

**PROYECCIÓN DE MOVIMIENTOS DE EFECTIVO EN CAJA COMO
MECANISMO PARA PREDECIR LA NECESIDAD DE EFECTIVO EN UNA
ENTIDAD FINANCIERA**

**MARÍA ISABEL MANTILLA VALCÁRCEL
MÓNICA VIVIANA RUBIANO MANTILLA**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICAS
BUCARAMANGA
2012**

**PROYECCIÓN DE MOVIMIENTOS DE EFECTIVO EN CAJA COMO
MECANISMO PARA PREDECIR LA NECESIDAD DE EFECTIVO EN UNA
ENTIDAD FINANCIERA**

**MARÍA ISABEL MANTILLA VALCÁRCEL
MÓNICA VIVIANA RUBIANO MANTILLA**

**Trabajo de Grado para optar al título de
ESPECIALISTA EN ESTADÍSTICA**

**Director
Ph.D. GERMAN MORENO ARENAS**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICAS
ESPECIALIZACIÓN EN ESTADÍSTICA
BUCARAMANGA**

2012

AGRADECIMIENTOS

En esta página queremos agradecer de manera muy especial a todas aquellas personas que hicieron posible la realización de esta investigación.

A Dios por iluminarnos y darnos la fuerza para no desfallecer en el camino.

A nuestras familias por darnos todo su amor, su apoyo incondicional, su paciencia, sus palabras de aliento y muy especialmente por creer en nosotras.

Al Dr. Gabriel Yáñez Canal por su invaluable contribución en la elaboración y realización de esta investigación, por su inmenso aporte en nuestra formación como especialistas en estadística y muy especialmente, por su valiosa e incondicional amistad.

A Jorge Armando Ortiz Sánchez, por su orientación y apoyo durante la realización de esta investigación.

DEDICATORIA

A nuestras Familias por su inmenso amor, su paciencia, comprensión, confianza y en especial por contribuir en la realización de nuestros sueños y nuestras metas.

TABLA DE CONTENIDO

	Pág.
INTRODUCCION	17
1. ANTECEDENTES	24
2. ANÁLISIS DESCRIPTIVO DE LOS DATOS	34
2.1 ANÁLISIS DESCRIPTIVO MENSUAL DE LA SERIE INGRESO	35
2.2 ANÁLISIS DESCRIPTIVO MENSUAL DE LA SERIE EGRESO	39
2.3 ANÁLISIS DESCRIPTIVO MENSUAL DE LA SERIE SALDOS	42
2.4 ANÁLISIS DESCRIPTIVO SEMANAL DE LA SERIE INGRESOS	46
2.5 ANÁLISIS DESCRIPTIVO SEMANAL DE LA SERIE EGRESOS	51
2.6 ANÁLISIS DESCRIPTIVO SEMANAL DE LA SERIE SALDO	57
3. SERIES DE TIEMPO	63
3.1 SERIES DE TIEMPO	63
3.2 MÉTODOS DE PRONÓSTICOS	65
3.2.1 Método de suavizamiento exponencial de Winters	65
3.2.2 Metodología Box-Jenkins	67
3.3 PROCESO DE PROMEDIO MÓVIL MA (Q)	74
3.4 PROCESOS AUTORREGRESIVOS AR (p)	76
3.5 PROCESOS MIXTOS ARMA (p, q)	77
3.6 PROCESOS ARIMA (p,d,q)	78
3.7 PROCESOS SARMA (P, Q) _s (p, q)	79
3.8 PROCESOS SARIMA (P, D, Q) _s (p, d, q)	80
4. ANÁLISIS DE MODELOS	83
4.1 IDENTIFICACIÓN DE MODELOS PARA LA SERIE INGRESOS	83
4.1.1 Análisis de la Funciones de Autocorrelación muestral (FAC) y de Autocorrelación Parcial (FACP) para la serie Ingreso	84
4.1.2 Identificación de modelos para la serie Ingreso	87
4.2 IDENTIFICACIÓN DE MODELOS PARA LA SERIE EGRESOS	88

4.2.1 Análisis de la Funciones de Autocorrelación muestral (FAC) y de Autocorrelación Parcial (FACP) para la serie Egreso	89
4.2.2 Identificación de modelos para la serie Egresos	91
4.3 ANÁLISIS DE MODELOS PARA LA SERIE INGRESO SEMANAL	92
4.3.1 Modelo de suavizamiento exponencial de Winters.	92
4.3.2 Modelos SARIMA	96
4.3.2.1 Modelo SARIMA $(2,1,2)_{13}(0,0,0)$	96
4.3.2.2 Modelo SARIMA $(2,1,2)_{13}(4,0,0)$.	100
4.3.2.3 Modelo SARIMA $(2,1,2)_{13}(0,0,4)$.	105
4.3.2.4 Comparación de modelos SARMA	109
4.3.2.5 Validación de las últimas cinco observaciones	110
4.3.3 Comparación de modelos: SARIMA $(2,1,2)_{13}(4,0,0)$ vs suavizamiento exponencial de Winters	111
4.3.3.1 Validación de las últimas cinco observaciones.	111
4.4 ANÁLISIS DE MODELOS PARA LA SERIE EGRESO SEMANAL	112
4.4.1 Modelo de suavizamiento exponencial de Winters	112
4.4.2 Modelo SARMA $(1,0,1)_{52}$	116
4.4.3 Comparación de modelos: SARMA $(1,0,1)_{52}$ vs suavizamiento exponencial de Winters para la serie egresos semanales	120
4.4.4 Validación de las últimas cinco observaciones	121
5. CONCLUSIONES	123
BIBLIOGRAFÍA	126

LISTA DE FIGURAS

	Pág.
Figura 1. Histograma para la serie ingresos mensual	36
Figura 2. Comparativo del comportamiento histórico para la serie ingresos mensual	39
Figura 3. Histograma para la serie egresos mensual	40
Figura 4. Comparativo del comportamiento histórico para la serie Egresos mensual	42
Figura 5. Histograma para la serie saldo mensual	43
Figura 6. Diagrama de caja para la serie saldo mensual	44
Figura 7. Comparativo del comportamiento histórico para la serie saldo mensual	45
Figura 8. Comparativo del comportamiento histórico para la serie saldo mensual	46
Figura 9. Histograma para la serie ingresos semanal	47
Figura 10. Diagrama de caja para la serie ingresos semanal	50
Figura 11. Comparativo del comportamiento histórico para la serie ingresos semanal	50
Figura 12. Gráfico de barras comparativo del comportamiento histórico para la serie ingresos semanal	51
Figura 13. Histograma para la serie egresos semanal	53
Figura 14. Diagrama de caja para la serie egresos semanal	55
Figura 15. Comparativo del comportamiento histórico para la serie Egresos semanal	56
Figura 16. Gráfico de barras comparativo del comportamiento histórico para la serie egresos semanal	56
Figura 17. Histograma para la serie saldo semanal	58
Figura 18. Diagrama de caja para la serie saldo semanal	60

Figura 19. Comparativo del comportamiento histórico para la serie saldo semanal	61
Figura 20. Comparativo del comportamiento semanal serie saldo	62
Figura 21. Comparativo del comportamiento trimestral para la serie saldo	62
Figura 22. Serie ingresos semanal (2009-2011)	84
Figura 23. Funciones de Autocorrelación muestral y parcial de la serie Ingreso	85
Figura 24. Funciones de Autocorrelación muestral y parcial de la serie D. ingreso	86
Figura 25. Comportamiento histórico para la serie egresos semanal (2009-2011)	89
Figura 26. Funciones de Autocorrelación muestral y parcial de la serie Egreso	90
Figura 27. Ajuste de la serie Ingreso usando Suavizamiento Exponencial de Winters.	92
Figura 28. Periodograma de Residuos	93
Figura 29. Contrastes de aleatoriedad de Residuos	94
Figura 30. Gráficos de Normalidad de Residuos	96
Figura 31. Ajuste de la serie Ingreso usando el modelo SARIMA(2,1,2) ₁₃ (0,0,0)	96
Figura 32. Peridograma de Residuos	98
Figura 33. Contrastes de aleatoriedad de Residuos	99
Figura 34. Gráficos de Normalidad de Residuos	100
Figura 35. Ajuste de la serie Ingreso usando el modelo SARIMA(2,1,2) ₁₃ (4,0,0)	101
Figura 36. Peridograma de Residuos	102
Figura 37. Contrastes de aleatoriedad de Residuos	103
Figura 38. Gráficos de Normalidad de Residuos	104
Figura 39. Ajuste de la serie Ingreso usando el modelo SARIMA(2,1,2) ₁₃ (4,0,0)	105
Figura 40. Peridograma de Residuos	106
Figura 41. Contrastes de aleatoriedad de Residuos	107
Figura 42. Gráficos de Normalidad de Residuos	109
Figura 43. Ajuste de la serie Egreso usando Suavizamiento Exponencial de Winters.	113

Figura 44. Peridograma de Residuos	114
Figura 45. Contrastes de aleatoriedad de Residuos	115
Figura 46. Gráficos de Normalidad de Residuos	116
Figura 47. Ajuste de la serie Egreso usando el modelo SARMA(1,0,1) ⁵²	117
Figura 48. Peridograma de Residuos	118
Figura 49. Contrastes de aleatoriedad de Residuos	119
Figura 50. Gráficos de Normalidad de Residuos	120

LISTA DE TABLAS

	Pág.
Tabla 1. Comparativo descriptivos para la serie Ingreso Mensual 2009 a 2011	36
Tabla 2. Tabla de Frecuencias para la serie ingresos mensual	37
Tabla 3. Comparativo descriptivos para la serie Egreso Mensual 2009 a 2011	39
Tabla 4. Tabla de Frecuencias para la serie egresos mensual	41
Tabla 5. Comparativo descriptivos para la serie Saldo Mensual 2009 a 2011	42
Tabla 6. Tabla de Frecuencias para la serie saldo mensual	44
Tabla 7. Descriptivos para la serie ingresos semanal	47
Tabla 8. Tabla de Frecuencias para la serie ingresos semanal	48
Tabla 9. Valores extremos para la serie ingreso semanal	49
Tabla 10. Descriptivos para la serie egresos semanal	52
Tabla 11. Tabla de Frecuencias para la serie egresos semanal	53
Tabla 12. Valores extremos para la serie egreso semanal	54
Tabla 13. Descriptivos para la serie saldo semanal	57
Tabla 14. Tabla de Frecuencias para la serie saldo semanal	59
Tabla 15. Valores extremos para la serie saldo semanal	59
Tabla 16. Coeficientes del modelo SARMA(2,1,2) ₅₂ (0,0,0) para la serie ingreso	97
Cuadro 6. Resultado del Test Shapiro-Wilk	100
Tabla 17. Coeficientes del modelo SARMA(1,0,3) ₅₂ para la serie ingreso	101
Tabla 18. Coeficientes del modelo SARMA(1,0,2) ₅₂ para la serie ingresos	106
Tabla 19. Comparación de modelos SARIMA para la serie ingreso	109
Tabla 20. Pronósticos de 5 semanas subsiguientes para cada modelo SARMA	110
Tabla 21. Comparación de modelos	111
Tabla 22. Pronósticos de 5 semanas subsiguientes para cada modelo	111
Tabla 23. Coeficientes del modelo SARMA(1,0,1) ₅₂ para la serie egresos	117
Tabla 24. Comparación de modelos para la serie egresos	121
Tabla 25. Matriz de correlaciones para los modelos de la serie egresos	121
Tabla 26. Pronósticos de 5 semanas subsiguientes para cada modelo	121

LISTA DE CUADROS

	Pág.
Cuadro 1. Resultado del Test Dickey-Fuller	87
Cuadro 2. Posibles Modelos	88
Cuadro 3. Resultado del Test Dickey-Fuller	90
Cuadro 4. Posibles Modelos	91
Cuadro 5. Resultado del Test Shapiro-Wilk	95
Cuadro 7. Resultado del Test Shapiro-Wilk	104
Cuadro 8. Resultado del Test Shapiro-Wilk	108
Cuadro 9. Resultado del Test Shapiro-Wilk	116
Cuadro 10. Resultado del Test Shapiro-Wilk	120

RESUMEN

TÍTULO: Proyección de movimientos de efectivo en caja como mecanismo para predecir la necesidad de efectivo en una entidad financiera *

AUTOR: María Isabel Mantilla Valcárcel
Mónica Viviana Rubiano Mantilla **

PALABRAS CLAVES: CLAVES Series de tiempo, Estacionariedad, Metodología Box-Jenkins, Ingresos, Egresos, Winters, SARMA

DESCRIPCIÓN O CONTENIDO

El manejo eficiente del efectivo se constituye en una de las herramientas financieras fundamentales con las que cuenta la administración de cualquier entidad para lograr sus objetivos trazados. El análisis de los ingresos y de los egresos en caja se considera importante, ya que ayuda a conocer las principales causas de las fluctuaciones del efectivo y a determinar con ayuda de métodos estadísticos los principales factores que inciden en el comportamiento de esta variable.

Esta investigación se interesa principalmente en el análisis del comportamiento del ingreso y el egreso del efectivo semanal por un periodo de tiempo específico a través del análisis descriptivo de los datos, buscando generar un modelo de pronóstico, que permita predecir con el mayor grado de confiabilidad posible, los ingresos y egresos de la entidad financiera. La metodología empleada consiste en la búsqueda de los mejores modelos para las series de ingresos y egresos, comparando el suavizamiento exponencial de Winters y los modelos SARMA identificados a través de la metodología de *Box-Jenkins (1970)*. Se comparan dichos modelos y se elige el mejor de ellos teniendo en cuenta la significancia de los parámetros en cada modelo, las estadísticas del error, el análisis de los residuos, y el menor promedio de error relativo en los pronósticos.

Posteriormente, se generan proyecciones de la tendencia de los ingresos y egresos teniendo en cuenta los modelos encontrados, con el objetivo de conocer el comportamiento futuro de dichos flujos de efectivo, lo cual permite establecer o ajustar las políticas administrativas y financieras de acuerdo con el comportamiento del mercado.

* Trabajo de grado

** Universidad Industrial de Santander. Facultad de ciencias. Licenciatura en Matemáticas. P.h.D. German Moreno

ABSTRACT

TÍTULO: Projected cash flows as a mechanism to predict the need for cash at a financial institution *

AUTHORS: María Isabel Mantilla Valcárcel
Mónica Viviana Rubiano Mantilla **

KEY WORDS: Time series, Stationarity, Box- Jenkins Methodology, Income, Expenditures, Winters, SARMA.

DESCRIPTION OR CONTENT

Efficient cash management constitutes a fundamental financial tools with which management has any entity to achieve its objectives. The analysis of income and expenses in cash is considered important, as it helps to know the main causes of fluctuations in cash and determine with the help of statistical methods the main factors affecting the behavior of this variable.

This research is primarily interested in analyzing the behavior of income and cash outflow weekly for a period of time specified by the descriptive analysis of the data, seeking to generate a forecast model, to predict with the highest degree of reliability possible , revenues and expenditures of the entity financiera. La methodology consists in finding the best models for the series of receipts and expenditures, comparing Winters exponential smoothing and SARMA models identified by the Box-Jenkins methodology (1970). We compare these models and chooses the best of them considering the significance of the parameters in each model, the error statistics, analysis of waste, and the lowest average relative error in the forecasts.

Subsequently, the trend projections generate income and expenses taking into account the models found, in order to know the future behavior of those cash flows, which allows you to set or adjust the administrative and financial policies in accordance with the behavior the market.

* Grade work

** Universidad Industrial de Santander. Faculty of Sciences. Licenciante in Mathematics. Ph.D. German Moreno

INTRODUCCION

Las finanzas y la contabilidad son áreas en las que los pronósticos se han mostrado especialmente valiosos en los últimos años. Estas áreas deben predecir los flujos de efectivo y las tasas a las cuales ocurrirían varios gastos e ingresos si pretenden mantener la liquidez y eficiencia de operación de la empresa. Así mismo, las funciones financiera y contable deben pronosticar las tasas de interés para apoyar la adquisición de capital nuevo, el pago de las cuentas por cobrar para ayudar en la planificación de las necesidades de capital de trabajo y las tasas de gasto en equipo de capital para contribuir a equilibrar el flujo de fondos de la organización.

La predicción en cuestión de finanzas en entidades financieras de ahorro y crédito, no alcanza un alto grado de exactitud debido a que se ve afectada por diversos factores como:

- Los incrementos de liquidez de los clientes, los cuales aumentan las captaciones de dinero.
- La realización de campañas masivas de cobro, por lo cual se generan temporadas de recaudo más altas que otras durante el año, por ejemplo los cierres de trimestre.
- Durante el mes de Noviembre por lo general bajan los saldos por retiros masivos previos a la temporada decembrina.
- Existen cambios generales en la economía con facilidad para la asignación de recursos económicos (aumento de las operaciones de crédito), que a su vez conlleva al desarrollo de otros procesos financieros y económicos, como el aumento de la inversión y el aumento del endeudamiento.

- Realización de campañas comerciales para lograr incrementos de saldos.
- Desembolsos masivos de créditos, ocasionan salidas de efectivo por el retiro de los mismos de las cuentas de ahorro donde son consignados.

Para reafirmar lo anteriormente dicho, Wheelwrightms y Makridakis (2006) señalan: *“...en las esferas económicas no se puede alcanzar tal grado de precisión, teniendo en cuenta que los pronósticos dependen de la identificación y extrapolación de modelos o relaciones establecidas entre los datos. Y que adicionalmente, los patrones y relaciones en economía y la empresa cambian a veces drásticamente. Aún más, las acciones de la gente pueden influir y de hecho afectan los sucesos futuros”.* (Pag.17)

El manejo eficiente del efectivo se constituye en una de las herramientas financieras fundamentales con las que cuenta la administración de cualquier entidad para lograr sus objetivos trazados. No se debe olvidar que el objetivo de la administración del efectivo es mantener la mínima inversión posible de efectivo, de modo que las entidades financieras de ahorro y préstamo pueda seguir operando sus actividades de manera eficaz y eficiente.

El departamento de tesorería de una entidad financiera tiene funciones específicas, siendo la más importante aplicar las medidas necesarias para la prevención de errores en cuanto al manejo de efectivo, la caja y los bancos, procurando al máximo mantener dinero inactivo, lo cual se refleja en el exceso de dinero en las cajas y a su vez, evitando que las agencias se vean en la necesidad de solicitar traslado de recursos que representen costos innecesarios para la entidad.

La principal fuente de recursos de la caja, proviene de los dineros captados a través de los diferentes productos y servicios ofrecidos como Cuentas de Ahorro,

CDATS (Certificado de Depósito de Ahorro a Término Fijo, orientado a inversiones a muy corto plazo, no son negociables), CDT (Certificado de Depósito a Término, orientado a inversiones a largo plazo, es negociable), convenios (Acuerdos comerciales realizados entre una empresa y una persona natural o jurídica, con el fin de ofrecer a productos y servicios que satisfagan sus necesidades), entre otros, junto con el recaudo de cartera de créditos.

Estos recursos se obtienen de los pagos o abonos de créditos vigentes e incluyen capital e intereses. Por su parte, los egresos se ven reflejados en los retiros de los clientes desde sus diferentes productos y servicios de captación. La liquidez de la empresa está compuesta por los recursos en caja y bancos que posean las agencias y Tesorería y a su vez, los recursos en caja están compuestos por el efectivo y cheques recibidos en la prestación del servicio.

Algunos de los motivos fundamentales para mantener efectivo son los siguientes:

- Transaccionales: necesidad de realizar operaciones comerciales en el “día a día”.
- Precautorio: los flujos de ingresos y egresos no son perfectamente predecibles.
- Especulativo: consiste en mantener efectivo para aprovechar oportunidades lucrativas de inversión que puedan surgir.
- Compensatorios: representa los saldos mínimos de efectivo que es conveniente mantener en la cuenta corriente bancaria o en caja.

Uno de los principales problemas dentro de este proceso radica en que no se cuenta con una medición que permita identificar con anticipación y con un

determinado nivel de confianza, tanto el dinero entrante como al saliente de la caja de la entidad, de manera que en forma semanal se pueda disponer de un valor que no sobrepase ni por exceso ni por defecto el monto requerido. Se requiere una medición semanal, dado que tanto la consecución como el transporte del efectivo requieren de un trámite dispendioso con diferentes entidades, llámese las mismas agencias, el banco con el cual se tiene la cuenta corriente o la transportadora de valores. Aún más, existen agencias que se encuentran ubicadas en municipios, que además de encontrarse alejadas de alguna ciudad, no cuentan con ninguna entidad bancaria, lo que hace necesario programar el envío o recolección del efectivo utilizando vuelos chárter.

Actualmente, en la entidad financiera objeto de esta investigación, para identificar la escasez o el excedente de recursos para el cumplimiento de compromisos, se monitorea de manual tanto los saldos en caja dentro de cada agencia como los saldos de las cuentas corrientes e inversiones en carteras colectivas, dado que no se cuenta con una herramienta que permita determinar el saldo final con el que cerrará cada agencia con un tiempo prudencial de anticipación, para de esta forma poder determinar los requerimientos o excedentes futuros de efectivo.

Se considera excedente de efectivo, cuando el valor recaudado por la agencia sobrepasa los topes asegurados, poniendo en riesgo el dinero en efectivo no cubierto. En estos casos, se informa al departamento de tesorería con el objeto de obtener provecho de estos dineros, bien sea para realizar inversiones o realizar traslados entre agencias de acuerdo a las opciones que más le convengan a la entidad. Todo lo anterior dependerá de las necesidades de efectivo inmediatas que presenten las agencias y la rentabilidad de las inversiones que se puedan constituir.

Por su parte, la escasez de efectivo en las agencias hace referencia a la falta de recursos para atender las transacciones en efectivo que se prevé serán realizadas

por los clientes, como son retiros de cuentas de ahorro, vencimientos de CDAT, entre otras. Cuando se presenta escasez de efectivo, se debe informar al departamento de tesorería, con el objeto de obtener recursos que permitan superarla.

En el momento en que una agencia se enfrenta a una situación de escasez de efectivo debe solicitarlo inmediatamente. Esto se puede realizar de dos formas diferentes:

- La primera consiste en solicitar dinero a alguna agencia que posea exceso de liquidez en ese momento, lo que ocasiona un sobrecosto por concepto del pago de la transportadora de valores.
- La segunda consiste en entregar un cheque al cliente para que lo haga efectivo en el banco con el cual la entidad financiera tenga convenio, obligando a la entidad a asumir el gravamen del cuatro por mil.

Adicionalmente a lo anterior se debe asumir el riesgo reputacional por no poder atender eficientemente los requerimientos de liquidez de los clientes ante la falta de los recursos en efectivo necesarios para la correcta operación del negocio.

Conscientes de esta problemática, se plantea un procedimiento por medio del análisis de series de tiempo, que permita realizar pronósticos alrededor de la cantidad de dinero en efectivo que queda en la caja de una entidad financiera. Más concretamente, este trabajo se propuso encontrar *un modelo de series de tiempo, que permitiera, de forma eficiente y oportuna, pronosticar los valores del dinero en efectivo que entran y salen semanalmente de la caja de una entidad financiera específica.*

La metodología empleada consistió en la búsqueda de los mejores modelos de suavizamiento y modelos ARIMA, para luego compararlos y seleccionar los mejores a la luz de su capacidad de predicción. Posteriormente, se generan proyecciones de la tendencia de los ingresos y egresos teniendo en cuenta los modelos encontrados, con el objetivo de conocer el comportamiento futuro de dichos flujos de efectivo, lo cual permite establecer o ajustar las políticas administrativas y financieras de acuerdo con el comportamiento del mercado.

La información utilizada en esta investigación se recopiló de los archivos transaccionales de la oficina principal de cierta entidad financiera de Bucaramanga, desde enero del año 2009, a octubre del año 2011. En ella se tuvo en cuenta la información semanal de las transacciones débito y crédito que mueven exclusivamente efectivo en esa oficina.

Este trabajo está conformado por cinco capítulos. En el primer capítulo, *Antecedentes*, se presenta una síntesis de algunos trabajos de investigación relacionados con la obtención de modelos que se ajustan a series de tiempo financieras, que permiten pronosticar valores óptimos de dinero en efectivo en la caja de una entidad y que además respaldan esta investigación.

En el segundo capítulo, *Análisis descriptivo de los datos*, se presenta un análisis descriptivo de las series de tiempo. Se analizan las series de ingresos y egresos semanales, para posteriormente realizar el estudio de las semanas agrupadas de forma mensual para poder determinar posibles comportamientos y tendencias similares en algunos períodos de tiempo específicos.

Adicionalmente se realiza la consolidación de datos para realizar el análisis comparativo de los ingresos y egresos mensuales para cada año, es decir el saldo y obtener conclusiones acerca de su comportamiento por medio del análisis descriptivo.

En el tercer capítulo, Series de tiempo, se presenta el grupo central de conceptos y resultados alrededor de series de tiempo univariadas, la metodología Box-Jenkins y modelos ARIMA, necesarias para formular y sustentar los modelos que se ajustan a las series estudiadas y que permiten realizar pronósticos óptimos de los ingresos y los egresos de efectivo en caja.

El cuarto capítulo, Análisis de modelos presenta una descripción detallada de los modelos que son significativos tanto en sus parámetros y análisis de residuos, como en los pronósticos que se obtienen alrededor de cada una de las series. Paso a paso se muestra la forma en que fueron desarrollados cada uno de estos modelos. Se presentan los mejores modelos de suavizamiento exponencial de Winter y los mejores ARIMA, para después, bajo criterios que son descritos en el tercer capítulo, escoger el que mejor se ajusta y pronostica en cada una de las series.

En el quinto capítulo, Conclusiones, se presentan las conclusiones generales obtenidas como resultado de la investigación y el planteamiento de los diferentes modelos. Se presentan, además, algunas sugerencias de ajuste y establecimiento de políticas financieras y de inversión teniendo en cuenta el conocimiento futuro de los ingresos y egresos a través del modelo encontrado.

1. ANTECEDENTES

A continuación se citan y resumen algunos trabajos relacionados con el análisis del saldo de efectivo en caja de diferentes entidades y que de una forma u otra influyeron en el diseño de nuestra investigación.

DÍAZ VELÁSQUEZ (2003)

En la investigación realizada por Díaz (2003) acerca de la determinación de las necesidades mínimas de efectivo en una empresa productiva, el objetivo consistió en determinar el saldo mínimo de efectivo a corto plazo con que debe operar una empresa de productos lácteos a partir del análisis de su situación financiera, basándose en métodos analíticos como el análisis del ciclo de caja.

Dentro de esta investigación se estudiaron los comportamientos del ciclo del efectivo en caja de periodos anteriores, así como el estudio de razones financieras. A través de este estudio se buscaba conocer la capacidad de la empresa para hacer frente a sus obligaciones más inmediatas y proyectar la estrategia para un futuro. Dado que la actividad de la Tesorería en una empresa es de vital importancia, constituyéndose en una herramienta de dirección, la idea era buscar invertir los excedentes temporales en activos financieros y buscar formas de endeudamiento poco costosas. Es por esto que lograr el equilibrio entre los beneficios y los costos de la liquidez es una parte esencial de la gestión de la Tesorería.

Teniendo en cuenta que la correcta administración del efectivo es un objetivo vital para la liquidez de una empresa, es necesario determinar las necesidades mínimas de liquidez, para lo cual pueden utilizarse diferentes métodos, tales como los analíticos.

Dentro de los métodos analíticos se encuentra el Método Corriente o Ciclo de Caja, que se define como el tiempo que transcurre desde el momento que la empresa realiza una salida de efectivo para la compra de materias primas, hasta el momento en que se cobra por concepto de la venta del producto terminado. Adicionalmente se utiliza el análisis de los Estados Financieros, el cual normalmente se refiere al cálculo de razones para evaluar el funcionamiento pasado, presente y proyectado de la empresa. De esta manera, con el uso de los métodos mencionados anteriormente, es posible determinar el funcionamiento de la empresa en cuanto a las posiciones de efectivo se refiere.

CERA GONZÁLEZ (2010)

En la investigación de Cera González (2010), se presenta el análisis de la administración del efectivo utilizando diferentes métodos, técnicas y herramientas propias del análisis financiero, tales como indicadores, razones y matrices, para finalmente formular políticas de administración del efectivo, las cuales, de ser aplicadas correctamente, contribuyen a mejorar la situación económica y financiera de la misma.

Este proyecto se enfoca principalmente en la aplicación del Modelo Baumol, el cual enfoca los problemas bajo la suposición de un alto grado de seguridad sobre las necesidades futuras de efectivo. El modelo Baumol es posiblemente el modelo más simple para la determinación de la posición óptima de efectivo. Es un modelo matemático ejemplo por excelencia, que explica la demanda de dinero, con el fin de intercambiarlo por bienes y servicios, teniendo en cuenta a la vez el factor "especulación". Mantener dinero en líquido, es decir, en efectivo, tiene un costo de oportunidad, consistente en el monto de intereses que se deja de ganar por no tenerlo colocado en activos. Pero también un aspecto positivo, ya que tener dinero líquido, permite hacer frente a cualquier transacción de una manera rápida.

SOTO (2007)

En la investigación realizada por María Soto (2007) se aborda el problema de la gestión adecuada de los fondos que la banca privada nacional mantiene en el Banco Central del Ecuador por medio de la formulación de un modelo de programación estocástica multietapa, observando que el tamaño del modelo determinístico para el problema estocástico crece exponencialmente al aumentar el número de periodos, lo que hace a estos métodos computacionalmente inaplicables en la práctica.

Teniendo en cuenta lo anterior, Soto propone nueve algoritmos en línea para el problema, obteniéndose la mejor solución en uno de ellos. Se presentan los principales resultados numéricos de la calibración de los parámetros, así como de la evaluación del desempeño de los algoritmos. Como criterio de comparación se emplea en cada instancia el resultado obtenido por la política óptima del modelo estocástico, la cual está dada por la solución de un problema de flujo de costo mínimo.

URQUIZA HERRERA Y BELLO CABALLERO (2008)

Urquiza y Bello, presentan un artículo referente a la administración del efectivo en una fábrica. Indican que es de gran importancia el análisis de factores que determinan la necesidad de tesorería en la empresa.

En primer lugar se debe señalar el problema de la sincronización de los flujos de caja, lo que exige que se establezca un horizonte de planeación de acuerdo con el cual se estiman las entradas y salidas de caja. Se plantea trabajar bajo el esquema de Miller- Orr o Modelo estocástico.

Este modelo fue diseñado para determinar el momento y el tamaño de las transferencias entre las cuentas de efectivo y los valores negociables según el proceso de decisión. Se basa en los elementos del costo de transferencia de valores negociables a efectivo y el costo de oportunidad de mantener efectivo.

Además, trabaja con dos límites, el inferior y superior, y con un nivel deseable que será más alto en la medida en que sea mayor la variabilidad diaria de los flujos esperados.

BECERRA (2003)

En este trabajo se presentan tres modelos para ayudar al administrador financiero a determinar los saldos óptimos de efectivo de una empresa, diseñados por William Baumol, Merton Miller y Daniel Orr, y William Beranek, y se conocen con estos mismos nombres.

Baumol (1952) aplica el modelo de la orden de la cantidad económica (EOQ) utilizada en la Administración del inventario al problema de administración del efectivo, por considerar que también el efectivo es un artículo, sólo que está representado en dinero. Aunque ese trabajo de William Baumol destacó las implicaciones macroeconómicas para la teoría monetaria, reconoció las implicaciones para las finanzas de las empresas y colocó el escenario para una mayor investigación en relación con este importante tema. En esencia, Baumol reconoció las semejanzas fundamentales de los inventarios y del efectivo desde un punto de vista financiero. En el caso de los inventarios, los costos de pedido y de faltantes de inventario hacen costoso mantener inventarios a un nivel de cero, colocando solamente órdenes (pedidos) para requerimientos inmediatos. Pero también intervienen ciertos costos en el nivel del inventario donde una política óptima balanceará los costos de comportamiento contrario de pedido y de mantenimiento de los inventarios.

La situación es muy parecida con el efectivo y con los valores negociables. Existen costos de mantenimiento (en este caso de oportunidad) como los intereses perdidos cuando se mantienen fuertes saldos de efectivo para evitar los costos de hacer las transferencias. Además, el faltante de efectivo ocasiona otros costos parecidos a cuando no hay inventarios. Así entonces, igual que en la administración de los inventarios, el saldo óptimo de efectivo minimiza estos costos. El modelo tiene en cuenta los costos fijos y variables de las transacciones de inversión y de retiro y los costos de mantener saldos de efectivo en caja.

Posteriormente, en 1966, Merton Miller y Daniel Orr, analizando el modelo de Baumol y consideraron que este es aplicable solo a algún tipo de empresas que tengan ese supuesto de comportamiento de entrada del dinero, siendo más aplicable a personas (sus ingresos mensuales) que a empresas, ya que en estas últimas los cambios diarios del saldo de efectivo (flujo de caja neto, es decir, Ingresos – Egresos) pueden subir o bajar siguiendo un comportamiento o patrón irregular e impredecible, a través de un determinado período, por lo cual ampliaron el modelo de Baumol introduciendo un proceso de generación aleatoria para los cambios diarios en el saldo de efectivo.

El modelo de Miller y Orr se basa, tal como en el modelo de Baumol, en una función de costos, que incluye el costo de hacer transferencia hacia y desde el efectivo (costos fijos y variables de transacción) y el costo de oportunidad por mantener efectivo en caja.

En 1963, el especialista en finanzas William Beranek, enfoca el problema de determinar la decisión óptima entre el saldo de efectivo y los valores negociables en relación con los fondos disponibles. Según él, para el análisis de los problemas de administración de efectivo es útil considerar las salidas (desembolsos) o egresos de efectivo, como si fueran directamente controlables por la administración, realizados todos en un determinado momento, mientras que los

ingresos se deben considerar como poco controlables y realizados de manera continua a través de un período determinado.

La técnica difiere de la de Baumol, en el sentido de que incluye una distribución de probabilidades de los flujos esperados de efectivo y una función de costos para la pérdida de descuentos por pago en efectivo y para el deterioro de la reputación de crédito cuando la empresa tiene insuficientes saldos de efectivo en un determinado momento que le es requerido (costos por corto).

HIGUEREY GÓMEZ (2007)

El docente Angel Higuerey de la Universidad de los Andes de Mérida (2007), inicialmente plantea en esta investigación una definición de efectivo como el dinero de curso legal del que dispone una empresa en caja o en los saldos de las cuentas corrientes en los bancos, es decir, que el dinero en efectivo no produce ninguna rentabilidad a la empresa, de tal forma que es necesario minimizar los saldos de efectivos ociosos que no producen ningún interés, pero adicionalmente contar con el efectivo suficiente para atender las necesidades que ocasiona la empresa para su normal funcionamiento.

Algunos de los factores que el autor tiene en cuenta para saber el saldo de efectivo son los siguientes: sincronización de las entradas y salidas de efectivo, costos de los saldos insuficientes, es decir todos aquellos gastos que sean necesarios para la consecución del efectivo, los costos por mantener saldos de efectivo excesivos, que está representado en el no aprovechamiento de invertir el dinero ocioso, aunque sea temporalmente a una determinada tasa de interés, los costos de manejo y administración del efectivo y por último la incertidumbre en la determinación de los flujos de efectivo, así como sus salidas. De aquí depende el saldo precautorio o colchón de seguridad que la empresa desea mantener.

Higuerey plantea que es necesario determinar el saldo óptimo de efectivo y lo realiza aplicando varios modelos, tales como el modelo Baumol y el modelo de Miller y Orr (MO) que se explicaron anteriormente. Adicionalmente, se basa en el modelo de Orgler, quien sugiere que es posible determinar una estrategia óptima de administración del efectivo mediante la utilización de un modelo de programación lineal múltiple. La ventaja de este modelo es que permite coordinar las estrategias óptimas de administración de efectivo con las estrategias operativas de la empresa.

Finalmente sugiere algunas técnicas y estrategias de administración de efectivo. Con el fin de utilizar mejor el efectivo en las empresas, existen algunas estrategias, tales como retrasar las cuentas por pagar al mayor plazo permitido sin dañar su reputación crediticia, administrar eficientemente los inventarios y la producción y/o acelerar el cobro de las cuentas. Dentro de las técnicas se encuentran las dirigidas a las entradas de efectivo, concentrándose en acelerar las entradas de efectivo y las dirigidas a las salidas de efectivo, tendientes a retrasar el desembolso de efectivo.

MISAS, LÓPEZ E, ARANGO A Y HERNÁNDEZ A. BANCO DE LA REPÚBLICA. SUBGERENCIA DE ESTUDIOS ECONÓMICOS (2003)

En la investigación realizada por el Banco de la República, titulada: “La demanda de efectivo en Colombia: Una caja negra a la luz de las redes neuronales” se presenta un análisis realizado a la demanda de efectivo en Colombia, cuyo principal objetivo es el de mejorar los pronósticos del efectivo para el caso colombiano, reconociendo la naturaleza no lineal de la relación entre el efectivo y sus variables explicativas, como también, explorar los modelos de redes neuronales ANN (artificial neural network) como modelos no lineales. Estos últimos permiten modelar cambios de régimen, observaciones atípicas y estacionalidad.

De otro lado, no son satisfactorios los resultados obtenidos con modelos univariados de series de tiempo (ARIMA y ARIMA con intervención y transferencia) en la medida en que se han presentado elevados errores de pronóstico. La dificultad que enfrentan los modelos tradicionales para predecir la futura demanda de efectivo es un fenómeno que se ha observado en muchas economías.

Resalta además que la limitada bondad de pronóstico, se produce no solo cuando estos modelos se enfrentan a cambios estructurales, sino también, por su naturaleza lineal. Por esta razón en la literatura contemporánea se explora la posibilidad de una relación no lineal entre las variables que conforman la demanda de efectivo.

Algunas conclusiones de esta investigación se presentan a continuación:

- Existe soporte teórico y evidencia empírica cada vez mayor sobre la existencia de comportamientos no lineales en algunas series macroeconómicas. Hecho que lleva a pensar que los modelos lineales no resultan apropiados para proyectar este tipo de variables. Este documento se ha centrado en el examen empírico de una de ellas: el efectivo real. Se han estimado para la variable dos modelos lineales univariados (ARIMA) y un modelo multivariado (VEC), este último con un mayor contenido teórico, y se ha comparado su desempeño en pronóstico vs. una red neuronal autorregresiva ampliada con la variable inflación, la mejor de las diferentes redes consideradas.
- Con el fin de reducir la subjetividad, principal defecto que se ha atribuido a las redes neuronales, la red seleccionada se somete a diversas pruebas.

- En primer lugar, se realiza un test con el fin de probar la existencia de una relación no lineal entre la variable dependiente y las variables que conforman el input, en este caso efectivo real e inflación.

- Se examina, de otro lado, la sensibilidad del modelo de red neuronal ampliado con el fin de determinar la relevancia de los distintos inputs que conforman tanto la parte lineal como la no lineal. En este caso se muestra que la desconexión de las variables que conforman el componente lineal de la red no produce cambios tan dramáticos como los observados bajo la desconexión de inputs no lineales. Si se desconectan las variables comunes en las dos componentes (lineal y no lineal), los resultados obtenidos son similares a los encontrados bajo desconexión no lineal.

- Una vez la red neuronal ampliada ha sido sometida a estos procedimientos, se compara su desempeño en pronóstico con los modelos lineales. De esta comparación surge el resultado central de este trabajo: la superioridad manifiesta que en el pronóstico tiene la red neuronal ampliada sobre todos los modelos lineales considerados.

Administrar el efectivo significa controlar el flujo de entradas (provenientes principalmente de los ingresos operativos) y el flujo de salidas (representado por los egresos de efectivo). Asegurar que el negocio tenga exactamente la cantidad de efectivo necesario y nunca sufra por escasez ni por defecto, es un objetivo difícil de lograr, pero el esfuerzo mejorará la administración del negocio. Siendo el efectivo un activo tan importante y a la vez volátil, cobra valor una buena administración, que pueda determinar el punto óptimo de entradas y salidas de efectivo, a fin de realizar la mínima inversión posible en este rubro, pero reduciendo el riesgo de incurrir en un faltante de efectivo necesario para operar.

Fruto del análisis de algunos documentos desarrollados por algunos autores en diferentes escenarios de la economía, citados anteriormente, se puede determinar que existen diferentes mecanismos, instrumentos y modelos, algunos de ellos matemáticos, financieros o estadísticos, que pueden utilizarse para facilitar el proceso de determinación del volumen óptimo de efectivo entrante y saliente de la caja. En el presente trabajo se utiliza un mecanismo estadístico de series de tiempo por medio de la metodología de Box y Jenkins para predecir la cantidad de ingresos y egresos que permita administrar los recursos de efectivo de manera eficiente en una entidad financiera.

2. ANÁLISIS DESCRIPTIVO DE LOS DATOS

Para alcanzar los objetivos planteados en esta investigación se utiliza la serie semanal de los ingresos, egresos y saldo en efectivo en caja de una entidad financiera durante un periodo de aproximadamente 3 años. Cada una de las series está compuesta por 151 observaciones semanales, correspondientes a 104 semanas del año 2009 y 2010 desde los meses de Enero hasta Diciembre y 47 semanas del año 2011 pertenecientes a los meses de Enero a Octubre.

La información se recopiló teniendo en cuenta todas las transacciones diarias que se realizan en la entidad (efectivo, cheque y nota contable), donde se utilizan para el estudio aquellas que mueven exclusivamente efectivo, ya que la idea es conocer el circulante como tal. Adicionalmente, la información se clasifica en ingresos y egresos. Dentro de los ingresos se encuentran todas aquellas transacciones que ocasionan una entrada de dinero a la caja de la entidad por concepto de apertura de productos o pago de obligaciones de crédito y las consignaciones provenientes de los diferentes productos que poseen los clientes. Los egresos corresponden a todas las salidas de dinero en efectivo por concepto de cancelación de productos o por retiros de los clientes de sus diferentes cuentas.

En promedio, se registran alrededor de 7 millones de transacciones mensuales externas e internas que corresponden al total de los movimientos que se realizan dentro del funcionamiento normal de una agencia, tomando la información desde 1 de enero de 2009 al 31 de Octubre de 2011.

Para iniciar con los análisis se procedió a consolidar la información de forma semanal y mensual. Para efectos óptimos de la empresa se requiere una medición semanal, ya que como se mencionó en la presentación del trabajo, tanto el hecho de conseguir el dinero como transportarlo por medio de las empresas de valores

requieren de un trámite dispendioso con diferentes entidades que, además de ser costoso, requiere disponer de cierto número de días y la actividad financiera no puede parar en ningún momento y mucho menos por ausencia de efectivo.

A continuación se presenta el análisis descriptivo semanal y mensual de los datos para cada una de las series con el fin de conocer el comportamiento del saldo en efectivo en caja dentro de la entidad financiera y, a través de esto, poder identificar y analizar valores atípicos, temporadas bajas, temporadas altas, impacto de las campañas tanto comerciales como de cartera que realiza la entidad durante el año, las cuales impactan directamente en los valores que ingresan o salen de caja.

Adicionalmente se requiere conocer las tendencias del comportamiento de las entradas y de las salidas de efectivo en ciertos períodos del año, las cuales reflejan la dinámica de la empresa y sirven para generar futuras políticas al interior de los procesos de la entidad. Es igualmente importante conocer la tendencia de la serie de los saldos que es, en últimas, quien refleja la diferencia de las entradas y salidas de efectivo.

2.1 ANÁLISIS DESCRIPTIVO MENSUAL DE LA SERIE INGRESO

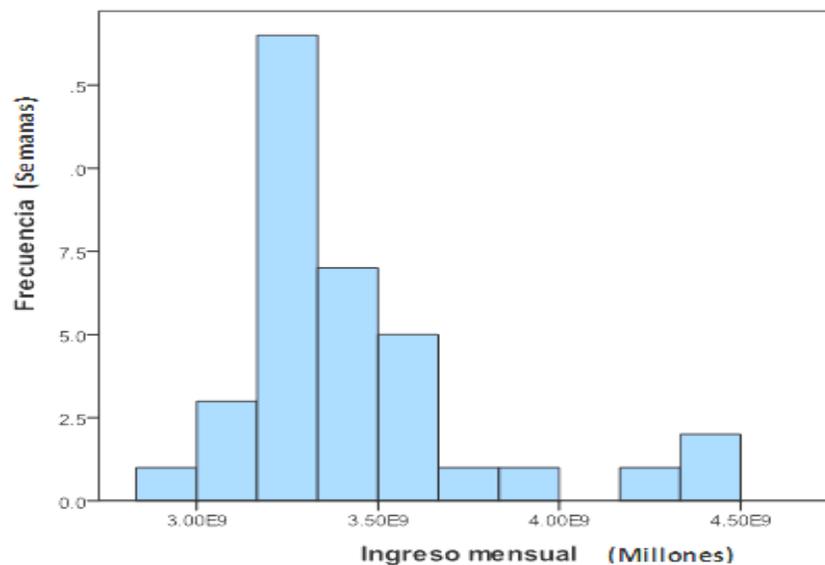
En la Tabla 1, se observa que los valores del ingreso mensual se mueven alrededor de la media, siendo esta de tres mil cuatrocientos millones de pesos (\$3.400.000.000); es de igual importancia resaltar que en el histograma de los datos (*Ver Figura 1*) se observa una ligera asimetría positiva en los datos (*coeficiente de asimetría de 1,16*), lo cual tal vez se deba a la presencia de valores fuera de los rangos normales en los cuales se mueve el saldo en efectivo. Se observa además que el valor de la mediana de tres mil trescientos treinta millones de pesos (\$3.330.676.750) es muy cercano al valor de la media, indicando como se mencionó anteriormente, una leve asimetría.

Tabla 1. Comparativo descriptivos para la serie Ingreso Mensual 2009 a 2011

Estadístico	2009	2010	2011	Total consolidado
Media	3.479.665.271,58	3.492.803.874,76	3.353.367.506,35	3.444.476.352,17
Mediana	3.371.378.948,01	3.325.748.074,14	3.315.723.444,00	3.330.676.750,81
Desv. típ.	388.730.235,14	391.751.740,44	204.198.223,16	338.749.896,35
Mínimo	2.949.755.464,48	3.131.947.284,39	3.009.216.315,91	2.949.755.464,48
Máximo	4.393.539.328,29	4.395.242.292,44	3.753.649.475,19	4.395.242.292,44
Asimetría	1,16	1,79	0,43	1,56

El histograma de la Figura 1 sustenta la idea de que los datos tienen una ligera distribución asimétrica hacia la derecha. Claramente nos muestra que hay más meses en los cuales los valores del ingreso se mueven entre los tres mil millones de pesos y los tres mil setecientos millones de pesos, mientras que hay pocos valores que se mueven entre los tres mil setecientos millones de pesos y los cuatro mil quinientos millones de pesos, razón por la cual la cola de la derecha del histograma se extiende más lejos que la cola de la izquierda.

Figura 1. Histograma para la serie ingresos mensual



Por otra parte, los ingresos mensuales presentan una variabilidad alrededor de la media de aproximadamente trescientos treinta y ocho millones de pesos (\$338.000.000). Adicionalmente se evidencia dispersión en los valores, prueba de lo cual es la diferencia existente entre el valor mínimo de efectivo (2,949 millones de pesos) y el máximo (4.395 millones de pesos).

Analizado la Tabla de Frecuencias Mensuales del total consolidado que se presenta en la Tabla 2, se deduce que el 80% de los Ingresos mensuales son inferiores a los tres mil setecientos millones de pesos (\$3.700.000.000), el 90% de los ingresos del total consolidado son inferiores a 4.000 millones de pesos. Se observa además que no hay valores entre los cuatro mil millones de pesos (\$4.000.000.000) y los cuatro mil ciento sesenta y seis millones de pesos (\$4.166.000.000) y por último un 10% son saldos superiores a los cuatro mil ciento sesenta y seis millones de pesos (\$4.166.000.000).

Tabla 2. Tabla de Frecuencias para la serie ingresos mensual

Frecuencias					
		2009	2010	2011	Total consolidado
N	Válidos	12	12	11	35
	Perdidos	0	0	0	0
Percentiles	10	2.999.062.256,65	3.172.713.947,33	3.043.803.128,35	3.124.812.945,99
	20	3.182.880.415,76	3.269.752.036,04	3.192.304.023,67	3.234.539.228,94
	30	3.270.830.872,89	3.273.658.990,85	3.237.688.122,00	3.273.366.801,44
	40	3.308.358.686,83	3.317.476.088,07	3.291.821.313,15	3.306.486.781,16
	50	3.371.378.948,01	3.325.748.074,14	3.315.723.444,00	3.330.676.750,81
	60	3.520.023.072,25	3.378.862.798,24	3.401.901.884,50	3.396.007.922,96
	70	3.635.281.069,48	3.472.634.682,11	3.428.552.915,43	3.481.551.820,02
	80	3.757.351.824,09	3.811.617.265,16	3.538.904.264,47	3.626.043.217,44
	90	4.250.822.524,40	4.335.314.761,76	3.722.079.649,11	4.028.883.531,94

En general el comportamiento de los ingresos en caja para los tres años, de acuerdo a la Figura 2, presenta un comportamiento similar durante los primeros 5 meses del año, posteriormente los saldos se incrementan para el mes de junio, nuevamente decrecen los siguientes 4 meses, para finalmente presentar un incremento finalizando año en el mes de diciembre.

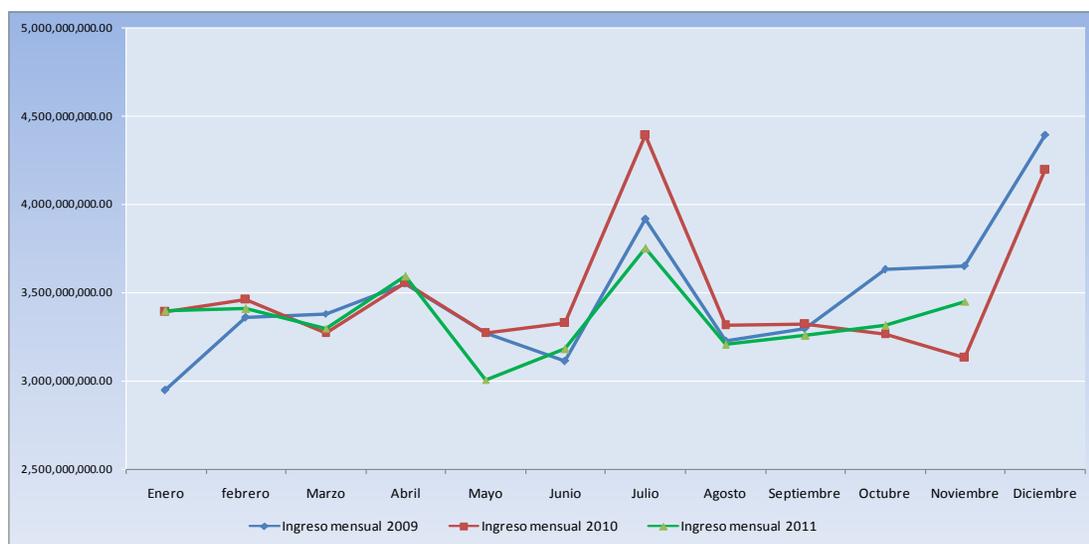
Tal comportamiento se debe a que los meses con mayores volúmenes de efectivo en caja son junio y diciembre para los 3 años, lo cual se da por cierres de semestre, con su correspondiente aumento de dinero por concepto de recuperación de cartera.

El mes de mayo, en particular, presenta los niveles más bajos de efectivo, donde se da un fenómeno de disminución de entradas de dinero por cierres de campaña de captación de efectivo, que se realizan en meses anteriores dependiendo de las necesidades de efectivo.

Dichas campañas, realizadas por la Gerencia Comercial, están encaminadas a incrementar los saldos de las diferentes cuentas de los clientes y consisten en el sorteo de elementos atractivos tales como vehículos, planes de viaje o sencillamente el aumento de las tasa de interés de los productos, a cambio que el cliente deposite sus recursos y los deje congelados en las cuentas (es decir, que no puede disponer de los mismos) durante un determinado lapso de tiempo dependiendo de las necesidades de efectivo de la entidad.

Como consecuencia de la culminación de dichas campañas, los clientes no continúan depositando masivamente dinero en sus cuentas.

Figura 2. Comparativo del comportamiento histórico para la serie ingresos mensual



2.2 ANÁLISIS DESCRIPTIVO MENSUAL DE LA SERIE EGRESO

En la Tabla 3 se observa que los valores del egreso mensual se mueven alrededor de la media, siendo esta de tres mil cien millones de pesos (\$3.100.000.000); es de igual importancia resaltar que en el histograma realizado a los datos (Ver Figura 3) se observa una leve asimetría positiva en los datos, lo cual tal vez se deba a la presencia de valores bastante fuera de los rangos normales en los cuales se mueve el saldo en efectivo, con un coeficiente de asimetría es de 1,12.

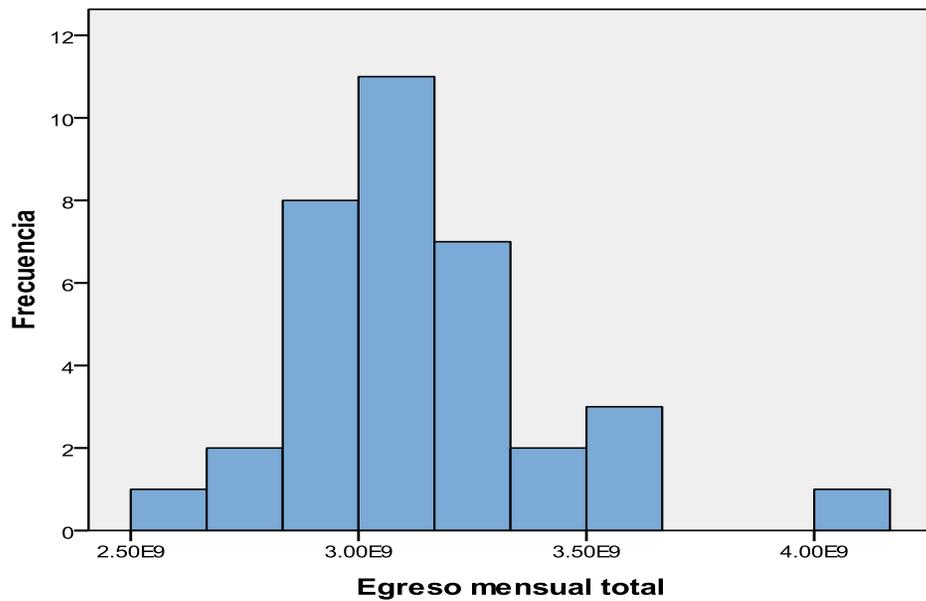
Tabla 3. Comparativo descriptivos para la serie Egreso Mensual 2009 a 2011

Estadístico	2009	2010	2011	Total consolidado
Media	3.072.948.081,17	3.262.299.850,02	3.069.771.637,46	3.136.870.376,75
Mediana	3.054.037.042,08	3.195.574.771,47	3.039.043.982,10	3.095.283.164,93
Desv. típ.	327.090.470,38	301.623.036,69	180.453.249,91	286.489.539,25
Mínimo	2.597.214.360,51	2.981.169.019,78	2.873.961.180,48	2.597.214.360,51
Máximo	3.621.994.452,09	4.078.380.986,47	3.495.345.291,90	4.078.380.986,47
Asimetría	0,48	2,05	1,44	1,12

El histograma de la figura 3, sustenta la idea de que los datos tienen una ligera distribución asimétrica hacia la derecha.

Existen mayor número de meses en los cuales los valores del egreso se mueven entre los dos mil quinientos millones de pesos y los tres mil ciento sesenta y seis millones de pesos, mientras que hay pocos valores que se mueven entre los tres mil ciento sesenta y seis millones de pesos y los cuatro mil millones de pesos.

Figura 3. Histograma para la serie egresos mensual



Analizado la Tabla de Frecuencias mensuales que se presenta en la Tabla 4, se deduce que el 30% de los egresos mensuales son inferiores a los tres mil millones de pesos (\$3.000.000.000) y un 70% son egresos entre los dos mil millones de pesos y los tres mil seiscientos millones de pesos (\$3.600.000.000).

Tabla 4. Tabla de Frecuencias para la serie egresos mensual

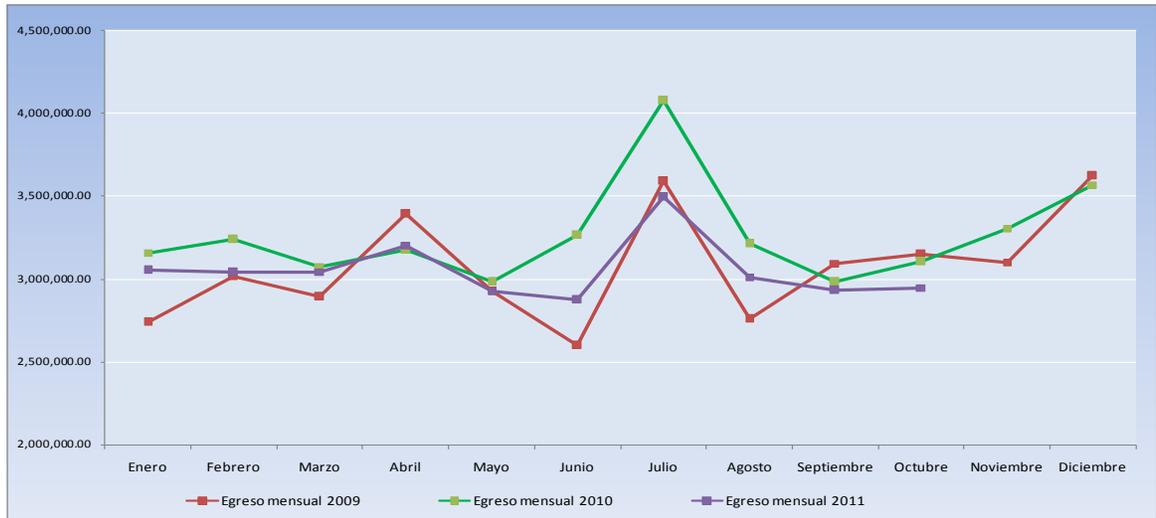
Frecuencias					
		2009	2010	2011	Total consolidado
N	Válidos	12	12	11	35
	Perdidos	0	0	0	0
Percentiles	10	2.639.560.736,44	2.981.398.948,47	2.884.675.423,08	2.827.630.670,15
	20	2.750.228.521,57	3.037.771.744,02	2.929.496.470,41	2.928.514.431,94
	30	2.878.112.425,32	3.100.476.470,92	2.941.172.522,68	2.981.782.162,95
	40	2.941.390.051,07	3.162.741.009,56	2.998.243.801,72	3.040.858.930,43
	50	3.054.037.042,08	3.195.574.771,47	3.039.043.982,10	3.095.283.164,93
	60	3.100.823.584,72	3.233.414.916,67	3.045.793.327,25	3.155.607.039,74
	70	3.174.163.325,32	3.267.747.109,95	3.113.897.241,61	3.219.401.422,67
	80	3.472.008.062,92	3.411.778.202,44	3.225.178.279,49	3.298.377.789,59
90	3.613.310.724,65	3.925.492.106,81	3.444.298.157,72	3.578.470.310,13	

Básicamente, el comportamiento de los egresos en caja para los tres años, de acuerdo a la Figura 4, presenta un comportamiento similar durante los primeros 6 meses del año, posteriormente las salidas de dinero (egresos) se incrementan para el mes de julio, nuevamente decaen para los siguientes 4 meses, para finalmente presentar un incremento finalizando año en el mes de diciembre.

Tal comportamiento se debe a que posterior al cierre de campañas de captación en el mes de mayo, se procede a descongelar los saldos de forma masiva entre la totalidad de los clientes que participaron en la campaña, por lo cual se da un fenómeno de retiro de saldos de las diferentes cuentas.

Para diciembre se observa un incremento de salidas de efectivo, debido a la temporada decembrina, donde la mayoría de los usuarios necesitan disponer de efectivo.

Figura 4. Comparativo del comportamiento histórico para la serie Egresos mensual



2.3 ANÁLISIS DESCRIPTIVO MENSUAL DE LA SERIE SALDOS

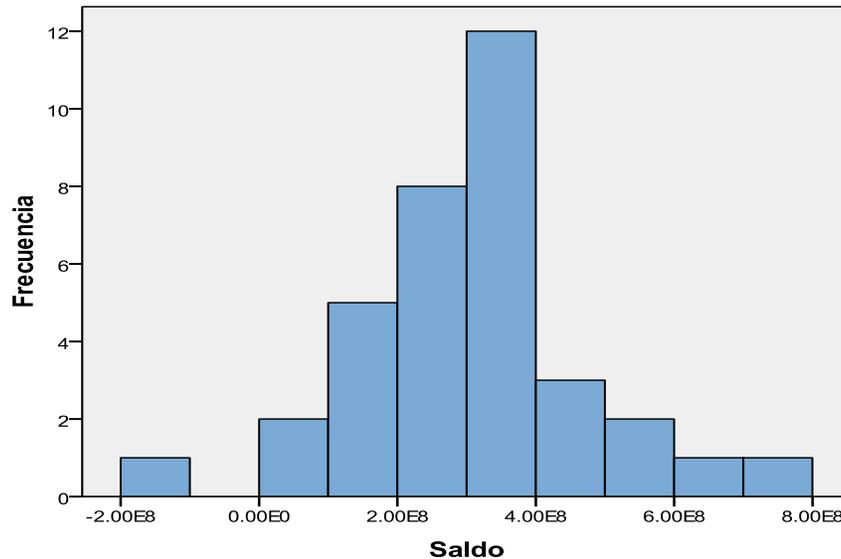
En la tabla 5 se observa que los valores del saldo mensual se mueven en torno a una media de trescientos siete millones de pesos (\$307.605.975). El saldo presenta una variabilidad alrededor de la media de aproximadamente ciento setenta y cuatro millones de pesos (\$174.435.934). Adicionalmente se observa dispersión en los datos, con un valor mínimo negativo de ciento setenta y cinco millones de pesos ((\$175.182.127)) y un máximo de setecientos setenta y un millones de pesos (\$771.544.876).

Tabla 5. Comparativo descriptivos para la serie Saldo Mensual 2009 a 2011

Estadístico	2009	2010	2011	Total consolidado
Media	406,717,190.40	230,504,024.73	283,595,868.89	307,605,975.41
Mediana	411,280,404.79	228,453,150.53	308,189,197.64	316,861,305.97
Desv. típ.	174,796,539.55	194,232,262.68	93,762,548.68	174,435,934.70
Mínimo	162,826,739.17	(175,182,127.95)	81,683,922.43	(175,182,127.95)
Máximo	771,544,876.20	626,732,469.26	393,019,077.56	771,544,876.20
Asimetría	0.46	(0.09)	(0.94)	0.08

En el histograma de la figura 5 es posible detectar una ligera asimetría en los datos, con un coeficiente de asimetría de 0.08. Se observa además que en la mayoría de los meses el saldo se mueve entre cien y seiscientos millones de pesos.

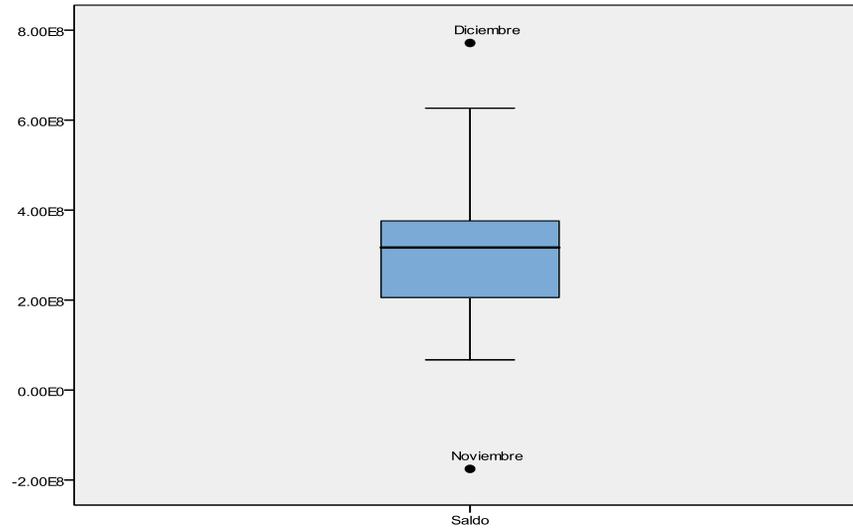
Figura 5. Histograma para la serie saldo mensual



Al observar el diagrama de caja de la figura 6, se observa como el valor de la mediana no se encuentra centrado dentro de la caja, lo cual confirma la asimetría mencionada anteriormente.

Adicionalmente se observan dos observaciones atípicas. La primera observación corresponde al mes de Noviembre (2010) la cual obedece a los saldos negativos en caja para dos semanas consecutivas durante ese mes, producto de la cancelación de dos cuentas representativas para la entidad. Por otra parte el valor atípico del mes de Diciembre (2009) corresponde al elevado volumen de transacciones efectuadas para pago de cartera en dicho año.

Figura 6. Diagrama de caja para la serie saldo mensual



Al analizar la tabla de frecuencias que se presenta en la tabla 6, se deduce que un 10% del saldo en caja es inferior a cerca de noventa cuatro millones de pesos (\$93.819.570), un 10% se encuentra entre cien y doscientos millones de pesos, un 60% presenta valores comprendidos entre doscientos y cuatrocientos cincuenta millones de pesos y tan solo un 10% son superiores a quinientos millones de pesos.

Tabla 6. Tabla de Frecuencias para la serie saldo mensual

Frecuencias					
		2009	2010	2011	Total consolidado
N	Válidos	12	12	11	35
	Perdidos	0	0	0	0
Percentiles	10	174,087,770.97	(102,435,853.89)	104,642,709.94	93,819,570.60
	20	206,977,315.26	88,068,182.53	203,266,925.20	196,973,335.88
	30	313,429,811.00	158,266,668.20	239,428,801.01	213,037,721.91
	40	347,526,508.96	204,227,877.84	257,992,877.23	257,370,265.12
	50	411,280,404.79	228,453,150.53	308,189,197.64	316,861,305.97
	60	480,973,097.08	278,205,786.99	329,231,519.69	342,413,043.69
	70	494,586,825.02	319,063,570.25	354,351,849.56	369,552,260.24

Frecuencias					
		2009	2010	2011	Total consolidado
	80	529,404,880.61	355,045,136.88	371,194,853.93	455,077,012.34
	90	704,531,388.75	552,498,804.21	388,982,751.57	529,404,880.61

Al analizar el saldo durante los tres años de estudio (*Ver figura 7*) se observa que presenta comportamientos muy diferentes y volátiles. Lo anterior es normal teniendo en cuenta que tanto las entradas como las salidas de efectivo actualmente no cuentan con un instrumento que permita monitorear permanentemente su comportamiento, para poder pronosticar el volumen de saldo de efectivo que quedará en la caja de la entidad financiera objeto de análisis.

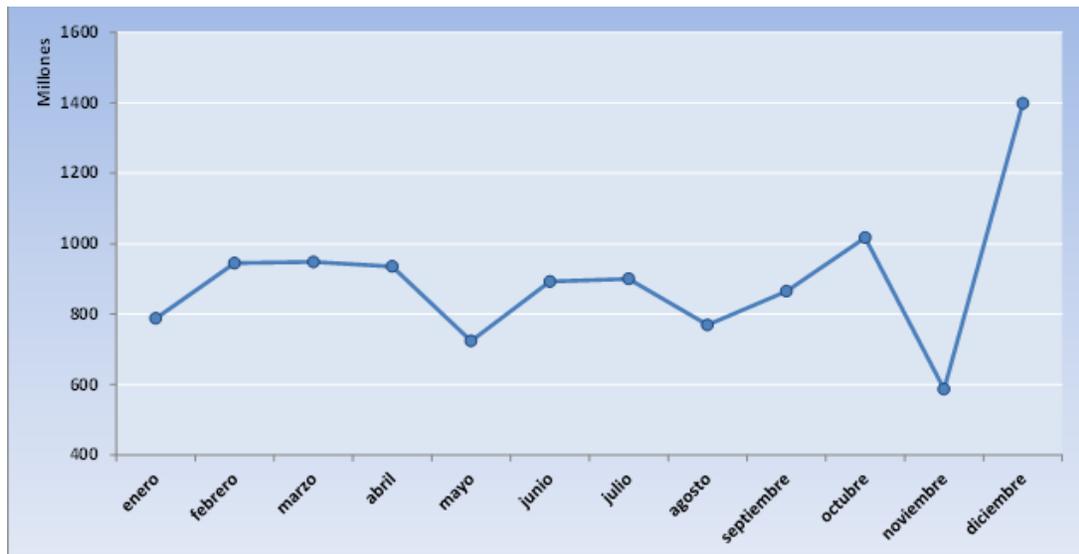
Figura 7. Comparativo del comportamiento histórico para la serie saldo mensual



En la figura 8 se observa el comportamiento de cada uno de los meses en estudio consolidados para cada año, evidenciando que existen periodos de tiempo donde el saldo se mantiene estable, sin embargo para los meses de mayo y en especial noviembre se aprecia una considerable disminución en el saldo, lo cual no se

traduce en situaciones de exposición para la empresa, sin embargo, el elevado aumento en el saldo de diciembre es importante tenerlo en cuenta, atendiendo a las razones expuestas en la presentación, por los sobrecostos que conlleva el exceso de efectivo.

Figura 8. Comparativo del comportamiento histórico para la serie saldo mensual



2.4 ANÁLISIS DESCRIPTIVO SEMANAL DE LA SERIE INGRESOS

En la Tabla 7 se observa que los valores del efectivo semanal se mueven alrededor de la media, siendo esta de ochocientos millones de pesos (\$800.000.000); es de igual importancia resaltar que en el histograma realizado a los datos (Ver Figura 7) se observa una distribución simétrica de los datos, con un coeficiente de asimetría de 0.23. Lo anterior se puede confirmar teniendo en cuenta que tanto el valor de la media (\$798.388.558) como el valor de la mediana (\$783.388.558) se encuentran bastante cercanos.

Puede observarse que los ingresos semanales presentaron valores muy cercanos para cada uno de los años de estudio, con una variabilidad alrededor de la media

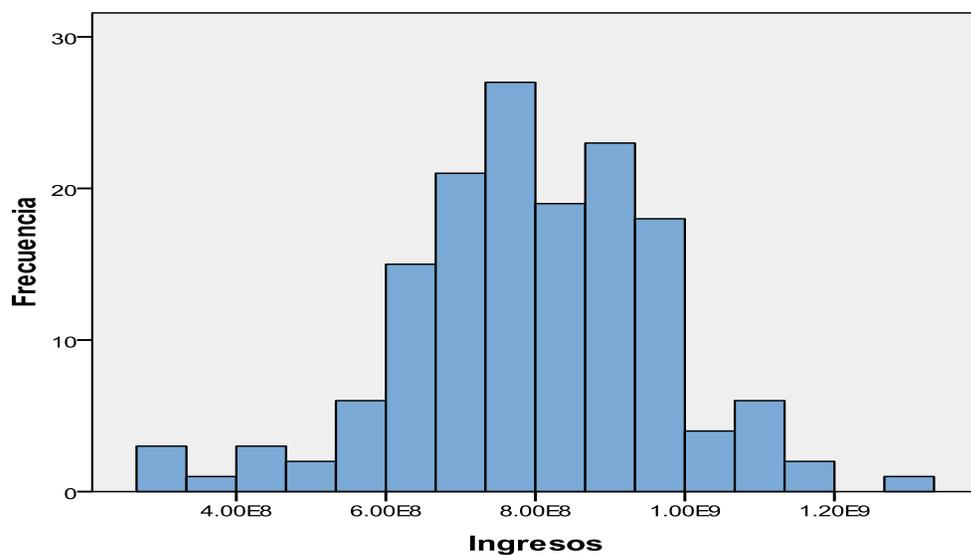
de aproximadamente ciento setenta y seis millones de pesos (\$176.000.000). Adicionalmente se evidencia una elevada dispersión en los valores, prueba de lo cual es la considerable diferencia existente entre el valor mínimo de efectivo (293 millones de pesos) y el máximo (1.294 millones de pesos).

Tabla 7. Descriptivos para la serie ingresos semanal

Estadístico	2009	2010	2011	Total consolidado
Media	802999678,1	806031663,4	784830693	798388558,5
Mediana	782446535,3	815318350,5	782332579,7	783776344,6
Desv. típ.	208928027,1	175543552	137685684,2	176684466
Mínimo	293366121,5	297328026,2	368202910,2	293366121,5
Máximo	1294448344	1189652334	1012365987	1294448344
Asimetría	-0,11	-0,41	-0,58	-0,23

El histograma de la Figura 9, sustenta la idea de que los datos tienen una distribución simétrica. Claramente se observa que la mayoría de las semanas se mueven alrededor de la media entre seiscientos cincuenta millones y los novecientos cincuenta millones de pesos.

Figura 9. Histograma para la serie ingresos semanal



Al analizar la información contenida en la Tabla 8 (se deduce que el 20% de los saldos de caja son superiores a los novecientos millones de pesos (\$900.000.000); un 50% de los saldos se encuentran entre setecientos y novecientos millones de pesos y por último un 30% son saldos inferiores a setecientos millones de pesos (\$700.000.000), para los tres años en estudio.

Tabla 8. Tabla de Frecuencias para la serie ingresos semanal

Frecuencias					
		2009	2010	2011	Total consolidado
N	Válidos	52	52	47	151
	Perdidos	52	47	0	0
Percentiles	10	555.627.175,79	608.239.046,65	642.184.189,59	601.025.816,69
	20	629.441.604,18	677.255.151,56	679.862.777,66	668.000.433,51
	30	703.447.361,60	727.415.768,25	715.175.801,35	721.310.765,15
	40	756.661.460,75	755.537.679,52	747.123.844,28	755.188.469,18
	50	782.446.535,25	815.318.350,45	782.332.579,69	783.776.344,57
	60	851.470.813,07	857.692.738,48	816.137.303,72	843.774.545,47
	70	911.943.435,57	918.123.886,74	862.513.581,87	905.782.292,07
	80	985.993.017,86	937.691.160,13	928.720.342,80	942.631.244,72
90	1.096.926.232,86	991.302.794,25	966.661.268,26	997.888.991,05	

Al realizar análisis a los valores extremos (*Ver Tabla 9*), es decir, aquellos valores del efectivo en caja que se encuentran distantes del resto de los datos, se observan datos máximos en las semanas 13 (para los tres años) y 52 (para 2009 y 2010), los cuales corresponden al recaudo de cartera, entendido como el dinero cancelado por los clientes correspondiente al pago de sus obligaciones financieras contraídas.

El fuerte recaudo se debe al intensivo plan de recuperación ejecutado de forma trimestral por parte del departamento de cartera de la entidad, con el objetivo de disminuir el indicador de mora, el cual se obtiene al realizar el cociente entre la

cartera que se encuentra vencida (es decir, con días de atraso en el pago) y el total de la cartera de la entidad.

Dicho indicador es un fuerte referente de la solvencia de la empresa, por lo que es de vital importancia mantenerlo lo más bajo posible. Para las semanas 52, se evidencia un fenómeno similar, correspondiente al pago de cartera, esta vez para cierre de año. En el año 2009 se observa un valor alto en la semana 47, correspondiente al prepago de una obligación crediticia.

Al analizar los datos mínimos, (*Tabla 9*) se puede concluir que los ingresos más bajos se dan durante la primera semana de cada año (*valores de las semanas 1, 53 y 105*), esto se debe a un período por lo general, de poca afluencia de clientes en las agencias.

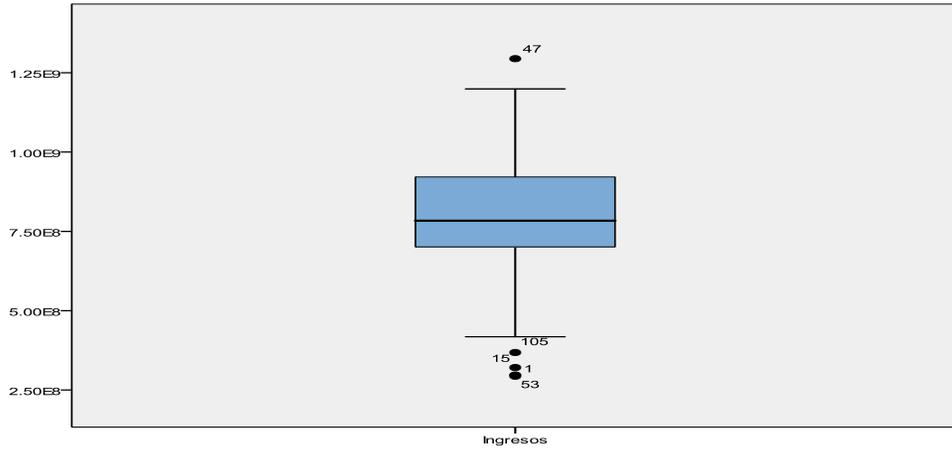
En el año 2009 se observa un valor extremo mínimo en la semana 15, correspondiente a la cancelación de un CDAT por parte de una reconocida institución educativa.

Tabla 9. Valores extremos para la serie ingreso semanal

	Semana 2009	Valor	Semana 2010	Valor	Semana 2011	Valor
Mayores	47	1.294.448.344,40	78	1.189.652.333,55	143	1.012.365.987,40
	52	1.199.056.759,23	65	1.118.865.572,10	130	999.589.106,24
	10	1.118.969.910,15	104	1.112.487.645,48	117	986.813.585,75
	43	1.117.401.659,48	57	1.106.628.569,28	110	968.333.638,35
	13	1.110.968.493,02	61	996.323.251,94	109	966.243.175,74
Menores	1	293.366.121,48	53	297.328.026,18	105	368.202.910,16
	15	320.971.280,89	54	417.851.260,46	121	466.786.663,46
	2	437.554.346,16	66	430.370.109,65	107	563.596.088,64
	16	491.746.232,76	77	589.476.473,00	133	630.904.238,01
	11	541.675.603,68	101	607.275.720,90	128	645.004.177,49

Los valores anteriormente mencionados es posible observarlos en el diagrama de caja de la Figura 10.

Figura 10. Diagrama de caja para la serie ingresos semanal



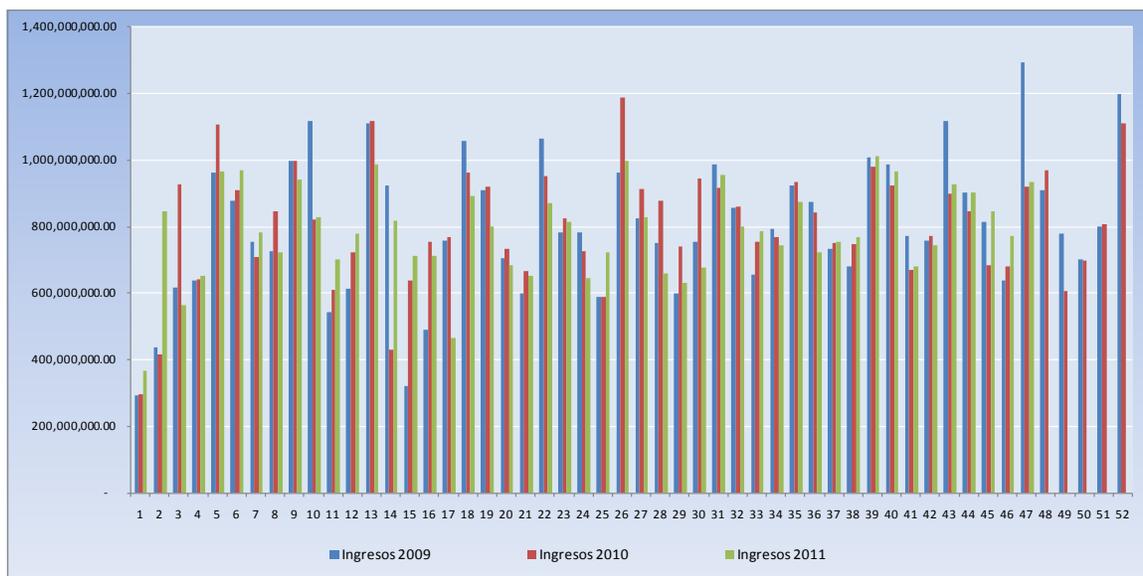
Los ingresos de efectivo en caja a lo largo de los años 2009, 2010 y 2011 (Ver Figura 11) presentan un comportamiento similar a lo largo de la mayoría de las semanas. Se observa además, que los ingresos de caja para el año 2011 han sido relativamente inferiores a los años 2009 y 2010, prueba de lo cual es la disminución tanto en el valor de la media (784 millones), como en la desviación estándar del saldo (137 millones). Lo anterior obedece a medidas internas tomadas por parte de la tesorería encaminadas a minimizar dicho saldo, en pro de la optimización de los recursos.

Figura 11. Comparativo del comportamiento histórico para la serie ingresos semanal



De acuerdo a la Figura 12, a lo largo de los años 2009, 2010 y 2011, es posible determinar que los menores ingresos de efectivo en caja se concentran en las primeras semanas de cada año, junto con las semanas que componen los días de semana santa; adicionalmente se observa incrementos de los ingresos para las semanas que corresponden a los cierres de trimestre de cartera, ocasionando incremento en los depósitos por captación de efectivo.

Figura 12. Gráfico de barras comparativo del comportamiento histórico para la serie ingresos semanal



2.5 ANÁLISIS DESCRIPTIVO SEMANAL DE LA SERIE EGRESOS

A continuación se analiza la serie de los egresos semanales. En la Tabla 10, se observa que los valores de los egresos semanales se mueven alrededor de la media, siendo esta de setecientos veintisiete millones de pesos (\$727.000.000); para el total consolidado. Puede observarse que los egresos semanales presentaron valores muy cercanos para cada uno de los años de estudio, con una variabilidad alrededor de la media de aproximadamente ciento treinta y cuatro millones de pesos (\$134.000.000). Adicionalmente se evidencia una elevada

dispersión en los valores, dada la considerable diferencia existente entre el valor mínimo (243 millones de pesos) y el máximo (998 millones de pesos).

Tabla 10. Descriptivos para la serie egresos semanal

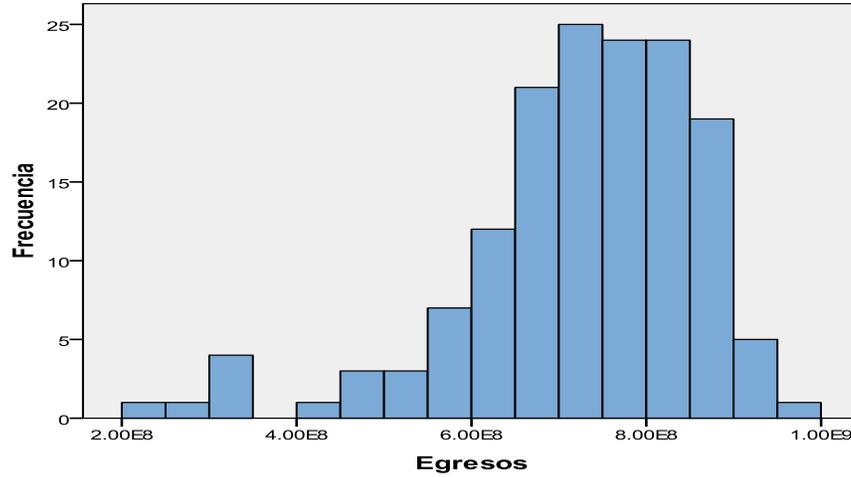
Estadístico	2009	2010	2011	Total consolidado
Media	709.141.864,89	752.838.426,93	718.457.191,75	727.089.160,17
Mediana	730.488.986,16	782.190.175,29	721.984.260,01	744.528.966,00
Desv. típ.	144.795.550,58	134.874.647,12	118.366.615,02	134.070.179,42
Mínimo	243.365.804,16	261.126.182,40	338.119.044,93	243.365.804,16
Máximo	949.942.018,69	998.023.286,28	874.125.632,00	998.023.286,28
Asimetría	-0,93	-1,59	-1,41	-1,22

En el histograma realizado a los datos (Ver Figura 13) se observa una distribución ligeramente asimétrica de los datos a la izquierda con un coeficiente de dicha asimetría de 1.22. Valores superiores a cero indican que existe este tipo de distribución.

Claramente se muestra que hay muchas más semanas en las que los valores del efectivo se mueven entre setecientos y los mil millones de pesos, mientras que hay pocos valores que se mueven entre doscientos cuarenta y setecientos millones de pesos; razón por la cual la cola de la izquierda del histograma se extiende más lejos que la cola de la derecha.

En estos casos, donde se presenta asimetría, es conveniente tomar como referente el valor de la mediana, ya que esta medida descriptiva tiene la ventaja de no verse afectada por las observaciones extremas, pues no depende de los valores que toma la variable, sino del orden de las mismas. Dicha mediana tiene un valor de setecientos treinta y cinco millones de pesos (\$735.000.000\$).

Figura 13. Histograma para la serie egresos semanal



Al analizar la información contenida en la Tabla 11, se observa que aproximadamente el 40% de los datos son inferiores a los setecientos millones de pesos (\$700.000.000), un 30% de los egresos se encuentran entre setecientos y ochocientos millones de pesos y finalmente, un 30% son saldos superiores a ochocientos millones de pesos (\$800.000.000), para los años en estudio.

Tabla 11. Tabla de Frecuencias para la serie egresos semanal

Frecuencias					
		2009	2010	2011	Total consolidado
N	Válidos	52	52	47	151
	Perdidos	0	47	0	0
Percentiles	10	543.993.004,32	597.132.233,02	591.563.299,48	566.160.975,50
	20	592.299.876,74	685.503.906,34	648.826.942,87	643.307.108,56
	30	653.403.052,66	724.047.600,51	689.056.048,67	687.322.054,19
	40	679.077.781,10	749.773.755,77	702.834.448,30	714.900.673,56
	50	730.488.986,16	782.190.175,29	721.984.260,01	744.528.966,00
	60	762.692.469,96	799.707.733,70	761.963.999,36	779.235.573,36
	70	805.327.129,60	819.822.846,43	784.262.183,24	801.504.057,84
	80	833.504.122,66	853.808.033,35	801.234.300,88	833.757.794,90
	90	885.591.564,61	890.665.716,82	861.369.546,19	874.091.697,70

Referente a los valores extremos (*Tabla 12*), se observan datos mínimos los cuales corresponden a inicio de año y a los días que componen la semana santa; lo que obedece a la baja cantidad de clientes que se presentan en las agencias a realizar transacciones.

En relación con los valores mínimos, se observan algunos datos extremos que corresponden a la primera semana al iniciar cada año (valores 1, 53 y 105), donde se realizan pocas transacciones por la baja cantidad de clientes que se acercan durante este periodo a la institución.

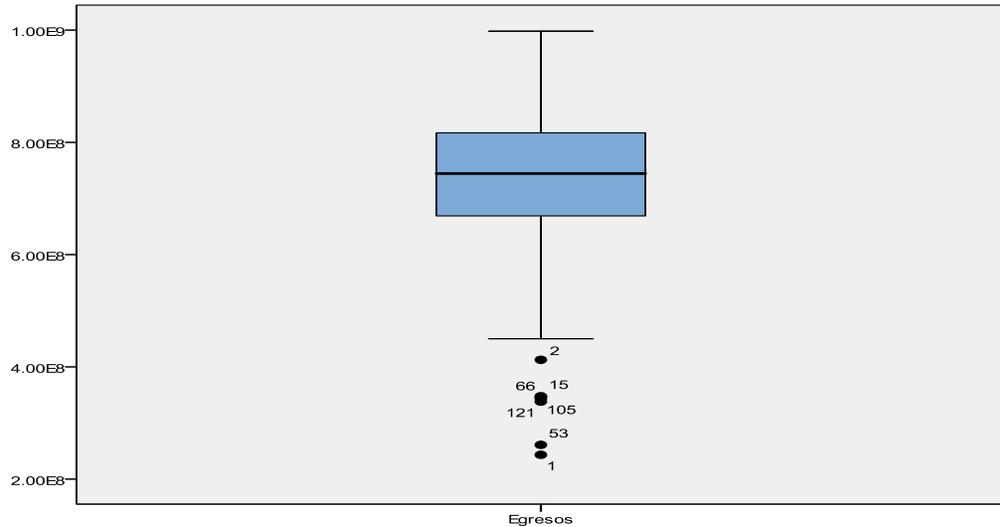
El mismo efecto sucede durante la semana que cobija la temporada de la Semana Santa (semanas 15, 66 y 121), durante las cuales se registran bajos niveles de egresos de dinero, dado que no se da atención al público desde el día jueves hasta el domingo.

Tabla 12. Valores extremos para la serie egreso semanal

	Semana 2009	Valor	Semana 2010	Valor	Semana 2011	Valor
Mayores	18	949.942.018,69	57	998.023.286,28	148	874.125.632,00
	10	915.377.246,39	78	915.394.193,55	110	873.955.960,49
	13	910.659.866,60	58	912.388.969,43	135	872.394.522,83
	40	897.326.412,40	79	896.396.387,52	151	869.525.366,00
	48	888.282.868,38	92	892.549.272,39	130	859.330.591,24
Menores	1	243.365.804,16	53	261.126.182,40	121	338.119.044,93
	15	347.751.765,89	66	346.565.678,00	105	343.337.452,67
	2	412.569.853,00	54	450.196.440,46	128	497.987.969,06
	12	499.652.084,28	104	558.949.245,21	107	527.704.499,04
	52	543.091.584,38	89	586.329.165,69	132	607.527.999,59

Los valores anteriormente mencionados es posible observarlos en el diagrama de caja de la Figura 14.

Figura 14. Diagrama de caja para la serie egresos semanal

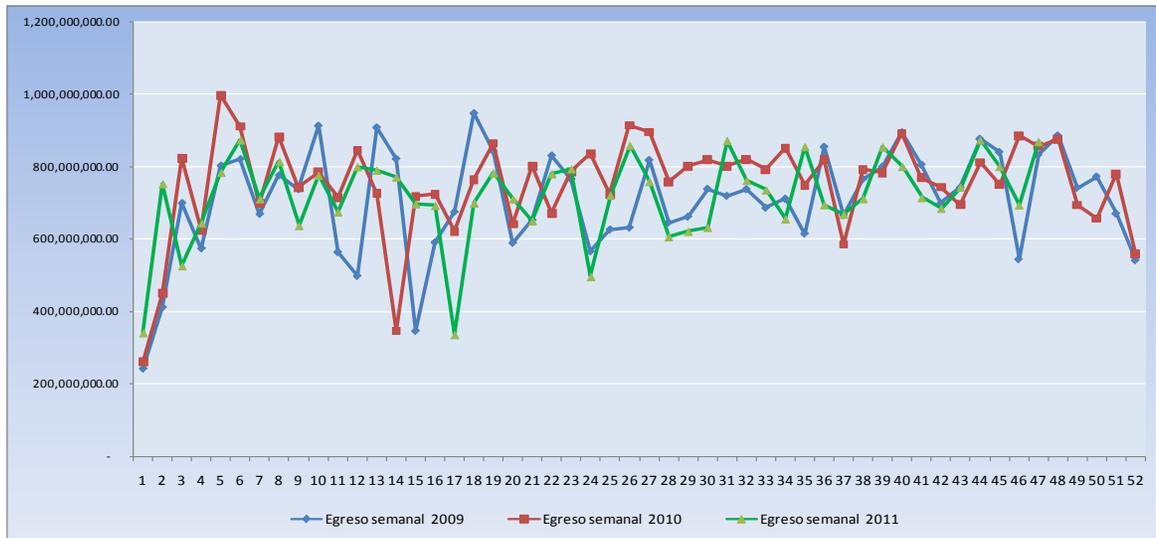


Los egresos de efectivo en caja a lo largo de los años 2009, 2010 y 2011 (Ver Figura 15) presentan un comportamiento similar a lo largo de la mayoría de las semanas, similar a lo observado con la serie ingresos.

Adicionalmente se observa, que los saldos de caja para el año 2011 han sido relativamente inferiores al año 2010, con una disminución tanto en el valor de la media (718 millones), como en la desviación del saldo (118 millones). Nuevamente, lo anterior obedece a medidas internas, en pro de la optimización de los recursos, tales como, el monitoreo del saldo diario en caja para tomar decisiones de inversión o provisión de efectivo al día siguiente.

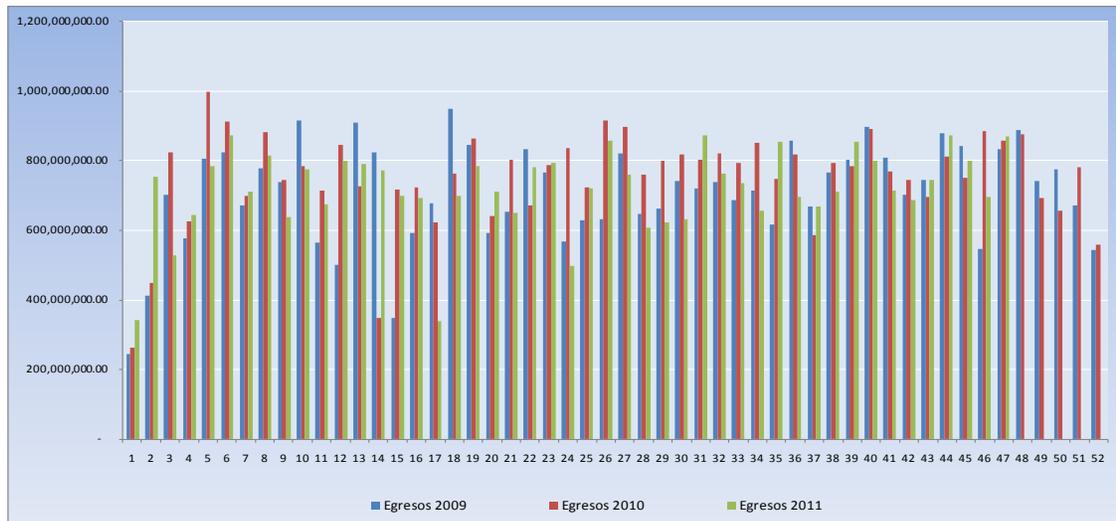
Se destaca que para los 3 años en estudio se observan descensos en dos períodos de tiempo: la primera semana de cada año y la Semana Santa. Este fenómeno es debido a que el número de transacciones disminuye para estas épocas del año, hecho que se traduce en disminución de los dineros que son retirados por parte de los clientes.

Figura 15. Comparativo del comportamiento histórico para la serie Egresos semanal



De acuerdo a la Figura 16, a lo largo de los años 2009, 2010 y 2011, es posible determinar que, al igual que en los ingresos, las menores salidas de efectivo en caja se concentran en las primeras semanas de cada año, junto con las semanas santas.

Figura 16. Gráfico de barras comparativo del comportamiento histórico para la serie egresos semanal



2.6 ANÁLISIS DESCRIPTIVO SEMANAL DE LA SERIE SALDO

En la tabla 13 se observa que los valores del saldo semanal se mueven en torno a una media de setenta y un millones de pesos (\$71.299.398). Los ingresos presentan una variabilidad alrededor de la media de aproximadamente ciento veintiséis millones de pesos (\$126.310.070). Existe adicionalmente una marcada dispersión de los datos con un valor mínimo negativo de doscientos cinco millones de pesos ((\$205.249.672)) y un máximo de setecientos cincuenta y cinco millones de pesos (\$655.965.174).

Tabla 13. Descriptivos para la serie saldo semanal

Estadístico	2009	2010	2011	Total consolidado
Media	93,857,813.17	53,193,236.48	66,373,501.23	71,299,398.27
Mediana	64,727,814.51	35,204,530.17	52,817,947.10	50,000,317.32
Desv. típ.	147,185,666.25	140,266,548.66	72,266,251.19	126,310,070.96
Mínimo	(101,150,945.80)	(205,249,672.83)	(93,134,039.83)	(205,249,672.83)
Máximo	655,965,174.85	553,538,400.27	301,389,652.44	655,965,174.85
Asimetría	1.53	1.09	0.88	1.40

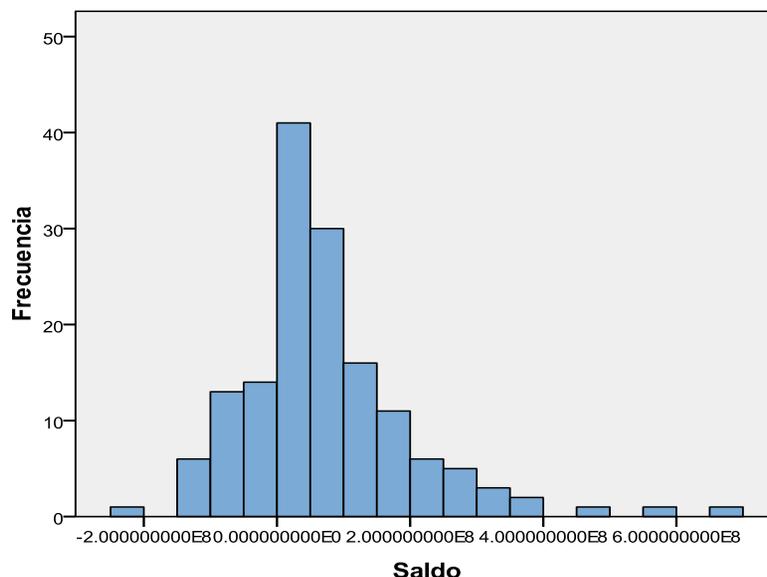
En el histograma de la figura 17 es posible detectar una ligera asimetría en los datos con un coeficiente de asimetría de 1,4. Lo anterior puede deberse a la presencia de valores fuera de los rangos normales en los cuales se mueve el saldo.

En estos casos, donde se presenta asimetría, es conveniente tomar como referente el valor de la mediana, ya que esta medida descriptiva tiene la ventaja de no verse afectada por las observaciones extremas, pues no depende de los valores que toma la variable, sino del orden de las mismas, la cual es de cincuenta millones de pesos (\$50.000.317).

Esta diferencia entre media y mediana sustenta la idea de la asimetría de los datos.

Claramente muestra que hay más meses en los cuales los valores del saldo se mueven entre los cinco y los doscientos millones de pesos, mientras que hay pocos valores que se mueven entre los doscientos millones de pesos y los seiscientos cincuenta millones de pesos, razón por la cual la cola de la derecha del histograma se extiende más lejos que la cola de la izquierda.

Figura 17. Histograma para la serie saldo semanal



Analizando la tabla de frecuencias de la tabla 14, se puede observar que un 20% del saldo semanal ha arrojado valores negativos, un 40% se encuentra entre los dieciséis y los setenta y siete millones de pesos, un 20% corresponde a montos entre cien y ciento cincuenta millones de pesos y tan solo un 10% corresponde a montos superiores a los cien millones de pesos.

Tabla 14. Tabla de Frecuencias para la serie saldo semanal

Frecuencias					
		2009	2010	2011	Total consolidado
N	Válidos	52	52	47	151
	Perdidos	52	47	0	0
Percentiles	10	(61,580,310.86)	(107,564,712.39)	(3,487,510.27)	(71,423,430.33)
	20	29,701,959.74)	(72,305,802.39)	15,568,684.29	(27,389,636.36)
	30	15,910,867.44	32,822,296.32)	25,079,136.09	16,453,945.08
	40	39,946,930.38	24,137,818.26	39,589,891.80	30,954,455.61
	50	64,727,814.51	35,204,530.17	52,817,947.10	50,000,317.32
	60	91,241,982.17	52,244,515.78	68,851,617.64	77,219,372.67
	70	112,880,047.77	103,066,932.77	87,237,531.28	101,402,454.09
	80	204,666,591.41	153,125,522.42	133,303,977.12	153,435,409.28
	90	293,419,747.61	237,333,507.18	181,218,969.16	211,917,979.10

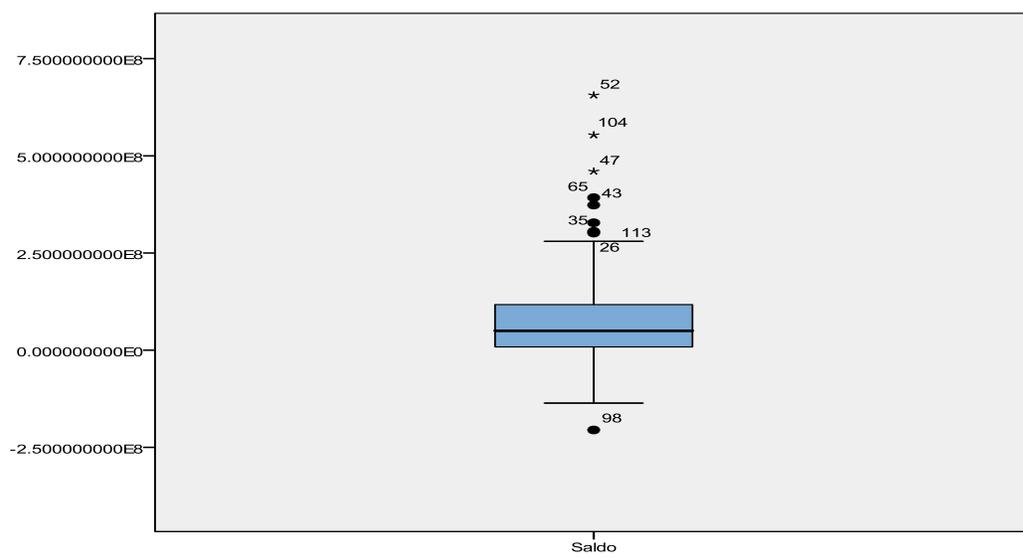
Al realizar análisis a los valores extremos que se presentan en la tabla 15, es decir aquellos valores del efectivo en caja que se encuentran distantes del resto de los datos, se observan datos máximos para diversas semanas para los tres años, los cuales corresponden al recaudo de cartera, explicado anteriormente. Al analizar los datos mínimos, no es posible determinar un patrón de comportamiento para el saldo de caja, dado que este depende de lo que ocurra tanto con los ingresos como con los egresos de efectivo.

Tabla 15. Valores extremos para la serie saldo semanal

	Semana 2009	Valor	Semana 2010	Valor	Semana 2011	Valor
Mayores	52	655,965,174.85	104	553,538,400.27	113	301,389,652.44
	47	460,183,205.00	65	392,452,598.00	117	196,026,866.00
	43	373,327,356.50	74	280,067,060.55	122	189,831,245.00
	26	327,796,999.67	78	274,258,140.00	109	181,415,157.78
	35	305,627,577.21	61	252,602,029.07	147	181,169,922.00
Menores	16	(101,150,945.80)	98	(205,249,672.83)	112	(93,134,039.83)
	38	(87,099,131.00)	73	(135,973,367.26)	145	(33,040,637.20)
	3	(84,748,055.45)	77	(135,651,048.00)	124	(27,795,737.27)
	50	(72,163,755.56)	64	(125,348,298.02)	116	(19,496,703.03)
	29	(64,134,718.66)	76	(108,506,801.76)	129	514,787.92

Sin embargo, al observar el diagrama de caja de la figura 18, se observa un valor mínimo atípico, correspondiente a Noviembre de 2010, correspondiente a la cancelación de dos cuentas representativas para la entidad, tal como se referencio anteriormente.

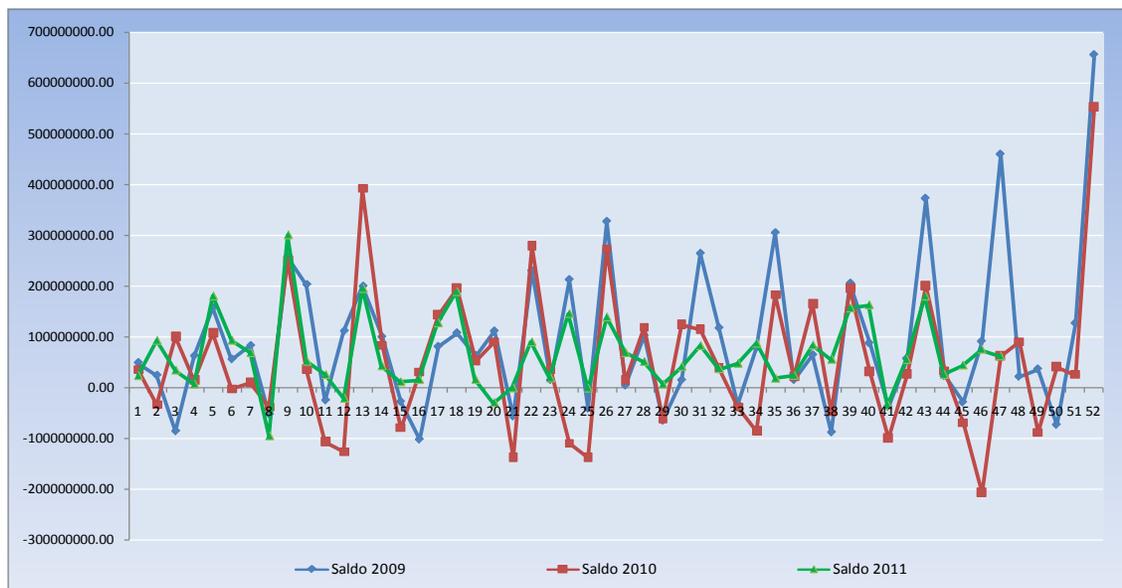
Figura 18. Diagrama de caja para la serie saldo semanal



El saldo en caja a lo largo de los años 2009, 2010 y 2011, presenta un comportamiento similar en algunos periodos de tiempo, tales como semana santa, mitad y fin de año, donde se incrementan los saldos, tal como se observa en la figura 19. Es importante observar que el 2009 fue el año donde se evidencia mayor fluctuación en los saldos, seguido de cerca por el 2010. Por otra parte, para el 2011 se observa una menor fluctuación en el movimiento de los saldos, producto de los controles que de forma manual se realizan para intentar controlar dichas oscilaciones. Prueba de lo anterior es la disminución de la desviación que se ha dado para los años 2009, 2010 y 2011 de ciento cuarenta y siete millones (\$147.185.666), ciento cuarenta millones (\$140.266.548) y setenta y dos millones (\$72.266.251) respectivamente.

El año donde se dio el mayor número de saldos negativos en caja fue el 2010, mientras que para el 2011 son pocas las semanas donde se detecta esta situación, poco conveniente para las finanzas de la entidad, por lo que representa económicamente mantener saldos negativos que deben cubrirse con los mecanismos señalados en el capítulo *Presentación*.

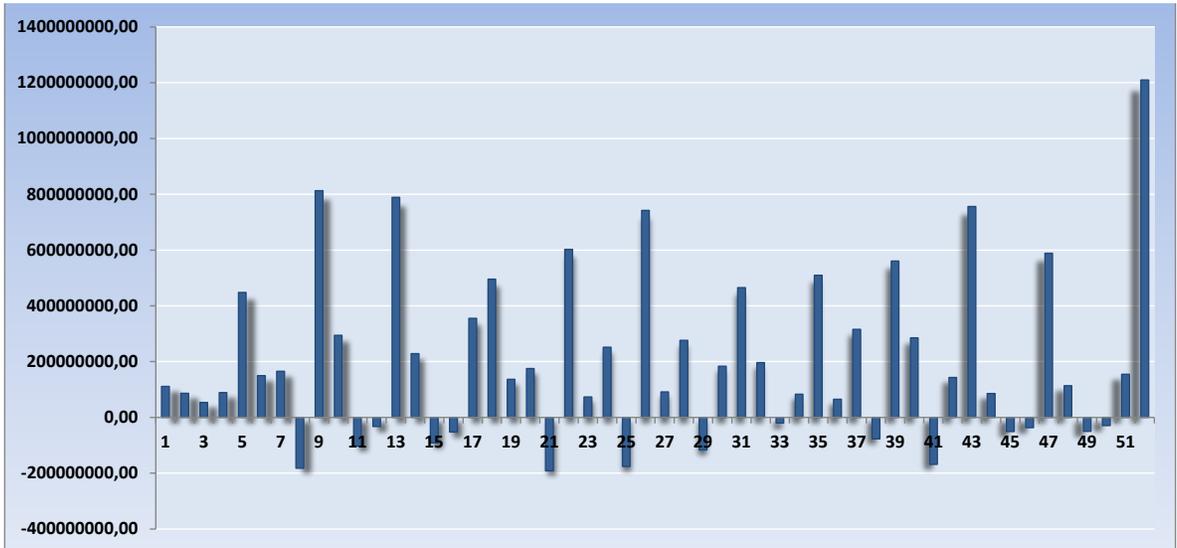
Figura 19. Comparativo del comportamiento histórico para la serie saldo semanal



La figura 20 permite observar claramente el comportamiento semanal del saldo de efectivo, donde alrededor del 29% de las semanas presentó saldo negativo, lo cual tal como se ha mencionado en repetidas oportunidades es contraproducente, debido a que esos faltantes de efectivo esporádicos es necesario compensarlos con transferencias, bien sea provenientes del banco convenio o de agencias con exceso de efectivo.

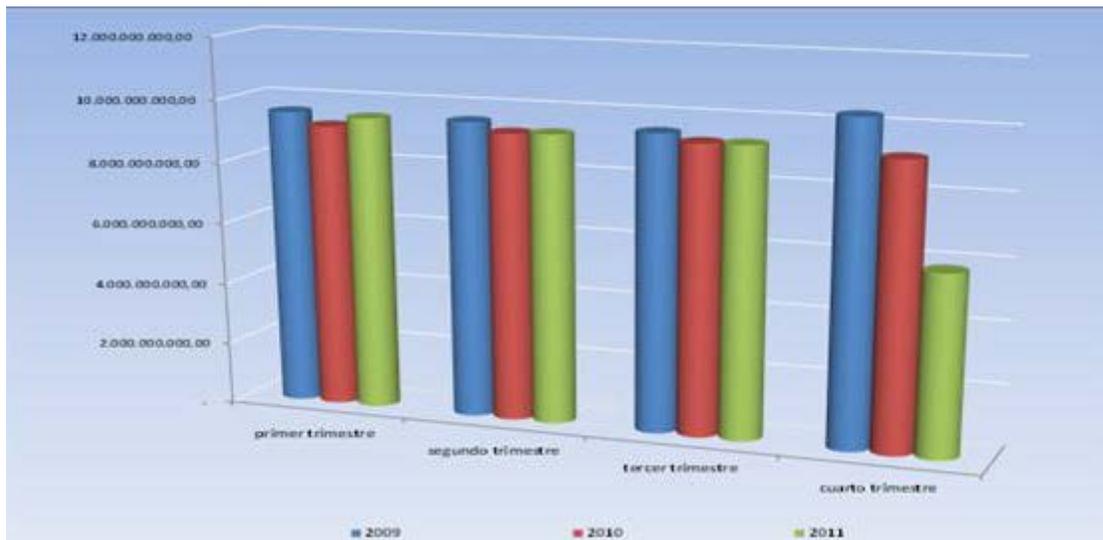
Cualquiera de los dos casos conlleva a un gasto de dinero adicional que debe tratar de evitarse al máximo en pro de optimizar los recursos de la organización.

Figura 20. Comparativo del comportamiento semanal serie saldo



Al realizar una comparación del comportamiento del saldo de efectivo de manera trimestral, es posible determinar que guarda el mismo comportamiento para cada uno de los años en estudio, tal como se aprecia en la figura 21, teniendo presente que se cuenta con información hasta octubre de 2011.

Figura 21. Comparativo del comportamiento trimestral para la serie saldo



3. SERIES DE TIEMPO

En este capítulo se presentan los conceptos y resultados alrededor de series de tiempo univariadas, la metodología Box-Jenkins, modelos de suavizamiento exponencial de Winter y modelos SARMA, necesarios para formular y sustentar los modelos que se ajustan a las series de ingresos, egresos y saldo y que a su vez permitirán realizar pronósticos óptimos de esas mismas series.

3.1 SERIES DE TIEMPO

Antes de empezar a analizar y construir modelos que permitan predecir valores de los ingresos y egresos de forma óptima, se define una serie de tiempo como “... *el resultado de observar los valores de una variable a lo largo del tiempo en intervalos regulares (cada día, cada mes, cada año, etc.)*.” Peña (2005).

Se observan series de tiempo en diferentes campos del conocimiento: en economía, mercadeo, demografía, meteorología, ingeniería, etc. Algunos ejemplos de series de tiempo que pueden citarse son:

- Las ventas mensuales de una empresa en un determinado periodo de tiempo.
- El número de automóviles producidos por año de determinada marca de 2005 a 2011.
- La temperatura promedio diaria promedio en Bucaramanga de los últimos seis meses.

Para describir lo que es una serie de tiempo dentro del contexto de procesos estocásticos, es necesario definir los procesos estocásticos como una familia de

variables aleatorias asociadas a un conjunto índice de números reales, de tal forma que a cada elemento del conjunto le corresponda una y sólo una variable aleatoria, lo que se describe como:

$\{Z(\tau); \tau \in T\}$, en donde T es el conjunto índice y $Z(\tau)$ es la variable aleatoria correspondiente al elemento τ de T .

Los métodos de análisis de series de tiempo consideran el hecho que los datos tomados en diversos periodos de tiempo pueden tener algunas características de autocorrelación, tendencia o estacionalidad que se deben tomar en cuenta.

Un supuesto en muchas técnicas de series de tiempo es que los datos son estacionarios, donde su media, varianza y autocorrelación no cambian en el tiempo, tampoco se presentan patrones de estacionalidad. Sin embargo, en la práctica sí se presentan estos patrones de tendencia y estacionalidad y es necesario contar con modelos que los consideren.

Una serie de tiempo frecuentemente es examinada con la intención de descubrir patrones históricos que puedan ser útiles en la predicción. Para identificar dichos patrones, una serie de tiempo presenta los siguientes componentes en su estructura (Bowerman y O'Connell, 1993)

- Tendencia: una serie de tiempo presenta tendencia cuando por largos periodos de tiempo los valores crecen o decrecen. También puede definirse como cambios en la media de los valores de la serie.
- Ciclos: se refiere a movimientos hacia arriba y hacia abajo alrededor del nivel de la tendencia. Esas fluctuaciones, medidas pico a pico, pueden tener una duración larga.

- Variaciones estacionales: son patrones periódicos que ocurren y se repiten cada determinado tiempo, por ejemplo: anualmente.

3.2 MÉTODOS DE PRONÓSTICOS

Existen métodos que permiten ajustar modelos a series de tiempo con diferentes características como lo son la estacionalidad, la estacionariedad de la serie e incluso su tendencia y ciclos. El presente apartado se concentra en describir aquellos métodos que permiten modelar series con estacionalidad y tendencia como lo son el método de suavizamiento exponencial de Winters y los modelos SARIMA; aunque teniendo en cuenta los resultados de los modelos obtenidos en esta investigación, se describirán los modelos SARMA.

3.2.1 Método de suavizamiento exponencial de Winters. El método de suavizamiento exponencial de Winters se desarrolló a principios de la década de 1960. Este método maneja datos estacionales junto con datos que tienen una tendencia. Se basa en tres ecuaciones, cada una de las cuales suaviza un factor asociado con uno de los tres componentes: aleatoriedad, tendencia y estacionalidad.

Hay tres ecuaciones de suavizamiento básicas implicadas en el método de Winters:

$$S_t = \alpha \frac{X_t}{I_{t-L}} + (1 - \alpha)(S_{t-1} + T_{t-1})$$

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1}$$

$$I_t = \gamma \frac{X_t}{S_t} + (1 - \gamma)I_{t-L}$$

En donde:

S = Valor suavizado de la serie desestacionalizada.

T = Valor suavizado de la tendencia.

I = Valor suavizado del valor estacional.

L = Duración de la estacionalidad.

La ecuación de I es comparable a un índice estacional. Dicho índice se calcula como la razón del valor actual de la serie X_t dividido entre el valor suavizado actual de la serie S_t . Si X_t es mayor que S_t la razón será mayor que 1. Si es menor que S_t la razón será menor que 1.

Es importante tener en cuenta que S_t es un valor (*promedio*) suavizado de la serie que incluye tendencia, pero no estacionalidad. Los valores de los datos X_t contienen estacionalidad. Por lo tanto la razón X_t / S_t dice algo acerca del nivel de estacionalidad de los datos. La estacionalidad en cada periodo no es perfecta, contiene aleatoriedad, por lo cual debe ser suavizada para eliminar tal estocasticidad. Para suavizar dicha estacionalidad, la ecuación de I pondera el factor estacional (X_t / S_t) con γ y el número estacional más reciente que corresponde a la misma estación I_{t-L} con $(1-\gamma)$

La ecuación de T_t suaviza la tendencia pues pondera la tendencia incremental ($S_t - S_{t-1}$) con β y el valor tendencial previo T_{t-1} con $(1-\beta)$. En cuanto al valor suavizado de S_t , el primer término se divide entre el factor estacional I_{t-L} . Esto se hace para desestacionalizar (eliminar fluctuaciones estacionales) X_t . Este ajuste se puede ejemplificar al considerar el caso cuando T_{t-1} es mayor que 1, lo cual ocurre cuando el valor de I en el periodo $T - L$ es mayor que el promedio de su

estacionalidad. Al dividir X_t entre I_{t-L} se tiene un valor que es menor que el valor original por un porciento precisamente igual a la cantidad en que la estacionalidad del periodo $t - L$ era mayor que el periodo. El ajuste opuesto ocurre cuando el factor de estacionalidad es menor que 1. El valor I_{t-L} se utiliza en estos cálculos porque I_t no se puede calcular hasta que se conoce S_t .

La predicción basada en el método de Winters se calcula como:

$$F_{t+m} = (S_t + T_t m) I_{t-L+m}$$

Donde m es el número de periodos adelantados pronosticados.

3.2.2 Metodología Box-Jenkins. La metodología Box y Jenkins, fue creada desde los años 70 por George Box y Gwilym Jenkins. Esta metodología propone un conjunto de procedimientos que permiten escoger entre varios modelos agrupados en tres clases distintas que se ajusten a los datos de una serie de tiempo observada, para posteriormente pronosticar valores futuros de la misma. Estos modelos están basados en funciones lineales de las observaciones. Como objetivo principal busca encontrar el modelo más simple que proporcione la mejor descripción y el mejor pronóstico teniendo en cuenta los datos de la serie.

Chayafield, (1989) afirma: *“La metodología de los modelos ARIMA también es conocida como modelos Box-Jenkins. Este enfoque parte del hecho de que la serie temporal que se trata de predecir es generada por un proceso estocástico cuya naturaleza puede ser caracterizada mediante un modelo. Básicamente, la metodología Box-Jenkins consiste en encontrar un modelo matemático que represente el comportamiento de una serie temporal de datos y permita hacer previsiones únicamente introduciendo el periodo de tiempo correspondiente”* (pag. 42).

La metodología de construcción de modelos consta de las siguientes etapas:

1. Identificación del modelo dentro de una clase general de modelos y de acuerdo con lo que los datos indiquen.
2. Estimación de los parámetros implícitos en el modelo, para lo cual se debe usar la técnica más eficiente que se conozca.
3. Verificación de supuestos, a fin de que los resultados que se deriven del modelo puedan considerarse válidos.
4. Uso del modelo para los fines que motivaron su construcción.

En la etapa de identificación del modelo se requiere definir primero una clase de modelos lo suficientemente general para que pueda representar a prácticamente cualquier serie de tiempo que llegue a ser observada en las aplicaciones.

Teniendo en cuenta la metodología Box – Jenkins, para ajustar un modelo a una serie de tiempo se requiere:

1. Decidir si los datos son estacionarios, es decir, si los datos poseen media y varianza constante. Para ello se debe:
 - Examinar la gráfica de serie de tiempo para determinar si es necesaria una transformación para tener varianza constante.
 - Examinar la función de autocorrelación (FAC) para observar si las autocorrelaciones no decaen, indicando que se deben requerir diferencias para dar una media constante.

Una herramienta clave para la identificación de patrones básicos y para determinar modelos apropiados que corresponden a la serie, es el coeficiente de autocorrelación. Dicho coeficiente de autocorrelación es similar a un coeficiente de correlación con la excepción de que describe la asociación (relación mutua) entre valores de la misma variable, pero en diferentes periodos.

Las autocorrelaciones proporcionan información importante acerca de la estructura de un conjunto de datos y de sus patrones. En un conjunto de datos completamente aleatorios la autocorrelación entre valores sucesivos estará cercana a cero, o será igual a cero, pero para valores de fuerte naturaleza estacional o cíclica estarán sumamente autocorrelacionados. Cabe resaltar que la identificación de un proceso de promedios móviles puro puede llevarse a cabo fácilmente con el empleo de la FAC muestral y la aproximación para la varianza de las autocorrelaciones.

Se tiene la serie de tiempo Y_1, Y_2, \dots, Y_n . La autocorrelación simple muestral en el retraso k , denotada por r_k , es:

$$r_k = \frac{\sum_{t=1}^{n-k} (Y_t \times Y_{t+k})}{\sum_{t=1}^n Y_t^2}$$

Esta cantidad mide la relación lineal entre las observaciones de la serie de tiempo separadas por un retraso de k unidades de tiempo. La autocorrelación r_k toma valores entre -1 y 1. Un valor de r_k cercano a 1 quiere decir que las observaciones separadas por un retraso de k unidades de tiempo, tienen una fuerte tendencia a moverse juntas en forma lineal con pendiente positiva. Por otro lado, un valor de r_k cercano a -1 significa que las observaciones separadas por un retraso de k

unidades de tiempo, tienen una fuerte tendencia a desplazarse juntas en forma lineal con pendiente negativa. (Bowerman y O'Connell, 1993, pág., 406).

La función de autocorrelación simple es el conjunto de autocorrelaciones simples muestrales en los retrasos $k=1,2,\dots$: a la representación gráfica de estas autocorrelaciones se le denomina correlograma simple.

Una medida estadística semejante a la de autocorrelación y que posee varias características que son de gran ayuda para identificar modelos de predicción apropiados es la autocorrelación parcial, entendida como la correlación entre conjuntos de pares ordenados de una serie de tiempo, mide la fuerza de la relación cuando otros términos son tomados en cuenta. La función de autocorrelación parcial proporciona la relación directa que existe entre observaciones separadas por k retardos.

La función de autocorrelación parcial (FACP) permite identificar los procesos autorregresivos (AR) ya que esta función adquiere determinadas características que dependen del orden del proceso y del tipo de parámetros involucrados.

La autocorrelación parcial muestral en el retraso k es:

$$r_{kk} \begin{cases} r_1 & \text{si } k=1 \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} & \text{si } k=2,3,\dots \end{cases}$$

Donde $r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j}$ para $j=1,2,\dots,k-1$

El valor de estas autocorrelaciones se puede pensar como la relación de las observaciones de la serie de tiempo separadas por un retraso de unidades de tiempo, eliminando el efecto de las observaciones intermedias.

Para conocer si el valor de r_{kk} es significativo se utiliza el estadístico $t_{r_{kk}}$:

$$t_{r_{kk}} = \frac{r_{kk}}{S_{r_{kk}}}$$

Donde $S_{r_{kk}}$ es el error estándar de r_{kk} dado por:

$$S_{r_{kk}} = \frac{1}{n^{1/2}}$$

La función de autocorrelación parcial es el conjunto de las autocorrelaciones parciales muestrales en los retrasos $k=1,2,\dots$; a la representación gráfica de estas autocorrelaciones se le denomina correlograma parcial.

2. Posteriormente se debe examinar la FAC y FACFP de los datos estacionarios para identificar el modelo autorregresivo o de promedio móvil según el caso.

- Una función FAC con picos altos iniciales que decaen a cero o una FACFP con picos altos en el primero y posiblemente en el segundo retraso indican un proceso autorregresivo.
- Una función FAC con pico alto inicial y posiblemente en el segundo retraso y una función FACP con picos altos en los primeros retrasos que decaen a cero indican un proceso de promedio móvil.

- Si las funciones ACF y ACFP tiene picos altos que gradualmente caen a cero indican que los procesos de promedios móviles y autorregresivos están presentes.

Para determinar los órdenes del componente autorregresivo (p) y promedio móvil (q) de un modelo ARMA (p, q), se hace uso de los patrones que se observan en los correlogramas simple y parcial:

Correlograma Simple	Correlograma Parcial	Modelo
Decae lentamente	Se corta después del retardo p	AR(p)
Se corta después del retardo q	Decae lentamente	MA(q)
Decae lentamente	Decae lentamente	ARMA(p,q)

3. Una vez que se ha identificado uno o más modelos a utilizar se continúa con el procedimiento.

- Ajustar el modelo y examinar la significancia de los parámetros y seleccionar el modelo que tenga el mejor ajuste.
- Verificar que los residuos indiquen un proceso aleatorio.

Se debe verificar el grado de ajuste del modelo en comparación con otros modelos alternativos. La bondad del ajuste puede evaluarse con la Desviación Estándar Residual (DER). La DER es un criterio de selección de modelo, donde un valor pequeño indica una mayor adecuación del modelo. Su expresión matemática es:

$$\hat{\sigma} = \sqrt{S / (T - r)}$$

Donde S es la suma cuadrática de los residuos; T es el número de las observaciones efectivas que se usa en el ajuste del modelo, recordando que se

pierden observaciones por diferenciación y r es el número de parámetros estimados en el modelo, incluyendo la constante.

Es importante juzgar la adecuación del modelo en función de qué tan bien se pronostican los datos no empleados para la estimación del modelo. Para evaluar la predicción, se utilizan los siguientes tipos de medición de error (Makridakis y Wheelwright, 1994)

- Error medio

$$ME = \frac{\sum_{i=1}^m y_{t+i} - \hat{y}_{t+i}}{m}$$

Donde y_{t+i} son los valores observados de la serie que pertenecen al conjunto de predicción y, \hat{y}_{t+i} son los valores pronosticados por el modelo.

- Error medio absoluto

$$MAE = \frac{\sum_{i=1}^m |y_{t+i} - \hat{y}_{t+i}|}{m}$$

Donde y_{t+i} son los valores observados de la serie que pertenecen al conjunto de predicción y, \hat{y}_{t+i} son los valores pronosticados por el modelo.

- Raíz del error cuadrático

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_{t+i} - \hat{y}_{t+i})^2}$$

- Error porcentual medio

$$MPE = \frac{\sum_{i=1}^m (y_{t+i} - \hat{y}_{t+i}) / y_{t+i}}{m}$$

- Porcentaje promedio absoluto de error

$$MAPE = \frac{\sum |(y_{t+i} - \hat{y}_{t+i}) / y_{t+i}|}{n} \times 100 \quad \text{con } y_{t+i} \neq 0$$

Mide la exactitud de los valores estimados de la serie de tiempo. La exactitud se expresa como un porcentaje con y_t igual al valor observado, \hat{y}_t es el valor estimado y n el número de observaciones.

Otra forma de evaluar la predicción es a través de la correlación entre los valores observados y los pronosticados por el modelo. Valores altos de esta correlación indican una buena adecuación del modelo.

4. Generar pronósticos, para validar el modelo.

3.3 PROCESO DE PROMEDIO MÓVIL MA (Q)

Una alternativa de modelización pasa por tratar de explicar el comportamiento de una variable y , no en función de los valores que tomó en el pasado (modelos AR) sino a través de los errores al estimar el valor de la variable en los períodos anteriores. Ello da lugar a los modelos de medias móviles (o Modelos MA, por sus siglas en inglés).

Un modelo de medias móviles de orden q , $MA(q)$ está definido de la siguiente forma:

$$Y_t = \mu - a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

Donde μ : valor de la constante alrededor del cual se mueve la variable, es decir, la media de la variable de la variable $\mu = E(Y_t)$

a_t, \dots, a_{t-q} : Errores observados en los periodos $t, \dots, t-q$

$\theta_1, \theta_2, \dots, \theta_q$: Coeficientes a estimar.

Una vez estimados los parámetros θ_t , el modelo expresa el patrón de comportamiento seguido por los términos de error en períodos anteriores.

Introduciendo la notación del operador de retardo B , el proceso de promedio móvil $MA(q)$ quedaría descrito por:

$$\tilde{Y}_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t$$

Un proceso $MA(q)$ siempre es estacionario, por ser la suma de procesos estacionarios. Se dice que el proceso es invertible si las raíces del operador $\theta_q(B) = 0$ son en módulo, mayores que la unidad.

En el caso de procesos de medias móviles, las condiciones similares a las de estacionariedad son las de invertibilidad. Cuando un proceso MA es invertible, entonces dicho proceso admite una representación autorregresiva, donde los valores pasados de la variable y_t reciben una ponderación cada vez menor.

3.4 PROCESOS AUTORREGRESIVOS AR (p)

Un modelo autorregresivo de orden p tiene la siguiente forma:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t$$

En donde:

$y_{t-1}, y_{t-2}, y_{t-3}, \dots, y_{t-p}$: Son las variables independientes. En este caso, estas variables independientes son valores de la misma variable (de aquí el nombre de auto) pero de períodos anteriores $(t-1), (t-2), (t-3), \dots, (t-p)$.

$\phi_1, \phi_2, \dots, \phi_p$: Coeficientes a estimar.

a_t : es el error o término residual que representa las perturbaciones aleatorias que no pueden ser explicadas por el modelo.

Un proceso autorregresivo lleva ese nombre porque se asemeja a la ecuación de regresión $Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_p X_p + e$, la única diferencia es que $X_1 = y_{t-1}$, $X_2 = y_{t-2}$, $X_3 = y_{t-3}$, ..., $X_p = y_{t-p}$ y así las variables independientes son simplemente valores rezagados de la variable dependiente con rezagos de tiempo $1, 2, 3, \dots, p$ periodos.

Introduciendo la notación de operadores:

$$a_t = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \tilde{Y}_t, \text{ donde } y_{t-k} = (B)^k y_t$$

3.5 PROCESOS MIXTOS ARMA (p, q)

La combinación de modelos autorregresivos $AR(p)$ y de medias móviles $MA(q)$ da lugar a modelos $ARMA(p,q)$, y que en forma general se define como:

$$Y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

O en forma de los operadores de manera abreviada:

$$\phi_p(B)Y_t = \theta_q(B)\varepsilon_t$$

Los modelos solo autorregresivos corresponden a modelos $ARMA(p, 0)$ y los de medias móviles a modelos $ARMA(0, q)$. En este sentido, es equivalente, por ejemplo, hablar de un modelo $AR(1)$ o de un $ARMA(1,0)$, de un $MA(2)$ o de un $ARMA(0,2)$.

Bajo ciertas condiciones los modelos AR y los MA pueden relacionarse. Estas condiciones se denominan de *invertibilidad* y de *estacionariedad*. Por ejemplo, un modelo $AR(1)$ sin término independiente como:

$$Y_t = \phi_1 y_{t-1} + a_t$$

Puesto que

$$Y_{t-1} = \phi_1 y_{t-2} + a_{t-1}$$

Se llega a través de sustituciones a la siguiente expresión:

$$Y_t = a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \phi_1^3 a_{t-3} + \dots$$

Es decir, un proceso autorregresivo es equivalente a una media móvil de infinitos términos con una ponderación decreciente en forma exponencial cuando $0 < |\phi_1| < 1$.

Este resultado es generalizable y se puede demostrar que bajo las condiciones de estacionariedad un modelo *AR* de orden reducido puede transformarse en un modelo *MA* de orden elevado o teóricamente infinito. De igual forma, bajo las condiciones de invertibilidad, modelos *MA* de orden reducido pueden aproximarse a modelos *AR* con un número elevado de términos.

La utilidad de esta relación radica en que, por aplicación del principio de parsimonia y sin olvidar que en la práctica puede favorecer la estimación del modelo, es preferible un modelo sencillo con el menor número posible de términos, y por lo tanto de parámetros a estimar, frente a un modelo con un gran número de coeficientes, siempre que conduzca a resultados similares.

3.6 PROCESOS ARIMA (p,d,q)

Los modelos de series de tiempo analizados anteriormente están basados en el supuesto de que las series de tiempo consideradas son estacionarias (media y varianza constantes). Pero se sabe que muchas series de tiempo son no estacionarias, es decir, son integradas.

Si se sabe diferenciar una serie de tiempo d veces para hacerla estacionaria y luego aplicar a esta el modelo ARMA (p,q), se dice que la serie de tiempo original sigue un proceso *autorregresivo integrado de promedio móvil* o *ARIMA(p,q,d)*, donde p denota el número de términos autorregresivos, d el número de veces que la serie debe ser diferenciada para hacerse estacionaria y q el número de términos de promedio móvil.

El modelo ARIMA (p,q,d) puede ser escrito como:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d z_t = c + (1 - \theta_1 B - \dots - \theta_q B^q) a_t$$

Este modelo permite describir una serie de observaciones después de que hayan sido diferenciadas **d** veces, a fin de extraer las posibles fuentes de no estacionariedad. Esta fórmula general se puede aplicar a cualquier modelo. Si hay alguna componente **p**, **q**, **d** igual a cero, se elimina el término correspondiente de la fórmula general.

Los procesos ARIMA son suficientes para explicar procesos con tendencia pero incapaces de representar procesos con estacionalidad y se hace necesaria una generalización de estos para lograr explicar los comportamientos estacionales. Los modelos estacionales consideran los retrasos del proceso y de la perturbación aleatoria periódicamente, es decir, cada **s** periodos.

3.7 PROCESOS SARMA (P, Q) _s(p, q)

Los procesos ARIMA son suficientes para explicar procesos con tendencia pero incapaces de representar procesos con estacionalidad y se hace necesaria una generalización de estos para lograr explicar los comportamientos estacionales. Los modelos estacionales consideran los retrasos del proceso y la perturbación aleatoria periódicamente, es decir, cada **s** periodos. Por ejemplo, cuando los datos son mensuales, es lógico considerar el periodo **s=12**. El objeto de estos retardos estacionales (**s**) es explicar la dependencia que tienen entre sí iguales periodos de años sucesivos, por ejemplo, marzo de 2010 con marzo de 2011.

Los modelos estacionales se denotan anteponiéndoles la letra **S** y el orden de sus parámetros se escribe con mayúscula, así: modelos autorregresivos estacionales SAR (**P**), promedios móviles estacionales SMA (**Q**) y autorregresivo y de

promedios móviles estacionales SARMA (P, Q). Los modelos SARMA son análogos al proceso ARMA pero considerando los retardos del ruido blanco* y del proceso de s en s.

La expresión que representa este tipo de modelos está dada por:

$$\phi(B)\Phi(B^s)(X_t - \mu) = \theta(B)\Theta(B^s)Z_t$$

Sin embargo estos modelos SARMA no son capaces de explicar todos los movimientos estacionales, pues si estos crecieran de año en año, los SARMA serían incapaces de recoger esta evolución, pues al igual que los ARMA son estacionarios. Esta dificultad se resuelve a través de los modelos autorregresivos de promedios móviles integrados estacionales SARIMA (P,D,Q)

3.8 PROCESOS SARIMA (P, D, Q) _s (p, d, q)

La unión de modelos estacionales con modelos no estacionales conduce a un modelo de gran capacidad de adaptación que puede reflejar tanto la tendencia, como la estacionalidad de una serie (*enfoque de Box-Jenkins*). La combinación de estos modelos se logra a través de la multiplicación de los operadores polinomiales que caracterizan a cada modelo obteniendo los modelos conocidos como SARIMA. La notación que se emplea para esta combinación de modelos es, en términos generales la siguiente, donde los parámetros no considerados en el modelo seleccionado serán de orden cero:

- p : orden de un modelo autorregresivo AR
- d : orden de diferenciación en la parte regular o no estacionaria de la serie.
- q : orden de un modelo promedio móvil MA

- P : orden de un modelo autorregresivo estacional SAR
- D : orden de diferenciación en la parte estacional de la serie
- Q : orden de un modelo promedio móvil estacional SMA

El modelo SARIMA $(P, D, Q)_s(p, d, q)$ puede escribirse como:

$$\Phi_p(B^s)\phi_p(B)\nabla_s^D\nabla^d z_t = \theta_q(B)\Theta_Q(B^s)a_t$$

Donde:

$\Phi_p(B^s) = (1 - \Phi_1 B^s - \dots - \Phi_p B^{sP})$ es el operador AR estacional de orden P .

$\phi_p = (1 - \phi_1 B - \dots - \phi_p B^p)$ es el operador AR regular de orden p .

$\nabla_s^D = (1 - B^s)^D$ representa las diferencias estacionales.

$\nabla^d = (1 - B)^d$ representa las diferencias regulares.

$\Theta_Q(B^s) = (1 - \Theta_1 B^s - \dots - \Theta_Q B^{sQ})$ es el operador media móvil estacional de orden Q .

$\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ es el operador de media móvil regular de orden q .

a_t es un proceso de ruido blanco¹.

¹ Está definido por las siguientes condiciones: 1. la esperanza siempre es constante e igual a cero. 2. la varianza es constante. 3. las variables del proceso están incorreladas para todos los retardos.

Esta clase de modelos introducidos por Box y Jenkins, representan bien muchas series estacionales que se encuentran en la práctica y se escriben de forma simplificada como el modelo $ARMA(P, D, Q)_s(p, d, q)$.

en estos procesos conocer los valores pasados no proporciona ninguna información sobre el futuro, ya que el proceso no tiene memoria.

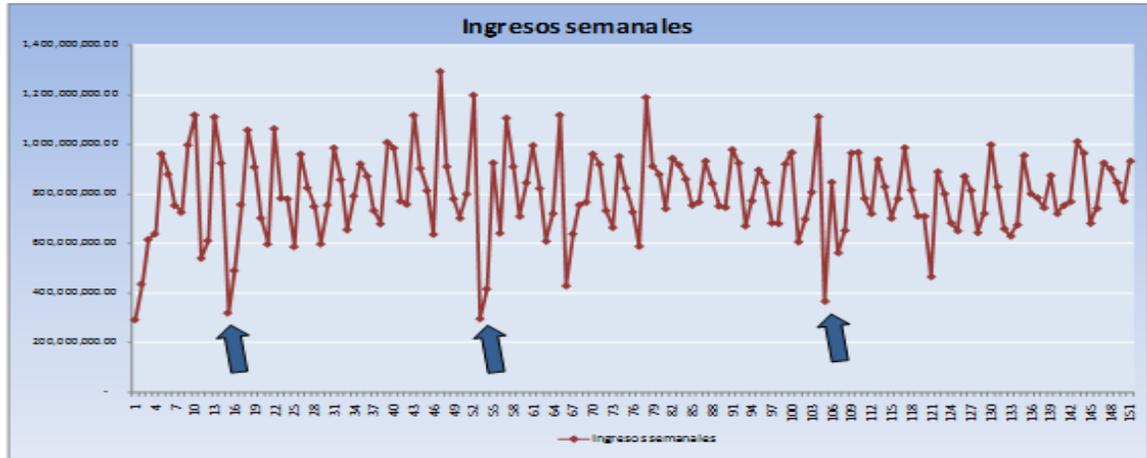
4. ANÁLISIS DE MODELOS

Esta etapa tiene como objetivo principal determinar los modelos de series temporales que mejor se ajustan a las series de ingresos, egresos y saldo. Es decir, se busca definir los modelos de suavizamiento, los modelos de regresión, modelos ARIMA, SARMA o SARIMA que mejor se ajusten a cada una de las series. Se presenta además, una descripción detallada de los modelos que son significativos tanto en sus parámetros y análisis de residuos, como en los pronósticos que se obtienen alrededor de cada una de las series. Paso a pasó se muestra la forma en que fueron desarrollados cada uno de estos modelos. Se presentan el análisis realizado a las series de Ingreso, Egreso y saldo con el modelo de suavizamiento exponencial de Winter, el mejor modelo de regresión y los mejores SARIMA, para después bajo criterios que son descritos en el tercer capítulo “*Series de tiempo*”, escoger el que mejor se ajusta y pronostica en cada una de las series.

4.1 IDENTIFICACIÓN DE MODELOS PARA LA SERIE INGRESOS

En la Figura 22 se observa el comportamiento de la serie Ingreso Semanal a través del tiempo, evidenciándose un comportamiento monótono, con algunos mínimos locales que se deben a ciertos fenómenos importantes durante el año, los cuales se expusieron en el capítulo 2 y que podrían dar lugar a un componente estacional observado cada 52 semanas.

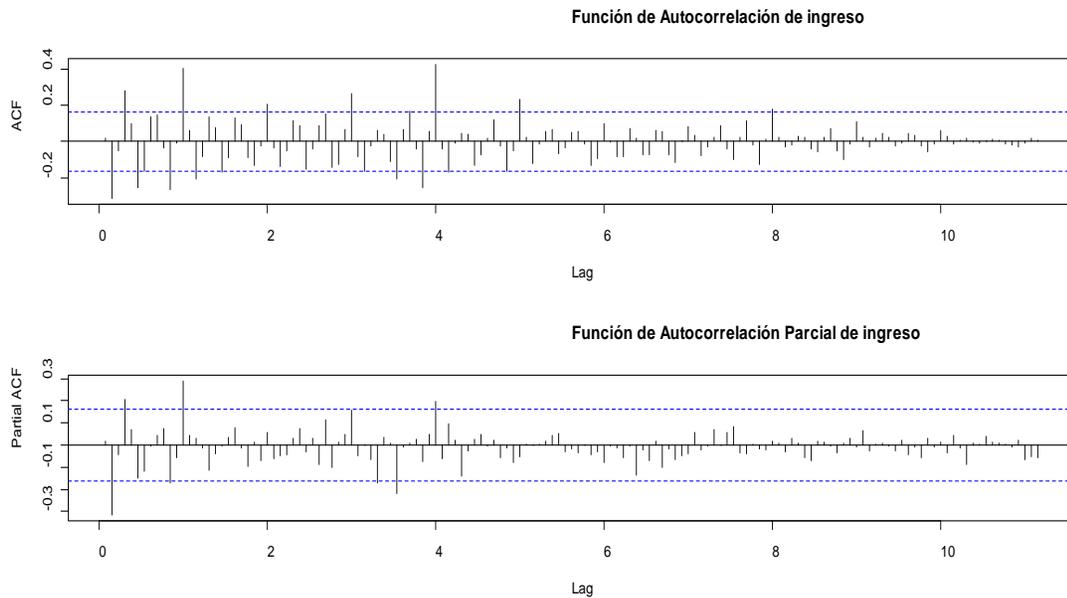
Figura 22. Serie ingresos semanal (2009-2011)



Es posible observar una serie estacionaria, con media y variabilidad constantes a lo largo del tiempo. La serie, a simple vista, muestra estabilidad en la varianza a lo largo del tiempo, es decir, no muestra valores ni muy altos, ni muy bajos; todos se encuentran casi simétricamente distribuidos alrededor de la media. Puede decirse entonces, que como sus propiedades no varían a lo largo del tiempo, no existen tendencias.

4.1.1 Análisis de la Funciones de Autocorrelación muestral (FAC) y de Autocorrelación Parcial (FACP) para la serie Ingreso. La función de autocorrelación muestral de la serie presenta estacionalidad de orden trece (13), lo cual se evidencia en los retrasos 13, 26, 39 y 52 del autocorrelograma de la serie (Ver Figura 23).

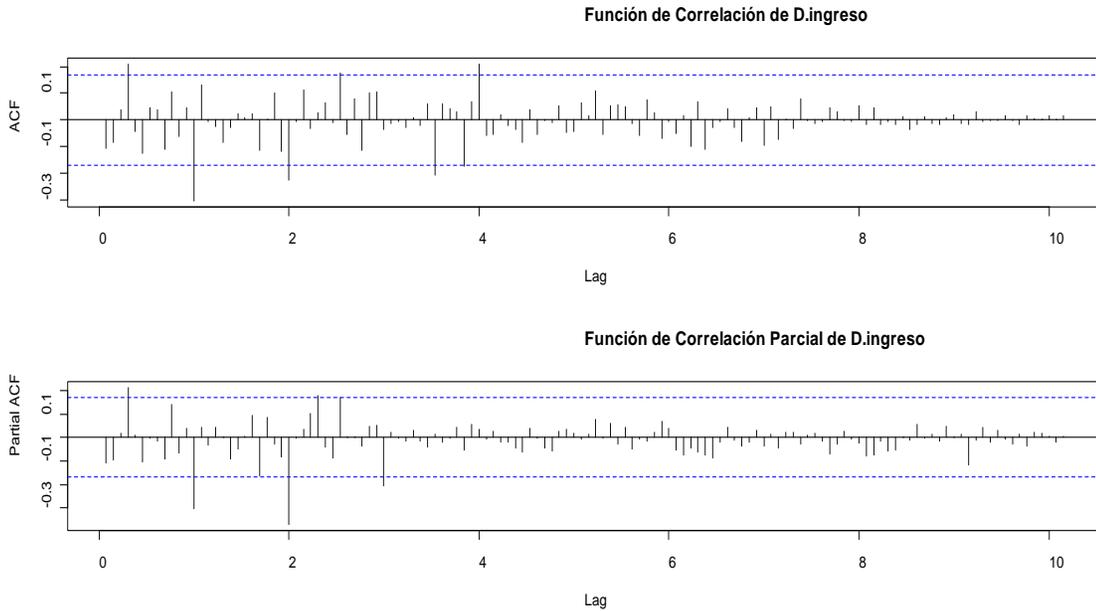
Figura 23. Funciones de Autocorrelación muestral y parcial de la serie Ingreso



Buscando que la serie sea estacionaria en media y en varianza, se procede a realizar una primera diferencia estacional, para de esta forma, eliminar el efecto de la estación en la serie.

Al figura 24 muestra las funciones de autocorrelación muestral y parcial de la serie Ingreso, llamada D.ingreso después de realizada la primera diferencia estacional.

Figura 24. Funciones de Autocorrelación muestral y parcial de la serie D. ingreso



La función de autocorrelación muestral de la serie después de realizada una primera diferencia estacional presenta un rápido decaimiento a cero, lo cual es indicativo de que la serie es estacionaria, en cuanto a nivel se refiere.

Con el propósito de corroborar la premisa de estacionariedad de la serie D.ingreso se aplica el Test de Dickey-Fuller. Es importante tener en cuenta que cuando una serie no es estacionaria en media, se dice que presenta al menos una raíz unitaria. Cuando esto ocurre, se sabe que es posible la obtención de una serie estacionaria mediante una sencilla transformación de la serie original, como es la diferenciación adecuada. Pues bien, el número de diferencias que habrá que tomar en la serie para convertirla en estacionaria en media viene dado, justamente, por el número de raíces unitarias que la serie original presente. Este número de raíces unitarias de la serie se determina formalmente mediante los procedimientos de Dickey-Fuller. El sustento de este test se amplía en el capítulo III "Series de Tiempo".

El test Dickey-Fuller plantea las siguientes hipótesis:

H_0 : La serie de tiempo Ingreso no es estacionaria

H_1 : La serie de tiempo Ingreso es estacionaria

Cuadro 1. Resultado del Test Dickey-Fuller

Augmented Dickey-Fuller Test
data: D.ingreso Dickey-Fuller = -4.392, Lag order = 5, p-value = 0.01 alternative hypothesis: stationary
Mensajes de aviso perdidos In adf.test(D.ingreso) : p-value smaller than printed p-value

Teniendo en cuenta los resultados de la prueba realizada en el paquete estadístico R (*The R Project for Statistical Computing*), el p-valor es menor a 0,05 por lo cual se rechaza la hipótesis de que la serie Ingreso es no estacionaria y se acepta la hipótesis alternativa: la serie Ingreso es estacionaria, tal como se había previsto con base en el análisis gráfico.

4.1.2 Identificación de modelos para la serie Ingreso. Teniendo en cuenta que la serie Ingreso es estacionaria en media y varianza, estacional (la tendencia la marca la media, aspecto que ya está claro cuando se dijo que es estacionaria) se inicia el proceso de identificación de modelos con el suavizamiento de Holt-Winter y suavizamiento de Brown.

Por otra parte, en las funciones de autocorrelación muestral (Ver la Figura 21) se observa un valor significativo en el retraso 1 (semana 1) de su parte regular, mientras que en la parte estacional se evidencian cortes significativos en los retrasos 13, 26 y 52 (semanas 13, 26 y 52). Lo anterior sugiere procesos MA de

orden 1 o 0 en su parte regular y procesos SMA de orden 3, 2, 1 o 0 en su parte estacional.

Analizando la función de autocorrelación parcial (Ver Figura 21) se observa un valor significativo en el retraso 4 (semana 4) de su parte regular, mientras que en la parte estacional se evidencian cortes significativos en los retrasos 13, 26 y 39. Lo anterior indica procesos AR de orden 4, 3, 2, 1 o 0 en su parte regular y procesos SAR de orden 3, 2, 1, o 0 en su parte estacional.

Teniendo en cuenta que la unión de procesos SMA y SAR con una diferencia estacional de orden 13 da lugar a modelos SARIMA se plantean los siguientes modelos:

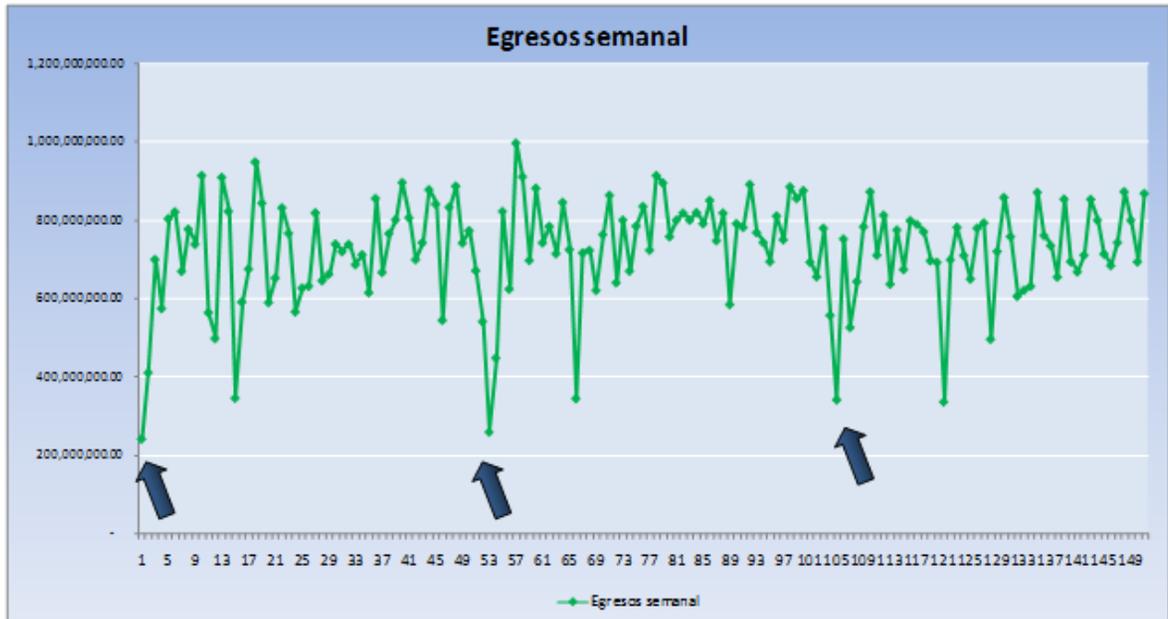
Cuadro 2. Posibles Modelos

POSIBLES MODELOS
SARIMA (2,1,2) ₁₃ (0,0,0)
SARIMA (2,1,2) ₁₃ (4,0,0)
SARIMA (2,1,2) ₁₃ (0,0,4)

4.2 IDENTIFICACIÓN DE MODELOS PARA LA SERIE EGRESOS

En la Figura 25 se observa el comportamiento de la serie Egreso semanal evidenciándose un comportamiento monótono, con algunos mínimos locales que se deben a ciertos fenómenos importantes durante el año, los cuales se expusieron anteriormente y que dan lugar a un posible componente estacional, observado cada 52 semanas.

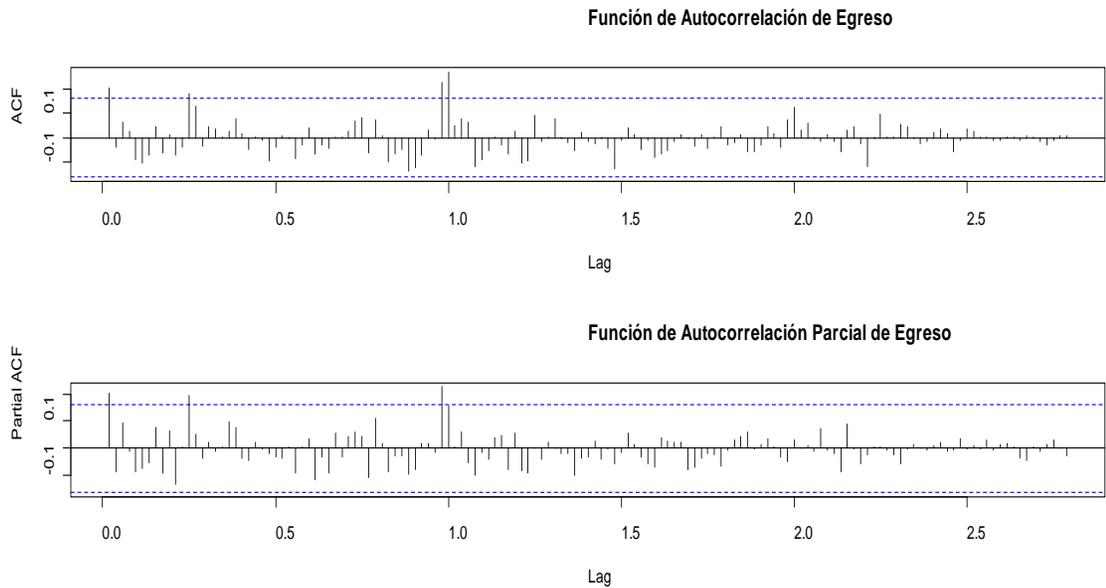
Figura 25. Comportamiento histórico para la serie egresos semanal (2009-2011)



La serie a simple vista muestra estabilidad en la varianza a lo largo del tiempo, es decir, no muestra valores ni muy altos, ni muy bajos; todos se encuentran casi simétricamente distribuidos alrededor de la media. Es posible observar una serie estacionaria, con media y variabilidad constantes a lo largo del tiempo. Puede decirse entonces, que como sus propiedades no varían a lo largo del tiempo, no existen tendencias.

4.2.1 Análisis de las Funciones de Autocorrelación muestral (FAC) y de Autocorrelación Parcial (FACP) para la serie Egreso. La función de autocorrelación muestral de la serie presenta un rápido decaimiento a cero, lo cual es indicativo de que la serie es estacionaria, en cuanto a nivel se refiere. Se observa además estacionalidad de orden cincuenta y dos (52), lo cual se evidencia en los retrasos 52 y 104 del autocorrelograma de la serie (Ver la Figura 26).

Figura 26. Funciones de Autocorrelación muestral y parcial de la serie Egreso



Con el propósito de corroborar la premisa de estacionariedad de la serie Egreso, se aplica el Test de Dickey-Fuller, que plantea las siguientes hipótesis:

H_0 : La serie de tiempo Egreso no es estacionaria

H_1 : La serie de tiempo Egreso es estacionaria

Cuadro 3. Resultado del Test Dickey-Fuller

Augmented Dickey-Fuller Test
Data: Egreso
Dickey-Fuller = -5.2374, Lag order = 5, p-value = 0.01
Alternative hypothesis: stationary
Mensajes de aviso perdidos
In <code>adf.test(egreso)</code> : p-value smaller than printed p-value

Teniendo en cuenta los resultados de la prueba, el p-valor es menor a 0,05 por lo cual se rechaza la hipótesis de que la serie Ingreso es no estacionaria y se acepta la hipótesis alternativa: la serie Egreso es Estacionaria.

4.2.2 Identificación de modelos para la serie Egresos. Teniendo en cuenta que la serie Egreso es estacionaria en media y varianza, estacional y sin tendencia se realiza Suavizamiento de Holt-Winter y suavizamiento de Brown.

Por otra parte, en las funciones de autocorrelación muestral (Ver la Figura 23) se observan valores significativos en el retraso 1 (semana 1) de su parte regular, al igual que en su parte estacional en el retraso 52 (*semana 52*). Lo anterior indica procesos MA de orden 1 en su parte regular y procesos SMA de orden 1 en su parte estacional.

Analizando la función de autocorrelación parcial (Ver figura 23) se observan valores significativos en el retraso 1 (*semana 1*) de su parte regular, mientras que en la parte estacional se evidencian un valor significativo en el retraso 1. Lo anterior indica procesos AR de orden 1 en su parte regular y procesos SAR de orden 1 en su parte estacional.

Teniendo en cuenta que la unión de procesos SMA y SAR da lugar a modelos SARMA se plantean los siguientes modelos:

Cuadro 4. Posibles Modelos

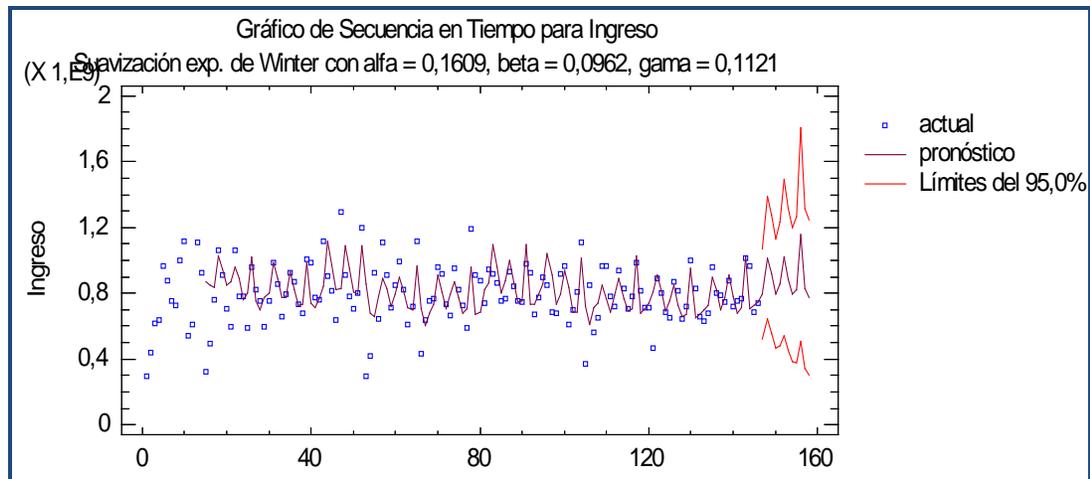
POSIBLE MODELO
SARMA (1,0,1) ₅₂

4.3 ANÁLISIS DE MODELOS PARA LA SERIE INGRESO SEMANAL

La serie de Ingresos semanales presenta estacionariedad en media y varianza, estacionalidad de orden 52 y no presenta tendencia. Teniendo en cuenta lo anterior se analizará el modelo exponencial de Winters y con ayuda de las funciones de autocorrelación descritos en el capítulo de descriptivo, se procede a analizar además, los modelos SARIMA planteados para la serie Ingresos.

4.3.1 Modelo de suavizamiento exponencial de Winters. En la figura 27 se observa el ajuste del modelo de suavizamiento exponencial de Winters realizado a la serie Ingresos.

Figura 27. Ajuste de la serie Ingreso usando Suavizamiento Exponencial de Winters.



El modelo de suavizamiento exponencial de Winters queda descrito por las siguientes ecuaciones:

$$S_t = 0,1609 \frac{X_t}{I_{t-L}} + (0,8391)(S_{t-1} + T_{t-1})$$

$$T_t = 0,0962(S_t - S_{t-1}) + (0,9038)T_{t-1}$$

$$I_t = 0,1121 \frac{X_t}{S_t} + (0,8879)I_{t-L}$$

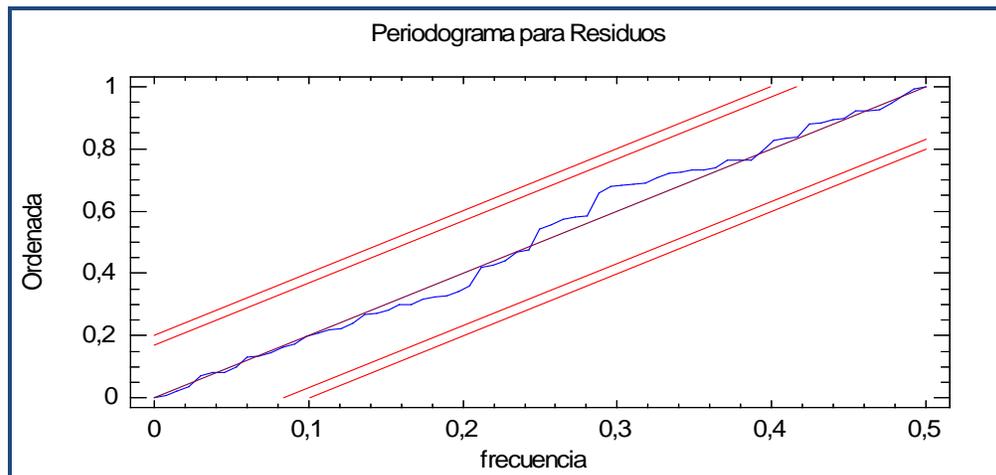
La predicción basada en el método de suavizamiento exponencial de Winters se calcula como:

$$F_{t+m} = \left[\left(0,1609 \frac{X_t}{I_{t-L}} + (0,8391)(S_{t-1} + T_{t-1}) \right) + (0,0962(S_t - S_{t-1}) + (0,9038)T_{t-1}) \right] \times \left[0,1121 \frac{X_t}{S_t} + (0,8879)I_{t-L} \right]$$

ANÁLISIS DE RESIDUOS

El periodograma de residuos de la figura 28 presenta un comportamiento de ruido blanco, lo que indica que el modelo de suavizamiento exponencial de Winters es eficiente al modelar el comportamiento de la serie.

Figura 28. Periodograma de Residuos



Lo anterior se puede comprobar analizando las pruebas o contrastes de aleatoriedad presentados en la figura 26, para determinar si los residuos forman, o no, una secuencia aleatoria de números.

La primera prueba cuenta el número de veces que la secuencia estuvo arriba o abajo de la mediana. El número de tales corridas es igual a 70, comparado con un valor esperado de 67,0 si la secuencia fuera aleatoria. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que los residuos son aleatorios, con un nivel de confianza del 95,0% o mayor.

La segunda prueba cuenta el número de veces que la secuencia ascendió o descendió. El número de tales corridas es igual a 83, comparado con un valor esperado de 87,6667 si la secuencia fuera aleatoria. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que la serie es aleatoria, con un nivel de confianza del 95,0% o mayor.

La tercera prueba está basada en la suma de cuadrados de los primeros 24 coeficientes de autocorrelación. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que la serie es aleatoria, con un nivel de confianza del 95,0% o mayor.

Figura 29. Contrastes de aleatoriedad de Residuos

(1) Corridas arriba o abajo de la mediana
Mediana = -3,96186E6
Número de corridas arriba o abajo de la mediana = 70
Número esperado de corridas = 67,0
Estadístico z para muestras grandes = 0,436865
Valor-P = 0,662206
(2) Corridas arriba y abajo
Número de corridas arriba y abajo = 83
Número esperado de corridas = 87,6667
Estadístico z para muestras grandes = 0,866095
Valor-P = 0,386436
(3) Prueba Box-Pierce
Prueba basada en las primeras 24 autocorrelaciones
Estadístico de prueba para muestras grandes = 15,5743
Valor-P = 0,79307

PRUEBAS DE NORMALIDAD DE RESIDUOS

Para validar el modelo de suavizamiento exponencial de Winters se realizaron las siguientes pruebas de normalidad a los residuos.

Test Shapiro – Wilk

El test de Shapiro-Wilk plantea las siguientes hipótesis:

H_0 = Los residuos del modelo presentan un comportamiento normal

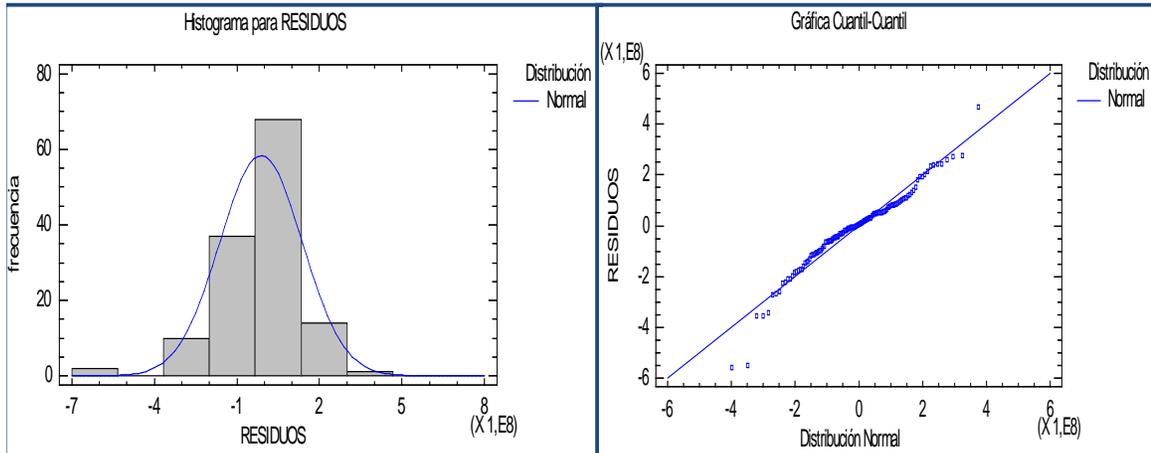
H_a = Los residuos del modelo No presentan un comportamiento normal

El p valor del test de Shapiro-Wilk (*ver cuadro 5*) es mayor que el nivel de significancia de 0,05 por tanto no se rechaza la hipótesis nula sobre la normalidad de los residuos. Esto también se puede observar en los gráficos de probabilidad normal (*ver figura 30*); por lo que se concluye que los residuos provienen de una distribución normal, es decir, que son aleatorios entre sí.

Cuadro 5. Resultado del Test Shapiro-Wilk

Pruebas de Normalidad para RESIDUOS		
Prueba	Estadístico	Valor-P
Estadístico W de Shapiro-Wilk	0,965352	0,025198

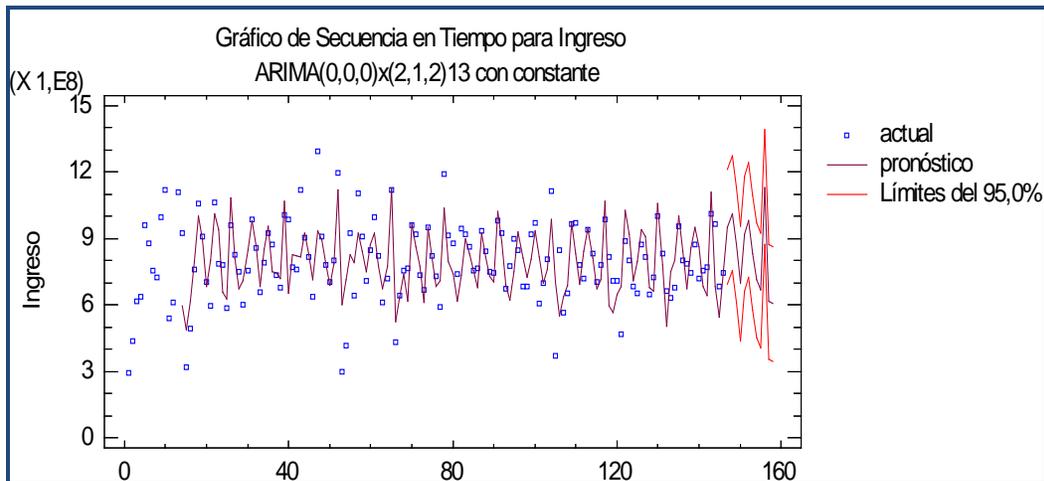
Figura 30. Gráficos de Normalidad de Residuos



4.3.2 Modelos SARIMA. En esta sección se presentará el análisis de cada uno de los modelos SARMA que resultaron significativos tanto en coeficientes como en pronósticos para la serie ingresos. Realizado este análisis se procederá a escoger el mejor de los modelos para compararlo con el modelo de suavizamiento exponencial de Winters.

4.3.2.1 Modelo SARIMA (2,1,2)₁₃(0,0,0). En la figura 31 se observa el ajuste del modelo SARIMA (2,1,2)₁₃(0,0,0) realizado a la serie Ingresos.

Figura 31. Ajuste de la serie Ingreso usando el modelo SARIMA(2,1,2)₁₃(0,0,0)



PARÁMETROS

La tabla 16 resume la significancia estadística de los términos en el modelo de pronósticos. Es importante resaltar que términos con valores-P menores que 0,05 son estadísticamente diferentes de cero con un nivel de confianza del 95,0%. En el modelo $SARIMA(2,1,2)_{13}(0,0,0)$ el valor-P para el término SAR(2) es menor que 0,05, de modo que es estadísticamente diferente de 0. El valor-P para el término SMA(2) es menor que 0,05, de modo que es estadísticamente diferente de 0. La desviación estándar estimada del ruido blanco de entrada es igual a 1,30643E8.

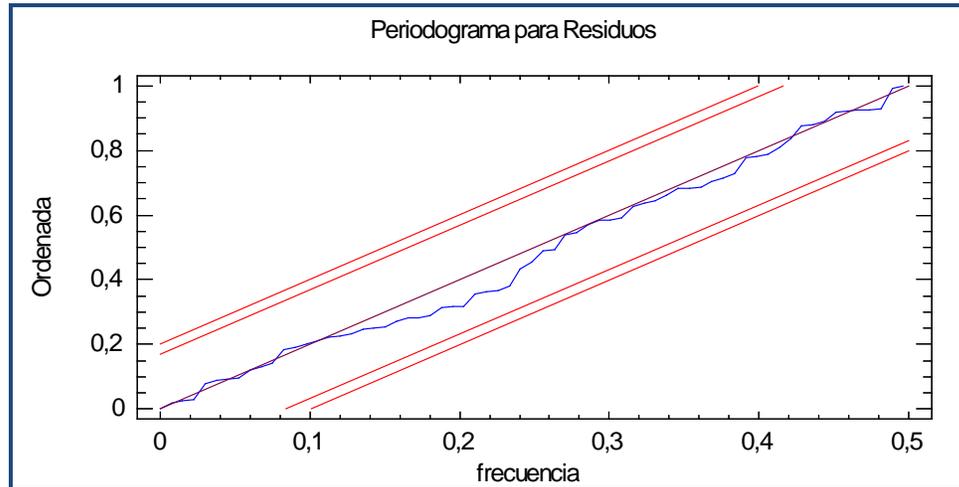
Tabla 16. Coeficientes del modelo $SARMA(2,1,2)_{52}(0,0,0)$ para la serie ingreso

Resumen de Modelo ARIMA				
Parámetro	Estimado	Error Estd.	t	Valor-P
SAR(1)	0,605706	0,0706924	8,56819	0
SAR(2)	-0,498172	0,069719	-7,14543	0
SMA(1)	1,61982	0,0517109	31,3246	0
SMA(2)	-0,683204	0,0488446	-13,9873	0

ANÁLISIS DE RESIDUOS

El periodograma de residuos de la figura 32 presenta un comportamiento de ruido blanco, lo que indica que el modelo $SARIMA(2,1,2)_{13}(0,0,0)$ es eficiente al modelar el comportamiento de la serie.

Figura 32. Peridograma de Residuos



Lo anterior se puede comprobar analizando las pruebas o contrastes de aleatoriedad presentados en la figura 33, para determinar si los residuos forman, o no, una secuencia aleatoria de números.

La primera prueba cuenta el número de veces que la secuencia estuvo arriba o abajo de la mediana. El número de tales corridas es igual a 62, comparado con un valor esperado de 67,0 si la secuencia fuera aleatoria. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que los residuos son aleatorios, con un nivel de confianza del 95,0% o mayor.

La segunda prueba cuenta el número de veces que la secuencia ascendió o descendió. El número de tales corridas es igual a 87, comparado con un valor esperado de 88,3333 si la secuencia fuera aleatoria. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que la serie es aleatoria, con un nivel de confianza del 95,0% o mayor.

La tercera prueba está basada en la suma de cuadrados de los primeros 24 coeficientes de autocorrelación. Puesto que el valor-P para esta prueba es mayor

o igual que 0,05, no se puede rechazar la hipótesis de que la serie es aleatoria, con un nivel de confianza del 95,0% o mayor.

Figura 33. Contrastes de aleatoriedad de Residuos

(1) Corridas arriba o abajo de la mediana
Mediana = 3,01223E6
Número de corridas arriba o abajo de la mediana = 62
Número esperado de corridas = 67,0
Estadístico z para muestras grandes = 0,786357
Valor-P = 0,431657
(2) Corridas arriba y abajo
Número de corridas arriba y abajo = 87
Número esperado de corridas = 88,3333
Estadístico z para muestras grandes = 0,172557
Valor-P = 0,862994
(3) Prueba Box-Pierce
Prueba basada en las primeras 24 autocorrelaciones
Estadístico de prueba para muestras grandes = 16,3825
Valor-P = 0,692627

PRUEBAS DE NORMALIDAD DE RESIDUOS

Para validar el modelo $SARIMA(2,1,2)_{13}(0,0,0)$ se realizaron las siguientes pruebas de normalidad a los residuos.

Test Shapiro – Wilk

El test de Shapiro-Wilk plantea las siguientes hipótesis:

H_0 = Los residuos del modelo presentan un comportamiento normal

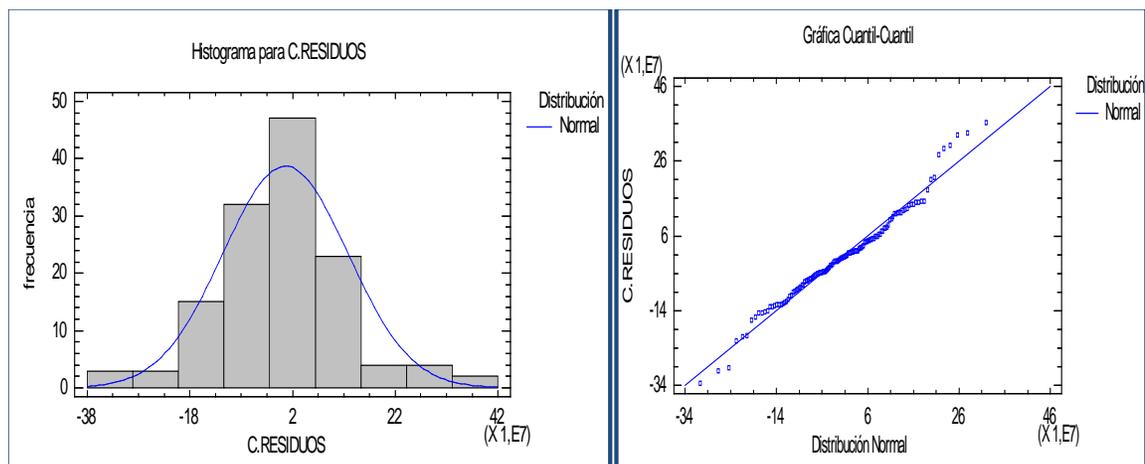
H_a = Los residuos del modelo No presentan un comportamiento normal

El p valor del test de Shapiro-Wilk (ver cuadro 6) es mayor que el nivel de significancia de 0,05 por tanto no se rechaza la hipótesis nula sobre la normalidad de los residuos. Esto también se puede observar en los gráficos de probabilidad normal (ver figura 33); por lo que se concluye que los residuos provienen de una distribución normal, es decir, que son aleatorios entre sí.

Cuadro 6. Resultado del Test Shapiro-Wilk

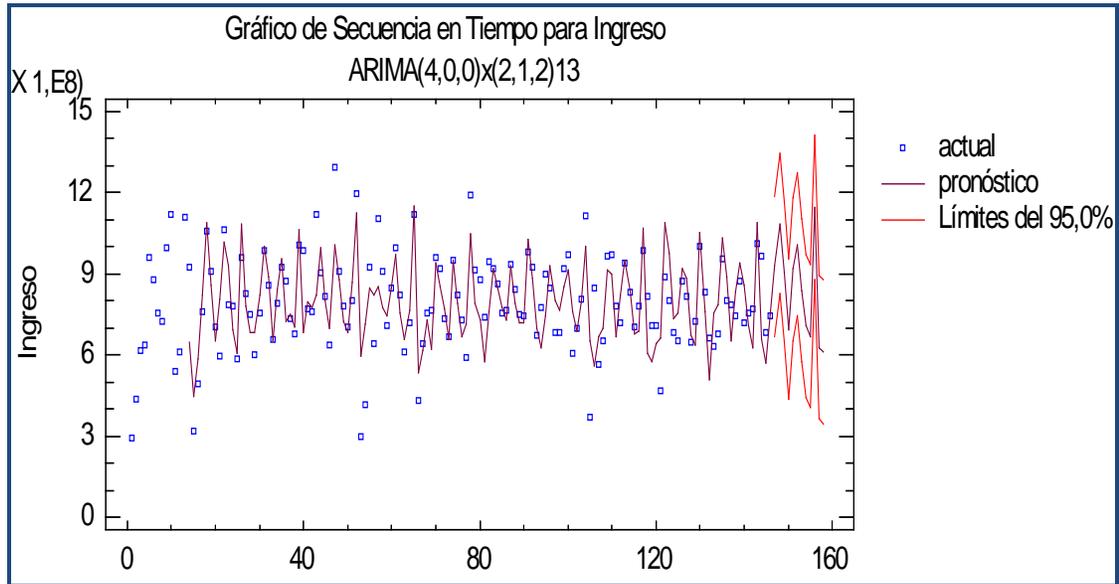
Pruebas de Normalidad para C.RESIDUOS		
Prueba	Estadístico	Valor-P
Estadístico W de Shapiro-Wilk	0,973553	0,155083

Figura 34. Gráficos de Normalidad de Residuos



4.3.2.2 Modelo SARIMA (2,1,2)₁₃(4,0,0). En la figura 35 se observa el ajuste del modelo $SARIMA(2,1,2)_{13}(4,0,0)$ realizado a la serie Ingresos.

Figura 35. Ajuste de la serie Ingreso usando el modelo SARIMA(2,1,2)₁₃(4,0,0)



PARÁMETROS

La tabla 17 resume la significancia estadística de los términos en el modelo de pronósticos. Es importante resaltar que términos con valores-P menores que 0,05 son estadísticamente diferentes de cero con un nivel de confianza del 95,0%. El valor-P para el término AR(4) es menor que 0,05, de modo que es estadísticamente diferente de 0. El valor-P para el término SAR(2) es menor que 0,05, de modo que es estadísticamente diferente de 0. El valor-P para el término SMA(2) es menor que 0,05, de modo que es estadísticamente diferente de 0. La desviación estándar estimada del ruido blanco de entrada es igual a 1,3025E8.

Tabla 17. Coeficientes del modelo SARMA(1,0,3)₅₂ para la serie ingreso

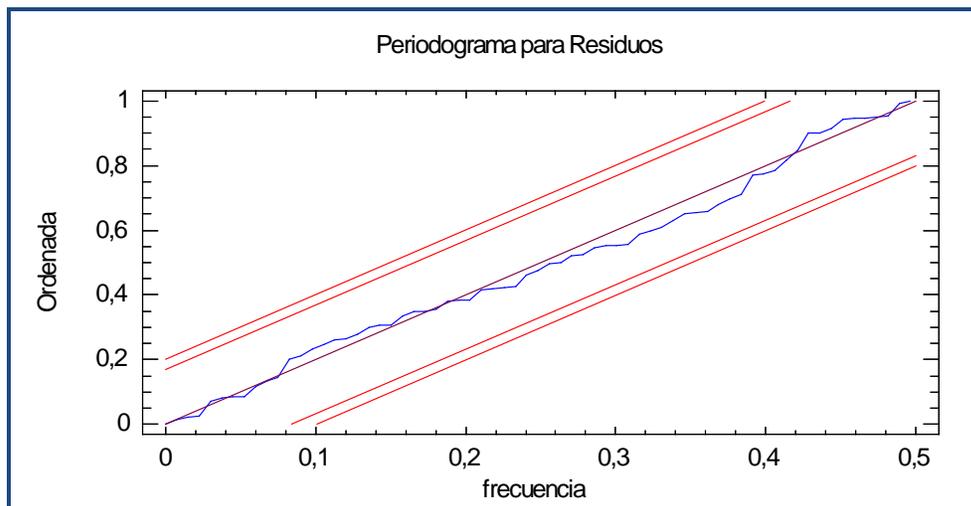
Resumen de Modelo ARIMA				
Parámetro	Estimado	Error Estd.	t	Valor-P
AR(1)	-0,0342456	0,0849821	-0,402974	0,687656
AR(2)	-0,0651486	0,0850186	-0,766287	0,44495
AR(3)	0,0198105	0,0854217	0,231914	0,816984
AR(4)	0,22339	0,0849229	2,6305	0,009598

Resumen de Modelo ARIMA				
Parámetro	Estimado	Error Estd.	t	Valor-P
SAR(1)	0,622476	0,0784221	7,93751	0
SAR(2)	-0,490083	0,0713225	-6,87136	0
SMA(1)	1,62881	0,0527865	30,8566	0
SMA(2)	-0,693048	0,0493915	-14,0317	0

ANÁLISIS DE RESIDUOS

El periodograma de residuos integrados de la figura 36 presenta un comportamiento de ruido blanco, lo que indica que el modelo *SARIMA* $(2,1,2)_{13}(4,0,0)$ es eficiente al modelar el comportamiento de la serie.

Figura 36. Peridograma de Residuos



Lo anterior se puede comprobar analizando las pruebas o contrastes de aleatoriedad presentados en la figura 37, para determinar si los residuos forman, o no, una secuencia aleatoria de números.

La primeraprueba cuenta el número de veces que la secuencia estuvo arriba o abajo de la mediana. El número de tales corridas es igual a 66, comparado con un valor esperado de 67,0 si la secuencia fuera aleatoria. Puesto que el valor-P para

esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que los residuos son aleatorios, con un nivel de confianza del 95,0% o mayor.

La segunda prueba cuenta el número de veces que la secuencia ascendió o descendió. El número de tales corridas es igual a 89, comparado con un valor esperado de 88,3333 si la secuencia fuera aleatoria. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que la serie es aleatoria, con un nivel de confianza del 95,0% o mayor.

La tercera prueba está basada en la suma de cuadrados de los primeros 24 coeficientes de autocorrelación. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que la serie es aleatoria, con un nivel de confianza del 95,0% o mayor.

Figura 37 Contrastes de aleatoriedad de Residuos

(1) Corridas arriba o abajo de la mediana
Mediana = 1,06294E6
Número de corridas arriba o abajo de la mediana = 66
Número esperado de corridas = 67,0
Estadístico z para muestras grandes = 0,087373
Valor-P = 0,93037
(2) Corridas arriba y abajo
Número de corridas arriba y abajo = 89
Número esperado de corridas = 88,3333
Estadístico z para muestras grandes = 0,0345115
Valor-P = 0,972464
(3) Prueba Box-Pierce
Prueba basada en las primeras 24 autocorrelaciones
Estadístico de prueba para muestras grandes = 13,169
Valor-P = 0,660363

PRUEBAS DE NORMALIDAD DE RESIDUOS

Para validar el modelo $SARIMA(2,1,2)_{13}(4,0,0)$ se realizaron las siguientes pruebas de normalidad a los residuos.

Test Shapiro – Wilk

El test de Shapiro-Wilk plantea las siguientes hipótesis:

H_0 = Los residuos del modelo presentan un comportamiento normal

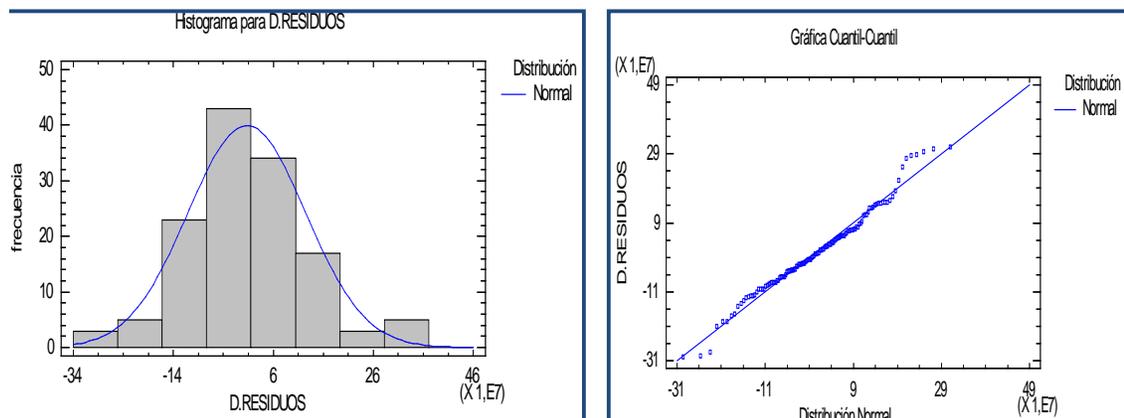
H_a = Los residuos del modelo No presentan un comportamiento normal

El p valor del test de Shapiro-Wilk (*ver cuadro 7*) es mayor que el nivel de significancia de 0,05 por tanto no se rechaza la hipótesis nula sobre la normalidad de los residuos. Esto también se puede observar en los gráficos de probabilidad normal (*ver figura 38*); por lo que se concluye que los residuos provienen de una distribución normal, es decir, que son aleatorios entre sí.

Cuadro 7. Resultado del Test Shapiro-Wilk

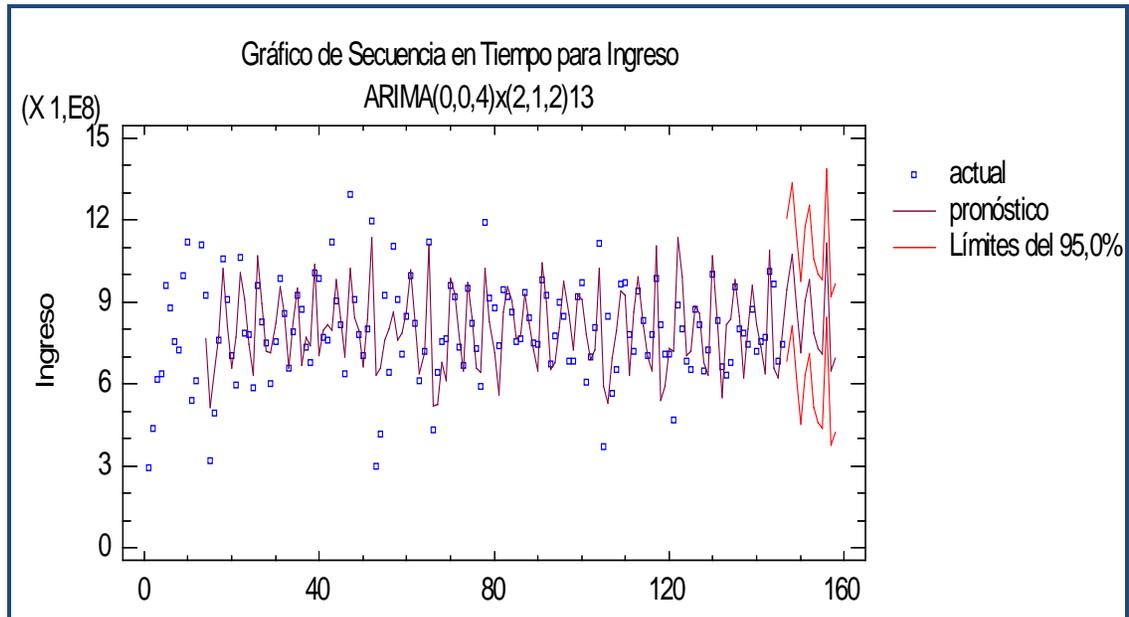
Pruebas de Normalidad para D.RESIDUOS		
Prueba	Estadístico	Valor-P
Estadístico W de Shapiro-Wilk	0,970412	0,0811406

Figura 38. Gráficos de Normalidad de Residuos



4.3.2.3 Modelo SARIMA (2,1,2)₁₃(0,0,4). En la figura 39 se observa el ajuste del modelo SARIMA (2,1,2)₁₃(0,0,4) realizado a la serie Ingresos.

Figura 39. Ajuste de la serie Ingreso usando el modelo SARIMA(2,1,2)₁₃(4,0,0)



PARÁMETROS

La tabla 18 resume la significancia estadística de los términos en el modelo de pronósticos. Es importante resaltar que términos con valores-P menores que 0,05 son estadísticamente diferentes de cero con un nivel de confianza del 95,0%.

El valor-P para el término MA(4) es menor que 0,05, de modo que es estadísticamente diferente de 0. El valor-P para el término SAR(2) es menor que 0,05, de modo que es estadísticamente diferente de 0. El valor-P para el término SMA(2) es menor que 0,05, de modo que es estadísticamente diferente de 0. La desviación estándar estimada del ruido blanco de entrada es igual a 1,31813E8.

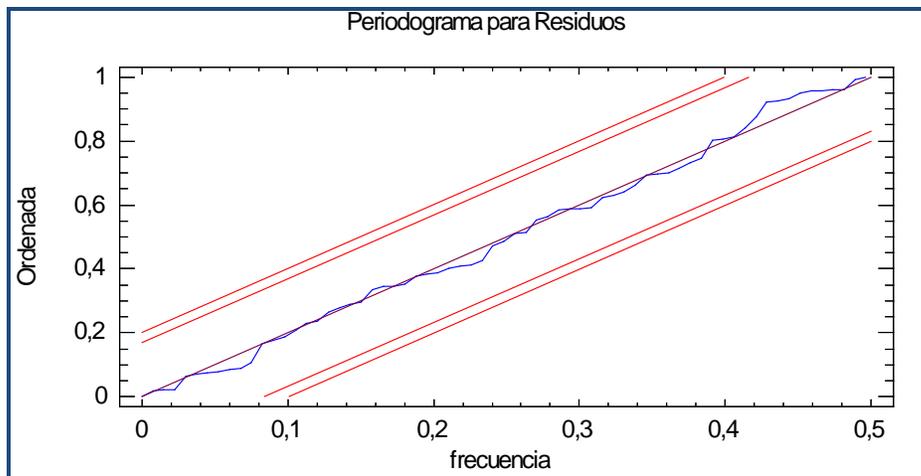
Tabla 18. Coeficientes del modelo SARMA(1,0,2)₅₂ para la serie ingresos

Resumen de Modelo ARIMA				
Parámetro	Estimado	Error Estd.	t	Valor-P
MA(1)	0,0560869	0,0866069	0,647603	0,518429
MA(2)	-0,0583541	0,0852838	-0,684235	0,495094
MA(3)	-0,0199975	0,0866802	-0,230704	0,817922
MA(4)	-0,28396	0,0852723	-3,33004	0,001142
SAR(1)	0,345595	0,125053	2,76358	0,006582
SAR(2)	-0,549604	0,0738128	-7,44592	0
SMA(1)	1,24524	0,141157	8,82169	0
SMA(2)	-0,367316	0,125447	-2,92805	0,004054

ANÁLISIS DE RESIDUOS

El periodograma de residuos integrados de la figura 40 presenta un comportamiento de ruido blanco, lo que indica que el modelo *SARIMA* $(2,1,2)_{13}(0,0,4)$ es eficiente al modelar el comportamiento de la serie.

Figura 40. Peridograma de Residuos



Lo anterior se puede comprobar analizando las pruebas o contrastes de aleatoriedad presentados en la figura 38, para determinar si los residuos forman, o no, una secuencia aleatoria de números.

La primera prueba cuenta el número de veces que la secuencia estuvo arriba o abajo de la mediana. El número de tales corridas es igual a 70, comparado con un valor esperado de 67,0 si la secuencia fuera aleatoria. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que los residuos son aleatorios, con un nivel de confianza del 95,0% o mayor.

La segunda prueba cuenta el número de veces que la secuencia ascendió o descendió. El número de tales corridas es igual a 85, comparado con un valor esperado de 88,3333 si la secuencia fuera aleatoria. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que la serie es aleatoria, con un nivel de confianza del 95,0% o mayor.

La tercera prueba está basada en la suma de cuadrados de los primeros 24 coeficientes de autocorrelación. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que la serie es aleatoria, con un nivel de confianza del 95,0% o mayor.

Figura 41. Contrastes de aleatoriedad de Residuos

(1) Corridas arriba o abajo de la mediana Mediana = $-9,89869E6$ Número de corridas arriba o abajo de la mediana = 70 Número esperado de corridas = 67,0 Estadístico z para muestras grandes = 0,436865 Valor-P = 0,662206
(2) Corridas arriba y abajo Número de corridas arriba y abajo = 85 Número esperado de corridas = 88,3333 Estadístico z para muestras grandes = 0,586695 Valor-P = 0,557405
(3) Prueba Box-Pierce Prueba basada en las primeras 24 autocorrelaciones Estadístico de prueba para muestras grandes = 15,553 Valor-P = 0,48457

PRUEBAS DE NORMALIDAD DE RESIDUOS

Para validar el modelo SARIMA $(2,1,2)_{13}(0,0,4)$ se realizaron las siguientes pruebas de normalidad a los residuos.

Test Shapiro – Wilk

El test de Shapiro-Wilk plantea las siguientes hipótesis:

H_0 = Los residuos del modelo presentan un comportamiento normal

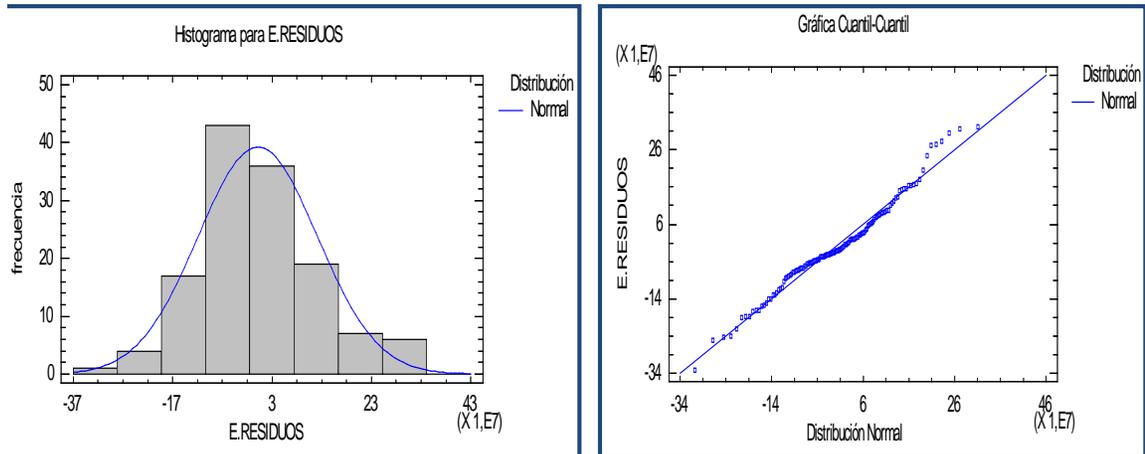
H_a = Los residuos del modelo No presentan un comportamiento normal

El p valor del test de Shapiro-Wilk (*ver cuadro 8*) es mayor que el nivel de significancia de 0,05 por tanto no se rechaza la hipótesis nula sobre la normalidad de los residuos. Esto también se puede observar en los gráficos de probabilidad normal (*ver figura 42*); por lo que se concluye que los residuos provienen de una distribución normal, es decir, que son aleatorios entre sí.

Cuadro 8. Resultado del Test Shapiro-Wilk

Pruebas de Normalidad para E.RESIDUOS		
Prueba	Estadístico	Valor-P
Estadístico W de Shapiro-Wilk	0,972152	0,117341

Figura 42. Gráficos de Normalidad de Residuos



4.3.2.4 Comparación de modelos SARMA. En el siguiente apartado se presenta la comparación de los modelos SARIMA analizados anteriormente, con el fin de definir cuál es el que mejor se ajusta a la serie y que sea un buen predictor para futuras eventualidades.

Tabla 19. Comparación de modelos SARIMA para la serie ingreso

Periodo de Estimación					
Modelo	RMSE	MAE	MAPE	ME	MPE
SARIMA(2,1,2) ₁₃ (0,0,0)	123.484.000,00	90.630.200,00	12,7679	6.151.670,00	-1,83886
SARIMA(2,1,2) ₁₃ (4,0,0)	121.549.000,00	90.022.300,00	12,5384	7.116.350,00	-1,48065
SARIMA(2,1,2) ₁₃ (0,0,4)	123.658.000,00	89.884.800,00	12,6706	1.431.770,00	-2,24746

La tabla anterior compara los resultados de cuatro diferentes modelos de pronósticos. Teniendo en cuenta las estadísticas del error, el modelo con la menor raíz cuadrada del error cuadrado medio (RMSE), con el menor porcentaje del error medio absoluto (MAPE) y el menor error porcentual durante el periodo de estimación del modelo es el modelo SARIMA(2,1,2)₁₃(4,0,0). Teniendo en cuenta

los anteriores resultados el modelo más apropiado es el modelo SARIMA(2,1,2)₁₃(4,0,0).

4.3.2.5 Validación de las últimas cinco observaciones. En la tabla 20 se observan los pronósticos de las 5 semanas subsiguientes a la fecha de corte para cada uno de los modelos.

Tabla 20. Pronósticos de 5 semanas subsiguientes para cada modelo SARMA

Modelo	Valor real	Valor pronosticado	Error relativo	Error relativo promedio
SARIMA(2,1,2) ₁₃ (0,0,0)	925.698.888,00	958.250.000,00	3%	6,0%
	901.281.412,00	1.020.240.000,00	12%	
	847.001.002,00	879.057.000,00	4%	
	772.025.205,00	704.604.000,00	10%	
	933.252.525,00	927.343.000,00	1%	
SARIMA(2,1,2) ₁₃ (4,0,0)	925.698.888,00	927.335.000,00	0%	7,0%
	901.281.412,00	1.085.490.000,00	17%	
	847.001.002,00	905.557.000,00	6%	
	772.025.205,00	695.386.000,00	11%	
	933.252.525,00	918.247.000,00	2%	
SARIMA(2,1,2) ₁₃ (0,0,4)	925.698.888,00	886.951.000,00	4%	7,0%
	901.281.412,00	1.068.910.000,00	16%	
	847.001.002,00	880.577.000,00	4%	
	772.025.205,00	746.278.000,00	3%	
	933.252.525,00	859.600.000,00	9%	

Aunque el modelo que proporciona un porcentaje de error más bajo en los pronósticos es el modelo SARIMA (2,1,2)₁₃(0,0,0), el modelo SARIMA (2,1,2)₁₃(4,0,0) presenta también un porcentaje bajo en los pronósticos, siendo este de 7%.

Por lo tanto, atendiendo al principio de parsimonia descrito en el capítulo III, se considera como mejor modelo el SARIMA (2,1,2)₁₃(4,0,0) ya que no solo cumple con tener estadísticas de error bajas, un promedio de error poco representativo en

los pronósticos, sino que además, posee un menor número de términos y por lo tanto un menor número de parámetros a estimar.

4.3.3 Comparación de modelos: SARIMA (2,1,2)₁₃(4,0,0) vs suavizamiento exponencial de Winters. La tabla 21 presenta la comparación de resultados de los modelos de pronósticos SARIMA (2,1,2)₁₃(4,0,0) y el suavizamiento exponencial de Winters. Teniendo en cuenta las estadísticas del error, el modelo con la menor raíz cuadrada del error cuadrado medio (RMSE), con el menor error medio absoluto (MAE) y con el menor porcentaje del error medio absoluto (MAPE) durante el periodo de estimación del modelo es el modelo SARIMA (2,1,2)₁₃(4,0,0).

Teniendo en cuenta los anteriores resultados el modelo más apropiado es el modelo SARIMA (2,1,2)₁₃(4,0,0) .

Tabla 21. Comparación de modelos

Periodo de Estimación					
Modelo	RMSE	MAE	MAPE	ME	MPE
Suavizamiento Exponencial de Winters	151.962.000,00	107.276.000,00	16,1141	-12.225.800,00	-5,96102
SARIMA(2,1,2) ₁₃ (4,0,0)	121.549.000,00	90.022.300,00	12,5384	7.116.350,00	-1,48065

4.3.3.1 Validación de las últimas cinco observaciones. En la tabla 22 se observan los pronósticos de las 5 semanas subsiguientes a la fecha de corte para cada uno de los modelos.

Tabla 22. Pronósticos de 5 semanas subsiguientes para cada modelo

Modelo	Valor real	Valor pronosticado	Error relativo	Error relativo promedio
SARIMA(2,1,2) ₁₃ (4,0,0)	925.698.888,00	927.335.000,00	0%	7,0%
	901.281.412,00	1.085.490.000,00	17%	
	847.001.002,00	905.557.000,00	6%	

Modelo	Valor real	Valor pronosticado	Error relativo	Error relativo promedio
	772.025.205,00	695.386.000,00	11%	
	933.252.525,00	918.247.000,00	2%	
Suavizamiento Exponencial de	925.698.888,00	792.641.000,00	17%	9,55%
	901.281.412,00	1.016.310.000,00	11%	
	847.001.002,00	915.472.000,00	7%	
	772.025.205,00	797.045.000,00	3%	
	933.252.525,00	856.112.000,00	9%	

Teniendo en cuenta la tabla anterior, se puede concluir que el modelo que proporciona los mejores pronósticos es el modelo SARIMA $(2,1,2)_{13}(4,0,0)$, ya que presenta los menores promedios de error relativo para las cinco semanas pronosticadas.

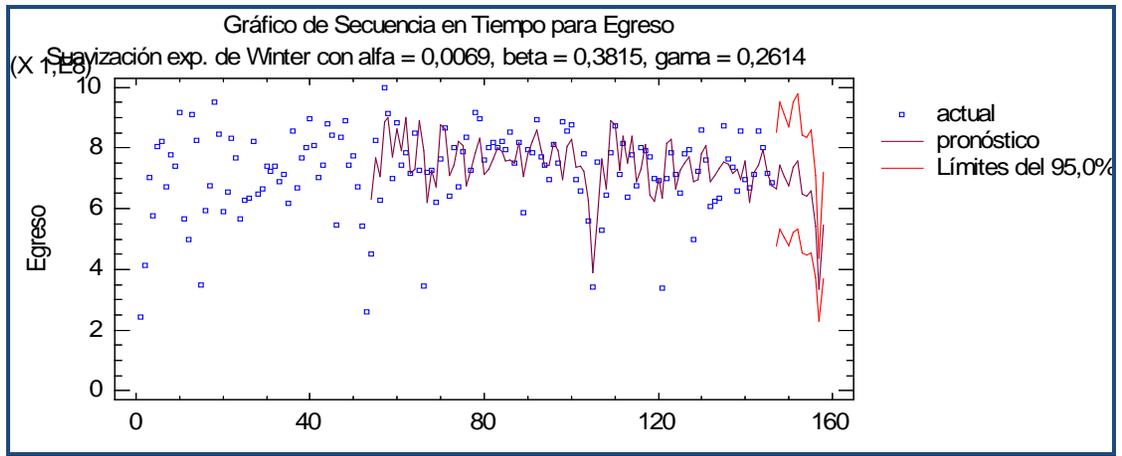
En conclusión, el modelo que mejor se ajusta y proporciona los mejores pronósticos para la serie ingreso es el modelo SARIMA $(2,1,2)_{13}(4,0,0)$.

4.4 ANÁLISIS DE MODELOS PARA LA SERIE EGRESO SEMANAL

La serie de Egresos semanales presenta estacionariedad en media y varianza, estacionalidad de orden 52 y no presenta tendencia. Teniendo en cuenta lo anterior se analizará el modelo exponencial de Winters y con ayuda de las funciones de autocorrelación descritos en el capítulo de descriptivo, se procede a analizar además, los modelos SARMA planteados para la serie Ingresos.

4.4.1 Modelo de suavizamiento exponencial de Winters. En la figura 43 se observa el ajuste del modelo de suavizamiento exponencial de Winters realizado a la serie Egresos.

Figura 43. Ajuste de la serie Egreso usando Suavizamiento Exponencial de Winters.



El modelo de suavizamiento exponencial de Winters queda descrito por las siguientes ecuaciones:

$$S_t = 0,0069 \frac{X_t}{I_{t-L}} + (0,9931)(S_{t-1} + T_{t-1})$$

$$T_t = 0,3815(S_t - S_{t-1}) + (0,6185)T_{t-1}$$

$$I_t = 0,2614 \frac{X_t}{S_t} + (0,7386)I_{t-L}$$

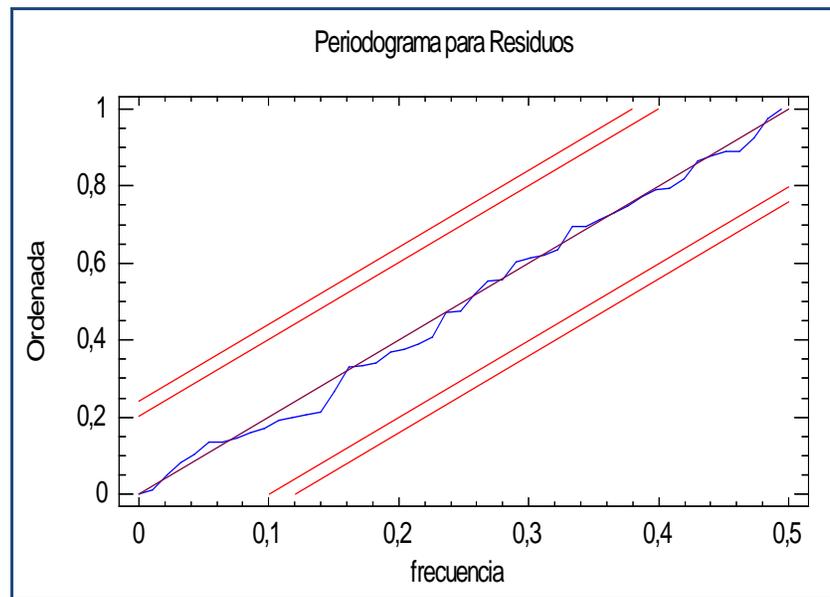
La predicción basada en el método de suavizamiento exponencial de Winters se calcula como:

$$F_{t+m} = \left[\left(0,0069 \frac{X_t}{I_{t-L}} + (0,9931)(S_{t-1} + T_{t-1}) \right) + (0,3815(S_t - S_{t-1}) + (0,6185)T_{t-1}) \right] \times \left[0,2614 \frac{X_t}{S_t} + (0,7386)I_{t-L} \right]$$

ANÁLISIS DE RESIDUOS

El periodograma de residuos de la figura 44 presenta un comportamiento de ruido blanco, lo que indica que el modelo de suavizamiento exponencial de Winters es eficiente al modelar el comportamiento de la serie Egresos.

Figura 44. Peridograma de Residuos



Lo anterior se puede comprobar analizando las pruebas o contrastes de aleatoriedad presentados en la figura 42, para determinar si los residuos forman, o no, una secuencia aleatoria de números.

La primera prueba cuenta el número de veces que la secuencia estuvo arriba o abajo de la mediana. El número de tales corridas es igual a 48, comparado con un valor esperado de 47,0 si la secuencia fuera aleatoria. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que los residuos son aleatorios, con un nivel de confianza del 95,0% o mayor.

La segunda prueba cuenta el número de veces que la secuencia ascendió o descendió. El número de tales corridas es igual a 63, comparado con un valor esperado de 61,6667 si la secuencia fuera aleatoria. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que la serie es aleatoria, con un nivel de confianza del 95,0% o mayor.

La tercera prueba está basada en la suma de cuadrados de los primeros 24 coeficientes de autocorrelación. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que la serie es aleatoria, con un nivel de confianza del 95,0% o mayor.

Figura 45. Contrastes de aleatoriedad de Residuos

(1) Corridas arriba o abajo de la mediana
Mediana = -1,77335E6
Número de corridas arriba o abajo de la mediana = 48
Número esperado de corridas = 47,0
Estadístico z para muestras grandes = 0,104835
Valor-P = 0,916502
(2) Corridas arriba y abajo
Número de corridas arriba y abajo = 63
Número esperado de corridas = 61,6667
Estadístico z para muestras grandes = 0,206972
Valor-P = 0,836027
(3) Prueba Box-Pierce
Prueba basada en las primeras 24 autocorrelaciones
Estadístico de prueba para muestras grandes = 13,8906
Valor-P = 0,874251

PRUEBAS DE NORMALIDAD DE RESIDUOS

Para validar el modelo de suavizamiento exponencial de Winters se realizaron las siguientes pruebas de normalidad a los residuos.

Test Shapiro – Wilk

El test de Shapiro-Wilk plantea las siguientes hipótesis:

H_0 = Los residuos del modelo presentan un comportamiento normal

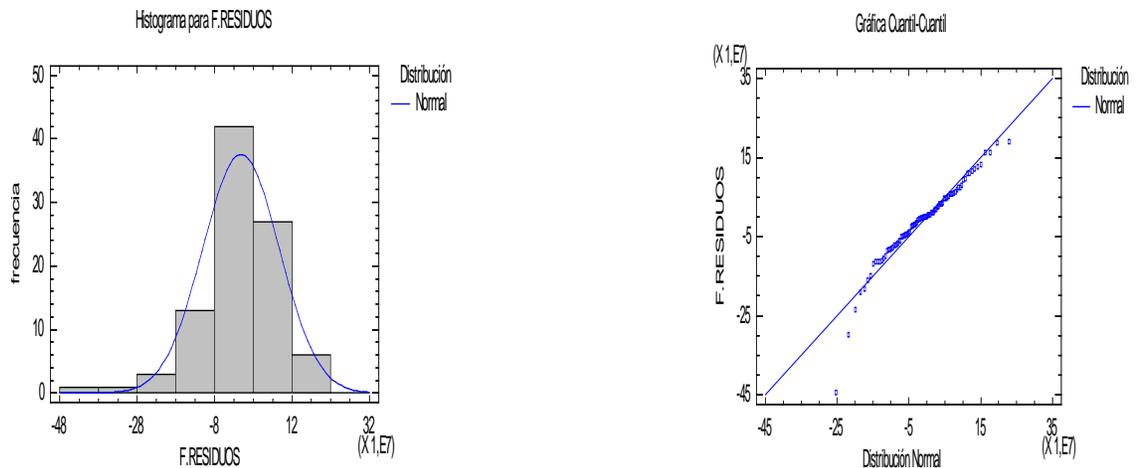
H_a = Los residuos del modelo No presentan un comportamiento normal

El p valor del test de Shapiro-Wilk (ver cuadro 9) es mayor que el nivel de significancia de 0,05 por tanto no se rechaza la hipótesis nula sobre la normalidad de los residuos. Esto también se puede observar en los gráficos de probabilidad normal (ver figura 46); por lo que se concluye que los residuos provienen de una distribución normal, es decir, que son aleatorios entre sí.

Cuadro 9. Resultado del Test Shapiro-Wilk

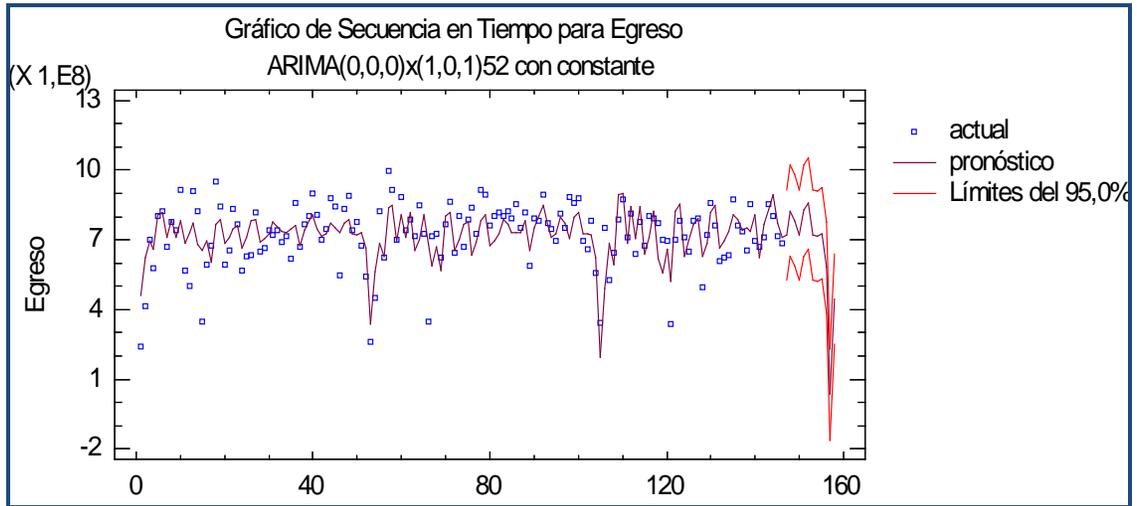
Pruebas de Normalidad para RESIDUOS		
Prueba	Estadístico	Valor-P
Estadístico W de Shapiro-Wilk	0,948121	0,00276207

Figura 46 Gráficos de Normalidad de Residuos



4.4.2 Modelo SARMA(1,0,1)₅₂. En la figura 47 se observa el ajuste del modelo SARMA (1,0,1)₅₂ realizado a la serie Egresos.

Figura 47. Ajuste de la serie Egreso usando el modelo SARMA(1,0,1)52



PARÁMETROS

La tabla 23 resume la significancia estadística de los términos en el modelo de pronósticos. Es importante resaltar que términos con valores-P menores que 0,05 son estadísticamente diferentes de cero con un nivel de confianza del 95,0%. El valor-P para los términos SAR(1) y SMA(1) es menor que 0,05, de modo que son estadísticamente diferentes de 0. El valor-P para el término de la constante es menor que 0,05, de modo que es estadísticamente diferente de 0. La desviación estándar estimada del ruido blanco de entrada es igual a 9,8516E7.

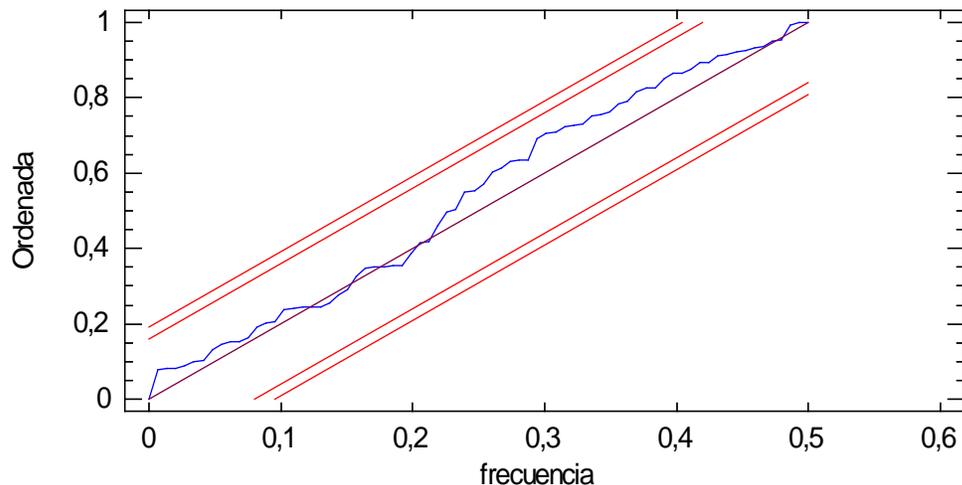
Tabla 23. Coeficientes del modelo SARMA(1,0,1)52 para la serie egresos

Resumen de Modelo SARMA				
Parámetro	Estimado	Error Estd.	T	Valor-P
SAR(1)	1,33928	0,0978602	13,6857	0,0000000
SMA(1)	1,20595	0,144171	8,36467	0,0000000
Media	7,28E+08	2,62E+07	27,8128	0,0000000
Constante	(246.967.000,00)			

ANÁLISIS DE RESIDUOS

El periodograma de residuos integrados de la figura 48 presenta un comportamiento de ruido blanco, lo que indica que el modelo $SARMA(1,0,1)_{52}$ es eficiente al modelar el comportamiento de la serie.

Figura 48. Peridograma de Residuos



Lo anterior se puede comprobar analizando las pruebas o contrastes de aleatoriedad presentados en la figura 46, para determinar si los residuos forman, o no, una secuencia aleatoria de números.

La primera prueba cuenta el número de veces que la secuencia estuvo arriba o abajo de la mediana. El número de tales corridas es igual a 73, comparado con un valor esperado de 74,0 si la secuencia fuera aleatoria. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que los residuos son aleatorios, con un nivel de confianza del 95,0% o mayor.

La segunda prueba cuenta el número de veces que la secuencia ascendió o descendió. El número de tales corridas es igual a 101, comparado con un valor

esperado de 97,0 si la secuencia fuera aleatoria. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que la serie es aleatoria, con un nivel de confianza del 95,0% o mayor.

La tercera prueba está basada en la suma de cuadrados de los primeros 24 coeficientes de autocorrelación. Puesto que el valor-P para esta prueba es mayor o igual que 0,05, no se puede rechazar la hipótesis de que la serie es aleatoria, con un nivel de confianza del 95,0% o mayor.

Figura 49. Contrastes de aleatoriedad de Residuos

(1) Corridas arriba o abajo de la mediana Mediana = 9,91659E6 Número de corridas arriba o abajo de la mediana = 73 Número esperado de corridas = 74,0 Estadístico z para muestras grandes = 0,0830475 Valor-P = 0,933808
(2) Corridas arriba y abajo Número de corridas arriba y abajo = 101 Número esperado de corridas = 97,0 Estadístico z para muestras grandes = 0,691298 Valor-P = 0,489376
(3) Prueba Box-Pierce Prueba basada en las primeras 24 autocorrelaciones Estadístico de prueba para muestras grandes = 22,7934 Valor-P = 0,413473

PRUEBAS DE NORMALIDAD DE RESIDUOS

Para validar el modelo $SARMA(1,0,1)_{52}$ se realizaron las siguientes pruebas de normalidad a los residuos.

Test Shapiro – Wilk

El test de Shapiro-Wilk plantea las siguientes hipótesis:

H_0 = Los residuos del modelo presentan un comportamiento normal

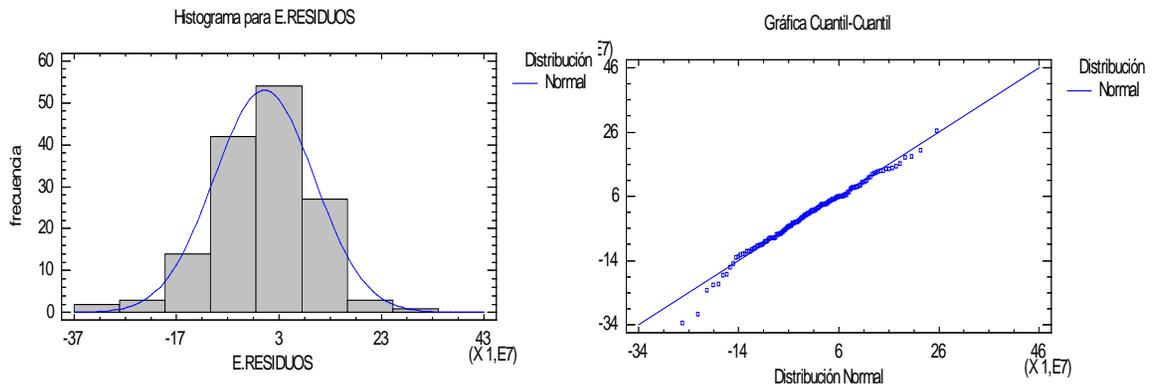
H_a = Los residuos del modelo No presentan un comportamiento normal

El p valor del test de Shapiro-Wilk (ver cuadro 10) es mayor que el nivel de significancia de 0,05 por tanto no se rechaza la hipótesis nula sobre la normalidad de los residuos. Esto también se puede observar en los gráficos de probabilidad normal (ver figura 50); por lo que se concluye que los residuos provienen de una distribución normal, es decir, que son aleatorios entre sí.

Cuadro 10. Resultado del Test Shapiro-Wilk

Pruebas de Normalidad para RESIDUOS		
Prueba	Estadístico	Valor-P
Estadístico W de Shapiro-Wilk	0,983137	0,613692

Figura 50 Gráficos de Normalidad de Residuos



4.4.3 Comparación de modelos: SARMA (1,0,1)₅₂ vs suavizamiento exponencial de Winters para la serie egresos semanales. La tabla 24 presenta la comparación de resultados de los modelos de pronósticos SARMA(1,0,1)₅₂ y suavizamiento exponencial de Winters para la serie de Egresos semanales. Teniendo en cuenta las estadísticas del error, el modelo con la menor raíz cuadrada del error cuadrado medio (RMSE), con el menor error medio (ME) y con el menor porcentaje del error medio (MPE) durante el periodo de estimación del

modelo es el modelo SARMA(1,0,1)₅₂. Por lo tanto, el modelo más apropiado, es decir el que mejor se ajusta a la serie de egresos semanales es el modelo SARMA(1,0,1)₅₂.

Tabla 24. Comparación de modelos para la serie egresos

Periodo de Estimación					
Modelo	RMSE	MAE	MAPE	ME	MPE
ARMA(1,0,1) ₅₂	98.496.100,00	76.331.700,00	12,30	1.309.430,00	(2,43)
Suavizamiento Exponencial de Winter	100.593.000,00	70.446.100,00	11,18	(11.352.700,00)	(3,98)

4.4.4 Validación de las últimas cinco observaciones. En la tabla a continuación se encuentra la matriz de correlación entre los valores reales y los valores pronosticados, observando la fuerte correlación existente para cada uno de los modelos, lo cual indica que son buenos predictores.

Tabla 25. Matriz de correlaciones para los modelos de la serie egresos

MATRIZ DE CORRELACIONES			
Modelo	Matriz	Valor real	valor pronosticado
SARMA (1,0,1) 52	Valor real	1	0,993866769
	valor pronosticado	0,99386677	1
Suavizamiento exponencial de Winters	Valor real	1	0,999559925
	valor pronosticado	0,99955993	1

En la tabla 26 se observan los pronósticos de las 5 semanas subsiguientes a la fecha de corte para cada uno de los modelos.

Tabla 26. Pronósticos de 5 semanas subsiguientes para cada modelo

Modelo	Valor real	Valor pronosticado	Error relativo	Error relativo promedio
SARMA (1,0,1) 52	744.528.966,00	720.765.000,00	3,297%	4,063%
	874.125.632,00	825.267.000,00	5,920%	
	801.222.258,00	784.473.000,00	2,135%	

Modelo	Valor real	Valor pronosticado	Error relativo	Error relativo promedio
	695.236.365,00	722.026.000,00	3,710%	
	869.525.366,00	826.116.000,00	5,255%	
Suavizamiento exponencial de Winter	744.528.966,00	664.950.000,00	11,968%	12,843%
	874.125.632,00	741.703.000,00	17,854%	
	801.222.258,00	706.824.000,00	13,355%	
	695.236.365,00	674.236.000,00	3,115%	
	869.525.366,00	737.375.000,00	17,922%	

Teniendo en cuenta la tabla anterior, se puede concluir que el modelo que proporciona los mejores pronósticos es el modelo $SARMA(1,0,1)_{52}$, ya que presenta los menores promedios de error relativo para las cinco semanas pronosticadas.

En conclusión, el modelo que mejor se ajusta y proporciona los mejores pronósticos para la serie ingreso es el modelo $SARMA(1,0,1)_{52}$.

5. CONCLUSIONES

En este capítulo se presentan las conclusiones generales de la investigación, en donde se analizan los efectos que trae consigo el uso de los diferentes modelos escogidos para cada una de las series (ingresos y egresos).

Los modelos seleccionados para las series de ingreso y egreso semanal SARIMA $(2,1,1)_{13}(0,0,4)$ y SARMA $(1,0,1)_{52}$ respectivamente, además de proporcionar un buen ajuste a la serie, brinda los mejores pronósticos. Lo anterior se evidencia en el bajo promedio de error relativo que estos presentan al ser comparados con los valores reales.

Los pronósticos generados por cada una de las series, guardan la misma tendencia de los datos reales, permitiendo determinar la tendencia futura tanto de los ingresos como de los egresos de caja, lo cual permite pronosticar de manera adecuada y con precisión, teniendo en cuenta que se está tratando con cifras monetarias que podrían implicar sobrecostos por una información inexacta.

El modelo seleccionado para cada una de las series fue seleccionado, ya que cumple con los siguientes requisitos:

En cuanto a los residuales la diferencia entre el valor original de la serie y el valor estimado por el modelo, se aproxima a un comportamiento de ruido blanco.

Al observar el periodograma integrado de residuos no se observan valores significativamente diferentes de cero, como indicativo de ausencia de correlación serial.

Los parámetros de los modelos seleccionados son significativamente diferentes de cero.

El grado de ajuste es elevado en comparación con otros modelos alternativos. La bondad del ajuste se evaluó con la Desviación estándar Residual (DER). Se juzgó la adecuación del modelo en función de los errores de predicción frente a los otros modelos.

La utilización de los modelos es viable dentro de la empresa para pronosticar las salidas y entradas de efectivo a la agencia, para tomar decisiones financieras de vital importancia para la entidad y para revisar políticas de Tesorería en cuanto al manejo de inversiones cortoplacistas tales como inversiones en operaciones repo o simultáneas, que son operaciones en corto. Teniendo en cuenta que dada la esencia del negocio que es captar y colocar recursos, cualquier mecanismo que permita, con cierto grado de precisión minimizar las pérdidas de dinero, es básico para su aplicación y utilización de forma constante.

Después de seleccionar el modelo SARMA en particular y de estimar sus parámetros, se observa que efectivamente el modelo seleccionado ajusta los datos en forma razonablemente buena. No obstante es importante tener presente, que es posible que exista otro modelo que también lo haga. Es por esto que “...el diseño de modelos de Box-Jenkins es un arte más que una ciencia...” [Gujarati, Pág. 722].

Teniendo en cuenta las funciones de autocorrelación muestral y parcial de la serie saldo, se plantearon diversos modelos en la búsqueda de aquel que mejor se ajustara a la serie y pronosticara de forma efectiva, sin obtener resultados favorables. De una parte los parámetros de los modelos encontrados no resultaron significativos, por otra, los pronósticos obtenidos arrojan errores relativos promedio considerablemente elevados y teniendo en cuenta el contexto financiero en el cual se mueven los datos, las cifras altas que se manejan conllevan a rechazar la utilización de estos modelos, por lo que se trabajará con los pronósticos

proporcionados por la serie ingreso y la serie egreso, de acuerdo a los buenos resultados obtenidos.

BIBLIOGRAFÍA

BECERRA, R. (2003). *Modelos de administración del efectivo*. Disponible en: <http://rigobertobecerra.tripod.com/modelosdeefectivo.htm>

BOWEMAN, B, y O'CONNEL, R. (1993). *Pronósticos, series de tiempo y regression: Un enfoque aplicado* (4ª ed.). Editorial CENGAGE, Learning. México, 2006.

CERA, Y.C. (2010). *Administración del efectivo en la Universidad de las ciencias informáticas*. Proyecto de investigación, Universidad de la Habana, La Habana, Cuba.

COLLANTES, D.V. (2001). *Predicción con redes neuronales: comparación con las metodologías de Box y Jenkins*. Tesis de grado, Universidad de los Andes, Mérida Venezuela.

CORREA, M.E (2004). *Series de tiempo: Conceptos Básicos*. Medellín: Universidad Nacional de Colombia, Facultad de Ciencias, Departamento de Matemáticas.

DIAZ, J. (2002). *Determinación de las necesidades mínimas de efectivo en la empresa "Productos lácteos Habana"*. Disponible en: <http://www.ilustrados.com/tema/5955/Determinacion-necesidades-minimas-efectivos-empresa-Productos.html>

GUERRERO, V. (2003). *Análisis Estadístico de Series de Tiempo Económicas*. Segunda Edición, Editorial Thompson, México.

MISAS, M., LOPEZ E., ARANGO C., XHERNANDEZ J. (2003) *La demanda de efectivo en Colombia: Una caja negra a la luz de las redes neuronales*. BANCO DE LA REPUBLICA, Subgerencia de estudios económicos.

PEÑA, D. (2005). *Análisis de Series Temporales*. Alianza Editorial, Madrid.

RANGEL, R. (2007). *Administración del efectivo*. Universidad de los Andes. Mérida, Venezuela.

SOTO, M. (2007). *Modelo estocástico y algoritmos en línea para el problema de gestión del saldo en caja en el Banco Central del Ecuador*. Cuestiones económicas Vol. 23, No 1:1-13.

URQUIZA, Y. y BELLO E. *Administración del efectivo en la empresa de muebles "Ludema" en las Tunas*. Disponible en: <http://www.eumed.net/cursecon/ecolat/cu/2010/uhbc.htm>

WHEELWRIGHTMS, y MAKRIDAKIS, S. (2006). *Métodos de Pronósticos*. Editorial Limusa, S.A de C.V. Versión autorizada en español de la obra publicada en inglés con el título: Forecasting Methods for Management. Mexico.