

**UNIVERSIDAD NACIONAL DEL ALTIPLANO**  
**ESCUELA DE POSGRADO**  
**DOCTORADO EN ESTADÍSTICA E INFORMÁTICA**



**TESIS**

**MODELO UNIVARIANTE PARA DESCRIBIR Y PREDECIR EL  
COMPORTAMIENTO DE LOS NIVELES MEDIOS DE AGUA DEL LAGO  
TITICACA – PUNO**

**PRESENTADA POR:**

**CONFESOR MILÁN VARGAS VALVERDE**

**PARA OPTAR EL GRADO ACADÉMICO DE:**

**DOCTOR EN ESTADÍSTICA E INFORMÁTICA**

**PUNO, PERÚ**

**2019**

UNIVERSIDAD NACIONAL DEL ALTIPLANO

ESCUELA DE POSGRADO



DOCTORADO EN ESTADÍSTICA E INFORMÁTICA

TESIS

MODELO UNIVARIANTE PARA DESCRIBIR Y PREDECIR EL  
COMPORTAMIENTO DE LOS NIVELES MEDIOS DE AGUA DEL LAGO  
TITICACA - PUNO

PRESENTADA POR:

CONFESOR MILÁN VARGAS VALVERDE

PARA OPTAR EL GRADO ACADÉMICO DE:

DOCTOR EN ESTADÍSTICA E INFORMÁTICA

APROBADA POR EL SIGUIENTE JURADO:

PRESIDENTE



Dr. MARCO ANTONIO QUISPE BARRA

PRIMER MIEMBRO



Dr. YALMAR PONCE ATENCIO

SEGUNDO MIEMBRO



Dr. MAXIMO MONTALVO ATCO

ASESOR DE TESIS



Dr. ALEJANDRO APAZA TARQUI

Puno, 13 de diciembre de 2019

**AREA:** Estadística e Informática.

**TEMA:** Modelo univariante del comportamiento de agua del lago Titicaca.

**LÍNEA:** Modelos predictivos uni y multivariante.

## DEDICATORIA

A la memoria de mis amados padres Vicente Vargas Mecola y Auxilia Valverde Briseño, por su inmenso cariño, sacrificio, comprensión y formación que me dieron durante mi vida.

A mis queridos hermanos Rómulo, Adriano, Mírope, Elodia, Ciria, Nora y sobrinos, quienes me han brindaron su apoyado para plasmar este trabajo de investigación.

A mi esposa Ana, por su apoyo incondicional y comprensión en la culminación de esta investigación.

A mis adoradas hijas Grisel, Eurídice y Nicol, que son la razón de mi vida y estuvieron alentándome en cada momento para la realización de este trabajo.

A mis colegas y amistades, quienes me brindaron su apoyo.

## AGRADECIMIENTOS

- A Dios todopoderoso. Por haberme dado fortaleza y sabiduría, para que fuera dable alcanzar este laurel y guiarme para seguir adelante.
- A la Universidad Nacional Del Altiplano Puno, sobre todo a la escuela profesional de Ingeniería Estadística e Informática, por darme la oportunidad de desarrollarme y permitirme formar parte de ella.
- A los Catedráticos del Doctorado en Estadística e Informática por sus enseñanzas impartidas durante mi formación de posgrado.
- A mis jurados de tesis, Dr. Marco Antonio Quispe Barra, Dr. Yalmar Ponce Atencio, Dr. Máximo Montalvo Atco y Dr. Alejandro Apaza Tarqui, por el apoyo en mi tesis doctoral.

## ÍNDICE GENERAL

	Pág.
DEDICATORIA	I
AGRADECIMIENTOS	II
ÍNDICE GENERAL	III
ÍNDICE DE TABLAS	VI
ÍNDICE DE FIGURAS	VII
ÍNDICE DE ANEXOS	IX
INTRODUCCIÓN	1

### CAPÍTULO I

#### REVISIÓN DE LITERATURA

1.1. Marco teórico	3
1.1.1. Pronóstico	3
1.1.2. La necesidad de pronosticar	3
1.1.3. Técnicas de pronósticos.	4
1.1.4. Pronósticos según plazos.	4
1.1.5. Series de tiempo.	5
1.1.5.1. Componentes de una serie temporal.	5
1.1.5.2. Análisis de series de tiempo.	6
1.1.5.3. Estimación de la tendencia	7
1.1.6. Proceso estocástico.	8
1.1.6.1. Proceso estocástico estacionario.	9
1.1.7. Ruido blanco.	10
1.1.8. Estacionalidad.	11
1.1.9. Coeficiente de correlación.	11
1.1.10. varianza y error estándar de ruido blanco.	11
1.1.11. Función de autocorrelación.	12
1.1.12. Función de autocorrelación parcial.	13
1.1.13. Modelo.	13
1.1.14. Modelos de series temporales.	14
1.1.15. Modelo univariante.	14
1.1.16. Operador de retardo y diferenciación de una serie.	15

1.1.17.	Elaboración de modelos AR, MA, ARMA y ARIMA.	15
1.1.18.	Metodología de Box-Jenkins.	33
1.1.19.	Dickey-Fuller aumentado (Test ADF).	33
1.1.20.	Estadístico Box Pierce.	34
1.2.	Antecedentes.	35

## CAPÍTULO II

### PLANTEAMIENTO DEL PROBLEMA

2.1	Identificación del problema.	42
2.2	Enunciados del problema.	44
2.2.1.	Interrogante del problema general.	44
2.2.2.	Interrogante de los problemas específicos.	44
2.3	Justificación.	44
2.4	Objetivos.	45
2.4.1.	Objetivo General.	45
2.4.2.	Objetivos Específicos.	45
2.5	Hipótesis.	45
2.5.1.	Hipótesis General.	45
2.5.2.	Hipótesis específica.	46

## CAPÍTULO III

### MATERIALES Y MÉTODOS

3.2.	Población.	47
3.3.	Muestra.	47
3.4.	Método de investigación.	47
3.5.	Metodología econométrica.	48

## CAPÍTULO IV

### RESULTADOS Y DISCUSIÓN

4.1.	Identificación del modelo.	52
4.2.	Estimación del modelo identificado.	60
4.3.	Validación o adecuación del modelo estimado.	61
4.4.	Pronóstico con el modelo validado.	64
	CONCLUSIONES	67



RECOMENDACIONES	68
BIBLIOGRAFÍA	69
ANEXOS	72

**ÍNDICE DE TABLAS**

	Pág.
1. Patrones teóricos de FAC y FACP	13
2. Nivel promedio mensual del lago Titicaca en la estación Huatajata por mes según año (Metros Sobre el Nivel del Mar) 1974-2018.	51
3. Resumen del modelo ARIMA (1,1,0)(0,1,1)	60
4. Resumen de los parámetros y estadísticos	60

## ÍNDICE DE FIGURAS

	Pág.
1. Coeficientes de autocorrelación y de autocorrelación parcial de los modelos AR (1) y AR (2)	30
2. Coeficientes de autocorrelación y de autocorrelación parcial de los Modelos MA (1) Y MA (2)	31
3. Coeficientes de autocorrelación y de autocorrelación parcial de un modelo ARIMA (1,1).	32
4. Fases de elaboración de un modelo ARIMA.	49
5. Secuencia de tiempo del nivel promedio mensual del lago Titicaca.	53
6. Autocorrelaciones estimadas del nivel promedio mensual del lago Titicaca.	53
7. Autocorrelaciones parciales estimadas del nivel promedio mensual del lago Titicaca.	54
8. Primera diferencia no estacional del nivel promedio mensual del lago Titicaca.	55
9. Autocorrelaciones estimadas para la primera diferencia no estacional del nivel promedio mensual del lago Titicaca.	55
10. Autocorrelaciones parcial estimada para la primera diferencia no estacional del nivel promedio mensual del lago Titicaca.	56
11. Primera Diferencia Estacional del nivel promedio mensual del lago Titicaca.	56
12. Autocorrelaciones Estimadas para la primera diferencia estacional del nivel promedio mensual del lago Titicaca.	57
13. Autocorrelaciones parciales estimadas para la primera diferencia estacional del nivel promedio mensual del lago Titicaca.	58
14. Primera diferencia no estacional por primera diferencia estacional del nivel promedio mensual del lago Titicaca.	58
15. Autocorrelaciones estimadas para la primera diferencia no estacional por primera diferencial estacional del nivel promedio mensual del lago Titicaca.	59
16. Autocorrelaciones parciales estimadas para la primera diferencia no estacionaria por primera diferencia estacional del nivel promedio mensual del lago Titicaca.	59
17. Probabilidad normal del nivel promedio mensual de agua del lago Titicaca.	62

18. Función de autocorrelación de residuales estimados del nivel promedio mensual de agua del lago Titicaca.	63
19. Función de autocorrelación parcial de residuales estimados del nivel promedio mensual del lago Titicaca.	63
20. Función de pronóstico con límites de confianza del 95% del nivel promedio mensual de agua del lago Titicaca.	64
21. Función pronosticada con límites de confianza del 95% del nivel promedio mensual de agua del lago Titicaca.	65
22. Función pronosticada con límites de confianza al 95% del nivel promedio mensual de agua del lago Titicaca.	65
23. Función pronosticada con límites de confianza al 95% del nivel promedio mensual de agua del lago Titicaca.	66

## ÍNDICE DE ANEXOS

	Pág.
1. Pronósticos de corto plazo para los niveles medios de agua en el lago Titicaca	73
2. Tabla de estimación de los parámetros del modelo.	74
3. Tabla de residuales del modelo	74

## RESUMEN

El presente trabajo de investigación comprende la aplicación y análisis de modelos univariantes, siguiendo la Teoría de WIENER KOLMOGOROV, más conocido como el enfoque de BOX-JENKINS, mediante el cual se elige el proceso adecuado para realizar predicciones. El ámbito geográfico de estudio se encuentra en la ciudad de Puno y busca responder a la siguiente interrogante: ¿Cuál es el modelo univariante que mejor se ajusta a la serie histórica de los niveles medios de agua del lago Titicaca para describir y pronosticar el comportamiento de las variaciones de los niveles medios mensual de agua del lago Titicaca de la región de Puno, periodo 1974-2018? El objetivo general es determinar el modelo univariante que mejor se ajusta a la serie histórica de los niveles medios de agua mensual del lago Titicaca, para describir y pronosticar el comportamiento de las variaciones fluctuantes del lago Titicaca en la región de Puno, periodo 1974 - 2018. La hipótesis que se plantea es que el modelo univariante que mejor se ajusta es el modelo ARIMA multiplicativo  $(p, d, q)*(P, D, Q)$ . Se llega a la conclusión después de identificar, estimar y validar el modelo que mejor se ajusta a la serie historia de niveles medios de agua mensual del lago Titicaca, es efectivamente el modelo ARIMA multiplicativo  $(1, 1, 0) (0, 1, 1)$ , cuya ecuación de pronóstico es la siguiente:  $\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + 0.32595Y_{t-1} - 0.81061\hat{\varepsilon}_{t-12}$ .

**Palabras clave:** Estimación de parámetros, función de autocorrelación, modelo ARIMA, pronóstico, validación del modelo.

## ABSTRACT

This research work includes the application and analysis of univariate models, following the WIENER KOLMOGOROV Theory, better known as the BOX-JENKINS approach, by which the appropriate process for making predictions is chosen. The geographical scope of study is in the city of Puno and seeks to answer the following question: What is the univariate model that best fits the historical series of the average water levels of Lake Titicaca to describe and predict the behavior of the variations of the average monthly water levels of Lake Titicaca of the Puno region, 1974 -2018? The general objective is to determine the univariate model that best fits the historical series of the average monthly water levels of Lake Titicaca, to describe and predict the behavior of fluctuating variations of Lake Titicaca in the Puno region, period 1974 - 2018. The hypothesis that arises is that the univariate model that best fits is the multiplicative ARIMA model  $(p, d, q) * (P, D, Q)$ . The conclusion is reached after identifying, estimating and validating the model that best fits the history series of average monthly water levels of Lake Titicaca, it is effectively the multiplicative ARIMA model  $(1, 1.0) (0, 1, 1)$ , whose forecast equation is as follows:  $\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + 0.32595Y_{t-1} - 0.81061\hat{\epsilon}_{t-12}$ .

**Keywords:** ARIMA model, autocorrelation function, forecast, model validation, parameter estimation.

## INTRODUCCIÓN

El lago Titicaca es el lago navegable más alto del mundo, situado en el altiplano andino, dentro de la meseta del Collao, a una altitud media de 3812 m s. n. m. entre los territorios de Perú y Bolivia. Conserva un área de 8562 km<sup>2</sup> de los cuales el 56 % (4772 km<sup>2</sup>) pertenecen a Perú y el 44 % (3790 km<sup>2</sup>) a Bolivia y 1125 km de ribera; su profundidad máxima se estima en 281 m y se calcula su profundidad media en 107 m, es de importancia, al ser el mayor lago de agua dulce con diversidad biológica de América Latina.

Así mismo también es relevante, saber el comportamiento de las variaciones futuras de los niveles de agua del lago Titicaca, ya que estas nos ayudaran a la toma de decisiones es decir a la prevención y planificación frente a los problemas ambientales. Es por ello que la finalidad de este trabajo es determinar mediante un modelo univariante los pronósticos de las variaciones de los niveles medios mensuales de agua del lago Titicaca de la región de Puno. Para ello recurriremos a la metodología Box-Jenkins, utilizando información de los niveles medios de agua del lago Titicaca por meses, desde el año 1974-2018 buscando luego un modelo adecuado que permitió ajustar el comportamiento de los niveles medios de agua.

La presente investigación presenta los siguientes Capítulos:

**Capítulo I.** En este capítulo se presenta la Revisión de Literatura. Y también se desarrolla el marco conceptual, y los antecedentes que contiene un mínimo de 15 estudios previos, que dan cuenta de los principales hallazgos y contribuciones a la investigación.

**Capítulo II.** En este capítulo se da a conocer el Planteamiento del Problema. Contiene la definición del problema, la justificación, los objetivos y las hipótesis de investigación.

El planteamiento del problema y la postura epistémica que son claros, concisos, confiables y concordantes. En este sentido, expresa la relación entre el problema de investigación, la pregunta de investigación y la intención para investigar.

En la justificación, se fundamentó la contribución a los vacíos de conocimiento y a la resolución del problema de investigación. Los objetivos de esta investigación comprenden la intención del estudio y permite describir el tipo de proceso que ha seguido

la investigación. Las hipótesis dan cuenta de las respuestas probables al problema objeto de investigación.

**Capítulo III.** En este capítulo se abarca sobre los materiales y métodos. En la parte de lugar de estudio, se debe identificó el lugar donde se realizó la investigación.

**Capítulo IV.** En este capítulo se describen los Resultados y Discusión. Los resultados se presentan por objetivos específicos, así mismo se da la interpretación de información contenida en tablas y/o figuras, demostrando la aceptación o rechazo de las hipótesis mediante la prueba estadística, según sea el caso.

**Conclusiones.** Se desarrolló de acuerdo a los objetivos específicos y se redactó en párrafos.

**Recomendaciones.** Se presentó en relación con los objetivos específicos, dando cuenta de las consecuencias, orientaciones y medidas a realizarse.

## CAPÍTULO I

### REVISIÓN DE LITERATURA

#### 1.1. Marco teórico

##### 1.1.1. Pronóstico

De acuerdo a Uriel (1985) los pronósticos son predicciones de lo que puede suceder o esperar, son premisas o suposiciones básicas en que se basan la planeación y la toma de decisiones.

El propósito del pronóstico consiste en reducir el margen de incertidumbre, haciendo el mejor uso de la información, ya que se tiene para guiar las actividades de la empresa hacia el cumplimiento de sus metas y objetivos. De esta forma los pronósticos son particularmente importantes en la asignación del uso de los recursos de la empresa.

Los pronósticos se basan en el uso de datos anteriores de una variable para predecir su desempeño futuro. A este respecto, los datos anteriores se dan generalmente en la forma de series de tiempo. Una hipótesis básica en la aplicación de las técnicas de pronóstico es que el desempeño de los datos anteriores continuará ocurriendo en el futuro inmediato. Evidencias empíricas indican que este supuesto es válido en muchas situaciones reales, sobre todo cuando las series de tiempo representan una larga historia de las variables analizadas.

##### 1.1.2. La necesidad de pronosticar

Hanke y Reitsch (1996) afirma que debido a que siempre ha sido cambiante el mundo en el que operan las organizaciones, siempre ha existido la necesidad de

hacer pronósticos. Sin embargo, en los últimos años se ha incrementado la confianza en las técnicas que abarcan una compleja manipulación de datos.

Las computadoras, junto con las técnicas cuantitativas, se han vuelto más que recomendables en las organizaciones modernas; se han vuelto esenciales.

¿Quién requiere hacer pronósticos?, Cualquier organización, grande y pequeña, pública y privada, utiliza el pronóstico, debido a que casi todas las organizaciones deben planear cómo enfrentar las condiciones futuras de las cuales tiene un conocimiento imperfecto. Además, la necesidad de hacer pronósticos cruza todas las líneas funcionales, lo mismo que todo tipo de organizaciones.

### **1.1.3. Técnicas de pronósticos.**

Bowerman y O'Connell (1987) refiere que se pueden emplear dos técnicas básicas de pronósticos: Las técnicas de pronóstico cualitativas y las técnicas de pronóstico cuantitativas.

#### **a) Técnicas de pronóstico cualitativas.**

Esta técnica es apropiada cuando los datos confidenciales son escasos o difíciles de utilizar. Se fundamentan en el juicio humano y en la intuición, más que en la manipulación de datos históricos anteriores. Las técnicas cualitativas comunes incluyen al método Delphi, curvas de crecimiento, escritura de escenarios, investigación de mercado y grupos de enfoque.

#### **b) Técnicas de pronóstico cuantitativas.**

Las técnicas de pronóstico cuantitativas se utilizan cuando existen suficientes datos históricos disponibles y cuando se juzga que estos datos son representativos de un futuro desconocido. Trabajan con modelos cuantitativos o modelos matemáticos que se basan en datos históricos, bajo el supuesto de que son relevantes para el futuro. Estos modelos se pueden utilizar con series de tiempo.

### **1.1.4. Pronósticos según plazos.**

Los pronósticos a largo plazo son necesarios para establecer el curso general de la organización para un largo periodo, sirven para tomar decisiones estratégicas y por lo general abarcan de tres a cinco años. Los pronósticos a mediano plazo abarcan

de uno a dos años. Los pronósticos a corto plazo se utilizan para diseñar estrategias inmediatas que ayuden en la toma de decisiones, sólo abarcan meses.

Las técnicas más complejas de Box-Jenkins resultan apropiadas para pronósticos de corto y mediano plazos.

### **1.1.5. Series de tiempo.**

Hanke y Reitsch (1996) afirman que las series de tiempo son un conjunto de observaciones ordenadas según una característica cuantitativa de un fenómeno individual en diferentes momentos del tiempo, en el cual las observaciones son realizadas.

Una serie de tiempo consta de datos que se reúnen, registran u observan sobre incrementos sucesivos de tiempo.

Una serie de tiempo es una secuencia cronológica de observaciones de una variable en particular.

Una serie de tiempo es un conjunto de observaciones producidas en determinados momentos durante un periodo, semanal, mensual, trimestral o anual, generalmente a intervalos iguales. El primer paso para analizar una serie de tiempo es graficarla, esto permite identificar la tendencia, la estacionalidad, las variaciones irregulares (componente aleatorio). Un modelo clásico para una serie de tiempo puede ser expresada como suma o producto de tres componentes: tendencia, estacional y un término de error aleatorio.

#### **1.1.5.1. Componentes de una serie temporal.**

De acuerdo con Chatfield (1987) en el análisis de series de tiempo de datos, una tentación inmediata consiste en intentar explicar o contabilizar el comportamiento de las series. La descomposición clásica es un método que se basa en la suposición de que se pueden descomponer en componentes como tendencia, ciclo, estacionalidad e irregularidad. Una predicción se hace mediante la combinación de las proyecciones de cada componente individual.

- **Tendencia.**

La tendencia es un movimiento de larga duración que muestra la evolución general de la serie en el tiempo. Es un movimiento que puede ser estacionario o ascendente o descendente, y su recorrido, una línea recta o una curva.

- **Componente cíclico o Variación cíclica.**

El componente cíclico es un conjunto de fluctuaciones en forma de onda o ciclos, de más de un año de duración. El ciclo sugiere la idea de que este tipo de movimiento se repite cada cierto periodo con características parecidas.

- **Componente estacional o Variación estacional.**

Se habla de este tipo de variaciones usualmente cuando el comportamiento de la variable en el tiempo en un periodo que está relacionado con la época o un periodo particular, por lo general en el espacio cronológico presente.

El componente estacional es un patrón de cambio que se repite, a sí mismo año tras año. Se encuentran típicamente en los datos clasificados por trimestre, mes o semana.

- **Componente aleatorio o Variación residual.**

El componente aleatorio mide la variabilidad de las series de tiempo después de que se retiran los otros componentes. Contabiliza la variabilidad aleatoria en una serie de tiempo ocasionada por factores imprevistos y no recurrentes. La mayoría de los componentes irregulares se conforman de variabilidad aleatoria. Sin embargo, ciertos sucesos a veces impredecibles como huelgas, cambios de clima (inundaciones, sequías o terremotos), elecciones, conflictos armados o la aprobación de asuntos legislativos, pueden causar irregularidades en una variable.

#### **1.1.5.2. Análisis de series de tiempo.**

Oliver (1976) refiere que el análisis de series de tiempo está dedicado al estudio de series; por lo general, los datos de dichas series son

independientes, pero están correlacionados; se puede decir que existe una relación entre observaciones contiguas.

Es el análisis de una secuencia de medidas hechas a intervalos específicos. El tiempo es usualmente la dimensión dominante de los datos. Sirven para establecer la efectividad de medidas que afectan a grupos poblacionales teniendo en cuenta las variaciones naturales que puede haber en el tiempo. Son muy comunes en la evaluación de leyes en la población. Permiten una visión parcial de la relación causa efecto, pero no pueden extrapolar los hallazgos de la población a individuos específicos.

El análisis de series de tiempo consiste en una descripción (generalmente matemática) de los movimientos y componentes presentes.

Chatfield (1987) son varios los objetivos por los cuales se desea analizar una serie de tiempo:

**Descripción:** Al tener una serie de tiempo, el primer paso en el análisis es graficar los datos y obtener medidas descriptivas simples de las propiedades principales de la serie.

**Explicación:** Cuando las observaciones son tomadas sobre dos o más variables, es posible usar la variación en una serie para explicar la variación en las otras series.

**Predicción:** Dada una serie de tiempo se intenta predecir los valores futuros de la serie. Este es el objetivo más frecuente en el análisis de series de tiempo.

**Control:** Si una serie de tiempo se genera por mediciones de calidad de un proceso, el objetivo del análisis puede ser el control del proceso.

### 1.1.5.3. Estimación de la tendencia

La tendencia se estima de la siguiente manera:

- ✓ Método de los Mínimos Cuadrados.

El método de los mínimos cuadrados se utiliza para calcular la ecuación de una recta o curva de tendencia apropiada que permite identificar el grado de correlación entre una variable dependiente y una o más variables independientes. El criterio de este método es usar la recta ( $Y = a + bX$ ) cuya suma de los cuadrados de los errores sea mínima.

Cálculo para b:

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

Cálculo para a:

$$a = \frac{\sum x^2 \sum y - \sum x \sum xy}{n \sum x^2 - (\sum x)^2}$$

$$a = \bar{Y} - Xb$$

Donde:

a: Intersección en el eje vertical.

b: Pendiente de la línea de regresión

n: Número de observaciones.

x: Valores de y que caen en la línea de tendencia.

y: valores de x que caen en la línea de tendencia.

✓ Método “a mano”.

El método consiste en trazar una recta o curva de tendencia simplemente observando la gráfica, se usa para estimar la ecuación de una recta o curva de tendencia apropiada, sin embargo tiene la desventaja de depender del juicio individual.

#### 1.1.6. Proceso estocástico.

“Un proceso estocástico se define como una familia de variables aleatorias que corresponden a momentos sucesivos del tiempo. Será designado por  $Y(t,u)$ , donde t es el tiempo y u es la variable aleatoria (De la Fuente, Pino, y Parreño, 1995)

Un proceso estocástico es una familia de variables aleatorias que corresponden a momentos sucesivos de tiempo.

Se denomina proceso estocástico a la sucesión infinita de variables aleatorias ordenadas.

$$\dots, Y_1, Y_2, \dots, Y_t, \dots$$

Si se dispone de un conjunto finito de estas variables,  $Y_1, Y_2, \dots, Y_t$ . se dice que esta sucesión de observaciones (realizaciones) forma una serie temporal.

### 1.1.6.1. Proceso estocástico estacionario.

De la Fuente, et al (1995) afirma que un proceso estocástico es estacionario en sentido estricto cuando al realizar un mismo desplazamiento en el tiempo de todas las variables de cualquier distribución conjunta finita, resulta que esta distribución no varía.

Considerando la función de distribución conjunta

$$F(Y_{t1}, Y_{t2}, \dots, Y_{tk})$$

Si se adopta el supuesto de que a todos los elementos de la anterior distribución se desplazan  $m$  periodos, la nueva función de distribución conjunta será

$$F(Y_{t1+m}, Y_{t2+m}, \dots, Y_{tk+m})$$

Si el proceso es estacionario en sentido estricto se deberá verificar que

$$F(Y_{t1}, Y_{t2}, \dots, Y_{tk}) = F(Y_{t1+m}, Y_{t2+m}, \dots, Y_{tk+m})$$

Y a su vez, se deberá obtener un resultado análogo para cualquier otra distribución conjunta que tenga carácter finito.

Se dice que un proceso es estacionario de primer orden, o en media, si se verifica que:  $E(Y_t) = \mu, \forall t$ .

Por tanto, en un proceso estacionario en media, la esperanza matemática, o media teórica, permanece constante a lo largo del tiempo.

Se dice que un proceso es estacionario de segundo orden (o en sentido amplio) cuando se verifican las dos condiciones siguientes:

- 1) La varianza es finita y permanece constante a lo largo del tiempo, es decir:

$$E(Y_t - \mu)^2 = \sigma^2 < \infty, \quad \forall t$$

- 2) La autocovarianza entre dos periodos distintos de tiempo únicamente viene afectada por el lapso de tiempo transcurrido entre esos dos periodos. Así:

$$E(Y_{t+k} - \mu)(Y_t - \mu) = \gamma_k, \quad \forall t$$

Que sería una autocovarianza de orden  $k$ , por ser éste el lapso de tiempo que separa a  $Y_t$  de  $Y_{t+k}$ . Su valor,  $\gamma_k$  es independiente de cuál sea el periodo  $t$  que se considere. Como puede verse, la varianza del proceso es simplemente la autocovarianza de orden 0.

Al definir un proceso estacionario en sentido amplio, se tiene en cuenta implícitamente que el proceso es también estacionario en media, ya que tanto en la varianza como en las autocovarianzas, el símbolo  $\mu$  no viene afectado por ningún subíndice.

En un proceso estacionario las autocorrelaciones quedan definidas de la siguiente forma:

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \quad \text{donde } k \geq 0$$

La representación gráfica de  $\rho_k$  para  $k=0, 1, 2, 3, \dots$ , recibe el nombre de correlograma.

### 1.1.7. Ruido blanco.

Se llama ruido blanco a una sucesión de variables aleatorias (proceso estocástico) con esperanza (media) cero, varianza constante e independientes para distintos valores de  $t$  (covarianza nula). Un tipo especial de proceso estocástico es el denominado ruido blanco.

Una variable  $\varepsilon_t$  se denomina “ruido blanco” si cumple las siguientes condiciones:

$$i) E[\varepsilon_t] = 0 \quad \forall t$$

$$ii) E[\varepsilon_t]^2 = \sigma^2 \quad \forall t$$

$$iii) \text{COV}[\varepsilon_t, \varepsilon_s] = E[\varepsilon_t, \varepsilon_s] = 0 \quad \forall t \neq s$$

### 1.1.8. Estacionalidad.

Estacionalidad es una variación repetitiva a lo largo de un intervalo de tiempo. Se dice que una serie de tiempo es estacional cuando los datos tienen oscilaciones estrictamente periódicas, donde el periodo es igual o inferior al año, por ejemplo la repetición del patrón puede ser cada 3 meses, 6 meses, cada año (12 meses), cada 4 años, etc.

### 1.1.9. Coeficiente de correlación.

El coeficiente de correlación mide el grado de independencia de una variable relacionada con otra variable. Es una cantidad que está entre -1 y +1, presenta el grado de correlación entre dichas variables; mientras este valor se aproxima a los límites, diremos que la correlación es buena, se expresa:

$$r = \sqrt{R^2}$$

Donde  $R^2$  es el coeficiente de determinación, y es una medida de ajuste de la regresión. Es determinado por la proporción de la suma de cuadrados del total. Se expresa como:

$$R^2 = \frac{\sum_{i=1}^n [\hat{Y}_i - E(Y)]^2}{\sum_{i=1}^n [Y_i - E(Y)]^2}$$

El error siempre existirá; en estadística es posible lograr una menor varianza de error  $\sigma_e^2$ , como lograr un  $R^2$  cercano a los límites.

### 1.1.10. Varianza y error estándar de ruido blanco.

La varianza de ruido blanco de la regresión es una medida que sirve para analizar la capacidad explicativa del modelo. Se expresa como:

$$\sigma_e^2 = \sum_{p=1}^T [Y_t - \hat{Y}_t]^2 / T - k$$

Dónde:

T es el número de residuos o errores.

K es el número de estimadores.

El error estándar o desviación estándar de la regresión está dado por:

$$\sigma_e = \sqrt{\sigma_e^2}$$

### 1.1.11. Función de autocorrelación.

Se usa para ver si la serie es estacionaria o no estacionaria. La función de autocorrelación mide la correlación entre los valores de la serie distanciados un lapso de tiempo  $k$ .

Dada una muestra  $Y_0, Y_1, \dots, Y_{n-1}$  de  $n$  observaciones, la función de autocorrelación muestral (FAC) de la muestra al rezago  $k$ , denotada por  $\rho_k$ , se define como:

$$\hat{\rho}_k = \begin{cases} \gamma_k & \text{para } k = 1, \dots, n-1 \\ \gamma_0 & \text{para } k = 0 \\ 1 & \text{para } k = 0 \end{cases}$$

Donde:

$\gamma_0$  es la varianza de la muestra dada por:

$$\gamma_0 = \frac{1}{n} \sum_{t=1}^n (Y_t - \bar{Y})^2$$

$\gamma_k$  es la covarianza al rezago  $k$  definida como:

$$\gamma_k = \frac{1}{n} \sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})$$

La función de autocorrelación indicacuánta correlación existe entre datos individuales contiguos en la serie  $Y_t$ . Conforme el valor del retraso aumenta, el número de observaciones comprendidas en la autocovarianza disminuye hasta el elemento final. Al graficar  $\rho_k$  frente a  $k$ , la gráfica obtenida se conoce como correlograma.

### 1.1.12. Función de autocorrelación parcial.

La Función de Autocorrelación Parcial se emplea para ayudar a identificar el grado de relación entre los valores reales de una variable y valores anteriores de la misma, mientras que se mantienen constantes los efectos de las otras variables (periodos retrasados).

La función de autocorrelación parcial (*FACP*) de la muestra  $\rho_k$  en el retraso  $k$  es la correlación entre observaciones (series de tiempo) que están separadas  $k$  periodos de tiempo, manteniendo constantes las correlaciones en los rezagos intermedios (es decir rezagos menores de  $k$ ). En otras palabras, la autocorrelación parcial es la correlación entre  $Y_t$  y  $Y_{t-k}$  después de eliminar el efecto de las  $Y$  intermedias.

En la siguiente tabla se dan algunos lineamientos generales acerca de los patrones típicos de las funciones de correlaciones muestral y parcial.

Tabla 1  
*Patrones teóricos de FAC y FACP*

Tipo de modelo	Patrón típico de FAC	Patrón típico de FACP
AR (p)	Disminuye exponencialmente o con un patrón decreciente o ambos.	Picos grandes a lo largo de los p rezagos.
MA (q)	Picos grandes a lo largo de los q rezagos.	Decrece exponencialmente.
ARMA (p,q)	Decrece exponencialmente.	Decrece exponencialmente.

Fuente: Gujarati (2004)

### 1.1.13. Modelo.

Un modelo es una expresión formalizada de una teoría, o la representación matemática de los datos observados. En el análisis estadístico un modelo es expresado en símbolos, en forma matemática.

Es un esquema teórico, generalmente en forma matemática, de un sistema o realidad compleja, como la evolución económica de un país, que se elabora para facilitar su comprensión y el estudio de su comportamiento.

#### 1.1.14. Modelos de series temporales.

Uriel (1985) afirma que son formas teóricas determinísticas y/o aleatorias o la combinación de ambas, para realizar el análisis de una serie de tiempo.

- Variables Temporales: Son variables que se observan a lo largo del tiempo.  $Y_t$  indica la variable  $Y$  en el momento  $t$ .
- Serie Temporal: Es el conjunto de  $t$  observaciones, una observación por cada una de las variables:  $Y_1, Y_2, \dots, Y_t$ . También es llamada serie cronológica.

A las observaciones de una variable temporal se les denomina “realizaciones”.

#### 1.1.15. Modelo univariante.

Los modelos univariantes en una serie de tiempo  $\{Y_t\}$ , son todos aquellos que solamente tienen una variable observada en el tiempo. Estos tipos de modelos se expresan en forma polinomial.

Son técnicas univariantes: el modelo autorregresivo de primer orden, el modelo de tendencia lineal o exponencial, entre otros.

Las técnicas más rigurosas para la predicción univariante son las denominadas técnicas o modelos Box-Jenkins, o más concretamente modelos ARIMA, pues las técnicas Box-Jenkins constituyen un conjunto más amplio, dentro del cual los modelos ARIMA univariantes son sólo una parte.

##### a) Modelo Univariante No Integrado.

Los procesos Autoregresivos AR ( $p$ ), de Medias Móviles MA ( $q$ ) y procesos Mixtos ARMA ( $p,q$ ) son considerados como los modelos No-Integrados debido a que no interviene el grado de diferenciación y la estacionalidad de la serie.

##### b) Modelo Univariante Integrado.

Son aquellos modelos que se pueden obtener mediante suma o integración de un proceso estacionario. A estos modelos se les denomina también modelos no estacionarios homogéneos.

Los procesos Mixtos Integrados ARIMA (p, d, q), los procesos Estacionales Mixtos Integrados ARIMA (p, d, q) \* (P, D, Q), procesos de Medias Móviles Integrado IMA, proceso de Medias Móviles Exponenciales EWMA, y los procesos de autorregresivos, son considerados como los modelos integrados.

#### 1.1.16. Operador de retardo y diferenciación de una serie.

Introduciremos a continuación el operador polinomial de retardos L. El operador L determina que:

$$LY_t = Y_{t-1}$$

Es decir, el resultado de aplicar el operador L corresponde a la observación en el período anterior de la variable (serie).

Aplicada dos veces sobre la variable  $Y_t$  es:

$$L(LY_t) = L^2Y_t = Y_{t-2}$$

En general:

$$L^k Y_t = Y_{t-k}$$

La diferencia de una serie es:

$$\Delta Y_t = Y_t - Y_{t-1} = (1 - L)Y_t$$

En general:

$$\Delta^k Y_t = \Delta(\Delta^{k-1} Y_t) = (1 - L)^k Y_t$$

#### 1.1.17. Elaboración de modelos AR, MA, ARMA y ARIMA.

Los modelos ARIMA o modelos de promedio móvil autorregresivo integrado son un tipo general de los modelos de Box-Jenkins para series de tiempo estacionarias. Una serie histórica estacionaria es aquella cuyo valor promedio no cambia a través del tiempo. Este grupo incluye a los modelos AR sólo con términos autorregresivos, los modelos MA sólo con términos de promedio móvil y los modelos ARIMA que comprenden tanto términos autorregresivos como de promedio móvil. La metodología de Box-Jenkins permite al analista seleccionar el modelo que mejor se ajuste a sus datos.

Dado el concepto de proceso estacionario anteriormente definido, los modelos de pronóstico se dividen en:

- Modelos Lineales Estacionarios
- Modelos Lineales No Estacionarios

**a) Modelos lineales estacionarios**

**a.1) Modelos autorregresivos (AR)**

Los modelos autorregresivos (AR) expresan  $Y_t$  como una función lineal de cierto número de valores anteriores reales de  $Y_t$ .

- Modelo AR(p): Un modelo autorregresivo de orden  $p$ , o abreviadamente un modelo AR(p), se define de la siguiente forma:

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

Donde:

$Y_t$ : Variable dependiente.

$Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ : Variables independientes que son variables dependientes desfasadas un número específico de periodos.

$C, \phi_1, \phi_2, \dots, \phi_p$ : Coeficientes de regresión.

$\varepsilon_t$  es el término de residuo que representa sucesos aleatorios no explicados por el modelo. También se le conoce como “ruido blanco”, y  $\varepsilon_t \sim (0, \sigma^2)$

Usando el operador polinomial de retardos L, se denota como:

$$Y_t = C + \phi_1 L Y_t + \phi_2 L^2 Y_t + \dots + \phi_p L^p Y_t + \varepsilon_t$$

Realizando algunas transformaciones tenemos:

$$Y_t - \phi_1 L Y_t - \phi_2 L^2 Y_t - \dots - \phi_p L^p Y_t = C + \varepsilon_t$$

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) Y_t = C + \varepsilon_t$$

La expresión entre paréntesis como un polinomio en el operador de retardos  $L$ , se puede expresar de forma compacta:

$$\phi(L)Y_t = C + \varepsilon_t$$

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

Para que el proceso sea estacionario se requiere que las raíces de la ecuación polinomial estén fuera del círculo unidad.

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

- Modelo AR(1): El caso más sencillo corresponde a un modelo autorregresivo de primer orden, donde el parámetro  $C$  es igual a cero:

El modelo autorregresivo de primer orden, viene definido por  $Y_t = \phi_1 Y_{t-1} + \varepsilon_t$ , donde  $|\phi_1| < 1$ , o utilizando el operador de retardos, por  $(1 - \phi_1 L)Y_t = \varepsilon_t$ .

Cada variable ruido blanco influye sobre los valores de  $Y$  correspondientes al mismo periodo, o a periodos posteriores, pero nunca ejerce influencia sobre los valores de  $Y$  correspondientes a periodos anteriores. Una consecuencia importante es que:

$$E[\varepsilon_t Y_{t-\tau}] = 0 \quad \forall \tau > 0.$$

Para que el proceso AR (1) definido, sea estacionario, la raíz del polinomio característico  $1 - \phi_1 L = 0$ , debe caer fuera del círculo unidad. Es decir;  $|L| = \left| \frac{1}{\phi_1} \right| > 1$ , lo que equivale a que  $|\phi_1| < 1$ .

- Modelo AR(2): Un modelo autorregresivo de segundo orden, viene definido por:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

O, utilizando el operador de retardos,

$$(1 - \phi_1 L - \phi_2 L^2)Y_t = \varepsilon_t$$

Para que el proceso anterior sea estacionario se requiere que las raíces de la ecuación estén situadas fuera del círculo unidad, es decir:

$$1 - \phi_1 L - \phi_2 L^2 = 0$$

Si se cumplen las condiciones de estacionariedad se verificará que  $E(Y_t) = 0$

- Condiciones de estacionariedad AR

Para que un modelo AR de orden p,

$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$ , sea estacionario, se debe cumplir lo siguiente:

Usando el operador polinomial de retardos L:

$$Y_t = \phi_1 L Y_t + \phi_2 L^2 Y_t + \dots + \phi_p L^p Y_t + \varepsilon_t$$

Realizando algunas transformaciones se tiene:

$$Y_t - \phi_1 L Y_t - \phi_2 L^2 Y_t - \dots - \phi_p L^p Y_t = \varepsilon_t$$

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

Y en forma compacta:

$$\phi(L) Y_t = \varepsilon_t$$

Para que el proceso sea estacionario se requiere que las raíces de la ecuación polinomial estén fuera del círculo unidad.

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

La estacionalidad de la serie  $Y_t$  requiere, entre otras condiciones una media que no varía; es decir que no debe existir una tendencia a lo largo del tiempo.

El polinomio autorregresivo usando el operador de retardo L, es:

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

**a.2) Modelos de medias móviles (MA).**

Los modelos autorregresivos (AR) expresan  $Y_t$  como una función lineal de cierto número de valores anteriores reales de  $Y_t$ , mientras que los modelos de promedio móvil (MA) proporcionan pronósticos de  $Y_t$  con base en una combinación lineal de errores anteriores de  $Y_t$ .

Un modelo de medias móviles explica el valor de una determinada variable en un período  $t$  en función de un término independiente y una sucesión de errores correspondientes a períodos precedentes, ponderados convenientemente. Se denotan normalmente con las siglas MA, como en el caso de los modelos autorregresivos, del orden entre paréntesis. Así, un modelo con  $q$  términos de error se denota como MA ( $q$ ).

- Modelo MA( $q$ ).

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

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Donde

$Y_t$ : Variable dependiente.

$\theta_1, \theta_2, \dots, \theta_q$ : Peso específico.

$\varepsilon_t$ : Residuo o error. También conocido como “ruido blanco”.

$\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$ : Valores previos de residuos.

Utilizando el operador polinomial de retardos,  $\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$ , el modelo de medias móviles se puede expresar de forma compacta  $Y_t = \theta(L)\varepsilon_t$ .

Donde la media es cero, cualesquiera que sean los valores de  $\theta_i$ , es decir:

$$E[Y_t] = \theta(L)E(\varepsilon_t) = 0$$

Si en el modelo MA (q) se incluye un término constante

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Donde:

$Y_t$ : Variable dependiente.

$\mu, \theta_1, \theta_2, \dots, \theta_q$ : Peso específico.

$\varepsilon_t$ : Error aleatorio o residuo. También conocido como “ruido blanco”.

$\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$ : Valores previos de residuos.

Entonces al tomar esperanzas matemáticas en la expresión anterior resulta

$$E[Y_t] = \mu$$

Así pues, en los modelos de medias móviles, la media del proceso coincide con el término independiente, que aparece en el segundo miembro. Sin pérdida de generalidad se supondrá en lo sucesivo que  $\mu = 0$ .

Para que un proceso MA (q) sea invertible se requiere que las raíces de la ecuación polinomial caigan fuera del círculo unidad.

$$1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q = 0$$

Calculando los momentos del proceso, a partir del operador de retardo L se tiene:

$$\begin{aligned} E(Y_t) &= E(\mu + \varepsilon_t - \theta_1 L \varepsilon_t - \theta_2 L^2 \varepsilon_t - \dots - \theta_q L^q \varepsilon_t) \\ &= E(\mu) + E(\Theta_q(L) \varepsilon_t) = \mu + \Theta_q(L) E(\varepsilon_t) \end{aligned}$$

Respecto de la varianza, se tiene:

$$V(Y_t) = E(Y_t^2) = E\left[\left(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}\right)^2\right]$$

- Modelo MA (1).

Un modelo MA (1) viene definido por:

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

Donde  $\varepsilon_t$  es un ruido blanco con las propiedades, ya definidas.

Un modelo MA (1) omitiendo la constante viene definido por:

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

$$Y_t = (1 - \theta_1 L) \varepsilon_t$$

Donde  $\varepsilon_t$  es un ruido blanco.

Para que la ecuación sea estable se requiere que la raíz del polinomio característico caiga fuera del círculo unidad, es decir

$$1 - \theta_1 L = 0$$

$$|L| = \left| \frac{1}{\theta_1} \right| > 1$$

O, de forma equivalente, que  $|\theta_1| < 1$

La condición de invertibilidad de un modelo MA (1) es equivalente en sentido formal a la condición de estacionariedad de un modelo AR (1). Un modelo MA (1) es siempre estacionario y la condición de invertibilidad se establece para poder pasar a un modelo AR ( $\infty$ ).

- Modelo MA (2).

Un modelo MA (2) omitiendo la constante viene definido por:

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} = (1 - \theta_1 L - \theta_2 L^2) \varepsilon_t$$

Donde  $\varepsilon_t$  es un ruido blanco.

Para que un proceso MA (2) sea invertible se requiere que las raíces del polinomio característico caigan fuera del círculo unidad.

$$1 - \theta_1 L - \theta_2 L^2 = 0$$

- Condiciones de estacionariedad e invertibilidad en los procesos MA

Los modelos de medias móviles finitos son siempre estacionarios. Como puede apreciarse de las deducciones anteriores, los momentos de los procesos (esperanza, varianza y autocovarianzas) son invariantes en el tiempo. A diferencia de los procesos AR, para la deducción de los momentos no es necesario suponer la estacionariedad de la serie. Si las raíces del polinomio de medias móviles caen fuera del círculo unidad, el proceso será también invertible.

Veamos el caso de un proceso MA (1) (omitiendo la constante para simplificar la exposición). Se tiene:

$$\begin{aligned}\varepsilon_t &= Y_t + \theta_1 \varepsilon_{t-1} \\ \varepsilon_{t-1} &= Y_{t-1} + \theta_1 \varepsilon_{t-2}\end{aligned}$$

Sustituyendo la segunda expresión en la primera:

$$\begin{aligned}\varepsilon_t &= Y_t + \theta_1 [y_{t-1} + \theta_1 \varepsilon_{t-2}] \\ \varepsilon_t &= Y_t + \theta_1 y_{t-1} + \theta_1^2 \varepsilon_{t-2}\end{aligned}$$

Y continuando con la sustitución recursiva:

$$\begin{aligned}\varepsilon_t &= Y_t + \theta_1 Y_{t-1} + \theta_1^2 Y_{t-2} + \dots + \theta_1^i Y_{t-i} + \dots \\ Y_t &= \varepsilon_t - (\theta_1 Y_{t-1} + \theta_1^2 Y_{t-2} + \dots + \theta_1^i Y_{t-i} + \dots) \\ \Rightarrow Y_t &= -\sum_{i=1}^{+\infty} \theta_1^i Y_{t-i} + \varepsilon_t\end{aligned}$$

Alternativamente:

$$\begin{aligned}\varepsilon_t &= Y_t + \theta_1 Y_{t-1} + \theta_1^2 Y_{t-2} + \dots + \theta_1^i Y_{t-i} + \dots \\ (1 + \theta_1 L + \theta_1^2 L^2 + \dots + \theta_1^i L^i + \dots) Y_t &= \varepsilon_t\end{aligned}$$

Es decir, en principio es posible expresar el modelo MA (1) como un AR( $\infty$ ).

Para que efectivamente ambos modelos sean equivalentes se requiere que el modelo AR sea estacionario, lo que impone la condición que  $|\theta| < 1$ .

De esta forma, cuando el polinomio de medias móviles  $\theta_q(L)$  tiene sus raíces fuera del círculo unidad, el proceso de medias móviles puede transformarse en un proceso AR estacionario.

Otra forma de mostrar la invertibilidad para procesos MA (1).

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} = (1 - \theta_1 L) \varepsilon_t \Rightarrow$$

$$\Rightarrow \varepsilon_t = \frac{1}{(1 - \theta_1 L)} Y_t$$

### a.3) Modelos mixtos autorregresivos - medias móviles (ARMA).

La combinación de modelos autorregresivos (AR) y de Medias Móviles (MA) da lugar al modelo ARMA. Un modelo ARMA  $(p,q)$  se define de la siguiente forma:

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

Donde  $\varepsilon_t$  es un ruido blanco.

Utilizando los operadores polinomiales de retardo, el modelo se expresará de forma compacta de la siguiente forma:

$$\theta(L)Y_t = \theta(L)\varepsilon_t$$

Para que el modelo sea estacionario se requiere que las raíces de la ecuación polinomial caigan fuera del círculo unidad.

$$\theta(L) = 1 - \phi_1 L - \dots - \phi_p L^p = 0$$

Si se cumplen las condiciones de estacionariedad, el modelo ARMA  $(p,q)$  se puede expresar como un  $MA(\infty)$ , pudiendo representarse de la siguiente forma:

$$Y_t = \frac{\theta(L)}{\phi(L)} \varepsilon_t = \psi(L)\varepsilon_t$$

Por tanto los coeficientes del operador polinomial  $\psi(L)$ , que tiene infinitos elementos, deben cumplir la siguiente identidad:

$$\phi(L)\psi(L) = \theta(L)$$

O, en notación más detallada, se establece la siguiente identidad:

$$(1 - \phi_1 L - \dots - \phi_p L^p)(1 + \psi_1 L + \psi_2 L^2 + \dots) \equiv (1 - \theta_1 L - \dots - \theta_q L^q)$$

A partir de la anterior identidad, se pueden deducir un conjunto de ecuaciones que nos permiten obtener los  $\psi_i$ , en función de los coeficientes  $\phi_h$  y  $\theta_j$ . Así en un modelo ARMA (1,1), la identidad anterior sería

$$(1 - \phi_1 L)(1 + \psi_1 L + \psi_2 L^2 + \dots) = (1 - \theta_1 L)$$

Para que un modelo ARMA ( $p, q$ ) sea invertible, se requiere que las raíces de la ecuación polinomial caigan fuera del círculo unidad.

$$\theta(L) = 1 - \theta_1 L - \dots - \theta_q L^q = 0$$

- Modelo ARMA(1,1)

Un proceso ARMA (1,1) (excluyendo la constante) se define:

$$Y_i = \phi_1 Y_{i-1} + \varepsilon_i - \theta_1 \varepsilon_{i-1}$$

El proceso ARMA (1,1) es estacionario cuando  $|\phi| < 1$ , e invertible cuando  $|\theta| < 1$ .

Multiplicando ambos miembros por  $Y_{t-k}$  y tomando esperanzas, tenemos:

$$Y_k = E[Y_t Y_{t-k}] = \phi_1 Y_{k-1} + E[\varepsilon_t Y_{t-k}] - \phi_1 E[\varepsilon_{t-1} Y_{t-k}]$$

Teniendo en cuenta que:

$$E[\varepsilon_t Y_t] = \sigma^2$$

$$E[\varepsilon_{t-1} Y_t] = E[\varepsilon_{t-1} (\phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1})] = (\phi_1 - \theta_1) \sigma^2$$

La expresión se deduce a las siguientes expresiones:

$$\text{Para } k=0 \rightarrow Y_0 = \phi_1 Y_1 + \sigma^2 - \theta_1 (\phi_1 - \theta_1) \sigma^2$$

$$\text{Para } k=1 \rightarrow Y_1 = \phi_1 Y_0 - \theta_1 \sigma^2$$

Sustituyendo este valor en la expresión de la varianza ( $Y_0$ ) se tiene: 0

$$Y_0 = \phi_1(\phi_1 Y_0 - \theta_1 \sigma^2) + \sigma^2 - \theta_1(\phi_1 - \theta_1)\sigma^2 =$$

$$\Rightarrow Y_0 = \frac{1 - 2\theta_1\phi_1 + \theta_1^2}{1 - \phi_1^2} \sigma^2$$

Que vuelve a sustituirse en la expresión de la covarianza de primer orden ( $Y_1$ )

$$\text{Para } k > 1 \rightarrow Y_k = \phi_1 Y_{k-1}$$

De esta forma, los coeficientes de autocorrelación quedan como:

$$P_k = \begin{cases} \frac{(1 - \phi_1\theta_1)(\phi_1\theta_1)}{1 - 2\theta_1\phi_1 + \theta_1^2} & \mathbf{k = 1} \\ \phi_1 P_{k-1} & \mathbf{k > 1} \end{cases}$$

Es decir, los coeficientes de autocorrelación de un ARMA(I,q) se comportan como un AR(1) puro para  $k > 1$ .

## b) Modelos lineales no estacionarios

### b.1) Modelos autorregresivo de promedio móvil integrado: ARIMA (p,d,q)

Es un modelo que permite describir un valor como una función lineal de datos anteriores y errores debidos al azar. Se analiza sobre una serie estacionaria.

Los modelos de promedio móvil autorregresivo integrado (ARIMA: Autorregresive integrated moving - average) son una clase especializada de técnicas de filtración que ignoran por completo a las variables independientes en la formulación de pronósticos. Estos modelos son dispositivos altamente refinados de ajuste de curvas que utilizan valores reales y anteriores de la variable dependiente, para producir pronósticos precisos de corto plazo.

En 1970, Box y Jenkins desarrollaron un cuerpo metodológico destinado a identificar, estimar y diagnosticar modelos dinámicos de series temporales en los que la variable tiempo juega un papel fundamental. Podemos decir que la consideración exclusiva de los valores pasados de

una determinada variable para explicar su evolución presente y futura supone, al mismo tiempo una ventaja y un inconveniente:

- La ventaja radica en el hecho de no necesitar distintas series de datos (distintas variables) referida al mismo periodo de tiempo (característica común a todos los modelos univariantes) y, al mismo tiempo, ahorramos la identificación y especificación del modelo en el sentido de la econometría tradicional.

- El inconveniente es que, al renunciar a la inclusión de un conjunto más amplio de variables explicativas, no atendemos a las relaciones que sin duda existen entre casi todas las variables económicas perdiendo capacidad de análisis de tiempo que renunciamos, implícitamente, al estudio teórico previo del fenómeno y a su indudable utilidad.

Los modelos ARIMA (p,d,q) constituyen una clase particular de procesos no estacionarios y se define como:

$$W_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} - \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \dots - \omega_q \varepsilon_{t-q}$$

Donde:

$$W_t = Y_t - Y_{t-1} \quad : \text{Primera diferencia}$$

$$d \quad : \text{Número de diferenciaciones}$$

La mayor parte de las series económicas corresponden a procesos no estacionarios. Así, si se desea obtener un tratamiento de las series basado en el análisis de series de tiempo (modelo ARMA), es necesario discutir mecanismos de transformación de las series a procesos estacionarios.

En principio pueden presentarse distintas (infinitas) formas por las que se introduce la no estacionariedad en un proceso estocástico. Sin embargo, interesa considerar solo algunas formas de la no estacionariedad que sean adecuados para describir el comportamiento de series económicas y, al mismo tiempo, posibles de ser transformados en procesos estacionarios. En primer lugar, analizaremos el proceso de "caminata aleatoria".

- Caminata aleatoria

Es una serie de tiempo estocástica en la que cada cambio sucesivo en  $Y_t$ , expresado como  $u_t$  es extraído en forma independiente de una distribución de probabilidad con media 0 y varianza  $\sigma_2$ . Por lo tanto,  $Y_t$  está determinada por:

$$Y_t = Y_{t-1} + \varepsilon_t$$

Donde  $\varepsilon_t$  es un ruido blanco.

- Caso general de un ARIMA

Dada una serie  $Y_t$ , que eventualmente corresponde a los logaritmos de los valores originales, si su diferencia de orden "d" puede ser representada por un proceso ARMA (p,q) estacionario, se dice que la serie  $Y_t$  sigue un proceso ARIMA(p,d,q).

La letra "I" en ARIMA corresponde a "Integración", la operación inversa a la diferenciación.

Si  $Z_t = \Delta^d Y_t$  y  $Z_t$  sigue un proceso ARMA (p,q) estacionarios:

$$\begin{aligned} (1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) Z_t &= (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) \varepsilon_t \\ \phi_p(L) Z_t &= \theta_q(L) \varepsilon_t \end{aligned}$$

Entonces  $Y_t$  sigue un proceso ARIMA (p,d,q). También se escribe en la variable original  $Y_t$  como:

$$\phi_p(L)(1-L)^d Y_t = \theta_q(L) \varepsilon_t$$

Generalmente, no son necesarias diferencias regulares de orden superior a 2, excepto en el caso de variables que presentan estacionalidad.

- Transformaciones Box Cox

Box y Cox (1964) definieron una transformación instantánea en el sentido de que no están involucrados simultáneamente varios periodos de tiempo de carácter más general que la transformación logarítmica. Esta transformación se define por:

$$Y_t^\lambda = \begin{cases} (Y_t^\lambda - 1) / \lambda & \lambda \neq 0 \\ \text{Ln } Y_t & \lambda = 0 \end{cases}$$

La transformación Box Cox requiere definir el parámetro  $\lambda$  de la transformación.

Cuando el parámetro es  $\lambda = 1$ , la transformación Box Cox consiste prácticamente en tomar logaritmos.

Cuando el parámetro es  $\lambda = 0$ , se define por la segunda igualdad (transformación logarítmica).

La primera igualdad vale también en el límite, el logarítmico de la serie original.

- Modelos ARIMA estacionales

Un modelo estacional puro se caracteriza porque sólo existe relación entre las observaciones que distan entre sí “s” periodos o múltiplos de “s”. Son series con ciclos u oscilaciones estrictamente periódicas, donde el periodo es igual o inferior al año. El periodo estacional se designa por “s”, así en datos trimestrales  $s=4$ , en datos anuales  $s=12$ , etc. La elaboración de modelos ARIMA estacionales presentan características análogas a las de los modelos ARIMA no estacionales.

Los métodos que emplean modelos multiplicativos ARIMA(p,d,q)(P,D,Q) suponen que el componente estacional es generado por un proceso estocástico, cuya identificación se realiza de manera similar a los modelos que representan la estructura regular de una serie, con la excepción de que para ello se examinan los “valores estacionales” de las funciones de autocorrelación (valores que corresponden a los rezagos 4,8,12, ... si los datos son trimestrales y 12, 24, 36, ... si los datos son mensuales). De este modo una serie podría requerir diferencias de orden estacional si los valores estacionales de la función de autocorrelación no tienden a cero rápidamente.

Un proceso ARIMA (p,d,q)(P,D,Q) se define así,

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

Donde:

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

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

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

$$\Theta_q(L^s) = 1 - \Theta_1 L^s - \Theta_2 L^{2s} - \dots - \Theta_q L^{qs}$$

- Intervalos de confianza para las predicciones.

La varianza del error de predicción puede utilizarse para obtener intervalos de confianza de las predicciones elaboradas, mediante la expresión:

$$\Pr[\tilde{y}_{T+k} \pm \lambda_\alpha \bar{\sigma}_{\varepsilon_T(k)}]$$

Donde, si se supone que la innovación  $\varepsilon_t$  sigue una distribución normal, el parámetro  $\lambda_\alpha$  se obtendrá de las tablas de dicha distribución, al nivel de confianza  $\alpha$  elegido.

- Distribuciones teóricas de los coeficientes de autocorrelación para algunos los modelos ARIMA más comunes.

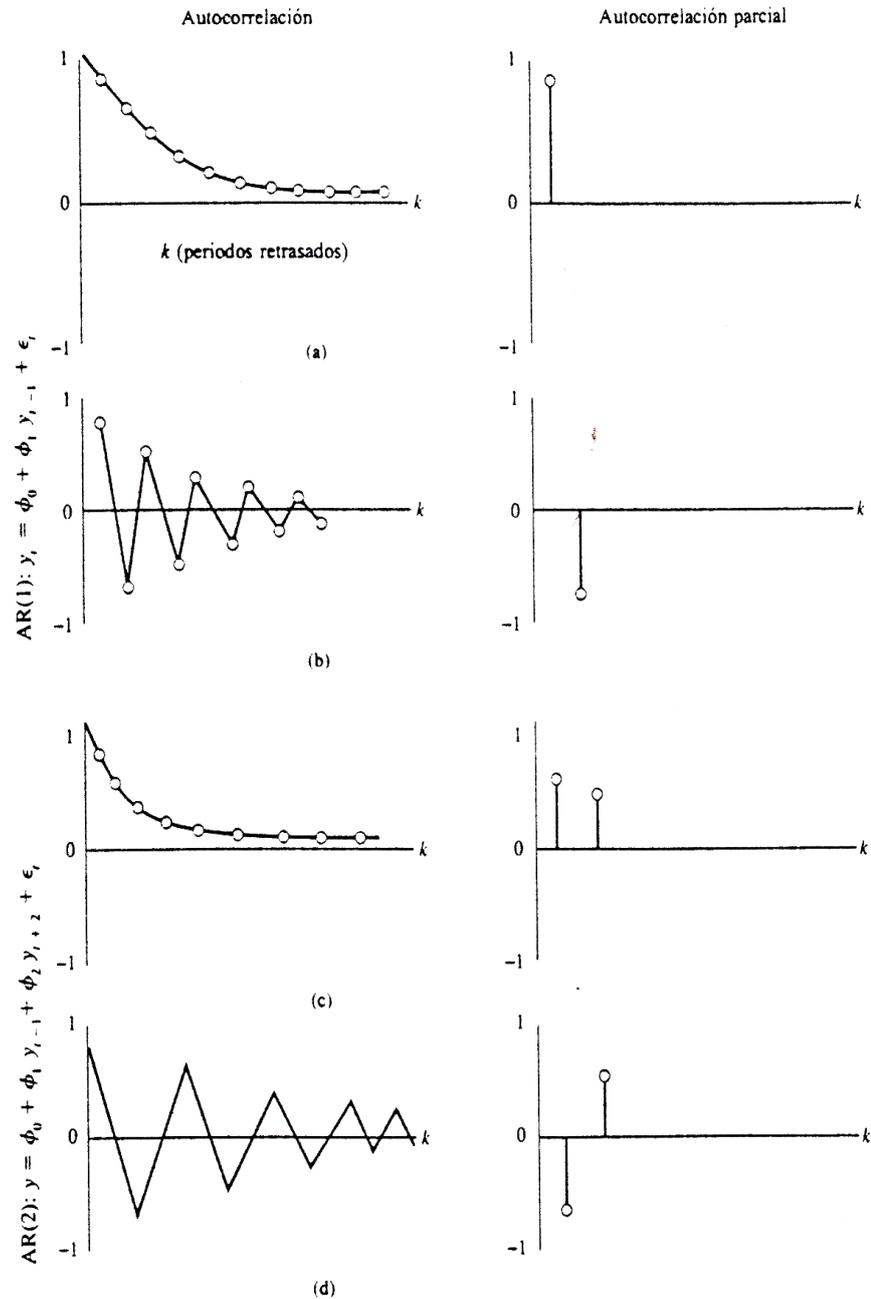


Figura 1. Coeficientes de autocorrelación y de autocorrelación parcial de los modelos AR (1) y AR (2)

Fuente: Hanke (1996)

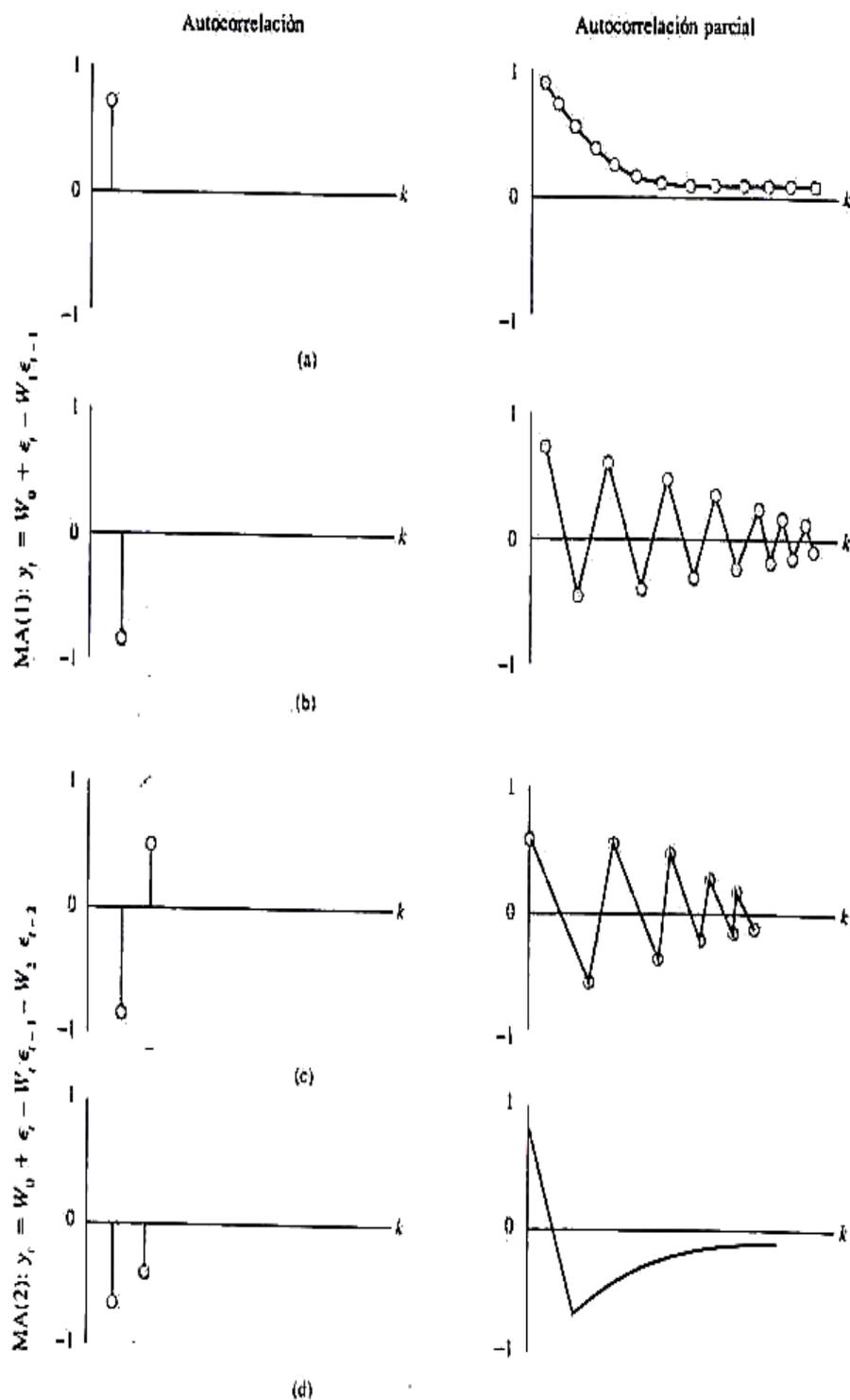


Figura 2. Coeficientes de autocorrelación y de autocorrelación parcial de los Modelos MA (1) Y MA (2)

Fuente: Hanke (1996)

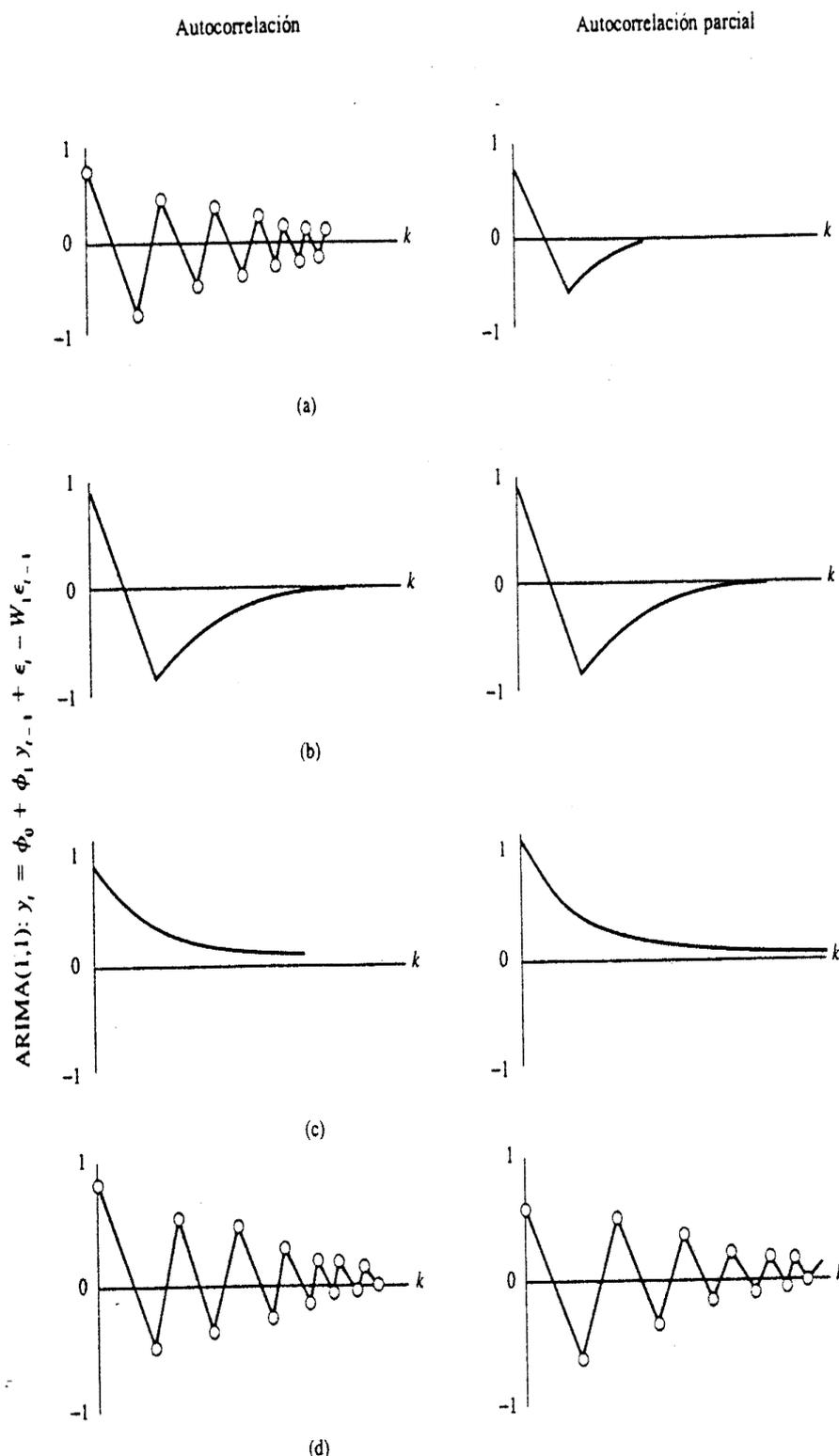


Figura 3. Coeficientes de autocorrelación y de autocorrelación parcial de un modelo ARIMA (1,1).

Fuente: Hanke (1996)

### 1.1.18. Metodología de Box-Jenkins.

El método de Box-Jenkins es uno de los métodos predictivos y se fundamenta en la estimación eficiente de los parámetros por medio de los procesos iterativos.

De acuerdo a Box-Jenkins (1970), el análisis de series de tiempo implica las siguientes etapas: (i) Identificación, (ii) Estimación, (iii) Verificación y, (iv) Pronóstico (Predicción). Si la serie es débilmente estacionaria, se procede de inmediato con la etapa (i); caso contrario, la serie debe ser "pre-procesada" a fin de ser transformada en realizaciones estacionarias. Asumiendo que se cuenta con series estacionarias, la identificación tiene por objeto determinar el tipo de modelo a aplicar (AR, MA ó ARMA) y el orden de los parámetros "p" y "q".

El procedimiento de la elaboración de un modelo ARIMA mediante la metodología Box Jenkins se muestra y explica más adelante en la figura n° 04.

Un aspecto importante en la modelación ARIMA de una serie de tiempo simple es el número de veces que ésta necesita de una diferencia antes de fijar el modelo.

### 1.1.19. Dickey-Fuller aumentado (Test ADF).

Sin duda alguna, el test más habitual a la hora de determinar la estacionariedad de una serie temporal, consiste en la aplicación del conocido como test de Dickey-Fuller (Test DF) o Dickey-Fuller Ampliado (Test ADF). Éste es un contraste de "No estacionariedad" ya que la hipótesis nula es precisamente la presencia de una raíz unitaria en el proceso generador de datos de la serie analizada.

Como modelo de partida para el análisis de una determinada serie  $Y_t$ , el de un proceso estacionario autorregresivo de orden uno:

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t$$

Como hipótesis nula  $H_0$ , el modelo alternativo de un paseo aleatorio no estacionario del tipo 1;  $Y_t = \phi_1 Y_{t-1} + \varepsilon_t$ .

El Test de Dickey-Fuller Ampliado (DFA): contrasta la presencia de una raíz unitaria en una serie que sigue un proceso AR(p), deberá aplicarse el procedimiento expuesto para el caso simple AR(1), pero suponiendo ahora del modelo:

$$\Delta Y_t = a_0 + \gamma Y_{t-1} + \sum_{i=2}^B B_i \Delta Y_{t-i+1} + \varepsilon_t$$

H0:  $\gamma = 0$  raíz unitaria (proceso no estacionario).

H1:  $\gamma < 0$  no existe raíz unitaria (proceso estacionario).

### 1.1.20. Estadístico Box Pierce.

El contraste de "Q" propuesto por Box-Pierce (1970) analiza la hipótesis nula:

$$H_0: \rho_1(a) = \rho_2(a) = \rho_3(a) = \dots = \rho_M(a) = 0$$

Cuya expresión es:

$$Q = T \sum_{k=1}^M \rho_k^2$$

La cual fue refinada a fin de disminuir el sesgo en pequeñas muestras, por Ljung y Box (1978) que propusieron el siguiente estadístico:

$$Q^* = T(T+2) \sum_{k=1}^M (T-k)^{-1} \rho_k^2$$

Donde

M: Número máximo de rezagos a analizar.

T: Número total de observaciones.

R<sub>j</sub>: La función de autocorrelación de los errores del proceso.

Que se distribuye con una  $\chi^2_{M-p-q}$  grados de libertad.

Si  $Q^* < \chi^2_{M-p-q}(\alpha)$  se acepta H<sub>0</sub>

Si  $Q^* > \chi^2_{M-p-q}(\alpha)$  se rechaza H<sub>0</sub>, ó

Si Prob (Q) >  $\alpha$ , se acepta H<sub>0</sub>. Los residuos son ruido blanco.

Si Prob (Q) <  $\alpha$ , se rechaza H<sub>0</sub>. Los residuos *no* son ruido blanco.

## 1.2. Antecedentes.

Huancantara (2010) en su trabajo de investigación tiene como objetivo determinar el modelo uniecuacional que mejor se ajusta para describir y predecir el comportamiento de las variaciones de los niveles medios mensuales de agua del lago Titicaca de la región de Puno. Concluye que los modelos univariantes integrados de Box-Jenkins proporcionan un mejor ajuste para describir y predecir el comportamiento de las variaciones de niveles medios de agua del Lago Titicaca de la región de Puno. El modelo univariante integrado que mejor se ajusta para describir y predecir el comportamiento de la serie de tiempo de niveles medios mensuales de agua del Lago Titicaca, es un ARIMA (1, 1,0) \* (2, 1,1).

Díaz (2008) en su investigación en el cual realiza pronósticos mediante modelos de series de tiempo para el consumo de agua potable de la Empresa Municipal de Saneamiento Básico de la ciudad de Puno, durante el periodo 2000-2007. Tiene como objetivo fue determinar el mejor modelo de series de tiempo para el consumo de agua potable de la Empresa municipal de Puno. Concluye que el mejor modelo para el consumo de agua potable es: ARIMA (0, 1,3).

Guerra, Sánchez, y Reyes (1997) su estudio tiene como objetivo fue proveer una serie de modelos con periodicidad mensual que permitan realizar proyecciones con distintas alternativas propuestas, de tal manera que se genere una visión adecuada sobre la inflación en el futuro inmediato. El periodo analizado corresponde a 1989:06 - 1997:07. Concluye que al ser modelada la variable únicamente en función de su propia evolución pasada, se plantea como supuesto básico que el proceso generador de datos de la inflación se mantiene invariable durante el período de proyección, no tomando en cuenta posibles alteraciones atribuidas a variables relacionadas en el sistema económico, mientras que las ecuaciones de regresión permiten incorporar cambios previsibles en las variables explicativas que afectarán los niveles futuros de la inflación.

Neira (1995) en su investigación tiene como objetivo fue determinar el modelo univariante más eficiente en las Series de Pacientes Hospitalizados por Neumonía y Bronconeumonía. Concluye que los Modelos Integrados de la Serie de Pacientes hospitalizados por Neumonía y Bronconeumonía proporcionan un mejor ajuste con respecto a los modelos No Integrados, y que los Modelos Integrados de la Serie de pacientes Niños menores de 5 años hospitalizados por Neumonía y Bronconeumonía proporcionan un mejor ajuste respecto a los Modelos Integrados.

Guevara, Vega, y Guevara (2010) en su trabajo realizaron un análisis estocástico para una serie cronológica de caudales mensuales del río Caroní (Venezuela) en la estación hidrométrica Guri para el periodo de 1950 a 2003, aplicando la metodología de Box – Jenkins. En la serie de tiempo original se observó periodicidad y una aparente tendencia media y para eliminarla se usó la diferenciación estacional, el modelo identificado fue una ARIMA (1, 0,0) (0,1,1)<sup>12</sup>.

Figueiredo y Blanco (2014) en su investigación realizaron una simulación de los niveles de agua del río Tapajos (Brasil) a través de modelos ARIMA, para los periodos de 1976 al 2008 (de los cuales en los periodos de 1976 a 2004 fueron utilizados para estimar el modelo, del 2005 al 2006 para calibrar el modelo y de 2007 al 2008 para validar el modelo). La identificación del modelo fue realizado a través de la función de autocorrelación y función de autocorrelación parcial. La estimación del modelo fue hecha a través de los criterios de verosimilitud de AIC (Criterio de Información de Akaike), AICC (Criterio de Información de Akaike Corregido) e BIC (Criterio de Información Bayesiano). El modelo de pronóstico escogido fue un SARIMA (1,0,0) (1,1,1).

Paula, Borbosa, y Bras (2015) en su trabajo realizaron un análisis comparativo de modelos de series de tiempo modelar y pronosticar los regímenes de valores medios mensuales del río Doce, Colatina en Espírito Santo (Brasil) a través de la metodología de Box – Jenkins. La serie de datos analizada corresponde a los valores medios mensuales de los periodos de enero de 1939 a diciembre del 2013 (de los cuales en los periodos de enero de 1939 a diciembre del 2012 fueron utilizados para estimar el modelo y del enero a diciembre del 2013 para comparar el modelo), caracterizada por datos faltantes en los periodos de 1978 a 1982 (que fueron estimados). El valor medio mensual fue de 856.90 m<sup>3</sup>/s, con una desviación estándar de 609.71 m<sup>3</sup>/s y un coeficiente de variación alto (indicando estacionalidad). El modelo de pronóstico escogido fue una SARIMA (1,1,1) (2,1,2).

Reyes y Mejia (2010) en su trabajo realizaron un análisis de las periodicidades de los caudales medios mensuales en la cuenca del río Santa. La serie de datos estuvo constituida por los registros históricos de las estaciones: Recreta, Pachacoto, Querecocha, Olleros, Quillcay, Chancos, Llanganuco, Parón, Cedros, Quitaracsa, Colcas, Condorcerro y Puente Carretera, para los periodo de 1956 a 1995. Se caracterizó cada una de las sub-cuencas del Río Santa, encontrándose que los caudales medios mensuales más altos correspondieron a las estaciones: Puente Carretera (151.10 m<sup>3</sup>/s), Condorcerro (146.44

m<sup>3</sup>/s) y La Balsa (87.88 m<sup>3</sup>/s). Se observó adicionalmente que los valores en el coeficiente de variación fueron superiores al 30% indicando estacionalidad en cada una de las estaciones de aforo. El modelo hallado fue un SARIMA (1,0,0) (0,1,1) para los caudales medios mensuales correspondientes a todas las estaciones consideradas, excepto a la estación del Puente Carretera que le correspondió el modelo SARIMA (1,0,1) (0,1,1).

Guevara (2012) en su investigación en el cual crea y aplica una Metodología para solucionar problemas resolubles de manera determinista mediante técnicas sencillas y en un tiempo razonable, como puede ser, por ejemplo la resolución de ecuaciones lineales, la realización de pronósticos basados en la ecuación de la línea, pudiendo acortar el tiempo de resolución, más o menos largo, de una manera aceptable. En esta investigación se presenta una nueva metodología, que utiliza la clase Box-Jenkins, para la predicción de la demanda de llamadas, que efectúan los clientes a los centros de llamadas más conocidos como call-center. Se registró, en el trabajo de campo, los aportes y los modelos de solución de diversos autores: orientación del asesor, aporte de investigadores de la UNMSM, investigaciones desarrolladas en diversas organizaciones y las propuestas del autor. El propósito de la tesis es generar un modelo de consenso con mayor eficiencia que las alternativas existentes y aplicar este modelo en la realidad administrativa. El estudio concluye que la propuesta metodológica para rediseñar procesos de atención al cliente, dentro del contexto de los modelos empresariales, empleando herramientas de serie de tiempos funciona de manera eficiente y reporta excelentes resultados en su aplicación, lo que ha de redundar en la mejora de la eficiencia y competitividad derivado de un uso racional de un recurso escaso clave, el humano.

Arias (2017) en la investigación, se analiza el comportamiento de la demanda por depósitos a plazo fijo en moneda nacional en las Cajas Municipales para el período enero 2004 a febrero 2016. La finalidad es tener un instrumento adicional que contribuya con la programación financiera que realizan dichas instituciones. Para ello se empleó métodos econométricos tradicionales y series de tiempo con el interés de contrastar los resultados, y de esta forma recomendar el método que mejor ayude a estimar esta demanda. Se evaluó la capacidad predictiva mediante criterios de información, evaluación de residuos y error de pronóstico. Luego de realizar las estimaciones y considerando la naturaleza no estacionaria de las series vistas, se concluye que la mejor alternativa es plantear un modelo ARIMA, ya que de considerar otra variable explicativa se obtiene una regresión espuria. El modelo elegido es un modelo SARIMA, debido a que la serie presentó la

estacionalidad de cada seis meses. Los coeficientes resultaron estadísticamente significativos y haciendo una evaluación de los residuos estos fueron aleatorios por lo que se concluye que el modelo elegido es apropiado para realizar pronósticos. El presente trabajo se estructuró de la siguiente forma: en el primer capítulo se presenta el marco teórico y algunos conceptos relevantes, así como la especificación teórica de la función general; en el segundo se caracteriza la evolución de los depósitos a plazo; en el tercero se presentan los resultados de la modelación de la demanda estudiada y en el cuarto las principales conclusiones y recomendaciones del estudio.

Aguilera (2015) en su investigación, tiene como objetivo principal, obtener un modelo adecuado de series de tiempo que describa el comportamiento de la masa mensual del río Jequetepeque en la región La Libertad, utilizando información de la Gerencia Regional de Agricultura desde el periodo de enero de 1944 a diciembre del 2014. La metodología aplicada en el trabajo fue la propuesta por Box-Jenkins y la serie se dividió en: enero del 1944 a diciembre del 2013 para la estimación de modelo y de enero a diciembre del 2014 para la validación del pronóstico. El modelo de pronóstico identificado fue una serie estacionaria estacional y el modelo de pronóstico estimado fue un SARIMA (1, 0, 0) (3, 0, 0)<sub>12</sub>, cuya ecuación es:  $Y_t = 0.5510Y_{t-1} - 0.7641Y_{t-12} - 0.5797Y_{t-24} - 0.3754Y_{t-36}$ , siendo probada la adecuación del modelo. Se pronosticó la masa mensual del Río Jequetepeque de enero a diciembre del 2014 y se evaluó los pronósticos con: la Desviación Absoluta Media (DAM de 26.825 millones de m<sup>3</sup>, el Error Medio Cuadrático (EMC) de 1,786.915, el Porcentaje de Error Medio Absoluto (PEMA) de 0.311% y el Porcentaje Medio del Error (PME) de 0.405%.

Barreras, Sánchez, Figueroa, y Olivas, (2014), en su estudio con el fin de pronosticar la producción de carne de bovino en Baja California, México, se utilizó el método de Box-Jenkins para seleccionar un modelo autorregresivo de promedios móviles (ARMA). Para ello se usó la información de canales de bovinos procesadas mensualmente en los rastros del estado durante el periodo de 2003 a 2010. Debido a que la inspección de la gráfica de la serie y el correlograma de la misma no permitieron establecer la estacionariedad, se aplicó la prueba de Dickey-Fuller aumentada, en la que se encontró que la serie era estacionaria. Como resultado del procedimiento de identificación se seleccionaron los modelos AR (1) y ARMA (2,1), los cuales se estimaron utilizando mínimos cuadrados; se compararon ambos modelos con base en la significancia de sus coeficientes de regresión y los estadísticos de Akaike y Schwartz. Se llevó a cabo una evaluación

diagnóstica para revisar la bondad de ajuste de los modelos mediante la gráfica de los residuales; el valor de los estadísticos  $Q$  se utilizó para determinar la ausencia de autocorrelación en los modelos propuestos. Debido a que los resultados fueron similares, se llevó a cabo una evaluación de la eficiencia predictiva de ambos modelos utilizando una serie de estadísticos. Los resultados de estas pruebas indicaron que el modelo ARMA (2,1) presentaba una mejor capacidad predictiva, que fue corroborada mediante una gráfica conjunta de la serie actual y la pronosticada, y una predicción para el mes de enero de 2011. Los resultados respaldan el uso de modelos ARMA para obtener, a corto plazo, predicciones de la producción de carne en Baja California.

Jaramillo, Gonzáles, Nuñez, Esperanza y Lucio (2006), en sus trabajo de investigación, presentan los resultados de la modelación para la predicción a corto plazo de la concentración de ozono troposférico en la zona urbana de la ciudad de Cali, Colombia, mediante el análisis univariante de series de tiempo. El método se aplicó a una serie de 2.496 datos horarios de concentración de ozono, obtenida en una de las estaciones de la Red de Monitoreo de Calidad del Aire (RMCA) de Cali, correspondiente a 104 días consecutivos durante el período abril a julio de 2003. Los datos de los primeros 93 días se utilizaron para la estimación del modelo, y los de los 11 días restantes para la validación del mismo. El modelo propuesto podría ser utilizado por las autoridades ambientales de la región para predecir y alertar a la población sobre posibles episodios de altas concentraciones de ozono que puedan poner en riesgo la salud pública con hasta 8 horas de anticipación.

Cuitiño, Ganón, Tiscordio, y Vicente (2010), en su trabajo de investigación, muestra las proyecciones de inflación ocupan un rol central en el diseño de la política monetaria, especialmente en instituciones que, como el Banco Central del Uruguay (BCU), adoptan en forma progresiva un esquema de metas de inflación. Este documento constituye el primer punto de una larga agenda de investigación en el campo de las predicciones de inflación, en el que se evalúa una batería de modelos univariados de series de tiempo del Índice de Precios al Consumo (IPC) y sus componentes, en función de su poder predictivo a distintos pasos, con foco en el corto plazo. Trabajando con una muestra total que abarca 1997.03-2009.10 y una muestra truncada a partir de 2003.01, se construyen modelos de predicción directa e indirecta (agregación por componentes) para el IPC nivel general y para un indicador de exclusión, y se realizan combinaciones lineales de modelos. Luego se evalúan los errores de predicción en el

horizonte 2009.11-2010.07, comparando estos modelos entre sí y con respecto a dos benchmarks: un modelo naif, constituido por el paseo al azar; y la mediana de la encuesta de expectativas de inflación que publica el BCU. Los principales resultados muestran que en predicciones un paso en adelante (al momento el único horizonte comparable), el mejor modelo de proyección es el modelo directo con la muestra truncada (M2), dominando a la proyección indirecta y a los dos benchmarks usados. En el análisis por componentes se encuentra que varios modelos muestran un buen desempeño, como los de transables y no transables de exclusión, mientras que las principales fuentes de error de la proyección indirecta son el modelo de frutas y verduras y la estimación con juicio de experto de los rubros administrados. Asimismo, los modelos para el indicador de exclusión son muy buenos, tanto los directos como los indirectos. Finalmente, la predicción trimestral usando la predicción mensual a 3, 2 o 1 paso en adelante resulta superior a la predicción a un paso del modelo estructural trimestral, confirmando la importancia de incluir estos modelos univariados en el marco general de predicciones que habitualmente se realizan en la institución. En el estado actual del trabajo, se sugiere usar el modelo M2 para el nivel general y sus componentes.

Mogni (2013) en su trabajo de investigación, tiene como objetivo pronosticar la demanda de pasajeros, ya que este tipo de información permite tomar acertadamente decisiones tales como aumento y/o disminución de frecuencias, cambios de equipos, inversión en equipos nuevos, aumento y/o disminución de tarifas, entre otros. Para ello, abordó un tipo de modelo de pronóstico elaborado por Box y Jenkins denominado ARMA/ARIMA. Llega a la conclusión que el mejor modelo que se ajusta para pronosticar es un ARIMA  $(1, 0, 0) * (0, 0, 1)$ .

Vázquez (2017) en su investigación tuvo como objetivo, estimar un modelo mediante la metodología de Box y Jenkins y Redes Neuronales Artificiales que permita pronosticar el precio del dólar del Sistema Bancario en moneda nacional año 2015. La investigación es de tipo descriptiva, predictiva-longitudinal. Contando con un periodo de 11 años desde 2004 hasta 2014. Las nociones básicas de análisis de Series de Tiempo y Redes Neuronales Artificiales se exponen brevemente en el desarrollo de la tesis. El modelo para el pronóstico que se obtuvo mediante la metodología de Box y Jenkins es un ARIMA  $(2, 2, 3)$ , con un RMSE igual 0.983 para la compra con un MAE de 0.008 y la venta un ARIMA  $(3, 1, 2)$  con un RMSE de 0.982y con un MAE de 0.008 indicando un buen

coeficiente. Por lo tanto, el modelo de Redes Neuronales Artificiales tiene una estructura de (12:1:5:1) con RMSE 0.042 para la compra y para la venta con estructura de (12:1:6:1) con RMSE 0.038.

## CAPÍTULO II

### PLANTEAMIENTO DEL PROBLEMA

#### 2.1 Identificación del problema.

El lago Titicaca es el segundo lago más grande de Sudamérica, el lago navegable más alto del mundo y el mayor lago de agua dulce de América Latina. Se extiende sobre territorios de Perú y Bolivia, y alberga islas flotantes que son visitadas por los turistas. Es el principal lago más importante del país debido a su tamaño, diversidad biológica e importancia socioeconómica. El nivel medio de agua es de 3810 metros sobre el nivel del mar.

La importancia de saber el comportamiento de las variaciones de los niveles de agua del lago Titicaca derivado en un futuro es muy importante para la prevención y planificación ante ciertos incidentes que puedan surgir producto de cambios climáticos.

Las técnicas de pronóstico de series de tiempo permiten llevar a cabo esta labor por medio de la metodología Box-Jenkins, la cual permite obtener buenas aproximaciones en el caso de que dicho método sea bien aplicado, ya que esta metodología tiene una parte de arte en el sentido de que el proceso es iterativo hasta alcanzar el mejor modelo.

En el trabajo de investigación se realiza el análisis de Series de Tiempo siguiendo la metodología de Box-Jenkins, utilizando información de los niveles medios de agua del lago Titicaca por meses y se busca un modelo adecuado que permita ajustar el comportamiento de los niveles medios de agua, el cual a su vez permite realizar pronósticos.

La información obtenida de los niveles medios mensuales de agua del lago Titicaca de la región de Puno corresponde al período 1974-2018, los cuales fueron agrupados

mensualmente, para evaluar el comportamiento de la serie histórica a fin de realizar proyecciones con la variable en estudio.

El lago Titicaca es un ecosistema de gran importancia tanto biológica como ecológica que juega un papel importante para los seres vivos que habitan sus cercanías.

Los ríos más importantes que desembocan en el lago, son entre ellos, el río Ramis, Coata, Ilave, Huancané, Suches, Zapatilla y Callacame, cuentan con un régimen irregular, bajos caudales en el periodo de Abril a Noviembre y con crecidas en el periodo de lluvias entre Diciembre y Marzo. En éste último periodo las superficies aledañas al lago Titicaca y a la desembocadura de los ríos, son muy sensibles al incremento de los caudales de los ríos, poniendo en peligro a los pobladores que se posicionan en estos lugares durante la época de estiaje y bajos caudales.

Según estudios realizados el lago es alimentado por los aportes de sus ríos afluentes (44%) y por las lluvias que caen directamente en su superficie (55%). Las pérdidas se deben principalmente a la evaporación (94%) y al desagüe superficial que sale por el río Desaguadero (5%).

Los niveles medios de agua del lago varían de manera irregular durante el transcurso de los años, existiendo largos periodos de sequías que ponen en riesgo la biodiversidad acuática, y algunos períodos de excesos que producen alteraciones importantes, inundaciones en algunas zonas o se constituyen en regeneradores de algunos sistemas degradados (lago Poopó, corrección de la salinidad, etc.).

Durante los últimos años, debido a las condiciones atmosféricas que se vienen presentando en la zona sur del territorio peruano, la falta de precipitaciones motivó que los ríos tributarios del Titicaca no aumenten su caudal y, por consiguiente, no aporten en la cota del lago; además la insolación incidió sobre el espejo de agua provocando evaporación constante. Sin embargo, desde hace muchos años han existido periodos de sequias e inundaciones que se han ido presentando alternadamente en la región del altiplano. Por ello, se necesita conocer con mayor precisión la variabilidad en el comportamiento de las variaciones de los niveles medios mensuales del lago Titicaca para que con estrategias de contingencia, los organismos de prevención local, regional y nacional puedan prevenir problemas que se puedan presentar a futuro y de esa manera prestar mejores servicios a los pobladores del departamento de Puno.

En la actualidad la predicción de series de tiempo ha adquirido gran atención dentro de las investigaciones empíricas, tanto que los estadistas y econométristas han empezado a dedicar mayores esfuerzos a este tipo de información.

En este sentido apoyándome en las herramientas estadísticas, en el análisis de series de tiempo, usando la metodología Box Jenkins, surgió la necesidad de conocer con mayor precisión la variabilidad en el comportamiento de los niveles medios de agua del lago Titicaca, teniendo como base la información de los niveles medios mensuales de agua del lago Titicaca a través del tiempo, lo cual nos permite tomar decisiones a futuro y realizar pronósticos acerca de dicha información.

## **2.2 Enunciados del problema.**

### **2.2.1. Interrogante del problema general.**

¿Cuál es el modelo univariante que mejor se ajusta a la serie histórica de los niveles medios de agua del lago Titicaca para describir y pronosticar el comportamiento de las variaciones de los niveles medios mensual de agua del lago Titicaca de la región de Puno, periodo 1974-2018?

### **2.2.2. Interrogante de los problemas específicos.**

- ✓ ¿Cuál es el modelo identificado que mejor se ajusta a la serie histórica para describir y pronosticar el comportamiento de las variaciones de los niveles medios mensual de agua del lago Titicaca de la región Puno, periodo 1974 – 2018?
- ✓ ¿Cuáles son los pronósticos alcanzados con el modelo adecuado para los niveles medios mensuales de agua del lago Titicaca de la región Puno, periodo 1974 – 2018?

## **2.3 Justificación.**

Una de las técnicas estadísticas tradicionales empleadas en el pronóstico de series de tiempo es la metodología Box-Jenkins, la cual tiene gran importancia en la toma de decisiones y permite obtener buenas aproximaciones cuando se aplica a series de tiempo estacionarias o que pueden convertirse en estacionarias mediante las transformaciones adecuadas (diferencia o Box-Cox).

El presente trabajo de investigación denominado “Modelo univariante para describir y predecir el comportamiento de los niveles medios de agua del lago Titicaca, periodo 1974 – 2018, Puno”, se realizó para monitorear el comportamiento de los niveles medios de agua del lago Titicaca con modelos de pronóstico que sigan la metodología Box-Jenkins, conocida también como técnica del modelo ARIMA, con el cual se describe el comportamiento de la serie y se realiza la formulación de pronósticos.

Además la presente investigación tiene como propósito mostrar la capacidad de realizar pronósticos en series de Tiempo siguiendo la metodología Box-Jenkins y de ese modo conocer el comportamiento de las variaciones de los niveles medios de agua del lago Titicaca, para que los organismos de prevención local, regional, nacional y público en general puedan tomar las medidas respectivas frente a cualquier problema.

## **2.4 Objetivos.**

### **2.4.1. Objetivo General.**

Determinar el modelo univariante que mejor se ajusta a la serie histórica niveles medios de agua mensual del lago Titicaca, para describir y pronosticar el comportamiento de las variaciones de los niveles medios mensual de agua del lago Titicaca de la región de Puno, periodo 1974 - 2018.

### **2.4.2. Objetivos Específicos.**

- ✓ Identificar, estimar y validar el modelo que mejor se ajusta a la serie histórica para describir y pronosticar el comportamiento de las variaciones de los niveles medios de agua mensual del lago Titicaca de la región Puno, periodo 1974 - 2018.
- ✓ Determinar los pronósticos con el modelo adecuado para los niveles medios de agua mensual del lago Titicaca de la región Puno, periodo 1974 - 2018.

## **2.5 Hipótesis.**

### **2.5.1. Hipótesis General.**

- ✓ El modelo univariante que mejor se ajusta a la serie histórica niveles medios de agua mensual del lago Titicaca para describir y pronosticar es un modelo ARIMA multiplicativo  $(p, d, q)*(P, D, Q)$ .

### 2.5.2. Hipótesis específica.

- ✓ El modelo identificado mediante la metodología de Box-Jenkins, se ajusta a la serie histórica para describir y pronosticar el comportamiento de las variaciones de los niveles medios mensuales de agua del lago Titicaca (sobre los 3810 m.s.n.m.) de la región Puno, periodo 1974 - 2018.
- ✓ Los pronósticos alcanzados para los niveles medios de agua mensual del lago Titicaca (sobre los 3810 m.s.n.m.) de la región Puno, se obtiene con el modelo identificado de Box-Jenkins.

## CAPÍTULO III

### MATERIALES Y MÉTODOS

#### **3.1. Lugar de estudio.**

El presente estudio se llevó a cabo en el departamento de Puno.

#### **3.2. Población.**

La población de estudio está constituida por la información correspondiente a los niveles medios de agua mensual del Lago Titicaca, desde que se empezó el registro de datos hasta junio de 2018.

#### **3.3. Muestra.**

Se utilizó una muestra no probabilística a criterio del investigador, por tratarse de una información temporal o de corte longitudinal y cuyo periodo de estudio comprendió desde 1974 hasta junio de 2018, cabe resaltar que los datos fueron obtenidos del SENAMHI y se trabajó específicamente con los datos registrados en la estación Huatajata.

#### **3.4. Método de investigación.**

El método cuantitativo es el utilizado en la presente investigación el cual parte de una idea que va acotándose y una vez delimitada, se derivan objetivos y preguntas de investigación, se revisa la literatura y se construye un marco o una perspectiva teórica. De las preguntas se establecen hipótesis y determinan variables; se traza un plan para probarlas (diseño); se miden las variables en un determinado contexto; se analizan las mediciones obtenidas utilizando métodos estadísticos, y se extrae una serie de conclusiones respecto de la o las hipótesis. (Hernandez, 2014).

### 3.5. Metodología econométrica.

La metodología que seguiremos es la propuesta por Box-Jenkins, que consta de cuatro etapas: Identificación, Estimación, Validación y Pronóstico.

La etapa de identificación reside en optar uno o más modelos ARIMA, como candidatos que representan apropiadamente el comportamiento de la serie. En ésta etapa debe determinarse las transformaciones necesarias para obtener estacionariedad, contraste de inclusión de un término de tendencia determinística ( $\theta_0$ ) y elegir los órdenes p y q para cada uno de los modelos finalistas.

La etapa de estimación consta en estimar los parámetros de cada uno de los modelos identificados en la fase anterior.

La etapa de validación consiste en determinar si el modelo identificado y estimado es adecuado para representar a los datos. Las insuficiencias encontradas en ésta etapa pueden utilizarse cómo información para reformular el modelo.

La etapa de pronóstico, ya con el modelo validado o adecuado, se realizar pronósticos.

La figura 4, muestra las fases de elaboración de un modelo ARIMA.

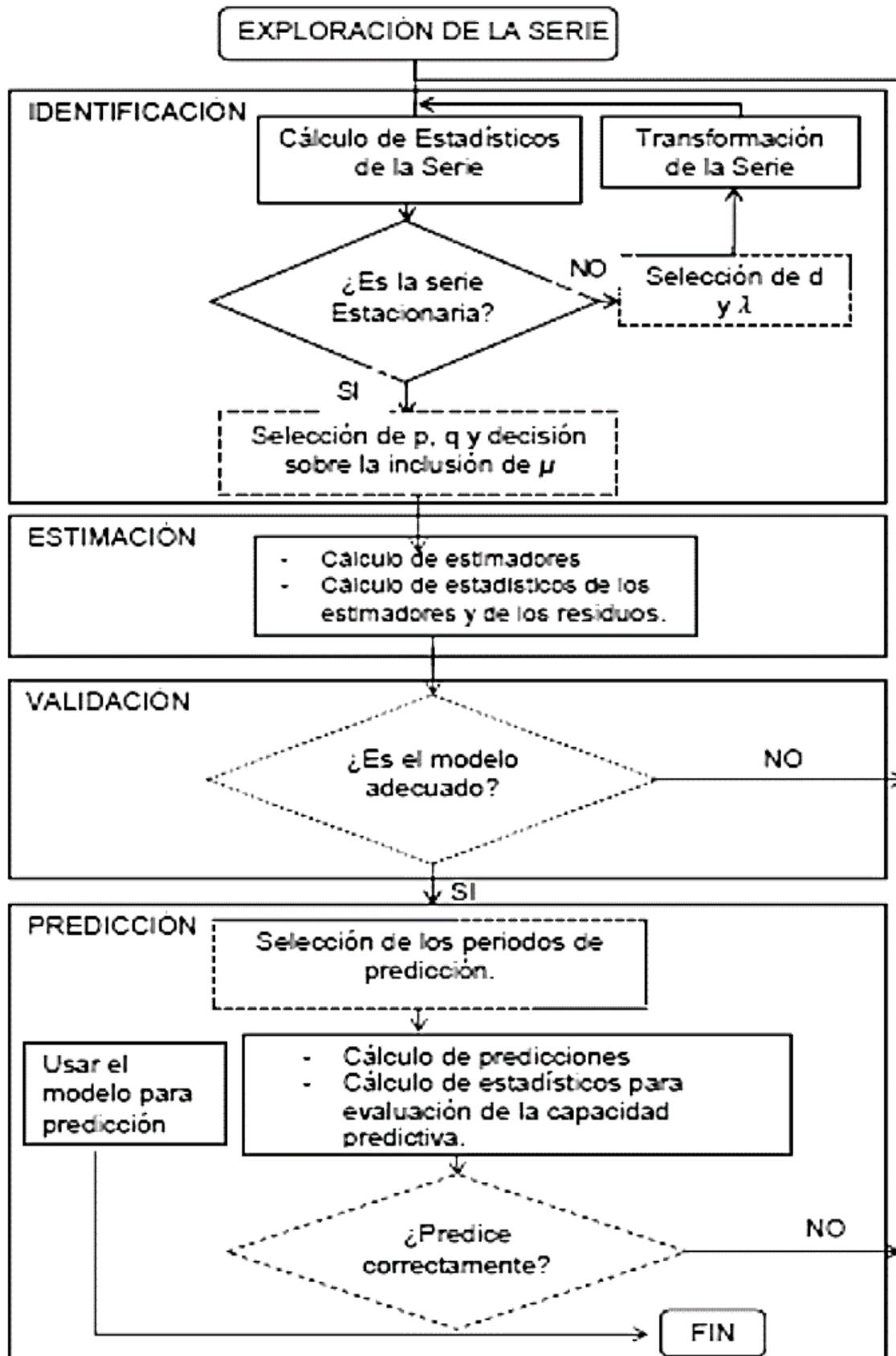


Figura 4. Fases de elaboración de un modelo ARIMA.

Fuente: Uriel (1985). Análisis de series temporales: Modelos ARIMA.

## **CAPÍTULO IV**

### **RESULTADOS Y DISCUSIÓN**

En este capítulo se expone los resultados obtenidos de la información de la serie histórica del nivel promedio mensual del lago Titicaca en la estación Huatajata – Puno (Metros Sobre el Nivel del Mar), durante el periodo 1974-2018.

Los resultados de esta investigación se obtuvieron mediante el software STATGRAPHICS v5.1.

Tabla 2

*Nivel promedio mensual del lago Titicaca en la estación Huatajata por mes según año (Metros Sobre el Nivel del Mar) 1974-2018.*

AÑO	Enc.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Ago.	Sep.	Oct.	Nov.	Dic.
1974	3808.1	3808.7	3809.1	3809.2	3809.2	3809.1	3809.0	3808.9	3808.9	3808.8	3808.7	3808.6
1975	3808.7	3809.1	3809.4	3809.5	3809.5	3809.4	3809.3	3809.2	3809.1	3809.1	3809.0	3809.0
1976	3809.3	3809.6	3809.8	3809.9	3809.8	3809.7	3809.6	3809.5	3809.4	3809.4	3809.2	3809.1
1977	3809.1	3809.1	3809.5	3809.6	3809.4	3809.4	3809.3	3809.2	3809.1	3809.0	3809.0	3809.0
1978	3809.1	3809.6	3809.8	3809.9	3809.8	3809.6	3809.5	3809.4	3809.3	3809.2	3809.1	3809.2
1979	3809.5	3809.9	3810.0	3810.1	3810.1	3809.9	3809.8	3809.7	3809.5	3809.5	3809.4	3809.3
1980	3809.4	3809.5	3809.6	3809.7	3809.6	3809.5	3809.4	3809.3	3809.2	3809.1	3809.1	3809.0
1981	3809.1	3809.3	3809.7	3809.9	3809.9	3809.8	3809.6	3809.5	3809.4	3809.3	3809.3	3809.2
1982	3809.5	3809.7	3809.8	3809.9	3809.8	3809.6	3809.5	3809.4	3809.3	3809.2	3809.2	3809.2
1983	3809.1	3809.1	3809.1	3809.0	3809.0	3808.8	3808.7	3808.5	3808.4	3808.3	3808.3	3808.3
1984	3808.4	3809.0	3809.6	3809.9	3809.8	3809.7	3809.5	3809.4	3809.3	3809.2	3809.2	3809.2
1985	3809.4	3809.6	3810.0	3810.2	3810.3	3810.2	3810.1	3809.9	3809.8	3809.7	3809.7	3809.9
1986	3810.3	3810.5	3811.0	3811.3	3811.2	3811.0	3810.9	3810.7	3810.6	3810.4	3810.3	3810.2
1987	3810.5	3810.7	3810.7	3810.6	3810.4	3810.3	3810.1	3810.0	3809.9	3809.8	3809.7	3809.7
1988	3809.7	3809.9	3810.0	3810.4	3810.4	3810.3	3810.1	3810.0	3809.8	3809.7	3809.6	3809.5
1989	3809.5	3809.7	3809.8	3809.9	3809.8	3809.7	3809.6	3809.5	3809.3	3809.2	3809.1	3809.0
1990	3809.1	3809.2	3809.1	3809.1	3809.0	3808.9	3808.8	3808.7	3808.6	3808.5	3808.5	3808.6
1991	3808.7	3808.7	3808.8	3808.9	3808.9	3808.8	3808.7	3808.6	3808.5	3808.3	3808.3	3808.2
1992	3808.3	3808.4	3808.5	3808.4	3808.3	3808.1	3807.9	3807.8	3807.8	3807.7	3807.7	3807.7
1993	3807.8	3808.0	3808.1	3808.2	3808.2	3808.1	3808.0	3807.9	3807.8	3807.7	3807.7	3807.8
1994	3807.9	3808.2	3808.4	3808.5	3808.5	3808.4	3808.3	3808.2	3808.1	3808.0	3807.9	3807.8
1995	3807.9	3808.0	3808.1	3808.2	3808.1	3808.0	3807.9	3807.8	3807.7	3807.6	3807.5	3807.4
1996	3807.5	3807.8	3807.8	3807.9	3807.8	3807.7	3807.5	3807.5	3807.4	3807.5	3807.5	3807.5
1997	3807.7	3808.1	3808.6	3808.8	3808.8	3808.6	3808.5	3808.4	3808.3	3808.3	3808.3	3808.2
1998	3808.2	3808.3	3808.4	3808.5	3808.4	3808.3	3808.1	3808.0	3807.9	3807.7	3807.7	3807.6
1999	3807.6	3807.7	3808.0	3808.4	3808.4	3808.3	3808.2	3808.0	3807.9	3807.9	3807.9	3807.8
2000	3807.8	3808.1	3808.5	3808.5	3808.4	3808.3	3808.1	3808.0	3808.0	3807.9	3807.9	3807.8
2001	3808.1	3808.7	3809.4	3809.8	3809.7	3809.6	3809.5	3809.3	3809.2	3809.2	3809.1	3809.0
2002	3809.0	3809.2	3809.6	3809.9	3810.0	3809.8	3809.7	3809.7	3809.6	3809.4	3809.4	3809.4
2003	3809.5	3809.8	3810.2	3810.3	3810.3	3810.2	3810.1	3810.0	3809.9	3809.8	3809.6	3809.5
2004	3809.8	3810.2	3810.3	3810.4	3810.3	3810.1	3810.0	3809.9	3809.8	3809.6	3809.5	3809.4
2005	3809.5	3809.6	3809.8	3809.8	3809.7	3809.5	3809.3	3809.3	3809.1	3809.1	3809.1	3809.0
2006	3809.8	3810.1	3810.3	3810.3	3810.1	3810.1	3810.0	3809.9	3809.8	3809.7	3809.6	3809.6
2007	3809.7	3809.7	3809.9	3810.1	3810.1	3809.9	3809.8	3809.7	3809.7	3809.6	3809.5	3809.5
2008	3809.5	3809.8	3809.9	3809.9	3809.8	3809.7	3809.5	3809.4	3809.3	3809.2	3809.1	3809.0
2009	3809.1	3809.2	3809.3	3809.3	3809.2	3809.1	3809.0	3808.8	3808.7	3808.6	3808.5	3808.6
2010	3808.7	3809.0	3809.3	3809.3	3809.2	3809.1	3809.0	3808.9	3808.8	3808.6	3808.5	3808.5
2011	3808.5	3808.7	3809.2	3809.3	3809.2	3809.1	3809.0	3808.9	3808.8	3808.7	3808.6	3808.6
2012	3808.8	3809.2	3809.7	3809.9	3809.9	3809.8	3809.7	3809.5	3809.4	3809.3	3809.2	3809.2
2013	3809.4	3809.8	3810.1	3810.1	3810.0	3809.9	3809.8	3809.7	3809.6	3809.5	3809.4	3809.4
2014	3809.6	3809.8	3809.9	3809.9	3809.8	3809.7	3809.6	3809.4	3809.4	3809.4	3809.3	3809.2
2015	3809.3	3809.5	3809.6	3809.7	3809.8	3809.7	3809.6	3809.4	3809.4	3809.3	3809.2	3809.1
2016	3809.2	3809.3	3809.4	3809.4	3809.3	3809.2	3809.1	3809.0	3808.8	3808.7	3808.6	3808.5
2017	3808.6	3808.7	3808.8	3809.0	3809.0	3808.9	3808.8	3808.7	3808.6	3808.6	3808.5	3808.4
2018	3808.5	3808.7	3809.1	3809.2	3809.1	3809.0						

Fuente: Servicio Nacional De Meteorología e Hidrología

Instituto Nacional de Estadística

msnm: Metros sobre el nivel del mar

Nota: El nivel de alerta del Lago Titicaca es a 3.812,057 msnm.

#### 4.1. Identificación del modelo.

A continuación, en la figura 5 presentamos la serie histórica del nivel promedio mensual del lago Titicaca en la estación Huatajata – Puno (Metros Sobre el Nivel del Mar), durante el periodo 1974-2018.

La serie histórica presenta un comportamiento irregular con tendencia creciente en ciertos periodos de tiempo y decreciente en otros, también presenta dos periodos largos de 22 años y 3 meses cada uno, así como también periodos cortos. La serie histórica muestra una varianza pronunciada entre los dos periodos de tiempo largos además dentro de cada periodo largo se muestra variaciones casi constantes lo que indica que estamos ante una serie con periodos estacionales y no estacionales, es decir combinada. En términos generales la serie histórica no presenta tendencia pronunciada o significativa, por tanto estamos ante una serie no estacionaria y por lo que debe ser diferenciada.

En la figura 6 los coeficientes de autocorrelación difieren de cero por varios periodos de desfase en el tiempo, es decir son significativos ya que sobrepasan el límite superior de autocorrelación (del 1 al 26), pero a partir del coeficiente 27 estos tienden a cero. La serie también presenta ondas regulares entre periodos de tiempo, esto indica que estamos ante una serie no estacionaria lo que corrobora a lo descrito en la figura 1.

En la figura 7 muestra la presencia de coeficientes de autocorrelación significativos en alternancia de signos y son los que sobrepasan los límites de la función de autocorrelación, es decir no tienden a cero, lo que confirma la presencia de tendencia de la serie en periodos.

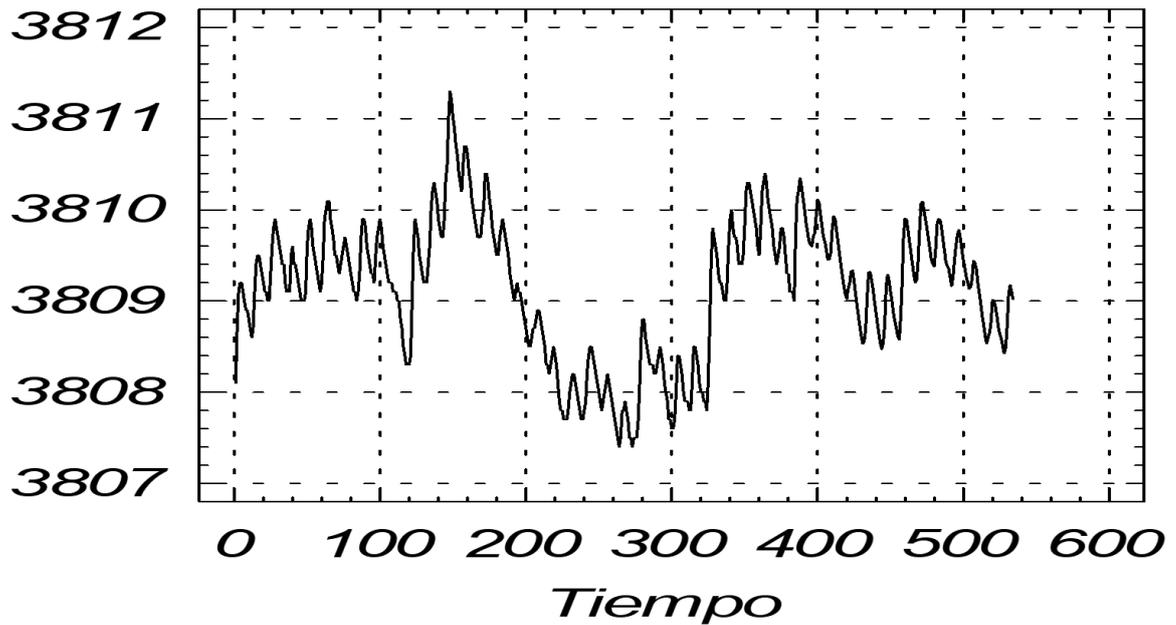


Figura 5. Secuencia de tiempo del nivel promedio mensual del lago Titicaca.

La primera diferencia no estacional de la figura 8 no muestra tendencia, tiene la forma de un ruido blanco, tampoco presenta patrón alguno, pero se perdió una observación por la diferenciación no estacional.

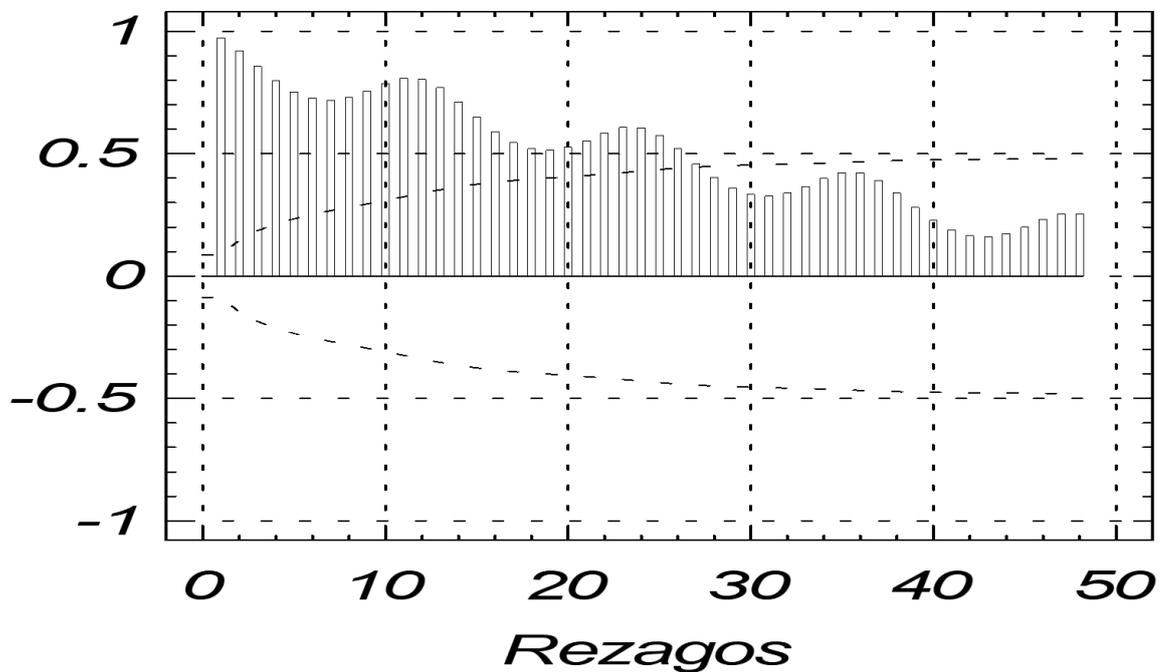
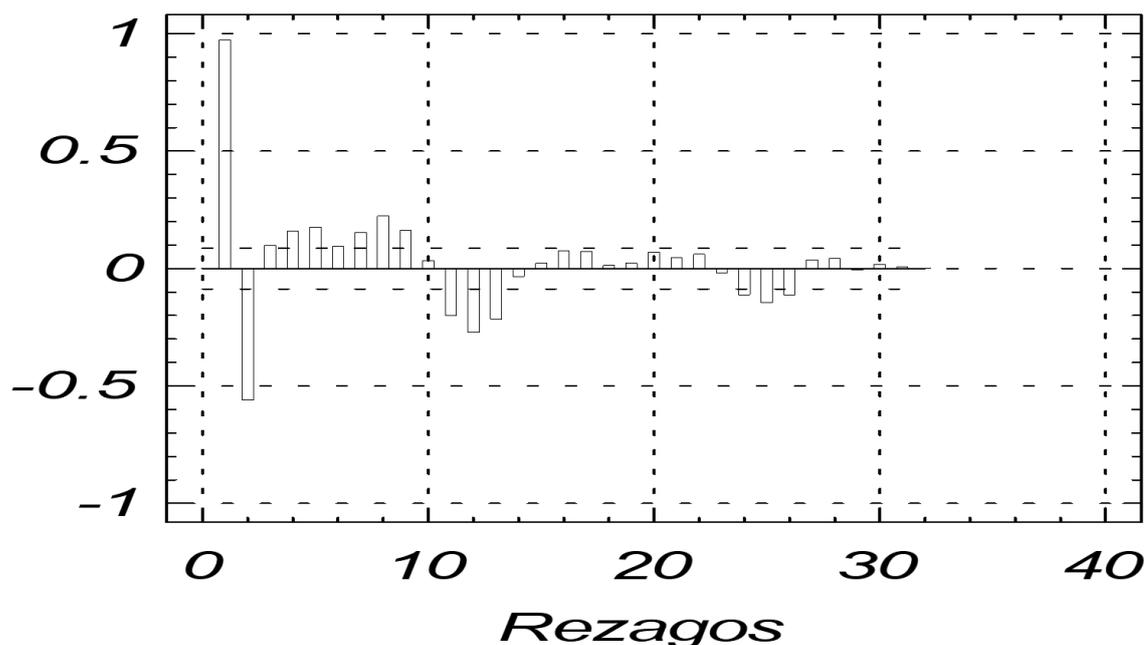


Figura 6. Autocorrelaciones estimadas del nivel promedio mensual del lago Titicaca.



*Figura 7.* Autocorrelaciones parciales estimadas del nivel promedio mensual del lago Titicaca.

Las autocorrelaciones estimadas de la primera diferencia no estacional de la figura 9, presentan coeficientes significativos alternos en periodos regulares, aproximadamente normales y con alternancia de signos en los coeficientes, esta última indica que la serie no es estacionaria. Esta función de autocorrelación estimada presenta un promedio de cero.

Las autocorrelaciones parciales estimadas para la primera diferencia no estacional en la figura 10, siguen mostrando coeficientes significativos en alternancia de signos, pero la forma como se presenta esta figura con el primer coeficiente significativo y positivo nos da la idea de un modelo autorregresivo de orden uno o AR (1) para la parte no estacional del modelo que se identificará.

En la figura 11 se observa que la nueva serie que es la primera diferencia estacional no muestra tendencia, además tiene la forma de ruido blanco con media cero y varianza significativa en periodos aproximadamente largos. En esta serie se perdió 12 observaciones ya que se trata de una estacionalidad de 12 periodos.

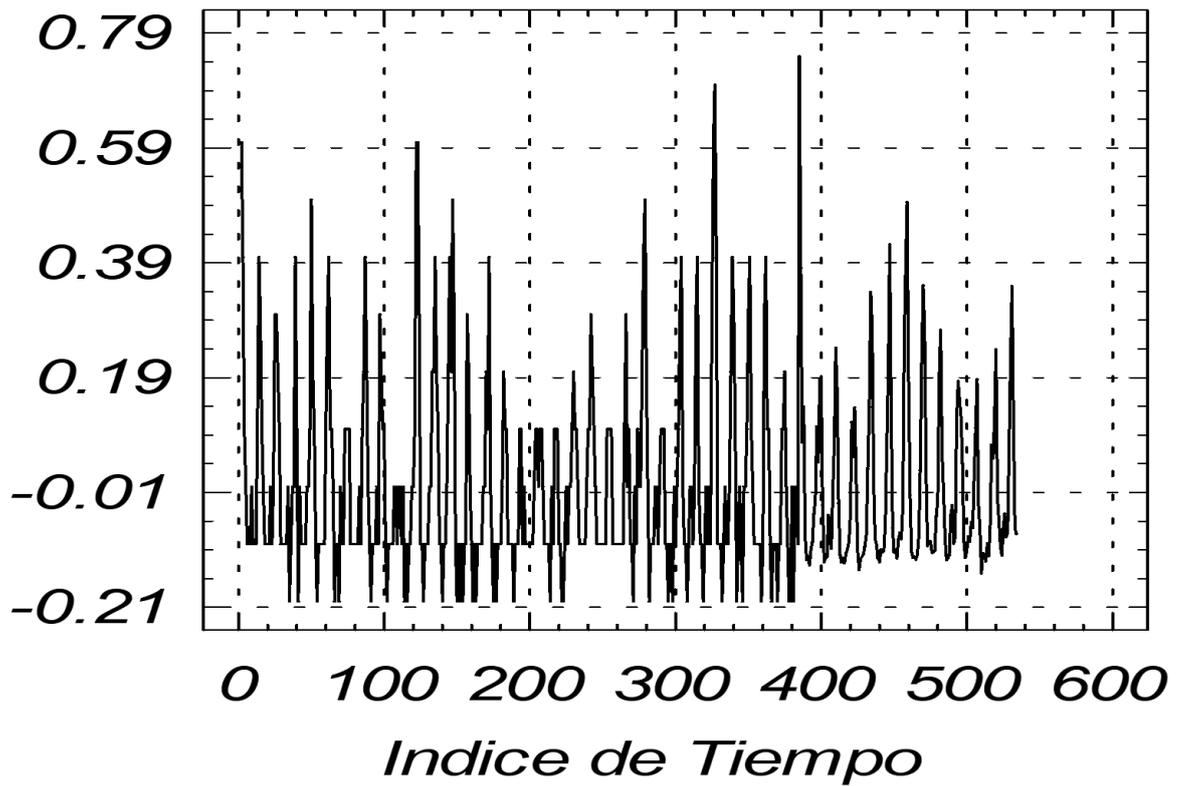


Figura 8. Primera diferencia no estacional del nivel promedio mensual del lago Titicaca.

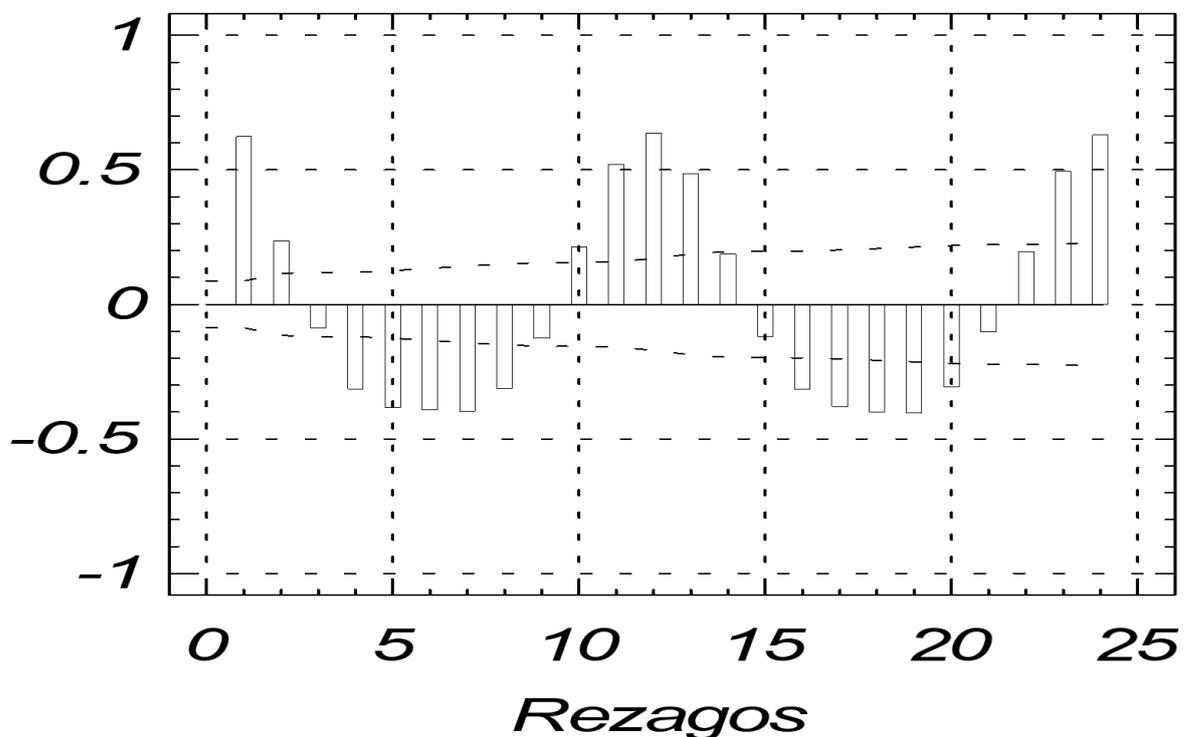


Figura 9. Autocorrelaciones estimadas para la primera diferencia no estacional del nivel promedio mensual del lago Titicaca.

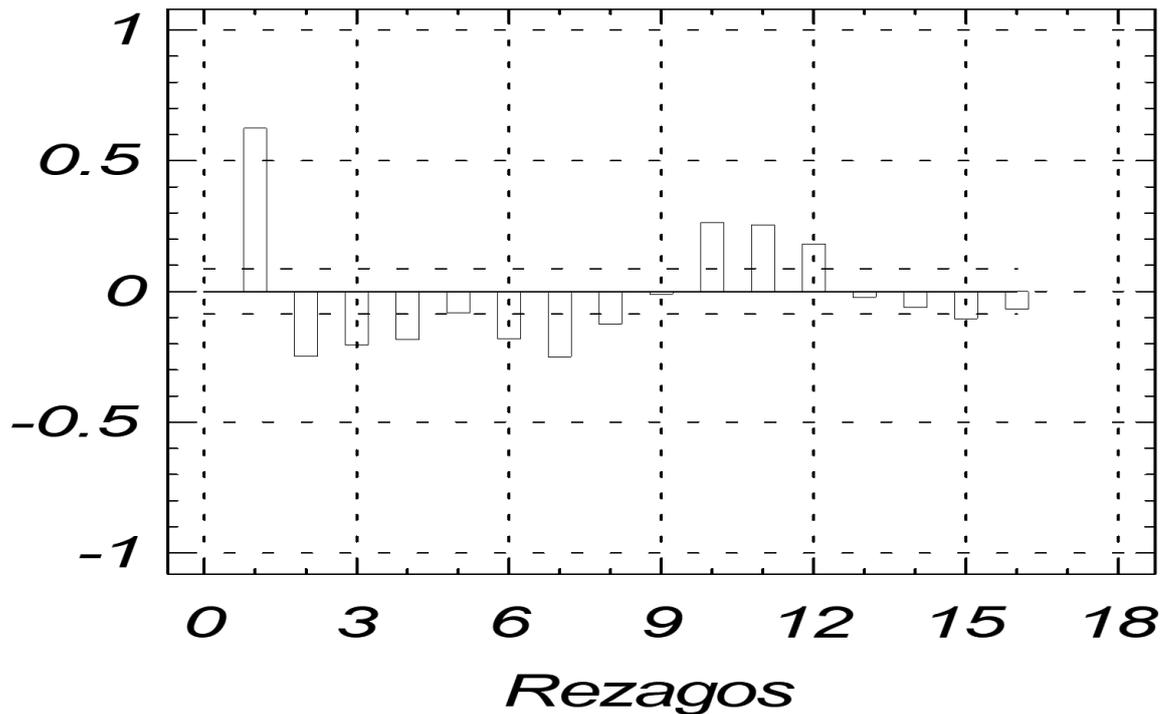


Figura 10. Autocorrelaciones parcial estimada para la primera diferencia no estacional del nivel promedio mensual del lago Titicaca.

Este correlograma de la figura 12 muestra los coeficientes de autocorrelación significativos positivos (1, 2, 3, 4, 5, 6,7 y 8), pero a partir del coeficiente 9 la función de autocorrelación tiende a cero rápidamente, lo que indica que estamos ante una serie no estacionaria en la parte estacional.

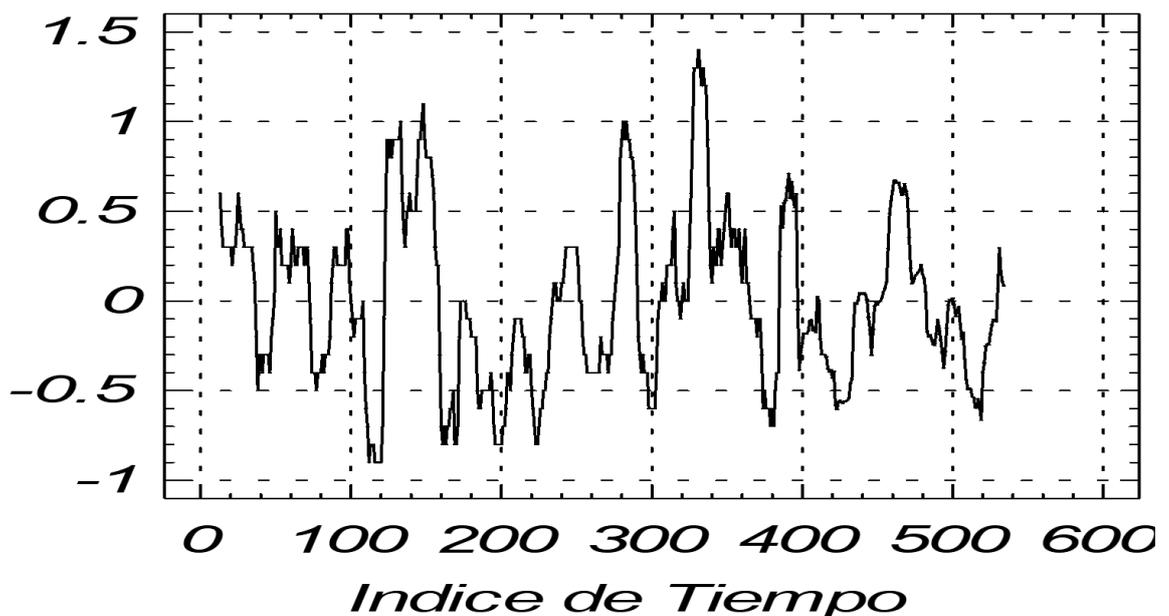


Figura 11. Primera Diferencia Estacional del nivel promedio mensual del lago Titicaca.

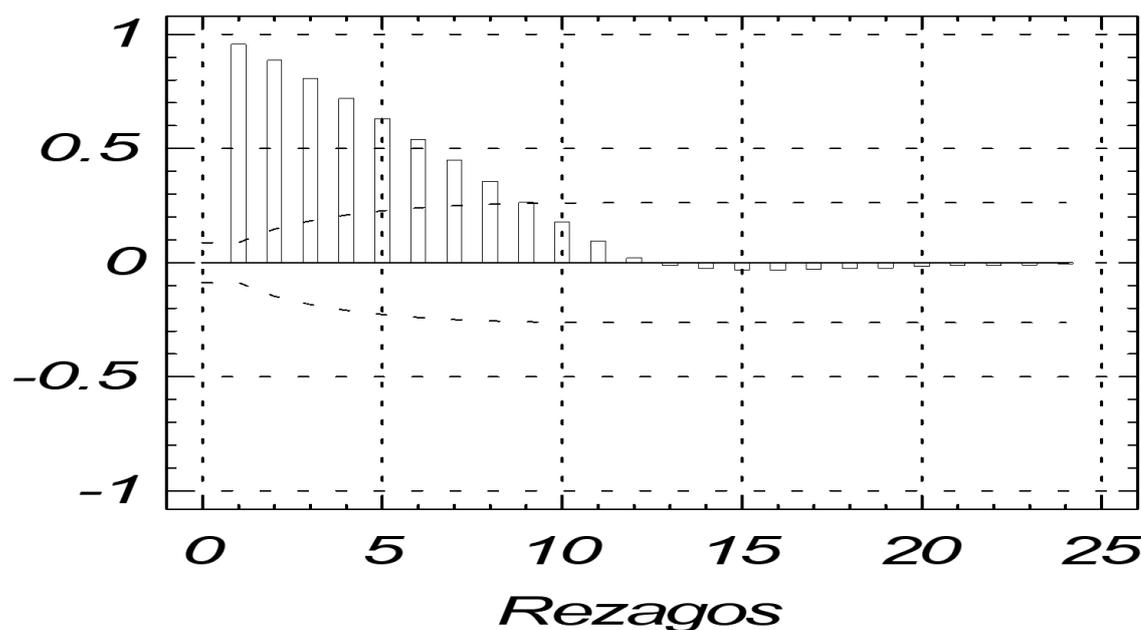


Figura 12. Autocorrelaciones Estimadas para la primera diferencia estacional del nivel promedio mensual del lago Titicaca.

Este correlograma de la figura 13 presenta coeficientes significativos (1, 2, 4,13), pero tiene una caída exponencial en el límite inferior lo que indica o sugiere un modelo de media móvil de orden uno o SMA (1) en la parte estacional.

La figura 14 muestra la nueva serie, conformada por la diferencia no estacional y estacional conjunta, tiene la forma de un ruido blanco con media cero, varianza constante y covarianza cero lo cual nos da la idea de que la serie historica se ajustará a un modelo ARIMA multiplicativo. En esta diferencia conjunta se perdieron 13 observaciones 1 por la diferencia no estacional y 12 observaciones por la diferencia estacional (12 periodos).

La figura 15 muestra el correlograma de las autocorrelaciones estimadas de la diferencia conjunta, en el cual se observa la presencia de 4 coeficientes significativos (1, 11, 12 y 13), el resto de los coeficientes caen a cero rápidamente dando la idea de un modelo multiplicativo SMA de orden uno por ser el primer coeficiente positivo en el periodo de 1 a 9. El correlograma también presenta una caída exponencial a partir del coeficiente 12 lo que confirma el modelo SMA de orden 1. Este correlograma también presenta alternancia de signos entre los coeficientes.

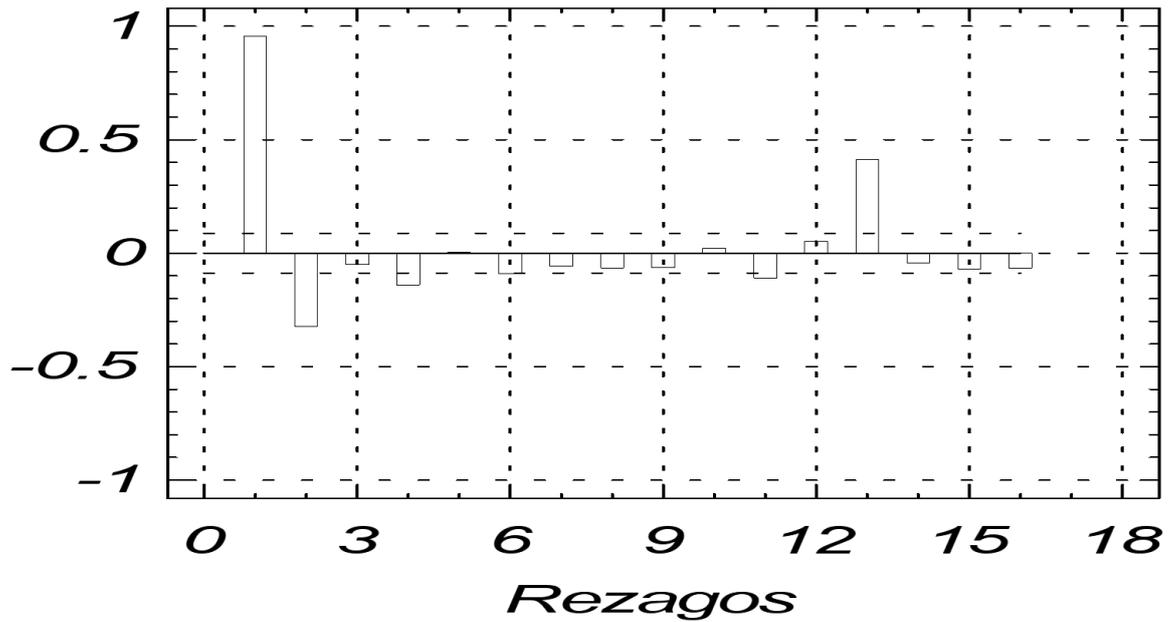


Figura 13. Autocorrelaciones parciales estimadas para la primera diferencia estacional del nivel promedio mensual del lago Titicaca.

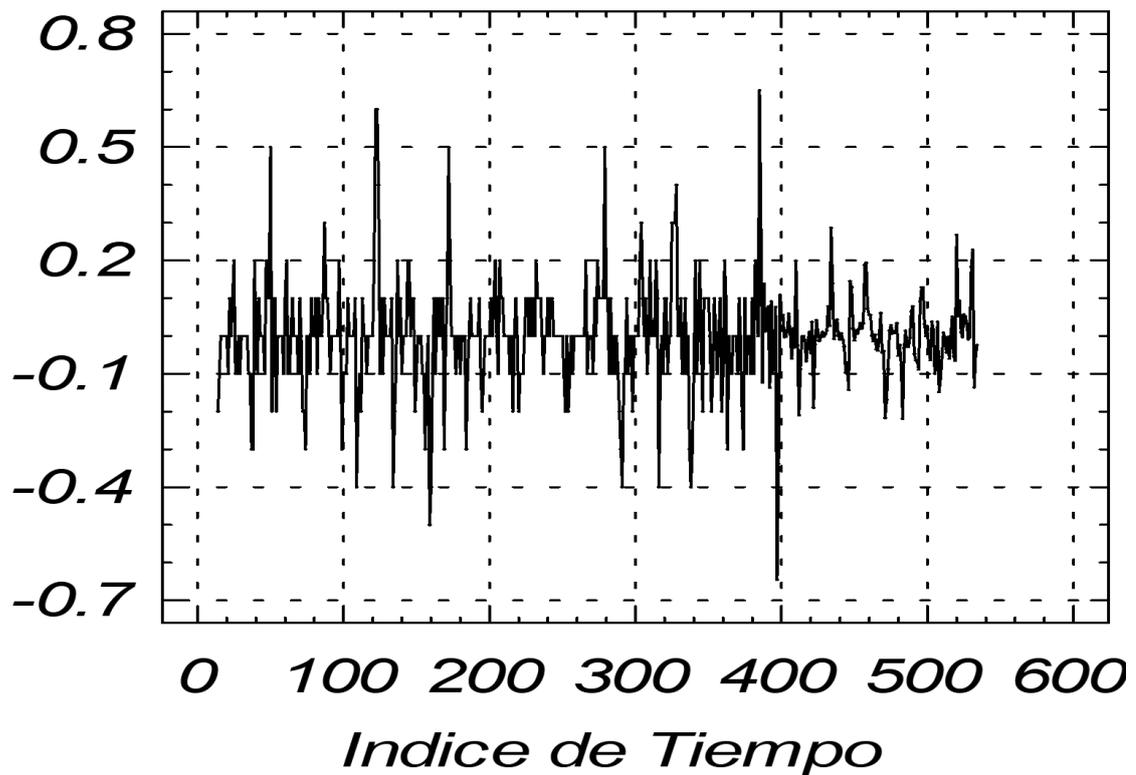


Figura 14. Primera diferencia no estacional por primera diferencia estacional del nivel promedio mensual del lago Titicaca.

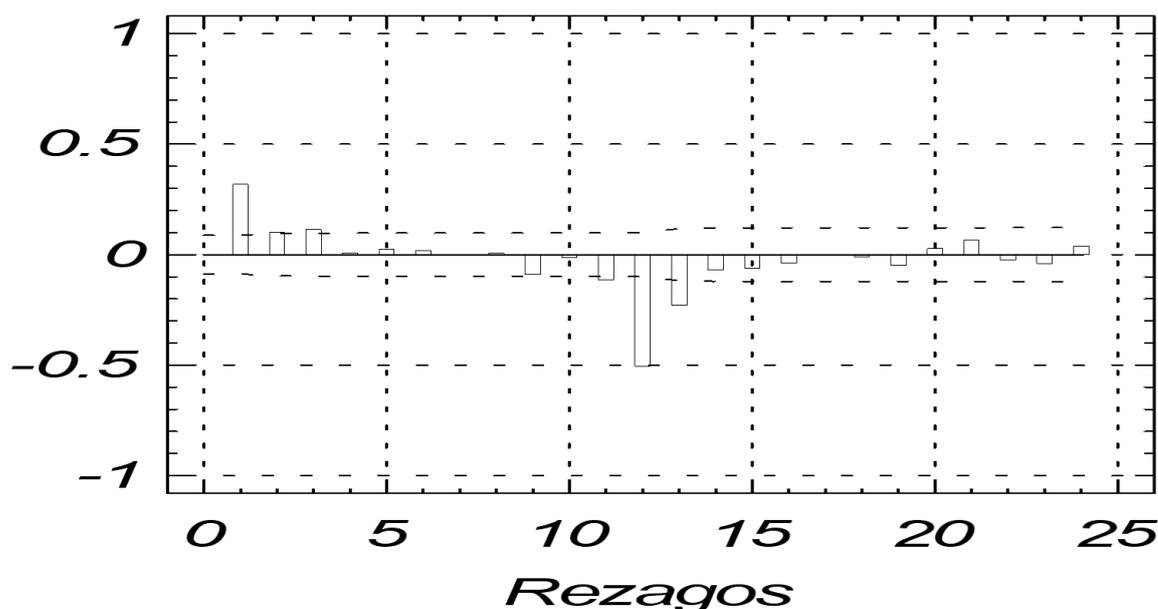


Figura 15. Autocorrelaciones estimadas para la primera diferencia no estacional por primera diferencial estacional del nivel promedio mensual del lago Titicaca.

La figura 16 muestra el correlograma de las autocorrelaciones parciales estimadas de la diferencia conjunta, en el cual se observa la presencia de 4 coeficientes significativos (1, 9, 11 y 12), a partir del coeficiente 13 la serie tiende a cero rápidamente. Este correlograma también presenta alternancia de signos entre los coeficientes.

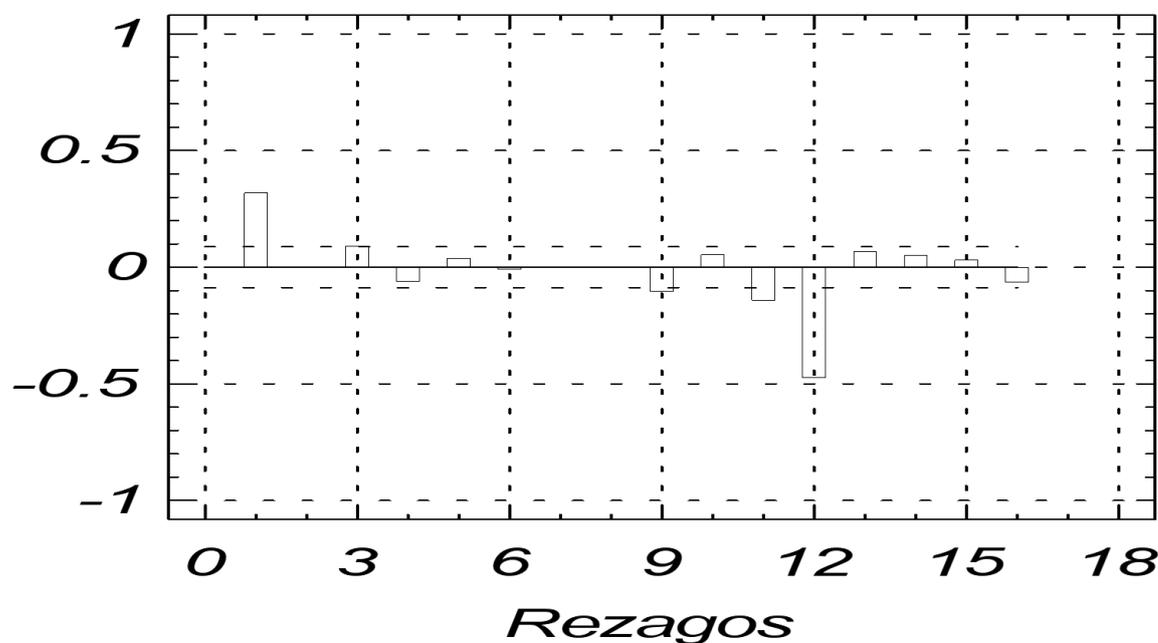


Figura 16. Autocorrelaciones parciales estimadas para la primera diferencia no estacionaria por primera diferencia estacional del nivel promedio mensual del lago Titicaca.

De todo lo discutido anteriormente se llegó a la conclusión de que el modelo que se ajusta a la serie histórica, niveles medios de agua del lago Titicaca, es un modelo multiplicativo ARIMA (1, 1,0) (0, 1,1), cuya ecuación es la siguiente:

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + \phi_1 Y_{t-1} + \varepsilon_t - \theta_{12} \varepsilon_{t-12} + \theta_{13} \varepsilon_{t-13} \quad (1)$$

**4.2. Estimación del modelo identificado.**

Una vez identificado el modelo ARIMA multiplicativo (p, d, q)\*(P, D, Q), para la serie histórica de los niveles medios de agua del lago Titicaca se estimaron los parámetros de dicho modelo.

A continuación se muestra la estimación de los parámetros del modelo, para más información ver anexo 2.

La tabla 3 muestra los estimadores de los parámetros, error estándar, valor T, valor-p (p-estadístico).

Tabla 3

*Resumen del modelo ARIMA (1,1,0)(0,1,1)*

Parámetro	Estimado	Error Estándar	Valor – T	Valor-p
AR (1)	0.32595	0.04159	783.786	0.00000
SMA (12)	0.81061	0.02654	3.053.886	0.00000

Fuente: Elaboración propia.

La tabla 4 muestra la varianza de ruido blanco, desviación estándar de ruido blanco, estadístico de Box-Pierce y probabilidad de aceptación.

Tabla 4

*Resumen de los parámetros y estadísticos*

Varianza de ruido blanco ( $\sigma_e^2$ )	Desviación estándar de ruido blanco ( $\sigma_e$ )	Estadístico de Box-Pierce ( $\chi^2$ )	Probabilidad de aceptación Box-Pierce (p)
0.0102579	0.101281	15.475	0.629135

Fuente: Elaboración propia.

Modelo estimado: Una vez estimado los parámetros se obtuvo la ecuación de pronóstico siguiente:

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + \hat{\phi}_1 Y_{t-1} - \hat{\theta}_{12} \hat{\varepsilon}_{t-12} + \hat{\theta}_{13} \hat{\varepsilon}_{t-13} \quad (2)$$

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + 0.32595Y_{t-1} - 0.81061\hat{\varepsilon}_{t-12} \quad (3)$$

### 4.3. Validación o adecuación del modelo estimado.

El modelo estimado debe cumplir las propiedades siguientes, para considerarse apto en la aplicación de los pronósticos:

- Los valores estimados en valor absoluto son menores que uno, es decir:

$$|\hat{\phi}_1| = |0.32595| < 1 \text{ y } |\hat{\theta}_1| = |0.81061| < 1$$

Por lo tanto la serie histórica de los niveles medios de agua mensual es estacionaria.

- El valor de los coeficientes estimados AR (1): 0.32595 y SMA (1): 0.81061 son significativos por ser el valor de  $P=0.00000 < 0.05$ , para ambos.
- El estadístico de Box- Pierce cuyo valor es:  $\chi^2 = 15.475$ , al nivel de significación  $p = 0.629135 > 0.05$ . Indica que los errores son aleatorios, entonces el modelo es el adecuado.
- La figura 17 de la probabilidad normal de los residuales es satisfactoria, ya que revela que hay una dispersión sensiblemente menor en los residuales debido a que se distribuyen alrededor de la línea sin mostrar signo de tendencia a excepción de los primeros y últimos valores que están un poco dispersos, pero que no implica mayormente alguna tendencia.
- Como ningún coeficiente de autocorrelación es significativo en la figura 18, es decir esta fuera de los límites de probabilidad al 95% de confianza del correlograma de la función de autocorrelación de residuales estimados, quiere decir que la serie temporal es completamente aleatoria (ruido blanco).
- Como ningún coeficiente de autocorrelación es significativo en la figura 19, es decir esta fuera de los límites de probabilidad al 95% de confianza del correlograma de la función de autocorrelación parcial de residuales estimados, quiere decir que la serie temporal es completamente aleatoria (ruido blanco).

Por tanto no hay evidencias de falta de adecuación del modelo. Los residuales de la función de autocorrelación estimada y función de autocorrelación parcial estimada

indican que dichas autocorrelaciones se aceptan individualmente como cero por encontrarse dentro de los límites confidenciales.

La Chi-cuadrado no da resultados significativos, con lo cual se acepta la hipótesis nula de que conjuntamente toda las autocorrelaciones son cero o aleatorias.

La figura 17 de probabilidad normal de residuales es satisfactoria ya que revela que no hay una dispersión menor.

El correlograma de la figura 18 muestra los coeficientes de autocorrelación no significativos ya que convergen a cero rápidamente, también presenta la forma de un ruido blanco con media igual a cero, varianza constante y covarianza cero. Por tanto la serie histórica es estacionaria.

La figura 19 muestra el correlograma de autocorrelación parcial donde se observa que los coeficientes de autocorrelación son no significativos ya que se encuentran dentro de los límites correspondientes, también presenta la forma de un ruido blanco con media igual a cero, varianza constante y covarianza cero. Por tanto la serie histórica es estacionaria.

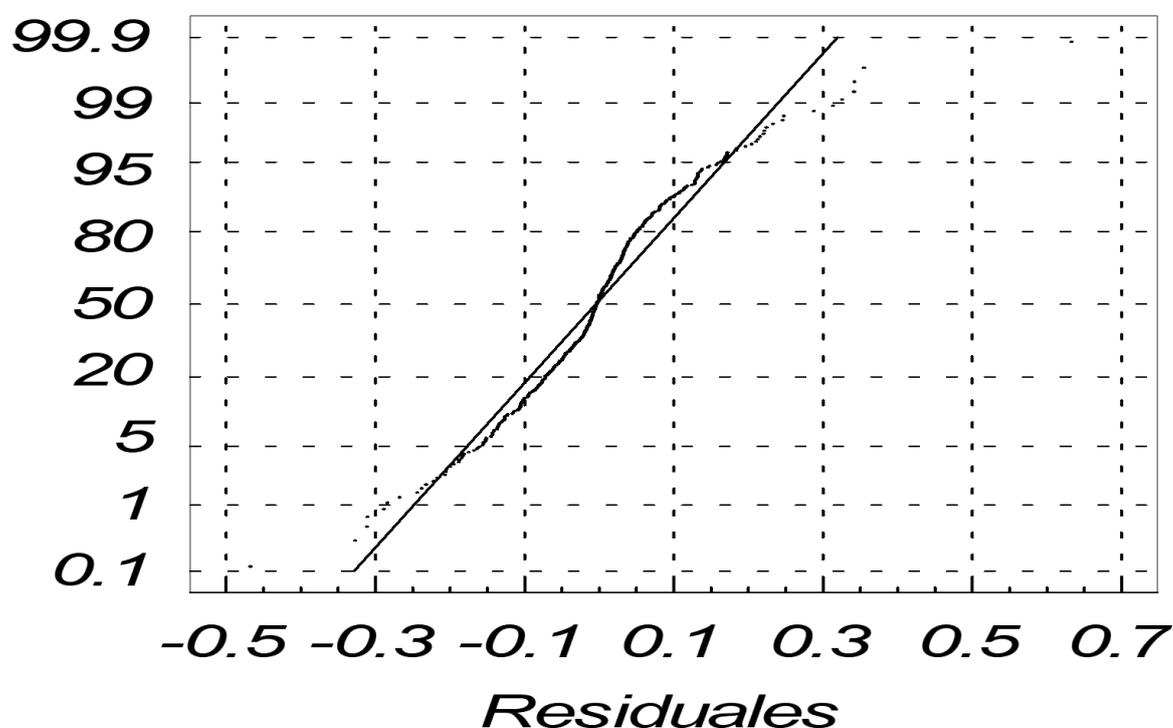


Figura 17. Probabilidad normal del nivel promedio mensual de agua del lago Titicaca.

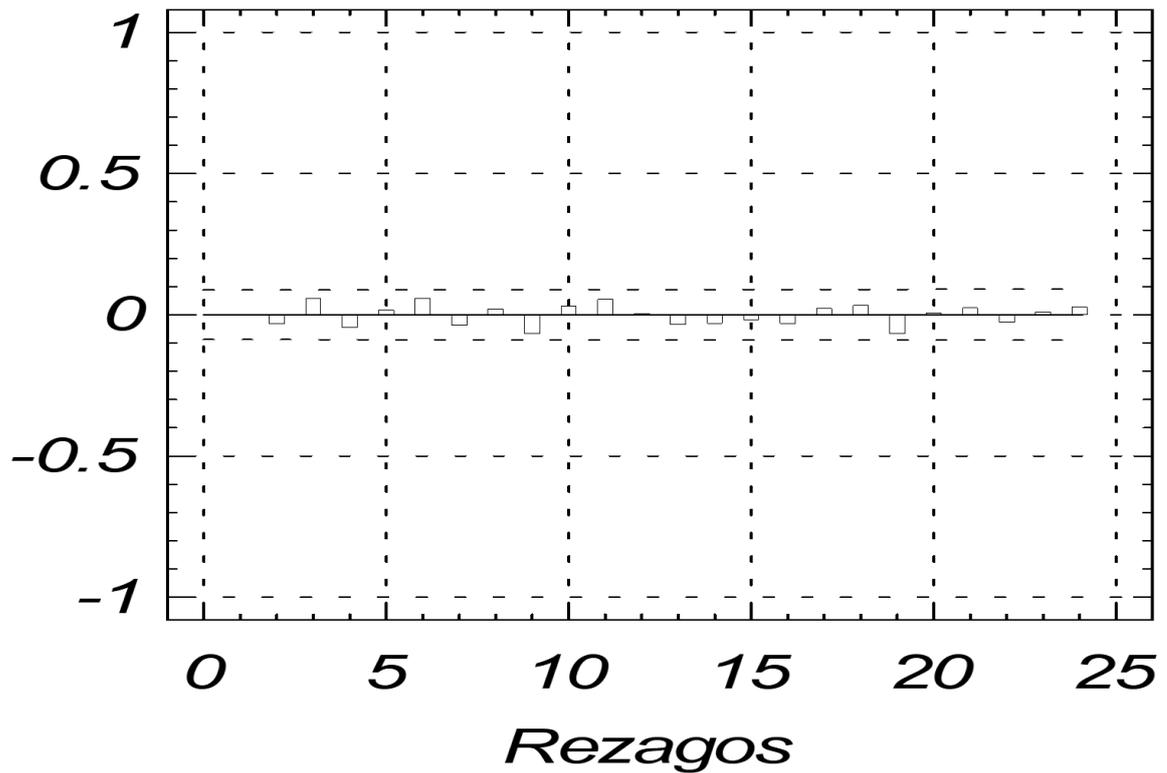


Figura 18. Función de autocorrelación de residuales estimados del nivel promedio mensual de agua del lago Titicaca.

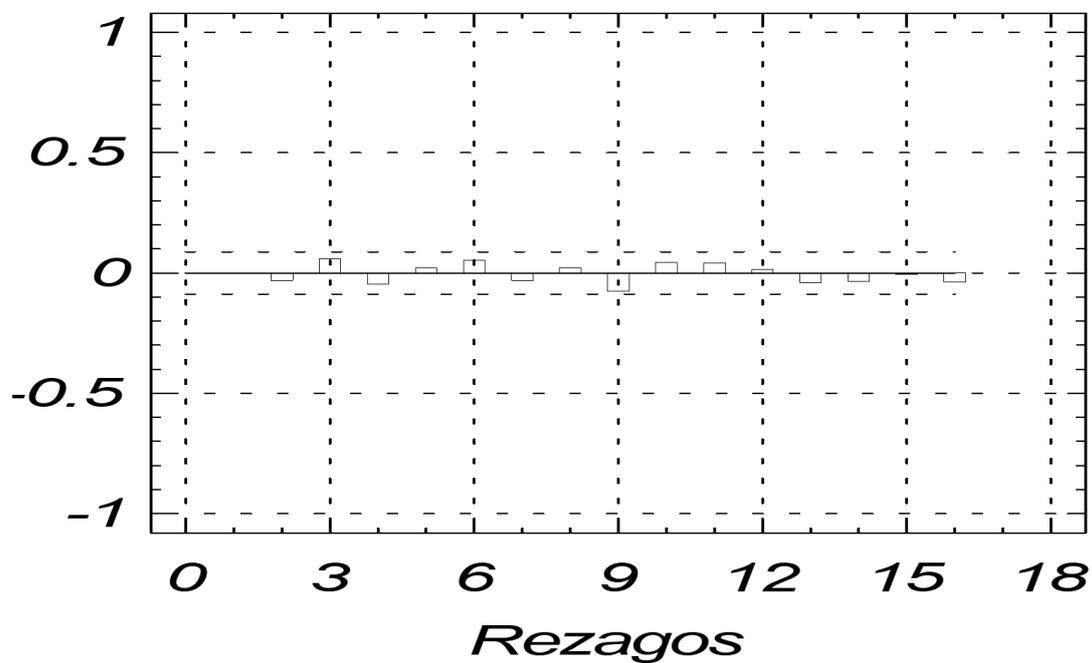


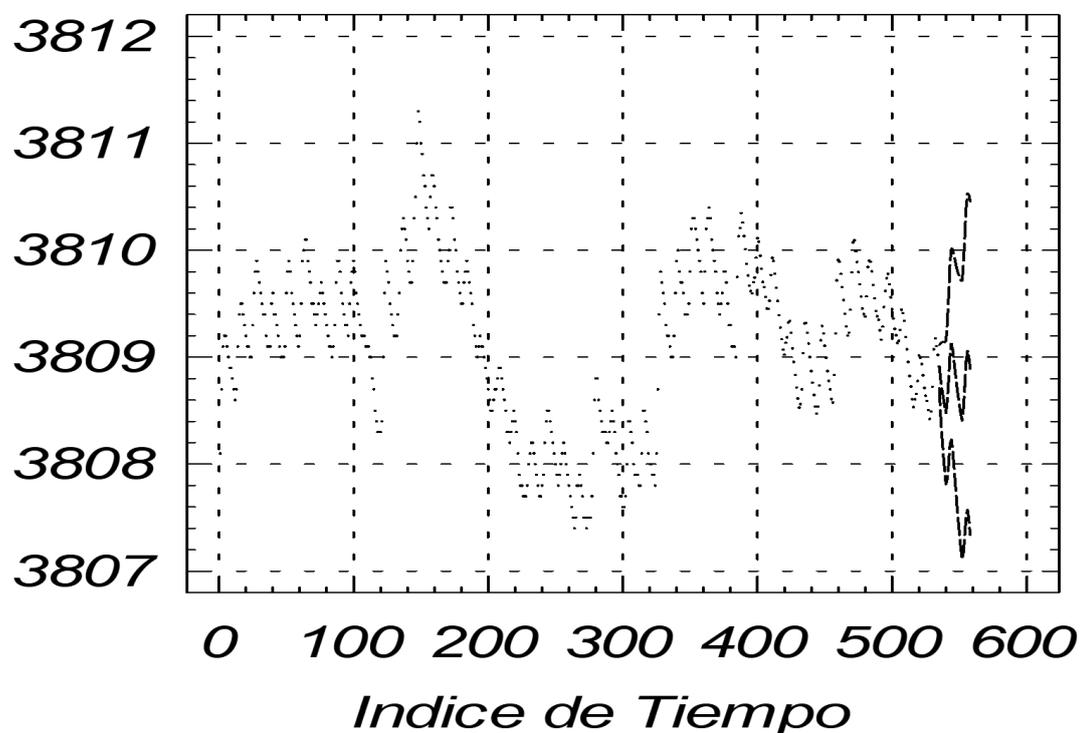
Figura 19. Función de autocorrelación parcial de residuales estimados del nivel promedio mensual del lago Titicaca.

#### 4.4. Pronóstico con el modelo validado.

A continuación, presentamos las figuras de pronósticos para los periodos de 2, 5, 10 y 20 años, para observar la secuencia de los cambios en cada periodo:

- En la figura 20 se muestran los pronósticos para el periodo de 2 años (24 meses), indicando que los niveles medios de agua siguen bajando levemente.

Los límites confidenciales al 95% siguen a los pronósticos obtenidos, notando una amplitud no muy estrecha, esto es debido a que tenemos un periodo de 2 años, mostrando que el nivel medio del agua sigue bajando.



*Figura 20.* Función de pronóstico con límites de confianza del 95% del nivel promedio mensual de agua del lago Titicaca.

- En la figura 21 se muestran los pronósticos para el periodo de 5 años (60 meses), se observa que dichos pronósticos siguen descendiendo pero los intervalos de confianza toman un tamaño más corto y siguiendo a los valores pronosticados, esto nos da a entender de que a medida que aumenta el tamaño del periodo, se ve que es más efectivo la técnica de este modelo ARIMA.

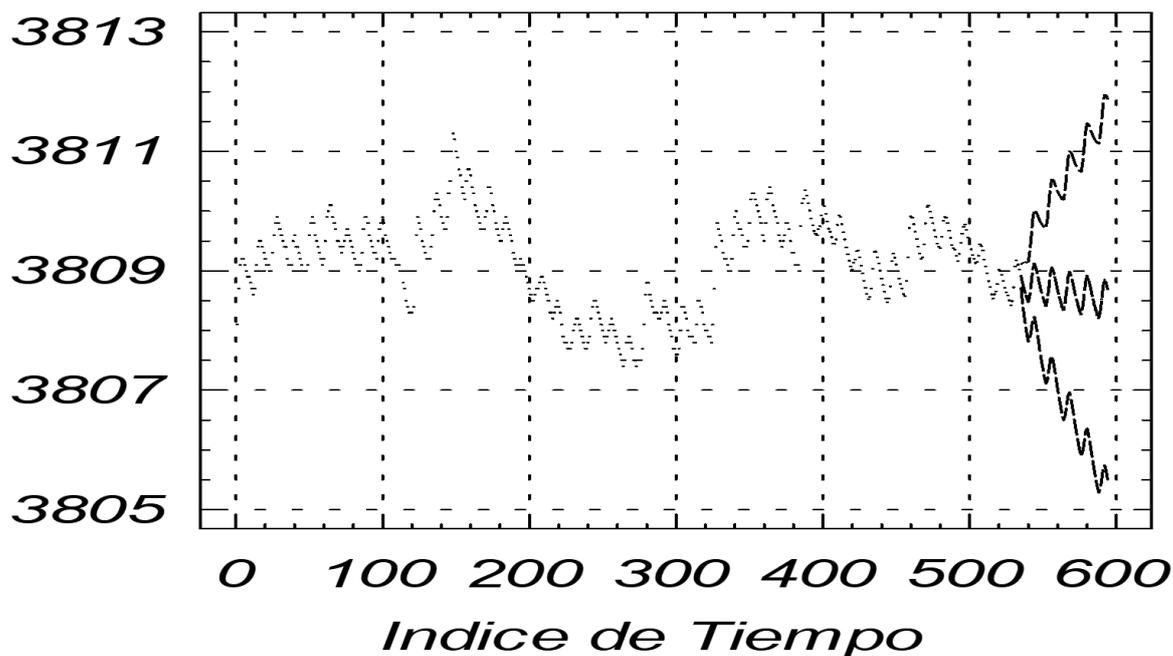


Figura 21. Función pronosticada con límites de confianza del 95% del nivel promedio mensual de agua del lago Titicaca.

- La figura 22 muestra los pronósticos para el periodo de 10 años (120 meses), de la misma manera se observa que mientras más grande es el periodo de tiempo los valores pronosticados siguen bajando su tendencia y los intervalos de confianza siguen reduciendo en tamaño y siguiendo en el sentido de los valores pronosticados.

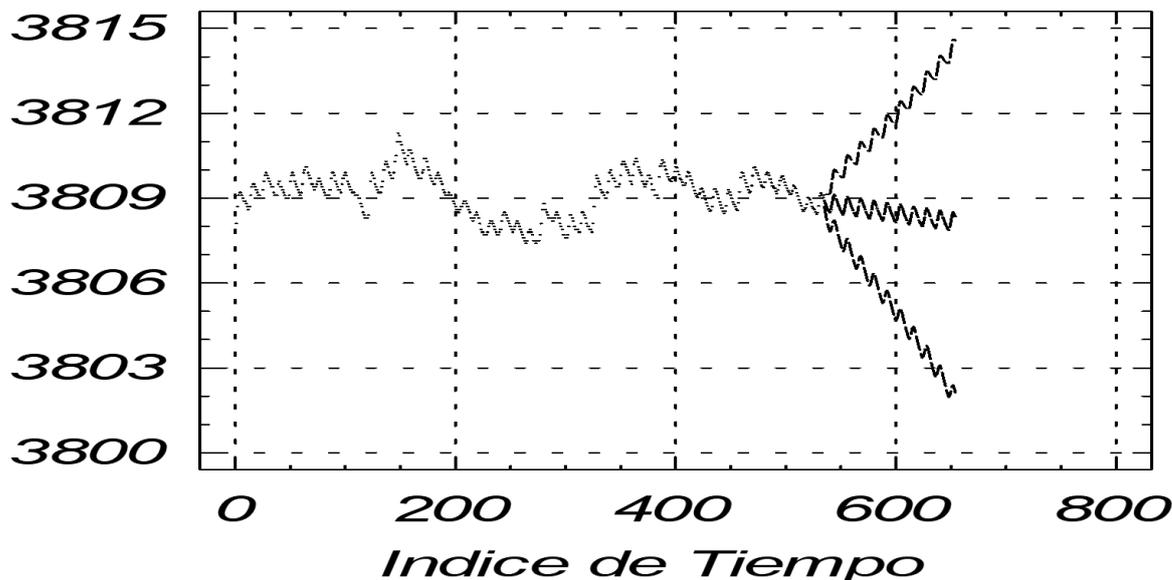
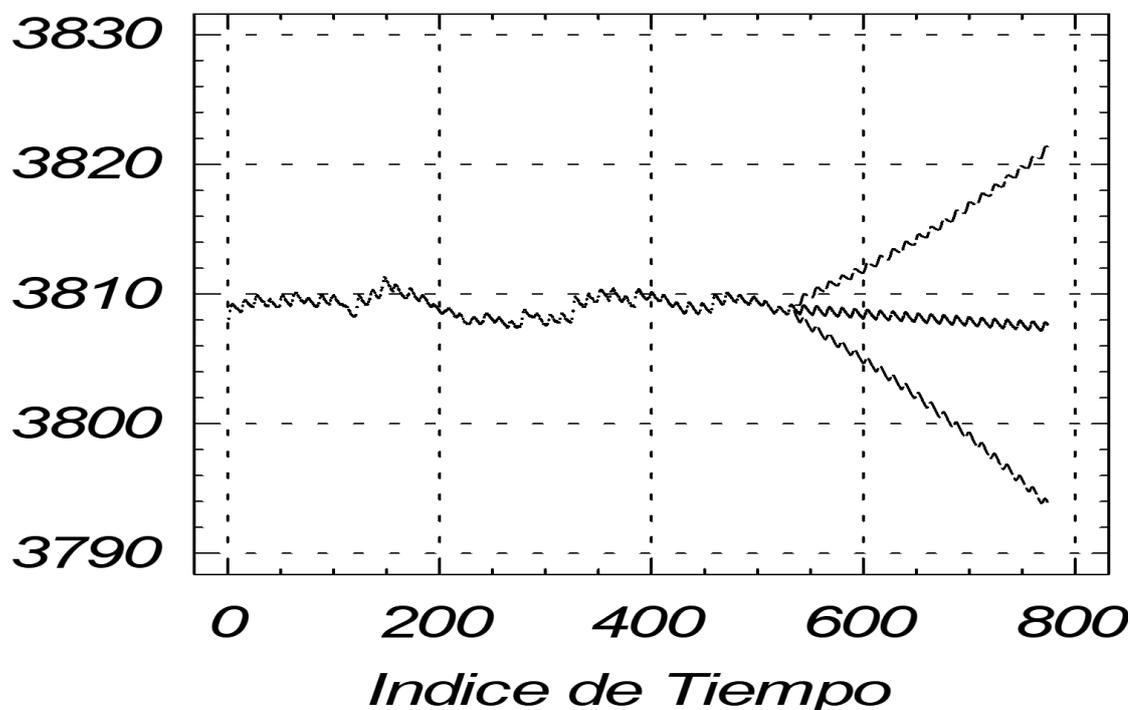


Figura 22. Función pronosticada con límites de confianza al 95% del nivel promedio mensual de agua del lago Titicaca.

- Finalmente en la figura 23 se observa la misma descendencia de los niveles medios de agua mensual del lago Titicaca, pero en una forma más pronunciada, asimismo los intervalos de confianza siguen a dichos pronósticos acortando su distancia, de tal manera que siguen descendiendo sin tener indicios de elevar su nivel medio de agua del lago Titicaca.



*Figura 23.* Función pronosticada con límites de confianza al 95% del nivel promedio mensual de agua del lago Titicaca.

Por tanto de acuerdo al monitoreo de este estudio de los niveles medios de agua mensuales del lago Titicaca a través de la técnica de los modelos ARIMA, se ve que siguen cayendo dichos pronósticos sin tener indicios de levantar su nivel, y esto hace ver que las precipitaciones en cada año van disminuyendo y afectando el caudal de los ríos que alimentan al lago Titicaca, y también a la agricultura y ganadería de la región altiplánica, lo cual causa preocupación en los agricultores y ganaderos, teniendo como consecuencia el encarecimiento de los productos agrícolas y ganaderos afectando a la economía de las personas.

En tal sentido la utilidad de estos modelos ARIMA en los pronósticos es de gran importancia para toma de decisiones orientado a la precaución en la solución de problemas en el futuro.

## CONCLUSIONES

- ✓ El modelo que mejor se ajusta a la serie historia de niveles medios de agua del lago Titicaca, es el modelo ARIMA multiplicativo (1, 1,0) (0, 1,1), donde el primer factor es la parte no estacional y el segundo factor es la parte estacional, cuya ecuación es la siguiente:

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + \phi_1 Y_{t-1} + \varepsilon_t - \theta_{12} \varepsilon_{t-12} + \theta_{13} \varepsilon_{t-13}$$

- ✓ Se identificó, estimó y validó el modelo que mejor se ajusta a la serie historia de niveles medios de agua mensual del lago Titicaca, el cual es un modelo ARIMA multiplicativo (1, 1,0) (0, 1,1), cuya ecuación de pronóstico es la siguiente:

$$\hat{Y}_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + 0.32595Y_{t-1} - 0.81061\hat{\varepsilon}_{t-12}.$$

- ✓ Se obtuvieron los pronósticos de la serie histórica de los niveles medios de agua mensual del lago Titicaca, con el modelo estimado para los periodos de 24, 60, 120 y 240, cuyos resultados se encuentra en anexo.

## RECOMENDACIONES

- ✓ Se debe realizar este tipo de investigación periódicamente a fin de tener conocimiento genérico de los niveles medios del agua del lago Titicaca, para de esta manera monitorear los cambios que se dan en diferentes periodos de tiempo.
- ✓ Se debe evitar la sobre-parametrización y sobre-diferenciación, debido a que no conducen a obtener un buen modelo.
- ✓ Se recomienda a las instituciones a aplicar estos modelos ARIMA para realizar sus pronósticos y tomar decisiones en la solución de problemas.

## BIBLIOGRAFÍA

- Aguilera, G. M. (2015). *Modelo SARIMA para el pronóstico de la masa mensual del río Jequetepeque*. (tesis de posgrado), Universidad Nacional de Trujillo, Trujillo, Perú.
- Arias, J. E. (2017). *Comportamiento de la demanda de depósitos a plazo fijo en moneda nacional en las cajas municipales períodos 2004-2006*. (tesis de posgrado), Universidad Agraria la Molina, Lima, Peru.
- Barreras, A., Sánchez, E., Figueroa, F., y Olivas, J. Á. (2014). *Uso de un modelo univariado de series de tiempo para la predicción, en el corto plazo, del comportamiento de la producción de la carne de bovino en Baja California*. . México: Veterinaria México, 1 (1), 1-9.
- Bowerman, B. L., y O'Connell, R. T. (1987). *Time series forecasting, unified concepts and computer implementation*. Boston: Duxbury Press.
- Chatfield, C. (1987). *The analysis of time series: theory and practice*. Londres: Chapman and Hall.
- Cuitiño, F., Ganón, E., Tiscordio, I., y Vicente, L. (2010). *Modelos univariados de series de tiempo para predecir la inflación de corto plazo*. . Uruguay: Banco Central del Uruguay, (8), 1-39.
- De la Fuente, G. D., Pino, D. R., y Parreño, F. J. (1995). *Métodos cuantitativos de previsión*. España: Universidad de Oviedo.
- Díaz, N. (2008). *Pronóstico mediante modelos de series de tiempo para el consumo de agua potable de la Empresa Municipal de Saneamiento Básico de la ciudad de*

- Puno, periodo 2000-2007* . (tesis de posgrado). Universidad Nacional del Altiplano, Puno, Peru, Perú.
- Figueiredo, N., y Blanco, C. J. (2014). *Simulação de Vazões e Níveis de Água Médios Mensais para o Rio Tapajós Usando Modelos ARIMA*. Brasil: Revista brasileira de recursos hídricos, 19 (3), 111-126.
- Guerra, J., Sánchez, G., y Reyes, B. (1997). *Modelos de series de tiempo para predecir la inflación en Venezuela*. Banco Central de Venezuela: Gerencia de Investigaciones Económicas, 13 (1), 1-31.
- Guevara, J. E., Vega, C. E., y Guevara, E. (2010). *Análisis estocástico de una serie cronológica de caudales mensuales. Caso de estudio: Río Caroní, Venezuela*. Venezuela: Revista IngenieríaUC-17(1),46-45.
- Guevara, R. R. (2012). *Metodología para la medición de la atención en una central telefónica usando Box-Jenkins*. (tesis de posgrado) Universidad Nacional Mayor de San Marcos, Lima, Perú.
- Hanke, J. E., y Reitsch, A. G. (1996). *Pronóstico en los negocios*. México: Editotial Prentice Hall.
- Hernandez, S. R. (2014). *Metodología de la Investigación*. México D.F.: Mc Graw Hill Education.
- Huancantara, K. (2010). *Modelo uniecuacional para describir y predecir el comportamiento de los niveles medios de agua del lago Titicaca, periodo 1984-2008*. (tesis de posgrado), Universidad Nacional del Altiplano, Puno, Perú.
- Jaramillo, M., Gonzáles, D. E., Nuñez, M. E., Esperanza, G., y Lucio, J. H. (2006). *Análisis de series de tiempo univariante aplicado metodología de Box-Jenkins para la predicción de ozono en la ciudad de Cali, Colombia*. Colombia: Facultad de Ingeniería, 39 (1), 79-88.
- Mogni, A. P. (2013). *Modelos de Series de Tiempo con aplicaciones en la industria aereocomercial*. Universidad Buenos Aires, Buenos Aires, Argentina.

- Neira, P. (1995). *Modelos Univariantes y Espectros de Series de Tiempo para los pacientes Hospitalizados por Neumonía Carlos Monje Medrano Juliaca 1980 – 1993*. (tesis de posgrado), Universidad Nacional del Altiplano, Puno-Perú.
- Oliver, D. A. (1976). *Time series analysis and forecasting: the Box-Jenkins approach*. Michigan: Butterworth.
- Paula, W., Borbosa, G., y Bras, J. (2015). *Análise comparativa de modelos de séries temporais para modelagem e previsão de regimes de vazões médias mensais do Rio Doce, Colatina - Espírito Santo*. Brasil: Centro de Ciencia Naturais e Exatas, 37 (4), 1-11.
- Reyes, T. M., y Mejia, T. R. (2010). *Análisis de las periodicidades de los caudales medios mensuales en la cuenca del Río Santa*. Rio Santa: Aporte Santiaguino, 3 (1), 9-17.
- Uriel, J. E. (1985). *Análisis de Series Temporales ARIMA*. Madrid-España: Paraninfo.
- Vázquez, M. F. (2017). *Modelo de Box-Jenkins y redes neuronales para pronósticar el precio del dólar del sistema bancario en moneda nacional año 2015*. (tesis de posgrado), Universidad Nacional Pedro Ruiz Gallo, Lambayeque, Perú.



**ANEXOS**

Anexo 1. Pronósticos de corto plazo para los niveles medios de agua en el lago Titicaca

Variable: WORKAREA.FORECASTS (length = 24 3)

---



---

( 1,1) 3808.91	( 1,2) 3808.71	( 1,3) 3809.11
( 2,1) 3808.8	( 2,2) 3808.47	( 2,3) 3809.13
( 3,1) 3808.7	( 3,2) 3808.27	( 3,3) 3809.14
( 4,1) 3808.62	( 4,2) 3808.1	( 4,3) 3809.15
( 5,1) 3808.54	( 5,2) 3807.93	( 5,3) 3809.14
( 6,1) 3808.48	( 6,2) 3807.81	( 6,3) 3809.15
( 7,1) 3808.6	( 7,2) 3807.86	( 7,3) 3809.33
( 8,1) 3808.8	( 8,2) 3808.01	( 8,3) 3809.59
( 9,1) 3809.03	( 9,2) 3808.19	( 9,3) 3809.87
(10,1) 3809.12	(10,2) 3808.23	(10,3) 3810.02
(11,1) 3809.07	(11,2) 3808.13	(11,3) 3810.01
(12,1) 3808.96	(12,2) 3807.98	(12,3) 3809.95
(13,1) 3808.85	(13,2) 3807.81	(13,3) 3809.89
(14,1) 3808.73	(14,2) 3807.64	(14,3) 3809.83
(15,1) 3808.64	(15,2) 3807.49	(15,3) 3809.79
(16,1) 3808.55	(16,2) 3807.35	(16,3) 3809.76
(17,1) 3808.47	(17,2) 3807.21	(17,3) 3809.72
(18,1) 3808.41	(18,2) 3807.11	(18,3) 3809.71

---



---

Variable: WORKAREA.FORECASTS (length = 24 3)

---



---

(19,1) 3808.53	(19,2) 3807.18	(19,3) 3809.88
(20,1) 3808.73	(20,2) 3807.34	(20,3) 3810.12
(21,1) 3808.96	(21,2) 3807.52	(21,3) 3810.4
(22,1) 3809.06	(22,2) 3807.58	(22,3) 3810.53
(23,1) 3809	(23,2) 3807.48	(23,3) 3810.52
(24,1) 3808.9	(24,2) 3807.34	(24,3) 3810.45

---



---

Anexo 2. Tabla de estimación de los parámetros del modelo.

---



---

Summary of Fitted Model for: LAGOTITI.nivelesmed

---

Parameter	Estimate	Std.error	T-value	P-value
AR (1)	0.32595	0.04159	7.83786	0.00000
SMA (12)	0.81061	0.02654	30.53886	0.00000

---

Model fitted to differences of order 1  
 Model fitted to seasonal differences of order 1 with seasonal length = 12  
 Estimated white noise variance = 0.0102579 with 519 degrees of freedom.  
 Estimated white noise standard deviation (std err) = 0.101281  
 Chi-square test statistic on first 20 residual autocorrelations = 15.475  
 with probability of a larger value given white noise = 0.629135  
 Backforecasting: no                      Number of iterations performed: 4

---



---

Anexo 3. Tabla de residuales del modelo

---



---

Variable: WORKAREA.RESIDS (length = 534)

---

( 1)	( 19) 0	( 37) -0.131705
( 2)	( 20) -4.54747E-13	( 38) -0.467537
( 3)	( 21) -0.1	( 39) 0.220275
( 4)	( 22) 0.132595	( 40) -0.0173506
( 5)	( 23) -0.0325954	( 41) -0.181061
( 6)	( 24) 0.1	( 42) 0.159018
( 7)	( 25) 0.167405	( 43) -0.0325954
( 8)	( 26) -0.327313	( 44) -3.84934E-13
( 9)	( 27) -0.0956212	( 45) -0.0657089
( 10)	( 28) 0.0590176	( 46) -0.0128729
( 11)	( 29) -0.1	( 47) 0.130116
( 12)	( 30) 0.0325954	( 48) 0.0458792
( 13)	( 31) 1.48227E-13	( 49) -0.0393566
( 14) -0.2	( 32) 8.61244E-14	( 50) 0.0884141
( 15) -0.0348091	( 33) -0.0810611	( 51) -0.18442
( 16) 0.0325954	( 34) 0.107483	( 52) 0.0511263
( 17) 0	( 35) -0.126422	( 53) -0.04677
( 18) 0	( 36) 0.0136565	( 54) -0.103694

---



---

( 55) 0.0387687	( 73) -0.02497	( 91) -0.105772
( 56) 1.42716E-13	( 74) -0.310619	( 92) 0.0325954
( 57) -0.0532644	( 75) -0.0780331	( 93) -0.0130188
( 58) -0.0104349	( 76) 0.0600168	( 94) -0.0259143
( 59) 5.47362E-3	( 77) -0.049671	( 95) 0.0889806
( 60) 0.169786	( 78) 0.038037	( 96) -0.0674049
( 61) 0.135502	( 79) -7.12095E-3	( 97) 0.183592
( 62) -0.0935215	( 80) 1.79901E-13	( 98) -0.188234
( 63) -0.216897	( 81) -0.0160605	( 99) -0.134514
( 64) 0.074039	( 82) -0.0319689	(100) 0.0390171
( 65) 0.0620877	( 83) 0.10977	(101) -0.045404
( 66) -0.116651	( 84) -0.0831532	(102) -0.0688331
( 67) 0.0314263	( 85) -0.020241	(103) 0.0468553
( 68) -3.3906E-13	( 86) -0.151791	(104) -6.17323E-3
( 69) -0.143177	( 87) 0.20415	(105) -0.0105532
( 70) 0.124137	( 88) 0.0508639	(106) -0.0210064
( 71) -0.0281585	( 89) 0.0271407	(107) 0.0721286
( 72) -0.06237	( 90) -1.76228E-3	(108) 0.0453609

---



---



---



---

(109) -0.283773	(127) -0.101807	(145) 0.196138
(110) -0.222203	(128) 0.047478	(146) -0.182795
(111) -0.143847	(129) -0.0131076	(147) 0.232692
(112) -0.135777	(130) -0.0138031	(148) 0.101194
(113) 0.128386	(131) 0.047395	(149) -0.127049
(114) -0.0883923	(132) 0.0298062	(150) -0.047608
(115) 0.0379814	(133) 0.0756576	(151) 0.04676
(116) -0.105004	(134) -0.145081	(152) -0.076286
(117) 0.0240409	(135) 0.163694	(153) 0.0178093

(118) -0.017028	(136) 0.0416843	(154) -0.10907
(119) 0.0584682	(137) 0.130207	(155) -0.0362618
(120) 0.03677	(138) -0.0157892	(156) -0.0856971
(121) -0.0300298	(139) 0.0174739	(157) 0.156778
(122) 0.354689	(140) -0.0941093	(158) -0.11558
(123) 0.287823	(141) 0.0219703	(159) -0.311377
(124) 0.0943652	(142) -0.0111889	(160) -0.154994
(125) -0.126311	(143) 0.0384189	(161) -0.0726052
(126) 0.0609437	(144) 0.224161	(162) 0.0940039

---



---

(163) -0.0946913	(181) -0.133993	(199) 0.0306243
(164) 0.0707572	(182) 3.32029E-3	(200) 0.0112662
(165) -0.018159	(183) -0.123542	(201) 0.0246188
(166) 0.0115868	(184) -0.0229616	(202) -5.00567E-3
(167) -0.0619897	(185) -0.0199107	(203) 0.0669816
(168) 0.030533	(186) 0.04152	(204) 0.117959
(169) -0.20551	(187) 0.0377793	(205) -0.0721751
(170) 4.09603E-3	(188) 0.0138984	(206) -0.205302
(171) -0.152406	(189) -0.0929931	(207) 0.0157177
(172) 0.341765	(190) 0.0340358	(208) -8.49536E-3
(173) -0.0218317	(191) -0.0407328	(209) 0.0807437
(174) 0.0110096	(192) -0.0609981	(210) -5.31311E-3
(175) -0.0767578	(193) -8.61605E-3	(211) 0.0248244
(176) 0.0573566	(194) -0.129904	(212) 9.13249E-3
(177) -0.11472	(195) -0.267549	(213) 0.0199562
(178) 0.0419878	(196) -0.0534221	(214) -0.104058
(179) -0.0502495	(197) 0.0164556	(215) 0.0868914
(180) -0.0752496	(198) 0.0336565	(216) -0.104381

---



---

---



---

(217) 6.68495E-3	(235) 0.0290776	(253) 0.0209112
(218) -0.0664196	(236) -1.72371E-4	(254) -0.088608
(219) -0.0198545	(237) -5.82591E-3	(255) -0.0121638
(220) -0.206886	(238) 0.0188591	(256) 0.0273944
(221) 0.0306425	(239) 0.0306732	(257) -0.0608057
(222) -0.0717114	(240) 0.112473	(258) 0.0386896
(223) -0.0472816	(241) -0.054625	(259) 0.0191066
(224) 0.0399983	(242) 0.137417	(260) -1.13263E-4
(225) 0.116177	(243) 0.0279361	(261) -3.82815E-3
(226) -0.0169457	(244) -6.41622E-3	(262) 0.0123921
(227) 0.0378397	(245) 0.0483515	(263) -0.060906
(228) 0.0153874	(246) 7.51803E-3	(264) -0.061795
(229) -0.0271765	(247) 0.0235706	(265) 0.0169509
(230) 0.0461596	(248) -1.39725E-4	(266) 0.128173
(231) -0.0486897	(249) -4.72255E-3	(267) -0.175051
(232) 0.0322957	(250) 0.0152874	(268) 0.0548016
(233) 0.0596483	(251) -0.075136	(269) -0.0492898
(234) 9.27453E-3	(252) -0.0762326	(270) 0.0313622

---



---



---



---

(271) -0.084512	(289) -0.101627	(307) 0.034629
(272) 0.132504	(290) -0.0959489	(308) -0.122723
(273) -0.0356986	(291) -0.0383548	(309) 0.0349989
(274) 0.210045	(292) 0.0153413	(310) 0.0840482
(275) -0.0145619	(293) -0.0451535	(311) 1.31534E-3
(276) 0.0173129	(294) 0.04572	(312) -0.0564873
(277) 0.0811451	(295) -0.0806441	(313) -0.0667783
(278) 0.171303	(296) 0.0121789	(314) 0.136953
(279) 0.325506	(297) 2.96505E-3	(315) 0.171729

(280) -0.0185544	(298) -0.143043	(316) -0.232176
(281) 0.0274497	(299) 0.0820446	(317) 2.5063E-3
(282) -0.107173	(300) -0.0696849	(318) 0.0362153
(283) 0.0640891	(301) -0.0823803	(319) -0.0719294
(284) -0.0251866	(302) -0.0777772	(320) 0.0331148
(285) 3.6578E-3	(303) 0.168909	(321) 0.0957751
(286) 0.0702648	(304) 0.247245	(322) -0.0644651
(287) 0.0207914	(305) -0.0343882	(323) 0.0336617
(288) -0.085966	(306) 4.46565E-3	(324) -0.0457892

---



---

(325) 0.245869	(343) 0.0663925	(361) 0.221068
(326) 0.313229	(344) 0.114276	(362) 0.0577061
(327) 0.341419	(345) -0.0568969	(363) -0.288617
(328) 0.114009	(346) -0.134876	(364) -5.06404E-3
(329) -0.12835	(347) 0.0798263	(365) -0.0437455
(330) 0.0293565	(348) 0.0637392	(366) -0.0528299
(331) 0.0416933	(349) -0.0142208	(367) 0.049799
(332) -0.105752	(350) 0.0282466	(368) -5.97141E-3
(333) 0.0102317	(351) 0.0542529	(369) -0.0109642
(334) 0.0803393	(352) -0.12688	(370) -0.107565
(335) -0.105309	(353) 0.0693977	(371) -3.49583E-3
(336) -4.52179E-3	(354) 0.0179798	(372) -0.0189298
(337) -0.100696	(355) 0.021223	(373) -0.0208
(338) -0.0483068	(356) -7.36656E-3	(374) -0.188032
(339) 0.107139	(357) -0.0135258	(375) -0.03617
(340) 0.0902035	(358) -9.33197E-3	(376) -0.1367
(341) 0.128553	(359) -0.167887	(377) -2.86513E-3
(342) -0.141394	(360) 0.0168585	(378) -0.0428245

---



---

(379) -0.0596324	(397) -0.156869	(415) 3.56652E-3
------------------	-----------------	------------------

(380) 0.127755	(398) -0.195276	(416) -0.0369919
(381) -0.141483	(399) 0.0214541	(417) -5.11684E-3
(382) 0.145402	(400) 0.0618208	(418) -0.0126918
(383) 0.0319754	(401) 0.0142383	(419) -0.0271498
(384) -0.0479401	(402) 5.53611E-3	(420) 0.0188216
(385) 0.633139	(403) 3.99769E-3	(421) -0.0810119
(386) -0.083291	(404) -0.0320646	(422) -0.161136
(387) -0.243913	(405) 0.0719182	(423) -0.0198168
(388) 0.0682816	(406) -0.0216265	(424) -0.0735039
(389) -0.0346302	(407) -0.0538989	(425) -0.0121813
(390) 0.0330486	(408) 0.083156	(426) 0.010471
(391) 6.19264E-3	(409) -0.163679	(427) 5.12847E-3
(392) -0.0588649	(410) 0.0504858	(428) -0.0183341
(393) 0.0219681	(411) -0.072474	(429) -8.73326E-3
(394) -9.12305E-3	(412) -0.151739	(430) -4.96212E-3
(395) 9.53715E-3	(413) -8.00783E-3	(431) -9.63765E-3
(396) 0.0537851	(414) 0.0321716	(432) 0.0896937

---



---

(433) -0.0644195	(451) 0.0117304	(469) 0.0212589
(434) 0.147581	(452) -0.0184962	(470) 0.0812376
(435) 0.0293874	(453) 0.018358	(471) -0.0544875
(436) -0.112891	(454) 0.0149166	(472) -0.0785842
(437) 0.0127333	(455) -0.0174701	(473) -0.0100212
(438) 0.0449688	(456) 0.0141867	(474) 0.0404818
(439) -8.85885E-3	(457) 0.103801	(475) 0.0205318
(440) -0.0241878	(458) 0.132273	(476) -8.47539E-3
(441) 2.85432E-3	(459) 0.184397	(477) -6.38286E-3
(442) -0.016304	(460) 0.0242826	(478) -8.8019E-3
(443) -0.0325528	(461) 0.0282411	(479) 0.0342706
(444) -0.0151667	(462) -2.71667E-3	(480) -5.63441E-3
(445) -0.0906015	(463) -4.81719E-3	(481) 0.0101823

(446) -5.52698E-4	(464) -7.42985E-3	(482) -3.32507E-3
(447) 0.215433	(465) -0.0210967	(483) -0.23807
(448) -9.7738E-3	(466) -0.0155	(484) -0.0233172
(449) -0.0117263	(467) 3.55076E-3	(485) 0.0159813
(450) 0.0179331	(468) 0.0708701	(486) -7.48369E-4

---



---

(487) -2.90405E-3	(505) -0.136736	(523) -6.69728E-3
(488) -1.41757E-3	(506) -0.105767	(524) -9.06128E-3
(489) 0.0611298	(507) 4.57654E-3	(525) 0.0305845
(490) 0.050678	(508) -0.0948572	(526) 0.0354288
(491) -0.0449703	(509) 0.0152265	(527) -0.0170697
(492) -0.0532475	(510) -0.0264781	(528) -0.0298905
(493) -0.038885	(511) 0.0155642	(529) -0.049573
(494) -0.0695304	(512) -3.05422E-3	(530) 0.058368
(495) -0.0729718	(513) -0.0337935	(531) 0.131861
(496) 0.080437	(514) 2.47741E-3	(532) -0.037965
(497) 0.0999065	(515) -0.0408582	(533) -0.0255142
(498) -0.0106548	(516) -0.0125017	(534) 0.0369342
(499) -0.0127846	(517) -0.0598395	
(500) 6.85091E-3	(518) -0.155359	
(501) 8.94483E-3	(519) -0.0450146	
(502) -0.0105336	(520) 0.212621	
(503) 0.0224077	(521) -0.022013	
(504) -0.0395492	(522) 0.054261	

---



---