de ruido blanco para seleccionar el modelo que presente el ajuste más adecuado.

El modelo que más se aproxima al proceso generador de la serie corresponde a un ARMA(24,4), cuya validación arroja un coeficiente de información ajustado  $R^2$  de 0,72, que puede interpretarse en forma muy general como un buen nivel de concordancia entre el modelo y los datos. <sup>6</sup> Después de revisar varios trabajos de investigación, se concluye que la metodología box Jenkins permite encontrar modelos eficientes con los cuales se puede predecir cualquier tipo de variable, por lo tanto, será utilizada con el fin de predecir la mora para Contactar.

---

<sup>6</sup> Garcia, Portilla, Nunez, Gonzales y Jaramillo (2007), Análisis de series de tiempo univariante aplicando metodología de Box-Jenkins para la predicción de ozono en la ciudad de Cali, Colombia. Revista Facultad de Ingeniería de la Universidad de Antioquia. Disponible en <<http://www.scielo.org.co/pdf/rfiua/n39/n39a07.pdf>>. Con acceso el 18 de octubre de 2019.

## CAPÍTULO 5

---

### Resultados

---

A continuación, se presentarán los modelos que se obtuvieron a partir de la base de datos de Contactar, bajo la metodología Box Jenkins. La base de datos disponible está comprendida entre mayo 2010 hasta Octubre 2019, sin embargo los datos son muy volátiles ya que la mora de la institución es muy sensible a cambios de política comercial, comportamientos de los clientes y variaciones en el mercado que afectan la economía, entre otras, por lo cual si se quiere datos confiables teniendo en cuenta que algunas decisiones financieras dependen en gran medida de este indicador, no es posible predecir más de 4 meses de mora, razón por lo cual dicha predicción se realiza en la institución cuatrimestralmente.

Para saber que tan bueno es el modelo frente a la realidad, se corrió la base desde mayo 2010 hasta abril 2019, con el fin de predecir la mora para los meses mayo 2019 a agosto 2019 que son datos que se conocen y hacer un comparativo frente a la mora real de la institución. Posteriormente se procedió a correr el modelo con la base de mayo 2010 hasta octubre 2019 con el fin de pronosticar los próximos 4 meses de mora par 30.

## 5.1 Metodología Box Jenkins

A continuación, se puede observar cómo ha sido el comportamiento de la mora par 30 en el tiempo. Los gráficos muestran una gran volatilidad de y algunos picos explicados por lo que se mencionaba anteriormente, políticas comerciales, factores externos entre otros.

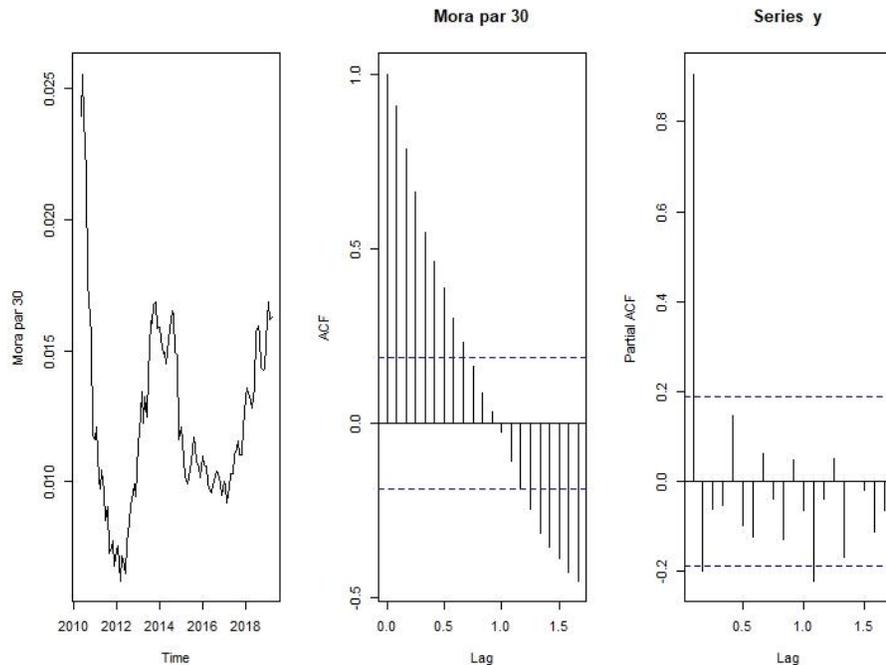


Figura 2. Mora par 30

### 5.1.1 Identificación del modelo más adecuado

#### Modelo 1. SARIMA (1,1,1) x (1,0,0)<sub>12</sub>

```
ajuste <- arima(y, order=c(1,1,1), seasonal = list(order=c(1,0,0), frequency=12))
```

```
ajuste$coef[ajuste$coef !=0]/sqrt(diag(ajuste$var.coef))
```

ar1	ma1	sar1
10.280.953	-6.590.780	3.311.666

Tabla 1. Modelo SARIMA (1,1,1) x (1,0,0)<sub>12</sub>

**Modelo 2. SARIMA (1,1,0) x (1,0,0)<sub>12</sub>**

```
ajuste <- arima(y, order=c(1,1,0), seasonal = list(order=c(1,0,0), frequency=12))
ajuste$coef[ajuste$coef !=0]/sqrt(diag(ajuste$var.coef))
```

ar1	sar1
1.702.943	3.116.688

Tabla 2. Modelo SARIMA (1,1,0) x (1,0,0)<sub>12</sub>**Modelo 3. SARIMA (1,1,0) x (1,0,1)<sub>12</sub>**

```
ajuste <- arima(y, order=c(1,1,0), seasonal = list(order=c(1,0,1), frequency=12))
ajuste$coef[ajuste$coef !=0]/sqrt(diag(ajuste$var.coef))
```

ar1	sar1	sma1
2.313.749	4.240.183	-2.031.194

Tabla 3. Modelo SARIMA (1,1,0) x (1,0,1)<sub>12</sub>**Modelo 4. SARIMA (1,1,1) x (2,1,0)<sub>12</sub>**

```
ajuste <- arima(y, order=c(1,1,1), seasonal = list(order=c(2,1,0), frequency=12))
ajuste$coef[ajuste$coef !=0]/sqrt(diag(ajuste$var.coef))
```

ar1	ma1	sar1	sar2
12.456.403	-7.064.284	-3.440.796	-1.283.598

Tabla 4. Modelo SARIMA (1,1,1) x (2,1,0)<sub>12</sub>

Después de la revisión exhaustiva de varios modelos y teniendo en cuenta la significancia estadística de los parámetros, el análisis de pruebas como autocorrelación serial, normalidad en los residuos, BIC, AIC y aplicando el principio de Parsimonia con el cual se esperaba utilizar la menor cantidad de parámetros para verificar que no siempre los modelos más complejos son los mejores, ya que con un

modelo sencillo si se puede predecir de forma más precisa la mora se determinó que el modelo 1 **SARIMA (1,1,1) x (1,0,0)**<sub>12</sub>, tiene mejores parámetros.

### 5.1.2 Estimación de los parámetros del modelo

SARIMA (1,1,1) x (1,0,0);  $p=1, d=1, q=1, P=1, D=0, Q=0$

Donde,

$d=1; z_t=y_t-y_{t-1}$  se refiere a restar el mes inmediatamente anterior

$D=1, z_t=y_T-y_{t-s}$  Se refiere a restar un mes de ese mismo mes pasado (ver tabla 1)

A continuación, se presenta la prueba de Raíz Unitaria, a partir de la cual se busca determinar que los datos tienen alguna tendencia.

### Prueba Raíz Unitaria

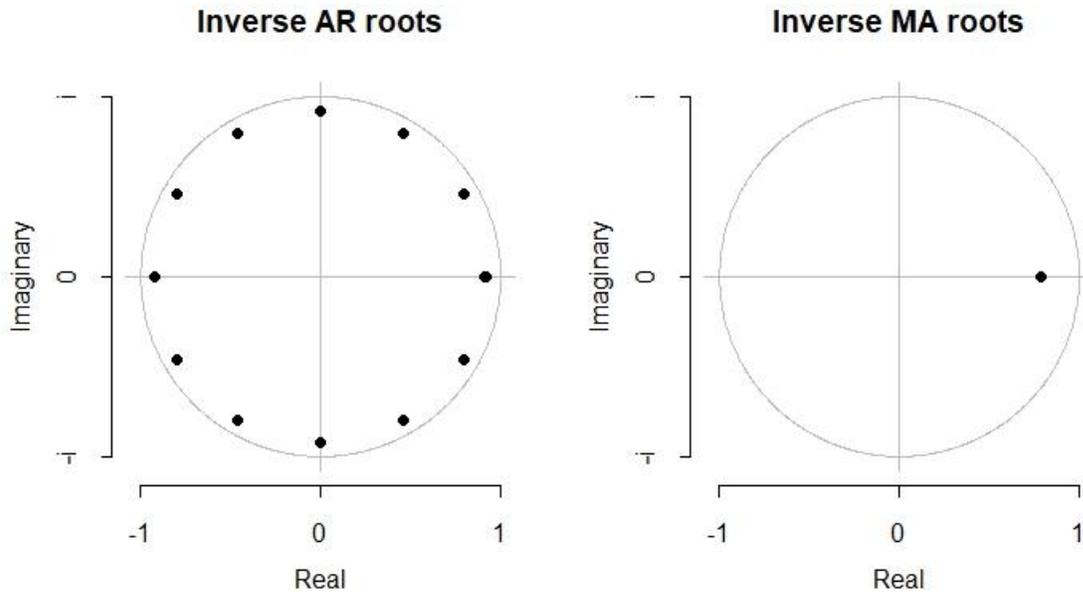


Figura 3. Prueba Raíz Unitaria

En el grafico se observa que todas las raíces unitarias están dentro del círculo, por lo cual se puede inferir que la serie es estacionaria.

## Prueba Dickey-Fuller Test

Dickey-Fuller = -4.2981, Lag order = 4, p-value = 0.01

El P valor de la prueba Dickey Fuller es menor al 5%, lo cual indica que la serie es estacionaria

### 5.1.3 Calculo de ACF y PACF

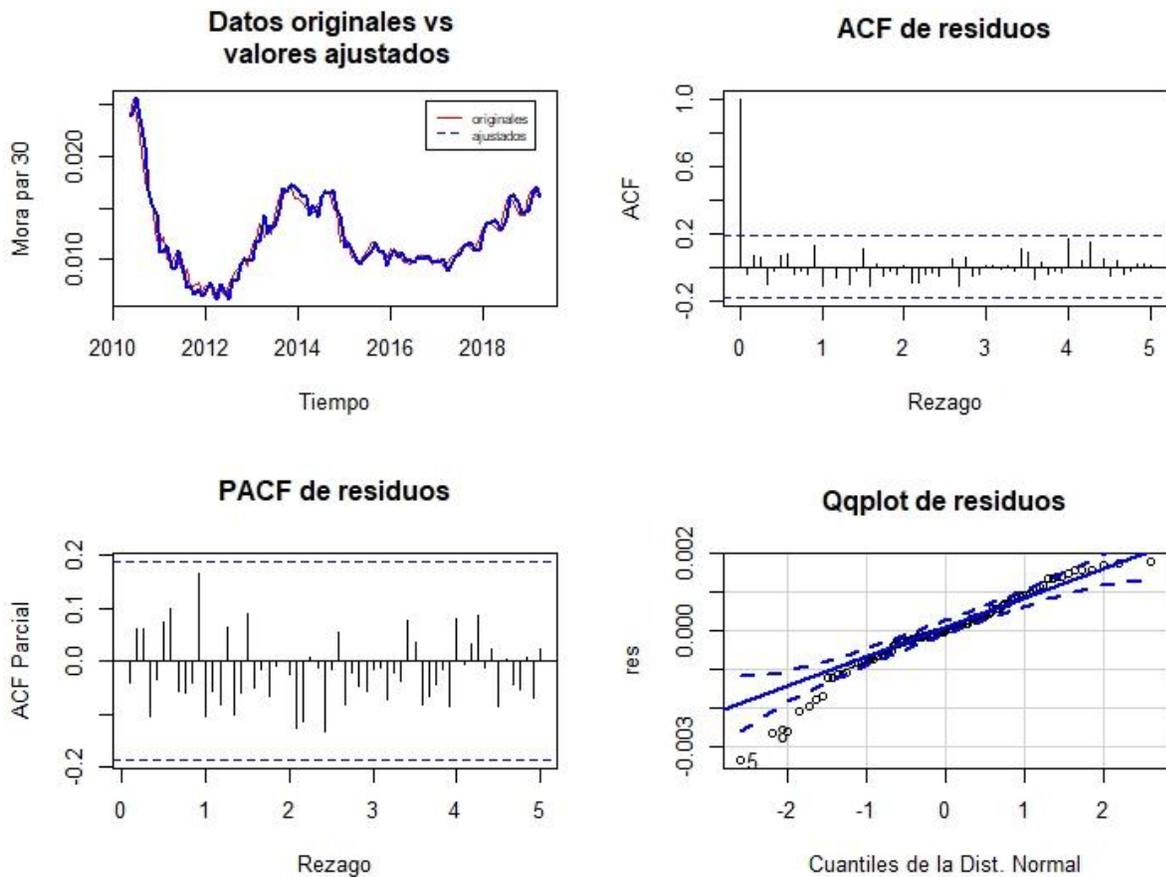


Figura 4. Diagnóstico del modelo SARIMA (1,1,1) x (1,0,0)<sub>12</sub>

En el diagnóstico del modelo se muestra en el primer gráfico la tendencia en el tiempo, en donde se observa que los datos originales frente a los datos ajustados tienen un buen ajuste, posteriormente en el gráfico del ACF y PACF de los residuos se observa que no hay problemas de autocorrelación, ya que las barras se encuentran dentro de las líneas punteadas, el gráfico Qqplot muestra algunos datos atípicos sugiere que los residuos se configuran como ruido blanco, pero no como ruido blanco gaussiano.

#### 5.1.4 Evaluación y diagnóstico del modelo

Modelos	Autocorrelación Box Pierce test	Normalidad Jarque Bera Test	BIC	AIC
Modelo 1	0.6656	0.0002265	-1.165.007	-1.175.698
Modelo 2	0.6829	0.0001067	-1.163.984	-1.172.002
Modelo 3	0.5818	0.0008662	-1.159.884	-1.170.576
Modelo 4	0.7739	0.01657	-1.014.863	1.027.632

Tabla 5. Comparativo test estadísticos modelos corridos

El test de Autocorrelación de Box Pierce es mayor al 5%, lo cual significa que con una confianza del 95% no hay evidencia de autocorrelación, por otro lado, la prueba de Jarque Bera, indica que los datos no se distribuyen de forma normal, como en la campana de Gauss, sin embargo, es entendible por la fluctuación de los datos, el BIC y AIC son mejores en el modelo 1, con lo cual se confirma que es el mejor modelo.

#### 5.1.5 Predicción

```
plot(forecast(ajuste, h=4))
forecast(ajuste, h=4)
```

Fecha	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
may-19	1,62%	1,50%	1,75%	1,44%	1,81%
jun-19	1,66%	1,48%	1,84%	1,38%	1,94%
jul-19	1,74%	1,51%	1,98%	1,38%	2,11%
ago-19	1,76%	1,47%	2,04%	1,32%	2,20%

Tabla 6. Predicción con niveles de confianza

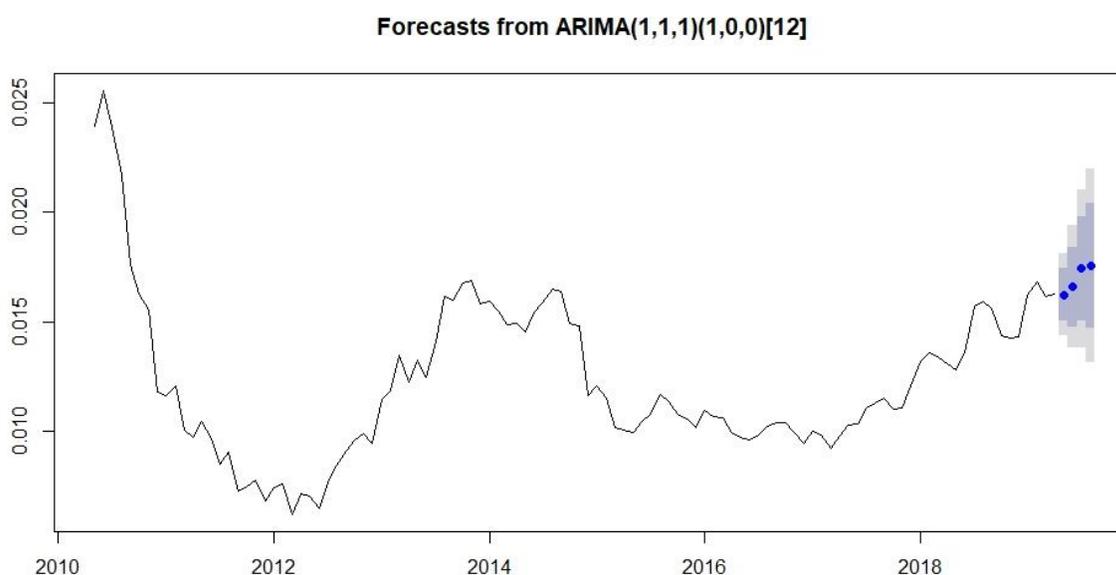


Figura 5. Predicción R studio hasta Agosto 2019

A continuación, se presenta la comparación entre las predicciones del área comercial, riesgos, la real y la que se muestra en la tabla 6 tomando la columna Point Forecast.

Fecha	Riesgos	Comercial	Point Forecast	Real
may-19	1,707%	1,746%	1,624%	1,566%
jun-19	1,719%	1,893%	1,659%	1,644%
jul-19	1,978%	1,907%	1,743%	1,881%
ago-19	2,055%	1,866%	1,757%	1,977%

Tabla 7. Comparativo Predicción áreas Contactar vs realizada vs Real

Posteriormente se procedió a correr el mismo modelo con la base de datos hasta octubre 2019 con el fin de conocer como cerrará la mora par 30 de Contactar a 2019 y como se presume será la mora para los meses de enero y febrero de 2020.

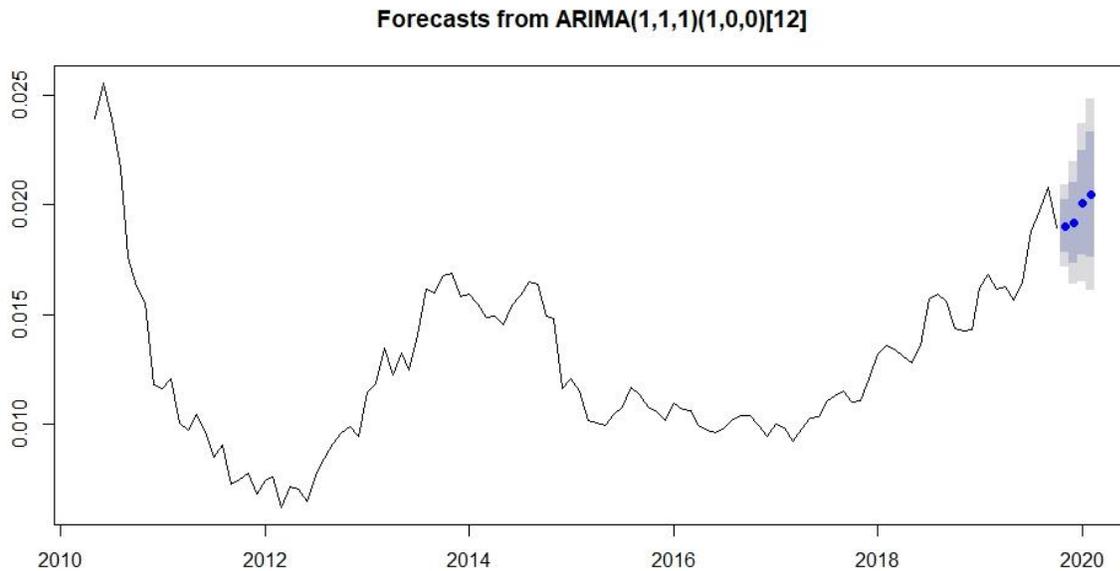


Figura 6. Predicción R studio hasta febrero 2020

<b>Fecha</b>	<b>Point Forecast</b>	<b>Lo 80</b>	<b>Hi 80</b>	<b>Lo 95</b>	<b>Hi 95</b>
nov-19	1,90%	1,78%	2,03%	1,72%	2,09%
dic-19	1,92%	1,73%	2,10%	1,64%	2,20%
ene-20	2,01%	1,77%	2,25%	1,65%	2,37%
feb-20	2,05%	1,76%	2,33%	1,61%	2,48%

Tabla 8. Predicción con niveles de confianza

## CAPÍTULO 6

---

### Conclusión y recomendaciones

---

- La calidad de la mora de la institución es muy buena, se entiende por calidad de la mora como un indicador por debajo del promedio de las instituciones a la par en el mercado, sin embargo, desde el área de Riesgos se busca monitorear constantemente los indicadores de mora par 30, con el fin de adelantarse ante posibles amenazas, por otra parte, la predicción es muy importante ya que con ella se toman decisiones financieras especialmente las concernientes al deterioro de la cartera, con lo cual la institución realiza una previsión para mitigar riesgos de liquidez.
- Por lo anterior la predicción realizada tiene que ser bastante robusta, por lo cual no se puede predecir demasiados meses ya que los datos son muy volátiles y a medida que se toman más meses la predicción se vuelve menos confiable.
- Bajo la metodología Box Jenkins se realizaron dos modelos, el primero con el fin de validar los datos que ya se conocen y analizar qué tan bueno es con los datos reales y el segundo con el fin de conocer que se espera de la mora para el cierre del año y con ello planear las metas comerciales del próximo.

- Después de correr distintos tipos de modelos AR, MA, SARIMA, y de realizar con ellos una serie de ejercicios, se determinó que el modelo más pertinente es el SARIMA ya que se ajusta muy bien a los datos reales de la mora par 30 de Contactar, se trata de un modelo sencillo y parsimonioso ya que no tiene muchos parámetros y los que tiene son significativos.
- Mas adelante en el comparativo de las predicciones se observa que los indicadores BIC y AIC son buenos para el modelo escogido por lo cual se procedió a realizar la predicción, arrojando datos muy parecidos a los que se conocen.
- Se recomienda realizar la predicción nuevamente en enero 2020, para empezar a analizar los cuatrimestres del siguiente año.

- Box, G; Jenkins, G. (1970). Time series analysis: forecasting and control.
- Brockwell, P; Davis, V (2002). Introduction to Time Series and forecasting.
- Diaz, Liens y Oliva de Colon Fidel (2016). "Aproximación a la metodología Box-Jenkins para la predicción de la tasa de cambio EUR/USD" Disponible en <[http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S2073-60612016000100005](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2073-60612016000100005)>. Con acceso el 18 de octubre de 2019.
- EIF. Entidad de intermediación financiera. Disponible en <[http://www.eco-finanzas.com/diccionario/I/Intermediacion\\_Financiera.htm](http://www.eco-finanzas.com/diccionario/I/Intermediacion_Financiera.htm)>. Con acceso el 07 de Septiembre de 2019.
- Farnsworth, V (2008). Econometrics in R.
- Forero-Lozano, S, Ballesteros-Ballesteros, V. y Toledo, J. (2018). Estadística de gradiente: una opción para realizar pruebas de hipótesis en escenarios de tamaño de muestra pequeño. Revista Internacional de Investigación de Ingeniería Aplicada, 13 (23), 16368-16375.
- Forero-Lozano, S, Ballesteros-Ballesteros, V. y Nisperuza-Toledo, J. (2018). Improved Likelihood Ratio Tests in Power Series Generalized Nonlinear Models. International Journal of Applied Engineering Research, 13(22), 15798-15805.
- García, Portilla, Nunez, Gonzales y Jaramillo (2007), Análisis de series de tiempo univariante aplicando metodología de Box-Jenkins para la predicción de ozono en la ciudad de Cali, Colombia. Revista Facultad de Ingeniería de la Universidad de Antioquia. Disponible en <<http://www.scielo.org.co/pdf/rfiua/n39/n39a07.pdf>>. Con acceso el 18 de octubre de 2019.

- 
- Montesinos Lopez, Abelardo. Estudio del AIC y BIC en la selección de modelos de vida con datos censurados. Disponible en <<https://probayestadistica.cimat.mx/sites/default/files/PDFs/TE414MontesinosLopez.pdf>>. Con acceso el 21 de oct. de 19.
  - Rodian Jasef, Saby Sanabria, Pronóstico de series de tiempo geodésicas mediante el uso del lenguaje de programación R y la implementación de los métodos estadísticos box-jenkins y holt-winters (2016). Disponible e <<https://repository.unimilitar.edu.co/bitstream/handle/10654/15481/SabySanabriaRodianJasef2016.pdf?sequence=1&isAllowed=y>>. Con acceso el 17 de oct de 2019.
  - Superintendencia Financiera De Colombia. Circular básica contable y financiera (Circular externa 100 de 1995, Cap. II. Disponible en <<https://www.superfinanciera.gov.co/jsp/loader.jsf?lServicio=Publicaciones&lTipo=publicaciones&lFuncion=loadContenidoPublicacion&id=15466>>. Con acceso el 06 de Septiembre de 2019.
  - Wei, William. (2006). Time series analysis: Univariate and multivariate methods.

