Nota técnica

Verificación de la homogeneidad regional mediante tres pruebas estadísticas

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

Resumen

Inicialmente se exponen conceptos generales sobre el análisis regional de frecuencia de crecientes. Enseguida se describe con detalle la prueba de Discordancia para detectar registros con datos anómalos y, por lo tanto, discordantes con el resto. A continuación se expone la prueba *H* de heterogeneidad, basada en la simulación de 500 regiones homogéneas, y el test de Wiltshire, el cual utiliza el coeficiente de variación local y regional para establecer un estadístico con distribución χ^2. Posteriormente se aplican las tres pruebas estadísticas citadas en cuatro regiones originalmente aceptadas como homogéneas y se analizan sus resultados. Por último se formulan las conclusiones, las cuales sugieren la aplicación sistemática de las tres pruebas para decidir sobre la homogeneidad de la región estudiada.

Palabras clave: momentos *L*, discordancia, simulación, prueba *H*, test de Wiltshire.

Introducción

De manera general, los eventos hidrológicos extremos de la naturaleza, como crecientes, sequías, tormentas severas y vientos fuertes causan daños en la sociedad. Por ello, estimar con cierta precisión qué tan frecuente es un evento de determinada magnitud es de enorme importancia; sin embargo, la estimación probabilística de tales eventos extremos es bastante difícil, pues por definición son raros y sus registros disponibles cortos (Hosking y Wallis, 1997).

El análisis regional de frecuencia de crecientes (ARFC) enfrenta el problema "comerciando espacio por tiempo", ya que los datos de varios sitios son utilizados para estimar los eventos extremos de una localidad en particular. Este enfoque es válido debido a que las muestras de crecientes utilizadas son típicamente observaciones de la misma variable en un número determinado de sitios de medición dentro de una *región* apropiadamente definida.

El término región sugiere una serie de sitios aledaños; sin embargo, la cercanía geográfica no necesariamente es un indicador de similaridad, por ejemplo, en la función de distribución de probabilidades (FDP). Entonces resulta razonable identificar regiones midiendo en cada sitio variables que tienen influencia en la FDP y después agrupar sitios que muestran características similares. Tales variables son de dos tipos: las asociadas con la cuenca (área, altitud media, lluvia anual o puntual, etcétera) y las relativas al registro de crecientes, como son sus propiedades estadísticas. Por ello, tanto para la integración de las regiones *homogéneas* como para su verificación se han formulado diversos procedimientos y pruebas estadísticas.

Ya que todo ARFC comienza con la revisión de los datos, este trabajo inicia describiendo con detalle la prueba estadística de Discordancia, la cual está basada en los momentos *L* y permite detectar registros anómalos. Enseguida se describen con sus pormenores las otras dos pruebas estadísticas de homogeneidad regional, la primera basada en simulación y la segunda en la dispersión de los coeficientes de variación.

Las tres pruebas estadísticas se aplican en cuatro regiones de México que han sido consideradas homogéneas para realizar diversos análisis hidrológicos, como son predicciones de crecientes, del volumen de sedimentos en cuencas sin aforos y de la precipitación máxima diaria. Como resultado de estas aplicaciones y de otras más que no fueron descritas por razones de espacio, se formulan dos conclusiones que recomiendan la aplicación sistemática de las pruebas descritas, para verificar la homogeneidad regional, antes de proceder con las técnicas del ARFC, basadas en conjunción de datos y ponderación de parámetros estadísticos.

Procedimientos

Revisión de los datos

En el inicio de todo ARFC, al menos se deben realizar las siguientes dos verificaciones: (1) cada muestra o serie de datos de un sitio debe ser revisada para buscar datos erróneos; es decir, valores demasiado grandes o extremadamente reducidos, así como repetidos, los cuales se pudieron originar en la transcripción; (2) se deben buscar tendencias en cada serie y comparar las muestras entre ellas y con las más cercanas. Además, los datos deben mostrar una "evolución" o cambio en magnitud, por ejemplo conforme el tamaño de la cuenca crece o su ubicación varía de una zona a otra de la región analizada.

Afortunadamente, los valores erróneos, los eventos dispersos (*outliers*), la tendencia, y los saltos o cambios en la media de los datos son reflejados en los momentos L de la muestra (ver apéndice). Por ello, una mezcla conveniente de los cocientes L en un solo estadístico (D_i) que mida la *discordancia* entre los cocientes L del sitio y los promedio de grupo se ha sugerido como prueba básica para detectar sitios que son discordantes con el grupo como un todo (Hosking y Wallis, 1997).

Test de Discordancia (primera prueba)

Suponiendo que existen N sitios en el grupo que se analiza, se define a $u_i = [t^i \; t_3^i \; t_4^i]^T$ como un vector que contiene los cocientes L: t, t_3 y t_4 de cada sitio i, definidos por las ecuaciones (A.6) a (A.8) del apéndice. El superíndice T significa transpuesto, ya que u_i es un vector renglón. El vector promedio (no ponderado) del grupo será (Hosking y Wallis, 1997):

$$\bar{\mathbf{u}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{u}_i \qquad (1)$$

La matriz **A** de suma de cuadrados y de productos cruzados estará definida como:

$$A = \sum_{i=1}^{N}\left(\mathbf{u}_i - \bar{\mathbf{u}}\right)\cdot\left(\mathbf{u}_i - \bar{\mathbf{u}}\right)^T \qquad (2)$$

Finalmente, la medida de la Discordancia de cada sitio será:

$$D_i = \frac{1}{3}N\left(\mathbf{u}_i - \bar{\mathbf{u}}\right)^T A^{-1}\left(\mathbf{u}_i - \bar{\mathbf{u}}\right) \qquad (3)$$

Entonces, cuando D_i es mayor que los valores críticos (D_c, cuadro 1), el sitio será discordante con el grupo. Una descripción más detallada de la prueba se tiene en Campos (2008c).

Test H de homogeneidad regional (segunda prueba)

El objetivo es estimar el grado de heterogeneidad en un grupo de sitios y evaluar

Cuadro 1. Valores críticos (D_c) para el estadístico de Discordancia (D_i).

N	D_c	N	D_c	N	D_c
5	1.333	9	2.329	13	2.869
6	1.648	10	2.491	14	2.971
7	1.917	11	2.632	≥15	3.000
8	2.140	12	2.757	–	–

cuando éstos pueden ser tratados como una región homogénea. Específicamente, la medida de heterogeneidad compara las variaciones entre sitios en los momentos *L* muestrales para el grupo, con la variación que cabe esperar en una región homogénea, la cual tiene igual distribución de probabilidades.

Suponiendo que la región propuesta tiene *N* sitios, cada localidad *j* tiene una longitud de registro de datos n_j, momentos *L* muestrales (l_k) y cociente *L–Cv* muestral t^j (ver apéndice). Entonces sus valores regionales ponderados serán:

$$l_k^R = \frac{\sum_{j=1}^{N} n_j \cdot l_k^j}{\sum_{j=1}^{N} n_j} \qquad \text{con } k = 1, 2, \ldots, 5 \qquad (4)$$

$$t^R = \frac{\sum_{j=1}^{N} n_j \cdot t^j}{\sum_{j=1}^{N} n_j} \qquad (5)$$

La desviación estándar ponderada de los cocientes *L–Cv* muestrales será:

$$V = \left[\frac{\sum_{j=1}^{N} n_j \cdot \left(t^j - t^R\right)^2}{\sum_{j=1}^{N} n_j} \right]^{1/2} \qquad (6)$$

Con base en los resultados de la ecuación (4) se calculan los cocientes de momentos t_3 y t_4 regionales, se llevan a la figura 1 y se define el mejor modelo probabilístico para simular 500 regiones homogéneas con *N* sitios, cada uno con n_j longitudes de registro. Estas regiones, por lo tanto, no tienen correlación cruzada ni serial. Para cada región sintética se calcula *V* (ecuación (6)), determinando al final de la simulación su media (μ_V) y su desviación estándar (σ_V). Finalmente, la medida de heterogeneidad será (Hosking y Wallis, 1997):

$$H = \frac{(V - \mu_V)}{\sigma_V} \qquad (7)$$

cuando $H \geq 2$, la región es definitivamente heterogénea; cuando $1 \leq H < 2$, la región es posiblemente heterogénea, y cuando $H < 1$, la región es aceptablemente homogénea. Valores de *H* cercanos a la unidad sugieren redefinir la región y cercanos a dos implican su modificación. Los valores negativos de *H* ocurren, indicando que hay menor dispersión entre los cocientes *L–Cv* muestrales de cada sitio y la que se debe esperar en una región homogénea con distribuciones de probabilidad independientes en cada sitio. Entonces ocurre correlación cruzada excesiva o existe una regularidad notable en los datos.

La simulación de las 500 regiones homogéneas se realiza generando números aleatorios con distribución uniforme u_i en el intervalo 0 a 1 y considerando tales valores iguales a la probabilidad de no excedencia, para estimar a través de la solución inversa del modelo probabilístico utilizado, el dato sintético que se busca. El algoritmo que se usó para generar los números aleatorios u_i se detalla en Metcalfe (1997).

Cuando los cocientes de momentos t_3 y t_4 regionales no definen en la figura 1 de manera clara o precisa una FDP idónea, se puede utilizar en las simulaciones un modelo probabilístico general, como la distribución Wakeby o la Kappa (Hosking y Wallis, 1997).

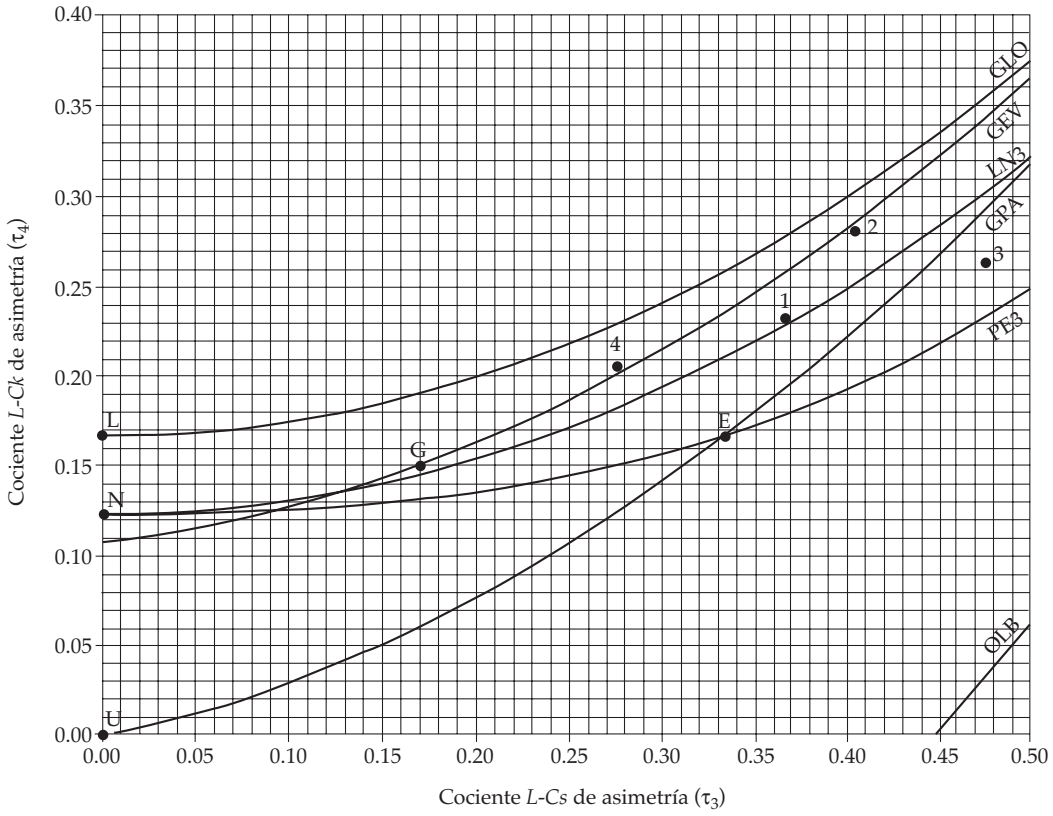
Test de Wiltshire (tercera prueba)

Basado en el coeficiente de variación (*CV*) de cada sitio *j* de la región estudiada, la cual incluye *N* sitios, se define como el cociente entre la desviación estándar del sitio y su media, es decir:

$$CV_j = \frac{S_j}{\bar{x}_j} \qquad (8)$$

Designando nuevamente por n_j la longitud del registro de datos en el sitio *j*, se establece como U_j la variancia muestral del CV_j:

Distribuciones de dos y tres parámetros de ajuste son mostradas como puntos y líneas, respectivamente



Simbología:

L	Logística	GLO	Logística generalizada
N	Normal	GEV	General de valores extremos
U	Uniforme	LN3	Log-normal de tres parámetros
G	Gumbel	GPA	Pareto generalizada
E	Exponencial	PE3	Pearson tipo III
		OLB	Límite inferior de τ_4 como función de τ_3

Figura 1. Diagrama de cocientes de momentos *L*.

$$U_j = \frac{V_R}{n_j} \tag{9}$$

donde V_R es la variancia regional del *CV*; se estima con la ecuación (13). La variación total del *CV* dentro de la región de *N* sitios se designa por el estadístico *S*, que tiene distribución χ^2 con $\nu = N - 1$ grados de libertad (Wiltshire, 1986a):

$$S = \sum_{j=1}^{N} \frac{\left(CV_j - CV_p\right)^2}{U_j} \tag{10}$$

siendo CV_p el valor ponderado del *CV* de la región, es decir:

$$CV_p = \frac{\sum_{j=1}^{N} CV_j / U_j}{\sum_{j=1}^{N} 1/U_j} \tag{11}$$

La variancia regional V_R de la ecuación (9) puede ser calculada como el promedio de las variancias truncadas calculadas en cada

sitio de la región. Entonces, siendo x_1, x_2,..., x_n los n_j datos independientes e idénticamente distribuidos de cada sitio, el $CV_{n\text{-}1}$ será el coeficiente de variación calculado de una muestra de tamaño $n_j - 1$, consistente en los x_1, x_2, x_{i-1}, x_{i+1},..., x_{nj} datos, es decir una serie con la i-ésima observación suprimida.

Entonces, la estimación de la variancia truncada de CV_j es la variancia del conjunto de n_j valores de CV_{n-1}, en los cuales cada valor ha sido calculado de las series con una observación diferente removida (Wiltshire, 1986b), es decir:

$$v_j = \frac{\left(n_j - 1\right)}{n_j} \cdot \sum_{i=1}^{N} \left[CV_{n-1}^{i} - \left(\frac{1}{n_j} \sum_{l=1}^{N} CV_{n-1}^{l} \right) \right]^2 \quad (12)$$

entonces la variancia regional V_R a través de los N sitios será:

$$V_R = \frac{1}{N} \sum_{j=1}^{N} n_j \cdot v_j \quad (13)$$

Los valores críticos de χ^2 con un nivel de significancia del 5% (95% de confiabilidad) para los grados de libertad empleados en los ARFC se muestran en el cuadro 2.

Finalmente, si S es menor que el valor crítico χ_c^2, la región será homogénea; en caso contrario, no homogénea (Wiltshire, 1986a, 1986b; Rao y Hamed, 2000).

Aplicaciones

Primera aplicación

Se realizó en la Región Hidrológica Número 10, utilizando once registros de gastos máximos anuales (m^3/s) obtenidos en sus estaciones de aforos de menor área de cuenca, de manera que éstas fluctuaron de 223 a 1 645 km^2. La información hidrométrica se obtuvo del sistema *BANDAS* (IMTA, 2003), con datos actualizados hasta el año 2002; sus características generales se pueden consultar en Campos (2008b). La prueba de Discordancia no detectó anomalías en los datos. Los cocientes de momentos L regionales (cuadro 3) determinan en la figura 1 como FDP más conveniente la log-normal de tres parámetros, por ello las simulaciones de la prueba H se realizaron con base en la FDP Wakeby ajustada por momentos L (Hosking y Wallis, 1997). Los resultados de las pruebas H y de Wiltshire se tienen en el cuadro 3, indicando que tales estaciones forman una región homogénea.

Cuadro 2. Valores críticos del estadístico χ^2 con $\alpha = 5\%$, según el número de grados de libertad (ν).

ν	χ_c^2	ν	χ_c^2	ν	χ_c^2	ν	χ_c^2
6	12.6	10	18.3	14	23.7	18	28.9
7	14.1	11	19.7	15	25.1	19	30.1
8	15.5	12	21.0	16	26.3	20	31.4
9	16.9	13	22.4	17	27.6	21	32.7

Cuadro 3. Resultados de las pruebas estadísticas de homogeneidad regional (H) y del test de Wiltshire (S) en las cuatro aplicaciones numéricas indicadas.

Núm. de aplicación	Prueba H								Prueba S			
	t^R	t_3^R	t_4^R	FDP	V	μ_V	σ_V	H	V_R	CV_p	S	χ_c^2
Primera	0.45794	0.36860	0.23138	WAK	0.06437	0.06268	0.01624	0.104	2.30741	0.95965	12.269	18.3
Segunda	0.43848	0.40435	0.28014	GVE	0.14439	0.04865	0.01713	5.589	1.46242	0.96523	37.269	12.6
Tercera	0.62103	0.47540	0.26270	PAG	0.08832	0.07090	0.01987	0.877	2.68344	1.31025	3.260	14.1
Cuarta	0.23336	0.27519	0.20546	GVE	0.03284	0.03243	0.00438	0.094	0.20320	0.44214	30.748	47.4

Segunda aplicación

En la cuenca del río Guayalejo, que pertenece a la Región Hidrológica Número 26 (Pánuco), existen diez estaciones hidrométricas, cuyos datos de gastos máximos anuales (m^3/s) fueron recabados del sistema *BANDAS* (IMTA, 2003), con datos hasta el 2002. Con base en la prueba de homogeneidad de Langbein se encontró que siete de tales hidrométricas forman una región homogénea. Los registros de estas estaciones se pueden consultar en Campos (2006). La aplicación de la prueba de Discordancia se muestra en el cuadro 4, indicando que ningún registro es discordante con el resto, pues no exceden el valor crítico de 1.917; sin embargo se observa que sus valores de Discordancia son elevados. Con base en los valores de los cocientes de momentos *L* regionales (cuadro 3) se definió en la figura 1 como distribución idónea la GVE, con la cual se efectuó la simulación de las 500 regiones homogéneas. Los resultados de la aplicación de las pruebas *H* y de Wiltshire se presentan en el cuadro 3, mostrando que la región es no homogénea.

Tercera aplicación

De acuerdo con el sistema *BANDAS* (IMTA, 2003), en la Región Hidrológica Número 25 (San Fernando-Soto La Marina) existen 14 estaciones de aforos que realizan muestreo de sedimentos, reportado como volumen de sólidos en suspensión anual (miles de m^3), algunas con datos hasta 1999 y características generales que se pueden consultar en Campos (2005). Los resultados de la prueba de Discordancia en las ocho estaciones hidrométricas que fueron procesadas como región homogénea se

Cuadro 4. Valores de los cocientes de momentos y de las discordancias en la cuenca del río Guayalejo, en Tamaulipas.

Estación de aforos	t	t_3	t_4	D_i
Mante	0.3808	0.3857	0.3882	1.04
Sabinas	0.2947	0.3102	0.1993	1.16
La Servilleta	0.3379	0.2650	0.2397	1.30
La Encantada	0.6139	0.6006	0.4134	1.11
San Gabriel II	0.6503	0.5638	0.2822	1.15
Magiscatzin II	0.3941	0.3241	0.1549	0.72
Tamesí	0.2480	0.2506	0.2238	0.51
Valor promedio	0.4170	0.3857	0.2716	Dc= 1.917

Cuadro 5. Valores de los cocientes de momentos y de las discordancias en la Región Hidrológica Número 25 (San Fernando-Soto La Marina).

Estación de aforos	t	t_3	t_4	D_i
Padilla II	0.6336	0.4412	0.1659	0.84
Paso del Aura	0.7121	0.6039	0.3706	1.28
Pablillo	0.5201	0.3669	0.1996	0.65
El Barretal II	0.7588	0.5628	0.1935	0.90
Barberena	0.7438	0.5924	0.2964	0.54
Cabezones	0.5908	0.4275	0.1431	1.73
Camacho	0.4927	0.4013	0.3428	1.15
Magueyes	0.6054	0.4788	0.3401	0.91
Valor promedio	0.6321	0.4843	0.2565	Dc = 2.140

exponen en el cuadro 5, indicando que ningún registro es anómalo.

Los cocientes de momentos *L* regionales (cuadro 3) definen en la figura 1 como FDP más conveniente la Pareto Generalizada, cuyo método de ajuste se describe en Hosking y Wallis (1997). Las simulaciones de la prueba *H* se realizaron con este modelo y sus resultados, mostrados en el cuadro 3, indican que los ocho registros de sedimentos procesados forman una región homogénea. A igual conclusión se llega con la prueba de Wiltshire.

Cuarta aplicación

En la zona Huasteca del estado de San Luis Potosí, la cual pertenece a la Región Hidrológica Número 26 (Pánuco), se localizan 34 estaciones pluviométricas, cuyos registros de precipitación máxima diaria anual tomados del sistema *ERIC II* (IMTA, 2000) abarcan desde 32 hasta 40 años, con las características generales que se pueden consultar en Campos (2008a). La prueba de Discordancia no detecta registros anómalos. Los valores regionales de los cocientes de momentos *L* (cuadro 3) conducen en la figura 1 a la distribución GVE como la más conveniente para realizar las simulaciones requeridas en la prueba *H*. Los resultados de esta prueba y del test de Wiltshire se presentan en el cuadro 3, los cuales indican que la región formada por los 34 registros procesados es homogénea.

Análisis de resultados

De manera general, los resultados de las tres pruebas estadísticas descritas y aplicadas en cuatro regiones consideradas originalmente como homogéneas son consistentes. Los resultados de las pruebas *H* y del test de Wiltshire en la tercera aplicación numérica demuestran que la prueba de Langbein propuesta por Dalrymple (1960) no es confiable, como lo han demostrado Fill y Stedinger (1995).

Conclusiones

Los resultados de las tres pruebas estadísticas descritas, aplicadas en diversas regiones consideradas originalmente homogéneas, aquí descritas únicamente cuatro, de manera general son coincidentes y se complementan para decidir sobre la homogeneidad de un grupo de sitios; por ello se recomienda su aplicación sistemática.

Con respecto a la prueba *H*, se recomienda llevar a cabo las simulaciones de las 500 regiones homogéneas con base en la distribución Wakeby y la definida en la figura 1 como idónea según los cocientes (t_3 y t_4) de momentos *L* regionales, sobre todo cuando el valor del estadístico *H* resultó cercano a la unidad.

Apéndice. Momentos *L* muestrales

Son un sistema alternativo para describir las formas de las FDP. Los momentos *L* son combinaciones lineales de los momentos de probabilidad pesada desarrollados por Greenwood *et al*. (1979), de manera que (Hosking y Wallis, 1997):

$$\lambda_1 = \beta_0 \qquad \text{(A.1)}$$

$$\lambda_2 = 2\cdot\beta_1 - \beta_0 \qquad \text{(A.2)}$$

$$\lambda_3 = 6\cdot\beta_2 - 6\cdot\beta_1 + \beta_0 \qquad \text{(A.3)}$$

$$\lambda_4 = 20\cdot\beta_3 - 30\cdot\beta_2 + 12\cdot\beta_1 - \beta_0 \qquad \text{(A.4)}$$

$$\lambda_5 = 70\cdot\beta_4 - 140\cdot\beta_3 + 90\cdot\beta_2 - 20\cdot\beta_1 + \beta_0 \qquad \text{(A.5)}$$

Además, se definen los cocientes (τ) de momentos *L*, comenzando con *L*–*Cv*, que es análogo a este coeficiente, y después los de similitud con los coeficientes de asimetría y de curtosis, que son:

$$\tau = \lambda_2 / \lambda_1 \qquad \text{(A.6)}$$

$$\tau_3 = \lambda_3 / \lambda_2 \qquad \text{(A.7)}$$

$$\tau_4 = \lambda_4 / \lambda_2 \qquad \text{(A.8)}$$

En una muestra de tamaño n, con sus elementos arreglados en orden ascendente ($x_1 \leq x_2 \leq ... \leq x_n$), los estimadores insesgados de β_r son:

$$b_0 = \frac{1}{n}\sum_{j=1}^{n} x_j \qquad \text{(A.9)}$$

$$b_1 = \frac{1}{n}\sum_{j=2}^{n} \frac{(j-1)}{(n-1)} x_j \qquad \text{(A.10)}$$

$$b_2 = \frac{1}{n}\sum_{j=3}^{n} \frac{(j-1)(j-2)}{(n-1)(n-2)} x_j \qquad \text{(A.11)}$$

con la expresión general siguiente:

$$b_r = \frac{1}{n}\sum_{j=r+1}^{n} \frac{(j-1)(j-2)\ldots(j-r)}{(n-1)(n-2)\ldots(n-r)} x_j \qquad \text{(A.12)}$$

Los estimadores muestrales de λ_r serán l_r, estando definidos por las ecuaciones (A.1) a (A.5), y los de los cocientes serán t, t_3 y t_4, según las ecuaciones (A.6) a (A.8).

Recibido: 01/06/2009
Aprobado: 02/06/2010

Referencias

CAMPOS-ARANDA, D.F. Predicciones de volúmenes de sólidos en suspensión en cuencas sin aforos en la Región Hidrológica No. 25 (San Fernando-Soto La Marina). *Tláloc*. Núm. 33, enero-abril, 2005, pp. 22-28.

CAMPOS-ARANDA, D.F. Contraste de métodos regionales de estimación de crecientes en la cuenca del río Guayalejo, en Tamaulipas. *Tláloc*. Núm. 37, mayo-agosto, 2006, pp. 14-24.

CAMPOS-ARANDA, D.F. Ajuste regional de la distribución *GVE* en 34 estaciones pluviométricas de la zona Huasteca de San Luis Potosí. *Agrociencia*. Vol. 42, núm. 1, 2008a, pp. 57-70.

CAMPOS-ARANDA, D.F. Calibración del método Racional en ocho cuencas rurales menores de 1,650 km^2 de la Región Hidrológica No. 10 (Sinaloa), México. *Agrociencia*. Vol. 42, núm. 6, 2008b, pp. 615-627.

CAMPOS-ARANDA, D.F. Aplicación de la prueba de Discordancias a las crecientes de la costa de Chiapas. Tema: Hidrología superficial y subterránea, ponencia 5. *Memorias del XX Congreso Nacional de Hidráulica*. 15 al 18 de octubre, Toluca, Estado de México, México, 2008c.

DALRYMPLE, T. Flood-Frequency Analyses. *Manual of Hydrology*. Part 3: Flood-Flow Techniques. Washington, D.C.: U.S. Geological Survey. Water-Supply Paper 1543-A, 1960.

FILL, H.D. and STEDINGER, J.R. Homogeneity test based upon Gumbel distribution and a critical appraisal of Dalrymple's test. *Journal of Hydrology*. Vol. 166, 1995, pp. 81-105.

GREENWOOD, J.A., LANWEHR, J.M., MATALAS, N.C. and WALLIS, J.R. Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research*. Vol. 15, 1979, pp. 1049-1054.

HOSKING, J.R.M. and WALLIS, J.R. *Regional Frequency Analysis. An approach based on L–moments*. Cambridge: Cambridge University Press, 1997, 224 pp.

IMTA. *ERIC II: Extractor Rápido de Información Climatológica 1920-1998*. 1 CD. Jiutepec, México: Comisión Nacional del Agua, Secretaría de Medio Ambiente y Recursos Naturales, Instituto Mexicano de Tecnología del Agua, 2000.

IMTA. *Banco Nacional de Datos de Aguas Superficiales (BANDAS)*. 8 CD's. Jiutepec, México: Comisión Nacional del Agua, Secretaría de Medio Ambiente y Recursos Naturales, Instituto Mexicano de Tecnología del Agua, 2003.

METCALFE, A.V. Probability distributions and Monte Carlo simulation (chapter 2), and Random number generation (appendix 2). In *Statistics in Civil Engineering*. London: Arnold Publishers, 1997, pp. 7-38, 319-320.

RAO, A.R. and HAMED, K.H. Regional homogeneity and regionalization (chapter 2, theme 2.5). In *Flood Frequency Analysis*. Boca Raton, USA: CRC Press LLC, 2000, pp. 47-52.

WILTSHIRE, S.E. Regional flood frequency analysis I: Homogeneity statistics. *Hydrological Sciences Journal*. Vol. 31, 1986a, pp. 321-333.

WILTSHIRE, S.E. Identification of homogeneous regions for flood frequency analysis. *Journal of Hydrology*. Vol. 84, 1986b, pp. 287-302.

Abstract

CAMPOS-ARANDA, D.F. Regional homogeneity verification through three statistical tests. Water Technology and Sciences, *formerly* Hydraulic engineering in Mexico *(in Spanish). Vol. I, No. 4, October-December, 2010, pp. 157-165.*

First, general concepts about regional flood frequency analysis are presented. Next, the Discordancy Test for the detection of records with anomalous data that are discordant with the group as a whole is described in detail. Two more tests are presented: the heterogeneity H test based on the simulation of 500 homogeneous regions and the Wiltshire Test, which used the local and regional coefficient of variation to establish a statistic value with χ^2 distribution. Then, the three statistical tests cited are used in four originally accepted homogeneous regions and the results are discussed. Lastly, conclusions are formulated, proposing the systematic application of the three tests to decide about the homogeneity of the studied region.

Keywords: *L moments, discordancy, simulation, H Test, Wiltshire Test.*

Dirección institucional del autor

Dr. Daniel Francisco Campos-Aranda

Profesor jubilado de la Universidad de San Luis Potosí
Genaro Codina 240, colonia Jardines del Estadio
78280 San Luis Potosí, San Luis Potosí, México
teléfono: +52 (444) 8151 431
campos_aranda@hotmail.com