

MODELACIÓN DEL RÉGIMEN DEL CAUDAL MENSUAL EN LA CUENCA BAJA DEL RÍO DAGUA MEDIANTE SERIES DE TIEMPO ESTACIONALES

Jefferson Valdés Basto

*Estudiante. Universidad del Valle, Escuela de Ingeniería Industrial y Estadística, Cali.
jeffer.valdes@gmail.com*

Diego A. Castro Llanos

*Estudiante. Universidad del Valle, Escuela de Ingeniería Industrial y Estadística, Cali.
dicastro05313@hotmail.com*

Wilmar A. Torres López

*Estudiante. Universidad del Valle, Escuela de Ingeniería Industrial y Estadística, Cali.
alexandertor@gmail.com*

Víctor M. González Rojas

*Profesor Asistente. Universidad del Valle, Escuela de Ingeniería Industrial y Estadística, Cali.
vimagor@gmail.com*

Resumen

Este trabajo plantea la elaboración de un modelo de series de tiempo Estacional para el régimen del caudal mensual en la cuenca baja del río Dagua para el periodo junio 1982- junio 2004. Donde se busca pronosticar valores futuros de la serie en los tres meses siguientes a partir de la fecha establecida. Con el modelo propuesto se obtuvieron resultados similares con respecto a los valores reales para la modelación del régimen mensual.

Palabras clave: Series de Tiempo, modelos SARIMA, Régimen caudal agua, Río Dagua, Pronósticos.

Abstract

This work raises the development of a Seasonal time series model for the monthly flow regime in the lower basin of the Dagua River for the period June 1982 - June 2004. Where is looking predict future values of the series within three months from the date established. With this proposed model were obtained similar results with regard to the actual values for modeling monthly regime.

Keywords: Time Series, SARIMA Models, Dagua River, Forecasting

1. Introducción

En la modelación del régimen del caudal del agua se han empleado diversas metodologías y técnicas para la predicción. En muchas ocasiones se realizaron modelos de regresión lineal múltiple, donde se evaluaba el comportamiento de esta serie a partir de otras variables físico-químicas y biológicas innatas en el régimen del caudal; pero son pocos los trabajos en donde se hace uso de la misma serie univariada para la predicción.

Para la obtención de esta serie se tuvo en cuenta la información suministrada por la CVC (Corporación Regional del Valle del Cauca) entidad que ha monitoreado la calidad del agua superficial de la cuenca baja del río Dagua en la estación Bendiciones con registros mensuales históricos que van desde Junio de 1982 hasta Mayo de 2004, para los períodos estacionales.

Este trabajo plantea la elaboración de un modelo de series de tiempo a partir del uso de herramientas estadísticas tales como los modelos SARIMA (Seasonal Autoregressive Intregred Moving Average) en donde se busca modelar la parte estacional con periodicidad anual y los demás componentes de la serie.

Los resultados de los pronósticos tres periodos adelante fueron muy similares comparados con los valores reales de la serie del régimen del caudal del agua.

2. Descripción de la zona de estudio

La cuenca del río Dagua limita con las cuencas del río Calima al norte y del río Anchicayá al sur y tiene la forma de un triángulo alargado hacia el oeste, con su punta cercana a Buenaventura. En esta cuenca se pueden distinguir tres (3) zonas morfológicamente diferentes:

- **La zona costera** (desde la desembocadura hasta Córdoba) con una extensión aproximada de 12 km. Tiene una pendiente mínima, lo cual permite que el cauce se separe en varios brazos ubicados en áreas inundables, sometidos a la acción de la marea (zonas de manglares).



Figura 1. Cuenca Dagua

- **La zona intermedia** es estrecha y muy encajonada, con un ancho de aproximadamente 17 km, el Valle se extiende altitudinalmente entre 150 y casi 2000 m; esta zona está comprendida entre Córdoba y Loboguerrero. En esta zona el río Dagua corre paralelo a la carretera que comunica a Buenaventura y Santiago de Cali. En el trayecto de Loboguerrero y la quebrada Bendiciones con longitud de 19 km se presentan problemas durante los períodos de crecientes con las quebradas que desembocan en el río Dagua.

- **La zona alta** ubicada arriba de Loboguerrero. En esta zona la cuenca está conformada por dos subcuencas: la del río Dagua y la del río Grande.

3. Metodología

Para obtener el modelo SARIMA se tuvo en cuenta la metodología propuesta por Box y Jenkins (1970), donde en la identificación del modelo estacional hace uso de gráficos simulados de diferentes Funciones de Autocorrelación Teóricas (FAC) que son comparadas con la FAC muestral obtenida por la serie de estudio. Para el desarrollo de este modelo se realizaron los siguientes pasos:

Identificación del modelo SARIMA, estimación de parámetros, Validación de supuestos y pronósticos.

En la figura 2 se presenta el gráfico de la serie mensual y la Función de Autocorrelación(FAC):

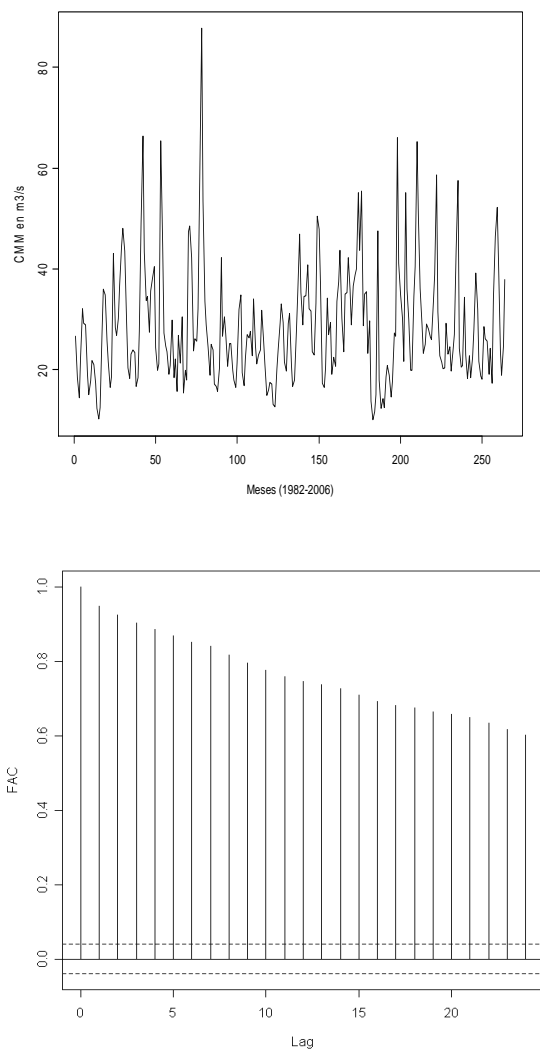


Figura 2. Grafico de la serie y Función de Autocorrelación del régimen del caudal de la Cuenca Baja del río Dagua

En el grafico de la serie se observa una gran variabilidad en el periodo de estudio pero no se encuentra alguna tendencia, además según la FAC la serie no es estacionaria en cuanto a nivel, por lo cual se hace necesario aplicar una transformación a la serie para volverla estacionaria.

3.1. Identificación del modelo SARIMA

Para la identificación del modelo primero se transformo la serie con logaritmo natural para estabilizarla en varianza, luego se diferencio una vez para estabilizarla en media al igual que en un

modelo ARIMA convencional. Se grafico la Función de Autocorrelación FAC de la serie transformada tal y como se muestra en la figura 3.

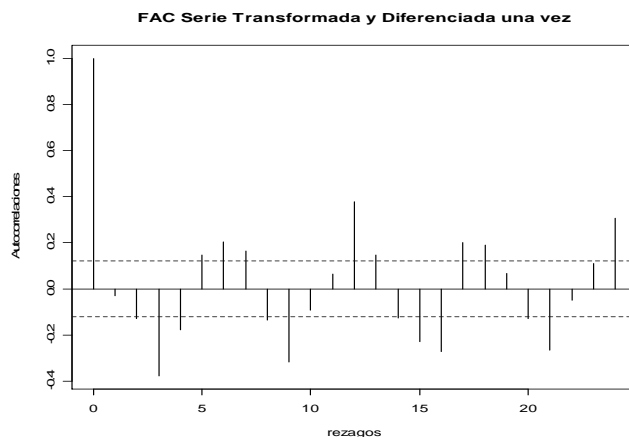


Figura 3. Función de Autocorrelacion de la serie transformada diferenciada una vez

En el gráfico de la FAC se observa que la serie transformada es estacionaria en nivel con la primera diferencia y con la transformación se logra la estabilidad en varianza. No se podría pensar en postular un modelo de medias móviles (MA) de orden bajo debido a que varios rezagos se salen de las bandas de confianza, además de ello se presenta una estacionalidad en la serie y razón por la cual se modelará mediante un ARIMA Estacional (SARIMA).

Siguiendo a Guerrero (2003) el modelo teórico que mejor se aproxima a la Figura 2 viene dado por

$$(1 - \phi B)W_t = (1 - \theta_1 B - \theta_E B^E - \theta_{E+1} B^{E+1})a_t \quad (1)$$

Donde W_t representa la serie transformada y diferenciada una vez y E es el periodo donde se presenta la estacionalidad.

El anterior modelo viene dado de manera general por:

$$ARIMA(1, d, E+1) \times (0, D, 0)_E \quad (2)$$

con $E \geq 3$

De esta forma se requieren d diferencias regulares y D diferencias estacionales para volver estacionaria la serie del Caudal Medio Mensual (CMM_t). En la práctica el grado de diferenciación estacional es comúnmente $D \leq 1$, aunque podría presentarse el caso, un tanto infrecuente, en que este valor sea igual a 2, el cual sucede cuando el patrón estacional sea creciente.

En primera instancia se hizo $D = 0$, luego $D = 1$ ya que se perderán solamente 12 observaciones que corresponde de Junio 1982 hasta Mayo de 1983. No se realizó $D = 2$, debido a que el patrón estacional no es creciente.

Con base en lo anterior se usó una diferenciación estacional, y a partir de esta nueva serie se graficó la FAC correspondiente para determinar si la serie necesitaba ser diferenciada o no para volverla estacionaria en nivel. En la literatura primero se determina el grado de diferenciación estacional y luego el grado de diferenciación para que la serie sea estacionaria en media.

En las figuras 4 y 5 se presentará las FAC para la serie estacional sin diferenciar y diferenciada una vez.

Se puede observar que en la FAC sin diferenciar las autocorrelaciones muestrales decaen lentamente por lo cual no se puede afirmar que la serie estacional es estacionaria en nivel, mientras que en la gráfica de la FAC diferenciada una vez los rezagos decaen rápidamente a cero a excepción del rezago 12, sin embargo esto puede suceder por la aleatoriedad. Por lo tanto se trabajará con una sola diferencia, considerando la serie estacional con periodo estacional igual a 12 ($E = 12$).

Un criterio para decidir los valores de D y d sin tener en cuenta los gráficos de la FAC muestral, consiste en calcular las desviaciones estándar $S(d, D)$; donde D corresponde al grado estacional y d el grado de estacionariedad de la serie.

La serie estacionaria será aquella que satisfaga el siguiente criterio:

$$S(d, D) = \min\{S(j, i); j = 0, 1, 2; i = 0, 1\} \quad (3)$$

Las desviaciones estándar de la serie transformada para el Caudal Medio Mensual, teniendo en cuenta lo anterior vienen dadas por:

$$\begin{aligned} S(0,0) &= 0,16527720 & S(0,1) &= 0,202465579 \\ S(1,0) &= 0,17616366 & S(1,1) &= 0,16150792 \\ S(2,0) &= 0,20972614 & S(2,1) &= 0,25487372 \end{aligned}$$

De las seis series consideradas, la que menor desviación estándar presenta es $S(1, 1)$, lo cual lleva a la misma conclusión usando las FAC muestrales, lo cual hace pensar que la serie es estacionaria en nivel usando $d = D = 1$.

Con lo dicho anteriormente se tiene un modelo de la forma:

$$ARIMA(1,1, E + 1)x(0,1,0)_E \quad \text{Donde } d = 1, \\ E = 12, D = 1.$$

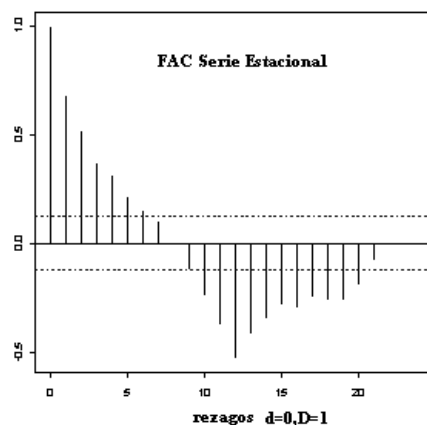


Figura 4. FAC de la serie estacional

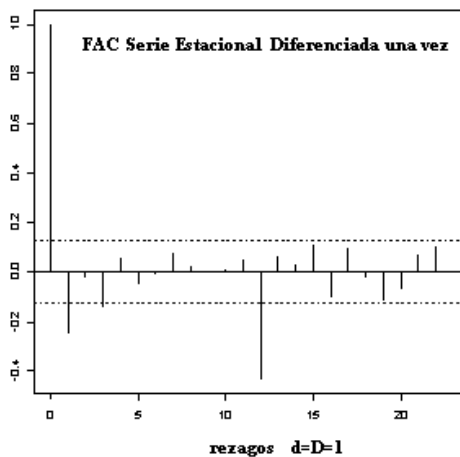


Figura 5. FAC de la serie estacional diferenciada una vez.

3.2. Estimación de los parámetros del modelo SARIMA

Las estimaciones de los parámetros para el modelo se obtuvieron a partir del software estadístico Minitab 14.3:

Final Estimates of Parameters

Type	Coef	SE Coef	T	Valor P
AR 1	-0.8552	0.3361	-2.54	0.013
MA 1	-0.5320	0.3461	-1.54	0.129
MA 12	0.2966	0.1198	2.48	0.016
MA 13	0.0331	0.2030	0.16	0.871

Se puede observar que dos de los cuatro parámetros estimados son mayores a un nivel de significancia del 5%, en consecuencia **no se rechaza la hipótesis nula** de que los parámetros θ_1 y θ_{13} son significativamente iguales a cero, en consecuencia se debe reducir el número de parámetros, y por ende se debe volver a correr el modelo sin incluir el último de ellos, además se debe tener en cuenta que el modelo con el cual se realizaran los pronósticos debe ser el más apropiado en aras de la parsimonia.

Se estiman nuevamente los parámetros del modelo considerado, sin tener en cuenta el parámetro θ_{13} , obteniendo:

Final Estimates of Parameters

Type	Coef	SE Coef	T	Valor P
AR 1	0.7962	0.1220	6.52	0.000
MA 1	0.5286	0.1528	3.46	0.001
MA 12	0.4402	0.1184	3.72	0.000

Se puede observar que los 3 parámetros asociados al modelo propuesto son significativamente diferentes de cero, ya que el p-valor de cada uno es menor que un nivel de significancia del 5%, por lo cual se rechaza la hipótesis nula de que estos son iguales a cero, por ende este es el modelo más adecuado para modelar la serie del Régimen del Caudal Medio Mensual

De esta manera el modelo estimado viene dado por:

$$(1 - 0.7962B)W_t = (1 - 0.5286B - 0.4402B^{12})a_t;$$

Donde $W_t = \nabla T(CMM_t)$

3.3. Validación de supuestos del modelo SARIMA

Se verificaron todos los supuestos del modelo según la metodología propuesta por Guerrero (2003) observando que se cumplen todos los supuestos estadísticos requeridos para que este modelo serie-temporal tenga validez en los pronósticos.

3.4. Predicciones

Las predicciones para los meses de Julio, Agosto y Septiembre de 2006 se calcularon en el paquete estadístico Minitab 14.3, obteniendo las siguientes estimaciones:

Period	Forecast	Limits 95 Percent		Predicción
		Lower	Upper	
Jul-06	1,497163	1,36411	1,630216	31,4169
Ago-06	1,305129	1,11434	1,495918	20,1897
Sep-06	1,242966	1,09500	1,390932	17,4971

Los valores originales obtenidos para el régimen de CMM en la estación bendiciones para Julio, Agosto y Septiembre de 2006 son 30.1514; 18.1987 y 19.1431 respectivamente. Obsérvese que los errores $e = y - \hat{y}$ son -1.26, -1.991 y -1.6442 respectivamente.

4. Conclusiones

- Se observa que los pronósticos puntuales para los tres periodos son muy parecidos con sus verdaderos valores (es decir, la discrepancia entre los valores reales y los pronósticos es pequeña), por lo tanto el modelo generador de la serie $ARIMA(1,1,E+1)x(0,1,0)_E$ es adecuado para pronosticar el régimen de CMM.
- Los intervalos de confianza al 95% de confianza contienen el verdadero valor del pronóstico, por lo cual los pronósticos no se

alejan tanto de la realidad, además el ancho del intervalo de confianza es pequeño, aunque este se incrementa cada vez un poco para los pronósticos de agosto y septiembre de 2006.

- En este trabajo se encontró que se puede hacer uso de modelos de series de tiempo estacionales para saber cómo será el comportamiento a futuro del régimen del caudal de la cuenca baja del río Dagua, con aproximaciones muy parecidas a lo que puede ser la realidad del comportamiento del caudal a diferencia de otros modelos donde en muchas ocasiones solo se pueden hacer predicciones dentro del intervalo de tiempo estudiado, cosa diferente ocurre con las series de tiempo en

los cuales se pueden hacer pronósticos a futuro.

5. Referencias Bibliográficas

1. Contreras, J; Espínola R.; Nogales, F.J.; Conejo A. (2003). ARIMA models to predict nextday electricity prices, IEEE Trans. Power Syst., Vol. 18, pp. 495-500.
2. Guerrero, V. M. (2003) Análisis Estadístico de Series de Tiempo Económicas. México, D.F. Ed. Thomson, Segunda Edición.
3. Uriel, E. (1992) Análisis de Series Temporales modelos ARIMA. España, Valencia. Ed. Paraninfo, Segunda Edición.