Ingeniería Investigación y Tecnología, ISSN 2594 0732, I.2, 59-71, 1999
(artículo arbitrado)
DOI: http://dx.doi.org/10.22201/fi.25940732e.1999.01n2.008

Uso de la correlación y la regresión lineal para ampliar registros de volúmenes escurridos anuales

D.F. Campos-Aranda
Facultad de Ingeniería, UASLP
Email: dcampos@deimos.tc.uaslp.mx

(recibido: febrero de 1998; aceptado: septiembre de 1998)

Resumen

Inicialmente se destaca la importancia de las estimaciones de la media y la desviación estándar como parámetros estadísticos básicos, los cuales caracterizan la magnitud y la variabilidad de los volúmenes escurridos anuales en los estudios hidrológicos de planeación del aprovechamiento de los recursos hidráulicos de una región. Enseguida, se describen con detalle las ecuaciones que permiten evaluar cuantitativamente si es conveniente o no, desde un punto de vista estadístico, ampliar el registro corto de escurrimientos, con base en datos comunes, adicionales y disponibles; esto en una o dos estaciones hidrométricas cercanas, con las cuales, el registro reducido guarda cierta correlación (dependencia o asociación). Lo anterior, significa evaluar si con base en el registro ampliado, las estimaciones de la media y la variancia mejoran estadísticamente. Posteriormente, se realizan dos aplicaciones numéricas a casos reales; una, para el modelo bidimensional que utiliza una estación hidrométrica cercana para ampliar el registro corto; y la otra, para el modelo tridimensional, que emplea dos estaciones de aforo auxiliares. Por último, se citan tres observaciones generales relativas al trabajo.

Abstract

Firstly, the estimates of standard deviation and arithmetic mean as basic statistical parameters are emphasised, which point out the variability and magnitude of annual streamflow records in the hydrological studies of planning water-resource developments inside a region. Then the equations for quantitative evaluations of statistical convenience of extending a short streamflow record are described in detail. The previous makes use of additional and common data in one or two closer hydrometric stations, with these the short observed record has a certain correlation (dependence or association). Later two numerical applications to real problems are given, the first one for the two dimensional model, which use a one close hydrometric station in order to extend the short record, and the second one application for the three dimensional model which makes use of two auxiliary hydrometric stations. Lastly, three general observations about the paper are cited.

Introducción

Los parámetros estadísticos básicos, la media (*Xm*), la desviación estándar (*S*) y el coeficiente de variación ($Cv = S / Xm$), permiten especificar las características del escurrimiento anual; el primero, en relación con su magnitud promedio y los segundos, respecto a sus características de dispersión o variabilidad. Por ello, tales parámetros son de uso común en la planeación del aprovechamiento de los recursos hidráulicos.

Estos parámetros son usualmente estimados con base en los registros disponibles de volúmenes escurridos anuales, de manera que si el registro es corto, los parámetros evaluados pueden tener poca garantía o cercanía a los valores poblacionales. Por lo anterior, y bajo ciertas circunstancias, la correlación con datos de otros sitios o estaciones hidrométricas cercanas puede ser empleada para mejorar las estimaciones de los parámetros en sitios o localidades con registros reducidos. El uso de la regresión lineal es la técnica básica para ampliar un registro corto con base en

DOI: http://dx.doi.org/10.22201/fi.25940732e.1999.01n2.008

los datos de otro cercano con el cual se tiene cierta dependencia o asociación (correlación).

Por otra parte, cuando se desea asignar probabilidades a las estimaciones del potencial hidráulico o hidrológico de un río, es común emplear distribuciones de probabilidad del escurrimiento anual, las cuales pueden ser las de los modelos Normal, Log–normal o de Weibull, cuyas características cuantitativas son definidas a través de los parámetros estadísticos citados.

Por lo tanto, resulta conveniente investigar si la ampliación de un registro corto con base en la información común y adicional disponible en una o dos estaciones hidrométricas cercanas es conveniente desde un punto de vista estadístico, esto es, si mejora la estimación de los parámetros estadísticos, base de futuras evaluaciones asociadas con el potencial hidrológico de la corriente y con el desarrollo futuro de los recursos hidráulicos de una zona o región.

Repaso de conceptos estadísticos básicos

Media y variancia

La media de una muestra de *N* términos, cada uno de los cuales tiene un valor numérico X_i, está dada por

$$\text{Media} = Xm = (1/N)(X_1 + X_2 + \cdots\cdots + X_N) = (1/N)\Sigma X_i \quad (1)$$

donde, la sumatoria abarca de *1* a *N*. La media aritmética es una cuantificación del promedio o tendencia central de los datos. La media de la población es comúnmente designada por μ.

En cambio, la variancia de una muestra es una medida del grado de dispersión de los términos alrededor de su valor medio y por ello se define como:

$$\text{Variancia} = \text{Var}(X) = (1/N)\ \Sigma\ (X_i - Xm)^2 . \quad (2)$$

Nuevamente, aquí la sumatoria comprende de *1* a *N*. La variancia de la población es designada generalmente por σ^2. Los datos que poseen una variancia reducida están agrupados alrededor de su media, mientras que los datos con una variancia grande están dispersos con respecto a su media. La raíz cuadrada de la variancia es la desviación estándar *(S)*.

Entonces, si se desea estimar un parámetro estadístico poblacional por ejemplo la media, y se dispone de varias series o secuencias de datos, cada una con una media, tales medias forman otra serie a partir de la cual es posible calcular la variancia de la media. Un valor reducido de esta variancia significa una estimación más confiable de la media, ya que implica menor dispersión y por lo tanto mayor precisión (Fiering,1963).

Regresión y correlación lineales

Una relación lineal de la forma

$$Y = \alpha + \beta \cdot X, \quad (3)$$

donde *Y* y *X* son los escurrimientos en dos corrientes, se denomina regresión lineal, y la técnica de mínimos cuadrados conduce a la expresión siguiente del coeficiente de regresión β:

$$\beta = \text{Cov}(X,Y) / \text{Var}(X) = (\Sigma X_i Y_i - N\ Xm\ Ym) / (\Sigma X_i^2 - N\ Xm^2) \quad (4)$$

La técnica de mínimos cuadrados también define al coeficiente de correlación (r_{xy}) como una medida de la calidad o bondad del ajuste (dependencia) entre las variables *X* y *Y*, estando definido como

$$r_{xy} = \text{Cov}(X,Y) / [\text{Var}(X)\ \text{Var}(Y)]^{1/2} \quad (5a)$$

$$r_{xy} = (\Sigma X_i\ Y_i - N\ Xm\ Ym) / [(\Sigma X_i^2 - N\ Xm^2) 1/2\ (\Sigma\ Y_i^2 - N\ Ym^2)^{1/2}] \quad (5b)$$

Un valor de r_{xy} igual a ±1 indica que las variables están perfectamente correlacionadas; en cambio, un valor de cero significa que no existe relación lineal entre ellas; los valores intermedios definen el grado de ajuste o dependencia entre *X* y *Y*. El coeficiente de correlación poblacional se designa por ρ. Entonces, la regresión corresponde a la ecuación matemática que relaciona las variables y la correlación, a la magnitud que indica el grado de asociación o dependencia entre las variables dependiente y la(s) independiente(s).

Hipótesis implícitas en la aplicación de la regresión

En lo que se denominará el ***modelo bidimensional***, X_i y Y_i son los escurrimientos anuales en dos corrientes cercanas (*X,Y*) o en dos estaciones hidrométricas próximas sobre el mismo río, con el arreglo siguiente en cuanto al tiempo (Fiering,1963):

$$\begin{array}{l} X_1, X_2, \ldots\ldots X_{n1}, \ldots\ldots X_{n1+n2} \\ Y_1, Y_2, \ldots\ldots Y_{n1,} \end{array} \quad (A.1)$$

donde *n1* es el número de años del registro disponible en la estación *Y*, que además es común con la estación *X*; en esta última existen *n2* datos o términos adicionales que pueden

DOI: http://dx.doi.org/10.22201/fi.25940732e.1999.01n2.008

ser utilizados para ampliar el registro *Y* de acuerdo con la siguiente ecuación de la línea de tendencia:

$$Y_j = Ym + \beta\,(X_j - Xm), \tag{6}$$

donde $j = 1, 2, \ldots, n2$ y *Ym* es la media de los *n1* datos Y_i. La aplicación de la ecuación anterior implica las siguientes hipótesis:

Primera. Un régimen hidrológico estable existe, de manera que una correlación entre X_i y Y_i, puede ser esperada.

Segunda. Los escurrimientos anuales, o alguna transformación de ellos, están distribuidos normalmente.

Tercera. La correlación serial de los escurrimientos anuales es cero.

¿Cuándo se debe usar la regresión?

Ya que los datos o una transformación de ellos están distribuidos normalmente, los primeros dos momentos estadísticos permiten especificar completamente su distribución. Entonces, un tópico importante será decidir cuándo los datos comunes y adicionales disponibles en la estación *X* pueden ser utilizados para alcanzar una mejor estimación de los parámetros poblacionales del registro *Y*.

En virtud de que la magnitud de la variancia de cada estimación es una medida de la precisión asociada a ésta, entonces, si la variancia de un parámetro calculada utilizando los *n1* datos originales excede a la estimada empleando también los *n2* adicionales obtenidos por regresión (según la ecuación 6), está claro que el uso de la correlación mejora la estimación del parámetro. En cambio, si la variancia de un parámetro calculada con el registro ampliado excede a la estimada con el registro original, entonces, el registro combinado proporciona una estimación menos precisa del parámetro.

Se puede definir, para un arreglo específico A.1 y un parámetro estadístico en particular, con base en los valores de *n1*, *n2* y ρ, el Cociente de Información Relativa (CIR) como la relación entre la variancia de un parámetro, estimada a partir del registro original y la estimada con el registro ampliado. Cuando el *CIR* excede la unidad, una estimación más precisa se logra con el registro ampliado. Por lo tanto, la regresión no debe ser utilizada para ampliar el registro original cuando el *CIR* es menor que uno (Fiering,1963).

Modelo bidimensional

Cociente de información relativa de la media

Para un arreglo como A.1 con datos *n1*, *n2* y r_{xy}, se tiene que (Fiering,1963)

$$CIRm = \{1 - [\,n2 / (n1+n2)]\,[(r_{xy}^2\,(n1-2) - 1) / (n1-3)]\,\}^{-1}. \tag{7}$$

Resolviendo la ecuación anterior para $CIRm = 1.0$, se obtiene el valor crítico

$$r_{xy} = [1/(n1-2)]^{1/2}, \tag{8}$$

por lo cual:

n1	5	10	15	20	25	30	40	50	75	100
r_{xy}	0.577	0.354	0.277	0.236	0.209	0.189	0.162	0.144	0.117	0.101

Cociente de información relativa de la variancia

Para un arreglo como el A.1 con datos *n1*, *n2* y r_{xy}, se tiene que (Fiering,1963)

$$CIRv = [\,2 / (n1-1)] / EMC[Var(Y)], \tag{9}$$

siendo *EMC* el error medio cuadrático de la variancia de las *Y*:

$$EMC[Var(Y)] = [2/(n1-1)] + D[2\,A + (n2+2)\,B + (n1+n2-1)\,C - F], \tag{10}$$

donde:

$$A = (n1-1)\,r_{xy}^4 + (n1+4)\,r_{xy}^2\,(1-r_{xy}^2) + [(n1+1)(1-r_{xy}^2)2 / (n1-3)]$$

$$B = r_{xy}^4 + [6\,r_{xy}^2\,(1-r_{xy}^2) / (n1-3)] + \{3(1-r_{xy}^2) / [(n1-3)(n1-5)]\,\}$$

$$C = [2(n1-4)\,(1-r_{xy}^2)] / (n1-3)$$

$$D = n2 / (n1+n2-1)^2$$

$$F = [(n1+1)(2\,n1+n2-2)] / (n1-1)$$

Para los valores típicos o de la práctica de *n1* y *n2*, *CIRv* excede la unidad cuando r_{xy} es igual o mayor de 0.80, de donde se deduce que las condiciones deseables que mejoran las estimaciones de la variancia son mucho más restrictivas que aquéllas que mejoran las estimaciones de la media (ecuación 8).

Modelo tridimensional

Arreglo y ecuación de regresión

En este caso, las estimaciones de media y variancia en un sitio *Z* pueden ser mejoradas por correlación con base en los escurrimientos anuales en dos sitios cercanos *X* y *Y*, cuyo arreglo en el tiempo es:

$$\begin{array}{l} X_1, X_2, \ldots\ldots X_{n1}, \ldots\ldots X_{n1+n2} \\ Y_1, Y_2, \ldots\ldots Y_{n1}, \ldots\ldots Y_{n1+n2} \\ Z_1, Z_2, \ldots\ldots Z_{n1} \end{array} \tag{A.2}$$

DOI: http://dx.doi.org/10.22201/fi.25940732e.1999.01n2.008

La ecuación de regresión trivariada utilizada para ampliar los datos es (Fiering,1963)

$$Z_j = Zm + bx\,(X_j - Xm) + by\,(Y_j - Ym) \tag{11}$$

en la cual, *j* varía de 1 a *n2*, *Zm* es la media de los datos *Zj*, *bx* y *by* son los coeficientes parciales de regresión calculados a partir de los *n1* datos comunes en las tres corrientes o los tres sitios; sus ecuaciones de acuerdo a la teoría de mínimos cuadrados son

$$bx = [Var(Y)\cdot Cov(X,Z) - Cov(X,Y)\cdot Cov(Y,Z)]\,/\Delta, \tag{12}$$

$$by = [Var(X)\cdot Cov(Y,Z) - Cov(X,Y)\cdot Cov(X,Z)]/\Delta, \tag{13}$$
$$\Delta = Var(X)\cdot Var(Y) - [Cov(X,Y)]^2 .$$

Cociente de información relativa de la media

Para un arreglo como el A.2 con datos *n1* y *n2*, se tiene que (Fiering,1963)

$$CIRM = \{1-(n2\,/\,N)\,[(R^2\,(n1-2)-2)\,/\,(n1-4)]\,\}^{-1}, \tag{14}$$

siendo $N = n1 + n2$. R^2 es el coeficiente de correlación total de *Z* sobre *X* y *Y*:

$$R^2 = (r_{zx}^2 + r_{zy}^2 - 2\,r_{xy}\,r_{zx}\,r_{zy})\,/\,(1 - r_{xy}^2) \tag{15}$$

Cociente de información relativa de la variancia

Para un arreglo como el A.2 con datos *n1* y *n2*, se tiene que (Fiering,1963)

$$CIRV = 2\,/\,[(n1-1)\,EMC(S^2)\,], \tag{16}$$

en la cual

$$EMC(S^2) = [1/(N-1)^2]\,\{2(N-1) + 4n2G + n2(n2-6)G^2 + 4n2(n2+3)G\,/\,(n1-4) - 4n2\,(2n2+1)G^2/(n1-4) + 8n2\,(n2+2)G^2\,/[(n1-4)\,(n1-6)]\,\},$$

siendo:

$$G = (1 - R^2)$$

Restricciones inherentes de las ecuaciones del *CIR*

Las ecuaciones del *CIR* de la variancia definen el número mínimo de *n1* que en el caso del modelo bidimensional es cinco según la ecuación 10 (último denominador del parámetro *B*), y en el del modelo tridimensional es seis como se deduce del último denominador de la expresión del *EMC*(S^2).

La comparación entre los valores de las expresiones del *CIRv* y *CIRV* conduce al hecho de que raramente la información estadística ganada en el modelo tridimensional excede, y generalmente es menor que la ganancia asociada con el modelo bidimensional (Fiering,1963).

Restricciones relativas a la homogeneidad hidrometeorológica

En el modelo bidimensional y principalmente en el trivariado es conveniente asegurarse de que las cuencas de los registros hidrométricos utilizados, formen parte de una región homogénea desde un punto de vista hidrometorológico, con lo cual se garantice de alguna manera su comportamiento similar y se acepte por lo tanto una relación causa-efecto semejante en todos los registros utilizados (Campos, 1991).

Aplicaciones numéricas

Modelo bidimensional en el río Piaxtla

Información disponible. La cuenca total del río Piaxtla es de aproximadamente 7 222 km² y su parteaguas define la mayor porción de la frontera sur de la región hidrológica No. 10 (Sinaloa), junto con la cuenca del río Quelite. La estación hidrométrica Piaxtla, operada por la Comisión Federal de Electricidad, tiene una cuenca de 5 307 km² e inició su operación en octubre de 1957. La estación de aforos Ixpalino drena 6 166 km², ésta pertenecía a la Secretaría de Recursos Hidráulicos e inició su operación en octubre de 1952 (SRH,1970). La información disponible sobre volúmenes escurridos anuales en Mm³ en las estaciones Piaxtla e Ixpalino, ambas en el río Piaxtla, ha sido tomada de la lámina II-04 (SRH, 1975) y abarca desde el comienzo de la operación hasta 1973, con el año 1961 incompleto para la estación Piaxtla; tal información se tiene en la tabla 1.

DOI: http://dx.doi.org/10.22201/fi.25940732e.1999.01n2.008

Tabla 1. Información hidrométrica en las estaciones del río Piaxtla que se indican. Sinaloa

Año	Vol. esc.anual (Mm^3)		Año	Vol. esc.anual (Mm^3)	
	Ixpalino	Piaxtla		Ixpalino	Piaxtla
1973	1 953.6	1 122.4	1962	948.8	803.1
1972	1 678.4	980.4	1961	2 208.2	1 570.7+
1971	1 413.7	829.8	1960	826.8	673.9
1970	1 699.5	1 081.3	1959	975.1	737.7
1969	1 101.6	800.7	1958	1 943.1	1 501.4
1968	2 754.4	2 364.9	1957	331.4	145.8*
1967	1 598.1	1 133.6	1956	632.1	381.9*
1966	1 906.7	1 360.9	1955	1 668.6	1 195.5*
1965	816.0	506.7	1954	1 145.7	785.0*
1964	1 472.2	988.3	1953	951.0	632.2*
1963	1 783.4	1 402.3			

+ Valor completado, ver inciso siguiente
** valores estimados, ver inciso: estimación de datos adicionales*

Deducción de datos faltantes. Al revisar el año de 1961 en la estación Piaxtla, con base en el resumen mensual (SRH,1975), se observó que solo existía un mes faltante, junio. Por lo anterior, el planteamiento para estimar el volumen escurrido de ese mes fue realizar un análisis probabilístico con base en los 15 datos disponibles, a través del ajuste de una distribución Pearson tipo III (Campos,1988) para obtener la mediana, la moda y el escurrimiento con 75% de probabilidad de excedencia; los resultados fueron: 14.504, 8.005 y 8.268 Mm^3, respectivamente. Se adoptó el valor de la moda y, entonces, el volumen escurrido anual del año 1961 se estimó en 1 570.652 Mm^3. Los datos disponibles en junio se presentan en la tabla 2.

Tabla 2. Volúmenes escurridos en junio en la estación hidrométrica Piaxtla

Año	Vol. esc.	Año	Vol. esc.	Año	Vol. esc.
1958	77.103	1964	22.288	1969	7.641
1959	13.824	1965	4.202	1970	15.021
1960	4.190	1966	12.684	1971	32.119
1962	9.035	1967	13.097	1972	15.059
1963	17.179	1968	9.894	1973	8.769

Verificación de las hipótesis implícitas en la regresión. El empleo de la ecuación 5b en los 16 datos comunes de las estaciones Ixpalino (X) y Piaxtla (Y) condujo al valor r_{xy} = 0.9377, por lo cual, la estadística z (apéndice) toma un valor de 6.1967, el cual es mayor que el valor crítico de 1.645 y por lo tanto el r_{xy} calculado es significativo. Por otra parte, el valor crítico r_o para 14 grados de libertad resulta de 0.497, el cual es menor que el r_{xy} calculado; entonces éste es significativo. De acuerdo con el procedimiento expuesto en el apéndice se obtuvieron los siguientes resultados:

Estación	Valor W calculado	W crítico	Resultado	Cs	K
Ixpalino (n=21)	0.977	0.908	normal	0.228	3.243
Piaxtla (n=16)	0.900	0.887	normal	1.378	6.093

DOI: http://dx.doi.org/10.22201/fi.25940732e.1999.01n2.008

donde $n = n1+n2$. Cs y K son los coeficientes de asimetría y curtosis, cuyos valores para la distribución Normal son 0.000 y 3.000, respectivamente; por lo tanto, se deduce que los datos muestrales de la estación hidrométrica Ixpalino son mucho más normales que los de la estación Piaxtla. Sin embargo, ambos registros cumplen con la condición de normalidad de acuerdo al *test W*. Por otra parte, el empleo de las ecuaciones I.5 y I.6 del apéndice condujo a los siguientes resultados:

Estación	r_l calculado	r_l crítico	Resultado
Ixpalino (n=21)	–0.177	–0.477	independiente
Piaxtla (n=16)	–0.276	–0.556	independiente

Por lo tanto, ambas muestras cumplen con la condición de independencia.

Cálculo de los cocientes de información relativa. El orden cronológico inverso en el que se ha expuesto la información disponible en la tabla 1 obedece al planteamiento del modelo bidimensional o arreglo A.1, por lo cual, $n1$ y $n2$ toman valores de 16 y 5, respectivamente. Para el valor calculado de $r_{xy} = 0.9377$, la aplicación de las ecuaciones 7 y 9 condujo a los valores siguientes, para los cocientes de información relativa de la media y de la variancia:

$$CIRm = 1.261;\ CIRv = 1.225,$$

de donde se deduce que la ampliación del registro de la estación hidrométrica Piaxtla, con base en los datos adicionales de la estación Ixpalino mejora la estimación de la media y la variancia.

Estimación de datos adicionales. Al aplicar la ecuación 4 se obtuvo un coeficiente de regresión de 0.785; el empleo de la ecuación 6 condujo a los valores presentados y destacados en la tabla 1 con un asterisco.

Modelo tridimensional en el río Fuerte

Información disponible. La cuenca del río Fuerte está localizada al norte de la región hidrológica No. 10. Sinaloa, su parteaguas define la frontera de esta región, es la más importante tanto por su extensión (33 590 km² hasta la estación hidrométrica San Blas), como por sus escurrimientos y la magnitud de las obras que en ella se han construido; abarca cuatro estados con los siguientes porcentajes: Chihuahua, 73.2; Sinaloa, 17.5; Sonora, 7.7; y Durango 1.6 (SRH,1970).

Las estaciones hidrométricas del río Fuerte utilizadas para ejemplificar el uso del modelo trivariado (arreglo A.2) fueron: *Huites* (X) en el río Fuerte cuya área de cuenca es de 26 020 km², *San Francisco* (Y) sobre el río San Miguel con un área drenada de 17 531 km² y *Palo dulce* (Z) en el río Chínipas con 6 323 km². La disponibilidad de datos de volúmenes escurridos anuales de acuerdo al concentrado de tal información en la lámina II-04 (SRH,1975) es la siguiente: las estaciones Huites y San Francisco tienen datos en el periodo 1942 a 1973, con los años 1943 y 1944 incompletos para San Francisco; la estación Palo Dulce presenta información desde 1958 hasta 1973, con el año 1961 faltante.

Deducción de datos faltantes. Al revisar la información disponible en la estación hidrométrica Palo Dulce en el sistema BANDAS (CNA,1997), se dedujo que existe información mensual en el periodo septiembre de 1957 a diciembre de 1976, con únicamente los meses de junio a noviembre faltantes en el año 1961. Por lo anterior, el planteamiento para la deducción de tales valores faltantes fue el análisis probabilístico por meses, con base en la distribución Pearson tipo III (Campos,1988) para estimar la moda, y cuando ésta no existe porque los datos definen una función de densidad tipo J invertida, se estimó el valor cuya probabilidad de excedencia es de 75%.

Bajo tal planteamiento, los valores estimados para los meses de junio, octubre y noviembre correspondieron al escurrimiento cuya probabilidad de no excedencia es de 25%, con los valores siguientes: 3.869, 24.271 y 7.746 Mm³, respectivamente. En el resto de los meses se empleó la moda con valores para julio, agosto y septiembre de: 109.087, 192.598 y 140.861 Mm³, respectivamente. Con base en tales estimaciones y los datos disponibles, el escurrimiento anual de 1961 se estimó en 797.204 Mm³. Como ejemplo de éstos análisis, en la tabla 3 se presentan los datos y resultados de los meses de junio y julio.

Tabla 3. Datos y resultados probalísticos en la estación Palo Dulce

Año	Vol. esc. (Mm³) junio	Vol. esc. (Mm³) julio	Año	Vol. esc. (Mm³) junio	Vol. esc. (Mm³) julio
1958	42.868	125.844	1963	2.879	94.279
1959	8.634	137.007	1964	4.561	122.452
1960	4.677	75.373	1965	1.546	93.538
1962	4.771	165.946	1966	26.512	237.343

continúa...

DOI: http://dx.doi.org/10.22201/fi.25940732e.1999.01n2.008

Tabla 3. Datos y resultados probalísticos en la estación Palo Dulce (continuación)

Año	Vol esc. (Mm^3) junio	Vol esc. (Mm^3) julio	Año	Vol esc. (Mm^3) junio	Vol esc. (Mm^3) julio
1967	31.727	198.649	1974	4.438	194.481
1968	10.883	306.711	1975	3.318	155.158
1969	1.869	364.868	1976	10.271	192.623
1970	1.507	117.701	Media	15.265	171.938
1971	62.905	270.443	Mediana [P(X<x)=0.50]	10.030	160.511
1972	47.382	178.784	Moda	---	136.869
1973	4.025	63.680	Esc. con P(X<x)=0.25	3.869	116.703

Planteamiento de los análisis. Nuevamente se utilizarán los datos disponibles en un orden cronológico inverso para cumplir con el arreglo A.2, de manera que ahora *n1* y *n2* serán igual a 12 (1973 -1962) y cuatro (1961-1958), respectivamente. Esta concepción en la aplicación del modelo tridimensional a la estación Palo Dulce permite tener un contraste de los cuatro valores estimados contra volúmenes anuales reales, uno de ellos, el del año 1961 estimado. Los volúmenes escurridos anuales (Mm^3) a utilizar son los mostrados en la tabla 4:

Tabla 4. Información hidrométrica en las estaciones que se indican del río Fuerte. Sinaloa

Año	Vol. esc. (X) Huites	Vol. esc. (Y) San Francisco	Vol. esc. (Z) Palo Dulce
1973	6 117.1	3 837.9	954.6
1972	5 546.7	3 566.3	1 044.4
1971	4 494.1	2 138.6	1 477.4
1970	3 252.5	2 083.7	632.8
1969	2 599.3	1 432.4	799.2
1968	6 725.6	4 238.1	1 205.8
1967	3 968.7	2 470.9	946.2
1966	5 991.9	3 764.3	1 330.5
1965	2 826.3	1 561.3	727.2
1964	3 326.1	2 239.6	677.5
1963	4 220.1	2 891.2	649.8
1962	2 679.5	1 992.3	637.9
1961	4 414.1	3 123.9	825.0*
1960	5 615.1	3 862.1	1 575.3*
1959	3 559.7	2 633.0	1 089.3*
1958	4 200.2	3 663.2	999.6*

** Valores reales que permitirán efectuar el contraste de los estimados por regresión*

DOI: http://dx.doi.org/10.22201/fi.25940732e.1999.01n2.008

Verificación de las hipótesis implícitas en la regresión. El empleo de la ecuación 5b en los 12 datos comunes de las estaciones Huites (X), San Francisco (Y) y Palo Dulce (Z) condujo a los valores: $r_{xy} = 0.962$, $r_{zx} = 0.689$ y $r_{zy} = 0.500$, por lo cual la estadística z (apéndice) toma un valor de 1.648 en el caso más desfavorable, el cual es escasamente mayor que el valor crítico de 1.645 y, por lo tanto, r_{zy} es aún significativo. Por otra parte, el valor crítico r_o para 10 grados de libertad resulta ser 0.576, el cual es mayor que el r_{zy} calculado, entonces éste no es diferente de cero.

Con base en el procedimiento expuesto en el apéndice se obtuvieron los siguientes resultados:

Estación	W calculado	W crítico	Resultado	Cs	K
Huites (n=16)	0.938	0.887	normal	0.368	2.518
San Fco. (n=16)	0.937	0.887	normal	−0.013	2.201
Palo Dulce (n=12)	0.890	0.859	normal	0.756	3.153

Nuevamente, $n = n1 + n2$. Todos los registros cumplen con la condición de normalidad y sus coeficientes de asimetría y curtosis se aproximan a los de la distribución Normal, cero y tres, respectivamente. El empleo de las ecuaciones I.5 y I.6 del apéndice condujo a los resultados siguientes:

Estación	r_1 calculado	r_1 crítico	Resultado
Huites (n=16)	−0.130	−0.556	independiente
San Fco. (n=16)	−0.168	−0.556	independiente
Palo Dulce (n=12)	0.047	0.382	independiente

Por lo tanto, los tres registros cumplen con la condición de independencia.

Cálculo de los cocientes de información relativa. Con base en los valores de r_{xy}, r_{zx} y r_{zy} la aplicación de la ecuación 15 aporta un valor de 0.915 para el coeficiente de correlación total y las ecuaciones 14 y 16 conducen a los valores siguientes, para los cocientes de información relativa de la media y variancia:

$$CIRM = 1.249;\ CIRV = 1.193,$$

de donde se deduce que la ampliación del registro de la estación hidrométrica Palo Dulce, con base en los datos de las estaciones Huites y San Francisco, mejora sensiblemente la estimación de la media y la variancia.

Deducción de datos adicionales. Al aplicar las ecuaciones 12 y 13 se obtuvieron los valores siguientes 0.5601 y −0.6763 para los coeficientes parciales de regresión. Por otra parte, con el uso de la ecuación 11 se obtienen las estimaciones de los valores de Z_j que se deducen; éstos fueron: 683.6, 857.0, 537.1 y 199.1 Mm^3.

Al comparar los valores anuales anteriores contra los volúmenes escurridos reales, la similitud es escasa; por ello, se realizaron los siguientes tres planteamientos y sus respectivos análisis numéricos:

1) Aplicar las ecuaciones 12, 13 y 11 a los logaritmos naturales de los datos X_i, Y_i y Z_i, lo cual implica realizar una regresión múltiple exponencial del tipo: $Z_j = a\ X^b\ Y^c$ (Campos, 1992a).

2) Aplicar el modelo bidimensional entre las estaciones Huites (X) y Palo Dulce (Y); los resultados fueron $r_{xy} = 0.689$, $\beta = 0.137$, $CIRm = 1.116$ y $CIRv = 1.050$.

3) Aplicar el modelo bidimensional entre las estaciones San Francisco (X) y Palo Dulce (Y); los resultados fueron $r_{xy} = 0.500$, $\beta = 0.152$, $CIRm = 1.043$ y $CIRv = 1.039$.

En la tabla 5 se contrastan los volúmenes escurridos anuales (VEA) reales, de la estación Palo Dulce en el periodo 1961-1958, contra los estimados con el modelo tridimensional y los obtenidos con cada uno de los tres planteamientos anteriores:

DOI: http://dx.doi.org/10.22201/fi.25940732e.1999.01n2.008

Tabla 5. Contraste de resultados en la estación hidrométrica Palo Dulce

VEA real (X)	VEA reg. lineal (Y1)	VEA reg. exponencial (Y2)	Mod. bidim. con Huites (Y3)	Mod. bidim. con San Fco. (Y4)
825.0	683.6	771.8	937.5	990.5
1 575.3	857.0	937.3	1 101.6	1 103.0
1 089.3	537.1	632.2	820.8	915.7
999.6	199.1	559.4	908.3	1 072.7
	$r_{xy1}=0.531$	$r_{xy2}=0.658$	$r_{xy3}=0.723$	$r_{xy4}=0.532$

Conclusiones

Primera. En este trabajo se han presentado las ecuaciones relativas a los modelos bi y tridimensionales de ampliación de registros de volúmenes escurridos anuales, haciendo uso de la correlación y la regresión lineal cuando ello es conveniente desde un punto de vista estadístico. Además, se han expuesto las herramientas estadísticas adicionales (apéndice) que son necesarias para la aplicación a casos reales.

Segunda. Se ha expuesto toda la información hidrométrica utilizada en las aplicaciones numéricas descritas (tablas 1 y 4) y se han detallado y comentado los resultados parciales y finales, incluyendo sus análisis complementarios (tabla 5) que se estimaron necesarios para llegar a resultados más confiables (caso trivariado).

Tercera. Debido a la complejidad numérica de algunas de las ecuaciones; por ejemplo, las asociadas al cociente de información relativa de la variancia, de ambos modelos (*CIRυ* y *CIRV*), así como las correspondientes a las ecuaciones de regresión, se desarrollaron programas de cómputo en *Basic* que las resuelven, los cuales están disponibles con el autor.

Apéndice. Pruebas para las hipótesis de existencia de correlación, normalidad e independencia

Prueba para el coeficiente de correlación

Para probar si existe la posibilidad de que el coeficiente de correlación poblacional (ρ) sea cero, aun cuando r_{xy}, r_{zx} y/o r_{zy} hayan resultado diferentes de cero, se evalúa la estadística

$$z = [(n1-3)^{1/2}/2]\ \mathrm{Ln}\,[(1+r_{xy})/(1-r_{xy})]\,, \qquad (I.1)$$

la cual se compara con el valor zc que tiene distribución normal para un cierto nivel de confianza, comúnmente 95%, por lo cual, $zc = 1.645$ (CFE,1970). Si $z>zc$ no hay posibilidad de que ρ sea igual a cero y por lo tanto, r_{xy} es significativo al 5% de nivel de significancia (Campos,1992a).

Otro criterio consiste en comparar el valor calculado del coeficiente de correlación (r_{xy}) contra el tabulado (r_o) con $\nu = n1-2$ grados de libertad; si el primero excede al segundo no existe posibilidad de que ρ sea igual a cero. Para un nivel de significancia de 5%, Maisel (1973) presenta los siguientes valores de r_o:

ν	r_0	ν	r_0	ν	r_0	ν	r_0
8	0 632	15	0.482	22	0.404	29	0.355
9	0.602	16	0.468	23	0.396	30	0.349
10	0.576	17	0.456	24	0.388	35	0.325
11	0.553	18	0.444	25	0.381	40	0.304
12	0.532	19	0.433	26	0.374	45	0.288
13	0.514	20	0.423	27	0.367	50	0.273
14	0.497	21	0.413	28	0.361	60	0.250

DOI: http://dx.doi.org/10.22201/fi.25940732e.1999.01n2.008

Prueba de normalidad

El *test W* de Shapiro & Wilk permite verificar si es probable que una muestra provenga de una población normal; su procedimiento es el siguiente (Ruiz-Maya,1977): en la muestra de n elementos ($n<50$) denominados x_i, se ordenan éstos de menor a mayor, siendo ahora el i-ésimo elemento X_i. En seguida se calculan las cantidades siguientes:

$$S^2 = \Sigma_1 (X_i - Xm)^2 \qquad (1.2)$$

$$b = \Sigma_2\, a_{n,k}\, (X_{n-i+1} - X_i) \qquad (1.3)$$

$$W = b^2 / S^2 \qquad (I.4)$$

La sumatoria 1 abarca de i=1 a i=n y la sumatoria 2 de i=1 a i=k. Si n es par, entonces $k = n / 2$ y si n es impar $k = (n-1) / 2$; $a_{n,k}$ son los coeficientes que se tienen en la tabla I.1. Si el valor calculado de W es mayor que el de la tabla I.2, se acepta la hipótesis que establece que la muestra proviene de una población normal en el nivel de significancia (α) adoptado, comúnmente 5% (Ruiz-Maya, 1977; Campos, 1992b).

Tabla I.1. Coeficientes $a_{n,k}$ del test W de Shapiro & Wilk (Ruiz-Maya, 1977)

k/n	10	11	12	13	14	15	16	17	18	19	20
1	0.5 739	0.5 601	0.5 475	0.5 359	0.5 251	0.5 150	0.5 056	0.4 968	0.4 886	0.4 808	0.4 734
2	.3 291	.3 315	.3 325	.3 325	.3 318	.3 306	.3 290	.3 273	.3 253	.3 232	.3 211
3	.2 141	.2 260	.2 347	.2 412	.2 460	.2 495	.2 521	.2 540	.2 553	.2 561	.2 565
4	.1 224	.1 429	.1 586	.1 707	.1 802	.1 878	.1 939	.1 988	.2 027	.2 059	.2 085
5	.0 399	.0 695	.0 922	.1 099	.1 240	.1 353	.1 447	.1 524	.1 587	.1 641	.1 686
6		0.0 000	0.0 303	0.0 539	0.0 727	0.0 880	0.1 005	0.1 109	0.1 197	0.1 271	0.1 334
7		-	-	.0 000	.0 240	.0 433	.0 593	.0 725	.0 837	.0 932	.1 013
8		-	-	-	-	.0 000	.0 196	.0 359	.0 496	.0 612	.0 711
9		-	-	-	-	-	-	.0 000	.0 163	.0 303	.0 422
10		-	-	-	-	-	-	-	-	.0 000	.0 140

k/n	21	22	23	24	25	26	27	28	29	30
1	0.4 643	0.4 590	0.4 542	0.4 493	0.4 450	0.4 407	0.4 366	0.4 328	0.4 291	0.4 254
2	.3 185	.3 156	.3 126	.3 098	.3 069	.3 043	.3 018	.2 992	.2 968	.2 944
3	.2 578	.2 571	.2 563	.2 554	.2 543	.2 533	.2 522	.2 510	.2 499	.2 487
4	.2 119	.2 131	.2 139	.2 145	.2 148	.2 151	.2 152	.2 151	.2 150	.2 148
5	.1 736	.1 764	.1 787	.1 807	.1 822	.1 836	.1 848	.1 857	.1 864	.1 870
6	0.1 399	0.1 443	0.1 480	0.1 512	0.1 539	0.1 563	0.1 584	0.1 601	0.1 616	0.1 630
7	.1 092	.1 150	.1 201	.1 245	.1 283	.1 316	.1 346	.1 372	.1 395	.1 415
8	.0 804	.0 878	.0 941	.0 997	.1 046	.1 089	.1 128	.1 162	.1 192	.1 219
9	.0 530	.0 618	.0 696	.0 764	.0 823	.0 876	.0 923	.0 965	.1 002	.1 036
10	.0 263	.0 368	.0 459	.0539	.0 610	.0 672	.0 728	.0 778	.0 822	.0 862
11	0.0 000	0.0 122	0.0 228	0.0 321	0.0 403	0.0 476	0.0 540	0.0 598	0.0 650	0.0 697
12	-	-	.0 000	.0 107	.0 200	.0 284	.0 358	.0 424	.0 483	.0 537
13	-	-	-	-	.0 000	.0 094	.0 178	.0 253	.0 320	.0 381
14	-	-	-	-	-	-	.0 000	.0 084	.0 159	.0 227
15	-	-	-	-	-	-	-	-	.0 000	.0 076

continúa...

DOI: http://dx.doi.org/10.22201/fi.25940732e.1999.01n2.008

Tabla I.1. Coeficientes $a_{n,k}$ del test W de Shapiro & Wilk (Ruiz-Maya,1977) (continuación)

k/n	31	32	33	34	35	36	37	38	39	40
1	0.4 220	0.4 188	0.4 156	0.4 127	0.4 096	0.4 068	0.4 040	0.4 015	0.3 989	0.3 064
2	.2 921	.2 898	.2 876	.2 854	.2 834	.2 813	.2 794	.2 774	.2 755	.2 737
3	.2 475	.2 463	.2 451	.2 439	.2 427	.2 415	.2 403	.2 391	.2 380	.2 368
4	.2 145	.2 141	.2 137	.2 132	.2 127	.2 121	.2 116	.2 110	.2 104	.2 098
5	.1 874	.1 878	.1 880	.1 882	.1 883	.1 883	.1 883	.1 881	.1 880	.1 878
6	0.1 641	0.1 651	0.1 660	0.1 667	0.1 673	0.1 678	0.1 683	0.1 686	0.1 689	0.1 691
7	.1 433	.1 449	.1 463	.1 475	.1 487	.1 496	.1 505	.1 513	.1 520	.1 526
8	.1 243	.1 265	.1 284	.1 301	.1 317	.1 331	.1 344	.1 356	.1 366	.1 376
9	.1 066	.1 093	.1 118	.1 140	.1 160	.1 179	.1 196	.1 211	.1 225	.1 237
10	.0 899	.0 931	.0 961	.0 988	.1 013	.1 036	.1 056	.1 075	.1 092	.1 108
11	0.0 739	0.0 777	0.0 812	0.0 844	0.0 873	0.0 900	0.0 924	0.0 947	0.0 967	0.0 986
12	.0 585	.0 629	.0 669	.0 706	.0 739	.0 770	.0 798	.0 824	.0 848	.0 870
13	.0 435	.0 485	.0 530	.0 572	.0 610	.0 645	.0 677	.0 706	.0 733	.0 759
14	.0 289	.0 344	.0 395	.0 441	.0 484	.0 523	.0 559	.0 592	.0 622	.0 651
15	.0 144	.0 206	.0 262	.0 314	.0 361	.0 404	.0 444	.0 481	.0 515	.0 546
16	0.0 000	0.0 068	0.0 131	0.0 187	0.0 239	0.0 287	0.0 331	0.0 372	0.0 409	0.0 444
17	-	-	.0 000	.0 062	.0 119	.0 172	.0 220	.0 264	.0 305	.0 343
18	-	-	-	-	.0 000	.0 057	.0 110	.0 158	.0 203	.0 244
19	-	-	-	-	-	-	.0 000	.0 053	.0 101	.0 146
20	-	-	-	-	-	-	-	-	.0 000	.0 049

k/n	41	42	43	44	45	46	47	48	49	50
1	0.3 940	0.3 917	0.3 894	0.3 872	0.3 850	0.3 830	0.3 808	0.3 789	0.3 770	0.3 751
2	.2 719	.2 701	.2 684	.2 667	.2 651	.2 635	.2 620	.2 604	.2 589	.2 574
3	.2 357	.2 345	.2 334	.2 323	.2 313	.2 302	.2 291	.2 281	.2 271	.2 260
4	.2 091	.2 085	.2 078	.2 072	.2 065	.2 058	.2 052	.2 045	.2 038	.2 032
5	.1 876	.1 874	.1 871	.1 868	.1 865	.1 862	.1 859	.1 855	.1 851	.1 847
6	0.1 693	0.1 694	0.1 695	0.1 695	0.1 695	0.1 695	0.1 695	0.1 693	0.1 692	0.1 691
7	.1 531	.1 535	.1 539	.1 542	.1 545	.1 548	.1 550	.1 551	.1 553	.1 554
8	.1 384	.1 392	.1 398	.1 405	.1 410	.1 415	.1 420	.1 423	.1 427	.1 430
9	.1 249	.1 259	.1 269	.1 278	.1 286	.1 293	.1 300	.1 306	.1 312	.1 317
10	.1 123	.1 136	.1 149	.1 160	.1 170	.1 180	.1 189	.1 197	.1 205	.1 212
11	0.1 004	0.1 020	0.1 035	0.1 049	0.1 062	0.1 073	0.1 085	0.1 095	0.1 105	0.1 113
12	.0 891	.0 909	.0 927	.0 943	.0 959	.0 972	.0 986	.0 998	.1 010	1 020
13	.0 782	.0 804	.0 824	.0 842	.0 860	.0 876	.0 892	.0 906	.0 919	.0 932
14	.0 677	.0 701	.0 724	.0 745	.0 765	.0 783	.0 801	.0 817	.0 832	.0 846

continúa...

DOI: http://dx.doi.org/10.22201/fi.25940732e.1999.01n2.008

Tabla I.1. Coeficientes $a_{n,k}$ del test W de Shapiro & Wilk (Ruiz-Maya,1977) (continuación)

k/n	41	42	43	44	45	46	47	48	49	50
15	.0 575	.0 602	.0 628	.0 651	.0 673	.0 694	.0 713	.0 731	.0 748	.0 764
16	0.0 476	0.0 506	0.0 534	0.0 560	0.0 584	0.0 607	0.0 628	0.0 648	0.0 667	0.0 685
17	.0 379	.0 411	.0 442	.0 471	.0 497	.0 522	.0 546	.0 568	.0 588	.0 608
18	.0 283	.0 318	.0 352	.0 383	.0 412	.0 439	.0 465	.0 489	.0 511	.0 532
19	.0 188	.0 227	.0 263	.0 296	.0 328	.0 357	.0 385	.0 411	.0 436	.0 459
20	.0 094	.0 136	.0 175	.0 211	.0 245	.0 277	.0 307	.0 335	.0 361	.0 386
21	0.0 000	0.0 045	0.0 087	0.0 126	0.0 163	0.0 197	0.0 229	0.0 259	0.0 288	0.0 314
22	-	-	.0 000	.0 042	.0 081	.0 118	.0 153	.0 185	.0 215	.0 244
23	-	-	-	-	.0 000	.0 039	.0 076	.0 111	0. 143	.0 174
24	-	-	-	-	-	-	.0 000	.0 037	.0 071	.0 104
25	-	-	-	-	-	-	-	-	.0 000	.0 035

Tabla I.2. Valores críticos de W en el test W de Shapiro & Wilk (Ruiz-Maya, 1977)

n	α=1%	α=2%	α=5%	α=10%	n	α=1%	α=2%	α=5%	α=10%
3	0.753	0.756	0.767	0.789	27	.894	.906	.923	.935
4	.687	.707	.748	.792	28	.896	.908	.924	.936
5	.686	.715	.762	.806	29	.898	.910	.926	.937
6	0.713	0.743	0.788	0.826	30	.900	.912	.927	.939
7	.730	.760	.803	.838	31	0.902	0.914	0.929	0.940
8	.749	.778	.818	.851	32	.904	.915	.930	.941
9	.764	.791	.829	.859	33	.906	.917	.931	.942
10	.781	.806	.842	.869	34	.908	.919	.933	.943
11	0.792	0.817	0.850	0.876	35	.910	.920	.934	.944
12	.805	.828	.859	.883	36	0.912	0.922	0.935	0.945
13	.814	.837	.866	.889	37	.914	.924	.936	.946
14	.825	.846	.874	.895	38	.916	.925	.938	.947
15	.835	.855	.881	.901	39	.917	.927	.939	.948
16	0.844	0.863	0.887	0.906	40	.919	.928	.940	.949
17	.851	.869	.892	.910	41	0.920	0.929	0.941	0.950
18	.858	.874	.897	.914	42	.922	.930	.942	.951
19	.863	.879	.901	.917	43	.923	.932	.943	.951
20	.868	.884	.905	.920	44	.924	.933	.944	.952
21	0.873	0.888	0.908	0.923	45	.926	.934	.945	.953
22	.878	.892	.911	.926	46	0.927	0.935	0.945	0.953
23	.881	.895	.914	.928	47	.928	.936	.946	.954
24	.884	.898	.916	.930	48	.929	.937	.947	.954

continúa...

DOI: http://dx.doi.org/10.22201/fi.25940732e.1999.01n2.008

Tabla I.2. Valores críticos de W en el test W de Shapiro & Wilk (Ruiz-Maya, 1977) (continuación)

25	.888	.901	.918	.931	49	.929	.937	.947	.955
26	0.891	0.904	0.920	0.933	50	.930	.938	.947	.955

α= nivel de significancia

Prueba de independencia

Se define como correlación serial o autocorrelación la dependencia existente entre un elemento de la muestra y otro distante k elementos. Entonces, el coeficiente de correlación serial unitario estará definido por la expresión (WMO,1971)

$$r_1 = [(n-1)\Sigma_1 x_i \cdot x_{i+1} - (\Sigma_1 x_i)(\Sigma_2 x_i)]/\{[(n-1)\Sigma_1 x_i^2 - (\Sigma_1 x_i)^2]^{1/2}[(n-1)\Sigma_2 x_i^2 - (\Sigma_2 x_i)^2]^{1/2}\}, \quad (I.5)$$

siendo x_i cada elemento de la serie cronológica compuesta de n términos. La sumatoria 1 abarca de $i=1$ a $i=n-1$ y la 2, de $i=2$ a $i=n$.

Si la distribución de frecuencias de la serie se aproxima a la normal o gaussiana, el valor crítico de r_1 estará dado por

$$(r_1)_c = [-1 + Z\,(n-2)^{1/2}]\,/\,(n-1), \quad (I.6)$$

donde, Z es la desviación normal estándar de una cola al 95% de probabilidad, es decir, Z = 1.645; si y solo si, r_1 es mayor que $(r_1)_c$, se concluye que la serie es no aleatoria, inconsistente o que sus términos no son independientes. Cuando llegan a ocurrir valores negativos de r_1, se debe a que la serie tiene una oscilación marcada; entonces es preferible usar una prueba de dos colas (Z = 1.960) y se debe sustituir el signo más de la ecuación 22 por uno negativo (WMO,1971).

Cuando se construye un *correlograma*, gráfica que relaciona los retrasos (k) en el eje de las abscisas contra los correspondientes coeficientes de correlación serial de orden k, su significancia se prueba a través del test gráfico de Anderson (Linsley *et al.*,1977)

Referencias

Campos A. (1988). Función de distribución de probabilidades Gamma Mixta: Soluciones y aplicaciones, en: *Memoria del 10° Congreso nacional de hidráulica*, Tomo II, pp. 318-322. Morelia, Michoacán.

Campos A. (1991). Propuesta de criterios para la elaboración de estudios hidrológicos (contexto hidrológico regional). *Ingeniería hidráulica en México*, Vol. VI, No. 3, 23-40.

Campos A. (1992a). *Procesos del ciclo hidrológico*. Universitaria Potosina, San Luis Potosí, S.L.P.

Campos A. (1992b). Estudios de homogeneidad en 34 estaciones pluviométricas del altiplano potosino. XII Congreso Nacional de Hidráulica (AMH). Tema 5: Investigación y Tecnología, Puerto Vallarta, Jal., octubre 1992.

CFE. Comisión Federal de Electricidad (1970). *Manual de diseño de obras civiles*. CFE, Ed, México.

CNA. Comisión Nacional del Agua (1997). *Banco Nacional de datos de aguas superficiales* (BANDAS). Subdirección General Técnica. México.

Fiering M.B. (1963). *Use of Correlation to Improve Estimates of the Mean and Variance*. Geological Survey Professional Paper 434-C. U.S. Department of the Interior, Washington, D.C.

Linsley R.K., M.A. Kohler y J.L.H. Paulus (1977). *Hidrología para ingenieros*. Editorial McGraw-Hill Latinoamericana, Bogotá, Colombia.

Maisel L. (1973). *Probabilidad y estadística*. Fondo Educativo Interamericano. Bogotá, Colombia.

Ruiz-Maya L. (1977). *Métodos estadísticos de investigación*. Instituto Nacional de Estadística. Madrid, España.

SHR. Secretaría de Recursos Hidráulicos (1970). Jefatura de Irrigación y Control de Ríos. *Boletín hidrológico No. 36. Región hidrológica No. 10*.

SRH. Secretaría de Recursos Hidráulicos (1975). *Actualización al boletín hidrológico No. 36. Región hidrológica No. 10. Sinaloa, periodo 1970-73*. Subsecretaría de Planeación. México.

WMO. World Meteorological Organization (1971). Annexe III: Standard Tests of Significance to be Recommended in Routine Analysis of Climatic Fluctuations, pp. 58-71, in *Climatic Change*. Technical Note No. 79, WMO-No. 195. T.P. 100. Geneva, Switzerland.

Semblanza del autor

Daniel Francisco Campos-Aranda. Concluyó la licenciatura de Ingeniería Civil en 1972 en la Escuela de Ingeniería de la UASLP. La División de Estudios de Posgrado de la Facultad de Ingeniería, UNAM le otorgó los títulos de maestro (1980); y doctor (1987) en ingeniería hidráulica y aprovechamientos hidráulicos, respectivamente.En 1989, obtuvo la medalla "Gabino Barreda" de la UNAM. Ha publicado en revistas nacionales como: Ingeniería hidráulica en México y Agrociencia. Asimismo, ha participado en congresos nacionales e internacionales. Actualmente, es investigador nacional y profesor de la División de Educación Continua en la Facultad de Ingeniería, UNAM.