



COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN
EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSGRADO EN HIDROCIENCIAS

**Estimación de Aportaciones Deducidas a Presas
Mediante Análisis de Series de
Tiempo Estacionales**

Isaías Ariza Hernández

T E S I S

PRESENTADA COMO REQUISITO PARCIAL PARA
OBTENER EL GRADO DE:

MAESTRO EN CIENCIAS

MONTECILLO, TEXCOCO, EDO. DE MÉXICO
2012

La presente tesis titulada: **Estimación de Aportaciones Deducidas a Presas Mediante Análisis de Series de Tiempo Estacionales**, realizada por el alumno: **Isaías Ariza Hernández**, bajo la dirección del Consejo Particular indicado ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

MAESTRO EN CIENCIAS

POSGRADO EN HIDROCIENCIAS

CONSEJO PARTICULAR

CONSEJERO _____
Dr. Jesús Chávez Morales

ASESOR _____
Dr. Juan M. González Camacho

ASESOR _____
Dra. Laura A. Ibáñez Castillo

Estimación de Aportaciones Deducidas a Presas Mediante Análisis de Series de Tiempo Estacionales

Isaías Ariza Hernández

Colegio de Postgraduados, 2012

Para el estudio de las aportaciones deducidas, que se pueden esperar en presas, se analizan con modelos matemáticos basados en la estadística, con los que se trata de inferir el comportamiento de las variables hidrológicas, que en este caso, son las aportaciones deducidas. Dichos modelos pueden ser modelos de series de tiempo. El objetivo fundamental de este trabajo es presentar una metodología de análisis de series de tiempo estacionales cuyo dominio es el tiempo propuesto por Box y Jenkins (1970). La importancia del estudio de las aportaciones radica principalmente en la estimación del volumen de agua que se espera escurra en una corriente con fines de planeación de los sistemas de irrigación. La metodología fue aplicada al registro histórico de aportaciones deducidas mensuales de la presa Adolfo Ruiz Cortines (1955-2011), que es la fuente de agua que abastece al distrito de Riego 038 Río Mayo, en el estado de Sonora, México. La estrategia de construcción de modelos autorregresivos e integrados de promedios móviles estacionales (SARIMA, por sus siglas en inglés), para series de tiempo univariadas consta de un proceso iterativo de cuatro etapas fundamentales: Identificación del modelo, Estimación de los parámetros del modelo, Verificación del modelo y Uso del modelo para fines de pronóstico. El modelo obtenido fue un $SARIMA(0, 1, 2)(3, 1, 0)_{12}$, los parámetros se estimaron por el método de máxima verosimilitud, y se realizaron pronósticos para dos ciclos agrícolas 2011-2012 y 2012-2013.

Palabras clave: Aportaciones Deducidas, Series de tiempo estacionales, Modelos Box-Jenkins.

Surface Inflows to Reservoirs Estimation Using Analysis of Seasonal Time Series

Isaías Ariza Hernández

Colegio de Postgraduados, 2005

The importance of studying the surface inflows to reservoirs, lies primarily in the estimation of water volume expected enter into reservoirs in coming times, with purposes of planning the irrigation districts operation planning. Surface inflows, which can be expected to enter into reservoirs, are analyzed with mathematical models based on statistical analyses. In Mexico, usually, frequency analysis is applied to estimate expected surface inflows for next planning time periods; however, another way to do that, could be with time series analysis. A methodology, as that proposed by Box and Jenkins, of application of seasonal time series modeling with prediction purpose of the behavior of the surface inflows to reservoirs, is presented. An iterative process for model building was followed in four fundamental stages: Identification, Estimation of model parameters, Diagnostic checking and Forecast with model. A seasonal, autoregressive, integrated, moving average univariate time-series, model SARIMA(0,1,2)(3,1,0)₁₂, was developed. The parameters of the model, which resulted statistically adequate, were estimated by the maximum likelihood method. The methodology was applied to the historical, (1955-2011), data of monthly deduced inflows to the Adolfo Ruiz Cortines reservoir which is the source of water for the 038 Mayo River Irrigation District, located in Sonora, Mexico. Predictions were made for two agricultural cycles 2011-2012 and 2012-2013; the results when compared with deduced inflows averages and those estimated by frequency analysis were acceptably consistent.

Key words: Surface Inflows, Reservoirs, SARIMA, Time-Series, Irrigation Districts.

AGRADECIMIENTOS

Al Colegio de Postgraduados, por haberme brindado la oportunidad de seguir mi formación académica y personal en sus aulas. Mi agradecimiento y reconocimiento SIEMPRE.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico brindado durante la realización de mis estudios.

Al Consejo Mexiquense de Ciencia y Tecnología (COMECYT) por el apoyo económico otorgado para realizar mi trabajo de tesis de grado.

A los integrantes de mi Consejo Particular:

Al Dr. Jesús Chávez Morales mi más sincero agradecimiento, por sus atinadas indicaciones y consejos mismos que me fueron de gran utilidad en la realización de mis estudios de maestría y por la magistral asesoría del presente trabajo de tesis.

Al Dr. Juan M. González Camacho y la Dra. Laura A. Ibáñez Castillo por la orientación, apoyo y colaboración desinteresada en el presente trabajo.

A mis profesores, compañeros y amigos de clase que de alguna manera fueron copartícipes de esta tarea, a todos gracias.

A mi familia mi más sincero agradecimiento:

Por el cariño y apoyo que durante mi estancia estando lejos de ellos he encontrado con enorme certeza. Es preciso destacar el profundo agradecimiento a mi Sr. padre Francisco Ariza, Hermanos, a mi querida Tía Petra y por supuesto a mi Esposa Coral, ya que cada uno de ustedes fueron, son y serán un aliciente en mi camino ¡GRACIAS!.

DEDICATORIA

A la memoria de mi adorada madre Francisca Hernández Cortés. . .

A mi padre Francisco Ariza Behéna. . .

A Coral, por su inmensurable amor. . .

Índice de Contenido

1. INTRODUCCIÓN	1
2. HIPÓTESIS y OBJETIVOS	3
2.1. HIPÓTESIS	3
2.2. OBJETIVOS	3
3. REVISIÓN DE LITERATURA	4
3.1. ANTECEDENTES	4
3.1.1. Uso de modelos estocásticos en análisis hidrológico	5
3.2. ESCURRIMIENTO SUPERFICIAL	6
3.2.1. Componentes del Esgurrimiento	7
3.2.2. Ciclo del Esgurrimiento	7
3.2.3. Aportaciones por Esgurrimiento	8
3.2.4. Cálculo de Aportaciones	9
3.2.5. Cálculo de aportaciones deducidas en los Distritos de Riego	10
3.3. MODELOS BOX-JENKINS	12
3.3.1. Definiciones y Notación	13
3.3.2. Procesos estocásticos lineales	15

Índice de Contenido

3.3.3. Procesos estacionarios	16
3.3.4. Medición del error de pronostico	18
3.3.5. Modelos Box-Jenkins o Modelos ARIMA	20
3.3.6. Procedimiento de elaboración de los modelos Box-Jenkins	21
3.3.7. Identificación del modelo	22
3.3.8. Estimación de parámetros	23
3.3.9. Verificación del modelo	24
3.4. ANÁLISIS DE SERIES DE TIEMPO ESTACIONALES.	26
3.4.1. Modelos ARIMA estacionales	27
3.4.2. Identificación de modelos estacionales	28
3.4.3. Construcción de modelos estacionales	29
3.5. PRONÓSTICO CON MODELOS ARIMA	30
4. MATERIALES Y MÉTODOS	31
4.1. MATERIALES	31
4.1.1. Distrito de Riego 038 Río Mayo	31
4.1.2. Presa Adolfo Ruiz Cortines	33
4.1.3. Estaciones hidrométricas en DR 038 Río Mayo	37
4.1.4. Red de distribución de aguas	38
4.1.5. Información hidrométrica	39
4.1.6. Software utilizado	39
4.2. MÉTODOS	40
4.2.1. Obtención de la información hidrométrica	40

Índice de Contenido

4.2.2. Desarrollo del modelo de series de tiempo: Box-Jenkins	41
5. RESULTADOS	59
5.1. INFORMACIÓN DE APORTACIONES DEDUCIDAS	59
5.1.1. Elaboración del Hidrograma de aportaciones históricas mensuales	60
5.2. MODELOS ESTACIONALES DE SERIES DE TIEMPO: Box-Jenkins .	61
5.2.1. Estacionaridad de la Serie	61
5.2.2. Identificación del modelo estacional	62
5.2.3. Estimación de parámetros del modelo seleccionado	62
5.2.4. Verificación del modelo	63
5.2.5. Predicción de aportaciones a la presa ARC	66
5.3. COMPARACIÓN ENTRE VALORES AJUSTADOS Y OBSERVADOS .	68
5.4. COMPARACIÓN CON ANÁLISIS DE FRECUENCIA	70
5.4.1. Pronóstico con la Función de Distribución de Probabilidad: Log Normal de 2 parámetros	70
6. CONCLUSIONES Y RECOMENDACIONES	72
6.1. CONCLUSIONES	72
6.2. RECOMENDACIONES	73
Nomenclatura	74
Referencias	75
Anexos	78
Anexo A: Estimación de Aportaciones Deducidas en distritos de Riego	78

Índice de Contenido

Anexo B: Registro Histórico de Aportaciones Deducidas a la Presa Adolfo Ruiz Cortines	85
Anexo C: Rutinas desarrolladas en R	87
Anexo D: Tablas utilizadas: t – <i>student</i> y J_i – <i>cuadrada</i>	92

Índice de Cuadros

3.1. Modelos AR(p), MA(q) y ARMA(p,q) generalizados a un modelo ARIMA(p,d,q)	21
3.2. Comportamiento de la FAC y FACP para procesos estacionarios	23
3.3. Condiciones de Estacionaridad e Invertibilidad para algunos Modelos	24
3.4. Comportamiento de la FAC y FACP para procesos estacionarios	29
4.1. Valores estimados de FAC y FACP para \tilde{Z}_t	44
4.2. Valores estimados de FAC y FACP para W_t	47
4.3. Valores estimados de FAC y FACP para W_t con $s = 12$	48
4.4. Estimadores de Máxima Verosimilitud de los modelos identificados	51
4.5. Valores de AIC de los modelos identificados	52
4.6. Valores de $\hat{\beta}_i$, $\hat{\sum}_{ii}$ y t_c para el modelo $SARIMA(0, 1, 1)(2, 1, 0)_{12}$	54
4.7. Valores de $\hat{\beta}_i$, $\hat{\sum}_{ii}$ y t_c para el modelo $SARIMA(0, 1, 2)(2, 1, 0)_{12}$	54
4.8. Valores de $\hat{\beta}_i$, $\hat{\sum}_{ii}$ y t_c para el modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$	55
4.9. Valores de $\hat{\rho}_k$, para residuales del modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$	57
5.1. Estadísticas Descriptivas de la serie de aportaciones Z_t	61
5.2. Coeficientes estimados del modelo $SARIMA(0, 1, 1)(2, 1, 0)_{12}$	64

Índice de Cuadros

5.3. Coeficientes estimados del modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$	64
5.4. Pronósticos de aportaciones a la presa Adolfo Ruiz Cortines.	67
5.5. Observados contra ajustados para el ciclo 2010-2011.	68
5.6. Comparación gráfica de la ST vs Análisis de frecuencia.	70
.1. Registro Historico de Aportaciones Deducidas a la Presa Adolfo Ruiz Cortines	85
.2. Continuación: Registro Historico de Aportaciones Deducidas a la Presa Adolfo Ruiz Cortines	86

Índice de figuras

3.1. Ciclo del Escurrimiento.	8
3.2. Serie de tiempo de aportaciones deducidas totales anuales a la presa Adolfo Ruiz Cortines en Hm^3	14
3.3. Procedimiento Iterativo de Modelos Box-Jenkins.	22
4.1. Componentes principales del DR 038.	32
4.2. Río Mayo, aguas abajo de la presa ARC.	34
4.3. Cortina de la presa ARC y Vaso de Almacenamiento de la presa Adolfo Ruiz Cortines.	34
4.4. Obra de toma de la presa Adolfo Ruiz Cortines.	35
4.5. Obra de excedencia de la presa Adolfo Ruiz Cortines.	36
4.6. Estación hidrométrica Túneles de Mocúzari.	37
4.7. Curva Áreas-Elevaciones del vaso de almacenamiento de la presa ARC.	38
4.8. Curva Capacidades-Elevaciones del vaso de almacenamiento de la presa ARC.	38
4.9. Hidrograma de Aportaciones Deducidas Totales Mensuales Z_t a la presa ARC.	41
4.10. Serie de aportaciones trasformada con desviaciones respecto a la media muestral, \tilde{Z}_t	42
4.11. Aportaciones mensuales $\tilde{Z}_t(2006-2011)$ $t = 612, \dots, 672$ a la presa ARC.	43

Índice de figuras

4.12. FAC (a) y FACP (b) estimadas para la serie \tilde{Z}_t	44
4.13. Serie $W_t = (1 - B)\tilde{Z}_t$	45
4.14. FAC (a) y FACP (b) estimadas para la serie $W_t = (1 - B)\tilde{Z}_t$	46
4.15. Serie $W_t = (1 - B^{12})(1 - B)\tilde{Z}_t$	47
4.16. FAC (a) y FACP (b) estimadas para la serie $W_t = (1 - B^{12})(1 - B)\tilde{Z}_t$	48
4.17. Función de AutoCorrelación de Residuales.	56
5.1. Hidrograma de Aportaciones Mensuales a la presa ARC.	60
5.2. Función de AutoCorrelación de Residuales.	65
5.3. Pronósticos para dos ciclos agrícolas.	67
5.4. Valores observados y ajustados para el ciclo 2011-2012.	69
5.5. Serie ajustada del modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$	69
5.6. comparación gráfica entre medias, Log Normal y Valores de pronóstico.	71

Capítulo 1

INTRODUCCIÓN

“ALL MODELS ARE WRONG BUT SOME ARE USEFUL”

George E. P. Box

Uno de los sectores que tiene los consumos de agua más altos a nivel mundial, *es la agricultura*, ya que ésta por sí misma requiere grandes cantidades de agua para el desarrollo sustancial de los cultivos. El riego es indispensable para satisfacer las necesidades hídricas de los cultivos en la mayor parte de nuestro país.

Dado que la precipitación es un fenómeno aleatorio, fuente primigenia de agua para la agricultura, y en el caso de la agricultura de riego, de los escurrimientos que constituyen las aportaciones hídricas a las presas; es necesario realizar una estimación confiable de dichas aportaciones a las presas de almacenamiento, con fines de planeación y operación de sistemas de irrigación, como son los distritos de riego, módulos de riego y unidades de riego.

Se entiende por aportaciones a vasos de almacenamiento que se forman al construir una presa, a los volúmenes de escurrimientos de una o varias corrientes superficiales o subterráneas que se presentan en un periodo y son almacenadas para su uso en diferentes actividades. Los escurrimientos se registran en los denominados puntos de control mediante la medida del flujo en estaciones hidrométricas en los mismos vasos. La importancia del estudio de las aportaciones deducidas radica principalmente en la estimación del volumen de agua que se espera entre al vaso, con la finalidad de realizar la planeación de su utilización en los distritos de riego y otros usos.

Para el estudio de las aportaciones, que se puede esperar en presas, se analizan con

1. INTRODUCCIÓN

modelos matemáticos basados en la estadística, con los que se trata de inferir el comportamiento de las variables hidrológicas, que en este caso, son las aportaciones. Dichos modelos pueden ser mediante análisis de frecuencia o modelos de series de tiempo. Este último tipo de análisis para aportaciones es el que se presentan en este trabajo.

Una serie de tiempo es un conjunto de observaciones de variables que se registran en forma sucesiva en el tiempo. Con frecuencia, estas observaciones se realizan en intervalos de tiempo regulares; por ejemplo, un año, un mes, una semana, un día y en algunos casos en periodos tan cortos como un segundo. Las series de tiempo aparecen en casi todas las disciplinas del conocimiento, tales como las ciencias naturales, las ciencias sociales, economía, ingeniería, medicina, meteorología, etc. y su análisis se puede realizar a partir de diferentes enfoques; por ejemplo, Análisis Espectral de series de tiempo, con modelos Autorregresivos y de Promedios Móviles, y el análisis armónico, entre otros.

Uno de los problemas que intenta resolver el análisis de series de tiempo es el de predecir. Esto es, dada una serie de datos $\{X_t\}$ en función del tiempo t , el objetivo es describir su comportamiento buscando posibles patrones temporales que permitan encontrar procesos de ajuste a los datos observados y así utilizar para predecir un comportamiento a futuro, con cierto grado de probabilidad.

El objetivo fundamental de este trabajo es presentar una metodología de análisis de series de tiempo aplicado a la estimación de aportaciones deducidas con fines de planeación de operación de Distritos de Riego, cuyo dominio es el tiempo. Se siguió la metodología clásica de análisis de series de tiempo propuesta por [Box y Jenkins \(1970\)](#). La metodología se aplicó, analizó y validó utilizando el registro histórico de aportaciones deducidas mensuales de la presa Adolfo Ruiz Cortines, que es la fuente de agua que abastece al distrito de Riego 038 Río Mayo, en el estado de Sonora, México.

Capítulo 2

HIPÓTESIS y OBJETIVOS

2.1. HIPÓTESIS

- El comportamiento y estimación de las aportaciones deducidas a presas de almacenamiento, como fenómenos hidrológicos que son, se pueden analizar con modelos de series de tiempo.

2.2. OBJETIVOS

- Desarrollar una metodología, para aplicar el análisis estadístico de series de tiempo, para estimar valores esperados de aportaciones deducidas esperadas a presas de almacenamiento.
- Aplicar y validar la metodología desarrollada para estimar las aportaciones de la presa Adolfo Ruiz Cortines, del Distrito de Riego 038 Río Mayo, Sonora, México.

Capítulo 3

REVISIÓN DE LITERATURA

Esta revisión tiene como finalidad el estudio del escurrimiento superficial, los fundamentos más comunes del análisis de series de tiempo bajo el enfoque metodológico de Box y Jenkins, así como la aplicación de éstas dentro del campo de la hidrología.

3.1. ANTECEDENTES

En toda ciencia, la disposición de datos básicos es esencial para una interpretación mejor del fenómeno a tratar. Una parte importante de la hidrología es precisamente la recolección y análisis de datos. [Chow \(1964\)](#), menciona que los procesos naturales que intervienen en los fenómenos hidrológicos son complejos por lo que resulta difícil examinarlos mediante un razonamiento deductivo y no siempre es posible la aplicación de una ley física para determinar un resultado hidrológico esperado. Por lo que se ha optado partir de una serie de datos observados y analizarlos estadísticamente para tratar de establecer la norma que gobierna dichos sucesos.

En el proceso de planeación de la operación de los sistemas de riego en operación se toman como base las observaciones históricas, del funcionamiento de los vasos de almacenamiento de las presas, las que se registran cuantitativamente en el tiempo. Tal es el caso de las aportaciones a presas de los Distritos de Riego del país. El análisis de estas series de datos históricos se utiliza para estimar los volúmenes esperados y con ello realizar los planes de riego para la operación del sistema en el plazo mediano, generalmente mensual y por ciclo agrícola.

De acuerdo a la literatura revisada, el problema de predicción de aportaciones hídricas ha sido estudiada bajo enfoques distintos que pueden resumirse en los siguientes grupos:

3.1. ANTECEDENTES

(a). **Ajuste de una Función de Distribución de Probabilidades.**

Desde este punto de vista se han realizado un gran número de trabajos, sobre todo para calcular probabilidades del monto de lluvia en un periodo dado. En el caso de las aportaciones [Guillen \(1990\)](#) propuso una metodología para realizar ajuste de funciones de distribución de probabilidad a aportaciones por escurrimiento. [Ariza \(2008\)](#) realizó el ajuste de cuatro funciones de distribución de probabilidad a las aportaciones de veinte presas de almacenamiento de Distritos de Riego del norte de México.

(b). **Procesos Estocásticos**

Este grupo puede dividirse a su vez en:

- *Cadenas de Markov*, En la teoría de la probabilidad, se conoce como cadena de Markov a un tipo especial de proceso estocástico discreto en el que la probabilidad de que ocurra un evento depende del evento inmediatamente anterior.
- *Análisis de Series de Tiempo*, cuyo dominio es el tiempo. Este trabajo esta enfocado desde este punto de vista y la metodología esta basada fundamentalmente por [Box y Jenkins \(1970\)](#). En la actualidad existen un gran número de referencias sobre las propiedades y aplicaciones de los modelos autoregresivos y de promedios móviles. [Quijano \(1977\)](#) realizó un modelo de predicción de lluvia bajo este enfoque.
- *Análisis Espectral*. Esta metodología se puede considerar como otra alternativa al análisis clásico de series de tiempo, en este caso se analiza la serie de tiempo utilizando como dominio la frecuencia, [Brockwell y Davis \(2002\)](#). Ambas metodologías se pueden considerar complementarias ya que algunas técnicas del análisis espectral se utilizan para determinar anomalías en el ajuste propuestos en el análisis clásico de series de tiempo.
- *Modelos de Espacio Estado*. Un enfoque mas reciente para el análisis de series de tiempo, son los modelos de espacio de estados, [Ariza \(2010\)](#). El análisis de estos modelos se basa principalmente, en la representación de los componentes de la serie (tendencias, estacionaridad y ruido) en dos ecuaciones, una de ellas dada por las observaciones y la otra por el proceso que forma.

En este trabajo son de interés el uso de modelos estocásticos, en particular los modelos de análisis de series de tiempo cuyo dominio es el tiempo.

3.1.1. **Uso de modelos estocásticos en análisis hidrológico**

El desarrollo de los modelos estocásticos ha llamado la atención de los hidrólogos por mucho tiempo, en especial los modelos para la representación de fenómenos hidrológicos

3.2. ESCURRIMIENTO SUPERFICIAL

como la precipitación, evaporación, escurrimiento, niveles freáticos, entre otros. En general un modelo estocástico es un modelo matemático cuya estructura esta regida por leyes de probabilidad, cuya inquietud básica ha sido la de predicción con fines de planeación [Cervantes \(1982\)](#).

[Chavasse y Santiago \(1997\)](#) propusieron las posibilidades de simulación conjunta de un modelo determinístico complejo y uno estocástico para representar caudales, en la cual se analiza la serie temporal de errores del modelo determinístico, mediante un modelo estocástico predecible. [Leon \(2000\)](#) propuso modelos ARIMA para explicar la estructura de la temperatura mensual.

El marco teórico para el estudio de las series de observada de aportaciones deducidas en este trabajo es la teoría de procesos estocásticos, en particular los procesos ARIMA (*Auto-Regressive Integrated Moving Average*), para el análisis univariado de series de tiempo estacionales.

Los modelos Box-Jenkins representan un método razonable para describir las series periódicas con comportamiento estocástico. Estos modelos son muy flexibles y capaces de modelar la estacionalidad y las tendencias de una manera más precisa que las técnicas convencionales de análisis [Pekárová et al \(2009\)](#).

Métodos diferentes como la Red Neuronal Artificial, Análisis de Regresión son usados para modelar datos hidroclimatológicos y a su vez se comparan con los modelos de series de tiempo ARIMA [Mohammadi1 et al \(2005\)](#).

[Pankratz \(1983\)](#) señala tres ventajas que tienen los modelos ARIMA propuestos por Box y Jenkins, sobre otros métodos tradicionales de análisis de series de tiempo. Primero, los conceptos y métodos asociados al ajuste de modelos ARIMA están respaldados por teorías sólidas de probabilidad dentro de la estadística matemática. Segundo, los modelos ARIMA son una familia de modelos, y tercero, la predicción que produce un modelo ARIMA tiene la propiedad de ser óptima, en el sentido que ningún otro modelo nos da una predicción con menor cuadrado medio de error de predicción.

3.2. ESCURRIMIENTO SUPERFICIAL

El termino escurrimiento es sinónimo de corriente. Es un proceso hidrológico, que se define como parte del agua proveniente de la precipitación que circula sobre o bajo la superficie terrestre y que llega a una corriente para ser drenada hasta la salida de una cuenca, este movimiento del agua está dado por la influencia de la gravedad. Una vez que la precipitación alcanza la superficie del suelo, el agua empieza a infiltrarse hasta que las capas del mismo se saturan, llenando algunas depresiones del terreno y al mismo tiempo el agua empieza a escurrir; debido a la diversidad de las fuentes de agua el escurrimiento

3.2. ESCURRIMIENTO SUPERFICIAL

puede ser superficial o directo; subsuperficial y subterráneo.

3.2.1. Componentes del Esgurrimiento

- (a). **El escurrimiento superficial.** Es aquel que proviene de la precipitación que no se infiltra en ningún momento, no queda retenido en depresiones naturales de los terrenos, escapa al fenómeno de la evapotranspiración y llega a la red de drenaje moviéndose sobre la superficie del terreno hasta la salida de la cuenca o almacenándose en un vaso de almacenamiento de una presa. El efecto sobre el escurrimiento total es directo y solo ocurrirá durante un evento de lluvia e inmediatamente después de que este cese. También suele denominarse flujo sobre la superficie.
- (b). **El escurrimiento subsuperficial.** Es la precipitación que habiéndose infiltrado en el suelo, se mueve lateralmente por los horizontes superiores del suelo para reaparecer e incorporarse a pequeños canalillos que lo conducirán a la red de drenaje.
- (c). **El escurrimiento subterráneo.** Es la precipitación que se infiltra hasta alcanzar la capa freática, circulando a través de los acuíferos y puede alcanzar la red de drenaje.

El componente final integrado por la suma de los tres escurrimientos anteriores que alcanza la red de drenaje de una cuenca se le llama *Esgurrimiento total o flujo superficial*.

3.2.2. Ciclo del Esgurrimiento

Los componentes del escurrimiento evolucionan según un ciclo, el cual se distingue de cuatro fases con relación al ritmo de las precipitaciones [Chávez \(2006\)](#), en la Figura 3.1 se muestran las fases del ciclo del escurrimiento.

- (a). **Primera fase.** Periodo sin precipitaciones, en el cual la evapotranspiración agota la humedad existente en las capas superficiales y extrae el agua de las partes subterráneas a través de la zona capilar. Las aguas subterráneas alimentan las corrientes superficiales descendiendo progresivamente su nivel piezométrico.
- (b). **Segunda fase.** Iniciación de las precipitaciones. En esta fase la evapotranspiración disminuye. La precipitación es interceptada por la vegetación, por las superficies de agua libre, depresiones naturales. Se infiltra una cantidad de agua importante que abastece la capacidad de abastecimiento del suelo; el excedente se mueve superficialmente en forma de escurrimiento y alimenta débilmente a los cursos de agua.

3.2. ESCURRIMIENTO SUPERFICIAL

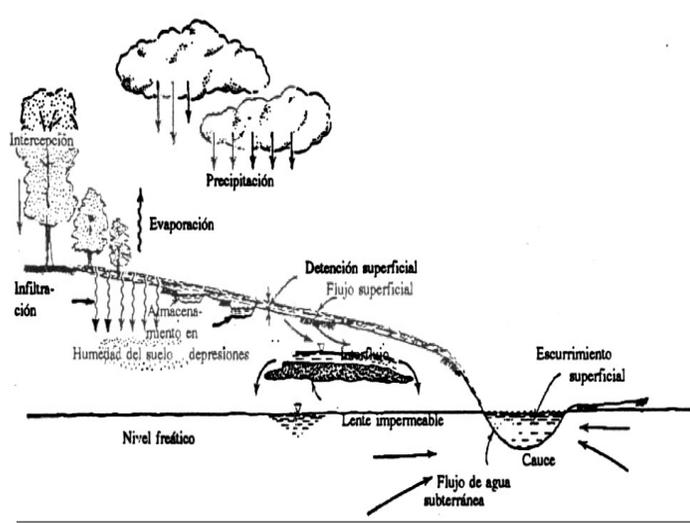


Figura 3.1. Ciclo del Escurrimiento.

- (c). **Tercera fase.** Precipitación máxima. Después del inicio de la precipitación y de una cierta duración de la misma, la cubierta vegetal apenas intercepta la lluvia, por lo que la totalidad del agua alcanza el suelo, las capas superficiales están casi saturadas. Parte del agua se infiltra alimentando el flujo subsuperficial y los acuíferos aumentando su nivel piezométrico. La precipitación no infiltrada ocasiona escurrimiento superficial que en esta fase alcanza el valor máximo, en esta fase el escurrimiento total alcanza su máximo valor, apareciendo las crecidas de las avenidas.
- (d). **Cuarta fase.** Posterior a la precipitación. En esta fase la lluvia cesa. El escurrimiento superficial desaparece rápidamente. El suelo y el subsuelo están casi saturados. Continúa la infiltración del agua estancada en depresiones superficiales alimentando a la humedad del suelo, al escurrimiento subsuperficial y subterráneo. Aparece nuevamente la evapotranspiración. El ciclo se cierra con la aparición de la nueva fase.

3.2.3. Aportaciones por Escurrimiento

Las aportaciones por escurrimiento superficiales son los volúmenes de una o varias corrientes superficiales que se presentan en un periodo y son almacenadas para su uso en diferentes actividades en vasos de almacenamiento, dichos escurrimientos se registran para su estudio en los denominados puntos de control, [Ariza \(2008\)](#).

El conocimiento de los escurrimientos en ríos para su almacenamiento o libre tránsito, es el elemento básico para planeación, diseño, construcción y operación de las obras

3.2. ESCURRIMIENTO SUPERFICIAL

hidráulicas. De igual forma el desconocimiento de la magnitud y época de ocurrencia de las aportaciones dan lugar a una planeación inadecuada de la operación de dichas obras y por ende de los sistemas de riego donde su principal fuente de abastecimiento es un almacenamiento.

La importancia de estudiar las aportaciones en la estimación de la cantidad de agua que escurre en una corriente para almacenarla y utilizarla adecuadamente a través del tiempo en vasos de almacenamiento de las presas. Para su estudio se han desarrollado modelos o funciones que representan variables aleatorias que a su vez simbolizan las aportaciones por lo que podemos conocer con un nivel de incertidumbre la presencia en tiempo y magnitud de ocurrencia de dichas aportaciones.

3.2.4. Cálculo de Aportaciones

Cuantificar de una manera precisa las aportaciones totales que llegan a un vaso de almacenamiento, debido a la complejidad de su naturaleza misma, resulta técnicamente imposible. Sin embargo se han implementado estrategias para obtener estimaciones aceptables de las aportaciones, además de las medidas en las estaciones hidrométricas en los ríos, que permiten realizar la planeación de operación según sea el uso de la presa.

Estas estrategias se basan en un balance hidrológico de la cuenca donde se encuentra ubicado el vaso de almacenamiento, esto permite conocer la disponibilidad de agua, por lo que es necesario contar con información básica: climatológica e hidrométrica. La ecuación de balance hídrico se basa en el principio de conservación de Lavoisier que establece *Nada se crea ni se destruye*, lo que se expresa que la variación de volumen en el almacenamiento es equivalente a la diferencia de las entradas de agua al sistema y las salidas de agua del sistema.

$$I = O \pm \Delta S \quad (3.1)$$

Donde:

I es la Entrada e agua a un sistema durante un periodo determinado en un área dada.
 O es la Salida de agua a un sistema durante un periodo determinado en un área dada.
 ΔS es el cambio de almacenamiento en el sistema.

Los parámetros de entrada y salida a considerar pueden llegar a resultar muy complejos para su cálculo, esto dependerá de la disponibilidad económica, espacial y temporal con que se cuente para realizar dicho balance.

3.2. ESCURRIMIENTO SUPERFICIAL

3.2.5. Cálculo de aportaciones deducidas en los Distritos de Riego

En los distritos de riego donde su fuente de abastecimiento son vasos de almacenamiento, se requiere conocer las aportaciones para la planificación del riego, estas pueden ser calculadas mediante la realización de un balance diario sobre el vaso, ya que cada presa cuenta con una gráfica de Elevaciones-Áreas-Capacidades, lo que permite conocer el almacenamiento diario del vaso y por consiguiente una variación del almacenamiento entre dos observaciones.

Las tablas, ecuaciones o gráficas de Elevaciones-Áreas-Capacidades, relaciona estas tres variables y se obtiene a partir del levantamiento topográfico del vaso, es decir que, a una elevación del nivel del agua corresponde un área de embalse y una capacidad de almacenamiento en el vaso.

Las salidas del vaso son controladas mediante la obra de toma. Otros parámetros a cuantificar como la evaporación y la precipitación son valores observados mediante estaciones meteorológicas ubicadas en el área de influencia del vaso. El régimen diario de almacenamiento de los vasos de las presas de los distritos de riego se registra debidamente y concentran a nivel mensual.

Si aplicamos la ecuación del balance hídrico, podemos conocer las entradas de agua al vaso de almacenamiento, lo que se le conocen como aportaciones deducidas. A este procedimiento realizado en los distritos de riego se le llama cálculo hidrométrico.

La información necesaria para realizar el cálculo hidrométrico diario es: la fecha, la meteorológica, información del vaso, entradas y salidas del vaso, a continuación se describen cada una de ellas.

FECHA. De este tipo de información se requiere saber el día i , el mes j y el año k .

$$i = 1, 2, \dots, NDM_j$$

$$j = 1, 2, \dots, 12$$

$$k = 1, 2, \dots, NAR$$

Donde:

NAM_j son el Número de Días del Mes j (28,29,30 ó 31).

NAR es el Número de Años de Registro Histórico.

3.2. ESCURRIMIENTO SUPERFICIAL

Información Meteorológica. Las principales son la precipitación Pr_{ijk} y la evaporación Ev_{ijk} , esta ultima modificada por un coeficiente de evaporación menor a la unidad.

Dichas variables generan un volumen de entrada o de salida de agua respectivamente según el área del vaso que registre el mismo día del cálculo.

Información del Vaso. En una escala ubicada en la cortina de la presa diariamente se lee un nivel de la elevación del agua L_{ijk} , con esa lectura de escala, se entra a las tablas o ecuaciones de las curvas de Elevación-Áreas-Capacidades elaboradas para el vaso de almacenamiento en cuestión, se obtiene el Área de embalse AE_{ijk} y el volumen almacenado en la presa VE_{ijk} .

$$\Delta S_{ijk} = VE_{(i+1)jk} - VE_{ijk} \quad (3.2)$$

Donde:

ΔS_{ijk} = Variación del volumen del día i , del mes j del año k .

VE_{ijk} = Volumen de embalse del día i , del mes j del año k .

$VE_{(i+1)jk}$ = Volumen de embalse del día $i + 1$, del mes j del año k .

Entradas de agua al vaso (I_{ijk}). Principalmente se tiene el volumen precipitado en la presa VL_{ijk} y el volumen de corrientes aforadas como ríos $Q1_{ijk}$ y manantiales inferiores $Q2_{ijk}$. lo que generan una entrada parcial agua que sumada al las aportaciones deducidas se tendrá la cantidad total de agua que entra al vaso.

$$I_{ijk} = VL_{ijk} + Q1_{ijk} + Q2_{ijk} \quad (3.3)$$

Salidas de agua en el vaso O_{ijk} . Se consideran el agua extraída por la toma ST_{ijk} , el desfogue D_{ijk} , la filtración en el vaso F_{ijk} , los derrames De_{ijk} y la evaporación en el vaso Ev_{ijk} .

$$O_{ijk} = ST_{ijk} + D_{ijk} + F_{ijk} + De_{ijk} + Ev_{ijk} \quad (3.4)$$

3.3. MODELOS BOX-JENKINS

Aportación deducida

Si aplicamos la ecuación 3.1 y sustituyendo las ecuaciones 3.2, 3.3 y 3.4, nuestro cálculo hidrométrico sería de la siguiente forma:

$$APD_{ijk} = (O_{ijk} \pm \Delta S_{ijk}) - I_{ijk} \quad (3.5)$$

Donde:

APD_{ijk} = Aportación deducida del día $i + 1$, del mes j del año k .

O_{ijk} = Salidas del vaso dadas por 3.4.

I_{ijk} = Entradas principales al vaso, dadas por 3.3.

ΔS_{ijk} = Variación del almacenamiento 3.2.

Esta información se registra y se concentra a diario para realizar el cálculo de la aportación diaria. La aportación deducida mensual se obtiene sumando las aportaciones deducidas diarias, en el Anexo A se da un ejemplo de cálculo de aportación para la presa Adolfo Ruiz Cortines.

En los siguientes secciones se exponen aspectos generales de los modelos ARIMA y del análisis espectral singular de series de tiempo.

3.3. MODELOS BOX-JENKINS

G.E.P. Box, profesor de Estadística de la Universidad de Wisconsin, y G.M. Jenkins, profesor de Ingeniería de Sistemas de la Universidad de Lancaster, introdujeron un nuevo enfoque en el análisis de series temporales, en sus trabajos sobre el comportamiento de la contaminación en la bahía de San Francisco, con el propósito de establecer mejores mecanismos de pronóstico y control. El libro [Box y Jenkins \(1970\)](#) en el que describen la metodología, se convirtió rápidamente en un clásico, y sus procedimientos se utilizan ampliamente desde entonces en diferentes ramas de la ciencia, conociéndose como modelos ARIMA, o bien, modelos Box-Jenkins.

3.3. MODELOS BOX-JENKINS

3.3.1. Definiciones y Notación

La teoría de procesos estocásticos estudia a los sistemas cuyo comportamiento, en el tiempo (o en el espacio), está determinado por leyes de probabilidad.

Proceso Estocástico. Un proceso estocástico es una colección de variables aleatorias $\{X_t\}$, referidas a un índice T , el cual puede ser discreto o continuo con una distribución común F_x . Note que, la definición no menciona si las variables aleatorias son independientes o no. Un aspecto importante es que el proceso debe estar referido al mismo espacio de probabilidad (Ω, B, P) . Detalles más avanzados sobre definiciones y conceptos de probabilidad y estadística en series de tiempo se pueden consultar en [Brockwell y Davis \(2002\)](#).

Serie de tiempo. Una serie de tiempo z_1, z_2, \dots, z_n es una *realización* de las variables aleatorias Z_1, Z_2, \dots, Z_N de un proceso estocástico con función de densidad conjunta $f(Z_1, \dots, Z_N)$.

En el análisis de series de tiempo no es muy frecuente expresar el modelo en forma de densidad conjunta $f(Z_1, \dots, Z_N)$ sino que en lugar de esto se expresa el mecanismo aleatorio que genera la función de probabilidad conjunta $f(Z_1, \dots, Z_N)$. Por ejemplo el modelo $Z_t = Z_{t-1} + U_t$ para $t = 1, 2, \dots$

Una serie de tiempo es un conjunto de observaciones z_t , cada una registrada en un tiempo específico t . Las series de tiempo discretas son aquellas en las cuales el conjunto de tiempos T_0 , en donde las observaciones se obtienen, es discreto, como en el caso cuando las observaciones se realizan en intervalos de tiempo fijos. Las series de tiempo continuas son aquellas cuando las observaciones son registradas continuamente sobre un intervalo de tiempo, por ejemplo, $T_0 = [0; 1]$. En el presente trabajo se considera el caso discreto y las observaciones igualmente espaciadas.

Un ejemplo de series de tiempo se da en la Figura 3.2 que muestra la serie $\{z_t, t = 1, 2, \dots, 68\}$, la cual representa el registro histórico (1942-2009) de aportaciones deducidas totales anuales a la presa Adolfo Ruiz Cortines, ubicada en el DR 038 Río Mayo en el estado de Sonora, México.

Existen fundamentalmente dos clases de modelos para describir las series de tiempo, que son los modelos estacionarios y los no estacionarios. Los primeros suponen que el proceso esta en un estado de equilibrio estadístico, es decir su media y varianza no cambian a través del tiempo. Los segundos se caracterizan por tener media y/o varianza que dependen del tiempo.

3.3. MODELOS BOX-JENKINS

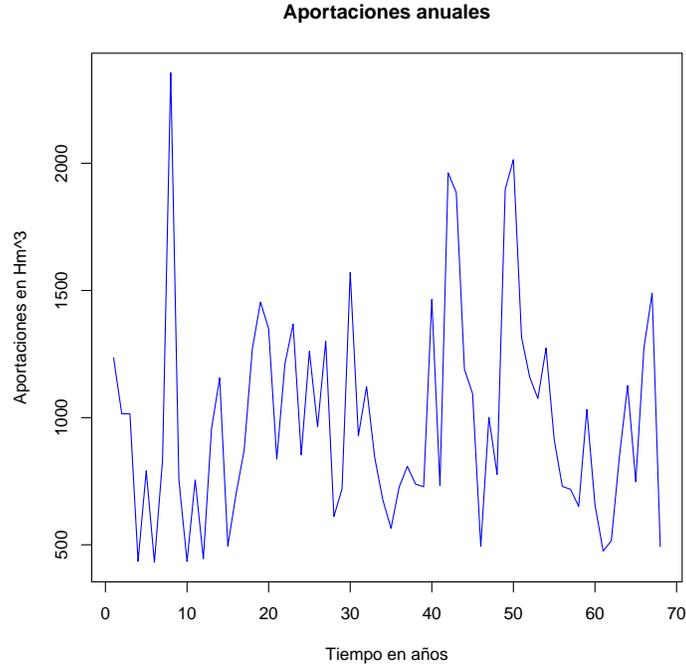


Figura 3.2. Serie de tiempo de aportaciones deducidas totales anuales a la presa Adolfo Ruiz Cortines en Hm^3

Uso de operadores de retraso B y ∇ . A continuación se presenta el uso de operadores de retraso B y ∇ incluyendo la relación entre ellos para poder mostrar la utilidad de las diferenciaciones en la transformación de series de tiempo.

El operador de retraso (*Backshift notation*) denotado por B , se define como:

$$BZ_t = Z_{t-1} \quad \forall t \quad (3.6)$$

se multiplica B^k por Z_t se obtiene la variable retrasada k períodos. Notese que el operador modifica toda la sucesión de valores $\{Z_1, Z_2, \dots, Z_t, \dots, Z_N\}$ y la transforma en la nueva sucesión $\{Z_{1-k}, Z_{2-k}, \dots, Z_{t-k}, \dots, Z_{N-k}\}$. La serie que originalmente constaba de N observaciones, se reduce a una serie de $N - k$ observaciones.

Otro operador de uso frecuente y que está íntimamente ligado con B es el operador diferencia ordinal ∇ . Y se define como:

$$\nabla Z_t = Z_t - Z_{t-1} = (1 - B)Z_t \quad \forall t \quad (3.7)$$

3.3. MODELOS BOX-JENKINS

Las potencias de B y ∇ son consistentes con el álgebra común:

$$B^k Z_t = Z_{t-k} \quad (3.8)$$

$$\nabla^k Z_t = (1 - B)^k Z_t \quad (3.9)$$

$$\nabla_k Z_t = (1 - B^k) Z_t = Z_t - Z_{t-k} \quad (3.10)$$

3.3.2. Procesos estocásticos lineales

La clase de procesos estacionarios más comunes para estudiar series de tiempo son los procesos lineales. Estos procesos son el punto de referencia para el estudio de los modelos más conocidos en las series de tiempo estacionarias. La idea se basa que una serie de tiempo, cuyos valores sucesivos pueden ser altamente *dependientes*, puede considerarse generada a partir de una serie de efectos aleatorios *independientes* $\{a_t\}$.

Estos efectos aleatorios se supone que son realizaciones independientes de una variable aleatoria cuya media es constante y varianza es σ^2 . A esta sucesión de variables aleatorias $\{a_t\}$. se le conoce como proceso de ruido blanco.

La idea expuesta previamente fue concebida por Yule (1927), quien propuso expresar al proceso $\{Z_t\}$ en función de $\{a_t\}$ mediante la relación lineal

$$Z_t = \mu + a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots = \mu + \psi(B) a_t \quad (3.11)$$

en donde μ es un parámetro que determina el nivel (no necesariamente la media) del proceso y $\psi(B) a_t$ el polinomio de retraso.

$$\psi(B) = 1 - \psi_1(B) - \psi_2(B^2) - \dots \quad (3.12)$$

que convierte al proceso $\{a_t\}$ en el proceso $\{Z_t\}$.

3.3. MODELOS BOX-JENKINS

3.3.3. Procesos estacionarios

El concepto de estacionariedad es de suma importancia para el análisis de series de tiempo. La propiedad de estacionaridad significa que el proceso esta en un estado de equilibrio estadístico *i.e.* su comportamiento estadístico no cambia a través del tiempo.

En general para caracterizar completamente un proceso estocástico es necesario conocer la función de densidad conjunta de sus variables aleatorias; sin embargo, en la práctica no es realista pensar que esto pueda lograrse con una serie de tiempo. No obstante, puesto que los primeros momentos de las variables aleatorias resumen en buena medida su distribución, conviene estudiar los momentos de primer y segundo orden, es decir, las medias, las varianzas y las covarianzas de las variables del proceso estocástico.

Función promedio. La función promedio de una serie de tiempo $\{Z_t\}$, para $t = 0, 1, 2, \dots$, denotada por μ , esta definida por:

$$E\{Z_t\} = \mu \quad (3.13)$$

La media del proceso $\{Z_t\}$, puede estimarse mediante la media muestral de la serie observada, denotada como:

$$\bar{Z} = \frac{1}{N} \sum_{t=1}^N Z_t \quad (3.14)$$

Notamos en (3.13) que la media del proceso no depende del tiempo, lo cual implica que aún cuando el proceso en un cierto periodo se aleje de la media, este siempre regresará a una vecindad de la misma.

Función de auto-covarianza. La varianza del proceso denotada como γ_0 , esta dada por la siguiente relación:

$$\gamma_0 = E\{(Z_t - \mu)^2\} = \sigma_Z^2 \quad (3.15)$$

Al considerar una serie de tiempo como un proceso estocástico, se debe tomar en cuenta todas las variables aleatorias Z_1, Z_2, \dots, Z_N , de ahí la importancia de estudiar la covarianza entre Z_t y Z_{t+k} . Dicha covarianza se denota como $\gamma_k = Cov(Z_t, Z_{t+k})$, es decir:

3.3. MODELOS BOX-JENKINS

$$\gamma_k = E\{(Z_t - \mu)(Z_{t+k} - \mu)\} \quad \forall t, k \quad (3.16)$$

Cuando $t = k$, la autocovarianza $\gamma_k = Cov(Z_t, Z_{t+k})$, es la varianza de la variable aleatoria en cuestión. La función de autocovarianza muestral C_k , se usa como estimador de γ_k dado en (3.16):

$$C_k = \frac{1}{N} \sum_{t=1}^{N-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z}) \quad k = 0, 1, 2, \dots \quad (3.17)$$

Note que cuando $k = 0$, entonces (3.17), es la varianza muestral dada en (3.15).

Estacionaridad estricta. Una serie de tiempo $\{Z_t\}$ con $t = 0, 1, 2, \dots$ se dice que es estrictamente estacionaria si para cualquier colección finita de variables aleatorias sobre el proceso se cumple que:

$$F_{z_1, z_2, \dots, z_k} = F_{z_{1+h}, z_{2+h}, \dots, z_{k+h}} \quad (3.18)$$

Es decir, la distribución conjunta de las variables aleatorias es la misma para $t = 0, 1, 2, \dots$ y las variables aleatorias desplazadas h unidades de tiempo $t = 0 + h, 1 + h, 2 + h, \dots$

Estacionaridad débil. Una serie de tiempo $\{Z_t\}$ con $t = 0, 1, 2, \dots$ se dice que es estacionaria en el sentido débil si para cualquier colección finita de variables aleatorias sobre el proceso aleatoria, cumple con las siguientes condiciones:

$$E\{Z_t\} = \mu \quad \forall t \quad (3.19)$$

$$Cov = (z_t, z_{t+k}) = E\{(z_t - \mu)(z_{t+k} - \mu)\} = \gamma_k \quad (3.20)$$

La estacionaridad débil asume que forzosamente debe satisfacer las dos condiciones expuestas en (3.19) y (3.20), es decir, la media y la covarianza no dependen de t sino de la diferencia (distancia) entre t_1 y $t_2 = t_1 + k$. La relación que existe entre estacionaridad débil y estricta es que para cualquier proceso que sea estrictamente estacionario también lo es en sentido débil. Lo contrario no es necesariamente cierto.

3.3. MODELOS BOX-JENKINS

Función de Autocorrelación (ACF). En la práctica es común suponer que la distribución asociada a series de tiempo es la normal, basta con conocer la media μ y la covarianza γ_k para caracterizar una serie estacionaria. Para evitar la influencia de las unidades de medida, es preferible trabajar con las autocorrelaciones, $\rho_0, \rho_1, \rho_2, \dots$ definidas por:

$$\rho_k = \frac{E\{(z_t - \mu)(z_{t+k} - \mu)\}}{E\{(Z_t - \mu)^2\}} = \frac{\gamma_k}{\gamma_0}, \quad k = \pm 0, \pm 1, \pm 2, \dots \quad (3.21)$$

que generan la *función de autocorrelación*. Para estimar a (3.21), se utiliza la Función de Autocorrelación Muestral r_k , dado por:

$$r_k = \frac{\sum_{t=1}^{N-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^{N-k} (Z_t - \bar{Z})^2} = \frac{C_k}{C_0}, \quad k = 0, 1, 2, \dots \quad (3.22)$$

La ecuación (3.22), mide la relación lineal entre las observaciones de la serie de tiempo separadas por un rezago (lag) de k unidades de tiempo. Note que, $-1 \leq r_k \leq 1$.

Para un valor cercano a 1, indica que las observaciones separadas por k unidades de tiempo tienen una fuerte tendencia a moverse juntas en forma lineal con pendiente positiva y con pendiente negativa para un valor cercano a -1.

3.3.4. Medición del error de pronóstico

El objetivo de un modelo matemático es representar alguna faceta de la realidad a través de una abstracción. Un modelo permite manipular artificialmente aspectos de la realidad, con el objeto de obtener respuestas. El modelo perfecto no existe, puesto que todos constituyen simplificaciones de la realidad y siempre son preferibles modelos con menos variables, puesto que además de ser más sencillos, son más estables y menos sometidos a sesgo.

Residuales. Los errores de pronóstico o residuales denotados por a_t , son la parte que no explica un modelo y se definen como:

$$a_t = y_t - \hat{y}_t \quad (3.23)$$

3.3. MODELOS BOX-JENKINS

Donde:

y_t es el valor observado en el período t .

\hat{y}_t es el valor estimado por el modelo en el período t .

Los errores en los pronósticos, sirven para comparar modelos y elegir aquel que mejor predice los valores. Algunas medidas de residuales están dadas a continuación.

Error Medio Absoluto (EMA), es definido como:

$$EMA = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (3.24)$$

Error Relativo Medio (ERM), es definido como:

$$ERM = \frac{100}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{y_t} \quad (3.25)$$

Error Cuadrado Medio (ECM), es definido como:

$$ECM = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad (3.26)$$

Criterio de Información de Akaike (AIC), Esta medida de alguna forma penaliza el número de parámetros del modelo, el AIC es usado como un criterio de selección en el análisis de series de tiempo cuando existen varios modelos [Kitagawa \(2010\)](#), se define como:

$$AIC = -2(\log \text{Verosimilitud} - p) \quad (3.27)$$

donde p es el número de parámetros del modelo. En principio el criterio de selección será escoger modelos con valores más bajos de AIC.

3.3.5. Modelos Box-Jenkins o Modelos ARIMA

Los modelos *autorregresivos e integrados de promedios móviles (ARIMA)*, pueden ser vistos como una generalización de los modelos *autorregresivos de promedios móviles (ARMA)*. Shumway y Stoffer (2006) menciona la posibilidad de que un cierto tipo de no estacionaridad mostrado por algunas series de tiempo, podría representarse mediante la simple toma sucesiva de diferencias de la serie original.

Esto permite gran flexibilidad de representación a los modelos ARMA, puesto que al aplicar el operador diferencia ∇^d dado por (3.7), es eliminar una posible tendencia polinomial de orden d , presente en la serie observada. Cuando la media de una serie de tiempo es estacionaria (constante a través del tiempo), entonces la media se puede considerar como un componente determinista de la serie, dado que nuestro interés es el comportamiento estocástico, es necesario expresar los datos $\{Z_t\}$ en desviaciones respecto a la media, definiéndose una nueva serie $\{\tilde{Z}_t\}$, es decir:

$$\tilde{Z}_t = Z_t - \bar{Z}_t \quad (3.28)$$

Donde \bar{Z}_t es el estimador de μ dado en (3.14), note que la media de esta nueva serie será exactamente cero. Por lo tanto es posible construir un proceso estacionario $\{W_t\}$, de tal manera que:

$$W_t = \nabla^d \tilde{Z}_t \quad \forall t \quad (3.29)$$

Para esta nueva serie $\{W_t\}$ ya podría ser posible representar un modelo ARMA:

$$\phi(B)Z_t = \theta(B)a_t \quad (3.30)$$

Donde B es un operador definido en (3.6). El modelo en (3.30) es equivalente a considerar el siguiente modelo ARIMA:

$$\phi(B)\nabla^d \tilde{Z}_t = \theta(B)a_t \quad d \geq 1 \quad (3.31)$$

El término *integrado* se refiere a que $\{\tilde{Z}_t\}$ se obtiene de la relación (3.29) aplicando la inversa de ∇^d . Así que un modelo ARIMA(p,d,q) indica que consta de un polinomio *autorregresivo* (AR) de orden p , con diferencia de orden d y de un polinomio de *promedios móviles* (MA) de orden q . De esta manera el modelo (3.31) se escribe como:

3.3. MODELOS BOX-JENKINS

$$(1 - \phi_1 B - \dots - \phi_p B^p)W_t = (1 - \theta_1 B - \dots - \theta_q B^q)a_t \quad (3.32)$$

O bien:

$$(W_t - \phi_1 W_{t-1} - \dots - \phi_p W_{t-p}) = (a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}) \quad (3.33)$$

Dado que el modelo (3.33) representa a una familia de modelos, algunos ejemplos de ello son los siguientes modelos mostrados en el Cuadro 3.1.

Cuadro 3.1. Modelos AR(p), MA(q) y ARMA(p,q) generalizados a un modelo ARIMA(p,d,q)

Nombre	Expresión
AR(1)	$(1 - \phi B)Z_t = a_t$
MA(1)	$Z_t = (1 - \theta B)a_t$
AR(2)	$(1 - \phi_1 B - \phi_2 B^2)Z_t = a_t$
MA(2)	$Z_t = (1 - \theta_1 B - \theta_2 B^2)a_t$
ARMA(1,1)	$(1 - \phi B)Z_t = (1 - \theta B)a_t$
ARIMA(1,d,0)	$(1 - \phi B)\nabla^d \tilde{Z}_t = a_t$
ARIMA(0,d,1)	$\nabla^d \tilde{Z}_t = (1 - \theta B)a_t$
ARIMA(1,d,1)	$(1 - \phi B)\nabla^d \tilde{Z}_t = (1 - \theta B)a_t$

En la amplia literatura referente al análisis de series de tiempo, es común presentar por separado cada proceso estacionario correspondiente al modelo general ARIMA, es decir, se detallan los modelos AR(p), MA(q), ARMA(p,q) por separado: Sin embargo la estrategia de ajuste de los modelos Box-Jenkins prosigue de la misma manera.

3.3.6. Procedimiento de elaboración de los modelos Box-Jenkins

Box y Jenkins (1970) propusieron un práctico procedimiento de cuatro etapas iterativas para encontrar un modelo adecuado a la serie observada. Dicho procedimiento es resumido esquemáticamente en la Figura 3.3 y se describe a continuación.

- (a) **Identificación** de un posible modelo dentro de la clase de modelos ARIMA, es decir, determinación del número de parámetros (p, d y q) que especifiquen el modelo ARIMA apropiado para la serie observada.
- (b) **Estimación** de parámetros involucrados en el modelo, a través de las técnicas de estimación no lineal.

3.3. MODELOS BOX-JENKINS

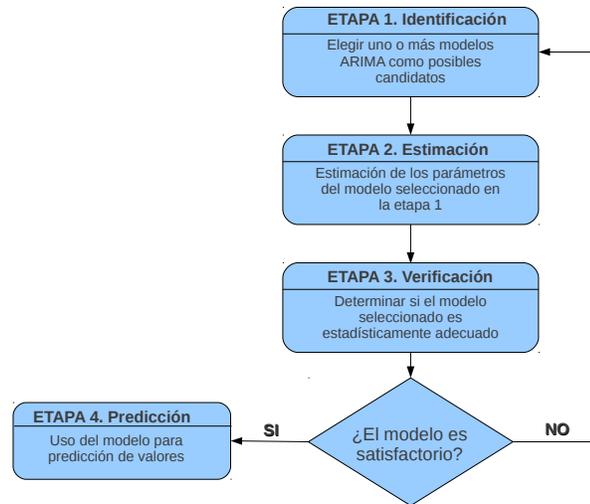


Figura 3.3. Procedimiento Iterativo de Modelos Box-Jenkins.

- (c) **Verificación** de que el modelo proporciona un ajuste adecuado y los supuestos implícitos en el modelo, se satisfacen. De no cumplirse los supuestos, se repiten las etapas anteriores.
- (d) **Uso del modelo** para pronóstico.

Un modelo ARIMA es una expresión algebraica que muestra como una variable de series de tiempo $\{Z_t\}$ se relaciona con sus propios valores pasados $\{Z_{t-k}\}$. En las siguientes secciones se presenta la forma algebraica de cada una de las etapas presentadas anteriormente.

3.3.7. Identificación del modelo

Esta etapa tiene como objetivo principal determinar los órdenes de los polinomios autorregresivos $AR(p)$ y de promedios móviles $MA(q)$, así como el número de veces que debería aplicarse el operador diferencia (∇^d) para eliminar la no estacionaridad homogénea de la serie observada (usualmente $d = 1$ o $d = 2$ son suficientes para lograr la estacionaridad). En vista de que en esta etapa no se tiene una formulación precisa de problema, se requiere del uso de métodos estadísticos aproximados.

En esta etapa los métodos gráficos y el sentido común son muy valiosos, no se debe olvidar que la identificación preliminar de modelos, solo sugiere modelos que más tarde serán validados estadísticamente.

3.3. MODELOS BOX-JENKINS

De manera más general podría decirse que la etapa de identificación consiste en determinar si la serie observada es estacionaria, de no ser así, es necesario hacer alguna transformación para lograr una serie estacionaria en función de la serie original, siendo esta nueva serie:

$$W_t = \nabla^d T(Z_t) \quad (3.34)$$

La principal herramienta para determinar si $\{Z_t\}$ es estacionaria es la Función de Autocorrelación Muestral dada en (3.22), un decaimiento rápido de las autocorrelaciones a cero es indicativo de que la serie es estacionaria.

Las principales herramientas para identificar el modelo adecuado son las gráficas de Z_t vs t , la función de autocorrelación muestral, FACM y la función de autocorrelación parcial muestral, FACPM. Se obtienen las FACM y FACPM de la serie observada y se compara con la Función de AutoCorrelación (FAC) y Función de Autocorrelación Parcial (FACP) teóricas de los modelos ARMA. Características de la FAC y FACP teóricas para series estacionarias, se resumen en el Cuadro 3.2.

Cuadro 3.2. Comportamiento de la FAC y FACP para procesos estacionarios

Proceso	FAC	FACP
AR(p)	Convergencia a cero.	Solo las primeras p autocorrelaciones parciales son diferentes de cero.
MA(q)	Sólo las primeras q autocorrelaciones son distintas de cero.	Sucesión infinita convergente a cero.
ARMA(p,q)	Comportamiento irregular de las primeras q autocorrelaciones y después converge a cero.	Sucesión infinita convergente a cero.

3.3.8. Estimación de parámetros

La etapa de estimación presupone que se ha identificado tentativamente un modelo ARIMA(p,d,q) y de ser éste adecuado para representar los datos. Ahora, paso siguiente es encontrar los mejores estimadores de los parámetros para que dicho modelo represente apropiadamente la serie en consideración.

Es decir, una vez conocidos los órdenes de los polinomios autorregresivos y de promedios móviles p y q , así como el grado de diferenciación d , se postula que el modelo

$$\phi(B)\nabla^d T(Z_t) = \theta_0 + \theta(B)a_t \quad (3.35)$$

resulta adecuado para representar la serie $\{T(Z_t)\}$. Entonces se requiere asignar valores a

3.3. MODELOS BOX-JENKINS

ϕ_1, \dots, ϕ_p y $\theta_1, \dots, \theta_q$. Métodos apropiados para la estimación de parámetros son el método de Mínimos Cuadrados (MC) y el método de Máxima Verosimilitud (MV), siendo este último el que emplearemos en este trabajo, para una presentación formal de los métodos consultar [Casella y Berger \(2002\)](#).

3.3.9. Verificación del modelo

Una vez que ya tenemos estimados en forma definitiva a los parámetros del modelo identificado, es conveniente verificar la adecuación del modelo a los datos. Pero antes de verificarse los parámetros del modelo identificado para la serie $\{W_t\}$ se debe cumplir con condiciones de estacionaridad e invertibilidad.

La condición de estacionaridad para un Modelo AR(p), es que las raíces de $\phi(B) = 0$ deben caer fuera del círculo unitario, o bien que el valor absoluto de cada una de las raíces es menor que uno. El proceso AR(p) siempre es invertible.

Una condición suficiente para un MA(q), será invertible si el valor absoluto de cada una de las raíces del polinomio $\theta(B) = 0$ son menor que uno. Un proceso MA(q) siempre es estacionario.

Un Modelo ARMA(p,q), es estacionario si el valor absoluto de cada una de las raíces de la ecuación característica $\phi(B) = 0$ caen fuera del círculo unitario. Similarmente, el proceso es invertible si los valores absolutos de las raíces de $\theta(B) = 0$ caen fuera del círculo unitario. En el Cuadro 3.3 se dan las condiciones que deben cumplir algunos modelos para que procesos estacionarios.

Cuadro 3.3. Condiciones de Estacionaridad e Invertibilidad para algunos Modelos

Modelo	Estacionaridad	Invertibilidad
AR(1)	$ \phi < 1$	Ninguna
AR(2)	$\phi_1 + \phi_2 < 1$ $\phi_2 - \phi_1 < 1$ $ \phi_2 < 1$	Ninguna
MA(1)	Ninguna	$ \theta < 1$
MA(2)	Ninguna	$\theta_1 + \theta_2 < 1$ $\theta_2 - \theta_1 < 1$ $ \theta_2 < 1$
ARMA(p,q)	$ \phi < 1$	$ \theta < 1$

Para realizar la verificación del modelo, es necesario verificar que los coeficientes de las variables incluidas en el modelo sean significativamente distintos de cero; aumentar términos al modelo y verificar si son significativos o no y finalmente comprobar que los residuales no estén correlacionados.

3.3. MODELOS BOX-JENKINS

Análisis de residuales. Una de las formas más claras y simples para detectar violaciones a los supuestos de los modelos es a través del análisis de residuales, ya que se considera residuo aquella parte de las observaciones que no explica el modelo. Los residuales para el modelo $\phi(B)W_t = \theta(B)a_t$ están dados mediante:

$$\hat{a}_t = W_t - \hat{W}_t \quad (3.36)$$

Los residuales miden la discrepancia entre los valores observados W_t y los valores estimados \hat{W}_t por el modelo. Cuando el tamaño de la muestra es grande, los errores aleatorios y los residuales son esencialmente iguales; por esta razón, al analizar los residuales observados $\{\hat{a}_t\}$ se analiza básicamente lo que debería ser una realización del proceso de ruido blanco $a_t \sim wn(0, \sigma_{w_t}^2)$.

Los supuestos acerca del proceso $\{a_t\}$ pueden verificarse y posiblemente corregirse mediante los siguientes supuestos.

SUPUESTO 1 $\{a_t\}$ **tiene media cero.**

Se establece el siguiente juego de hipótesis:

$$H_0 : \mu_{a_t} = 0 \quad \text{VS} \quad H_1 : \mu_{a_t} \neq 0$$

Verificación. Se Calcula la media aritmética y la desviación estándar muestral de los residuales.

$$m(\hat{a}) = \frac{1}{(N - d - p)} \sum_{t=t'}^N \hat{a}_t \quad (3.37)$$

siendo (3.37) la media de los residuales y para la desviación se tiene que:

$$\hat{\sigma}_a = \sqrt{\frac{\sum_{t=t'}^N [\hat{a}_t - m(\hat{a})]^2}{(N - d - p - q)}} \quad (3.38)$$

para $t' = d + p + q$.

Regla de decisión. Rechazar H_0 si:

3.4. ANÁLISIS DE SERIES DE TIEMPO ESTACIONALES.

$$\left| \frac{m(\hat{a})\sqrt{N-d-p}}{\hat{\sigma}_a} \right| < 2$$

SUPUESTO 2 $\{a_t\}$ tiene varianza constante.

Verificación. Se grafican residuales versus el tiempo para observar, si la varianza se comporta constante a través del tiempo. Esta verificación visual podría pensarse burda, pero la idea es que solamente las violaciones muy notorias de este supuesto son las que realmente llegan a causar problemas.

SUPUESTO 3. Las variables aleatorias $\{\hat{a}_t\}$ son mutuamente independientes.

Debido que independencia implica no autocorrelación, se requiere que $\rho_k(a) = 0$ para toda $k \neq 0$. Por ello se establecen el siguiente juego de hipótesis:

$$H_0 : \rho_k(a_t) = 0 \quad \text{VS} \quad H_1 : \rho_k(a_t) \neq 0$$

Verificación. Se calcula FAC muestral de los residuos $\{r_k(\hat{a})\}$ y bajo el supuesto de que su media es cero, está dado por:

$$r_k(\hat{a}) = \frac{\sum_{t=t'}^{N-k} [\hat{a}_t(\hat{a}_{t+k})]}{\sum_{t=t'}^{N-k} [\hat{a}_t^2]} \quad k = 0, 1, 2, \dots \quad (3.39)$$

para $t' = d + p + q$.

Regla de decisión. Rechazar H_0 si:

$$|\rho_k(\hat{a})| < \frac{2}{\sqrt{N-d-p}}$$

3.4. ANÁLISIS DE SERIES DE TIEMPO ESTACIONALES.

Aun cuando los modelos ARIMA mostrados en sección anterior permiten representar una gran cantidad de fenómenos reales debido a su gran flexibilidad, existen todavía series que no son posibles modelar como simples modelos ARIMA. Por este motivo conviene extender la aplicabilidad de los modelos ARIMA para cubrir la necesidad de análisis de

3.4. ANÁLISIS DE SERIES DE TIEMPO ESTACIONALES.

las series de tiempo estacionales las cuales aparecen con gran frecuencia en la práctica.

3.4.1. Modelos ARIMA estacionales

Una serie de tiempo que aparte de contener una tendencia (y/o ciclos) de larga duración, muestre también fluctuaciones que se repitan en base a períodos de tiempo, trimestral, semestral o anual se conocen como series de tiempo estacionales o modelos **SARIMA**. El procedimiento consta de los siguiente pasos:

- (a) **Identificación** de un posible modelo dentro de la clase de modelos SARIMA, es decir, determinación del número de parámetros (p , d y q) que definan la parte no estacional de la serie y los parámetros (P, D, Q) para la parte estacional de la serie, que especifiquen el modelo SARIMA apropiado para la serie observada.
- (b) **Estimación** de parámetros involucrados en el modelo, a través de las técnicas de estimación no lineal.
- (c) **Verificación** de que el modelo proporciona un ajuste adecuado y los supuestos implícitos en el modelo, se satisfacen. De no cumplirse los supuestos, se repiten las etapas anteriores.
- (d) **Uso del modelo** para pronóstico.

Un modelo SARIMA es una expresión algebraica que muestra como una variable de series de tiempo $\{Z_t\}$ se relaciona con sus propios valores pasados $\{Z_{t-k}\}$ y además tiene un comportamiento estacional.

Una característica fundamental de las series de tiempo estacionales con observaciones mensuales, es la similitud de comportamiento entre observaciones para el mismo año y consecutivos. De igual manera si la serie estacional es de observaciones trimestrales, existirán similitudes entre los mismos trimestres para los años consecutivos. Es importante notar que aun cuando la estacionalidad se considera un fenómeno repetitivo anual, esto no implica que no pueda existir un cierto modelo de comportamiento periódico con duración menor a un año.

Por lo anterior, es conveniente hablar genéricamente de un periodo estacional que comprenda E observaciones contiguas.

Modelo multiplicativo estacional Para tener en cuenta ambos tipos de efectos, estacionales y no estacionales, [Box y Jenkins \(1970\)](#) propusieron el modelo general:

3.4. ANÁLISIS DE SERIES DE TIEMPO ESTACIONALES.

$$\Phi(B^s)\nabla_s^D Z_t - \mu = \Theta(B^E)\alpha_t \quad (3.40)$$

donde las variables α_t no se suponen como un ruido blanco, sino son generadas por un proceso $ARIMA(p; d; q)$, o sea

$$\phi(B)\nabla^d \alpha_t = \theta(B^s)a_t \quad (3.41)$$

con a_t un proceso de ruido blanco. De estas dos últimas expresiones se obtiene el modelo multiplicativo estacional

$$\phi(B)\Phi(B^s)\nabla^d\nabla_s^D Z_t - \mu = \theta(B)\Theta(B^s)\alpha_t \quad (3.42)$$

denotado por $ARIMA(p, d, q)(P, D, Q)_s$.

Los parámetros para la parte no estacional de la serie de tiempo se identifican con las letras minúsculas p, d, q y los parámetros para la parte estacional de la serie se denotan con las letras mayúsculas P, D, Q .

3.4.2. Identificación de modelos estacionales

El proceso de construcción de modelos para series de tiempo estacionales, es similar para el caso de series no estacionales, ahora lo que se pretende es transformar una serie estacional a una serie estacionaria.

Estabilización de la varianza y Nivel. Si la variabilidad de una serie de tiempo incrementa cuando el tiempo avanza (lo cual implica que la serie de tiempo no es estacionaria), entonces como primer paso se hace uso del método para seleccionar una transformación que estabilice la varianza, dicho método sugiere elegir la potencia λ para la transformación de $T(Zt) = Z_t^\lambda$.

Una vez que se ha elegido la potencia λ , como segundo paso es decidir el grado de diferenciación regular dy estacional D requerido para volver estacionaria la serie $\{T(Zt)\}$.

Empleo de FAC y FACP. Una vez que se ha hecho la transformación apropiada para estabilizar la varianza y elegido los valores de d y D para estabilizar el nivel, esto significa que la nueva serie W_t ya es estacionaria definida como:

3.4. ANÁLISIS DE SERIES DE TIEMPO ESTACIONALES.

$$W_t = \nabla^d \nabla^D {}_E T(Z_t) \quad (3.43)$$

Entonces lo que procede es analizar la Función de Autocorrelación (FAC) y la Función de Autocorrelación Parcial (FACP) para W_t , y de manera visual se identifica a que modelo tentativo corresponde. El comportamiento de las FAC y FACP para modelos ARMA estacionales están dadas en el Cuadro 3.4.

Cuadro 3.4. Comportamiento de la FAC y FACP para procesos estacionarios

Proceso	FAC	FACP
AR(P)	Convergencia a cero en $s = 1, 2, \dots$	Solo las primeras P_s autocorrelaciones parciales son diferentes de cero.
MA(Q)	Sólo las primeras Q_s autocorrelaciones son distintas de cero.	Sucesión infinita convergente a cero en $s = 1, 2, \dots$
ARMA(P,Q)	Comportamiento irregular de las primeras Q_s autocorrelaciones y después converge a cero.	Sucesión infinita convergente a cero en $s = 1, 2, \dots$

3.4.3. Construcción de modelos estacionales

La construcción de modelos para series estacionales también se realiza de manera eficiente de acuerdo con la estrategia sugerida por Box y Jenkins para el caso de series no estacionales.

No obstante, importa notar que en la etapa de verificación existen algunos pequeños cambios originados por la aparición de parámetros estacionales, así por ejemplo la expresiones:

$$m(\hat{a}) = \frac{1}{(N - d - p - (D + P)_E)} \sum_{t=t'}^N \hat{a}_t \quad (3.44)$$

siendo (3.44) la media de los residuales y para la desviación se tiene que:

$$\hat{\sigma}_a = \sqrt{\frac{\sum_{t=t'}^N [\hat{a}_t - m(\hat{a})]^2}{(N - d - p - q - (D + P + Q)_E)}} \quad (3.45)$$

para $t' = d + p + (D + P)_E + 1$. Por lo que le estadístico Q' de Ljung y Box se convierte en

3.5. PRONÓSTICO CON MODELOS ARIMA

$$Q' = [N-d-p-(D+P)_E][N-d-p-(D+P)+2]X \frac{[\sum_{k+1}^K r_k^2(\hat{a})]}{(N-d-p-(D+P)_E-k)} \quad (3.46)$$

3.5. PRONÓSTICO CON MODELOS ARIMA

Después de haber considerado las etapas de Identificación, estimación y verificación de los modelos, La última etapa para la estrategia de construcción de modelos ARIMA es el uso del modelo para predecir valores futuros de una serie de tiempo observada, es decir con fines de pronóstico.

Supóngase que $\{W_t\}$ es una serie de tiempo estacionaria con media cero, obtenida a partir de la serie original $\{Z_t\}$ (con N observaciones) como $W_t = \nabla^d T(Z_t)$, para algún valor de d y para una cierta transformación T , con W_t medida como desviación respecto a su media (3.28). Supóngase además que W_t admite una representación $W_t = \psi(B)a_t$, para la cual existe un modelo ARMA equivalente, modelo que se desea utilizar en la obtención de pronóstico de la serie observada.

En particular, si a partir del origen t , se desea pronosticar la observación W_{t+k} , un pronóstico cualquiera de esta observación, que se obtenga como una combinación lineal de los valores de la serie $\{W_t\}$, y en consecuencia de los errores $\{a_t\}$, sería denotado por $\tilde{W}_t(h)$, mientras que su pronóstico óptimo es $\hat{W}_t(h)$.

El criterio que se usa para determinar la optimalidad del pronóstico será el del error cuadrático medio mínimo (3.26) y el criterio de información de Akaike (3.27). Para lo primero $\hat{W}_t(h)$ debe satisfacer:

$$E[W_t - \hat{W}_t(h)]^2 = \min E[W_t - \tilde{W}_t(h)]^2 \quad (3.47)$$

En la mayoría de los casos, los pronósticos serán requeridos para la serie observada $\{Z_t\}$ y no para la transformada $\{T(Z_t)\}$. Por ello se usa $T^{-1}[\tilde{T}(Z_t)(h)]$ como pronóstico de Z_{t+h} .

Capítulo 4

MATERIALES Y MÉTODOS

Con el fin de ilustrar la metodología de análisis estadístico desarrollada para la construcción de modelos de series de tiempo, cuyo dominio es el tiempo (**modelos Box-Jenkins**) y así poder realizar el pronóstico de las aportaciones deducidas esperadas a presas de almacenamiento, se presenta como *estudio de caso* el de las aportaciones deducidas mensuales totales del Río Mayo a la presa de almacenamiento Adolfo Ruiz Cortines, que es la fuente de agua que abastece al Distrito de Riego 038 Río Mayo, en el estado de Sonora, México.

4.1. MATERIALES

Para el desarrollo de este trabajo, se utilizó el registro histórico de aportaciones deducidas mensuales de la presa Adolfo Ruiz Cortines, los datos fueron proporcionados por la Subdirección de Infraestructura Hidroagrícola, Gerencia de Distritos de Riego, Subgerencia de operación **CONAGUA** (2011) y la oficina de Hidrometría del Distrito de Riego 038. El registro histórico de la presa Adolfo Ruiz Cortines, **ARC**, se presenta en el anexo B.

4.1.1. Distrito de Riego 038 Río Mayo

En esta sección se presenta de manera simplificada la descripción e información referente al DR 038 Río Mayo, la idea principal es presentar una noción del área geográfica y descripción física de la infraestructura hidroagrícola, de donde se obtuvieron los registros históricos de aportaciones utilizados en este trabajo.

El distrito tiene una superficie regable de 96,951 ha y se encuentra organizado por una

4.1. MATERIALES

Sociedad de Responsabilidad Limitada (SRL) que opera la red mayor de canales y en 16 módulos de riego. La presa Adolfo Ruiz Cortines es operada por la CONAGUA, a través de la jefatura del Distrito de Riego 038 Río Mayo, Sonora, México, como se aprecia en la Figura 4.1.

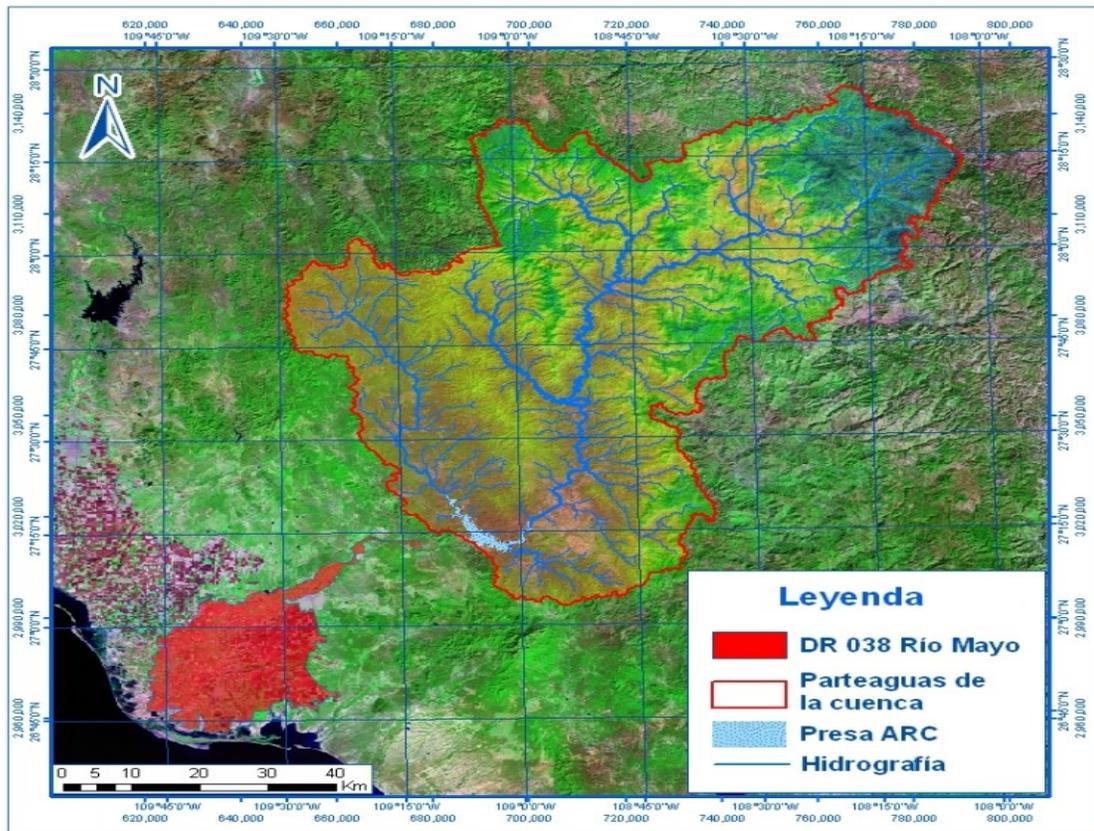


Figura 4.1. Componentes principales del DR 038.

Ubicación. El Distrito de Riego (DR) 038 Río Mayo, Sonora, se localiza al sur del Estado de Sonora entre los paralelos $26^{\circ}45'$ y $27^{\circ}15'$ y los meridianos $109^{\circ}30'$ y $110^{\circ}00'$, en los municipios de Navojoa, Etchojoa y Huatabampo. Limita al norte con el río Mayo, al este con la Sierra Madre Occidental, al sur con el estado de Sinaloa y al oeste con el Golfo de California.

Climatología. La clasificación climática de Thornthwaite para la región del Valle del Mayo es: E d B' a; lo que implica una provincia de humedad E. árida con vegetación de desierto, humedad deficiente en todas las estaciones, provincia de temperatura B' mesotérmica; subprovincia de temperatura a. La precipitación media anual (estación Etchojoa) es de 335.6 mm y una evaporación media anual de 2 118.1 mm.

4.1. MATERIALES

Las lluvias en la región se concentran en los meses de julio, agosto, septiembre y octubre, las precipitaciones son de poca importancia. Las temperaturas medias tienen una variación de $30.0^{\circ}C$ en agosto a $15.8^{\circ}C$ en enero.

Características agronómicas del DR 038. El ciclo agrícola empieza el primero de octubre y finaliza el 30 de septiembre. Dentro de este año agrícola se pueden diferenciar varios subciclos, siendo el más importante de ellos el subciclo de Invierno, donde se siembra aproximadamente el 85 % de la superficie del Distrito. El subciclo de Primavera sigue en importancia, sembrándose el 15 % restante de la superficie aproximadamente.

Las necesidades hídricas de los demás cultivos son variables. Algunos necesitan más de $6,000m^3/ha$ (alfalfa, maíz), otro menos (cártamo, garbanzo). Para solucionar este desequilibrio, los módulos de riego manejan alternativas de carácter interno.

Sumado a la escasez de agua los principales problemas en el distrito de riego son las dificultades en la operación y grandes pérdidas de agua por la falta de revestimiento en los canales y las bajas eficiencias de aplicación en el riego por gravedad ?.

Fuente de abastecimiento de agua. La principal corriente que alimenta a la presa Adolfo Ruiz Cortines es el Río Mayo, Figura 4.2, que nace en las faldas de la Sierra Madre Occidental, a elevaciones superiores a 2 500 m, en el estado de Chihuahua. Aguas abajo penetra el estado de Sonora y lo cruza siguiendo un rumbo de norte a sur. Recibe en su recorrido las aportaciones del arroyo Guajaray por su margen derecha y las de los arroyos Algodones, San Bernardo, Taymuco, Techobampo y Piedras Negras por su margen izquierda. Después de recibir las aportaciones del arroyo Quiriego, sale a la planicie costera, conservando un caudal de estiaje bastante importante. Toda la topografía de la cuenca, hasta el sitio de la presa es muy accidentada, factor que contribuye a que el río tenga un régimen torrencial.

4.1.2. Presa Adolfo Ruiz Cortines

El Área de la cuenca hasta el sitio de la presa es de $10,934.02km^2$. Se encuentra ubicada a 40 km, aproximadamente al noreste de la ciudad de Navojoa, sobre el cauce del río Mayo, ligeramente aguas abajo de la confluencia del arroyo Quiriego, en el municipio de Álamos, Estado de Sonora. Partiendo de la ciudad de Navojoa, con destino a la presa Adolfo Ruiz Cortines, se recorren 29 km por la carretera Navojoa - Álamos, para continuar hacia el norte un camino de terracería de 20 km de longitud que conduce al sitio de la presa.

Con objeto de aprovechar las aguas del río Mayo en riego y generación de energía eléctrica, se construyó la presa Adolfo Ruiz Cortines. La Figura 4.3 muestra una panorámica de la cortina y del vaso de almacenamiento de la presa.

4.1. MATERIALES



Figura 4.2. Río Mayo, aguas abajo de la presa ARC.



Figura 4.3. Cortina de la presa ARC y Vaso de Almacenamiento de la presa Adolfo Ruiz Cortines.

La presa Adolfo Ruiz Cortines se construyó en el periodo 1952 a 1955, registrándose los almacenamientos a partir del mes de julio de 1955. La presa fue diseñada con capacidad total de 1 376 millones de m^3 (azolves: 100; útil 914 y superalmacenamiento: 362),

4.1. MATERIALES

destinada a regar una superficie de 70 000 ha de terrenos de ambas márgenes, a generar energía eléctrica y otros propósitos secundarios. Está constituida por una cortina del tipo de materiales graduados, construida en la boquilla y 3 diques del mismo tipo situados, el primero en el lado derecho y en el izquierdo de la cortina, una obra de toma consistente en 2 túneles alojados en la margen izquierda de la cortina, utilizados para el desvío de la corriente durante la construcción y una obra de excedencias situada en el puerto del margen derecho, en el dique número 1.

Obra de toma. La obra de toma está alojada en la margen izquierda de la cortina, Figura 4.4 y consiste de 2 túneles denominados número 1 y número 2, de 6.00 m de diámetro y 342.00 m de longitud, 0.005 de pendiente, el primero y de 368.00 m e iguales características el segundo, revestidos de concreto simple, que se utilizaron para la obra de desvío y se destinan respectivamente para hacer las extracciones destinadas a riego y para alimentar a la planta hidroeléctrica.



Figura 4.4. Obra de toma de la presa Adolfo Ruiz Cortines.

En cada una de las entradas se apoya una estructura de rejillas que conectan con los túneles. En la zona central de los túneles se construyeron tapones de concreto en los que están ahogadas tuberías de acero. La del túnel número 1 tiene 4.00 m de diámetro y cuenta con una compuerta de rodillos de 4.00 por 4.00 m para emergencias, que se opera a través de una lumbrera y una válvula de chorro divergente de 2.50 m de diámetro para servicio, instalada en el extremo de la galería de válvulas. La del túnel número 2 tiene 3.70 m de diámetro y está provista de una válvula de emergencia del tipo de mariposa de 3.57 de diámetro, instalada en la galería de válvulas situada inmediatamente aguas

4.1. MATERIALES

abajo del tapón, que liga a una tubería de acero que cuenta a la salida del túnel con una bifurcación para abastecer, por una parte, a la planta hidroeléctrica y por otra para hacer descargas por medio de una válvula de servicio de chorro divergente.

Obra de excedencias. Está situada en el puerto del lado derecho, cerrado parcialmente por el dique número 1. Consiste en un vertedor de concreto del tipo de cresta libre con cimacio de perfil Creager y planta en curva, de 7.43 m de altura y 330.00 m de longitud, que se encuentra a la elevación 135.00. Al pie del cimacio está un plano a nivel que se prolonga en un canal de descarga de 250.00 m de longitud con ancho en la plantilla variable de 164.00 a 90.00 m, pendiente variable de 0.05 a 0.03 y taludes de 1:1, que remata en un dentellón profundo y termina en un canal sin revestimiento, la obra de excedencias puede verse en la Figura 4.5.



Figura 4.5. Obra de excedencia de la presa Adolfo Ruiz Cortines.

La obra de excedencias con capacidad de diseño para un gasto de $8\,000\text{ m}^3/\text{s}$, que pasaban sobre la cresta con una carga de 5.11 m dejando un bordo libre de 3.39 m.

Sobreelevación. La presa Adolfo Ruiz Cortines se sobre elevó con el propósito básico de controlar las avenidas del río Mayo para proteger contra inundaciones a las localidades asentadas a lo largo de la corriente. La sobreelevación de la presa consistió en aumentar 4.50 m la altura de la cortina y los diques, así como el piso de maniobras de la toma, así mismo la cresta del vertedor se sobre elevó 5.23 m, y en la parte central del cimacio se alojó una estructura de control de 51.90 m de longitud, provista de 5 compuertas radiales

4.1. MATERIALES

sobre una cresta vertedora situada a la elevación 130.82 m. el vertedor tiene capacidad de diseño para descargar un gasto máximo de $7\,450\text{ m}^3/\text{s}$.

4.1.3. Estaciones hidrométricas en DR 038 Río Mayo

Las aportaciones a la presa se registran en las estaciones hidrométricas San Bernardo y Tezocoma; la primera localizada sobre el río Mayo, donde se tienen datos desde el año de 1960 y la segunda sobre el cauce del arroyo Quiriego, que empezó a funcionar a partir del año de 1961. Las extracciones de la presa se miden en la estación hidrométrica Túneles de Mocúzari, ubicada aproximadamente a 150 m aguas Abajo de la planta hidroeléctrica Figura 4.6 , contándose con observaciones desde el año de 1964. Los aforos de los derrames se efectúan en la estación Tres Hermanos, localizada aproximadamente a 7 km aguas debajo de la cortina de la presa y donde se tienen datos desde el año de 1959, [IMTA \(2006\)](#).



Figura 4.6. Estación hidrométrica Túneles de Mocúzari.

Curva Áreas-Elevaciones-Capacidades. Otra forma de medir las aportaciones es en el vaso de almacenamiento a través de la Curva Áreas-Elevaciones-Capacidades, ya que relaciona los valores correspondientes de la elevación del nivel de agua, el área de embalse y el volumen almacenado cubicado. La Figura 4.7 y Figura 4.8 [Chávez \(2006\)](#), muestran las relaciones entre Área-Elevaciones y Volúmenes-Elevaciones, respectivamente.

4.1. MATERIALES

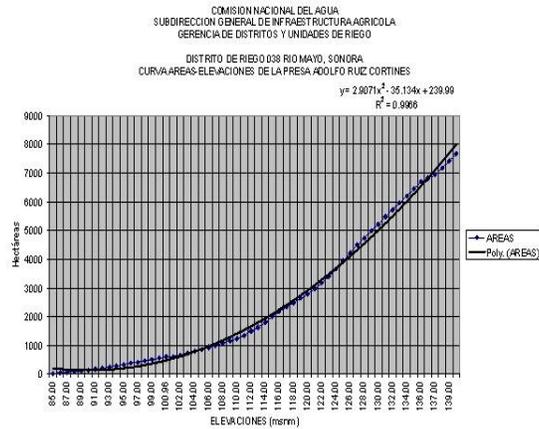


Figura 4.7. Curva Áreas-Elevaciones del vaso de almacenamiento de la presa ARC.



Figura 4.8. Curva Capacidades-Elevaciones del vaso de almacenamiento de la presa ARC.

4.1.4. Red de distribución de aguas

La red mayor del Distrito tiene una capacidad de distribución de $60m^3/s$. El canal Principal Margen Izquierda con $40m^3/s$ y el canal Principal Margen Derecha con $20m^3/s$. De acuerdo con el inventario de la Comisión Nacional del Agua, el DR 038 cuenta con 290.29 km de canales en red mayor, de los cuales sólo 58.85 km están revestidos y con 965.42 km de canales en red menor, de los cuales 175.57 km están revestidos. La red de drenaje está integrada por 809 km de drenes colectores y secundarios que son operados y conservados por la SRL del distrito.

4.1. MATERIALES

4.1.5. Información hidrométrica

El registro histórico de aportaciones deducidas a la presa ARC, utilizado en este trabajo, cuenta con un total 56 años de registro. Los datos mensuales de aportaciones se presentan en el Anexo B, están organizados por año agrícola (Octubre a Septiembre).

Las aportaciones deducidas se obtienen mediante el balance hídrico diario y mensual del vaso, Anexo A.

La serie de aportaciones mensuales empieza en el mes de octubre del año agrícola 1955-1956, y termina en el mes de septiembre del año agrícola 2010-2011, por lo que se tiene un registro de 672 meses observados. La información hidrométrica corresponde a la serie observada $\{Z_t\}$ para $t = 1, 2, \dots, 672$, que se utilizó para la construcción del modelo SARIMA que se utilizó para el pronóstico de aportaciones mensuales deducidas de los años agrícolas 2011-2012 y 2012-2013.

4.1.6. Software utilizado

Para el manejo de las aportaciones deducidas, $\{Z_t\}$ para $t = 1, 2, \dots, 672$, a la presa Adolfo Ruiz Cortines y procesamiento de los resultados del análisis de series de tiempo de modelos Box-Jenkins realizado a esta serie observada de aportaciones, se utilizaron los siguientes programas de cómputo:

LibreOffice 3.3.4. LibreOffice es una suite ofimática libre y gratuita, que funciona en muchos tipos de ordenadores y sistemas operativos. Dispone de un procesador de texto (Writer), un editor de hojas de cálculo (Calc), un creador de presentaciones (Impress), un gestor de bases de datos (Base), un editor de gráficos vectoriales (Draw) y un editor de fórmulas matemáticas (Math). Este programa se utilizó para el manejo de información hidrométrica, básicamente en el **LibreOffice Calc**. www.libreoffice.org.

Kile, es un Entorno Integrado (IDE) de **Tex/LaTeX**. \LaTeX es un sistema de composición de textos, orientado especialmente a la creación de libros, documentos científicos y técnicos que contengan fórmulas matemáticas. Este programa se utilizó para la edición del documento. www.kile.sourceforge.net/index.php

RKward, es una interfaz gráfica para el lenguaje de programación **R**. R es un lenguaje y entorno de programación para análisis estadístico y gráfico. Utilizado para hacer el análisis de series de tiempo propuesta en este trabajo. www.r-project.org/.

4.2. MÉTODOS

Los métodos utilizados en este trabajo corresponden a la obtención de la información hidrométrica, posteriormente a la aplicación de la metodología de modelos Box-Jenkins de Series de Tiempo, con el fin de realizar predicción de volúmenes de aportaciones esperadas en el vaso de almacenamiento de la presa Adolfo Ruiz Cortines.

Puesto que la metodología Box Jenkins es para analizar series de tiempo, generalmente, para aplicar técnicas de series de tiempo, requiere de datos de tamaño $N > 50$.

El análisis de series de tiempo para la metodología Box-Jenkins se hizo en el Lenguaje de programación **R** usando la serie de tiempo de aportaciones deducidas mensuales a la presa ARC, con un total de 672 datos observados. Las rutinas elaboradas para dicho análisis se muestran en el ANEXO C.

4.2.1. Obtención de la información hidrométrica

La medida del caudal, gasto o escurrimiento medio diario, se expresa como medio porque generalmente se realizan varios aforos en estaciones hidrométricas al día, se puede sintetizar a varios periodos, como mensual o anual; también se utiliza para expresar el escurrimiento, como volúmenes escurridos tanto a periodos diarios, mensuales o anuales. Esta información forma el registro histórico de los escurrimientos observados en las estaciones hidrométricas de los ríos.

En este caso se utilizan aportaciones deducidas con la observación de la variación diaria del almacenamiento del vaso.

En hidrología, la gráfica del registro histórico de aportaciones se le denomina hidrograma de aportaciones totales mensuales, para el vaso de almacenamiento de la presa Adolfo Ruiz Cortines. La serie observada del registro histórico de aportaciones mensuales se muestra en la Figura 4.9.

Cálculo de aportaciones deducidas

Las aportaciones deducidas son los volúmenes estimados de agua mediante un balance hidrológico realizado en el vaso de almacenamiento de una presa. Como ejemplo se realizó el cálculo de la aportación deducida para el 1 de junio de 1984 y la mensual del mes de junio de 1984 de la presa Adolfo Ruíz Cortines, siguiendo el procedimiento descrito en el Anexo A.

4.2. MÉTODOS

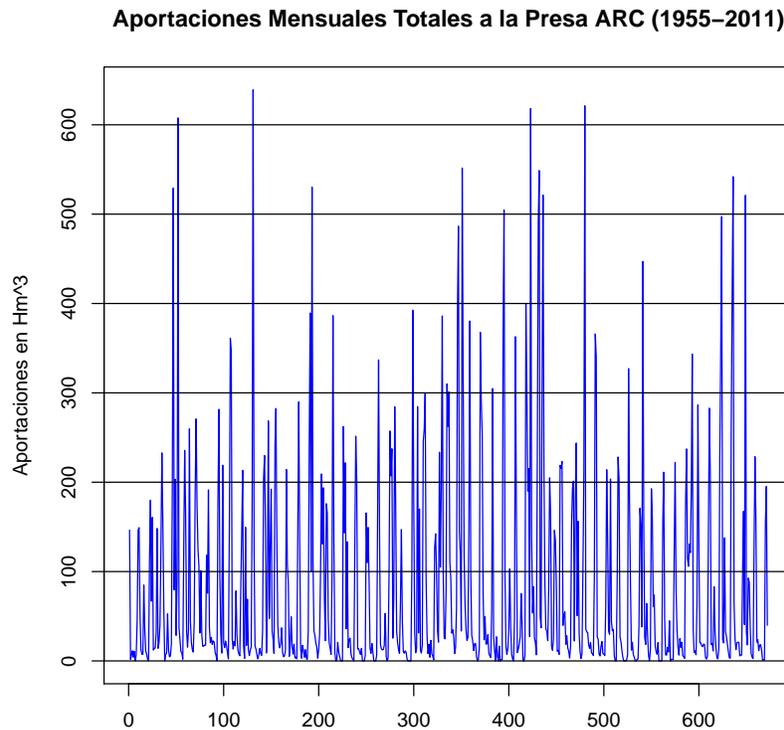


Figura 4.9. Hidrograma de Aportaciones Deducidas Totales Mensuales Z_t a la presa ARC.

4.2.2. Desarrollo del modelo de series de tiempo: Box-Jenkins

Muchas series de tiempo encontradas en la práctica presentan un comportamiento periódico. Una serie periódica tiene un patrón el cual se repite cada s periodos de tiempo, donde $s > 1$.

Uno de los tipos más comunes de comportamiento periódico es la *variación estacional*. Los modelos ARIMA para series estacionales se construyen usando el mismo procedimiento iterativo descrito anteriormente.

En esta sección se presenta el análisis estadístico para la construcción de modelos estacionales de serie de tiempo, basados en la metodología Box-Jenkins: Identificación, Estimación, Verificación y Uso del modelo. Así mismo se discuten las condiciones de estacionaridad e Invertibilidad para el modelo estacional de los datos de aportaciones deducidas mensuales de la presa Adolfo Ruiz Cortines.

4.2. MÉTODOS

Identificación del modelo ARIMA estacional

Como primer paso del análisis de la serie, se presenta la gráfica de Z_t vs t con los datos del Cuadro del anexo B, con la finalidad de observar el comportamiento de la media y varianza de la serie a lo largo del tiempo. La gráfica se muestra en la Figura 4.9, en la que notamos claramente que existen datos que se desvían muchos de los demás, esto se puede deber a dos cosas; a la naturaleza misma de los datos o bien, que sean observaciones aberrantes en la serie.

Con el fin de eliminar una tendencia determinista causada por la media ya que esta es una constante estadísticamente diferente de cero, se presenta la gráfica de la serie trasformada con desviaciones respecto a su media muestral, cuyo valor fue $\bar{Z}_t = 83.031$, estimado por 3.28. La serie \tilde{Z} se obtiene mediante:

$$\tilde{Z}_t = Z_t - \bar{Z}_t \quad t = 1, 2, \dots, 672$$

Note que esta nueva serie tiene $\mu = 0$. La serie \tilde{Z}_t se presenta en la Figura 4.10.

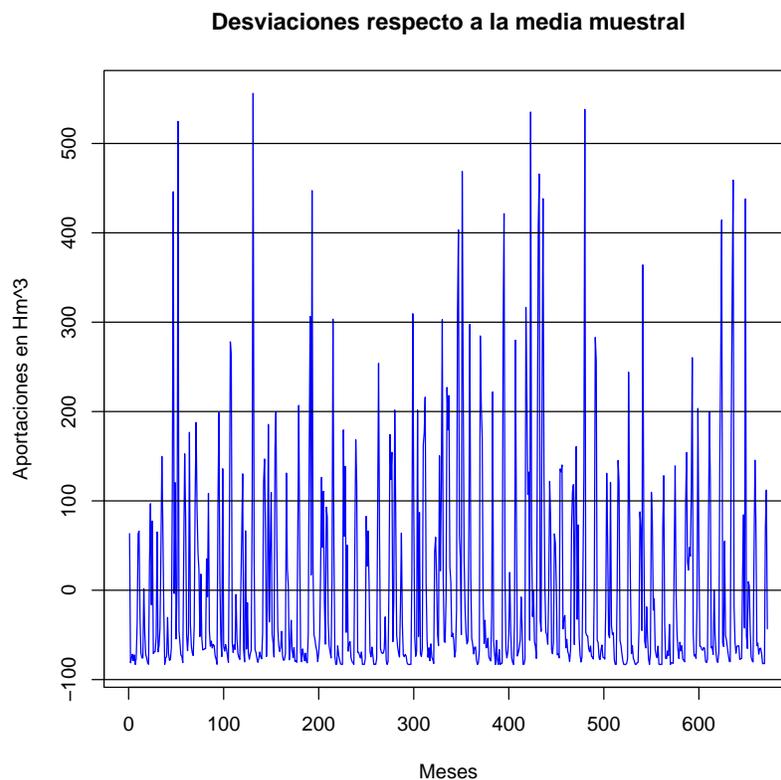


Figura 4.10. Serie de aportaciones trasformada con desviaciones respecto a la media muestral, \tilde{Z}_t .

4.2. MÉTODOS

Para mostrar de manera visual la estacionalidad de la serie \tilde{Z}_t , se gráfico un periodo de 5 años (2006-2011) de \tilde{Z}_t , los datos graficados en la Figura 4.11 corresponden a \tilde{Z}_t para $t = 612, \dots, 672$, es decir las últimas 60 observaciones de la serie (octubre de 2006 a septiembre de 2011).

Los valores de aportaciones en los meses de verano, principalmente julio y agosto, son regularmente superiores a los otros meses dentro del mismo año, mientras que las aportaciones para los meses de invierno son regularmente mas bajos.

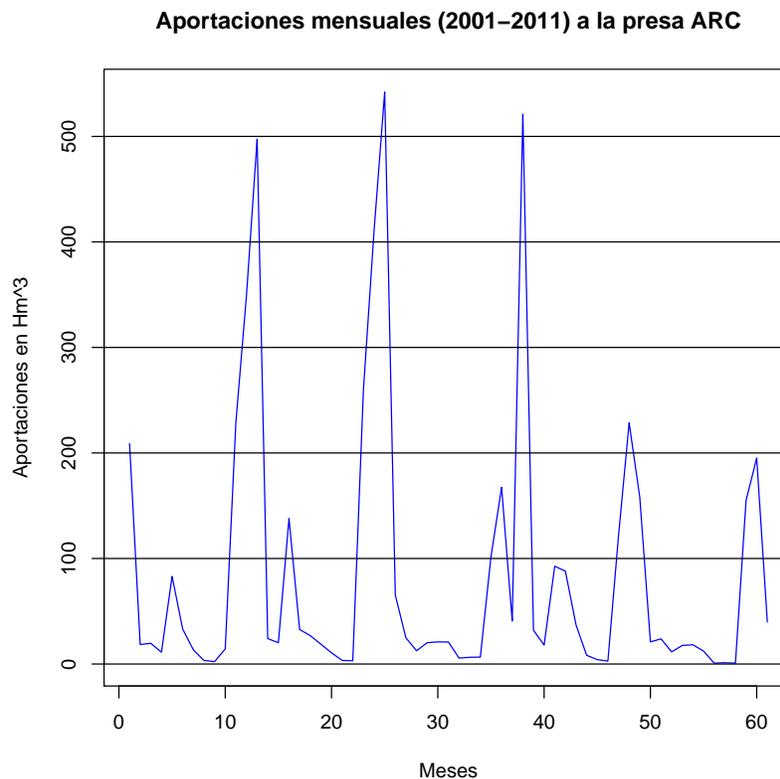


Figura 4.11. Aportaciones mensuales \tilde{Z}_t (2006-2011) $t = 612, \dots, 672$ a la presa ARC.

Lo anterior sugiere que los valores de aportaciones deducidas de un determinado mes a la presa Adolfo Ruiz Cortines son similares a los valores en el mismo mes para otros años; es decir, el valor de Julio del año uno es similar a los valores de Julio en otros años, y así sucesivamente para cada mes. En base a esto suponemos un patrón estacional en la serie de aportaciones a la presa ARC que se repite cada 12 observaciones, $s = 12$.

Para realizar el ajuste de modelos Box-Jenkins es necesario que la serie sea estacionaria; es decir, que la media y la varianza no dependan del tiempo. La función estimada de autocorrelación (FAC) y Función de AutoCorrelación Parcial (FACP) de la serie \tilde{Z}_t , ayudan a tener certeza si la serie es estacionaria o no, ya que un rápido decaimiento a

4.2. MÉTODOS

cero en la ACF significa que la serie es estacionaria.

En la Figura 4.12, se muestran la FAC y FACP para la serie \tilde{Z}_t . Note que la FAC confirma el carácter no estacionario de la variación estacional en los datos observados de aportaciones, dado a los ciclos y similitudes que se forman en las *lag* estacionales (12,24,...), las cuales a su vez decaen lentamente a cero.

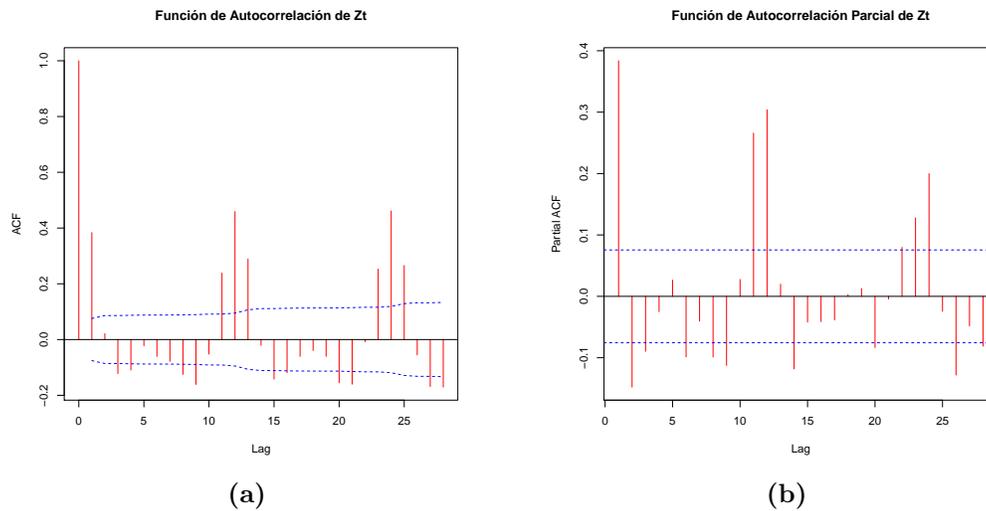


Figura 4.12. FAC (a) y FACP (b) estimadas para la serie \tilde{Z}_t .

Los valores estimados de FAC y FACP graficados en la figura 4.12 para la serie \tilde{Z}_t se presentan en el Cuadro 4.1.

Cuadro 4.1. Valores estimados de FAC y FACP para \tilde{Z}_t

lag	FAC	FACP	Lag	FAC	FACP
0	1.000		14	-0.021	-0.118
1	0.384	0.384	15	-0.142	-0.042
2	0.021	-0.148	16	-0.119	-0.041
3	-0.122	-0.089	17	-0.061	-0.039
4	-0.109	-0.025	18	-0.040	0.003
5	-0.022	0.026	19	-0.060	0.013
6	-0.060	-0.099	20	-0.155	-0.083
7	-0.078	-0.040	21	-0.160	-0.004
8	-0.125	-0.099	22	-0.008	0.080
9	-0.161	-0.113	23	0.253	0.128
10	0.052	0.028	24	0.462	0.200
11	0.238	0.266	25	0.266	-0.024
12	0.459	0.304	26	-0.055	-0.128
13	0.289	0.020	27	-0.169	-0.048
			28	-0.170	-0.081

4.2. MÉTODOS

Series observadas en la práctica, tal como es el caso de las aportaciones deducidas a la presa ARC, contienen patrones tanto estacionales como no estacionales. Cuando la media de una realización cambia acorde a un patrón estacional, como en el caso de la Figura 4.11, una *diferenciación* con frecuencia induce a tener una media constante y por tanto convierte a la serie en estacionaria.

En la Figura 4.13 se muestra la serie \tilde{Z}_t diferenciada a distancia $d = 1$, entonces la nueva serie W_t se obtuvo mediante:

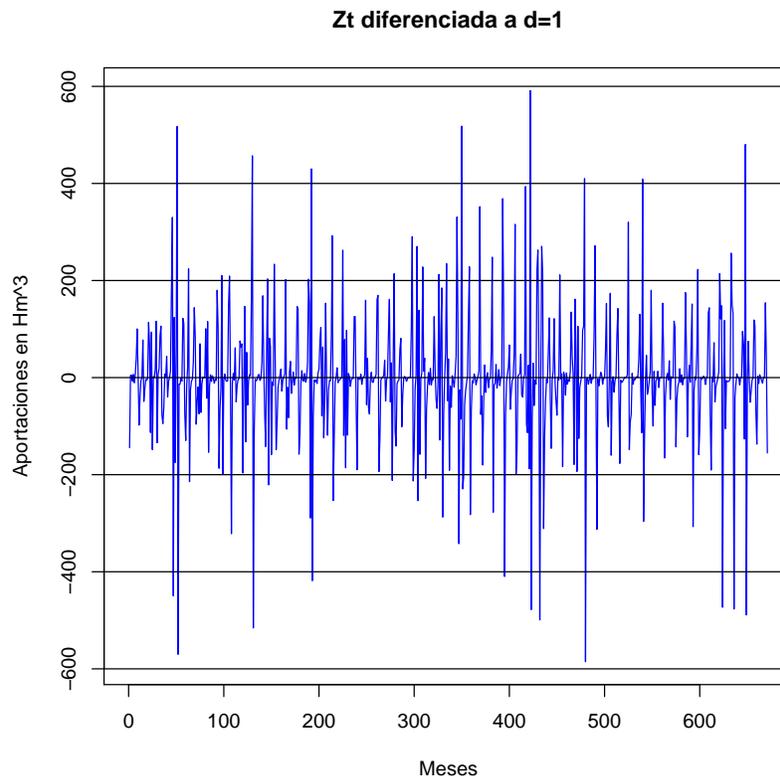


Figura 4.13. Serie $W_t = (1 - B)\tilde{Z}_t$.

$$\nabla \tilde{Z}_t = \tilde{Z}_t - \tilde{Z}_{t-1}$$

O bien:

$$W_t = (1 - B)\tilde{Z}_t$$

De un análisis visual a la Figura 4.13 parece ser que se ha estabilizado la media y la varianza de la serie; en las FAC y FACP de la serie W_t mostradas en la Figura 4.14 se puede observar que para las *lag* 1, 2, ..., 11 que están dentro de un ciclo estacional convergen rápidamente a cero; sin embargo hay correlaciones altas (picos) en las *lag*

4.2. MÉTODOS

estacionales 12, 24, 36, ... y estas decaen lentamente a cero, lo que sugiere la necesidad de una *diferencia estacional* a distancia, $s = 12$, entonces esta nueva transformación de W_t esta dada por:

$$\nabla_{12}\nabla\tilde{Z}_t = \tilde{Z}_t - \tilde{Z}_{t-1}$$

O bien:

$$W_t = (1 - B^{12})(1 - B)\tilde{Z}_t$$

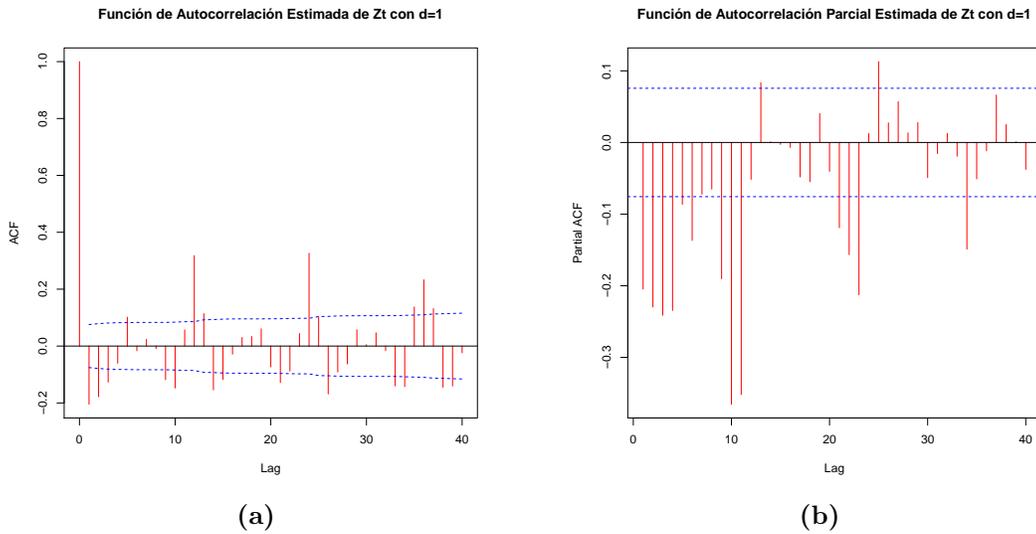


Figura 4.14. FAC (a) y FACP (b) estimadas para la serie $W_t = (1 - B)\tilde{Z}_t$.

Los valores estimados de FAC y FACP graficados en la figura 4.14 para la serie W_t se presentan en el Cuadro 4.2.

4.2. MÉTODOS

Cuadro 4.2. Valores estimados de FAC y FACP para W_t

lag	FAC	FACP	Lag	FAC	FACP
0	1.000		14	-0.154	0.001
1	-0.205	-0.205	15	-0.118	-0.003
2	-0.178	-0.230	16	-0.028	-0.007
3	-0.127	-0.241	17	0.030	-0.048
4	-0.060	-0.235	18	0.034	-0.055
5	0.102	-0.086	19	0.061	0.041
6	-0.017	-0.137	20	-0.073	-0.040
7	0.024	-0.072	21	-0.128	-0.119
8	-0.009	-0.065	22	-0.088	-0.157
9	-0.118	-0.190	23	0.044	-0.213
10	-0.147	-0.365	24	0.326	0.013
11	0.057	-0.352	25	0.103	0.113
12	0.318	-0.051	26	-0.168	0.027
13	0.114	0.084	27	-0.091	0.057
			28	-0.063	0.014

La serie $W_t = (1 - B^{12})(1 - B)\tilde{Z}_t$, donde \tilde{Z}_t representan las aportaciones deducidas a la presa ARC corregidas por la media muestral dadas en la Figura 4.10, se muestra en la Figura 4.15. La FAC y FACP estimadas para esta nueva serie $W_t = (1 - B^{12})(1 - B)\tilde{Z}_t$, aparecen en la Figura 4.16.

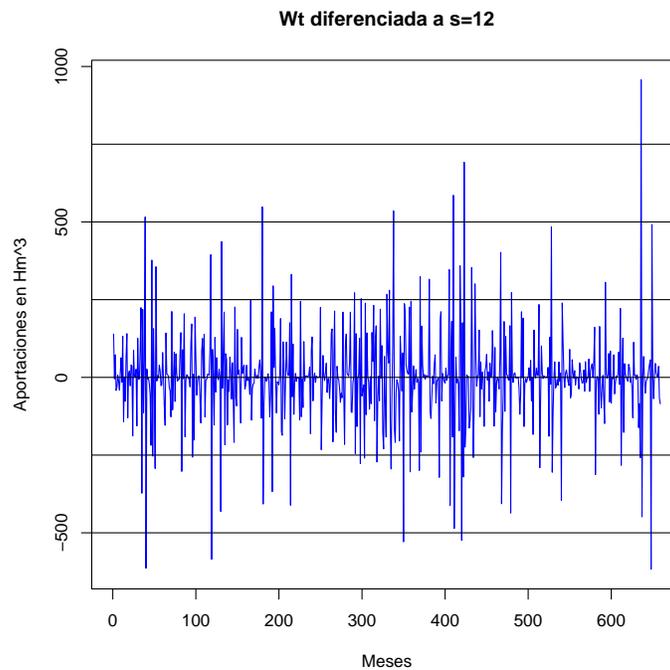


Figura 4.15. Serie $W_t = (1 - B^{12})(1 - B)\tilde{Z}_t$.

4.2. MÉTODOS

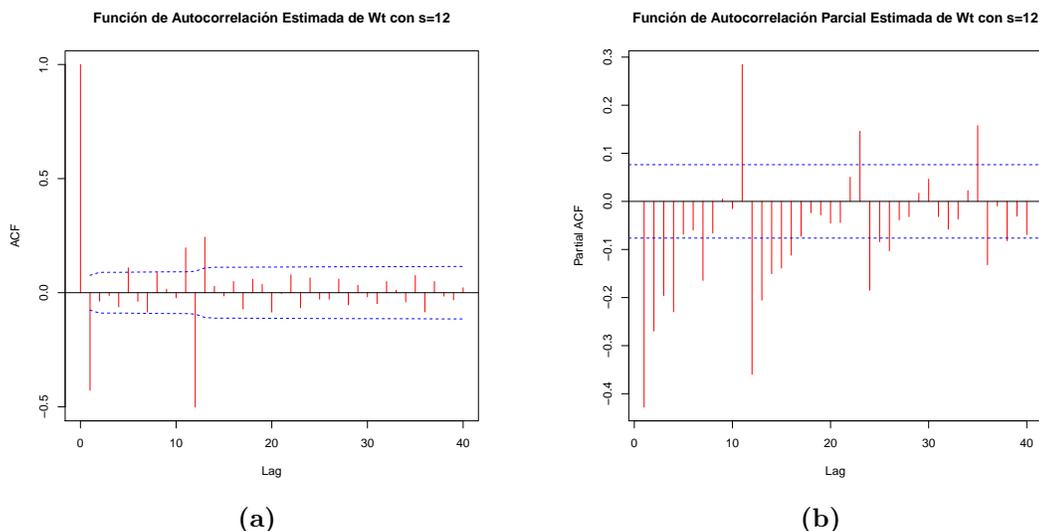


Figura 4.16. FAC (a) y FACP (b) estimadas para la serie $W_t = (1 - B^{12})(1 - B)\tilde{Z}_t$.

Los valores estimados de FAC y FACP graficados en la figura 4.16 para la serie W_t se presentan en el Cuadro 4.3.

Cuadro 4.3. Valores estimados de FAC y FACP para W_t con $s = 12$

lag	FAC	FACP	Lag	FAC	FACP
0	1.000		14	-0.150	0.001
1	-0.428	-0.428	15	-0.139	-0.003
2	-0.037	-0.269	16	-0.112	-0.007
3	-0.013	-0.196	17	-0.073	-0.048
4	-0.062	-0.230	18	-0.024	-0.055
5	0.109	-0.069	19	-0.028	0.041
6	-0.038	-0.060	20	-0.045	-0.040
7	-0.084	-0.164	21	-0.044	-0.119
8	0.087	-0.066	22	0.058	-0.157
9	0.016	-0.005	23	0.146	-0.213
10	-0.022	-0.015	24	-0.185	0.013
11	0.196	0.285	25	-0.084	0.113
12	-0.502	-0.360	26	-0.103	0.027
13	0.243	-0.205	27	-0.039	0.057
			28	-0.032	0.014

Debido a la diferenciación, tanto como la parte estacional $s = 12$ y no estacional $d = 1$ de la serie \tilde{Z}_t se han perdido observaciones, para $d = 1$ la serie $W_2 = Z_2 - Z_1 = -146.600 \text{ Hm}^3$, siendo $Z_1 = 63.368 \text{ Hm}^3$ que corresponde al mes de octubre de 1955 y $Z_2 = -81.231 \text{ Hm}^3$ al mes de noviembre de 1956 y así sucesivamente para todo

4.2. MÉTODOS

valor de t . Con respecto a $s = 12$ la serie W_t para $t = 1, 2, \dots, 671$ la cual debido a la diferenciación perdió una observación, se volvió a diferenciar $W_t = W_t - W_{t-12}$, por lo tanto 12 observaciones más se perdieron; entonces se obtuvo W_t para $t = 1, 2, \dots, 659$ y en total suman 13 observaciones perdidas debido a la diferenciación $\nabla_{12}\nabla\tilde{Z}_t$ realizada a \tilde{Z}_t .

Propuesta del modelo estacional. Las características de la FAC mostrada en la figura 4.16 para la serie $\nabla_{12}\nabla\tilde{Z}_t$ muestra una correlación alta para la $lag=12$ con pequeños picos en los valores de $lag=24, 36$ las cuales están dentro de las bandas de confianzas; por lo que se consideran no significativas. Con respecto a la FACP de la figura 4.16 se notan picos altos en las correlaciones para $lag=12, 24, 36$.

Tomando como referencia el Cuadro 3.4 para analizar la Figura 4.16, la FAC sugiere que existen componentes estacionales de *promedios móviles* de orden $Q = 1$ asociado a $lag=12$. Por otra parte la FACP sugiere componentes estacionales *autoregresivos* de un posible orden $P = 2$ asociados $lag=12, 24$. Debido al hecho de que tanto FAC y FACP pueden ser relacionados con los componentes estacionales de la serie, se necesitan ambos componentes estacionales $AR(2)$ y $MA(1)$ en el modelo.

Inspeccionando la FAC y la FACP de la Figura 4.16 dentro de un ciclo estacional $lag=1, 2, \dots, 11$, parece que tanto la FAC y la FACP están en decremento. Basados en el Cuadro 3.2, estos resultados indican que ambos componentes no estacionales $MA(q)$ y $AR(p)$ de orden $p > 0$ y $q > 0$ son necesarios para la obtención de un modelo adecuado, entonces consideramos $p = 1$ y $q = 1$.

De la ecuación 3.42 que expresa el modelo multiplicativo estacional conocido como $ARIMA(p, d, q)(P, D, Q)_s$. Para este caso los órdenes de la parte Estacional son $P = 2$, $D = 1$ y $Q = 1$, la $D = 1$ es debido a la diferenciación estacional $s = 12$. Y para la parte no estacional son $p = 1$, $d = 1$ y $q = 1$, ya que primeramente se diferenció a distancia $d = 1$ la serie para estabilizar a la media y la varianza en la serie de aportaciones deducidas a la presa ARC \tilde{Z}_t dada en la Figura 4.10.

Entonces sustituyendo los valores, nuestro modelo ARIMA estacional propuesto queda como:

$$\phi(B)\Phi(B^E)\nabla^d\nabla_E^D Z_t - \mu = \theta(B)\Theta(B^E)\alpha_t$$

Desarrollando los operadores $\phi(B)$, $\Phi(B^E)$, $\theta(B)$ y $\Theta(B^E)$ para los ordenes correspondientes $P = 2$, $D = 1$, $Q = 1$, y $p = 1$, $d = 1$ y $q = 1$, tenemos que:

$$(1 - \phi_1 B)(1 - \Phi_{12} B^{12} - \Phi_{24} B^{24})\nabla^1\nabla_{12}^1 \tilde{Z}_t = (1 - \theta_1 B)(1 - \Theta_{12} B^{12})\alpha_t$$

Pero $\nabla^1\nabla_{12}^1 \tilde{Z}_t = (1 - B^{12})(1 - B)\tilde{Z}_t = W_t$, entonces el modelo propuesto se escribe como

4.2. MÉTODOS

un modelo estacional tipo $ARIMA(1, 1, 1)(2, 1, 1)_{12}$ en su forma multiplicativa expresada en la ecuación 3.42, tiene la forma:

$$(1 - \phi_1 B)(1 - \Phi_{12} B^{12} - \Phi_{24} B^{24})W_t = (1 - \theta_1 B)(1 - \Theta_{12} B^{12})\alpha_t$$

Observe que la parte estacional esta compuesto por dos componentes *autoregresivos* AR(2), una diferencia estacional $D = 1$, tomadas a distancia $s = 12$ y un componente de *promedios móviles* MA(1). La parte no estacional del modelo se integra por una diferencia $d = 1$ y se compone por un componente *autorregresivo* AR(1) y un componente de *promedios móviles* MA(1).

Dado que en esta etapa la identificación del modelo; es prácticamente una inspección visual, la FAC y FACP mostradas en la Figura 4.16 pudieran ocultar información de interés para la elección de un modelo adecuado, por lo que es conveniente proponer varios modelos más. Dadas las características de la FAC y la FACP de la Figura 4.16 se proponen los siguientes dos modelos a partir del modelo $ARIMA(1, 1, 1)(2, 1, 1)_{12}$, en uno se elimina el componente de MA(Q) en la parte estacional y en el otro se elimina el componente *autorregresivo* AR(p) en la parte no estacional. Es decir:

$$(1 - \phi_1 B)(1 - \Phi_{12} B^{12} - \Phi_{24} B^{24})W_t = (1 - \theta_1 B)\alpha_t$$

y para el modelo sin AR(p) y MA(Q):

$$(1 - \Phi_{12} B^{12} - \Phi_{24} B^{24}) = (1 - \theta_1 B)\alpha_t$$

Los tres modelos estacionales propuestos para realizar el pronóstico de aportaciones deducidas a la presa Adolfo Ruiz Cortines son los siguientes:

Un modelo inicial con estructura completa:

$$ARIMA(1, 1, 1)(2, 1, 1)_{12}$$

Un modelo a partir del inicial, sin el componente MA(Q=1) en la parte estacional:

$$ARIMA(1, 1, 1)(2, 1, 0)_{12}$$

Y finalmente otro modelo sin MA(Q) en la parte estacional y sin componentes Autorregresivos AR(p) en la parte no estacional:

$$ARIMA(0, 1, 1)(2, 1, 0)_{12}$$

4.2. MÉTODOS

Este tipo de modelos estacionales también son conocidos en la práctica con el nombre de modelos $SARIMA(p, d, q)(P, D, Q)_s$ *Seasonal Autoregressive Integrated Moving Avarange*.

Estimación de parámetros

En la etapa de identificación se ha seleccionado tentativamente tres modelos que probablemente proporcionan representaciones parsimoniosas y adecuadas estadísticamente de los datos observados. Para lograr esta selección se calcularon varias estadísticos (Función de Autocorrelación y Función de Autocorrelación Parcial) que nos ayudaron a lograr la identificación de los modelos.

En esta etapa se estimaron los parámetros correspondientes para cada uno de los tres modelos propuestos en la etapa de identificación: $ARIMA(1, 1, 1)(2, 1, 1)_{12}$, el modelo sin MA(Q) $ARIMA(1, 1, 1)(2, 1, 0)_{12}$ y el modelo sin MA(Q) y AR(p) $ARIMA(0, 1, 1)(2, 1, 0)_{12}$. El método que se utilizó en la estimación de los parámetros fue el de máxima verosimilitud, bajo el supuesto de una distribución Normal con media cero y función de autocovarianzas $\gamma_{(h)}$ dada en la Ecuación 3.16. Los Valores de los parámetros correspondientes a cada modelo sugerido se presentan en el Cuadro 4.4.

Cuadro 4.4. Estimadores de Máxima Verosimilitud de los modelos identificados

MODELO	Componentes No-estacionales		Componentes Estacionales		
	AR(p)	MA(q)	AR(P1)	AR(P2)	MA(Q)
—					
SARIMA(1,1,1)(2,1,1)	0.197	-0.978	0.041	0.059	-1.000
SARIMA(1,1,1)(2,1,0)	0.218	-1.000	-0.630	-0.243	-
SARIMA(0,1,1)(2,1,0)	-	-0.885	-0.666	-0.268	-

Del Cuadro 4.4 los modelos propuestos tienen la siguiente forma:

Para el modelo completo $SARIMA(1, 1, 1)(2, 1, 1)_{12}$:

$$(1 - 0.197B)(1 - 0.041B^{12} - 0.059B^{24})W_t = (1 - 0.978B)(1 - 1.000B^{12})\alpha_t$$

Para el modelo sin el componente MA(Q) en la parte estacional $SARIMA(1, 1, 1)(2, 1, 0)_{12}$:

$$(1 - 0.218B)(1 - 0.630B^{12} - 0.243B^{24})W_t = (1 - 1.000B^{12})\alpha_t$$

y para el modelo sin MA(Q) y AR(p) $SARIMA(0, 1, 1)(2, 1, 0)_{12}$:

$$(1 - \Phi_{12}B^{12} - \Phi_{24}B^{24}) = (1 - 0.666B^{12} - 0.2687B^{24})W_t = (1 - 0.8853B)\alpha_t$$

4.2. MÉTODOS

El Criterio de Información de Akaike (AIC) dado en la ecuación 3.27 para cada modelo estimado se presentan en el Cuadro 4.5.

Cuadro 4.5. Valores de AIC de los modelos identificados

MODELOS	Valor de AIC
$SARIMA(1, 1, 1)(2, 1, 1)_{12}$	7804.51
$SARIMA(1, 1, 1)(2; 1, 0)_{12}$	7948.91
$SARIMA(0, 1, 1)(2, 1, 0)_{12}$	7973.80

En base al Cuadro 4.5 se sugiere que el modelo adecuado es el $SARIMA(1, 1, 1)(2, 1, 1)_{12}$ dado que tiene el menor valor de AIC, sin embargo los parámetros del modelo seleccionado deben cumplir las condiciones de estacionaridad e invertibilidad dadas en el Cuadro 3.3.

Elección del modelo. Para el modelo $SARIMA(1, 1, 1)(2, 1, 1)_{12}$ las condiciones de Estacionaridad para los componentes autoregresivos se cumplen para la parte no estacional: $|\phi_1| = |0.197| < 1$ y para la parte estacional del modelo los componentes autoregresivos estacionales cumplen con las tres condiciones: $\Phi_{12} + \Phi_{24} = 0.041 + 0.059 < 1$, $\Phi_{24} - \Phi_{12} = 0.059 - 0.041 < 1$ y $|\Phi_{24}| = |0.059| < 1$. Sin embargo los componentes de promedios móviles cumplen solo con la condición de Invertibilidad en la parte no estacional dado que $|\theta| = |-0.978| < 1$, pero para la parte estacional la condición de invertibilidad no se cumple dado que el parámetro estimado es $\Theta_{12} = -1.000$ y la condición es $|\Theta| < 1$.

En lo que corresponde al modelo $SARIMA(1, 1, 1)(2, 1, 0)_{12}$, las condiciones de estacionaridad para los componentes autoregresivos tanto en la parte estacional como en la no estacional se cumplen para los parámetros estimados, no obstante la condición de invertibilidad para el componente de promedios móviles en la parte no estacional del modelo no se cumple, ya que $\theta_1 = -1.000$ invertibilidad es $|\theta| < 1$.

Finalmente para el modelo $SARIMA(0, 1, 1)(2, 1, 0)_{12}$ los parámetros de los componentes autoregresivos en la parte estacional $\Phi_{24} - \Phi_{12} = 0.243 - (-0.630) < 1$ cumplen con la condición de estacionaridad. Y el parámetro del componente de promedios móviles en la parte no estacional cumple con la condición de invertibilidad $|\theta| < |-0.885| < 1$ invertibilidad 3.3.

En base a lo anterior, el modelo adecuado hasta el momento para la predicción de aportaciones a la presa Adolfo Ruiz Cortines es el modelo $SARIMA(0, 1, 1)(2, 1, 0)_{12}$. En la siguiente sección se realiza la verificación de este modelo.

4.2. MÉTODOS

Verificación del modelo

Una vez que se han estimado los parámetros en forma definitiva del modelo identificado, es conveniente verificar la adecuación del modelo a los datos observados.

Para verificar que el modelo $SARIMA(0, 1, 1)(2, 1, 0)_{12}$ fuera adecuado para la predicción de aportaciones, se realizaron los siguientes tres pasos:

- (a) Verificar que los coeficientes de las variables incluidas en el modelo sean significativamente distintos de cero.
- (b) Aumentar términos en el modelo y verificar si son significativos estadísticamente.
- (c) Verificar que los residuales no estén correlacionados

Calidad de los coeficientes. En esta etapa el modelo $SARIMA(1; 1; 0)(2; 1; 0)_{12}$ se verificó que los coeficientes θ_1 , Φ_{12} y Φ_{24} fueran estadísticamente distintos de cero, lo que es equivalente a contrastar los siguientes juegos de hipótesis:

$$\begin{array}{lll} H_0 : \theta_1 = 0 & vs & H_a : \theta_1 \neq 0 \\ H_0 : \Phi_{12} = 0 & vs & H_a : \Phi_{12} \neq 0 \\ H_0 : \Phi_{24} = 0 & vs & H_a : \Phi_{24} \neq 0 \end{array}$$

Dadas las propiedades de los estimadores de máxima verosimilitud para el modelo, se usó el estadístico de prueba $t - student$, con el siguiente criterio:

$$Rechazar \quad H_0 : t_c = \frac{|\hat{\beta}_i|}{\sqrt{\hat{\Sigma}_{ii}}} > t_{\alpha, \nu}$$

Donde $\hat{\beta}_i$ es el estimador de β_i y $\hat{\Sigma}_{ii}$ es la varianza del estimador i . La prueba se realizó con un nivel de significancia de $\alpha = 0.05$. En el Cuadro $t_{\alpha, \nu}$ dada en el Anexo D, el valor de t_{tablas} es 1.96, por lo que se rechaza H_0 con valores de t_c mayores que 1.96. En el Cuadro 4.6 se muestran los valores del estimador $\hat{\beta}_i$, su varianza $\hat{\Sigma}_{ii}$ y el estadístico de prueba t_c para cada parámetro del modelo $SARIMA(0, 1, 1)(2, 1, 0)_{12}$.

Dados los valores de t_c para cada estimador de máxima verosimilitud en el Cuadro 4.6 se rechazó la H_0 para todos los parámetros; puesto que $\theta_1, \Phi_{12}, \Phi_{24} > 1.96$, concluyendo que tanto para los parámetros autorregresivos de la parte no estacional, así como para la

4.2. MÉTODOS

Cuadro 4.6. Valores de $\hat{\beta}_i$, $\hat{\Sigma}_{ii}$ y t_c para el modelo $SARIMA(0, 1, 1)(2, 1, 0)_{12}$.

Parámetros	$\hat{\beta}_i$	$\hat{\Sigma}_{ii}$	t_c
θ_1	-0.8853	0.00093	28.9313
Φ_{12}	-0.6666	0.00140	17.7760
Φ_{24}	-0.2687	0.00141	7.1273

parte estacional y el componente de promedios móviles del modelo son significativamente distintos de cero, es decir se acepta H_a ; lo que significa que los parámetros estimados son estadísticamente adecuados.

Aumento de términos. En la parte de estimación de parámetros se determinó que los componentes de promedios móviles no cumplían con las condiciones de invertibilidad en la parte estacionaria, por lo que se descartaban los modelos que contenían dichos componentes. Para validar nuestro modelo se aumentó un componente a la vez en la parte autorregresiva y de promedios móviles. Se verificó que fueran estadísticamente significativos.

Aumento en el componente MA(q). $SARIMA(0, 1, 1)(2, 1, 0)_{12} \rightarrow SARIMA(0, 1, 2)(2, 1, 0)_{12}$.

En el Cuadro 4.7 se muestran los valores del estimador $\hat{\beta}_i$, varianza $\hat{\Sigma}_{ii}$ y el estadístico de prueba t_c para cada parámetro del modelo $SARIMA(0, 1, 2)(2, 1, 0)_{12}$.

Cuadro 4.7. Valores de $\hat{\beta}_i$, $\hat{\Sigma}_{ii}$ y t_c para el modelo $SARIMA(0, 1, 2)(2, 1, 0)_{12}$.

Parámetros	$\hat{\beta}_i$	$\hat{\Sigma}_{ii}$	t_c
θ_1	-0.8086	0.0013	22.3988
θ_2	-0.1914	0.0012	5.3463
Φ_{12}	-0.6304	0.0014	16.7214
Φ_{24}	-0.2422	0.0014	6.3569

Los valores de t_c para cada estimador en el Cuadro 4.7 rechazan H_o para todos los parámetros; puesto que $\theta_1, \theta_2, \Phi_{12}, \Phi_{24} > 1.96$, concluyendo que tanto para los parámetros autorregresivos de la parte estacional y los parámetros de promedios móviles en la parte no estacional del modelo son significativamente distintos de cero, también los parámetros de este modelo cumplen con las condiciones de estacionaridad dadas en el Cuadro 3.3.

Aumento en el componente AR(P). $SARIMA(0, 1, 2)(2, 1, 0)_{12} \rightarrow SARIMA(0, 1, 2)(3, 1, 0)_{12}$.

En el Cuadro 4.8 se muestran los valores del estimador $\hat{\beta}_i$, varianza $\hat{\Sigma}_{ii}$ y el estadístico de prueba t_c para cada parámetro del modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$.

Los valores de t_c para cada estimador en el Cuadro 4.8 rechazan H_o para todos los parámetros; puesto que $\theta_1, \theta_2, \Phi_{12}, \Phi_{24}, \Phi_{36} > 1.96$, entonces los parámetros estimados

4.2. MÉTODOS

Cuadro 4.8. Valores de $\hat{\beta}_i$, $\hat{\sum}_{ii}$ y t_c para el modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$.

Parámetros	$\hat{\beta}_i$	$\hat{\sum}_{ii}$	t_c
θ_1	-0.8074	0.00134	22.0601
θ_2	-0.1925	0.00131	5.3176
Φ_{12}	-0.6903	0.00142	18.2619
Φ_{24}	-0.3963	0.00195	8.9660
Φ_{36}	-0.2457	0.00152	6.3000

en este modelo son significativamente distintos de cero. De igual forma los parámetros estimados cumplen con las condiciones de estacionaridad e invertibilidad presentadas en el Cuadro 3.3.

Los tres modelos anteriores son estadísticamente igual de validos, ya que sus parámetros estimados cumplen con las condiciones de estacionaridad e invertibilidad y sus coeficientes fueron estadísticamente significativos.

El modelo con menor AIC fue $SARIMA(0, 1, 2)(3, 1, 0)_{12}$ con un valor de $AIC = 7917.09$, seguido del modelo $SARIMA(0, 1, 2)(2, 1, 0)_{12}$ con un $AIC = 7953.37$ y para el modelo $SARIMA(0, 1, 1)(2, 1, 0)_{12}$ el valor fue $AIC = 7973.80$. Siendo el modelo con más bajo AIC al que se verificaron sus residuales.

Residuales no correlacionados. La herramienta analítica base para el análisis en esta etapa es la *FAC de residuales*. Los residuales \hat{a}_t del modelo se obtuvieron mediante la ecuación 3.23. La FAC de residuales es básicamente la misma que para cualquier otra FAC, la diferencia es que ahora usamos los residuales (\hat{a}_t) de un modelo estimado en lugar de las realizaciones Z_t de modelo. Los coeficientes de autocorrelación para los residuales se obtuvieron mediante la ecuación 3.22.

Para verificar si los residuales del modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$ no están correlacionados se realizó la siguiente prueba de hipótesis:

$$H_0 : \hat{\rho}_1 = \hat{\rho}_2 = \dots = \hat{\rho}_k = 0 \quad vs \quad H_a : \text{Al menos una es } \neq 0$$

El estadístico de prueba es el de Ljung-Box Q_k , con el criterio de rechazar H_0 si:

$$Q_k > \chi_{gl, (1-\alpha)}$$

Donde el estadístico de Ljung-Box Q_k se calculó mediante la ecuación 3.46, para $k = 20$. El valor de $\chi_{(\nu, (1-\alpha))}$ se obtuvo del Cuadro del anexo D con $\alpha = 0.01$ y $gl = 15$, el cual fue: $\chi_{15, 0.99} = 30.6$. La FAC estimada para los residuales se muestran de manera gráfica en

4.2. MÉTODOS

la Figura 5.2, en la que se observa que las correlaciones de residuales para las $lag = 2, 5$ salen ligeramente de las bandas de confianza lo que nos indica de manera visual que los residuales no están correlacionados.

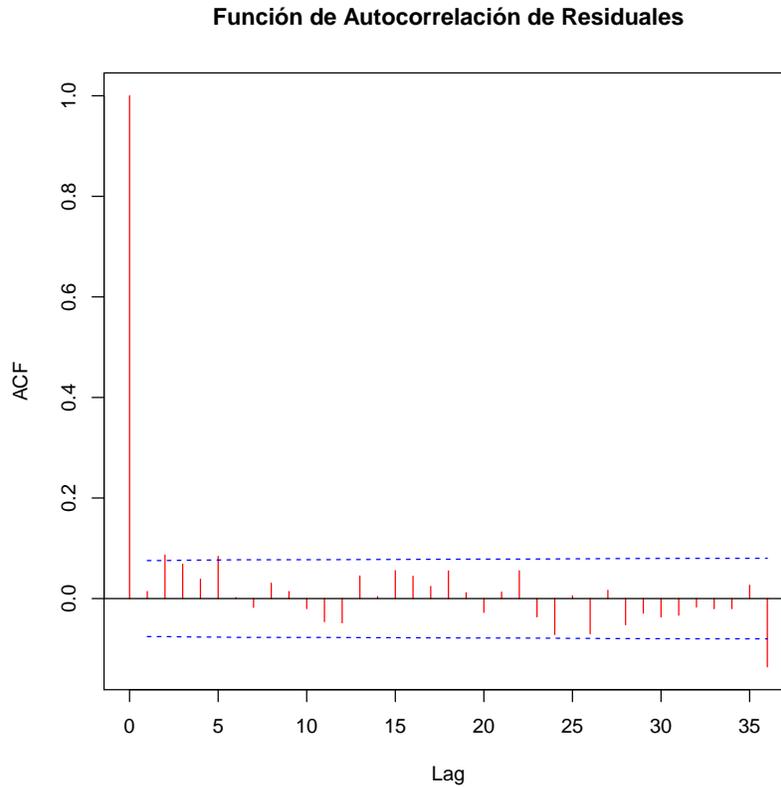


Figura 4.17. Función de AutoCorrelación de Residuales.

Los valores estimados de FAC para los residuales (\hat{a}_t), graficados en la figura 5.2 se presentan en el Cuadro 4.9.

4.2. MÉTODOS

Cuadro 4.9. Valores de $\hat{\rho}_k$, para residuales del modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$.

Lag	$\hat{\rho}_k$	Lag	$\hat{\rho}_k$
1	0.014	13	0.045
2	0.087	14	0.004
3	0.068	15	0.056
4	0.039	16	0.044
5	0.084	17	0.024
6	0.002	18	0.055
7	-0.018	19	0.012
8	0.031	20	-0.027
9	0.014	21	0.013
10	-0.020	22	0.050
11	-0.046	23	-0.036
12	-0.048	24	-0.072

Entonces, el valor del estadístico de Box-Ljung fue:

$$Q_{20} = (672)(674) \sum_{k=1}^{20} (n-k)^{-1} r_k^2$$

$$Q_{20} = (672)(674)[(1/671)0.014^2 + \dots + (1/652) - 0.027^2]$$

$$Q_{20} = 26.5061$$

Dado que $Q_k = 26.5061$ es menor que $\chi_{15,0.99} = 30.6$ nos indica que los residuales del modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$ no están correlacionados; es decir, se distribuyen como un proceso de ruido blanco $\hat{a}_t \sim WN(0, \sigma^2)$.

Finalmente nuestro el cumple con las tres condiciones de verificación por lo que es estadísticamente adecuado, por lo tanto el modelo que se adopto para la predicción de las aportaciones deducidas mensuales a la presa ARC es:

$$SARIMA(0, 1, 2)(3, 1, 0)_{12}$$

La forma algebraica del modelo anterior se presentan en la siguiente sección.

Uso del modelo para predicción

Después de haber considerado las etapas de Identificación, Estimación y Verificación del modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$, se utiliza el modelo para predecir valores futuros de la serie de tiempo de aportaciones deducidas observada a la presa Adolfo Ruiz Cortines.

4.2. MÉTODOS

La **Forma de la Ecuación de Diferencia** del modelo es la más apropiada para obtener las predicciones de la serie. El proceso consiste en desarrollar los binomios usados para volver estacionaria la serie $(1 - B)^D(1 - B)^d$. Entonces el modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$, en forma de una ecuación de diferencia se expresa como:

$$(1 - \Phi_{12}B^{12} - \Phi_{24}B^{24} - \Phi_{36}B^{36})(1 - B^{12})(1 - B)\tilde{Z}_t = (1 - \theta_1B - \theta_2B^2)\alpha_t$$

Expandiendo los polinomios $\Phi_P(B^P)$, $\theta_q(B^q)$, $(1 - B)^D(1 - B)^d$ y despejando \tilde{Z}_t se tiene:

$$[1 - B - B^{12} + B^{13} - \Phi_{12}(B^{12} - B^{13} - B^{24} + B^{25}) - \Phi_{24}(B^{24} - B^{25} - B^{36} + B^{37}) - \Phi_{36}(B^{36} - B^{37} - B^{48} + B^{49})]\tilde{Z}_t = \alpha_t - \theta_1\alpha_{t-1} - \theta_2\alpha_{t-2}$$

Sustituyendo el valor de los coeficientes estimados en la Tabla 4.8:

$$\begin{aligned} &\tilde{Z}_t - \tilde{Z}_{t-1} - \tilde{Z}_{t-12} + \tilde{Z}_{t-13} + 0.60903(\tilde{Z}_{t-12} - \tilde{Z}_{t-13} - \tilde{Z}_{t-24} + \tilde{Z}_{t-25}) + \\ &0.3963(\tilde{Z}_{t-24} - \tilde{Z}_{t-25} - \tilde{Z}_{t-36} + \tilde{Z}_{t-37}) + 0.2457(\tilde{Z}_{t-36} - \tilde{Z}_{t-37} - \tilde{Z}_{t-48} + \tilde{Z}_{t-49}) = \\ &\alpha_t - \theta_1\alpha_{t-1} - \theta_2\alpha_{t-2} \end{aligned}$$

Despejando \tilde{Z}_t :

$$\begin{aligned} \tilde{Z}_t = &\tilde{Z}_{t-1} + \tilde{Z}_{t-12} - \tilde{Z}_{t-13} - 0.60903(\tilde{Z}_{t-12} - \tilde{Z}_{t-13} - \tilde{Z}_{t-24} + \tilde{Z}_{t-25}) - \\ &0.3963(\tilde{Z}_{t-24} - \tilde{Z}_{t-25} - \tilde{Z}_{t-36} + \tilde{Z}_{t-37}) - 0.2457(\tilde{Z}_{t-36} - \tilde{Z}_{t-37} - \tilde{Z}_{t-48} + \tilde{Z}_{t-49}) + \\ &\alpha_t - \theta_1\alpha_{t-1} - \theta_2\alpha_{t-2} \end{aligned}$$

En el pronóstico de series de tiempo son de interés los valores futuros denotados como Z_{t+L} donde $L > 1$, entonces el modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$ en la forma de ecuación de diferencia para realizar de aportaciones tiene la forma:

$$\begin{aligned} \tilde{Z}_{t+L} = &\tilde{Z}_{t+L-1} + \tilde{Z}_{t+L-12} - \tilde{Z}_{t+L-13} - 0.60903(\tilde{Z}_{t+L-12} - \tilde{Z}_{t+L-13} - \tilde{Z}_{t+L-24} + \tilde{Z}_{t+L-25}) - \\ &0.3963(\tilde{Z}_{t+L-24} - \tilde{Z}_{t+L-25} - \tilde{Z}_{t+L-36} + \tilde{Z}_{t+L-37}) - 0.2457(\tilde{Z}_{t+L-36} - \tilde{Z}_{t+L-37} - \tilde{Z}_{t+L-48} + \tilde{Z}_{t+L-49}) + \\ &\alpha_{t+L} - \theta_1\alpha_{t+L-1} - \theta_2\alpha_{t+L-2} \end{aligned}$$

Note que $\tilde{Z}_t = Z_t - \mu$. El periodo t se llama pronostico de origen, para nuestro caso es $t = 672$. En este trabajo se realizaron los pronósticos $L = 1, 2, \dots, 24$ para las siguientes veinticuatro observaciones de la serie de aportaciones de la presa Adolfo Ruiz Cortines, $Z_{t+L} = 673, 674, \dots, 696$. Los resultados se muestran en el siguiente capitulo.

Capítulo 5

RESULTADOS

La aplicación de la metodología descrita en el Capítulo anterior, para el análisis de modelos de series de tiempo estacionales cuyo dominio es el tiempo, conocidos como modelos Box-Jenkins, se usó con el fin de realizar pronósticos para estimar las aportaciones deducidas a la presa Adolfo Ruíz Cortines. En esta parte se reseñan los resultados obtenidos para el modelo estacional generado a partir de la serie de aportaciones deducidas y sus pronósticos.

5.1. INFORMACIÓN DE APORTACIONES DEDUCIDAS

Se utilizaron las aportaciones deducidas mensuales totales del Río Mayo a la presa de almacenamiento “Adolfo Ruiz Cortines”, que es la fuente de agua que abastece al Distrito de Riego 038 Río Mayo, en el estado de Sonora, México.

El registro histórico de aportaciones deducidas cuenta con 56 años de registro desde el año 1955 hasta el año 2011. Los datos que se utilizaron en este trabajo fueron proporcionados por Subdirección de Infraestructura Hidroagrícola. Gerencia de Distritos de Riego. Subgerencia de operación, CONAGUA (2011) y la jefatura del Distrito de Riego 038 Río Mayo, Sonora, México.

5.1. INFORMACIÓN DE APORTACIONES DEDUCIDAS

5.1.1. Elaboración del Hidrograma de aportaciones históricas mensuales

El hidrograma correspondiente a las aportaciones históricas mensuales se muestra en la Figura 5.1.

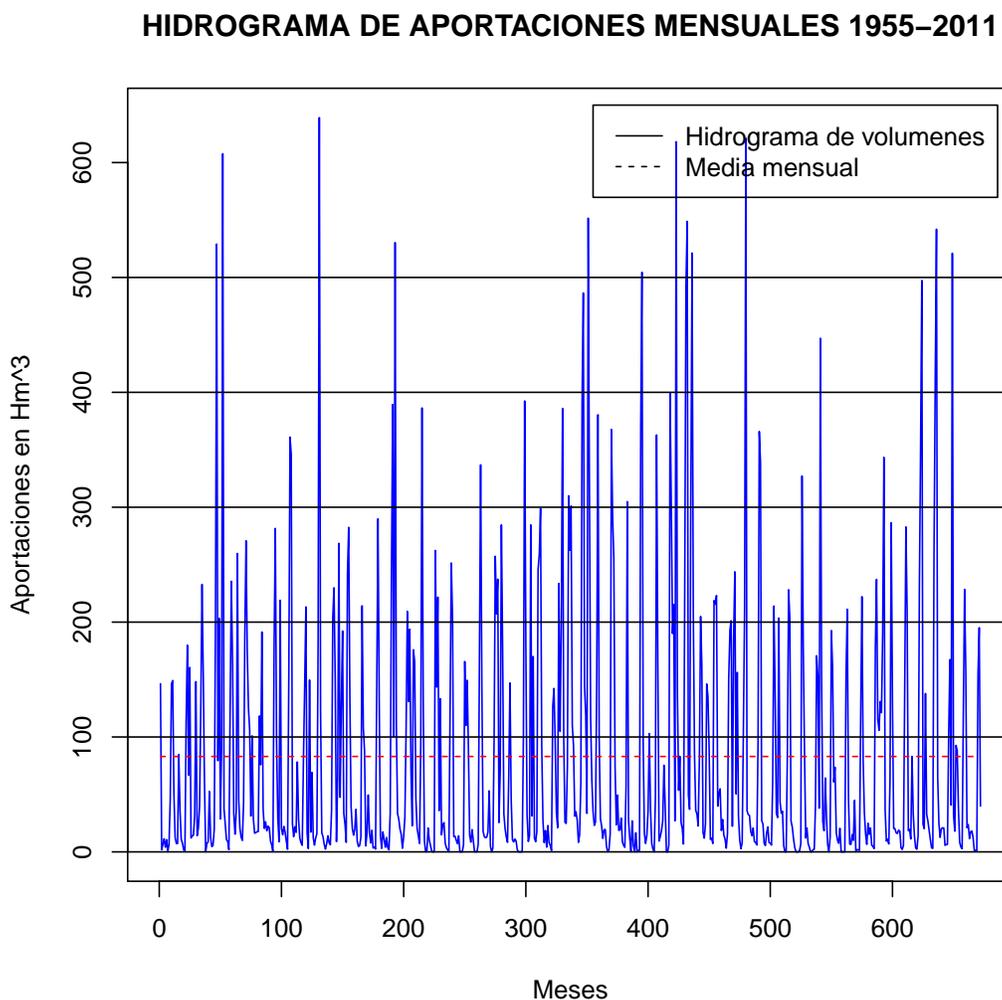


Figura 5.1. Hidrograma de Aportaciones Mensuales a la presa ARC.

En el hidrograma mostrado en la Figura 5.1, se representa la variación de los valores mensuales para cada año del registro histórico de aportaciones deducidas mensuales a la presa Adolfo Ruíz Cortines. En el Cuadro 5.1 se muestran estadísticas descriptivas de interés para la serie mensual de las aportaciones deducidas a la presa ARC registradas en el periodo de 1955-2011.

5.2. MODELOS ESTACIONALES DE SERIES DE TIEMPO: Box-Jenkins

Cuadro 5.1. Estadísticas Descriptivas de la serie de aportaciones Z_t

ESTADÍSTICA	VALOR Hm^3
Número	672
Suma	55,796.840
Máximo	639.100
Mínimo	0
Promedio	83.031
Varianza	13,637.570
Desv. Est.	116.780

Se observan para algunos años que los volúmenes de aportación son mayores en relación a otros años debido a la variación estacional presente en la serie, los valores altos de aportaciones corresponden a los meses de temporal y en contra parte los niveles más bajos corresponden a los meses de estiaje.

5.2. MODELOS ESTACIONALES DE SERIES DE TIEMPO: Box-Jenkins

Los modelos estaciones de Box-Jenkins se utilizaron ya que la serie presenta variación estacional, la metodología desarrollada para la obtención de este tipo de modelos consta de cuatro etapas: Identificación del modelo, Estimación de sus parámetro, Verificación estadística del modelo estimado y Uso del modelo con fines de predicción. Los resultados de cada etapa se muestran en seguida.

5.2.1. Estacionaridad de la Serie

Una de las condiciones fundamentales para el análisis de series de tiempo en base a la metodología desarrollada por Box y Jenkins, es que la media y la varianza de la serie no dependan del tiempo; es decir, que sea estacionaria.

La serie original observada de aportaciones dada en la Figura 5.1 se diferenció a distancia $d = 1$ para lograr la estacionaridad, la serie diferenciada se muestra en la Figura 4.13, en base a la Función de AutoCorrelación y Función de AutoCorrelación Parcial para la serie diferenciada presentadas en la Figura 4.14, se detecto un comportamiento estacional por lo que una segunda diferenciación a distancia $D = 12$ fue requerida, su gráfica se muestra en la Figura 4.15.

Para lograr la estacionaridad de la serie de aportaciones deducidas fueron necesarias dos diferencias $d = 1$ y $D = 12$, por lo que $Z_t = \nabla^D \nabla^d Z_t = W_t$.

5.2.2. Identificación del modelo estacional

La herramienta analítica fundamental para la identificación del modelo fue la Función de Autocorrelación (FAC) y Función de Autocorrelación Parcial (FACP) de la serie W_t . Retomando la Figura 4.16 que muestra la FAC y FACP se identificaron tentativamente los siguientes tres modelos:

$$ARIMA(1, 1, 1)(2, 1, 1)_{12}$$

$$ARIMA(1, 1, 1)(2, 1, 0)_{12}$$

$$ARIMA(0, 1, 1)(2, 1, 0)_{12}$$

Los modelos anteriores se propusieron en base a un análisis visual a la FAC y FACP de la serie diferenciada para eliminar la estacionalidad y homogeneizar la media y la varianza de la serie mostrada en la Figura 4.15, donde se observó los comportamientos de las lag en la parte estacional $lag = 12, 24, \dots$ para la parte estacional del modelo y dentro de un mismo ciclo $lag = 1, 2, \dots, 12$ para la parte no estacional de cada modelo.

5.2.3. Estimación de parámetros del modelo seleccionado

En el Cuadro 4.4 se encuentran los valores de los estimadores de máxima verosimilitud para cada uno de los modelos propuesto, los estimadores de los parámetros de un modelo integrado ARMA deben cumplir condiciones de estacionaridad e Invertibilidad para que el modelo sea adecuado, dichas condiciones se reseñan en el Cuadro 3.3.

Estacionaridad e Invertibilidad para $ARIMA(1, 1, 1)(2, 1, 1)_{12}$

Las condiciones de Estacionaridad para los componentes autoregresivos se cumplen para la parte no estacional y para la parte estacional del modelo: $|\phi_1| = |0.197| < 1$ y $\Phi_{12} + \Phi_{24} = 0.041 + 0.059 < 1$, $\Phi_{24} - \Phi_{12} = 0.059 - 0.041 < 1$ y $|\Phi_{24}| = |0.059| < 1$. Sin embargo los componentes de promedios móviles cumplen solo con la condición de Invertibilidad en la parte no estacional dado que $|\theta| = |-0.978| < 1$ pero no para la parte estacional.

Estacionaridad e Invertibilidad para $ARIMA(1, 1, 1)(2, 1, 0)_{12}$

En lo que corresponde al modelo $SARIMA(1, 1, 1)(2, 1, 0)_{12}$, las condiciones de estacionaridad para los componentes autoregresivos tanto en la parte estacional como

5.2. MODELOS ESTACIONALES DE SERIES DE TIEMPO: Box-Jenkins

en la no estacional se cumplen para los parámetros estimados.

La condición de invertibilidad para el componente de promedios móviles en la parte no estacional del modelo no se cumple, dado que $|\theta| = 1$ y debe ser menor.

Estacionaridad e Invertibilidad para $ARIMA(0, 1, 1)(2, 1, 0)_{12}$

Finalmente para el modelo $SARIMA(0, 1, 1)(2, 1, 0)_{12}$ los parámetros de los componentes autorregresivos en la parte estacional $\Phi_{24} - \Phi_{12} = < 1$ cumplen con la condición de estacionaridad. Y el parámetro del componente de promedios móviles en la parte no estacional cumple con la condición de invertibilidad $|\theta| < 1$ dada en el Cuadro 3.3.

Elección del modelo

El modelo adecuado para la predicción de aportaciones a la presa Adolfo Ruiz Cortines, en base a las condiciones de estacionaridad e Invertibilidad de los estimadores, fue el modelo $SARIMA(0, 1, 1)(2, 1, 0)_{12}$, en su forma algebraica:

$$(1 - \Phi_{12}B^{12} - \Phi_{24}B^{24})W_t = (1 - \theta_1B)\alpha_t$$

Sustituyendo el valor de los estimadores:

$$(1 - 0.666B^{12} - 0.2687B^{24})W_t = (1 - 0.8853B)\alpha_t$$

5.2.4. Verificación del modelo

Para verificar que el modelo $SARIMA(0, 1, 1)(2, 1, 0)_{12}$ fuera adecuado para la predicción de aportaciones, se realizaron los siguientes pasos, *Checar que los coeficientes de las variables incluidas en el modelo sean significativamente distintos de cero, Aumentar términos en el modelo y verificar si son significativos estadísticamente y Verificar que los residuales no estén correlacionados.*

Calidad de los coeficientes

En esta etapa de verificación del modelo $SARIMA(0, 1, 2)(2, 1, 0)_{12}$ se probó que los se usó el estadístico de prueba *t-student*, rechazando todos aquellos coeficientes cuyo valor fuera menor a 1.96, los resultados se muestran en el Cuadro 5.2:

5.2. MODELOS ESTACIONALES DE SERIES DE TIEMPO: Box-Jenkins

Cuadro 5.2. Coeficientes estimados del modelo $SARIMA(0, 1, 1)(2, 1, 0)_{12}$.

Parámetros	$\hat{\beta}_i$	t_c
θ_1	-0.8853	28.9313
Φ_{12}	-0.6666	17.7760
Φ_{24}	-0.2687	7.1273

Dados los valores de t_c para cada estimador de máxima verosimilitud en el Cuadro 5.2 notamos que $\theta_1, \Phi_{12}, \Phi_{24} > 1.96$, concluyendo que tanto para los parámetros autorregresivos de la parte no estacional, así como para la parte estacional y el componente de promedios móviles del modelo son significativamente distintos de cero, por lo que los parámetros estimados son estadísticamente adecuados.

Aumento de términos al modelo $SARIMA(0, 1, 1)(2, 1, 0)_{12}$

Para validar nuestro modelo se aumentó primeramente un componente en la parte autoregresiva estacional, se verificó que los estimadores de los nuevos parámetros cumplieran con las condiciones de Estacionaridad e Invertibilidad y además que fueran estadísticamente significativos. Como ambas condiciones se cumplían se aumentó un coeficiente de promedios móviles en la parte no estacionaria, por lo que nuestro modelo estimado tuvo el siguiente cambio:

$$SARIMA(0, 1, 1)(2, 1, 0)_{12} \rightarrow SARIMA(0, 1, 2)(3, 1, 0)_{12}.$$

En el Cuadro 5.3 se muestran los valores del estimador $\hat{\beta}_i$, varianza $\hat{\Sigma}_{ii}$ y el estadístico de prueba t_c para cada parámetro del modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$.

Cuadro 5.3. Coeficientes estimados del modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$.

Parámetros	$\hat{\beta}_i$	t_c
θ_1	-0.8074	22.0601
θ_2	-0.1925	5.3176
Φ_{12}	-0.6903	18.2619
Φ_{24}	-0.3963	8.9660
Φ_{36}	-0.2457	6.3000

Los valores de t_c para cada estimador en el Cuadro 5.3 notamos que $\theta_1, \theta_2, \Phi_{12}, \Phi_{24}, \Phi_{36} > 1.96$, entonces los parámetros estimados en este modelo son significativamente distintos de cero. De igual forma los parámetros estimados cumplen con las condiciones de estacionaridad e invertibilidad presentadas en el Cuadro 3.3.

5.2. MODELOS ESTACIONALES DE SERIES DE TIEMPO: Box-Jenkins

Residuales no correlacionados

La herramienta analítica base para el análisis en esta etapa fue la *FAC de residuales*. Para verificar si los residuales del modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$ no están correlacionados se realizó la siguiente prueba de Box-Ljung, con el criterio de aceptar correlación entre los datos estimados si:

$$Q_k > \chi_{gl, (1-\alpha)}$$

La FAC estimada para los residuales se muestran de manera gráfica en la Figura 5.2, en la que se observa que las correlaciones de residuales para las $lag = 2, 5$ salen ligeramente de las bandas de confianza lo que nos indica de manera visual que los residuales no están correlacionados.

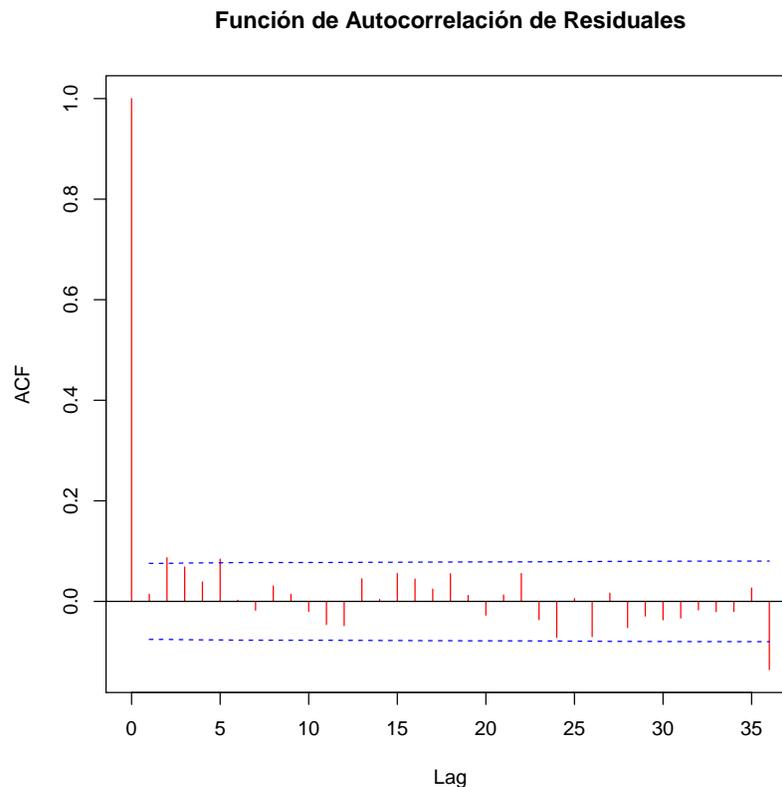


Figura 5.2. Función de AutoCorrelación de Residuales.

Entonces, el valor del estadístico de Box-Ljung para 20 autocorrelaciones y el valor de $\chi_{15,0.99}$ fue:

5.2. MODELOS ESTACIONALES DE SERIES DE TIEMPO: Box-Jenkins

$$Q_{20} = 26.5061 \quad y \quad \chi_{15,0.99} = 30.6$$

Dado que $Q_k = 26.5061$ es menor que $\chi_{15,0.99} = 30.6$ nos indica que los residuales del modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$ no están interrelacionados; es decir, se distribuyen como un proceso de ruido blanco $\hat{a}_t \sim WN(0, \sigma^2)$.

Finalmente nuestro modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$ cumple con las tres condiciones de verificación por lo que es estadísticamente adecuado.

5.2.5. Predicción de aportaciones a la presa ARC

Una vez realizadas y verificadas las primeras tres etapas de la metodología de análisis de series de tiempo (Identificación, Estimación y Verificación del modelo) bajo en enfoque Box y Jenkins, se determino que para los serie observada de aportaciones se puede modelar y así predecir valores futuros mediante el ajuste del modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$.

$$(1 - \Phi_{12}B^{12} - \Phi_{24}B^{24} - \Phi_{36}B^{36})(1 - B^{12})(1 - B)\tilde{Z}_t = (1 - \theta_1B - \theta_2B^2)\alpha_t$$

En el pronóstico de series de tiempo son de interés los valores futuros denotados como Z_{t+L} donde $L > 1$, entonces nuestro modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$ en la forma de ecuación diferencia para realizar de aportaciones tiene la forma:

$$\begin{aligned} \tilde{Z}_{t+L} = & \tilde{Z}_{t+L-1} + \tilde{Z}_{t+L-12} - \tilde{Z}_{t+L-13} - 0.60903(\tilde{Z}_{t+L-12} - \tilde{Z}_{t+L-13} - \tilde{Z}_{t+L-24} + \tilde{Z}_{t+L-25}) - \\ & 0.3963(\tilde{Z}_{t+L-24} - \tilde{Z}_{t+L-25} - \tilde{Z}_{t+L-36} + \tilde{Z}_{t+L-37}) - 0.2457(\tilde{Z}_{t+L-36} - \tilde{Z}_{t-37} - \tilde{Z}_{t-48} + \tilde{Z}_{t-49}) + \\ & \alpha_{t+L} - \theta_1\alpha_{t+L-1} - \theta_2\alpha_{t+L-2} \end{aligned}$$

En este trabajo se realizaron los pronósticos $L = 1, 2, \dots, 24$ para las siguientes veinticuatro observaciones de la serie de aportaciones de la presa Adolfo Ruiz Cortines, $Z_{t+L} = 673, 674, \dots, 696$. Los resultados se muestran en el Cuadro 5.4.

5.2. MODELOS ESTACIONALES DE SERIES DE TIEMPO: Box-Jenkins

Cuadro 5.4. Pronósticos de aportaciones a la presa Adolfo Ruiz Cortines.

L	MES	Pronóstico	L	MES	Pronóstico
1	OCT 2011	135.198	13	OCT 2012	143.102
2	NOV 2011	26.004	14	NOV 2012	26.502
3	DIC 2011	45.1438	15	DIC 2012	23.700
4	ENE 2012	44.316	16	ENE 2013	38.369
5	FEB 2012	41.818	17	FEB 2013	37.253
6	MAR 2012	23.018	18	MAR 2013	21.991
7	ABR 2012	6.684	19	ABR 2013	5.483
8	MAY 2012	3.936	20	MAY 2013	4.299
9	JUN 2012	3.357	21	JUN 2013	3.855
10	JUL 2012	163.296	22	JUL 2013	139.913
11	AGO 2012	254.163	23	AGO 2013	212.191
12	SEP 2012	198.649	24	SEP 2013	107.586

Los valores del Cuadro 5.4, se presentan gráficamente en la Figura 5.3

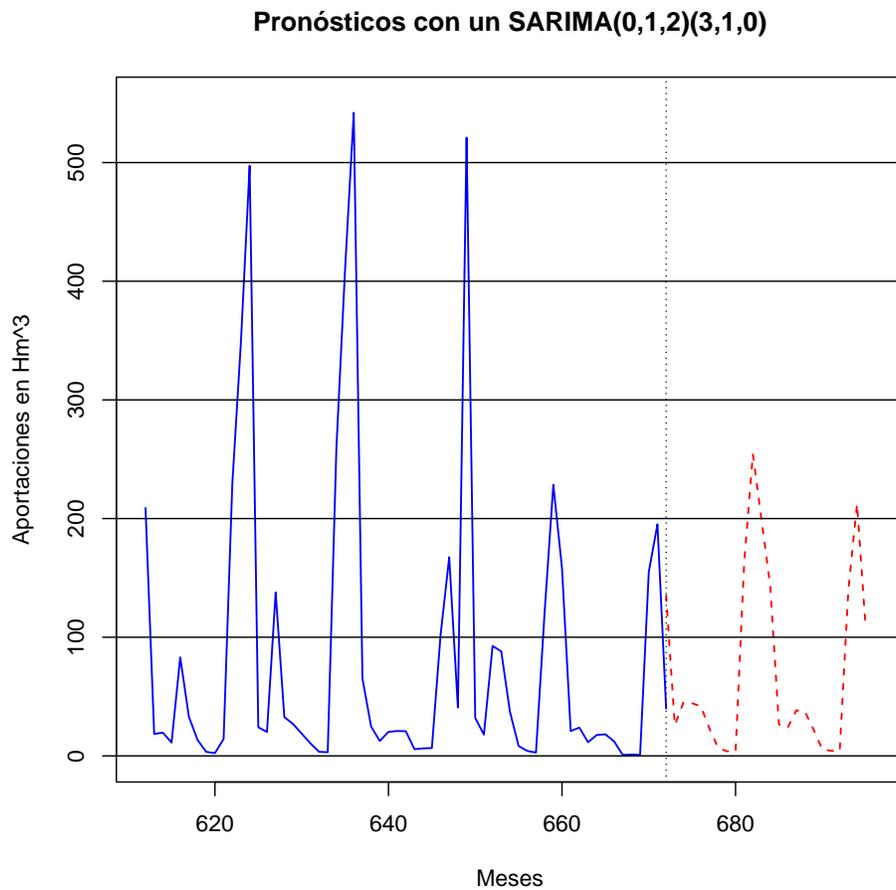


Figura 5.3. Pronósticos para dos ciclos agrícolas.

5.3. COMPARACIÓN ENTRE VALORES AJUSTADOS Y OBSERVADOS

Del Cuadro 5.4 se observa que los valores mínimos pronosticados corresponden a los meses de mayo y junio, y los meses con volúmenes mayores de aportaciones corresponden a los meses de agosto; lo anterior nos muestra que los valores pronosticados tienen una tendencia estacional.

5.3. COMPARACIÓN ENTRE VALORES AJUSTADOS Y OBSERVADOS

Una alternativa útil para verificar la adecuación del modelo, es a partir de la serie observada de aportaciones deducidas a la presa ARC $Z_t = 1, \dots, 672$, tomar una parte de la serie $Z_t = 1, \dots, 660$ equivalente a tener una serie observada de 1955 al 2010, entonces el procedimiento consiste en ajustar esta parte de la serie al modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$ y realizar pronósticos para el ciclo agrícola 2010-2011 y comparar los valores pronosticados con los valores observados.

El Cuadro 5.5 muestra los comparativos entre los valores pronosticados de octubre del 2010 al septiembre del 2011 contra los valores observados para el mismo ciclo.

Cuadro 5.5. Observados contra ajustados para el ciclo 2010-2011.

Mes	Observado	Ajustado
Oct	20.972	169.315
Nov	23.822	27.315
Dic	11.532	35.982
Ene	17.607	63.186
Feb	18.208	48.376
Mar	12.272	26.085
Abr	0.804	8.983
May	1.196	6.482
Jun	0.844	9.042
Jul	155.002	167.100
Ago	195.171	273.205
Sep	39.809	274.306

De manera gráfica los valores observados se comparan con los pronosticados en la Figura 5.4. De igual forma en la Figura 5.5 se observa la serie ajustada contra la serie observada para el registro histórico de aportaciones deducidas.

5.3. COMPARACIÓN ENTRE VALORES AJUSTADOS Y OBSERVADOS

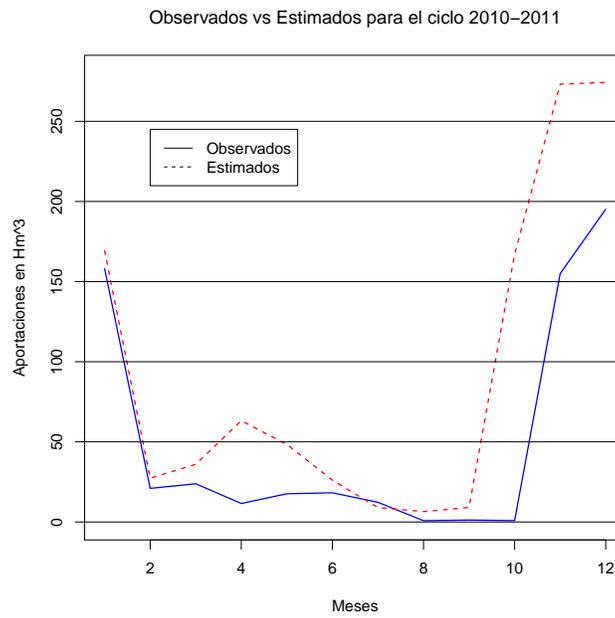


Figura 5.4. Valores observados y ajustados para el ciclo 2011-2012.

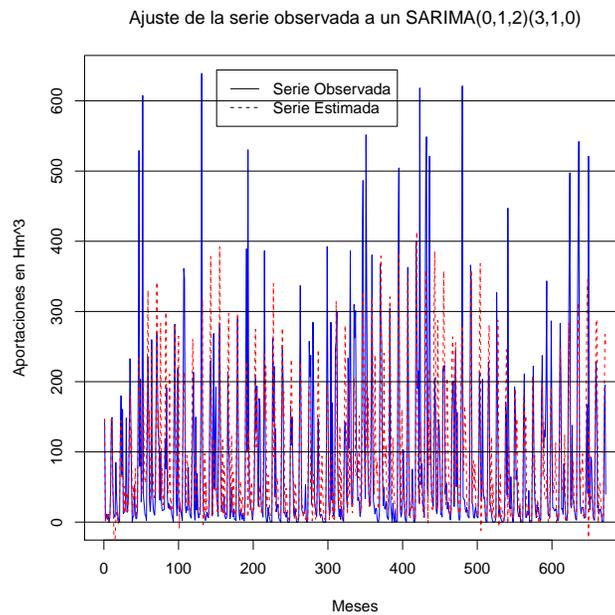


Figura 5.5. Serie ajustada del modelo $SARIMA(0,1,2)(3,1,0)_{12}$.

5.4. COMPARACIÓN CON ANÁLISIS DE FRECUENCIA

Estos valores estimados de aportaciones deducidas esperadas pronosticados a la presa Adolfo Ruiz Cortines con el modelo $SARIMA(0, 1, 2)(3, 1, 0)_{12}$ para los ciclos agrícolas 2011-2012 y 2012-2013 se pueden comparar con los valores obtenidos por el método de análisis de frecuencias, lo que proporcionará un elemento de juicio teórico más para la planeación de la operación del Distrito de Riego.

5.4.1. Pronóstico con la Función de Distribución de Probabilidad: Log Normal de 2 parámetros

Los valores de aportaciones pronosticados mediante el análisis de series de tiempo estacionales presentados en el Cuadro 5.4 se comparan gráficamente con los valores esperados de aportaciones obtenidos del ajuste de la Función de Distribución de Probabilidad Log Normal de 2 parámetros al 50 % nivel de probabilidad. Los datos del ajuste de la Función Log Normal fueron tomados de Ariza (2008) y se presentan en el Cuadro 5.6.

Cuadro 5.6. Comparación gráfica de la ST vs Análisis de frecuencia.

Mes	Medias	LN 50 %	ST 1	ST 2
Oct	82.299	68.246	135.198	143.103
Nov	33.814	31.054	26.004	26.503
Dic	72.643	61.868	45.144	23.700
Ene	74.558	28.954	44.316	38.370
Feb	53.424	29.721	41.818	37.254
Mar	37.874	21.897	23.018	21.991
Abr	10.151	7.000	6.684	5.484
May	7.270	4.170	3.936	4.300
Jun	12.134	17.235	3.357	3.855
Jul	166.256	201.834	163.296	139.913
Ago	271.162	288.629	254.163	212.191
Sep	174.787	152.640	198.649	107.587
Anual	996.372	913.248	945.583	764.251

Los valores del Cuadro 5.6 de manera gráfica se presentan en la Figura 5.6.

5.4. COMPARACIÓN CON ANÁLISIS DE FRECUENCIA

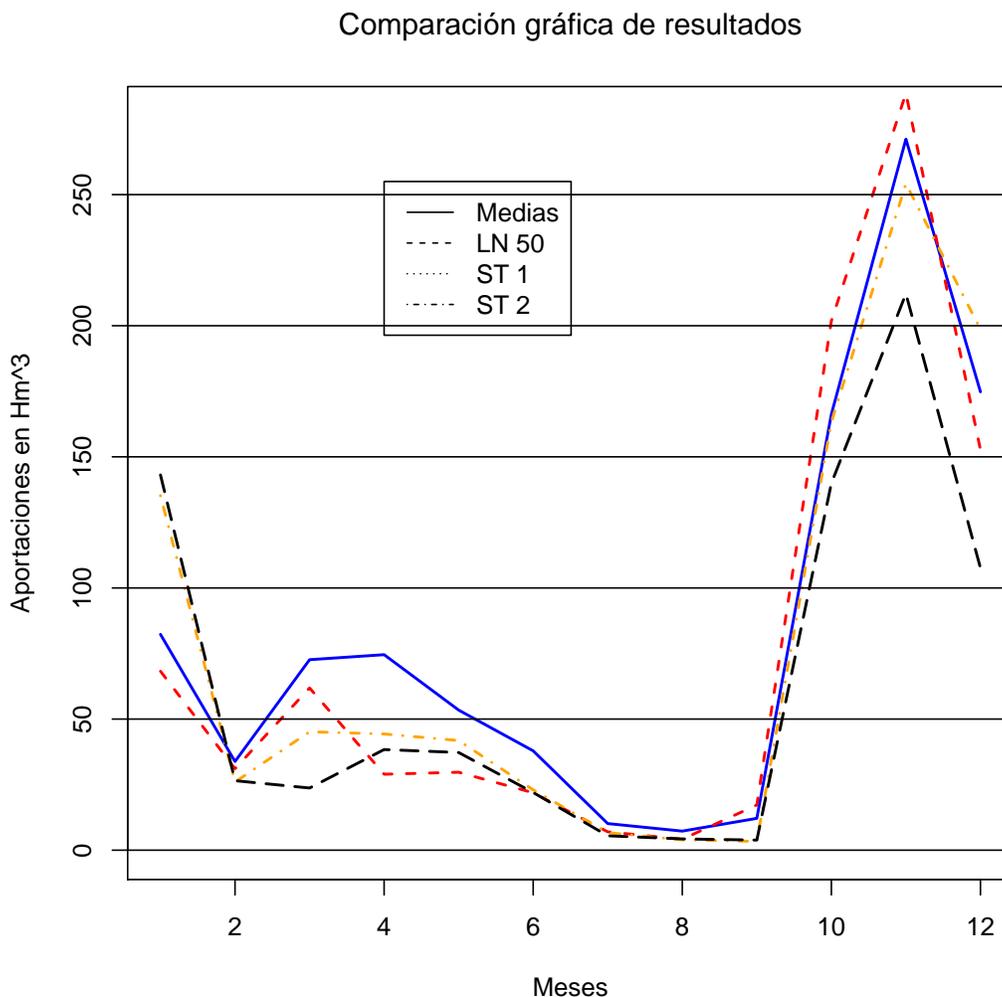


Figura 5.6. comparación gráfica entre medias, Log Normal y Valores de pronóstico.

Como puede observarse en la Figura 5.6, los valores pronosticados con el modelo de series de tiempo para el mes de octubre en los dos ciclos agrícolas, son relativamente altos en comparación con los valores que se obtienen con la Función de Distribución de Probabilidad Log Normal de 2 parámetros. Sin embargo estos valores pronosticados no son mayores a los observados en el registro histórico, es decir, en el Cuadro del Anexo B para el mes de octubre existen valores superiores a los pronosticados con el modelo de series de tiempo.

Capítulo 6

CONCLUSIONES Y RECOMENDACIONES

La aplicación de la metodología probada en este trabajo, permitió obtener las conclusiones y recomendaciones siguientes:

6.1. CONCLUSIONES

La estimación de las aportaciones que se espera escurran hacia una presa de almacenamiento son de gran interés para una mejor planeación de la operación de los sistemas de irrigación, el análisis de series de tiempo realizado en este trabajo da una alternativa más para la estimación de las mismas.

En el estudio de caso analizado, la clase de modelos estacionales de series de tiempo conocidos como modelos $SARIMA(p, d, q)(P, D, Q)$, propuestos en este trabajo reproducen el comportamiento cíclico de aportaciones deducidas a la presa Adolfo Ruiz Cortines, por lo que sus predicciones son estadísticamente validas.

El modelo seleccionado para la predicción de aportaciones deducidas mensuales a la presa ARC en la etapa de validación fue un $SARIMA(0, 1, 2)(3, 1, 0)$, dicho modelo se comparó con varios modelos más siendo este modelo el de menor varianza y menor Criterio de Información de Akaike (AIC). Por lo que se concluye que las predicciones con este modelo son óptimas ya que tienen el mínimo error cuadrado medio.

Se pronosticaron veinticuatro aportaciones para el ciclo agrícola 2011-2012 y 2012-2013 con este modelo y se observó que estas siguen un comportamiento estacional, siendo los valores de aportaciones más altos para los meses lluviosos y con los niveles mas bajos para

6.2. RECOMENDACIONES

los meses de estiaje, este comportamiento es similar al encontrado en la serie observada.

Se concluye que se pueden aplicar la estimación de aportaciones deducidas mediante el análisis y elaboración de modelos de series de tiempo junto con el análisis de frecuencia usual par tener un elemento de juicio en la toma de decisión sobre que serie de aportaciones esperadas emplear en la planeación de la operación de Distritos de Riego.

6.2. RECOMENDACIONES

La clase de modelos propuesta en este trabajo es rígida, es decir no permite detectar las variaciones con respecto a la media mensual ni tener una varianza heterogénea. en este caso, se puede utilizar una generalización de los modelos ARIMA.

La generalización consiste en incluir funciones lineales de seno y cosenos que dependan del tiempo, con las cuales se intenta captar la variación de la media y desviación estándar mensuales. Un análisis comparativo de este tipo de modelos puede ser una alternativa adecuada para explicar el comportamiento estacional de las aportaciones.

El comportamiento cíclico de aportaciones puede ser explicado usando las medias mensuales. Como recomendación a una posible posible alternativa de análisis es trabajar dichas aportaciones con un modelo multivariado de series de tiempo, incluyendo nuevas variables que permitan explicar las variaciones dentro de los ciclos.

Nomenclatura

SERIES DE TIEMPO:

Z_t , Serie observada en función del tiempo t .

z_1, z_2, \dots , Realizaciones de la serie Z_t .

\bar{Z}_t , Media muestral de la serie Z_t .

\tilde{Z}_t , Desviaciones de la serie Z_t respecto a la media \bar{Z}_t

$B\tilde{Z}_t$, Serie retrasada \tilde{Z}_{t-1}

$(1 - B)\tilde{Z}_t$, Primera diferencia $\nabla Z_t = \tilde{Z}_t - \tilde{Z}_{t-1}$

$(1 - B)^d \tilde{Z}_t$, Diferencia de \tilde{Z}_t a distancia d

MODELOS BOX-JENKINS:

$AR(p)$, Modelo Autorregresivo de orden p

$MA(q)$, Modelo de Promedios Móviles de orden q

$ARMA(p,q)$, Modelo Autorregresivo de Promedios Móviles de orden p, q

$ARIMA(p,d,q)$, Modelo Autorregresivo Integrado de Promedios Móviles de orden p, d, q .

MODELOS ESTACIONALES:

Los modelos Box-Jenkins estacionales de series de tiempo se componen de dos partes la No estacional y la Estacional, se denotan como modelos SARIMA (*Seasonal Autoregressive Integrated Moving Avarange*, por sus siglas en Inglés).

$AR(P)$, Parte Autorregresiva estacional de orden P .

$MA(Q)$, Parte Promedios Móviles estacional de orden Q .

D , Diferencia estacional D , $(1 - B^s)\tilde{Z}_t$ a distancia s

$SARIMA(p, d, q)(P, D, Q)_s$, Modelo Autorregresivo ($AR(p)$ y $AR(P)$) Integrado ($d = 1$ y $D = s$) de Promedios Móviles ($MA(q)$ y $MA(Q)$) Estacional s .

Referencias

- Ariza, H. F. J. (2010). *Modelo de Espacio de Estados con Observaciones Censuradas*. Colegio de Postgraduados, Tesis Doctoral. Posgrado en Socioeconomía, Estadística e Informática estadística, Montecillos México.
- Ariza, H. I. (2008). *Estimación Estadística de Aportaciones Deducidas a Presas de Distritos de Riego*. UACH, Tesis Profesional. Departamento de Irrigación, Chapingo, México.
- Box, G. E. P. and Jenkins, G. M. (1970). *Time Series analysis, forecasting and control*. Holden-Day, U.S.A.
- Brockwell P. J. and Davis R. A. (2002). *Introduction to Time Series and forecasting*. Springer, 2nd ed. USA.
- Broomhead, R. J. and King, G.P. (1987). *Topological Dimension and Local Coordinates from Time Series Data*. J. Phys. A 20 L 563 - 569.
- Casella, G. and Berger, R. L. (2002). *Statistical Inference*. Duxbury. Thomson Learning, 2nd ed. USA.
- Cervantes, M. E.(1982). *Uso de modelos estocásticos en el análisis de excedencias fluviales*. Boletín técnico No. 31. Universidad Autónoma Chapingo. México.
- Chavasse, D. I. y Santiago, R. S (1997). *Asociación Determinístico-Estocástica para la Predicción de Caudales*. Ingeniería del Agua. Vol.4 Num.2 (junio 1997), páginas 55-64.
- Chávez, M. Jesús (2006). *11. Apuntes de Hidrología Superficial*. Comunicación personal.
- Chow, Ven Te. (1964). *Handbook of Applied Hydrology; A Compendium of Water Resource Technology*. Mc Graw Hill Book Company, USA.
- Comisión Nacional del Agua (2011). *Subdirección de Infraestructura Hidroagrícola. Gerencia de Distritos de Riego. Subgerencia de operación*. Comunicación personal.
- Guillen, G. J. (1990). *Selección de Funciones de Distribución de Probabilidad para Estimación de Aportaciones por Esguerramientos*. UACH, Tesis Profesional. Departamento de Irrigación, Chapingo, México.

Referencias

- Instituto Mexicano de Tecnología del Agua (2006). *11. Atlas de Presas de Almacenamiento de agua de México*. IMTA, México.
- Kitagawa Genshiro (2010). *Introduction to Time Series Modeling*. CRC Press, Tokyo, Japan.
- Leon, A. G. (2010). *Modelos Univariados en las Series de Temperatura del Aire*. Meteorología Colombiana, No. 2, pp. 43–46. Octubre, 2000. Bogotá D.C. ISSN-0124-6984.
- Mohammadi1, K., Eslami, H. R. and Dardashti, Sh. D. (2005). *Comparison of Regression, ARIMA and ANN Models for Reservoir Inflow Forecasting using Snowmelt Equivalent (a Case study of Karaj)*. J. Agric. Sci. Technol. (2005) Vol. 7: 17-30.
- Pekárová, P, Onderka, M. and Mikláněkand, P. (2009). Prediction of Water Quality in the Danube River under Extreme Hydrological and Temperature Conditions. J. Hydrol. Hydromech., 57 (2009), 1, 3 15.
- Pankratz, Alan. (1983). *Forecasting whith Univariate Box-Jenkins Models; Concepts and Cases*. Jonh Wiley and Sons, USA.
- Quijano, L. José (1977). *Predicción de Lluvia Mediante el Análisis de Series de Tiempo*. Colegio de Postgraduados, Tesis de Maestría. Posgrado en Socioeconomía, Estadística e Informatica estadística, Montecillos México.
- Shumway, R. H. and Stoffer, D. S. (2006). *Time Series Analysis and Its Applications, With R Examples*. Springer, 2nd ed. USA.

Anexos

Anexo A: Estimación de Aportaciones Deducidas en distritos de Riego

En los Distritos de Riego se utiliza un formato de cálculo para estimar las aportaciones deducidas, Figura A.1. La información necesaria para realizar este cálculo se agrupa y se clasifica en cinco apartados: a) Fecha, b) Meteorológica, c) Vaso, d) Entradas y e) Salidas.

Esta información se registra y se concentra a diario para realizar el cálculo de la aportación diaria a partir de la lectura de la escala del nivel de agua en el vaso. La aportación deducida mensual se obtiene sumando las aportaciones deducidas diarias.

El cálculo diario, se realiza siguiendo el proceso que se ilustra en la Figura A.1, misma que corresponde al mes de junio del año 1984 para la presa de almacenamiento Adolfo Ruiz Cortines ubicada en el estado de Sonora. El cálculo se realiza por columnas C_i , como se explica a continuación:

Se adopta un coeficiente de evaporación CE_v , necesario para modificar la evaporación observada en la estación meteorológica y estimar la evaporación en el vaso de almacenamiento de la presa.

FECHA i , es el día en que se realiza el cálculo de la aportación deducida.

C_1 . i = Día del mes j del año k .

SECRETARIA DE AGRICULTURA Y RECURSOS HIDRAULICOS
DIRECCION GENERAL DE HIDROLOGIA
DEPTO. DE HIDROMETRIA

CALCULO HIDROMETRICO
REGISTRO DE REGIMEN DE ALMACENAMIENTO
EN VASO PRESA "ADOLFO RUIZ COZZINES"

FEY. CUA PRASE: 135,000 M.
CAPACIDAD MAXIMA: 1015 000 000 M. CUB.
LONG. CRETA VERTIBRA: 330 000 M.
COEFICIENTE DE EMPERACION: 0.77 m.

MUS DE JUNIO DE 1964
COMIENZO R I O M A Y O
QUENCA R I O M A Y O

DIA	METEOROLOGIA		V A S O		E N T R A D A S		S A L I D A S		A P O R T A C I O N	
	PLU. (mm)	W.M. (mm)	Alt. (m)	Area (ha)	Vol. (m ³)	RECURSOS	DEBIDA			
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										
29										
30										
31										
Suma										
Prom.										

DATOS DE CAMPO
Observador de Meteorología: Edúlio Becerra Rosas.
Observador de Tránsito: Poo. Casco Miras.
Ahoradores: Salvador Becerra Ochoa.
Principió el derrame por el vertedero: *11111.
Terminó el derrame por el vertedero: *11111.

DATOS DE EMPLEADOS EN EL CALCULO
Curso de Areas y Capacidades: Para el No. Según Abans Ním. al No.
Curso de Cálculo Ním. De la Nota Ním. 2
Curso de Cálculo Ním. De el Decimo 2
Curso de Cálculo Ním. De el Decimo

DATOS DE GABINETE
Calculo: Perdomo Valencuela M.
Revisó: Rafael Leykolt N.
Fecha: *11111.
Aprobó: Jefe de Cálculo Hidrométrico
Fecha: *11111.
Director Gen. de Hidrología

Figura A.1. Formato de cálculo de aportaciones deducidas

INFORMACIÓN METEOROLÓGICA, corresponden a los valores observados de precipitación Pr_{ijk} y evaporación Ev_{ijk} , en mm, observados en la estación metereológica de la presa.

C_2 . Pr_{ijk} = Precipitación en el vaso en el día i del mes j del año k .

C_3 . Ev_{ijk} = Evaporación en el vaso en el día i del mes j del año k .

OBSERVACIONES EN EL VASO, el nivel de agua L_{ijk} , Área de embalse A_{ijk} , almacenamiento en el vaso S_{ijk} y la variación del almacenamiento ΔS_{ijk} , los primeros tres valores se determinan con la grafica Elevaciones-Áreas-Capacidades y la variación de almacenamiento es calculada con la ecuación [A.1](#).

$$\Delta S_{ijk} = S_{(i+1)jk} - S_{ijk} \quad (\text{A.1})$$

C_4 . L_{ijk} = Lectura de la escala del nivel de agua en el vaso en el día i del mes j del año k .

C_5 . A_{ijk} = Área de embalse del vaso en el día i del mes j del año k .

C_6 . S_{ijk} = Almacenamiento en el vaso en el día i del mes j del año k .

C_7 . ΔS_{ijk} = Variación de almacenamiento en el vaso, en el día $(i + 1)jk$ y ijk .

ENTRADAS AL VASO, son las entradas de agua al vaso definidas por:

C_8 . VPr_{ijk} = volumen de entrada al vaso por precipitación en el día i del mes j del año k . Que es igual a la ecuación [A.2](#):

$$VPr_{ijk} = A_{ijk}xPr_{ijk} \quad (\text{A.2})$$

C_9 . VE_{ijk} = Entradas por escurrimiento en ríos en el día i del mes j del año k .

C_{10} . VM_{ijk} = Entrada por escurrimientos locales y manantiales, día i del mes j del año k .

C_{11} . ET_{ijk} = Entradas totales en el día i del mes j del año k , dado por la ecuación [A.3](#):

$$ET_{ijk} = VPr_{ijk} + VE_{ijk} + VM_{ijk} \quad (\text{A.3})$$

SALIDAS DEL VASO, son las salidas de agua en el vaso de almacenamiento, ya sea de manera controlada o de forma natural, las variables son:

C_{12} . $ST1_{ijk}$ = Volumen de extracción de la toma 1, en el día i del mes j del año k .

C_{13} . $ST2_{ijk}$ = Volumen de extracción de la toma 2, en el día i del mes j del año k .

C_{14} . SD_{ijk} = Volumen de salidas por desfogue, en el día i del mes j del año k .

C_{15} . SF_{ijk} = Volumen de salida por filtraciones, en el día i del mes j del año k .

C_{16} . $SDER_{ijk}$ = Volumen de salidas por derrame, en el día i del mes j del año k .

C_{17} . VE_{ijk} = Volumen de salida por evaporación en el vaso, en el día i del mes j del año k , dado por la ecuación A.4:

$$VEv_{ijk} = CEv_{ijk} + Ev_{ijk} + A_{ijk} \quad (\text{A.4})$$

C_{18} . ST_{ijk} = Volumen de extracción total, en el día i del mes j del año k . Se calcula con la siguiente ecuación A.5:

$$ST_{ijk} = ST1_{ijk} + ST2_{ijk} + SD_{ijk} + SF_{ijk} + SDER_{ijk} + VEv_{ijk} \quad (\text{A.5})$$

APORTACIÓN DEDUCIDA DIARIA. Finalmente la aportación deducida diaria, se calcula con la siguiente ecuación A.6:

$$APD_{ijk} = \Delta S_{ijk} - ET_{ijk} + ST_{ijk} \quad (\text{A.6})$$

En el formato de cálculo de aportaciones de la figura A.1. La aportación deducida diaria se calcula de la siguiente forma: $C_{19} = C_7 - C_{11} + C_{18}$.

APORTACIÓN DEDUCIDA MENSUAL. Conocida la aportación deducida diaria se calcula la aportación deducida mensual con la ecuación A.7:

$$APD_{jk} = \sum_{i=1}^{NDM_j} APD_{ijk} \quad (\text{A.7})$$

Donde:

APD_{jk} = Aportación deducida del mes j , del año k .

APD_{ijk} = Aportación deducida en el día i del mes j del año k .

NDM_j = Numero de días del mes j .

Ejemplo de cálculo

En el formato de cálculo de aportaciones de la figura A.1. La aportación deducida diaria se calcula de la siguiente forma: $C_{19} = C_7 - C_{11} + C_{18}$.

Coefficiente de evaporación, $CE_v = 0.77m$

C_1 . $i = 1$

INFORMACIÓN METEOROLÓGICA:

C_2 . $Pr_{ijk} = 0mm$.

C_3 . $Ev_{ijk} = 12.20mm$.

INFORMACIÓN DEL VASO:

De acuerdo a la lectura de la escala ubicada en el vaso de almacenamiento se obtiene el nivel de embalse:

C_4 . $L_{ijk} = 125.04m$.

Con esta elevación se entra a la gráfica de Elevaciones-Áreas y se determina el área de embalse:

C_5 . $A_{ijk} = 3948ha$.

Con el área de embalse se entra a la grafica Área-Capacidades para conocer el volumen almacenado:

C_6 . $S_{ijk} = 495,714.0$ Miles de m^3 para $i = 1$ de junio y $S_{(i+1)jk} = 492,642.0$ Miles de m^3 para $i = 2$ de junio de 1984.

La variación del almacenamiento es:

Anexos

$$C_7. \Delta S_{ijk} = -3,072.0 \text{ Miles de } m^3.$$

Si el valor es negativo indica que disminuyó, si es positivo que el volumen almacenado aumento.

ENTRADAS DEL VASO:

$$C_8. VPr_{ijk} = 0.0m^3$$

$$C_9. VE_{ijk} = 0.0m^3$$

$$C_{10}. VM_{ijk} = 0.0m^3$$

Las entradas totales:

$$C_{11}. ET_{ijk} = 0.0m^3$$

SALIDAS DE AGUA DEL VASO:

$$C_{12}. ST1_{ijk} = 1209.6 \text{ Miles de } m^3.$$

$$C_{13}. ST2_{ijk} = 2160.0 \text{ Miles de } m^3.$$

$$C_{14}. SD_{ijk} = 0.0 \text{ Miles de } m^3.$$

$$C_{15}. SF_{ijk} = 0.0 \text{ Miles de } m^3.$$

$$C_{16}. SDER_{ijk} = 0.0 \text{ Miles de } m^3.$$

$$C_{17}. VE_{ijk} = 370.9 \text{ Miles de } m^3.$$

El volumen de extracción total es:

$$C_{18}. ST_{ijk} = 3740.5 \text{ Miles de } m^3.$$

Finalmente la aportación deducida diaria:

$$APD_{ijk} = -3072.0 - 0.0 + 3740.5$$

$$C_{19}. APD_{ijk} = +668.5 \text{ Miles de } m^3.$$

Conocida la aportación deducida diaria se calcula la aportación deducida mensual, con la ecuación [A.7](#) se tiene:

$$APD_{jk} = \sum_{i=1}^{NDM_j} APD_{ijk} = 60,178.0m^3$$

Anexo B: Registro Histórico: Presa Adolfo Ruiz Cortines

Cuadro .1. Registro Histórico de Aportaciones Deducidas a la Presa Adolfo Ruiz Cortines

Año	Oct	Nov	Dic	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Total
1955-1956	146.400	1.800	6.500	11.400	4.100	10.700	0.000	6.700	46.000	146.400	149.200	51.300	580.500
1956-1957	12.600	7.300	7.300	84.800	35.600	11.100	7.400	1.800	0.200	114.100	179.900	66.800	528.900
1957-1958	160.600	12.100	14.000	14.500	32.100	148.200	14.100	20.600	40.000	126.800	232.800	155.700	971.500
1958-1959	61.000	0.000	7.300	8.800	52.600	12.600	4.600	5.500	19.500	199.000	529.100	79.400	979.400
1959-1960	203.300	28.600	90.400	607.700	37.600	24.200	10.600	9.100	2.000	124.100	235.600	162.900	1536.100
1960-1961	33.400	15.400	35.600	259.800	45.500	20.600	12.800	9.900	42.200	186.300	270.800	174.700	1107.000
1961-1962	125.100	105.700	31.200	101.000	29.200	16.300	16.600	17.700	17.600	118.300	75.800	191.400	845.900
1962-1963	37.500	20.400	26.200	18.400	22.400	21.200	9.600	6.100	0.200	180.200	281.600	94.800	718.600
1963-1964	29.900	8.800	219.000	20.300	14.900	22.200	17.000	9.300	2.300	151.900	361.100	346.200	1202.900
1964-1965	24.600	13.300	21.900	16.900	78.300	28.400	12.100	7.900	5.700	81.200	143.800	213.300	647.400
1965-1966	16.700	2.900	149.600	17.200	69.100	12.400	5.900	12.100	15.800	182.100	639.100	123.400	1246.300
1966-1967	16.600	13.200	6.500	2.400	7.200	13.900	7.900	6.200	35.800	204.300	229.900	150.900	694.800
1967-1968	8.900	65.300	268.600	47.500	128.400	192.300	33.500	24.300	8.300	241.900	282.500	134.000	1435.500
1968-1969	39.400	20.300	14.500	19.700	37.000	9.200	4.600	5.600	12.400	214.200	107.900	87.800	572.600
1969-1970	5.200	16.100	49.500	16.900	7.700	18.800	3.300	4.300	2.400	149.000	289.900	132.100	695.200
1970-1971	14.000	3.000	17.400	11.200	4.100	12.600	3.500	1.800	33.900	236.700	389.400	100.200	827.800
1971-1972	530.300	112.000	33.400	28.500	19.800	15.400	3.100	13.100	30.300	106.100	209.300	130.900	1232.200
1972-1973	193.800	69.700	22.700	176.000	165.500	46.700	20.600	14.000	7.300	94.200	386.400	133.100	1330.000
1973-1974	9.600	0.000	0.400	20.800	10.200	5.200	0.700	0.000	0.000	262.400	143.400	221.700	674.400
1974-1975	35.900	133.400	14.800	23.400	25.500	7.400	3.600	1.700	0.000	125.800	251.500	203.200	826.200
1975-1976	13.300	13.700	10.200	7.200	14.000	0.000	0.000	0.000	6.400	165.700	110.000	149.400	489.900
1976-1977	89.500	14.400	8.500	19.400	9.800	0.000	0.200	0.000	0.000	167.400	336.900	143.400	796.100
1977-1978	18.000	13.300	12.300	12.800	17.400	53.200	4.800	0.000	3.200	95.900	257.300	207.100	695.300
1978-1979	237.400	25.500	70.600	284.600	174.100	33.000	20.200	12.900	8.300	66.200	147.000	45.600	1125.400
1979-1980	10.600	8.600	11.200	9.500	1.700	0.000	0.000	0.000	0.000	102.000	392.500	179.500	715.600
1980-1981	31.600	9.100	14.300	284.600	31.300	170.100	12.000	8.800	17.900	245.800	259.000	299.000	1383.500
1981-1982	91.400	27.200	8.100	18.600	4.100	23.000	7.900	5.600	1.100	126.600	142.300	97.600	553.500
1982-1983	34.100	21.000	233.600	104.900	201.300	386.000	98.700	25.700	24.700	75.000	310.000	262.500	1777.500
1983-1984	301.000	110.000	92.400	31.000	35.000	28.000	8.200	15.000	60.200	391.000	486.500	144.300	1702.600

Cuadro .2. Continuación: Registro Histórico de Aportaciones Deducidas a la Presa Adolfo Ruiz Cortines

Año	Oct	Nov	Dic	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Total
1984-1985	119.000	33.500	551.600	322.200	117.000	62.100	34.800	23.000	26.000	152.300	380.500	98.100	1920.100
1985-1986	29.000	22.600	11.700	19.300	19.500	4.700	0.000	1.600	15.400	367.700	292.000	254.500	1038.000
1986-1987	74.400	23.600	49.500	18.700	19.100	29.400	8.600	5.700	3.900	56.900	304.900	27.400	622.100
1987-1988	12.100	0.000	27.000	5.500	0.000	16.600	0.000	1.200	1.300	370.100	504.500	94.700	1033.000
1988-1989	16.100	7.300	14.100	35.800	103.000	37.700	7.400	2.400	0.000	47.200	362.900	164.100	798.000
1989-1990	9.400	13.200	18.500	27.100	75.400	40.100	0.000	0.000	5.900	399.600	303.500	190.100	1082.800
1990-1991	215.300	27.100	618.200	140.300	53.600	83.200	25.600	20.300	6.800	229.000	491.700	548.800	2459.900
1991-1992	49.600	37.000	307.500	521.300	210.200	85.100	36.100	33.100	22.600	81.900	204.900	162.800	1752.100
1992-1993	17.000	11.900	24.700	146.100	135.600	88.300	10.800	11.900	7.300	218.800	215.500	223.200	1111.100
1993-1994	39.800	51.200	54.900	18.600	28.200	14.700	11.000	3.200	16.000	150.900	193.300	201.200	783.000
1994-1995	22.300	184.000	243.800	50.500	156.200	30.900	10.300	2.700	7.900	105.600	211.200	621.200	1646.600
1995-1996	35.800	32.400	31.800	18.000	14.100	20.700	9.000	8.200	6.000	94.300	365.900	339.600	975.800
1996-1997	27.100	24.400	7.400	5.500	16.400	21.700	8.100	7.600	5.900	61.200	214.000	135.100	534.400
1997-1998	33.500	29.800	203.600	43.600	33.600	35.400	5.400	0.000	0.000	85.900	228.200	203.700	902.700
1998-1999	26.900	21.600	11.700	4.000	0.000	0.000	0.000	1.100	7.100	327.200	178.900	87.600	666.100
1999-2000	12.100	20.900	7.077	4.100	0.000	0.000	1.900	2.200	40.300	170.800	151.500	38.000	448.877
2000-2001	447.100	150.700	26.200	18.500	64.400	29.900	6.300	0.000	12.800	192.700	160.543	60.800	1169.943
2001-2002	73.700	16.600	10.200	7.500	20.500	0.000	0.000	0.000	0.000	153.500	211.200	45.500	538.700
2002-2003	6.700	6.500	15.100	9.971	44.845	0.000	2.000	1.800	1.234	117.544	222.142	79.418	507.254
2003-2004	34.201	15.730	6.537	24.908	15.234	20.937	6.073	4.693	2.870	177.906	237.209	115.200	661.498
2004-2005	105.741	130.870	121.133	191.794	343.372	36.331	9.367	11.183	6.906	63.582	286.487	127.858	1434.624
2005-2006	20.361	20.219	16.671	15.912	18.928	18.022	3.534	2.086	5.535	138.953	282.816	208.950	751.987
2006-2007	18.444	19.620	11.200	83.097	33.017	13.431	3.357	2.321	14.232	228.687	349.453	497.379	1274.238
2007-2008	24.098	20.145	137.839	32.685	26.883	18.763	10.602	3.441	3.079	259.489	410.175	542.004	1489.203
2008-2009	65.253	24.772	12.562	20.212	20.957	20.805	5.641	6.358	6.555	101.489	167.434	40.723	492.761
2009-2010	521.060	31.991	17.938	92.686	88.025	37.190	8.301	4.157	2.761	121.474	228.526	158.108	1312.217
2010-2011	20.972	23.822	11.532	17.607	18.208	12.272	0.804	1.196	0.844	155.002	195.171	39.809	497.239
Media	82.299	33.814	72.643	74.558	53.424	37.874	10.151	7.270	12.134	166.256	271.162	174.787	996.372
Desv. Est.	120.073	40.997	124.968	123.247	66.024	61.415	14.751	7.607	13.972	85.404	115.682	128.630	436.451
Max	530.300	184.000	618.200	607.700	343.372	386.000	98.700	33.100	60.200	399.600	639.100	621.200	2459.900
Min	5.200	0.000	0.400	2.400	0.000	0.000	0.000	0.000	0.000	47.200	75.800	27.400	448.877

Anexo C: Rutinas desarrolladas en R

Se desarrollaron algunas rutinas, se usaron funciones y Paquetes programadas propias de R para el análisis de la serie de aportaciones a la Presa Adolfo Ruiz Cortines, el script de R es el siguiente:

```
#-----#
# C O L E G I O   D E   P O S T G R A D U A D O S   #
#           Campus Montecillo                       #
#           PROGRAMA DE HIDROCIENCIAS               #
#"Estimación de Aportaciones Deducidas mediante el #
#           Análisis de Series de Tiempo"           #
#           Tesis de Maestría                       #
#           Programó: Isaías Ariza Hdez.            #
#-----#

#----- ANALISIS DE DATOS. PARTE 1. MODELO BOX AND JENKINS-----

#--Importar la matriz de datos de aportaciones
datos2<-c()
datos2<-read.csv("/home/isariza/Documentos/TesisColposIAH/
                ANALISISAPRESAS/st_matriz.csv",header=TRUE)
datos3<-as.matrix(datos2)
datos3
class(datos3)
n<-nrow(datos3)

st_mensual<-c()
for(i in 1:n)
{
  st_mensual<-c(st_mensual, datos3[i,])
}
st_mensual
length(st_mensual)
class(st_mensual)

#-----Desviaciones respecto a la media

Z<-st_mensual-mean(st_mensual)

#-----Covierte los datos en un Objeto tipo 'time series'
X<-as.ts(st_mensual)
length(X)
plot(X,main="Aportaciones Mensuales Totales a la Presa ARC (1955-2011)",
      xlab="",ylab="Aportaciones",col="blue")
x11()
```

Anexos

```
Zt<-as.ts(Z)
plot(Zt,main="Desviaciones respecto a la media muestral",
      xlab="Meses",ylab="Aportaciones",col="blue")
x11()

#-----Hidrograma de aportaciones  totales mensuales
n<-length(X)
U<-rep(mean(st_mensual),n)
plot(X,type="l",xlab="Meses",ylab="Aportaciones en Hm^3",col="blue",)
lines(U,lty=2,col="Red")
legend(355,650,c("Hidrograma de volumenes", "Media mensual"),lty=1:2)
title(main = list("HIDROGRAMA DE APORTACIONES MENSUALES 1955-2011", cex=1.2,
                  col="black", font=2))

#-----Funcion de Autocorrelacion para la serie Zt
fac<-acf(Zt,ci.type="ma", main="Función de Autocorrelación de Zt", col="red")
fac
x11()

#-----Funcion de autocorrelacion parcial
facp<-pacf(Zt,main="Función de Autocorrelación Parcial de Zt",col="red")
facp

#-----Para visualizar la estacionalidad de la serie
S<-Z[612:672]
St<-as.ts(S)
St
plot(St,main="Aportaciones mensuales (2001-2011) a la presa ARC",xlab="Meses",
      ylab="Aportaciones en Hm^3",col="blue")
x11()

#-----Funcion que realiza diferencias en una serie

dif.serie<-function(datos,D) #datos:seria a diferenciar y D=distancia
{
  k<-D+1
  n<-length(datos)
  z.1<-rep(0,n-D)
  for(i in k:n)
  {
    z.1[i-D]<-datos[i]-datos[i-D]
  }
  z.1<-as.ts(z.1)
  z.1
}

#-----Primera diferencia de la serie a distancia d=1,
```

```
d<-1
Z1<-dif.serie(Zt,d)
Z1
x11()
#-----Graficamos la nueva serie diferenciada
plot(Z1,main="Zt diferenciada a d=1", xlab="Meses", ylab="Aportaciones",
      col="blue")

#-----FAC y FACP de Z1
fac1<-acf(Z1,ci.type="ma", main="Función de Autocorrelación Estimada de
      Zt con d=1", col="red", lag.max=c(40))
fac1
x11()
facp1<-pacf(Z1,main="Función de Autocorrelación Parcial Estimada de
      Zt con d=1", col="red", lag.max=c(40))
facp1

#-----Segunda diferencia estacional s=12
D<-12
Z2<-dif.serie(Z1,D)
Z2
plot(Z2,main="Wt diferenciada a s=12",xlab="Meses",ylab="Aportaciones",
      col="blue")

#-----FAC y FACP de Z1
x11()
fac2<-acf(Z2,ci.type="ma", lag.max=c(40), main="Función de Autocorrelación
      Estimada de Wt con s=12", col="red")
fac2
x11()
facp2<-pacf(Z2,main="Función de Autocorrelación Parcial Estimada de
      Wt con s=12", col="red",lag.max=c(40))
facp2

#-----ETAPA 2: ESTIMACION DE LOS PARAMETROS DEL MODELO
#Entonces dadas la fac.2 y facp.2, se propone el modelo los siguientes modelos:

M1<-arima(Zt, order=c(1,1,1), seasonal=list(order=c(2,1,1),period=12),method="ML")
M1

M2<-arima(Zt, order=c(1,1,0), seasonal=list(order=c(2,1,0),period=12),method="ML")
M2

M3<-arima(Zt, order=c(0,1,1), seasonal=list(order=c(2,1,0),period=12),method="ML")
M3
```

#ETAPA 3. VALIDACION DEL MODELO SARIMA(1,1,0)(2,1,0)_12

#3.1. ---Significancia de los estimadores de los parámetros

```
varianza<-M32$var.coef
varianza
#var<-c( 1.235183e-03,1.405634e-03,0.0014771810)
#var2<-var^(1/2)
#tc<-abs(p/var2)
#tc
#[1] 12.331709 16.979719 6.881907
```

#3.2. --- Aumentamos parámetros

```
M4<-arima(Zt, order=c(0,2,2), seasonal=list(order=c(3,1,0),period=12),method="ML")
M4
varianza2<-M4$var.coef
varianza2
```

#3.----- Verificación de los residuales no correlacionados

```
n<-length(Zt)
Residuales<-M4$residuals
Residuales
res<-acf(Residuales,ci.type="ma", lag.max=c(36),
        main="Función de Autocorrelación de Residuales", col="red")
res
#acf de Residuales obtenidos
b7<-c(-0.065,-0.140,-0.225,-0.015,0.056,-0.046,-0.059,0.035,0.034,-0.030,-0.084,-0.071,
      0.054,0.018,0.035,0.024,0.020,0.042,-0.032,-0.066)
```

#---Prueba de Ljung-Box

```
b7<-res$acf[1:21]
b8<-b7[2:21]
c1<-(b8)^2
N_k<-c((n-1):(n-20))
c2<-c1/N_k
Q1<-sum(c2)
Q2<-{(n*(n+2))}*Q1
Q2
```

```
tsdiag(M4, gof.lag=12)
final<-Box.test(b8,lag=12,type=c("Ljung"))
final
```

#-----4.Prediccion con el modelo

```
pronos<-predict(M4,n.ahead=12)
LS<-c(pronos$pred+2*pronos$se)
LI<-c(pronos$pred-2*pronos$se)

#-Sumando la media extradia
media<-mean(X)
pronostico<-c(pronos$pred+media)
LS2<-c(pronostico+2*pronos$se)
LI2<-c(pronostico-2*pronos$se)
pronostico
LS2
LI2

S2<-c(S+mean(X))
S2
plot(S2,type="l",xlab="Meses",ylab="Aportaciones",col="blue",
      xlim=c(612,684),ylim=c(0,600))
lines(pronostico,lty=2,col="Red")
legend(451,567,c("Serie Observada", "Pronosticos"),lty=1:2)
title(main = list("Pronosticos con un SARIMA(0,1,2)(3,1,0)", cex=1.2,
                  col="black", font=2))

#-----Ajuste

Residuales2<-c(Residuales+mean(X))
Residuales2

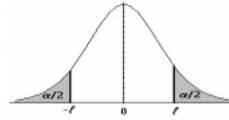
plot(X,type="l",xlab="Meses",ylab="Aportaciones en Hm^3",col="blue",)
lines(X-Residuales,lty=2,col="Red")
legend(451,567,c("Serie Observada", "Serie Estimada"),lty=1:2)
title(main = list("Ajuste de la serie observada con un SARIMA(0,1,2)(3,1,0)", cex=1.2,
                  col="black", font=1.8))
```

Anexo D: Tablas utilizadas: *t* – student y *Chi* – cuadrada

La prueba de *t* – student Cuadro D.1, se utilizó para valorar que los coeficientes estimados fueran estadísticamente diferentes de cero.

Tabla de la *t* de Student.

Contiene los valores *t* tales que $p[|T| > t] = \alpha$,
donde *n* son los grados de libertad.



<i>n</i> \ α	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,1584	0,3249	0,5095	1,0000	1,9626	3,0777	6,3137	12,7062	31,8210	63,6559	636,5776
2	0,1421	0,2887	0,4447	0,8165	1,3862	1,8856	2,9200	4,3027	6,9645	9,9250	31,5998
3	0,1366	0,2767	0,4242	0,7649	1,2498	1,6377	2,3534	3,1824	4,5407	5,8408	12,9244
4	0,1338	0,2707	0,4142	0,7407	1,1896	1,5332	2,1318	2,7765	3,7469	4,6041	8,6101
5	0,1322	0,2672	0,4082	0,7267	1,1558	1,4759	2,0150	2,5706	3,3649	4,0321	6,8685
6	0,1311	0,2648	0,4043	0,7176	1,1341	1,4398	1,9432	2,4469	3,1427	3,7074	5,9587
7	0,1303	0,2632	0,4015	0,7111	1,1192	1,4149	1,8946	2,3646	2,9979	3,4995	5,4081
8	0,1297	0,2619	0,3995	0,7064	1,1081	1,3968	1,8595	2,3060	2,8965	3,3554	5,0414
9	0,1293	0,2610	0,3979	0,7027	1,0997	1,3830	1,8331	2,2622	2,8214	3,2498	4,7809
10	0,1289	0,2602	0,3966	0,6998	1,0931	1,3722	1,8125	2,2281	2,7638	3,1693	4,5868
11	0,1286	0,2596	0,3956	0,6974	1,0877	1,3634	1,7959	2,2010	2,7181	3,1058	4,4369
12	0,1283	0,2590	0,3947	0,6955	1,0832	1,3562	1,7823	2,1788	2,6810	3,0545	4,3178
13	0,1281	0,2586	0,3940	0,6938	1,0795	1,3502	1,7709	2,1604	2,6503	3,0123	4,2209
14	0,1280	0,2582	0,3933	0,6924	1,0763	1,3450	1,7613	2,1448	2,6245	2,9768	4,1403
15	0,1278	0,2579	0,3928	0,6912	1,0735	1,3406	1,7531	2,1315	2,6025	2,9467	4,0728
16	0,1277	0,2576	0,3923	0,6901	1,0711	1,3368	1,7459	2,1199	2,5835	2,9208	4,0149
17	0,1276	0,2573	0,3919	0,6892	1,0690	1,3334	1,7396	2,1098	2,5669	2,8982	3,9651
18	0,1274	0,2571	0,3915	0,6884	1,0672	1,3304	1,7341	2,1009	2,5524	2,8784	3,9217
19	0,1274	0,2569	0,3912	0,6876	1,0655	1,3277	1,7291	2,0930	2,5395	2,8609	3,8833
20	0,1273	0,2567	0,3909	0,6870	1,0640	1,3253	1,7247	2,0860	2,5280	2,8453	3,8496
21	0,1272	0,2566	0,3906	0,6864	1,0627	1,3232	1,7207	2,0796	2,5176	2,8314	3,8193
22	0,1271	0,2564	0,3904	0,6858	1,0614	1,3212	1,7171	2,0739	2,5083	2,8188	3,7922
23	0,1271	0,2563	0,3902	0,6853	1,0603	1,3195	1,7139	2,0687	2,4999	2,8073	3,7676
24	0,1270	0,2562	0,3900	0,6848	1,0593	1,3178	1,7109	2,0639	2,4922	2,7970	3,7454
25	0,1269	0,2561	0,3898	0,6844	1,0584	1,3163	1,7081	2,0595	2,4851	2,7874	3,7251
26	0,1269	0,2560	0,3896	0,6840	1,0575	1,3150	1,7056	2,0555	2,4786	2,7787	3,7067
27	0,1268	0,2559	0,3894	0,6837	1,0567	1,3137	1,7033	2,0518	2,4727	2,7707	3,6895
28	0,1268	0,2558	0,3893	0,6834	1,0560	1,3125	1,7011	2,0484	2,4671	2,7633	3,6739
29	0,1268	0,2557	0,3892	0,6830	1,0553	1,3114	1,6991	2,0452	2,4620	2,7564	3,6595
30	0,1267	0,2556	0,3890	0,6828	1,0547	1,3104	1,6973	2,0423	2,4573	2,7500	3,6460
40	0,1265	0,2550	0,3881	0,6807	1,0500	1,3031	1,6839	2,0211	2,4233	2,7045	3,5510
80	0,1261	0,2542	0,3867	0,6776	1,0432	1,2922	1,6641	1,9901	2,3739	2,6387	3,4164
120	0,1259	0,2539	0,3862	0,6765	1,0409	1,2886	1,6576	1,9799	2,3578	2,6174	3,3734
∞	0,126	0,253	0,385	0,674	1,036	1,282	1,645	1,96	2,326	2,576	3,291

Cuadro D.1. Valores de t-Student

Anexos

La prueba *Chi – cuadrada* Cuadro D.2, se utilizó para verificar la no correlación de los residuales.

P = Probabilidad de encontrar un valor mayor o igual que el chi cuadrado tabulado, v = Grados de Libertad

v/p	0,001	0,0025	0,005	0,01	0,025	0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5
1	10,8274	9,1404	7,8794	6,6349	5,0239	3,8415	2,7055	2,0722	1,6424	1,3233	1,0742	0,8735	0,7083	0,5707	0,4549
2	13,8150	11,9827	10,5965	9,2104	7,3778	5,9915	4,6052	3,7942	3,2189	2,7726	2,4079	2,0996	1,8326	1,5970	1,3863
3	16,2660	14,3202	12,8381	11,3449	9,3484	7,8147	6,2514	5,3170	4,6416	4,1083	3,6649	3,2831	2,9462	2,6430	2,3660
4	18,4662	16,4238	14,8602	13,2767	11,1433	9,4877	7,7794	6,7449	5,9886	5,3853	4,8784	4,4377	4,0446	3,6871	3,3567
5	20,5147	18,3854	16,7496	15,0863	12,8325	11,0705	9,2363	8,1152	7,2893	6,6257	6,0644	5,5731	5,1319	4,7278	4,3515
6	22,4575	20,2491	18,5475	16,8119	14,4494	12,5916	10,6446	9,4461	8,5581	7,8408	7,2311	6,6948	6,2108	5,7652	5,3481
7	24,3213	22,0402	20,2777	18,4753	16,0128	14,0671	12,0170	10,7479	9,8032	9,0371	8,3834	7,8061	7,2832	6,8000	6,3458
8	26,1239	23,7742	21,9549	20,0902	17,5345	15,5073	13,3616	12,0271	11,0301	10,2189	9,5245	8,9094	8,3505	7,8325	7,3441
9	27,8767	25,4625	23,5893	21,6660	19,0228	16,9190	14,6837	13,2880	12,2421	11,3887	10,6564	10,0060	9,4136	8,8632	8,3428
10	29,5879	27,1119	25,1881	23,2093	20,4832	18,3070	15,9872	14,5339	13,4420	12,5489	11,7807	11,0971	10,4732	9,8922	9,3418
11	31,2635	28,7291	26,7569	24,7250	21,9200	19,6752	17,2750	15,7671	14,6314	13,7007	12,8987	12,1836	11,5298	10,9199	10,3410
12	32,9092	30,3182	28,2997	26,2170	23,3367	21,0261	18,5493	16,9893	15,8120	14,8454	14,0111	13,2661	12,5838	11,9463	11,3403
13	34,5274	31,8830	29,8193	27,6882	24,7356	22,3620	19,8119	18,2020	16,9848	15,9839	15,1187	14,3451	13,6356	12,9717	12,3398
14	36,1239	33,4262	31,3194	29,1412	26,1189	23,6848	21,0641	19,4062	18,1508	17,1169	16,2221	15,4209	14,6853	13,9961	13,3393
15	37,6978	34,9494	32,8015	30,5780	27,4884	24,9958	22,3071	20,6030	19,3107	18,2451	17,3217	16,4940	15,7332	15,0197	14,3389
16	39,2518	36,4555	34,2671	31,9999	28,8453	26,2962	23,5418	21,7931	20,4651	19,3689	18,4179	17,5646	16,7795	16,0425	15,3385
17	40,7911	37,9462	35,7184	33,4087	30,1910	27,5871	24,7690	22,9770	21,6146	20,4887	19,5110	18,6330	17,8244	17,0646	16,3382
18	42,3119	39,4220	37,1564	34,8052	31,5264	28,8693	25,9894	24,1555	22,7595	21,6049	20,6014	19,6993	18,8679	18,0860	17,3379
19	43,8194	40,8847	38,5821	36,1908	32,8523	30,1435	27,2036	25,3289	23,9004	22,7178	21,6891	20,7638	19,9102	19,1069	18,3376
20	45,3142	42,3358	39,9969	37,5663	34,1696	31,4104	28,4120	26,4976	25,0375	23,8277	22,7745	21,8265	20,9514	20,1272	19,3374
21	46,7963	43,7749	41,4009	38,9322	35,4789	32,6706	29,6151	27,6620	26,1711	24,9348	23,8578	22,8876	21,9915	21,1470	20,3372
22	48,2676	45,2041	42,7957	40,2894	36,7807	33,9245	30,8133	28,8224	27,3015	26,0393	24,9390	23,9473	23,0307	22,1663	21,3370
23	49,7276	46,6231	44,1814	41,6383	38,0756	35,1725	32,0069	29,9792	28,4288	27,1413	26,0184	25,0055	24,0689	23,1852	22,3369
24	51,1790	48,0336	45,5584	42,9798	39,3641	36,4150	33,1962	31,1325	29,5533	28,2412	27,0960	26,0625	25,1064	24,2037	23,3367
25	52,6187	49,4351	46,9280	44,3140	40,6465	37,6525	34,3816	32,2825	30,6752	29,3388	28,1719	27,1183	26,1430	25,2218	24,3366
26	54,0511	50,8291	48,2898	45,6416	41,9231	38,8851	35,5632	33,4295	31,7946	30,4346	29,2463	28,1730	27,1789	26,2395	25,3365
27	55,4751	52,2152	49,6450	46,9628	43,1945	40,1133	36,7412	34,5736	32,9117	31,5284	30,3193	29,2266	28,2141	27,2569	26,3363
28	56,8918	53,5939	50,9936	48,2782	44,4608	41,3372	37,9159	35,7150	34,0266	32,6205	31,3909	30,2791	29,2486	28,2740	27,3362
29	58,3006	54,9662	52,3355	49,5878	45,7223	42,5569	39,0875	36,8538	35,1394	33,7109	32,4612	31,3308	30,2825	29,2908	28,3361

Cuadro D.2. Valores de Chi Cuadrada