



UNIVERSIDAD NACIONAL DEL ALTIPLANO
FACULTAD DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA
ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA E
INFORMÁTICA



**MODELO DE SERIES DE TIEMPO, PARA EL PRONÓSTICO
MENSUAL DE CONSUMO DE ENERGÍA ELÉCTRICA DE LA
REGIÓN DE PUNO, ELECTRO PUNO S.A.A PERÍODO 2014-2020**

TESIS

PRESENTADA POR:

Bach. DENNIS URIEL AÑASCO CHATA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ESTADÍSTICO E INFORMÁTICO

PUNO - PERÚ

2023



Reporte de similitud

NOMBRE DEL TRABAJO

MODELO DE SERIES DE TIEMPO, PARA EL PRONÓSTICO MENSUAL DE CONSUMO DE ENERGÍA ELÉCTRICA DE LA REGIÓN DE PUNO, ELECTRO PUNO S.A.A PERÍODO 2014-2020

AUTOR

DENNIS URIEL AÑASCO CHATA

RECuento DE PALABRAS

21418 Words

RECuento DE CARACTERES

120009 Characters

RECuento DE PÁGINAS

107 Pages

TAMAÑO DEL ARCHIVO

4.4MB

FECHA DE ENTREGA

Jun 2, 2023 4:19 PM GMT-5

FECHA DEL INFORME

Jun 2, 2023 4:21 PM GMT-5

● 3% de similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para cada base c

- 2% Base de datos de Internet
- Base de datos de Crossref
- 2% Base de datos de trabajos entregados
- 0% Base de datos de publicaciones
- Base de datos de contenido publicado de Crossr

● Excluir del Reporte de Similitud

- Material bibliográfico
- Material citado
- Fuentes excluidas manualmente
- Material citado
- Coincidencia baja (menos de 15 palabras)
- Bloques de texto excluidos manualmente



UNA
PUNO

Firmado digitalmente por TITO LIPA
Jose Pantilo FAU 20145496170 hard
Motivo: Soy el autor del documento
Fecha: 02.06.2023 16:29:24 -05:00



Universidad
Nacional
del Altiplano

Firmado digitalmente por HJATA
PANCA Percy FAU 20145496170 soft
Motivo: Soy el autor del documento
Fecha: 02.06.2023 16:25:38 -05:00

Resumen



DEDICATORIA

Dedico mi tesis como muestra de amor y agradecimiento a mi Señor padre José Antonio Añasco Humpiri y a mi Señora madre Ermenegilda Filomena Chata Huamán, por su apoyo absoluto en cada paso de mi vida, les estaré eternamente agradecido.

A mi hermana Gladys Cynthia Añasco Chata y familia, quien me apoyó incondicionalmente en cada etapa de mi vida.

A mi esposa Amalia Maribel Quispe Tuni, por su aliento, amor y apoyo absoluto para seguir adelante.

A la memoria de mi tía Carmen Añasco Humpiri (†), que desde el cielo guía cada paso de sus hijos y de la familia Añasco como en vida fue.

Dennis Uriel Añasco Chata



AGRADECIMIENTOS

Agradecer a Dios por darme la fuerza, sabiduría y paciencia para lograr mis objetivos.

A mi alma mater la Universidad Nacional del Altiplano de Puno y a la Escuela Profesional de Ingeniera Estadística e Informática.

Un merecido agradecimiento al ing. Zenon Ronal Jara Cajigas, por su valioso aporte a mi formación profesional y la amistad brindada.

Un especial agradecimiento a mis primos Rosmery D. Cárdenas Añasco, Ivan G. Cárdenas Añasco, Ramiro Cárdenas Añasco, Bilfredo Cárdenas Añasco, Guido E. Cárdenas Añasco, quienes contribuyeron con su experiencia y sus valiosos consejos a mi formación personal y profesional.

Un agradecimiento a mi primo Bill B. Cárdenas Añasco, por su amistad y apoyo desinteresado en nuestra vida universitaria.

A mi tía Hermelinda Añasco Humpiri, por sus valiosos consejos para seguir cumpliendo mis objetivos.

A mi director de tesis al D.Sc. Percy Huata Panca y al jurado calificador presidido por el Dr. Samuel D. Pérez Quispe, la Dra. Emma O. Azañero De Aguirre y al Mg. Elqui Y. Pari Condori, quienes contribuyeron a la mejora de este trabajo de investigación.

A mis demás familiares, compañeros de estudio y estimados amigos.

Dennis Uriel Añasco Chata



ÍNDICE GENERAL

DEDICATORIA

AGRADECIMIENTOS

ÍNDICE GENERAL

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

ÍNDICE DE ACRÓNIMOS

RESUMEN 11

ABSTRACT..... 12

CAPITULO I

INTRODUCCIÓN

1.1 PLANTEAMIENTO DEL PROBLEMA..... 15

1.1.1 Descripción del problema 15

1.2 FORMULACIÓN DE PROBLEMA 17

1.2.1 Problema general 17

1.2.2 Problema específicos 17

1.3 JUSTIFICACIÓN DEL PROBLEMA..... 17

1.3.1 Delimitación Del Proyecto: 19

1.4 OBJETIVOS DE LA INVESTIGACIÓN..... 20

1.4.1 Objetivo general..... 20

1.4.2 Objetivos específicos 20

CAPITULO II

REVISIÓN DE LITERATURA

2.1 ANTECEDENTES DE LA INVESTIGACIÓN 21

2.1.1 Nivel internacional..... 21



2.1.2	Nivel nacional	22
2.1.3	Nivel local.....	23
2.2	MARCO TEÓRICO	24
2.2.1	Pronóstico	24
2.2.2	Series De Tiempo.....	25
2.2.3	Proceso Estocástico.....	28
2.2.4	Ruido Blanco	30
2.2.5	Operador De Retador Y Diferencias.....	31
2.2.6	Modelos De La Metodología Box Y Jenkins.....	33
2.2.7	Fases De La Metodología Box-Jenkins	37
2.3	MARCO CONCEPTUAL	47
2.3.1	Energía eléctrica	47
2.3.2	Instituciones involucradas en el sector eléctrico.....	48
2.3.3	Horizonte de tiempo.....	49
2.4	HIPÓTESIS DE LA INVESTIGACIÓN	51
2.4.1	Hipótesis General.....	51
2.4.2	Hipótesis Especifica.....	51
CAPITULO III		
MATERIALES Y MÉTODOS		
3.1	UBICACIÓN GEOGRAFICA DE LA INVESTIGACIÓN.....	52
3.1.1	Localización de la investigación.....	52
3.1.2	Ámbito de estudio.....	53
3.2	POBLACIÓN Y MUESTRA DE LA INVESTIGACIÓN.....	56
3.2.1	Población de estudio	56
3.3	RECOLECCIÓN DE DATOS	56



3.4	TIPO DE INVESTIGACIÓN.....	57
3.4.1	Diseño y tipo de investigación.....	57
3.4.2	Metodología para la predicción de series de tiempo.....	57
3.4.3	Procesamiento de datos.....	60
CAPITULO IV		
RESULTADOS Y DISCUSIÓN		
4.1	RESULTADOS.....	63
4.1.1	Identificación	63
4.1.2	Estimación	77
4.1.3	Validación.....	79
4.1.4	Pronóstico	84
4.2	DISCUSIÓN.....	90
V.	CONCLUSIONES.....	93
VI.	RECOMENDACIONES	94
VII.	REFERENCIAS BIBLIOGRÁFICAS.....	95
ANEXOS.....		99

Área : Estadística
Tema : Modelos Predictivos Uni y Multivariados

FECHA DE SUSTENTACIÓN: 06 de junio de 2023



ÍNDICE DE FIGURAS

Figura 1. Serie de Ruido Blanco (White noise).....	31
Figura 2. ACF Para El Cambio Porcentual Trimestral en el Consumo de EE.UU.....	39
Figura 3. PAC Para el Cambio Porcentual Trimestral en el Consumo de EE. UU	40
Figura 4. Proceso de Cálculo Mediante Modelos ARIMA.....	46
Figura 5. Funcionamiento Del Sistema Eléctrico En El Perú.....	47
Figura 6. Diagrama de flujo metodología Box-Jenkins.....	59
Figura 7. Diagrama Para El Procesamiento De La Información	61
Figura 8. Consumo de Energía Eléctrica de la Región Puno 2014-2020	64
Figura 9. Comportamiento del Consumo de Energía Eléctrica por Meses y Año.....	65
Figura 10. Descomposición de la Serie Consumo de Energía Eléctrica.....	66
Figura 11. Histograma de la Serie Consumo De Energía Eléctrica De La Región Puno	67
Figura 12. Correlograma Simple y Parcial de la Serie Consumo de Energía Eléctrica.	70
Figura 13. Serie Estacionaria de Consumo De Energía Eléctrica	72
Figura 14. Distribución De La Serie Diferenciada	73
Figura 15. Función de autocorrelación simple y parcial de la primera diferencia de la serie	75
Figura 16. Autocorrelación Simple y Parcial de los Residuos del Modelo Seleccionado	82
Figura 17. Serie Generada por el Modelo <i>ARIMA</i> validada, frente a la Serie Original	87
Figura 18. Pronóstico del Consumo de Energía Eléctrica para los periodos 2021 – 2023	90



ÍNDICE DE TABLAS

Tabla 1 Subestaciones Eléctricas de Transmisión	54
Tabla 2 Sistemas Eléctricos de ELECTRO PUNO S.A.A.....	55
Tabla 3 Consumo Mensual de Energía Eléctrica de la Región Puno 2014-2020	63
Tabla 4 Test Aumentado de Dickey – Fuller	69
Tabla 5 Serie Diferenciada del Consumo de Energía Eléctrica de la Región Puno	71
Tabla 6 Test Aumentado de Dickey – Fuller	74
Tabla 7 Modelos ARIMA(p, d, q) \times P, D, Q12 identificados	77
Tabla 8 Coeficientes del modelo ARIMA.....	78
Tabla 9 Modelos Candidato Para el Pronóstico de Energía Eléctrica de la Región Puno	80
Tabla 10 Test de Q Ljung-Box	83
Tabla 11 Modelos ARIMA y sus Métricas de Precisión del Pronóstico.....	85
Tabla 12 Escala de Lewis(1982) Para el Análisis de Precisión MAPE.....	85
Tabla 13 Pronóstico consumo de Energía Eléctrica Para la Región de Puno Durante 3 Años	88



ÍNDICE DE ACRÓNIMOS

ACF:	: Función de Autocorrelación Simple
AIC	: Criterio de Información de Akaike
ARIMA	: Modelos Autorregresivos Integrados De Promedios Móviles
BIC	: Criterio de Información de Bayes
COES	: Comité De Operación Económica Del Sistema
COES-SINAC	: Comité de Operación Económica del Sistema Interconectado Nacional
ELECTRO PUNO	: Empresa Regional De Servicio Público De Electricidad de Puno
ELPU	: Electro Puno
INEI:	: Instituto Nacional de Estadística e Informática
KWh	: Kilowatt por hora
MAE	: Error Medio Absoluto
MAPE	: Error Medio Absoluto Porcentual
MINEM	: Ministerio De Energía Y Minas
MWh	: Megawatt por hora
OIEA	: Organización Internacional De Energía Nuclear
OSINERMIN	: Organismo Supervisor de la Inversión en Energía y Minería
PACF	: Función de Autocorrelación parcial
PBI	: Producto Bruto Interno
RMSE	: Raíz Media De Los Errores Al Cuadrado
SEIN	: Sistema Eléctrico Interconectado Nacional



RESUMEN

El consumo de energía eléctrica en los hogares, comercio e industria se ha ido incrementando continuamente en la región de Puno, este incremento está relacionado con el crecimiento de la población y estilo de vida, ELECTRO PUNO S.A.A como proveedora del servicio de energía eléctrica es la responsable de garantizar el suministro de energía eléctrica hacia los usuarios, de manera continua y segura. Por tal motivo, el objetivo de este trabajo fue determinar un modelo de series de tiempo basado en la metodología de Box-Jenkins, que se ajuste apropiadamente a la serie original en el pronóstico del consumo de energía eléctrica de la región de Puno para los próximos tres años, esto permitiría a la empresa contar con una metodología detallada que respalde la toma de decisiones, que se ajusten a sus objetivos corporativos en un horizonte de tiempo a mediano plazo, que garanticen la distribución y la comercialización de energía eléctrica con decisiones basadas en datos. Para tal propósito se contó con datos históricos mensuales de consumo de energía eléctrica del 2014 al 2020 obtenidos del área de facturación de la empresa ELECTRO PUNO S:A:A, se aplicó las fases de la metodología Box-Jenkins como: la fase de identificación, estimación y validación del modelo, llegando a determinar un modelo multiplicativo estacional $ARIMA (0,1,3) \times (2,0,0)_{12}$ que se ajustó perfectamente a la serie en estudio, siendo el grado de precisión de los pronósticos de “alta precisión” con un valor *MAPE* del 2.79% , finalmente en la última fase de la metodología se realizó los pronósticos mensuales de consumo de energía eléctrica para los años 2021,2022 y 2023, con una desviación en promedio por cada pronóstico de un valor $MAE = 625.80$ MWh.

Palabras Clave: Energía eléctrica, Consumo, Puno, Box-Jenkins, Pronóstico.



ABSTRACT

The consumption of electricity in homes, commerce and industry has been increasing continuously in the Puno region, this increase is related to population growth and lifestyle, ELECTRO PUNO S.A.A. as a provider of electric power service is responsible for ensuring the supply of electricity to users, continuously and safely. For this reason, the objective of this work was to determine a time series model based on the Box-Jenkins methodology, which is appropriately adjusted to the original series in the forecast of electricity consumption in the region of Puno for the next three years, this would allow the company to have a detailed methodology to support decision making, which conforms to its corporate objectives in a medium-term time horizon, to ensure the distribution and marketing of electricity with data-based decisions. For such purpose, monthly historical data of electricity consumption from 2014 to 2020 obtained from the billing area of the company ELECTRO PUNO S:A:A, the phases of the Box-Jenkins methodology were applied as: the phase of identification, estimation and validation of the model, reaching to determine a multiplicative seasonal $ARIMA (0,1,3) \times (2,0,0,0)_{12}$ model that was perfectly adjusted to the series under study, being the degree of accuracy of the forecasts of "high accuracy" with a MAPE value of 2.79%. Finally, in the last phase of the methodology, monthly forecasts of electricity consumption were made for the years 2021, 2022 and 2023, with an average deviation for each forecast of a MAE value =625.80 MWh.

Key words: Electricity, Demand, Puno, Box-Jenkins, Forecasting.



CAPITULO I

INTRODUCCIÓN

El Banco Mundial señala que, la energía es un elemento fundamental en el desarrollo de una sociedad, promueve la inversión, innovación y la aparición de nuevas industrias que constituyen una fuente de creación de empleos; por otra parte, para las familias el acceso a la electricidad significa un desarrollo en su calidad de vida.

Desde su aparición de la energía eléctrica en la humanidad, la demanda de energía eléctrica se encuentra en constante crecimiento, instituciones del rubro eléctrico y empresas eléctricas dedican recursos para la realización de estudios que contribuyan a determinar la energía que se necesitará en un futuro con alcances de tiempo en el corto, mediano y largo plazo, como parte de un trabajo de planificación energética que guiarán al establecimiento de políticas y facilitará el proceso en la toma de decisiones a los gerentes, en los diferentes niveles (Hanke & Wichern, 2018).

En ese sentido, consideramos que en un proceso de planificación para mantener en operación un sistema de distribución de energía hacia los usuarios de manera continua, segura y con el menor número de posibles interrupciones, es importante conocer el comportamiento futuro del consumo de energía eléctrica, para el establecimiento de políticas principalmente en el mantenimiento preventivo del sistema, el mismo que debe estar respaldado por un estudio que sirva como soporte para la toma de decisiones, logrando reducir de esa forma la incertidumbre sobre los eventos que vayan a ocurrir en un futuro. Por esta razón, se hace necesario desarrollar un modelo matemático apropiado, que garantice la confiabilidad de los resultados en el consumo futuro de energía para los próximos tres años.



En ese marco, la pregunta central es ¿Cuál es el modelo determinado con la metodología Box - Jenkins para el pronóstico mensual de consumo de energía eléctrica para los años 2021 – 2023 en la región de Puno?, como hipótesis principal se ha planteado que el modelo ARIMA proporciona un mejor ajuste para validar y pronosticar el consumo de energía eléctrica de la región de Puno, y como objetivo general es determinar un modelo de series de tiempo que pronostique el consumo de energía eléctrica a mediano plazo para la región de Puno, empleando modelos estocásticos univariados *ARIMA* conjuntamente con la metodología de Box – Jenkins.

Para tal propósito, se utilizó datos históricos de la venta mensual de energía eléctrica de los periodos comerciales desde enero-2014 a diciembre-2020 que ascienden a un total de 84 datos mensuales que corresponden a 38 zonas de facturación, tomando en cuenta la totalidad en ventas de energía eléctrica representados en (MWh), a cargo de la empresa ELECTRO PUNO S.A.A..

Nuestra investigación consta de cinco capítulos; en el capítulo I se explica los fundamentos del problema de investigación, para dar lugar a la formulación del problema; en la sección 1.3 se expone la justificación del problema y la fijación de objetivos e hipótesis de investigación; el Capítulo II aborda la descripción del marco teórico, donde se presentan investigaciones relacionadas con nuestro tema de estudio, seguido de la teoría de series temporales, modelos univariantes y la metodología Box Jenkins, que son conceptos necesarios para el entendimiento de este trabajo; el capítulo III muestra los materiales y métodos que serán empleados para la obtención de los modelos ARIMA con la metodología Box Jenkins; en el Capítulo IV se exponen los resultados y discusión de la investigación que responden a los objetivos e hipótesis planteadas, en la parte final se encuentran las conclusiones y recomendaciones, por último los anexos.



1.1 PLANTEAMIENTO DEL PROBLEMA

1.1.1 Descripción del problema

La energía eléctrica es un bien necesario para el uso de los hogares, comercio e industria, por ende, su estudio en la etapa de planificación es de vital importancia en todos los niveles del sector eléctrico (generación, transmisión, distribución), por esa razón la proyección de consumo de energía eléctrica es parte fundamental para el estudio de planificación energética, sin embargo en el campo de la Ingeniería Eléctrica no existe literatura que ayude a proyectar la demanda de energía eléctrica en escenarios futuros, es ahí donde la teoría estadística y econometría desempeñan un rol importante (Marcelo Barreto & Villareal Escate, 2021).

La predicción de energía eléctrica se ha convertido en una herramienta fundamental en los procesos de planificación energética (Chong Fuentes & Aguilar, 2016), es considerado como primer paso de estudio en la etapa de planificación y análisis operacional en el sector eléctrico (Cervan, 2017), sus resultados sirven como indicadores para la formulación de planes estratégicos y operacionales en la industria de generación, transmisión y distribución de energía eléctrica.

Al respecto, la OIEA es un Organismo Internacional De Energía Atómica que está conformada por 176 países miembros, este organismo ha desarrollado el Modelo De Análisis De La Demanda De Energía, que sirve para proyectar la demanda de electricidad a mediano ó largo plazo, los resultados que se obtiene de dicho análisis permite a sus países miembros a tomar decisiones energéticas inteligentes, que van desde la compra de energía hasta la construcción de nuevas infraestructuras eléctricas.

Los estudios realizados por el MINEM en el 2008 para la expansión del Sistema Interconectado Nacional del 2008 al 2017, demuestran que las metodologías de series de



tiempo y los modelos econométricos son ampliamente utilizados en el sector eléctrico, precisamente en las etapas de estudio previo para determinar la viabilidad de un proyecto energético.

Las proyecciones en la etapa de planificación están asociadas a horizontes de tiempo, Cervan (2017) menciona la clasificación de los horizontes de planificación en el sector eléctrico de acuerdo al tiempo de alcance, el tiempo de horizonte a mediano plazo es conocido como “Pronóstico de carga a mediano plazo” con un tiempo de duración de semanas a meses, que se emplea para mantener el equilibrio entre la demanda y la oferta de generación de energía, para la programación de mantenimiento del sistema eléctrico, además para pronosticar los picos de demanda en los siguientes meses.

La empresa ELECTRO PUNO S.A.C actualmente no cuenta con un procedimiento metodológico que permita predecir la demanda de energía eléctrica a mediano plazo; que contribuiría en los planes de programación de mantenimiento a los sistemas de distribución oportunamente, sin embargo, las políticas de mantenimiento a los sistemas de distribución que adopta la empresa se dan en base a factores como:

- Vida útil de los equipos electromecánicos
- Mantenimiento por indicación del fabricante
- Experiencia del personal a cargo del mantenimiento del sistema eléctrico
- Incidentes no previstos por fallas técnicas, atmosféricas y otros

La no aplicación de metodologías estadísticas y/o econométricas como parte de un estudio previo en la programación de mantenimientos a los sistemas de distribución, incrementa la incertidumbre en la toma de decisiones generando posibles interrupciones



en el servicio del suministro de energía eléctrica hacia los usuarios, por no anticiparse a hechos futuros.

1.2 FORMULACIÓN DE PROBLEMA

1.2.1 Problema general

- ¿Cuál es el modelo de series de tiempo a determinar con la metodología Box – Jenkins, para el pronóstico mensual de consumo de energía eléctrica para los años 2021 – 2023 en la región de Puno?

1.2.2 Problema específicos

- ¿Cuál es el comportamiento del consumo de energía eléctrica de la Región Puno en los periodos 2014-2020?
- ¿Cuál es el modelo estimado de consumo de energía eléctrica de la Región Puno para los periodos 2014-2020?
- ¿Cuáles son las proyecciones de consumo mensual de energía eléctrica de la Región Puno para los años 2021-2023?

1.3 JUSTIFICACIÓN DEL PROBLEMA

Electro Puno S-A-A es una empresa Regional de Servicio Público de Electricidad, cuyo objetivo es realizar todas las actividades de distribución y comercialización de energía eléctrica dentro de su zona de concesión que comprende en la Región de Puno, así mismo desarrolla en menor medida actividades de sub -transmisión y de generación, a través de pequeños sistemas, siendo el más importante la Central Hidráulica de Chijisia, ubicada en la ciudad de Sandia.

Los reportes realizados por ELECTRO S.A.A en el año 2021 señalan que en la región de Puno desde enero del 2014 a Diciembre 2020 el número de clientes se ha



incrementado en 66,176 usuarios durante 7 años, teniendo a diciembre de 2020 un total de 403,141 usuarios; anualmente el incremento de clientes en promedio es de 11,029 representado por una tasa promedio de crecimiento del 3% anual, el reporte también revela que el consumo promedio anual de un cliente residencial en el año 2014 era de 46.14 Kwh, en cambio para el año 2020 el consumo promedio anual se elevó a 95.76 Kwh, reflejando un incremento del 52% durante 7 años, así mismo el sector industrial como la minería ha registrado incrementos durante los 7 años obteniendo en el año 2020 un incremento del 82.82 MW con respecto a los años anteriores hasta el año 2014.

Investigadores como Quispe Pacco (2015) y Cervan (2017) conceptúan que el incremento en la demanda de energía eléctrica tanto en la zona urbana como en la zona rural se fundamenta en el crecimiento del estilo de vida de la población y el crecimiento poblacional, que hacen que la utilización de energía sea constante en los hogares, en el caso del sector industrial la demanda de bienes y servicios, la creación de nuevas empresas en el sector implica mayor consumo de energía, por ende el servicio de suministro de energía eléctrica debe ser continua y de calidad.

Esto implica que la distribución y la comercialización de energía eléctrica hacia los usuarios debe llevarse a cabo de manera continua y segura, por lo tanto la empresa deberá contar con una eficiente planeación, mantenimiento y control de los equipos del sistema de distribución; de ocasionarse una falla en el sistema de distribución este genera una paralización en las actividades cotidianas de las familias, en las empresas del sector industrial que se dedican a ofrecer bienes y servicios generaría pérdidas económicas y materiales; en relación a ELECTRO PUNO S.A.A por encontrarse en un sector bastante regulado por el OSINERMIN se haría acreedor de multas afectando la estabilidad económica de la empresa. (Maque Tinta, 2017).



Para una adecuada planificación y toma de decisiones energéticas se requiere de una adecuada proyección de la demanda de energía eléctrica, para contribuir a la reducción de incertidumbre de los hechos que ocurrirán en un futuro (Ducos Ramiez, 2009).

Esta tesis se justifica en la medida que el pronóstico a mediano plazo de consumo de energía eléctrica realizados en base a los modelos ARIMA, permitirá detallar la demanda de energía futura de la región de Puno en un horizonte de tiempo de tres años, en donde la empresa ELECTRO PUNO S.A.A. podrá conocer el comportamiento del mercado a mediano plazo, definir la programación de manteamientos al sistema en zonas de mayor consumo y para pronosticar los picos de demanda en los siguientes meses, con la finalidad de garantizar un servicio con el menor número de interrupciones o fallas imprevistas.

1.3.1 Delimitación Del Proyecto:

El proyecto presenta una gran complejidad, puesto que se requiere la opinión en temas técnicos y económicos de otras especialidades; por esta razón el proyecto se centrará en gran medida en el análisis estadístico, econométrico de la variable consumo de energía eléctrica de la Región de Puno, tomando como fuente de información los datos históricos de consumos mensuales de energía eléctrica (carga vegetativa) representados en la unidad de medida (Mwh, Kwh), correspondientes a los periodos comerciales 2014 a 2020.



1.4 OBJETIVOS DE LA INVESTIGACIÓN

1.4.1 Objetivo general

- Determinar un modelo de series de tiempo con la metodología de Box-Jenking modelo ARIMA, para el pronóstico mensual de consumo de energía eléctrica de la Región de Puno, a fin de pronosticar los periodos 2021 a 2023.

1.4.2 Objetivos específicos

- Analizar el comportamiento de los consumos mensuales de Energía Eléctrica de la Región de Puno 2014-2023.
- Estimar un modelo de series de tiempo con la metodología de Box-Jenkins para el consumo mensual de energía eléctrica de la Región de Puno.
- Proyectar el consumo mensual de energía eléctrica de la Región de Puno, para los años 2021 a 2023.



CAPITULO II

REVISIÓN DE LITERATURA

2.1 ANTECEDENTES DE LA INVESTIGACIÓN

2.1.1 Nivel internacional

Barahona Bonifaz (2018) En Ecuador – Riobamba realizó la investigación “Pronóstico del consumo y demanda de potencia máxima de energía eléctrica”, el autor plantea como objetivo pronosticar el consumo y demanda de la potencia máxima de energía eléctrica para un periodo de 2 años, para este estudio se tomó información proporcionada por la empresa eléctrica RIOBAMBA S.A de los periodos 2013 al 2017. Aplicando la ciencia de la geo-estadística se hace el análisis y la modelación de las variables espaciales, identificando el sector comercial como principal consumidor de energía eléctrica entre los años 2015-2017 situadas en el centro de la ciudad de Riobamba, para el pronóstico de consumo de energía eléctrica el investigador llega a obtener el modelo $ARIMA(0,1,1) \times (0,1,1)_{12}$, finalmente para el pronóstico de la potencia máxima se aplicó el método de Holt-Winter.

Mejía Vasquez & Gonzales Chávez (2019) en Cuba publicó la revista científica “Pronóstico de consumo de energía eléctrica de usuarios residenciales de la Región Cajamarca”, la investigación plantea como objetivo pronosticar el consumo de energía eléctrica para la región de Cajamarca de los usuarios residenciales, aplicando el modelo Holt-Winters. Para la ejecución del estudio se utilizó la data histórica de consumo de tres empresas eléctricas de la Región Cajamarca - Perú, Hidrandina S.A., Electronorte S.A. y Electro Oriente S.A., de enero 2005 a diciembre 2017, haciendo un total de 156 observaciones, el método aplicado para el pronóstico es el modelo de Holt-Winter, los autores destacan como principal fortaleza que, los modelos de Holt- Winters son fáciles



de implementar y que usa bajos recursos computacionales para hacer pronósticos. La investigación concluye con la selección del modelo Holt-Winters Aditivo como mejor modelo de pronóstico, obteniendo un error porcentual absoluto medio (MAPE) de 1.70%.

Rey Onzaga (2018) En Colombia se ejecutó la investigación “Predicción del consumo de energía eléctrica para la ciudad de Bogotá”, investigación que tuvo como objetivo elaborar un pronóstico de consumo de energía eléctrica aplicando herramientas investigativas, matemáticas y computacionales. La investigación fue desarrollada en base a la utilización de 6 variables socio- económicos de la ciudad de Bogotá, estos son: temperatura superficial, ingreso per cápita, producto bruto interno, suscriptores, precio y consumo de energía eléctrica, precio del servicio del gas natural, información que fue recogida desde el año 2005 hasta el año 2016. Concluyendo que, la investigación presenta al método de regresión lineal Múltiple, como un método que explora y pronostica con mayor efectividad las variables económicas (demanda de energía eléctrica (...)), por otra parte, las variables sociales presentan dificultades para ser tomados en cuenta en los modelos de pronóstico, incluso con los modelos más sofisticados no se podría explicar, ya que estos son fenómenos de la vida real e infinitos.

2.1.2 Nivel nacional

Vázquez Díaz & Gamonal Sánchez (2019) En Trujillo realizó la investigación “Pronóstico De Consumo Mensual De Energía Eléctrica, De La Provincia De Bagua Grande”, los autores plantean determinar un modelo que pronostique el consumo de energía eléctrica mensual de la ciudad de Bagua Grande. la investigación cuenta con información histórica de consumos de energía eléctrica de la Empresa Municipal de Servicios de Utcubamba SAC. de los periodos 2009 a 2015, Información que fue sometida a técnicas y pruebas estadísticas para identificar el modelo *ARIMA*, finalmente



se concluye determinando que el mejor modelo para el pronóstico de consumo de energía eléctrica es el modelo $ARIMA(4,2,0)$, con un coeficiente de determinación de $R^2 = 0,996$.

Arévalo Cabrera & Santisteban Morales (2020) en Lambayeque realizó la investigación “ Pronóstico de la producción de energía eléctrica de la región Lambayeque”, el estudio tiene como objetivo determinar un modelo optimo con la metodología Box Jenkins para pronosticar la producción de energía eléctrica para el periodo 2018,. para el análisis y determinación del modelo se utilizó información de la producción mensual de energía eléctrica de los periodos 2005 al 2017 obtenido del portal web del Banco Central de Reserva del Perú (BCRP), obteniendo como resultado el modelo multiplicativo estacional $ARIMA(2,1,2) \times (1,1,0)_{12}$, que fue de mejor ajuste a la serie original con un error promedio histórico de $MAE = 1.6154$ y un error promedio porcentual de $MAPE$ de 19.3201%.

Marcelo Barreto & Villareal Escate (2021) logran determinan que el modelo multiplicativo estacional $ARIMA(0,1,1) \times (1,0,2)_{12}$ predice eficientemente la demanda mensual de energía eléctrica residencial de la región de Cusco, el modelo es respaldado por la métrica de evaluación MAPE con un valor porcentual del 2.296%, además señala que la aplicación de la metodología de Box- Jenkins se ajusta perfectamente a las series energía eléctrica históricas, confirmando que el modelo ARIMA es una herramienta que puede ser aplicado en la etapa de planificación energética.

2.1.3 Nivel local

Quispe Colque (2019) realizó un modelo de series de tiempo con la metodología Box - Jenkins , para pronosticar la cantidad de ventas mensuales de bolsas de cemento rumi producidas por la planta cemento sur para los años 2017 a 2018, el investigador



plantea como objetivo determinar un modelo univariante para predecir y describir la cantidad de ventas de bolsas de cemento rumi a 24 meses, para tal fin la información es recogida del portal del INEI con un periodo de ventas mensuales del 2005 al 2016, la investigación concluye determinando el modelo multiplicativo estacional $ARIMA(0,2,1) \times (0,1,1)_{12}$ como modelo de mejor ajuste a la serie.

Quispe Pacco (2015) Realizó la investigación “Modelo de pronóstico para el Consumo De Energía Eléctrica para el distrito De Ayaviri ”, la investigación plantea como objetivo determinar un modelo Univariante que describa y pronostique el consumo de energía eléctrica para la ciudad de Ayaviri, el autor concluye que, el modelo que permite describir y predecir el comportamiento de consumo de energía eléctrica es el modelo multiplicativo estacional $ARIMA(3,1,1) \times (1,0,1)_{12}$. Además, considera que los modelos $ARIMA$ son una alternativa eficaz para describir y pronosticar las variables de consumo de energía eléctrica.

2.2 MARCO TEÓRICO

2.2.1 Pronóstico

En el siglo XXI con la evolución de las computadoras y la informática, los pronósticos han tomado un papel fundamental dentro de las organizaciones para la toma de decisiones basadas en datos, que son aplicadas en la industria de bienes y servicios, finanzas, comercialización, producción, entre otros. Hanke & Wichern(2018) resalta la importancia y la necesidad de hacer un buen pronóstico que ayuden a los gerentes a tomar decisiones, que guíen al establecimiento de políticas y a la planeación dentro de sus organizaciones.

La palabra “pronóstico” proviene del griego “prognostikon” que significa “enunciado de lo que es probable que ocurra en el futuro”. En el campo de la estadística



los pronósticos son una estimación cualitativa o cuantitativa de una o varias variables que conforman un evento futuro, tomando como base información actual o del pasado (Villarreal, 2016), que tiene por objetivo reducir la incertidumbre de hechos que puedan acontecer en el futuro, proporcionando una información lo más cercana posible a la realidad, respaldada por toda la comunidad científica.

De acuerdo a Hanke & Wichern (2018) existen dos tipos de pronóstico, los pronósticos cualitativos que, es una técnica antigua que no requiere de la manipulación de datos, por lo que solo se usa la “experiencia e intuición” y “juicio” del profesional encargado. En el otro extremo los pronósticos cuantitativos se basan en la manipulación compleja de datos, técnicas como los métodos de series de tiempo, regresión lineal, modelo de promedios móviles, el método de Holt – Winters, entre otros métodos de análisis cuantitativo, requieren el análisis de datos pasados para pronosticar valores futuros de una o varias variables de interés. Los pronósticos están acompañados de un horizonte de tiempo, que se deben tomar en cuenta con el objetivo de trazar la visión de la organización.

2.2.2 Series De Tiempo

Mauricio (2007) define una serie de tiempo como una secuencia de N observaciones (datos numéricos); medidos, ordenados y recolectados cronológicamente en diferentes intervalos regulares del tiempo(día, mes, año, etc.), éstas pueden contener una (serie univariante) o varias características (series multivalentes) de una variable observable, que están asociadas con un instante específico del tiempo.

Las observaciones de series de tiempo son representadas matemáticamente por Y del cual se deriva los valores que toma la variable desde el momento 1 hasta t , teniendo como conjunto de información representados por $\{Y_1, Y_2, \dots, Y_t\}$ donde Y_t es el valor que



toma una variable aleatoria de una unidad considerada en el momento t tiempo (Mauricio, 2007).

Su principal objetivo es comprender su estructura, analizar, y determinar patrones de comportamiento en los datos recolectados a través del tiempo, para realizar pronósticos a partir de valores pasados (Claudia, 2010), en diferentes campos de estudio como el sector eléctrico, economía, meteorología, geofísica, química, demografía, procesos de control, etc..

2.2.2.1 Componentes De Una Serie De Tiempo.

Para ejecutar un análisis de series de tiempo se debe identificar y analizar los componentes de la serie, a este proceso de identificación se le denomina “descomposición”, esta herramienta es fundamental para la comprensión de las series de tiempo, estos componentes son identificados como: Tendencia(T), Ciclo(C), Estacionalidad(E) y Aleatoriedad(A) (Hanke & Wichern, 2018).

Tendencia(T): La tendencia representa el incremento o declinación de una serie con respecto a un periodo de tiempo, estas se pueden manifestar mediante rectas lineales y no lineales, cuyo puntos caen de manera alternante de arriba y debajo de la línea de la tendencia (Hanke & Wichern, 2018).

Estacionalidad(E): La estacionalidad es un patrón que se repite de un periodo a otro, está asociada a factores externos de una variable en estudio como el clima (invierno, primavera, otoño y verano) o las costumbres, este componente se manifiesta en observaciones con longitudes de intervalo de tiempo semanales, trimestrales o mensuales (Hanke & Wichern, 2018).

Ciclo(C): Son movimientos hacia arriba y hacia abajo que fluctúan alrededor de los niveles de tendencia (Borwerman, O'Connell, & Koehler, 2006). Fluctuaciones que contienen una duración de 2 a 10 años, causados por factores que no están relacionados necesariamente a factores económicos.

Aleatoriedad(A): Son movimientos aleatorios de una serie de tiempo que siguen un patrón indefinido o irregular (Borwerman et al., 2006). Son causados por hechos inusuales o fortuitos (cambios climáticos, desastres naturales, huelgas, etc.), que no se pueden predecir, estos errores también forman parte del análisis, aparecen después de haber explicado los componentes de la tendencia, el ciclo, y la estacionalidad (Villarreal, 2016).

Existen dos métodos de descomposición para una serie de tiempo, se trata de los métodos de descomposición aditiva y multiplicativa (Borwerman et al., 2006).

Asimismo J Hyndman & Athanasopoulos(2018) aclara que, estos métodos de descomposición son métodos clásicos originados en el año 1920, que sirven como base o punto de partida para la mayoría de otros métodos de descomposición de una series de tiempo.

2.2.2.2 Método de descomposición aditiva

Este método de descomposición sostiene que, los datos de una serie de tiempo, son una función de la suma de sus componentes, es útil para modelar series de tiempo que manifiestan variaciones estacionales creciente o decreciente (Borwerman et al., 2006).

Así.

$$Y_t = T_t + E_t + T_t + A_t \quad (1)$$

Donde:

$Y_t =$ Valor observado de la serie temporal en el periodo t

$T_t =$ Componente de la tendencia en periodo t .

$E_t =$ Componente estacional en el periodo t .

$C_t =$ Componente ciclo en el periodo t .

$A_t =$ Componente aleatorio en el periodo t

2.2.2.3 Método de descomposición multiplicativa:

A diferencia del método aditivo, el método multiplicativo sostiene que, los datos de una serie de tiempo, son el resultado de la multiplicación de sus componentes, se puede aplicar cuando la serie muestra una variación estacional constante (los parámetros de la serie cambian en el tiempo) (Borwerman et al., 2006).

Así.

$$Y_t = T_t \times E_t \times C_t \times A_t \quad (2)$$

Donde:

$Y_t =$ Valor observado de la serie temporal en el periodo t

$T_t =$ Componente de la tendencia en periodo t .

$E_t =$ Componente estacional en el periodo t .

$C_t =$ Componente ciclo en el periodo t .

$A_t =$ Componente aleatorio en el periodo t

2.2.3 Proceso Estocástico

Un proceso estocástico es un conjunto de variables aleatorias, que corresponden a distintos instantes en el tiempo y espacio, que evolucionan en función a otra variable, las mismas que se encuentran ordenadas en el tiempo, estos procesos intentan predecir fenómenos de forma segura. En resumen, un proceso estocástico es una colección de variables aleatorias ordenadas en el tiempo (Parra, 2019).

Variabes aleatorias $\{Y_1, Y_2, Y_3\}$ ó $\{Y_t\}$

Para $t = -\infty, \dots, -2, -1, 0, 1, 2, 3, \dots, \infty$

Si consideramos el proceso Y_t , donde $t = 1$, obtendremos una variable aleatoria de (Y_1) que adquiere diferentes valores con diferente probabilidades. Gujarati & Porter (2010) clasifica los procesos estocásticos en dos grupos : procesos estocásticos estacionarios y procesos estocásticos no estacionarios.

2.2.3.1 Procesos Estocásticos Estacionarios.

Un proceso estocástico es estacionario, cuando su media, varianza y covarianza son constantes con respecto al tiempo (son invariantes a lo largo del tiempo). Sanchez de Rivera (2005) como ejemplo nos muestra a las variables climatológicas como un proceso estacionario, considerando a las estaciones del año (primavera, otoño invierno y verano).

Sea (Y_t) un proceso de series de tiempo estocástico es estable en:

Media

$$E(Y_t) = \mu \quad (3)$$

Varianza

$$Var(Y_t) = E(Y_t - \mu)^2 = \sigma^2 \quad (4)$$

Covarianza

$$\gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)] \quad (5)$$

2.2.3.2 *Procesos estocásticos no estacionarios*

Por otro lado, un proceso estocástico es “no estacionario”, cuando la serie contenga una media o una varianza que cambia con el tiempo, o ambas, o en su defecto

que cuya distribución de probabilidad varía de forma no constante, es decir, que el comportamiento de los datos es desordenado y poseen tendencia, este proceso se compara con el modelo clásico modelo de caminata aleatoria (MCA) (Aragón & Parra Royón, 2017)

$$Y_t = Y_{t-1} + \mu_t \quad (6)$$

Donde, el valor de Y en el tiempo t es igual al valor en el tiempo $(t - 1)$.

2.2.4 Ruido Blanco

Se denomina ruido blanco a las series que no muestren correlación entre sus elementos, además que cuya media es igual a cero y su varianza es constante, esta denotada por (ε_t) (J Hyndman & Athanasopoulos, 2018). Como ejemplo se muestra una serie de ruido blanco en la **Figura 1**.

$$E(\omega_t) = 0(\text{media cero})$$

$$\text{Var}(\omega_t) = \sigma^2(\text{varianza constante})$$

$$\forall k \neq 0, \text{Cov}(\omega_t, \omega_{t+k}) = 0(\text{no correlacion})$$

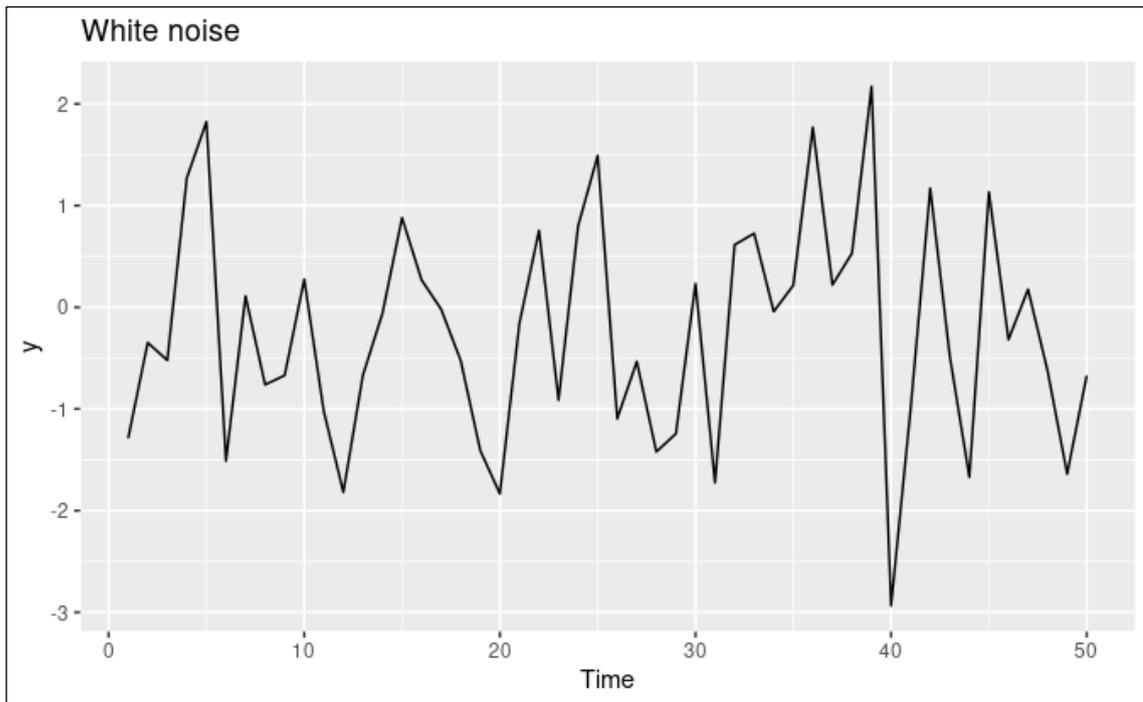


Figura 1. Serie de Ruido Blanco (White noise)

Nota. Simulación de ruido blanco en el software R. Fuente: (J Hyndman & Athanasopoulos, 2018).

Para detectar un comportamiento de ruido blanco en la serie original, se analiza visualmente la función de autocorrelación estos deben ser cercanos a cero, aplicando la prueba Q de Ljung-Box se determinará si la serie tiene un comportamiento de ruido blanco o no (Víctor Morales, 2021).

2.2.5 Operador De Retador Y Diferencias

Con frecuencia para ejecutar un análisis de series temporales se hace uso de dos nomenclaturas específicas: los operadores de retardos representados por la letra (B) u (L), y el operador de diferencias con el símbolo del triángulo invertido (∇) (Villagarcia, 2020).

2.2.5.1 Operador de Retardos

Denotado por la letra B (*Backshif*) o hace de veces por la letra L (*Lag*), se define como una función matemática de operador lineal que, al ser aplicada a una serie temporal este desfasa una posición la observación (Sanchez de Rivera, 2005).

Aplicando (B) a (z_t) está se desfasa (un período) y la transforma en z_{t-1} , ecuación (7).

$$Bz_t = z_{t-1} \quad (7)$$

Aplicando el operador de retardos dos veces obtendremos

$$\begin{aligned} B^2 z_t &= B(Bz_t = Bz_{t-1} = z_{t-2}) \\ B^2 z_t &= z_{t-2} \end{aligned} \quad (8)$$

Si aplicamos este operador en k número de veces, tendremos la expresión.

$$B^k z_t = z_{t-k} \quad (9)$$

2.2.5.2 Operador de Diferencias

Es un método general que se aplica para estabilizar una serie temporal, eliminando la tendencia y la estacionalidad, mediante una diferencia regular $(1 - B)$ ó también conocido como operador de diferencia $\nabla = (1 - B)$, que consiste en restar cada observación de la serie (z_t) con la observación anterior de la misma serie (z_{t-1}) (J Hyndman & Athanasopoulos, 2018). Es decir:

$$w_t = z_t - z_{t-1} \quad (10)$$

Puede escribirse con la siguiente expresión:

$$w_t = z_t - z_{t-1} = z_t - Bz_t = (1 - B)z_t \quad (11)$$

Donde se multiplica z_t por $(1 - B)$, que es equivalente a tomar una diferencia, obteniendo la ecuación (12).

$$z_t - z_{t-1} = \nabla z_t \quad (12)$$

Cuando usamos el símbolo ∇z_t , indica que se ha realizado una diferencia (d) a la serie, de aplicarse dos diferencias de orden 2, será representada por $\nabla^2 z_t$ (Villagarcía, 2020).

2.2.6 Modelos De La Metodología Box Y Jenkins

Los modelos de series de tiempo desarrollados en la metodología Box - Jenkins, sirve para representar formalmente los procesos aleatorios (procesos estocásticos) generadas por las series reales, las misma que se pueden realizar por medio de modelos lineales de series de tiempo, que buscan describir los posibles modelos teóricos los cuales permitan explicar y detectar el comportamiento de la serie y de su proceso generador, tomando en cuenta la dependencia que existe entre los datos (Parra, 2019). Como principal característica es examinar el modelo más adecuado analizando errores recientes de pronóstico para seleccionar el ajuste apropiado, cumpliendo la condición de estacionariedad.

Los procesos aleatorios estacionarios lineales se clasifican en tres tipos: modelos Autorregresivos (*AR*), modelos de Medias Móviles (*MA*) y modelos Mixtos *ARMA*, en los modelos con procesos no estacionarios se encuentran los modelos *ARIMA* y los modelos estacionales *ARIMA* (Court Monteverde & williams Rengifo, 2011).

2.2.6.1 Proceso Autorregresivo *AR*(*q*)

Es un proceso estacionario que representan los valores de una variable en un instante del tiempo (*t*) en función a valores anteriores y_{t-1} , que explican el comportamiento temporal de una variable aleatoria (y_t), se representa como la suma de observaciones pasadas de la variable, donde (*p*) es el número de rezagos que determina el orden del modelo.

La forma general de un proceso Autorregresivo se presenta por:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (13)$$

Donde:

$$\begin{aligned} Y_t &= \text{Variable de respuesta} \\ Y_{t-1}, Y_{t-2}, \dots, Y_{t-p} &= \text{Periodos anteriores} \\ \emptyset_0, \emptyset_1, \emptyset_2, \dots, \emptyset_p &= \text{Parámetros que serán estimados} \\ \varepsilon_t &= \text{Termino residual o error} \end{aligned}$$

2.2.6.2 Proceso De Medias Móviles $MA(p)$

Es un modelo generado por un proceso de ruido blanco, que toma en cuenta la parte estacionaria del modelo, supone que una variable puede ser explicada por un cierto número de rezagos (q) de los errores, donde (q) determina el orden del modelo.

La forma general del modelo está representada por:

$$Y_t = \mu + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \dots - \omega_q \varepsilon_{t-q} \quad (14)$$

Donde

$$\begin{aligned} Y_t &= \text{Variable independiente en el tiempo } t \text{ (dependiente)} \\ \mu &= \text{Promedio} \\ \omega_1, \omega_2, \dots, \omega_q &= \text{coeficientes} \\ \varepsilon_t &= \text{Error} \\ \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q} &= \text{Error en periodos anteriores} \end{aligned}$$

2.2.6.3 Proceso Autorregresivo De Promedios Móviles $ARMA(p,q)$.

Los modelos $ARMA$, es la unión de los modelos Autorregresivos $AR(p)$ y modelos de Media Móvil $MA(q)$, como resultado de esta combinación se obtiene un modelo “mixto” denominado “Procesos Autorregresivos De Medias Móviles” abreviado en sus siglas $ARMA(p, q)$, donde (p) es el orden de la parte Autorregresivo y (q) es el orden de la Media Móvil.

Un modelo $ARMA(p, q)$ se representa en su forma general como:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \dots - \omega_q \varepsilon_{t-q} \quad (15)$$

Con los modelos $AMA(p, q)$ es posible explicar una gran variedad de comportamientos de series estacionarias, que deben ser determinados por los patrones de autocorrelaciones simples y autocorrelaciones parciales de la muestra. Los procesos que genera los modelos $ARMA(p, q)$ dependen de valores actuales y pasados de la respuesta (y).

2.2.6.4 Modelo Autorregresivo Integrado De Medias Móviles $ARIMA(p, d, q)$.

Los modelos $ARIMA$ de orden (p, d, q) se construye a partir de los modelos $ARMA(p, q)$, ambos modelos son la combinación de los procesos Autorregresivos y de Medias Móviles, con la diferencia que al modelo $ARMA(p, q)$ se le incluye un componente de estacionariedad denominada Integración $I(d)$, obteniendo como resultado el Modelo Autorregresivo Integrado De Medias Móviles conocido en su forma abreviada como $ARIMA(p, d, q)$ (Hernandez, 2014), donde (p) es el orden Autorregresivo, (d) el número de veces que ha sido diferenciada la serie y (q) el orden de las Medias Móviles (Gujarati & Porter, 2010).

El componente de Integración $I(d)$ que se incluye en el modelo $ARMA$ viene a ser la diferenciación regular (d) del número de veces que sufre una “serie original no estacionaria” para convertirla en “estacionaria” (eliminación de tendencia).

Para determinar el orden en que debe ser diferenciada una serie no estacionaria hasta alcanzar la estacionariedad, existe dos test de detección que logran determinar el número de diferencias para la serie: el “Test Dickey Fuller ó Test Dick Fuller Aumentado (DF o ADF) y el Test de Phillips-Perron (PP)” (Ducos Ramiez, 2009).

La expresión general del modelo $ARIMA(p, d, q)$ está dada por:

$$y'_t = \varphi_1 y'_{t-1} + \varphi_2 y'_{t-2} + \varphi_3 y'_{t-3} + \dots + \varphi_p y'_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (16)$$

En donde (y'_t) es el número de (d) diferencias que se aplica a la series original para producir una series estacionaria (Ducos Ramiez, 2009).

2.2.6.5 Modelos Estacionales Autorregresivos De Medias Móviles Integrados $ARIMA(p, d, q) \times (P, D, Q)_S$

Los modelos multiplicativos $ARIMA$, resulta ser la combinación de términos ordinarios de los procesos $ARMA$ y términos estacionales, de igual forma se incorpora las diferencias regulares y diferencias estacionales, la parte regular o de tendencia en la estructura $ARIMA(p, d, q)$ es encargada de modelar las dependencias regulares, que están asociadas a observaciones consecutivas, por otro lado la parte estacional $ARIMA(P, D, Q)_S$ es la encargada de modelar la dependencia estacional que están asociadas a observaciones separadas por (s) intervalos de tiempo (Villavicencio, 2010), donde $S = \text{intervalo de tiempo de la observación}$, estas pueden ser observaciones diarias, mensuales, anuales, semestrales, etc.(...), estas variaciones son captadas por el modelo multiplicativo estacional $ARIMA(p, d, q) \times (P, D, Q)_S$, dentro de los parámetros estacionales (Barahona Bonifaz, 2018).

Su estructura general del modelo multiplicativo estacional $ARIMA(p, d, q) \times (P, D, Q)_S$ esta dada por la ecuación (17).

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \theta_1 Y_{t-s} + \dots + \theta_p Y_{t-ps} + \varepsilon_t - \varphi_1 \varepsilon_{t-1} - \dots - \varphi_p \varepsilon_{t-p} - \vartheta_1 \varepsilon_{t-s} - \dots - \vartheta_Q \varepsilon_{t-Qs} \quad (17)$$



Donde:

$p =$ orden AR no estacional

$d =$ orden diferencias no estacional

$q =$ orden MA no estacional

$P =$ orden AR estacional

$D =$ orden diferencia estacional

$Q =$ orden MA estacional

$s =$ tiempo del patrón estacional

El procedimiento para modelar este proceso mixto estacional, es el mismo que plantea la metodología de Box –Jenkins.

2.2.7 Fases De La Metodología Box-Jenkins

En la década de los 70' en el año de 1973 George Edward Pelham Box y Gwilym Mirion Jenkins publican su libro titulado “Análisis De Series Temporales Pronóstico y Control”, que años después se conocería como “Metodología de Box-Jenkins” o técnicamente conocida como “metodología *ARIMA*”, donde combinan los enfoques de modelos Autorregresivos y el modelo de Medias Móviles (estudiados inicialmente por Yule), desarrollando un proceso iterativo que consta de cuatro fases: Identificación del modelo, estimación de los parámetros o coeficientes , validación de los modelos seleccionados y predicción de valores, este proceso iterativo sirve como herramienta para la selección de un modelo que mejor se ajuste a la serie temporal obteniendo como resultado pronósticos con mayor grado de exactitud (Sanchez de Rivera, 2005).

2.2.7.1 Fase De Identificación

Consiste en la identificación de los posibles modelos *ARIMA*(p, d, q), para ello se realiza un análisis visual de la serie que está siendo sometida a estudio, para tal fin se dispone a construir un gráfico de serie temporal o también denominado grafico temporal



(Marcelo Barreto & Villareal Escate, 2021), con el objetivo de identificar el comportamiento de los datos si estos presentan tendencia(estacionalidad) u otro comportamiento, el análisis de normalidad es una herramienta estadística fundamental para analizar la normalidad de los datos de la serie que también se deberá emplear.

Otra técnica a hacer aplicada para determinar la presencia de tendencia, es la observación visual del comportamiento de retardos en las funciones de autocorrelación simple(*ACF*) y parcial(*PAC*), cabe señalar que esta técnica se aplica de igual forma para determinar el orden de los modelos *ARIMA*(p, d, q)(Hernandez, 2014); de no existir estacionariedad en la serie original este deberá ser transformada a una serie estacionaria aplicando métodos de diferenciación, transformación de Box Cox y/o transformación logarítmica, para comprobar la no estacionariedad de la serie se hace uso del estadístico de prueba de Q de Ljung-Box o el test de Box-Pierce, ambas pruebas buscaran determinar analíticamente la existencia de estacionariedad o no.

Análisis de normalidad. Tiene por objetivo analizar y comprobar a través de contrastes de hipótesis(test de Shapiro-wilk, test de Kolmogorov-Smirnov) y gráficas (histogramas, gráficos Q-Q), si los valores de la serie siguen o no una distribución normal, caso contrario obtendremos estimadores que no serán eficientes, y los parámetros de los modelos no serán exactos(Amat Rodrigo, 2016).

Función de Autocorrelación simple (ACF). La Función De Autocorrelación Simple (*ACF*) es una herramienta fundamental para describir un proceso estocástico con fines de modelado de series de tiempo, además prueba si una serie es estacionaria, también se aplica para indicar el número de veces que debe ser diferenciado un proceso no estacionario, “se aplica como indicativo para determinar el orden de los procesos de Media Móvil *MA*(p)”(Pindyck & Rubinfeld, 2018, p.567).

Esta función indica cuanta de correlación o interdependencia existe entre los datos de la serie $y_t y_{t-1} \Gamma_2 y_t y_{t-2}$, porque mide la relación lineal entre valores rezagados de una serie, la representación gráfica (correlograma) se puede observar en la **Figura 2**, y está representada por la ecuación (18).

$$p_k = \frac{cov(y_t, y_{t-k})}{\sqrt{var(y_t)} \sqrt{var(y_{t-k})}} \quad (18)$$

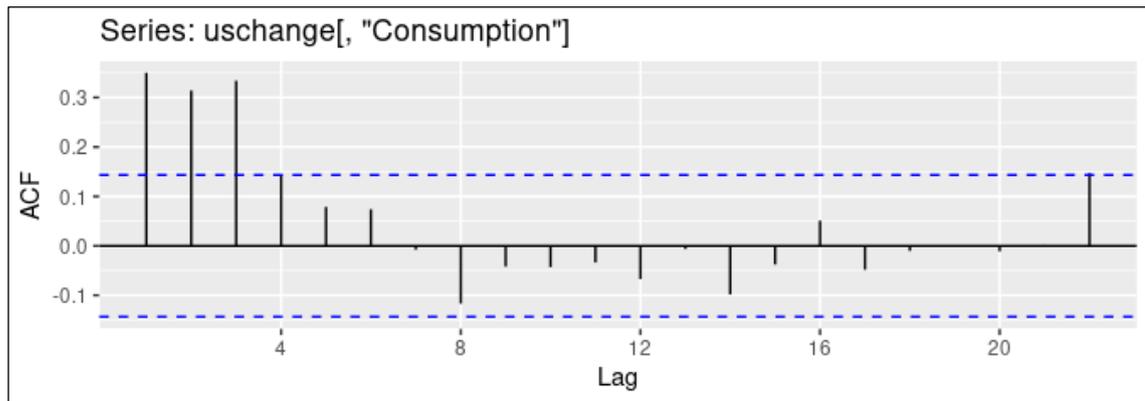


Figura 2. ACF Para El Cambio Porcentual Trimestral en el Consumo de EE.UU

Nota. Correlograma Simple de la serie cambio porcentual trimestral de consumo de los EE.UU. Fuente:(J Hyndman & Athanasopoulos, 2018)

La Función de Autocorrelación Parcial(PACF). Mide la relación entre dos variables independientes del proceso en diferentes periodos de tiempo y_t Y y_{tk} luego de haber eliminado los efectos de retrasos $1,2,3, \dots, k - 1$ (J Hyndman & Athanasopoulos, 2018). Ambas autocorrelaciones son idénticas una de la otra, el objetivo de la (PAC) es identificar y determinar el orden de un proceso Autorregresivo AR(q) que debe incluirse en el modelo(parámetros) (Pindyck & Rubinfeld, 2018). Su representación gráfica véase en la **Figura 3**, y su ecuación (19).

$$\alpha(s, t) = \frac{cov(X_s - \hat{X}_s^{(s,t)}, X_t - \hat{X}_t^{(s,t)})}{\sqrt{varX_s - \hat{X}_s^{(s,t)}} \sqrt{varX_s - \hat{X}_t^{(s,t)}}} \quad (19)$$

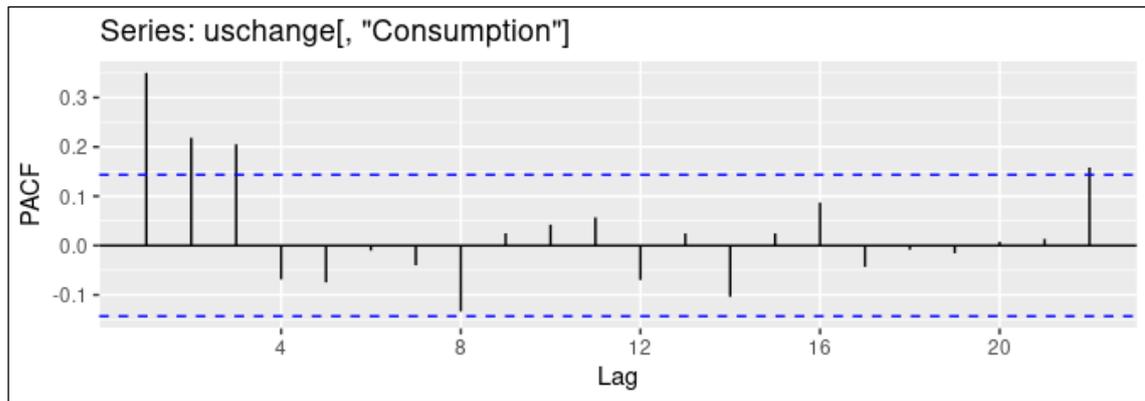


Figura 3. PAC Para el Cambio Porcentual Trimestral en el Consumo de EE. UU

Nota. Correlograma Parcial de la serie cambio porcentual trimestral de consumo de los EE.UU. Fuente: (J Hyndman & Athanasopoulos, 2018)

Prueba de Q de Ljung – Box. Q de Ljung-Box denominado así por los estadísticos Greta M. Ljung y George Edward Pelham Box, proponen un tipo de prueba estadística que se aplica para verificar la existencia de ruido blanco en una serie de tiempo, es decir “determina si una serie es o no un proceso de ruido blanco”(Víctor Morales, 2021), enfocando su análisis en los coeficientes de autocorrelación de residuos. Para ello se plantean las siguientes hipótesis:

H_0 : Los datos se distribuyen de forma independiente (ruido blanco).

H_a : No se distribuyen los datos de forma independiente (no son ruido blanco).

En el supuesto caso que la hipótesis nula sea rechazada, se tendrá que volver a identificar nuevamente el modelo (Pichua Serna, 2017).

La prueba de Q de Ljung-Box está representada por la ecuación (20)

$$Q = n(n + 2) \times \sum_{k=1}^h \frac{\hat{p}_k^2}{(n - k)} \sim \chi_k^2 \text{ gl. Donde: } r_j = \frac{\sum y_t y_{t-j}}{\sum y_t^2} \quad (20)$$

Donde n corresponde al tamaño de la muestra, \hat{p}_k simboliza la autocorrelación en el retraso k , y h expresa los retardos que están siendo probados (Víctor Morales, 2021).

Test de Box-Pearce. Desarrollado por Box y Pierce en (1970), el test indica la significancia estadística de los coeficientes de autocorrelación de forma conjunta, a fin de probar que los primeros rezagos (espigas) p_k sean iguales a cero, también se aplica para probar si los residuos son ruido blanco (Gujarati & Porter, 2010, p.753).

$$Q = n \times \sum_{k=1}^m \hat{p}_j^2 = \sum_{k=1}^m (\sqrt{n} \times \hat{p}_j)^2 \sim \chi^2(m) \quad (21)$$

En donde $n =$ tamaño de muestra y $m =$ longitud del rezago (Court Monteverde & Williams Rengifo, 2011, p.427).

2.2.7.2 Fase De Estimación

Identificado el modelo tentativo con los valores p, q se procede a estimar los parámetros, aplicando el procedimiento no lineal de mínimos cuadrados, “los parámetros se estiman minimizando la suma de cuadrados de los errores de ajuste” (Hanke & Wichern, 2018), donde el principal objetivo es evaluar y comparar modelos diferentes además de calcular los límites de error de pronóstico, determinada las estimaciones se considerará como significativos a los parámetros diferentes de cero que son conservados en el modelo ajustado, el resto que no son significativos se eliminan.

Error cuadrático medio de residuos se define como:

$$S^2 = \frac{\sum_{t=1}^n e_t^2}{n-r} = \frac{\sum_{t=1}^n (Y - \hat{Y}_t)^2}{n-r} \quad (22)$$

Donde, $e_t = Y_t - \hat{Y}_t$ residuo para el tiempo t $n =$ es el número de residuos y $r =$ número total de parámetros estimados.

Box – Jenkins además resalta la importancia del principio de parsimonia, es decir al momento de estimar el modelo, se deberá optar por aquel que contiene el menor número



de coeficientes. Un modelo con estas características realiza mejores pronósticos, con lo cual el modelo seleccionado explicará el comportamiento de la serie de forma adecuada y práctica(Gujarati & Porter, 2010).

2.2.7.3 Fase De Validación

Consiste en verificar el modelo que ha sido seleccionado en la fase de estimación, con el objetivo de que el modelo sea el más apropiado para representar la serie, para tal fin se aplica diferentes pruebas estadísticas para demostrar que se ajusta correctamente a la serie.

Análisis de residuos. Se analiza las funciones de autocorrelación simple (*ACF*) y parcial (*PACF*), como resultado se obtiene un diagnóstico de los residuos, que deben aproximarse al comportamiento de “ruido blanco” (media constante e igual a cero y varianza constante) (Jimenez, 1985), para ello se toma en cuenta los estadísticos de prueba de Q de Ljung-Box(1976) o Box-Pierce(1970) que estadísticamente determinará si los residuos de la serie de tiempo tiene un comportamiento de “ruido blanco”(puramente aleatorio) (Hanke & Wichern, 2018, p.410).

Análisis de bondad del ajuste. Haciendo uso del Coeficiente de Determinación ó el Coeficiente De Determinación Corregido se analiza la adecuación que existe entre la serie original y la serie estimada, en cuanto los coeficientes sean más próximos a 1 el modelo seleccionado será el mejor.

En el caso de seleccionar un modelo entre varios modelos candidato, se aplicará dos criterios: el Criterio De Información De Akaike (*AIC*), que es una herramienta estadística para determinar el orden de un modelo *ARIMA*, propuesto por Akaike en (1974), este criterio selecciona un mejor modelo entre muchos candidatos, considerando

el mejor ajuste considerando el menor número de parámetros (Hanke & Wichern, 2018).

Se define por la siguiente ecuación (23).

$$AIC = \ln \hat{\sigma}^2 + \frac{2}{n} r \quad (23)$$

Donde

\ln = *Logaritmo natural*

$\hat{\sigma}^2$ = *Suma residual de cuadrados dividida entre número de observaciones*

n = *Número de observaciones(residuos)*

r = *Número total de parámetros en el modelo ARIMA*

El indicador de criterio de Información De Bayes (*BIC*), propuesto por Schwarz en (1978), se emplea para la comparación del desempeño de un pronóstico dentro y fuera de la muestra de un modelo, este criterio de información supone que “mientras más pequeño sea el valor de BIC, mejor será el modelo para el pronóstico”(Mauricio, 2007,p.107).Está definida por la siguiente ecuación (24).

$$BIC = \ln \hat{\sigma}^2 + \frac{\ln * n}{n} r \quad (24)$$

Donde

\ln = *Logaritmo natural*

$\hat{\sigma}^2$ = *Suma residual de cuadrados dividida entre número de observaciones*

n = *Número de observaciones(residuos)*

r = *Número total de parámetros en el modelo ARIMA*

Ambas pruebas seleccionará un solo modelo óptimo de varios modelos candidatos, teniendo en cuenta el valor mínimo obtenido de las pruebas de (AIC) ó (BIC) (Jimenez, 1985).

Test de Dickey y Fuller. El test mide la presencia significativa de tendencia en una serie temporal a través de un contraste de hipótesis, determinando si existe o no raíces unitarias en la serie (Víctor Morales, 2021), para ello se plantea dos hipótesis.

H_0 : Tendencia estocástica en la serie, $\delta = 0$

H_a : La serie no tiene tendencia estocástica, $\delta < 0$

Prueba de Dickey-Fuller aumentada(DFA). Es una prueba de raíz unitaria que mide la estacionariedad o no de la serie, tomando como referencia el $p - value$ (probabilidad de hipótesis tomada), si el valor es menor que el $p - value = 0.05$ este demuestra que la serie es estacionaria a un nivel de confianza del 95%, caso contrario es no estacionaria (Aragón & Parra Royón, 2017).

2.2.7.4 Fase De Predicción

Finalmente, una vez que el modelo seleccionado *ARIMA* haya superado la fase de validación, es factible elaborar pronósticos a periodos futuros sea uno ó varios, considerando las restricciones de niveles de confianza asociados a los intervalos de predicción, que suponen “mientras más larga sea el horizonte de tiempo del pronóstico, mayor será el intervalo de predicción(incertidumbre)” (Hanke & Wichern, 2018), así mismo el modelo seleccionado se considerará como predictor optimo a aquel que cumpla con el análisis de error, justificándose en las métricas de evaluación estadísticas como: la Raíz de la Media de los Errores al Cuadrado (RMSE), la Media Absoluta de errores (MAE) y La Media Absoluta De Errores Porcentuales (MAPE) (Pindyck & Rubinfeld, 2018,p.586), en la **Figura 4** se resume en un diagrama de flujo el proceso de cálculo de modelos con la metodología de Box y Jenkins.

Evaluación de los pronósticos. Se emplea para determinar, comparar y comprobar la efectividad del modelo frente a otros modelos utilizados, buscando el menor valor posible de estos indicadores (RMSE, MAE, MAPE) se comprueba la precisión del modelo, siendo el más relevante y factor decisivo el indicador de error de Mean Absolute Percentage Error (MAPE), que representa el cambio porcentual de los errores (Court Monteverde & Williams Rengifo, 2011, p.428).

Raíz media de los errores al cuadrado (RMSE-Mean Square Error)

$$RMSE = \sqrt{\frac{\sum_{j=t+1}^{t+r} (y_i - g_{t,j})^2}{r}} \quad (25)$$

Media Absoluta De Los Errores (MAE-Mean Absolute Error)

$$MAE = \frac{\sum_{j=t+1}^{t+r} |y_i - g_{t,j}|}{r} \quad (26)$$

Media Absoluta De Errores Porcentuales (MAPE-Mean Absolute Percentage Error)

$$MAPE = \frac{\sum_{j=t+1}^{t+r} \left| \frac{y_i - g_{t,j}}{g_{t,j}} \right|}{r} \times 100 \quad (27)$$

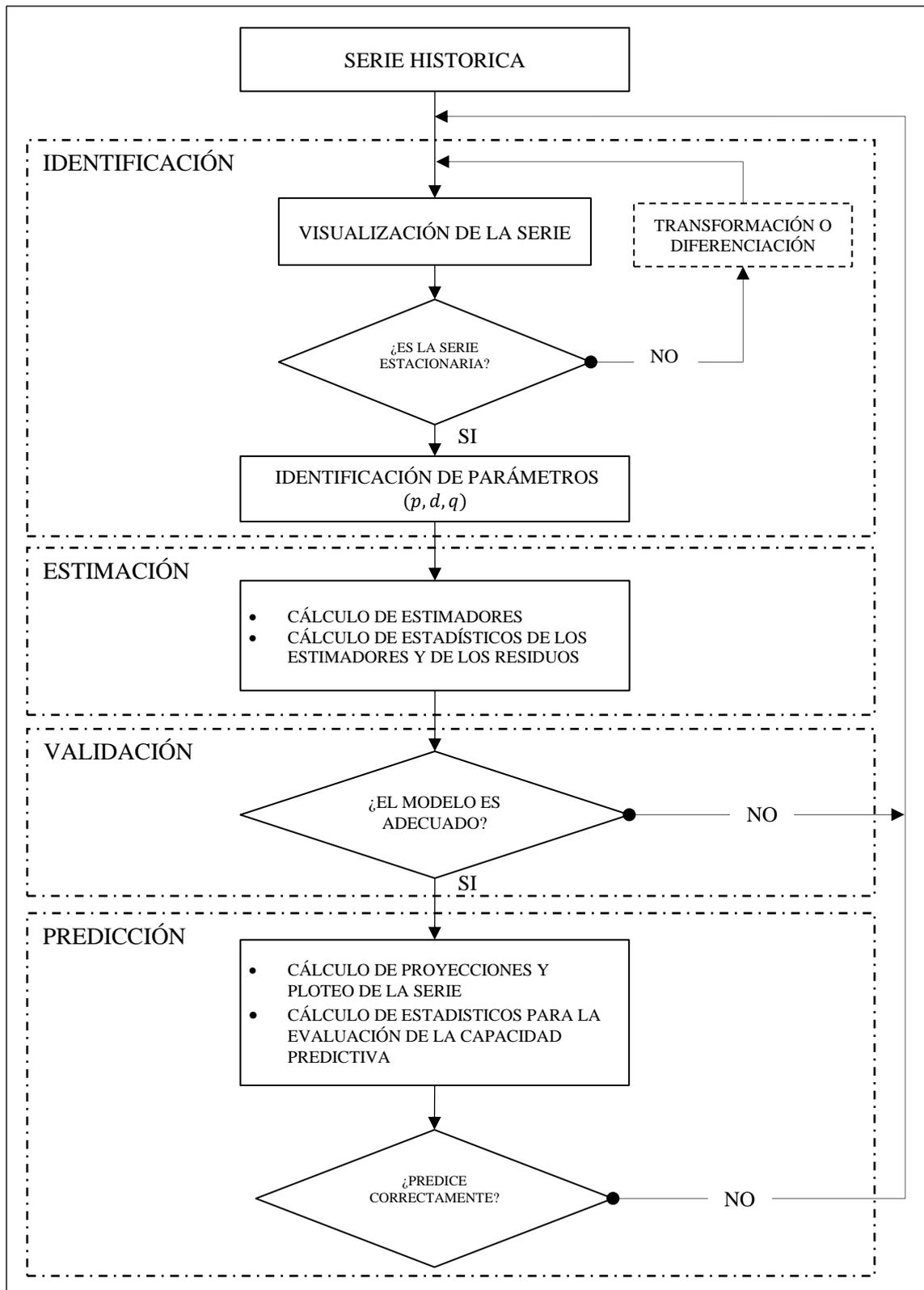


Figura 4. Proceso de Cálculo Mediante Modelos ARIMA

Nota. Diagrama de flujo de la metodología Box - Jenkins Fuente: Análisis de series de tiempo modelos ARIMA (Jimenez, 1985, p. 72).

2.3 MARCO CONCEPTUAL

2.3.1 Energía eléctrica

La energía eléctrica se define como “el movimiento de electrones que se trasladan por un conductor eléctrico durante un determinado periodo, la fuerza física que induce este movimiento se le denomina voltaje (V), mientras que el amperio(A) mide la intensidad de la corriente eléctrica” (OSINERGMIN, 2016, p. 28)

La energía eléctrica es producida en las centrales eléctricas, eólicas, hidroeléctricas (...), desde ahí es transportada a través de redes eléctricas hacia las empresas distribuidoras de energía eléctrica, que finalmente distribuirán la energía hacia los usuarios finales (clientes), en el Perú el sector eléctrico está estructurado y basa su funcionamiento de la siguiente forma, véase la **Figura 5**.

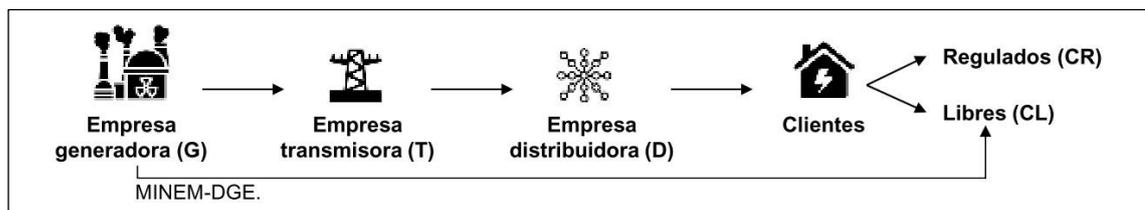


Figura 5. Funcionamiento Del Sistema Eléctrico En El Perú
Fuente: Ministerio de Energía y Minas (2020)

2.3.1.1 Empresas Generadoras(G)

La generación de energía eléctrica es la primera actividad en la cadena productiva de la industria eléctrica, que se encarga de transformar la energía primaria (solar, térmica, hidráulica, nuclear, eólica) en energía eléctrica, a través de centrales eléctricas.

En el Perú Participan 54 empresas generadoras de energía eléctrica, que está conformada por centrales hidráulicas, solares, térmicas y eólicas (Ministerio de Energía y Minas, 2020).



2.3.1.2 Empresas transmisoras(T)

Este segmento permite el transporte de la electricidad en grandes cargas de energía desde los centros de generación hacia las empresas distribuidoras de energía eléctrica, a través de torres de transmisión, líneas de transmisión, entre otros componentes que participan en este proceso.

En este segmento de transporte de energía eléctrica en el Perú participan 19 empresas transmisoras, el cual poseen una longitud de 28,263 kilómetros de línea de transmisión tendido a lo largo de nuestro país (Ministerio de Energía y Minas, 2020).

2.3.1.3 Empresas distribuidoras(D)

Es la etapa final de la cadena productiva de la industria eléctrica, donde se hace posible que la energía llegue a los consumidores finales(clientes) desde las líneas de transporte de alta tensión.

En este segmento para la distribución de energía, en el Perú participan 23 empresas distribuidoras de energía eléctrica ubicadas en distintas regiones, de los cuales 10 son empresas públicas, dentro de ellas se encuentra la empresa ELECTRO PUNO S.A.A ubicada en la región de Puno (Ministerio de Energía y Minas, 2020).

2.3.1.4 Usuarios o clientes (CR y CL)

Es el consumidor final de energía, está dividido según la demanda de potencia entre clientes libres y clientes regulados (OSINERGMIN, 2016).

2.3.2 Instituciones involucradas en el sector eléctrico

2.3.2.1 Ministerio De Energía Y Minas

Es el ministerio que conforma el poder ejecutivo encargado de legislar en temas relacionados en el sector energético y minero del Perú (OSINERGMIN, 2016).



2.3.2.2 *OSINERMIN*

Es una institución adscrita a la presidencia de consejo de ministros del Perú, se encarga de la regulación y supervisión de las empresas del sector eléctrico, hidrocarburos y minería.

2.3.2.3 *COES*

El COES, es una entidad privada conformada por todos los agentes del SEIN (generadoras, transmisoras, distribuidoras y usuarios Libres), administra el adecuado equilibrio entre la demanda y la oferta de energía.

2.3.3 **Horizonte de tiempo**

Todo pronóstico está asociado a un horizonte de tiempo, este puede ser de corto, mediano, y de largo plazo (Ríos, 2008). Para el pronóstico de valores futuros, es importante definir el tiempo de alcance a pronosticar, porque con ello la organización define la visión de los objetivos a alcanzar, mitigando los efectos negativos que puedan ocurrir durante su ejecución.

En la teoría estadística el pronóstico se sustenta en un análisis numérico, Ríos (2008) señala que cuando mayor es el alcance del pronóstico, mayor será el nivel de incertidumbre que se deberá enfrentar.

Corto plazo Tiene un tiempo de alcance hasta de 1 año, se utiliza para la programación de planta, asignaciones de trabajo, niveles de producción y compra, con el fin de diseñar estrategias inmediatas.

Mediano plazo El tiempo de alcance es desde 3 meses hasta 3 años, es utilizado en la evaluación y control de resultados de la gestión y de los negocios de la empresa, esto para satisfacer las necesidades de un futuro inmediato.



Largo plazo Su tiempo de alcance se considera desde los 3 años a más, estos pronósticos señalan el curso general de una organización para un tiempo de funcionamiento a largo plazo como la planificación de la producción y evaluación de proyectos.

Por otra parte Cervan (2017), en el sector eléctrico para ejecutar una proyección de la demanda eléctrica, se conoce a los horizontes de tiempo con el nombre de “Horizontes de planificación”, y lo clasifica en tres grupos.

Pronóstico de carga a corto plazo en ingles **Short-Term Load Forecasting (STLF)**. El tiempo de proyección va desde minutos, horas, días y una semana, es empleada para predecir la demanda de energía del día a la semana siguiente, teniendo en cuenta factores como: la demanda histórica, los precios de electricidad y condiciones meteorológicas.

Previsión de carga a mediano plazo en ingles **Medium-Term Load Forecasting (MTLF)**. El periodo de proyección como mínimo es una semana hasta 2 años, tiene como objetivo **la programación de mantenimiento**, la coordinación de los despachos de carga y liquidación de precios para que la demanda y la generación estén equilibrados, adicionalmente es empleada para pronosticar los picos de demanda de electricidad en las siguientes semanas y meses.

Previsión de carga a largo plazo en ingles **Long-Term Load Forecasting (LTLF)**, tiene un periodo de proyección que varía entre 10 a 20 años, el objetivo es la expansión del sistema eléctrico, generación, transmisión y distribución, dentro de las cuales también se considera la inversión para la adquisición de nuevas unidades de generación.



En ambos conceptos basan el modelado de datos aplicando las técnicas de pronóstico cuantitativas.

2.4 HIPÓTESIS DE LA INVESTIGACIÓN

2.4.1 Hipótesis General

- El modelo ARIMA proporciona un mejor ajuste para pronosticar el consumo de energía eléctrica de la Región de Puno.

2.4.2 Hipótesis Especifica

- El comportamiento de la serie Consumo de Energía Eléctrica de la Región Puno presenta una tendencia creciente y positiva con variaciones cíclicas y estacionales.
- Los modelos integrados aplicados para pronosticar, se ajustan mejor a la serie de Consumo de Energía Eléctrica de la Región Puno.
- El consumo mensual de energía eléctrica en la Región de Puno se incrementará en los años 2021 a 2023.



CAPITULO III

MATERIALES Y MÉTODOS

3.1 UBICACIÓN GEOGRAFICA DE LA INVESTIGACIÓN

3.1.1 Localización de la investigación

Geográficamente la investigación está ubicada en la Región de Puno al Suroeste de la República del Perú; situada en la meseta del Collao a una altitud de 3,827 *m. s. n. m.*, cuenta con una extensión territorial de 71,999.00 *Km²* (representa el 6,0 % del territorio nacional) siendo la quinta Región más grande de nuestro territorio Nacional.

La región de Puno, esta subdividida en 13 Provincias y 110 distritos, cuenta con una población estimada al año 2017 de 1,172,697 habitantes de acuerdo al Censo Nacional de Población y Vivienda 2017 (Fuente INEI, 2018, p. 23), ubicándose en el noveno lugar del total de número de habitantes a nivel nacional, cuyas tres provincias que más destacan a nivel poblacional es San Román (capital de Juliaca, principal centro comercial e industrial de la Región) que concentra el 26.2% de habitantes de la Región, en segundo lugar se ubica la provincia de Puno (Capital de la región de Puno) con el 18.7%, luego se ubica la provincia de Azángaro con un 9.4% de habitantes y el resto de provincias con un total de 45.7 por ciento de habitantes.

Puno tiene un clima variante, en cuanto mayor sea la altura el clima será frío y glacial; mientras que en la selva puneña el clima es cálido con precipitaciones pluviales, considerando estos factores de ubicación geográfica y condiciones climatológicas, la agricultura, ganadería, caza y silvicultura, son las principales actividades económicas de la región de Puno con una contribución del 18.5% al Valor Agregado Bruto (VAB) (BCRP, 2022).



3.1.2 Ámbito de estudio

La “Empresa Regional de Servicio Público de Electricidad de Puno-Sociedad Anónima Abierta” o en su defecto ELECTRO PUNO S.A.A., es una empresa concesionaria de distribución de energía eléctrica en la Región Puno, fue creada y aprobada mediante Junta General de Accionistas un 9 de Julio del año 1999, iniciando sus operaciones el 01 de noviembre de 1999, teniendo como objeto social prestar servicio de distribución y comercialización de energía eléctrica dentro de su área de concesión otorgado por el estado peruano (S.A.A, 2020).

La empresa basa su funcionamiento en amparo a la Ley de Concesiones Eléctricas D.L N°25844 y su reglamento D.S 009-93-EM; con fecha 7 de diciembre del año 2000 el ministerio de energía y minas mediante Resolución Suprema N°106-2000-EM y sus modificaciones, otorga a la empresa ELECTRO PUNO S.A.A. el área de concesión para el desarrollo de sus actividades dentro de la región de Puno, en donde se ubican sus instalaciones electromecánicas de distribución, transmisión, y también de generación, a cargo de la operación de la Central Hidroeléctrica de Chijisia localizada en la Provincia de Sandia (Electro Puno S.A.A, 2021).

Para garantizar el suministro de energía eléctrica a sus 403,141 clientes (al cierre del periodo diciembre 2020), ELECTRO PUNO S.A.A., realiza la compra de energía eléctrica en (Kv) para sus 10 sistemas eléctricos **Tabla 2**, a las empresas de generación eléctrica: SAN GABÁN, ELECTRO PERÚ, EDEGEL, SNPOWER, TERMOCHILCA, EEPSA, CHINANGO, EGASA, ENERSUR, STATKRAFT - CHEVES, EGEMSA, ENERSUR, ELECTRO PERÚ - PRO INVERSIÓN y complementando con la producción propia de la Central Hidroeléctrica de Chijisia, compras de energía que son captados a través de las Sub Estaciones Eléctricas de Transmisión a cargo de ELECTRO

PUNO S.S.A., ubicadas en las provincias de Azángaro, Ayaviri, Juliaca y Puno **Tabla 1** (S.A.A, 2020, p. 21).

En la **Tabla 1** se aprecia la barra de compra de energía eléctrica, hasta el cierre del periodo 2020 la compra fue de 396,627 *MWh*, obteniendo un incremento de compra del 13,347 *MWh* que representa un aumento de 3.48%, en comparación a años anteriores (S.A.A, 2020).

Tabla 1

Subestaciones Eléctricas de Transmisión

Compra de energía en (MWh)	2014	2015	2016	2017	2018	2019	2020
Juliaca	148,109	155,165	158,188	158,695	164,255	167,719	168,791
Puno	99,454	103,384	107,171	108,607	110,292	111,450	116,042
Azángaro	61,831	64,543	78,722	83,313	80,155	89,976	95,433
Ayaviri	9,427	10,990	11,349	10,551	10,839	11,173	13,020
San Gabán	2,201	2,395	2,884	3,115	3,035	2,962	3,340
TOTAL	321,022	336,477	358,314	364,281	368,576	383,280	396,627

Nota. Compra de energía a las empresas generadoras que pertenecen al COES y SEIN, captadas por las Sub Estaciones Eléctricas de Transmisión. Fuente: (S.A.A, 2020)

OSINERMIN mediante la Resolución OSINERMIN N°075-2018-OS/CD establece y determina el alcance de 10 sistemas eléctricos **Tabla 2** divididas en dos sectores: Urbano y Rural (SER), que son distribuidos en distintas provincias de la región de Puno, de donde ELECTRO PUNO S.A.A. provee del suministro de energía eléctrica atendiendo a 403,141 clientes (al cierre del periodo 2020), adicionalmente cumple con brindar el alumbrado público a las distintas provincias de la región que se encuentren dentro de las zonas de concesión a cargo de la empresa.



Tabla 2

Sistemas Eléctricos de ELECTRO PUNO S.A.A

N°	Código	Sistema eléctrico	Sector típico	Área de influencia	Barra de compra
1	SE0025	JULIACA	2	Juliaca (5004, 5005, 5006, 5008, 5009)	Juliaca 10 kV
2	SE0238	JULIACA RURAL	4	Juliaca (5007, 5010, 5011, 5021, 5022, 5023, 5024, 5025)	Juliaca 22.9 kV
	SR0124	SER001 JULIACA	4	Huancané (4001, 4002, 4003)	
3	SE0026	PUNO	2	Puno (0102, 0103, 0104)	Puno 22.9 kV
4	SE0239	PUNO BAJA DENSIDAD	4	Puno (0101, 0105, 0106)	Puno 60 kV
	SR0125	SER002-PUNO	4	Totorani (0201, 0202)	
5	SE0030	ILAVE-POMATA	4	Ilave (1001, 1002, 1003) Pomata (2001, 2002, 2003, 2004)	
6	SE0027	AZÁNGARO	4	Azángaro (8002) Ananea (3001, 3002, 3003, 3004, 3005, 3006, 3007) Chijisia (7501, 7502)	Azángaro 22.9 kV
	SE0237	AZÁNGARO RURAL	4		
7	SR0122	SER003 AZÁNGARO	4	Azángaro (8001)	
8	SE0028	ANTAUTA	4		Azángaro 60 kV
	SR0135	SER006 ANTAUTA	4	Antauta (9001, 9002)	
9	SE0029	AYAVIRI	4		Ayaviri 10 kV
	SR0123	SER004 AYAVIRI	4	Ayaviri (6001, 6002, 6003, 6004)	Ayaviri 22.9 kV
10	SE0220	SAN GABÁN	4		San Gabán 13.8 kV
	SR0292	SER007 SAN GABÁN	4	San Gabán (9501)	

Nota. Información recogida de la memoria anual 2020 publicado por ELECTRO PUNO S.A.A. Fuente: ELECTRO PÚNO



3.2 POBLACIÓN Y MUESTRA DE LA INVESTIGACIÓN

3.2.1 Población de estudio

La población está constituida por los reportes de consumo mensual de energía eléctrica representados en (*MWh/mes*), que corresponde al área de concesión a cargo de la empresa de servicio público ELECTRO PUNO S.A. para la distribución y comercialización de energía eléctrica en la Región de Puno; considerando un periodo comercial de 7 años que comprende desde enero-2014 a diciembre-2020, obteniendo un total de 84 observaciones mensuales(S.A.A, 2020).

3.3 RECOLECCIÓN DE DATOS

Para la recolección de la información primeramente se realizó las coordinaciones y tramites respectivos, acto seguido se aplicó *la técnica de la entrevista*, formulando preguntas verbales al Ingeniero encargado del área de la Gerencia de Operaciones, con el fin de obtener mayor conocimiento acerca del funcionamiento técnico del sector eléctrico, del mismo modo se entrevistó al Ingeniero encargado de la facturación zona Sur del área de la Gerencia Comercial, obteniendo con más detalle información relevante concerniente al proceso de facturación de recibos de energía eléctrica.

Para la obtención de la data se aplicó la técnica de *revisión documental*, teniendo acceso al sistema (software) SIELSE COMERCIAL propiedad de ELECTRO-PUNO S.A.A, el software de acuerdo a nuestro requerimiento nos proporcionó los resúmenes de facturación que se encuentran registrados mensualmente por cada periodo comercial de facturación, garantizando de este modo la efectividad y la originalidad de los datos para nuestra investigación.

3.4 TIPO DE INVESTIGACIÓN

3.4.1 Diseño y tipo de investigación

Este estudio tiene un enfoque cuantitativo, cuya finalidad es “ explicar, predecir, verificar y controlar los fenómenos”(Gallardo Echenique, 2017, p. 22). que se define como un conjunto de procesos que son secuenciales y probatorios, que inicia con la concepción de una idea, y continua con el planteamiento de objetivos y preguntas de investigación que dará lugar a la formulación de hipótesis y determinación de las variables, además de revisar la teoría correspondiente para la construcción del marco teórico; Los datos son recolectados en función a procedimientos estandarizados aprobados por una comunidad científica, los mismos que serán puestos a prueba y analizados a través de métodos estadísticos apropiados que establecerán una serie de conclusiones a las hipótesis planteadas (Hernandez Sampieri et al., 2010).

La investigación es de diseño no experimental, porque se observará el fenómeno tal y como se muestran en su contexto natural sin la manipulación intencionada de las variables, para posteriormente analizarlos, el tipo de investigación no experimental es de diseño Longitudinal de tendencia (trend), este tipo de diseño se basa en la recolección de los datos a través del tiempo en puntos o periodos, al cual analizaremos los cambios que sufrieron en el tiempo dentro de la población, Hernandez Sampieri et al. (2010) “para realizar inferencias acerca de la evolución, causas y sus efectos” (p.158).

3.4.2 Metodología para la predicción de series de tiempo

Para el desarrollo de esta investigación de series de tiempo, se aplicó la metodología de Box-Jenkins, esta metodología se refiere a “un conjunto de procedimientos para identificar, estimar, validar los modelos *ARIMA* y pronosticar con el modelo determinado” (Hanke & Wichern, 2018), con el único objetivo de determinar



el mejor ajuste de una serie temporal de valores, a fin de que los pronósticos sean lo más acertado posible.

Según Hanke & Wichern (2018) la metodología de Box – Jenkins contempla cuatro fases:

Fase de Identificación. Como primera fase para la identificación de los parámetros y el orden de los modelos $ARIMA(p, d, q)$, consiste en identificar si la serie es estacionaria o no, para ello es útil graficar la serie temporal, conjuntamente con las gráficas de distribución normal y correlogramas de las funciones de autocorrelación simple (ACF) y parcial (PACF), que nos ayudaran a determinar y analizar visualmente si la serie histórica contiene un comportamiento aleatorio o no, posteriormente se aplicará pruebas estadísticas que determinará objetivamente el comportamiento de la serie, si la serie es “no estacionaria” se aplicará técnicas de transformación o de diferenciación a la serie original para convertirla en una serie estacionaria (Court Monteverde & Williams Rengifo, 2011).

Fase de Estimación. Seleccionado el modelo con los parámetros tentativos de (p, q) en la fase de identificación, se debe estimar los parámetros para el modelo, para ello se hará uso de un procedimiento no lineal de mínimos cuadrados simples o el procedimiento de máximo verosimilitud, donde se analizará si los parámetros son significativos, como resultado del procedimiento obtendremos una serie que contiene los residuos del ajuste del modelo más adecuado (Hanke & Wichern, 2018).

Fase de Validación. Antes de hacer uso del modelo seleccionado para fase de pronóstico, este deberá comprobarse qué tan adecuado es, sometiendo a prueba el modelo a las evaluaciones de: análisis de coeficientes o parámetros del modelo, evaluación de

bondad de ajuste y análisis de residuos(Aragón & Parra Royón, 2017). En esencia, un modelo es adecuado si los residuos son aleatorios (ruido blanco).

Fases de Predicción. Finalmente, una vez seleccionado el modelo adecuado, se evaluará el modelo en cuanto a precisión del modelo se refiere, con las métricas de medición de RMSE,MAE y MAPE, una vez logrado superar estas pruebas de evaluación es factible elaborar los pronósticos considerando los objetivos del estudio y el horizonte de tiempo a pronosticar, a partir de la serie original (Gujarati & Porter, 2010).

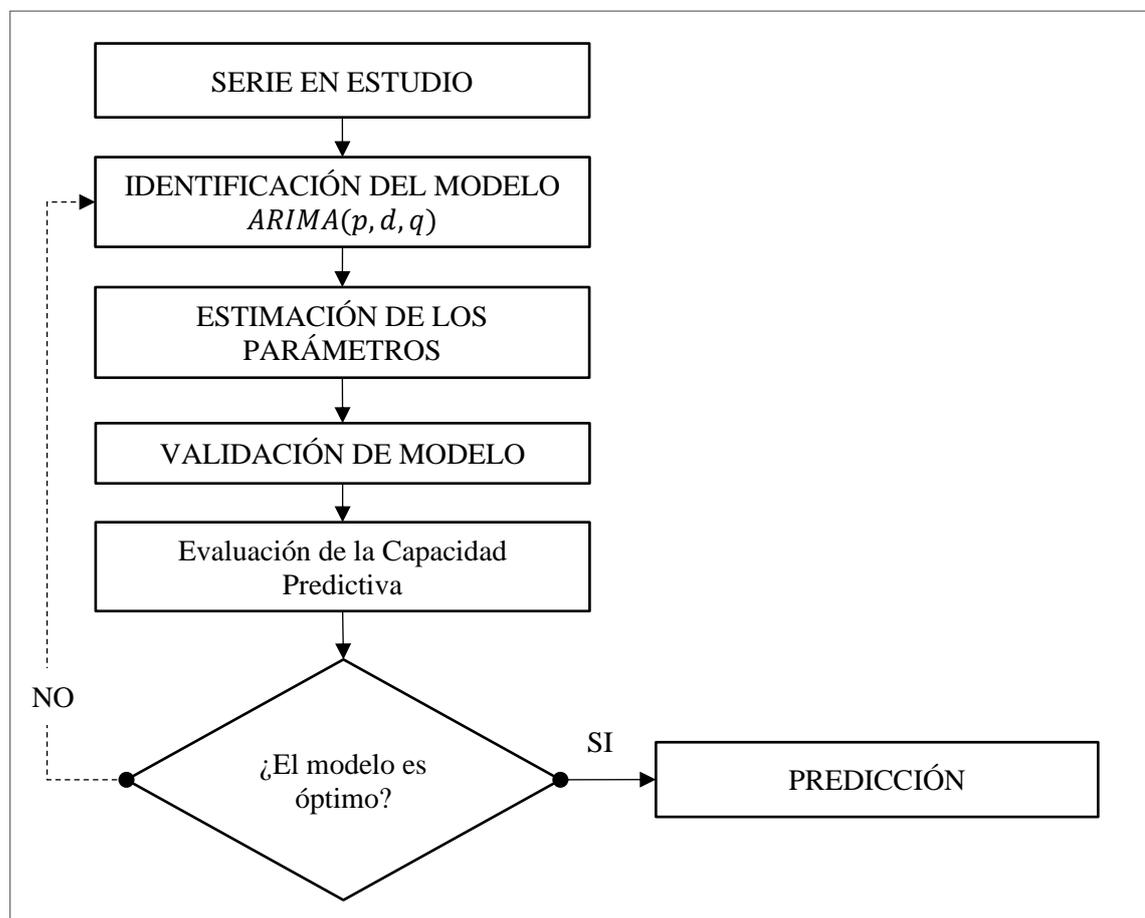


Figura 6. Diagrama de flujo metodología Box-Jenkins

Nota. Metodología de Box –Jenkins en un diagrama de flujo basado en (Jimenez, 1985) Fuente: Elaboración propia

3.4.3 Procesamiento de datos

Para el procesamiento y análisis de información de la serie de consumo de energía eléctrica, se cuenta con el software libre para análisis estadístico R versión 4.2.1 y R studio (actúa como entorno de desarrollo integrado (IDE) de R), R fue desarrollado en 1993 por Robert Gentleman y Ross Ihak en el Departamento de Estadística de la Universidad de Auckland en Nueva Zelanda, orientando su aplicación para el análisis estadístico y la visualización de gráficos (Wickham & Grolemond, 2017), este software en la actualidad es una herramienta muy potente para el procesamiento, manipulación, y análisis de grandes volúmenes de datos, extendiendo su uso como herramienta de análisis estadístico y estadística espacial en diferentes campos de investigación.

En nuestro caso, en el análisis de los datos de la serie consumo de energía eléctrica de la región Puno se utilizó las librerías (también llamadas paquetes) del software R para el pronóstico de consumo de energía eléctrica, véase el **Anexo H**, estas librerías son algoritmos desarrollados con el propósito de facilitar el trabajo del investigador al momento de aplicar y desarrollar diversos análisis estadísticos; siendo como principales librerías del software R a considerar en nuestro código(script) de las más de 12, las librerías: library (forecast) y library (tseries) ambos paquetes nos ayudarán a desarrollar y obtener los modelos que propone la metodología Box-Jenkins, y cómo principal motor de visualización de gráficos se usó la librería: library (ggplot).

En la **Figura 7** se muestra el proceso de análisis de información en el software R, básicamente inicia con la importación de los datos a la interfaz de R studio, continua con la limpieza de la data (eliminación de errores), seguido con la creación del objeto time-series con la función ts() para la manipulación de la información, posteriormente se analiza, aplica las pruebas estadísticas(modelar) y culmina con la comunicación de

resultados a través de la obtención de modelos con la función `auto.arima()` y gráficos con la función `ggplot()`.

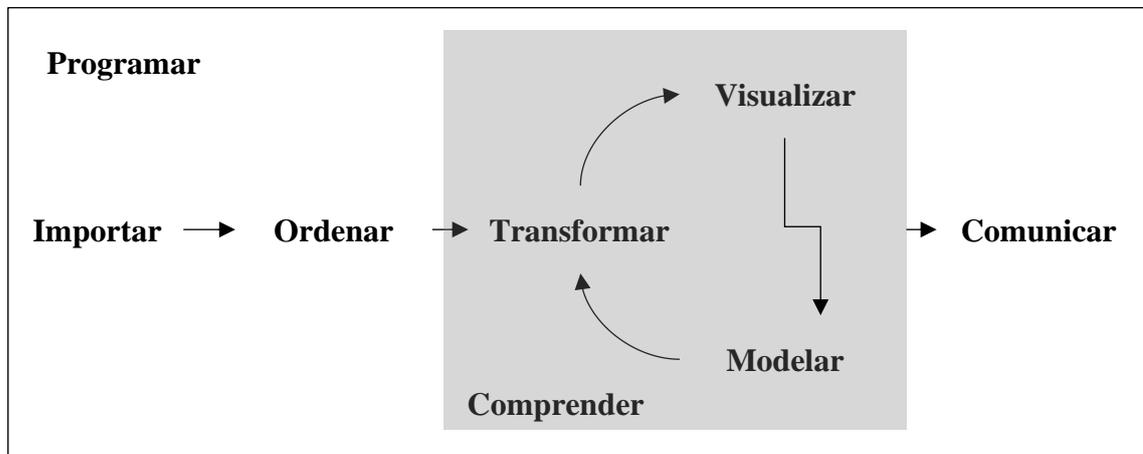


Figura 7. Diagrama Para El Procesamiento De La Información

Nota. Proceso para el análisis de información con el software R 4.2.1. Fuente: (Wickham & Golemund, 2017)



CAPITULO IV

RESULTADOS Y DISCUSIÓN

La Empresa Regional de Servicio Público de Electricidad de Puno Sociedad Anónima Abierta, conocido como “ELECTRO PUNO S.A.A”, es una empresa de distribución y comercialización de energía eléctrica, cuyo objetivo es de suministrar y garantizar el servicio de energía eléctrica a sus consumidores dentro de su área de concesión. Por tal razón con la finalidad de proveer el servicio de energía eléctrica a sus 403,141 clientes, la empresa cuenta con una central Hidroeléctrica ubicado en la Provincia de Sandía, así mismo tiene en operación 10 sistemas eléctricos distribuidos en toda su área de concesión de la región de Puno con el objetivo de garantizar el suministro de energía eléctrica.

La información que se muestra en la **Tabla 3** proviene de la base de datos de la empresa ELECTRO PUNO S.A.A, producto del consumo de energía eléctrica por parte de sus clientes, información que fue proporcionada por intermedio de la oficina de facturación – Puno; al cierre del periodo Diciembre 2020 la empresa ELECTRO PUNO S.A.A cuenta con 403,141 clientes, dentro de ellos se tiene a: clientes con tarifa residencial, clientes con tarifa comercial, clientes colectivos y clientes mayores, la información proporcionada del consumo energía eléctrica de la región de Puno se encuentra en la unidad de medida Megavatios por hora (MWh), consta de 84 registros mensuales de facturación que inicia desde enero 2014 y concluye en diciembre 2020, haciendo un total de 7 periodos comerciales(años).

Tabla 3*Consumo Mensual de Energía Eléctrica de la Región Puno 2014-2020*

Meses Años	2014	2015	2016	2017	2018	2019	2020
enero	21095	22285	22848	23627	23988	23793	24580
febrero	20942	21949	22186	24361	23144	24267	24468
marzo	18513	21441	21772	21084	22499	21659	22807
abril	21410	23165	23485	24222	24682	24810	21163
mayo	21643	24112	22888	23518	23841	23968	17639
junio	22841	23873	24323	25249	25628	25850	19225
julio	22394	23359	23787	24985	25141	25366	20316
agosto	22669	23130	23943	24144	24583	25400	22847
setiembre	22351	24061	24200	24505	24982	25788	22759
octubre	21594	23338	24154	24463	24134	24933	23911
noviembre	22746	23411	23592	24611	24519	24615	24465
diciembre	22021	22770	23974	24150	23916	24138	24973

Nota. Datos mensuales del consumo de energía eléctrica en Megavatios por hora (MWh) de la Región de Puno, período 2014 a 2020. Fuente: ELECTRO PUNO S.A.A.- Área de Facturación

4.1 RESULTADOS

4.1.1 Identificación

Analizaremos el comportamiento de la serie consumo de energía eléctrica, identificando a cuatro componentes de la serie (tendencia, ciclo, aleatoriedad y estacionalidad) utilizando métodos de descomposición; mediante un análisis de normalidad de datos, correlogramas y el test aumentado de Dickey – Fuller, se determinará la estacionariedad de la serie, de ser el caso se estabilizará la serie en media y varianza para la identificación de posibles modelos.

4.1.1.1 Primera etapa de la fase de identificación

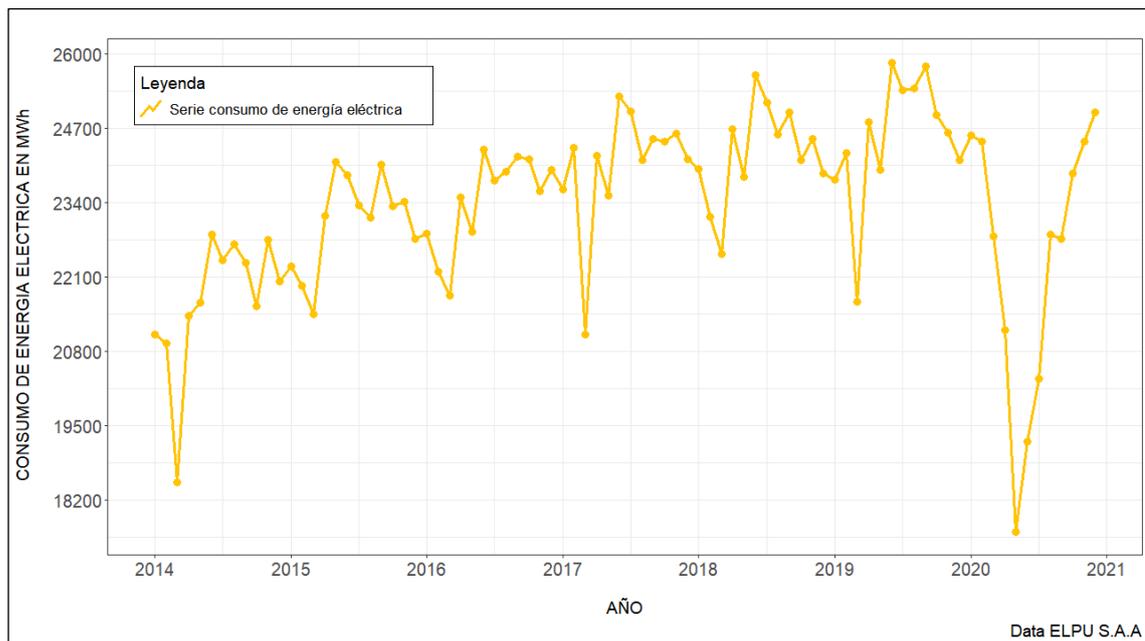


Figura 8. Consumo de Energía Eléctrica de la Región Puno 2014-2020

Nota. Periodograma de la serie consumo de energía eléctrica periodos 2014 a 2020, realizado con el software R versión 4.2.1. Fuente: Data ELP U S.A.A. – Elaboración propia del autor

Se presenta la **Figura 8** para analizar el comportamiento de la serie consumo de energía eléctrica en (MWh) de la región de Puno que inicia desde enero 2014 a diciembre 2020, en donde se observa un incremento de consumo de energía eléctrica durante 6 años de manera ininterrumpida, podemos afirmar que, según el comportamiento de los datos la serie presenta una tendencia creciente con variaciones cíclicas desde el año 2014 al 2019 tal como se muestra en la **Figura 8**, en el año 2020 específicamente en los meses de marzo a mayo existen picos alterados muy bajos, esto se debe a consecuencia del estado de emergencia sanitaria promulgado por el estado peruano a causa del COVID -19, en el mes de junio del mismo año se observa una ligera recuperación de acuerdo a los registros en la facturación de consumos de energía eléctrica hasta el mes de diciembre del año 2020.

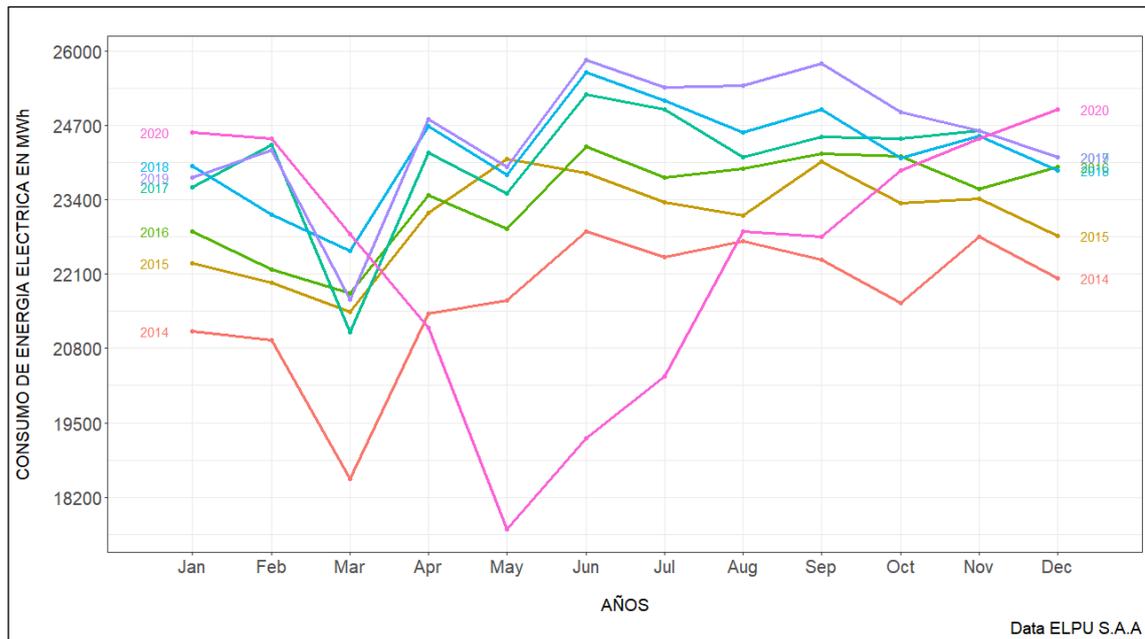


Figura 9. Comportamiento del Consumo de Energía Eléctrica por Meses y Año

Nota. Análisis mensual del comportamiento de consumo de energía eléctrica de la Región Puno durante 7 años desde el 2014 al 2020. Fuente: Elaboración propia del autor

Continuando con el análisis del comportamiento de la serie, se presenta la **Figura 9** donde se observa con más detalle el consumo mensual de energía eléctrica en Megavatios por hora (*MWh*) de la Región Puno , en el año 2020 la serie presenta una variación irregular desde el mes de Marzo hasta el mes de Mayo obteniendo un pico de consumo de facturación muy bajo, este factor imprevisible es provocado por la pandemia del COVID-19 ocurrido en el año 2020, como se recuerda el estado peruano emitió varias normas y decretos de urgencia durante el estado de emergencia sanitaria, ocasionando el cese de diferentes actividades económicas como: comerciales, educativas , sociales y culturales, en todas las regiones del Perú, en la región de Puno los principales sectores de consumo de energía eléctrica(industria, minería, comercio) el 16 de marzo en adelante paralizan sus actividades económicas, hecho que origina un decrecimiento en el consumo energía eléctrica en los meses de marzo, abril y mayo, tal como se aprecia en la **Figura 9**, en el mes de Junio del mismo año la serie presenta una recuperación de consumo de energía eléctrica, cerrando el mes diciembre del año 2020 con un consumo de 24,973(*MWh*) mayor a los 6 años anteriores del mismo mes.

El análisis sobre el comportamiento de la serie de consumo de energía eléctrica guarda relación con el trabajo realizado por Barahona Bonifaz (2018), quien refiere que “el consumo de energía eléctrica en la ciudad de Riobamba-Ecuador, presenta una tendencia creciente, además tiene un efecto estacional causando fluctuaciones con picos altos y bajos en el consumo de energía eléctrica, además la serie presenta un componente de aleatoriedad en los datos”.

Bajo la presentación de estos resultados afirmamos que el comportamiento de la serie consumo de energía eléctrica en ambos estudios presenta una tendencia creciente positiva, con variaciones cíclicas y estacionales, análisis que se corrobora a través de la descomposición de la serie con el método multiplicativo que se presenta en la **Figura 10** donde se puede identificar los 4 componentes de la serie de tiempo.

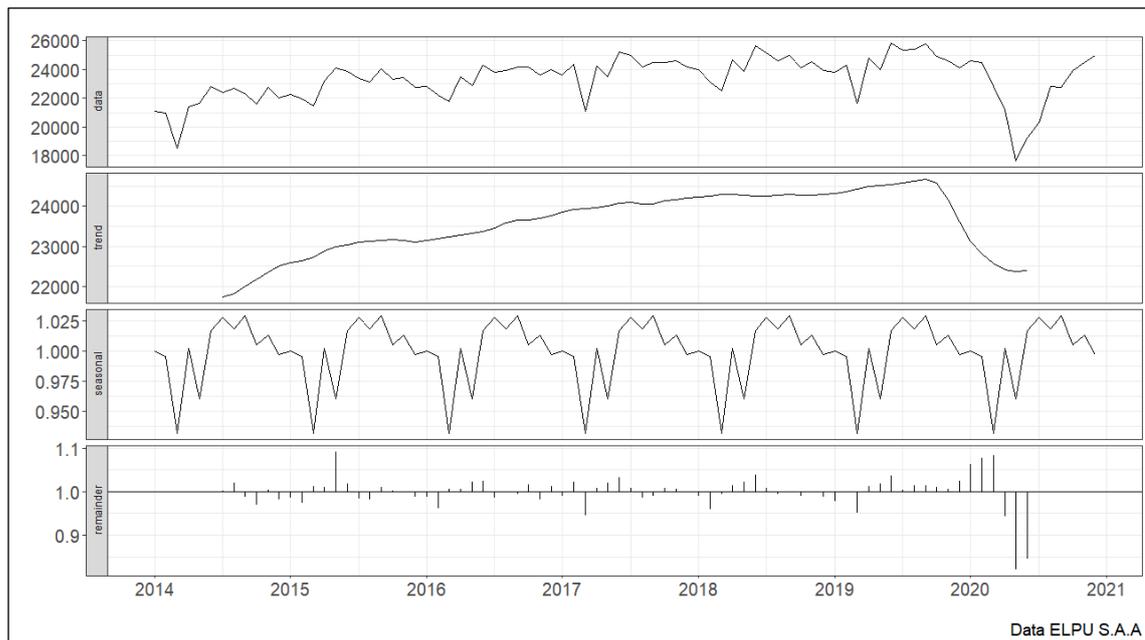


Figura 10. Descomposición de la Serie Consumo de Energía Eléctrica

Nota. Descomposición de la serie original con el método multiplicativo, realizado en el software R versión 4.2.1. Fuente: Elaboración propia del autor

De acuerdo a la metodología que propone Box – Jenkins, para aplicar los modelos *ARIMA*, como primer paso se debe determinar si la serie es “estacionaria”, es decir, la serie debe mantenerse constante a lo largo del tiempo con respecto a su media y varianza;

en caso la serie sea “no estacionaria” la metodología propone métodos de transformación como la diferenciación o la transformación de Box-Cox, para convertirla en una serie “estacionaria”.

Conforme a nuestro criterio de análisis al comportamiento de la serie en estudio, llegamos a afirmar que la serie consumo de energía eléctrica presenta una tendencia creciente positiva, con variaciones cíclicas y de aleatoriedad en los datos por lo tanto es una serie no estacionaria, tal como se muestra en la descomposición multiplicativa de la serie **Figura 10**, para determinar la estacionariedad de la serie antes de ser diferenciada en d o D número de veces, se analizará a través del histograma la normalidad de los datos conjuntamente con el test aumentado de Dickey – Fuller.

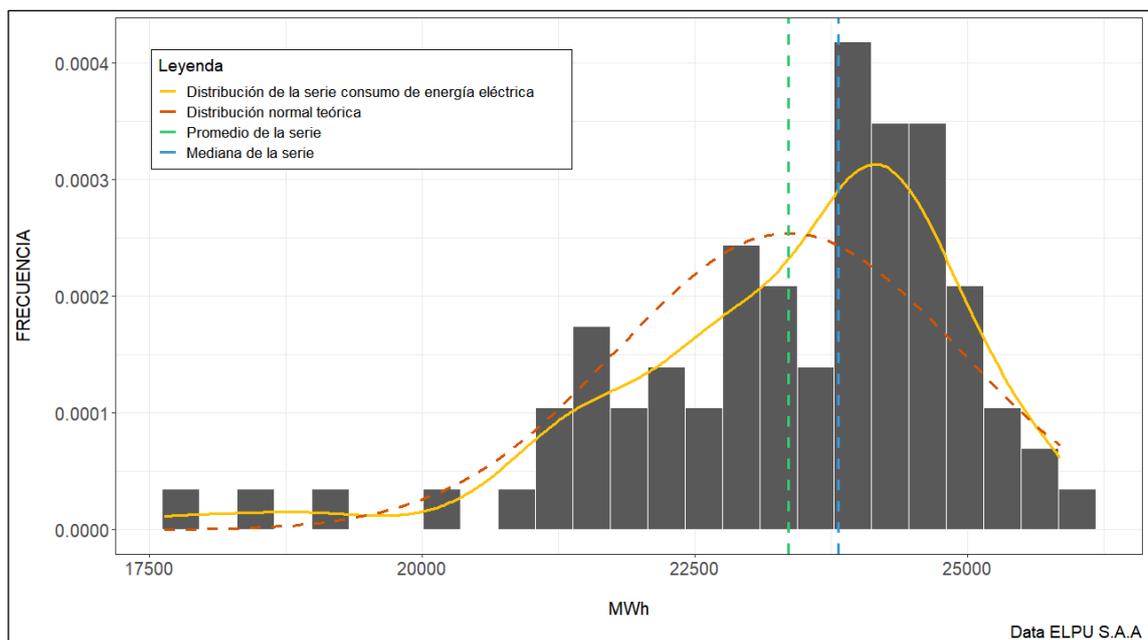


Figura 11. Histograma de la Serie Consumo De Energía Eléctrica De La Región Puno
Nota. Histograma de la serie consumo de energía eléctrica, frente al histograma teórico de la serie, realizado con el software R v.4.2.1. Fuente: Elaboración propia del autor

Se presenta la **Figura 11**, donde se observa la distribución de los valores individuales de la serie consumo de energía eléctrica; de la **Figura 11** señalamos que, durante los 7 periodos comerciales(años) el consumo máximo de energía en la región de Puno fue de $Max = 25,850.55 MWh$, seguido de un consumo mínimo de energía

eléctrica de $Min = 17,639.05 MWh$; el consumo promedio obtenido de la serie es de $\bar{\chi} = 23,356.93 MWh$ (Línea horizontal verde) con una desviación estándar de $\sigma = 1,570.98$ y una mediana de $\hat{\chi} = 23,817.16$ (Línea horizontal azul), de acuerdo a los datos presentados afirmamos que los valores de la serie sigue una distribución asimétrica negativa hacia la izquierda tal como se muestra en la **Figura 11**, en donde los valores se distribuyen hacia la izquierda con una tendencia variante hacia arriba y abajo de $1,570.98 MWh$ con respecto a su promedio, todo indica que los valores de la serie consumo de energía eléctrica no se distribuyen según una distribución normal, en otras palabras, los valores individuales de la serie no giran en torno a su promedio y varianza, por lo expuesto la serie es “no estacionaria”.

De lo mencionado en el párrafo anterior, la **Figura 11** indica visualmente con los datos proporcionados que la serie de consumo de energía eléctrica es “no estacionaria”, sin embargo, existe otros métodos estadísticos para determinar si la serie es estacionaria o no, entre varios métodos aplicaremos el Test Aumentado de Dickey – Fuller, el test comprobará a través de un contraste de hipótesis si la serie presenta una tendencia significativa.

A. Prueba

Test Aumentado de Dickey – Fuller

B. Formulación de hipótesis

H_0 : La serie consumo de energía eléctrica es no estacionaria.

H_α : La serie consumo de energía eléctrica es estacionaria.

C. Regla

Se acepta H_0 , Si, $p - value \geq 0.05$

Se rechaza H_0 , Si $p - value < 0.05$

D. Nivel de significancia

$$\alpha = 0.05$$

E. Prueba estadística

Tabla 4

Test Aumentado de Dickey – Fuller

statistic	p.value	parameter	method	alternative
-2.9	0.1917	4.0	Augmented Dickey-Fuller Test	stationary

Signif. codes: 0 <= '****' < 0.001 < '***' < 0.01 < '**' < 0.05

Nota. Prueba de hipótesis con el Test Aumentado de Dickey – Fuller. Fuente: R versión 4.2.1.

F. Resultado

$$0.1917 \geq 0.05$$

G. Decisión

A la luz de los datos con un nivel de significancia del 5% y tras el resultado obtenido a través del Test Aumentado de Dickey – Fuller, se dispone de evidencia suficiente para aceptar la hipótesis nula H_0 . Por lo tanto, la serie consumo de energía eléctrica es “no estacionaria”; de acuerdo a los resultados se confirma que la serie presenta estacionalidad, bajo estos resultados se deberá diferenciar o transformar la serie a una serie estacionaria.

Vázquez Díaz & Gamonal Sánchez (2019) en su estudio indican que la serie consumo de energía eléctrica de la provincia de Bagua Grande, presenta un comportamiento no estacionario, afirmando que los valores de la serie varían en un 45.8% con respecto a su promedio, para comprobar la estacionariedad de la serie los

investigadores aplican la prueba estadística de Levene, obteniendo como resultado un valor de significancia del $11.006 > 0.05$ concluyendo que la serie es no estacionaria, hallazgo que concuerda con nuestro estudio.

Con el objetivo de cumplir con la estacionariedad de la serie que exige la metodología *ARIMA* de Box y Jenkins, se procederá a graficar los correlogramas simples (*ACF*) y parciales (*PAC*) de la serie para identificar el número de veces que deberá ser diferenciada, en diferencias regulares o estacionales para convertirla a una serie estacionaria.

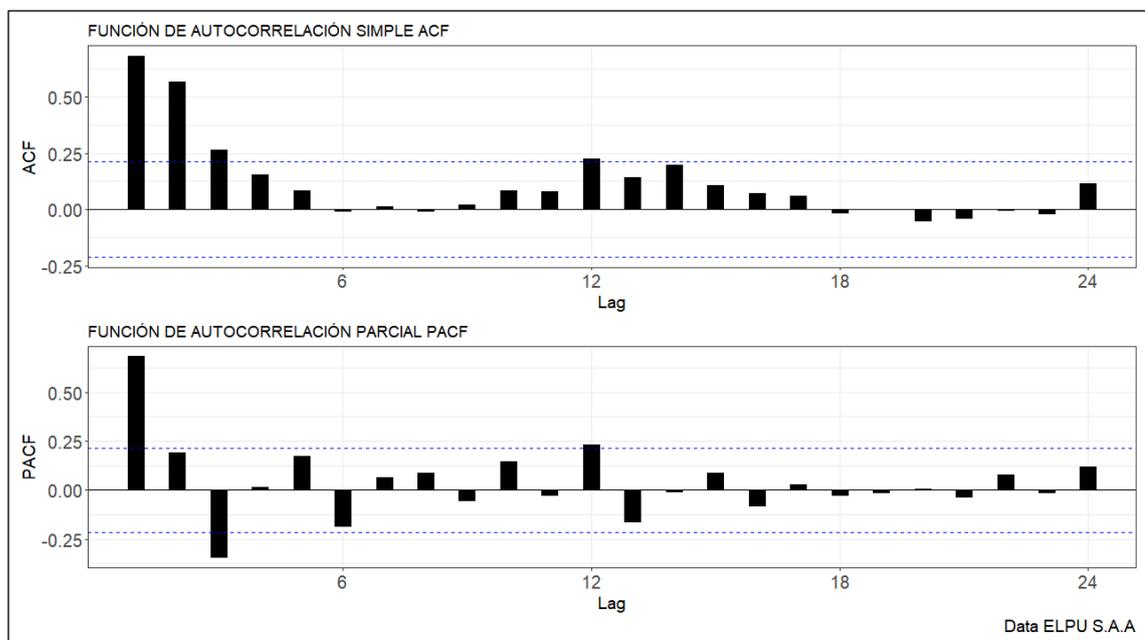


Figura 12. Correlograma Simple y Parcial de la Serie Consumo de Energía Eléctrica
Nota. Función de autocorrelación simple y parcial, de la serie original consumo de energía eléctrica de la Región Puno 2014-2020, realizado en el software R versión 4.2.1. Fuente: Elaboración propia del autor

De la **Figura 12**, se observa los correlogramas simple (*ACF*) y parcial (*PACF*) de la serie consumo de energía eléctrica de la región Puno en (MWh), en donde se presenta veinte cuatro rezagos de la correlación simple y parcial más significativas en su representación gráfica; con la prueba del test Aumentado Dickey Fuller se llegó determinar que la serie es “no estacionaria” presentado estacionalidad en su componente,

en ese sentido todo indica que estamos frente a un modelo multiplicativo estacional $ARIMA(p, d, q) \times (P, D, Q)_s$ con un orden estacional $s = 12$.

Ahora bien, dada la función de autocorrelación simple (ACF), nótese que en los primeros rezagos 1,2 y 3 presenta un decrecimiento lento indicando la presencia de estacionalidad en la parte regular de la serie; con relación a la función de autocorrelación parcial (PACF) se aprecia que los rezagos 1 y 3 sobrepasan los límites significativos de confianza seguido de correlaciones no significativas, de acuerdo al patrón que presenta los rezagos en ambos correlogramas **Figura 12** no existe similitud una de la otra, concluyendo que la serie debe ser diferenciada en una diferencia ($d = 1$) en su parte regular del modelo multiplicativo estacional $ARIMA(p, d, q) \times (P, D, Q)_s$, para la eliminación de la tendencia.

4.1.1.2 Segunda etapa de la fase identificación

Diferenciación de la serie. Para el cálculo de la primera diferencia haremos uso del comando `diff(data= , differences = 1)`, este comando del software R realizará la primera diferencia a la serie original.

Tabla 5

Serie Diferenciada del Consumo de Energía Eléctrica de la Región Puno

Mes \ Año	2014	2015	2016	2017	2018	2019	2020
enero		264	78	-348	-162	-123	441
febrero	-153	-336	-663	734	-844	473	-112
marzo	-2430	-507	-414	-3277	-645	-2608	-1660
abril	2898	1724	1713	3138	2184	3151	-1644
mayo	233	946	-597	-704	-841	-843	-3524
junio	1198	-238	1434	1730	1787	1882	1586
julio	-447	-514	-536	-264	-487	-484	1091
agosto	275	-229	156	-840	-558	34	2531

setiembre	-318	931	257	361	399	388	-88
octubre	-757	-723	-46	-42	-848	-855	1153
noviembre	1152	73	-563	149	385	-318	554
diciembre	-725	-640	383	-462	-603	-477	508

Nota. Serie consumo de energía eléctrica diferenciada en un periodo ($d=1$), realizado con el software R versión 4.2.1. Fuente: Elaboración propia del autor

En la **Tabla 5**, podemos observar la serie diferenciada de consumo de energía eléctrica de la región Puno, después de haberse aplicado la primera diferencia ($d = 1$), en el proceso de diferenciación la serie sufre la pérdida de un valor inicial y los valores de la serie cambian en su totalidad, esto debió a la aplicación de la fórmula de diferenciación $y_t = x_t - x_{t-1}$ con el objetivo de estabilizar la serie en media y varianza, para eliminar la tendencia y estacionalidad de la serie.

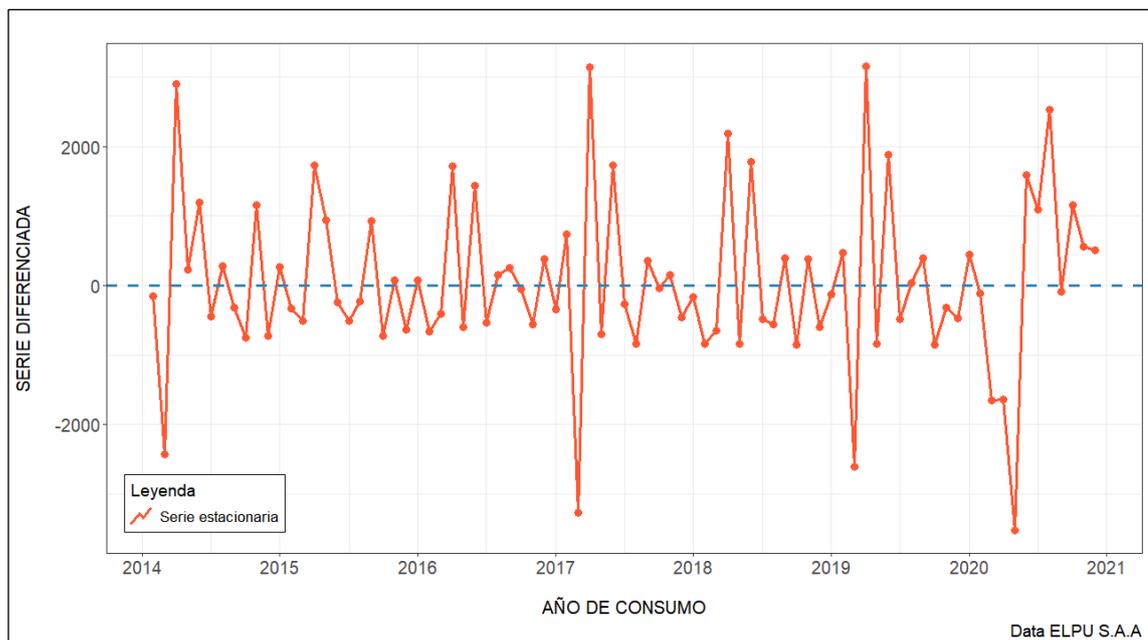


Figura 13. Serie Estacionaria de Consumo De Energía Eléctrica

Nota. Serie estacionaria después de haberse aplicado la primera diferencia a la serie original, hecho en el software R versión 4.2.1. Fuente: Elaboración propia del autor

Se presenta la **Figura 13**, donde se aprecia la serie de consumo de energía eléctrica en su forma diferenciada, del cual identificamos que la varianza y la media se mantiene constante a lo largo del tiempo, en la **Figura 13** la serie tiene un recorrido lineal mensual con valores diferenciados que van de arriba hacia abajo entorno a su media sin mostrar

signos de tendencia, claramente luego de haber diferenciado la serie original esta se ha convertido en una serie estacionaria tal como lo exige la metodología de Box – Jenkins.

Este análisis es respaldado por Quispe Pacco (2015) quien refiere que, luego de haber estudiado la serie original de consumo energía eléctrica de la provincia de Ayaviri, se determinó que es necesario diferenciar la serie en un periodo $d = 1$, para conseguir estabilizar la serie original en una serie estacionaria.

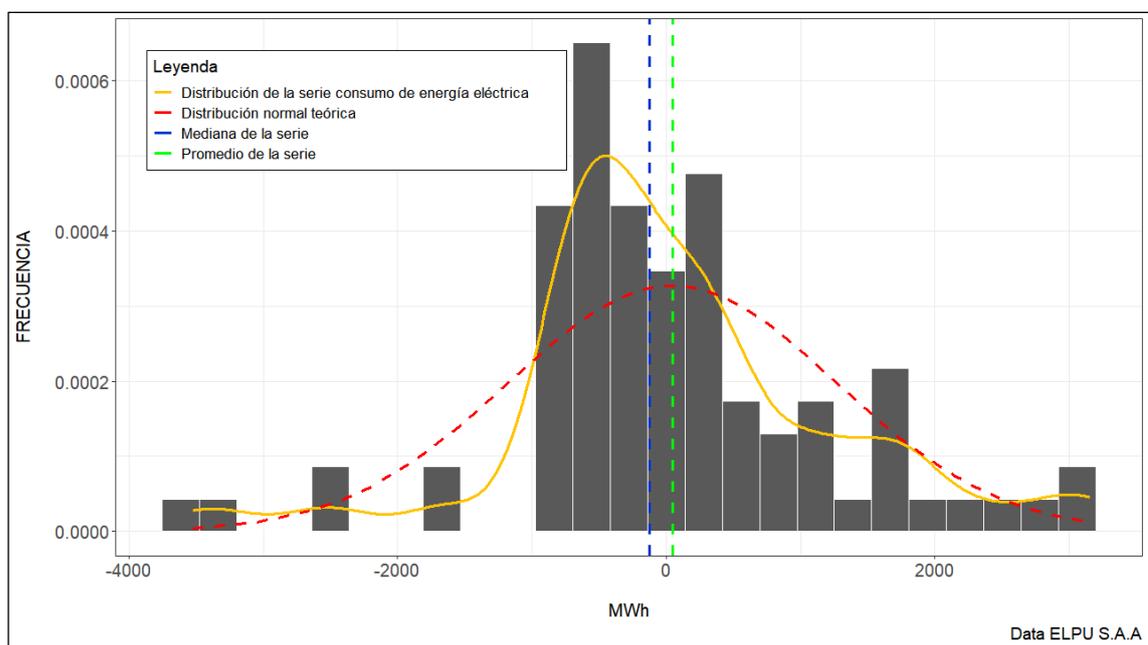


Figura 14. Distribución De La Serie Diferenciada

Nota. Distribución de la serie consumo de energía eléctrica de la región puno con una diferencia regular, realizado con el software r versión 4.2. 1.. Fuente: Elaboración propia del autor

En la **Figura 14** se aprecia la distribución de la serie Diferenciada, en donde se observar visualmente que sus valores individuales fluctúan entorno a su media que es $\hat{\chi} = 46.73$ (línea horizontal verde) y su mediana $\hat{\chi} = -122.53$ (línea horizontal azul), además los valores de la serie transformada se distribuyen según una distribución normal, por lo tanto podemos afirmar que estamos frente a una serie estacionaria, por lo mencionado, procederemos a determinar y confirmar estadísticamente con el test Aumentado de Dickey – Fuller la estacionariedad de la serie.



A. Prueba

Test aumentado de Dickey – Fuller

B. Formulación de hipótesis

H_0 : La serie diferenciada de consumo de energía eléctrica es no estacionaria.

H_α : La serie diferenciada de consumo de energía eléctrica es estacionaria.

C. Regla

Se acepta H_0 , Si, $p - value \geq 0.05$

Se rechaza H_0 , Si $p - value < 0.05$

D. Nivel de significancia

$$\alpha = 0.05$$

E. Prueba estadística

Tabla 6

Test Aumentado de Dickey – Fuller

statistic	p.value	parameter	method	alternative
-4.1	0.0107 *	4.0	Augmented Dickey-Fuller Test	stationary

Signif. codes: 0 <= '****' < 0.001 < '***' < 0.01 < '**' < 0.05

Nota. Prueba de hipótesis con el Test Aumentado de Dickey – Fuller. Fuente: R versión 4.2.1

F. Resultado

$$0.0107 < 0.05$$

G. Decisión

A la luz de los datos con un nivel de significancia del 5% y tras el resultado obtenido a través del Test Aumentado de Dickey – Fuller, se dispone de evidencia

suficiente para rechazar la hipótesis nula H_0 . Por lo tanto, la serie transformada consumo de energía eléctrica es estacionaria; de acuerdo a los resultados se confirma que la serie no presenta tendencia ni estacionalidad, bajo estos resultados es posible la identificación del orden de los parámetros del modelo multiplicativo $ARIMA(p, d, q) \times (P, D, Q)_{12}$.

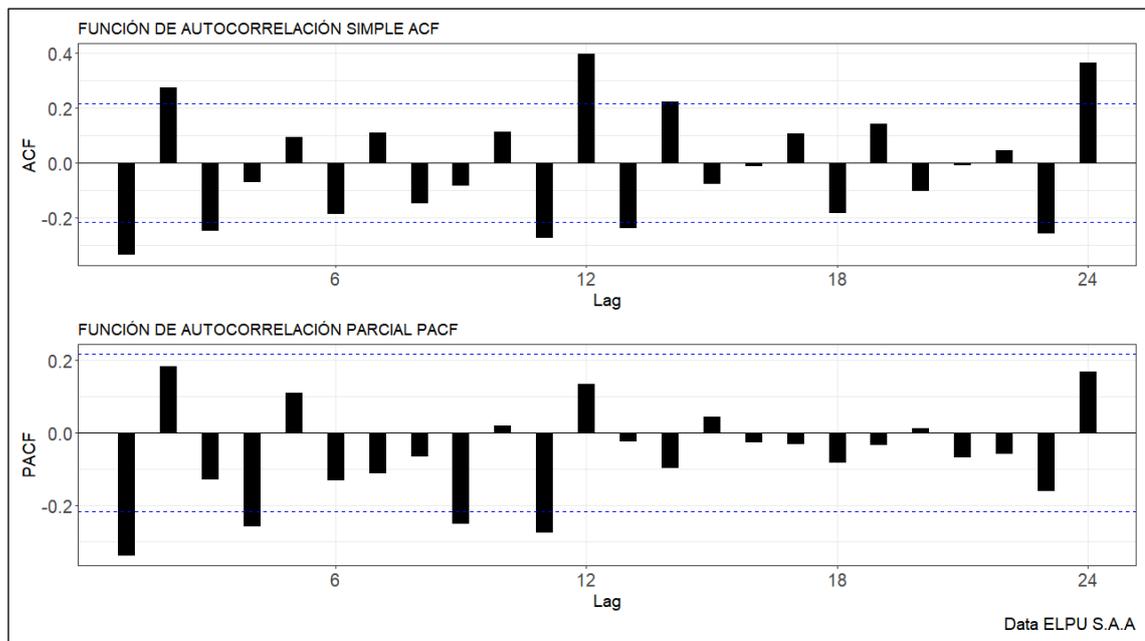


Figura 15. Función de autocorrelación simple y parcial de la primera diferencia de la serie
Nota. Función de autocorrelación simple y parcial, aplicado con la primera diferencia a la serie original consumo de energía eléctrica, realizado con el software R versión 4.2.1. Fuente: Elaboración propia del autor

Se presenta la **Figura 15**, donde se observa la función de autocorrelación simple y parcial de la serie diferenciada en 1 periodo, claramente se evidencia que la serie no presenta tendencia de acuerdo a los rezagos mostrados en el correlograma, tal como se ha confirmado con el test Aumentado de Dickey Fuller. En relación a la función de autocorrelación simple deducimos que los rezagos 1,2 y 3 sobresalen del intervalo del límite de confianza siendo estos rezagos significativos de cero, seguido de rezagos no significativos. En referencia a la función se autocorrelación parcial se identifica a los rezagos 1 y 4, ambos rezagos se encuentran fuera del intervalo del límite de confianza, también seguidos de rezagos no significativos.

Con los correlogramas mostrados en la **Figura 15** es posible identificar el orden de los parámetros del modelo multiplicativo estacional $ARIMA(p, d, q) \times (P, D, Q)_{12}$, realizando un contraste con el orden de los parámetros teóricos identificados por Box - Jenkins, sin embargo existe en el software *r* versión 4.2.1 la librería *forecast* para la identificación de modelos *ARIMA* automáticamente y esta función se denomina “*auto.arima()*”, esta función está desarrollada en base a la metodología Box – Jenkins, siendo una herramienta fundamental para la identificación de modelos *ARIMA* eficientemente.

Por su parte Donado & Huertas (2018) realizó una evaluación crítica a la función *auto.arima()* concluyendo que “ a pesar de existir métodos automáticos para la identificación de los modelo de series de tiempo, es necesario tener conocimiento teórico de las series temporales y conocimiento técnico de la función *auto.arima()*, considerando estos factores se reducirá el tiempo de trabajo y se obtendrá resultados confiables para la generación de pronósticos”.

4.1.1.3 Tercera etapa de la fase de identificación

Para la identificación de modelos *ARIMA* con la función *auto.arima()*, se ha configurado los parámetros de la función de acuerdo a la metodología en la fase de identificación del modelo, la función *auto.arima()* ha generado 71 modelos candidatos, véase en la **Tabla 7**, para el pronóstico de consumo de energía eléctrica de la Región Puno del 2021 al 2023, en la selección del modelo este debe cumplir con las condiciones del Criterio De Información De Akaike y el criterio de Información De Bayes (BIC) ambos criterios señalan que “mientras más pequeño sea el valor del AIC y BIC, mejor será el modelo para el pronóstico “ teniendo en consideración ambos criterios el modelo óptimo es el modelo 1:

$$ARIMA(0,1,3) \times (2,0,0)_{12}$$

entre 71 modelos candidatos tal como se muestra en la **Tabla 7**.

Tabla 7

Modelos $ARIMA(p, d, q) \times (P, D, Q)_{12}$ identificados

Modelos	$ARIMA(p, d, q) \times (P, D, Q)_{12}$	AIC	BIC
Modelo 1	$ARIMA(0,1,3) \times (2,0,0)_{12}$	1383.29	1397.80
Modelo 2	$ARIMA(0,1,3) \times (1,0,1)_{12}$	1384.80	1399.32
Modelo 3	$ARIMA(1,1,3) \times (1,0,0)_{12}$	1386.29	1400.81
Modelo 4	$ARIMA(0,1,3) \times (1,0,0)_{12}$	1387.30	1399.40
Modelo 5	$ARIMA(3,1,0) \times (2,0,0)_{12}$	1387.56	1402.07
⋮	⋮	⋮	⋮
Modelo 71	$ARIMA(0,1,0) \times (0,0,1)_{12}$	1406.24	1411.078

Nota. 71 modelos identificados en el software r con la función `auto.arima()`
Fuente: Elaboración propia del autor

4.1.2 Estimación

En esta fase se estiman los coeficientes $\theta, \varphi, \phi, \omega$ del modelo 1.

4.1.2.1 Estimación de los parámetros

Seleccionado el modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$ procedemos a estimar los parámetros, el modelo contiene los siguientes parámetros: Para estabilizar la serie en media y varianza se aplicó la diferencia de un periodo en su componente regular $d = 1$, el parámetro seleccionado para la parte regular de orden (q) es una Media Móvil de orden 3 $MA(3)$, del mismo que se deriva los coeficientes $MA(1), MA(2)$ y $MA(3)$, en cuanto a la parte estacional del modelo se identifica el parámetro Autorregresivo $AR(2)$, del cual se deriva los coeficientes $AR(1)$ y $AR(2)$, identificados los coeficientes del modelo se estima la ecuación (28).

$$(1 - \Phi_{12}B^{12} - \Phi_{24}B^{24})(1 - B)y_t = (1 - \omega_1B - \omega_2B^2 - \omega_3B^3)\varepsilon_t \quad (28)$$

Donde:

$y_t =$ Variable independiente

$\Phi_p =$ Coeficiente Autorregresivo estacional

$B =$ Operador de retardador

$\varepsilon_t =$ Error

$\omega_q =$ Coeficiente de Media Móvil regular

$(1 - B) =$ Diferencia regular

En la **Tabla 8**, se presenta los estimadores y coeficientes del modelo $ARIMA(0,1,3) \times (2,0,0)_{12}$, donde la probabilidad para los coeficientes $MA(1)$, $MA(3)$, $SAR(1)$ y $SAR(2)$ son significativos a un nivel de confianza del 5% ($p - value < 0.05$); con respecto al coeficiente $MA(2)$ la probabilidad no es significativa, sin embargo el estimador es menor que 1 con un valor de $0.0312 < 1$ considerandose bueno, este coeficiente podría ser considerado para ser excluido del modelo, no obstante el modelo ha sido seleccionado utilizando la técnica de estimación de máxima verosimilitud (MLE) (J Hyndman & Athanasopoulos, 2018), por lo tanto el coeficiente si aporta al modelo para el ajuste de la serie.

$$(1 - \Phi_{12}B^{12} - \Phi_{24}B^{24})(1 - B)y_t = (1 - \omega_1B - \omega_2B^2 + \omega_3B^3)\varepsilon_t \quad (29)$$

Tabla 8

Coefficientes del modelo ARIMA

<i>ARIMA(0,1,3) × (2,0,0)₁₂</i>				
Coefficients:				
term	estimate	std.error	statistic	p.value
ma1	-0.2258	0.1137	-1.9863	0.046997 *
ma2	0.0312	0.1230	0.2534	0.799962
ma3	-0.5168	0.0943	-5.4817	0.00000004212 ***
sar1	0.2631	0.1240	2.1222	0.033821 *
sar2	0.3696	0.1382	2.6741	0.007492 **

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

sigma^2 = 859900: log likelihood = -685.64

AIC=1383.29 AICc=1384.39 BIC=1397.80

Nota. Resultados del modelo identificado *ARIMA*. Fuente: Elaboración propia del autor

Una vez logrado estimar los parámetros del modelo y la ecuación, reemplazamos los estimadores de los coeficientes del modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$ en la ecuación (29), la estimación del modelo para la serie consumo de energía eléctrica de la región Puno, es la siguiente:

$$(1 - 0.2631B^{12} - 0.3696B^{24})(1 - B)y_t = (1 + 0.2258B - 0.0312B^3 + 0.5168B^3)\varepsilon_t \quad (30)$$

4.1.3 Validación

En esta fase pondremos en evaluación el modelo 1 que ha sido estimado en las secciones 4.1.1 y 4.1.2 del capítulo de resultados, con el objetivo de que el modelo sea adecuado para representar el comportamiento de la serie en estudio antes de ser empleado en la fase de pronóstico.

4.1.3.1 Análisis de Bondad de Ajuste

La prueba del criterio de akaike además de ser empleado en la fase de identificación del modelo *ARIMA* también es utilizado en la fase de validación, en donde el modelo identificado y estimado deberán cumplir con el análisis de bondad de ajuste; por ello haremos uso del criterio de información de akaike para validar el modelo mediante el valor estadístico *AIC* más bajo, “penalizando los modelos con mayor número de parámetros”(Jimenez, 1985, p. 136).

En la **Tabla 9**, se muestra 5 modelos candidato para la serie de consumo de energía eléctrica, de los 5 modelos identificados el modelo 1 cumple con el criterio de información de Akaike obteniendo el valor de *AIC* más bajo entre los 4 modelos candidato por lo que también ha sido seleccionado en la fase de identificación, sin embargo debemos señalar que los 4 modelos restantes también son adecuados para la representación de la serie original, esta afirmación se sustenta en el análisis de

correlogramas simple (ACF) y parcial (PACF) pues los rezagos de los 4 modelos están dentro de los límites de confianza.

No obstante, para tener mayor certeza que el modelo 1 es el adecuado hemos aplicado el criterio de información de Bayes donde el modelo 1 obtiene un valor *BIC* bajo frente a los 4 modelos **Tabla 9** (Hanke & Wichern, 2018, p.431).

Por lo expuesto, para la validación del modelo por medio del análisis de bondad de ajusta, el modelo 1 cumple con ambos criterios obteniendo los valores más bajos, determinando que los coeficientes del modelo 1 son suficientes para la representación de la serie consumo de energía eléctrica.

Tabla 9

Modelos Candidato Para el Pronóstico de Energía Eléctrica de la Región Puno

Modelos	$ARIMA(p, d, q) \times (P, D, Q)_{12}$	AIC	BIC
Modelo 1	$ARIMA(0,1,3) \times (2,0,0)_{12}$	1383.29	1397.80
Modelo 2	$ARIMA(0,1,3) \times (1,0,1)_{12}$	1384.80	1399.32
Modelo 3	$ARIMA(1,1,3) \times (1,0,0)_{12}$	1386.29	1400.81
Modelo 4	$ARIMA(0,1,3) \times (1,0,0)_{12}$	1387.30	1399.40
Modelo 5	$ARIMA(3,1,0) \times (2,0,0)_{12}$	1387.56	1402.07

Nota. Modelos identificados con el software R versión 4.2.1. Fuente: Elaboración propia del autor

4.1.3.2 *Análisis de residuos*

En el proceso para la validación del modelo seleccionado haremos uso del análisis de los Correlogramas Simple (ACF) y Parcial (PACF), donde analizaremos los residuos del modelo 1 a través de la significancia de los rezagos, para determinar si el modelo se ajusta razonablemente a los datos de la serie original.

La **Figura 16** muestra los correlograma simple y parcial del modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$ donde se observa que los rezagos en ambos



correlogramas se encuentran dentro del umbral de los límites de confianza al 95%, esto quiere decir que los residuos del modelo seleccionado son puramente aleatorios y que los datos no están correlacionados una de la otra.

Por su parte Barahona Bonifaz (2018) comprueba la existencia de independencia de su modelo $ARIMA(0,1,1) \times (1,0,0)_{12}$ por el mismo análisis, indicando que no existen retardos significativos en los correlogramas, corroborando que los residuos del modelo son independientes.

Vázquez Díaz & Gamonal Sánchez (2019) realizan un análisis similar a los residuos mediante las Autocorrelaciones Simples y Parciales del modelo $ARIMA(4,2,0)$, donde los retardos en ambos correlogramas se encuentran dentro de los límites de confianza, en otras palabras el modelo presenta aleatoriedad en su residuo siendo estos de ruido blanco, por lo tanto según refieren los investigadores el modelo es adecuado para la representación de la serie.

En ambas investigaciones los hallazgos encontrados concuerdan con lo analizado en este proceso de validación mediante el análisis de correlogramas simples(ACF) y parciales(PACF), en cuanto podemos afirmar que el modelo es el adecuado y el más óptimo para la representación de la serie original.

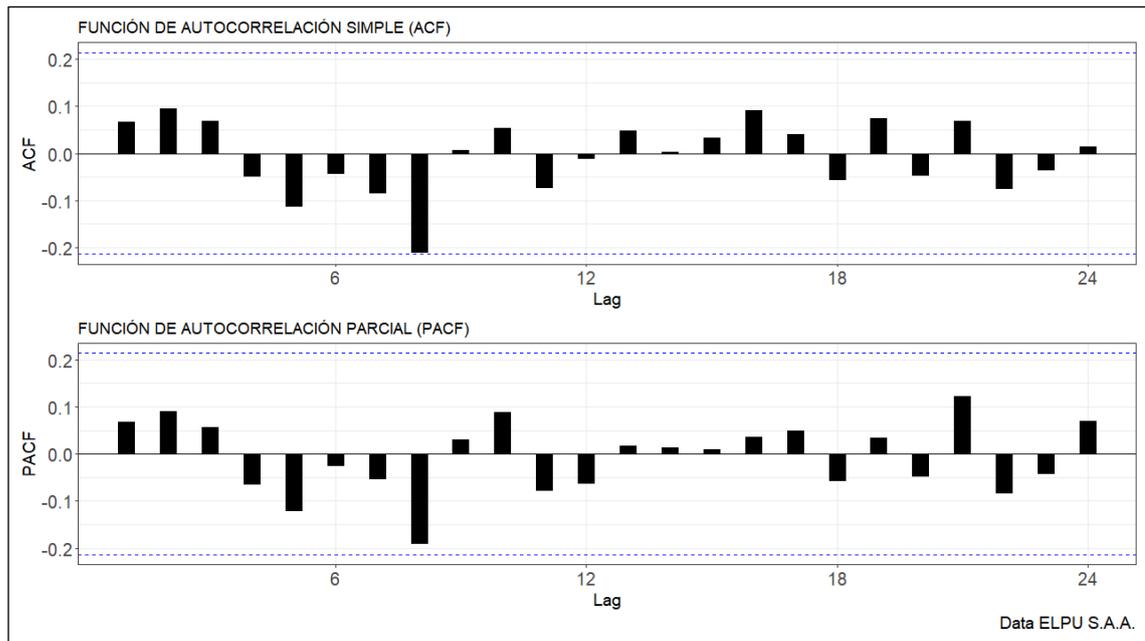


Figura 16. Autocorrelación Simple y Parcial de los Residuos del Modelo Seleccionado

Nota. Función de autocorrelación simple y parcial de los residuos del modelo $ARIMA(0,1,3) \times (2,0,0)_{12}$. Fuente: Elaboración propia del autor

Por otra parte, las funciones de autocorrelación simple (ACF) y parcial (PACF) nos indican visualmente en el análisis de correlogramas que los residuos de la serie son independientes una de la otra, sin embargo, con la prueba del test de Q de Box-Ljung comprobaremos estadísticamente a través del planteamiento de una hipótesis el supuesto de independencia.

A. Prueba

Test de Q de Box – Lung

B. Formulación de hipótesis

H_0 : Los residuos se distribuye de forma independiente (ruido blanco).

H_{α} : Los residuos no se distribuyen de forma independiente (no son ruido blanco).

C. Nivel de significancia

$$\alpha = 0.05$$

D. Regla

Se acepta H_0 , Si, $p - value \geq 0.05$

Se rechaza H_0 , Si $p - value < 0.05$

E. Prueba estadística

Tabla 10

Test de Q Ljung-Box

statistic	p.value	parameter	method
0.4	0.5316	1.0	Box-Ljung test

Signif. codes: 0 \leq '****' $<$ 0.001 $<$ '***' $<$ 0.01 $<$ '**' $<$ 0.05

Nota. Prueba de hipótesis con el Test de Q de Box – Lung. Fuente: R versión 4.2.1

F. Resultado

$$0.5316 \geq 0.05$$

G. Decisión

A la luz de los datos con un nivel de significancia del 5% y tras el resultado obtenido a través del Test de Q de Ljung-Box, se dispone de evidencia suficiente para aceptar la hipótesis nula H_0 . Por lo tanto, los residuos del modelo seleccionado cumplen con el supuesto de independencia esto quiere decir que no existe correlación en los residuos.

En los resultados obtenidos por Quispe Pacco (2015) en la fase de validación de su modelo multiplicativo estacional $ARIMA(3,1,1) \times (1,0,1)_{12}$ realiza el Test de Q de Ljung-Box para probar la independencia de los datos, obteniendo como resultado una probabilidad superior al $\alpha = 0.05$, comprobando su hipótesis nula que los residuos de la serie son independientes, es decir son ruido blanco.

Obteniendo estos resultados estadísticamente favorables, el modelo $ARIMA(0,1,3) \times (2,0,0)_{12}$ ha superado con éxito el proceso de validación del modelo, por tanto el modelo es óptimo para ser implementado en la etapa de generación de pronósticos de energía eléctrica para la Región de Puno.

4.1.4 Pronóstico

Analizaremos la precisión del pronóstico que tiene el modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$ a través de los residuos del modelo, que viene hacer la diferencia entre los valores reales y valores pronosticados, con el fin de obtener pronósticos precisos que respondan a la toma de decisiones.

4.1.4.1 Evaluación de capacidad predictiva

En la **Tabla 11** se presenta las métricas para la medición del grado de acierto de un pronóstico, entre el más utilizado es la RMSE (Raíz media de los errores al cuadrado) así como también se usa para comparar modelos seleccionando a aquel modelo que contenga el menor valor estadístico. En este caso, el RMSE del modelo 1 es de 893.57 siendo el valor más bajo entre los cuatro modelos, asegurando que el modelo tiene un mejor ajuste a la serie original por contener el promedio mínimo de errores (diferencia entre la serie original y el modelo ajustado), siendo este un valor aceptable considerando que el modelo seleccionado cumple con el principio de parsimonia apoyado por los criterios de AIC y BIC, ajustándose adecuadamente a la serie original.

La métrica de evaluación MAE (error absoluto medio) mide la exactitud de los pronósticos, promediando el conjunto de errores del pronóstico, en este caso para el modelo que hemos seleccionado, el valor de medición de error MAE es de 625.80, indicando que por cada pronóstico de consumo de energía eléctrica este estará desviado en un promedio de 625.80 MWh, véase la **Tabla 11**.

Tabla 11

Modelos ARIMA y sus Métricas de Precisión del Pronóstico

Modelos	$ARIMA(p, dk, q) \times (P, D, Q)_{12}$	AIC	BIC	RMSE	MAE	MAPE
Modelo 1	$ARIMA(0,1,3) \times (2,0,0)_{12}$	1383.29	1397.80	893.576	625.800	2.790
Modelo 2	$ARIMA(0,1,3) \times (1,0,1)_{12}$	1384.80	1399.32	908.244	624.464	2.787
Modelo 3	$ARIMA(1,1,3) \times (1,0,0)_{12}$	1386.29	1400.81	924.697	643.741	2.864
Modelo 4	$ARIMA(0,1,3) \times (1,0,0)_{12}$	1387.30	1399.40	943.120	656.495	2.929
Modelo 5	$ARIMA(3,1,0) \times (2,0,0)_{12}$	1387.56	1402.07	907.978	649.547	2.880

Nota. Modelos con sus respectivas métricas de evaluación de capacidad predictiva. Fuente: Elaboración propia del autor

El valor MAPE (Error de porcentaje medio absoluto) es un indicador de desempeño del pronóstico que mide el error en porcentajes; (Lewis, 1982, citado en Sudhan et al., 2020) en su libro “International and Business Forecasting Methods. Butterworths, London.”, define una escala de precisión de los pronósticos en base al valor MAPE, véase la **Tabla 12**.

Tabla 12

Escala de Lewis(1982) Para el Análisis de Precisión MAPE

Porcentaje MAPE	Precisión del pronóstico
Menos de 10%	Alta precisión
11% a 20%	Buen pronóstico
21% a 50%	Pronóstico razonable
51% a más	Pronóstico inexacto

Nota. Escala de Lewis para la evaluación del valor MAPE Fuente: Lewis (1982)

En este caso el valor MAPE obtenido por nuestro modelo en términos porcentuales es del 2.79%, ubicando el valor en la escala de Lewis se traduce que los pronósticos serán de alta precisión, tal como se muestra en la **Tabla 12**.

Barahona Bonifaz, (2018) compara la precisión de su modelo $ARIMA(0,1,1) \times (1,0,0)$ frente al modelo Holt – Winters, encontrado las métricas de evaluación RMSE de



5,3434, una MAE del 4,1358 y un MAPE de 2,4206%, estos indicadores son inferiores a los obtenidos por el método de Holt – Winters, asegurando mediante estas métricas de evaluación que el modelo *ARIMA* en cuanto a capacidad predictiva es precisa.

Marcelo Barreto & Villareal Escate, (2021) demuestran que el modelo $ARIMA(0,1,1) \times (1,0,2)_{12}$ es mejor en comparación con el método de suavización exponencial Multiplicativo de Winter, afirmación que se apoya bajo los resultados obtenidos mediante el error absoluto porcentual promedio (MAPE), el modelo *ARIMA* presenta un valor MAPE del 2.296% y el modelo Suavizado exponencial de Holt Winter presenta un valor MAPE del 2.4%, demostrando de esta manera que el modelo *ARIMA* es quien mejor se ajusta al pronóstico de demanda de energía eléctrica sistema eléctrico Cusco.

Ambas investigaciones coinciden en que el modelo *ARIMA* es el adecuado para el ajuste y pronóstico de sus series tiempo, las métricas de evaluación de capacidad predictiva para la comparación de sus modelos fue el indicador MAPE, los resultados encontrados con el indicador fueron inferiores al 10%, considerando ese valor según a la escala de Lewis los modelos seleccionados realizaran pronósticos con alta precisión; hallazgos que se relacionan con nuestro valor MAPE calculado.

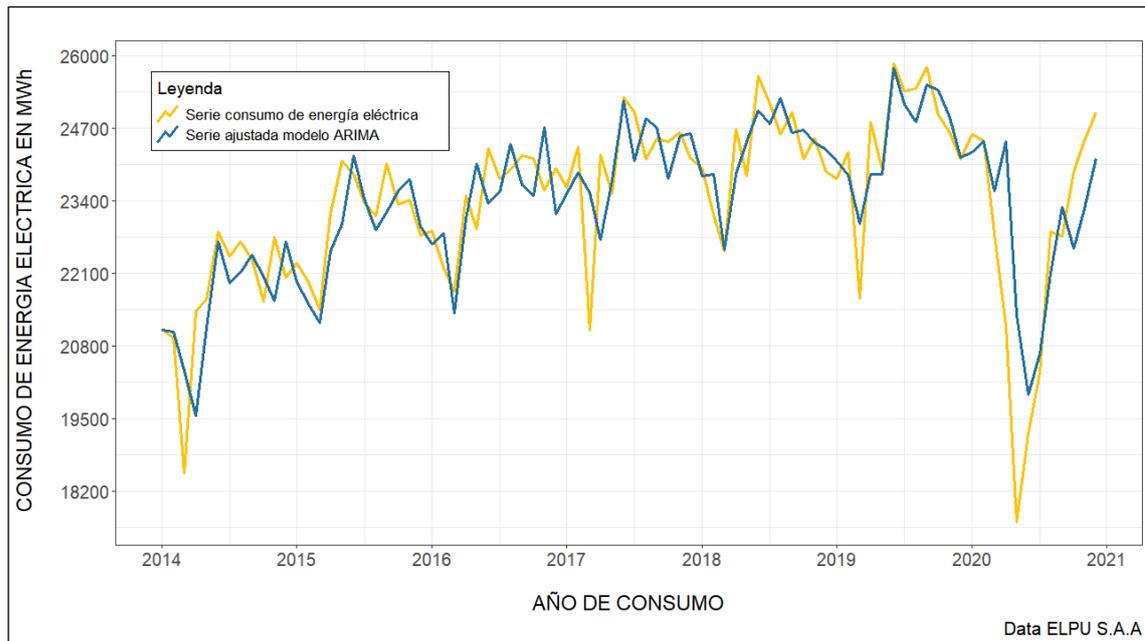


Figura 17. Serie Generada por el Modelo *ARIMA* validada, frente a la Serie Original
Nota. Serie ajustada $ARIMA(0,1,3) \times (2,0,0)_{12}$ frente a la serie original, realizado en el software R.4.2.1.
Fuente. Propio del autor. Fuente: Elaboración propia del autor

En la **Figura 17** se muestra la serie original consumo de energía eléctrica (línea amarilla) y la serie generada por el modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$ (línea azul), como se aprecia el modelo generador *ARIMA* se ajusta (es semejante) perfectamente a la serie original de consumo de energía eléctrica.

4.1.4.2 Ejecución de pronósticos

En la **Tabla 13** se muestra los valores proyectados de consumo de energía eléctrica para la región Puno, para los próximos 3 años aun nivel del 95% de confianza con el modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$ cuya ecuación general es:

$$(1 - 0.2631B^{12} - 0.3696B^{24})(1 - B)y_t = (1 + 0.2258B - 0.0312B^2 + 0.5168B^3)\varepsilon_t \quad (31)$$

Tabla 13*Pronóstico consumo de Energía Eléctrica Para la Región de Puno Durante 3 Años*

Años	Meses	Pronóstico (MWh)	Intervalos de confianza al 95%	
			Límite inferior(MWh)	Límite superior(MWh)
2021	enero	24187.44	22369.95	26004.93
2021	febrero	23732.08	21433.56	26030.61
2021	marzo	21907.94	19182.90	24632.97
2021	abril	22640.03	19864.98	25415.08
2021	mayo	21401.37	18577.19	24225.54
2021	junio	22514.16	19641.69	25386.62
2021	julio	22622.58	19702.62	25542.54
2021	agosto	23300.90	20334.21	26267.59
2021	setiembre	23421.17	20408.47	26433.86
2021	octubre	23408.42	20350.41	26466.43
2021	noviembre	23436.83	20334.17	26539.49
2021	diciembre	23394.12	20247.44	26540.80
2022	enero	23350.54	20047.97	26653.11
2022	febrero	23189.25	19767.63	26610.86
2022	marzo	22095.68	18555.22	25636.14
2022	abril	21680.63	18078.73	25282.54
2022	mayo	20052.28	16389.95	23714.60
2022	junio	20931.17	17209.41	24652.93
2022	julio	21363.09	17582.83	25143.36
2022	agosto	22476.88	18639.01	26314.75
2022	setiembre	22475.82	18581.19	26370.45
2022	octubre	22898.51	18947.93	26849.08
2022	noviembre	23110.81	19105.08	27116.55
2022	diciembre	23287.20	19227.05	27347.35
2023	enero	22985.34	18670.66	27300.03
2023	febrero	22774.61	18274.09	27275.14
2023	marzo	21812.70	17126.84	26498.55
2023	abril	21974.07	17203.95	26744.19
2023	mayo	21087.84	16234.92	25940.76
2023	junio	21730.36	16796.03	26664.69



2023	julio	21884.08	16869.65	26898.50
2023	agosto	22427.82	17334.57	27521.08
2023	setiembre	22471.99	17301.11	27642.88
2023	octubre	22578.49	17331.13	27825.86
2023	noviembre	22644.85	17322.11	27967.60
2023	diciembre	22675.48	17278.40	28072.55

Nota. Pronóstico de consumo de energía eléctrica en Megavatios por hora (MWh) para la región de Puno, en los próximos 3 años (2021 al 2023), hecho en el software R versión 4.2. 1.. Fuente: Elaboración propia del autor

En la **Figura 18** con los valores de la **Tabla 13** se muestra el pronóstico de energía eléctrica de la región de Puno para los años 2021,2022 y 2023 al 95% de confianza, con un grado de error promedio por cada pronóstico de 625.80 MWh; consideramos que existen tres escenarios para la toma de decisiones, el primer escenario es el escenario conservador (línea roja) donde indica que el consumo de energía eléctrica decrece en los próximos 3 años, el segundo escenario es el escenario optimista (Límite superior de la línea roja) donde se señala que el consumo de energía eléctrica en la región de Puno se incrementará en los próximos 3 años, y el tercer escenario es el escenario pesimista donde ocurre todo lo contrario (Límite inferior de la línea roja) indicando que el consumo de energía eléctrica en los próximos tres años decrece.

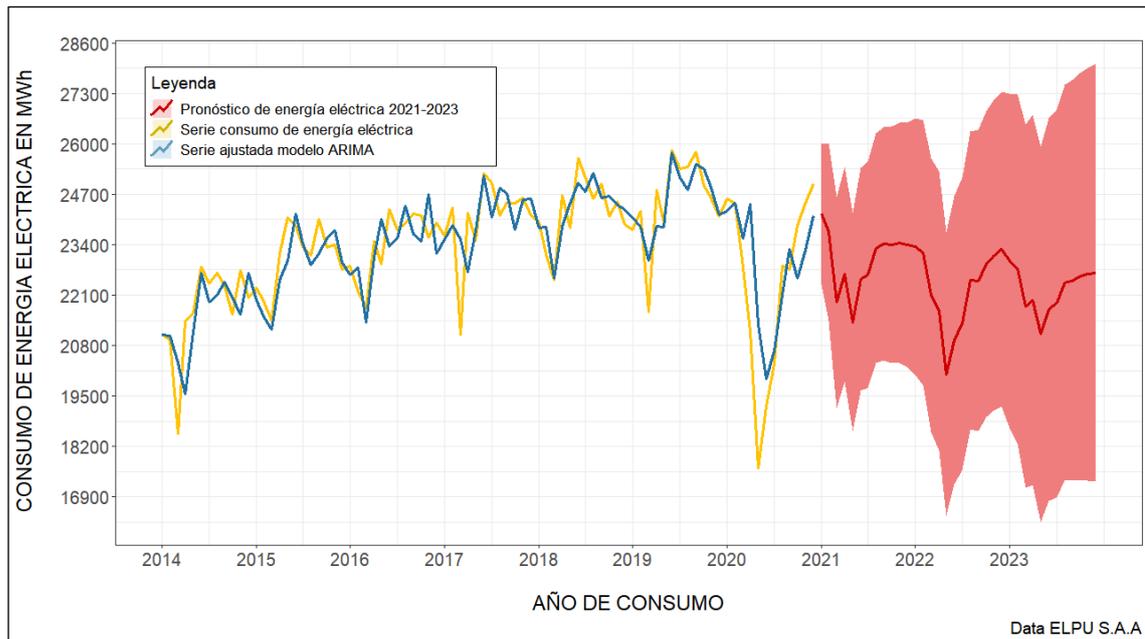


Figura 18. Pronóstico del Consumo de Energía Eléctrica para los periodos 2021 – 2023

Nota. Pronóstico del consumo de energía eléctrica en Megavatios hora (MWh) para la Región de Puno, para los próximos 3 años (2021 al 2023), realizado en el software R versión 4.2.1. Fuente: Elaboración propia del autor

4.2 DISCUSIÓN

La metodología de Box – Jenkins, exige el cumplimiento de cuatro fases para determinar un modelo *ARIMA*, de esta forma se garantiza una mejor descripción, ajuste y proyección de la variable en estudio (Marcelo Barreto & Villareal Escate, 2021). El modelo al que hemos llegado a determinar es el modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$, este modelo es respaldado por los criterios de $AIC = 1383.29$ y $BIC = 1397.80$ valores que son inferiores frente a otros modelos, demostrando que el modelo es el apropiado cumpliendo con el principio de parsimonia, Barahona Bonifaz (2018) en su investigación determina el modelo multiplicativo $ARIMA(0,1,1) \times (1,0,0)_{12}$ reportando indicadores de $AIC = 364.89$ y un $BIC = 371$ valores que son inferiores frente a otros modelos, mientras que Quispe Pacco (2015) determina su modelo únicamente a través del indicador $BIC = 2476.5$ siendo este el menor valor frente a otros 3 modelos, en ambas investigaciones determina su modelo seleccionando el menor valor obtenido en los indicadores AIC y BIC ; de esta manera se

prueba que el modelo proporciona un mejor ajuste a la serie de tiempo consumo de energía eléctrica; en atención a la hipótesis general. Véase la **Figura 17**.

El comportamiento de la serie consumo de energía eléctrica como se menciona en la fase de identificación del modelo, de acuerdo a las **Figura 8** y **Figura 9** presenta una tendencia creciente, con variaciones cíclicas propias de una serie no estacionaria con media y varianza que varían en el tiempo, también se pudo identificar que la serie presenta ciclos estacionales a lo largo de los 7 periodos, y finalmente la serie en el último periodo presenta aleatoriedad por un hecho externo que afecta a la serie, Quispe Colque (2019) señala en su investigación que, la serie ventas mensuales de cemento rumi presenta una tendencia positiva muy irregular, con variabilidad y estacionalidad en la serie, hallazgo que concuerda con nuestro análisis, y confirma la segunda hipótesis de investigación, donde señalamos que el comportamiento de la serie es creciente y positiva, con variaciones cíclicas y estacionales.

Los modelos *ARIMA* permiten detectar el comportamiento de la serie y del proceso generador, esto posibilita que se describa una gran variedad de comportamientos de series de tiempo no estacionarias, el modelo estimado para este estudio es el modelo $ARIMA(0,1,3) \times (2,0,0)_{12}$ cuya expresión general es la ecuación (29), que ha sido estimado de acuerdo a la técnica de máxima verosimilitud que se emplea para ajustar los parámetros, encontrándose coeficientes significativos a un nivel del 5%, es decir que los coeficientes son adecuados y aportan al modelo, Arévalo Cabrera & Santisteban Morales (2020) considera que un parámetro es significativo cuando el estadístico T-Student (t) posea un valor menor de 2 y/o un (p-value) menor a 0.05, conforme a lo señalado los coeficientes estimados en este estudio son significativos, véase la **Tabla 8**; por lo tanto el



modelo que hemos estimado es el modelo que mejor ajuste tiene a la serie, confirmando la tercera hipótesis de investigación, véase la **Figura 17**.

Así mismo, se ejecutaron los pronósticos de consumo mensual de energía eléctrica a 36 meses (3 años), con una capacidad de predicción del 2.79% considerándose de alta precisión, al respecto Marcelo Barreto & Villareal Escate (2021) realizan una comparación de capacidad predictiva entre el modelo ARIMA contra el modelo Multiplicativo Winter, seleccionando el modelo ARIMA por presentar un valor MAPE del 2.96%, siendo inferior frente al otro modelo, resultado que coincide con nuestra investigación. Por otra parte Arévalo Cabrera & Santisteban Morales (2020) seleccionaron el modelo multiplicativo estacional $ARIMA(2,1,2) \times (1,0,0)_{12}$ con una capacidad de predicción del valor $MAPE = 19.32\%$ resultado que difiere con nuestra investigación, pues los pronósticos del modelo serán buenos pero no precisos. Con respecto al incremento en el consumo de energía en los próximos tres años, los pronósticos señalan que habrá un incremento para un escenario optimista, véase la **Tabla 13**, de plantearse un escenario conservador la tasa de crecimiento es a la baja, al respecto (Mejía Vasquez & Gonzales Chávez, 2019) señala que este fenómeno se debe a que el modelo ARIMA recoge las últimas tendencias; como hemos señalado en la fase de identificación el modelo presenta una anomalía en los meses de marzo, abril y Mayo del año 2020, véase la **Figura 9**, haciendo que su efecto en los pronósticos tiendan a la baja, de seleccionarse un escenario optimista se cumple con lo señalado en nuestra cuarta hipótesis planteada.

V. CONCLUSIONES

PRIMERO: El modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$, es el modelo apropiado que mejor se ajusta a la serie en el pronóstico de consumo de energía eléctrica para los próximos tres años.

SEGUNDO: En el análisis a la serie consumo de energía eléctrica se muestra que factores atípicos como la pandemia en el año 2020 distorsionan la serie de consumo de energía eléctrica, este fenómeno afecta directamente en la identificación del modelo ARIMA, pues estos modelos capturan los últimos datos de la serie para su implementación en el modelo, ello desfavorece al pronóstico con el menor margen de error.

TERCERO: Se ha estimado los parámetros con el método de máxima verosimilitud del modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$, obteniendo los siguientes coeficientes a un nivel de confianza del 5 %, indicando que son óptimos: $MA(1) = -0.2258$, $MA(2) = 0.0312$, $MA(3) = -0.5168$, $SAR(1) = 0.2631$, $SAR(2) =$, y está representada por la siguiente ecuación general:

$$(1 - 0.2631B^{12} - 0.3696B^{24})(1 - B)y_t = (1 + 0.2258B - 0.0312B^3 + 0.5168B^3)\varepsilon_t$$

CUARTO: Los pronósticos generados por el modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$ tienen un grado de alta precisión, de acuerdo a las pruebas realizadas al modelo en cuanto se refiere a la evaluación de la capacidad predictiva, el modelo ha obtenido valores inferiores en los indicadores $MAPE = 2.79\%$ y un $MAE = 625.80$ frente a otros modelos candidato, demostrando de esta forma que el modelo es apropiado en todos los aspectos.



VI. RECOMENDACIONES

PRIMERO: Recomendamos a la empresa ELECTRO PUNO S.A.A., considerar este estudio como referente para ser implementado como metodología en la planificación de programación de mantenimientos preventivos en zonas de concesión de mayor consumo de energía eléctrica en la región de Puno.

SEGUNDO: Recomendamos para las próximas investigaciones realizar pronósticos por cada sistema eléctrico con los que cuenta Electro Puno, de esta forma la empresa podrá determinar planes de mantenimiento preventivo a los sistemas de distribución de energía eléctrica en sectores específicos con alta demanda de energía eléctrica.

TERCERO: En las fases de identificación y estimación de parámetros de los modelos ARIMA, se recomienda a los investigadores, docentes y estudiantes de la Facultad de Ingeniería Estadística E Informática, el uso de la función `auto.arima()` del paquete `forecast()` en el software R, el cual seleccionará un modelo ARIMA automáticamente, con ello se reducirá el tiempo de trabajo que emplea el investigador para la identificación de un modelo apropiado.

CUARTO: En las próximas investigaciones se recomienda desarrollar modelos multivariados para el pronóstico de energía eléctrica, considerando horizontes de tiempo a largo plazo, empleando variables como la población, producto bruto interno(PBI), perfil de consumo de usuario y otras variables macroeconómicas, que servirán para determinar la demanda máxima de energía eléctrica en los próximos 4 a 20 años, los resultados contribuirían en estudios de planificación energética para la región de Puno.



VII. REFERENCIAS BIBLIOGRÁFICAS

- Amat Rodrigo, J. (2016). *Análisis de normalidad: gráficos y contrastes de hipótesis*.
https://www.cienciadedatos.net/documentos/8_analisis_normalidad
- Aragón, F., & Parra Royón, M. (2017). *Series Temporales*. GitHub.
<https://github.com/manuparra/seriestemporales/blob/master/README.md>
- Arévalo Cabrera, S. E., & Santisteban Morales, C. R. (2020). Universidad Nacional Pedro Ruiz Gallo [Universidad Nacional Pedro Ruiz Gallo]. In *Universidad Nacional Pedro Ruiz Gallo*.
[https://repositorio.unprg.edu.pe/bitstream/handle/20.500.12893/5092/BC-3893 BANCES PISCOYA-ROJAS PUICON.pdf?sequence=3&isAllowed=y](https://repositorio.unprg.edu.pe/bitstream/handle/20.500.12893/5092/BC-3893_BANCES_PISCOYA-ROJAS_PUICON.pdf?sequence=3&isAllowed=y)
- Barahona Bonifaz, J. I. (2018). *Pronósticos Del Consumo Y Demanda De Potencia Máxima De Energía Eléctrica En La Ciudad De Riobamba Para El Periodo 2017-2020, Mediante Modelos Arima*. Escuela Superior Politécnica De Chimborazo.
- BCRP. (2022). *Caracterización Del Departamento De Puno* (Departamen, p. 15).
<https://www.google.com/maps/place/Puno/@-15.0833935,->
- Borwerman, B. L., O'Connell, R. T., & Koehler, A. B. (2006). *Pronosticos, Series de Tiempo y Regresion Un Enfoque Aplicado* (D.R.2007 p).
- Cervan, D. (2017). *Proyección de la demanda eléctrica: clasificación, variables y métodos empleados*. <https://dheybicervan.com/proyeccion-de-la-demanda-electrica/>
- Chong Fuentes, A., & Aguilar, R. (2016). Proyección de Series de Tiempo para el Consumo de la Energía Eléctrica a Clientes Residenciales en Ecuador. *Revista Tecnológica ESPOL-RTE*, 29(1), 1–21.
- Claudia, S. (2010). *Series de Tiempo y pronósticos* (pp. 673–685).
- Court Monteverde, E., & Williams Rengifo, E. (2011). *Estadísticas Y Econometría Financiera* (Impreso en MetroColor S.A. (ed.); Vol. 1, Issue 1).
https://drive.google.com/file/d/11wROxx_eixpxm-6_0GiBbWtGfgzK_fd-/view?fbclid=IwAR0KU3FwQ-RR3gv1CSGh0_Jf4gPsKGVpwxgiMc69zaqAw06x_yGHK1CNB24
- Donado, S., & Huertas, N. (2018). *Evaluación Crítica De La Función auto.arima De R En La Identificación De Procesos De Series De Tiempo* (Issue 201326730).
- Ducos Ramiez, A. (2009). *Metodología Para El Cálculo De La Demanda De Mediano Y Largo Plazo Del Sistema Eléctrico Interconectado Nacional*. Universidad Nacional de Ingeniería.
- Electro Puno S.A.A. (2021). *Plan Estratégico Institucional 2017-2021* (Issue 1, p. 74).
http://www.electropuno.com.pe/web/elpuadmin/res/files/2020/07/20/trans_3529.pdf



- Gallardo Echenique, E. E. (2017). Metodología de la Investigación. Manual Autoformativo Interactivo I. In *Universidad Continental* (Vol. 1). <http://repositorio.continental.edu.pe/>
- Gujarati, D. N., & Porter, D. C. (2010). *Econometría* (S. A. D. C. . McGRAW-HILL/INTERAMERICANA EDITORES (ed.); Quinta edi). https://drive.google.com/file/d/1lwROxx_eixpxm-6_0GiBbWtGfgzK_fd-/view?fbclid=IwAR0KU3FwQ-RR3gv1CSGh0_Jf4gPsKGVpwxgiMc69zaqAw06x_yGHK1CNB24
- Hanke, J. E., & Wichern, D. W. (2018). Pronósticos en los negocios. In *Journal of Materials Processing Technology* (Vol. 9, Issue 1). <http://dx.doi.org/10.1016/j.cirp.2016.06.001><http://dx.doi.org/10.1016/j.powtec.2016.12.055><https://doi.org/10.1016/j.ijfatigue.2019.02.006><https://doi.org/10.1016/j.matlet.2019.04.024><https://doi.org/10.1016/j.matlet.2019.127252><http://dx.doi.org>
- Hernandez, J. (2014). Modelación ARIMA. In *UNAM. Economía Mexicana*. <http://www.ptolomeo.unam.mx:8080/jspui/bitstream/132.248.52.100/363/7/A7.pdf>
- Hernandez Sampieri, R., Fernandez Collado, C., & Baptista Lucio, M. del P. (2010). Metodología de la Investigación. In *Metodología de la investigación* (Quinta edi). <http://www.casadellibro.com/libro-metodologia-de-la-investigacion-5-ed-incluye-cd-rom/9786071502919/1960006>
- INEI. (2018). *Censos Nacionales 2017: XII de Población, VII de Vivienda y III de Comunidades Indígenas*. https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1539/libro.pdf
- J Hyndman, R., & Athanasopoulos, G. (2018). *Forescasting: Principles and Practice* (Otexts (ed.)). <https://otexts.com/fpp2/what-can-be-forecast.html>
- Jimenez, E. U. (1985). *Análisis de series temporales Modelos ARIMA*. PARANINFO S.A.
- Maque Tinta, R. S. (2017). *Análisis, Diagnóstico Y Propuesta De Mejora De Calidad De Servicio A Causa De Fallas Imprevistas En El Suministro Eléctrico En El Distrito De Macusani-Carabaya*. Universidad Nacional Del Altiplano - Puno.
- Marcelo Barreto, E. A., & Villareal Escate, L. G. (2021). *El Modelo Estocástico Univariante ARIMA Como Herramienta Predictiva De La Demanda De Energía Eléctrica Residencial Del Sistema Eléctrico Cusco*. Universidad Nacional de Ingeniería.
- Mauricio, J. A. (2007). Análisis De Series Temporales. In *Universidad Complutense De Madrid* (Vol. 8, Issue 1).
- Mejía Vasquez, E. J., & Gonzales Chávez, S. (2019). Predicción del consumo de energía eléctrica residencial de la Región Cajamarca mediante modelos Holt -



- Winters. *Ingeniería Energética*, 3, 1–2019.
<http://rie.cujae.edu.cu/index.php/RIE/article/view/569>
- MINEM. (2008). *Plan Referencial De Electricidad 2008-2017* (Electricidad).
https://www.academia.edu/8568453/MINISTERIO_DE_ENERGIA_Y_MINAS_PLAN_REFERENCIAL_DE_ELECTRICIDAD_2006_2015_165
- Ministerio de Energía y Minas. (2020, March). Anuario Ejecutivo De Electricidad 2019. *Dirección General De Electricidad, Dirección De Estudios Y Promoción Eléctrica, Suministro*, 1–65. www.minem.gob.pe
- OIEA. (n.d.). *Proyecciones de electricidad, cálculo de la demanda a largo plazo*. Retrieved March 6, 2023, from <https://www.iaea.org/es/temas/proyecciones-energeticas>
- OSINERGMIN. (2016). La Industria de la Electricidad en el Perú. *Osinergmin*, 178.
http://www.osinergmin.gob.pe/seccion/centro_documental/Institucional/Estudios_Economicos/Libros/Osinergmin-Industria-Electricidad-Peru-25anos.pdf
- Parra, F. (2019). *Estadística y machine learning con R*. Rpubs.
<https://bookdown.org/content/2274/series-temporales.html>
- Pichua Serna, J. (2017). Universidad Nacional Agraria La Molina. In *Universidad Nacional Agraria La Molina*.
- Pindyck, R. S., & Rubinfeld, D. L. (2018). *Econometría Modelos y Pronóstico* (R. del B. Alayón, A. G. Maya, & Z. G. García (eds.); Impresora, Vol. 4, Issue 1).
<http://dx.doi.org/10.1016/j.cirp.2016.06.001><http://dx.doi.org/10.1016/j.powtec.2016.12.055><https://doi.org/10.1016/j.ijfatigue.2019.02.006><https://doi.org/10.1016/j.matlet.2019.04.024><https://doi.org/10.1016/j.matlet.2019.127252><http://dx.doi.org/10.1016/j.cirp.2016.06.001>
- Quispe Colque, J. E. (2019). *Modelo Univariante Para Pronosticar La Cantidad De Ventas Mensuales De Bolsas De Cemento Rumi Producidos En La Planta Cemento Sur Del Distrito De Caracoto, Periodo 2005 - 2018* [Universidad Nacional del Altiplano-Puno].
http://repositorio.unap.edu.pe/bitstream/handle/UNAP/7104/Molleapaza_Mamani_Joel_Neftali.pdf?sequence=1&isAllowed=y
- Quispe Pacco, Y. L. R. (2015). *Modelo Univariado Para El Consumo De Energía Eléctrica Doméstica En El Distrito De Ayaviri - Electro Puno, Periodo 2004 - 2013* (Vol. 0, Issue 0). Universidad Nacional Del Altiplano-Puno.
- Rey Onzaga, J. (2018). *Pronóstico Del Consumo De Energia Electrica Residencial Para La Ciudad De Bogotá*. Universidad Santo Tomás.
- Ríos, G. (2008). *Series de Tiempo*. https://www.ucursos.cl/ingenieria/2010/1/CC52A/1/material_docente/objeto/296003
- S.A.A, E. P. (2020). Memoria Anual 2020. In *Empresa Regional de Servicios Públicos*



de Electricidad - Electro Puno S.A.A.

Sanchez de Rivera, D. P. (2005). *Análisis de series temporales* (2005 Alianza Editorial, S.A., Madrid (ed.)).

Sudhan, S. P., Mathew, B. S., Rose, S., & Isaac, K. P. (2020). Development of pavement deterioration prediction models for low volume roads using system dynamics. *Journal of Stomatology*, 146(3).
<https://doi.org/10.1061/JPEODX.0000170>

Vázquez Díaz, K. Y., & Gamonal Sánchez, M. E. (2019). *Modelo Para El Pronóstico Del Consumo Mensual De Energía Eléctrica, De La Provincia Bagua Grande, Mediante La Metodología De Box y Jenkins, Para El Año 2016*. Universidad Nacional Pedro Ruiz Gallo - Lambayeque.

Víctor Morales, O. (2021). *Series De Tiempo*.
https://bookdown.org/victor_morales/SeriesdeTiempo/modelización-de-series-univariantes-sarima.html#procesos-sarima

Villagarcía, T. (2020). *Series Temporales*. 1–40.
<http://www.fao.org/faostat/es/#data/OA>

Villarreal, F. (2016). Introducción a los modelos de pronósticos. In *Universidad Nacional del Sur* (pp. 1–121).
https://www.matematica.uns.edu.ar/uma2016/material/Introduccion_a_los_Modelos_de_Pronosticos.pdf

Villavicencio, J. (2010). *Introducción a Series de Tiempo*.
http://www.estadisticas.gobierno.pr/iepr/LinkClick.aspx?fileticket=4_BxecUaZmg%3D

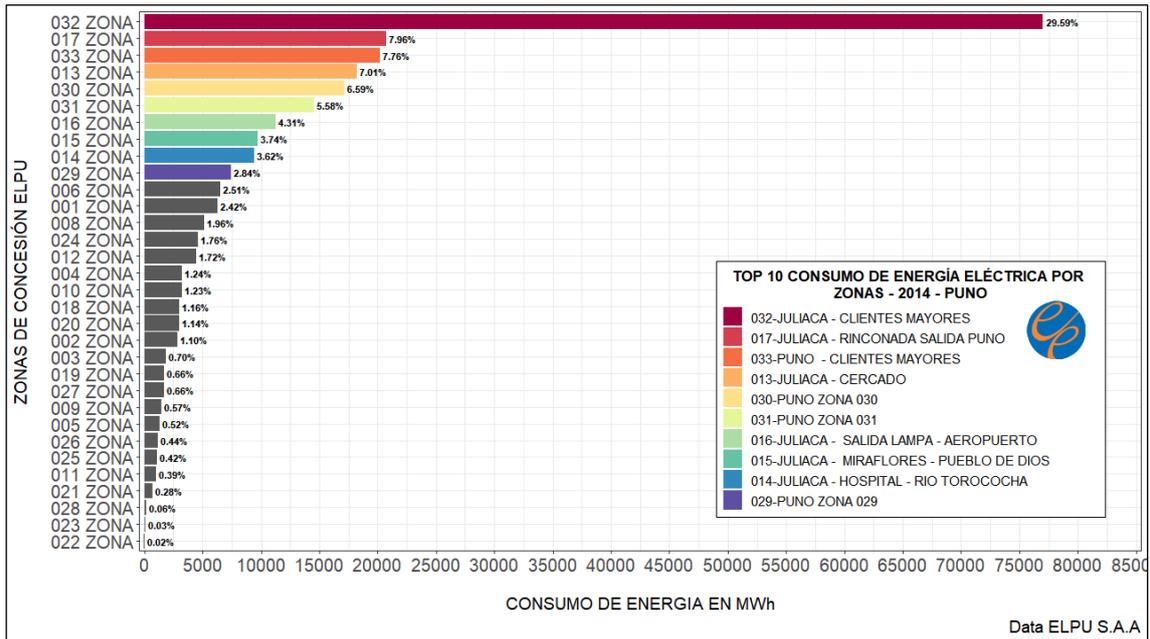
Wickham, H., & Grolemund, G. (2017). *R for data Science* (M. Beaugureau & M. Loukides (eds.)). <https://es.r4ds.hadley.nz/visualización-de-datos.html>



ANEXOS

Anexo A

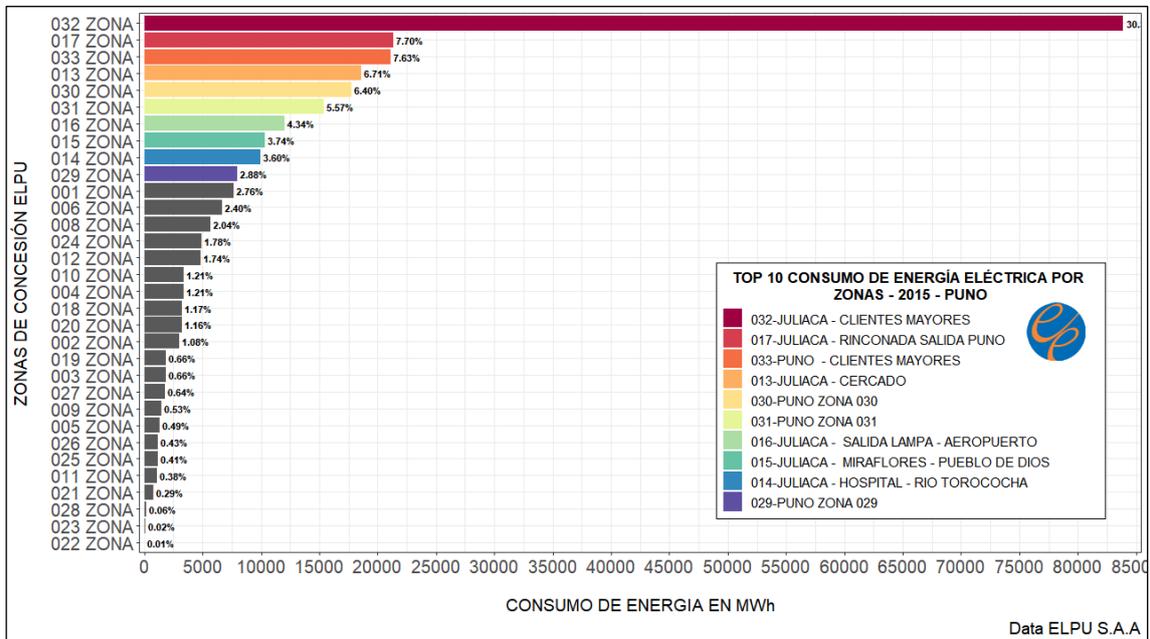
Consumo Energía Eléctrica de La Región De Puno del Año 2014-ELPU S.A.A



Nota. Realizado por el autor en el software R v.4.2.1 Fuente: Data ELPU S.A.A.

Anexo B

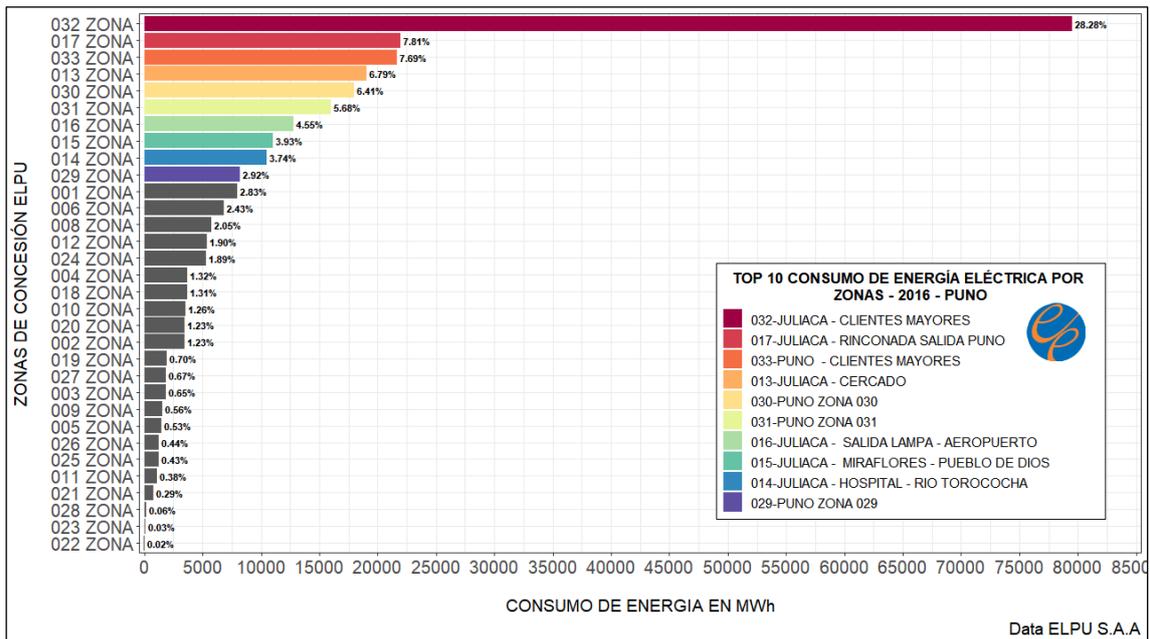
Consumo Energía Eléctrica de La Región De Puno del Año 2015-ELPU S.A.A



Nota. Realizado por el autor en el software R v.4.2.1 Fuente: Data ELPU S.A.A.

Anexo C

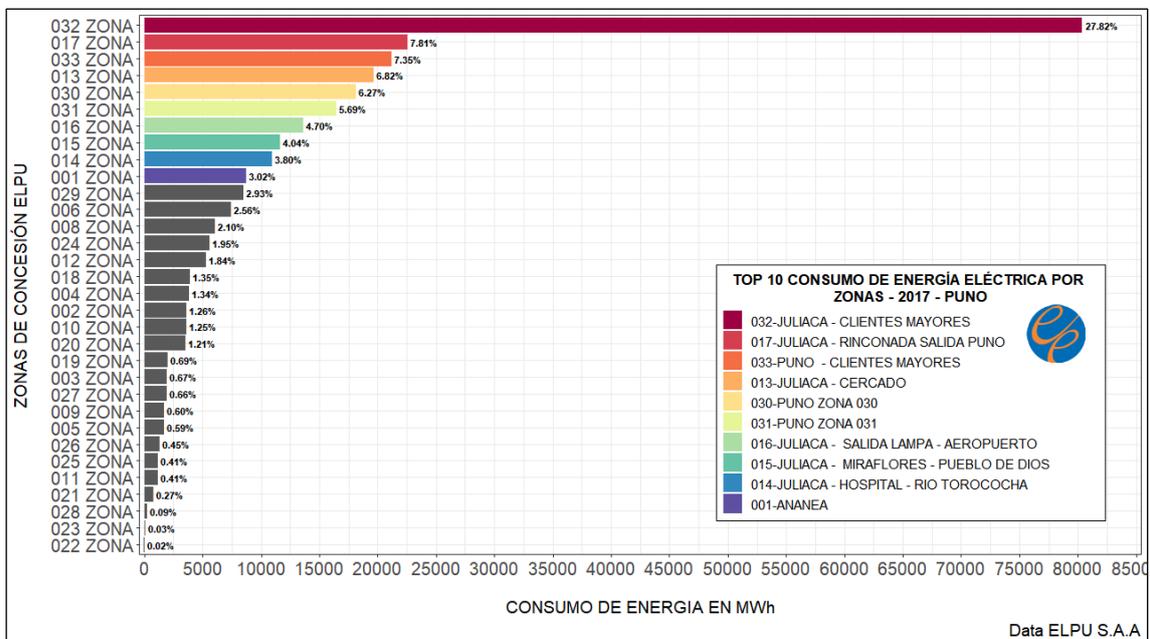
Consumo Energía Eléctrica de La Región De Puno del Año 2016-ELPU S.A.A



Nota. Realizado por el autor en el software R v.4.2.1 Fuente: Data ELPUS.A.A.

Anexo D

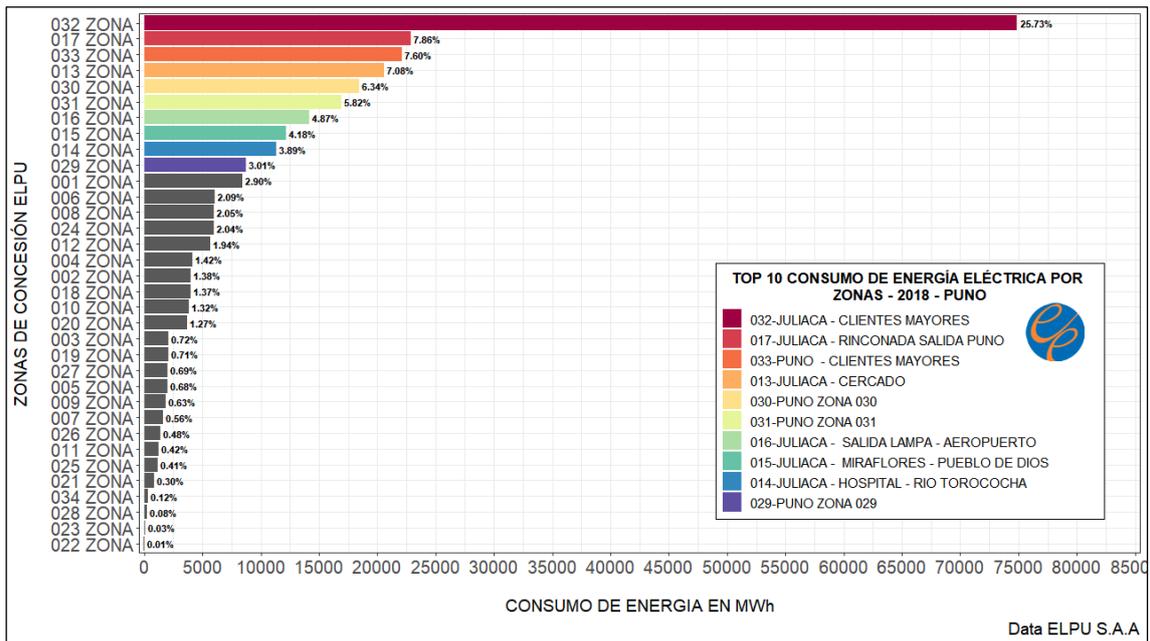
Consumo Energía Eléctrica de La Región De Puno del Año 2017-ELPU S.A.A



Nota. Realizado por el autor en el software R v.4.2.1 Fuente: Data ELPUS.A.A.

Anexo E

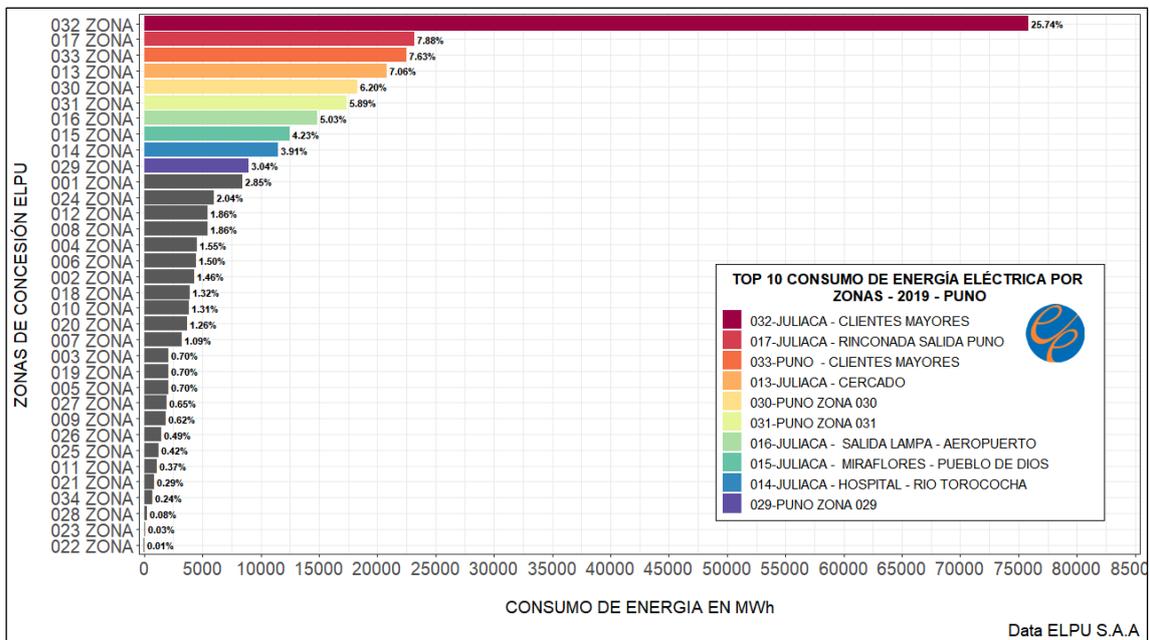
Consumo Energía Eléctrica de La Región De Puno del Año 2018-ELPU S.A.A



Nota. Realizado por el autor en el software R v.4.2.1 Fuente: Data ELPUS.A.A.

Anexo F

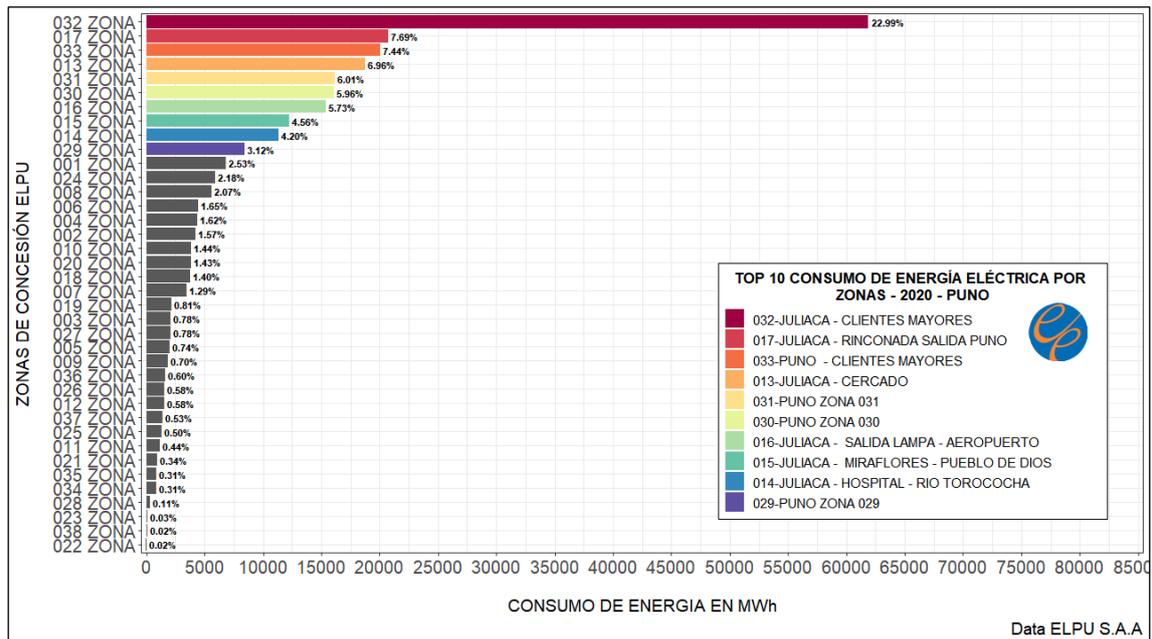
Consumo Energía Eléctrica de La Región De Puno del Año 2019-ELPU S.A.A



Nota. Realizado por el autor en el software R v.4.2.1 Fuente: Data ELPUS.A.A..

Anexo G

Consumo Energía Eléctrica de La Región De Puno del Año 2020-ELPU S.A.A



Nota. Realizado por el autor en el software R v.4.2.1 Fuente: Data ELPU S.A.A.

Anexo H

Código fuente para la determinación del modelo ARIMA en el software r

```
rm(list = ls())
options(scipen = 999)
#Cargamos Las Librerías que vamos a utilizar en el desarrollo del modelo

library(forecast)
library(tseries)
library(ggplot2)
library(tidyverse)
library(tidyr)
library(dplyr)
library(readxl)
library(readxl)
library(astsa)
library(flextable)
library(nortest)
library(lmtest)

#importacion de La data_general
energia<- read_excel("~/rrr.xlsx")
head(energia,10)

## # A tibble: 10 × 9
##   clave mes      año zona      cliente consumo mon_c...1 consu...2 monto...3
##   <dbl> <chr>    <dbl> <chr>      <dbl> <dbl> <dbl> <dbl> <dbl>
```



```
## 1      1 enero      2014 006 AZANGARO U... 18760 568209 365550. 549880. 353758.  
## 2      2 febrero   2014 006 AZANGARO U... 18760 532484 347935. 515307. 336711.  
## 3      3 marzo     2014 006 AZANGARO U... 18776 469836 320682. 503396. 343588.  
## 4      4 abril     2014 006 AZANGARO U... 18767 513663 345734. 497093. 334582.  
## 5      5 mayo      2014 006 AZANGARO U... 18760 554411 373952. 554411 373952.  
## 6      6 junio     2014 006 AZANGARO U... 18749 544526 364642. 526961. 352879.  
## 7      7 julio     2014 006 AZANGARO U... 18744 561369 380185. 561369 380185.  
## 8      8 agosto    2014 006 AZANGARO U... 18756 565059 378329. 546831. 366125.  
## 9      9 setiembre 2014 006 AZANGARO U... 18787 572269 376444. 553809. 364300.  
## 10    10 octubre   2014 006 AZANGARO U... 18793 524585 352028. 524585 352028.  
## # ... with abbreviated variable names 1mon_consumo, 2consumo_mes, 3monto_mes
```

```
#Cambio de formato en nuestra data base
```

```
energia$mes<- factor(energia$mes)  
energia$año<- factor(energia$año)  
energia$zona <- factor(energia$zona)  
energia$consumo <- as.numeric(energia$consumo, options(digits = 10))  
energia$consumo <- round(energia$consumo,2)
```

```
#Conversion de nuestra variable consumo de energia de KWh a MWh
```

```
energia$consumo <- energia$consumo/1000  
str(energia)
```

```
## tibble [2,795 × 9] (S3: tbl_df/tbl/data.frame)  
## $ clave      : num [1:2795] 1 2 3 4 5 6 7 8 9 10 ...  
## $ mes        : Factor w/ 12 levels "abril","agosto",...: 4 5 8 1 9 7  
6 2 12 11 ...  
## $ año        : Factor w/ 7 levels "2014","2015",...: 1 1 1 1 1 1 1 1  
1 1 ...  
## $ zona       : Factor w/ 38 levels "006 AZANGARO URBANO",...: 1 1 1 1  
1 1 1 1 1 1 ...  
## $ cliente    : num [1:2795] 18760 18760 18776 18767 18760 ...  
## $ consumo    : num [1:2795] 568 532 470 514 554 ...  
## $ mon_consumo: num [1:2795] 365550 347935 320682 345734 373952 ...  
## $ consumo_mes: num [1:2795] 549880 515307 503396 497093 554411 ...  
## $ monto_mes  : num [1:2795] 353758 336711 343588 334582 373952 ...
```

```
View(energia)
```

```
#Selección y filtración de consumos mensuales de energía eléctrica de La Base de datos
```

```
energia_anual <- energia %>%  
  select( ,clave,año,mes,consumo) %>%  
  group_by(año,mes) %>%  
  summarise(  
    clave = sum(clave),  
    consumo = sum(consumo)) %>%  
  arrange(clave)
```

```
#transformacion del tipo de variable a una serie temporal
```

```
serie_energia<- ts(energia_anual$consumo, start = c(2014,1), frequency =  
12 )  
view(serie_energia)
```

```
#Grafico de la serie de tiempo
```

```
autoplot(serie_energia)
```



```
#Descomposicion de La serie
desc.energia <- decompose(serie_energia, type="multiplicative")
autoplot(desc.energia)

#Aplicamos la primera diferencia en 1 periodo a la serie original
serie_energia_1 <- diff(serie_energia)

#Serie estacionaria en primera diferencia
autoplot(serie_energia_1)

#Prueba de Dikey - fuller poara comprobar la estaionariedad en media y v
arianza
prueba_adf <- adf.test(serie_energia_1, alternative = "stationary")
prueba_adf %>% as_flextable()

#Determinacion del modelo ARIMA
modelo_determinado <- auto.arima(serie_energia, seasonal = T, ic = "aic"
, stepwise = FALSE, allowdrift = F , approximation=FALSE, trace = F)
summary(modelo_determinado)

## Series: serie_energia
## ARIMA(0,1,3)(2,0,0)[12]
##
## Coefficients:
##          ma1          ma2          ma3          sar1          sar2
##      -0.2257977  0.0311740 -0.5167989  0.2631069  0.3695910
## s.e.   0.1136758  0.1230245  0.0942766  0.1239784  0.1382097
##
## sigma^2 = 859900: log likelihood = -685.64
## AIC=1383.29  AICc=1384.39  BIC=1397.8
##
## Training set error measures:
##              ME          RMSE          MAE          MPE          MAPE
## Training set -1.942247335  893.5762801  625.7998838 -0.161869089  2.7897
14349
##              MASE          ACF1
## Training set 0.5977807825  0.06704700804

#Coeficientes de modelo determinado

coefptest(modelo_determinado)
## z test of coefficients:
##
##          Estimate  Std. Error  z value      Pr(>|z|)
## ma1  -0.225797741  0.113675795 -1.98633    0.0469966 *
## ma2   0.031173960  0.123024513  0.25340    0.7999620
## ma3  -0.516798943  0.094276583 -5.48173  0.000000042118 ***
## sar1  0.263106924  0.123978389  2.12220    0.0338210 *
## sar2  0.369591024  0.138209711  2.67413    0.0074923 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#Generacion de pronóstico de energía eléctrica para Los próximos tres añ
os
modelo_pronosticado <- forecast(modelo_determinado, level = c(95), h =36
```



)

*#Grafico del pronóstico de La demanda de energía en Los próximos 3 años
autoplot(modelo_pronosticado)*

*#Valores generados por el modelo ARIMA
print(modelo_pronosticado)*

##		Point Forecast	Lo 95	Hi 95
##	Jan 2021	24187.44051	22369.95037	26004.93066
##	Feb 2021	23732.08210	21433.55761	26030.60660
##	Mar 2021	21907.93599	19182.90058	24632.97140
##	Apr 2021	22640.03225	19864.98199	25415.08250
##	May 2021	21401.36547	18577.18597	24225.54497
##	Jun 2021	22514.15510	19641.68652	25386.62369
##	Jul 2021	22622.58396	19702.62476	25542.54315
##	Aug 2021	23300.90284	20334.21316	26267.59252
##	Sep 2021	23421.16891	20408.47350	26433.86431
##	Oct 2021	23408.42088	20350.41180	26466.42996
##	Nov 2021	23436.82933	20334.16830	26539.49035
##	Dec 2021	23394.12296	20247.44353	26540.80238
##	Jan 2022	23350.53785	20047.96919	26653.10652
##	Feb 2022	23189.24698	19767.63076	26610.86320
##	Mar 2022	22095.68179	18555.22240	25636.14118
##	Apr 2022	21680.63473	18078.72752	25282.54194
##	May 2022	20052.27787	16389.95368	23714.60205
##	Jun 2022	20931.16848	17209.40797	24652.92899
##	Jul 2022	21363.09362	17582.83118	25143.35606
##	Aug 2022	22476.88049	18639.00778	26314.75321
##	Sep 2022	22475.82393	18581.19303	26370.45483
##	Oct 2022	22898.50845	18947.93473	26849.08217
##	Nov 2022	23110.81355	19105.07822	27116.54888
##	Dec 2022	23287.20202	19227.05443	27347.34960
##	Jan 2023	22985.34457	18670.66311	27300.02603
##	Feb 2023	22774.61145	18274.08592	27275.13697
##	Mar 2023	21812.69884	17126.84273	26498.55495
##	Apr 2023	21974.07330	17203.95321	26744.19338
##	May 2023	21087.84121	16234.92005	25940.76236
##	Jun 2023	21730.36047	16796.02750	26664.69344
##	Jul 2023	21884.07730	16869.65410	26898.50049
##	Aug 2023	22427.82291	17334.56873	27521.07708
##	Sep 2023	22471.99418	17301.11067	27642.87768
##	Oct 2023	22578.49384	17331.12933	27825.85836
##	Nov 2023	22644.85229	17322.10559	27967.59900
##	Dec 2023	22675.47743	17278.40131	28072.55354



AUTORIZACIÓN PARA EL DEPÓSITO DE TESIS O TRABAJO DE INVESTIGACIÓN EN EL REPOSITORIO INSTITUCIONAL

Por el presente documento, Yo DENNIS URIEL DÍASCO CHATA
identificado con DNI 47479350 en mi condición de egresado de:

Escuela Profesional, Programa de Segunda Especialidad, Programa de Maestría o Doctorado
INGENIERÍA ESTADÍSTICA E INFORMÁTICA

informo que he elaborado el/la Tesis o Trabajo de Investigación para la obtención de Grado

Título Profesional denominado:

"MODELO DE SERIES DE TIEMPO PARA EL PRONÓSTICO MENSUAL DE CONSUMO DE ENERGÍA ELÉCTRICA DE LA REGIÓN DE PUNO, ELECTRO PUNO S.A.A PERÍODO 2014-2020"

Por medio del presente documento, afirmo y garantizo ser el legítimo, único y exclusivo titular de todos los derechos de propiedad intelectual sobre los documentos arriba mencionados, las obras, los contenidos, los productos y/o las creaciones en general (en adelante, los "Contenidos") que serán incluidos en el repositorio institucional de la Universidad Nacional del Altiplano de Puno.

También, doy seguridad de que los contenidos entregados se encuentran libres de toda contraseña, restricción o medida tecnológica de protección, con la finalidad de permitir que se puedan leer, descargar, reproducir, distribuir, imprimir, buscar y enlazar los textos completos, sin limitación alguna.

Autorizo a la Universidad Nacional del Altiplano de Puno a publicar los Contenidos en el Repositorio Institucional y, en consecuencia, en el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto, sobre la base de lo establecido en la Ley N° 30035, sus normas reglamentarias, modificatorias, sustitutorias y conexas, y de acuerdo con las políticas de acceso abierto que la Universidad aplique en relación con sus Repositorios Institucionales. Autorizo expresamente toda consulta y uso de los Contenidos, por parte de cualquier persona, por el tiempo de duración de los derechos patrimoniales de autor y derechos conexos, a título gratuito y a nivel mundial.

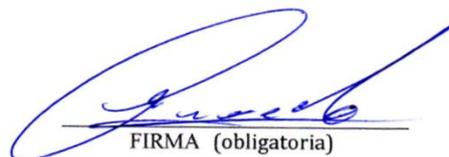
En consecuencia, la Universidad tendrá la posibilidad de divulgar y difundir los Contenidos, de manera total o parcial, sin limitación alguna y sin derecho a pago de contraprestación, remuneración ni regalía alguna a favor mío; en los medios, canales y plataformas que la Universidad y/o el Estado de la República del Perú determinen, a nivel mundial, sin restricción geográfica alguna y de manera indefinida, pudiendo crear y/o extraer los metadatos sobre los Contenidos, e incluir los Contenidos en los índices y buscadores que estimen necesarios para promover su difusión.

Autorizo que los Contenidos sean puestos a disposición del público a través de la siguiente licencia:

Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional. Para ver una copia de esta licencia, visita: <https://creativecommons.org/licenses/by-nc-sa/4.0/>

En señal de conformidad, suscribo el presente documento.

Puno 02 de JUNIO del 2023


FIRMA (obligatoria)



Huella



DECLARACIÓN JURADA DE AUTENTICIDAD DE TESIS

Por el presente documento, Yo DENNIS URIEL APASCO CHATA
identificado con DNI 47479350 en mi condición de egresado de:

Escuela Profesional, Programa de Segunda Especialidad, Programa de Maestría o Doctorado

INGENIERÍA ESTADÍSTICA E INFORMÁTICA

,informo que he elaborado el/la Tesis o Trabajo de Investigación para la obtención de Grado

Título Profesional denominado:

"MODELO DE SERIES DE TIEMPO, PARA EL PROMOSTICO MENSUAL DE CONSUMO DE ENERGIA

ELECTRICO DE LA REGION DE PUNO, ELECTRO PUNO S.A.A PERIODO 2014 - 2020 .

" Es un tema original.

Declaro que el presente trabajo de tesis es elaborado por mi persona y **no existe plagio/copia** de ninguna naturaleza, en especial de otro documento de investigación (tesis, revista, texto, congreso, o similar) presentado por persona natural o jurídica alguna ante instituciones académicas, profesionales, de investigación o similares, en el país o en el extranjero.

Dejo constancia que las citas de otros autores han sido debidamente identificadas en el trabajo de investigación, por lo que no asumiré como tuyas las opiniones vertidas por terceros, ya sea de fuentes encontradas en medios escritos, digitales o Internet.

Asimismo, ratifico que soy plenamente consciente de todo el contenido de la tesis y asumo la responsabilidad de cualquier error u omisión en el documento, así como de las connotaciones éticas y legales involucradas.

En caso de incumplimiento de esta declaración, me someto a las disposiciones legales vigentes y a las sanciones correspondientes de igual forma me someto a las sanciones establecidas en las Directivas y otras normas internas, así como las que me alcancen del Código Civil y Normas Legales conexas por el incumplimiento del presente compromiso

Puno 02 de JUNIO del 2023


FIRMA (obligatoria)



Huella