

**UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
ESCUELA DE MATEMÁTICA**



**Trabajo de Graduación Titulado:
Análisis e identificación del patrón de ingresos económicos y su afectación
debido a veda por marea roja en el sector hotelero y restaurantes
en el municipio de La Libertad (1998-2014)**

**Presentado por:
Elizabeth Noemy Alvarenga Carbajal
Mirna Elizabeth Ortega de Sánchez**

**Para Optar al grado de:
Licenciada en Estadística.**

**Asesores:
Msc. Ricardo Salvador Ríos Márquez
Msc. Mario Giovanni Molina Masferrer**

Septiembre, 2015

San Salvador

El Salvador

Centroamérica

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
ESCUELA DE MATEMÁTICA

Trabajo de Graduación Titulado:
Análisis e identificación del patrón de ingresos económicos y su afectación
debido a veda por marea roja en el sector hotelero y restaurantes
en el municipio de La Libertad (1998-2014)

Presentado por:
Elizabeth Noemy Alvarenga Carbajal
Mirna Elizabeth Ortega de Sánchez

Asesor:
Msc. Ricardo Salvador Ríos Márquez

Asesor:
Msc. Mario Giovanni Molina Masferrer

Ciudad Universitaria, San Salvador, Septiembre de 2015.

Universidad de El Salvador

Rector Universitario:

Ing. Mario Roberto Nieto Lovo

Secretaria General:

Dra. Ana Leticia Zavaleta de Amaya

Fiscal General:

Lic. Francisco Cruz Letona

Facultad de Ciencias Naturales y Matemática

Decano:

Msc. Martín Enrique Guerra Cáceres

Secretario:

Lic. Carlos Antonio Quintanilla Aparicio

Escuela de Matemática

Director:

Dr. José Nerys Funes Torres

Secretaria:

Msc. Alba Idalia Córdova Cuéllar

Agradecimientos

A Dios nuestro padre celestial, por darnos el apoyo incondicional de principio a fin, por la sabiduría dada a cada una de nosotras, por darnos la fuerza para levantarnos aun cuando habíamos caído, gracias porque en su voluntad y propósito estaba el permitirnos culminar la licenciatura.

A nuestras familias por su apoyo moral y aporte económico, por animarnos en los momentos difíciles, por confiar en que podíamos cumplir esta meta y por ser nuestra compañía durante todo el proceso.

A todos los docentes que han sido parte de nuestra formación académica.

A nuestros asesores Msc. Ricardo Salvador Ríos Márquez y Msc. Mario Giovanni Molina Masferrer por ayudarnos en la etapa final de nuestra carrera con sus consultas y sugerencias, su aporte ha sido esencial para la elaboración del trabajo de grado.

A todas nuestras amigas y amigos por apoyarnos.

Mirna Ortega y Noemy Alvarenga.

Índice de Contenidos

1. Resumen.....	1
2. Introducción	2
2.1. Justificación del problema	4
2.2. Objetivos	5
2.2.1. General.....	5
2.2.2. Específicos	5
2.3. Planteamiento del Problema	7
3. Metodología	9
4. Análisis Teórico y Resultados	14
4.1. Antecedentes de Evaluaciones de Impacto de Floraciones Algales Nocivas: internacionales y nacionales.....	14
4.2. Floraciones Algales Nocivas	16
4.2.1. ¿Qué es una Floración Algal Nociva y cuáles son los factores que la generan?	17
4.2.2. Especies que Provocan las Floraciones Algales Nocivas	19
4.2.3. Duración de una Floración Algal Nociva (<i>Torres Chuquimarca, 2011</i>)	21
4.2.4. ¿Dónde ocurren las Floraciones Algales Nocivas y cuál es su Frecuencia?	22
4.2.5. Efectos de las Floraciones Algales Nocivas.....	22
4.3. Modelo de Series de Tiempo.	29
4.3.1. Introducción de las series temporales	29
4.3.2. Clasificaciones de las series temporales	30
4.3.3. Objetivos del análisis de series temporales	30
4.3.4. Componentes de una serie temporal	31
4.3.5. Clasificación descriptiva de las series temporales.....	32
4.3.6. Estimación de la tendencia.....	34
4.3.7. Diferenciación de la serie.....	35
4.3.8. Estimación de la estacionalidad	35
4.3.9. Modelos Autorregresivos (Modelos AR).....	36
4.3.10. Modelos de Medias Móviles (Modelos MA)	39
4.3.11. Modelos ARMA	43
4.3.12. Modelos estocásticos lineales no estacionarios homogéneos. Modelos ARIMA.....	45

4.4.	Análisis de series temporales con outliers e intervenciones	50
4.4.1.	Análisis de Intervención	50
4.4.2.	Modelos de Intervención	50
4.5.	Resultados de la aplicación teórica	74
4.5.1.	Análisis descriptivo de la serie temporal de ingresos mensuales en restaurantes del municipio de La Libertad.....	74
4.5.2.	Análisis descriptivo de la serie temporal de Ingresos mensuales en hoteles del municipio de La Libertad.	102
5.	Conclusiones y Recomendaciones.....	126
6.	Referencias bibliográficas y fuentes consultadas.	129
7.	Anexos.	131

Índice de Figuras

Figura 1. Sitios donde se han registrado Floraciones Algales o Mareas Rojas en la costa salvadoreña (Líneas Punteadas); 2009-2013. (Espinoza Navarrete, Amaya Monterrosa, & Quintanilla, 2013)	8
Figura 2. Esquema de la metodología BOX-JENKINS con intervención. (Jaimes Berrios & Quintanilla Aparicio, 2008).....	13
Figura 3. Floración Algal Nociva	16
Figura 4. Tipos de coloraciones de las mareas rojas	18
Figura 5. Imagen de una especie de microalgas: dinoflagelado	19
Figura 6. Imagen de una especie de microalgas: Surirella, una diatomea gigante.	20
Figura 7. Muertes de tortugas marinas causadas por especies de microalgas	28
Figura 8. Ejemplo de una Serie de tiempo equiespaciada (anual)	29
Figura 9. Análisis descriptivo de una Serie Temporal	30
Figura 10. Serie de tiempo con predicciones.....	31
Figura 11. Serie temporal estacionaria.	33
Figura 12. Serie temporal no estacionaria.	33
Figura 13. Comportamiento de la variable impulso y escalón en una serie temporal	54

Índice de Gráficos

Gráfico 1. Serie original. Ingresos económicos mensuales en dólares en los Restaurantes del municipio de La Libertad. (Enero 1998 – Junio 2014)	74
Gráfico 2. Descomposición de la serie de datos de los ingresos mensuales en los restaurantes del municipio de La Libertad. (Enero 1998 – Junio 2014)	75
Gráfico 3. Serie de ingresos mensuales de los restaurantes, pre intervención (Enero de 1998 – Agosto de 2008).....	76
Gráfico 4. FAS. Serie Restaurantes, pre intervención (Enero 1998 – Agosto 2008)	76
Gráfico 5. FAP. Serie de los ingresos de restaurantes, pre intervención (Enero 1998 – Agosto 2008)	77
Gráfico 6. Serie de ingresos económicos de los restaurantes, pre intervención transformada (log)	78
Gráfico 7. Serie de ingresos de los restaurantes, pre intervención transformada con una diferencia regular (d=1).....	80
Gráfico 8. FAS de la serie pre intervención de ingresos económicos de restaurantes transformada y diferenciada en la parte regular.....	81
Gráfico 9. FAP de la serie pre intervención de ingresos económicos de los restaurantes, transformada y diferenciada en la parte regular.....	82
Gráfico 10. FAS de la serie pre intervención de restaurantes, transformada y diferenciada en la parte regular y estacional.....	82
Gráfico 11. FAP de la serie pre intervención de los ingresos económicos de los restaurantes, transformada y diferenciada en la parte regular y estacional.	83
Gráfico 12. Residuos de la serie de los ingresos económicos de los restaurantes. Modelo ARI (2,1) x ARI (2,1) ₁₂ . (Enero 1998 a Junio de 2014).	86
Gráfico 13. FAS de los residuos. Modelo ARI (2,1) x ARI (2,1) ₁₂ . Serie de los ingresos económicos de los restaurantes. (Enero 1998 a Junio de 2014).	87
Gráfico 14. FAP de los residuos. Modelo ARI (2,1) x ARI (2,1) ₁₂ . Serie de los ingresos económicos de los restaurantes. (Enero 1998 a Junio de 2014).	87
Gráfico 15. Histograma de los residuos. Modelo ARI (2,1) x ARI (2,1) ₁₂ . Serie de los ingresos económicos de los restaurantes. (Enero 1998 a Junio de 2014).	88

Gráfico 16. Probabilidad normal de los residuos. Serie de los ingresos económicos de los restaurantes. Modelo ARI (2,1) x ARI (2,1) ₁₂	88
Gráfico 17. Residuos del modelo ARI (2,1) x ARI (2,1) ₁₂ con intervención Comienzo brusco y duración temporal. Serie de ingresos económicos de restaurantes.	91
Gráfico 18. FAS de los residuos del modelo ARI (2,1) x ARI (2,1) ₁₂ con intervención Comienzo brusco y duración temporal. Serie de los ingresos económicos mensuales de los restaurantes. ...	92
Gráfico 19. FAP de los residuos del modelo ARI (2,1) x ARI (2,1) ₁₂ con intervención Comienzo brusco y duración temporal. Serie de los ingresos económicos de los restaurantes.	92
Gráfico 20. Histograma residual del modelo ARI (2,1) x ARI (2,1) ₁₂ con intervención Comienzo brusco y duración temporal. Serie de ingresos económicos de los restaurantes.	93
Gráfico 21. Probabilidad normal de los residuos. Modelo ARI (2,1) x ARI (2,1) ₁₂ con intervención Comienzo brusco y duración temporal. Serie de ingresos económicos de los restaurantes	93
Gráfico 22. Residuos del modelo ARI (2,1) x ARI (2,1) ₁₂ con intervención Comienzo brusco y duración temporal comienzo gradual duración temporal. Serie de ingresos económicos de los restaurantes	94
Gráfico 23. FAS del modelo ARI (2,1) x ARI (2,1) ₁₂ con intervención comienzo gradual duración temporal. Serie de ingresos económicos de los restaurantes.....	95
Gráfico 24. FAS de los residuos del modelo ARI (2,1) x ARI (2,1) ₁₂ con intervención comienzo gradual duración temporal. Serie de ingresos económicos de los restaurantes.	95
Gráfico 25. Histograma residual del modelo ARI (2,1) x ARI (2,1) ₁₂ con intervención comienzo gradual y duración temporal. Serie de ingresos económicos de los restaurantes.	96
Gráfico 26. Probabilidad normal de los residuos. Modelo ARI (2,1) x ARI (2,1) ₁₂ con intervención comienzo gradual y duración temporal. Serie de ingresos económicos de los restaurantes.	96
Gráfico 27. Serie de ingresos económicos de los restaurantes transformada y serie superpuesta del modelo ARI (2,1) x ARI (2,1) ₁₂ con intervención Comienzo brusco y duración temporal.....	97
Gráfico 28. Detección de datos atípicos en los residuos del modelo ARI (2,1) x ARI (2,1) ₁₂ con intervención Comienzo brusco y duración temporal. Serie de ingresos mensuales de los restaurantes.	100

Gráfico 29. Detección de vedas por marea roja en los residuos del modelo ARI (2,1) x ARI (2,1) ₁₂ con intervención comienzo brusco y duración temporal. Serie de ingresos mensuales de los restaurantes.....	101
Gráfico 30. Serie original. Ingresos económicos mensuales en dólares en los Hoteles del municipio de La Libertad. (Enero 1999- Mayo 2014).....	102
Gráfico 31. Descomposición de la serie de datos de los ingresos mensuales en los hoteles del municipio de La Libertad. (Enero 1999- Mayo 2014).....	103
Gráfico 32. Serie de los ingresos económicos mensuales de los hoteles, antes de la intervención (Enero de 1999 – Agosto de 2008).....	104
Gráfico 33. FAS. Serie de los ingresos económicos mensuales de los hoteles, pre intervención (Enero 1999 - Agosto 2008).....	104
Gráfico 34. FAP. Serie de los ingresos económicos mensuales de los hoteles, pre intervención (Enero 1999- Agosto 2008).....	105
Gráfico 35. Serie de los ingresos económicos mensuales de los hoteles, pre intervención con una diferencia regular (d=1).....	106
Gráfico 36. FAS de la serie de los ingresos económicos mensuales de los hoteles pre intervención con una diferencia en la parte regular.	107
Gráfico 37. FAP de la serie de los ingresos económicos mensuales de los hoteles pre intervención con una diferencia en la parte regular.....	108
Gráfico 38. FAS de la serie de los ingresos económicos mensuales de los hoteles. Serie pre intervención, diferenciada en la parte regular (d=1) y en la parte estacional (D=1).....	109
Gráfico 39. FAP de la serie de los ingresos económicos mensuales de los hoteles, serie pre intervención, diferenciada en la parte regular (d=1) y en la parte estacional (D=1).....	109
Gráfico 40. Residuos del Modelo IMA (1,1) x ARI (2,1) ₁₂ . Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014.	112
Gráfico 41. FAS Residual del Modelo IMA (1,1) x ARI (2,1). Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014.....	113
Gráfico 42. FAP Residual del Modelo IMA (1,1) x ARI (2,1) ₁₂ . Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014.....	113
Gráfico 43. Histograma Residual del Modelo IMA (1,1) x ARI (2,1) ₁₂ . Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014.	114

Gráfico 44. Probabilidad Residual del Modelo IMA (1,1) x ARI (2,1) ₁₂ . Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014.	114
Gráfico 45. Residuos del Modelo IMA (1,1) x ARI (2,1) ₁₂ con intervención. Serie de los ingresos económicos mensuales de los hoteles (Enero de 1999 a Mayo de 2014).	117
Gráfico 46. FAS Residual. Modelo IMA (1,1) x ARI (2,1) ₁₂ con intervención. Serie de los ingresos económicos mensuales de los hoteles. Enero 1999 a Mayo de 2014.	118
Gráfico 47. FAP Residual. Modelo IMA (1,1) x ARI (2,1) ₁₂ con intervención. Serie de los ingresos económicos mensuales de los hoteles. Enero 1999 a Mayo de 2014.	118
Gráfico 48. Histograma Residual del Modelo IMA (1,1) x ARI (2,1) ₁₂ con Intervención. Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014.	119
Gráfico 49. Probabilidad Normal de los Residuos del Modelo IMA (1,1) x ARI (2,1) ₁₂ con Intervención. Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014.	120
Gráfico 50. Serie de los ingresos económicos mensuales de los hoteles superpuesto con el modelo IMA (1,1) x ARI (2,1) ₁₂ con intervención (Comienzo brusco y duración permanente)	121
Gráfico 51. Representación de datos atípicos del modelo residual IMA (1,1) x ARI (2,1) ₁₂ con intervención. Serie de los ingresos económicos mensuales de los hoteles.	123
Gráfico 52. Representación de períodos de vedas por mareas rojas en el modelo residual IMA x (1,1) x ARI (2,1) ₁₂ con intervención. Serie de los ingresos económicos mensuales de los hoteles.	124
Gráfico 53. Representación de períodos de vedas por mareas rojas y datos outliers sobre el modelo residual IMA x (1,1) x ARI (2,1) ₁₂ con intervención. Serie de los ingresos económicos mensuales de los hoteles.	125

Índice de Tablas

Tabla 1. Resumen de eventos recurrentes de la mareas roja en El Salvador.....	23
Tabla 2. Toxinas comunes y sus efectos. (MARES, 2005).....	25
Tabla 3. Pérdidas económicas mundiales causadas por marea roja	27
Tabla 4. Conducta de datos de la tendencia de una serie temporal según clasificación de Pegel .	34
Tabla 5. Correspondencia entre un modelo AR (p) y un modelo MA (q)	42
Tabla 6. Diferentes formas del Comportamiento de la variable escalón en una serie temporal continúa y sus respectivas ecuaciones.....	52
Tabla 7. Diferentes formas del Comportamiento de la variable impulso en una serie temporal continúa y sus respectivas ecuaciones.....	53
Tabla 8. Valores de b, r y s de modelos de intervención.....	55
Tabla 9. Test de Levene para homogeneidad de varianza (serie pre intervención)	78
Tabla 10. Test de Levene para homogeneidad de varianza (serie de restaurantes transformada pre intervención)	79
Tabla 11. Test de Dickey-Fuller para homogeneidad de medias (serie de ingresos de los restaurantes, pre intervención transformada).....	80
Tabla 12. Test de Dickey-Fuller para homogeneidad de medias (serie de ingresos de los restaurantes, pre intervención transformada con una diferencia regular)	81
Tabla 13. Análisis de estacionariedad e invertibilidad de los modelos identificados reducidos. Serie de los ingresos económicos mensuales de los restaurantes.	85
Tabla 14. Modelos propuestos con los estadísticos AIC y BIC (serie ingresos económicos de los restaurantes).....	89
Tabla 15. Detección de outliers y comparación de residuos del modelo $ARI(2,1) \times ARI(2,1)_{12}$ con intervención Comienzo brusco y duración temporal. Serie de Restaurantes.	98
Tabla 16. Test de Levene para Homogeneidad de Varianza (Serie de los ingresos económicos mensuales de los hoteles, pre intervención)	105
Tabla 17. Test de Dickey- Fuller para Homogeneidad de Medias (Serie de los ingresos económicos mensuales de los hoteles, pre intervención)	106
Tabla 18. Test de Dickey- Fuller para Homogeneidad de Medias (Serie de los ingresos económicos mensuales de los hoteles, pre intervención con una diferencia en la parte regular)	107

Tabla 19. Análisis de estacionariedad e invertibilidad de los modelos identificados reducidos. Serie de los ingresos económicos mensuales de los hoteles.....	111
Tabla 20. Modelos con los estadísticos AIC y BIC. Serie de los ingresos económicos mensuales de los hoteles.....	115
Tabla 21. Detección de outliers y comparación de los residuos del modelo de la serie de los ingresos económicos mensuales de los hoteles con intervención comienzo brusco y duración permanente: IMA (1,1) x ARI (2,1) ₁₂	122

1. Resumen

La presente investigación se basó en la creación de dos modelos estadísticos de series temporales ARIMA (p,d,q) x ARIMA (P,D,Q)₁₂, que permiten explicar la estructura de los ingresos en el sector restaurantes y hoteles del municipio de La Libertad, departamento de La Libertad, El Salvador. Específicamente para investigar si existe alguna asociación entre un evento de veda por marea roja y disminuciones en los ingresos de los sectores ya antes mencionados, así también analizar valores extremos (outliers) para identificar los eventos donde hubo disminuciones económicas debido a veda por marea roja o por otras causas.

La investigación se realiza indagando primeramente sobre investigaciones previas a ésta, referentes a mediciones de impacto de vedas por marea roja, posteriormente solicitando una base de datos de IVA al Ministerio de Hacienda para poder deducir una base de ingresos mensuales tanto para la Serie de Restaurantes como también para la de Hoteles. Luego se realizó el análisis de cada serie por medio de la metodología Box Jenkins, obteniendo un modelo que mejor se ajustaba, después se analizó los diferentes tipos de intervención y tomar la decisión de elegir el mejor modelo para cada una de las series en estudio. El software utilizado para el análisis de la información es R Studio (versión 0.98.1091).

De esta manera se concluye que analizando los residuos de los modelos que mejor se ajustan a las series de datos de restaurantes y hoteles, las disminuciones económicas de los ingresos en dichos sectores no están asociadas a veda por marea roja en el municipio de La Libertad, sino más bien, tienen relación con períodos estacionales, ya sea por fechas fuera de períodos vacacionales o de días feriados.

2. Introducción

El presente trabajo de investigación está referido a medir el impacto económico causado por una veda de marea roja en el sector hotelero y restaurantes¹ del municipio de La Libertad.

A nivel internacional se han llevado a cabo una serie de investigaciones relacionadas al impacto económico de las Floraciones Algales Nocivas (FAN), principalmente en Florida (EEUU). En El Salvador existe la Comisión Nacional de Marea Roja (CONAMAR)², éste, es un organismo que evalúa de forma permanente el comportamiento de las floraciones algales a través de la colecta de muestras a nivel nacional. Las muestras son analizadas por los laboratorios del Ministerio de Salud y de Toxina Marinas de la Universidad de El Salvador (LABTOX). (*Castellanos, 2012*), pero aún no se tiene conocimiento de estudios que cuantifiquen las pérdidas y daños generadas por las FAN en diferentes rubros: salud, economía, medio ambiente, negocios y otros, es por esa razón que en primer lugar se enviaron solicitudes para saber si existen antecedentes referente a la investigación en los distintos Ministerios Gubernamentales y de esta manera respaldar el trabajo.

La investigación se desarrolló haciendo uso de una base de datos proporcionada por el Ministerio de Hacienda referente al Impuesto al Valor Agregado (IVA) mensual que declaran el sector hotelero y restaurantes en el municipio de La Libertad, a partir de esta se deducen los ingresos mensuales de cada negocio y posteriormente se crea un modelo que permite medir el impacto económico del sector hotelero y restaurantes de la zona.

Respecto al contenido teórico de las FAN, se agrega información general: definiciones, factores y/o especies que la generan, tipos de FAN, duración, lugar de ocurrencia y la frecuencia, así como también los efectos que la provoca. Seguidamente se incluye teoría del modelo a elaborar para estimar el impacto económico ocasionado por veda de una FAN en el municipio de La

¹ Sector Hotelero y Restaurantes: Negocios formales, que declaran IVA en el Ministerio de Hacienda.

² La Comisión Nacional de Marea Roja (CONAMAR), conformada por el Ministerio de Medio Ambiente y Recursos Naturales(MARN), Ministerio de Agricultura y Ganadería (MAG) y el Ministerio de Salud (MINSAL).

Libertad. Se estima y valida el modelo, posteriormente se presentan los resultados obtenidos a través del uso del paquete R³.

³ Software de análisis de datos libres y de código abierto, es una base para el trabajo innovador e importante en la ciencia, la educación y la industria, fué escrito inicialmente por Robert Gentleman y Ross Ihaka. <http://www.r-project.org/contributors.html>

2.1. Justificación del problema

La investigación a realizar contribuirá a generar un modelo estadístico para medir las disminuciones económicas causadas por una Floración Algal Nociva en el sector hotelero y restaurantes en el municipio de La Libertad, ya que estos negocios tienen una importancia en la economía del país y en el turismo de la zona. Estos se ven afectados en épocas de marea roja debido a que las personas toman medidas muy radicales con respecto a las visitas turísticas y al ingerir los alimentos, especialmente el consumo de mariscos, dejando de comprar incluso aquellos que están libres de toxinas.

En el país, se investigará si hasta el momento existen investigaciones previas a esta, que estén relacionadas con la medición de impactos económicos causados por marea roja. Esta problemática se considera necesaria medirla, debido a que se ven afectados los inversionistas. Para ello, se iniciará este estudio en el municipio de La Libertad, por ser uno de los sitios afectados por la marea roja según registros del LABTOX⁴ y por su gran afluencia de turistas tanto nacionales como extranjeros. Esta investigación servirá de base metodológica para posteriores estudios sobre la restante zona costera del país.

Asimismo, los resultados del estudio ayudarán a crear una mayor conciencia de parte de los propietarios de los negocios antes mencionados informando a través de medios de comunicación sobre las medidas preventivas a tomar cuando sucede un evento de éstos y explicar de que no todos los productos pesqueros están contaminados por la sustancia tóxica, de esta manera se podrá orientar a la población y evitar en la medida de lo posible la disminución de ventas en el sector de restaurantes y alojamiento.

⁴ (Espinoza Navarrete, Amaya Monterrosa, & Quintanilla, 2013), páginas 8 y 9

2.2. Objetivos

2.2.1. General

Crear un modelo estadístico que permita explicar la estructura de los ingresos en el sector restaurantes y hoteles ante eventos de veda por marea roja en el municipio La Libertad, departamento de La Libertad, El Salvador.

2.2.2. Específicos

- Indagar sobre investigaciones realizadas por los diferentes Ministerios de El Salvador, para medir las disminuciones económicas que causa una marea roja en: salud, medio ambiente, agricultura, turismo y comercio; en las entidades correspondientes: Ministerio de Salud, Ministerio de Medio Ambiente y Recursos Naturales, Ministerio de Agricultura y Ganadería, Ministerio de Turismo, Ministerio de Economía, Banco Central de Reserva y Ministerio de Hacienda.
- Construir un cuadro de eventos de veda de marea roja registrados en Ministerios Nacionales, ONG y noticias periodísticas de diversos medios de comunicación
- Obtener en el Ministerio de Hacienda las dos bases de datos de la declaración del Impuesto al Valor Agregado (IVA) mensual que declara el sector hotelero y restaurantes del municipio de La Libertad.
- Preparar la base proporcionada por el Ministerio de Hacienda para deducir y estudiar los ingresos económicos del sector hotelero y restaurantes en el municipio de La Libertad.
- Determinar la existencia de períodos que pudieran influenciar los ingresos en el sector de restaurantes y hoteles. En caso de existir, modelizarlos por medio de un análisis de intervención.

- Estimar y validar un modelo estadístico de series temporales que permita captar el comportamiento de los ingresos en el sector restaurante y hoteles en el municipio de La Libertad.
- Realizar un análisis de valores extremos (outliers) para identificar los eventos donde hubo disminuciones económicas debido a veda por marea roja y por otras causas.
- Realizar un análisis de los residuos del modelo final de ambas series para luego comparar estos residuos con eventos de veda por marea roja y observar si existe una asociación entre ellos.
- Ajustar un modelo ARIMA para los ingresos económicos de los hoteles y restaurantes del municipio de La Libertad, utilizando la metodología de Box Jenkins.

2.3. Planteamiento del Problema

La periodicidad de ocurrencia de los fenómenos naturales marinos es más frecuente en los últimos años, estos ocasionan grandes estragos, a tal grado que han despertado el interés de investigadores y observadores en los cambios de estos eventos. Dentro de esos cambios están los de color en el mar y la fosforescencia nocturna. Estos eventos están asociados a masivas floraciones de micro algas, resultando uno de los principales problemas las Floraciones Algas Nocivas (FAN), conocidas comúnmente como Mareas Rojas, las cuales en el peor de los casos pueden llegar hasta causar la muerte.

Históricamente, el primer caso registrado de intoxicación humana por consumo de mariscos tóxicos, se produjo el 15 de junio de 1793 en la costa oeste de los Estados Unidos. Durante los dos siglos posteriores en esa zona, incluyendo Alaska, se han registrado numerosos casos de envenenamiento que causan la muerte por parálisis respiratoria. El 17 de octubre de 1885, ocurrió otra intoxicación masiva en el puerto de Wilhelshaven en el norte de Alemania. En julio de 1927, se produjo otra en San Francisco (EEUU) que causó la muerte con severos síntomas paralizantes de varias personas.

Las FAN ocurren con frecuencia creciente en las áreas costeras, no siendo la excepción El Salvador, donde hay especies Fito planctónicas que se clasifican por sus efectos, ya sean inocuas (ayudan a la productividad de la zona costera), nocivas (afectan peces y moluscos) y tóxicas (producen diferentes tipos de toxinas que afectan a los seres humanos). Según fuentes del Laboratorio de Toxinas Marinas de la Universidad de El Salvador (LABTOX) los últimos eventos registrados son: Agosto 2001, Noviembre 2005, Noviembre 2009, Abril 2010, Marzo 2012 y Septiembre 2013.

Los lugares que se ven afectados por este fenómeno ya están mapeados (1. Golfo de Fonseca, 2. Bahía de Jiquilisco, 3. Puerto de La Libertad, 4. Los Cóbano). Se tiene información confirmada por Taxonomía de las especies que en El Salvador hay más de 200 especies tóxicas de micro algas que producen venenos amnésicos, diarreicos y paralizantes.

Figura 1. Sitios donde se han registrado Floraciones Algales o Mareas Rojas en la costa salvadoreña (Líneas Punteadas); 2009-2013. (Espinoza Navarrete, Amaya Monterrosa, & Quintanilla, 2013)



Con esta problemática, el ámbito económico resulta afectado debido a la disminución de ingresos en los negocios de la zona costera, una de las áreas productivas son los restaurantes y hoteles, en particular los negocios que están situados en el municipio de La Libertad.

Por tanto, es necesario crear mediciones de impacto en el país, a través de procedimientos estadísticos después de ocurrido un evento de veda por marea roja con el objetivo de determinar específicamente como se ven afectadas los ingresos monetarios en restaurantes y hoteles ante un fenómeno de marea roja en el municipio de La Libertad, departamento de La Libertad. La base de datos que se utilizará para analizar este fenómeno, se proporcionó por el Ministerio de Hacienda, deduciéndose a través de la declaración del IVA por parte de los negocios formales (hoteles y restaurantes del municipio de La Libertad) los ingresos económicos mensuales de los restaurantes y hoteles de la zona antes mencionada.

Esta investigación tiene como finalidad proponer un modelo estadístico, en el que se pretende dar respuesta al impacto generado por la marea roja en la economía de negocios como: hoteles y restaurantes del sitio ya especificado. En este contexto, es relevante plantearse la siguiente interrogante: ¿Está asociada una disminución económica del sector hotelero y restaurantes en el municipio de La Libertad con una veda por marea roja?

3. Metodología

Motivados por una reunión con el director del LABTOX UES, se planteó la necesidad de crear algún modelo estadístico que pueda medir los costos económicos generados por una FAN en El Salvador; se consideró esta propuesta como posible problema para abordar en el trabajo de investigación. Para iniciar dicho estudio fué necesario revisar información relacionada con las FAN y de esta manera se conocieron los impactos negativos que ocasionan en nuestra nación, los cuales se ven reflejados en las noticias de los diferentes medios de comunicación del país y en páginas web de instituciones gubernamentales.

A continuación se describen los pasos para el desarrollo de esta investigación.

- Reunión con el director de LABTOX- UES, donde se propuso como tema de investigación estimar el impacto económico que causa una FAN en El Salvador, dadas las disminuciones que se han obtenido después de uno de estos eventos.
- Búsqueda de referencias bibliográficas de Floraciones Algales Nocivas.
- Revisión de libros y documentos sobre modelos para estimar disminuciones económicas de Floraciones Algales Nocivas a nivel mundial.
- Solicitud de base de datos al Ministerio de hacienda y petición de información a cerca de investigaciones previas a Ministerios como: Medio Ambiente y Recursos Naturales, Agricultura y Ganadería, de Economía y al Ministerio de Salud, así como también al Banco Central de Reserva.
- Estudiar, comprender y sistematizar la teoría referente al modelo para estimar disminuciones económicas ocasionadas por una FAN en Restaurantes y Hoteles del municipio de La Libertad.
- Identificar las variables que se tomaran en cuenta para la creación del modelo estadístico.

Ingresos: total de ventas mensuales en dólares.

Se obtendrá a través del Impuesto al Valor Agregado (IVA) mensual que declaran en el Ministerio de Hacienda el sector hotelero y restaurantes del municipio de La Libertad.

$$\text{Ingreso mensual} = \frac{\text{impuesto mensual}}{0.13}$$

Variable cuantitativa continúa. Representada en dólares.

Esta fórmula se aplica sobre cada uno de los impuestos mensuales declarados de hoteles y restaurantes.

Disminución económica: diferencia entre la media de los ingresos y los ingresos observados. Variable cuantitativa continua representada en dólares.

- Una vez considerada la base a estudiar, se procede a un Análisis descriptivo: se analizará la serie de tiempo por cada una de las veces en que se observe disminución de ingresos. Luego, se hará una revisión en medios de comunicación o registros de Ministerios competentes del motivo de dicha disminución, posiblemente sea ocasionada por una FAN u otro factor como: huracanes, tormentas, lluvias y otros, es posible que se vean interrumpidas por intervenciones, datos atípicos(outliers) o simplemente es un comportamiento estacional en la serie.

Se debe tomar en cuenta que en el caso de series temporales interrumpidas, las funciones de Autocorrelación simple y parcial estarán distorsionadas por el efecto de las mismas. Por lo que no serán adecuadas para poder identificar el modelo generador de la serie, entonces será necesario para tratar dichas interrupciones dos métodos que se aplicarán en la etapa de identificación:

1. Si el subconjunto de datos anteriores o posteriores a las intervenciones es suficientemente grande, se identifica el Modelo Autorregresivo Integrado de Media Móvil o ARIMA (acrónimo del inglés *Autoregressive Integrated Moving Average*) de la forma tradicional, tomando como base de datos el subconjunto más grande, incorporando después la modelización de las intervenciones.
2. Cuando antes o después de las intervenciones no hayan suficientes datos, o si se tiene alguna intuición respecto a la forma en que han afectado al patrón de comportamiento de la serie, se especifica la forma de la intervención y se estima. Luego, la intervención se elimina de la serie original obteniendo los residuos de la serie, esto se conoce como: “valores limpios”. Con los nuevos valores se identifica el modelo ARIMA de la forma tradicional, seguidamente, se combina el modelo de intervención y el modelo residual

para obtener el modelo de intervención global, el cual se estima conjuntamente por máxima verosimilitud.

Para averiguar qué modelo $ARIMA(p, d, q)$ x $ARIMA(P, D, Q)_s$ representa adecuadamente el comportamiento de la serie, se aplicará la metodología propuesta por Box y Jenkins (1970, 1976).

Esquema de la metodología

La metodología de Box Jenkins nos propone primeramente examinar si la serie temporal es estacionaria y/o estacional usando herramientas estadísticas como: Función de Autocorrelación Simple (FAS) y Función de Autocorrelación Parcial (FAP) de la serie original de los datos o gráfico de secuencia.

Si la serie no es estacionaria en varianza, se transforma la serie (raíz cuadrada, logaritmo natural y raíz cuadrada inversa), el Test de Levene permitirá detectar la mayor o menor aproximación a la homocedasticidad para cada una de las transformaciones ya mencionadas (a menor estadístico y mayor significancia, mejor es la transformación).

Se aplica el test de Dickey Fuller para ver si la serie es estacionaria en media, si resulta no estacionaria en media, se diferencia las veces que sea necesario hasta que pueda admitirse la estacionariedad, se suavizará la serie por medio de diferenciaciones regulares (d) y en caso presente estacionalidad, se diferenciará estacionalmente (D), lo normal es hacer una o dos diferenciaciones. Si la serie transformada y/o diferenciada tiene varianza constante y media cero, se dirá que la serie es estacionaria

Una vez la serie sea estacionaria tanto en media como en varianza, se presentan la FAS y la FAP, el comportamiento no debe tener decaimiento lento a lo largo del tiempo, además según el número de retardos de las autocorrelaciones fuera de los límites de confianza en cada una de las gráficas, se definen los posible modelos. Posiblemente exista más de un modelo candidato, pero estos se van depurando tomando de base el principio de parsimonia.

Por medio del método de Mínimos Cuadrados Incondicional, se estimarán los coeficientes con valores iniciales automáticos para la estimación. El estadístico t es útil para determinar si los parámetros son significativos, una vez se tienen los valores de los parámetros, se verifica si los

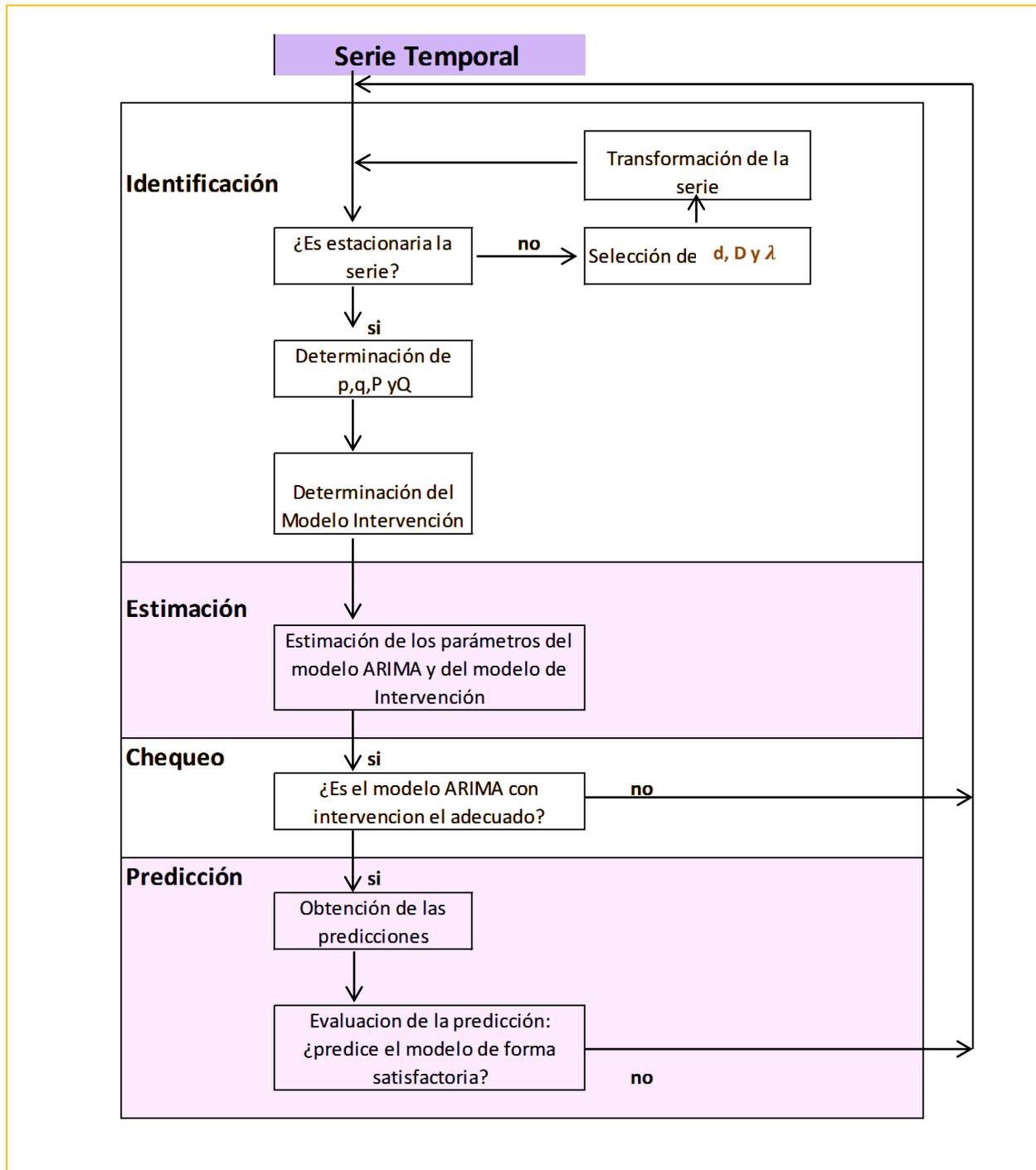
modelos cumplen las condiciones de estacionariedad e invertibilidad consideradas en el análisis teórico. Realizado el análisis anterior, se descartan algunos modelos.

Se analizan los residuos (diferencia entre el valor realmente observado y el valor previsto por el modelo) del modelo \hat{u}_t , observando si tienen un comportamiento similar al de las perturbaciones del mismo (ver si son semejantes a un ruido blanco), con el fin de comprobar si el ajuste del modelo a nuestros datos es adecuado. Para seleccionar el mejor modelo hay otros criterios: Criterio de Información Akaike (AIC), Criterio de Información Bayesiana (BIC), donde se seleccionará el modelo que tenga el menor valor de éstos estadísticos y el criterio de parsimonia, es decir, el modelo que tenga menos parámetros.

Los modelos que superen la etapa de Chequeo, tanto en la parte del ARIMA clásica y la parte de Intervenciones, pueden ya ser utilizados para el fin básico de la etapa de predicción, una vez obtenidas las predicciones del modelo con intervenciones, se trata de volver a chequear la adecuación del mismo, pudiendo utilizar métodos no paramétricos (Error Cuadrático Medio), como paramétricos. Mientras que para estimar los efectos de los outliers se hará por medio de los procesos iterativos mencionados en el marco teórico, recordando que es necesario tener los valores de los parámetros ARMA (caso más simple, puede ser otro modelo), conocer la forma del impacto y el tiempo en que ocurrió un outliers. Una vez se encuentran los parámetros, se procede a realizarlo como en las intervenciones con la diferencia que estos datos solamente serán estimados.

Todo el análisis de ambas bases de datos se hará por medio del paquete estadístico R.

Figura 2. Esquema de la metodología BOX-JENKINS con intervención. (*Jaimes Berrios & Quintanilla Aparicio, 2008*)



4. Análisis Teórico y Resultados

4.1. Antecedentes de Evaluaciones de Impacto de Floraciones Algales Nocivas: internacionales y nacionales.

En los últimos años se han generado una serie de estudios acerca de los Impactos Económicos de las Floraciones de Algas Nocivas (FAN) en diferentes países, tal es el caso de Canadá, Europa (Finlandia, Holanda y Bulgaria) y Estados Unidos (Maine, Texas, Carolina del Norte, Washington, Florida, New York, Maryland y Hawái); siendo estos estudios realizados con más frecuencia en el estado de Florida.

Además, existen investigaciones que preceden al presente estudio:

- 29 Trabajos de investigación
- 20 Documentos resúmenes de investigaciones de especialistas en el área y
- 7 Trabajos de investigación intermedia

Todos los anteriores realizados en el periodo comprendido entre 1974 a 2012.

Hay colaboradores en investigación de las FAN de la Universidad de Florida que tuvieron a bien realizar un resumen de todas las investigaciones, estas son realizadas por Charles M. Adams y Sherry L. Larkin, ambos profesores de la Universidad de Florida.

Se conoce que el estudio y análisis de los impactos económicos de las FAN es extenso, ya sea para estimar disminuciones de un evento o de varios eventos en el tiempo. Se puede agregar que las investigaciones han sido realizadas en lagos, bahías, condados e inclusive en un ámbito nacional por períodos de tiempo diferentes, sean éstos: días, semanas, meses o años. Los estudios suelen utilizar una serie temporal de datos para identificar los cambios en actividades empresariales durante una FAN.

De cualquier forma, existen limitaciones para las diferentes investigaciones internacionales que se han desarrollado hasta el momento, ya sea por la disponibilidad de datos para evaluar el impacto de un determinado evento o por qué la FAN sucede en un lugar que no se considera como sitio de muestreo y en una fecha donde no corresponde monitorear el área. Los investigadores reconocen que conlleva mucho tiempo la recolección de los datos primarios en la

investigación y es costoso para los residentes afectados, empresas, agencias gubernamentales e investigadores. Con respecto a las investigaciones sobre efectos directos e indirectos, éstos últimos no han sido explorados a fondo, ya que los efectos son retardados. Pocos estudios han investigado la intensidad de una FAN y ninguno parece abordar el potencial de una relación no lineal entre las disminuciones económicas y la duración de una FAN; así mismo ningún estudio ha utilizado series de tiempo y datos de corte transversal a la vez.

De modo idéntico a muchos países, El Salvador, con una actividad costera importante, está continuamente amenazado por la aparición de FAN, las cuales han aumentado su frecuencia e intensidad en la última década, causando daños en la salud de la población por consumo de alimentos marinos contaminados con biotóxicas marinas, además, de causar impactos negativos en la economía y en el ambiente marino.

Es importante agregar que en el país, el problema es muy complejo, ya que la población desconoce el significado de una FAN y los efectos que genera. Los habitantes de la zona costera, comerciantes del lugar o de mercados importantes de la región y pescadores, creen que una FAN conocida comúnmente como “Marea Roja” es invento de empresarios o de instituciones de gobierno para que los consumidores compren carnes o pollo, dándole poca relevancia al fenómeno.

Ahora bien, se conoce que cuando hay presencia en las playas de una FAN, se debe decretar veda, los afectados en su mayoría son comerciantes, extractores y las personas que consumen el producto del mar, incluyendo cadenas hoteleras, restaurantes, mercados, centros turísticos y otros. Debido al poco conocimiento de los productos que se pueden consumir cuando esto sucede, las personas descartan comprar cualquier producto que provenga del mar, en el peor de los casos si hay una persona fallecida, los familiares de este renuncian muchas veces de por vida a comer productos marinos.

Además, en la nación Salvadoreña aún no se tiene conocimiento de registros de investigaciones acerca de la medición del daño que una Floración Algal Nociva provoca; sin embargo, hay un único laboratorio en implementar un método de detección de toxinas marinas, este es el Laboratorio de Toxinas Marinas de la Universidad de El Salvador (LABTOX); las investigaciones que desarrolla van encaminadas a determinar cuándo y bajo qué condiciones se

produce las FAN, pero ante la necesidad de cuantificar el costo económico que esta genera, nace la idea de crear un modelo estadístico para medir los costos de dichos fenómenos, de manera que puedan ser útiles en la toma de decisiones con respecto a la prevención, mitigación y control de las disminuciones económicas ocasionadas por una FAN.

Por último, señalar que a pesar que es relevante y necesario realizar investigaciones sobre impactos económicos de las FAN, a nivel gubernamental, se le da muy poca importancia al fenómeno. Ahora bien, para hacer este tipo de estudios, se necesitan registros de datos de efectos (económicos, de salud y ambientales) del fenómeno de las diferentes instituciones gubernamentales. La base de datos con que se cuenta esta registrada en el Ministerio de Hacienda sobre el Impuesto al Valor Agregado (IVA) mensual que declaran el sector hotelero y restaurantes en el municipio de La Libertad y otros datos de instituciones que sean necesarios para la creación del modelo, tomando como referencia para el trabajo de grado una publicación realizada por Sherry L. Larkin y Charles M. Adams referente a Consecuencias Económicas en negocios de Florida causadas por Floraciones Algales Nocivas (Adams L. &., 2007).

4.2. Floraciones Algales Nocivas

Las Floraciones Algales Nocivas (FAN, en inglés HAB = *Harmful Algal Blooms*), es un término adoptado por la Comisión Oceanográfica Intergubernamental (COI) de la UNESCO y aceptado internacionalmente para denominar cualquier proliferación de microalgas (independientemente de su concentración).

Figura 3. Floración Algal Nociva



Microalgas: son el alimento natural de moluscos bivalvos filtradores (ostras, ostiones, choritos, cholgas, huepos, almejas, entre otros), de peces herbívoros y de larvas de diferentes organismos

marinos, tales como moluscos y crustáceos, que a su vez son alimento marino esencial para los seres humanos. (Reguera, Alonso, Moreira, & Méndez, 2011)

4.2.1. ¿Qué es una Floración Algal Nociva y cuáles son los factores que la generan?

Las Floraciones Algales Nocivas, es un fenómeno que se produce por un aumento de la población de algún tipo de microalga, debido a diferentes factores oceánicos como temperatura, luminosidad, salinidad, corrientes y por otros factores como la contaminación producida por el ser humano y que finalmente es eliminada en el mar.

Las "floraciones de algas", son beneficiosas para la acuicultura y la producción de los recursos marinos que se alimentan de estos organismos. En los últimos años, son numerosos los ejemplos del cultivo a escala industrial de algunas de estas microalgas, precisamente como alimento de peces, moluscos y crustáceos o como materia prima para el aislamiento y la purificación de compuestos químicos naturales de uso farmacéutico. (Suarez Isla & Guzmán Méndez)

Sin embargo, no todas estas floraciones son beneficiosas, son numerosos los casos de floraciones de algas que han sido asociadas a severas intoxicaciones y muerte de seres humanos, de mamíferos marinos, aves y peces. De las casi 4.000 especies marinas del fitoplancton microscópico viviente estimadas, unas 300 especies pueden proliferar en tan alto número que alcanzan densidades de millones de células por litro de agua, siendo tóxicas alrededor de 60.

Sobre la base de los daños que ocasionan, se hace una clasificación de las FAN:

1. **Floraciones algales de elevada biomasa, no tóxicas**, que pueden causar daños de tipo físico-químico.
2. **Floraciones productoras de toxinas** que se transfieren a través de la cadena trófica.
3. **Floraciones algales ictiotóxicas** estas son productoras de toxinas hemolíticas o de sustancias oxidantes que se liberan al medio y causan mortandades de peces silvestres o en cultivo.
4. **Floraciones de microalgas productoras de toxinas** que se transfieren a través del aerosol marino y causan irritaciones.
5. **Floraciones de Cianobacterias** (CHAB = *Cyanobacterial Harmful Algal Blooms*)

Tipos de coloraciones de una floración algal nociva.

La marea roja son rápidos aumentos en la concentración de microalgas tóxicas o no tóxicas que colorean el agua. El tono o color de la marea roja, dependerá de las especies involucradas (poseen diversos pigmentos que imparten su color cuando la abundancia del organismo es elevada), en su mayor frecuencia son de color rojo y por eso se llaman mareas rojas, pero algunas forman parches o manchas superficiales con pocos metros de espesor, otras se pueden extender por cientos de kilómetros como una carretera, adquiriendo diferentes colores (por ejemplo, verde, pardo, amarillo, café, anaranjado) y en ocasiones no son visibles, la ausencia de la coloración es una de las características de mareas rojas tóxicas siendo muy peligrosas por las toxinas que contienen o por el daño físico que causan a los seres vivos, a continuación se muestra la Figura 4 que ilustra lo antes mencionado. *(Torres Chuquimarca, 2011)*

Figura 4. Tipos de coloraciones de las mareas rojas



4.2.2. Especies que Provocan las Floraciones Algales Nocivas

Algunas especies de microalgas producen potentes venenos o toxinas. Cuando estas microalgas son filtradas por los mejillones y otros bivalvos, las toxinas se acumulan en sus tejidos y se transmiten a niveles superiores de la red alimentaria, incluyendo a los seres humanos.

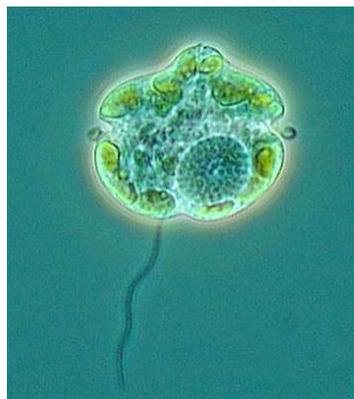
Todas estas toxinas son metabolitos que están presentes en el interior de las microalgas y no son eliminados al exterior. Estas microalgas son concentradas en los moluscos bivalvos (de dos conchas) como los choritos, almejas, cholgas, ostras, ostiones, culengues, entre otros. Cuando los moluscos filtran el agua para obtener sus alimentos, las microalgas quedan alojadas en el intestino de estos mariscos; por lo tanto, al consumir el marisco uno está ingiriendo también la toxina.

Las microalgas que causan las mareas rojas, corresponden a cianobacterias, diatomeas, dinoflagelados, silicoflagelados, primnesiofitas y rafidofitas; siendo los dinoflagelados y diatomeas los grupos más representativos y capaces de provocar eventos nocivos (tóxico o dañino) e impactos ecológicos y económicos reportados a nivel mundial.

Dinoflagelados y diatomeas. *(Suarez Isla & Guzmán Méndez)*

Los dinoflagelados en conjunto con las diatomeas son los principales componentes del fitoplancton microscópico. Los dinoflagelados marinos constituyen un grupo exitoso de microorganismos adaptados a hábitats *pelágicos* y *bentónicos*, con muchas especies cosmopolitas identificables en todos los mares.

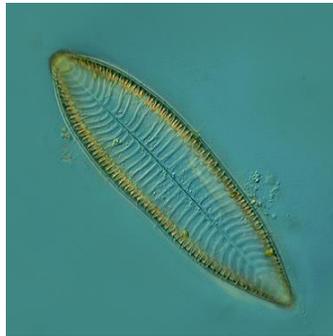
Figura 5. Imagen de una especie de microalgas: dinoflagelado



Los dinoflagelados son principalmente organismos unicelulares móviles propulsados por flagelos (Dodge, 1984). Aproximadamente la mitad de ellos poseen cloroplastos (como las plantas), organelos con los que pueden desarrollar fotosíntesis y ser autotróficos. El resto de los dinoflagelados que no tiene cloroplastos incorporan materia orgánica en suspensión o fagocitan bacterias u otros flagelados (son claramente *fagotróficos* o *saprotróficos*). No es sorprendente entonces que algunos autores los hayan clasificado como plantas y otros como animales.

Los dinoflagelados en general viven como células aisladas, tienen forma esférica y su diámetro mayor varía desde 5 milésimas de milímetro (o 5 micrómetros) hasta 2 mm. Desde el punto de vista evolutivo, estos organismos descienden de los más antiguos ancestros de los animales y plantas, las llamadas algas azul verdosas o cianófitas y que se remontan a más de 3.300 millones de años.

Figura 6. Imagen de una especie de microalgas: *Surirella*, una diatomea gigante.



Del total de especies de diatomeas, solamente tres han sido relacionadas con intoxicaciones humanas por consumo de mariscos y con muerte de aves. Las diatomeas tóxicas tienen formas alargadas (se les denomina *diatomeas pennadas*), son muy pequeñas y para su identificación taxonómica debe usarse el microscopio electrónico de barrido, aunque ciertos rasgos característicos pueden ser apreciados con un buen microscopio de campo luminoso dotado de contraste de fase.

Existen otros géneros de diatomeas (*Chaetoceros*, *Skeletonema*, *Leptocilindrus*) que han cobrado relevancia con el aumento de la piscicultura, técnica con la que se crían peces en espacios confinados en el medio marino. Estas diatomeas de formas aguzadas, al proliferar en un espacio limitado, dañan las branquias de los peces en cultivo, causándoles hemorragias, infecciones y eventualmente la muerte.

Tipos de venenos de diatomea y dinoflagelados.

Los grupos de toxinas de origen microalgal son:

- **Toxinas de tipo paralizante (PST = Paralytic Shellfish Toxins)**, producidas por dinoflageladas de los generos *Alexandrium*, *Gymnodinium* y *Pyrodinium*.
- **Toxinas lipofílicas (DST = Diarrhetic Shellfish Toxins, pectenotoxinas y yesotoxinas)**. Las toxinas diarreicas (ácido ocadaico y dinophysistoxinas) son producidas por dinoflageladas del genero *Dinophysis* y especies bentónicas del genero *Prorocentrum*; las pectenotoxinas por *Dinophysis* spp. y las yesotoxinas por *Gonyaulax spinifera*, *Lingulodinium polyedrum* y *Protoceratium reticulatum*.
- **Toxinas de tipo amnésico (AST = Amnesic Shellfish Toxins)**, producidas por diatomeas del genero *Pseudo-nitzschia*.
- **Toxinas neurotóxicas (NST = Neurotoxic Shellfish Toxins)**, producidas por especies del genero *Karenia*, tales como *K. brevis* en el Golfo de México.
- **Ciguatoxinas (CTX = Ciguatoxins)**, producidas por dinoflagelados bentónicos del genero *Gambierdiscus*.
- **Cianotoxinas** producidas por cianobacterias.

4.2.3. Duración de una Floración Algal Nociva (Torres Chuquimarca, 2011)

A pesar que el ciclo de vida de las Floraciones Algas Nocivas es muy corto pueden mantenerse desde pocos días a algunos meses, si las condiciones ambientales persisten. Suárez y Guzmán (2005), mencionan que la estrategia reproductiva de los dinoflagelados, incluyendo a las especies tóxicas, es reproducirse dos veces en 24 horas hasta una vez cada cinco días, lo que permite que de unas pocas células en el medio natural lleguen a producir una floración masiva al cabo de pocos días, estimulados por los factores ambientales y nutricionales.

4.2.4. ¿Dónde ocurren las Floraciones Algales Nocivas y cuál es su Frecuencia?

Las microalgas planctónicas ocurren tanto en agua salada como en agua dulce, la mayoría tiene hábitat en los primeros metros de agua (<5m); otras algas habitan en el sedimento de fondos someros se denominan bentónicas (Fogg, 2002) y sobre otros organismos (microalgas y arrecifes de coral) se designan epifitas o epizoicas. Se encuentran en estuarios, bahías o áreas costeras semi-encerradas donde no exista mucha corriente. Su ocurrencia es aperiódica e impredecible, aunque en algunos sectores se presentan con cierta periodicidad y en otros en forma ocasional. Se relacionan cuando aparece aguas superficiales estratificadas sobre aguas más frías y ricas en nutrientes como consecuencia del calentamiento o por desplazamientos de agua dulce (FAO, 2005). Las floraciones de las cianobacterias son más abundantes en los ecosistemas acuáticos continentales que en estuarios y ecosistemas marino costeros. (Del Valle, 2009)

4.2.5. Efectos de las Floraciones Algales Nocivas

Las Floraciones Algales Nocivas son percibidas como un daño por su impacto negativo en la salud pública, la acuicultura, el medio ambiente y las actividades recreativas. Cabe mencionar que algunas FAN pueden causar múltiples efectos nocivos y otras solo en determinada área. Esto se confirma en la opinión de Deborah Cedeño, candidata doctoral del Departamento de Ciencias Marinas de la Universidad de Puerto Rico en Mayagüez, cuyo trabajo investigativo está relacionado al dinoflagelado ictiotóxico (nocivo para los peces) *Cocholodinium polykrikoides*, indicó que este fenómeno de la marea roja acarrea consecuencias nocivas para la salud del ser humano, para el medio ambiente y para la economía (Del Valle, 2009). A continuación se muestra una tabla resumen con información sobre los registros de mareas rojas en El Salvador, esta ha sido elaborada de datos contenidos en el Atlas del fitoplancton marino del LABTOX-UES.

Tabla 1. Resumen de eventos recurrentes de la mareas roja en El Salvador⁵

FECHA	LUGAR	ESPECIE CAUSANTE DE LA MAREA ROJA	EFFECTOS
Noviembre 1989	desde Panamá hasta México, PRINCIPALMENTE Guatemala y México	Pyrodinium bahamense (Cortes- Altamirano, et al. 1993.)	99 intoxicados y 3 fallecidos
desde agosto de 2001 hasta enero de 2002	La Libertad uno de los más afectados, principalmente las playas desde Mizata hasta el puerto de La Libertad	Gymnodinium	41 intoxicados
Noviembre de 2005 a Marzo de 2006	El Icacal, La Unión; La Playa El Cuco, San Miguel; La Isla Menéndez en Puerto el Triunfo, Usulután	Pyrodinium bahamense var.compressum(Licea et al. 2007)	Se reportaron siete personas intoxicadas, tres fallecidos y al menos 206 tortugas marinas muertas
Agosto/Septiembre 2007 (un mes)	Desde La Libertad hasta la Bahía de Jiquilisco, Usulután	Pyrodinium bahamense	No hay registros
Noviembre de 2009 abril de 2010	Nov 2009. A 46 millas de la línea de costa en dirección de la bahía de Jiquilisco hasta Acajutla	Pyrodinium bahamense	5 tortugas muertas
	Abril 2010. intensas manchas alcanzaron por completo la línea de la costa	Pyrodinium bahamense	17 tortugas muertas
24 de Octubre de 2010(una semana)	Los Cobanos	Alexandrium cf peruvianum/Ostenfeldii	muerte de una docena de tortugas
Agosto/Septiembre de 2011 (20 días)	Puerto de La Libertad	Pyrodinium bahamense	No hay registros

⁵ Tomado de (Espinoza Navarrete, Amaya Monterrosa, & Quintanilla, 2013)

Marzo/Mayo 2012	Desde El Puerto de La Libertad hasta La Perla, Libertad. Se generalizo en toda la línea Costera del país, desplazándose hasta Guatemala.	Cochlodinium polykrikoides	Mortandad de peces bentónicos
18 Septiembre de 2012	Playas de Taquillo, El Zonte y Mizata	Pyrodinium bahamense	No hay registros
Desde el 28 de Septiembre al 21 de Octubre de 2013	Diferentes playas de la Costa Salvadoreña, principalmente en La Libertad	Alexandrium monilatum, Gymnodium catenatum, Pyrodinium bahamense, Alexandrium cf peruvianum/Ostenfeldii, Gyrodinium Instriatum, Cochlodinium polykrikoides	233 tortugas muertas

Impacto de salud

En los seres humanos, la Floración Algal Nociva puede provocar síndromes tóxicos que llegan al organismo por vías de la alimentación. Al ingerir peces y mariscos como camarones, almejas, vieiras, ostras y mejillones, el ser humano puede sufrir de algún tipo de intoxicación como por ejemplo el envenenamiento paralítico, diarreico, neurotóxico y amnésico. (MARES, 2005)

Tres de las microalgas más comunes son las que producen distintas toxinas nocivas

Tabla 2. Toxinas comunes y sus efectos. (MARES, 2005)

Nombre de Toxina	Efecto
Alexandrium catenella:	Produce la Toxina Paralizante de los Mariscos (VPM). En el hombre y animales provoca parálisis de las extremidades y posteriormente la muerte por paro cardio-respiratorio.
Dinophysis acuta:	Produce la Toxina Diarreica de los Mariscos (VDM), en el ser humano provoca graves diarreas.
Pseudonitzchia sp:	Productora de la Toxina Amnésica de los Mariscos (VAM), en el ser humano provoca pérdida de la memoria de corto plazo y puede causar la muerte.

Impacto económico

Otra consecuencia que acarrearán los FAN para el ser humano además de la salud, es el daño económico, tanto para empresas exportadoras como para pescadores artesanales y población en general.

Existen grandes extensiones en el mar que están en veda, lo que hace que muchas familias vean disminuidos sus ingresos económicos. En el planeta hay regiones cuya única fuente de trabajo y subsistencia son los recursos que da el mar. Incluso, en lugares aislados es la única fuente de alimento. Esto hace que sea necesario estudiar el fenómeno para así poder comprender y convivir de mejor manera con la naturaleza que nos rodea. *(Tovar R, 2006)*

En el marco de la Alianza Estratégica de ARCAL-OIEA (2008), en la clasificación de los problemas ambientales de Latino América y el Caribe, se mencionan los impactos negativos económicos causados por los florecimientos algales nocivos, incluyendo el costo de programas de monitoreo rutinarios en mariscos y otros recursos afectados, cierre por corto tiempo o permanente de los stocks de peces y mariscos, reducción en las ventas de alimentos marinos, mortalidad de peces silvestres y cultivados, vegetación acuática sumergida y arrecifes coralinos, impacto en el turismo y negocios asociados, tratamientos médicos a poblaciones expuestas. *(Torres Chuquimarca, 2011)*

Hay que tener en cuenta que no sólo estamos hablando de pesca, sino también de acuicultura. Por citar un ejemplo, en el cultivo del mejillón (miticultura), cuando en las aguas donde se cultivan aumenta el número de peridinas, pequeños organismos que viven en el plancton, que producen mareas rojas, estos organismos se acumulan en el mejillón y entonces lo convierten en un alimento tóxico y peligroso.

Respecto a las personas dedicadas a la extracción de moluscos y que se ven fuertemente afectadas en su actividad, Chile ha buscado soluciones en cuanto a sus fuentes de trabajo y se piensa que podría desarrollarse con mayor fuerza el cultivo de los moluscos en aguas defendidas de la contaminación.

El siguiente cuadro muestra pérdidas económicas causadas por las Floraciones Algales Nocivas en las Pesquerías e Instalaciones de Acuicultura en diferentes años y países a nivel mundial.

Tabla 3. Pérdidas económicas mundiales causadas por marea roja⁶

Año	Lugar	Especie	Pérdida (millones de dólares)
1972	Japón	Jurel	~47
1977	Japón	Jurel	~20
1978	Japón	Jurel	~22
1978	Corea	Ostra	4.6
1979	Maine	Muchas especies	2.8
1980	Nueva Inglaterra	Muchas especies	7
1981	Corea	Ostra	>60
1985	Long Island. N.Y	Conchas	2
1986	Chile	Salmon Rojo	21
1987	Japón	Jurel	15
1988	Noruega, Suecia	Salmon	5
1989	Noruega	Salmon, Trucha, Arco Iris	4.5
1989/90	Puget Sound, W.A.	Salmon de Estanque	4.5
1991	Washington State	Ostra	133
1996	Texas	Ostra	24
1998	Hong Kong	Peces de Estanque	32

La Floración Algal Nociva es una de las principales amenazas que se cierne sobre todas las costas del mundo, se está impulsando la búsqueda de mecanismos para combatir sus efectos y mitigar los daños sociales y económicos, ya se han desarrollado métodos para la detección de toxinas y se trabaja en procedimientos para la desintoxicación y antídotos, basados en el uso de la biotecnología. (Torres Chuquimarca, 2011)

⁶ Tomado de <http://mareasrojas.asp.blogspot.com/2006/10/15-consecuencias-economicas-de-las.html>

Impacto ambiental

Por otra parte, Floración Algal Nociva también tiene efectos negativos sobre el medio ambiente. A nivel ecológico dichas floraciones causan: desplazamiento de especies endógenas, alteración de hábitats y de estructura trófica, condiciones de anoxia (falta casi total del oxígeno en un tejido), deterioro en la calidad del agua, mortandades masivas de peces y de organismos filtradores, tales como los bivalvos, y muerte de aves y de mamíferos marinos.

Figura 7. Muertes de tortugas marinas causadas por especies de microalgas⁷



La tetradotóxina y la ciguatera son tipos de intoxicaciones que afectan principalmente a ciertos peces y han sido descritas para zonas tropicales. A pesar de que la ciguatera es ocasionada por un dinoflagelado tóxico, ésta no forma floraciones. Estos efectos sobre el medio ambiente también tienen severas repercusiones sobre la industria turística, la acuicultura y la recreación, ya que si la floración es de color, se forman manchas en pequeñas playas y bahías visitadas por los turistas, creando contaminación visual, y alarma social si el público no está debidamente informado. (*Del Valle, 2009*)

El ejemplo más notorio lo constituyen las floraciones del dinoflagelado *Karenia brevis* en el Golfo de México: sus neurotoxinas (NSP = *Neurotoxic Shellfish Poisoning*) son transferidas a través de la cadena trófica, causando incluso la muerte de mamíferos marinos; sus floraciones de elevada biomasa causan muertes masivas de organismos marinos de fondo (ver Informe de México en IOCARIBE-ANCA 2007).

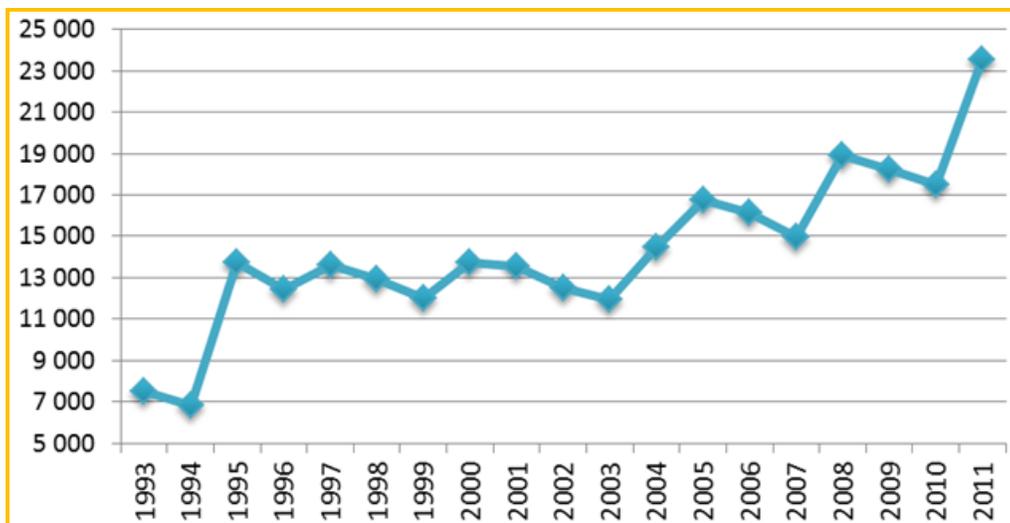
⁷ 233 tortugas muertas en las diferentes playas de la costa salvadoreña, principalmente en La Libertad. Desde el 28 de Septiembre al 21 de Octubre de 2013

4.3. Modelo de Series de Tiempo.⁸

4.3.1. Introducción de las series temporales

Una serie temporal se define como una colección de observaciones de una variable recogidas secuencialmente en el tiempo. Estas observaciones se suelen recoger en instantes de tiempo equiespaciados. Si los datos se recogen en instantes temporales de forma continua, es necesario digitalizar la serie, es decir, recoger los valores en instantes de tiempo equiespaciados, o bien acumular los valores sobre intervalos de tiempo.

Figura 8. Ejemplo de una Serie de tiempo equiespaciada (anual)



La característica fundamental de las series temporales es que las observaciones sucesivas no son independientes entre sí, y el análisis debe llevarse a cabo teniendo en cuenta el orden temporal de las observaciones. Los métodos estadísticos basados en la independencia de las observaciones no son válidos para el análisis de series temporales porque las observaciones en un instante de tiempo dependen de los valores de la serie en el pasado.

⁸ Aznar, A.: Métodos de Predicción en Economía, Tomo II, Análisis de Series Temporales, Ariel Economía, 1993.

4.3.2. Clasificaciones de las series temporales

Una serie temporal puede ser discreta o continua dependiendo de cómo sean las observaciones.

- Si se pueden predecir exactamente los valores, se dice que las series son determinísticas.
- Si el futuro sólo se puede determinar de modo parcial por las observaciones pasadas y no se pueden determinar exactamente, se considera que los futuros valores tienen una distribución de probabilidad que está condicionada a los valores pasados. Las series son así estocásticas.

4.3.3. Objetivos del análisis de series temporales

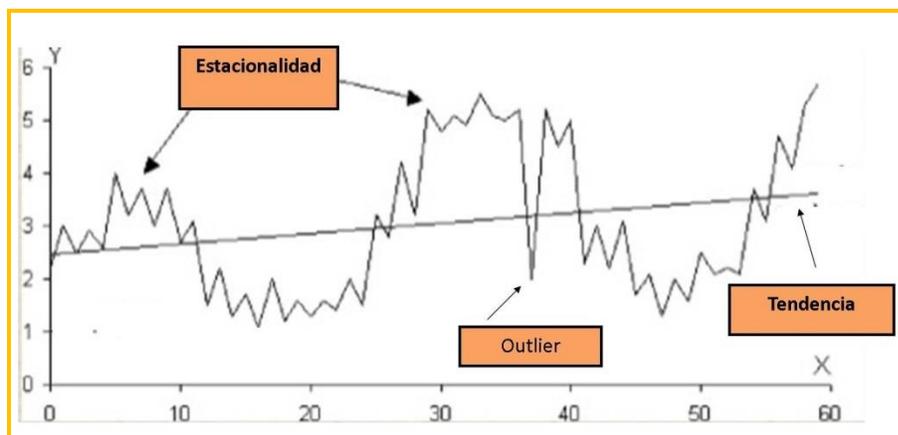
Se pueden considerar varios posibles objetivos:

1. Descripción

Cuando se estudia una serie temporal, lo primero que se tiene que hacer es la representación gráfica y luego considerar las medidas descriptivas básicas. Así, se tiene que considerar:

- a) Si los datos presentan forma creciente (tendencia).
- b) Si existe influencia de ciertos períodos de cualquier unidad de tiempo (estacionalidad).
- c) Si aparecen outliers (observaciones extrañas o discordantes).

Figura 9. Análisis descriptivo de una Serie Temporal



2. Predicción

Cuando se observan los valores de una serie, se pretende normalmente no sólo explicar el pasado, sino también predecir el futuro.

Figura 10. Serie de tiempo con predicciones



4.3.4. Componentes de una serie temporal

El estudio descriptivo de series temporales se basa en la idea de descomponer la variación de una serie en varias componentes básicas. Este enfoque no siempre resulta ser el más adecuado, pero es interesante cuando en la serie se observa cierta tendencia o cierta periodicidad. Hay que resaltar que esta descomposición no es en general única.

Este enfoque descriptivo consiste en encontrar componentes que correspondan a una tendencia a largo plazo, un comportamiento estacional y una parte aleatoria.

Las componentes o fuentes de variación que se consideran habitualmente son las siguientes:

1. **Tendencia:** Se puede definir como un cambio a largo plazo que se produce en relación al nivel medio, o el cambio a largo plazo de la media. La tendencia se identifica con un movimiento suave de la serie a largo plazo.
2. **Efecto Estacional:** Muchas series temporales presentan cierta periodicidad o dicho de otro modo, variación de cierto periodo (ejemplo: anual, mensual). Estos tipos de efectos

son fáciles de entender y se pueden medir explícitamente o incluso se pueden eliminar del conjunto de los datos, desestacionalizando la serie original.

3. **Componente Aleatoria:** Una vez identificados los componentes anteriores y después de haberlos eliminado, persisten unos valores que son aleatorios. Se pretende estudiar qué tipo de comportamiento aleatorio presentan estos residuos, utilizando algún tipo de modelo probabilístico que los describa.

De las tres componentes reseñadas, las dos primeras son componentes determinísticas, mientras que la última es aleatoria. Así, se puede denotar que

$$X_t = T_t + E_t + I_t$$

Donde T_t es la tendencia, E_t es la componente estacional, que constituyen la señal o parte determinística, e I_t es el ruido o parte aleatoria.

Es necesario aislar de alguna manera la componente aleatoria y estudiar qué modelo probabilístico es el más adecuado. Conocido éste, podremos conocer el comportamiento de la serie a largo plazo.

Este aislamiento de la componente aleatoria se suele abordar de dos maneras.

1. Enfoque descriptivo: Se estima T y E_t y se obtiene I_t como

$$I_t = X_t - T_t - E_t$$

2. Enfoque de Box-Jenkins: Se elimina de X_t la tendencia y la parte estacional (mediante transformaciones o filtros) y queda sólo la parte probabilística. A esta última parte se le ajustan modelos paramétricos

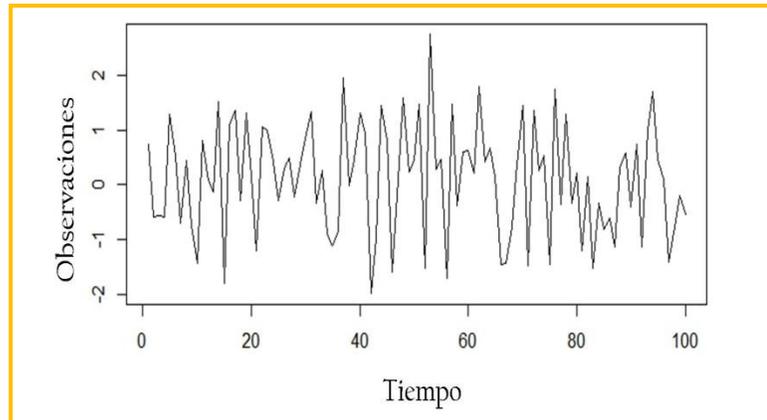
4.3.5. Clasificación descriptiva de las series temporales

Las series temporales se pueden clasificar en:

- **Estacionarias:** Una serie es estacionaria cuando la media y la variabilidad son constantes a lo largo del tiempo. Esto puede determinarse en la representación gráfica que los valores de la serie tienden a oscilar alrededor de una media constante y la variabilidad con respecto a esa media que también permanece constante en el tiempo. Es una serie básicamente estable a lo largo del tiempo, sin que se aprecien aumentos o disminuciones

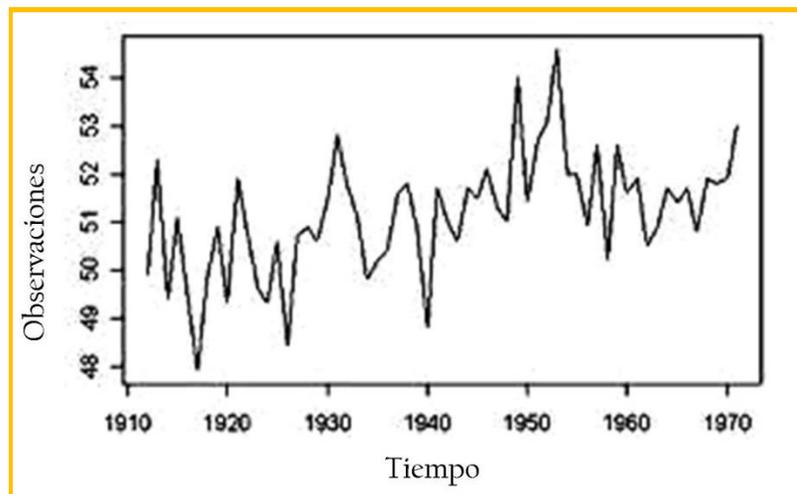
sistemáticos de sus valores. Para este tipo de series tiene sentido conceptos como la media y la varianza. Sin embargo, también es posible aplicar los mismos métodos a series no estacionarias si se transforman previamente en estacionarias.

Figura 11. Serie temporal estacionaria.



- **No Estacionarias:** Son series en las cuales la media y/o variabilidad cambian en el tiempo. Los cambios en la media determinan una tendencia a crecer o decrecer a largo plazo, por lo que la serie no oscila alrededor de un valor constante.

Figura 12. Serie temporal no estacionaria.

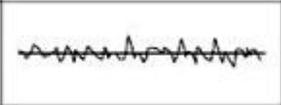
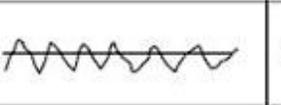
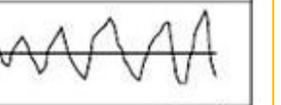
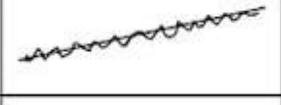
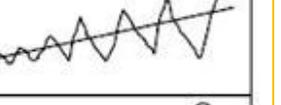
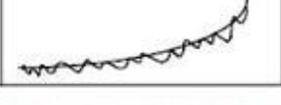
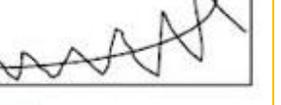


4.3.6. Estimación de la tendencia

Para estimar la tendencia supondremos que tenemos una serie no estacionaria sin componente estacional, es decir, que la serie se puede descomponer en $X_t = T_t + I_t$

Para estimar T_t debemos formular algunas hipótesis sobre su forma. Se analizan e ilustran a continuación algunos casos.

Tabla 4. Conducta de datos de la tendencia de una serie temporal según clasificación de Pegel⁹

	1 No efecto estacional	2 Estacionalidad Aditiva	3 Estacionalidad multiplicativa
A No efecto tendencia			
B Tendencia aditiva			
C Tendencia multiplicativa			

Tendencia determinista

En este caso supondremos que la tendencia es una función determinística. La función más sencilla posible es una recta, es decir, $T_t = a + b_t$ donde a y b son dos constantes a determinar. La forma de estimar estas constantes es mediante un modelo de regresión lineal entre las variables X_t y el tiempo $t = 1, 2, 3, \dots$. De esta forma, si estimamos los parámetros \hat{a} y \hat{b} , entonces la componente irregular será $I_t = X_t - \hat{a} - \hat{b}t$

Ahora I_t sería una serie estacionaria que tendríamos que modelizar.

En algunos casos, no es posible ajustar la tendencia mediante una recta. En estos casos, lo mejor sería ajustar la tendencia a un polinomio o a la curva que mejor se pueda ajustar. Para ello tendríamos que ajustar una regresión no lineal. Otra opción es describir la tendencia de manera evolutiva o diferenciar la serie.

⁹ <http://statisticalecology-st.blogspot.com/2010/12/identificacion-de-modelos-de-series.html>, consultado 13 de octubre de 2014.

Tendencia evolutiva (medias móviles)

Se supone que la tendencia es una función que evoluciona lentamente y que puede aproximarse en intervalos muy cortos (por ejemplo de 3 ó 5 datos) por una función simple del tiempo. En general se supone una recta, pero ahora sus coeficientes van cambiando suavemente en el tiempo.

Suponemos que la representación de la tendencia por una recta es válida para tres períodos consecutivos de tiempo, $t - 1$, t , $t + 1$, y representamos las tendencias en los tres periodos consecutivos de la siguiente manera:

$$\begin{aligned}T_{t-1} &= T_t - \text{crecimiento} \\ &T_t \\ T_{t+1} &= T_t + \text{crecimiento}\end{aligned}$$

4.3.7. Diferenciación de la serie

Un método más general para eliminar la tendencia consiste en suponer que la tendencia evoluciona lentamente en el tiempo, de manera que en el instante t la tendencia debe estar próxima a la tendencia en el instante $t - 1$. De esta forma, si restamos a cada valor de la serie el valor anterior, la serie resultante estará aproximadamente libre de tendencia. Esta operación se denomina diferenciación de la serie y consiste en pasar de la serie original x_t a la serie y_t mediante: $y_t = x_t - x_{t-1}$. De este modo, la serie diferenciada resulta ser estacionaria.

4.3.8. Estimación de la estacionalidad

Se elimina de la serie la componente estacional. Si la serie presenta una estacionalidad mensual, se colocan los datos en una tabla de doble entrada. Si hay efecto estacional mensual, por tanto, existen 12 coeficientes estacionales, uno para cada mes del año. Para estimarlos, se calcula primero la media de las observaciones para cada mes, M_1, \dots, M_{12} , y el coeficiente estacional resulta ser: $S_i = M_i - M$ para $i = 1, \dots, 12$. donde M es la media total de las observaciones. Así se puede observar que los meses con observaciones más pequeñas (por debajo de la media

general) tienen coeficientes estacionales negativos, mientras que los meses con observaciones mayores tienen coeficientes estacionales positivos. La suma de los coeficientes estacionales tiene que ser cero.

Se denomina serie desestacionalizada a una serie donde se ha eliminado el efecto de cada mes y que se obtiene restando al valor de cada mes el coeficiente estacional de dicho mes.

4.3.9. Modelos Autorregresivos (Modelos AR)

Un modelo autorregresivo de orden p , o abreviadamente AR (p), se define como:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \delta + u_t \quad (1)$$

En términos teóricos, no hay nada que nos limite el número de períodos del pasado que influyen en el valor actual del proceso, sin embargo tal planteamiento provoca problemas no justificables en términos prácticos; es decir, no es necesario tener que estimar infinitos coeficientes de regresión cuando en la práctica se observa que solo los períodos más recientes tienen influencia significativa en el valor actual del proceso.

Por tanto, la ecuación (1) puede escribirse de manera abreviada como: $\phi(L)y_t = \delta + u_t$ donde $\phi(L)$ es un operador polinomial de retardos, siendo $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$.

Estos modelos, deben cumplir la condición de estacionariedad.

Una condición necesaria, aunque no suficiente, para que un modelo autorregresivo de cualquier orden sea estacionario es la siguiente

$$\phi_1 + \phi_2 + \dots + \phi_p < 1$$

Al cumplir el supuesto de estacionariedad, la esperanza matemática viene dada por

$$E(y_t) = \mu = \frac{\delta}{1 - \phi_1 - \phi_2 - \dots - \phi_p}$$

Su varianza

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \dots + \phi_p\gamma_p + \sigma_u^2$$

Función de autocovarianzas

$$\gamma_j = \phi_1\gamma_{j-1} + \phi_2\gamma_{j-2} + \dots + \phi_p\gamma_{j-p} \quad \text{para } j > 0$$

Considerando γ_0 y γ_j para $j = 1, 2, 3, \dots, p$ se obtiene el sistema de $p + 1$ ecuaciones:

$$\begin{aligned} \gamma_0 &= \phi_1\gamma_1 + \phi_2\gamma_2 + \dots + \phi_p\gamma_p + \sigma_u^2 \\ \gamma_1 &= \phi_1\gamma_0 + \phi_2\gamma_1 + \dots + \phi_p\gamma_{p-1} \\ \gamma_2 &= \phi_1\gamma_1 + \phi_2\gamma_0 + \dots + \phi_p\gamma_{p-2} \\ &\dots \\ \gamma_p &= \phi_1\gamma_{p-1} + \phi_2\gamma_{p-2} + \dots + \phi_p\gamma_0 \end{aligned}$$

Si se conocen los valores de los parámetros $\phi_1, \phi_2, \dots, \phi_p$ y σ_u^2 , pueden calcularse la varianza y autocovarianzas de orden no superior al del proceso. A partir de dicho orden, es decir, para $j > p$, las autocovarianzas se obtendrán de forma recursiva según las ecuaciones siguientes:

$$\begin{aligned} \gamma_{p+1} &= \phi_1\gamma_p + \phi_2\gamma_{p-1} + \dots + \phi_p\gamma_1 \\ \gamma_{p+2} &= \phi_1\gamma_{p+1} + \phi_2\gamma_p + \dots + \phi_p\gamma_2 \\ \gamma_{p+3} &= \phi_1\gamma_{p+2} + \phi_2\gamma_{p+1} + \dots + \phi_p\gamma_3 \end{aligned}$$

Función de autocorrelación. Dividiendo por la función de autocovarianzas, γ_0 , se tiene

$$\rho_j = \phi_1\rho_{j-1} + \phi_2\rho_{j-2} + \dots + \phi_p\rho_{j-p} \quad \text{para } j > 0$$

Las ecuaciones de Yule-Walker, que se obtienen directamente de la función de autocorrelación, serán

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2\rho_1 + \dots + \phi_p\rho_{p-1} \\ \rho_2 &= \phi_1\rho_1 + \phi_2 + \dots + \phi_p\rho_{p-2} \\ &\dots \end{aligned}$$

$$\rho_p = \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p$$

Resolviendo dicho sistema para $\phi_1, \phi_2, \dots, \phi_p$ y expresado en forma de matrices se tiene

$$\begin{pmatrix} \phi_1 \\ \phi_2 \\ \dots \\ \phi_p \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \dots & \dots & \dots & \dots \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \dots \\ \rho_p \end{pmatrix}$$

Función de Autocorrelaciones Parciales

$$\phi_{jj} = \begin{cases} \rho_1 & \text{para } j = 1 \\ \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \frac{\phi_2[(1 - \phi_2)^2 - \phi_1^2]}{(1 - \phi_2)^2 - \phi_1^2} & \text{para } j = 2 \\ \left| \begin{array}{ccccc} 1 & \rho_1 & \dots & \rho_{p-2} & \rho_1 \\ \rho_1 & 1 & \dots & \rho_{p-3} & \rho_2 \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{p-1} & \rho_{p-2} & \dots & \rho_1 & \rho_p \end{array} \right| & \text{para } j = p \\ \left| \begin{array}{ccccc} 1 & \rho_1 & \dots & \rho_{p-2} & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-3} & \rho_{p-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{p-1} & \rho_{p-2} & \dots & \rho_1 & 1 \end{array} \right| & \text{para } j > p \\ 0 & \end{cases}$$

Propiedades de los modelos AR (p):

1. Un modelo autorregresivo siempre es invertible.
2. Un modelo AR estacionario debe satisfacer la condición de que las raíces de la ecuación $\phi(L) = 0$ caigan fuera del círculo unidad.
3. El correlograma típico de los modelos AR presenta siempre una tendencia amortiguada hacia cero, sin llegar a anularse. En el caso de que las raíces del polinomio característico sean complejas, la forma del correlograma es sinusoidal, pero también amortiguada hacia cero.
4. Las autocorrelaciones parciales se anulan para retardos superiores al orden del modelo.

4.3.10. Modelos de Medias Móviles (Modelos MA)

Un modelo de medias móviles MA describe una serie temporal estacionaria. En este modelo el valor actual puede predecirse a partir de la componente aleatoria de este momento y, en menor medida, de los impulsos aleatorios anteriores.

El proceso de medias móviles de orden q , representado por MA (q), se define como:

$$y_t = \mu + u_t - \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q} \quad (2)$$

Donde:

y_t : Variable en estudio, la cual ha sido observada durante cierto periodo de tiempo

μ : Esperanza matemática del modelo

θ_i : Parámetros o coeficientes desconocidos del modelo de medias móviles de orden q

u_t : Variables aleatorias distribuidas idéntica e independientemente con media cero y varianza σ^2 .

La expresión dada en la ecuación (2) puede escribirse de manera simplificada como

$$y_t = \mu + \theta(L)u_t,$$

donde $\theta(L)$ es un operador polinomial de retardos, definido como:

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q,$$

siendo L el operador de retardo, tal que el valor de su exponente indica el orden del retardo de la variable a la que multiplica; es decir: $L^s y_t = y_{t-s}, \forall s \geq 0$.

Este modelo siempre será estacionario, ya que es un caso particular del proceso estocástico discreto general, definido como:

$$y_t = \mu + u_t + \Psi_1 u_{t-1} + \Psi_2 u_{t-2} + \dots$$

Donde μ y los Ψ_i 's son parámetros desconocidos, y $u_t, u_{t-1}, u_{t-2}, \dots$ es una secuencia de perturbaciones aleatorias distribuidas idéntica e independientemente con media cero y varianza σ_u^2 .

Con las restricciones: $\Psi_i = 0, \forall i \geq q$, $\Psi_i = -\theta_i$ para $1 \leq i \leq q$ y los $\Psi_0 = 1$, debe cumplirse entonces: $\sum_{i=0}^{\infty} \Psi_i = 1 - \theta_1 - \theta_2 - \dots - \theta_q = 1 - \sum_{i=0}^{\infty} \theta_i = k$, donde k es una constante, q el número de parámetros de medias móviles.

$$\sum_{i=0}^{\infty} \Psi_i^2 = 1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2 < \infty$$

$$\sum_{i=0}^{\infty} \Psi_i \Psi_{i+j} = \Psi_0 \Psi_j + \Psi_1 \Psi_{j+1} + \Psi_2 \Psi_{j+2} + \dots + \Psi_{q-j} \Psi_q = -\theta_j - \theta_1 \theta_{j+1} + \theta_2 \theta_{j+2} + \dots + \theta_{q-j} \theta_q < \infty$$

Además, un modelo de medias móviles es invertible cuando se puede escribir como un proceso autorregresivo de orden infinito, para lo cual debe cumplir que las raíces de

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q,$$

caigan fuera del círculo unitario, es decir si L_1, L_2, \dots, L_q son las q raíces, estas deben ser mayores que uno en valor absoluto, cuando las raíces son reales; si las raíces son complejas el módulo debe ser mayor a la unidad.

La esperanza matemática del modelo de medias móviles, MA (q) viene dada por:

$$E(y_t) = \mu$$

La varianza del modelo MA (q)

$$\gamma_0 = \sigma_u^2 (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)$$

La función de autocovarianzas

$$\gamma_j = \begin{cases} (-\theta_j + \theta_1 \theta_{j+1} + \dots + \theta_{q-j} \theta_q) \sigma_u^2 & \text{para } j = 1, 2, 3, \dots, q \\ 0 & \text{para } j > q \end{cases}$$

Y la función de autocorrelación

$$\rho_j = \begin{cases} \frac{-\theta_j + \theta_1\theta_{j+1} + \dots + \theta_{q-j}\theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & \text{para } j = 1, 2, 3, \dots, q \\ 0 & \text{para } j > q \end{cases}$$

Finalmente, La función de autocorrelación parcial

$$\phi_{11} = \rho_1$$

$$\phi_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

$$\phi_{33} = \frac{\rho_1^3 - \rho_1\rho_2(2 - \rho_2) + \rho_3(1 - \rho_1^2)}{1 - \rho_1^2 - 2\rho_1^2(1 - \rho_2)}$$

...

$$\phi_{jj} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{j-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{j-3} & \rho_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_{j-1} & \rho_{j-2} & \rho_{j-3} & \dots & \rho_1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{j-2} & \rho_{j-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{j-3} & \rho_{j-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_{j-1} & \rho_{j-2} & \rho_{j-3} & \dots & \rho_1 & 1 \end{vmatrix}}$$

Propiedades del modelo MA (q):

1. Un modelo de medias móviles siempre es estacionario.
2. Un proceso de medias móviles siempre es estacionario, esto es, puede escribirse como un modelo AR (∞) siempre y cuando las raíces de la ecuación $\theta(L) = 0$ caigan fuera del círculo unidad.
3. Las autocovarianzas, y consecuentemente las autocorrelaciones (y el correlograma), se anulan cuando el desfase temporal considerado es superior al orden del modelo. Consecuentemente, diremos que la memoria de un proceso de medias móviles es igual al orden del proceso.
4. Las autocorrelaciones parciales no se anulan, presentando un comportamiento amortiguado hacia cero (caso de raíces reales) o sinusoidal también tendente a cero (caso de raíces complejas).

Al observar las características de los modelos AR (p) y MA (q) se aprecia cierta simetría que se denomina dualidad, la cual es consecuencia del hecho de que un mismo proceso podemos representarlo como autorregresivo o como de medias móviles según nuestra conveniencia, dado que un AR(p) puede representarse como un MA (∞) y un MA (q) como un AR (∞). La ventaja de usar en cada caso la representación adecuada reside en tener un modelo con pocos parámetros.

En resumen, las correspondencias entre ambos modelos se muestran en la tabla siguiente:

Tabla 5. Correspondencia entre un modelo AR (p) y un modelo MA (q)

Concepto	AR(p)	MA(q)
Expresión	$\phi(L)y_t = \delta + u_t$	$y_t = \mu + \theta(L)u_t$
Condición de estacionariedad	Raíces de $\phi(L)$ en módulo mayores que 1	Siempre es estacionario
Condición de invertibilidad	Siempre es invertible	Raíces de $\theta(L)$ en módulo mayores que 1
Correlograma simple	Infinitos valores no nulos decreciendo según mezcla de exponenciales y sinusoidales amortiguadas	Coefficientes no nulos hasta el retardo q, el resto nulos
Correlograma parcial	Coefficientes no nulos hasta el retardo p, el resto nulos	Infinitos valores no nulos decreciendo según mezcla de exponenciales y sinusoidales amortiguadas
Modelo Dual	MA(∞)	AR(∞)

4.3.11. Modelos ARMA

Una extensión natural de los modelos AR (p) y MA (q) es un tipo de modelos que incluyen tanto términos autorregresivos como de medias móviles. Se representa por la ecuación:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \delta + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q}$$

Este modelo puede escribirse utilizando los operadores polinomiales de retardos como:

$$\phi(L)y_t = \delta + \theta(L)u_t$$

Dado que el modelo ARMA contiene una parte AR y una parte MA, deberá cumplir las condiciones de estacionariedad como de invertibilidad. En concreto, diremos que es estacionario cuando lo es su parte AR, y diremos que es invertible cuando lo es su parte MA.

Diremos que dicho modelo es estacionario e invertible, si tanto las raíces de

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p = 0$$

como las de $\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q = 0$ caen fuera del círculo unidad.

Una condición necesaria, aunque no suficiente, para que un modelo mixto autorregresivo medias móviles de cualquier orden sea estacionario e invertible es la siguiente:

$$\theta_1 + \theta_2 + \dots + \theta_q < 1$$

$$\phi_1 + \phi_2 + \dots + \phi_p < 1$$

Adicionalmente a la estacionariedad e invertibilidad supondremos que las p raíces de $\phi(L) = 0$ y las q raíces de $\theta(L) = 0$ no se cancelan, esto es, no hay raíces comunes. Bajo estas condiciones, la esperanza matemática y la varianza del modelo ARMA (p,q) se definen así:

Esperanza matemática

$$\mu = \frac{\delta}{1 - \phi_1 - \phi_2 - \dots - \phi_p}$$

Varianza

$$\gamma_0 = \frac{\sigma_u^2(1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2 - 2\phi_1\theta_1 - 2\phi_2\theta_2 - \dots - 2\phi_h\theta_h)}{1 - \phi_1^2 - \phi_2^2 - \dots - \phi_p^2}$$

Función de autocovarianzas

$$\gamma_j = \begin{cases} \phi_1\gamma_0 + \phi_2\gamma_1 + \dots + \phi_p\gamma_{p-1} - \theta_1\sigma_u^2 - \theta_2E(\tilde{y}_{t-1}u_{t-2}) - \dots - \theta_qE(\tilde{y}_{t-1}u_{t-q}) & \text{para } j = 1 \\ \phi_1\gamma_1 + \phi_2\gamma_0 + \dots + \phi_p\gamma_{p-2} - \theta_2\sigma_u^2 - \theta_3E(\tilde{y}_{t-1}u_{t-3}) - \dots - \theta_qE(\tilde{y}_{t-2}u_{t-q}) & \text{para } j = 2 \\ \dots & \dots \\ \phi_1\gamma_{q-1} + \phi_2\gamma_{q-2} + \dots + \phi_p\gamma_{p-q} - \theta_q\sigma_u^2 & \text{para } j = q \\ \phi_1\gamma_{j-1} + \phi_2\gamma_{j-2} + \dots + \phi_p\gamma_{p-q} & \text{para } j > q \end{cases}$$

Función de Autocorrelaciones

$$\rho_j = \begin{cases} \phi_1 + \phi_2\rho_1 + \dots + \phi_p\rho_{p-1} + \frac{1}{\gamma_0}[-\theta_1\sigma_u^2 - \theta_2E(\tilde{y}_{t-2}u_{t-2}) - \dots - \theta_qE(\tilde{y}_{t-1}u_{t-q})] & \text{para } j = 1 \\ \phi_1\rho_1 + \phi_2 + \dots + \phi_p\rho_{p-2} + \frac{1}{\gamma_0}[-\theta_2\sigma_u^2 - \theta_3E(\tilde{y}_{t-2}u_{t-3}) - \dots - \theta_qE(\tilde{y}_{t-2}u_{t-q})] & \text{para } j = 2 \\ \dots & \dots \\ \phi_1\rho_{q-1} + \phi_2\rho_{p-2} + \dots + \phi_p\rho_{p-q} + \frac{1}{\gamma_0}\theta_q\sigma_u^2 & \text{para } j = q \\ \phi_1\rho_{j-1} + \phi_2\rho_{j-2} + \dots + \phi_p\rho_{j-q} & \text{para } j > q \end{cases}$$

Propiedades del modelo ARMA (p,q)

1. El modelo ARMA (p,q) será estacionario e invertible siempre y cuando las raíces de $\phi(L) = 0$ y $\theta(L) = 0$ caigan fuera del círculo unidad. Adicionalmente supondremos que no hay raíces comunes.
2. La función de autocorrelación no se anula, si bien tiende a cero a medida que aumentan los desfases temporales considerados. Consecuentemente la memoria de este modelo es infinita. La presencia del término de medias móviles en el modelo ARMA (p,q) afecta la determinación de las q primeras autocorrelaciones $(\rho_1, \rho_2, \dots, \rho_q)$ estando el resto de las

autocorrelaciones afectadas únicamente por la parte autorregresivo. Concretamente, el análisis de la función de autocorrelación muestra que una vez superado el orden de la parte de medias móviles, las autocorrelaciones del modelo ARMA (p,q) se comportan como las de un modelo AR(p).

3. El correlograma, representación gráfica de la función de autocorrelación, presentara q picos, hasta el retardo q, en los que influye tanto la parte autorregresiva como la de medias móviles del modelo y, a continuación su forma coincide con la de los modelos AR (p), esto es, comportamiento amortiguado hacia cero sin llegar a anularse, si las raíces son reales, o comportamiento sinusoidal, pero también amortiguado hacia cero, si las raíces son complejas.
4. La función de autocorrelación parcial no se anula, como consecuencia de que el modelo ARMA (p,q) contiene el modelo MA(q) como un caso especial. La representación gráfica de la función de autocorrelación parcial tendrá también un comportamiento amortiguado hacia cero.

4.3.12. Modelos estocásticos lineales no estacionarios homogéneos. Modelos ARIMA

Muy pocas de las series temporales económicas son estacionarias. No parece, por ello, muy conveniente suponer que aquellas series que no son estacionarias vengan generadas por cualquiera de los procesos anteriores, en los cuales se han impuesto las condiciones de estacionariedad. Afortunadamente, sin embargo, muchas de estas series temporales no estacionarias tienen la propiedad de convertirse en estacionarias al diferenciar una o más veces la serie original.

Resulta conveniente, por ello, considerar junto a los procesos estocásticos analizados, aquellos que aun no cumpliendo originalmente las condiciones de estacionariedad, son susceptibles de transformarse en estacionarios efectuando alguna diferenciación. A dichos procesos estocásticos se les denomina como homogéneos.

En concreto, diremos que un proceso estocástico no estacionario es homogéneo cuando al diferenciar en el proceso original, el proceso transformado resultante es estacionario. El número de veces que debe diferenciarse el proceso original para transformarse en estacionario constituye el grado u orden de homogeneidad.

Dada una serie y_t y supóngase que no cumple con los requisitos de estacionariedad en media y en varianza. Para obtener un modelo ARIMA será necesario aplicar ciertas transformaciones para convertir dicha serie temporal en estacionaria.

Estas transformaciones se concretan en diferenciaciones sucesivas en la serie (la determinación del parámetro d).

Si diferenciamos una vez y_t , se obtiene un modelo transformado de la serie $w_t = u_t$

$$\text{Dónde: } w_t = \Delta y_t = y_t - y_{t-1} = (1 - L)y_t$$

Este proceso transformado es un ruido blanco, el cual siempre es estacionario.

Al diferenciar la serie temporal se pasa de y_t al proceso w_t . Ahora considérese el caso inverso de obtener y_t a partir del proceso w_t . Por sustituciones sucesivas se obtiene que

$$\begin{aligned} y_t &= w_t + y_{t-1} \\ &= w_t + w_{t-1} + y_{t-2} \\ &= w_t + w_{t-1} + w_{t-2} + y_{t-3} \\ &= w_t + w_{t-1} + w_{t-2} + w_{t-3} + \dots \end{aligned}$$

Luego se obtiene:

$$y_t = w_t + w_{t-1} + w_{t-2} + w_{t-3} + \dots$$

Como puede notarse, que el proceso y_t se obtiene sumando, o lo que es lo mismo “integrando” el proceso w_t .

A este proceso integrado, y_t se le denomina proceso autorregresivo-medias móviles integrado o más abreviadamente ARIMA (p,d,q), si tomando diferencias de orden d se obtiene un proceso estacionario w_t del tipo ARMA (p,q).

La forma general del modelo ARIMA (p,d,q) es

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q}$$

Donde: $w_{t-1} = \Delta^d y_{t-j} = (1 - L)^d y_{t-j}$ para $j = 0, 1, 2, 3, \dots, p$

Abreviadamente, utilizando los operadores de polinomios de retardos definidos, en los apartados anteriores podemos escribir:

$$\phi(L)w_t = \theta(L)u_t$$

Donde:

$$w_t = \Delta^d y_t = (1 - L)^d y_t$$

El modelo ARIMA (p, d, q) representa un modelo estocástico general. Así, si $p = d = 0$ estamos ante el modelo ARMA (0, 0, q) o bien MA (q); si $q = d = 0$ se tiene el modelo ARMA (p, 0, 0) o bien un modelo AR (p) abordados anteriormente. Por otro lado, si $d = 0$ se tiene el modelo ARMA (p, 0, q) o bien ARMA (p, q).

Otros casos posibles son:

- a) $q = 0 \rightarrow$ ARIMA (p, d, 0) o ARI (p, d)
- b) $p = 0 \rightarrow$ ARIMA (0, d, q) o IMA (d, q)
- c) $p = q = 0; d = 1 \rightarrow$ ARIMA (0, 1, 0) o “paseo aleatorio”
- d) $p = d = q = 0$ (y $\delta = 0$) \rightarrow ARIMA (0, 0, 0) o “ruido blanco”

Al analizar la evolución de la mayoría de las series temporales presentan tendencia, creciente o decreciente. La eliminación de esta tendencia (no estacionaria en media) de la serie se puede conseguir mediante las diferenciaciones implícitas en los modelos ARIMA. También puede observarse tendencia en la varianza; es decir, la dispersión de las observaciones respecto a la media no es constante a lo largo del tiempo. Este tipo de tendencia no se elimina mediante el proceso de diferenciación, siendo lo adecuado transformar la serie aplicando logaritmos naturales.

Esta posibilidad de transformar la serie puede concretarse de la forma más general mediante la transformada de Box-Cox. De esta forma, el modelo ARIMA más general puede enunciarse como:

$$\phi(L)\Delta^d y_t^{(\lambda)} = \delta + \theta(L)u_t$$

O bien

$$\phi(L)(1-L)^d(y_t^{(\lambda)} - \mu) = \theta(L)u_t$$

Donde μ es la media de $y_t^{(\lambda)}$, siendo:

$$y_t^{(\lambda)} = \begin{cases} \frac{y_t^{(\lambda)} - 1}{\lambda} & \text{para } \lambda \neq 0 \\ \ln y_t & \text{para } \lambda = 0 \end{cases}$$

Restringiéndonos a los casos $\lambda = 1$ y $\lambda = 0$, tendremos respectivamente la serie sin transformar (la serie original) o la serie transformada aplicando logaritmos neperianos, con esto se logra una mejor interpretación que con otros valores de la transformación Box- Cox.

4.3.13. Modelos estacionales no estacionarios homogéneos. Modelos ARIMA (P,D,Q)_s

Al igual que sucede con los modelos no estacionales, dado que la mayoría de las series económicas son no estacionarias, resulta de utilidad considerar los modelos estacionales no estacionarios, pero homogéneos. Concretamente, el modelo ARIMA (P,D,Q)_s será:

$$w_t = \Phi_1 w_{t-s} + \Phi_2 w_{t-2s} + \dots + \Phi_P w_{t-Ps} + \delta + u_t - \Theta_1 u_{t-s} - \dots - \Theta_Q u_{t-Qs}$$

donde: $w_t = \Delta_s^D y_t = (1 - L^s)^D y_t$,

el cual se puede representar como:

$$\Phi_P(L^s)w_t = \delta + \Theta_Q(L^s)u_t; \Theta_Q(L^s) = 1 - \Theta_1 L^s - \Theta_2 L^{2s} - \dots - \Theta_Q L^{Qs},$$

es el operador de retardo para el modelo estacional de medias móviles estacional MA(Q)_s y $\Phi_p(L^s) = 1 - \Phi_1 L^s - \Phi_2 L^{2s} - \dots - \Phi_p L^{ps}$ es el operador de retardo para el modelo autorregresivo estacional AR(P)_s.

Si escribimos el modelo ARIMA (P,D,Q)_s como:

$$(1 - \Phi_1 L^s - \Phi_2 L^{2s} - \dots - \Phi_p L^{ps})(1 - L^s)^D y_t = (1 - \Theta_1 L^s - \Theta_2 L^{2s} - \dots - \Theta_q L^{qs}) \varepsilon_t,$$

entonces asumamos que ε_t viene generado por un proceso ARIMA (p,d,q), es decir:

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)(1 - L)^d \varepsilon_t = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) u_t,$$

Siendo u_t un ruido blanco. Si sustituimos la segunda en la primera de las ecuaciones, obtendremos el modelo estacional general que denotaremos abreviadamente como ARIMA (p,d,q) x (ARIMA (P,D,Q)_s y que escribiremos como:

$$\begin{aligned} & (1 - \Phi_1 L^s - \Phi_2 L^{2s} - \dots - \Phi_p L^{ps})(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)(1 - L^s)^D (1 - L)^d y_t \\ & = (1 - \Theta_1 L^s - \Theta_2 L^{2s} - \dots - \Theta_q L^{qs})(1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) u_t \end{aligned}$$

O, abreviadamente:

$$\Phi_p(L^s) \phi_p(L) (1 - L^s)^D (1 - L)^d y_t = \Theta_q(L^s) \theta_q(L) u_t$$

En su forma más general, esta expresión puede escribirse como:

$$\Phi_p(L^s) \phi_p(L) [(1 - L^s)^D (1 - L)^d y_t - \mu] = \Theta_q(L^s) \theta_q(L) u_t$$

Siendo μ la media de $w_t = (1 - L^s)^D (1 - L)^d y_t$.

4.4. Análisis de series temporales con outliers e intervenciones

4.4.1. Análisis de Intervención

Cuando se analizan los datos de series temporales, es normal encontrar datos outliers (atípicos) que tienen su origen en intervenciones incontroladas o inesperadas. Las series de tiempo se ven influenciadas por ciertos sucesos externos, tales como: problemas sociales (vacaciones, huelgas, cambios de política), fenómenos naturales (huracanes, tormentas tropicales, marea roja), y otros. A los sucesos anteriores Box y Tiao les denominan: Intervención.

El análisis de intervención tiene su propia técnica, esta consiste en evaluar el efecto de las intervenciones en el proceso de comportamiento de una serie temporal.

Antes de aplicar esta técnica se debe identificar dos características de los modelos de intervención:

- a) El periodo de comienzo de dichos sucesos externos
- b) La forma general del impacto de los sucesos

Para analizar detenidamente las intervenciones, se detalla los modelos en los que se puntualizan las variables de intervención que se pueden dar en una serie de tiempo.

4.4.2. Modelos de Intervención

Estos modelos son utilizados cuando un evento ocurre y hace que los valores de la ST en estudio se vean aumentados o disminuidos a partir de la ocurrencia del evento. La disminución de la ST puede ser continua o de un cambio brusco en el instante de ocurrencia del evento.

Para determinar si el cambio de la serie es continuo o temporal se hace a través de dos tipos de variables de intervención: Variable escalón y variable impulso.

Variable escalón: Tiene efectos a partir del tiempo t_0 en el que ocurre, se emplea solamente cuando los eventos externos tienen una duración permanente.

Se puede representar de la siguiente manera:

$$S_t^{(t_0)} = \begin{cases} 0, & t < t_0 \\ 1, & t \geq t_0 \end{cases}$$

El otro tipo de variable de intervención es la **variable impulso**, que tienen efectos solo en el tiempo t_0 de ocurrencia y se emplea cuando los sucesos externos tienen una duración temporal.

Este tipo de variable se representa de la siguiente forma:

$$P_t^{(t_0)} = \begin{cases} 1, & t = t_0 \\ 0, & t \neq t_0 \end{cases}$$

Nótese, que las variables escalón e impulso se relacionan mediante la siguiente expresión:

$$P_t^{(t_0)} = (1 - L) S_t^{(t_0)}, \text{ donde } L \text{ es el operador de retardo.}$$

Entonces, un modelo de intervención puede representarse igualmente, bien con la variable escalón o con la variable impulso. El uso de una forma específica de estas variables es usualmente basado en la conveniencia de la interpretación en el modelo.

Algunos modelos pueden estructurarse, tomando en cuenta algunos criterios como los siguientes:

1. El modelo más simple sería aquel en que el impacto de una intervención es fijo y de magnitud desconocida. Así, dependiendo del tipo de intervención, y tomando en cuenta que el modelo se puede generalizar transcurridos b períodos, el impacto es:

$$X_t = \omega L^b S_t^{(t_0)} \quad \text{ó} \quad X_t = \omega L^b P_t^{(t_0)}$$

2. Si se tiene un impacto de una intervención con comienzo gradual, para los dos tipos de duración del efecto, con b períodos transcurridos después de la intervención se tiene para la variable escalón el modelo siguiente:

$$X_t = \frac{\omega L^b}{1 - \delta L} S_t^{(t_0)}$$

Y para la variable impulso $b=1$, se tiene:

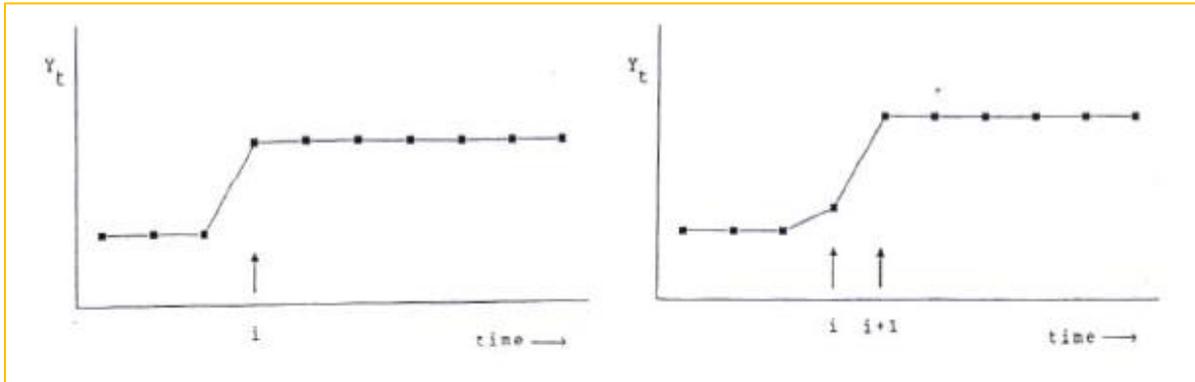
$$X_t = \frac{\omega L^b}{1 - \delta L} P_t^{(t_0)}; \text{ donde } 0 \leq \delta \leq 1.$$

En la tabla 6 y 7, se muestran las gráficas para ambas variables de intervención con $b=1$ y $0 < \delta < 1$.

Tabla 6. Diferentes formas del Comportamiento de la variable escalón en una serie temporal continua y sus respectivas ecuaciones.

a) $X_t = \omega S_t^{t_0}$

b) $X_t = \omega L S_t^{t_0}$



c) $X_t = \frac{\omega}{1-\delta_1 L} S_t^{t_0}, \quad |\delta_1| < 1$

d) $X_t = \frac{\omega}{1-\delta_1 L} S_t^{t_0}, \quad |\delta_1| = 1$

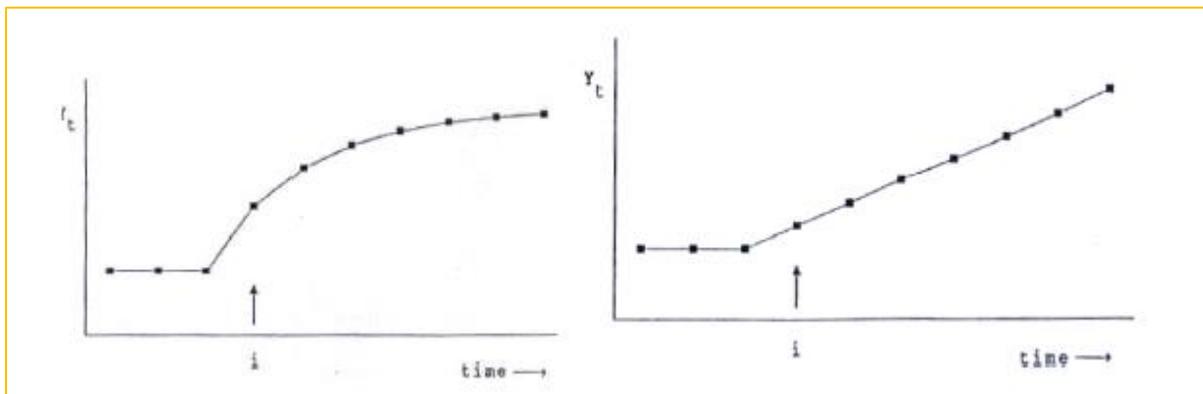
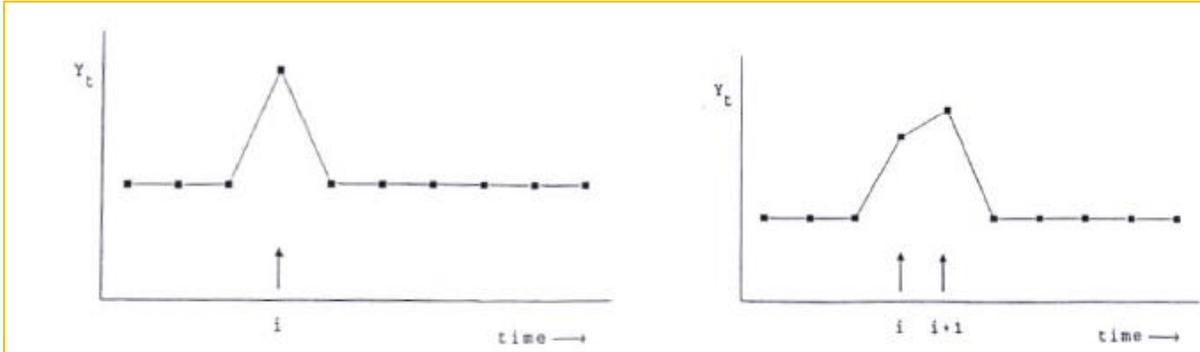


Tabla 7. Diferentes formas del Comportamiento de la variable impulso en una serie temporal continua y sus respectivas ecuaciones.

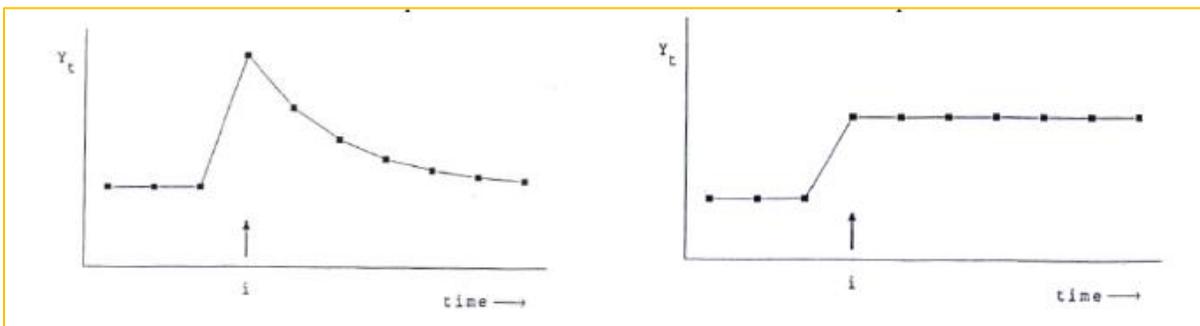
a) $X_t = \omega P_t^{t_0}$

b) $X_t = \omega L P_t^{t_0}$



c) $X_t = \frac{\omega}{1-\delta_1 L} P_t^{t_0}, \quad |\delta_1| < 1$

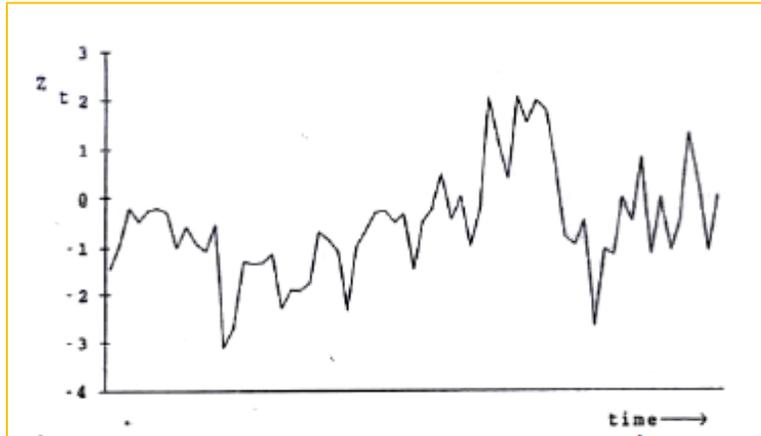
d) $X_t = \frac{\omega}{1-\delta_1 L} P_t^{t_0}, \quad |\delta_1| = 1$



Es de hacer notar que en ocasiones suelen combinarse la variable escalón e impulso, en cuyo caso puede obtenerse el siguiente modelo de intervención, y resultará de utilidad cuando se espera que el efecto externo deje un efecto residual permanente después de una declinación gradual del impacto.

$$X_t = \left(\frac{\omega_1 L}{1 - \delta L} + \frac{\omega_2 L}{1 - L} \right) P_t^{t_0}$$

Figura 13. Comportamiento de la variable impulso y escalón en una serie temporal



Sin embargo, teniendo en cuenta la relación: $P_t^{t_0} = (1 - L)S_t^{t_0}$, la última expresión, puede expresarse de la forma siguiente:

$$X_t = \left(\frac{\omega_1 L}{1 - \delta L} P_t^{t_0} + \omega_2 L S_t^{t_0} \right)$$

Finalmente, se puede señalar que los modelos de intervención pueden especificarse mediante la siguiente expresión general: $X_t = v(L)I_t^{t_0}$

Siendo $I_t^{t_0}$ una variable impulso o escalón $v(L) = \frac{\omega(L)L^b}{\delta(L)}$

donde:

$$\begin{aligned} \omega(L) &= \omega_0 - \omega_1 L - \omega_2 L^2 - \dots - \omega_s L^s \\ \delta(L) &= 1 - \delta_1 L - \delta_2 L^2 - \dots - \delta_r L^r \end{aligned}$$

En la tabla 8, se recogen los valores que deberán tener b, r y s (subíndice de la expansión de cada uno de los parámetros) en la expresión anterior, para poder representar en $X_t = v(L)I_t^{t_0}$ cada uno de los modelos expuestos.

Tabla 8. Valores de b, r y s de modelos de intervención.

	$I_t^{t_0} = S_t^{t_0}$	$I_t^{t_0} = P_t^{t_0}$
$r = 0$ $b = 0$	$X_t = \omega S_t^{t_0}$	Modelo no significativo
$r = 0$ $b = 1$	$X_t = \omega L S_t^{t_0}$	$X_t = \omega L P_t^{t_0}$
$r = 1$ $b = 0$	$X_t = \frac{\omega}{1 - \delta L} S_t^{t_0}$	$X_t = \frac{\omega}{1 - \delta L} P_t^{t_0}$
$r = 1$ $b = 1$	$X_t = \frac{\omega L}{1 - \delta L} S_t^{t_0}$	$X_t = \frac{\omega L}{1 - \delta L} P_t^{t_0}$
$r = 2$ $b = 0$	Modelo no significativo	$X_t = \frac{\omega}{1 - \delta_1 L - \delta_2 L^2} P_t^{t_0}$

Obsérvese que en todos ellos el valor de s es cero porque $\omega(L) = \omega_0 = \omega$.

Entonces, el modelo del análisis para k intervenciones, viene dada por:

$$X_t = \sum_{j=1}^k v_j(L) I_{jt}^{t_0j},$$

donde:

$$v_j(L) = \frac{\omega_j(L) L^{bj}}{\delta_j(L)}$$

4.4.3. Identificación de los modelos ARIMA con intervención

Dada una serie temporal $\{y_t\}$, en la que algún suceso externo produce modificaciones sustanciales en el comportamiento de la misma, será preciso tener en cuenta en su modelización la inclusión de alguna intervención (o varias). Veamos ahora como identificar un modelo ARIMA con intervención o intervenciones, para ello consideramos el modelo ARIMA univariante: $\phi(L)(1 - L)^d y_t = \theta(L)u_t$

Dónde: $\phi(L) = 1 - \phi_1 L - \dots - \phi_1 L^p$ y $\theta(L) = 1 - \theta_1 L - \dots - \theta_1 L^p$

De manera que las raíces de ambos polinomios, $\phi(L)$ y $\theta(L)$, deben caer fuera del círculo unitario para cumplir la invertibilidad; otra condición que deben de cumplir estos polinomios es la estacionariedad, teniendo en cuenta que $(1 - L)^d$, es el operador de diferencias consecutivas

donde el exponente d indica el número de diferencias que se realizaron para suavizar la serie y u_t un ruido blanco.

Considerando que la media de $\{y_t\}$ es igual a cero y si el modelo fuese estacional multiplicativo ARIMA (p,d,q) x ARIMA (P,D,Q)s, la forma general viene dada por:

$$y_t = \frac{\Theta(L^s)\theta(L)}{\Phi(L^s)\phi(L)(1-L^s)^D(1-L)^d} u_t$$

Si quisiéramos incorporar ahora el modelo de intervención, la ecuación debería modificarse así:

$$y_t = v(L)I_t^{t_0} + \frac{\Theta(L^s)\theta(L)}{\Phi(L^s)\phi(L)(1-L^s)^D(1-L)^d} u_t$$

La ecuación puede generalizarse al modelo con k intervenciones, en cuyo caso escribiremos:

$$y_t = \sum_{j=1}^K v_j(L)I_{jt}^{t_0j} + N_t$$

Dónde:

$$N_t = \frac{\Theta(L^s)\theta(L)}{\Phi(L^s)\phi(L)(1-L^s)^D(1-L)^d} u_t$$

Ahora, la tarea consiste en identificar el modelo recién escrito. Caben dos métodos posibles de identificación que se exponen a continuación:

- 1) Si el subconjunto de datos anteriores o posteriores a las intervenciones es suficientemente grande, podríamos identificar el modelo ARIMA, tomando como base de datos dicho subconjunto, incorporando posteriormente la modelización de las intervenciones.

Para aclarar el método, supongamos que analizando la función de autocorrelación muestral (FACM) y la función de autocorrelación parcial muestral (FAPM) de ese subconjunto de datos identificamos que no hay patrón estacional y que $p = 1, d = 1$; esto es:

$$(1 - \phi_1 L)(1 - L)y_t^* = (1 - \theta_1 L)u_t$$

donde y_t^* es la serie $\{y_t\}$ truncada en las observaciones anteriores o posteriores a t_{0j} , o lo que es igual: $N_t = \frac{(1-\theta_1 L)}{(1-\phi_1 L)(1-L)} u_t$

A continuación, supongamos que solo hay una intervención y que esta tiene un comienzo gradual (a partir del periodo t_0) y una duración permanente. Es decir, que el modelo de intervención adecuado es

$$x_t = \frac{\omega}{1-\delta L} S_t^{t_0}.$$

Al combinar $N_t = \frac{(1-\theta_1 L)}{(1-\phi_1 L)(1-L)} u_t$ y $x_t = \frac{\omega}{1-\delta L} S_t^{t_0}$ de la manera siguiente:

$$y_t = \frac{\omega}{1-\delta L} S_t^{t_0} + \frac{(1-\theta_1 L)}{(1-\phi_1 L)(1-L)} u_t.$$

Una vez identificado el modelo anterior, se deberá estimar conjuntamente todos los parámetros de dicho modelo y analizar los residuos del mismo para establecer si este modelo supera la etapa de chequeo (cuando dichos residuos pueda aceptarse que se comportan como un ruido blanco), permitiéndonos en este caso con dicho modelo realizar predicciones.

Por otra parte, debe tenerse en cuenta que el modelo supuesto para la intervención $x_t = \frac{\omega}{1-\delta L} S_t^{t_0}$, será tentativo; esto es, podríamos probar con varias especificaciones alternativas, analizando posteriormente en la etapa de chequeo cuál de ellas resulta más aceptable.

- 2) Cuando no disponemos de suficientes datos antes o después de las intervenciones, o bien al inspeccionar la serie tenemos alguna intuición respecto a la forma en que las intervenciones han afectado al patrón de comportamiento de la serie, podemos especificar la forma de $v_j(L)$ en $y_t = \sum_{j=1}^k v_j(L) I_{jt}^{t_0^j} + N_t$ y estimar solo este modelo, obteniendo:

$$\hat{y}_t = \sum_{j=1}^k \hat{v}_j(L) I_{jt}^{t_0^j}$$

A continuación calcularíamos los residuos como: $\hat{N}_t = y_t - \sum_{j=1}^k \hat{v}_j(L)I_{jt}^{t_0j}$

Los cuales pueden interpretarse como los valores “limpios” de la serie original; es decir, los valores de la serie que esperabamos observar sin la intervención.

Estos residuos pueden ya utilizarse para identificar el proceso ARIMA univariante por medio de la ecuación:

$$N_t = \frac{\theta(L^s)\theta(L)}{\Phi(L^s)\phi(L)(1-L^s)^D(1-L)^d} u_t$$

utilizando para ello los instrumentos tradicionales; esto es, la FACM y FAPM.

Posteriormente, combinaremos el modelo de intervención y el modelo residual para obtener el modelo de intervención global $y_t = \sum_{j=1}^k v_j(L)I_{jt}^{t_0j} + N_t$, el cual debería ser estimado conjuntamente por el método de máxima verosimilitud. Una vez estimado, nuevamente se debe comprobar si este modelo supera la etapa de chequeo y, en caso afirmativo, se utiliza el modelo para obtener predicciones de la serie temporal y_t .

4.4.4. Análisis de outliers en series temporales

Cuando el tiempo y la causa de los factores externos creadores de observaciones atípicas del comportamiento de la ST son desconocidos, provocan repercusiones de interrupción y crean observaciones espurias que son inconsistentes por lo que tales observaciones hacen referencia al análisis de outliers de series temporales.

Justificada la relevancia del tratamiento de las observaciones (outliers) en el análisis de ST, se pueden describir los tipos de outliers, así como también los efectos de los mismos. La forma en que pueden detectarse y darle tratamiento a estos.

4.4.4.1. Tipos de outliers

Denominaremos por $\{y_t\}$ la serie observada, y por $\{z_t\}$ la serie libre de outliers. Suponiendo que $\{z_t\}$, sigue un proceso estacional multiplicativo ARIMA (p,d,q) x ARIMA(P,D,Q)_s, y considerando sin pérdida de generalidad que la media es cero, escribiremos:

$$\Phi(L^s)\phi(L)(1-L^s)^D(1-L)^d z_t = \theta(L^s)\theta(L)u_t$$

Dónde:

$$\Phi(L^s) = 1 - \Phi_1 L^s - \Phi_2 L^{2s} - \dots - \Phi_p L^{ps}$$

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$$

$$\theta(L^s) = 1 - \theta_1 L^s - \theta_2 L^{2s} - \dots - \theta_q L^{qs}$$

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$$

Los operadores anteriormente expandidos, son operadores estacionales e invertibles, sin factores comunes, siendo $\{u_t\}$ una secuencia de ruidos blancos. En el caso particular en que z_t venga generada por un proceso ARMA (p,q), la ecuación anterior será igual a: $z_t = \frac{\theta(L)}{\phi(L)} u_t$

De ahora en adelante supondremos que $\{z_t\}$, viene generada por un proceso ARMA (p,q), (para simplificar la exposición y dado que la generalización resulta inmediata).

Respecto a Hillmer, Bell y Tiao, 1983, Muirhead, 1986, Tsay, 1988, Chen y Liu, 1990, Chen, Liu y Hudak, 1990, y Chen y Tiao, 1990, los cuatro tipos de outliers que se han considerado en la literatura son: el outliers aditivo (additive outliers, AO), el outliers innovacional (innovational outliers, IO), el cambio en nivel (level shift, LS) y el cambio temporal (temporary change, TC).

a) Outliers Aditivo (AO)

Un outliers aditivo es un suceso que afecta a una serie en un solo instante, temporal ($t=t_0$), y puede expresar como: $y_t = z_t + \omega I_t^{t_0}$

Donde
$$I_t^{t_0} = \begin{cases} 1 & t = t_0 \\ 0 & t \neq t_0 \end{cases}$$

$I_t^{t_0}$: Presencia o ausencia de Outliers en el periodo t_0

ω : Efecto de dicho Outliers

z_t : Serie Libre de Outliers

y_t : Serie Observada

b) Outliers Innovacional (IO)

Un outlier innovacional es un suceso cuyo efecto se propaga en conformidad con el modelo ARIMA del proceso, afectando a todos los valores observados después de su ocurrencia.

Se puede representar como sigue:

$$y_t = z_t + \frac{\theta(L)}{\phi(L)} \omega I_t^{t_0}$$

c) Cambio de Nivel (LS)

Un cambio en nivel (LS) es un suceso que afecta a una serie en un período dado y cuyo efecto es permanente, se puede representar como sigue:

$$y_t = z_t + \frac{1}{(1-L)} \omega I_t^{t_0}$$

d) Cambio Temporal (TC)

Un cambio temporal (TC) es un suceso que tiene un impacto inicial y cuyo efecto decae exponencialmente con un factor de amortiguación, que denotaremos mediante el parámetro δ por lo que se puede expresar como:

$$y_t = z_t + \frac{1}{(1-\delta L)} \omega I_t^{t_0} \quad \text{para } 0 < \delta < 1$$

Debe tenerse en cuenta que una serie temporal puede contener varios Outliers (por ejemplo k) de diferentes tipos pudiéndose expresar el modelo general de outliers como:

$$y_t = \sum_{j=1}^k \omega_j v_j(L) I_{t_j} + z_t$$

$$z_t = \frac{\theta(L)}{\phi(L)} u_t$$

donde:

$$v_j(L) = \begin{cases} 1 & \text{Para un AO en } t = t_{0j} \\ \frac{\theta(L)}{\phi(L)} & \text{Para un IO en } t = t_{0j} \\ \frac{1}{(1-L)} & \text{Para un LS en } t = t_{0j} \\ \frac{1}{(1-\delta L)} & \text{Para un TC en } t = t_{0j} \end{cases}$$

4.4.4.2. Efecto de los Outliers

Se distinguirá entre los efectos de los outliers en la serie original y en la serie residual. Para el caso de los efectos de los outliers en la serie original resultan inmediatos a partir de las expresiones de los Outliers Aditivos, Innovacional, Cambio de Nivel, y Cambio Temporal. En concreto, las observaciones respecto a los Outliers son los siguientes:

- 1) Los efectos de los Outliers son independientes de la estructura ARIMA de las series, a excepción del Outliers Innovacional.
- 2) Los Outliers AO y LS son caso límites del Outlier de Cambio Temporal. Dada la definición del cambio temporal, cuando $\delta = 0$ estamos ante el OA, y cuando $\delta = 1$ ante el cambio en nivel.
- 3) El Outlier Aditivo causa un efecto inmediato y único en la serie observada en el período t_0 , de magnitud ω .
- 4) En el caso de un Cambio Temporal, el Outlier produce un efecto inicial ω en el período t_0 , y este efecto decae gradualmente en el tiempo, con un factor de amortiguación δ (aconseja utilizar el valor $\delta = 0.7$ según el analista Chen y Liu(1990))

- 5) En un Cambio de Nivel, el Outlier introduce un cambio brusco y permanente en la serie observada de tipo escalón
- 6) El efecto del Outlier Innovacional es el más complicado, por cuanto depende de la estructura ARIMA de la serie.

A partir de la expresión $y_t = Z_t + \frac{\theta(L)}{\phi(L)} \omega I_t^{t_0}$ podemos concluir que cuando un outlier IO ocurre en $t = t_0$, el efecto de ese outlier en $t_0 + j$ para $j \geq 0$ es igual a $\omega \psi_j$, donde ω es el efecto inicial y ψ_j es el coeficiente j-ésimo del polinomio $\psi(L)$, el cual tomando como modelo de z_t el $ARMA_{(p,q)}$, la expresión $z_t = \frac{\theta(L)}{\phi(L)} u_t$ es igual a:

$$\psi(L) = \frac{\theta(L)}{\phi(L)} = 1 + \psi_1 L + \psi_1 L^2 + \dots$$

Y para el caso de la serie residual, para analizar los efectos Outliers, se comenzará suponiendo el caso ideal (e irreal) de que los parámetros del modelo ARMA de la serie son conocidos. En este caso, se examinará las series filtradas, estas se escriben de la siguiente manera: $\hat{u}_t = \pi(L) Y_t$, donde

$$\pi(L) = \frac{\theta(L)}{\phi(L)} = 1 - \pi_1 L - \pi_2 L^2 - \dots$$

Considerando la expresión anterior y que tenemos un solo Outliers que se produce en el período $t = t_0$. Podemos reescribir \hat{u}_t de acuerdo con el tipo de Outlier presente como sigue:

$$AO: \hat{u}_t = \omega \pi(L) I_t^{t_0} + u_t$$

$$IO: \hat{u}_t = \omega I_t^{t_0} + u_t$$

$$LS: \hat{u}_t = \frac{\omega}{1-L} \pi(L) I_t^{t_0} + u_t$$

$$TC: \hat{u}_t = \frac{\omega}{1-\delta L} \pi(L) I_t^{t_0} + u_t$$

En forma conjunta, las expresiones anteriores se pueden escribir como sigue:

$$\hat{u}_t = \omega x_{it} + u_t, \text{ Siendo } x_{it} = \begin{cases} x_{1t} = \pi(L)I_t^{t_0} \\ x_{2t} = I_t^{t_0} \\ x_{3t} = \frac{1}{1-L}\pi(L)I_t^{t_0} \\ x_{4t} = \frac{1}{1-\delta L}\pi(L)I_t^{t_0} \end{cases} \quad \text{donde } i=1,2,3,4.$$

Y donde:

$$x_{it} = \begin{cases} 0 & \text{para } t < t_0 \text{ y } \forall_i \\ 1 & \text{para } t = t_0 \text{ y } \forall_i \\ -\pi_j & \text{para } t = t_0 + j (j = 1, 2, \dots, T - t_0), i = 1 \text{ (AO)} \\ 0 & \text{para } t = t_0 + j (j = 1, 2, \dots, T - t_0), i = 2 \text{ (IO)} \\ 1 - \sum_{h=1}^j \pi_h & \text{para } t = t_0 + j (j = 1, 2, \dots, T - t_0), i = 3 \text{ (LS)} \\ -\delta^j \sum_{h=1}^{j-1} \delta^{j-h} \pi_h - \pi_j & \text{para } t = t_0 + j (j = 1, 2, \dots, T - t_0), i = 4 \text{ (TC)} \end{cases}$$

En todos los outliers, excepto en el innovacional se observa, que acota sus efectos en los residuos en el instante temporal $t = t_0$, propaga su efecto en todos los residuos correspondientes a los períodos de tiempo posteriores a t_0 , dependiendo además de dichos coeficientes π_j , los cuales están en función de los parámetros del modelo ARMA identificado para la serie ($\hat{u}_t = \pi(L)Y_t$).

4.4.5. Estimación de los efectos de los outliers

Si solamente hubiese un outlier del cual conociésemos tanto el tipo, como el período de ocurrencia, y además también se conocen los parámetros del modelo ARMA subyacente a la serie temporal, la estimación de los efectos de los outliers podría llevarse a cabo a partir de las expresiones anteriores, obteniendo:

$$\tilde{\omega}_m(t_0) = \frac{\sum_{t=t_0}^T \hat{u}_t x_{it}}{\sum_{t=t_0}^T x_{it}^2}, \text{ siendo } m = \begin{cases} AO & \text{cuando } i = 1 \\ IO & \text{cuando } i = 2 \\ LS & \text{cuando } i = 3 \\ TC & \text{cuando } i = 4 \end{cases}$$

Las estimaciones de los efectos de los efectos de los distintos tipos de outliers en $t = t_0$, serán iguales a:

$$\tilde{\omega}_{AO}(t_0) = \frac{\hat{u}_{t_0} - \sum_{j=1}^{T-t_0} \pi_j \hat{u}_{t_0+j}}{\tau^2}$$

$$\tilde{\omega}_{IO}(t_0) = \hat{u}_{t_0}$$

$$\tilde{\omega}_{LS}(t_0) = \frac{\hat{u}_{t_0} + \sum_{j=1}^{T-t_0} (1 - \sum_{h=1}^j \pi_h) \hat{u}_{t_0+j}}{\tau^{*2}}$$

$$\tilde{\omega}_{TC}(t_0) = \frac{\hat{u}_{t_0} + \sum_{j=1}^{T-t_0} (\delta^j - \sum_{h=1}^{j-1} \delta^{j-h} \pi_h - \pi_h) \hat{u}_{t_0+j}}{\tau^{**2}}$$

Dónde:

$$\tau^2 = \sum_{j=0}^{T-t_0} \pi_j^2, \quad \pi_0 = 1$$

$$\tau^{*2} = 1 + \sum_{j=1}^{T-t_0} (1 - \sum_{h=1}^j \pi_h)^2$$

$$\tau^{**2} = 1 + \sum_{j=1}^{T-t_0} (\delta^j - \sum_{h=1}^{j-1} \delta^{j-h} \pi_h - \pi_j)^2$$

Las varianzas de los estimadores $\tilde{\omega}_m(t_0)$ serán iguales, en consecuencia a:

$$Var[\tilde{\omega}_m(t_0)] = \frac{\sigma_u^2}{\sum_{t=t_0}^T x_{it}^2}$$

Esto es:

$$Var[\tilde{\omega}_{AO}(t_0)] = \frac{\sigma_u^2}{\tau^2}$$

$$Var[\tilde{\omega}_{IO}(t_0)] = \sigma_u^2$$

$$Var[\tilde{\omega}_{LS}(t_0)] = \frac{\sigma_u^2}{\tau^{*2}}$$

$$Var[\tilde{\omega}_{TC}(t_0)] = \frac{\sigma_u^2}{\tau^{**2}}$$

A partir de las expresiones anteriores pueden definirse los cuatro estadísticos siguientes:

$$\lambda_{1,t_0} = \frac{\tilde{\omega}_{AO}(t_0)}{DT[\tilde{\omega}_{AO}(t_0)]} = \frac{\tau \tilde{\omega}_{AO}(t_0)}{\sigma_u}$$

$$\lambda_{2,t_0} = \frac{\tilde{\omega}_{IO}(t_0)}{DT[\tilde{\omega}_{IO}(t_0)]} = \frac{\tilde{\omega}_{IO}(t_0)}{\sigma_u}$$

$$\lambda_{3,t_0} = \frac{\tilde{\omega}_{LS}(t_0)}{DT[\tilde{\omega}_{LS}(t_0)]} = \frac{\tau^* \tilde{\omega}_{LS}(t_0)}{\sigma_u}$$

$$\lambda_{4,t_0} = \frac{\tilde{\omega}_{TC}(t_0)}{DT[\tilde{\omega}_{TC}(t_0)]} = \frac{\tau^{**} \tilde{\omega}_{TC}(t_0)}{\sigma_u}$$

Los cuales, bajo la hipótesis nula de que y_{t_0} no es un outlier, siguen una distribución $N(0,1)$.

En consecuencia, si desconociésemos el periodo de ocurrencia de los outliers, pero conociésemos los valores de los parámetros del proceso ARMA subyacente a la serie temporal, esto es, los coeficientes π_j podríamos contrastar directamente si la observación t_0 corresponde a un outlier, así como el tipo del mismo, actuando como sigue, una vez fijado un nivel de significación ε :

- ✓ Si $|\lambda_{i,t_0}| \leq N_{\varepsilon/2} \Rightarrow$ aceptaremos la hipótesis nula, es decir, la observación t_0 -ésima no es un outlier
- ✓ Si $|\lambda_{i,t_0}| > N_{\varepsilon/2} \Rightarrow$ rechazaremos la hipótesis nula, diciendo que estamos ante un outlier AO si $i = 1$, un IO si $i = 2$, un LS si $i = 3$ o un TC si $i = 4$.

4.4.6. Detección y tratamiento de los Outliers

Un tratamiento adecuado de los outliers permite mejorar cuatro aspectos, por eso es muy importante la detección de los valores atípicos (Chen, Liu y Hudak (1990))¹⁰. Los aspectos son los siguientes:

- 1) Comprensión de la serie objeto de análisis
- 2) Modelización y estimación
- 3) Análisis de intervención
- 4) Calidad de las predicciones

En la práctica, t_0 y los valores de los parámetros $(\phi_j, \theta_j, \sigma_u^2)$ no son conocidos, por lo cual estos deben estimarse, es notorio además, que las existencias de los outliers causa sesgos importantes en las estimaciones de estos parámetros, concretamente, σ_u^2 tenderá a estar sobreestimado.

Se han propuesto procedimientos iterativos para detectar outliers, como el de Chang y Tiao (1983) o el de Tsay (1986), los cuales se centran en detectar los posibles outliers y especificar posteriormente de acuerdo con el tipo de outliers identificado un modelo de intervención, tal como se expresa en términos genéricos en la siguiente expresión:

$$y_t = \sum_{j=1}^k \omega_j v_j(L) I_{tj} + z_t,$$

y finalmente estimar dicho modelo.

A continuación se describe cada uno de estos enfoques, centrándose en el primero de ellos, en el procedimiento iterativo de Chang y Tiao (1983).

4.4.6.1. Procedimiento Iterativo de Detección de Outliers

Chang y Tiao (1983) han propuesto un procedimiento iterativo para detectar posibles outlier y el tipo de ellos, que consta de las siguientes etapas.

Etapas 1: Modelizamos la serie $\{y_t\}$ suponiendo que no hay outliers y estimamos el modelo identificado: $\tilde{\phi}(L)y_t = \tilde{\theta}(L)\hat{u}_t$ o lo que es igual a: $y_t = \frac{\tilde{\theta}(L)}{\tilde{\phi}(L)}\hat{u}_t$

¹⁰ Antonio Aznar y Francisco Javier Trávez, Métodos de Predicción en economía(II), pág. 328

Donde:

$$\tilde{\theta}(L) = 1 - \tilde{\theta}_1 L - \tilde{\theta}_2 L^2 - \dots - \tilde{\theta}_q L^q$$

$$\tilde{\phi}(L) = 1 - \tilde{\phi}_1 L - \tilde{\phi}_2 L^2 - \dots - \tilde{\phi}_q L^q$$

A partir de $y_t = \frac{\tilde{\theta}(L)}{\tilde{\phi}(L)} \hat{u}_t$ los residuos son iguales a: $\hat{u}_t = \tilde{\pi}(L) Y_t$

Donde:

$$\tilde{\pi}(L) = \frac{\tilde{\theta}(L)}{\tilde{\phi}(L)} = 1 - \tilde{\pi}_1 L - \tilde{\pi}_2 L^2 \dots$$

y la estimación de σ_u^2 será: $\tilde{\sigma}_u^2 = \frac{1}{T} \sum_{t=1}^T (\hat{u}_t - \bar{\hat{u}})^2$

Etapa 2: Se calcula $\tilde{\lambda}_{i,t}$ donde $i = 1,2,3,4$ utilizando el modelo estimado, esto es, teniendo en cuenta los estadísticos calculados previamente.

$$\tilde{\lambda}_{1,t_0} = \frac{\tilde{\tau} \tilde{\omega}_{AO}(t_0)}{\tilde{\sigma}_u}; \tilde{\lambda}_{2,t_0} = \frac{\tilde{\omega}_{IO}(t_0)}{\tilde{\sigma}_u}; \tilde{\lambda}_{3,t_0} = \frac{\tilde{\tau}^* \tilde{\omega}_{LS}(t_0)}{\tilde{\sigma}_u}; \tilde{\lambda}_{4,t_0} = \frac{\tilde{\tau}^{**} \tilde{\omega}_{TC}(t_0)}{\tilde{\sigma}_u}$$

Entonces definiendo $\tilde{\lambda}_{t_0} = \text{Max}_t \text{Max}_i \{|\tilde{\lambda}_{i,t}|\}$

Donde t_0 denota el período para el que tiene lugar el máximo valor de $\tilde{\lambda}$, si:

a) $\tilde{\lambda}_{t_0} < c$, donde c es una constante positiva predeterminada a la que se tomara de referencia mas adelante, diremos que no hay ningún outliers, finalizado el procedimiento.

b. 1) $\tilde{\lambda}_{t_0} = |\tilde{\lambda}_{1,t_0}| > c$,

Se dirá que en el periodo t_0 ocurre un outlier AO, cuyo efecto estimado es $\tilde{\omega}_{AO}(t_0)$. En este caso, podremos modificar los valores de la serie, utilizando $y_t = z_t + \omega I_t^{t_0}$, como sigue:

$$\tilde{y}_t = y_t - \tilde{\omega}_{AO}(t_0) I_t^{t_0} \text{ y definir los nuevos residuos: } \tilde{u}_t = \tilde{\pi}(L) \tilde{y}_t = \hat{u}_t - \tilde{\omega}_{AO}(t_0) \tilde{\pi}(L) I_t^{t_0}$$

A partir de estos nuevos residuos (\tilde{u}_t), se calculará una nueva estimación de σ_u^2 :

$$\tilde{\sigma}_u^2 = \frac{1}{T} \sum_{t=1}^T (\tilde{u}_t - \bar{\tilde{u}})^2$$

b. 2) $\tilde{\lambda}_{t_0} = |\tilde{\lambda}_{2,t_0}| > c$

Para este caso diremos que en el periodo t_0 ocurre un outlier IO, cuyo efecto es $\tilde{\omega}_{IO}(t_0)$.

También en este caso puede eliminarse el efecto IO modificando la serie utilizando

$$y_t = z_t + \frac{\theta(L)}{\phi(L)} \omega I_t^{t_0} \text{ Obteniendo así } \tilde{y}_t = y_t - \frac{\tilde{\theta}(L)}{\tilde{\phi}(L)} \tilde{\omega}_{IO}(t_0) I_t^{t_0},$$

Siendo los nuevos residuos $\tilde{u}_t = \tilde{\pi}(L)\tilde{y}_t = \hat{u}_t - \tilde{\omega}_{IO}(t_0)\tilde{\pi}(L)I_t^{t_0}$

b. 3) $\tilde{\lambda}_{t_0} = |\tilde{\lambda}_{3,t_0}| > c$

Diremos que en el periodo t_0 ocurre un cambio en nivel (LS), cuyo efecto es $\tilde{\omega}_{LS}(t_0)$. También

en este caso puede eliminarse el efecto LS modificando la serie utilizando

$$y_t = z_t + \frac{1}{(1-L)} \omega I_t^{t_0} \text{ Obteniendo así } \tilde{y}_t = y_t - \frac{1}{(1-L)} \tilde{\omega}_{LS}(t_0) I_t^{t_0},$$

Siendo los nuevos residuos a partir de los cuales estimaremos nuevamente

$$\sigma_u^2: \tilde{u}_t = \tilde{\pi}(L)\tilde{y}_t = \hat{u}_t - \tilde{\omega}_{LS}(t_0)\tilde{\pi}(L)I_t^{t_0}$$

b. 4) $\tilde{\lambda}_{t_0} = |\tilde{\lambda}_{4,t_0}| > c$

Finalmente para esta estimación en el periodo t_0 ocurre un Cambio Temporal (TC), cuyo efecto

es $\tilde{\omega}_{TC}(t_0)$. También, en este caso se puede eliminar el efecto TC modificando la serie, utilizando

$$y_t = z_t + \frac{1}{(1-\delta L)} \omega I_t^{t_0}, \text{ Obteniendo así } \tilde{y}_t = y_t - \frac{1}{(1-\delta L)} \tilde{\omega}_{TC}(t_0) I_t^{t_0},$$

Siendo los nuevos residuos a partir de los cuales estimaremos nuevamente

$$\sigma_u^2: \tilde{u}_t = \tilde{\pi}(L)\tilde{y}_t = \hat{u}_t - \frac{\tilde{\omega}_{TC}(t_0)}{(1-\delta L)} \tilde{\pi}(L) I_t^{t_0}$$

Etapa 3: Volvemos a calcular los $\tilde{\lambda}_{i,t}$ donde $i = 1,2,3,4$. Basandonos ahora en los residuos \tilde{u}_t y, consecuentemente, en $\tilde{\sigma}_u^{*2}$, y repetimos la etapa 2 hasta que todos los outliers son identificados (esto es, hasta que en alguna iteración nos encontremos con el caso a) de la etapa 2). Debe tenerse en cuenta que las estimaciones iniciales en $\pi(L)$ permanecen invariables.

Etapa 4: Supongamos que la etapa 3 finaliza con la detección de outliers en los periodos $t_{01}, t_{02}, t_{03}, \dots, t_{0k}$. Entonces, trataremos estos periodos como si fueran conocidos estimando los parámetros de los outliers $\omega_1, \omega_2, \dots, \omega_k$ y los parámetros ϕ_j y θ_j simultáneamente utilizando el modelo: $y_t = \sum_{j=1}^k \omega_j v_j(L) I_{t_j} + z_j$

$$\text{Dónde: } v_j(L) = \begin{cases} 1; & \text{Para un AO en } t = t_{0j} \\ \frac{\theta(L)}{\phi(L)}; & \text{Para un IO en } t = t_{0j} \\ \frac{1}{(1-L)}; & \text{Para un LS en } t = t_{0j} \\ \frac{1}{(1-\delta L)}; & \text{Para un TC en } t = t_{0j} \end{cases}$$

Desde la estimación de $y_t = \sum_{j=1}^k \omega_j v_j(L) I_{t_j} + z_t$ podemos calcular unos nuevos residuos como sigue:

$$\hat{u}_t^{(1)} = \tilde{\pi}^{(1)}(1) [y_t - \sum_{j=1}^k \tilde{\omega}_j \tilde{v}_j(L) I_t^{t_{0j}}],$$

obteniendo por tanto una nueva estimación de σ_u^2

$$\tilde{\sigma}_u^{2(1)} = \frac{1}{T} \sum_{t=1}^T (\hat{u}_t^{(1)} - \bar{\hat{u}}^{(1)})^2$$

Las etapas 2 a 4 son repetidas sucesivamente hasta que todos los outliers son identificados y sus impactos son estimados simultáneamente.

El modelo de los outliers estimado será:

$$y_t = \sum_{j=1}^k \tilde{\omega}_j \tilde{v}_j(L) I_t^{t_{0j}} + \frac{\tilde{\theta}(L)}{\tilde{\phi}(L)} u_t,$$

donde $\tilde{\omega}_j$, $\tilde{\phi}(L)$ y $\tilde{\theta}(L)$ son los estimadores obtenidos en la iteración final.

Respecto a este procedimiento de detección de outliers cabe efectuar dos observaciones:

- 1) El valor de la constante positiva c , que constituye el punto crítico para aceptar o rechazar la hipótesis nula de ausencia de outliers, se recomienda a partir de determinados experimentos de simulación que debe tomarse valores comprendidos entre 3 y 4.5. lo habitual es tomar el valor 3, sobre todo cuando el tamaño muestral disponible no es muy elevado.
- 2) Si bien el procedimiento se ha explicado considerando los cuatro tipos de outliers, solamente había sido desarrollado para los outliers AO e IO.

Finalmente, hay que destacar que si bien, este procedimiento iterativo es efectivo cuando se trata de detectar las localizaciones y estimar los efectos de grandes outliers aislados, Chen y Liu (1990) encuentran tres problemas que surgen a partir del mismo:

- a) La presencia de outliers puede hacer que el modelo no se especifique adecuadamente.
- b) Incluso, si el modelo esta adecuadamente especificado, los outliers pueden producir sesgos importantes en los estimadores de los parámetros.
- c) Puede que no se identifiquen algunos outliers debido a un efecto de enmascaramiento.

Con el fin de paliar estos problemas (y más concretamente los dos últimos referidos), se ha propuesto el método alternativo de detección de outliers que describimos en el apartado siguiente:

4.4.6.2. Procedimiento de detección de outliers y estimación conjunta de los parámetros y los efectos de los Outliers

Chen y Liu (1990) se centran en los problemas b) y c) mencionados anteriormente, diseñando un procedimiento que, según sus palabras, “este libre del efecto espureo y ocultamiento de la detección de outliers, y sea capaz de estimar conjuntamente los parámetros del modelo y los efectos de los outliers”

El punto de partida para este procedimiento viene dado por el modelo general de los outliers múltiples, $y_t = \sum_{j=1}^k \omega_j v_j(L) I_{jt}^{t_0^j} + z_t$. Y la siguiente expresión de los residuos, que se deriva y directamente a partir de $\hat{u}_t = \pi(L)y_t$, entonces $\hat{u}_t = \sum_{j=1}^k \omega_j \pi(L) v_j(L) I_{jt}^{t_0^j} + u_t$.

Esto es $y_t = \sum_{j=1}^k \omega_j v_j(L) I_{jt}^{t_0^j} + z_t$, multiplicado en ambos lados por $\pi(L)$, y conociendo que $\hat{u}_t = \pi(L)y_t$, entonces tenemos:

$$\pi(L)y_t = \pi(L) \sum_{j=1}^k \omega_j v_j(L) I_{jt}^{t_0^j} + z_t = \sum_{j=1}^k \omega_j \pi(L) v_j(L) I_{jt}^{t_0^j} + z_t \pi(L)$$

$$\hat{u}_t = \sum_{j=1}^k \omega_j \pi(L) v_j(L) I_{jt}^{t_0^j} + u_t$$

Obviamente, si el efecto de un outlier y su localización se conocieran, se trataría de especificar adecuadamente la expresión: $y_t = \sum_{j=1}^k \omega_j v_j(L) I_{jt}^{t_0^j} + z_t$ y estimar los parámetros de dicho modelo, lo habitual sin embargo, es que no se conocen, por lo cual debe diseñar un procedimiento que permita tanto detectar los posibles outliers, como estimar los parámetros del modelo y los efectos de dichos outliers.

El procedimiento desarrollado con este fin por Chen y Liu (1990) consta de las tres etapas siguientes:

Etapla 1: estimación de los parámetros y detección inicial de los outliers.

- 1.1. Estimamos el modelo ARMA identificado utilizando las series observadas ajustadas más recientemente (el procedimiento comienza sin ajuste, tal como se señala en la etapa 1 del procedimiento iterativo de Chang y Tiao, 1993). Calculamos una serie de residuos.
- 1.2. Si descubrimos un outlier potencial, eliminamos su efecto de los residuos y de las observaciones, teniendo en cuenta el tipo de outliers detectado, y repetimos la búsqueda de outliers. En caso contrario pasamos a la siguiente etapa.
- 1.3. Si no se ha encontrado ningún outlier, podemos concluir que la serie está libre de outliers, ahora bien, si hemos encontrado outliers a partir de la estimación inicial de los parámetros,

revisamos la estimación del modelo con los nuevos datos ajustados y repetimos las etapas iniciales. El procedimiento se repite hasta que no se encuentra ningún nuevo outlier.

Etapla 2: Estimación conjunta de efectos de outliers y parámetros del modelo.

2.1. Supongamos que hemos identificado k posibles outliers de diferentes tipos.

Entonces, los efectos de los outliers, ω'_j s, pueden estimarse conjuntamente utilizando el modelo de regresión múltiple $\hat{u}_t = \sum_{j=1}^k \omega_j \pi(L) v_j(L) I_{jt}^{t_0^j} + u_t$, siendo \hat{u}_t la variable output y $v_j(L) I_{jt}^{t_0^j}$ las variables inputs.

2.2. Calculamos los estadísticos $\tilde{\lambda}$ de los $\tilde{\omega}'_j$ s estimados, siendo:

$$\tilde{\lambda}_j = \frac{\tilde{\omega}_j}{DT(\tilde{\omega}_j)} \quad j = 1, 2, \dots, k$$

Si el mínimo valor, en términos absolutos, de estos estadísticos ($\tilde{\lambda}_j$) es inferior a un determinado valor crítico, esto es:

$$\min_j |\tilde{\lambda}_j| = \tilde{\lambda}_L \leq c$$

Siendo c el valor crítico utilizado en el procedimiento iterativo, habitualmente igual a 3, entonces el outlier del periodo t_{0L} es eliminado del conjunto de outliers identificados, volviendo a repetir la etapa 2.1 con los $k-1$ outliers restantes.

Repetimos 2.1 y 2.2 hasta que encontramos que el valor mínimo, en términos absolutos del estadístico $\tilde{\lambda}$ correspondiente a los outliers no eliminados, supera c , en cuyo caso, pasamos a 2.3.

2.3. Ajustamos la serie eliminando sólo los efectos de los outliers significativos, entendiendo por tales los así detectados a partir de 2.1 y 2.2, esto es, los que no han sido eliminados.

2.4. Calculamos las estimaciones máximo verosímiles de los parámetros del modelo en base a la serie ajustada obtenida en 2.3. si la variación en la estimación de σ_u con respecto a la estimación previa, es mayor que un determinado valor prefijado λ (habitualmente, $\lambda = 0.001$), volvemos a 2.1 para efectuar iteraciones adicionales, en otro caso, pasamos a 3.1.

Etapa 3. Detección de outliers a partir de las estimaciones finales de los parámetros.

- 3.1. Calculamos los residuos filtrando la serie original, teniendo en cuenta las estimaciones de los parámetros calculadas en 2.4.
- 3.2. Utilizamos los residuos obtenidos en 3.1 para volver a repetir las etapas 1 y 2, teniendo en cuenta: a) que las estimaciones de los parámetros utilizadas para repetir 1.2-1.4 son fijadas en 2.4, y b) que se omiten las etapas 2.3 y 2.4

Los $\tilde{\omega}'_j$ s estimados en 2.1 en esta última iteración, son las estimaciones finales de los efectos de los outliers detectados.

El procedimiento descrito puede aplicarse directamente a un modelo de intervención.

Las principales diferencias entre este procedimiento y el referido en el procedimiento iterativo de detección de outliers son las siguientes:

- a) La detección de outliers se efectúa iterativamente considerando tanto residuos como observaciones ajustadas.
- b) No se precisa llevar a cabo la estimación de modelos de intervención
- c) Los outliers se detectan en base a estimaciones robustas de los parámetros del modelo, es decir, las estimaciones de los parámetros del modelo son poco sensibles a variaciones ocasionales en la muestra, así como a la presencia de outliers.
- d) Los efectos de los outliers se estiman conjuntamente utilizando un modelo de regresión múltiple.

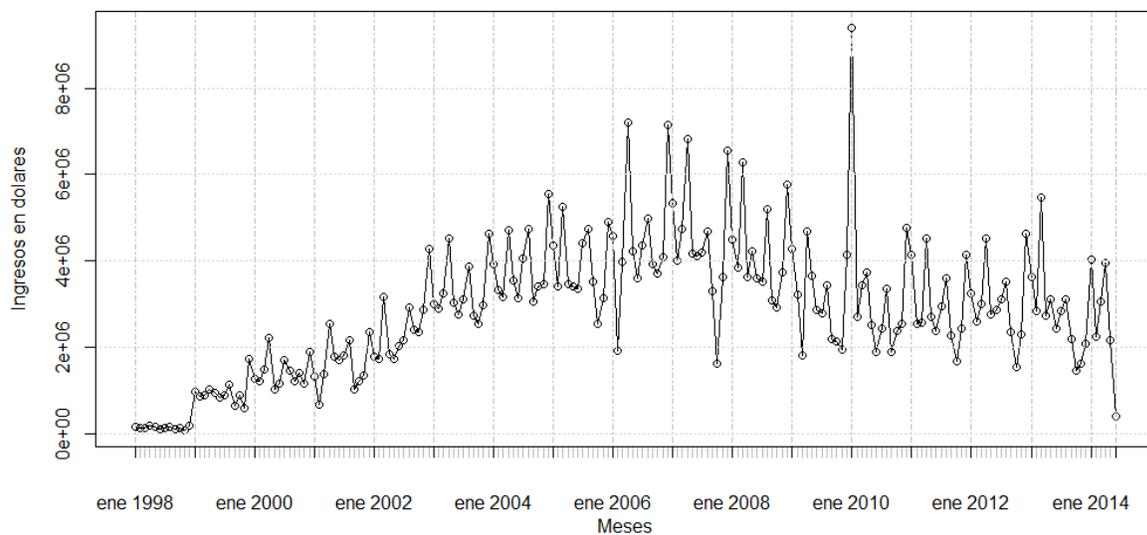
Para finalizar, cabe señalar que si bien los dos procedimientos de detección de outliers descritos son complejos, especialmente el último; pero la aplicabilidad de los mismos es inmediata, más cuando se conoce algún software en el que estén disponibles dichos procedimientos.

4.5. Resultados de la aplicación teórica

4.5.1. Análisis descriptivo de la serie temporal de ingresos mensuales en restaurantes del municipio de La Libertad

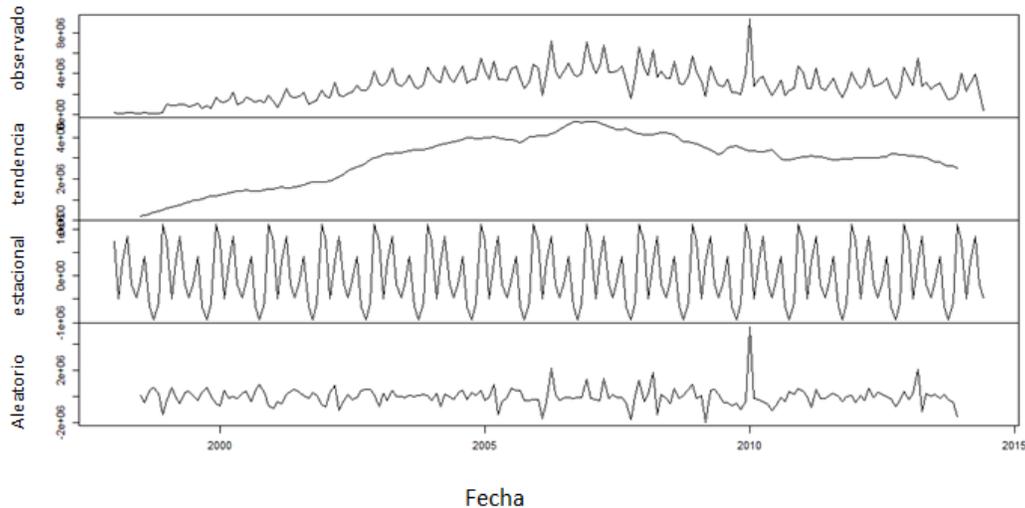
La Serie Temporal de ingresos en dólares de los restaurantes del municipio de La Libertad, consta de 198 datos con periodicidad mensual desde Enero de 1998 a Junio de 2014. Se inicia el análisis de la serie, representando gráficamente los ingresos de restaurantes, luego observando si es estacionaria, es decir que la media, la varianza y las autocorrelaciones no dependen del tiempo.

Gráfico 1. Serie original. Ingresos económicos mensuales en dólares en los Restaurantes del municipio de La Libertad. (Enero 1998 – Junio 2014)



Seguidamente, se realiza la descomposición de la serie para analizar cada una de sus componentes principales, como se muestra a continuación:

Gráfico 2. Descomposición de la serie de datos de los ingresos mensuales en los restaurantes del municipio de La Libertad. (Enero 1998 – Junio 2014)

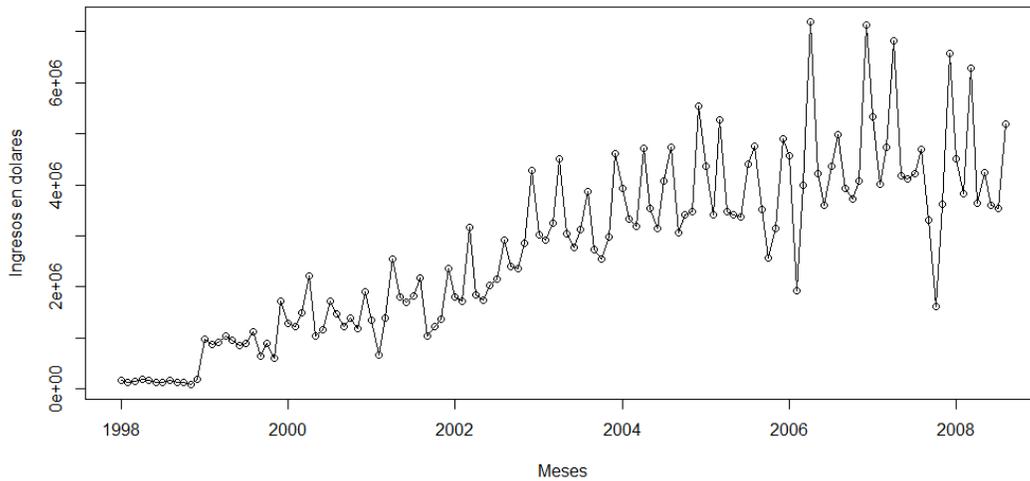


Puede observarse la evolución de los ingresos a largo plazo, en la cual se observa una tendencia creciente, sin embargo en ciertos valores (ver anexo 7.1) se observa que existe un comportamiento irregular a partir del mes de septiembre de 2008 ($t_0=129$), la cual posiblemente se trata de una intervención, está podría ser que se debe a la crisis económica mundial. Respecto a la componente estacional, se evidencian tres valores altos dentro de cada año, los primeros dos picos corresponden a los periodos vacacionales de abril y agosto, y el tercer pico con mayor valor a diciembre, la componente aleatoria muestra lo que no es explicado por las demás componentes.

Para identificar el modelo ARIMA que más se aproxima a los datos, se toma el subconjunto de datos anterior o posterior de la intervención que sea más representativo (mayor que el 60%)¹¹. Para los datos de ingresos económicos de los restaurantes se tiene el 64.97% antes de la intervención y 35.03% después de la intervención. A continuación se presenta la grafica de la serie con los datos antes de la intervención, ya que esta es la que supera el 60% de la serie completa.

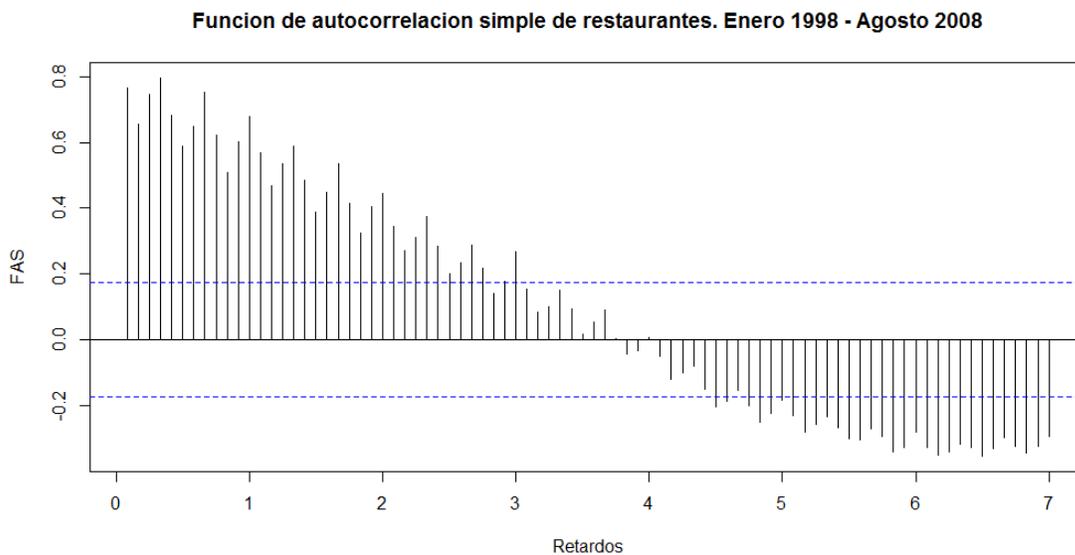
¹¹ Peña Sánchez, Daniel: Estadística. Modelos y Métodos. Volumen II, Alianza Editorial, 1989.

Gráfico 3. Serie de ingresos mensuales de los restaurantes, pre intervención¹² (Enero de 1998 – Agosto de 2008)



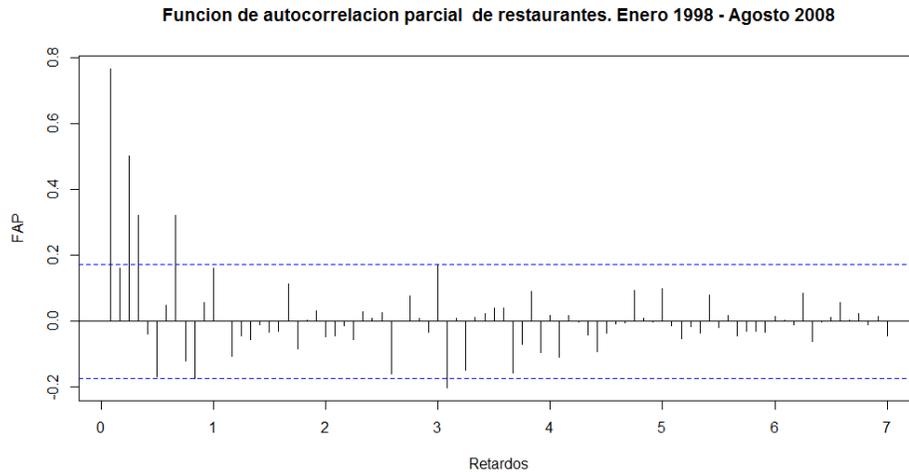
El gráfico 3 no presenta estacionariedad en varianza ni en media, puede observarse en los gráficos 4 y 5 de la Función de Autocorrelación Simple (FAS) y la Función de Autocorrelación Parcial (FAP) de la serie tomada (antes de la intervención).

Gráfico 4. FAS. Serie Restaurantes, pre intervención (Enero 1998 – Agosto 2008)



¹² “Pre intervención” sinónimo de “antes de la intervención”.

Gráfico 5. FAP. Serie de los ingresos de restaurantes, pre intervención (Enero 1998 – Agosto 2008)



En el gráfico 4, se observa que la función de autocorrelación simple decae lentamente a cero y en el gráfico 5, se observan que algunos retardos son significativos y los demás decaen lentamente a cero, este es un indicativo de que la serie no es estacionaria. Lo anterior lo verificaremos analizando el test de Levene y el de Dickey Fuller respectivamente.

Para verificar la estacionariedad en varianza se aplica el test de Levene que establece las hipótesis siguientes:

- $H_0: \sigma_1 = \sigma_2 = \dots = \sigma_k$, es decir, las varianzas son iguales para k submuestras.
- $H_1: \sigma_i \neq \sigma_j$, las varianzas son distintas para al menos un par (i, j) .

La hipótesis nula se rechaza si la significancia del test de Levene es menor que 0.05 (valor de significación nominal), es decir, en este caso las varianzas son significativamente diferentes, caso contrario se acepta la hipótesis.

Para nuestro caso la serie se dividió en 10 submuestras, tomando un grupo por año. Se aplicó el test utilizando el Software Estadístico R, obteniendo como resultado:

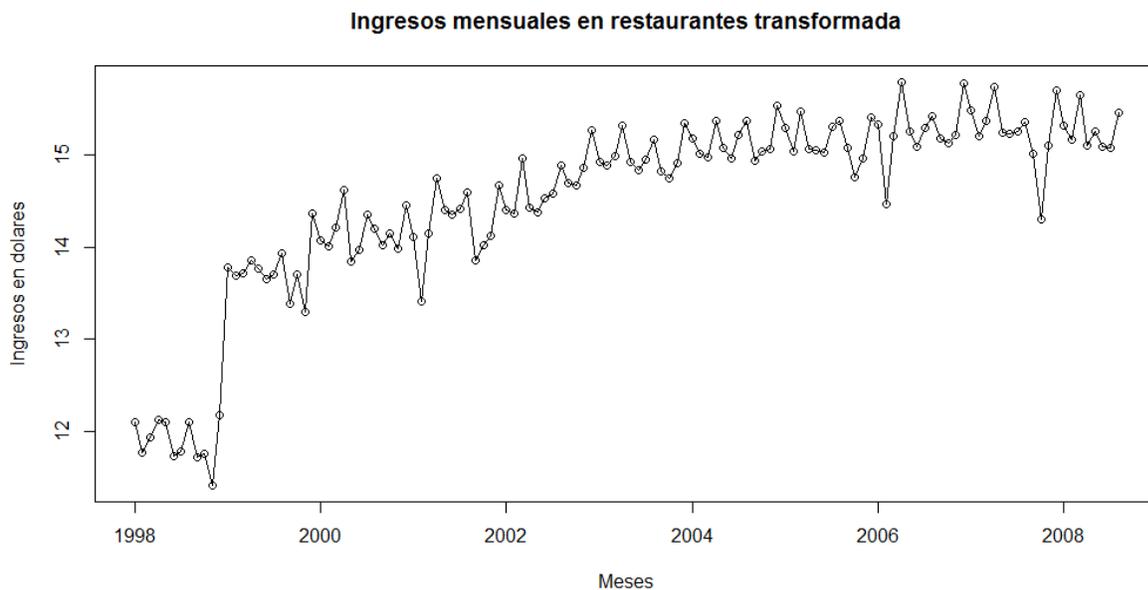
Tabla 9. Test de Levene para homogeneidad de varianza (serie pre intervención)

Levene's Test for Homogeneity of Variance (center = median)			
	Df	F value	Pr(>F)
group	10	2.6349	0.006254 **
	117		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

En la tabla 9, se observa que el estadístico de Levene con valor de 2.6349 tiene una significancia de 0.006254, por tanto se puede rechazar la hipótesis nula y se concluye que la serie no es estacionaria en varianza para conseguir la estacionariedad en varianza se aplica una transformación (log). Para ello, se elaboro el grafico con su respectiva transformación obteniendo:

Gráfico 6. Serie de ingresos económicos de los restaurantes, pre intervención transformada (log)



Se aplica nuevamente el test de Levene para verificar si ya es estacionaria en varianza y se tienen el siguiente resultado:

Tabla 10. Test de Levene para homogeneidad de varianza (serie de restaurantes transformada pre intervención)

Levene's Test for Homogeneity of Variance (center = median)			
	Df	F value	Pr(>F)
group	10	0.7478	0.6782
	117		

En la tabla 10, se observa que el estadístico de Levene con valor de 0.7478 tiene una significancia de 0.6782, por tanto, no se puede rechazar la hipótesis nula y se concluye que la serie es estacionaria en varianza.

Para confirmar que la serie presenta estacionariedad en media, se procede a realizar el test de raíces unitarias de Dickey-Fuller (DF)¹³.

Este test plantea las siguientes hipótesis:

- Hipótesis nula H_0 : existencia de raíz unitaria
- Hipótesis alterna H_1 : no existencia de raíz unitaria.

De donde:

- Si $t^* > DF$ no se rechaza H_0 , es decir existen raíces unitarias, por tanto la serie no es estacionaria.
- Si $t^* < DF$ se rechaza H_0 , es decir no existen raíces unitarias, por tanto la serie es estacionaria.

Donde t^* son los valores críticos que Dickey-Fuller (1979) y MacKinnon (1994) construyeron mediante simulaciones de Montecarlo, y establecieron unas tablas especiales en las que la t es superior. Se utilizara un nivel de significancia de 0.05.

Se realiza el contraste de hipótesis y se obtienen los resultados que se muestran a continuación:

¹³ Aznar, A., Trivez, Francisco, "Métodos de predicción en Economía, Tomo I, Capítulo 4, apartado 6.2, Ariel Economía, 1993.

Tabla 11. Test de Dickey-Fuller para homogeneidad de medias (serie de ingresos de los restaurantes, pre intervención transformada)

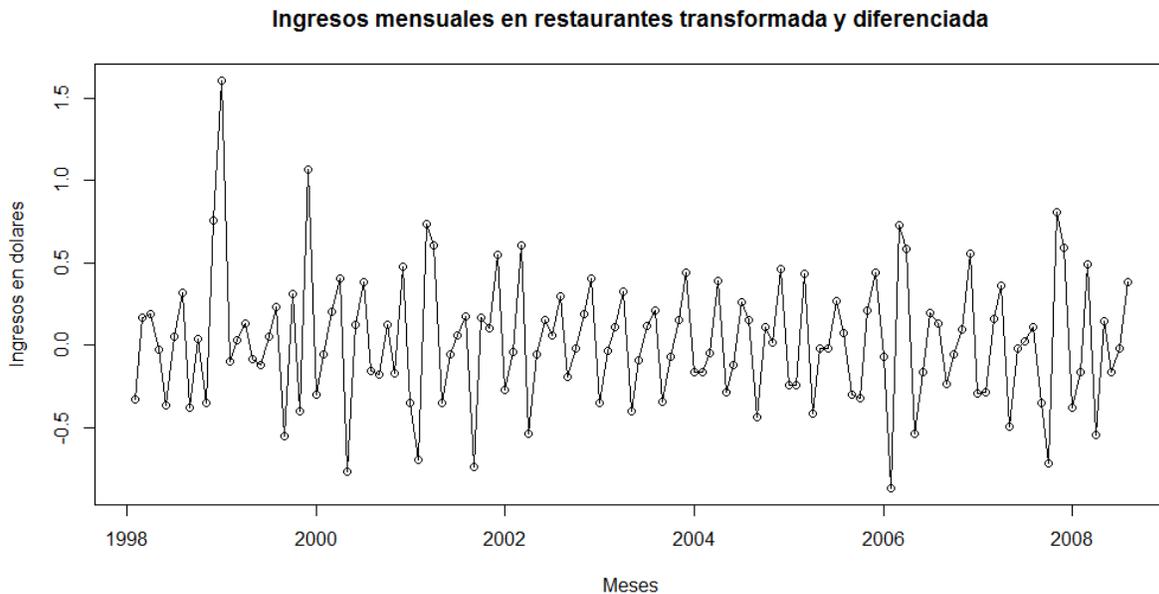
Augmented Dickey-Fuller Test

```
data: log.pre.int.restaurantes
Dickey-Fuller = -3.0376, Lag order = 5, p-value = 0.1452
alternative hypothesis: stationary
```

En la tabla 11 se observa que el estadístico de Dickey Fuller con valor de -3.0376 tiene una significancia de 0.1452, por lo que no se rechaza la hipótesis nula de la existencia de una raíz unitaria y por lo tanto, la serie no es estacionaria en media.

Para lograr la estacionariedad en media, se aplica una diferencia. Esto se muestra en el siguiente gráfico, donde se visualiza que la serie cumple la condición de estacionariedad en media:

Gráfico 7. Serie de ingresos de los restaurantes, pre intervención transformada con una diferencia regular ($d=1$)



En el gráfico 7, se observa que la serie es estacionaria en media, sin embargo, se aplica nuevamente el test para comprobar si la serie es estacionaria en media con una diferencia, los resultados se muestran a continuación:

Tabla 12. Test de Dickey-Fuller para homogeneidad de medias (serie de ingresos de los restaurantes, pre intervención transformada con una diferencia regular)

Augmented Dickey-Fuller Test

```
data: diff.log.pre.int.restaurantes
Dickey-Fuller = -6.0725, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary
```

En la tabla 12 se observa el estadístico de Dickey Fuller con valor de -6.0725 tiene un p-valor de 0.01, por lo que a un nivel de significancia de $\alpha = 0.05$ se rechaza la hipótesis nula de la existencia de una raíz unitaria y por lo tanto, la serie es estacionaria en media en la parte regular.

Para identificar los posibles modelos, se elaboran los gráficos de las funciones de autocorrelación muestral simple y la función de autocorrelación muestral parcial.

Gráfico 8. FAS de la serie pre intervención de ingresos económicos de restaurantes transformada y diferenciada en la parte regular.

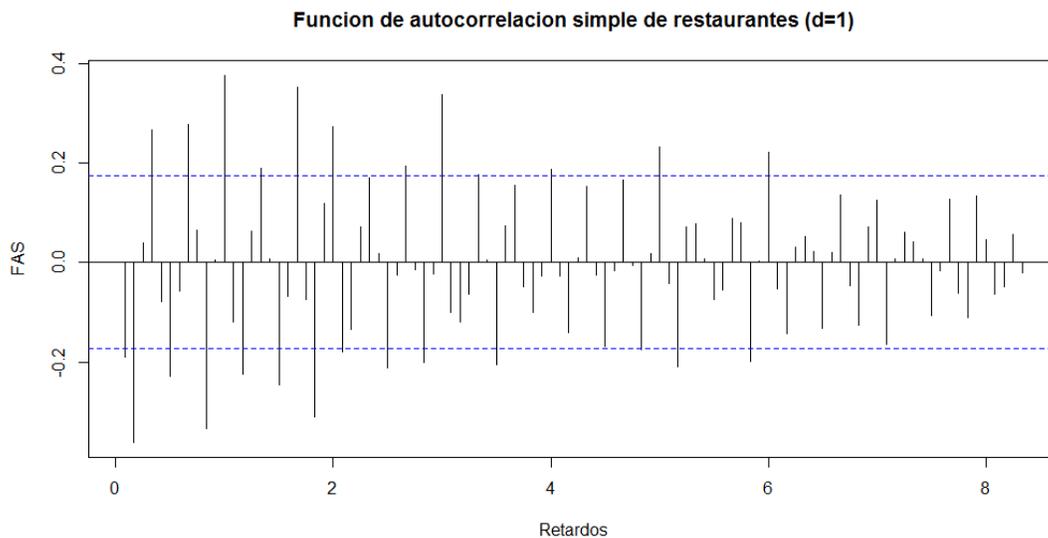
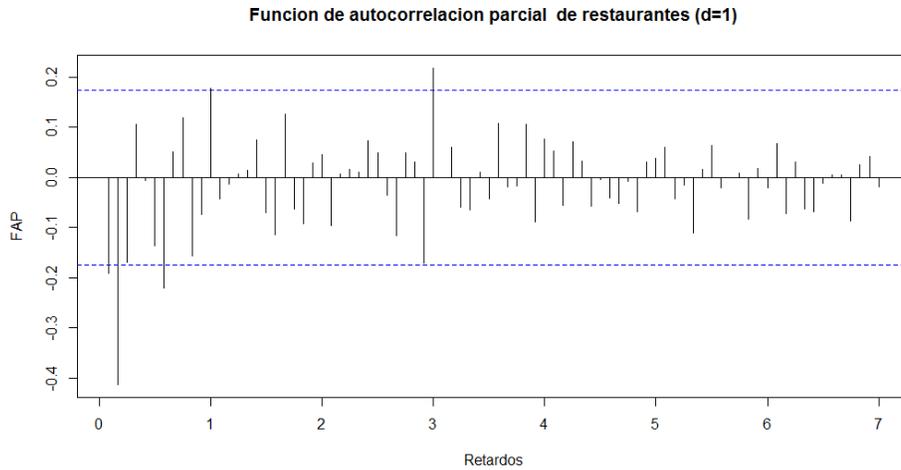


Gráfico 9. FAP de la serie pre intervención de ingresos económicos de los restaurantes, transformada y diferenciada en la parte regular.



En los gráficos anteriores (gráfico 8 y 9) se observa que existen varios retardos estacionales significativos y además existe una tendencia lentamente a anularse. Por lo tanto es necesario aplicar una diferencia estacional, con el objetivo de volver la serie estacionaria que es uno de los supuestos de los modelos ARIMA.

A continuación se presenta el grafico con una diferenciación en la parte regular y estacional.

Gráfico 10. FAS de la serie pre intervención de restaurantes, transformada y diferenciada en la parte regular y estacional.

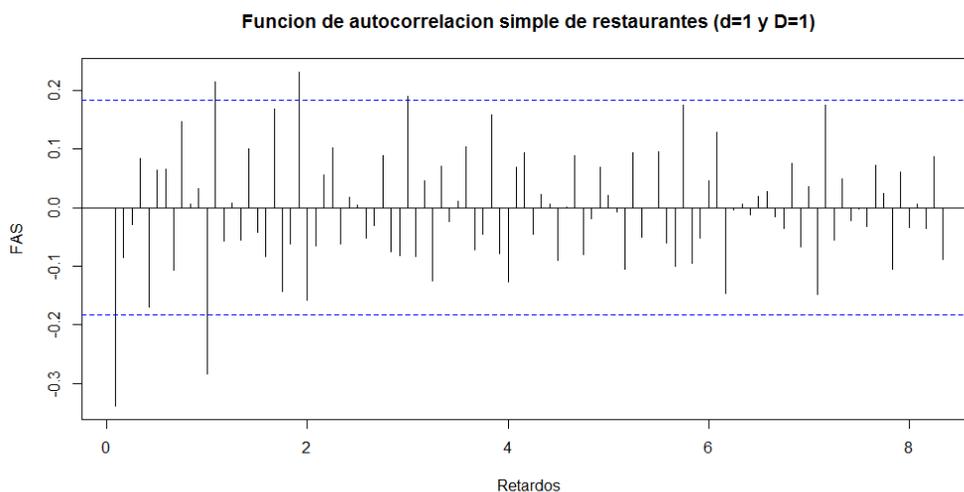
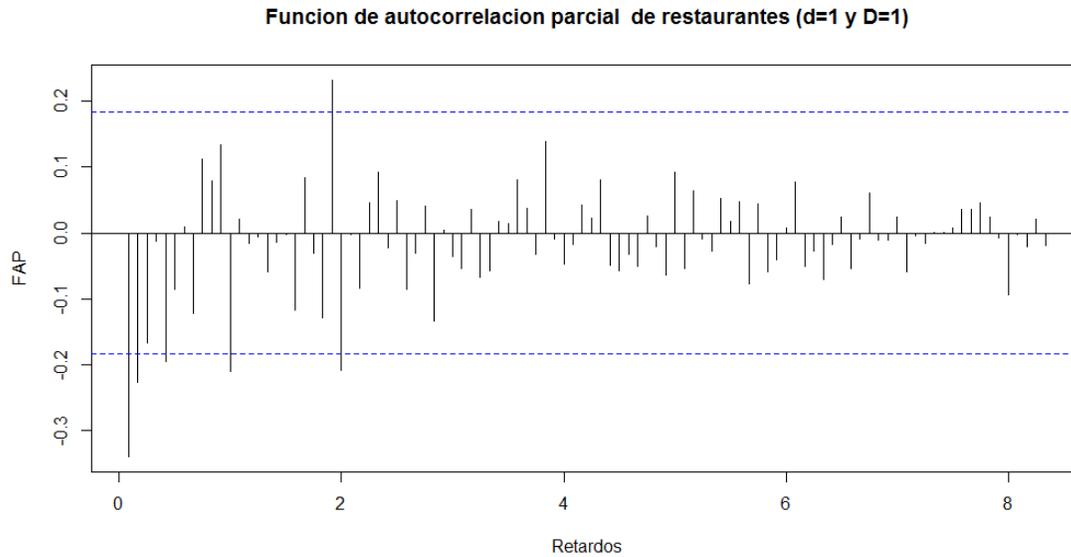


Gráfico 11. FAP de la serie pre intervención de los ingresos económicos de los restaurantes, transformada y diferenciada en la parte regular y estacional.



En los gráficos 10 y 11, se observa que las autocorrelaciones tienen menos retardos significativos, además tienden rápidamente a cero; por lo que se podría pensar que el modelo buscado es del tipo ARIMA (p,d,q) x (P,D,Q)_s.

Luego de observar los gráficos de la FAS y la FAP de la serie pre-intervención transformada (d=1 y D=1), se analizan ambas autocorrelaciones con el fin de identificar los posibles modelos, los cuales serán examinados en la etapa de estimación.

4.5.1.1. Identificación del modelo. Serie de Restaurantes

De acuerdo al análisis de la FAS y la FAP, se determinaron los posibles modelos:

1. ARI (2,1) x ARI (2,1)₁₂ $(1 - \phi_1 L - \phi_2 L^2)(1 - \phi_1 L^{12} - \phi_2 L^{24})(1 - L)(1 - L^{12}) \ln y_t = u_t$
2. ARI (2,1) x ARIMA(2,1,1)₁₂ $(1 - \phi_1 L - \phi_2 L^2)(1 - \phi_1 L^{12} - \phi_2 L^{24})(1 - L)(1 - L^{12}) \ln y_t = (1 - \theta_1 L^{12}) u_t$
3. ARI (2,1) x ARIMA(2,1,2)₁₂ $(1 - \phi_1 L - \phi_2 L^2)(1 - \phi_1 L^{12} - \phi_2 L^{24})(1 - L)(1 - L^{12}) \ln y_t = (1 - \theta_1 L^{12} - \theta_2 L^{24}) u_t$
4. ARIMA (2,1,1) x ARIMA(2,1,1)₁₂ $(1 - \phi_1 L - \phi_2 L^2)(1 - \phi_1 L^{12} - \phi_2 L^{24})(1 - L)(1 - L^{12}) \ln y_t = (1 - \theta_1 L)(1 - \theta_1 L^{12}) u_t$
5. ARI (1,1) x ARIMA(2,1,1)₁₂ $(1 - \phi_1 L)(1 - \phi_1 L^{12} - \phi_2 L^{24})(1 - L)(1 - L^{12}) \ln y_t = (1 - \theta_1 L^{12}) u_t$

4.5.1.2. Estimación de los parámetros

Una vez identificados los posibles modelos (ver anexo 7.2), se pretende en la estimación cuantificar los parámetros de cada modelo seleccionado. Después de revisar la significancia de los parámetros de los modelos identificados se muestran los siguientes modelos reducidos junto con sus respectivas expresiones, excluyendo los que no tuvieron parámetros significativos. Los modelos se muestran a continuación:

1. ARI (2,1) x ARI (2,1)₁₂ $(1 + 0.46L + 0.31L^2)(1 + 0.66L^{12} + 0.38L^{24})(1 - L)(1 - L^{12})lny_t = u_t$
2. ARI (2,1) x IMA(1,1)₁₂ $(1 + 0.39L + 0.24L^2)(1 - L)(1 - L^{12})lny_t = (1 + 0.71L^{12})u_t$
3. ARI (2,1) x I (1)₁₂ $(1 + 0.40L + 0.24L^2)(1 - L)(1 - L^{12})lny_t = u_t$
4. IMA (1,1) x ARIMA(2,1,1)₁₂ $(1 + 0.21L^{24})(1 - L)(1 - L^{12})lny_t = (1 + 0.64L)(1 + 0.68L^{12})u_t$, con $\Phi_1 = 0$
5. ARI (1,1) x ARIMA(2,1,1)₁₂ $(1 + 0.31L)(1 + 0.21L^{24})(1 - L)(1 - L^{12})lny_t = (1 + 0.76L^{12})u_t$, con $\Phi_1 = 0$

Tabla 13. Análisis de estacionariedad e invertibilidad de los modelos identificados reducidos. Serie de los ingresos económicos mensuales de los restaurantes.

Modelo	Estacionariedad	Invertibilidad
ARI (2,1) x ARI (2,1) ₁₂	<p>Debe cumplirse que</p> $\phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1 \text{ y } \phi_2 < 1,$ $\phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1, \phi_2 < 1,$ <p>siendo</p> $\phi_1 = -0.46, \phi_2 = -0.31, \phi_1 = -0.66,$ $\phi_2 = -0.38.$ <p><i>El modelo es estacionario</i></p>	<p>El modelo es un AR y una de las características de estos modelos es que siempre es invertible.</p>
ARI (2,1) x IMA(1,1) ₁₂	<p>Debe cumplirse que</p> $\phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1, \phi_2 < 1,$ <p>siendo</p> $\phi_1 = -0.39, \phi_2 = -0.24.$ <p><i>El modelo es estacionario</i></p>	<p>Debe cumplirse que</p> $ \theta_1 < 1$ <p>siendo</p> $\theta_1 = -0.71$ <p><i>El modelo es invertible</i></p>
ARI (2,1) x I (1) ₁₂	<p>Debe cumplirse que</p> $\phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1, \phi_2 < 1,$ <p>siendo</p> $\phi_1 = -0.40, \phi_2 = -0.24.$ <p><i>El modelo es estacionario</i></p>	<p>El modelo es un AR y una de las características de estos modelos es que siempre es invertible.</p>
IMA (1) x ARIMA (2,1,1) ₁₂	<p>Debe cumplirse que</p> $\phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1,$ $ \phi_2 < 1$ <p>siendo</p> $\phi_1 = 0, \phi_2 = -0.21.$ <p><i>El modelo es estacionario</i></p>	<p>Debe cumplirse que</p> $ \theta_1 < 1 \text{ y } \theta_1 < 1,$ <p>siendo</p> $\theta_1 = -0.64, \theta_1 = -0.68$ <p><i>El modelo es invertible</i></p>
ARI (1,1) x ARIMA(2,1,1) ₁₂	<p>Debe cumplirse que</p> $ \phi_1 < 1, \phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1,$ $ \phi_2 < 1,$ <p>siendo</p> $\phi_1 = -0.31,$ $\phi_1 = 0, \phi_2 = -0.21.$ <p><i>El modelo es estacionario</i></p>	<p>Debe cumplirse que</p> $ \theta_1 < 1,$ <p>siendo</p> $\theta_1 = -0.76,$ <p><i>El modelo es invertible</i></p>

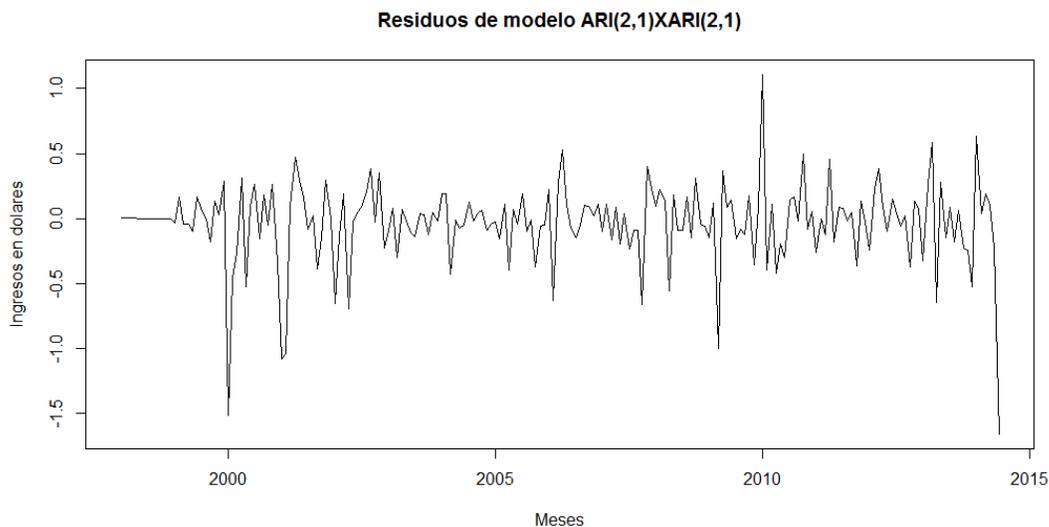
Se observa que los cinco modelos cumplen las condiciones de estacionariedad e invertibilidad. Entonces, para poder seleccionar el mejor modelo de ajuste se hace necesario el análisis de los residuos, esto es, verificar si se comportan como un ruido blanco (es decir si tienen un comportamiento normal).

4.5.1.3. Análisis de los residuos

Sometiendo los modelos al análisis de los residuos, en todos se tiene media alrededor de cero y la varianza constante, estas condiciones tiene mayor exactitud para el modelo ARI (2,1) x ARI (2,1)₁₂, de igual manera la condición de incorrelación, además de tener el mayor número de parámetros significativos.

Para analizar si la media de los residuos oscila alrededor de cero, se presentan gráficamente la serie de los residuos del modelo ARI (2,1) x ARI (2,1)₁₂.

Gráfico 12. Residuos de la serie de los ingresos económicos de los restaurantes. Modelo ARI (2,1) x ARI (2,1)₁₂. (Enero 1998 a Junio de 2014).



En el gráfico 12 se observa que la media está alrededor de cero, es decir que los residuos del modelo tienen un comportamiento constante a lo largo del tiempo; además se ve un ruido blanco con algunos outliers. Los gráficos de la FAS y la FAP de los residuos para el modelo ARI (2,1) x ARI (2,1)₁₂ se muestran a continuación:

Gráfico 13. FAS de los residuos. Modelo ARI (2,1) x ARI (2,1)₁₂. Serie de los ingresos económicos de los restaurantes. (Enero 1998 a Junio de 2014).

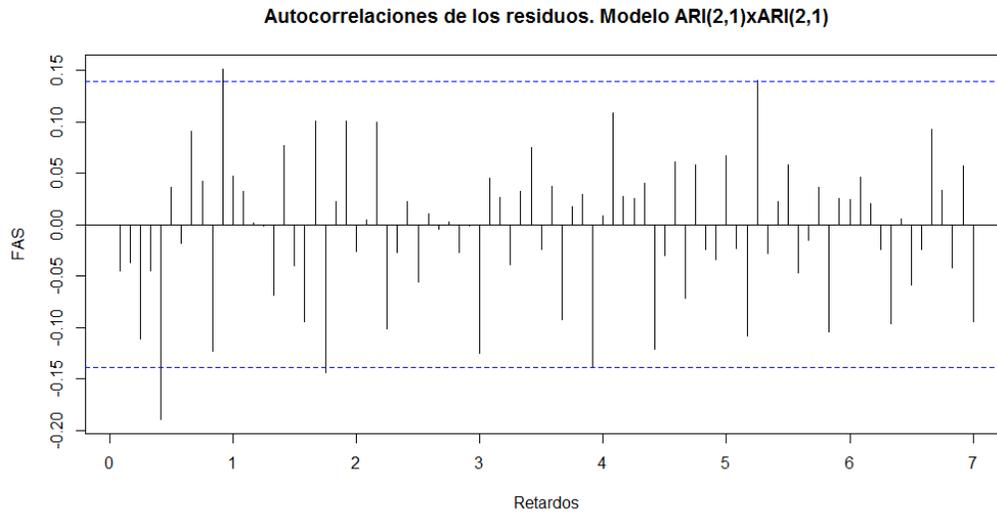
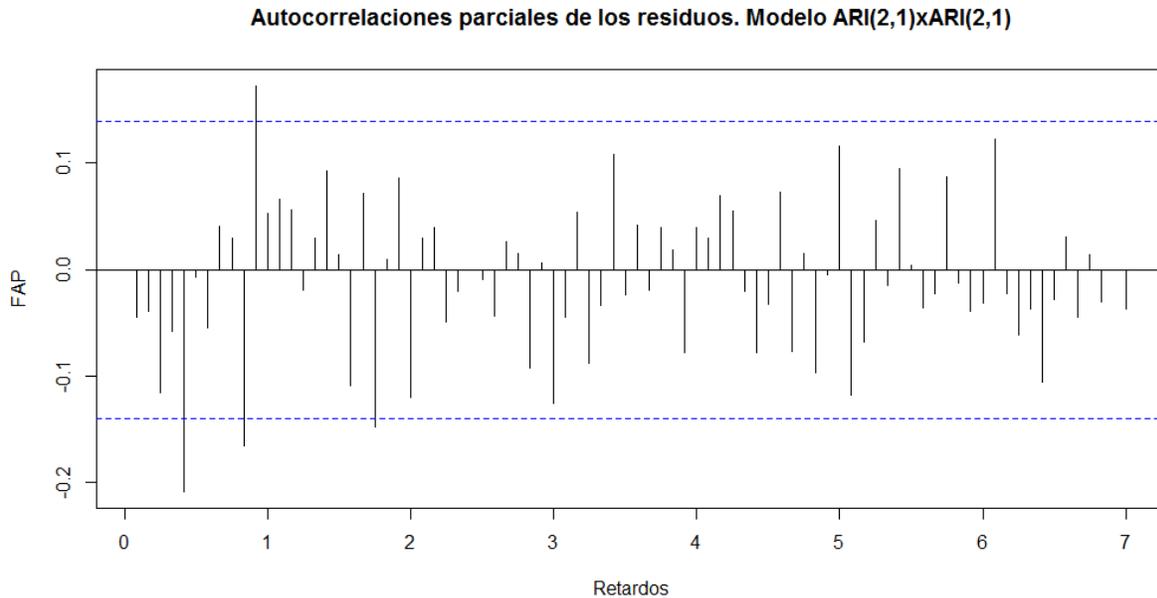


Gráfico 14. FAP de los residuos. Modelo ARI (2,1) x ARI (2,1)₁₂. Serie de los ingresos económicos de los restaurantes. (Enero 1998 a Junio de 2014).



En el gráfico 13 y 14 de las autocorrelaciones simple y parcial obedecen aproximadamente al intervalo de aceptación propuesto. Las autocorrelaciones de los residuos están incorrelados, es decir, se evidencia que las autocorrelaciones muestrales y parciales obedecen al intervalo de

aceptación, para afirmar que no son significativas, es decir próximas a cero consideradas a un nivel de significancia al 0.05, no hay ninguna relación entre los residuos. A continuación se muestran los gráficos para detectar la normalidad en los residuos.

Gráfico 15. Histograma de los residuos. Modelo ARI (2,1) x ARI (2,1)₁₂. Serie de los ingresos económicos de los restaurantes. (Enero 1998 a Junio de 2014).

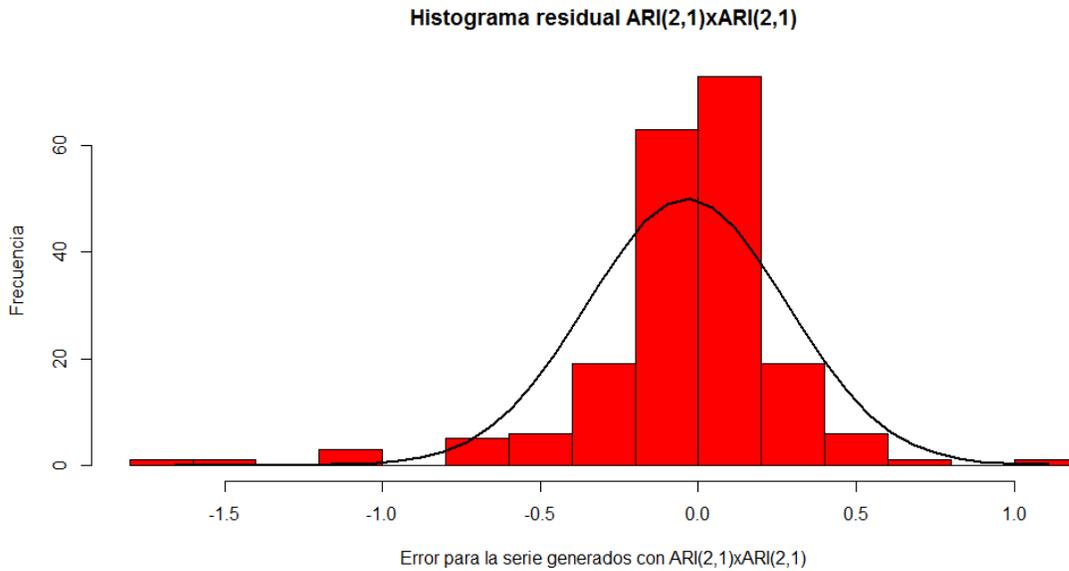
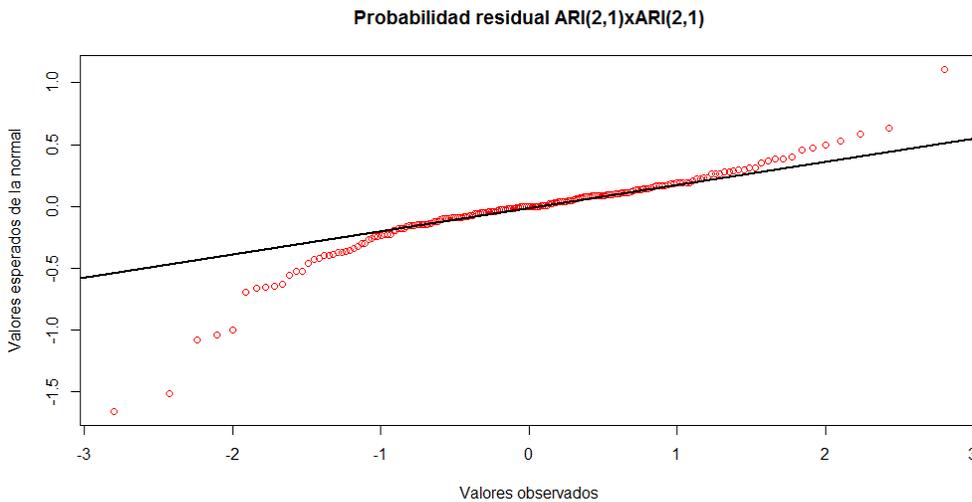


Gráfico 16. Probabilidad normal de los residuos. Serie de los ingresos económicos de los restaurantes. Modelo ARI (2,1) x ARI (2,1)₁₂



En el gráfico del histograma de los residuos y en el gráfico de probabilidad residual, se observa un comportamiento aproximadamente normal, pero existen algunos puntos extremos en las colas que pueden considerarse datos atípicos.

Al analizar los modelos con los estadísticos AIC (El Criterio de Información de Akaike) y BIC (El Criterio de Información Bayesiana), permite verificar en mayor solidez que se tenga un criterio más para identificar el modelo generador de la serie, ya que se ha planteado teóricamente que el modelo adecuado será aquel que cumple mayores condiciones y criterios.

Tabla 14. Modelos propuestos con los estadísticos AIC y BIC (serie ingresos económicos de los restaurantes)

Nº	Modelo	AIC	BIC
1	ARI (2,1) x ARI (2,1) ₁₂	127.91	144.01
2	ARI (2,1) x IMA(1,1) ₁₂	119.6	138.92
3	ARI (2,1) x I(1) ₁₂	121.41	143.96
4	IMA (1,1) x ARIMA(2,1,1) ₁₂	118.61	141.15
5	ARI (1,1) x ARIMA(2,1,1) ₁₂	126.54	142.64

Para el caso de nuestra investigación, el modelo que cumple con esta condición es el que ya se ha identificado en el análisis de los residuos, por lo tanto se acepta el modelo ARI (2,1) x ARI (2,1)₁₂ como el que más se ajusta al comportamiento de los datos de la serie de los ingresos mensuales de restaurantes, ya que este cumple con la mayoría de los criterios que nos conducen a la metodología de Box- Jenkins a pesar de tener un mayor AIC y BIC el modelo a pesar de tener un mayor AIC y BIC.

Una vez se ha identificado el mejor ARIMA (p,d,q) x (P,D,Q)₁₂ para el mayor subconjunto de los datos, el siguiente paso es el modelo de la intervención.

4.5.1.4. Análisis de intervención

Entonces, una vez se define el modelo ARIMA se realiza el análisis de intervención, es necesario identificar dos características: 1) el periodo de comienzo de las intervenciones y 2) la forma general del impacto. Como ya se ha mencionado anteriormente, que el período de comienzo de la intervención es en el mes de septiembre de 2008 ($t_0=129$). En cuanto a la forma general del impacto, en base al artículo “Crisis económica de 2008-2015”¹⁴ el cual menciona que la crisis financiera de Estados Unidos de América afecto a la económica mundial, no siendo la excepción El Salvador, se estudian las cuatro formas generales de modelo de intervención, debiendo observar que cuando la duración del efecto se considera permanente se debe incluir una variable escalón, mientras que cuando la duración del efecto se considere temporal la variable incluida será tipo impulso.

Para el análisis de la base de ingresos económicos de los restaurantes, se muestra el resultado del análisis con solo dos tipos de intervención (Comienzo brusco y duración temporal, comienzo gradual y duración temporal), ya que los parámetros asociados a estos modelos resultaron ser significativos. Respecto al modelo de intervención comienzo brusco y duración permanente, resultaron ser no significativos (anexo 7.3.1), por tanto se excluyen el análisis de estos modelos. El otro modelo de intervención comienzo gradual y duración permanente no se analiza por limitaciones en el proceso de optimización del software R (anexo 7.3.2).

4.5.1.4.1. Modelo de intervención: Comienzo brusco y duración temporal

El modelo que se puede identificar en este tipo de intervención es el siguiente:

$$y_t = \frac{\omega}{1 - \delta L} P_t^{(t_0=129)} + \exp \left(\frac{u_t}{(1 - \phi_1 L - \phi_1 L^2)(1 - \Phi_1 L^{12} - \Phi_2 L^{24})(1 - L^{12})(1 - L)} \right)$$

Se calcula la estimación de los parámetros (ver anexo 7.3.3) y se observan los significativos, haciendo notar que todos aportan al modelo. Por lo tanto el modelo conjuntamente estimado es:

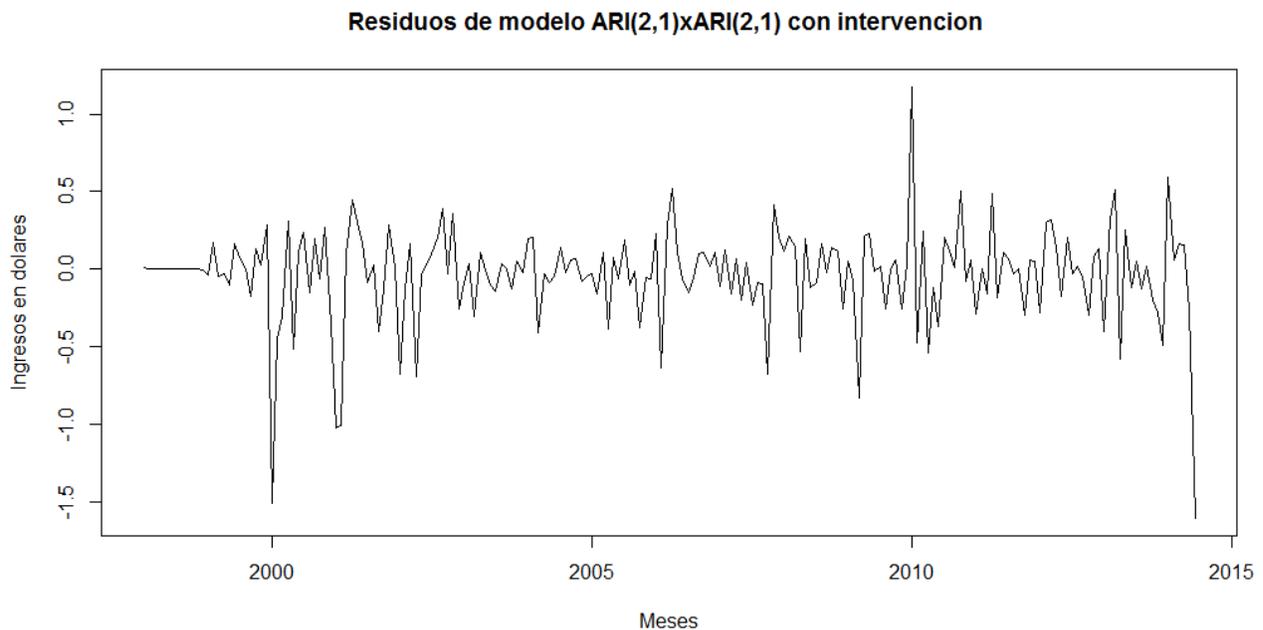
¹⁴ http://es.wikipedia.org/wiki/Crisis_econ%C3%B3mica_de_2008-2015, consultado el 12 de mayo de 2015.

$$y_t = \frac{-0.1285}{1 - 0.9711L} P_t^{(t_0=129)} + \exp\left(\frac{u_t}{(1 + 0.4478L + 0.3235L^2)(1 + 0.6446L^{12} + 0.4054L^{24})(1 - L^{12})(1 - L)}\right)$$

Análisis de los residuos del modelo ARIMA con intervención Comienzo brusco y duración temporal

Para llevar a cabo este análisis, a continuación se muestra el primer gráfico de este análisis:

Gráfico 17. Residuos del modelo ARI (2,1) x ARI (2,1)₁₂ con intervención Comienzo brusco y duración temporal. Serie de ingresos económicos de restaurantes.



En el gráfico 17 se observa que la media de los residuos esta en torno a cero, es decir que los residuos del modelo tienen un comportamiento constante a lo largo del tiempo.

Los gráficos de la FAS y la FAP de los residuos para el modelo ARI (2,1) x ARI (2,1)₁₂ con intervención comienzo brusco y duración temporal se muestran a continuación:

Gráfico 18. FAS de los residuos del modelo $ARI(2,1) \times ARI(2,1)_{12}$ con intervención Comienzo brusco y duración temporal. Serie de los ingresos económicos mensuales de los restaurantes.

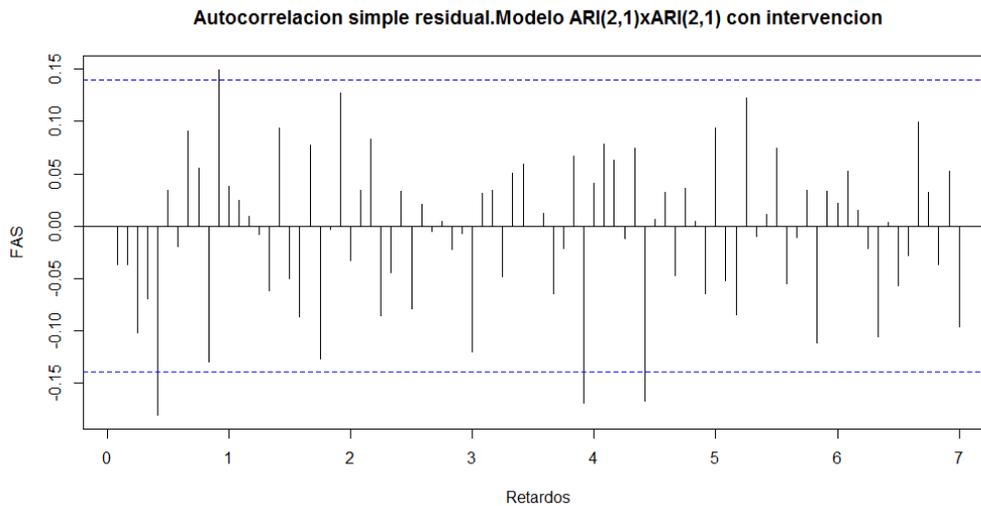
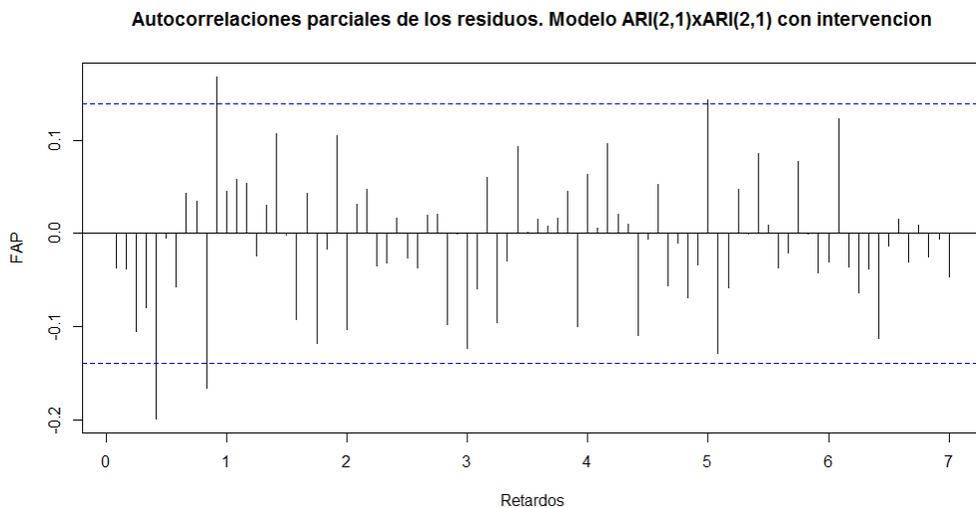


Gráfico 19. FAP de los residuos del modelo $ARI(2,1) \times ARI(2,1)_{12}$ con intervención Comienzo brusco y duración temporal. Serie de los ingresos económicos de los restaurantes.



En el gráfico 18 y 19 de la FAS y FAP de los residuos, obedecen aproximadamente al intervalo de aceptación propuesto. Las autocorrelaciones de los residuos están incorrelados, es decir no hay ninguna relación.

A continuación se presentan los gráficos: histograma residual y Probabilidad normal de los residuos para el modelo en estudio.

Gráfico 20. Histograma residual del modelo $ARI(2,1) \times ARI(2,1)_{12}$ con intervención Comienzo brusco y duración temporal. Serie de ingresos económicos de los restaurantes.

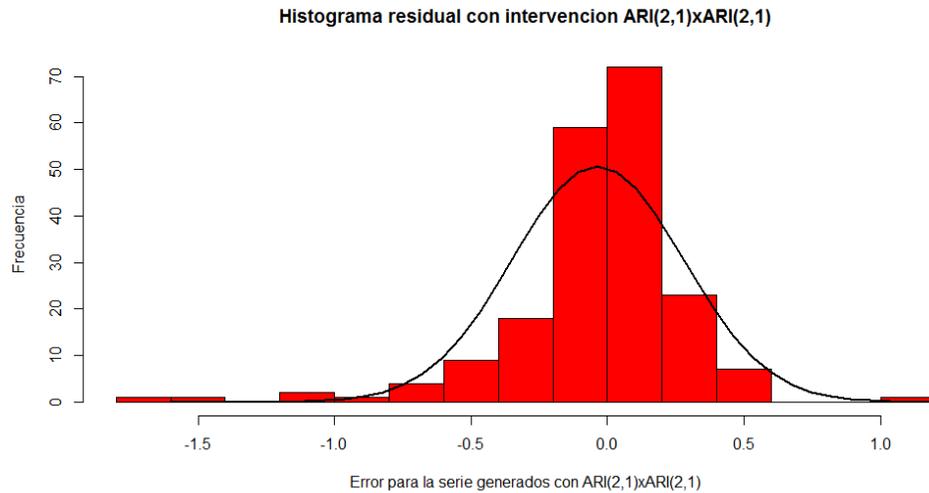
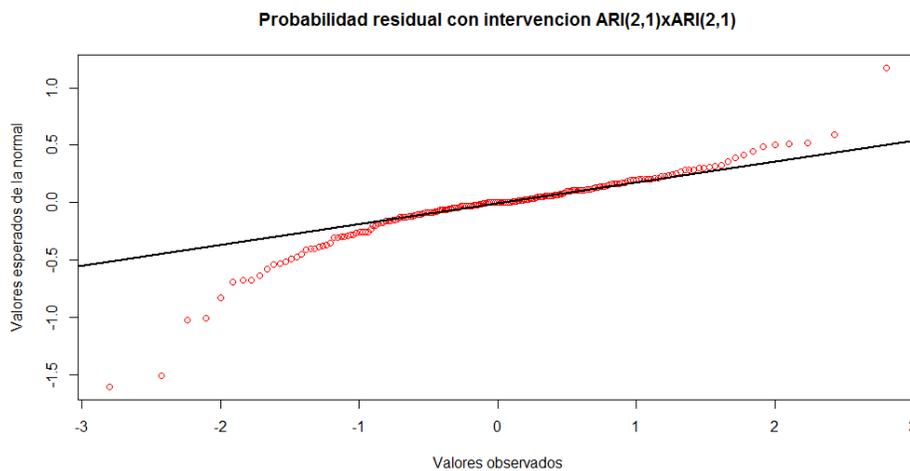


Gráfico 21. Probabilidad normal de los residuos. Modelo $ARI(2,1) \times ARI(2,1)_{12}$ con intervención Comienzo brusco y duración temporal. Serie de ingresos económicos de los restaurantes



En el gráfico 20 y 21, correspondientes al histograma de los residuos y gráfico de probabilidad residual, se observa un comportamiento aproximadamente normal, pero existen algunos puntos extremos en las colas que pueden considerarse datos atípicos.

4.5.1.4.2. Modelo con intervención: Comienzo gradual duración temporal

El modelo que se puede identificar en este tipo de intervención es el siguiente:

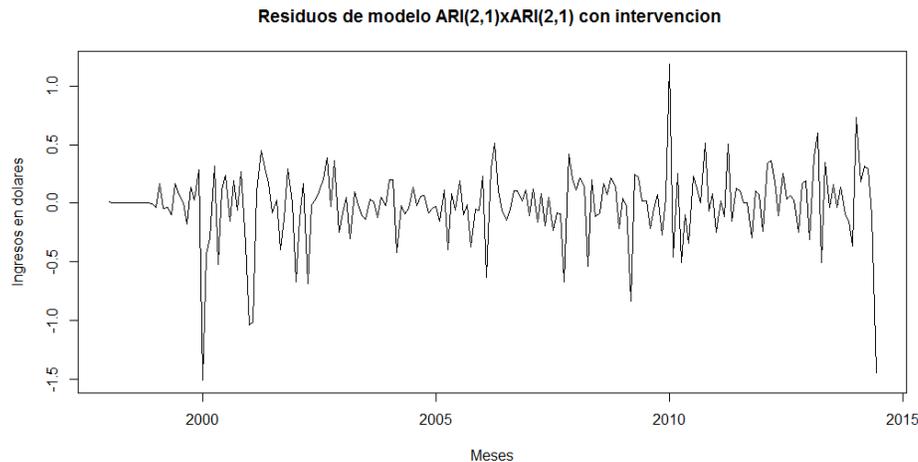
$$y_t = \frac{\omega}{1 - \delta_1 L - \delta_2 L^{12}} P_t^{(t_0=129)} + \exp \left[\frac{u_t}{(1 - \phi_1 L - \phi_1 L^2)(1 - \phi_1 L^{12} - \phi_2 L^{24})(1 - L^{12})(1 - L)} \right]$$

Se calcula la estimación de los parámetros (ver anexo 7.3.4) y se observan los significativos, haciendo notar que todos aportan al modelo. Por lo tanto el modelo conjuntamente estimado es:

$$y_t = \frac{-0.2294}{1 - 0.0740L - 1.0222L^{12}} P_t^{(t_0=129)} + \exp \left[\frac{u_t}{(1 + 0.4440L + 0.3220L^2)(1 + 0.6550L^{12} + 0.3999L^{24})(1 - L^{12})(1 - L)} \right]$$

Análisis de los residuos del modelo ARIMA con intervención Comienzo gradual duración temporal. Serie de Restaurantes.

Gráfico 22. Residuos del modelo ARI (2,1) x ARI (2,1)₁₂ con intervención Comienzo brusco y duración temporal comienzo gradual duración temporal. Serie de ingresos económicos de los restaurantes



En el gráfico 22 se observa que la media está alrededor de cero, es decir que los residuos del modelo tienen un comportamiento constante a lo largo del tiempo; además podemos ver un ruido blanco con algunos outliers. Los gráficos de la FAS y la FAP de los residuos para el modelo ARI

(2,1) x ARI (2,1)₁₂ con intervención comienzo gradual y duración temporal se muestran a continuación:

Gráfico 23. FAS del modelo ARI (2,1) x ARI (2,1)₁₂ con intervención comienzo gradual duración temporal. Serie de ingresos económicos de los restaurantes.

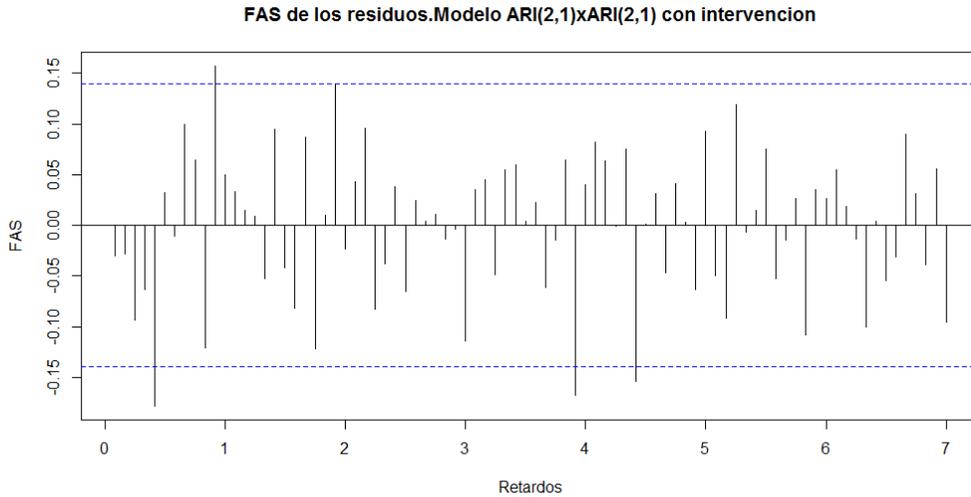
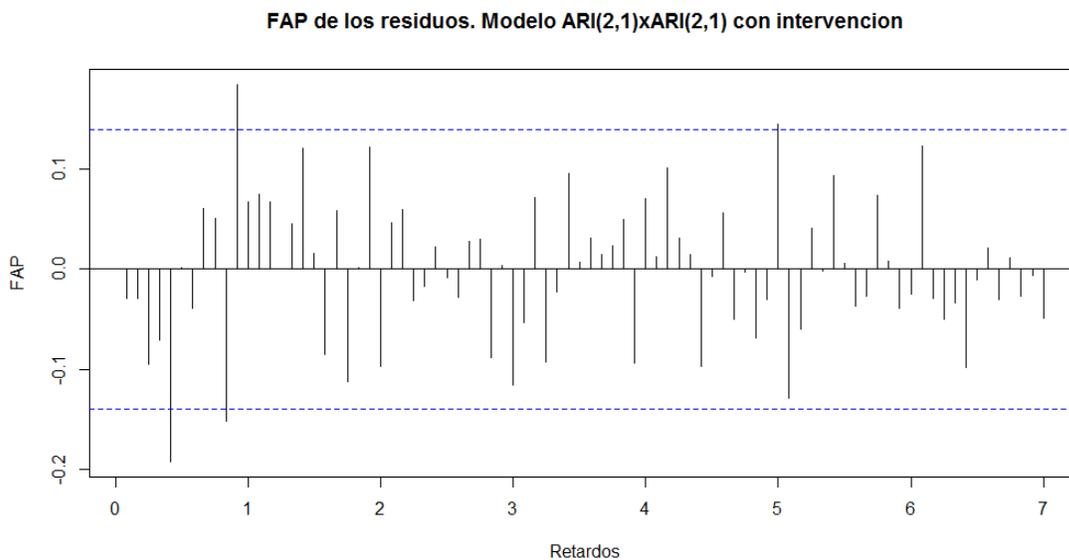


Gráfico 24. FAS de los residuos del modelo ARI (2,1) x ARI (2,1)₁₂ con intervención comienzo gradual duración temporal. Serie de ingresos económicos de los restaurantes.



En el grafico 23 y 24 de las autocorrelaciones simple y parcial de los residuos, obedecen aproximadamente al intervalo de aceptación propuesto. Las autocorrelaciones de los residuos

están incorrelados, es decir no hay ninguna relación. A continuación se muestran el histograma residual y el de probabilidad residual del mismo modelo de intervención.

Gráfico 25. Histograma residual del modelo ARI (2,1) x ARI (2,1)₁₂ con intervención comienzo gradual y duración temporal. Serie de ingresos económicos de los restaurantes.

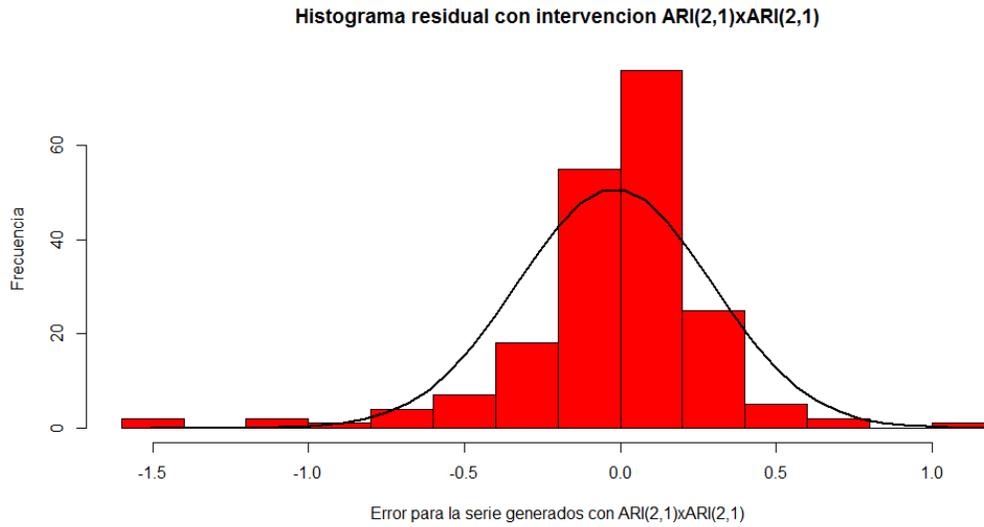
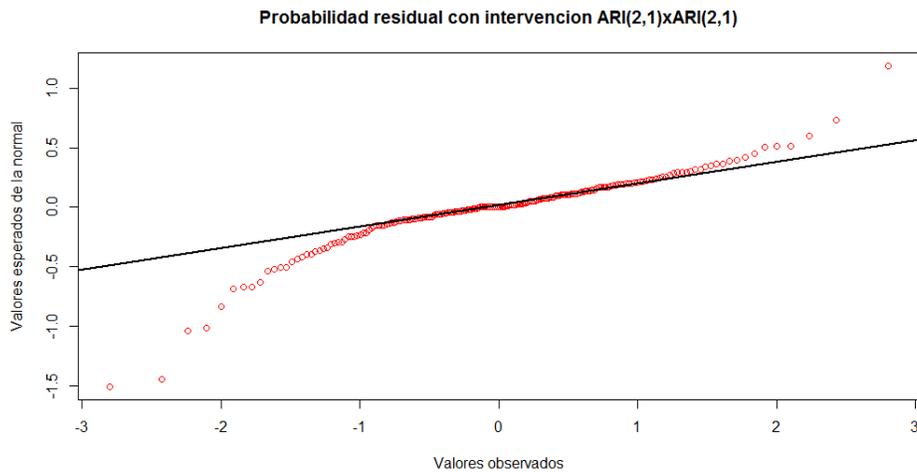


Gráfico 26. Probabilidad normal de los residuos. Modelo ARI (2,1) x ARI (2,1)₁₂ con intervención comienzo gradual y duración temporal. Serie de ingresos económicos de los restaurantes.

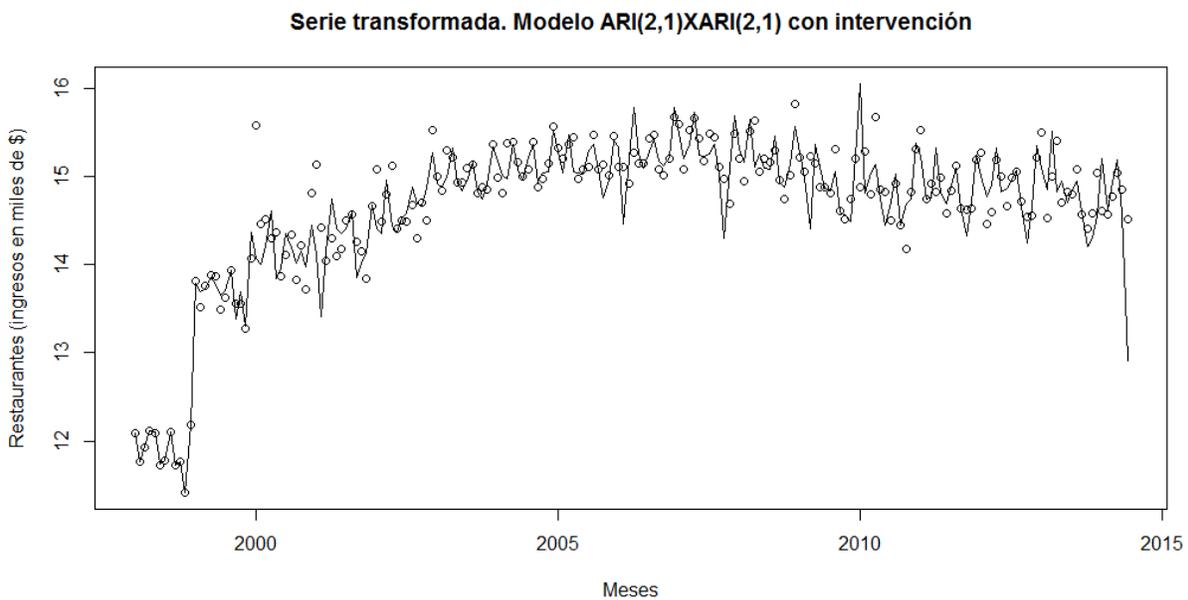


En el gráfico 25 y 26, correspondientes al histograma de los residuos y gráfico de probabilidad residual, se observa un comportamiento aproximadamente normal, pero existen algunos puntos extremos en las colas que pueden considerarse datos atípicos.

Se observa que ambos modelos de intervención cumplen las condiciones de incorrelación (media cero, varianza constante y covarianza cero en los residuos), entonces se elige el mejor modelo de intervención tomando en cuenta el principio de parsimonia, con lo que resulta el modelo $ARI(2,1) \times ARI(2,1)_{12}$ con intervención comienzo brusco y duración temporal.

Luego se presenta el gráfico de la serie de restaurantes transformada con la serie superpuesta del modelo $ARI(2,1) \times ARI(2,1)_{12}$ con intervención Comienzo brusco y duración temporal.

Gráfico 27. Serie de ingresos económicos de los restaurantes transformada y serie superpuesta del modelo $ARI(2,1) \times ARI(2,1)_{12}$ con intervención Comienzo brusco y duración temporal



En el gráfico 27 se observa que los puntos del modelo $ARI(2,1) \times ARI(2,1)_{12}$ con intervención comienzo brusco y duración temporal se ajusta bastante bien a la serie de datos transformadas, sin embargo, se tienen algunos puntos que discrepan con el modelo planteado, por lo que se realizará un análisis de outliers.

4.5.1.5. Análisis de Outliers.

Para mejorar la comprensión de la serie objeto de análisis, se detectan los valores atípicos del modelo residual, identificando el tipo de dato atípico, también se determina el valor residual y el dato real de la serie. Luego, se verifica si en la fecha que ocurrió el outlier se presentó el evento de veda por marea roja. A continuación se detallan estos resultados:

Tabla 15. Detección de outliers y comparación de residuos del modelo ARI (2,1) x ARI (2,1)₁₂ con intervención Comienzo brusco y duración temporal. Serie de Restaurantes.

Nº de casos	Tipo	Fecha	Valor de residuo	Dato real (ingresos en \$)	Presencia de Veda por marea roja
24	AO	Diciembre 1999	0.2888774	1,728,401.31	No
25	AO, TC, LS	Enero 2000	-1.512017	1,284,888.00	No
36	LS	Diciembre 2000	-0.3545282	1,896,099.77	No
37	TC, LS	Enero 2001	-1.026002	1,340,015.38	No
38	AO, TC	Febrero 2001	-1.009792	668,050.08	No
40	TC, LS	Abril 2001	0.4487686	2,550,315.77	No
98	AO	Febrero 2006	-0.6352143	1,922,167.31	Si
99	LS	Marzo 2006	0.2824988	3,993,824.46	Si
118	AO, TC	Octubre 2007	-0.6764983	1,618,856.92	No
135	AO, TC, LS	Marzo 2009	-0.8277968	1,811,366.38	No
145	AO, TC, LS	Enero 2010	1.17452	9,394,325.85	No
154	AO, TC	Octubre 2010	0.5067519	2,389,772.00	Si
184	AO	Abril 2013	-0.5798407	2,740,353.00	No
192	AO	Diciembre 2013	-0.4921321	2,097,408.85	No
193	TC	Enero 2014	0.5956615	4,040,213.38	No
198	AO, TC, LS	Junio 2014	-1.605897	403,842.85	No

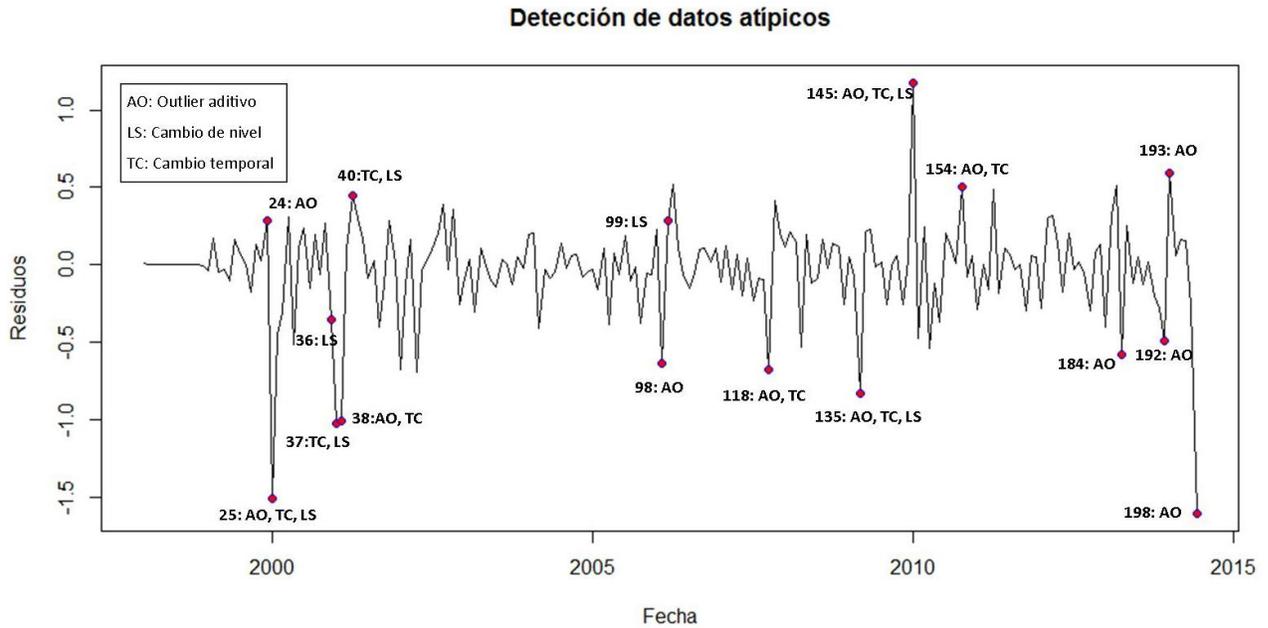
Los resultados del programa R nos muestran los 16 datos atípicos, se observa que 3 están asociados a veda por marea roja:

1. N° de caso 98, que corresponden a Febrero 2006 (valor del atípico bajo), en esta fecha según la tabla 1 del Marco Teórico, la veda por marea roja corresponde al periodo de Noviembre 2005 hasta Marzo de 2006, este atípico bajo está asociado a los cinco meses de veda que había hasta ese momento.
2. N° de caso 99, que corresponden a Marzo 2006 (valor del atípico alto), la veda por marea roja corresponde al periodo de Noviembre 2005 hasta Marzo de 2006, a pesar que había veda hubieron ingresos altos en los restaurantes, los turistas asisten a los restaurantes a consumir alimentos provenientes del mar que no hayan sido prohibidos por la veda u optan por consumir otros platos alternativos.
3. N° de caso 154, que corresponden a octubre 2010 (valor del atípico alto), la veda por marea roja corresponde a octubre de 2010 (una semana de duración), los ingresos no fueron afectados por el fenómeno natural posiblemente por el periodo corto de duración de veda por marea roja.

Respecto al resto de los demás datos atípicos que no están asociados con veda por marea roja posiblemente son debido a otros factores; en caso de ser outliers bajos, generalmente se nota que es un mes antes o un mes después del período vacacional y en caso de ser outliers altos es porque son días festivos o períodos vacacionales (semana santa, agosto y diciembre).

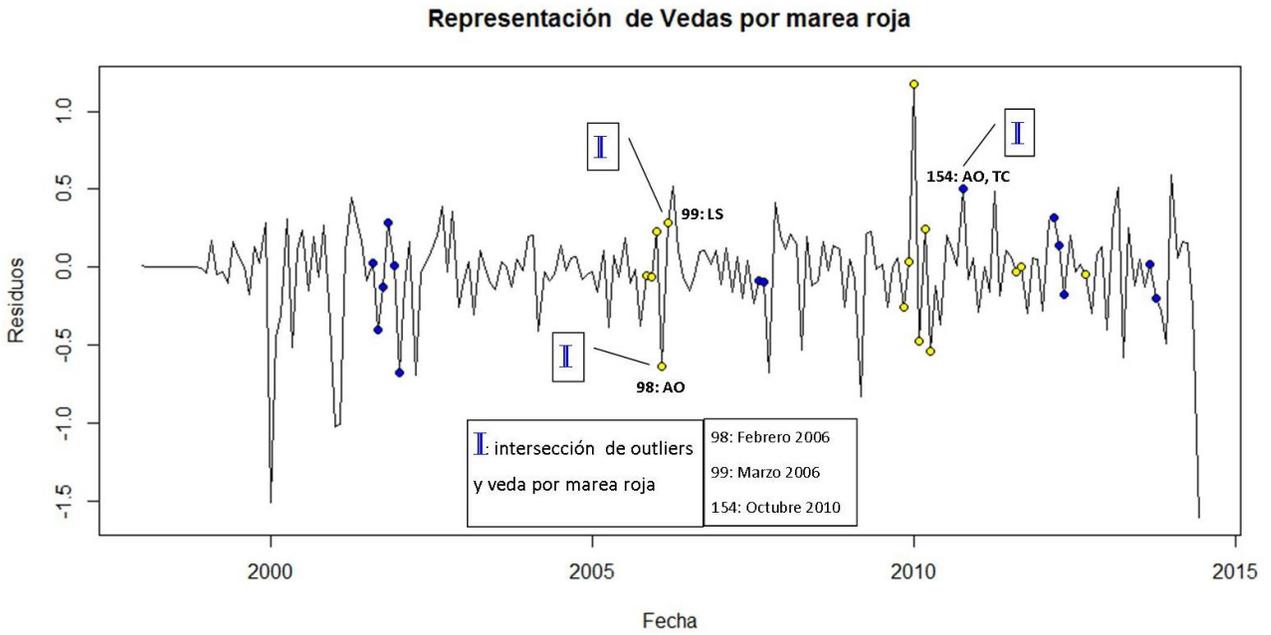
A continuación se muestra la asociación por medio del gráfico de los residuos del modelo con intervención, los datos considerados como atípicos se han representado por puntos de color rojo, para los cuales ya se ha realizado un análisis descriptivo.

Gráfico 28. Detección de datos atípicos en los residuos del modelo $ARI(2,1) \times ARI(2,1)_{12}$ con intervención Comienzo brusco y duración temporal. Serie de ingresos mensuales de los restaurantes.



En el gráfico 28, se detectan 16 datos atípicos, en los que se señala el tipo de outlier al que corresponde cada uno, estos ya fueron descritos anteriormente, sin embargo, es importante observar cuales de estos atípicos se interceptan con Veda por marea roja, estos se muestran en el gráfico siguiente:

Gráfico 29. Detección de vedas por marea roja en los residuos del modelo ARI (2,1) x ARI (2,1)₁₂ con intervención comienzo brusco y duración temporal. Serie de ingresos mensuales de los restaurantes.

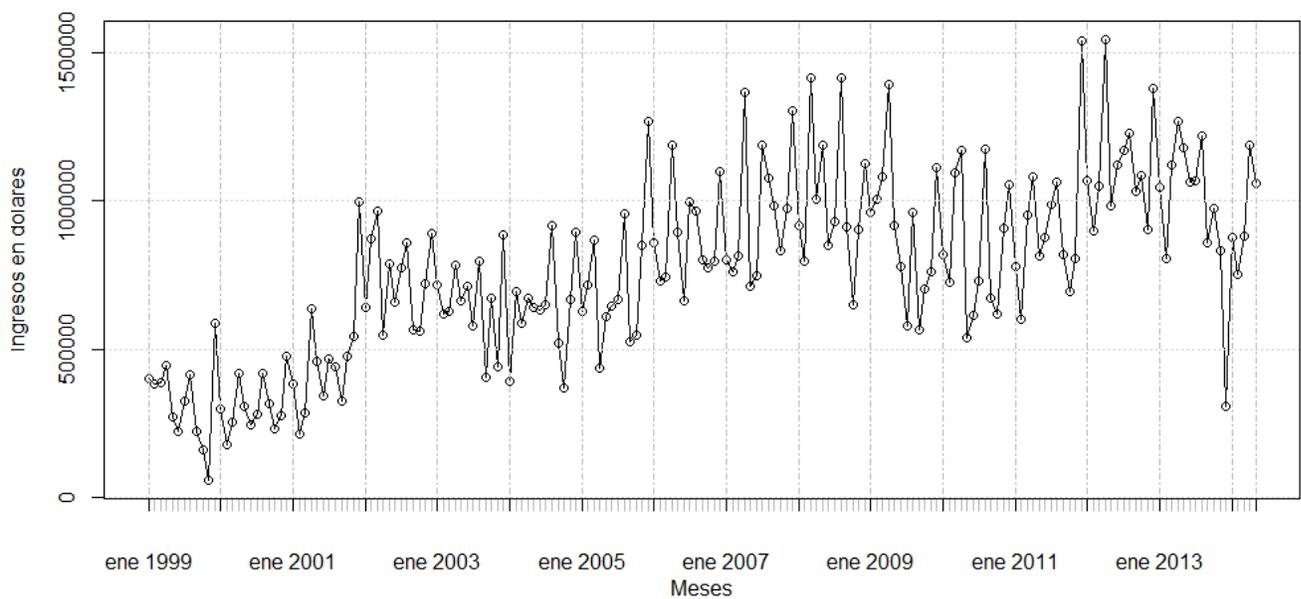


En el gráfico 29 se detectan 9 mareas rojas por vedas en diferentes fechas, y tres de los datos atípicos que se detectaron anteriormente coinciden con meses donde hay presencia de un evento de veda por marea roja, estos también son señalizados en el gráfico.

4.5.2. Análisis descriptivo de la serie temporal de Ingresos mensuales en hoteles del municipio de La Libertad.

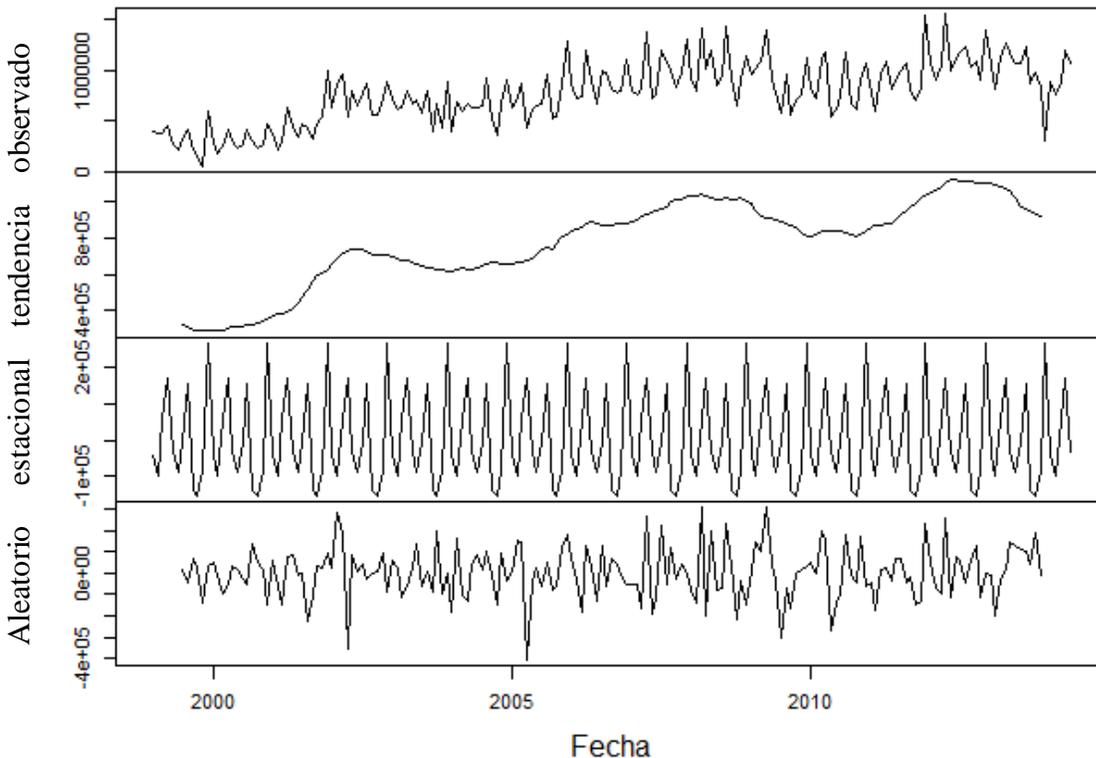
La serie temporal de Ingresos en dólares de hoteles del municipio de La Libertad, consta de 185 datos con periodicidad mensual desde Enero de 1999 a Mayo de 2014. Se inicia el análisis de la serie, observando si la serie es estacionaria, es decir que la media, la varianza y las autocorrelaciones no dependan del tiempo. A continuación se muestra el gráfico 30 de secuencia de la serie original.

Gráfico 30. Serie original. Ingresos económicos mensuales en dólares en los Hoteles del municipio de La Libertad. (Enero 1999- Mayo 2014)



Seguidamente, se realiza la descomposición de la serie original, esta se muestra a continuación en el gráfico 31:

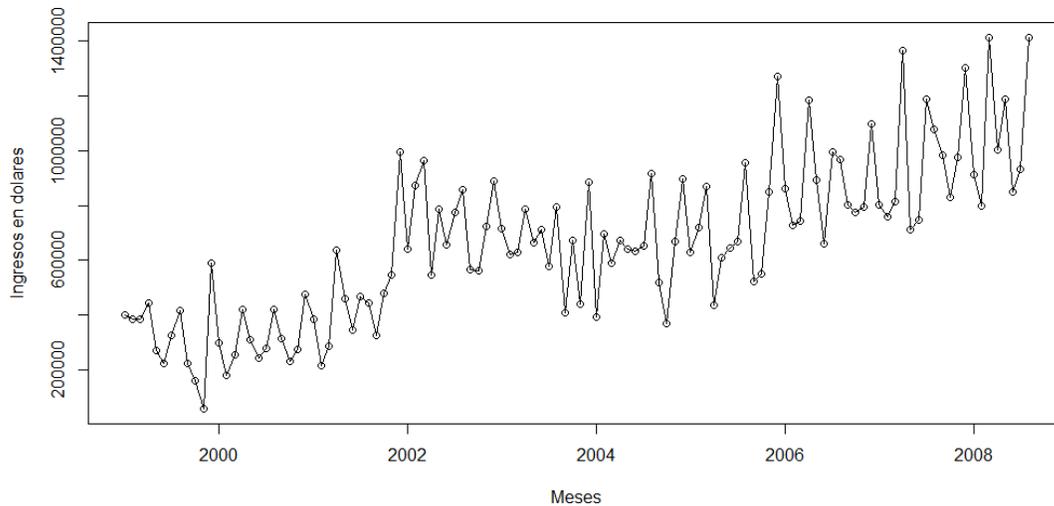
Gráfico 31. Descomposición de la serie de datos de los ingresos mensuales en los hoteles del municipio de La Libertad. (Enero 1999- Mayo 2014)



La componente tendencia representa la evolución a largo plazo de los ingresos, en la cual se observa una tendencia creciente. Sin embargo, a partir de los valores de la serie (ver anexo 7.4) se nota que existe un comportamiento irregular a partir de septiembre de 2008 ($t_0=117$), la cual se podría tratar de una intervención, esta posiblemente se debe a la crisis económica mundial. En la componente estacional se evidencian tres picos dentro de cada año, los primeros dos picos corresponden a los periodos vacacionales de Abril y Agosto, y el tercer pico con mayor valor a Diciembre, la componente aleatoria muestra lo que no es explicado por las demás componentes.

Para identificar el modelo ARIMA que más se aproxima a los datos, se toma el subconjunto de datos anterior o posterior de la intervención que sea más representativo (mayor que el 60%). Para los datos de los ingresos de hoteles, se tiene el 62.7% (116 datos) antes de la intervención y 37.3% (69 datos) después de la intervención, se analizan los datos antes de la intervención, ya que esta supera el 60% de la serie completa, según los criterios de Peña (ver marco teórico).

Gráfico 32. Serie de los ingresos económicos mensuales de los hoteles, antes de la intervención (Enero de 1999 – Agosto de 2008)



Observando el gráfico anterior (gráfico 32) se observa que presenta estacionariedad en varianza pero no en media, esto puede verse en los gráficos 33 y 34 de la Función de Autocorrelación Simple (FAS) y en la Función de Autocorrelación Parcial (FAP) de la serie antes de la intervención.

Gráfico 33. FAS. Serie de los ingresos económicos mensuales de los hoteles, pre intervención (Enero 1999 - Agosto 2008)

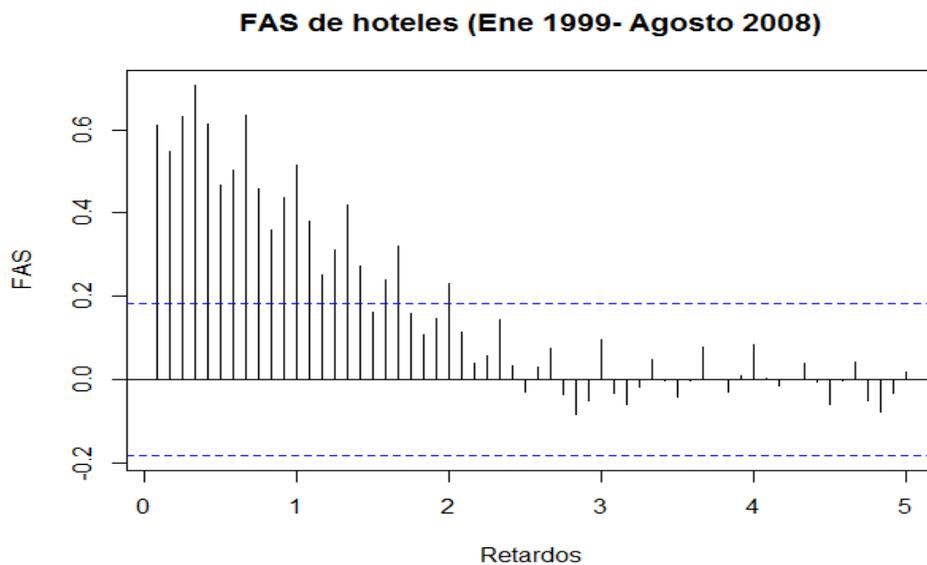
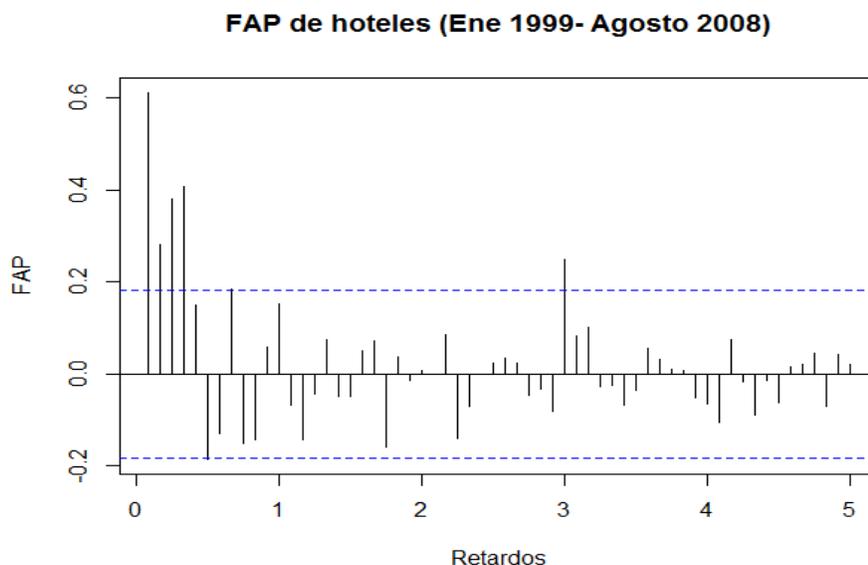


Gráfico 34. FAP. Serie de los ingresos económicos mensuales de los hoteles, pre intervención (Enero 1999- Agosto 2008)



En los gráficos anteriores, se observa que los correlogramas decaen lentamente a cero, es decir, la serie no es estacionaria. Para comprobar la estacionariedad en varianza y en media, se analiza el test de Levene¹⁵ y el test de Dickey Fuller respectivamente.

Para verificar la estacionariedad en varianza, se utiliza el test de Levene. En este caso, la serie se dividió en 9 submuestras, tomando un grupo por año. El test se obtiene aplicando el Software Estadístico R, obteniendo como resultado:

Tabla 16. Test de Levene para Homogeneidad de Varianza (Serie de los ingresos económicos mensuales de los hoteles, pre intervención)

Levene's Test for Homogeneity of Variance (center = median)			
	Df	F value	Pr(>F)
group	9	1.1779	0.3167
	106		

En el resultado anterior se observa que el estadístico de Levene con valor de 1.1779 tiene un p-valor de 0.3167, a un nivel de significancia de $\alpha=0.05$ no se rechaza la hipótesis nula de la igualdad de varianzas de las 9 submuestras y se concluye que la serie es estacionaria en varianza, por lo tanto no es necesario transformar la serie. En seguida se realiza el análisis de

¹⁵ Las Hipótesis del Test de Levene se plantean en el caso práctico de la serie de Restaurantes

estacionariedad en media, para esto se procede a realizar el test de raíces unitarias de Dickey-Fuller (DF)¹⁶. Se realiza el contraste de hipótesis y se obtienen los resultados que se muestran a continuación:

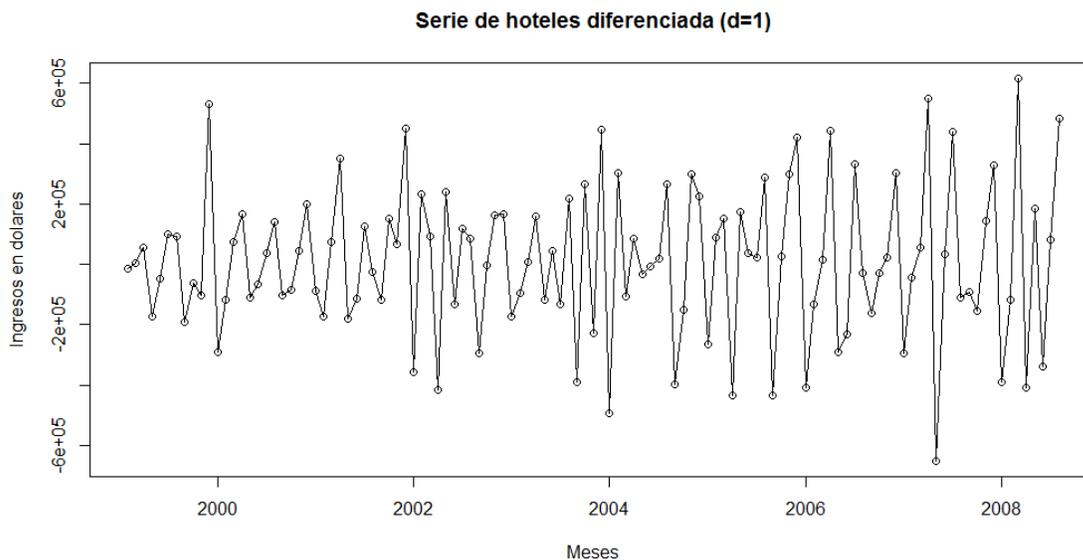
Tabla 17. Test de Dickey- Fuller para Homogeneidad de Medias (Serie de los ingresos económicos mensuales de los hoteles, pre intervención)

Augmented Dickey-Fuller Test

```
data: pre.int.hoteles
Dickey-Fuller = -2.4258, Lag order = 4, p-value = 0.3998
alternative hypothesis: stationary
```

En los resultados anteriores (Tabla 17), se observa que el estadístico de Dickey Fuller con valor de -2.4258 tiene un p-valor de 0.3998 y a un nivel de significación de $\alpha=0.05$ no se rechaza la hipótesis nula de la existencia de una raíz unitaria, por lo tanto, la serie no es estacionaria en media. Para lograrlo, se aplica una diferencia en la parte regular, esto se muestra en el siguiente gráfico.

Gráfico 35. Serie de los ingresos económicos mensuales de los hoteles, pre intervención con una diferencia regular ($d=1$)



¹⁶ Las Hipótesis del Test de Dickey Fuller se plantean en el caso práctico de la serie de Restaurantes

Según el gráfico 35, a simple vista se observa que la serie es estacionaria en media, sin embargo, se aplica una vez más el test para comprobarlo si la serie es estacionaria en media con una diferencia regular, los resultados son los siguientes:

Tabla 18. Test de Dickey- Fuller para Homogeneidad de Medias (Serie de los ingresos económicos mensuales de los hoteles, pre intervención con una diferencia en la parte regular)

```
Augmented Dickey-Fuller Test
data: diff.pre.int.hoteles
Dickey-Fuller = -7.2106, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary
```

En la tabla 18 se observa que el estadístico de Dickey Fuller con valor de -7.2106 tiene un p-valor de 0.01, por lo que a un nivel de significancia de $\alpha=0.05$ se rechaza la hipótesis nula de la existencia de una raíz unitaria y por lo tanto, la serie es estacionaria en media en la parte regular. A continuación se presentan los correlogramas con $d=1$ de la serie de los ingresos pre intervención para identificar los posible modelos.

Gráfico 36. FAS de la serie de los ingresos económicos mensuales de los hoteles pre intervención con una diferencia en la parte regular.

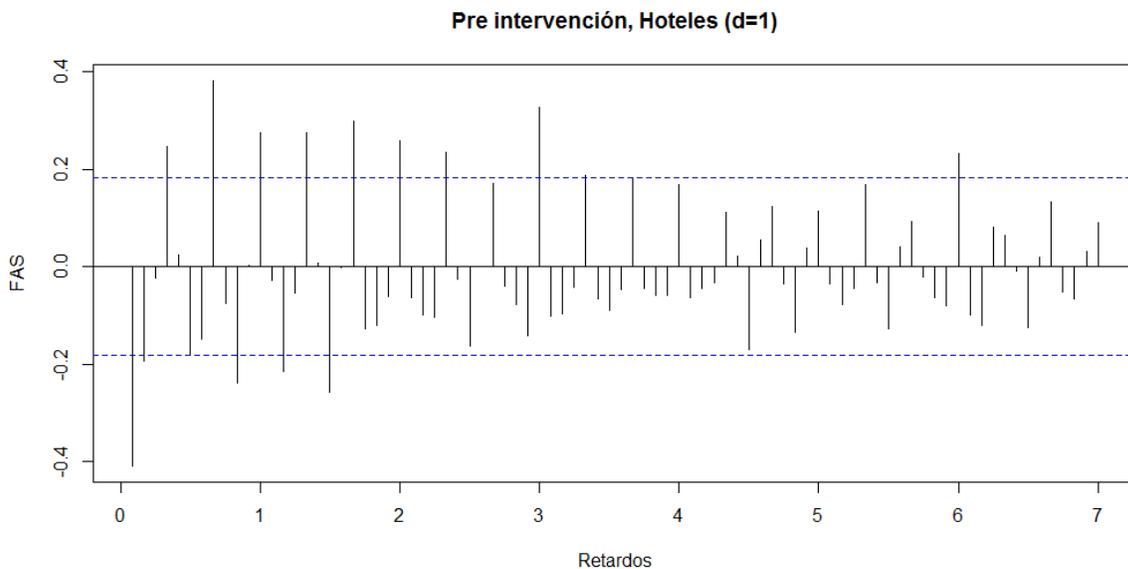
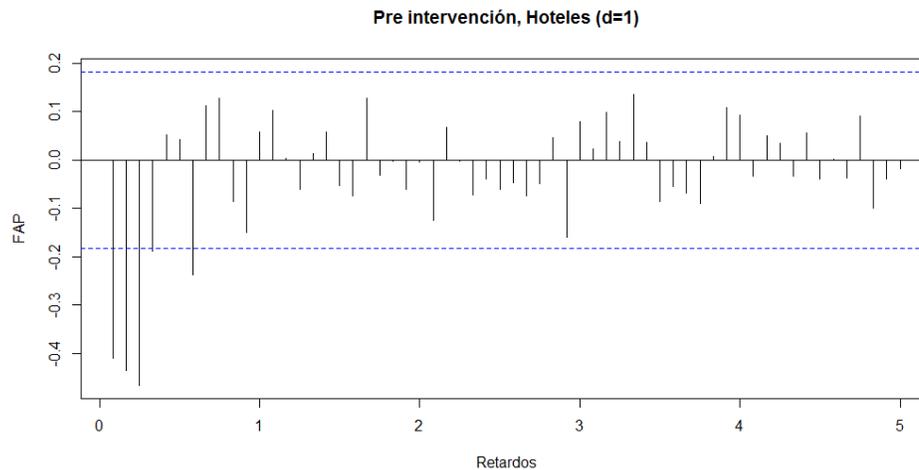


Gráfico 37. FAP de la serie de los ingresos económicos mensuales de los hoteles pre intervención con una diferencia en la parte regular



En los correlogramas anteriores (gráfico 36 y gráfico 37), se observan muchos retardos estacionales significativos ($s=1, s=2, \dots$) y una tendencia lentamente a anularse de dichos retardos, por lo que es necesario diferenciar también en la parte estacional.

Seguidamente se presentan los gráficos de la FAS y la FAP, en este caso se presentan los correlogramas de la serie con dos diferencias: una diferencia en la parte regular y una en la parte estacional, esto con el objetivo de que se cumplan los supuestos para el ajuste de los modelos ARIMA (ya mencionados en análisis anteriores).

A continuación se presentan los correlogramas con una diferenciación en la parte regular y una diferencia en la parte estacional.

Gráfico 38. FAS de la serie de los ingresos económicos mensuales de los hoteles. Serie pre intervención, diferenciada en la parte regular ($d=1$) y en la parte estacional ($D=1$)

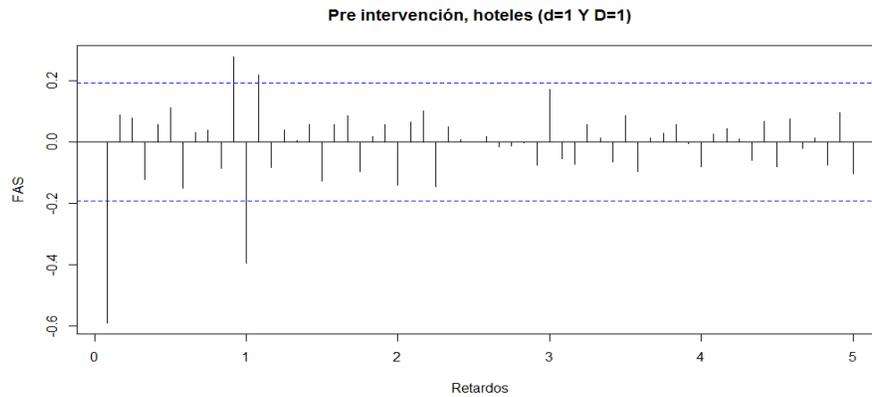
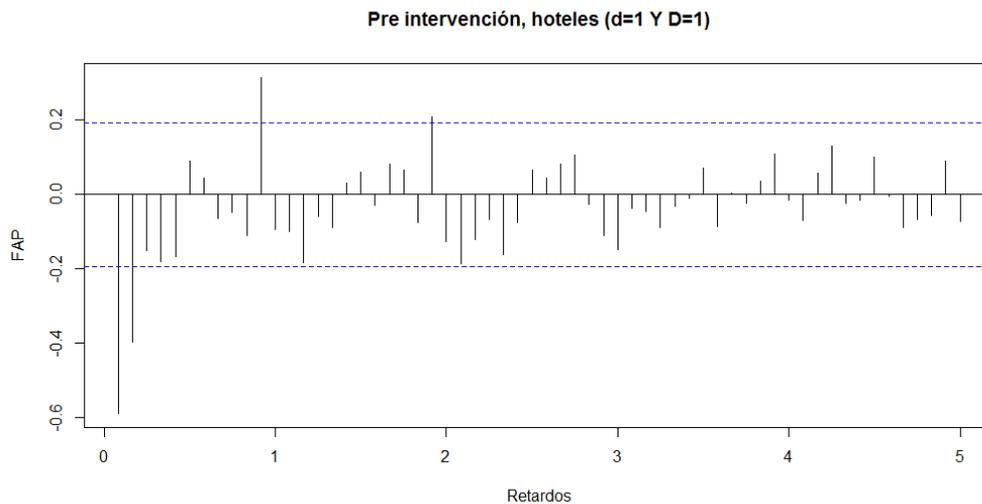


Gráfico 39. FAP de la serie de los ingresos económicos mensuales de los hoteles, serie pre intervención, diferenciada en la parte regular ($d=1$) y en la parte estacional ($D=1$)



En la gráfica 38 de la FAS, se observa que existen al menos 4 retardos significativos y que existe una tendencia a anularse rápidamente. Por otro lado en la gráfica 39 de la FAP se ven también al menos 4 retardos significativos con una tendencia rápida a anularse. Entonces por la interpretación anterior se podría pensar que el modelo buscado es del tipo:

$$\text{ARIMA}(p,d,q) \times \text{ARIMA}(P,D,Q)_s$$

Luego de observar los gráficos de la FAS y la FAP de la serie pre-intervención diferenciada ($d=1$ y $D=1$), se identifican de los posibles modelos, los cuales serán examinados en la etapa de estimación tal como los propone Box- Jenkins. Estos modelos propuestos son presentados a continuación.

4.5.2.1. Identificación del modelo.

Los modelos identificados son:

1. ARIMA (2,1,1) x ARIMA(2,1,1)₁₂

$$(1 - \phi_1 L - \phi_2 L^2)(1 - \Phi_1 L^{12} - \Phi_2 L^{24})(1 - L)(1 - L^{12})y_t = (1 - \theta_1 L)(1 - \theta_1 L^{12})u_t$$
2. ARIMA (2,1,1) x ARIMA(2,1,2)₁₂

$$(1 - \phi_1 L - \phi_2 L^2)(1 - \Phi_1 L^{12} - \Phi_2 L^{24})(1 - L)(1 - L^{12})y_t = (1 - \theta_1 L)(1 - \theta_1 L^{12} - \theta_2 L^{24})u_t$$
3. ARIMA (2,1,1) x ARIMA(1,1,1)₁₂

$$(1 - \phi_1 L - \phi_2 L^2)(1 - \Phi_1 L^{12})(1 - L)(1 - L^{12})y_t = (1 - \theta_1 L)(1 - \theta_1 L^{12})u_t$$
4. ARIMA (0,1,1) x ARIMA(2,1,0)₁₂

$$(1 - \Phi_1 L^{12} - \Phi_2 L^{24})(1 - L)(1 - L^{12})y_t = (1 - \theta_1 L)u_t$$

4.5.2.2. Estimación de los parámetros

Una vez identificados los posibles modelos y observado la significancia de los parámetros de los modelos identificados (anexo 7.5), se pueden expresar los siguientes modelos reducidos junto con sus respectivas expresiones, excluyendo los que no tuvieron parámetros significativos. Los modelos se muestran a continuación:

1. IMA (1,1) x ARIMA(2,1,1)₁₂

$$(1 + 0.2066L^{24})(1 - L)(1 - L^{12})y_t = (1 + 0.6907L)(1 + 0.7477L^{12})u_t ; \text{ con } \Phi_1 = 0$$
2. IMA (1,1) x ARIMA(2,1,2)₁₂

$$(1 + 0.5743 + 0.1948L^{12})(1 - L)(1 - L^{12})y_t = (1 + 0.6736L)(1 + 0.4405L^{12})u_t ; \text{ con } \theta_1 = 0$$
3. IMA (1,1) x IMA(1,1)₁₂

$$(1 - L)(1 - L^{12})y_t = (1 + 0.6636L)(1 + 0.8708L^{12})u_t$$
4. IMA (1,1) x ARI(2,1)₁₂

$$(1 + 0.6253 L^{12} + 0.4507 L^{24})(1 - L)(1 - L^{12})y_t = (1 + 0.7384 L)u_t$$

En base a los modelos mostrados anteriormente se hace una verificación de coeficientes para estacionariedad e invertibilidad.

Tabla 19. Análisis de estacionariedad e invertibilidad de los modelos identificados reducidos. Serie de los ingresos económicos mensuales de los hoteles.

Modelo	Estacionariedad	Invertibilidad
IMA(1,1) x ARIMA(2,1,1) ₁₂	Debe cumplirse que $\Phi_1 + \Phi_2 < 1, \Phi_2 - \Phi_1 < 1, \Phi_2 < 1$ <i>siendo</i> $\Phi_1 = 0, \Phi_2 = 0.2066.$ El modelo es estacionario	Debe cumplirse que $ \theta_1 < 1, \theta_1 < 1$ <i>siendo</i> $\theta_1 = -0.6907, \theta_1 = 0.7477.$ El modelo es invertible
IMA(1,1) x ARIMA(2,1,2) ₁₂	Debe cumplirse que $\Phi_1 + \Phi_2 < 1, \Phi_2 - \Phi_1 < 1, \Phi_2 < 1$ <i>siendo</i> $\Phi_1 = 0.5743, \Phi_2 = -0.1948.$ El modelo es estacionario	Debe cumplirse que $ \theta_1 < 1,$ $\theta_1 + \theta_2 < 1, \theta_2 - \theta_1 < 1, \theta_2 < 1$ <i>siendo</i> $\theta_1 = -0.6736, \theta_1 = 0, \theta_2 = -0.4405$ El modelo es invertible
IMA (1,1) x IMA(1,1) ₁₂	El modelo es un MA y una de las características de estos modelos es que siempre es Estacionario	Debe cumplirse que $ \theta_1 < 1, \theta_1 < 1$ <i>siendo</i> $\theta_1 = -0.6636, \theta_1 = -0.8708.$ El modelo es invertible
IMA (1,1) x ARI(2,1) ₁₂	Debe cumplirse que $\Phi_1 + \Phi_2 < 1, \Phi_2 - \Phi_1 < 1, \Phi_2 < 1$ <i>siendo</i> $\Phi_1 = -0.6253, \Phi_2 = 0.4507.$ El modelo es estacionario	Debe cumplirse que $ \theta_1 < 1,$ <i>siendo</i> $\theta_1 = -0.7384$ El modelo es invertible

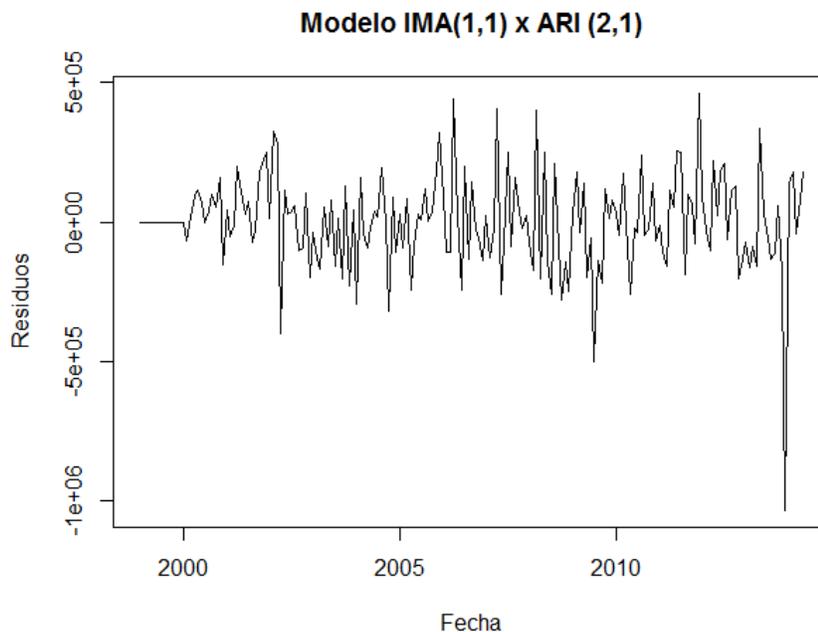
Se observa que los cuatro modelos anteriores cumplen las condiciones de estacionariedad e invertibilidad. Entonces para poder seleccionar el mejor modelo de ajuste, se hace necesario el análisis de los residuos, esto es, verificar si se comportan como un ruido blanco (es decir si tienen un comportamiento normal).

4.5.2.3. Análisis de los residuos

Sometiendo los modelos al análisis de los residuos, en todos se tiene media cero y la varianza constante, esta condición tiene mayor exactitud para el modelo IMA (1,1) x ARI (2,1)₁₂, de igual manera la condición de incorrelación, así como también el que tiene mayor número de parámetros significativos.

Se representan gráficamente la serie de los residuos del modelo con intervención y a partir de este gráfico se puede ver si los valores de los residuos oscilan alrededor de una media cero.

Gráfico 40. Residuos del Modelo IMA (1,1) x ARI (2,1)₁₂. Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014.



Observando el gráfico anterior (gráfico 40), se observa que la media está en torno a cero ($\mu = 0$), es decir que los residuos del modelo tienen un comportamiento constante a lo largo del tiempo. Los gráficos de la FAS y la FAP de los residuos para el modelo IMA (1,1), x ARI (2,1)₁₂ se muestran en seguida:

Gráfico 41. FAS Residual del Modelo IMA (1,1) x ARI (2,1). Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014

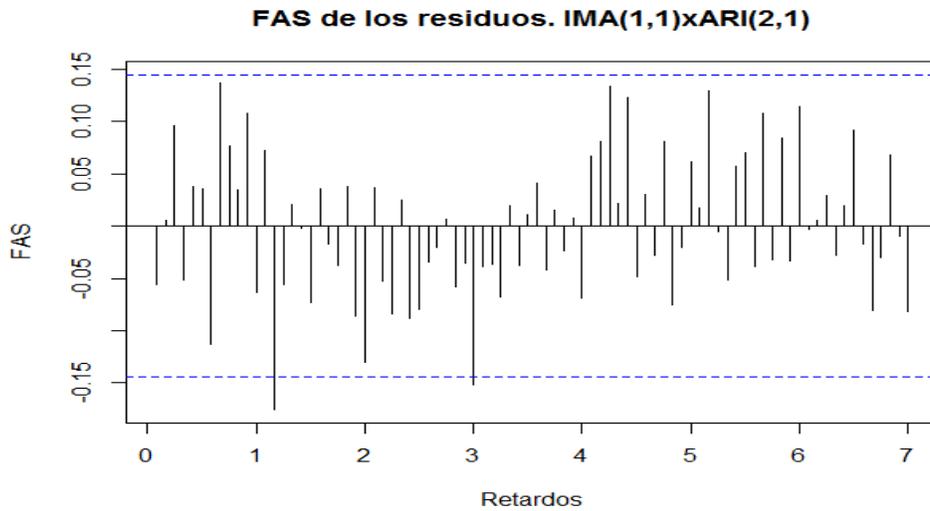
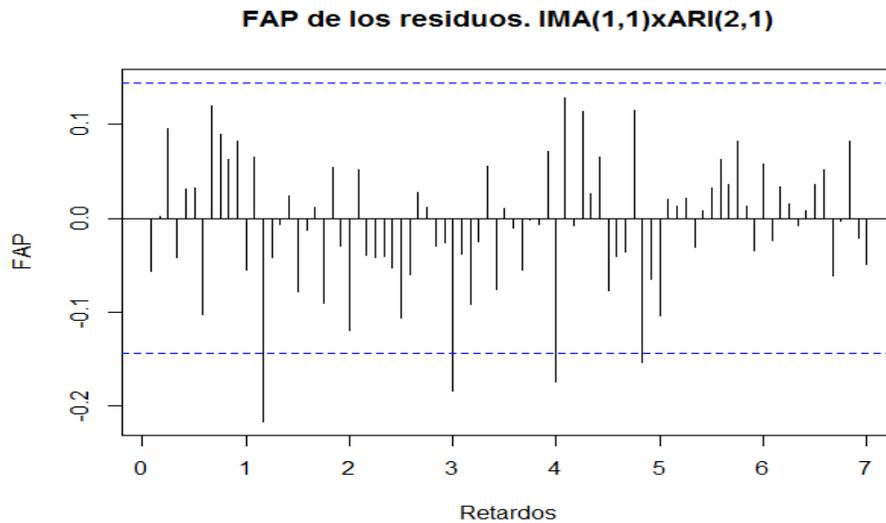


Gráfico 42. FAP Residual del Modelo IMA (1,1) x ARI (2,1)₁₂. Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014



Los dos gráficos anteriores obedecen aproximadamente al intervalo de aceptación propuesto. Las autocorrelaciones de los residuos son no significativas o que están muy próximas a cero para un nivel de significación $\alpha = 0.05$ y se acepta que no existe correlación entre los residuos del modelo.

Gráfico 43. Histograma Residual del Modelo IMA (1,1) x ARI (2,1)₁₂. Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014.

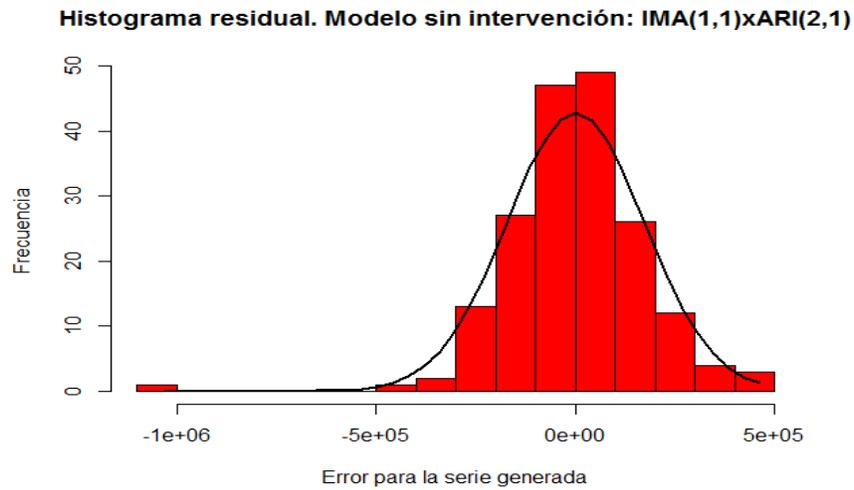
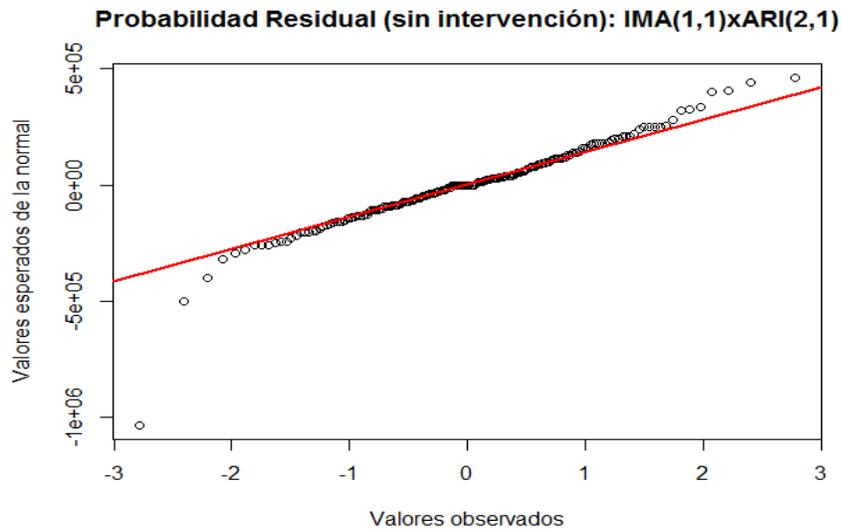


Gráfico 44. Probabilidad Residual del Modelo IMA (1,1) x ARI (2,1)₁₂. Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014.



En el gráfico 43 y 44 correspondientes al histograma de los residuos y probabilidad normal de los residuos, se comprueba que estos se comportan de una manera aproximadamente normal, en el histograma es una normal sesgada hacia la izquierda, esto es debido a los valores atípicos que tiene la serie.

Al analizar los modelos con los estadísticos AIC (El Criterio Akaike) y BIC (El Criterio de información Bayesiana); permite verificar que se tenga un criterio más para identificar el modelo generador de la serie, ya que se ha planteado teóricamente que el modelo adecuado será aquel que cumpla mayores condiciones y criterios.

Tabla 20. Modelos con los estadísticos AIC y BIC. Serie de los ingresos económicos mensuales de los hoteles.

N°	Modelo	AIC	BIC
1	IMA (1,1) x ARIMA(2,1,1) ₁₂	4656.15	4678.18
2	IMA (1,1) x ARIMA(2,1,2) ₁₂	4656.05	4681.23
3	IMA (1,1) x IMA(1,1) ₁₂	4657.77	4676.66
4	IMA (1,1) x ARI (2,1) ₁₂	4662.99	4675.58

Para este caso, el modelo que cumple con esta condición es el que ya se ha identificado en el análisis de residuo. Por lo tanto se acepta el modelo IMA (1,1) x ARI (2,1)₁₂, como el que más se ajusta al comportamiento de los datos de la serie de los ingresos mensuales, pues cumple con la mayoría de los criterios para ser seleccionado. Luego de haber identificado el modelo adecuado, a continuación se incorpora el modelo de intervención.

4.5.2.4. Análisis de intervención.

Una vez definido el modelo ARIMA, se realiza el análisis de intervención. Primero se identifica el periodo de comienzo y posteriormente el tipo de intervención. El período de comienzo de la intervención es en Septiembre de 2008 ($t_0=117$). Se debe analizar los cuatro tipos de intervención propuestos en la teoría para ver cuál de ellos se ajusta mejor a la base de datos en estudio, sin embargo se presentan solo las salidas del modelo con intervención comienzo brusco y duración permanente, ya que por limitaciones en el proceso de optimización del software R, no se pueden calcular los otros tres tipos de modelos de intervención, esto se solventa con otros trabajos interdisciplinarios. A continuación se muestra el único modelo de intervención a analizar.

4.5.2.4.1. Modelo con intervención: Comienzo brusco y duración permanente

Anteriormente se ha calculado el mejor modelo estacional ARIMA (p,d,q) x ARIMA (P,D,Q)₁₂, ahora se incorpora el modelo de intervención identificado, este se puede especificar de la manera siguiente:

$$y_t = \omega S_t^{(t_0=117)} + \frac{(1 - \theta_1 L)}{(1 - \Phi_1 L^{12} - \Phi_2 L^{24})(1 - L^{12})(1 - L)}$$

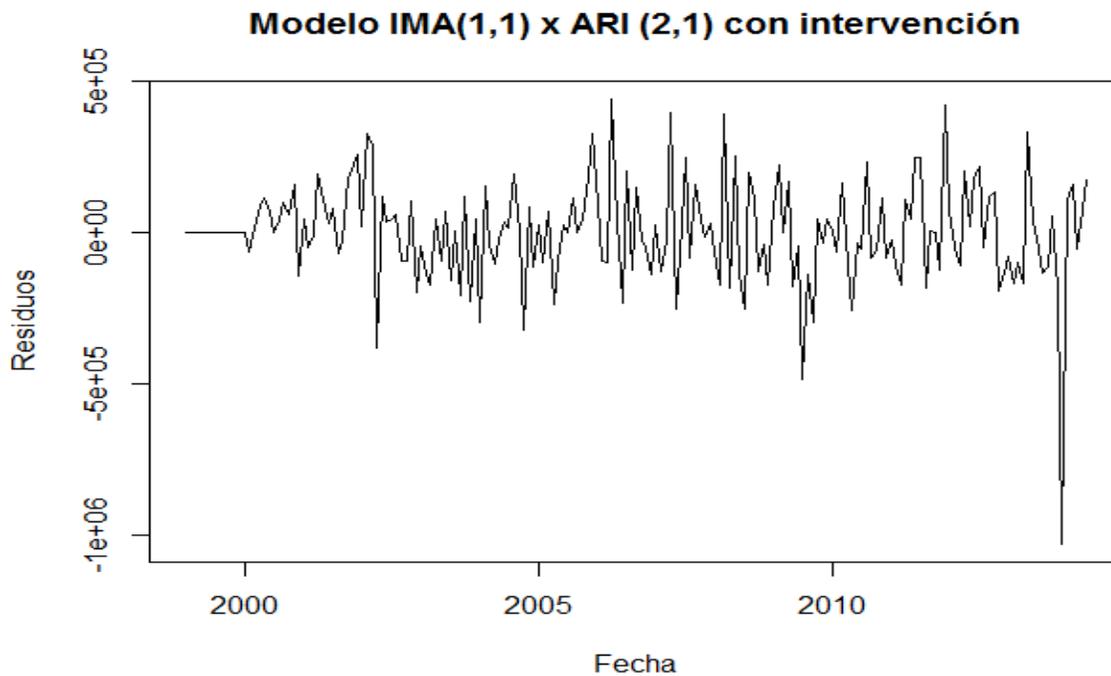
Se calcula la estimación de los parámetros (anexo 7.6) por medio del software R y en seguida se observan los que son significativos, notando que todos aportan al modelo. Por lo tanto, el modelo conjuntamente estimado es:

$$y_t = -20000.53 S_t^{(t_0=117)} + \frac{(1 + 0.7632L)}{(1 + 0.6261L^{12} + 0.4688L^{24})(1 - L^{12})(1 - L)}$$

Análisis de los residuos del Modelo IMA (1,1) x ARI (2,1)₁₂ con intervención comienzo brusco y duración permanente. Serie de los ingresos económicos mensuales de los hoteles

Para llevar a cabo este análisis, a continuación se muestran los gráficos de secuencia y las autocorrelaciones residuales del modelo con intervención.

Gráfico 45. Residuos del Modelo IMA (1,1) x ARI (2,1)₁₂ con intervención. Serie de los ingresos económicos mensuales de los hoteles (Enero de 1999 a Mayo de 2014).



Observando el gráfico anterior (gráfico 45), se observa que la media se ubica en torno a cero ($\mu = 0$), es decir que los residuos del modelo tienen un comportamiento constante a lo largo del tiempo. Los gráficos de la FAS y la FAP de los residuos para el modelo IMA (1,1), x ARI (2,1)₁₂ con intervención se muestran en seguida:

Gráfico 46. FAS Residual. Modelo IMA (1,1) x ARI (2,1)₁₂ con intervención. Serie de los ingresos económicos mensuales de los hoteles. Enero 1999 a Mayo de 2014.

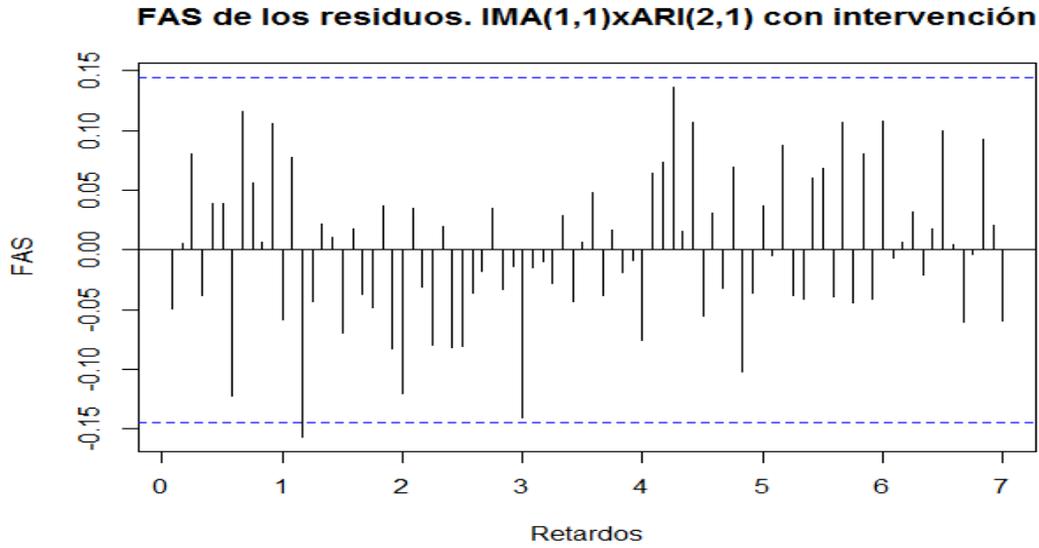
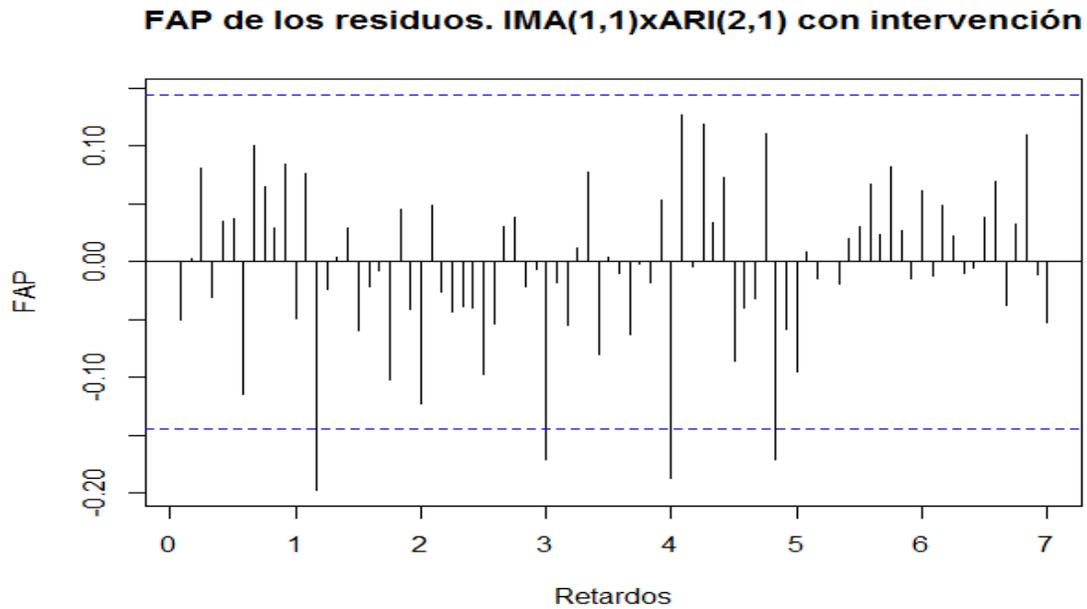


Gráfico 47. FAP Residual. Modelo IMA (1,1) x ARI (2,1)₁₂ con intervención. Serie de los ingresos económicos mensuales de los hoteles. Enero 1999 a Mayo de 2014.



Los gráficos 46 y 47 correspondientes a los correlogramas, obedecen aproximadamente al intervalo de aceptación propuesto. Las autocorrelaciones de los residuos son no significativas o están muy próximas a cero para un nivel de significación $\alpha = 0.05$ y se acepta que no existe correlación entre los residuos del modelo, aunque se observa que aparecen algunos valores elevados de coeficientes de la FAS, así como también en la FAP, estos no son sistemáticos por lo que sería un caso de correlación espúrea. Seguidamente se presenta el histograma residual y probabilidad normal del modelo con intervención.

Gráfico 48. Histograma Residual del Modelo IMA (1,1) x ARI (2,1)₁₂ con Intervención. Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014.

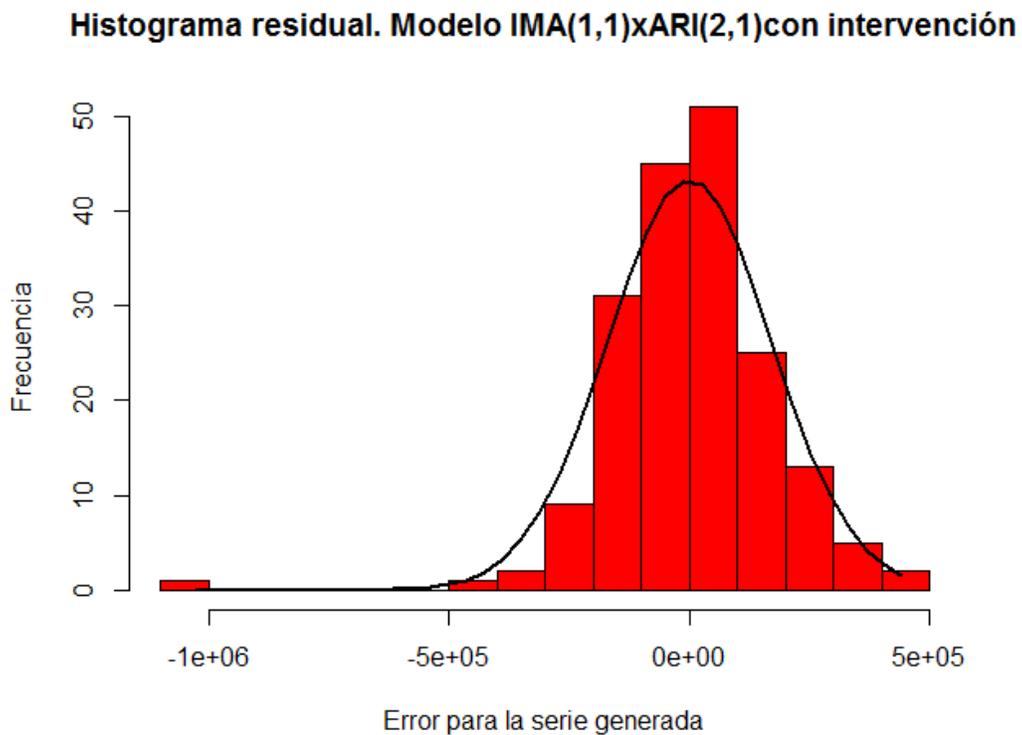
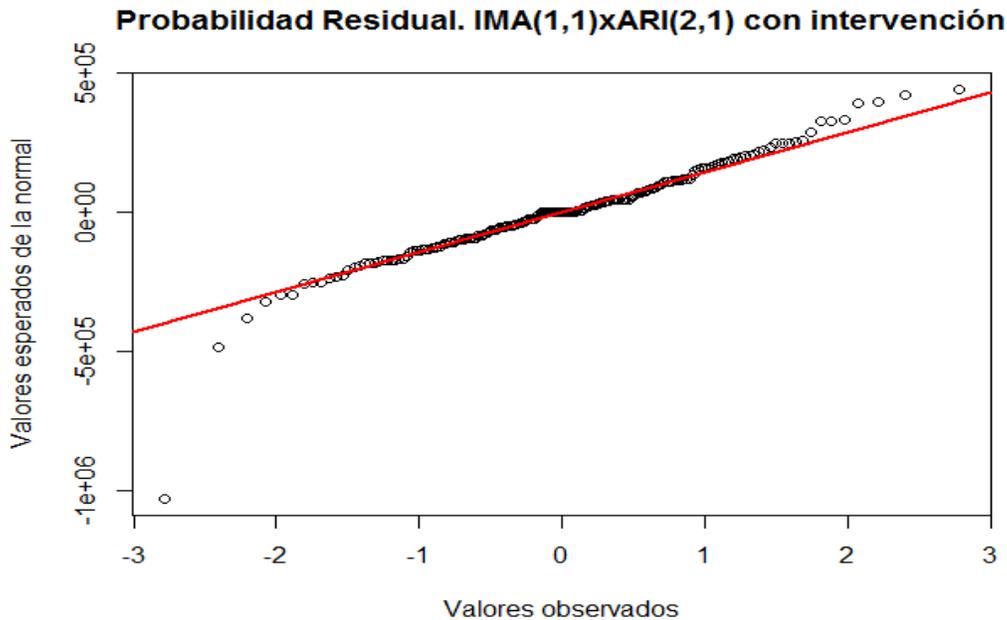


Gráfico 49. Probabilidad Normal de los Residuos del Modelo IMA (1,1) x ARI (2,1)₁₂ con Intervención. Serie de los ingresos económicos mensuales de los hoteles. Enero de 1999 a Mayo de 2014.

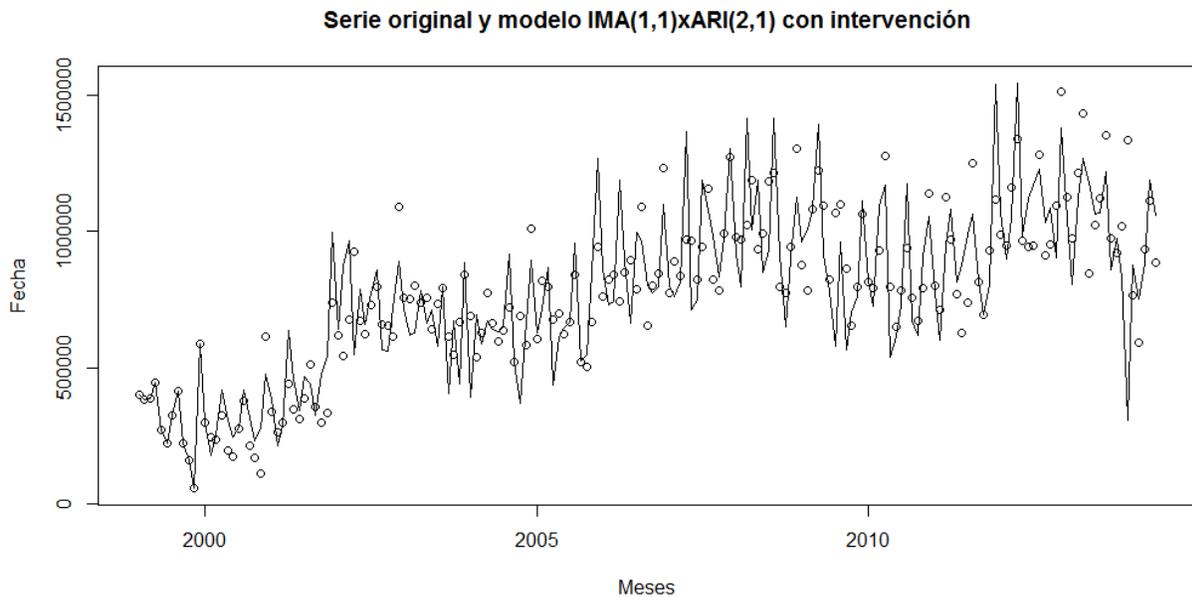
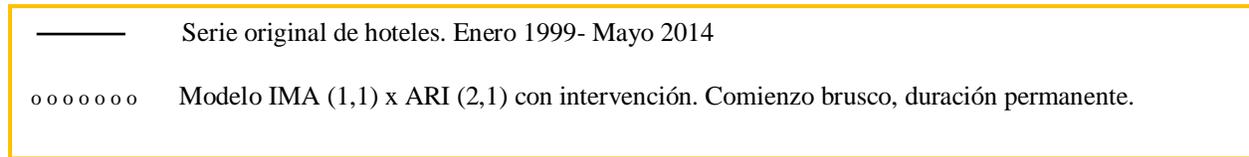


En el gráfico 48 y 49 correspondientes al histograma de los residuos y gráfico de probabilidad residual, se observa que se comportan de forma acampanada con un sesgo a la izquierda, lo cual es debido a los valores atípicos de la serie.

Al finalizar el análisis de los residuos, se decide cuál de los dos modelos se ajusta más a la serie en estudio (el modelo IMA (1,1) x ARI (2,1)₁₂ sin intervención o el IMA (1,1) x ARI (2,1)₁₂ con intervención), en este caso, se elige el modelo con intervención ya que mejora los supuestos de los residuos (media cero, varianza constante y covarianza cero en los residuos). Además el parámetro de intervención es significativo al 5%.

A continuación se muestra la serie de ingresos económicos mensuales de los hoteles con el modelo superpuesto IMA (1,1) x ARI (2,1)₁₂ con intervención. Comienzo brusco y duración permanente.

Gráfico 50. Serie de los ingresos económicos mensuales de los hoteles superpuesto con el modelo IMA (1,1) x ARI (2,1)₁₂ con intervención (Comienzo brusco y duración permanente)



En el gráfico 50 se observa que el modelo IMA (1,1) x ARI (2,1)₁₂ con intervención (Comienzo brusco y duración permanente) representado a través de puntos, se ajusta aproximadamente a la serie de datos de la serie original de hoteles, sin embargo, se tienen algunos puntos que discrepan con el modelo planteado por lo que se realiza un análisis de outliers.

4.5.2.5. Análisis de Outliers.

La detección de outliers puede proporcionar información respecto a la ocurrencia de determinados factores externos que afectan a la serie, señalando tanto el momento en que se producen como la manera en que actúan sobre la misma.

Para mejorar la comprensión de la serie objeto de análisis, se detectan primeramente todos los valores atípicos del modelo residual IMA (1,1) x ARI (2,1)₁₂ con intervención por medio del software R, identificando el tipo de dato atípico, también se determina el valor residual y el dato

real de la serie. Luego, se verifica si en la fecha que ocurrió el outlier se presentó el evento de veda por marea roja. A continuación se detallan estos resultados:

Tabla 21. Detección de outliers y comparación de los residuos del modelo de la serie de los ingresos económicos mensuales de los hoteles con intervención comienzo brusco y duración permanente: IMA (1,1) x ARI (2,1)₁₂.

N° de caso	Tipo	Fecha	Valor del residuo	Dato real (ingreso en \$)	Presencia de veda por marea roja ¹⁷
34	LS	Octubre 2001	177715.8	477248.62	Si
127	TC	Julio 2009	-487730.5	578931.23	No
156	AO	Diciembre 2011	423563.8	1539000.00	No
179	TC, LS	Noviembre 2013	-185313	831501.00	No
180	AO, TC, LS	Diciembre 2013	-1029997	305513.54	No

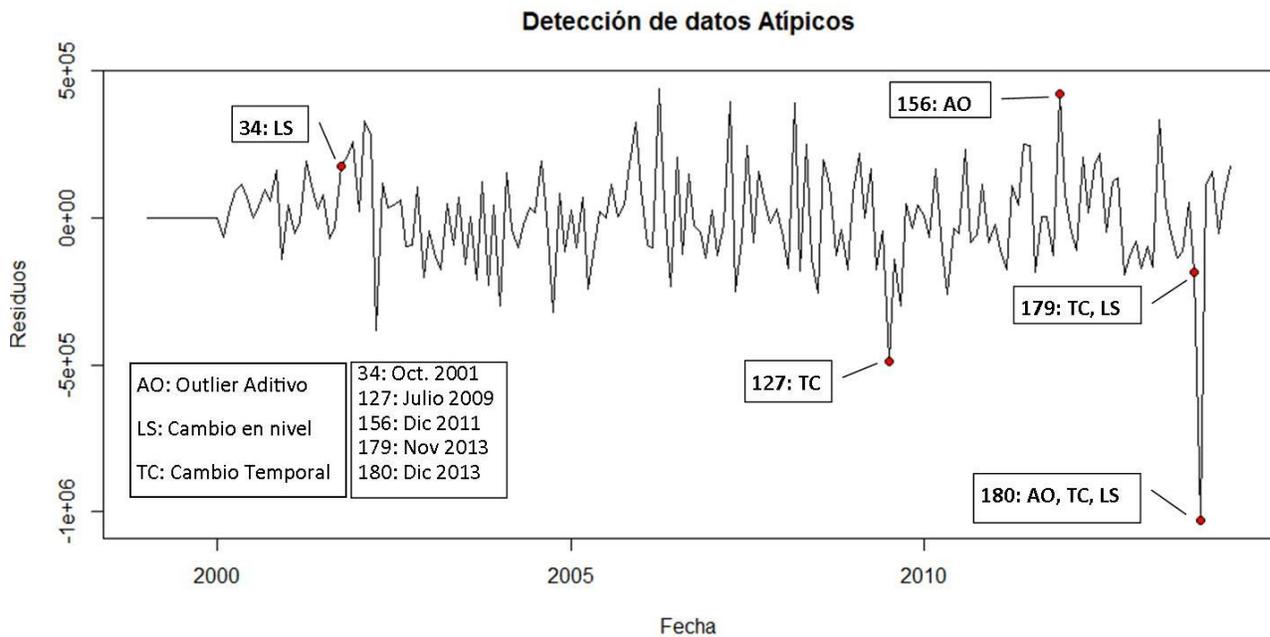
Los resultados del software muestra 5 datos atípicos, representados en la tabla anterior (tabla 21), de los cuales solo uno está asociado a un evento de veda por marea roja (N° de caso 34) correspondiente a Octubre de 2001, recordando que en esta fecha según la tabla 1 del Marco Teórico, la veda por marea roja corresponde al periodo de Agosto 2001 hasta Enero de 2002, a pesar que el dato atípico se intercepta con la veda, el valor del residuo es alto, lo que indica que la veda decretada en ese período, no perjudico a los ingresos de los Hoteles del municipio de La Libertad.

Respecto a los 4 cuatro datos atípicos que no están asociados a vedas por marea roja posiblemente son debido a otros factores. En caso de ser outliers bajos, generalmente se nota que es un mes antes de un período vacacional, es decir que un mes antes los turistas están en espera de la vacación para ir a hospedarse a los hoteles del municipio. Respecto a valores altos en los outliers corresponde en la mayoría de los casos a días festivos o períodos vacacionales (semana santa, agosto y diciembre)

¹⁷ Ver tabla resumen de eventos recurrentes de las mareas rojas en El Salvador.

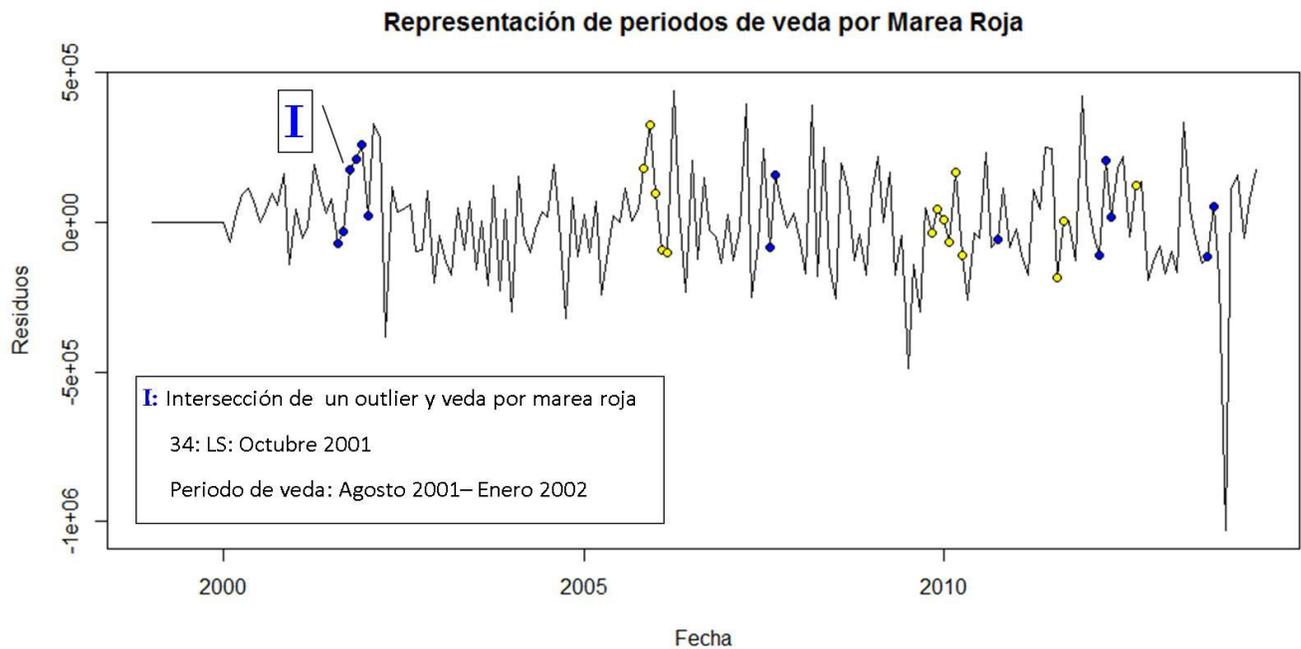
A continuación se muestra el gráfico de los residuos del modelo con intervención, en la cual se han representado por puntos de color rojo los datos considerados como atípicos, sobre los cuales ya se ha realizado un análisis descriptivo.

Gráfico 51. Representación de datos atípicos del modelo residual IMA (1,1) x ARI (2,1)₁₂ con intervención. Serie de los ingresos económicos mensuales de los hoteles.



El gráfico 51 muestra los 5 outliers de los residuos del modelo con intervención, el número de caso, la fecha a la que corresponde, así como el tipo de outlier al que corresponde cada uno, es importante recordar que uno de estos atípicos está asociado a un evento de veda por marea roja, este se muestra en el grafico siguiente:

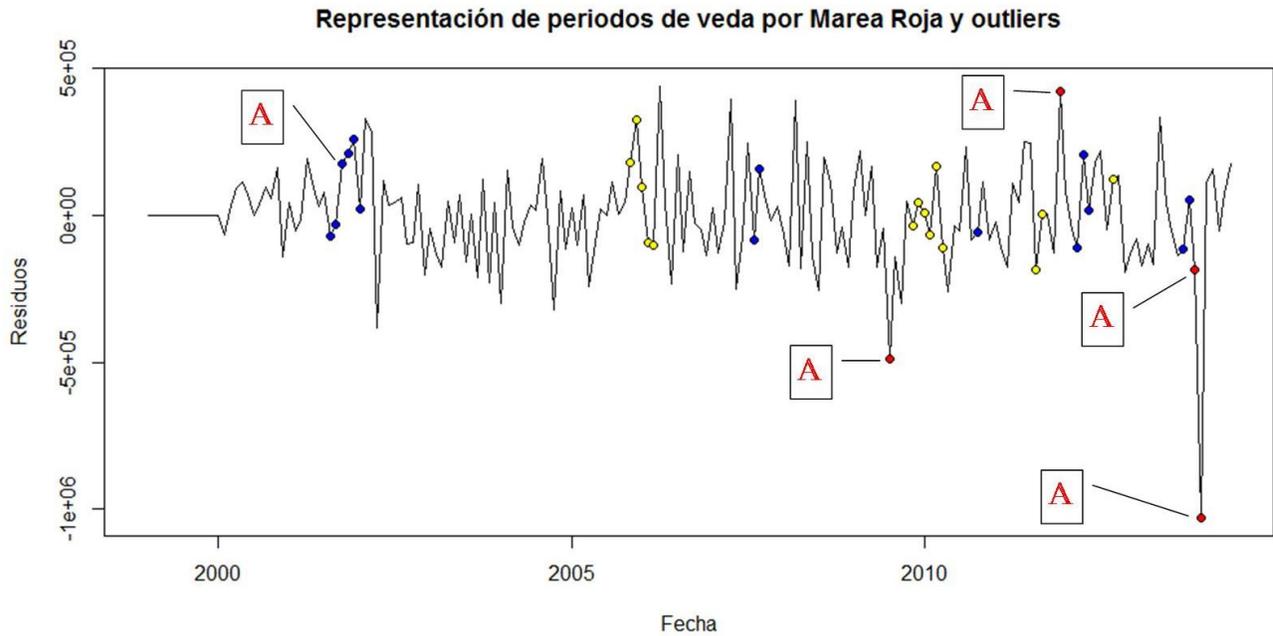
Gráfico 52. Representación de períodos de vedas por mareas rojas en el modelo residual IMA x (1,1) x ARI (2,1)₁₂ con intervención. Serie de los ingresos económicos mensuales de los hoteles.



En el gráfico 52 se muestran los 9 períodos de veda por marea roja, diferenciando cada periodo con colores amarillo o azul, se observa que la serie tiene un comportamiento muy similar durante el periodo en estudio, no se puede asociar con cambios en los ingresos de los hoteles del municipio de La Libertad. También se representa un dato atípico de los 5 que se identificaron anteriormente, dicho dato está asociado a un evento de veda, así que es representado gráficamente por una “I”.

Seguidamente se muestra una superposición de los puntos que representan los outliers juntamente con los periodos de veda para un mejor análisis gráfico.

Gráfico 53. Representación de períodos de vedas por mareas rojas y datos outliers sobre el modelo residual IMA $x(1,1) \times \text{ARI}(2,1)_{12}$ con intervención. Serie de los ingresos económicos mensuales de los hoteles.



En el gráfico 53 se representan por la letra “A” los valores atípicos de la serie de hoteles, también se muestra por medio de puntos amarillos y azules los períodos de veda por marea roja, estos se han representado juntos con el fin de observar si gráficamente existe alguna asociación entre la intervención y los datos atípicos.

5. Conclusiones y Recomendaciones

Al finalizar el trabajo de graduación, se concluye que:

- En la indagación referente a: 1) Investigaciones realizadas por medio de los diferentes ministerios (MAG, MINEC, MARN y de Hacienda) así como también en el Banco Central de Reserva sobre mediciones de impacto económico causado por mareas rojas en El Salvador y 2) Base de datos de las ventas diarias, semanales y/o mensuales de mariscos por los pescadores artesanales de las diferentes zonas costeras del país; no se encontró ninguna investigación previa a esta, ni base alguna para realizar la investigación.
- Las instituciones consultadas se responsabilizan unas con otras de tener los registros de datos mencionados anteriormente y otras que claramente dicen que dicha información es inexistente. Por las razones antes mencionadas, se optó por hacer la investigación sobre una base de ingresos mensuales de negocios formales del sector hotelero y de restaurantes deducida a través de la declaración de IVA proporcionada por el Ministerio de Hacienda.
- Se detectó la existencia de una intervención (crisis económica, septiembre 2008) que pudo influenciar los ingresos en el sector de restaurantes y hoteles, esto se trató con un análisis de intervención.
- En la serie de restaurantes se realizó un análisis de los residuos para el modelo $ARI(2,1) \times ARI(2,1)_{12}$ con intervención y sin intervención, quedándonos con el modelo $ARI(2,1) \times ARI(2,1)_{12}$ con intervención comienzo brusco y duración temporal, siendo este el que mejor se ajusta.
- En la serie de hoteles se realizó un análisis de los residuos para el modelo $IMA(1,1) \times ARI(2,1)_{12}$ con intervención y sin intervención, quedándonos con el modelo con intervención: comienzo brusco y duración permanente, siendo este el que mejor se ajusta, a la serie de datos.

- Se realizó en ambas bases un análisis de valores extremos, para identificar los eventos donde hubo disminuciones económicas debido a veda por marea roja o por otras causas. Y se observa que la veda por marea roja no afecta los ingresos económicos de los restaurantes y hoteles, ya que aunque haya veda los turistas siempre consumen otros alimentos y deciden hospedarse tranquilamente.

Se recomienda que:

- En vista que las FAN no afectan en los ingresos económicos de los negocios formales de hoteles y restaurantes en el municipio de La Libertad, se les recomienda a las instituciones gubernamentales primeramente que se cree una base de datos en la que se registren ingresos diarios, semanales y/o mensuales del comercio de alimentos proveniente del mar, que puedan ser útiles posteriormente para análisis estadísticos similares o superiores al presentado en este trabajo de investigación. Por ejemplo, se recomienda que CENDEPESCA junto a las cooperativas registren ingresos exactos en ventas de productos extraídos del mar, no solo en el municipio de La Libertad¹⁸, sino que en toda la zona costera del país.
- Que el resultado de las actas impresas donde se decreta las vedas y el periodo de duración sean puestos a disposición del público a través del Ministerio de Medio Ambiente y Recursos Naturales por medio de cuadros resúmenes en formato digital, que se registren días exactos de ocurrencia de mareas rojas en el país para poder utilizarse en investigaciones posteriores, además que se diferencie si fué marea roja con veda o sin veda y que se incluyan si hubieron daños en los animales y en la población.
- Que el Ministerio de Salud registre los cuadros clínicos de pacientes por intoxicación causada por marea roja, ya que en los registros actuales solo estas clasificados como intoxicación, pero no especifica la causa.

¹⁸ La oficina de CENDEPESCA ubicado en El Puerto de La Libertad, es la única sede a nivel nacional que registra los ingresos de las cooperativas pesqueras, esto lo han empezado a hacer recientemente (desde el año 2013). Información proporcionada por: Norma Idalia Lobo Martel, Norma.lobom@mag.gob.sv y Lic. Numa Hernández, numa.hernandez@mag.gob.sv, el día 10 de Junio de 2014.

- En caso fallezca una persona por causa de intoxicación al consumir productos contaminados por marea roja, esta se debe de clasificar por el Instituto De Medicina Legal de acuerdo a la causa que determina esa muerte.
- Que el Ministerio de Turismo registre periodicamente el ingreso de turistas a parques recreativos y piscinas, restaurantes y hoteles de las zonas costeras. Además, que invierta en campañas de concientización en medios televisivos, radiales y redes sociales, donde se mencionen las diferentes formas en que se puede intoxicar una persona y sobre los productos específicos que no se deben consumir cuando se decreta la veda.
- Si las instituciones pertinentes toman en cuenta las recomendaciones expuestas en este trabajo de investigación, pasados unos años se podrán realizar investigaciones puntuales sobre el impacto de veda por marea roja en diversos sectores del país, empezando por algunos municipios específicos, pueden ser referentes a ventas (de preferencia diarias), ya sea en mercados o ventas directas de los pescadores artesanales en determinados puertos de El Salvador. Estos estudios se van replicando hasta que finalmente se podrá estar realizando la investigación recomendada por LABTOX de la UES donde se pueda responder ¿cuanto le cuesta al estado una Marea Roja?. Lo anterior se logra a través de la creación de un modelo multivariado que estime los costos económicos causado por una marea roja en El Salvador, donde incluya el sector salud, economía, turismo, medio ambiente y otros.

6. Referencias bibliográficas y fuentes consultadas.

- Adams, L. &. (2007). HABS and Coastal Business: Economic Consequences in Florida. *Society and Natural Resources*.
- Adams, S. L. (Agosto de 2013). *Economic Consequences of Harmful Algal Blooms: Literature Summary*. Recuperado el Abril de 2014, de Universidad de Florida:
<http://www.fred.ifas.ufl.edu/pdf/Adams-Larkin-LitRev-April2013.pdf>
- Aznar, A., & Trivez, F. J. (1993). *Métodos de Predicción en economía (II). Análisis de Series Temporales. 1a Edición*. Barcelona: ARIEL, S.A.
- Castellanos, J. (23 de marzo de 2012). *Ministerio de Agricultura y Ganadería*. Recuperado el 13 de septiembre de 2014, de
http://www.mag.gob.sv/index.php?option=com_k2&view=item&id=105:conamar-descarta-presencia-de-marea-roja-t%C3%B3xica&Itemid=168
- Del Valle, L. Y. (18 de Febrero de 2009). *Marea roja: Un fenómeno para investigar. Alerta Ambiental*. Recuperado el 12 de Marzo de 2014, de <http://seagrantpr.org/marejada/alerta-ambiental/marea-roja-un-fenomeno-para-investigar/>
- Espinoza Navarrete, J. J., Amaya Monterrosa, O. A., & Quintanilla, C. R. (2013). *Atlas de Fitoplancton Marino. 1a Edición*. Ciudad Universitaria, San Salvador: LABTOX-UES; Facultad de Ciencias Naturales y Matemática.
- Fernández, S. d. (s.f.). Recuperado el 13 de septiembre de 2014, de <http://www.fuenterrebollo.com/Economicas/SERIES-TEMPORALES/modelo-arima.pdf>
- Fonseca, M. (s.f.). *CONICYT, Ministerio de Educación, Gobierno de Chile*. Recuperado el 17 de Mayo de 2014, de ¿Cuáles son las consecuencias de la marea roja para las personas? : <http://www.explora.cl/descubre/respuestas-de-un-cientifico/oceanos-preguntas/oceanografia-preguntas/1349-cuales-son-las-consecuencias-de-la-marea-roja-para-las-personas>
- Jaimes Berrios, A. A., & Quintanilla Aparicio, M. I. (2008). *Análisis de Series Temporales con outliers e intervenciones y sus aplicaciones*. Ciudad Universitaria, San Salvador, El Salvador.: Trabajo de Graduación.
- MARES, A. d. (24 de Agosto de 2005). *PORTAL DE LOS 7 MARES*. Recuperado el 14 de Marzo de 2014, de
http://mardechile.cl/index.php?option=com_content&view=article&id=123:marea-roja&catid=21:ecologia&Itemid=66.

- Reguera, B., Alonso, R., Moreira, Á., & Méndez, S. (2011). *Guía para el diseño y puesta en marcha de un plan de seguimiento de microalgas productoras de toxinas. Proyecto ARCAL RLA 7/014. Manual de la UNESCO. Manuals and Guides 59. Intergovernment Oceanographic Commission*. <http://unesdoc.unesco.org/images/0021/002145/21450s.pdf>.
- Suarez Isla, D., & Guzmán Méndez, D. (s.f.). *Floraciones de Algas Nocivas. Mareas Rojas y Toxinas Marinas. Instituto de Ciencias Biomedicas, Facultad de Medicina; Universidad de Chile. Instituto de Fomento Pesquero*. Recuperado el 12 de Marzo de 2014, de http://www.ifop.cl/mr/images/upload/file/8_%20FLORACIONES%20DE%20ALGAS%20NOCIVAS%20Mareas%20Rojas%20y%20Toxinas%20Marinas%20Guzman%20y%20Suarez%201998.pdf
- Torres Chuquimarca, G. M. (2011). *Eventos de mareas rojas: estrategias de manejo preventivas en Ecuador. Tesis de grado para la obtención de título de Magister en Ciencias*. Recuperado el 11 de Marzo de 2014, de Guayaquil, Ecuador: PARTE I: <http://201.218.27.134/bitstream/redug/812/1/Eventos%20de%20mareas%20rojas%20Parte%201.pdf> y PARTE 2: <http://201.218.27.134/bitstream/redug/812/3/Eventos%20de%20mareas%20rojas%20Parte%202.pdf>
- Tovar R, T. (03 de Octubre de 2006). *Mareas rojas y toxina amnésica de los moluscos: Posibilidades en el mar Mediterraneo*. Recuperado el 14 de Marzo de 2014, de <http://mareasrojasyp.blogspot.com/2006/10/15-consecuencias-econmicas-de-las.html>

7. Anexos.

Anexo 7.1. Base de datos de los ingresos económicos mensuales de los restaurantes del municipio de La Libertad

Fecha	Rest IVA	Fecha	Rest IVA								
ene-98	23.252,23	oct-00	181.784,41	jul-03	406.282,03	abr-06	935.407,24	ene-09	556.379,71	oct-11	217.004,31
feb-98	16.753,83	nov-00	153.296,47	ago-03	502.043,56	may-06	548.768,06	feb-09	417.989,25	nov-11	316.093,67
mar-98	19.802,74	dic-00	246.492,97	sep-03	355.281,99	jun-06	466.874,14	mar-09	235.477,63	dic-11	540.095,58
abr-98	23.887,66	ene-01	174.202,00	oct-03	331.476,52	jul-06	568.307,30	abr-09	610.513,35	ene-12	423.848,36
may-98	23.280,80	feb-01	86.846,51	nov-03	386.370,62	ago-06	648.262,13	may-09	475.617,65	feb-12	337.358,64
jun-98	16.162,40	mar-01	181.140,01	dic-03	600.799,29	sep-06	510.846,61	jun-09	372.528,95	mar-12	389.531,40
jul-98	17.004,91	abr-01	331.541,05	ene-04	510.300,51	oct-06	483.168,45	jul-09	361.403,25	abr-12	587.102,47
ago-98	23.372,23	may-01	234.195,48	feb-04	432.178,57	nov-06	530.792,02	ago-09	448.669,75	may-12	358.805,47
sep-98	16.012,46	jun-01	221.869,66	mar-04	413.194,65	dic-06	928.483,17	sep-09	284.406,33	jun-12	372.575,29
oct-98	16.677,14	jul-01	236.315,88	abr-04	612.322,16	ene-07	692.613,64	oct-09	277.861,56	jul-12	406.619,55
nov-98	11.796,11	ago-01	282.681,24	may-04	460.008,86	feb-07	521.238,32	nov-09	254.832,26	ago-12	457.888,38
dic-98	25.229,60	sep-01	134.960,28	jun-04	407.987,00	mar-07	614.549,10	dic-09	538.916,78	sep-12	305.763,19
ene-99	125.953,99	oct-01	159.322,79	jul-04	529.465,73	abr-07	885.741,67	ene-10	1.221.262,36	oct-12	200.906,62
feb-99	114.441,67	nov-01	176.846,13	ago-04	616.483,00	may-07	542.513,59	feb-10	353.871,42	nov-12	298.163,95
mar-99	118.253,78	dic-01	307.285,89	sep-04	398.387,07	jun-07	534.243,63	mar-10	446.029,96	dic-12	602.332,53
abr-99	134.727,54	ene-02	233.862,24	oct-04	444.088,71	jul-07	547.475,04	abr-10	487.225,56	ene-13	470.934,20
may-99	124.352,39	feb-02	224.084,83	nov-04	452.463,00	ago-07	609.968,72	may-10	328.973,94	feb-13	369.636,42
jun-99	110.363,24	mar-02	411.325,22	dic-04	721.430,57	sep-07	431.031,52	jun-10	245.826,50	mar-13	711.820,68
jul-99	116.028,94	abr-02	240.608,23	ene-05	567.506,13	oct-07	210.451,40	jul-10	318.587,01	abr-13	356.245,89
ago-99	146.943,12	may-02	227.290,30	feb-05	444.292,99	nov-07	470.996,55	ago-10	436.882,66	may-13	405.855,76
sep-99	84.452,81	jun-02	265.066,17	mar-05	684.423,39	dic-07	853.600,57	sep-10	245.352,87	jun-13	318.350,76
oct-99	115.626,47	jul-02	280.825,02	abr-05	451.768,29	ene-08	585.035,81	oct-10	310.670,36	jul-13	369.072,79
nov-99	77.455,39	ago-02	378.776,76	may-05	444.971,77	feb-08	498.532,90	nov-10	331.323,18	ago-13	404.505,27
dic-99	224.692,17	sep-02	313.167,52	jun-05	437.173,64	mar-08	816.805,23	dic-10	618.040,77	sep-13	283.968,32
ene-00	167.035,44	oct-02	306.854,87	jul-05	573.822,57	abr-08	473.346,67	ene-11	538.286,25	oct-13	192.203,30
feb-00	157.724,83	nov-02	371.719,52	ago-05	617.526,47	may-08	549.908,83	feb-11	331.631,14	nov-13	211.501,64
mar-00	193.320,16	dic-02	557.276,66	sep-05	458.244,19	jun-08	467.194,95	mar-11	334.185,91	dic-13	272.663,15
abr-00	289.274,57	ene-03	391.643,83	oct-05	332.361,35	jul-08	459.633,48	abr-11	586.361,46	ene-14	525.227,74
may-00	134.304,77	feb-03	378.379,18	nov-05	409.265,66	ago-08	675.308,00	may-11	350.654,88	feb-14	293.502,94
jun-00	151.766,79	mar-03	422.191,90	dic-05	638.138,58	sep-08	400.947,86	jun-11	310.540,35	mar-14	397.660,90
jul-00	222.668,48	abr-03	587.296,09	ene-06	594.472,86	oct-08	379.208,98	jul-11	385.505,47	abr-14	515.050,86
ago-00	190.524,13	may-03	394.228,31	feb-06	249.881,75	nov-08	487.562,54	ago-11	468.188,53	may-14	281.924,75
sep-00	159.910,81	jun-03	360.262,59	mar-06	519.197,18	dic-08	749.103,64	sep-11	294.738,51	jun-14	52.499,57

Anexo 7.2. Resultados del estadístico R de los modelos propuestos. Serie de los ingresos económicos mensuales de los restaurantes.

7.2.1. ARIMA (2,1,0) x ARIMA (2,1,0)₁₂

```
Series: x
ARIMA(2,1,0)(2,1,0)[12]
Coefficients:
      ar1      ar2      sar1      sar2
      -0.4635  -0.3063  -0.6632  -0.3842
s.e.      0.0755   0.0766   0.0864   0.0833
sigma^2 estimated as 0.1076:  log likelihood=-59.95
AIC=127.91  AICC=128.24  BIC=144.01

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1  -0.463530   0.075463  -6.1425  8.125e-10 ***
ar2  -0.306349   0.076617  -3.9984  6.376e-05 ***
```

```

sar1 -0.663215 0.086430 -7.6734 1.675e-14 ***
sar2 -0.384163 0.083319 -4.6108 4.012e-06 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

7.2.2. ARIMA (2,1,0) x ARIMA (2,1,1)₁₂

```

Series: x
ARIMA(2,1,0)(2,1,1)[12]
Coefficients:
      ar1      ar2      sar1      sar2      sma1
-0.3938 -0.2409 -0.1495 -0.1827 -0.7104
s.e.    0.0791  0.0794  0.1399  0.1169  0.1320
sigma^2 estimated as 0.09816: log likelihood=-54.8
AIC=119.6  AICC=120.07  BIC=138.92

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1  -0.393780  0.079111 -4.9776 6.439e-07 ***
ar2  -0.240929  0.079449 -3.0325 0.002425 **
sar1 -0.149484  0.139926 -1.0683 0.285383
sar2 -0.182700  0.116943 -1.5623 0.118217
sma1 -0.710442  0.132026 -5.3811 7.404e-08 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

7.2.3. ARIMA (2,1,0) x ARIMA (2,1,2)₁₂

```

Series: x
ARIMA(2,1,0)(2,1,2)[12]
Coefficients:
      ar1      ar2      sar1      sar2      sma1      sma2
-0.3989 -0.2416 -0.5396 -0.1734 -0.3108 -0.3416
s.e.    0.0789  0.0796  0.4429  0.1172  0.4551  0.3907
sigma^2 estimated as 0.09799: log likelihood=-54.71
AIC=121.41  AICC=122.05  BIC=143.96

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1  -0.398899  0.078922 -5.0543 4.319e-07 ***
ar2  -0.241579  0.079604 -3.0348 0.002407 **
sar1 -0.539591  0.442949 -1.2182 0.223157
sar2 -0.173430  0.117229 -1.4794 0.139030
sma1 -0.310819  0.455055 -0.6830 0.494584
sma2 -0.341644  0.390709 -0.8744 0.381890
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

7.2.4. ARIMA (2,1,1) x ARIMA (2,1,1)₁₂

```

Series: x
ARIMA(2,1,1)(2,1,1)[12]
Coefficients:
      ar1      ar2      ma1      sar1      sar2      sma1
0.2009 -0.0302 -0.6406 -0.1408 -0.2095 -0.6805
s.e.    0.1828  0.1154  0.1663  0.1315  0.1120  0.1242
sigma^2 estimated as 0.09697: log likelihood=-53.3
AIC=118.61  AICC=119.24  BIC=141.15

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1  0.200905  0.182803  1.0990 0.2717560
ar2 -0.030197  0.115399 -0.2617 0.7935701
ma1 -0.640588  0.166346 -3.8509 0.0001177 ***
sar1 -0.140753  0.131511 -1.0703 0.2844967
sar2 -0.209496  0.111979 -1.8709 0.0613657 .
sma1 -0.680460  0.124156 -5.4807 4.237e-08 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

7.2.5. ARIMA (1,1,0) x ARIMA (2,1,1)₁₂

```
Series: x
ARIMA(1,1,0)(2,1,1)[12]
Coefficients:
      ar1      sar1      sar2      sma1
      -0.3100 -0.1006 -0.2090 -0.7642
s.e.    0.0753  0.1213  0.1084  0.1111
sigma^2 estimated as 0.1018: log likelihood=-59.27
AIC=126.54  AICC=126.88  BIC=142.64

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1  -0.310041  0.075263 -4.1194 3.798e-05 ***
sar1  -0.100554  0.121291 -0.8290  0.40709
sar2  -0.208962  0.108433 -1.9271  0.05397 .
sma1  -0.764214  0.111083 -6.8797 5.999e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Anexo 7.3. Estimación de los parámetros del mejor modelo de los modelos propuestos para la serie de los ingresos económicos mensuales de los hoteles. Modelos con intervención

7.3.1. ARIMA (2,1,0) x ARIMA (2,1,0)₁₂ con intervención: comienzo brusco y duración permanente

```
Series: log.restaurantes.ts
ARIMA(2,1,0)(2,1,0)[12]
Coefficients:
      ar1      ar2      sar1      sar2  step129-MA0
      -0.4637 -0.3063 -0.6629 -0.3842  0.0067
s.e.    0.0758  0.0766  0.0870  0.0833  0.2518
sigma^2 estimated as 0.1076: log likelihood=-59.95
AIC=129.91  AICC=130.38  BIC=149.23

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1  -0.4637113  0.0757634 -6.1205 9.327e-10 ***
ar2  -0.3063049  0.0766070 -3.9984 6.377e-05 ***
sar1  -0.6629366  0.0869656 -7.6230 2.479e-14 ***
sar2  -0.3842269  0.0833237 -4.6113 4.002e-06 ***
step129-MA0  0.0067128  0.2517889  0.0267  0.9787
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Con el coeficiente de step129-MA0 se puede ver que este término no es significativo, por tanto no se analizan los residuos.

7.3.2. ARIMA (2,1,0) x ARIMA (2,1,0)₁₂ con intervención: comienzo gradual y duración permanente

```
Series: log.restaurantes.ts
```

```

ARIMA(2,1,0)(2,1,0)[12]
Coefficients:
      ar1      ar2      sar1      sar2  step129-AR1  step129-MA0
s.e.  -0.4637  -0.3063  -0.6629  -0.3842   -0.0013     0.0065
      0.0754   0.0766   0.0863   0.0833          NaN          NaN

sigma^2 estimated as 0.1076: log likelihood=-59.95
AIC=131.91  AICC=132.54  BIC=154.45
Warning message:
In sqrt(diag(x$var.coef)) : NaNs produced

```

El modelo anterior, no se incluye en el análisis.

7.3.3. ARIMA (2,1,0) x ARIMA (2,1,0)₁₂ con intervención: comienzo brusco y duración temporal

```

Series: log.restaurantes.ts
ARIMA(2,1,0)(2,1,0)[12]
Coefficients:
      ar1      ar2      sar1      sar2  step129-AR1  step129-MA0
s.e.  -0.4478  -0.3235  -0.6446  -0.4054   -0.9711     -0.1285
      0.0749   0.0757   0.0878   0.0822    0.0224     0.0593
sigma^2 estimated as 0.1046: log likelihood=-57.44
AIC=126.88  AICC=127.51  BIC=149.42

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1      -0.447808  0.074894  -5.9792  2.242e-09 ***
ar2      -0.323509  0.075736  -4.2715  1.941e-05 ***
sar1     -0.644568  0.087796  -7.3416  2.110e-13 ***
sar2     -0.405391  0.082238  -4.9295  8.244e-07 ***
step129-AR1 -0.971140  0.022432 -43.2928 < 2.2e-16 ***
step129-MA0 -0.128536  0.059324  -2.1667  0.03026 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

7.3.4. ARIMA (2,1,0) x ARIMA (2,1,0)₁₂ con intervención: comienzo gradual y duración temporal

```

Series: log.restaurantes.ts
ARIMA(2,1,0)(2,1,0)[12]
Coefficients:
      ar1      ar2      sar1      sar2  step129-AR1  step129-AR2
s.e.  -0.4440  -0.3220  -0.6550  -0.3999    0.0740    1.0222
      0.0744   0.0754   0.0863   0.0820    0.0212    0.0253
step129-MA0
s.e.  -0.2294
      0.1127
sigma^2 estimated as 0.1034: log likelihood=-56.33
AIC=126.66  AICC=127.48  BIC=152.43

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1      -0.444015  0.074391  -5.9686  2.393e-09 ***
ar2      -0.321983  0.075366  -4.2722  1.935e-05 ***
sar1     -0.655024  0.086275  -7.5923  3.143e-14 ***
sar2     -0.399895  0.082002  -4.8766  1.079e-06 ***
step129-AR1  0.074015  0.021226  3.4869  0.0004886 ***
step129-AR2  1.022222  0.025321  40.3707 < 2.2e-16 ***
step129-MA0 -0.229403  0.112734  -2.0349  0.0418594 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Anexo 7.4. Base de datos de ingresos económicos mensuales de los hoteles del municipio de La Libertad.

Fecha	Hot_IVA	Fecha	Hot_IVA	Fecha	Hot_IVA	Fecha	Hot_IVA	Fecha	Hot_IVA	Fecha	Hot_IVA	Fecha	Hot_IVA
ene-99	51.899,80	abr-01	82.868,03	jul-03	75.081,24	oct-05	71.427,98	ene-08	118.879,56	abr-10	152.060,28	jul-12	151.915,71
feb-99	49.817,54	may-01	59.612,07	ago-03	103.517,93	nov-05	110.505,89	feb-08	103.642,89	may-10	70.043,80	ago-12	159.903,01
mar-99	50.186,65	jun-01	44.615,20	sep-03	52.789,22	dic-05	165.109,51	mar-08	183.722,52	jun-10	79.887,83	sep-12	134.287,39
abr-99	57.626,95	jul-01	60.897,54	oct-03	87.191,91	ene-06	111.828,91	abr-08	130.705,31	jul-10	95.084,35	oct-12	141.148,18
may-99	35.247,94	ago-01	57.580,94	nov-03	57.276,51	feb-06	94.711,44	may-08	154.467,10	ago-10	152.672,03	nov-12	117.330,02
jun-99	28.960,01	sep-01	42.154,09	dic-03	115.198,94	mar-06	96.494,45	jun-08	110.544,71	sep-10	87.448,36	dic-12	179.207,68
jul-99	42.018,82	oct-01	62.042,32	ene-04	51.117,21	abr-06	154.144,84	jul-08	121.033,14	oct-10	80.283,61	ene-13	136.154,31
ago-99	54.011,30	nov-01	70.836,59	feb-04	90.294,40	may-06	116.269,75	ago-08	183.790,39	nov-10	117.873,59	feb-13	104.431,04
sep-99	29.043,81	dic-01	129.473,75	mar-04	76.559,49	jun-06	86.063,62	sep-08	118.506,95	dic-10	137.042,11	mar-13	145.561,32
oct-99	20.777,93	ene-02	83.185,79	abr-04	87.536,30	jul-06	129.358,92	oct-08	84.279,30	ene-11	101.281,34	abr-13	164.678,01
nov-99	7.491,71	feb-02	113.483,49	may-04	83.179,59	ago-06	125.658,46	nov-08	117.637,20	feb-11	78.390,79	may-13	153.391,91
dic-99	76.611,54	mar-02	125.289,29	jun-04	82.450,47	sep-06	104.390,90	dic-08	146.615,24	mar-11	123.640,12	jun-13	138.300,25
ene-00	38.706,70	abr-02	71.067,48	jul-04	84.747,14	oct-06	100.515,35	ene-09	125.030,90	abr-11	140.346,58	jul-13	139.047,14
feb-00	23.445,37	may-02	102.488,31	ago-04	119.281,00	nov-06	103.292,56	feb-09	130.734,90	may-11	105.862,85	ago-13	158.184,11
mar-00	32.984,70	jun-02	85.501,44	sep-04	67.630,36	dic-06	142.753,97	mar-09	140.662,96	jun-11	114.187,65	sep-13	111.418,96
abr-00	54.368,32	jul-02	100.727,78	oct-04	47.952,28	ene-07	104.248,66	abr-09	180.861,72	jul-11	128.180,79	oct-13	126.605,17
may-00	39.986,16	ago-02	111.600,03	nov-04	86.988,20	feb-07	98.716,67	may-09	119.126,28	ago-11	138.452,62	nov-13	108.095,13
jun-00	31.620,90	sep-02	73.435,04	dic-04	116.360,07	mar-07	105.967,70	jun-09	101.290,73	sep-11	106.263,66	dic-13	39.716,76
jul-00	36.247,89	oct-02	72.909,51	ene-05	81.785,64	abr-07	177.476,35	jul-09	75.261,06	oct-11	90.394,08	ene-14	113.713,94
ago-00	54.449,26	nov-02	93.880,56	feb-05	93.296,33	may-07	92.756,58	ago-09	125.003,01	nov-11	104.632,18	feb-14	97.498,21
sep-00	40.881,06	dic-02	115.659,15	mar-05	112.896,64	jun-07	97.336,55	sep-09	73.431,90	dic-11	200.070,00	mar-14	114.340,37
oct-00	29.933,37	ene-03	93.166,40	abr-05	56.596,25	jul-07	154.407,52	oct-09	91.489,14	ene-12	138.673,03	abr-14	154.318,55
nov-00	35.664,11	feb-03	80.713,18	may-05	79.148,05	ago-07	140.058,08	nov-09	98.943,77	feb-12	117.038,92	may-14	137.930,35
dic-00	61.661,44	mar-03	81.539,52	jun-05	84.029,36	sep-07	127.979,99	dic-09	144.410,42	mar-12	136.567,67		
ene-01	50.014,38	abr-03	102.020,61	jul-05	86.950,83	oct-07	107.972,77	ene-10	106.701,63	abr-12	200.888,02		
feb-01	27.740,17	may-03	86.482,65	ago-05	124.448,90	nov-07	126.791,63	feb-10	94.592,17	may-12	128.099,28		
mar-01	37.251,80	jun-03	92.376,53	sep-05	68.130,68	dic-07	169.664,90	mar-10	142.453,26	jun-12	145.785,12		

Anexo 7.5. Estimación de los parámetros de los modelos propuestos para la serie de ingresos económicos mensuales de los hoteles

7.5.1. ARIMA (2,1,1) x ARIMA (2,1,1)₁₂

```
Series: x
ARIMA(2,1,1)(2,1,1)[12]
Coefficients:
      ar1      ar2      ma1      sar1      sar2      sma1
    -0.1159  -0.0954  -0.6907  -0.0818  -0.2066  -0.7477
s.e.   0.1233   0.1059   0.1032   0.1207   0.1032   0.1084
sigma^2 estimated as 2.845e+10: log likelihood=-2322.07
AIC=4656.15  AICC=4656.83  BIC=4678.18

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1  -0.115862   0.123259  -0.9400   0.3472
ar2  -0.095406   0.105942  -0.9006   0.3678
ma1  -0.690674   0.103240  -6.6900  2.232e-11 ***
sar1  -0.081788   0.120674  -0.6778   0.4979
sar2  -0.206619   0.103212  -2.0019   0.0453 *
sma1  -0.747684   0.108428  -6.8957  5.361e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

7.5.2. ARIMA (2,1,1) x ARIMA (2,1,2)₁₂

```

Series: x
ARIMA(2,1,1)(2,1,2)[12]
Coefficients:
      ar1      ar2      ma1      sar1      sar2      sma1      sma2
-0.1224 -0.0932 -0.6736 -0.5743 -0.1948 -0.2413 -0.4405
s.e.    0.1301  0.1097  0.1122  0.2271  0.1089  0.2246  0.2039
sigma^2 estimated as 2.8e+10: log likelihood=-2321.02
AIC=4656.05  AICC=4656.93  BIC=4681.23

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1  -0.122357  0.130121 -0.9403  0.34704
ar2  -0.093191  0.109727 -0.8493  0.39571
ma1  -0.673573  0.112173 -6.0048  1.916e-09 ***
sar1 -0.574267  0.227135 -2.5283  0.01146 *
sar2 -0.194750  0.108907 -1.7882  0.07374 .
sma1 -0.241269  0.224558 -1.0744  0.28264
sma2 -0.440496  0.203874 -2.1606  0.03072 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

7.5.3. ARIMA (2,1,1) x ARIMA (1,1,1)₁₂

```

Series: x
ARIMA(2,1,1)(1,1,1)[12]
Coefficients:
      ar1      ar2      ma1      sar1      sma1
-0.1344 -0.0972 -0.6636  0.0325 -0.8708
s.e.    0.1294  0.1084  0.1106  0.1060  0.0917
sigma^2 estimated as 2.883e+10: log likelihood=-2323.89
AIC=4657.77  AICC=4658.28  BIC=4676.66

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1  -0.134389  0.129377 -1.0387  0.2989
ar2  -0.097169  0.108436 -0.8961  0.3702
ma1  -0.663637  0.110554 -6.0028  1.939e-09 ***
sar1  0.032462  0.105971  0.3063  0.7594
sma1 -0.870849  0.091700 -9.4967 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

7.5.4. ARIMA (2,1,1) x ARIMA (2,1,1)₁₂

```

Series: x
ARIMA(0,1,1)(2,1,0)[12]
Coefficients:
      ma1      sar1      sar2
-0.7384 -0.6253 -0.4507
s.e.    0.0487  0.0769  0.0733
sigma^2 estimated as 3.193e+10: log likelihood=-2328.5
AIC=4662.99  AICC=4663.23  BIC=4675.58

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ma1  -0.738433  0.048749 -15.1478 < 2.2e-16 ***
sar1 -0.625320  0.076901 -8.1315  4.240e-16 ***
sar2 -0.450741  0.073266 -6.1521  7.645e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Anexo 7.6. Estimación de los parámetros del mejor modelo de los modelos propuestos para la serie de ingresos económicos mensuales de los hoteles. Modelo con intervención: comienzo brusco y duración permanente.

7.6.1. ARIMA (0,1,1) x ARIMA (2,1,0)₁₂. Modelo con intervención :comienzo brusco y duración permanente

```
Series: hoteles.ts
ARIMA(0,1,1)(2,1,0)[12]
Coefficients:
      mal      sar1      sar2  step117-MA0
      -0.7632 -0.6261 -0.4688 -200008.53
s.e.    0.0492  0.0760  0.0732   98836.77
sigma^2 estimated as 3.113e+10:  log likelihood=-2326.59
AIC=4661.17  AICC=4661.54  BIC=4676.91

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
mal      -7.6318e-01  4.9242e-02 -15.4985 < 2.2e-16 ***
sar1      -6.2611e-01  7.5956e-02  -8.2431 < 2.2e-16 ***
sar2      -4.6879e-01  7.3173e-02  -6.4066  1.488e-10 ***
step117-MA0 -2.0001e+05  9.8837e+04  -2.0236  0.04301 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Anexo. 7.7. Código fuente del software libre R para cada una de las series estudiadas.

Anexo 7.7.1. Código fuente en R. Serie de ingresos económicos mensuales de los restaurantes.

```
library(lmtest)
library(astsa)
library(forecast)
library(TSA)
library(astsa)
library(xts)
library(car)
library(graphics)
library(tseries)
library(timeSeries)
library(tsoutliers)
#####
##### Modelos ARIMA #####
#####

## Se analiza los datos desde Enero de 1998 hasta junio de 2014
## Se observa la serie de los datos originales

restaurantes <- scan("restaurantes_ingresos.csv",sep=";", dec=",")
restaurantes.ts <- ts(restaurantes, frequency=12, start = c(1998,1))
restaurantes.ts
Dates <- seq(as.Date("1998-01-01"), as.Date("2014-06-01"), "month")
restaurantes.xts <- xts(restaurantes.ts, order.by=Dates)
plot(restaurantes.xts, type="o", main="Ingresos mensuales en restaurantes del municipio de La
Libertad. Enero 1998-Junio 2014", ylab="Ingresos en dólares", xlab="Meses")

## se descompone la serie original de restaurantes, se estima la tendencia, la componente
##estacional y aleatoria de la serie
restaurantes.ts.components <- decompose(restaurantes.ts)

## Coeficientes estacionales
round(restaurantes.ts.components$figure,2)
## Se observa que los mayores coeficientes estacionales están en Abril, Agosto y Diciembre.

## Gráfico de las componentes
plot(restaurantes.ts.components, main="Descomposición aditiva de la serie restaurantes")

## Se plantea un modelo de intervención, la crisis económica mundial de 2008-2015, comenzó en
## Septiembre de 2008, se clasifica como comienzo brusco y duración temporal
## se particiona la serie antes de la intervención t0 = 129, se analizan los datos antes de la
```

```

## intervención que corresponde al 64.65% de los datos. Suponiendo que la crisis fué en
## septiembre del 2008, tomando desde enero de 1998 hasta agosto de 2008
pre.int.restaurantes<- ts(restaurantes[1:128], frequency=12, start = c(1999,1))
pre.int.restaurantes.ts <- ts(pre.int.restaurantes, frequency=12, start = c(1998,1))
pre.int.restaurantes.ts
plot.ts(pre.int.restaurantes.ts, type="o", main="Ingresos mensuales en restaurantes del municipio
de La Libertad. Enero 1998- Agosto 2008", ylab="Ingresos en dolares", xlab="Meses")

## Se tiene la FAS y la FAP de la serie desde enero de 1998 hasta agosto 2008
acf(pre.int.restaurantes.ts,xlab="Retardos", ylab="FAS", main="Funcion de autocorrelacion simple
de restaurantes. Enero 1998 - Agosto 2008", lag.max=84)
pacf(pre.int.restaurantes.ts,xlab="Retardos", ylab="FAP",main="Funcion de autocorrelacion parcial
de restaurantes. Enero 1998 - Agosto 2008",lag.max=84)

## Para ver si la serie es estacionaria en varianza, se aplica el test de levenne.
temporal1 <- c(rep(1998:2007, each = 12) , rep(2008,8) )
temporal2 <- data.frame( x=pre.int.restaurantes.ts, group = temporal1 )
temporal2$group <- as.factor(temporal2$group)
leveneTest(x ~ group, data=temporal2)
## Interpretación: se rechaza Ho, es decir la serie no es estacionaria en varianza. Para ello se
##transforma la serie aplicando logaritmo.

log.pre.int.restaurantes.ts <- ts(log(pre.int.restaurantes.ts), frequency=12, start = c(1998,1))
#gráfico de la serie transformada
plot.ts(log.pre.int.restaurantes.ts, type="o", main="Ingresos mensuales en restaurantes
transformada", ylab="Ingresos en dólares", xlab="Meses")

## después de transformarla una vez, se aplica de nuevo el test de levenne para verificar
## si la serie es estacionaria en variancia.
temporal1 <- c(rep(1998:2007, each = 12) , rep(2008,8) )
temporal2 <- data.frame( x=log.pre.int.restaurantes.ts, group = temporal1 )
temporal2$group <- as.factor(temporal2$group)
leveneTest(x ~ group, data=temporal2)
## Interpretación: No se rechaza Ho, es decir la serie es estacionaria en varianza.

## Luego se aplica el tes de dickey fuller de raíces unitarias para corroborar estacionariedad en
##media.
adf.test(log.pre.int.restaurantes.ts)
## Interpretación: no se rechaza H0, por lo tanto no es estacionaria en media. Entonces es
## necesario aplicar una diferencia.
diff.log.pre.int.restaurantes.ts <- diff(log.pre.int.restaurantes.ts, differences=1)
#grafica de la serie pre-intervencion, transformada y diferenciada
plot.ts(diff.log.pre.int.restaurantes.ts, type="o", main="Ingresos mensuales en restaurantes
transformada y diferenciada", ylab="Ingresos en dolares", xlab="Meses")

## Se aplica nuevamente el test de Dikey Fuller
adf.test(diff.log.pre.int.restaurantes.ts)
## Interpretación: se rechaza H0, por lo tanto la serie es estacionaria en media.
## Parece que el proceso es estacionario en media y varianza.

#####
### ANÁLISIS DESCRIPTIVO DE LA SERIE###
#####
## Se observa la función de autocorrelación simple y parcial de la base tranformada y
## diferenciada en la parte regular
acf(diff.log.pre.int.restaurantes.ts,xlab="Retardos", ylab="FAS",main="Función de autocorrelación
simple de restaurantes (d=1)", lag.max=100)
pacf(diff.log.pre.int.restaurantes.ts,xlab="Retardos", ylab="FAP",main="función de
autocorrelación parcial de restaurantes (d=1)",lag.max=84)
## En la FAS se observan muchos picos significativos y ademas
## decae lentamente a cero, así que se aplica una diferencia estacional.

## Se observa la función de autocorrelación simple y parcial con una diferencia regular y una
## estacional
acf( diff(diff.log.pre.int.restaurantes.ts, lag = 12, differences = 1),xlab="Retardos",
ylab="FAS",main="Funcion de autocorrelacion simple de restaurantes (d=1 y D=1)", lag.max=100)
pacf(diff(diff.log.pre.int.restaurantes.ts, lag = 12, differences = 1),xlab="Retardos",
ylab="FAP",main="Funcion de autocorrelacion parcial de restaurantes (d=1 y D=1)",lag.max=100)

## Se tiene la serie transformada

```

```

log.restaurantes.ts <- ts(log(restaurantes.ts), frequency=12, start = c(1998,1))

#####
### modelo.nulo es el modelo sin la intervención #####
#####
modelo.nulo <- TSA::arima(log.restaurantes.ts, order = c(2,1,0) ,
                        seasonal = list(order = c(2, 1, 0), period = 12 )
                        ,method = 'ML')

print(modelo.nulo)
coefptest(modelo.nulo)

#####
##### modelos propuestos #####
#####
#Modelo principal
modelo.int <- arimax(log.restaurantes.ts, order = c(2, 1, 0), seasonal = list(order = c(2, 1, 0),
period = 12 ))
print(modelo.int)
coefptest(modelo.int)

##Modelo 2
modelo.int2 <- arimax(log.restaurantes.ts, order = c(2, 1, 0), seasonal = list(order = c(2, 1,
1), period=12 ))
print(modelo.int2)
coefptest(modelo.int2)

##Modelo 3
modelo.int3 <- arimax(log.restaurantes.ts, order = c(2, 1, 0), seasonal = list(order = c(2, 1,
2), period=12 ))
print(modelo.int3)
coefptest(modelo.int3)

##Modelo 4
modelo.int4 <- arimax(log.restaurantes.ts, order = c(2, 1, 1), seasonal = list(order = c(2, 1,
1), period=12 ))
print(modelo.int4)
coefptest(modelo.int4)

## Modelo 5
modelo.int5 <- arimax(log.restaurantes.ts, order = c(1, 1, 0), seasonal = list(order = c(2, 1,
1), period=12 ))
print(modelo.int5)
coefptest(modelo.int5)

#####
### ANÁLISIS DE RESIDUOS###
#####

## Gráfico de Residuos
plot(residuals(modelo.int), main="Residuos de modelo ARI(2,1)XARI(2,1)", ylab="Ingresos en
dólares", xlab="Meses")
## FAS y FAP de los residuos
acf(residuals(modelo.int),xlab="Retardos", ylab="FAS",main="Autocorrelaciones de los residuos.
Modelo ARI(2,1)XARI(2,1)", lag.max=84)
pacf(residuals(modelo.int),xlab="Retardos", ylab="FAP",main="Autocorrelaciones parciales de los
residuos. Modelo ARI(2,1)XARI(2,1)", lag.max=84)
## Los residuos están incorrelados

## Histograma de los residuos
h<-hist(modelo.int$residuals, breaks=10, col="red",ylab="Frecuencia", xlab="Error para la serie
generados con ARI(2,1)XARI(2,1)", main="Histograma residual ARI(2,1)XARI(2,1)")
xfit<-seq(min(modelo.int$residuals),max(modelo.int$residuals),length=40)
yfit<-dnorm(xfit,mean=mean(modelo.int$residuals),sd=sd(modelo.int$residuals))
yfit <- yfit*diff(h$mids[1:2])*length(modelo.int$residuals)
lines(xfit, yfit, col="black", lwd=2)

qqnorm(modelo.int$residuals, col=2, xlab="Valores observados", ylab="Valores esperados de la
normal", main="Probabilidad residual ARI(2,1)XARI(2,1)")
qqline(modelo.int$residuals, lwd=2,lty=1)

```

```

#####
##### Modelos con intervención #####
#####

#####
#### Comienzo brusco duración permanente #####
#####

modelo1.int <- arimax(log.restaurantes.ts, order = c(2, 1, 0), seasonal = list(order = c(2, 1,
0), period = 12 ),
                    xtrans=data.frame(step129=1*(seq(log.restaurantes.ts))>=129),
                    transfer=list(c(0,0)), method = 'ML' )

print(modelo1.int)

#significancia de los parámetros
coefstest(modelo1.int)

## Con el coeficiente de step129-MA0 se puede ver que este termino
## no es significativo, por tanto no se analizan los residuos.

#####
#### Comienzo gradual duración permanente #####
#####

### Empieza en el periodo t
modelo2.int <- arimax(log.restaurantes.ts, order = c(2, 1, 0), seasonal = list(order = c(2, 1,
0), period = 12 ),
                    xtrans=data.frame(step129=1*(seq(log.restaurantes.ts))>=129),
                    transfer=list(c(1,0)) , method = 'ML' )
print(modelo2.int)
## Se tienen problemas con el software, por tanto no se calcula

#####
#### Comienzo brusco duración temporal #####
#####

### Nota como es funcion pulso se cambia el mayor o igual con igual

modelo3.int <- arimax(log.restaurantes.ts, order = c(2, 1, 0), seasonal = list(order = c(2, 1,
0), period = 12 ),
                    xtrans=data.frame(step129=1*(seq(log.restaurantes.ts))==129),
                    transfer=list(c(1,0)) , method = 'ML' )

print(modelo3.int)
coefstest(modelo3.int)

## Gráfico de los residuos del mejor modelo con intervención
plot(residuals(modelo3.int), main="Residuos de modelo ARI(2,1)xARI(2,1) con intervención ",
      ylab="Ingresos en dólares", xlab="Meses")

## FAS y FAP de los residuos del mejor modelo con intervención
acf(residuals(modelo3.int),xlab="Retardos", ylab="FAS",main="Autocorrelación simple
residual.Modelo ARI(2,1)xARI(2,1) con intervención", lag.max=84)
pacf(residuals(modelo3.int),xlab="Retardos", ylab="FAP",main="Autocorrelaciones parciales de los
residuos. Modelo ARI(2,1)xARI(2,1) con intervención", lag.max=84)

## Histograma de los residuos del mejor modelo con intervención
h<-hist(modelo3.int$residuals, breaks=10, col="red",ylab="Frecuencia", xlab="Error para la serie
generados con ARI(2,1)xARI(2,1)", main="Histograma residual con intervencion ARI(2,1)xARI(2,1)")
xfit<-seq(min(modelo3.int$residuals),max(modelo3.int$residuals),length=40)
yfit<-dnorm(xfit,mean=mean(modelo3.int$residuals),sd=sd(modelo3.int$residuals))
yfit <- yfit*diff(h$mids[1:2])*length(modelo3.int$residuals)
lines(xfit, yfit, col="black", lwd=2)

#gráfico de probabilidad residual
qqnorm(modelo3.int$residuals, col=2, xlab="Valores observados", ylab="valores esperados de la
normal", main="Probabilidad residual con intervencion ARI(2,1)xARI(2,1)")
qqline(modelo3.int$residuals, lwd=2,lty=1)

```

```
#####
### Comienzo gradual duración temporal #####
#####
### Nota: como es función pulso se cambia el mayor o igual

modelo4.int <- arimax(log.restaurantes.ts, order = c(2, 1, 0), seasonal = list(order = c(2, 1,
0), period = 12 ),
                    xtrans=data.frame(step129=1*(seq(log.restaurantes.ts))==129),
                    transfer=list(c(2,0)), method = 'ML' )

print(modelo4.int)
coefstest(modelo4.int)

plot(residuals(modelo4.int), main="Residuos de modelo ARI(2,1)xARI(2,1) con intervención ",
ylab="Ingresos en dólares", xlab="Meses")
acf(residuals(modelo4.int),xlab="Retardos", ylab="FAS", main="FAS de los residuos. Modelo
ARI(2,1)xARI(2,1) con intervención", lag.max=84)
pacf(residuals(modelo4.int),xlab="Retardos", ylab="FAP",main="FAP de los residuos. Modelo
ARI(2,1)xARI(2,1) con intervención", lag.max=84)

## Histograma residual
h<-hist(modelo4.int$residuals, breaks=10, col="red", ylab="Frecuencia",xlab="Error para la serie
generados con ARI(2,1)xARI(2,1)", main="Histograma residual con intervención ARI(2,1)xARI(2,1)")
xfit<-seq(min(modelo4.int$residuals),max(modelo4.int$residuals),length=40)
yfit<-dnorm(xfit,mean=mean(modelo4.int$residuals),sd=sd(modelo4.int$residuals))
yfit <- yfit*diff(h$mids[1:2])*length(modelo4.int$residuals)
lines(xfit, yfit, col="black", lwd=2)

#probabilidad residual
qqnorm(modelo4.int$residuals, col=2, xlab="Valores observados", ylab="Valores esperados de la
normal", main="Probabilidad residual con intervención ARI(2,1)xARI(2,1)")
qqline(modelo4.int$residuals, lwd=2,lty=1)

### El mejor modelo comienzo brusco -- duración temporal ###

## Gráfico de la serie transformada
plot(log.restaurantes.ts,ylab="Restaurantes (ingresos en miles de $)", xlab="Meses", main="Serie
transformada. Modelo ARI(2,1)xARI(2,1) con intervención")
points(fitted(modelo3.int))

## Detección de Outliers
locate.outliers(residuals(modelo3.int), pars, cval = 3.5, types = c("AO", "LS", "TC"),
                delta = 0.7, n.start = 50)
which(abs(tstats[,"AO","tstat"]) > 3.5)
which(abs(tstats[,"TC","tstat"]) > 3.5)
which(abs(tstats[,"LS","tstat"]) > 3.5)

plot(residuals(modelo3.int), xlab="Fecha", ylab="Residuos", main="Representación de vedas por
marea roja")

#Periodos de veda por marea roja
outlier_a <- stats::window(residuals(modelo3.int), start = c(2001,8), end=c(2002,1) )
points( outlier_a , pch = 21, col = "black" , bg = "blue")

outlier_b <- stats::window(residuals(modelo3.int), start = c(2005,11), end=c(2006,3) )
points( outlier_b , pch = 21, col = "black" , bg = "yellow")

outlier_c <- stats::window(residuals(modelo3.int), start = c(2007,8), end=c(2007,9) )
points( outlier_c , pch = 21, col = "black" , bg = "blue")

outlier_d <- stats::window(residuals(modelo3.int), start = c(2009,11), end=c(2010,4) )
points( outlier_d , pch = 21, col = "black" , bg = "yellow")

outlier_e <- stats::window(residuals(modelo3.int), start = c(2010,10), end=c(2010,10) )
points( outlier_e , pch = 21, col = "black" , bg = "blue")

outlier_f <- stats::window(residuals(modelo3.int), start = c(2011,8), end=c(2011,9) )
points( outlier_f , pch = 21, col = "black" , bg = "yellow")

outlier_g <- stats::window(residuals(modelo3.int), start = c(2012,3), end=c(2012,5) )
```

```

points( outlier_g , pch = 21, col = "black" , bg = "blue")

outlier_h <- stats::window(residuals(modelo3.int), start = c(2012,9), end=c(2012,9) )
points( outlier_h , pch = 21, col = "black" , bg = "yellow")

outlier_i <- stats::window(residuals(modelo3.int), start = c(2013,9), end=c(2013,10) )
points( outlier_i , pch = 21, col = "black" , bg = "blue")

## Detección de Outliers
plot(residuals(modelo3.int), xlab="Fecha", ylab="Residuos", main="Detección de datos atípicos")
outlier24 <- stats::window(residuals(modelo3.int), start = c(1999,12), end=c(1999,12) )
points( outlier24 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[24]

outlier25 <- stats::window(residuals(modelo3.int), start = c(2000,1), end=c(2000,1) )
points( outlier25 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[25]

outlier36 <- stats::window(residuals(modelo3.int), start = c(2000,12), end=c(2000,12) )
points( outlier36 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[36]

outlier37 <- stats::window(residuals(modelo3.int), start = c(2001,1), end=c(2001,1) )
points( outlier37 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[37]

outlier38 <- stats::window(residuals(modelo3.int), start = c(2001,2), end=c(2001,2) )
points( outlier38 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[38]

outlier40 <- stats::window(residuals(modelo3.int), start = c(2001,4), end=c(2001,4) )
points( outlier40 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[40]

outlier98 <- stats::window(residuals(modelo3.int), start = c(2006,2), end=c(2006,2) )
points( outlier98 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[98]

outlier99 <- stats::window(residuals(modelo3.int), start = c(2006,3), end=c(2006,3) )
points( outlier99 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[99]

outlier118 <- stats::window(residuals(modelo3.int), start = c(2007,10), end=c(2007,10) )
points( outlier118 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[118]

outlier135 <- stats::window(residuals(modelo3.int), start = c(2009,3), end=c(2009,3) )
points( outlier135 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[135]

outlier145 <- stats::window(residuals(modelo3.int), start = c(2010,1), end=c(2010,1) )
points( outlier145 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[145]

outlier154 <- stats::window(residuals(modelo3.int), start = c(2010,10), end=c(2010,10) )
points( outlier154 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[154]

outlier184 <- stats::window(residuals(modelo3.int), start = c(2013,4), end=c(2013,4) )
points( outlier184 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[184]

outlier192 <- stats::window(residuals(modelo3.int), start = c(2013,12), end=c(2013,12) )
points( outlier192 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[192]

outlier193 <- stats::window(residuals(modelo3.int), start = c(2014,1), end=c(2014,1) )
points( outlier193 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[193]

```

```

outlier198 <- stats::window(residuals(modelo3.int), start = c(2014,6), end=c(2014,6) )
points( outlier198 , pch = 21, col = "blue" , bg = "red")
residuals(modelo3.int)[198]

## Gráfico de intersección de outliers y vedas
plot(residuals(modelo3.int), xlab="Fecha", ylab="Residuos", main="Interseccion de outliers y
Vedas por marea roja")

outlier98 <- stats::window(residuals(modelo3.int), start = c(2006,2), end=c(2006,2) )
points( outlier98 , pch = 21, col = "red" , bg = "red")
residuals(modelo3.int)[98]

outlier99 <- stats::window(residuals(modelo3.int), start = c(2006,3), end=c(2006,3) )
points( outlier99 , pch = 21, col = "red" , bg = "red")
residuals(modelo3.int)[99]

outlier154 <- stats::window(residuals(modelo3.int), start = c(2010,10), end=c(2010,10) )
points( outlier154 , pch = 21, col = "red" , bg = "red")
residuals(modelo3.int)[154]

```

Anexo 7.7.2. Código fuente en R. Serie de ingresos económicos mensuales de los hoteles.

```

#se cargan las librerías
library(astsa)
library(forecast)
library(TSA)
library(astsa)
library(fma)
library(xts)
library(lme4)
library(SparseM)
library(pbkrtest)
library(quantreg)
library(car)
library(graphics)
library(tseries)
library(timeSeries)
library(tsoutliers)
library(lmtest)
#####
# Modelos ARIMA #####
#####
# # Se analizan los datos desde Enero de 1999 hasta mayo de 2014

#se carga la serie de los datos originales
hoteles <- scan("hoteles_ingresos.csv",sep=";", dec=",")
hoteles.ts <- ts(hoteles, frequency=12, start = c(1999,1))
hoteles.ts
Dates <- seq(as.Date("1999-01-01"), as.Date("2014-05-01"), "month")
hoteles.xts <- xts(hoteles.ts, order.by=Dates)
par(mar = rep(2, 4))

#gráfico de la serie original
plot(hoteles.xts, type="o", main="Ingresos mensuales en hoteles del municipio de La Libertad.
Enero 1999-Mayo 2014", ylab="Ingresos en dólares", xlab="Meses")

# se descompone la serie original de hoteles
# se estima la tendencia, la componente estacional y aleatoria de la serie
hoteles.ts.components <- decompose(hoteles.ts)
# Coeficientes estacionales
round(hoteles.ts.components$figure,2)
# se observa que los mayores coeficientes estacionales
# están en Marzo, Agosto y Diciembre.

# Gráfico de las componentes
plot(hoteles.ts.components,xlab="Fecha")

```

```

# se plantea un modelo de intervención
# la crisis económica mundial de 2008-2015, empezó en Septiembre de 2008
# se clasifica como comienzo brusco y duración permanente
## se particiona la serie antes y después de la intervención, t0 = 117
## se analizan los datos antes de la intervención que corresponde al 62.7% de los datos.
## suponiendo que la crisis fue en septiembre del 2008, tomando desde enero de 1999 hasta agosto
## de 2008
pre.int.hoteles<- ts((hoteles.ts[1:116]), frequency=12, start = c(1999,1))
pre.int.hoteles.ts <- ts(pre.int.hoteles, frequency=12, start = c(1999,1))
pre.int.hoteles.ts
plot.ts(pre.int.hoteles.ts, type="o", main="Ingresos mensuales en hoteles del municipio de La
Libertad. Enero 1999- Agosto 2008", ylab="Ingresos en dólares", xlab="Meses")

#correlogramas de la serie (serie antes de la intervención)
acf(pre.int.hoteles.ts,xlab="Retardos", ylab="FAS",main="FAS de hoteles (Ene 1999- Agosto
2008)",lag.max=60)
pacf(pre.int.hoteles.ts,xlab="Retardos", ylab="FAP",main="FAP de hoteles (Ene 1999- Agosto
2008)",lag.max=60)

## Para ver si la serie es estacionaria en varianza, se aplica el test de levenne.
temporal1 <- c(rep(1999:2007, each = 12) , rep(2008,8) )
temporal2 <- data.frame( x=pre.int.hoteles.ts, group = temporal1 )
temporal2$group <- as.factor(temporal2$group)
leveneTest(x ~ group, data=temporal2)
rm(temporal1, temporal2)
## Interpretación: no se rechaza Ho, es decir la serie es estacionaria en varianza, no se
## necesita transformar.

## se aplica el tes de dickey fuller de raices unitarias
## para corroborar estacionariedad en media.
adf.test(pre.int.hoteles.ts)
## Interpretación: no se rechaza H0, por lo tanto no es estacionaria en media. Entonces es
##necesario aplicar una diferencia en la parte regular.

diff.pre.int.hoteles.ts <- diff(pre.int.hoteles.ts, differences=1)
#gráfico de la serie con una diferencia
plot.ts(diff.pre.int.hoteles.ts, type="o", main="Serie de hoteles diferenciada (d=1)",
ylab="Ingresos en dolares", xlab="Meses")

## se aplica nuevamente el test de Dikey Fuller
adf.test(diff.pre.int.hoteles.ts)
## Interpretación: se rechaza H0, por lo tanto es estacionaria en media.
# Parece que el proceso es estacionario en media y varianza.

#correlogramas con una diferencia en la parte regular (serie pre-intervención)
acf(diff.pre.int.hoteles.ts,xlab="Retardos", ylab="FAS",main="Pre intervención, Hoteles (d=1)",
lag.max=84)
pacf(diff.pre.int.hoteles.ts,xlab="Retardos", ylab="FAP",main="Pre intervención, Hoteles
(d=1)",lag.max=60)

## En la FAS se observan muchos picos significativos y además
## decae lentamente a cero, así que se aplica una diferencia estacional.
#gráfico con d=1 y D=1
acf( diff(diff.pre.int.hoteles.ts, lag=12, differences = 1),xlab="Retardos", ylab="FAS",main="Pre
intervención, hoteles (d=1 Y D=1)",lag.max=60)
pacf(diff(diff.pre.int.hoteles.ts, lag=12, differences = 1),xlab="Retardos", ylab="FAP",main="Pre
intervención, hoteles (d=1 Y D=1)",lag.max=60)
#es estacionaria en la parte estacional, al menos ya no evidencia picos significativos en los
retardos estacionales.

#####
##Propuesta de modelos
#####

hoteles.ts <- ts(hoteles.ts, frequency=12, start = c(1999,1))

#modelo 1, ARIMA(2,1,1)X(2,1,1)12
modelo.int <- arimax(hoteles.ts, order = c(2,1,1) ,
seasonal = list(order = c(2, 1, 1), period = 12))

```

```

print(modelo.int)
#significancia de los parámetros
coefstest(modelo.int)

#modelo 2
#ARIMA(2,1,1)x(2,1,2)12
modelo2.int <- arimax(hoteles.ts, order = c(2,1,1) ,
                    seasonal = list(order = c(2, 1, 2), period = 12))
print(modelo2.int)
coefstest(modelo2.int)

#modelo 3
#ARIMA(2,1,1)x(1,1,1)12
Modelo3.int <- arimax(hoteles.ts, order = c(2,1,1) ,
                    seasonal = list(order = c(1, 1, 1), period = 12))
print(modelo3.int)
coefstest(modelo3.int)

##modelo 4 (modelo.nulo)
hoteles.ts <- ts(hoteles.ts, frequency=12, start = c(1999,1))
modelo.nulo <- arimax(hoteles.ts, order = c(0,1,1) ,
                    seasonal = list(order = c(2, 1, 0), period = 12))
print(modelo.nulo)
coefstest(modelo.nulo)

#####
## Gráfico de los residuos del modelo.nulo, sin intervención, mejor modelo de los propuestos
#####

hoteles.ts <- ts(hoteles.ts, frequency=12, start = c(1999,1))
modelo.nulo <- arimax(hoteles.ts, order = c(0,1,1) ,
                    seasonal = list(order = c(2, 1, 0), period = 12))

#gráfico residual
plot(residuals(modelo.nulo),xlab="Fecha",ylab="Residuos",main="Modelo IMA(1,1) x ARI (2,1)")

#FAS Residual con el modelo IMA(1,1)xARI(2,1)
acf(residuals(modelo.nulo), xlab="Retardos",ylab="FAS",main="FAS de los residuos.
IMA(1,1)xARI(2,1)", lag.max = 84)

#FAP Residual con el modelo IMA(1,1)xARI(2,1)
pacf(residuals(modelo.nulo),xlab="Retardos",ylab="FAP", main="FAP de los residuos.
IMA(1,1)xARI(2,1)",lag.max = 84)

#Histograma Residual
h<-hist(modelo.nulo$residuals, breaks=20, col="red",xlab="Error para la serie
generada",ylab="Frecuencia",
main="Histograma residual. Modelo sin intervención: IMA(1,1)xARI(2,1)")
xfit<-seq(min(modelo.nulo$residuals),max(modelo.nulo$residuals),length=40)
yfit<-dnorm(xfit,mean=mean(modelo.nulo$residuals),sd=sd(modelo.nulo$residuals))
yfit <- yfit*diff(h$mids[1:2])*length(modelo.nulo$residuals)
lines(xfit, yfit, col="black", lwd=2)

#Probabilidad residual
qqnorm(modelo.nulo$residuals,xlab="Valores observados",ylab="Valores esperados de la normal",
main="Probabilidad Residual (sin intervención): IMA(1,1)xARI(2,1)")
qqline(modelo.nulo$residuals, col = 2,lwd=2,lty=1)

#####
###Comienzo brusco, duración permanente #####
#####
modelo.int <- arimax(hoteles.ts, order = c(0, 1, 1), seasonal = list(order = c(2, 1, 0), period =
12 ),
                    xtrans=data.frame(step117=1*(seq(hoteles.ts))>=117),
                    transfer=list(c(0,0)), method = 'ML')
print(modelo.int)
### La siguiente función es mejor para ver la significancia
coefstest(modelo.int)

```

```
#####
## Gráfico de los residuos, modelo con intervención
#####
### El mejor modelo comienzo brusco, duración permanente ###

#residuos del modelo
plot(residuals(modelo.int),xlab="Fecha",ylab="Residuos",main="Modelo IMA(1,1) x ARI (2,1) con
intervención")

#FAS Residual con el modelo IMA(1,1)xARI(2,1)
acf(residuals(modelo.int), xlab="Retardos",ylab="FAS",main="FAS de los residuos.
IMA(1,1)xARI(2,1) con intervención", lag.max = 84)

#FAP Residual con el modelo IMA(1,1)xARI(2,1)
pacf(residuals(modelo.int),xlab="Retardos",ylab="FAP", main="FAP de los residuos.
IMA(1,1)xARI(2,1) con intervención",lag.max = 84)

#Histograma Residual
h<-hist(modelo.int$residuals, breaks=20, col="red",xlab="Error para la serie
generada",ylab="Frecuencia",
main="Histograma residual. Modelo IMA(1,1)xARI(2,1)con intervención")
xfit<-seq(min(modelo.int$residuals),max(modelo.int$residuals),length=40)
yfit<-dnorm(xfit,mean=mean(modelo.int$residuals),sd=sd(modelo.int$residuals))
yfit <- yfit*diff(h$mids[1:2])*length(modelo.int$residuals)
lines(xfit, yfit, col="black", lwd=2)

#Probabilidad residual
qqnorm(modelo.int$residuals,xlab="Valores observados",ylab="Valores esperados de la normal",
main="Probabilidad Residual. IMA(1,1)xARI(2,1) con intervención")
qqline(modelo.int$residuals, col = 2,lwd=2,lty=1)

#####
plot(hoteles.ts,ylab="Fecha", xlab="Meses",main="Serie original y modelo IMA(1,1)xARI(2,1) con
intervención")
points(fitted(modelo.int))
#####
### Detección de outliers
## Almacena los coeficientes del modelo en pars
pars <- coefs2poly(modelo.int)
resid <- residuals(modelo.int)
tstats <- outliers.tstatistics(pars, resid)

locate.outliers(residuals(modelo.int), pars, cval = 3.5, types = c("AO", "LS", "TC"),
delta = 0.7, n.start = 50)

#se confirman los outliers identificados anteriormente por medio de las siguientes instrucciones
which(abs(tstats[,"AO","tstat"]) > 3.5)
which(abs(tstats[,"TC","tstat"]) > 3.5)
which(abs(tstats[,"LS","tstat"]) > 3.5)

#valor del residuo de cada outlier
residuals(modelo.int)
residuals(modelo.int)[34]
residuals(modelo.int)[127]
residuals(modelo.int)[156]
residuals(modelo.int)[179]
residuals(modelo.int)[180]

#asignación de la fecha respectiva de los outliers
outlier1 <- stats:window(residuals(modelo.int), start = c(2001,10), end=c(2001,10) )
outlier2 <- stats:window(residuals(modelo.int), start = c(2009,7), end=c(2009,7) )
outlier3 <- stats:window(residuals(modelo.int), start = c(2011,12), end=c(2011,12) )
outlier4 <- stats:window(residuals(modelo.int), start = c(2013,11), end=c(2013,11) )
outlier5 <- stats:window(residuals(modelo.int), start = c(2013,12), end=c(2013,12) )

#gráfico de los residuos con los outliers
plot(residuals(modelo.int),xlab="Fecha", ylab="Residuos", main="Detección de datos Atípicos")
points( outlier1 , pch = 21, col = "black" , bg = "red")
points( outlier2 , pch = 21, col = "black" , bg = "red")
points( outlier3 , pch = 21, col = "black" , bg = "red")
```

```

points( outlier4 , pch = 21, col = "black" , bg = "red")
points( outlier5 , pch = 21, col = "black" , bg = "red")

#gráfico de periodos de veda por marea roja
plot(residuals(modelo.int),xlab="Fecha", ylab="Residuos", main="Representación de periodos de
veda por Marea Roja y outliers")

outlier_a <- stats::window(residuals(modelo.int), start = c(2001,8), end=c(2002,1) )
points( outlier_a , pch = 21, col = "black" , bg = "blue")

outlier_b <- stats::window(residuals(modelo.int), start = c(2005,11), end=c(2006,3) )
points( outlier_b , pch = 21, col = "black" , bg = "yellow")

outlier_c <- stats::window(residuals(modelo.int), start = c(2007,8), end=c(2007,9) )
points( outlier_c , pch = 21, col = "black" , bg = "blue")

outlier_d <- stats::window(residuals(modelo.int), start = c(2009,11), end=c(2010,4) )
points( outlier_d , pch = 21, col = "black" , bg = "yellow")

outlier_e <- stats::window(residuals(modelo.int), start = c(2010,10), end=c(2010,10) )
points( outlier_e , pch = 21, col = "black" , bg = "blue")

outlier_f <- stats::window(residuals(modelo.int), start = c(2011,8), end=c(2011,9) )
points( outlier_f , pch = 21, col = "black" , bg = "yellow")

outlier_g <- stats::window(residuals(modelo.int), start = c(2012,3), end=c(2012,5) )
points( outlier_g , pch = 21, col = "black" , bg = "blue")

outlier_h <- stats::window(residuals(modelo.int), start = c(2012,9), end=c(2012,9) )
points( outlier_h , pch = 21, col = "black" , bg = "yellow")

outlier_i <- stats::window(residuals(modelo.int), start = c(2013,9), end=c(2013,10) )
points( outlier_i , pch = 21, col = "black" , bg = "blue")

```