

UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO

FACULTAD DE CIENCIAS FISICAS Y MATEMATICAS

ESCUELA PROFESIONAL DE ESTADISTICA



TESIS

“Pronóstico de Las Recaudaciones Tributarias de la Actividad Económica Comercio a Nivel Nacional, Mediante La Metodología de Box - Jenkins y Redes Neuronales Artificiales para el año 2018”

PARA OPTENER EL TÍTULO PROFESIONAL DE:

LICENCIADA EN ESTADÍSTICA.

Investigadora:

Bach. Jacksumy Deisy Gamonal Ortíz

Asesor:

Dr. Alfonso Tesen Arroyo

Lambayeque, febrero 2021

UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS ESCUELA PROFESIONAL DE ESTADÍSTICA



Dr. Jorge Antonio Acosta Piscocoya.
Presidente



Dr. Wilver Omero Rodríguez López
Secretario



M.Sc. Grimaldo Dermalí Benavides Campos
Vocal



Dr. Alfonso Tesen Arroyo
Asesor

ACTA DE SUSTENTACION



UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
FACULTAD DE CIENCIAS FISICAS Y MATEMATICAS
DECANATO
Ciudad Universitaria - Lambayeque



ACTA DE SUSTENTACIÓN VIRTUAL N°038-2021-D/FACFyM

Siendo las 4:00 pm del día lunes 22 de noviembre del 2021, se reunieron vía plataforma virtual, <http://meet.google.com/nqo-brvs-cjr> los miembros del jurado evaluador de la Tesis titulada: "Pronóstico de Las Recaudaciones Tributarias de la Actividad Económica Comercio a Nivel Nacional, Mediante La Metodología de Box - Jenkins y Redes Neuronales Artificiales para el año 2018"

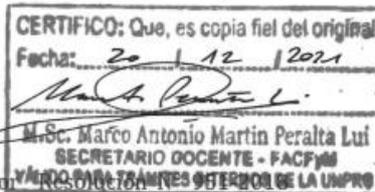
Designados por Resolución N° 1464-2018 D/FACFyM de fecha 16 de noviembre del 2021

Con la finalidad de evaluar y calificar la sustentación de la tesis antes mencionada, conformada por los siguientes docentes:

Dr. Jorge Antonio Acosta Piscoya Presidente

Dr. Wilver Omero Rodríguez López Secretario

M.Sc. Grimaldo Dermalí Benavides Campos Vocal



La tesis fue asesorada por el Dr. Alfonso Tesen Arroyo, nombrado por Resolución N° 830-2021-virtual D/FACFyM de fecha 9 de agosto del 2018.

El Acto de Sustentación fue autorizado por Resolución N° 830-2021-virtual D/FACFyM de fecha 16 de noviembre del 2021.

La Tesis fue presentada y sustentada por la Bachiller: Jacksumy Deisy Gamonal Ortíz, y tuvo una duración de 30 minutos.

Después de la sustentación, y absueltas las preguntas y observaciones de los miembros del jurado se procedió a la calificación respectiva, otorgándole el Calificativo de 17 (Diecisiete) en la escala vigesimal, mención Bueno.

Por lo que queda apto para obtener el Título Profesional de Licenciada en Estadística de acuerdo con la Ley Universitaria 30220 y la normatividad vigente de la Facultad de Ciencias Físicas y Matemáticas y la Universidad Nacional Pedro Ruiz Gallo.

Siendo las 17:10 pm se dio por concluido el presente acto académico, dándose conformidad al presente acto con la firma de los miembros del jurado.

Dr. Jorge Antonio Acosta Piscoya
Presidente

M.Sc. Grimaldo Dermalí Benavides Campos
Vocal

Dr. Wilver Omero Rodríguez López
Secretario

Dr. Alfonso Tesén Arroyo
Asesor

Declaración jurada de Originalidad

Yo, Jacksumy Deisy Gamonal Ortiz, investigador principal, y Dr. Alfonso Tesen Arroyo, asesor del trabajo de investigación “Pronóstico de Las Recaudaciones Tributarias de la Actividad Económica Comercio a Nivel Nacional, Mediante La Metodología de Box - Jenkins y Redes Neuronales Artificiales para el año 2018”, declaramos bajo juramento que este trabajo no ha sido plagiado, ni contiene datos falsos. En caso se demostrará lo contrario, asumo responsablemente la anulación de este informe y por ende el proceso administrativo a que hubiera lugar. Que puede conducir a la anulación del título o grado emitido como consecuencia de este informe.

Lambayeque, 17 de febrero del 2020



Bach. Jacksumy Deisy Gamonal Ortiz
Investigadora



Dr. Alfonso Tesen Arroyo
Asesor

DEDICATORIA

A mis padres Margarita y Medardo por haberme forjado como la persona que soy con su ejemplo, sus consejos y por haberme dejado la herencia más valiosa, mi educación; los logros que hasta hoy he alcanzado se lo debo a ustedes entre los que se incluye éste.

A mi hermana Mónica, con quién he festejado alegrías, logros y dado la mano en las vicisitudes de la vida.

A mi pequeña hermosa hija Libi Scarlett, mi motor y motivo de seguir y escalar por las sendas de la vida.

Jacksumy Deisy

AGRADECIMIENTO

A Dios, el ser omnipotente que guía día a día mi camino.

A mis queridos padres por su apoyo incondicional, moral y económico en mi formación profesional, gracias por tanto.

A mis formadores, personas de gran sabiduría que en aulas impartieron conocimientos y experiencias de aprendizaje.

A mi querida hermana por su comprensión y apoyo constante.

A mi hija por llenar de alegría mis días.

Al Dr. Alfonso Tesen Arroyo por el asesoramiento de la presente tesis.

Jacksumy Deisy.

RESUMEN

Esta investigación tuvo como objetivo estimar el pronóstico de las recaudaciones tributarias de la actividad económica comercio a nivel nacional para el año 2018.

El diseño de investigación empleado para este estudio fue no experimental de tipo longitudinal, predictivo y descriptivo; la población estuvo conformada por el registro de los ingresos tributarios de Impuesto General a las ventas (IGV) interno, obtenido de la SUNAT, de igual modo la muestra está compuesta por los ingresos Tributarios recaudados del IGV internas, por la SUNAT en el periodo enero 2007 – diciembre del 2017.

Para lograr el objetivo planteado se aplicaron la metodología de Box y Jenkins y redes neuronales artificiales; encontrándose un modelo $SARIMA(0,1,1)(1,1,2)_{12}$, y un modelo $NNAR(2,1,2)_{12}$ respectivamente.

Finalmente se concluyó que el modelo que mejor pronostica IGV según actividad económica Comercio, enero 2007 a diciembre 2017, es el modelo $SARIMA(0,1,1)(1,1,2)_{12}$.

Palabras clave: IGV, pronóstico, series de tiempo, red neuronal.

ABSTRACT

The objective of this research was to estimate the forecast of tax revenue from commercial economic activity in Peru, using the box Jenkins methodologies and neural networks for the year 2018.

The research design used for this study was longitudinal, predictive and descriptive non-experimental; The population was made up of the record of internal General Sales Tax (IGV) tax revenues, obtained from SUNAT, in the same way the sample is composed of Tax revenues collected from internal VAT, by SUNAT in the January period 2007 - December 2017.

To achieve the proposed objective, the Box and Jenkins methodology and artificial neural networks were applied; finding a model $SARIMA(0,1,1)(1,1,2)_{12}$, and a model $NNAR(2,1,2)_{12}$ respectively.

Finally, it was concluded that the model that best predicts IGV according to economic activity Commerce, January 2007 to December 2017, is the model $SARIMA(0,1,1)(1,1,2)_{12}$.

Keywords: IGV, forecast, time series, neural network.

INDICE

INTRODUCCIÓN	1
CAPÍTULO I: DISEÑO TEÓRICO.....	3
1.1 Antecedentes.....	3
1.1.1 Antecedentes locales	3
1.1.2 Antecedentes Nacionales.....	3
1.2 Series de tiempo.....	4
1.2.1 Clasificación de una serie de tiempo.....	4
Estacionaria.	4
No estacionaria.	4
1.2.2 Componentes de una serie de tiempo	4
1.2.3 Procesos estocásticos.....	4
Procesos estocásticos estacionarios.....	4
Procesos estocásticos no estacionarios.....	5
1.2.4 Ruido blanco	5
1.2.5 Prueba de Dickey-Fuller aumentada (DFA).....	5
1.2.6 Modelos univariantes de series temporales	5
Proceso autorregresivo $AR(p)$	5
Proceso de medias móviles $MA(q)$	6
Modelos $ARMA(p, q)$	6
Modelo $ARIMA(p, d, q)$	6
Modelo $SARIMA(p, d, q)(P, D, Q)$	7
1.2.7 Metodología para el modelamiento de series temporales	8
Identificación del modelo.....	9
Estimación de los parámetros del modelo.....	9
Diagnóstico del modelo.....	10
Prueba de Ljung-Box	10

Prueba de Kolmogorov- Smirnov	11
Pronóstico.....	11
1.3 Redes neuronales.....	12
1.3.1 Objetivo de las Redes Neuronales.....	13
1.3.2 Elementos que componen una red neuronal.....	13
Entradas	13
Salidas	13
1.3.3 Estructura de una red neuronal.....	13
El perceptrón	13
El perceptrón multicapa	14
1.3.4 Funciones de las Redes Neuronales Artificiales	15
Backpropagation.....	16
1.3.5 Métodos de aprendizaje.....	17
Redes con aprendizaje supervisado.....	17
Redes con aprendizaje no supervisado.....	17
1.3.6 Redes neuronales aplicadas a las series de tiempo.....	18
Modelo NAR caso univariado.....	18
CAPÍTULO II. MÉTODOS Y MATERIALES	19
2.1 Tipo y diseño de investigación	19
2.2 Población y muestra.....	19
2.3 Técnica e instrumento de recolección de datos	19
2.4 Análisis y procesamiento estadístico de los datos	19
CAPÍTULO III. RESULTADOS Y DISCUSIÓN	21
3.1 Descripción del comportamiento de los ingresos tributarios.....	21
3.2 Estimación del modelo para el pronóstico de los ingresos tributarios, mediante la metodología de Box y Jenkins.	29
3.3 Estimación del modelo para el pronóstico de los ingresos tributarios, mediante la metodología de redes neuronales.	69

3.4	Estimación a través de las diferentes metodologías utilizadas para la predicción de los ingresos tributarios para el año 2018.....	72
3.5	Identificar el mejor modelo de series de tiempo para pronosticar los ingresos tributarios para el año 2018.....	73
CAPÍTULO IV. CONCLUSIONES		75
REFERENCIAS BIBLIOGRÁFICAS		76
APÉNDICE Y ANEXOS		80
ANEXO 1		80

Lista de tablas

Tabla 1. Estadísticos descriptivos del IGV según actividad económica Comercio, enero 2007 a diciembre 2017	21
Tabla 2. Test de Dickey-Fuller para IGV actividad económica Comercio, enero 2007 – diciembre 2017	24
Tabla 3. Test de Dickey-Fuller de la primera diferencia al IGV actividad económica Comercio, enero 2007 – diciembre 2017	28
Tabla 4. Modelos propuestos que explican el comportamiento de la serie de tiempo IGV actividad económica Comercio	29
Tabla 5. Test de Ljung-Box de los residuos del modelo SARIMA(0,1,1)(1,1,0) ₁₂	31
Tabla 6. Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA(0,1,1)(1,1,0) ₁₂	33
Tabla 7. Test de Ljung-Box de los residuos del modelo SARIMA(0,1,1)(1,1,0) ₁₂	35
Tabla 8. Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA(0,1,1)(1,1,2) ₁₂	37
Tabla 9. Test de Ljung-Box de los residuos del modelo SARIMA(0,1,1)(1,1,2) ₁₂	39
Tabla 10. Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA(0,1,1)(1,1,2) ₁₂	41
Tabla 11. Test de Ljung-Box de los residuos del modelo SARIMA(2,1,0)(0,1,1) ₁₂	43
Tabla 12. Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA(2,1,0)(0,1,1) ₁₂	45
Tabla 13. Test de Ljung-Box de los residuos del modelo SARIMA(2,1,0)(1,1,0) ₁₂	47
Tabla 14. Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA(2,1,0)(1,1,0) ₁₂	49
Tabla 15. Test de Ljung-Box de los residuos del modelo SARIMA(2,1,0)(1,1,2) ₁₂	51
Tabla 16. Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA(2,1,0)(1,1,2) ₁₂	53
Tabla 17. Test de Ljung-Box de los residuos del modelo SARIMA(3,1,0)(0,1,1) ₁₂	55

Tabla 18. Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA(3,1,0)(0,1,1) ₁₂	57
Tabla 19. Test de Ljung-Box de los residuos del modelo SARIMA(3,1,0)(1,1,0) ₁₂	59
Tabla 20. Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA(3,1,0)(1,1,0) ₁₂	61
Tabla 21. Test de Ljung-Box de los residuos del modelo SARIMA(3,1,0)(1,1,2) ₁₂	63
Tabla 22. Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA(3,1,0)(1,1,2) ₁₂	65
Tabla 23. Pronósticos del IGV actividad económica Comercio con la metodología Box y Jenkins para el año 2018	66
Tabla 24. Comparación de la evaluación de los errores de pronósticos de los modelos propuestos.....	67
Tabla 25. Información sobre la red neuronal artificial - MPL, aplicado a los Ingresos Tributarios por Impuesto General a la Ventas a nivel nacional.....	70
Tabla 26. Pronósticos del IGV actividad económica Comercio para el año 2018, usando la metodología de Box y Jenkins y redes neuronales.....	72
Tabla 27. Comparación de los modelos de redes neuronales y SARIMA	73
Tabla 28. IGV a nivel nacional, recaudados por la SUNAT, periodo enero 2007 – diciembre 2017.....	80

Lista de figuras

Figura 1. Esquema del modelado de series de tiempo	¡Error! Marcador no definido.
Figura 2. Estructura de una red neuronal	12
Figura 3. Modelo de una red neuronal artificial.....	13
Figura 4. Modelo de una red monocapa.....	14
Figura 5. Arquitectura de un Perceptrón simple	14
Figura 6. Arquitectura de un Perceptrón multicapa	15
Figura 7. Función de activación	16
Figura 8. Backpropagation Network	17
Figura 9. Comportamiento de la recaudación del IGV según actividad económica Comercio, enero 2007 a diciembre 2017	21
Figura 10. Diagrama de cajas del IGV según actividad económica Comercio, enero 2007 a diciembre 2017.....	22
Figura 11. Función de Autocorrelación simple a la recaudación IGV según actividad económica Comercio	23
Figura 12. Función de autocorrelación parcial de la recaudación del IGV según actividad económica Comercio, enero 2007 – diciembre 2017	24
Figura 13. Primera diferencia aplicada al IGV actividad económica Comercio, enero 2007 – diciembre 2017.....	25
Figura 14. Grafica de cajas de la primera diferencia al IGV según actividad económica Comercio, enero 2007 a diciembre 2017	26
Figura 15. Función de autocorrelación simple de la primera diferencia del IGV actividad económica Comercio, enero 2007 – diciembre 2017	27
Figura 16. Función de autocorrelación parcial de la primera diferencia del IGV actividad económica Comercio, enero 2007 – diciembre 2017.....	27
Figura 17. Residuos del modelo $SARIMA(0,1,1)(1,1,0)_{12}$	30
Figura 18. Función de autocorrelación de los residuos del modelo $SARIMA(0,1,1)(1,1,0)_{12}$	30

Figura 19. Función de autocorrelación parcial de los residuos del modelo $SARIMA(0,1,1)(1,1,0)_{12}$	31
Figura 20. Histograma de los residuos del modelo $SARIMA(0,1,1)(1,1,0)_{12}$	32
Figura 21. Grafica Q-Q normal de los residuos del modelo $SARIMA(0,1,1)(1,1,0)_{12}$	32
Figura 22. Residuos del modelo $SARIMA(1,1,0)(1,1,0)_{12}$	34
Figura 23. Función de autocorrelación de los residuos del modelo $SARIMA(1,1,0)(1,1,0)_{12}$	34
Figura 24. Función de autocorrelación parcial de los residuos del modelo $SARIMA(1,1,0)(1,1,0)_{12}$	35
Figura 25. Histograma de los residuos del modelo $SARIMA(1,1,0)(1,1,0)_{12}$	36
Figura 26. Grafica Q-Q normal de los residuos del modelo $SARIMA(1,1,0)(1,1,0)_{12}$	36
Figura 27. Residuos del modelo $SARIMA(0,1,1)(1,1,2)_{12}$	38
Figura 28. Función de autocorrelación de los residuos del modelo $SARIMA(0,1,1)(1,1,2)_{12}$	38
Figura 29. Función de autocorrelación parcial de los residuos del modelo $SARIMA(0,1,1)(1,1,2)_{12}$	39
Figura 30. Histograma de los residuos del modelo $SARIMA(0,1,1)(1,1,2)_{12}$	40
Figura 31. Grafica Q-Q normal de los residuos del modelo $SARIMA(0,1,1)(1,1,2)_{12}$	40
Figura 32. Residuos del modelo $SARIMA(2,1,0)(0,1,1)_{12}$	42
Figura 33. Función de autocorrelación de los residuos del modelo $SARIMA(2,1,0)(0,1,1)_{12}$	42
Figura 34. Función de autocorrelación parcial de los residuos del modelo $SARIMA(2,1,0)(0,1,1)_{12}$	43
Figura 35. Histograma de los residuos del modelo $SARIMA(2,1,0)(0,1,1)_{12}$	44
Figura 36. Grafica Q-Q normal de los residuos del modelo $SARIMA(2,1,0)(0,1,1)_{12}$	44
Figura 37. Residuos del modelo $SARIMA(2,1,0)(1,1,0)_{12}$	46
Figura 38. Función de autocorrelación de los residuos del modelo $SARIMA(2,1,0)(1,1,0)_{12}$	46

Figura 39. Función de autocorrelación parcial de los residuos del modelo $SARIMA(2,1,0)(1,1,0)_{12}$	47
Figura 40. Histograma de los residuos del modelo $SARIMA(2,1,0)(1,1,0)_{12}$	48
Figura 41. Grafica Q-Q normal de los residuos del modelo $SARIMA(2,1,0)(1,1,0)_{12}$	48
Figura 42. Residuos del modelo $SARIMA(2,1,0)(1,1,2)_{12}$	50
Figura 43. Función de autocorrelación de los residuos del modelo $SARIMA(2,1,0)(1,1,2)_{12}$	50
Figura 44. Función de autocorrelación parcial de los residuos del modelo $SARIMA(2,1,0)(1,1,2)_{12}$	51
Figura 45. Histograma de los residuos del modelo $SARIMA(2,1,0)(1,1,2)_{12}$	52
Figura 46. Grafica Q-Q normal de los residuos del modelo $SARIMA(2,1,0)(1,1,2)_{12}$	52
Figura 47. Residuos del modelo $SARIMA(3,1,0)(0,1,1)_{12}$	54
Figura 48. Función de autocorrelación de los residuos del modelo $SARIMA(3,1,0)(0,1,1)_{12}$	54
Figura 49. Función de autocorrelación parcial de los residuos del modelo $SARIMA(3,1,0)(0,1,1)_{12}$	55
Figura 50. Histograma de los residuos del modelo $SARIMA(3,1,0)(0,1,1)_{12}$	56
Figura 51. Grafica Q-Q normal de los residuos del modelo $SARIMA(3,1,0)(0,1,1)_{12}$	56
Figura 52. Residuos del modelo $SARIMA(3,1,0)(1,1,0)_{12}$	58
Figura 53. Función de autocorrelación de los residuos del modelo $SARIMA(3,1,0)(1,1,0)_{12}$	58
Figura 54. Función de autocorrelación parcial de los residuos del modelo $SARIMA(3,1,0)(1,1,0)_{12}$	59
Figura 55. Histograma de los residuos del modelo $SARIMA(3,1,0)(1,1,0)_{12}$	60
Figura 56. Grafica Q-Q normal de los residuos del modelo $SARIMA(3,1,0)(1,1,0)_{12}$	60
Figura 57. Residuos del modelo $SARIMA(3,1,0)(1,1,2)_{12}$	62
Figura 58. Función de autocorrelación de los residuos del modelo $SARIMA(3,1,0)(1,1,2)_{12}$	62

Figura 59. Función de autocorrelación parcial de los residuos del modelo $SARIMA(3,1,0)(1,1,2)_{12}$	63
Figura 60. Histograma de los residuos del modelo $SARIMA(3,1,0)(1,1,2)_{12}$	64
Figura 61. Grafica Q-Q normal de los residuos del modelo $SARIMA(3,1,0)(1,1,2)_{12}$	64
Figura 62. Pronósticos con el modelo $SARIMA(0,1,1)(1,1,2)_{12}$ del IGV actividad económica Comercio para el año 2018.....	68
Figura 63. Arquitectura de Redes Neuronales Artificiales (RNA) del IGV actividad económica Comercio, enero 2007 – diciembre 2017	69
Figura 64. Evolución del Error de Generalización de la red neuronal	70
Figura 65. Pronóstico del IGV actividad económica Comercio, mediante redes neuronales, para el año 2018	71
Figura 66. Pronóstico del IGV actividad económica Comercio, mediante el modelo SARIMA y RNA, para el año 2018.....	74

INTRODUCCIÓN

Cabe mencionar que una entidad pública o privada, busca planificar sus recursos, por lo que es fundamental conocer el comportamiento de ciertos acontecimientos que representan dificultad e importancia para las organizaciones, debido a que afectan en forma directa o indirectamente a sus actividades y labores (Villalobos, 2014).

El pronóstico de los ingresos tributarios es de suma importancia para el Estado, más aún si este pronóstico abarca la actividad económica, quien es la principal fuente de donde se recaudan los tributos; por ello, es necesario conocer el comportamiento inmediato o a futuro de estos ingresos para el Estado, así mismo se podrá observar la evolución futura que tendrán, a fin de estimar y conocer la cantidad de recursos internos disponibles y de realizar una evaluación de la repercusión de la política tributaria.

Para el modelo de pronósticos existen una serie de técnicas y metodologías; siendo la más conocida, el Box Jenkins que se requiere del cumplimiento de ciertos supuestos, y hace que se limite en algunos aspectos con respecto a las variables. Otro método de pronóstico son Las Redes Neuronales Artificiales; esta técnica pertenece a las no paramétricas, y no necesita de algún supuesto teórico.

Según la Superintendencia Nacional de Aduanas y Administración Tributaria (SUNAT, 2018), el IGV total se incrementó en 15.9%, viéndose reflejado en el IGV interno (13.9%); además se resalta que los sectores que más incremento han tenido son los sectores de minería, hidrocarburos, agropecuario y comercio, mientras que las devoluciones obtuvieron una reducción de 30.4% respecto a lo reflejado en el año 2017.

Por lo antes mencionado se planteó el siguiente problema: ¿Cuál es el pronóstico de las recaudaciones tributarias de la actividad económica comercio a nivel nacional, mediante la metodología de Box - Jenkins y redes neuronales artificiales para el año 2018?

El objetivo general de este trabajo es estimar el pronóstico de las recaudaciones tributarias de la actividad económica comercio a nivel nacional, mediante la metodología de Box - Jenkins y redes neuronales artificiales para el año 2018.

Para la solución del problema mencionado anteriormente, se describió el comportamiento de los ingresos tributarios, luego se estimó un modelo para el pronóstico de los ingresos tributarios mediante la metodología de Box Jenkins, así mismo se estimó un modelo para el pronóstico de ingresos tributarios mediante la metodología de redes neuronales, también se comparó las estimaciones de las diferentes metodologías para la predicción de los ingresos tributarios de la actividad económica comercio para el año 2018, y finalmente se identificó el mejor modelo de series de tiempo para pronosticar los ingresos tributarios para el año 2018.

Esta investigación es importante ya que los resultados obtenidos serán de gran utilidad para el Estado Peruano, para elaborar el presupuesto general de la república, especialmente para el Poder Ejecutivo. Así mismo dichas estimaciones tendrán incidencia directa en la planificación del presupuesto para el año fiscal 2019.

CAPÍTULO I: DISEÑO TEÓRICO

1.1 Antecedentes

1.1.1 Antecedentes locales

Ayala & Samillan (2019), concluyen que la metodología de Redes Neuronales Artificiales Recurrentes (RNAR) arroja un menor error de pronóstico respecto a la metodología de Box Jenkins, para la serie del precio de compra de cierre mensual de las acciones del Banco de Crédito del Perú en la Bolsa de Valores de Lima, abril de 2005 hasta febrero de 2018.

Vásquez (2017), concluye que el modelo Box Jenkins, permite pronosticar mejor el precio del dólar de la entidad bancaria.

1.1.2 Antecedentes Nacionales

Neira (2017), concluye que la Red Neuronal Artificial presenta la mejor estimación de los alumnos matriculados para el año 2017 en Instituciones Educativas públicas del nivel secundaria a nivel nacional.

De la Cruz & Lazaro (2015), concluyen que la metodología de Box Jenkins comete menor error en el pronóstico, por lo tanto pronostica con mayor precisión.

Cárdenas (2015), concluye que la mejor estimación para la serie Número de peruanos retornantes según año de regreso mediante medio de transporte aéreo, es obtenida mediante la metodología de RNA.

1.2 Series de tiempo.

Gujarati (2009) define que: “Una serie de tiempo es un conjunto de valores observados durante un lapso de tiempo determinado”. (p. 22)

1.2.1 Clasificación de una serie de tiempo.

Estacionaria.

Villavicencio (2018) afirma que: “Tiende a no variar con el tiempo, y presenta media y varianza permanente, es decir, refleja valores de una serie de estación, las cuales tienden a estar rondando en torno al promedio y a una varianza constante”. (p. 2)

No estacionaria.

Villavicencio (2018) define que: “Es una serie con tendencia en el tiempo, es decir, existen cambios en la media y varianza, los cuales determinan crecimiento o decrecimiento a largo plazo, por lo que no se encuentra alrededor de un valor medio”. (p. 2)

1.2.2 Componentes de una serie de tiempo

De la Fuente (2014) refiere que los componentes son: “Tendencia, movimiento regular de la serie a largo plazo; variación cíclica, movimientos a mediano plazo respecto a la tendencia; variación estacional, oscilaciones a corto plazo del periodo regular, variación irregular, son fluctuaciones producidas por factores eventuales”. (p. 2)

1.2.3 Procesos estocásticos

Procesos estocásticos estacionarios.

Gujarati y Porter (2010) manifiestan que: “Son aquellos cuya media y varianza son constantes en el tiempo y si la covarianza entre dos periodos depende sólo de la

distancia entre estos dos periodos, y no del tiempo en el cual se calculó la covarianza”. (p. 740)

Procesos estocásticos no estacionarios.

Gujarati y Porter (2010) precisan que: “Son aquellos cuya distribución de probabilidad varía de forma no constante”. (p. 741)

1.2.4 Ruido blanco

Villalobos (2014) define que: “Es una secuencia de variables aleatorias con esperanza nula, varianzas constantes y covarianzas nulas para diferentes valores de t ”. (p. 16)

1.2.5 Prueba de Dickey-Fuller aumentada (DFA)

Gujarati y Porter (2010) definen que: “Consiste en estimar la siguiente regresión: $\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^m \alpha_i \Delta Y_{t-i} + \varepsilon_t$, donde ε_t es un término de error puro de ruido blanco y donde $\Delta Y_{t-1} = (Y_{t-1} - Y_{t-2})$, $\Delta Y_{t-2} = (Y_{t-2} - Y_{t-3})$, etc”. (p. 757)

1.2.6 Modelos univariantes de series temporales

Proceso autorregresivo AR(p)

Los procesos autorregresivos son, como sugiere su nombre, regresiones sobre sí mismos. Específicamente, un proceso autorregresivo de p -ésimo orden $\{Y_t\}$ satisface la ecuación: (Cryer, 2008, p. 66)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t$$

“Donde Y_t es una combinación lineal de los p valores pasados más recientes entre sí más un término e_t que incorpora todo lo nuevo en la serie en el tiempo t que no se explica por los valores pasados”. (p. 66)

Proceso de medias móviles $MA(q)$

En el caso de que solo un número finito de los pesos θ_i sea distinto de cero, tenemos lo que se llama un proceso de promedio móvil. En este caso, cambiamos un poco la notación y escribimos:

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

Llamamos a esta serie una media móvil de orden q y abreviamos el nombre a $MA(q)$. (Cryer, 2008, p. 57)

Modelos $ARMA(p, q)$

Son modelos mixtos y se representa por medio de la ecuación:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

Entonces, Y_t sigue un modelo de promedio móvil autorregresivo, o $ARMA(p, q)$, donde p y q son las órdenes de las partes promedio autoregresivas y móviles, respectivamente. (Prado, West, 2010, p. 65)

Modelo $ARIMA(p, d, q)$

Se dice que una serie de tiempo $\{Y_t\}$ sigue un modelo de promedio móvil autorregresivo integrado si la diferencia d -ésima $W_t = \nabla^d Y_t$ es un proceso $ARMA$ estacionario. Si $\{W_t\}$ sigue un modelo $ARMA(p, q)$, decimos que $\{Y_t\}$ es un proceso $ARIMA(p, d, q)$. Afortunadamente, a efectos prácticos, normalmente podemos tomar $d = 1$ o como máximo 2. (Cryer, 2008, p. 92)

Guzmán (2014), manifiesta que: “Un modelo $ARIMA(p, d, q)$ responde a la siguiente ecuación:

$$\phi_p(L)\nabla^d Y_t = c + \theta_q(L)\varepsilon_t$$

Donde ∇^d es el operador de diferencias de orden d , definido en general como:

$$\nabla^d z_t = (1 - L)^d z_t = z_t - z_{t-d}$$

(p. 16)

Modelo SARIMA(p, d, q)(P, D, Q)

El modelo de media móvil integrado autorregresivo estacional multiplicativo o modelo SARIMA viene dado por:

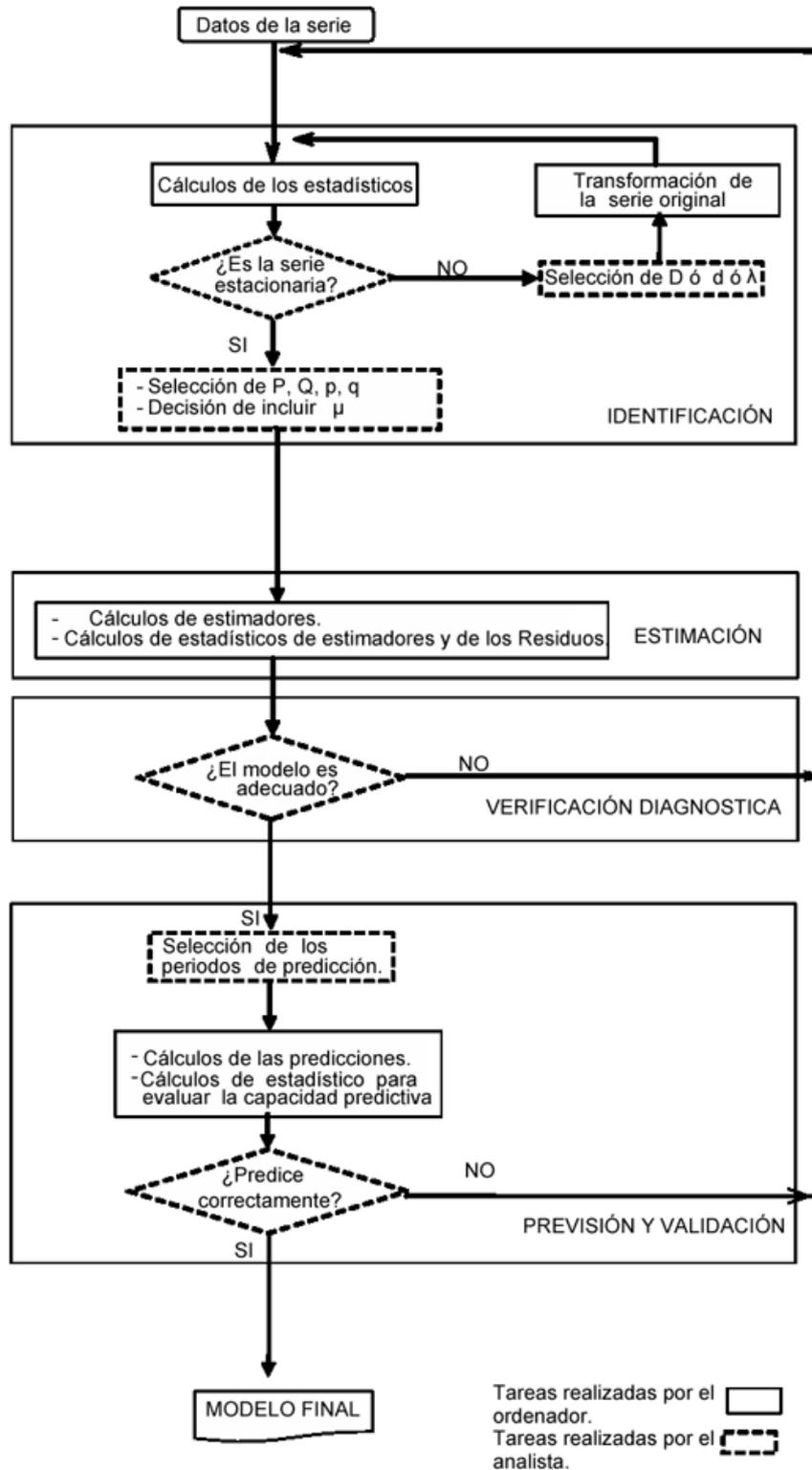
$$\Phi_p(B^S)\phi(B)\nabla_s^D\nabla^d Y_t = \delta + \theta_q(B^S)\theta(B)W_t$$

donde W_t es el proceso habitual de ruido blanco gaussiano. El modelo general se denota como $ARIMA(p, d, q)(P, D, Q)_s$. Los componentes de media móvil y autorregresiva ordinaria están representados por polinomios $\phi(B)$ y $\theta(B)$ de órdenes p y q , respectivamente, y los componentes de media móvil y autorregresiva estacional por $\Phi_p(B^S)$ y $\theta_q(B^S)$ de órdenes P y Q y componentes de diferencia ordinaria y estacional por $\nabla^d = (1 - B)^d$ y $\nabla_s^D = (1 - B^S)^D$. (Shumway y Stoffer, 2011, pp. 152)

1.2.7 Metodología para el modelamiento de series temporales

Figura 1

Esquema del modelado de series de tiempo. (Uriel y Peiró, 2000)



Identificación del modelo.

Uriel y Peiró (2000) precisan: “Primero procede a efectuar un análisis de estacionariedad de la serie. En el caso de que se trate de una serie no estacionaria se aplican las transformaciones adecuadas al objeto de convertirla en estacionaria”. (p. 66)

Así, si se dispone de una serie Y_t , se elegirán unos valores de d y λ de forma que la serie:

$$\varepsilon_t = (1 - L)^d Y_t^{(\lambda)}$$

goce de la propiedad de estacionariedad. Tanto a Y_t como a ε_t se le puede restar el valor de una media si se considera necesario. (Uriel y Peiró ,2000. pp. 65 - 66)

“En la segunda etapa de la fase se procede a determinar el orden de la parte autorregresiva y el orden de la parte de medias móviles del proceso ARMA que se considere haya podido generar a la serie estacionaria”. (pp. 66)

Estimación de los parámetros del modelo.

En esta fase, el objetivo es la obtención de estimadores de los parámetros $\alpha_1, \alpha_2, \dots, \alpha_p, \beta_1, \beta_2, \dots, \beta_q, \mu_\varepsilon$ y σ_ε^2 . Para designar a los parámetros autorregresivos y de medias móviles se utilizará el vector

$$\hat{\beta} = (\alpha_1, \alpha_2, \dots, \alpha_p; \beta_1, \beta_2, \dots, \beta_q)'$$

Los supuestos que se adoptarán en la exposición del proceso de estimación son los siguientes:

- a) $\varepsilon \sim N(0, \sigma_\varepsilon^2 1)$.
- b) El proceso ε_t es estacionario.
- c) El proceso ε_t es invertible.
- d) δ_ε y, en consecuencia, μ_ε son nulos. (Uriel y Peiró, 2000, p. 108)

Para realizar la estimación se dispone de una muestra de tamaño T de la variable Y . Al tomar diferencias de orden d , la variable ε_t constará de $T - d$ elementos. Con objeto de simplificar la notación, a la diferencia $T - d$ se le designará por N , y a los elementos de la muestra de ε_t se les designará por el vector

$$\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N)'$$

Si se adopta el criterio mínimo – cuadrático, los estimadores se obtienen minimizando la expresión:

$$S = \sum \varepsilon_t^2$$

Si se adopta el criterio máximo – verosímil, entonces los estimadores se obtendrán mediante maximización de la función de verosimilitud

$$L = f(\varepsilon_1, \dots, \varepsilon_N)$$

Cualquiera que sea el criterio adoptado, en la estimación de un modelo ARMA se plantea por una parte el problema de determinación de los valores iniciales y por otra parte el problema de que, en general, los modelos de este tipo son no lineales. (Uriel y Peiró, 2000, p. 109)

Diagnóstico del modelo.

Uriel y Peiró (2000) puntualizan: “Un modelo ideal sería si: sus residuos se comportan como ruido blanco, si es estacionario, los coeficientes son estadísticamente significativos y suficientes para representar la serie, el grado de ajuste es elevado en comparación al de otros modelos”. (p. 127).

Prueba de Ljung-Box

Esta prueba, Villavicencio (2007) afirma: “Permite probar en forma conjunta de que todos los coeficientes de autocorrelación son independientes”. (p. 5)

Está definida como:

$$LB = n(n + 2) \sum_{k=1}^m \left(\frac{\hat{\rho}_k^2}{n - k} \right) \sim \chi_{(m)}^2$$

Donde

n = tamaño de la muestra, m longitud del rezago.

k = número de rezagos que se analizan ($k = 12, 24, 36, 48$)”.

Prueba de Kolmogorov- Smirnov

Siegel y Castellan (1998) definen que: “Es una prueba estadística no paramétrica, la cual determina si las puntuaciones en una muestra pueden razonablemente provenir de una población que tiene una distribución teórica”. (p. 75)

Pronóstico.

Hanke y Wichern (2010) precisa: “Una vez que se ha encontrado un modelo adecuado, es factible elaborar los pronósticos de uno o varios periodos futuros. Con base en los pronósticos también se pueden construir intervalos de predicción”. (p. 411)

Es buena idea hacer un seguimiento de los errores de pronóstico, usando los siguientes indicadores:

- Error cuadrático medio (RMSE), Error medio absoluto (MAE) y Error absoluto medio porcentual (MAPE).
- Solo para el caso de la metodología de Box y Jenkins se calcularán los coeficientes de información de Akaike (AIC) y el bayesiano (BIC).

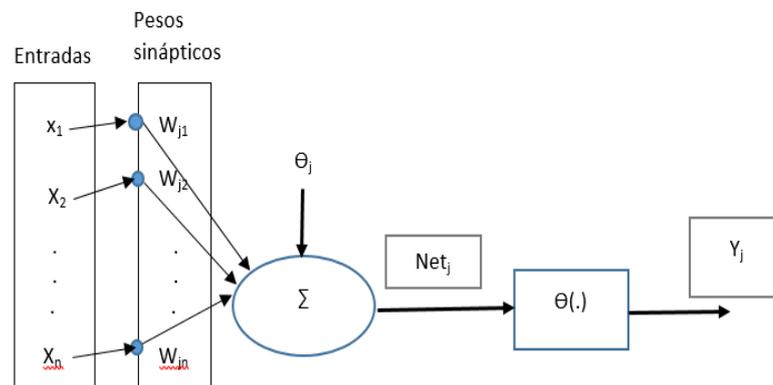
1.3 Redes neuronales.

Salas (2014) define que: “Son un sistema computacional inspirada en el sistema nervioso humano. Su estructura conecta diferentes procesadores elementales que conforman un sistema que posee un algoritmo que ajusta pesos que se requieren en la solución del problema, basándose en muestras representativas”. (p. 3)

De igual modo, Damián (2001) afirma que: “Son redes interconectadas y simplificadas que emulan el modo en que el cerebro humano procesa la información con un orden jerárquico, lo facilitará la interacción con el mundo actual de forma parecida en que lo hace el sistema”. (p. 4)

Figura 2

Estructura de una red neuronal (IBM)



Su arquitectura se caracteriza por:

- Agrupaciones de unidades elementales, las cuales cuentan con una capacidad muy reducida.
- Está interconectada por enlaces ponderados
- Parámetros libres que tienen que ser ajustados con el fin de que satisfagan los requerimientos del desempeño.
- Posee alto grado de paralelismo

1.3.1 Objetivo de las Redes Neuronales

Salas (2014) afirma que “El objetivo de esta técnica es buscar replicar el comportamiento de las neuronas en la naturaleza similar a un cerebro caracterizado por su generalización y su robustez”. (p. 3)

1.3.2 Elementos que componen una red neuronal

Entradas

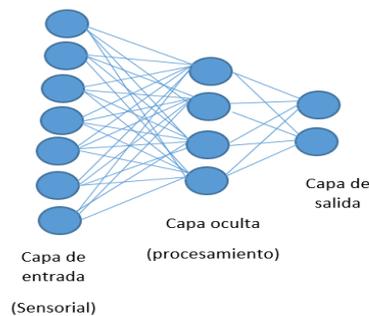
La capa de entrada, “Funciona como un regulador de las entradas, transfiriendo datos a la siguiente capa. Esto puede requerir un pre procesamiento para que la red sólo maneje datos numéricos”. (Cruz, 2007, p. 20)

Salidas

Cruz (2007) manifiesta: “Funciona de una forma similar a la capa de entrada excepto que transfiere la información de la red al exterior”. (p. 20)

Figura 3

Modelo de una red neuronal artificial (IBM)



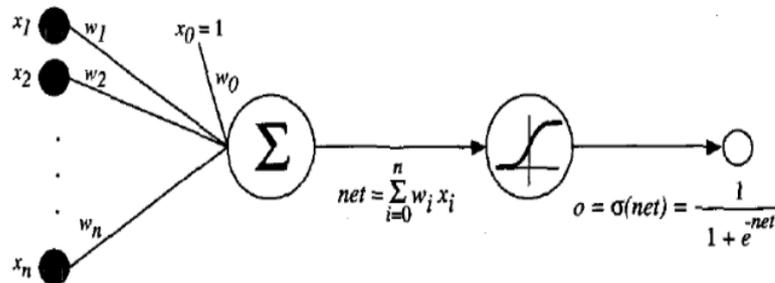
1.3.3 Estructura de una red neuronal

El perceptrón

Kriesel (2007) define que: “La suma ponderada ya discutida y una función de activación no lineal como componentes del perceptrón”. (p. 71)

Figura 4

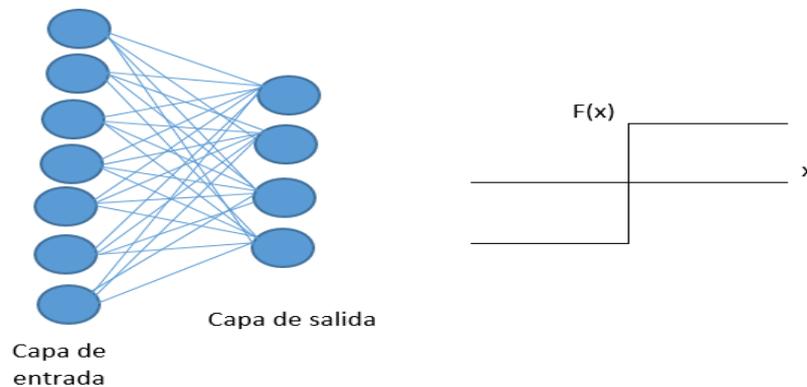
Modelo de una red monocapa



Así mismo afirma que: “No existe una definición establecida para un perceptrón, pero la mayoría de las veces el término percepción se usa para describir una red de retroalimentación con conexiones de acceso directo”. (p. 71)

Figura 5

Arquitectura de un Perceptrón simple. (Damián, 2001)



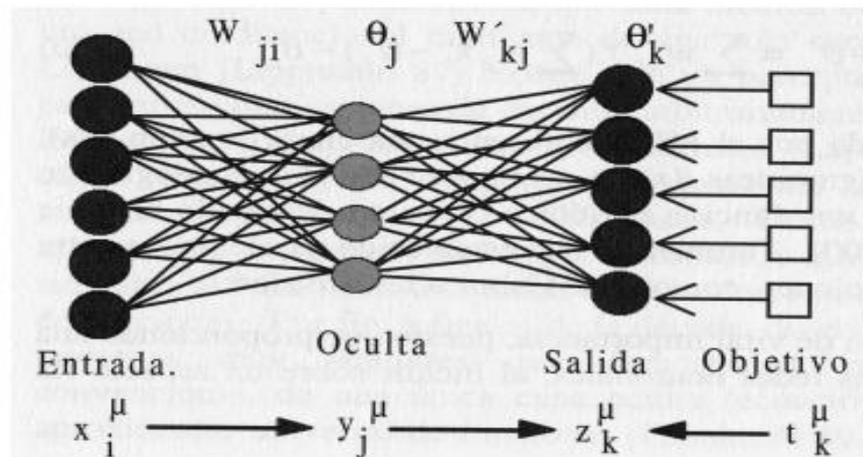
El perceptrón multicapa

Es una modificación del perceptrón lineal estándar porque utiliza tres o más capas de neuronas (nodos) con funciones de activación no lineales y es más potente que el perceptrón; además de ser modelos no lineales, flexibles y de propósito general, que constan de varias unidades organizadas en múltiples capas. La

complejidad de la red MLP se puede cambiar variando el número de capas y el número de unidades en cada capa. (Alsmadi, Omar y Noah, 2009, p. 378)

Figura 6

Arquitectura de un Perceptrón multicapa



1.3.4 Funciones de las Redes Neuronales Artificiales

Función de propagación

Villalobos (2014), afirma que: “Llamada también función de excitación, se origina sumando las señales de entrada, para luego multiplicarlo por el total del peso de su interconexión”. (p.32)

$$h_i(t) = \sum_j w_{ij}x_{ij}$$

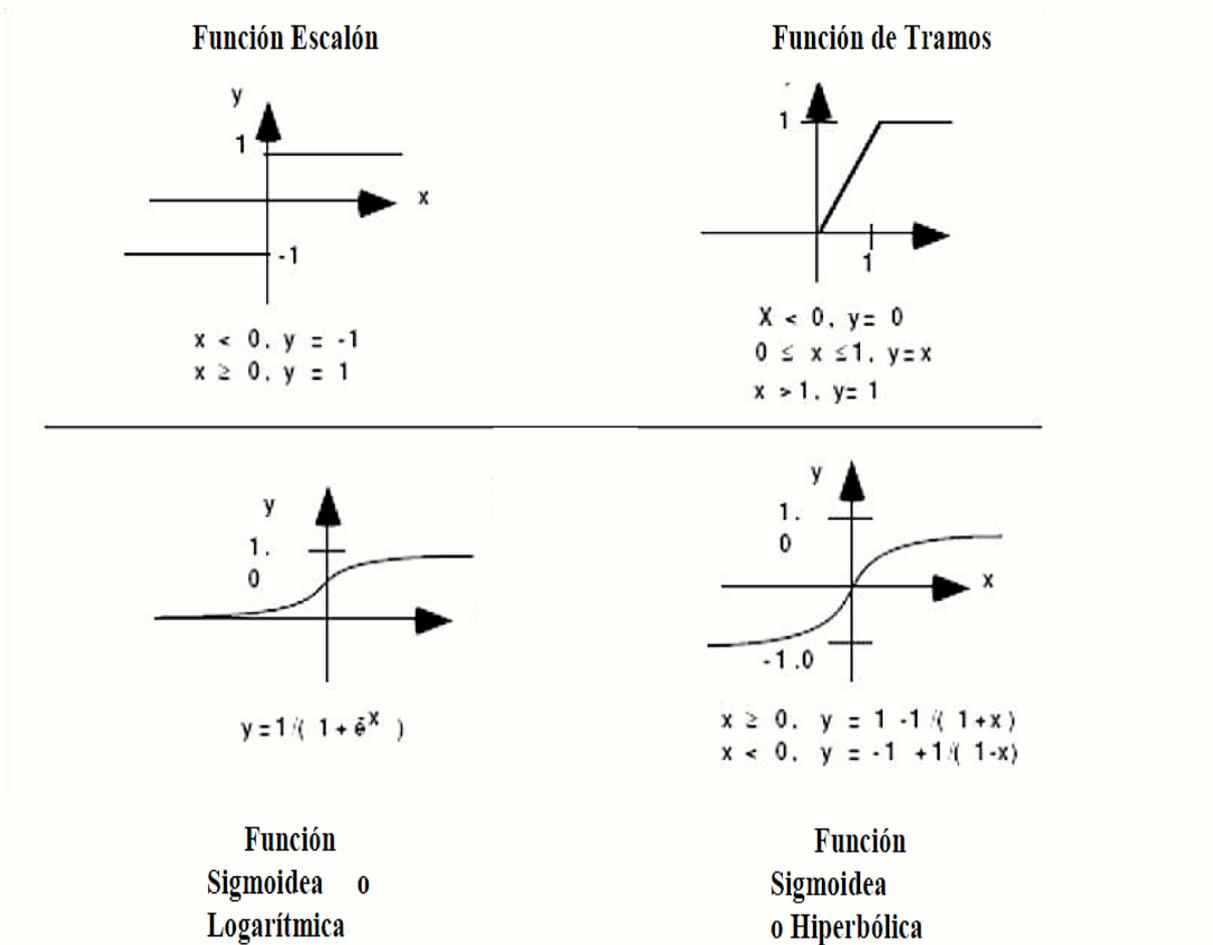
Función de Activación o Función de transferencia

Es considerada como un elemento de topología con el fin de ayudar al cálculo de la situación de la actividad en una neurona artificial, en donde los rangos más empleados son la función sigmoidea (para obtener valores de (0 a 1), por el cual debe estar dependiendo de su estado activo o inactivo. (Villalobos, 2014, p. 33).

$$a_i(t) = f_i(n_it)$$

Figura 7

Función de activación

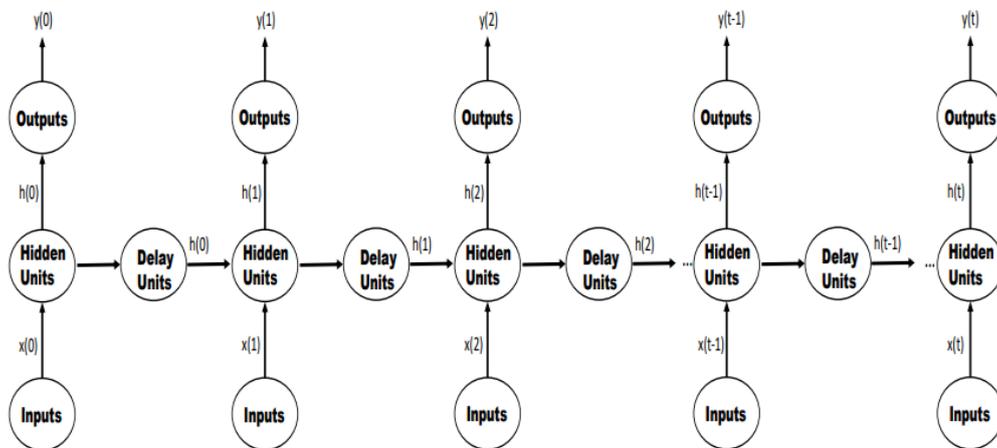


Backpropagation.

Lewis (2017) manifiesta que: “Cualquier red neuronal de alimentación hacia adelante La red puede ser entrenada usando el algoritmo de propagación hacia atrás. La idea básica es desplegar el RNR a través del tiempo, esto suena complicado, pero el concepto es bastante simple”. (p. 37).

Figura 8

Backpropagation Network. (Lewis, 2017)



1.3.5 Métodos de aprendizaje

Khare y Nagendra (2007), definen que un proceso de "aprendizaje", o "formación", forma las interconexiones (correlaciones) entre las neuronas. Se logra proporcionando datos de entrada y salida conocidos de una manera ordenada a la red neuronal. (p. 35)

Redes con aprendizaje supervisado

Granados (2010) puntualiza que: "Necesitan un conjunto de datos de entrada previamente clasificados; es decir la respuesta objetivo se conoce. Ejemplos de este tipo de redes son el perceptrón (simple y multicapa), la red Adaline y la memoria asociativa bidireccional". (p. 47)

Redes con aprendizaje no supervisado

Granados (2010) afirma que: "No necesitan de tal conjunto previo. Ejemplos de estas redes son las memorias asociativas, redes de Hopfield; máquina de Boltzmann y las redes de Kohonen". (p. 47)

1.3.6 Redes neuronales aplicadas a las series de tiempo

Para predecir series de tiempo se han diseñado las redes neuronales alimentadas hacia delante (feed-forward) y las redes recurrentes (recurrent). Dentro de las redes alimentadas hacia adelante existen algoritmos como: Linear-Associator, Backpropagation, Cascade Correlation, Cascade 2, RAN (Resource Allocating Network). Entre las redes recurrentes se tiene: Recurrent Cascade Correlation, Simple Recurrent Networks, Real Time Recurrent Learning, Sequential Cascaded Network. (Granados, 2010, p. 52)

Modelo NAR caso univariado.

Ortiz (2012) manifiesta que: “Es un modelo neuronal autorregresivo no lineal, basado en una única serie de datos; que predice valores de la serie utilizando las observaciones pasadas. Es considerada una de las mejores para analizar y predecir información de series de tiempo”. (p. 87).

CAPÍTULO II. MÉTODOS Y MATERIALES

2.1 Tipo y diseño de investigación

El diseño usado en esta investigación fue no experimental **retrospectivo-proyectista de**, tipo descriptivo, predictivo y longitudinal.

2.2 Población y muestra

Población

Estuvo conformada por el registro de los ingresos tributarios de Impuesto General a las ventas (IGV) interno, obtenido de la SUNAT.

Muestra

Estuvo compuesta por los ingresos Tributarios recaudados del IGV internas, por la SUNAT en el periodo enero 2007 – diciembre del 2017. La cuál estuvo conformada por 132 datos de IGV.

2.3 Técnica e instrumento de recolección de datos

Se utilizó la técnica del análisis documental y como instrumento, el registro automatizado de los ingresos tributarios recaudados del Impuesto general de las ventas internas por mes y año.

2.4 Análisis y procesamiento estadístico de los datos

Se siguió los pasos de la metodología de Box y Jenkins; mediante el cual se identificó el modelo, se estimó los parámetros del modelo seleccionado, realizar el diagnostico correspondiente y finalmente realizar el pronóstico.

En la identificación se presentó el gráfico de líneas de serie, las autocorrelaciones simple y parcial, así mismo se utilizó la prueba de Dickey – Fuller para probar que la serie es estacionaria.

En el diagnóstico, para probar que los residuos sean una variable de ruido blanco y además que los errores provengan de una distribución normal, se utilizó las pruebas de Ljung-Box y Kolmogorov Smirnov respectivamente con un nivel de confianza del 95%.

En la etapa de estimación, se obtuvieron los posibles modelos que mejor se ajustan a la serie en estudio.

Finalmente se calcularon los pronósticos de IGV según actividad económica para el año 2018.

Así mismo se aplicó la metodología de redes neuronales, la cual consistió en entrenar, validar, probar y construir una red neuronal para obtener los pronósticos correspondientes.

Finalmente se realizó el análisis comparativo de los pronósticos obtenidos mediante los métodos de Box y Jenkins y redes neuronales de la serie utilizada.

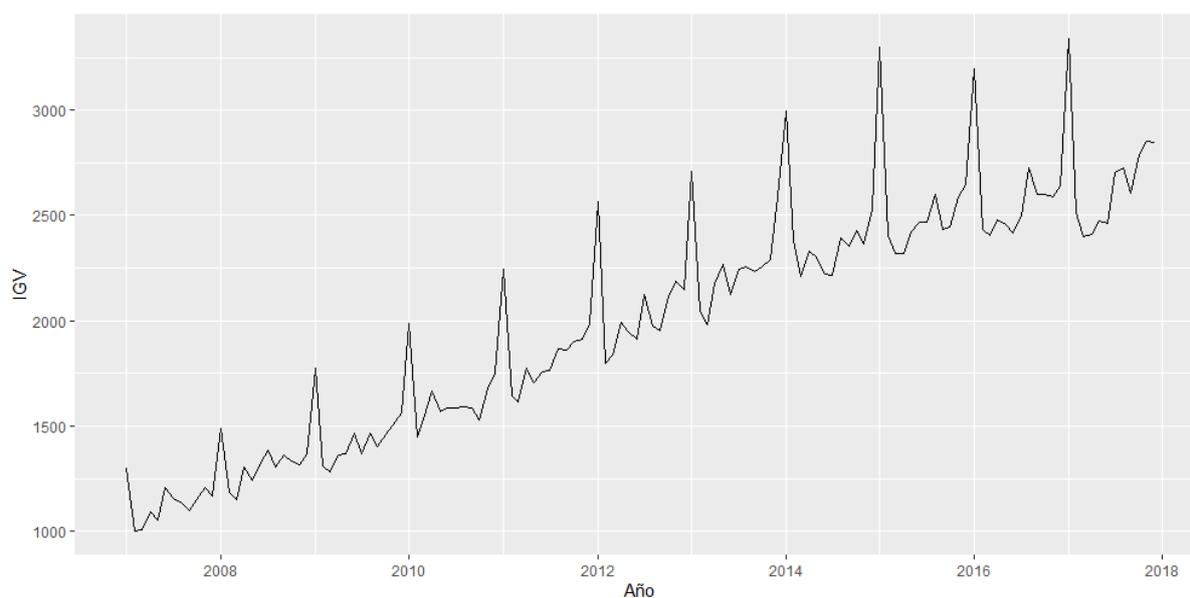
Para el procesamiento y análisis de hizo uso del software estadístico RStudio.

CAPÍTULO III. RESULTADOS Y DISCUSIÓN

3.1 Descripción del comportamiento de los ingresos tributarios

Figura 9

Comportamiento de la recaudación del IGV según actividad económica Comercio, enero 2007 a diciembre 2017



En la figura 9, se muestra la tendencia creciente, así como la variabilidad a través del tiempo, por lo que se puede afirmar que la serie no es estacionaria.

Tabla 1

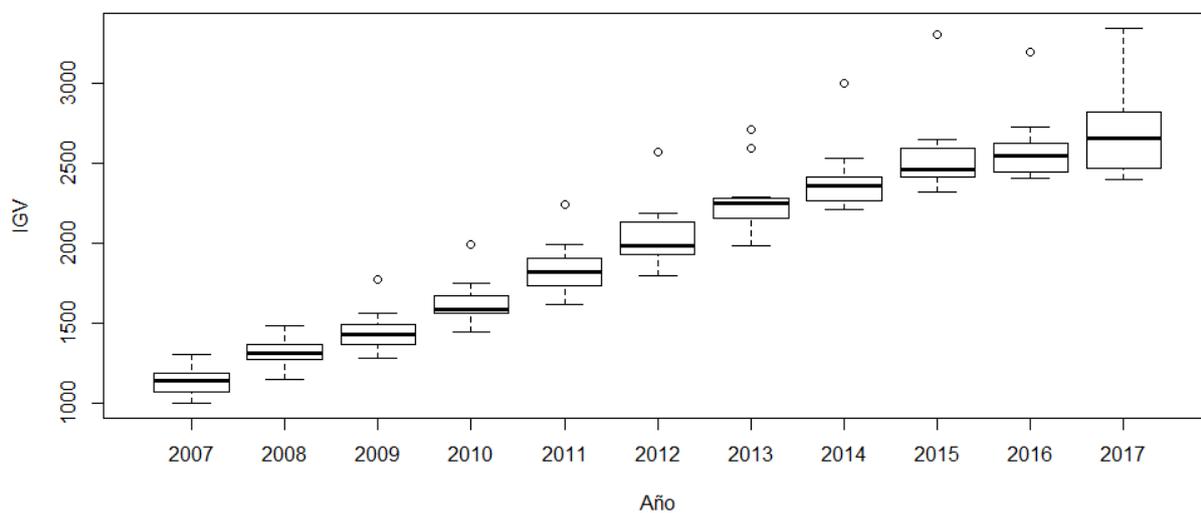
Estadísticos descriptivos del IGV según actividad económica Comercio, enero 2007 a diciembre 2017

Estadístico	Valor
Promedio	1986.44
Mínimo	1003.15
Máximo	3337.98

En la tabla 1 se aprecian los estadísticos descriptivos como el mínimo ingreso tributario, el cual es 1003.15 millones de soles, y un máximo ingreso de 3337.98 millones de soles. Se encontró también que el promedio es igual a 1986,44 millones de soles.

Figura 10

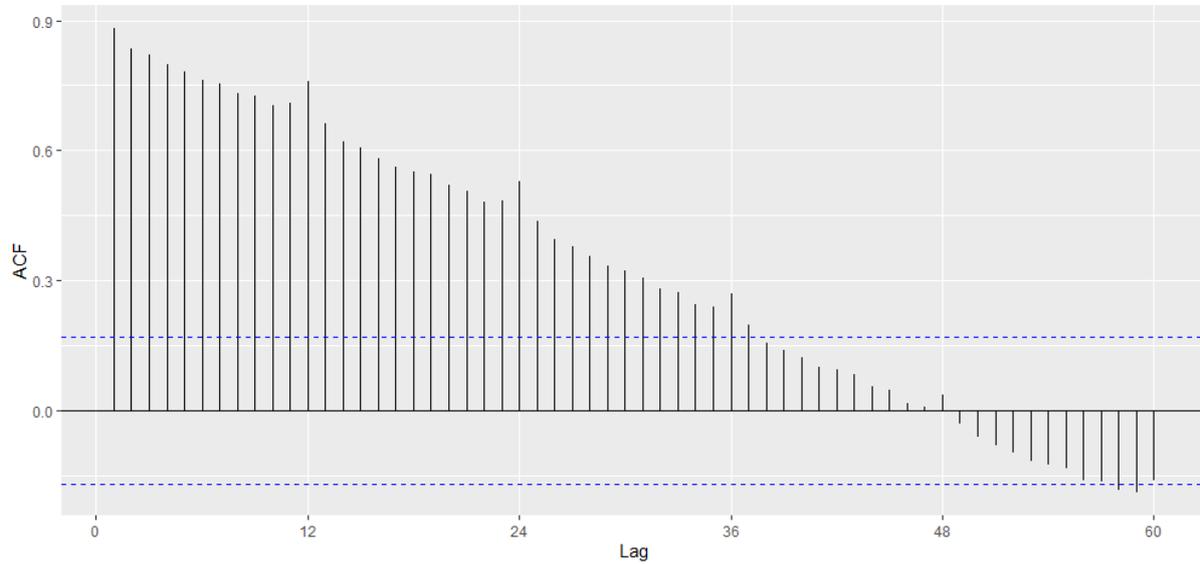
Diagrama de cajas del IGV según actividad económica Comercio, enero 2007 a diciembre 2017



En la figura 10 se presenta el diagrama de cajas, en la cual se puede visualizar mejor la variabilidad por año de los ingresos tributarios.

Figura 11

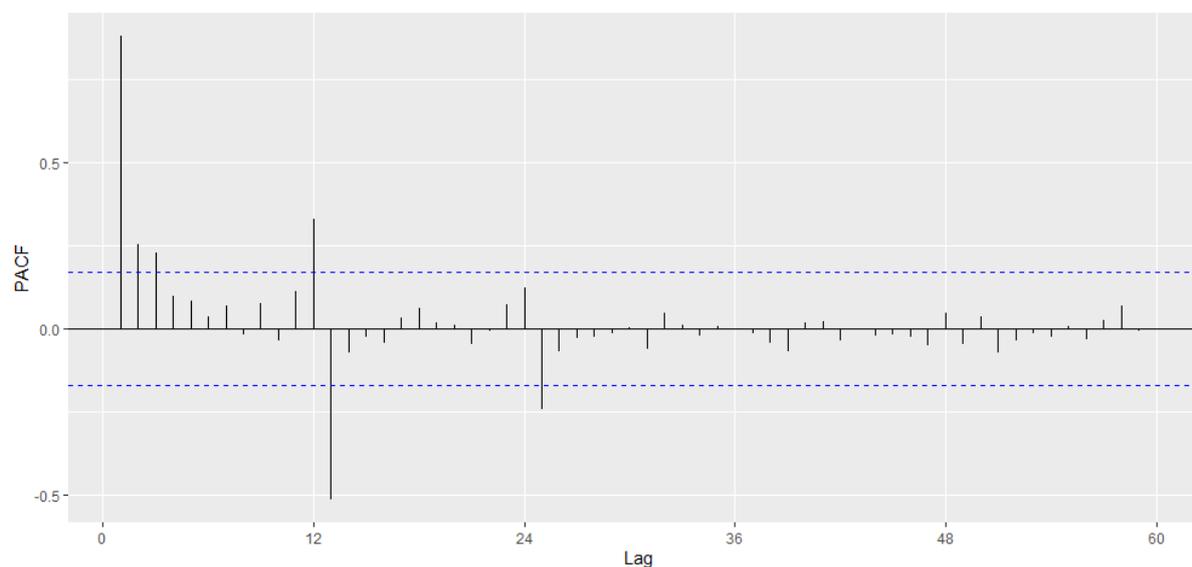
Función de Autocorrelación simple a la recaudación IGV según actividad económica Comercio



En la figura 11 se aprecia la función de autocorrelación simple aplicada a la serie actividad económica comercio, en donde se aprecia que la serie presenta tendencia, por lo tanto la serie no es estacionaria.

Figura 12

Función de autocorrelación parcial de la recaudación del IGV según actividad económica Comercio, enero 2007 – diciembre 2017



En la figura 12 se aprecia la función de autocorrelación parcial, la cual muestra que la actividad económica comercio durante los periodos de estudios mencionados, presenta decaimiento brusco, además los retardos son significativos.

Tabla 2

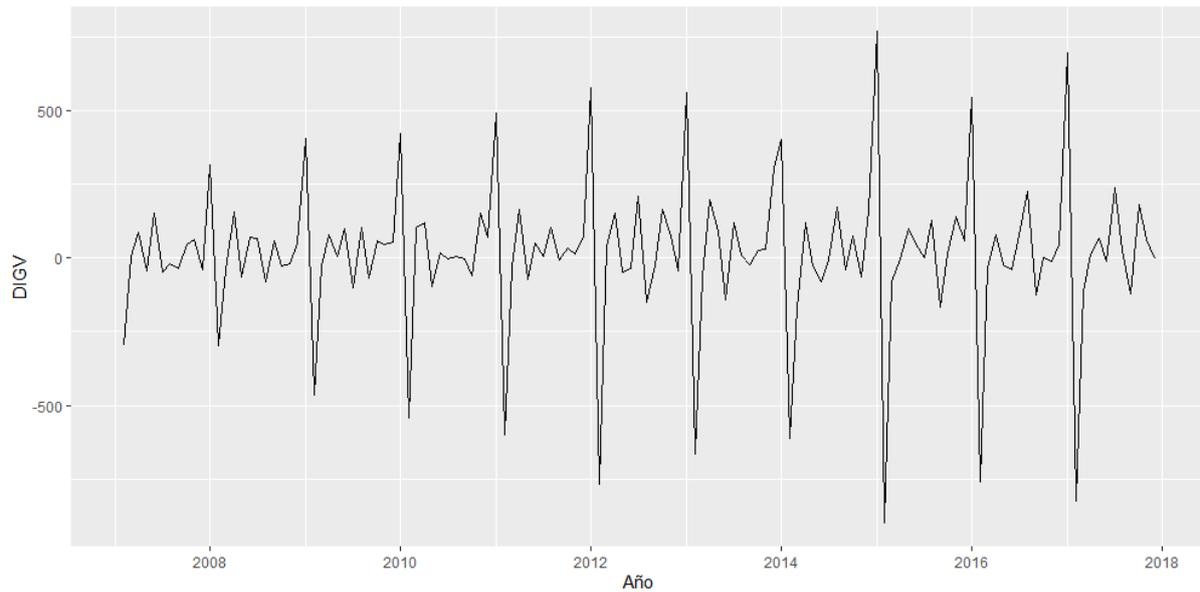
Test de Dickey-Fuller para IGV actividad económica Comercio, enero 2007 – diciembre 2017

Dickey-Fuller	Lag order	p-value
-0.954	5	0.702

En la tabla 2 se aprecia el valor del estadístico Dickey – Fuller, el cual es -0.954, y un valor $p = 0.702$, lo que permite corroborar que no existe presencia de raíz unitaria, por lo que la serie debe ser diferenciada.

Figura 13

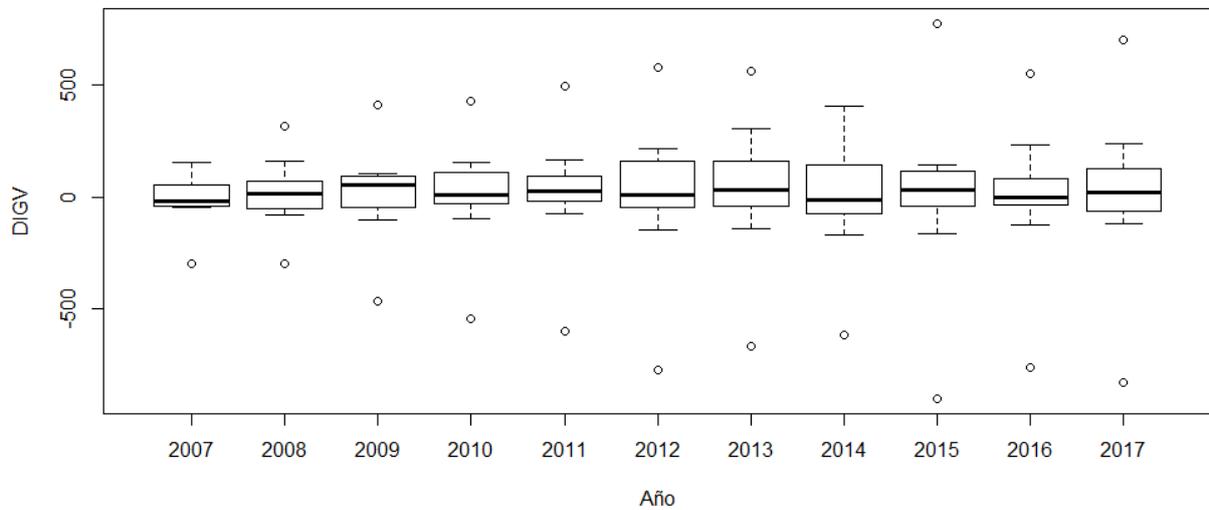
Primera diferencia aplicada al IGV actividad económica Comercio, enero 2007 – diciembre 2017



En la figura 13 se muestra la serie diferenciada en un periodo al IGV; se aprecia que no cuenta con tendencia, por lo cual se observa que la serie se estabiliza respecto a su media, además presenta varianza constante, por lo que se puede deducir que la serie es estacionaria.

Figura 14

Grafica de cajas de la primera diferencia al IGV según actividad económica Comercio, enero 2007 a diciembre 2017



Como se puede apreciar en la figura 14, después de aplicar una diferencia en un periodo al IGV según actividad económica Comercio, se puede apreciar que las medianas se estabilizan; se puede decir que la serie ya es estacionaria.

Figura 15

Función de autocorrelación simple de la primera diferencia del IGV actividad económica Comercio, enero 2007 – diciembre 2017

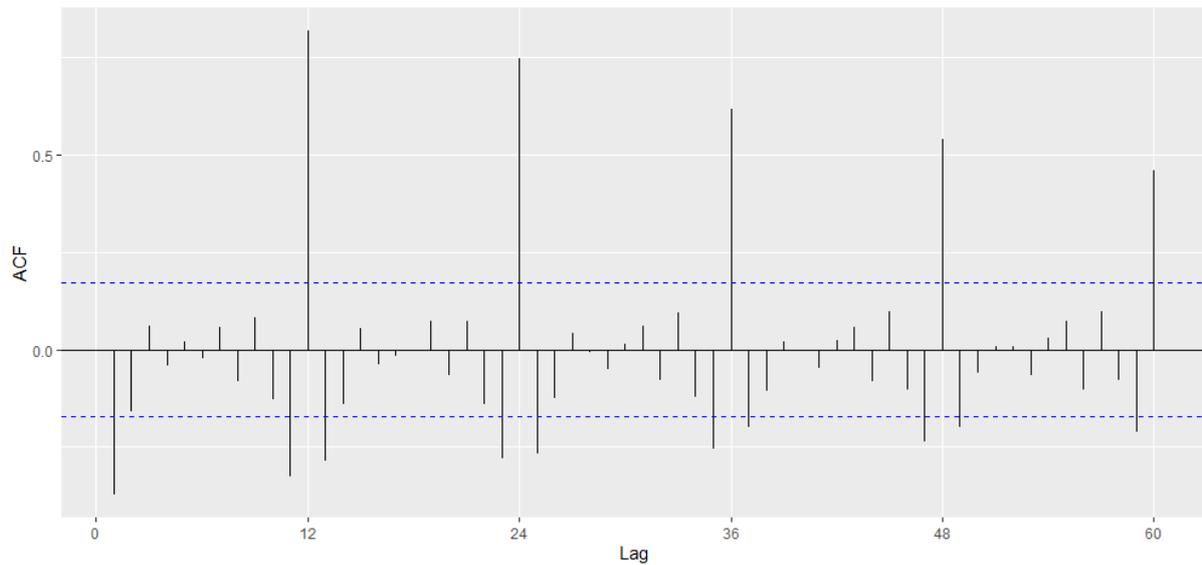
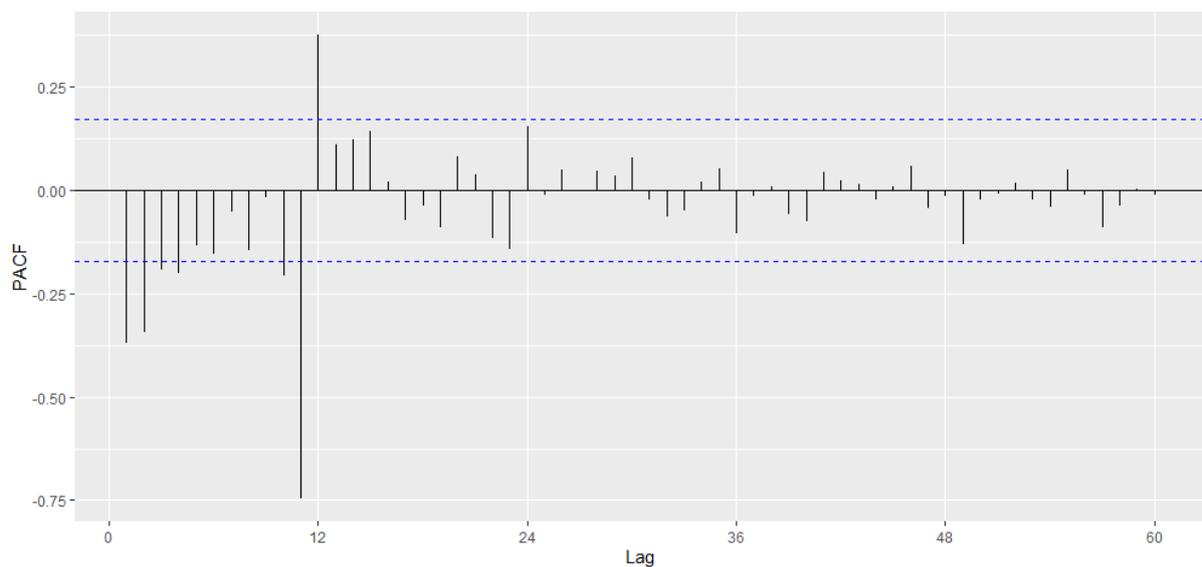


Figura 16

Función de autocorrelación parcial de la primera diferencia del IGV actividad económica Comercio, enero 2007 – diciembre 2017



En la función de autocorrelación simple de la serie en primera diferencia se observa que los retardos tienen una estructura dispersa, es decir existen retardos que están fuera de la banda de confianza; aun así la serie aparentemente es estacionaria.

Así mismo, en la función de autocorrelación parcial de la serie en primera diferencia también se presentan 2 retardos significativos.

Tabla 3

Test de Dickey-Fuller de la primera diferencia al IGV actividad económica Comercio, enero 2007 – diciembre 2017

Dickey-Fuller	Lag order	p-value
-7.668	5	0.01

En esta tabla 3 se aprecia el valor del estadístico Dickey – Fuller, el cual es -7.668, y un valor $p = 0.01$, por lo que se concluye que existe presencia de raíz unitaria, lo que permite establecer que la serie diferenciada es estacionaria.

3.2 Estimación del modelo para el pronóstico de los ingresos tributarios, mediante la metodología de Box y Jenkins.

Tabla 4

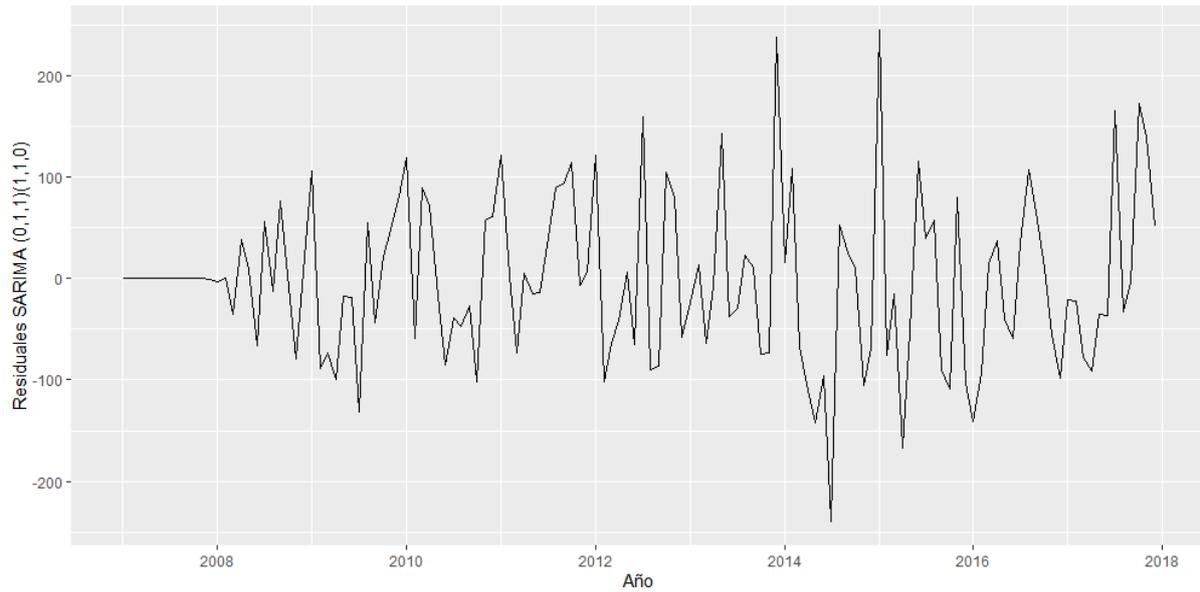
Modelos estimados de la serie de tiempo IGV actividad económica Comercio

MODELO	PARAMETRO	ESTIMADO	Pr (> z)
SARIMA (0,1,1) (1,1,0)	ma1	-0.763458	2.20E-16
	sar1	-0.438753	1.27E-07
SARIMA (1,1,0) (1,1,0)	ar1	-0.393458	2.92E-06
	sar1	-0.483305	1.61E-09
SARIMA (0,1,1) (1,1,2)	ma1	-0.72442	2.20E-16
	sar1	0.8345	3.49E-04
	sma1	-1.38674	4.41E-08
	sma2	0.56307	1.51E-05
SARIMA (2,1,0) (0,1,1)	ar1	-0.494023	2.08E-08
	ar2	-0.304556	4.86E-04
	sma1	-0.465613	4.11E-10
SARIMA (2,1,0) (1,1,0)	ar1	-0.514937	4.11E-09
	ar2	-0.30253	5.14E-04
	sar1	-0.473525	3.14E-09
SARIMA (2,1,0) (1,1,2)	ar1	-0.515464	1.22E-08
	ar2	-0.320141	2.37E-04
	sar1	0.805315	9.94E-05
	sma1	-1.404788	6.76E-10
	sma2	0.608149	2.60E-06
SARIMA (3,1,0) (0,1,1)	ar1	-0.572347	1.78E-10
	ar2	-0.432807	6.98E-06
	ar3	-0.252302	5.25E-03
	sma1	-0.439332	1.94E-08
SARIMA (3,1,0) (1,1,0)	ar1	-0.582247	7.69E-11
	ar2	-0.422169	1.55E-05
	ar3	-0.227244	1.41E-02
	sar1	-0.427424	3.91E-07
SARIMA (3,1,0) (1,1,2)	ar1	-0.580731	2.06E-10
	ar2	-0.438376	5.82E-06
	ar3	-0.232245	1.15E-02
	sar1	0.88283	4.05E-05
	sma1	-1.45087	9.20E-09
	sma2	0.59269	4.48E-06

MODELO SARIMA (0,1,1) (1,1,0)₁₂

Figura 17

Residuos del modelo SARIMA (0,1,1)(1,1,0)₁₂



En la figura 17 se presentan los residuos del modelo $SARIMA (0,1,1)(1,1,0)_{12}$, los mismos que siguen un ruido blanco.

Figura 18

Función de autocorrelación de los residuos del modelo SARIMA (0,1,1)(1,1,0)₁₂

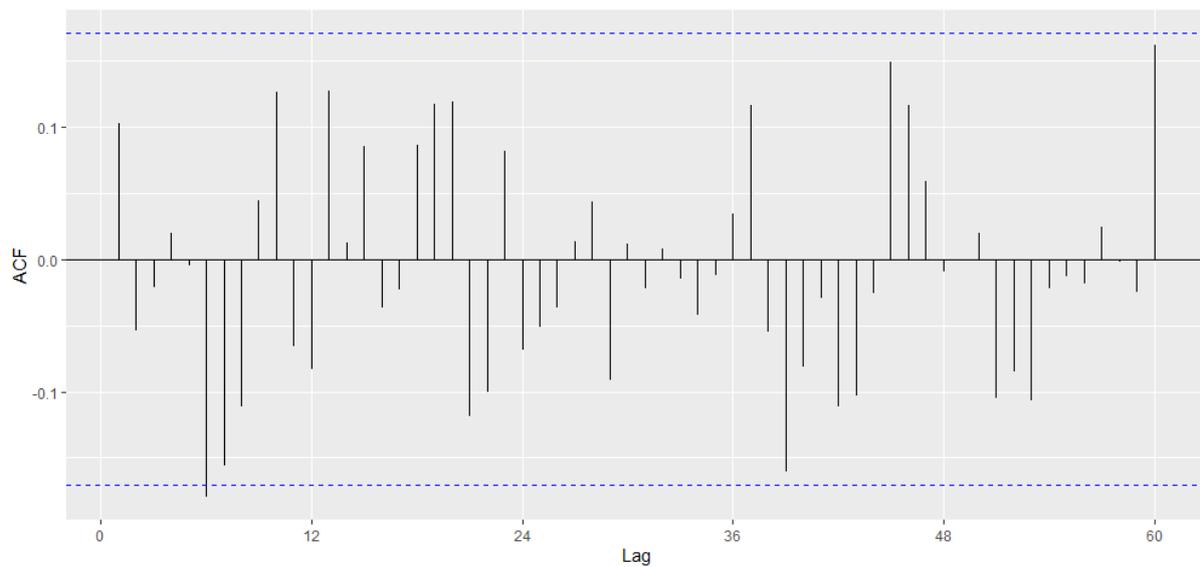
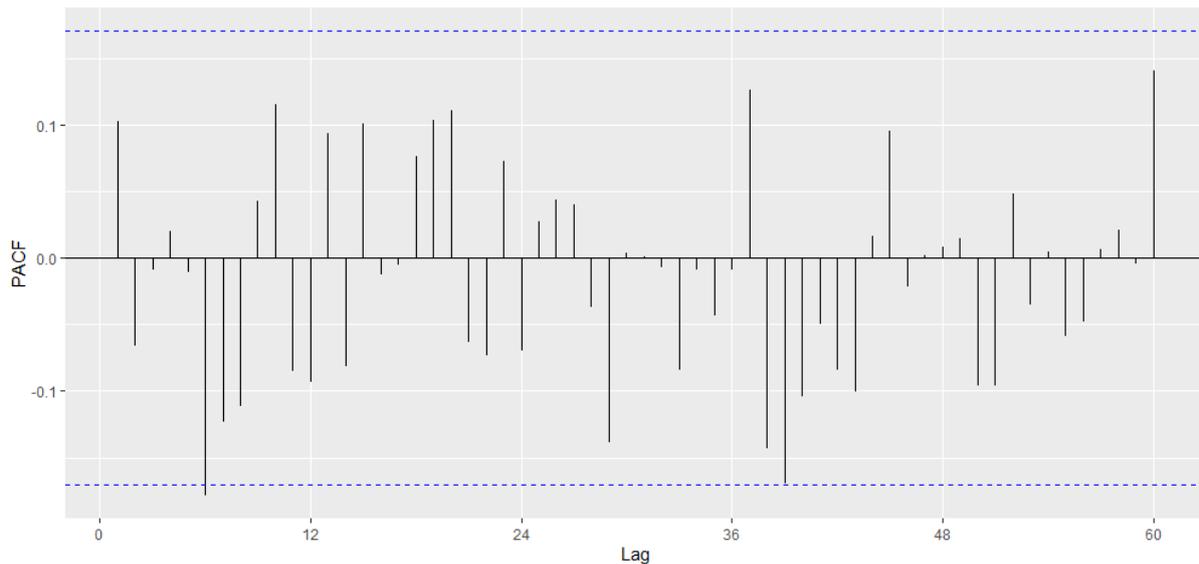


Figura 19

Función de autocorrelación parcial de los residuos del modelo SARIMA (0,1,1)(1,1,0)₁₂



En las funciones de autocorrelación, simple y parcial de los residuos del modelo $SARIMA (0,1,1)(1,1,0)_{12}$, se observa que existe 1 retardo que sobrepasa los límites de confianza, aun así se deduce que los residuos cumplirían con las condiciones de ruido blanco.

Tabla 5

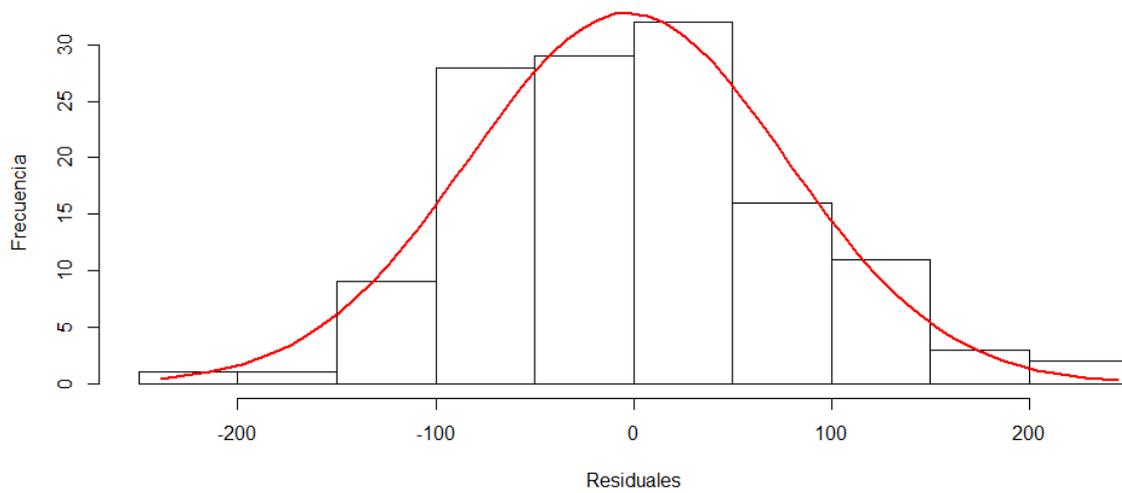
Test de Ljung-Box de los residuos del modelo SARIMA (0,1,1)(1,1,0)₁₂

Chi-Cuadrado	GL	P-Valor
1.433	1	0.231

En la tabla 5 se observa el test de Ljung-Box, el cual muestra el valor chi-cuadrado de 1.433 y un valor $p = 0.231$, por lo que se determinó que los coeficientes de autocorrelación son simultáneamente igual a cero, lo que permite concluir que los residuos del modelo $SARIMA (0,1,1)(1,1,0)_{12}$ se distribuyen de forma independiente.

Figura 20

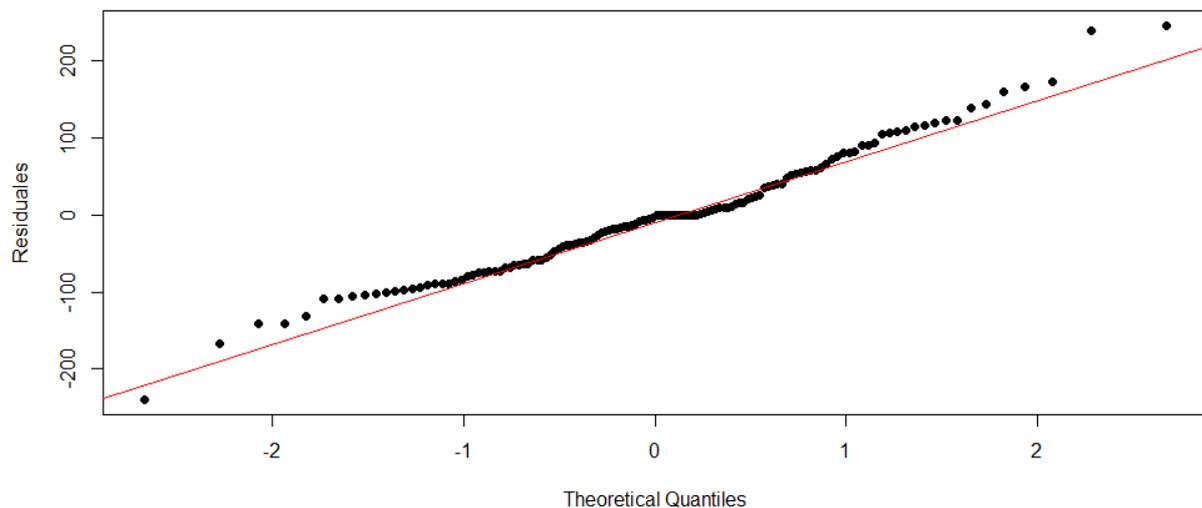
Histograma de los residuos del modelo SARIMA (0,1,1)(1,1,0)₁₂



En el histograma se observa que los residuos del modelo $SARIMA (0,1,1)(1,1,0)_{12}$, tienden a seguir una distribución normal.

Figura 21

Gráfica Q-Q normal de los residuos del modelo SARIMA (0,1,1)(1,1,0)₁₂



En el gráfico de probabilidad normal (gráfico Q-Q) se muestra que los residuos del modelo $SARIMA (0,1,1)(1,1,0)_{12}$, tienden a seguir una línea recta, por lo cual aparentemente siguen una distribución normal.

Tabla 6

Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA (0,1,1)(1,1,0)₁₂

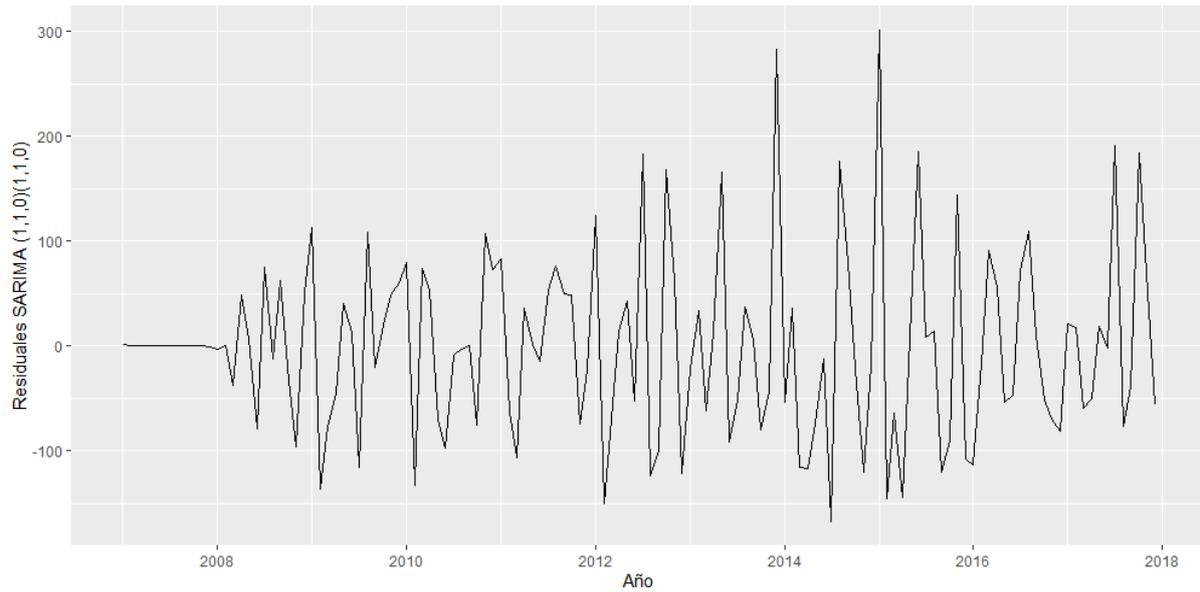
Prueba de normalidad	Kolmogorov-Smirnov	
	Estadístico	Sig
Residuales	0.092	0.213

En la tabla 6 se aprecia el estadístico de Kolmogorov - Smirnov, el cual arroja un valor de 0.092 y $p = 0.213$, por lo que se concluye que los residuos del modelo *SARIMA (0,1,1)(1,1,0)₁₂* siguen una distribución normal.

MODELO SARIMA (1,1,0) (1,1,0)₁₂

Figura 22

Residuos del modelo SARIMA (1,1,0)(1,1,0)₁₂



En la figura 22 se presentan los residuos del modelo $SARIMA (1,1,0)(1,1,0)_{12}$, los mismos que siguen un ruido blanco.

Figura 23

Función de autocorrelación de los residuos del modelo SARIMA (1,1,0)(1,1,0)₁₂

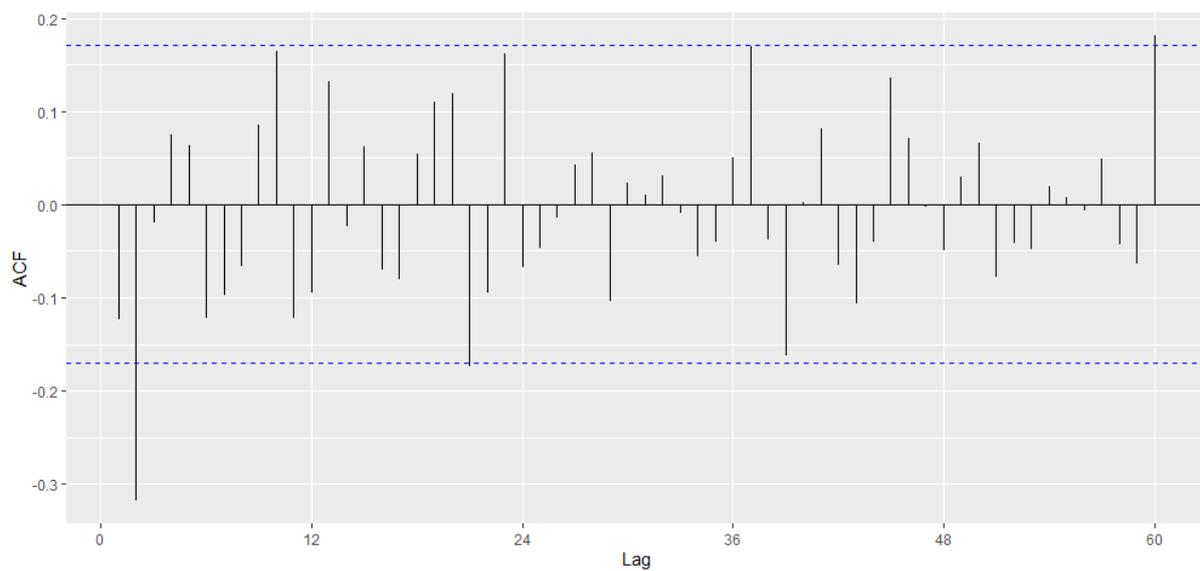
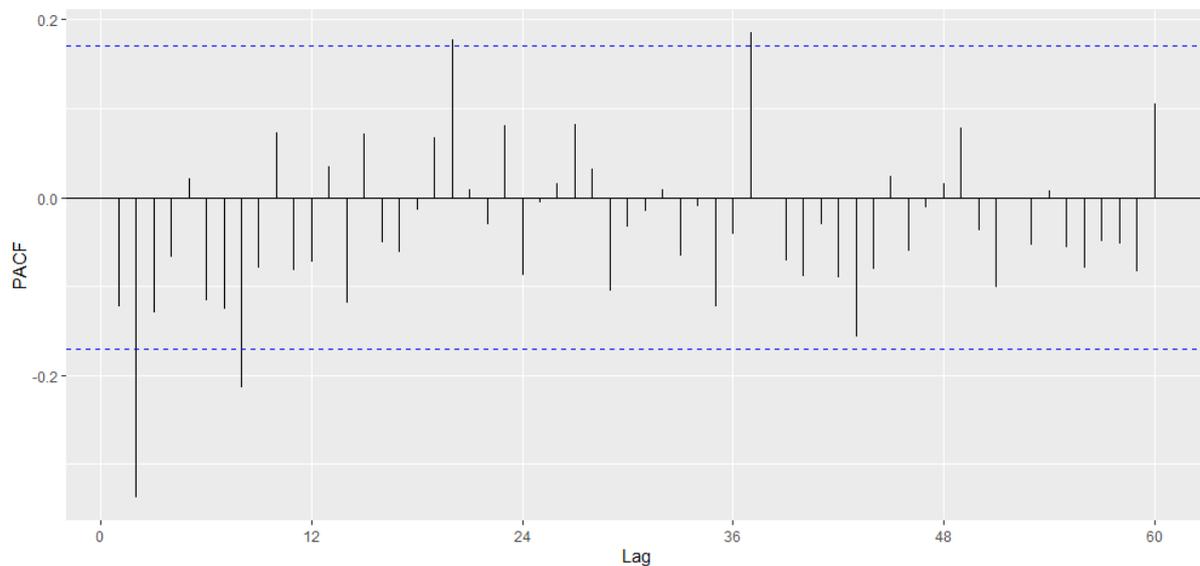


Figura 24

Función de autocorrelación parcial de los residuos del modelo SARIMA (1,1,0)(1,1,0)₁₂



En las funciones de autocorrelación, tanto simple como parcial de los residuos del modelo *SARIMA (1,1,0)(1,1,0)₁₂*, se observan retardos que sobrepasan los límites de confianza, pero aun así los residuos cumplirían con las condiciones de ruido blanco.

Tabla 7

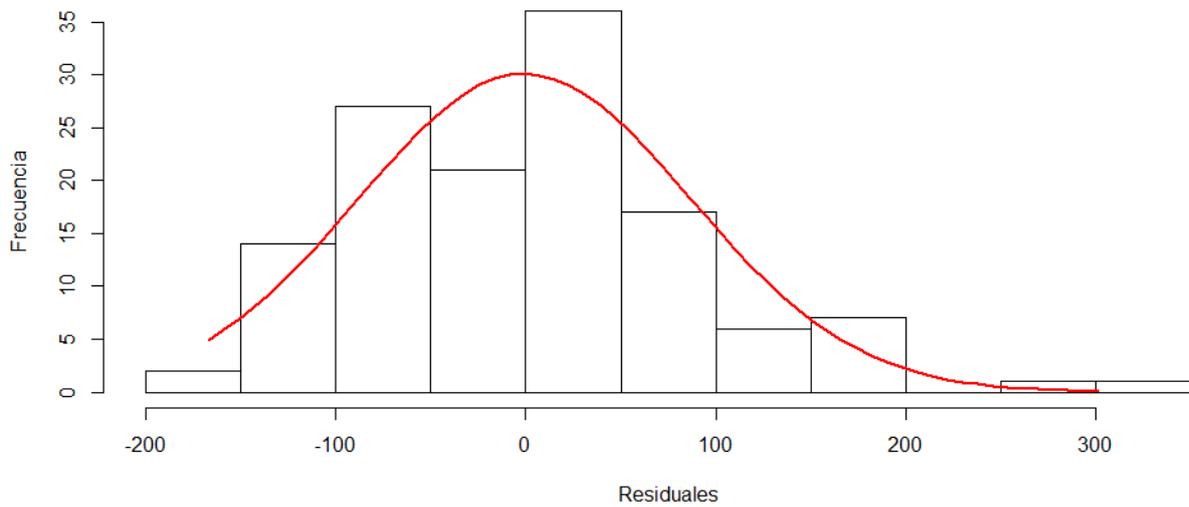
Test de Ljung-Box de los residuos del modelo SARIMA (1,1,0)(1,1,0)₁₂

Chi-Cuadrado	GL	P-Valor
2.037	1	0.154

En la tabla 7 se observa el test de Ljung-Box, el cual arroja un valor chi-cuadrado de 2.037 y $p = 0.093$, por lo se determinó que los coeficientes de autocorrelación son igual a cero, lo que permite concluir que los residuos del modelo *SARIMA (1,1,0)(1,1,0)₁₂* se distribuyen de forma independiente.

Figura 25

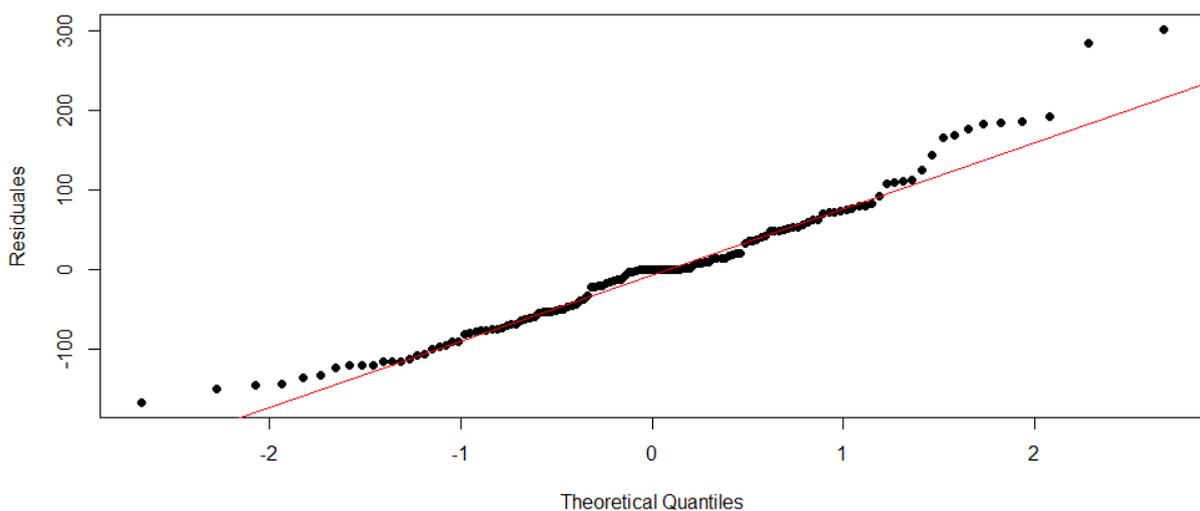
Histograma de los residuos del modelo SARIMA (1,1,0)(1,1,0)₁₂



En el histograma se observa que los residuos del modelo *SARIMA* (1,1,0)(1,1,0)₁₂, no tienden a seguir una distribución normal.

Figura 26

Grafica Q-Q normal de los residuos del modelo SARIMA (1,1,0)(1,1,0)₁₂



En el gráfico de probabilidad normal (gráfico Q-Q) se muestra que los residuos del modelo *SARIMA* (1,1,0)(1,1,0)₁₂, tienden a seguir una línea recta, por lo cual aparentemente siguen una distribución normal.

Tabla 8

Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA (1,1,0)(1,1,0)₁₂

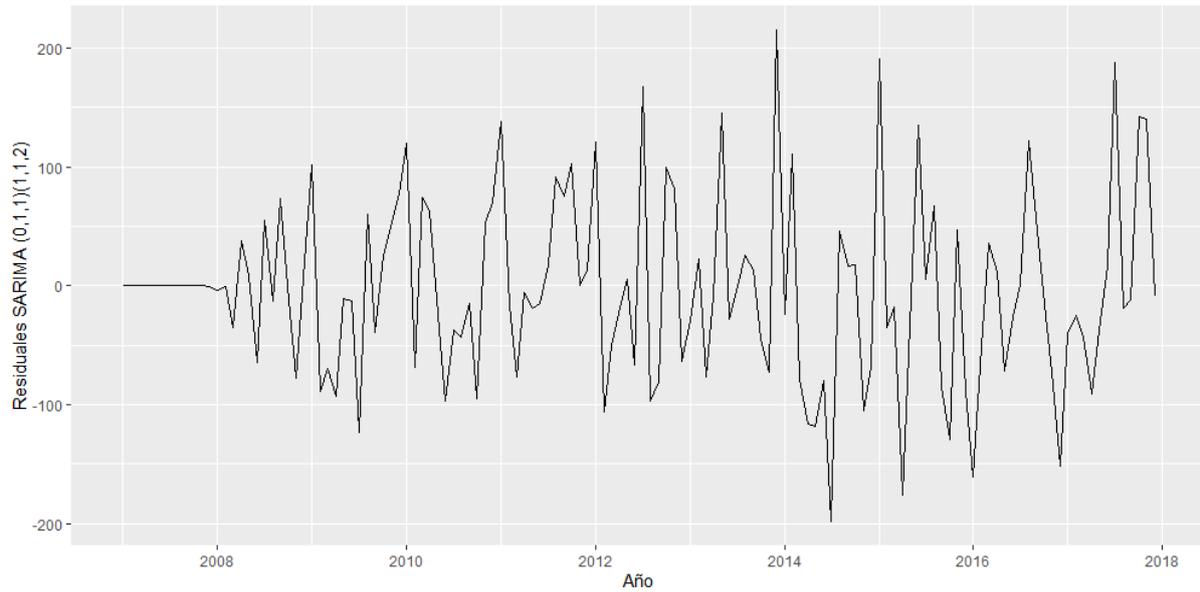
Prueba de normalidad	Kolmogorov-Smirnov	
	Estadístico	Sig
Residuales	0.087	0.268

En la tabla 8 se aprecia el estadístico de Kolmogorov - Smirnov, el cual arroja un valor de 0.087 y un valor $p = 0.268$, por lo que se determinó que las perturbaciones tienen una distribución conjunta normal, concluyendo que los residuos del modelo SARIMA (1,1,0)(1,1,0)₁₂ siguen una distribución normal.

MODELO SARIMA (0,1,1) (1,1,2)₁₂

Figura 27

Residuos del modelo SARIMA (0,1,1)(1,1,2)₁₂



En la figura 27 se presentan los residuos del modelo $SARIMA (0,1,1)(1,1,2)_{12}$, los mismos que siguen un ruido blanco.

Figura 28

Función de autocorrelación de los residuos del modelo SARIMA (0,1,1)(1,1,2)₁₂

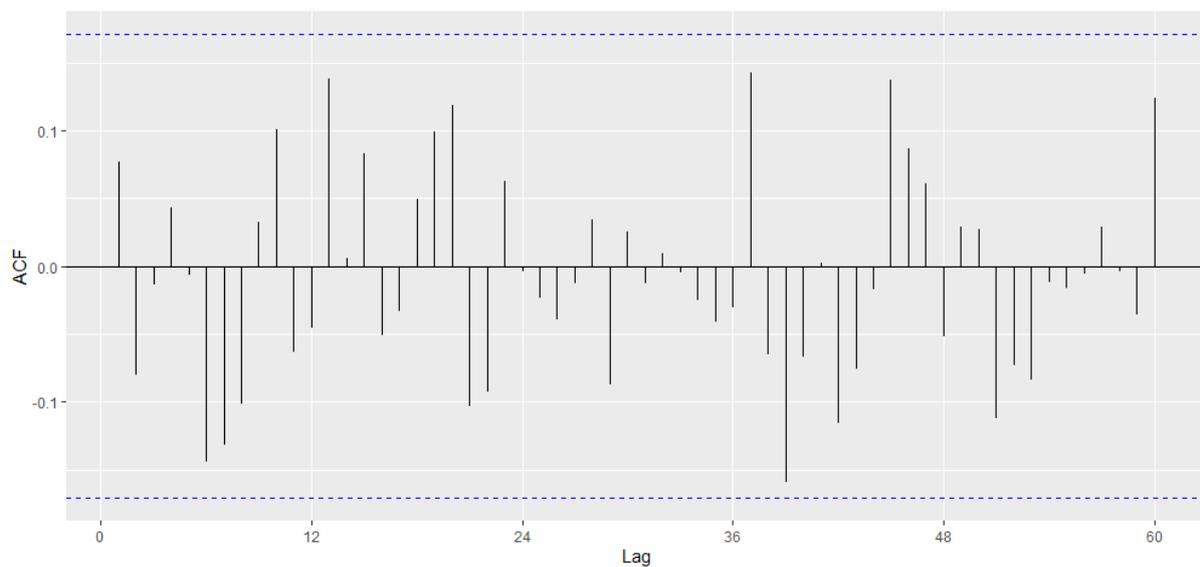
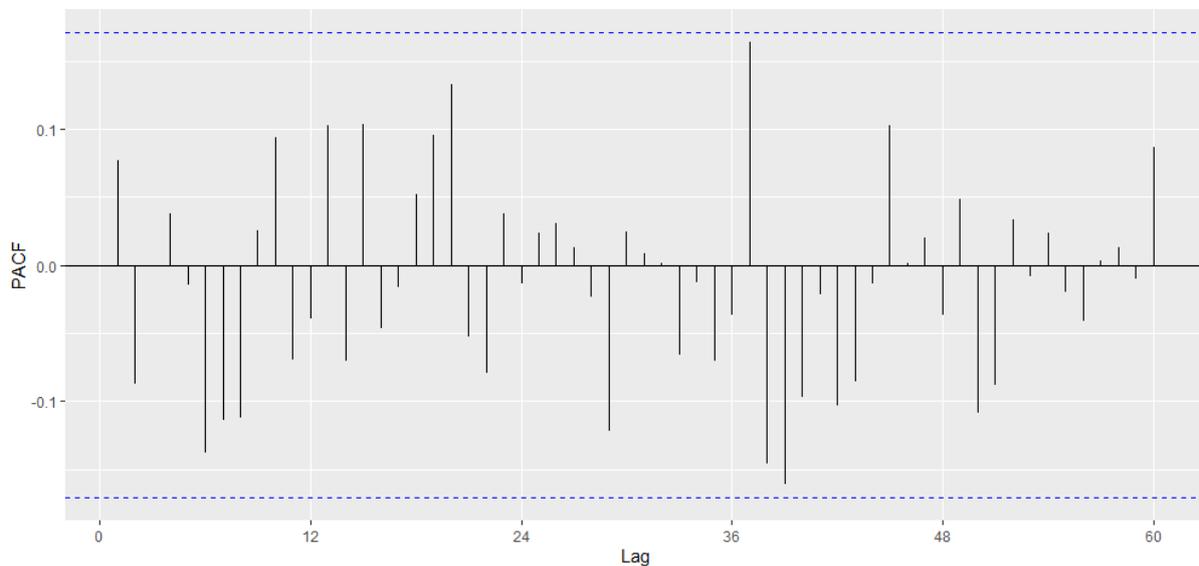


Figura 29

Función de autocorrelación parcial de los residuos del modelo SARIMA (0,1,1)(1,1,2)₁₂



En las funciones de autocorrelación, tanto simple como parcial de los residuos del modelo $SARIMA (0,1,1)(1,1,2)_{12}$, se observa que ningún retardo sobrepasa los límites de confianza, por lo cual los residuos cumplirían con las condiciones de ruido blanco.

Tabla 9

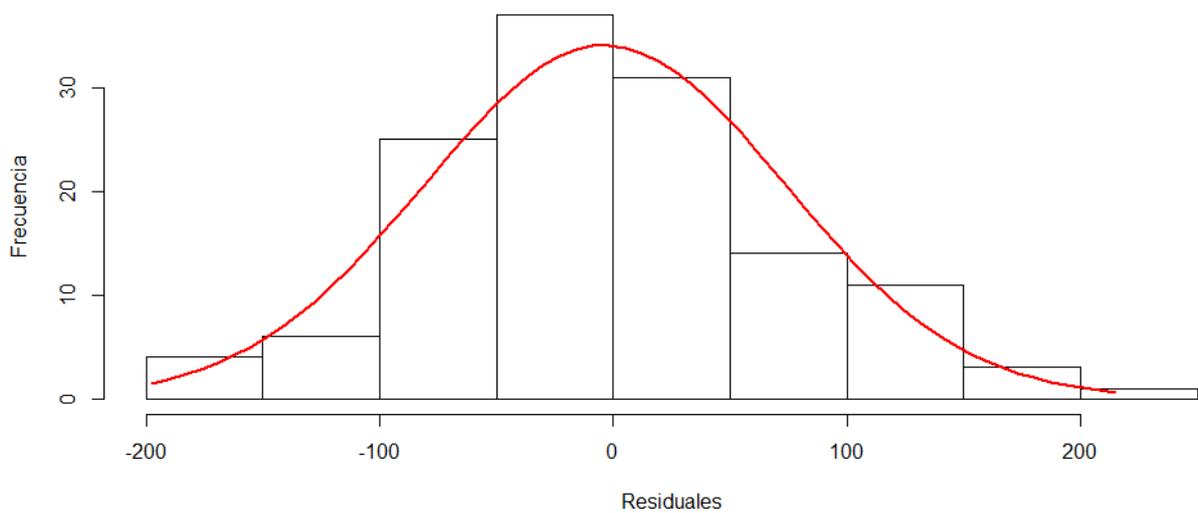
Test de Ljung-Box de los residuos del modelo SARIMA (0,1,1)(1,1,2)₁₂

Chi-Cuadrado	GL	P-Valor
0.797	1	0.372

En la tabla 9 se observa el test de Ljung-Box, presenta un valor chi-cuadrado de 0.797 y un valor $p = 0.372$, por lo se determinó que los coeficientes de autocorrelación son igual a cero, lo que permite concluir que los residuos del modelo $SARIMA (0,1,1)(1,1,2)_{12}$ se distribuyen de forma independiente.

Figura 30

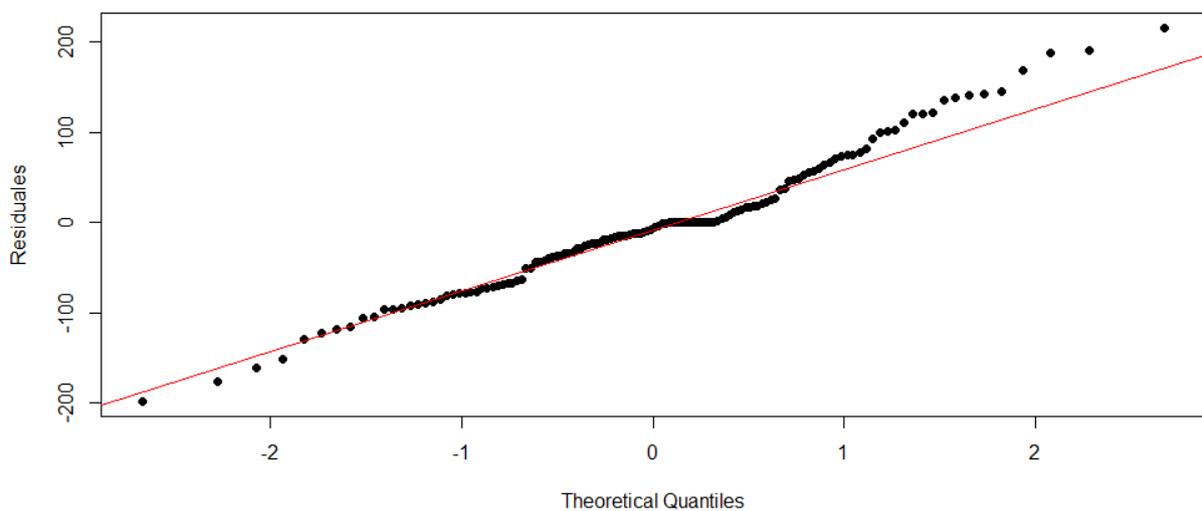
Histograma de los residuos del modelo SARIMA (0,1,1)(1,1,2)₁₂



En el histograma se observa que los residuos del modelo $SARIMA (0,1,1)(1,1,2)_{12}$, tienden a seguir una distribución normal.

Figura 31

Gráfica Q-Q normal de los residuos del modelo SARIMA (0,1,1)(1,1,2)₁₂



En el gráfico de probabilidad normal (gráfico Q-Q) se muestra que los residuos del modelo $SARIMA (0,1,1)(1,1,2)_{12}$, tienden a seguir una línea recta, por lo cual aparentemente siguen una distribución normal.

Tabla 10

Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA (0,1,1)(1,1,2)₁₂

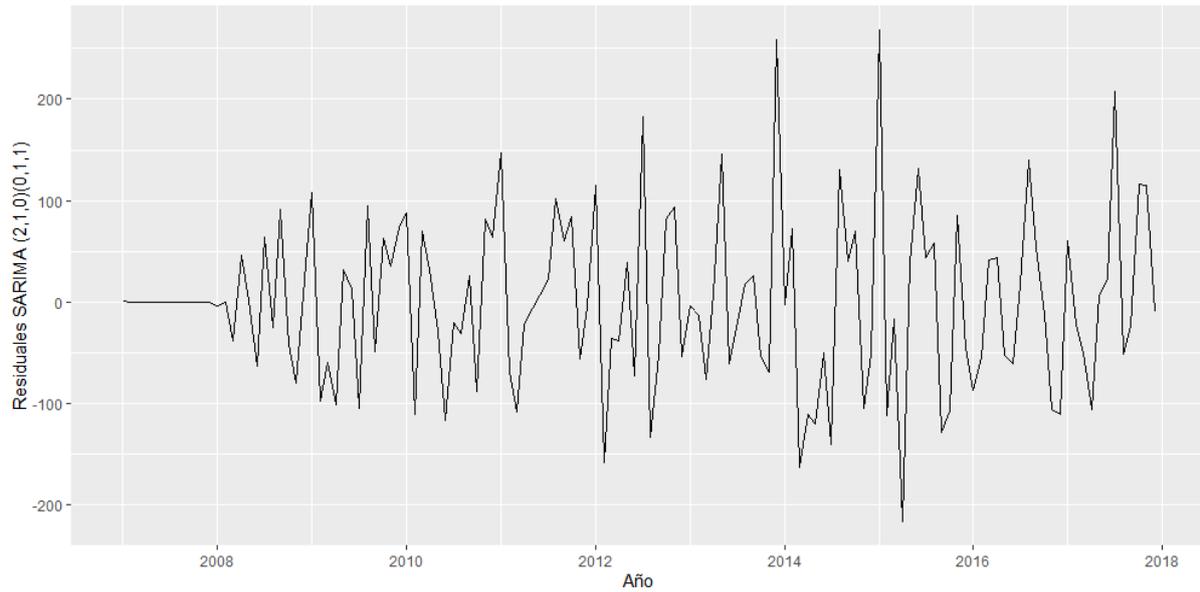
Prueba de normalidad	Kolmogorov-Smirnov	
	Estadístico	Sig
Residuales	0.108	0.090

En la tabla 10 se aprecia el estadístico de Kolmogorov-Smirnov, el cual arroja un valor de 0.108 y un valor $p = 0.090$, por lo que se determinó que las perturbaciones tienen una distribución conjunta normal, concluyendo que los residuos del modelo SARIMA (0,1,1)(1,1,2)₁₂ se distribuyen normalmente.

MODELO SARIMA (2,1,0) (0,1,1)₁₂

Figura 32

Residuos del modelo SARIMA (2,1,0)(0,1,1)₁₂



En la figura 32 se presentan los residuos del modelo $SARIMA (2,1,0)(0,1,1)_{12}$, los mismos que siguen un ruido blanco.

Figura 33

Función de autocorrelación de los residuos del modelo SARIMA (2,1,0)(0,1,1)₁₂

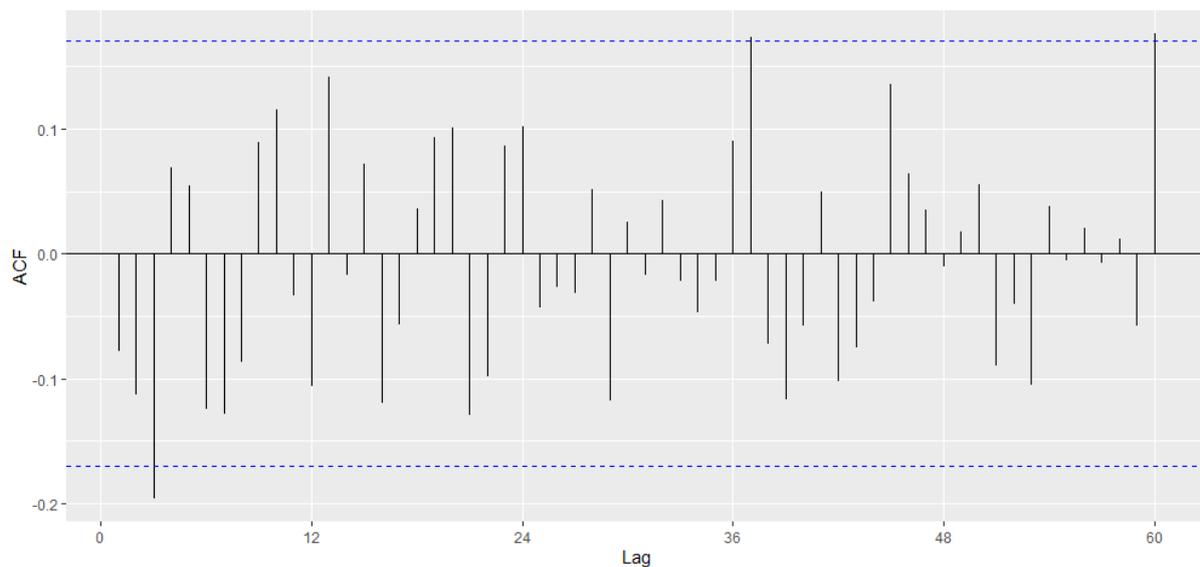
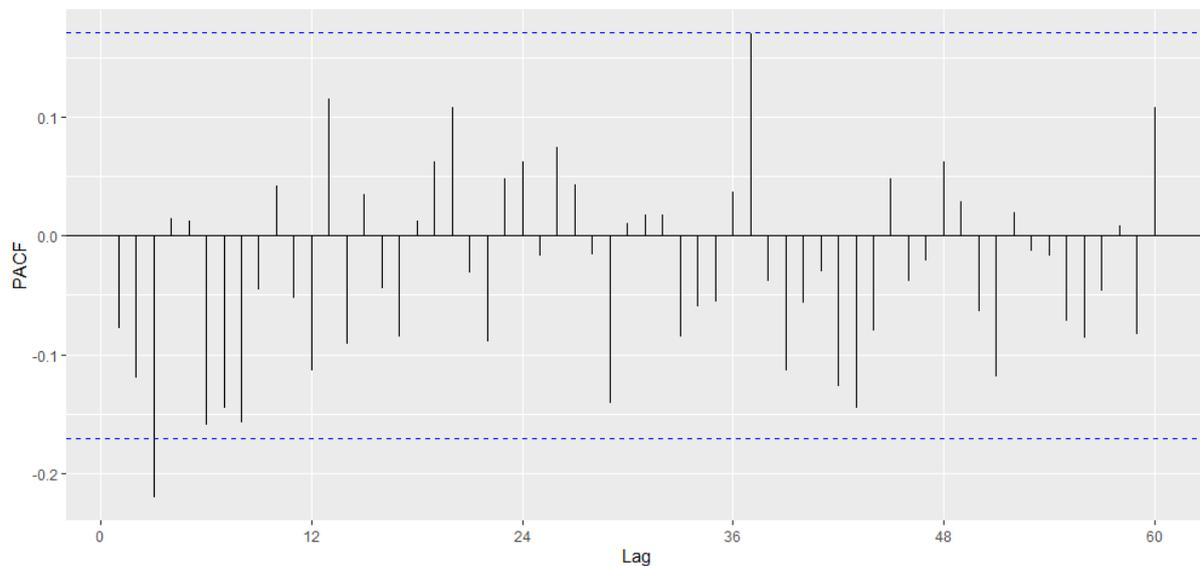


Figura 34

Función de autocorrelación parcial de los residuos del modelo SARIMA (2,1,0)(0,1,1)₁₂



En las funciones de autocorrelación, tanto simple como parcial de los residuos del modelo $SARIMA (2,1,0)(0,1,1)_{12}$, se observa que existen retardos que sobrepasan los límites de confianza, aun así los residuos cumplirían con las condiciones de ruido blanco.

Tabla 11

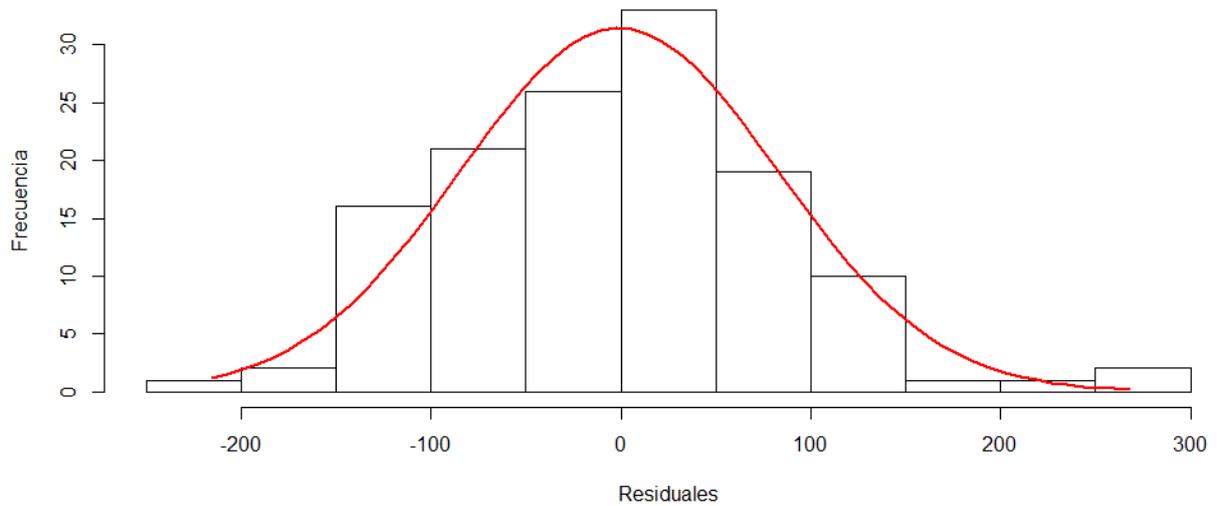
Test de Ljung-Box de los residuos del modelo SARIMA (2,1,0)(0,1,1)₁₂

Chi-Cuadrado	GL	P-Valor
0.832	1	0.362

En la tabla 11 se observa el test de Ljung-Box, el cual arroja un valor chi-cuadrado de 0.832 y un valor $p = 0.362$, por lo que se determinó que los coeficientes de autocorrelación son igual a cero, lo que permite concluir que los residuos del modelo $SARIMA (2,1,0)(0,1,1)_{12}$ se distribuyen de forma independiente.

Figura 35

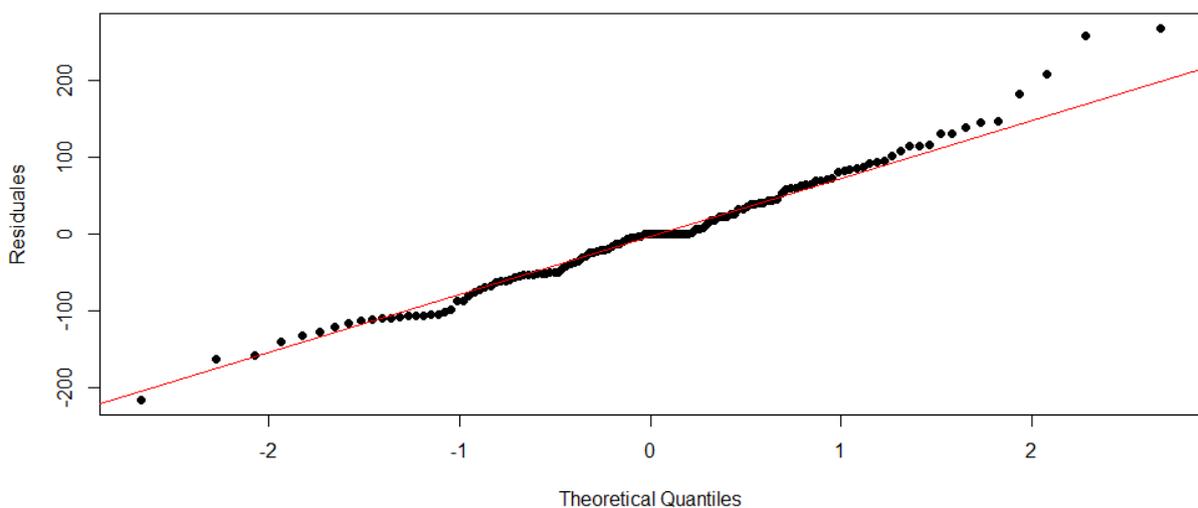
Histograma de los residuos del modelo SARIMA (2,1,0)(0,1,1)₁₂



En el histograma se observa que los residuos del modelo $SARIMA (2,1,0)(0,1,1)_{12}$, tienden a seguir una distribución normal.

Figura 36

Gráfica Q-Q normal de los residuos del modelo SARIMA (2,1,0)(0,1,1)₁₂



En el gráfico de probabilidad normal (gráfico Q-Q) se muestra que los residuos del modelo $SARIMA (2,1,0)(0,1,1)_{12}$, tienden a seguir una línea recta, por lo cual aparentemente siguen una distribución normal.

Tabla 12

Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA (2,1,0)(0,1,1)₁₂

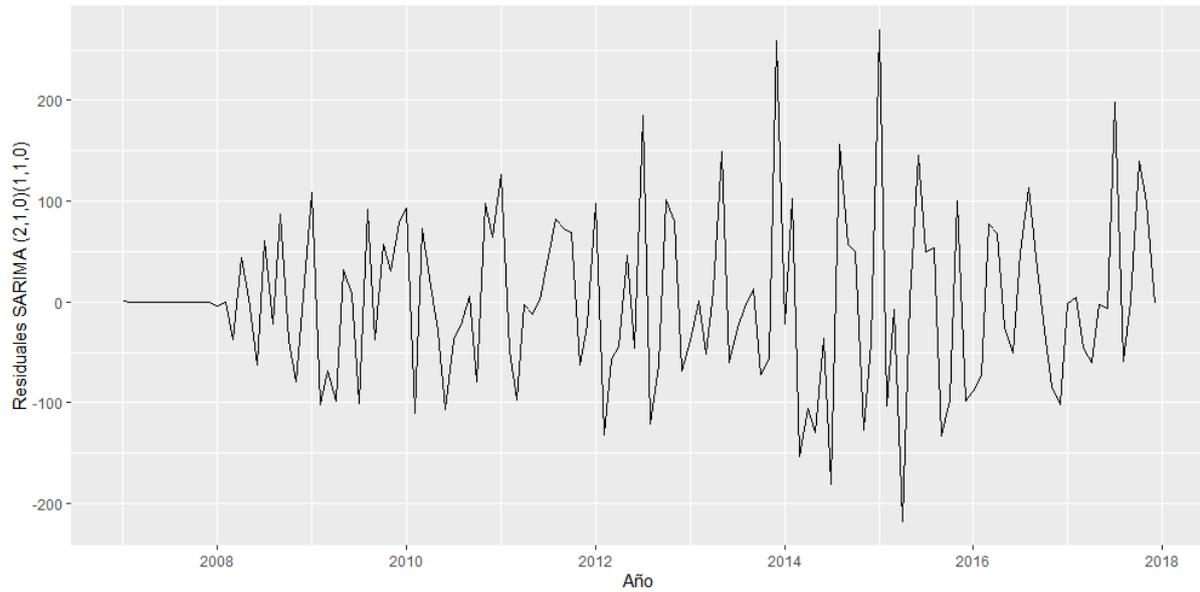
Prueba de normalidad	Kolmogorov-Smirnov	
	Estadístico	Sig
Residuales	0.079	0.382

En la tabla 12 se aprecia el estadístico de Kolmogorov-Smirnov, el cual arroja un valor de 0.079 y un valor $p = 0.382$, por lo que se determinó que las perturbaciones tienen una distribución conjunta normal, concluyendo que los residuos del modelo SARIMA (2,1,0)(0,1,1)₁₂ siguen una distribución normal.

MODELO SARIMA (2,1,0) (1,1,0)₁₂

Figura 37

Residuos del modelo SARIMA (2,1,0)(1,1,0)₁₂



En la figura 37 se presentan los residuos del modelo $SARIMA (2,1,0)(1,1,0)_{12}$, los mismos que siguen un ruido blanco.

Figura 38

Función de autocorrelación de los residuos del modelo SARIMA (2,1,0)(1,1,0)₁₂

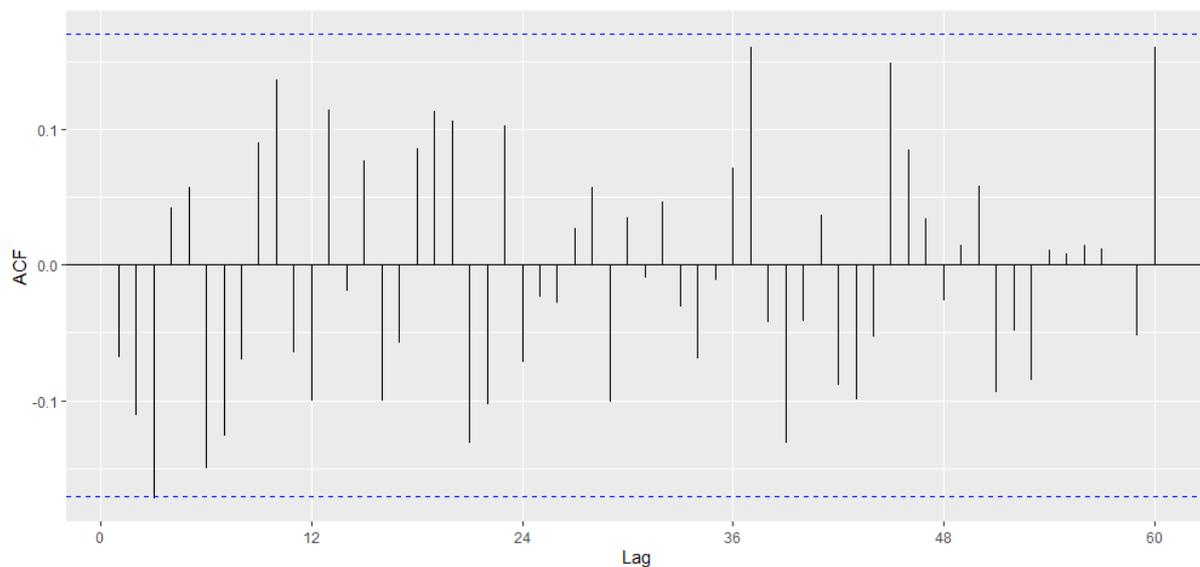
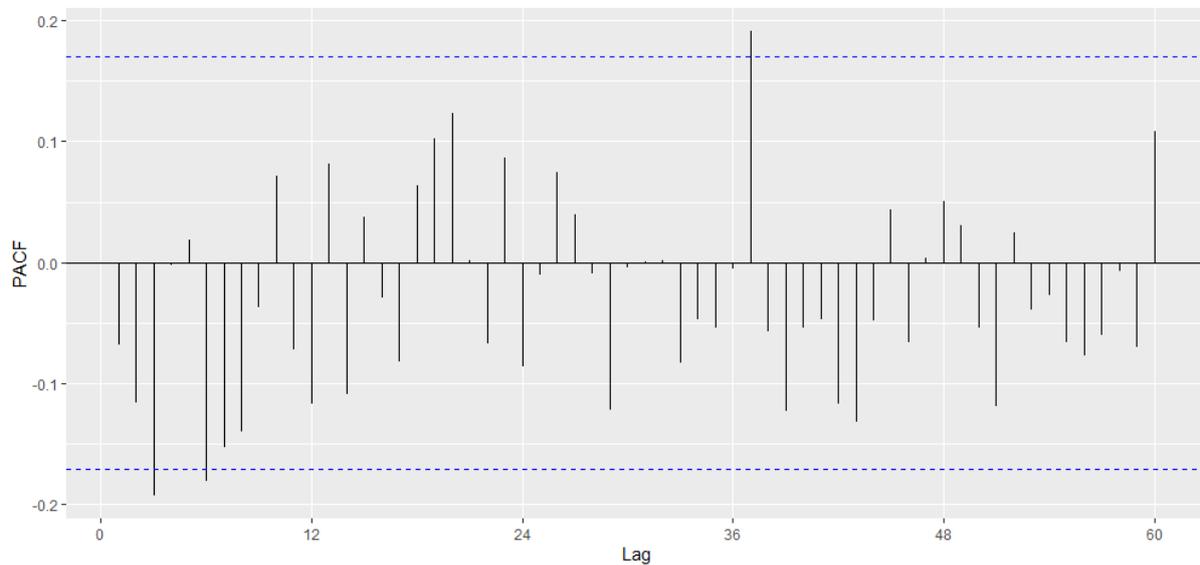


Figura 39

Función de autocorrelación parcial de los residuos del modelo SARIMA (2,1,0)(1,1,0)₁₂



En las funciones de autocorrelación, tanto simple como parcial de los residuos del modelo SARIMA (2,1,0)(1,1,0)₁₂, se observa que existen retardos que sobrepasan los límites de confianza, aun así los residuos cumplirían con las condiciones de ruido blanco.

Tabla 13

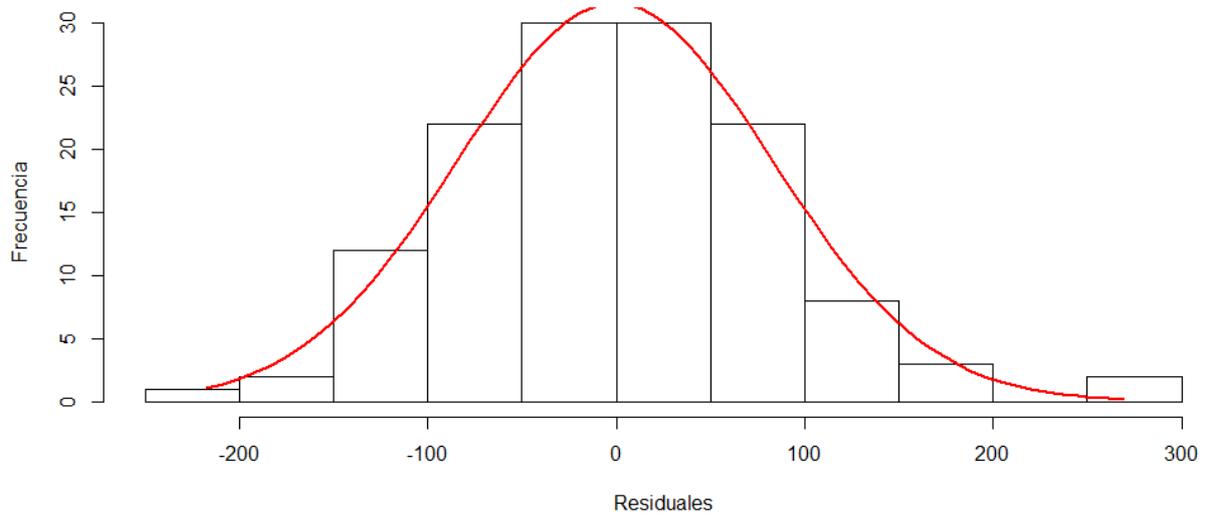
Test de Ljung-Box de los residuos del modelo SARIMA (2,1,0)(1,1,0)₁₂

Chi-Cuadrado	GL	P-Valor
0.624	1	0.429

En la tabla 13 se observa el test de Ljung-Box, el cual presenta un chi-cuadrado de 0.624 y un valor $p = 0.429$, por lo que se determinó que los coeficientes de autocorrelación son igual a cero, lo que permite concluir que los residuos del modelo SARIMA (2,1,0)(1,1,0)₁₂ se distribuyen de forma independiente.

Figura 40

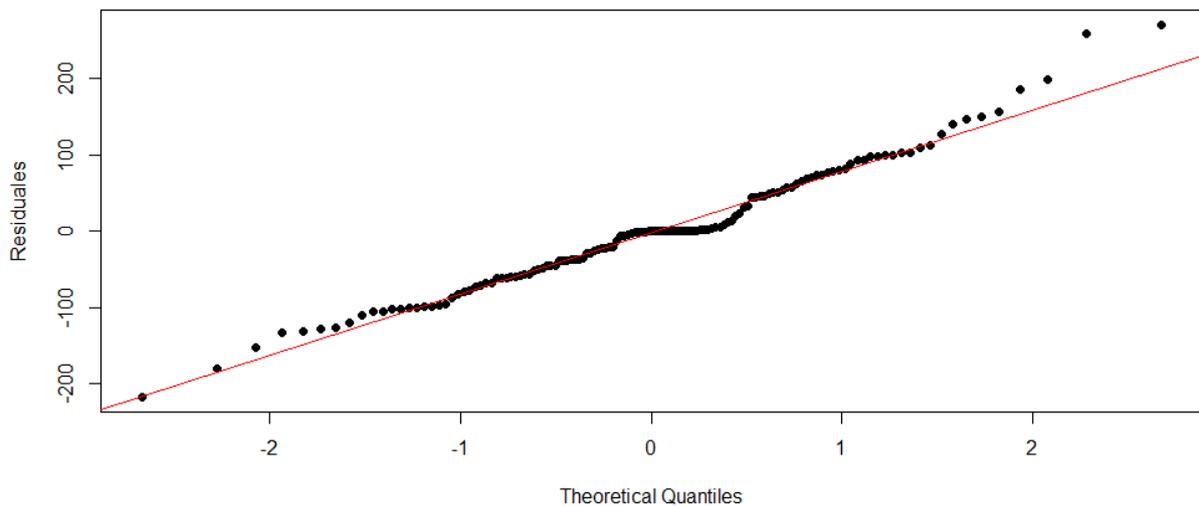
Histograma de los residuos del modelo SARIMA (2,1,0)(1,1,0)₁₂



En el histograma se observa que los residuos del modelo $SARIMA (2,1,0)(1,1,0)_{12}$, tienden a seguir una distribución normal.

Figura 41

Gráfica Q-Q normal de los residuos del modelo SARIMA (2,1,0)(1,1,0)₁₂



En el gráfico de probabilidad normal (gráfico Q-Q) se muestra que los residuos del modelo $SARIMA (2,1,0)(1,1,0)_{12}$, tienden a seguir una línea recta, por lo cual aparentemente siguen una distribución normal.

Tabla 14

Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA (2,1,0)(1,1,0)₁₂

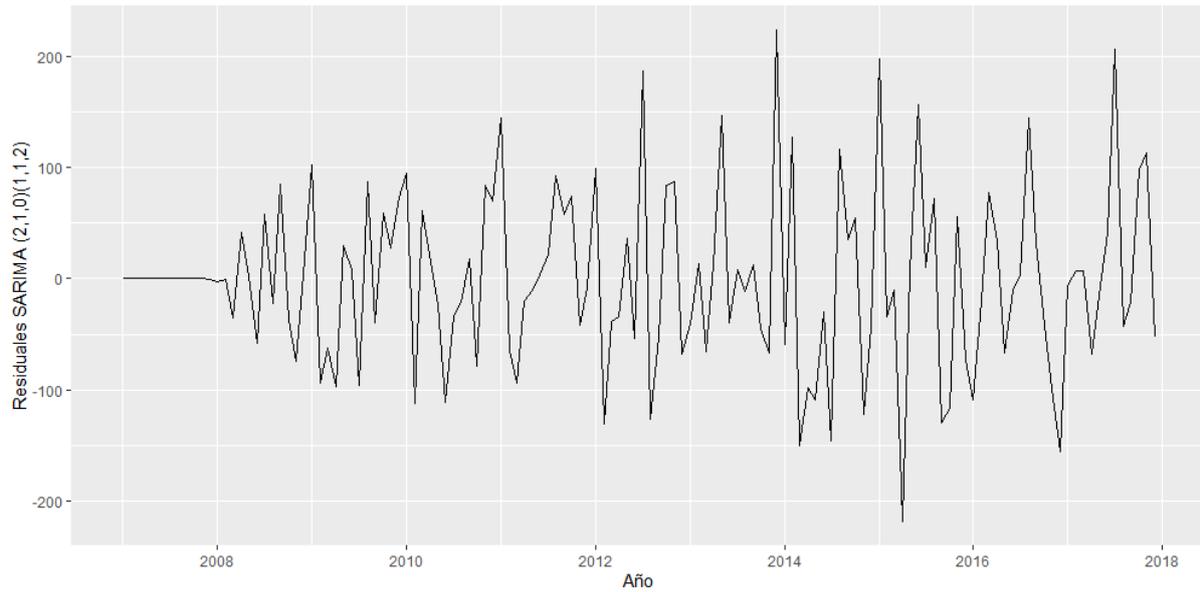
Prueba de normalidad	Kolmogorov-Smirnov	
	Estadístico	Sig
Residuales	0.114	0.063

En la tabla 14 se aprecia el estadístico de Kolmogorov-Smirnov, el cual arroja un valor de 0.114 y un valor $p = 0.554$, por lo que se determinó que las perturbaciones tienen una distribución conjunta normal, concluyendo que los residuos del modelo SARIMA (2,1,0)(1,1,0)₁₂ siguen una distribución normal.

MODELO SARIMA (2,1,0) (1,1,2)₁₂

Figura 42

Residuos del modelo SARIMA (2,1,0)(1,1,2)₁₂



En la figura 42 se presentan los residuos del modelo $SARIMA (2,1,0)(1,1,2)_{12}$, los mismos que siguen un ruido blanco.

Figura 43

Función de autocorrelación de los residuos del modelo SARIMA (2,1,0)(1,1,2)₁₂

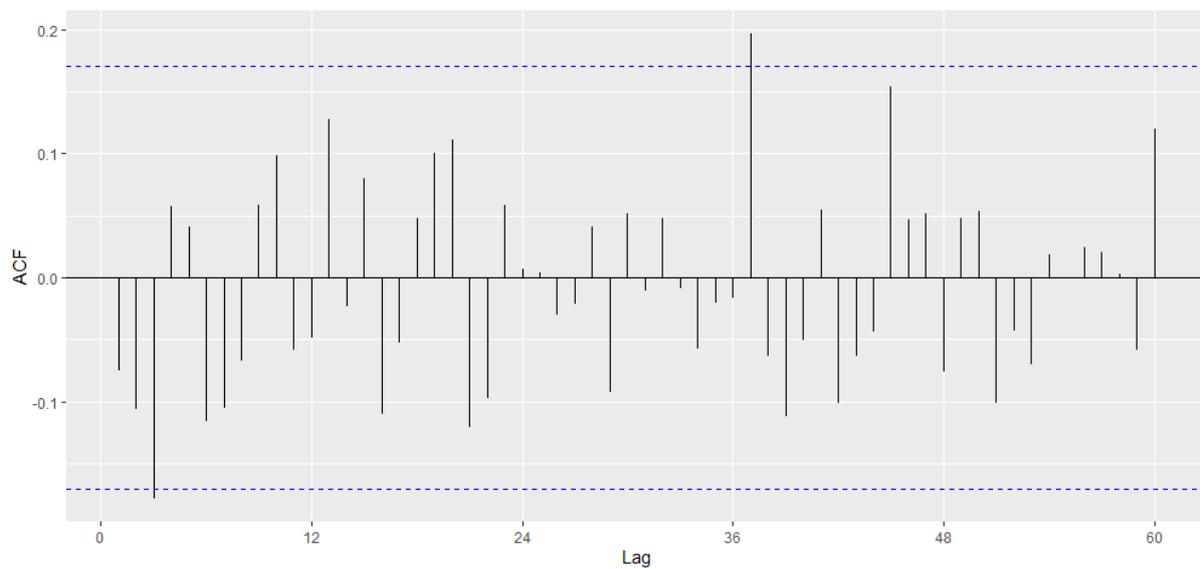
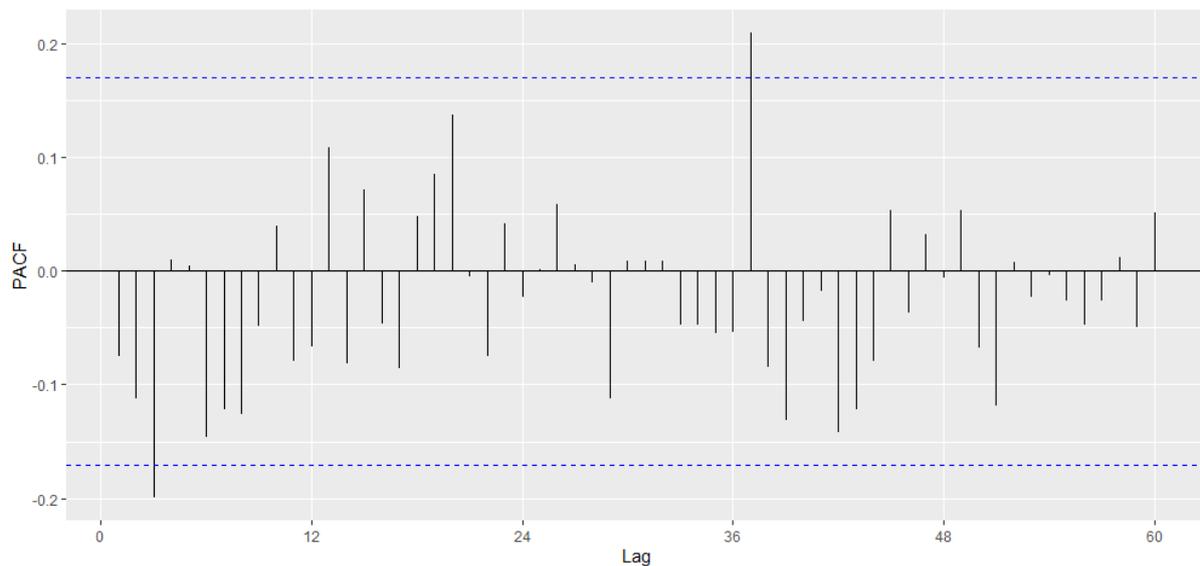


Figura 44

Función de autocorrelación parcial de los residuos del modelo SARIMA (2,1,0)(1,1,2)₁₂



En las funciones de autocorrelación, tanto simple como parcial de los residuos del modelo SARIMA (2,1,0)(1,1,2)₁₂, se observa que existen 2 retardos que sobrepasan los límites de confianza, aun así los residuos cumplirían con las condiciones de ruido blanco.

Tabla 15

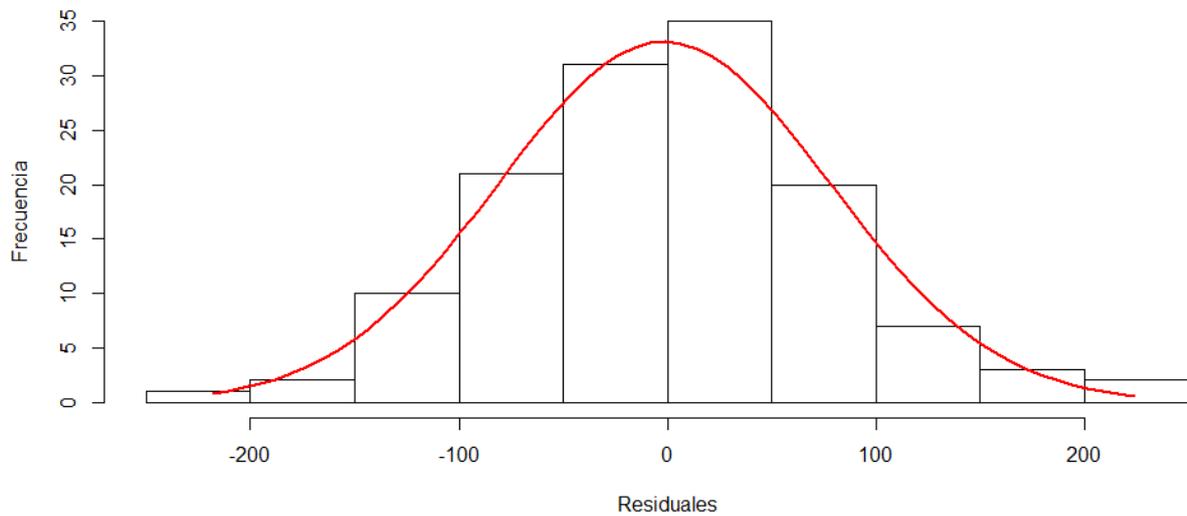
Test de Ljung-Box de los residuos del modelo SARIMA (2,1,0)(1,1,2)₁₂

Chi-Cuadrado	GL	P-Valor
0.765	1	0.382

En la tabla 15 se observa el test de Ljung-Box, el cual arroja un chi-cuadrado de 0.765 y un valor p = 0.382, por lo se determinó que los coeficientes de autocorrelación son igual a cero, lo que permite concluir que los residuos del modelo SARIMA (2,1,0)(1,1,2)₁₂ se distribuyen de forma independiente.

Figura 45

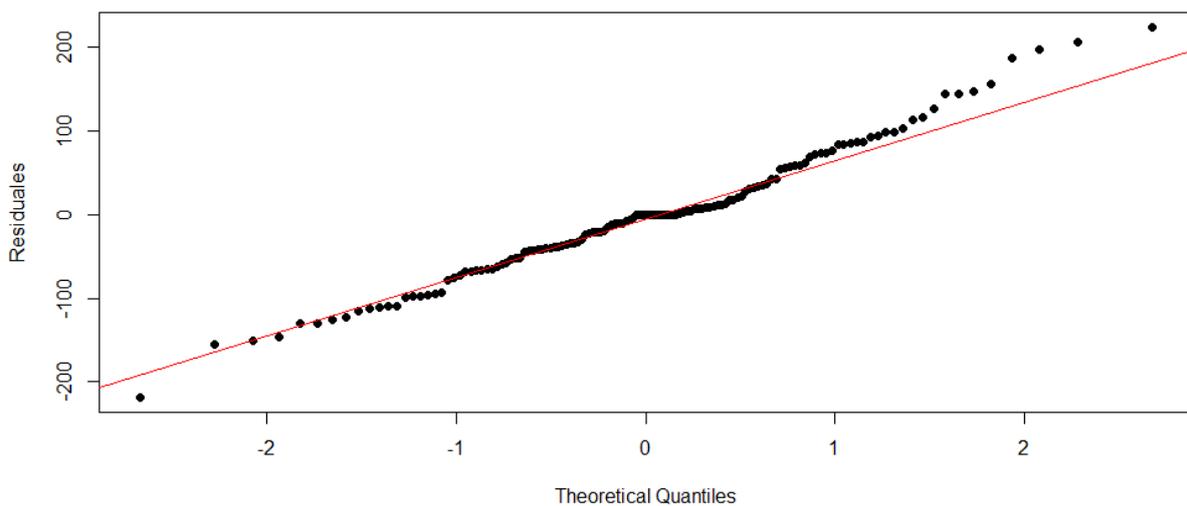
Histograma de los residuos del modelo SARIMA (2,1,0)(1,1,2)₁₂



En el histograma se observa que los residuos del modelo $SARIMA (2,1,0)(1,1,2)_{12}$, tienden a seguir una distribución normal.

Figura 46

Grafica Q-Q normal de los residuos del modelo SARIMA (2,1,0)(1,1,2)₁₂



En el gráfico de probabilidad normal (gráfico Q-Q) se muestra que los residuos del modelo $SARIMA (2,1,0)(1,1,2)_{12}$, tienden a seguir una línea recta, por lo cual aparentemente siguen una distribución normal.

Tabla 16

Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA (2,1,0)(1,1,2)₁₂

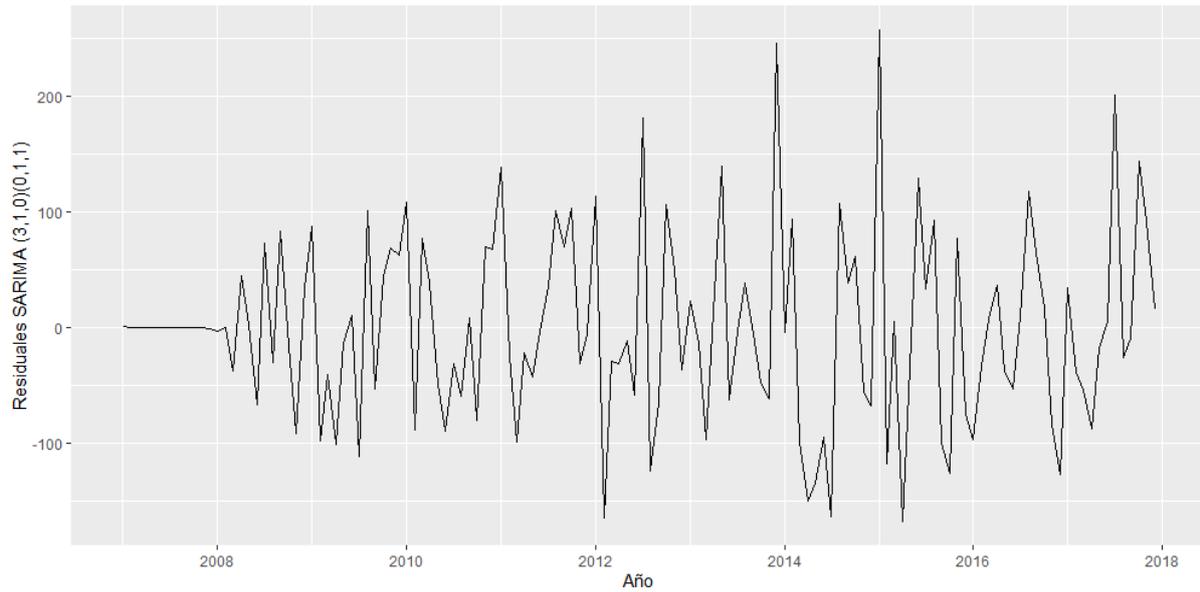
Prueba de normalidad	Kolmogorov-Smirnov	
	Estadístico	Sig
Residuales	0.094	0.198

En la tabla 16 se aprecia el estadístico de Kolmogorov-Smirnov, el cual arroja un valor de 0.094 y un valor $p = 0.198$, por lo que se determinó que las perturbaciones tienen una distribución conjunta normal, concluyendo que los residuos del modelo *SARIMA (2,1,0)(1,1,2)₁₂* siguen una distribución normal.

MODELO SARIMA (3,1,0) (0,1,1)₁₂

Figura 47

Residuos del modelo SARIMA (3,1,0)(0,1,1)₁₂



En la figura 47 se presentan los residuos del modelo $SARIMA (3,1,0)(0,1,1)_{12}$, los mismos que siguen un ruido blanco.

Figura 48

Función de autocorrelación de los residuos del modelo SARIMA (3,1,0)(0,1,1)₁₂

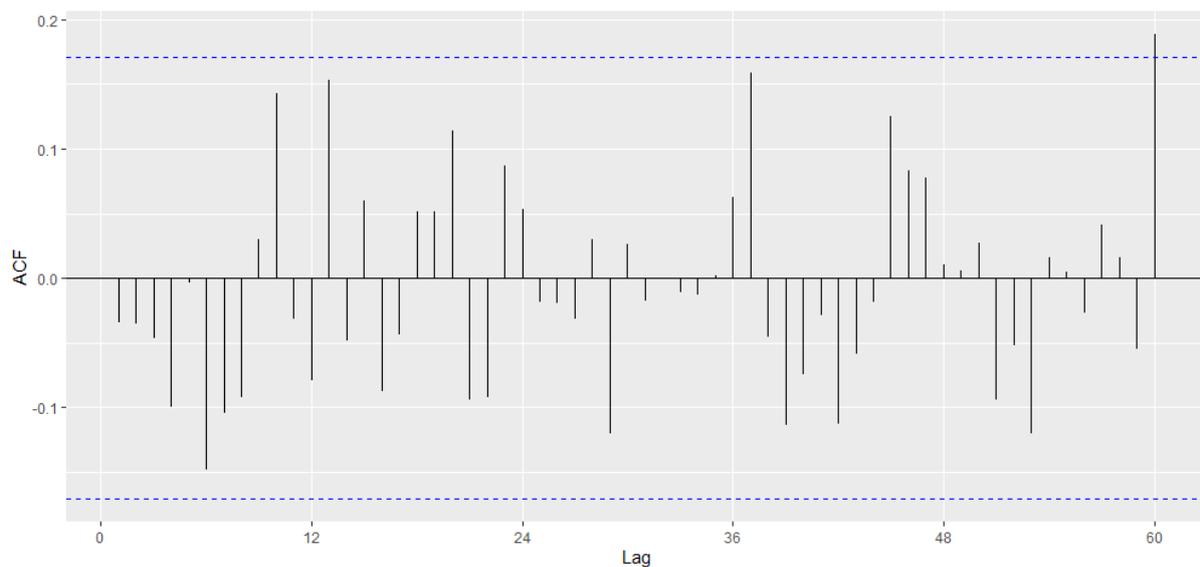
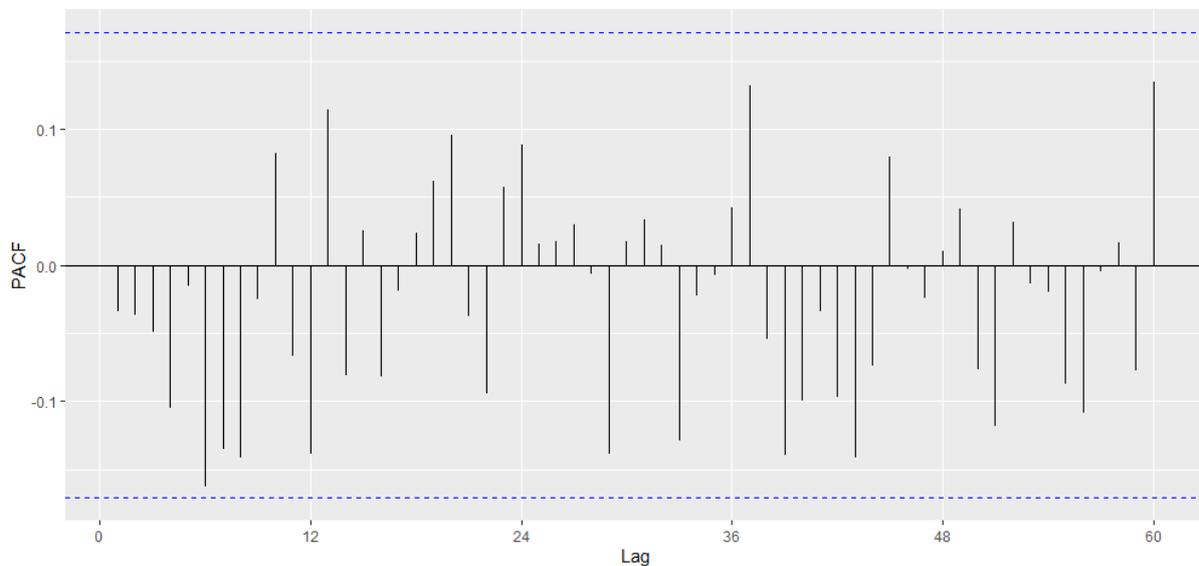


Figura 49

Función de autocorrelación parcial de los residuos del modelo SARIMA (3,1,0)(0,1,1)₁₂



En la función de autocorrelación simple de los residuos del modelo $SARIMA (3,1,0)(0,1,1)_{12}$, se observa que existe 1 retardo que sobrepasa los límites de confianza, y en la función de autocorrelación parcial ningún retardo sobrepasa los límites de confianza, aun así los residuos cumplirían con las condiciones de ruido blanco.

Tabla 17

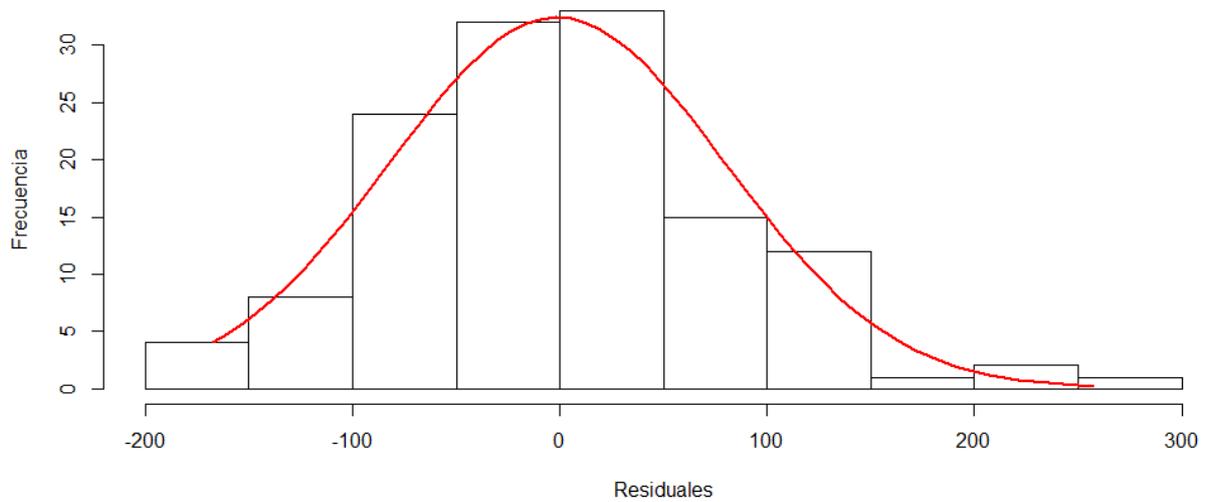
Test de Ljung-Box de los residuos del modelo SARIMA (3,1,0)(0,1,1)₁₂

Chi-Cuadrado	GL	P-Valor
0.158	1	0.691

En la tabla 17 se observa el test de Ljung-Box, el cual arroja un chi-cuadrado de 0.158 y un valor $p = 0.691$, por lo que se determinó que los coeficientes de autocorrelación son igual a cero, lo que permite concluir que los residuos del modelo $SARIMA (3,1,0)(0,1,1)_{12}$ se distribuyen de forma independiente.

Figura 50

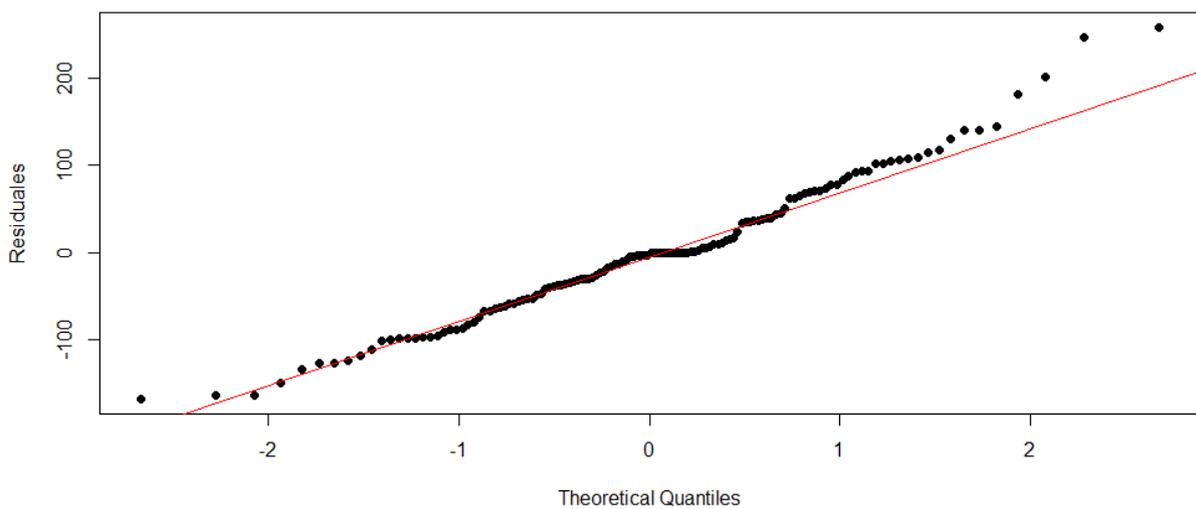
Histograma de los residuos del modelo SARIMA (3,1,0)(0,1,1)₁₂



En el histograma se observa que los residuos del modelo $SARIMA (3,1,0)(0,1,1)_{12}$, posiblemente tienden a seguir una distribución normal.

Figura 51

Grafica Q-Q normal de los residuos del modelo SARIMA (3,1,0)(0,1,1)₁₂



En el gráfico de probabilidad normal (gráfico Q-Q) se muestra que los residuos del modelo $SARIMA (3,1,0)(0,1,1)_{12}$, tienden a seguir una línea recta, por lo cual aparentemente siguen una distribución normal.

Tabla 18

Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA (3,1,0)(0,1,1)₁₂

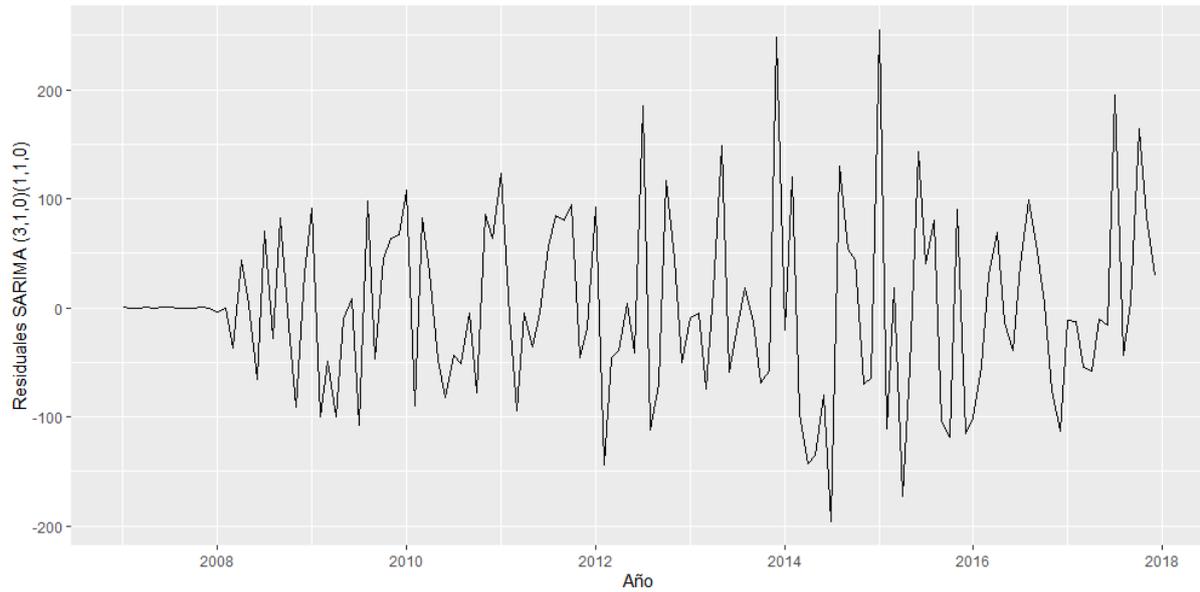
Prueba de normalidad	Kolmogorov-Smirnov	
	Estadístico	Sig
Residuales	0.093	0.202

En la tabla 32 se aprecia el estadístico de Kolmogorov-Smirnov, el cual arroja un valor de 0.093 y un valor $p = 0.202$, por lo que se determinó que las perturbaciones tienen una distribución conjunta normal, concluyendo que los residuos del modelo *SARIMA (3,1,0)(0,1,1)₁₂* siguen una distribución normal.

MODELO SARIMA (3,1,0) (1,1,0)₁₂

Figura 52

Residuos del modelo SARIMA (3,1,0)(1,1,0)₁₂



En la figura 52 se presentan los residuos del modelo $SARIMA (3,1,0)(1,1,0)_{12}$, los mismos que siguen un ruido blanco.

Figura 53

Función de autocorrelación de los residuos del modelo SARIMA (3,1,0)(1,1,0)₁₂

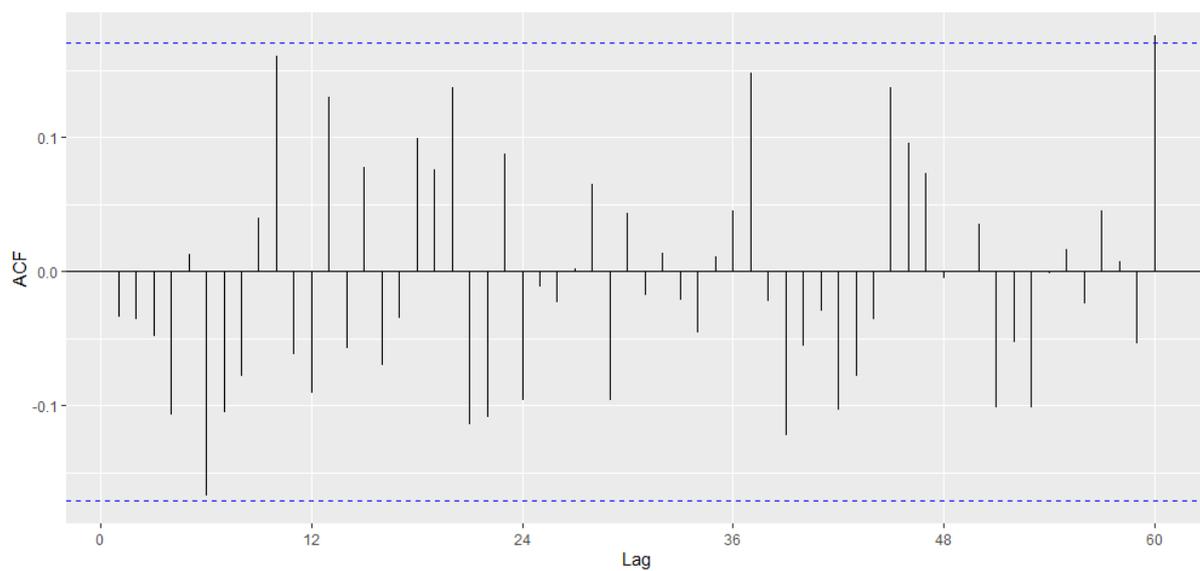
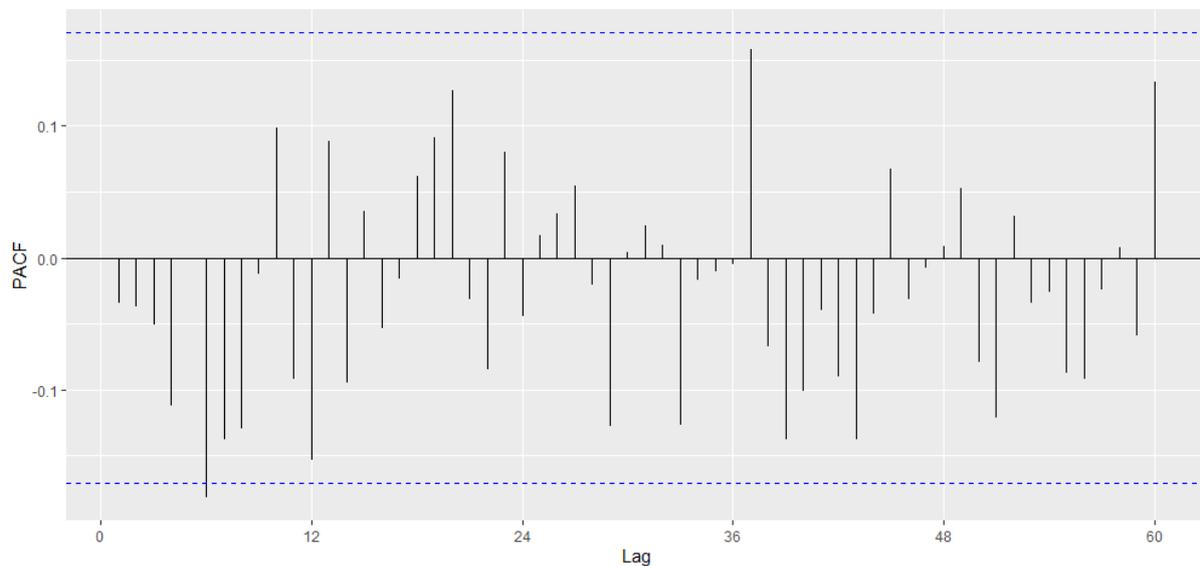


Figura 54

Función de autocorrelación parcial de los residuos del modelo SARIMA (3,1,0)(1,1,0)₁₂



En las funciones de autocorrelación, tanto simple como parcial de los residuos del modelo $SARIMA (3,1,0)(1,1,0)_{12}$, se observa que existe 1 retardo que sobrepasa los límites de confianza, aun así los residuos cumplirían con las condiciones de ruido blanco.

Tabla 19

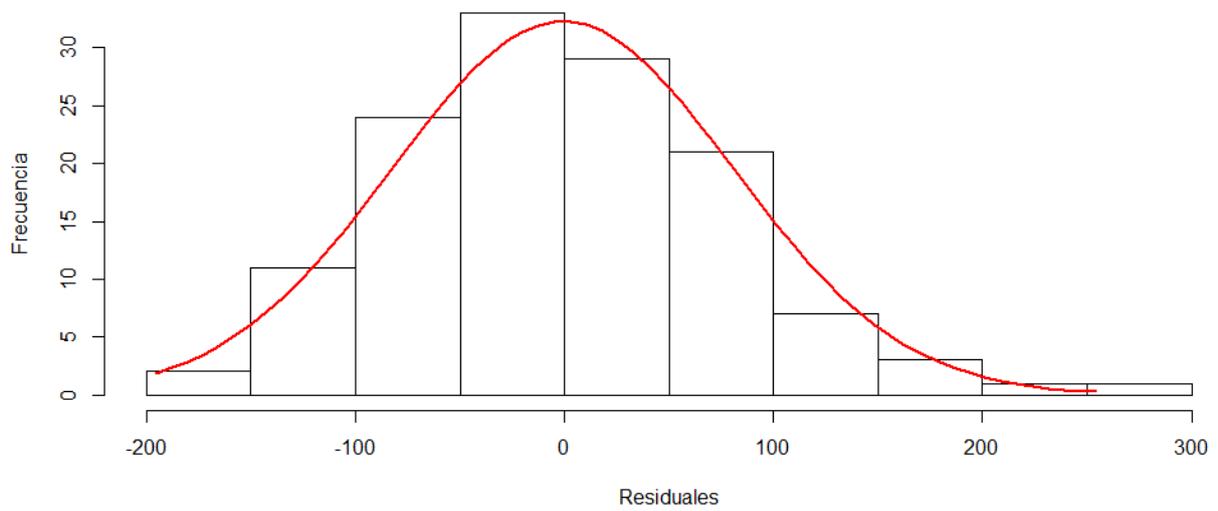
Test de Ljung-Box de los residuos del modelo SARIMA (3,1,0)(1,1,0)₁₂

Chi-Cuadrado	GL	P-Valor
0.156	1	0.693

En la tabla 19 se observa el test de Ljung-Box, el cual arroja un chi-cuadrado de 0.156 y un valor $p = 0.693$, por lo que se determinó que los coeficientes de autocorrelación son igual a cero, lo que permite concluir que los residuos del modelo $SARIMA (3,1,0)(1,1,0)_{12}$ se distribuyen de forma independiente.

Figura 55

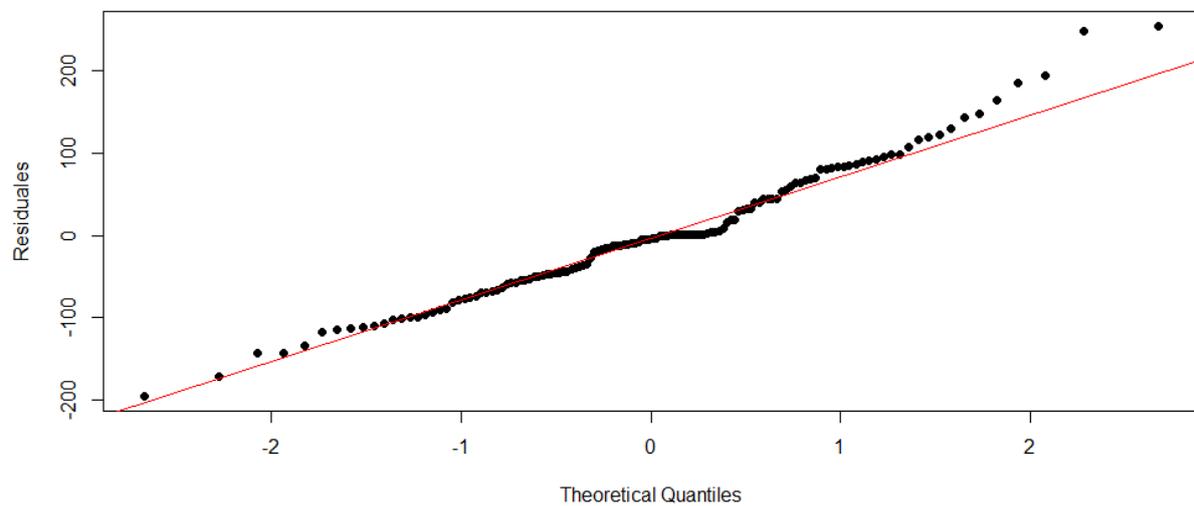
Histograma de los residuos del modelo SARIMA (3,1,0)(1,1,0)₁₂



En el histograma se observa que los residuos del modelo $SARIMA (3,1,0)(1,1,0)_{12}$, posiblemente tienden a seguir una distribución normal.

Figura 56

Gráfica Q-Q normal de los residuos del modelo SARIMA (3,1,0)(1,1,0)₁₂



En el gráfico de probabilidad normal (gráfico Q-Q) se muestra que los residuos del modelo $SARIMA (3,1,0)(1,1,0)_{12}$, tienden a seguir una línea recta, por lo cual aparentemente siguen una distribución normal.

Tabla 20

Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA (3,1,0)(1,1,0)₁₂

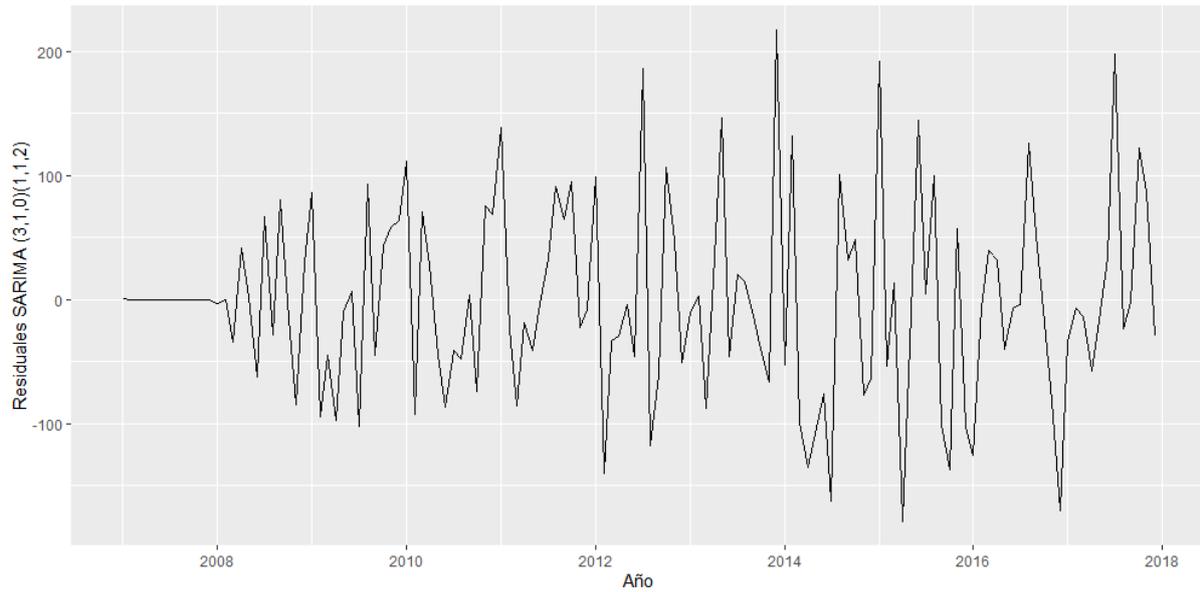
Prueba de normalidad	Kolmogorov-Smirnov	
	Estadístico	Sig
Residuales	0.112	0.071

En la tabla 20 se aprecia el estadístico de Kolmogorov-Smirnov, el cual arroja un valor de 0.112 y un valor $p = 0.071$, por lo que se determinó las perturbaciones tienen una distribución conjunta normal, concluyendo que los residuos del modelo *SARIMA (3,1,0)(1,1,0)₁₂* siguen una distribución normal.

MODELO SARIMA (3,1,0) (1,1,2)₁₂

Figura 57

Residuos del modelo SARIMA (3,1,0)(1,1,2)₁₂



En la figura 57 se presentan los residuos del modelo $SARIMA (3,1,0)(1,1,2)_{12}$, los mismos que siguen un ruido blanco.

Figura 58

Función de autocorrelación de los residuos del modelo SARIMA (3,1,0)(1,1,2)₁₂

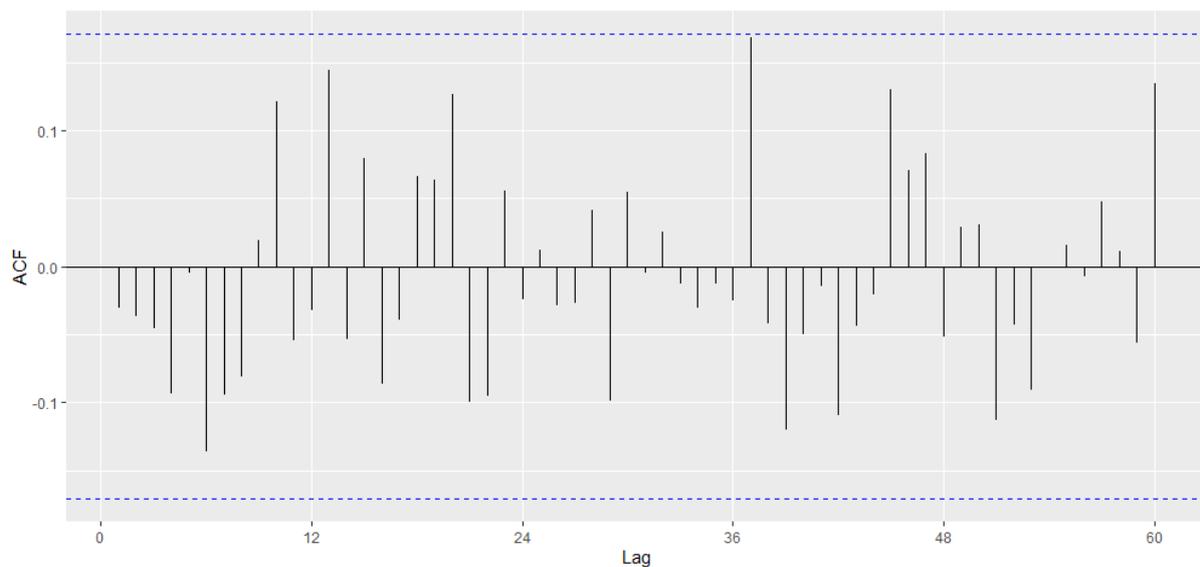
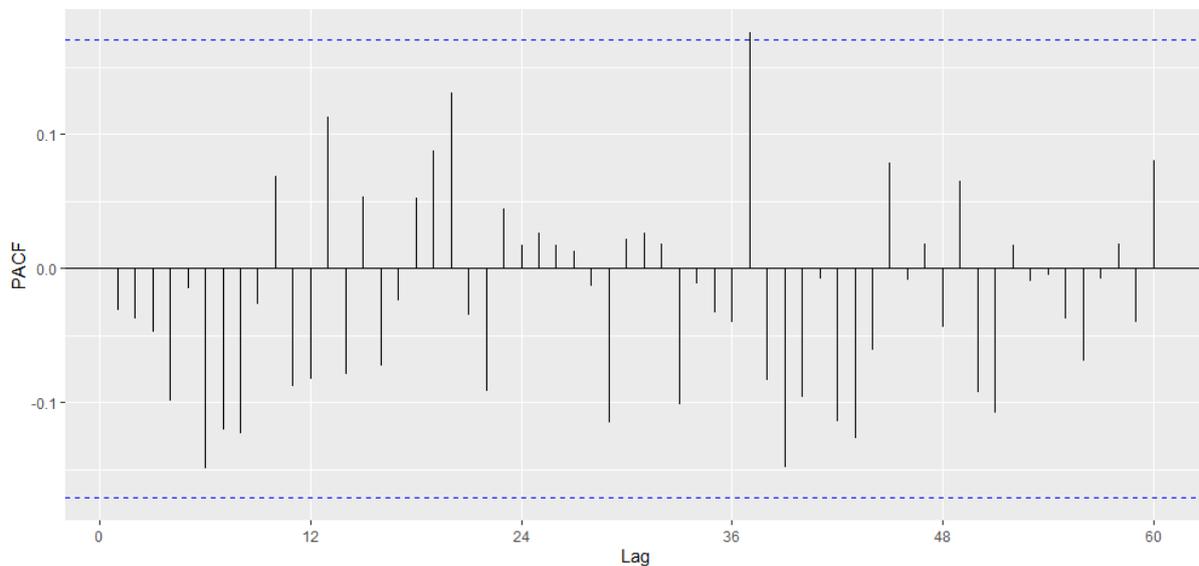


Figura 59

Función de autocorrelación parcial de los residuos del modelo SARIMA (3,1,0)(1,1,2)₁₂



En la función de autocorrelación simple de los residuos del modelo $SARIMA (3,1,0)(1,1,2)_{12}$, ningún retardo sobrepasa los límites de confianza, y en la función de autocorrelación parcial solo un retardo sobrepasa los límites de confianza, aun así los residuos cumplirían con las condiciones de ruido blanco.

Tabla 21

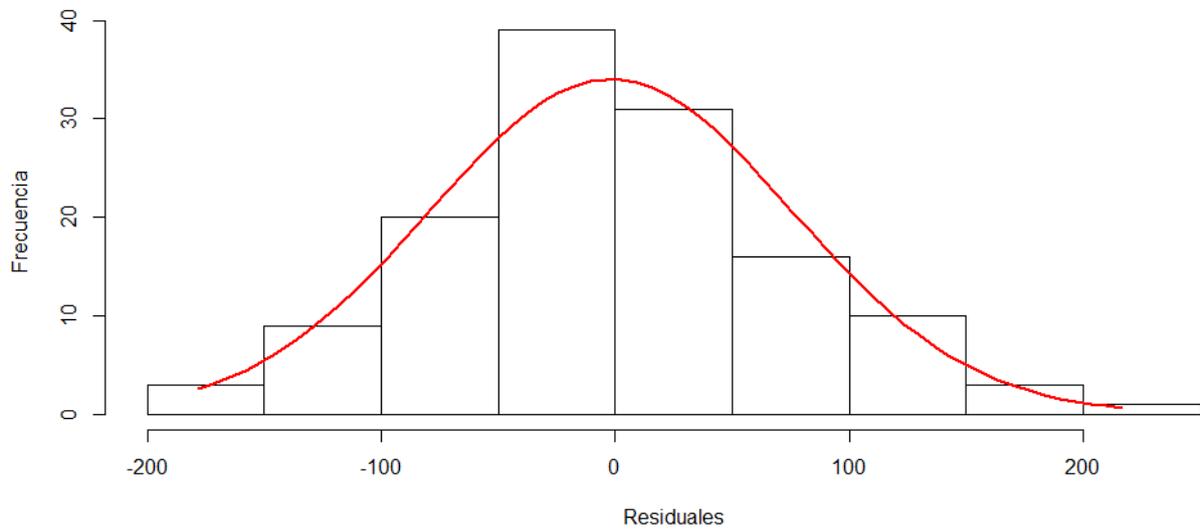
Test de Ljung-Box de los residuos del modelo SARIMA (3,1,0)(1,1,2)₁₂

Chi-Cuadrado	GL	P-Valor
0.130	1	0.719

En la tabla 21 se observa el test de Ljung-Box, el cual arroja un chi-cuadrado de 0.130 y un valor $p = 0.719$, por lo que se determinó los coeficientes de autocorrelación son igual a cero, lo que permite concluir que los residuos del modelo $SARIMA (3,1,0)(1,1,2)_{12}$ se distribuyen de forma independiente.

Figura 60

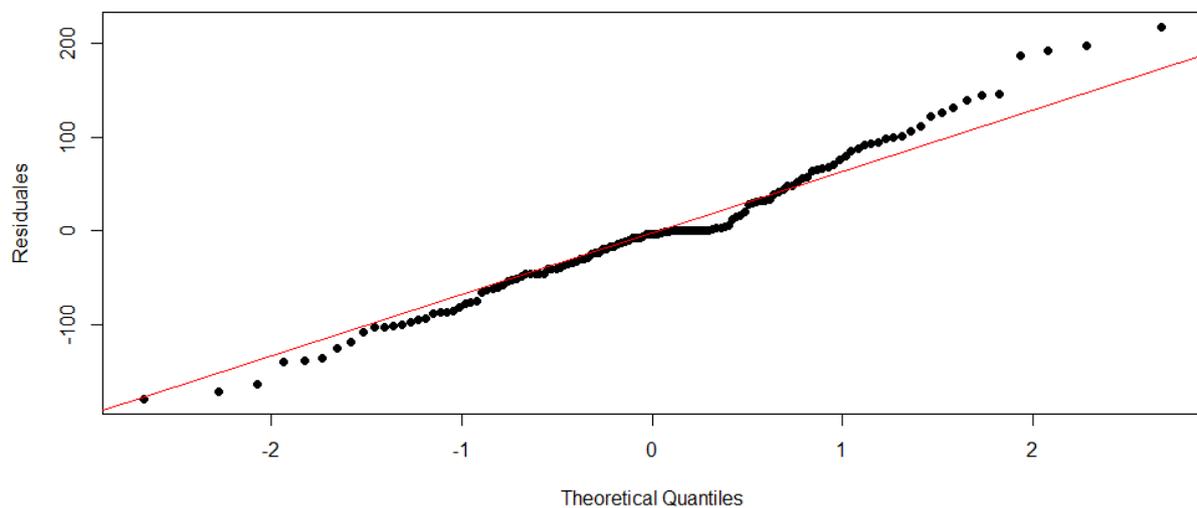
Histograma de los residuos del modelo SARIMA (3,1,0)(1,1,2)₁₂



En el histograma se observa que los residuos del modelo $SARIMA (3,1,0)(1,1,2)_{12}$, tienden a seguir una distribución normal.

Figura 61

Gráfica Q-Q normal de los residuos del modelo SARIMA (3,1,0)(1,1,2)₁₂



En el gráfico de probabilidad normal (gráfico Q-Q) se muestra que los residuos del modelo $SARIMA (3,1,0)(1,1,2)_{12}$, tienden a seguir una línea recta, por lo cual aparentemente siguen una distribución normal.

Tabla 22

Test de Kolmogorov-Smirnov de los residuos del modelo SARIMA (3,1,0)(1,1,2)₁₂

Prueba de normalidad	Kolmogorov-Smirnov	
	Estadístico	Sig
Residuales	0.120	0.044

En la tabla 22 se aprecia el estadístico de Kolmogorov-Smirnov, el cual arroja un valor de 0.120, por lo que se determinó que las perturbaciones no tienen una distribución conjunta normal, concluyendo que los residuos del modelo *SARIMA (3,1,0)(1,1,2)₁₂* no siguen una distribución normal.

Se puede observar que todos los modelos propuestos cumplen con los requisitos, los cuales son que sus residuos se distribuyan como un ruido blanco, sean independientes y sigan una distribución normal; ahora se realizara el cálculo de sus pronósticos para el año 2018.

Tabla 23

Pronósticos del IGV actividad económica Comercio con la metodología Box y Jenkins para el año 2018

	SARIMA (0,1,1) (1,1,0)	SARIMA (1,1,0) (1,1,0)	SARIMA (0,1,1) (1,1,2)	SARIMA (2,1,0) (0,1,1)	SARIMA (2,1,0) (1,1,0)	SARIMA (2,1,0) (1,1,2)	SARIMA (3,1,0) (0,1,1)	SARIMA (3,1,0) (1,1,0)
Ene-18	3442.147	3496.125	3529.839	3522.274	3503.733	3570.491	3483.587	3470.906
Feb-18	2645.819	2694.154	2678.853	2734.239	2712.135	2732.746	2714.387	2701.919
Mar-18	2567.308	2622.708	2571.785	2632.030	2626.761	2598.423	2631.811	2623.624
Abr-18	2608.053	2665.037	2602.745	2675.830	2674.070	2636.340	2655.565	2653.769
May-18	2634.230	2687.785	2658.112	2718.208	2697.896	2698.234	2693.696	2679.821
Jun-18	2610.420	2662.560	2621.962	2697.324	2671.469	2656.316	2681.583	2660.877
Jul-18	2778.930	2824.049	2761.452	2850.203	2835.074	2786.547	2840.578	2831.256
Ago-18	2893.629	2947.765	2900.299	2944.634	2957.014	2935.888	2930.015	2941.950
Set-18	2770.177	2824.193	2760.379	2825.851	2833.173	2792.381	2809.759	2818.461
Oct-18	2873.627	2919.534	2853.667	2932.627	2930.392	2879.064	2921.314	2924.718
Nov-18	2904.196	2946.807	2892.725	2978.777	2958.425	2915.450	2968.188	2956.108
Dic-18	2923.810	2968.666	2945.253	3012.649	2979.743	2974.327	2998.577	2974.863

En la tabla 23 se muestran los pronósticos del IGV actividad económica Comercio para el año 2018, de acuerdo a los modelos que describen la serie original.

Tabla 24*Comparación de la evaluación de los errores de pronósticos de los modelos propuestos*

	RMSE	MAE	MAPE	AIC	BIC
SARIMA (0,1,1) (1,1,0)	79.96202	61.06305	2.973431	1402.33	1410.66
SARIMA (1,1,0) (1,1,0)	87.22676	65.8711	3.229383	1422.91	1431.25
SARIMA (0,1,1) (1,1,2)	77.07083	58.00843	2.833775	1402.08	1415.98
SARIMA (2,1,0) (0,1,1)	83.42134	63.97243	3.146545	1414.21	1425.33
SARIMA (2,1,0) (1,1,0)	83.11222	62.85222	3.084072	1413.46	1424.57
SARIMA (2,1,0) (1,1,2)	79.2035	59.91093	2.939881	1411.75	1428.43
SARIMA (3,1,0) (0,1,1)	80.88531	61.67269	3.039521	1408.7	1422.59
SARIMA (3,1,0) (1,1,0)	81.25254	62.10175	3.048636	1409.6	1423.49

En la tabla 24 se observa la evaluación de los errores de pronósticos de todos los modelos que se ajustan a la serie original, a través de la cual se observa que el modelo $SARIMA(0,1,1)(1,1,2)_{12}$ es el modelo que tiene menor error y buenos criterios de información como son los de Akaike y Bayesiano.

SARIMA (0,1,1) (1,1,2)₁₂

Figura 62

Pronósticos con el modelo SARIMA (0,1,1)(1,1,2)₁₂ del IGV actividad económica Comercio para el año 2018

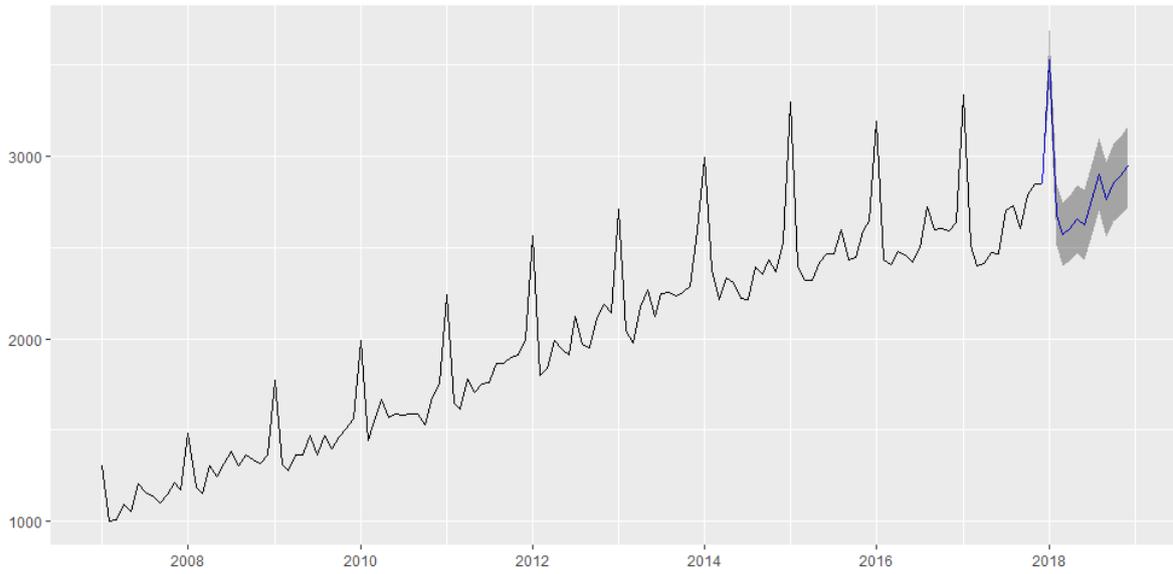


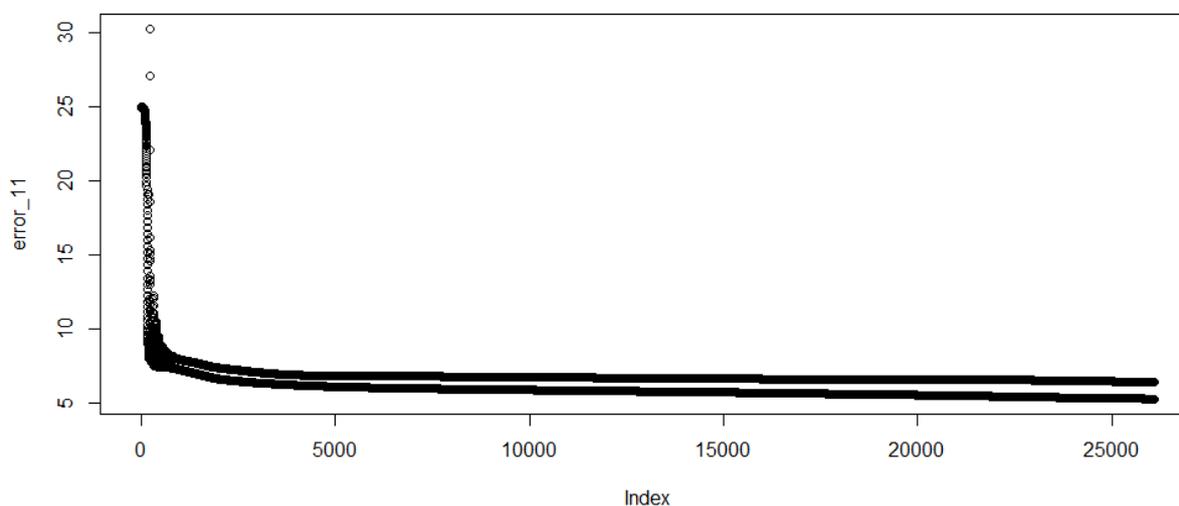
Tabla 25

Información sobre la red neuronal artificial - MPL, aplicado a los Ingresos Tributarios por Impuesto General a la Ventas a nivel nacional.

Aprendizaje	0.05
Nodos en la Capa de Entrada	12
Nodos en la Capa Oculta	18, 2
Nodos en la Capa de Salida	1
Dimensiones de la Región (Sinapsis)	12, 18, 2, 1
Entradas	120
Función de Activación: sigmoidea	Logistic
Tipo de Red	Rnn
Valores de Salida (nodos)	1
Número de Épocas	26087
Regla de Aprendizaje Sigmoidea	Sgd
Épocas actuales	26087

Figura 64

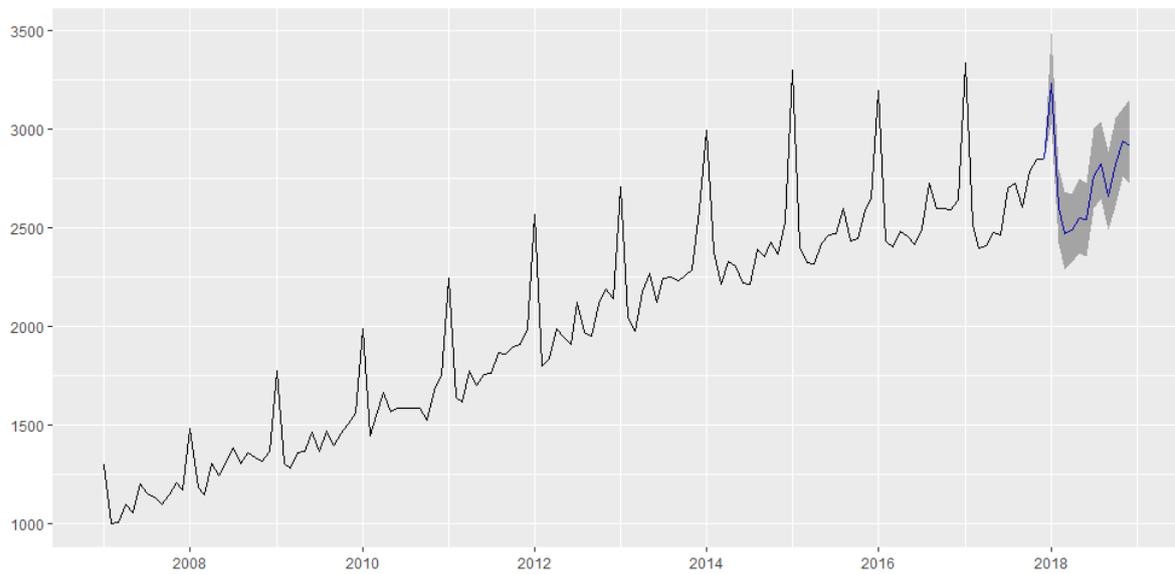
Evolución del Error de Generalización de la red neuronal



Se aprecia que el modelo logró aprender y finalizar su funcionamiento con el menor error de generalización posible, en el proceso se observó un incremento del error referenciado en las fases iniciales, pese a esto la red se repuso ajustando sus pesos sinápticos a través de las épocas o ciclos previamente definidos logrando quedarse con la mejor posibilidad.

Figura 65

Pronóstico del IGV actividad económica Comercio, mediante redes neuronales, para el año 2018



En la figura 65 se observa los pronósticos calculados para el año 2018 utilizando redes neuronales artificiales, estos pronósticos siguen el comportamiento de los años anteriores, lo que demuestra que el procedimiento utilizado y los pronósticos obtenidos son adecuados.

3.4 Estimación a través de las diferentes metodologías utilizadas para la predicción de los ingresos tributarios para el año 2018

Tabla 26

Pronósticos del IGV actividad económica Comercio para el año 2018, usando la metodología de Box y Jenkins y redes neuronales

Fecha	SARIMA (0,1,1) (1,1,2)	Límite inferior 95%	Límite superior 95%	Redes Neuronales	Límite inferior 95%	Límite superior 95%
Ene-18	3529.839	3367.948	3691.729	3235.696	3026.595	3488.426
Feb-18	2678.853	2510.929	2846.778	2602.413	2420.369	2798.799
Mar-18	2571.785	2398.036	2745.534	2468.965	2285.951	2683.304
Abr-18	2602.745	2423.36	2782.129	2491.160	2326.327	2674.809
May-18	2658.112	2473.263	2842.96	2550.939	2371.640	2752.413
Jun-18	2621.962	2431.807	2812.118	2544.306	2356.205	2724.276
Jul-18	2761.452	2566.134	2956.77	2765.948	2596.005	3003.347
Ago-18	2900.299	2699.951	3100.647	2825.759	2648.983	3037.578
Sep-18	2760.379	2555.124	2965.633	2661.748	2488.919	2887.961
Oct-18	2853.667	2643.62	3063.713	2826.097	2610.665	3056.954
Nov-18	2892.725	2677.993	3107.457	2937.569	2762.800	3107.474
Dic-18	2945.253	2725.936	3164.57	2913.694	2721.487	3152.350

En la tabla 26 se aprecia los valores pronosticados para el IGV actividad económica Comercio, para el año 2018; también se presentan los límites de confianza al 95% para cada mes pronosticado.

3.5 Identificar el mejor modelo de series de tiempo para pronosticar los ingresos tributarios para el año 2018

Tabla 27

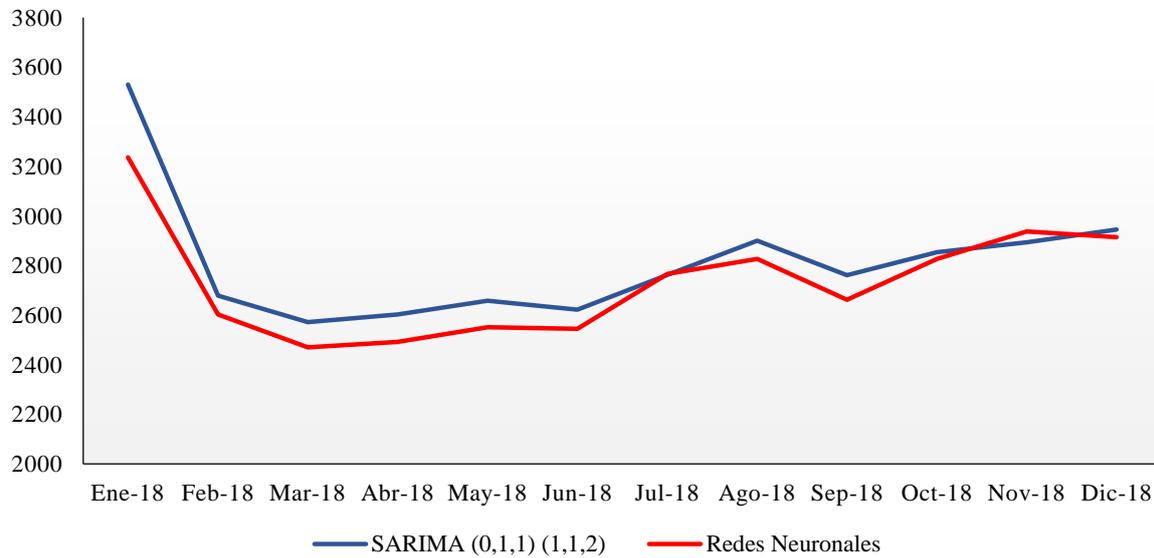
Comparación de los modelos de redes neuronales y SARIMA

	RMSE	MAE	MAPE	MASE
SARIMA (0,1,1) (1,1,2)	77.071	58.008	2.834	0.877
RNA	82.309	66.164	3.222	0.413

En la presente tabla se muestra los indicadores para identificar el mejor modelo, estos son la raíz del error cuadrático medio (RMSE), para el modelo de redes neuronales es 82.309 millones de soles y para el modelo SARIMA es 77.071. El MAPE, error absoluto porcentual promedio, para redes neuronales es 3.22% y para el modelo SARIMA es 2.83%, siendo este último mucho menor que el modelo de redes neuronales.

Figura 66

Pronóstico del IGV actividad económica Comercio, mediante el modelo SARIMA y RNA, para el año 2018



En esta figura se observa la comparación grafica de los pronósticos obtenidos con el modelo SARIMA y con RNA, ambos valores pronosticados se ajustan a la serie original, pero los pronósticos obtenidos con la metodología de Box y Jenkins son los que mejor se ajustan.

CAPÍTULO IV. CONCLUSIONES

1. La serie presenta tendencia creciente y variabilidad a través del tiempo por lo cual la serie no es estacionaria.
2. El modelo estimado mediante la metodología de Box y Jenkins para el IGV según actividad económica Comercio, enero 2007 – diciembre 2017, fue el modelo *SARIMA* $(0, 1, 1) (1, 1, 2)_{12}$.
3. A través de la metodología de redes neuronales artificiales (RNA), para el pronóstico del IGV según actividad económica Comercio, enero 2007 – diciembre 2017; el modelo es *NNAR* $(2, 1, 2)_{12}$; donde 2 son las covariables, 1 capa oculta, 2 nodos en la capa oculta.
4. Con la metodología de Box y Jenkins se estima que para el año 2018, habrá 33777.071 millones de soles de IGV actividad económica comercio; mientras que con RNA habrá 32824.294 millones de soles.
5. El modelo obtenido a través de la metodología de Box y Jenkins es el modelo que mejor se ajusta a la serie IGV actividad económica comercio, por tener menor error cuadrático medio.

REFERENCIAS BIBLIOGRÁFICAS

1. Alam, M., Kumar, M., Chowdhury, M., Siddique, A. (2015). An Algorithm for Training Multilayer Perceptron (MLP) For Image Reconstruction Using Neural Network without Overfitting. *International Journal of Scientific and Technology Research*,4(2), 271-275.
2. Alsmadi, M., Omar, K., y Noah, S. (2009). Backpropagation algorithm: The best algorithm among the multi-layer perceptron algorithm. *International Journal of Computer Science and Network Security*, 9(4), 378–383.
3. Barrio, D. (5 de julio de 2017). El cáncer de mama sigue siendo una preocupación en el Perú. El Comercio. Recuperado de <http://elcomercio.pe>
4. Barrio, D. (11 de julio de 2017). Cáncer: el alto riesgo de morir si naciste en el Perú. El Comercio. Recuperado de <http://elcomercio.pe>
5. Cardenas, J. (2015). Pronósticos y comparación de una serie de tiempo con cambios estructurales mediante la red neuronal artificial de retropropagación resiliente y modelos no lineales (tesis de pregrado). Universidad Nacional Mayor de San Marcos, Lima, Perú.
6. Carlos, S., Lima, A., Brito, C., Rodrigues, J., Escórcio, L., Monte, L. y Teixeira, M. (2012). *Oncología Básica*. Teresina, Brasil: FundaçãoQuixote
7. Castillo, R., & Varela, R. (2011). *Econometría práctica: Fundamentos de serie de tiempo*. Universidad Autónoma de Baja California, Baja California, México.
8. Chatfield, C., & Xing, H. (2019). *The Analysis of Time Series: An Introduction with R*. Florida, EE.UU: Chapman & Hall/CRC.
9. Correo (11 de Octubre del 2017). Más de 400 casos de cáncer en lo que va del año en Lambayeque. Correo. Recuperado de <https://diariocorreo.pe>

10. Cowpertwait, P. y Metcalfe, A. (2009). *Introductory Time Series with R*, EE.UU: Springer.
11. Cruz, I. (2007). *Pronósticos en el mercado de derivados utilizando redes neuronales y modelos ARIMA: una aplicación al Cete de 91 días en el MexDer* (tesis de maestría). Universidad Nacional Autónoma de México, Coyoacán, México.
12. Damian, M., Sandoval, N. (2018). *Modelo óptimo de pronóstico del índice mensual de producción de electricidad* (tesis de pregrado). Universidad Nacional Pedro Ruiz Gallo, Lambayeque, Perú.
13. Enders, W. (2015). *Applied econometric time series*. New Jersey, EE.UU: Wiley.
14. *Gestión* (5 de febrero de 2016). *Cáncer en Perú: El 85% de casos se detectan en estadios avanzados*. *Gestión*. Recuperado de <http://gestion.pe>
15. Gómez, W. (2010). *Pronóstico de una serie temporal usando redes neuronales* (tesis de pregrado). Universidad Tecnológica de la Mixteca, Oaxaca, México.
16. Goss, P., Lee, B., Badovinac, T., Strasser, K., Chavarri, Y., St Louis, J.,..., Azenha, G. (2013). *La planificación del control del cáncer en América Latina y el Caribe*, *Lancet Oncol*, 14, 391–436.
17. Gujarati, D. y Porter, D. (2010). *Econometría*. Santa Fe, México: McGraw-Hill.
18. Guzman, J. (2014). *Pronóstico de demanda y precios de energía en un mercado eléctrico*. (tesis de maestría). Universidad Michoacana de San Nicolás de Hidalgo, Michoacán, México.
19. Haykin, S. (1999). *Neural Network: A comprehensive foundation*. India: Pearson.
20. Hanke, J. y Wichern, D. (2010). *Pronósticos en los negocios*. México: Pearson.
21. Khare, M. y Nagendra, S. (2007). *Artificial neural networks in vehicular pollution modelling*. EE.UU: Springer.
22. Kozhan, R. (2010). *Financial Econometrics - with EVIEWS*, Dinamarca: Ventus.

23. Mercado, D., Pedraza, L. y Martínez, E. (2015). Comparison of Neural Network applied to prediction of times Series, *Prospect*, 13(2), 88-95.
24. Ministerio de Salud (2013). Análisis de la situación del cáncer en el Perú. Recuperado de <http://www.dge.gob.pe>
25. Novalés, A. (2000). *Econometría*, Madrid, España: McGraw-Hill.
26. Organización mundial de la salud (2012). *Globocan 2012: Cancer Incidence, Mortality and Prevalence Worldwide*. EU.: Organización mundial de la salud. Recuperado de <http://globocan.iarc.fr>
27. Ortiz, C. (2012). *Modelos Econométricos y de Redes Neuronales para predecir la Oferta Maderera en México: ARIMA vs NAR y ARMAX vs NARX* (tesis de maestría). Universidad Nacional Autónoma de México, Coyoacán, México.
28. Prado, R y West, M. (2010). *Time Series Modeling, Computation, and Inference*. EE.UU: CHAPMAN & HALL/CRC.
29. Roth, B. (2010). *Pautas en ONCOLOGIA Diagnóstico, tratamiento y seguimiento del cáncer*. Argentina: Instituto “Angel H. Roffo”.
30. Serrano, A., Soria, E., y Martín, J. (2009-2010). *Open courseware. Redes neuronales artificiales*. Llevado a cabo en el tercer ciclo del programa de doctorado del departamento de ingeniería electrónica. Valencia, España.
31. Shumway, R y Stoffer, D. (2011). *Time Series Analysis and Its Applications with R Examples*, EE.UU:Springer.
32. Siegel, S., yCastellan, J. (1998). *Estadística no paramétrica: aplicada a las ciencias de la conducta*, Mexico: Trillas.
33. Siegel, R., Miller, K., y, Jemal, A. (2017). *Cancer Statistics, 2017*, CA: A Cancer Journal for Clinicians, 67, 7-30.

34. Villavicencio, J. (2007). Introducción a series de tiempo. Recuperado de http://www.estadisticas.gobierno.pr/iepr/LinkClick.aspx?fileticket=4_BxecUaZmg%3D
35. Uriel, E Perió, A. (2000). Introducción al análisis de series temporales. Madrid, España: Editorial AC.
36. Wei, W. (2006). Time series analysis: Univariate and multivariate methods. EE.UU: Pearson.

APÉNDICE Y ANEXOS

ANEXO 1

Tabla 28

IGV a nivel nacional, recaudados por la SUNAT, periodo enero 2007 – diciembre 2017

MES/AÑO	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
ENERO	1301.63	1485.35	1774.14	1987.72	2243.17	2566.19	2706.76	2993.33	3299.36	3193.76	3337.98
FEBRERO	1003.15	1186.65	1307.76	1445.52	1644.21	1797.00	2042.97	2380.37	2402.03	2435.17	2512.15
MARZO	1009.37	1149.21	1280.70	1548.41	1613.81	1839.00	1978.73	2211.89	2321.62	2402.85	2397.53
ABRIL	1095.24	1307.21	1362.08	1666.79	1777.34	1991.06	2177.44	2332.34	2320.72	2480.39	2409.51
MAYO	1052.73	1244.09	1366.02	1568.76	1704.79	1944.35	2268.00	2304.51	2420.63	2456.74	2474.64
JUNIO	1204.71	1314.43	1466.51	1584.90	1755.89	1910.68	2124.99	2222.23	2466.49	2417.72	2462.72
JULIO	1155.74	1383.08	1365.55	1584.33	1762.34	2122.87	2245.75	2216.91	2468.51	2496.75	2701.18
AGOSTO	1136.49	1303.20	1467.62	1588.39	1868.22	1972.75	2253.71	2391.10	2596.88	2725.45	2726.76
SEPTIEMBRE	1100.39	1362.37	1399.67	1586.22	1862.40	1950.08	2231.40	2353.16	2432.45	2600.32	2604.62
OCTUBRE	1147.71	1334.78	1460.38	1527.08	1898.57	2115.45	2256.96	2430.21	2446.50	2601.69	2787.87
NOVIEMBRE	1208.95	1315.61	1507.87	1679.62	1910.81	2189.23	2287.49	2366.91	2587.64	2590.71	2850.92
DICIEMBRE	1169.67	1365.94	1563.51	1750.86	1987.28	2144.44	2590.04	2528.90	2646.75	2638.60	2848.43