

Stochastic Analysis of the Flow of the Vilcanota River in the Machupicchu Hydroelectric Power Plant using ARIMA models 2023.

Margarita Chevarria Moscoso¹, Josue Eliezer Alata Rey¹, Jorge Bojórquez Segura², José Serna-Landivar³

¹Universidad Nacional Agraria La Molina, Perú, mchevarria@lamolina.edu.pe, josue@lamolina.edu.pe.

²Universidad Privada del Norte SAC, Perú, bojorquez.ja@gmail.com.

³Universidad Tecnológica del Perú, Perú, U18101283@utp.edu.pe.

Abstract— The constant demand for energy consumption suggests that its supply be duly attended to for the economic and social development of the country. Since fossil fuels generate a large amount of carbon emissions and are also being depleted; The change towards renewable energies is essential.

However, the creation and commissioning of clean energy plants is a crucial challenge because it is difficult to obtain natural resources for continuous energy generation. Consequently, a forecast model is an important tool to anticipate energy generation, consumption and reserve. This document presents a flow forecasting model for hydroelectric energy production, using the autoregressive integrated moving average (ARIMA) model. The data involved was obtained from the Machupicchu Hydroelectric Power Plant.

Although the energy capacity of the hydroelectric plant is greatly affected by environmental variability, having a forecast model and a long-term plan that will greatly benefit the production of renewable energy to meet the continuous increase in the demand.

Keywords— ARIMA, Machupicchu Hydroelectric Power Plant, flow of the Vilcanota River.

Digital Object Identifier: (only for full papers, inserted by LACCEI).
ISSN, ISBN: (to be inserted by LACCEI).
DO NOT REMOVE

Análisis Estocástico del Caudal del Río Vilcanota en la central Hidroeléctrica de Machupicchu mediante modelos ARIMA 2023

Margarita Chevarria Moscoso¹, Josue Eliezer Alata Rey¹, Jorge Bojórquez Segura², José Serna-Landivar³

¹Universidad Nacional Agraria La Molina, Perú, mchevarria@lamolina.edu.pe, josue@lamolina.edu.pe.

²Universidad Privada del Norte SAC, Perú, bojorquez.ja@gmail.com.

³Universidad Tecnológica del Perú, Perú, U18101283@utp.edu.pe.

Resumen– La constante demanda de consumo de energía, sugiere que su abastecimiento sea atendido debidamente, para el desarrollo económico y social del país. Dado que los combustibles fósiles generan una gran cantidad de emisiones de carbono y además se están agotando; el cambio hacia energías renovables es fundamental.

No obstante, la creación y puesta en marcha de plantas de energía limpia es un desafío crucial debido a que es difícil conseguir recursos naturales para la generación de energía de manera continua. En consecuencia, un modelo de pronóstico es una herramienta importante para anticipar la generación, el consumo y la reserva de energía. En este documento se presenta un modelo de pronóstico del caudal para la producción de energía hidroeléctrica, mediante el modelo de media móvil integrada autorregresiva (ARIMA). Los datos involucrados se obtuvieron de la central Hidroeléctrica de Machupicchu.

A pesar de que la capacidad de energía de la central hidroeléctrica se ve sumamente afectada por la variabilidad ambiental, disponer de un modelo de pronóstico y un plan a largo plazo que beneficiará en gran medida la producción de energía renovable para atender el continuo aumento de la demanda.

Palabras clave: ARIMA, Central Hidroeléctrica de Machupicchu, caudal del río Vilcanota.

I. INTRODUCCIÓN

El empleo de energías renovables resulta esencial para disminuir las emisiones de gases de efecto invernadero, con el propósito de prevenir el cambio climático. La energía hidroeléctrica es, en la actualidad, la principal fuente renovable que contribuye al suministro de electricidad y se prevé que su contribución aumente significativamente. No obstante, la energía hidroeléctrica es una de las fuentes de energía con mayor probabilidad de verse afectadas por las condiciones ambientales y la variabilidad climática, ya que la cantidad de electricidad generada está íntimamente vinculada con la cantidad de agua [1]–[4]. Para mantener una generación de energía sostenible y suficiente, es importante realizar una gestión eficiente de las centrales hidroeléctricas, así como realizar una eficaz planificación energética a corto y largo plazo [5][6].

La predicción de caudal es importante para la generación de energía, uno de los métodos más relevantes y eficientes

para la predicción del caudal es la metodología ARIMA. La investigación [7] que estudió los caudales de la “Presa Three Georges” para la planificación y gestión eficiente de los recursos hídricos utilizando la metodología ARIMA. De manera similar [8], [9], realizaron estudios de predicción del caudal.

La Central Hidroeléctrica de Machupicchu es una de las centrales de generación hidroeléctrica más importante del sur del Perú, esta se encuentra ubicada en la provincia de Urubamba del departamento del Cusco y utiliza los recursos hídricos que el río Vilcanota le proporciona. El río Vilcanota nace en la cordillera del Nudo del Vilcanota y discurre a través de todo el valle Sur de Cusco, para luego discurrir por el Valle Sagrado de los Incas e internándose hacia la ceja de selva llegando al distrito de Machupicchu [10]. (Ver fig. 1).



Fig. 1. Río Vilcanota. C. H. de Machupicchu [10].

En el 2001 su operación contaba con tres grupos Pelton de 30 MW cada uno; totalizando 90 MW. En el 2015, el COES aprueba la Puesta en Operación del Grupo N° 04 de la II fase de la Central Hidroeléctrica de Machupicchu, de una potencia instalada de 102 MW, con lo que la potencia instalada de la central se habría incrementado hasta 192 MW como podemos observar en la tabla I.

TABLA I
CARACTERÍSTICAS OPERATIVAS DE LA C. H. DE MACHUPICCHU [10].

Cantidad	Unidades Generadoras	Potencia Instalada	Caída Neta	Caudal de Diseño
3	Turbinas Pelton	90 MW (3*30MW)	345 m	30 m ³ /s
1	Francis Vertical	102 MW	345 m	32 m ³ /s

La central hidroeléctrica de Machupicchu (ver Fig.2), abastece de energía eléctrica a los departamentos de Cusco, Puno y Apurímac; además se consolida como un importante

Digital Object Identifier: (only for full papers, inserted by LACCEI).
ISSN, ISBN: (to be inserted by LACCEI).
DO NOT REMOVE

polo energético que podría crecer en un futuro cercano con la instalación de la central hidroeléctrica de Santa Teresa II. El mejoramiento de la generación eléctrica en la central hidroeléctrica de Machupicchu y la evaluación del comportamiento de la cuenca del río Vilcanota es de suma importancia, debido a que los caudales naturales aforados son muy variables, pues históricamente en época de estiaje, el caudal es de 30 m³/s, mientras que la central hidroeléctrica demanda 62 m³/s. (ver Fig. 3).



Fig. 2. C. H. de Machupicchu [10].

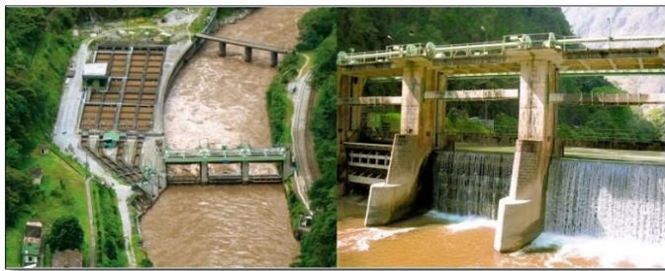


Fig. 3. Descarga del río Vilcanota [10].

La finalidad de esta investigación es la modelización y predicción del caudal del río Vilcanota que llega a la Central Hidroeléctrica de Machupicchu, mediante análisis cuantitativo ARIMA de Series Temporales en resolución mensual, el cual nos permitirá conocer el comportamiento de sus caudales en un horizonte predictivo de no menor de 2 años; así como sugerir y prever el suministro de caudales provenientes de otros embalses en la época de estiaje a fin de garantizar la disponibilidad operativa de la Central Hidroeléctrica de Machupicchu.

II. METODOLOGÍA

2.1. Especificación del modelo

Los Modelos ARIMA son modelos paramétricos que representan la serie de variables aleatorias y se analizan en función de la interrelación temporal de sus elementos [11]. Box-Jenkins son los autores de la modelización ARIMA y ellos recomiendan al menos 50 datos y así poder determinar el modelo ARIMA adecuado [12], [13].

Se dice que un proceso es estocástico no estacionario cuando al diferenciar el proceso original, el proceso transformado resultante es estacionario y el número de veces

que debe diferenciarse el proceso original para transformarse en estacionario constituye el grado de orden de homogeneidad.

A un proceso integrado X_t se le denomina proceso autorregresivo de medias móviles integrado, ARIMA (p,d,q), si tomando diferencias de orden d se obtiene un proceso estacionario Z_t . El modelo ARIMA se expresa de la siguiente forma [14]:

$$Z_t = \phi_1 Z_t + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + \delta + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q} \quad (1)$$

Abreviadamente se tiene:

$$\phi(L)Z_t = \delta + \theta(L)u_t \text{ siendo } Z_t = \Delta^d X_t = (1-L)^d X_t$$

Quedando:

$$\phi(L)(1-L)^d X_t = \delta + \theta(L)u_t \quad (2)$$

Donde:

$\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$: Son los coeficientes

Z_t : Es la variable

Z_{t-p} : Los datos previos a la etapa

Se puede observar que existe una tendencia en la varianza, esto es, que la dispersión de las observaciones no es constante a lo largo del tiempo. Cuando se presenta este hecho la transformación adecuada puede consistir en tomar logaritmos neperianos [14]:

$$\phi(L)\Delta^d X_t^\lambda = \delta + \theta(L)u_t \quad (3)$$

$$\phi(L)(1-L)^d (X_t^\lambda - u) = \theta(L)u_t \quad (4)$$

Otra fuente de no estacionariedad en muchas de las series reales del ámbito energético lo constituye la estacionalidad. Para desestacionalizar las series se procede a la diferenciación estacional. Los modelos estacionales no estacionarios pero homogéneos, ARIMA(P,D,Q).

$$Z_t = \Phi_1 Z_{t-s} + \Phi_2 Z_{t-2s} + \dots + \Phi_p Z_{t-ps} + \delta + u_t - \Theta_1 u_{t-s} - \dots - \Theta_q u_{t-qs} \quad (5)$$

$$Z_t = \Delta_s^D X_t = (1-L^s)^D X_t \quad (6)$$

Abreviadamente se tiene:

$$\phi_p(L^s)(1-L^s)^D X_t = \delta + \theta_q(L^s)u_t \quad (7)$$

2.2. Procedimiento para la elaboración de un modelo ARIMA

Para identificar el modelo adecuado se debe cumplir con las siguientes etapas (ver fig.4).

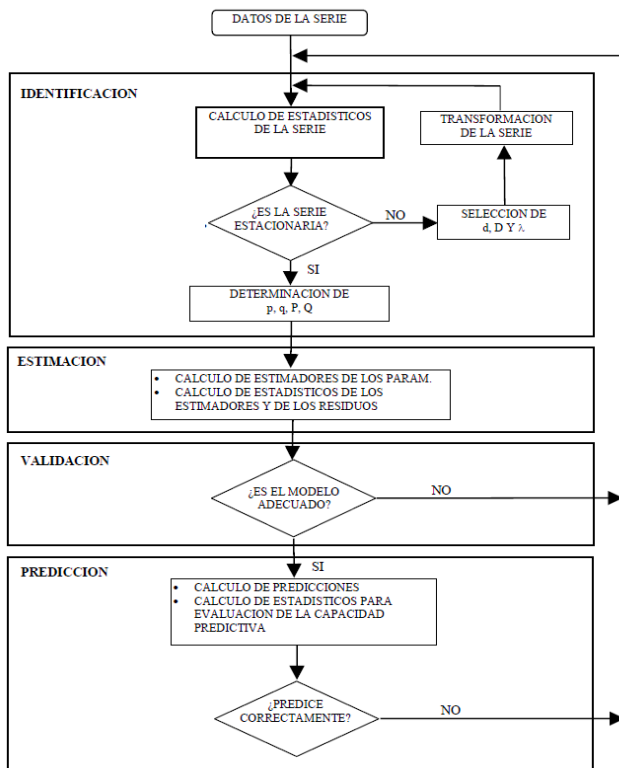


Fig. 4. Esquema de análisis de modelos ARIMA [14].

2.3. Tratamiento de la información de entrada

La variable de estudio Caudal está representada por la serie de caudales tomados en la bocatoma de la Central Hidroeléctrica de Machupicchu en valores mensuales en m^3/s .

2.3.1. Tamaño y característica de la muestra

De acuerdo a la información obtenida de los caudales naturalizados desde el año 2015 al 2023 [10], se tienen un total de 108 observaciones cuyo comportamiento se muestra a continuación (Ver fig.5).

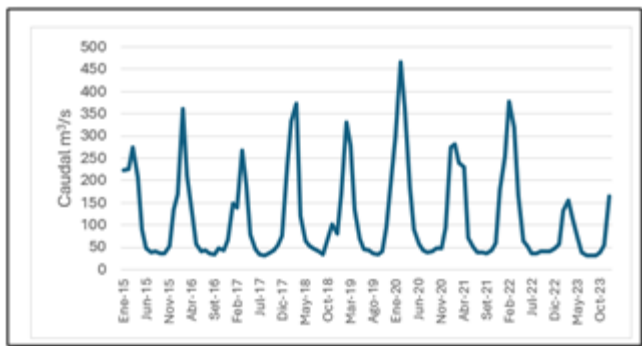


Fig. 5. Serie temporal de caudales del 2015 al 2023

2.3.2. Estadísticos descriptivos del Caudal

Podemos observar el caudal máximo y mínimo que se dan entre los años del 2015-2023 en la tabla 2.

TABLA II
CAUDAL MÁXIMO Y MÍNIMO.

	N	Mínimo	Máximo	Media	Desviación
Caudal	108	30.25	465.8	115.7731	102.84596

La serie caudal, de acuerdo con los registros obtenidos no presenta una tendencia marcada, pero sin variaciones estacionales mensuales. La serie muestra un marcado patrón estacional, ya que los caudales alcanzan su punto máximo en los meses de noviembre de un año a mayo del año siguiente. Mientras que los caudales menores ocurren en los meses de junio a octubre. Se puede decir que la serie exhibe un comportamiento periódico con un periodo de 12 meses. También, se caracteriza por ser una serie no estacionaria y se sabe que el modelo ARIMA requiere que la serie sea estacionaria en varianza y en media regular y estacional.

2.3.3. Análisis de estabilidad de la varianza de la serie

Para analizar la estabilidad de las varianzas de la serie Caudal se utiliza el Test de Levene. Esta prueba consiste en evaluar la posible igualdad o desigualdad de las varianzas de la serie. Para determinarla se prueba la Hipótesis Nula de que las varianzas sean iguales.

La significancia (p valor) Sig=0.032 es menor de 0.05, por lo tanto, no existe homogeneidad de varianzas. El estadístico de Levene indica que la serie no es estacionaria en varianza, se puede observar en la tabla 3.

TABLA III
PRUEBA DE HOMOGENEIDAD EN VARIANZA

Prueba de Homogeneidad de Varianza					
		Estadístico de Levene	gl1	gl2	Sig
Caudal	Se basa en la media	2.227	8	99	0.032
	Se basa en la media	0.517	8	99	0.841
	Se basa en la media y con gl ajustado	0.517	8	73.164	0.84
	Se basa en la media recortada	1.783	8	99	0.089

2.3.4. Componente estacional y cíclico de la serie original

La estacionalidad de la serie la podemos analizar mediante el periodograma por frecuencia de la serie temporal, a cada amplitud le corresponde una frecuencia cuya inversa es el periodo estacional o cíclico [15].

El valor más alto de 0.08 de frecuencia, de acuerdo con el periodograma de Frecuencia se observa una fuerte variación estacional cada 12 meses que corresponde a un periodo de un

año o cada 12 meses. Frecuencia: $0,080 \quad 1/0.080= 12$ meses (ver fig.6).

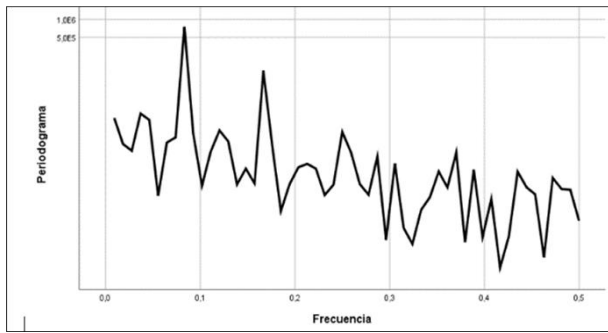


Fig. 6. Periodograma de frecuencia de la serie temporal de caudal.

2.3.5. Prueba de Levene

La prueba de Levene, que es una prueba que permite comprobar la hipótesis de que los grupos anuales de datos mensuales formados proceden de poblaciones con varianza común. En esta situación existe una familia de transformaciones de Box-Cox, que puede estabilizar la varianza, y es dado por:

$$\text{Transformación de } X_t: \begin{cases} X_t^\lambda, & \lambda \neq 0 \\ \ln(X_t), & \lambda = 0 \end{cases} \quad (8)$$

El poder de transformación es igual a uno menos la pendiente de la recta de regresión mínimo-cuadrática ajustada a los puntos generados de la gráfica de dispersión vs. nivel. Debe tenerse en cuenta que esta familia de transformaciones no solo permite estabilizar las varianzas, sino que además puede proporcionar normalidad. Así, la prueba de Levene permitirá contrastar la hipótesis nula de que no existen diferencias significativas entre las varianzas de la serie caudales en los grupos anuales conformados. Con el programa SPSS se realiza esta prueba estimando el poder de transformación λ mediante el comando exploración dispersión nivel con test de Levene. Dado que la potencia de transformación estimada se obtiene restando a uno el valor de la pendiente de la recta de regresión: $\lambda= 1-1.273 = -0.272$ (ver fig. 7).

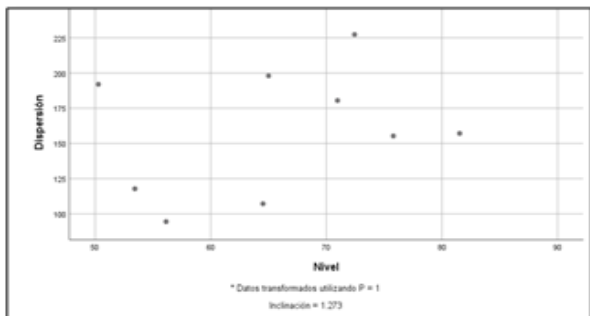


Fig. 7. Gráfico de dispersión vs Nivel de caudal por año.

Se sabe que es habitual utilizar potencias redondeadas a múltiplos de 0.5. Dado que la potencia de transformación estimada es de -0.272 , de las potencias posibles que lograrán la estabilización en varianza de la serie original Caudal, serían $\lambda=0$ ó $\lambda=-0.5$, se recomienda un $\lambda=0$ equivale a una transformación de logaritmo natural, y será la transformación que efectuaremos para estabilizar las varianzas de la serie temporal Caudal.

2.4. Evaluación

Para poder lograr estabilidad en media regular y media estacional debemos de encontrar los valores adecuados de las diferenciaciones “d” y “D” de la serie. Generalmente los valores de d y D se encuentran entre 0, 1 ó 2. A partir de la serie Caudal transformada (LnCaudal), se analizó su correspondiente función de autocorrelación simple ACF (ver fig. 8), se observa que tiene un decaimiento rápido en los primeros retardos, por lo que se deduce que la serie no presenta una tendencia marcada.

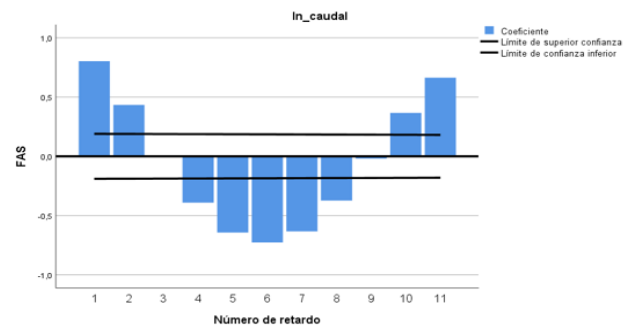


Fig. 8. Gráfico de autocorrelaciones simple para la serie.

Con respecto a la estacionalidad de la serie LnCaudal, se evaluó la correspondiente función de autocorrelación simple ACF con más retardos (ver Fig. 9), se visualiza que cada 12 periodos la autocorrelación vuelve a predominar, por lo que se deduce que la serie presenta estacionalidad, por lo tanto, la serie será diferenciada estacionalmente.

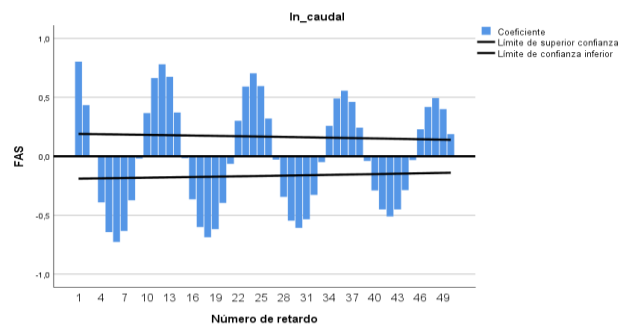


Fig. 9. Gráfico de autocorrelaciones simple para la serie LnCaudal con más retardos.

2.4.1. Determinación de ordenes autorregresivos y medias móviles.

Para analizar la parte regular de la serie (ver fig.10), se muestran las autocorrelaciones estimadas ACF de la serie LnCaudal con $d=0$ y $D=1$, y la figura 11 muestra las correspondientes autocorrelaciones parciales estimadas ACFP. A partir de la función de autocorrelación de la serie, se observa que las autocorrelaciones caen aproximadamente exponencialmente (rápidamente), y en la función de autocorrelación parcial, se ve que después del retardo 1 las correlaciones parciales disminuyen con bastante regularidad, estas características son de un modelo autorregresivo, esto sugiere que esta serie temporal podría describirse sólo mediante un modelo AR(1), por lo tanto, una alternativa para la parte regular de la serie será un modelo de la forma ARMA(1,0) con $p=1$ y $q=0$. Para analizar la parte estacional de la serie, también se emplean autocorrelaciones estimadas ACF y autocorrelaciones parciales estimadas ACFP, de la serie LnCaudal con $d=0$ y $D=1$, vistas en la figura 10 y figura 11. Analizando los rezagos estacionales, en la función de autocorrelación de la serie, se puede observar que hay un decaimiento de las correlaciones de los rezagos estacionales, y en la función de autocorrelación parcial, se ve que después del primer rezago estacional (periodo 12) la autocorrelación parciales estacionales (24, 36, ...) son pequeñas, esto sugiere que la serie se puede describir mediante un modelo estacional AR(1)s, por lo tanto, una alternativa para la parte estacional de la serie será un modelo de la forma ARMA(1,0)s con $P=1$ y $Q=0$. Luego, el modelo obtenido es de la forma ARIMA (1,0,0)x(0,1,1)s. Se propuso además el modelo ARIMA (1,0,1)x(0,1,1)s. Estos modelos serán sometidos a análisis para primero ver si sus parámetros son significativos y evaluar las pruebas de residuos, escogiendo el modelo que consiga el mejor ajuste.

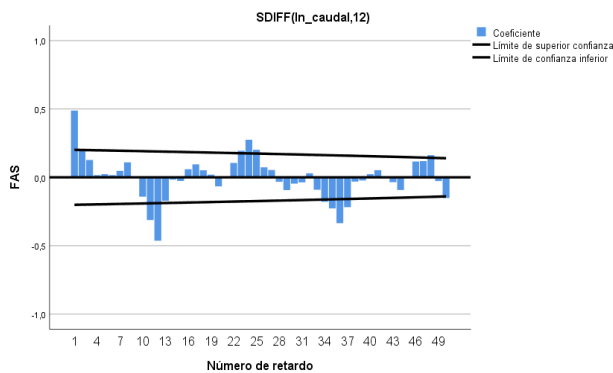


Fig. 10. Gráfico de autocorrelaciones simple para la serie LnCaudal con $d=0$ y $D=1$

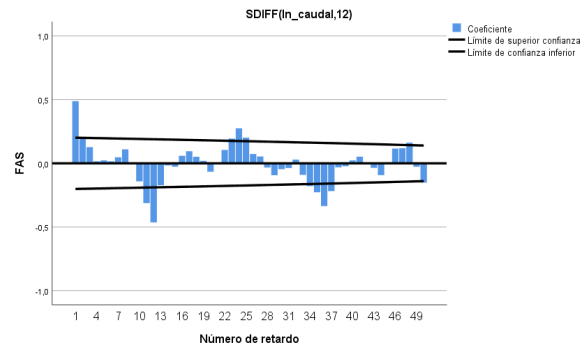


Fig. 11. Gráfica de autocorrelaciones Parciales para la serie LnCaudal con $d=0$ y $D=1$.

Teniendo en cuenta las características de la serie, su transformación de logaritmo natural, sus características de estacionalidad y tendencia proponemos evaluar los siguientes modelos ARIMA como se puede observar en la tabla 4.

TABLA IV
MODELOS ARIMA IDENTIFICADOS

Modelo	Tipo de transformación	Orden regular d	Orden Estacional D	Modelo ARIMA	Condición
M1	Logaritmo natural del caudal	0	1	(1,0,0)x(1,1,0)	Tentativo
M2	Logaritmo natural del caudal	0	1	(1,0,1)x(1,1,0)	Alternativo

III. RESULTADOS

3.1. Estimación de parámetros.

Un indicativo del grado de ajuste del modelo ARIMA a los datos es la raíz cuadrada de las varianzas de los residuales, cuando menor sea este valor será mejor el ajuste de los datos de la serie temporal al modelo. Los modelos identificados; ARIMA (1,0,0)x(0,1,1)s y ARIMA (1,0,0)x(1,1,1)s serán analizados, utilizando el programa SPSS, que nos permite verificar la significancia de sus parámetros, así como efectuar las pruebas de residuos como se puede ver en la tabla 5.

TABLA V
MODELOS ARIMA OPTIMO.

Modelo	Modelo ARIMA	Contraste de Significancia de parámetros	Evaluación
M1	(1,0,0)x(1,1,0)	Coefficientes estables y significativos	Óptimo
M2	(1,0,1)x(1,1,0)	Coefficientes estables y significativos	

3.2. Resultados de las estimaciones de modelos ARIMA representativos.

Las estimaciones y ajuste, calculados con el Programa SPSS, de los coeficientes autorregresivos y medias móviles regulares y estacionales se muestran a continuación para cada modelo:

Modelo (M1) ARIMA (1,0,0) x (1,1,0) de la serie LnCaudal. En la tabla 6, se muestra las estimaciones de los parámetros

del primer modelo ARIMA (1,0,0) x (1,1,0) s, con sus respectivos errores estándar y significancia. Se puede observar que los parámetros autorregresivos AR tanto de la parte regular como de la parte estacional son significativos (con **Sig.=0.000 < 0.05**).

TABLA VI
ESTIMACIONES DE LOS PARÁMETROS DEL MODELO (M1) ARIMA (1,0,0) X (1,1,0)S DE LA SERIE LNCAUDAL.

		Estimación	SE	t	Sig	
Modelo 1	Caudal	Constante	-0.022	0.041	-0.55	0.584
		AR Retardo1	0.511	0.09	5.696	0.0
	Logaritmo Natural	AR, estacional Retardo1	-0.597	0.089	-6.683	0.0
		Diferencia Estacional	1			

Luego este modelo M1 tiene la siguiente estructura: ARIMA (1,0,0)x(1,1,0)s para LnCaudal.

$$(1 - \phi L)(1 - \Phi L^{12})(1 - L^{12})LnCaudal = \theta_0 + a_t \quad (8)$$

$$(1 - 0.511L)(1 + 0.597L^{12})(1 - L^{12})LnCaudal = a_t \quad (9)$$

Donde $\phi = 0.511$ $\Phi = -0.597$ $\theta_0 = -0.022$

En la tabla 7, se muestra las estimaciones de los parámetros del primer modelo ARIMA (1,0,1) x (1,1,0)s, con sus respectivos errores estándar y significancia. Se puede observar que los parámetros autorregresivos AR tanto de la parte regular como de la parte estacional son significativos (con Sig.=0.000 < 0.05), pero el parámetro MA de la parte regular no es significativo (**Sig.=0.5872 > 0.05**).

TABLA VII
ESTIMACIONES DE LOS PARÁMETROS DEL MODELO (M2) ARIMA (1,0,1)x(1,1,0)S DE LA SERIE LNCAUDAL

		Estimación	SE	t	Sig	
Modelo 2	Caudal	Constante	-0.022	0.043	-0.513	0.609
		AR Retardo1	0.593	0.166	3.568	0.001
	Logaritmo Natural	MA Retardo1	-0.114	0.206	0.552	0.6
		AR, estacional Retardo1	-0.604	0.09	-6.714	0.0
		Diferencia Estacional	1			

Luego este modelo M2 tiene la siguiente estructura: ARIMA (1,0,1)x(1,1,0)s para LnCaudal.

$$(1 - \phi L)(1 - \Phi L^{12})(1 - L^{12})LnCaudal = \theta_0 + (1 - \theta L)a_t \quad (9)$$

$$(1 - 0.593L)(1 - 0.604L^{12})(1 - L^{12})LnCaudal = (1 - 0.114L)a_t \quad (10)$$

Donde $\phi = 0.593$ $\Phi = -0.604$ $\theta_0 = 0$

En la tabla 8 se resume los modelos ajustados incluyendo el indicador BIC para evaluar los modelos M1 y M2.

TABLA VIII
MODELOS ARIMA

Modelo	Modelo ARIMA (p,d,q)x(P,D,Q)	Ecuación del Modelo	BIC
M1	(1,0,0)x(1,1,0)	$(1 - 0.511L)(1 + 0.597L^{12})(1 - L^{12})LnCAUDAL = a_t$	8.021
M2	(1,0,1)x(1,1,0)	$(1 - 0.593L)(1 + 0.604L^{12})(1 - L^{12})LnCAUDAL = (1 - 0.114L)a_t$	8.044

3.3. Validación del modelo ARIMA optimo

La validación de los mejores modelos ARIMA calculados y seleccionados consiste en comprobar si satisface la hipótesis relativa de que los residuos del modelo (at), obedecen al comportamiento de ruido blanco. Si Xt en la serie en estudio y X, es su estimación mediante modelo ajustado, una estimación de la realización de ruido blanco será el resultado de $\hat{a}_t = X_t - X$ [8].

En el programa SPSS, las estimaciones de los valores de la serie temporal Caudales mediante modelo ARIMA elegido, conforman una serie denominada FIT acompañada de la serie residual ERR.

$$ERR = X_t - FIT$$

Por lo tanto, la validación consistirá en comprobar que la serie ERR cumple:

- Tiene media cero
- Es estable en varianza:
- En lo que se refiere a la estabilidad en varianza de ERR, es consecuencia de la estabilidad en varianza de la serie original Caudal, realizada en la etapa de identificación del modelo.
- Las observaciones están incorrelacionadas.
- Procede de una distribución normal.

Se debe de comprobar la idoneidad del ajuste del modelo examinado y los residuos del modelo ajustado.

3.4. Prueba de incorrelación de residuos

Para efectuar una verificación que los residuos están incorrelacionados, se utiliza el gráfico de la función de autocorrelación simple de dichos residuos, si todos los coeficientes de la ACF se encuentran dentro de la banda de confianza al 95% tendiendo a cero.

En las siguientes figuras se presentan las salidas del SPSS, en la figura 12 se muestra los residuos del modelo M1 ajustado ARIMA (1,0,0)x(1,1,0)s, se puede notar que tiene un comportamiento alrededor de cero y mostrando una aleatoriedad. En la figura 12 se observa que las autocorrelaciones residuales no presentan evidencia de falta de ajuste, ya que ninguno de los valores queda fuera de las bandas de confianza del 95%. En la figura 13 se muestra las autocorrelaciones con la estadística de **Q** de Box-Ljung, y se presenta que todos los valores pasan la prueba de que los residuos son incorrelacionados. Por lo tanto, ante estas evidencias se puede concluir para el modelo M1 ajustado ARIMA (1,0,0)x(1,1,0)s que los residuos tienen media cero y son incorrelacionados, esto es los residuos son un ruido blanco.

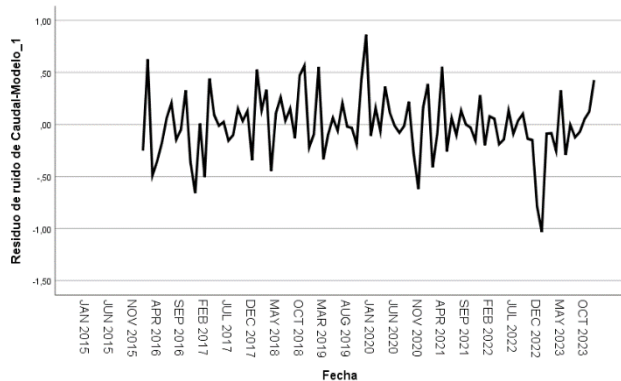


Fig. 12. Gráfico de residuos del modelo M1 ajustado ARIMA (1,0,0)x(1,1,0)s

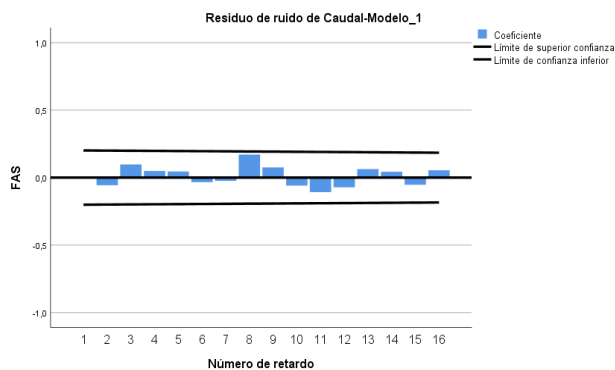


Fig. 13. Gráfico de autocorrelaciones de los residuos del modelo M1 ajustado ARIMA (1,0,0) x (1,1,0)s

En la tabla 9, se puede observar las autocorrelaciones simples y la estadística de Q de BOX-Ljung.

TABLA IX
AUTOCORRELACIONES SIMPLES Y LA ESTADÍSTICA DE Q DE BOX-LJUNG

Serie: Residuo de Ruido de Caudal-Modelo_1					
Retardo	Autocorrelación	Desv. Error ^a	Estadístico de Box-Ljung		
			Valor	gl	Sig. B
1	-0.002	0.100	0.000	1	0.985
2	-0.056	0.100	0.317	2	0.853
3	0.097	0.099	1.275	3	0.735
4	0.049	0.099	1.521	4	0.823
5	0.046	0.098	1.736	5	0.884
6	-0.034	0.098	1.855	6	0.933
7	-0.024	0.097	1.917	7	0.964
8	0.170	0.097	5.016	8	0.756
9	0.076	0.096	5.638	9	0.776
10	-0.060	0.096	6.025	10	0.813
11	-0.108	0.095	7.312	11	0.773
12	-0.072	0.094	7.894	12	0.793
13	0.063	0.094	8.341	13	0.821
14	0.044	0.093	8.558	14	0.858
15	-0.053	0.093	8.886	15	0.883
16	-0.054	0.092	9.233	16	0.904

3.5. Prueba de normalidad de residuos

Es una prueba que permite contrastar la hipótesis nula de distribución normal es la prueba de bondad de ajuste de Kolmogorov-Smirnov, donde si el p-valor asociado al valor del estadístico de contraste es menor que α , se rechazara la hipótesis nula al nivel de significancia α (0.05). La salida en SPSS mostrada en la tabla 10, de esta prueba para contrastar la hipótesis nula de normalidad de la serie residual del modelo M1 ajustado ARIMA (1,0,0) x (1,1,0)s, se da a continuación.

TABLA X
PRUEBA DE NORMALIDAD PARA LOS RESIDUOS DEL MODELO M1 AJUSTADO ARIMA (1,0,0) x (1,1,0)s.

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig
Residuo de Ruido de Caudal-Modelo_1	0.000	1	0.985	0.000	1	0.985
a. Corrección de Significación de Lilliefors						

De ello se observa que el p-valor (Sig.=0.085>0.05) asociado al estadístico de contraste es grande; por tanto, no se puede rechazar la hipótesis nula de distribución normal. También en la figura 14, se muestra el Gráfico Q-Q normal de los residuos del modelo M1 ajustado ARIMA (1,0,0)x(1,1,0)s, se observa que casi los todos puntos están sobre la recta, es evidencia que los residuos presentan una distribución normal.

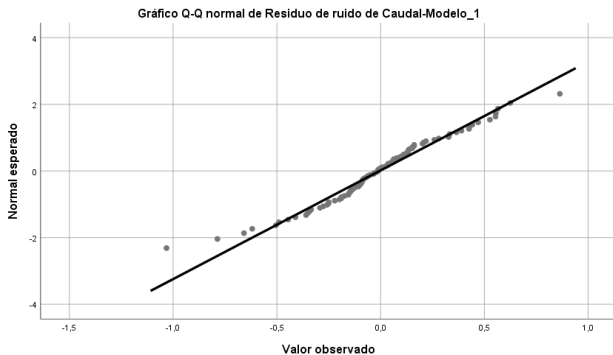


Fig. 14. Gráfico Q-Q normal de los residuos del modelo M1 ajustado ARIMA (1,0,0) x (1,1,0)s.

Luego, al analizar el modelo M1 propuesto ARIMA (1,0,0) x (1,1,0)s de la serie LnCaudal, se comprobó la idoneidad del ajuste del modelo examinando los residuos del modelo ajustado. Esto comprobó que los residuos están incorrelacionados y proceden de una distribución normal. Por lo tanto, este modelo M1 propuesto será utilizado con fines de pronóstico

3.5.1. Prueba de normalidad de residuos

Cálculo de Predicciones ARIMA

El cálculo de las predicciones de la serie temporal Caudal, utilizando modelos ARIMA, se realiza a partir de las expresiones matemáticas obtenidas en el proceso de estimación y ajuste, hemos seleccionado el modelo M1, cuyos coeficientes son significativos y es un buen modelo predictivo, la ecuación tiene la siguiente forma.

El modelo M1 tiene la siguiente estructura:

ARIMA (1,0,0)x(1,1,0)s para LnCaudal. Y utiliza la ecuación 9 y se asume que $Z_t = L_n$ Caudal. La ecuación obtenida para nuestro modelo es:

$$z_t = 0.511z_{t-1} + 0.403z_{t-12} - 0.206z_{t-13} + 0.597z_{t-24} - 0.305z_{t-25} + a_t$$

Resultados predictivos

En nuestro caso se trata de 108 observaciones de la serie temporal caudal. Las predicciones deben ser procesos dinámicos que deben ser validados año a año, por lo que definimos que nuestro horizonte predicción será de dos años.

En la figura 15, se observa la capacidad predictiva del modelo ARIMA y coincidió con el modelo original (fig.5). La serie se ajusta al modelo M1. Donde Lncaudal es el logaritmo natural de la serie. LCL es el límite de confianza inferior y UCL es el límite de confianza superior. Se observa que las predicciones caen dentro del límite de confianza. Tanto en la zona de validación como la zona predictiva, lo cual justifica que nuestro modelo sea el mejor.

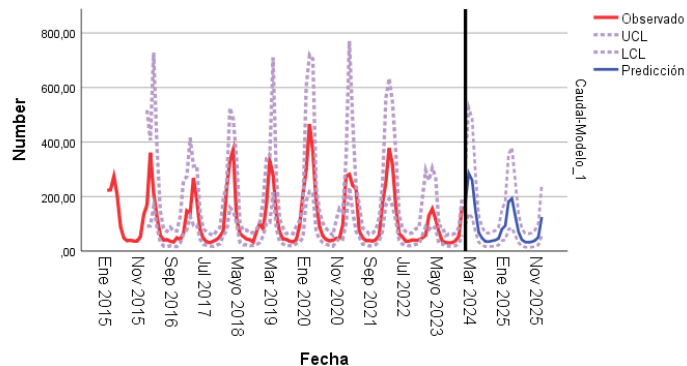


Fig. 15. Gráfico de caudal observado y ajustado.

3.5.2 Análisis comparativo con modelos determinístico de alisados exponenciales.

Los alisados exponenciales están definidos como métodos de predicciones incondicionales, autoproyectivos, basados en el enfoque determinista, adecuados para efectuar predicciones a corto plazo. El modelo a elegir será:

Modelo WINTERS. - Asume que la serie temporal tiene tendencia lineal y variación estacional multiplicativa.

El paquete estadístico SPSS nos permite efectuar el análisis de series temporales con modelos de suavizado exponencial, dándonos la posibilidad de escoger modelos; como el Winters Multiplicativo y Winters Aditivo.

Se seleccionará el modelo que presente el menor error estadístico MAPE (error porcentual absoluto medio). También se elegirán los parámetros: alfa, gamma y delta que correspondan al modelo seleccionado. A continuación, se muestran los indicadores estadísticos de los modelos predictivos analizados con el SPSS, se observa en la tabla XI.

TABLA XI
Resultado de los modelos predictivos determinísticos analizados

Indicadores Estadísticos	Aditivo de Winters	Multiplicativo de Winters
MAPE	27.76%	17.51%

El mejor modelo determinístico es el Suavizado exponencial Multiplicativo de Winters que muestra el menor error porcentual absoluto medio (MAPE) con el valor de 17.51%.

A continuación, en la tabla XII podemos observar los parámetros del modelo predictivo Multiplicativo de Winters

TABLA XII
Parámetros del modelo predictivo Multiplicativo de Winters

Modelo		Estimación	SE	t	Sig
Caudal modelo	Alfa (nivel)	.320	0.053	6.023	.000
	Gamma (tendencia)	.000	0.020	0.010	.992
	Delta (estacionalidad)	.039	0.080	0.492	0.624

Así mismo se procedió a efectuar el pronóstico con el modelo multiplicativo de Winters hasta diciembre del año 2025, cuyo gráfico del comportamiento del caudal, se muestra en la siguiente figura 16.

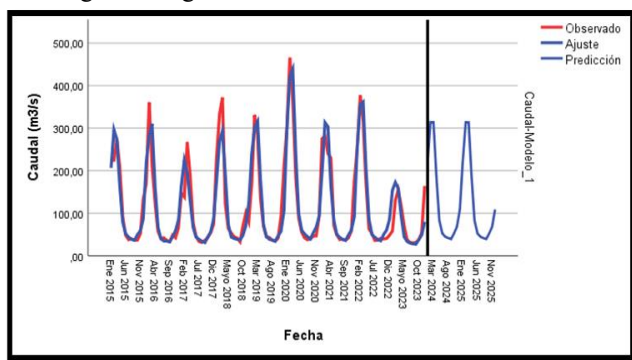


Fig. 166. Gráfico de la predicción al 2025 con Modelo Multiplicativo de Winters.

IV. CONCLUSIONES

1.- La serie temporal fue analizada para un periodo de 9 años desde enero del 2015 hasta diciembre del 2023, el cual hace un total de 108 meses. Se pudo observar que los datos son estacionarios en media, pero no en varianza.

2.- Al realizar el análisis mediante modelos ARIMA, se encontró que el mejor modelo de predicción es el ARIMA (1,0,0)x(1,1,0)s con transformación de la serie Caudal a logaritmo natural. Debido a que cumple con todas las exigencias de la etapa de validación además cuenta con un BIC igual 8.02.

3.- Los errores se encuentran incorrelados dentro de la banda de confianza al 95%, también pasan la prueba de normalidad de residuos y muestran los menores valores en la comparación por estadísticos de error.

4.- Los valores de las predicciones del comportamiento de la variable caudal, permiten predecir el comportamiento del río y planear acciones de tipo operativo para los años analizados 2024 y 2025.

5.- Al hacer una comparación entre el modelo de alisados exponenciales Winters y ARIMA. Observamos que la

predicción en ARIMA es más exacta y sus valores son más conservadores.

REFERENCIAS

- [1] G. Harrison and H. Whittington, *The Impact of Climatic Change on Hydropower Investment*. 2001.
- [2] C. Bello, W. Suarez, F. Drenkhan, and F. Vega-Jácome, "Hydrological impacts of dam regulation for hydropower production: The case of Lake Sibinacocha, Southern Peru," *J. Hydrol. Reg. Stud.*, vol. 46, p. 101319, 2023, doi: <https://doi.org/10.1016/j.ejrh.2023.101319>.
- [3] M. A. D. Larsen *et al.*, "Challenges of data availability: Analysing the water-energy nexus in electricity generation," *Energy Strateg. Rev.*, vol. 26, p. 100426, 2019, doi: <https://doi.org/10.1016/j.esr.2019.100426>.
- [4] A. Bhaduri, C. Ringler, I. Dombrowski, R. Mohtar, and W. Scheumann, "Sustainability in the water-energy-food nexus," *Water Int.*, vol. 40, no. 5-6, pp. 723-732, 2015, doi: 10.1080/02508060.2015.1096110.
- [5] S. Ozturk and F. Ozturk, "Forecasting Energy Consumption of Turkey by Arima Model," *J. Asian Sci. Res.*, vol. 8, no. 2, pp. 52-60, 2018, doi: 10.18488/journal.2.2018.82.52.60.
- [6] J. Polprasert, V. A. H. Nguyen, and S. Nathanael Charoensook, "Forecasting Models for Hydropower Production Using ARIMA Method," in *Proceeding of the 2021 9th International Electrical Engineering Congress, iEECON 2021*, 2021, pp. 197-200, doi: 10.1109/iEECON51072.2021.9440293.
- [7] D. Gupta, A. Chakrabarti, and J. Gautam, "ARIMA Based Forecasting of stream flows of Three Georges Dam for efficient Water Resource Planning and Management," 2021, doi: 10.1109/ISAECT53699.2021.9668539.
- [8] X. Zhang, X. Wu, G. Zhu, X. Lu, and K. Wang, "A seasonal ARIMA model based on the gravitational search algorithm (GSA) for runoff prediction," *Water Supply*, vol. 22, no. 8, pp. 6959-6977, 2022, doi: 10.2166/ws.2022.263.
- [9] Z.-Y. Wang, J. Qiu, and F.-F. Li, "Hybrid models combining EMD/EEMD and ARIMA for long-term streamflow forecasting," *Water (Switzerland)*, vol. 10, no. 7, 2018, doi: 10.3390/w10070853.
- [10] Egemsa, "Central Hidroeléctrica Machupicchu," *Egemsa*, 2024, <https://www.egemsa.com.pe/central-hidroelectrica-machupicchu>.
- [11] M. P. González Casimiro, *Análisis de series temporales: Modelos ARIMA*. 2009.
- [12] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: Forecasting and control, 5th Edition*. 2015.
- [13] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time series analysis: Forecasting and control: Fourth edition*. 2013.
- [14] S. González Chávez, "Proyección de la demanda de energía metodologías y aplicaciones," 2020.
- [15] V. M. Rueda, J. D. Velásquez Henao, and C. J. Franco Cardona, "Recent advances in load forecasting using nonlinear models," *DYNA*, vol. 78, no. 167, pp. 36-43, 2011, [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-83155175827&partnerID=40&md5=4967267fbb82e52d01979bb1f533c99>.