

Búsqueda armónica para optimizar la función Gumbel Mixta univariada

Juan Pablo Molina-Aguilar¹

Martín Alfonso Gutiérrez-López²

Francisco Javier Aparicio-Mijares³

¹ Centro de Investigaciones del Agua, División de Investigación y Posgrado Facultad de Ingeniería, Universidad Autónoma de Querétaro, Santiago de Querétaro, Querétaro, México,
valnahar@hotmail.com

² Centro de Investigaciones del Agua, División de Investigación y Posgrado Facultad de Ingeniería, Universidad Autónoma de Querétaro, Santiago de Querétaro, Querétaro, México,
alfonso.gutierrez@uaq.mx, <https://orcid.org/0000-0003-2770-8642>

³ División de Estudios de Postgrado de la Facultad de Ingeniería, Universidad Nacional Autónoma de México, Jiutepec, Morelos, México,
francisco.aparicio@conagua.gob.mx

Autor para correspondencia: Martín Alfonso Gutiérrez-López,
alfonso.gutierrez@uaq.mx

Resumen

Los fenómenos que se analizan en el campo de la ingeniería con un análisis de frecuencia son habitualmente eventos extremos como inundaciones y sequías. Con frecuencia estos eventos tienen como origen la ocurrencia simultánea de dos fenómenos, por lo que se analizan con dos distribuciones de probabilidad en forma simultánea. Tal es el caso de las series climatológicas e hidrométricas que México, por su exposición ante fenómenos como los huracanes, se analizan con distribuciones de probabilidad para dos poblaciones. El buen uso de estas distribuciones depende de la correcta estimación de sus parámetros de ajuste. Se presenta la aplicación de un algoritmo meta-heurístico de búsqueda armónica para la estimación de los parámetros que optimizan la función Gumbel Mixta univariada. Se utilizan datos hidrométricos máximos anuales para comparar el ajuste de la distribución univariada con las metodologías tradicionales como máxima verosimilitud y el algoritmo de Rosenbrock. Los resultados muestran

que existe una disminución en el error de ajuste y una rápida convergencia cuando se utiliza una búsqueda armónica. Con la disminución del error de ajuste se mejora la estimación de los valores de los caudales de diseño. Se presenta el pseudocódigo del algoritmo para la implementación de la metodología propuesta.

Palabras clave: Búsqueda Armónica, Rosenbrock, Optimización, Meta heurístico, Gumbel Mixta, Análisis de Frecuencias.

Publicado por invitación

Introducción

El objetivo general de un análisis de frecuencias consiste en utilizar los datos históricos de alguna variable hidrológica como precipitación o caudal para estimar el régimen de un evento hidrológico asociado a un periodo de retorno empleando diversas distribuciones de probabilidad. En el campo de la ingeniería civil la selección del período de retorno depende de la naturaleza del proyecto y del riesgo que se quiera asumir (Escalante-Sandoval, 2007). La mayoría de los análisis de frecuencia se realizan con distribuciones univariadas, es decir con una sola variable de análisis que se ajusta a una función de probabilidad con dos o más parámetros (Ashkar *et al.*, 1991). Los fenómenos que se analizan en el campo de la ingeniería con un análisis de frecuencia son diversos. Habitualmente, el análisis de extremos hidrológicos como inundaciones (Arnaud *et al.*, 2017; Al Aamery *et al.*, 2018) y sequías son los más estudiados (Ahn y Merwade, 2016; Ayantobo, *et al.*, 2018; Yang *et al.*, 2018). Sin embargo, estudios del fenómeno del Niño (Karim, *et al.*, 2015), cambio climático (Schulz y Bernhardt, 2016; Smitha *et al.*, 2018) o incluso eventos sísmicos se estudian con distribuciones de probabilidad (Grünthal *et al.*, 2006; Öztürk *et al.*, 2008). En la actualidad son numerosas las funciones que se utilizan en hidrología y gracias a la disposición operacional de los procesadores modernos es posible utilizar modelos probabilísticos que consideran para su solución más de tres parámetros, lo que aumenta la precisión cuando se cuenta con muestras históricas con más de 50 años de registros (Díaz-Delgado *et al.*, 1999). En México y en general en

Latinoamérica el uso de la distribución Gumbel (Gumbel, 1960) es muy difundido, ya que es una distribución versátil y fácil de ajustar. Por ejemplo, comúnmente se utiliza la distribución Gumbel en la estimación de las curvas intensidad-duración-periodo de retorno (Borga *et al.*, 2005; So *et al.*, 2017; Mélèse *et al.*, 2018) y en la validación de dichas curvas con modelos atmosféricos (Vu *et al.*, 2016; Van de Vyver, 2018).

La regionalización hidrológica es un conjunto de metodologías que se utilizan para estimar eventos de diseño en sitios no medidos o con información escasa; estas técnicas también utilizan la distribución de probabilidad univariada de Gumbel (Huang *et al.*, 2016; Zhang *et al.*, 2017). Tal es el caso del Método Regional de la Avenida Índice (Campos-Aranda, 1994), que utiliza la distribución de Gumbel para transferir información hidrométrica dentro de una región considerada hidrológicamente homogénea (Lilenthal *et al.*, 2018). Tal ha sido el uso de este Método que se ha modificado para que pueda ser utilizado en regiones en donde se tiene una afectación directa por fenómenos naturales extremos (Gutiérrez-López y Ramírez, 2005). Por ejemplo, las costas de la república mexicana se ven afectadas todos los años por huracanes que generan precipitaciones fuera de lo normal; es decir, en estas zonas un registro climatológico o hidrométrico estará formado por dos tipos de datos: las mediciones que provienen de los eventos de la temporada normal de lluvias y por los registros que provienen de los eventos extremos (Phillips *et al.*, 2018; Yan *et al.*, 2016). Esto crea la necesidad de analizar una variable con dos distribuciones de probabilidad en forma simultánea; es así como aparece la distribución Gumbel univariada para dos poblaciones o también conocida como Gumbel Mixta univariada (Rossi *et al.*, 1984; Yue, 2000). Sin embargo, los problemas hidrológicos son cada vez más complejos y se requieren estudios que puedan emplear en forma simultánea no solo dos distribuciones de probabilidad, sino también dos variables de análisis; a lo que se le conoce como distribución mezclada bivariada. Por ejemplo, para el caso de la Gumbel, sería una distribución Gumbel Mixta bivariada, la cual puede estimar en forma simultánea los volúmenes y el caudal máximo de un hidrograma (Qi y Liu, 2018) o estimar la duración y lámina de una tormenta (Qian *et al.*, 2017). En México, el precursor en el estudio y aplicación de las distribuciones bivariadas y trivariadas fue Escalante-Sandoval (2007), quien desarrolló el planteamiento teórico para la solución de distribuciones mezcladas, así como su estimación de parámetros. La estimación de parámetros de una distribución bivariada es compleja y requiere cumplir cuatro condiciones: (i) las distribuciones marginales deben ser asintóticas extremas; (ii) la distribución debe ser estable para los valores más grandes de la muestra; (iii) debe existir una función de densidad y (iv) se elimina el caso trivial en donde la

distribución multivariada es el producto de sus distribuciones marginales extremas (Escalante-Sandoval y Domínguez-Esquivel, 2001).

En una distribución de probabilidad univariada o bivariada, cualquiera que sea el caso, la estimación de sus parámetros es uno de los temas más extensos y complejos de la hidrología moderna y requiere de un exhaustivo análisis (Strupczewski *et al.*, 2017; Tosunoglu y Can, 2016). En sus inicios la estimación de parámetros se realizaba con la técnica de mínimos cuadrados, promedios pesados, momentos y recientemente se perfeccionó la teoría de máxima verosimilitud. Actualmente los algoritmos de optimización o la técnica de máxima entropía son sin duda los procedimientos de mayor precisión, incluso para las distribuciones de extremos máximos y mínimos en sus versiones mezcladas bivariadas (Chen y Singh, 2018; Montaseri *et al.*, 2018; Ahn y Palmer, 2016; Tao *et al.*, 2013). Adicionalmente a la selección del tipo de distribución a utilizar, queda perfeccionar la técnica adecuada para la estimación de parámetros. Es por eso que se recurre a los procesos de optimización los cuales son cada vez más aplicados a los problemas de ingeniería (Escalante-Sandoval, 2013). En la actualidad se utiliza una extensa variedad de algoritmos, muchos de ellos basados en métodos de programación numérica lineal y no lineal, que requieren calcular gradientes para buscar la solución en la vecindad de un dato inicial; esto se convierte en una estrategia útil en la obtención del óptimo global mediante modelos simples (Lee y Geem, 2005). Si bien la hidrología moderna avanza hacia la inteligencia artificial, muchos de los problemas reales de optimización ingenieril son de naturaleza compleja y difíciles de resolver utilizando dichos algoritmos (Bardsley, 2016). En particular, la búsqueda de un óptimo local en donde el resultado depende de la selección del dato inicial, de las condiciones de frontera y sobre todo de encontrar la solución óptima en un espacio de solución global es difícil si la función objetivo y sus restricciones presentan inestabilidades numéricas. De lo anterior es evidente que se debe seguir trabajando en la optimización de los modelos y esquemas de solución de las distribuciones univariadas o bivariadas de probabilidad. En este estudio se presenta un algoritmo de búsqueda armónica para la optimización de los parámetros de la función Gumbel Mixta univariada. Los resultados son comparados con los algoritmos de Rosenbrock y los métodos de estimación de parámetros de Momento y de Máxima Verosimilitud, siendo la búsqueda armónica una nueva opción para minimizar el error estándar de ajuste.

Métodos

Algoritmos metaheurísticos

Los inconvenientes computacionales de los modelos numéricos existentes motivaron el desarrollo de algoritmos metaheurísticos basados en simulaciones de fenómenos naturales o artificiales para resolver problemas de optimización ingenieril cuyo factor común es su combinación de reglas de aleatoriedad. Los algoritmos metaheurísticos se utilizan para describir fenómenos muy variados. Por ejemplo, se utilizan algoritmos evolutivos (Fogel *et al.*, 1966) y algoritmos genéticos (Holland, 1975) para revelar procesos de evolución biológica. Glover (1977) presentó algoritmos que se conocen como de búsquedas tabú para explicar la conducta animal. Existen otros procesos complejos como la simulación llamada de recocido simulado (*Simulated Annealing, SA*), que simula y optimiza una función en un gran espacio de búsqueda, cuando los valores son principalmente discretos (Kirkpatrick *et al.*, 1983). De igual manera la interpretación musical definida como el arte de ejecutar una obra musical mediante instrumentos utiliza algoritmos de búsqueda armónica (Geem *et al.*, 2001).

Distribución General de Valores Extremos (DGVE)

En 1960 Gumbel desarrolló la teoría de distribución de valores extremos en la cual la función convergerá a una asíntota de las tres posibles conforme se incrementa la cantidad de datos en la serie de máximos (Koutsoyiannis, 2004; Amin *et al.*, 2014). Las asíntotas se representan por la expresión conocida como Distribución General de

Valores Extremos (Jenkinson, 1969) que usa tres parámetros: ubicación (ε), escala (λ) y forma (k).

$$F(x_i) = e^{\left\{-\left[1+\frac{x_i-\varepsilon}{\lambda}k\right]^{-\frac{1}{k}}\right\}} \quad (1)$$

La función de densidad de probabilidad (f_{dp}) es de la forma:

$$f(x_i) = \frac{dF(x_i)}{dx} = \frac{1}{\lambda} e^{\left\{-\left[1-\frac{x_i-\varepsilon}{\lambda}k\right]^{\frac{1}{k}}\right\}} \left[1 - \frac{x_i-\varepsilon}{\lambda}k\right]^{\frac{1}{k}-1} \quad (2)$$

Se denomina DGVE Tipo I (Gumbel) cuando $k=0$, $-\infty < x_i < \infty$, DGVE Tipo II (Frechet) cuando $k < 0$, $\varepsilon + \lambda/k < x_i < \infty$ y DGVE Tipo III (Weibull) cuando $k > 0$, $-\infty < x_i < \varepsilon + \lambda/k$. En los tres casos $\lambda > 0$.

De igual forma se puede establecer para esta f_{dp} la variable reducida

$$y_i = -\ln \left[1 - \frac{x_i-\varepsilon}{\lambda}k\right]^{\frac{1}{k}} \quad (3)$$

Función de distribución Gumbel

Cuando el parámetro de forma en la Distribución General de Valores Extremos es cero, se tiene la distribución Tipo I, también conocida como Función de distribución de Gumbel y la forma de la función de distribución acumulada en términos de $-\infty < x_i < \infty$, $-\infty < \beta < \infty$ y $\alpha > 0$ es

$$F(x_i) = e^{\left\{-e^{\left[-\frac{x_i-\varepsilon}{\lambda}\right]}\right\}} \quad (4)$$

Su respectiva *fdp* es

$$f(x_i) = \frac{dF(x_i)}{dx} = \frac{1}{\lambda} \left[-e^{\left\{-\left[\frac{x_i-\varepsilon}{\lambda}\right]\right\}} e^{-e^{-\left[\frac{x_i-\varepsilon}{\lambda}\right]}} \right] \quad (5)$$

y su variable reducida:

$$y_i = \frac{x_i-\varepsilon}{\lambda} \quad (6)$$

Gracias a su comportamiento, esta *fdp* es ampliamente utilizada en muchas regiones para el análisis estadístico tanto de caudales como de precipitaciones máximas anuales.

Función de distribución Gumbel Mixta

La expresión matemática que representa la función de distribución de probabilidad para dos poblaciones mezcladas considera la estimación de la probabilidad de no excedencia del gasto máximo anual (x_i), la probabilidad (P) de ocurrencia de eventos no ciclónicos, dos parámetros de ubicación (λ_1, λ_2) y dos parámetros de escala ($\varepsilon_1, \varepsilon_2$) para los cuales los subíndices 1 se asocian a la población no ciclónica mientras que los subíndices 2 se relacionan con la población ciclónica:

$$P(X \leq x_i) = F(x_i) \quad (7)$$

$$F(x_i) = Pe^{-e^{\left\{-\left[\frac{x_i-\varepsilon_1}{\lambda_1}\right]\right\}}} + (1-P)e^{-e^{\left\{-\left[\frac{x_i-\varepsilon_2}{\lambda_2}\right]\right\}}} \quad (8)$$

la *fdp* correspondiente es:

$$f(x_i) = \frac{dF(x_i)}{dx} = \frac{P}{\lambda_1} e^{\left[-\left[\frac{x_i-\varepsilon_1}{\lambda_1}\right] - e^{-\left[\frac{x_i-\varepsilon_1}{\lambda_1}\right]}\right]} + \frac{(1-P)}{\lambda_2} e^{\left[-\left[\frac{x_i-\varepsilon_2}{\lambda_2}\right] - e^{-\left[\frac{x_i-\varepsilon_2}{\lambda_2}\right]}\right]} \quad (9)$$

La estimación de los parámetros de ajuste de las ecuaciones 5 y 9 puede realizarse por los métodos de máxima verosimilitud, momentos o estadísticos convencionales; no obstante, la eficiencia disminuye con el incremento de parámetros por ajustar. Para evitarlo, se minimiza la expresión de la suma de errores cuadráticos pesados E de los valores estimados $F(x_i)$ respecto de los valores empíricos $F'(x_i)$ aplicando la formulación de Weibull

$$E = \sum_{i=1}^n [F'(x_i) - F(x_i)]^2 w_i \quad (10)$$

$$F'(x_i) = \frac{m}{n+1} \quad (11)$$

En la ecuación 10 se recurre al peso asignado al error generado (w_i) que corresponde a la estimación mediante de la función de distribución, mientras que en la ecuación 11 se considera el orden creciente m del caudal máximo anual (x_i) y el total de caudales en el estudio n .

La minimización de la suma de errores cuadráticos pesados se ha llevado a cabo mediante el método de máximo ascenso (Joshi *et al.*, 2012) que es una técnica de gradientes basada en evaluaciones de la función a minimizar y sus derivadas (Raynal-Villaseñor, 1997). No obstante, este proceso es lento, pues requiere de frecuentes cambios de dirección, con lo que se vuelve ineficiente desde el punto de vista computacional y falla al considerar las segundas derivadas.

Algoritmo de Rosenbrock

Para minimizar E en la ecuación 10 se puede aplicar el algoritmo de Rosenbrock para múltiples variables no restringidas (Rosenbrock, 1960; Kuester y Mize, 1973; Campos-Aranda, 2003). Esta técnica se cataloga como una de búsqueda directa. Para aplicarla se requieren valores iniciales de los parámetros, para lo cual se procede al cálculo como si fuese una sola población mediante la técnica de momentos (Chow, 1964) considerando el valor de la media (x_1, x_2), la desviación estándar (s_1, s_2) de las ambas poblaciones, no ciclónica y ciclónica respectivamente, así como el número de caudales de origen ciclónico (n_c) considerados. Los parámetros se pueden entonces estimar como:

$$\varepsilon_1 = \bar{x}_1 - 0.5772\lambda_1 \quad (12)$$

$$\lambda_1 = \frac{\sqrt{6}}{\pi} s_1 \quad (13)$$

$$\varepsilon_2 = \bar{x}_2 - 0.5772\lambda_2 \quad (14)$$

$$\lambda_2 = \frac{\sqrt{6}}{\pi} s_2 \quad (15)$$

$$P = \frac{n-n_c}{n} \quad (16)$$

La bondad de ajuste mediante el Error Estándar de Ajuste (EE) propuesta por Raynal (1997) considera la comparación de los caudales máximos anuales registrados $F'(x_i)$ y estimados $F(x_i)$ así como la comparación de la cantidad de caudales (n) respecto de los parámetros de ajuste (n_p) de la función de distribución:

$$EE = \left[\frac{[F'(x_i) - F(x_i)]^2}{n-n_p} \right]^{\frac{1}{2}} \quad (17)$$

Este algoritmo permite minimizar la ecuación 17, la cual considera en su estructura las ecuaciones 5 y 9, que se trata de una función no lineal de múltiples variables no restringidas de la forma $F(x_1, x_2, \dots, x_n)$ ejecutándose el seudocódigo mostrado en la Tabla 1. El detalle y código de programación se puede consultar en Campos-Aranda (1989).

Tabla 1. Seudocódigo del algoritmo de Rosenbrock.

Paso	Actividad
1	Definir los parámetros iniciales punto inicial (x_i), criterio de parada ($(c_p > 0)$, factor de expansión ($f_E > 1$) y factor de contracción ($(-1 < f_c < 0)$)
2	Para $j = 1, \dots, n$ considerar $d_j = e^j$, el vector unitario de la variable (x_j) y las longitudes de paso en cada una de las direcciones (h_j), tomando $y_1 = x_1$ siendo $k = j = 1$. Ir al paso 3
3	Evaluar $f(y_j + h_j d_j)$; si $f(y_j + h_j d_j) < f(y_j)$ entonces $y_{j+1} = y_j + f_E h_j d_j$; en caso contrario, $y_{j+1} = y_j + f_c h_j d_j$. Ir al paso 4
4	Si $j < n$ entonces $j = j + 1$ y regresar al paso 2; en caso contrario, ir al paso 5
5	Calcular $f(y_{n+1})$ y revisar las siguientes condiciones Si $f(y_{n+1}) < f(y_1)$ entonces $y_1 = y_{n+1}$ haciendo $j = 1$ y regresar al paso 2. Si $f(y_{n+1}) < f(y_1)$ y $f(y_{n+1}) < f(x_k)$ ir al paso 6. Si $f(y_{n+1}) < f(y_1)$ y $f(y_{n+1}) = f(x_k)$ seleccionar uno de los dos casos siguientes, el que corresponda. Si $ h_j < c_p$ entonces x_k es aproximación del óptimo para la función objetivo y concluye el algoritmo. En caso contrario $y_1 = y_{n+1}$ asignando $j = 1$. Ir al paso 3
6	Considerar $x_{k+1} = y_{n+1}$ y revisar las siguientes condiciones Si $\ x_{k+1}\ < c_p$ entonces x_{k+1} es aproximación del óptimo para la función objetivo y concluye el algoritmo. En caso contrario calcular $\lambda_1, \dots, \lambda_n$ de la relación $x_{k+1} - x_k = \sum_{j=1}^n \lambda_j d_j$ formando un nuevo conjunto de direcciones $\bar{d}_1, \dots, \bar{d}_n$ de la forma

$$w_j = \begin{cases} d_j & \lambda_j = 0 \\ \sum_{j=1}^n \lambda_j d_j & \lambda_j \neq 0 \end{cases}$$

$$y_j = \begin{cases} w_j & y = 1 \\ \sum_{i=1}^{j-1} (w_j^t \bar{d}_i) \bar{d}_i & j \geq 2 \end{cases}$$

Tomando $\bar{d}_j = \frac{y_j}{\|y_j\|}$. Considerar las longitudes del paso 2, tomar $y_1 = x_{k+1}$, con $k = k + 1$ y $j = 1$ regresar al paso 2.

Algoritmo Búsqueda Armónica

Este algoritmo metaheurístico conceptualiza el uso de procesos musicales para perfeccionar armonías las cuales de manera análoga en las técnicas de optimización son vectores de solución y las improvisaciones realizadas por los músicos que hacen alusión a los esquemas de búsqueda local o global (Geem et al., 2001).

No requiere de valores iniciales para las variables de decisión, es decir, en lugar de una búsqueda de gradientes realiza una búsqueda estocástica aleatoria basada en la tasa de consideración de memoria armónica y la tasa de ajuste de tono lo cual elimina el uso de información de derivadas, reduciendo los requerimientos matemáticos y facilitando su implementación en problemas de optimización en problemáticas de ingeniería (Lee y Geem, 2005). El seudocódigo se muestra en la Tabla 2.

Para realizar la optimización de la función objetivo (CF) se consideran soluciones candidatas que constituyen las armonías y se representan matemáticamente por vectores n-dimensionales; dichas soluciones constituyen la población inicial generada de manera aleatoria y almacenada en la memoria armónica (HM), a partir de la cual se genera una nueva armonía partiendo de los elementos contenidos por reiniciación estocástica o bien ajustando el tono de un vector existente. Con ello HM se actualiza al comparar la nueva solución candidata respecto de la peor almacenada sustituyéndola si mejoró la solución de la función objetivo y en caso contrario se elimina la solución candidata. Este proceso se realiza hasta alcanzar el criterio de paro establecido (Cuevas y Ortega-Sánchez, 2013)

Cada solución candidata está comprendida dentro del rango definido por los límites superior (xU) e inferior (xL) denominado ancho de banda ($B = xU - xL$) de las variables de decisión (N), la cantidad de vectores en HM define el tamaño de la memoria (HMS) la cual almacena los mejores desempeños armónicos para la función objetivo, por su parte al puntualizar la tasa de consideración de memoria armónica para el ajuste de tono ($HMCR$) se establece la posibilidad de las nuevas armonías tomando elementos almacenados en HM y la tasa de ajuste de tono (PAR) da la posibilidad de modificar una solución previa sin exceder la magnitud máxima B , todo lo anterior dentro de un proceso definido por el número total de improvisaciones (NI) que corresponden a las iteraciones propuestas.

Para garantizar la solución manteniendo características particulares del problema se pueden establecer funciones de restricción (RF) con coeficientes de penalización (PC) que corrige el valor de CF previo a la actualización de HM .

Tabla 2. Seudocódigo del algoritmo de Búsqueda de armonías.

Paso	Actividad
1	Definidas CF y RF acorde al problema, establecer HMS , $HMCR$, B , PAR , NI y PC
2	Definir $HM = [x_i(j)]$ para $x_i(j) = xL(j) + [xU(j) - xL(j)] * rand(0,1)$ donde $i = 1,2, \dots, HMS$ y $j = 1,2, \dots, N$
3	Para $j = 1,2, \dots, N$ si $rand_1(0,1) < HMCR$ entonces $x_{new}(j) = x_i(j)$ continuar al paso 4, caso contrario ir al paso 5.
4	Si $rand_2(0,1) < PAR$ entonces $x_{new}(j) = x_{new}(j) + B * rand_3(0,1)$, caso contrario $x_{new}(j) = x_i(j)$.
5	Actualizar la memoria armónica si $f(x_{new}) + PC * f(x_{new}) < f(x_w)$ entonces $x_w = x_{new}$ donde x_w es la peor armonía en HM , caso contrario ir al paso 3.
6	Si $k = NI$ concluye la optimización y la solución es x_{best} en HM caso contrario ir al paso 3

Materiales

En primera instancia se obtuvo la información hidrométrica histórica de la estación 10040 Santa Cruz ubicada en el estado de Sinaloa (106° 57'10", 24° 29'05") en el periodo de 1944 a 1980 del Banco Nacional de Datos de Aguas Superficiales (BANDAS) de la Comisión Nacional del Agua (CONAGUA) (Tabla 3 y Figura 1). Como caso adicional para verificar la capacidad de optimización en la función de distribución Gumbel Mixta utilizando el algoritmo de Búsqueda Armónica, se comparó el registro de caudales máximos anuales (Gómez *et al.*, 2010) correspondientes a la estación hidrométrica número 12504 nombrada La Cuña en el estado de Jalisco, México (102° 49'59", 21° 00'15") para el periodo de 1947 a 2004, contándose con 58 datos registrados en el Tabla 4 (Figura 2).

Tabla 3. Registro de caudales máximos anuales de la estación hidrométrica Santa Cruz, Sinaloa, México.

año	Q (m ³ /s)								
1944	2142	1952	677	1960	1074	1968	7000	1976	1495
1945	1023.4	1953	807	1961	1280	1969	484	1977	836
1946	837.6	1954	553	1962	1002	1970	920.6	1978	940
1947	1161.2	1955	1252	1963	3680	1971	812	1979	3080
1948	1062	1956	369.5	1964	861	1972	3332.4	1980	1550
1949	784.2	1957	293	1965	888.8	1973	898		
1950	1086.3	1958	1157.2	1966	1166.4	1974	2790		
1951	487.8	1959	762.222	1967	950	1975	620		

Tabla 4. Registro de caudales máximos anuales de la estación hidrométrica La Cuña, Jalisco, México.

año	Q (m ³ /s)										
1947	784.00	1957	199.00	1967	1474.86	1977	439.72	1987	184.68	1997	78.44
1948	736.80	1958	690.00	1968	323.00	1978	280.22	1988	595.20	1998	261.86
1949	510.00	1959	340.60	1969	160.40	1979	267.20	1989	110.18	1999	196.30
1950	461.00	1960	249.60	1970	763.75	1980	287.28	1990	523.86	2000	46.81
1951	411.00	1961	350.00	1971	578.00	1981	280.70	1991	1636.33	2001	313.81
1952	326.00	1962	317.00	1972	191.75	1982	156.50	1992	1168.00	2002	319.60
1953	349.80	1963	732.56	1973	2440.00	1983	455.50	1993	295.00	2003	621.07
1954	130.40	1964	265.05	1974	238.35	1984	501.20	1994	212.80	2004	824.50
1955	690.00	1965	743.60	1975	622.08	1985	385.00	1995	367.42		
1956	266.00	1966	463.90	1976	1374.00	1986	698.19	1996	144.57		

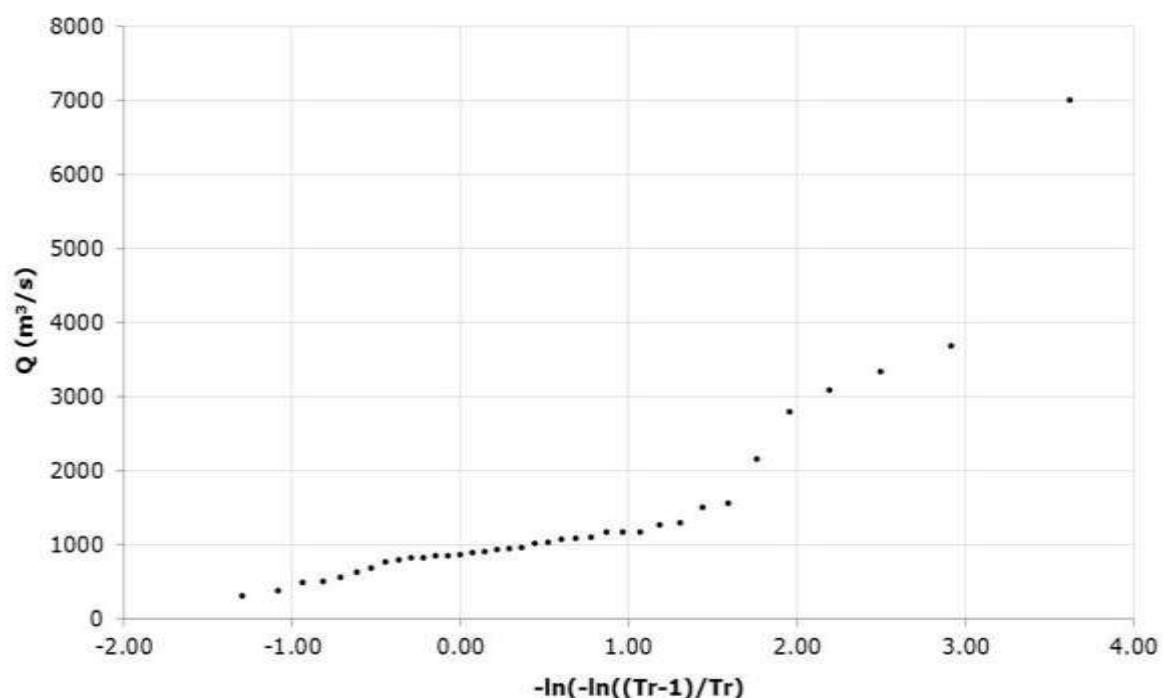


Figura 1. Variable reducida de caudales máximos anuales de la estación Santa Cruz, Sinaloa.

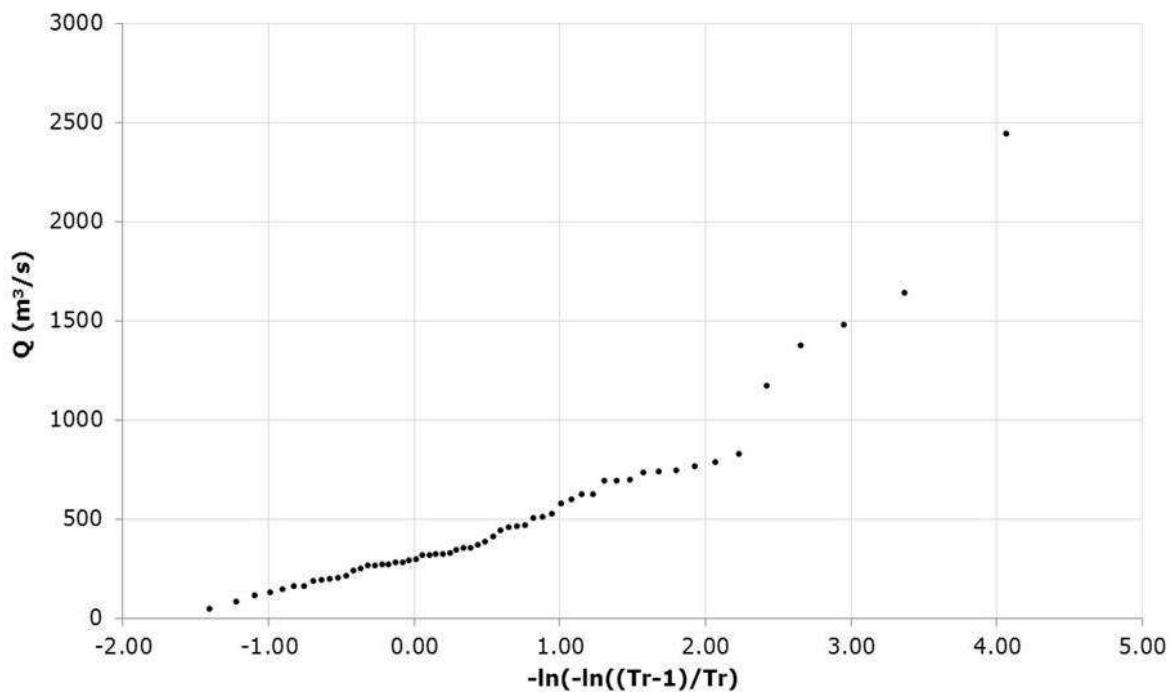


Figura 2. Variable reducida de caudales máximos anuales de la estación La Cuña, Jalisco.

Resultados

Gumbel para una población

Primeramente, se realizó el ajuste de la función Gumbel de una población univariada, por las metodologías de Momentos, Máxima Verosimilitud y Búsqueda Armónica, a los 37 datos históricos de la estación Santa Cruz. La programación del algoritmo de Búsqueda Armónica se llevó a cabo en Matlab para lo cual se definieron los rangos de las variables de decisión, las cuales corresponden a los parámetros del algoritmo recordando que los valores iniciales se establecen estocásticamente $xU=[1000,10000]$, $xL=[0;0]$, $HMS=40$, $N=10000$,

$HMCR=0.90$, $PAR=0.35$; $B=(xU-xL)/NI$. Una vez verificada la programación realizada en Matlab se modificó la función objetivo y se calculó el error estándar de ajuste, para optimizar la función Gumbel Mixta, contrastando los resultados respecto al algoritmo de Rosenbrock, (Figura 3) estableciendo los parámetros iniciales considerando nueve datos ciclónicos $P=0.756757$, $\lambda_1=194.24 \text{ m}^3/\text{s}$, $\varepsilon_1=736.678 \text{ m}^3/\text{s}$, $\lambda_2=1368.31 \text{ m}^3/\text{s}$ y $\varepsilon_2=2137.95 \text{ m}^3/\text{s}$. De igual manera se definieron los rangos para las variables de decisión que corresponden a los parámetros del algoritmo $xU=[1, 10000, 10000, 10000, 10000]$, $xL=[0,0,0,0,0]$, $HMS=40$, $NI=10000$, $HMCR=0.90$, $PAR=0.35$; $B=(xU-xL)/NI$. La Tabla 5 y la Figura 3 muestran los valores de los errores de ajuste para los tres procedimientos de estimación de parámetros por los métodos mencionados.

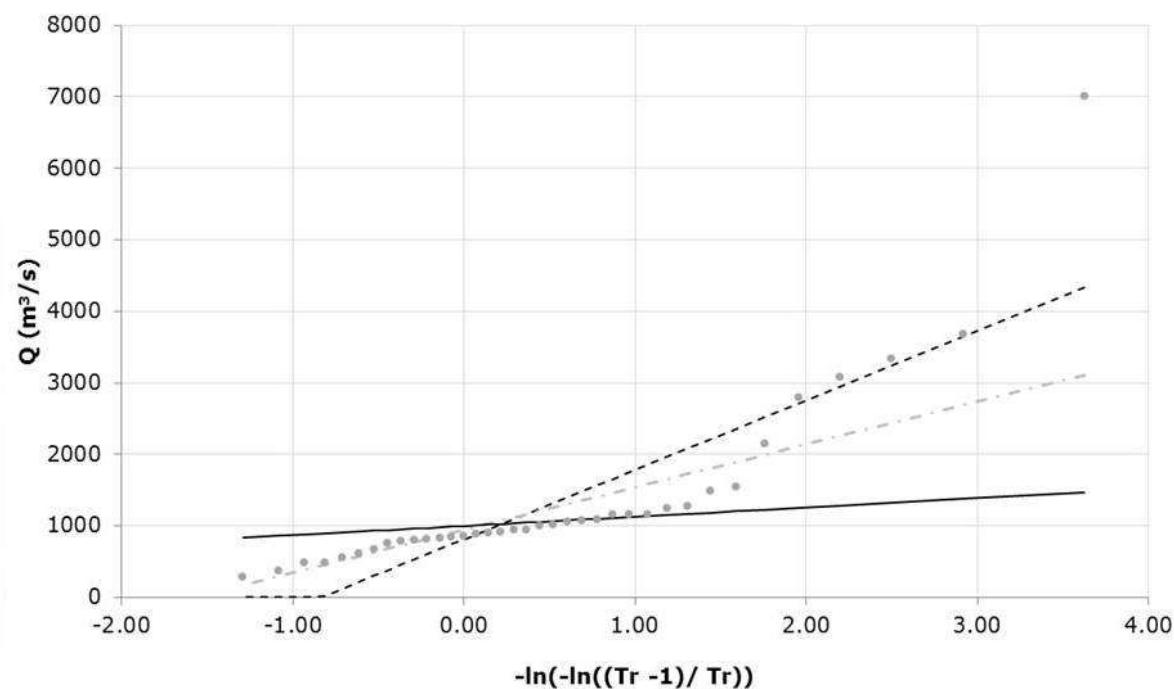


Figura 3. Ajuste de la función Gumbel de una población para los datos de la estación Santa Cruz, Sinaloa por el método de momentos (—), el Método de Máxima verosimilitud (---) y Búsqueda Armónica (—)

Tabla 5. Parámetros de ajuste para la función Gumbel de una población para el registro de caudales máximos anuales de la estación hidrométrica Santa Cruz, Sinaloa, México.

Metodología	$\lambda_1 (\text{m}^3/\text{s})$	$\varepsilon_1 (\text{m}^3/\text{s})$	EE
-------------	-----------------------------------	---------------------------------------	----

Momentos	972.2	814.40	598.77
Máxima Verosimilitud	597.05	946.58	737.63
Búsqueda Armónica	1000	128.98	82.82

A lo largo del número total de iteraciones se registró el valor de la función de optimización representada por la ecuación 17 mostrado en la Figura 4; de igual manera, se promedió la función de optimización evaluada en las 40 armonías de la memoria armónica como se muestra en la Figura 5.

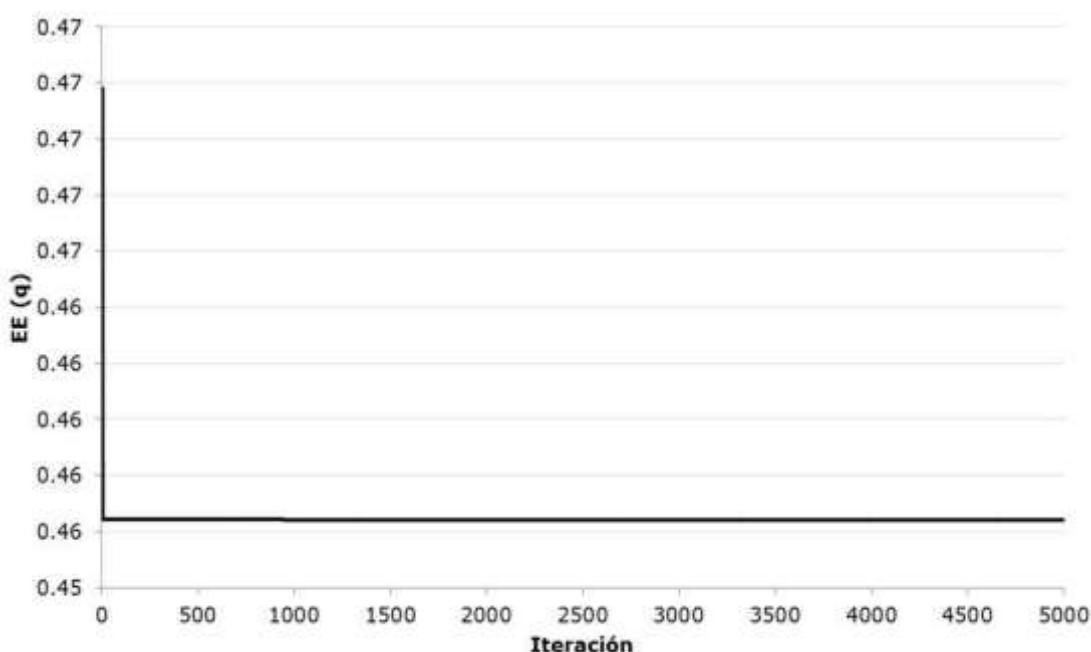


Figura 4. Evaluación de la función de optimización para Gumbel de una población.

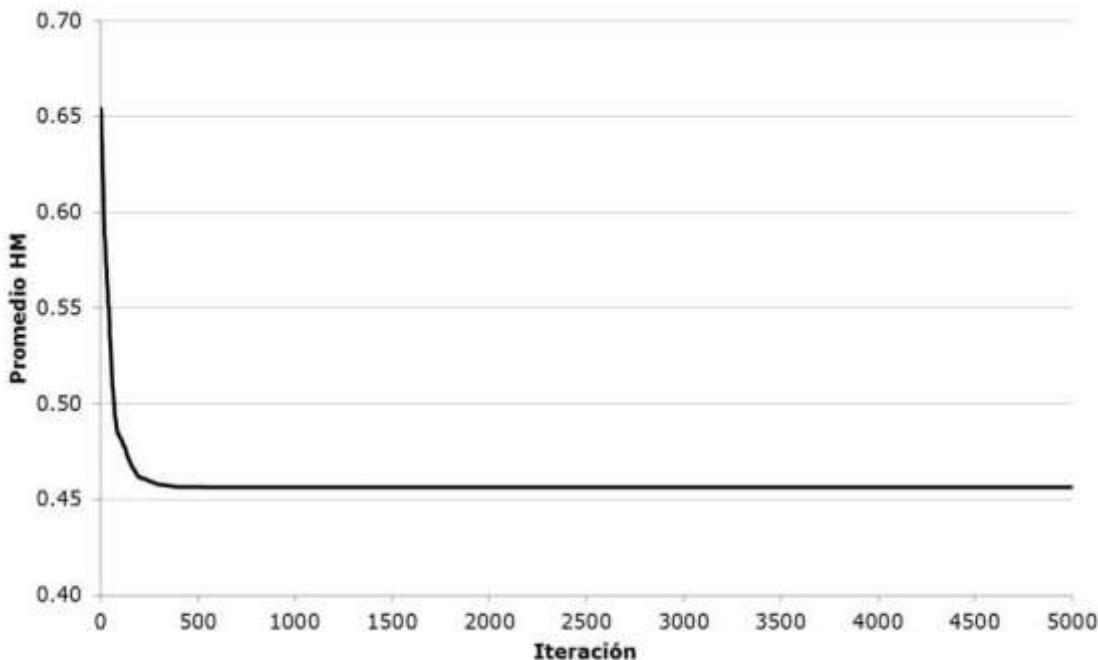


Figura 5. Promedio de la función de optimización en la memoria armónica.

Gumbel Mixta

A continuación, se utiliza el mismo algoritmo para llevar a cabo el proceso de optimización de la distribución de probabilidad Gumbel Mixta o de dos poblaciones. El algoritmo realiza la modificación armónica más desfavorable contenida en la memoria armónica y estocásticamente modifica una de las variables de decisión. Es decir, modifica los cuatro parámetros de escala y ubicación, así como el parámetro de asociación P . El algoritmo lleva a cabo la exploración del espacio de solución como se observa en las Figuras 6 y 7 definiéndose una tendencia horizontal por la repetición sucesiva del valor óptimo del parámetro en la armonía evaluada, mientras que los valores aislados representan los valores estocásticos generados en el proceso de convergencia. La Tabla 6 muestra los valores ajustados de los cinco parámetros que resultan de este procedimiento, así como el error de ajuste. En la Figura 8 se muestra la comparación de los ajustes por momentos, Rosenbrock y por Búsqueda Armónica, para los datos históricos de la estación hidrométrica Santa Cruz.

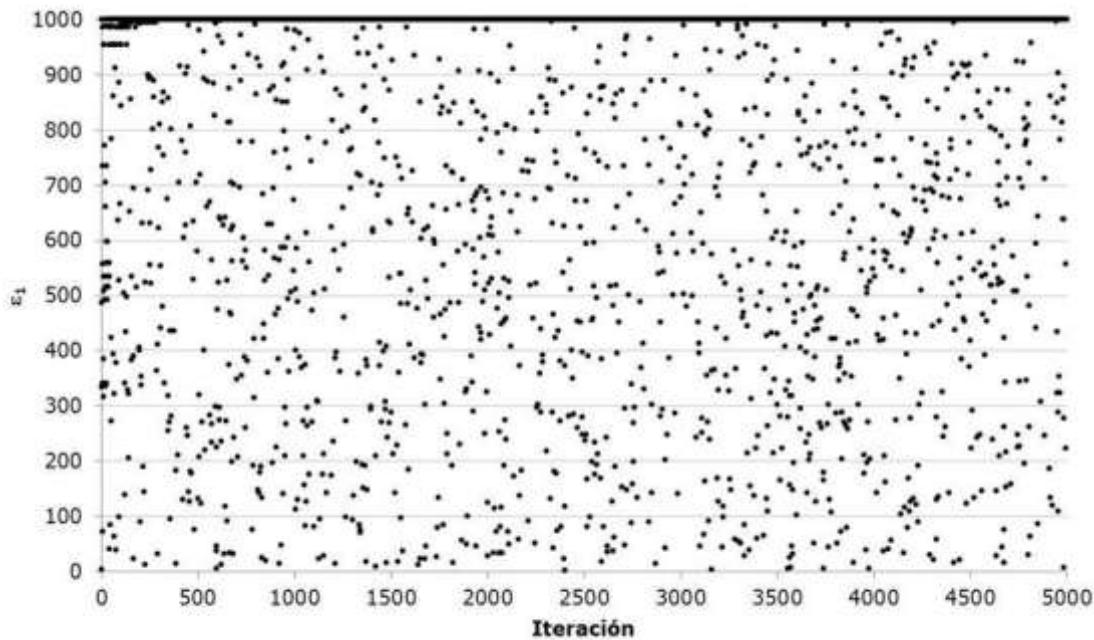


Figura 6. Comportamiento estocástico del parámetro de escala.

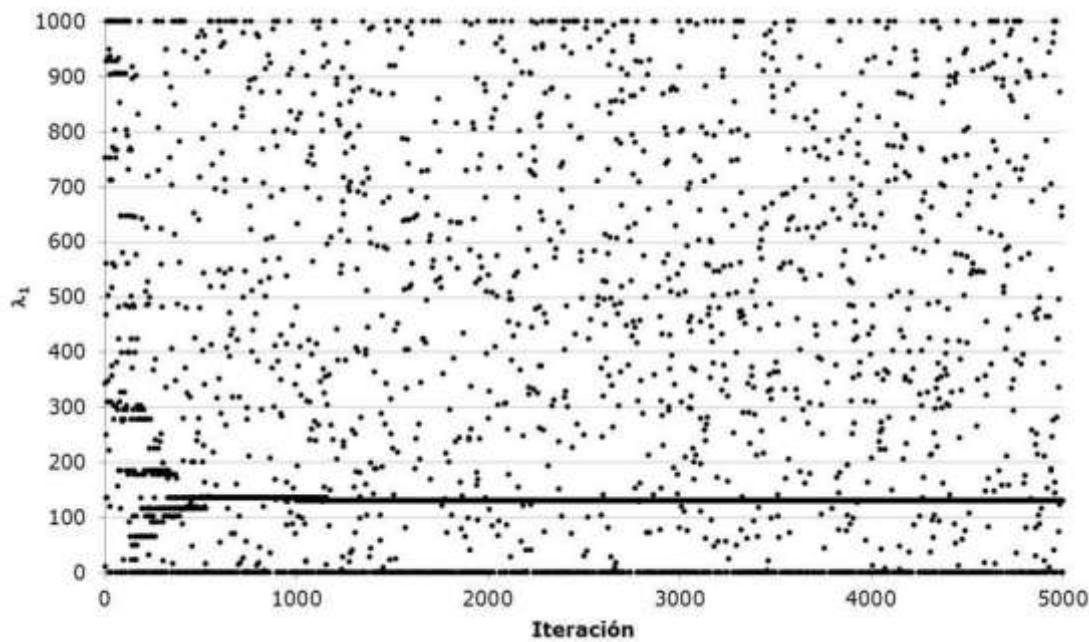


Figura 7. Comportamiento estocástico del parámetro de ubicación.

Tabla 6. Parámetros de ajuste para la función Gumbel Mixta para el registro de caudales máximos anuales de la estación hidrométrica Santa Cruz, Sinaloa. México.

Metodología	P	λ_1 (m ³ /s)	ε_1 (m ³ /s)	λ_2 (m ³ /s)	ε_2 (m ³ /s)	EE
Momentos	0.784	863.652	200.780	3,133.625	1,369.655	542.022
Rosenbrock	0.850	769.051	296.580	3,037.430	1,022.059	407.024
B. Armónica	0.824	788.796	281.276	2,850.544	1,436.975	400.008

Las Figuras 9 y 10 muestran la evolución del valor de la función de optimización representada por la ecuación 17 para el caso de la distribución Gumbel Mixta univariada. La evolución en el espacio de solución de la búsqueda aleatoria de los armónicos que lleva a cabo el algoritmo propuesto se presenta en una secuencia en las Figuras 11 a 15.

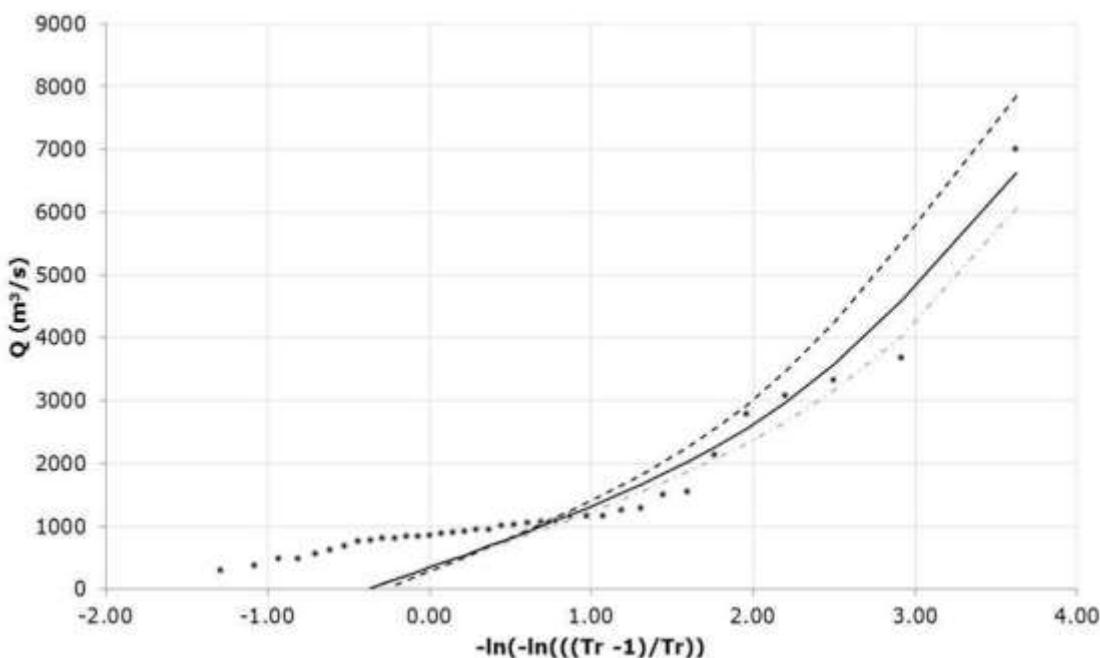


Figura 8. Ajuste de la función Gumbel Mixta para los datos de la estación Santa Cruz, Sinaloa. Método de momentos (—), algoritmo de Rosenbrock (...) y Búsqueda Armónica (—)

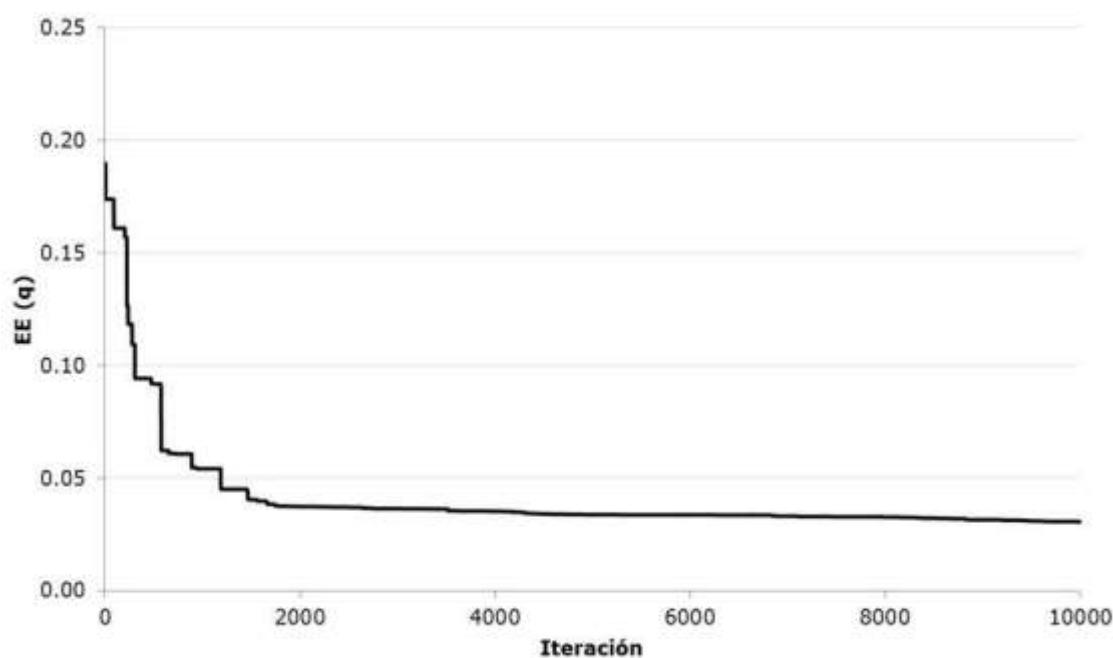


Figura 9. Evaluación de la función de optimización para Gumbel Mixta.

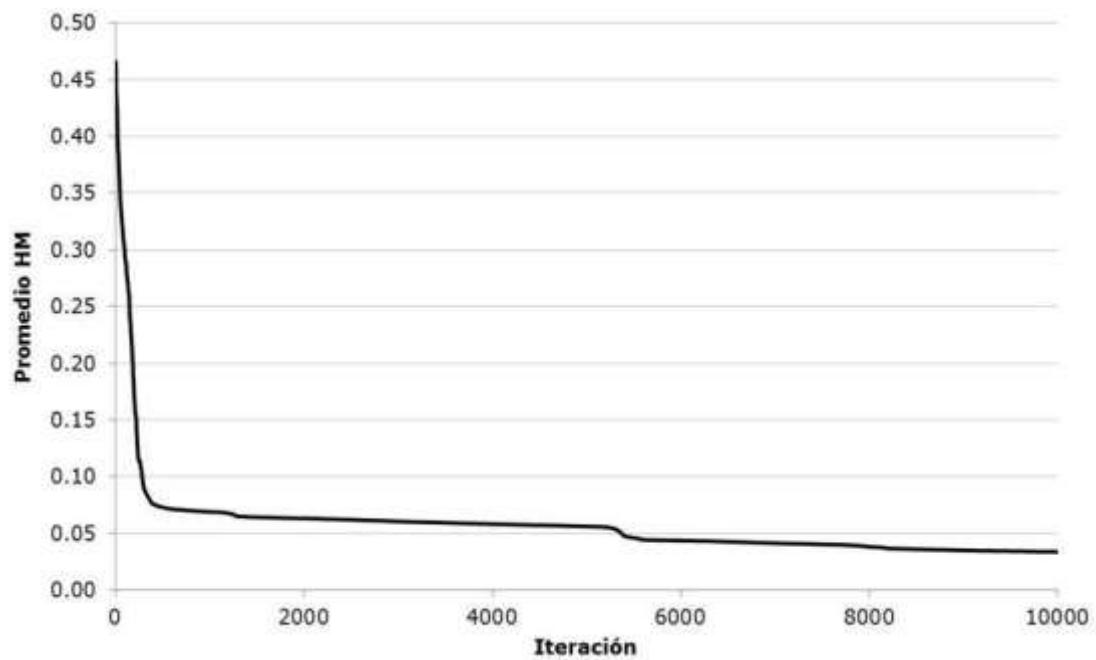


Figura 10. Promedio de las 40 armonías evaluadas en la función de optimización.

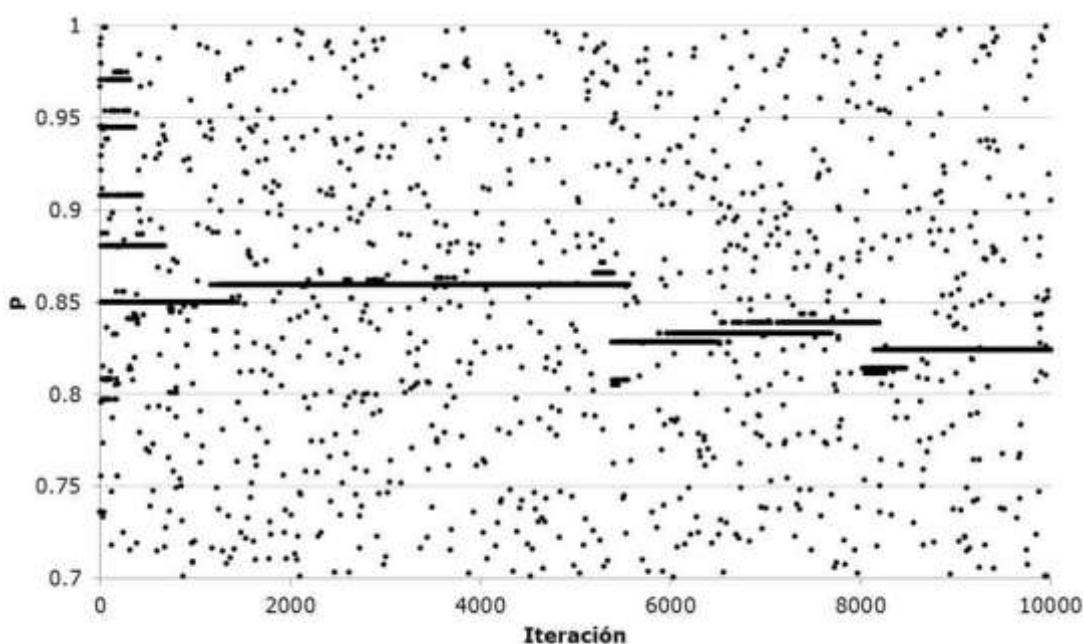


Figura 11. Comportamiento aleatorio de la probabilidad.

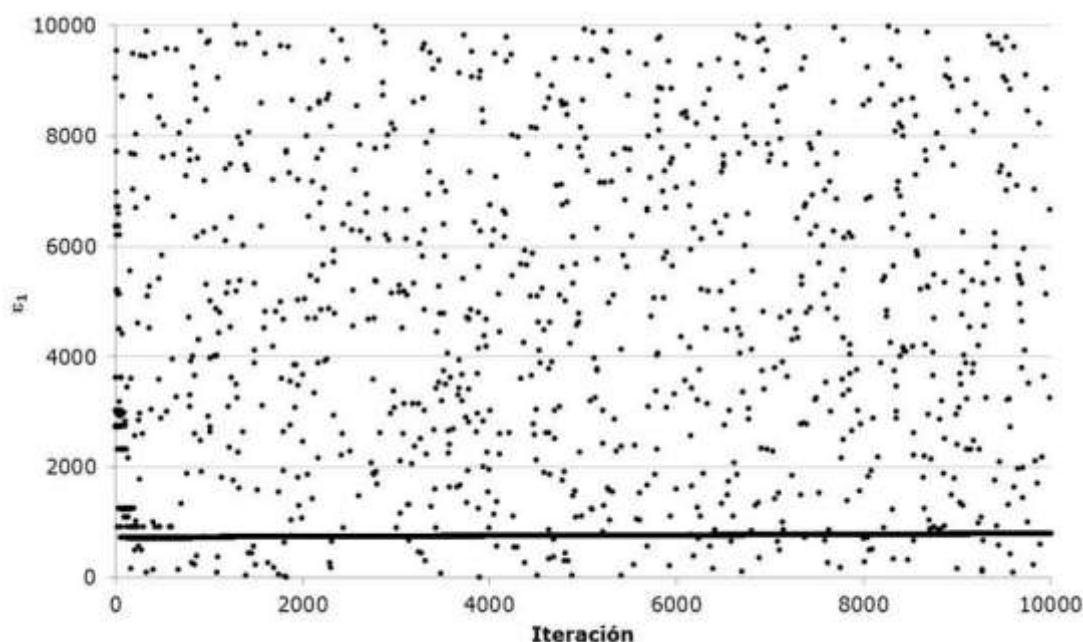


Figura 12. Comportamiento aleatorio del parámetro de escala no ciclónico.

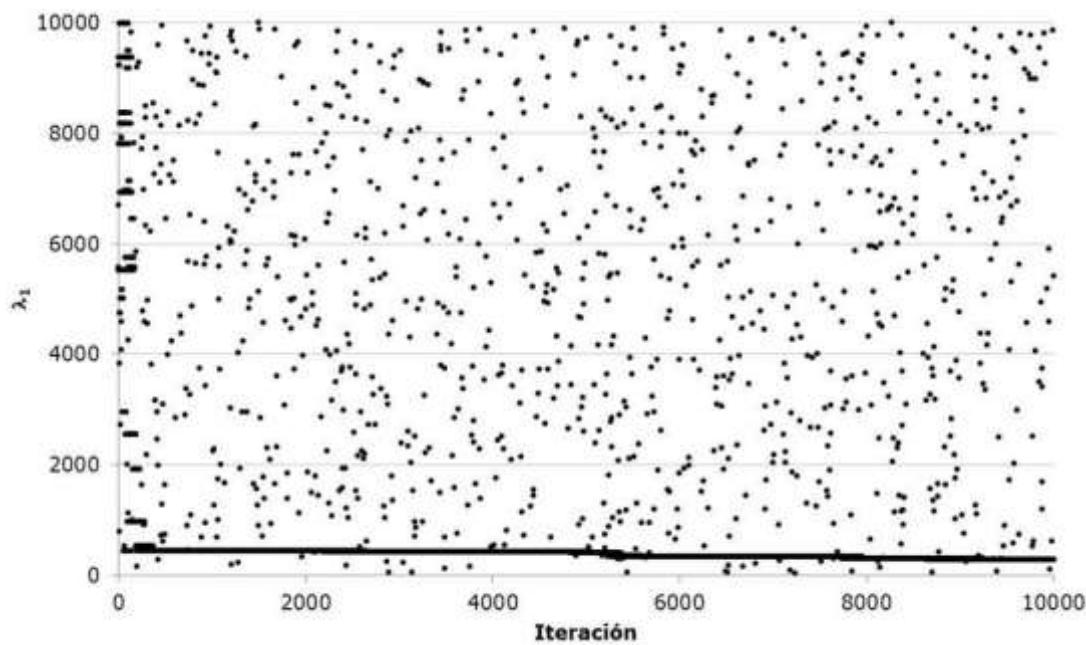


Figura 13. Comportamiento aleatorio del parámetro P de ubicación no ciclónico.

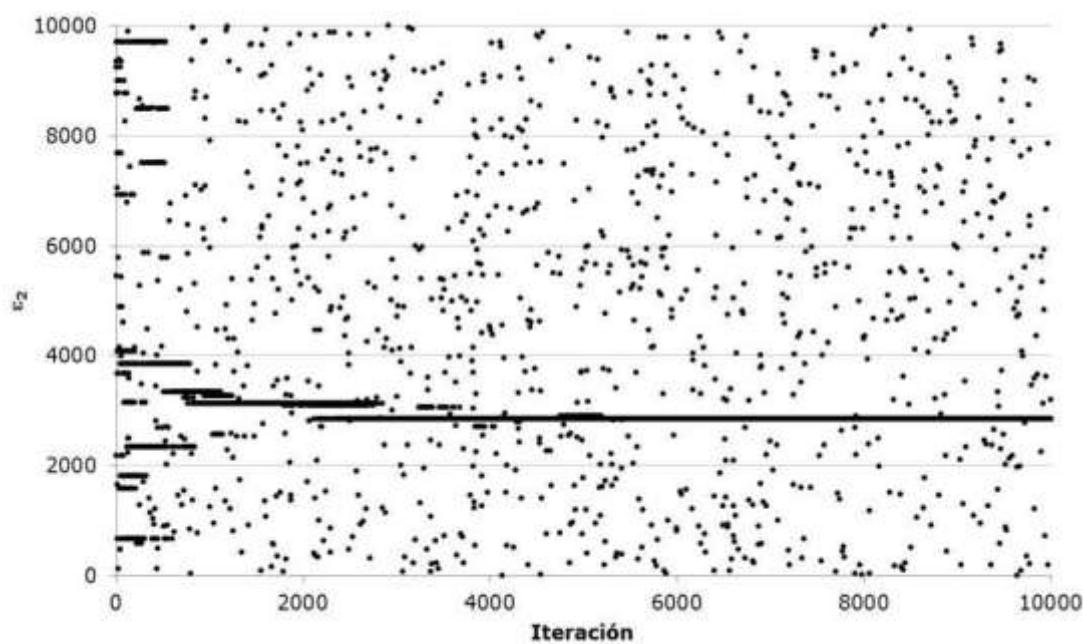


Figura 14. Comportamiento aleatorio del parámetro P de escala ciclónico.

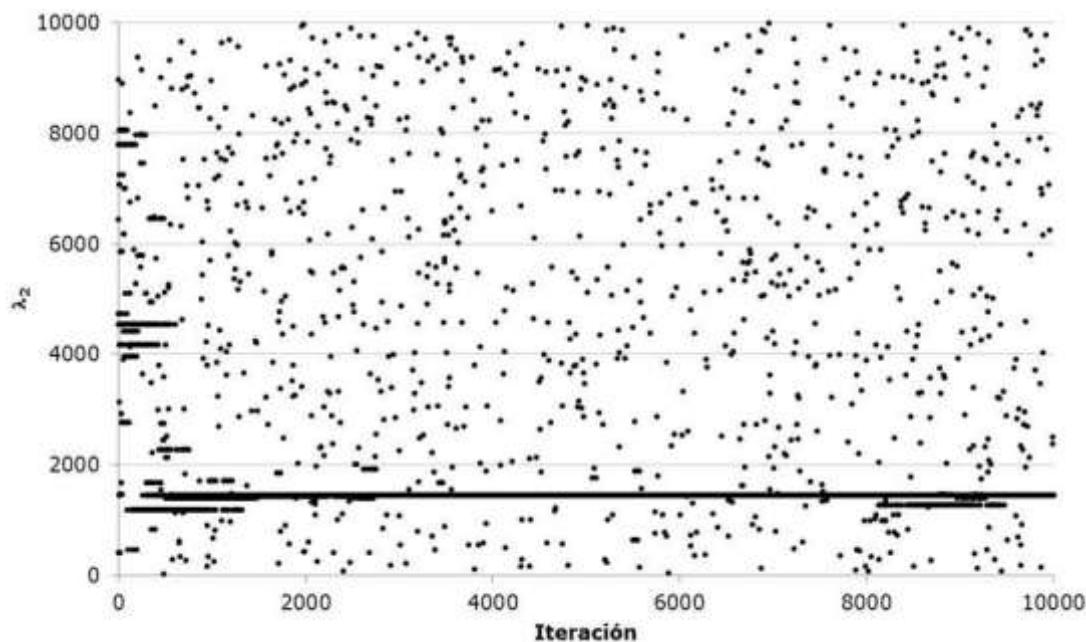


Figura 15. Comportamiento aleatorio del parámetro de ubicación ciclónico.

Gómez *et al.* (2010) presentan un ejemplo clásico del ajuste de numérico de una distribución Gumbel Mixta. Con el objeto de exemplificar el procedimiento propuesto se utilizan los datos históricos de la estación 12504 denominada La Cuña en el estado de Jalisco. Los parámetros del algoritmo fueron $xU=[1, 3000, 3000, 3000, 3000]$, $xL=[1, 0, 0, 0, 0]$, $HMS=30$, $NI=10000$, $HMCR=0.90$, $PAR=0.38$; $B=(xU-xL)/NI$ para iniciar el seudocódigo de la Tabla 2.

La Tabla 7 muestra el valor de los cinco parámetros de ajuste para la función Gumbel Mixta para el registro de caudales máximos anuales de la estación hidrométrica La Cuña, Jalisco. La Figura 16 muestra la precisión en el ajuste de los valores extremos cuando se utiliza la búsqueda armónica en la estimación de los parámetros de la distribución Gumbel Mixta univariada. De la misma forma que para los datos de la estación Santa Cruz, en las Figuras 17 y 18 se muestra la evolución en la convergencia de la función de optimización.

Tabla 7. Parámetros de ajuste para la función Gumbel Mixta para el registro de caudales máximos anuales de la estación hidrométrica La Cuña, Jalisco. México

Metodología	P	λ_1 (m ³ /s)	ε_1 (m ³ /s)	λ_2 (m ³ /s)	ε_2 (m ³ /s)	EE
Gomez <i>et al.</i> 2010	0.8965	292.311	157.143	1271.095	424.786	80.200
Búsqueda Armónica	0.8674	296.599	170.766	713.726	782.383	66.082

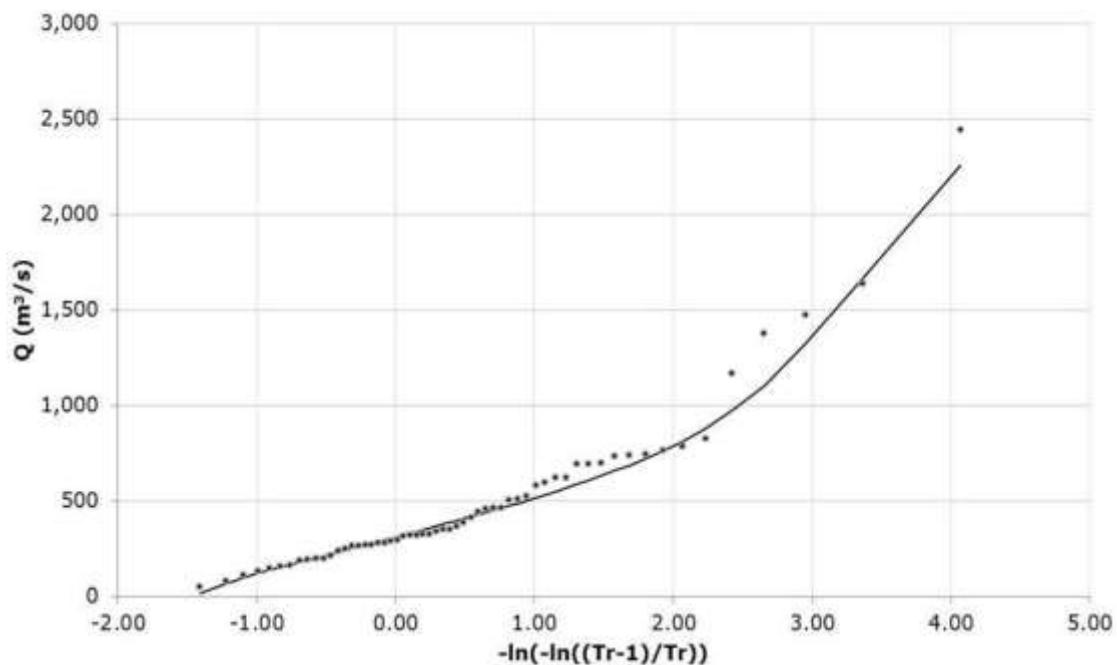


Figura 16. Optimización de la función Gumbel Mixta para la estación La Cuña, Jalisco utilizando el algoritmo de Búsqueda Armónica.

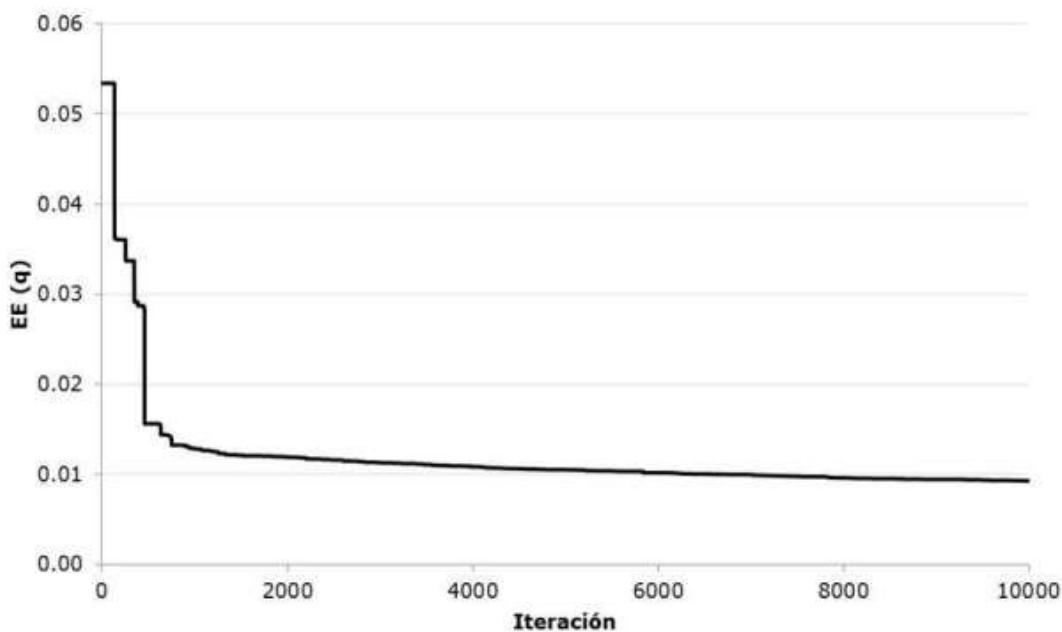


Figura 17. Evaluación de la función de optimización para Gumbel Mixta.

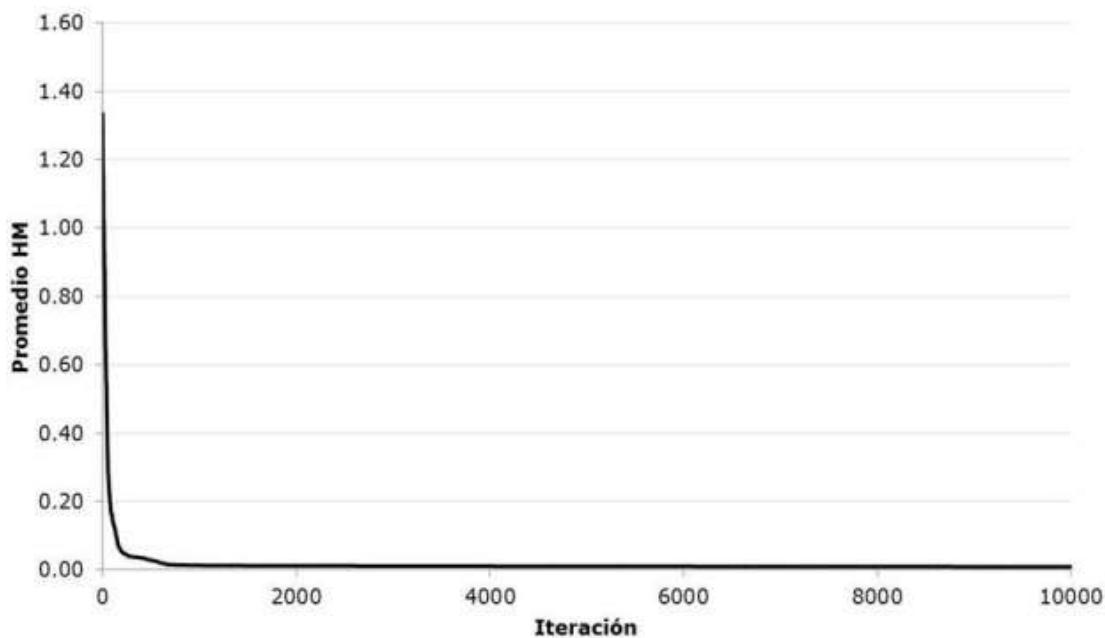


Figura 18. Promedio de las 40 armonías evaluadas en la función de optimización.

Por otro lado, en las Figuras 19 a 23 se muestra el comportamiento aleatorio de los parámetros de ubicación y escala para las muestras ciclónicas y no ciclónicas.

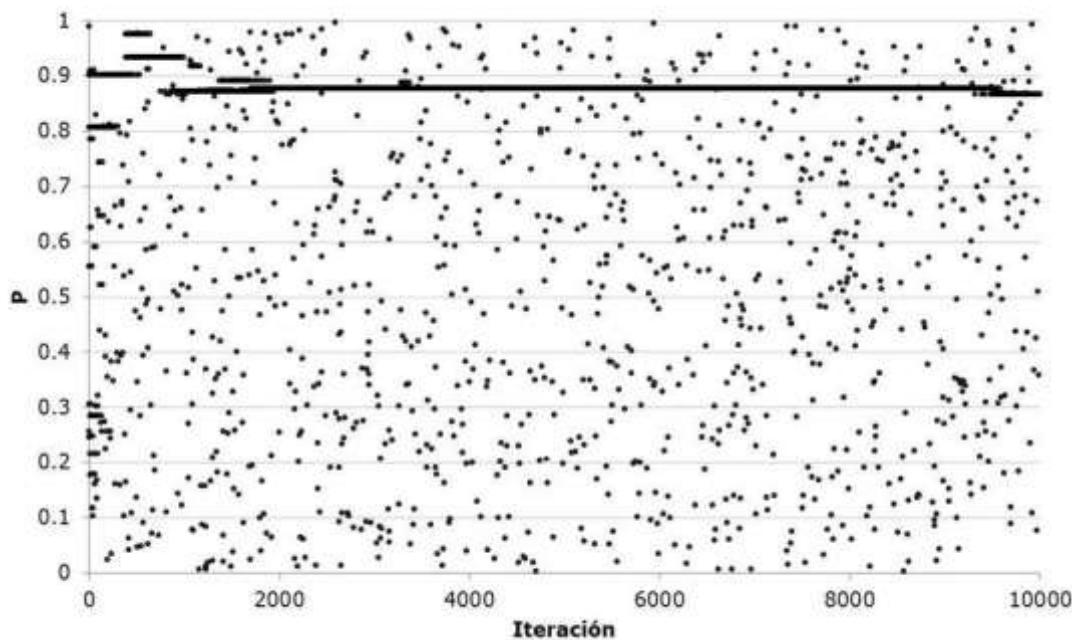


Figura 19. Comportamiento aleatorio de la probabilidad.

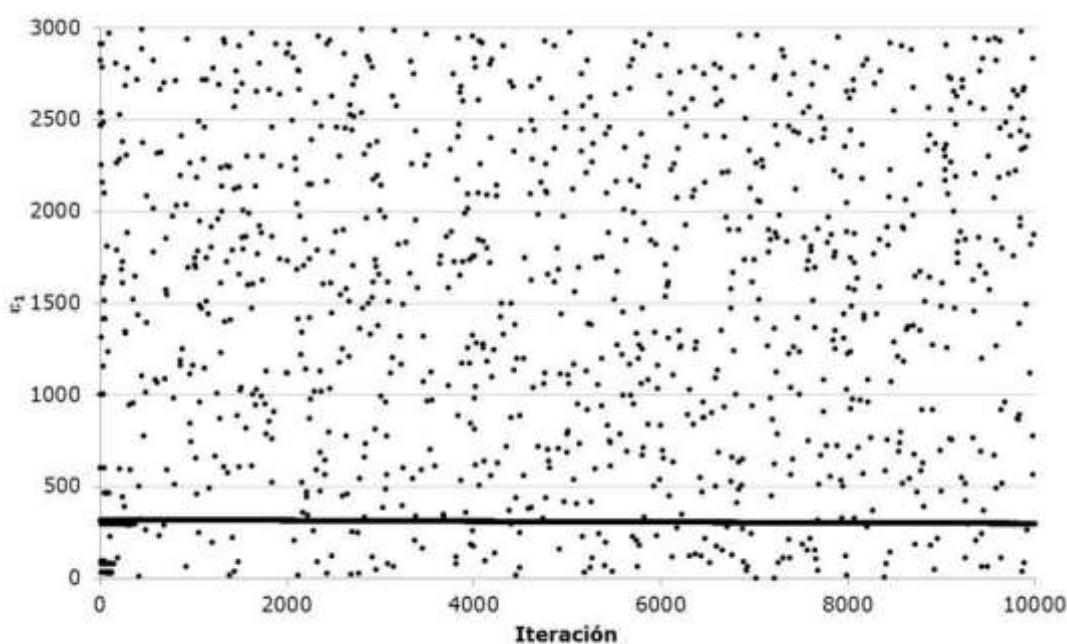


Figura 20. Comportamiento aleatorio del parámetro de escala no ciclónico.

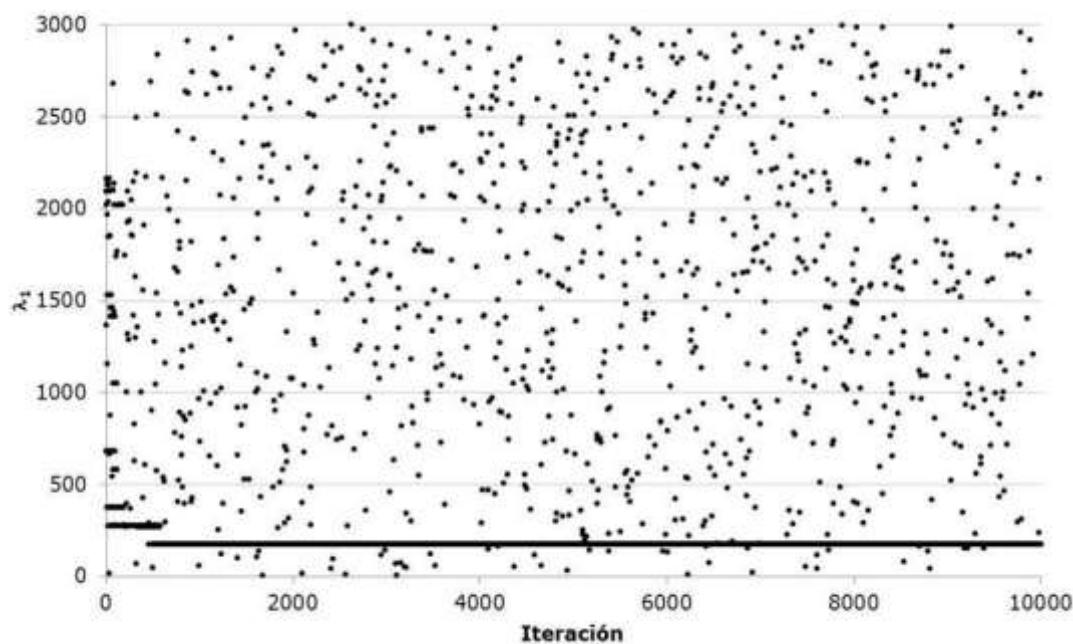


Figura 21. Comportamiento aleatorio del parámetro de ubicación no ciclónico.

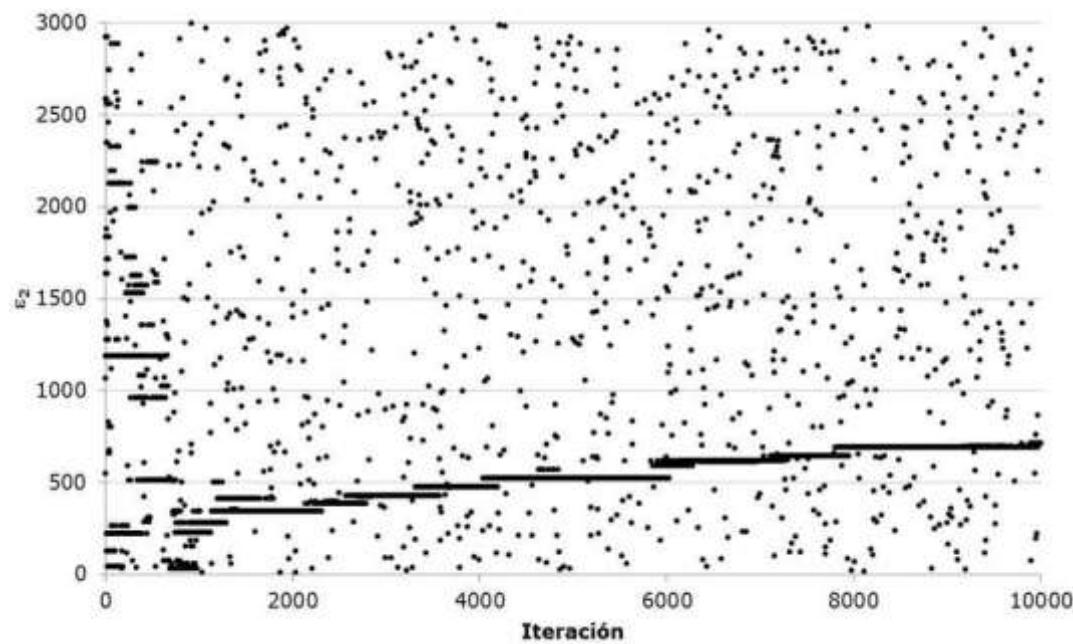


Figura 22. Comportamiento aleatorio del parámetro de escala ciclónico.

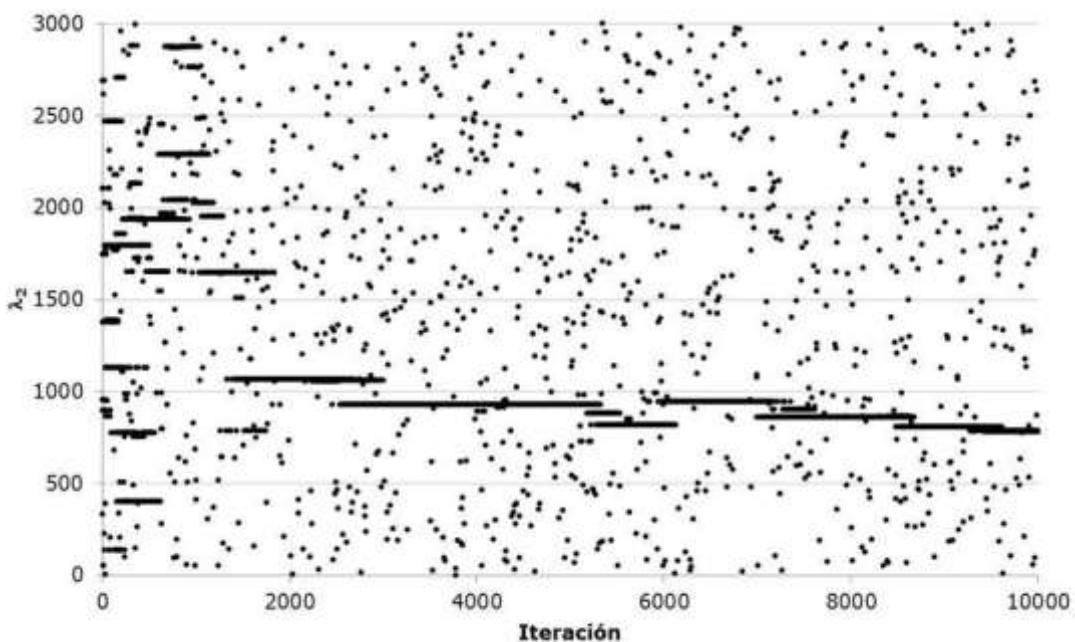


Figura 23. Comportamiento aleatorio del parámetro de ubicación ciclónico.

Discusión

El primer caso de estudio se refirió a la estación hidrométrica Santa Cruz, en la cual se ajustó en primera instancia la función de distribución Gumbel de una población para contrastar el algoritmo de Búsqueda Armónica respecto las metodologías convencionales de Momentos y Máxima Verosimilitud. Se pudo observar una disminución importante en el valor del Error Estándar de Ajuste, aunque gráficamente se aprecia un mejor desempeño del ajuste por Momentos e inclusive por Máxima Verosimilitud. Sin embargo, debe observarse que hubo una disminución en la magnitud de *EE* para el algoritmo de Búsqueda Armónica. Este mejor ajuste se debe principalmente a que se presenta un buen desempeño en la zona de caudales con valor de la variable reducida menor a 1.5. Es decir, no cumple para los caudales históricos

máximos de la serie, generando un error importante de pronóstico en la zona final de la serie.

En cuanto a la convergencia del algoritmo se observa en la Figura 4 que se requiere de muy pocos procesos iterativos para llegar a la solución óptima. De igual manera en la Figura 5 se puede apreciar el promedio de las 40 armonías o combinaciones de valores de las variables evaluadas donde al irse eliminando la peor armonía las 39 restantes empiezan a reducir el valor de la función de optimización al compararse entre ellas, lo que garantiza una rápida convergencia, a pesar de que continua la exploración estocástica del espacio de solución definido por los valores de umbral en las variables de decisión y el número de iteraciones. En las Figuras 6 y 7 se aprecia cómo durante la búsqueda se alinean valores que corresponden a la mejor armonía o solución al problema generándose visualmente una línea de tendencia.

Al desarrollarse la optimización de la función Gumbel Mixta podemos apreciar de nueva cuenta que disminuye el valor del Error Estándar de Ajuste respecto de los métodos de momento y el Algoritmo de Rosenbrock. Con respecto a este último, la variación es mínima en el resultado pero visualmente podemos observar que el ajuste mejora su desempeño en la región de los caudales máximos de la serie pero presenta dificultades en la zona de los caudales mínimos. De igual manera presenta un buen desempeño de convergencia requiriendo aproximadamente 2000 iteraciones de las 10000 realizadas para optimizar los cinco parámetros de la función objetivo, (la evolución de las iteraciones se pueden ver en las Figuras 9 y 10) dado que rápidamente disminuye la diferencia de la ecuación 17 al evaluar las 40 armonías de la memoria armónica, con lo cual los parámetros ε_1 , λ_1 y λ_2 muestran mayor estabilidad en el proceso de exploración estocástica para definir su valor optimizado a diferencia de ε_2 y P siendo este último parámetro el que requiere de más iteraciones para identificar el mínimo en correlación con los restantes.

Finalmente, al incrementar el tamaño de la serie temporal de caudales máximos podemos apreciar en la Figura 16 que el ajuste mediante el algoritmo de Búsqueda Armónica mejora significativamente en la zona de los caudales más pequeños sin descuidar la región de caudales mayores con cual se tiene un valor reducido del Error Estándar de Ajuste. en este caso particular podemos observar mayor estabilidad en la determinación de los parámetros P , ε_1 y λ_1 , los dos últimos asociados a los datos no ciclónicos, mientras que los parámetros ε_2 y λ_2 correspondientes a los datos ciclónicos manifiestan un mayor desequilibrio, particularmente en las primeras 2000 iteraciones reduciéndose paulatina y gradualmente.

Conclusiones

Los resultados muestran que el algoritmo de búsqueda armónica puede optimizar la función de distribución Gumbel tanto de una como de dos poblaciones (Gumbel Mixta), en el primero de los caso sin una diferencia notable respecto a los métodos convencionales de ajuste inclusive sin mejorar la tendencia ordenada mediante la variable reducida. Por otro lado, al considerar las poblaciones no ciclónica y ciclónica se aprecia un incremento importante en el desempeño del algoritmo propuesto para optimizar el ajuste y disminuir el error de ajuste; lo cual se ve potenciado en el uso de series históricas extensas de datos, por lo que su área de oportunidad se localiza precisamente en registros amplios. La Búsqueda Armónica debe ser considerada una opción viable y de aplicación generalizada para la futura estimación de parámetros de distribuciones de probabilidad empleadas en hidrología.

Referencias

- Ahn, K., & Merwade, V. (2016). The effect of land cover change on duration and severity of high and low flows. *Hydrological Processes*, 31(1), 133-149. doi: 10.1002/hyp.10981
- Ahn, K., & Palmer, R. (2016). Use of a nonstationary copula to predict future bivariate low flow frequency in the Connecticut River basin. *Hydrological Processes*, 30(19), 3518-3532. doi: 10.1002/hyp.10876
- Al Aamery, N., Fox, J., Snyder, M., & Chandramouli, C. (2018). Variance analysis of forecasted streamflow maxima in a wet temperate climate. *Journal of Hydrology*, 560, 364-381. doi: 10.1016/j.jhydrol.2018.03.038
- Amin, M., Rizwan, M., & Alazba, A. (2014). Comparison of mixed distribution with EV1 and GEV components for analyzing hydrologic data containing outlier. *Environmental Earth Sciences*, 73(3), 1369-1375. doi: 10.1007/s12665-014-3490-4
- Arnaud, P., Cantet, P., & Odry, J. (2017). Uncertainties of flood frequency estimation approaches based on continuous simulation using

- data resampling. *Journal of Hydrology*, 554, 360-369. doi: 10.1016/j.jhydrol.2017.09.011
- Ashkar, F., El-Jabi, N., & Sarraf, S. (1991). Study of hydrological phenomena by extreme value theory. *Natural Hazards*, 4(4), 373-388. doi: 10.1007/bf00126645
- Ayantobo, O., Li, Y., Song, S., Javed, T., & Yao, N. (2018). Probabilistic modelling of drought events in China via 2-dimensional joint copula. *Journal of Hydrology*, 559, 373-391. doi: 10.1016/j.jhydrol.2018.02.022
- Bardsley, W. (2016). Cautionary note on multicomponent flood distributions for annual maxima. *Hydrological Processes*, 30(20), 3730-3732. doi: 10.1002/hyp.10886
- Borga, M., Vezzani, C., & Fontana, G. (2005). Regional Rainfall Depth-Duration-Frequency Equations for an Alpine Region. *Natural Hazards*, 36(1-2), 221-235. doi: 10.1007/s11069-004-4550-y
- Campos-Aranda, D. F. (2003). *Introducción a los Métodos Numéricos: Software en Basic y aplicaciones en hidrología superficial*. Capítulo 5: Ajuste de curvas, pp. 93-127 y Capítulo 9. Optimización numérica, pp. 172-211, San Luis Potosí, Librería Universitaria Potosina.
- Campos-Aranda, F. (1989). Estimación de los parámetros óptimos de la distribución Gumbel mixta por medio del algoritmo de Rosenbrock. *Tecnología y Ciencias del Agua*, vol. Ene-Abr., pp. 9-18
- Campos-Aranda, F. (1994). Aplicación del Método del Índice de Crecientes en la Región Hidrológica Número 10, Sinaloa. *Tecnología y Ciencias del Agua*, vol. IX, núm. 3, pp. 41-55
- Chen, L., & Singh, V. (2018). Entropy-based derivation of generalized distributions for hydrometeorological frequency analysis. *Journal of Hydrology*, 557, 699-712. doi: 10.1016/j.jhydrol.2017.12.066
- Chow, V. T. (1964). *Statistical and probability analysis of hydrologic data*. Section 8, part I frequency analysis. In: *Handbook of applied hydrology*. McGraw-Hill, Nueva York.
- Cuevas, E., Ortega-Sánchez, N. (2013). El algoritmo de búsqueda armónica y sus usos en el procesamiento digital de imágenes. *Computación y sistemas* vol. 17 No. 4, pp. 543-560. Doi: 10.13053/CyS-17-4-2013-007
- Díaz-Delgado C., Bâ K., y Trujillo-Flores E. (1999). Las funciones Beta-Jacobi y Gamma-Laguerre como métodos de análisis de valores hidrológicos extremos. Caso de precipitaciones máximas anuales. *Tecnología y Ciencias del Agua*, vol. XIV, núm. 2, pp. 39-48
- Escalante-Sandoval, C. (2007). Application of bivariate extreme value distribution to flood frequency analysis: a case study of Northwestern

Mexico. *Natural Hazards*, 42(1), 37-46. doi: 10.1007/s11069-006-9044-7

Escalante-Sandoval, C. (2013). Estimation of Extreme Wind Speeds by Using Mixed Distributions. *Ingeniería Investigación y Tecnología*, XIV(2), 153-162

Escalante-Sandoval, C., y Domínguez-Esquivel, J. (2001). Análisis regional de precipitación con base en una distribución bivariada ajustada por máxima entropía. *Tecnología y Ciencias del Agua*, vol. XVI, núm. 3, pp. 91-102

Fogel, L., Owens, A., & Walsh, M. (1966). *Artificial intelligence through simulated evolution*. Chichester, WS, UK: Wiley.

Geem, Z. W., Kim, J. H., & Loganathan, G. V. (2001). A new heuristic optimization algorithm: Harmony search. *Simulation*, 76(2), 60-68. doi: 10.1177/003754970107600201

Glover, F. (1977). Heuristics for integer programming using surrogate constraints. *Decision Sciences*, 8(1), 156-166. doi: 10.1111/j.1540-5915.1977.tb01074.x

Gómez, J., Aparicio, J., y Patiño, C. (2010). *Manual de análisis de frecuencias en hidrología*. Instituto Mexicano de Tecnología del Agua. Jiutepc Morelos.

Grünthal, G., Thielen, A., Schwarz, J., Radtke, K., Smolka, A., & Merz, B. (2006). Comparative Risk Assessments for the City of Cologne – Storms, Floods, Earthquakes. *Natural Hazards*, 38(1-2), 21-44. doi: 10.1007/s11069-005-8598-0

Gumbel, E. J. (1960). *Distributions des valeurs extremes en plusieurs dimensions*, Publications de L'Institute de Statistique, Paris, Vol. 9, 171-173.

Gutiérrez-López A., y Ramírez, A. (2005). Predicción hidrológica mediante el Método de la Avenida Índice para dos poblaciones. *Tecnología y Ciencias del Agua*, vol. XX, núm. 2, pp. 37-47

Holland, J.H. (1975) *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor. (2nd Edition, MIT Press, 1992.)

Huang, J., Li, Y., Yin, J., & Mcbean, E. (2016). Precipitation regional extreme mapping as a tool for ungauged areas and the assessment of climate changes. *Hydrological Processes*, 30(12), 1940-1954. doi: 10.1002/hyp.10743

Jenkinson, A. F. (1969). *Estimation of maximum floods*. World Meteorological Organization, Technical Note No. 98, ch. 5, 183-257.

Joshi, S., Jayadeva, Ramakrishnan, G., & Chandra, S. (2012). Using Sequential Unconstrained Minimization Techniques to simplify SVM

solvers. *Neurocomputing*, 77(1), 253-260. doi: 10.1016/j.neucom.2011.07.010

Karim, F., Petheram, C., Marvanek, S., Ticehurst, C., Wallace, J., & Hasan, M. (2015). Impact of climate change on floodplain inundation and hydrological connectivity between wetlands and rivers in a tropical river catchment. *Hydrological Processes*, 30(10), 1574-1593. doi: 10.1002/hyp.10714

Kirkpatrick, S., Gelatt, C., & Vecchi, M. (1983). Optimization by Simulated Annealing. *Science*, 220(4598), 671-680. doi: 10.1126/science.220.4598.671

Koutsoyiannis, D. (2004). On the appropriateness of the Gumbel distribution for modelling extreme rainfall (solicited). *Hydrological Risk: recent advances in peak river flow modelling, prediction and real-time forecasting. Assessment of the impacts of land-use and climate changes*. Edited by A. Brath, A. Montanari, and E. Toth, Bologna, 303-319, doi:10.13140/RG.2.1.3811.6080, Editoriale Bios, Castrolibero, Italy.

Kuester, J., & Mize, J. (1973). *Optimization Techniques with Fortran*, ROSENBL algorithm, New York, McGraw-Hill Book Co. pp:320-330

Lee, K., & Geem, Z. (2005). A new meta-heuristic algorithm for continuous engineering optimization: harmony search theory and practice. *Computer Methods in Applied Mechanics and Engineering*, 194(36-38), 3902-3933. doi: 10.1016/j.cma.2004.09.007

Lilienthal, J., Fried, R., & Schumann, A. (2018). Homogeneity testing for skewed and cross-correlated data in regional flood frequency analysis. *Journal of Hydrology*, 556, 557-571. doi: 10.1016/j.jhydrol.2017.10.056

Mélèse, V., Blanchet, J., & Molinié, G. (2018). Uncertainty estimation of Intensity–Duration–Frequency relationships: A regional analysis. *Journal of Hydrology*, 558, 579-591. doi: 10.1016/j.jhydrol.2017.07.054

Montaseri, M., Amirataee, B., & Rezaie, H. (2018). New approach in bivariate drought duration and severity analysis. *Journal of Hydrology*, 559, 166-181. doi: 10.1016/j.jhydrol.2018.02.018

Öztürk, S., Bayrak, Y., Çınar, H., Koravos, G., & Tsapanos, T. (2008). A quantitative appraisal of earthquake hazard parameters computed from Gumbel I method for different regions in and around Turkey. *Natural Hazards*, 47(3), 471-495. doi: 10.1007/s11069-008-9234-6

Phillips, R., Samadi, S., & Meadows, M. (2018). How extreme was the October 2015 flood in the Carolinas? An assessment of flood frequency analysis and distribution tails. *Journal of Hydrology*, 562, 648-663. doi: 10.1016/j.jhydrol.2018.05.035

- Qi, W., & Liu, J. (2018). A non-stationary cost-benefit based bivariate extreme flood estimation approach. *Journal of Hydrology*, 557, 589-599. doi: 10.1016/j.jhydrol.2017.12.045
- Qian, L., Wang, H., Dang, S., Wang, C., Jiao, Z., & Zhao, Y. (2017). Modelling bivariate extreme precipitation distribution for data-scarce regions using Gumbel-Hougaard copula with maximum entropy estimation. *Hydrological Processes*, 32(2), 212-227. doi: 10.1002/hyp.11406
- Raynal-Villaseñor, J. (1997). Sobre el uso del dominio de atracción para la identificación de distribuciones de valores extremos para máximos. *Tecnología y Ciencias del Agua*, vol. XII, núm. 2, pp. 57-62
- Rosenbrock, H. H. (1960). An automatic method for finding the greatest or least value of a function. *The Computer Journal*, 3:175-184. doi:10.1093/comjnl/3.3.175
- Rossi, F., Florentino, M., & Versace, P. (1984). Two-Component Extreme Value Distribution for Flood Frequency Analysis. *Water Resources Research* 20(7). 847-856. doi:10.1029/WR020i007p00847.
- Schulz, K., & Bernhardt, M. (2016). The end of trend estimation for extreme floods under climate change?. *Hydrological Processes*, 30(11), 1804-1808. doi: 10.1002/hyp.10816
- Smitha, P., Narasimhan, B., Sudheer, K., & Annamalai, H. (2018). An improved bias correction method of daily rainfall data using a sliding window technique for climate change impact assessment. *Journal of Hydrology*, 556, 100-118. doi: 10.1016/j.jhydrol.2017.11.010
- So, B., Kim, J., Kwon, H., & Lima, C. (2017). Stochastic extreme downscaling model for an assessment of changes in rainfall intensity-duration-frequency curves over South Korea using multiple regional climate models. *Journal of Hydrology*, 553, 321-337. doi: 10.1016/j.jhydrol.2017.07.061
- Strupczewski, W., Kochanek, K., & Bogdanowicz, E. (2017). Historical floods in flood frequency analysis: Is this game worth the candle?. *Journal of Hydrology*, 554, 800-816. doi: 10.1016/j.jhydrol.2017.09.034
- Tao, S., Dong, S., Wang, N., & Guedes Soares, C. (2013). Estimating storm surge intensity with Poisson bivariate maximum entropy distributions based on copulas. *Natural Hazards*, 68(2), 791-807. doi: 10.1007/s11069-013-0654-6
- Tosunoglu, F., & Can, I. (2016). Application of copulas for regional bivariate frequency analysis of meteorological droughts in Turkey. *Natural Hazards*, 82(3), 1457-1477. doi: 10.1007/s11069-016-2253-9

- Van de Vyver, H. (2018). A multiscaling-based intensity-duration-frequency model for extreme precipitation. *Hydrological Processes*. doi: 10.1002/hyp.11516
- Vu, M., Raghavan, V., & Liong, S. (2016). Deriving short-duration rainfall IDF curves from a regional climate model. *Natural Hazards*, 85(3), 1877-1891. doi: 10.1007/s11069-016-2670-9
- Yan, L., Xiong, L., Liu, D., Hu, T., & Xu, C. (2016). Frequency analysis of nonstationary annual maximum flood series using the time-varying two-component mixture distributions. *Hydrological Processes*, 31(1), 69-89. doi: 10.1002/hyp.10965
- Yang, J., Chang, J., Wang, Y., Li, Y., Hu, H., & Chen, Y. et al. (2018). Comprehensive drought characteristics analysis based on a nonlinear multivariate drought index. *Journal of Hydrology*, 557, 651-667. doi: 10.1016/j.jhydrol.2017.12.055
- Yue, S. (2000). The Gumbel Mixed Model Applied to Storm Frequency Analysis. *Water Resources Management*, 14(5), 377-389. doi: 10.1023/a:1011124423923
- Zhang, R., Chen, X., Wang, H., Cheng, Q., Zhang, Z., & Shi, P. (2017). Temporal change of spatial heterogeneity and its effect on regional trend of annual precipitation heterogeneity indices. *Hydrological Processes*, 31(18), 3178-3190. DOI: 10.1002/hyp.11225

Harmony search to optimize the univariate Gumbel Mixed Model

Abstract

Phenomena analyzed in the field of engineering with frequency analyses are usually extreme events such as floods and droughts. Such events frequently have as origin the simultaneous occurrence of two phenomena, and therefore are analyzed with two simultaneous probability distributions. Such is the case of the climatological and hydrometric series in Mexico which, due to its exposure to phenomena such as hurricanes, are analyzed with two-population probability distributions. The good use of these distributions depends on the correct estimation of their parameters. The application of a harmonic search meta-heuristic algorithm for the estimation of the parameters that optimize the univariate Gumbel Mixed function is presented. Annual maximum hydrometric data are used to compare the best-fit of the univariate distribution with traditional methodologies, such as maximum likelihood and the Rosenbrock algorithm. The results show that there is a decrease in the root mean square error and a fast convergence when a harmonic search algorithm is used. With the decrease of the error of fit, estimation of the design flows values is improved. The pseudocode of the algorithm for the implementation of the proposed methodology is presented.

Keywords: Harmony Search, Rosenbrock, Optimization, Meta heuristic, Gumbel Mixed Model, Flow frequency analysis.

Posted by invitation

Introduction

The general objective of a frequency analysis is to use the historical data of some hydrological variable such as precipitation or flow to estimate the regime of a hydrological event associated with a return period using different probability distributions. In the field of civil engineering, the selection of the return period depends on the nature of the project and the risk that is to be assumed (Escalante-Sandoval, 2007). Most frequency analyzes are performed with univariate distributions, that is, with a single analysis variable that fits a probability function with two or more parameters (Ashkar *et al.*, 1991). The phenomena that are analyzed in the field of engineering with a frequency analysis are diverse. Usually, the analysis of hydrological extremes such as floods (Arnaud *et al.*, 2017, Al Aamery *et al.*, 2018) and droughts are the most studied (Ahn and Merwade, 2016, Ayantobo, *et al.*, 2018, Yang *et al.*, 2018). However, studies of the El Niño phenomenon (Karim, *et al.*, 2015), climate change (Schulz & Bernhardt, 2016, Smitha *et al.*, 2018) or even seismic events are studied with probability distributions (Grünthal *et al.*, 2006; Öztürk *et al.*, 2008). Currently there are many functions that are used in hydrology and thanks to the operational disposition of modern processors it is possible to use probabilistic models that consider more than three parameters for their solution, which increases the precision when there are historical samples with more of 50 years of records (Díaz-Delgado *et al.*, 1999). In Mexico and in general in Latin America the use of the Gumbel distribution (Gumbel, 1960) is very widespread, since it is a versatile and easy to adjust distribution. For example, the Gumbel distribution is commonly used in the estimation of intensity-duration-return period curves (Borga *et al.*, 2005; So *et al.*, 2017; Mélèse *et al.*, 2018) and in the validation of these curves with atmospheric models (Vu *et al.*, 2016; Van de Vyver, 2018).

Hydrological regionalization is a set of methodologies that are used to estimate design events at unmeasured sites or with scarce information; these techniques also use Gumbel univariate probability distribution function (Huang *et al.*, 2016; Zhang *et al.*, 2017). Such is the case of the Index Flood Method (Campos-Aranda, 1994), which uses the Gumbel distribution to transfer hydrometric information within a region considered hydrologically homogeneous (Lilienthal *et al.*, 2018). Such has been the use of this Method that has been modified so that it can be used in regions where there is direct affectation due to extreme natural phenomena (Gutiérrez-López & Ramírez, 2005). For example, the coasts of the Mexican Republic are affected every year by hurricanes that generate out of normal rainfall; that is, in these zones a climatological or hydrometric record will be formed by two types of data: the measurements that come from the events of the normal rainy season and by the records that come from the extreme events (Phillips *et al.*, 2018; Yan *et al.*, 2016). This creates the need to analyze a

variable with two probability distributions simultaneously; this is how the univariate Gumbel distribution appears for two populations or also known as Univariate Mixed Gumbel (Rossi *et al.*, 1984, Yue, 2000). However, hydrological problems are increasingly complex and studies are needed that can simultaneously employ not only two probability distributions, but also two analysis variables; to what is known as bivariate mixed distribution. For example, in the case of Gumbel, it would be a mixed bivariate Gumbel distribution, which can simultaneously estimate the volumes and maximum flow of a hydrograph (Qi & Liu, 2018) or estimate the duration and sheet of a storm (Qian *et al.*, 2017). In Mexico, the forerunner in the study and application of bivariate and trivariate distributions was Escalante-Sandoval (2007), who developed the theoretical approach for the solution of mixed distributions, as well as its estimation of parameters. The estimation of parameters of a bivariate distribution is complex and requires fulfilling four conditions: (i) the marginal distributions must be extreme asymptotic; (ii) the distribution must be stable for the largest values in the sample; (iii) there must be a density function and (iv) the trivial case where the multivariate distribution is the product of its extreme marginal distributions is eliminated (Escalante-Sandoval & Domínguez-Esquivel, 2001).

In an univariate or bivariate probability distribution, whatever the case, the estimation of its parameters is one of the most extensive and complex topics of modern hydrology and requires an exhaustive analysis (Strupczewski *et al.*, 2017; Tosunoglu & Can, 2016). In the beginning, the estimation of parameters was done with the technique of least squares, heavy means, moments and recently the theory of maximum likelihood was perfected. Currently optimization algorithms or the maximum entropy technique are undoubtedly the most accurate procedures, even for the distributions of maximum and minimum extremes in their mixed bivariate versions (Chen & Singh, 2018, Montaseri *et al.*, 2018, Ahn & Palmer, 2016; Tao *et al.*, 2013). In addition to the selection of the type of distribution to be used, it is necessary to perfect the appropriate technique for parameter estimation. That is why we resort to optimization processes which are increasingly applied to engineering problems (Escalante-Sandoval, 2013). Currently, a wide variety of algorithms are used, many of them based on linear and non-linear numerical programming methods, which require calculating gradients to look for the solution in the vicinity of an initial data; this becomes a useful strategy in obtaining the global optimum through simple models (Lee & Geem, 2005). Although modern hydrology advances towards artificial intelligence, many of the real problems of engineering optimization are complex in nature and difficult to solve using such algorithms (Bardsley, 2016). In particular, the search for a local optimum where the result depends on the

selection of the initial data, the boundary conditions and above all to find the optimal solution in a global solution space is difficult if the objective function and its restrictions present numerical instabilities. From the above, it is evident that work should continue to be done on the optimization of the models and solution schemes of univariate or bivariate probability distributions. In this study we present a harmonic search algorithm for optimizing the parameters of the Univariate Mixed Gumbel function. The results are compared with the algorithms of Rosenbrock and the estimation methods of Moment parameters and Maximum Likelihood, being the harmonic search a new option to minimize the standard error of adjustment.

Methods

Metaheuristic Algorithms

The computational disadvantages of the existing numerical models motivated the development of metaheuristic algorithms based on simulations of natural or artificial phenomena to solve engineering optimization problems whose common factor is their combination of randomness rules. Metaheuristic algorithms are used to describe very varied phenomena. For example, evolutionary algorithms (Fogel *et al.*, 1966) and genetic algorithms (Holland, 1975) are used to reveal processes of biological evolution. Glover (1977) presented algorithms that are known as taboo search to explain animal behavior. There are other complex processes such as simulation called simulated annealing (*Simulated Annealing*, SA), which simulates and optimizes a function in a large search space, when the values are mainly discrete (Kirkpatrick *et al.*, 1983). Similarly, musical performance defined as the art of performing a musical work using instruments uses harmonic search algorithms (Geem *et al.*, 2001).

Generalized extreme value (GEV) distribution

In 1960 Gumbel developed the theory of distribution of extreme values in which the function will converge to an asymptote of the three possible ones as the quantity of data increases in the series of maximums (Koutsoyiannis, 2004; Amin *et al.*, 2014). The asymptotes are represented by the expression known as generalized extreme value (GEV) distribution (Jenkinson, 1969) that uses three parameters: location (ε), scale (λ) and form (k).

$$F(x_i) = e^{\left\{-\left[1 + \frac{x_i - \varepsilon}{\lambda} k\right]^{-\frac{1}{k}}\right\}} \quad (1)$$

The probability density function (*fdp*) is of the form:

$$f(x_i) = \frac{dF(x_i)}{dx} = \frac{1}{\lambda} e^{\left\{-\left[1 - \frac{x_i - \varepsilon}{\lambda} k\right]^{\frac{1}{k}}\right\}} \left[1 - \frac{x_i - \varepsilon}{\lambda} k\right]^{\frac{1}{k}-1} \quad (2)$$

It is called GEV Type I (Gumbel) when $k = 0$, $-\infty < x_i < \infty$, GEV Type II (Frechet) when $k < 0$, $\varepsilon + \lambda/k < x_i < \infty$ and GEV Type III (Weibull) when $k > 0$, $-\infty < x_i < \varepsilon + \lambda/k$. In all three cases $\lambda > 0$.

In the same way, the reduced variable can be established for this *fdp*

$$y_i = -\ln \left[1 - \frac{x_i - \varepsilon}{\lambda} k\right]^{\frac{1}{k}} \quad (3)$$

Gumbel distribution function

When the parameter of the form in the generalized extreme value (GEV) distribution is zero, we have the Type I distribution, also known as the Gumbel distribution function and the shape of the cumulative distribution function in terms of $-\infty < x_i < \infty$, $-\infty < \beta < \infty$ and $\alpha > 0$ is

$$F(x_i) = e^{\left\{-e^{\left[-\frac{x_i-\varepsilon}{\lambda}\right]}\right\}} \quad (4)$$

Your respective *fdp* is

$$f(x_i) = \frac{dF(x_i)}{dx} = \frac{1}{\lambda} \left[-e^{\left\{-\left[\frac{x_i-\varepsilon}{\lambda}\right]\right\}} e^{-e^{-\left[\frac{x_i-\varepsilon}{\lambda}\right]}} \right] \quad (5)$$

and its reduced variable

$$y_i = \frac{x_i-\varepsilon}{\lambda} \quad (6)$$

Thanks to its behavior, this fdp is widely used in many regions for the statistical analysis of both flows and annual maximum rainfall.

Gumbel Mixed distribution function

The mathematical expression that represents the probability distribution function for two mixed populations considers the estimation of the probability of non-exceedance of the maximum annual expenditure (x), the probability (P) of the occurrence of non-cyclonic events, two location parameters (λ_1, λ_2) and two scale parameters ($\varepsilon_1, \varepsilon_2$) for which the subscripts 1 are associated with the non-cyclonic population while the subscripts 2 are related to the cyclonic population:

$$P(X \leq x_i) = F(x_i) \quad (7)$$

$$F(x_i) = Pe^{-e^{\left\{-\left[\frac{x_i-\varepsilon_1}{\lambda_1}\right]\right\}}} + (1-P)e^{-e^{\left\{-\left[\frac{x_i-\varepsilon_2}{\lambda_2}\right]\right\}}} \quad (8)$$

the corresponding *fdp* is:

$$f(x_i) = \frac{dF(x_i)}{dx} = \frac{P}{\lambda_1} e^{\left[-\left[\frac{x_i-\varepsilon_1}{\lambda_1}\right]-e^{\left[-\frac{x_i-\varepsilon_1}{\lambda_1}\right]}\right]} + \frac{(1-P)}{\lambda_2} e^{\left[-\left[\frac{x_i-\varepsilon_2}{\lambda_2}\right]-e^{\left[-\frac{x_i-\varepsilon_2}{\lambda_2}\right]}\right]} \quad (9)$$

The estimation of the adjustment parameters of equations 5 and 9 can be performed by conventional maximum likelihood, moments or

statistical methods; however, the efficiency decreases with the increase of parameters to be adjusted. To avoid this, the expression of the sum of the heavy quadratic errors E of the estimated values $F(x_i)$ with respect to the empirical values $F'(x_i)$ is minimized by applying the Weibull formulation

$$E = \sum_{i=1}^n [F'(x_i) - F(x_i)]^2 w_i \quad (10)$$

$$F'(x_i) = \frac{m}{n+1} \quad (11)$$

In equation 10 we use the weight assigned to the generated error (w_i) that corresponds to the estimation by means of the distribution function, while in equation 11 we consider the increasing order m of the annual maximum flow (x) and the total flow in study n .

The minimization of the sum of heavy quadratic errors has been carried out by the method of maximum ascent (Joshi *et al.*, 2012) which is a gradient technique based on evaluations of the function to be minimized and its derivatives (Raynal-Villaseñor, 1997). However, this process is slow, since it requires frequent changes of direction, with which it becomes inefficient from the computational point of view and fails to consider the second derivatives.

Rosenbrock Algorithm

To minimize E in equation 10, the Rosenbrock algorithm can be applied for multiple unrestricted variables (Rosenbrock, 1960, Kuester & Mize, 1973, Campos-Aranda, 2003). This technique is cataloged as a direct search. To apply it, initial values of the parameters are required, for which the calculation is made as if it were a single population using the moments technique (Chow, 1964) considering the value of the mean (x_1, x_2), the standard deviation (S_1, S_2) of the two populations, not cyclonic and cyclonic respectively, as well as the number of flows of cyclonic origin (n_c) considered. The parameters can then be estimated as:

$$\varepsilon_1 = \bar{x}_1 - 0.5772\lambda_1 \quad (12)$$

$$\lambda_1 = \frac{\sqrt{6}}{\pi} S_1 \quad (13)$$

$$\varepsilon_2 = \bar{x}_2 - 0.5772\lambda_2 \quad (14)$$

$$\lambda_2 = \frac{\sqrt{6}}{\pi} s_2 \quad (15)$$

$$P = \frac{n-n_c}{n} \quad (16)$$

The goodness of adjustment through the Standard Adjustment Error (*EE*) proposed by Raynal (1997) considers the comparison of the registered annual maximum flows $F'(x_i)$ and estimates $F(x_i)$ as well as the comparison of the quantity of flows (*n*) with respect to the adjustment parameters (n_p) of the distribution function:

$$EE = \left[\frac{[F'(x_i) - F(x_i)]^2}{n - n_p} \right]^{\frac{1}{2}} \quad (17)$$

This algorithm allows us to minimize equation 17, which considers equations 5 and 9 in its structure, which is a non-linear function of multiple unrestricted variables of the form $F(x_1, x_2, \dots, x_n)$ running the pseudocode shown in Table 1. The detail and programming code can be consulted in Campos-Aranda (1989).

Table 1. Pseudocode of the Rosenbrock algorithm

Step	Activity
1	Define the initial parameters initial point (x_i), stop criterion ($(C_p > 0)$), expansion factor ($f_E > 1$) and contraction factor ($(-1 < f_c < 0)$)
2	For $j = 1, \dots, n$ consider $d_j = e^j$, the unit vector of the variable (x_j) and the step lengths in each of the directions (h_j), taking $y_1 = x_1$ being $k = 1$. Go to step 3.
3	Evaluate $f(y_j + h_j d_j)$; if $f(y_j + h_j d_j) < f(y_j)$ then $y_{j+1} = y_j + f_E h_j d_j$; otherwise, $y_{j+1} = y_j + f_c h_j d_j$. Go to step 4.
4	If $j < n$ then $j = j + 1$ and return to step 2; otherwise, go to step 5.
5	Calculate $f(y_{n+1})$ and check the following conditions.

If $f(y_{n+1}) < f(y_1)$ then $y_1 = y_{n+1}$ making $j = 1$ and return to step 2.

If $f(y_{n+1}) < f(y_1)$ and $f(y_{n+1}) < f(x_k)$ go to step 6.

If $f(y_{n+1}) < f(y_1)$ and $f(y_{n+1}) = f(x_k)$ select one of the following two cases, the one that corresponds.

If $|h_j| < c_p$ then x_k is approximation of the optimum for the objective function and concludes the algorithm.

Otherwise $y_1 = y_{n+1}$ assigning $j = 1$. Go to step 3.

6 Consider $x_{k+1} = y_{n+1}$ and check the following conditions.

If $\|x_{k+1}\| < c_p$ then x_{k+1} is approximation of the optimum for the objective function and concludes the algorithm. Otherwise calculate $\lambda_1, \dots, \lambda_n$ of the relationship $x_{k+1} - x_k = \sum_{j=1}^n \lambda_j d_j$ forming a new set of directions $\bar{d}_1, \dots, \bar{d}_n$ in the way

$$w_j = \begin{cases} d_j & \lambda_j = 0 \\ \sum_{j=1}^n \lambda_j d_j & \lambda_j \neq 0 \end{cases}$$

$$y_j = \begin{cases} w_j & j = 1 \\ w_j - \sum_{i=1}^{j-1} (w_j^t \bar{d}_i) \bar{d}_i & j \geq 2 \end{cases}$$

Taking $\bar{d}_j = \frac{y_j}{\|y_j\|}$. Consider the lengths of step 2, take $y_1 = x_{k+1}$, with $k = k + 1$ and $j = 1$ go back to step 2.

Harmonic Search Algorithm

This metaheuristic algorithm conceptualizes the use of musical processes to perfect harmonies which in an analogous way in optimization techniques are solution vectors and improvisations made by musicians that allude to local or global search schemes (Geem et al., 2001).

It does not require initial values for the decision variables, that is, instead of a gradient search it performs a random stochastic search based on the rate of consideration of harmonic memory and the pitch adjustment rate which eliminates the use of information from derived, reducing the mathematical requirements and facilitating their implementation in problems of optimization in engineering problems (Lee & Geem, 2005). The pseudocode is shown in Table 2.

To perform the optimization of the objective function (CF) candidate solutions are considered that constitute the harmonies and are represented mathematically by n-dimensional vectors; these solutions constitute the initial population generated in a random way and stored in the harmonic memory (HM), from which a new harmony is generated starting from the elements contained by stochastic reinitiation or by adjusting the tone of an existing vector. With this, HM is updated by comparing the new candidate solution with respect to the worst stored solution, replacing it if the solution of the objective function was improved and, otherwise, the candidate solution is eliminated. This process is carried out until reaching the criterion of established unemployment (Cuevas & Ortega-Sánchez, 2013).

Each candidate solution is within the range defined by the upper (xU) and lower (xL) limits called the bandwidth ($B = xU - xL$) of the decision variables (N), the number of vectors in HM defines the size of the memory (HMS) which stores the best harmonic performances for the objective function, on the other hand when punctuating the rate of consideration of harmonic memory for the adjustment of tone ($HMCR$) the possibility of the new harmonies is established by taking elements stored in HM and the tone adjustment rate (PAR) gives the possibility of modifying a previous solution without exceeding the maximum magnitude B , all of the above within a process defined by the total number of improvisations (NI) that correspond to the proposed iterations.

To guarantee the solution maintaining specific characteristics of the problem, it is possible to establish restriction functions (RF) with penalty coefficients (PC) that correct the CF value prior to the HM update.

Table 2. Pseudocode of the Harmony Search Algorithm.

Step	Activity
1	Define CF and RF according to the problem, set HMS , $HMCR$, B , PAR , NI and PC
2	Define $HM = [x_i(j)]$ for $x_i(j) = xL(j) + [xU(j) - xL(j)] * rand(0,1)$ where $i = 1, 2, \dots, HMS$ and $j = 1, 2, \dots, N$
3	For $j = 1, 2, \dots, N$ if $rand_1(0,1) < HMCR$ then $x_{new}(j) = x_i(j)$ continue to step 4, otherwise go to step 5.
4	If $rand_2(0,1) < PAR$ then $x_{new}(j) = x_{new}(j) + B * rand_3(0,1)$, in case $x_{new}(j) = x_i(j)$.
5	Update the harmonic memory if $f(x_{new}) + PC * f(x_{new}) < f(x_w)$ then $x_w = x_{new}$ where x_w is the worst harmony in HM , otherwise go to step 3.
6	If $k = NI$ concludes the optimization and the solution is x_{best} in HM otherwise go to step 3.

Materials

In the first instance, the historical hydrometric information of the 10040 Santa Cruz station located in the state of Sinaloa ($1060^{\circ}57'10''$, $240^{\circ}29'05''$) was obtained in the period from 1944 to 1980 of the National Flow Data Bank (BANDAS) of the National Water Commission (CONAGUA) (Table 3 and Figure 1). As an additional case to verify the optimization capacity in the Gumbel Mixed distribution function using the Harmonic Search algorithm, the record of annual maximum flows (Gómez *et al.*, 2010) corresponding to the hydrometric station number 12504 named La Cuña in the state of Jalisco, Mexico ($1020^{\circ}49'59''$, $210^{\circ}00'15''$) for the period from 1947 to 2004, with 58 data recorded in Table 4 (Figure 2).

Table 3. Annual maximum flows data of the Santa Cruz hydrometric station, Sinaloa, Mexico.

year	Q (m ³ /s)								
1944	2142	1952	677	1960	1074	1968	7000	1976	1495
1945	1023.4	1953	807	1961	1280	1969	484	1977	836
1946	837.6	1954	553	1962	1002	1970	920.6	1978	940
1947	1161.2	1955	1252	1963	3680	1971	812	1979	3080
1948	1062	1956	369.5	1964	861	1972	3332.4	1980	1550
1949	784.2	1957	293	1965	888.8	1973	898		
1950	1086.3	1958	1157.2	1966	1166.4	1974	2790		
1951	487.8	1959	762.222	1967	950	1975	620		

Table 4. Annual maximum flows data of the La Cuña hydrometric station, Jalisco. Mexico.

year	Q (m ³ /s)										
1947	784.00	1957	199.00	1967	1474.86	1977	439.72	1987	184.68	1997	78.44
1948	736.80	1958	690.00	1968	323.00	1978	280.22	1988	595.20	1998	261.86
1949	510.00	1959	340.60	1969	160.40	1979	267.20	1989	110.18	1999	196.30
1950	461.00	1960	249.60	1970	763.75	1980	287.28	1990	523.86	2000	46.81
1951	411.00	1961	350.00	1971	578.00	1981	280.70	1991	1636.33	2001	313.81
1952	326.00	1962	317.00	1972	191.75	1982	156.50	1992	1168.00	2002	319.60
1953	349.80	1963	732.56	1973	2440.00	1983	455.50	1993	295.00	2003	621.07
1954	130.40	1964	265.05	1974	238.35	1984	501.20	1994	212.80	2004	824.50
1955	690.00	1965	743.60	1975	622.08	1985	385.00	1995	367.42		
1956	266.00	1966	463.90	1976	1374.00	1986	698.19	1996	144.57		

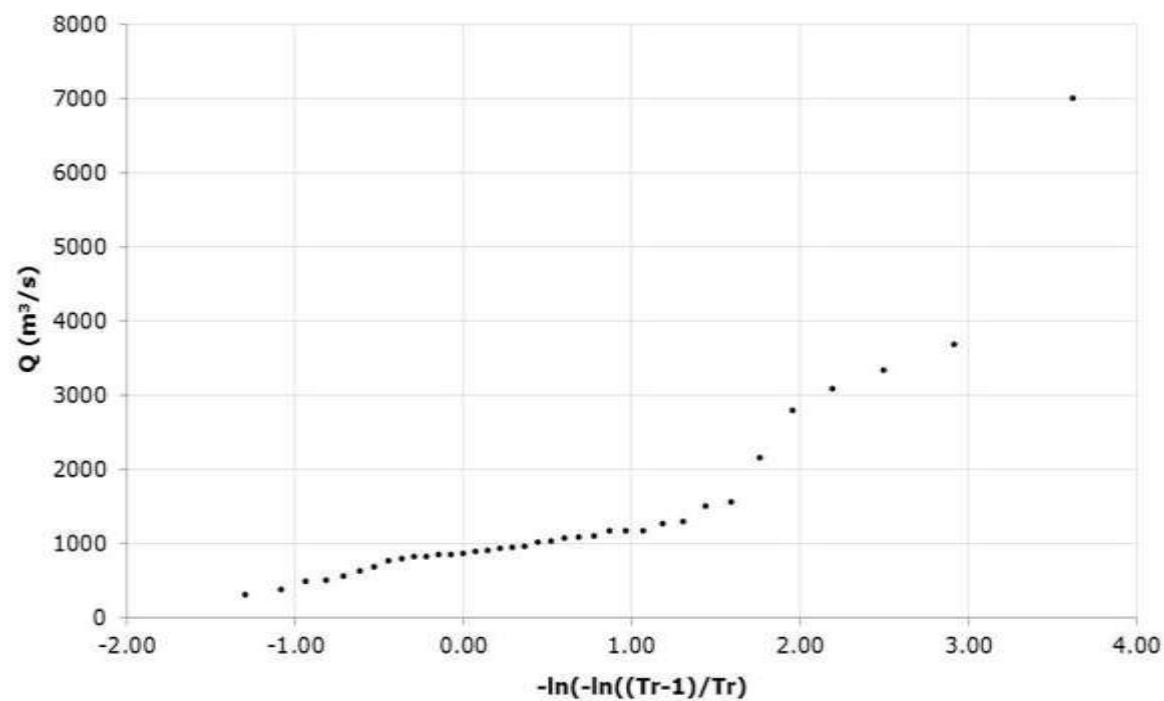


Figure 1. Maximum annual flows of the Santa Cruz, Sinaloa station.

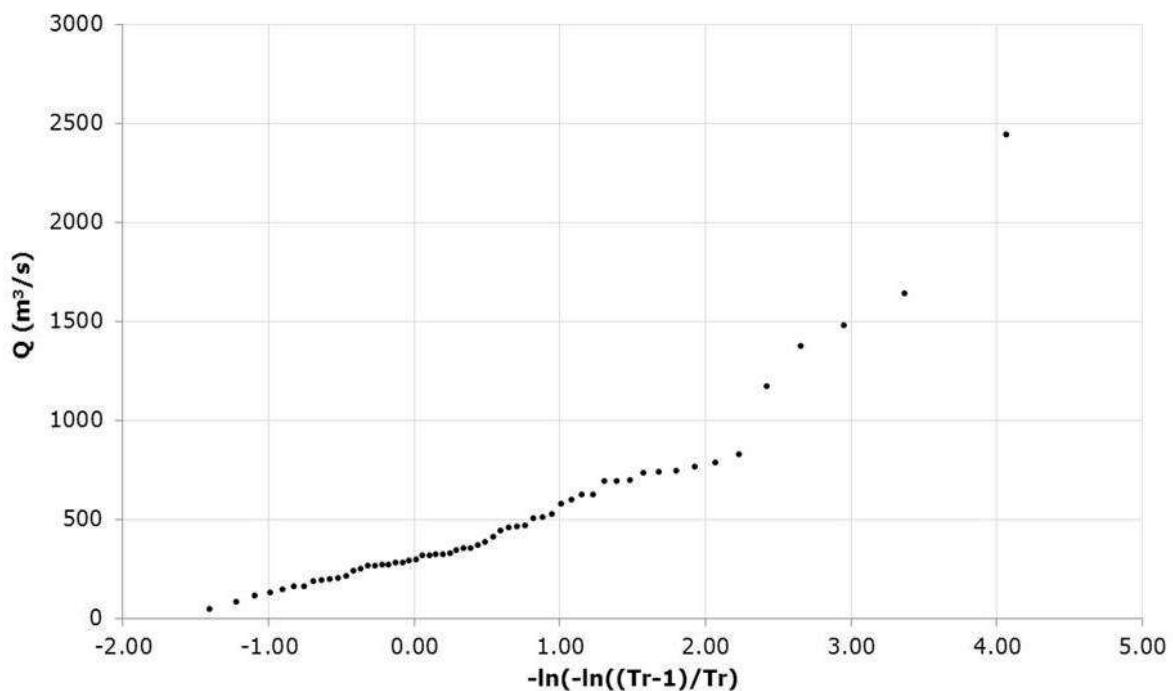


Figure 2. Maximum annual flows of the La Cuña, Jalisco station.

Results

Univariate Gumbel

Firstly, the adjustment of the univariate Gumbel function of one population was made by the methodologies of Moments, Maximum Likelihood and Harmonic Search, to the 37 historical data of the Santa Cruz station. The programming of the Harmonic Search algorithm was carried out in Matlab for which the ranges of the decision variables were defined, which correspond to the parameters of the algorithm remembering that the initial values are established stochastically $xU=[1000,10000]$, $xL=[0;0]$, $HMS=40$, $NI=10000$, $HMCR=0.90$, $PAR=0.35$; $B=(xU-xL)/NI$. Once the programming made in Matlab was verified, the objective function was modified and the standard adjustment error was calculated to optimize the Gumbel Mixed function, contrasting the results with respect to the Rosenbrock algorithm, (Figure 3) establishing the initial parameters considering nine cyclonic data $P = 0.756757$, $\lambda_1=194.24 \text{ m}^3/\text{s}$, $\varepsilon_1=736,678 \text{ m}^3/\text{s}$, $\lambda_2=1368.31 \text{ m}^3/\text{s}$ and $\varepsilon_2=2137.95 \text{ m}^3/\text{s}$. Similarly, the ranges were defined for the decision variables that correspond to the parameters of the algorithm $xU=[1, 10000, 10000, 10000, 10000, 10000]$, $xL=[0,0,0,0,0]$, $HMS=40$, $NI=10000$, $HMCR=0.90$, $PAR=0.35$; $B=(xUxL)/NI$. The Table 5 and Figure 3 show the values of the adjustment errors for the three procedures for estimating parameters by the mentioned methods.

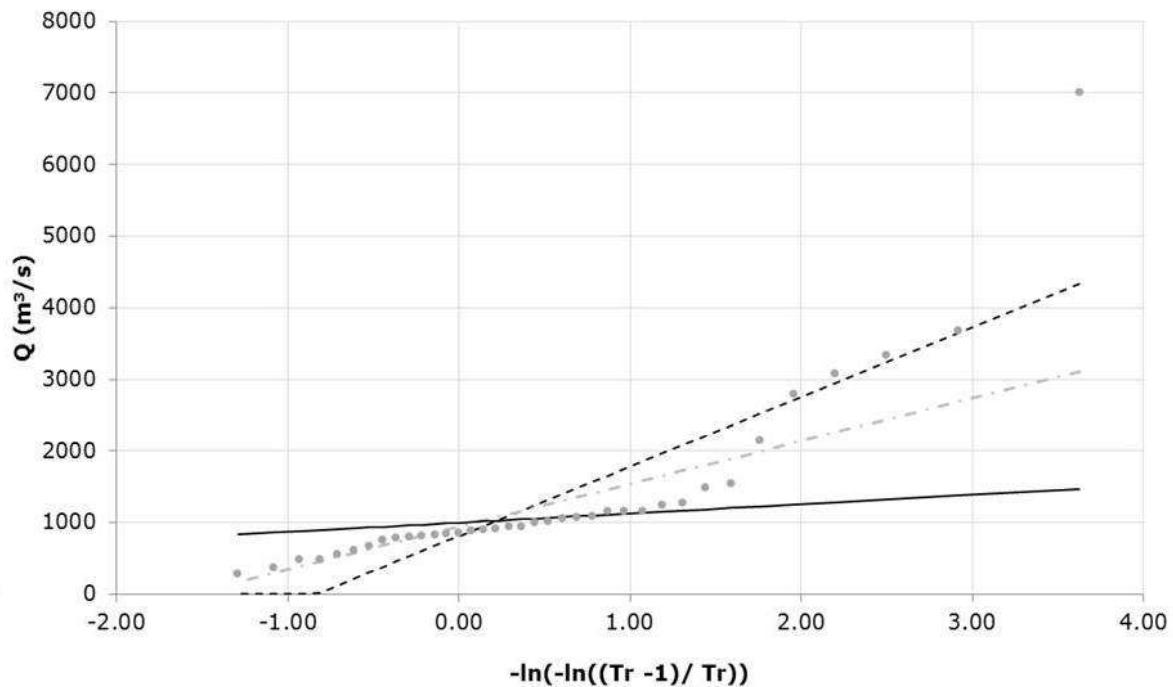


Figure 3. Gumbel function of one population fit for the Santa Cruz, Sinaloa station data by the moments method (—), the Maximum Likelihood Method (---), and Harmonic Search (—).

Table 5. Parameters for the Gumbel function of one population for the record of annual maximum flows of the Santa Cruz hydrometric station, Sinaloa, Mexico.

Methodology	$\lambda_I (m^3/s)$	$\varepsilon_I (m^3/s)$	EE
Moments	972.2	814.40	598.77
Maximum Likelihood	597.05	946.58	737.63
Harmonic Search	1000	128.98	82.82

Throughout the total number of iterations the value of the optimization function represented by equation 17 shown in Figure 4 was recorded; in the same way, the optimization function evaluated in the 40 harmonies of the harmonic memory was averaged as shown in Figure 5.

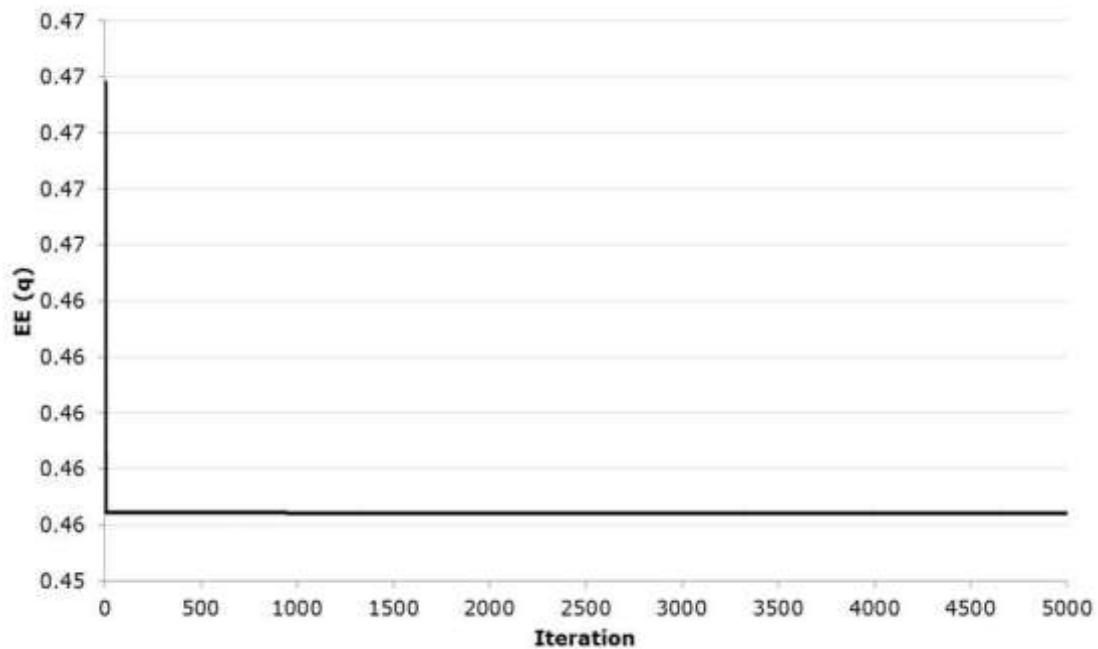


Figure 4 Evaluation of the optimization function for Univariate Gumbel.

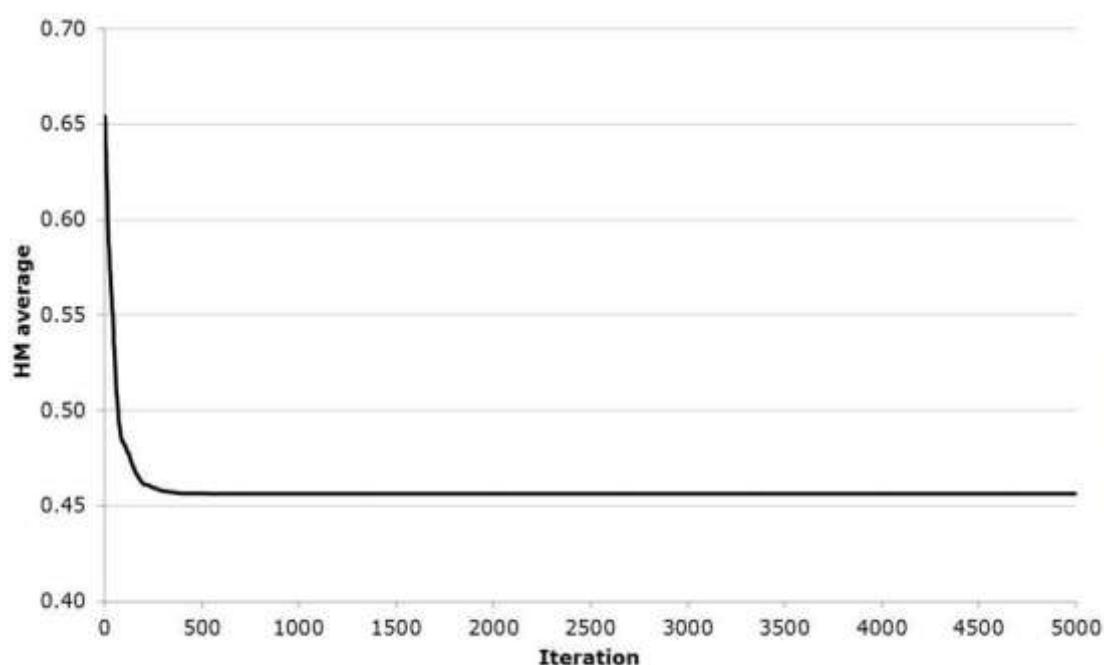


Figure 5 Mean of the optimization function in the harmonic memory.

Gumbel Mixed

Next, the same algorithm is used to carry out the optimization process of the Gumbel Mixed probability distribution or of two populations. The algorithm performs the most unfavorable harmonic modification contained in the harmonic memory and stochastically modifies one of the decision variables. That is, it modifies the four parameters of scale and location as well as the association parameter P . The algorithm carries out the exploration of the solution space as shown in Figures 6 and 7, defining a horizontal trend by the successive repetition of the optimal value of the parameter in the evaluated harmony, while the isolated values represent the stochastic values generated in the convergence process. Table 6 shows the adjusted values of the five parameters that result from this procedure, as well as the adjustment error. Comparing adjustments moments, Rosenbrock and Harmonica search for historical data of Santa Cruz hydrometric station shown in Figure 8.

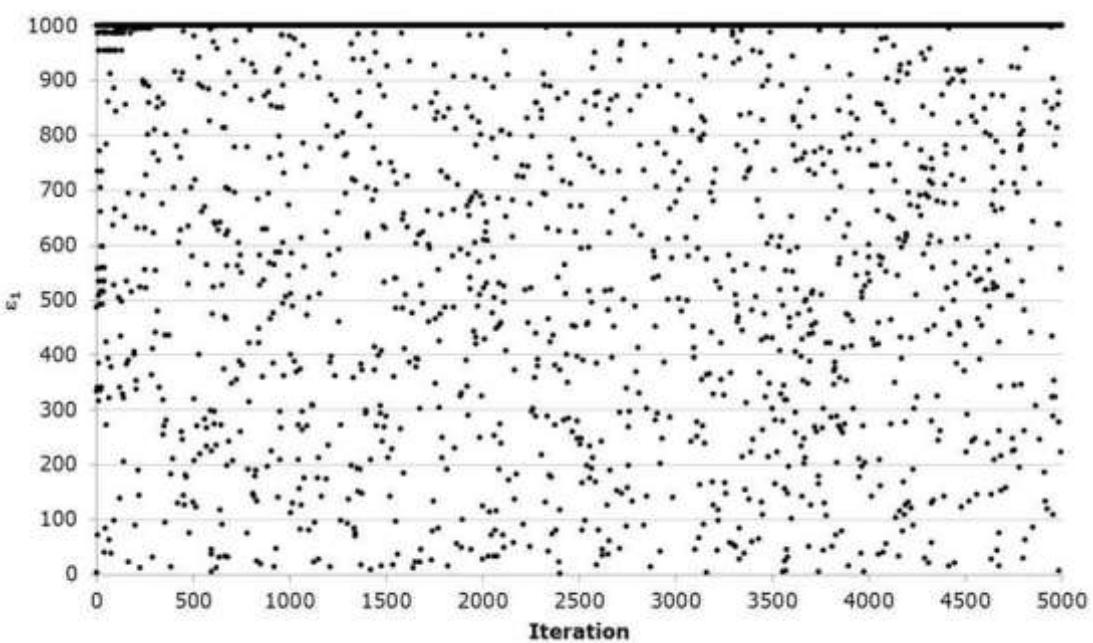


Figure 6. Stochastic behavior of the scale parameter.

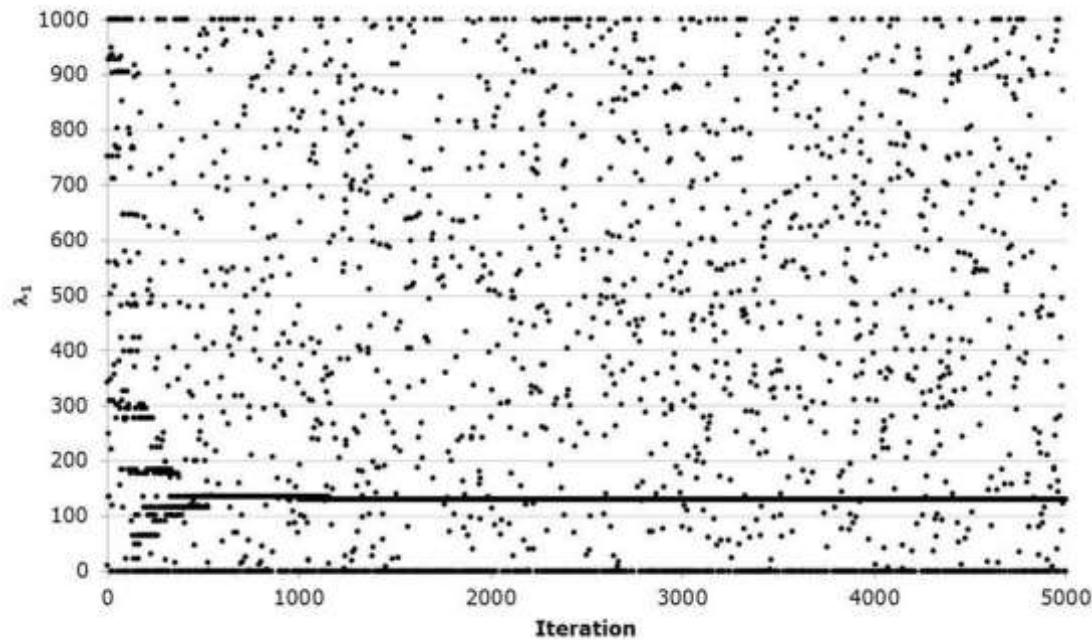


Figure 7. Stochastic behavior of the location parameter.

Table 6. Parameters for the Gumbel Mixed function fit for the recording of annual maximum flows of the Santa Cruz hydrometric station, Sinaloa. Mexico.

Methodology	P	λ_1 (m³/s)	ε_1 (m³/s)	λ_2 (m³/s)	ε_2 (m³/s)	EE
Moments	0.784	863.652	200.780	3,133.625	1,369.655	542.022
Rosenbrock	0.850	769.051	296.580	3,037.430	1,022.059	407.024
B. Harmonic	0.824	788.796	281.276	2,850.544	1,436.975	400.008

Figures 9 and 10 show the evolution of the value of the optimization function represented by equation 17 for the case of the Gumbel Mixed univariate distribution. The evolution in the solution space of the aleatory search of the harmonics carried out by the proposed algorithm is presented in a sequence in Figures 11 to 15.

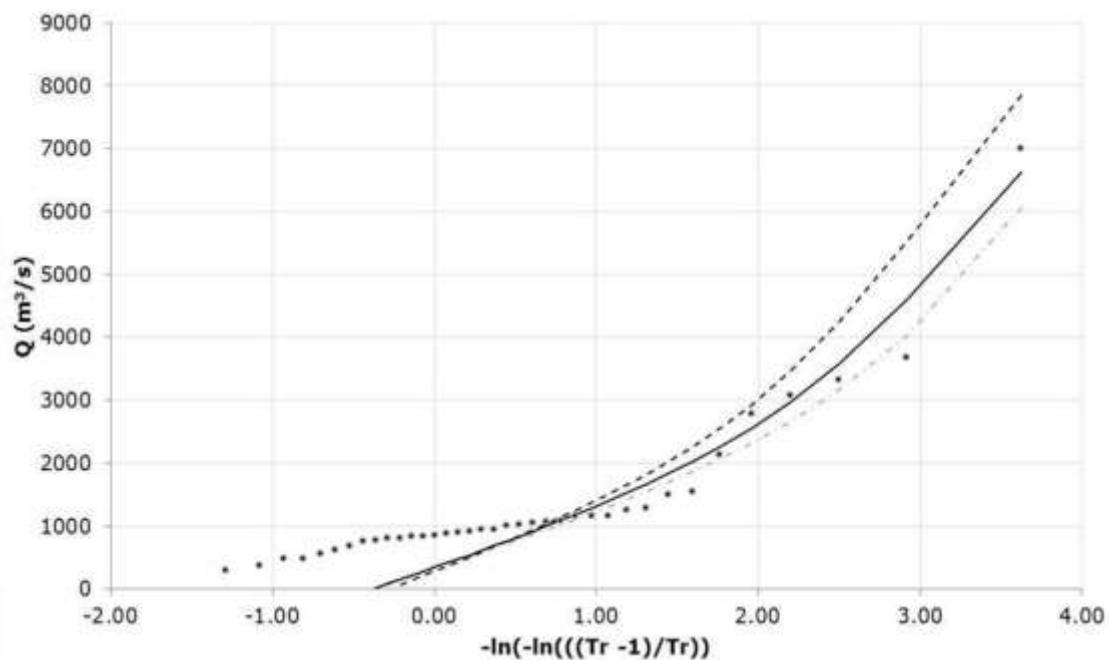


Figure 8. Gumbel Mixed function fit for data from the Santa Cruz, Sinaloa station. Method of moments (—), Rosenbrock's algorithm (---) and Harmonic Search (—).

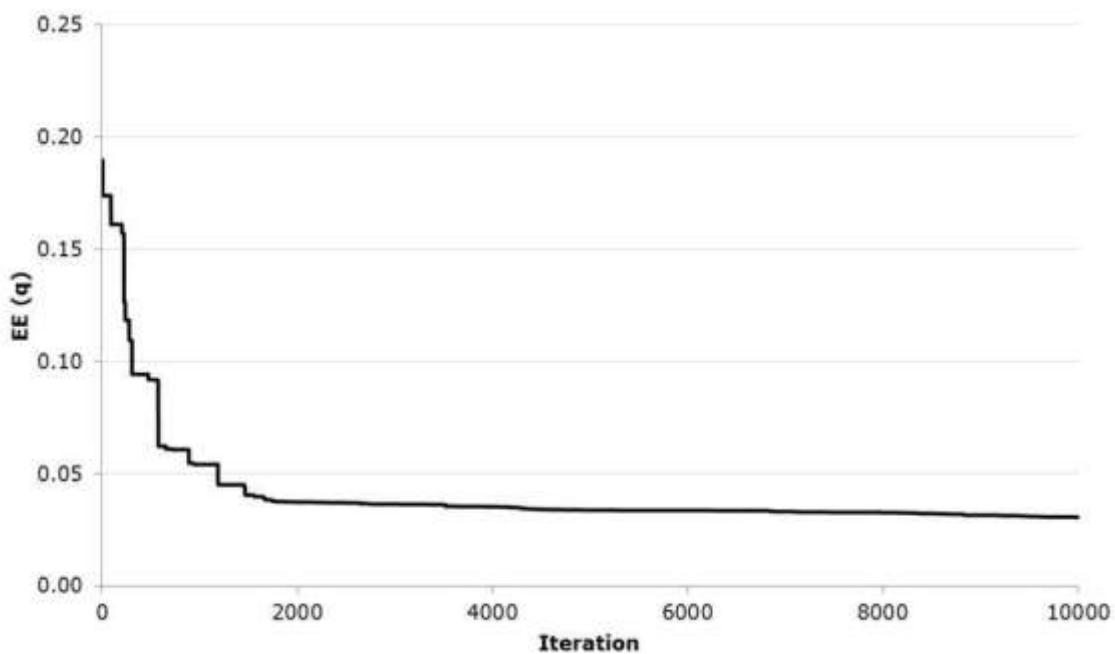


Figure 9. Evaluation of the optimization function for Gumbel Mixed

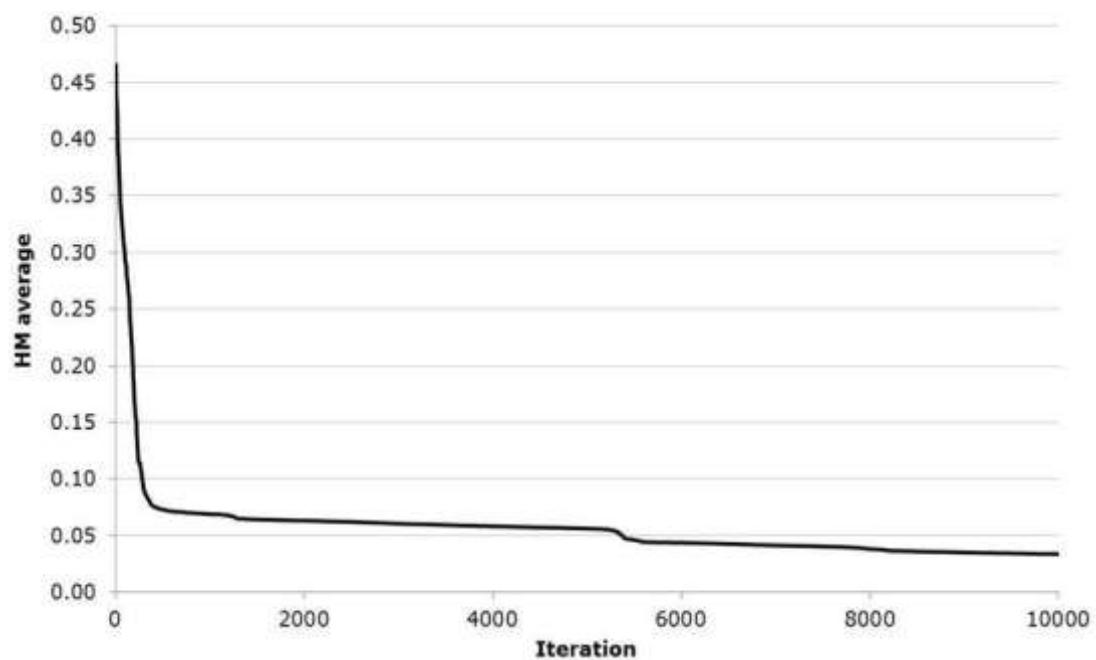


Figure 10. Mean of the 40 harmonies evaluated in the optimization function.

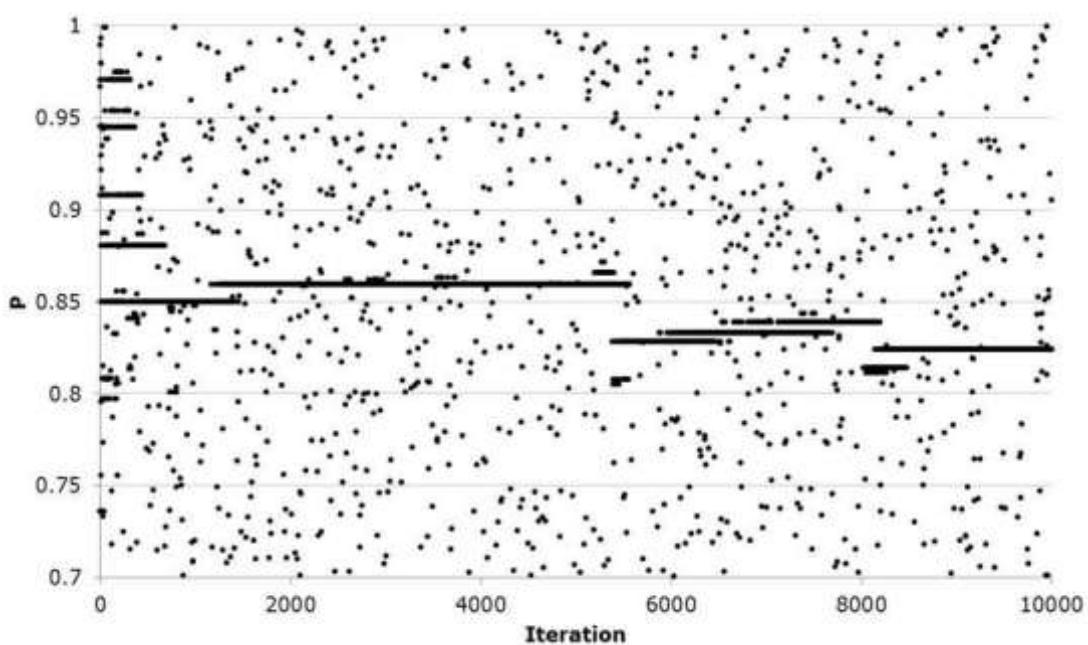


Figure 11. Aleatory probability behavior.

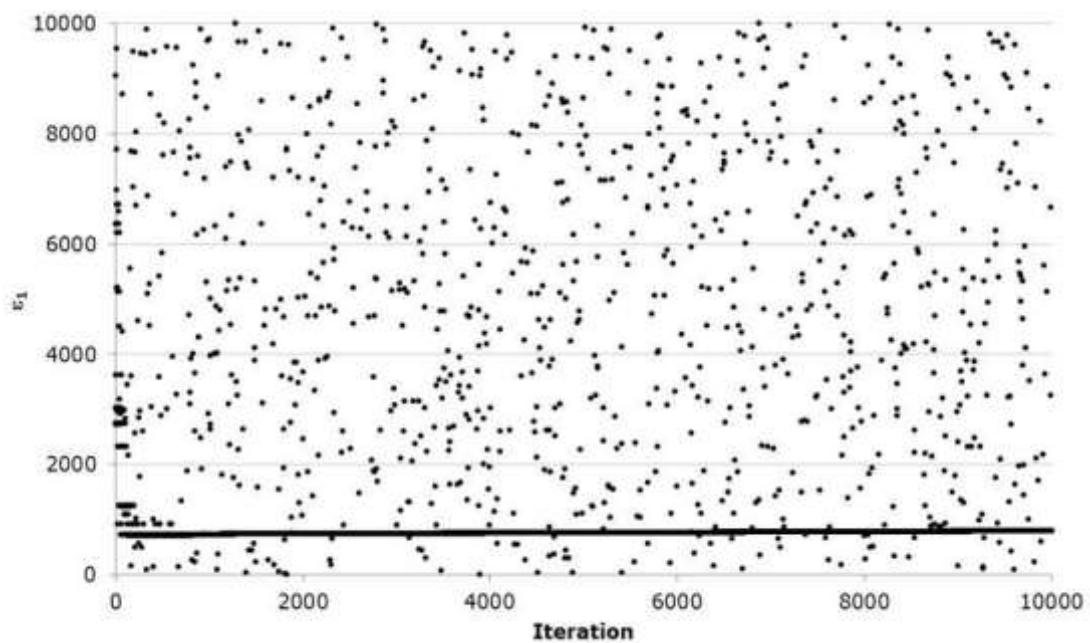


Figure 12. Aleatory behavior of the non-cyclonic scale parameter.

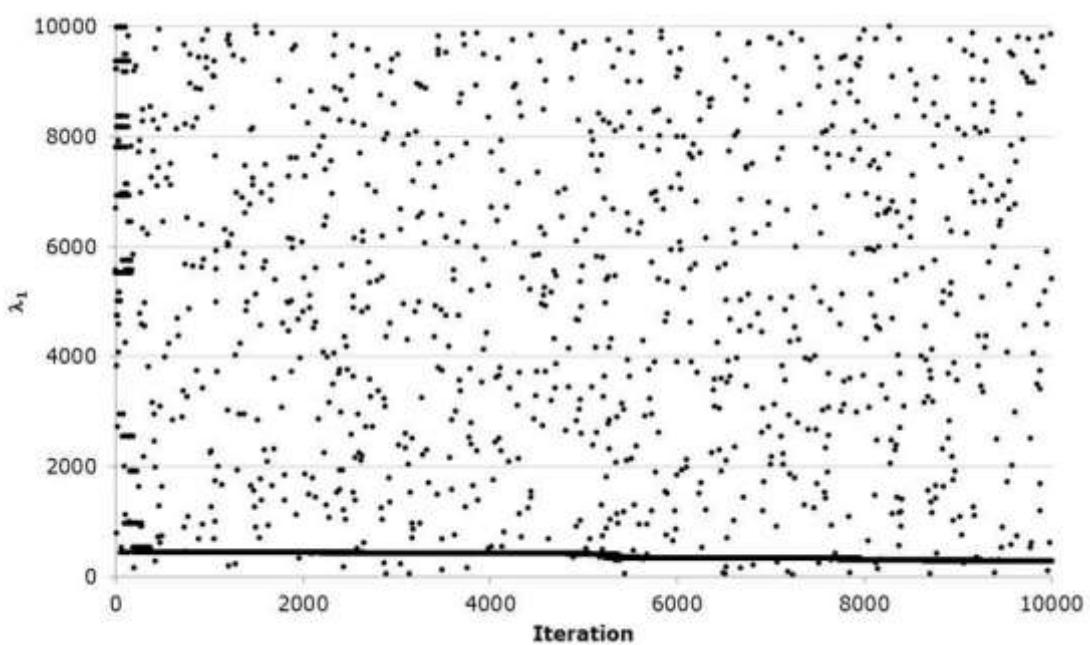


Figure 13. Aleatory behavior of parameter P of non-cyclonic location.

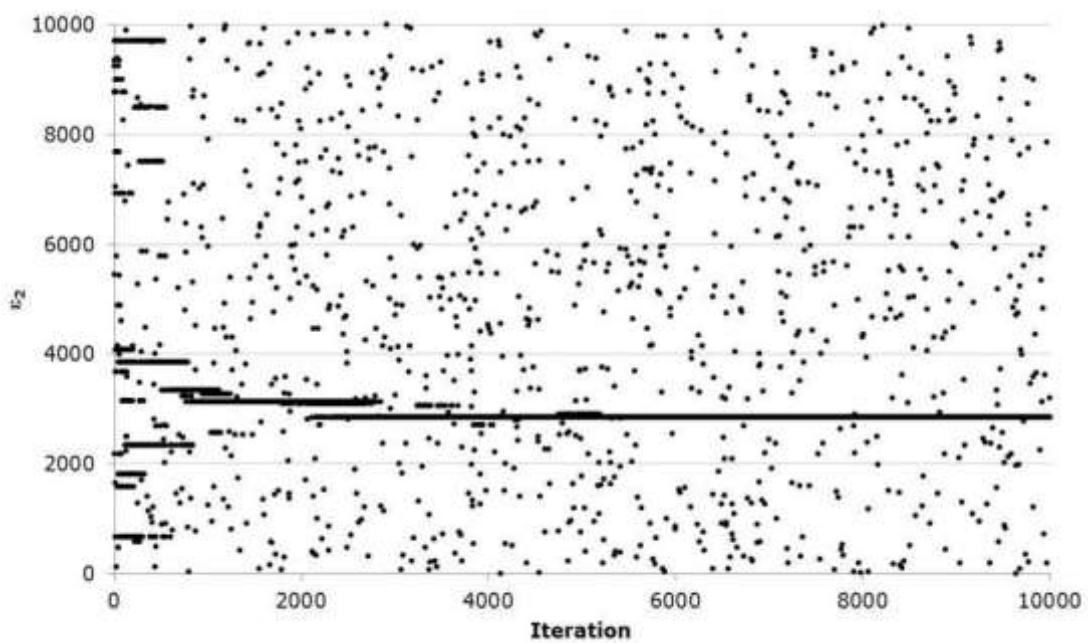


Figure 14. Aleatory behavior of the P parameter of the cyclonic scale.

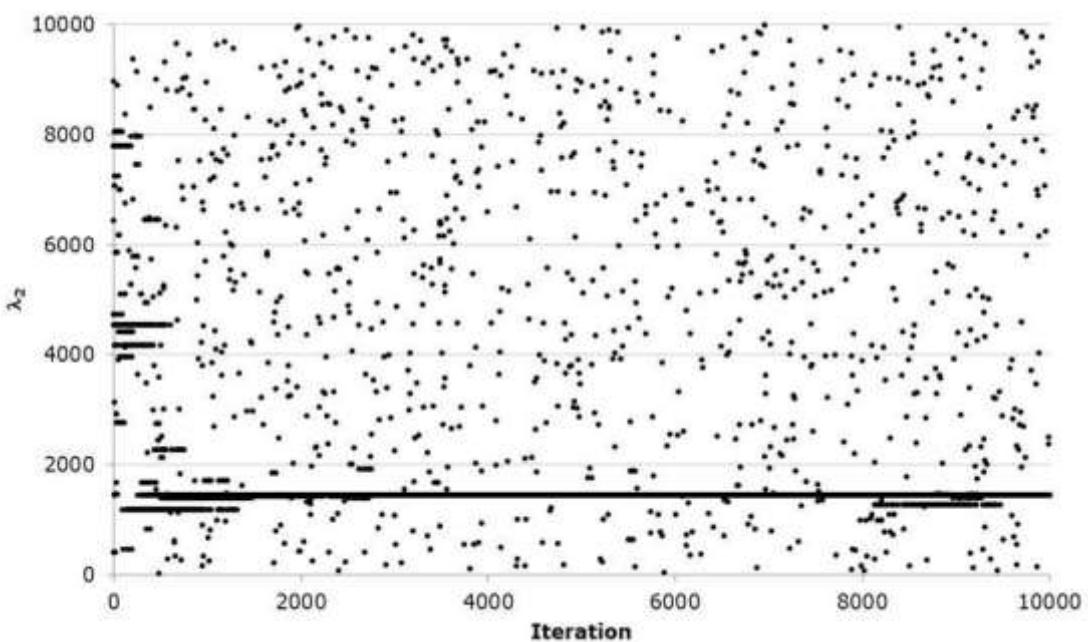


Figure 15. Aleatory behavior of the cyclonic location parameter.

Gómez *et al.*, (2010) present a classic example of the numerical adjustment of a Gumbel Mixed distribution. In order to exemplify the proposed procedure, the historical data of station 12504 called La Cuña in the state of Jalisco are used. The parameters of the algorithm were $xU=[1, 3000, 3000, 3000, 3000]$, $xL=[1, 0, 0, 0, 0]$, $HMS=30$, $NI=10000$, $HMCR=0.90$, $PAR=0.38$; $B=(xU-xL)/NI$ to start the *pseudo-code* of Table 2.

Table 7 shows the value of the five adjustment parameters for the Gumbel Mixed function for the record of annual maximum flows of the La Cuña hydrometric station, Jalisco. Figure 16 shows the accuracy in adjusting the extreme values when the harmonic estimation search parameters Gumbel Mixed univariate distribution is used. In the same way as for the data of the Santa Cruz station, Figures 17 and 18 show the evolution in the convergence of the optimization function.

Table 7. Parameters for the Gumbel Mixed function fit for the recording of maximum annual flows of the La Cuña hydrometric station, Jalisco Mexico.

Methodology	P	λ_1 (m ³ /s)	ε_1 (m ³ /s)	λ_2 (m ³ /s)	ε_2 (m ³ /s)	EE
Gomez <i>et al.</i> 2010	0.8965	292.311	157.143	1271.095	424.786	80.200
Harmonic Search	0.8674	296.599	170.766	713.726	782.383	66.082

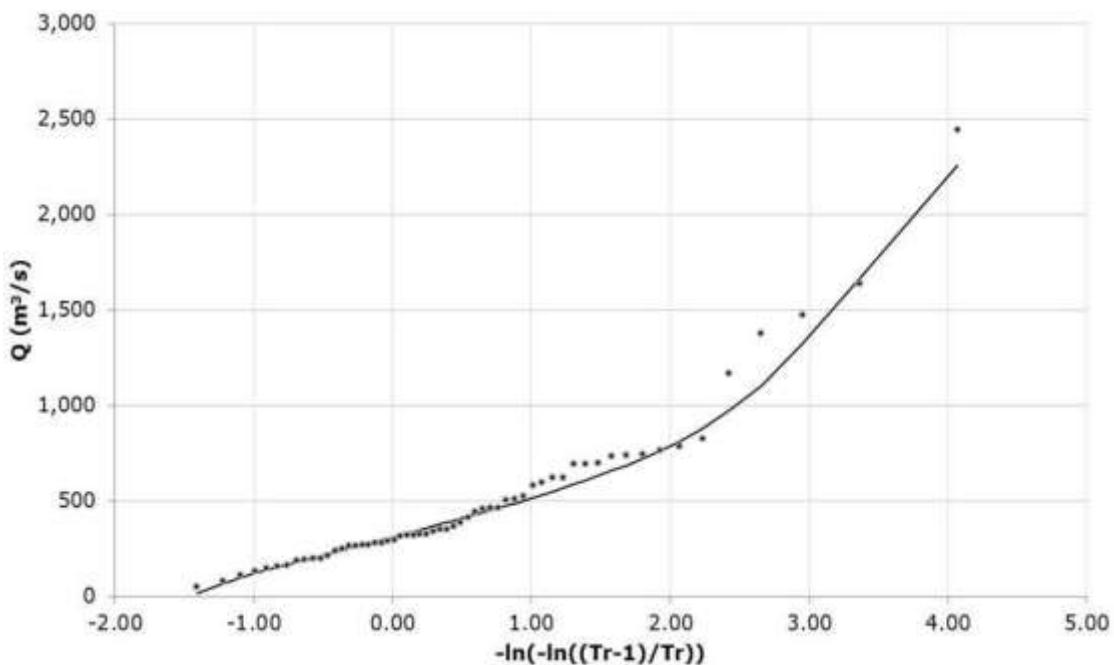


Figure 16. Optimization of the Mixed Gumbel function for the Cuña, Jalisco station using the Harmonic Search algorithm.

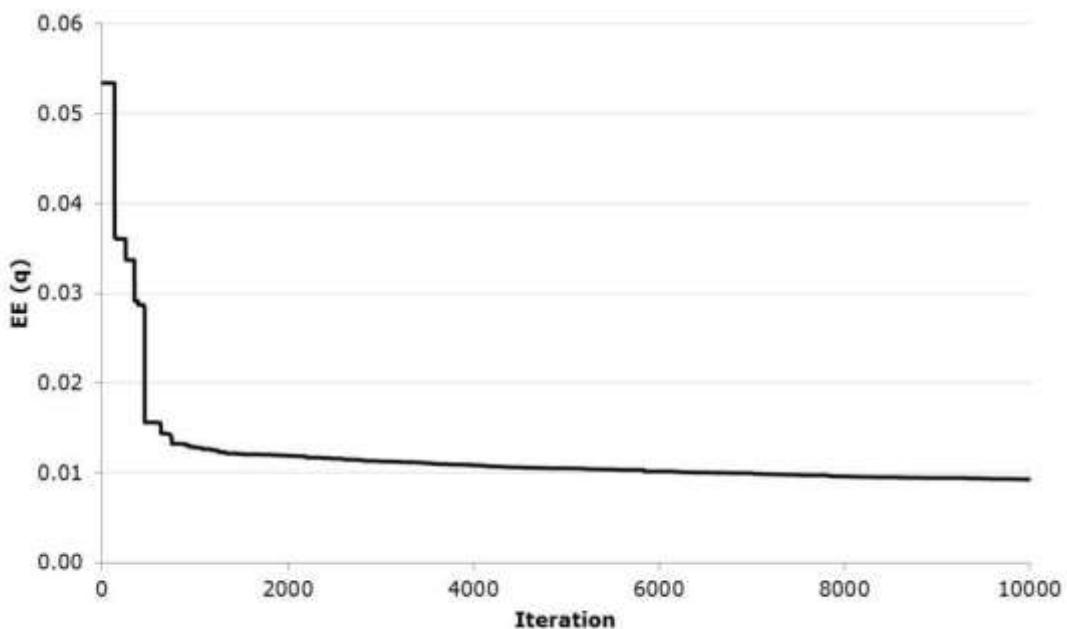


Figure 17. Evaluation of the optimization function for Gumbel Mixed.

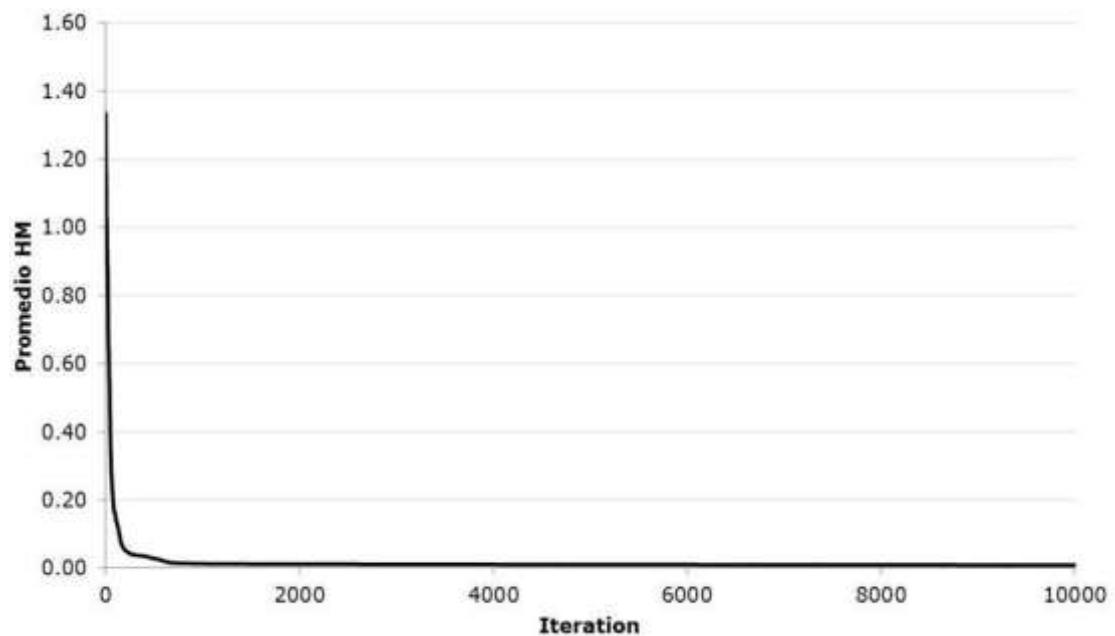


Figure 18. Mean of the 40 harmonies evaluated in the optimization function.

On the other hand, Figures 19 to 23 show the aleatory behavior of the location and scale parameters for the cyclonic and non-cyclonic samples.

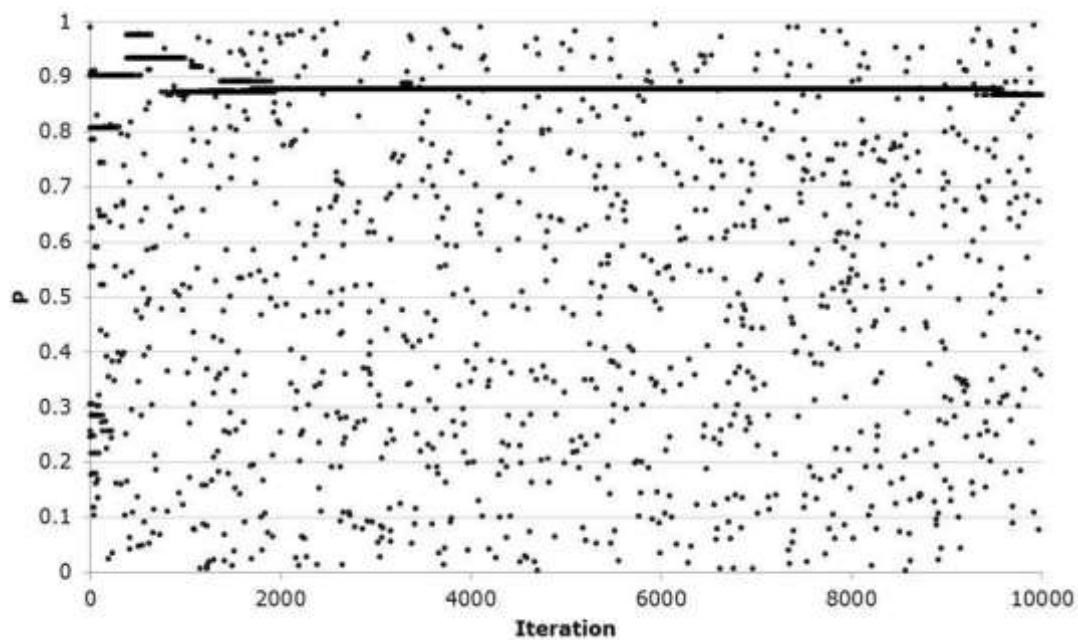


Figure 19. Aleatory probability behavior.

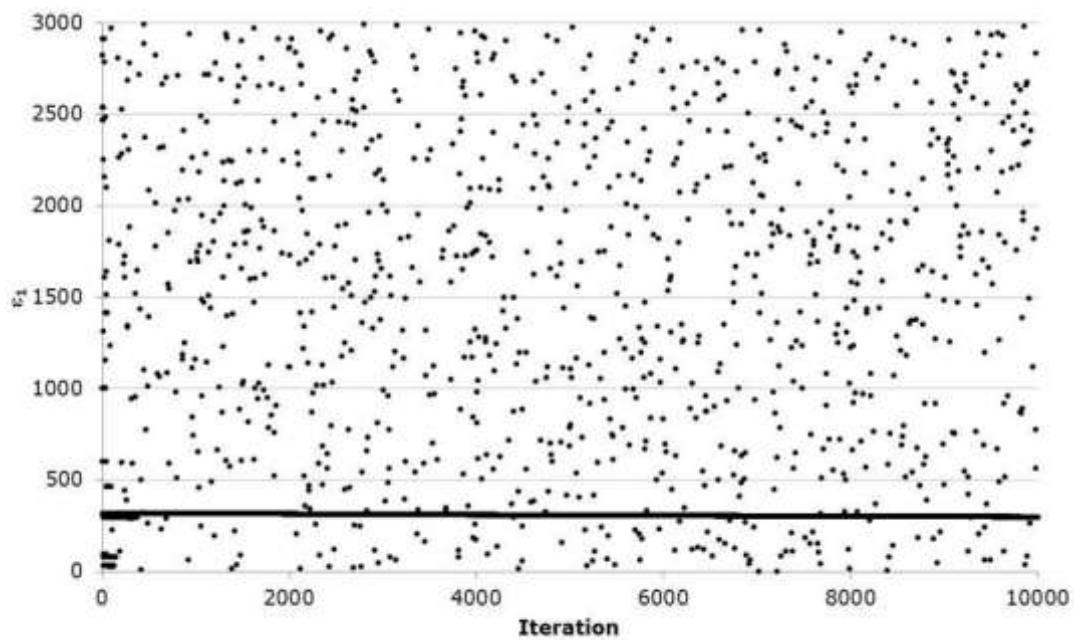


Figure 20. Aleatory behavior of the non-cyclonic scale parameter.

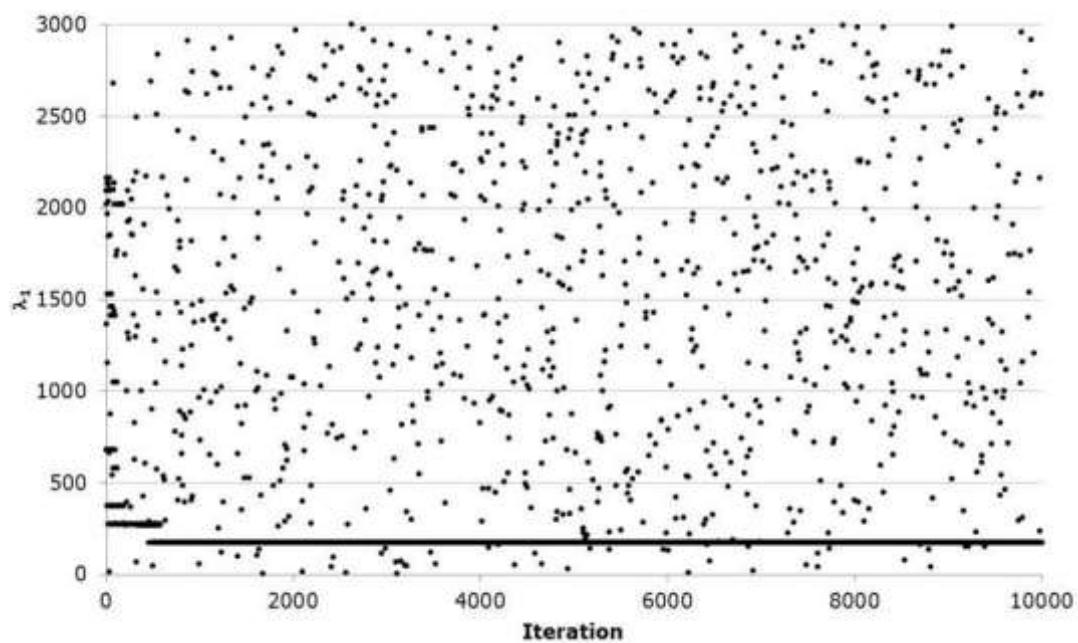


Figure 21 Aleatory behavior of the non-cyclonic location parameter.

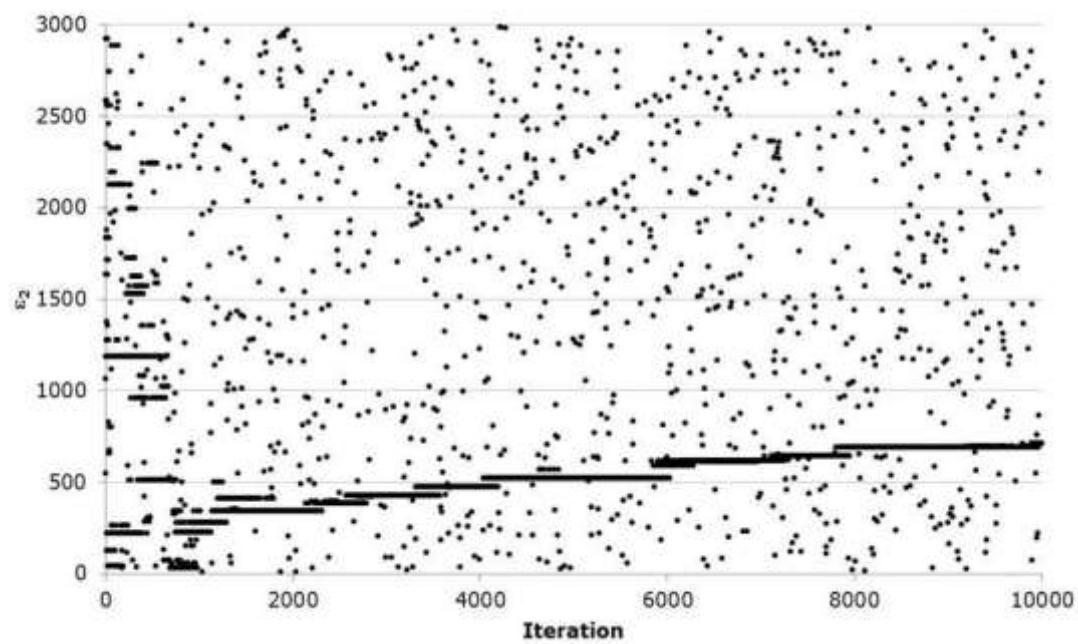


Figure 22. Aleatory behavior of the cyclonic scale parameter.

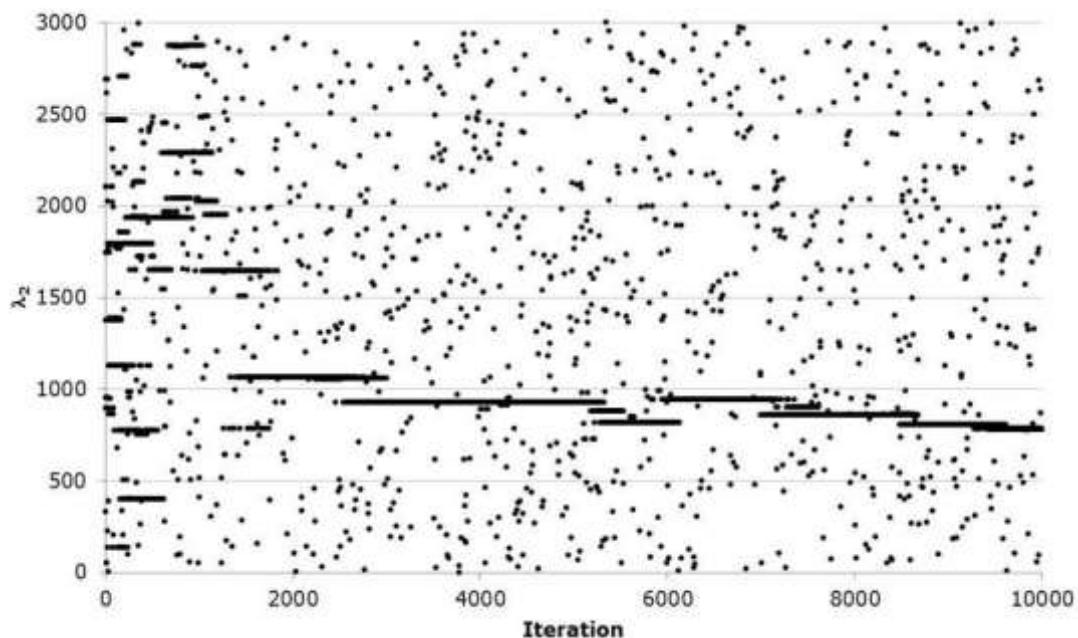


Figure 23. Aleatory behavior of the cyclonic location parameter.

Discussion

The first case study referred to the Santa Cruz hydrometric station, in which the Gumbel distribution function of one population was first adjusted to contrast the Harmonic Search algorithm with the conventional methodologies of Moments and Maximum Likelihood. An important decrease in the value of the Standard Adjustment Error could be observed, although graphically a better performance of the adjustment by Moments and even by Maximum Likelihood is appreciated. However, it should be noted that there was a decrease in the magnitude of *EE* for the Harmonic Search algorithm. This better adjustment is mainly due to the fact that good performance is shown in the area of flows with a value of the reduced variable of less than 1.5. That is, it does not comply with the maximum historical flows of the series, generating an important forecast error in the final zone of the series.

Regarding the convergence of the algorithm, it is observed in Figure 4 that very few iterative processes are required to reach the optimal solution. Similarly in Figure 5 you can see the mean of the 40 harmonies or combinations of values of the evaluated variables where, when the worst harmony is eliminated, the remaining 39 begin to reduce the value of the optimization function when compared with each other. Which guarantees a fast convergence, although the stochastic exploration of the solution space defined by the threshold values in the decision variables and the number of iterations continues. In Figures 6 and 7 shows how during the search values corresponding to the best solution to the problem harmony or visually generating a trend line aligned.

As the optimization of the Gumbel Mixed function develops, we can again appreciate that it decreases the value of the Standard Adjustment Error with respect to the methods at the moment and the Rosenbrock Algorithm. Regarding the latter, the variation is minimal in the result but visually we can see that the adjustment improves its performance in the region of the maximum flows of the series but presents difficulties in the zone of the minimum flows. Likewise, it presents a good convergence performance requiring approximately 2000 iterations of the 10000 performed to optimize the five parameters of the objective function, (the evolution of the iterations can be seen in the Figures 9 and 10) since the difference of equation 17 quickly decreases when evaluating the harmonic harmonics of 40, whereby the parameters ε_1 , λ_1 and λ_2 show greater stability in the stochastic exploration process to define its value optimized unlike ε_2 and P , this last parameter being the one that requires more iterations to identify the minimum in correlation with the rest.

Finally, by increasing the size of the time series of maximum flows we can see in Figure 16 that the adjustment by means of the Harmonic Search algorithm improves significantly in the area of the smaller flows without neglecting the region of higher flows with which there is a reduced value of the Standard Adjustment Error. in this particular case we can observe greater stability in the determination of the parameters P , ε_1 and λ_1 , the last two associated with the non-cyclonic data, while the parameters ε_2 and λ_2 corresponding to cyclonic data show a greater imbalance, particularly in the first 2000 iterations gradually and gradually reducing.

Conclusions

The results show that the harmonic search algorithm can optimize the Gumbel distribution function of both one and two populations (Gumbel Mixed), in the first case without a notable difference compared to conventional methods of adjustment even without improving the trend ordered by the reduced variable. On the other hand, when considering non-cyclonic and cyclonic populations, a significant increase in the performance of the proposed algorithm is observed to optimize the adjustment and decrease the adjustment error; which is enhanced by the use of extensive historical data series, so that their area of opportunity is located precisely in large registers. The Harmonic Search must be considered a viable option and of generalized application for the future estimation of parameters of probability distributions used in hydrology.

Referencias

- Ahn, K., & Merwade, V. (2016). The effect of land cover change on duration and severity of high and low flows. *Hydrological Processes*, 31(1), 133-149. doi: 10.1002/hyp.10981
- Ahn, K., & Palmer, R. (2016). Use of a nonstationary copula to predict future bivariate low flow frequency in the Connecticut River basin. *Hydrological Processes*, 30(19), 3518-3532. doi: 10.1002/hyp.10876
- Al Aamery, N., Fox, J., Snyder, M., & Chandramouli, C. (2018). Variance analysis of forecasted streamflow maxima in a wet temperate climate. *Journal of Hydrology*, 560, 364-381. doi: 10.1016/j.jhydrol.2018.03.038
- Amin, M., Rizwan, M., & Alazba, A. (2014). Comparison of mixed distribution with EV1 and GEV components for analyzing hydrologic data containing outlier. *Environmental Earth Sciences*, 73(3), 1369-1375. doi: 10.1007/s12665-014-3490-4
- Arnaud, P., Cantet, P., & Odry, J. (2017). Uncertainties of flood frequency estimation approaches based on continuous simulation using data resampling. *Journal of Hydrology*, 554, 360-369. doi: 10.1016/j.jhydrol.2017.09.011
- Ashkar, F., El-Jabi, N., & Sarraf, S. (1991). Study of hydrological phenomena by extreme value theory. *Natural Hazards*, 4(4), 373-388. doi: 10.1007/bf00126645

Ayantobo, O., Li, Y., Song, S., Javed, T., & Yao, N. (2018). Probabilistic modelling of drought events in China via 2-dimensional joint copula. *Journal of Hydrology*, 559, 373-391. doi: 10.1016/j.jhydrol.2018.02.022

Bardsley, W. (2016). Cautionary note on multicomponent flood distributions for annual maxima. *Hydrological Processes*, 30(20), 3730-3732. doi: 10.1002/hyp.10886

Borga, M., Vezzani, C., & Fontana, G. (2005). Regional Rainfall Depth-Duration-Frequency Equations for an Alpine Region. *Natural Hazards*, 36(1-2), 221-235. doi: 10.1007/s11069-004-4550-y

Campos-Aranda, D. F. (2003). *Introducción a los Métodos Numéricos: Software en Basic y aplicaciones en hidrología superficial*. Capítulo 5: Ajuste de curvas, pp. 93-127 y Capítulo 9. Optimización numérica, pp. 172-211, San Luis Potosí, Librería Universitaria Potosina.

Campos-Aranda, F. (1989). Estimación de los parámetros óptimos de la distribución Gumbel mixta por medio del algoritmo de Rosenbrock. *Tecnología y Ciencias del Agua*, vol. Ene-Abr., pp. 9-18

Campos-Aranda, F. (1994). Aplicación del Método del Índice de Crecientes en la Región Hidrológica Número 10, Sinaloa. *Tecnología y Ciencias del Agua*, vol. IX, núm. 3, pp. 41-55

Chen, L., & Singh, V. (2018). Entropy-based derivation of generalized distributions for hydrometeorological frequency analysis. *Journal of Hydrology*, 557, 699-712. doi: 10.1016/j.jhydrol.2017.12.066

Chow, V. T. (1964). *Statistical and probability analysis of hydrologic data*. Section 8, part I frequency analysis. In: *Handbook of applied hydrology*. McGraw-Hill, Nueva York.

Cuevas, E., Ortega-Sánchez, N. (2013). El algoritmo de búsqueda armónica y sus usos en el procesamiento digital de imágenes. *Computación y sistemas* vol. 17 No. 4, pp. 543-560. Doi: 10.13053/CyS-17-4-2013-007

Díaz-Delgado C., Bâ K., y Trujillo-Flores E. (1999). Las funciones Beta-Jacobi y Gamma-Laguerre como métodos de análisis de valores hidrológicos extremos. Caso de precipitaciones máximas anuales. *Tecnología y Ciencias del Agua*, vol. XIV, núm. 2, pp. 39-48

Escalante-Sandoval, C. (2007). Application of bivariate extreme value distribution to flood frequency analysis: a case study of Northwestern Mexico. *Natural Hazards*, 42(1), 37-46. doi: 10.1007/s11069-006-9044-7

Escalante-Sandoval, C. (2013). Estimation of Extreme Wind Speeds by Using Mixed Distributions. *Ingeniería Investigación y Tecnología*, XIV(2), 153-162

Escalante-Sandoval, C., y Domínguez-Esquivel, J. (2001). Análisis regional de precipitación con base en una distribución bivariada ajustada por máxima entropía. *Tecnología y Ciencias del Agua*, vol. XVI, núm. 3, pp. 91-102

Fogel, L., Owens, A., & Walsh, M. (1966). *Artificial intelligence through simulated evolution*. Chichester, WS, UK: Wiley.

Geem, Z. W., Kim, J. H., & Loganathan, G. V. (2001). A new heuristic optimization algorithm: Harmony search. *Simulation*, 76(2), 60-68. doi: 10.1177/003754970107600201

Glover, F. (1977). Heuristics for integer programming using surrogate constraints. *Decision Sciences*, 8(1), 156-166. doi: 10.1111/j.1540-5915.1977.tb01074.x

Gómez, J., Aparicio, J., y Patiño, C. (2010). *Manual de análisis de frecuencias en hidrología*. Instituto Mexicano de Tecnología del Agua. Jiutepc Morelos.

Grünthal, G., Thieken, A., Schwarz, J., Radtke, K., Smolka, A., & Merz, B. (2006). Comparative Risk Assessments for the City of Cologne – Storms, Floods, Earthquakes. *Natural Hazards*, 38(1-2), 21-44. doi: 10.1007/s11069-005-8598-0

Gumbel, E. J. (1960). *Distributions des valeurs extrêmes en plusieurs dimensions*, Publications de L'Institute de Statistique, Paris, Vol. 9, 171-173.

Gutiérrez-López A., y Ramírez, A. (2005). Predicción hidrológica mediante el Método de la Avenida Índice para dos poblaciones. *Tecnología y Ciencias del Agua*, vol. XX, núm. 2, pp. 37-47

Holland, J.H. (1975) *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor. (2nd Edition, MIT Press, 1992.)

Huang, J., Li, Y., Yin, J., & Mcbean, E. (2016). Precipitation regional extreme mapping as a tool for ungauged areas and the assessment of climate changes. *Hydrological Processes*, 30(12), 1940-1954. doi: 10.1002/hyp.10743

Jenkinson, A. F. (1969). *Estimation of maximum floods*. World Meteorological Organization, Technical Note No. 98, ch. 5, 183-257.

Joshi, S., Jayadeva, Ramakrishnan, G., & Chandra, S. (2012). Using Sequential Unconstrained Minimization Techniques to simplify SVM solvers. *Neurocomputing*, 77(1), 253-260. doi: 10.1016/j.neucom.2011.07.010

Karim, F., Petheram, C., Marvanek, S., Ticehurst, C., Wallace, J., & Hasan, M. (2015). Impact of climate change on floodplain inundation and hydrological connectivity between wetlands and rivers in a tropical

river catchment. *Hydrological Processes*, 30(10), 1574-1593. doi: 10.1002/hyp.10714

Kirkpatrick, S., Gelatt, C., & Vecchi, M. (1983). Optimization by Simulated Annealing. *Science*, 220(4598), 671-680. doi: 10.1126/science.220.4598.671

Koutsoyiannis, D. (2004). On the appropriateness of the Gumbel distribution for modelling extreme rainfall (solicited). *Hydrological Risk: recent advances in peak river flow modelling, prediction and real-time forecasting. Assessment of the impacts of land-use and climate changes*. Edited by A. Brath, A. Montanari, and E. Toth, Bologna, 303-319, doi:10.13140/RG.2.1.3811.6080, Editoriale Bios, Castrolibero, Italy.

Kuester, J., & Mize, J. (1973). *Optimization Techniques with Fortran*, ROSEN B algorithm, New York, McGraw-Hill Book Co. pp:320-330

Lee, K., & Geem, Z. (2005). A new meta-heuristic algorithm for continuous engineering optimization: harmony search theory and practice. *Computer Methods in Applied Mechanics and Engineering*, 194(36-38), 3902-3933. doi: 10.1016/j.cma.2004.09.007

Lilienthal, J., Fried, R., & Schumann, A. (2018). Homogeneity testing for skewed and cross-correlated data in regional flood frequency analysis. *Journal of Hydrology*, 556, 557-571. doi: 10.1016/j.jhydrol.2017.10.056

Mélèse, V., Blanchet, J., & Molinié, G. (2018). Uncertainty estimation of Intensity–Duration–Frequency relationships: A regional analysis. *Journal of Hydrology*, 558, 579-591. doi: 10.1016/j.jhydrol.2017.07.054

Montaseri, M., Amirataee, B., & Rezaie, H. (2018). New approach in bivariate drought duration and severity analysis. *Journal of Hydrology*, 559, 166-181. doi: 10.1016/j.jhydrol.2018.02.018

Öztürk, S., Bayrak, Y., Çınar, H., Koravos, G., & Tsapanos, T. (2008). A quantitative appraisal of earthquake hazard parameters computed from Gumbel I method for different regions in and around Turkey. *Natural Hazards*, 47(3), 471-495. doi: 10.1007/s11069-008-9234-6

Phillips, R., Samadi, S., & Meadows, M. (2018). How extreme was the October 2015 flood in the Carolinas? An assessment of flood frequency analysis and distribution tails. *Journal of Hydrology*, 562, 648-663. doi: 10.1016/j.jhydrol.2018.05.035

Qi, W., & Liu, J. (2018). A non-stationary cost-benefit based bivariate extreme flood estimation approach. *Journal of Hydrology*, 557, 589-599. doi: 10.1016/j.jhydrol.2017.12.045

Qian, L., Wang, H., Dang, S., Wang, C., Jiao, Z., & Zhao, Y. (2017). Modelling bivariate extreme precipitation distribution for data-scarce

regions using Gumbel-Hougaard copula with maximum entropy estimation. *Hydrological Processes*, 32(2), 212-227. doi: 10.1002/hyp.11406

Raynal-Villaseñor, J. (1997). Sobre el uso del dominio de atracción para la identificación de distribuciones de valores extremos para máximos. *Tecnología y Ciencias del Agua*, vol. XII, núm. 2, pp. 57-62

Rosenbrock, H. H. (1960). An automatic method for finding the greatest or least value of a function. *The Computer Journal*, 3:175-184. doi:10.1093/comjnl/3.3.175

Rossi, F., Florentino, M., & Versace, P. (1984). Two-Component Extreme Value Distribution for Flood Frequency Analysis. *Water Resources Research* 20(7). 847-856. doi:10.1029/WR020i007p00847.

Schulz, K., & Bernhardt, M. (2016). The end of trend estimation for extreme floods under climate change?. *Hydrological Processes*, 30(11), 1804-1808. doi: 10.1002/hyp.10816

Smitha, P., Narasimhan, B., Sudheer, K., & Annamalai, H. (2018). An improved bias correction method of daily rainfall data using a sliding window technique for climate change impact assessment. *Journal of Hydrology*, 556, 100-118. doi: 10.1016/j.jhydrol.2017.11.010

So, B., Kim, J., Kwon, H., & Lima, C. (2017). Stochastic extreme downscaling model for an assessment of changes in rainfall intensity-duration-frequency curves over South Korea using multiple regional climate models. *Journal of Hydrology*, 553, 321-337. doi: 10.1016/j.jhydrol.2017.07.061

Strupczewski, W., Kochanek, K., & Bogdanowicz, E. (2017). Historical floods in flood frequency analysis: Is this game worth the candle?. *Journal of Hydrology*, 554, 800-816. doi: 10.1016/j.jhydrol.2017.09.034

Tao, S., Dong, S., Wang, N., & Guedes Soares, C. (2013). Estimating storm surge intensity with Poisson bivariate maximum entropy distributions based on copulas. *Natural Hazards*, 68(2), 791-807. doi: 10.1007/s11069-013-0654-6

Tosunoglu, F., & Can, I. (2016). Application of copulas for regional bivariate frequency analysis of meteorological droughts in Turkey. *Natural Hazards*, 82(3), 1457-1477. doi: 10.1007/s11069-016-2253-9

Van de Vyver, H. (2018). A multiscaling-based intensity-duration-frequency model for extreme precipitation. *Hydrological Processes*. doi: 10.1002/hyp.11516

Vu, M., Raghavan, V., & Liong, S. (2016). Deriving short-duration rainfall IDF curves from a regional climate model. *Natural Hazards*, 85(3), 1877-1891. doi: 10.1007/s11069-016-2670-9

Yan, L., Xiong, L., Liu, D., Hu, T., & Xu, C. (2016). Frequency analysis of nonstationary annual maximum flood series using the time-varying two-component mixture distributions. *Hydrological Processes*, 31(1), 69-89. doi: 10.1002/hyp.10965

Yang, J., Chang, J., Wang, Y., Li, Y., Hu, H., & Chen, Y. et al. (2018). Comprehensive drought characteristics analysis based on a nonlinear multivariate drought index. *Journal of Hydrology*, 557, 651-667. doi: 10.1016/j.jhydrol.2017.12.055

Yue, S. (2000). The Gumbel Mixed Model Applied to Storm Frequency Analysis. *Water Resources Management*, 14(5), 377-389. doi: 10.1023/a:1011124423923

Zhang, R., Chen, X., Wang, H., Cheng, Q., Zhang, Z., & Shi, P. (2017). Temporal change of spatial heterogeneity and its effect on regional trend of annual precipitation heterogeneity indices. *Hydrological Processes*, 31(18), 3178-3190. doi: 10.1002/hyp.11225