Nota técnica

Una aplicación hidrológica de la regresión lineal múltiple ponderada

• Daniel Francisco Campos-Aranda •
Profesor jubilado de la Universidad Autónoma de San Luis Potosí, México

Resumen

Campos-Aranda, D. F. (julio-agosto, 2016). Una aplicación hidrológica de la regresión lineal múltiple ponderada. *Tecnología y Ciencias del Agua*, *7*(4), 161-173.

La regresión lineal múltiple tiene dos aplicaciones hidrológicas básicas: (1) ampliar registros cortos con base en series largas cercanas, y (2) deducir ecuaciones empíricas que permiten estimar, en sitios de interés sin aforos, crecientes de diseño (Q_{Tr}). Como ambas aplicaciones se realizan en un *contexto regional*, siempre está presente la multicolinealidad en el primer caso, y la falta de homocedasticidad en el segundo. Para corregir la no uniformidad que tienen las varianzas de la variable dependiente (Y_i) se usa una función de ponderado (w_i) en el ajuste de mínimos cuadrados, lo cual conduce a la técnica de mínimos cuadrados ponderados (MCP). En este trabajo se exponen con detalle dos procedimientos para estimar las w_i óptimas. El primero toma en cuenta la teoría de los residuales y el error medio del ajuste de mínimos cuadrados ordinarios y el segundo se basa en los datos que son vecinos cercanos, para buscar la manera en que varían las varianzas de Y_i. Ambos métodos se aplican a las ecuaciones empíricas que permiten estimar el gasto máximo medio anual (*Qma*) de la Región Hidrológica 10 (Sinaloa, México). Con base en los resultados, se concluye que es recomendable aplicar siempre el método de MCP, al obtener ecuaciones empíricas que estiman el *Qma*, o bien las Q_{Tr}, pues sus indicadores de desempeño evaluados en el dominio real muestran mejoría de ajuste.

Palabras clave: homocedasticidad, mínimos cuadrados ordinarios, mínimos cuadrados ponderados, desviación absoluta media, error estándar medio, error relativo estándar medio, creciente media anual, Región Hidrológica 10.

Abstract

Campos-Aranda, D. F. (July-August, 2016). An Hydrological Application of Weighted Multiple Linear Regression. Water Technology and Sciences *(in Spanish), 7(4), 161-173.*

Multiple linear regression has two basic hydrological applications: (1) to extend short records based on long series that are close and (2) to derive empirical equations to estimate flood design (Q_{Tr}) at sites of interest where records are not available. Since both applications are made in a regional context, multicollinearity is always present in the first case, and the lack of homoscedasticity in the second. In order to correct the non-uniformity with the variances of the dependent variable (Y_i) a weighting function (w_i) in the least squares fit is used, which leads to the weighted least squares (WLS) technique. In this work two methods to estimate the optimal w_i are discussed in detail; the first one takes into account the theory of residuals and the mean error of setting ordinary least squares and the second one is based on data that are close neighbors, seeking for changes on the variances of Y_i. Both methods are applied to empirical equations that estimate the average annual maximum flow (Qaa) of Hydrological Region No. 10 (Sinaloa). Based on the results it is concluded that it is advisable to always apply the method of WLS, to obtain empirical equations that estimate the Qaa or the Q_{Tr}, due to the improvement of adjustment of the performance indicators evaluated in the real domain.

***Keywords**: Homoscedasticity, ordinary least squares, weighted least squares, mean absolute deviation, mean standard error, means standard relative error, average annual flood, Hydrological Region No. 10.*

Recibido: 20/01/2015
Aceptado: 16/02/2016

Introducción

La regresión lineal múltiple (RLM) es una técnica estadística básica empleada por los hidrólogos para transferir información de las características del escurrimiento, calculadas en cuencas con hidrometría hacia sitios o cuencas donde tal información es necesaria y no existen aforos. Por ejemplo, la RLM se emplea para completar series de volumen escurrido anual (Y) en sitios con pocos datos, con base en varios registros amplios y cercanos (Xp); también se ha utilizado para encontrar ecuaciones empíricas que relacionan la creciente media anual (Qma) o de cierto periodo de retorno (Q_{Tr}) con las características físicas de las cuencas con aforos, por lo común el tamaño de la cuenca y ciertas propiedades del cauce, o tormentas de la zona (Pandey & Nguyen, 1999; Griffis & Stedinger, 2007; Salas *et al*., 2008; Wilks, 2011).

Desde el inicio del siglo XXI se han comenzado a notar los efectos negativos del *cambio climático* a través de observar eventos extremos, como tormentas, crecientes y sequías más severos y más frecuentes. Ante tal evidencia se han sugerido varias acciones (Campos-Aranda, 2015), una de ellas, quizá la más importante, sea aumentar los sitios de registro de lluvias, gastos y niveles en lagos, embalses y acuíferos. Sin embargo, la realidad demuestra que el número de estaciones pluviográficas, pluviométricas e hidrométricas ha disminuido notablemente, hasta llegar a una situación crítica, como lo han destacado Lafragua-Contreras, González-Rojas y Solís-Alvarado (2006). En este escenario, es prioritario hacer un uso eficiente de la información hidrológica disponible y la RLM es la técnica estadística que lo permite, al trabajar los datos regionalmente tanto en el transporte de datos como en la estimación de ecuaciones empíricas.

Los parámetros de ajuste de un modelo definido por una RLM se estiman a través del método o técnica de mínimos cuadrados ordinarios (MCO), la cual acepta o asume que las varianzas de la variable dependiente (Y_i) son las mismas para toda i, lo cual se conoce como condición de *homocedasticidad*. Lo anterior implica que todas las observaciones de Y sean "igualmente confiables". En cualquier aplicación hidrológica *regional* de la RLM, tal condición seguramente será violada debido a que la confiabilidad en la estimación de las características del escurrimiento depende de la amplitud de su registro, o bien de las condiciones de medición en las estaciones de aforos. Cuando la hipótesis de homocedasticidad no es satisfecha, los parámetros estimados no tienen varianza mínima y todas las estimaciones asociadas con la RLM no son exactas (Tasker, 1980; Stedinger & Tasker, 1985; Tasker & Stedinger, 1989; Pandey & Nguyen, 1999; Kottegoda & Rosso, 2008).

Como las varianzas de la variable dependiente $\sigma^2(Y_i)$ no son iguales, se puede aplicar una función de ponderado (w_i) que corrija tal variación, según se expone en la teoría del método de mínimos cuadrados ponderados (MCP). Tal función es óptima cuando $w_i = 1/\sigma^2(Y_i)$.

El *objetivo* de este trabajo consiste en exponer con detalle dos procedimientos para encontrar y aplicar la función de ponderado óptima. El primero utiliza el algoritmo desarrollado por Tasker (1980), que toma en cuenta la teoría de los residuales y los resultados del método de MCO. El segundo se basa en una propiedad de los valores cercanos del regresor ($X1$), que son considerados puntos de repetición. Ambos procedimientos se aplican en la obtención de dos ecuaciones de regresión potencial, que permiten estimar el gasto máximo medio anual (Qma) de la Región Hidrológica 10 (Sinaloa, México). Como los resultados del método de MCP mejoran el ajuste de MCO, según indicadores evaluados en el dominio real, se recomienda su aplicación sistemática al estimar ecuaciones empíricas por RLM.

Resumen de la teoría operativa

Regresión lineal múltiple (RLM)

Algunas veces se puede establecer una relación de tipo lineal entre la variable dependiente (Y) y varias (p) independientes $X1$, $X2$..., Xp o

• ISSN 0187-8336

regresores, la cual es la generalización o extensión natural de la regresión lineal simple. Su expresión es (Ryan, 1998):

$$Y = \beta_0 + \beta_1 X1 + \beta_2 X2 + \beta_3 X3 + \cdots + \beta_p Xp + \varepsilon \quad (1)$$

Entonces, los principios que rigen la regresión lineal se aplican a la RLM, por ejemplo, que tanto *Y* como las *Xp* estén normalmente distribuidas, y que los errores ε sean independientes y tengan distribución normal de media cero y misma varianza (σ^2) para cada *X*. La *solución de mínimos cuadrados de los residuos* de forma matricial para el caso general expuesto y con *n* observaciones o datos de *Y*, y de los regresores es la siguiente (Ryan, 1998):

$$\mathbf{Y} = \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

siendo:

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & X1_1 & X2_1 & \cdots & Xp_1 \\ 1 & X1_2 & X2_2 & \cdots & Xp_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & X1_n & X2_n & \cdots & Xp_n \end{bmatrix},$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

El planteamiento de esta solución implica que la sumatoria de 1 a *n* de los residuos al cuadrado debe ser minimizada, es decir que:

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i \right)^2$$

$$= \sum_{i=1}^{n} \left(Y_i - \beta_0 - \beta_1 X1_i - \beta_2 X2_i - \cdots - \beta_p Xp_i \right)^2 = 0 \quad (3)$$

Entonces, diferenciando el lado derecho de la ecuación anterior con respecto a β_0, β_1, β_2,.. β_p por separado, se originan las ecuaciones llamadas *normales*, función de los parámetros desconocidos. En notación matricial, estas ecuaciones son:

$$\left(\mathbf{X}' \cdot \mathbf{X} \right) \cdot \hat{\boldsymbol{\beta}} = \mathbf{X}' \cdot \mathbf{Y} \quad (4)$$

cuya solución es:

$$\hat{\beta} = \left(\mathbf{X}' \cdot \mathbf{X} \right)^{-1} \cdot \left(\mathbf{X}' \cdot \mathbf{Y} \right) \quad (5)$$

en la cual $\mathbf{X}'$ es la matriz transpuesta de $\mathbf{X}$ y $(\mathbf{X}' \cdot \mathbf{X})^{-1}$ indica la matriz inversa de $\mathbf{X}' \cdot \mathbf{X}$.

Solución de mínimos cuadrados ponderados

Como ya se indicó, las suposiciones que por lo común se establecen en relación con RLM (ecuación (2)) son que $E(\varepsilon) = 0$ y que $\text{Var}(\varepsilon) = \sigma^2 \cdot \mathbf{I}$, siendo $\mathbf{I}$ la matriz unitaria o identidad. Con frecuencia, tales premisas son irrazonables, pues se tiene que $\text{Var}(\varepsilon) = \sigma^2 \cdot \mathbf{V}$, siendo $\mathbf{V}$ una matriz conocida de $n \times n$. Si $\mathbf{V}$ es diagonal, con elementos diagonales distintos, las observaciones $\mathbf{Y}$ no están correlacionadas, pero tienen varianzas desiguales; en cambio, si existen algunos elementos fuera de la diagonal principal de $\mathbf{V}$, las observaciones están correlacionadas y la solución de mínimos cuadrados de los residuos es (Montgomery, Peck, & Vining, 2002):

$$\hat{\beta}_g = \left(\mathbf{X}' \cdot \mathbf{V}^{-1} \cdot \mathbf{X} \right)^{-1} \cdot \left(\mathbf{X}' \cdot \mathbf{V}^{-1} \cdot \mathbf{Y} \right) \quad (6)$$

En la ecuación anterior, $\hat{\beta}_g$ es el estimador de mínimos cuadrados generalizados (MCG) de β. Cuando $\mathbf{V}$ es una matriz diagonal, con elementos $1/w_1$, $1/w_2$... $1/w_n$, se hace $\mathbf{W} = \mathbf{V}^{-1}$, siendo $\mathbf{W}$ una matriz también diagonal con elementos que son los pesos o factores de ponderación (w_1, w_2... w_n), la solución de mínimos cuadrados de los residuos será:

$$\hat{\beta}_p = \left(\mathbf{X}' \cdot \mathbf{W} \cdot \mathbf{X} \right)^{-1} \cdot \left(\mathbf{X}' \cdot \mathbf{W} \cdot \mathbf{Y} \right) \quad (7)$$

ISSN 0187-8336

en la cual $\hat{\beta}_p$ es el estimador de mínimos cuadrados ponderados (MCP) de β. Conviene notar que los datos con varianzas grandes tienen menos peso que los de varianzas pequeñas (Montgomery *et al*., 2002). Los estimados $\hat{\beta}_p$ se pueden obtener con facilidad modificando ligeramente el programa de cómputo de mínimos cuadrados ordinarios (MCO). Se multiplica cada uno de los datos o valores de la *i*-ésima observación, incluyendo el uno de la ordenada al origen por la raíz cuadrada del peso (w_i) correspondiente a tales datos. Así, se obtendrán las siguientes *matrices de datos transformados* (Montgomery *et al*., 2002):

$$\mathbf{B} = \begin{bmatrix} 1\sqrt{w_1} & X1_1\sqrt{w_1} & \cdots & Xp_1\sqrt{w_1} \\ 1\sqrt{w_2} & X1_2\sqrt{w_2} & \cdots & Xp_2\sqrt{w_2} \\ \vdots & \vdots & & \vdots \\ 1\sqrt{w_n} & X1_n\sqrt{w_n} & \cdots & Xp_n\sqrt{w_n} \end{bmatrix}$$

$$\mathbf{Z} = \begin{bmatrix} Y_1\sqrt{w_1} \\ Y_2\sqrt{w_2} \\ \vdots \\ Y_n\sqrt{w_n} \end{bmatrix}$$

Aplicando MCO a los datos transformados, se obtiene el estimador de mínimos cuadrados ponderados, que será:

$$\hat{\beta}_p = \left(\mathbf{B}'\mathbf{B}\right)^{-1}\mathbf{B}'\mathbf{Z} = \left(\mathbf{X}'\mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W}\mathbf{X} \qquad (8)$$

Para usar la técnica de mínimos cuadrados ponderados se deben conocer los pesos w_i. Con frecuencia se puede recurrir a la experiencia o conocimiento previo, a la información de un modelo teórico, o bien el análisis de los residuos puede indicar que la varianza de los errores puede ser una función de uno de los regresores; por ejemplo, si $\mathrm{Var}(\varepsilon) = \sigma^2 \cdot X1_i$, entonces $w_i = 1/X1_i$. Incluso, en aplicaciones prácticas, se pueden suponer los pesos y hacer iteraciones para mejorar la regresión y/o minimizar algunos residuos (Montgomery *et al*., 2002).

Indicadores de calidad del ajuste

Cuando se obtienen ecuaciones de RLM con mínimos cuadrados ponderados es necesario compararlas para escoger la de mejor ajuste, lo cual implica evaluar indicadores basados en los residuos. El más común de tales indicadores es el coeficiente de determinación (R^2), que indica la proporción de la varianza de la variable dependiente que es explicada por la ecuación de regresión. Por ello su expresión es (Ryan, 1998):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2} \qquad (9)$$

en la cual $\hat{Y}_i$ es el valor estimado de la variable dependiente con la ecuación de regresión y $\overline{Y}$ es su media aritmética observada. Cuando se usan los logaritmos de las variables, el R^2 no es confiable, pues los residuos se evalúan en el dominio logarítmico, como han señalado McCuen, Leahy y Johnson (1990). Por lo anterior, las ecuaciones de RLM que se analizan o contrastan se evaluarán por medio de los tres índices de desempeño (ID) siguientes, calculados en el dominio real (Pandey & Nguyen, 1999). *DAM* es la *desviación absoluta media* con las mismas unidades que Y_i:

$$DAM = \frac{1}{n - npa}\sum_{i=1}^{n}\left|Y_i - \hat{Y}\right| \qquad (10)$$

siendo *npa* el número de parámetro de ajuste de la ecuación de RLM. *EEM* es el *error estándar medio* también con las unidades de Y_i:

$$EEM = \left[\frac{1}{n - npa}\sum_{i=1}^{n}\left(Y_i - \hat{Y}\right)^2\right]^{1/2} \qquad (11)$$

Por último, *EREM* es el *error relativo estándar medio*, que es adimensional:

$$EREM = \left[\frac{1}{n - npa}\sum_{i=1}^{n}\frac{\left(Y_i - \hat{Y}\right)^2}{Y_i^2}\right]^{1/2} \qquad (12)$$

• ISSN 0187-8336

Aplicación hidrológica

Antecedentes

Clarke (1994) presenta los datos de 23 estaciones hidrométricas ubicadas dentro del sistema del río Itajaí-Acú en Brasil, relativos al gasto máximo medio anual (*Qma*, m^3/s), áreas de cuenca (*A*, km^2) y número de años de registro (*NA*). Los valores de *Qma* variaron de 31 a 3194 m^3/s; los de *A* de 105 a 11 719 km^2, y los de *NA* de 3 a 118 años. Realiza una regresión lineal del tipo $Qma = b_0 + b_1 \cdot A$, cuyo coeficiente de determinación (R^2) resultó de 0.828. Para mejorar la estimación, se emplea la regresión potencial $Qma = b_0 \cdot A^{b_1}$, cuyo R^2 resultó de 0.909. El análisis de los residuos de esta última ecuación indica que los mayores errores corresponden a los datos que tienen las menores amplitudes de registro, por ello aplica la técnica de MCP utilizando $w_i = NA$, obteniendo un R^2 de 0.935 y observando que los residuos disminuyeron. La función de ponderado aplicada por Clarke (1994) es en realidad una versión bastante simplificada de la función óptima, como se deduce a continuación en su primer procedimiento de búsqueda. Este enfoque simple para la aplicación del ajuste de MCP también ha sido aplicado por Vogel, Wilson y Daly (1999).

Primera aproximación

Procedimiento que divide la varianza residual

Tasker (1980) desarrolló un procedimiento para estimar la función de ponderado óptima $w_i = 1/\sigma^2(Y_i)$, a utilizar en el ajuste por MCP. Partió de los resultados de Matalas y Gilroy (1968), que establecieron que la varianza de la variable dependiente (Y_i) se puede dividir en dos componentes: la primera originada por el error del modelo $\sigma^2(\delta_i)$ y la segunda debida al error de muestreo $\sigma^2(e_i)$; es decir:

$$\sigma^2\left(Y_i\right) = \sigma^2\left(\delta_i\right) + \sigma^2\left(e_i\right) \quad \text{para toda } i \qquad (13)$$

El subíndice *i* varía de uno al número de datos o valores de Y_i, y también es igual al número de estaciones hidrométricas (*NE*) o registros procesados en los análisis regionales. La condición de homocedasticidad requiere que tanto $\sigma^2(\delta_i)$ como $\sigma^2(e_i)$ sean independientes de *i*. En el procedimiento desarrollado por Tasker (1980), sólo se acepta la independencia del error del modelo. Para estimar la varianza del error de muestreo, se considera que las crecientes asociadas con un cierto periodo de retorno (*Tr*) siguen una distribución Pearson tipo III y entonces de acuerdo con Bobée (1973) se tiene:

$$\sigma^2(e_i) \cong \frac{\sigma^2}{n_i}\left[1 + K_p \cdot \gamma + \frac{K_p^2}{2}\left(1 + \frac{3}{4}\gamma^2\right)\right] \qquad (14)$$

en la cual σ es la desviación estándar de los gastos máximos anuales; *n*, el número de gastos anuales observados en la estación hidrométrica *i*; γ, el coeficiente de asimetría de los gastos máximos anuales, y *Kp* es la desviación estandarizada con distribución Pearson tipo III asociada con el valor de γ y de la probabilidad de no excedencia *p*. Algunas veces K_p se designa por K_{Tr}, pues $Tr = 1/(1 - p)$. En estudios hidrológicos regionales se puede aceptar que γ y σ^2 son aproximadamente constantes en todos los sitios de tal zona o región debido precisamente a la homogeneidad regional verificada previamente (Hosking & Wallis, 1997) y entonces la varianza de Y_i (ecuación 13) se expresa como (Tasker, 1980):

$$\sigma^2\left(Y_i\right) = c_o + c_1\left(1/n_i\right) \qquad (15)$$

siendo $c_o = \sigma^2(\delta_i)$ una constante y c_1 otra, que se logra estimar con base en la información regional disponible, según la expresión:

$$\hat{c}_1 = \hat{\sigma}^2\left[1 + \hat{K}_p \cdot \hat{\gamma} + \left(K_p^2/2\right)\cdot\left(1 + 3\hat{\gamma}^2/4\right)\right] \qquad (16)$$

en la cual $\hat{\sigma}$, $\hat{\gamma}$ y $\hat{K}_p$ son estimaciones regionales de las mismas variables descritas en la ecuación (14). La aproximación de la constante c_o será, según la ecuación (15):

$$\hat{c}_o = \hat{\sigma}^2\left(Y_i\right) - \hat{c}_1\left(1/\bar{n}\right) \qquad (17)$$

ISSN 0187-8336

en donde $\hat{\sigma}^2(Y_i)$ es el cuadrado medio residual, cuya raíz cuadrada es el error estándar de la regresión obtenido usando MCO y $\overline{n}$ es la media aritmética de las amplitudes de registro en las estaciones hidrométricas que integran la región; $\hat{\sigma}^2(Y_i)$ equivale al numerador de la ecuación (9) divido entre *NE - np*, siendo *np* el número de parámetros que se estiman en la regresión. Finalmente, la estimación de la función de ponderado ($\hat{w}_i$) para el ajuste de MCP será:

$$\hat{w}_i = \frac{1}{\hat{c}_o + \hat{c}_1\left(1/n_i\right)} \tag{18}$$

siendo n_i el número de datos de cada registro procesado. Tasker (1980) indica que cuando el error debido al modelo es grande, $\hat{c}_o$ tiende a ser mayor que $\hat{c}_1$ y $\hat{w}_i$ se aproxima a un valor constante para toda *i*, lo cual es el caso de ajuste de MCO. Para errores de modelo pequeños ocurre lo contrario y entonces $\hat{w}_i$ tiende a una relación directa con n_i, que es la función de ponderado apropiada cuando no existe error debido al modelo. Tasker y Stedinger (1986) exponen otra aplicación del procedimiento anterior para estimar el coeficiente de asimetría regional de la distribución Log-Pearson tipo III.

Cuando se disponga de los datos anuales de cada variable dependiente (Y_i), el procedimiento anterior puede ser mejorado, calculando la magnitud de la constante $\hat{c}_1$ (ecuación (16)), con base en tales valores, de manera que el error de muestreo se estime para cada registro procesado. También se puede aplicar el equivalente de la ecuación (14) para adoptar otro modelo probabilístico, como los descritos en Kite (1977), Rao y Hamed (2000), o Asquith (2011).

Aspectos operativos previos

Para aplicar el procedimiento de Tasker (1980), descrito en las ecuaciones (13) a (18), primero se deben definir las expresiones de las estimaciones regionales $\hat{\sigma}$, $\hat{\gamma}$ y $\hat{K}_p$ de la ecuación (16). De inicio se aclara que en la aplicación hidrológica que se describirá, relativa al gasto máximo medio anual (*Qma*), se tiene que el *Tr* es de dos años y por tanto la probabilidad de excedencia es de 0.50 en tal evento; además, *NE* = 22 (ver cuadro 1). Como ya se indicó, *n* varía de 21 a 56 datos, con $\overline{n}$ = 37. Para los valores de *Qma*, las expresiones de $\hat{\sigma}$ y $\hat{\gamma}$ son:

$$\hat{\sigma} = \left[\frac{\sum_{i=1}^{NE}\left(Qma_i - \overline{Qma}\right)^2}{NE-1}\right]^{1/2} \tag{19}$$

$$\hat{\gamma} = \frac{NE\sum_{i=1}^{NE}\left(Qma_i - \overline{Qma}\right)^3}{\left(NE-1\right)\cdot\left(NE-2\right)\cdot\hat{\sigma}^3} \tag{20}$$

siendo:

$$\overline{Qma} = \frac{\sum_{i=1}^{NE} Qma_i}{NE} \tag{21}$$

Para la estimación del valor de $\hat{K}_p$ se usará la aproximación polinomial obtenida por Bobée y Ashkar (1991) para los valores tabulados por Harter (1969), cuya expresión es:

$$K_p = b_0 + b_1\cdot\gamma + b_2\cdot\gamma^2 + b_3\cdot\gamma^3 + b_4\cdot\gamma^4 \tag{22}$$

La ecuación anterior está limitada a $|\gamma| \le 4$; algunos de los valores de los coeficientes *b* se tienen en el cuadro 1. Cuando el coeficiente de asimetría γ es negativo, se aplica la siguiente igualdad: $K_p(\gamma) = -K_{1-p}(-\gamma)$. Por ejemplo, para p = 99% y γ = -1.5, se calcula K_p para p = 1% y γ = 1.5, que es -1.25611; por lo tanto, el valor buscado es 1.25611.

Análisis de resultados

Los datos que serán procesados se muestran en el cuadro 2 en sus primeras seis columnas; proceden de Campos-Aranda (2013), y corresponden a los valores del gasto máximo medio anual (*Qma*) o creciente media anual, así como de varias propiedades fisiográficas de 22 cuencas de las estaciones hidrométricas

• ISSN 0187-8336

Cuadro 1. Coeficientes b_i de la ecuación (22), función de periodo de retorno (*Tr*) en años.

***Tr* (años)**	b_0	b_1	b_2	b_3	b_4
2	-0.983388E-03	-0.158259E+00	-0.137341E-01	+0.949401E-02	-0.646736E-03
5	+0.841362E+00	-0.469111E-01	-0.383029E-01	+0.863879E-03	+0.599519E-03
10	+0.128142E+01	+0.107660E+00	-0.474931E-01	-0.182259E-02	+0.646707E-03
20	+0.164473E+01	+0.284920E+00	-0.493567E-01	-0.420491E-02	+0.761456E-03
50	+0.205352E+01	+0.538133E+00	-0.427890E-01	-0.780118E-02	+0.985350E-03
100	+0.232598E+01	+0.738681E+00	-0.323114E-01	-0.110577E-01	+0.121961E-02

de la Región Hidrológica 10 (Sinaloa), que no presentan régimen hidrológico modificado, las cuales fueron tomadas de Escalante-Sandoval y Reyes-Chávez (2002). La regresión del tipo $Qma = b_0 \cdot A^{b_1}$ se obtuvo con el método de MCO, b_0 = 11.6751 y b_1 = 0.5258, y las estimaciones y residuales mostrados en las columnas 2 y 3 del cuadro 3. Este método condujo a un valor de R^2 de 0.813, con el resto de indicadores de ajuste mostrados al final de la citada columna 3 y un error estándar de ajuste de 309.8 m^3/s.

Para los datos de la columna 3 del cuadro 2 se obtiene que su coeficiente de asimetría es 2.0303, lo cual conduce a un valor de K_p = –0.3104. Con base en estos valores y los citados anteriormente, se obtuvo que c_0 = 88 468.7 y c_1 = 27 7787.1, ambos con unidades de varianza (m^6/s^2). Los correspondientes factores de ponderación (ecuación (18)) se muestran en la columna 7 del cuadro 2.

El método de MCP aporta b_0 = 11.7339 y b_1 = 0.5251, con las estimaciones de *Qma*, los residuos y sus indicadores de desempeño (ID) que se tienen en las columnas 4 y 5 del cuadro 3. Se observa en la porción final de la columna 5 que los ID del método de MCP son casi iguales a los del ajuste de MCO. Este resultado se considera congruente, pues en esta aplicación numérica, las amplitudes de los registros procesados son semejantes (ver cuadro 2), variando de 21 a 56, con una media de 37 años.

También se revisó la regresión potencial del tipo $Qma = b_0 \cdot A^{b_1} \cdot L_{cp}^{b_2}$, cuyos resultados del ajuste de MCO fueron (Campos-Aranda, 2013): b_0 = 11.7598, b_1 = 0.5334 y b_2 = –0.0148, con las estimaciones y residuos expuestos en las columnas 8 y 9 del cuadro 3; así como los ID mostrados al final de la columna 9. Empleando los resultados citados, el error estándar de ajuste resulta de 312.5 m^3/s, pues ahora np = 3; con tal valor y los ya citados se obtuvieron c_0 = 94113.8 y c_1 = 131 764.9, con los factores de ponderación (ecuación (18)), que se exponen en la columna 8 del cuadro 2. El método de MCP conduce a b_0 = 11.7911, b_1 = 0.5334 y b_2 = -0.0153, con las estimaciones de *Qma*, los residuos y sus ID que se tienen en las columnas 10 y 11 del cuadro 3. Se observa al final de la columna 11 que dos ID del método de MCP aumentan con respecto a los del ajuste de MCO y sólo disminuye el error estándar medio.

Segunda aproximación

Procedimiento basado en datos cercanos

Sugerido por Draper y Smith (1998), y por Montgomery *et al*. (2002), comienza por definir conjuntos de valores del regresor *X*1 que son "vecinos cercanos", por tener observaciones con magnitudes semejantes de *X*1. El procedimiento supone que tales conjuntos pueden considerarse "puntos de repetición" y por lo tanto se puede usar la varianza promedio de sus respuestas (Y_i) para estimar la forma en que, de manera aproximada, cambia Var(*Y*) en función de *X*1.

En la columna 9 del cuadro 2 se indica el renglón donde comienza cada conjunto de datos repetidos, el número de elementos que incluye entre paréntesis y el valor promedio de *X*1, es decir, del área de cuenca (*A*). En la columna 10 se indica la varianza muestral (promedio, aproximadamente) de las Y_i de cada conjunto, estimada con la expresión siguiente:

ISSN 0187-8336

$$\mathrm{Var}(Qma) = \frac{\sum_{i=1}^{n}\left(Qma_i - \overline{Qma}\right)^2}{n-1} \qquad (23)$$

Con base en las seis parejas de valores calculados de $\overline{A}$ y de Var(Qma) se realiza una regresión polinomial casi perfecta ($R^2 > 0.9999$) del tipo:

$$\mathrm{Var}(Qma) = a_0 + a_1 \cdot A + a_2 \cdot A^2 + a_3 \cdot A^3 + a_4 \cdot A^4 \quad (24)$$

con a_0 = -2025.336, a_1 = 26.55365, a_2 = -1.122622E-02, a_3 = 1.481098E-06, a_4 = -4.118003E-11 y error estándar de la estimación de 9 315 m^3/s. Con base en tal ecuación se obtuvieron los factores de ponderación (w_i) mostrados en la columna 11 del cuadro 2.

Cuando se tienen dos o más regresores, resulta muy difícil la identificación visual de los datos que son vecinos cercanos y por ello se debe aplicar una técnica analítica para buscar pares de puntos cercanos entre sí en el espacio de *Xp* (Montgomery *et al*., 2002), o bien aplicar la ecuación (8), con la función de ponderado estimada con el primer regresor (*X*1), que es el más importante.

Análisis de resultados

Al aplicar la técnica de MCP, con la función de ponderado mostrada en la columna final del cuadro 2 y la regresión del tipo $Qma = b_0 \cdot A^{b_1}$, se obtienen b_0 = 7.7896 y b_1 = 0.5784, con una ligera mejoría en el ajuste, pues ahora se tiene un R^2 de 0.824 y los valores del error disminuyen (ver *EEM* y *EREM*) al final de la columna 7 del cuadro 3. Las estimaciones de *Qma* y sus residuos se tienen en las columnas 6 y 7 del citado cuadro 3. Este ajuste reduce notablemente los residuos positivos de las estaciones hidrométricas Huites y Guatenipa II, lo cual se ve reflejado en el valor del *EEM*. Lo anterior se puede observar al comparar las figuras 1 y 2.

Al aplicar la técnica de MCP a la segunda regresión potencial, se obtuvieron estos resultados: b_0 = 8.0767, b_1 = 0.6563 y b_2 = -0.1422. Las estimaciones y los residuos de este método se tienen en las dos columnas finales del cuadro 3. Los ID relacionados con el error muestran una mejoría de ajuste, como se observa al comparar los tres últimos renglones de las columnas 9 y 13. Al igual que la ecuación potencial anterior, los residuos de las estaciones Huites y Guatenipa II se reducen de modo sustancial y ello se aprecia en los ID del error.

Conclusiones

En los experimentos numéricos realizados por Tasker (1980) variaron: (1) el error del modelo de 0 a 100%; (2) la correlación entre estaciones tomó valores de 0.0, 0.4 y 0.8, y (3) la amplitud de los registros fluctuó de 10 a 50 años, con tres formas aleatorias de variación. Concluye, con base en la simulación numérica, que cuando se aplica la función de ponderado (w_i) definida por la ecuación (18), en el método de mínimos cuadrados ponderados (MCP), la ecuación de regresión resultante siempre tiene mejores indicadores de desempeño (ecuaciones (9) a (12)), que la obtenida con mínimos cuadrados ordinarios (MCO); excepto cuando n_i no varía y/o existe correlación cruzada importante entre los eventos anuales de las variables dependientes. En este último caso, habrá que aplicar la técnica de mínimos cuadrados generalizados (Griffis & Stedinger, 2007).

Con base en la aplicación hidrológica descrita, se pudo verificar que siempre alguno o varios de los indicadores de desempeño (ID) mostraron un mejor ajuste, es decir, se redujeron, al aplicar la técnica de MCP, en comparación con los ID obtenidos por MCO. Para el caso mostrado, en general los errores residuales se reducen más (se obtienen valores menores de los ID) con la segunda función de ponderado, la cual se obtiene con base en los datos cercanos.

Por lo anterior, se recomienda revisar las ecuaciones potenciales obtenidas para estimar *Qma*, o bien Q_{Tr}, a través del ajuste de MCP, empleando al menos la primera función de ponderado descrita, pues la segunda requiere la ocurrencia de datos cercanos. Incluso en

• ISSN 0187-8336

Cuadro 2. Datos de las 22 estaciones hidrométricas procesadas de la Región Hidrológica 10 (Sinaloa) y estimación de las funciones de ponderado (w_i) según procedimiento indicado.

1	2	3	4	5	6	7	8	9	10	11
Núm.	**Nombre**	**$Qma = Y_i$ (m^3/s)**	***NA*** **(años)**	***A*** **(km^2)**	**L_{cp} (km)**	**$\hat{w}_i = 1/\sigma^2(Y_i)$ (var. residual dividida)**	**$\hat{w}_i = 1/\sigma^2(Y_i)$ (var. residual dividida)**	**(Núm. de datos) $\overline{A}$ en km^2**	**Var(*Qma*)**	**$w_i = 1/\text{Var}(Y_i)$ (vecinos cercanos)**
1	Huites	3 328.3	51	26 057	267	1.064788E-05	1.034153E-05	(2) 21 794.0	1 285 927.0	3.482385E-06
2	San Francisco	1 724.6	33	17 531	246	1.032136E-05	1.019298E-05			9.061188E-07
3	Santa Cruz	1 037.6	52	8 919	180	1.065976E-05	1.034685E-05	(3) 8 450.0	105 239.5	7.574589E-06
4	Guatenipa II	1 590.9	32	8 252	178	1.029341E-05	1.018004E-05			1.064360E-05
5	Jaina	1 020.8	56	8 179	177	1.070330E-05	1.036626E-05			1.107969E-05
6	Palo Dulce	1 129.2	21	6 439	184	9.833169E-06	9.961316E-06	(3) 6 251.3	17 205.0	3.556125E-05
7	Ixpalino	1 199.0	45	6 166	174	1.056617E-05	1.030482E-05			4.429667E-05
8	La Huerta	945.1	28	6 149	117	1.016367E-05	1.011944E-05			4.491600E-05
9	Toahayana	1 078.7	27	5 281	131	1.012586E-05	1.010162E-05	(2) 5 189.5	19 129.7	8.908066E-05
10	Chinipas	883.1	24	5 098	131	9.995691E-06	1.003975E-05			9.997204E-05
11	Tamazula	596.9	32	2 241	58	1.029341E-05	1.018004E-05	(3) 2 063.0	13 425.7	5.976290E-05
12	Naranjo	633.3	45	2 064	107	1.056617E-05	1.030482E-05			5.803047E-05
13	Acatitán	813.3	43	1 884	115	1.053421E-05	1.029038E-05			5.701151E-05
14	Guamúchil	702.3	32	1 645	65	1.029341E-05	1.018004E-05			5.692032E-05
15	Choix	349.0	38	1 403	82	1.044072E-05	1.024786E-05			5.860841E-05
16	Badiraguato	1 224.3	26	1 018	59	1.008545E-05	1.008250E-05			6.715656E-05
17	El Quelite	479.1	33	835	42	1.032136E-05	1.019298E-05			7.597623E-05
18	Zopilote	351.9	56	666	68	1.070330E-05	1.036626E-05			9.001415E-05
19	El Bledal	289.0	56	371	31	1.070330E-05	1.036626E-05			1.573383E-04
20	Pericos	250.8	30	270	31	1.023246E-05	1.015167E-05	(3) 249.0	5 357.0	2.296374E-04
21	La Tina	105.0	24	254	30	9.995691E-06	1.003975E-05			2.488107E-04
22	Bamícori	189.2	33	223	20	1.032136E-05	1.019298E-05			2.981353E-04

ISSN 0187-8336

• *Tecnología y Ciencias del Agua*, vol. VII, núm. 4, julio-agosto de 2016, pp. 161-173

• ISSN 0187-8336

Cuadro 3. Gasto máximo medio anual (*Qma*) estimado con la ecuación de regresión indicada y sus residuos, en las 22 estaciones hidrométricas procesadas de la Región Hidrológica 10 (Sinaloa).

1			2	3	4	5	6	7	8	9	10	11	12	13
Datos observados e indicadores de desempeño (ID)			**Ajuste de MCO[1]**		**Ajuste de MCP[2]**		**Ajuste de MCP[3]**		**Ajuste de MCO[1]**		**Ajuste de MCP[2]**		**Ajuste de MCP[3]**	
			$Qma = b_0 \cdot A^{b_1}$		$Qma = b_0 \cdot A^{b_1}$		$Qma = b_0 \cdot A^{b_1}$		$Qma = b_0 \cdot A^{b_1} \cdot L_{cp}^{b_2}$		$Qma = b_0 \cdot A^{b_1}$		$Qma = b_0 \cdot A^{b_1} \cdot L_{cp}^{b_2}$	
Núm.	**Estación hidrométrica**	***Qma*** **(m³/s)**	$Q_{estimado}$ **(m³/s)**	**Residuo (m³/s)**	$Q_{estimado}$ **(m³/s)**	**Residuo (m³/s)**	$Q_{estimado}$ **(m³/s)**	**Residuo (m³/s)**	$Q_{estimado}$ **(m³/s)**	**Residuo (m³/s)**	$Q_{estimado}$ **(m³/s)**	**Residuo (m³/s)**	$Q_{estimado}$ **(m³/s)**	**Residuo (m³/s)**
1	Huites	3 328.3	2 449.1	879.2	2 445.5	882.8	2 789.9	538.4	2 455.9	872.4	2 602.6	725.7	2 886.5	441.8
2	San Francisco	1 724.6	1 988.4	-263.8	1 986.0	-261.4	2 218.4	-493.8	1 990.3	-265.7	2 107.2	-382.6	2 251.5	-526.9
3	Santa Cruz	1 037.6	1 393.8	-356.2	1 392.7	-355.1	1 500.7	-463.1	1 394.3	-356.7	1 470.8	-433.2	1 510.6	-473.0
4	Guatenipa II	1 590.9	1 338.0	252.9	1 337.0	253.9	1 434.7	156.2	1 337.9	253.0	1 411.1	179.8	1 437.7	153.2
5	Jaina	1 020.8	1 331.7	-310.9	1 330.8	-310.0	1 427.4	-406.6	1 331.7	-310.9	1 404.5	-383.7	1 430.5	-409.7
6	Palo Dulce	1 129.2	1 174.4	-45.2	1 173.7	-44.5	1 242.9	-113.7	1 171.5	-42.3	1 236.1	-106.9	1 216.0	-86.8
7	Ixpalino	1 199.0	1 147.9	51.1	1 147.3	51.7	1 212.2	-13.2	1 145.7	53.3	1 208.1	-9.1	1 191.3	7.7
8	La Huerta	945.1	1 146.3	-201.2	1 145.7	-200.6	1 210.2	-265.1	1 150.7	-205.6	1 207.8	-262.7	1 258.2	-313.1
9	Toahayana	1 078.7	1 058.1	20.6	1 057.7	21.0	1 108.3	-29.6	1 059.2	19.5	1 113.2	-34.5	1 120.5	-41.8
10	Chinipas	883.1	1 038.7	-155.6	1 038.3	-155.2	1 085.9	-202.8	1 039.5	-156.4	1 092.5	-209.4	1 094.8	-211.7
11	Tamazula	596.9	674.2	-77.3	674.3	-77.4	675.0	-78.1	678.6	-81.7	706.7	-109.8	716.8	-119.9
12	Naranjo	633.3	645.7	-12.4	645.8	-12.5	643.7	-10.4	643.6	-10.3	674.9	-41.6	622.5	10.8
13	Acatitán	813.3	615.4	197.9	615.6	197.7	610.6	202.7	612.4	200.9	642.7	170.6	580.3	233.0
14	Guamúchil	702.3	573.1	129.2	573.2	129.1	564.5	137.8	574.5	127.8	599.0	103.3	575.7	126.6
15	Choix	349.0	527.1	-178.1	527.3	-178.3	514.9	-165.9	525.9	-176.9	549.8	-200.8	501.8	-152.8
16	Badiraguato	1 224.3	445.3	779.0	445.6	778.7	427.7	796.6	445.4	778.9	463.9	760.4	426.0	798.3
17	El Quelite	479.1	401.2	77.9	401.5	77.6	381.4	97.7	402.7	76.4	417.9	61.2	392.6	86.5
18	Zopilote	351.9	356.2	-4.3	356.6	-4.7	334.6	17.3	354.4	-2.5	369.8	-17.9	316.0	35.9
19	El Bledal	289.0	261.9	27.1	262.2	26.8	238.6	50.4	262.4	26.6	271.5	17.5	240.7	48.3
20	Pericos	250.8	221.6	29.2	221.9	28.9	198.5	52.3	221.5	29.3	229.2	21.6	195.4	55.4
21	La Tina	105.0	214.6	-109.6	214.9	-109.9	191.6	-86.6	214.5	-109.5	221.9	-116.9	188.6	-83.6
22	Bamícori	189.2	200.4	-11.2	200.7	-11.5	177.7	11.5	201.3	-12.1	207.4	-18.2	183.4	5.8|
DAM (m³/s)			-	208.5	-	208.3	-	219.5	-	219.4	-	229.9	-	232.8
EEM (m³/s)			-	309.8	-	310.1	-	301.2	-	317.4	-	312.6	-	308.4
EREM			-	0.3412	-	0.3414	-	0.3307	-	0.3496	-	0.3747	-	0.3395

Simbología:
1 mínimos cuadrados ordinarios.
2 mínimos cuadrados ponderados, con función de ponderado con procedimiento que divide la varianza residual.
3 mínimos cuadrados ponderados, con función de ponderado con procedimiento basado en datos cercanos.

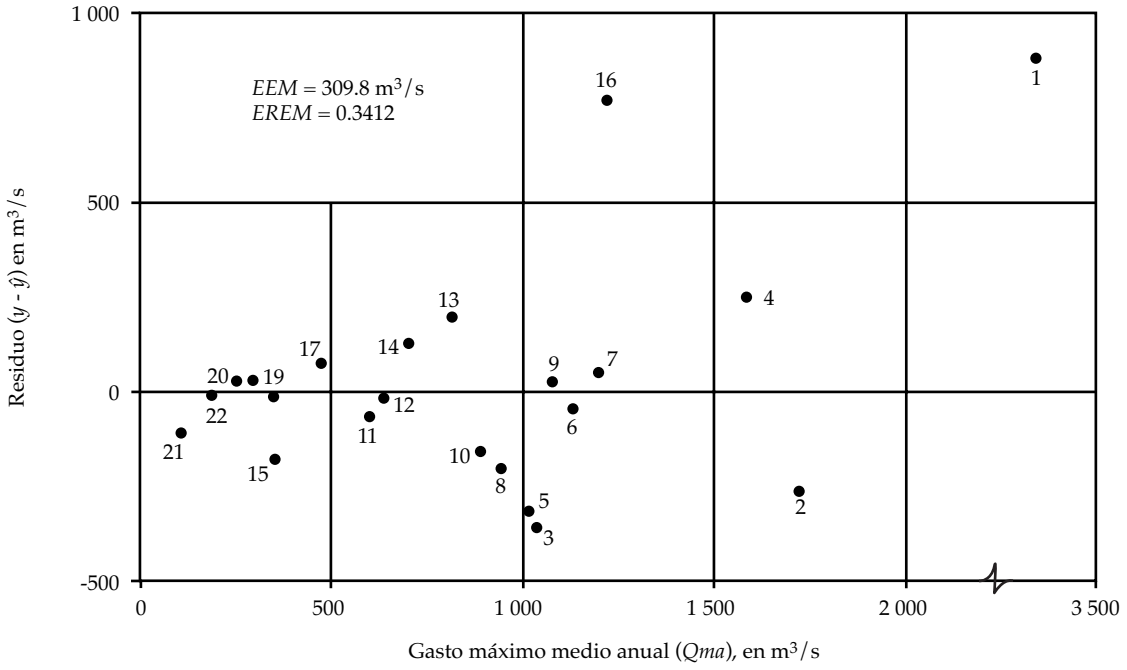

Figura 1. Gráfica de residuales para el ajuste de MCO con la ecuación potencial $Qma = b_0 A^{b_1}$.

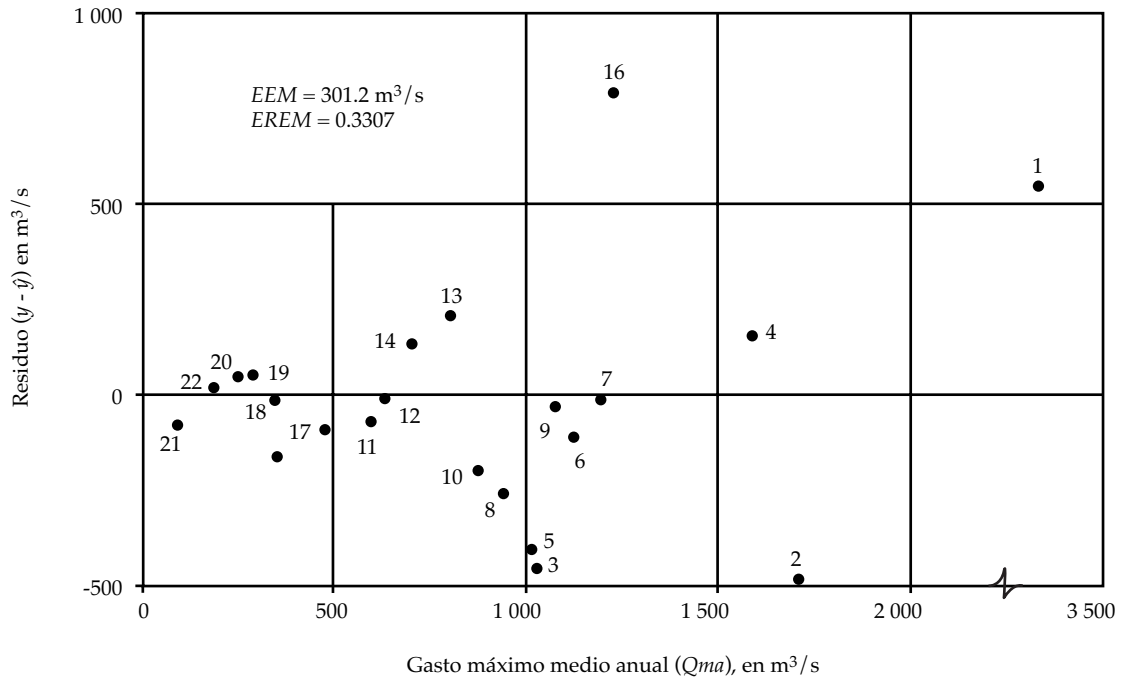

Figura 2. Gráfica de residuales para el ajuste de MCP con la ecuación potencial $Qma = b_0 A^{b_1}$ y la función de ponderado basada en datos cercanos.

ISSN 0187-8336

su versión simplificada, como la utilizó Clarke (1994) y Vogel *et al*. (1999), es decir, con $w_i = NA$, podrá aportar una mejoría estadística, esto es, reducir los ID o cuando menos verificar la similitud numérica de resultados del ajuste.

Agradecimiento

Se agradecen las observaciones y sugerencias de árbitro anónimo, las cuales permitieron volver explícito el texto y ayudaron a mostrar mejor sus alcances prácticos.

Referencias

Asquith, W. H. (2011). Chapter 8: L-Moments of Three Parameter Univariate Distributions (pp. 169-201). In: *Distributional Analysis with L-moment Statistics using the R Environment for Statistical Computing*. Texas, USA: Author edition.

Bobée, B. (1973). Sample Error of T-Year Events Computed by Fitting a Pearson Type 3 Distribution. *Water Resources Research*, *9*(5), 1264-1270.

Bobée, B., & Ashkar, F. (1991). Chapter 4: The One Parameter Gamma, G1(λ) and Standardized Pearson Type 3, SP(λ) Distributions (pp. 27-39). In: *The Gamma Family and Derived Distributions Applied in Hydrology*. Littleton, USA: Water Resources Publications.

Campos-Aranda, D. F. (2013). Estimación de la creciente media anual en la Región Hidrológica No. 10 de México con base en ecuaciones de regresión potencial. *Ingeniería. Investigación y Tecnología*, *14*(4), 563-574.

Campos-Aranda, D. F. (2015). Aplicación de la distribución de probabilidades no acotada del Sistema Johnson para estimación de crecientes. *Ingeniería. Investigación y Tecnología, 16*(4), 527-537.

Clarke, R. T. (1994). Chapter 4: Linear Relationships with explanatory variables (pp. 86-163). In: *Statistical Modelling in Hydrology*. Chichester, England: John Wiley & Sons, Ltd.

Draper, N. R., & Smith, H. (1998). Chapter 9: Multiple Regression: Special Topics (pp. 217-234). In: *Applied Regression Analysis*. Third edition. New York: John Wiley & Sons, Inc.

Escalante-Sandoval, C., & Reyes-Chávez, L. (2002). Apéndice A (pp. 291-298). En: *Técnicas estadísticas en hidrología.* México, DF: Facultad de Ingeniería de la UNAM.

Griffis, V. W., & Stedinger, J. R. (2007). The Use of GLS Regression in Regional Hydrologic Analyses. *Journal of Hydrology*, *344*(1-2), 82-95.

Harter, H. L. (1969). A New Table of Percentage Points of the Pearson Type III Distribution. *Technometrics*, *11*(1), 177-187.

Hosking, J. R. M., & Wallis, J. R. (1997). Chapter 4: Identification of Homogeneous Regions (pp. 54-72). In: *Regional Frequency Analysis. An Approach Based on L-Moments*. Cambridge: Cambridge University Press.

Kite, G. W. (1977). *Frequency and Risk Analyses in Hydrology* (224 pp.). Fort Collins, USA: Water Resources Publications.

Kottegoda, N. T., & Rosso, R. (2008). Chapter 6: Methods of Regression and Multivariate Analysis (pp. 326-404). *Applied Statistics for Civil and Environmental Engineers*. Second edition. Oxford, United Kingdom: Blackwell Publishing Ltd.

Lafragua-Contreras, J., González-Rojas, D., & Solís-Alvarado, Y. (2006). *Cantidad y calidad de la información climatológica e hidrométrica para el cálculo de la disponibilidad de agua superficial*. Tema: Hidrología, Ponencia 7. XIX Congreso Nacional de Hidráulica, del 8 al 10 de noviembre, Cuernavaca, Morelos, México.

Matalas, N. C., & Gilroy, E. J. (1968). Some Comments on Regionalization in Hydrologic Studies. *Water Resources Research, 4*(6), 1361-1369.

McCuen, R. H., Leahy, R. B., & Johnson, P. A. (1990). Problems with Logarithmic Transformations in Regression. *Journal of Hydraulic Engineering*, *116*(3), 414-428.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2002). Capítulo 5: Transformaciones y ponderación para corregir inadecuaciones del modelo (pp. 155-184). En: *Introducción al análisis de regresión lineal*. México, DF: Compañía Editorial Continental.

Pandey, G. R., & Nguyen, V. T. V. (1999). A Comparative Study of Regression Based Methods in Regional Flood Frequency Analysis. *Journal of Hydrology*, *225*(1-2), 92-101.

Rao, A. R., & Hamed, K. H. (2000). *Flood Frequency Analysis*. Boca Raton, USA: CRC Press.

Ryan, T. P. (1998). Chapter 14: Linear Regression (pp. 14.1-14.43). In: *Handbook of Statistical Methods for Engineers and Scientists*. Wadsworth, H. W. (ed.). Second edition. New York: McGraw-Hill Co.

Salas, J. D., Raynal, J. A., Tarawneh, Z. S., Lee, T. S., Frevert, D., & Fulp, T. (2008). Chapter 20: Extending Short Record of Hydrologic Data (pp. 717-760). In: *Hydrology and Hydraulics*. Singh, V. P. (ed.). Highlands Ranch, USA: Water Resources Publications.

Stedinger, J. R., & Tasker, G. D. (1985). Regional Hydrologic Analysis 1. Ordinary, Weighted and Generalized Least Squares Compared. *Water Resources Research*, *21*(9), 1421-1432.

Tasker, G. D. (1980). Hydrologic Regression with Weighted Least Squares. *Water Resources Research*, *16*(6), 1107-1113.

Tasker, G. D., & Stedinger, J. R. (1986). Regional Skew with Weighted LS Regression. *Journal of Water Resources Planning and Management*, *112*(2), 225-237.

Tasker, G. D., & Stedinger, J. R. (1989). An Operational GLS Model for Hydrologic Regression. *Journal of Hydrology*, *111*(1-4), 361-375.

Vogel, R. M., Wilson, I., & Daly, C. (1999). Regional Regression Models of Annual Streamflow for the United States. *Journal of Irrigation and Drainage Engineering*, *125*(3), 148-157.

Wilks, D. S. (2011). Chapter 7: Statistical forecasting (pp. 215-300). *Statistical Methods in the Atmospheric Sciences*. Third edition. San Diego: Academic Press (Elsevier).

Dirección institucional del autor

Dr. Daniel Francisco Campos Aranda

Profesor Jubilado de la Universidad Autónoma de San Luis Potosí
Genaro Codina 240, Colonia Jardines del Estadio
78280 San Luis Potosí, San Luis Potosí, México
campos_aranda@hotmail.com

Tecnología y Ciencias del Agua, vol. VII, núm. 4, julio-agosto de 2016, pp. 161-173
ISSN 0187-8336 •