

Método de Box-Jenkins para el pronóstico de consumo de energía eléctrica del 2021-2023

Box-Jenkins method for electricity consumption forecast for 2021-2023

Dennis Uriel Añasco-Chata^A, Percy Huata-Panca^B, Ecler Mamani-Vilca^C, Adolfo Carlos Jimenez-Chura^D, y Pablo Cesar Tapia-Catacora^E

ORCID 0009-0006-7254-5837^A, ORCID 0000-0002-1624-5526^B, ORCID 0000-0002-5205-3660^E, ORCID 0000-0003-3125-8197^D y ORCID 0000-0002-8608-4203^E

(Recepción: 23/09/2023 y aceptación 06/10/2023)

Resumen— El trabajo fue determinar un modelo de serie de tiempo basado en el método de Box-Jenkins, que se ajuste apropiadamente a la serie original en el pronóstico del consumo de energía eléctrica de la región de Puno – Perú, para los próximos tres años (2021-2023), esto permitiría a la empresa Electro Puno contar con el modelo que respalde la toma de decisiones, que se ajusten a sus objetivos corporativos en un horizonte de tiempo a mediano plazo, garantizando la distribución y la comercialización de energía eléctrica. El estudio fue cuantitativo con diseño no experimental, de corte longitudinal de tendencia (trend), la población fue, 84 reportes de consumo mensual de energía eléctrica representados en (MWh/mes), considerando un periodo comercial de 7 años, se aplicó las tres fases de la metodología Box-Jenkins como: identificación, estimación y validación del modelo, llegando a determinar un modelo multiplicativo estacional ARIMA (0,1,3)×(2,0,0)₁₂ que se ajustó perfectamente a la serie en estudio, siendo 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, con una desviación en promedio por cada pronóstico de un valor MAE =625.80 MWh.

Palabras Clave: Box-Jenkins, Pronóstico, Energía eléctrica

Abstract— The work was to determine a time series model based on the Box-Jenkins method, which is appropriately adjusted to the original series in the forecast of electricity consumption in the region of Puno Peru, for the next three years (2021-2023), this would allow the company Electro Puno to have the model to support decision making, which conform to its corporate objectives in a medium-term time horizon, ensuring the distribution and marketing of electricity. The study was quantitative with non-experimental design, trend longitudinal cut, the population was 84 reports of monthly electricity consumption represented in (MWh/month), considering a commercial period of 7 years, the three phases of the Box-Jenkins methodology were applied as: identification, estimation and validation of the model, arriving to determine a multiplicative seasonal ARIMA (0,1,3)×(2,0,0)₁₂ model that adjusted perfectly to the series under study, being of "high precision" with a MAPE value of 2.79%, finally in the last phase of the methodology, monthly forecasts of electricity consumption were made, with an average deviation for each forecast of a MAE value =625.80 MWh.

Keywords: Box-Jenkins, Forecasting, Electric Power

- A. Dennis Uriel Añasco-Chata, Facultad de Ingeniería Estadística e Informática-UNA-Puno, rusbel92@hotmail.com
B. Percy Huata-Panca, Facultad de Ingeniería Estadística e Informática-UNA-Puno, phuata@unap.edu.pe
C. Ecler Mamani-Vilca, Departamento de Ing. Informática y Sistemas- UNAMBA-Apurímac, eclervirtual@unamba.edu.pe
D. Adolfo Carlos Jimenez-Chura, Escuela Profesional de Ingeniería de Sistemas -UNA-Puno, ajimenez@unap.edu.pe
E. Pablo Cesar Tapia-Catacora, Escuela Profesional de Ingeniería de Sistemas -UNA-Puno, pctapia@unap.edu.pe

1 INTRODUCCIÓN

La energía eléctrica es un elemento fundamental en el desarrollo de una sociedad, según el Banco Mundial [1]. La energía es esencial para hacer posibles las inversiones, las innovaciones y las nuevas industrias, que son los motores del empleo, incluido el crecimiento, y la prosperidad compartida de economías enteras [1]. Sin embargo, casi 675 millones de personas todavía viven sin electricidad en todo el mundo, y alrededor de 2.3 millones de personas cocinan o calientan sus hogares con combustibles contaminantes que dañan su salud

y el medio ambiente. Las crisis energéticas afectan a la mayoría de los países, pero los países en vías desarrollo enfrentan las mayores cargas, especialmente los países importadores de energía [2]. Desde su aparición de la energía eléctrica en la humanidad, su demanda se encuentra en constante crecimiento el crecimiento de la vivienda y la población puede aumentar la demanda de electricidad, ya que más viviendas y edificios requieren electricidad para iluminación, calefacción y refrigeración [3].

Es fundamental tener información acerca del comporta-

miento futuro del consumo de energía eléctrica, resulta imperativo desarrollar un modelo matemático adecuado y confiable, que nos proporcione resultados certeros sobre el consumo de energía en los próximos tres años. Este modelo nos permitirá contar con estimaciones confiables y realistas, lo cual resulta esencial para poder planificar y tomar decisiones frente a una necesidad a futuro. En ese marco, la pregunta es ¿Cuál es el modelo determinado con la metodología Box - Jenkins para el pronóstico 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.

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.

En la región de Puno, el incremento en la demanda de energía eléctrica en la zona urbana y rural se fundamenta en el crecimiento 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 [4].

Así también se encontraron diversos trabajos anteriores como: En Bagua Vázques y Gamonal [5] realizaron el pronóstico de consumo mensual de energía eléctrica, de la provincia 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, se concluye 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 $R^2=0,996$. En Cusco Marcelo Barreto y Villareal Escate [6] logran determinar el modelo multiplicativo estacional ARIMA(0,1,1)×(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. En Puno, distrito de Ayaviri [4] realizó el modelo de pronóstico para el consumo de energía eléctrica para el distrito de Ayaviri, el objetivo fue determinar un modelo univariante que describa y pronostique el consumo de energía eléctrica, indicando que el modelo permite describir y predecir el comportamiento de consumo de energía eléctrica, es el modelo multiplicativo estacional ARIMA(3,1,1)×(1,0,1)₁₂.

Para entender ampliaremos algunas definiciones:

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 [7].

La Metodología Box Y Jenkins es un modelo de series de tiempo, cuya función representa 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 [8]. La 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 [9].

En la década de los 70' en el año de 1973 George Edward Pelham Box y Gwilym Mirion Jenkins publican el libro “Análisis de Series Temporales Pronóstico y Control”, 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 [10], resumimos dichas fases en el diagrama de la Fig 1.

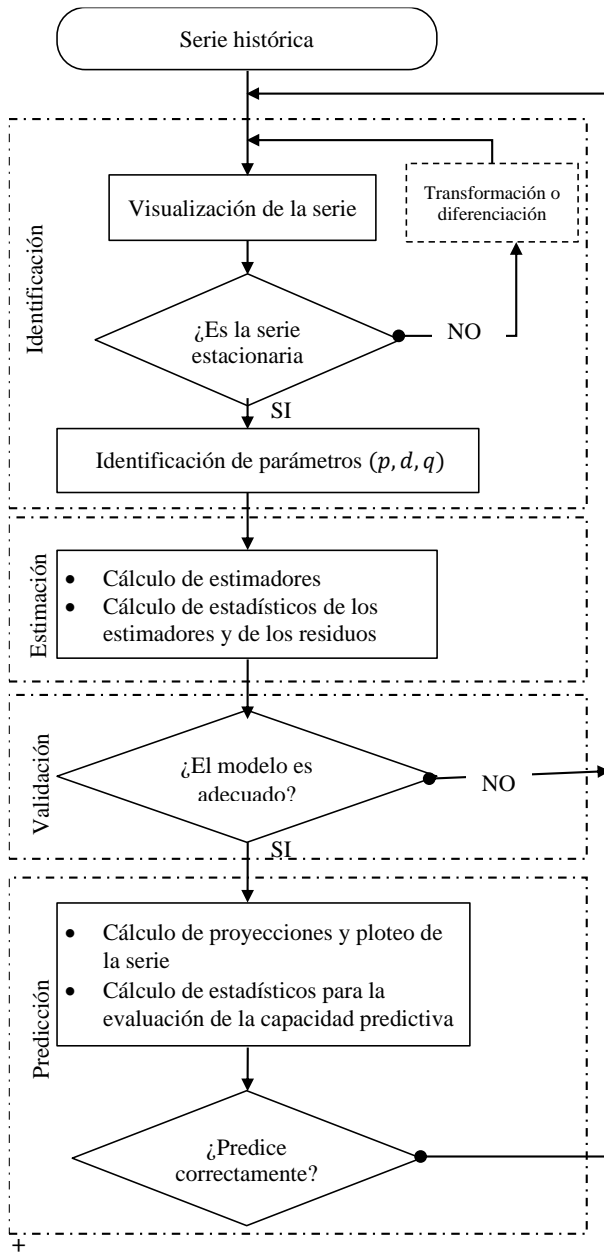


Fig. 1. Diagrama de flujo de la metodología Box-Jenkins.

2 MATERIALES Y METODOS

La investigación se desarrolló 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, 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 [11]. EL ámbito de estudio fue en la empresa ELECTRO PUNO S.A.A quien 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 la concesión dentro de la región de Puno, a cargo de la operación de la Central Hidroeléctrica de Chijisia localizada en la Provincia de Sandia [12].

El estudio es cuantitativo, cuya finalidad es “ explicar, predecir, verificar y controlar los fenómenos” [13], con diseño no experimental, de corte longitudinal de tendencia (trend), este tipo de diseño se basa en la recolección de los datos a través del tiempo en puntos o periodos [14]. Para el procesamiento y análisis de información de la serie de consumo de energía eléctrica, se usó el software libre R con las siguientes librerías: forecast, tseries, ggplot2, tidyverse, tidyr, dplyr, readxl, readxl, astsa, flextable, nortest, lmttest siendo las dos primeras las utilizadas para la estimación. La población fue, 84 reportes de consumo mensual de energía eléctrica representados en (MWh/mes), considerando un periodo comercial de 7 años que comprende desde enero-2014 a diciembre-2020 [13] ver Tabla 1. Para la obtención de la data se aplicó la técnica de revisión documental, teniendo acceso al software SIELSE COMERCIAL propiedad de ELECTRO-PUNO S.A.A, proporcionado los resúmenes de facturación mensualmente por cada periodo comercial.

TABLA 1
Consumo Mensual de Energía Eléctrica de la Región Puno 2014-2020

Mes	2014	2015	2016	2017	2018	2019	2020
Años							
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

Datos mensuales del consumo de energía eléctrica en Megavatios por hora (MWh) de la Región de Puno, período 2014 a 2020 tomados de Electro Puno.

3 RESULTADOS

3.1 Identificación

En el comportamiento de la serie consumo de energía eléc-

trica se identificaron 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. El análisis al comportamiento de la serie en estudio, se afirma que la serie consumo de energía eléctrica presenta una tendencia creciente positiva (cumpliendo el requisito) Figura 11, donde se observa la distribución de los valores individuales de la serie consumo de energía eléctrica; de el consumo máximo de energía 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 $\bar{x} = 23,356.93 MWh$ (Línea horizontal verde) con una desviación estándar de $\sigma = 1,570.98$ y una mediana de $\hat{x} = 23,817.16$ (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 2, 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 los valores individuales de la serie no giran en torno a su promedio y varianza, por lo expuesto la serie es "no estacionaria".

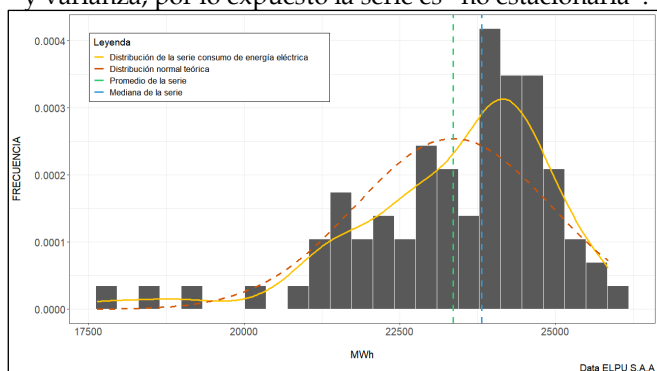


Fig. 2. Histograma de la Serie Consumo de energía eléctrica de la Región Puno, frente al histograma teórico de la serie.

Por el análisis visual se determinó "No estacionaria" para validar usaremos el Test Aumentado de Dickey - Fuller, este comprobará a través de un contraste de hipótesis si la serie presenta una tendencia significativa. El resultado con un nivel de di significancia $\alpha=0.05$ la decisión basados en al Tabla 3 fue, se dispone de evidencia suficiente para aceptar la hipótesis nula. 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.

TABLA 2

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

Vázquez [5] 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.

3.2 Segunda etapa de la fase identificación

Diferenciación de la serie 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.



Fig. 3. En la parte superior serie irregular el grafico inferior 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 validado por Test aumentado de Dickey - Fuller donde se obtuvo $P=0.0107 < 0.05$ con un nivel de significancia del 5% se dispone de evidencia suficiente para rechazar la hipótesis nula. Por lo tanto, la serie transformada consumo de energía eléctrica es estacionaria; 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 $ARIMA(p,d,q) \times (P,D,Q)_{12}$.

Este análisis es respaldado por Quispe [4] quien refiere que, luego de haber estudiado la serie original de consumo energía eléctrica de la provincia de Ayaviri, se determinó que es ne-

cesario diferenciar la serie en un periodo $d = 1$, para conseguir estabilizar la serie original en una serie estacionaria. Con los correlogramas mostrados en la **Figura 4** 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.

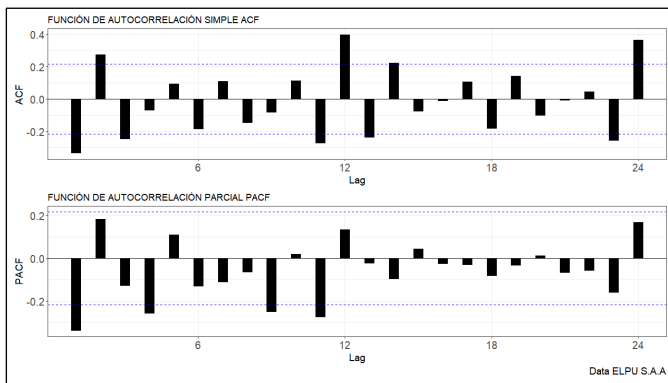


Fig. 4. Función de autocorrelación simple y parcial de la primera diferencia de la serie.

Se presenta la **Fig. 4**, 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 de 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. Por su parte Donado & Huertas 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 [16]" .

3.3 Tercera etapa de la fase de identificación

Se generaron 71 modelos candidatos, véase en la **Tabla 3**, para el pronóstico de consumo de energía eléctrica, 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:

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

TABLA 3

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

71 modelos identificados en el software R con la función auto.arima()

3.4 Estimación

Para la estimación de parámetros según el modelo multiplicativo estacional $ARIMA(0,1,3) \times (2,0,0)_{12}$, 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 (1).

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

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 4**, 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) [17], por lo tanto el coeficiente si aporta al modelo para el ajuste de la serie.

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

TABLA 4
Coeficientes del modelo ARIMA

ARIMA(0,1,3) × (2,0,0) ₁₂				
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	0.0000004212 ***
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

Logrado estimar los parámetros del modelo y reemplazando los estimadores de los coeficientes del modelo multiplicativo estacional ARIMA(0,1,3) × (2,0,0)₁₂ en la ecuación **¡Error! No se encuentra el origen de la referencia.**, la estimación del modelo, es la siguiente:

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

El modelo seleccionado fue validado mediante el análisis de los correlogramas simples (ACF) y Parcial (PACF), donde los residuos del modelo 1 a través de la significancia de los rezagos, así determinar si el modelo se ajusta razonablemente a los datos de la serie original.

La Fig. 5 muestra los correlogramas simples y parcial del modelo multiplicativo estacional ARIMA(0,1,3) × (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 indica que los residuos del modelo seleccionado son puramente aleatorios y que los datos no están correlacionados una de la otra.

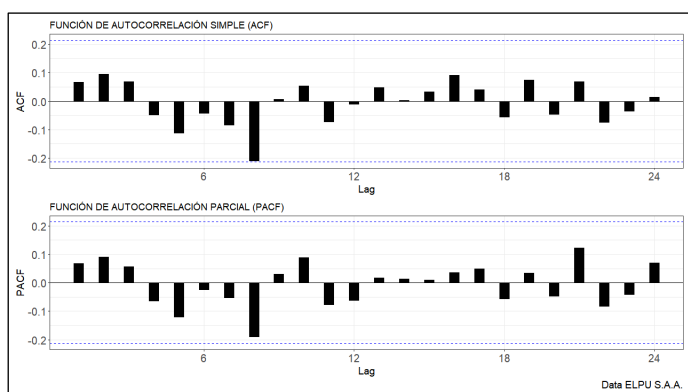


Fig. 5. Autocorrelación Simple y Parcial de los Residuos del Modelo Seleccionado Función de autocorrelación simple y parcial de los residuos del modelo ARIMA(0,1,3) × (2,0,0)₁₂

Por su parte Barahona [18] comprueba la existencia de independencia de su modelo ARIMA(0,1,1) × (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 [19] 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, 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.

Comprobamos estadísticamente a través del planteamiento de una hipótesis el supuesto de independencia mediante el Test de Q de Box - Lung donde se obtuvo P = 0.5316 ≥ 0.05 con un nivel de significancia del 5%, se acepta la hipótesis nula. Por lo tanto, los residuos del modelo seleccionado cumplen con el supuesto de independencia esto quiere decir que no existe correlación en los residuos, superado con éxito el proceso de validación del modelo, por tanto, el modelo es óptimo para ser implementado.

En los resultados obtenidos por Quispe [4] en la fase de validación de su modelo multiplicativo estacional ARIMA(3,1,1) × (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 α = 0.05, comprobando su hipótesis nula que los residuos de la serie son independientes, es decir son ruido blanco.

3.5 Pronóstico

La evaluación de la capacidad predictiva se observa en la Tabla 5 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 usó 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.

TABLA 5
Modelos ARIMA y sus Métricas de Precisión del Pronóstico

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

El valor seleccionado 625.80 en la métrica de evaluación MAE (error absoluto medio) es el indicando. El valor MAPE obte-

nido en 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 6**. Barahona [18] 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.

TABLA 6
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

Se determinó 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 son de alta precisión.

Barreto [6] 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 en Cusco, ambas investigaciones coinciden.

Ejecución de pronóstico, se observa en la Fig. 6 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 $ARIMA(0,1,3) \times (2,0,0)_{12}$ de la ecuación (3) es:

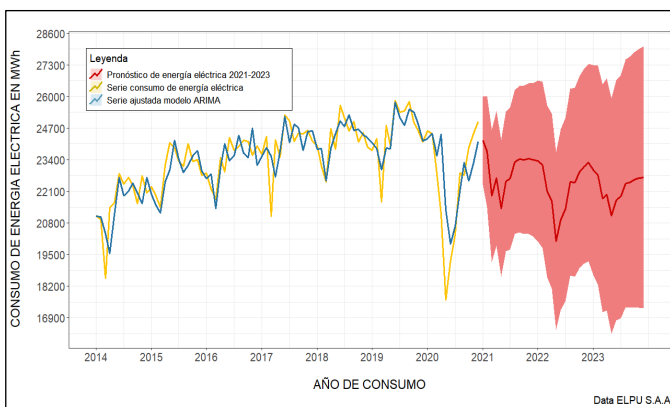


Fig. 6 Pronóstico del consumo de energía eléctrica para los próximos 3 años (2021 al 2023).

4 DISCUSIONES

El modelo determinado es 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 [18] 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 [4] 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. 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 Barreto [6] 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.

5 CONCLUSIONES

En el análisis a la serie consumo de energía eléctrica se muestra que factores atípicos como la pandemia en el año 2020 que 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.

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 (3)

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.

REFERENCIAS

- [1] Banco Mundial, «<https://www.worldbank.org>,» 09 09 2023. [En línea]. Available: <https://www.worldbank.org>.
- [2] B. Mundial, «<https://www.worldbank.org>,» 12 09 2022. [En línea]. Available: <https://www.worldbank.org/en/news/feature/2018/04/18/access-energy-sustainable-development-goal-7>.
- [3] M. Phillips, «<https://marketrealist.com>,» Market Realist, 10 10 2022. [En línea]. Available: <https://marketrealist.com/2014/09/must-know-factors-impact-electricity-demand/>.
- [4] Y. L. R. Quispe Pacco, «Modelo univariante para el consumo de energía eléctrica Doméstica en el Distrito de Ayaviri-Electro Puno, periodo 2004-2013,» UNA , Puno, 2015.
- [5] K. Y. Vásquez Díaz y M. E. Gamonal Sánchez , «Modelo para el pronóstico del consumo mensual de energía eléctrica, de la provincia de Bagua Grande, mediante la metodología Box Jenkins par eñ año 2016,» Universidad de Pedro Ruíz Gallo, Lambayeque, 2019.
- [6] L. G. Villarreal Escate y E. A. Marcelo Barreto, «El modelo estocástico univariante ARIMA como herramienta predictiva de la demanda de energía eléctrica residencial del sistema eléctrico Cusco,» Universidad Nacional de Ingeniería, Lima, 2021.
- [7] F. Villareal, «Introducción a los Modelos de Pronósticos,» Universidad Nacional del Sur, Bahia Blanca, 2016.
- [8] F. Parra, *Estadística y Machine Learning con R*, Madrid: <https://bookdown.org/>, 2019.
- [9] E. Court Monteverde y E. Williams Rengifo, *Estadísticas y econometría financiera*, Argentina: Cengage Learning Argentina, 2005.
- [10] D. Peña Sanchez de Rivera, *Análisis de series temporales*, Madrid: Alianza, 2005.
- [11] Instituto Nacioanal de Estadística e Informática, *Censos Nacionales 2017: XII de población, VII de vivienda y III de Comunidades Indígenas*, Lima: INEI, 2022.
- [12] Electro Puno S.A.A, *Plan estratégico Institucional 2017-2021*, Electro Puno S.A.A, 2021, p. 74.
- [13] E. E. Gallardo Echenique, *Metodología de la Onvestigación. Manual autoinformativo interactiva I*, vol. 1, Huancayo: Universidad Continental, 2017, p. 22.
- [14] R. Hernadez Sampieri, C. Fernández Collado y P. Batista Lucio, *Metodología de la Investigación*, Cuarta ed., Mexico: MACGRAW-HILL, 2010.
- [15] Electro Puno - S.A.A, *Memoria Anual 2020*, Empresa Regional de Servicios Públicos de Electricidad , 2020.
- [16] N. Huertas Gnecco y S. Donado Torres, «Evaluación crítica de la función Auto. Arima de R en la identificación de procesos de series de tiempo,» Universidad de los Andes, Bogotá , 2018.
- [17] R. J. Hyndman y G. Athanasopoulos, «orecasting: principles and practice,» OTexts, 2018.
- [18] J. I. Barahona Bonifaz, «Pronósticos del consumo y demanda de potencia máxima de energía eléctrica en la ciudad de Riobamba para el periodo 2017-2020,» Escuela Superior Politécnica de Chimborazo, Chimborazo, 2018.
- [19] K. Y. Vásquez Díaz y M. Estefany Gam, «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 Pedro Ruíz Gallo, Lambayeque, 2019.

Biografías:

Dennis Uriel Añasco-Chata, Universidad Nacional del Altiplano Puno – Perú. Ingeniero Estadístico e Informático, desarrollador de aplicaciones estadísticas e informáticas.

Percy Huata-Panca, Universidad Nacional del Altiplano Puno – Perú, Doctoris Scientiae en Economía y Gestión, Doctorado en Estadística Aplicada, Magister Scientiae en Informática, Ingeniero Estadístico, Docente nombrado a tiempo completo en la Universidad Nacional del Altiplano Puno – Perú.

Ecler Mamani-Vilca, Universidad Nacional Micaela Bastidas de Apurímac - Perú, Dr. en Ciencias de la Computación, desarrollador de aplicaciones multimedia y Software Educativo Intercultural, docente nombrado a tiempo completo en la Universidad Nacional Micaela Bastidas de Apurímac.

Adolfo Carlos Jimenez-Chura, Universidad Nacional del Altiplano Puno – Perú, Doctoris Scientiae en Ciencias de la Computación, Magister Scientiae en Informática, Ingeniero de Sistemas, Docente nombrado a tiempo completo en la Universidad Nacional del Altiplano Puno – Perú.

Pablo Cesar Tapia-Catacora, Universidad Nacional del Altiplano Puno – Perú, Doctoris Scientiae en Ciencias de la Computación, Magister en Contabilidad y Administración, Ingeniero de Sistemas, Docente nombrado a tiempo completo en la Universidad Nacional del Altiplano Puno – Perú.