

Ajuste de la distribución no estacionaria GVE_{11} a través de momentos L

Fitting of the non-stationary distribution GEV_{11} through L moments

Daniel Francisco Campos-Aranda¹

¹ Profesor jubilado de la Universidad Autónoma de San Luis Potosí, San Luis Potosí, México, campos_aranda@hotmail.com

Autor para correspondencia: Daniel Francisco Campos-Aranda, campos_aranda@hotmail.com

Resumen

El dimensionamiento o la revisión hidrológica de las obras hidráulicas y la elaboración de mapas de riesgo por inundación se realizan con base en las llamadas *crecientes de diseño*, que son gastos máximos del río asociados con bajas probabilidades de excedencia. La manera más confiable de estimar tales *predicciones* es a través del Análisis de Frecuencias de Crecientes (AFC), cuya suposición fundamental es que el proceso estocástico bajo estudio es *estacionario*, es decir, que no cambia

con el tiempo. La construcción de embalses pequeños, la urbanización y los cambios de uso del suelo en la cuenca, así como el cambio climático global o regional alteran los procesos hidrológicos y generan registros de crecientes anuales que son *no estacionarios*, al mostrar tendencias y cambio en su variabilidad. Para el AFC de tales registros se ha extendido la teoría de valores extremos, a fin de aplicar su distribución clásica, la General de Valores Extremos (GVE), con parámetros de ubicación (u) y escala (a) variables con el tiempo (t), que se introduce como *covariable*. En este trabajo se expone el método de momentos L para el ajuste del modelo probabilístico GVE₁₁, cuyos parámetros u y a varían linealmente con el tiempo. Se describen tres aplicaciones numéricas. Las conclusiones destacan la sencillez del método expuesto y su importancia en la estimación de las predicciones buscadas en series de datos máximos anuales no estacionarios.

Palabras clave: covariable, momentos L, distribución GVE, error estándar de ajuste, tendencia, regresión lineal, residuales.

Abstract

Hydrological dimensioning or revision of the hydraulic works and the elaboration of flood risk maps, is based on the so-called *design floods*, that are maximum flows of the river associated with low probabilities of exceedance. The most reliable way of estimating such *predictions* is through the Flood Frequency Analysis (FFA), its fundamental assumption is that the stochastic process under study is *stationary*, that is, it does not change with time. Construction of small reservoirs, urbanization and changes in land use in the basin, as well as global or regional climate

change, alter the hydrological processes and generate records of annual flows that are *non-stationary*, showing trends and changes in variability. For the FFA of such registries, the extreme value theory has been extended to apply its classical distribution, the General of Extreme Values (GEV), with parameters of location (u) and scale (a) varying with time (t), introducing it as a *covariate*. In this work the L moments method is presented for the fit of the probabilistic model GVE₁₁ whose parameters u and a vary linearly over time. Three numerical applications are described. Conclusions highlight the simplicity of the exposed method and its importance in the estimation of the sought predictions in non-stationary annual maximum data series.

Keywords: Covariate, L moments, GEV distribution, standard error of fit, trend, linear regression, residuals.

Recibido: 28/08/2019

Aceptado: 11/08/2020

Introducción

La estimación del riesgo de inundaciones de una cierta área y el dimensionamiento hidrológico de las obras hidráulicas, como embalses,

puentes, muros y diques de protección y canalizaciones, se realiza con base en las llamadas *crecientes de diseño*, que son gastos extremos del río asociados con bajas probabilidades de excedencia. La estimación más confiable de las crecientes de diseño se lleva a cabo a través del Análisis de Frecuencias de Crecientes (AFC) del registro o serie disponible de gastos máximos anuales. El AFC asume que el proceso aleatorio que genera tales crecientes es *estacionario* y por ello sus propiedades estadísticas no cambian a lo largo del tiempo (Khaliq, Ouarda, Ondo, Gachon, & Bobée, 2006; El-Adlouni, Ouarda, Zhang, Roy, & Bobée, 2007; El-Adlouni & Ouarda, 2008).

Todos los cambios físicos que ocurren en las cuencas de drenaje, como son la construcción de embalses, la urbanización, la rectificación de cauces, la deforestación y el cambio de uso del suelo, alteran los procesos hidrológicos, dando origen a series o registros de gastos máximos anuales *no estacionarios*, ya que presentan tendencias y cambios en la variabilidad. Aunado a lo anterior, los impactos del cambio climático global pueden exacerbar las alteraciones de los procesos hidrológicos, favoreciendo la formación de series de datos hidrológicos extremos no estacionarias. Desde comienzos de la década de 1990, la no estacionariedad debida a la urbanización y/o al cambio climático fue tomada en cuenta, corrigiendo el registro disponible, aplicando factores de incremento a las predicciones, o bien realizando predicciones en la cuenca rural y luego ajustando los resultados por su grado de urbanización (Jakob, 2013; Prosdocimi, Kjeldsen, & Miller, 2015).

Actualmente, la teoría estadística de valores extremos ha sido extendida a las condiciones no estacionarias para ajustar la distribución

clásica que siguen asintóticamente las series de datos hidrológicos máximos, es decir el modelo probabilístico *General de Valores Extremos* estacionario (GVE_0) con tres parámetros de ajuste (u, a, k). Lo anterior con base en la introducción de las llamadas *covariables*. Una de las más utilizadas es el tiempo (t) en años, a través de la cual se puede tomar en cuenta la tendencia observada en la serie de datos, adoptando variable el parámetro de ubicación (u_t). Cuando la tendencia es lineal, $u_t = \mu_0 + \mu_1 \cdot t$ y entonces se ajusta el modelo GVE_1 de cuatro parámetros de ajuste (μ_0, μ_1, a, k). Si la tendencia es curva, $u_t = \mu_0 + \mu_1 \cdot t + \mu_2 \cdot t^2$ y se aplica el modelo GVE_2 de cinco parámetros de ajuste ($\mu_0, \mu_1, \mu_2, a, k$). En el modelo GVE_{11} tanto el parámetro de ubicación como el de escala ($a_t = \sigma_0 + \sigma_1 \cdot t$) son funciones lineales del tiempo t y entonces se toma en cuenta, además de la tendencia, el cambio a través del tiempo de la variabilidad de la serie (Khaliq *et al.*, 2006; El-Adlouni & Ouarda, 2008; Aissaoui-Fqayeh, El-Adlouni, Ouarda, & St. Hilaire, 2009; Jakob, 2013; Gado & Nguyen 2016).

Khaliq *et al.* (2006), y Meylan, Favre y Musy (2012) han expuesto una revisión de los diversos métodos utilizados en el AFC, en el contexto no estacionario. Tales técnicas incluyen la incorporación de tendencias en los parámetros de la distribución o en los momentos estadísticos, el método del cuantil de la regresión, los procesos de verosimilitud local y el enfoque Bayesiano. En los modelos probabilísticos del primer enfoque, el uso de *covariables* pretende integrar los cambios que han ocurrido en el pasado directamente en las técnicas del AFC, con el objeto de extrapolarlos al futuro. En este contexto, los modelos que usan el tiempo como covariable, caso del aquí expuesto, son simples y útiles para

reproducir el comportamiento no estacionario de un registro de crecientes, pero la aceptación de sus predicciones no es plenamente correcta, ya que como indican López y Francés (2013), las tendencias pueden cambiar en el corto o largo plazos debido a la variabilidad climática o a la intensificación de las actividades humanas.

Debido a lo anterior, en las dos últimas décadas se ha explorado la posibilidad de incorporar índices climáticos como forzamientos externos en los modelos del AFC, asumiendo dependencia lineal o no lineal. Este enfoque ha mostrado que el uso de tales *covariables* conduce a modelos que describen mejor los cambios temporales ocurridos en los registros de crecientes. En realidad, se ha encontrado una teleconexión entre los cambios observados en los regímenes de crecientes y las anomalías en los índices (ENSO: *El Niño-Southern Oscillation*; PDO: *Pacific Decadal Oscillation*, and NAO: *North Atlantic Oscillation*) que describen la evolución temporal de los patrones de circulación atmosférica de baja frecuencia (Khaliq *et al.*, 2006; López-de-la-Cruz & Francés, 2014).

El *objetivo* de este trabajo consistió en exponer la teoría del ajuste de la distribución GVE a las condiciones no estacionarias por medio de la generalización del método de los momentos L, planteada por Gado y Nguyen (2016) para el modelo GVE₁₁. Se exponen y procesan tres series de datos hidrológicos extremos tomadas de la literatura especializada y se discuten sus resultados, destacando la sencillez y utilidad de método de los momentos L para obtener *predicciones* dentro del registro histórico y a futuro, cuando ellas son importantes, debido a la tendencia ascendente de la serie y/o su variabilidad en aumento.

Teoría operativa

Distribución General de Valores Extremos

Los valores extremos anuales son eventos raros que ocurren en la extremidad o cola derecha de la función de distribución de probabilidades (FDP), que es el modelo probabilístico que define el comportamiento de la variable aleatoria hidrológica bajo estudio, como son crecientes y valores máximos de lluvias, temperaturas, vientos y niveles del mar. Los valores extremos se pueden *predecir* con base en la FDP, como el máximo de la variable aleatoria que corresponde a un cierto intervalo *promedio* de recurrencia o periodo de retorno (Tr), cuya probabilidad de excedencia es $p = 1 / Tr$.

La teoría de valores extremos justifica y establece que los datos extremos siguen de manera *asintótica* alguno de los tres tipos de distribuciones denominadas Gumbel, Fréchet o Weibull (Clarke, 1973; Stedinger, Vogel, & Foufoula-Georgiou, 1993; Coles, 2001; Khaliq *et al.*, 2006). Estos tres modelos probabilísticos se pueden representar en uno solo, denominado *distribución General de Valores Extremos* (GVE), cuya aplicación ha sido ampliamente recomendada para modelar gastos

máximos anuales (q) y otros datos extremos (Hosking & Wallis, 1997; Papalexiou & Koutsoyiannis, 2013; Gado & Nguyen, 2016; Stedinger, 2017). La FDP de la GVE con una probabilidad de no excedencia [$F(q)$] es:

$$F(q) = \exp\left\{-\left[1 - \frac{k(q-u)}{\alpha}\right]^{1/k}\right\} \text{ cuando } k \neq 0 \quad (1)$$

En la expresión anterior, u , a y k son los parámetros de ubicación, escala y forma de la distribución GVE. Cuando $k = 0$, se obtiene la distribución Gumbel, que es una línea recta en la papel de probabilidad Gumbel-Powell (Chow, 1964), por lo cual el intervalo de la variable es $-\infty < q < \infty$. Cuando $k > 0$, la distribución es Weibull, que es una curva con concavidad hacia abajo y límite superior, por lo cual $-\infty < q \leq u + a / k$. Por último, si $k < 0$, la distribución es Fréchet, que también es una curva, pero con concavidad hacia arriba y frontera inferior, por lo que $u + a / k \leq q < \infty$. Las *predicciones* buscadas (Q_{Tr}) se obtienen con la solución inversa de la Ecuación (1):

$$Q_{Tr} = u + \frac{\alpha}{k} \{1 - [-\ln(1 - p)]^k\} \text{ cuando } k \neq 0 \quad (2)$$

Momentos L de la muestra de datos

El método de los momentos L es quizás el más simple y se ha convertido en uno de los procedimientos confiables para estimar los parámetros de ajuste de las FDP utilizadas en hidrología. Lo anterior debido a que los momentos L, que son combinaciones lineales (Hosking & Wallis, 1997) de los momentos de probabilidad ponderada (β_r), no son afectados de manera significativa por los valores dispersos (*outliers*) de la muestra. Los primeros tres momentos L de una muestra (l_1, l_2, l_3) y el cociente L de asimetría (t_3) se estiman a través del estimador insesgado (b_r) de los β_r , como sigue:

$$l_1 = b_0 \quad (3)$$

$$l_2 = 2 \cdot b_1 - b_0 \quad (4)$$

$$l_3 = 6 \cdot b_2 - 6 \cdot b_1 + b_0 \quad (5)$$

$$t_3 = l_3/l_2 \quad (6)$$

El estimador insesgado de los β_r es (Hosking & Wallis, 1997):

$$b_r = \frac{1}{n} \sum_{j=r+1}^n \frac{(j-1)(j-2)\dots(j-r)}{(n-1)(n-2)\dots(n-r)} q_j \quad (7)$$

donde $r = 0, 1, 2, \dots$ y q_j son los datos de la muestra o serie disponible de tamaño n , ordenados de menor a mayor ($q_1 \leq q_2 \leq \dots \leq q_n$).

Parámetros de ajuste de la distribución GVE estacionaria

Para el modelo GVE_0 con el método de los momentos L se obtienen las ecuaciones siguientes para sus tres parámetros de ajuste (Stedinger *et al.*, 1993; Hosking & Wallis, 1997; Rao & Hamed, 2000; Stedinger, 2017):

$$k \cong 7.8590 \cdot c + 2.9554 \cdot c^2 \quad (8)$$

siendo:

$$c = \frac{2}{3+t_3} - 0.63093 \quad (9)$$

$$\alpha = \frac{l_2 \cdot k}{(1-2^{-k}) \cdot \Gamma(1+k)} \quad (10)$$

$$u = l_1 - \frac{\alpha}{k} [1 - \Gamma(1+k)] \quad (11)$$

Para la estimación de la función Gamma $\Gamma(\omega)$ se utilizó la fórmula de Stirling (Davis, 1972):

$$\Gamma(\omega) = e^{-\omega} \cdot \omega^{\omega-1/2} \cdot \sqrt{2\pi} \cdot \left(1 + \frac{1}{12 \cdot \omega} + \frac{1}{288 \cdot \omega^2} - \frac{139}{51840 \cdot \omega^3} - \frac{571}{2488320 \cdot \omega^4} + \dots\right) \quad (12)$$

Planteamiento del análisis no estacionario

Tiene dos suposiciones inherentes. La primera acepta que la no *estacionariedad* de las series hidrológicas de valores anuales extremos es causada por cambios graduales del uso del suelo o por el cambio climático global o de escala regional, generando una alteración ligera de sus parámetros estadísticos. El segundo supuesto acepta que la FDP es independiente del tiempo, entonces la distribución GVE, con parámetros de ajuste variables con el tiempo, es aceptable para modelar datos extremos no estacionarios.

Debido a la no estacionariedad existe gran variabilidad estadística asociada con la muestra o serie disponible de datos anuales extremos y por ello conviene adoptar que la media (μ) y la desviación estándar (σ) son funciones del tiempo. La manera más simple se realiza por medio del modelo probabilístico GVE₁₁ con variaciones lineales, por lo cual $\mu_t = \mu_0 +$

$\mu_1 \cdot t$ y $\ln(\sigma_t) = \sigma_0 + \sigma_1 \cdot t$. En las expresiones anteriores, t es la *covariable* tiempo en años que abarca la serie de datos de tamaño n . El uso de los logaritmos naturales de σ asegura valores positivos del parámetro de escala (α) en la distribución GVE (Gado & Nguyen, 2016). Las expresiones generales de la media y de la desviación estándar del modelo GVE₀, obtenida la segunda a través del segundo momento central, son (Rao & Hamed, 2000):

$$\mu = u + \frac{\alpha}{k} [1 - \Gamma(1 + k)] \quad (13)$$

$$\sigma^2 = \frac{\alpha^2}{k^2} [\Gamma(1 + 2 \cdot k) - \Gamma^2(1 + k)] \quad (14)$$

Ajuste de la distribución GVE₁₁

Como la media y el logaritmo natural de la desviación estándar son funciones lineales del tiempo, primero se introduce la expresión $\mu_t = \mu_0 + \mu_1 \cdot t$ en la Ecuación (13) para despejar la expresión de parámetro de ubicación variable con el tiempo (u_t), en la cual también es variable el parámetro de escala (α_t) y donde FK_1 es el primer factor función del parámetro de forma (k). Tal ecuación es (Gado & Nguyen, 2016):

$$u_t = \mu_0 + \mu_1 \cdot t - \frac{\alpha_t}{k} [1 - \Gamma(1+k)] = \mu_0 + \mu_1 \cdot t - FK_1 \cdot \alpha_t \quad (15)$$

en donde:

$$FK_1 = \frac{1 - \Gamma(1+k)}{k} \quad (16)$$

De la Ecuación (14) se obtiene la expresión de parámetro de escala variable con el tiempo:

$$\alpha_t^2 = \frac{k^2 \cdot \sigma_t^2}{\Gamma(1+2 \cdot k) - \Gamma^2(1+k)} \quad (17)$$

Por lo cual:

$$\alpha_t = FK_2 \cdot \sigma_t = FK_2 \cdot \exp(\sigma_0 + \sigma_1 \cdot t) \quad (18)$$

siendo:

$$FK_2 = \sqrt{\frac{k^2}{\Gamma(1+2 \cdot k) - \Gamma^2(1+k)}} \quad (19)$$

Sustituyendo la Ecuación (18) en la Ecuación (15) se obtiene la expresión final del parámetro de ubicación variable con el tiempo:

$$u_t = \mu_0 + \mu_1 \cdot t - FK_1 \cdot FK_2 \cdot \exp(\sigma_0 + \sigma_1 \cdot t) \quad (20)$$

Las magnitudes μ_0 y μ_1 de la recta que representa la tendencia lineal de la muestra o serie de datos q_t se obtienen con base en las ecuaciones de la recta de regresión lineal de la manera siguiente: se considera que la variable dependiente (y) son los datos hidrológicos máximos anuales q_t y los tiempos o años t son las abscisas (x), en este caso iguales al i -ésimo valor de i . La pendiente (μ_1) de la recta de regresión ajustada por mínimos cuadrados de los residuos y la ordenada al origen (μ_0) se obtienen con las ecuaciones siguientes (Campos-Aranda, 2003):

$$q_t = \mu_0 + \mu_1 \cdot t \quad (21)$$

$$\mu_1 = \frac{\text{cov}(q,t)}{\text{var}(t)} = \frac{\frac{1}{n} \sum_{i=1}^n q_i i - \bar{q} \cdot \bar{t}}{\frac{1}{n} \sum_{i=1}^n i^2 - \bar{t}^2} \quad (22)$$

$$\mu_0 = \bar{q} - \mu_1 \cdot \bar{t} \quad (23)$$

En las ecuaciones anteriores, \bar{q} y \bar{t} son las medias aritméticas de la serie de datos y del tiempo. El coeficiente de correlación lineal (r_{xy}) mide el grado de dependencia o asociación entre las variables q y t , y varía de cero a uno, indicando con la unidad la regresión perfecta. Su ecuación es:

$$r_{xy} = \frac{\text{cov}(q,t)}{\sqrt{\text{var}(t) \cdot \text{var}(q)}} \quad (24)$$

siendo:

$$\text{var}(q) = \frac{1}{n} \sum_{i=1}^n q_i^2 - \bar{q}^2 \quad (25)$$

Para estimar la tendencia lineal en la desviación estándar (σ_t), primero se obtienen los residuales (ε_t) con la ecuación siguiente (Cunderlik & Burn, 2003; Gado & Nguyen, 2016):

$$\varepsilon_t = q_t - \mu_1 \cdot t \quad (26)$$

Se obtiene su valor medio ($\bar{\varepsilon}$) y se calculan las desviaciones absolutas (da) de los residuales, con la expresión siguiente:

$$\varepsilon_t^{da} = |\varepsilon_t - \bar{\varepsilon}| \quad (27)$$

Se aplican logaritmos naturales a ε_t^{da} y se representan por una línea recta, cuya ecuación es:

$$\ln(\varepsilon_t^{da}) = \sigma_0 + \sigma_1 \cdot t \quad (28)$$

Los valores de σ_1 y σ_0 se calculan con base en las ecuaciones (22) y (23). El grado de asociación entre $\ln(\varepsilon_t^{da})$ y t se estima con la Ecuación (24) y se designa r_y , ya que la serie y es el lado izquierdo de la Ecuación (28). Para obtener la serie estacionaria sin tendencias en la media y en la desviación estándar ($q_e^{\mu,\sigma}$) se aplican las ecuaciones siguientes (Cunderlik & Burn, 2003; Khaliq *et al.*, 2006; Gado & Nguyen, 2016):

$$q_e^{\mu,\sigma} = \varepsilon_t - \sigma_t \text{ si } \varepsilon_t \geq \bar{\varepsilon} \text{ cuando } \sigma_t > 0 \quad (29)$$

$$q_e^{\mu,\sigma} = \varepsilon_t + \sigma_t \text{ si } \varepsilon_t < \bar{\varepsilon} \text{ cuando } \sigma_t > 0 \quad (30)$$

$$q_e^{\mu,\sigma} = \varepsilon_t + \sigma_t \text{ si } \varepsilon_t \geq \bar{\varepsilon} \text{ cuando } \sigma_t < 0 \quad (31)$$

$$q_e^{\mu,\sigma} = \varepsilon_t - \sigma_t \text{ si } \varepsilon_t < \bar{\varepsilon} \text{ cuando } \sigma_t < 0 \quad (32)$$

En las expresiones anteriores, $\sigma_t = \exp(\sigma_0 + \sigma_1 \cdot t)$. Se aplican las ecuaciones (9) y (8) a la serie de $q_e^{\mu,\sigma}$ para estimar el parámetro de forma (k) de la distribución GVE_0 estacionaria. Por último, para realizar las predicciones buscadas se aplica la Ecuación (2), utilizando en ella las expresiones de los parámetros de ubicación y escala variables con el tiempo, ecuaciones (20) y (18), respectivamente.

Error estándar de ajuste

Desde mediados de la década de 1970 se formuló el error estándar de ajuste (*EEA*) como una medida cuantitativa que estima la habilidad descriptiva del modelo probabilístico ajustado (Meylan *et al.*, 2012). El *EEA* permite la comparación objetiva entre las diversas FDP que se prueban o ajustan a una serie o muestra de datos, ya que tiene las unidades de los datos (q_i). Su expresión es la siguiente (Kite, 1977):

$$EEA = \sqrt{\frac{\sum_{i=1}^n (q_i - \hat{q}_i)^2}{n - npa}} \quad (33)$$

en donde n es el número de datos de la serie disponible; npa , el número de parámetros de ajuste de la FDP que se prueba (cinco para el modelo GVE₁₁); q_i , los datos ordenados de menor a mayor, y \hat{q}_i son los valores estimados con la solución inversa de la FDP (Ecuación (2)), para la probabilidad de no excedencia $P(X \leq x)$ estimada con la fórmula de Weibull (Benson, 1962):

$$P(X \leq x) = \frac{m}{n+1} \quad (34)$$

en la cual m es el número de orden del dato, con 1 para el menor y n para el mayor.

Planteamiento de los análisis probabilísticos

Para que la distribución GVE_{11} sea aplicable al registro no estacionario, su gráfica de valores de los datos (q_i) contra el tiempo (t_i) deberá mostrar tendencia lineal, y variabilidad o desviación estándar que aumenta o disminuye con el tiempo. El cálculo del error estándar de ajuste (EEA) con la Ecuación (33) permite la comparación o confrontación de los diversos modelos no estacionarios probados en la serie que se procesa. Cuando los valores del EEA son similares, se puede adoptar un modelo no estacionario de manera subjetiva, por ejemplo, el que conduce a las predicciones más desfavorables.

Respecto a otras distribuciones no estacionarias, existen varias, por ejemplo la Log-Normal (Vogel, Yaindl, & Walter, 2011; Aissaoui-Fqayeh *et al.*, 2009) y los modelos Logística Generalizada (Kim, Nam, Ahn, Kim, & Heo, 2015) y Pareto Generalizada, que son susceptibles de un tratamiento idéntico al descrito en este trabajo para la distribución GVE, cambiando las ecuaciones (13) y (14) por las correspondientes a estos modelos (Rao & Hamed, 2000). También se pueden usar otras covariables en lugar del tiempo (t), como algunos índices climáticos globales o regionales (López-de-la-Cruz & Francés, 2014; Álvarez-Olguín & Escalante-Sandoval, 2016; Campos-Aranda, 2018a).

Con base en la Ecuación (2) se estiman las *predicciones* con periodos de retorno (Tr) de 2, 10, 25, 50 y 100 años, a través del periodo de registro, aplicando las ecuaciones (20) y (18). La primera predicción corresponde a la *mediana*, ya que su probabilidad de no excedencia ($1-p$) es del 50 % y las cuatro siguientes se calculan para los valores siguientes: 0.90, 0.96, 0.98 y 0.99, respectivamente. Además, en una de las series procesadas se hacen predicciones a futuro, en los años 2025 y 2050. Se considera que extrapolar más de 30 años el comportamiento observado en la tendencia histórica y su variabilidad es extremadamente arriesgado.

Al respecto, se considera mucho más confiable el método de los *momentos L móviles* sugerido por Mudersbach y Jensen (2010), y aplicado por Campos-Aranda (2018b), cuya limitante principal es requerir registros bastante amplios. También se muestran en los gráficos de datos y predicciones las estimaciones de los Tr extremos (2 y 100 años) con el modelo estacionario GVE_0 , las cuales son líneas rectas horizontales punteadas.

Series hidrológicas procesadas

Serie 1: con tendencia descendente y disminución de variabilidad

Este registro fue expuesto por Katz (2013) como ejemplo impresionante de la reducción de la precipitación máxima diaria (*PMD*) del invierno (mayo a octubre), medida en la estación Manjimup del extremo suroeste de Australia. El registro abarca 75 valores anuales (de 1930 a 2004), con tendencia descendente y variabilidad mayor en el inicio del registro. Sus magnitudes aproximadas se exponen en la Tabla 1, ya que fueron leídas de una gráfica de barras y se muestran en la Figura 1.

Tabla 1. Datos máximos anuales por procesar de precipitación máxima diaria (*PMD*, en milímetros) y de gasto (q , m³/s) en las estaciones indicadas.

Núm. del dato	(<i>PMD</i>) Est. Manjimup	(q) río Aberjona	(q) Mercer Creek	Núm. del dato	(<i>PMD</i>) Est. Manjimup	(q) río Aberjona	(q) Mercer Creek
1	35.7	9.1	6.8	39	34.7	18.7	5.2
2	49.0	5.9	5.1	40	24.8	39.6	8.9
3	37.8	7.4	6.7	41	33.6	7.1	15.5
4	47.1	6.8	6.2	42	24.8	9.3	17.6
5	84.0	5.0	5.9	43	29.9	24.4	9.1

6	44.9	6.2	5.4	44	28.3	12.7	12.3
7	54.1	9.1	4.8	45	34.8	22.9	11.0
8	45.9	9.1	4.3	16	36.1	8.5	8.6
9	64.8	10.2	6.4	47	34.1	14.7	15.5
10	35.1	3.4	5.4	48	34.9	24.6	18.1
11	52.3	7.1	5.2	49	35.0	8.8	21.3
12	63.9	6.8	7.2	50	30.0	8.8	10.9
13	42.5	6.2	5.0	51	35.0	19.8	13.1
14	31.4	7.1	7.1	52	28.1	13.9	-
15	54.6	13.9	5.3	53	28.0	6.8	-
16	97.5	19.5	5.7	54	51.5	13.9	-
17	49.6	10.8	11.5	55	32.1	15.0	-
18	95.9	4.8	6.9	56	37.0	5.9	-
19	50.0	11.0	9.1	57	24.1	9.1	-
20	42.9	4.8	9.6	58	32.0	34.0	-
21	51.8	5.4	7.8	59	63.7	31.1	-
22	35.6	6.8	7.5	60	33.3	13.3	-
23	37.2	10.8	11.4	61	46.3	11.0	-
24	57.0	22.4	13.3	62	51.8	45.3	-
25	53.9	6.5	14.7	63	37.9	6.8	-
26	54.8	8.8	11.9	64	41.2	8.8	-

27	43.5	2.8	19.0	65	34.1	28.3	-
28	47.0	10.8	17.4	66	35.0	6.5	-
29	36.7	18.4	11.5	67	35.0	36.8	-
30	29.2	13.0	10.0	68	43.4	15.3	-
31	34.3	21.2	23.6	69	36.0	8.2	-
32	48.2	4.0	14.2	70	51.4	-	-
33	50.7	15.6	9.5	71	30.0	-	-
34	37.6	8.5	6.5	72	34.1	-	-
35	42.7	7.4	18.9	73	37.0	-	-
36	31.9	6.5	15.1	74	36.0	-	-
37	38.8	13.0	7.8	75	27.3	-	-
38	29.4	10.2	7.9	-	-	-	-

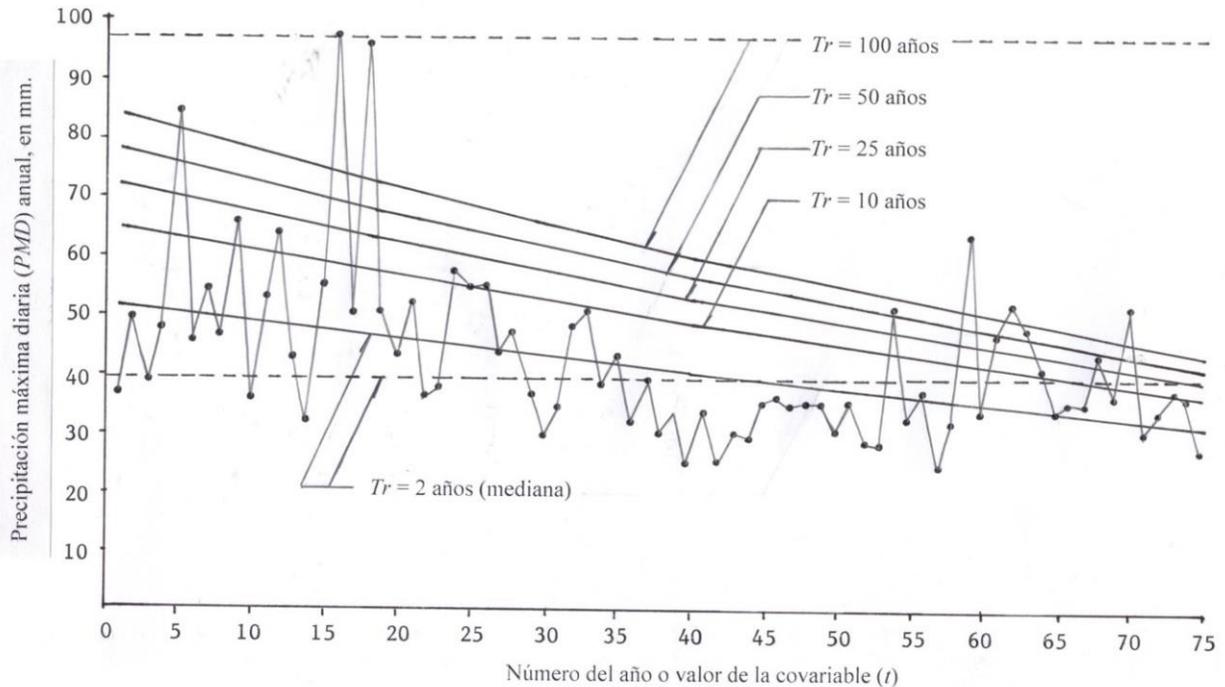


Figura 1. Serie cronológica de los datos (*PMD*) y curvas de predicciones estimadas con la distribución GVE_{11} en la estación pluviométrica Manjimup, Australia.

Serie 2: con tendencia ascendente y variabilidad en aumento

Este registro de gastos máximos anuales fue procesado por Vogel *et al.* (2011), tiene 69 valores en el lapso de 1940 a 2008 y procede de una

estación de aforos localizada en el río Aberjona, justo después de la ciudad de Boston, Massachusetts, EUA, de manera que su cuenca siempre ha estado bajo el impacto en aumento del desarrollo urbano. Sus valores aproximados se muestran en la Tabla 1, ya que fueron leídos en un diagrama de dispersión y se exponen en la Figura 2.

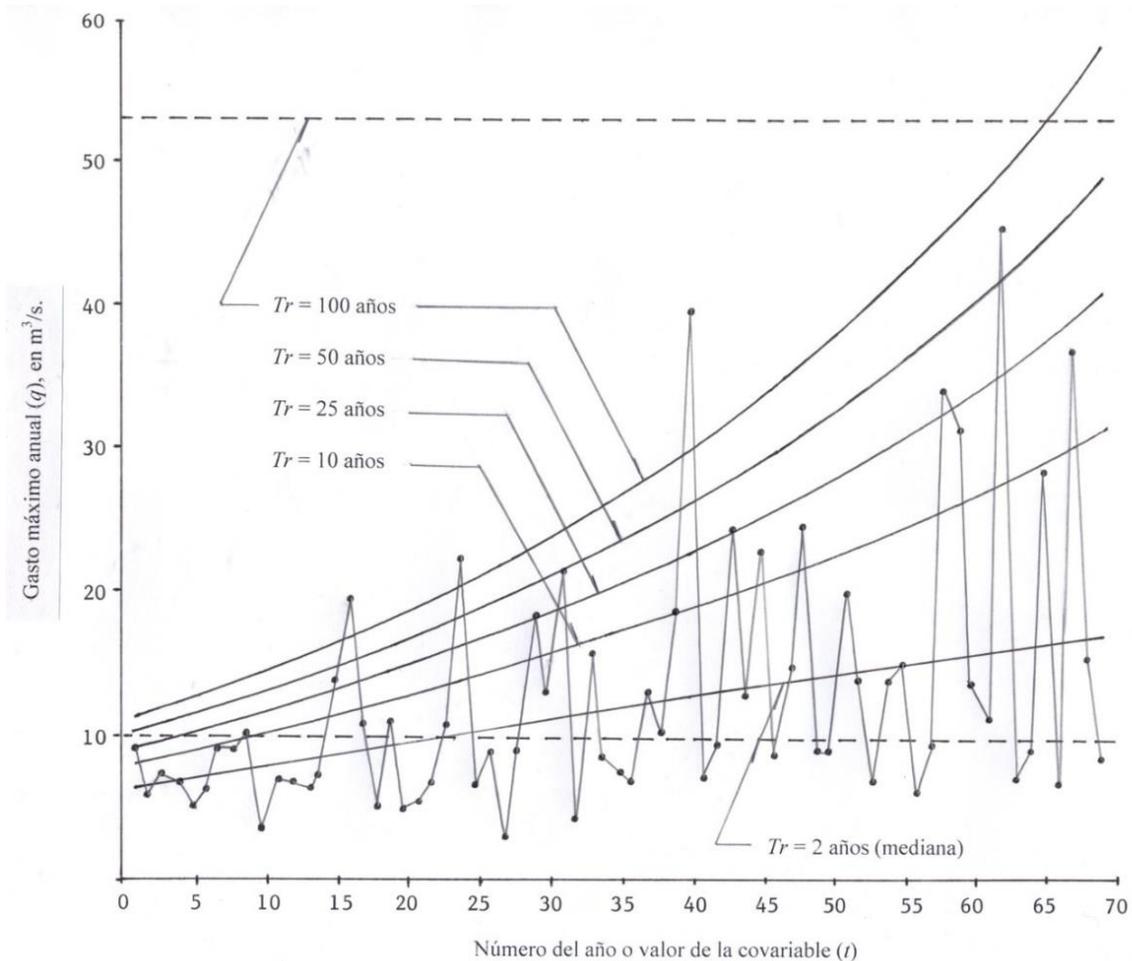


Figura 2. Serie cronológica de los datos (q) y curvas de predicciones estimadas con la distribución GVE_{11} en una estación de aforos del río Aberjona, EUA.

Serie 3: tramo con tendencia ascendente y variabilidad en aumento

Este registro de gastos máximos anuales fue expuesto por Gilleland y Katz (2011), y Katz (2013), y procede de una cuenca en el estado de Washington, EUA, cuya estación de aforos Mercer Creek tiene 51 valores en el periodo de 1956 a 2006. En el intervalo de 15 años entre 1971 y 1985, este registro estuvo influenciado por una urbanización en aumento e intensa. Debido a ello, únicamente tal lapso debe ser considerado no estacionario. Los datos aproximados de tal serie se tienen en la Tabla 1, ya que fueron leídos de una gráfica de barras (Katz, 2013) y se ilustran en la Figura 3.

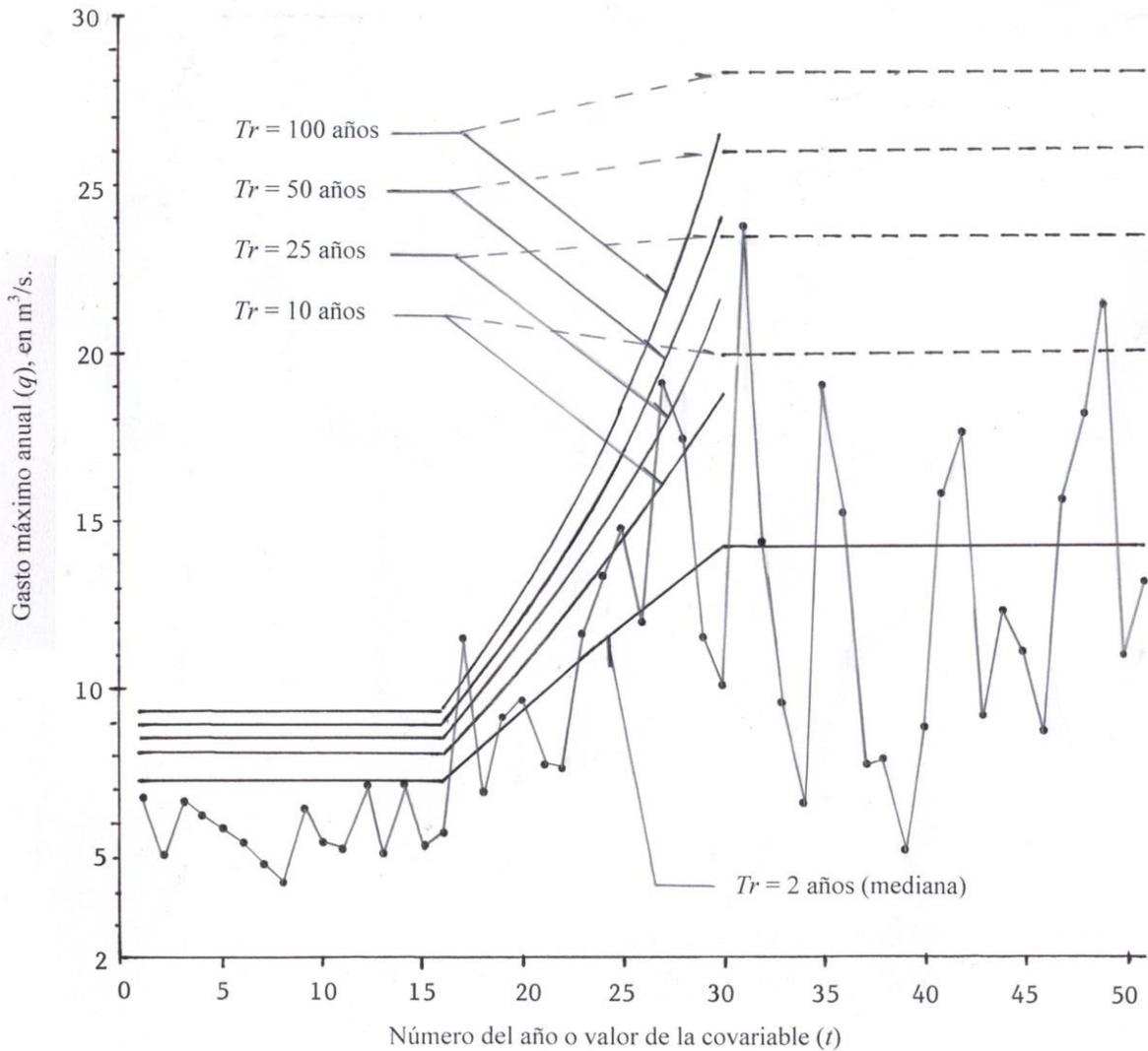


Figura 3. Serie cronológica de los datos (q) y curvas de predicciones estimadas con la distribución GVE_{11} en la estación de aforos Mercer Creek, Washington, EUA.

Descripción de los resultados

Predicciones en la estación Manjimup

La aplicación de las ecuaciones (22) a (25) a los 75 datos de *PMD* de la Tabla 1 condujo a los resultados siguientes: $\mu_0 = \bar{\varepsilon} = 52.8382$, $\mu_1 = -0.2811$, $r_{xy} = -0.4313$, $\sigma_0 = 2.2767$, $\sigma_1 = -0.0132$ y $r_y = -0.2701$. La serie estacionaria $q_e^{\mu, \sigma}$ según las ecuaciones (29) a (32) acepta un modelo GVE_0 con parámetro de forma: $k = -0.0580$.

La Ecuación (2), al utilizar los parámetros de ubicación (ecuación 20) y escala (Ecuación (18)) variables conduce a un error estándar de ajuste (*EEA*) de 15.5 mm. En la Tabla 2 se muestra una parte de las predicciones dentro de registro histórico; a futuro sus predicciones no son importantes debido a la tendencia lineal descendente y a la disminución de la variabilidad. Las predicciones con el modelo estacionario GVE_0 de periodos de retorno 2 y 100 años son 38.7 y 96.6 mm. En la Figura 1 se muestran los datos y las curvas de predicciones descendentes.

Tabla 2. Predicciones (mm) en el periodo histórico en la estación Manjimup con base en la distribución no estacionaria GVE_{11} .

Núm.	Año	Dato	Periodos de retorno en años
------	-----	------	-----------------------------

(t)		(mm)	2	10	25	50	100
1	1930	35.7	50.9	64.4	71.8	77.6	83.5
10	1939	35.1	48.5	60.6	67.1	72.2	77.5
20	1949	42.9	45.9	56.4	62.2	66.7	71.3
30	1959	29.2	43.2	52.5	57.5	61.5	65.5
40	1969	24.8	40.6	48.7	53.1	56.5	60.1
50	1979	30.0	37.9	45.0	48.9	51.9	55.0
60	1989	33.3	35.2	41.4	44.8	47.5	50.2
70	1999	51.4	32.5	37.9	40.9	43.2	45.6
75	2004	27.3	31.1	36.2	39.0	41.2	43.4

Predicciones en el río Aberjona

La aplicación de las ecuaciones (22) a (25) a los 69 gastos máximos anuales (q) de la Tabla 1 condujo a los resultados siguientes: $\mu_0 = \bar{\varepsilon} = 6.1849$, $\mu_1 = 0.1902$, $r_{xy} = 0.4256$, $\sigma_0 = 0.2563$, $\sigma_1 = 0.0304$ y $r_y = 0.5496$. La serie estacionaria $q_e^{\mu, \sigma}$ según las ecuaciones (29) a (32) acepta un modelo GVE₀ con parámetro de forma: $k = -0.1857$; en cambio, la Ecuación (2), al utilizar los parámetros de ubicación (Ecuación (20)) y escala (Ecuación (18)) variables conduce a un error estándar de ajuste

(EEA) de 2.39 m³/s. En la Tabla 3 se muestra una parte de las predicciones dentro de registro histórico; a futuro, sus predicciones son importantes debido a la tendencia lineal ascendente y al aumento de variabilidad. Debido a que este registro de 69 datos concluye en el año 2008, entonces el valor del tiempo t en 2025 es 86 y en 2050 es 111. Las predicciones con el modelo estacionario GVE₀ de periodos de retorno de 2 y 100 años son 10.1 y 53.3 m³/s, respectivamente. En la Figura 2 se exponen los datos y las curvas de predicciones ascendentes.

Tabla 3. Predicciones (m³/s) en el periodo histórico y a futuro en el río Aberjona con base en la distribución no estacionaria GVE₁₁.

Núm. (t)	Año	Dato (m ³ /s)	Periodos de retorno en años				
			2	10	25	50	100
1	1940	9.1	6.1	7.9	9.1	10.1	11.3
10	1949	3.4	7.7	10.1	11.7	13.0	14.6
20	1959	4.8	9.4	12.7	14.8	16.7	18.8
30	1969	13.0	11.1	15.6	18.5	21.0	23.8
40	1979	39.6	12.8	18.8	22.7	26.1	29.9
50	1989	8.8	14.3	22.5	27.8	32.4	37.6
60	1999	13.3	15.7	26.8	34.0	40.2	47.2
69	2008	8.2	16.8	31.4	40.8	49.0	58.3
86	2025	-	18.4	42.7	58.6	72.4	87.9

111	2050	-	18.4	70.5	104.6	134.0	167.2
-----	------	---	------	------	-------	-------	-------

Predicciones en la estación Mercer Creek

La aplicación de las ecuaciones (22) a (25) a los 15 datos de 1971 a 1985, es decir los datos con números 16 a 30 de la Tabla 1, condujo a los resultados siguientes: $\mu_0 = \bar{\varepsilon} = 6.9048$, $\mu_1 = 0.5311$, $r_{xy} = 0.6299$, $\sigma_0 = -0.7769$, $\sigma_1 = 0.1292$ y $r_y = 0.3835$. La serie estacionaria $q_e^{\mu,\sigma}$ según las ecuaciones (29) a (32) acepta un modelo GVE₀ con parámetro de forma: $k = -0.1451$; en cambio, la Ecuación (2), al utilizar los parámetros de ubicación (Ecuación (20)) y escala (Ecuación (18)) variables conduce a un error estándar de ajuste (EEA) de 0.85 m³/s.

En la Tabla 4 se muestra una parte de las predicciones dentro de registro histórico y en la Figura 3 se ilustran los datos, las curvas ascendentes de predicciones y sus extrapolaciones respectivas.

Tabla 4. Predicciones (m³/s) en el periodo histórico en la estación Mercer Creek con base en la distribución no estacionaria GVE₁₁.

Núm. (t)	Año	Dato (m ³ /s)	Periodos de retorno en años				
			2	10	25	50	100

1	1971	5.7	7.3	8.1	8.5	8.9	9.3
4	1974	9.1	8.9	9.9	10.6	11.2	11.8
7	1977	7.5	10.4	12.0	13.0	13.8	14.7
10	1980	14.7	11.8	14.2	15.7	16.9	18.2
13	1983	17.4	13.3	16.7	18.9	20.7	22.7
15	1985	10.0	14.2	18.7	21.5	23.8	26.3

Para esta serie de datos, las predicciones con el modelo estacionario GVE_0 son de importancia vital para *aprobar* la extrapolación hacia el inicio y el final del registro de las predicciones del modelo GVE_{11} que serán las más críticas (Gilleland & Katz, 2011). Por ello, se realizaron dos ajustes: (1) con el registro inicial estacionario de 15 datos (1956-1970), y (2) con el lapso de 21 valores (1986-2006) que ocurren después del tramo con tendencia y variabilidad en aumento (1971-1985). Las predicciones del primer ajuste fueron 5.7, 7.0, 7.6, 7.9 y 8.2 m^3/s para los periodos de retorno (Tr) de 2, 10, 25, 50 y 100 años, respectivamente, con un $EEA = 0.193 m^3/s$; las predicciones correspondientes al segundo ajuste fueron 12.2, 19.9, 23.4, 25.9 y 28.2 m^3/s , respectivamente, con un $EEA = 0.691 m^3/s$.

Entonces, según resultados de la Tabla 4, todas las predicciones de la GVE_{11} son extrapolables hacia el inicio del registro por ser mayores que las de la GVE_0 . Hacia el final del registro, sólo se extrapola la mediana ($Tr = 2$ años), que es 14.2 m^3/s , debido a las predicciones de los otros Tr ;

son más extremas las obtenidas con la GVE_0 , que las del último renglón de la Tabla 4. Lo anterior se muestra en la Figura 3.

Discusión de los resultados

Desde un punto de vista práctico, las tres aplicaciones numéricas descritas cubren los casos comunes en que resulta *aplicable* el modelo GVE_{11} a registros de datos hidrológicos extremos no estacionarios y que son los siguientes:

1. Series con tendencia y variabilidad descendentes. Se presentan en registros de crecientes cuando se han construido varios embalses pequeños dentro de la cuenca de drenaje. Tal comportamiento también puede estar asociado con el cambio climático global, como en el caso de la estación pluviométrica Manjimup.
2. Registros de crecientes con tendencia ascendente. Por lo general se asocian con cuencas que tienen desarrollo urbano o deforestación intensas, como en el caso de la estación Mercer Creek.

Si ambos procesos ocurren en la cuenca, o se ven acrecentados por los impactos del cambio climático, la variabilidad también aumenta hacia el final de la muestra, como en el caso del río Aberjona.

Las *predicciones*, más allá del registro histórico ($t > n$), asociadas con bajas probabilidades de excedencia (*crecientes de diseño*) carecen de importancia en los registros con tendencia y variabilidad descendentes, ya que son menores en el futuro. Lo contrario ocurre con las series con tendencia ascendente y variabilidad en aumento, pero en ellas es necesario explicar o justificar el origen físico probable de tal comportamiento para aceptar las predicciones a futuro e intentar discernir sobre el alcance real de ellas. Es extremadamente riesgoso aceptar el comportamiento de variabilidad en aumento, cuando, por ejemplo, no se sabe si seguirá el desarrollo urbano en la cuenca de drenaje.

Para ampliar lo anterior, se indica que Meylan *et al.* (2012) han señalado que los diversos enfoques o nuevos métodos del AFC que aceptan la no estacionariedad han generado técnicas que permiten hacer predicciones asociadas con bajas probabilidades de excedencia para afrontar el diseño de las obras hidráulicas. Sin embargo, existe la necesidad imperiosa de remplazar el concepto de periodo de retorno (Tr , en años), para adecuarlo al contexto de la no estacionariedad. Lo anterior debido a que el Tr está definido con el intervalo de recurrencia promedio, medido a lo largo de un gran número de ocurrencias. Como ahora, en la no estacionariedad, la media cambia a través del tiempo, el concepto de Tr carece de sentido. Detalles y soluciones al respecto se han propuesto en las referencias de Sivapalan y Samuel (2009); Salas y Obeysekera (2014); Serinaldi (2015), y Salas, Obeysekera y Vogel (2018).

Conclusiones

Los análisis de frecuencias de crecientes (AFC) en registros no estacionarios que muestran tendencia y/o variabilidad no constante serán, en el futuro inmediato, cada vez más comunes debido al aumento de la demanda de agua potable y de alimentos, así como a los impactos del cambio climático y del desarrollo urbano. Un enfoque simple y sin dificultades computacionales que permite realizar el AFC en tales registros se basa en la extensión de la teoría de valores extremos a través del ajuste con momentos L , de la distribución no estacionaria GVE_{11} con parámetro de ubicación y escala variables linealmente con el tiempo que se introduce como *covariable*. Por lo anterior, el modelo GVE_{11} resulta adecuado cuando la serie de datos hidrológicos extremos presenta tendencia, y aumento o disminución de variabilidad que pueden ser aceptadas lineales.

Con base en los resultados de las tres aplicaciones numéricas en registros no estacionarios se observa la simplicidad del método expuesto, además de la facilidad para obtener las *predicciones* asociadas con probabilidades de no excedencia. El contraste gráfico es básico para validar la habilidad descriptiva de las predicciones dentro del registro histórico y en el futuro cercano. Los resultados numéricos del error estándar de ajuste permitirán el contraste y la aceptación de otros modelos probabilísticos no estacionarios. La aplicación del modelo

probabilístico GVE₁₁ a otros datos hidrológicos extremos, como vientos, temperaturas, sequías y niveles del mar, es plenamente factible.

Agradecimientos

Se agradecen las correcciones sugeridas por los tres árbitros anónimos A, C y E, las cuales permitieron ampliar los conceptos relativos a la no estacionariedad y las covariables, además de ubicar el tipo de método expuesto y sus ventajas, entre los disponibles para realizar los AFC no estacionarios.

Referencias

- Aissaoui-Fqayeh, I., El-Adlouni, S., Ouarda, T. B. M. J., & St.-Hilaire, A. (2009). Développement du modèle log-normal non-stationnaire et comparaison avec le modèle GEV non-stationnaire. *Hydrological Science Journal*, 54(6), 141-1156.
- Álvarez-Olguín, G., & Escalante-Sandoval, C. A. (2016). Análisis de frecuencias no estacionario de series de lluvia anual. *Tecnología y ciencias de agua*, 7(1), 71-88.
- Benson, M. A. (1962). Plotting positions and economics of engineering planning. *Journal of Hydraulics Division*, 88(6), 57-71.
- Campos-Aranda, D. F. (2003). Capítulo 5: Ajuste de Curvas. En: *Introducción a los métodos numéricos: Software en Basic y aplicaciones en hidrología superficial* (pp. 93-127). San Luis Potosí, México: Editorial Universitaria Potosina.

- Campos-Aranda, D. F. (2018a). Ajuste con momentos L de las distribuciones GVE, LOG y PAG no estacionarias en su parámetro de ubicación, aplicado a datos hidrológicos extremos. *Agrociencia*, 52(2), 169-189.
- Campos-Aranda, D. F. (2018b). Ajuste con momentos L móviles de la distribución GVE con parámetros variables de ubicación y escala. *Agrociencia*, 52(7), 933-950.
- Chow, V. T. (1964). Statistical and probability analysis of hydrologic data. Frequency analysis. In: Chow, V. T. (ed.). *Handbook of applied hydrology* (pp. 8.1-8.42). New York, USA: McGraw-Hill Book Co.
- Clarke, R. T. (1973). Chapter 5: The estimation of floods with given return period. In: *Mathematical models in hydrology* (pp. 130-146) (Irrigation and Drainage Paper 19). Rome, Italy: FAO.
- Coles, S. (2001). Chapter 3: Classical Extreme Value Theory and Models. In: *An introduction to statistical modeling of extreme values* (pp. 45-73). London, UK: Springer-Verlag.
- Cunderlik, J. M., & Burn, D. H. (2003). Non-stationary pooled flood frequency analysis. *Journal of Hydrology*, 276(1-4), 210-223.
- Davis, P. J. (1972). Gamma function and related functions. In: Abramowitz, M., & Stegun, I. A. (eds.). *Handbook of mathematical functions* (pp. 253-296). New York, USA: Dover Publications.
- El-Adlouni, S., Ouarda, T. B. M. J., Zhang, X., Roy, R., & Bobée, B. (2007). Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resources Research*, 43(3), 1-13.

- El-Adlouni, S., & Ouarda, T. B. M. J. (2008). Comparaison des méthodes d'estimation des paramètres du modèle GEV non stationnaire. *Revue des Sciences de L'Eau*, 21(1), 35-50.
- Gado, T. A., & Nguyen, V. T. V. (2016). An at-site flood estimation in the context of non-stationarity. I: A simulation study. *Journal of Hydrology*, 535, 710-721.
- Gilleland, E., & Katz, R. W. (2011). New software to analyze how extremes change over time. *Eos*, 92(2), 13-14.
- Hosking, J. R., & Wallis, J. R. (1997). Chapter 2: L-moments. Regional frequency analysis and Appendix: L-moments for some specific distributions. In: *An approach based on L-moments* (pp. 14-43, 191-209). Cambridge, UK: Cambridge University Press.
- Jakob, D. (2013). Non-stationarity in extremes and engineering design. In: AghaKouchak, A., Easterling, D., Hsu, K., Schubert, S., & Sorooshian, S. (eds.). *Extremes in a changing climate* (pp. 363-417). Dordrecht, The Netherlands: Springer.
- Katz, R. W. (2013). Statistical methods for non-stationary extremes. In: AghaKouchak, A., Easterling, D., Hsu, K., Schubert, S., & Sorooshian, S. (eds.). *Extremes in a changing climate* (pp. 15-37). Dordrecht, The Netherlands: Springer.
- Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J. C., Gachon, P., & Bobée, B. (2006). Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review. *Journal of Hydrology*, 329(3-4), 534-552.

- Kim, S., Nam, W., Ahn, H., Kim, T., & Heo, J. H. (2015). Comparison of nonstationary generalized logistic models based on Monte Carlo simulation. *Proceedings of the IAHS* (pp. 65-68) (No. 371). 22 June to 2 July, Prague, Czech Republic.
- Kite, G. W. (1977). Chapter 12: Comparison of frequency distributions. In: *Frequency and risk analyses in hydrology* (pp. 156-168). Fort Collins, USA: Water Resources Publications.
- López, J., & Francés, F. (2013). Non-stationary flood frequency analysis in continental Spanish rivers, using climate and reservoir indices as external covariates. *Hydrology and Earth System Sciences*, 17(8), 3189-3203.
- López-de-la-Cruz, J., & Francés, F. (2014). La variabilidad climática de baja frecuencia en la modelación no estacionaria de los regímenes de las crecidas en las regiones hidrológicas Sinaloa y Presidio-San Pedro. *Tecnología y ciencias del agua*, 5(4), 79-101.
- Meylan, P., Favre, A. C., & Musy, A. (2012). Chapter 6: Validation of the model. Chapter 9: Perspectives. In: *Predictive hydrology. A frequency analysis approach*. (pp. 103-117, pp. 161-187). Boca Raton, USA: CRC Press.
- Mudersbach, C., & Jensen, J. (2010). Nonstationary extreme value analysis of annual maximum water levels for designing coastal structures on the German North Sea coastline. *Journal of Flood Risk Management*, 3(1), 52-62.

- Papalexiou, S. M., & Koutsoyiannis, D. (2013). Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, 49(1), 187-201.
- Prosdocimi, I., Kjeldsen, T. R., & Miller, J. D. (2015). Detection and attribution of urbanization effect on flood extremes using nonstationary flood frequency models. *Water Resources Research*, 51(6), 4244-4262.
- Rao, A. R., & Hamed, K. H. (2000). Chapter 7: Extreme Value Distributions. In: *Flood frequency analysis* (pp. 207-257). Boca Raton, USA: CRC Press.
- Salas, J. D., & Obeysekera, J. (2014). Revisiting the concepts of return period and risk for nonstationary hydrology extreme events. *Journal of Hydrologic Engineering*, 19(3), 554-568.
- Salas, J. D., Obeysekera, J., & Vogel, R. M. (2018). Techniques for assessing water infrastructure for nonstationary extreme events: A review. *Hydrological Sciences Journal*, 63(3), 325-352.
- Serinaldi, F. (2015). Dismissing return periods! *Stochastic Environmental Research and Risk Assessment*, 29(4), 1179-1189.
- Sivapalan, M., & Samuel, J. M. (2009). Transcending limitations of stationary and the return period: Process-based approach to flood estimation and risk. *Hydrological Processes*, 23(11), 1671-1675.
- Stedinger, J. R. (2017). Chapter 76. Flood frequency analysis. In: Singh, V. P. (ed.). *Handbook of applied hydrology*, 2nd ed. (pp. 76.1-76.8). New York, USA: McGraw-Hill Education.

Stedinger, J. R., Vogel, R. M., & Foufoula-Georgiou, E. (1993). Frequency analysis of extreme events. In: Maidment, D. R. (ed.). *Handbook of Hydrology* (pp. 18.1-18.66) New York, USA: McGraw-Hill, Inc.

Vogel, R. M., Yaindl, C., & Walter, M. (2011). Nonstationarity: Flood magnification and recurrence reduction factors in the United States. *Journal of the American Water Resources Association*, 47(3), 464-474.