

MODELACIÓN DE LOS ESCURRIMIENTOS DEL RÍO YAQUI

Runoff Modeling of the Yaqui River

Isidro Roberto Cruz Medina^{1‡} y Fernando García Páez¹

RESUMEN

El agua es uno de los factores que limitan la agricultura extensiva en las regiones áridas del noroeste de México. La estimación del volumen de agua captado anualmente por los vasos de almacenamiento es de gran importancia porque ayuda a calcular la disponibilidad de agua para usos doméstico e industrial y permite la planeación de actividades tan importantes como la agricultura. Decisiones sustentables sobre el volumen de agua a extraer de los almacenamientos requiere la predicción del escurrimiento. El objetivo de este estudio fue obtener un modelo confiable para la predicción de las escorrentías. Con este propósito, se modelaron las escorrentías del río Yaqui en el estado de Sonora, México, con varios modelos de “caja negra”. Algunos de los modelos estudiados caracterizan la dependencia entre los escurrimientos y, otros, suponen independencia en la serie de datos disponibles, como es el caso de las distribuciones de probabilidad univariadas. Se utilizaron también modelos de regresión dinámica que incorporan variables auxiliares para incluir la posible influencia de las perturbaciones del Pacífico Central bautizadas con los nombres “El Niño” (El Niño-Oscilación del Sur o ENOS) y “La Niña”. Se concluyó que la inclusión de las variables auxiliares que caracterizan a los fenómenos de “El Niño” y “La Niña” no resultó significativa, la serie de escurrimientos anuales puede considerarse independiente, y de los modelos estudiados se eligió una distribución de probabilidad no paramétrica con ventana gaussiana.

Palabras clave: *distribución no paramétrica, regresión dinámica, ENOS.*

¹ Instituto Tecnológico de Sonora. Cinco de Febrero 818 Sur, Obregón, Sonora, México.

[‡] Autor responsable (rcruz@itson.mx)

SUMMARY

Fresh water is one of the limiting factors for extensive agriculture in arid regions of northwestern Mexico. Runoff volume estimation is very important to calculate water availability for municipal, industrial, and agricultural use. Sustainable decisions on the volume of water to be extracted from the reservoirs require prediction of runoff. The objective of this paper was to obtain a prediction model for runoff. The annual runoff of the Yaqui River in the State of Sonora, Mexico, was modeled with various “black box” methods. Some of these models characterize the dependence of the runoff time series, but others assume that the runoff series is independent. Dynamic regression models were used to include the possible effect of the surface water warming phenomenon in the central Pacific Ocean known as the “El Niño Southern Oscillation” (ENSO). It is concluded that dynamic regression, with auxiliary variables that characterize the ENSO event, was not significant, the annual runoff time series can be considered independent and a non-parametric probability distribution with a Gaussian window was chosen.

Index words: *nonparametric distribution, dynamic regression, ENSO.*

INTRODUCCIÓN

La cuenca del río Yaqui comprende 71 452 km²; de esta superficie, el 74% se encuentra dentro del estado de Sonora, el 21% en el estado de Chihuahua y el 5% en el estado de Arizona en Estados Unidos. A lo largo del cauce del río Yaqui se han construido tres presas: “La Angostura”, “El Novillo” y “Oviachic” que en conjunto tienen una capacidad útil de 5900 hm³ (5900 millones de m³) (INIFAP, 2001), con las características que se presentan en el Cuadro 1.

Las escorrentías de la cuenca se utilizan, en su mayoría, para irrigar el valle del Yaqui en el sur de Sonora con una población cercana al medio millón de habitantes y un total de 255 000 ha de riego. El valle del Yaqui es

una de las zonas agrícolas con mayor superficie irrigada en México. Los cultivos más importantes en este valle han sido trigo, algodón, soya, cártamo y maíz. En la última década, por problemas económicos y de sequía, el patrón de cultivos se ha modificado y se ha incrementado la superficie sembrada con hortalizas; sin embargo, el cultivo más importante sigue siendo el trigo con una superficie de siembra usualmente mayor de las 150 000 ha, y un rendimiento promedio de 6 Mg ha⁻¹. En la actualidad este valle, que hace dos décadas fue un emporio agrícola, está en crisis, tanto por el efecto económico del Tratado de Libre Comercio de América del Norte suscrito por México, como por el periodo de sequía y las extracciones que prácticamente vaciaron el sistema de presas y que ocasionó que en el ciclo de invierno 2003-2004 se sembraran menos de 40 000 ha de trigo. El volumen de extracción, para el ciclo agrícola siguiente, lo determinan los administradores y los usuarios del sistema de presas del valle del Yaqui a principios de octubre de cada año (razón por la cual las escorrentías se contabilizan de acuerdo con el ciclo agrícola de octubre a septiembre). Esta decisión, que se toma en condiciones de incertidumbre, tendría mayor justificación técnica si se basara en un modelo estadístico para las escorrentías. Lo anterior justifica el objetivo de este estudio que fue la selección de un modelo probabilístico para la predicción de las escorrentías anuales que son las que proporcionan la disponibilidad hidráulica para el sistema de riego del valle del Yaqui.

MATERIALES Y MÉTODOS

En la modelación de las escorrentías del río Yaqui, se analizaron tres tipos de modelos: ARIMA, regresiones dinámicas y funciones de probabilidad, clasificados como modelos de caja negra por Baratti *et al.* (2003). Estos modelos tienen la ventaja de proporcionar predicciones confiables sin involucrar conocimientos profundos del fenómeno físico; como principal desventaja

sus parámetros usualmente no están relacionados con las características físicas de la cuenca. El análisis inició con la descripción de las escorrentías anuales proporcionadas por la Comisión Nacional del Agua (CNA, 2006). Posteriormente, se analizó la dependencia de la serie, porque la selección del modelo para las escorrentías depende si éstas se declaran o no independientes; con este propósito se analizaron las autocorrelaciones de la serie anual, así como las correlaciones entre y dentro de los periodos de verano y de invierno. Estos análisis condujeron a las siguientes alternativas: 1) Si se acepta que la serie anual de escorrentías es independiente, la mejor predicción, en términos del cuadrado medio del error (CME), para la escorrentía de cualquier año futuro, será el valor promedio de las escorrentías (3090 hm³); la modelación, en este caso consiste en la selección de una función de probabilidad univariada que proporcione las escorrentías, sus probabilidades de ocurrencia y sus periodos de retorno. 2) Si se acepta que la serie es dependiente, esta propiedad se utilizará para mejorar la predicción de las escorrentías futuras con la información de las escorrentías del pasado por medio de los modelos autoregresivos de promedios móviles integrados (autoregressive integrated moving average o modelos ARIMA) (Box y Jenkins, 1972) que permiten modelar la dependencia de una serie, o bien, ajustando distribuciones multivariadas y, posteriormente, utilizando las distribuciones condicionales que proporcionan predicciones que dependen de las escorrentías pasadas. 3) Si la serie depende de fenómenos externos, como por ejemplo, del calentamiento anormal de la superficie del mar en el Océano Pacífico central, fenómeno que se ha bautizado con el nombre de “El Niño” (ENOS en español) y ENSO (El Niño Southern Oscillation, en inglés) y con el enfriamiento de esta misma superficie, fenómeno conocido con el nombre de “La Niña” (Glantz, 1996). Puede ser posible mejorar la predicción de las escorrentías con la información de las variables que

Cuadro 1. Características del sistema de presas del valle del Yaqui.

Año de terminación	Presas	Capacidad total	Superficie de la subcuenca
		hm ³	ha
1941	Lázaro Cárdenas, “La Angostura”	921	19 292
1965	Plutarco E. Calles, “El Novillo”	3 020	40 368
1952	Álvaro Obregón, “Oviachic”	3 227	11 792
Total		7 168	71 452

caracterizan al fenómeno ENOS por medio de modelos de regresión (si las series no están autocorrelacionadas) o por modelos de regresión dinámica (RD) o funciones de transferencia (si las series están autocorrelacionadas). Si los modelos ARIMA y las RD, aun con su mayor grado de complejidad no mejoran significativamente las predicciones proporcionadas por la media aritmética, la modelación consistirá, como en la Alternativa 1, en la selección de una función de probabilidad univariada. Se pueden utilizar innumerables funciones de probabilidad para la modelación de variables hidrológicas y no existe una distribución universalmente aceptada (Silverman, 1986; Mon y Lall, 1994). Yue *et al.* (2001) mencionan que la distribución gamma es la distribución más utilizada en hidrología. En el pasado, por la falta de programas computacionales para el ajuste de distribuciones de probabilidad, las distribuciones empírica y lognormal fueron las más utilizadas. Por ejemplo, De la Peña (2004) utilizó las distribuciones empírica y lognormal y un año después De la Peña (2005) utilizó las distribuciones normal y lognormal para modelar las escorrentías del río Yaqui. En la actualidad, sin la limitación de cómputo, existe una gran variedad de distribuciones utilizadas; Coras *et al.* (2005) ajustaron una distribución Gumbel a las lluvias máximas para periodos de cinco días en Tabasco, y Campos (2008) una distribución de valores extremos a las lluvias máximas diarias en San Luis Potosí. Durante la última década, los métodos no paramétricos se han consolidado como una alternativa a los métodos paramétricos y las densidades no paramétricas con núcleos normales y de Epanechnikov son bastante utilizadas por su facilidad de cálculo y el buen ajuste que proporcionan (Adamowski, 1985; 1996).

RESULTADOS Y DISCUSIÓN

Las escorrentías anuales en el sistema de presas del río Yaqui (CNA, 2006) se presentan en el Cuadro 2. El registro de la serie de datos para las escorrentías de

la presas “La Angostura” inició en el ciclo agrícola octubre-septiembre 1941-1942, el de las dos presas restantes en el ciclo 1964-1965 y el registro de la escorrentía total del sistema en el ciclo 1955-1956.

Las escorrentías totales del ciclo 1955-1956 (1955) al ciclo 2002-2003 (2002), que se calculan por medio de un balance de entradas (escurrimientos, aportes de otros vasos y precipitaciones) y salidas (extracciones y pérdidas) se presentan en el Cuadro 3. La escorrentía total, que proporciona la disponibilidad hidráulica y permite la planeación del sistema de riego del valle del Yaqui (De la Peña, 2004; 2005), tiene un valor mínimo de 999 hm³ que se registró en el ciclo 2002-2003 (en un periodo de sequía que ha ocasionado la mayor crisis económica del valle del Yaqui), ciclo en el cual se registraron los mínimos históricos en todas las presas del sistema, un máximo de 6390 hm³ en el ciclo 1984-1985 y un valor promedio de 3090 hm³ (que es el valor máximo posible para el promedio de los volúmenes anuales de extracción por la ecuación de continuidad). Al analizar la dependencia de las escorrentías mensuales de los datos reportados por CNA (2006), se encontraron tres grupos de escorrentías mensuales independientes pero con correlaciones internas altamente significativas. Estos grupos son: octubre-noviembre, diciembre-junio y julio-septiembre, las dos últimas series de meses se denotarán como periodos de invierno-primavera y verano, respectivamente. En estos periodos se acumula en promedio el 9, 35 y 56%, respectivamente, de las escorrentías anuales. Es importante notar que las escorrentías de estos grupos de meses no se correlacionan en forma significativa, lo cual impide incrementar la precisión de los pronósticos para uno de estos periodos con la información de las series restantes; estos resultados justifican que se analicen y modelen las escorrentías anuales acumuladas del Cuadro 3. La prueba aproximada de Ljung y Box (1978) muestra que la serie de escorrentías del río Yaqui está autocorrelacionada ($\alpha = 0.05$), las autocorrelaciones para

Cuadro 2. Escurrimientos en las presas del sistema de riego del río Yaqui.

Escorrentía	Mínimo	Máximo	Mediana	Q ₁ [†]	Q ₃ [‡]	Media	Desviación estándar
	----- hm ³ -----						
Angostura	97	1424	420	268	659	503	310
Novillo	674	4963	1879	1500	3090	2249	1107
Oviachic	319	1185	695	499	786	675	208
Total	999	6390	2636	2192	3957	3092	1275

[†] Q₁ = primer cuartil; [‡] Q₃ = tercer cuartil.

Cuadro 3. Escorrentías del río Yaqui, ordenadas por columna, para los ciclos agrícolas (octubre-septiembre) 1955-2002.[†]

55-60	61-66	67-72	73-78	79-84	85-90	91-96	97-02
----- hm ³ -----							
1561	2712	6360	2483	2334	4402	4467	2551
2185	2976	2135	2968	3644	2485	4681	2341
3407	3117	2519	2187	1519	3750	2366	1504
3572	1978	2143	2461	5130	2534	3542	2757
3447	4255	3019	2531	4330	4723	2560	1413
1943	2206	4096	4026	6390	5924	1769	999

[†] Las cifras mencionadas en los renglones de cada columna corresponden al seguimiento de los años dentro de los intervalos mencionados. Por ejemplo: 55 (1955), 56, 57, 58, 59, 60 para el intervalo 55-60, etc.

los desfases de 1 y 2 años (lags 1 y 2) son 0.1838 y 0.3059, respectivamente, aunque la significancia aproximada de la autocorrelación no es alta (Valor $P = 0.04$ con la distribución χ^2).

Modelos que Utilizan la Autocorrelación en las Predicciones

En esta sección se utilizan dos métodos para incrementar la precisión de los pronósticos cuando las observaciones están autocorrelacionadas: la esperanza o media de la distribución condicional (que es el estimador de mínimos cuadrados de la variable de interés) y los modelos ARIMA. Cuando las variables tienen una distribución normal bivariada, la media de la distribución condicional, $m_{y/x}$, corresponde a la ecuación de regresión (Cassella y Berger, 1990):

$$m_{y/x} = m_y + r(S_y/S_x)(x - m_x)$$

donde: m_y es la media y S_y es la desviación estándar de la variable que se pronostica Y , m_x es la media y S_x es la desviación estándar de la variable auxiliar en el pronóstico X , y r el coeficiente de correlación entre X y Y . Para predecir las escorrentías de un año determinado utilizando las escorrentías del año anterior, se sustituyen, en esta ecuación, las medias poblacionales m_x y m_y por las medias observadas de los escurrimientos, el coeficiente de correlación r se estima con el coeficiente de autocorrelación r (las desviaciones estándar se cancelan por ser iguales). Después de las sustituciones, el estimador \hat{Y}_t de la escorrentía en el ciclo o año t es:

$$\hat{Y}_t = \bar{Y} + r(Y_{t-1} - \bar{Y})$$

donde: \bar{Y} es la media observada de las escorrentías, Y_{t-1} es la escorrentía observada en el ciclo o año $t-1$ y r es el coeficiente de autocorrelación. Los pronósticos con las distribuciones condicionales, con periodos de desfase (lag) de un año y dos años se obtuvieron con las ecuaciones del Cuadro 4.

Los modelos ARIMA se utilizaron para las escorrentías porque permiten modelar series autocorrelacionadas (Box y Jenkins, 1972) con diferentes tipos de dependencia. La comparación de modelos se efectuó por medio del cuadrado medio del error del modelo (CME); para el cálculo del CME se obtuvieron los pronósticos con los modelos seleccionados y los errores se calcularon, como es usual, por medio de la diferencia entre los pronósticos y los valores observados. El Cuadro 5 muestra, en la segunda columna, el CME de los pronósticos de los modelos estudiados en donde se observa que las distribuciones condicionales no mejoran significativamente el pronóstico que proporciona la media. El modelo ARIMA que produjo el menor CME (1250²), fue un modelo ARMA (1,1), en el cual no se incluye la I en la sigla porque no se efectuaron diferencias en la serie. Este modelo de mayor complejidad tampoco superó significativamente a la media cuando se evaluó por medio del CME. Como los modelos hasta ahora descritos, pueden haber fallado porque no utilizan la información de variables climáticas externas, a continuación se analiza el efecto del fenómeno ENOS en las escorrentías del río Yaqui.

Modelos que Incluyen la Información del Fenómeno ENOS

Varios investigadores han encontrado relaciones significativas entre la precipitación y el fenómeno ENOS. Por ejemplo, Rasmusson y Carpenter (1983)

Cuadro 4. Ecuaciones para la predicción de las escorrentías con las distribuciones normal y log normal.[†]

Desfasamiento	Distribución normal	Distribución log normal
años		
1	$\hat{Y}_t = \bar{Y} + 0.184(Y_{t-1} - \bar{Y})$	$\hat{Y}_t = \text{Anti log} \left[\bar{L} + 0.1715 (L_{t-1} - \bar{L}) \right]$
2	$\hat{Y}_t = \bar{Y} + 0.306 (Y_{t-2} - \bar{Y})$	$\hat{Y}_t = \text{Anti log} \left[\bar{L} + 0.28 (L_{t-2} - \bar{L}) \right]$
1 y 2	$\hat{Y}_t = \bar{Y} + 0.184 (Y_{t-1} - \bar{Y}) + 0.306 (Y_{t-2} - \bar{Y})$	$\hat{Y}_t = \text{Anti log} \left[\bar{L} + 0.1715 (L_{t-1} - \bar{L}) + 0.28 (L_{t-2} - \bar{L}) \right]$

[†] Los símbolos utilizados en este cuadro son los siguientes: Y_{t-i} es la escorrentía observada en el ciclo $t-i$ ($i=1, 2$), \bar{Y} es la media de las escorrentías, \hat{Y}_t es el estimador de la escorrentía en el periodo t , la variable L representa al logaritmo (\log) de las escorrentías $L=\log(Y)$, \bar{L} es la media de los logaritmos de las escorrentías, L_{t-i} es el logaritmo de la escorrentía observada en el ciclo $t-i$ ($i=1, 2$), y Anti log es la función antilogaritmo.

mencionan la influencia del fenómeno ENOS en la India y Sri Lanka. En el año 2003, la revista Geofísica Internacional dedicó un número especial a los efectos de este fenómeno en México, en el cual Magaña *et al.* (2003) describen el impacto de “El Niño” y de “La Niña” en la precipitación en México con mapas, obtenidos por Kriging (que es una técnica de interpolación geoestadística). De estos trabajos es posible inferir que la escorrentía también puede estar relacionada con el fenómeno ENOS. Para modelar la dependencia entre las escorrentías de las series julio-septiembre, octubre-noviembre, diciembre-junio y las escorrentías anuales con el fenómeno ENOS, se utilizó el índice de oscilación del sur (SOI index) proporcionado por Ropelewsky y Jones (1987) y el índice multivariado (MEI) proporcionado por Wolter y Timlin (1998) en modelos de regresión. En estos modelos, se utilizaron los promedios de los índices SOI y MEI de diversos grupos

mensuales con varios periodos de desfase como variables independientes. Se observó que únicamente las escorrentías de la serie diciembre-junio están correlacionadas significativamente con los índices de octubre a marzo con un periodo de desfase de dos meses. El máximo coeficiente de determinación $R^2 = 0.212$ se obtuvo con el promedio del índice SOI de octubre a marzo con la ecuación de regresión:

$$Y = 10237 - 4842 I \tag{1}$$

donde: Y es el escurrimiento de diciembre a junio e I es el promedio del índice SOI de octubre a marzo. Esta ecuación muestra que después de un año fuerte de “El Niño” (con un índice SOI, $I = -2$) se puede esperar un incremento de 968.4 hm³ en la escorrentía del periodo diciembre-junio y que después de un año fuerte de la “La Niña” ($I = +2$) se puede esperar un decremento de la misma magnitud en la escorrentía de este periodo. La Figura 1 muestra la recta de regresión con un intervalo de predicción del 95% para las escorrentías. Las dos escorrentías mayores de la serie diciembre-junio (que coinciden con la ocurrencia de las mayores escorrentías anuales) caen fuera del intervalo de predicción. El CME de los modelos (índices SOI y MEI) del Cuadro 5, se obtuvo ajustando una regresión a la escorrentía anual. La ecuación obtenida es:

$$Y = 3007 - 522.4 I \tag{2}$$

con una $R^2 = 0.145$, que indica que el modelo explica únicamente el 14.5% de la variabilidad de las escorrentías. Observe que el coeficiente de regresión

Cuadro 5. Comparación de los modelos ajustados por medio del cuadrado medio del error (CME).

Modelo	CME
Media	12752
Distribución lognormal condicional lag 1	12702
Distribución lognormal condicional lag 2	12332
Distribución lognormal condicional lag 1 and 2	12482
Distribución normal condicional lag 1	12452
Distribución normal condicional lag 2	12052
Distribución normal condicional lag 1 y 2	11952
AR (1)	12632
ARMA (1,1)	12502
Índice de oscilación del sur	11782
Índice multivariado	12282

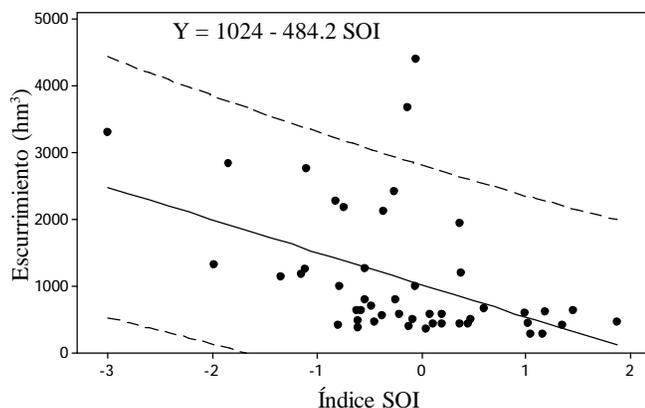


Figura 1. Línea de regresión de la escorrentía de la serie diciembre-junio con el índice de oscilación del sur (SOI) con un intervalo de predicción del 95%.

es similar al obtenido con las escorrentías de diciembre a junio, lo cual confirma que las escorrentías de las series restantes no son influenciadas por el fenómeno ENOS. Las ecuaciones de regresión para las escorrentías de “La Angostura” $Y = 507 - 159 I$ con una $R^2 = 0.22$, del “El Novillo” $Y = 2211 - 337 I$ con una $R^2 = 0.091$ y del “Oviachic” $Y = 670 - 62.9 I$ con una $R^2 = 0.093$, todas significativas a un $\alpha = 0.01$, indican que el fenómeno ENOS afecta en mayor proporción (mayor coeficiente de regresión) a las escorrentías captadas en las presas de “La Angostura” y de “El Novillo”. Las ecuaciones de regresión pueden interpretarse de dos formas: 1) si se supone que la escorrentía y el promedio del índice tienen una distribución normal bivariada, la ecuación de regresión proporciona el estimador de mínimos cuadrados de la media de la distribución condicional; 2) si se rechazara la suposición de que estas variables tienen una distribución normal bivariada, la ecuación de regresión ya no proporciona los estimadores de mínimos cuadrados, aunque se puede utilizar con fines de predicción.

El análisis de regresión supone que la variable de respuesta es independiente y, en este caso, la serie anual de las escorrentías está autocorrelacionada. A favor de la ecuación de regresión se menciona que los residuales del modelo de regresión se declararon no correlacionados por la prueba de Durbin-Watson. Sin embargo, como la autocorrelación de la serie puede afectar la validez del análisis de regresión, se ajustaron regresiones dinámicas (RD) a las escorrentías. Antes de ajustar las RD, Pankratz (1991) recomienda ajustar modelos ARIMA, tanto para la variable de respuesta (las escorrentías), como para la variable auxiliar o función de transferencia;

en este caso, el índice SOI (I), con el propósito múltiple de tener un modelo de pronóstico para Y (cuando el modelo de regresión dinámica no sea adecuado) y para establecer un modelo base para comparar el modelo RD. Los modelos base son importantes, porque es posible, como sucedió en este trabajo, que los modelos de regresión dinámica, con su mayor grado de complejidad, no mejoren significativamente el ajuste obtenido por modelos más simples (evaluación efectuada por el CME). Como se mencionó anteriormente, el modelo ARMA (1,1) no disminuyó significativamente el CME y, al incluir el índice SOI como función de transferencia, la reducción del CME no fue significativa con respecto a la disminución lograda por el modelo ARMA, lo que significa que, al ajustar las escorrentías por autoregresión, el efecto del índice SOI ya no es significativo. Los modelos ARIMA y los modelos de regresión dinámica se obtuvieron por medio de los paquetes estadísticos Minitab (Minitab, 2003) y SAS (SAS Institute, 1999).

Como los modelos estudiados anteriormente no incrementaron significativamente la precisión de los pronósticos y debido a que la autocorrelación con lag 2 como máximo explica el 9.5% (0.306^2) de la variación de los datos, a continuación se ajustan funciones de probabilidad para reproducir la distribución de la serie.

Funciones de Probabilidad

Las distribuciones de probabilidad seleccionadas para las escorrentías del río Yaquí (con las pruebas de Kolmogorov-Smirnov y Ji-cuadrada) fueron: la distribución inversa de Gauss, IG (3090, 17495):

$$f(y) = \left[\frac{17495}{2p y^3} \right]^{1/2} \text{Exp} \left(\frac{-17495 (y - 3091)^2}{2 * 3091^2 y} \right) \text{ para } y > 0 \quad (3)$$

la distribución lognormal, LogN (7.956, 0.404):

$$f(y) = \frac{1}{0.404 y \sqrt{2} p} \text{Exp} \left(\frac{- (\ln y - 7.956)^2}{2 * 0.404^2} \right) \text{ para } y > 0 \quad (4)$$

y una distribución no paramétrica con ventana gaussiana. Aquí $f(y)$ es la función de densidad del escurrimiento y, \ln es la función logaritmo natural y Exp es la función exponencial.

Las distribuciones lognormal y la no paramétrica se muestran en la Figura 2.

Con respecto a escorrentías captadas en cada presa, para “La Angostura”, con las escorrentías del ciclo

agrícola 1941-1942 al ciclo 2002-2003, la distribución de mejor ajuste (con las pruebas de Kolmogorov-Smirnov y Ji-cuadrada) es una distribución inversa de Gauss, IG (503.91, 1027.27).

Para “El Novillo”, con las escorrentías del ciclo 1964-1965 al ciclo 2002-2003, la distribución de mejor ajuste es una distribución Weibull, W (2.2026, 2550.4) con función de densidad:

$$f(y) = \frac{2.2026 y^{2.2026-1}}{2550.4^{2.2026}} \text{Exp}\left[-(y / 2550.4)^{2.2026}\right] \quad \text{para } y > 0, \quad (5)$$

donde: la simbología se explica en las ecuaciones anteriores.

Finalmente, para la presa de “El Oviachic”, con las escorrentías del ciclo 1964-1965 al ciclo 2002-2003, la distribución de mejor ajuste es una distribución Weibull, W (3.5889, 749.47). El Cuadro 6 muestra que las escorrentías proporcionadas por estas distribuciones son semejantes para periodos de retorno menores de 40 años y que las distribuciones paramétricas proporcionan escorrentías mayores para periodos de retorno superiores; esta situación se explica debido a que las distribuciones continuas teóricamente proporcionan escorrentías infinitas y, por esta razón, el área en su cola derecha es mayor.

Para el cálculo de las escorrentías para los periodos de retorno y de las probabilidades necesarias para analizar cualquier propuesta de administración del sistema de presas, se recomienda la distribución no paramétrica por tres razones: no efectúa suposiciones matemáticas, únicamente suaviza al histograma con el núcleo y ventana

Cuadro 6. Escorrentías con las distribuciones seleccionadas para algunos periodos de retorno.

Distribución y escorrentía			Periodo de retorno
Inversa de Gauss	Log normal	No paramétrica	
	hm ³		años
2844	2852	2786	2
4806	4787	4870	10
5559	5543	5838	20
6293	6296	6330	40
7247	7301	6665	100

elegidos (Silverman, 1986) y por el buen ajuste que proporciona (Figura 2).

CONCLUSIONES

La serie de escorrentías anuales del río Yaqui está ligeramente autocorrelacionada con un periodo de desfase de dos años. Esta autocorrelación se utilizó para mejorar la predicción proporcionada por la media anual de las escorrentías por medio de modelos ARIMA y de distribuciones condicionales (que incluyen los valores observados en el pasado). Sin embargo, las predicciones de estos modelos no proporcionaron resultados significativamente mejores [utilizando el criterio del cuadrado medio del error (CME)] que los proporcionados por la media. Los modelos de regresión que incluyen como variable independiente al índice que

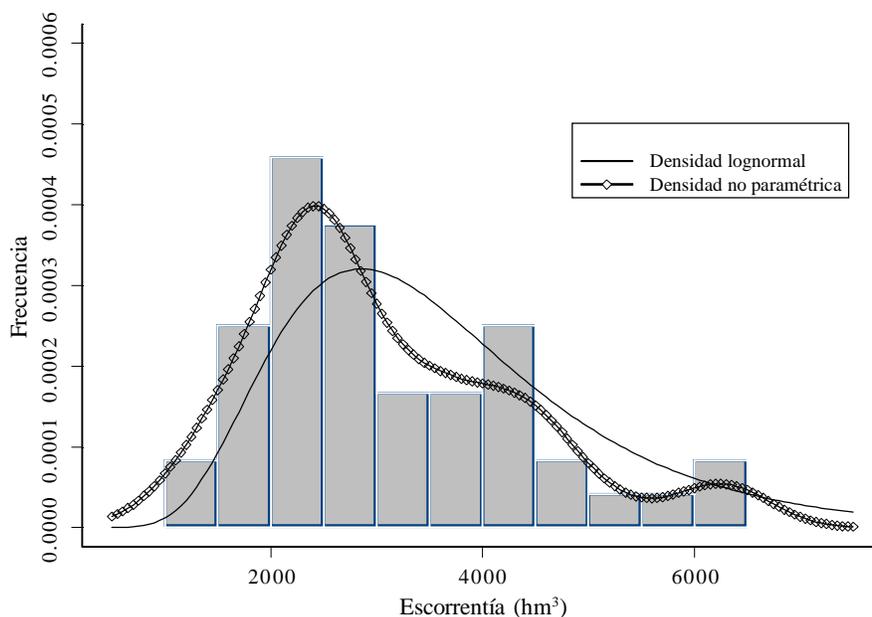


Figura 2. Funciones de probabilidad ajustadas a los escurrimientos del río Yaqui.

caracteriza al fenómeno de “La Niña” y de “El Niño” [fenómenos El Niño Southern Oscillation (ENOS)] permiten concluir que: a) la serie de escorrentías de julio-noviembre (en donde se acumula el 65% de la escorrentía anual) y que corresponde al periodo de lluvias conocido como el “monzón mexicano” no está correlacionado con el fenómeno ENOS, b) después de un año fuerte de “El Niño” (con un índice SOI, $I = -2$) se puede esperar un incremento de 968.4 hm³ en las escorrentías del periodo diciembre-junio y después de un año fuerte de “La Niña” ($I = +2$) se puede esperar un decremento de la misma magnitud en las escorrentías de este periodo que si están relacionadas significativamente con el fenómeno ENOS, c) el fenómeno ENOS actúa con un retraso de dos meses en las escorrentías de la serie diciembre-junio, d) el fenómeno ENOS influye más en las escorrentías captadas en las presas de “La Angostura y de “El Novillo”. Finalmente, como las distribuciones condicionales, los modelos de regresión (Ecuaciones 1 y 2), los modelos ARIMA, AR (1) y ARMA (1,1) del Cuadro 5 y el modelo RD, con su mayor complejidad, no mejoraron significativamente las predicciones proporcionadas por la media aritmética, se ajustaron distribuciones de probabilidad paramétricas (Ecuaciones 3, 4 y 5) y no paramétricas, de las cuales se seleccionó una distribución no paramétrica con ventana gaussiana (Figura 2) para modelar las escorrentías del río Yaqui. Con la selección de esta distribución, que sin suposiciones matemáticas proporciona las escorrentías para diversos periodos de retorno y estimaciones similares a las distribuciones paramétricas de mejor ajuste, se cumple el objetivo de este estudio.

LITERATURA CITADA

- Adamowski, K. 1985. Nonparametric kernel estimation of flood frequencies. *Water Resour. Res.* 21: 1885-1890.
- Adamowski, K. 1996. Nonparametric estimation of low-flow frequencies. *J. Hydraulic Eng.* 122: 46-49.
- Baratti, R., B. Cannas, and A. Fanni. 2003. River flow forecast for reservoir management through neural networks. *Neurocomputing* 55: 421-447.
- Box, G. E. P. and G. M. Jenkins. 1972. *Time series analysis forecasting and control*. Holden-Day. Oakland, CA.
- Campos-Aranda, D. F. 2008. Ajuste regional de la distribución GVE en 34 estaciones pluviométricas de la zona Huasteca de San Luis Potosí, México. *Agrociencia* 42: 57-70.
- Cassella, G. and R. L. Berger. 1990. *Statistical inference*. Duxbury Press. Pacific Grove, CA, USA.
- CNA (Comisión Nacional del Agua). 2006. Datos estadísticos del Departamento de Hidrometría, Distrito de Riego No. 041, Río Yaqui, Sonora, México. Cd. Obregón, Sonora, México.
- Coras-Merino, P. M., R. Arteaga Ramírez y M. Vázquez Peña. 2005. Análisis de frecuencias de lluvias máximas con fines de drenaje superficial local, Tabasco, México. *Terra Latinoamericana* 23: 113-120.
- De la Peña, T. I. 2004. Disponibilidad hidráulica, superficial y subterránea y manejo y operación del distrito de riego del Río Yaqui. Publicación del Distrito de Riego del Río Yaqui. S de R.L. de I. P. y C. V. Cd. Obregón, Sonora, México.
- De la Peña, T. I. 2005. Manual de operación distribución y administración del uso eficiente del agua en el distrito de riego 041, Río Yaqui. Publicación del Distrito de Riego del Río Yaqui. S de R.L. de I. P. y C. V. Cd. Obregón, Sonora, México.
- Glantz, M. H. 1996. *Currents of change, El Niño's impact on climate and society*. Cambridge University Press. Cambridge, UK.
- INIFAP (Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias). 2001. Guía técnica para los cultivos del área de influencia del campo experimental del Valle del Yaqui. Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias-SAGARPA. Cd. Obregón, Sonora México.
- Ljung, G. G. and G. E. P. Box. 1978. On a measure of lack of fit in time series models. *Biometrika* 62: 297-303.
- Magaña, V. O., J. L. Vázquez, J. L. Pérez, and J. B. Pérez. 2003. Impact of El Niño on precipitation in Mexico. *Geofísica Int.* 42: 313-330.
- Minitab. 2003. *Meet Minitabs/Stat Release 14*. Minitab. State College, PA, USA.
- Mon, Y. and U. Lall. 1994. Kernel quantile function estimator for flood frequency analysis. *Water Resour. Res.* 30: 3095-3103.
- Pankratz, A. 1991. *Forecasting with dynamic regression models*. John Wiley. New York, NY, USA.
- Rasmusson, E. M. and T. H. Carpenter. 1983. The relationship between eastern equatorial Pacific sea surface temperatures and rainfall over India and Sri Lanka. *Monthly Weather Rev.* 111: 517-528.
- Ropelewsky, C. F. and P. D. Jones. 1987. An extension of the Tahiti-Darwin southern oscillation index. *Monthly Weather Rev.* 115: 2161-2165.
- SAS Institute. 1999. *SAS/STAT user's guide. Version 8*. SAS Institute. Cary, NC, USA.
- Silverman, B. W. 1986. *Density estimation for statistics and data analysis*. Chapman & Hall. London, UK.
- Wolter, K. and M. S. Timlin. 1998. Measuring the strength of ENSO - How does 1997/98 Rank? *Weather* 53: 315-324.
- Yue, S., T. B. M. J. Quarda, and B. Bobée. 2001. A review of bivariate gamma distributions for hydrological applications. *J. Hydrol.* 246: 1-18.