

**ESTIMACIÓN DEL RÉGIMEN HIDROLÓGICO DE EVENTOS
EXTREMOS EN LA CUENCA DEL RÍO BOGOTÁ CONSIDERANDO
LAS FASES DEL ENSO**

Presentado por:

Christian Camilo Orozco Ramírez

Director:

Ing. Darwin Mena Rentería

UNIVERSIDAD SANTO TOMÁS DE AQUINO

FACULTAD DE INGENIERÍA AMBIENTAL

TESIS DE GRADO

BOGOTÁ D.C

2015

Contenido

Resumen.....	3
Introducción.....	3
Objetivos.....	4
General.....	4
Específicos.....	4
Marco de referencia.....	4
Marco contextual.....	4
<i>Amenaza por inundaciones en la cuenca</i>	7
Marco teórico.....	8
<i>Distribuciones de probabilidad mixtas</i>	10
<i>Hipótesis estadísticas</i>	11
Desarrollo central.....	17
Metodología.....	17
<i>Pruebas estadísticas de hipótesis</i>	19
<i>Función de distribución de probabilidad mixta</i>	19
Resultados y análisis.....	20
<i>Pruebas estadísticas de hipótesis</i>	20
<i>Función de distribución de probabilidad mixta</i>	21
Conclusiones.....	26
Recomendaciones.....	27
Bibliografía.....	27

Resumen

Las metodologías tradicionales de estimación de eventos extremos en hidrología se basan en hipótesis de independencia y estacionalidad en las series de caudales extremos (máximos ó mínimos) anuales. Estas hipótesis no son ciertas dada la persistencia hidrológica introducida por la influencia de fenómenos macro-climáticos como el fenómeno El Niño/Oscilación del Sur (ENSO), y sus dos fases: El Niño y La Niña, así como por los efectos del cambio y la variabilidad climática en las variables del ciclo hidrológico, y de los cambios en el uso del suelo y la deforestación. En éste estudio se aplicó una metodología alternativa para la estimación de caudales máximos en la cuenca del rio Bogotá, teniendo en consideración el fenómeno ENSO como parte determinante en la hidrología del país. Dicha metodología se basa en la aplicación de una función de distribución de probabilidad conjunta mediante la ponderación, en este caso el factor de ponderación se obtuvo del número de datos de la serie que corresponden a cada fase del ENSO, así como los periodos No ENSO o Normales; Se utilizó para este fin la distribución Gumbel y se estimaron los caudales máximos para los periodos de retorno de 2.33, 10, 25, 50 y 100 años. Haciendo uso de distintas pruebas estadísticas se comprobó que las hipótesis de independencia, estacionalidad y homogeneidad para el tratamiento de datos no se cumplen en ninguna de las 15 estaciones pertenecientes al estudio. Se encontró además, que en el 53% de los casos el caudal estimado para el periodo de 2.33 años mediante la distribución mixta es superior que las estimaciones hechas mediante el método tradicional, aunque igualmente se pudo observar que en los siguientes periodos de retorno los caudales del método tradicional superan levemente a los estimados mediante la distribución mixta.

Palabras clave: Distribución de probabilidades, Análisis de frecuencia, ENSO, caudales extremos.

Introducción

El cálculo de caudales extremos se ha llevado a cabo a través de modelos estadísticos clásicos de frecuencias, los cuales no son adecuados o no reflejan la situación real de los eventos hidrometeorológicos, ya que asumen que las series de registros máximos o mínimos anuales son independientes, estacionarios y homogéneos.

En las dos últimas décadas se ha evidenciado que las variables hidrológicas están fuertemente influenciadas por fenómenos macro-climáticos de escala global. En particular, a la escala de tiempo interanual el sistema El Niño/Oscilación del Sur (ENSO) es el principal mecanismo regulador de la variabilidad de la hidrología a escala global. [1]

Lo planteado anteriormente permite deducir que las metodologías tradicionales de análisis de frecuencias de eventos extremos en hidrología, adolecen de muchas limitaciones tanto desde el punto de vista de los supuestos metodológicos como de la física que gobierna los eventos extremos. Por ello, este trabajo investiga y aplica en la cuenca del río Bogotá una metodología reciente como es el análisis de frecuencias de probabilidad mixta, para enfrentar los problemas discutidos anteriormente.

Objetivos

General

- Calcular la ocurrencia de caudales extremos (máximos) para la cuenca del río Bogotá.

Específicos

- Evaluar la Hipótesis de independencia en series de duración anual.
- Evaluar la Hipótesis de estacionariedad en series de duración anual.
- Aplicar las Funciones de distribución de probabilidad (FDP) mixtas.

Marco de referencia

Marco contextual

La cuenca del río Bogotá se encuentra localizada en el departamento de Cundinamarca y junto con los ríos Sumapaz, Magdalena, Negro, Minero, Suárez, Blanco, Gacheta y Mchetá. Tiene una superficie total de 589.143 hectáreas que corresponden a aproximadamente el 32% del total de la superficie departamental de los 46 municipios que la conforman, los cuales están representados en la tabla 1. [2]

A su vez, el río Bogotá constituye la corriente principal de la cuenca, recorriendo desde su nacimiento a los 3300 msnm en el municipio de Villapinzón, que pertenece a la subcuenca río Alto Bogotá, hasta su desembocadura al río Magdalena a los 280 msnm en el municipio de Girardot, el cual hace parte de la subcuenca río Bajo Bogotá Apulo – Girardot, un total de 308 kilómetros. A continuación se encuentran las diferentes subcuencas que conforman la cuenca del Río Bogotá con sus respectivas áreas dentro de la misma. [2]

No.	SUBCUENCA	AREA HA	%
1	Rio Bogota (Sector Tibitoc-Soacha)	71284	12.1
2	Rio Balsillas	62442	10.6
3	Rio Bajo Bogota	54431	9.2
4	Rio Apulo	48505	8.2
5	Rio Neusa	44735	7.6
6	Rio Tunjuelito	41535	7.0
7	Embalse Tomine	37428	6.4
8	Rio Teusaca	35818	6.1
9	Rio Medio Bogota (Sector Salto-Apulo)	31650	5.4
10	Rio Alto Bogota	27615	4.7
11	Rio Calandaima	26840	4.6
12	Rio Bogota (Sector Sisga - Tibitoc)	25397	4.3
13	Rio Frio	20160	3.4
14	Embalse Sisga	15526	2.6
15	Rio Chicu	14189	2.4
16	Embalse del Muña	13422	2.3
17	Rio Bogota (Sector Soacha - Salto)	10725	1.8
18	Rio Soacha	4052	0.7
19	Rio Negro	3390	0.6
	TOTALES	589143	100.0

Tabla 1. Subcuencas pertenecientes a la cuenca Hidrográfica del Río Bogotá [2]
Fuente: Corporación Autónoma Regional de Cundinamarca

La cuenca del río Bogotá cuenta con una población aproximada de 1.297.752 habitantes, de los cuales el 75 % corresponde a población urbana y el 25% restante a población rural. Las subcuencas río Apulo, río Calandaima, río Medio Bogotá sector Salto Tequendama – Apulo, río Teusacá, río Neusa, río Bogotá sector Sisga – Tibitoc, Embalse Tominé y embalse Sisga, se caracterizan por su predominio de población rural. [2]

Por otra parte las subcuencas río Balsillas, río Bogotá sector Tibitoc – Soacha y río Bajo Bogotá Apulo – Girardot, son las que mayor número de habitantes aportan al total de población de la cuenca. [2]

Con relación a la densidad poblacional que se encuentran en las diferentes subcuencas, río Negro y río Soacha, las cuales son dos de las que tienen menor tamaño en relación a las demás, presentan el mayor número de habitantes por unidad de superficie, con 2305 y 2500 habitantes por kilómetro cuadrado, respectivamente. Al contrario se encuentran las subcuencas Embalse Sisga y Embalse Tominé, que a su vez presentan las menores densidades poblacionales con 36 y 44 habitantes por kilómetro cuadrado. [2]

MUNICIPIO	AREA MUN ha	AREA EN CUENCA ha	% MUNIC EN CUENCA
AGUA DE DIOS	8567	7033	82
ANAPOIMA	12377	12359	100
ANOLAIMA	12045	11010	91
BOGOTA D.C.	163617	84531	52
BOJACA	10061	10061	100
CACHIPAY	5388	5388	100
CAJICA	5157	5157	100
CHIA	7928	7928	100
CHOCONTA	30178	25383	84
COGUA	13289	13278	100
COTA	6041	6041	100
CUCUNUBA	11163	1280	11
EL COLEGIO	11767	11753	100
EL ROSAL	9791	8492	87
FACATATIVA	15405	15164	98
FUNZA	6731	6731	100
GACHANCIPA	4165	4165	100
GIRARDOT	12979	7691	59
GRANADA	6739	1689	25
GUASCA	36457	21144	58
GUATAVITA	24561	15028	61
LA CALERA	33239	19223	58
LA MESA	14338	14295	100
MADRID	11829	11829	100
MOSQUERA	10822	10822	100
NEMOCON	9906	9906	100
OSPINA PEREZ	11876	11875	100
QUIPILE	12619	3130	25
RICARTE	12810	8491	66
SAN ANTONIO DEL TEQUENDAMA	8845	8837	100
SESQUILE	14125	14093	100
SIBATE	12269	9313	76
SOACHA	18148	16950	93
SOPO	11045	11045	100
SUBACHOQUE	19473	17374	89
SUESCA	17282	12217	71
TABIO	7583	7583	100
TAUSA	19281	14219	74
TENA	5114	5114	100
TENJO	11200	11200	100
TOCAIMA	24722	24243	98
TOCANCIPA	7321	7321	100
VILLAPINZON	22596	12772	57
VIOTA	20339	20130	99
ZIPACON	5872	5872	100
ZIPAQUIRA	19513	18438	94
Totales (ha) de 46 municipios	776573	587597	

Tabla 2. Municipios pertenecientes a la Cuenca del Río Bogotá [2]
Fuente: Corporación Autónoma Regional de Cundinamarca

Amenaza por inundaciones en la cuenca

Para la amenaza por inundación se tiene en cuenta el estudio de hidrología e hidráulica [2], donde se determinan los niveles de agua para diferentes periodos de retorno, la zona de amenaza alta está dada por el nivel correspondiente al periodo de retorno de 100 años, la amenaza media está dada por una franja paralela al límite de la zona de amenaza alta que dependiendo de la legislación que se aplique puede tener 1.5 m en Bogotá o más, la zona de amenaza alta más la zona de amenaza media en el caso del presente estudio se define como zona de ronda. Como para los objetos del presente estudio no es posible obtener secciones transversales a los cursos de los ríos, esta actividad se debe dirigir hacia el cruce de mapas cualitativos como son los de suelos, en particular lo referente a los suelos que se inundan periódicamente, el mapa de geomorfología en donde se identifican formas de relieve y formaciones superficiales susceptibles de sufrir inundaciones y el mapa de pendientes. [3] [2]

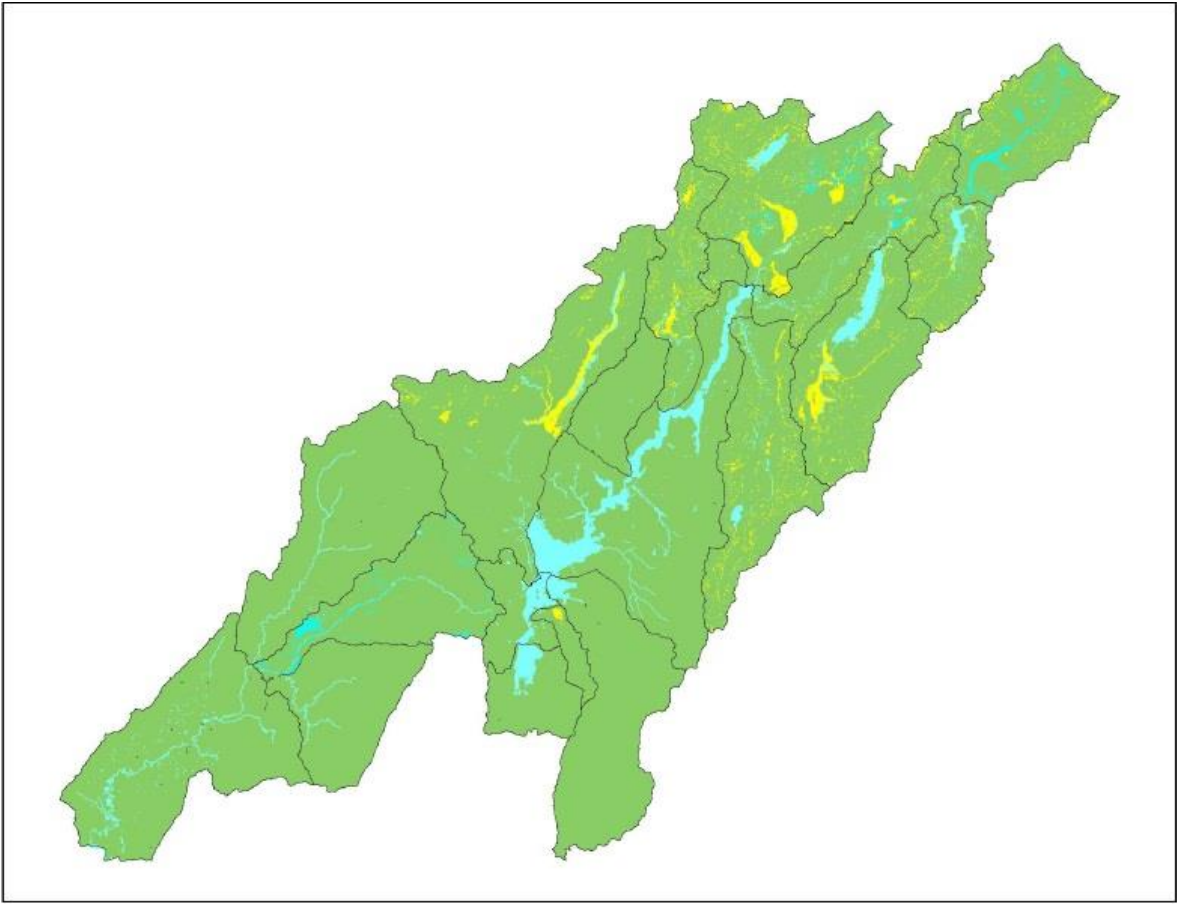


Ilustración 1. Zonas de amenaza por inundación
Fuente: Corporación Autónoma Regional de Cundinamarca [2]

El color azul en el mapa indican las áreas de alta amenaza por inundaciones, las cuales se ubican principalmente en las subcuencas río Bogotá sector Tibitoc – Soacha, río Bogotá sector Salto Soacha, Embalse Muña y Embalse Tominé. [2] [3]

Marco teórico

Las metodologías tradicionales para el cálculo de caudales extremos que están hechas con funciones de distribución de probabilidad, están limitadas por la hipótesis de que los conjuntos de datos son homogéneos y que no varían considerablemente en el tiempo [4]. Dichas suposiciones caducan desde uno de los mismos precursores de éstos métodos para la aplicación y predicciones en la realidad, en el trabajo de E. J. Gumbel [5] se afirma que “Con el propósito de aplicar cualquier teoría suponemos que los datos son homogéneos, por ejemplo, que no hay cambios sistemáticos como el cambio climático ni cambios importantes en la cuenca durante el periodo de observación y estos se darán en el periodo para el cual las extrapolaciones son hechas”.

Un análisis de frecuencia puede estar basado en datos de una sola estación hidrometeorológica o de una red completa de estaciones, aunque también puede realizarse usando los dos tipos.

Las distribuciones de probabilidad se utilizan en diferentes tipos de estudios hidrológicos, particularmente en estudios en los cuales se hace la estimación de caudales extremos, tanto máximos como mínimos, crecidas, volúmenes en embalses, precipitación; entre otros. [6]

El análisis de frecuencias es una herramienta primordial para la predicción de eventos extremos en el contexto hidrológico. Se debe tener en cuenta que cada muestra en un análisis de frecuencia tiene su propia función de distribución o de densidad de probabilidad, la cual comúnmente no se conoce antes de realizar el estudio. [7]

Distribución normal (N)

Es útil en hidrología para describir fenómenos de comportamiento moderado, como el caudal total anual. No existe coeficiente de asimetría en este tipo de distribuciones, ya que estas son simétricas. La función de distribución acumulativa de la distribución normal no tiene una expresión conocida, pero los libros de estadística contienen tablas de la variable aleatoria normal normalizada (z_p). La cantidad de z_p es un ejemplo de factor de frecuencia, ya que el p-ésimo cuantil x_p de una distribución normal cuya media es μ y cuya varianza es s^2 puede describirse así: $x_p = \mu + s z_p$ [6]

Distribución log-normal

Las distribuciones de crecida presentan una asimetría positiva y no son adecuadamente descritas por una distribución normal. Frecuentemente, los logaritmos de una variable aleatoria X no se ajustan a una distribución normal. En tales casos, el problema puede resolverse introduciendo un parámetro de contorno z antes de calcular los logaritmos, obteniéndose así una distribución log-normal triparamétrica (LN3) [6]

A continuación se hará una breve descripción de tres de las distribuciones más frecuentemente usadas para el estudio de extremos [6].

Distribución de Gumbel

Los flujos presentados en el año corresponden al valor máximo de todos los flujos de crecida que suceden anualmente. La definición nos dice que si X_1, \dots, X_n son un conjunto de caudales anuales máximos, y $X = \max\{X_i\}$. Si los valores de X_i son independientes y son variables aleatorias sin un límite superior, independientes y distribuidas idénticamente, con un extremo superior de forma aproximadamente exponencial, para valores grandes de n , la variable aleatoria X presentará una distribución de valores extremos de tipo I o una distribución de Gumbel con una función de distribución acumulativa. [8]

Distribución de valores extremos generalizada

La distribución de valores extremos generalizada abarca los tres tipos de distribuciones de valores extremos respecto de máximos.

La distribución de Gumbel es un caso especial de distribución de valores extremos generalizada, que corresponde a $k = 0$. En esta expresión, x es un parámetro de ubicación, α es un parámetro de escala, y k es el importante parámetro de configuración. Para $k > 0$, tiene una medida superior finita o parámetro de forma, el cual es $\xi + \alpha/k$; mientras que para $k < 0$, no tiene dicho límite superior [6].

Distribución biparamétrica de Weibull

Cuando W_i es el caudal mínimo en diferentes días del año, el mínimo anual será el menor de los W_i , cada uno de los cuales está limitado por el valor cero como mínimo. En tales casos, la variable aleatoria $X = \min\{W_i\}$ puede ser adecuadamente descrita mediante una distribución de valores extremos de tipo III para mínimos, o distribución de Weibull. Para $k < 1$, la probabilidad de Weibull tiende a infinito cuando X se aproxima a cero, y cae lentamente para valores de X muy elevados. Para $k = 1$, la distribución de Weibull reduce la distribución exponencial correspondiente a $y = 2$. Para $k > 1$, la función de densidad de Weibull es similar a una función de densidad de una distribución de Pearson de tipo III para valores pequeños de X y de $\alpha_{P3} = k$, pero tiende a cero más rápidamente para valores de x elevados. [6].

Aunque de estas tres, Las distribuciones de Gumbel y la de valores extremos generalizada, se utilizan ampliamente para el análisis de frecuencia de crecida en todo el mundo [9].

En la mayoría de los análisis de frecuencias se considera que los datos provienen de una sola población básica. Sin embargo, es muy frecuente encontrar que los fenómenos son el resultado de factores causantes distintos. En estos casos, la distribución de la variable aleatoria es una combinación de dos o más distribuciones básicas. [10]

Distribuciones de probabilidad mixtas

El concepto de probabilidad se apoya en las hipótesis de que el espacio de eventos S es finito y que los eventos elementales del espacio S son igualmente posibles. Como es sabido, la probabilidad de que ocurra un cierto evento está entre [0,1] siendo estos extremos los casos de un evento imposible y uno seguro, respectivamente. [10]

Si A y B son conjuntos de eventos elementales en S, es posible asociar dos nuevos eventos dados por las condiciones "tanto A como B ocurren" y "o bien A, o bien B, o bien ambos ocurren". En el primer caso el nuevo evento contiene todos los eventos elementales que son comunes a A y B y lo llamamos intersección de eventos; en el segundo caso el nuevo evento contiene todos los puntos muestrales, con la excepción de los que no pertenecen a A ni a B y lo llamamos unión de eventos. Si dos eventos no pueden ocurrir simultáneamente, se dice que son mutuamente excluyentes y se cumple que $A \cap B = \emptyset$. Existen dos leyes básicas para la determinación de las probabilidades de eventos resultantes de la unión y la intersección de eventos elementales, y son conocidas como las leyes de adición y multiplicación de probabilidades, las cuales se generalizan para el caso de n eventos como se indica a continuación. [10]

Ley de adición de probabilidades

La generalización de esta ley ha sido referida como la fórmula de inclusión-exclusión.

Sean E_1, E_2, \dots, E_n , n eventos arbitrarios. Entonces la fórmula de inclusión-exclusión resulta

$$P\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n P(E_i) - \sum_{1 \leq i_1 < i_2 \leq n} P(E_{i_1} \cap E_{i_2}) + \sum_{1 \leq i_1 < i_2 < i_3 \leq n} P(E_{i_1} \cap E_{i_2} \cap E_{i_3}) \cdot \dots \\ + (-1)^{n+1} \sum_{1 \leq i_1 < i_2 < \dots < i_n \leq n} P(E_{i_1} \cap E_{i_2} \dots \cap E_{i_n})$$

Llamando

$$S_{k,n} = \sum_{1 \leq i_1 \leq i_2 < \dots < i_n \leq n} P(E_{i_1} \cap E_{i_2} \cap \dots \cap E_{i_n})$$

Y recordando que

$$P\left(\bigcup_{i=1}^n E_i\right) = 1 - P\left(\bigcap_{i=1}^n E'_i\right)$$

Donde E'_i es el evento complementario a E_i . Se puede expresar también que:

$$P\left(\bigcap_{i=1}^n E'_i\right) = 1 - P\left(\bigcap_{i=1}^n E_i\right) = \sum_{i=0}^n (-1)^i S_{i,n}$$

Donde $S_{0,n}=1$.

Ley de multiplicación de probabilidades

La generalización de la ley de multiplicación de probabilidades para n eventos $E_{1,i}$, $i = 1, 2, \dots, n$ está dada por la siguiente expresión, válida para $n \geq 2$:

$$P\left(\bigcap_{i=1}^n E_i\right) = \prod_{i=1}^n P\left(E_i \mid \bigcap_{j=1}^{i-1} E_j\right)$$

En donde $P(B|A)=P(A \cap B)/P(A)$ es la probabilidad condicional de B dado que haya ocurrido A .

Para el caso de n eventos estadísticamente independientes $E_i, i=1, 2, \dots, n$ se cumple que:

$$P\left(\bigcap_{i=1}^n E_i\right) = \prod_{i=1}^n P(E_i)$$

Para los fines de éste estudio, se realizó el cálculo haciendo uso de distribuciones de probabilidad mixta, las cuales son el resultado de la suma ponderada de las funciones de distribución de probabilidades de cada una de las muestras seleccionadas. El factor de ponderación se estima como la relación entre los datos de cada muestra y los datos de la serie completa. [4]

Hipótesis estadísticas

Para el desarrollo de éste tipo de estudios debe tenerse en cuenta el desarrollo de pruebas para confirmar o desmentir las hipótesis de independencia, homogeneidad y estacionalidad de las series de datos, para lo cual existen ciertas pruebas

recomendadas que tienen gran aceptación en las aplicaciones de estadística en hidrología [11] [6]. A continuación se explican algunas de éstas pruebas:

Independencia y estacionalidad de Wald-Wolfowitz [12] [13]

Dada una muestra de datos de tamaño N (x_1, \dots, x_N), la prueba de Wald-Wolfowitz considera la estadística R de tal modo que:

$$R = \sum_{i=1}^{N-1} x_i x_{i+1} + x_1 x_N$$

Cuando los elementos de la muestra son independientes, R se ajusta asintóticamente a una distribución normal, cuya media y varianza vienen dadas por las ecuaciones siguientes:

$$\bar{R} = (s^2_2 - s_2)/(N - 1)$$

$$Var(R) = \frac{s^2_2 - s_4}{N - 1} - \bar{R}^2 + \frac{(s^4_1 - 4s^2_1 s_2 + 4s_1 s_3 + s^2_2 - 2s_4)}{(N - 1)(N - 2)}$$

Donde $s_r = N m^r$ y m^r es el r-ésimo momento de la muestra en torno al origen.

La magnitud $(R - \bar{R}) / \sqrt{Var(R)}$ presenta una distribución normal normalizada (media = 0 y varianza = 1), y puede utilizarse para comprobar al nivel α a la hipótesis de independencia comparando $|n|$ con la variable aleatoria normal normalizada $u_{\alpha/2}$ correspondiente a una probabilidad de excedencia $\alpha/2$.

Prueba de Mann-Whitney [14]

Este procedimiento es una buena alternativa cuando no se puede utilizar la prueba t de Student, por no cumplir con los requisitos que esta prueba exige. La hipótesis nula es que ambas muestras provienen de una misma población o distribución.

H_0 = Ambas muestras provienen de una misma distribución

H_1 = Las muestras no provienen de la misma distribución

Para el estadístico de prueba se requiere conocer:

n_1 = tamaño de la muestra del grupo 1.

n_2 = tamaño de la muestra del grupo 2.

Se ordenan todos los datos del menor al mayor valor, y se le asignan rangos. La idea es que si ambas muestras provienen de la misma distribución, los rangos de uno y otro grupo van a tener una magnitud global similar (relativa al tamaño de las muestras).

Se definen:

R_1 = sumatoria de los rangos del grupo 1.

R_2 = sumatoria de los rangos del grupo 2.

Si n_1 y n_2 son menores que 20, (muestras pequeñas) se definen los valores estadísticos de Mann-Whitney:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - \sum R_1$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - \sum R_2$$

Se elige el más pequeño de ambos para comparar con los valores críticos de U de la tabla de probabilidades asociadas al test de Mann-Whitney

En caso de que n_1 y n_2 sean mayores que 20 (muestras grandes), los estadísticos U se distribuye en forma aproximadamente gaussiana y se calcula el valor $Z_{emp} \sim N(0,1)$

$$Z_{emp} = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (N + 1)}{12}}}$$

Luego se decide si se acepta o rechaza la hipótesis, de acuerdo al nivel de significación elegido.

Prueba de Rachas para Independencia [15]

Esta prueba se basa en el número de rachas que presenta una muestra. Una racha se define como una secuencia de valores muestrales con una característica común precedida y seguida por valores que no presentan esa característica. Así, se considera una racha la secuencia de k valores consecutivos superiores o iguales a la media muestral (o a la mediana o a la moda, o a cualquier otro valor de corte) siempre que estén precedidos y seguidos por valores inferiores a la media muestral (o a la mediana o a la moda, o a cualquier otro valor de corte).

Si la muestra es suficientemente grande y la hipótesis de independencia es cierta, la distribución muestral del número de rachas, R , puede aproximarse mediante una distribución normal de parámetros:

$$\mu_R = \frac{2n_1 n_2}{n} + 1$$

$$\sigma_R = \sqrt{\frac{2n_1n_2(2n_1n_2 - n)}{n^2(n - 1)}}$$

Donde n_1 es el número de elementos de una clase, n_2 es el número de elementos de la otra clase y n es el número total de observaciones.

$$Z = \frac{R + c - \mu_R}{\sigma_R}$$

Donde $c = 0,5$ si $R < \mu_R$ y $c = -0,5$ si $R > \mu_R$

Prueba de independencia de Anderson [13] [12]

Con el uso del coeficiente de autocorrelación serial r_k^j para diferentes tiempos de retraso k . En el caso de analizar un solo registro, entonces $j = 1$.

$$r_k^j = \frac{\sum_{i=1}^{n_j-k} (Q_i^j - \bar{Q}^j) \cdot (Q_{i+k}^j - \bar{Q}^j)}{\sum_{i=1}^{n_j} (Q_i^j - \bar{Q}^j)^2}$$

Para: $k=1, 2, \dots, n_j/3$; donde:

$$\bar{Q}^j = \sum_{i=1}^{n_j} \frac{Q_i^j}{n_j}$$

Además, los límites al 95% de confianza para r_k^j se pueden obtener con:

$$r_k^j(95\%) = \frac{-1 \pm 1,96\sqrt{(n_j - k - 1)}}{n_j - k}$$

Si la gráfica de los valores estimados para r_k^j versus los tiempos de retraso k , junto con sus correspondientes límites de confianza, se le llama correlograma de la muestra.

Si no más del 10% de los valores r_k^j sobrepasan los límites de confianza, se dice que la serie Q es independiente y por lo tanto es una variable que sigue las leyes de la probabilidad.

Coefficiente de correlación de Spearman

El coeficiente de correlación de Spearman (ρ) es una medida de la correlación entre dos variables aleatorias continuas. Para calcular ρ , los datos son ordenados y reemplazados por su respectivo orden.

El estadístico ρ viene dado por la expresión:

$$\rho = 1 - \frac{6 \sum D^2}{N(N^2 - 1)}$$

Donde D es la diferencia entre los correspondientes estadísticos de orden de x - y. N es el número de parejas.

Para muestras mayores de 20 observaciones, podemos utilizar la siguiente aproximación a la distribución t de Student

$$t = \frac{\rho}{\sqrt{(1 - \rho^2)/(n - 2)}}$$

La interpretación de coeficiente de Spearman está entre -1 y +1, indicando asociaciones negativas o positivas respectivamente, 0 cero, significa no correlación pero no independencia.

Prueba de homogeneidad de Helmer [13] [12]

Consiste en analizar el signo de las desviaciones de cada evento Q_i^j de la serie j para $i = 1, 2, 3, \dots, n_j$, con respecto a su valor medio Q^j .

Si una desviación de un cierto signo es seguida por otra del mismo signo, entonces se dice que se forma una secuencia: S, de lo contrario se considera un cambio: C.

La serie se considera homogénea si se cumplen las siguientes condiciones:

$$-\sqrt{n_j - 1} \leq (S - C) \leq \sqrt{n_j - 1}$$

Prueba T de student para Homogeneidad [13] [12]

Si se considera una serie Q_i^j para $i = 1, 2, 3, \dots, n_j$, del sitio j, la cual se divide en dos conjuntos de tamaño $n_1 = n_2 = n_j/2$, entonces el estadístico de prueba se define con:

$$t_d = \frac{\bar{x}_1 - \bar{x}_2}{\left[\frac{n_1 \cdot s_1^2 + n_2 \cdot s_2^2}{n_1 + n_2 - 2} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \right]^{1/2}}$$

\bar{x}_1, s_1^2 = Son la media y la varianza de la primera parte del registro de tamaño n_1

\bar{x}_2, s_2^2 = Son la media y la varianza de la segunda parte del registro de tamaño n_2

El valor absoluto de t_d se compara con el valor de la distribución t de Student de dos colas y con $n = n_1 + n_2 - 2$ grados de libertad y para un nivel de significancia: $\alpha = 0,05$

Si y solo si el valor absoluto de t_d es mayor que aquel de la distribución t de Student, se concluye que la diferencia entre las medias es evidencia de inconsistencia, y por lo tanto la serie Q^j se considera no homogénea.

Prueba de Homogeneidad de Cramer [13] [12]

Esta prueba se utiliza con el propósito de verificar homogeneidad en el registro Q^j de la serie j para $i = 1, 2, 3, \dots, n_j$, y también para determinar si el valor medio no varía significativamente de un período de tiempo a otro. Con este propósito se consideran tres bloques, el primero del tamaño total de la muestra, n_j , el segundo de tamaño n_{60} (últimos 60% de los valores de la muestra) y el tercero de tamaño n_{30} (últimos 30% de los valores de la muestra).

La prueba compara el valor Q^j del registro total con cada una de las medias de los bloques elegidos Q_{60}^j y Q_{30}^j .

Para que se considere la serie analizada como estacionaria en la media, se deberá cumplir que no existe una diferencia significativa entre las medias de los dos bloques.

$$\bar{Q}^j = \sum_{i=1}^{n_j} \frac{Q_i^j}{n_j}$$

Para una sola muestra analizada $j = 1$.

$$S_Q^j = \left[\frac{1}{(n_j - 1)} \cdot \sum_{i=1}^{n_j} (Q_i^j - \bar{Q}^j)^2 \right]^{1/2}$$

$$\overline{Q}_{60}^j = \sum_{k=1}^{n_{60}} \frac{Q_k^j}{n_{60}}$$

$$\overline{Q}_{30}^j = \sum_{k=1}^{n_{30}} \frac{Q_k^j}{n_{30}}$$

$$\tau_{60}^j = \frac{\overline{Q}_{60}^j - \overline{Q}^j}{S_Q^j}$$

$$\tau_{30}^j = \frac{\overline{Q}_{30}^j - \overline{Q}^j}{S_Q^j}$$

$$t_w = \left\{ \frac{n_w \cdot (n_j - 2)}{n_j - n_w \cdot [1 + (\tau_w^j)^2]} \right\}^{\frac{1}{2}} |\tau_w^j|$$

El estadístico t_w tiene distribución t de Student de dos colas con $v = n_1 + n_2 - 2$ grados de libertad y para un nivel de significancia $\alpha = 0,05$.

Si y solo si el valor absoluto de t_w para $w = 60$ y $w = 30$, es mayor que el de la distribución t de Student se concluye que la diferencia entre las medias es evidencia de inconsistencia y por lo tanto la serie Q^j se considera no homogénea. En caso contrario la serie es Homogénea.

Desarrollo central

Metodología

Se seleccionaron las estaciones contenidas en las bases de datos de caudales máximos mensuales multianuales tanto del Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) y de la Corporación Autónoma Regional de Cundinamarca (CAR) pertenecientes a la cuenca del Río Bogotá; que a su vez cumplieran con los siguientes parámetros:

- Registro mínimo de 15 años.
- Años con menos de 6 datos mensuales faltantes.

De acuerdo con estos parámetros se obtuvo un total de 15 estaciones aptas para realizar el estudio, de las cuales 14 pertenecen a la CAR y la restante al IDEAM, respectivamente, las cuales se presentan en la tabla 3.

Código Estación	Nombre Estación	Tipo de Estación	Entidad
2120742	BALSA LA	LG	CAR
2120920	CAMPIÑA LA	LM	CAR
2120793	ESPIÑO EL	LG	CAR
2120882	PORTILLO EL	LM	CAR
2120917	PTE CHOCONTA	LM	CAR

2120714	PTE CUNDINAMARCA	LM	CAR
2120767	PTE FLORENCIA	LG	CAR
21207960	PTE PORTILLO	LG	IDEAM
2120734	PTE VARGAS	LG	CAR
2120719	SAUCIO	RN	CAR
2120891	SILENCIO EL	LM	CAR
2120816	STA ROSITA	LM	CAR
2120890	TOCAIMA	LM	CAR
2120792	TOCANCIPA	LG	CAR
2120815	VILLAPINZÓN	LM	CAR

Tabla 3. Estaciones aptas para el estudio

Una vez depurados los datos, se procedió a realizar las pruebas estadísticas para verificar las hipótesis de independencia, homogeneidad y estacionalidad; Mediante el uso del software Excel 2013 y SPSS y cuyos resultados fueron registrados en las tablas 5 y 6.

Para la elaboración del análisis de frecuencias mixtas los caudales se discriminaron en tres subconjuntos, así:

- Caudales máximos en periodo ENSO niño.
- Caudales mínimos en periodo ENSO niña.
- Caudales extremos normales o NO ENSO.

Los tres subconjuntos se obtuvieron teniendo en cuenta la tabla 4, sobre el consenso para clasificación de diferentes fases del ENSO [16]

El Niño			La Niña			Normal o NO ENSO 1	
Débil	Moderado	Fuerte	Débil	Moderado	Fuerte		
1963-64	1951-52	1965-66	1954-55	1950-51	1955-56	1952-53	1953-54
1968-69	1957-58	1972-73	1956-57	1970-71	1973-74	1958-59	1959-60,
1969-70	1977-78	1982-83	1962-63	1998-99	1975-76	1960-61	1961-62
1976-77	1987-88	1991-92	1964-65	2000-01	1988-89	1966-67	1978-79
1986-87	1994-95	1997-98	1967-68	2007-08	2010-11	1979-80	1980-81
1992-93	2002-03	1971-72				1981-82	1983-84
2004-05	2009-10	1974-75				1985-86	1989-90
2006-07	1984-85					1990-91	1993-94
1995-96						1996-97	2001-02
1999-00						2003-04	2005-06
2008-09							

Tabla 4. Consenso para clasificación de diferentes fases del ENSO
Fuente: IDEAM [16]

Pruebas estadísticas de hipótesis

Es de suma importancia que los datos cumplan con las hipótesis mencionadas anteriormente de independencia, homogeneidad y estacionalidad;

Con la independencia se busca que la probabilidad de ocurrencia de cualquier evento no dependa de la ocurrencia de los datos que le preceden y que tampoco influya en los datos futuros. Mientras que con la estacionalidad se intenta comprobar que en la serie de datos no se presente una tendencia temporal; Y finalmente para el caso de la homogeneidad se revisa que en los datos no se presente una gran variabilidad en el tiempo, es decir que mantengan cierto tipo de tendencia. [17] [18]

Debido a esta necesidad, se realizaron las pruebas estadísticas de Mann-Whitney, Rachas y el coeficiente de correlación de Spearman para independencia, cuyos resultados se encuentran en la tabla 5. Mientras que para la hipótesis de homogeneidad se realizaron las pruebas de Hermert, T de student y la prueba estadística de Cramer. Finalmente, para la estacionalidad, el resultado de la misma prueba de independencia de Mann-Whitney, realizada para independencia arrojaba el veredicto acerca de éste supuesto.

Función de distribución de probabilidad mixta

La función de distribución de probabilidades mixta se expresa como la suma ponderada de las funciones de probabilidad separadas. Para este caso, los factores de ponderación se obtuvieron a partir de la discriminación sobre las fases del ENSO que determina la relación entre los datos de cada muestra y los datos de la serie completa de cada una de las estaciones.

La representación de la mezcla de funciones de distribución de probabilidades está dada según,

$$F_T(X \geq x) = \rho_1 F_1(X \geq x) + \rho_2 F_2(X \geq x) + \dots + \rho_i F_i(X \geq x)$$

$$\sum_{i=1}^n \rho_i = 1$$

En donde F_T representa la función de distribución de probabilidad de toda la muestra de caudales extremos, F_i es la función de distribución de probabilidad de cada muestra resultante de la discriminación, y ρ_i representa el factor de ponderación de cada muestra.

Posteriormente, haciendo uso del software de análisis hidrológico HidroEsta, se aplicó la distribución Gumbel [8], que fue la seleccionada para éste estudio con periodos de retorno de 2.33, 10, 25, 50 y 100 años; Los cuales con los típicos para

estudios hidrológicos a nivel de manejo de cuencas y diseños de obras hidráulicas. Esto se hizo cada uno de los subconjuntos para todas las estaciones y obteniendo así las estimaciones de los caudales en cada uno de los periodos de retorno mediante la ponderación de los mismos.

Como último paso y con la ayuda del software HidroEsta de nuevo, se procedió a hacer una estimación de los caudales para los mismos periodos de retorno mencionados anteriormente pero sin la discriminación de las diferentes fases del ENSO, es decir, dicha estimación se realizó para cada una de las estaciones con los registros tal cual se encontraban en la base de datos.

Resultados y análisis

Pruebas estadísticas de hipótesis

Con el fin de corroborar que no en todos los casos se cumple la premisa teórica de E.J Gumbel [5] en la cual indica que los datos para ser manejados mediante las distribuciones de probabilidad comunes, las series de registros deben ser independientes, homogéneos y no estacionales se realizaron las pruebas estadísticas mencionadas anteriormente, cuyos resultados están distribuidos en las tablas 5 y 6.

ESTACIÓN	INDEPENDENCIA		
	Mann-Whitney	Rachas	Spearman
BALSA LA	Independiente	Independiente	Positivo
CAMPIÑA LA	Independiente	Independiente	Positivo
ESPIÑO EL	No independiente	No independiente	Positivo
PORTILLO EL	No independiente	No independiente	Positivo
PTE CHOCONTA	Independiente	Independiente	Positivo
PTE CUNDINAMARCA	No independiente	Independiente	Positivo
PTE FLORENCIA	No independiente	No independiente	Positivo
PTE PORTILLO	Independiente	Independiente	Positivo
PTE VARGAS	Independiente	Independiente	Positivo
SAUCIO	Independiente	Independiente	Positivo
SILENCIO EL	No independiente	No independiente	Positivo
STA ROSITA	Independiente	No independiente	Positivo
TOCAIMA	Independiente	No independiente	Positivo
TOCANCIPA	Independiente	Independiente	Positivo
VILLAPINZÓN	Independiente	Independiente	Positivo

Tabla 5. Resultados para las pruebas de independencia

ESTACIÓN	HOMOGENEIDAD	ESTACIONALIDAD
----------	--------------	----------------

	Helmert	T de Student	Cramer	Mann-Whitney
BALSA LA	Homogénea	No homogénea	No homogénea	Estacional
CAMPIÑA LA	No homogénea	Homogénea	No homogénea	Estacional
ESPIÑO EL	No homogénea	Homogénea	No homogénea	No estacional
PORTILLO EL	No homogénea	No homogénea	No homogénea	No estacional
PTE CHOCONTA	No homogénea	No homogénea	No homogénea	Estacional
PTE CUNDINAMARCA	No homogénea	Homogénea	No homogénea	No estacional
PTE FLORENCIA	Homogénea	Homogénea	No homogénea	No estacional
PTE PORTILLO	Homogénea	No homogénea	No homogénea	Estacional
PTE VARGAS	Homogénea	No homogénea	No homogénea	Estacional
SAUCIO	Homogénea	No homogénea	No homogénea	Estacional
SILENCIO EL	Homogénea	No homogénea	No homogénea	No estacional
STA ROSITA	No homogénea	No homogénea	No homogénea	Estacional
TOCAIMA	Homogénea	No homogénea	No homogénea	Estacional
TOCANCIPA	Homogénea	No homogénea	No homogénea	Estacional
VILLAPINZÓN	No homogénea	No homogénea	Homogénea	Estacional

Tabla 6. Resultados para las pruebas de homogeneidad y estacionalidad

Evidentemente se observa que los resultados de las pruebas indican que sin importar que varias de éstas arrojen un resultado positivo a las hipótesis de independencia, homogeneidad y estacionalidad en todos los casos alguna de ellas desarma el argumento teórico de Gumbel acerca del manejo de los datos. Esto se puede ver claramente en que para varias estaciones las pruebas para independencia hayan sido positivas, los resultados de al menos 2 pruebas de homogeneidad resultan descartando esta hipótesis, por ejemplo el caso de las estaciones: BALSA LA, CAMPIÑA LA, PTE CHOCONTA, PTE PORTILLO, PTE VARGAS, SAUCIO, TOCANCIPÁ y VILLAPINZÓN.

El caso contrario no se presenta, es decir, que todas las pruebas de homogeneidad sean positivas y la premisa teórica se derrumbe debido a la independencia, apenas para el caso de la estación PTE FLORENCIA las pruebas de homogeneidad de Cramer y T de Student arrojan un resultado positivo, mientras que las pruebas Mann-Whitney y la prueba de Rachas son negativas para independencia.

Un caso excepcional es el de la estación PORTILLO EL, para la cual los resultados de las pruebas fueron casi en su totalidad negativos, únicamente tuvo un coeficiente de correlación de Spearman positivo, es decir, cercano a 1.

Función de distribución de probabilidad mixta

El análisis tradicional de eventos extremos en hidrología asume las hipótesis de independencia, estacionariedad y homogeneidad en los registros de caudales extremos. Como se ha mencionado con anterioridad, dichas suposiciones son violadas por la variabilidad climática a la que están sometidos los registros históricos.

Centrándose en el objeto de estudio de éste trabajo, los caudales máximos están principalmente influenciados por la ocurrencia de La Niña, durante el cual se presentan lluvias extremas e inundaciones en la mayoría del país. Debido a que el comportamiento de los caudales colombianos está influenciado a su vez por la ZCIT y la geomorfología local, este análisis se realizará por zonas del país.

El primer paso para realizar la función de probabilidad mixta es la discriminación de los datos en los 3 subconjuntos (Caudales máximos en periodos ENSO niña, ENSO niño y años normales o NO ENSO), para los cuales se realizó el conteo y se organizó en tablas separadas para cada evento, el resumen de los datos se encuentra en la tabla 7.

ESTACIÓN	TOTAL DE DATOS POR FASE			
	NIÑO	NIÑA	NO ENSO	TOTAL
BALSA LA	12	6	8	26
CAMPIÑA LA	4	6	6	16
ESPINO EL	12	4	9	25
PORTILLO EL	8	6	5	19
PTE CHOCONTA	7	5	4	16
PTE CUNDINAMARCA	12	5	8	25
PTE FLORENCIA	14	4	12	30
PTE PORTILLO	16	5	16	37
PTE VARGAS	12	6	11	29
SAUCIO	16	6	13	35
SILENCIO EL	8	6	5	19
STA ROSITA	15	6	13	34
TOCAIMA	7	5	4	16
TOCANCIPA	10	5	5	20
VILLAPINZÓN	10	5	9	24

Tabla 7. Total de datos de los 3 subconjuntos de trabajo

Una vez realizada esta etapa, se procede a introducir los datos de cada estación, para todos los subconjuntos en el software HidroEsta haciendo uso de la distribución Gumbel, se registran los resultados de cada una teniendo en cuenta los periodos de retorno seleccionados y por último se aplica la misma distribución a los datos sin discriminar para tener el caudal estimado del modo tradicional.

ESTACIÓN	Periodo de Retorno = 2.33 años
----------	--------------------------------

	NIÑO	NIÑA	NO ENSO	MIXTO (m/s)	TRADICIONAL (m/s)
BALSA LA	34.50	42.7	31.46	35.46	35.31
CAMPIÑA LA	168.65	233.73	82.38	160.70	163.01
ESPINO EL	19.75	22.24	18.5	19.70	20.15
PORTILLO EL	119.47	123.85	93.95	114.14	105.69
PTE CHOCONTA	4.9	5.43	6.91	5.57	5.52
PTE CUNDINAMARCA	48.57	75.59	47.75	53.71	53.58
PTE FLORENCIA	31.13	22.71	24.28	27.27	27.21
PTE PORTILLO	387.71	424.86	339.59	371.92	354.96
PTE VARGAS	34.27	33.74	31.88	33.25	33.90
SAUCIO	30.68	22.47	30.46	29.19	29.57
SILENCIO EL	5.59	5.35	5.03	5.37	5.30
STA ROSITA	22.57	20.37	23.28	22.45	23.71
TOCAIMA	0.93	1.14	0.94	1.00	1.02
TOCANCIPA	31.37	22	24.3	27.26	29.25
VILLAPINZÓN	7.55	6.39	8.35	7.61	8.84

Tabla 8. Caudales obtenidos a partir de la distribución Gumbel para periodo de retorno de 2.33 años

El modelo de la tabla 8 también se aplicó para los periodos de retorno de 10, 25, 50 y 100 años; a continuación se presentan los resultados de las estimaciones de los caudales máximos por el método de las distribuciones de probabilidad mixta contra las estimaciones mediante el método tradicional.

ESTACIÓN	T= 10 años		T= 25 años	
	MIXTO (m/s)	TRADICIONAL (m/s)	MIXTO (m/s)	TRADICIONAL (m/s)
BALSA LA	48.01	49.50	55.12	57.55
CAMPIÑA LA	311.62	338.72	394.39	438.38
ESPINO EL	29.12	29.91	34.46	35.44
PORTILLO EL	164.06	166.10	192.37	200.36
PTE CHOCONTA	7.87	7.90	9.18	9.25
PTE CUNDINAMARCA	76.93	80.85	90.09	96.32
PTE FLORENCIA	41.05	45.53	48.86	51.22
PTE PORTILLO	558.19	506.40	663.84	592.30
PTE VARGAS	45.67	50.82	52.71	60.43
SAUCIO	44.34	47.92	52.94	58.32
SILENCIO EL	6.74	6.91	7.52	7.82
STA ROSITA	31.08	39.19	35.98	47.97
TOCAIMA	1.37	1.39	1.58	1.61
TOCANCIPA	47.37	53.80	58.78	67.72
VILLAPINZÓN	14.57	18.28	18.52	23.63

Tabla 9. Caudales obtenidos a partir de la distribución Gumbel para periodo de retorno de 2.33 años

ESTACIÓN	T= 50 años		T= 100 años	
	MIXTO (m/s)	TRADICIONAL (m/s)	MIXTO (m/s)	TRADICIONAL (m/s)
BALSA LA	60.40	63.52	65.64	69.45
CAMPIÑA LA	455.78	512.32	516.73	585.70
ESPIÑO EL	38.90	39.54	42.36	43.62
PORTILLO EL	213.37	225.78	234.22	251.01
PTE CHOCONTA	10.15	10.26	11.12	11.25
PTE CUNDINAMARCA	99.86	107.80	109.56	119.19
PTE FLORENCIA	54.66	57.66	60.41	64.06
PTE PORTILLO	742.21	656.02	820.01	719.27
PTE VARGAS	57.94	67.55	63.12	74.62
SAUCIO	59.31	66.04	65.64	73.71
SILENCIO EL	8.10	8.50	8.67	9.17
STA ROSITA	39.61	54.49	43.22	60.96
TOCAIMA	1.74	1.77	1.89	1.92
TOCANCIPA	67.24	78.05	75.64	88.30
VILLAPINZÓN	21.45	27.60	24.36	31.54

Tabla 10. Caudales obtenidos a partir de la distribución Gumbel para los periodos de retorno de 50 y 100 años

El análisis tradicional de eventos extremos que se usa en hidrología supone que la función de distribución de probabilidades de los eventos extremos es estacionaria en el tiempo. El conocimiento del ENSO y sus efectos pone de presente que esta suposición no es válida, por lo cual se hace necesario incorporar este efecto en la estimación de caudales extremos de distintos periodos de retorno.

A grandes rasgos se observa que en diferentes estaciones hay un fenómeno particular, en el cual, el primero de los periodos de retorno con los cuales se realizó la estimación (2.33 años) arroja un caudal con una magnitud mayor, para el método de función de distribución de probabilidad mixto en comparación con los caudales estimados mediante el método tradicional; Siendo éste el caso concreto de las estaciones BALSA LA, PORTILLO EL, PTE CHOCONTA, PTE CUNDINAMARCA, PTE FLORENCIA, PTE PORTILLO y SILENCIO EL.

Una posibilidad frente al fenómeno mencionado anteriormente es que se produzca por la incertidumbre que hay en grandes periodos de tiempo con relación a la ocurrencia de alguna de las fases del ENSO, ya sea el niño o la niña, ya que por ejemplo, durante las épocas de la niña en la región andina, se presentan caudales extremadamente altos. Aunque también se debe tener en cuenta que según el consenso para clasificación de las diferentes fases del ENSO [16] se han presentado con más frecuencia periodos de NO ENSO y el niño en relación a los periodos ENSO niña; lo cual también hace suponer que los factores de ponderación para los casos de la niña son muy bajos. En la tabla 8, se aprecia que en 9 de las 15 estaciones las estimaciones para el periodo niña son más las altas, como era de

suponerse, pero posiblemente debido a los factores de ponderación tan bajos, las estimaciones al realizar el cálculo final no son de gran magnitud.

Por otra parte, la única de las estaciones en la cual su estimación mediante el método de probabilidad mixta es superior en todos los periodos de retorno a la estimación mediante el método tradicional es la estación de PTE PORTILLO. Cabe resaltar que las diferencias en las magnitudes de éste caso son las más altas de todo el estudio; se podría presumir que dicho resultado se encuentra bastante influenciado debido a que los registros en ese punto son los más completos, con un total de 37 años, aunque la mayoría de ellos no pertenezcan al periodo ENSO niña, puesto que si fuese así, como se mencionaba anteriormente los resultados podrían ser más abultados. En la figura 1 se muestran los resultados para ésta estación.

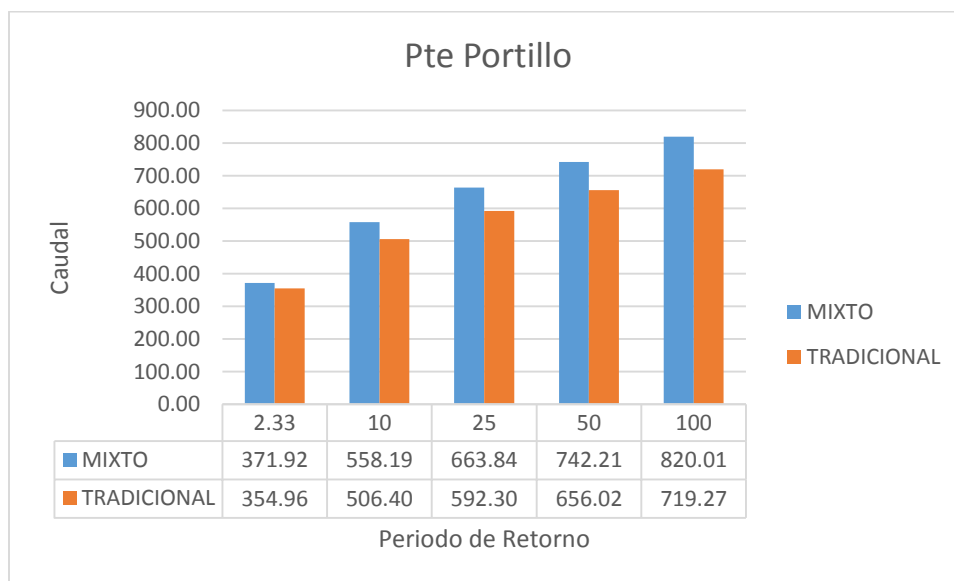


Figura 1. Resultados compilados de la estación PTE PORTILLO

Respecto a los resultados que arrojaron las pruebas estadísticas, se observó que para la estación mencionada anteriormente, solo las pruebas de T de Student y Cramer para homogeneidad fueron negativas, es decir, que PTE PORTILLO es independiente y estacional; El caso contrario es el que se tiene en la estación PORTILLO EL, la cuál fue la única que obtuvo resultados negativos en casi todas las pruebas, como se mencionaba con anterioridad.

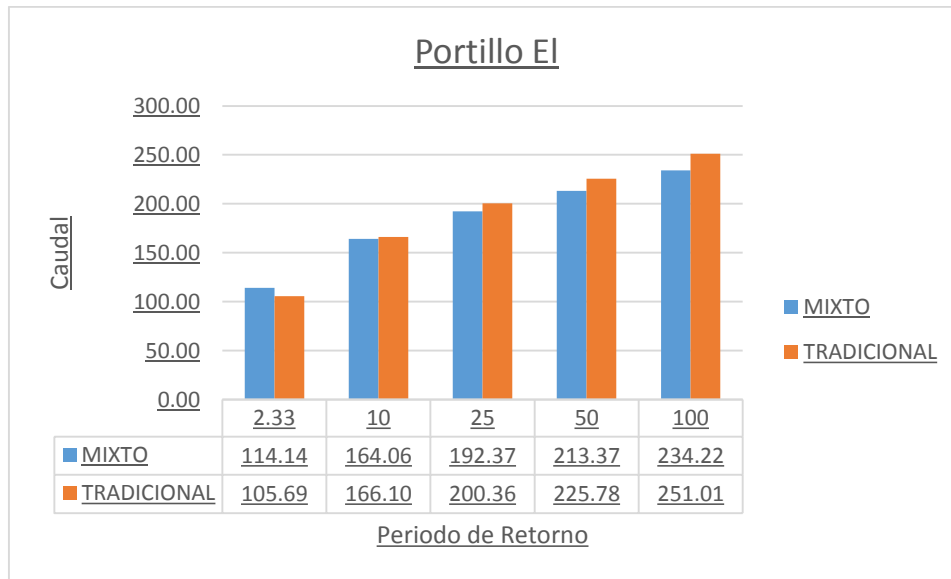


Figura 2. Resultados compilados de la estación PORTILLO EL

La estación PORTILLO EL es un gran ejemplo del fenómeno al que se hacía referencia anteriormente acerca las estimaciones de los caudales para los periodos de retorno a los 2.33 años, y la progresión de los CALCULOS mediante el método tradicional que a partir de los 25 años en periodo de retorno abre una diferencia relativamente amplia sobre el método de estimación mixto, aunque en ésta estación solo se tienen 19 registros.

Conclusiones

1. Este trabajo hace parte de un compendio de estudios que buscan superar las limitaciones de los métodos que se usan tradicionalmente en tareas de ingeniería hidrológica relacionadas con la estimación de caudales extremos, en específico de los máximos, en distintos periodos de retorno. Concretamente, dichos limitantes tienen que ver con las hipótesis de independencia, homogeneidad y estacionalidad de las series de caudales extremos anuales; Y se hace necesario comenzar a tener en cuenta estudios de éste tipo para los trabajos de diseño hidrológico.
2. Para la mayoría de las estaciones pertenecientes a éste estudio, es posible rechazar las hipótesis de independencia de los datos. Por lo tanto, antes de realizar el análisis de frecuencia de cualquier serie hidrológica de caudal máximo, se hace necesario evaluar la hipótesis de independencia de los datos.
3. En todas las estaciones que fueron objeto de estudio en éste trabajo, se rechaza la hipótesis de homogeneidad, lo cual indica que los fenómenos de variabilidad climática en Colombia son bastante influyentes en el régimen hidrológico de la zona de la cuenca del Río Bogotá. Esto significa que las

series de datos tienen diferentes fluctuaciones, por lo cual no presentan un comportamiento determinado.

4. Teniendo en cuenta las evaluaciones que se hicieron a priori, se decidió trabajar con la distribución de Gumbel, debido a que ya ha sido ampliamente usada en estimaciones hidrológicas a través de los años y es una de las distribuciones que mejor se ajusta a los eventos extremos, independientemente que se trabaje con series de datos de caudales máximos y/o mínimos.
5. Las estimaciones realizadas con las respectivas discriminaciones de las fases del ENSO, podrían llegar a ser más altas si los factores de ponderación para los eventos de La Niña, fueran más altos, debido a que en dichos periodos es donde se presentan los caudales más grandes en la zona de la cuenca del Río Bogotá.

Recomendaciones

1. Es de suma importancia conocer a priori los datos que se van a trabajar, debido a que se pueden presentar problemas por no contar con un grueso registro de información hidrológica, con eso es posible tener en cuenta si el estudio se deberá modificar o no.
2. Con el fin de reducir el tiempo para realizar los diferentes cálculos, se sugiere hacer uso de software de análisis estadístico como Minitab o SPSS, sin dejar de lado el tener la comprensión de qué se está haciendo y su respectiva finalidad.
3. Un aporte interesante a éste estudio sería la realización de un nuevo mapa de amenaza por inundación a lo largo de la cuenca, debido a que los cálculos varían considerablemente en puntos como a la altura el municipio de Tocaima

Bibliografía

- [1] D. M. Álvarez, *Nuevos Enfoques Para la Estimación de Caudales Extremos en Colombia*, Medellín, 2007.
- [2] Corporación Autónoma Regional de Cundinamarca (CAR), «Plan de Ordenación y Manejo de la Cuenca Hidrográfica del Río Bogotá,» Bogotá, 2006.
- [3] Corporación Autónoma Regional de Cundinamarca, «Documento Técnico Proyecto: "Adecuación Hidráulica y Recuperación Ambiental Río Bogotá",» Bogotá, 2009.

- [4] G. Poveda y D. M. Álvarez, «El colapso de la hipótesis de estacionariedad por cambio y variabilidad climática: implicaciones para el diseño hidrológico en ingeniería,» *Revista de ingeniería de la Universidad de los Andes*, nº 36, 2012.
- [5] E. J. Gumbel, «The return period of flood flows,» *Ann Math Stat*, pp. 163-190, 1941.
- [6] Organización Meteorológica Mundial, «Guía de Prácticas Hidrológicas,» Ginebra, 2011.
- [7] N. Suárez Buenrostro, *Caracterización de eventos extremos de precipitación, empleando distribuciones mezcladas*, Querétano, 2011.
- [8] E. J. Gumbel, *Statistics of extremes*, Nueva York: Universidad de Columbia, 1958.
- [9] C. Cunnane, «Unbiased plotting positions - A review,» *Journal of hydrology*, 1989.
- [10] A. I. Ramírez Orozco y Á. A. Aldama, *Análisis de Frecuencias Conjunto Para la Estimación de Avenidas de Diseño*, México, 2007.
- [11] D. M. Álvarez, *Nuevos Enfoques Para la Estimación de Caudales Extremos en Colombia*, Medellín, 2007.
- [12] F. Ganancias Martínez, «Cátedra de Hidrología y Procesos Hidráulicos: Clase de Estadística Hidrológica,» Córdoba, 2009.
- [13] F. M. Ganancias Martínez, «Evaluación de Metodologías de Regionalización Hidrológica: Aplicación a los Caudales Máximos de Cuencas Representativas de la Región Sur-Oeste de la Provincia de Córdoba,» *Universidad Nacional de Córdoba*, 2010.
- [14] Universidad de Salamanca, «USAL MOOC,» 7 Noviembre 2013. [En línea]. Available: <https://www.youtube.com/watch?v=dCG3VAfa11Y>.
- [15] Universidad de Barcelona, «Universidad de Barcelona,» 2014. [En línea]. Available: http://www.ub.edu/aplica_infor/spss/cap5-4.htm. [Último acceso: 2015].
- [16] Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM), «Estudio Nacional del Agua,» Bogotá, 2010, p. 286.
- [17] D. F. Martínez Díaz, «Identificación De Modelos Estadísticos Para La Regionalización De Los Caudales Máximos De La Cuenca Magdalena-Cauca,» Bogotá, 2005.
- [18] M. J. Lencinas, «Validación de la Tormenta de Diseño Para la Presa Los Molinos - Jujuy - Utilizando Técnicas de Regionalización,» Córdoba, 2013.
- [19] P. R. Waylen y C. Caviedes, «Annual and Seasonal Fluctuations of Precipitation and Streamflow in the Aconcagua River Basin, Chile,» *Journal of Hydrology*, pp. 79-102, 1990.