

Análisis estocástico de una serie cronológica de caudales mensuales. Caso de estudio: Río Caroní, Venezuela

Jorge E. Guevara Bello^a, Cristóbal E. Vega González^{*b}, Edilberto Guevara Pérez^a

^a(CIHAM-UC) Centro de Investigaciones Hidrológicas y Ambientales, Facultad de Ingeniería, Universidad de Carabobo, Barbula, Carabobo, Venezuela

^bInstituto de Matemáticas y Cálculo Aplicado (IMYCA), Facultad de Ingeniería, Barbula, Carabobo, Venezuela

Resumen.-

Las series de tiempo se analizan, tanto en el tiempo, como en el espacio y la frecuencia. El estudio en el tiempo y espacio, es más común entre científicos e investigadores; mientras que el estudio en el dominio de la frecuencia se usa en aplicaciones prácticas de la ingeniería. En esta investigación se lleva a cabo un análisis estocástico de la serie de tiempo de caudales mensuales del Río Caroní en la estación hidrométrica Guri para el período 1950–2003, utilizando el “método clásico” de Box y Jenkins. Para eliminar la tendencia y la periodicidad se usó la diferenciación estacional. Los modelos estocásticos de menor longitud de descripción fueron: SARIMA(1, 0, 0)(1, 1, 1)₁₂, SARIMA(0, 0, 0)(0, 1, 1)₁₂ y AR(1). Un análisis de residuos condujo a la selección del modelo SARIMA(1, 0, 0)(0, 1, 1)₁₂, como el modelo que está mejor ajustado a los datos.

Palabras clave: Modelos ARIMA estacionales, mínima longitud de descripción, Río Caroní.

Stochastic analysis of a monthly flows time series. Study case: Caroní River.

Abstract.-

Time series are analyzed in the domains of time, space and frequency. Time and space analysis is common between researchers; while frequency analysis is common in practical engineering applications. In this study, following the classic Box-Jenkins methodology, a stochastic analysis of monthly flows of Caroní River is done using the data of Guri station for the period of 1950–2003. The trend and periodicity of series was eliminated by seasonal differentiation. The stochastic models of minimum description length were: SARIMA(1, 0, 0)(1, 1, 1)₁₂, SARIMA(0, 0, 0)(0, 1, 1)₁₂ and AR(1). A residuals analysis led to SARIMA(1, 0, 0)(0, 1, 1)₁₂ as the model that fits better to data.

Keywords: Seasonal ARIMA models, minimum description length, Caroní River.

Recibido: 15 marzo 2010

Aceptado: 09 abril 2010

1. Introducción

Uno de los aspectos fundamentales del proceso de planificación, diseño y operación de sistemas de obras hidráulicas es el de conocer la variabilidad de las disponibilidades de agua, de los usos y demandas correspondientes. Tradicionalmente, la variabilidad de las disponibilidades del agua era representada por el histórico de los eventos hidrológicos y particularmente por un periodo crítico de este registro.

*Autor para correspondencia

Correos-e: jorgeguevara85@gmail.com (Jorge E. Guevara Bello), vega.cristobal@gmail.com, cvega@uc.edu.ve (Cristóbal E. Vega González), eguevara@uc.edu.ve (Edilberto Guevara Pérez)

Durante las últimas tres décadas, la teoría de la estadística, probabilidades y procesos estocásticos se ha utilizado con el objeto de tener una representación más adecuada de la variabilidad de los datos hidrológicos. Esta representación generalmente se hace mediante un modelo estocástico que describa todas las características de la serie histórica. La relevancia del uso de modelos estocásticos de series hidrológicas en el análisis de las obras hidráulicas que se proyectan, diseñan u operan, está en el hecho de poder estudiar las diferencias estructurales bajo diversas posibles condiciones hidrológicas que puedan presentarse durante el período de vida de las obras.

Los datos de series históricas proporcionan una valiosa información sobre la disponibilidad y usos de agua en el pasado, y si se analizan debidamente, dan una buena indicación de las disponibilidades y demandas de agua que pueden ser utilizadas en el futuro. El análisis estocástico de las series de caudales, así como el de las diferentes series de uso de agua, proporcionan modelos matemáticos que según el intervalo de tiempo escogido para el análisis, pueden reflejar las variaciones diarias, semanales, mensuales y anuales de dichas series.

Numerosos estudios se han desarrollado para investigar las características estocásticas de series hidrológicas y de usos de agua, así como la determinación de modelos que describan tales características. La identificación, estimación y selección del modelo de la serie cronológica que mejor se ajuste a los de los caudales mensuales del río Caroní en la estación Guri se llevará a cabo mediante el enfoque de Box y Jenkins descrito en su clásica obra "Time Series Analysis: Forecasting and Control" [1]. Este trabajo considera los modelos factibles explicados en Peña [2]. Entre los modelos factibles, será seleccionado como el de mejor ajuste a los datos el modelo de menor longitud de descripción según Rissanen [3], tal como desarrolló Vega [4].

El modelo obtenido es importante ya que la cuenca del río Caroní es emblemática para Venezuela desde el punto de vista de la generación de energía. Esa cuenca genera más del 65 % de la energía total que consume el país y más del 80 % de la energía hidroeléctrica que se produce.

2. Fundamento Teórico

Un proceso estocástico discreto es una sucesión de variables aleatorias Y_t ordenadas por t , donde t puede tomar cualquier valor entero entre $-\infty$ y ∞ ,

$$\{y_t\}_{t \in \mathbb{Z}}, \dots, y_{-3}, y_{-2}, y_{-1}, y_0, y_1, y_2, y_3, \dots,$$

donde

$$\mu_t = E[y_t] \tag{1}$$

$$\sigma_t^2 = \text{Var}(y_t) = E[(y_t - \mu_t)^2] \tag{2}$$

$$\begin{aligned} \gamma(t, t+k) &= \text{Cov}(y_t, y_{t+k}) \\ &= E[(y_t - \mu_t)(y_{t+k} - \mu_{t+k})] \end{aligned} \tag{3}$$

con *media*, *varianza* y *función de autocovarianza* como en las Ecuaciones (1), (2) y (3), respectivamente. La descripción que sigue se basa en la literatura [1], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14] y [15].

Un proceso *Estacionario en Sentido Amplio o Débil* es aquel en el cual, la esperanza matemática y la varianza son constantes y la función de autocovarianza depende sólo del resago k y no del tiempo

$$\begin{cases} \mu_t = \mu, & \text{constante} \\ \sigma_t^2 = \sigma^2 = \gamma_0 & \text{constante} \\ \gamma(t, t+k) = \gamma_k & \text{no depende de } t. \end{cases} \tag{4}$$

Los detalles de la definición dada por la Ecuación (4) están en el libro de Peña [2].

La *Función de Auto Correlación* (FAC) y la *Función de Auto Correlación Parcial* (FACP) dadas en las Ecuaciones (5) y (6), respectivamente, para una serie de datos dada $\{X_t\}_{t=1, \dots, T}$, ayudan a identificar el proceso por comparación con un conjunto de patrones que tipifican a los distintos modelos (ver [2]).

$$\widehat{\rho}_k = \frac{\gamma_k}{\gamma_0} \tag{5}$$

Los coeficientes $\widehat{\phi}_k = \phi_{kk}$ de la FACP se estiman utilizando las Ecuaciones de Yule-Walker (6) a partir de los coeficientes estimados de FAC $\widehat{\rho}_k$

$$\begin{aligned} \widehat{\rho}_0 &= 1. \\ \widehat{\rho}_j &= \phi_{k1}\widehat{\rho}_{j-1} + \phi_{k2}\widehat{\rho}_{j-2} + \dots + \phi_{kk}\widehat{\rho}_{j-j} \end{aligned} \tag{6}$$

3. Metodología

La serie de tiempo que se utilizó en este trabajo está constituida por los caudales mensuales del río Caroní en la estación hidrométrica Guri, periodo 1950–2003, ubicada en la Latitud N: 7.51.00 y Longitud O:-63.02.00; a 271 msnm (Figura 1). Se usó el software libre “GNU Regression Econometric Time series Library” (Gretl 1.9.0), para la identificación y estimación de los modelos factibles. Para la selección del modelo de mejor ajuste fue usado el principio de longitud de descripción mínima (MDL) y el estudio de los residuos mediante el software XploRe version 4 (bajo licencia del Research Data Center (RDC), HUMBOLDT-UNIVERSITÄT) “EXploratory Regression”.

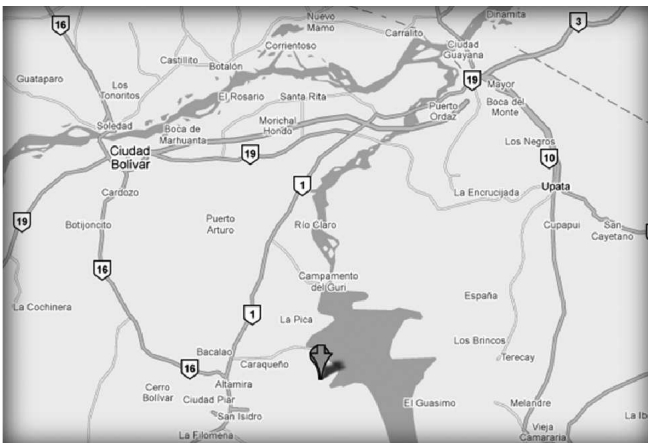


Figura 1: Estación Hidrométrica Caroní en Guri Fuente: Google Earth

Para la identificación del modelo fue estimado el correlograma de la serie y el periodograma. Por la naturaleza del correlograma y del periodograma y las pruebas de raíces unitarias, se determinó si son necesarias las diferencias estacionales Δ_s y no estacionales Δ para obtener una serie estacional. A partir de este análisis se determinó la naturaleza de la familia de modelos que identifican el modelo.

Dentro de la familia de modelos identificados se procedió a eliminar aquellos modelo que no son factibles; es decir, fueron eliminados los modelos en los cuales los t -valores de los parámetros no son significativos a un nivel de confianza de 90 %. También fueron eliminados los modelos en los cuales los polinomios de los operadores

de rezagos, tanto de la parte autorregresiva, como de la parte media móvil tuvieran raíces fuera del círculo unidad.

En este conjunto de modelos factible fueron seleccionados los modelo que mejor se adaptan a los datos por el Principio MDL de Rissanen [3]. El principio MDL está basado en minimizar la función de Longitud de Descripción (DL) [16],

$$DL(y_t) = f_{\text{res}} + f_{\text{pen}}. \quad (7)$$

Para un modelo \mathcal{M} La DL, dada en la Ecuación (7) consta de dos partes. La primera, una función de energía residual

$$f_{\text{res}}(y_t) = \frac{T}{2} \log \left(\sum (\hat{\varepsilon}_t^{(\mathcal{M})})^2 \right), \quad (8)$$

donde T es la longitud de los datos y $\{\hat{\varepsilon}_t^{(\mathcal{M})}\}_t$ son los residuos de la estimación del modelo \mathcal{M} . La segunda, una función de penalización del número de parámetros

$$f_{\text{pen}} = \eta \frac{n_{\text{par}}}{2} \log(T), \quad (9)$$

donde n_{par} es el número del parámetros no nulos del modelo, y η vale 2,5 para las series temporales. La función de la penalización tiene la finalidad de que el modelo sea lo más parsimoniosos posible. De las Ecuaciones (8) y (9) queda

$$LD(y_t) = \eta \frac{n_{\text{par}}}{2} \log(T) + \frac{T}{2} \log \left(\sum (\hat{\varepsilon}_t^{(\mathcal{M})})^2 \right). \quad (10)$$

Los detalles del cálculo de la Ecuación (10) y su implementación en XploRe están en [4].

Con un análisis de los residuos fueron validados los modelos de menor DL, a fin de determinar el modelo que mejor se adapta a los datos del caudal (ver [17]).

4. Resultados

En la Figura 2 se presenta la serie de tiempo original, en la cual se observa una periodicidad y una aparente tendencia en la media. En la Figura 3 se representa el correlograma de la serie, conjuntamente con los límites de confianza al 95 %. Se observa un alto grado de correlación para retardos cíclicos, lo cual implica que la media

también varía en forma periódica; es decir, cada mes tiene una media distinta. Esta afirmación es confirmada en el espectro de la muestra presentado en la Figura 4

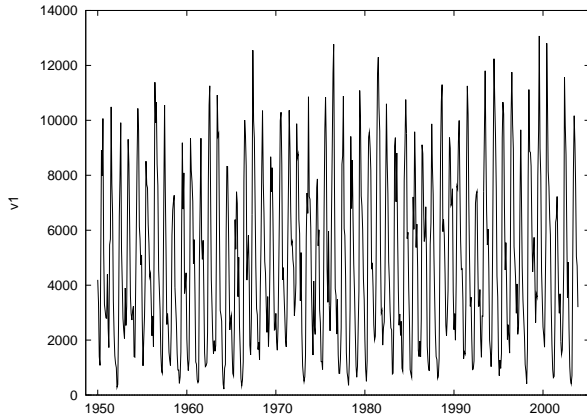


Figura 2: Serie de tiempo de caudales mensuales de la estación Guri

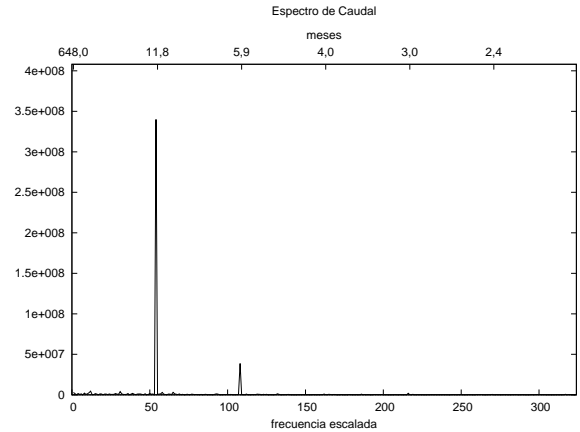


Figura 4: Periodograma muestral de los caudales mensuales de la estación Guri

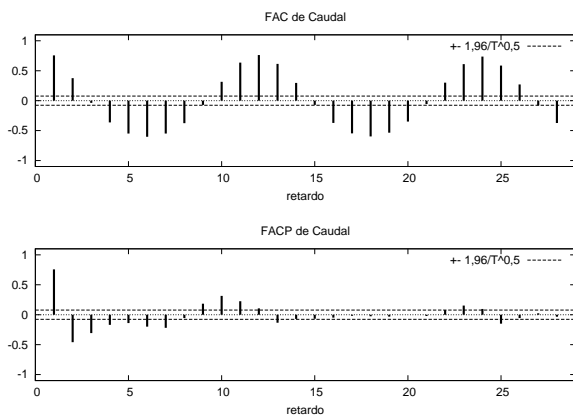


Figura 3: FAC y FACP de los caudales mensuales de la estación Guri

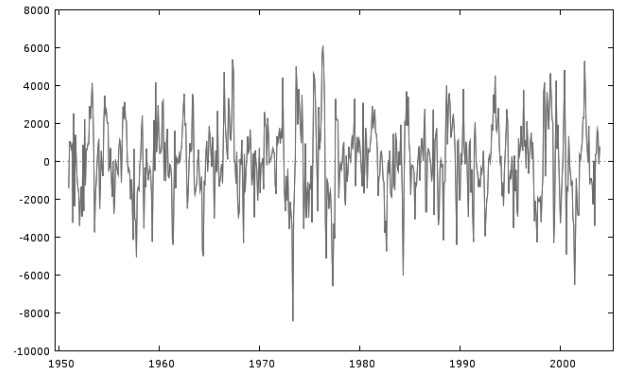


Figura 5: Serie de tiempo de la diferencia estacional

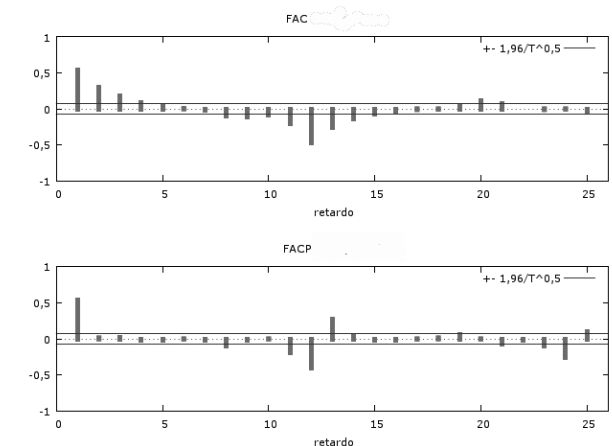


Figura 6: Correlograma de la diferencia estacional

En la Figura 4 esta representada la densidad espectral en función de la frecuencia escalada ($f_e = T/s$), T es la longitud de la muestra y s la frecuencia; puede observarse que para una velocidad angular de $\omega = 2\pi/s = 0,5236$, $f_e = 54$, $s = 12,00$ se obtiene un valor máximo representativo de $3,4008 \cdot 10^{08}$, lo cual garantiza una longitud de periodo anual.

Para remover la media cíclica, se aplicó la diferenciación estacional, $\Delta_s Caudal_t$, la cual está representada en la Figura 5. En las Figuras 6 y 7 están el

correlograma y el periodograma muestral, respectivamente, de la serie de la diferencia estacional.

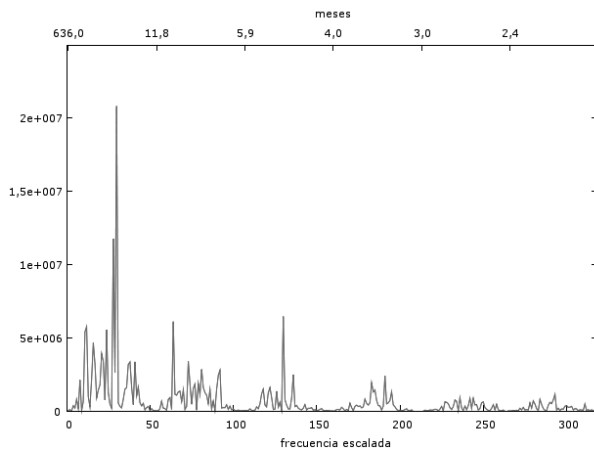


Figura 7: Espectro de la diferencia estacional

De las Figuras 6 y 7 se puede deducir que la serie de la diferencia estacional es estacionaria con orden autorregresivo no-estacional menor o igual a 1, orden media móvil no-estacional menor o igual a 4, orden de la diferencia no-estacional 0, orden autorregresivo estacional menor o igual a 1, orden media móvil estacional menor o igual a 3 y orden de la diferencia estacional 0 (ver [2]). Afirmando que el orden de la diferencia no-estacional es 0, implica que el proceso es estacionario o $I(0)$. Para confirmar esta aseveración fue realizada una prueba de raíces unitarias basada en el contraste aumentado de Dickey-Fuller (ADF) para la serie $\Delta_s Caudal_t$, la cual incluye un retardo de $(1 - B)\Delta_s Caudal_t$, donde B es el operador de retardo $B(y_t) = y_{t-1}$, para un tamaño muestral de 634, con la hipótesis nula de raíz unitaria: $a = 1$, lo cual da el siguiente resultado presentado en la Tabla 1.

Tabla 1: ADF contraste sin constante de la diferencia estacional del caudal

modelo: $(1 - B)y_t = (a - 1) \star y_{t-1} + \dots + e_t$	
Coef. AR(1) de e :	-0,001
valor estimado de $(a - 1)$:	-0,4585
Estadístico de contraste: $\tau_{nc(1)}$	-11,9623
p -valor asintótico	$9,827 \cdot 10^{-025}$

Lo cual garantiza que no existe raíz unitaria, por lo tanto la serie $\Delta_s Caudal_t$ es $I(0)$.

Del análisis anterior surge que el modelo inicial

a estimar, para la serie del Caudal, será de la forma

$$SARIMA(1, 0, 4)(1, 1, 3)_{12}. \tag{11}$$

En un modelo $SARIMA(p, d, q)(P, D, Q)_s$, se tiene que: p es el orden autorregresivo no-estacional, q es el orden media móvil no-estacional, d es el orden de la diferencia no-estacional, P es el orden autorregresivo estacional, Q es el orden media móvil estacional, D es el orden de la diferencia estacional y s es la longitud de período.

Tabla 2: SARIMA(1, 0, 4)(1, 1, 3)₁₂

	Coef.	σ	z	p -valor
const	13,7048	6,9554	1,9704	0,0488
ϕ_1	0,2485	0,2283	1,0885	0,2764
Φ_1	-0,5242	0,114185	-4,5908	0,0000
θ_1	0,2835	0,2267	1,2505	0,2111
θ_2	0,1954	0,1233	1,5850	0,1130
θ_3	0,0895	0,0807	1,1093	0,2673
θ_4	0,0396	0,0488	0,8120	0,4168
Θ_1	-0,4590	0,1163	-3,9463	0,0001
Θ_2	-0,4878	0,1156	-4,2202	0,0000
Θ_3	-0,0532	0,0394	-1,3491	0,1773
$\mu(V. D.)$	-0,5462	$\sigma(V. D.)$	2075,764	
$\mu(\epsilon_t)$	-2,3065	$\sigma(\epsilon_t)$	1243,318	
Log-veros.	-5457,917			

	r	Real	Imag.	Mód.	Frec.
AR					
	1	4,0249	0,0000	4,0249	0,0000
AR(sd)					
	1	-1,9077	0,0000	1,9077	0,5000
MA					
	1	0,7137	1,9325	2,0601	0,1937
	2	0,7137	-1,9325	2,0601	-0,1937
	3	-1,8422	1,5970	2,4381	0,3863
	4	-1,8422	-1,5970	2,4381	-0,3863
MA(sd)					
	1	1,0000	0,0000	1,0000	0,0000
	2	-2,4283	0,0000	2,4283	0,5000
	3	-7,7434	0,0000	7,7434	0,5000

La salida del Modelo de la Ecuación (11) está en la Tabla 2. Este modelo no es factible ya que varios de los coeficientes no son significativos al 90 %. Un procesamiento exhaustivo de búsqueda de los modelos factibles según el p -valor de los coeficientes llevó a los modelos tales como el $SARIMA(1, 0, 0)(1, 1, 2)_{12}$ y el $SARIMA(0, 0, 4)(2, 1, 3)_{12}$, que aunque el p -valor es menor a 0,10, no son factibles porque tienen raíces fuera del círculo unidad.

En la Tabla 3 se presentan los modelos factibles obtenidos con menor longitud de descripción. Para

Tabla 3: Longitud de Descripción de los Modelos Factibles para la Data de Caudal

Modelo	LD
1 SARIMA(1, 0, 0)(0, 1, 1) ₁₂	2868,1835
2 SARIMA(0, 0, 0)(0, 1, 1) ₁₂	2913,6736
3 ARMA(1, 4) $\theta_1 = 0$	3038,8034
4 AR(1)	3061,1399

una selección entre estos fue realizado un análisis de residuos. Los modelos seleccionados por el principio MDL son los siguientes

$$Q_t = 13,7421 + 0,5434Q_{t-1} - 0,9998\Delta_s \varepsilon_{t-1} \quad (12)$$

$$Q_t = 14,3551 - 0,9998\Delta_s \varepsilon_{t-1} \quad (13)$$

$$Q_t = 4837,47 + 0,8440Q_{t-1} - 0,1531\varepsilon_{t-2} - 0,3157\varepsilon_{t-3} - 0,3413\varepsilon_{t-4} \quad (14)$$

$$Q_t = 4824,75 + 0,755361Q_{t-1} \quad (15)$$

5. Análisis de Residuos

Para el análisis de los residuos de los modelos fueron realizados los siguientes estudios:

1. Normalidad de los residuos. Basado en los contrastes: Doornik–Hansen, el cual será graficado; W de Shapiro–Wilk; Lilliefors y Jarque–Bera.
2. Correlación de los residuos. Un análisis de la FAC y LA FACP con la prueba de Box–Pierce basada en las primeras 25 autocorrelaciones, para garantizar que los residuos no están correlacionados.
3. Análisis del Espectro. Una prueba basada en el periodograma muestral, para verificar si los residuos no son periódicos.

5.1. Residuos del Modelo SARIMA(1, 0, 0)(0, 1, 1)₁₂

La Figura 8 muestra el contraste de normalidad de residuos del modelo SARIMA(1, 0, 0)(0, 1, 1)₁₂, de la Ecuación 12, con los resultados: contraste de Doornik–Hansen = 21,460, con p -valor = 0,00002; W de Shapiro–Wilk = 0,986, con p -valor = 0,00001;

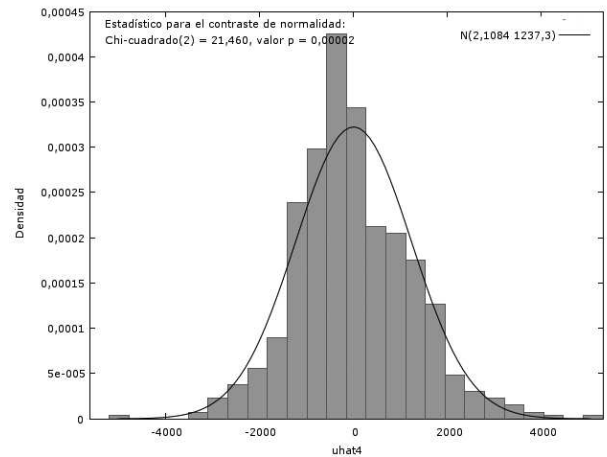


Figura 8: Contraste de Normalidad de los Residuos del Modelo 1

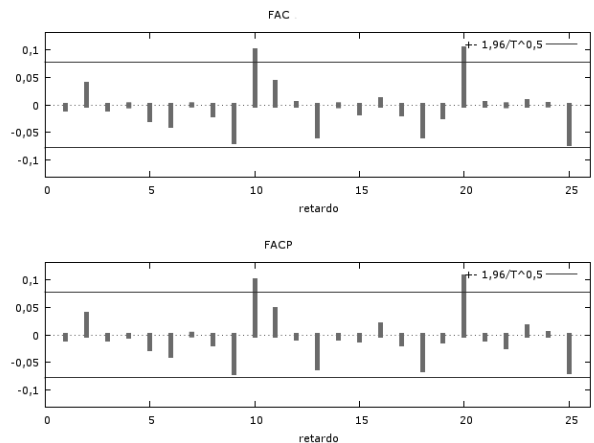


Figura 9: Correlograma de los Residuos del Modelo 1

contraste de Lilliefors = 0,065, con p -valor = 0,0000; contraste de Jarque–Bera = 35,535, con p -valor = 0,0000

La Figura 9 muestra la FAC y la FACP de los residuos del modelo. En la prueba de Box–Pierce el estadístico de prueba para muestras grandes es 27,003, con un p -valor = 0,3558. Como el p -valor para esta prueba es mayor o igual que 0,10, no se puede rechazar la hipótesis de que los residuos son aleatorios, con un nivel de confianza del 90,0 % o mayor, lo cual confirma que no existe correlación significativa en los residuos.

En la Figura 10 se observa el periodograma muestral de los residuos del modelo. Como ninguno de los valores de la densidad espectral

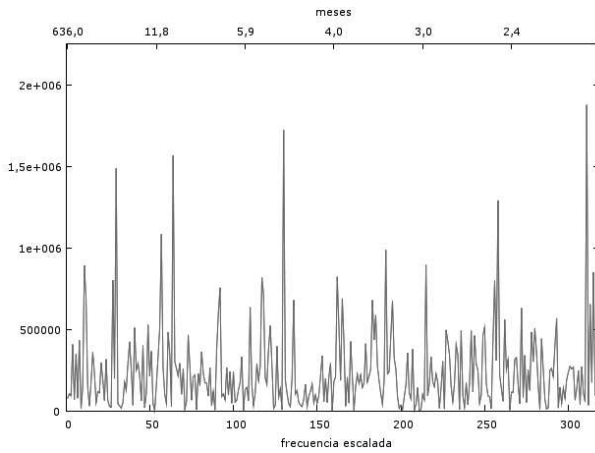


Figura 10: Espectro de los Residuos del Modelo 1

es superior en orden de magnitud a los valores de la demás frecuencias, podemos garantizar que los residuos no son periódicos.

5.2. Residuos del Modelo SARIMA(0, 0, 0)(0, 1, 1)₁₂

Los resultados obtenidos con el modelo SARIMA(0, 0, 0)(0, 1, 1)₁₂ de la Ecuación 13 fueron similares a los del modelo anterior, en los referido a las pruebas de normalidad y periodicidad, pero la prueba de autocorrelación de los residuos arrojó resultados significativos.

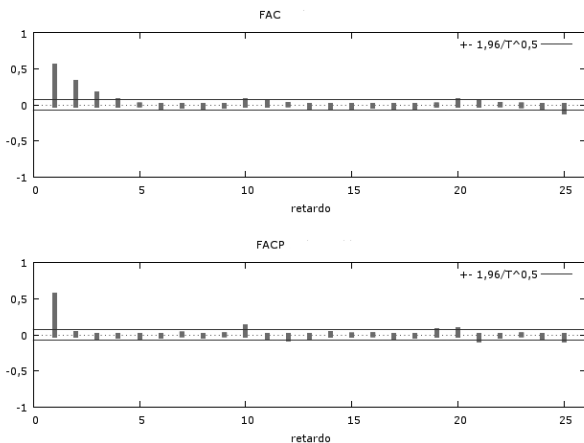


Figura 11: Correlograma de los Residuos del Modelo 2

La Figura 11 representa la FAC y la FACP de los residuos del modelo en evaluación. El estadístico de prueba para muestras grandes en la prueba de

Box–Pierce es 287,177, con un p -valor = 0,000. Como el p -valor para esta prueba es menor o igual que 0,01, se puede rechazar la hipótesis de que los residuos son aleatorios, con un nivel de confianza del 99,0 % o mayor, lo cual confirma que este modelo no recogió la autocorrelación del Caudal.

5.3. Residuos de los Modelos ARMA(1,4) y AR(1)

Si bien, las pruebas de normalidad no marcan diferencias significativas entre la distribución de los residuos de los modelos de las Ecuaciones 14 y 14, y la distribución normal; dichos modelos no recogen la estacionalidad de la serie del Caudal, como puede verse en el análisis de correlogramas y periodogramas dados en las Figuras 12 a 15.

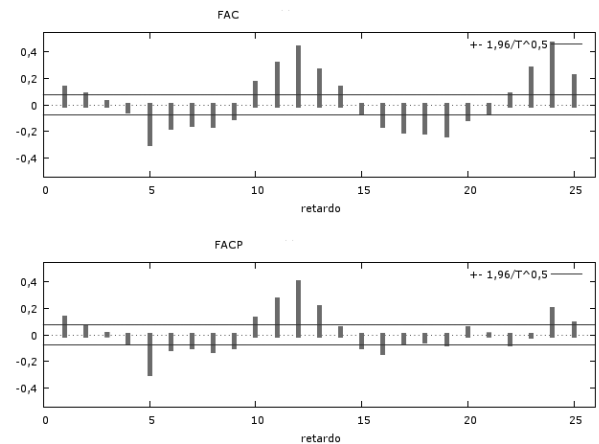


Figura 12: Correlograma de los Residuos del Modelo ARMA(1,4)

Las Figuras 12 y 13 muestran que los modelos ARMA(1,4) y AR(1) no recogen la correlación estacional de los datos. Este resultado puede verificarse mediante el periodograma muestral de los residuos.

Las Figuras 14 y 15 presentan los periodogramas muestrales de los residuos de los modelos, donde se destaca que en ambas series de residuos existe un período, $s = 12$ no incluido por estos modelos.

Los resultados de este análisis de residuos confirman que el modelo de menor longitud de descripción, SARIMA(1, 0, 0)(0, 1, 1)₁₂ es el que modelo se ajusta a los datos.

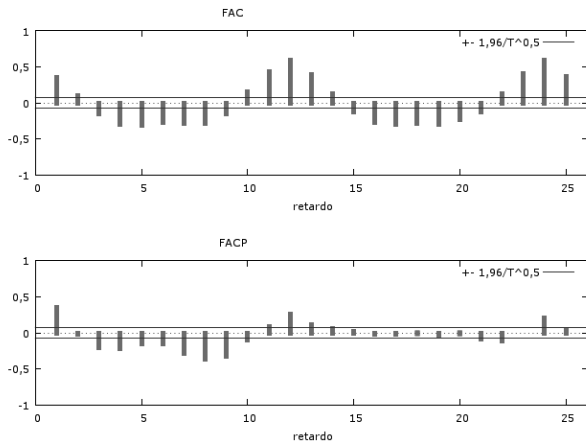


Figura 13: Correlograma de los Residuos del Modelo AR(1)

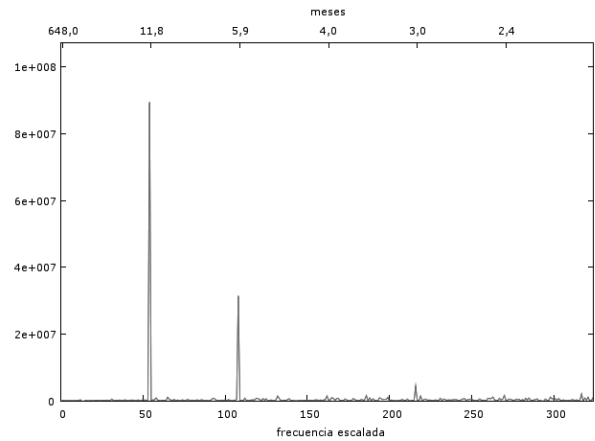


Figura 15: Espectro de los Residuos del Modelo AR(1)

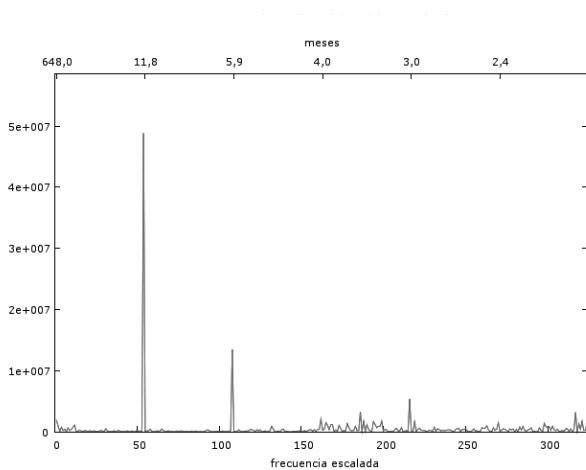


Figura 14: Espectro de los Residuos del Modelo ARMA(1, 4)

La salida de la estimación del modelo SARIMA(1, 0, 0)(0, 1, 1)₁₂ está tabulada en la Tabla 4.

6. Conclusiones y Recomendaciones

En esta investigación se modeló la serie de tiempo de los caudales mensuales del Río Caroní en la estación hidrométrica Guri, mediante la metodología de Box y Jenkins. Para eliminar la periodicidad se usó la diferencia estacional. De los modelos factibles se seleccionó el modelo de menor longitud de descripción, con lo cual se llegó a que el modelo que mejor está ajustado a los datos es SARIMA(1, 0, 0)(0, 1, 1)₁₂, con ecuación

Tabla 4: SARIMA(1, 0, 0)(0, 1, 1)₁₂

	Coef.	σ	z	p -valor
const	13,7421	6,8383	2,0096	0,0445
ϕ_1	0,5434	0,0321	16,9485	0,0000
Θ_1	-0,9998	0,0162	-61,7415	0,0000
μ (V. D.)		-0,5462	σ (V. D.)	2075,764
μ (ϵ_t)		2,1084	σ (ϵ_t)	1245,715
Log-veros.		-5459,617		

	r	Real	Imag.	Mód.	Frec.
AR					
	1	1,8403	0,0000	1,8403	0,0000
MA(sd)					
	1	1,0001	0,0000	1,0001	0,0000

$Q_t = 13,7421 + 0,5434Q_{t-1} - 0,9998\Delta_s\epsilon_{t-1}$. Con el análisis de los residuos se verificó la selección del modelo realizada por el principio MDL.

El modelo presentado facilitará las predicciones de caudal, así como el nivel esperado de la Represa de Guri, en la producción de potencia eléctrica.

Se recomienda usar los procedimientos de inteligencia artificial, como algoritmos genéticos, redes neuronales artificiales y vectores de máquinas de soporte para minimizar el tiempo de computo.

7. Agradecimientos

Los autores agradecen el financiamiento parcial de este trabajo a las ayudas menores otorgadas por el CDCH-UC, así como el Financiamiento

otorgado por el FONACYT en el proyecto de contrato número 201100178.

grado para optar al grado de magister en ingeniería eléctrica, Área de Estudios de Posgrado, Facultad de Ingeniería, Universidad de Carabobo, Valencia, Venezuela, 2010.

Referencias

- [1] G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden Day, San Francisco, USA, 1976.
- [2] Daniel Peña. *Análisis de series Temporales*. Alianza, Madrid, España, 2005.
- [3] J. Rissanen. Modelling by shortest data description. *Automatica*, 14:465–471, 1978.
- [4] Cristóbal E. Vega. *Aplicación de Técnicas Wavelet al Análisis de Series Temporales*. Tesis doctoral, Universidad de Granada, Granada, España, Julio 2003.
- [5] V. T. Chow. *Handbook of Applied Hydrology*. Mac Graw–Hill, New York, USA, 1964.
- [6] Edilberto Guevara Pérez. Análisis estocástico de series de tiempo univariadas. *El Agua*, 25:43 – 53, Julio 1982.
- [7] Edilberto Guevara Pérez. Modelos estocásticos para la simulación de series hidrológicas. *CIDITA. Series Hidrología*, H(13):156, 1983.
- [8] D. R. Maidment. *Handbook of Applied Hydrology*. Mac Graw–Hill, New York, USA, 1993.
- [9] Saeid Amiri and Dietrich Von Rosen. The svm aproach for box–jenkins models. *REVSTAT Statistical Journal*, 7(1), April 2009.
- [10] Cesar O. Seijas. Predicción de series de tiempo usando redes neuronales artificiales. caso de estudio: Series de tiempo econométricas venezolanas. Trabajo de grado para optar al grado de magister en ingeniería eléctrica, Área de Estudios de Posgrado, Facultad de Ingeniería, Universidad de Carabobo, Valencia, Venezuela, 2001.
- [11] J. Salas, C. Saada, and C. Chung. Stochastic modeling and simulation of the Nile river system monthly flows. Technical Report 25, Computing Hydrology Laboratory. ERC. CSU, 1995.
- [12] J. Salas. *Analysis and modeling of hydrologic time series*. *Hadbook of Hydrology*. Mac Graw–Hill, 1993.
- [13] J. Salas, J. Delleuer, W. Lane, and V. Yevjevich. *Applied modeling of hydrologic time series*. Water Resources Publication, Littleton, Colorado, USA, 1980.
- [14] H. Shen and P. Box. *Stochastic approaches to water resources*, volume I AND II. Fort Collins, Colorado, USA, 1976.
- [15] V. Yevjevich. *Stochastic processes in hydrology*. Water Resources Publication, Littleton, Colorado, USA, 1972.
- [16] Cristóbal E. Vega. Un método automático para la elección de la base de ondícula y la selección del umbral para la estimación de señales. *Revista Ingeniería U. C.*, page in press, 2009.
- [17] Jorge E. Guevara. Estudio predictivo del caudal del río caroní utilizando máquinas de vectores de soportes para la planificación de recursos hídricos. Trabajo de