



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ciencias Matemáticas

Escuela Profesional de Estadística

**Modelos de cambio de régimen de transición
determinística: Modelos SETAR (Self-Exciting
Threshold Autoregressive)**

TESIS

Para optar el Título Profesional de Licenciado en Estadística

AUTOR

Luis Alberto BAUTISTA BAUTISTA

ASESOR

Rosa Fátima MEDINA MERINO

Lima, Perú

2016



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Bautista, L. (2016). *Modelos de cambio de régimen de transición determinística: Modelos SETAR (Self-Exciting Threshold Autoregressive)*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Escuela Profesional de Estadística]. Repositorio institucional Cybertesis UNMSM.

1226



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú, DECANA DE AMÉRICA
FACULTAD DE CIENCIAS MATEMÁTICAS



ESCUELA PROFESIONAL DE ESTADÍSTICA

ACTA DE SUSTENTACIÓN DE TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE
LICENCIADO EN ESTADÍSTICA

149

En la Ciudad Universitaria, Facultad de Ciencias Matemáticas, siendo las 16:35 horas del día 31 de agosto del año 2016 se reunieron los docentes designados como miembros del Jurado:

- Mg. Wilfredo Eugenio Domínguez Cirilo (Presidente)
- Lic. Grabiela Yolanda Montes Quintana (Miembro)
- Mg. Rosa Fátima Medina Merino (Miembro Asesor)

Para la sustentación de la Tesis intitulada "MODELOS DE CAMBIO DE RÉGIMEN DE TRANSICIÓN DETERMINÍSTICA: MODELOS SETAR (SELF-EXCITING THRESHOLD AUTOREGRESSIVE)", presentada por el Bachiller Luis Alberto Bautista Bautista, para obtener el Título Profesional de Licenciado en Estadística.

Luego de la exposición de la Tesis, el Presidente invitó al expositor a dar respuesta a las preguntas formuladas.

Realizada la evaluación correspondiente por los miembros del jurado, el expositor mereció la aprobación SOBRESALIENTE, con un calificativo promedio de Dieciocho (18) (letras y números).

A continuación los miembros del jurado, dan manifiesto que el participante Bachiller Luis Alberto Bautista Bautista, en virtud de haber aprobado la sustentación de su tesis, será propuesto para que se le otorgue el Título Profesional de Licenciado en Estadística.

Siendo las 17:20 horas, se levantó la Sesión, firmando para constancia la presente Acta en tres (3) copias originales.

Lic. Grabiela Yolanda Montes Quintana
Miembro

Mg. Rosa Fátima Medina Merino
Miembro Asesor

Mg. Wilfredo Eugenio Domínguez Cirilo
Presidente

MODELOS DE CAMBIO DE RÉGIMEN DE TRANSICIÓN DETERMINÍSTICA: MODELOS SETAR

LUIS ALBERTO BAUTISTA BAUTISTA

Tesis presentada a consideración del Cuerpo Docente de la Facultad de Ciencias Matemáticas, de la Universidad Nacional Mayor de San Marcos, como parte de los requisitos para obtener el Título Profesional de Licenciado en Estadística.

Aprobada por:

Mg. Wilfredo Eugenio Domínguez Cirilo
(Presidente)

Lic. Grabiela Yolanda Montes Quintana
(Miembro – Jurado)

Mg. Fátima Medina Merino
(Miembro – Asesor)

Lima – Perú

Octubre – 2016

FICHA CATALOGRÁFICA

BAUTISTA BAUTISTA, LUIS ALBERTO

Modelos de cambio de régimen de transición determinística: Modelos

SETAR (Self-Exciting Threshold Autoregressive), (Lima) 2016.

vii, 149 p., 29.7 cm, (UNMSM, Licenciado, Estadística, 2016).

Tesis, Universidad Nacional Mayor de San Marcos,

Facultad de Ciencias Matemáticas 1. Estadística

I. UNMSM/FACULTAD DE CIENCIAS MATEMÁTICAS

DEDICATORIA

A mi familia por el apoyo que me brindan día tras día
para seguir adelante en este largo camino de la vida universitaria.

A mi hermanita María Lourdes, por ser la mayor inspiración
para creer en los sueños.

AGRADECIMIENTOS

Quiero expresar mi más sincero agradecimiento a mi asesora Mg. Fátima Medina Merino. Por su apoyo, voluntad y dedicación en el desarrollo de la presente tesis.

Mi agradecimiento a todos mis profesores de la Facultad de Ciencias Matemáticas de la UNMSM por sus enseñanzas.

Finalmente mi agradecimiento a la UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS, por permitir sentirme orgulloso de llamarla Alma Mater.

RESUMEN

MODELOS DE CAMBIO DE RÉGIMEN DE TRANSICIÓN DETERMINÍSTICA: MODELOS SETAR (SELF-EXCITING THRESHOLD AUTOREGRESSIVE)

BR. LUIS ALBERTO BAUTISTA BAUTISTA

OCTUBRE – 2016

ASESORA : Mg. Fátima Medina Merino

TÍTULO OBTENIDO : Licenciado en Estadística

En el presente trabajo de investigación se desarrolló los modelos de cambio de régimen de transición determinística denominado modelos SETAR, cuya principal característica es su aplicación en series que presentan ciclos límites, frecuencias dependientes de amplitud y fenómenos de salto, que no pueden ser estimadas mediante los modelos lineales de series de tiempo. La metodología empleada se basa en la propuesta por (Tong, 1978) y (Tsay, 1989) que incluye la identificación y estimación de los parámetros estructurales. Por otra parte en la hipótesis de investigación se plantea una comparación entre los modelos SETAR y los modelos GARCH, a fin de determinar el modelo más apropiado para modelar las series económicas.

Una contribución del trabajo fue la implementación de una metodología para estimar el umbral óptimo mediante la minimización de criterios de información (AIC y BIC).

El trabajo presenta el desarrollo de ambos métodos y ha finalizado con la aplicación en la cual se compararon las dos metodologías, demostrando que las mejores estimaciones son obtenidas con los modelos SETAR.

PALABRAS CLAVE: **MODELOS SETAR**
 CICLO LÍMITE
 UMBRAL
 RÉGIMEN
 MODELOS GARCH

ABSTRACT

REGIME CHANGE MODELS DETERMINISTIC TRANSITION: (SELF-EXCITING THRESHOLD AUTOREGRESSIVE)

BR. LUIS ALBERTO BAUTISTA BAUTISTA

OCTOBER – 2016

ADVISOR : Mg. Fatima Medina Merino
DEGREE OBTAINED : Bachelor's degree in Statistics

In the present research is developing models of regime change deterministic transition called SETAR models, whose main characteristic is its application in series exhibiting limit cycles, dependent frequency and amplitude phenomena jump, which cannot be estimated by the linear time series models. The methodology is based on a proposal by (Tong, 1978) and (Tsay, 1989) which includes the identification and estimation of structural parameters. Moreover, in the research hypothesis it raises a comparison between SETAR models and GARCH models, to determine the most appropriate model to model economic series.

A contribution of this work was the implementation of a methodology to estimate the optimal threshold by minimizing information criteria (AIC and BIC).

The paper presents the development of both methods and finished with the application in which the two methods were compared, showing that the best estimates are obtained with SETAR models.

KEYWORDS: SETAR MODELS
LIMIT CYCLE
THRESHOLD
REGIME
GARCH MODELS

CONTENIDO

RESUMEN

ABSTRACT

INTRODUCCIÓN

CAPÍTULO I. PLANTEAMIENTO DE LA INVESTIGACIÓN18

1.1 SITUACIÓN PROBLEMÁTICA.....	18
1.2 FORMULACIÓN DEL PROBLEMA.....	19
1.3 JUSTIFICACIÓN DE LA INVESTIGACIÓN.....	19
1.4 LIMITACIONES DE LA INVESTIGACIÓN	20
1.5 CONTRIBUCIÓN E IMPACTO.....	20
1.6 OBJETIVOS DE LA INVESTIGACIÓN	21
1.6.1 Objetivo general.....	21
1.6.2 Objetivos específicos	21

CAPÍTULO II. MARCO TEÓRICO23

2.1 ANTECEDENTES DEL PROBLEMA	23
2.1.1 Antecedentes internacionales.....	23
2.1.2 Antecedentes nacionales	25
2.2 CONCEPTOS BÁSICOS.....	27
2.2.1 Procesos estocásticos	27
2.2.2 Características de un proceso estocástico	28
2.2.3 Tipo de procesos estacionarios	28
2.2.4 Función de autocovarianza	29
2.2.5 Función de autocorrelación.....	29
2.2.6 Serie temporal.....	30
2.2.7 Procesos autorregresivos.....	30
2.3 MODELO LINEAL Y NO LINEAL EN EL TIEMPO.....	35
2.3.1 Proceso de ruido blanco.....	35
2.3.2 Proceso lineal general	36
2.3.3 Modelo lineal en el tiempo	37
2.3.4 Modelo no lineal en el tiempo	38
2.4 HIPÓTESIS.....	41
2.5 TERMINOLOGÍA	41

CAPÍTULO III. METODOLOGÍA.....	43
3.1 DISEÑO DE LA INVESTIGACIÓN.....	43
3.2 TIPO DE INVESTIGACIÓN.....	43
3.3 POBLACIÓN DE ESTUDIO.....	43
3.4 UNIDAD DE ANÁLISIS.....	43
3.5 TAMAÑO DE MUESTRA	44
3.6 SELECCIÓN DE LA MUESTRA	44
3.7 TÉCNICAS DE RECOLECCIÓN DE DATOS.....	44
3.7.1 Fuentes de información.....	44
3.7.2 Procedimiento de recolección de datos.....	45
3.8 PROCESAMIENTO Y ANÁLISIS DE LA INFORMACIÓN.....	46
3.9 MÉTODOS ESTADÍSTICOS.....	46
3.9.1 Análisis de la linealidad en el tiempo	47
3.9.2 Modelos por umbrales	51
3.9.3 Modelos de cambio de régimen de transición determinística	52
3.10 MODELO SETAR	54
3.10.1 Definición del modelo SETAR.....	55
3.10.2 Estimación del modelo SETAR.....	56
3.10.3 Validación del modelo SETAR	62
3.10.4 Pronósticos con el modelo SETAR.....	62
3.11 MODELO ARCH.....	67
3.12 MODELO GARCH.....	69
CAPÍTULO IV. APLICACIÓN.....	72
4.1 ANÁLISIS EXPLORATORIO DE LA SERIE	73
4.2 PRUEBA DE LINEALIDAD	77
4.2.1 Análisis exploratorio de la linealidad	77
4.2.2 Prueba de la linealidad de tipo umbral (Prueba de Tsay).....	79
4.2.3 Prueba de la linealidad (Prueba de Hansen).....	80
4.3 DETERMINACIÓN DE LOS PARÁMETROS ESTRUCTURALES	81
4.3.1 Determinación de la variable umbral	81
4.3.2 Región de variación para los valores del umbral	82
4.3.3 Localización óptima de los valores umbrales	84
4.3.4 Localización óptima del valor umbral usando la prueba de Hansen.....	85
4.4 ESTIMACIONES MÍNIMOS CUADRÁTICOS DEL MODELO SETAR	87
4.5 VALIDACIÓN DEL MODELO SETAR	88

4.5.1 Prueba de normalidad de los residuos.....	89
4.5.2 Prueba de incorrelación de los residuos.....	91
4.6 PRONÓSTICOS EN LOS MODELOS SETAR.....	94
4.7 APLICACIÓN MEDIANTE EL MODELO GARCH.....	96
4.7.1 Descripción de la serie.....	96
4.7.2 Identificación del modelo ARMA.....	97
4.7.3 Evaluación de la presencia de heterocedasticidad condicional.....	99
4.7.4 Identificación del modelo GARCH.....	100
4.7.5 Estimación del modelo.....	101
4.7.6 Validación del modelo.....	102
4.7.7 Pronósticos del modelo.....	108
CAPÍTULO IV. DISCUSIÓN DE LOS RESULTADOS.....	112
CAPÍTULO VI. CONCLUSIONES.....	116
CAPÍTULO VII. RECOMENDACIONES.....	119
CAPÍTULO VIII. REFERENCIAS BIBLIOGRÁFICAS.....	121
CAPÍTULO IX. ANEXOS.....	125

INDICE DE CUADROS

Cuadro 2.1: Modelos no lineales.....	39
Cuadro 4.1: Identificación del orden del proceso.....	77
Cuadro 4.2: Prueba de linealidad de Tsay.....	80
Cuadro 4.3: Prueba de linealidad de Hansen.....	81
Cuadro 4.4: Determinación de la variable umbral mediante S-PLUS.....	82
Cuadro 4.5: Determinación de la variable umbral mediante R-PROJECT.....	82
Cuadro 4.6: Localización óptima de los valores del umbral.....	85
Cuadro 4.7: Valor del umbral según la prueba de Hansen.....	86
Cuadro 4.8: Estimación de los coeficientes de cada régimen.....	88
Cuadro 4.9: Normalidad de los residuos.....	89
Cuadro 4.10: Incorrelación de los residuos.....	91
Cuadro 4.11: Pronósticos usando el método de Montecarlo.....	95
Cuadro 4.12: Identificación del orden del proceso autorregresivo.....	98
Cuadro 4.13: Estimación del modelo AR(6).....	99
Cuadro 4.14: Efecto ARCH.....	99
Cuadro 4.15: Identificación del modelo GARCH.....	101
Cuadro 4.16: Estimación del modelo ARMA (6,0)-GARCH (1,1).....	101
Cuadro 4.17: Efecto ARCH de los residuos del modelo.....	105
Cuadro 4.18: Criterio de bondad de ajuste.....	106
Cuadro 4.19: Pronósticos del modelo ARMA (6,0)-GARCH (1,1).....	110

Cuadro 5.1: Comparación de pronósticos.....	113
Cuadro 5.2: Indicadores de bondad de ajuste.....	114
Cuadro 9.1: Prueba de Dickey-Fuller aumentado (ADF).....	142
Cuadro 9.2: Periodograma de la serie Z_t	143

INDICE DE FIGURAS

Figura 2.1: Círculo unitario.....	31
Figura 2.2: Identificación del orden del proceso autorregresivo.....	34
Figura 2.3: Filtro lineal.....	36
Figura 3.1: Diagramas de dispersión directa	50
Figura 3.2: Suavizado no paramétrico.....	51
Figura 3.3: Esquematización del algoritmo de Tsay (1989).....	60
Figura 3.4: Algoritmo para la obtención del umbral óptimo.....	61
Figura 4.1: PBI trimestral, 1980-2014.....	73
Figura 4.2: Variación porcentual del PBI trimestral, 1981-2014.....	74
Figura 4.3: Periodograma de la serie Z_t	75
Figura 4.4: Función de autocorrelación de la serie Z_t	76
Figura 4.5: Diagramas de dispersión directa de la serie Z_t	78
Figura 4.6: Suavizado no paramétrico de la serie Z_t-d	79
Figura 4.7: Posibles valores umbrales para el retardo $d=1$	83
Figura 4.8: Intervalo de confianza para el valor umbral.....	86
Figura 4.9: Prueba Q-Q Normal de los residuos.....	90
Figura 4.10: Histograma de los residuos y curva de densidad normal.....	90
Figura 4.11: Correlograma de los residuos.....	92
Figura 4.12: Espectro de los residuos.....	93
Figura 4.13: Serie original vs Serie estimada.....	94

Figura 4.14: Estimación de la serie y pronósticos.....	95
Figura 4.15: Variación porcentual del PBI, 1981-2014.....	97
Figura 4.16: Función de autocorrelación de la serie Z_t	98
Figura 4.17: Autocorrelaciones de los residuos al cuadrado.....	100
Figura 4.18: Prueba de normalidad e histograma de los residuos.....	103
Figura 4.19: Correlograma de los residuos.....	104
Figura 4.20: Correlograma de los residuos al cuadrado.....	104
Figura 4.21: Espectro de los residuos del modelo ARMA(6,0)-GARCH(1,1).....	107
Figura 4.22: Serie original vs Serie estimada.....	108
Figura 4.23: Predicción de la volatilidad (Predicción de la media y la varianza).....	109
Figura 9.1: Densidad espectral (Ejemplo).....	126
Figura 9.2: Densidad espectral del ruido blanco.....	127

ANEXOS

ANEXO 1: DESARROLLO DE ALGUNOS CONCEPTOS TRATADOS EN LA INVESTIGACIÓN.....	125
ANEXO 2: SISTEMA DE CUENTAS NACIONALES TRIMESTRALES.....	132
ANEXO 3: METODOLOGÍA DE MEDICIÓN DEL PBI TRIMESTRAL.....	136
ANEXO 4: PROGRAMAS USADOS PARA EL MODELAMIENTO DE DATOS.....	139
ANEXO 5: PRUEBA DE DICKEY-FULLER AUMENTADO (ADF).....	142
ANEXO 6: PERIODOGRAMA DE LA SERIE Z_t	143
ANEXO 7: CÓDIGO EN S-PLUS 8.0.....	144
ANEXO 8: CÓDIGO EN R-PROJECT 3.3.4.....	147

INTRODUCCIÓN

Actualmente la metodología de Box & Jenkins (1970) es la más utilizada para modelar variables que representan situaciones reales de la economía o de las finanzas. Sin embargo, hay una serie de fenómenos presentes en la evolución de series temporales encontradas en la práctica que no pueden ser representadas satisfactoriamente mediante modelos lineales. Bajo este contexto surgen los modelos no lineales cuya principal característica es mejorar la calidad del ajuste y en algunos casos la capacidad predictiva.

En el presente trabajo de investigación se implementará la metodología del modelo SETAR propuesta por (Tong, 1978) y (Tsay, 1989) para modelar la serie “Variación porcentual del PBI trimestral en el Perú, 1981-2014”, y a partir de estos resultados, modelar el comportamiento no lineal de la serie usando los aplicativos estadísticos R y S-PLUS.

Por último el presente trabajo de investigación comprende de cuatro capítulos principales, el Capítulo I comprende el Planteamiento de la investigación: Situación problemática, formulación del problema, justificación, contribución e impacto y los objetivos, el Capítulo II contiene el Marco teórico: Antecedentes y algunos conceptos básicos para la comprensión de los modelos no lineales, el Capítulo III comprende la metodología de investigación usada y el desarrollo de los métodos estadísticos (Modelos SETAR, ARCH y GARCH) y por último en el Capítulo IV se muestran los principales resultados obtenidos.

CAPÍTULO I
PLANTEAMIENTO DE LA
INVESTIGACIÓN

CAPÍTULO I. PLANTEAMIENTO DE LA INVESTIGACIÓN

1.1 SITUACIÓN PROBLEMÁTICA

Un gran número de series de tipo económico o financiero presentan cambios de nivel, así como períodos de estabilidad seguidos de otros de alta volatilidad, entendiendo por esta última como la variabilidad de la serie en torno a un valor medio, medida por la varianza, es decir, es muy alta en determinados tramos de la muestra frente a otros períodos en los que, al menos aparentemente, es menor (Pérez & Velazqués, 2004).

La metodología clásica para el análisis de las series temporales “Enfoque de dominio del tiempo” desarrollado por Box y Jenkins (1970) permite la identificación y estimación de los modelos ARMA (Autoregresivos y de medias móviles). Estos modelos parten de los supuestos de estacionariedad y linealidad.

Ante esta situación es natural que la investigación en el campo de las series temporales se haya extendido a modelos más generales en los que no se asumen los supuestos de linealidad y estacionariedad. Estudios realizados sobre diferentes conjuntos de datos indican que los modelos no lineales y no estacionarios son más flexibles para capturar las características de los datos y, en algunos casos, pueden mejorar la calidad del ajuste y la capacidad de predicción del modelo.

(Tiao & Tsay, 1994) consideran que estos avances no significan que los modelos lineales sean menos útiles, al contrario, los modelos lineales nos proporcionan una primera aproximación que puede resultar de gran ayuda como primera estimación de algunos parámetros.

Por otra parte, se han propuesto diferentes métodos estadísticos para modelar no linealidades resultantes de conductas asimétricas a lo largo del ciclo económico. En particular se ha empleado los modelos SETAR introducido por (Tong, 1978).

A pesar de las dificultades aún presentes en los algoritmos propuestos, relacionados con la estimación de los parámetros estructurales, la modelación a través de un modelo SETAR permite reproducir las características no lineales de la serie y en algunos casos, elaborar mejores pronósticos. Es por este motivo que en el presente trabajo de investigación se desarrollará el modelo SETAR.

1.2 FORMULACIÓN DEL PROBLEMA

Por lo planteado anteriormente, el presente trabajo de investigación pretende desarrollar la siguiente interrogante:

¿Es posible implementar la metodología de Tsay para aplicar el modelo SETAR en series económicas?

La propuesta de solución aborda el problema de identificación, estimación, validación y la obtención de los pronósticos bajo la misma metodología propuesta por (Tsay, 1989).

1.3 JUSTIFICACIÓN DE LA INVESTIGACIÓN

Justificación teórica: Debido a los problemas de los años 80 a causa de la crisis financiera cuyas repercusiones fueron mundiales, las series de tiempo económicas presentaron comportamientos no lineales con alteraciones en el cambio de régimen, los cuales no fueron tomados en cuenta y fueron modeladas como modelos ARIMA primero y más adelante como modelos de la familia ARCH (Autoregressive conditional heteroskedasticity).

Justificación práctica: En nuestro medio se ha venido aplicando de forma errónea los modelos de heterocedasticidad condicional a toda serie que presenta alta volatilidad, incluso con deficiencias en la incorporación de periodicidades a largo plazo o ciclos. Esta situación ocasiona pronósticos con horizontes no sostenibles con pobre modelamiento de ciclos, lo cual implica el deficiente uso del modelamiento de procesos estocásticos.

Justificación metodológica: Desde el punto de vista metodológico, el desarrollo del presente trabajo de investigación servirá como guía para estudios posteriores a todo investigador en las distintas áreas que pretende aplicar este modelo, para los pronósticos o en su defecto para profundizar el tema bajo otros enfoques.

1.4 LIMITACIONES DE LA INVESTIGACIÓN

Si bien puede considerarse una justificación, la presente investigación no cuenta con evidencia empírica a nivel nacional o estudios anteriores que respalden el problema principal de la investigación. Además, la escasez de programas estadísticos que incorporan los modelos SETAR dificulta el modelamiento de los datos.

1.5 CONTRIBUCIÓN E IMPACTO

La contribución metodológica del presente trabajo de investigación es la propuesta de desarrollo sistemático y metódico del modelo SETAR para estimación y pronóstico de distintas series de tiempo, debido a los beneficios y ventajas que posee el modelo SETAR frente a otros modelos de pronóstico. Aun cuando la aplicación se da en series económicas, los resultados de esta investigación son aplicables para series de tiempo de todo índole, ecológicas, económicos, biológicos, financieros, etc.

Además se propone una metodología para determinar los valores umbrales óptimos mediante la utilización de los criterios de información AIC y BIC.

Finalmente se desarrollarán adaptaciones con los programas estadísticos R y S-PLUS a fin de realizar modelizaciones SETAR con aplicación de umbrales óptimos.

Se considera que debido a su aplicabilidad, facilidad para modelar y necesidad de aplicación los modelos SETAR y la adaptación desarrollada serán de gran impacto tanto a nivel científico como aplicativo.

1.6 OBJETIVOS DE LA INVESTIGACIÓN

1.6.1 Objetivo general

Implementar la metodología de Tsay para aplicar el modelo SETAR en series económicas.

1.6.2 Objetivos específicos

- Estudiar las características del modelo SETAR en comparación con los modelos de la familia ARCH.
- Analizar los métodos de identificación, estimación, validación y pronósticos de los modelos SETAR.
- Proponer una metodología para determinar los valores umbrales óptimos mediante la utilización de los criterios de información AIC y BIC.
- Innovar el empleo de los modelos SETAR para el modelamiento y pronósticos de las series económicas con ayuda de los aplicativos estadísticos R y S-PLUS.

CAPÍTULO II

MARCO TEÓRICO

CAPÍTULO II. MARCO TEÓRICO

2.1 ANTECEDENTES DEL PROBLEMA

Los modelos autorregresivos por umbrales (TAR), fueron propuestos por (Tong, 1978) y discutidos en detalle por (Tong & Lim, 1980) y (Tong, 1983), permiten la modelización de series de tiempo no lineales. En estos modelos, la variable dependiente tiene un comportamiento diferente en función del régimen en que se encuentre la variable umbral (variable que define los regímenes).

En los modelos TAR la variable umbral es una variable observable. Es decir, el régimen observado en el momento t es determinado por el valor que asume una variable umbral, en relación a un valor umbral (r) previamente estimado de esa variable.

Su principal características es que permite usar diferentes modelos autorregresivos (AR) separados por una variable umbral.

2.1.1 Antecedentes internacionales

El comportamiento de las variables económicas, ambientales u otras tienen comportamientos asimétricos durante las fases de expansión y contracción, en cuanto a variables económicas fueron importante los estudios de la década de los años 80, como consecuencia de las dificultades para explicar episodios de crisis y del inapropiado uso de métodos lineales en: el análisis de los regímenes del ciclo, la caracterización de sus etapas y la descripción de procesos de cambio estructural.

La existencia de asimetrías y no linealidades en el comportamiento de las variables económicas, independientemente de la fuente que la produzca, implica que su estructura probabilística durante la fase de crecimiento no corresponde a la de su fase recesiva.

(Pérez & Velazqués, 2004) publicaron un artículo denominado: “Análisis de cambio de régimen en series de tiempo no lineales utilizando modelos TAR aplicado al PBI trimestral de España”, se consideró el comportamiento dinámico del crecimiento trimestral del PBI español de 1970 a 1998. En épocas de recesión, la variable “crecimiento del PBI” es más aguda que en épocas de expansión; esto justifica la aplicación de modelos no lineales, los cuales ayudan a captar tal característica y permiten hacer una interpretación económica más adecuada para esta variable.

(Hoyos, Ramos, & Vivas, 2009) publicaron un artículo en el cual aplicaron el modelo SETAR para el PBI Colombiano. El presente artículo implementó la metodología SETAR propuesta por (Tsay, 1989) y (Tong, 1978) para probar la hipótesis de no linealidad del ciclo económico del PBI colombiano, y a partir de estos resultados, modelar el comportamiento no lineal de la serie.

(Moreno Lopez, 2010) desarrolló una tesis de maestría en estadística denominada: “Una aplicación del modelo TAR (Threshold Autoregressive) en series de tiempo financieras”. La tesis investigó el desempeño de los modelos TAR y de la metodología de análisis propuesta por (Nieto, 2005), y además realizar una comparación de este modelo con los modelos GARCH en términos de algunos de los momentos marginales y condicionales del proceso estocástico subyacente.

2.1.2 Antecedentes nacionales

(**Bazan Palomino, 2009**) desarrolló una tesis de licenciatura en economía denominada: “Modelos de cambio de régimen: Relación no-lineal entre el crédito bancario al sector privado y el producto bruto interno para Perú 1994-2008”. En el presente trabajo se investigó la relación no-lineal entre el crédito privado agregado del sector bancario y el producto bruto interno para Perú con data trimestral.

Se examinó la existencia de diferentes regímenes que gobiernen ambas variables de manera independiente y conjunta, asociados a diferentes estados de la economía y cambios en el sistema financiero. Se utilizaron dos modelos de cambio de régimen, LSTAR y Markov Switching, porque permiten estimar parámetros cambiantes acorde a las diferentes etapas de la economía. Los resultados muestran que existen por lo menos dos regímenes, el primero de contracción y segundo de expansión, los cuales condicionan el comportamiento del crédito y el producto. La tesis presentó resultados y aplicaciones no incluye desarrollo estadístico del método.

(**Romero Madrid, 2013**) desarrolló una tesis de licenciatura en economía denominada: “Modelación del comportamiento de la tasa de interés activa promedio en moneda nacional del sistema bancario peruano”. El presente trabajo de investigación realizó un análisis econométrico del comportamiento dinámico de la Tasa de Interés Activa Promedio en Moneda Nacional del sistema bancario peruano (TAMN), para el periodo comprendido desde enero de 1996 hasta diciembre de 2011.

El objetivo fue demostrar que la introducción de no linealidades en la modelación de la TAMN, nos permite entender de mejor manera su comportamiento dinámico, a diferencia de la modelación lineal, ya que al incorporar la no linealidad se captura la idea intuitiva de que el comportamiento de los agentes cambia en función del estado de la economía. A fin de lograr los objetivos planteados, se estimó un modelo Autorregresivo (AR), como la representación lineal de la TAMN, el cual servirá de punto de referencia y comparación, y dos modelos no lineales, un modelo de cambio de régimen de transición suave (STAR) y un modelo de Cambio de Régimen del tipo Markov Switching (MS). La tesis presentó resultados y aplicaciones no incluye desarrollo estadístico del método.

(Agüero Palacios & Elescano Rojas, 2004) publicaron un artículo denominado: “MODELOS ARCH: UNA APLICACIÓN EN EL PRONÓSTICO DE LA VOLATILIDAD DE ACCIONES COTIZADAS EN LA BOLSA DE VALORES DE LIMA” en el que se propone un esquema para realizar el análisis de datos generados por una familia de procesos estocásticos denominados Procesos con Heterocedasticidad Condicional Autorregresiva-ARCH, los cuales son ampliamente utilizados para la predicción de la volatilidad de series financieras. Se ajustó un modelo ARCH para pronosticar la volatilidad de las cotizaciones de las acciones de la empresa minera Atacocha, utilizando la serie de datos observados desde 1992 a 2003.

2.2 CONCEPTOS BÁSICOS

2.2.1 Procesos estocásticos

Definición: Conjunto de variables aleatorias $\{ Z_t ; t \in T \}$, definidas en un espacio de probabilidades (Ω, α, \wp) , donde el índice t toma valores en un conjunto T , que en general se suele identificar con el tiempo (Morettin & Maria de Castro Toloí, 2006).

Si T es discreto, el proceso es en tiempo discreto. Si T es continuo, el proceso es en tiempo continuo.

Un proceso estocástico queda determinado si conocemos la función de distribución de las variables aleatorias que lo componen y todas las posibles funciones de distribución conjuntas de dichas variables.

$$\begin{aligned} & \cdot F(Z_t) \quad \forall t \in \mathbb{Z} \\ & \cdot F(Z_{t_1} \dots Z_{t_n}) \quad \forall t_1 \dots t_n \in \mathbb{Z} \quad \forall n \in \mathbb{N} \end{aligned}$$

La determinación de las características de un proceso estocástico a partir de las funciones de distribución es, en general, un procedimiento complicado, por lo que se acostumbra a caracterizarlos a partir de los momentos de primer y segundo orden (media y covarianza).

Momento de primer orden: $\mu_t = E(Z_t) \quad t \in \mathbb{Z}$

Momento de segundo orden: $\gamma_{t,s} = Cov(Z_t, Z_s) \quad t, s \in \mathbb{Z}$

2.2.2 Características de un proceso estocástico

Gaussiano: Las variables aleatorias que conforman el proceso estocástico $\{Z_t\}$ provienen de una distribución normal.

No Markoviano: La ocurrencia en el momento “t” se relaciona por la ocurrencia en el momento “t-k”. Es decir existe asociación entre el momento actual y el pasado del proceso.

Estacionario: Este término está asociado con el concepto de “estabilidad” en el tiempo. Entonces, para poder estimar las características del proceso (medias, varianzas o autocovarianzas) necesitamos suponer que son estables a lo largo del tiempo; es decir, que el proceso sea estacionario. Se conocen dos formas de estacionariedad.

2.2.3 Tipo de procesos estacionarios

Proceso estacionario en sentido estricto

Un proceso estocástico $\{Z_t\}_{t \in \mathbb{Z}}$ es estacionario en sentido estricto si se mantienen estables todas las posibles relaciones entre las variables aleatorias que lo componen, de manera que solo dependen de la distancia que las separa:

$$\begin{aligned} \cdot F(Z_t) &= F(Z) \quad \forall t \in \mathbb{Z} \\ \cdot F(Z_{t_1} \dots Z_{t_n}) &= F(Z_{t_1+k} \dots Z_{t_n+k}) \quad \forall t_1 \dots t_n \in \mathbb{Z}, \forall k \in \mathbb{Z}, \forall n \in \mathbb{N} \end{aligned}$$

Proceso estacionario en sentido débil

Un proceso estocástico es estacionario en sentido débil cuando su media y su varianza permanecen constantes a lo largo del tiempo y las relaciones lineales entre las variables solo dependen de la distancia que las separa:

- $E(Z_t) = \mu \quad \forall t \in \mathbb{Z}$
- $Var(Z_t) = \sigma^2 \quad \forall t \in \mathbb{Z}$
- $Cov(Z_{t_1}, Z_{t_2}) = Cov(Z_{t_1+k}, Z_{t_2+k}) \quad t_1, t_2 \in \mathbb{Z}, \forall k \in \mathbb{Z}$

2.2.4 Función de autocovarianza

Es una función de dos argumentos que describe las covarianzas entre dos variables del proceso en dos instantes cualesquiera, se define como:

$$\gamma_k = Cov(Z_t, Z_{t+k}) = E(Z_t - \mu)(Z_{t+k} - \mu) \quad t \in \mathbb{Z}, \forall k \in \mathbb{Z}$$

2.2.5 Función de autocorrelación

Función de autocorrelación simple

Se denomina autocorrelación de orden k , ρ_k , a la correlación de cualesquiera de dos variables aleatorias del proceso estocástico, distanciadas k instantes de tiempo.

$$\rho_{(k)} = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)}\sqrt{Var(Z_{t+k})}} = \frac{\gamma_{(k)}}{\gamma_{(0)}} \quad \forall k \in \mathbb{Z}, \forall t \in \mathbb{Z}$$

Propiedades:

- $\rho_{(k)} = \rho_{(-k)}$
- $\rho_{(0)} = 1$
- $|\rho_{(k)}| \leq 1$

Función de autocorrelación parcial

Se denomina autocorrelación parcial de orden k , ϕ_{kk} , a la correlación entre dos variables aleatorias del proceso estocástico distanciadas k instantes de tiempo, pero sin considerar la existencia de las variables aleatorias intermedias. Es decir, para calcular la autocorrelación parcial entre Z_t y Z_{t+k} eliminamos la influencia que sobre ambas tienen $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$.

$$\phi_{kk} = Corr(Z_t, Z_{t+k} | Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1})$$

Las funciones de autocorrelación simple y parcial, tal y como veremos más adelante, constituyen uno de los instrumentos clave para ajustar el modelo que genera una serie temporal.

2.2.6 Serie temporal

Una serie temporal es la realización de un proceso estocástico en tiempo discreto, donde los elementos de T están ordenados y corresponden a instantes equidistantes del tiempo, que tienen un inicio y un fin (Medina Merino, 2011).

Una serie temporal describe la evolución aleatoria de una variable en el tiempo.

2.2.7 Procesos autorregresivos

Los procesos autorregresivos (AR) parten de identificar qué tipo de proceso estocástico genera la serie de tiempo utilizando rezagos de la misma variable sin recurrir a variables explicativas exógenas. Estos modelos fueron aplicados con fines de pronóstico por Box y Jenkins (1970) y se han desempeñado adecuadamente en comparación con otros modelos más complejos debido a la capacidad para describir fenómenos donde los eventos producen un efecto inmediato que solo duran un periodo corto de tiempo.

Definición de un proceso autorregresivo de orden p:

Un proceso se dirá autorregresivo de orden p o AR (p) si podemos representar a la serie de tiempo como:

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + \varepsilon_t, \text{ donde } \varepsilon_t \sim RB(0, \sigma_\varepsilon^2)$$

$$\text{De forma equivalente } \tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + \varepsilon_t$$

Si definimos el operador $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$ podemos escribir el modelo

$$\text{AR}(p) \text{ como: } \phi(B)Z_t = \varepsilon_t$$

Es decir, el valor de la variable Z_t depende de sus valores pasados y de la innovación al tiempo t .

Teorema 1: Un proceso autorregresivo finito AR (p) es estacionario si y solo si el módulo de las raíces del polinomio autorregresivo $\phi_p(B)$ está fuera del círculo unitario. Las p raíces de la ecuación $\phi_p(z_i) = 0$, $i = 1, 2, \dots, p$ deben cumplir que: $|z_i| > 1$, donde z_i son raíces del polinomio autorregresivo.

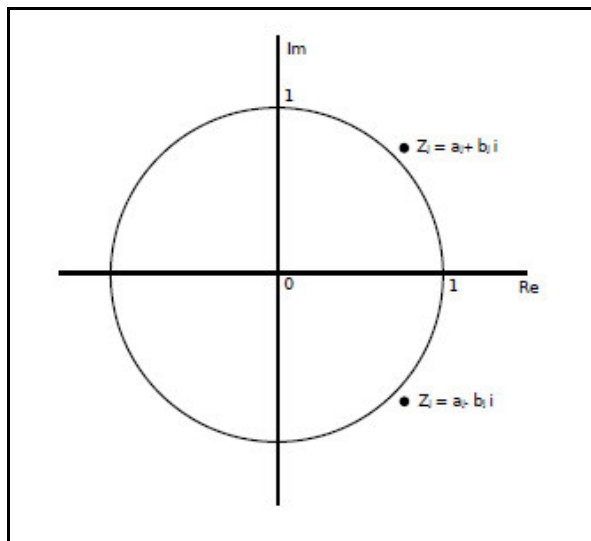


Figura 2.1: Círculo unitario.

Fuente: Elaboración propia.

Demostración:

Dado las raíces de $\phi(z)$; $\zeta_1, \zeta_2, \dots, \zeta_p$, tener en cuenta que algunas de las raíces pueden ser iguales, por este supuesto debemos ordenar según sus magnitudes en el orden de aumento, de modo que: $1 < |\zeta_1| < \dots < |\zeta_p|$

Definamos $|\zeta_1| = 1 + \varepsilon$, para algún $\varepsilon > 0$, entonces Z se puede escribir como:

$$|z| < 1 + \varepsilon, \phi(z) \neq 0$$

En consecuencia, hay una expansión en series de potencia para $\phi(z)^{-1}$ con $|z| < 1 + \varepsilon$, es decir:

$$\frac{1}{\phi(z)} = \sum_{i=0}^{\infty} \psi_i Z^i \quad (2.1)$$

La expresión anterior es una serie convergente para $|z| < 1 + \varepsilon$, ahora elegimos

$0 < \delta < \varepsilon$ y se reemplaza $z = 1 + \delta$ en la ecuación (2.1). Entonces:

$$\frac{1}{\phi(1 + \delta)} = \sum_{i=0}^{\infty} \psi_i (1 + \delta)^i < \infty$$

Por lo tanto, existe una constante $M > 0$ de manera que $\forall i$

$$|\psi_i (1 + \delta)^i| < M, \text{ es decir } \forall i \quad |\psi_i| < M (1 + \delta)^{-i}$$

De esta manera; $\sum_{i=0}^{\infty} |\psi_i| < \infty$ como resultado de ello el siguiente proceso se puede

$$\text{escribir como: } Y_t = \frac{1}{\phi(B)} Z_t = \sum_{i=0}^{\infty} \psi_i Z_{t-i}$$

Bajo esta última expresión el proceso está bien definido y por lo tanto es estacionario.

Función de autocorrelación simple de un AR (p)

La función de autocorrelación simple de un proceso autorregresivo de orden p es:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad k > 0$$

A continuación se detalla el proceso de obtención del resultado anterior.

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t \quad (2.2)$$

Multiplicando (2.2) por \tilde{Z}_{t-k} se obtiene

$$\tilde{Z}_{t-k} \tilde{Z}_t = \phi_1 \tilde{Z}_{t-k} \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-k} \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-k} \tilde{Z}_{t-p} + \tilde{Z}_{t-k} a_t \quad (2.3)$$

Al tomar los valores esperados en (2.3) se obtiene la ecuación de diferencia

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \quad k > 0 \quad (2.4)$$

Dividiendo (2.4) por γ_0 se obtiene

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad k > 0 \quad (2.5)$$

Función de autocorrelación parcial de un AR (p)

La función de autocorrelación parcial es una herramienta que explota el hecho de que, mientras un proceso AR (p) tiene una función de autocorrelación que es infinito en extensión, se puede describir por su propia naturaleza en términos de p funciones no nulas de las autocorrelaciones.

$$\rho_j = \phi_{k1} \rho_{j-1} + \dots + \phi_{k(k-1)} \rho_{j-k+1} + \phi_{kk} \rho_{j-p} \quad j = 1, 2, \dots, k \quad (2.6)$$

A continuación se muestra dos ejemplos, el primer ejemplo consiste en la caracterización de un AR (1) y el segundo ejemplo consiste en identificar el orden de un proceso autorregresivo.

Ejemplo1: Caracterización de un AR (1)

Sea un proceso $\{Z_t\}$ AR (1):

$$\begin{aligned} Z_t &= \delta + \phi Z_{t-1} + \varepsilon_t \\ (1 - \phi B)Z_t &= \delta + \varepsilon_t \end{aligned}$$

- Un proceso estacionario se puede escribir como un proceso MA(∞)

$$Z_t = \frac{\delta + \varepsilon_t}{(1 - \phi B)} = (1 + \phi B + \phi^2 B^2 + \dots)(\delta + \varepsilon_t) = \frac{\delta}{1 - \phi} + \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$$

- El momento de primer orden es:

$$\mu = E(Z_t) = E\left(\frac{\delta}{1 - \phi} + \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}\right) = \frac{\delta}{1 - \phi}$$

- La varianza del proceso es:

$$\gamma_0 = E(Z_t - \mu)^2 = E\left(\sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}\right)^2 = \sum_{j=0}^{\infty} \phi^{2j} \sigma_{\varepsilon}^2 = \frac{\sigma_{\varepsilon}^2}{1-\phi^2}$$

- La autocovarianza del proceso es:

$$\gamma_1 = E(\tilde{Z}_t \tilde{Z}_{t-1}) = E((\phi \tilde{Z}_{t-1} + \varepsilon_t) \tilde{Z}_{t-1}) = \phi \gamma_0$$

$$\gamma_2 = E(\tilde{Z}_t \tilde{Z}_{t-2}) = E((\phi \tilde{Z}_{t-1} + \varepsilon_t) \tilde{Z}_{t-2}) = \phi \gamma_1 = \phi^2 \gamma_0$$

...

$$\gamma_k = E(\tilde{Z}_t \tilde{Z}_{t-k}) = E((\phi \tilde{Z}_{t-1} + \varepsilon_t) \tilde{Z}_{t-k}) = \phi \gamma_k = \phi^k \gamma_0$$

- La función de autocorrelación simple es: $\rho_k = \frac{\gamma_k}{\gamma_0} = \phi^k$ y tiene un decrecimiento exponencial.
- FAP, sólo tiene un coeficiente diferente de cero. Se puede demostrar con la ecuación (2.6).

Ejemplo2: Identificación de un proceso autorregresivo

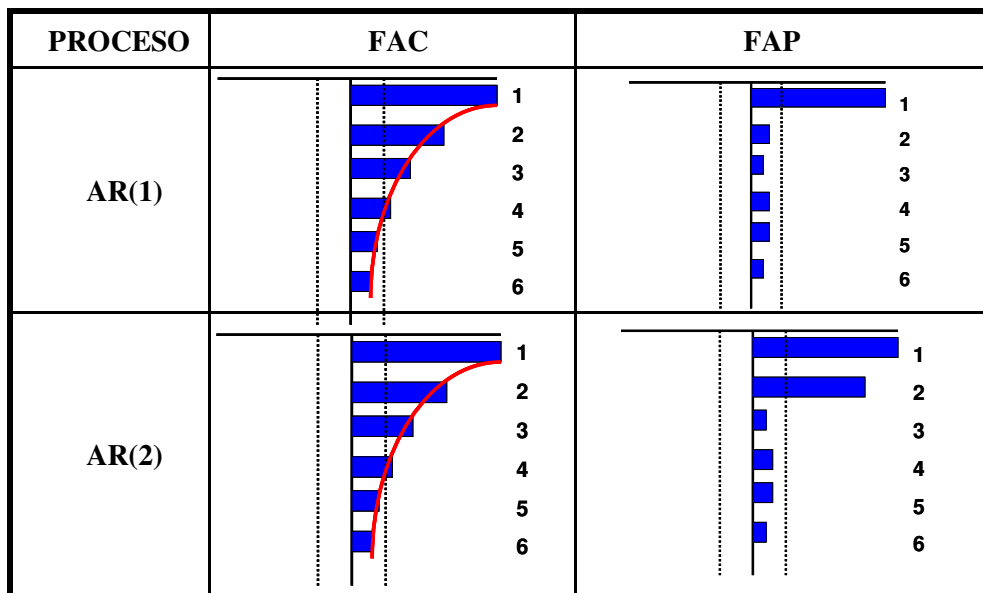


Figura 2.2: Identificación del orden del proceso autorregresivo.

Fuente: Elaboración propia.

2.3 MODELO LINEAL Y NO LINEAL EN EL TIEMPO

2.3.1 Proceso de ruido blanco

El proceso de ruido blanco se considera como una serie de choques que conducen el sistema. Se compone de una serie de variables aleatorias no correlacionadas con media cero y varianza constante (George, Gwilym, & Reinsel, 2008). Además si la distribución es normal se denomina Ruido Blanco Gaussiano.

$$E(\varepsilon_t) = 0 \quad \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2$$

Desde el momento que se supone como variables aleatorias no correlacionadas se deduce que su función de autocovarianza es:

$$\gamma_k = E(\varepsilon_t \varepsilon_{t-k}) = \begin{cases} \sigma_\varepsilon^2 & ; k = 0 \\ 0 & ; k \neq 0 \end{cases}$$

Por lo tanto, la función de autocorrelación del ruido blanco tiene una forma particularmente simple.

$$\rho_k = \begin{cases} 1 & ; k = 0 \\ 0 & ; k \neq 0 \end{cases}$$

Un proceso de ruido blanco no es necesariamente estacionario en sentido estricto ni debe estar formado por variables independientes, ya que sólo exige la incorrelación. Si suponemos la condición adicional de que las variables del proceso sean independientes, y no solo incorrelacionadas, llamaremos al proceso ruido blanco estricto (Peña, 2005).

2.3.2 Proceso lineal general

Los modelos estocásticos empleados se basan en la idea de Yule que consiste en que una serie de tiempo Z_t y sus valores sucesivos son altamente dependientes y con frecuencia puede ser considerado como el generado a partir de una serie de choques independientes ε_t , denominado proceso de ruido blanco.

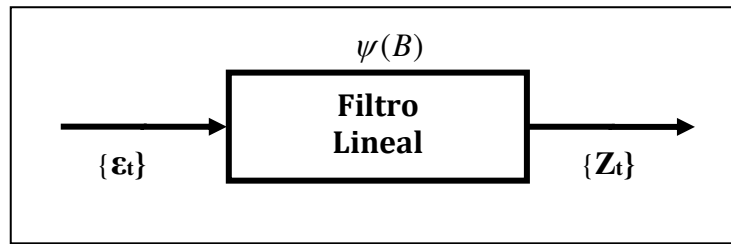


Figura 2.3: Filtro lineal

Fuente: Elaboración propia

Se observa que un proceso estocástico se genera como producto de la aplicación de un filtro lineal al proceso ruido blanco, es decir, nos permite representar Z_t como una suma ponderada de los valores actuales y pasados del proceso de ruido blanco ε_t .

$$Z_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \psi_3 \varepsilon_{t-3} \dots$$

$$Z_t = \mu + \psi(B) \varepsilon_t$$

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$$

$$Z_t = \mu + \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k} \quad (2.7)$$

En general μ es el parámetro que determina el nivel del proceso y $\Psi(B)$ es un operador lineal que transforma ε_t en Z_t y es llamado función filtro de transferencia.

El proceso definido en (2.7) es estacionario si se cumplen las siguientes condiciones:

$$\sum_{k=0}^{\infty} \psi_k < \infty \quad \sum_{k=0}^{\infty} \psi_k^2 < \infty \quad \sum_{k=0}^{\infty} \psi_k \psi_{k+l} < \infty$$

2.3.3 Modelo lineal en el tiempo

Una manera habitual de modelar una serie temporal consiste en establecer las relaciones existentes entre la serie temporal Z_t y la secuencia de perturbaciones asociadas ε_t . En el estudio de las series temporales se considera que las perturbaciones están incorrelacionadas, pero no es necesario asumir que son independientes y están idénticamente distribuidas (IID).

Es bien conocido que los modelos lineales solo pueden generar cuatro tipos de movimientos: oscilatorio y estable, oscilatorio y explosivo, no oscilatorio y estable, y no oscilatorio y explosivo. Sin embargo los modelos no lineales pueden generar un conjunto mucho más amplio de movimientos.

Definición de modelo lineal en el tiempo

Los modelos lineales en el tiempo generalmente son considerados modelos estacionarios de segundo orden, que se pueden expresar por la descomposición de (Wold, 1939) como una combinación lineal con coeficientes constantes ψ_k de las perturbaciones, donde éstas, ε_t son variables aleatorias no correlacionadas.

$$Z_t = \mu + \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k} \quad (2.8)$$

Donde μ es la media de Z_t , $\varepsilon_0 = 1$ y ε_t es un proceso ruido blanco.

Esta representación nos asegura que una serie temporal lineal puede expresarse como un MA (∞) de las perturbaciones, y es la responsable de la no condicionalidad de la varianza y de las autocovarianzas de la serie Z_t .

2.3.4 Modelo no lineal en el tiempo

En muchas situaciones una serie temporal puede exhibir comportamientos incompatibles con lo formulado de un proceso lineal. Tales comportamientos incluyen, por ejemplo, varianza condicional, irreversibilidad temporal, volatilidades, etc. Aunque los modelos lineales estudiados proporcionan aproximaciones útiles a la verdadera relación, podremos mejorar la predicción si utilizamos un modelo más apropiado.

Definición de modelo no lineal en el tiempo

Cualquier proceso estocástico que no satisfaga la condición de la ecuación (2.8) se dice que es un modelo no lineal. Para un mejor entendimiento de los modelos no lineales escribiremos la serie Z_t en términos de sus momentos condicionales.

Sea Z_t una serie y Ω_{t-1} el conjunto generado por la información disponible hasta el instante (t-1). El conjunto Ω_{t-1} denota la colección de combinaciones lineales de elementos en $\{Z_{t-1}, Z_{t-2}, Z_{t-3}, \dots\}$ y $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3}, \dots\}$.

La media y la varianza condicional de Z_t dado Ω_{t-1} son:

$$\mu_{Z_t} = E(Z_t | \Omega_{t-1}) = g(\Omega_{t-1})$$

$$\sigma_{Z_t}^2 = \text{Var}(Z_t | \Omega_{t-1}) = h(\Omega_{t-1})$$

Donde $g(\cdot)$ y $h(\cdot)$ son funciones bien definidas con $h(\cdot) > 0$. Si la serie Z_t es lineal, $g(\cdot)$ es una función lineal de elementos de Ω_{t-1} y $h(\cdot) = \sigma_\varepsilon^2$ (varianza constante).

Si $g(\cdot)$ es una función no lineal, se dice que Z_t es no lineal en media. Si $h(\cdot)$ varía en el tiempo, entonces se dice que Z_t es no lineal en la varianza.

Tipos de modelos no lineales

(Tong, 1990) realiza una exhaustiva clasificación de los modelos no lineales para series temporales. En ella, estos modelos se estructuran en 13 grupos diferentes:

Cuadro 2.1: Modelos no lineales

MODELOS NO LINEALES (Tong 1990):
Modelos Umbral
Modelos EXPAR
Modelos FAR
Modelos PAR
Modelos RCA
Modelos NEAR
Modelos autoregresivos con espacio de estados discreto
Modelos bilineales
Modelos no lineales de medias móviles
Modelos ARCH
Modelos de segunda generación
Modelos doblemente estocásticos
Modelos SDM

Fuente: (Tong, 1990).

Modelos ARCH

Los modelos ARCH fueron propuestos por (Engle, 1982). Estos modelos suponen que la varianza marginal o incondicional es constante y que la covarianza entre las variables en periodos de tiempo diferentes sólo depende del tiempo que transcurre entre ellas, hecho que es necesario para que el proceso sea débilmente estacionario o estacionario en covarianzas, mientras que la varianza condicional no es constante.

La familia de modelos ARCH, han sido diseñadas para modelar la varianza condicional, y resultan apropiados cuando los residuos de las estimaciones presentan heterocedasticidad y son útiles para capturar los agrupamientos de volatilidad observada en las series financieras: inflación, rendimientos de índices de bolsa, etc.

(Bollerslev, 1986) introdujo el modelo ARCH generalizado (GARCH), cuya forma más sencilla, la de un GARCH (1,1), es equivalente a un proceso ARCH de un orden infinito, ARCH (∞), pero más fácil de estimar dado que depende solo de tres parámetros.

Modelos Bilineales

Los modelos bilineales fueron introducidos por (Granger & Anderson, 1978). Son una generalización de los modelos ARMA a los que se le añade una parte de interacción entre los instantes pasados y las innovaciones.

Son especialmente útiles para series temporales en las que ocasionalmente se produzcan grandes alejamientos de la media.

Modelos FAR

Los modelos Modelos Autorregresivos con Coeficiente Funcional (FAR) fueron introducidos por primera vez por (Chen & Tsay, 1993). Tiene la forma de un proceso autorregresivo sin constante en el que los parámetros del modelo son funciones evaluadas en la variable Z_{t-d} .

Modelos de segunda generación

Los modelos mostrados hasta ahora intentan explicar un tipo de fenómeno no lineal, por ello Tong los denomina modelos no lineales de primera generación. Para aprovechar mejor sus ventajas, Tong propone crear modelos más complejos que unen a varios de estos modelos, creado así modelos de segunda generación, tercera generación, y así.

Un posible modelo de segunda generación que empieza a ser empleado es el modelo por umbrales con heterocedasticidad condicional denominado modelo SETAR-ARCH.

2.4 HIPÓTESIS

El modelo SETAR bajo la metodología de Tsay es más apropiado que los modelos de la familia ARCH, cuando se analizan series de tiempo no lineales.

2.5 TERMINOLOGÍA

Régimen: El régimen o estado es la transición de un ciclo a otro, debido a la ocurrencia de ciertos factores que cambian la estructura del proceso estocástico.

Umbral: Es un punto de referencia respecto de los cuales se puede comparar una medición. La Real Academia Española lo define como el “Valor mínimo de una magnitud a partir del cual se produce un efecto determinado”. Para nuestro caso es el valor límite en el cual se realiza un cambio de estado o régimen cuya notación es r .

Variable umbral: Es aquella variable que define el comportamiento de cada estado o régimen. Para nuestro caso será la variable Z_t rezagada d veces, es decir, Z_{t-d} .

Volatilidad: La volatilidad es una medida de la intensidad de los cambios aleatorios o impredecibles de la serie entorno a un valor medio denominado varianza.

Ciclo límite: Son soluciones de equilibrio que en vez de colapsar a un punto límite simple, convergen a un conjunto de puntos, creando una órbita periódica.

CAPÍTULO III

METODOLOGÍA

CAPÍTULO III. METODOLOGÍA

3.1 DISEÑO DE LA INVESTIGACIÓN

El diseño de investigación a usar es no experimental (observacional) pues se observan los fenómenos tal y como se dan en su contexto natural. Es además, longitudinal, por tratarse de series de tiempo, retrospectivo y cuantitativo (Hernández Sampieri, Fernández Collado, & Baptista Lucio, 2010).

3.2 TIPO DE INVESTIGACIÓN

El tipo de investigación es aplicada, pues se propone transformar el conocimiento 'puro' en conocimiento útil. Tiene por finalidad la búsqueda y consolidación del saber y la aplicación de los conocimientos para el enriquecimiento del acervo cultural y científico (Universidad Nacional Abierta y a Distancia (UNAD), 2015).

3.3 POBLACIÓN DE ESTUDIO

La población de estudio es el conglomerado total de bienes y servicios finales producidos en el país en un periodo específico, es decir, comprende el valor de los bienes producidos, como viviendas, comercio, servicios, Gobierno, transporte, etc. Cada uno de estos bienes y servicios se valora a su precio de mercado y los valores se suman para obtener el PBI.

3.4 UNIDAD DE ANÁLISIS

Un conglomerado de bienes y servicios finales producidos en el país en un periodo específico.

3.5 TAMAÑO DE MUESTRA

Se consideró el PBI trimestral del periodo 1980-2014, el cual consistía de 140 observaciones, cabe resaltar que el análisis estadístico se realizó considerando las variaciones porcentuales del PBI, es decir, se consideró 136 observaciones.

3.6 SELECCIÓN DE LA MUESTRA

En series de tiempo la muestra corresponde a la realización de un proceso estocástico, por lo que no se puede hablar de aleatoriedad en la selección, lo cual es un requisito fundamental para la obtención de una muestra.

3.7 TÉCNICAS DE RECOLECCIÓN DE DATOS

Las técnicas de recolección de datos son las distintas formas o maneras de obtener la información. Son ejemplos de técnicas de información; la observación directa, la encuesta en sus dos modalidades: oral o escrita (cuestionario), la entrevista, el análisis documental, análisis de contenido, etc. (Arias, 2006). Los datos de la serie provienen de fuentes secundarias. A continuación se detallan la fuente de información y el procedimiento de recolección.

3.7.1 Fuentes de información

Los datos de la serie en estudio provienen del registro de estadísticas trimestrales del Banco Central de Reserva (Gerencia Central de Estudios Económicos). Para dicha estimación también se usó información del Instituto Nacional de Estadística e Informática – INEI.

3.7.2 Procedimiento de recolección de datos

La recolección de los datos para el cálculo del PBI proviene de diferentes fuentes. A continuación mencionaremos la metodología del cálculo del PBI y las principales variables económicas que intervienen en el cálculo.

El cálculo del PBI se realiza desde el punto de vista del Gasto o destino de la producción. Se mide el valor de las diferentes utilizaciones finales de la producción en la Economía, deducido el valor de las importaciones de los bienes y servicios (producción no generada en el territorio económico).

$$\text{PBI} = \text{GCFP} + \text{GCFG} + \text{FBKF} + \text{VE} + \text{X} - \text{M}$$

Para el cálculo del PBI trimestral el procedimiento dependerá de la variable que intervenga en el cálculo, a continuación mencionaremos solo algunos de ellos:

Gasto de Consumo Fijo Privado (GCFP):

- Importaciones y exportaciones de la SUNAT.
- Indicadores de la Producción Nacional.
- Encuesta Nacional de Hogares (ENAHU).

Formación Bruta de Capital Fijo (FBCF):

- Índice de Precios al Consumidor – INEI.
- Índice de Precios de Maquinaria y Equipo - INEI.
- Demanda Interna de Cemento-ASOCEM.

3.8 PROCESAMIENTO Y ANÁLISIS DE LA INFORMACIÓN

En base a las estadísticas elaboradas por el Banco Central de Reserva del Perú (BCRP) y el Instituto Nacional de Estadística e Informática (INEI); fuentes del sector para el PBI trimestral, se eligió el período 1981-2014 por ser el más extenso y confiable. Si bien hay datos disponibles del 2015, esta información se utilizará para la comparación de los pronósticos.

Para modelar nuestra serie en estudio, se utilizó el programa S-PLUS 8.0 para su tratamiento, el cual consistió en probar la no linealidad de tipo umbral mediante la prueba de Tsay, obtención de los parámetros estructurales, estimación del modelo, validación y los pronósticos respectivos.

Para la realización del análisis exploratorio se usó el programa R-PROJECT 3.3.4, es decir, se obtuvo la función de autocorrelación de la serie, el espectro, diagramas de dispersión directa, estimación Kernel de la linealidad, etc.

3.9 MÉTODOS ESTADÍSTICOS

Los métodos estadísticos que se desarrollarán están enfocados a los modelos no lineales en el tiempo, los cuales son el marco de desarrollo para los modelos GARCH y los modelos SETAR. A continuación se desarrollan las pruebas que permiten determinar si un modelo es lineal o no lineal, los modelos autorregresivos por umbral dentro de los cuales se encuentra el modelo SETAR y finalmente se desarrollan los modelos de heterocedasticidad condicional, específicamente los modelos GARCH.

3.9.1 Análisis de la linealidad en el tiempo

Frente a la modelización de una serie de tiempo, una primera interrogante que se plantea es si el comportamiento de la serie es lineal o no lineal. Aunque la observación a través de gráficos de la serie original ofrece una primera aproximación sobre el tipo de dinámica del proceso, son las pruebas estadísticas de contraste las que nos permiten asegurar si la serie tiene un comportamiento lineal o no.

Existen diferentes pruebas diseñados para probar la linealidad de una serie de tiempo, pero en esta investigación nos vamos a centrar en aquellos que presentan como alternativa a la hipótesis de linealidad, la no linealidad por umbrales o tipo “threshold”.

- **Prueba TAR-F de Tsay**

La prueba TAR-F se basa en la prueba de linealidad de (Keenan, 1985). Al igual que la prueba de Keenan, se obtiene a partir de una autorregresión ordenada y utiliza los residuos predictivos, pero es más potente, más simple y ampliamente aplicable. Una de las mayores ventajas del TAR-F es su especificidad para capturar no linealidad tipo umbral y además permite determinar el valor del parámetro de retardo d .

Se desea probar lo siguiente.

$$\left\{ \begin{array}{l} \mathbf{H}_0: \text{La serie } Z_t \text{ presenta linealidad.} \\ \mathbf{H}_1: \text{La serie } Z_t \text{ presenta no linealidad de tipo umbral.} \end{array} \right.$$

En esencia (Tsay, 1989) propone la siguiente prueba:

Supongamos que la variable Z_t sigue un proceso lineal estacionario de orden p ; entonces para valores de n grandes y para un valor de d fijado, el estadístico definido como:

$$F(p, d) = \frac{\left(\sum_{t=b+1}^{n-d-h+1} \hat{\varepsilon}_t^2 - \sum_{t=b+1}^{n-d-h+1} \varepsilon_t^2 \right) / p + 1}{\sum_{t=b+1}^{n-d-h+1} \varepsilon_t^2 / (n - d - b - p - h)}$$

$$\text{Con } b = \frac{n}{10} + p \text{ y } h = \max\{1, p + 1 - b\}$$

Donde los $\hat{\varepsilon}_t^2$ son los residuales de cada autorregresión tipificada y los ε_t^2 son aquellos que se obtienen de las regresiones de los residuales de cada autorregresión tipificada sobre los regresores $Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$.

$\hat{F}(p, d)$ sigue aproximadamente una distribución F de Fisher con $(p+1)$ y $(n-d-b-p-h)$ grados de libertad.

La aplicación de esta prueba tiene como inconveniente que tan solo acepta que la serie de observaciones siga un modelo no lineal del tipo TAR; de manera que si la serie resulta ser no lineal pero no del tipo TAR, es necesario utilizar otro método para averiguar el modelo a identificar.

- **Prueba de Hansen**

Esta subsección introduce otra prueba de la no linealidad de umbral y otro procedimiento para la estimación de los modelos SETAR según lo propuesto por (Hansen, 1997). La ventaja de este procedimiento es que los umbrales pueden ser estimados junto con otros parámetros del modelo y los intervalos de confianza

válidos se pueden construir para los umbrales estimados. La desventaja es que la implementación actual sólo es compatible con el modelo SETAR de dos regímenes y por lo tanto sólo un valor umbral puede ser estimado.

Para nuestro caso solo usaremos la prueba de Hansen para verificar la no linealidad de nuestra serie en estudio.

En esencia Hansen propone la siguiente prueba:

Supongamos que la variable Z_t sigue un proceso lineal estacionario de orden p y para un valor de d fijado, se desea verificar si se cumplen las siguientes hipótesis.

$$\left\{ \begin{array}{l} \mathbf{H}_0: \text{La serie } Z_t \text{ presenta linealidad.} \\ \mathbf{H}_1: \text{La serie } Z_t \text{ presenta no linealidad de tipo umbral.} \end{array} \right.$$

El estadístico de decisión se calcula mediante:

$$F_n^B = \sup_{r^* \in I_r} (F_n(r^*))$$

donde:

$$F_n(r^*) = n \frac{\hat{\sigma}^2(r^*) - \hat{\sigma}_0^2}{\hat{\sigma}^2(r^*)}$$

Siendo $\hat{\sigma}^2(r^*)$ la varianza residual de la regresión de los números aleatorios \mathbf{u}_t sobre el proceso Z_t , donde \mathbf{u} es un vector de números aleatorios con distribución $N(0, I_n)$.

- **Pruebas gráficas**

Diagrama de dispersión directa: Se trata de gráficos dispersión del proceso $\{Z_t\}$ respecto a su retardo $\{Z_{t-k}\}$, en el cual los puntos consecutivos se van uniendo mediante líneas. En la figura 3.1 vemos un ejemplo del diagrama para datos generados mediante un modelo AR lineal y un modelo SETAR no lineal.

El diagrama perteneciente al proceso lineal tiene un comportamiento completamente aleatorio, mientras que el proceso no lineal dibuja una especie de elipse, sin datos en el centro. (Tong, 1990) explica cómo este tipo de resultados muestran la existencia de ciclos.

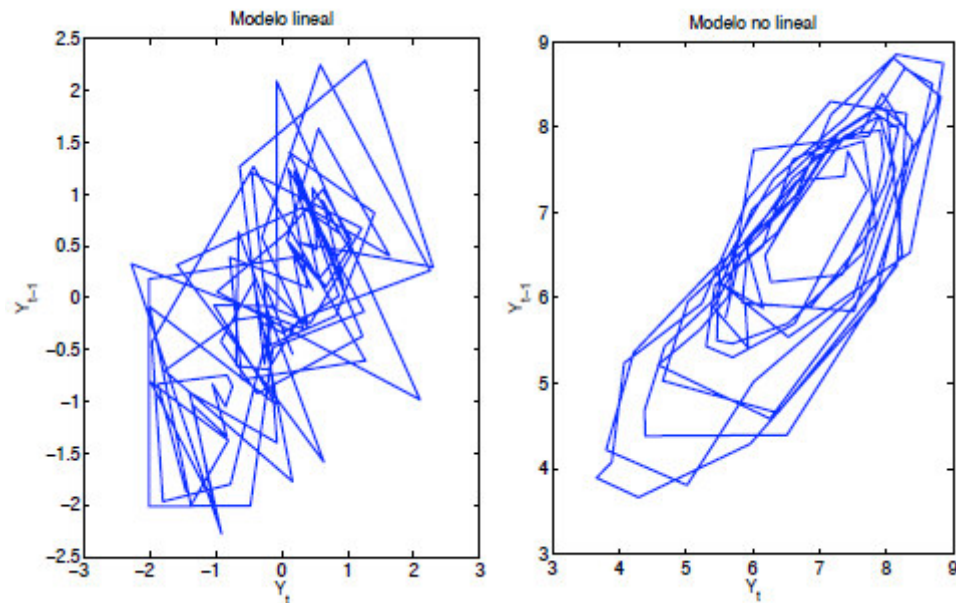


Figura 3.1: Diagramas de dispersión directa.

Fuente: Elaboración propia.

Suavizado no paramétrico: El estudio del suavizado no paramétrico realizado sobre la relación existente proceso $\{Z_t\}$ respecto a su retardo $\{Z_{t-k}\}$ puede ser una herramienta útil para identificar una posible relación no lineal. En la figura 3.2 vemos la diferencia entre realizar un ajuste lineal y un suavizado no paramétrico, y cómo este último ayuda a explicar mejor los datos, pudiendo indicar por lo tanto no linealidad.

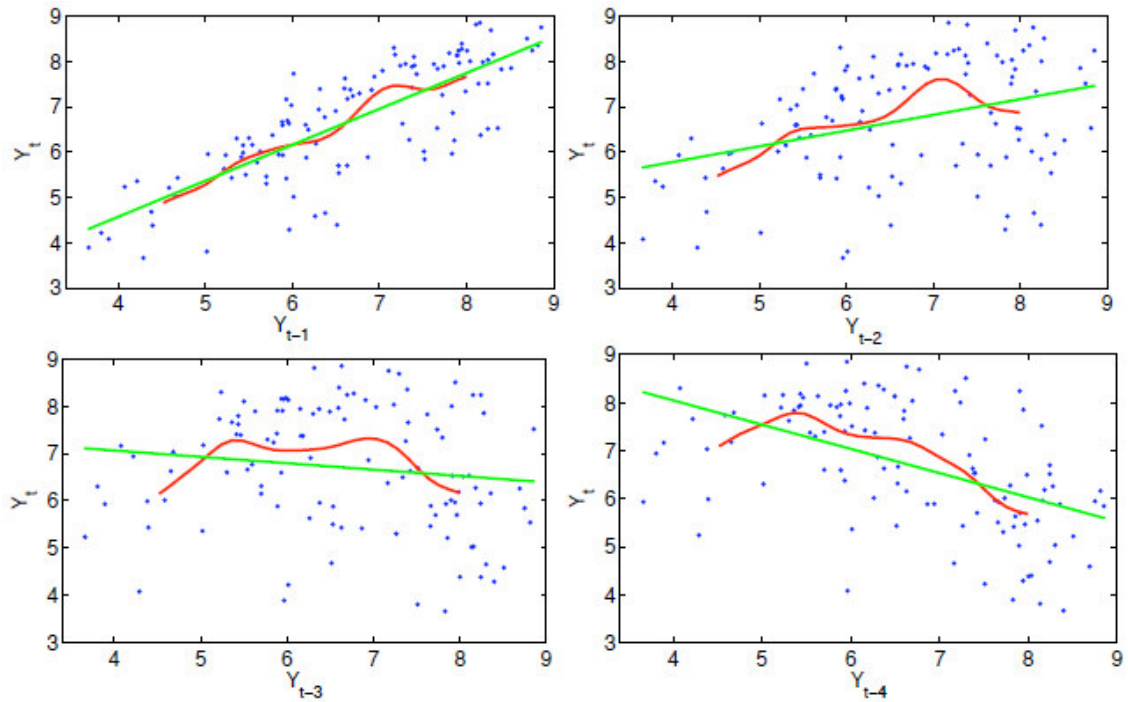


Figura 3.2: Suavizado no paramétrico.

Fuente: Elaboración propia.

3.9.2 Modelos por umbrales

Los modelos por umbrales, mencionados por vez primera por Tong (1977), asumen la existencia de diferentes funciones lineales en diferentes espacios de estados. La división de los espacios de estados viene dada por la variable umbral Z_{t-d} , con $d \geq 1$. Basándose en esta idea Tong formuló el principio del umbral, el cual dice que “El análisis de un sistema estocástico complejo puede ser realizado mediante su descomposición en subsistemas más sencillos” (Tong, 1990).

Los modelos por umbrales son también denominados, modelos de cambio de régimen, pues consideran la existencia de estados o regímenes. El comportamiento dinámico de la serie está determinado por el régimen que ocurre en un momento dado del tiempo, y el régimen a su vez, por una variable que puede ser observable o por una variable latente.

Además difieren en sus supuestos sobre el proceso estocástico que genera el régimen y de forma genérica, pueden clasificarse en Modelos de Transición Determinística, si el régimen fuera observable, y Modelos de Transición Estocástica, si el régimen fuera no observable.

Las principales características de esta clase de modelos son ciclos límite, frecuencias dependientes de amplitud y fenómenos de salto, que no puede ser capturada por un modelo lineal de series de tiempo. Gran parte de la motivación original es que el modelo es capaz de producir ciclos límites asimétricos (Tsay, 1989).

Los modelos autorregresivos por umbral, sin embargo, no ha recibido mucha atención en la aplicación. Esto es debido a la falta de un adecuado procedimiento de modelado y la incapacidad de identificar la variable de umbral y estimar los valores del umbral.

3.9.3 Modelos de cambio de régimen de transición determinística

Un enfoque natural para el modelado de series temporales (económicos, hidrológicos, financieros, etc.) con los modelos no lineales de cambio de régimen es considerar diferentes estados o regímenes que permitan la posibilidad de que el comportamiento dinámico de la variable en estudio dependa del régimen de que se produce en cualquier momento dado del tiempo. Por comportamiento dinámico dependiente del estado se entiende que ciertas propiedades de la serie de tiempo, tales como su media, la varianza y autocorrelación, son diferentes en cada régimen.

En los últimos años se han propuesto varios modelos de series de tiempo que formalizan la idea de la existencia de diferentes regímenes generados por un proceso estocástico (principalmente en la economía y las finanzas).

A continuación, se discuten las representaciones de los diferentes modelos de cambio de régimen y los métodos de estimación y análisis de la presencia de los efectos de cambio de estado.

- **Modelos autorregresivos por umbrales (TAR)**

Este modelo está motivado por diversas características que comúnmente son observadas en la práctica, tales como la asimetría en la disminución y el aumento de la estructura de un proceso. Este modelo usa modelos lineales a trozos para obtener una mejor aproximación de la media condicional. Sin embargo, en comparación con el modelo lineal tradicional a trozos que permite los cambios en el modelo se produzcan en el espacio “tiempo”, el modelo autorregresivo con umbral, denotado por TAR (Threshold Autorregressive), usa el espacio umbral para mejorar la aproximación lineal. Como ya se ha mencionado, en los modelos TAR se asume que el régimen queda determinado por una variable observable denominada variable umbral (Monsalve & Harmath, 2015).

Los modelos TAR son simples y fácil de entender, pero lo suficientemente útil como para generar complejos dinámicos no lineales. Por ejemplo, se puede demostrar que los modelos TAR pueden tener ciclos límite y por lo tanto ser usados para modelar series de tiempo periódicos, o producir asimetría y fenómenos de salto que no pueden ser capturados por un modelo de serie de tiempo lineal (Tong & Lim, 1980).

- **Modelos de auto-activación de umbral (SETAR)**

Los modelos auto-excitables denotados por SETAR (Self-Exciting Threshold Autoregressive) son un caso particular de los modelos TAR en el cual la variable observable que determina los distintos estados o regímenes, Z_{t-d} , es un retardo de la propia serie.

Los modelos SETAR dan lugar a patrones de comportamientos muy diversos, dependiendo de los parámetros seleccionados.

- **Modelos de transición suave (STAR)**

(Teräsvirta, Granger, & Tjøstheim, 2010), mencionan que el Modelo de Cambio de Régimen de Transición Suave denotados por STAR (Smooth Transition Autoregressive Model) tiene como supuesto que los cambios de regímenes son generados por una variable de transición observable Z_t , ya sea endógena rezagada, exógena o ambas. Además de ello se define la función de transferencia exponencial que determina el peso del régimen y que está, usualmente, limitada entre 0 y 1.

3.10 MODELO SETAR

El modelo SETAR es un modelo lineal por tramos que, globalmente, presenta un comportamiento no lineal. Un modelo SETAR puede representar ciertas características que uno lineal no puede: ciclos límite, amplitud dependiente de la frecuencia, ciclos asimétricos, cambios de nivel, etc. (Zivot & Wang, 2006).

3.10.1 Definición del modelo SETAR

Una serie de tiempo Z_t es un proceso SETAR (Self-Exciting Threshold Autoregressive) si tiene la siguiente estructura:

$$Z_t = \Phi_0^{(j)} + \sum_{i=1}^p \Phi_i^{(j)} Z_{t-i} + \varepsilon_t^{(j)} \quad ; \text{ si } r_{j-i} \leq Z_{t-d} \leq r_j \quad (3.1)$$

donde $j=1,2,\dots,k$ y d es un entero positivo.

Los valores umbrales son $-\infty < r_0 < r_1 < \dots < r_k < \infty$. La partición $r_{j-i} \leq Z_{t-d} \leq r_j$ se corresponde con el j -ésimo régimen del modelo y los $\{\varepsilon_t^{(j)}\}$ son sucesiones de variables aleatorias normales e independientes e idénticamente distribuidas con media cero y varianza σ^2 .

Para resolver el problema de detectar los cambios estructurales que supone un modelo como (3.1), en cada umbral se adoptara la idea de autorregresión tipificada de Tsay.

Pero si para un modelo como el descrito anteriormente se plantea la autorregresión clásica:

$$Z_t = \Phi_0 + \Phi_1 Z_{t-1} + \Phi_2 Z_{t-2} + \dots + \Phi_p Z_{t-p} + \varepsilon_t \quad (3.2)$$

Entonces, ésta no será de utilidad, ya que los Φ_i son distintos en cada régimen definido. Por otra parte, si las observaciones se ordenan de menor a mayor, de acuerdo con el valor de su variable umbral, puede considerarse la autorregresión

$$Z_{(t-d)+d} = \Phi_0 + \Phi_1 Z_{(t-d)+d-1} + \Phi_2 Z_{(t-d)+d-2} + \dots + \Phi_p Z_{(t-d)+d-p} + \varepsilon_{(t-d)+d} \quad (3.3)$$

3.10.2 Estimación del modelo SETAR

A continuación adoptaremos la estrategia de modelación diseñada por (Tsay, 1989), para especificar un modelo SETAR. Los siguientes pasos para la identificación y estimación de los modelos SETAR, se consideró del artículo “Análisis de cambio de régimen en series de tiempo no lineales utilizando modelos TAR aplicado al PBI trimestral de España” desarrollado por (Pérez & Velazqués, 2004).

Paso 1:

Especificar un modelo AR lineal de orden p apropiado para la serie de tiempo bajo estudio y determinar el conjunto de posibles umbrales rezagados S . El orden p puede seleccionarse considerando el ACF y ACP de Z_t . Un modelo AR de orden alto puede tener razonables aproximaciones a un modelo no lineal; se espera que la selección de p en este paso pueda reflejar tal aproximación.

Tsay recomienda que la especificación de p se realice mediante las funciones de autocorrelación simple y parcial, con preferencia frente a criterios de penalización como el Criterio de Información de Akaike (AIC) o el Criterio de Información Bayesiano (BIC).

El conjunto S de posibles umbrales rezagados lo denotamos por $d = 1, 2, \dots, p$, para el p seleccionado. Es decir, se generan las siguientes variables umbrales.

$$Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$$

Paso 2:

Seleccionar la variable umbral Z_{t-d} para $d = 1, 2, \dots, p$, se tienen $Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$ variables umbrales. Para cada variable umbral y p seleccionados en el paso 1, se realiza la estimación de mínimos cuadrados ordinarios (MCO) en forma recursiva para la autorregresión tipificada considerada en (3.3), teniendo en cuenta que las observaciones se ordenan de menor a mayor.

Tsay propone iniciar el proceso recursivo de estimación por MCO con $b = \frac{n}{10} + p$ observaciones, donde n es el número de observaciones.

Paso 3:

Prueba de no linealidad y estimación de la autorregresión tipificada para valores fijos de p y d , el número efectivo de observaciones en la autorregresión tipificada es $n-d-h+1$, donde $h = \text{Max} \{1, p+1-d\}$. Como se asume que la recursión para estimar las autorregresiones tipificadas se inicia con b observaciones, entonces se tienen $n-d-b-h+1$ residuales disponibles. Utilizaremos los residuales de (3.3) para realizar el contraste de no linealidad, basándonos en el estadístico con distribución asintótica F .

Para la prueba de no linealidad, se estima la regresión de los residuales predichos, que se obtienen de la autorregresión tipificada en (3.3) sobre las variables regresoras $\{Z_{(t-d)+d-v} \mid v = 1, 2, \dots, p\}$. Dicha regresión de mínimos cuadrados ordinarios es:

$$\hat{e}_{(t-d)+d} = w_0 + \sum_{v=1}^p w_v Z_{(t-d)+d-v} + \varepsilon_{(t-d)+d}$$

Para los residuales predichos $b+1, \dots, n-d-h+1$, luego, el estadístico F se define como:

$$F(p, d) = \frac{\left(\sum_{t=b+1}^{n-d-h+1} \hat{\varepsilon}_t^2 - \sum_{t=b+1}^{n-d-h+1} \varepsilon_t^2 \right) / p + 1}{\sum_{t=b+1}^{n-d-h+1} \varepsilon_t^2 / (n-d-b-p-h)} \quad (3.4)$$

Donde los $\hat{\varepsilon}_t^2$ son los residuales de cada autorregresión tipificada y los ε_t^2 son aquellos que se obtienen de las regresiones de los residuales de cada autorregresión tipificada sobre los regresores $Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$. Para elegir la variable umbral Z_{t-d} , se considera el F más significativo, que es el F tal que:

$$F(p, d) = \text{Max}_{v \in S} \{F(p, v)\}, \text{ donde } F(p, v), \text{ es el estadístico definido en (3.4)}$$

La prueba $F(p, d)$ además de comprobar la no linealidad umbral, sirve como indicador de la intensidad de los efectos no lineales y permite, en consecuencia, determinar el valor más apropiado para el retardo del umbral.

Paso 4:

Localizar los valores umbrales. Para los valores de p y d seleccionados en los pasos anteriores, se localizan los valores umbrales, usando diagramas de dispersión de los estadísticos T , de estimaciones recursivas de un coeficiente AR contra la variable umbral $Z_{(t-d)}$. Los mayores cambios en la pendiente de estos valores T sugieren cambios de régimen.

Los estadísticos T tienen dos funciones: a) muestran la significancia de un coeficiente AR en particular y b) cuando el coeficiente es significativo, convergen gradualmente y suavemente a un valor fijo de la recursión continua. Para ilustrar esto, consideremos el modelo SETAR:

$$Z_t = \begin{cases} \Phi_1^{(1)} Z_{t-1} + \varepsilon_t^{(1)} & \text{si } X_{t-d} \leq r_1 \\ \Phi_1^{(2)} Z_{t-1} + \varepsilon_t^{(2)} & \text{si } X_{t-d} > r_1 \end{cases}$$

Sea $\hat{\Phi}_1$ la estimación recursiva de un coeficiente AR (1) en la autorregresión tipificada, los estadísticos T de $\hat{\Phi}_1$ se comportan exactamente como los de un modelo lineal antes de que la recursión investigue el valor umbral r_1 .

Una vez se encuentra r_1 , el estimador $\hat{\Phi}_1$ cambia y la razón T empieza a desviarse. En ese momento, el patrón de convergencia gradual de la razón T se daña. De hecho, la razón T cambia de dirección en el valor umbral. Para el modelo SETAR dado, $\hat{\Phi}_1$ empieza a cambiar cuando $Z_{(t-d)}$ encuentra a r_1 .

Paso 5:

De ser necesario redefinir el orden AR en cada régimen usando técnicas de autorregresión lineal y luego proseguir con la estimación de los parámetros de cada régimen.

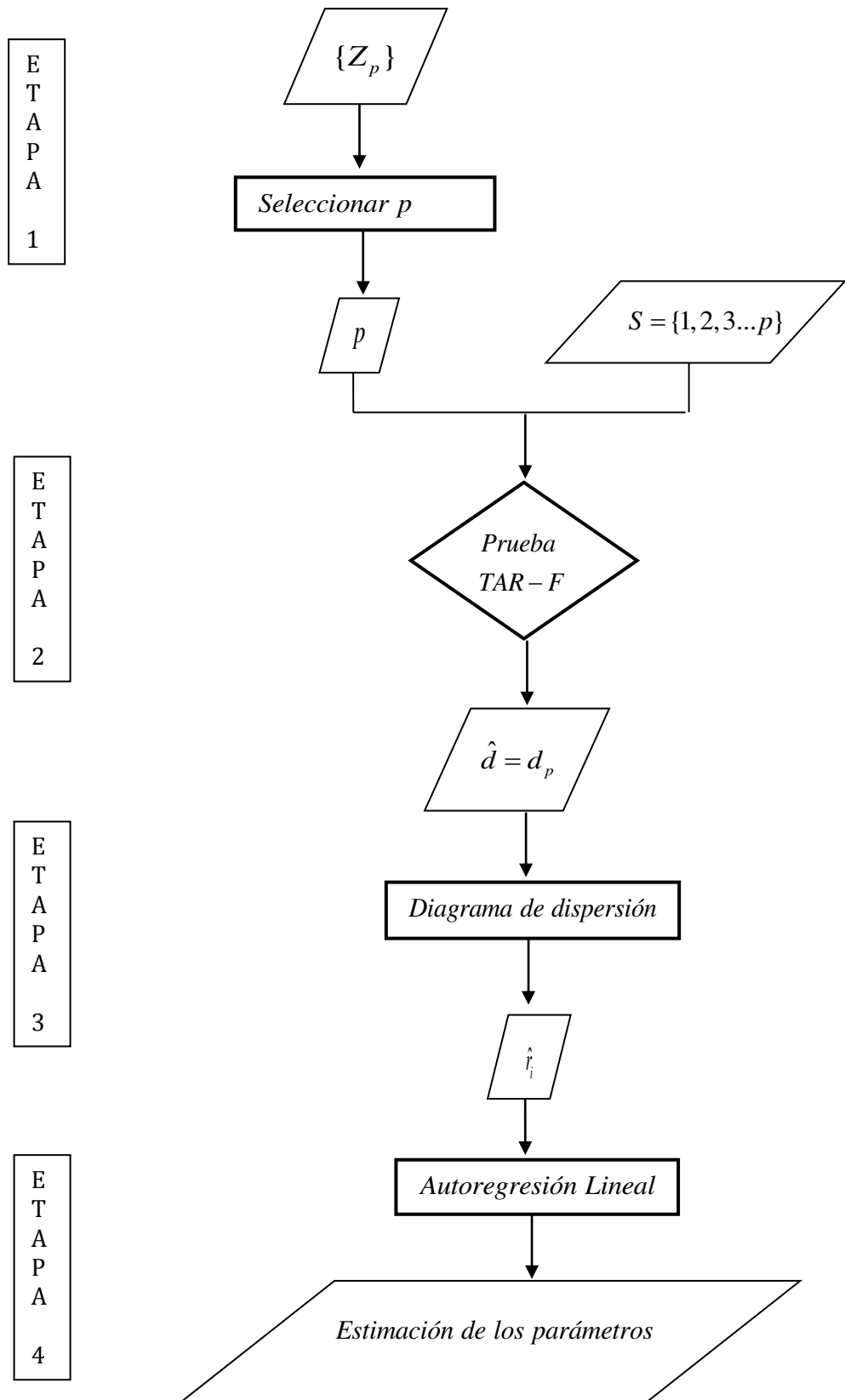


Figura 3.3: Esquematzación del algoritmo de Tsay (1989).

Fuente: Elaboración propia.

La metodología de Tsay nos permite localizar los valores umbrales para los valores de p y d seleccionados en los pasos anteriores, la localización de los umbrales fue considerando los diagramas de dispersión de los estadísticos T. Sin embargo no se conocía con exactitud el valor del umbral óptimo. A continuación se muestra un algoritmo para la identificación del valor umbral óptimo usando los criterios de información AIC (Criterio de información de Akaike) y BIC (Criterio de información Bayesiano).

El algoritmo consiste en elegir una vecindad de posibles valores umbrales y descartar los valores usando los criterios de información antes mencionados.

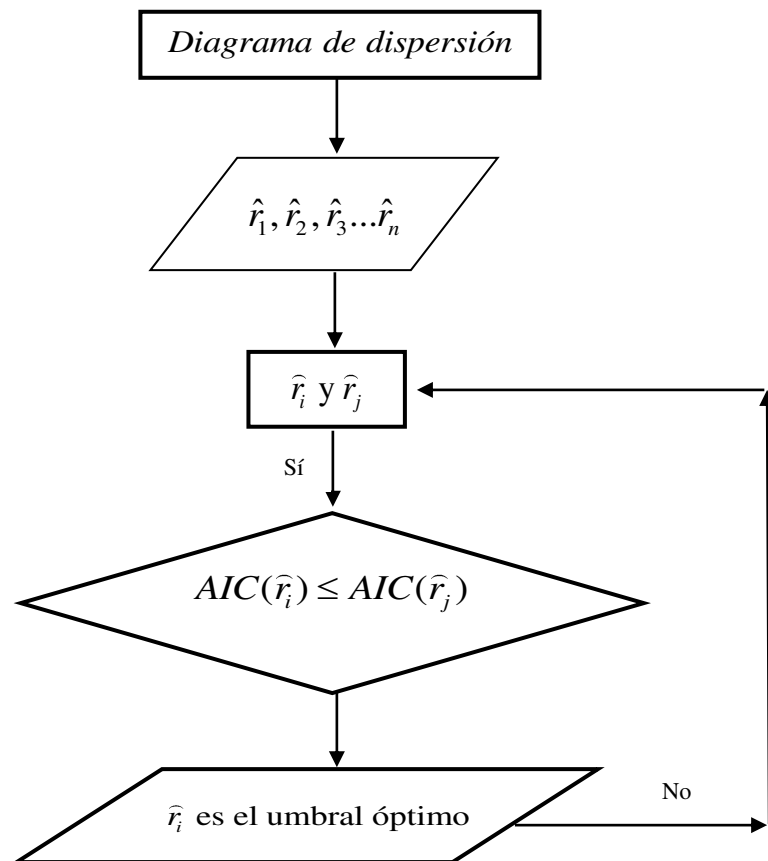


Figura 3.4: Algoritmo para la obtención del umbral óptimo.

Fuente: Elaboración propia.

3.10.3 Validación del modelo SETAR

En el artículo de (Tsay, 1989) denominado “Testing and Modeling Threshold Autoregressive Processes” propone una metodología para la estimación de los modelos SETAR, sin embargo no propone ningún método de validación, por este motivo se validara el modelo SETAR bajo la misma metodología de Box-Jenkins.

Una vez identificado el modelo es necesario validar el modelo, es decir, verificar el supuesto básico de los residuales del modelo, tales como la incorrelación, normalidad y aleatoriedad de los residuos. Por otra parte también se puede usar las pruebas gráficas como el correlograma y el espectro de los residuos.

Para verificar si los residuales se comportan como ruido blanco [media cero y varianza constante], es común analizar el comportamiento de los correlogramas, el gráfico de los residuales con respecto al tiempo y de los residuales cuadráticos. En el primero se examina si las correlaciones son estadísticamente significativas, o si forman algún patrón, lo cual es un indicio de incorrelación en los residuos; en el segundo se evalúa si los residuales se distribuyen uniformemente alrededor de cero; y en el tercero se analiza la variabilidad de los residuales, la presencia de altas volatilidades es reflejo de una varianza no constante.

3.10.4 Pronósticos con el modelo SETAR

Los modelos de series de tiempo no lineales pueden ser considerados para diversos fines. A veces el objetivo principal es la obtención de una descripción adecuada de los patrones dinámicos que están presentes en una variable en particular. Sin embargo, un objetivo adicional es emplear el modelo para la predicción de valores futuros de la serie de tiempo.

Por otra parte, fuera de la muestra de previsión también se puede considerar como una manera de evaluar los modelos de estimación de cambio de régimen. En particular, la comparación de las previsiones de los modelos no lineales con los de un modelo lineal podría permitir a uno determinar el valor añadido de las características no lineales del modelo.

En esta sección se discuten varias formas de obtener puntos y el intervalo previsiones de los modelos no lineales. Esto es seguido por algunas observaciones sobre la forma de evaluar los pronósticos de los modelos no lineales, y sobre la manera de comparar las previsiones a partir de modelos lineales y no lineales.

A continuación se desarrollara algunas definiciones sobre los métodos de pronósticos usados para los modelos SETAR, se va considerar las definiciones propuestas del libro: “Non-Linear Time Series Models in Empirical Finance” desarrollado por (Franses & Dijk, 2000).

Los métodos de pronósticos usados son:

- Método de errores de predicción normales propuesto por Al-Qassam y Lane (1989). La predicción se estima mediante una ponderación de las predicciones correspondientes a cada uno de los regímenes del modelo SETAR empleado.
- Método de predicción adaptativa. La idea es emplear diferentes estimaciones de los parámetros para diferentes horizontes de predicción, como sugirió Cox (1961). En series temporales esta aproximación ha sido empleada para su uso en modelos lineales por (Tiao & Tsay, 1994).

Para calcular la predicción a horizonte h , el modelo es estimado minimizando la suma de los cuadrados de los errores de predicción dentro de la muestra a horizonte h .

- Métodos de Monte Carlo. Dado un modelo SETAR, es claro que la predicción a horizonte $h \leq d$ puede ser calculada sin ningún tipo de complicación, puesto que la variable umbral siempre será observada. Cuando el horizonte sea $h > d$ las predicciones serán calculadas simulando diferentes trayectorias posibles, que luego serán empleadas para estimar la predicción.
- Métodos Bootstrap. El funcionamiento es igual al detallado para el método de Monte Carlo, salvo en la generación del número aleatorio $\varepsilon^{(i)}$. En este caso, en vez de emplear una distribución, se realiza un muestreo con reemplazamiento entre los residuos del i -ésimo régimen obtenidos en la estimación del modelo.

Lin & Granger (1994) y Clements & Smith (1997) compararon varios de estos métodos para modelos STAR y SETAR, respectivamente. Estos autores encontraron que el método de Monte Carlo y el método de Bootstrap muestran un buen desempeño en la elaboración de pronósticos.

En este documento se utilizará el método Montecarlo para el cómputo de los pronósticos de los modelos considerados. Para ello, se simularán N valores para $Z_{t+1}, Z_{t+2}, \dots, Z_{t+h}$ y se obtendrá el pronóstico h pasos adelante como el promedio de estos valores. Para tener una idea de las complicaciones que surgen en este contexto, considere el proceso Z_t que obedece a un modelo SETAR es:

$$Z_t = F(1, Z_{t-1}, \dots, Z_{t-p_j}; \theta) + \varepsilon_t \quad (3.5)$$

donde el esqueleto F es una función no lineal de los parámetros θ .

El pronóstico óptimo h pasos adelante está dado por:

$$\hat{Z}_{t+h|t} = E[Z_{t+h} | \Omega_t] \quad (3.6)$$

donde Ω_t es el conjunto de información disponible hasta el periodo t.

El pronóstico óptimo un paso adelante equivale a:

$$\begin{aligned} \hat{Z}_{t+1|t} &= E[Z_{t+1} | \Omega_t] \\ E[Z_{t+1} | \Omega_t] &= E[F(1, Z_t, \dots, Z_{t-p_j+1}; \theta) + \varepsilon_{t+1} | \Omega_t] \\ E[Z_{t+1} | \Omega_t] &= F(1, Z_t, \dots, Z_{t-p_j+1}; \theta) \end{aligned} \quad (3.7)$$

y el pronóstico óptimo dos pasos adelante es:

$$\begin{aligned} \hat{Z}_{t+2|t} &= E[Z_{t+2} | \Omega_t] \\ E[Z_{t+2} | \Omega_t] &= E[F(1, Z_{t+1}, Z_t, \dots, Z_{t-p_j+2}; \theta) + \varepsilon_{t+2} | \Omega_t] \\ E[Z_{t+2} | \Omega_t] &= E[F(1, \hat{Z}_{t+1|t} + \varepsilon_{t+1}, Z_t, \dots, Z_{t-p_j+2}; \theta) | \Omega_t] \end{aligned} \quad (3.8)$$

Como puede verse, el pronóstico un paso adelante es equivalente al de un modelo AR lineal, sin embargo, no existen expresiones analíticas exactas para horizontes de predicción mayores a 1. Es posible aproximar la esperanza condicional usando técnicas de integración numérica, aunque esta solución no suele usarse mucho en la práctica debido a que se requieren secuencias computacionales intensivas.

A continuación se detallara el método de Montecarlo y el método de Bootstrap.

Método de Montecarlo

Una alternativa para calcular pronósticos h-pasos hacia adelante consiste en aproximar la esperanza condicional respectiva por medio de un procedimiento computacional recursivo del tipo Montecarlo, por ejemplo, el pronóstico dos pasos adelante se aproxima mediante la siguiente ecuación:

$$\hat{Z}_{MC, t+2} = \frac{1}{N} \sum_{j=1}^N F(\hat{Z}_{T+1} + G_j)$$

donde G_j son extraídos de una función de distribución D conocida.

Bajo este esquema los pronósticos de Montecarlo para Z_{t+h} con $h>1$ es de la forma:

$$\hat{Z}_{MC, t+h} = \frac{1}{N} \sum_{j=1}^N F(F(\dots F(\hat{Z}_{T+1} + G_j^1)\dots) + G_j^{h-1})$$

Con $Z_j^1, Z_j^2, \dots, Z_j^{h-1}$ generados de la distribución D, para $j=1, 2, \dots, N$. Cabe mencionar que este método es el de mayor uso en la práctica, dada la facilidad de implementación y las propiedades de convergencia para N suficientemente grande.

Método de Bootstrap

Similar al método de Montecarlo, el método de remuestreo consiste en aproximar la esperanza condicional, mediante la expresión:

$$\hat{Z}_{B, T+2} = \frac{1}{n-1} \sum_{j=1}^{n-1} F(\hat{Z}_{T+1} + \hat{\varepsilon}_j)$$

donde a diferencia del método de Montecarlo los $\hat{\varepsilon}_j, j=1, \dots, n-1$ son valores extraídos de los residuales observados $\hat{\varepsilon}_j$ sobre el periodo muestral.

Equivalentemente tenemos que el pronóstico h-pasos adelante se aproxima por medio de la siguiente igualdad:

$$\hat{Z}_{B,T+h} = \frac{1}{n-1} \sum_{j=1}^{n-1} F(F(\dots F(\hat{Z}_{T+1} + \hat{\varepsilon}_j^1)\dots) + \hat{\varepsilon}_j^{h-1}), \quad h > 1$$

donde $\hat{\varepsilon}_j^1, \hat{\varepsilon}_j^2, \dots, \hat{\varepsilon}_j^{h-1}$ son (h-1) valores aleatorios extraídos de los residuales observados, $j=1,2,\dots, n-1$.

En la práctica la función $F(\cdot)$ no es conocida y deberá ser especificada y estimada por medio de algún procedimiento de estimación. La función estimada $\hat{F}(\cdot)$ se sustituye en cada una de los métodos anteriores para obtener los pronósticos de interés.

Los pronósticos mediante el método exacto son computacionalmente costosos y complicados, mientras que los pronósticos mediante el método de Monte Carlo, pueden ser sesgados si seleccionamos una función de distribución incorrecta, sin embargo en condiciones generales suele ser la mejor técnica.

3.11 MODELO ARCH

Los modelos ARCH tienen su origen en un artículo de (Engle, 1982), en el que se pretendía obtener una predicción adecuada para la inflación en el Reino Unido, sujeta a fuerte volatilidad y con períodos de especial calma o de especial agitación.

Una de las contribuciones importantes de la literatura de procesos ARCH es mostrar que las variaciones que aparentemente se producen en la volatilidad de las series temporales económicas pueden explicarse mediante una determinada forma de dependencia no lineal, que permite además predecir dichos cambios en volatilidad sin necesidad de recurrir a la modelización explícita de cambios estructurales en la varianza.

Definición: Un modelo ARCH(r) está definido por

$$\begin{cases} Y_t = \sqrt{h_t} \varepsilon_t \\ h_t = \alpha_0 + \alpha_1 Y_{t-1}^2 + \dots + \alpha_r Y_{t-r}^2 \end{cases}$$

donde ε_t es una secuencia de variables aleatorias independientes e idénticamente distribuidas (iid) con media cero y varianza uno. Además se debe cumplir que $\alpha_0 > 0, \alpha_i \geq 0, i > 0$.

La condición $\alpha_0 > 0$ corresponde a la mínima varianza condicional a ser observada mientras que $0 < \alpha_0 < 1$ es una condición necesaria y suficiente para la existencia de la varianza incondicional y la condicional.

Procedimiento para la construcción de un modelo ARCH

- Construir un modelo ARMA para la serie dada con el fin de eliminar alguna dependencia lineal en la media y usar los residuales del modelo para determinar si existe efecto ARCH. Los residuos del modelo ARMA podrían estar correlacionados y posiblemente no serán independientes. Este efecto será visible en la función de autocorrelación de los residuos al cuadrado, que mostrarían la correlación serial.
- Especificar el orden del modelo ARCH, mediante la observación del gráfico de autocorrelación de residuos al cuadrado.
- Realizar el ajuste del modelo ARCH y validarlo.

3.12 MODELO GARCH

Un rasgo común a muchas de las primeras aplicaciones empíricas de los modelos ARCH es que requieren un gran número de parámetros autorregresivos y, para representar adecuadamente el comportamiento dinámico de la varianza, se imponía una estructura fija de retardos. (Bollerslev, 1986) con el fin de flexibilizar estas restricciones propuso el modelo ARCH generalizado o GARCH.

La generalización del modelo ARCH al modelo GARCH tiene gran similitud con la extensión de los procesos autorregresivos, AR, a los autorregresivos de medias móviles, ARMA, permitiendo una representación más parsimoniosa de la volatilidad.

Bollerslev considera que la varianza, h_t , además de depender de las observaciones pasadas de Y_t , depende también de su propio pasado. Esta dependencia se expresa incluyendo cierto número de retardos p de h_t , de forma que la varianza condicional se defina como:

$$h_t = \alpha_0 + \sum_{i=1}^r \alpha_i Y_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j}$$

Se pueden aplicar modelos GARCH en cualquiera de los diversos campos de la administración del riesgo, de la administración de portafolios, en la colocación de activos, en la valoración de opciones, en las tasas de cambio y en la estructura de tasas de interés.

Definición: Un modelo GARCH(r,s) está definido por

$$\begin{cases} Y_t = \sqrt{h_t} \varepsilon_t \\ h_t = \alpha_0 + \sum_{i=1}^r \alpha_i Y_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j} \end{cases}$$

donde ε_t es una secuencia de variables aleatorias independientes e idénticamente distribuidas (iid) con media cero y varianza uno. Además se debe cumplir que $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $\sum_{i=1}^q (\alpha_i + \beta_i) < 1$, $q = \max(r, s)$.

Procedimiento para la construcción de un modelo GARCH

- Construir un modelo ARMA para la serie dada con el fin de eliminar alguna dependencia lineal en la media y usar los residuales del modelo para determinar si existe efecto ARCH. Los residuos del modelo ARMA podrían estar incorrelacionados y posiblemente no serán independientes. Este efecto será visible en la función de autocorrelación de los residuos al cuadrado, que mostrarían correlación serial.
- Especificar el orden del modelo GARCH, mediante los criterios de información AIC y BIC.
- Validar el ajuste del modelo GARCH.

CAPÍTULO IV

APLICACIÓN

CAPÍTULO IV. APLICACIÓN

El PBI es una serie que posee un comportamiento muy característico, es decir, en épocas de recesión, la variable “Variación porcentual del PBI trimestral” es más aguda que en épocas de expansión; esto justifica la aplicación de modelos no lineales, los cuales ayudan a captar tal característica y permiten hacer una interpretación económica más adecuada para esta variable. Por otro lado la variación del PBI es de mucha utilidad para la economía de los países debido a que es considerado como el mejor indicador económico, que mide el crecimiento de un país basándose en la producción dentro de ella, tanto de empresas nacionales y extranjeras. De ahí que su modelización es de interés no reciente.

En diversos trabajos se presentan modelizaciones por series de tiempo mediante modelos ARIMA, ARCH y otros. En este capítulo se presentan los resultados obtenidos mediante la modelización SETAR se discuten los resultados obtenidos mediante estos en contraste con la estimación GARCH. Los gráficos fueron trabajados con el programa R 3.3.4 debido a que este programa cuenta con una librería especial para realizar los gráficos. La aplicación metodológica se realizó con el programa S-PLUS 8.0 versión demostración. Se eligió este programa porque tiene incorporado la librería **Finmetrics** la cual es especializada para el desarrollo de modelos de series de tiempo no lineal y por último la aplicación de los modelos GARCH se realizó con el programa EVIEWS 8.0.

4.1 ANÁLISIS EXPLORATORIO DE LA SERIE

Se observa que el PBI trimestral no es estacionario en nivel y ni en variabilidad (Ver figura 4.1). En lo que respecta al análisis de las componentes, se observa una tendencia creciente, estacionalidad determinística y existe una ligera sospecha de la presencia de la componente ciclo. Este tipo de comportamiento en series de tiempo se modela frecuentemente mediante modelos SARIMA.

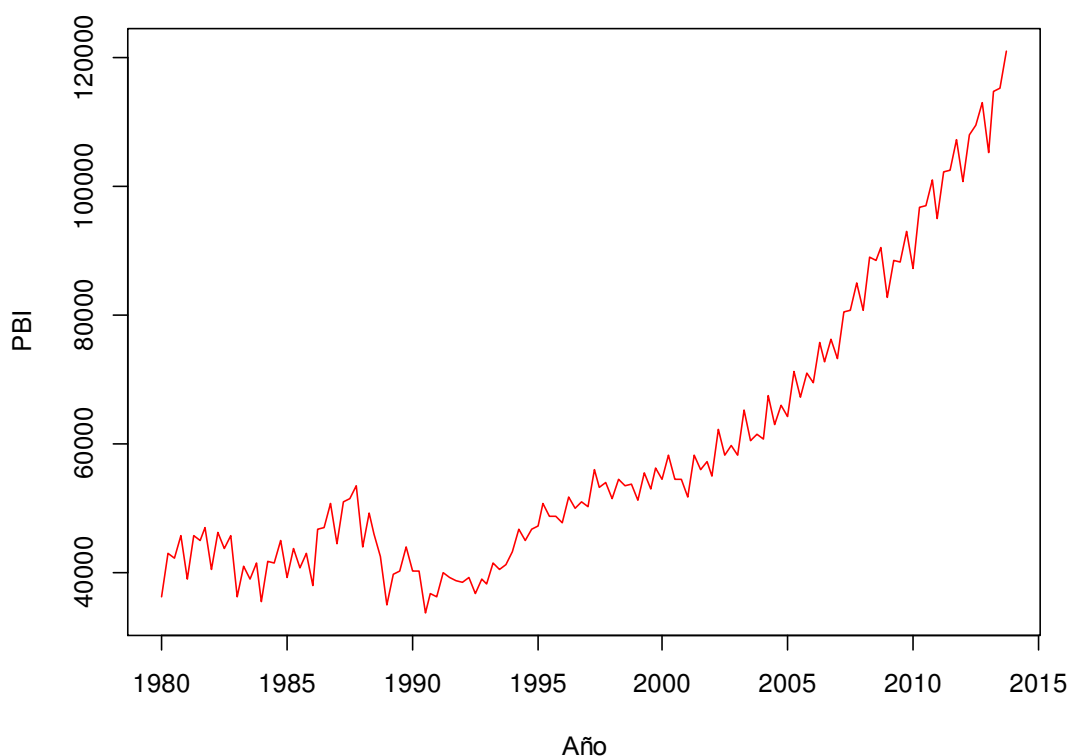


Figura 4.1: PBI trimestral, 1980-2014.

Fuente: INEI y BCRP.

La variación porcentual del PBI trimestral (Z_t) es de mayor uso debido a su gran aplicación en economía por la interpretación que se le da al crecimiento económico del país. Por este motivo la aplicación se realizará con esta serie.

Se observa que la serie variación porcentual del PBI trimestral (Z_t) es estacionario en nivel pero no en variabilidad (Ver figura 4.2). Otra forma de verificar la estacionariedad de la serie, es usando la prueba de Dickey-Fuller Aumentando (Ver anexo 5).

En lo que respecta a las componentes se observa una tendencia estable, existe una ligera sospecha de la presencia de la componente ciclo y es clara la presencia de la componente aleatoria.

El hecho de que nuestra serie no presente estacionariedad en la varianza, es un claro indicio de la no linealidad de nuestra serie, nuestro objetivo es modelar este tipo de comportamientos usando los modelos SETAR.

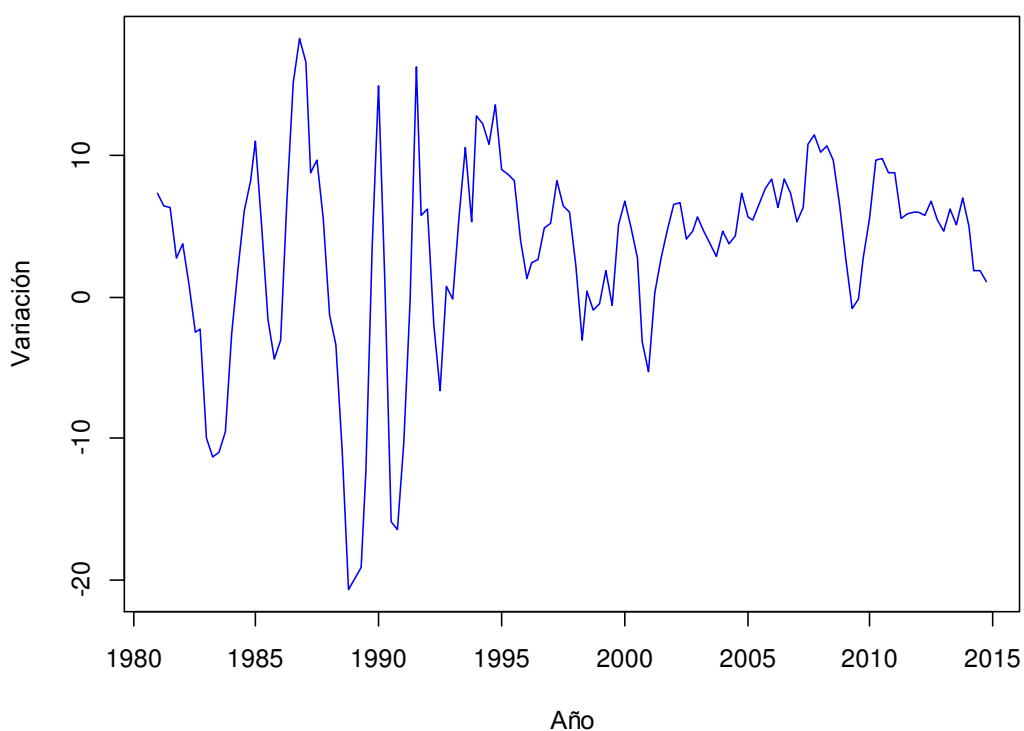


Figura 4.2: Variación porcentual del PBI, 1981-2014.

Fuente: INEI y BCRP.

La representación gráfica del periodograma nos proporciona un primer conocimiento de la estructura del modelo y nos sugiere modelos más apropiados: si existieran realmente periodos en la serie temporal, en las frecuencias correspondientes a dichos periodos (periodo=1/frecuencia) presentaría la gráfica picos muy pronunciados con grandes concentraciones de masas espectrales.

Según la figura 4.3 se puede observar que la intensidad máxima del periodograma no se presenta en la frecuencia (0,0) es decir que la tendencia no es la componente más importante de nuestra serie. Otra característica importante que se observa es una frecuencia máxima en el punto 13, cuyo periodo es 10 trimestres, es decir, la presencia de ciclos es cada 40 meses (3 años y medio). Para una mejor comprensión del periodograma revisar el anexo 6.

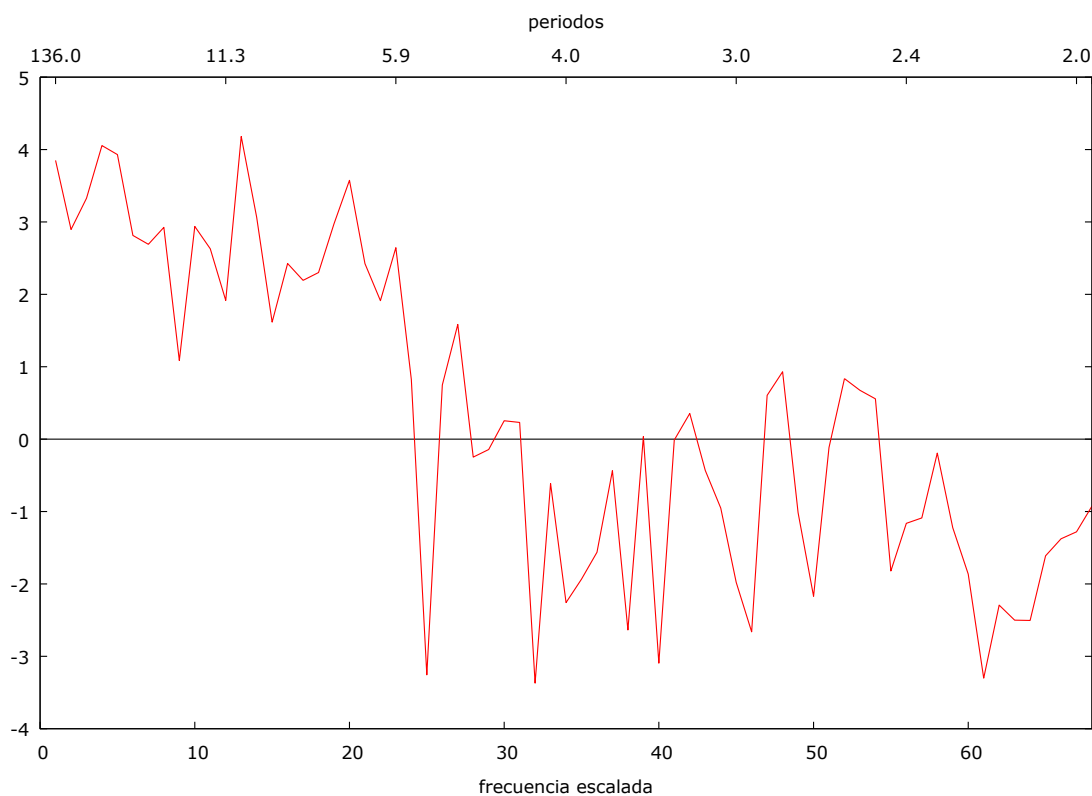


Figura 4.3: Periodograma de la serie Z_t .

Fuente: Elaboración propia.

La función de autocorrelación simple (ACF) de la figura 4.4 presenta un comportamiento sinusoidal, el cual es un típico comportamiento de los procesos autorregresivos, mientras que la función de autocorrelación parcial (ACP) es significativo para los primeros seis rezagos. Esto sugiere un modelo de AR con orden de 5 o 6 puede ser considerado como un punto de partida para el modelado de los datos.

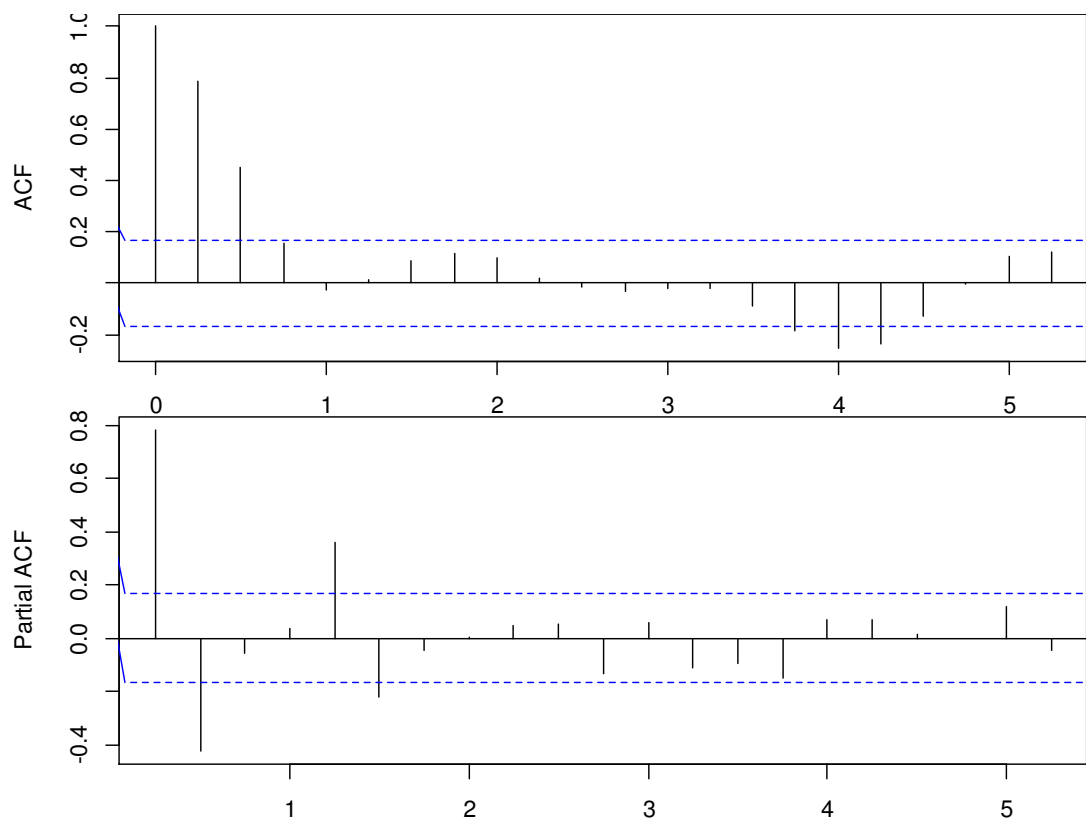


Figura 4.4: Función de autocorrelación de la serie Z_t .

Fuente: Elaboración propia.

Al analizar la función de autocorrelación parcial (ACP) de la serie, nos indicó que la serie se puede modelar con un AR (6), pues existe un valor significativo en el orden 6. Otra forma de identificar el proceso autorregresivo en usando los criterios de información AIC y BIC (Ver cuadro 4.1).

Cuadro 4.1: Identificación del orden del proceso

Proceso AR(p)	AIC	BIC
AR(1)	793.4649	802.2029
AR(2)	768.2445	779.8951
AR(3)	769.8777	784.4410
AR(4)	771.7574	789.2333
AR(5)	754.5349	774.9235
AR(6)	749.0301	772.3313
AR(7)	750.8408	777.0547
AR(8)	752.8393	781.9659
AR(9)	754.7003	786.7395
AR(10)	756.2538	791.2057

Fuente: Elaboración propia.

4.2 PRUEBA DE LINEALIDAD

Identificar un fenómeno no lineal particular suele ser una tarea compleja. Para ello se pueden emplear diferentes métodos como el análisis de los datos mediante métodos gráficos y la realización de contrastes de linealidad.

4.2.1 Análisis exploratorio de la linealidad

Mediante el análisis exploratorio se trata de detectar la no linealidad mediante el uso de métodos gráficos. Se presentan los resultados de los gráficos de dispersión directa que se desarrollan para diferentes retardos y los gráficos comparativos entre la serie original y sus diferentes desfases en el tiempo. En ambos casos se trabajan hasta 3 retardos debido a que en $d=3$ es evidente el comportamiento no lineal.

En los gráficos de dispersión directa, se espera que el comportamiento de los puntos unidos por líneas deje espacios vacíos dibujando una elipse (Ver figura 3.1), en tanto que el gráfico para los modelos lineales muestra un comportamiento caótico por la ausencia de espacios vacíos.

Observando los gráficos de dispersión directa para $d=1,2,3$ (Ver figura 4.5) encontramos un típico comportamiento de los modelos no lineales, pues se puede observar espacios vacíos de diversas formas elípticas.

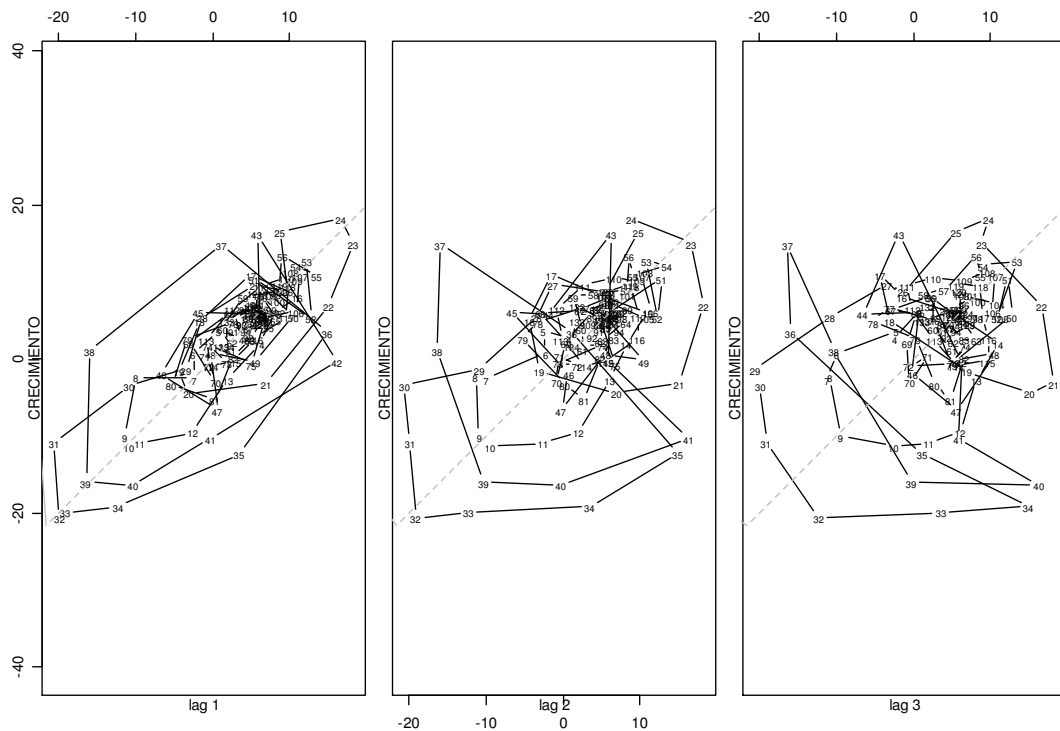


Figura 4.5: Diagramas de dispersión directa de la serie Z_t .

Fuente: Elaboración propia.

Una herramienta útil para identificar una posible relación no lineal es usando el suavizado no paramétrico y el ajuste lineal. En la figura 4.6 se observa que la serie temporal Z_t y la serie desfasada Z_{t-d} , (donde $d=1, 2, 3, 4$) poseen una relación no lineal, es decir, la serie en estudio es un modelo no lineal en el tiempo, además la no linealidad se hace más significativa al aumentar el orden del retardo.

Como vemos, si realizamos un ajuste lineal será complicado que se pueda predecir su comportamiento futuro adecuadamente. Por ello, es importante identificar si existe algún fenómeno no lineal subyacente en los datos para así poder aplicar métodos más adecuados.

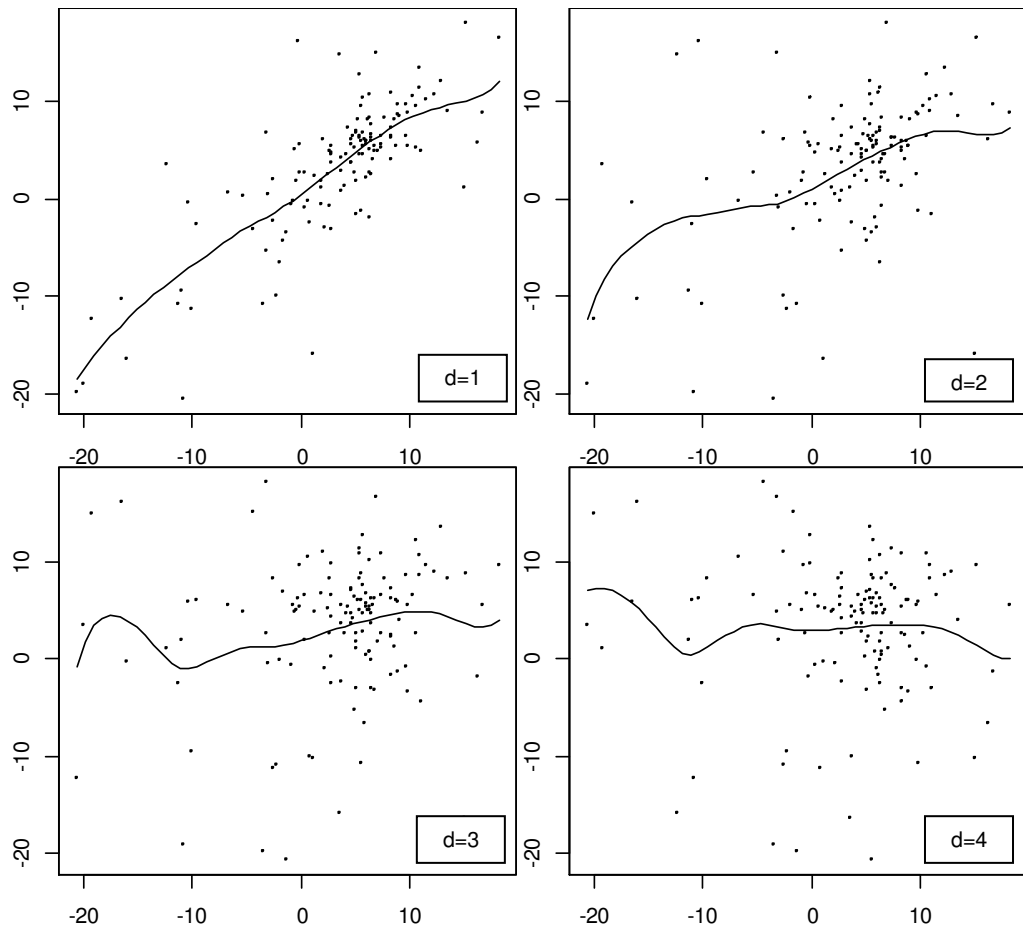


Figura 4.6: Suavizado no paramétrico de la serie Z_{t-d} .

Fuente: Elaboración propia.

4.2.2 Prueba de la linealidad de tipo umbral (Prueba de Tsay)

Una herramienta más formal para comprobar la existencia de no linealidad en el proceso es el uso de contrastes de hipótesis. Existen dos clases de contrastes de linealidad, por una parte están los contrastes generales que no asumen ningún tipo de estructura no lineal en los datos, sino que solo contrastan la falta de linealidad en los mismos. Por otra parte, están aquellos contrastes específicos que nos permiten comprobar si un tipo de fenómeno no lineal en particular es correcto para el proceso estudiado.

A continuación se muestra la prueba de no linealidad de tipo umbral propuesta por (Tsay, 1989).

H₀: La serie Z_t presenta linealidad.

H₁: La serie Z_t presenta no linealidad de tipo umbral.

En el cuadro 4.2 se observa la salida de las estadísticas F y su p-valores correspondientes para todos los valores seleccionados del parámetro de retardo d, evidencian que la no linealidad umbral es muy fuerte con la especificación AR (6). Concluimos que a un nivel de significación del 5% se rechaza la hipótesis nula, es decir, nuestra serie bajo estudio presenta no linealidad de tipo umbral.

Cuadro 4.2: Prueba de linealidad de Tsay cuando p=6

Retardo	Estadístico F	P-valor
d=1	4.9307	0.0001
d=2	3.3254	0.0031
d=3	2.1614	0.0436
d=4	4.4815	0,0002
d=5	5.0627	0.0001
d=6	8.5019	0.0000

Fuente: Elaboración propia.

4.2.3 Prueba de la linealidad (Prueba de Hansen)

H₀: La serie Z_t presenta linealidad.

H₁: La serie Z_t presenta no linealidad de tipo umbral.

En el cuadro 4.3 se observa la salida de los estadísticos y sus correspondientes valores para el parámetro de retardo d=1 y p=6, evidencian que la no linealidad umbral es muy fuerte con la especificación AR (6). Concluimos que a un nivel de significación del 5% se rechaza la hipótesis nula, es decir, nuestra serie bajo estudio presenta no linealidad de tipo umbral.

Cuadro 4.3: Prueba de linealidad de Hansen cuando $p=6$ y $d=1$

Estadísticos	Valores
Number of Bootstrap Replications	1000
Trimming percentage	0.1
F-test for no threshold	62.0128
Bootstrap P-Value	0.0

Fuente: Elaboración propia.

4.3 DETERMINACIÓN DE LOS PARÁMETROS ESTRUCTURALES

Después de comprobar la no linealidad umbral de nuestra serie en estudio, el siguiente procedimiento es identificar el parámetro de retardo d y los umbrales r , y posteriormente, utilizar los mínimos cuadrados (MC) para estimar los parámetros desconocidos dados en (3.1) con valores dados del retardo d y del valor umbral r .

4.3.1 Determinación de la variable umbral

El criterio propuesto para seleccionar la variable umbral es buscar el retardo d que maximice el estadístico definido en (3.4), es decir $F(p, d) = \text{Max}_{v \in S} \{F(p, v)\}$, donde S es el conjunto de los posibles retardos y $d \leq p$.

Del cuadro 4.4 se observa que el retardo $d=6$ posee el mayor estadístico F sin embargo este caso no es factible debido a que debe cumplir que $d < p$. La segunda opción sería el retardo $d=5$ sin embargo las ganancias en cuanto a estimación son muy bajas. La otra opción es considerar el retardo $d=1$ el cual nos proporciona un estadístico F significativo.

Esta última afirmación se comprueba usando la prueba de razón de verosimilitud (Ver cuadro 4.5), el cual es una propuesta para determinar la variable umbral incorporada en la librería TSA del programa R-PROJECT, con lo cual concluimos que la variable umbral es Z_{t-1} .

Cuadro 4.4: Determinación de la variable umbral mediante S-PLUS

Retardo	Estadístico F	P-valor	Variable umbral
d=1	4.9307	0.0029	Zt-1
d=2	3.3254	0.0334	Zt-2
d=3	2.1614	0.0436	Zt-3
d=4	4.4815	0.0002	Zt-4
d=5	5.0627	0.0001	Zt-5
d=6	8.5019	0.0000	Zt-6

Fuente: Elaboración propia.

Cuadro 4.5: Determinación de la variable umbral mediante R-PROJECT

Retardo	Estadístico_RV	P-valor	Variable umbral
d=1	62.4550	0.0000	Zt-1
d=2	22.9030	0.0270	Zt-2
d=3	10.4660	0.3500	Zt-3
d=4	18.7520	0.0950	Zt-4
d=5	22.5720	0.0310	Zt-5
d=6	35.9730	0.0000	Zt-6

RV: Razón de verosimilitud.

Fuente: Elaboración propia.

4.3.2 Región de variación para los valores del umbral

Para determinar los valores umbrales, (Tsay, 1989) considera de utilidad los gráficos que representan los estadísticos T de los distintos coeficientes autorregresivos en las autorregresiones tipificadas respecto a la variable umbral Z_{t-d} , ya que los mayores cambios en la pendiente de estos valores T sugieren cambios de régimen.

En la figura 4.7 se observa la tendencia en las estadísticas T parece tener dos descansos: uno se produce cuando la variable de umbral esta alrededor de 5; y el otro se produce cuando la variable de umbral es alrededor de 7, lo cual indica que el valor umbral se encontrara en la vecindad de 5 a 7. De manera gráfica los posibles valores del umbral serian $r_1=5$, $r_2=5.5$, $r_3=6$, $r_3=6.5$ y $r_4=7$.

Más adelante se obtendrá el valor umbral óptimo mediante la utilización de los criterios de información AIC y BIC.

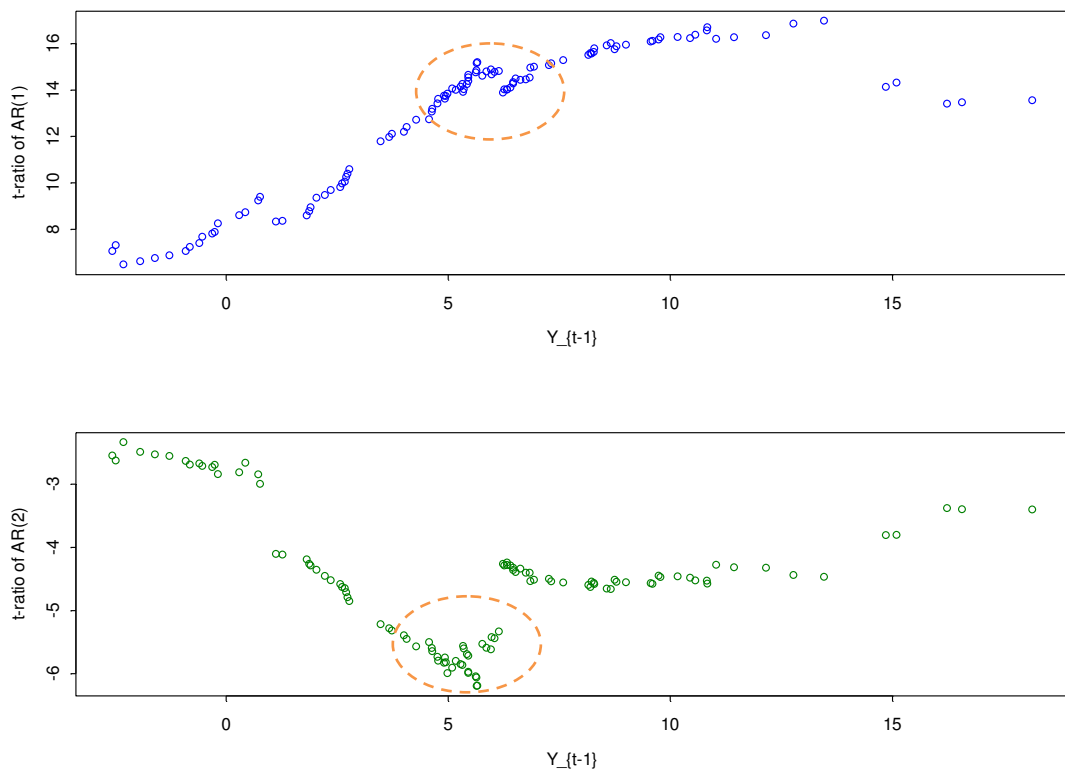


Figura 4.7: Posibles valores umbrales para el retardo $d=1$.

Fuente: Elaboración propia.

(Tsay, 1989) propuso la detección de los umbrales solo de forma gráfica, tal como hemos venido trabajando hasta este punto de la investigación. Sin embargo como una contribución propia procederemos a ampliar la detección de umbrales mediante los criterios AIC y BIC ya que contamos con las herramientas computacionales que permiten aplicar esta metodología.

4.3.3 Localización óptima de los valores umbrales

Una vez definido el valor del retardo (d) y la variable umbral (Z_{t-d}) se procede a calcular el siguiente parámetro estructural denominado valor umbral (r). Se usara el criterio el criterio AIC (Criterio de Información de Akaike) y el criterio BIC (Criterio de Información Bayesiano) para determinar el valor umbral óptimo.

Como se mencionó anteriormente Tsay propuso un método gráfico para determinar el valor umbral, sin embargo la contribución de Tsay solo nos proporciona una aproximación del valor umbral, pero gracias a esta contribución se puede obtener una vecindad de posibles valores umbrales.

A continuación se determinara el valor umbral óptimo considerando los criterios de información AIC y BIC, es decir, se elegirá aquel valor umbral que minimice los criterios de información antes mencionados.

Del cuadro 4.6 se observa que el valor umbral óptimo es $r=5.6$, debido a presentar los menores valores en los criterios de información.

Cuadro 4.6: Localización óptima de los valores del umbral

Umbral	AIC	BIC
5.0	685.2411	725.3866
5.1	686.2350	726.3805
5.2	687.8962	728.0417
5.3	688.0942	728.2397
5.4	688.8582	729.0037
5.5	681.5813	721.7268
5.6	681.5810	721.7268
5.7	681.7960	721.9415
5.8	687.8415	727.9869
5.9	687.8372	727.9827
6.0	688.4312	728.5767
6.1	688.4779	728.6234
6.2	687.8809	728.0263
6.3	695.6837	735.8292
6.4	698.8986	739.0440
6.5	695.8893	736.0348

Fuente: Elaboración propia.

4.3.4 Localización óptima del valor umbral usando la prueba de Hansen

Otro procedimiento para la estimación del valor umbral es la prueba de Hansen, el proceso anterior requiere procedimientos que dificultan la elección del valor umbral en general de los parámetros del modelo. Este inciso introduce otra prueba para la estimación de los parámetros estructurales de los modelos SETAR.

Del cuadro N°4.7 se observa que el valor umbral óptimo es $r=5.648$, la precisión del valor umbral dependerá del número de réplicas. Esta prueba se realizó bajo el supuesto de que los residuos son homocedásticos.

Cuadro 4.7: Valor del umbral según la prueba de Hansen

Estadísticos	Valores
Number of Bootstrap Replications	1000
Trimming percentage	0.1
Threshold Estimate	5.648
F-test for no threshold	62.0128
Bootstrap P-Value	0.0

Fuente: Elaboración propia.

Otra contribución de Hansen fue proponer una forma de determinar el valor aproximado del umbral usando el grafico de intervalo de confianza para el valor umbral mediante la inversión de las estadísticas de razón de verosimilitud, el cual nos permite determinar la vecindad de posibles valores umbrales (Ver figura 4.8).

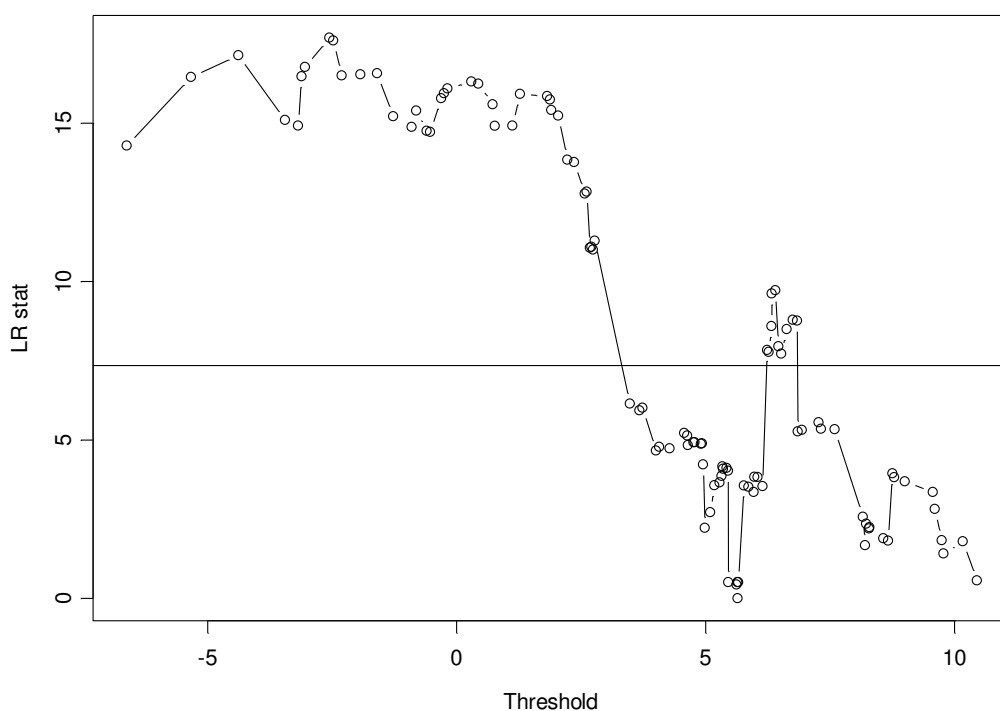


Figura 4.8: Intervalo de confianza para el valor umbral.

Fuente: Elaboración propia.

De lo anterior podemos concluir que la estimación del valor umbral óptimo (r), se basó en la propuesta inicial de Tsay, el cual consistía en determinar el posible valor umbral de manera gráfica. En el presente trabajo de investigación se determinó el valor umbral óptimo considerando una vecindad de posibles valores umbrales y se procedió a descartar las posibilidades usando los criterios de información AIC y BIC.

La validez de los resultados se comprobó usando la prueba de Hansen, el cual nos proporcionó el mismo resultado $r=5.648$ (Ver cuadro 4.7).

4.4 ESTIMACIONES MÍNIMOS CUADRÁTICOS DEL MODELO SETAR

Es necesario emplear un método de estimación recursiva para calcular los estimadores de los parámetros variables en el tiempo. Existen varios algoritmos que permiten hacerlo, como son mínimos cuadrados recursivos (MCR) o el filtro de Kalman. El método MCR es especialmente útil cuando los cambios en los parámetros ocurren lentamente a lo largo del índice en el que estén ordenadas las observaciones, es decir, los cambios se pueden producir de manera temporal, o como será nuestro caso, se producirán debido a que la variable que puede producir dichos cambios en el parámetro está ordenada. El inconveniente del filtro de Kalman es que para emplearlo se necesita conocer el modelo que sigue la variación del parámetro.

En el cuadro 4.8 se observa que los coeficientes para el primer régimen estimado son significativos, mientras que los coeficientes para el segundo régimen existen coeficientes no significativos, los cuales serán considerados en el análisis, pues descartar coeficientes en función del p-valor no siempre es recomendable, se recomienda adecuarse y verificar con la realidad.

Cuadro 4.8: Estimación de los coeficientes de cada régimen

Régimen	Coefficientes	Valor	Error std.	Estadístico t	P-valor
I	C	1.6933	0.4271	3.9647	0.0001
	Yt-1	1.5986	0.1055	15.1481	0.0000
	Yt-2	-0.9195	0.1484	-6.1968	0.0000
	Yt-3	0.3480	0.1463	2.3786	0.0188
	Yt-4	-0.6547	0.1270	-5.1547	0.0000
	Yt-5	0.7799	0.1265	6.1657	0.0000
	Yt-6	-0.4642	0.0921	-5.0426	0.0000
II	C	2.1034	1.4986	1.4036	0.1642
	Yt-1	0.4763	0.1758	2.7098	0.0079
	Yt-2	0.2293	0.1673	1.3709	0.1741
	Yt-3	-0.0591	0.1645	-0.3592	0.7206
	Yt-4	-0.2358	0.2096	-1.1250	0.2641
	Yt-5	0.1844	0.2071	0.8905	0.3762
	Yt-6	0.1556	0.1379	1.1284	0.2627

Fuente: Elaboración propia.

Por lo tanto la ecuación del modelo es un SETAR (2, 6,6):

$$Z_t = \begin{cases} 1.6933 + 1.5986Z_{t-1} - 0.9195Z_{t-2} + 0.3480Z_{t-3} - 0.6547Z_{t-4} \\ + 0.7799Z_{t-5} - 0.4642Z_{t-6} + a_{t,1} ; Z_{t-1} \leq 5.648 \\ 2.1034 + 0.4763Z_{t-1} + 0.2293Z_{t-2} - 0.0591Z_{t-3} - 0.2358Z_{t-4} \\ + 0.1844Z_{t-5} + 0.1556Z_{t-6} + a_{t,2} ; Z_{t-1} > 5.648 \end{cases}$$

4.5 VALIDACIÓN DEL MODELO SETAR

Después de estimar los parámetros del modelo se evalúan las propiedades del modelo ajustado y la validación de algunos supuestos básicos, para tal propósito se usaran los residuos del modelo.

Para evaluar la normalidad, se usará la prueba de Kolmogorov- Smirnov y el gráfico de cuantiles (Q-Q Normal). Para evaluar la incorrelación de los residuos se trabajará con la prueba de Ljung & Box además de las pruebas gráficas (correlogramas residuales).

Para evaluar la distribución de los residuos se realizó el gráfico de densidad espectral. Finalmente para verificar la adecuación del modelo se graficó la serie original versus la serie estimada.

4.5.1 Prueba de normalidad de los residuos

H₀: Los residuos del modelo SETAR (2, 6,6) poseen distribución normal.

H₁: Los residuos del modelo SETAR (2, 6,6) no poseen distribución normal.

Realizando las pruebas paramétricas respectivas se concluye que a un nivel de significancia del 5%, no existe suficientes evidencias para rechazar la hipótesis nula (Ver cuadro 4.9), es decir, los residuos del modelo SETAR (2, 6, 6) poseen distribución normal.

De manera gráfica también se puede observar la normalidad de los residuos, el gráfico Q-Q Normal (Ver figura 4.9) nos muestra un conjunto de datos concentrados alrededor de la línea de tendencia, lo cual es un indicio para suponer la normalidad. Similar comportamiento se puede observar en la figura 4.10, los residuos se ajustan a la curva normal.

Cuadro 4.9: Normalidad de los residuos

Prueba	Estadístico	P-valor
Kolmogorov-Smirnov	0.0722	0.0928
Shapiro-Wilk	0.9754	0.2234

Fuente: Elaboración propia.

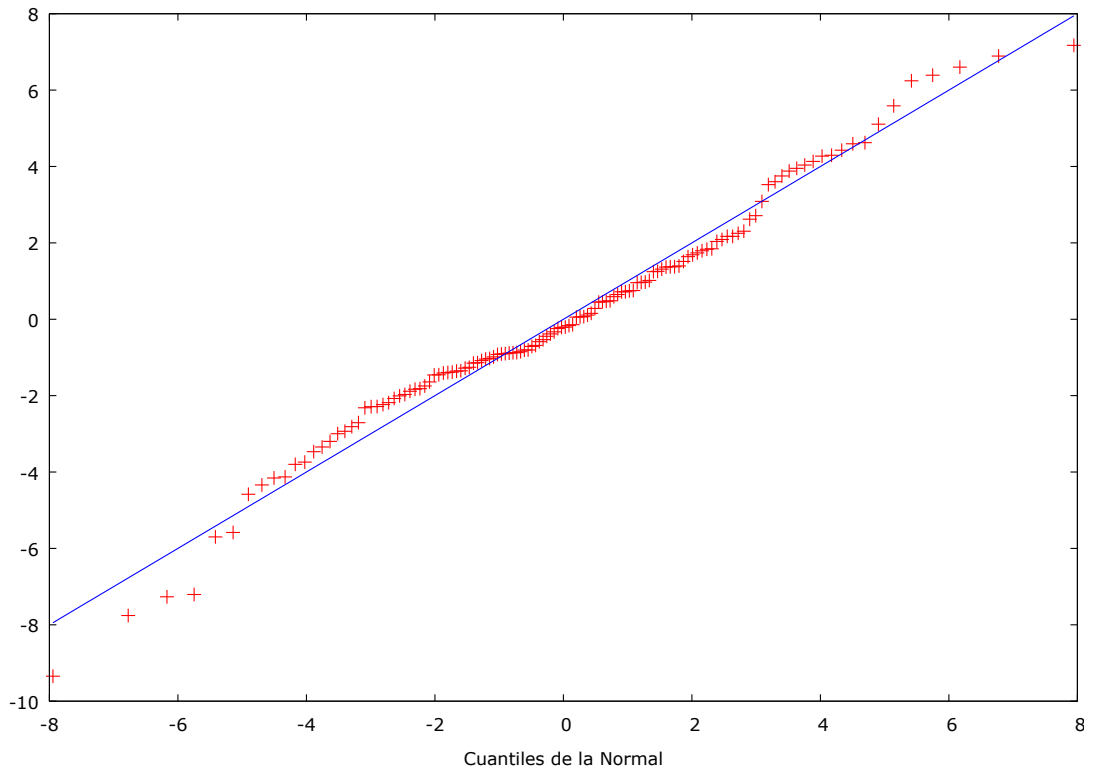


Figura 4.9: Prueba Q-Q Normal de los residuos.

Fuente: Elaboración propia.

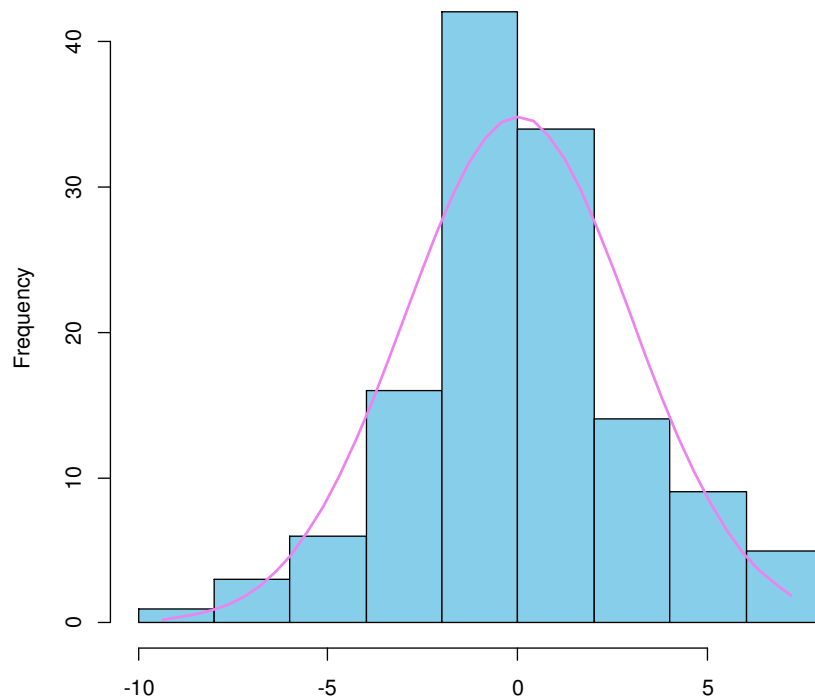


Figura 4.10: Histograma de los residuos y curva de densidad normal.

Fuente: Elaboración propia.

4.5.2 Prueba de incorrelación de los residuos

Uno de los supuestos básicos en el análisis de series temporales es la incorrelación de los residuos, bajo este contexto comprobaremos la incorrelación de los residuos del modelo SETAR (2, 6,6).

H₀: Los residuos del modelo SETAR (2, 6,6) son incorrelacionados.

H₁: Los residuos del modelo SETAR (2, 6,6) son correlacionados.

Según las prueba paramétricas de Box & Pierce y Box & Ljung y considerando un nivel de significancia del 5%, no se rechaza la hipótesis nula, es decir los residuos del modelo SETAR (2, 6,6) cumplen con el supuesto de incorrelación de los residuos (Ver cuadro 4.10).

Cuadro 4.10: Incorrelación de los residuos

Prueba	Estadístico	P-valor
Box & Pierce	8.0904	0.7052
Box & Ljung	8.7395	0.6459

Fuente: Elaboración propia.

Otra forma de validar el modelo es usando el correlograma de los residuos, se espera que los residuos se comporten como ruido blanco, es decir las autocorrelaciones deben estar dentro de las bandas de confianza.

Para nuestro caso se observa que las autocorrelaciones están dentro de las bandas de confianza, lo cual nos indica que los residuos se comportan como ruido blanco (Ver figura 4.11).

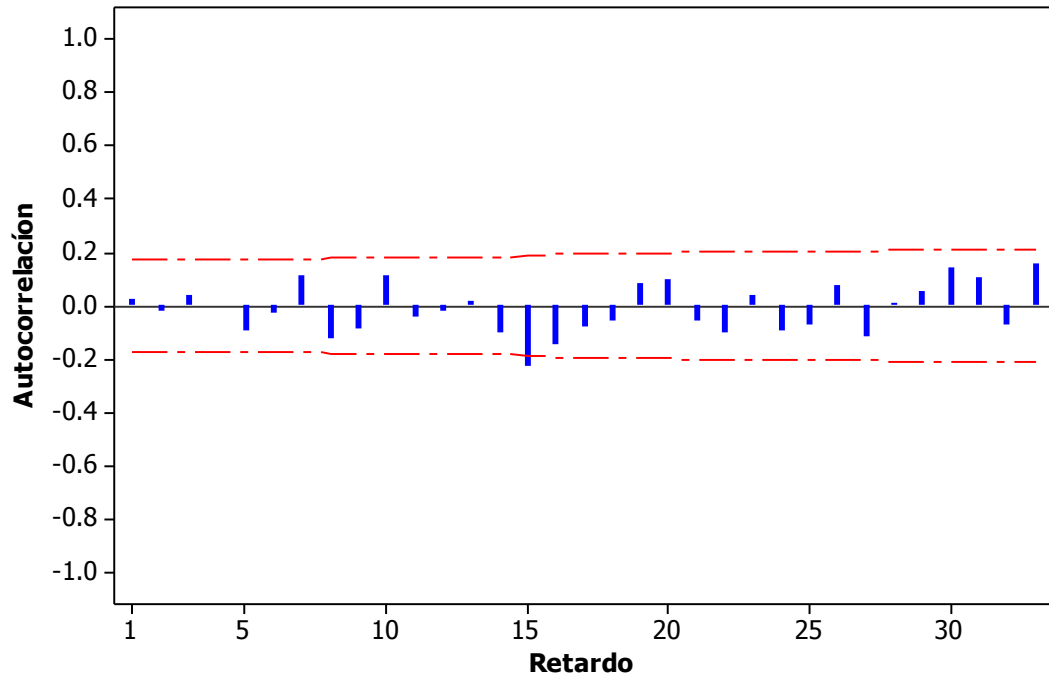


Figura 4.11: Correlograma de los residuos.

Fuente: Elaboración propia.

A continuación se validará el modelo ajustado haciendo uso del espectro de los residuos, la idea básica es que el espectro no debe de presentar patrones de comportamiento, es decir, debe presentar un comportamiento constante.

De la figura 4.12 se concluye que el resultado es un espectro en el que prevalece la parte aleatoria, además podemos observar que los residuos no presentan información respecto a la tendencia y ciclos, sin embargo hay algo de la estacionalidad que no ha sido explicada por el modelo, lo cual no es tan significativo. En consecuencia el modelo es apropiado.

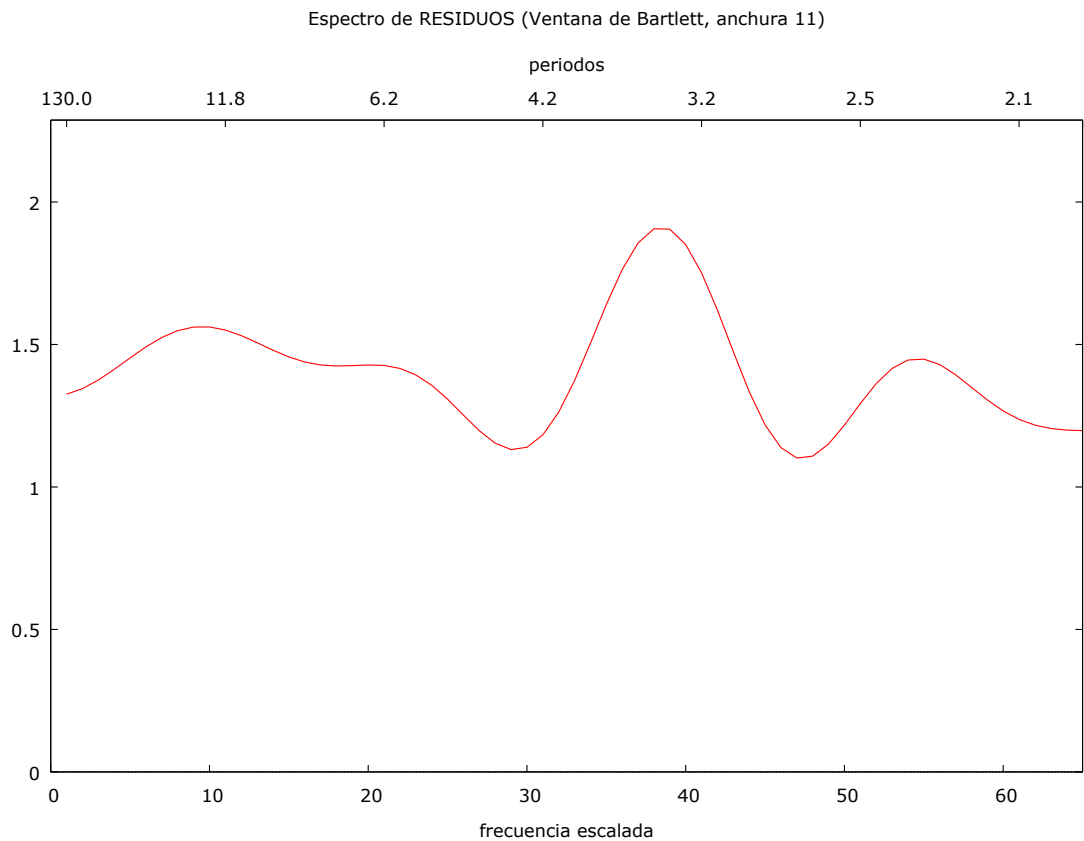


Figura 4.12: Espectro de los residuos del modelo SETAR (2, 6,6).

Fuente: Elaboración propia.

Algunos investigadores plantean validar el modelo observando el grado de ajuste de la serie estimada versus la serie original, según la figura 4.13 se observa que la propiedad más importante del modelo propuesto es la capacidad de adaptación al cambio de proceso de nuestra serie, lo cual nos indica que el modelo es adecuado para realizar las estimaciones y los pronósticos relacionados a la variación porcentual del PBI trimestral.

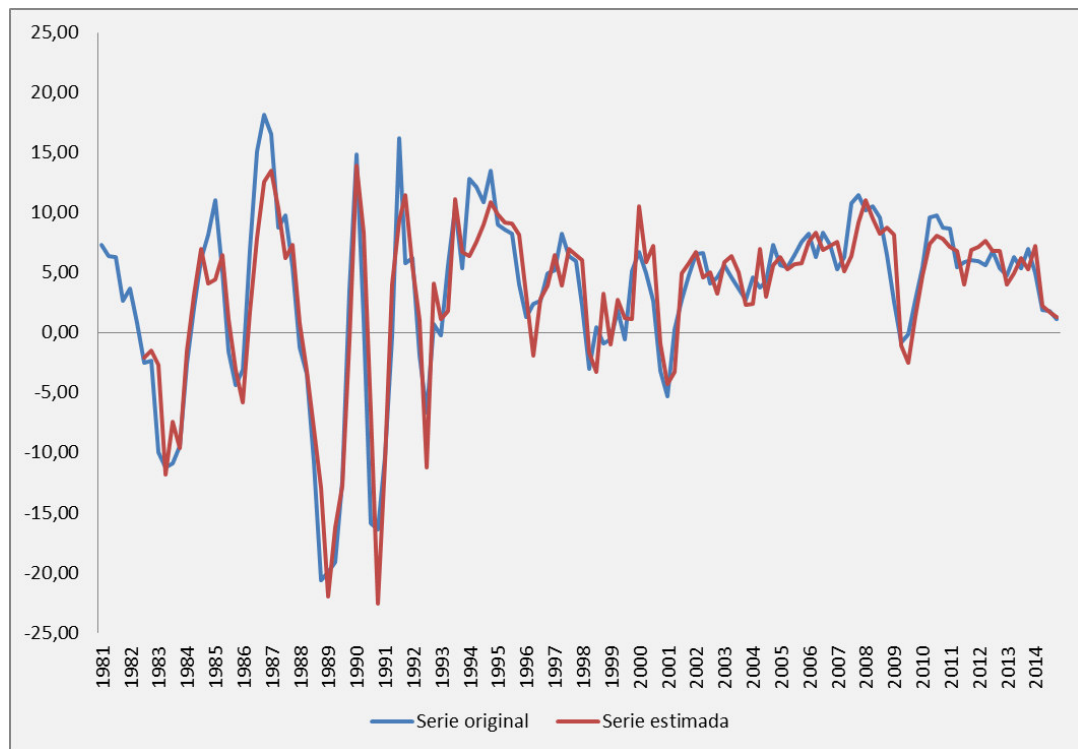


Figura 4.13: Serie original vs Serie estimada.

Fuente: Elaboración propia.

4.6 PRONÓSTICOS EN LOS MODELOS SETAR

La utilidad de los modelos con umbrales va más allá de su uso como función a trozos. Una de sus principales ventajas es que dependiendo de la complejidad de la estructura identificada, un modelo con umbrales puede servir para modelar funciones no lineales más complejas. Por este motivo los modelos de series de tiempo no lineales han sido ampliamente usados para pronosticar.

Para evaluar la capacidad de predictiva del modelo se computaron pronósticos fuera de la muestra por medio de la simulación de Montecarlo. La evaluación fue realizada contrastando pronósticos de valores estimados, con sus valores reales, para diferentes orígenes y horizontes de predicción. Se computaron pronósticos cuatro pasos adelante, es decir, para los cuatro trimestres del año 2015.

Los pronósticos obtenidos según el método de Montecarlo proporcionaron valores muy similares a los valores reales, esto debido a que el método trabaja con un número elevado de réplicas (Ver cuadro 4.11).

Cuadro 4.11: Pronósticos usando el método de Montecarlo

Período	Valor real	Pronóstico
2015-I	1.8467	2.2176
2015-II	3.1258	4.2386
2015-III	3.2359	3.5606
2015-IV	4.6792	2.8672

Fuente: Elaboración propia.

A continuación se muestra la serie original comparado con la serie estimada (incluido los pronósticos) a fin de determinar el buen ajuste de nuestro modelo. Se observa que la serie estimada posee similar comportamiento que la serie original, más aun los pronósticos conservan el patrón de la serie original (Ver figura 4.14), lo cual nos indica que el modelo SETAR (2, 6,6) es un buen modelo.

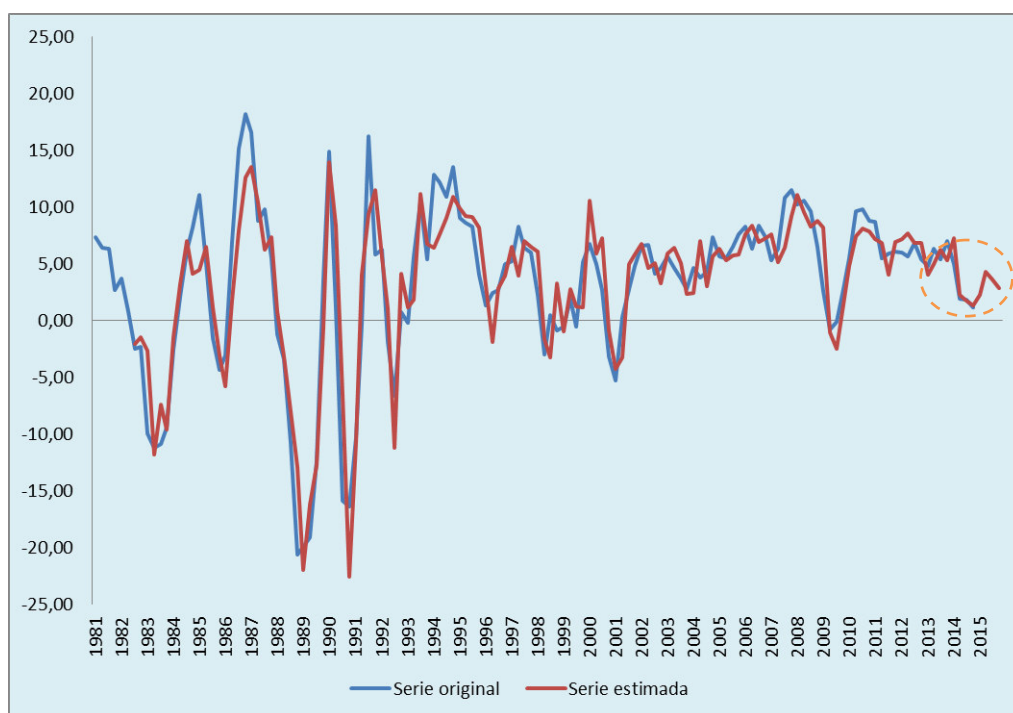


Figura 4.14: Estimación de la serie y pronósticos.

Fuente: Elaboración propia.

4.7 APLICACIÓN MEDIANTE EL MODELO GARCH

Una serie puede tener estructura no lineal no solo en la media sino también en la varianza, por ejemplo con datos financieros, económicos y ambientales donde se pueden observar rachas de alta volatilidad

El modelo GARCH ha aumentado de popularidad porque reconoce que la varianza no es constante, como ya está reconocido; la volatilidad tiende a tener períodos tranquilos de volatilidad y periodos de alta volatilidad. La conclusión obtenida de esto es que la volatilidad depende por sí misma de la volatilidad pasada y es, por ello, diferente a los modelos de movimiento Brownianos, que asumen que las fluctuaciones futuras no están influidas por eventos pasados.

A continuación se ajustara un modelo GARCH a nuestro conjunto de datos a fin de comparar con los resultados obtenidos por el modelo SETAR.

4.7.1 Descripción de la serie

En la figura 4.15 se observa que la serie variación porcentual del PBI trimestral (Z_t) es estacionario en nivel pero no en variabilidad, en lo que respecta a los componentes se observa una tendencia estable, existe una ligera sospecha de la presencia de la componente ciclo y se observa claramente la presencia de la componente aleatoria.

El hecho de que nuestra serie no presente estacionariedad en la varianza, es un claro indicio de la no linealidad de nuestra serie, nuestro objetivo es modelar este tipo de comportamientos usando los modelos de heterocedasticidad condicional.

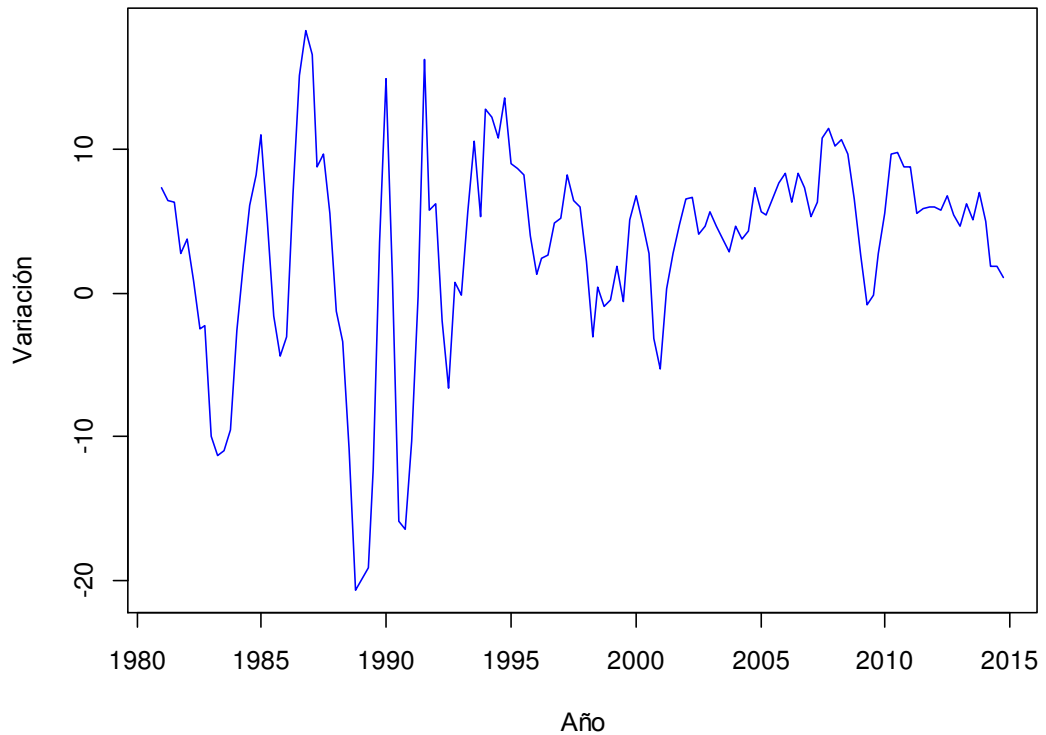


Figura 4.15: Variación porcentual del PBI, 1981-2014.

Fuente: INEI y BCRP.

4.7.2 Identificación del modelo ARMA

La función de autocorrelación simple (ACF) de la figura 4.16 es un típico comportamiento de los procesos autorregresivos y presenta un comportamiento sinusoidal, mientras que la función de autocorrelación parcial (ACP) es significativo para los primeros seis rezagos. Esto sugiere un modelo AR de orden 6 como un punto de partida para el modelado de los datos.

Al analizar los criterios de información AIC y BIC (Ver cuadro 4.12) encontramos que el proceso AR (6) presenta los menores criterios de información, motivo por el cual se proseguirá con la modelización considerando un proceso autorregresivo de orden seis.

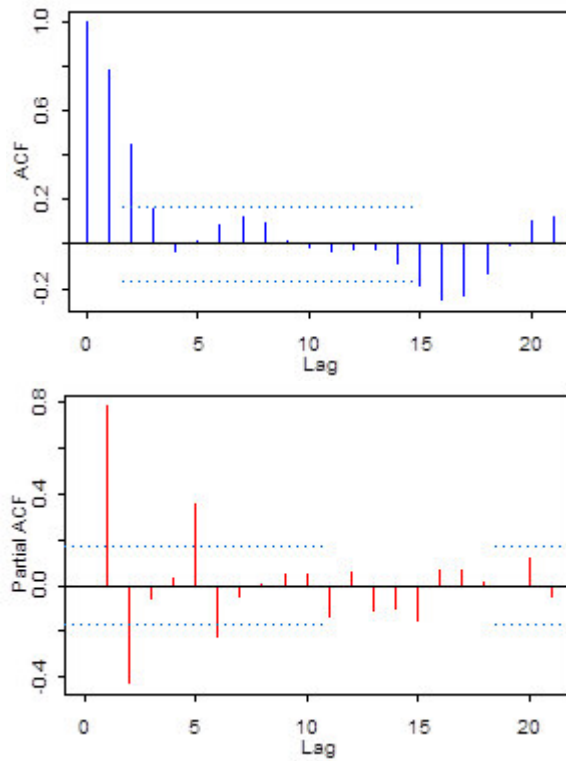


Figura 4.16: Función de autocorrelación de la serie Z_t .

Fuente: Elaboración propia.

Cuadro 4.12: Identificación del orden del proceso

Proceso AR(p)	AIC	BIC
AR(1)	793.4649	802.2029
AR(2)	768.2445	779.8951
AR(3)	769.8777	784.4410
AR(4)	771.7574	789.2333
AR(5)	754.5349	774.9235
AR(6)	749.0301	772.3313
AR(7)	750.8408	777.0547
AR(8)	752.8393	781.9659
AR(9)	754.7003	786.7395
AR(10)	756.2538	791.2057

Fuente: Elaboración propia.

Del cuadro 4.13 se observa que los coeficientes del modelo AR (6) son significativos a excepción del tercer retardo. Se optó por este modelo debido a que posee los menores criterios de información AIC y BIC (Ver cuadro 4.12).

Cuadro 4.13: Estimación del modelo AR (6)

Variable	Coefficiente	Std. Error	t-Statistic	P-valor
C	3.332158	1.349628	2.468946	0.0149
AR(1)	1.182651	0.087309	13.54557	0.0000
AR(2)	-0.426571	0.125201	-3.407091	0.0009
AR(3)	0.066617	0.124025	0.537127	0.5922
AR(4)	-0.455217	0.123742	-3.678754	0.0003
AR(5)	0.625672	0.124891	5.009741	0.0000
AR(6)	-0.234286	0.087157	-2.688087	0.0082

Fuente: Elaboración propia.

4.7.3 Evaluación de la presencia de heterocedasticidad condicional

Una vez ajustado el modelo sugerido por la estructura de autocorrelación de la serie en estudio, se procederá a evaluar la adecuación del modelo mediante el análisis de los residuos. Nuestro principal interés es averiguar si los residuos son condicionalmente heterocedásticos. Si lo son, procederemos a identificar un proceso ARCH o GARCH. En general se plantea la siguiente hipótesis.

$$\begin{array}{ll}
 H_0 : \alpha_1 = \alpha_2 = \dots \alpha_r = 0 & H_0 : \text{No existe efecto ARCH} \\
 H_1 : \alpha_i \neq 0, i = 1, 2, \dots, r & H_1 : \text{Existe efecto ARCH}
 \end{array}$$

Según el cuadro 4.14 los resultados de la prueba de Multiplicadores de Lagrange permiten concluir que los residuos del modelo AR (6) son condicionalmente heterocedásticos ($p < 0.05$), hecho que también se puede observar en el correlograma de los residuos al cuadrado (Ver figura 4.17), el cuál presenta autocorrelaciones significativas ($p < 0.05$).

Cuadro 4.14: Efecto ARCH

F-statistic	29.92691	Prob. F(1,127)	0.0000
Obs*R-cuadrado	24.60108	Prob. Chi-Square(1)	0.0000

Fuente: Elaboración propia.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.436	0.436	25.319	0.000
		2	0.217	0.033	31.625	0.000
		3	0.047	-0.073	31.923	0.000
		4	0.116	0.142	33.753	0.000
		5	0.280	0.246	44.527	0.000
		6	0.529	0.403	83.212	0.000
		7	0.320	-0.070	97.512	0.000
		8	0.169	-0.031	101.51	0.000
		9	0.112	0.115	103.30	0.000
		10	0.050	-0.098	103.66	0.000
		11	0.073	-0.148	104.42	0.000
		12	0.085	-0.204	105.46	0.000

Figura 4.17: Autocorrelación de los residuos al cuadrado.

Fuente: Elaboración propia.

4.7.4 Identificación del modelo GARCH

Usualmente, la identificación de modelos ARCH y especialmente de los GARCH mediante las FAC y FACP estimadas de los residuos al cuadrado es difícil y requiere cierta experiencia. Por lo general, se proponen varios modelos de orden bajo tales como GARCH (1, 1), GARCH (1,2), GARCH (2, 1), GARCH (2,2) y luego escoger el modelo adecuado en base a algunos indicadores tales como los criterios de información AIC y BIC.

Siguiendo este procedimiento y luego de evaluar varios modelos se eligió finalmente el modelo GARCH(1,1) con base en los criterios de bondad de ajuste AIC y BIC.

Cuadro 4.15: Identificación del modelo GARCH

Modelo Ajustado	AIC	BIC
ARMA(6,0) - ARCH(1)	702.056	728.2699
ARMA(6,0) - ARCH(2)	702.6202	731.7467
ARMA(6,0) - ARCH(3)	704.4891	736.5283
ARMA(6,0) - ARCH(4)	699.3211	734.2729
ARMA(6,0) - ARCH(5)	699.8488	737.7133
ARMA(6,0) - ARCH(6)	701.077	741.8542
ARMA(6,0) - GARCH(1,1)	696.9138	726.0404
ARMA(6,0) - GARCH(2,1)	696.6664	728.7056
ARMA(6,0) - GARCH(1,2)	697.2262	729.2654
ARMA(6,0) - GARCH(2,2)	696.7902	731.7421

Fuente: Elaboración propia.

4.7.5 Estimación del modelo

En adelante, la estimación, validación y predicción estarán referidas al mejor modelo identificado, esto es, al modelo ARMA (6,0)-GARCH (1,1).

El modelo ARMA (6,0)-GARCH (1,1) fue identificado sobre los residuos al cuadrado de un modelo ARMA (6,0) por lo que no es posible utilizar directamente la metodología de estimación descrita por (Bollerslev, 1986). En este caso utilizaremos una variante propuesta por (Weiss, 1984).

Cuadro 4.16: Estimación del modelo ARMA (6,0) – GARCH (1,1)

Variable	Coefficient	Std. Error	Z-Statistic	P-value
C	5.277476	1.010207	5.224151	0.0000
AR(1)	1.064544	0.094266	11.29300	0.0000
AR(2)	-0.261860	0.155028	-1.689116	0.0912
AR(3)	0.052588	0.167290	0.314352	0.7533
AR(4)	-0.434987	0.142101	-3.061119	0.0022
AR(5)	0.457122	0.126646	3.609436	0.0003
AR(6)	-0.110817	0.082692	-1.340114	0.1802
Variance Equation				
C	0.259761	0.258498	1.004888	0.3150
RESID(-1)^2	0.211333	0.096196	2.196903	0.0280
GARCH(-1)	0.761259	0.083914	9.071901	0.0000

Fuente: Elaboración propia.

Todos los coeficientes estimados de la varianza son significativos ($p < 0.05$) a excepción de la constante y cumplen con las restricciones impuestas al modelo (Ver cuadro 4.16). Por otro lado, si comparamos los coeficientes estimados para la parte ARMA (Cuadro 4.12) con los obtenidos en el cuadro 4.16 notamos leves diferencias debido a que el proceso de re-estimación de parámetros incluye al modelo GARCH (1,1).

$$\left\{ \begin{array}{l} Z_t = 5.277476 + 1.064544Z_{t-1} - 0.261860Z_{t-2} + 0.052588Z_{t-3} \\ \quad - 0.434987Z_{t-4} + 0.457122Z_{t-5} - 0.110817Z_{t-6} + a_t \\ a_t = \sqrt{h_t} \varepsilon_t \\ h_t = 0.259761 + 0.211333a_{t-1}^2 + 0.761259h_{t-1} \end{array} \right.$$

4.7.6 Validación del modelo

Una vez estimado los parámetros procedemos a evaluar la bondad de ajuste y la adecuación del modelo. Este punto es similar al realizado con procesos ARIMA.

Un modelo estimado válido es el que cumple los siguientes requisitos:

- i)** Los residuos del ajuste se comportan aproximadamente como un proceso ruido blanco gaussiano.
- ii)** No se observa heterocedasticidad condicional en los residuos.
- iii)** El ajuste del modelo es estadísticamente significativo.
- iv)** De manera opcional usaremos el gráfico de la serie original versus la serie estimada, además del gráfico espectral de los residuos.

A continuación se comprobara que los residuos se comportan como ruido blanco Gaussiano, es decir se debe verificar que los residuos del modelo ARMA (6,0) – GARCH (1,1) poseen distribución normal y que las autocorrelaciones de los residuos se encuentren dentro de las bandas de confianza.

Prueba de normalidad de los residuos

H₀: Los residuos del modelo poseen distribución normal.

H₁: Los residuos del modelo no poseen distribución normal.

Realizando las pruebas paramétricas respectivas se concluye que a un nivel de significancia del 5%, no existe suficientes evidencias para rechazar la hipótesis nula (Ver figura 4.18), es decir, los residuos del modelo ARMA (6,0) – GARCH (1,1) poseen distribución normal.

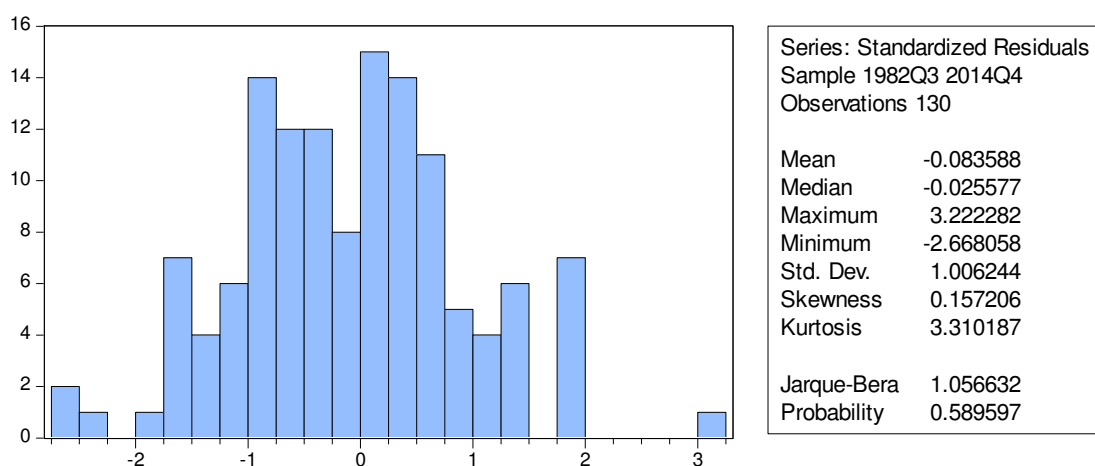


Figura 4.18: Prueba de normalidad e histograma de los residuos.

Fuente: Elaboración propia.

Según la figura 4.19 ninguna de las autocorrelaciones de los residuos (ε_t) del modelo ARMA (6,0)-GARCH (1,1) son significativas, éste hecho evidencia que la información relevante fue capturada por el modelo.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.063	0.063	0.5302	
		2	-0.019	-0.023	0.5784	
		3	0.009	0.012	0.5890	
		4	-0.023	-0.025	0.6591	
		5	-0.092	-0.089	1.8228	
		6	-0.042	-0.032	2.0641	
		7	0.027	0.029	2.1668	0.141
		8	-0.009	-0.013	2.1779	0.337
		9	-0.070	-0.072	2.8817	0.410
		10	0.118	0.119	4.8776	0.300
		11	0.076	0.054	5.7035	0.336
		12	0.009	0.009	5.7142	0.456

Figura 4.19: Correlograma de los residuos.

Fuente: Elaboración propia.

Prueba de heterocedasticidad condicional de los residuos

Una herramienta que se utiliza para probar que ya no existe heterocedasticidad condicional en la serie de residuos es el correlograma de los residuos al cuadrado del modelo ajustado (ε_t^2), el cual no debe presentar ninguna autocorrelación significativa.

En la figura 4.20 se observa que las autocorrelaciones de los residuos al cuadrado no son significativas, lo cual nos indica que no existe heterocedasticidad condicional en los residuos.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.019	0.019	0.0461	0.830
		2	-0.114	-0.114	1.7763	0.411
		3	-0.070	-0.066	2.4315	0.488
		4	0.027	0.016	2.5280	0.640
		5	0.066	0.051	3.1344	0.679
		6	0.124	0.125	5.2727	0.509
		7	0.076	0.093	6.0867	0.530
		8	-0.039	-0.005	6.2961	0.614
		9	0.035	0.069	6.4719	0.692
		10	0.030	0.026	6.5982	0.763
		11	0.019	0.008	6.6482	0.827
		12	-0.090	-0.108	7.8343	0.798

Figura 4.20: Correlograma de los residuos al cuadrado (ε_t^2).

Fuente: Elaboración propia.

Otra forma de evaluar la presencia de heterocedasticidad en los residuos es utilizando la prueba de multiplicadores de Lagrange la cual evidencia que no existe presencia de heterocedasticidad en los residuos del modelo ajustado.

H_0 : No existe efecto ARCH en los residuos

H_1 : Existe efecto ARCH en los residuos

Los resultados de la prueba de Multiplicadores de Lagrange permiten concluir que los residuos del modelo ARMA (6,0)-GARCH (1,1) no presentan heterocedasticidad condicional (Ver cuadro 4.17).

Cuadro 4.17: Efecto ARCH de los residuos del modelo

F-statistic	0.044180	Prob. F(1,127)	0.8339
Obs*R-cuadrado	0.044861	Prob. Chi-Square(1)	0.8323

Fuente: Elaboración propia.

Bondad de ajuste

La bondad de ajuste de un modelo estadístico describe lo bien que se ajusta un conjunto de observaciones. Por este motivo es necesario elegir el mejor modelo en términos de bondad de ajuste. En el cuadro 4.18 se muestran valores de los estadísticos AIC, BIC y el valor del Log-verosimilitud asociado con los posibles modelos ajustados a los datos.

Según el cuadro 4.18 se observa que el mejor modelo según los indicadores de bondad de ajuste es el ARMA (6,0) - GARCH (1,1).

Cuadro 4.18: Criterio de bondad de ajuste

Modelo Ajustado	AIC	BIC	Log-Verosimilitud
ARMA(6,0) - ARCH(1)	702.0560	728.2699	-2.5149
ARMA(6,0) - ARCH(2)	702.6202	731.7467	-2.5099
ARMA(6,0) - ARCH(3)	704.4891	736.5283	-2.5092
ARMA(6,0) - ARCH(4)	699.3211	734.2729	-2.4828
ARMA(6,0) - ARCH(5)	699.8488	737.7133	-2.4774
ARMA(6,0) - ARCH(6)	701.0770	741.8542	-2.4746
ARMA(6,0) - GARCH(1,1)	696.9138	726.0404	-2.4887
ARMA(6,0) - GARCH(2,1)	696.6664	728.7056	-2.4804
ARMA(6,0) - GARCH(1,2)	697.2262	729.2654	-2.4825
ARMA(6,0) - GARCH(2,2)	696.7902	731.7421	-2.4735

Fuente: Elaboración propia.

Una manera gráfica de validar el modelo ajustado es haciendo uso del espectro de los residuos, si los residuos se comportan como ruido blanco, se debe esperar que el espectro de los residuos sea una línea constante, es decir, el espectro no debe de presentar patrones de comportamiento.

De la figura 4.21 se concluye que el resultado es un espectro en el que prevalece la parte aleatoria, además podemos observar que los residuos no presentan información respecto a la tendencia y ciclos, sin embargo hay algo de la estacionalidad que no ha sido explicada por el modelo, lo cual no es tan significativo. En consecuencia el modelo es apropiado.

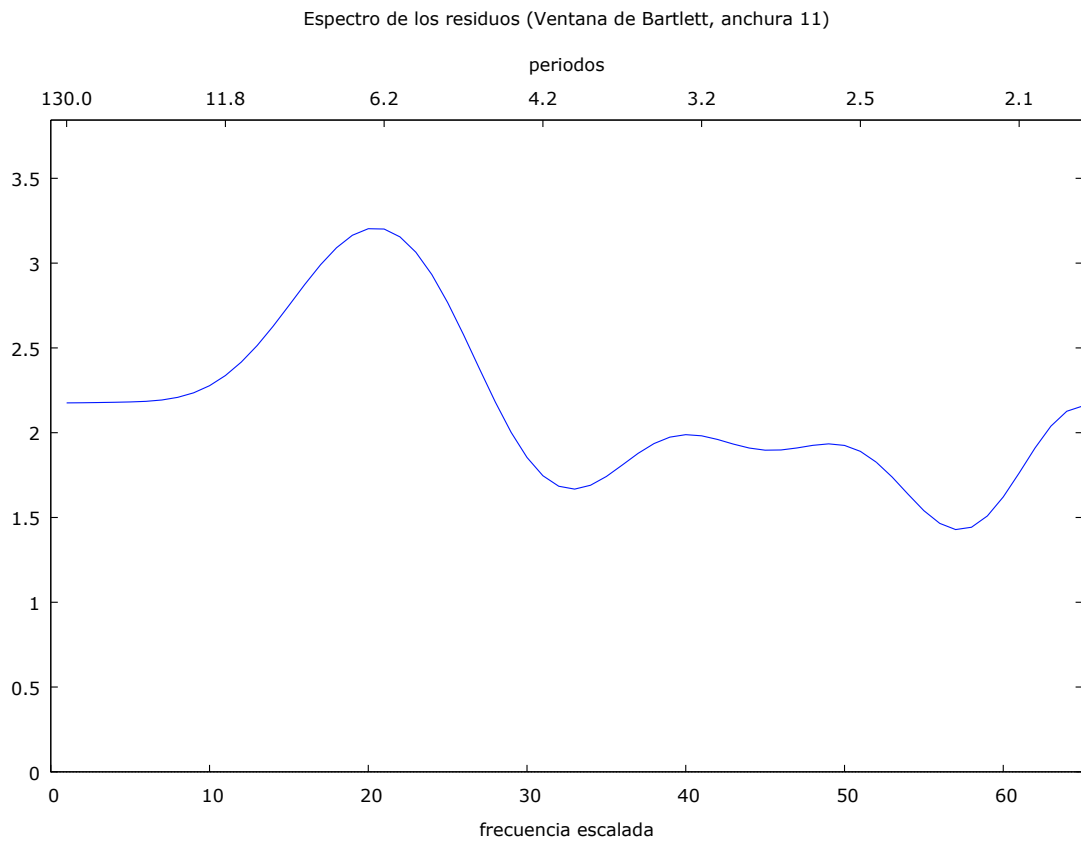


Figura 4.21: Espectro de los residuos del modelo ARMA (6,0)-GARCH (1,1).

Fuente: Elaboración propia.

Otra forma gráfica de validar el modelo propuesto es observando el grado de ajuste de la serie estimada versus la serie original, según la figura 4.22 se observa que las estimaciones de la serie bajo estudio poseen un buen ajuste, pues se comporta de manera similar a la serie original.

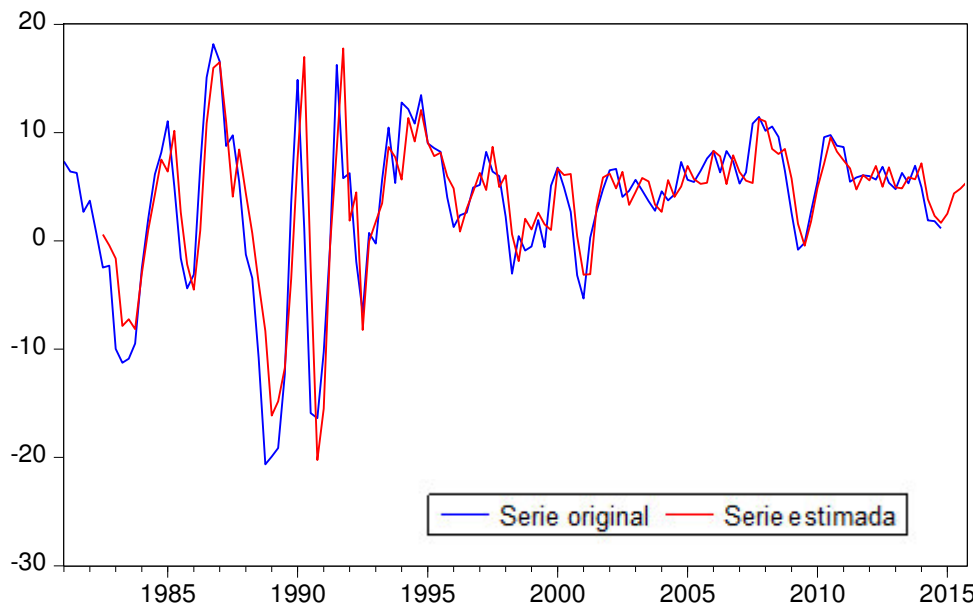


Figura 4.22: Serie original vs la serie estimada.

Fuente: Elaboración propia.

4.7.7 Pronósticos del modelo

La identificación, estimación y evaluación de los modelos de heterocedasticidad condicional ARCH, constituyen un proceso iterativo; cuyo objetivo final es la obtención de un modelo consistente con la estructura de los datos y que permitan obtener los pronósticos respectivos, por lo que antes de utilizar el modelo para los pronósticos es importante evaluar la capacidad predictiva del modelo utilizando datos crudos.

Luego de elegir el modelo ajustado ARMA (6,0)-GARCH (1,1) como el más adecuado estamos en condiciones de predecir la volatilidad futura de la variación porcentual del PBI trimestral en el Perú.

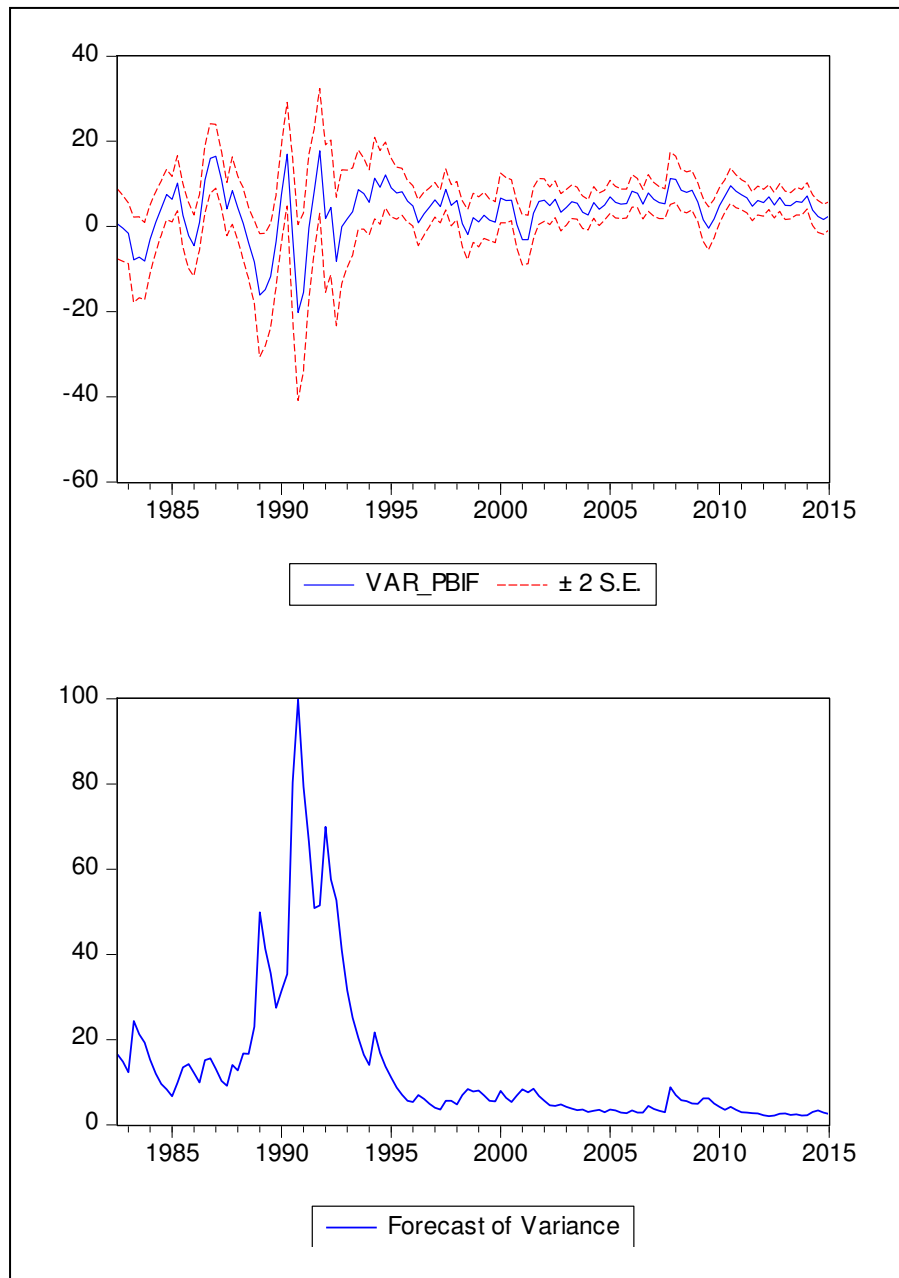


Figura 4.23: Predicción de la volatilidad (Predicción de la media y la varianza).

Fuente: Elaboración propia.

Se puede observar en la figura 4.23 dos gráficos una para la predicción de la media con intervalos de confianza al 95% y la otra para la predicción de la varianza, en el primero se puede observar que las bandas de confianza son adecuadas por lo que otorgarían buenos pronósticos, y el segundo muestra el ajuste de la varianza en el cual se observa picos de alta intensidad para los periodos de 1990 a 1995, lo cual nos hace suponer la presencia de un efecto de intervención.

A continuación se realizarán los pronósticos respectivos a fin de verificar el buen ajuste de nuestro modelo, para ello se compara el valor real y el valor pronosticado. La idea básica de esta comparación es que no existan diferencias significativas entre los dos valores de un mismo periodo.

Los pronósticos obtenidos mediante el modelo ARMA (6,0)-GARCH (1,1) proporcionan valores muy similares a los valores reales, aunque en comparación con los pronósticos del modelo SETAR (2,6,6), los modelos GARCH tienden a sobrestimar los pronósticos (Ver cuadro 4.19).

Cuadro 4.19: Pronósticos del modelo ARMA (6,0)-GARCH (1,1)

Periodo	Valor real	Valor pronosticado
2015-I	1.8467	2.5111
2015-II	3.1258	4.3758
2015-III	3.2359	4.8119
2015-IV	4.6792	5.4618

Fuente: Elaboración propia.

CAPÍTULO V
DISCUSIÓN DE LOS
RESULTADOS

CAPÍTULO IV. DISCUSIÓN DE LOS RESULTADOS

Esta investigación tuvo como propósito desarrollar el modelo SETAR bajo la metodología de Tsay, el cual consistió en el modelamiento (Identificación, estimación, validación y pronósticos) de la variación porcentual del PBI trimestral en el Perú, 1981-2014.

A fin de comparar los resultados obtenidos mediante los modelos SETAR, se procedió al desarrollo de los modelo GARCH, la elección de este modelo no lineal, se debió básicamente al uso masivo por parte de los investigadores del campo económico y financiero. Esto principalmente se debe a la capacidad para modelar series de alta volatilidad.

La comparación de los modelos, no solo se realizó a nivel de los supuestos o los pronósticos, sino que también se analizaron, otros aspectos como el análisis de las componentes de la serie en estudio, los cambios de régimen y la practicidad al momento de modelar los datos.

Modelo SETAR

El modelo SETAR (2,6,6) cumple con todos los supuestos de los residuos, tales como la prueba de normalidad, incorrelación de los residuos y algunas pruebas gráficas como el correlograma y el espectro de los residuos (Ver capítulo IV, validación de los supuestos).

En lo que respecta a los pronósticos, se obtuvieron resultados muy similares comparados con los valores reales (Ver cuadro 5.1). Por otro lado, las ganancias en cuanto a la interpretación de nuestra serie en estudio son muy interesantes ya que el modelamiento incorpora los cambios de estado o régimen que se producen en la serie y por otro lado son fáciles de contextualizar por cuanto guardan relación con eventos reales.

Modelo GARCH

El modelo ARMA (6,0) - GARCH (1,1) cumple con los supuestos habituales de los residuos, otra prueba potente de validación, fue la prueba LM (Multiplicador de Lagrange) de Engle, el cual indicó que los residuos del modelo no presentan heterocedasticidad condicional (Ver cuadro 4.16). También se realizó las pruebas gráficas como el correlograma y espectro de los residuos.

En lo que respecta a los pronósticos, se obtuvieron resultados muy similares comparados con los valores reales (Ver cuadro 5.1), sin embargo se pierde demasiado en cuanto a las interpretaciones, sobre todo en las componentes de la serie bajo estudio. Esto es debido a que los modelos GARCH fueron desarrollados con el propósito de modelar la volatilidad de las series financieras, sin considerar las componentes de la series, como son los ciclos, estacionalidad y tendencia.

Cuadro 5.1: Comparación de pronósticos

Período	Valor real	Pronóstico SETAR	Pronóstico GARCH
2015-I	1.8467	2.2176	2.5111
2015-II	3.1258	4.2386	4.3758
2015-III	3.2359	3.5606	4.8112
2015-IV	4.6792	2.8673	5.4618

Fuente: Elaboración propia.

Comparación de los indicadores de bondad de ajuste

Para determinar si un modelo o método de pronóstico es superior a otro existen diversos criterios de evaluación, como por ejemplo los criterios de información AIC, BIC, SSE (Suma de Cuadrado de los Residuos) y el MAD (Desviación Media Absoluta).

Como se puede observar en el cuadro 5.2, el modelo SETAR presenta menor criterio de información de Akaike (AIC) y el Bayesiano (BIC) comparándolo con el modelo GARCH, también nos otorga menor Suma de Cuadrados de Residuos (SSE) y menor Desviación media absoluta (MAD).

Para fines comparativos también se consideró el modelo AR (6), el cual proporcione indicadores deficientes de bondad de ajuste que los demás modelos, lo cual es un indicio para suponer un pobre modelamiento de los datos.

Cuadro 5.2: Indicadores de bondad de ajuste

Indicador	AR(6)	SETAR	GARCH
AIC	749.5480	680.7176	696.9138
BIC	772.8492	720.8631	726.0404
SSE	1736.1300	1146.4380	1780.9039
MAD	2.4344	2.2551	2.5460

Fuente: Elaboración propia.

CAPÍTULO VI

CONCLUSIONES

CAPÍTULO VI. CONCLUSIONES

- ❖ El modelo SETAR de dos regímenes no solo ofrece ganancias importantes respecto a los pronósticos, sino también muestra ventajas frente a los modelos GARCH en términos de facilidad de interpretación.
- ❖ En lo que respecta a la identificación, estimación, validación y pronósticos usando los modelos SETAR, se presentan muchas ganancias debido a lo práctico que es modelamiento de los datos, pues solo se basa en procesos autorregresivos, los cuales son conceptos manejables por la mayoría de los investigadores.
- ❖ Se amplió la metodología para la determinación del valor umbral óptimo (r) mediante la incorporación de los criterios de información AIC y BIC lo cual complementa la propuesta gráfica de Tsay.
- ❖ Se verificó el valor umbral óptimo usando la prueba de Hansen, llegando a resultados equivalentes con los obtenidos mediante los criterios de información (AIC y BIC).
- ❖ Se desarrolló e implementó la modelización de los modelos SETAR mediante el uso de los programas S-PLUS 8.0 y el R-PROJECT 3.3.4, específicamente para la identificación, estimación, validación y los pronósticos del modelo. Por otro lado se usó el programa EVIEWS 8.0 para la identificación, estimación, validación y los pronósticos del modelo GARCH.

- ❖ Se comprobó que el modelo SETAR es más eficiente en el modelamiento que los modelos GARCH, obteniendo menores valores en los criterios de información AIC y BIC. Además los residuos del modelo se comportan como ruido blanco.
- ❖ En cuanto a los pronósticos se observó que tanto el modelo SETAR (pronósticos usando el método de Montecarlo) y el modelo GARCH, generan pronósticos similares al verdadero valor.
- ❖ En el presente trabajo de investigación se ajustó un modelo SETAR de dos regímenes para la variación porcentual del PBI trimestral, con base en la metodología propuesta por (Tsay, 1989) y (Tong, 1990). Los principales resultados muestran evidencia empírica de no linealidad de umbral en la serie asociada.
- ❖ Se comprueba que los modelos SETAR son útiles para capturar periodicidades límite. Para nuestro caso (“Variación porcentual del PBI trimestral en el Perú, 1981-2014”) se observó en el periodograma una frecuencia máxima en el punto 13, cuyo periodo es 10 trimestres, es decir, la presencia de ciclos es cada 40 meses (3 años y medio).

CAPÍTULO VII

RECOMENDACIONES

CAPÍTULO VII. RECOMENDACIONES

- ❖ Los programas comerciales no tienen incorporado la modelización SETAR, por lo cual se recomienda investigar y desarrollar aplicaciones académicas con los programas S-PLUS y R-PROJECT a fin de familiarizar a la comunidad investigadora y a los alumnos con estos modelos.
- ❖ Incorporar el aporte de este trabajo, en cuanto a la metodología ampliada para la determinación de umbrales óptimos, en las aplicaciones que se desarrollen como parte de la rutina de modelización.
- ❖ Finalmente, este trabajo constituye un aporte académico, principalmente en la metodología de series de tiempo, la cual es una de las ramas más dinámicas de la estadística, pues este modelo nos ayudan a obtener una mejor estimación y unos buenos pronósticos. Por esta razón, es importante que se continúe una agenda de investigación que complemente estos resultados mediante otros métodos o modelos, como son los modelos SETAR de cambio régimen Markoviano o los modelos SETAR bajo en enfoque bayesiano.

CAPÍTULO VIII
REFERENCIAS
BIBLIOGRÁFICAS

CAPÍTULO VIII. REFERENCIAS BIBLIOGRÁFICAS

- Agüero Palacios, Y., & Elescano Rojas, A. (2004). MODELOS ARCH: UNA APLICACIÓN EN EL PRONÓSTICO DE LA VOLATILIDAD DE ACCIONES COTIZADAS EN LA BOLSA DE VALORES DE LIMA. *PESQUIMAT*, 64-79.
- Arias, F. G. (2006). *Técnicas e instrumentos de recolección de datos*. Caracas-Venezuela: Episteme.
- Asenjo Godoy, P., & Praetorius Batalla, S. (Viernes de Junio de 2016). *OPTIMIZACIÓN DE MODELOS GARCH A TRAVÉS DE ALGORITMO GENÉTICO*. Obtenido de http://repositorio.uchile.cl/tesis/uchile/2006/ec-asenjo_g/pdfAmont/ec-asenjo_g.pdf
- Bazan Palomino, W. N. (2009). *MODELOS DE CAMBIO DE RÉGIMEN: RELACIÓN NO-LINEAL ENTRE EL CRÉDITO BANCARIO AL SECTOR PRIVADO Y EL PRODUCTO INTERNO BRUTO PARA PERÚ 1994 – 2008*. Lima: USMP-FACULTAD DE CIENCIAS CONTABLES, ECONÓMICAS Y FINANCIERAS.
- Bollerslev, T. (1986). GENERALIZED AUTOREGRESSIVE CONDITIONAL. *Elsevier Science Publishers B.V.*, 307-327.
- Chen, R., & Tsay, R. S. (1993). Functional-coefficient autoregressive models. *Journal of the American Statistical Association*, 298-308.
- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *The Econometric Society*, 987-1007.
- Franses, P. H., & Dijk, D. v. (2000). *Non-Linear Time Series Models in Empirical Finance*. Cambridge: Cambridge University Press.
- George, B., Gwilym, J., & Reinsel, G. (2008). *Times series analysis: Forecasting and control*. New Jersey: Wiley.
- Gerencia Central de Estudios Económicos. (s.f.). *Banco Central de Reserva del Perú*. Recuperado el 23 de Noviembre de 2015, de <http://www.bcrp.gob.pe/estadisticas/cuadros-trimestrales-historicos.html>
- Granger, C. W., & Anderson, A. P. (1978). *An Introduction to Bilinear Models*. Gottingen: Van derhoeck & Ruprecht.
- Hansen, B. E. (1997). Inference in TAR models. *The MIT Press*, 1-14.
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, M. d. (2010). *Metodología de la investigación*. México: McGraw-Hill.
- Hoyos, M., Ramos, J., & Vivas, L. (2009). UN MODELO SETAR PARA EL PBI COLOMBIANO. *Cuadernos de economía*, 63-78.
- Keenan, D. M. (1985). A Tukey Nonadditivity-Type Test for Time Series Nonlinearity. *Biometrika*, 39-44.
- Martinez, W. L., & Martinez, A. R. (2002). *Computational Statistics Handbook with MATLAB*. New York: CHAPMAN & HALL/CRC.

- Medina Merino, F. (2011). *Análisis de series cronológicas, aplicadas a series económicas*. Lima: Centro de investigación y desarrollo (CIDE-INEI).
- Monsalve, A., & Harmath, P. (2015). *INTRODUCCIÓN AL ANÁLISIS DE SERIES DE TIEMPO CON APLICACIONES A LA ECONOMETRÍA Y FINANZAS*. Mérida-Venezuela: Instituto Venezolano de Investigaciones Científicas.
- Moreno Lopez, E. C. (2010). *Una aplicación del modelo TAR en series de tiempo financieras*. Bogotá: Universidad Nacional de Colombia-Facultad de Ciencias-Departamento de Estadística.
- Morettin, P. A., & Maria de Castro Toloi, C. (2006). *Modelos para previsão de séries temporais*. Rio de Janeiro: Blucher.
- Nieto, F. H. (2005). Modeling bivariate threshold autoregressive processes in the presence of missing data. *Communications in Statistics, Theory and Methods*, 905-930.
- Peña, D. (2005). *Análisis de series temporales*. Madrid: Alianza editorial.
- Peréz, F., & Velazqués, H. (2004). Análisis de cambio de régimen en series de tiempo no lineales utilizando modelos TAR. *Lecturas de economía*, 101-119.
- Romero Madrid, V. (2013). *MODELACIÓN DEL COMPORTAMIENTO DE LA TASA DE INTERÉS ACTIVA PROMEDIO EN MONEDA NACIONAL DEL SISTEMA BANCARIO PERUANO*. Piura: UNIVERSIDAD DE PIURA-FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES.
- Russell Noriega, M. G. (2006). *Pronósticos en modelos autorregresivos con umbral*. Guanajuato: Centro de investigación en Matemáticas, A.C.
- Teräsvirta, T., Granger, C. W., & Tjøstheim, D. (2010). *Modelling Nonlinear Economic Time Series*. Oxford: OUP Oxford.
- Tiao, G. C., & Tsay, R. S. (1994). Some advances in non linear and adaptative modelling in time series. *Journal of Forecasting*, 109-140.
- Tong, H. (1978). *On a Thresold Model*. In *Pattern Recognition and Signal Processing*. Manchester: Sijthoff and Noordhof.
- Tong, H. (1983). *Threshold Models in Non-Linear Time-Series Analysis*. New York: Lecture Notes in Statistics.
- Tong, H. (1990). *Non-linear Time Series: A Dynamical System Approach*. Oxford: Clarendon Press.
- Tong, H., & Lim, K. S. (1980). Threshold Autoregression, Limit Cycles and Cyclical Data. *Journal of the Royal Statistical Society*, 245-292.
- Tsay, R. S. (1989). Testing and Modeling Threshold Autoregressive Processes. *Journal of the American Statistical Association*, 231-240.
- Universidad Nacional Abierta y a Distancia (UNAD). (18 de Diciembre de 2015). *Lección 5: Investigación pura, investigación Aplicada, Investigación profesional*. Obtenido de http://dateca.unad.edu.co/contenidos/100104/100104_EXE/leccin_5_investigacin_pura_investigacin_aplicada_investigacin_profesional.html

- Weiss, A. A. (1984). ARMA Models with ARCH Errors. *Journal of Times Series Analysis*, 303-310.
- Wold, H. (1939). A Study in the Analysis of Stationary Time Series by Herman Wold. *Journal of the Institute of Actuaries*, 113-115.
- Zivot, E., & Wang, J. (2006). *Modelling Financial Time Series with S-PLUS*. Washington: Springer.

CAPÍTULO IX

ANEXOS

CAPÍTULO IX. ANEXOS

ANEXO 1: DESARROLLO DE ALGUNOS CONCEPTOS TRATADOS EN LA INVESTIGACIÓN

Análisis espectral

Los primeros estudios de las series temporales se realizaron en el dominio del tiempo y se utilizaron diversos métodos ingeniosos de difícil justificación científica. Arthur Schuster introdujo un método denominado **periodograma**, el cual estudia las series temporales en el dominio de la frecuencia.

La posibilidad de estudiar el análisis de una serie temporal, en el dominio del tiempo o en el de la frecuencia, se justifica científicamente por la equivalencia entre la función de covarianza (definida en el dominio del tiempo) y el espectro o función de densidad espectral (definida en el dominio de la frecuencia), y esta relación nos la proporciona las transformadas de Fourier por ser una de ellas la transformada de la otra función.

Empecemos por definir el espectro como la transformada de Fourier de la función de autocovarianza de la serie Z_t . Recordemos que el espectro “reparte” la varianza entre varias bandas de frecuencias haciéndonos ver cuáles son las más importantes.

$$f(w) = \frac{1}{2\pi} \left(\gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j \cos(jw) \right) ; w \in (0, \pi)$$

La representación gráfica del espectro estimado nos proporciona un primer conocimiento de la estructura del modelo y nos sugiere modelos más apropiados: si existieran realmente periodos en la serie temporal, en las frecuencias correspondientes a dichos periodos ($\text{periodo}=1/\text{frecuencia}$) presentaría la gráfica picos muy pronunciados con grandes concentraciones de masas espectrales.

También sirve como herramienta para analizar las consecuencias de cosas que se hacen en el dominio del tiempo como la aplicación de filtros (se puede averiguar lo que hacen las diferencias finitas, los promedios móviles y otros). En la literatura especializada se encuentran otros casos en que ayuda a trabajar en el dominio del tiempo en cuanto a demostraciones o a descubrir detalles o problemas. A continuación se muestran algunos ejemplos del análisis espectral.

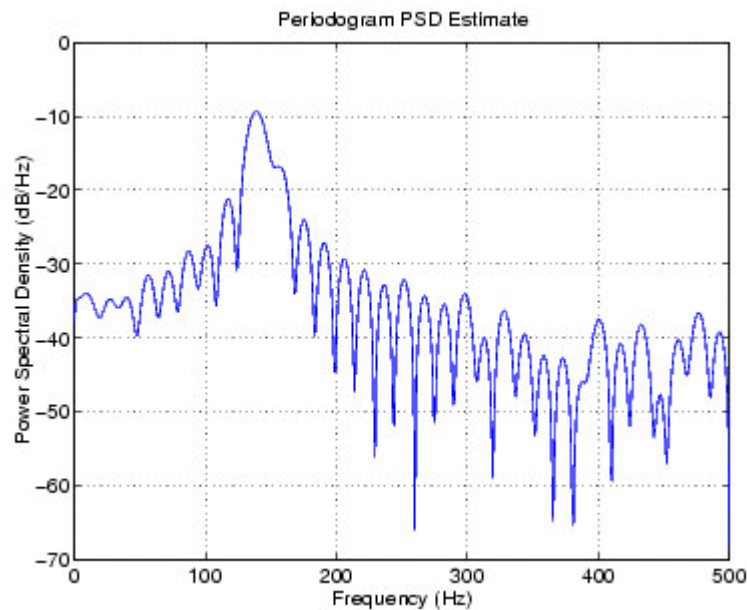


Figura 9.1: Densidad espectral (Ejemplo).

Fuente: Elaboración propia.

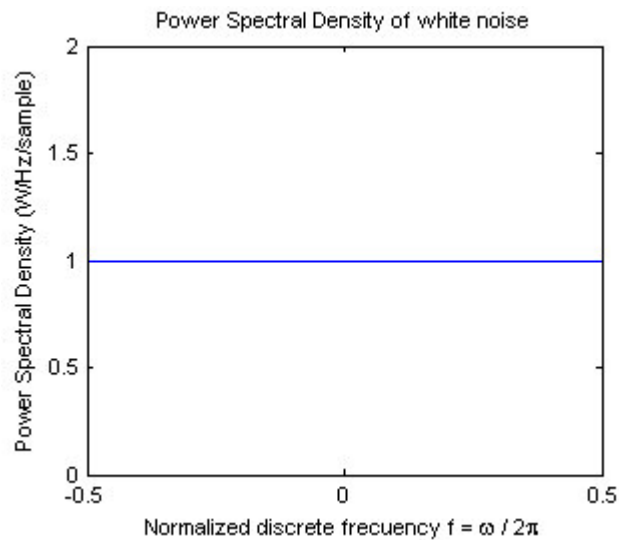


Figura 9.2: Densidad espectral del ruido blanco.

Fuente: Elaboración propia.

Prueba LM de Engle

El objetivo de esta prueba es ver si es que una serie de tiempo posee heterocedasticidad condicional, en otras palabras, analiza si es que tiene sentido aplicar un modelo del tipo ARCH/GARCH sobre la serie en estudio. Esta prueba fue creada por Engle en el año 1982 y se basa en el principio del Multiplicador de Lagrange (LM).

El método de Multiplicador de Lagrange también puede ser utilizado para realizar un contraste de especificación basado sobre la homocedasticidad, pero éste sólo puede ser aplicado en modelos ARCH (q) y no en modelos GARCH, ya que estos últimos presentarían problemas de multicolinealidad (Asenjo Godoy & Praetorius Batalla, 2016).

Prueba de Jarque-Bera (1987)

En estadística, la prueba de Jarque-Bera es una prueba de bondad de ajuste para comprobar si una muestra de datos tiene la asimetría y la curtosis de una distribución normal. La prueba recibe el nombre de Carlos Jarque y Anil K. Bera.

La prueba de Jarque-Bera se define como:

$$JB = \frac{n-k+1}{6} \left(S^2 + \frac{1}{4} (C-3)^2 \right)$$

donde n es el número de observaciones (o grados de libertad en general); S es la asimetría de la muestra, C la curtosis de la muestra y k el número de regresores:

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^3}{\left(\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right)^{3/2}} \quad C = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^4}{\left(\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right)^2}$$

Donde $\hat{\mu}_3$ y $\hat{\mu}_4$ son las estimaciones de los momentos centrales tercer y cuarto, respectivamente, \bar{X} es la media de la muestra y σ^2 es la estimación del segundo momento central, es decir, la varianza.

El estadístico de Jarque-Bera se distribuye asintóticamente como una distribución chi-cuadrado con dos grados de libertad y puede usarse para probar la hipótesis nula de que los datos pertenecen a una distribución normal.

La hipótesis nula es una hipótesis conjunta de que la asimetría y la curtosis son nulas.

$$H_0 : \alpha_3 = 0 \wedge \alpha_4 - 3 = 0$$

$$H_1 : \alpha_3 \neq 0 \wedge \alpha_4 - 3 \neq 0$$

Criterios de información de Akaike (AIC)

El criterio de información de Akaike (AIC) es una medida de la calidad relativa de un modelo estadístico, para un conjunto dado de datos. Como tal, el AIC proporciona un medio para la selección del modelo.

Definición:

En el caso general, el criterio AIC se define como:

$$AIC = 2k - 2\ln(L)$$

donde k es el número de parámetros en el modelo estadístico, y L es el máximo valor de la función de verosimilitud para el modelo estimado.

Dado un conjunto de modelos candidatos para los datos, el modelo preferido es el que tiene el valor mínimo en el AIC. Por lo tanto AIC no sólo recompensa la bondad de ajuste, sino también incluye una penalidad, que es una función creciente del número de parámetros estimados. Esta penalización desalienta el sobreajuste (aumentando el número de parámetros libres en el modelo mejora la bondad del ajuste, sin importar el número de parámetros libres en el proceso de generación de datos).

Criterios de información de Bayesiano (BIC)

En estadística, el criterio de información bayesiano (BIC) es un criterio para la selección de modelos entre un conjunto finito de modelos. Se basa, en parte, de la función de probabilidad y que está estrechamente relacionado con el Criterio de Información de Akaike (AIC).

Cuando el ajuste de modelos, es posible aumentar la probabilidad mediante la adición de parámetros, pero si lo hace puede resultar en sobreajuste.

Tanto el BIC y AIC resuelven este problema mediante la introducción de un término de penalización para el número de parámetros en el modelo, el término de penalización es mayor en el BIC que en el AIC.

Definición:

El BIC es una consecuencia derivada asintótica bajo los supuestos de que la distribución de los datos se encuentra en la familia exponencial. A continuación se muestran las variables que definen el criterio Bayesiano.

- x : Son los datos observados.
- n : El número de datos u observaciones X o equivalentemente, el tamaño de la muestra.
- k : El número de parámetros libres a ser estimados. Si el modelo está bajo el supuesto de que es lineal, k es el número de regresores, incluyendo el intercepto.
- $P(x|M)$: La probabilidad marginal de los datos observados dado el modelo M .
- \widehat{L} : El máximo valor de la función de verosimilitud del modelo M .

La fórmula para el BIC es:

$$-2 * \ln p(x | M) \approx BIC = -2 * \ln(L) + k * \ln(n)$$

Bajo la suposición de que los errores de modelo o perturbaciones son independientes e idénticamente distribuidos según una distribución normal y que la condición límite de que la derivada de la probabilidad de registro con respecto a la varianza real es cero, esto se convierte en:

$$BIC = n * \ln(\widehat{\sigma}_e^2) + k * \ln(n); \text{ donde } \widehat{\sigma}_e^2 \text{ es la varianza del error}$$

Método de Montecarlo

La simulación de Montecarlo es una técnica que combina conceptos estadísticos (muestreo aleatorio) con la capacidad que tienen los ordenadores para generar números pseudo-aleatorios y automatizar cálculos (Martinez & Martinez, 2002).

La simulación de Montecarlo también fue creada para resolver integrales que no se pueden resolver por métodos analíticos, para solucionar estas integrales se usaron números aleatorios. Posteriormente se utilizó para cualquier esquema que emplee números aleatorios, usando variables aleatorias con distribuciones de probabilidad conocidas, el cual es usado para resolver ciertos problemas estocásticos y determinísticos, donde el tiempo no juega un papel importante.

Algoritmo:

1. Determinar el pseudo-población o modelo que representa la verdadera población de interés.
2. Utilice un procedimiento de muestreo para tomar muestras de la pseudo-población.
3. Calcular un valor para el estadístico de interés y almacenarla.
4. Repita los pasos 2 y 3 para los ensayos M .
5. Utilice los valores de M que se encuentran en el paso 4 para estudiar la distribución de la estadística.

ANEXO 2: SISTEMA DE CUENTAS NACIONALES TRIMESTRALES

Definición

Las cuentas trimestrales constituyen un sub conjunto totalmente coherente de las cuentas anuales, constituyendo una síntesis de la información coyuntural. Las Cuentas Nacionales Trimestrales se originaron principalmente, debido a la necesidad de seguir analizando en el corto plazo los movimientos de los fenómenos económicos.

Finalidad

El sistema de cuentas nacionales trimestrales, tiene por finalidad conocer en el corto plazo, los cambios macroeconómicos de la economía.

El uso de las cuentas nacionales trimestrales, tiene cada vez más importancia como un instrumento en la toma de decisiones en materia de política económica, sobre todo en la gestión de la política monetaria y fiscal.

Objetivos

Entre los principales objetivos de las cuentas trimestrales podemos mencionar:

- Presentar la situación económica general del país y de una manera resumida la estructura de la economía en su conjunto. Ofrecen además, cifras sintéticas a partir de las cuales es posible determinar el comportamiento general de la economía.
- Explicar detalladamente la evolución de las actividades económicas.

- Establecer los movimientos coyunturales de la economía, reflejando sus puntos de inflexión. Las cuentas anuales calculan la evolución promedio que se sintetiza en una tasa anual de crecimiento, pero ésta oculta movimientos coyunturales importantes. En períodos de recesión o auge, los cambios de tendencia se producen en el corto plazo y es necesario detectarlos con rapidez.
- Sintetizar la información coyuntural; reuniendo y haciendo coherentes los indicadores parciales de las diferentes actividades económicas. Quienes utilizan directamente datos de sectores particulares en el análisis de coyuntura, deben hacer el esfuerzo de reunir la información, confrontarla con otros indicadores, llegando en algunos casos, a resultados contradictorios. Las estimaciones trimestrales cumplen esta labor.
- Servir de guía para el desarrollo del Sistema Estadístico de corto plazo. La implementación del programa de cuentas trimestrales debe llevar necesariamente a desarrollar nuevas investigaciones y encuestas, revisando y rediseñando las existentes.

Enfoques para el cálculo del PBI

- **Enfoque de la Producción**

Se basa en el cálculo de la producción y el consumo intermedio de las diversas industrias de la economía. El valor agregado bruto de una industria se define como la diferencia entre el Valor Bruto de la Producción (Precios básicos) y el valor del Consumo Intermedio (Precios comprador).

- **Enfoque del Gasto**

Se basa en el cálculo de los componentes de la demanda final del PBI:

- = Gasto de consumo final (hogares, instituciones sin fines de lucro que sirven a los hogares (IPSFL), a precios de comprador
- + Gasto de consumo final del gobierno
- + Formación Bruta de Capital Fijo a precios de comprador
- + Variación de existencias a precios de comprador
- + Exportaciones (F.O.B.)
- Importaciones (C.I.F.)

- **Enfoque del Ingreso**

Se realiza a partir de cálculos separados de los componentes del valor agregado de industrias ramas o sectores.

- = Remuneraciones de los empleados
- + Excedente de explotación bruto/ingreso mixto
- + Impuestos a la producción e importación
- Subvenciones

Características de las cuentas trimestrales

- Las cuentas trimestrales tienen como punto de referencia las cuentas nacionales anuales, utilizan los mismos conceptos, nomenclaturas y definiciones.

- En términos generales el contenido de las cuentas trimestrales es más limitado que el anual. La información disponible en el corto plazo no permite implementar el modelo total. No se tienen balances y estados de pérdidas y ganancias para la mayor parte de las sociedades; las encuestas coyunturales no investigan todas las variables.
- El nivel de cálculo de las Cuentas Nacionales Trimestrales en el Perú, se inicia con el cálculo del PBI por el lado de la producción y el gasto, a valores corrientes y constantes, así mismo con la implementación de encuestas especializadas para obtener información de las variables del ingreso, se complementa el cálculo del PBI por los tres métodos.

ANEXO 3: METODOLOGÍA DE MEDICIÓN DEL PBI TRIMESTRAL

La elaboración del PBI trimestral responde a la necesidad de tener una visión de la economía del país actualizada y oportuna, que permita conocer el estado de la economía agregada en cada trimestre, usando toda la información cualitativa y cuantitativa disponible en el corto plazo.

Definición del PBI

El Producto Bruto Interno es el valor de todos los bienes y servicios finales producidos por un país en un determinado período. Comprende el valor de los bienes producidos, como viviendas, comercio, servicios, Gobierno, transporte, etc. Cada uno de estos bienes y servicios se valora a su precio de mercado y los valores se suman para obtener el PBI.

Calculo del PBI trimestral por el método del gasto

El cálculo del PBI por este método se realiza desde el punto de vista del Gasto o destino de la producción. Se mide el valor de las diferentes utilizaciones finales de la producción en la Economía, deducido el valor de las importaciones de los bienes y servicios (producción no generada en el territorio económico).

Los diferentes usos finales a los cuales se hace referencia son: consumo final privado, el consumo final del gobierno general, la formación bruta de capital fijo, la variación de existencias y las exportaciones. Cada uno de estos elementos se estima a precio de comprador.

$$\text{PBI} = \text{GCFP} + \text{GCFG} + \text{FBKF} + \text{VE} + \text{X} - \text{M}$$

Elementos que intervienen en el cálculo del PBI

- **Gasto de consumo final privado (GCFP)**

El Gasto de Consumo Final Privado (GCFP) es el valor de los bienes y servicios adquiridos por los hogares para la satisfacción directa de sus necesidades ya sean éstas individuales o colectivas. Incluye además, el valor de los gastos realizados por las instituciones sin fines de lucro que sirven a los hogares en bienes o servicios de consumo individual suministrados a los hogares en forma de transferencias sociales en especie.

- **Gasto de consumo final del gobierno (GCFG)**

Es el gasto realizado por el Gobierno en bienes no duraderos y servicios de uso intermedio, proporcionados a la comunidad en forma de servicios individuales o colectivos. Asimismo, considera las remuneraciones a los trabajadores de la administración pública y servicios de consumo individual y servicios de consumo colectivo.

El GCFG guarda relación con el Valor Bruto de Producción de la actividad del Gobierno, una vez deducidos sus ingresos por la venta de bienes y servicios, al cual se le incluye los bienes y servicios comprados por el gobierno para su transferencia final, sin transformación a los hogares.

- **Formación bruta de capital fijo (FBCF)**

La Formación Bruta de Capital Fijo (FBCF) agrupa todos los bienes y servicios adquiridos por los productores y cuya duración excede el año, interviene en el proceso productivo e incrementa el capital fijo de la empresa.

- **Variación de existencias (VE)**

La Variación de Existencias (VE) registra los cambios en el valor de los stocks, es decir, es la diferencia entre el volumen de mercancías existente entre principio y fin de cada periodo, valuada a precios vigentes durante el lapso de referencia. Se incluyen los bienes producidos, los que se encuentran en proceso de elaboración y los bienes terminados que estén en poder de los productores.

- **Exportaciones de bienes y servicios (E)**

Las exportaciones de bienes y servicios consisten en ventas, trueques, regalos o donaciones, de bienes y servicios de los residentes a los no residentes.

- **Importaciones de bienes y servicios (M)**

Las importaciones de bienes y servicios consisten en compras, trueques, regalos o donaciones, de bienes y servicios que realizan los residentes de un determinado país con el resto del mundo.

ANEXO 4: PROGRAMAS USADOS PARA EL MODELAMIENTO DE LOS DATOS

R-PROJECT 3.3.4

R es un entorno y lenguaje de programación con un enfoque al análisis estadístico.

R es una implementación de software libre del lenguaje S pero con soporte de alcance estático. Se trata de uno de los lenguajes más utilizados en investigación por la comunidad estadística, siendo además muy popular en el campo de la minería de datos, la investigación biomédica, la bioinformática y las matemáticas financieras. A esto contribuye la posibilidad de cargar diferentes bibliotecas o paquetes con funcionalidades de cálculo o graficación.

Librerías usadas de R-Project

- **tsDyn:**

Cuyos autores son Antonio Fabio Di Narzo y José Luis Aznarte. Esta librería agrupa una selección de modelos no lineales para series temporales como los modelos SETAR, STAR y RNA (Red Neuronal Artificial).

Para la modelización estadística se usó la librería tsDyn, específicamente para para la parte descriptiva así como para realizar las estimaciones, los pronósticos y comparaciones entre modelos.

- **TSA:** Es una librería del R cuyo objetivo es el modelado de series de tiempo no lineales (modelos TAR, prueba de no linealidad de Tsay) Contiene funciones R y conjuntos de datos que se detallan en el libro "Time Series Analysis with Applications in R (second edition)" por **Jonathan Cryer y Kung-Sik Chan**, los cuales son los creadores de la librería.

S-PLUS 8.0

Es una implementación comercial del lenguaje de programación S vendido por TIBCO Software Inc. Cuenta con capacidades de programación orientada a objetos y algoritmos avanzados de análisis.

S - PLUS es la mejor solución para el análisis exploratorio de datos y modelización estadística. Con más de 4 mil 200 funciones de análisis de datos, incluyendo el conjunto más amplio de métodos robustos y modernos disponibles en cualquier lugar.

Librería usada del S-PLUS

- **S+FinMetrics:** Es un módulo S - PLUS concebido por Andrew Bruce, Doug Martin, Jiahui Wang, Eric Zivot y desarrollado por [Insightful Corporation](#), que ofrece software avanzado de análisis y visualización de datos del mercado financiero. El software ofrece la herramienta más completa, moderna, flexible y analítica disponible para la modelización econométrica precisa, predictiva de series de tiempo financieras. La versión 2.0 del módulo (que opera bajo S - PLUS 8.0) incluye las siguientes funcionalidades:
 - Modelos GARH univariados.
 - Modelos de cointegración
 - Modelos GARCH multivariados.
 - Modelos de espacio estado.
 - Modelos REGARIMA.
 - Modelos VAR.
 - Modelos TAR, SETAR, LSTAR.
 - Etc.

EViews 8.0

EViews es un paquete estadístico para Microsoft Windows, usado principalmente para análisis econométrico. Ha sido desarrollado por Quantitative Micro Software (QMS). Combina la tecnología de hoja de cálculo con tareas tradicionales encontradas en software estadístico tradicional, empleando una interfaz de usuario gráfica. Estas características se combinan con un lenguaje de programación propio.

EViews puede ser empleado para análisis estadístico general, pero es especialmente útil para realizar análisis econométrico, como modelos de corte transversal, datos en panel y estimación y predicción con modelos de series de tiempo. Entre los tipos de archivo con los que es compatible destacan el Excel, SPSS, SAS, Stata, RATS, y TSP.

ANEXO 5: PRUEBA DE DICKEY- FULLER (DFA)

En estadística, la prueba de Dickey-Fuller confirma si una raíz unitaria está presente en un modelo autorregresivo. Lleva el nombre de los estadísticos David Dickey y Wayne Fuller, quienes desarrollaron la prueba en 1979.

Con el fin de determinar las propiedades de estacionariedad de las series se pueden utilizar distintos procedimientos: Prueba de Dickey Fuller (DF), el prueba de Dickey Fuller Aumentado (ADF), prueba de Phillips-Perron (PP) y las prueba KPSS (Kwiatkowski–Phillips–Schmidt–Shin).

A continuación se realizara la prueba de Dickey Fuller Aumentado (ADF) para nuestra serie en estudio (Variación porcentual del PBI trimestral, 1981-2014).

Cuadro 9.1: Prueba de Dickey-Fuller aumentado (ADF)

```
Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root
Type of Test: t-test
Test Statistic: -3.588
P-value: 0.007249

Coefficients:
      Value Std. Error t value Pr(>|t|)
lag1 -0.2411  0.0672   -3.5884  0.0005
lag2  0.4238  0.0955    4.4373  0.0000
lag3 -0.0028  0.0857   -0.0327  0.9740
lag4  0.0638  0.0839    0.7608  0.4482
lag5 -0.3914  0.0814   -4.8060  0.0000
lag6  0.2343  0.0872    2.6882  0.0082
constant 0.8035  0.3935    2.0421  0.0433

Regression Diagnostics:

      R-Squared 0.4101
Adjusted R-Squared 0.3814
Durbin-Watson Stat 2.0239

Residual standard error: 3.71 on 126 degrees of freedom
F-statistic: 14.25 on 6 and 123 degrees of freedom, the p-value is 2.781e-012
```

Fuente: Elaboración propia.

ANEXO 6: PERIODOGRAMA DE LA SERIE Z_t

Cuadro 9.2: Periodograma de la serie Z_t

Orden	Frecuencia	Período	Transformación seno	Transformación coseno	Periodograma
1	0.0000	-	0.0000	3.3284	0.000
2	0.0074	136.0000	-2.9292	-0.3374	591.186
3	0.0147	68.0000	-1.7921	0.3244	225.559
4	0.0221	45.3333	2.1472	-0.7539	352.159
5	0.0294	34.0000	-2.4088	-2.1977	723.002
6	0.0368	27.2000	-2.7520	1.3454	638.109
7	0.0441	22.6667	-0.6399	1.6337	209.348
8	0.0515	19.4286	1.2284	1.1021	185.194
9	0.0588	17.0000	1.8452	0.2039	234.342
10	0.0662	15.1111	-0.5886	-0.4552	37.650
11	0.0735	13.6000	-1.0836	1.5266	238.313
12	0.0809	12.3636	0.9995	1.2556	175.131
13	0.0882	11.3333	0.9702	-0.5608	85.402
14	0.0956	10.4615	3.2389	-1.2511	819.810
15	0.1029	9.7143	-0.4865	-1.9258	268.296
16	0.1103	9.0667	-0.9215	0.2880	63.386
17	0.1176	8.5000	-0.6020	1.3112	141.554
18	0.1250	8.0000	0.2771	-1.2552	112.357
19	0.1324	7.5556	-1.3523	-0.1416	125.714
20	0.1397	7.1579	0.2317	1.8892	246.343
21	0.1471	6.8000	2.5667	0.0624	448.255
22	0.1544	6.4762	-0.1335	-1.4365	141.542
23	0.1618	6.1818	-1.1134	0.0343	84.377
24	0.1691	5.9130	0.3307	1.5827	177.764
25	0.1765	5.6667	0.4270	-0.4872	28.539
26	0.1838	5.4400	-0.0834	-0.0142	.487
27	0.1912	5.2308	0.0266	0.6240	26.526
28	0.1985	5.0370	0.9412	0.1030	60.962
29	0.2059	4.8571	-0.0632	-0.3737	9.766
30	0.2132	4.6897	-0.3629	-0.1811	11.184
31	0.2206	4.5333	-0.2785	0.3991	16.109
32	0.2279	4.3871	0.4752	0.0993	16.023
33	0.2353	4.2500	-0.0025	-0.0760	.393
34	0.2426	4.1212	-0.2276	0.2183	6.763
35	0.2500	4.0000	-0.0243	0.1316	1.218
36	0.2574	3.8857	0.1476	-0.0712	1.827
37	0.2647	3.7778	0.1540	-0.1348	2.850
38	0.2721	3.6757	-0.1189	0.3264	8.204
39	0.2794	3.5789	-0.0724	0.0855	.854
40	0.2868	3.4872	0.3531	-0.2649	13.249
41	0.2941	3.4000	-0.0758	-0.0456	.532
42	0.3015	3.3171	-0.3467	0.2508	12.451

Fuente: Elaboración propia.

ANEXO 7: CÓDIGO EN SPLUS 8.0

```
=====
MODELO SETAR APLICADO A LA VARIACIÓN PORCENTUAL DEL PBI, PERÚ 1981-2014
=====

# GRÁFICO DE LA SERIE: VARIACIÓN PORCENTUAL DEL PBI 1981-2014
#-----

SDF6.rts <- as.rts(SDF6)
SDF6time <- time(SDF6.rts)
ts.plot(SDF6.rts,reference.grid=F,main="Variación porcentual del
      PBI")

# GRÁFICO DE LAS AUTOCORRELACIONES
#-----

par(mfrow=c(2,2))
ndx.acf = acf(SDF6.rts,col="blue",main="ACF")
ndx.pacf = acf(SDF6.rts, type="partial", col="red",main="ACP")

# ORDEN AR(p) DE LA SERIE ORIGINAL
#-----

library(MASS)
lh.arima1 <- arima(SDF6.rts, order = c(6,0,0))

# DENSIDAD ESPECTRAL Y PERIODOGRAMA DE LA SERIE
#-----

par(mfrow = c(2, 2))
spectrum(SDF6.rts)
spectrum(SDF6.rts, spans = 3)
spectrum(SDF6.rts, spans = c(3, 3))
spectrum(SDF6.rts, spans = c(3, 5))

# PRUEBA DE NO LINEALIDAD
#-----

library(finmetrics)
nonlinearTest(SDF6.rts, method="threshold", p=6, d=1:6)
nonlinearTest(SDF6.rts, method="threshold", p=5, d=1:5)
nonlinearTest(SDF6.rts, method="threshold", p=4, d=1:4)
nonlinearTest(SDF6.rts, method="threshold", p=3, d=1:3)
nonlinearTest(SDF6.rts, method="threshold", p=2, d=1:2)

ndx.test = nonlinearTest(SDF6.rts, method="threshold", p=6, d=1,
save.RLS=T)
names(ndx.test)

# REGIÓN DEL VALOR UMBRAL
#-----

par(mfrow=c(2,1))
plot(ndx.test$yd, ndx.test$tRatio[,1], xlab="Y_{t-1}",ylab="t-ratio
of AR(1)",col="blue")
plot(ndx.test$yd, ndx.test$tRatio[,2], xlab="Y_{t-1}",ylab="t-ratio
of AR(2)",col="green")
```

```

# PRUEBA DE HANSEN
#-----
library(finmetrics)
nonlinearTest(SDF6.rts, method="sup-LR", p=6, d=1, trim.pct=0.1,
n.boot=1000)
nonlinearTest(SDF6.rts, method="sup-LR", p=6, d=1, trim.pct=0.1,
n.boot=1000, hetero=T)
ndx.setar.r = TAR(SDF6.rts, p=6, d=1, trim.pct=0.1)
ndx.setar.r

# INTERVALOS DE CONFIANZA DEL UMBRAL
#-----
ndx.setar.r=TAR(SDF6.rts,p=6,d=1,trim.pct=0.1)
names(ndx.setar.r$LR.q)
plot(ndx.setar.r$LR.q$Threshold, ndx.setar.r$LR.q$LR,type="b",
xlab="Threshold", ylab="LR stat")
abline(h=ndx.setar.r$LR.q$Critical)

# ESTIMACIÓN DEL MODELO SETAR
#-----
library(finmetrics)
ndx.setar = SETAR(SDF6.rts, c(5.648), p=6, d=1)
summary(ndx.setar)

# VALIDACIÓN DEL MODELO SETAR
#-----

#Estimación de los residuos
ndx.resid=resid(ndx.setar)
ndx.resid

#Histograma y densidad normal estimada de los residuos
histogram(ndx.resid,probability=T,ylim=c(0,30))
dens<- density(ndx.resid,n=130)
points(dens,type="l")

#Gráfico Q-QPLOT
tsplot(cbind(ndx.resid))
qqnorm(ndx.resid)
qqline(ndx.resid)

#Espectro de los residuos
par(mfrow = c(2, 2))
spectrum(ndx.resid)
spectrum(ndx.resid, spans = 3)
spectrum(ndx.resid, spans = c(3, 5))

#Gráfico de autocorrelación de los residuos
AutoccoResid = acf(ndx.resid,col="blue",main="ACF de los residuos")

#Gráfico de la serie original vs serie estimada
ndx.fitted = fitted(ndx.setar)
tsplot(cbind(ndx.fitted))
tsplot(SDF6.rts,ndx.fitted)

#Prueba de normalidad de los residuos
ks.gof(ndx.resid, distribution = "normal")
normalTest(ndx.resid,method="sw")
normalTest(ndx.resid,method="jb")

```

```

#Prueba de incorrelación de los residuos
autocorTest (ndx.resid,lag.n=12,method="lb")
autocorTest (ndx.resid,lag.n=12,method="bp")

# PREDICCIONES CON EL MODELO SETAR
#-----
class (ndx.setar)
ndx.pred = predict (ndx.setar, n.predict=4, CI.alpha=0.6,n.sim=10000)
tsplot (cbind (ndx.pred$values))
tsplot (cbind (ndx.pred$values, ndx.pred$CI), lty=c (1,6,6))

# ANEXOS: PRUEBA DE RAÍCES UNITARIAS (DFA)
#-----

adf.out=unitroot (SDF6.rts,trend="c",statistic="t",method="adf",lags=
6)
class (adf.out)
adf.out
summary (adf.out)

```

ANEXO 8: CÓDIGO EN R-PROJECT 3.3.4

MODELO SETAR APLICADO A LA VARIACIÓN PORCENTUAL DEL PBI, PERÚ 1981-2014

```
# SERIE ORIGINAL 1980-2014: Datos Trimestrales
#-----
data<-read.csv(file="D:/work/PBI.csv")
PBI<-ts(data[,1],start=1980,freq=4)
plot(PBI,col='red',ylab='PBI',xlab='Año')

# VARIACIÓN TRIMESTRAL DEL PBI 1981-2014
#-----
data<-read.csv(file="D:/work/CRECIMIENTO.csv")
CRECIMIENTO<-ts(data[,1],start=1981,freq=4)
plot(CRECIMIENTO,col='blue',ylab='Variación',xlab='Año')

# ESPECTRO DE LA SERIE
#-----
par(mfrow=c(2,1), mar=c(2,4,0,0))
spectrum(CRECIMIENTO, pad=1, spans=3, col='red')
p2<- spectrum(CRECIMIENTO, method="ar", plot=F)
plot(p2,col='blue')

# SERIE DESFASADA
#-----
library(tsDyn)
library(sm)
par(mfrow=c(2,2), mar=c(2,2,0,0))
autopairs(CRECIMIENTO, lag=1, type="regression")
autopairs(CRECIMIENTO, lag=2, type="regression")
autopairs(CRECIMIENTO, lag=3, type="regression")
autopairs(CRECIMIENTO, lag=4, type="regression")

# CORRELOGRAMA DE LA SERIE
#-----
par(mfrow=c(2,1), mar=c(2,4,0,0))
acf(CRECIMIENTO)
pacf(CRECIMIENTO)

# ANÁLISIS EXPLORATORIO DE LA NO LINEALIDAD
#-----
library(tseriesChaos)
mutual(CRECIMIENTO)
recurr(CRECIMIENTO, m=3, d=1, levels=c(0,0.2,1))
lag.plot(CRECIMIENTO, lags=3, layout=c(1,3))
```

```

# PRUEBA DE NO LINEALIDAD DE TSAY
#-----
Tsay.test<-function(x, order, ...) {
  if(missing(order)) order=ar(x, ...)$order
  CRECIMIENTO=as.vector(x)
  m=order
  X=NULL
  for (i in 1:m) X=cbind(X, zlag(x, i))
  X1=cbind(x, X)
  for (j in 1:m) { for (k in 1:j) X=cbind(X, zlag(x, j)*zlag(x, k)) }
  X2=cbind(x, X)
  X1=na.omit(X1)
  X2=na.omit(X2)
  y=X1[, 1]
  lm1=lm(y~X1[, -1])
  lm2=lm(y~X2[, -1])
  a1=anova(lm1, lm2)

  list(test.stat=signif(a1[[5]][2], 4), p.value=signif(a1[[6]][2], 4), order=order)
}
Tsay.test(CRECIMIENTO, order=6)

# DETERMINACIÓN DE LA VARIABLE UMBRAL
#-----
library(TSA)
pvaluem=NULL
for (d in 1:6) {
  res=tlrt(CRECIMIENTO, p=6, d=d, a=0.25, b=0.75)
  pvaluem= cbind( pvaluem, round(c(d, signif(c(res$test.statistic, res$p.value))), 3))
}
rownames(pvaluem)=c('d', 'test statistic', 'p-value')
pvaluem

# VALIDACIÓN DEL MODELO
#-----

#Residuos
data<-read.csv(file="D:/work/RESIDUOS_SETAR.csv")
RESIDUOS_SETAR<-ts(data[, 1], start=1982, freq=3)
plot(RESIDUOS_SETAR, col='red', ylab='PBI', xlab='Año')

#Histograma y curva de densidad normal
h <- hist(RESIDUOS_SETAR, breaks = 10, col = "skyblue", xlab =
"Residuos")
xfit <- seq(min(RESIDUOS_SETAR), max(RESIDUOS_SETAR), length = 40)
yfit <- dnorm(xfit, mean = mean(RESIDUOS_SETAR), sd =
sd(RESIDUOS_SETAR))
yfit <- yfit * diff(h$mids[1:2]) * length(RESIDUOS_SETAR)
lines(xfit, yfit, col = "violet", lwd = 2)

```

```
#Q-QPLOT
qqnorm(RESIDUOS_SETAR)
qqline(RESIDUOS_SETAR)

#Prueba de normalidad
library(fBasics)
shapiroTest(RESIDUOS_SETAR)
lillieTest(RESIDUOS_SETAR)

#Prueba de incorrelación de los residuos
Box.test(RESIDUOS_SETAR, lag = 11, type="Box-Pierce")
Box.test(RESIDUOS_SETAR, lag = 11, type="Ljung")

#Prueba de aleatoriedad de los residuos(opcional)
library(tseries)
runs.test(as.factor(RESIDUOS_SETAR>median(RESIDUOS_SETAR)))
```