

Modelos estocásticos para el pronóstico de sequías en la microcuenca del río Chulco en Ecuador

Stochastic models for droughts forecasting in the Chulco river micro-basin in Ecuador

Darío Zhiña¹, Jheimy Pacheco^{2,3}, Alex Avilés^{1,4}

¹Departamento de Recursos Hídricos y Ciencias Ambientales, Universidad de Cuenca,

xavidx316@gmail.com

²Escuela de Ingeniería de Sistemas, Facultad de Administración y Escuela de Biología, Ecología y Gestión, Facultad de Ciencia y Tecnología. Universidad del Azuay

³Instituto de Estudios de Régimen Seccional del Ecuador (IERSE)

[jlpacheco@uazuay.edu.ec](mailto:jl Pacheco@uazuay.edu.ec)

⁴Carrera de Ingeniería Ambiental, Facultad de Ciencias Químicas. Universidad de Cuenca,

alex.aviles@ucuenca.edu.ec

Recibido: 7-10-2016. Aceptado después de revisión: 15-12-2016

Resumen. Las sequías son un fenómeno natural que afecta tanto a los ecosistemas como a las personas que se benefician de diversos servicios hídricos. Anticipar la ocurrencia de este fenómeno podría ayudar a planificar y gestionar los recursos hídricos de una manera adecuada para disminuir los impactos negativos en los sectores urbanos, agrícolas y energéticos. Este estudio tiene el propósito de desarrollar pronósticos de sequías para seis meses adelante mediante los modelos Autoregresivos Integrados de Media Móvil (ARIMA) y utilizando el índice estandarizado de precipitación (SPI) en tres ventanas de tiempo. Para probar la metodología se escogió la microcuenca del río Chulco ubicada en el sistema hidrográfico Machángara-Paute-Santiago al sur del Ecuador. Los resultados muestran que los mejores pronósticos se obtuvieron para tres meses adelante cuando se utilizó los índices SPI3 y SPI6. Este estudio podría ser de utilidad para los tomadores de decisiones en cuencas hidrográficas con respecto a la planificación de medidas y políticas más adecuadas para la gestión de recursos hídricos en épocas de escasez de agua.

Palabras Claves: ARIMA, Cuencas Andinas, Pronóstico de Sequías, SPI.

Abstract. Drought is a natural phenomenon that affects both ecosystems and people benefiting from various water services. Anticipate the occurrence of this phenomenon could help to plan and manage water resources in an appropriate way to reduce the negative impacts in urban, agricultural and energy sectors. This study aims to develop drought forecasts six months ahead through the Autoregressive Integrated Moving Average models (ARIMA) and using the standardized precipitation index (SPI) in three time windows. To test the methodology, Chulco river basin was chosen located in Machángara-Paute-Santiago river system south of Ecuador. The results show that the best predictions were obtained for three months ahead when the SPI3 and SPI6 indices was used. This study could be useful for decision makers in watersheds regarding to plan policies and measures appropriate for the management of water resources in times of scarcity of water.

Keywords: ARIMA, Andean basins, Drought Forecasting, SPI.

1. Introducción

Si se consulta un cierto número de enciclopedias o diccionarios, podemos darnos cuenta que todos ellos tienen un concepto similar sobre las sequías el cual es “deficiencia de precipitaciones durante un período de tiempo relativamente prolongado” [1]. Sin embargo la sequía es un fenómeno relativamente complejo, del cual resulta difícil poder dar un enfoque genérico que abarque todos sus aspectos y satisfaga toda expectativa [2]. Según [2] y [3] entre las causas que pueden originar las sequías están: (a) cambios en la temperatura superficial de los mares y océanos en el que se da un recalentamiento de la superficie del mar, dando como resultado una alteración general de la atmósfera, (b) existencia de una capa de polvo espesa en la atmósfera, la cual impide que la Tierra se enfríe, dando como resultado un aumento lento pero constante de la temperatura, esta capa de polvo se origina debido a la erosión eólica, (c) la quema de combustibles fósiles, el cambio de uso del suelo y la actividad antropogénica que contribuyen a la modificación de la atmósfera dando como resultado alteraciones en los patrones de precipitación. En el Ecuador, las sequías se originan debido a situaciones climáticas anómalas como son: (a) la situación septentrional anómala ocasionada por la corriente fría de Humboldt, (b) La influencia del fenómeno del Niño sobre los regímenes hidro-pluviométricos del Ecuador, en la cual las precipitaciones se ven afectadas y (c) situación anticiclónica que bloquea la acción de los frentes, sistemas convectivos y depresiones, como es el caso de la sierra ecuatoriana y de las estribaciones orientales de la cordillera (Cadier et al. 1997; Rossel, 1997). Además, se ha determinado una relación entre la lluvia del Ecuador y algunos índices climáticos como son la Temperatura Superficial del Mar (TSM), el Índice Oceánico del Niño (ONI) y el índice de Oscilación Sur (IOS) [5].

Las sequías se caracterizan habitualmente por medio de índices, los cuales son indicadores de diferentes variables ambientales e hidrometeorológicas relacionadas con el fenómeno (caudales, precipitación, producción de cultivos, nivel de aguas subterráneas, etc.). Entre los índices de sequía comúnmente utilizados tenemos: El Índice de sequía de Palmer (PDSI), basado en la precipitación y temperatura, mide la duración, intensidad y extensión espacial de la sequía, además es uno de los pocos índices estandarizados que facilita la comparación directa entre PDSI de diferentes regiones [6], su máxima limitación radica en que el índice está basado en reglas y supuestos arbitrarios [7], [8]; El Índice de humedad de cultivos (CMI) evalúa las condiciones de humedad a corto plazo, utiliza datos medios semanales de temperatura y precipitación, sus valores pueden ser utilizados para comparar las condiciones de humedad entre diferentes localidades, sin embargo no permite monitorear las condiciones de humedad que afectan a los cultivos en desarrollo a largo plazo [9]; El Índice de precipitación estandarizado (SPI), utiliza un registro de la precipitación de un lugar, el cual es ajustado a una función de probabilidad gamma que es transformada en una distribución de probabilidad normal estandarizada, su principal ventaja radica en que puede ser calculado para una gran variedad de ventanas de tiempo, es uno de los métodos más sencillos de efectuar y se encuentra entre los índices más usados para el pronóstico de sequías [7], su desventaja radica en el uso de únicamente la precipitación para la obtención del índice [10]; El Índice de caudal estandarizado (SSI), basado en el ajuste de cada serie mensual de caudal a distribuciones de probabilidad distintas que son transformadas en distribuciones de probabilidad normal estandarizadas, permite monitorear las sequías hidrológicas, el método para su cálculo es similar al del SPI, por lo que tiene la ventaja de poder calcularse para varias ventanas de tiempo, su desventaja reside en el uso de los datos de caudal únicamente [11], [12].

Una de las deficiencias en la prevención de los impactos causados por las sequías es la incapacidad de poder predecir con precisión las épocas de escasez de agua; esta incapacidad puede ser debido a la amplia variabilidad espacio temporal de las variables hidrometeorológicas. Por otro lado, las épocas con déficit de agua activan el reto de operar

sistemas de recursos hídricos en situaciones de riesgo de desabastecimiento de agua, ya que esta reducción del recurso podría perjudicar a los diferentes sectores de la población desde el punto de vista ambiental, social y económico. El desarrollo de metodologías que adviertan los tiempos de escasez brinda un buen soporte a la hora de tomar una decisión con respecto al suministro de agua a los diferentes usos y además aumenta la confianza de los gestores del agua en una cuenca para minimizar los impactos en la población en épocas de sequía. Existen diversas metodologías para pronosticar las sequías. Entre las cuales tenemos: Los comúnmente utilizados modelos de Regresión que se basan en la relación entre 2 o más variables cuantitativas, una dependiente (cuyo valor va a determinarse, también llamado predictando) y otras independientes cuyo valor ya es conocido (también llamados predictores) [12], [13]. Los modelos de series de tiempo usados en muchas aplicaciones científicas [14], [15], [16], [17], entre sus ventajas tenemos la capacidad para la identificación, estimación y comprobación de bondad ajuste para el desarrollo de un modelo [12], [15]. Los modelos basados en probabilidades, cuales son útiles en el pronóstico de sequías debido que cuantifican las incertidumbres asociadas con las variables hidrometeorológicas que causan las sequías [7], [12], [18]. Los modelos basados en redes neuronales artificiales, los cuales son un tipo de modelo no lineal flexible que puede descubrir de forma adaptativa los patrones de los datos, pueden aprender de la experiencia y estimar cualquier relación funcional compleja con una alta precisión [19], etc. De todas las metodologías descritas se ha elegido los modelos de series de tiempo para la predicción de sequías mediante la utilización de los modelos Autoregresivos Integrados de Media Móvil (ARIMA) y utilizando el índice SPI en varias ventanas de temporales. Estos modelos tienen la ventaja de proporcionar óptimas predicciones en un plazo inmediato, no requieren de otras variables para su cálculo, además su metodología es sencilla y útil, su única desventaja radica en que al utilizar una sola variable para su desarrollo no toma en cuenta las relaciones que podrían existir entre otras variables [20]. Los modelos ARIMA fueron desarrollados en 1970 por Box y Jenkins, los cuales desarrollaron una metodología destinada a identificar, estimar y diagnosticar modelos de series temporales. Este tipo de modelos generalizan los modelos ARMA permitiendo trabajar con series no estacionarias. Su metodología deja que los datos de la serie temporal de la variable a ser estudiada indique las características de la estructura probabilística que se encuentra en la serie y ayude a pronosticar su futuro. [21].

Para probar la metodología se ha escogido la microcuenca del río Chulco, perteneciente a la subcuenca del río Machángara, la cual tiene una gran importancia debido a que en ella se encuentran el complejo Hidroeléctrico Machángara, llamado así por los proyectos hidroeléctricos más importantes de la empresa ELECAUSTRO como son la las centrales de Saucay y Saymirín abastecidas por los embalses de El Labrado (microcuenca del río Chulco) y Chanlud (microcuenca del río Machángara Alto). El agua almacenada en los embalses también sirve para solventar a la mitad de la demanda de consumo humano de la ciudad de Cuenca y para proveer de recursos hídricos a demandas agrícolas e industriales. Asimismo la zona de estudio tiene una importancia biológica ya que tiene una gran biodiversidad de flora y de fauna en la cuenca media y alta [22]. En la subcuenca del río Machángara ya se han realizado algunos estudios como [24], el cual estuvo enfocado en la predicción de lluvia y caudal. Otros estudios como [7], [18] que desarrollaron un índice de sequía para predecir la ocurrencia y magnitud de los eventos de escasez de agua en la subcuenca mediante cadenas de Markov (MC) y redes Bayesianas (BN), los resultados revelaron que los modelos basados en MC predicen mejor los períodos húmedos y secos, mientras que los modelos basados en BN generan ligeramente pronósticos de las sequías más severas de forma más precisa.

Por lo expuesto se puede apreciar el interés explorativo en esta cuenca estratégica de la región, es por esto la razón del presente estudio que pretende aportar con nuevo conocimiento en la región mediante el desarrollo herramientas de pronóstico de eventos de sequías que podrían

servir para que los gestores del agua formulen oportunas y equitativas medidas de prevención y mitigación para disminuir los impactos de la escasez de agua en el sector urbano, agrícola, energético e industrial.

2. Materiales y métodos

2.1 Descripción de la zona de estudio

La subcuenca del río Machángara, perteneciente a la cuenca del Paute se encuentra ubicada en el austro del Ecuador, en la provincia del Azuay, en el sistema hidrográfico del Santiago. Tiene un área de 355 km² y está dividida por tres microcuencas (Chulco, Machángara Alto y Machángara Bajo). De acuerdo con su ubicación altitudinal, la cuenca corresponde a la clasificación bosque seco montano bajo (bs-MB) con parches de bosques de *Polylepis*, además presenta un clima tropical típico de alta montaña. [25], [26]. Para este estudio escogió la microcuenca del río Chulco para probar la metodología y obtener resultados que podrían servir para la gestión de sequías. Esta microcuenca se encuentra a una altitud entre 3300 – 4000 m.s.n.m. y cruza el río que lleva el mismo nombre (Figura 1).

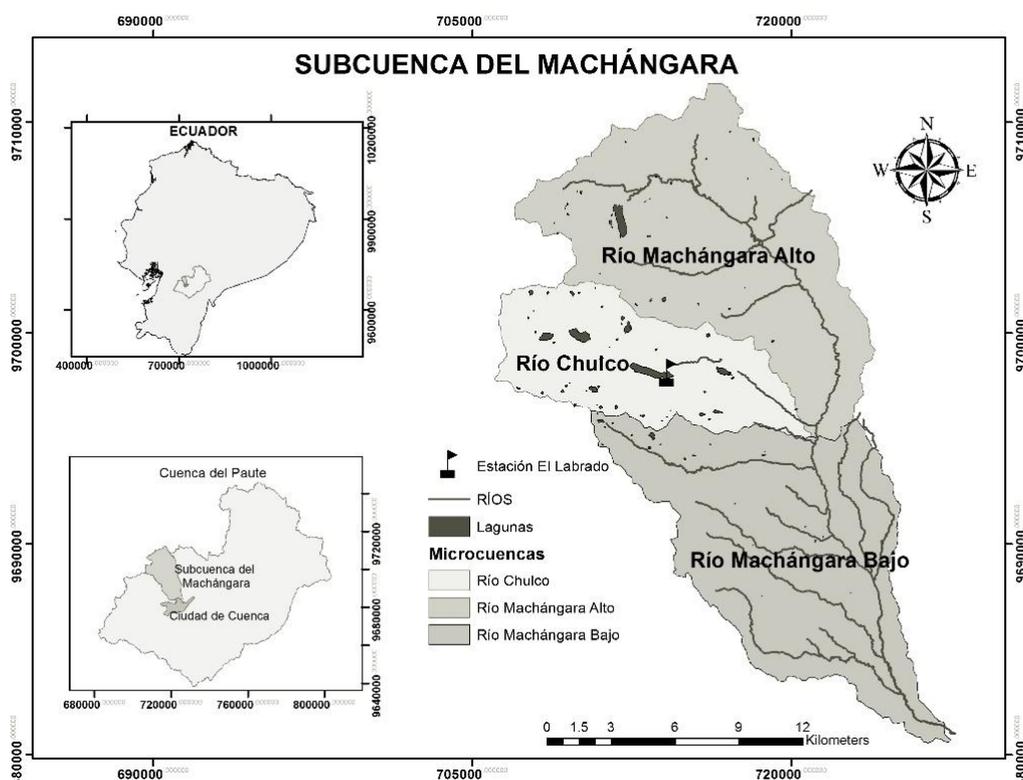


Figura 1. Ubicación de Microcuenca del Río Chulco

Fuente: Instituto Geográfico Militar

De acuerdo a [27] cerca de 13000 hectáreas son irrigadas a lo largo de la subcuenca del Machángara, abasteciendo alrededor de 1900 usuarios, los mismos que se verían afectados de ocurrir sequías en el área. En la cuenca también existen dos plantas de generación hidroeléctrica (Saucay y Saymirín) que generan conjuntamente alrededor de 38420 kW. Además, esta cuenca abastece de agua a más de la mitad de la población de la ciudad de Cuenca (aproximadamente 350000 habitantes).

Series de tiempo mensuales de precipitación (1965-2012) fueron usadas en este estudio para aplicar la metodología. Esta información corresponde a la estación meteorológica de El Labrado proporcionada por el INAMHI (Instituto Nacional de Meteorología e Hidrología). En la Figura 2 se muestra el promedio mensual de la serie de tiempo de precipitación en donde se puede observar mayores lluvias en el mes de Abril, una estación seca durante los meses de junio, agosto y septiembre, y lluvias menos intensas en el mes de Octubre.

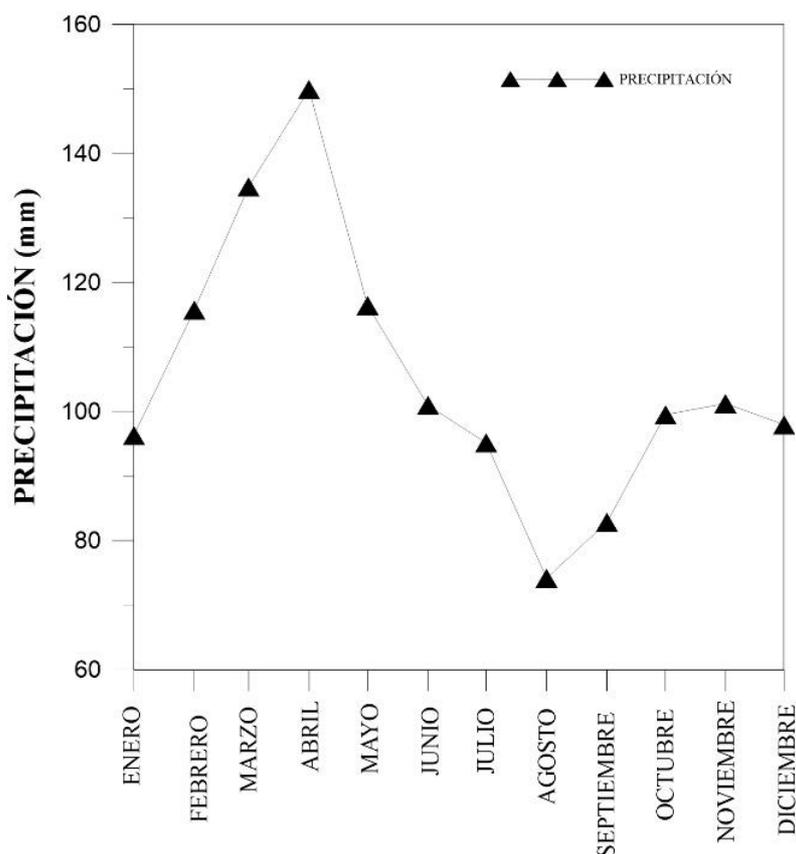


Figura 2. Promedios mensuales de precipitación de la serie de tiempo
Fuente: INAMHI

2.2 Índice SPI

El índice SPI desarrollado por [10] cuantifica las condiciones de déficit o exceso de precipitación en un lugar, para un lapso determinado de tiempo. Para el cálculo del SPI se debe definir el nivel de agrupación de los registros de datos, estos pueden ser mensual, bimensual, trimestral, semestral, etc., de esta manera se forman los SPI como SPI1, SPI2, SPI3, SPI6, etc. [28]. Por ejemplo, si se desea obtener el índice SPI3, implica que deba calcularse para cada mes el valor acumulado trimestral, es decir el valor del mes en consideración sumado a los valores de los 2 meses anteriores. Las series de datos resultantes son ajustados a una función de probabilidad gamma que es transformada a una distribución de probabilidad normal estandarizada.

La distribución función de densidad de probabilidad gamma se define como:

$$g(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \quad \text{Ecuación (1)}$$

Donde x es la variable (en este caso la precipitación), $x > 0$; β es un parámetro de escala, $\beta > 0$; α es un parámetro de forma de la curva, $\alpha > 0$; $\Gamma(\alpha)$ es la función gamma, la cual está definida como:

$$\Gamma(\alpha) = \int_0^{\alpha} y^{\alpha-1} e^{-y} dy \quad \text{Ecuación (2)}$$

[29] sugieren la estimación de los parámetros α y β utilizando la aproximación de Thom [30] para una máxima probabilidad por medio de las siguientes ecuaciones:

$$\hat{\alpha} = \frac{1}{4A} \left(1 + \sqrt{1 + \frac{4A}{3}} \right) \quad \text{Ecuación (3)}$$

$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}} \quad \text{Ecuación (4)}$$

En donde A es igual a:

$$A = \ln(\bar{x}) - \frac{\sum \ln(x)}{n} \quad \text{Ecuación (5)}$$

En donde n es el número observaciones. Los parámetros que se obtienen se los utiliza para encontrar la función de distribución de probabilidad acumulada de un evento de precipitación para un mes o cualquier otra ventana de tiempo.

$$G(x) = \int_0^x g(x) dx = \frac{1}{\hat{\beta}^{\hat{\alpha}} \Gamma(\hat{\alpha})} \int_0^x x^{\hat{\alpha}-1} e^{-x/\hat{\beta}} dx \quad \text{Ecuación (6)}$$

Debido a que la función gamma es indefinida para valores de $x = 0$ y dado que la precipitación puede tener valores nulos, la probabilidad acumulada se convierte en:

$$H(x) = u + (1 - u) G(x) \quad \text{Ecuación (7)}$$

Donde u es la probabilidad de que ocurra una precipitación nula, $1 - u$ es la probabilidad de que no ocurra precipitación nula. La función de probabilidad acumulada $H(x)$, se transforma a la función de probabilidad normal estándar acumulada (Z) con media cero y varianza la unidad mediante una aproximación utilizando las Ecuaciones 8 y 9 [29], [31], [32]. Los valores resultantes de esta transformación se convertirían en la serie de datos del SPI.

$$Z = SPI = - \left(k - \frac{c_0 + c_1 k + c_2 k^2}{1 + d_1 k + d_2 k^2 + d_3 k^3} \right) \quad \text{Ecuación (8)}$$

Para $0 < H(x) \leq 0.5$

$$Z = SPI = + \left(k - \frac{c_0 + c_1k + c_2k^2}{1 + d_1k + d_2k^2 + d_3k^3} \right) \quad \text{Ecuación (9)}$$

Para $0.5 < H(x) < 1$

Donde:

$$k = \sqrt{\ln \left[\frac{1}{(H(x))^2} \right]} \quad \text{Para } 0 < H(x) \leq 0.5 \quad \text{Ecuación (10)}$$

$$k = \sqrt{\ln \left[\frac{1}{(1-H(x))^2} \right]} \quad \text{Para } 0.5 < H(x) < 1 \quad \text{Ecuación (11)}$$

Los valores de las constantes según la aproximación de Abramowitz & Stegun [32], son:

$$c_0 = 2.515517, \quad c_1 = 0.802853, \quad c_2 = 0.010328, \\ d_1 = 1.432788, \quad d_2 = 0.189269, \quad d_3 = 0.001308.$$

El valor promedio de la serie del SPI es cero, valores positivos del SPI son mayores al valor medio; mientras que valores negativos son menores al valor medio. En la Tabla 1 se muestra una categorización de los estados de sequías según los valores del SPI [15].

Tabla 1. Clasificación de las sequías basado en el SPI

VALORES SPI	CATEGORÍA DE SEQUÍA
> 2	Extremadamente húmedo
1.99 – 1.50	Muy húmedo
1.49 – 1.00	Moderadamente húmedo
0.99 – -0.99	Normal
-1.00 – -1.49	Sequía moderada
-1.50 – -1.99	Sequía severa
< -2.00	Sequía extrema

Fuente: [15]

En este estudio, se utilizó 3 ventanas de tiempo 1, 3 y 6 meses (SPI1, SPI3 y SPI6) para probar los modelos ARIMA.

2.3 Modelos ARIMA

Los modelos Autoregresivos (AR), como su nombre lo dice, indica regresiones de sí mismo. Modelos Autoregresivos de p-ésimo orden $\{Y_t\}$ satisfacen la ecuación:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t \quad \text{Ecuación (12)}$$

El valor actual de la serie Y_t , resulta de la combinación lineal de los p valores pasados más recientes de la serie, más un término residual (e_t), que añade todo lo nuevo en la serie en el tiempo que no es posible explicar por los valores pasados [33].

En cambio los modelos de Media Móvil (MA) son una combinación lineal ponderada entre los términos presentes y los q valores pasados de los residuales (e_t), y satisfacen la siguiente ecuación:

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad \text{Ecuación (13)}$$

Los modelos Autoregresivos (AR) se los puede acoplar a los modelos de Media Móvil (MA) formando así un nuevo modelo llamado Autoregresivo de Media Móvil (ARMA), el mismo que satisface la siguiente ecuación:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad \text{Ecuación (14)}$$

Los órdenes del modelo p y q representan los retrasos de la parte autoregresiva y de media móvil. A este tipo de modelos se los conoce como ARMA (p, q) [33].

Modelos ARMA pueden ser usados cuando los datos son estacionarios (parámetros estadísticos constantes) y pueden extenderse a series no estacionarias mediante la diferenciación (d) de las series de datos $W_t = \nabla^d Y_t$, dando como resultado los Modelos Autoregresivos Integrados de Media Móvil (ARIMA). Matemáticamente los modelos ARIMA son modelos AR de orden p y MA de orden q , además operan en d diferencias de una serie de tiempo Y_t . Las tres órdenes (p, d, q) pueden ser valores cero o valores positivos [20]. La ecuación general para los modelos ARIMA es:

$$W_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad \text{Ecuación (15)}$$

En donde $\phi_1, \phi_2, \dots, \phi_p$ son los parámetros de la parte autoregresiva y $\theta_1, \theta_2, \dots, \theta_q$ son los parámetros de la parte de medias móviles.

El desarrollo del modelo ARIMA implicó 3 etapas descritas por [20], las cuales se detallan a continuación:

(i) Identificación del modelo. En esta etapa se determinó qué modelo puede ser el más apropiado para aplicar a la serie de datos. Se analiza la estacionariedad de la serie de datos mediante el test de Dickey Fuller Aumentado, donde la hipótesis nula confirma si una raíz unitaria está presente en un modelo autoregresivo y por lo tanto corroborando que la serie es no estacionaria [34]. Si la serie es no estacionaria se realiza la diferenciación apropiada para lograr la estacionariedad.

Para identificar el mejor modelo es necesario realizar la Función de Autocorrelación (ACF) que para una proceso estocástico $\{Y_t\}$ mide la correlación existente entre dos variables Y_t e Y_{t-k} y la Función de Autocorrelación Parcial (PACF) que mide la correlación entre dos variables Y_t e Y_{t-k} , pero eliminando el efecto de las variables intermedias ($Y_{t-1}, Y_{t-2}, \dots, Y_{t-k-1}$) [33]. En base a estas dos funciones se obtienen los órdenes “ p ” y “ q ” del modelo, la ACF determina el orden de la parte de medias móviles y la PACF determina el orden de la parte autoregresiva. Para la determinación del orden “ d ” es necesario saber si los datos son estacionarios. Si es el caso de estacionariedad la orden “ d ” es cero, resultando un modelo

ARIMA ($p, 0, q$) o también llamados ARMA (p, q). En caso contrario, es decir cuando no existe estacionariedad, es necesario aplicar d diferenciaciones a la serie de datos hasta conseguir los datos sean estacionarios.

Pueden existir varios modelos con diferentes órdenes que se ajusten a las series de datos. En este caso se escoge el mejor modelo utilizando dos criterios utilizados frecuentemente en la literatura [35]. Estos son el Criterio de Información de Akaike (AIC) [36] y el Criterio de Información Bayesiano (BIC) [37]. Estos criterios están basados en una penalización de los modelos con mayor número de parámetros y su diferencia radica en que el BIC toma en cuenta el número de datos n para su cálculo.

La formulación matemática para el AIC y BIC se definen como:

$$AIC = -2 \log L + 2m \quad \text{Ecuación (16)}$$

$$BIC = -2 \log L + m \ln(n) \quad \text{Ecuación (17)}$$

En donde m es el número de parámetros estimados en los modelos ($m = p + q$), L denota la función de verosimilitud del modelo ARMA y n es el número de observaciones. Los mejores modelos son los que tienen el menor valor de AIC y BIC.

(ii) Estimación de los parámetros. Para determinar los parámetros AR y MA se utilizó el método de máxima verosimilitud [33], el cual consiste en encontrar los valores de parámetros que maximicen la probabilidad conjunta de la muestra observada.

(iii) Comprobación de bondad de ajuste. En esta etapa se determinó si el modelo ajustado cumple las hipótesis de normalidad e independencia de los residuos y se estableció los mejores modelos de pronóstico mediante el coeficiente de eficiencia de Nash-Sutcliffe (E) [38]. La normalidad se comprobó mediante el Test de Shapiro [39] y para comprobar la independencia se utilizó la ACF.

Los valores pronosticados y observados del SPI1, SPI3 y SPI6 sirvieron para encontrar los modelos con mejor desempeño mediante el E. Este coeficiente varía de $-\infty$ hasta 1, en donde valores cercanos a la unidad reflejan un mejor comportamiento de los modelos. El cálculo del E está definido por:

$$E = 1 - \frac{\sum_{m=1}^n (Q_m - Q_p)^2}{\sum_{m=1}^n (Q_m - \bar{Q})^2} \quad \text{Ecuación (18)}$$

En donde Q_m es el valor observado, Q_p es el valor pronosticado y \bar{Q} es valor promedio de las observaciones.

2. Resultados y discusión

3.1 Índice SPI

La Figura 3 muestra los índices SPI con diferentes ventanas de tiempo (1, 3 y 6 meses). Un análisis de frecuencias de las series de tiempo del SPI indica que los meses de mayor ocurrencia de sequías son los meses de Junio y Octubre para el SPI1, los meses de Agosto y Noviembre para el SPI3 y los meses de Febrero y Junio para el SPI6. De acuerdo a los tres SPI se puede observar que el mes de Marzo de 1985 fue la época con la sequía más extrema.

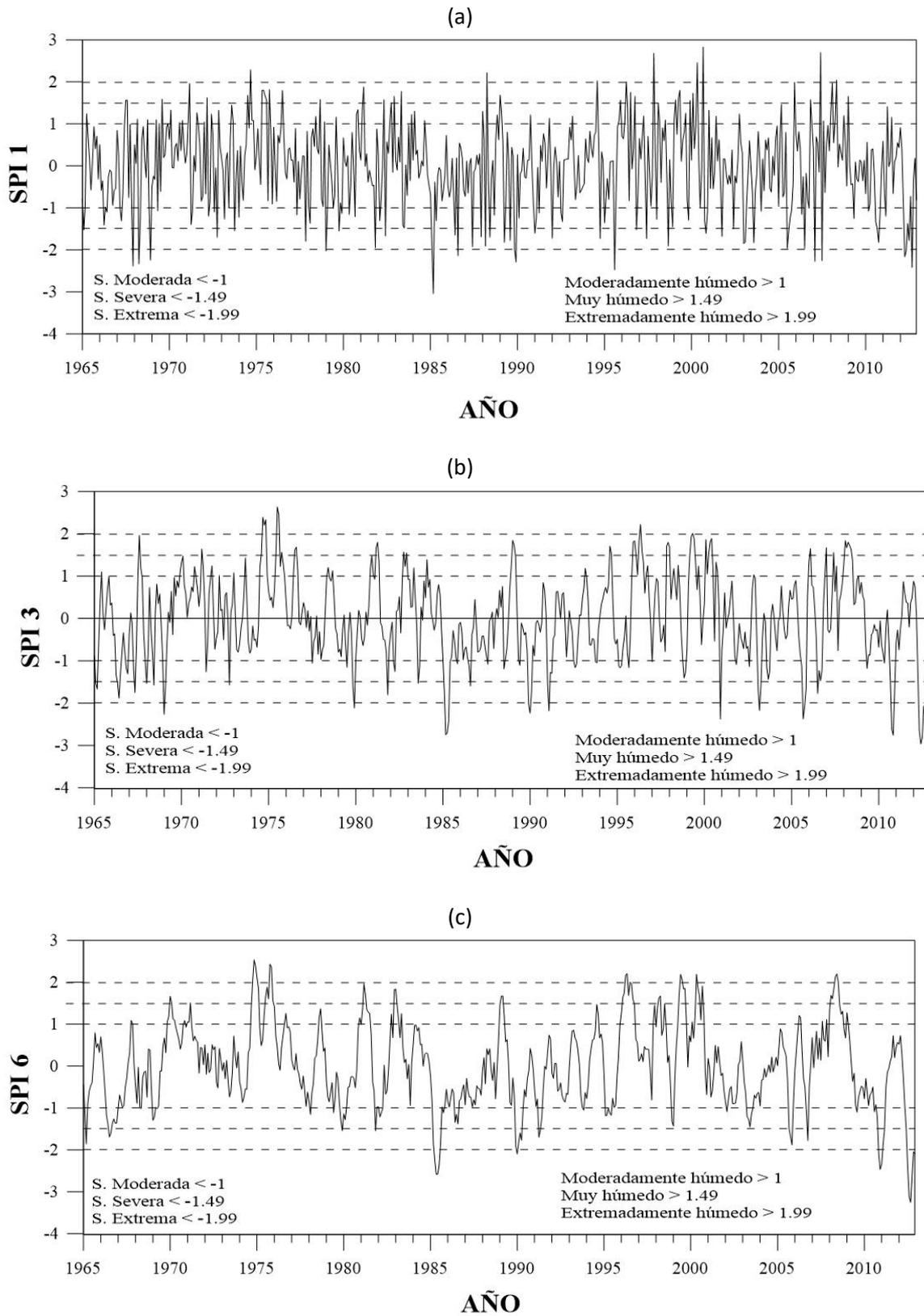


Figura 3. Índices (a) SPI1, (b) SPI3 y (c) SPI6

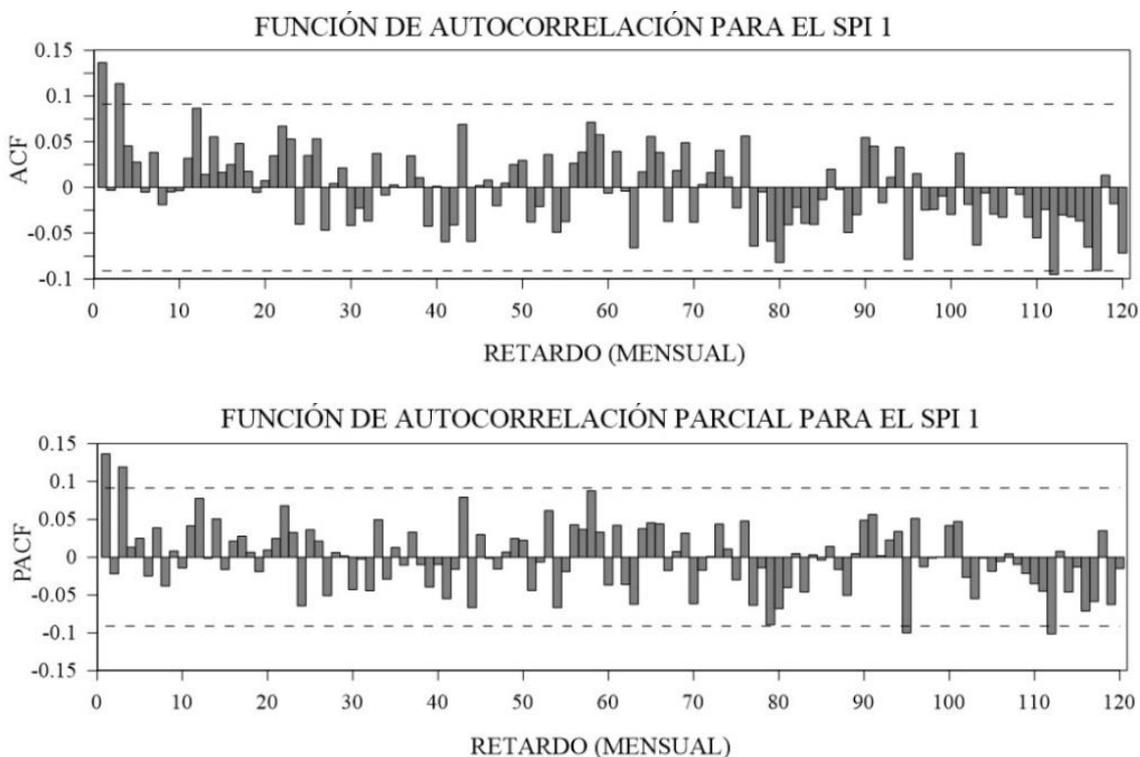
3.2 Modelo ARIMA

Para el periodo de calibración de los modelos ARIMA se tomó en cuenta la serie de tiempo del SPI desde el año 1965 hasta el año 2004 y para el período de validación se tomó en cuenta los valores del SPI desde el año 2005 hasta el año 2012.

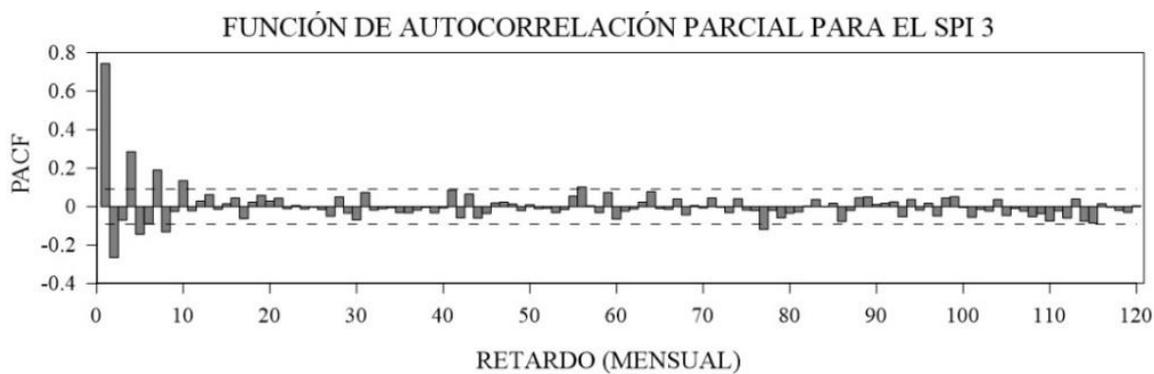
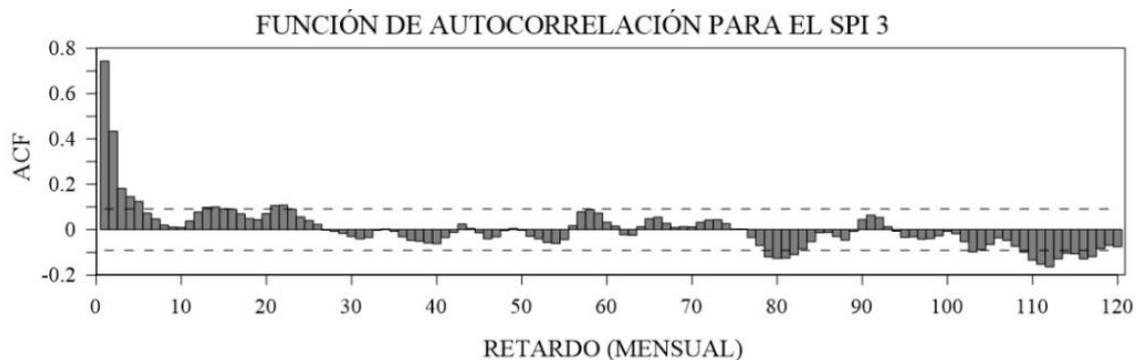
3.2.1 Identificación del modelo

Primero, se determinó si las series de tiempo de los índices SPI1, SPI3 y SPI6, son estacionarias; para ello se aplicó la prueba de Dickey Fuller Aumentado dando como resultado que cada series de tiempo son estacionarias y no es necesario aplicar ninguna diferenciación, por lo tanto se realizó todo el proceso con modelos ARMA (p, q). Para la identificación de las órdenes de los modelos ARMA se obtuvo las ACF y PACF de los índices SPI. Los valores de la ACF muestran que la orden p más probable varía de 0 a 3 para el SPI1, 0 a 4 para el SPI3 y 0 a 4 para el SPI6. Mientras que los valores de la PACF muestran que los valores más probables de q varían de 0 a 3 para el SPI1, 0 a 4 para el SPI3 y 0 a 7 para el SPI6. La figura 4 muestra las gráficas de ACF y PACF de la serie de tiempo del SPI1, SPI3 y SPI6. Con la combinación de los valores de las órdenes de p y q se construyeron 131 modelos ARMA, en los cuales se comprobó la normalidad (mediante el Test de Saphiro) y la independencia (mediante la ACF) de los residuales, resultando 41 modelos ARMA que cumplieron las dos hipótesis (19 modelos para el SPI1, 7 modelos para el SPI3 y 15 modelos para el SPI6). Estos modelos se sometieron a los criterios AIC y BIC para determinar los mejores modelos para cada SPI. La Tabla 2 muestra los mejores modelos con los valores más bajos de AIC y BIC de cada una de las series de tiempo del SPI1, SPI3 y SPI6.

(a)



(b)



(c)

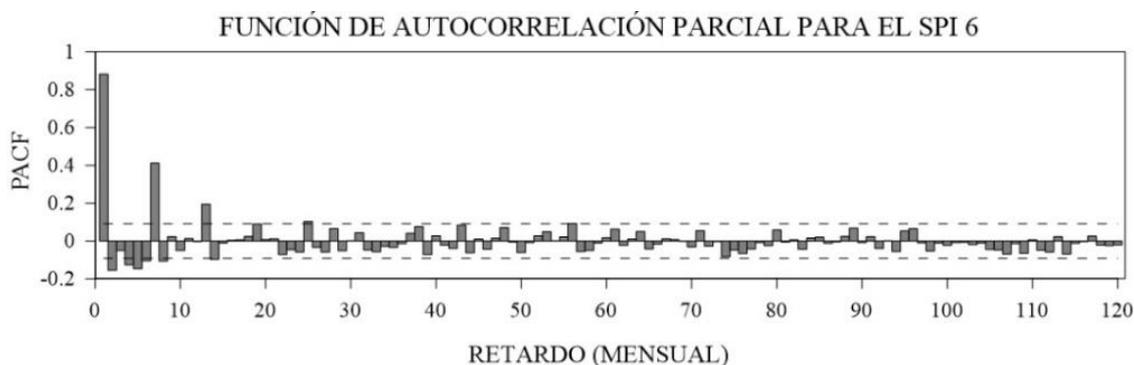
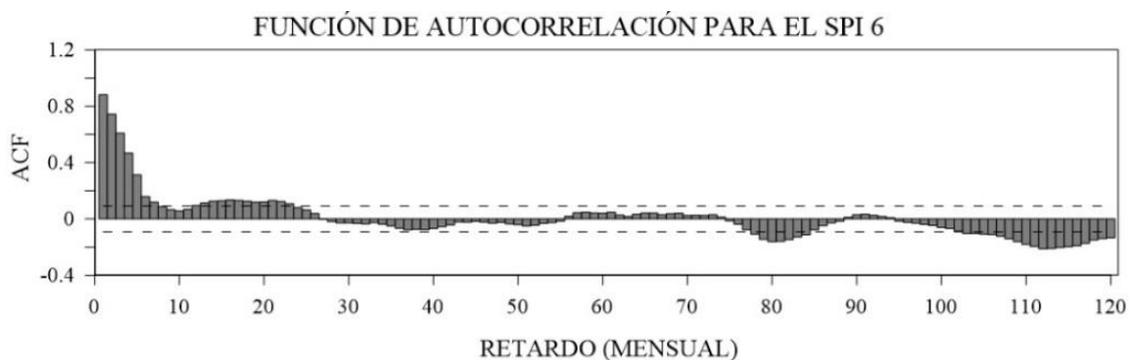


Figura 4. Gráficas ACF y PACF para la serie de tiempo (a) SPI1, (b) SPI3 y (c) SPI6

Tabla 2. AIC y BIC de los mejores modelos ARMA

SPI	MODELO	AIC	BIC
SPI 1	ARMA (3,0)	1374.119	1395.112
SPI 3	ARMA (3,1)	892.201	917.243
SPI 6	ARMA (3,2)	557.283	586.499

3.2.2 Estimación de parámetros del modelo

Utilizando el método de máxima verosimilitud [33] se calculó los parámetros de los mejores modelos ARMA presentados en la Tabla 3, los cuales se detallan en la Tabla 4.

Tabla 3. Parámetros de los mejores modelos ARMA de las series de tiempo del SPI1, SPI3 y SPI6

ÍNDICE	MODELO	PARÁMETROS				
		$\phi 1$	$\phi 2$	$\phi 3$	$\theta 1$	$\theta 2$
SPI1	ARMA (3,0)	0.144	-0.038	0.122	-	-
SPI3	ARMA (3,1)	0.347	0.406	-0.309	0.573	-
SPI6	ARMA (3,2)	0.805	0.628	-0.545	0.296	-0.624

3.2.3 Comprobación de la bondad de ajuste del modelo ARMA

En este punto se muestra el análisis de los residuales de las series de tiempo de los mejores modelos ARMA para el SPI1, SPI3 y SPI6.

La Figura 6 muestra la ACF de las series de tiempo de los residuales para el SPI1, SPI3 y SPI6 para comprobar la independencia entre sus valores. Como se aprecia en las figuras, existen algunos valores en los residuales que sobrepasan la línea de control de autocorrelación (0.09 y -0.09); sin embargo al ser pocos casos con esta particularidad se podría concluir que los residuales son independientes. De igual forma para el SPI1 y SPI6 muestran un similar comportamiento, presentando muy pocos valores que superan los límites de control de la autocorrelación

En la Tabla 4 se muestra los valores p del test de Shapiro de los residuales de los mejores modelos ARMA para cada serie de tiempo del SPI, en donde se puede observar que todos los modelos cumplen la hipótesis de normalidad con un 5 % de significancia.

Tabla 4. Valores p del Test de Shapiro de los residuales de los mejores modelos ARMA para cada serie de tiempo del SPI

SERIE	MODELO	TEST DE SHAPIRO
SPI 1	ARMA (3,0)	0.079
SPI 3	ARMA (3,1)	0.058
SPI 6	ARMA (3,2)	0.067

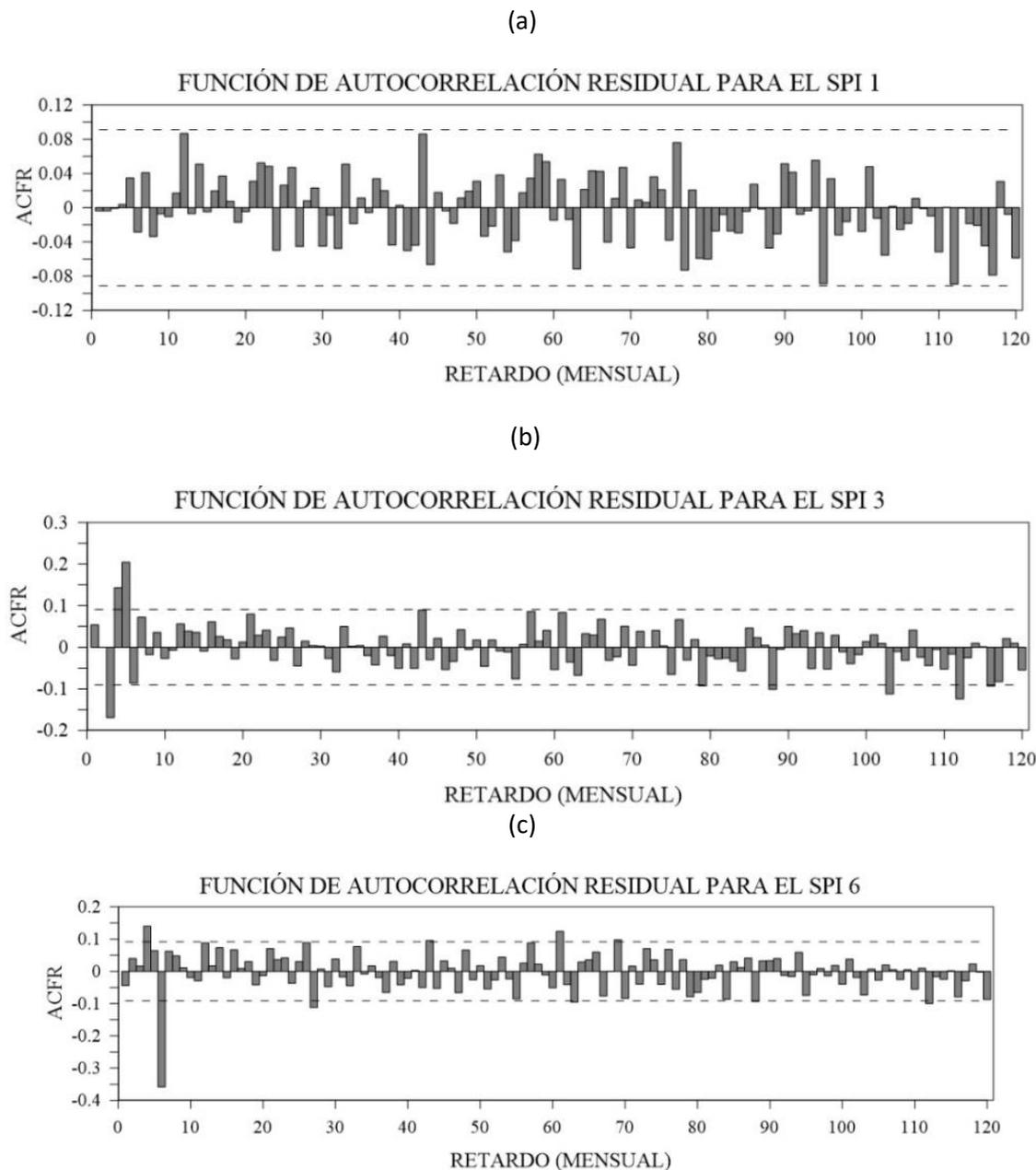


Figura 5. ACF de los residuales del modelo ARMA para el (a) SPI1, (b) SPI3 y (c) SPI 6

Como conclusión general se puede indicar que las series de tiempo de los residuales de los mejores modelos ARMA cumplen las hipótesis de normalidad e independencia.

3.3 Pronóstico y Validación del modelo ARMA

Mediante la utilización de la Ecuación 14 y los parámetros anteriormente identificados se procedió a determinar los pronósticos para el SPI1, SPI3 y SPI6 en el periodo 2005 a 2012. Las gráficas de los valores pronosticados y observados para cada una de las series de tiempo se presentan en las Figuras 6, 7 y 8.

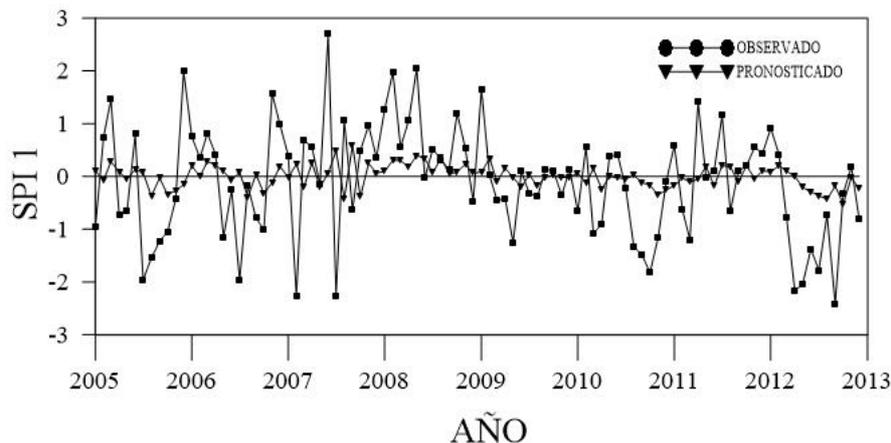


Figura 6. SPI1 observado y pronosticado

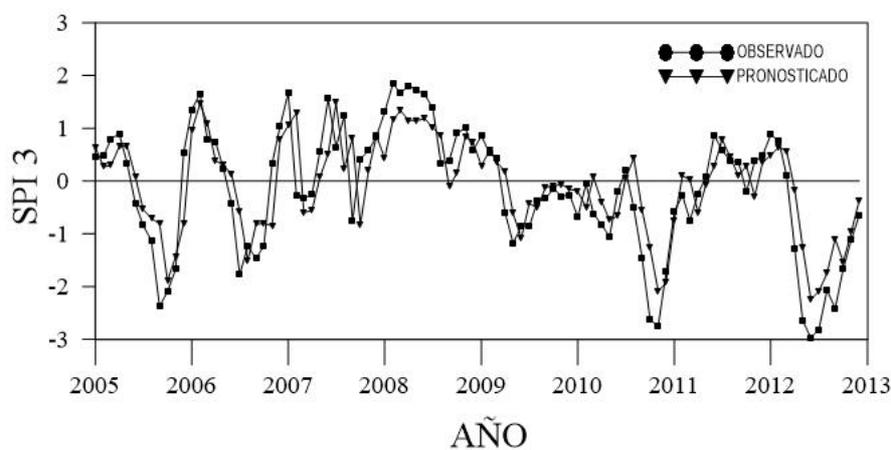


Figura 7. SPI3 observado y pronosticado

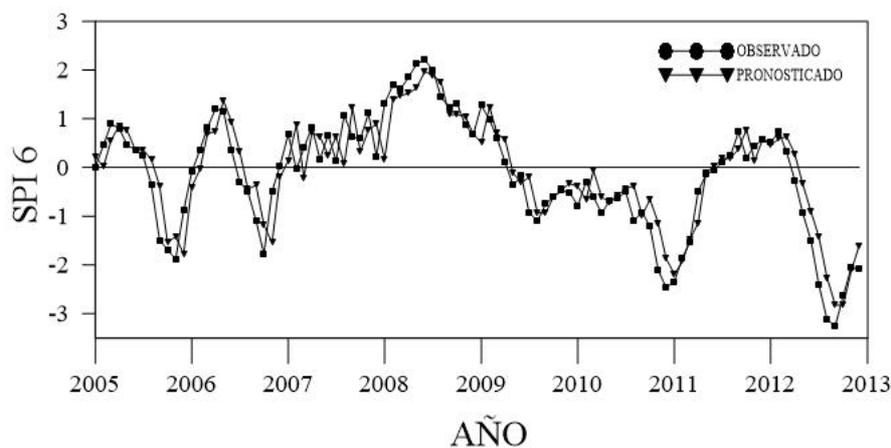


Figura 8. SPI6 observado y pronosticado

Estas figuras muestran que los modelos ARMA se ajustan mejor a las series de tiempo del SPI3 y SPI6, y para la serie de tiempo del SPI1 el modelo no se ajusta de una forma adecuada; esto puede deberse a que las series de tiempo del SPI3 y SPI6 tienen mayor relación temporal entre sus valores (más memoria) en comparación con la serie del SPI1. La Tabla 5 se puede apreciar los valores de los coeficientes de Nash-Sutcliffe para los pronósticos de 6 meses en el futuro, en donde se muestra que los modelos para las series de tiempo SPI3 y SPI6

resultaron ser mejores que el modelo de la serie de tiempo del SPI1 por tener coeficientes más cercanos a la unidad.

Tabla 5. Valores de los coeficientes de Nash-Sutcliffe

Serie	Nash-Sutcliffe					
	1 mes	2 meses	3 meses	4 meses	5 meses	6 meses
SPI1	0.047	0.019	0.024	-0.005	-0.016	-0.018
SPI3	0.704	0.618	0.431	0.263	-0.140	-0.197
SPI6	0.844	0.783	0.657	0.546	0.082	-0.260

También se puede observar que los coeficientes decrecen a medida que las escalas de tiempo de pronóstico son mayores. Lo que nos indica que los mejores pronósticos se realizan a una escala de tiempo de 1 hasta 3 meses.

Haciendo un análisis general se puede observar que los mejores pronósticos se dan con las series de tiempo SPI3 y SPI6 hasta un máximo de 3 meses en el futuro, una mayor escala de tiempo de pronóstico implica obtener valores menos precisos.

Al comparar estos resultados con otros estudios como [15] y [16], hay una coincidencia de que los mejores pronósticos son los que se realizan para los 3 primeros meses en el futuro. Sin embargo estos estudios realizan modelos de pronóstico tomando en cuenta la estacionalidad.

3. Conclusiones

Este estudio se enfocó en pronosticar las sequías en la microcuenca del río Chulco, para lo cual se utilizó los índices SPI en tres ventanas de tiempo 1,3 y 6 meses. Utilizando estos índices se construyeron seis modelos ARIMA no estacionales para el pronóstico de sequías para 1 hasta 6 meses adelante. Los mejores pronósticos se obtuvieron cuando se utilizaron los índices SPI3 y SPI6 y para tres meses en el futuro. Esta metodología se puede aplicar a cualquier otra cuenca de características similares, como es el caso de la microcuenca del río Chanlud, que conjuntamente con la microcuenca del río Chulco alimentan el sistema hídrico del río Machángara.

La modelación de la dependencia temporal de series de datos mediante la formulación de modelos estocásticos, constituyen una manera sencilla y rápida de modelar una realidad donde predominan los procesos aleatorios. Es por esto que la exploración de modelos ARIMA para captar los patrones de sequías meteorológicas mediante el uso del SPI en diferentes escalas podrían ser una buena opción para prevenir y adaptarse a tiempos de escasez con la finalidad de reducir los impactos por la disminución de la disponibilidad de agua en cuencas andinas.

El proceso de identificación, ajuste y comprobación de los modelos ARIMA constituyen una metodología robusta, sencilla y de fácil aplicación por parte de los gestores del agua, comparada con otros modelos estocásticos que tienen métodos más complejos que pueden ser de difícil comprensión y por ende poco aplicados en casos reales.

Finalmente indicar que los resultados del presente estudio podría ser una fuente de información para la toma de decisiones de los gestores de agua en la cuenca con respecto a la aplicación de medidas prevención y mitigación en épocas de sequía para el abastecimiento normal del recurso hídrico a los distintos usos.

Referencias

- [1] Ó. M. Valiente, "Sequía: definiciones, tipologías y métodos de cuantificación," *Investig. geográficas*, no 26, 2001; pp. 59-80, vol. 26, pp. 59–80, 2001.
- [2] I. Velasco, L. Ochoa, and C. Gutiérrez, "Sequía, un problema de perspectiva y gestión," *Región y Soc.*, vol. XVII, no. 34, pp. 35–71, 2005.

- [3] O. U. Martínez and M. R. Carrillo, “Caracterización y análisis de la sequía en el estado de Coahuila.” Nov-2005.
- [4] É. Cadier, G. Gornez, R. Calvez, and F. Rossel, “Inundaciones y Sequias en el Ecuador: el proyecto INSEQ,” Quito, 1997.
- [5] F. Rossel, *Influencia de El Niño sobre los regímenes hidro-pluviométricos del Ecuador, serie INSEQ No.18, tomos I, II y III. INAMHI-ORSTOM, 1997.*
- [6] W. C. Palmer, “Meteorological Drought,” U.S. Weather Bureau, Res. Pap. No. 45. p. 58, 1965.
- [7] A. Avilés, R. Célleri, J. Paredes, and A. Solera, “Evaluation of Markov Chain Based Drought Forecasts in an Andean Regulated River Basin Using the Skill Scores RPS and GMSS,” *Water Resour. Manag.*, pp. 1949–1963, 2015.
- [8] A. K. Mishra and V. P. Singh, “A review of drought concepts,” *J. Hydrol.*, vol. 391, no. 1–2, pp. 202–216, 2010.
- [9] W. C. Palmer, “Keeping Track of Crop Moisture Conditions, Nationwide: The New Crop Moisture Index,” *Weatherwise*, vol. 21, no. 4, pp. 156–161, 1968.
- [10] T. B. Mckee, N. J. Doesken, and J. Kleist, “The relationship of drought frequency and duration to time scales,” no. January, pp. 17–22, 1993.
- [11] S. V. Serrano, J. I. López Moreno, S. Beguería, J. L. Lacruz, C. A. Molina, and E. M. Tejada, “Accurate Computation of a Streamflow Drought Index,” *J. Hydrol. Eng.*, no. February, pp. 318–332, 2012.
- [12] A. K. Mishra and V. P. Singh, “Drought modeling – A review,” *J. Hydrol.*, vol. 403, no. 1–2, pp. 157–175, Jun. 2011.
- [13] V. Kumar and U. Panu, “PREDICTIVE ASSESSMENT OF SEVERITY OF AGRICULTURAL DROUGHTS BASED ON AGRO-CLIMATIC FACTORS,” *J. Am. Water Resour. Assoc.*, vol. 33, no. 6, pp. 1255–1264, Dec. 1997.
- [14] R. Modarres, “Streamflow drought time series forecasting,” *Stoch. Environ. Res. Risk Assess.*, vol. 21, pp. 223–233, 2007.
- [15] A. K. Mishra and V. R. Desai, “Drought forecasting using stochastic models,” *Stoch. Environ. Res. Risk Assess.*, vol. 19, no. 5, pp. 326–339, Jun. 2005.
- [16] Ö. F. Durdu, “Application of linear stochastic models for drought forecasting in the Büyük Menderes river basin, western Turkey,” *Stoch. Environ. Res. Risk Assess.*, vol. 24, no. 8, pp. 1145–1162, Jan. 2010.
- [17] Fernández Cristina, J. Vega, T. Fonturbe, and E. Jiménez, “Streamflow drought time series forecasting: A case study in a small watershed in North West Spain,” *Stoch. Environ. Res. Risk Assess.*, vol. 23, pp. 1063–1070, 2009.
- [18] A. Avilés, R. Célleri, A. Solera, and J. Paredes, “Probabilistic Forecasting of Drought Events Using Markov Chain- and Bayesian Network-Based Models: A Case Study of an Andean Regulated River Basin,” *Water*, vol. 8, no. 2, p. 37, 2016.
- [19] S. Morid, V. Smakhtin, and K. Bagherzadeh, “Drought forecasting using artificial neural networks and time series of drought indices,” *Int. J. Climatol.*, vol. 27, no. 15, pp. 2103–2111, Dec. 2007.
- [20] G. Box and G. Jenkins, *Time series analysis forecasting and control*. San Francisco : Holden-Day, 1976.
- [21] A. Espasa, “PERSPECTIVA HISTÓRICA DE LOS MODELOS ARIMA Y SU UTILIDAD EN EL ANÁLISIS ECONÓMICO,” *Rev. Hist. Económica*, vol. Año IX, Ot, pp. 541–552, 1991.
- [22] ETAPA EP, “COMITÉ DE GESTIÓN DE LA CUENCA DEL MACHÁNGARA,” 2014. [Online]. Available: <http://www.etapa.net.ec/Productos-y-servicios/Gesti%C3%B3n-ambiental/Manejo-de-Cuencas-Hidrogr%C3%A1ficas/Comit%C3%A9-de-Gesti%C3%B3n-de-la-Cuenca-del-Mach%C3%A1ngara>. [Accessed: 15-Apr-2015].
- [23] E. Carchi, “ELABORACIÓN DE UN BALANCE HÍDRICO DE LA CUENCA DEL RÍO MACHANGARA,” Universidad de Cuenca, 2015.
- [24] N. Quito, “PREDICCIÓN DE CAUDALES EN LA CUENCA DEL MACHÁNGARA,” Universidad de Cuenca, 2015.
- [25] Universidad del Azuay - Instituto Geográfico Militar, “INFORMACIÓN TEMÁTICA DIGITAL DE LA CUENCA DEL RÍO PAUTE,” 2004. [Online]. Available: <http://www.uazuay.edu.ec/promsa/paute/home.htm>. [Accessed: 10-Jan-2016].
- [26] W. Buytaert, R. Célleri, B. De Bièvre, F. Cisneros, G. Wyseure, J. Deckers, and R. Hofstede, “Human impact on the hydrology of the Andean páramos,” *Earth-Science Rev.*, vol. 79, no. 1–2, pp. 53–72, 2006.
- [27] Consejo de la cuenca del Machángara, “Informe hidrológico del Plan de Manejo Integral de la subcuenca del Río Machángara,” Cuenca, 2014.
- [28] F. Paredes, J. Luis, and E. Guevara, “Análisis espacial de las sequías meteorológicas en la región de Los Llanos de Venezuela durante el período 1961-1996,” *Rev. Climatol.*, vol. 8, no. February 2016, pp. 15–27, 2008.
- [29] D. C. Edwards and T. B. Mckee, “Characteristics of 20th century drought in the United States at multiple time scales,” *Climatol. Rep* 97-2, pp. 4–29, 1997.

- [30] H. C. S. Thom, "A NOTE ON THE GAMMA DISTRIBUTION," vol. 86, no. 4, pp. 117–122, 1958.
- [31] B. Hughes and M. A. Saunders, "A drought climatology for Europe," *Int. J. Climatol.*, vol. 22, no. 13, pp. 1571–1592, 2002.
- [32] M. Abramowitz and I. Stegun, "Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables (National Bureau of Standards Applied Mathematics Series No. 55)," *Journal of Applied Mechanics*, vol. 32, no. 1, p. 239, 1965.
- [33] J. D. Cryer and K.-S. Chan, *Time Series Analysis. With Applications to R*, no. January, 2008.
- [34] D. A. Said, S. E. ; Dickey, "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order," *Biometrika*, vol. 71, pp. 599–607, 1984.
- [35] F. Díaz, "Selección de modelos mediante criterios de información en análisis factorial. Aspectos teóricos y computacionales," Universidad de Granada, 2011.
- [36] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 716–723, 1974.
- [37] G. Schwartz, "Estimating the dimension of a model," *Ann Stat*, vol. 6, pp. 461–464, 1978.
- [38] J. E. Nash and J. V. Sutcliffe, "River flow forecasting through conceptual models part I - A discussion of principles," *J. Hydrol.*, vol. 10, no. 3, pp. 282–290, 1970.
- [39] S. S. Shapiro and M. B. Wilk, "An Analysis of Variance Test for Normality (Complete Samples)," *Biometrika*, vol. 52, no. 3/4, pp. 591–611, 1965.