COMPARACIÓN DEL MODELO ARIMA Y MARKOVIANO EN LA DETERMINACIÓN DE PRONÓSTICOS DE CAUDALES EN UNA ESTACIÓN HIDROMÉTRICA

Ing. Mónica Arellano Ontiveros, M.C. David Clemente López Pérez
y M.C. Juan Carlos Salinas López

Universidad Autónoma de Nuevo León, México.
E-mail: monica.arellanoon@uanl.edu.mx, david.lopezpr@uanl.edu.mx, juan.salinaslp@uanl.edu.mx

Marco teórico

El efecto que puede contraer el cambio climático es la elevación gradual de la temperatura media del planeta, perjudicando significativamente el ciclo hidrológico lo cual puede contraer ciertos factores como la elevación del nivel del mar, sequías más graves y prolongadas, ondas de calor más severas y frecuentes, aumento en los ciclones tropicales, entre otros.

El análisis de frecuencia de una tormenta nos proporciona información sobre el nivel de intensidad que podría registrarse a lo largo de los años, con base en las condiciones actuales o pronosticadas; existen dos aplicaciones básicas de éste estudio, la primera es determinar la magnitud de una tormenta en un tiempo definido, la segunda es obtener el periodo de retorno de cierta envergadura; para estimar lo anteriormente mencionado se necesita de datos históricos disponibles sobre un caudal o punto de interés, cuanto mayor el número de datos, más representativas serán las probabilidades de continuidad de la crecida.

De acuerdo con la guía de prácticas climatológicas se recomienda utilizar como mínimo 30 años, a pesar de que en la predicción difiera la cantidad de datos a requerir por ejemplo si se dispone a visualizar la temperatura podría utilizarse un periodo de tiempo más corto, refiriéndonos a la precipitación contar con un tiempo mayor a los 30 años sería lo más óptimo (OMM, 2011).

Modelo ARIMA

El modelo ARIMA comprende de un desglose óptimo del comportamiento a largo plazo, al igual que incluye la estructura temporal al tender regresar al camino de un crecimiento ecuánime, en términos generales presenta un comportamiento estocástico dado a quedar determinado por las condiciones iniciales del método.

Lo que conlleva a realizar lo siguiente:

-La recolección de información: se recomienda alrededor de 50 o más datos para el análisis y de esta manera no tener tanta incertidumbre en la estacionalidad de los datos.

-Representación gráfica de la distribución temporal de la información, esto nos será de gran utilidad para determinar si la serie de tiempo es estacionaria; en caso de que en la representación gráfica identifiquemos que no es estacionaria se requerirá de una modificación de la serie para no obtener incertidumbre de los datos obtenidos.

-La identificación del modelo más adecuado se determina con base en las funciones de autocorrelación y autocorrelación parcial tanto en la parte normal como estacional; se puede seleccionar varios modelos para compararlos entre si y poder determinar el modelo que se adapte mejor.

-Se prosigue a la estimación de los parámetros de los coeficientes del modelo elegido.

-Se utilizan diversos procedimientos para juzgar el modelo seleccionado, de igual manera se tendrá que analizar la probabilidad de errores importantes.

-Con base a lo descrito anteriormente se prosigue a seleccionar el modelo final.

-El modelo selecto se usará para la predicción inicial (De la Fuente Fernández).

Proceso de Markov

Se realizará un proceso de Markov discreto en el cual consiste en un proceso estocástico, además de tener la propiedad de que la probabilidad de comportamiento en un futuro no se ve afectada debido a que está totalmente definida si se conoce el estado actual, es decir, el saber cómo se encuentra actualmente no altera la probabilidad esperada en un futuro (Conesa Garcia & Alonso Sarria, 1992). Se expresa que un proceso es a tiempo discreto cuando el compuesto de parámetros es un grupo discreto es decir si una variable aleatoria la llamáramos "T" dicho movible sería igual a 0, 1, 2…n por lo que podría ser una sucesión infinita de variables aleatorias.

Una cadena de Markov es un proceso estocástico a tiempo discreto X_n: n=0, 1,…} con espacio de estados discretos de manera que satisfaga la propiedad de Markov, es decir que cualquier entero n≥0 para estados $x_0,x_1,\ldots,x_{n+1}$ cumpla:

$$P(X_{n+1}=x_{n+1}|X_n=x_n,\ldots,X_1=x_1,X_0=x_0)=P(X_{n+1}=x_{n+1}|X_n=x_n) \quad [1]$$

Con base a lo anterior se dispone a elaborar una matriz de probabilidades de transición entre estados.

$$P=\begin{pmatrix} p_{11} & \cdots & p_{1k} \\ \vdots & \ddots & \vdots \\ p_{k1} & \cdots & p_{kk} \end{pmatrix} \quad [2]$$

La suma de cada renglón de la matriz de transición deberá de ser 1; la cual nos ayudará junto con una distribución inicial para X_0 a elaborar una cadena de Markov (Rincón, 2012).

Definiremos el vector de condiciones iniciales a la matriz que se elaborara con base en el dato terminal registrado en la serie temporal y se localizará en el estado correspondiente.

$$V_t=\left(0,0,\ldots,0,\underbrace{1}_{j-\acute{e}simo},0,\ldots,0\right) \quad [3]$$

Con lo descrito anteriormente se obtiene un proceso de Markov de k pasos. Para poder realizar los pronósticos correspondientes se obtiene la matriz de transición de k pasos de la cadena (p^m) (Ramírez).

Los pronósticos se elaboraron de la siguiente manera:

Primer pronóstico $S_1=V_t P$ [4]

Segundo pronóstico $S_2=V_t P^2=(V_tP)P=S_1P$ [5]

Y así sucesivamente hasta donde sea requerida.

Metodología

El lugar elegido es una estación hidrométrica de la Subcuenca RH24Bf de la Red Hidrográfica Nacional de México, es una superficie que estima una subdivisión de la cuenca hidrográfica Bravo- San Juan perteneciente a la región hidrográfica Bravo-Conchos; dicha subcuenca cuenta con características singulares tanto de escurrimientos como de expansión.

Se obtuvo la ubicación de las estaciones hidrométricas pertenecientes a la cuenca RH24Bf; en la tabla 1 se muestran dichas estaciones con la información disponible de la base de datos del Banco Nacional de Datos de Aguas Superficiales

(BANDAS), el cual fue elaborado por la CONAGUA (Comisión Nacional del Agua) y del Instituto Mexicano de Tecnología del Agua (IMTA)

Tabla 1.- Datos de las estaciones hidrométricas en la zona.

Estación [ID]	Cantidad de datos	Caudal maximo [m^3/s]	Caudal minimo [m^3/s]	Rango [m^3/s]
24327	54	3329,14	5,43	3323,71
24384	23	171,96	0,00	171,96
24198	32	98,94	0,00	98,94

Con lo establecido en el marco teórico mencionado y por la cantidad de datos disponibles se utiliza la estación 24327 "CADEREYTA" (Figura 1), en la cual se aplicó el análisis correspondiente al Modelo ARIMA y Proceso de Markov, obteniendo los siguientes resultados representados en la Figura 2.

Figura 1.- Gastos máximos diarios anuales de la estación hidrométrica "CADEREYTA".

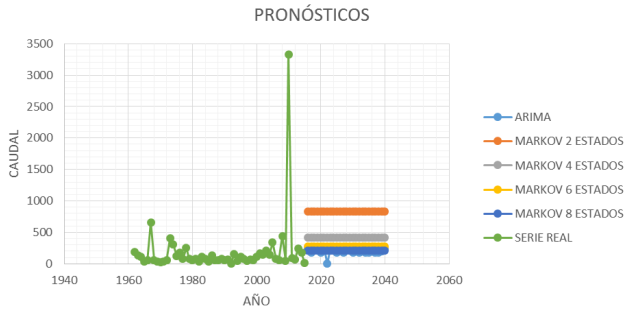

Figura 2.- Serie de tiempo registrado y los pronósticos realizados de la estación 24327 "Cadereyta".

Tabla 2.- Comparación de las metodologías de ARIMA y proceso Markov 8 estados.

Año	Tr (años)	
	ARIMA	MARKOV
2010	80	80
1967	40	3
2008	27	3
1973	20	3
2005	16	3
1974	13	3
1978	11	3
2013	9	2
2003	4	2

Conclusiones

Al obtener los resultados de los métodos analizados en esta investigación, se procedió a realizar un contraste con el fin de tener una comparativa de los periodos de retorno (TR), esto nos ayudaría a poder diseñar obras de mitigación más eficientes ante la probabilidad de eventos extraordinarios y de esta manera poder salvaguardar el bienestar humano.

En conclusión, se pude observar que, con base a la comparación, el método Markov tiene la característica de no alterar la probabilidad de comportamiento futuro por lo que no es necesario una homogenización de datos lo que hace al proceso más exacto al contemplar todos los datos, por lo que se considera más eficaz su predicción que el método ARIMA.

Por lo tanto, utilizar el método Markov es de gran ayuda para la determinación de un Caudal asociado a un TR más preciso, por lo que al comparar las figuras 1 y 3, se aprecia que en el modelo Markov de 8 estados y los gastos máximos diarios anuales de la estación hidrométrica "CADEREYTA" tienen cierta semejanza en sus datos.

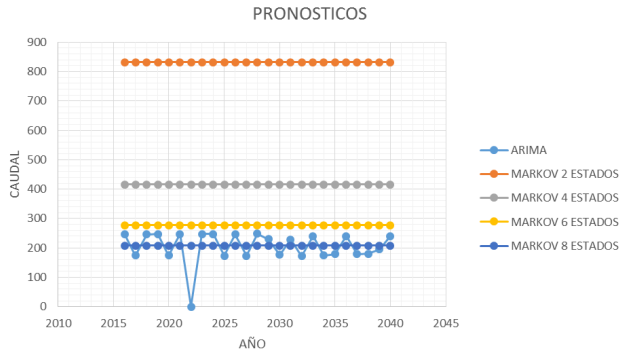

Figura 3.- Pronósticos método ARIMA y proceso Markov.

Se recomienda la realización de más estados en cuestión del proceso de Markov para poder tener mayor aproximación en la estimación del TR al igual que el caudal máximo diario anual esperado, con el fin de tener un ajuste más preciso. El calentamiento global, las condiciones del suelo, la falta de datos de estaciones hidroclimatológicas son factores que pueden ocasionar una gran incertidumbre al momento de obtener la probabilidad del periodo de retorno correspondiente, por lo que se recomienda que en el futuro utilizar otra metodología o un factor de seguridad previniendo este tipo de cambios.

Referencias bibliográficas

(OMM), O. M. (2011). Recuperado el 23 de Diciembre de 2016, de http://www.wmo.int/pages/prog/wcp/ccl/guide/documents/wmo_100_es.pdf

CONAGUA, F. G. (2012). Recuperado el 30 de Diciembre de 2016, de http://biblioteca.semarnat.gob.mx/janium/Documentos/Ciga/Libros2014/229246.pdf

Conesa Garcia, C., & Alonso Sarria, F. (1992). Aplicación de las cadenas de Markov en el análisis de conversión lluvia escorrentia. Actas del V Coloquio de Geografía Cuantitativa, (págs. 373-391). Zaragoza, España. Recuperado el 08 de junio de 2016

De la Fuente Fernández, S. (s.f.). Recuperado el 20 de Octubre de 2016, de http://www.estadistica.net/ECONOMETRIA/SERIES-TEMPORALES/modelo-arima.pdf

INEGI. (04 de diciembre de 2014). Instituto Nacional de Estadística y Geografía. Recuperado el 05 de junio de 2016, de Instituto Nacional de Estadística y Geografía: http://www.inegi.org.mx/geo/contenidos/recnat/hidrologia/regiones_hidrograficas.aspx

Mexicana, S. M. (2002). Constantin Tudor Procesos Estocásticos. Aportaciones Matemáticas.

Rincón, L. (2012). Introducción a la teoría de riesgo. Ciudad Universitaria. Delegación Coyoacán: Facultad de Ciencias UNAM.