



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA  
INGENIERÍA CIVIL – HIDRÁULICA

**APLICACIÓN DE FUNCIONES CÓPULA PARA EL  
ANÁLISIS DE FRECUENCIAS BIVARIADO PARA ESTIMAR AVENIDAS DE DISEÑO  
CON GASTO Y VOLUMEN PICO**

**T E S I S**  
PARA OPTAR POR EL GRADO DE:  
**MAESTRO EN INGENIERÍA**

PRESENTA  
**ING. LUIS EUSEBIO RAMÍREZ SALAZAR**

TUTOR PRINCIPAL  
**DRA. MARITZA LILIANA ARGANIS JUÁREZ** INSTITUTO DE INGENIERÍA UNAM

COMITÉ

DR. RAMÓN DOMÍNGUEZ MORA	INSTITUTO DE INGENIERÍA UNAM
DR. MOISÉS BEREZOWSKY VERDUZCO	INSTITUTO DE INGENIERÍA UNAM
DR. VÍCTOR HUGO ALCOCER YAMANAKA	FACULTAD DE INGENIERÍA UNAM
DR. JOSÉ MARÍA GONZÁLEZ-BARRIOS MURGUÍA	INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS APLICADAS Y EN SISTEMAS UNAM

CIUDAD UNIVERSITARIA, CD. MX., DICIEMBRE DE 2019.



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



## **JURADO ASIGNADO:**

Presidente: DR. RAMÓN DOMÍNGUEZ MORA  
Secretario: DR. MOISÉS BEREZOWSKY VERDUZCO  
Vocal: DRA. MARITZA LILIANA ARGANIS JUÁREZ  
1er. Suplente: DR. VÍCTOR HUGO ALCOCER YAMANAKA  
2do. Suplente: DR. JOSÉ MARÍA GONZÁLEZ-BARRIOS MURGUÍA

Lugar donde se realizó la tesis:

INSTITUTO DE INGENIERÍA UNAM, CIUDAD DE MÉXICO

TUTOR DE TESIS

---

**DRA. MARITZA LILIANA ARGANIS JUÁREZ**





*A mis hijos Jesús Cristian y Luis Ángel,  
a mis padres Luis Ramírez y María E. Salazar,  
a mi familia.*



# Agradecimientos

Esta tesis de maestría, plasma el resultado de muchos meses de trabajo, investigación y sacrificio en muchos aspectos. Durante esos momentos, la vida ha dado muchos giros, tanto en lo profesional como en lo personal, de tal forma que han sido muchas las personas que de alguna manera y en algún momento se han visto involucradas en el desarrollo de la misma, a la cuales tengo mucho que agradecer, desde ayuda en los aspectos técnicos hasta el apoyo humano, moral y de superación.

Agradezco sobretodo la orientación y apoyo técnico-administrativo de mi directora de tesis, la Dra. Maritza Liliana Arganis Juárez, quien fue la responsable y guía de la idea del tema de cópulas aplicadas a hidrología, que a partir de esta motivación he desarrollado esta investigación. Dra. Maritza, muchas gracias por la buena disposición e insistencia para el exhorto de esta travesía, que la verdad fue un gran reto a lograr.

Quiero destacar la colaboración copiosamente valiosa del Dr. José María González-Barrios Murguía, quien por sus magistrales comentarios sobre la teoría de funciones cópula y por su infinita paciencia y disposición, puede consolidar los elementos de esta teoría y llegar a ser capaz de aplicarlo al análisis hidrológico.

Asimismo destacar la colaboración del Dr. Humberto Marengo Mogollón, quien fue una persona que me animó y condujo con éxito en lo laboral y durante mis estudios de maestría, gran persona y excelente líder para el desarrollo de proyectos en ingeniería de presas, ahora jubilado de la *Comisión Federal Electricidad*.

Destacar el apoyo facilitado por parte del equipo del área técnico-administrativo de la *Comisión Nacional del Agua*; SG T y S G A A, de los cuales he aprendido a ver con otra visión la problemática del agua a lo largo y ancho de México. Gracias CONAGUA por aprender y permitirme conocer gran parte de los estados de la República Mexicana y formar parte de las decisiones para una mejor administración del agua.

No quisiera pasar por alto mencionar, las cátedras provechosas del Dr. Moisés Berezowsky en Análisis Numérico y Flujo Permanente, del Dr. Víctor Alcocer Yamanaka en modelación hidráulica de ríos y urbana, al Dr. Javier Aparicio por su interesante clase de Hidráulica Fluvial, y en general a todos mis maestros del posgrado del campus C.U. e I.M.T.A.

También un reconocimiento para mis sinodales que muy amablemente aceptaron a formar parte de este comité: Dra. Maritza Arganis, Dr. Ramón Domínguez, Dr. Moisés Berezowsky, Dr. Víctor Alcocer y al Dr. José María González-Barrios.

He de dar las gracias a mis padres principalmente, ustedes me enseñaron lo más importante; me inculcaron valores, propósito de trabajo, esfuerzo para lograr objetivos, éxito en la vida, ser líder en mis proyectos, y junto con mis hijos a tener carácter para enfrentar las adversidades.

Finalmente a mi *Alma Mater*, que son dos; a la *Universidad La Salle* quien me inculcó valores morales, así como valores de éxito, liderazgo y con gran visión para los proyectos; a la *Universidad Nacional Autónoma de México* quien me ha formado y destacado como Ingeniero Civil especializado en el sector hidráulico, gracias por su exquisita formación en matemáticas aplicadas y ciencias de la ingeniería.

# Índice general

**Resumen**

**Abstract**

**Introducción**

## **I. Antecedentes**

<b>Capítulo 1. Estudios preliminares de análisis de frecuencia bivariados de avenidas: Un enfoque con Funciones Cópula.....</b>	<b>1</b>
1.1 Estudios realizados bajo este enfoque.....	1
1.2 Desventajas de los modelos bivariados tradicionales para el análisis de frecuencia de avenidas.....	4
1.3 Ventajas de los modelos bivariados utilizando Funciones Cópulas para el análisis de frecuencia de avenidas .....	5

## **II. Estado del Conocimiento**

<b>Capítulo 2. Análisis de Frecuencia Univariado de Avenidas .....</b>	<b>7</b>
2.1 Introducción.....	7
2.2 Funciones de Distribución de Probabilidad Univariadas.....	8
2.2.1 Función de distribución de probabilidad de Valores Extremos Gumbel tipo I .....	8
2.2.2 Función de distribución de probabilidad mezclada Gumbel Dos Poblaciones .....	10
2.2.3 Función de distribución de probabilidad mezclada Gumbel Mixta (Doble Gumbel).....	12
2.2.4 Función de distribución de probabilidad Normal (Gaussiana).....	14

---

2.2.5	Función de distribución de probabilidad Log Normal con tres parámetros .....	16
2.3	Estimación de parámetros por medio del programa AX.exe.....	19
2.4	Prueba de Bondad de Ajuste para modelos de FDA univariados ....	19
2.5	Periodo de retorno univariado .....	20
2.5.1	Selección del Periodo de Retorno.....	21
<b>Capítulo 3. Teoría de Funciones Cópula .....</b>		<b>25</b>
3.1	Introducción.....	25
3.2	Conceptos y resultados básicos .....	27
3.3	El Teorema de Sklar .....	29
3.4	Cópulas y variables aleatorias.....	31
3.5	Familias de Cópulas.....	32
3.5.1	Cópulas Elípticas .....	33
3.5.2	Cópulas Arquimedianas.....	39
3.5.3	Cópulas de Valores Extremos.....	42
3.5.4	Otros modelos cópula .....	46
3.6	Medidas de Concordancia .....	50
3.6.1	Coefficiente de correlación de Pearson .....	50
3.6.2	Concepto de concordancia y discordancia.....	51
3.6.3	Medidas de correlación de rango.....	52
3.6.3.1	Tau de Kendall .....	52
3.6.3.1	Rho de Spearman .....	55
3.7	Prueba de Bondad de Ajuste para Cópulas .....	63
3.7.1	Definiciones Básicas y Resultados Preliminares.....	64
3.7.2	Aproximación Checkerboard de la cópula $C_\theta$ .....	68
3.7.3	Cópula Muestral de Orden m $C_n^{(m)}$ .....	71
3.7.4	Cópula del modelo $C_\theta$ .....	76
3.7.5	Selección del mejor modelo cópula $C_\theta$ .....	78
<b>Capítulo 4. Análisis de Frecuencia Bivariado de Avenidas: Un enfoque con Funciones Cópula.....</b>		<b>83</b>
4.1	Modelos bivariados tradicionales.....	83
4.1.1	Función de Distribución Normal Bivariada .....	84

---

4.1.2	Funciones de Distribución Bivariada de Valores Extremos .....	86
4.1.3	Función de distribución logístico bivariado de valores extremos.....	87
4.2	Ajuste de las distribuciones marginales X, Y .....	90
4.3	Elección de la Función Cópula bivariada.....	90
4.4	Periodo de retorno bivariado usando cópulas.....	90
4.4.1	Periodo de retorno condicional.....	92
4.4.2	Periodo de retorno conjunto .....	94
4.4.3	Periodo de retorno Kendall.....	98
<b>Capítulo 5. Generación de hidrogramas sintéticos: Función Gamma 2P .....</b>		<b>101</b>
5.1	Introducción.....	101
5.2	Parametrización del hidrograma a través de la Función Pearson Tipo III (sin conocer el tiempo pico) (Método 1) .....	102
5.3	Parametrización del hidrograma a través de la Función Gamma con dos parámetros (sin conocer $t_p$ ) (Método 2) .....	106
5.4	Parametrización del hidrograma a través de la Función Gamma (se conoce $t_p$ ) (Método 3).....	112
5.5	Elección del método para construir el hidrograma de avenida.....	116
 <b>III. Metodología</b>		
<b>Capítulo 6. Cálculo de Avenidas de Diseño: Un enfoque con Funciones Cópula .....</b>		<b>119</b>
6.1	Introducción.....	119
6.2	Metodología General.....	120
6.2.1	Recopilación de información hidrométrica y contexto de la presa.....	122
6.2.2	Selección de las distribuciones marginales .....	122
6.2.3	Selección de la Función Cópula Bivariada.....	124
6.2.4	Generación de pares de puntos de la distribución bivariada	



---

$(q=F^{-1}(u_1),v=F^{-1}(u_2))$ .....	127
6.2.5 Estimación de los periodos de retorno bivariados asociados a los datos observados .....	128
6.2.6 Generación de muestras bivariadas y asociadas a distintos periodos de retorno de diseño. Cálculo de volúmenes .....	129
6.2.7 Generación de curvas de periodo de retorno bivariado $(Q, V; T)$ .....	129
6.2.8 Generación de hidrogramas sintéticos .....	130
6.2.9 Simulación de Tránsitos de avenidas de los hidrogramas sintéticos .....	131
6.2.10 Construcción de la Avenida de diseño más desfavorable	132
6.3 Software disponible.....	133

#### **IV. Caso de Estudio**

##### **Capítulo 7. Aplicación del Análisis de frecuencia bivariado para estimar Avenidas de Diseño: Un enfoque con Funciones Cópula**

7.1 Introducción.....	136
7.2 Caso de Estudio .....	136
7.2.1 Recopilación de información hidrométrica .....	137
7.2.2 Cálculo de volúmenes asociados a los gastos pico.....	154
7.2.3 Selección de las distribuciones marginales .....	156
7.2.4 Selección de las Cópulas Bivariadas .....	162
7.2.5 Extensión de muestras sintéticas .....	176
7.2.6 Estimación de periodos de retorno conjuntos asociados a los datos observados .....	178
7.2.7 Generación de curvas de periodo de retorno conjunto $(Q, V; T)$ .....	185
7.2.8 Generación de hidrogramas sintéticos .....	188
7.2.9 Simulación de Tránsitos de avenidas de los hidrogramas sintéticos de la presa Huites .....	192
7.2.10 Avenidas de diseño más desfavorables en el embalse de la presa Huites.....	194

7.3 Avenidas de diseño de la presa Huites obtenidas en el libro de Ramírez y Aldama (2000).....	199
7.4 Avenidas de diseño de la presa Huites obtenidas por el IIUNAM.....	200
7.5 Comparación de resultados con las tres metodologías del análisis de frecuencias bivariado de avenidas.....	206
7.6 Conclusiones sobre los resultados obtenidos.....	209

## **V. Conclusiones y recomendaciones**

<b>Capítulo 8. Conclusiones, recomendaciones y futuras líneas de investigación.....</b>	<b>213</b>
-----------------------------------------------------------------------------------------	------------

8.1 Conclusiones.....	213
8.2 Aportaciones relevantes.....	216
8.3 Recomendaciones para próximas investigaciones.....	217
8.4 Futuras líneas de investigación.....	218

<b>Apéndice A. Coeficiente de Correlación de Spearman basado en Orden-Rango.....</b>	<b>219</b>
--------------------------------------------------------------------------------------	------------

<b>Apéndice B. Métodos gráficos para detectar dependencia.....</b>	<b>239</b>
--------------------------------------------------------------------	------------

<b>Apéndice C. Determinación de la mínima distancia supremo "D".....</b>	<b>243</b>
--------------------------------------------------------------------------	------------

<b>Apéndice D. Compendio de códigos fuente utilizados en la investigación.....</b>	<b>251</b>
------------------------------------------------------------------------------------	------------

<b>Apéndice E. Registros de Gastos medios diarios de la Estación Huites, Sinaloa.....</b>	<b>267</b>
-------------------------------------------------------------------------------------------	------------

<b>Apéndice F. Análisis de Frecuencias Univariado de Avenidas.....</b>	<b>277</b>
------------------------------------------------------------------------	------------

<b>Bibliografía.....</b>	<b>285</b>
--------------------------	------------



# Resumen

*El análisis del riesgo ante avenidas es cada vez más una necesidad y una exigencia en las legislaciones de los países desarrollados. Las variables consideradas para ello son, el gasto pico y el volumen del hidrograma. La adecuada estimación de avenidas de diseño asociadas a altos periodos de retorno es necesaria para el diseño y gestión de estructuras hidráulicas como presas. En la práctica, la estimación de estos parámetros se realiza normalmente a través de análisis de frecuencia univariados, basados en su mayoría en el estudio de caudales punta. Sin embargo, la naturaleza de las avenidas es multivariada, siendo esencial tener en cuenta características representativas de las avenidas, tales como caudal punta, volumen y duración del hidrograma, con el fin de llevar a cabo un análisis apropiado. Recientemente, el uso de cópulas se ha extendido en hidrología debido a sus beneficios en relación al contexto multivariado, permitiendo superar los inconvenientes de las técnicas tradicionales. La estimación de periodos de retorno multivariados, y por lo tanto, de cuantiles multivariados, también se facilita debido a la manera en la que las cópulas están formuladas. La presente tesis de maestría proporciona metodologías que mejoran las técnicas tradicionales usadas por profesionales para estimar cuantiles de avenida más adecuados para el diseño y la gestión de presas, así como para la evaluación del riesgo de avenida, mediante análisis de frecuencia de avenidas bivariados basados en cópulas. El procedimiento se ha verificado con los datos del embalse de la C.H. Huites, en el río Fuerte de México, con aforos de 52 años.*



# Abstract

Systematic analysis of food risk is becoming a necessity and a requirement in the legislation of developed countries. The flood variables considered for that goal are peak flow and hydrograph volume. Accurate design flood estimates associated with high return periods are necessary to design and manage hydraulic structures such as dams. In practice, the estimate of such parameters is usually done via univariate flood frequency analyses, mostly based on the study of peak flows. Nevertheless, the nature of floods is multivariate, being essential to consider representative flood characteristics, such as flood peak, hydrograph volume and hydrograph duration to carry out an appropriate analysis. Recently, the use of copulas has been extended in hydrology because of their benefits regarding dealing with the multivariate context, as they overcome the drawbacks of the traditional approach. The estimate of multivariate return periods, and multivariate parameters ( $Q_p$ ,  $V$ ), is also facilitated by the way in which copulas are formulated. The present thesis seeks to provide methodologies that improve traditional techniques used by practitioners, in order to estimate more appropriate flood quantiles for dam design, dam management and flood risk assessment, through bivariate flood frequency analyses based on the copula approach. The method is verified with data from the H.C. Huites reservoir, at the Fuerte River of Mexico, and 52 years of gauged data.



# Introducción

## Motivación

Las inundaciones se catalogan como la catástrofe más dañina de todos los desastres naturales. Según datos informados por la UNESCO (2003), Implican repercusiones socio-económicas sustanciales, causando víctimas humanas [Candela *et al.* (2014)]. Por lo tanto, el desarrollo de modelos que permiten obtener estimaciones adecuadas de la magnitud de los eventos extremos es esencial, a fin de lograr un adecuado diseño y gestión de las estructuras hidráulicas como las presas para control de avenidas o Centrales Hidroeléctricas.

Los enfoques clásicos para el estudio de análisis de frecuencia univariado o en muy pocos casos *n-variado*, han permitido hasta el momento llevar a cabo con aceptable seguridad el diseño, la construcción y el mantenimiento de una gran cantidad de presas en sus diferentes clases. Investigadores, Servidores Públicos Gubernamentales, Iniciativa Privada y Docentes, deben de tomar en cuenta que, en estos tiempos existen variaciones climáticas extremas, lo que implica una exigencia para el uso de nuevas herramientas de estadística avanzada para administrar con liderazgo la seguridad de presas en sus diferentes clases.

Fenómenos naturales complejos como: las sequías prolongadas en el Norte de México; o como las avenidas importantes que se presentan en el Sureste de México, son eventos que deben ser estudiados con carácter multivariado y sustentados por medio de un análisis estadístico riguroso, ya que se caracterizan por tener variables aleatorias correlacionadas entre sí; como por ejemplo, gasto pico, volumen de escurrimiento y duración del hidrograma de la avenida.



Cuando las avenidas extremas afectan a las presas, la relación entre el gasto pico y el nivel máximo de agua en el embalse de la presa no es sencilla, el nivel de agua en el embalse no puede deducirse directa y eficazmente del gasto pico. De modo que, es necesario contar con la caracterización del hidrograma de la avenida y por lo tanto, un enfoque multivariado es indispensable para el análisis frecuencia de avenidas. A través de los años, se ha demostrado que el enfoque univariado puede conducir a subestimar o sobreestimar el riesgo asociado a este tipo de eventos [De Michele *et al.* (2005)], y como consecuencia, aumentar los costos de construcción en las estructuras hidráulicas.

En la actualidad, es posible construir funciones de distribución multivariada con diferentes distribuciones marginales **exclusivamente** para cada variable aleatoria, la estructura de dependencia para abordar este tipo de problemas, se ha formalizado matemáticamente a través de la ***Función Cópula***.

La motivación contundente e iniciativa que se tuvo, se resume en introducir un nuevo enfoque que se ajuste a la evaluación óptima del riesgo ante avenidas, esto se logra con la incorporación de funciones cópula para el análisis de frecuencia de avenidas *n-variado*. La teoría de cópulas es considerada una herramienta de gran potencia y moderna hoy en día para análisis de frecuencia *n-variado* de avenidas.

## Objetivos de la tesis

El Objetivo de esta investigación es proponer una metodología con base en un análisis de frecuencia bivariado, que mejore las prácticas del cálculo de avenidas diseño en el embalse de una presa, tomando en cuenta la evaluación el riesgo de sobrevertido en la misma. A partir de una muestra de tamaño  $n$ , de datos observados provenientes de registros históricos medidos en una estación hidrométrica (Gastos pico o instantáneos (Q)), inferir el volumen de escurrimiento (V) a través de los gastos pico mediante la caracterización de un hidrograma de avenida, determinar una función de distribución de probabilidad bivariada que haga dependiente a las variables Q y V respectivamente, proponer diversas clases de Funciones Cópula para la modelación bivariada, y seleccionar el modelo que mejor se ajuste a las variables en estudio,

haciendo uso de una prueba de bondad de ajuste para cópulas, con el modelo seleccionado extender las muestras a 10,000 pares de puntos  $(Q, V)$  y generar las curvas de periodo de retorno conjunto, luego generar hidrogramas sintéticos a través de la Función Gamma 2P, finalmente determinar la avenida de diseño más desfavorable para un caso de estudio, tomando en cuenta el riesgo de desbordamiento en la presa.

## Importancia de la tesis

El uso de la teoría de cópulas en la ingeniería es relativamente nuevo y de rápido crecimiento. Durante los últimos 15 años, se ha presentado un gran avance en lo que se refiere al diseño de avenidas con un enfoque multivariado de cópulas, mediante la aplicación de tópicos de estadística avanzada, dichos logros radican en la implementación del concepto de función cópula adaptado al estudio de hidrología con un grado significativo de difusión [Bender *et al.* (2014)]. Para fines de esta investigación, se emplea un enfoque de cópulas para el análisis de frecuencia de avenidas únicamente para el caso bivariado.

La presente investigación bajo este enfoque, es pionera en la Facultad de Ingeniería de la UNAM, este trabajo servirá como un parámetro a seguir en diversas aplicaciones en el área hidrológica y de las Ingenierías como: Civil, Energía Renovable, Geológica, Geofísica y Petrolera, por mencionar algunas. El enfoque mediante el uso de funciones cópula para el análisis de frecuencia *n-variado* de avenidas promete ser de gran envergadura; para el cálculo de cuantiles, extensión de muestras sintéticas y la estimación del periodo de retorno *2-variado*.

## Organización de la tesis

En el primer capítulo, se comentan los estudios realizados bajo la teoría de Funciones Cópula aplicado al análisis de frecuencia bivariado de avenidas, se enuncian algunas ventajas y desventajas sobre los modelos bivariados tradicionales con los modelos bajo un enfoque a Funciones Cópula, para el tratamiento de análisis de frecuencia bivariado.

En lo sucesivo se identificará al acrónimo FDA como la Función de distribución acumultiva.

Lo que respecta al Capítulo 2, se expone un resumen con los resultados básicos referentes a las funciones de distribución de probabilidad univariadas más utilizadas en los análisis de frecuencia univariado de avenidas, se hace referencia de la forma de estimar los parámetros estadísticos por medio del programa AX y los criterios para la evaluación de la prueba de bondad de ajuste para modelos de FDA univariados, se aborda el tema de periodo de retorno univariado y la forma en como es selección el periodo de retorno para diferentes obras hidráulicas.

Referente al Capítulo 3, se introduce la teoría de cópulas de manera formal, se hace referencia sobre las principales definiciones y teoremas que sustentan esta teoría, se incluyen cuatro clases de familias cópula; **Elípticas** (Gaussiana y t-Student), **Arquimedianas** (Clayton, Frank, Ali-Mikhail-Haq, Gumbel-Hougaard y Joe/B5), **Valores Extremos** [Gumbel-Hougaard, Galambos (modelo logístico negativo), Hüsler-Reiss, t-VE (t-Student), Joe/BB5 y Tawn], y **Otros modelos** como Plackett y Farlie-Gumbel-Morgenstern. Se describen las medidas de concordancia más importantes como la  $\rho_s$  de Spearman y la  $\tau_\theta$  de Kendall, se muestra la Tabla 3.4, la cual contiene las relaciones entre las medidas citadas y el parámetro de asociación  $\theta$ . En lo que respecta a la prueba de bondad de ajuste para cópulas, se aplica la metodología **Cópula Muestral de Orden  $m$** , propuesta por el Dr. José María González-Barrios Murguía, investigador del Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS-UNAM).

El Capítulo 4, comprende el estudio sobre el análisis de frecuencia bivariado de avenidas con enfoque a Funciones Cópula, se hace mención de algunas funciones de distribución de probabilidad bivariada (modelos tradicionales bivariados) como la normal bivariada en su forma extendida y para el caso cuando el vector de medias es cero y con matriz de correlación  $\Sigma$  (coincide con la cópula gaussiana), el modelo logístico bivariado de Gumbel (modelo positivo) y se hace la observación de que este modelo es similar al de la cópula de Gumbel-Hougaard, una gran diferencia es el parámetro de asociación. Se señala la diferencia de entre el periodo de retorno univariado y los aspectos fundamentales que involucran la ocurrencia simultanea de eventos, mejor conocido como periodo de retorno bivariado, esta probabilidad de ocurrencia resulta cuando, en el diseño o revisión de una presa de determinada clase, la

descarga máxima que escurre sobre el vertedor, no solo depende de la variable gasto pico de entrada a la presa, sino también de las variables de volumen de escurrimiento, duración de la avenida y la forma que caracteriza el hidrograma de entrada. De modo que, el análisis de frecuencia bivariado para dos variables aleatorias correlacionadas entre sí, están definidas por un período de retorno conjunto, se analizan tres metodologías.

En el Capítulo 5, se desarrolla el tema referente a la Generación y parametrización de hidrograma sintéticos, se plantean tres métodos muy versátiles, básicamente el tema se centra en la construcción de hidrogramas a través de la Función Gamma de dos parámetros (2P) o también llamada Función Pearson Tipo III de dos parámetros.

Acerca del Capítulo 6, aquí se pone en práctica la teoría vista en los capítulos anteriores, de modo que se presenta la metodología general para la construcción de avenidas de diseño haciendo uso de la teoría de cópulas para el análisis de frecuencias bivariado, además de señala el empleo de dos softwares mediante los cuales se realizaron los cálculos y gráficos para el estudio.

El Capítulo 7, comprende los hallazgos y resultados sobresalientes del caso de estudio, conforma el logro de los propósitos estipulados en los objetivos de esta tesis. Se muestran las características principales de la cuenca en donde se ubican las estaciones hidrométricas, las funciones de distribución de probabilidad univariadas que se probaron y las distribuciones marginales que mejor se ajustaron a los datos observados, se muestra una tabla con los modelos cópula que se probaron y conforme a una prueba de bondad de ajuste para cópulas se elige el mejor modelo, se estima la avenida de diseño más desfavorable, en función de la máxima elevación en el embalse de la presa por conducto de una simulación del tránsito de la avenida. Se muestran gráficos con las curvas de periodo de retorno conjunto.

Finalmente el cierre esta investigación se basa en las conclusiones y recomendaciones para futuras líneas de investigación en la teoría de cópulas aplicado a análisis de frecuencias *n-varidos*.



# Capítulo 1

## Estudios preliminares de análisis de frecuencia bivariados de avenidas: Un enfoque con Funciones Cópula

### 1.1. Estudios realizados bajo este enfoque

Históricamente investigadores como Hiemstra *et al.* (1976) realizaron un trabajo pionero en análisis de frecuencia; Singh (1987) muestra una revisión exhaustiva de este primer análisis. En la mayoría de las situaciones, tradicionalmente el análisis del período de retorno involucra casos univariados; desafortunadamente, esto puede llevar a una sobre/subestimación del riesgo asociado con un evento dado [ver, por ejemplo, Raynal-Villasenor y Salas (1987) y De Michele *et al.* (2004)]. De hecho, los eventos hidrológicos a menudo se caracterizan por presentar un comportamiento conjunto de varias variables aleatorias, y estos generalmente no son independientes: por mencionar algunos ejemplos, gasto pico, volumen de la avenida, duración de la avenida, intensidad de la tormenta [De Michele *et al.* (2004)] y eventos de sequía.

Otros trabajos significativos sobre la aplicación de distribuciones de probabilidad conjunta para modelar eventos hidrológicos extremos son de Yue *et al.* (1999, 2000 y 2001). Como consecuencia, los eventos relevantes deben definirse mejor en términos de dos o más variables.

Autores en la época más reciente como Salvadori *et al.* 2004; Favre *et al.* (2004); Grimaldi y Serinaldi (2006, 2007) *destacaron la necesidad de aplicar nuevos enfoques como son las funciones cópula al análisis frecuencia de avenidas multivariadas*. Favre *et al.* (2004) identificaron la cópula Frank como la mejor para el modelado de avenidas pico y el volumen. Zhang y Singh (2006) encontraron que la cópula Gumbel es la mejor representación para la relación de gasto pico y volumen. En este estudio, se analizaron las relaciones entre el pico, el volumen y duración del hidrograma. Se utilizaron cópulas bivariadas para caracterizar las relaciones de pico-volumen y volumen-duración. Grimaldi y Serinaldi (2006) aplicaron cópulas Arquimedianas asimétricas (cópulas arquimedianas anidadas) para abordar el análisis de frecuencia de avenidas trivariadas mediante el estudio de avenidas pico, el volumen y la duración. Genest y Favre (2007) también estudiaron la relación entre el gasto pico y el volumen. Consideraron alrededor de 20 tipos de cópulas, el uso de cópulas para obtener distribuciones multivariadas supera los inconvenientes del enfoque tradicional. Seleccionaron cinco familias de cópulas, de las cuales cuatro pertenecen a la clase de valores extremos (Gumbel-Hougaard, Galambos, Hüsler-Reiss y BB5/Joe). La quinta (BB1/Joe) pertenece a la clase de cópulas arquimedianas de dos parámetros. La mejor familia de cópula fue la BB5 de Joe. Genest *et al.* (2007) aplicaron cópulas meta-elípticas debido a su flexibilidad para el modelado de gastos pico, volumen y la duración. Justificaron lo anterior al obtener un resultando de dependencia negativa entre la relación pico-duración (en este caso las cópulas arquimedianas asimétricas no pueden ser aplicadas), además se obtuvo un menor grado de dependencia para la relación de volumen-duración.

Zhang y Singh (2007) utilizaron la cópula Gumbel trivariada para el modelado de avenidas pico, volumen y duración. En este caso de estudio se encontró una dependencia negativa para la relación pico-duración y una relación positiva entre el resto de las variables.

En la década de los 2010, Salvadori *et al.* (2011) analizaron un caso de estudio trivariado donde se consideró el pico de la avenida, el volumen y el nivel inicial ( $z$ ) de agua en una presa (antes de la avenida). A medida que el nivel inicial era independiente del pico y el volumen, se empleó el uso de cópulas arquimedianas anidadas considerando la Cópula de Gumbel bivariada para caracterizar la relación entre el pico-volumen, y la distribución marginal del nivel inicial. Chowdhary *et al.* (2011) trabajaron con la cópula Clayton, la cual fue seleccionada como la más apta debido a las características de su estudio. Ganguli y Reddy (2013) compararon el comportamiento de tres cópulas arquimedianas asimétricas (Clayton, Frank y Gumbel) y una cópula meta-elíptica (t de Student). El caso de estudio mostró una dependencia positiva de menor grado en la relación pico-duración, mientras que en las otras dos resultaron dependencias positivas y similares entre sí, para las relaciones pico-volumen y duración-volumen. La cópula meta-elíptica se mostró flexible y fue elegida como la mejor cópula para modelar las variables, dichos resultados fueron sustentados bajo la aplicación de diversas pruebas estadísticas.

Sin embargo, Brechmann y Schepsmeier (2013) señalaron la falta de flexibilidad de las cópulas elípticas y arquimedianas para el modelado de la estructura de dependencia multivariada, por lo que propusieron el uso de Cópulas Vine. Las cópulas Vine fueron introducidos por Joe (1996) y estudiadas por varios autores [Aas *et al.* (2009)]. Las cópulas Vine son modelos flexibles capaces de representar complejas estructuras de dependencia multivariadas para altas dimensiones, se construyen usando cópulas bivariadas (2-cópulas). Esta familia de cópulas debe de estudiarse con detenimiento en futuras investigaciones en hidrología.

El uso de cópulas para obtener distribuciones multivariadas supera los inconvenientes del enfoque tradicional. Una cópula es una función de distribución de  $m$  variables con distribuciones marginales distribuidas uniformemente en  $[0,1]^m$  que representa la estructura de dependencia entre variables correlacionadas, y permite la obtención de una función de distribución multivariada a través de las funciones de distribución univariadas.



Conceptos generales sobre análisis de frecuencia con enfoque a cópulas se incluyen en Salvadori *et al.* (2007), las siguientes dos publicaciones se consideran hasta la fecha como las más contundentes para diversas aplicaciones en hidrología, [Zhang y Singh (2019)] y [Chen y Guo (2019)].

## 1.2. Desventajas de los modelos bivariados tradicionales para el análisis de frecuencia de avenidas

Los modelos *n-variados* tradicionales para administrar la seguridad hidrológica de presas, marcó un nuevo rumbo para el análisis de frecuencia de avenidas. La mayoría de los modelos bivariados tradicionales utilizan la distribución normal bivariada para describir la correlación de fenómenos que involucran variables aleatorias como gasto pico, volumen y duración [Favre *et al.* (2004)].

Una distribución normal bivariada requiere de una transformación previa de la serie de datos (normalización), ya que *las variables aleatorias que intervienen en los datos observados habitualmente no son normales*, conforme a ello, los resultados que se obtendrían estarían fuera de la realidad. Sin embargo, otro inconveniente de estos modelos, es que tras llevar a cabo la transformación (normalización) de los datos, no se puede asegurar que el resultado siga una distribución normal.

En lo que respecta al modelo logístico bivariado de Gumbel (modelo positivo), aunque ofrece aceptables resultados, la desventaja radica en que únicamente pueden ser empleadas distribuciones marginales de valores extremos (Gumbel, Exponencial, GVE, Fréchet y Weibull), el empleo de esta distribución bivariada se acota únicamente a un caso particular.

### ▪ Limitaciones de modelos bivariados tradicionales:

- 1) Las variables aleatorias que caracterizan un hidrograma de avenida como gasto pico, volumen y duración, tienen que seguir una misma función de distribución marginal. Esto no sucede así en la realidad, ya que las variables anteriores pueden tener un comportamiento diferente.

- 2) Las variables anteriores son implementadas bajo la suposición de que siguen una distribución normal, y conforme a la naturaleza de las mismas, éstas no siguen una distribución normal, a menos que se haga una transformación (normalización) previa de los datos observados.

### **1.3. Ventajas de los modelos bivariados utilizando Funciones Cópulas para el análisis de frecuencia de avenidas**

Los inconvenientes que presentan los modelos bivariados tradicionales, han sido solucionados bajo el estudio de funciones cópula para el análisis de frecuencias de avenidas, las cuales ofrecen una amplia gama de modelos y el empleo de diversas distribuciones marginales que previamente han sido ajustadas conforme a la naturaleza de los datos observados.

La modelización de la distribución bivariada con el uso de cópulas, disminuye en gran medida la limitación que se tenía anteriormente para el desarrollo de análisis de frecuencia tradicional, en el sentido de que, existía cierto grado de restricción en las variables aleatorias que integraban la distribución bivariada, debido a la exigencia de hacer normales dichas variables, así como la necesidad de utilizar las mismas marginales para todas las variables que intervenían [Salvadori *et al.* (2007)]. Por medio de esta teoría, es factible implementar distribuciones marginales más adecuadas, y no solo restringirse al caso gaussiano o de valores extremos.

Las expresiones matemáticas (Funciones Cópula) son fácilmente manejables, por lo que la aportación de las cópulas ha venido a revolucionar un avance destacado en la hidrología de presas.

Las cópulas se utilizan ampliamente para investigar diversas áreas como: finanzas, análisis de riesgos, hidrología, tormentas, sequías, petrofísica, contaminantes y análisis de supervivencia [Requena (2015)].



# Capítulo 2

## Análisis de Frecuencia Univariado de Avenidas

### 2.1. Introducción

El análisis de frecuencia de avenidas se clasifica en univariados o multivariados, dependiendo del número de variables a estudiar; pueden ser de tipo local o regional. El análisis de frecuencia, considera para su desarrollo el tratamiento de Funciones de Distribuciones de Probabilidad (o Función de Distribución Acumulativa (FDA)), las cuales se determinan mediante el ajuste empírico a los datos. El ajuste requiere de la estimación de los parámetros estadísticos de la distribución en turno y la evaluación de la bondad del ajuste para la elección del mejor modelo. El objetivo de este capítulo es discutir brevemente estos elementos básicos, que son necesarios para el análisis de frecuencia.

En el pasado, los investigadores se centraron en el análisis de frecuencia de avenidas univariadas, generalmente en cuanto al pico de avenida [Kite (1988) y Stedinger *et al.* (1993)]. El enfoque univariado comúnmente se utilizó debido a que requiere un análisis matemático menos elaborado que el enfoque multivariado para la descripción de las variables hidrológicas correlacionadas [Shiau (2003)]. Sin embargo, *el modelo univariado no es capaz de caracterizar completamente la probabilidad de ocurrencia de la avenida* [Chebana y Ouarda (2011)].

A continuación se comentarán las funciones de distribución de probabilidad más comunes para el análisis de frecuencias univariado, que a través del ejercicio profesional se emplean con regularidad en proyectos de ingeniería acorde al sector hídrico.

## 2.2. Funciones de Distribución de Probabilidad Univariadas

Entre las distribuciones de probabilidad univariadas continuas más comúnmente aplicadas en análisis de frecuencia hidrológica son, exponencial, de valores extremos tipo I [Gumbel (1958)], Gumbel Mixta (o Doble Gumbel) [Raynal y Guevara (1997)], Gumbel de dos Poblaciones [González (1970)], Normal y Log-Normal (2 y 3 parámetros).

Se denotará a  $X$  como una Variable Aleatoria con Función Densidad de Probabilidad (FDP)  $f(x)$  y su respectiva Función de Distribución Acumulativa (FDA)  $F(x)$ .

### 2.2.1. Función de distribución de probabilidad de Valores Extremos Gumbel tipo I

La distribución de valores extremos tipo I de Gumbel se expresa como

$$F(x) = e^{-e^{-\left(\frac{x-\beta}{\alpha}\right)}}, \quad (2.1)$$

y a su correspondiente función densidad de probabilidad es [NERC (1975)]

$$f(x) = \frac{1}{\alpha} \cdot e^{-e^{-\left(\frac{x-\beta}{\alpha}\right)}} \cdot e^{-\left(\frac{x-\beta}{\alpha}\right)}, \quad (2.2)$$

para todo  $-\infty < x < \infty$ ,  $\alpha > 0$ ,

donde  $\alpha$  denota el parámetro de escala y  $\beta$  es el parámetro de ubicación.

- **Gráfico en papel tipo Gumbel**

Para dibujar a la probabilidad de valores extremos (eje vertical) de una FDA Gumbel T1, se emplea la *variable reducida* ( $z$ ) en el eje horizontal, la cual se define como:

$$z = -\ln \left[ \ln \left( \frac{Tr}{Tr - 1} \right) \right], \quad (2.3)$$

de donde  $Tr$  (*periodo de retorno*) resulta

$$Tr = \frac{e^{1/e^z}}{e^{1/e^z} - 1}, \quad (2.4)$$

la *probabilidad de no excedencia* definida a partir del periodo de retorno  $Tr$  es

$$P(X \leq x) = 1 - \frac{1}{Tr}. \quad (2.5)$$

Observación: Cuando en la expresión (2.3),  $z = 0$  (*moda*), ocurre una probabilidad  $P = 0.3678794412$  y un periodo de retorno  $Tr = 1.5819767069$ ; la media se ubica en  $z = 0.5986749645$ , con probabilidad  $P = 0.5772156649$  y  $Tr = 2.3652721186$ , la Figura 2.1 muestra el bosquejo de la FDA Gumbel, en papel tipo Gumbel.

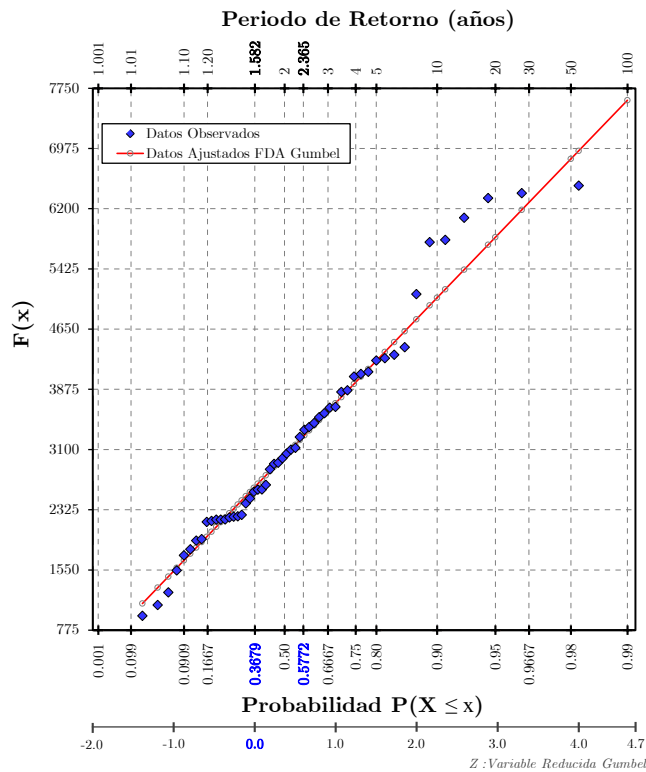


Figura 2.1. FDA Gumbel referenciada en papel Gumbel

## 2.2.2. Función de distribución de probabilidad mezclada Gumbel Dos Poblaciones

Conforme a las investigaciones realizadas por González (1970), el autor en su momento argumentó que, en México el principal modelo para la construcción de funciones de probabilidad para poblaciones mezcladas, estaba dado para el caso de dos poblaciones por la siguiente relación:

$$F(x) = F_1(x) \{ p + [(1-p)F_2(x)] \}, \quad (2.6)$$

donde

$F_1(x)$  FDA asociada con la población 1,

$F_2(x)$  FDA asociada con la población 2,

$F(x)$  FDA de la mezcla de dos poblaciones  $F_1(x)$  y  $F_2(x)$ .

$p$  probabilidad de que la variable pertenezca a la población 1

González (1970), estableció que las dos poblaciones siguen una función de distribución de probabilidad de valores extremos Gumbel tipo I (2.1), resultando el siguiente modelo (en la Figura 2.2 se bosqueja su gráfica):

$$F(x) = e^{-e^{-\left(\frac{x-\beta_1}{\alpha_1}\right)}} \left\{ p + \left[ (1-p) e^{-e^{-\left(\frac{x-\beta_2}{\alpha_2}\right)}} \right] \right\}, \quad (2.7)$$

y su respectiva función densidad de probabilidad es [Ramírez (2011)]

$$f(x) = e^{-e^{-\left(\frac{x-\beta_1}{\alpha_1}\right)}} \left\{ \frac{p}{\alpha_1} e^{-\left(\frac{x-\beta_1}{\alpha_1}\right)} + \frac{(1-p)}{\alpha_1 \alpha_2} e^{-e^{-\left(\frac{x-\beta_2}{\alpha_2}\right)}} \left[ \alpha_2 e^{-\left(\frac{x-\beta_1}{\alpha_1}\right)} + \alpha_1 e^{-\left(\frac{x-\beta_2}{\alpha_2}\right)} \right] \right\}, \quad (2.8)$$

para toda  $x > 0$ ,  $\alpha_1 = \alpha_2 > 0$ ,  $p \in (0,1)$ ,

donde

$x$  variable aleatoria, en este caso es el gasto máximo, en  $(m^3/s)$ ,

$F(x)$  FDA

$f(x)$  FDP

- $\alpha_1$  parámetro de escala asociado a la población 1 (no ciclónico),
- $\beta_2$  parámetro de ubicación asociado a la población 1 (no ciclónico),
- $\alpha_2$  parámetro de escala asociado a la población 2 (ciclónico),
- $\beta_1$  parámetro de ubicación asociado a la población 2 (ciclónico),
- $p$  probabilidad de que se presenten eventos no ciclónicos (población 1).

Respecto a la expresión (2.7), una importantísima cita enunciada en el libro de Ramírez y Aldama (2000), toma en cuenta la siguiente aclaración:

[...] Con la formulación (2.6), González intento tomar en cuenta la posibilidad de haber considerado un evento como perteneciente a la población 1 cuando realmente corresponde a la población 2 [...] Es evidente que para el caso  $p = 1$ , las dos formulaciones coinciden, pero a medida que el grado de segregación del fenómeno aumenta, las formulaciones pueden arrojar resultados muy distintos. La distribución mezclada propuesta por González presenta una clara inconsistencia, ya que para el caso en que  $p = 0$ , es decir, que toda la muestra provenga de la población 2, se tiene que  $F(x) = F_1(x) \cdot F_2(x)$ , lo cual carece de sentido. Al introducir el concepto de periodo de retorno en la expresión (2.7) [...] la formulación propuesta por González Villareal genera gastos mayores que la correspondiente a eventos excluyentes. Para fines de diseño, esta sobrestimación parece no ser un problema grave, sin embargo, la utilización de la distribución Gumbel propuesta por González puede conducir a un sobrediseño sistemático de las obras. [...] la construcción de funciones de distribución multivariadas de valores extremos con marginales de dos poblaciones no se encuentra completamente documentada [...]

Ramírez, L. (2011, pp. 87, 110 y 257) ajustó una muestra de Gastos máximos anuales, utilizando los registros históricos de la estación hidrométrica Angostura, ubicada en el Estado de Chiapas, México, en la cual se identificaron dos poblaciones, el ajuste se realizó utilizando la FDA Gumbel Dos Poblaciones (expresión 2.7 de este trabajo), conforme a los resultados que presentó y comparó con los reportados por Domínguez *et al.* (2006), llegó a la conclusión de que para la estimación de eventos de diseño (cuantiles), los gastos máximos asociados a periodos de retorno mayores a 5000 años, se sobreestiman o subestiman



notablemente en relación a la comparación de resultados. La FDA Gumbel Dos Poblaciones es otra alternativa para considerar un mejor ajuste a los datos observados, es importante tomarla en cuenta para diversos análisis de frecuencias d-variados, así como en los análisis de frecuencia con enfoque a cópulas.

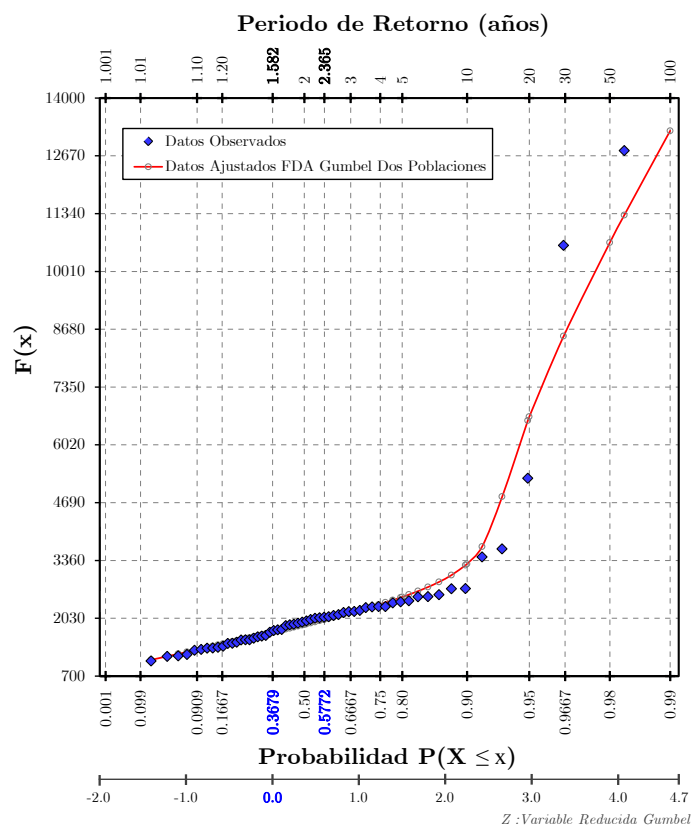


Figura 2.2. Ajuste de datos mediante la FDA Gumbel Dos Poblaciones.

### 2.2.3. Función de distribución de probabilidad mezclada Gumbel Mixta (Doble Gumbel)

El uso de Funciones de distribución de probabilidad univariadas bajo el concepto de mezcla de distribuciones, se convirtió en una herramienta ampliamente eficiente para solucionar problemas, cuya modelación considera el tratamiento una muestra proveniente de dos poblaciones. Mood *et al.* (1974, p. 64) señala la existencia de una distribución de probabilidad mezclada, contemplando dos poblaciones, definiéndola como:

$$F(x) = p F_1(x) + (1 - p) F_2(x), \quad (2.9)$$

donde  $p$  es la proporción de la segunda población en la mezcla,  $F_1(x)$  y  $F_2(x)$  son las componentes de la mezcla.

Recientemente, se han propuesto algunas distribuciones de probabilidad de valores extremos Gumbel tipo I para dos poblaciones, bajo la idea de (2.9) [Raynal y Guevara (1997)], en este sentido, si  $F_1(x)$  y  $F_2(x)$  siguen una distribución Gumbel tipo I, entonces se conforma un modelo conocido como FDA Gumbel Mixta o Doble Gumbel, y se define como:

$$F(x) = p e^{-e^{-\left(\frac{x-\beta_1}{\alpha_1}\right)}} + (1 - p) e^{-e^{-\left(\frac{x-\beta_2}{\alpha_2}\right)}}, \quad (2.10)$$

donde  $\alpha_1$ ,  $\beta_1$  y  $\alpha_2$ ,  $\beta_2$  son los parámetros estadísticos de escala y ubicación de la primera y segunda población, respectivamente.

La correspondiente función densidad de probabilidad FDP es

$$f(x) = \frac{p}{\alpha_1} e^{-e^{-\left(\frac{x-\beta_1}{\alpha_1}\right)}} e^{-\left(\frac{x-\beta_1}{\alpha_1}\right)} + \frac{(1-p)}{\alpha_2} e^{-e^{-\left(\frac{x-\beta_2}{\alpha_2}\right)}} e^{-\left(\frac{x-\beta_2}{\alpha_2}\right)}. \quad (2.11)$$

El modelo (2.10) ofrece excelentes resultados derivado de un análisis de frecuencias univariado, ha sido aplicado por diversos investigadores, compañías, consultorías y en la estructura gubernamental en México.

Más adelante se empleará este modelo para trabajar con funciones cópula. A continuación se presenta en la Figura 2.3, la configuración del modelo (2.10), utilizando datos de la estación hidrométrica Huites (gastos máximos), ubicada en el Estado de Sinaloa, México.

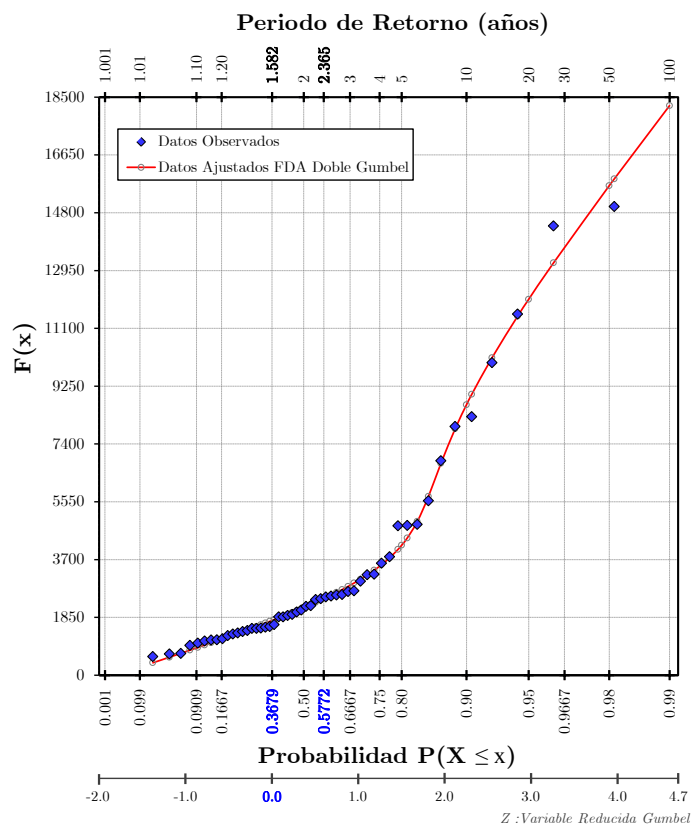


Figura 2.3. Ajuste de datos mediante la FDA Doble Gumbel.

## 2.2.4. Función de distribución de probabilidad Normal (Gaussiana)

La función de distribución de probabilidad Normal es simétrica respecto a  $\mu$  (media), las correspondientes FDA y FDP, están dadas por las siguientes expresiones:

$$F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx, \quad (2.12)$$

para toda  $x \in (-\infty, \infty)$ ,  
y su correspondiente FDP

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad (2.13)$$

donde

$\sigma$  y  $\mu$  son los parámetros de escala y ubicación.

▪ **Función de Distribución de Probabilidad Normal Estándar**

La expresión (2.12) es evidente que se evalúa numéricamente, sin embargo, una simple transformación de variables, es posible lograr que la evaluación sea independiente de  $\sigma$  y  $\mu$ , esto es,

sea la variable normal estándar definida como

$$z = \frac{x - \mu}{\sigma} \quad (2.14)$$

de donde

$$F(x) = P\left(Z \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right), \quad \mu \in \mathbb{R}, \sigma > 0, \quad (2.15)$$

y resulta la distribución normal estándar

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{u^2}{2}} du; \text{ para toda } z \in (-\infty, \infty), \quad (2.16)$$

donde  $\Phi(z)$  representa la función de distribución de probabilidad normal estándar, con media cero y varianza uno; esto es  $Z \sim N(0,1)$ .

La correspondiente función densidad de probabilidad, resulta

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}; \text{ para toda } z \in (-\infty, \infty). \quad (2.17)$$

Soranzo y Epure (2014) desarrollaron una aproximación numérica de la función de distribución de probabilidad normal estándar  $\Phi(z) = P(Z \leq z)$ , expresión (2.16), y su inversa  $\Phi^{-1}(P(Z \leq z)) = z$ , con error absoluto menor a  $1.28 E - 04$ , los resultados fueron los siguientes:

la aproximación de la distribución normal estándar es

$$\Phi(z) \simeq 2^{(-22)} \left[ 1 - 41^{(z/10)} \right], \text{ para toda } z \geq 0. \quad (2.18)$$

y la aproximación de la inversa de  $\Phi(z)$ , es el cuantil

$$\Phi^{-1}(P(Z \leq z)) = z \approx \frac{10}{\ln 41} \cdot \ln \left\{ 1 - \frac{\ln \left[ \frac{-\ln \Phi(z)}{\ln 2} \right]}{\ln 22} \right\}, \quad (2.19)$$

para toda  $\Phi(z) \in \left[ \frac{1}{2}, 1 \right)$ .

Observación: La aproximación de  $\Phi(z) = P(Z \leq z)$  cuando  $z < 0$  y su respectiva inversa  $\Phi^{-1}(P(Z \leq z)) = z$  cuando  $\Phi(z) \in \left[ 0, \frac{1}{2} \right]$ , se tienen las siguientes propiedades:

$$\Phi(-z) = 1 - \Phi(z), \quad (2.20)$$

$$\Phi^{-1}(P(Z \leq z) \leq 1/2) = -\Phi^{-1}(P(Z \leq z) \leq 1/2). \quad (2.21)$$

Abramowitz y Stegun (1965) propusieron una aproximación numérica de  $\Phi(z)$  con error absoluto menor a  $7.5 E - 05$ , como sigue:

$$\Phi(z) \approx 1 - \phi(z) \left[ a_1 z + a_2 z^2 + a_3 z^3 + a_4 z^4 + a_5 z^5 \right] + \varepsilon(z), \quad (2.22)$$

donde

$$a_1 = 0.319381530, a_2 = 0.356563782, a_3 = 1.781477937, a_4 = 1.821255978,$$

$$a_5 = 1.330274429, \varepsilon(z) \text{ es el error de aproximación y } \phi(z) \text{ resulta de (2.17).}$$

### 2.2.5. Función de distribución de probabilidad Log Normal con tres parámetros

La FDA Log Normal se define como:

$$F(x) = \frac{1}{\alpha \sqrt{2\pi}} \int_{\delta}^x \frac{1}{(x-\delta)} \cdot e^{-\frac{1}{2} \left[ \frac{\ln(x-\delta) - \beta}{\alpha} \right]^2} dx, \quad (2.23)$$

para toda  $x \in (\delta, +\infty)$ ,

donde

$\alpha$  es el parámetro de escala,

$\delta$  es el parámetro de ubicación,

$\beta$  es el parámetro de forma.

y su correspondiente función densidad de probabilidad es

$$f(x) = \frac{1}{(x - \delta) \alpha \sqrt{2\pi}} \cdot e^{-\frac{1}{2} \left[ \frac{\ln(x - \delta) - \beta}{\alpha} \right]^2} \quad (2.24)$$

Cuando se considera a  $\delta = 0$ , en la ecuación (2.23), se dice que se trata de una función de distribución de probabilidad Log Normal con dos parámetros.

▪ **Estimación de estadísticos muestrales**

La media muestral se define como

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (2.25)$$

Varianza insesgada

$$S_{insesg}^2 = \frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2, \quad (2.26)$$

Coefficiente de asimetría insesgado

$$g_{insesg} = \frac{n}{(n-1)(n-2)} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{[S_{insesg}^2]^{3/2}}, \quad (2.27)$$

Coefficiente de curtosis insesgado

$$k_{insesg} = \frac{n^2}{(n-1)(n-2)(n-3)} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{[S_{insesg}^2]^2}. \quad (2.28)$$

Escalante y Reyes (2005) puntualizan la recomendación de usar estadísticos muestrales insesgados, ya que generalmente se trabaja con muestras relativamente pequeñas.

- **Estimación de parámetros estadísticos por el método de momentos**

Parámetro de ubicación

$$\delta = \bar{x} - \frac{S_{insseg}}{z}, \quad (2.29)$$

y toma en cuenta los siguientes parámetros

$$\omega = \frac{\sqrt{g_{insseg}^2 + 4} - g_{insseg}}{2} \quad (2.30)$$

$$z = \frac{1 - \omega^{2/3}}{\omega^{1/3}} \quad (2.31)$$

$$\alpha = \sqrt{\ln(z^2 + 1)} \quad (2.32)$$

$$\beta = \ln\left(\frac{S_{insseg}}{z}\right) - \frac{1}{2} \ln(z^2 + 1) \quad (2.33)$$

- **Estimación de parámetros estadísticos utilizando el método de máxima verosimilitud**

El logaritmo natural de la función de verosimilitud para resolver la distribución de probabilidad Log Normal incluyendo a tres parámetros, se cita en Kite (1978), y la define como:

$$\ln L = -\frac{n}{2} \ln(2\pi) - n \ln(\alpha) - \sum_{i=1}^n (x_i - \delta) - \frac{1}{2\alpha^2} \sum_{i=1}^n [\ln(x_i - \delta) - \beta]^2, \quad (2.34)$$

al derivar la expresión (2.34) con respecto a  $\alpha$ ,  $\beta$  y  $\delta$ , e igualando a cero cada una de las respectivas derivadas, resulta lo siguiente

$$\alpha^2 = \sum_{i=1}^n \frac{[\ln(x_i - \delta) - \beta]^2}{n}, \quad (2.35)$$

$$\beta = \sum_{i=1}^n \frac{\ln(x_i - \delta)}{n}, \quad (2.36)$$

$$\sum_{i=1}^n \frac{\beta - \alpha^2}{x_i - \delta} - \sum_{i=1}^n \frac{\text{Ln}(x_i - \delta)}{x_i - \delta} = 0, \quad (2.37)$$

al evaluar las expresiones (2.35) y (2.36) en la expresión (2.37), conduce a una ecuación no lineal en términos únicamente de  $\delta$ , y conforme a las recomendaciones de Berezowsky *et al.* (1983) es factible resolverla empleando el método de la secante. Otra alternativa es utilizar algoritmos genéticos [Fuentes *et al.* (2005)].

### 2.3. Estimación de parámetros por medio del programa **AX.exe**

Dos métodos de estimación de parámetros se consideran: el método de máxima verosimilitud y el método de momentos [Laio (2004); Rao y Hamed (1999)]. En general, se aplica el método de momentos cuando se consideran muestras de longitud corta, debido a sus buenos resultados [Hosking y Wallis (1997)]; mientras que el método de máxima verosimilitud se utiliza para otros casos. En esta investigación, la estimación de parámetros y en general el ajuste de funciones de distribución de probabilidad (FDA) univariadas, se utiliza el programa **AX.exe**, este programa cuenta con una paquetería de hasta seis FDP [Jiménez (1997)]: Normal, LogNormal 2p y 3P, Gamma 2p y 3P, Exponencial, Gumbel y Doble Gumbel 5p. Una forma resumida del manejo de este programa se puede consultar en Ramírez, L. (2011).

### 2.4. Prueba de Bondad de Ajuste para modelos de FDA univariados

Las pruebas de bondad de ajuste se aplican para identificar formalmente a las distribuciones apropiadas para el ajuste de los datos, estas son la prueba Kolgomorov-Smirnov [Klein et al. (2010)] y la prueba de Anderson-Darling transformado [Laio (2004)]. Otro método de gran aceptación en el argot hidrológico es el Error Estándar de Ajuste (**EEA**), y tiene el propósito de comparar la eficiencia del ajuste realizado a la muestra con otros modelos de FDA [Kite (1988)]. El EEA se obtiene con la siguiente expresión:



$$EEA = \left[ \frac{\sum_{i=1}^n (Q_{observado} - Q_{calculado})^2}{n - np} \right]^{1/2}, \quad (2.38)$$

donde

$n$  es el tamaño de la muestra,

$np$  es el número de parámetros de la distribución ajustada.

El criterio para la selección del mejor modelo univariado, es aquel que posea el valor mínimo de EEA respectivamente, esta comparación se realiza para todas las FDA en estudio.

## 2.5. Periodo de retorno univariado

Para el diseño de estructuras hidráulicas como la obra de excedencia, canal de llamada, obra de desvío, obra de contención, y demás aditamentos que conforman un arreglo en una clase de presa, se requieren obtener estimaciones de la magnitud de las avenidas ocurridas en determinada cuenca y con cierta probabilidad de ocurrencia o mejor conocido como período de retorno.

El periodo de retorno se define como el tiempo promedio transcurrido entre ocurrencias de un evento de magnitud dado o uno de magnitud mayor (Haan, 1986). El período de retorno proporciona un medio simple y eficiente para la evaluación de riesgos, porque concentra una gran cantidad de información en un solo número.

Los eventos más probables tienen períodos de retorno más cortos, mientras que los eventos menos probables tienen períodos de retorno más largos. En la práctica de la ingeniería, la elección del período de retorno depende de la importancia de uso que tendrá la estructura hidráulica en cuestión y las consecuencias de su falla [Favre *et al.* (2016)].

El ajuste de distribuciones de probabilidad a las series de datos históricos, es el método más común para el análisis de frecuencia. Con

esto, la magnitud de los eventos se determina a partir de la línea de mejor ajuste (que es la distribución considerada). Algunos hidrólogos han identificado en sus estudios, que registros históricos menores a diez años no son factibles para desarrollar un análisis de frecuencia [Viessman *et al.* (1977)]. Con forme a la evolución en que se ha presentado el análisis de frecuencia en hidrología, diversos autores recomiendan contar con al menos 30 años de registros históricos para sus análisis [Escalante (2008)].

El periodo de retorno ( $T$ ) matemáticamente está definido como el inverso de la probabilidad de excedencia, es decir:

$$T = \frac{1}{P(X \geq x)}, \quad (2.39)$$

o bien como

$$T = \frac{1}{1 - P(X \leq x)} = \frac{1}{1 - F(x)}, \quad (2.40)$$

donde

$T$ , periodo de retorno univariado,

$X$ , variable aleatoria,

$x$ , evento,

$P(X \geq x) = 1/T$ , es la probabilidad de excedencia,

$P(X \leq x) = 1 - 1/T$ , es la probabilidad de no excedencia,

$F(x)$ , función de distribución de probabilidad.

A partir de (2.40), la determinación de eventos para diferentes periodos de retorno, una vez que se ha elegido una función de distribución de probabilidad, es prácticamente directa, ya que para un  $T$  dado, se calcula  $F(x)$  y con la FDA considerada se determina el valor del cuantil.

### 2.5.1. Selección del Periodo de Retorno

La legislación nacional en materia de agua, en la mayoría de los países generalmente se establece un período de retorno para llevar a cabo una obra hidráulica.

La selección del periodo de retorno para diseñar una obra es un procedimiento complejo, para lo cual se debe realizar un estudio hidrológico detallado y un análisis económico del problema. Para la elección óptima se considera que al incrementar la magnitud del evento de diseño, el costo de la obra aumenta y al mismo tiempo, la probabilidad de excedencia disminuye. Si disminuye notablemente el evento de diseño aumenta el riesgo de falla.

El primer concepto que se debe tomar en cuenta para proponer un periodo de retorno es el tipo de obra hidráulica que se va a realizar y la caracterización de la zona que se desea proteger, es decir, conocer las condiciones en las que se va a construir nuestra obra hidráulica y las consecuencias directas que se tendrían al seleccionar cierto periodo de retorno. En las tablas 2.1, 2.2 y 2.3 se muestran algunas recomendaciones para elegir un adecuado periodo de retorno conforme a las características de la obra hidráulica que se planea construir [Domínguez *et al.* (2003)].

**Tabla 2.1.** Períodos de retorno para el diseño de presas de almacenamiento.

CATEGORÍA	CARACTERÍSTICAS DE LA PRESA		POTENCIAL DE DAÑOS		PERIODO DE RETORNO
	Almacenamiento (Hm <sup>3</sup> )	Altura (m)	Pérdida de vidas	Daños materiales	
<i>Pequeña</i>	Menor de 1.5	Menor de 15	Ninguna	Menor que el costo de la presa	100 a 500 años
			Moderada	Del orden del costo de la presa	250 a 1000 años
			Considerable	Mayor que el costo de la presa	500 a 10,000 años
<i>Mediana</i>	Entre 1.5 y 60	Entre 15 y 30	Ninguna	Dentro de la capacidad financiera	500 a 10,000 años
			Moderada	Ligeramente mayor que la capacidad financiera	1000 a 10,000 años
			Considerable	Mayor que la capacidad financiera	≥ 10,000 años*
<i>Mayor (No falla)</i>	Mayor de 60	Mayor de 30	Considerable	Mayor que la capacidad financiera	≥ 10,000 años*

\*Para las presas cuya cuenca se encuentra protegida del ingreso directo de huracanes se recomienda utilizar avenida de 10,000 años de periodo de retorno estimada, utilizando métodos estadísticos: en caso contrario es necesario verificar el cálculo de la avenida máxima probable.

**Tabla 2.2.** Períodos de retorno para presas derivadoras.

CATEGORÍA	SUPERFICIE (Ha)	PERIODO DE RETORNO (Años)
Zona de riego pequeña	menor de 1,000	50 – 100
Zona de riego mediana	1,000 a 10,000	100 – 500
Zona de riego grande	más de 10,000	500 – 1000

Nota. La clasificación y el tamaño se refieren a la zona afectable en caso de falla.

**Tabla 2.3.** Períodos de retorno para obras de desvío temporal.

CATEGORÍA	TIPO DE OBRA	PERIODO DE RETORNO (Años)
Pequeña	Presa	10 – 25
Mediana	Presa	25 – 50
Grande	Presa	50 – 100
- -	Cauce de alivio en corriente	25 – 50 (o mayor, según importancia)

Para el caso de grandes presas, la selección del periodo de retorno está influenciada por un factor económico, ya que también esta obra hidráulica puede satisfacer otros usos, como por ejemplo, el abastecimiento de agua potable, agua para riego y la generación de energía. También la protección de vidas humanas, bienes materiales e infraestructura, representan beneficios económicos.

El periodo de retorno en presas se refiere sobre todo al diseño de la obra de excedencia (vertedores). Este diseño a su vez está relacionado con el gasto máximo extraordinario (avenida de diseño) que determina su tamaño, es decir, la magnitud con la que se diseña la capacidad de la obra de excedencia (vertedores), la altura de la obra de contención (cortina) y el volumen de regulación del vaso de la presa.



# Capítulo 3

## Teoría de Funciones Cópula

### 3.1. Introducción

Las exigencias de los eventos extremos actualmente demandan modelos estadísticos cada vez más precisos. Exigencias que van desde normativas más estrictas y asertivas para la legislación en la elaboración de proyectos o revisión de las estructuras hidráulicas de presas, o como es el caso, de fenómenos hidrológicos que cada vez son más complejos.

En la última década se han elaborado estudios con mayor frecuencia, utilizando modelos estadísticos en donde no sólo se explique por separado el comportamiento de cada variable, sino modelos más flexibles que representen la asociación entre éstas, es decir *modelos conjuntos*. El hacer modelos conjuntos requiere tomar en consideración el concepto de *dependencia*. El análisis de dependencia entre variables aleatorias es un tema que suele tomarse muy a la ligera, o que simplemente, por facilidad, se opta por suponer independencia.

Un problema que desde hace mucho tiempo ha interesado a los investigadores, indistintamente de la rama a la que se dediquen, ha sido la relación entre una función de distribución multivariada y sus marginales. Este problema ha sido abordado por diversos autores a partir de los años 50 y se han obtenido resultados importantes.

Han sido varios autores los que han contribuido a solucionar este problema, entre los que destacan Fréchet (1951, 1958) y Sklar (1959).

La relación de una función de distribución multivariada con sus marginales uniformes, está dada mediante una función con determinadas características y a la que Sklar (1959) denomina *cópula*. *El propósito que implica esta función es unir las marginales uniformes para obtener una función de distribución conjunta.*

Las cópulas se han convertido en la actualidad en una herramienta de gran potencia para la modelación de datos multivariados, en los que la dependencia entre variables aleatorias existe y el supuesto de normalidad multivariada no se tiene. Dicho esto, el interés en modelar problemas multivariados que involucran variables dependientes se generaliza en hidrología, lo que convierte a ésta teoría en una forma conveniente para modelar la estructura de dependencia en distribuciones conjuntas de variables aleatorias (gasto, volumen, duración, precipitación).

La implementación de las funciones cópula pueden llegar a implicar procesos computacionales arduos, más aún, cuando las dimensiones de éstas aumentan.

Actualmente, existen diversos softwares con aplicación estadística que ofrecen un catálogo para llevar a cabo un análisis de cópulas de manera convincente, por ejemplo: R, Matlab, Mathematica, S-plis, SAS, XploRe y Microsoft Excel (complemento adicional). Sin embargo, la gran mayoría de estos softwares cuentan con una paquetería limitada de cópulas, en general sólo disponen con un conjunto de familias más comunes: Clayton, Frank, Gumbel, Gaussiana y t-Student. El único paquete que, en cierto sentido, ofrece una mayor variedad de familias cópula es **R**, por su naturaleza de "*software libre*".

A pesar del amplio grado de interés acerca de las cópulas, todavía creemos que este concepto se encuentra en "su infancia" y otras tantas investigaciones pueden (y deben) ser conducidas con el fin de subrayar si las cópulas, o conceptos relacionados con cópulas, pueden ser realmente considerados como un "fuerte" concepto matemático de interés en diversas aplicaciones [Durante y Sempi (2010)].

## 3.2. Conceptos y resultados básicos

La noción de cópula fue introducida por **Abe Sklar** en 1959, interesado en investigar acerca de la relación entre una función de probabilidad multidimensional y sus marginales de menor dimensión. Estudios preliminares abordaron este problema, pero Sklar (1959) es quien obtiene el resultado más contundente al respecto, introduciendo la noción y el nombre de **cópula**, mediante el "teorema de Sklar" [Nelsen (2006)]. Las cópulas son funciones que "unen" o "juntan" funciones de distribución marginales de una dimensión. Dicho de otro modo, *las cópulas son funciones de distribución multivariada tales que sus marginales de una dimensión son uniformes en el intervalo  $[0,1]$ .*

En este trabajo se restringirá al caso bivariado. Si es requerido extender un problema a una dimensión mayor, consulte el libro de Nelsen (2006).

**Definición 3.1.** *Una subcópula bidimensional (2-subcópula) es una función  $C'$  con las siguientes propiedades [Nelsen (2006)]*

1.  $Dom(C') = S_1 \times S_2$  donde  $S_1$  y  $S_2$  son subconjuntos de  $I = [0,1]$  tal que  $0,1 \in S_1$  y  $0,1 \in S_2$ ;
2.  $C'$  es una función fija y 2-creciente;
3. Para todo  $u \in S_1$  y para todo  $v \in S_2$ ,

$$C'(u,1) = u \quad \text{y} \quad C'(1,v) = v \quad (3.1)$$

El rango de la función  $C'$  es el subconjunto  $I = [0,1]$ .

**Definición 3.2.** *Una cópula bidimensional (2-cópula) es una función  $C$  cuyo dominio es  $I^2 = [0,1]^2 = S_1 \times S_2 = [0,1] \times [0,1]$ . De manera equivalente, una cópula  $C(u,v)$  es una función  $C : [0,1]^2 \rightarrow [0,1]$  con las siguientes propiedades [Nelsen (2006)]*

1. Para cada  $u, v \in I = [0,1]$ ,

$$C(u,0) = C(0,v) = 0 \quad (3.2)$$



y

$$C(u, 1) = u \wedge C(1, v) = v; \quad (3.3)$$

2. Para cada  $u_1, u_2, v_1, v_2$  en  $I$ , tal que

$$0 \leq u_1 \leq u_2 \leq 1 \wedge 0 \leq v_1 \leq v_2 = 1 \quad \Rightarrow \quad V_C([u_1, u_2] \times [v_1, v_2]),$$

se cumple que

$$V_C = C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0. \quad (3.4)$$

Observación: la segunda condición se extiende de manera natural a dimensiones mayores a dos. En la expresión (3.4)  $C$  es 2-creciente.

**Teorema 3.3.** *Sea  $C$  una cópula. Entonces para todo  $(u_1, v_1)$  y  $(u_2, v_2) \in \text{Dom } C = S_1 \times S_2$ , se cumple que [Nelsen (2006)]:*

$$|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|, \quad (3.5)$$

por lo tanto  $C$  es uniformemente continua sobre su dominio.

**Teorema 3.4.** *Sea  $C$  una cópula, entonces para cada  $(u, v) \in I^2$ , se cumple que [Nelsen (2006)]:*

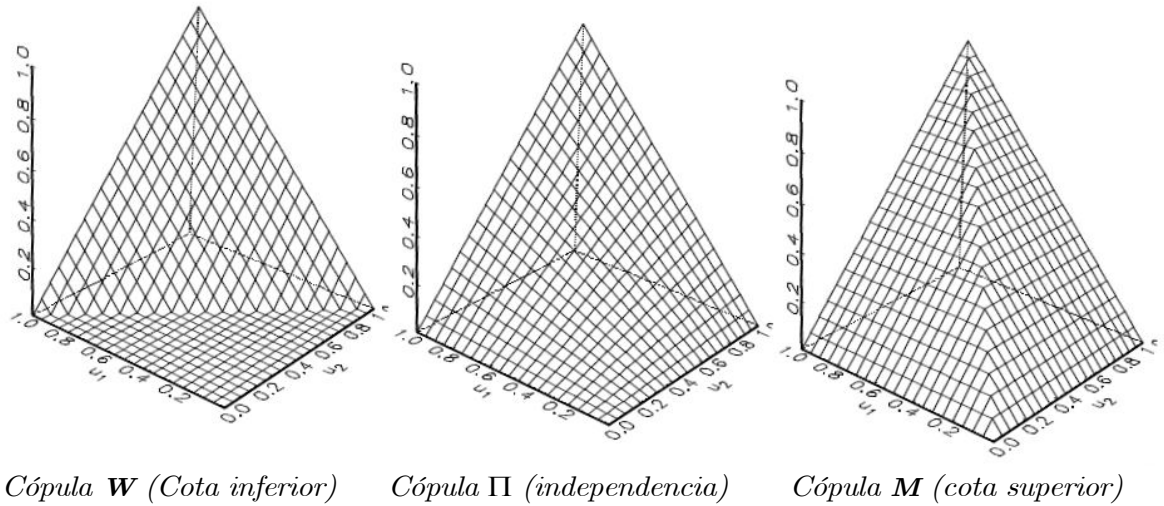
$$W(u, v) = \max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v) = M(u, v). \quad (3.6)$$

Las funciones  $W(u, v)$  y  $M(u, v)$  son cópulas, y comúnmente son llamadas **cotas de Fréchet-Hoeffding** para las cópulas.

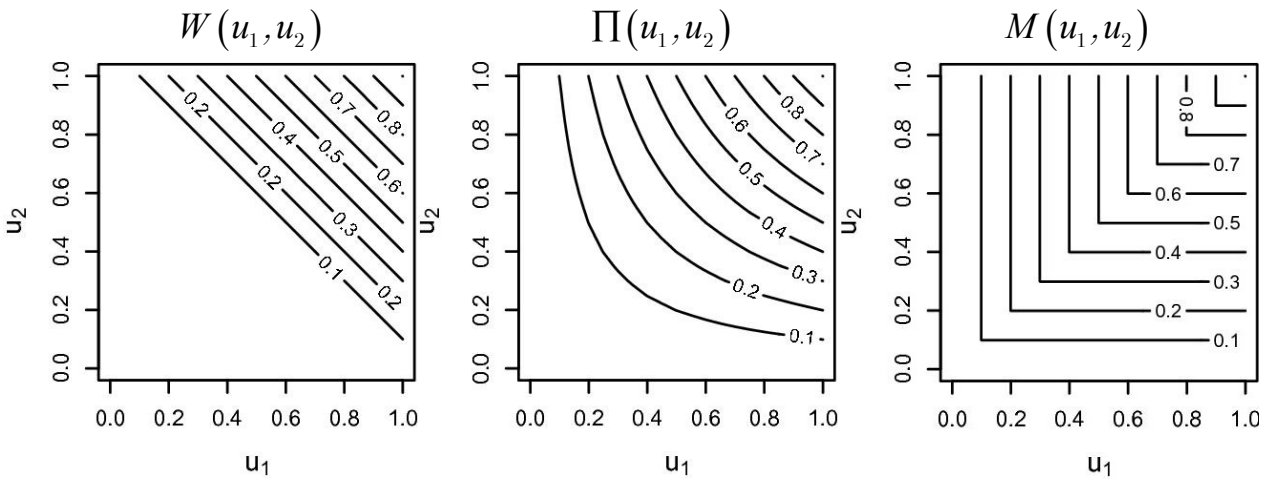
$W$  y  $M$  representan las cotas inferior y superior de Fréchet-Hoeffding, respectivamente. Otra cópula muy frecuente de utilizar es la cópula producto  $\Pi(u, v) = u \cdot v$ , también conocida como cópula independencia.

Como consecuencia de la Definición 3.2, y del Teorema 3.3, la gráfica de cualquier cópula es una superficie continua dentro del cubo unitario  $I^3$ ; y conforme al Teorema 3.4, se deduce que ésta superficie se encuentra dentro de las cotas de Fréchet-Hoeffding (véase Figura 3.1).

Una manera simple pero útil de presentar la gráfica de una cópula, son con los diagramas de contorno [Conway (1979)], que son curvas de nivel, es decir, se bosqueja la gráfica del conjunto en  $I^2$  tal que  $C(u, v) = a$ , consulte la Figura 3.2.



**Figura 3.1.** Superficies posicionadas dentro de las cotas de Fréchet-Hoeffding. Representación de las Gráficas de las Cópulas  $W$ ,  $\Pi$ ,  $M$ .



**Figura 3.2.** Gráficas de Contorno (curvas de nivel) de las cópulas  $W$ ,  $\Pi$ ,  $M$ .

### 3.3. El Teorema de Sklar

El teorema de Sklar exhibe el papel que desempeñan las cópulas en la relación entre funciones de distribución conjunta y sus marginales

univariadas. En esta sección, en primera instancia se facilita una breve explicación sobre las funciones de distribución.

**Definición 3.5.** Una *función de distribución* es una función  $F$  con dominio  $\overline{\mathbb{R}} : \mathbb{R} \cup \{-\infty, \infty\}$  tal que:

1.  $F$  es no decreciente y continua por la derecha.
2.  $F(-\infty) = 0$  y  $F(\infty) = 1$ .

**Definición 3.6.** Una *función de distribución conjunta* es una función  $H$  con dominio  $\overline{\mathbb{R}}^2 := \overline{\mathbb{R}} \times \overline{\mathbb{R}}$  tal que:

1.  $H$  es 2-creciente,
2.  $H(x, -\infty) = 0 = H(-\infty, y)$ ,  $H(\infty, \infty) = 1$  y  $H$  es continua por la derecha.

Observación: La definición anterior implica que  $H$  es fija y como el  $\text{Dom } H = \overline{\mathbb{R}}^2$ , entonces  $H$  tiene marginales  $F$  y  $G$  dadas por  $F(x) = H(x, \infty)$  y  $G(y) = H(\infty, y)$ . Entonces  $F$  y  $G$  son funciones de distribución.

**Teorema 3.7. (Teorema de Sklar).** Sea  $H$  una función de distribución conjunta con marginales  $F$  y  $G$ . Entonces existe una cópula  $C$  tal que para toda  $x, y$  en  $\overline{\mathbb{R}}$

$$H(x, y) = C(F(x), G(y)). \quad (3.7)$$

Si  $F$  y  $G$  son continuas, entonces  $C$  es única; de otra manera  $C$  está determinada en  $\text{Rango } F \times \text{Rango } G$  [siendo,  $\text{Rango } F := \{F(x) \mid x \in \mathbb{R}\}$  y  $\text{Rango } G := \{G(y) \mid y \in \mathbb{R}\}$ ]. De manera inversa, si  $C$  es una cópula y  $F, G$  son funciones de distribución, entonces la función  $H$  definida en (3.7) es una función de distribución conjunta con marginales  $F$  y  $G$ .

**Definición 3.8.** Sea  $F$  una función de distribución. Una *cuasi inversa* de  $F$  es cualquier función  $F^{(-1)}$  con dominio  $I = [0, 1]$  tal que,

1. Si  $t$  está en  $\text{Rang } F$ , entonces  $F^{(-1)}(t) = x$  es cualquier número  $x$  en  $\overline{\mathbb{R}}$  tal que  $F(x) = t$ , es decir, para todo  $t$  en  $\text{Rang } F$

$$F(F^{(-1)}(t)) = t ;$$

2. Si  $t$  no está en Rang  $F$ , entonces

$$F^{(-1)}(t) = \inf \{x | F(x) \geq t\} = \sup \{x | F(x) \leq t\}.$$

Sí  $F$ , es estrictamente creciente, entonces tiene una única cuasi-inversa, que de hecho es  $F^{(-1)} = F^{-1}$ , en donde  $F^{-1}$  es la inversa usual.

**Corolario 3.9.** *Sea  $H$  una función de distribución conjunta bivariada con marginales continuas  $F$  y  $G$ , y sea  $C$  la única cópula tal que (3.7) se cumple. Entonces para cualquier  $(u, v)$  en  $[0, 1]^2$*

$$C(u, v) = H\left(F^{(-1)}(u), G^{(-1)}(v)\right). \quad (3.8)$$

Siendo  $F^{(-1)}(u) = x$  y  $G^{(-1)}(v) = y$ , las cuasi-inversas de  $F(x) = u$  y  $G(y) = v$  respectivamente. Este corolario ilustra un método para la construcción de cópulas cuando  $F$  y  $G$  son funciones de distribución continuas.

*En resumen, dada una función de distribución conjunta  $H$ ; correspondiente a un par de variables aleatorias continuas con funciones de distribución (marginal)  $F$  y  $G$ ; es posible "extraer" la cópula subyacente  $C_{XY}$  mediante el Corolario 3.9, y en seguida se construye una nueva función de distribución conjunta  $H_1$ ; con la misma cópula pero distintas funciones de distribución (marginal)  $F_1$  y  $G_1$ ; esto es  $H_1(x, y) = C(F_1(x), G_1(y))$  [Erdely (2009)].*

### 3.4. Cópulas y variables aleatorias

**Definición 3.10.** *La función de distribución de probabilidad de una variable aleatoria  $X$  es la función  $F : \overline{\mathbb{R}} \rightarrow [0, 1]$ , si para todo  $x \in \mathbb{R}$*

$$F(x) = P\{\omega : X(\omega) \leq x\}, \quad (3.9)$$

*entonces  $F$  es creciente y continua por la derecha, con  $F(\infty) = 1$  y  $F(-\infty) = 0$ .*

Una variable aleatoria  $X$  se dice que es continua si y sólo si la función de distribución es continua para todo  $x \in \mathbb{R}$ . Equivalentemente,  $X$  es continua si  $P[X = x] = 0$  para todo  $x$ .

El teorema siguiente restablece el *Teorema de Sklar en términos de variables aleatorias* y sus funciones de distribución.

**Teorema 3.11.** *Sean  $X$  e  $Y$  variables aleatorias con funciones de distribución  $F$  y  $G$ , respectivamente, y con una función de distribución conjunta  $H$ . Entonces existe una cópula  $C$  tal que  $H(x, y) = C(F(x), G(y))$ . Si  $F$  y  $G$  son continuas entonces  $C$  es única, de lo contrario  $C$  es única solamente en  $\text{Rang } F \times \text{Rang } G$ .*

**Teorema 3.12.** *Las variables aleatorias continuas  $X$  e  $Y$  son independientes si y sólo si  $C_{XY} = \Pi$ .*

**Demostración:**

$$\begin{aligned} X \text{ e } Y \text{ independientes} &\Leftrightarrow H(x, y) = F(x) \cdot G(y), \\ &\Leftrightarrow H(x, y) = C(F(x), G(y)), \end{aligned}$$

donde  $C(u, v) = \Pi(u, v) = u \cdot v$ , y como  $F$  y  $G$  son continuas, entonces  $C$  es única.

### 3.5. Familias de Cópulas

En los problemas de modelización estadística en hidrología, las funciones cópula se han implementado como una excelente alternativa para análisis bivariado. Entre la amplia variedad de familias de cópulas que existen, las más conocidas son: las Cópulas Elípticas, Arquimedianas, Valores Extremos y otras familias como Plackett, Farlie-Gumbel-Morgenstern (FGM), etc. [Véase cuadro 3.1].

Varios autores han trabajado en el desarrollo de familias de cópulas, un extenso listado de clases de cópulas están disponibles para su aplicación en los libros de Nelsen (2006) y Joe (2015), además de las investigaciones

que han realizado Salvadori, De Michele, Genest y Favre, por mencionar algunos autores.

Las familias más populares con aplicación en Ingeniería Civil (hidrología) son las Arquimedianas y las cópulas de Valores Extremos. Los alcances en este trabajo, únicamente *se hará uso de funciones cópula con un parámetro* ( $\theta$ ), debido a su versatilidad y ahorro en procesos de cálculos.

### 3.5.1. Cópulas Elípticas

Esta familia de cópulas se caracteriza por compartir propiedades de la función de distribución de probabilidad normal multivariada, tal es el caso como la *simetría radial*, o el hecho de que la estructura de dependencia esté determinada por la matriz de correlación. Este tipo de familias están asociadas únicamente a las distribuciones elípticas, por ejemplo es el caso de dos cópulas; la Cópula Gaussiana (Normal) y t-Student [Embrechts *et al.* (2001)].

**Definición 3.13.** (Distribución Elíptica) *Sea  $X$  un vector aleatorio  $n$ -dimensional. Se dice que  $X$  tiene distribución elíptica con parámetros  $\mu(n \times 1)$  y  $\Sigma(n \times n)$  si tiene representación estocástica*

$$X \stackrel{d}{=} \mu + rAY, \quad (3.10)$$

donde  $r \geq 0$  es una variable aleatoria,  $Y$  tiene distribución esférica uniforme en  $\mathbb{R}^n$ ,  $A$  es una matriz constante de  $n \times n$  tal que  $AA^T = \Sigma$  con  $\text{rango}(\Sigma) = k$  y el signo  $\stackrel{d}{=}$  significa que ambos lados de la igualdad tienen la misma distribución [Fang *et al.* (1998)].

- **Cópula Gaussiana**

Nelsen (2006) señala que conforme al teorema de Sklar es posible conocer la existencia de una cópula única, en vista de lo anterior, Embrechts *et al.* (2001) señalaron que la Cópula de una función de distribución de probabilidad normal bivariada se expresa como

$$C(u, v; R) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v); R), \quad u, v \in [0, 1]^2, \quad (3.11)$$

donde

$\Phi_2(\cdot, \cdot)$  denota la función de distribución conjunta de una función de distribución de probabilidad normal **estándar** bivariada con media  $(0, 0)^T$  y matriz de covarianza  $R$ , siendo igual a una matriz positiva y definida de  $2 \times 2$  cuyos elementos fuera de la diagonal son cada uno iguales al coeficiente de correlación  $\rho$ , con  $\rho \in [-1, 1]$ , y los elementos en la diagonal son iguales a uno,

$\Phi^{-1}(\cdot)$  denota la inversa de la función de distribución de probabilidad normal **estándar** univariada.

Embrechts *et al.* (2001, p. 25) presentaron una versión de la **Cópula Gaussiana** para el caso bivariado y se expresa como

$$C(u, v; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \exp\left[-\frac{t_1^2 - 2\rho t_1 t_2 + t_2^2}{2(1-\rho^2)}\right] dt_1 dt_2, \quad (3.12)$$

donde

$t_1 = \Phi^{-1}(u)$  Cuantil asociado a la función de distribución de probabilidad normal estándar univariada  $u$ ,

$t_2 = \Phi^{-1}(v)$  Cuantil asociado a la función de distribución de probabilidad normal estándar univariada  $v$ .

La función densidad de la Cópula gaussiana está dada por:

$$c(u, v; \rho) = \frac{\phi_2(\Phi^{-1}(u), \Phi^{-1}(v); \rho)}{\phi(\Phi^{-1}(u)) \cdot \phi(\Phi^{-1}(v))}, \quad u, v \in [0, 1]^2 \quad (3.13)$$

donde

$\phi_2(\cdot, \cdot)$  denota la función densidad de probabilidad normal estándar bivariada,

$\phi(\cdot)$  denota la función densidad de probabilidad normal estándar univariada,

$\Phi^{-1}(\cdot)$  denota la inversa de la función de distribución de probabilidad normal estándar univariada.

Meyer (2009, p. 4) presenta una fórmula para el caso bivariado de la función **densidad de la Cópula Gaussiana**, escrita como

$$c(u, v; \rho) = \frac{1}{\sqrt{1 - \rho^2}} \exp \left[ -\frac{\rho^2 t_1^2 - 2\rho t_1 t_2 + \rho^2 t_2^2}{2(1 - \rho^2)} \right], \quad (3.14)$$

donde

$$t_1 = \Phi^{-1}(u), \quad t_2 = \Phi^{-1}(v).$$

El uso de la cópula Gaussiana bivariada es atractivo ya que codifica la dependencia en la misma forma en que la distribución normal bivariada lo hace, usando el coeficiente  $\rho$ . La notable diferencia radica en que es factible utilizar variables aleatorias con cualesquiera marginales arbitrarias.

Esta cópula tiene la capacidad de capturar el rango completo de dependencia, ya que incluye las cópulas de cota superior e inferior de Fréchet-Hoeffding como el modelo de independencia. Este último caso se obtiene cuando  $\rho = 0$  y define la cópula de independencia.

#### ▪ Cópula t-Student (t-cópula)

La d-variada densidad de la *función de distribución de probabilidad t-Student* con  $\nu$  grados de libertad se define como [Kotz y Nadarajah (2004)]



$$f_{\Sigma, \nu}(X; \nu, \mu, \Sigma) = \frac{\Gamma\left(\frac{\nu + d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{(\pi\nu)^d |\Sigma|}} \left(1 + \frac{(X - \mu)^T \Sigma^{-1} (X - \mu)}{\nu}\right)^{-\frac{\nu+d}{2}}, \quad (3.15)$$

donde  $X \in \mathbb{R}^d$  es un vector aleatorio,  $\mu \in \mathbb{R}^d$  es el vector de medias,  $\Sigma \in \mathbb{R}^{d \times d}$  es la matriz de dispersión (covarianza) y  $\Gamma$  es la función Gama.

Demarta y McNeil (2004) señalan la valiosa observación de que la cópula de una distribución d-variada t-Student  $f_{\Sigma, \nu}(\nu, \mu, \Sigma)$  es idéntica a la de una distribución d-variada t-Student **estándar**  $f_{R, \nu}(\nu, 0, R)$ , con media cero y matriz de correlación  $R$ , entonces se define lo siguiente.

**Definición 3.14.** *Sea  $R$  una matriz simétrica y definida positiva con diagonal  $R = 1$  y  $T_{R, \nu}$  la función de distribución de probabilidad d-variada t-Student estándar con  $\nu$  grados de libertad y matriz de correlación  $R$ , con base en ello, la cópula d-variada t-student se define como en [Bouyé (2000)]:*

$$C(u_1, \dots, u_d; R, \nu) = T_{R, \nu}\left(T_{\nu}^{-1}(u_1), \dots, T_{\nu}^{-1}(u_d)\right), \quad u_1, \dots, u_d \in [0, 1]^d, \quad (3.16)$$

siendo  $T_{\nu}^{-1}(\cdot)$  la inversa de la función de distribución de probabilidad univariada t-Student estándar.

Mediante el uso de las ecuaciones (3.15) y (3.16), Demarta y McNeil (2004) han propuesto una versión no cerrada de la t-cópula d-variada en términos de  $\nu$ ,  $\Gamma$ ,  $X^T$  y el determinante de  $R$  como

$$C(u_1, \dots, u_d; R, \nu) = \int_{-\infty}^{a_1} \dots \int_{-\infty}^{a_d} \frac{\Gamma\left(\frac{\nu + d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) (\pi\nu)^{d/2} \sqrt{\det R}} \left(1 + \frac{X^T R^{-1} X}{\nu}\right)^{-\frac{\nu+d}{2}} dX, \quad (3.17)$$

donde  $a_i = T_{\nu}^{-1}(u_i)$  denota la función cuantil asociada a la función de distribución de probabilidad univariada t-Student estándar.

La densidad de la t-cópula d-variada es [Demarta y McNeil (2004)]

$$c(u_1, \dots, u_d; R, \nu) = \frac{f_{R, \nu}(T_\nu^{-1}(u_1), \dots, T_\nu^{-1}(u_d))}{\prod_{i=1}^d f_\nu(T_\nu^{-1}(u_i))}, \quad u_1, \dots, u_d \in [0, 1]^d, \quad (3.18)$$

donde  $f_{R, \nu}$  es la densidad de la función de distribución de probabilidad d-variada t-Student **estándar** con  $\nu$  grados de libertad, media cero y matriz de correlación  $R$  y  $f_\nu$  la densidad de la función de distribución de probabilidad univariada t-Student **estándar** con  $\nu$  grados de libertad.

Con base en 3.18, Bouyé (2000) presenta una versión de la función densidad de la t-cópula d-variada en términos de  $\nu$ ,  $\Gamma$  y el determinante de  $R$  como:

$$c(u_1, \dots, u_d; R, \nu) = \frac{\Gamma\left(\frac{\nu}{2}\right)^{d-1} \Gamma\left(\frac{\nu+d}{2}\right) \left(1 + \frac{\zeta^T R^{-1} \zeta}{\nu}\right)^{-\frac{\nu+d}{2}}}{\Gamma\left(\frac{\nu+1}{2}\right)^d \sqrt{\det R} \prod_{i=1}^d \left(1 + \frac{\zeta_i^2}{\nu}\right)^{-\frac{\nu+d}{2}}}, \quad (3.19)$$

donde

$$\zeta = [T_\nu^{-1}(u_1), \dots, T_\nu^{-1}(u_d)].$$

*Para el caso de cópulas bivariadas ( $d = 2$ ), las expresiones<sup>1</sup> de la cópula y su respectiva densidad se han deducido conforme a (3.17) y (3.19).*

**La Cópula bivariada t-Student es**

$$C(u_1, u_2; \rho, \nu) = \int_{-\infty}^{T_\nu^{-1}(u_1)} \int_{-\infty}^{T_\nu^{-1}(u_2)} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \pi \nu \sqrt{1-\rho^2}} \cdot \left[1 + \frac{t_1^2 - 2\rho t_1 t_2 + t_2^2}{\nu(1-\rho^2)}\right]^{-\frac{\nu+2}{2}} dt_1 dt_2, \quad (3.20)$$

donde

$$t = [T_\nu^{-1}(u_1) \quad T_\nu^{-1}(u_2)]$$

---

<sup>1</sup>Se hace la observación para que el lector tome las debidas precauciones en consultar el caso bivariado, ya que diversas publicaciones presentan errores de escritura en ambas expresiones, véase en Alexander (2008, p. 268) y por Chen y Guo (2019, p. 24).

$t_1 = T_v^{-1}(u_1)$  Cuantil asociado a la función de distribución de probabilidad univariada t-Student estándar  $u_1$ ,

$t_2 = T_v^{-1}(u_2)$  Cuantil asociado a la función de distribución de probabilidad univariada t-Student estándar  $u_2$ .

**La Función Densidad de la Cópula bivariada t-Student es**

$$c(u_1, u_2; \rho; \nu) = \frac{\Gamma\left(\frac{\nu}{2}\right) \cdot \Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)^2 \sqrt{1-\rho^2}} \cdot \left[1 + \frac{t_1^2 - 2\rho t_1 t_2 + t_2^2}{\nu(1-\rho^2)}\right]^{-\frac{\nu+2}{2}} \cdot \left[(1 + \nu^{-1} t_1^2)(1 + \nu^{-1} t_2^2)\right]^{\frac{\nu+1}{2}}, \quad (3.21)$$

donde

$$t = \left[ T_v^{-1}(u_1) \quad T_v^{-1}(u_2) \right],$$

$t_1 = T_v^{-1}(u_1)$  y  $t_2 = T_v^{-1}(u_2)$  Cuantiles asociados a la función de distribución univariada t-Student estándar [Joe (2015, p. 181)].

Cuando el parámetro  $\nu$  que representa los grados de libertad de la cópula *t-Student* tiene un valor muy elevado, la cópula tiende a ser una cópula *Gaussiana*.

▪ **Aplicaciones de copulas elípticas en hidrología**

Chen y Guo (2019) exponen una breve reseña sobre las aplicaciones que han tenido las cópulas gaussiana y t-student; se han aplicado en análisis de sequías en cuencas, asimismo en el análisis de frecuencia d-variado de avenidas: Genest y Favre (2007) implementaron las pruebas de bondad de ajuste para cópulas elípticas y analizaron datos como gasto pico, volumen y duración del río Romaine, Canadá; Wang *et al.* (2010) construyeron un modelo trivariado entre el volumen, duración y la intensidad máxima de eventos de lluvia extrema en 12 estaciones hidrométricas en Connecticut, Estados Unidos; Chen *et al.* (2013) midió la correlación de gastos en diferentes ríos de China, logrando construir distribuciones conjuntas en diversas estaciones hidrométricas.

Demarta y McNeil (2004, p. 12) presentaron una nueva cópula de valores extremos, la que denominaron t-VE o t-cópula límite, en futuros trabajos será de utilidad implementar esta opción para el análisis de eventos extremos. En la Tabla 3.1 se muestra un resumen de cópulas elípticas.

### 3.5.2. Cópulas Arquimedianas

Todas las cópulas Arquimedianas encontradas comúnmente son expresiones de forma cerrada. Su popularidad también proviene del hecho de que permiten una gran variedad de estructuras de dependencia.

Las cópulas arquimedianas disponen de una amplia gama de facilidades para su aplicación, las cuales ofrecen lo siguiente [Nelsen (2006)]:

1. Fácil implementación para construirlas;
2. Fácil acotación y simulación (*función generadora*);
3. Flexibilidad;
4. Propiedades y aplicaciones más estudiadas;
5. Variedad de familias que pertenecen a esta clase.

**Definición 3.15.** Sea  $\varphi : I \rightarrow [0, \infty]$  una función continua, estrictamente decreciente tal que  $\varphi(1) = 0$ . La pseudo-inversa de  $\varphi$  es la función  $\varphi^{[-1]}$  con  $\text{Dom } \varphi^{[-1]} = [0, \infty]$  y  $\text{Rango } \varphi^{[-1]} = I$ , y está parametrizada por

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & \text{si } 0 \leq t \leq \varphi(0), \\ 0, & \text{si } \varphi(0) \leq t \leq \infty. \end{cases} \quad (3.22)$$

Note que  $\varphi^{[-1]}$  es continua y decreciente en  $[0, \infty]$ , y es estrictamente decreciente en  $[0, \varphi(0)]$ .

Además,  $\varphi^{[-1]}(\varphi(u)) = u$  en  $I$ , y

$$\begin{aligned} \varphi(\varphi^{[-1]}(t)) &= \begin{cases} t & , \quad \text{si } 0 \leq t \leq \varphi(0), \\ \varphi(0), & \text{si } \varphi(0) \leq t \leq \infty, \end{cases} \\ &= \min(t, \varphi(0)). \end{aligned}$$

Finalmente, si  $\varphi(0) = \infty$ , entonces  $\varphi^{[-1]} = \varphi^{-1}$ , en donde  $\varphi^{-1}$  es la inversa usual.

**Lema 3.16.** *Sea  $\varphi : I \rightarrow [0, \infty]$  una función continua, estrictamente decreciente tal que  $\varphi(1) = 0$ , y sea  $\varphi^{[-1]}$  la pseudo-inversa de  $\varphi$  definida en (3.21). Sea  $C : I^2 \rightarrow I$  dada por*

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)). \quad (3.23)$$

Entonces  $C$  satisface las condiciones de frontera (3.2) y (3.3) para una cópula.

En el siguiente lema, se obtiene una condición necesaria y suficiente para que la función  $C$  en (3.23) sea 2-creciente.

**Lema 3.17.** *Sean  $\varphi$ ,  $\varphi^{[-1]}$  y  $C$  bajo la hipótesis del Lema 3.16. Entonces  $C$  es 2-creciente si y solo si cuando  $u_1 \leq u_2$ ,*

$$C(u_2, v) - C(u_1, v) \leq u_2 - u_1. \quad (3.24)$$

**Teorema 3.18.** *Sea  $\varphi$  una función continua, estrictamente decreciente de  $I$  a  $[0, \infty]$  tal que  $\varphi(1) = 0$ , y sea  $\varphi^{[-1]}$  la pseudo-inversa de  $\varphi$  definida en (3.22). Entonces la función  $C : I^2 \rightarrow I$  dada por (3.23) es una cópula si y solo si  $\varphi$  es convexa.*

Rodríguez (2003, p. 67) resume la idea de cópula arquimediana y la función generadora como: Sea  $\Phi$  el conjunto de funciones  $\varphi : [0, 1] \rightarrow [0, \infty]$  cuyas características son **continuas, estrictamente decrecientes, convexas**, tal que  $\varphi(1) = 0$  y **con pseudo-inversa** de  $\varphi$ . Cada  $\varphi \in \Phi$  genera una cópula  $C$ , es decir, se genera una función de distribución bivariada con marginales uniformes  $[0, 1]$  dada por

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)), \quad 0 \leq u, v \leq 1.$$

Nelsen (2006, p. 112) señala que la familia de cópulas cuyas características sean de la forma como en (3.23), se les llama "**Cópulas Arquimedianas**", asimismo, la función  $\varphi$  se le llama "**Generador de la**

**Cópula  $C$  "**. Si  $\varphi(0) = \infty$ , se dice que  $\varphi$  es un generador estricto. En este caso,  $\varphi^{[-1]} = \varphi^{-1}$  (inversa usual) y  $C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$  es una **cópula arquimediana estricta**. Si  $\varphi(0) < \infty$  entonces  $\varphi$  es un generador no estricto.

**Teorema 3.19.** *Sea  $C$  una cópula Arquimediana con generador  $\varphi$ . Entonces se consideran las siguientes propiedades:*

1.  $C$  es simétrica, es decir  $C(u, v) = C(v, u)$  para todo  $u, v \in I$ ;
2.  $C$  es **asociativa**, es decir  $C(C(u, v), w) = C(u, C(v, w))$  para toda  $u, v, w \in I$ ;
3. Si  $c > 0$  es cualquier constante, entonces  $c \cdot \varphi$  es también generador de  $C$ .

En Tabla 3.2 se muestra un conjunto de familias de cópulas arquimedianas más conocidas para su implementación en fenómenos ocurridos en Ingeniería Civil (hidrología), estas son: Clayton, Frank, Ali-Mikhail-Haq, Gumbel-Hougaard y Joe.

Para las cópula de Clayton, Frank y Gumbel-Hougaard:  $C_\theta$ ,  $\varphi_\theta(t)$  y  $\varphi_\theta^{-1}(t)$  son citadas en Nelsen (2006), asimismo todas las  $c_\theta$  están referenciadas en Joe (2015). En lo que respecta a la cópula de Ali-Mikhail-Haq:  $C_\theta$ ,  $\varphi_\theta(t)$  se citan en Nelsen (2006),  $\varphi_\theta^{-1}(t)$  está citada en Mächler (2014) y  $c_\theta$  se menciona en Joe (1997). En el caso de la cópula de Joe, todas las funciones están citadas en el libro de Joe (2015).

Donde  $C_\theta$  es la cópula del modelo,  $c_\theta$  es la correspondiente función densidad de la cópula del modelo,  $\varphi_\theta(t)$  es el generador de la cópula y  $\varphi_\theta^{-1}(t)$  es la inversa usual del generador.

La cópula de Gumbel-Hougaard es también considerada una cópula de valores extremos, cabe señalar que dicha cópula posee una función de dependencia de Pickands, en vez de una función generador.

### 3.5.3. Cópulas de Valores Extremos

En esta sección se introducen aquellas familias de cópulas que permiten modelar eventos extremos bivariados. En este trabajo no se profundizará con los elementos teóricos que caracterizan a esta clase de cópulas. Por definición, las cópulas de valores extremos son las cópulas con distribución de valores extremos. A continuación se describen algunos elementos básicos que sustentan al teorema de cópulas de valores extremos, las cuales están asociadas a una *Función de Dependencia de Pickands*.

**Definición 3.20.** (Cópula de valores extremos) *Una cópula  $C$  es una cópula de valores extremos si existe otra cópula  $C_F$  tal que, para  $n \rightarrow \infty$ ,*

$$C_F(u_1^{1/n}, \dots, u_d^{1/n})^n \rightarrow C(u_1, \dots, u_d), \quad (3.25)$$

para todo  $(u_1, \dots, u_d) \in [0, 1]^d$ .

En este sentido, la cópula límite  $C$  es determinada de forma única por la cópula  $C_F$ . Mediante un paralelismo con la teoría de valores extremos ya desarrollada, se dice que la cópula  $C_F$  pertenece al dominio de atracción de  $C$ . Esta definición es concebida de forma que las distribuciones marginales estén modeladas por un máximo. La representación de cópulas de valor extremo puede simplificarse si se utiliza el concepto de máxima estabilidad.

**Definición 3.21.** (Cópula Máx-estable) *Se dice que una cópula  $C$  es máx-estable si satisface la siguiente relación*

$$C(u_1, \dots, u_d) = C(u_1^{1/m}, \dots, u_d^{1/m})^m, \quad (3.26)$$

para todo entero  $m \geq 1$  y  $(u_1, \dots, u_d) \in [0, 1]^d$ .

**Teorema 3.22.** *Una cópula  $d$ -variada  $C$  es cópula de valores extremos si y solo si existe una medida finita Borel  $H$  en la unidad simple  $\Delta_{d-1} = \{(u_1, \dots, u_d) \in [0, \infty)^d : \sum_j u_j = 1\}$ , llamada medida espectral, entonces*

$$C(u_1, \dots, u_d) = \exp[-\ell(-\ln u_1, -\ln u_2, \dots, -\ln u_d)], \quad (3.27)$$

para todo  $(u_1, \dots, u_d)^T \in (0, 1]^d$ ,

donde la función de dependencia de cola estable  $\ell : [0, \infty)^d \rightarrow [0, \infty)$  está dada por

$$\ell(x_1, \dots, x_d) = \int_{\Delta_{d-1}} \max_{j=1, \dots, d} \{w_j x_j\} dH(w_1, \dots, w_d),$$

para todo  $(x_1, \dots, x_d)^T \in (0, 1]^d$ .

La medida espectral  $H$  es arbitraria, excepto por las restricciones de los d-momentos

$$\int_{\Delta_{d-1}} w_j dH(w_1, \dots, w_d) = 1, \quad j \in \{1, \dots, d\}.$$

La siguiente definición exhibe una caracterización de las cópulas de valor extremo, lo cual permite realizar la construcción de muchas de ellas.

**Definición 3.23.** (Función de dependencia de Pickands (1981)) *La función de dependencia de Pickands  $A : \Delta_{d-1} \rightarrow [1/d, 1]$  está definida como [Eschenburg, (2013)]*

$$A(w_1, \dots, w_d) := \frac{\ell(x_1, \dots, x_d)}{\sum_{i=1}^d x_i}, \quad (3.28)$$

donde  $w_j := \frac{x_j}{\sum_{j=1}^d x_j}$ ,  $j = 1, \dots, d$

para todo  $(x_1, \dots, x_d)^T \in [0, \infty)^d \setminus \{0\}$ .

Sea  $t_j = \frac{\ln u_j}{\sum_{j=1}^d \ln u_j}$ ,  $j = 1, \dots, d$ , y junto con (3.27), se obtiene la representación más común de la cópula de valores extremos d-variada

$$C(u_1, \dots, u_d) = \exp \left\{ \left( \sum_{j=1}^d \ln u_j \right) A \left( \frac{\ln u_1}{\sum_{j=1}^d \ln u_j}, \dots, \frac{\ln u_{d-1}}{\sum_{j=1}^d \ln u_j} \right) \right\}. \quad (3.29)$$



**Teorema 3.24.** (*Cópula bivariada de valores extremos*) Una cópula bivariada  $C$  es una cópula de valores extremos si y solo si se cumple que [Gudendorf y Segers (2010)]

$$\begin{aligned} C(u_1, u_2) &= \exp \left\{ \left[ \ln(u_1) + \ln(u_2) \right] \cdot A \left( \frac{\ln u_2}{\ln(u_1 \cdot u_2)} \right) \right\}, \quad (3.30) \\ &= \exp \{ A(t) \cdot \ln(u_1 \cdot u_2) \} \\ &= (u_1 \cdot u_2)^{A(t)} \\ &= (u_1 \cdot u_2)^{A \left( \frac{\ln u_2}{\ln(u_1 \cdot u_2)} \right)} \end{aligned}$$

para todo  $(u_1, u_2) \in (0, 1]^2 \setminus \{1, 1\}$ ,

donde  $A: [0, 1] \rightarrow [1/2, 1]$  es convexa y satisface  $\max\{1-t, t\} \leq A(t) \leq 1$ ,

además  $t = \frac{\ln u_2}{\ln(u_1 \cdot u_2)}$ , para todo  $t \in [0, 1]$ .

**Ejemplo 1:** Sea  $C(u, v)$  una cópula de valores extremos (VE) y  $A(t)$  la función de dependencia de Pickands asociada a la cópula de Gumbel-Hougaard. Obtener la cópula de Gumbel-Hougaard.

$$C(u, v) = \exp \{ \ln(u \cdot v) \cdot A(t) \},$$

y

$$A(t) = \left[ t^\theta + (1-t)^\theta \right]^{\frac{1}{\theta}},$$

si se realiza la sustitución  $A(t)$  en la cópula de VE

$$C(u, v; \theta) = \exp \left\{ \ln(u \cdot v) \cdot \left[ t^\theta + (1-t)^\theta \right]^{\frac{1}{\theta}} \right\}$$

si  $t = \ln v / \ln(u \cdot v)$ , entonces resulta lo siguiente

$$C(u, v; \theta) = \exp \left\{ \ln(u \cdot v) \cdot \left[ \left( \frac{\ln v}{\ln(u \cdot v)} \right)^\theta + \left( 1 - \frac{\ln v}{\ln(u \cdot v)} \right)^\theta \right]^{\frac{1}{\theta}} \right\},$$

Si se realiza algebra de funciones logarítmicas, finalmente se obtiene

$$C(u, v; \theta) = \exp \left\{ - \left[ (-\ln u)^\theta + (-\ln v)^\theta \right]^{\frac{1}{\theta}} \right\}$$

La cópula anterior se le conoce como la **Cópula de Valores Extremos de Gumbel-Hougaard**, con  $\theta \in [1, \infty)$ , dicha expresión está asociada a la función de dependencia de Pickands. *Cabe destacar que esta cópula también pertenece a la familia de cópulas Arquimedianas asociada a una función generadora* (consulte las tablas 3.2 y 3.3).

Las cópulas de valores extremos más usadas son; Gumbel-Hougaard (modelo logístico), Galambos (1975) (modelo logístico negativo) y Hüsler-Reiss (1989). En la literatura, diversos investigadores con orientación en hidrología, han utilizado otras cópulas de valores extremos como son: la cópula de Tawn con *dos parámetros* [Tawn (1998)], cópula t-EV (forma similar a la cópula de Hüsler-Reiss y utiliza la FDA t-Student estándar) [Demarta y McNeil (2004)] y la cópula Joe/BB5 con *dos parámetros* [Joe (2015)].

Las respectivas funciones de dependencia de Pickands para las cópulas bivariadas de Gumbel-Hougaard, Galambos, Hüsler-Reiss y BB5/Joe son citadas en Joe (2015) y mostradas en la Tabla 3.3 de este trabajo. Para fines de esta tesis, no se emplearán las funciones de Pickands, únicamente se trabajará con la cópula bivariada de Gumbel-Hougaard, Galambos y Hüsler-Reiss.

### 3.5.4. Otros modelos cópula

- **Cópula bivariada Plackett**

Para  $0 \leq \theta < \infty$ , la cópula es [Joe (2015)]

$$C(u, v; \theta) = \frac{[1 + (\theta - 1)(u + v)] - \sqrt{[1 + (\theta - 1)(u + v)]^2 - 4(\theta - 1)(\theta uv)}}{2(\theta - 1)}, \quad (3.31)$$

y la densidad de la cópula es

$$c(u, v; \theta) = \frac{\theta [1 + (\theta - 1)(u - 2uv + v)]}{\sqrt{\{[1 + (\theta - 1)(u + v)]^2 - 4(\theta - 1)(\theta uv)\}^3}}. \quad (3.32)$$

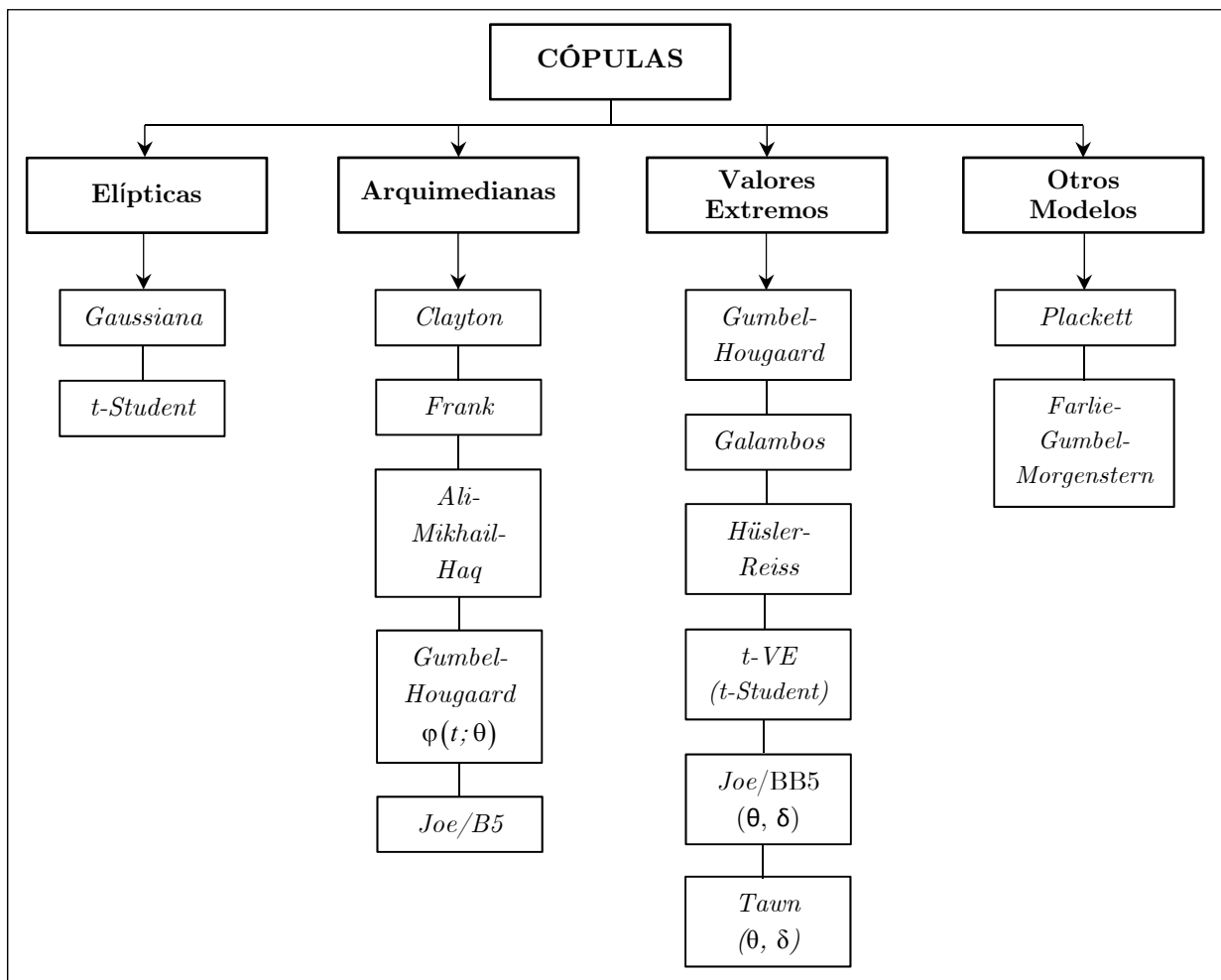
▪ **Cópula bivariada Farlie-Gumbel-Morgenstern (FGM)**

Para  $-1 \leq \theta \leq 1$ , la cópula es [Joe (2015)]

$$C(u, v; \theta) = uv [1 + \theta(1 - u)(1 - v)], \quad (3.33)$$

y la densidad de la cópula se expresa como

$$c(u, v; \theta) = 1 + \theta(1 - 2u) \cdot (1 - 2v). \quad (3.34)$$



**Cuadro 3.1.** Clasificación de Familias de Funciones Cópula más comunes.

**Tabla 3.1.** Cópulas Elípticas bivariadas más comunes, de un parámetro

Función	Coefficiente $\rho$	Cópula Bivariada Gaussiana
Cópula $C(u_1, u_2; \rho)$	$[1, 1]$	$\frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \exp\left[-\frac{t_1^2 - 2\rho t_1 t_2 + t_2^2}{2(1-\rho^2)}\right] dt_1 dt_2$
Densidad de la Cópula $c(u_1, u_2; \rho)$	$[1, 1]$	$\frac{1}{\sqrt{1-\rho^2}} \exp\left[-\frac{\rho^2 t_1^2 - 2\rho t_1 t_2 + \rho^2 t_2^2}{2(1-\rho^2)}\right]$
$t_i = \Phi^{-1}(u_i)$ función cuantil asociada a la función de distribución de probabilidad normal estándar (univariada).		
Función	Coefficiente $\rho$	Cópula Bivariada t-Student
Cópula $C(u_1, u_2; \rho, \nu)$	$[1, 1]$	$\frac{1}{\pi\nu\sqrt{1-\rho^2}} \int_{-\infty}^{T_v^{-1}(u_1)} \int_{-\infty}^{T_v^{-1}(u_2)} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \cdot \left[1 + \frac{t_1^2 - 2\rho t_1 t_2 + t_2^2}{\nu(1-\rho^2)}\right]^{-\frac{\nu+2}{2}} dt_1 dt_2$
$\nu$ es el número de grados de libertad y $\Gamma$ es la función gama.		
Densidad de la Cópula $c(u_1, u_2; \rho, \nu)$	$[1, 1]$	$\frac{1}{\sqrt{1-\rho^2}} \cdot \frac{\Gamma\left(\frac{\nu}{2}\right) \cdot \Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)^2} \cdot \left[1 + \frac{t_1^2 - 2\rho t_1 t_2 + t_2^2}{\nu(1-\rho^2)}\right]^{-\frac{\nu+2}{2}} \cdot \left[\left(1 + \frac{t_1^2}{\nu}\right) \cdot \left(1 + \frac{t_2^2}{\nu}\right)\right]^{\frac{\nu+1}{2}}$
$t_i = T_v^{-1}(u_i)$ función cuantil asociada a la función de distribución de probabilidad t-Student estándar (univariada).		

**Tabla 3.2.** Familias de Cópulas Arquimedianas bivariadas más comunes, de un parámetro

Familia	Parámetro $\theta$	Función Generador $\varphi(t; \theta)$	Inversa del Generador $\varphi^{-1}(t; \theta)$	Función Cópula $C(u, v; \theta)$	Densidad de la Cópula $c_\theta(u, v; \theta)$
Clayton	$[-1, \infty) \setminus \{0\}$	$t^{-\theta} - 1$	$(1+t)^{-\frac{1}{\theta}}$	$(u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$	$(1+\theta)(u \cdot v)^{-\theta-1} (u^{-\theta} + v^{-\theta} - 1)^{\frac{1}{\theta}-2}$
Frank	$(-\infty, \infty) \setminus \{0\}$	$-\ln \left[ \frac{e^{-\theta t} - 1}{e^{-\theta} - 1} \right]$	$-\frac{1}{\theta} \ln [1 - e^{-t} (1 - e^{-\theta})]$	$-\frac{1}{\theta} \ln \left[ 1 + \frac{(e^{-\theta u} - 1) \cdot (e^{-\theta v} - 1)}{e^{-\theta} - 1} \right]$	$-\frac{\theta(e^{-\theta} - 1) \cdot e^{-\theta(u+v)}}{[(e^{-\theta} - 1) + (e^{-\theta u} - 1) \cdot (e^{-\theta v} - 1)]^2}$
Ali- Mikhail -Haq	$[-1, 1)$	$\ln \left[ \frac{1 - \theta(1-t)}{t} \right]$	$\frac{1 - \theta}{e^t - \theta}$	$\frac{u \cdot v}{1 - \theta(1-u) \cdot (1-v)}$	$\frac{[1 - \theta(1-u) \cdot (1-v)] \cdot (1-\theta) + 2\theta uv}{[1 - \theta(1-u) \cdot (1-v)]^3}$
Gumbel- Hougaard	$[1, \infty)$	$(-\ln t)^\theta$	$\exp \left( -t^{\frac{1}{\theta}} \right)$	$\exp \left\{ -[(-\ln u)^\theta + (-\ln v)^\theta]^{\frac{1}{\theta}} \right\}$	$\frac{C(u, v; \theta)}{u \cdot v} \cdot \left\{ (\theta - 1) + [(-\ln u)^\theta + (-\ln v)^\theta]^{\frac{1}{\theta}} \right\} \cdot \left\{ [(-\ln u)^\theta + (-\ln v)^\theta]^{\frac{1}{\theta}-2} \right\} \cdot \{(-\ln u) \cdot (-\ln v)\}^{\theta-1}$
Joe/B5	$[1, \infty)$	$-\ln [1 - (1-t)^\theta]$	$1 - (1 - e^{-t})^{\frac{1}{\theta}}$	$1 - [(1-u)^\theta + (1-v)^\theta - (1-u)^\theta \cdot (1-v)^\theta]^{\frac{1}{\theta}}$	$[(1-u) \cdot (1-v)]^{\theta-1} \cdot [(\theta - 1) + w] w^{\frac{1}{\theta}-2},$ $w = (1-u)^\theta - [(1-u) \cdot (1-v)]^\theta + (1-v)^\theta$

**Tabla 3.3.** Familias de Cópulas bivariadas de Valores Extremos y sus funciones de dependencia de Pickands.

Familia	Parámetros	Función de Dependencia $A(t)$	Función Cópula $C(u, v; \theta)$	Densidad de la Cópula $c_\theta(u, v; \theta)$
Gumbel-Hougaard	$\theta \in [1, \infty)$ $t \in [0, 1]$	$[t^\theta + (1-t)^\theta]^{\frac{1}{\theta}}$	$\exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right\}$	$\frac{C(u, v; \theta)}{u \cdot v} \cdot \left\{(\theta - 1) + \left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right\} \cdot \left\{\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}-2}\right\} \cdot \{(-\ln u) \cdot (-\ln v)\}^{\theta-1}$
Galambos	$\theta \in (0, \infty)$ $t \in [0, 1]$	$1 - \left[t^{-\theta} + (1-t)^{-\theta}\right]^{-\frac{1}{\theta}}$	$u \cdot v \exp\left\{\left[(-\ln u)^{-\theta} + (-\ln v)^{-\theta}\right]^{\frac{1}{\theta}}\right\}$	$\frac{C(u, v; \theta)}{u \cdot v} \cdot \left\{(\theta + 1) + \left[(-\ln u)^{-\theta} + (-\ln v)^{-\theta}\right]^{\frac{1}{\theta}}\right\} \cdot \left[(-\ln u)^{-\theta} + (-\ln v)^{-\theta}\right]^{\frac{1}{\theta}-2} \cdot [(-\ln u) \cdot (-\ln v)]^{-\theta-1} - \left[(-\ln u)^{-\theta} + (-\ln v)^{-\theta}\right]^{\frac{1}{\theta}-1} \cdot [(-\ln u)^{-\theta-1} + (-\ln v)^{-\theta-1}] + 1\right\}$
Hüsler-Reiss	$\theta \in (0, \infty)$ $t \in [0, 1]$	$t \cdot \Phi(Z_t) + (1-t)\Phi(Z_{1-t});$ $Z_t = \frac{1}{\theta} + \frac{\theta}{2} \ln\left(\frac{t}{1-t}\right)$ $Z_{1-t} = \frac{1}{\theta} + \frac{\theta}{2} \ln\left(\frac{1-t}{t}\right)$	$\exp\left\{\ln u \cdot \Phi\left[\frac{1}{\theta} + \frac{\theta}{2} \ln\left(\frac{\ln u}{\ln v}\right)\right] + \ln v \cdot \Phi\left[\frac{1}{\theta} + \frac{\theta}{2} \ln\left(\frac{\ln v}{\ln u}\right)\right]\right\}$ $\Phi$ FDA Normal Estándar	$\frac{C(u, v; \theta)}{u \cdot v} \cdot \left\{\Phi(z_1) \cdot \Phi(z_2) - \left(\frac{\theta}{2 \ln v}\right) \cdot \phi(z_1)\right\}$ $z_1 = \frac{1}{\theta} + \frac{\theta}{2} \ln\left(\frac{\ln u}{\ln v}\right)$ ; $z_2 = \frac{1}{\theta} + \frac{\theta}{2} \ln\left(\frac{\ln v}{\ln u}\right)$ $\phi$ FDP Normal Estándar
t-VE	$\rho \in [-1, 1]$ $\nu \in (0, \infty)$ $t \in [0, 1]$	$t \cdot T_{\nu+1}(Z_t) + (1-t)T_{\nu+1}(Z_{1-t});$ $Z_t = \left\{\frac{[t/(1-t)]^{\nu} - \rho}{\sqrt{(1-\rho^2)/(v+1)}}\right\}$ $Z_{1-t} = \left\{\frac{[(1-t)/t]^{\nu} - \rho}{\sqrt{(1-\rho^2)/(v+1)}}\right\}$	$\exp\{\ln(u \cdot v) \cdot A(t)\};$ $t = \frac{\ln v}{\ln(u \cdot v)}$ $T_{\nu+1}$ FDA t-Student Estándar	$\frac{C(u, v; \rho, \nu)}{u \cdot v} \cdot \left\{\frac{\partial A(x, y; \rho, \nu)}{\partial x} \cdot \frac{\partial A(x, y; \rho, \nu)}{\partial y} - \frac{\partial A(x, y; \rho, \nu)}{\partial x \partial y}\right\}$ $A(x, y; \rho, \nu) = x \cdot T_{\nu+1}(z_{xy}) + y \cdot T_{\nu+1}(z_{yx})$ $x = \ln u$ ; $y = \ln v$ <i>[consulte Joe (2015)]</i>

## 3.6. Medidas de Concordancia

Se ha dicho que la cópula es en sí misma una caracterización del tipo de dependencia de un par de variables aleatorias  $(X, Y)$ , por lo que tendría sentido usarlas para medir cual sería ese grado de dependencia. En este apartado se explica el papel que adoptan las cópulas en el estudio de la dependencia y cuál es la relación que guardan con los conocidos *coeficientes no paramétricos de concordancia, rho de Spearman y tau de Kendall*, ambos relacionados a una estructura de dependencia conocida como concordancia. La propiedad más importante que las dos medidas anteriormente citadas guardan, es que permanecen invariantes por medio de aplicaciones estrictamente crecientes que actúan sobre las variables aleatorias. Antes de profundizar en ello, se toma en cuenta lo siguiente:

- El término coeficiente de correlación es exclusivo para medidas de correlación lineal entre dos variables aleatorias, como por ejemplo el coeficiente de Pearson.
- El término medida de concordancia es usado para las medidas tales como la tau de Kendall y la rho de Spearman, las cuales "miden" una forma de dependencia conocida como concordancia.

### 3.6.1. Coeficiente de correlación de Pearson

El coeficiente de correlación lineal de Pearson es uno de los más utilizados por su facilidad de cálculo, por su naturalidad como una medida de dependencia en distribuciones normales multivariadas.

Sean  $X$  e  $Y$  variables aleatorias reales con  $0 < \text{var}(X) < \infty$  y  $0 < \text{var}(Y) < \infty$ , el coeficiente de correlación entre  $X$  e  $Y$  se define como:

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}} = \frac{E(X \cdot Y) - E(X)E(Y)}{\sqrt{\text{var}(X)}\sqrt{\text{var}(Y)}} \quad (3.35)$$

donde  $E$  es la esperanza matemática de  $X$ ,  $Y$  y  $XY$ .

No obstante, se hace necesario implementar otro tipo de indicadores de la dependencia debido a que este coeficiente de correlación presenta los siguientes inconvenientes:

- Las varianzas de  $X$  e  $Y$  deben ser finitas y positivas de lo contrario la correlación lineal no está definida.
- Independencia de dos variables aleatorias implica que no están correlacionadas, es decir  $\rho(X, Y) = 0$  pero una correlación de cero no implica en general independencia, pues esto ocurre solo en el caso de la distribución normal multivariada.
- La correlación lineal no es invariante bajo transformaciones no lineales estrictamente crecientes  $T: \mathbb{R} \rightarrow \mathbb{R}$  ya que para dos variables aleatorias de valor real se tiene en general:

$$\rho(T(X), T(Y)) \neq \rho(X, Y)$$

Debido a los problemas que presenta el coeficiente de correlación de Pearson, otras medidas de dependencia son utilizadas como los coeficientes rho de Spearman y tau de Kendall. Estas medidas se expresan mediante funciones cópulas.

### 3.6.2. Concepto de concordancia y discordancia

Supongamos que se tienen dos variables aleatorias  $X$  e  $Y$ , entonces si valores “grandes” (pequeños) de una variable tienden a estar asociados con valores “grandes” (pequeños) de la otra variable, entonces se dice que dichas variables son concordantes, por otro lado si valores “grandes” (pequeños) de una variable tienden a estar asociados con valores “pequeños” (grandes) de la otra variable, entonces se dice que estas variables son discordantes. A continuación se definen conceptos de concordancia y discordancia:

**Definición 3.25.** (Concordancia) Sean  $(x_i, y_i)$  y  $(x_j, y_j)$  dos observaciones de un vector de variables aleatorias continuas  $(X, Y)$  se dice que  $(x_i, y_i)$  y  $(x_j, y_j)$  son concordantes si  $(x_i < x_j$  y  $y_i < y_j)$  o  $(x_i > x_j$  y  $y_i > y_j)$ .



Una forma alterna de presentar la anterior definición es:  $(x_i, y_i)$  y  $(x_j, y_j)$  son *concordantes* sí  $(x_i - x_j)(y_i - y_j) > 0$ .

**Definición 3.26.** (Discordancia) Sean  $(x_i, y_i)$  y  $(x_j, y_j)$  dos observaciones de un vector de variables aleatorias continuas  $(X, Y)$  se dice que  $(x_i, y_i)$  y  $(x_j, y_j)$  son *discordantes* si  $(x_i < x_j$  y  $y_i > y_j)$  o  $(x_i > x_j$  y  $y_i < y_j)$ .

Una forma alterna de presentar la anterior definición es:  $(x_i, y_i)$  y  $(x_j, y_j)$  son discordantes sí  $(x_i - x_j)(y_i - y_j) < 0$ .

### 3.6.3. Medidas de correlación de rango

A continuación se discuten dos importantes medidas de dependencia conocidas como  $\rho$  de Spearman y  $\tau$  de Kendall. Estos coeficientes de correlación de rango miden el grado de dependencia monótona entre dos variables aleatorias, además ofrecen una medida de correlación para variables que no pertenecen a la familia de distribuciones elípticas. Estas dos medidas son de gran importancia para la teoría de cópulas, a partir de éstas se puede estimar el parámetro  $\theta$  el cual mide la estructura de dependencia entre variables aleatorias.

#### 3.6.3.1. Tau de Kendall

Este coeficiente está íntimamente relacionado con los conceptos de concordancia/discordancia.

**Definición 3.27.** (Tau de Kendall muestral) Sea una muestra aleatoria  $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$  de  $n$  observaciones de un vector aleatorio continuo  $(X, Y)$ ; se tienen  $\binom{n}{2}$  pares distintos de  $(x_i, y_i)$  y  $(x_j, y_j)$  de observaciones de la muestra. Sea  $c$  el número de pares concordantes y  $d$  el número de pares discordantes. Se define la tau de Kendall como [Nelsen (2006)]

$$\tau = \frac{c - d}{c + d} = \frac{c - d}{\binom{n}{2}} \quad (3.36)$$

Equivalentemente,  $\tau$  es la probabilidad estimada de la concordancia menos la probabilidad estimada de la discordancia para un par de observaciones  $(x_i, y_i)$  y  $(x_j, y_j)$ , escogidas aleatoriamente de la muestra.

La versión poblacional del tau de Kendall para un vector  $(X, Y)$  de variables aleatorias continuas con función de distribución conjunta  $H_{XY}(x, y)$  se define de la siguiente forma.

**Definición 3.28.** (Tau de Kendall poblacional) Sean  $(X_1, Y_1)$  y  $(X_2, Y_2)$  vectores aleatorios independientes e idénticamente distribuidos con función de distribución conjunta  $H_{XY}(x, y)$  igual para ambos. La tau de Kendall para el vector  $(X, Y)$  está definida como la probabilidad de la concordancia menos la probabilidad de la discordancia, es decir

$$\tau = \tau_{X,Y} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]. \quad (3.37)$$

**Definición 3.29.** Sean  $(X_1, Y_1)$  y  $(X_2, Y_2)$  dos vectores aleatorios continuos con diferentes funciones de distribución conjunta  $H_{1_{XY}}$  y  $H_{2_{XY}}$  respectivamente pero con marginales comunes  $F_X$  para  $X_1$  y  $X_2$ , y  $G_Y$  para  $Y_1$  e  $Y_2$ . Definimos la función de concordancia  $Q$  como la probabilidad de la concordancia menos la probabilidad de la discordancia de los dos vectores  $(X_1, Y_1)$  y  $(X_2, Y_2)$ , es decir

$$Q = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]. \quad (3.38)$$

**Teorema 3.30.** Sean  $(X_1, Y_1)$  y  $(X_2, Y_2)$  vectores aleatorios continuos e independientes con función de distribución conjunta  $H_{1_{XY}}$  y  $H_{2_{XY}}$  respectivamente, y marginales comunes  $F_X$  (de  $X_1$  y  $X_2$ ) y  $G_Y$  (de  $Y_1$  y  $Y_2$ ). Sean  $C_1$  y  $C_2$  las cópulas de  $(X_1, Y_1)$  y  $(X_2, Y_2)$  respectivamente, así

que, por el teorema de Sklar se tiene que  $H_{1_{XY}}(x_1, y_1) = C_1(F_X(x_1), G_Y(y_1))$  y  $H_{2_{XY}}(x_2, y_2) = C_2(F_X(x_2), G_Y(y_2))$ . Sea  $Q$  la diferencia entre la probabilidades de concordancia y discordancia entre  $(X_1, Y_1)$  y  $(X_2, Y_2)$ , Entonces la versión poblacional de la tau de Kendall para  $X$  e  $Y$  viene dada por la expresión

$$Q = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0].$$

de donde

$$\tau = \tau_{X,Y} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0] =$$

$$\tau_{X,Y} = 4 \iint_{I^2} C_2(u, v) dC_1(u, v) - 1. \quad (3.39)$$

Aunque, la fórmula del Teorema 3.30 no siempre es fácil de calcular, ya que  $C$  puede ser singular o tener una componente singular. El resultado que a continuación se presenta, brinda una forma más fácil de calcular la tau de Kendall para estos casos.

**Teorema 3.31.** *Sea  $C$  una cópula tal que el producto  $(\partial C/\partial u)(\partial C/\partial v)$  es integrable en  $I^2$ , entonces [Nelsen (2006)]:*

$$\iint_{I^2} C(u, v) dC(u, v) = \frac{1}{2} - \iint_{I^2} \frac{\partial}{\partial u} C(u, v) \frac{\partial}{\partial v} C(u, v) dudv. \quad (3.40)$$

**Corolario 3.32.** *Sean  $X$  e  $Y$  variables aleatorias con cópula Arquimediana  $C$ , generada por  $\varphi$ , la versión poblacional de la tau de Kendall  $\tau_C$  para  $X$  e  $Y$ , está dada por*

$$\tau_C = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt. \quad (3.41)$$

Para la demostración de este Corolario consulte el libro de Nelsen (2006).

**Ejemplo 2:** Sea  $C_\theta$  una cópula Arquimediana de la familia de cópulas Gumbel-Hougaard, con  $\theta \geq 1$

$$C_{\theta}(u, v) = \exp \left\{ - \left[ (-\ln u)^{\theta} + (-\ln v)^{\theta} \right]^{1/\theta} \right\},$$

la función generadora

$$\varphi(t) = (-\ln t)^{\theta},$$

la derivada del generador  $\varphi$

$$\varphi'(t) = -\frac{\theta}{t} (-\ln t)^{\theta-1},$$

por el Corolario 3.32 para calcular la tau de Kendall, se tiene que

$$\frac{\varphi(t)}{\varphi'(t)} = \frac{(-\ln t)^{\theta}}{-\frac{\theta}{t} (-\ln t)^{\theta-1}} = \frac{t \ln t}{\theta}.$$

si se integra por partes la expresión (3.41) y se toma  $u = \ln t$  y  $dv = tdt$

$$\begin{aligned} \tau_{\theta} &= 1 + 4 \int_0^1 \frac{t \ln t}{\theta} dt \\ &= 1 + \frac{4}{\theta} \left[ \left[ \frac{t^2}{2} \ln t \right]_0^1 - \int_0^1 \frac{t}{2} dt \right] \\ &= 1 + \frac{4}{\theta} \left[ -\frac{t^2}{4} \right]_0^1 = 1 - \frac{4}{\theta} \left[ \frac{1}{4} \right] \\ &= 1 - \frac{1}{\theta} \end{aligned}$$

### 3.6.3.2. Rho de Spearman

Esta cantidad está definida también a partir de la concordancia/discordancia, al igual que la tau de Kendall. Para obtener la versión poblacional del coeficiente rho de Spearman, introducimos  $(X_1, Y_1)$ ,  $(X_2, Y_2)$  y  $(X_3, Y_3)$  vectores aleatorios e independientes con distribución conjunta  $H_{XY}$  (cuyas marginales son de nuevo  $F_X$  y  $G_Y$  para

$X_i, 1 \leq i \leq 3, Y_j, 1 \leq j \leq 3$ ) y cuya cópula es  $C$ ; la versión poblacional de dicho coeficiente está caracterizada por ser proporcional a la probabilidad de la concordancia menos la probabilidad de la discordancia de dos vectores aleatorios  $(X_1, Y_1)$  y  $(X_2, Y_3)$ ,

$$\rho_{X,Y} = 3\left(P\left[(X_1 - X_2)(Y_1 - Y_3) > 0\right] - P\left[(X_1 - X_2)(Y_1 - Y_3) < 0\right]\right) \quad (3.42)$$

donde el 3 en la parte derecha de la igualdad actúa como constante de normalización.

Es importante resaltar que la función de distribución conjunta  $(X_2, Y_3)$  es  $F_X(x)G_Y(y)$ , puesto que  $X_2$  e  $Y_3$  son independientes (mientras que la distribución conjunta de  $(X_1, Y_1)$  es  $H_{XY}(x, y)$ ). Así la cópula de  $(X_2, Y_3)$  es la cópula producto.

**Teorema 3.33.** (Rho de Spearman poblacional a partir de la cópula) *Sean  $X$  e  $Y$  variables aleatorias continuas cuya cópula es  $C$ . Entonces la versión poblacional del coeficiente rho de Spearman para  $X$  e  $Y$  viene dada por*

$$\begin{aligned} \rho_{X,Y} &= 3\left(P\left[(X_1 - X_2)(Y_1 - Y_3) > 0\right] - P\left[(X_1 - X_2)(Y_1 - Y_3) < 0\right]\right) = \\ &= 3Q(C, \Pi) = 12 \iint_{I^2} uv dC(u, v) - 3 \\ \rho_{X,Y} &= 12 \iint_{I^2} C(u, v) dudv - 3 \end{aligned} \quad (3.43)$$

donde las dos últimas igualdades se obtienen si se toma en cuenta que  $Q$  es simétrica respecto a sus dos variables.

### 3.6.3.3. Relación entre el parámetro $\theta$ y las medidas de dependencia $\rho_s$ y $\tau_\theta$

Algunos resultados de las medidas de dependencia como la rho de Spearman  $\rho_s$ , tau de Kendall  $\tau_\theta$  y su relación que existe con el parámetro  $\theta$ , se presentan la Tabla 3.4. Para estimar las medidas de algunos modelos cópula, su forma es no cerrada, para dar solución a estos

casos es conveniente tener en cuenta las generalidades de la cópula bivariada de valores extremos [Joe (2015)], la cual se estableció en el Teorema 3.24 de este trabajo.

**Teorema 3.34.** (Coeficientes  $\rho_s$ ,  $\tau_\theta$ , para una cópula bivariada de valores extremos). *Para una cópula bivariada de valores extremos  $C(u, v) = \exp\{-A(x + y)\} = \exp\{-(x + y) \cdot B(x/(x + y))\}$  con  $x = -\ln u$ ,  $y = -\ln v$  y  $B(w) = A(w, 1 - w)$ , sus medidas de dependencia de acuerdo con Joe (2015), son*

$$\rho_s = 12 \int_0^1 [1 + B(w)]^{-2} dw - 3 \quad (3.44)$$

$$\tau_\theta = \int_0^1 \frac{(2w - 1)B'(w) + w(1 - w)[B'(w)]^2}{[B'(w)]^2} dw. \quad (3.45)$$

Por lo que respecta a la **Cópula de Gumbel**, para estimar la medida  $\rho_s$  se toma en cuenta que, para cópulas bivariadas de valores extremos, sus exponentes se consideran como [Joe (2015)]:

$$A = (x, y; \theta) = (x^\theta + y^\theta)^{1/\theta} \quad (3.46)$$

$$B(w; \theta) = A(w, 1 - w; \theta) = [w^\theta + (1 - w)^\theta]^{1/\theta} \quad (3.47)$$

de modo que

$$\rho_s = 12 \int_0^1 \left[ 1 + [w^\theta + (1 - w)^\theta]^{1/\theta} \right]^{-2} dw - 3 \quad (3.48)$$

En el caso de la **Cópula de Frank** se usa la función de Debye  $D_k(x)$  para ambas medidas de dependencia ( $\rho_s$  y  $\tau_\theta$ ) [Genest (1987) y Nelsen (1986)], y está definida como

$$D_k(x) = \frac{k}{x^k} \int_0^x \frac{t^k}{e^t - 1} dt. \quad (3.49)$$

para  $k=1, 2$

La estimación del parámetro  $\theta$  utilizando medidas de dependencia  $\rho_s$  y  $\tau_\theta$ , se puede consultar en Genest (1987).

Nelsen (2006), ha proporcionado una fórmula para estimar la medida  $\rho_s$  de Spearman para la *Cópula de Ali-Mikhail-Haq*, como

$$\rho_s = \frac{12(1+\theta)}{\theta^2} \cdot \text{dilog}(1-\theta) - \frac{24(1-\theta)}{\theta^2} \cdot \log(1-\theta) - \frac{3(\theta+12)}{\theta}, \quad (3.50)$$

donde  $\text{dilog}(x) = Li_2(1-x) = \text{polilog}(1-x, 2)$ , y  $Li_2(x)$  es la definición habitual del "dilogaritmo" (también llamada "Función de Spence").

$$Li_2(z) = -\int_0^z \frac{\ln(1-u)}{u} du = \sum_{k=1}^{\infty} \frac{z^k}{k^2}, \quad z \in \mathbb{C} \setminus [1, \infty), \quad (3.51)$$

donde la suma infinita solo es aplicable para  $|z| < 1$ .

Con límites para  $\theta \in [-1, 1)$ , esto conduce a un rango de la medida  $\rho_s$  en el intervalo  $\rho_s \in [33 - 48 \ln(2), 4\pi^2 - 39] \cong [-0.2711, 0.4784]$ , lo que comprende una variación pequeña de dependencia.

Para calcular  $\rho_s$  con precisión para todos los valores de  $\theta$ , se desarrolla la serie de Taylor en los términos respectivos de (3.51), por lo que se obtendrá una fórmula con una serie infinita para  $\rho_s$  [Mächler (2014)],

$$\rho_s = \sum_{k=1}^{\infty} \frac{3}{\binom{k+2}{2}} \cdot \theta^k = \frac{\theta}{3} + \frac{\theta^2}{12} + \frac{3\theta^3}{100} + \frac{\theta^4}{75} + \dots \quad (3.52)$$

El rango de valores para la medida  $\tau$  de Kendall para esta cópula, oscila en el intervalo  $\tau_\theta \in \left[ \frac{5 - 8 \ln(2)}{3}, 1/3 \right] \cong [-0.1817, 0.3333]$ , de igual manera que  $\rho_s$ , se tiene una variación pequeña de dependencia [Mächler (2014)].

Para la **Cópula de Galambos**, la medida  $\rho_s$ , se consideran exponentes como [Joe (2015)]:

$$A(x, y; \theta) = x + y - (x^{-\theta} + y^{-\theta})^{-1/\theta}, \quad (3.53)$$

$$B(w; \theta) = A(w, 1 - w; \theta) = 1 - \left[ w^{-\theta} + (1 - w)^{-\theta} \right]^{-1/\theta}, \quad (3.54)$$

$$B'(w; \theta) = \left[ 1 + \left( \frac{1 - w}{w} \right)^\theta \right]^{-1/\theta - 1} - \left[ 1 + \left( \frac{w}{1 - w} \right)^\theta \right]^{-1/\theta - 1}, \quad (3.55)$$

de donde

$$\rho_s = 12 \int_0^1 \left[ 2 - \left[ w^{-\theta} + (1 - w)^{-\theta} \right]^{-1/\theta} \right]^{-2} dw - 3. \quad (3.56)$$

En lo que respecta a la **Cópula de Husler-Reiss**, sus exponentes están dados como [Joe (2015)]:

$$A(x, y; \theta) = x \Phi \left[ \frac{1}{\theta} + \frac{1}{2} \theta \ln \left( \frac{x}{y} \right) \right] + y \Phi \left[ \frac{1}{\theta} + \frac{1}{2} \theta \ln \left( \frac{y}{x} \right) \right], \quad (3.57)$$

$$B(w; \theta) = A(w, 1 - w; \theta), \quad (3.58)$$

$$B(w; \theta) = w \Phi \left[ \frac{1}{\theta} + \frac{1}{2} \theta \ln \left( \frac{w}{1 - w} \right) \right] + (1 - w) \Phi \left[ \frac{1}{\theta} + \frac{1}{2} \theta \ln \left( \frac{1 - w}{w} \right) \right]$$



$$B'(w; \theta) = \Phi \left[ \frac{1}{\theta} + \frac{1}{2} \theta \ln \left( \frac{w}{1-w} \right) \right] - \Phi \left[ \frac{1}{\theta} + \frac{1}{2} \theta \ln \left( \frac{1-w}{w} \right) \right]. \quad (3.59)$$

por lo tanto

$$\rho_s = 12 \int_0^1 \left\{ 1 + w \Phi \left[ \frac{1}{\theta} + \frac{1}{2} \theta \ln \left( \frac{w}{1-w} \right) \right] + (1-w) \Phi \left[ \frac{1}{\theta} + \frac{1}{2} \theta \ln \left( \frac{1-w}{w} \right) \right] \right\}^{-2} dw - 3 \quad (3.60)$$

donde  $\Phi_i(z)$  es la función de distribución de probabilidad normal estándar.

Según Joe (2015), para estimar la medida  $\tau_\theta$  de Kendall de la **Cópula Joe/B5**, toma en cuenta consideraciones relativamente complicadas, lo que en la práctica ingenieril para el desarrollo de proyectos significa una labor ardua en cuanto a tiempo, el autor antes mencionado propone lo siguiente

$$\tau_\theta = 1 + 2(2 - \theta)^{-1} \cdot \left[ \psi(2) - \psi \left( \frac{2}{\theta + 1} \right) \right] \quad (3.61)$$

con límite en  $1 - trigama(2)$  cuando  $\theta \rightarrow 2$ ,

siendo

$$digama(z) = \psi(z) = \frac{d}{dz} \ln \Gamma(z) \quad (3.62)$$

$$trigama(z) = \psi'(z) = \frac{d^2}{dz^2} \ln \Gamma(z) \quad (3.63)$$

En el *Apéndice A* de este trabajo, se expone el procedimiento para calcular el coeficiente de correlación de Spearman  $\rho_s$  basado en rangos, se analizan dos escenarios; una primera forma consiste en el análisis de una muestra cuyas observaciones no se repiten, en el segundo escenario la muestra presenta diversas repeticiones en sus observaciones, las cuales se identifican como "empates".

Una vez estimado el coeficiente  $\rho_s$ , se calcula el parámetro de asociación  $\theta$  para las distintas familias cópula. En caso de que  $\theta$  sea estimado numéricamente, hágalo a partir de simulaciones y proponga un conjunto de 10,000 pares de puntos (ejemplo; Joe y t-Student para obtener  $\rho_s$ ).

### 3.6.3.4. Relaciones entre Rho de Spearman y Tau de Kendall

La relación entre  $\rho_s$  y  $\tau_\theta$  varía de una familia a otra. En la Figura 3.3, se observa que existe relación uno a uno entre los coeficientes  $\rho_s$  y  $\tau_\theta$ . El siguiente teorema, escrito por Daniels (1950) resulta con desigualdades universales para estas medidas comentadas.

**Teorema 3.35.** *Sean  $X$  e  $Y$  variables aleatorias continuas, y sean  $\rho_s$  y  $\tau_\theta$  denotadas como rho de Spearman y tau de Kendall respectivamente, entonces*

$$-1 \leq 3\tau_\theta - 2\rho \leq 1. \quad (3.64)$$

**Teorema 3.36.** *Sean  $X$  e  $Y$  variables aleatorias continuas, y sean  $\rho_s$  y  $\tau_\theta$  denotadas como rho de Spearman y tau de Kendall respectivamente, entonces*

$$\frac{1 + \rho}{2} \geq \left( \frac{1 + \tau_\theta}{2} \right)^2 \quad (3.65)$$

y

$$\frac{1 - \rho}{2} \geq \left( \frac{1 - \tau_\theta}{2} \right)^2. \quad (3.66)$$

**Corolario 3.37.** *Sean  $X$  e  $Y$  variables aleatorias continuas, y sean  $\rho_s$  y  $\tau_\theta$  denotadas como rho de Spearman y tau de Kendall respectivamente, entonces*

$$\frac{3\tau_\theta - 1}{2} \leq \rho \leq \frac{1 + 2\tau_\theta - \tau_\theta^2}{2}, \quad \tau_\theta \geq 0, \quad (3.67)$$

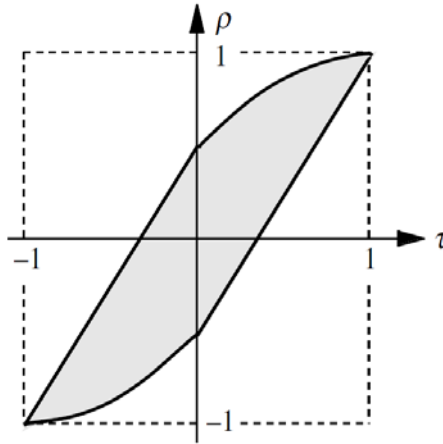
$$\frac{\tau_\theta^2 + 2\tau_\theta - 1}{2} \leq \rho \leq \frac{1 + 3\tau_\theta}{2}, \quad \tau_\theta \leq 0. \quad (3.68)$$

Los límites de los valores de  $\rho_s$  y  $\tau_\theta$  se ilustran en la Figura 3.3. Para cualquier par de variables aleatorias continuas  $X$  e  $Y$ , los valores de las versiones poblacionales de  $\rho_s$  y  $\tau_\theta$  deben estar en la región sombreada, referida como la región  $\rho_s - \tau_\theta$ .

**Tabla 3.4.** Familias de cópulas y sus medidas de dependencia.

<i>Familia</i>	<i>Parámetro</i> $\theta$	<i>Rho de Spearman</i> $\rho_s$	<i>Tau de Kendall</i> $\tau_\theta$
<i>Clayton</i>	$[0, \infty)$	<i>Forma complicada</i>	$\frac{\theta}{\theta + 2}$
<i>Frank</i>	$(-\infty, \infty) \setminus \{0\}$	$1 - \frac{12}{\theta} [D_2(\theta) - D_1(\theta)]$	$1 - \frac{4}{\theta} [D_1(\theta) - 1]$
<i>Ali-Mikhail-Haq</i>	$[-1, 1)$	$\sum_{k=1}^{\infty} \frac{3\theta^k}{\binom{k+2}{2}}$	$1 - \frac{2}{3} \left[ \frac{(1-\theta)^2 \cdot \ln(1-\theta) + \theta}{\theta^2} \right]$
<i>Joe/B5</i>	$[1, \infty)$	<i>Forma complicada</i>	$1 + 2(2-\theta)^{-1} \cdot \left[ \psi(2) - \psi\left(\frac{2}{\theta+1}\right) \right]$
<i>Gumbel</i>	$[1, \infty)$	$12 \int_0^1 \left[ 1 + [w^\theta + (1-w)^\theta]^{1/\theta} \right]^{-2} dw - 3$	$1 - \frac{1}{\theta}$
<i>Galambos</i>	$[0, \infty)$	$12 \int_0^1 \left[ 2 - [w^{-\theta} + (1-w)^{-\theta}]^{-1/\theta} \right]^{-2} dw - 3$	$\int_0^1 \frac{(2w-1)B'(w) + w(1-w)[B'(w)]^2}{[B'(w)]^2} dw$
<i>Husler-Reiss</i>	$\theta \geq 0$	$12 \int_0^1 \left\{ \begin{array}{l} 1 + w\Phi\left[\frac{1}{\theta} + \frac{1}{2}\theta \ln\left(\frac{w}{1-w}\right)\right] \\ + (1-w)\Phi\left[\frac{1}{\theta} + \frac{1}{2}\theta \ln\left(\frac{1-w}{w}\right)\right] \end{array} \right\}^{-2} dw - 3$	$\int_0^1 \frac{(2w-1)B'(w) + w(1-w)[B'(w)]^2}{[B'(w)]^2} dw$
<i>Gaussiana</i>	$[-1, 1]$	$\frac{6}{\pi} \arcsen\left(\frac{\rho}{2}\right)$	$\frac{2}{\pi} \arcsen(\rho)$
<i>t-Student</i>	$[-1, 1]$	<i>Forma complicada</i>	$\frac{2}{\pi} \arcsen(\rho)$
<i>Plackett</i>	$[0, \infty)$	$\frac{\theta+1}{\theta-1} - \frac{2\theta \ln(\theta)}{(\theta-1)^2}$	<i>Sin forma cerrada</i>
<i>FGM</i>	$[-1, 1]$	$\frac{1}{3}\theta \quad ; \quad \rho_s \in \left[-\frac{1}{3}, \frac{1}{3}\right]$	$\frac{2}{9}\theta \quad ; \quad \tau_\theta \in \left[-\frac{2}{9}, \frac{2}{9}\right]$

En la Tabla 4.2, se muestran las primeras derivadas de la Función Generadora  $\phi'(\mathbf{t}; \theta)$  de las Cópula Arquimedianas, para estimar periodos de retorno Kendall.



**Figura. 3.3.** Cotras  $\rho_s$  y  $\tau_\theta$  de un par de variables aleatorias continuas con cópula  $C$ , visto como una relación 1 a 1 [Nelsen (2006)].

### 3.7. Prueba de Bondad de Ajuste para Cópulas

Las cópulas se han convertido en la actualidad en una herramienta de gran potencia para la modelación de datos multivariados, en los que la dependencia entre variables aleatorias existe y el supuesto de normalidad multivariada no se tiene. Dicho esto, el interés en modelar problemas multivariados que involucran variables dependientes se generaliza en hidrología, lo que convierte a ésta teoría en una forma conveniente para modelar la estructura de dependencia en distribuciones conjuntas de variables aleatorias (gasto, volumen, duración).

Sin embargo, en la práctica no existe un método estándar para seleccionar una cópula entre diversos posibles modelos cópula, por lo que, la elección de una cópula adecuada es uno de los grandes retos al que se enfrenta el investigador. En este trabajo se discute un método para seleccionar una cópula, propuesto por el Dr. José María González-Barrios Murguía, investigador del Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS-UNAM).

Uno de los problemas al que se enfrenta el modelador de fenómenos hidrológicos cuando se trabaja con funciones cópula, una vez que se ha detectado dependencia entre marginales, es el de elegir el modelo cópula que mejor se ajusta a los datos observados. Mediante la metodología que

se propone en este trabajo es posible seleccionar el modelo óptimo, determinando de la *distancia de variación total o distancia supremo* (versión alterna del teorema de *Glivenko-Cantelli*) entre la d-cópula muestral de orden  $m$   $C_n^{(m)}$  y la cópula verdadera  $C$  (modelos propuestos).

La metodología que se propone consiste en obtener tres copulas; Cópula muestral de orden  $m$   $C_n^{(m)}$  conformada por el conteo de puntos de la muestra de orden  $n$ , aproximación checkerboard conocida como cópula  $C^{(m)}$  la cual se entiende como una interpolación bilineal derivado del teorema de Sklar, y cópula real o verdadera  $C$  que está restringida a determinados modelos propuestos. El propósito es comparar dos cópulas, mediante la distancia supremo, estas son  $C_n^{(m)}$  y  $C^{(m)}$ , dicho procedimiento se realiza comparando todos los modelos propuestos. Para la elección del mejor modelo se identifica la mínima distancia supremo entre los modelos propuestos y la cópula muestral de orden  $m$   $C_n^{(m)}$ .

### 3.7.1. Definiciones Básicas y Resultados Preliminares

La prueba del Teorema de Sklar (Teorema 3.7 de este trabajo) con dimensión  $d=2$ , es posible consultarlo en Nelsen (2006), consiste en la formulación de dos Lemas (*Lema 3.3.5*, p. 18 y *Lema 3.3.5*, p. 19) y fácilmente se puede extender a dimensiones mayores. El segundo Lema, considerado el más contundente y mencionado más adelante, indica en términos generales que, cualquier d-subcópula  $C'$  puede extenderse a una d-copula  $C$ . La prueba del segundo Lema se realiza mediante una simple interpolación d-lineal, que en otras palabras, este procedimiento proporciona una masa de probabilidad la cual se distribuye uniformemente en cada una de las d-cajas que conforman un grid, con vértices en el Dominio ( $C'$ ).

A continuación se obtendrá una cópula  $C$  bajo el sustento del segundo Lema, enunciado en Nelsen (2006), mediante una interpolación bilineal y la generación de un grid.

Sea  $m \geq 1$  un número entero y  $C$  una d-cópula, entonces en primera instancia, definimos un grid uniforme de tamaño  $m$  de  $[0,1]$  dado por

$J_m = \{0, 1/m, \dots, m/m\}$ . De modo que  $C$  está restringida al grid  $J_m$  y conforme a la Definición 3.1, se observa que  $\left\{C(\bar{i}) \mid \bar{i} \in J_m^d\right\}$  es una  $d$ -subcópula con dominio en  $J_m^d$ .

Por otra parte, sea  $\lambda^d$  la medida de Lebesgue en el espacio Borel  $(I^d, \mathcal{B}(I^d))$  y conforme a Nelsen (2006, Lema 2.3.5, p. 19), toma en cuenta el siguiente criterio de interpolación bilineal para hallar la cópula  $C(a, b)$ .

***Dada cualquier  $d$ -subcópula  $C'$ , existe una  $d$ -cópula  $C$  tal que  $C(u_1, \dots, u_d) = C'(u_1, \dots, u_d)$  para cada  $(u_1, \dots, u_d) \in \text{Dom}(C')$ , es decir, cualquier  $d$ -subcópula  $C'$  puede extenderse a una  $d$ -cópula  $C$ .***

Para  $(a, b) \in [0, 1]^2$  y una subcópula  $C'$  con dominio finito del conjunto  $S_1 \times S_2$ , sean  $a_1$  y  $a_2$  el mayor y el menor elemento de  $S_1$  que satisfaga  $a_1 \leq a \leq a_2$ , y sea  $b_1$  y  $b_2$  el mayor y menor elemento de  $S_2$  que satisfaga  $b_1 \leq b \leq b_2$ . Nótese que si  $a \in S_1$ , entonces  $a_1 = a = a_2$ , y si  $b \in S_2$ , entonces  $b_1 = b = b_2$  (obsérvese la Figura 3.4), por lo que se tiene

$$\lambda_1(a, b) = \begin{cases} (a - a_1)/(a_2 - a_1), & \text{si } a_1 < a_2 \\ 1, & \text{si } a_1 = a_2 \end{cases} \quad (3.69)$$

y

$$\mu_1(a, b) = \begin{cases} (b - b_1)/(b_2 - b_1), & \text{si } b_1 < b_2 \\ 1, & \text{si } b_1 = b_2 \end{cases} \quad (3.70)$$

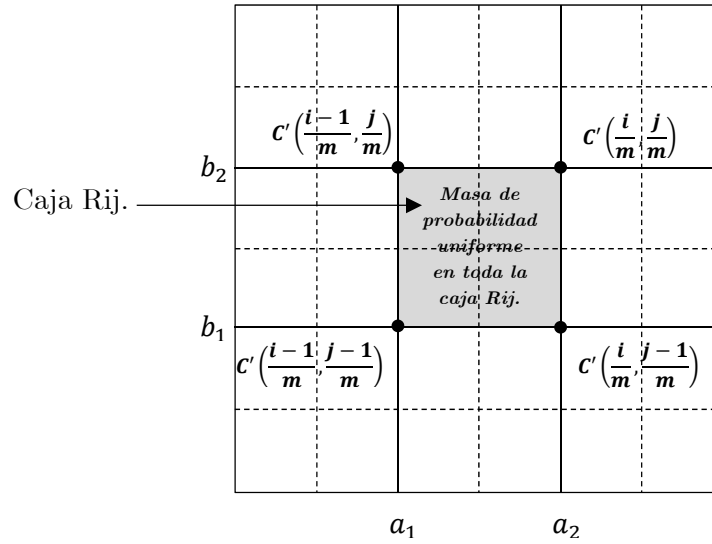
de donde

$$a_1 = \frac{i-1}{m}; \quad a_2 = \frac{i}{m}; \quad b_1 = \frac{j-1}{m}; \quad b_2 = \frac{j}{m}$$

$$\lambda_1 = \frac{a - \left(\frac{i-1}{m}\right)}{\frac{i}{m} - \left(\frac{i-1}{m}\right)} = \frac{a - \left(\frac{i-1}{m}\right)}{\frac{1}{m}}; \quad \mu_1 = \frac{b - \left(\frac{j-1}{m}\right)}{\frac{j}{m} - \left(\frac{j-1}{m}\right)} = \frac{b - \left(\frac{j-1}{m}\right)}{\frac{1}{m}} \quad (3.71)$$

y resulta la cópula

$$\begin{aligned}
 C(a, b) &= (1 - \lambda_1(a, b))(1 - \mu_1(a, b))C'(a_1, b_1) \\
 &= + (1 - \lambda_1(a, b))\mu_1(a, b)C'(a_1, b_2) \\
 &= + \lambda_1(a, b)(1 - \mu_1(a, b))C'(a_2, b_1) \\
 &= + \lambda_1(a, b)\mu_1(a, b)C'(a_2, b_2).
 \end{aligned} \tag{3.72}$$



**Figura 3.4.** Interpolación bilineal para hallar la cópula  $C(a, b)$

Los siguientes resultados que se muestran a continuación (3.38-3.44), son citados en el artículo de González-Barrios y Hoyos (2019).

**Definición 3.38.** (Aproximación Checkerboard) *La cópula  $C(a, b)$  definida en 3.72 es llamada **Aproximación Checkerboard de orden  $m$**  de la cópula  $C'$ . Esta aproximación distribuye uniformemente una masa de probabilidad en cada una de las 2-cajas representadas como  $R_{i,j} = \langle (i-1)/m, i/m \rangle \times \langle (j-1)/m, j/m \rangle$ , para cada  $i, j \in \{1, \dots, m\}$ . Dicha definición se puede extender a dimensiones mayores para  $d$ -cópulas  $C$ .*

**Definición 3.39.** (Función de Rango) *Sea  $X_1, \dots, X_n$  una muestra aleatoria de tamaño  $n$  de una variable aleatoria continua  $X$  y  $X_{(1)}, \dots, X_{(n)}$  sus estadísticas de orden. En este sentido, se define la función de rango  $R: \{1, 2, \dots, n\} \times \mathbb{R}^n \rightarrow \{1, 2, \dots, n\}$  como*

$$R(j, X_1, \dots, X_n) = k, \text{ si y solo si } X_j = X_{(k)} \text{ donde } j, k \in \{1, \dots, n\}. \quad (3.73)$$

**Definición 3.40.** (Muestra Modificada o Pseudomuestra) Sea  $\underline{X}_1, \dots, \underline{X}_n$  una muestra aleatoria de tamaño  $n$  de un vector aleatorio  $\underline{X}$  de dimensión  $d$ , donde  $\underline{X}_i = (X_{i,1}, \dots, X_{i,d}) \in \mathbb{R}^d$ , para todo  $i = 1, \dots, n$ . Sea  $i \in I_n$ , la  $i$ -ésima muestra modificada  $\underline{Y}_i = (Y_{i,1}, \dots, Y_{i,d}) \in [0, 1]^d$ , la cual se define como

$$Y_{i,j} = \frac{1}{n} R(i, X_{i,j}, \dots, X_{n,j}) \text{ para todo } j \in I_d. \quad (3.74)$$

Se observa que la muestra modificada (transformada por rangos)  $\underline{Y}_1, \dots, \underline{Y}_n$  es siempre un subconjunto de  $I^d$ .

**Definición 3.41.** (Cópula Empírica) Sea  $\underline{X}_1, \dots, \underline{X}_n$  una muestra aleatoria de tamaño  $n$  de un vector aleatorio  $\underline{X}$  de dimensión  $d$ , con distribución conjunta y continua  $H$ , donde  $\underline{X}_i = (X_{i,1}, \dots, X_{i,d}) \in \mathbb{R}^d$  para todo  $i = 1, \dots, n$ . Sea  $\underline{Y}_1, \dots, \underline{Y}_n$  la correspondiente muestra modificada. Entonces la cópula empírica denotada como  $C_n : I^d \rightarrow I$  se define como

$$C_n(u_1, \dots, u_d) = \frac{1}{n} \sum_{i=1}^n 1_{\{Y_{i,1} \leq u_1, \dots, Y_{i,d} \leq u_d\}}(u_1, \dots, u_d), \quad (3.75)$$

para todo  $(u_1, \dots, u_d) \in I^d$ .

**Lema 3.42.** Sea  $H$  una distribución conjunta y continua  $d$ -dimensional,  $d \geq 2$ , y sea  $C$  la única  $d$ -copula dada en la ecuación (3.7) del Teorema de Sklar. Sea  $\underline{X}_1, \dots, \underline{X}_n$  una muestra aleatoria de tamaño  $n$  de un vector aleatorio  $\underline{X}$  de dimensión  $d$ , con distribución conjunta y continua  $H$ , y sea  $\underline{Y}_1, \dots, \underline{Y}_n$  la correspondiente muestra modificada. Entonces la cópula empírica se define como en (3.75) de la siguiente manera

$$\sup_{(x_1, x_2, \dots, x_d) \in I^d} |C(x_1, x_2, \dots, x_d) - C_n(x_1, x_2, \dots, x_d)| \geq \frac{1}{n} \text{ a.s.}[Pc], \quad (3.76)$$

donde  $Pc$  es la medida de probabilidad inducida por la cópula  $C$  en  $\mathbb{R}^d$ .

---



Un resultado muy importante acerca de la cópula empírica es el Teorema de Glivenko-Cantelli, que establece que la cópula empírica  $C_n$  se aproxima a la cópula real o verdadera  $C$  por medio de su distancia supremo.

**Teorema 3.43.** (Glivenko-Cantelli). *Sea  $C_n$  la cópula empírica construida a partir de una muestra de tamaño  $n$  de una distribución conjunta continua  $H$  con  $d$ -copula  $C$ , entonces*

$$\lim_{n \rightarrow \infty} \sup_{(u_1, u_2, \dots, u_d) \in I^d} |C_n(u_1, u_2, \dots, u_d) - C(u_1, u_2, \dots, u_d)| = 0 \quad a.s., \quad (3.77)$$

### 3.7.2. Aproximación Checkerboard de la cópula $C_\theta$

La aproximación checkerboard, denotada como  $C^{(m)}$ , es simplemente una interpolación bilineal utilizada en la prueba del teorema de Sklar, describe la convergencia de la cópula  $C^{(m)}$  y la cópula del modelo  $C_\theta$  (verdadera cópula, proviene de las  $r$  familias propuestas),  $C^{(m)}$  resulta ser una excelente aproximación de la cópula del modelo  $C_\theta$ , aun para valores pequeños de  $m$ . De modo que, sí  $m$  es relativamente grande,  $C^{(m)}$  es una muy buena aproximación de la cópula del modelo  $C_\theta$ . A continuación se muestra un resultado contundente para obtener la cópula  $C^{(m)}(u, v)$  por medio de una interpolación bilineal.

De acuerdo con González-Barrios y Hoyos (2019), para  $d = 2$  y conforme a la notación de Nelsen (2006, Lema 2.3.5, p. 19). Para toda subcópula  $S$  con dominio finito del conjunto  $S_1 \times S_2$ , donde  $S_1 = S_2 = \{0, 1/m, 2/m, \dots, (m-1)/m, 1\}$ , y para toda  $(u, v) \in I^2 = [0, 1]^2$ , existen únicos  $i, j \in \{1, \dots, m\}$  tal que  $(u, v) \in R_{i,j} = \langle (i-1)/m, i/m \rangle \times \langle (j-1)/m, j/m \rangle$ , entonces se define la aproximación checkerboard de la cópula del modelo  $C_\theta$ , como

$$u_1 = \frac{i-1}{m}; \quad u_2 = \frac{i}{m}; \quad v_1 = \frac{j-1}{m}; \quad v_2 = \frac{j}{m}$$

$$\lambda_1(u, v) = \frac{u - \left(\frac{i-1}{m}\right)}{\frac{i}{m} - \left(\frac{i-1}{m}\right)} = \frac{u - \left(\frac{i-1}{m}\right)}{\frac{1}{m}};$$

(3.78)

$$\mu_1(u, v) = \frac{v - \left(\frac{j-1}{m}\right)}{\frac{j}{m} - \left(\frac{j-1}{m}\right)} = \frac{v - \left(\frac{j-1}{m}\right)}{\frac{1}{m}},$$

Si  $n$  es múltiplo de  $m$ , entonces se considera la siguiente representación como la aproximación Checkerboard

$$\begin{aligned} C^{(m)}(u, v) = \sum_{j=1}^m \sum_{i=1}^m & \left[ 1_{\left(\frac{i-1}{m}, \frac{i}{m}\right] \times \left(\frac{j-1}{m}, \frac{j}{m}\right]}(u, v) \right. \\ & \cdot \left( (1 - \lambda_1(u, v))(1 - \mu_1(u, v)) C(u_1, v_1) \right. \\ & + (1 - \lambda_1(u, v)) \mu_1(u, v) C(u_1, v_2) \\ & + \lambda_1(u, v)(1 - \mu_1(u, v)) C(u_2, v_1) \\ & \left. \left. + \lambda_1(u, v) \mu_1(u, v) C(u_2, v_2) \right) \right] \end{aligned}$$

(3.79)

y la aproximación Checkerboard resulta

$$\begin{aligned} C^{(m)}(u, v) = \sum_{j=1}^m \sum_{i=1}^m & \left[ 1_{\left(\frac{i-1}{m}, \frac{i}{m}\right] \times \left(\frac{j-1}{m}, \frac{j}{m}\right]}(u, v) \right. \\ & \cdot \left( (1 - \lambda_1(u, v))(1 - \mu_1(u, v)) C\left(\frac{i-1}{m}, \frac{j-1}{m}\right) \right. \\ & + (1 - \lambda_1(u, v)) \mu_1(u, v) C\left(\frac{i-1}{m}, \frac{j}{m}\right) \\ & + \lambda_1(u, v)(1 - \mu_1(u, v)) C\left(\frac{i}{m}, \frac{j-1}{m}\right) \\ & \left. \left. + \lambda_1(u, v) \mu_1(u, v) C\left(\frac{i}{m}, \frac{j}{m}\right) \right) \right] \end{aligned}$$

(3.80)

si  $(u, v) \in R_{i,j}$ , para todo  $i, j \in I_m$

**La cópula  $C^{(m)}$  definida anteriormente se le llama aproximación checkerboard de la cópula del modelo  $C_\theta$ .** Esta aproximación otorga una masa de probabilidad, que a su vez, se distribuye uniformemente en cada una de las superficies de las d-cajas (grid) representadas como  $R_{i,j} = \left[\frac{i-1}{m}, \frac{i}{m}\right] \times \left[\frac{j-1}{m}, \frac{j}{m}\right]$  para todo  $i, j \in \{1, \dots, m\}$ .

Por lo tanto se dice que,  $C^{(m)}$  es producida de una interpolación bilineal, de modo que  $\lambda_1$  y  $\mu_1$  son lineales en  $u$  y  $v$  respectivamente, el dominio de  $C^{(m)}$  es  $I^2$  y si  $(u, v)$  está en el conjunto  $S_1 \times S_2$ , entonces  $\lambda_1 = \mu_1 = 1$  por lo que  $C^{(m)}(u, v) = C_\theta(u, v)$ , entonces se afirma que  $C^{(m)}$  es una cópula.

El siguiente lema establece una forma del teorema de *Glivenko-Cantelli*, es aplicado para determinar la distancia supremo (distancia de variación total) entre la aproximación checkerboard  $C^{(m)}$  y la cópula del modelo  $C_\theta$ .

**Lema 3.44.** *Sea  $X$  un vector aleatorio con cópula  $C$ . Sea  $m \geq 1$  y  $C^{(m)}$  es una cópula obtenida con la aproximación checkerboard y se aproxima a la cópula verdadera  $C$ , por lo tanto,  $C^{(m)}$  no es aleatoria. Entonces, con  $d \geq 2$  y para cada  $m \geq 1$  se obtiene la distancia supremo como [Cuberos, Masiello y Maume-Deschamps (2019)]*

$$\sup_{(u_1, u_2, \dots, u_d) \in I^d} \left| C^{(m)}(u_1, u_2, \dots, u_d) - C(u_1, u_2, \dots, u_d) \right| \leq \frac{d}{2m} \quad (3.81)$$

para el caso bivariado  $d = 2$  se tiene

$$\sup_{(u, v) \in I^d} \left| C^{(m)}(u, v) - C(u, v) \right| \leq \frac{1}{m} \quad (3.82)$$

por lo tanto,

$$\lim_{m \rightarrow \infty} \sup_{(u, v) \in I^2} \left| C^{(m)}(u, v) - C(u, v) \right| = 0 \quad (3.83)$$

Sí  $m \geq 2$ , y se considera la partición uniforme  $R_{ij}, i, j \in \{1, \dots, m\}$  de  $I^2 = [0, 1]^2$ , y dada una 2-cópula  $C$ , para todo  $i, j \in \{1, \dots, m\}$  se tiene

$$0 \leq V_{C^{(m)}}(R_{ij}) \leq \frac{1}{m}, \quad (3.84)$$

$$\sum_{j=1}^m \sum_{i=1}^m V_{C^{(m)}}(R_{i,j}) = 1, \quad (3.85)$$

y la aproximación checkerboard  $C^{(m)}$  cumple con

$$C^{(m)}\left(\frac{i}{m}, \frac{j}{m}\right) = C_{\theta}\left(\frac{i}{m}, \frac{j}{m}\right);$$

por lo tanto

$$V_{C^{(m)}}(R_{ij}) = V_{C_{\theta}}(R_{ij}) \quad (3.86)$$

Como se comentó, la aproximación checkerboard otorga una masa de probabilidad uniforme para cada una de las d-cajas  $R_{ij}$ , al ocurrir esto, se puede entender como la distribución volumen constante  $V_{C^{(m)}}(R_{ij})$  alrededor de la superficie de cada caja  $R_{ij}$ . Los volúmenes contenidos en las cajas  $R_{ij}$  de la cópula  $C^{(m)}$  son comparados con los volúmenes en las cajas  $R_{ij}$  de la cópula del modelo  $C_{\theta}$  (expresión 3.86).

### 3.7.3. Cópula Muestral de Orden $m$ $C_n^{(m)}$

En la teoría de probabilidad, las funciones llamadas cópulas son en realidad funciones de distribución que representan las relaciones de dependencia entre marginales uniformes.

En este apartado, se comentará las generalidades de la metodología para estimar la *d-cópula muestral de orden  $m$* , presentada por [González-Barrios y Hernández (2013)], y [González-Barrios y Hoyos (2019)].

La cópula muestral de orden  $m$ , denotada como  $C_n^{(m)}$ , es una d-cópula basada en una muestra de tamaño  $n$ , y es un estimador de la aproximación checkerboard conocida como cópula  $C^{(m)}$ , ésta última

resulta ser una excelente aproximación de la verdadera cópula  $C$  (restringida a determinados modelos cópula propuestos), aun para valores pequeños de  $m$ .

El parámetro  $m$ , es un número entero en  $2 \leq m \leq n$ , siendo  $n$  el tamaño de la muestra. El estimador  $C_n^{(m)}$  se aproxima rápidamente a  $C^{(m)}$  cuando aumenta el tamaño de la muestra  $n$ , para cada  $m \geq 2$ . Si  $m$  es relativamente grande, entonces  $C^{(m)}$  es una muy buena aproximación de  $C$ .

La cópula muestral de orden  $m$   $C_n^{(m)}$ , proviene de una muestra bivariada de tamaño  $n$  de datos observados, que al ser transformada por rangos, es decir modificada y mostrados en un plano bidimensional de coordenadas  $(u, v)$  además de considerar un grid uniforme de tamaño  $m \geq 2$  en cada una de sus cajas, forma particiones uniformes  $R_{ij}$  en  $I^2 = [0, 1]^2$  sobre el mismo plano, entonces, *la cópula  $C_n^{(m)}$  está conformada mediante el conteo de puntos que se localizan en cada una de las 2-cajas  $R_{ij}$* , dicho conteo se convierte en una masa de probabilidad, que a su vez, se distribuye uniformemente en cada una de las superficies de las 2-cajas  $R_{ij}$  del grid, para finalmente conformar un volumen  $V_{C_n^{(m)}}$  en cada  $R_{ij}$ . Posterior a ello, se determina la distancia supremo entre los volúmenes de la aproximación checkerboard  $C^{(m)}$  y la cópula de muestral  $C_n^{(m)}$  para  $m \geq 2$ .

La evaluación de la cópula muestral  $C_n^{(m)}$  es bastante simple y computacionalmente es más eficiente en cuanto a la capacidad de memoria que la cópula empírica, que usualmente se empleaba.

### 3.7.3.1. Generación del Grid de tamaño $m$

Se propone un grid de tamaño uniforme  $m \geq 2$  en cada una de sus 2-cajas, para adecuar una partición uniforme  $R_{ij}, i, j \in \{1, \dots, m\}$ , de  $I^2 = [0, 1]^2$ . Cada caja  $R_{ij}$  contara con un área de  $1/m^2$ , siendo  $m^2$  el número de cajas totales que conforman el grid.

El parámetro  $m$ , es un número entero en  $2 \leq m \leq n$ , donde  $n$  es el tamaño de la muestra. La cantidad de puntos (masa de probabilidad) que se localizan en cada caja  $R_{ij}$ , se refiere al volumen que existe de manera uniforme en cada caja  $R_{ij}$ . El volumen total de cada fila o columna del Grid, dado por  $\{0, 1/m, 2/m, \dots, (m-1)/m, m/m = 1\}$  es  $V_C = 1/m$  (achurado vertical para columnas y horizontal para filas), véase Figura 3.5.

De tal forma, el volumen total del Grid dada una partición uniforme es

$$\sum_{j=1}^m \sum_{i=1}^m V_C(R_{i,j}) = 1 \quad (3.87)$$

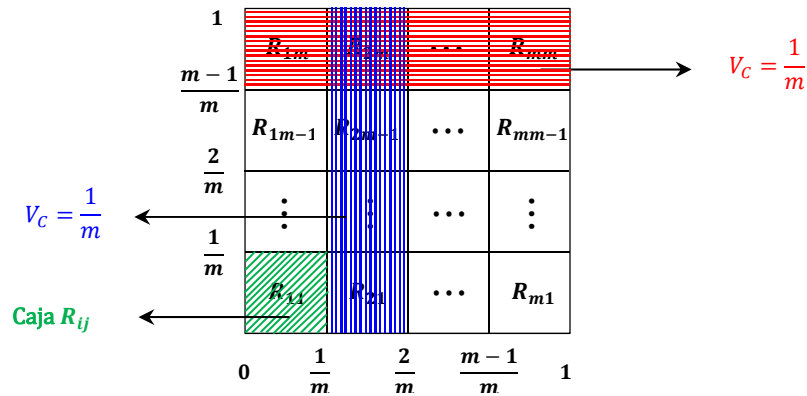


Figura 3.5. Grid de tamaño  $2 \leq m \leq n$  y partición uniforme en  $I^2 = [0,1]^2$

### 3.7.3.2. Conteo de puntos observados en las 2-cajas representadas por $R_{ij}$

Una vez generado el Grid de tamaño  $m \geq 2$ , además de contar con una muestra bivariada de datos observados de tamaño  $n$  y transformados por rangos, lo cual se convierte en una muestra modificada, se procede a graficarlos en un plano bidimensional de coordenadas  $(u, v)$ . Para generar la cópula muestral de orden  $m$   $C_n^{(m)}$ , es preciso realizar el conteo de puntos que se localizan en cada una de las 2-cajas, dicho conteo se convierte en una masa de probabilidad, que a su vez, se distribuye uniformemente en cada una de las superficies de las 2-cajas  $R_{ij}$  del grid, para finalmente conformar un volumen  $V_{C_n^{(m)}}$  en cada  $R_{ij}$ .

**Definición 3.45.** (Cópula muestral de orden  $m$ ). Sea  $2 \leq m \leq n$  y  $\underline{X}_1, \dots, \underline{X}_n$  una muestra aleatoria de tamaño  $n$  de un vector aleatorio  $\underline{X}$  de dimensión  $d$ , con distribución conjunta y continua  $H$  llamada  $d$ -cópula  $C$ , donde  $\underline{X}_i = (X_{i,1}, \dots, X_{i,d}) \in \mathbb{R}^d$ , para todo  $i = 1, \dots, n$ . Sea  $U_n = \{\underline{Y}_1, \dots, \underline{Y}_n\}$  la correspondiente muestra modificada.

La partición uniforme de tamaño  $m$  de  $I^d$ , donde para cada  $\bar{i} = (i_1, \dots, i_d) \in I_m^d$ , se define como

$$R_{\bar{i}}^m = \left\langle \frac{i_1 - 1}{m}, \frac{i_1}{m} \right\rangle \times \dots \times \left\langle \frac{i_d - 1}{m}, \frac{i_d}{m} \right\rangle. \quad (3.88)$$

La frecuencia relativa de los **puntos** en  $R_{\bar{i}}^m$  (**conteos**) se define por

$$s_{i_1, \dots, i_d}^{n, (m)} = \frac{\text{card}(R_{\bar{i}}^m \cap U_n)}{n}, \quad (3.89)$$

donde  $\text{card}(\cdot)$  denota la cardinalidad de un conjunto. Y sea

$$S_m^n = \left( s_{i_1, \dots, i_d}^{n, (m)} \right)_{(i_1, \dots, i_d) \in I_m^d}, \quad (3.90)$$

entonces  $S_m^n$  es siempre una matriz de transformación generalizada de dimensión  $d$ . Por medio de la matriz  $S_m^n$ , es posible realizar el conteo de puntos observados en las  $d$ -cajas representadas por  $R_{ij}$ .

*El conteo realizado en las  $d$ -cajas  $R_{ij}$ , la forma de interpretarlo es que la masa de probabilidad contenida en las diversas  $R_{ij}$  se distribuye de manera uniforme en la superficie de las  $d$ -cajas, lo que origina a obtener el volumen en las  $d$ -cajas  $R_{ij}$ .*

El siguiente lema establece que la suma de volúmenes obtenidos en cada renglón o columna del grid (resultado de los conteos previamente obtenidos), es igual al volumen  $V_{C_n^{(m)}}(R_{ij}) \leq 1/m$ , consulte la Figura 3.6.

**Lema 3.46.** Sea  $m \geq 2$ , además de considerar la partición uniforme  $R_{ij}, i, j \in \{1, \dots, m\}$  de  $I^2 = [0, 1]^2$ , y dada una 2-cópula  $C_n^{(m)}$ , para todo  $i, j \in \{1, \dots, m\}$  se define

$$P_{ij} = V_{C_n^{(m)}}(R_{ij}) \tag{3.91}$$

entonces se cumplen las siguientes relaciones

$$0 \leq \sum_{i=1}^{m-1} P_{ij} \leq \frac{1}{m} \quad (j \in \{1, \dots, m-1\}) \tag{3.92}$$

$$0 \leq \sum_{j=1}^{m-1} P_{ij} \leq \frac{1}{m} \quad (i \in \{1, \dots, m-1\}) \tag{3.93}$$

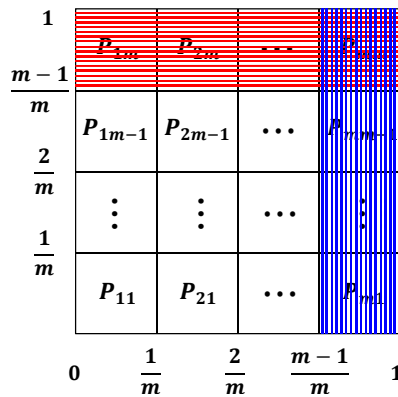
$$\frac{m-2}{m} \leq \sum_{i,j=1}^{m-1} P_{ij} \leq \frac{m-1}{m} \tag{3.94}$$

para todo  $i, j \in \{1, 2, \dots, m-1\}$

$$P_{im} = \frac{1}{m} - \sum_{j=1}^{m-1} p_{ij}, \quad P_{mj} = \frac{1}{m} - \sum_{i=1}^{m-1} p_{ij} \tag{3.95}$$

$$P_{mm} = \frac{1}{m} - \sum_{j=1}^{m-1} p_{mj} = \frac{1}{m} - \sum_{i=1}^{m-1} p_{im} \tag{3.96}$$

Como se ha dicho anteriormente, para seleccionar el modelo cópula que mejor se ajuste a los datos bivariados observados, se tiene que analizar para cada modelo, la distancia supremo entre la cópula muestral de orden  $m$   $C_n^{(m)}$  y la aproximación checkerboard  $C^{(m)}$  de la cópula de cada modelo  $C$ .



**Figura 3.6.** Suma de volúmenes  $V_{C_n^{(m)}}(R_{ij}) \leq 1/m$  posterior a los conteos.



### 3.7.4. Cópula del modelo $C_\theta$

En primera instancia, posterior a transformar por rangos  $[0,1]$  para modificar los datos observados de una muestra bivariada de tamaño  $n$ , se calcula el coeficiente de correlación rho de Spearman  $(\rho_s)$  o tau de Kendall  $(\tau_\theta)$  directamente de la muestra modificada u original (consulte el Apéndice A). De acuerdo con Nelsen (2006), existe una relación entre los coeficientes antes mencionados y el parámetro de asociación  $\theta$  para distintas familias cópula, conforme a lo anterior, es posible obtener el parámetro de la cópula  $\theta$  de forma analítica para algunos modelos, como por ejemplo las cópulas arquimedianas. Para algunas otras familias, la estimación del parámetro  $\theta$  se realiza numéricamente en función de los coeficientes  $(\rho_s)$  o  $(\tau_\theta)$ , a partir de simulaciones, proponiendo un conjunto de 10,000 pares de puntos (por ejemplo, copula Joe y t-Student).

Una vez estimado el parámetro  $\hat{\theta}$  de cada familia cópula en estudio, se construyen los  $r$  modelos cópula  $C_{\hat{\theta}}$  y se calcula el volumen  $V_{C_{\hat{\theta}}}(R_{ij})$  en las  $d$ -cajas para cada modelo. El volumen (masa de probabilidad) que está contenido en las cajas  $R_{ij}$ , se calcula obteniendo la copula en los cuatro puntos de las cajas  $R_{ij}$  (véase Figura 3.7):

sea la  $2$ -cópula del modelo

$$\left\{ C_{\hat{\theta}}\left(\frac{i}{m}, \frac{j}{m}\right) / i, j \in \{0, 1, 2, \dots, m\} \right\} \quad (3.97)$$

entonces el volumen bajo la cópula del modelo, expresado como el producto cartesiano es

$$V_{C_{\hat{\theta}}} = V_{C_{\hat{\theta}}}([u_1, u_2] \times [v_1, v_2]) = V_{C_{\hat{\theta}}}(R_{i,j}) \quad (3.98)$$

y el volumen bajo la cópula del modelo en los cuatro puntos es

$$V_{C_{\hat{\theta}}} = C_{\hat{\theta}}(u_2, v_2) - C_{\hat{\theta}}(u_2, v_1) - C_{\hat{\theta}}(u_1, v_2) + C_{\hat{\theta}}(u_1, v_1) \geq 0 \quad (3.99)$$

por definición de *2-volumen* de una subcópula, se tiene que el volumen es

$$V_{C_{\hat{\theta}}}(R_{i,j}) = C_{\hat{\theta}}\left(\frac{i}{m}, \frac{j}{m}\right) - C_{\hat{\theta}}\left(\frac{i}{m}, \frac{j-1}{m}\right) - C_{\hat{\theta}}\left(\frac{i-1}{m}, \frac{j}{m}\right) + C_{\hat{\theta}}\left(\frac{i-1}{m}, \frac{j-1}{m}\right). \quad (3.100)$$

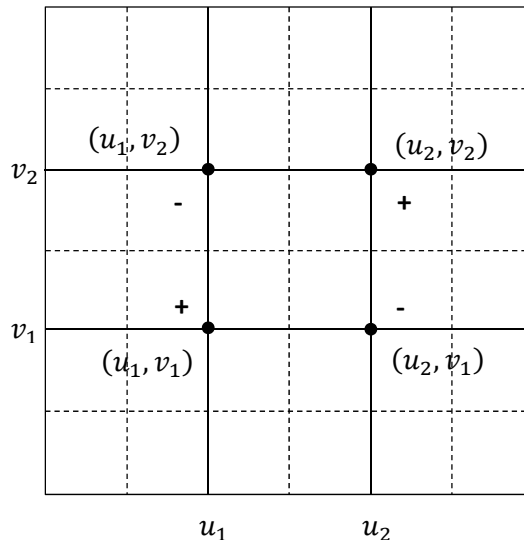
Sí  $m \geq 2$ , y se considera la partición uniforme  $R_{ij}, i, j \in \{1, \dots, m\}$  de  $I^2 = [0, 1]^2$ , y dada una *2-cópula*  $C$ , para todo  $i, j \in \{1, \dots, m\}$  se tiene que

$$0 \leq V_{C_{\hat{\theta}}}(R_{ij}) \leq \frac{1}{m}$$

$$\sum_{j=1}^m \sum_{i=1}^m V_{C_{\hat{\theta}}}(R_{ij}) = 1$$

por lo que

$$V_{C_{\hat{\theta}}}(R_{ij}) = V_{C^{(m)}}(R_{ij}). \quad (3.101)$$



**Figura 3.7.** Volumen de la caja  $R_{ij}$  (*masa de Probabilidad*).

*Cabe señalar que, las masas de probabilidad contenidas en las cajas  $R_{ij}$  de la cópula del modelo  $C_{\hat{\theta}}$  no están expandidas de manera uniforme.*

*Por otro parte, en la aproximación Checkerboard  $C^{(m)}$  de la cópula del modelo  $C_{\hat{\theta}}$ , las masas de probabilidad (volúmenes) que están contenidas*

---

en las cajas  $R_{ij}$ , sí son expandidas de manera uniforme al realizar la interpolación bilineal de  $C_{\hat{\theta}}$ .

### 3.7.5. Selección del mejor modelo cópula $C_{\theta}$

En la literatura hay diversas metodologías que se han propuesto para seleccionar un modelo que mejor se ajuste a datos bivariados, a la fecha no se conoce de alguno que sea más efectivo que otro.

El criterio de selección en principio, se basa en la comparación de *volúmenes* contenidos en las cajas  $R_{ij}$  entre la *cópula muestral*  $C_n^{(m)}$  y la *aproximación Checkerboard*  $C^{(m)}$  de los  $r$  modelos cópula  $C_{\hat{\theta}}$ , mediante la *distancia supremo* entre los volúmenes de las cópulas, se elige como mejor modelo quién posea la mínima distancia supremo (MDS), obtenida como  $D = \text{MDS}/2$ .

Con fines ilustrativos, en la Figura 3.8 se muestra un esquema con las generalidades de la metodología propuesta para la elección del mejor modelo cópula, y en la Tabla 7.15 se ha desarrollado un ejemplo aplicando esta metodología a los datos observados de la Estación Las Perlas, ubicada en la cuenca del Río Coatzacoalcos en Estado de Veracruz, México.

A continuación, se ha elaborado un resumen con las etapas más significativas para la selección del modelo cópula que mejor se ajusta a los datos bivariados observados.

1. Los datos observados de una *muestra bivariada* de tamaño  $n$ , se *transforman por rangos* de  $[0,1]$ , tome en cuenta que, en los desarrollos posteriores se trabajará con la muestra modificada, llamada también pseudo-muestra.
2. Se calcula el *coeficiente de correlación rho de Spearman*  $\rho_s$ , directamente de la pseudo-muestra bivariada (Apéndice A), o de la muestra de datos observados. En seguida se calcula el *parámetro de asociación*  $\theta$  para las distintas familias cópula seleccionadas

para el estudio (Tabla 3.4, relación  $\rho_s$  con  $\theta$ ). En caso que, el parámetro  $\theta$  sea estimado numéricamente, hágalo a partir de simulaciones y proponga un conjunto de 10,000 pares de puntos.

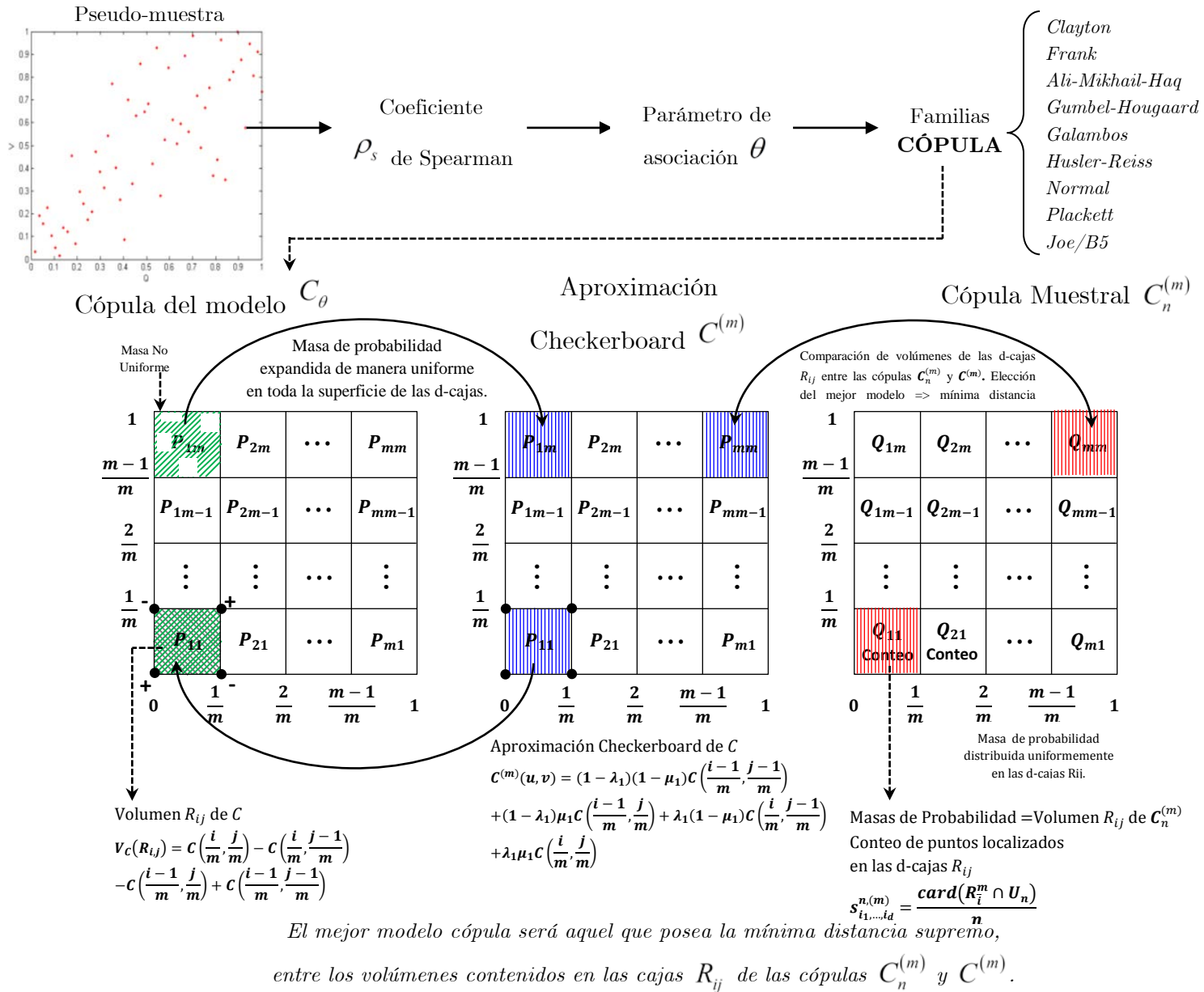
3. Se *propone un Grid de tamaño uniforme*  $m$  en  $I^2 = [0,1]^2$  con  $2 \leq m \leq n$  y partición uniforme  $R_{ij}$  en cada una de sus cajas. Siendo  $1/m^2$  el área en cada caja,  $m^2$  el número de cajas totales que conforman el grid y  $n$  el tamaño de la muestra.
4. Con base en la muestra bivariada de datos observados de tamaño  $n$  es posible obtener la **Cópula Muestral**  $C_n^{(m)}$ , para lograrlo, una vez definido el grid, el procedimiento continúa mediante la representación gráfica de valores de la muestra bivariada en un plano bidimensional, en seguida se realiza el conteo de puntos que se localizan en cada una de las cajas  $R_{ij}$ . *Por lo tanto, con el conteo de puntos, que visto como una masa de probabilidad (volumen) y expandida de manera uniforme en las cajas  $R_{ij}$ , es posible obtener los volúmenes de  $C_n^{(m)}$  en cada caja.*
5. Se calcula la **Aproximación Checkerboard**  $C^{(m)}$  de los  $r$  modelos cópula  $C_\theta$  que se pretenden estudiar, el análisis de aproximación se realiza mediante una interpolación bilineal en cada una de las cajas  $R_{ij}$  de la cópula del modelo  $C_\theta$ , hecho que conduce a obtener los respectivos volúmenes de  $C^{(m)}$ . Cabe señalar que, las masas de probabilidad (volúmenes) que están contenidas en las cajas  $R_{ij}$  de  $C^{(m)}$ , son expandidas de manera uniforme en toda la superficie de la caja  $R_{ij}$ .
6. Se realiza una **comparación de volúmenes** que están contenidos en las cajas  $R_{ij}$  de las cópulas siguientes: *Cópula Muestral*  $C_n^{(m)}$  y la *Aproximación Checkerboard*  $C^{(m)}$  de los  $r$  modelos cópula  $C_{\hat{\theta}}$  que se pretenden estudiar. Bajo este contexto, la comparación se realiza obteniendo la *distancia supremo* de todos los volúmenes contenidos

en las cajas  $R_{ij}$  entre las copulas  $C_n^{(m)}$  y  $C^{(m)}$ , por medio de la ecuación 3.83.

7. Finalmente, la *elección del mejor modelo*  $C_{\hat{\theta}}$ , será quien posea la mínima diferencia de volúmenes  $V_{C_n^{(m)}}(R_{ij})$  y  $V_{C^{(m)}}(R_{ij})$  respectivamente, o sea la *mínima distancia supremo (MDS)* y se calcula como  $D=MDS/2$ . En otras palabras, una vez que se ha *detectado* la mínima diferencia de volúmenes entre la cópula  $C_n^{(m)}$  y la aproximación  $C^{(m)}$  para todo el conjunto de familias, es evidente deducir que los volúmenes de la cópula del modelo  $C_{\hat{\theta}}$  vinculados con  $C^{(m)}$  posterior a la *detección*, poseen un grado importante de aproximación con los volúmenes de la cópula  $C^{(m)}$  elegida y  $C_n^{(m)}$ . Derivado de la explicación anterior, implica entonces seleccionar al mejor modelo cópula.

El concepto de mínima distancia supremo se entiende como la mejor aproximación del volumen contenido en las cajas  $R_{ij}$  entre la copula  $C_n^{(m)}$  y la aproximación  $C^{(m)}$  de todo el conjunto de familias en estudio, por lo tanto,  $C^{(m)}$  tiende a presentar volúmenes muy similares a los obtenidos en la cópula muestral  $C_n^{(m)}$ . Bajo este escenario, es evidente que la diferencia de volúmenes  $V_{C_n^{(m)}}(R_{ij})$  y  $V_{C^{(m)}}(R_{ij})$  resulta ser la mínima.

La metodología antes descrita está programada en lenguaje R, software que a la fecha está considerado de dominio público. El código fuente del programa TestCmvsC.R está disponible en el Apéndice C, fue proporcionado por el Dr. José María González-Barrios Murguía, investigador del IIMAS-UNAM. El programa es una segunda versión del programa TestH0SvsH1S-2Datos.R, al cual se ha contribuido agregando los modelos cópula siguientes: Galambos, Hüsler-Reiss y Galambos Decreasing-Decreasing (forma invertida del modelo Galambos).



**Figura 3.8.** Metodología para la Selección de un Modelo Cópula



# Capítulo 4

## Análisis de Frecuencia Bivariado de Avenidas: Un enfoque con Funciones Cópula

### 4.1. Modelos bivariados tradicionales

En hidrología, las distribuciones bivariadas más comúnmente aplicadas son: la normal, exponencial, gamma y de valores extremos [Yue et al. (1999)]. En la mayoría de los casos, el uso de una distribución normal bivariada no es apropiada para el modelado de avenidas, ya que las distribuciones marginales son asimétricas y de cola pesada [Favre (2012)]. Como se comentó en la introducción, los inconvenientes que conlleva el hacer uso de estas distribuciones bivariadas tradicionales para modelar la dependencia entre variables hidrológicas son que:

- i. Las variables aleatorias deben forzarse de tener la misma distribución marginal;
- ii. En principio de cuenta, las variables aleatorias no siguen una distribución normal en términos generales, a menos que se haga una transformación previa (Box-Cox);
- iii. Se utiliza los mismos parámetros de las distribuciones marginales para modelar la dependencia entre las 2 variables aleatorias [Favre



et al., 2004), en el caso del modelo bivariado de valores extremos, el parámetro de asociación que se emplea es de baja potencia, ya que está definido con un enfoque lineal, lo que significa subestimar o sobreestimar la dependencia entre variables.

Por otra parte, Salvadori y De Michele (2004) señalaron que los estudios realizados con distribuciones bivariadas se pueden facilitar mediante el uso de **Funciones Cópula**. Mediante este enfoque ya no es una restricción el hecho de que todas las variables que intervienen (gasto pico, volumen, duración, etc.) no tienen por qué seguir la misma función de distribución univariada.

A continuación se presenta algunos modelos bivariados para el análisis de dependencia entre variables aleatorias.

#### 4.1.1. Función de Distribución Normal Bivariada

La distribución normal 2-variada, también llamada distribución gaussiana bivariada, es una generalización de la distribución normal unidimensional con dimensión  $d = 2$ .

Un vector aleatorio  $X = [X_1, X_2]^T$  sigue una distribución normal bivariada con vector de medidas  $\mu = [\mu_1, \mu_2]^T$  y una matriz de covarianzas simétrica y definida positiva como

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \rho \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^2 \end{bmatrix}, \quad (4.1)$$

donde  $\sigma_x, \sigma_y \neq 0$  y  $|\rho| < 1$  el coeficiente de correlación entre  $x$  e  $y$ . Esta distribución es descrita con la notación  $X \sim N(\mu, \Sigma)$ .

Entonces, la *función densidad de probabilidad de la distribución normal bivariada*, se escribe de la siguiente forma:

$$f(x, y; \rho) = \frac{1}{2\pi\sqrt{\det \Sigma}} \exp\left[-\frac{1}{2}(X - \mu)\Sigma^{-1}(X - \mu)^T\right], \quad (4.2)$$

donde  $\det \Sigma$  es el determinante de la matriz  $\Sigma$  definido como

$$\det \Sigma = (\sigma_x \sigma_y)^2 (1 - \rho^2) \quad (4.3)$$

y la inversa de la matriz  $\Sigma$  está dada por

$$\Sigma^{-1} = \frac{1}{1 - \rho^2} \begin{bmatrix} 1/\sigma_x^2 & -\rho/(\sigma_x \sigma_y) \\ -\rho/(\sigma_x \sigma_y) & 1/\sigma_y^2 \end{bmatrix}, \quad (4.4)$$

si se sustituye (4.3) y (4.4) en (4.2), se tiene la **FDP de la distribución normal bivariada**, en forma explícita como

$$f(x, y; \rho) = \frac{1}{2\pi(\sigma_x \sigma_y)\sqrt{1 - \rho^2}} \exp \left\{ -\frac{1}{2(1 - \rho^2)} \left[ \left( \frac{x - \mu_x}{\sigma_x} \right)^2 - 2\rho \left( \frac{x - \mu_x}{\sigma_x} \right) \left( \frac{y - \mu_y}{\sigma_y} \right) + \left( \frac{y - \mu_y}{\sigma_y} \right)^2 \right] \right\}. \quad (4.5)$$

Para el caso bivariado, cuando la media es cero y la matriz de covarianza  $\Sigma$  es definida positiva con dimensión  $2 \times 2$  y cuyas características presentan elementos iguales al coeficiente de correlación  $\rho$  fuera de la diagonal, con  $\rho \in [-1, 1]$ , además de que los elementos posicionados en la diagonal principal son iguales a uno, con forme a ello, entonces la FDP de la Distribución Normal Bivariada que presenta estas características, es

denotada como  $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right)$ .

La matriz de covarianza  $\Sigma$ , su determinante  $\det \Sigma$  e inversa  $\Sigma^{-1}$  resultan

$$\begin{aligned} \Sigma &= \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \\ \det \Sigma &= 1 - \rho^2 \\ \Sigma^{-1} &= \frac{1}{1 - \rho^2} \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}, \end{aligned} \quad (4.6)$$

si se realiza la sustitución de las expresiones de (4.6) en (4.2), o de otro modo, si hacemos que  $\sigma_x = \sigma_y = 1$  y  $\mu_x = \mu_y = 0$  en (4.5) resulta lo siguiente para el caso bivariado:

$$f(x, y; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right]. \quad (4.7)$$

La expresión (4.7) representa la **FDP de la Distribución Normal Bivariada** con  $X \sim N(0, \Sigma)$ . Otra característica que presenta la FDA de (4.7), es que coincide con la Función Cópula Gaussiana (FDA), la cual se analizará en el Capítulo 3, expresión (3.12).

### 4.1.2. Funciones de Distribución Bivariada de Valores Extremos

Los artículos pioneros de Finkelstein (1953), Gumbel (1958) y Tiago de Oliveira (1958) dieron los fundamentos teóricos para el análisis de las distribuciones bivariadas de valores extremos.

Después de estos trabajos, se ha reportado el desarrollo y aplicación de la aproximación bivariada de valores extremos ajustada por máxima entropía y la extensión del modelo logístico trivariado, ajustado por máxima verosimilitud, a la solución de algunos problemas hidrológicos [Escalante (1998)].

- **Condiciones básicas que definen a las distribuciones asintóticas extremas**
  - a) Las observaciones de la cual los valores extremos son tomados deben ser independientes.
  - b) Las observaciones deben ser confiables y estar hechas bajo condiciones idénticas.
  - c) El número de observaciones extremas,  $n$ , debe ser grande. El qué tan grande debe ser  $n$  depende de la distribución inicial y del grado de precisión buscado.

Gumbel (1958) estableció que uno de los principales hallazgos en la teoría de valores extremos, es que las distribuciones marginales univariadas deben ser de los siguientes tipos:

Tipo I. Distribución Gumbel o doble exponencial (primera asíntota),

Tipo II. Distribución Fréchet (segunda asíntota),

Tipo III. Distribución Weibull (tercera asíntota).

Gumbel (1962) demostró que existen tres posibles funciones para distribuciones bivariadas de valores extremos:

1. Tipo Gumbel-Gumbel, en donde las dos funciones marginales son Gumbel.
2. Tipo Gumbel-GVE, en donde una de las funciones marginales es Gumbel y la otra es general de valores extremos.
3. Tipo GVE-GVE, en donde ambas funciones marginales son tipo GVE.

Entre las distribuciones bivariadas de valores extremos, se encuentran los llamados modelos diferenciales (*aquellos que cuentan con una función de densidad*) y los no diferenciables, estos modelos no cuentan con una función densidad de probabilidad explícita, lo cual hace que la estimación de parámetros sea muy complicada [Raynal (1985)].

Los resultados obtenidos por Tiago de Oliveira (1962, 1963) y Gumbel-Mustafi (1967) consideran explícitamente, para el *caso bivariado*, dos posibles funciones de distribución de probabilidades diferenciables: el modelo mezclado o mixto [Gumbel (1960)] y el *modelo logístico bivariado de Gumbel*.

### 4.1.3. Función de distribución logístico bivariado de valores extremos

El modelo logístico bivariado propuesto por Gumbel (1960) tiene mayor versatilidad que el modelo mixto, este modelo recibió este nombre ya que la diferencia entre dos marginales Gumbel reducidas sigue una distribución logística. El modelo logístico es más apropiado para su aplicación a la solución de problemas de análisis de frecuencia conjunto en Hidrología [Raynal (1985)], el modelo logístico está definido como:

$$F(x_1, x_2; m) = \exp \left\{ - \left[ (-\ln F(x_1))^m + (-\ln F(x_2))^m \right]^{\frac{1}{m}} \right\}, \quad (4.8)$$

donde

$F(x_1)$  y  $F(x_2)$  son algunas posibles marginales como Gumbel T1, Fréchet, Weibull, exponencial o GVE.

$m$  es el parámetro de asociación, con  $m \geq 1$ .

La Función Densidad de Probabilidad Logística Bivariada es

$$f(x_1, x_2; m) = \frac{F(x_1, x_2; m)}{F(x_1) \cdot F(x_2)} \cdot \left\{ (m-1) + \left[ (-\ln F(x_1))^m + (-\ln F(x_2))^m \right]^{\frac{1}{m}} \right\} \cdot \left\{ \left[ (-\ln F(x_1))^m + (-\ln F(x_2))^m \right]^{\frac{1}{m}-2} \right\} \cdot \{ (-\ln F(x_1)) \cdot (-\ln F(x_2)) \}^{m-1} \cdot \{ f(x_1) \cdot f(x_2) \}. \quad (4.9)$$

La relación entre el coeficiente de correlación lineal de Pearson  $\rho$  y el parámetro de asociación  $m$ , se define como [Gumbel y Mustafi (1967)]

$$\rho = \left( 1 - \frac{1}{m^2} \right) \quad (4.10)$$

donde

$$m = \frac{1}{\sqrt{1 - \rho}} \quad (4.11)$$

En general, de la expresión (4.8) se tiene que sí  $m = 1$ , conlleva a tener independencia estadística, y se cumple que:

$$F(x_1, x_2; 1) = F(x_1) \cdot F(x_2),$$

para  $m = \infty$  y  $\rho = 1$ , conduce al caso de dependencia total, y la función de distribución bivariada es:

$$F(x_1, x_2; \infty) = \min[F(x_1) \cdot F(x_2)].$$

El modelo definido en (4.8), coincide con el modelo Cópula de la familia de Gumbel-Hougaard, mostrado en la Tabla 3.2, no ocurre así con la FDP Logística Bivariada (4.9) y la densidad de la cópula de Gumbel-Hougaard (Tabla 3.2), el término que la diferencia es el producto de las densidades de cada marginal  $\{f(x_1) \cdot f(x_2)\}$  de la expresión (4.9).

*Galambos (1978) estableció que una distribución multivariada es considerada de valores extremos bajo la premisa de que debe ser necesario, pero no suficiente, que las distribuciones marginales sean de valores extremos.*

*Como ejemplo, se realizó un ejercicio empleando la FDP logística bivariada, considerando marginales diferentes; Gumbel Tipo I y Log Normal con tres parámetros, para este caso se hizo la comparación con el mismo modelo bivariado, pero ahora empleando marginales de valores extremos en ambos casos; Gumbel Tipo I. Se optimizaron los parámetros estadísticos de la FDP bivariada mediante algoritmos genéticos y se concluye que el primer caso la optimización no es mejor que el segundo caso, además de que el tiempo para calcular los parámetros del primer caso es sumamente costosa (4 horas). Por lo tanto se comprueba una limitante bajo este enfoque.*

Anteriormente, los análisis elaborados para estudiar la ocurrencia de eventos extremos mediante el modelo logístico bivariado, generalmente se utilizaron marginales con distribuciones tipo Gumbel T1, Exponencial, Fréchet, Weibull o General de Valores Extremos GVE. En términos generales, este modelo conjunto no es tan potente para analizar datos que provengan de eventos extremos, debido a que se restringe únicamente a marginales de valores extremos, lo cual se puede provocar a tener resultados fuera de la realidad.

Hasta antes del año 2004, el modelo logístico era utilizado frecuentemente para el análisis de frecuencia conjunto, posterior a este año, algunos investigadores en el mundo con orientación en hidrología comenzaron a introducir una herramienta poderosa para el análisis de frecuencia conjunto, estas son las llamadas **Funciones Cópula**, dichas funciones se estudiaron a detalle en el Capítulo 3, cabe destacar que la teoría de cópulas ofrece hoy en día una amplia gama de modelos, con el propósito de seleccionar un mejor y adecuado ajuste a los datos observados.

Las marginales que se emplean en las funciones cópula, en ningún momento deben ser idénticas o que pertenezcan a una determinada familia, lo que las convierte en versátiles y flexibles para el caso bivariado. Hasta nuestros días, es el análisis más eficiente y moderno para el estudio de las técnicas estadísticas en hidrología.

---

## 4.2. Ajuste de las distribuciones marginales X, Y

El análisis de frecuencia bivariado comienza por ajustar las muestras X e Y a una distribución marginal que represente cada variable estudiada, y se lleva a cabo por las siguientes etapas:

- i. Estimación de parámetros estadísticos de las distribuciones marginales,  $F(X)=x$  y  $F(Y)=y$ .
- ii. Estimación de la evaluación de bondad de ajuste a diversos modelos de distribución univariado.
- iii. Aplicación de criterios para la selección del mejor modelo de ajuste.

## 4.3. Elección de la Función Cópula bivariada

La selección de la mejor cópula para caracterizar la estructura de dependencia de las variables X e Y se resume en los siguientes pasos (metodología que se detalla en el Capítulo 3):

- i. Obtención de la pseudo-muestra  $(R_i/n, S_i/n)$ ;
- ii. Evaluación de las medidas de dependencia de los datos observados de la muestra bivariada:  $\rho$  de Spearman y  $\tau$  de Kendall
- iii. Estimación del parámetro de la cópula ( $\theta$ )
- iv. Evaluación de la bondad de ajuste para cópulas
- v. Aplicación de criterios para la selección del mejor modelo

## 4.4. Periodo de retorno bivariado usando cópulas

En el análisis de frecuencias tradicional, el periodo de retorno está relacionado con una sola variable, el gasto máximo por ejemplo. Sin embargo, cuando más de una variable aleatoria se encuentra involucrada en el control del comportamiento de una avenida, se debe considerar una extensión del concepto del periodo de retorno, dando lugar al periodo de retorno multivariado.

Los eventos hidrológicos no solo se describen mediante una variable, sino por un conjunto de variables aleatorias correlacionadas entre sí, y que generalmente consisten en el gasto pico ( $Q_p$ ), volumen de la avenida ( $V$ )

y duración de la avenida ( $d$ ). Conforme a este contexto, si más de una de estas variables ( $Q_p, V, d$ ), se considera significativa en el proceso de diseño de una estructura hidráulica, el estudio de un análisis de frecuencia univariado, no proporcionará una evaluación completa y convincente de la probabilidad de ocurrencia de este evento ante una avenida, lo que probablemente conduciría a una estimación inadecuada del riesgo asociado con ese evento. En tal caso, un análisis multivariado dará lugar a estimaciones más apropiadas que un análisis univariado.

Las variables gasto pico y volumen tienen una determinada dependencia, por lo que se deberá considerar el periodo de retorno de ambas variables conjuntamente.

El modelado de la dependencia entre dos o más variables se realiza mediante un análisis basado en Funciones Cópula, esta teoría facilita la estimación de los períodos de retorno multivariados (Capítulo 3), como apoyo se puede consultar la página web de la Comisión Internacional de Hidrología Estadística.

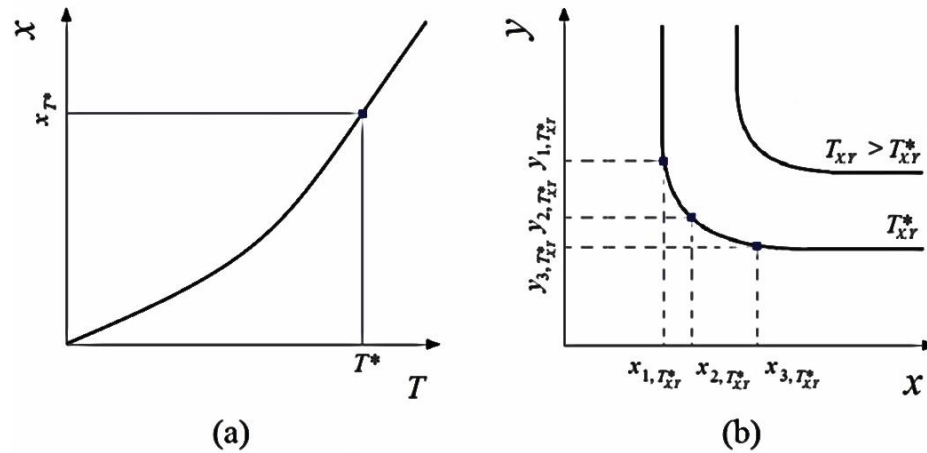
En la Figura 4.1 se muestra las diferencias entre el enfoque univariado y bivariado para obtener el periodo de retorno. Para el caso univariado, un determinado valor del período de retorno  $T^*$  se asocia con uno y solo un valor de cuantil  $x_{T^*}$  (Figura 4.1(a)).

En la Figura 4.1(b) se esquematiza el enfoque del período de retorno bivariado  $T_{X,Y}^*$  y está asociado mediante el conjunto de valores  $(X, Y)$  como  $(X_{1,T_{X,Y}^*}, Y_{1,T_{X,Y}^*})$ ,  $(X_{2,T_{X,Y}^*}, Y_{2,T_{X,Y}^*})$  y  $(X_{3,T_{X,Y}^*}, Y_{3,T_{X,Y}^*})$ .

Todos los pares  $(X, Y)$  que están contenidos en el mismo nivel de probabilidad de la cópula (isolínea o curva de nivel de la cópula) tendrán el mismo período de retorno conjunto.

En este sentido se observa claramente que el periodo de retorno bajo un enfoque bivariado, proporciona una infinidad de soluciones contenidas en una curva de nivel asociada a una probabilidad.





**Figura 4.1.** Enfoques entre periodos de retorno univariado y biviariado.

La estimación del período de retorno biviariado se puede determinar mediante tres enfoques [Favre *et al.* (2016)]:

- a) Probabilidad condicional
- b) Distribución de probabilidad conjunta
- c) Función de distribución de Kendall

#### 4.4.1. Período de retorno condicional

El enfoque del período de retorno condicional se aplica cuando una de las variables de diseño  $(Q_p, V)$  se considera más importante que la otra. El período de retorno condicional se basa en una función de distribución de probabilidad condicional de una variable, dado que se cumpla alguna condición. Este enfoque se aplica a eventos condicionales particulares que se eligen según el problema en cuestión [Favre *et al.* (2016)].

Existen dos tipos de eventos con este enfoque para diseñar una estructura hidráulica,

$$E_{x|y}^> = \{X > x | Y > y\}, \tag{4.12}$$

$$E_{y|x}^> = \{Y > y | X > x\}, \tag{4.13}$$

con probabilidad asociada  $P [X > x / Y > y]$  y  $P [Y > y / X > x]$ .

El período de retorno condicional de estos dos eventos se puede escribir como

$$T(x|y) = \frac{\mu}{1 - \frac{F_X(x) - F_{XY}(x, y)}{1 - F_Y(y)}}, \quad (4.14)$$

$$T(y|x) = \frac{\mu}{1 - \frac{F_Y(y) - F_{XY}(x, y)}{1 - F_X(x)}}, \quad (4.15)$$

donde  $\mu$  es el tiempo medio de llegada entre dos eventos sucesivos, y se define como uno dividido por el número de eventos de avenida ( $Q_p$ ) ocurridos por año. En esta tesis se trabajará con el gasto máximo ocurrido anualmente, por lo tanto  $\mu$  corresponde a 1 año ( $\mu = 1$ ).

El primer evento, ecuaciones (4.12 y 4.14) corresponde a la situación cuando la descarga máxima X supera un umbral x, dado que el volumen de la avenida Y excede un umbral y. Este evento se usaría si el volumen de la avenida se considerara como la variable más representativa para el proceso de diseño.

El segundo evento, ecuaciones (4.13 y 4.15) corresponde a la situación en la que el volumen de la avenida Y supera un umbral y, dado que la descarga máxima X excede un umbral x. Este evento se usaría si la descarga máxima se considerara la variable más importante en el proceso de diseño.

#### 4.4.1.1. Periodo de retorno condicional utilizando Funciones Cópula

El estudio del periodo de retorno condicional se puede facilitar utilizando cópulas,

$$T(u|v) = \mu \frac{1 - v}{1 - u - v + C(u, v)}, \quad (4.16)$$

$$T(v|u) = \mu \frac{1-u}{1-u-v+C(u,v)}, \quad (4.17)$$

donde

$u, v$  son distribuciones marginales uniformes,

$C(u, v)$  es la función cópula,

$\mu$  tiempo medio de llegada entre dos eventos sucesivos en un año.

#### 4.4.2. Periodo de retorno conjunto

El período de retorno conjunto de un evento multivariado se puede calcular utilizando diferentes funciones de distribución de probabilidad conjunta. En la Figura 4.2(a) se ilustran cuatro formas diferentes de definir los valores de la función de distribución de probabilidad conjunta. Los cuadrantes I a IV muestran diferentes formas de definir una probabilidad conjunta:

$$\text{Cuadrante I} : P[X > x, Y > y] = 1 - F_X(x) - F_Y(y) + F_{XY}(x, y)$$

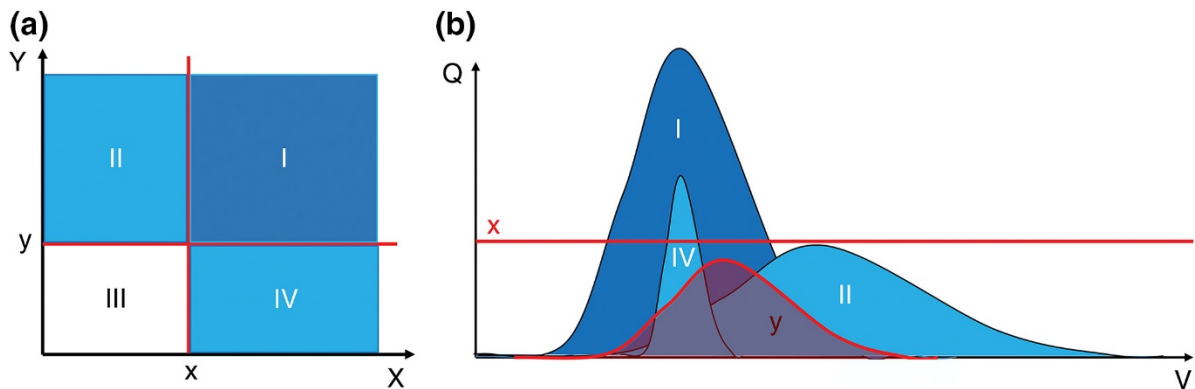
$$\text{Cuadrante II} : P[X \leq x, Y > y]$$

$$\text{Cuadrante III} : P[X \leq x, Y \leq y] = F_{XY}(x, y)$$

$$\text{Cuadrante IV} : P[X > x, Y \leq y]$$

En la Figura 4.2(a), la línea vertical en color rojo representa el umbral del gasto pico  $x$ , mientras que la línea horizontal en color rojo representa el umbral del volumen de la avenida. En la Figura 4.2(b), la línea horizontal en color rojo representa el umbral del gasto pico  $x$ , en esta misma ilustración el hidrograma delimitado con línea en color rojo representa el umbral del volumen total de la avenida  $y$ .

En la Figura 4.2(a), se observa que en el cuadrante I se presenta el caso cuando ambas variables  $X$  e  $Y$  exceden los valores  $x$  e  $y$ . El cuadrante II muestra el caso donde  $Y$  pero no  $X$  excede el valor de referencia. El cuadrante III muestra el caso cuando  $X$  e  $Y$  no superan sus valores de referencia. Finalmente, el cuadrante IV muestra el caso cuando  $X$  pero no  $Y$  excede el valor de referencia.



**Figura 4.2.** Ilustración de probabilidades conjuntas distribuidas en cuatro cuadrantes y restringido por los umbrales  $x$  e  $y$  [Favre *et al.* (2016)].

Para comprender mejor aún las probabilidades distribuidas en los cuadrantes de la Figura 4.2(a), se proporciona un evento como ejemplo. El evento de inundación en el Cuadrante I tiene un pico de descarga más alto y un volumen de avenida más alto que el dado por los umbrales. El evento en el cuadrante II tiene un volumen mayor que el umbral pero una descarga pico más baja. El evento en el cuadrante IV tiene un volumen más bajo que el umbral pero una descarga máxima más alta.

En lo que respecta a la Figura 4.2(b), los eventos ubicados en el Cuadrante I corresponden a eventos en los que tanto la descarga pico  $X$  como el volumen de la avenida  $Y$  superan los umbrales dados " $x$ " e " $y$ ". Los eventos ubicados en el cuadrante II como el volumen de la avenida  $Y$  superan el umbral determinado " $y$ ", pero no así el umbral de la descarga pico " $x$ ". Por el contrario, los eventos ubicados en el Cuadrante IV como la descarga máxima que sobrepasa el umbral fijado, pero no ocurre así con el volumen de la avenida, ya que no supera un cierto umbral " $y$ ".

Los cuatro posibles eventos conjuntos anteriormente citados, utilizan operadores " $O$ " e " $Y$ ", identificados con los símbolos,  $(\vee)$  y  $(\wedge)$  respectivamente, al añadir estos símbolos se generan cuatro combinaciones más como posibles eventos, en la Tabla 4.1 se muestran las ocho posibles combinaciones de probabilidad conjunta.

El período de retorno conjunto de los eventos situados en los cuadrantes I, II o IV donde el gasto pico o el volumen de la avenida (o ambos) exceden un umbral dado, se define como  $T^{\vee}(x, y)$ ;

**Tabla 4.1.** Ocho posibles eventos conjuntos con los operadores "O" ( $\vee$ ) e "Y" ( $\wedge$ )

Eventos		$(X \leq x)$		$(X > x)$	
"O"	$(Y > y)$	$(X \leq x) \vee (Y > y)$	II	$(X > x) \vee (Y > y)$	I
(V)	$(Y \leq y)$	$(X \leq x) \vee (Y \leq y)$	III	$(X > x) \vee (Y \leq y)$	IV
"Y"	$(Y > y)$	$(X \leq x) \wedge (Y > y)$	II	$(X > x) \wedge (Y > y)$	I
(Λ)	$(Y \leq y)$	$(X \leq x) \wedge (Y \leq y)$	III	$(X > x) \wedge (Y \leq y)$	IV

$$T^\vee(x, y) = \frac{\mu}{P[X > x \vee Y > y]} = \frac{\mu}{1 - F_{XY}(x, y)}. \quad (4.18)$$

El período de retorno conjunto de los eventos situados en el cuadrante I, donde tanto el gasto pico y el volumen de la avenida exceden un umbral, se puede expresar como  $T^\wedge(x, y)$ .

$$T^\wedge(x, y) = \frac{\mu}{P[X > x \wedge Y > y]} = \frac{\mu}{1 - F_X(x) - F_Y(y) + F_{XY}(x, y)}. \quad (4.19)$$

#### 4.4.2.1. Período de retorno conjunto utilizando Funciones Cópula

La distribución conjunta bivariada se pueden obtener utilizando un modelo de cópula bivariado. El período de retorno conjunto solo depende del valor de cópula y no de las distribuciones marginales. Los periodos de retorno "O" ( $\vee$ ) e "Y" ( $\wedge$ ) se calculan de la siguiente manera:

$$T^\vee(u, v) = \frac{1}{1 - C(u, v; \theta)}, \quad (4.20)$$

$$C(u, v; \theta) - \left[ 1 - \frac{1}{T^\vee(u, v; \theta)} \right] = 0, \quad (4.21)$$

$$T^\wedge(u, v) = \frac{1}{1 - u - v + C(u, v; \theta)}, \quad (4.22)$$

$$-C(u, v; \theta) + u + v - \left[ 1 - \frac{1}{T^\wedge(u, v; \theta)} \right] = 0, \quad (4.23)$$

donde

$u, v$  son distribuciones marginales uniformes,  
 $C(u, v)$  es la función cópula.

Como ejemplo, en la Figura 4.3 se muestra una gráfica que representa a una familia de curvas de nivel con distintos contornos de probabilidad  $t$ , o mejor dicho, periodos de retorno bivariados que involucran a variables  $(Q, V)$ , dichas curvas de periodo de retorno conjunto se construyeron por medio del análisis "O" ( $\vee$ ) e "Y" ( $\wedge$ ). La cópula que se utilizó para modelar los periodos de retorno conjuntos, es la Cópula Joe/B5 de un parámetro, en la misma gráfica también se representa el comportamiento de los periodos de retorno univariados asociados a  $Q$  y  $V$ .

En este sentido, al existir un número infinito de eventos posibles para un único periodo de retorno dado, es necesario analizar diferentes combinaciones de conjuntos  $(Q, V)$  que produzcan los efectos más desfavorables sobre la estructura u obra en particular bajo estudio. En el Capítulo 6, se detalla la metodología que se empleará para determinar cuál evento en particular genera el o los eventos más desfavorables, tomando en cuenta las características propias de la obra.

Derivado de las expresiones (4.20 a 4.23) y conforme a la Figura 4.3, resulta interesante notar que para una curva de periodo de retorno dado  $T(u, v; \theta)$ , evidentemente existe un conjunto infinito de soluciones  $(Q, V)$ . Conforme a la Figura en comento, los contornos en color rojo indican la familia de curvas de periodo de retorno conjunto  $T^{\wedge}(u, v; \theta)$  mediante un análisis "Y" ( $\wedge$ ), los contornos en color azul son las curvas de periodo de retorno conjunto  $T^{\vee}(u, v; \theta)$  mediante un análisis "O" ( $\vee$ ), y las curvas en color verde representan el comportamiento de los periodos de retorno univariados;  $Q$  y  $V$  respectivamente.

El modelo cópula que adopta la mejor dependencia entre variables, previamente justificado mediante una prueba de bondad de ajuste, es la Cópula de Joe/B5 de un parámetro. Las distribuciones marginales que mejor representan a las variables  $Q$  y  $V$ , son Doble Gumbel respectivamente.

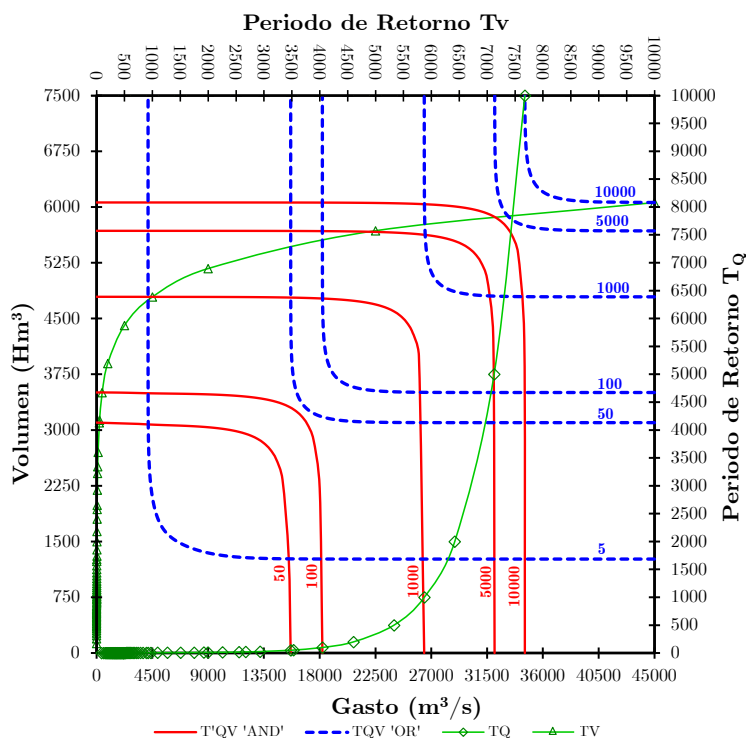


Figura 4.3. Contornos de Periodos de retorno conjuntos mediante análisis "O" (v) e "Y" (Λ); modelados con la Cópula de Joe/B5.

### 4.4.3. Periodo de retorno Kendall

Salvadori y De Michele (2004) introdujeron la función de distribución de Kendall (4.24), que depende solo de la función cópula  $C$ , y por lo tanto, divide el espacio muestral en una región supercrítica y no crítica. La función de distribución de Kendall  $K_C$  representa la función de distribución acumulada de las curvas de nivel de la cópula o isolíneas, para el caso bivariado se define como:

$$K_C(t) = P[W \leq t] = P[C(U, V) \leq t] \tag{4.24}$$

La función de Kendall  $K_C$  en términos de la función generadora de una cópula arquimediana se expresa como

$$K_C(t) = t - \frac{\varphi(t)}{\varphi'(t)} \tag{4.25}$$

donde

$t = C(u, v; \theta)$  es el nivel de probabilidad de una cópula arquimediana

$\varphi(t)$  es la función generadora de una cópula arquimediana

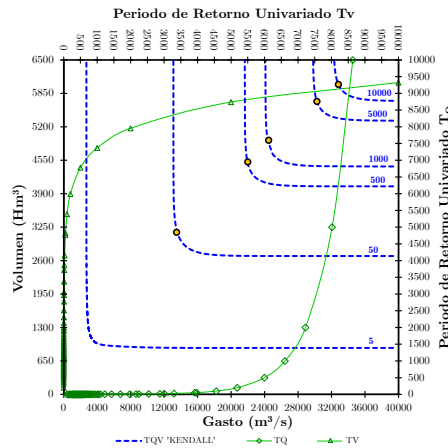
$\varphi'(t)$  es la derivada de la función generadora de una cópula arquimediana

El período de retorno Kendall, por lo tanto, corresponde al tiempo medio de llegada entre eventos críticos que se encuentran en el nivel de probabilidad  $t$ , y está dado por:

$$T_{Kc} = \frac{\mu}{1 - K_C(t)}, \quad (4.26)$$

$$C(u, v; \theta) - \frac{\varphi(t)}{\varphi'(t)} - \left[ 1 - \frac{1}{T_{Kc}(u, v; \theta)} \right] = 0, \quad (4.27)$$

El periodo de retorno Kendall, únicamente se puede emplear para Cópulas Arquimedianas. En la Figura 4.4 se muestra un ejemplo de curvas de periodo de retorno mediante la Función Kendall.



**Figura 4.4.** Contornos de Periodos de retorno conjunto mediante la función Kendall; modelados con la Cópula de Joe/B5.

**Tabla 4.2.** Primera derivada de la Función Generadora de una Cópula Arquimediana

Derivada del Generador	Clayton	Frank	Ali-Mikhail-Haq	Gumbel-Hougaard	Joe/B5
$\varphi'(t; \theta)$	$-\frac{\theta}{t^{\theta+1}}$	$\frac{\theta}{1 - e^{\theta \cdot t}}$	$\frac{\theta - 1}{t - \theta t + \theta t^2}$	$-\frac{\theta}{t} (-\ln t)^{\theta-1}$	$-\frac{\theta(1-t)^{\theta-1}}{1 - (1-t)^\theta}$





# Capítulo 5

## Generación de hidrogramas sintéticos: Función Gamma 2P

### 5.1. Introducción

La parametrización de hidrogramas resulta útil en la definición de avenidas de diseño, por lo que, diversas opciones de parametrización han sido propuestas en la literatura. No obstante, éstas exhiben ciertas limitaciones que restringen su aplicación en la representación de hidrogramas naturales.

Si bien un hidrograma puede caracterizarse a través de su gasto pico, tiempo pico y su volumen, el conocimiento de  $Q_p$ ,  $t_p$  y  $V$  no basta para determinar los efectos que una avenida que ingresa al vaso. Para propósitos de diseño, es necesario transitar la avenida por el vaso, lo cual supone el conocimiento del hidrograma en la forma de una relación funcional gasto-tiempo  $Q(t; Q_p, V, t_p)$ , donde  $Q$  representa el gasto y tiene el papel de la variable dependiente; el tiempo  $t$  es la variable independiente y  $Q_p, t_p$  y  $V$  son parámetros de la función. El proceso de plantear ecuaciones como  $Q(t; Q_p, V, t_p)$  es conocido como parametrización de hidrogramas [Ramírez y Aldama (2000)].

La parametrización de hidrogramas se puede realizar de diferentes formas, existen diversos métodos en la literatura, por ejemplo uno muy conocido en el mundo es el hidrograma adimensional del Servicio de Conservación de Suelos de Estados Unidos (USSCS, por sus siglas en inglés, Mockus (1957)); el hidrograma sintético con la forma de la función de distribución Gamma de dos parámetros [Croley II (1980)]; el hidrograma analítico basado en la función Gamma [Ponce (1988), Reich (1962)]; el hidrograma con la forma de la función Pearson tipo III [Reich (1962), Hiemstra y Francis (1981)], por mencionar algunos. Singh (1988) presenta un desarrollo importante al respecto, propone métodos gráficos para estimar los parámetros de la distribución Gamma.

A continuación se ofrece una breve descripción de las tres últimas parametrizaciones, las cuales son utilizadas con frecuencia en el análisis de sistemas hidrológicos.

## 5.2. Parametrización del hidrograma a través de la Función Pearson Tipo III (sin conocer el tiempo pico) (Método 1)

Reich (1962) construyó una función que se asemeja a la forma promedio de un hidrograma de avenida, con base en la Función Pearson tipo III. Esta función es más flexible con respecto a otras opciones que ofrece la literatura sobre este tema y se puede utilizar como una primera aproximación del hidrograma de avenida. La propuesta realizada por el autor es muy práctica y versátil, aunque no la más favorable en comparación con los resultados que proporcionan otros métodos.

Para ajustar la forma del hidrograma de avenida, una forma de la función Pearson tipo III se puede expresar como,

$$Q(t) = Q_p \left( 1 + \frac{t}{t_p} \right)^{t_p/t_G} e^{-t/t_G}, \quad (5.1)$$

$$t \in [-t_p, \infty)$$

donde

$Q$  gasto en cualquier momento  $t_i$

$Q_p$  gasto pico

$t_p$  tiempo pico

$t_G$  tiempo desde el gasto pico hasta el centroide del hidrograma.

En la Figura 5.1 se muestran los parámetros de la Ec. (5.1) que describen a la forma del hidrograma de avenida utilizando la función Pearson tipo III.

Si se integra la ecuación (5.1) resulta el volumen total ( $V$ ) de la avenida y resulta lo siguiente

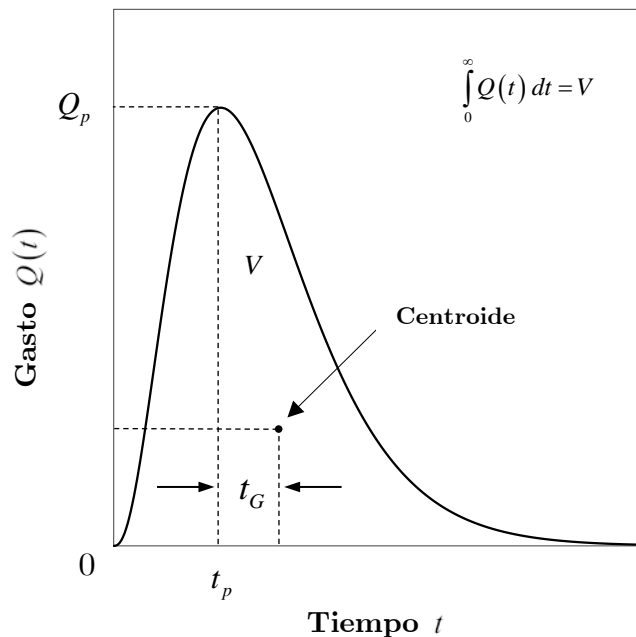
$$\int_0^{\infty} Q(t) dt = V,$$

entonces

$$V = Q_p e^{t_p/t_G} \left(\frac{t_G}{t_p}\right)^{t_p/t_G} t_G \Gamma\left(1 + \frac{t_p}{t_G}\right), \quad (5.2)$$

donde  $\Gamma(\cdot)$  es la función gamma.

Supóngase que en una avenida, el volumen  $V$  y el gasto pico  $Q_p$ , son conocidos, entonces  $t_p$  y  $t_G$  están relacionados de manera única, por consiguiente, solo un parámetro adicional es necesario para describir un hidrograma de manera única, ese parámetro puede ser  $t_p$  o  $t_G$ .



**Figura 5.1.** Parámetros del Hidrograma de avenida (Función Pearson T-III)

---

Hiemstra y Francis (1981) presentaron una ecuación para calcular el tiempo al centroide  $t_G$  en función de  $V$  y  $Q_p$ , con correlación  $R^2 = 0.72$ , y está basada en 1,112 hidrogramas probados con  $t_G$  positivo únicamente, finalmente la ecuación resulta

$$t_G = \exp \left\{ -18.1129 + 8.336 \operatorname{Ln} \left[ \operatorname{Ln}(V) - \operatorname{Ln}(Q_p) \right] \right\}, \quad (5.3)$$

donde,  $V$  está en metros cúbicos,  $Q_p$  en metros cúbicos por segundo, y  $t_G$  está en horas.

Una vez encontrado  $t_G$ , es posible hallar el tiempo pico  $t_p$ , a partir de la ecuación (5.2).

La ecuación (5.2) se puede reescribir como

$$\frac{V}{Q_p \cdot t_G} = e^{t_p/t_G} \left( \frac{t_G}{t_p} \right)^{t_p/t_G} \Gamma \left( 1 + \frac{t_p}{t_G} \right) \quad (5.4)$$

Si el primer miembro de (5.4) la nombramos como  $\alpha'$

$$\alpha' = \frac{V}{Q_p \cdot t_G} \quad (5.5)$$

entonces

$$\alpha' = e^{t_p/t_G} \left( \frac{t_G}{t_p} \right)^{t_p/t_G} \Gamma \left( 1 + \frac{t_p}{t_G} \right) \quad (5.6)$$

Por lo tanto, dados  $Q_p$  y  $V$ , se encuentra  $t_G$  de (5.3), posteriormente se calcula  $\alpha'$  (5.5). El tiempo pico  $t_p$  se obtiene a partir de (5.6) proponiendo distintos valores de  $t_p$ . Las ordenadas para construir la forma del hidrograma de la avenida se deduce de la ecuación (5.1).

**Ejemplo 1:** Para aplicar la presente metodología, se utiliza la información proporcionada en el ejemplo propuesto en Hiemstra y Francis (1981, p. 772), utilizar datos de las parejas ordenadas  $(Q_p, V)$  de cinco los hidrogramas. Determinar el gasto pico  $t_p$  y demás parámetros que caracterizan a los cinco hidrogramas, luego construir los hidrogramas.

En la Tabla 5.1 se muestran los resultados y en la Figura 5.2 se muestran la familia de hidrogramas.

**Tabla 5.1.** Parámetros de una familia de hidrogramas (T=50 años)

Hidrograma	V (Hm <sup>3</sup> )	Q <sub>p</sub> (m <sup>3</sup> /s)	t <sub>G</sub> (h)	α'	t <sub>p</sub> (h)	t <sub>p</sub> / t <sub>G</sub>	Γ(1+ t <sub>p</sub> /t <sub>G</sub> )	α'
1	308.839182485148	2231	12.0380	3.1943	<b>17.4552</b>	1.4500	1.2842	3.1943
2	435.931185459339	2120	15.8338	3.6074	<b>30.0608</b>	1.8985	1.8250	3.6074
3	583.839237294971	1878	20.8874	4.1344	<b>53.2438</b>	2.5491	3.5096	4.1344
4	692.260507873469	1507	26.9125	4.7413	<b>91.7042</b>	3.4075	10.2403	4.7413
5	736.966599324117	1164	32.9584	5.3361	<b>143.7725</b>	4.3622	42.0047	5.3361

Para bosquejar la forma del hidrograma con la Ec. (5.1), es de interés observar lo siguiente respecto al artículo de Hiemstra y Francis (1981) acerca de esta ecuación: basándonos en (5.1) de este trabajo, un primer escenario consiste cuando  $t_i = 0$  y es de esperarse que resultara  $Q(t_i) = 0$ , en este caso, con estas evaluaciones no es así, por lo que resulta  $Q(t_i) = Q_p$ , en un segundo escenario se observa que cuando  $t_i = t_p$  es de esperarse que resultara  $Q(t_i = t_p) = Q_p$ , lo cual no fue cierto, al momento de la evaluación resultó  $Q(t_i) \neq Q_p$ . Realmente lo que sucede en ambos escenarios, es el hecho que, la Ec. (5.1) únicamente está representando la mitad del hidrograma, es decir, hacia la cola derecha, en  $t_i \in [0, \infty)$ .

Para dar solución práctica a esta observación y por lo tanto bosquejar un hidrograma completo se propone lo siguiente: proponer un intervalo de tiempo  $\Delta t$  e implementar  $(t_0 = -t_p), (t_1 = t_0 + \Delta t), \dots, (t_i = t_{i-1} + \Delta t)$  en la serie de tiempo  $t_i$ , con esta distribución de tiempo se calculan los  $Q(t_i)$ .

Si se desea bosquejar la gráfica con estos últimos pares de valores ordenados  $(t_i, Q(t_i))$ , se obtendría un hidrograma con cola negativa en la izquierda y positiva a la derecha, mejor dicho, la escala de la serie de tiempo  $t_i$  se ubicaría en el intervalo  $t_i \in [-t_p, \infty)$ , la ubicación del gasto pico se posicionaría a una distancia  $t_p$  del eje vertical  $Q$  con respecto al origen del sistema coordenado  $tQ$ .

En este sentido, si se desea bosquejar el hidrograma completo cuya escala de la serie de tiempo  $t_i$  se ubique en el intervalo  $t_i \in [0, \infty)$ , basta con

utilizar el mismo intervalo de tiempo  $\Delta t$  definido anteriormente y modificar la serie de tiempo  $t_i$ , como  $(t_0 = 0), (t_1 = t_0 + \Delta t), \dots, (t_i = t_{i-1} + \Delta t)$ , y junto con la serie de gastos  $Q(t_i)$  calculados anteriormente con (5.1), finalmente se bosqueja el hidrograma completo en dicho intervalo.

En la Figura 5.2, se muestra la familia de cinco hidrogramas del Ejemplo 1, utilizando los parámetros resultantes (Tabla 5.1) con la metodología antes descrita, los hidrogramas están bosquejados en el intervalo  $t_i \in [0, \infty)$  para el eje de las abscisas  $t$ .

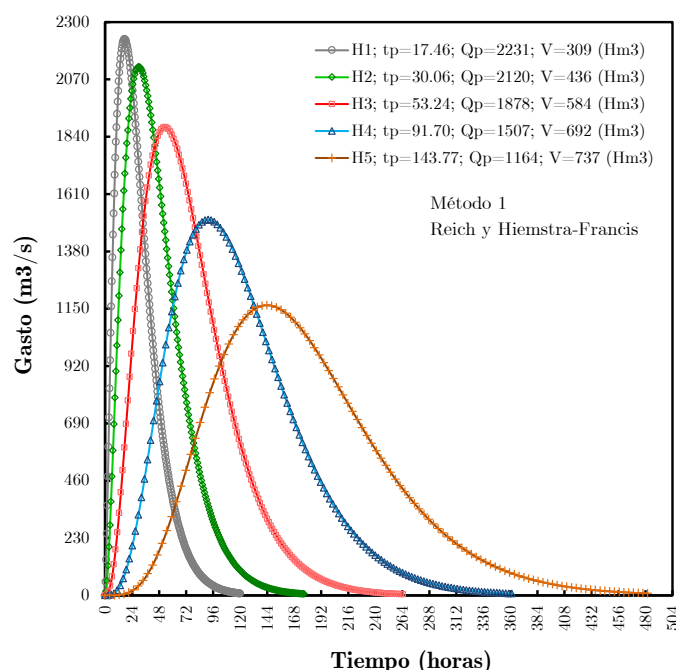


Figura 5.2. Familia de hidrogramas de avenida (Función Pearson T-III)

### 5.3. Parametrización del hidrograma a través de la Función Gamma con dos parámetros (sin conocer $t_p$ ) (Método 2)

Croley II (1980) propuso a la distribución Gamma de dos parámetros para dar la forma de un hidrograma de avenida, la forma de presentar a la distribución Gamma es la misma que emplea el Modelo de Nash (1957) para el análisis de reservorios lineales en cascada [Singh (1988)]. Para aplicar esta metodología es necesario contar mínimo con los valores de las variables de gasto pico  $Q_p$  y volumen de la avenida  $V$ .

La función de distribución Gamma de dos parámetros se define como

$$Q(t) = \frac{1}{\alpha \Gamma(\beta)} \left( \frac{t}{\alpha} \right)^{(\beta-1)} \cdot e^{-\frac{t}{\alpha}}, \quad (5.7)$$

Donde,  $\alpha > 0$  es el parámetro de escala y  $\beta > 0$  es el parámetro de forma de la distribución Gamma, los cuales satisfacen un conjunto de dos condiciones límite y  $\Gamma(\beta)$  es la función gamma.

Para obtener el tiempo máximo  $t = t_p$ , o tiempo pico, se deriva con respecto a  $t$  la Ec. (5.7) y se iguala a cero, esto es  $dQ/dt = 0$ , entonces

$$t_p = \alpha(\beta - 1). \quad (5.8)$$

Si se combinan las Ecs. (5.7) y (5.8) se obtiene el gasto máximo o gasto pico, entonces si se evalúa  $Q(t_p) = Q_p$ , por lo tanto se obtiene la siguiente expresión:

$$Q_p = \frac{1}{\alpha \Gamma(\beta)} (\beta - 1)^{(\beta-1)} \cdot e^{-(\beta-1)}. \quad (5.9)$$

La Ec. (5.9) se puede reescribir de la siguiente manera

$$\frac{1}{\alpha \Gamma(\beta)} = Q_p (\beta - 1)^{-(\beta-1)} \cdot e^{(\beta-1)} \quad (5.10)$$

Para construir la forma del hidrograma de avenida sin tomar en cuenta la función gamma  $\Gamma(\beta)$ , se combinan las Ecs. (5.7) y (5.10), las siguientes expresiones son conocidas como hidrogramas adimensionales estándar, y se escriben en términos de sus parámetros  $\alpha$  y  $\beta$  como:

$$Q(t) = Q_p \left( \frac{t}{t_p} \right)^{(\beta-1)} e^{(\beta-1) \left( 1 - \frac{t}{t_p} \right)} \quad (a), \quad (5.11)$$

$$Q(t) = Q_p \left( \frac{t}{t_p} \right)^{t_p/\alpha} e^{(t_p-t)/\alpha} \quad (b).$$



*Croley II (1980) ha contribuido con una alternativa muy interesante para estimar dichos parámetros.* Al multiplicar la distribución Gamma de dos parámetros Ec. (5.7) por el volumen de la venida  $V$ , resulta

$$Q(t) = \frac{V}{\alpha \Gamma(\beta)} \left( \frac{t}{\alpha} \right)^{(\beta-1)} \cdot e^{-\frac{t}{\alpha}}. \quad (5.12)$$

Conforme a la expresión (5.12), si se desea averiguar el tiempo máximo o tiempo pico, resulta la misma expresión (5.8), en el caso del gasto máximo o gasto pico, resulta una expresión similar a la (5.9), como

$$Q_p = \frac{V}{\alpha \Gamma(\beta)} (\beta - 1)^{(\beta-1)} \cdot e^{-(\beta-1)}. \quad (5.13)$$

Si se despeja  $\alpha$  de la expresión (5.8) y se emplea en (5.13), realizando algebra, resulta la siguiente expresión en términos de  $\beta$  como

$$\frac{Q_p t_p}{V} = \frac{1}{\Gamma(\beta)} (\beta - 1)^\beta \cdot e^{-(\beta-1)}. \quad (5.14)$$

Por medio de la Ec. (5.14) es posible estimar el parámetro  $\beta$  numéricamente, el empleo de esta ecuación únicamente es operable si se conocen tres variables  $t_p$ ,  $Q_p$  y  $V$ . Donde, el tiempo pico  $t_p$  tiene unidades<sup>1</sup> en millones de segundos [Mseg], el volumen tiene unidades en hectómetros cúbicos [Hm<sup>3</sup>], el gasto pico  $Q_p$  en unidades de hectómetros cúbicos por millones de segundos [Hm<sup>3</sup>/Mseg], y  $\beta$  es adimensional.

Se recomienda utilizar las unidades anteriores, ya que generalmente el volumen se maneja en, (Hm<sup>3</sup>) y el tiempo en horas, por lo que únicamente para homogeneizar el tiempo se convierte en [Mseg], también es muy válido emplear unidades como,  $Q_p$  en [m<sup>3</sup>/s],  $V$  en [m<sup>3</sup>],  $t_p$  en [s].

La Ec. (5.14) en términos de  $\alpha$ , resulta como

$$\frac{Q_p t_p}{V} = \frac{\left[ \frac{e^{(1)}}{(t_p/\alpha)} \right]^{-(t_p/\alpha)}}{\Gamma(t_p/\alpha)}, \quad (5.15)$$

<sup>1</sup>Conversiones.  $1Hm^3 = 1E06 m^3$ ;  $1Mseg = 1E06 seg$ ;  $1Mseg = (1E06/3600) hrs$ ;  $1(m^3/s) = 1(Hm^3/Mseg)$ ;  $1(m^3/s) = 0.0036(Hm^3/h)$ ;  $1(m^3/s hora) = 127116 (0.3048)^3 m^3$ .

donde, el parámetro  $\alpha$  tiene unidades de tiempo, depende de las unidades que se empleen para el tiempo pico  $t_p$ , generalmente  $t_p$  se expresa en horas, por lo tanto  $\alpha$  tendrá unidades de horas.

Las expresiones (5.14) o (5.15) se pueden usar para determinar  $\alpha$  o  $\beta$ , posteriormente, una vez obteniendo un parámetro, con ayuda de la Ec. (5.8), se obtiene el segundo parámetro. Debido a los casos que he analizado, se recomienda hacer uso de la expresión (5.14).

La función gamma,  $\Gamma(\beta)$  puede aproximarse con la siguiente expresión, con  $\beta > 0$  [Abramowitz y Stegun (1967); Croley (1977)]:

$$\Gamma(\beta) \approx \frac{\left[ \frac{2\pi}{(\beta+m)} \right]^{1/2} (\beta+m)^{(\beta+m)} \exp \left\{ -\beta - m + \frac{1}{12(\beta+m)} - \frac{1}{360(\beta+m)^3} \right\}}{(\beta+m-1)(\beta+m-2)\dots(\beta+1)\beta} \quad (5.16)$$

La exactitud de la aproximación en la Ec. (5.16) aumenta con el valor de m, pero también lo hace el número de cálculos. Los valores de m mayores que 5 no mejoran en gran medida la precisión, con m=5 la aproximación es precisa, arrojando entre 6 y 7 dígitos significativos.

Al calcular  $\alpha$  y  $\beta$ , y conociendo  $t_p$ ,  $Q_p$  y  $V$ , el hidrograma adimensional estándar se construye como se indica en la ecuación (5.11 (a)).

Conforme a la ecuación (5.11 (a)), el volumen de la avenida queda definido de la siguiente forma

$$V = Q_p t_p \int_0^{\infty} \left( \frac{t}{t_p} \right)^{(\beta-1)} e^{-\left(1 - \frac{t}{t_p}\right)^{\beta}} dt = Q_p t_p e^{(\beta-1)} \cdot (\beta-1)^{-\beta} \cdot \Gamma(\beta) \quad (5.17)$$

Croley II (1980) usó el enfoque anterior para caracterizar el hidrograma de avenida, por conducto de los parámetros  $\alpha$  y  $\beta$ , y dando por hecho que se tiene conocimiento de las variables  $t_p$ ,  $Q_p$  y  $V$ . Como resumen, las ecuaciones para hallar dichos parámetros y construir el hidrograma de avenida, son en este orden: (5.14), (5.16), (5.8) y (5.11 (a)).

*En un principio se comentó que, en esta sección era posible determinar un hidrograma sin conocer el tiempo pico  $t_p$ , según el análisis propuesto*

por Croley II (1980), no es posible llegar a este objetivo. Estudiando con mayor detenimiento las ecuaciones planteadas anteriormente, se llega a la conclusión de que si es factible bosquejar un hidrograma sin conocer el tiempo pico  $t_p$ , aunque en el artículo de Croley II (1980) no se hace comentario alguno.

A continuación se desarrolla un planteamiento breve para dar a conocer las expresiones que implican calcular el tiempo pico  $t_p$  y las demás características que definen al hidrograma.

Partiendo de la Ec. (5.13) ésta se puede reorganizar de la siguiente forma

$$Q_p = \left[ \frac{V}{\alpha \Gamma(\beta)} \right] \cdot (\beta - 1)^{(\beta-1)} \cdot e^{-(\beta-1)}, \quad (5.18)$$

es bien conocido que el gasto se relacione de la siguiente forma como

$$Q = \frac{V}{t}. \quad (5.19)$$

Ahora bien, en la Ec. (5.18), el término separado entre corchetes relaciona a una cantidad de volumen y en el denominador al parámetro  $\alpha$ , que anteriormente se ha comentado que tiene unidades de tiempo, por lo que respecta al parámetro  $\beta$  es adimensional, en este sentido, el término entre corchetes se puede ver como una relación de gasto, lo que equivale a tener

$$\frac{V}{\alpha \Gamma(\beta)} = \frac{V}{t} = Q. \quad (5.20)$$

Al realizar diversas estimaciones con (5.18), y generando dos escenarios se tiene: el primer caso donde se considera al término de (5.20), y para el segundo escenario no se incluye al término. Los resultados que produjeron estas pruebas, se llega a la conclusión de que son totalmente idénticos los resultados en ambos escenarios, por lo que se observa que el termino de (5.20) es igual a la unidad (véase la Tabla 5.2), esto es

$$\frac{V}{\alpha \Gamma(\beta)} = 1 \quad (5.21)$$

Conforme a lo anterior, las ecuaciones (5.12) y (5.13) se simplifican como:

$$Q(t) = \left(\frac{t}{\alpha}\right)^{(\beta-1)} \cdot e^{-\frac{t}{\alpha}}, \quad (5.22)$$

$$Q_p = (\beta - 1)^{(\beta-1)} \cdot e^{-(\beta-1)}. \quad (5.23)$$

El tiempo pico se calcula por medio de la expresión (5.14), lo que resulta

$$t_p = \frac{V}{Q_p \Gamma(\beta)} (\beta - 1)^\beta \cdot e^{-(\beta-1)}, \quad (5.24)$$

donde, el tiempo pico está en términos de  $\beta$ , y se supone que se conoce a  $Q_p$  y  $V$ . El parámetro  $\alpha$  se estima con la expresión (5.8). Se hace un recordatorio que *esta simplificación (Ec. (5.21)) se emplea únicamente cuando no se conoce al tiempo pico  $t_p$* .

A continuación se presentará un ejemplo en el cual se analizan los dos casos anteriores. Para aplicar ambas metodologías, se hace uso de los datos concentrados en la Tabla 5.1. Determinar los parámetros que definen a un hidrograma, tomando en cuenta en principio que no se conoce el tiempo pico, después construya el hidrograma correspondiente, para un segundo caso haga lo propio, la diferencia es que en este caso si se conoce al tiempo pico. En la Tabla 5.2 y Figura 5.3 se presentan los resultados para el primer caso (no conoce  $t_p$ ), para el segundo caso, en la Tabla 5.3 y Figura 5.4 se muestran los respectivos resultados.

**Tabla 5.2.** Parámetros de una familia de hidrogramas (no se conoce  $t_p$ )

Hi	$Q_p$ ( $m^3/s$ )	V ( $Hm^3$ )	$\beta$	$\Gamma(\beta)$	$t_p$ (h)	$\alpha$ (h)	$\alpha \Gamma(\beta)$ (Mseg)	$V/\alpha \Gamma(\beta)$ ( $m^3/s$ )
1	2231	308.839182485148	8.5487737	15535.285	<b>41.6856976</b>	5.5221814	308.839	1
2	2120	435.931185459339	8.5235058	14738.168	<b>61.8147591</b>	8.2162174	435.931	1
3	1878	583.839237294971	8.4633234	13004.629	<b>93.0732915</b>	12.4707569	583.839	1
4	1507	692.260507873469	8.3534231	10360.164	<b>136.4865928</b>	18.5609601	692.261	1
5	1164	736.966599324117	8.2234014	7932.697	<b>186.4087030</b>	25.8062224	736.967	1

**Tabla 5.3.** Parámetros de una familia de hidrogramas (si se conoce  $t_p$ )

Hi	$Q_p$ ( $m^3/s$ )	V ( $Hm^3$ )	$t_p$ (h)	$\beta$	$\Gamma(\beta)$	$\alpha$ (h)	$\alpha \Gamma(\beta)$ (Mseg)	$V/\alpha \Gamma(\beta)$ ( $m^3/s$ )
1	2231	308.839182485148	<b>17.4552095</b>	2.4500143	1.284	12.0379566	55653.40	5549.33
2	2120	435.931185459339	<b>30.0608290</b>	2.8985244	1.825	15.8337858	104026.34	4190.58
3	1878	583.839237294971	<b>53.2437945</b>	3.5490717	3.510	20.8875230	263904.57	2212.31
4	1507	692.260507873469	<b>91.7041997</b>	4.4074773	10.240	26.9126367	992112.67	697.76
5	1164	736.966599324117	<b>143.7725032</b>	5.3622544	42.006	32.9583036	4983952.09	147.87

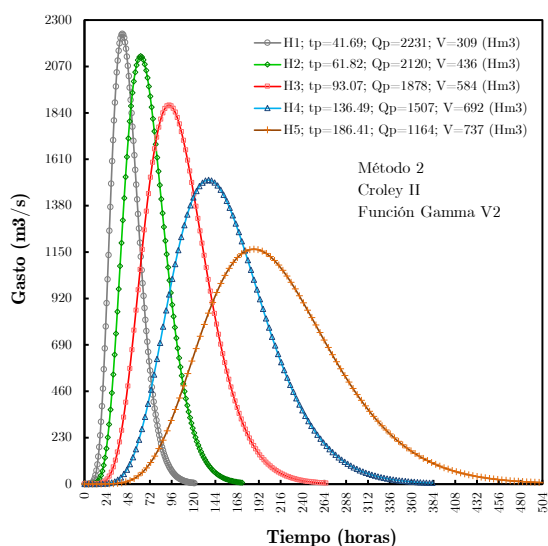


Figura 5.3. Hidrogramas,  $t_p$  calculado (Función Gamma 2P)

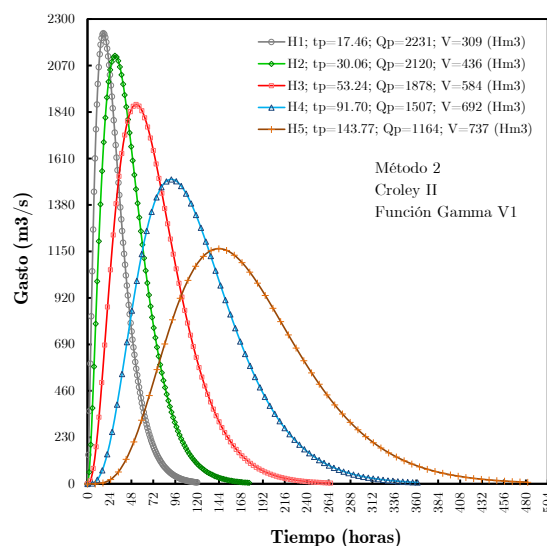


Figura 5.4. Hidrogramas,  $t_p$  conocido (Función Gamma 2P)

### 5.4. Parametrización del hidrograma a través de la Función Gamma (se conoce $t_p$ ) (Método 3)

Se presenta otra forma de parametrizar hidrogramas, es una versión modificada del Método 1, son cambios no tan contundentes, básicamente se modifica la forma en construir el hidrograma de la avenida. Se recuerda que la naturaleza de la Ec. (5.1), el dominio en el cual está parametrizada esta función es  $t \in [-t_p, \infty)$ , situación que originó a proporcionar determinadas recomendaciones para para construir el hidrograma completo con dominio en  $t \in [0, \infty)$ . La actual metodología corrige esa situación, además de que evita el uso de expresiones de regresión, la metodología que se planteará, requiere contar con las variables  $(t_p, Q_p, V)$ .

Considérese el hidrograma adimensional estándar dado en la expresión (5.11 (b)), la cual se deriva de la Función Gamma de dos parámetros,

$$Q(t) = Q_p \left( \frac{t}{t_p} \right)^{t_p/\alpha} e^{-(t_p-t)/\alpha}$$

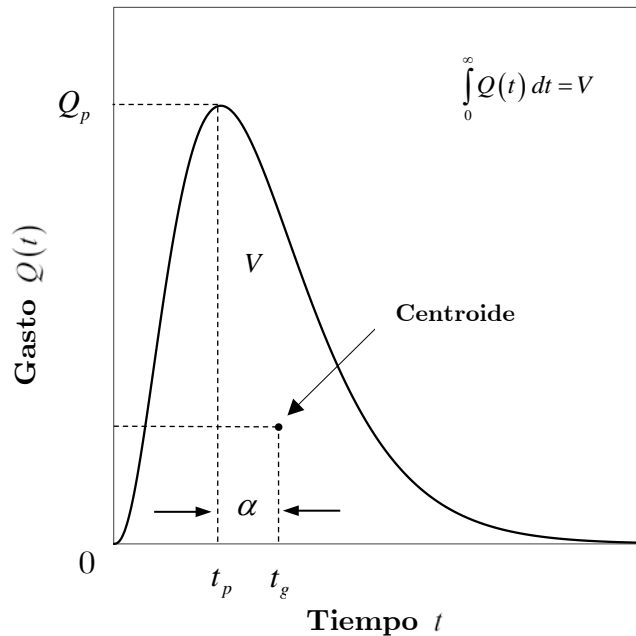
En base a la Figura 5.1,  $t_G$  representa al tiempo desde que se presenta el gasto pico hasta el tiempo en donde se ubica el centroide del hidrograma, por otro lado,  $t_G$  está relacionado directamente con el parámetro  $\alpha$ , que en este sentido es el parámetro de escala y se mueve alrededor del eje del tiempo, por lo que evidentemente este parámetro tiene unidades de tiempo, entonces geoméricamente resulta que

$$t_G = \alpha. \quad (5.25)$$

Ahora bien, si se desea que el hidrograma esté definido con dominio en el tiempo  $t \in [0, \infty)$ , lo que se requiere es desplazar el parámetro del tiempo referenciado al centroide del hidrograma hacia el primer cuadrante del sistema coordenado  $tQ$  (véase Figura 5.5).

De este modo, el tiempo al centroide estará en el intervalo  $t \in [0, t_g)$ , véase Figura 5.5, y con ello, se definen los parámetros del hidrograma en  $t \in [0, \infty)$ , entonces conforme a la figura se tiene la siguiente relación

$$\alpha = t_g - t_p. \quad (5.26)$$



**Figura 5.5.** Parámetros del Hidrograma (Función Gamma)

Si se combinan las ecuaciones (5.11 (b)) y (5.26), se obtiene el siguiente hidrograma adimensional estándar,

$$Q(t) = Q_p \left( \frac{t}{t_p} \right)^{t_p/(t_g-t_p)} \cdot e^{[(t_p-t)/(t_g-t_p)]}, \quad (5.27)$$

para todo  $t \in [0, \infty)$ ,  $t_b \rightarrow \infty$ ,

donde

- $Q$  gasto en cualquier instante  $t_i$
- $Q_p$  gasto pico
- $t_p$  tiempo pico
- $t_g$  tiempo desde  $t_0 = 0$  hasta el centroide del hidrograma.

La cuantificación del volumen se logra integrando la ecuación (5.27), y representa el volumen total de la avenida, esto es

$$\int_0^{\infty} Q(t) dt = V,$$

la solución es la misma que la expresión (5.17), en términos de  $\beta$

$$V = Q_p t_p e^{(\beta-1)} \cdot (\beta-1)^{-\beta} \cdot \Gamma(\beta),$$

la solución anterior se puede escribir en términos de  $t_p$  y  $t_g$  como,

$$V = Q_p e^{t_p/(t_g-t_p)} \left( \frac{t_g-t_p}{t_p} \right)^{t_p/(t_g-t_p)} \cdot (t_g-t_p) \Gamma \left( 1 + \frac{t_p}{t_g-t_p} \right), \quad (5.28)$$

donde,  $\Gamma(\cdot)$  es la función gamma, y se resuelve con (5.16).

De las Ecs. (5.8) y (5.26) se deduce el parámetro  $\beta$  en términos de  $t_p$  y  $t_g$ , como

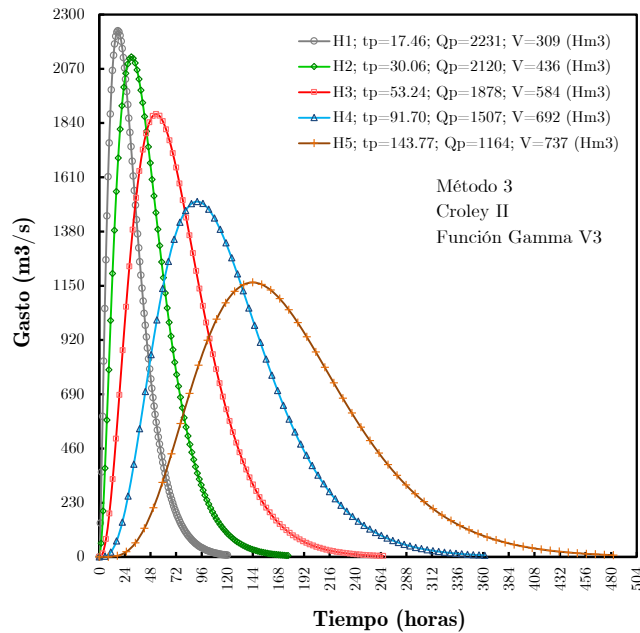
$$\beta = 1 + \frac{t_p}{t_g - t_p}. \quad (5.29)$$

Cuando  $t_p$ ,  $Q_p$  y  $V$  son conocidos, (5.28) representa una función trascendente para  $t_g$ , la cual se puede resolver de forma numérica, se recomienda el método de bisección, y en la sección de apéndices se puede consultar el código programado en Matlab.

Para fines de comparación entre los métodos 1 y 3, nuevamente se realizará un ejercicio utilizando los datos de la Tabla 5.1, aplicando ahora ésta última metodología, sí se conoce  $t_p$ , véase Tabla 5.4 y Figura 5.6.

**Tabla 5.4.** Parámetros de una familia de hidrogramas (T=50 años)

$H_i$	$Q_p$ ( $m^3/s$ )	$V$ ( $Hm^3$ )	$t_p$ (h)	$t_g$ (h)	$t_p/(t_g-t_p)$	$\Gamma[1+(t_p/(t_g-t_p))]$	$V_{calculado}$ ( $m^3$ )
1	2231	308.839182485148	<b>17.455</b>	<b>29.493</b>	1.45001	1.28421	308839182.485
2	2120	435.931185459339	<b>30.061</b>	<b>45.895</b>	1.89852	1.82497	435931185.459
3	1878	583.839237294971	<b>53.244</b>	<b>74.131</b>	2.54907	3.50960	583839237.295
4	1507	692.260507873469	<b>91.704</b>	<b>118.617</b>	3.40748	10.24005	692260507.873
5	1164	736.966599324117	<b>143.773</b>	<b>176.731</b>	4.36225	42.00553	736966599.324



**Figura 5.6.** Familia de hidrogramas de avenidas (Función Gamma)



---

## 5.5. Elección del método para construir el hidrograma de avenida

Finalmente se concluye que, los métodos 1 y 3, los resultados son verdaderamente similares, la diferencia es que el Método 1 emplea una ecuación de regresión con determinada correlación, por lo que disminuye la precisión de sus resultados, lo que se percibe como una desventaja, pero tiene la ventaja de que se obtiene el tiempo pico, y sus cálculos son más versátiles. El Método 3, requiere como dato de entrada el tiempo pico, tiene la ventaja de ofrecer una mayor precisión en sus resultados, ya que su procedimiento es puramente numérico.

En lo que respecta al Método 2, es una excelente alternativa, según sus resultados, la obtención del tiempo pico ofrece valores de alrededor del doble que en los métodos 1 y 3, esto se debe a que el tiempo al centroide es bastante cercano al tiempo pico.

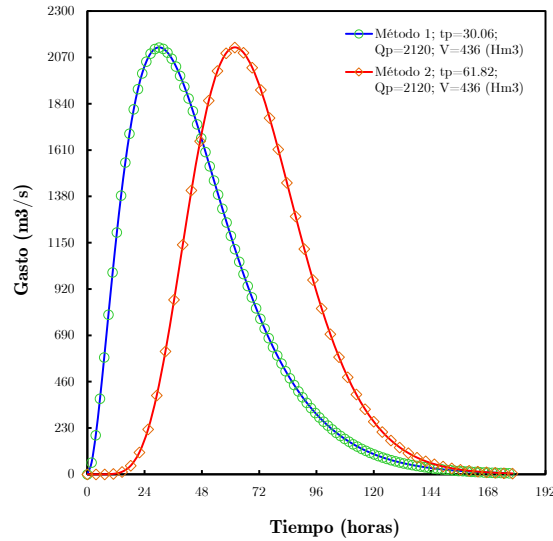
En general, la elección para construir la forma del hidrograma es probando estas tres alternativas y comparándolas con la forma que mejor describa a las avenidas en la cuenca, por ejemplo, alguna avenida que se considere más representativa entre los registros históricos de gastos máximos, por lo tanto, ésta idea es lo que dará pauta para seleccionar la mejor alternativa.

En la Figura 5.7 se comparan los hidrogramas con los métodos 1 y 2, ambos con los mismos parámetros  $Q_p$  y  $V$ , en este caso la incógnita es el tiempo pico  $t_p$ . En la gráfica se observa que la estimación de  $t_p$  conforme al Método 1 es menor que el obtenido con el Método 2, mediante esta observación no quiere decir que se descarta por completo alguno de los dos, simplemente se elige el que mejor se ajuste a un hidrograma que caracterice a una cuenca en particular.

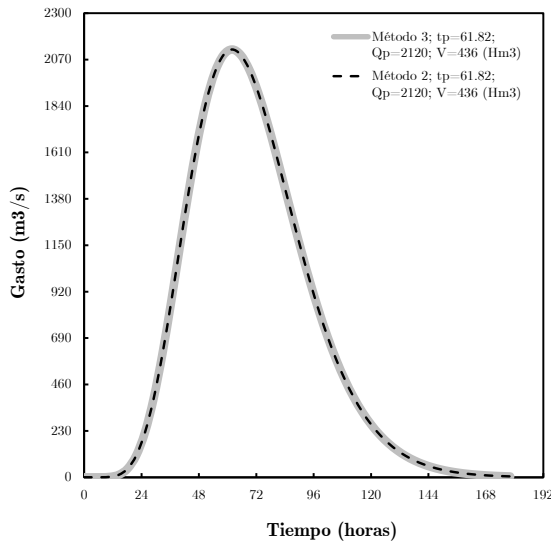
En la Figura 5.8 (a), se comparan los hidrogramas con los métodos 2 y 3, teniendo en cuenta que ambos poseen el mismo tiempo pico  $t_p$ . En la Figura 5.8 (b), se comparan los hidrogramas con los métodos 2 y 3, se muestra el efecto de los parámetros  $\alpha$  (escala) y  $\beta$  (forma) del Método 2, para ambos casos se ha fijado el tiempo pico  $t_p$ , es evidente dichos parámetros del Método 2 varían considerablemente si disminuye  $t_p$ .

(véase Tabla 5.3), para este caso es más notorio el tiempo al centroide  $t_g$ .

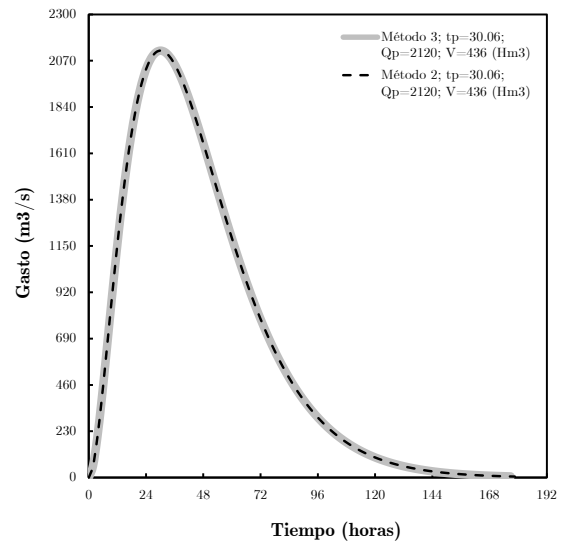
En la Figura 5.9 se muestra la comparación entre los métodos 1 y 3, usando los mismos parámetros  $Q_p$  y  $V$  para ambos casos, prácticamente son los mismos hidrogramas, de modo que  $t_p$  se obtuvo con el Método 1 y se utilizó como dato para el Método 2.



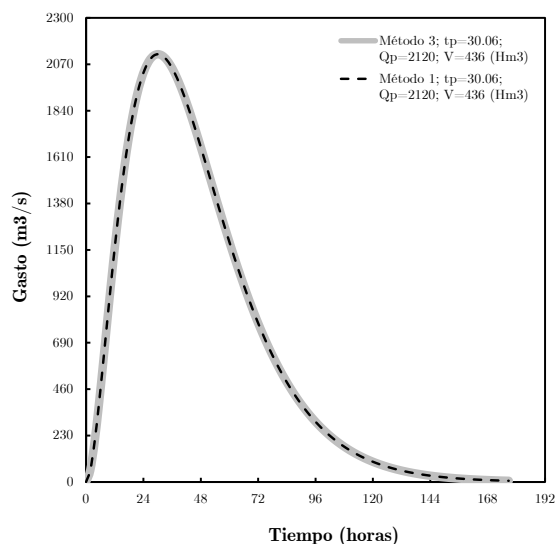
**Figura 5.7.** Comparación entre métodos 1 y 2, obteniendo (Método 2) > (Método 1);  $t_p = 61.82$  (h) >  $t_p = 30.06$  (h)



**Figura 5.8 (a).** Métodos 2 y 3, fijando  $t_p = 61.82$  (h)



**Figura 5.8 (b).** Métodos 2 y 3, fijando  $t_p = 30.06$  (h)



**Figura 5.9.** Comparación entre métodos 1 y 3

A continuación se presenta un resumen con las diferentes representaciones de la Función Gamma para construir hidrogramas, en todo momento se preserva el volumen y gasto.

**Tabla 5.5.** Hidrogramas Paramétricos con la Función Gamma (2P);  $t \in [0, \infty)$

<i>FDA Gamma</i> (2 Parámetros)	<i>Función Gamma</i> (Croley II V1) [ $t_p, Q_p, V$ ]	<i>Función Gamma</i> (Croley II V2) [ $Q_p, V$ ]	<i>Función Gamma</i> [ $t_p, Q_p, V$ ]
$Q(t) = \frac{1}{\alpha \Gamma(\beta)} \left(\frac{t}{\alpha}\right)^{(\beta-1)} \cdot e^{-\frac{t}{\alpha}}$  $t_p = \alpha(\beta - 1)$	$Q(t) = \frac{V}{\alpha \Gamma(\beta)} \left(\frac{t}{\alpha}\right)^{(\beta-1)} \cdot e^{-\frac{t}{\alpha}}$  $t_p = \alpha(\beta - 1)$	$Q(t) = \left(\frac{t}{\alpha}\right)^{(\beta-1)} \cdot e^{-\frac{t}{\alpha}} ; \frac{V}{\alpha \Gamma(\beta)} = 1$  $t_p = \frac{V}{Q_p \Gamma(\beta)} (\beta - 1)^\beta \cdot e^{-(\beta-1)}$	$Q(t) = Q_p \left(\frac{t}{t_p}\right)^{t_p/\alpha} e^{-(t_p-t)/\alpha}$
$Q_p = \frac{1}{\alpha \Gamma(\beta)} (\beta - 1)^{(\beta-1)} \cdot e^{-(\beta-1)}$	$Q_p = \frac{V}{\alpha \Gamma(\beta)} (\beta - 1)^{(\beta-1)} \cdot e^{-(\beta-1)}$	$Q_p = (\beta - 1)^{(\beta-1)} \cdot e^{-(\beta-1)}$	$\alpha = t_g - t_p ; \beta = 1 + \frac{t_p}{t_g - t_p}$
$V = \frac{1}{\alpha \Gamma(\beta)} \int_0^\infty \left(\frac{t}{\alpha}\right)^{(\beta-1)} e^{-\frac{t}{\alpha}} dt$	$V = Q_p t_p e^{(\beta-1)} \cdot (\beta - 1)^{-\beta} \cdot \Gamma(\beta)$	$V = Q_p t_p e^{(\beta-1)} \cdot (\beta - 1)^{-\beta} \cdot \Gamma(\beta)$	$V = Q_p e^{t_p/(t_g-t_p)} \left(\frac{t_g - t_p}{t_p}\right)^{t_p/(t_g-t_p)} \cdot (t_g - t_p) \Gamma\left(1 + \frac{t_p}{t_g - t_p}\right)$
	$Q(t) = Q_p \left(\frac{t}{t_p}\right)^{(\beta-1)} e^{-(\beta-1)\left(1 - \frac{t}{t_p}\right)}$		$Q(t) = Q_p \left(\frac{t}{t_p}\right)^{t_p/(t_g-t_p)} \cdot e^{[(t_p-t)/(t_g-t_p)]}$
<b>Reich y Hiemstra-Francis. Función Pearson Tipo III, [Qp, V]; <math>t \in [-t_p, \infty)</math></b>			
$Q(t) = Q_p \left(1 + \frac{t}{t_p}\right)^{t_p/t_g} e^{-t/t_g}$	$t_g = \exp\{-18.1129 + 8.336 \text{Ln}[\text{Ln}(V) - \text{Ln}(Q_p)]\}$		$V = Q_p e^{t_p/t_g} \left(\frac{t_g}{t_p}\right)^{t_p/t_g} t_g \Gamma\left(1 + \frac{t_p}{t_g}\right)$

# Capítulo 6

## Cálculo de Avenidas de Diseño: Un enfoque con Funciones Cópula

### 6.1. Introducción

En el tercer capítulo correspondiente a la Teoría de Funciones Cópula, se ha analizado en detalle el enfoque para llevar a cabo un modelo multivariado de funciones de ajuste, que hoy en día es de suma exigencia para evaluar estudios y toma de decisiones en lo que se refiere a la seguridad hidrológica de presas. En esta sección, se tiene como objetivo el desarrollo de una metodología que nos lleve a obtener un modelo bivariado, que sea capaz de proporcionar una función cópula para estimar periodos de retorno conjuntos bivariados y por consiguiente calcular la avenida de diseño más desfavorable (Gasto pico-Volumen de la avenida). Para dar respuesta con un mayor nivel de confianza a la seguridad hidrológica de cualquier presa administrada por la Comisión Federal de Electricidad (Empresa productiva del estado) y la Comisión Nacional del Agua (Órgano desconcentrado de SEMARNAT). El modelo es robusto y riguroso en el aspecto probabilístico y estadístico, estándares que demanda la hidrología moderna.

La metodología general para llevar a cabo esta investigación se describe en este capítulo. Explicaciones y diagramas resumidos en relación con las

metodologías específicas se esquematizan. Además, se incluye un resumen con la aplicación de las metodologías para cada caso de estudio considerado, y finalmente se exponen las generalidades de los softwares utilizados. La metodología aquí presentada sirve no sólo para el caso bivariado sino para el caso multivariado, puesto que ésta es una de las propiedades de las funciones cópulas.

## 6.2. Metodología General

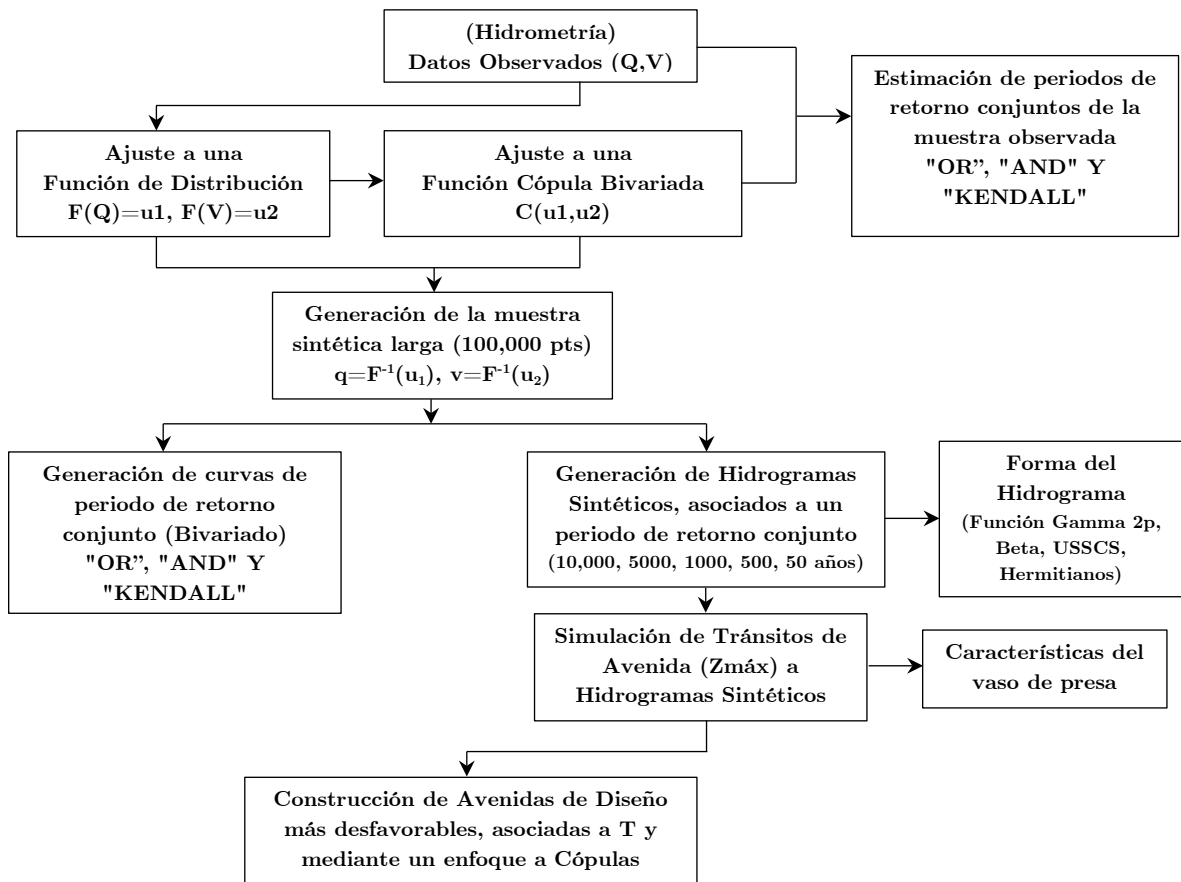
En hidrología los estudios bivariados para el ajuste de datos históricos se realizan mediante el análisis de las observaciones de las variables involucradas en una avenida o creciente, que previamente fueron medidas en una estación de aforo (hidrométrica) de una determinada cuenca hidrológica: Gasto pico máximo anual de la avenida ( $Q$ ) y su volumen asociado (estimado) a dicha avenida ( $V$ ). Con el fin de obtener la función de distribución conjunta de estas dos variables aleatorias  $Q$  y  $V$ , esto implica en primera instancia la identificación de las distribuciones marginales uniformes que mejor representan a cada variable, y a su vez, la función cópula que caracteriza la estructura de dependencia entre las dos variables  $Q$  y  $V$  (consulte la Figura 6.1).

En caso de no existir una estación de aforo, se deberá proceder a través de un estudio hidrometeorológico que determinara los hidrogramas (avenida de diseño) de entrada al embalse de la presa en estudio.

Las etapas a seguir para la obtención de: análisis de frecuencia bivariado de avenidas basado en la teoría de cópulas bivariadas; y la avenida de diseño más desfavorable basado en el nivel máximo de agua en el embalse de la presa, se detalla a continuación:

- i. Recopilación de información hidrométrica, climatológica y topográfica de la cuenca en estudio, y de contexto del conjunto de obras de la presa;
- ii. Selección de las distribuciones marginales de mejor ajuste a  $Q$  y  $V$ ;
- iii. Selección de la Función Cópula bivariada de mejor ajuste a  $(Q, V)$ ;
- iv. Generación de pares de puntos  $(q=F^{-1}(u_1), v=F^{-1}(u_2))$  de la distribución bivariada, basada y calculada mediante la función cópula de mejor ajuste;

- v. Estimación de los periodos de retorno bivariados asociados a los datos observados (tres enfoques: "OR", "AND" y "KENDALL").
- vi. Generación de las muestras bivariadas y asociadas a los periodos de retorno de diseño (10,000, 5,000, 1,000, 500, 50 y 5 años). Cálculo de volúmenes.
- vii. Generación de curvas de periodo de retorno bivariado (Q, V; T)
- viii. Generación de hidrogramas sintéticos
- ix. Simulación de Tránsitos de Avenida de los hidrogramas sintéticos
- x. Construcción de la Avenida de Diseño más desfavorable



**Figura 6.1.** Diagrama de flujo del Análisis de frecuencia bivariado de avenidas (Cópulas) y la construcción de Avenidas de diseño óptima.

---

### **6.2.1. Recopilación de información hidrométrica y contexto de la presa**

Al iniciar un estudio hidrológico para realizar un anteproyecto de planeación o proyecto de revisión, la Dirección encargada debe adquirir la información de carácter hidrológico, climatológico y topográfico (tecnología LIDAR, levantamiento con DRON) de la cuenca en estudio, así como del catálogo de planos en donde se detalla el conjunto de obras que conforman el tipo de presa en estudio, incluyendo las características del vaso de almacenamiento de ésta última, tales como las curvas E-A, E-V y E- $Q_{s_{\text{vertedor}}}$ . De ser posible, recopilar los estudios hidrológicos que se hayan elaborado con anterioridad, como informes técnicos o modelos físicos que simulen el comportamiento de la presa en un laboratorio de hidráulica.

### **6.2.2. Selección de las distribuciones marginales**

La selección de la distribución marginal univariada (Capítulo 2) para representar cada variable estudiada, es decir,  $Q$  y  $V$ , para lo cual se disponen los valores  $q$  y  $v$  respectivamente, se lleva a cabo por las siguientes etapas:

- i. Estimación de parámetros estadísticos de las distribuciones marginales,  $F(Q)=u_1$  y  $F(V)=u_2$ .
- ii. Estimación de la evaluación de bondad de ajuste a diversos modelos de distribución univariada.
- iii. Aplicación de criterios para la selección del mejor modelo de ajuste.

#### **6.2.2.1. Estimación de parámetros de las distribuciones marginales**

Dos métodos de estimación de parámetros se consideran: el método de máxima verosimilitud y el método de momentos [Laio (2004); Rao y Hamed (1999)]. En general, se aplica el método de momentos cuando se consideran muestras de longitud corta, debido a sus buenos resultados [Hosking y Wallis (1997)]; mientras que el método de máxima verosimilitud se utiliza para otros casos. En esta investigación, la

estimación de parámetros y en general el ajuste de funciones de distribución de probabilidad (FDP) univariadas, se utiliza el programa **AX.exe**, este programa cuenta con una paquetería de hasta seis FDP [Jiménez (1997)].

### 6.2.2.2. Evaluación de bondad de ajuste

Las pruebas de bondad de ajuste se aplican para identificar formalmente a las distribuciones apropiadas para el ajuste de los datos, estas son la prueba Kolgomorov-Smirnov [Klein et al. (2010)] y la prueba de Anderson-Darling transformado [Laio (2004)]. Otro método de gran aceptación en el argot hidrológico es el Error Estándar de Ajuste (**EEA**), y tiene el propósito de comparar la eficiencia del ajuste realizado a la muestra con otros modelos de FDP [Kite (1988)].

### 6.2.2.3. Aplicación de criterios para la selección del mejor modelo de ajuste de la FDP

El criterio del EEA se aplica en este trabajo para la selección del modelo y considera como la mejor distribución de probabilidad, aquella que posea el valor mínimo de EEA correspondiente. En la Figura 6.2 se muestra un ejemplo del ajuste de una FDP, obteniendo el mínimo EEA.

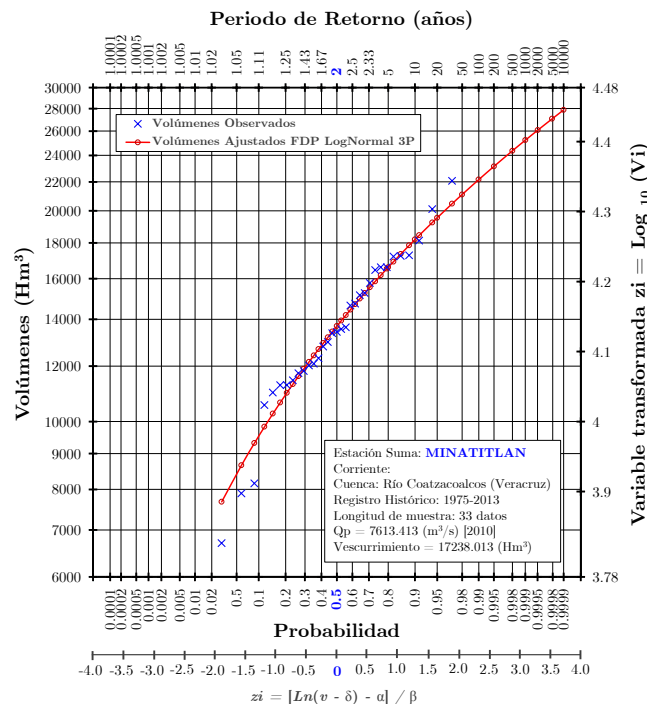


Figura 6.2. FDP de mejor ajuste a los datos observados. LogNormal 3P



---

### 6.2.3. Selección de la Función Cópula Bivariada

La selección de la mejor cópula para caracterizar la estructura de dependencia de las variables  $Q$  y  $V$  se lleva a cabo teniendo en cuenta la información proporcionada en el Capítulo 3 y se resume en los siguientes pasos:

- i. Obtención de la muestra modificada o pseudo-muestra  $(R_i/n, S_i/n)$ ;
- ii. Evaluación de las medidas de dependencia de los datos observados de la muestra bivariada:  $\rho$  de Spearman y  $\tau$  de Kendall (Figura 6.3 (a) y (b));
- iii. Estimación del parámetro de la cópula  $(\theta)$  para las distintas familias de funciones cópula en estudio;
- iv. Evaluación de la bondad de ajuste para cópulas, método de la Aproximación Checkerboard;
- v. Aplicación de criterios para la selección del mejor modelo, mínima distancia supremo  $D$ .

#### 6.2.3.1. Obtención de la muestra modificada o pseudo-muestra

Los datos originales observados de la muestra bivariada  $Q, V$  de tamaño  $n$ , se transforman por rangos de  $[0,1]$ , a la cual se le llama muestra modificada o pseudo-muestra, véase ec. 3.74, tabla 7.21, col. 7, 8, y tabla A.3 de Apéndice A. Es decir, se obtienen las pseudo-observaciones mediante los pares  $(R_i/n, S_i/n)$ , donde  $R_i$  es el rango de  $q_i$  entre el tamaño de los datos observados ( $n$ ) y de manera análoga,  $S_i$  es el rango de  $v_i$  entre  $n$ .

#### 6.2.3.2. Evaluación de las medidas de dependencia de los datos observados de la muestra

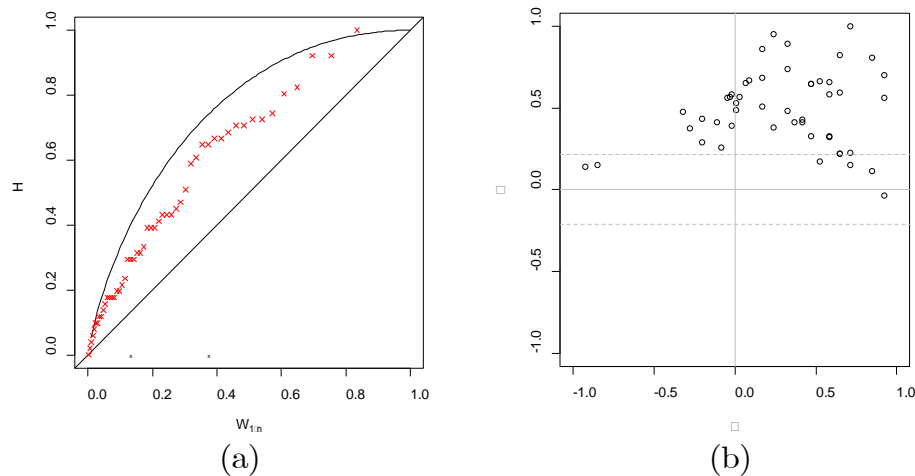
La  $\tau$  de Kendall y la  $\rho$  de Spearman están basadas en rangos, y se consideran medidas no paramétricas de dependencia, se utilizan para adquirir un valor cuantitativo de la relación de dependencia entre las variables involucradas  $Q, V$ , estas medidas se pueden obtener con la muestra bivariada original o pseudo-muestra, consulte el Apéndice A.

La estructura de la dependencia de los datos bivariados disponibles se evalúa generalmente mediante herramientas gráficas y medidas de dependencia. Las herramientas gráficas más utilizadas con frecuencia son el gráfico de dispersión de las pseudo-observaciones [Genest et al. (2009)], el Chi-plot [Fisher y Switzer (1985, 2001)] y el K-plot [Genest y Boies (2003)], consulte la Figura 6.3 (a) y (b).

**Gráfico de dispersión:** es la representación de los pares  $(R_i/n, S_i/n)$  (pseudo-observaciones). A través de este gráfico de dispersión, se da una idea del tipo de dependencia existente entre las variables.

**Gráfico Chi-plot:** muestra una medida de la ubicación de una observación  $\lambda_i$  en relación con el conjunto de observaciones, en contra de una medida de la conocida prueba estadística de independencia chi-cuadrado  $\chi_i$ . En consecuencia, cuanto mayor es la distancia entre los puntos y el valor cero en el eje vertical, mayor es la dependencia. La dependencia es positiva si los puntos están por encima del límite superior de control y negativa si se encuentran por debajo del límite inferior.

**Gráfico K-plot:** relata las estadísticas de orden estimados a partir de los datos observados  $H_i$  para el valor esperado de estas estadísticas generadas bajo la hipótesis nula de independencia entre las distribuciones marginales  $W_i/n$ . Por lo tanto, cuanto mayor es la distancia entre los puntos y la diagonal, mayor es la dependencia. Si los puntos están situados por encima de la diagonal y tienden al límite representado por una línea curva, la dependencia es positiva, y si están situados por debajo de la diagonal, la dependencia es negativa (consultar el Apéndice B).



**Figura 6.3.** Medidas de dependencia de los datos observados de la muestra bivariada Gráfico, K-plot (a) y Gráfico Chi-plot (b).

### 6.2.3.3. Estimación del parámetro de la Cópula $\theta$

La estimación del parámetro  $\theta$  se logra mediante el uso de los siguientes métodos:

- i. Método de Inversión (relación de  $\rho_s$  y  $\tau$  con  $\theta$ ).
- ii. Método de pseudo-máxima verosimilitud;
- iii. A través del paquete R, en función de  $\tau$  y  $\rho$ .

La estimación del parámetro de la cópula ( $\theta$ ) para las distintas familias en estudio, se puede efectuar mediante las relaciones que se enuncian en la Tabla 3.4 (relación  $\rho_s$  y  $\tau$  con  $\theta$ ), a ésta forma se le conoce como método de inversión y se utiliza para cópulas de un parámetro, se basa en la existencia de una ecuación, mediante la cual, el parámetro de la cópula se estima en términos de las medidas de dependencia  $\tau$  o  $\rho$  previamente obtenidas. El método de pseudo-máxima verosimilitud se aplica para cópulas de uno y dos parámetros [Chowdhary et al. (2011)].

Para el caso de las cópulas t-student y Joe/B5, el parámetro  $\theta$  con relación a  $\rho$  de Spearman, es posible estimarlo numéricamente a través del paquete R: se propone un vector de valores de  $\theta$  con intervalo entre valores relativamente pequeño, posteriormente se calculan los valores de  $\rho$ , esto se logra a partir de simulaciones y proponiendo un conjunto de 10,000 puntos de una muestra aleatoria que servirá para calcular cada valor de  $\rho$ ;

### 6.2.3.4. Evaluación de la bondad de ajuste para Cópulas

Una vez que se ha detectado dependencia entre marginales, el siguiente paso es la elección del modelo cópula que mejor se ajusta a los datos bivariados observados. Mediante la metodología propuesta en la sección 3.7 del Capítulo 3, se ha desarrollado un método de selección para el modelo cópula de mayor éxito [González-Barrios y Hoyos (2019)].

La metodología que se propone consiste en obtener y comparar las siguientes copulas; *Cópula muestral* de orden  $m$   $C_n^{(m)}$  conformada por el

conteo de puntos de la muestra de orden  $n$ , **Aproximación Checkerboard** (Cópula  $C^{(m)}$ ) de los  $r$  modelos cópula  $C_\theta$  que se pretenden estudiar, y la **Cópula real o verdadera**  $C$  que está restringida a determinados modelos propuestos. Se compara los volúmenes contenidos en las cajas  $R_{ij}$  de la *Copula Muestral*  $C_n^{(m)}$  y la *Aproximación Checkerboard*  $C^{(m)}$  **para los  $r$  modelos cópula  $C_{\hat{\theta}}$  que se pretenden estudiar**. Bajo este contexto, la comparación se realiza obteniendo la distancia supremo  $D$  de todos los volúmenes contenidos en las cajas  $R_{ij}$  entre las copulas  $C_n^{(m)}$  y  $C^{(m)}$ .

### 6.2.3.5. Aplicación de criterios para la selección del mejor modelo Cópula

La **elección del mejor modelo**  $C_{\hat{\theta}}$ , será quien posea la mínima diferencia de volúmenes  $V_{C_n^{(m)}}(R_{ij})$  y  $V_{C^{(m)}}(R_{ij})$  respectivamente (*mínima distancia supremo  $D$* ), en otras palabras, una vez que se ha **detectado** la mínima diferencia de volúmenes entre la cópula  $C_n^{(m)}$  y la aproximación  $C^{(m)}$  **de todo el conjunto de familias**, es evidente deducir que los volúmenes de la cópula del modelo  $C_{\hat{\theta}}$  (vinculados con  $C^{(m)}$  posterior a la **detección**) poseen un grado importante de aproximación con los volúmenes de la cópula  $C^{(m)}$  elegida y  $C_n^{(m)}$ .

En el momento cuando ya se ha elegido al mejor modelo Cópula, se está en condiciones de poder generar muestras de distinto tamaño (extensión de la muestra).

### 6.2.4. Generación de pares de puntos de la distribución bivariada ( $\mathbf{q}=\mathbf{F}^{-1}(\mathbf{u}_1), \mathbf{v}=\mathbf{F}^{-1}(\mathbf{u}_2)$ )

Para la realización de análisis de frecuencia bivariado de avenidas, es necesario contar una longitud de datos extensa, por lo tanto, una gran muestra sintética consiste en generar  $n$  observaciones de pares de puntos

$(q, v)$ , a partir de los datos observados y medidos en la estación de aforo  $(Q, V)$ , cuyas marginales ya ajustadas son respectivamente  $F(Q)=u_1$  y  $F(V)=u_2$ , las cuales conforman la función cópula bivariada  $C(u_1, u_2)$  seleccionada previamente, a partir de esta función y con apoyo de un algoritmo se logra generar una muestra aleatoria de tamaño  $n$  [Devroye (1986)] y [Johnson (1987)].

Basándonos en el teorema de Sklar, se necesita generar parejas  $(u_1, u_2)$  y mediante las expresiones  $\mathbf{q}=F^{-1}(u_1)$  y  $\mathbf{v}=F^{-1}(u_2)$  (inversas de las FDP) se logra transformar los conjuntos  $(u_1, u_2)$  en  $(q, v)$  respectivamente, manteniendo las propiedades de los datos observados, como son los parámetros estadísticos de las marginales ajustadas, los cuales se emplean para calcular las inversas de las FDP.

Lo establecido anteriormente se resume en las siguientes etapas:

- i. Una muestra sintética de longitud larga de pares  $(u_1, u_2) \in (0,1)$  se genera a partir de la cópula ajustada a los datos observados (se recomienda de 25,000 a 100,000 pares de puntos).
- ii. El conjunto de valores obtenido de la muestra extendida  $(u_1, u_2) \in (0,1)$  vistas como probabilidades, deben de ser transformados como cuantiles  $(q, v)$  a través de las funciones inversas de las FDP de cada marginal, empleando los mismos parámetros estadísticos de las marginales ajustadas. Mediante este procedimiento se logra conseguir las unidades originales de cada variable, que para este caso es el Gasto pico y Volumen de la avenida.

### **6.2.5. Estimación de los periodos de retorno bivariados asociados a los datos observados**

Conforme a los estudios que actualmente se están elaborando en el tema, en esta investigación se consideran tres maneras de cuantificar el período de retorno bivariado (Capítulo 4, ecs. 4.20-4.23 y 4.26-4.27, expresadas en términos de  $u_1$  y  $u_2$ ), ya sea asociándolo a los datos observados o para calcular los eventos de diseño (extremos) y vinculándolos con determinado periodo de retorno (10,000, 5,000, 1,000, 500, 50 y 5 años),

estos son: " $T_{OR}$ " (primario), período de retorno en relación con la probabilidad de excedencia de una u otra variable,  $Q$  o  $V$ ; " $T_{AND}$ " (primario), período de retorno en relación con la probabilidad de excedencia de ambas variables,  $Q$  y  $V$ ; y período de retorno mediante la función Kendall (secundario), que es el tiempo medio entre llegadas de un evento con un período de retorno primario superior a un umbral. Cabe señalar que este último enfoque es aplicable a cópulas arquimedianas [Salvadori y De Michele (2004)].

### **6.2.6. Generación de muestras bivariadas y asociadas a distintos periodos de retorno de diseño. Cálculo de volúmenes**

Para obtener las muestras bivariadas que deberán estar asociadas a diferentes periodos de retorno conjuntos (10,000, 5,000, 1,000, 500, 50 y 5 años), se procede a lo siguiente:

- i. Los valores de la variable de gastos pico  $Q$  de la muestra extendida se fijan, y utilizando las ecuaciones 4.21, 4.23 y 4.27, es posible obtener el par de valores  $(Q, V; T_{CONJUNTO})$  para un determinado periodo de retorno conjunto.
- ii. Obtener los volúmenes asociados  $V$ , de forma numérica (se recomienda el método de bisección).
- iii. El procedimiento anterior se realiza para todos los periodos de retorno de diseño (10,000, 5,000, 1,000, 500, 50 y 5 años).

### **6.2.7. Generación de curvas de periodo de retorno bivariado $(Q, V; T)$**

La generación de muestras bivariadas asociando al gasto pico y volumen de la venida de un periodo de retorno, conforman las curvas de período de retorno bivariado  $(Q, V; T)$ . En otras palabras, los conjuntos de pares  $(q, v)$  ubicados en una curva de nivel de la cópula  $C(u_1, u_2)$  están asociados con un valor de período de retorno de diseño definido previamente, en este sentido, dicha curva de nivel tendrá la misma probabilidad de ocurrencia. Como es de observarse, una curva de nivel estará conformada por una infinidad de soluciones, en este caso pares de puntos  $(Q, V; T)$ , que a diferencia del enfoque de periodo de retorno univariado, únicamente se tiene una solución. Una de las interrogantes

que se tiene es preguntarse ¿Es posible obtener una solución de eventos extremos asociado a un periodo de retorno conjunto bajo el enfoque de análisis de cópulas bivariadas?, la respuesta es sí. Éste problema se ha solucionado mediante una simulación del tránsito de avenida por el embalse de la presa ( $z_{\text{máx}}$ ), planteamiento que se abordará más adelante.

### 6.2.8. Generación de hidrogramas sintéticos

Una vez generada una muestra suficientemente grande de parejas ( $Q, V; T_{\text{CONJUNTO}}$ ) capaz de representar la variabilidad del fenómeno que se pretende estudiar y asociada a un periodo de retorno de diseño, se tiene que transformar en hidrogramas. Nótese que cada pareja ( $Q, V; T_{\text{CONJUNTO}}$ ) representa un hidrograma de avenida con determinado periodo de retorno, de tal forma que se pasa de un conjunto de parejas ( $Q, V; T_{\text{CONJUNTO}}$ ) a un hidrograma sintético. En este sentido, para poder emitir un juicio de cómo debería ser la forma del hidrograma de entrada, generalmente lo que se realiza es: analizar los registros históricos de gastos (datos observados), y se selecciona un conjunto de avenidas observadas más representativas del registro histórico, por ejemplo las avenidas máximas, de este conjunto seleccionado se analiza la forma que presentan estos hidrogramas, por esta razón, para atribuir la forma de los hidrogramas sintéticos de cada par ( $Q, V$ ) que se emplearán en el análisis, el procedimiento siguiente es comparar la forma de la avenida observada que se ha seleccionado como la más crítica con la forma de un hidrograma sintético construido mediante una función de distribución, en el Capítulo 5 se presenta la descripción general de la Función Gamma  $2p$  como probable hidrograma sintético, más sin en cambio, se da la recomendación de hacer la comparación con otras metodologías como: la Función Beta, USSCS, y los Hermitianos basados en los polinomios de Hermite [Ramírez y Aldama (2000)].

La descripción general para construir un hidrograma sintético es el siguiente:

- i. Se generan muestras bivariadas ( $Q, V; T_{\text{CONJUNTO}}$ ) asociadas a diferentes periodos de retorno y calculadas con base en el mejor modelo cópula.
- ii. Se selecciona la forma del hidrograma de avenida observado con base en el registro histórico (avenida más desfavorable).

- iii. Construir los hidrogramas sintéticos, tomando en cuenta la relación más cercana en cuanto a la forma geométrica del hidrograma de avenida observado. Se recomienda comparar diversos criterios de hidrogramas sintéticos, y seleccionar aquel cuya similitud sea lo más aproximada al hidrograma histórico observado.
- iv. La cantidad de hidrogramas sintéticos que se generaran corresponde al tamaño de la muestra extendida (de 25,000 a 100,000), esto con el propósito de obtener una muestra grande para tener un amplio margen de análisis en el estudio de períodos de retorno de eventos extremos.
- v. En el caso que se desee construir un hidrograma sintético con la forma de la Función Gamma 2P, es necesario verificar si se cuenta con el **tiempo pico  $t_p$**  en la zona de estudio, ya que de esta información dependerá aplicar uno de los escenarios que se presentan en el Capítulo 5.

### **6.2.9. Simulación de Tránsitos de avenidas de los hidrogramas sintéticos**

Para el diseño hidrológico de vasos (obras de excedencia y capacidad de control de avenidas), es necesario conocer la evolución del flujo en el tiempo. El proceso para la determinación de la evolución del almacenamiento y la elevación de la superficie libre del agua en un vaso, dado un hidrograma de avenida a la entrada del mismo, se conoce como tránsito de avenidas en vasos. Algunos modelos para su solución pueden ser consultados en las publicaciones de Arganis (1998), y de Ramírez y Aldama (2000).

El conjunto de hidrogramas de avenida sintéticos se transitan a través del embalse para evaluar el riesgo de sobrevertido de la presa, y su solución está en función del nivel máximo de agua ( $Z_{m\acute{a}x}$ ) alcanzado durante el proceso de tránsito de la avenida. La simulación de los Tránsitos de Avenida, toma en cuenta las características del vaso de almacenamiento (curva E-V) y la obra de excedencia de la presa (curva E-Qs, elevación y longitud de la cresta del vertedor). La cantidad de hidrogramas sintéticos a transitar por el vaso de la presa corresponde al tamaño de pares de puntos de Gasto pico – Volumen de la avenida ( $Q, V$ ) de la muestra extendida, aproximadamente un conjunto de 25,000 a



100,000 hidrogramas sintéticos son transitados. Los hidrogramas con diferentes combinaciones de Gasto pico y Volumen de la avenida pueden alcanzar el mismo nivel máximo de agua en la presa, lo que implica un riesgo similar de sobrevertido en la presa; en consecuencia, se puede considerar que tienen el mismo período de retorno [Requena (2015)].

Las simulaciones de los tránsitos de avenida en embalses se ejecutaron con el programa **TRATE.bas**, proporcionado por el Instituto de Ingeniería de la UNAM, cuya versión del código fuente se encuentra programado en lenguaje QuickBASIC 4.5. El manejo del programa puede ser consultado en Ramírez, L. (2011, p. 111). Para la obtención de las múltiples simulaciones de tránsitos de avenida, el código fuente original (QuickBASIC 4.5) se tradujo a lenguaje de programación de MATLAB.

El procedimiento para la simulación de Tránsitos de avenidas de los hidrogramas sintéticos es:

- i. Contar con los hidrogramas de entrada al vaso y su tiempo pico.
- ii. Niveles en el vaso de la presa: NAMO, NAME, Bordo Libre
- iii. Curvas E-V y E-Qsalida
- iv. Intervalo de tiempo  $\Delta t$ .
- v. Selección del conjunto de valores  $Q$ ,  $V$  más desfavorables, estimados en función del nivel máximo alcanzado en el embalse de la presa ( $Z_{\text{máx}}$ ).

### **6.2.10. Construcción de la Avenida de Diseño más desfavorable**

Para prevenir que la cortina de una presa sea sobrepasada, se diseñan estructuras de desfogue para el desalojo del exceso de agua, llamadas Obras de Excedencia. Estas estructuras son esenciales para el funcionamiento y seguridad de una presa y en muchas ocasiones son estructuras separadas de la cortina. Las obras de excedencia están formadas generalmente por vertedores de cresta libre o controlada.

El proceso de tránsito de la avenida depende de las características y tipo de obra de excedencias, así como de su política de operación, y de la topografía del vaso de almacenamiento

La estimación de avenidas para el diseño de vasos consiste en la determinación de las características del hidrograma de entrada al vaso, de tal forma que dicha avenida tenga una probabilidad de ser igualada o excedida.

En este trabajo se presenta la aplicación del análisis de frecuencias bivariado para el diseño de la capacidad de control de presas. El proceso se ha llevado al punto en el que se dispone de un método para determinar el hidrograma de diseño, descrito por tres parámetros  $Q$ ,  $V$ , y  $t_p$ , para un periodo de retorno conjunto dado y basado en el análisis de funciones cópula, todo ello se realiza para un vaso en particular.

Con el objeto de determinar el hidrograma de entrada (Avenida de Diseño) que produzca los efectos más desfavorables sobre el vaso, se plantea un problema de optimización no lineal, en otras palabras, se trata de simular múltiples tránsitos y seleccionar aquel que resulte con el máximo nivel en la superficie libre del embalse.

En resumen, la construcción de la avenida de diseño más desfavorable está caracterizada por tres parámetros de gran importancia ( $Q_p$ ,  $V$ ,  $t_p$ ), además deberá estar asociado a un periodo de retorno  $T$ . Se comienza estimando a  $t_p$  y con la combinación de parejas de valores ( $Q$ ,  $V$ ;  $T$ ) que se obtienen de la muestra extendida se genera el hidrograma, éste se transita por el vaso de la presa en función del tiempo para lograr obtener el máximo nivel en la superficie libre del embalse, este procedimiento se realiza para un determinado periodo de retorno y bajo un enfoque bivariado mediante un modelo cópula.

### **6.3. Software disponible**

Últimamente la comunidad estadística está proporcionando una gran cantidad de funciones para ayudar en la realización de los análisis de frecuencia de avenidas a través del software R (<https://www.r-project.org/>), lo que implica un entorno de programación libre para realizar cálculos y gráficos estadísticos. Las funciones que proporciona este software se actualizan con regularidad, y su manejo se puede consultar en los tutoriales del mismo proyecto R o en papers científicos que proporcionan descripciones e ilustraciones.

Un resumen de los R-packages usados como base para la realización del presente trabajo de investigación se incluye a continuación:

- i. **copula** (Yan, 2007; Kojadinovic y Yan, 2010; Hofert y Mächler, 2011; Hofert et al., 2012). Utilizado para la construcción de la cópula. <http://copula.r-forge.r-project.org/>
- ii. **CDVine** (Brechmann y Schepsmeier, 2013). Utilizado para la construcción de los gráficos Chi-Plot y K-plot, ambos para el modelado de dependencia de la cópula. <https://cran.r-project.org/web/packages/CDVine>
- iii. **lcopula** (Leo Belzile, Christian Genest, Alexander J. McNeil, Johanna G. Neslehova). Utilizado especialmente para para los gráficos de k-plot. <https://cran.r-project.org/web/packages/lcopula>
- iv. **asbio** (Ken Aho). Utilizado especialmente para para los gráficos de Chi-Plot. <https://cran.r-project.org/web/packages/asbio>
- v. **lattice** (Deepayan Sarkar). Utilizado para visualizar gráficos en R. <https://cran.r-project.org/web/packages/lattice>

El software comercial Matlab (Matlab 2015a, The MathWorks, Inc.) también incluye algunas librerías para llevar a cabo la modelación con Cópulas, hasta este momento solo se limitan a las más populares Gaussiana, t-student, Clayton, Frank y Gumbel.

## Capítulo 7

# Aplicación del Análisis de frecuencia bivariado para estimar Avenidas de Diseño: Un enfoque con Funciones Cópula

El Objetivo primordial de esta investigación es proponer una metodología con base en un análisis estadístico n-variado, que mejore las prácticas del cálculo de avenidas diseño en el embalse de una presa, tomando en cuenta la evaluación el riesgo de sobrevertido en la misma. Derivado de esta motivación, se utiliza un modelo cópula 2-variado para obtener la distribución conjunta bivariada con variables hidráulicas de Gato pico y Volumen, a fin de obtener con base en funciones cópulas, períodos de retorno conjuntos teóricos, en este sentido, la estimación de hidrogramas sintéticos de entrada al vaso con la misma probabilidad de periodo de retorno (T) se dibujan en familias de curvas Gasto-Volumen. Los hidrogramas sintéticos se transitan a través del vaso de la presa, la simulación se define en términos de riesgo de sobrevertido de la presa, y toma en cuenta las características del vertedor. Finalmente se construye la Avenida de diseño más desfavorable, utilizando el hidrograma ya transitado que reporte el mayor volumen, entendido de otra manera, el nivel máximo de agua alcanzado en el vaso de la presa.

## 7.1. Introducción

Antes de iniciar en la elaboración de cálculos numéricos, se debe de tener una clara concepción de lo que se pretende hacer, como es caso de la disponibilidad de información hidrométrica y de contexto de la presa, así como de los recursos computacionales con que se cuente en el momento.

Mediante el sustento teórico que se ha expuesto en los capítulos anteriores, bajo esta lógica, es recomendable para diseñar algunas estructuras hidráulicas como las presas, efectuar un análisis multivariado sobre las variables que actúan en una avenida o creciente, ya que la complejidad del proceso del tránsito de flujo de agua en el embalse de una presa requiere una representación del hidrograma completo. En este sentido, el período de retorno debe definirse en términos de riesgo de sobrevertido de la presa o daños aguas debajo de la misma, donde se deba de tomar en cuenta la influencia de las características del embalse y la presa para el proceso del tránsito del hidrograma de avenida de entrada.

En caso de riesgo de sobrevertido de la presa, el nivel máximo de agua alcanzado durante el proceso del tránsito de la avenida debe usarse para definir el período de retorno. Sin embargo, la relación entre un hidrograma de entrada dado y su nivel máximo de agua no es sencilla, ya que depende del volumen del vaso, la longitud y elevación de la cresta del vertedor. En consecuencia, el proceso del tránsito de la avenida debe estudiarse en cada presa. En este sentido, varios autores han estudiado la seguridad de las presas a través de funciones cópula, aplicando esta teoría en el proceso de transitar los hidrogramas de avenida a través del vaso para evaluar el nivel máximo de agua alcanzado (por ejemplo, De Michele et al., 2005; Klein et al., 2010).

## 7.2. Caso de Estudio

La metodología propuesta para cada parte de la presente investigación (consulte la Figura 6.1) y descrita en el Capítulo 6, se aplicó en un caso práctico dado, debido a la disponibilidad de datos. La información hidrométrica utilizada se recopiló del Banco Nacional de Datos de Aguas

Superficiales (BANDAS), que es la base de datos de la Comisión Nacional del Agua (2014a), constituida como un órgano desconcentrado de SEMARNAT y máxima autoridad en el sector hídrico del Gobierno Federal de México.

### 7.2.1. Recopilación de información hidrométrica

Un total de dos casos de estudio se consideran (Figura 7.1): el primer caso se encuentra en el Sureste de México, con información hidrométrica concerniente a la cuenca del Río Coatzacoalcos; el segundo se localiza en el Noroeste de México, cuya información hidrométrica pertenece a la cuenca del Río Fuerte, específicamente al vaso de la presa Huítes. Las características de las estaciones hidrométricas de cada caso de estudio se resumen en la Tabla 7.1.



**Figura 7.1.** Localización de los casos estudio de para la investigación

Cabe destacar que la información hidrométrica alusiva al *primer caso*, **Cuenca del Río Coatzacoalcos**, únicamente se necesitará para fines didácticos, puesto que la metodología se aplicará a la selección del modelo cópula más apropiado, la cual fungirá como distribución de probabilidad bivariada que haga dependiente a las variables Q y V. Para

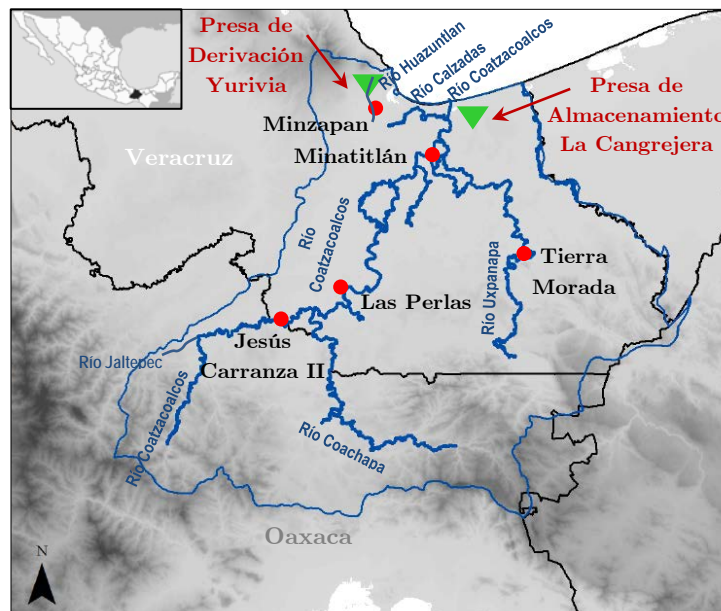
el análisis de este caso, se toma en cuenta las estaciones *Las Perlas*, *Jesús Carranza II*, *Tierra Morada* y *Minzapan*. A este primer análisis se adiciona la información hidrométrica de *Minatitlán*, como tal no es una estación de medición, simplemente se realizó la suma de las matrices de gastos medios diarios correspondientes a las estaciones Las Perlas y Tierra Morada, llamada "virtual". En la Figura 7.2 se muestra la delimitación de la cuenca del Río Coatzacoalcos, incluyendo las estaciones hidrométricas y añadiéndose la estación virtual de Minatitlán.

**Tabla 7.1.** Generalidades de las Estaciones Hidrométricas [CONAGUA (2014a)].

<i>Clave</i>	<i>Estación</i>	<i>Corriente</i>	<i>Cuenca</i>	<i>Estado</i>	<i>Coordenadas Geográficas</i>	<i>Reg. Hist.</i>
10037	<i>Huites</i> (Gastos instantáneos)	<i>Río Fuerte</i>	<i>Río Fuerte</i>	<i>Sinaloa</i>	26° 53' 00"N, -108° 22' 00"W	<b>52</b> (1941-1992)
29005	<i>Las Perlas</i> (Gastos medios/día)	<i>Río Coatzacoalcos</i>	<i>Río Coatzacoalcos</i>	<i>Veracruz</i>	17° 26' 17"N, -94° 52' 00"W	<b>57</b> (1953-2014)
29006	<i>Jesús Carranza II</i> (Gastos medios/día)	<i>Río Jaltepec</i>	<i>Río Coatzacoalcos</i>	<i>Veracruz</i>	17° 23' 30"N, -95° 03' 15"W	<b>55</b> (1953-2014)
29008	<i>Tierra Morada</i> (Gastos medios/día)	<i>Río Uxpanapa</i>	<i>Río Coatzacoalcos</i>	<i>Veracruz</i>	17° 34' 18"N, -94° 09' 47"W	<b>37</b> (1975-2014)
29009	<i>Minzapan</i> (Gastos medios/día)	<i>Río Huazuntlan</i>	<i>Río Coatzacoalcos</i>	<i>Veracruz</i>	18° 09' 36"N, -94° 47' 20"W	<b>39</b> (1974-2014)

La cuenca de Coatzacoalcos se encuentra geográficamente entre los 16°38' y 18°22' latitud norte, y los 94°11' y 95°45' longitud oeste [CONAGUA (2014b)]. Tiene un área aproximada de 21,091 km<sup>2</sup>, distribuida entre los estados de Oaxaca y Veracruz. El río Coatzacoalcos nace en el estado de Oaxaca, en la Sierra Atravesada, a una altura de 2,000 msnm; tras recorrer unos 37 km hacia el noroeste cambia su dirección hacia el oeste y la conserva hasta Sta. María Chimalapa. Aguas abajo de este poblado continúa hacia el norte a través de un cauce muy sinuoso y a la altura de Suchiapa. Adquiere una dirección NNE que conserva hasta su desembocadura en la Barra de Coatzacoalcos, junto a la ciudad de este nombre. El sistema hidrográfico del río Coatzacoalcos es el tercero de importancia en el país por su caudal, después de los sistemas Grijalva-Usumacinta y Papaloapan. Este sistema está constituido por importantes afluentes; Río Uxpanapa, Jaltepec, Coachapa y Calzadas.

La cuenca tiene un potencial hidráulico mediante dos presas: La primera está sobre el arroyo Teapa, afluente del Río Coatzacoalcos, ahí se sitúa la presa de almacenamiento La Cangrejera ( $18^{\circ} 06' 34.41''\text{N}$ ,  $-94^{\circ} 19' 50.91''\text{W}$ ), cuya finalidad es el abastecimiento de agua a los Complejos Petroquímicos Pajaritos, Cangrejera y Morelos, así como a las ciudades de Coatzacoalcos, Minatitlán, Nanchital y Allende. El área de la subcuenca del arroyo Teapa es de  $60 \text{ km}^2$  y el área de embalse de la presa La Cangrejera es de 750 hectáreas; La segunda obra hidráulica es la presa de derivación Yurivia ( $18^{\circ} 16' 13.46''\text{N}$ ,  $-94^{\circ} 46' 10.77''\text{W}$ ), se ubica en la cuenca del Río Huazuntlán y en el municipio de Tatahuicapan de Juárez [CONAGUA (2014c)].



**Figura 7.2.** Cuenca del Río Coatzacoalcos, en el Estado de Veracruz.

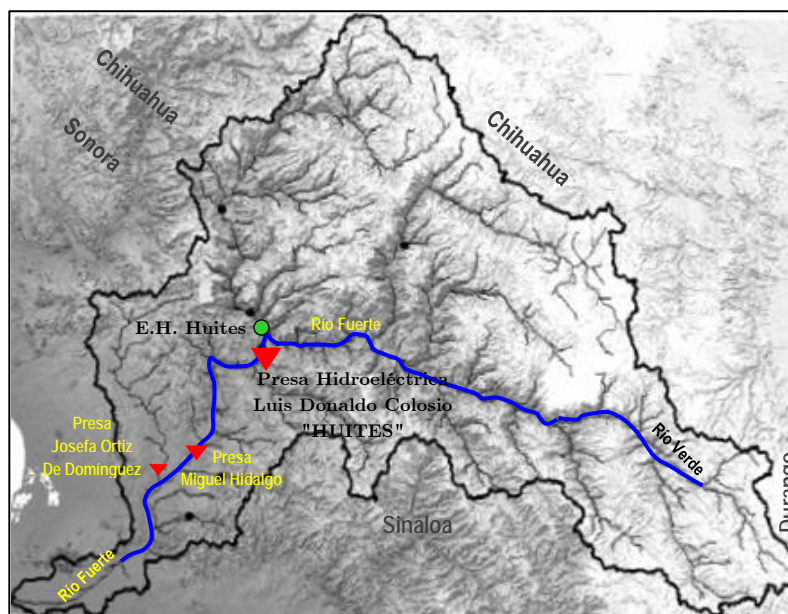
El *segundo caso* de estudio corresponde a la cuenca del Río Fuerte, especialmente para este escenario se aplicará la metodología general anteriormente planteada, con el objetivo de calcular la Avenida de Diseño y los eventos extremos que pudiesen ocurrir en determinado periodo de retorno en el vaso de la **Presa Hidroeléctrica Luis Donaldo Colosio** menor conocida como "**Huites**" (Figura 7.8), bajo un enfoque bivariado conforme a la teoría de Funciones Cópula.

La cuenca del Río Fuerte pertenece a la Región Hidrológica No. 10; se localiza al noroeste de la República Mexicana entre los paralelos  $25^{\circ} 55'$  y



28° 17' de latitud norte y los meridianos 106° 14' y 109° 02' de longitud oeste, abarca un área de aproximadamente 33,590 km<sup>2</sup> dentro de los estados de Chihuahua (24,574 km<sup>2</sup>), Sinaloa (5919 km<sup>2</sup>) Sonora (2570 km<sup>2</sup>) y Durango (527 km<sup>2</sup>) [CNA (1997)], la generalidades de la cuenca en estudio se pueden consultar en la Figura 7.3.

La presa Huites está ubicada en el sitio denominado Huites, al noroeste de la República Mexicana en el estado de Sinaloa, aproximadamente a cinco kilómetros aguas abajo de la confluencia del río Fuerte con el Río Chínipas, la presa colinda con los límites de los estados de Sonora y Chihuahua. La Presa Huites forma parte del sistema de desarrollo agrícola y energético del Río Fuerte, en el norte del estado de Sinaloa, véase Figura 7.5. Adicionalmente se encuentran dos almacenamientos de usos múltiples: las presas Miguel Hidalgo y Josefa Ortiz de Domínguez.



**Figura 7.3.** Cuenca del Río Fuerte, en el Noroeste de México

En la Figura 7.4 se esquematizan las Subcuenclas que conforman la cuenca del Río Fuerte, para el estudio que se llevará a cabo, se utilizará la información hidrométrica de la Subcuencla de Huites, que en lo sucesivo se denominará la cuenca de la Presa Huites. La cuenca abarca un área de 26,020 km<sup>2</sup>, el desnivel del Río Fuerte desde su nacimiento hasta su desembocadura es de 3,100 m, con un recorrido de 540 km. Para el caso de la cuenca de la presa Huites, el desnivel hasta el sitio de la presa es de 2,950 m, con un recorrido de 330 km [CNA (1997)].

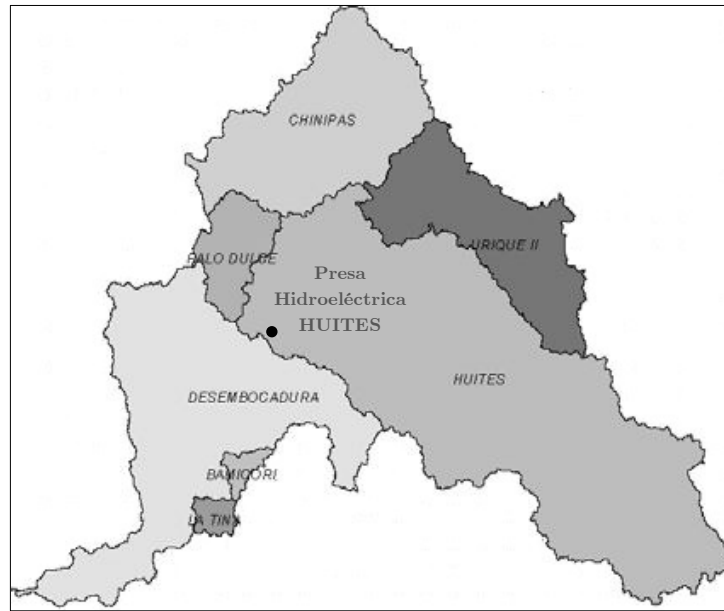


Figura 7.4. Subcuencas del Río Fuerte

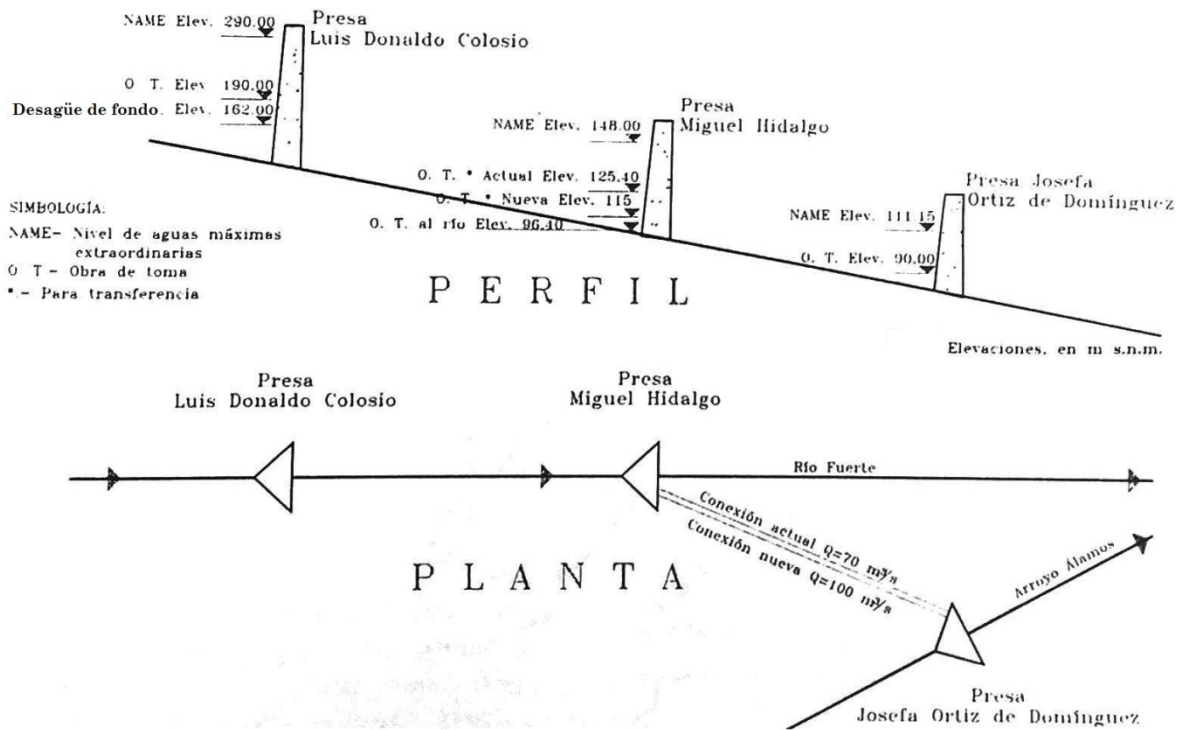


Figura 7.5. Sistema de presas en la cuenca del Río Fuerte [CNA (1997)]

Adicionalmente se obtuvieron las características fisiográficas de la cuenca en comento, por medio de un Modelo Digital de Elevaciones (MDE), las cuales son: el desnivel hasta el sitio de la presa es 2,730 m, con un recorrido de 337.783 km.

### **7.2.1.1. Información Hidrológica**

#### **▪ Cuenca del Río Coatzacoalcos**

La cuenca del río Coatzacoalcos es considerada una de las cinco con mayor incidencia de inundaciones, lo que provoca que esta zona sea considerada de alto riesgo. En la cuenca se originan constantes inundaciones que se registran en la parte baja de la región, y se deben principalmente a la continua presencia de ondas tropicales o ciclones y a que no existen suficientes obras de control. La frecuencia con que se presenta este fenómeno es anual, con fuerte impacto en las zonas urbanas e industriales de Minatitlán y Coatzacoalcos principalmente [CONAGUA (2014c)].

La precipitación media anual es de 2,450 mm, superior al promedio nacional de 777 mm (Figura 10). Las precipitaciones más elevadas se presentan en las zonas del norte de la cuenca, contiguas a la Laguna del Ostión (donde llegan a oscilar entre 3,000 y 4,000 mm). Todas las subcuencas formadas por las áreas de drenaje de los afluentes del Río Coatzacoalcos, están sujetas a un régimen muy intenso de lluvias, dando lugar a muy altos coeficientes de escurrimiento (hasta 80% y posiblemente mayores) [CONAGUA (2014c)].

La región cuenta con una red de estaciones climatológicas convencionales (que se han convertido en estaciones meteorológicas por la frecuencia con la que transmiten la información) e hidrométricas. Seis estaciones son hidrometeorológicas, es decir, monitorean tanto el agua superficial como las variables meteorológicas, tres son estaciones climatológicas y cuatro son estaciones hidrométricas [CONAGUA (2014c)].

En este estudio se utilizará la información de cuatro estaciones hidrométricas [CONAGUA (2014a)]. La Tabla 7.1 concentra el resumen de los gastos máximos anuales de cinco estaciones, con información hasta el año de 2014.

**Tabla 7.1.** Gastos máximos anuales; Cuenca del Río Coatzacoalcos.

i	Las Perlas		Jesús Carranza II		Tierra Morada		Minzapan		Minatitlán (virtual)	
	t	Qmáx	t	Qmáx	t	Qmáx	t	Qmáx	t	Qmáx
	Año	m <sup>3</sup> /s	Año	m <sup>3</sup> /s	Año	m <sup>3</sup> /s	Año	m <sup>3</sup> /s	Año	m <sup>3</sup> /s
1	1953	1947.374	1953	463.000	1975	1171.189	1974	164.662	1975	3960.814
2	1954	2181.895	1954	1043.875	1976	1271.945	1975	119.742	1976	3157.922
3	1955	3098.750	1955	844.249	1977	1419.625	1976	336.925	1978	7037.829
4	1956	2200.750	1956	1295.875	1978	1716.212	1977	94.206	1981	3924.047
5	1957	1738.249	1957	929.625	1979	1527.332	1978	156.555	1982	5375.723
6	1958	2406.000	1958	1158.125	1980	1347.556	1979	91.247	1983	4082.258
7	1959	1815.312	1959	563.437	1981	1485.977	1980	309.789	1984	4193.450
8	1960	3439.749	1960	1610.187	1982	1792.200	1981	197.514	1987	2627.356
9	1961	4101.375	1961	1464.520	1983	1704.797	1982	206.141	1988	3750.334
10	1962	2235.041	1962	919.583	1984	1409.545	1983	195.031	1989	5156.578
11	1963	6497.666	1963	1178.874	1987	1322.793	1986	175.553	1990	2628.518
12	1964	2239.448	1964	863.267	1988	1429.465	1987	178.910	1991	5632.659
13	1965	2197.403	1965	742.389	1989	1600.253	1988	91.131	1992	3145.874
14	1966	3566.507	1966	1099.945	1990	1458.875	1989	64.639	1993	3650.578
15	1967	2555.979	1967	1074.025	1991	1624.675	1990	63.021	1994	1675.660
16	1968	6082.677	1968	1903.280	1992	1234.824	1991	120.847	1995	5275.151
17	1969	5769.958	1969	1924.668	1993	1379.116	1992	131.166	1996	5110.423
18	1970	4417.460	1970	1353.315	1994	993.897	1993	152.583	1997	4153.879
19	1971	5798.591	1971	1032.827	1995	1625.988	1994	73.302	1998	4097.641
20	1972	3862.104	1972	1021.627	1996	1748.050	1995	92.768	1999	6503.965
21	1973	5099.650	1973	2908.776	1997	1633.918	1996	172.440	2000	3832.553
22	1975	2986.931	1976	709.586	1998	2961.060	1997	236.153	2001	3926.905
23	1976	2226.037	1979	1967.920	1999	4017.361	1998	87.736	2002	4642.538
24	1978	6336.781	1981	1035.372	2000	1815.872	1999	119.539	2003	5285.819
25	1981	2929.631	1982	2041.277	2001	1574.334	2000	82.235	2004	4729.402
26	1982	4245.799	1983	1065.129	2002	1356.309	2001	126.194	2005	4941.280
27	1983	3048.729	1984	954.854	2003	1317.131	2002	80.424	2006	3024.572
28	1984	3260.750	1985	679.601	2004	1387.614	2003	180.764	2008	5982.003
29	1985	4319.889	1986	1387.380	2005	1329.181	2004	52.738	2009	3650.161
30	1987	1928.050	1987	968.670	2006	1235.790	2005	186.558	2010	7613.413
31	1988	2916.014	1988	844.267	2008	2374.517	2006	141.514	2011	2731.955
32	1989	4274.984	1993	513.018	2009	1987.238	2007	283.417	2012	2032.395
33	1990	1542.900	1994	228.800	2010	1213.470	2008	99.005	2013	2014.331
34	1991	4073.784	1995	1387.100	2011	1715.450	2009	116.796	2014	1943.851
35	1992	2644.343	1996	1165.443	2012	919.033	2010	145.523		
36	1993	2582.860	1997	1034.651	2013	1427.064	2011	109.505		
37	1994	958.511	1998	662.346	2014	1267.942	2012	129.288		
38	1995	3649.163	1999	685.447			2013	142.282		
39	1996	2168.700	2000	757.476			2014	70.540		
40	1997	2586.347	2001	656.626						
41	1998	2198.073	2002	529.254						
42	1999	3353.591	2003	1091.737						
43	2000	2258.943	2004	843.185						
44	2001	3122.859	2005	316.693						
45	2002	3389.663	2006	331.694						
46	2003	4038.793	2007	239.622						
47	2004	3516.746	2008	371.301						
48	2005	3638.544	2009	609.083						
49	2006	2845.812	2010	791.234						
50	2007	1261.411	2011	504.806						
51	2008	3842.252	2012	549.551						
52	2009	2470.312	2013	411.467						
53	2010	6399.943	2014	439.955						
54	2011	1098.361								
55	2012	1228.784								
56	2013	1172.337								
57	2014	1065.888								

---

- **Cuenca del Río Fuerte**

La cuenca se caracteriza hidrométricamente por tres periodos: uno comprende los meses de mayo a octubre, en el que se presentan avenidas provocadas por precipitaciones pluviales convectivas y de tipo ciclónico; otro periodo va de noviembre a febrero, en éste periodo los escurrimientos son muy variados y los gastos máximos son provocados principalmente por los deshielos asociados a las precipitaciones (equipatas) en la zona alta de la cuenca, las cuales son erráticas en tiempo, magnitud e intensidad; finalmente, los meses de marzo a mayo representan un periodo más o menos bien definido, durante el cual, la corriente conduce gastos bajos [CNA (1997)].

El periodo de registro histórico de datos para la estación hidrométrica Huites, que abarca de 1941 a 1992, se estima un gasto medio de 132.963 m<sup>3</sup>/s en dicho periodo, y un volumen de escurrimiento medio de 4,273.544 Hm<sup>3</sup>, en la Tabla 7.2 se concentra la información de los registros de Gastos medios anuales y de Volúmenes de escurrimiento medios anuales. El 12 de enero de 1960 se presentó una avenida que registró una lectura de escala de 17.72 m; la extinta dependencia, Secretaría de Recursos Hidráulicos, indirectamente estimó el gasto pico de 15,000 m<sup>3</sup>/s, este parámetro se considera como el gasto medio diario máximo anual en la estación [CNA (1997)]. En la Tabla 7.3, se presentan los registros de gastos máximos instantáneos, y los registros de gastos máximos anuales. En la Tabla E.1 del Apéndice E, se puede consultar la información de gastos medios diarios, durante el periodo de 1941-1992 de la estación Huites. En la Figura 7.6 se muestra el hidrograma de la avenida máxima [CNA (1997)].

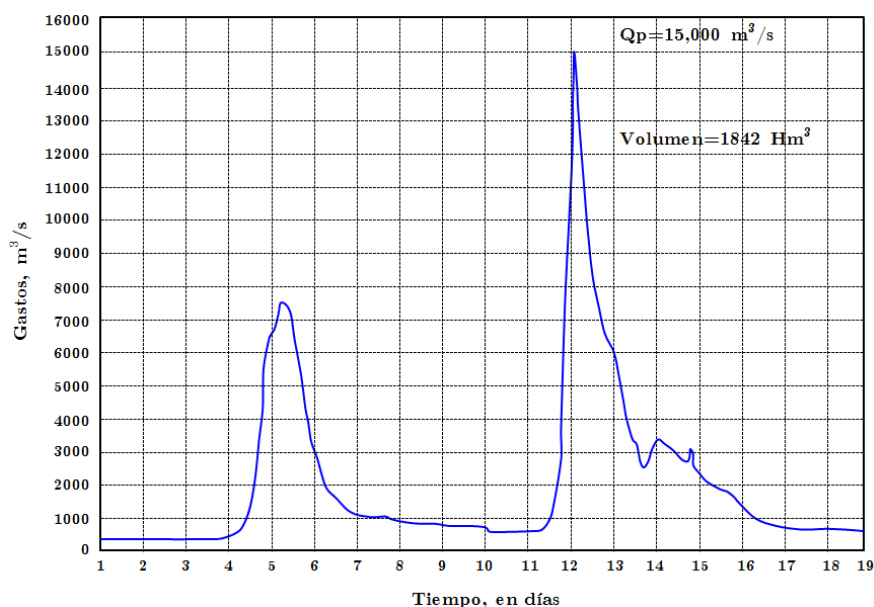
En 1975 la SRH, a través de la Dirección General de Estudios, realizó un levantamiento topográfico por métodos aerofotogramétricos del vaso de almacenamiento de Huites a escala 1:10000 con equidistancia de 5 m entre curvas de nivel, con ello se obtuvieron las *curvas elevaciones-áreas-capacidades* del vaso, las cuales se utilizaron para los estudios del funcionamiento de vaso y de tránsito de avenidas. En la Tabla 7.4, se presentan los puntos que conforman las curvas E-A-V, y en la Figura 7.7 se muestra la gráfica con las curvas. El IIUNAM ha proporcionado una actualización de las curvas E-V y E-Q<sub>s<sub>vertedor</sub></sub>, ver la Tabla 7.5.

**Tabla 7.2.** Qmed, Vmed; Huites

i	t Año	Qmed m <sup>3</sup> /s	Vesc. Hm <sup>3</sup>
1	1941	67.508	2169.763
2	1942	151.095	4856.317
3	1943	183.112	5885.353
4	1944	116.940	3758.560
5	1945	90.846	2919.875
6	1946	75.229	2417.925
7	1947	95.522	3070.163
8	1948	79.763	2563.655
9	1949	274.486	8822.188
10	1950	105.524	3391.610
11	1951	43.666	1403.458
12	1952	78.301	2516.662
13	1953	65.112	2092.750
14	1954	95.316	3063.548
15	1955	124.981	4016.976
16	1956	59.826	1922.871
17	1957	50.833	1633.806
18	1958	130.680	4200.155
19	1959	110.754	3559.716
20	1960	174.702	5615.068
21	1961	137.335	4414.050
22	1962	83.367	2679.476
23	1963	131.298	4220.025
24	1964	103.483	3326.026
25	1965	87.934	2826.283
26	1966	186.424	5991.831
27	1967	123.477	3968.656
28	1968	209.254	6725.587
29	1969	80.872	2599.277
30	1970	101.195	3252.478
31	1971	139.823	4494.029
32	1972	172.575	5546.691
33	1973	190.321	6117.058
34	1974	156.128	5018.073
35	1975	116.182	3734.184
36	1976	120.006	3857.082
37	1977	97.365	3129.395
38	1978	155.350	4993.086
39	1979	162.004	5206.951
40	1980	113.202	3638.402
41	1981	205.576	6607.391
42	1982	91.272	2933.559
43	1983	181.654	5838.496
44	1984	210.594	6768.663
45	1985	159.759	5134.781
46	1986	141.290	4541.169
47	1987	90.349	2903.902
48	1988	141.255	4540.044
49	1989	157.912	5075.411
50	1990	222.988	7167.014
51	1991	275.279	8847.678
52	1992	194.367	6247.112
Media		<b>132.963</b>	<b>4273.544</b>

**Tabla 7.3.** Qmáx, Qinst.; Huites

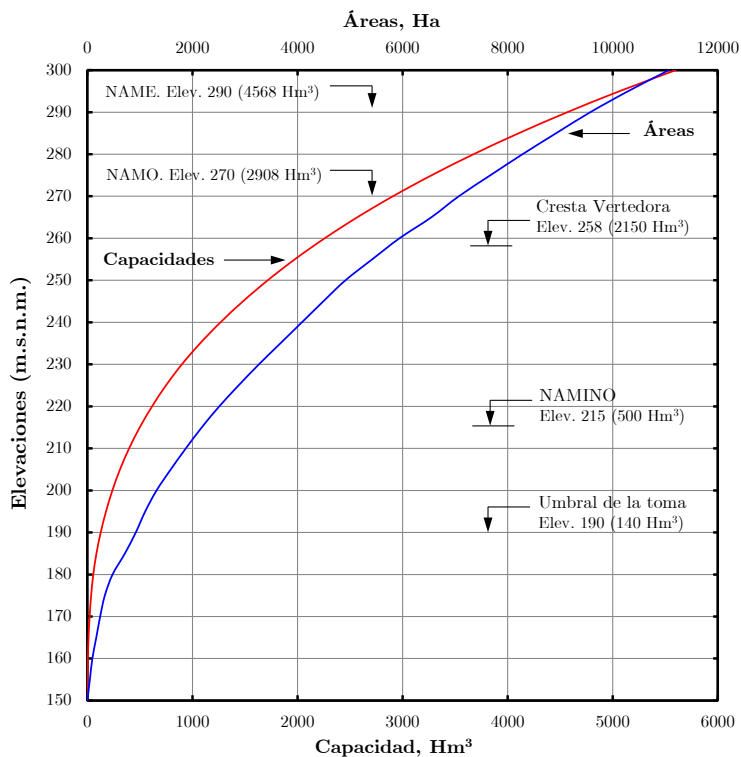
t Fecha	Qmáx m <sup>3</sup> /s	t Fecha	Qinst. m <sup>3</sup> /s
12/12/1941	1772.472	12/12/1941	2085
20/08/1942	1894.69	19/08/1942	2531
09/12/1943	8600.374	09/12/1943	14376
28/02/1944	2115.452	25/02/1944	2580
30/07/1945	1363.603	30/07/1945	1499.2
30/08/1946	842.541	30/08/1946	1164.8
29/08/1947	1001.259	29/08/1947	1127.333
05/02/1948	1820.182	05/02/1948	3215
14/01/1949	6758.359	15/01/1949	10000
13/01/1950	2722.403	13/01/1950	3229.25
14/12/1951	475.316	14/12/1951	677
14/08/1952	915.709	14/08/1952	1266
25/08/1953	728.588	22/07/1953	1025
19/08/1954	853.6	19/08/1954	954.8
15/01/1955	4007.343	15/01/1955	4780.278
23/08/1956	629.715	23/08/1956	695.7
03/09/1957	421.79	05/10/1957	593
21/09/1958	2613.1	21/09/1958	3010
03/10/1959	1627.472	03/10/1959	1908
12/01/1960	8991.811	12/01/1960	15000
24/01/1961	1314.925	24/01/1961	1396.25
05/10/1962	1176.088	05/10/1962	1620
12/12/1963	2509.375	12/12/1963	2702
02/08/1964	1133.86	01/08/1964	1319.133
03/09/1965	1414.697	23/08/1965	1944
17/08/1966	2050.475	19/08/1966	2420
16/12/1967	2193.251	16/12/1967	2505.8
01/08/1968	1411.75	01/08/1968	1534.333
17/07/1969	1084.447	17/07/1969	1508
13/08/1970	1312.732	13/08/1970	1558
26/10/1971	2018.832	26/10/1971	2200
24/11/1972	1910.898	24/11/1972	2225
22/02/1973	5801.07	22/02/1973	7960
09/11/1974	2908.613	24/09/1974	3790
05/09/1975	991.101	05/09/1975	1095
25/10/1976	1627.691	24/10/1976	2677
01/09/1977	1071.329	01/09/1977	1135
27/09/1978	3977.063	27/09/1978	4790
25/01/1979	4346.931	25/01/1979	6860
07/09/1980	1351.902	07/09/1980	1496
09/10/1981	3110.33	09/10/1981	4828.125
01/10/1982	1775.724	01/10/1982	2450
04/03/1983	6478.583	04/03/1983	8275
15/12/1984	4082.736	15/12/1984	5580
25/01/1985	2660.549	25/01/1985	3585
17/07/1986	1298.273	17/07/1986	1348.785
31/07/1987	948.496	31/07/1987	1429.237
11/07/1988	1702.549	11/07/1988	1866.324
01/09/1989	1342.256	01/09/1989	1868.488
29/12/1990	10129.264	29/12/1990	11558.610
21/12/1991	2014.055	21/12/1991	2563.147
12/01/1992	1913.774	12/01/1992	2025.341
Máximo	<b>10129.264</b>	Máximo	<b>15000</b>



**Figura 7.6.** Hidrograma de la avenida máxima observada en el Río Fuerte en la estación Huites, en enero de 1960 [CNA (1997)]

**Tabla 7.4.** Curva E-A-V

Elev. (msnm)	Áreas Ha	Capacidades $Hm^3$
150	0.00	0.00
155.00	48.40	1.20
160.00	95.50	4.80
165.00	170.40	11.40
170.00	244.10	21.80
175.00	332.80	36.20
180.00	479.90	56.60
185.00	714.60	86.60
190.00	922.20	127.30
195.00	1101.70	177.90
200.00	1317.30	238.40
205.00	1587.90	311.00
210.00	1874.60	397.60
215.00	2177.80	498.90
220.00	2510.90	616.10
225.00	2879.30	748.40
230.00	3265.40	899.50
235.00	3674.30	1073.00
240.00	4083.30	1266.90
245.00	4491.60	1481.30
250.00	4912.50	1716.40
255.00	5422.40	1974.80
260.00	5941.10	2258.90
265.00	6548.10	2569.30
270.00	7070.20	2908.10
275.00	7671.80	3276.60
280.00	8282.60	3675.50
285.00	8922.80	4105.60
290.00	9572.60	4568.00
295.00	10293.00	5064.60
300.00	11066.80	5598.60



**Figura 7.7.** Curvas E-A-V del vaso de Huites [CNA (1997)]

**Tabla 7.5.** Curva E-V y política de operación del vaso de Huites (C.F.E.)

Elev. (msnm)	Volumen (Mm <sup>3</sup> )	Q <sub>salida</sub> (m <sup>3</sup> /s)	Elev. (msnm)	Volumen (Mm <sup>3</sup> )	Q <sub>salida</sub> (m <sup>3</sup> /s)	Elev. (msnm)	Volumen (Mm <sup>3</sup> )	Q <sub>salida</sub> (m <sup>3</sup> /s)
270.0	2409.15	0.00	278.0	3017.02	7354.18	286.0	3699.18	18465.36
270.5	2446.00	1366.55	278.5	3056.91	7700.00	286.5	3745.42	18962.74
271.0	2482.86	1925.00	279.0	3096.80	9850.00	287.0	3791.66	19460.12
271.5	2519.72	2562.50	279.5	3136.68	12000.00	287.5	3837.90	19957.50
272.0	2556.57	3200.00	280.0	3176.57	12496.80	288.0	3884.13	20454.88
272.5	2593.43	3546.48	280.5	3219.58	12994.18	288.5	3930.37	20952.26
273.0	2630.28	3892.63	281.0	3262.60	13491.56	289.0	3976.61	21449.64
273.5	2667.14	4238.79	281.5	3305.61	13988.94	289.5	4022.85	21947.02
274.0	2703.99	4584.94	282.0	3348.62	14486.32	290.0	4069.09	22445
274.5	2740.84	4931.10	282.5	3391.64	14983.70	290.5	4112.30	22445
275.0	2777.70	5277.25	283.0	3434.65	15481.08	291.0	4154.85	22445
275.5	2817.59	5623.40	283.5	3477.66	15978.46	291.5	4195.64	22445
276.0	2857.47	5969.56	284.0	3520.67	16475.84	292.0	4234.48	22445
276.5	2897.36	6315.71	284.5	3563.69	16973.22	292.5	4271.24	22445
277.0	2937.25	6661.87	285.0	3606.70	17470.60	293.0	4305.89	22445
277.5	2977.14	7008.02	285.5	3652.94	17967.98			

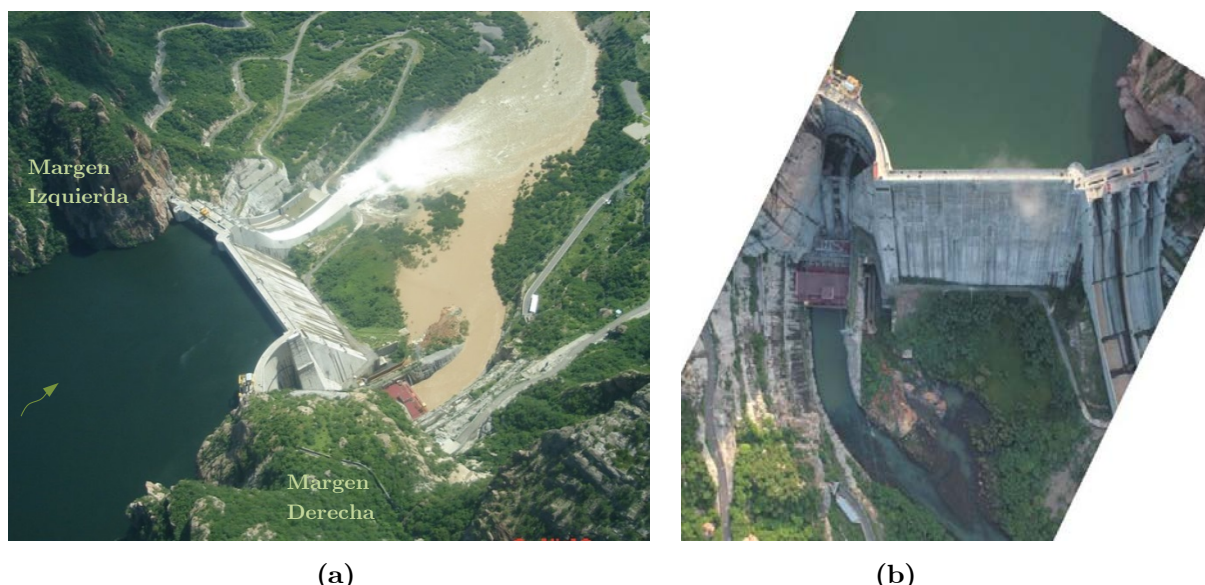
La CNA (1997), realizó un estudio hidrológico previo a la ejecución de la obra (ingeniería de detalle), entre sus generalidades destacan que; para el cálculo de la avenida de diseño se contó con información de gastos máximos instantáneos anuales (Tabla 7.3), de la estación hidrométrica Huites ubicada sobre el Río Fuerte (1942-1990); para la determinación de la avenida máxima probable se utilizó la *distribución Doble Gumbel*, ya que en el análisis estadístico univariado de los gastos máximos anuales (Tabla 7.3) se demostró la existencia de dos poblaciones; de los resultados obtenidos por el Instituto de Ingeniería, UNAM (IIUNAM) y por la Gerencia de Aguas Superficiales e Ingeniería de Ríos (GASIR) de CNA ahora CONAGUA, se determinó que con la estructura seleccionada (obra de excedencia) y su política de operación asociada, la *avenida máxima probable* con periodo de retomo de 10,000 años y gasto pico de entrada de **30,000 m<sup>3</sup>/s**, se regula un **gasto de salida de 22,445 m<sup>3</sup>/s** por el vertedor de la Presa Huites.

Se aceptó una altura de seguridad adicional del bordo libre (B.L), lo cual significó un B.L.=2 m contados a partir de la elevación del NAME (290 m.s.n.m.). De esta forma, la *elevación de la corona del parapeto es de 292 m.s.n.m.* La *elevación de la corona de la cortina* se ubicó en la **290.75 m.s.n.m.** El NAME no debía superar la elevación 290 m.s.n.m. que es la elevación de la parte más baja del puente del ferrocarril que alcanza el río aguas arriba en relación a la cortina de la Presa Huites.



### 7.2.1.2. Descripción de la Presa Hidroeléctrica Luis Donaldo Colosio "Huites"

La Presa Huites comprende un conjunto de obras que permiten controlar y aprovechar las aguas del Río Fuerte, ver Figura 7.8. Las obras consisten en una cortina de concreto tipo gravedad, un canal de desvío del río para ejecutar las obras en el cauce (recinto de la obra), un vertedor con descarga controlada, una toma y planta de generación eléctrica, y un desagüe de fondo.



**Figura 7.8.** (a) Presa Hidroeléctrica Huites en desfogue por su vertedor, (b) vista general de la presa.

#### ▪ **Obra de contención o cortina de la presa**

La cortina se compone de tres tramos diferentes: el primero, ubicado sobre el cauce principal del río, llamado cortina de sección tipo gravedad no vertedora, otro tramo sobre la ladera del macizo rocoso de la margen izquierda, llamado sección tipo gravedad vertedora y en el extremo de la margen derecha una cortina de arco-bóveda que cierra el espacio entre el extremo derecho de la sección tipo gravedad no vertedora y el macizo rocoso de la margen derecha.

#### ▪ **Obra de Desvío**

El canal de desvío se diseñó con la capacidad necesaria para conducir los escurrimientos mayores del río, establecidos con base en el estudio

hidrológico del cual resultó un gasto máximo de  $9000 \text{ m}^3/\text{s}$  para el periodo de retorno de 50 años. Las ataguías protegen el espacio necesario del recinto de la obra de construcción. Su altura se determinó de acuerdo con los niveles máximos del agua, obtenidos con el cálculo del funcionamiento hidráulico del canal empleando el gasto de diseño.

▪ **Obra de control y excedencias**

La obra de control y excedencias se sitúa en la margen izquierda y consiste en una estructura para alojar cuatro compuertas radiales de 15.50 m de ancho por 21 m de altura, sobre un cimacio que se liga a un canal de descarga que está dividido por un muro central. Con ello se define en función del sentido del flujo, por el lado izquierdo, un canal de servicio para el manejo de avenidas ordinarias (más frecuentes) y por el lado derecho, otro canal para el manejo de las avenidas extraordinarias.

Los elementos que integran el vertedor son; el cimacio, con perfil de tipo Creager diseñado para la lámina vertiente; las compuertas radiales, sus cabezales de apoyo y los servomotores que las operan; las pilas que seccionan el gasto y sirven de apoyo para el puente carretero y de maniobras; los canales de descarga, y la cubeta deflectora, consulte las Figuras 7.9. a 7.12. La cresta del cimacio está situada a la elevación 258 y tiene una longitud vertedora de 62 m, el gasto máximo de descarga por la obra de excedencias es de  $22,445 \text{ m}^3/\text{s}$ .

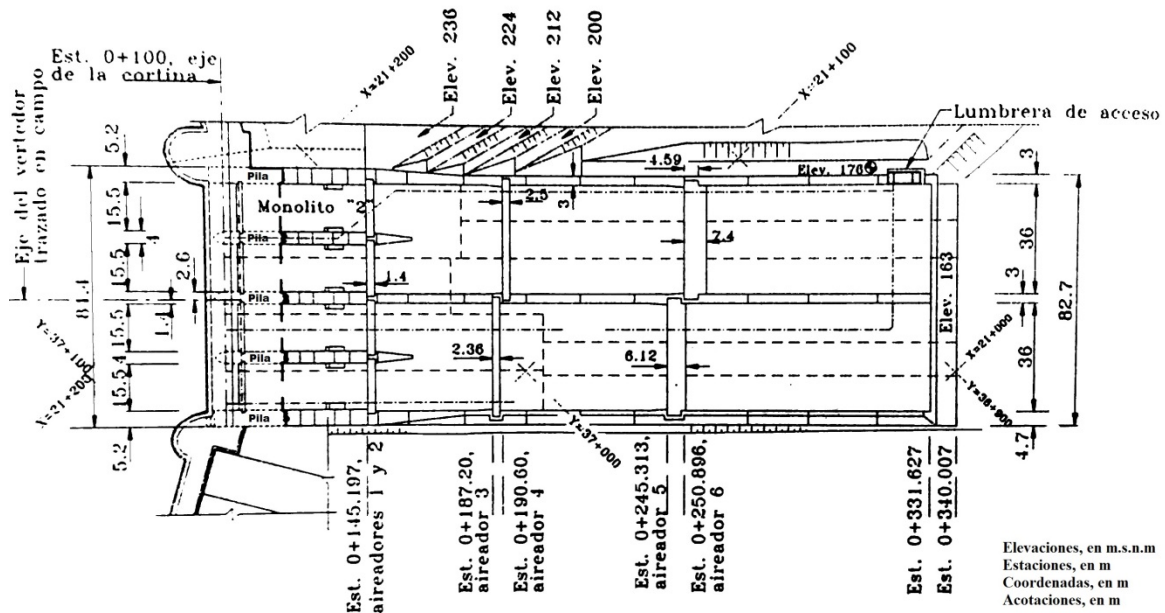


Figura 7.9. Planta del Vertedor de la Presa Huites [CNA (1997)]

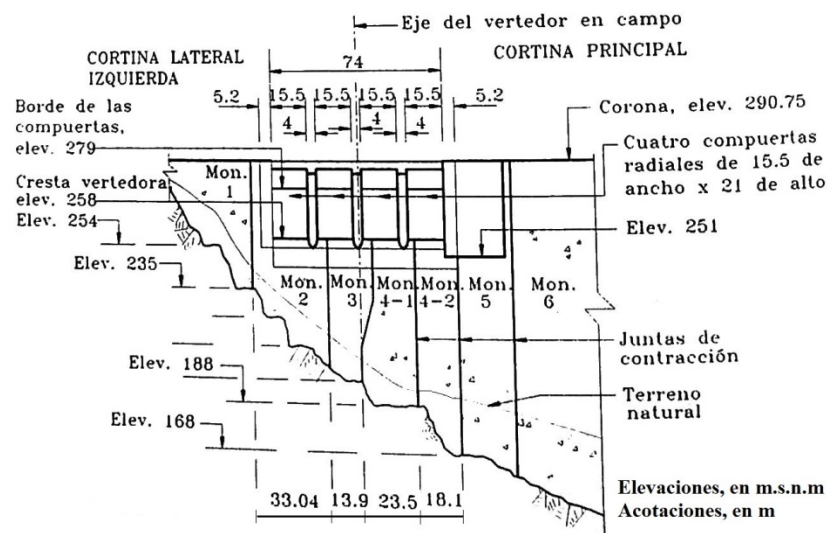


Figura 7.10. Sección longitudinal del vertedor (eje de cortina) [CNA (1997)]

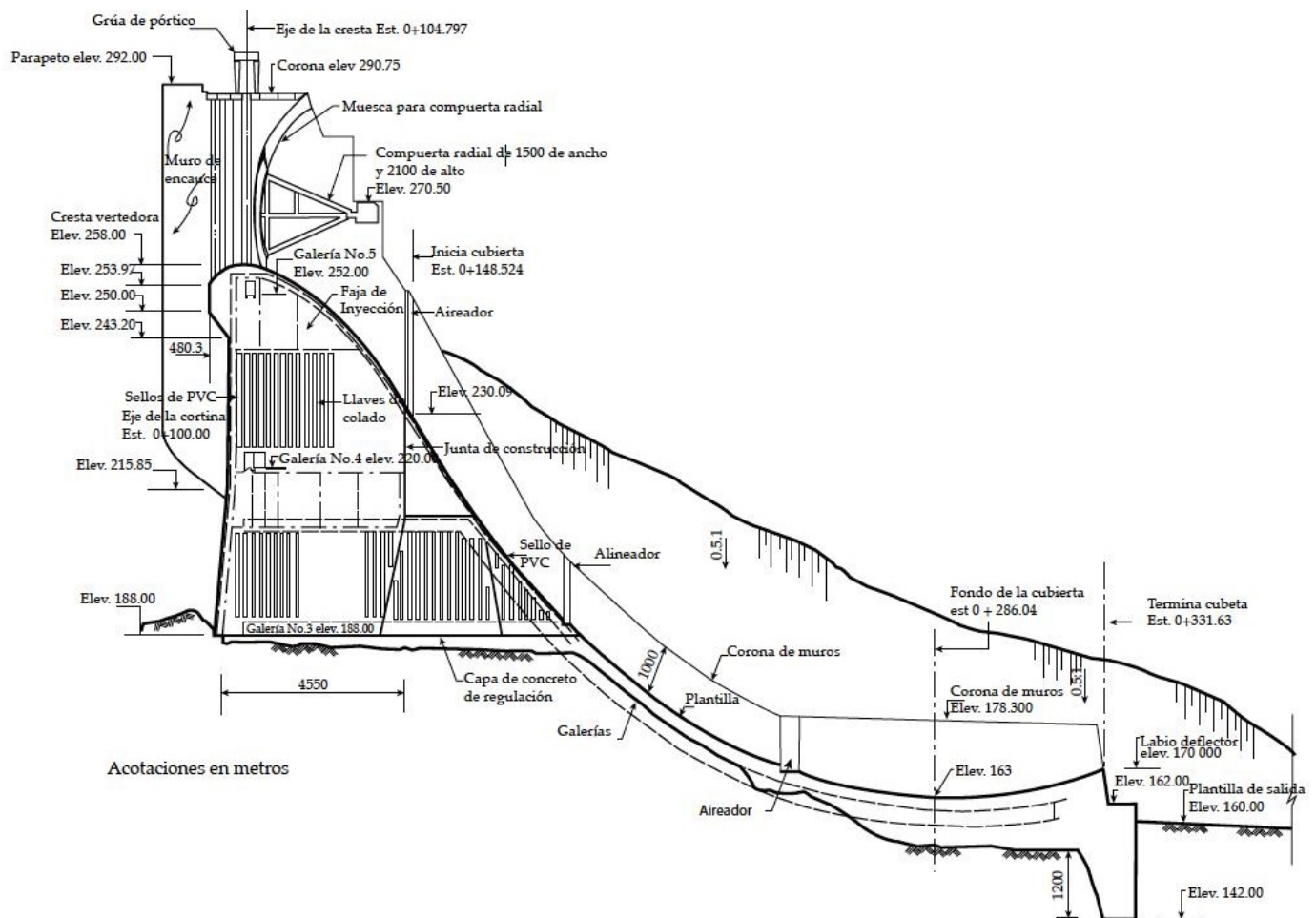
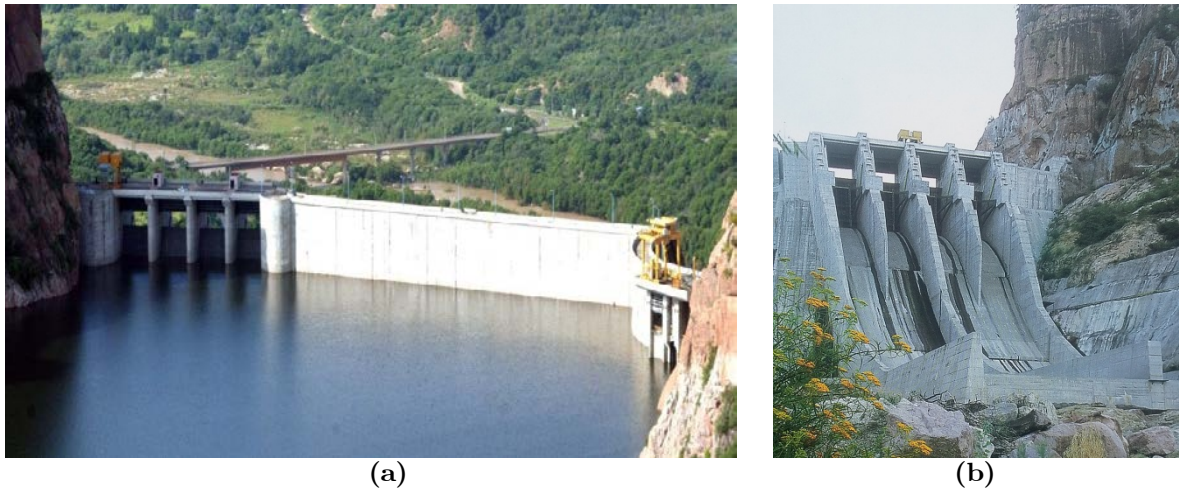


Figura 7.11. Sección transversal del vertedor por el centro del vano. [Arreguín, F. y Alcocer, V. (2011)]



**Figura 7.12.** Obra de excedencia (a) aguas arriba, (b) aguas abajo  
[Aldama *et al.* (2006)]

### ▪ **Obra de Generación**

La central hidroeléctrica de Huites está localizada en el noroeste del Sistema Eléctrico Nacional (SEN). Su operación representa un incremento del 95% de la potencia instalada y un 55% de la generación adicional en esa área.

La obra de generación se localiza en la margen derecha de la presa, al pie del macizo rocoso y aguas abajo de la cortina de arco. Está compuesta por dos unidades de generación independientes pero en una sola estructura, y se integra por: obra de toma, tubería a presión, casa de máquinas, desfogue, subestación elevadora de 230 kV y líneas de transmisión de 230 kV.

La casa de máquinas es de tipo exterior con playa de montaje en caverna; sus dimensiones exteriores son 32 m de ancho, 53 m de longitud y 42 m de altura. En ella se alojan dos grupos turbogeneradores de eje vertical, equipados con turbinas Francis de 211 MW de potencia nominal cada uno, y separados entre sí 22 m. Entonces su capacidad es de 422 MW. En la Figura 7.13, se muestran las características generales de la obra de excedencia.

En la Tabla 7.6, se resume las características generales de la Presa Huites, y en la Figura 7.14 se presenta el arreglo general.



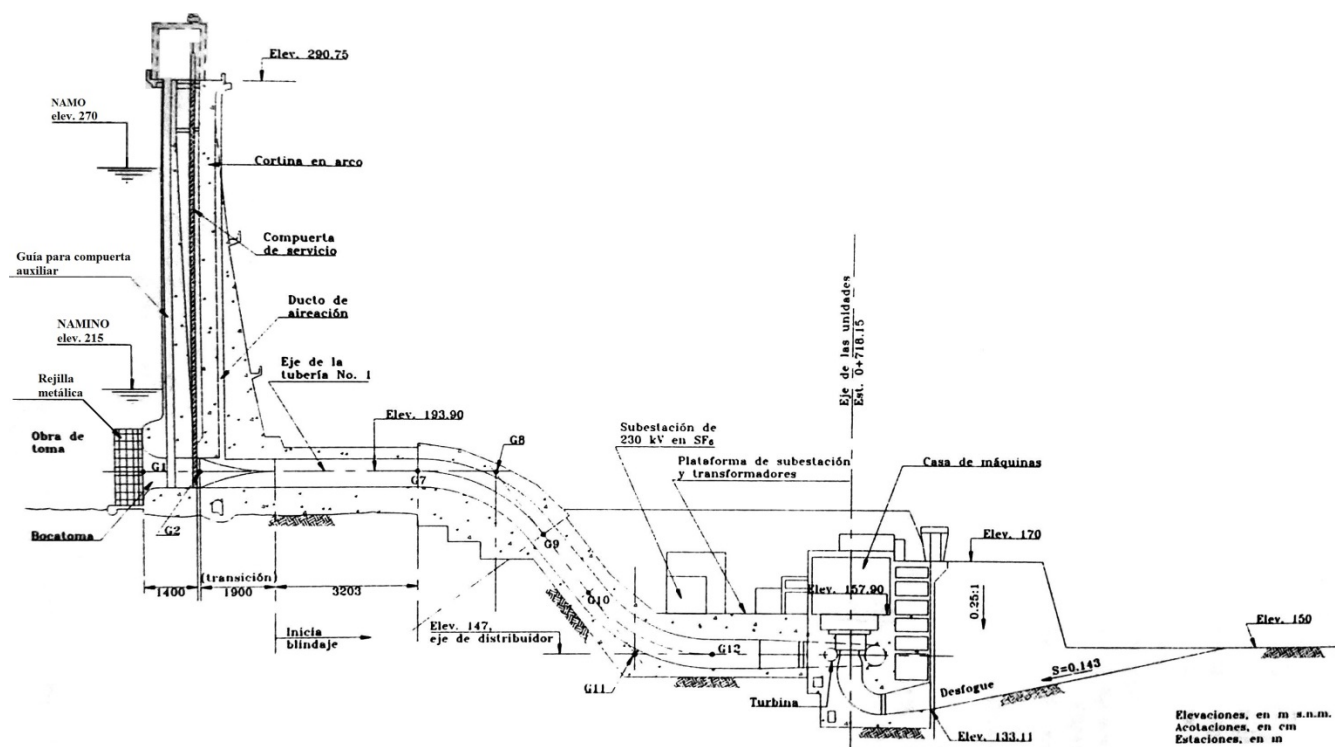


Figura 7.13. Obra de generación. Perfil eje de la toma No.1 [CNA (1997)]

Tabla 7.6. Datos generales de la Presa Luis Donaldo Colosio "Huites" [CNA (1997)]

Hidrología		Obra de excedencias	
Periodo de registro (1941-1992)	52 años	Gasto máximo de avenida de diseño	30,000 m <sup>3</sup> /s
Escurrimiento medio en el periodo	4,274 Hm <sup>3</sup>	Gasto máximo de salida del vertedor	22,445 m <sup>3</sup> /s
Gasto medio anual en el periodo	133 m <sup>3</sup> /s	Periodo de retorno	10,000 años
Gasto máximo anual en el periodo	10,129 m <sup>3</sup> /s	Longitud de la cresta vertedora	62 m
Gasto máximo instantáneo	15,000 m <sup>3</sup> /s	Carga máxima sobre la cresta	32 m
<b>Capacidad</b>		Compuerta radial de 15.5 x 21 m	4 pza.
Total del vaso al NAME	4,568 Hm <sup>3</sup>	<b>Obra de toma</b>	
Conservación al NAMO	2,908 Hm <sup>3</sup>	Gasto de diseño	470 m <sup>3</sup> /s
Muerta al NAMINO	500 Hm <sup>3</sup>	Compuertas de servicio (6.13 x 7.8)	2 pza.
Control de avenidas	1,102 Hm <sup>3</sup>	Compuertas de emergencia ( " )	1 pza.
Superalmacenamiento	558 Hm <sup>3</sup>	<b>Obra de desvío</b>	
Útil para riego y generación	2,408 Hm <sup>3</sup>	Gasto máximo de descarga (tajo)	9,000 m <sup>3</sup> /s
Cresta del vertedor	2150 Hm <sup>3</sup>	Ancho de la plantilla	28 m
<b>Elevación</b>		Elevación de la plantilla	150 m.s.n.m.
Parapeto ciego de la cortina	292 m.s.n.m.	Elev. Corona, ataguía aguas arriba	183.70 m.s.n.m.
Corona de la cortina	290.75 m.s.n.m.	Elev. Corona, ataguía aguas abajo	171.50 m.s.n.m.
NAME	290 m.s.n.m.	<b>Obra de generación</b>	
NAMO	270 m.s.n.m.	Casa de máquinas exterior	32 x 53 x 42 m
Cresta vertedora	258 m.s.n.m.	Turbogeneradores	2 (Francis)
NAMINO	215 m.s.n.m.	Capacidad del turbogenerador	211 MW c/u
Umbral de la toma	190 m.s.n.m.	<b>Obra de contención</b>	
Bordo libre	2 m	Cortina tipo gravedad de concreto	164.75 m

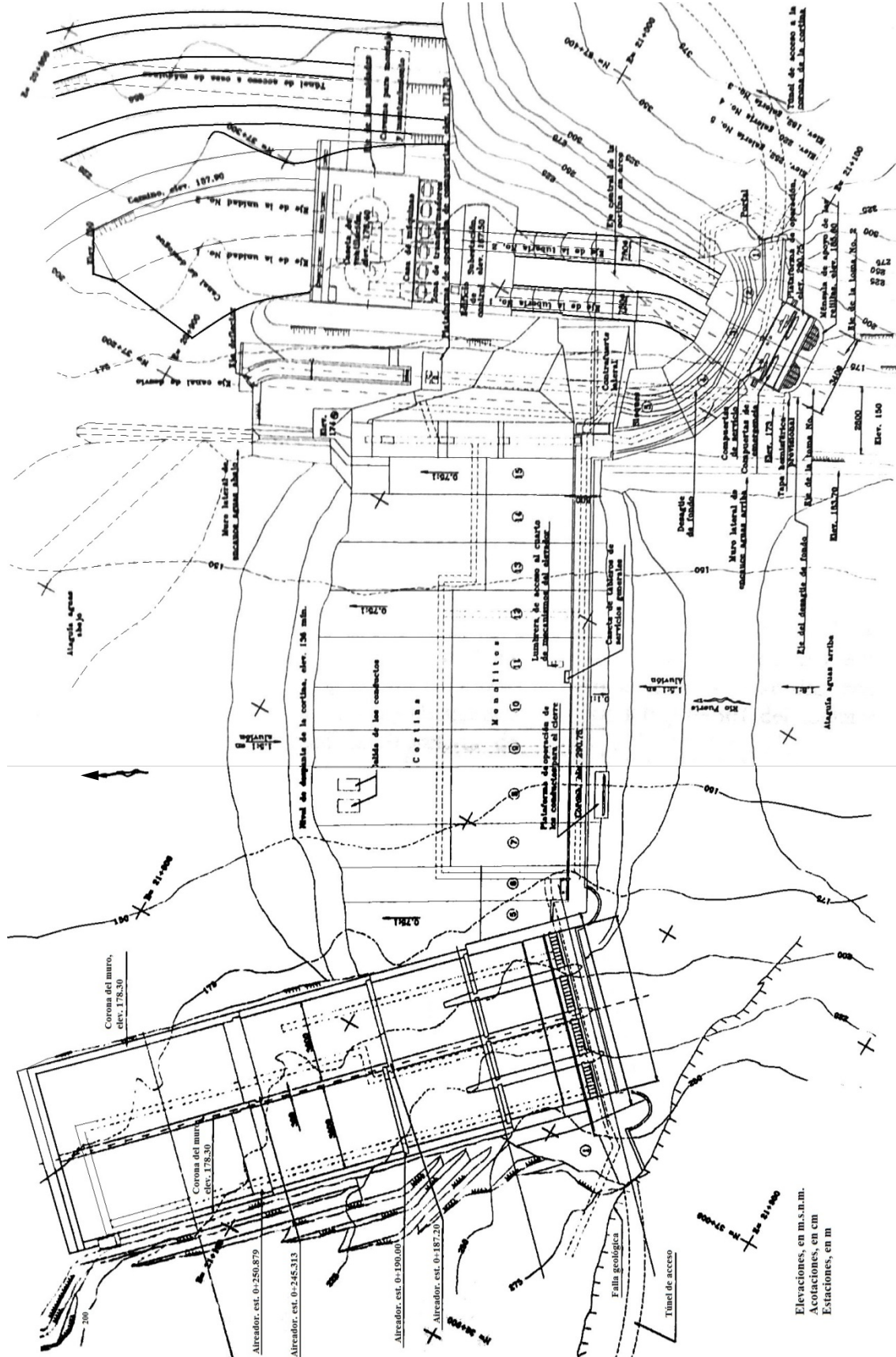


Figura 7.14. Arreglo general de la Presa Huites [CNA (1997)]

---

## 7.2.2. Cálculo de volúmenes asociados a los gastos pico

### ▪ Cuenca del Río Coatzacoalcos

El contenido de la Tabla 7.1, referente a gastos máximos anuales, fueron obtenidos de las mediciones de gastos medios diarios en las distintas estaciones, en este sentido, la forma en que se calculó el volumen de la avenida asociado a cada gasto máximo anual de cada estación hidrométrica consistió en:

- i. Emplear la información de gastos medios diarios de una determinada estación hidrométrica, ubicar el gasto máximo de cada año del registro histórico, a partir de esta referencia, se toma en cuenta 30 días de mediciones de gastos medios diarios a la derecha y a la izquierda;
- ii. Bosquejar la gráfica de  $t$  vs  $Q_{med/día}$ , o sea 60 puntos;
- iii. El área acumulada por debajo de la curva que conforman los 61 puntos de gastos, será el volumen asociado al gasto máximo de la avenida.

En la Tabla 7.7 se observa las muestras bivariadas de gastos pico (máximos anuales) con sus volúmenes asociados, dichas muestras se utilizarán para el análisis univariado (FDP) y bivariado (cópulas).

### ▪ Cuenca del Río Fuerte

Como se comentó, la información de gastos máximos instantáneos, proporcionada por GASIR-CONAGUA, no es suficiente para el análisis de frecuencia bivariado de avenidas. En la literatura, existen algunas variantes para estimar el volumen asociado a la avenida que está gobernada por un gasto pico; por ejemplo mediante filtros digitales [Eckhardt (2005)], [Ladson *et al.* (2015)], [IGME (2008)] por mencionar algunos. Para fines de esta investigación, los volúmenes asociados a los gastos máximos instantáneos se obtuvieron de los trabajos elaborados por Ramírez y Aldama (2000) y Aldama *et al.* (2006). En la Tabla 7.8, se observa la muestra bivariada de gastos pico (instantáneos) con sus volúmenes asociados.

**Tabla 7.7.** Gastos pico y Volúmenes. Cuenca del Río Coatzacoalcos

i	Las Perlas		J. Carranza II		Tierra Morada		Minzapan		Minatitlán (virtual)	
	Qmáx	V	Qmáx	V	Qmáx	V	Qmáx	V	Qmáx	V
	m <sup>3</sup> /s	Hm <sup>3</sup>	m <sup>3</sup> /s	Hm <sup>3</sup>	m <sup>3</sup> /s	Hm <sup>3</sup>	m <sup>3</sup> /s	Hm <sup>3</sup>	m <sup>3</sup> /s	Hm <sup>3</sup>
1	1947.374	4581.725	463.000	264.317	1171.189	6237.837	164.662	61.925	3960.814	12792.656
2	2181.895	5857.019	1043.875	1027.261	1271.945	4527.039	119.742	76.034	3157.922	11719.436
3	3098.750	6115.796	844.249	528.853	1419.625	3680.570	336.925	57.392	7037.829	16613.017
4	2200.750	7542.663	1295.875	1223.053	1716.212	5928.174	94.206	14.912	3924.047	15148.540
5	1738.249	5220.917	929.625	416.056	1527.332	5478.731	156.555	39.990	5375.723	20107.457
6	2406.000	7247.795	1158.125	797.909	1347.556	5116.982	91.247	28.777	4082.258	18129.470
7	1815.312	5219.094	563.437	230.696	1485.977	5365.303	309.789	83.327	4193.450	15261.417
8	3439.749	8390.354	1610.187	969.758	1792.200	5706.877	197.514	39.357	2627.356	12093.419
9	4101.375	7332.841	1464.520	911.077	1704.797	4960.950	206.141	61.518	3750.334	13368.214
10	2235.041	6651.430	919.583	650.001	1409.545	4354.193	195.031	38.272	5156.578	11453.083
11	6497.666	9504.804	1178.874	710.303	1322.793	3107.467	175.553	46.713	2628.518	8157.220
12	2239.448	8055.928	863.267	459.628	1429.465	4028.189	178.910	51.685	5632.659	12025.152
13	2197.403	5597.068	742.389	295.910	1600.253	4359.807	91.131	35.167	3145.874	13431.360
14	3566.507	8313.473	1099.945	668.909	1458.875	4872.658	64.639	16.405	3650.578	14736.010
15	2555.979	5094.033	1074.025	363.241	1624.675	4863.636	63.021	18.874	1675.660	6704.137
16	6082.677	14061.219	1903.280	885.963	1234.824	5231.154	120.847	64.333	5275.151	16606.251
17	5769.958	12055.792	1924.668	1124.572	1379.116	5146.921	131.166	27.607	5110.423	12986.947
18	4417.460	11217.620	1353.315	540.811	993.897	2878.562	152.583	65.436	4153.879	12313.410
19	5798.591	8372.399	1032.827	535.701	1625.988	7186.606	73.302	18.229	4097.641	11271.250
20	3862.104	9225.439	1021.627	662.981	1748.050	5276.912	92.768	12.643	6503.965	17211.417
21	5099.650	16356.012	2908.776	1231.761	1633.918	4366.273	172.440	51.616	3832.553	16455.574
22	2986.931	7298.026	709.586	478.387	2961.060	5989.554	236.153	41.322	3926.905	13551.066
23	2226.037	7156.266	1967.920	630.104	4017.361	8501.140	87.736	27.781	4642.538	10557.501
24	6336.781	10730.080	1035.372	479.946	1815.872	6164.973	119.539	54.658	5285.819	15771.178
25	2929.631	9253.425	2041.277	566.359	1574.334	5893.518	82.235	26.328	4729.402	17271.696
26	4245.799	14272.462	1065.129	833.070	1356.309	2883.500	126.194	35.934	4941.280	22075.669
27	3048.729	13621.620	954.854	666.248	1317.131	5563.175	80.424	14.761	3024.572	14644.227
28	3260.750	11251.216	679.601	733.054	1387.614	5033.078	180.764	45.892	5982.003	13633.401
29	4319.889	10437.004	1387.380	703.318	1329.181	7129.018	52.738	13.351	3650.161	11807.150
30	1928.050	7430.401	968.670	563.598	1235.790	3441.189	186.558	47.734	7613.413	17238.013
31	2916.014	8913.535	844.267	737.115	2374.517	6264.537	141.514	26.142	2731.955	11272.745
32	4274.984	7113.683	513.018	572.835	1987.238	5250.013	283.417	48.083	2032.395	7893.388
33	1542.900	3398.817	228.800	235.820	1213.470	4716.113	99.005	25.710	2014.331	10998.466
34	4073.784	7127.026	1387.100	556.790	1715.450	5874.413	116.796	37.464	1943.851	9513.334
35	2644.343	8705.360	1165.443	720.520	919.033	2790.896	145.523	50.598		
36	2582.860	9383.771	1034.651	812.417	1427.064	4977.438	109.505	20.501		
37	958.511	3627.363	662.346	399.939	1267.942	4023.714	129.288	31.287		
38	3649.163	9419.645	685.447	551.801			142.282	31.176		
39	2168.700	6504.538	757.476	290.698			70.540	37.452		
40	2586.347	7046.819	656.626	141.249						
41	2198.073	5771.812	529.254	219.770						
42	3353.591	8526.117	1091.737	628.348						
43	2258.943	10230.013	843.185	421.695						
44	3122.859	7845.769	316.693	201.572						
45	3389.663	7605.741	331.694	144.191						
46	4038.793	10144.786	239.622	136.322						
47	3516.746	12115.124	371.301	364.398						
48	3638.544	14949.500	609.083	262.742						
49	2845.812	11959.499	791.234	537.973						
50	1261.411	3786.419	504.806	292.760						
51	3842.252	7599.782	549.551	316.397						
52	2470.312	5922.941	411.467	478.026						
53	6399.943	12521.900	439.955	360.799						
54	1098.361	5392.886								
55	1228.784	5157.358								
56	1172.337	5786.891								
57	1065.888	5613.939								



**Tabla 7.8.** Gastos pico y Volúmenes. Estación Huites [Ramírez y Aldama (2000)]

t años	Qp m <sup>3</sup> /s	Volumen (Hm <sup>3</sup> )	t años	Qp m <sup>3</sup> /s	Volumen (Hm <sup>3</sup> )	t años	Qp m <sup>3</sup> /s	Volumen (Hm <sup>3</sup> )
1941	2085	458	1959	1908	352	1977	1135	601
1942	2531	1302	1960	15000	1842	1978	4790	1245
1943	14376	1928	1961	1396.25	689	1979	6860	986
1944	2580	871	1962	1620	437	1980	1496	1076
1945	1499.2	684	1963	2702	885	1981	4828.125	992
1946	1164.8	720	1964	1319.133	384	1982	2450	351
1947	1127.333	435	1965	1944	305	1983	8275	1625
1948	3215	344	1966	2420	2716	1984	5580	1258
1949	10000	2966	1967	2505.8	593	1985	3585	1092
1950	3229.25	644	1968	1534.333	1706	1986	1348.785	1185
1951	677	111	1969	1508	837	1987	1429.237	766
1952	1266	474	1970	1558	1001	1988	1866.324	683
1953	1025	163	1971	2200	905	1989	1868.488	428
1954	954.8	596	1972	2225	442	1990	11558.61	2930
1955	4780.278	787	1973	7960	1250	1991	2563.147	653
1956	695.7	513	1974	3790	607	1992	2025.341	601
1957	593	69	1975	1095	1768			
1958	3010	740	1976	2677	565			

### 7.2.3. Selección de las distribuciones marginales

- **Cuenca del Río Coatzacoalcos**

Las distribuciones marginales más apropiadas que representan a las variables Q y V, se eligieron con base en un análisis de ajuste de funciones de distribución de probabilidad (FDP). El ajuste se realiza empleando el programa AX.exe. Se aplicaron tres métodos para estimar los parámetros estadísticos de las funciones: momentos, máxima verosimilitud y mínimos cuadrados para la función Doble Gumbel.

En la Tabla 7.9 se exhiben las FDP que se aplicaron a las muestras bivariadas (Q, V) para los cinco conjuntos de datos observados. En esta tabla, se muestra la relación del Error Estándar de Ajuste que obtuvo cada FDP vinculada con los datos observados. Con el objetivo de validar el mejor ajuste, se aplicó la evaluación de bondad de ajuste a los modelos de distribución univariados, fijando como criterio para la selección del mejor modelo, aquel que posea el valor mínimo de EEA correspondiente.

En la Tabla 7.10, se expone un resumen con los modelos univariados de mejor ajuste a las variables Q y V (marginales), asimismo se muestran los parámetros de escala, ubicación y de forma respectivamente.

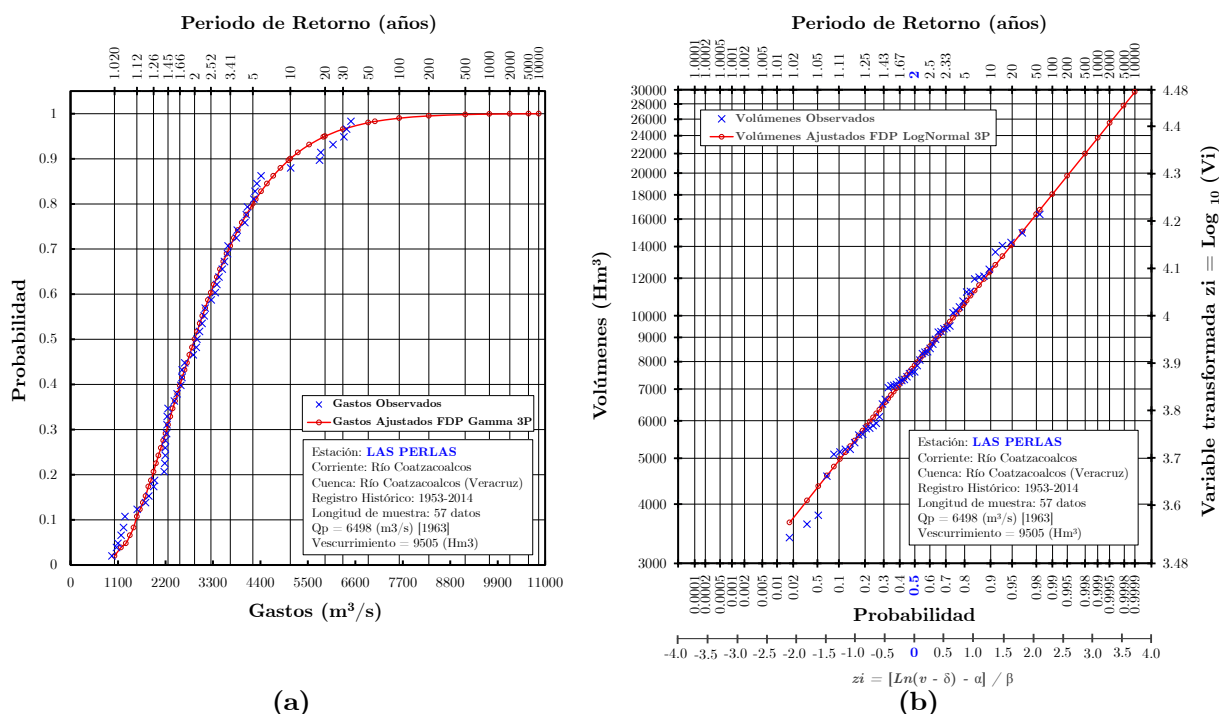
**Tabla 7.9.** Error Estándar de Ajuste de cada FDP

Est.	Var.	Método		Normal	LogNormal	Gamma	Exponencial	Gumbel	Doble Gumbel	
Las Perlas	Gasto pico	Momentos	2p	334.311	251.998	215.778	359.319	228.168	-	
			3p	-	227.114	221.252	-	-	-	
		Máxima	2p	334.311	237.973	219.836	1496.374	221.307	-	
			3p	-	241.728	208.720	-	-	-	
	Volumen	Momentos	2p	604.562	334.443	321.569	688.951	340.116	-	
			3p	-	337.931	325.008	-	-	-	
		Máxima	2p	604.562	287.374	358.341	4677.656	5736.107	-	
			3p	-	272.634	288.030	-	-	-	
	Mínimos Cuadrados			-	-	-	-	-	221.944	
	Jesús Carranza II	Gasto pico	Momentos	2p	165.916	93.972	102.236	109.839	97.631	-
				3p	-	95.394	97.855	-	-	-
			Máxima	2p	165.916	76.89	110.783	379.075	109.79	-
3p				-	75.279	99.270	-	-	-	
Mínimos Cuadrados			-	-	-	-	-	82.976		
Volumen		Momentos	2p	49.603	49.348	34.168	73.496	37.959	-	
			3p	-	33.352	32.661	-	-	-	
		Máxima	2p	49.603	58.957	31.651	249.06	31.583	-	
			3p	-	66.802	40.019	-	-	-	
Mínimos Cuadrados			-	-	-	-	-	36.843		
Tierra Morada		Gasto pico	Momentos	2p	329.274	264.373	288.295	215.786	255.369	-
				3p	329.274	211.371	223.992	-	-	-
	Máxima		2p	-	292.819	300.243	910.491	292.378	-	
			3p	-	416.684	270.077	-	-	-	
	Mínimos Cuadrados			-	-	-	-	-	58.622	
	Volumen	Momentos	2p	256.994	283.924	267.92	475.334	313.863	-	
			3p	-	258.195	5467.461	-	-	-	
		Máxima	2p	256.994	274.756	261.517	3327.897	323.051	-	
			3p	-	275.415	290.464	-	-	-	
	Mínimos Cuadrados			-	-	-	-	-	311.04	
	Minzapan	Gasto pico	Momentos	2p	20.645	11.377	12.418	12.797	11.757	-
				3p	-	11.822	11.245	-	-	-
Máxima			2p	20.645	11.678	14.726	60.649	14.85	-	
			3p	-	12.652	10.241	-	-	-	
Mínimos Cuadrados			-	-	-	-	-	8.637		
Volumen		Momentos	2p	3.154	3.293	2.3	4.996	2.723	-	
			3p	-	2.304	2.254	-	-	-	
		Máxima	2p	3.154	3.136	2.149	17.68	2.27	-	
			3p	-	3.042	2.485	-	-	-	
Mínimos Cuadrados			-	-	-	-	-	1.948		
Minatitlán		Gasto pico	Momentos	2p	226.373	238.182	199.92	437.956	247.279	-
				3p	-	195.314	194.669	-	-	-
	Máxima		2p	226.373	203.848	198.674	2351.115	194.350	-	
			3p	-	201.7	202.364	-	-	-	
	Mínimos Cuadrados			-	-	-	-	-	238.914	
	Volumen	Momentos	2p	583.571	641.738	591.541	1194.217	741.494	-	
			3p	-	573.127	14887.3	-	-	-	
		Máxima	2p	583.571	595.299	581.012	8890.669	7867.77	-	
			3p	-	593.623	610.101	-	-	-	
	Mínimos Cuadrados			-	-	-	-	-	717.511	

**Tabla 7.10.** Modelos univariados (Marginales) de mejor ajuste y sus parámetros

Estación	Marginal	FDP	Método	$\alpha 1$	$\beta 1$	$\alpha 2$		P	Mín. EEA
						$\delta$	$\beta 2$		
Las Perlas	F(Q)	Gamma 3p	Máx. Ver	806.365	3.278	494.232	-	-	208.72
	F(V)	LogNormal 3p	Máx. Ver	8.977	0.356	-73.709	-	-	272.634
Jesús Carranza II	F(Q)	LogNormal 3p	Máx. Ver	6.779	0.536	-23.409	-	-	75.279
	F(V)	Gumbel	Máx. Ver	227.273	432.273	-	-	-	31.583
Tierra Morada	F(Q)	Doble Gumbel	Mín. Cuad.	194.326	1344.006	1138.952	2412.853	0.88	58.622
	F(V)	Normal	Momentos	5087.428	1250.419	-	-	-	256.994
Minzapan	F(Q)	Doble Gumbel	Mín. Cuad.	37.244	103.795	61.512	250.498	0.87	8.637
	F(V)	Doble Gumbel	Mín. Cuad.	10.537	24.754	10.453	54.747	0.70	1.948
Minatitlán	F(Q)	Gumbel	Máx. Ver	1241.185	3553.278	-	-	-	194.35
	F(V)	LogNormal 3p	Momentos	10.872	0.064	-38977.621	-	-	573.127

Es común en el análisis de frecuencia univariado presentar gráficos en donde se reproduzca el ajuste de los datos observados mediante el comportamiento de una curva a través de su DFP; asimismo la gráfica se construye en función de su probabilidad asociado al periodo de retorno. No se tomará en cuenta este análisis, únicamente es para ejemplificar el ajuste, a continuación se presentan en las Figuras 7.15 a 7.19 las curvas.



**Figura 7.15.** Ajuste de datos observados en la estación Las Perlas. FDP Gamma 3p para Gastos (a) y FDP LogNormal3p para Volúmenes (b)

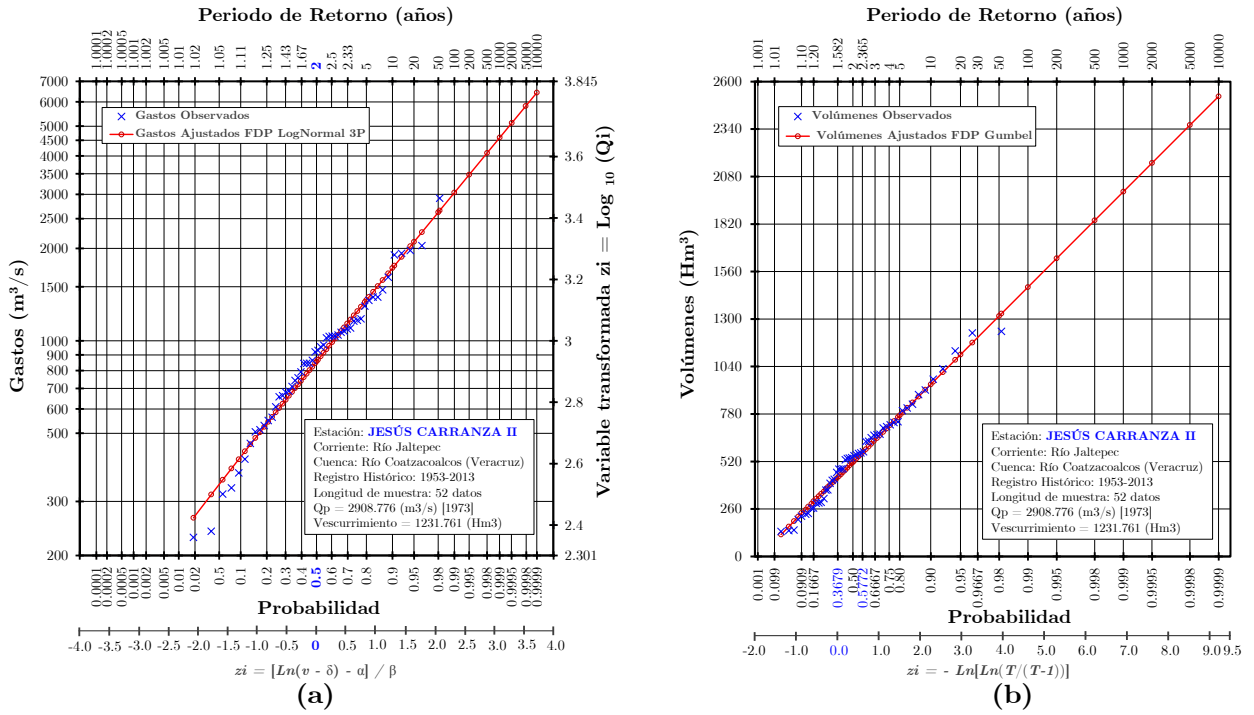


Figura 7.16. Ajuste de datos observados en la estación Jesús Carranza II. FDP LogNormal3p para Gastos (a) y FDP Gumbel para Volúmenes (b)

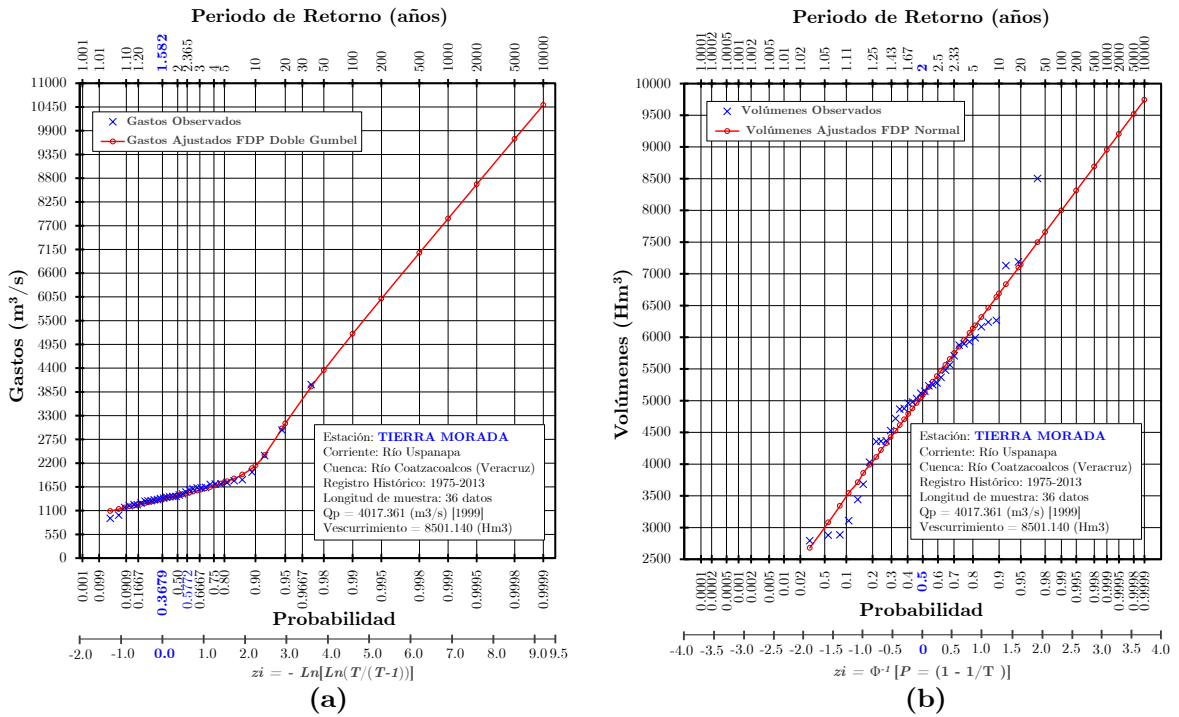


Figura 7.17. Ajuste de datos observados en la estación Tierra Morada. FDP Doble Gumbel para Gastos (a) y FDP Normal para Volúmenes (b)

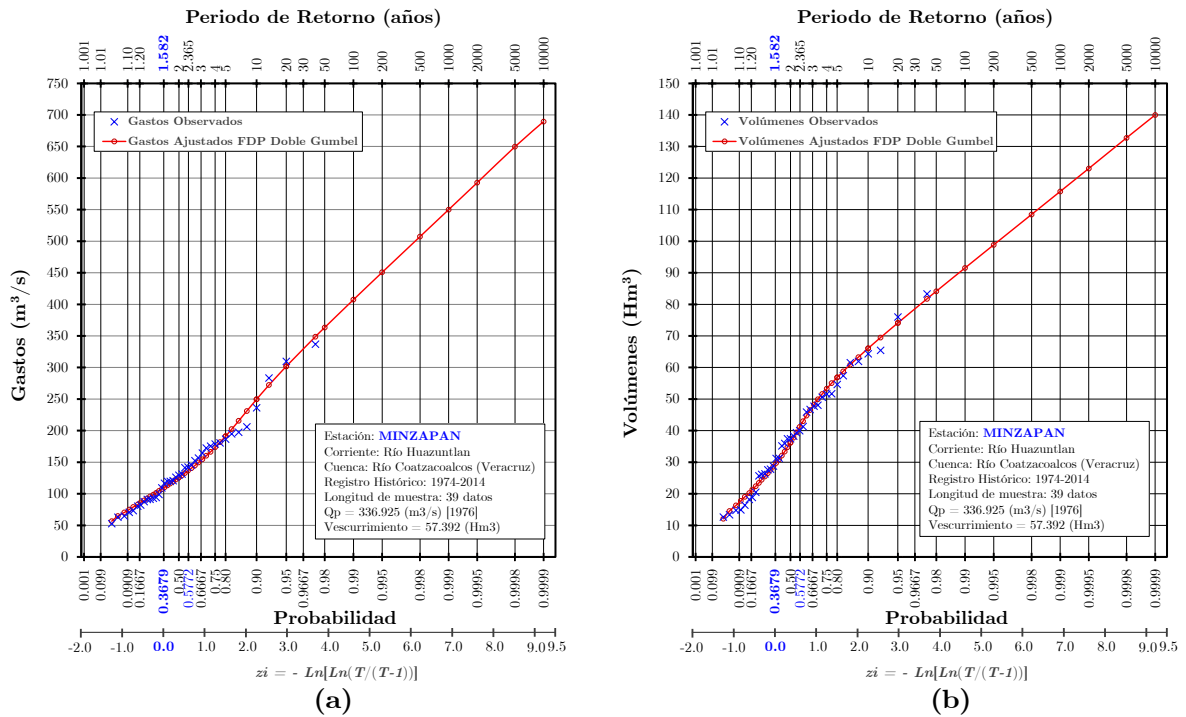


Figura 7.18. Ajuste de datos observados en la estación Minzapan. FDP Doble Gumbel para Gastos (a) y FDP Doble Gumbel para Volúmenes (b)

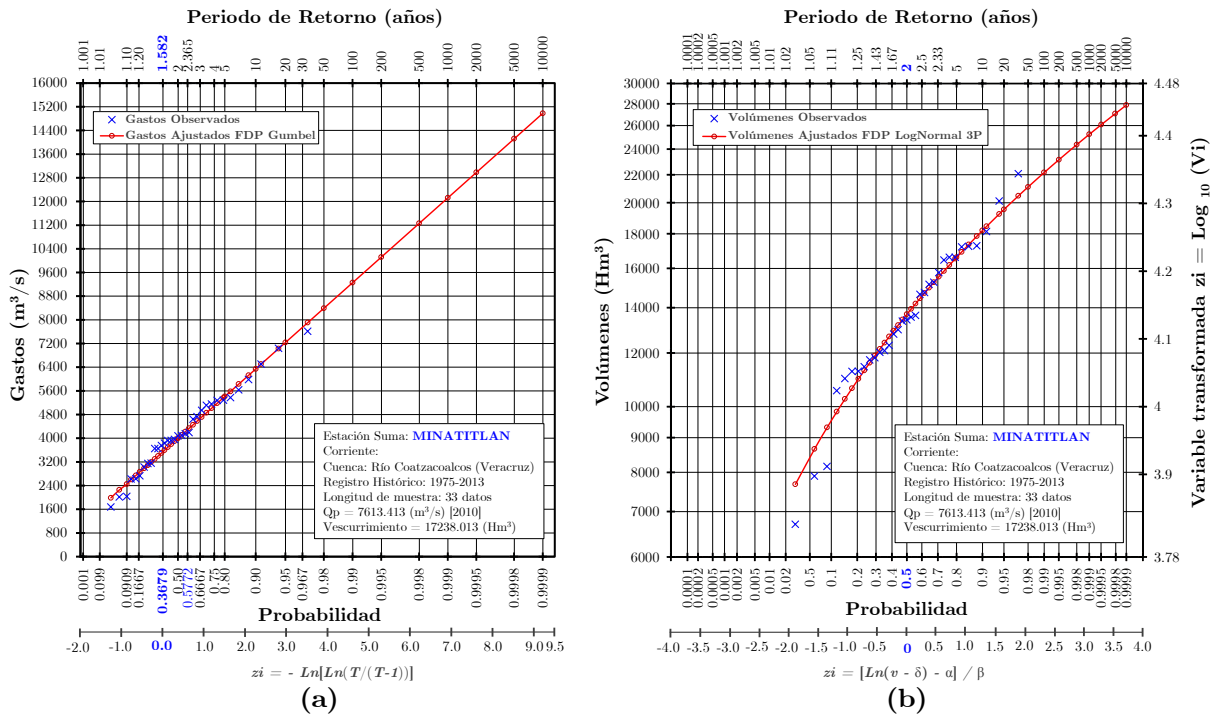


Figura 7.19. Ajuste de datos observados en la estación Minatitlán. FDP Gumbel para Gastos (a) y FDP LogNormal3p para Volúmenes (b)

▪ **Cuenca del Río Fuerte**

El ajuste de datos observados a través de distribuciones marginales que mejor representan a las variables Q y V de la estación Huites, se realizaron empleando el programa AX.exe. En la Tabla 7.11 se visualizan las FDP que se aplicaron a la muestra bivariadas (Q, V), a través del parámetro Error Estándar de Ajuste (EEA) que obtuvo cada FDP vinculada a los datos observados. Se aplicó la evaluación de bondad de ajuste a los modelos de distribución univariados, llamados Marginales, fijando como criterio para la selección del mejor modelo, aquel que posea el mínimo EEA correspondiente. En la Tabla 7.12, se expone un resumen con las marginales de mejor ajuste a las variables Q y V, asimismo se muestran los parámetros de escala, ubicación y de forma respectivamente. En este apartado no se tomará en cuenta el análisis de frecuencia univariado, en el Apéndice F se presentará un resumen del mismo en las Tablas F.1 a F.6. En este sentido, como propósito de validación, se presentarán en este trabajo las curvas de mejor ajuste a los datos observados, por medio de su FDP. En la Figura 7.20 se presentan las curvas de mejor ajuste a los datos observados de Huites. Las gráficas se efectuaron con ayuda del software estadístico R.

**Tabla 7.11.** Error Estándar de Ajuste de cada FDP

Est.	Var.	Método	Normal	LogNormal	Gamma	Exponencial	Gumbel	Doble Gumbel	
<b>Huites</b>	<i>Gasto pico</i>	Momentos	2p	1888.912	944.017	959.534	897.554	1314.936	-
			3p	-	1018.579	905.808	-	-	-
		Máxima	2p	1888.912	1225.297	1367.299	892.996	8594.782	-
			Verosimilitud	3p	-	1446.471	1286.359	-	-
		Mínimos Cuadrados	-	-	-	-	-	-	286.367
		<i>Volumen</i>	Momentos	2p	270.534	140.787	145.63	126.361	166.361
	3p			-	145.946	138.145	-	-	-
	Máxima		2p	270.534	102.492	167.487	203.413	215.944	-
			Verosimilitud	3p	-	115.982	152.511	-	-
	Mínimos Cuadrados		-	-	-	-	-	-	84.328

**Tabla 7.12.** Modelos univariados (Marginales) de mejor ajuste y sus parámetros

Estación	Mar-ginal	FDP	Método	$\alpha 1$	$\beta 1$	$\alpha 2$	$\beta 2$	P	Mín. EEA
Huites	F(Q)	Doble Gumbel	Mín. Cuad.	888.100	1581.371	3521.127	8569.593	0.84	286.367
	F(V)	Doble Gumbel	Mín. Cuad.	315.956	558.434	550.358	2207.009	0.89	84.328

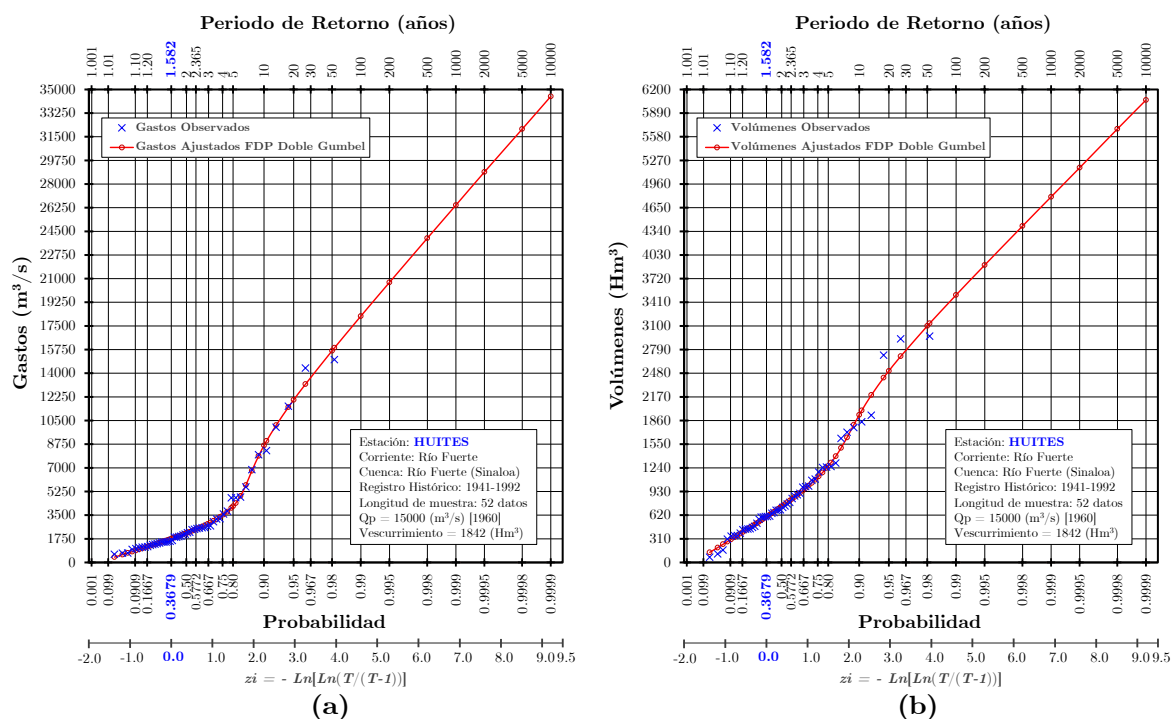


Figura 7.20. Ajuste de datos observados en la estación Huites.

FDP Doble Gumbel para Gastos (a) y FDP Doble Gumbel para Volúmenes (b)

## 7.2.4. Selección de las Cópulas Bivariadas

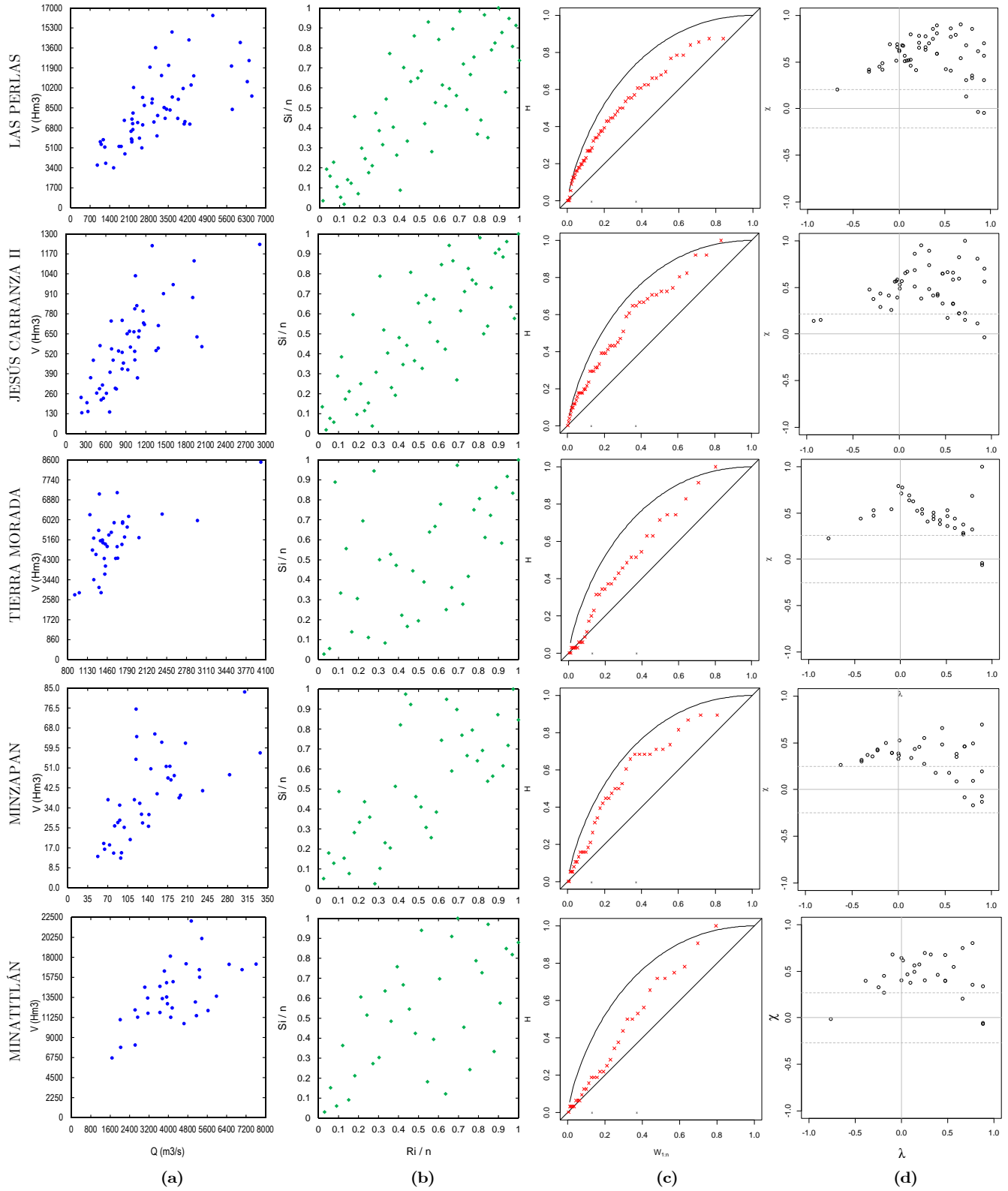
Una vez conocidas las marginales  $F(Q)$  y  $F(V)$ , el siguiente paso consiste en estudiar la dependencia entre las marginales asociadas al gasto pico y volumen de la avenida, por medio de una distribución de probabilidad bivariada, definida por una función cópula, o gráficamente. El proceso inicia con la obtención de las muestras modificadas o pseudo-muestras, es decir, transformadas en rangos  $(R_i/n, S_i/n)$ , obtenidas con los datos observados de las estaciones objeto de estudio (vea tablas 7.21 a 7.26).

### 7.2.4.1. Cuenca del Río Coatzacoalcos

- Dependencia mediante gráficos

Este análisis indicará a la vista una aceptable relación positiva de dependencia entre las variables  $Q$  y  $V$ , para cada estación objeto de análisis. Los diagramas de dispersión, de los datos observados en unidades originales, y de las pseudo-observaciones (transformadas por rangos  $(R_i/n, S_i/n)$ ), asimismo los diagramas de Chi-plot y de K-plot, para cada estación de análisis, se exhiben en la Figura 7.21.

## Aplicación del Análisis de frecuencia bivariado para estimar Avenidas de diseño



**Figura 7.21.** Diagramas de dispersión: (a) datos observados ( $Q_i$ ,  $V_i$ ); (b) pseudo-observaciones ( $R_i/n$ ,  $S_i/n$ ). Diagramas K-plot (c) y Chi-plot (d)



Los gráficos de dispersión presentados en la Figura 7.21, en términos generales muestran una relación positiva de dependencia entre variables, con excepción de las estaciones de Tierra Morada y Minatitlán en el caso de las pseudo-observaciones ( $R_i/n$ ,  $S_i/n$ ), los puntos se encuentran relativamente más dispersos que los otros casos. Estas inconsistencias que a simple vista se observan, pueden esclarecerse mediante el apoyo de los diagramas de K-plot y Chi-plot respectivamente. Los diagramas de K-plot para las cinco estaciones, los puntos se posicionaron por encima de la diagonal y tienden a ubicarse sobre la línea curva, lo que indica que sí existe dependencia positiva en los cinco casos. En los diagramas de Chi-plot, los valores se ubican por encima del límite superior derecho, lo que supone dependencia positiva en todos los casos. Los gráficos se elaboraron por medio del apoyo del software estadístico R.

- **Cálculo de medidas de dependencia  $\tau$  y  $\rho$**

Para corroborar los resultados proporcionados por la información gráfica, se tiene que calcular las medidas de dependencia basadas en rangos,  $\tau$  de Kendall y la  $\rho$  de Spearman (consulte el Apéndice A), los cuales *se obtienen con la muestra bivariada de datos observados (Tabla 7.7) o con las pseudo-muestras* (consulte las tablas 7.21 a 7.26). Las medidas  $\tau$  y  $\rho$  para las cinco muestras bivariadas se presentan en la Tabla 7.17, se observa que en cuatro casos, la dependencia oscila entre 60% y 70% para la medida  $\rho$ , mientras que en la estación Tierra Morada se obtuvo la medida de dependencia más baja para ambos casos,  $\tau = 0.40$  y  $\rho = 0.53$ .

- **Dependencia mediante Funciones Cópula**

El segundo análisis de dependencia entre marginales es de carácter numérico. Para el análisis se consideran cuatro conjuntos de familias de Funciones Cópula de un parámetro ( $\theta$ ): Elípticas, Arquimedianas, de Valor Extremo y otras familias (Cuadro 3.1), se recuerda que el sustento teórico se puede consultar en el Capítulo 3. Cabe mencionar que la cópula de Gumbel-Hougaard pertenece al grupo de cópulas arquimedianas y de valores extremos, depende su enfoque (generador o Pickands). El cuarto conjunto de cópulas, hace referencia a los modelos Plackett, Farlie-Gumbel-Morgenstern y dos modelos que fueron propuestos por el Dr. González-Barrios; Gumbel-Hougaard Decreasing-Decreasing y Gamabos Decreasing-Decreasing, estos últimos modelos

aplican funciones decrecientes a las muestras bivariadas (datos observados), en otras palabras, las cópulas resultantes son los modelos de Gumbel-Hougaard y Galambos de forma invertida.

El conjunto de modelos cópula considerados se redujo después de analizar los rangos admisibles de dependencia a través del valor de Kendall. Como resultado, se descartaron las cópulas de Ali-Mikhail-Haq ( $\tau \in [-0.1817, 1/3]$ ), Farlie-Gumbel-Morgenstern ( $\tau \in [-2/9, 2/9]$ ) y Tawn ( $\tau \in [0, 0.4184]$ ). Finalmente, los modelos cópula que se consideraron para el trabajo de esta tesis son: **Normal, Clayton, Frank, Joe/B5, Gumbel-Hougaard, Galambos, Hüsler-Reiss, Plackett, Gumbel-H DD y Galambos DD**. Cabe señalar que *para toda clase de simulación referente a cópulas es obligatorio trabajar con las pseudo-observaciones*, es decir, transformadas en rangos  $(R_i/n, S_i/n) \in [0,1]$ . Las simulaciones de los modelos cópula se efectuaron mediante el software estadístico R.

*¿Qué se obtiene al simular una cópula?*, primero: lograr fusionar dos distribuciones marginales y hacerlas dependientes por medio de su parámetro  $\theta$ ; segundo: caracterizar la estructura de dependencia de la muestra original (datos observados), al elegir el modelo más apropiado al fenómeno; tercero: extender la longitud de las muestras de datos observados al tamaño que uno desee.

- **Estimación del parámetro de la cópula ( $\theta$ )**

La estimación del parámetro de la cópula ( $\theta$ ) para los distintos modelos en estudio, se efectuaron mediante las relaciones que se enuncian en la Tabla 3.4 (relación  $\rho_s$  y  $\tau$  con  $\theta$ ), método de inversión. Los cálculos pertinentes de  $\theta$  se efectuaron mediante el software estadístico R, y se muestran en la Tabla 7.17.

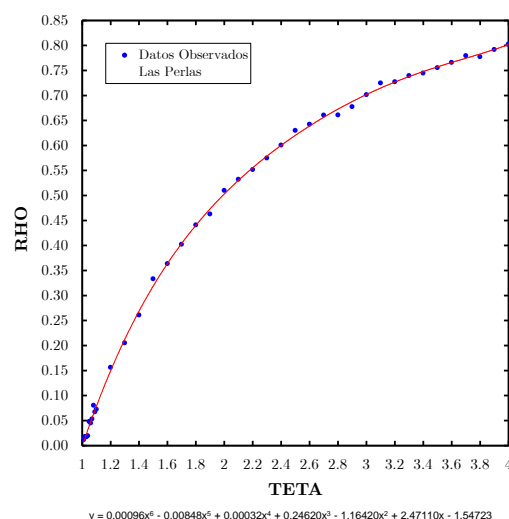
Para el caso de la cópula Joe/B5, el parámetro  $\theta$  con relación a  $\rho$  de Spearman tiene una forma complicada de obtener, la paquetería que ofrece el software estadístico R tampoco es factible obtenerlo de forma directa. Con base en esta dificultad, es posible estimarlo numéricamente a través de programación R: se propone un vector de parámetros  $\theta$  con intervalo relativamente pequeño; posteriormente se calculan los valores de  $\rho$ , esto se logra a partir de simulaciones y proponiendo un conjunto de 10,000 puntos de una muestra aleatoria que servirá para calcular cada

valor de  $\rho$ ; se bosqueja una gráfica de  $\theta$  vs  $\rho$ , y se realiza un ajuste por medio de una función; de la curva generada y conociendo el valor de  $\rho$  (previamente calculado de la muestra bivariada) es fácil obtener el parámetro utilizando la función de ajuste.

Para ejemplificar lo anterior, se obtendrá el valor del parámetro  $\theta$  de la cópula Joe/B5 con relación a  $\rho$  de Spearman. Si se utiliza la metodología del Apéndice A, la medida  $\rho = 0.7636764324604615$ . En la Tabla 7.13, se visualiza el vector de parámetros  $\theta$  propuesto y su correspondiente  $\rho$  calculada. En la Figura 7.22, se muestra la gráfica de  $\theta$  vs  $\rho$ , y su curva de ajuste representada por una función polinómica de grado seis. Finalmente se obtiene un valor de  $\theta = 3.5803104$ .

**Tabla 7.13.** Valores propuestos de  $\theta$  y cálculo de  $\rho$  con la Cópula Joe/B5

$\theta$	$\rho$	$\theta$	$\rho$	$\theta$	$\rho$
1.01	0.0113326	1.50	0.3332129	2.80	0.6611518
1.02	0.0180305	1.60	0.3635284	2.90	0.6775144
1.03	0.0171949	1.70	0.4023485	3.00	0.7014593
1.04	0.0194442	1.80	0.4411172	3.10	0.7249101
1.05	0.0480237	1.90	0.4630793	3.20	0.7272637
1.06	0.0449225	2.00	0.5100477	3.30	0.7395789
1.07	0.0529924	2.10	0.5322283	3.40	0.7447843
1.08	0.0805206	2.20	0.5519220	3.50	0.7556693
1.09	0.0672887	2.30	0.5749017	3.60	0.7660852
1.10	0.0727179	2.40	0.6005983	3.70	0.7795260
1.20	0.1564216	2.50	0.6301434	3.80	0.7775106
1.30	0.2052158	2.60	0.6422674	3.90	0.7919417
1.40	0.2610622	2.70	0.6608561	4.00	0.8027762



**Figura 7.22.** Curva  $\theta$  vs  $\rho$

▪ **Simulación de estructuras de dependencia mediante cópulas**

La dependencia entre variables Q y V, se demostró en primera instancia por los gráficos de dispersión y los gráficos de K-Plot y Chi-Plot. Ésta visualización a priori, se verificó numéricamente por conducto de los coeficientes Rho de Spearman, Tau de Kendall y el parámetro de la Cópula (depende de la muestra objeto de análisis). En este sentido, se ha comprobado la dependencia que existe entre las parejas de valores (Q, V), de tal forma que son susceptibles de tratarlas de forma conjunta, es decir, no generando gastos pico y volúmenes de forma independiente porque no lo son (en caso de extender la muestra para modelar periodos de retorno), sino mediante un tratamiento bivariado por medio de funciones cópula.

Los datos observados de cada estación, han servido para llevar a cabo el ajuste de funciones de distribución marginales que conforman la función cópula, que en lo sucesivo se denominarán como  $F(Q)=u_1$  para gastos pico y  $F(V)=u_2$  para volúmenes.

La simulación de un modelo cópula bivariado  $C(u_1, u_2)$  consiste en trabajar con las distribuciones marginales  $u_1$  y  $u_2$ , de forma extendida, Aranda (2014) recomienda de 10,000 a 100,000 conjuntos de probabilidades  $(u_1, u_2)$ . En este trabajo se generó un conjunto de 10,000 pares sintéticos  $(u_1, u_2)$  (asociados a determinada pseudo-muestra bivariada objeto de análisis), para cada cópula simulada.

En consecuencia, se simuló un total de diez modelos cópula  $C(u_1, u_2)$  para cada conjunto de distribuciones marginales extendidas  $(u_1, u_2)$  objeto de análisis (cinco estaciones), o sea, un total de 50 simulaciones se realizaron con el objetivo de seleccionar que modelo cópula caracteriza de manera apropiada a cada una de las cinco muestras bivariadas en su forma original (datos observados). En la Tabla 7.17, se muestran los modelos simulados.

#### ▪ Evaluación de la bondad de ajuste de los modelos Cópula

La siguiente etapa del análisis es la elección del modelo cópula que mejor se ajusta a los datos observados de cada una de las cinco muestras. En el subtema 3.7.5 y en la Figura 3.8, el lector puede consultar un resumen con las generalidades de la metodología propuesta.

La metodología que se propone consiste en obtener y comparar los volúmenes contenidos en las cajas  $R_{ij}$  del grid de tamaño uniforme  $m$  en  $I^2=[0,1]^2$ , de las siguientes copulas; **Cópula muestral** de orden  $m$   $C_n^{(m)}$  conformada por el conteo de puntos de la muestra de datos observados de orden  $n$  y la **Aproximación Checkerboard** (Cópula  $C^{(m)}$ ) de los  $r$  modelos cópula  $C_\theta$  que se pretenden estudiar; ( $C^{(m)}$  no es aleatoria). Se calcula la diferencia de volúmenes entre estas dos cópulas por medio del Lema 3.44, ec. 3.83. Bajo este contexto, el criterio de selección del mejor modelo cópula  $C_\theta$ , será quien posea la mínima distancia supremo  $D$ . En la Tabla 7.17, se muestra el concentrado de la evaluación de bondad de ajuste para las cinco estaciones.

**Ejemplo 1.** Aplicar la metodología propuesta referente a la prueba de bondad de ajuste. Se utilizará la pseudo-muestra bivariada en su forma original de la estación Las Perlas, con tamaño de 57 parejas ( $R_{iQ}$ ,  $S_{iV}$ ) (Tabla 7.14). Sus medidas de dependencia que resultaron son  $\rho = 0.7636764$  y  $\tau = 0.5614035$ . El modelo cópula que se empleará es Clayton (Figura 7.24), cuyo parámetro es  $\theta = 2.723024$ .

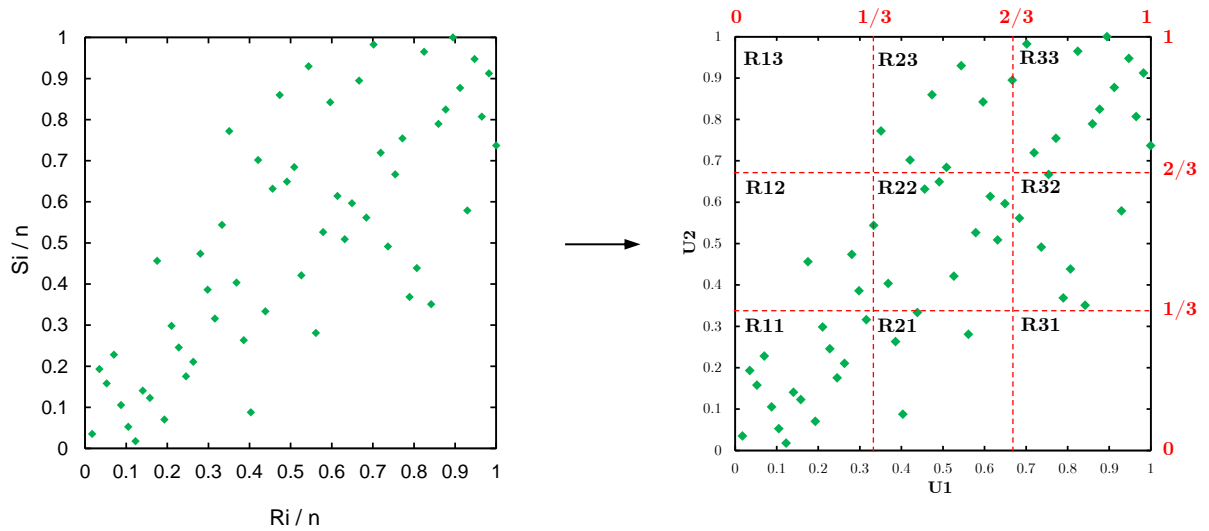
**Tabla 7.14.** Pseudo-muestra bivariada de la estación Las Perlas

$R_{iQ}$	$S_{iV}$	$R_{iQ}$	$S_{iV}$	$R_{iQ}$	$S_{iV}$	$R_{iQ}$	$S_{iV}$	$R_{iQ}$	$S_{iV}$	$R_{iQ}$	$S_{iV}$
0.19298	0.07018	1	0.73684	0.89474	1	0.49123	0.64912	0.26316	0.21053	0.73684	0.49123
0.22807	0.24561	0.33333	0.54386	0.52632	0.42105	0.84211	0.35088	0.61404	0.61404	0.38596	0.26316
0.56140	0.28070	0.24561	0.17544	0.29825	0.38596	0.12281	0.01754	0.35088	0.77193	0.98246	0.91228
0.28070	0.47368	0.68421	0.56140	0.96491	0.80702	0.78947	0.36842	0.57895	0.52632	0.05263	0.15789
0.14035	0.14035	0.40351	0.08772	0.50877	0.68421	0.45614	0.63158	0.63158	0.50877	0.08772	0.10526
0.36842	0.40351	0.94737	0.94737	0.82456	0.96491	0.42105	0.70175	0.77193	0.75439	0.07018	0.22807
0.15789	0.12281	0.91228	0.87719	0.54386	0.92982	0.01754	0.03509	0.66667	0.89474	0.03509	0.19298
0.64912	0.59649	0.87719	0.82456	0.59649	0.84211	0.71930	0.71930	0.70175	0.98246		
0.80702	0.43860	0.92982	0.57895	0.85965	0.78947	0.21053	0.29825	0.47368	0.85965		
0.31579	0.31579	0.75439	0.66667	0.17544	0.45614	0.43860	0.33333	0.10526	0.05263		

Proponer un grid de tamaño uniforme  $m$  en cada una de sus  $2$ -cajas. El parámetro  $m$ , debe ser un número entero en  $2 \leq m \leq n$ , donde  $n$  es el tamaño de la pseudo-muestra ( $m$  debe ser múltiplo de  $n$ ). El volumen total de cada fila o columna del grid es  $Vc = 1/m$ . Cada caja  $R_{ij}$  contara con un área de  $1/m^2$ , siendo  $m^2$  el número de cajas totales que conforman el grid. De tal forma, el volumen total del grid es  $Vc = 1$ . Si se requiere hacer el conteo de puntos (cardinalidad) de cada fila o columna del grid, la relación es  $n/m$ , asimismo debe ser un entero.

**Cópula muestral de orden  $C_n^{(m)}$**

Para una muestra de  $n = 57$  datos, se propone  $m = 3$ , de donde la relación  $n/m = 57/3 = 19$ , es un entero. En este contexto, se utilizará un grid de 9 cajas  $R_{ij}$ , los volúmenes totales de cada fila y columna son de  $1/3$  respectivamente y el volumen total de las 9 cajas que conforman el grid es de 1, Figuras 7.23a y 7.23b. La Cópula muestral de orden  $m$ ,  $C_n^{(m)}$ , a la cual se le ha implementado una malla a partir del origen y se ha calculado el conteo de puntos  $S_{ij}/n$  que caen en cada caja  $R_{ij}$ , (ec. 3.89), Figura 7.23c. Los volúmenes obtenidos en las cajas de la malla, representan una masa de probabilidad expandida en toda la superficie de la caja.  $C_n^{(m)}$ , dicha cópula será, a la cual se tendrá que aproximar el modelo en estudio  $C_\theta$ .



(a) Pseudo-muestra. Estación Las Perlas      (b) Generación del grid;  $n=57$  y  $m=3$

**Figura 7.23**

Dese cuenta que en el gráfico de la generación del grid, la distribución de los elementos  $R_{ij}$  se organizan a partir del origen de coordenadas,  $R_{11}$ , de esta posición verticalmente hacia arriba hasta llegar a  $R_{13}$ , y de esta forma para los demás arreglos respectivamente. Este procedimiento se programó en lenguaje R, y la forma de analizar estos arreglos es de forma matricial, de modo que, el arreglo es el común que se usa para cálculo de matrices.

Conteo de puntos ( $S_{ij}$ ) en las cajas  $R_{ij}$ , donde  $R_{ij} = S_{ij}$

S11	S12	S13	→	15	4	0	19 = n/m
S21	S22	S23		4	8	7	19
S31	S32	S33		0	7	12	19
				19	19	19	57

Volúmenes resultantes en las cajas  $R_{ij}$  de la Cópula muestral  $C_n^{(m)}$

S11/n	S12/n	S13/n	→	0.26315789	0.07017544	0	$1/3 = 1/m$
S21/n	S22/n	S23/n		0.07017544	0.14035088	0.1228070	$1/3$
S31/n	S32/n	S33/n		0	0.12280702	0.2105263	$1/3$
				$1/3$	$1/3$	$1/3$	1

Los volúmenes contenidos en las cajas  $R_{ij}$  de  $C$ , representan masas de probabilidad y se expanden de manera uniforme, véase la Figura 7.23c.

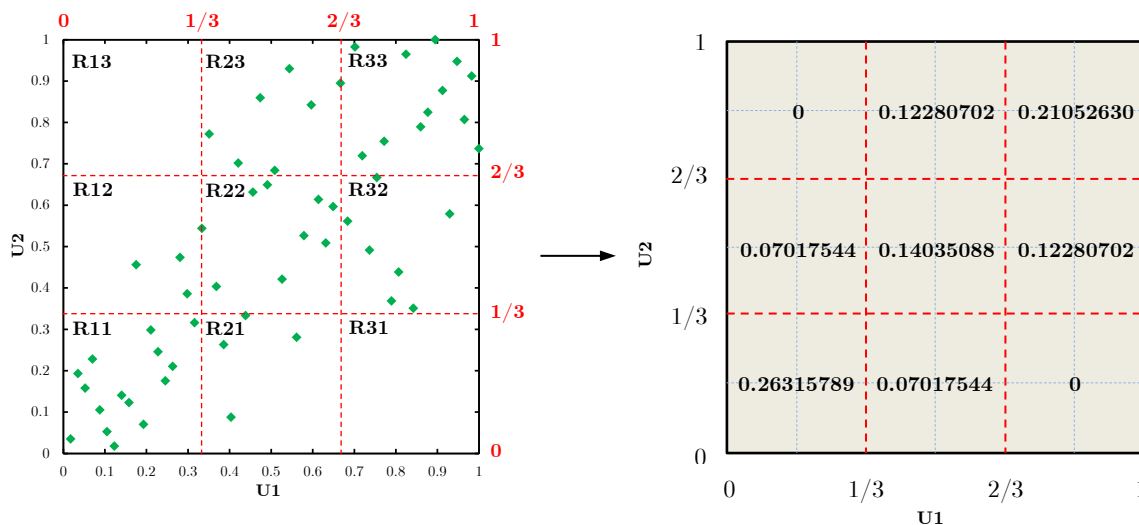


Figura 7.23c. Cópula muestral de orden  $m$ ,  $C_n^{(m)}$

**Modelos cópula  $C_\theta$**

Con ayuda del programa estadístico R, se simuló un total de diez modelos: Normal, Clayton, Frank, Joe/B5, Gumbel-Hougaard, Galambos, Hüsler-Reiss, Plackett, Gumbel-H DD y Galambos DD, consulte la Tabla 7.15. Con la misma lógica que la cópula muestral de orden  $m$ , se genera un grid con las mismas características que se plantearon y se toma en cuenta como parte de la información que proporciona cada simulación. El volumen de las cajas  $R_{ij}$ , se calculan a partir de los cuatro puntos que conforman cada caja (ec. 3.100 y figura 3.7). En estricto sentido, las probabilidades (volúmenes) que se han generado en cada caja  $R_{ij}$ , no están expandidas de manera uniforme, para hacer esto posible es necesario calcular una aproximación de cada modelo cópula.

**Aproximación Checkerboard  $C^{(m)}$  de los  $r$  modelos cópula  $C_\theta$**

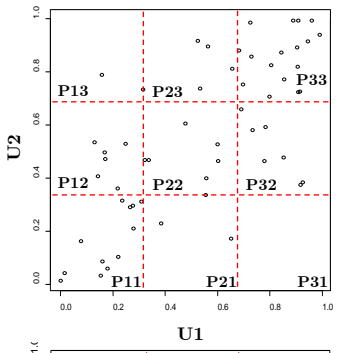
Con el propósito de uniformizar las masas de probabilidad en cada caja  $R_{ij}$  que conforman el grid de los  $r$  modelos cópula  $C_\theta$ , se estima la aproximación checkerboard  $C^{(m)}$  para cada modelo, y es producida por una interpolación bilineal. Lo que consigue la cópula  $C^{(m)}$ , es estimar la probabilidad de puntos que caen en las cajas  $R_{ij}$  del grid generado en la cópula del modelo, y hacer uniforme esa masa de probabilidad en toda la superficie de la caja  $R_{ij}$ , en la Tabla 7.15 se muestran las aproximaciones de los  $r$  modelos cópula .

Tabla 7.15. Metodología de la Prueba de bondad de ajuste para la Estación Las Perlas (González-Barrios y Hoyos).

		CÓPULA DEL MODELO $C_\theta$				CÓPULA $C^{(m)}$				CÓPULAS $C^{(m)} - C_n^{(m)}$							
		MODELO SIMULADO				APROXIMACIÓN CHECKERBOARD (VOLÚMENES)				CONTEO				DIFERENCIA DE VOLÚMENES			
		CONTEO				$C^{(m)}$				CONTEO				$C^{(m)} - C_n^{(m)}$			
CLAYTON		13	2	1	16	0.26085	0.06089	0.01160	1/3	15	3	1	19	0.00231	0.00929	0.01160	0.02320
		2	12	4	18	0.06089	0.16979	0.10266	1/3	3	10	6	19	0.00929	0.02944	0.02015	0.05888
		1	10	12	23	0.01160	0.10266	0.21908	1/3	1	6	12	19	0.01160	0.02015	0.00855	0.04030
		16	24	17	57	1/3	1/3	1/3	1	19	19	19	57	0.02320	0.05888	0.04030	<b>0.12238</b>
															D = <b>0.06119</b>		
FRANK		13	4	0	17	0.24214	0.08038	0.01082	1/3	14	5	1	19	0.02102	0.01020	0.01082	0.04204
		2	10	5	17	0.08038	0.17258	0.08038	1/3	5	10	5	19	0.01020	0.03223	0.04243	0.08486
		0	6	17	23	0.01082	0.08038	0.24214	1/3	1	5	14	19	0.01082	0.04243	0.03161	0.08486
		15	20	22	57	1/3	1/3	1/3	1	19	19	19	57	0.04204	0.08486	0.08486	<b>0.21177</b>
														D = <b>0.10588</b>			
JOE/B5		15	6	1	22	0.21553	0.10480	0.01300	1/3	12	6	1	19	0.04763	0.03463	0.01300	0.09526
		6	15	5	26	0.10480	0.16870	0.05983	1/3	6	10	3	19	0.03463	0.02835	0.06298	0.12595
		0	2	7	9	0.01300	0.05983	0.26050	1/3	1	3	15	19	0.01300	0.06298	0.04997	0.12595
		21	23	13	57	1/3	1/3	1/3	1	19	19	19	57	0.09526	0.12595	0.12595	<b>0.34717</b>
														D = <b>0.17359</b>			
GUMBEL-HOUGAARD		17	5	1	23	0.22889	0.09028	0.01416	1/3	13	5	1	19	0.03426	0.02010	0.01416	0.06853
		7	7	6	20	0.09028	0.17086	0.07220	1/3	5	10	4	19	0.02010	0.03051	0.05061	0.10122
		0	2	12	14	0.01416	0.07220	0.24698	1/3	1	4	14	19	0.01416	0.05061	0.03645	0.10122
		24	14	19	57	1/3	1/3	1/3	1	19	19	19	57	0.06853	0.10122	0.10122	<b>0.27097</b>
														D = <b>0.13548</b>			
GALAMBOS		14	2	1	17	0.22808	0.09155	0.01370	1/3	13	5	1	19	0.03507	0.02137	0.01370	0.07015
		7	9	7	23	0.09155	0.16838	0.07341	1/3	5	10	4	19	0.02137	0.02803	0.04940	0.09879
		1	3	13	17	0.01370	0.07341	0.24622	1/3	1	4	14	19	0.01370	0.04940	0.03569	0.09879
		22	14	21	57	1/3	1/3	1/3	1	19	19	19	57	0.07015	0.09879	0.09879	<b>0.26773</b>
														D = <b>0.13386</b>			
HÜSLER-REISS		13	5	0	18	0.22653	0.09387	0.01293	1/3	13	5	1	19	0.03663	0.02369	0.01293	0.07325
		3	5	6	14	0.09387	0.16382	0.07564	1/3	5	9	4	19	0.02369	0.02347	0.04717	0.09433
		2	4	19	25	0.01293	0.07564	0.24476	1/3	1	4	14	19	0.01293	0.04717	0.03423	0.09433
		18	14	25	57	1/3	1/3	1/3	1	19	19	19	57	0.07325	0.09433	0.09433	<b>0.26192</b>
														D = <b>0.13096</b>			
GAUSSIANA		13	6	0	19	0.23529	0.08531	0.01273	1/3	13	5	1	19	0.02786	0.01514	0.01273	0.05573
		7	10	2	19	0.08531	0.16271	0.08531	1/3	5	9	5	19	0.01514	0.02236	0.03749	0.07499
		0	7	12	19	0.01273	0.08531	0.23529	1/3	1	5	13	19	0.01273	0.03749	0.02477	0.07499
		20	23	14	57	1/3	1/3	1/3	1	19	19	19	57	0.05573	0.07499	0.07499	<b>0.20571</b>
														D = <b>0.10285</b>			
PLACKETT		7	3	1	11	0.24295	0.07367	0.01671	1/3	14	4	1	19	0.02021	0.00350	0.01671	0.04041
		3	12	7	22	0.07367	0.18599	0.07367	1/3	4	11	4	19	0.00350	0.04564	0.04914	0.09827
		2	3	19	24	0.01671	0.07367	0.24295	1/3	1	4	14	19	0.01671	0.04914	0.03243	0.09827
		12	18	27	57	1/3	1/3	1/3	1	19	19	19	57	0.04041	0.09827	0.09827	<b>0.23695</b>
														D = <b>0.11848</b>			



GUMBEL-HOUGAARD DD



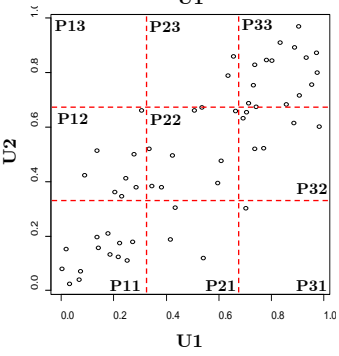
12	6	1	19
2	7	5	14
0	7	17	24
14	20	23	57

0.24698	0.07220	0.01416	$1/3$
0.07220	0.17086	0.09028	$1/3$
0.01416	0.09028	0.22889	$1/3$
$1/3$	$1/3$	$1/3$	<b>1</b>

14	4	1	19
4	10	5	19
1	5	13	19
19	19	19	<b>57</b>

0.01618	0.00202	0.01416	0.03236
0.00202	0.03051	0.03253	0.06505
0.01416	0.03253	0.01837	0.06505
0.03236	0.06505	0.06505	<b>0.16247</b>
			D = <b>0.08124</b>

GALAMBOS DD



13	7	1	21
3	9	2	14
1	7	14	22
17	23	17	57

0.24622	0.07341	0.01370	$1/3$
0.07341	0.16838	0.09155	$1/3$
0.01370	0.09155	0.22808	$1/3$
$1/3$	$1/3$	$1/3$	<b>1</b>

14	4	1	19
4	10	5	19
1	5	13	19
19	19	19	<b>57</b>

0.01694	0.00324	0.01370	0.03388
0.00324	0.02803	0.03126	0.06252
0.01370	0.03126	0.01756	0.06252
0.03388	0.06252	0.06252	<b>0.15893</b>
			D = <b>0.07946</b>

***Selección del mejor modelo Cópula. Distancia supremo D***

La siguiente etapa consiste en obtener la diferencia de volúmenes contenidos en las cajas  $R_{ij}$  de las cópulas  $C_n^{(m)}$  y  $C^{(m)}$ , mejor dicho  $V_{C_n^{(m)}}(R_{ij})$  y  $V_{C^{(m)}}(R_{ij})$ , asimismo para cada diferencia de volúmenes de cópulas, se obtiene la distancia supremo (distancia de variación total) entre las dos cópulas, en la Tabla 7.15 se observan las diferencias entre ambas cópulas y su distancia supremo D, para la Estación Las Perlas.

La *elección del mejor modelo cópula*  $C_\theta$ , será quien posea la *mínima distancia supremo D*, entre todos los modelos aplicados. En otras palabras, la cópula diferencial  $C^{(m)} - C_n^{(m)}$  con la mínima D, compara sus volúmenes de las cajas  $R_{ij}$  con los de la cópula muestral  $C_n^{(m)}$ , los cuales deben ser similares. Cabe señalar que la cópula diferencial  $C^{(m)} - C_n^{(m)}$  vincula a la mejor aproximación de un determinado modelo propuesto, por consiguiente, de esta forma es como se elige al mejor modelo bivariado que representará a la distribución de datos observados. Conforme a los resultados de la Tabla 7.15, se observa que la mejor Función Cópula es la de Clayton, al obtener la mínima  $D = 0.06118965$ .

En resumen, la Tabla 7.15 detalla la metodología para llevar a cabo la evaluación de la prueba de bondad de ajuste de los diez modelos cópula propuestos, por medio del método de González-Barrios y Hoyos, para la estación Las Perlas.

En términos generales, se simuló un total de diez modelos para cada una de las cinco estaciones, a los cuales se le obtuvo la cópula  $C^{(m)}$  por medio de la aproximación checkerboard, cincuenta aproximaciones en total. En la Tabla 7.17 se resume la evaluación de la prueba de bondad de ajuste de los diez modelos propuestos para cada una de las cinco estaciones de la cuenca del Río Coatzacoalcos y la correspondiente selección del mejor modelo que representa a cada estación de datos hidrométricos.

Asimismo, en la Tabla 7.17, se presentan los resultados de un método alternativo para evaluar la prueba de bondad de ajuste, empleando la metodología del mínimo error estándar de ajuste (EEA) aplicado a cópulas. Consiste en comparar la cópula empírica con la cópula del

modelo en estudio; se obtiene el error cuadrático entre ambas cópulas y posteriormente se calcula el EEA aplicando la expresión 2.38, tomando en cuenta que las cópulas que se han propuesto son de un parámetro; el modelo más apropiado es aquel que presente el mínimo error estándar de ajuste (EEA).

La cópula empírica se construye por medio de la frecuencia relativa acumulada, o sea la Ley empírica de Weibull en su versión ascendente  $k/(n+1)$ , donde  $k$  es la frecuencia absoluta acumulada y  $n$  el tamaño de la muestra. La frecuencia  $k_i$ , está conformada por los volúmenes menores o iguales que los valores de la serie  $V_i$  de la muestra de datos observados, sin tomar en cuenta repeticiones de  $V_i$  durante el proceso. Las etapas son las siguientes, apóyese en la Tabla 7.16:

- i. Ordenar de mayor a menor la muestra bivariada de datos observados, tomando como referencia la serie de gastos pico  $Q_i$ , nótese que al ordenar la muestra, la serie de volúmenes  $V_i$  no corresponden a una forma ascendente, véase Tabla 7.16.
- ii. Obtener la clasificación por rangos  $S_i$ , de la serie de volúmenes  $V_i$ .
- iii. Obtener la frecuencia absoluta acumulada  $k$ , contabilizando los volúmenes que sean menores o iguales a los valores  $V_i$  de la muestra de datos ordenada (de acuerdo a su clasificación por rangos  $S_i$ ), sin tomar en cuenta repeticiones de  $V_i$ .
- iv. Obtener la Cópula Empírica  $C_n$ .

**Tabla 7.16.** Obtención de la Cópula Empírica  $C_n$ . Estación Minatitlán

i	$Q_i$ (descendente) (m <sup>3</sup> /s)	$V_i$ (asociado) (Hm <sup>3</sup> )	$S_{vi}$	f (>)	k (<)	$C_n$ k/(n+1)	i	$Q_i$ (descendente) (m <sup>3</sup> /s)	$V_i$ (asociado) (Hm <sup>3</sup> )	$S_{vi}$	f (>)	k (<)	$C_n$ k/(n+1)
1	7613.413	17238.013	5	4	29	<b>0.853</b>	18	3960.814	12792.656	20	7	9	<b>0.265</b>
2	7037.829	16613.017	7	5	27	<b>0.794</b>	19	3926.905	13551.066	16	4	11	<b>0.324</b>
3	6503.965	17211.417	6	4	27	<b>0.794</b>	20	3924.047	15148.540	12	1	13	<b>0.382</b>
4	5982.003	13633.401	15	11	19	<b>0.559</b>	21	3832.553	16455.574	9	0	13	<b>0.382</b>
5	5632.659	12025.152	23	18	11	<b>0.324</b>	22	3750.334	13368.214	18	3	9	<b>0.265</b>
6	5375.723	20107.457	2	1	27	<b>0.794</b>	23	3650.578	14736.010	13	0	11	<b>0.324</b>
7	5285.819	15771.178	10	5	22	<b>0.647</b>	24	3650.161	11807.150	24	3	7	<b>0.206</b>
8	5275.151	16606.251	8	3	23	<b>0.676</b>	25	3157.922	11719.436	25	3	6	<b>0.176</b>
9	5156.578	11453.083	26	17	8	<b>0.235</b>	26	3145.874	13431.360	17	1	7	<b>0.206</b>
10	5110.423	12986.947	19	11	13	<b>0.382</b>	27	3024.572	14644.227	14	0	7	<b>0.206</b>
11	4941.280	22075.669	1	0	23	<b>0.676</b>	28	2731.955	11272.745	27	1	5	<b>0.147</b>
12	4729.402	17271.696	4	1	21	<b>0.618</b>	29	2628.518	8157.220	31	2	3	<b>0.088</b>
13	4642.538	10557.501	30	17	4	<b>0.118</b>	30	2627.356	12093.419	22	0	4	<b>0.118</b>
14	4193.450	15261.417	11	2	18	<b>0.529</b>	31	2032.395	7893.388	32	1	2	<b>0.059</b>
15	4153.879	12313.410	21	9	10	<b>0.294</b>	32	2014.331	10998.466	29	0	2	<b>0.059</b>
16	4097.641	11271.250	28	13	5	<b>0.147</b>	33	1675.660	6704.137	33	0	1	<b>0.029</b>
17	4082.258	18129.470	3	0	17	<b>0.500</b>							

**Tabla 7.17.** Prueba de bondad de ajuste para las estaciones de la Cuenca del Río Coatzacoalcos; D y EEA.

FAMILIA	CÓPULA	LAS PERLAS				JESÚS CARRANZA II				TIERRA MORADA				MINZAPAN				MINATITLÁN			
		$n$	$\rho$	$\tau$	$\theta$	$n$	$\rho$	$\tau$	$\theta$	$n$	$\rho$	$\tau$	$\theta$	$n$	$\rho$	$\tau$	$\theta$	$n$	$\rho$	$\tau$	$\theta$
Argumédianas (AR)	Clayton	57	0.7636764	0.5614035	52	0.7790489	0.5867270	36	0.5315315	0.4	39	0.7143725	0.5040779	33	0.5999332	0.4431818					
	Frank	7.01445	0.10588	0.35171	7.36624	0.12280	0.28651	3.74278	0.14401	0.40171	6.06058	0.10907	0.30955	4.46509	0.09901	0.34257					
	Joe/B5	3.58031	0.17359	0.36193	3.76350	0.21559	0.29484	2.10442	0.13569	0.41135	3.09890	0.17764	0.31811	2.39876	0.12423	0.35939					
	Gumbel-H	2.35548	0.13548	0.35637	2.44291	0.15817	0.29009	1.60368	0.13391	0.40695	2.12305	0.14063	0.31187	1.75699	0.10424	0.35494					
Valores Extremos (VE)	Galambos	1.63599	0.13386	0.35646	1.72424	0.15770	0.29009	0.87509	0.13814	0.40696	1.40163	0.13889	0.31181	1.03121	0.10569	0.35500					
	Husler-Reiss	2.19311	0.13096	0.35664	2.29289	0.15667	0.29010	1.31812	0.14146	0.40698	1.92692	0.13631	0.31174	1.50079	0.10711	0.35506					
Elíptica	Normal	0.77858	0.10285	0.35226	0.79338	0.11916	0.28642	0.54946	0.15145	0.40263	0.73077	0.10733	0.30741	0.61797	0.10756	0.35064					
	Plackett	18.73537	0.11848	0.35159	17.13950	0.12752	0.28664	5.11566	0.14061	0.40255	13.82111	0.12903	0.30928	5.11566	0.11668	0.35110					
Otras	Gumbel-H DD	2.35548	0.08124	0.35009	2.44291	0.09237	0.28357	1.60368	0.15961	0.40795	2.12305	0.08476	0.30591	1.75699	0.10424	0.34549					
	Galambos DD	1.63599	0.07946	0.34501	1.72424	0.09134	0.28099	0.87509	0.16235	0.41250	1.40163	0.08292	0.30123	1.03121	0.10569	0.35099					

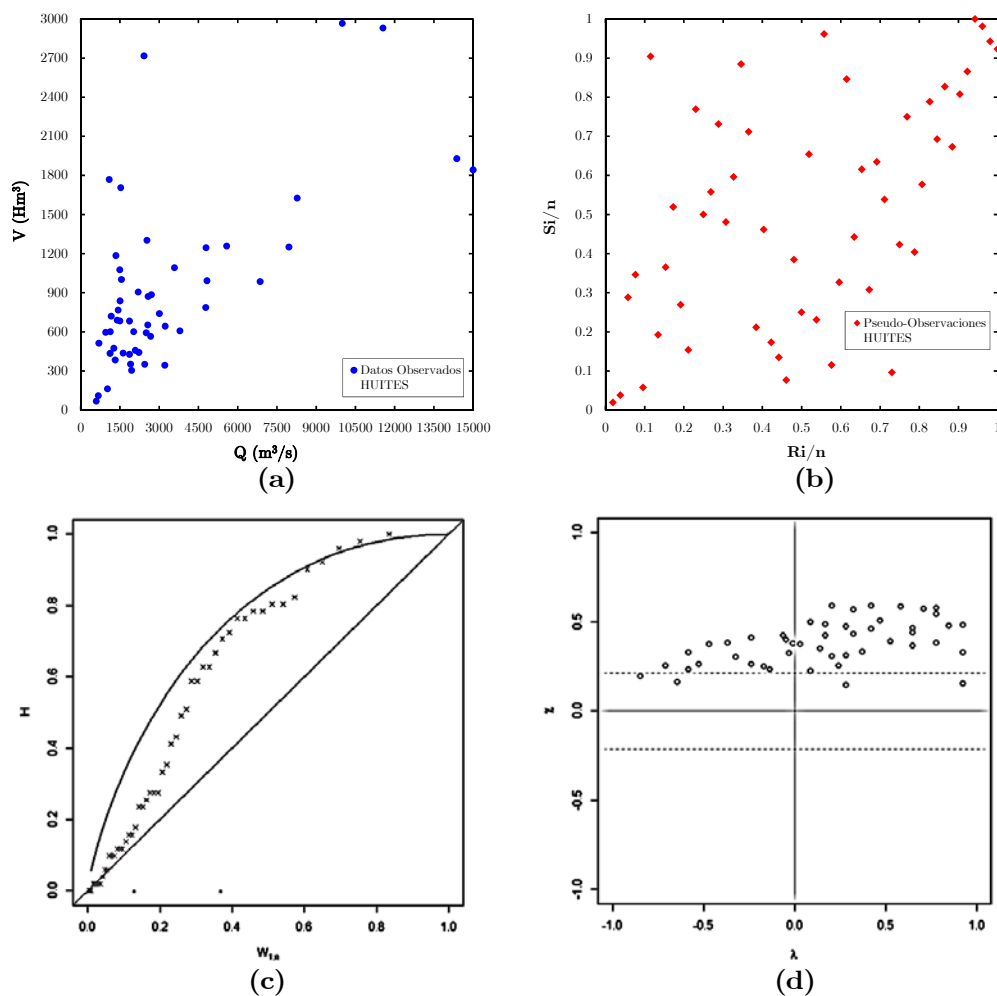
**Tabla 7.18.** Modelos seleccionados que mejor representan a las estaciones de la Cuenca del Río Coatzacoalcos

Estación	González-Barrios				Error Estándar de Ajuste			
	Cópula	$\theta$	D	Cópula	$\theta$	EEA	$\theta$	EEA
Las Perlas	Clayton	2.72302	0.06119	Clayton	2.72302	0.34253		
Jesús Carranza II	Clayton	2.90419	0.07009	Clayton	2.90419	0.27933		
Tierra Morada	Gumbel-H	1.60368	0.13391	Clayton	1.19649	0.39564		
Minzapan	Clayton	2.24595	0.06714	Clayton	2.24595	0.29950		
Minatitlán	Frank	4.46509	0.09901	Frank	4.46509	0.34257		

En la Tabla 7.18 se muestran los modelos cópula que se han seleccionado para representar de manera apropiada a las cinco estaciones en estudio, correspondientes a la cuenca del Río Coatzacoalcos, dicha selección está sustentada bajo un análisis como lo es la prueba de bondad de ajuste para cópulas. En los análisis posteriores, se empleará los modelos cópula seleccionados por medio del método de González-Barrios y Hoyos (2019).

### 7.2.4.2. Cuenca del Río Fuerte

A continuación se presentan los gráficos de dispersión con una aceptable dependencia entre  $Q_p$  y  $V$ , y los diagramas de Kendall-plot y Chi-plot, para la estación Huites, en el primer caso los puntos tienden a posicionarse en la línea curva, y en el segundo la nube de puntos se ubica por encima de las bandas y cargada en la parte superior derecha.



**Figura 7.24.** Diagramas de dispersión: (a) datos observados ( $Q_i, V_i$ ); (b) pseudo-observaciones ( $R_i/n, S_i/n$ ). Diagramas K-plot (c) y Chi-plot (d).

El conjunto de funciones cópula que se utilizará para ajustar los datos bivariados de la estación Huites, será el mismo que se utilizó para la cuenca del Río Coatzacoalcos. En la Tabla 7.19, se presenta un resumen de los modelos simulados, a través de sus medidas de dependencia Rho de Sperman y Tau de Kendall, asimismo sus respectivos parámetros de asociación. En este mismo contenido, se compara la evaluación de la prueba de bondad de ajuste con los métodos de González-Barrios (Distancia Supremo D) y del Error Estándar de Ajuste (EEA) para la estación Huites. En la Tabla 7.20, se muestra el modelo que mejor se ajusta a los datos observados de Huites, función cópula de Joe/B5. Las medidas de dependencia  $\rho$  (53%) y  $\tau$ , demuestran que existe una aceptable correlación entre variables  $Q_p$  y  $V$ .

**Tabla 7.19.** Pruebas de bondad de ajuste para la estación Huites

FAMILIA	CÓPULA	HUITES		
		n	$\rho$	$\tau$
		52	0.5260821	0.3831071
		$\theta$	D	EEA
Arquimedianas (AR)	Clayton	1.17479	0.20692	0.03663
	Frank	3.69008	0.15398	0.02930
	Joe/B5	2.08369	0.11000	0.02382
VE / AR	Gumbel-H	1.59303	0.13506	0.02500
Valores Extremos (VE)	Galambos	0.86403	0.13335	0.02514
	Hüsler-Reiss	1.30508	0.13230	0.02524
Elíptica	Normal	0.54397	0.15473	0.02838
	Plackett	5.11566	0.16079	0.02949
Otras	Gumbel-H DD	1.59303	0.18250	0.03238
	Galambos DD	0.86403	0.18050	0.02578

**Tabla 7.20.** Modelo seleccionado que mejor representa a la estación Huites

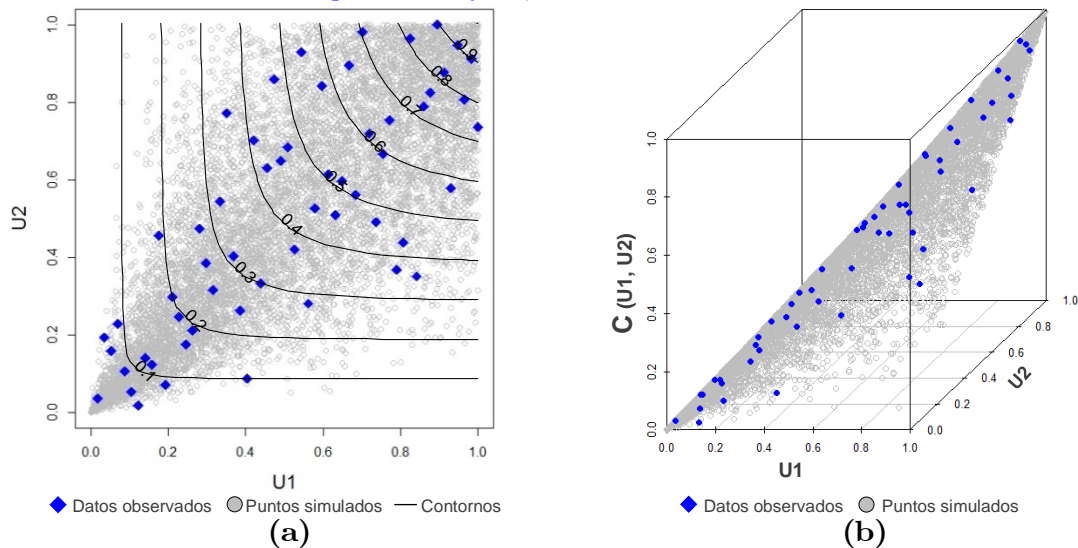
Estación	González-Barrios			Error Estándar de Ajuste		
	Cópula	$\theta$	$D_{\min}$	Cópula	$\theta$	EEA <sub>min</sub>
HUITES	Joe/B5	2.08369	0.11000	Joe/B5	2.08369	0.02382

### 7.2.5. Extensión de muestras sintéticas ( $F(Q)=u_1, F(V)=u_2$ )

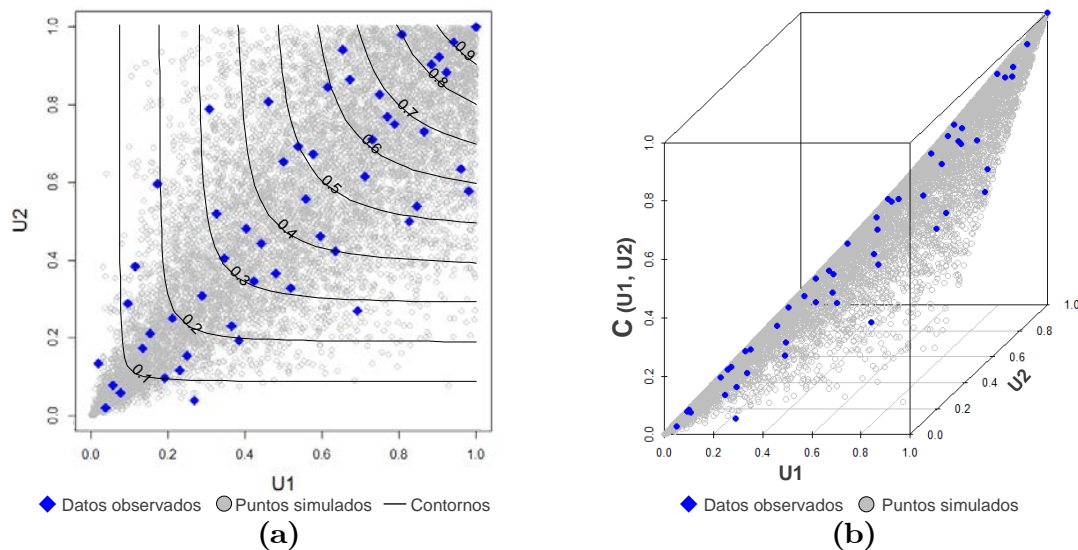
Una muestra sintética de longitud larga de pares  $(u_1, u_2) \in (0,1)$ , se genera a partir de la cópula ajustada a los datos observados de cada una de las estaciones por analizar, se considerarán generaciones de 10,000 pares de puntos (probabilidades).

▪ Cuenca del Río Coatzacoalcos

Cópula de Clayton; Estación Las Perlas



Cópula de Clayton; Estación Jesús Carranza II

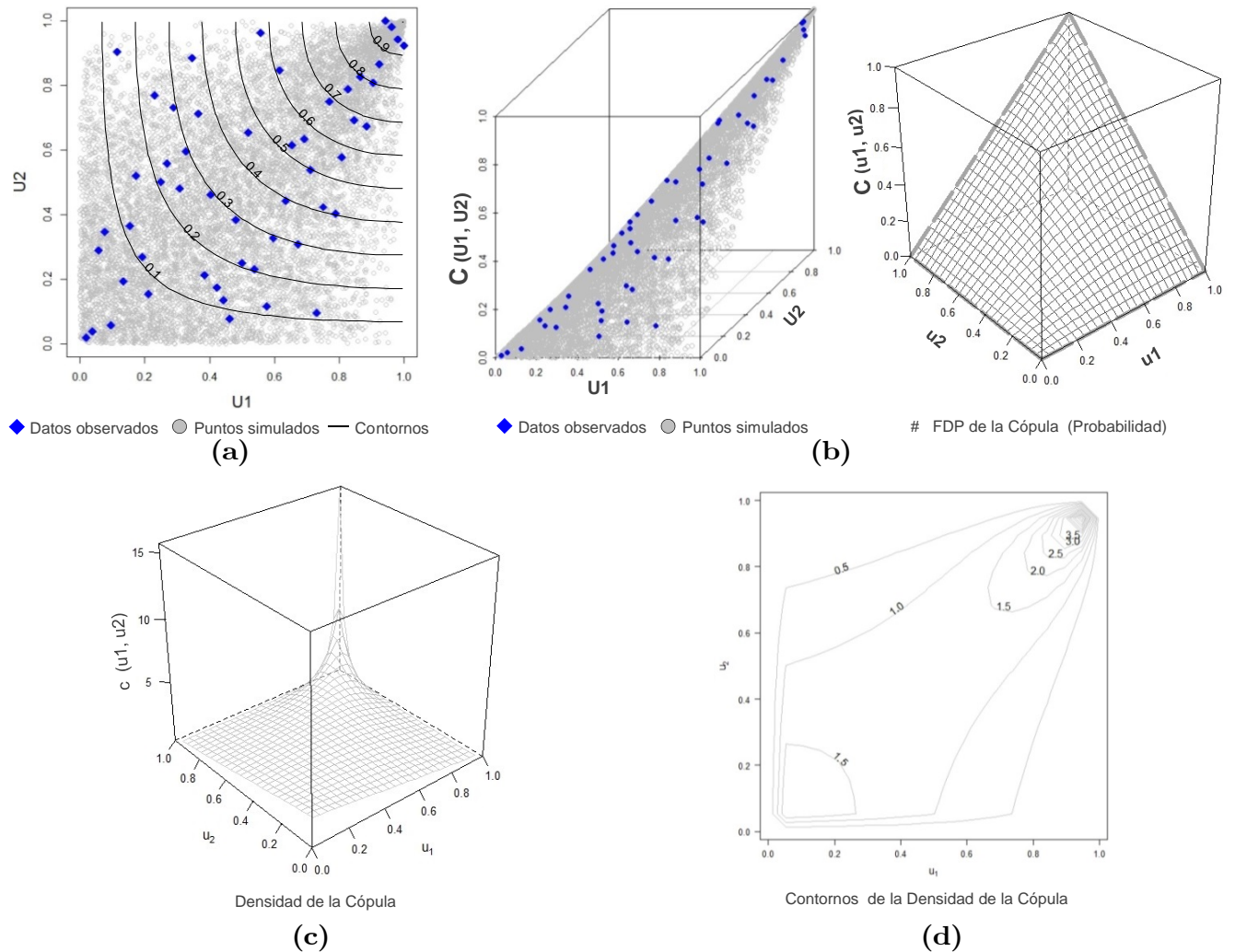


**Figura 7.25.** Comparación entre los datos observados y una muestra generada a partir de un modelo cópula, asimismo se trazan los contornos de la cópula (a); representación de la cópula en 3D (b) *Continúa...*





▪ Cuenca del Río Fuerte



**Figura 7.26.** Comparación entre los datos observados de la estación Huites y una muestra generada a partir del modelo Cópula Joe/B5, asimismo se trazan sus contornos (a); FDP de la cópula Joe/B5 (b); Densidad de la Cópula (c) y sus contornos (d).

### 7.2.6. Estimación de periodos de retorno conjuntos asociados a los datos observados

Se consideran tres maneras de cuantificar el período de retorno conjunto a través de la cópula apropiada a cada estación: " $T_{OR}$ " (primario), " $T_{AND}$ " (primario), y " $T_{KENDALL}$ " (secundario). Consulte ecuaciones 4.20 a 4.23, 4.26 a 4.27; tablas 3.2 y 4.2; resultados de tablas 7.21 a 7.25, utilizando gastos medios diarios máximos anuales y en tabla 7.26 gastos instantáneos anuales.

**Tabla 7.21.** Análisis de Frecuencia Bivariado de Avenidas de la Estación Las Perlas; Cuenca del Río Coatzacoalcos

ORDEN	DATOS OBSERVADOS			RANKING		DATOS TRANSFORMADOS		CÓPULA	MARGINALES UNIFORMES		CÓPULA	PERIODOS DE RETORNO BIVARIADOS OBTENIDOS CON CÓPULA CLAYTON					
	ESTACIÓN: LAS PERLAS			CLASIFICACIÓN		$Y_{ij} \in [0, 1]$		EMPÍRICA	GAMMA 3P	LOGNORMAL 3P	CLAYTON	"OR"	"AND"	GENERADOR	DERIVADA	FUN. KENDALL	"KENDALL"
	i	Año	Q m³/s	V Hm³	R <sub>Qi</sub>	R <sub>Vi</sub>	R <sub>Qi</sub> /N	R <sub>Vi</sub> /N	C(R <sub>Qi</sub> /N, R <sub>Vi</sub> /N) (Probabilidad)	F(Q)=u <sub>1</sub> (Probabilidad)	F(V)=u <sub>2</sub> (Probabilidad)	C(u <sub>1</sub> ,u <sub>2</sub> ) (Probabilidad)	T <sup>V</sup> (u <sub>1</sub> ,u <sub>2</sub> ) (años)	T <sup>A</sup> (u <sub>1</sub> ,u <sub>2</sub> ) (años)	φ (t)	φ' (t)	Kc (t) (Probabilidad)
1	1953	1947.374	4581.725	11	4	0.1929825	0.0701754	0.0686206	0.2141509	0.0674183	0.0663997	1.0711	1.2742	1610.6473	-66092.9879	0.0908	1.0998
2	1954	2181.895	5857.019	13	14	0.2280702	0.2456140	0.1838130	0.2861937	0.2078088	0.1834398	1.2246	1.4505	100.2803	-1503.4301	0.2501	1.3336
3	1955	3098.750	6115.796	32	16	0.5614035	0.2807018	0.2692581	0.5631111	0.2438800	0.2370045	1.3106	2.3255	49.4142	-579.2252	0.3223	1.4756
4	1956	2200.750	7542.663	16	27	0.2807018	0.4736842	0.2617947	0.2921124	0.4561374	0.2681790	1.3665	1.9233	35.0092	-365.6281	0.3639	1.5722
5	1957	1738.249	5220.917	8	8	0.1403509	0.1403509	0.1089044	0.1543536	0.1285474	0.1080725	1.1212	1.2119	426.7763	-10778.3690	0.1477	1.1733
6	1958	2406.000	7247.795	21	23	0.3684211	0.4035088	0.3022366	0.3568951	0.4124410	0.2993606	1.4273	1.8867	25.6891	-242.7680	0.4052	1.6812
7	1959	1815.312	5219.094	9	7	0.1578947	0.1228070	0.1057882	0.1757201	0.1283441	0.1127977	1.1271	1.2365	379.7232	-9190.9540	0.1541	1.1822
8	1960	3439.749	8390.354	37	34	0.6491228	0.5964912	0.5078831	0.6492539	0.5740784	0.4953072	1.9814	3.6768	5.7743	-37.2427	0.6504	2.8600
9	1961	4101.375	7332.841	46	25	0.8070175	0.4385965	0.4257876	0.7805547	0.4251350	0.4113641	1.6988	4.8621	10.2324	-74.3526	0.5490	2.2172
10	1962	2235.041	6651.430	18	18	0.3157895	0.3157895	0.2467980	0.3029025	0.3226366	0.2439246	1.3226	1.6171	45.6142	-520.3720	0.3316	1.4961
11	1963	6497.666	9504.804	57	42	1	0.7368421	0.7368421	0.9706565	0.7036104	0.6954104	3.2831	47.2959	1.6889	-10.5291	0.8558	6.9356
12	1964	2239.448	8055.928	19	31	0.3333333	0.5438596	0.3104710	0.3042913	0.5292302	0.2861332	1.4008	2.2094	29.1841	-287.2506	0.3877	1.6333
13	1965	2197.403	5597.068	14	10	0.2456140	0.1754386	0.1554023	0.2910609	0.1735668	0.1605636	1.1913	1.4369	144.5599	-2468.5743	0.2191	1.2806
14	1966	3566.507	8313.473	39	32	0.6842105	0.5614035	0.4993226	0.6780794	0.5639895	0.4990506	1.9962	3.8913	5.6368	-36.2132	0.6547	2.8961
15	1967	2555.979	5094.033	23	5	0.4035088	0.0877193	0.0872614	0.4039379	0.1147715	0.1135425	1.1281	1.6811	372.9606	-8968.4781	0.1551	1.1836
16	1968	6082.677	14061.219	54	54	0.9473684	0.9473684	0.9037646	0.9573407	0.9484210	0.9130263	11.4977	137.6538	0.2812	-3.8210	0.9866	74.6883
17	1969	5769.958	12055.792	52	50	0.9122807	0.8771930	0.8206874	0.9437781	0.8847748	0.8480743	6.5822	51.2258	0.5663	-5.0291	0.9607	25.4316
18	1970	4417.460	11217.620	50	47	0.8771930	0.8245614	0.7588943	0.8274947	0.8407916	0.7390595	3.8323	14.1296	1.2782	-8.3939	0.8913	9.2027
19	1971	5798.591	8372.399	53	33	0.9298246	0.5789474	0.5687720	0.9451691	0.5717345	0.5643140	2.2952	21.0924	3.7491	-22.9163	0.7279	3.6753
20	1972	3862.104	9225.439	43	38	0.7543860	0.6666667	0.5918754	0.7384706	0.6741948	0.5899098	2.4385	5.6419	3.2088	-19.4277	0.7551	4.0829
21	1973	5099.650	16356.012	51	57	0.8947368	1	0.8947368	0.9003815	0.9799660	0.8867010	8.8262	157.3936	0.3874	-4.2607	0.9776	44.6974
22	1975	2986.931	7298.026	30	24	0.5263158	0.4210526	0.3673661	0.5322844	0.4199461	0.3682333	1.5829	2.4038	14.1865	-112.3015	0.4946	1.9785
23	1976	2226.037	7156.266	17	22	0.2982456	0.3859649	0.2596480	0.3000664	0.3987174	0.2635752	1.3579	1.7706	36.7477	-389.9762	0.3578	1.5572
24	1978	6336.781	10730.080	55	46	0.9649123	0.8070175	0.7907619	0.9660416	0.8088105	0.7929657	4.8301	55.2071	0.8808	-6.4585	0.9293	14.1519
25	1981	2929.631	9253.425	29	39	0.5087719	0.6842105	0.4636802	0.5160377	0.6772352	0.4670648	1.8764	3.6524	6.9487	-46.3413	0.6170	2.6110
26	1982	4245.799	14272.462	47	55	0.8245614	0.9649123	0.8069900	0.8031719	0.9526858	0.7814072	4.5747	39.1397	0.9575	-6.8214	0.9218	12.7832
27	1983	3048.729	13621.620	31	53	0.5438596	0.9298246	0.5357562	0.5494688	0.9383213	0.5421734	2.1842	18.3880	4.2960	-26.5988	0.7037	3.3748
28	1984	3260.750	11251.216	34	48	0.5964912	0.8421053	0.5673502	0.6055661	0.8428106	0.5750240	2.3531	7.8959	3.5121	-21.3671	0.7394	3.8372
29	1985	4319.889	10437.004	49	45	0.8596491	0.7894737	0.7236152	0.8140013	0.7870237	0.6971208	3.3016	10.4063	1.6710	-10.4332	0.8573	7.0069
30	1987	1928.050	7430.401	10	26	0.1754386	0.4561404	0.1714048	0.2084049	0.4396122	0.2001042	1.2502	1.8113	78.9275	-1087.6553	0.2727	1.3749
31	1988	2916.014	8913.535	28	37	0.4912281	0.6491228	0.4431406	0.5121341	0.6389034	0.4542903	1.8325	3.2976	7.5721	-51.3815	0.6017	2.5104
32	1989	4274.984	7113.683	48	20	0.8421053	0.3508772	0.3465396	0.8074990	0.3923151	0.3837682	1.6228	5.4361	12.5703	-96.2881	0.5143	2.0590
33	1990	1542.900	3398.817	7	1	0.1228070	0.0175439	0.0175119	0.1047749	0.0101659	0.0101594	1.0103	1.1170	267511.6819	-71701319.5457	0.0139	1.0141
34	1991	4073.784	7127.026	45	21	0.7894737	0.3684211	0.3606741	0.7760006	0.3943224	0.3834731	1.6220	4.6915	12.5988	-96.5643	0.5139	2.0574
35	1992	2644.343	8705.360	26	36	0.4561404	0.6315789	0.4149119	0.4312493	0.6139335	0.3937635	1.6495	2.8688	11.6527	-87.4983	0.5269	2.1139
36	1993	2582.860	9383.771	24	40	0.4210526	0.7017544	0.3994805	0.4122859	0.6911215	0.3909887	1.6420	3.4773	11.8987	-89.8326	0.5234	2.0984
37	1994	958.511	3627.363	1	2	0.0175439	0.0350877	0.0166585	0.0123210	0.0161444	0.0106714	1.0108	1.0181	233989.5979	-59707340.3179	0.0146	1.0148
38	1995	3649.163	9419.645	41	41	0.7192982	0.7192982	0.6063445	0.6959250	0.6948639	0.5814204	2.3890	5.2457	3.3782	-20.5050	0.7462	3.9397
39	1996	2168.700	6504.538	12	17	0.2105263	0.2982456	0.1873921	0.2820589	0.3006684	0.2269063	1.2935	1.5524	55.7604	-681.1612	0.3088	1.4467
40	1997	2586.347	7046.819	25	19	0.4385965	0.3333333	0.2927974	0.4133664	0.3822454	0.3122112	1.4539	1.9357	22.8029	-207.6023	0.4221	1.7303
41	1998	2198.073	5771.812	15	12	0.2631579	0.2105263	0.1800743	0.2912712	0.1963402	0.1768293	1.2148	1.4509	110.9253	-1723.5564	0.2412	1.3178
42	1999	3353.591	8526.117	35	35	0.6140351	0.6140351	0.5015487	0.6286575	0.5915530	0.4968715	1.9876	3.6145	5.7164	-36.8080	0.6522	2.8750
43	2000	2258.943	10230.013	20	44	0.3508772	0.7719298	0.3435559	0.3104392	0.7704019	0.3056984	1.4403	4.4473	24.2092	-224.5523	0.4135	1.7051
44	2001	3122.859	7845.769	33	30	0.5789474	0.5263158	0.4431591	0.5695986	0.4998559	0.4256103	1.7410	2.8078	9.2379	-65.5010	0.5666	2.3076
45	2002	3389.663	7605.741	36	29	0.6315789	0.5087719	0.4500829	0.6373788	0.4653508	0.4225999	1.7319	3.1263	9.4377	-67.2551	0.5629	2.2879
46	2003	4038.793	10144.786	44	43	0.7719298	0.7543860	0.6540206	0.7701148	0.7632513	0.6581843	2.9256	8.0117	2.1235	-12.9225	0.8225	5.6342
47	2004	3516.746	12115.124	38	51	0.6666667	0.8947368	0.6400640	0.6669738	0.8874217	0.6382034	2.7640	11.9320	2.3970	-14.4942	0.8036	5.0912
48	2005	3638.544	14949.500	40	56	0.7017544	0.9824561	0.6969651	0.6936744	0.9641663	0.6840958	3.1655	38.0878	1.8118	-11.1922	0.8460	6.4924
49	2006	2845.812	11959.499	27	49	0.4736842	0.8596491	0.4625987	0.4917609	0.8803568	0.4813422	1.9281	9.1555	6.3229	-41.4270	0.6340	2.7320
50	2007	1261.411	3786.419	6	3	0.1052632	0.0526316	0.0499804	0.0484307	0.0215665	0.0207525	1.0212	1.0518	38251.3348	-5019253.2883	0.0284	1.0292
51	2008	3842.252	7599.782	42	28	0.7368421	0.4912281	0.4612300	0.7347105	0.4644827	0.4394306	1.7839	4.1625	8.3847	-58.1540	0.5836	2.4016
52	2009	2470.312	5922.941	22	15	0.3859649	0.2631579	0.2372505	0.3771462	0.2168300	0.2024194	1.2538	1.6435	76.4627	-1042.0584	0.2758	1.3808
53	2010	6399.943	12521.900	56	52	0.9824561	0.9122807	0.8997261	0.9679294	0.9041048	0.8817550	8.4570	102.8722	0.4087	-4.3503	0.9757	41.1558
54	2011	1098.361	5392.886	3	9	0.0526316	0.1578947	0.0516991	0.0256774	0.1483940	0.0255989	1.0263	1.1744	21598.9595	-2297648.5383	0.0350	1.0363
55	2012	1228.784	5157.358	5	6	0.0877193	0.1052632	0.0736891	0.0432631								

**Tabla 7.22.** Análisis de Frecuencia Bivariado de Avenidas de la Estación Jesús Carranza II; Cuenca del Río Coatzacoalcos

ORDEN	DATOS OBSERVADOS			RANKING		DATOS TRANSFORMADOS		CÓPULA	MARGINALES UNIFORMES		CÓPULA	PERIODOS DE RETORNO BIVARIADOS		OBTENIDOS CON CÓPULA CLAYTON			
	ESTACIÓN: J. CARRANZA II			CLASIFICACIÓN		$Y_{ij} \in [0, 1]$		EMPÍRICA	LOGNORMAL 3P	GUMBEL	CLAYTON	"OR"	"AND"	GENERADOR	DERIVADA	FUN. KENDALL	"KENDALL"
	i	Año	Q m <sup>3</sup> /s	V Hm <sup>3</sup>	R <sub>Qi</sub>	R <sub>Vi</sub>	R <sub>Qi</sub> /N	R <sub>Vi</sub> /N	C(R <sub>Qi</sub> /N, R <sub>Vi</sub> /N) (Probabilidad)	F(Q)=u <sub>1</sub> (Probabilidad)	F(V)=u <sub>2</sub> (Probabilidad)	C(u <sub>1</sub> ,u <sub>2</sub> ) (Probabilidad)	T <sup>V</sup> (u <sub>1</sub> ,u <sub>2</sub> ) (años)	T <sup>A</sup> (u <sub>1</sub> ,u <sub>2</sub> ) (años)	φ (t)	φ' (t)	Kc (t) (Probabilidad)
1	1953	463.000	264.317	7	9	0.1346154	0.1730769	0.1176435	0.1347132	0.1232123	0.1012324	1.1126	1.1858	773.0019	-22204.8414	0.1360	1.1575
2	1954	1043.875	1027.261	34	49	0.6538462	0.9423077	0.6419359	0.6412127	0.9296452	0.6274744	2.6844	17.6627	2.8710	-17.9163	0.7877	4.7107
3	1955	844.249	528.853	23	23	0.4423077	0.4423077	0.3541895	0.4901745	0.5200663	0.4070087	1.6864	2.5204	12.6077	-97.0973	0.5369	2.1592
4	1956	1295.875	1223.053	42	51	0.8076923	0.9807692	0.7991927	0.7755267	0.9696453	0.7639219	4.2359	53.3337	1.1860	-8.3104	0.9066	10.7104
5	1957	929.625	416.056	27	17	0.5192308	0.3269231	0.3051133	0.5598021	0.3416532	0.3214054	1.4736	2.3812	26.0154	-244.1089	0.4280	1.7482
6	1958	1158.125	797.909	39	43	0.7500000	0.8269231	0.6817236	0.7093321	0.8186262	0.6496610	2.8544	8.2167	2.4994	-15.6434	0.8094	5.2475
7	1959	563.437	230.696	12	6	0.2307692	0.1153846	0.1105726	0.2253722	0.0882415	0.0863543	1.0945	1.2941	1227.1020	-41302.4305	0.1161	1.1313
8	1960	1610.187	969.758	47	48	0.9038462	0.9230769	0.8500411	0.8761300	0.9103236	0.8196146	5.5437	30.1559	0.7819	-6.3141	0.9435	17.6854
9	1961	1464.520	911.077	46	47	0.8846154	0.9038462	0.8216835	0.8368529	0.8854723	0.7744111	4.4328	19.1990	1.1011	-7.8795	0.9142	11.6487
10	1962	919.583	650.001	26	34	0.5000000	0.6538462	0.4537956	0.5519956	0.6813648	0.4959766	1.9840	3.8078	6.6637	-44.8748	0.6445	2.8127
11	1963	1178.874	710.303	41	39	0.7884615	0.7500000	0.6629094	0.7203607	0.7450920	0.6235189	2.6562	6.3265	2.9427	-18.3642	0.7838	4.6245
12	1964	863.267	459.628	25	19	0.4807692	0.3265384	0.3256568	0.5063111	0.4120552	0.3605589	1.5639	2.2615	18.3475	-155.8385	0.4783	1.9168
13	1965	742.389	295.910	19	12	0.3653846	0.2307692	0.2137535	0.3983521	0.1616847	0.1580701	1.1877	1.6721	211.1745	-3898.2402	0.2122	1.2694
14	1966	1099.945	668.909	38	37	0.7307692	0.7115385	0.6114265	0.6762464	0.7025561	0.5784545	2.3722	5.0087	3.9025	-24.6134	0.7370	3.8024
15	1967	1074.025	363.241	36	14	0.6923077	0.2692308	0.2654214	0.6604441	0.2579688	0.2540330	1.3405	2.9796	52.4949	-611.5722	0.3399	1.5149
16	1968	1903.280	885.963	48	46	0.9230769	0.8846154	0.8348059	0.9283659	0.8729774	0.8288962	5.8444	36.2938	0.7246	-6.0425	0.9488	19.5371
17	1969	1924.668	1124.572	49	50	0.9423077	0.9615385	0.9114508	0.9311386	0.9535693	0.8954063	9.5608	93.4722	0.3783	-4.4704	0.9800	50.0715
18	1970	1353.315	540.811	43	26	0.8269231	0.5000000	0.4841005	0.7986146	0.5377881	0.5121990	2.0500	5.6884	5.9798	-39.5760	0.6633	2.9700
19	1971	1032.827	535.701	31	24	0.5961538	0.4615385	0.4141582	0.6339333	0.5302566	0.4680368	1.8798	3.2911	8.0692	-56.2749	0.6114	2.5735
20	1972	1021.627	662.981	30	35	0.5769231	0.6730769	0.5092520	0.6264253	0.6960312	0.5474106	2.2095	4.4454	4.7542	-30.5281	0.7031	3.3686
21	1973	2908.776	1231.761	52	52	1	1	1.0000000	0.9876866	0.9707695	0.9597812	24.8640	754.6974	0.1266	-3.4090	0.9969	324.9132
22	1976	709.586	478.387	18	21	0.3461538	0.4038462	0.2948642	0.3671945	0.4420396	0.3172496	1.4647	1.9684	27.0560	-256.8329	0.4226	1.7319
23	1979	1967.920	630.104	50	33	0.9615385	0.6346154	0.6277260	0.9364071	0.6578618	0.6443114	2.8114	19.9830	2.5844	-16.1567	0.8043	5.1092
24	1981	1035.372	479.946	33	22	0.6346154	0.4230769	0.3944336	0.6356213	0.4445133	0.4106203	1.6967	3.0258	12.2630	-93.8053	0.5413	2.1803
25	1982	2041.277	566.359	51	30	0.9807692	0.5769231	0.5746086	0.9443959	0.5744517	0.5674735	2.3120	20.5652	4.1831	-26.5258	0.7252	3.6386
26	1983	1065.129	833.070	35	45	0.6730769	0.8653846	0.6385476	0.6548656	0.8424491	0.6170219	2.6111	8.3537	3.0645	-19.1307	0.7772	4.4885
27	1984	954.854	666.248	28	36	0.5384615	0.6923077	0.4898418	0.5789522	0.6996402	0.5191477	2.0796	4.1570	5.7120	-37.5477	0.6713	3.0420
28	1985	679.601	733.054	16	41	0.3076923	0.7884615	0.3043301	0.3382595	0.7662718	0.3326142	1.4984	4.3844	23.4554	-213.5304	0.4425	1.7936
29	1986	1387.380	703.318	45	38	0.8653846	0.7307692	0.6843752	0.8111885	0.7382803	0.6664067	2.9977	8.5515	2.2501	-14.1638	0.8253	5.7231
30	1987	968.670	563.598	29	29	0.5576923	0.5576923	0.4540741	0.5891575	0.5705730	0.4738469	1.9006	3.1835	7.7500	-53.6286	0.6184	2.6203
31	1988	844.267	737.115	24	42	0.4615385	0.8076923	0.4479040	0.4901900	0.7698930	0.4680847	1.8800	4.8077	8.0665	-56.2525	0.6115	2.5739
32	1993	513.018	572.835	9	31	0.1730769	0.5961538	0.1718187	0.1783239	0.5834676	0.1767968	1.2148	2.4096	152.2775	-2517.8462	0.2373	1.3111
33	1994	228.800	235.820	1	7	0.0192308	0.1346154	0.0192076	0.0099099	0.0931486	0.0099048	1.0100	1.1027	661371.7578	-193921328.7249	0.0133	1.0135
34	1995	1387.100	556.790	44	28	0.8461538	0.5384615	0.5205165	0.8110883	0.5609211	0.5335919	2.1440	6.1888	5.1978	-33.7328	0.6877	3.2018
35	1996	1165.443	720.520	40	40	0.7692308	0.7692308	0.6639644	0.7132669	0.7547923	0.6241226	2.6604	6.4077	2.9317	-18.2949	0.7844	4.6375
36	1997	1034.651	812.417	32	44	0.6153846	0.8461538	0.5860415	0.6351438	0.8288206	0.5975203	2.4846	7.4875	3.4618	-21.6864	0.7572	4.1178
37	1998	662.346	399.939	15	16	0.2884615	0.3076923	0.2355158	0.3214739	0.3157249	0.2524897	1.3378	1.6252	53.4501	-626.2963	0.3378	1.5102
38	1999	685.447	551.801	17	27	0.3269231	0.5192308	0.3051133	0.3439274	0.5537679	0.3224746	1.4760	2.3542	25.7561	-240.9642	0.4294	1.7524
39	2000	757.476	290.698	20	10	0.3846154	0.1923077	0.1846501	0.4124619	0.1549927	0.1522198	1.1796	1.7101	235.7338	-4516.6290	0.2044	1.2569
40	2001	656.626	141.249	14	2	0.2692308	0.0384615	0.0384161	0.3158952	0.0273660	0.0273586	1.0281	1.4618	34591.3197	-3672075.4830	0.0368	1.0382
41	2002	529.254	219.770	10	5	0.1923077	0.0961538	0.0921220	0.1932064	0.0782974	0.0764462	1.0828	1.2423	1748.6346	-66468.5968	0.1028	1.1145
42	2003	1091.737	628.348	37	32	0.7115385	0.6153846	0.5464771	0.6713145	0.6557285	0.5529732	2.2370	4.4261	4.5877	-29.3466	0.7093	3.4400
43	2004	843.185	421.695	22	18	0.4230769	0.3461538	0.3001940	0.4892616	0.3507647	0.3177351	1.4657	2.0933	26.9317	-255.3040	0.4232	1.7338
44	2005	316.693	201.572	3	4	0.0576923	0.0769231	0.0509653	0.0382031	0.0633178	0.0355701	1.0369	1.0706	16139.7765	-1317844.2570	0.0478	1.0502
45	2006	331.694	144.191	4	3	0.0769231	0.0576923	0.0509653	0.0453799	0.0286625	0.0264463	1.0272	1.0500	38171.8290	-4191938.0818	0.0356	1.0369
46	2007	239.622	136.322	2	1	0.0384615	0.0192308	0.0184182	0.0121814	0.0252907	0.0117158	1.0119	1.0264	406118.8240	-100671628.3918	0.0157	1.0160
47	2008	371.301	364.398	5	15	0.0961538	0.2884615	0.0948629	0.0675720	0.2597502	0.0671213	1.0720	1.3517	2551.8337	-110455.5795	0.0902	1.0992
48	2009	609.083	262.742	13	8	0.2500000	0.1538462	0.1428712	0.2694669	0.1214311	0.1176366	1.1333	1.3760	499.4058	-12353.9288	0.1581	1.1877
49	2010	791.234	537.973	21	25	0.4038462	0.4807692	0.3487695	0.4434383	0.5336135	0.3865553	1.6301	2.4420	14.8058	-118.7490	0.5112	2.0460
50	2011	504.806	292.760	8	11	0.1538462	0.2115385	0.1372784	0.1709156	0.1576241	0.1290947	1.1482	1.2491	381.0248	-8594.2539	0.1734	1.2098
51	2012	549.551	316.397	11	13	0.2115385	0.2500000	0.1797521	0.2121884	0.1891814	0.1573159	1.1867	1.3228	214.1421	-3971.7135	0.2112	1.2678
52	2013	411.467	478.026	6	20	0.1153846	0.3846154	0.1142769	0.0945377	0.4414660	0.0942040	1.1040	1.7915	952.8915	-29407.2851	0.1266	1.1450

**Tabla 7.23.** Análisis de Frecuencia Bivariado de Avenidas de la Estación Tierra Morada; Cuenca del Río Coatzacoalcos

ORDEN	DATOS OBSERVADOS			RANKING		DATOS TRANSFORMADOS		CÓPULA	MARGINALES UNIFORMES			CÓPULA	PERIODOS DE RETORNO BIVARIADOS			OBTENIDOS CON CÓPULA GUMBEL - H		
	ESTACIÓN: T. MORADA			CLASIFICACIÓN		$Y_{ij} \in [0, 1]$		EMPÍRICA	DOBLE GUMBEL		NORMAL	GUMBEL - H	"OR"	"AND"	GENERADOR	DERIVADA	FUN. KENDALL	"KENDALL"
	i	Año	Q m <sup>3</sup> /s	V Hm <sup>3</sup>	R <sub>Qi</sub>	R <sub>Vi</sub>	R <sub>Qi</sub> /N	R <sub>Vi</sub> /N	C(R <sub>Qi</sub> /N, R <sub>Vi</sub> /N) (Probabilidad)	F(Q)=u <sub>1</sub> (Probabilidad)	F(V)=u <sub>2</sub> (Probabilidad)	C(u <sub>1</sub> ,u <sub>2</sub> ) (Probabilidad)	T <sup>V</sup> (u <sub>1</sub> ,u <sub>2</sub> ) (años)	T <sup>A</sup> (u <sub>1</sub> ,u <sub>2</sub> ) (años)	φ (t)	φ' (t)	Kc (t) (Probabilidad)	T <sub>K</sub> (u <sub>1</sub> ,u <sub>2</sub> ) (años)
1	1975	1171.189	6237.837	3	32	0.0833333	0.8888889	0.0831606	0.0833309	0.8212187	0.0811514	1.0883	5.6625	4.3787	-34.4543	0.2082	1.2630	
2	1976	1271.945	4527.039	7	11	0.1944444	0.3055556	0.1358960	0.2145237	0.3270185	0.1271501	1.1457	1.7076	3.1926	-19.5244	0.2907	1.4098	
3	1977	1419.625	3680.570	16	6	0.4444444	0.1666667	0.1424039	0.4578502	0.1302714	0.1001474	1.1113	1.9530	3.8057	-26.4832	0.2438	1.3225	
4	1978	1716.212	5928.174	29	29	0.8055556	0.8055556	0.7396643	0.7784725	0.7493265	0.6597839	2.9393	7.5766	0.2448	-1.4311	0.8309	5.9126	
5	1979	1527.332	5478.731	21	24	0.5833333	0.6666667	0.5139163	0.6098440	0.6228362	0.4743479	1.9024	4.1379	0.6248	-2.8322	0.6950	3.2782	
6	1980	1347.556	5116.982	11	18	0.3055556	0.5000000	0.2571910	0.3390350	0.5094283	0.2529131	1.3385	2.4725	1.6659	-7.6839	0.4697	1.8858	
7	1981	1485.977	5365.303	20	23	0.5555556	0.6388889	0.4822022	0.5562094	0.5879310	0.4224466	1.7314	3.5932	0.7876	-3.4699	0.6494	2.8526	
8	1982	1792.200	5706.877	31	26	0.8611111	0.7222222	0.7012630	0.8179521	0.6898389	0.6359903	2.7472	7.8004	0.2804	-1.5625	0.8155	5.4192	
9	1983	1704.797	4960.950	27	15	0.7500000	0.4166667	0.4001985	0.7713854	0.4597164	0.4239575	1.7360	5.1852	0.7824	-3.4489	0.6508	2.8638	
10	1984	1409.545	4354.193	15	8	0.4166667	0.2222222	0.1789201	0.4417842	0.2788062	0.1946212	1.2417	2.1096	2.2036	-11.0942	0.3933	1.6481	
11	1987	1322.793	3107.467	9	4	0.2500000	0.1111111	0.0767496	0.2973425	0.0566608	0.0368408	1.0382	1.4645	6.7885	-89.5144	0.1127	1.1270	
12	1988	1429.465	4028.189	18	7	0.5000000	0.1944444	0.1714626	0.4732834	0.1984676	0.1501931	1.1767	2.0901	2.7893	-15.7097	0.3277	1.4875	
13	1989	1600.253	4359.807	23	9	0.6388889	0.2500000	0.2350342	0.6890443	0.2803165	0.2516396	1.3363	3.5426	1.6757	-7.7399	0.4681	1.8802	
14	1990	1458.875	4872.658	19	14	0.5277778	0.3888889	0.3242424	0.5177375	0.4318138	0.3137639	1.4572	2.7456	1.2672	-5.5876	0.5405	2.1765	
15	1991	1624.675	4863.636	24	13	0.6666667	0.3611111	0.3369791	0.7113490	0.4289793	0.3813317	1.6164	4.1493	0.9430	-4.1136	0.6106	2.5679	
16	1992	1234.824	5231.154	5	20	0.1388889	0.5555556	0.1289339	0.1595248	0.5457545	0.1322333	1.1524	2.3422	3.0959	-18.5576	0.2991	1.4267	
17	1993	1379.116	5146.921	13	19	0.3611111	0.5277778	0.3047518	0.3919864	0.5189741	0.2899069	1.4083	2.6389	1.4087	-6.2932	0.5137	2.0565	
18	1994	993.897	2878.562	2	2	0.0555556	0.0555556	0.0177548	0.0057671	0.0386559	0.0013915	1.0014	1.0450	20.5065	-3593.1640	0.0071	1.0071	
19	1995	1625.988	7186.606	25	35	0.6944444	0.9722222	0.6938600	0.7124915	0.9534026	0.7060811	3.4023	24.8836	0.1840	-1.2010	0.8593	7.1080	
20	1996	1748.050	5276.912	30	22	0.8333333	0.6111111	0.5937093	0.7965568	0.5602238	0.5184080	2.0764	6.1871	0.5098	-2.4005	0.7308	3.7146	
21	1997	1633.918	4366.273	26	10	0.7222222	0.2777778	0.2682603	0.7192725	0.2820607	0.2579630	1.3476	3.8967	1.6276	-7.4679	0.4759	1.9081	
22	1998	2961.060	5989.554	35	30	0.9722222	0.8333333	0.8318535	0.9444709	0.7646867	0.7542132	4.0686	22.1948	0.1314	-0.9904	0.8869	8.8399	
23	1999	4017.361	8501.140	36	36	1	1	1	0.9739756	0.9968338	0.9734428	37.6546	379.7339	0.0030	-0.1858	0.9898	97.8584	
24	2000	1815.872	6164.973	32	31	0.8888889	0.8611111	0.8287589	0.8278775	0.8055869	0.7316357	3.7263	10.1863	0.1548	-1.0861	0.8742	7.9487	
25	2001	1574.334	5893.518	22	28	0.6111111	0.7777778	0.5784898	0.6631094	0.7404251	0.5758260	2.3575	5.8041	0.3856	-1.9454	0.7740	4.4250	
26	2002	1356.309	2883.500	12	3	0.3333333	0.0833333	0.0676680	0.3537881	0.0389881	0.0284068	1.0292	1.5732	7.6660	-121.5274	0.0915	1.1007	
27	2003	1317.131	5563.175	8	25	0.2222222	0.6944444	0.2141110	0.2878703	0.6482018	0.2506596	1.3345	3.1788	1.6833	-7.7834	0.4669	1.8759	
28	2004	1387.614	5033.078	14	17	0.3888889	0.4722222	0.3041642	0.4060603	0.4826654	0.2837474	1.3962	2.5315	1.4480	-6.4969	0.5066	2.0269	
29	2005	1329.181	7129.018	10	34	0.2777778	0.9444444	0.2775159	0.3080672	0.9487366	0.3065246	1.4420	20.1123	1.3084	-5.7888	0.5325	2.1392	
30	2006	1235.790	3441.189	6	5	0.1666667	0.1388889	0.0721235	0.1608724	0.0939948	0.0389682	1.0405	1.2753	6.6043	-83.7559	0.1178	1.1336	
31	2008	2374.517	6264.537	34	33	0.9444444	0.9166667	0.9022708	0.9182914	0.8267429	0.8013297	5.0335	17.7634	0.0892	-0.8056	0.9120	11.3638	
32	2009	1987.238	5250.013	33	21	0.9166667	0.5833333	0.5799866	0.8765093	0.5517264	0.5340495	2.1462	9.4506	0.4733	-2.2660	0.7429	3.8901	
33	2010	1213.470	4716.113	4	12	0.1111111	0.3333333	0.0878328	0.1310765	0.3832514	0.0913441	1.1005	1.7331	4.0526	-29.7309	0.2277	1.2948	
34	2011	1715.450	5874.413	28	27	0.7777778	0.7500000	0.6860631	0.7780097	0.7354491	0.6496049	2.8539	7.3450	0.2597	-1.4861	0.8243	5.6931	
35	2012	919.033	2790.896	1	1	0.0277778	0.0277778	0.0067527	0.0030505	0.0331337	0.0007243	1.0007	1.0368	23.8673	-7308.3435	0.0040	1.0040	
36	2013	1427.064	4977.438	17	16	0.4722222	0.4444444	0.3363545	0.4695416	0.4649533	0.3096443	1.4485	2.6656	1.2904	-5.7008	0.5360	2.1552	

**Tabla 7.24.** Análisis de Frecuencia Bivariado de Avenidas de la Estación Minzapan; Cuenca del Río Coatzacoalcos

ORDEN	DATOS OBSERVADOS			RANKING		DATOS TRANSFORMADOS		CÓPULA	MARGINALES UNIFORMES			CÓPULA	PERIODOS DE RETORNO BIVARIADOS OBTENIDOS CON CÓPULA CLAYTON				
	ESTACIÓN: MINZAPAN			CLASIFICACIÓN		Y <sub>ij</sub> ∈ [ 0, 1 ]		EMPÍRICA	DOBLE GUMBEL	DOBLE GUMBEL	CLAYTON	"OR"	"AND"	GENERADOR	DERIVADA	FUN. KENDALL	"KENDALL"
	i	Año	Q m <sup>3</sup> /s	V Hm <sup>3</sup>	R <sub>Qi</sub>	R <sub>Vi</sub>	R <sub>Qi</sub> /N	R <sub>Vi</sub> /N	C(R <sub>Qi</sub> /N , R <sub>Vi</sub> /N) (Probabilidad)	F(Q)=u <sub>1</sub> (Probabilidad)	F(V)=u <sub>2</sub> (Probabilidad)	C(u <sub>1</sub> ,u <sub>2</sub> ) (Probabilidad)	T <sup>V</sup> (u <sub>1</sub> ,u <sub>2</sub> ) (años)	T <sup>^</sup> (u <sub>1</sub> ,u <sub>2</sub> ) (años)	φ (t)	φ' (t)	Kc (t) (Probabilidad)
1	1974	164.662	61.925	27	35	0.6923077	0.8974359	0.6581206	0.7180947	0.8611051	0.6646457	2.9819	11.7033	1.5030	-8.4579	0.8423	6.3429
2	1975	119.742	76.034	17	38	0.4358974	0.9743590	0.4341033	0.4534400	0.9579269	0.4500188	1.8182	25.8719	5.0093	-29.9914	0.6170	2.6113
3	1976	336.925	57.392	39	33	1	0.8461538	0.8461538	0.9700533	0.8071021	0.7918862	4.8051	67.8854	0.6889	-4.7900	0.9357	15.5526
4	1977	94.206	14.912	12	4	0.3076923	0.1025641	0.0991584	0.2386144	0.0549453	0.0540992	1.0572	1.3149	699.1532	-29067.1840	0.0782	1.0848
5	1978	156.555	39.990	26	23	0.6666667	0.5897436	0.4992296	0.6839327	0.5580678	0.4861073	1.9459	4.0966	4.0534	-23.3482	0.6597	2.9387
6	1979	91.247	28.777	10	14	0.2564103	0.3589744	0.2191648	0.2144070	0.3537134	0.1912019	1.2364	1.6049	40.0898	-482.6609	0.2743	1.3779
7	1980	309.789	83.327	38	39	0.9743590	1	0.9743590	0.9553368	0.9784402	0.9366856	15.7942	343.8028	0.1582	-2.7772	0.9937	157.8376
8	1981	197.514	39.357	34	22	0.8717949	0.5641026	0.5407177	0.8147029	0.5489490	0.5154352	2.0637	6.5883	3.4304	-19.3050	0.6931	3.2587
9	1982	206.141	61.518	35	34	0.8974359	0.8717949	0.8031767	0.8326417	0.8567277	0.7472208	3.9560	17.2857	0.9241	-5.7833	0.9070	10.7535
10	1983	195.031	38.272	33	21	0.8461538	0.5384615	0.5133225	0.8091148	0.5328952	0.5010975	2.0044	6.2859	3.7202	-21.1562	0.6769	3.0954
11	1986	175.553	46.713	29	26	0.7435897	0.6666667	0.5775560	0.7565158	0.6528140	0.5741276	2.3481	6.0680	2.4774	-13.6034	0.7562	4.1025
12	1987	178.910	51.685	30	31	0.7692308	0.7948718	0.6677004	0.7668850	0.7262463	0.6257277	2.6719	7.5417	1.8662	-10.2878	0.8071	5.1848
13	1988	91.131	35.167	9	17	0.2307692	0.4358974	0.2125741	0.2134722	0.4828888	0.2022689	1.2536	1.9766	35.2118	-402.0883	0.2898	1.4081
14	1989	64.639	16.405	3	5	0.0769231	0.1282051	0.0681085	0.0497502	0.0769052	0.0431764	1.0451	1.0911	1160.9086	-60440.2743	0.0624	1.0665
15	1990	63.021	18.874	2	7	0.0512821	0.1794872	0.0499956	0.0438144	0.1219976	0.0420046	1.0438	1.1413	1234.9741	-66086.5500	0.0607	1.0646
16	1991	120.847	64.333	18	36	0.4615385	0.9230769	0.4545841	0.4621657	0.8849869	0.4511292	1.8219	9.6175	4.9762	-29.7525	0.6184	2.6204
17	1992	131.166	27.607	21	12	0.5384615	0.3076923	0.2822717	0.5387085	0.3264642	0.2962475	1.4210	2.3198	14.3687	-116.5155	0.4196	1.7229
18	1993	152.583	65.436	25	37	0.6410256	0.9487179	0.6282557	0.6652080	0.8947848	0.6341085	2.7331	13.4924	1.7818	-9.8530	0.8150	5.4040
19	1994	73.302	18.229	5	6	0.1282051	0.1538462	0.1024695	0.0900945	0.1092458	0.0722025	1.0778	1.1457	365.1313	-11388.9878	0.1043	1.1164
20	1995	92.768	12.643	11	1	0.2820513	0.0256410	0.0255919	0.2267671	0.0298195	0.0296861	1.0306	1.2935	2694.0915	-203901.9489	0.0429	1.0448
21	1996	172.440	51.616	28	30	0.7179487	0.7692308	0.6217624	0.7463166	0.7252145	0.6143276	2.5929	7.0030	1.9871	-10.9205	0.7963	4.9088
22	1997	236.153	41.322	36	24	0.9230769	0.6153846	0.5980725	0.8822351	0.5768972	0.5541689	2.2430	10.5223	2.7650	-15.2590	0.7354	3.7789
23	1998	87.736	27.781	8	13	0.2051282	0.3333333	0.1820411	0.1866835	0.3305592	0.1687833	1.2031	1.5348	53.3727	-723.5225	0.2426	1.3202
24	1999	119.539	54.658	16	32	0.4102564	0.8205128	0.3971535	0.4518271	0.7696162	0.4271515	1.7457	4.8613	5.7561	-35.5233	0.5892	2.4342
25	2000	82.235	26.328	7	11	0.1794872	0.2820513	0.1574933	0.1461222	0.2958520	0.1351420	1.1563	1.4427	88.5784	-1488.7225	0.1946	1.2417
26	2001	126.194	35.934	19	18	0.4871795	0.4615385	0.3635142	0.5030016	0.4959214	0.3854018	1.6271	2.5875	7.5117	-49.6026	0.5368	2.1591
27	2002	80.424	14.761	6	3	0.1538462	0.0769231	0.0707245	0.1336935	0.0529663	0.0502875	1.0530	1.1579	824.0096	-36846.7913	0.0727	1.0783
28	2003	180.764	45.892	31	25	0.7948718	0.6410256	0.5807078	0.7723500	0.6410017	0.5723701	2.3385	6.2886	2.5014	-13.7395	0.7544	4.0722
29	2004	52.738	13.351	1	2	0.0256410	0.0512821	0.0235498	0.0169400	0.0365984	0.0157534	1.0160	1.0393	11183.3793	-1594551.4829	0.0228	1.0233
30	2005	186.558	47.734	32	27	0.8205128	0.6923077	0.6279653	0.7883347	0.6676658	0.5971407	2.4823	7.0852	2.1836	-11.9742	0.7795	4.5352
31	2006	141.514	26.142	22	10	0.5641026	0.2564103	0.2434886	0.6053919	0.2913444	0.2758272	1.3809	2.6379	17.0427	-146.9148	0.3918	1.6443
32	2007	283.417	48.083	37	28	0.9487179	0.7179487	0.6996705	0.9354121	0.6727914	0.6537980	2.8885	21.9325	1.5972	-8.9220	0.8328	5.9814
33	2008	99.005	25.710	13	9	0.3333333	0.2307692	0.1986170	0.2790103	0.2808555	0.2082736	1.2631	1.5422	32.9090	-365.6633	0.2983	1.4251
34	2009	116.796	37.464	15	20	0.3846154	0.5128205	0.3303557	0.4297486	0.5205399	0.3587486	1.5595	2.4482	8.9982	-62.5936	0.5025	2.0101
35	2010	145.523	50.598	24	29	0.6153846	0.7435897	0.5442545	0.6284021	0.7100687	0.5396410	2.1722	4.9709	2.9965	-16.6332	0.7198	3.5688
36	2011	109.505	20.501	14	8	0.3589744	0.2051282	0.1853293	0.3689460	0.1566279	0.1483073	1.1741	1.6058	71.6991	-1100.9492	0.2134	1.2713
37	2012	129.288	31.287	20	16	0.5128205	0.4102564	0.3455213	0.5254914	0.4088042	0.3480963	1.5340	2.4166	9.6985	-69.0275	0.4886	1.9554
38	2013	142.282	31.176	23	15	0.5897436	0.3846154	0.3462804	0.6099212	0.4064604	0.3654876	1.5760	2.8645	8.5889	-58.9243	0.5112	2.0460
39	2014	70.540	37.452	4	19	0.1025641	0.4871795	0.1014778	0.0756638	0.5203464	0.0753251	1.0815	2.0863	331.9193	-9926.5772	0.1088	1.1220

**Tabla 7.25.** Análisis de Frecuencia Bivariado de Avenidas de la Estación Minatitlán; Cuenca del Río Coatzacoalcos

ORDEN	DATOS OBSERVADOS			RANKING		DATOS TRANSFORMADOS		CÓPULA	MARGINALES UNIFORMES			CÓPULA	PERIODOS DE RETORNO BIVARIADOS OBTENIDOS CON CÓPULA FRANK					
	ESTACIÓN: MINATITLÁN			CLASIFICACIÓN		Y <sub>ij</sub> ∈ [ 0, 1 ]		EMPÍRICA	GUMBEL	LOGNORMAL3P		FRANK	"OR"	"AND"	GENERADOR	DERIVADA	FUN. KENDALL	"KENDALL"
	i	Año	Q m <sup>3</sup> /s	V Hm <sup>3</sup>	R <sub>Qi</sub>	R <sub>Vi</sub>	R <sub>Qi</sub> /N	R <sub>Vi</sub> /N	C(R <sub>Qi</sub> /N , R <sub>Vi</sub> /N) (Probabilidad)	F(Q)=u <sub>1</sub> (Probabilidad)	F(V)=u <sub>2</sub> (Probabilidad)	C(u <sub>1</sub> ,u <sub>2</sub> ) (Probabilidad)	T <sup>V</sup> (u <sub>1</sub> ,u <sub>2</sub> ) (años)	T <sup>^</sup> (u <sub>1</sub> ,u <sub>2</sub> ) (años)	φ (t)	φ' (t)	Kc (t) (Probabilidad)	T <sub>K</sub> (u <sub>1</sub> ,u <sub>2</sub> ) (años)
1	1975	3960.814	12792.656	16	14	0.4848485	0.4242424	0.3203893	0.4866963	0.3941551	0.3036340	1.4360	2.3653	0.2865	-1.5505	0.4884	1.9547	
2	1976	3157.922	11719.436	9	9	0.2727273	0.2727273	0.1559251	0.2528127	0.2758711	0.1485114	1.1744	1.6134	0.7125	-4.7459	0.2986	1.4258	
3	1978	7037.829	16613.017	32	27	0.9696970	0.8181818	0.8044150	0.9414262	0.8000194	0.7745700	4.4360	30.1892	0.0204	-0.1451	0.9152	11.7971	
4	1981	3924.047	15148.540	14	22	0.4242424	0.6666667	0.3787133	0.4762722	0.6647286	0.4178352	1.7177	3.6123	0.1566	-0.8177	0.6093	2.5598	
5	1982	5375.723	20107.457	28	32	0.8484848	0.9696970	0.8327408	0.7942840	0.9634569	0.7792369	4.5297	46.5204	0.0197	-0.1420	0.9182	12.2310	
6	1983	4082.258	18129.470	17	31	0.5151515	0.9393939	0.5079967	0.5204859	0.8964087	0.5067862	2.0275	11.1245	0.0983	-0.5186	0.6964	3.2933	
7	1984	4193.450	15261.417	20	23	0.6060606	0.6969697	0.5203107	0.5504398	0.6764834	0.4746635	1.9035	4.0365	0.1164	-0.6095	0.6656	2.9906	
8	1987	2627.356	12093.419	4	12	0.1212121	0.3636364	0.0928747	0.1214144	0.3154192	0.0863172	1.0945	1.5397	1.1284	-9.4957	0.2051	1.2581	
9	1988	3750.334	13368.214	12	16	0.3636364	0.4848485	0.2842647	0.4260505	0.4617179	0.31115184	1.4525	2.3599	0.2746	-1.4791	0.4971	1.9886	
10	1989	5156.578	11453.083	25	8	0.7575758	0.2424242	0.2327082	0.7597320	0.2491111	0.2391036	1.3142	4.3429	0.4098	-2.3396	0.4142	1.7072	
11	1990	3628.518	8157.220	5	3	0.1515152	0.0909091	0.0406350	0.1216542	0.0417097	0.0167461	1.0170	1.1718	2.6189	-57.5106	0.0623	1.0664	
12	1991	5632.659	12025.152	29	11	0.8787879	0.3333333	0.3270042	0.8292378	0.3080434	0.2993995	1.4273	6.1683	0.2932	-1.5907	0.4837	1.9368	
13	1992	3145.874	13431.360	8	17	0.2424242	0.5151515	0.2062653	0.2494445	0.4691892	0.2029149	1.2546	2.0649	0.5062	-3.0282	0.3701	1.5875	
14	1993	3650.578	14736.010	11	21	0.3333333	0.6363636	0.2996072	0.3966896	0.6202374	0.3462407	1.5296	3.0366	0.2281	-1.2092	0.5349	2.1499	
15	1994	1675.660	6704.137	1	1	0.0303030	0.0303030	0.0036582	0.0106814	0.0132281	0.0006060	1.0006	1.0239	5.9021	-1647.8748	0.0042	1.0042	
16	1995	5275.151	16606.251	26	26	0.7878788	0.7878788	0.6824899	0.7789930	0.7994876	0.6838644	3.1632	9.4891	0.0368	-0.2212	0.8501	6.6728	
17	1996	5110.423	12986.947	24	15	0.7272727	0.4545455	0.4168944	0.7518637	0.4167966	0.3897875	1.6388	4.5223	0.1813	-0.9501	0.5807	2.3847	
18	1997	4153.879	12313.410	19	13	0.5757576	0.3939394	0.3324559	0.5398959	0.3396114	0.2840662	1.3968	2.4718	0.3187	-1.7475	0.4665	1.8742	
19	1998	4097.641	11271.250	18	6	0.5454545	0.1818182	0.1612220	0.5246890	0.2315923	0.1993112	1.2489	2.2572	0.5172	-3.1116	0.3655	1.5761	
20	1999	6503.965	17211.417	31	28	0.9393939	0.8484848	0.8159779	0.9113763	0.8434269	0.7955332	4.8908	24.5519	0.0175	-0.1318	0.9284	13.9743	
21	2000	3832.553	16455.574	13	25	0.3939394	0.7575758	0.3707324	0.4499986	0.7874097	0.4248126	1.7386	5.3361	0.1510	-0.7882	0.6164	2.6068	
22	2001	3926.905	13551.066	15	18	0.4545455	0.5454545	0.3624554	0.4770856	0.4833557	0.3478820	1.5335	2.5810	0.2261	-1.1980	0.5366	2.1581	
23	2002	4642.538	10557.501	21	4	0.6363636	0.1212121	0.1137183	0.6598249	0.1693624	0.1590934	1.1892	3.0312	0.6646	-4.3152	0.3131	1.4559	
24	2003	5285.819	15771.178	27	24	0.8181818	0.7272727	0.6583738	0.7806598	0.7269958	0.6395319	2.7742	7.5829	0.0477	-0.2725	0.8145	5.3899	
25	2004	4729.402	17271.696	22	30	0.6666667	0.9090909	0.6435945	0.6786313	0.8474036	0.6333840	2.7276	9.3154	0.0494	-0.2806	0.8094	5.2453	
26	2005	4941.280	22075.669	23	33	0.6969697	1	0.6969697	0.7212000	0.9893677	0.7181750	3.5483	131.4520	0.0298	-0.1884	0.8761	8.0730	
27	2006	3024.572	14644.227	7	20	0.2121212	0.6060606	0.1931736	0.2163035	0.6100452	0.1971955	1.2456	2.6965	0.5239	-3.1620	0.3629	1.5695	
28	2008	5982.003	13633.401	30	19	0.9090909	0.5757576	0.5605284	0.8682177	0.4930940	0.4769047	1.9117	8.6510	0.1150	-0.6026	0.6678	3.0101	
29	2009	3650.161	11807.150	10	10	0.3030303	0.3030303	0.1819899	0.3965664	0.2849508	0.2076046	1.2620	1.9008	0.4922	-2.9244	0.3759	1.6023	
30	2010	7613.413	17238.013	33	29	1	0.8787879	0.8787879	0.9627498	0.8451903	0.8259721	5.7462	55.4571	0.0138	-0.1146	0.9461	18.5642	
31	2011	2731.955	11272.745	6	7	0.1818182	0.2121212	0.0945194	0.1439725	0.2317337	0.0828652	1.0904	1.4141	1.1620	-9.9726	0.1994	1.2490	
32	2012	2032.395	7893.388	3	2	0.0909091	0.0606061	0.0186792	0.0331952	0.0344781	0.0044982	1.0045	1.0674	3.9062	-220.0856	0.0222	1.0228	
33	2013	2014.331	10998.466	2	5	0.0606061	0.1515152	0.0280991	0.0315787	0.2065380	0.0187081	1.0191	1.2811	2.5124	-51.2513	0.0677	1.0727	

**Tabla 7.26.** Análisis de Frecuencia Bivariado de Avenidas de la Estación Huites; Cuenca del Río Fuerte

ORDEN	DATOS OBSERVADOS			RANKING		DATOS TRANSFORMADOS		CÓPULA	MARGINALES UNIFORMES			CÓPULA	PERIODOS DE RETORNO BIVARIADOS OBTENIDOS CON CÓPULA JOE/B5				
	ESTACIÓN: HUITES			CLASIFICACIÓN		Y <sub>ij</sub> ∈ [ 0, 1 ]		EMPÍRICA	DOBLE GUMBEL	DOBLE GUMBEL	JOE/B5	"OR"	"AND"	GENERADOR	DERIVADA	FUN. KENDALL	"KENDALL"
	i	Año	Q m <sup>3</sup> /s	V Hm <sup>3</sup>	R <sub>Qi</sub>	R <sub>Vi</sub>	R <sub>Qi</sub> /N	R <sub>Vi</sub> /N	C(R <sub>Qi</sub> /N , R <sub>Vi</sub> /N) (Probabilidad)	F(Q)=u <sub>1</sub> (Probabilidad)	F(V)=u <sub>2</sub> (Probabilidad)	C(u <sub>1</sub> ,u <sub>2</sub> ) (Probabilidad)	T <sup>V</sup> (u <sub>1</sub> ,u <sub>2</sub> ) (años)	T <sup>A</sup> (u <sub>1</sub> ,u <sub>2</sub> ) (años)	φ (t)	φ' (t)	Kc (t) (Probabilidad)
1	1941	2085	458	26	13	0.5000000	0.2500000	0.1834821	0.4766767	0.2252070	0.1604512	1.1911	2.1807	1.1861	-5.6449	0.3706	1.5888
2	1942	2531	1302	32	44	0.6153846	0.8461538	0.5925053	0.5965621	0.8099245	0.5637451	2.2922	6.3590	0.1955	-1.0311	0.7533	4.0538
3	1943	14376	1928	51	49	0.9807692	0.9423077	0.9395789	0.9720172	0.8993220	0.8960542	9.6204	40.4612	0.0090	-0.1808	0.9457	18.4217
4	1944	2580	871	34	32	0.6538462	0.6153846	0.5047851	0.6076940	0.6136203	0.4755744	1.9068	3.9330	0.3019	-1.4001	0.6912	3.2381
5	1945	1499	684	16	25	0.3076923	0.4807692	0.2165638	0.2804924	0.4544863	0.1903606	1.2351	2.1960	1.0329	-4.6563	0.4122	1.7012
6	1946	1165	720	9	27	0.1730769	0.5192308	0.1322529	0.1699549	0.4885981	0.1245898	1.1423	2.1458	1.4182	-7.4496	0.3150	1.4598
7	1947	1127	435	7	10	0.1346154	0.1923077	0.0459835	0.1584856	0.2030097	0.0563223	1.0597	1.4392	2.1735	-17.1979	0.1827	1.2235
8	1948	3215	344	38	5	0.7307692	0.0961538	0.0895739	0.7182352	0.1239579	0.1144809	1.1293	3.6726	1.4970	-8.1614	0.2979	1.4243
9	1949	10000	2966	49	52	0.9423077	1	0.9423077	0.9221247	0.9750768	0.9187416	12.3064	46.4251	0.0054	-0.1380	0.9576	23.6041
10	1950	3229	644	39	22	0.7500000	0.4230769	0.3909782	0.7200486	0.4150776	0.3758264	1.6021	4.1545	0.4692	-1.9989	0.6106	2.5679
11	1951	677	111	2	2	0.0384615	0.0384615	0.0029601	0.0527258	0.0144413	0.0015305	1.0015	1.0702	5.7488	-652.8257	0.0103	1.0104
12	1952	1266	474	10	14	0.1923077	0.2692308	0.0867813	0.2018136	0.2410192	0.0822817	1.0897	1.5638	1.8090	-11.5890	0.2384	1.3130
13	1953	1025	163	5	3	0.0961538	0.0576923	0.0106781	0.1293661	0.0269901	0.0066927	1.0067	1.1760	4.2762	-148.8727	0.0354	1.0367
14	1954	955	596	4	18	0.0769231	0.3461538	0.0443724	0.1109636	0.3662522	0.0663684	1.0711	1.6974	2.0149	-14.5075	0.2053	1.2583
15	1955	4780	787	42	30	0.8076923	0.5769231	0.5454623	0.8259082	0.5479190	0.5245505	2.1033	6.6347	0.2388	-1.1820	0.7266	3.6572
16	1956	696	513	3	15	0.0576923	0.2884615	0.0288393	0.0559079	0.2804991	0.0273165	1.0281	1.4474	2.8810	-36.0588	0.1072	1.1201
17	1957	593	69	1	1	0.0192308	0.0192308	0.0007549	0.0400638	0.0080381	0.0006538	1.0007	1.0498	6.5989	-1528.8740	0.0050	1.0050
18	1958	3010	740	37	28	0.7115385	0.5384615	0.4764202	0.6888763	0.5069050	0.4417163	1.7912	4.0661	0.3522	-1.5756	0.6652	2.9871
19	1959	1908	352	23	7	0.4423077	0.1346154	0.0924824	0.4205808	0.1302203	0.0862574	1.0944	1.8676	1.7640	-11.0277	0.2462	1.3266
20	1960	15000	1842	52	48	1	0.9230769	0.9230769	0.9762040	0.8906086	0.8884666	8.9659	46.1809	0.0104	-0.1955	0.9417	17.1572
21	1961	1396	689	13	26	0.2500000	0.5000000	0.1834821	0.2450829	0.4593072	0.1691085	1.2035	2.1518	1.1387	-5.3231	0.3830	1.6208
22	1962	1620	437	20	11	0.3846154	0.2115385	0.1281503	0.3225754	0.2049118	0.1075738	1.1205	1.7239	1.5553	-8.7243	0.2858	1.4003
23	1963	2702	885	36	33	0.6923077	0.6346154	0.5403492	0.6336748	0.6235904	0.4980530	1.9922	4.1530	0.2716	-1.2953	0.7077	3.4213
24	1964	1319	384	11	8	0.2115385	0.1538462	0.0568496	0.2192002	0.1567033	0.0597208	1.0635	1.4624	2.1168	-16.1865	0.1905	1.2353
25	1965	1944	305	24	4	0.4615385	0.0769231	0.0550681	0.4323163	0.0956755	0.0651382	1.0697	1.8617	2.0330	-14.7924	0.2026	1.2540
26	1966	2420	2716	29	50	0.5576923	0.9615385	0.5566244	0.5698406	0.9630239	0.5688139	2.3192	27.8169	0.1903	-1.0129	0.7567	4.1099
27	1967	2506	593	31	17	0.5961538	0.3269231	0.2672114	0.3269231	0.2672114	0.2932007	1.4148	2.9471	0.6641	-2.7793	0.5321	2.1374
28	1968	1534	1706	18	46	0.3461538	0.8846154	0.3412094	0.2926438	0.8759222	0.2880197	1.4045	8.3714	0.6787	-2.8425	0.5268	2.1132
29	1969	1508	837	17	31	0.3269231	0.5961538	0.2672114	0.2836134	0.5882361	0.2313629	1.3010	2.7815	0.8626	-3.7120	0.4637	1.8648
30	1970	1558	1001	19	37	0.3653846	0.7115385	0.3303637	0.3009929	0.6956336	0.2705181	1.3708	3.6511	0.7304	-3.0732	0.5082	2.0333
31	1971	2200	905	27	34	0.5192308	0.6538462	0.4360046	0.5107158	0.6373474	0.4220308	1.7302	3.6501	0.3843	-1.6893	0.6495	2.8532
32	1972	2225	442	28	12	0.5384615	0.2307692	0.1790315	0.5178395	0.2096917	0.1589581	1.1890	2.3179	1.1946	-5.7039	0.3684	1.5833
33	1973	7960	1250	47	42	0.9038462	0.8076923	0.7877453	0.8880854	0.7960307	0.7707314	4.3617	11.5453	0.0476	-0.4429	0.8782	8.2078
34	1974	3790	607	41	21	0.7884615	0.4038462	0.3824722	0.7762506	0.3775498	0.3557265	1.5521	4.9523	0.5110	-2.1570	0.5926	2.4547
35	1975	1095	1768	6	47	0.1153846	0.9038462	0.1144460	0.1490770	0.8827965	0.1472045	1.1726	8.6707	1.2645	-6.2098	0.3508	1.5405
36	1976	2677	565	35	16	0.6730769	0.3076923	0.2714781	0.6285498	0.3342165	0.2822794	1.3933	3.1298	0.6952	-2.9151	0.5208	2.0866
37	1977	1135	601	8	19	0.1538462	0.3653846	0.0908426	0.1608742	0.3713937	0.0960615	1.1063	1.7737	1.6619	-9.8417	0.2649	1.3604
38	1978	4790	1245	43	41	0.8269231	0.7884615	0.7323731	0.8262290	0.7945911	0.7365457	3.7957	8.6411	0.0641	-0.5235	0.8590	7.0909
39	1979	6860	986	46	35	0.8846154	0.6730769	0.6573196	0.8693052	0.6873460	0.6659305	2.9934	9.1509	0.1074	-0.7071	0.8178	5.4886
40	1980	1496	1076	15	38	0.2884615	0.7307692	0.2659548	0.2794527	0.7328441	0.2581446	1.3480	4.0676	0.7695	-3.2546	0.4946	1.9786
41	1981	4828	992	44	36	0.8461538	0.6923077	0.6620737	0.8274271	0.6906963	0.6530029	2.8819	7.4140	0.1168	-0.7437	0.8100	5.2632
42	1982	2450	351	30	6	0.5769231	0.1153846	0.0950562	0.1294294	0.5772694	0.1064951	1.1192	2.5013	1.5648	-8.8187	0.2839	1.3965
43	1983	8275	1625	48	45	0.9230769	0.8653846	0.8469377	0.8934944	0.8662609	0.8317777	5.9445	13.8846	0.0247	-0.3095	0.9115	11.3018
44	1984	5580	1258	45	43	0.8653846	0.8269231	0.7846297	0.8461951	0.7982946	0.7512202	4.0196	9.3694	0.0567	-0.4883	0.8673	7.5337
45	1985	3585	1092	40	39	0.7692308	0.7500000	0.6685894	0.7590559	0.7399217	0.6550466	2.8989	6.4074	0.1152	-0.7378	0.8112	5.2977
46	1986	1349	1185	12	40	0.2307692	0.7692308	0.2182289	0.2292066	0.7757355	0.2174794	1.2779	4.7051	0.9160	-3.9925	0.4469	1.8081
47	1987	1429	766	14	29	0.2692308	0.5576923	0.2125115	0.2563510	0.5299539	0.1958336	1.2435	2.4418	1.0078	-4.5077	0.4194	1.7224
48	1988	1866	683	21	24	0.4038462	0.4615385	0.2680451	0.4067053	0.4535191	0.2658600	1.3621	2.4653	0.7449	-3.1396	0.5031	2.0125
49	1989	1868	428	22	9	0.4230769	0.1730769	0.1140732	0.4073703	0.1963982	0.1250202	1.1429	1.9185	1.4150	-7.4219	0.3157	1.4613
50	1990	11558	2930	50	51	0.9615385	0.9807692	0.9574276	0.9442821	0.9735809	0.9389275	16.3740	47.4732	0.0030	-0.1010	0.9682	31.4406
51	1991	2563	653	33	23	0.6346154	0.4423077	0.3694371	0.6038777	0.4240647	0.3435953	1.5235	3.1680	0.5378	-2.2608	0.5815	2.3893
52	1992	2025	601	25	19	0.4807692	0.3653846	0.2533871	0.4581311	0.3713937	0.2474547	1.3288	2.3927	0.8052	-3.4256	0.4825	1.9324

### **7.2.7. Generación de curvas de periodo de retorno conjunto (Q, V; T)**

Los conjuntos de probabilidades que conforman la muestra sintética extendida  $(u_1, u_2) \in (0,1)$  de cada estación, deben ser transformados a cuantiles  $q_i = F^{-1}(u_1)$  y  $v_i = F^{-1}(u_2)$ , a través de las funciones inversas de sus marginales, con ello se logra conseguir que los conjuntos  $(q, v)$  sigan manteniendo sus unidades originales de cada variable, Gasto pico ( $m^3/s$ ) y Volumen ( $Hm^3$ ).

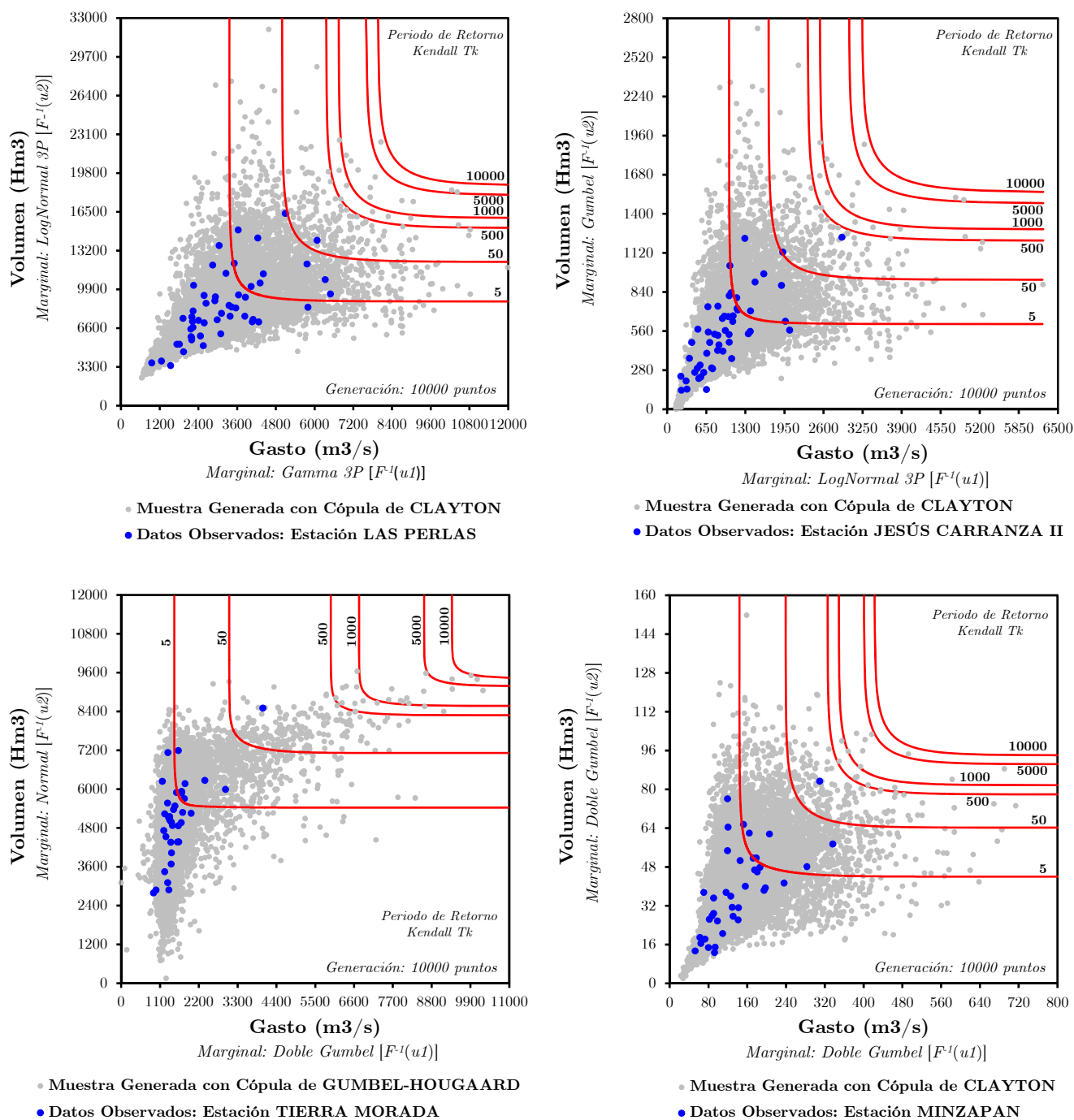
De las seis FDP que se implementaron, únicamente la FDP de Gumbel tiene forma analítica para obtener su inversa, las funciones restantes su inversa se obtiene de forma numérica. Por otra parte, se generaron muestras sintéticas vinculadas a distintos periodos de retorno de diseño (10,000, 5,000, 1,000, 500, 50 y 5 años) para cada una de las estaciones y bajo el enfoque de funciones cópula que previamente fueron seleccionadas como las más apropiadas.

Los valores de la variable de gastos pico Q de la muestra extendida se fijan, y por medio del periodo de retorno de "Kendall" basado en cópulas (ec. 4.27), se calculan los volúmenes asociados. En la Figuras 7.27, se presentan las muestras extendidas (10,000 pares) en unidades originales de las variables Q y V, para las estaciones de las cuencas del Río Coatzacoalcos y del Río Fuerte, respectivamente.

La generación de muestras sintéticas asociando al gasto pico y volumen de la avenida de un determinado periodo de retorno, conforman las curvas de período de retorno bivariado (Q, V; T). Los conjuntos  $(q, v)$  ubicados en una curva de nivel de la cópula  $C(u_1, u_2)$  están asociados con un valor de período de retorno. Una curva de nivel estará conformada por una infinidad de soluciones (Q, V; T). En la Figura 7.27 se muestran las curvas de periodo de retorno conjunto calculadas a través del mejor modelo cópula que representa a la muestra de datos observados de cada estación.

En la Figura 7.28, se representa el comportamiento de cada cópula en el espacio, asociada a los datos observados de cada estación para las cuencas del Río Coatzacoalcos y del Río Fuerte, respectivamente. En el eje de las abscisas se mide los gastos, en el eje de las ordenadas se mide los volúmenes y en el eje de las cotas se mide las probabilidades, o sea la cópula  $C(u_1, u_2)$ .





**Figura 7.27.** Comparación de las muestras de datos observados con la generación de muestras sintéticas extendidas a 10,000 puntos y obtenidas en unidades originales; Curvas teóricas de período de retorno conjunto obtenidas a través del enfoque de Kendall (líneas rojas), para las estaciones de las cuencas del Río Coatzacoalcos y del Río Fuerte *continúa...*

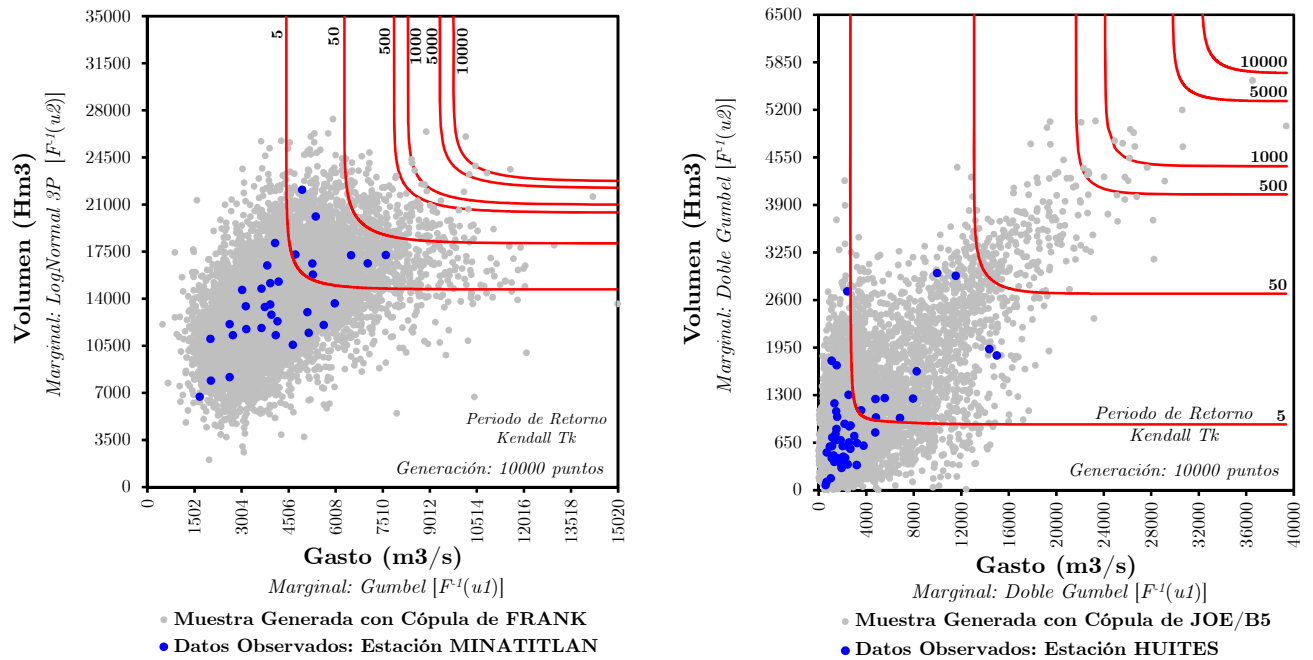


Figura 7.27. Continuación

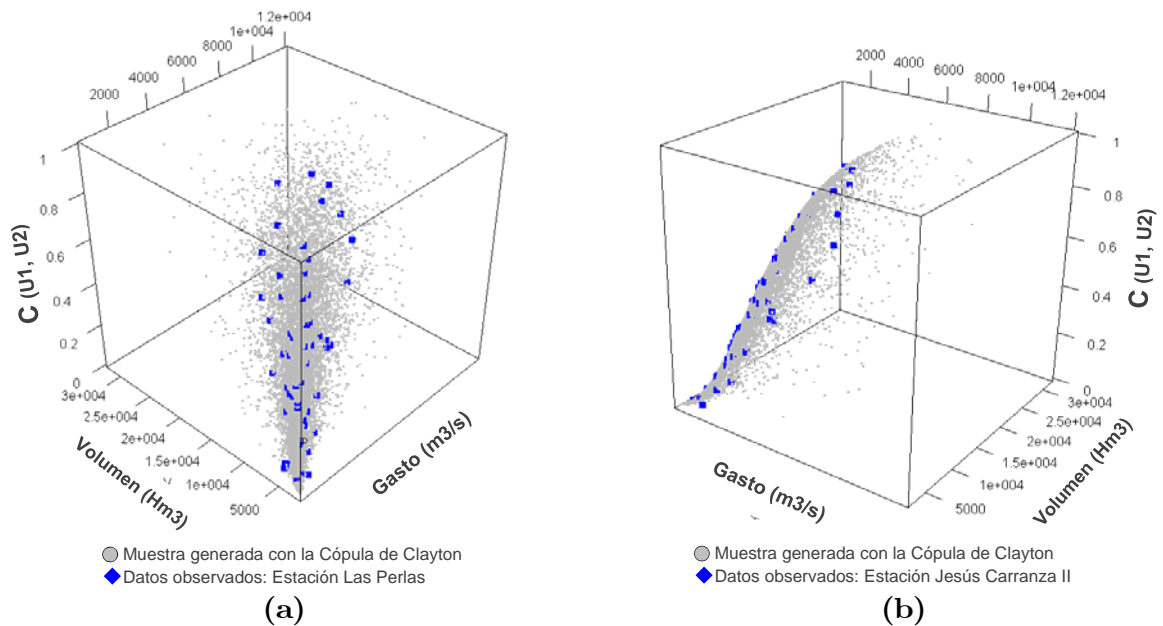


Figura 7.28. Representación de las Funciones Cópula en el espacio, asociadas a los datos observados de cada estación para las cuencas del Río Coatzacoalcos (a, b, c, d, e) y del Río Fuerte (f), respectivamente.

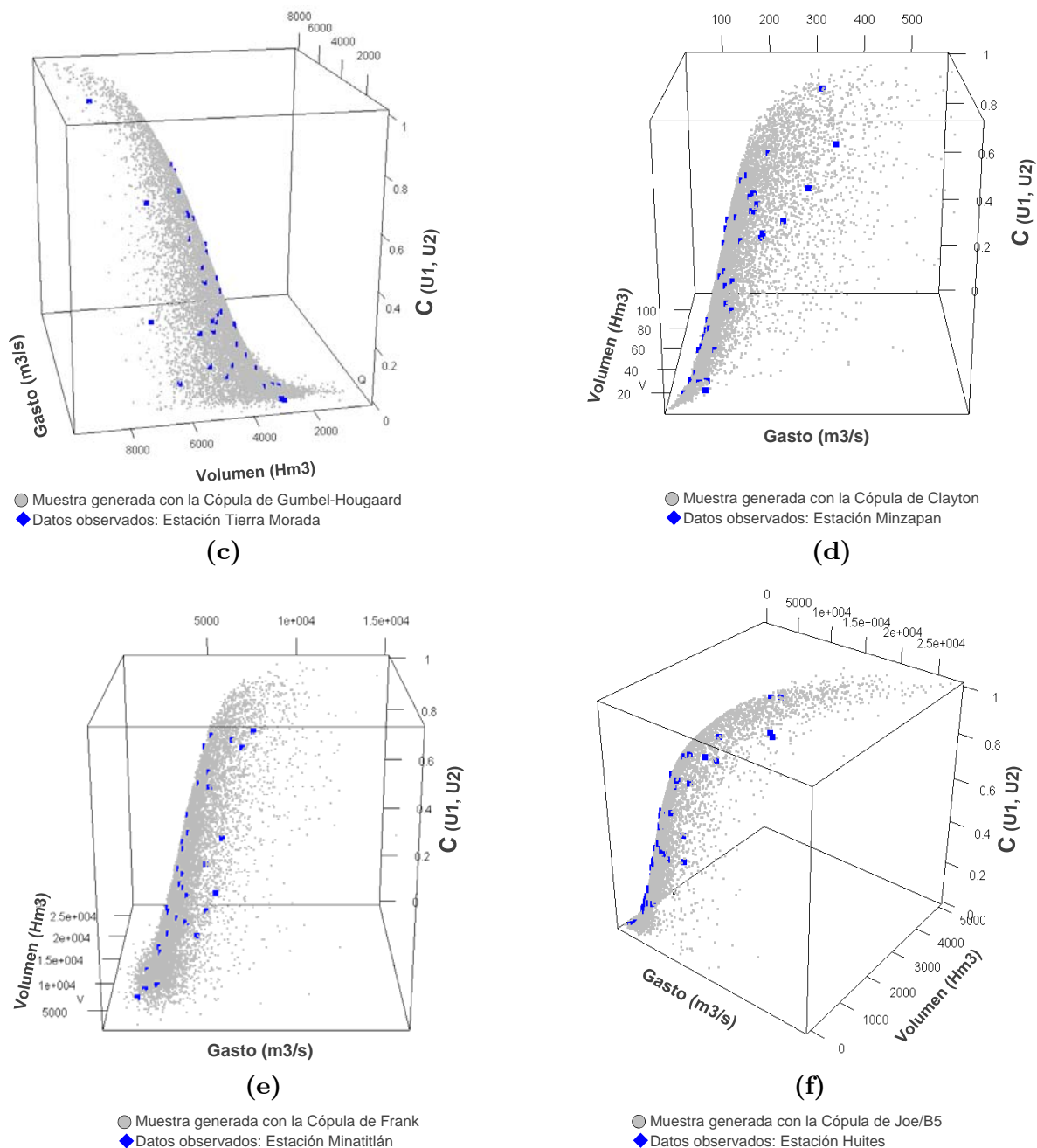


Figura 7.28. Continuación

### 7.2.8. Generación de hidrogramas sintéticos

Cada pareja  $(Q, V; T_{\text{CONJUNTO}})$  representa un hidrograma de avenida con determinado periodo de retorno, de tal forma que se pasa de un conjunto de parejas  $(Q, V; T_{\text{CONJUNTO}})$  a un hidrograma sintético. Se analizaron los registros históricos de gastos medios diarios (datos observados) para cada estación, y se seleccionó las avenidas máximas del registro histórico. De este

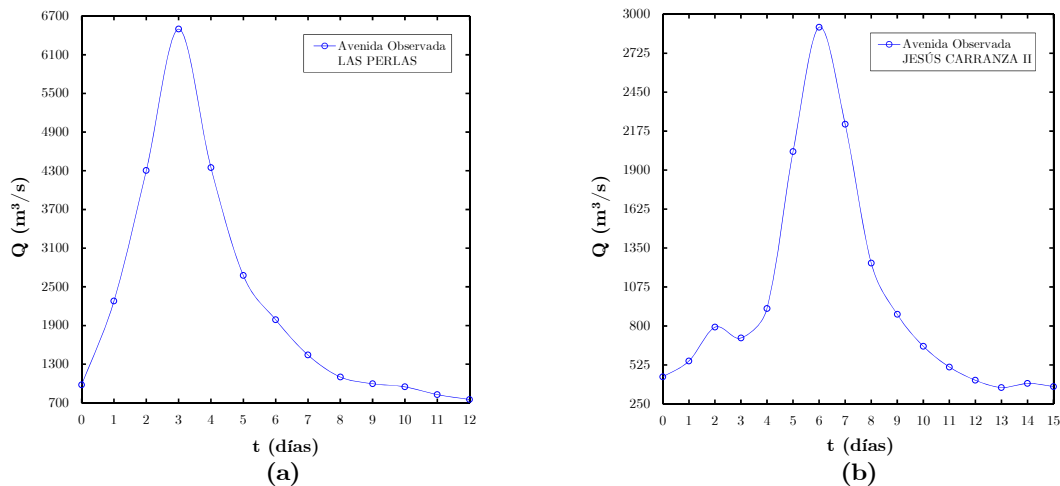
conjunto se analizó la forma que presentan los hidrogramas, con el objetivo de atribuir la similitud de forma que tendrán los hidrogramas sintéticos generados (figuras 7.29 y 7.30). Se optó por construir hidrogramas sintéticos mediante la Función Gamma de dos parámetros (2P). Se generaron 10,000 hidrogramas para cada estación, lo que corresponde al tamaño de la muestra extendida. En la Tabla 7.27, se muestran la relación de gastos pico de cada estación.

En las figuras 7.29 y 7.30, se muestran las avenidas máximas ocurridas en cada estación, correspondiente a las cuencas del Río Coatzacoalcos y Fuerte. Cabe señalar que, los gastos pico del primer y segundo caso se obtuvieron de los gastos máximos documentados en los registros de gastos medios diarios.

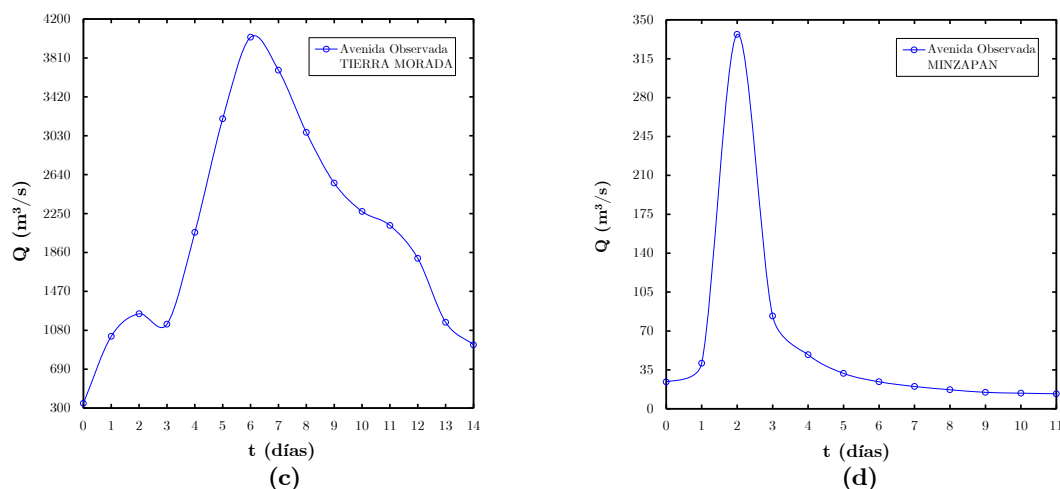
Para construir los hidrogramas con la forma de la Función Gamma 2P es necesario pero no suficiente, calcular el tiempo pico, para el caso de la estación Huites se ha estimado este parámetro por medio de las características físicas de la cuenca de la presa Huites [Mendoza *et al.* (2014)].

**Tabla 7.27.** Relación de gastos pico de la cuenca del Río Coatzacoalcos

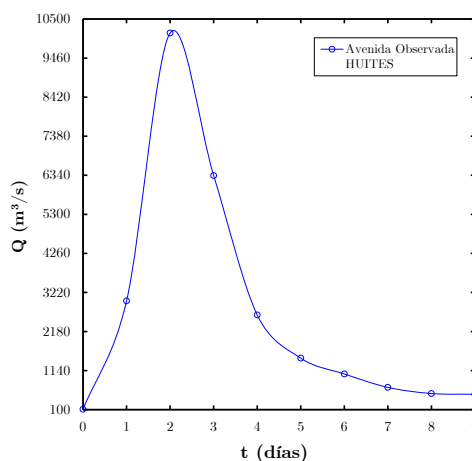
Estación	Duración Avenida	Fecha Qp	Qp (m <sup>3</sup> /s)
Las Perlas	21/09/1963 - 03/10/1963	24/09/1963	6497.666
Jesús Carranza II	26/08/1973 - 13/09/1973	01/09/1973	2908.776
Tierra Morada	27/09/1999 - 11/10/1999	03/10/1999	4017.361
Minzapan	12/09/1976 - 23/09/1976	14/09/1976	336.925



**Figura 7.29.** Avenidas históricas registradas en la cuenca del Río Coatzacoalcos, se nota una importante similitud de forma con de la Función Gamma 2P (Método 3) en (a) y (d); para el caso de (b) y (c) existe una notoria similitud con la Función Gamma 2P (Método 2).



**Figura 7.29.** *Continuación*



**Figura 7.30.** Avenida histórica registrada en el periodo del 27/12/1990 al 05/01/1991 en la estación Huites de la cuenca del Río Fuerte, con un  $Q_{\text{máx}}=10,129.264$  (m<sup>3</sup>/s) [29/12/1990], se nota una importante similitud de forma con la Función Gamma 2P (Método 1 y 3).

▪ **Cálculo del tiempo pico ( $t_p$ )**

Se empleó la cartografía de la zona de Huites a escala 1:250,000, obteniendo una longitud del cauce de 330 km (desde el origen del Río Fuerte hasta la presa Huites) y una superficie de 26,020 km<sup>2</sup>, con un desnivel entre los extremos  $D = 2,950$  m [CNA (1997)].

Se aplicó el método de Taylor-Schwarz [Aparicio (2000)] para tramos equidistantes (ec. 7.1), la pendiente media del cauce principal es  $S = 0,00408$  (0,408%).

$$S = \left[ \frac{m}{1/S_1 + 1/S_2 + \dots + 1/S_m} \right]^2 \quad (7.1)$$

donde

S, Pendiente media del cauce principal, adimensional;

m, Número de tramos en que se dividió el cauce;

S1, S2, Sm, Pendiente de cada tramo, adimensional.

El tiempo de concentración  $t_c$  (ec. 7.2), se calculó con la ecuación del *Soil Conservation Service* (SCS), [SARH (1987)], el cual resultó de  $t_c = 33.20125$  (h).

$$t_{c_{SCS}} = \frac{L^{1.15}}{3085(D^{0.385})} \quad (7.2)$$

donde

L, Longitud del cauce principal, en m;

D, Desnivel entre los extremos del cauce, en m.

El tiempo de retraso  $tr$  (ec. 7.3), se calculó con la ecuación de Víctor Mockus (1957) y citado en [Aparicio (2000)], se obtuvo un valor de  $tr = 19.92075$  (h).

$$tr = 0.6 t_c \quad (7.3)$$

La duración en exceso de la lluvia de diseño ( $de$ ) [Chow (1964)], considera la siguiente expresión para cuencas grandes,

$$de = 2 \sqrt{t_c}, \quad (7.4)$$

se obtuvo un valor de  $de = 11.52410$  (h).

Finalmente el tiempo pico  $tp$  (ec. 7.5) se calcula de la siguiente manera,

$$tp = \frac{de}{2} + tr, \quad (7.5)$$

se obtuvo un tiempo pico  $tp = 25.68280$  (h).

El tiempo pico calculado, es el que se utilizará para generar los hidrogramas sintéticos correspondientes a la estación Huites. Otras alternativas de obtener  $tp$  es, utilizar la parametrización de hidrogramas

---

con base en los Método 1 y 2 del Capítulo 5, Otra forma de obtener  $tp$ , es aproximando los tiempos pico de avenidas históricas más importantes.

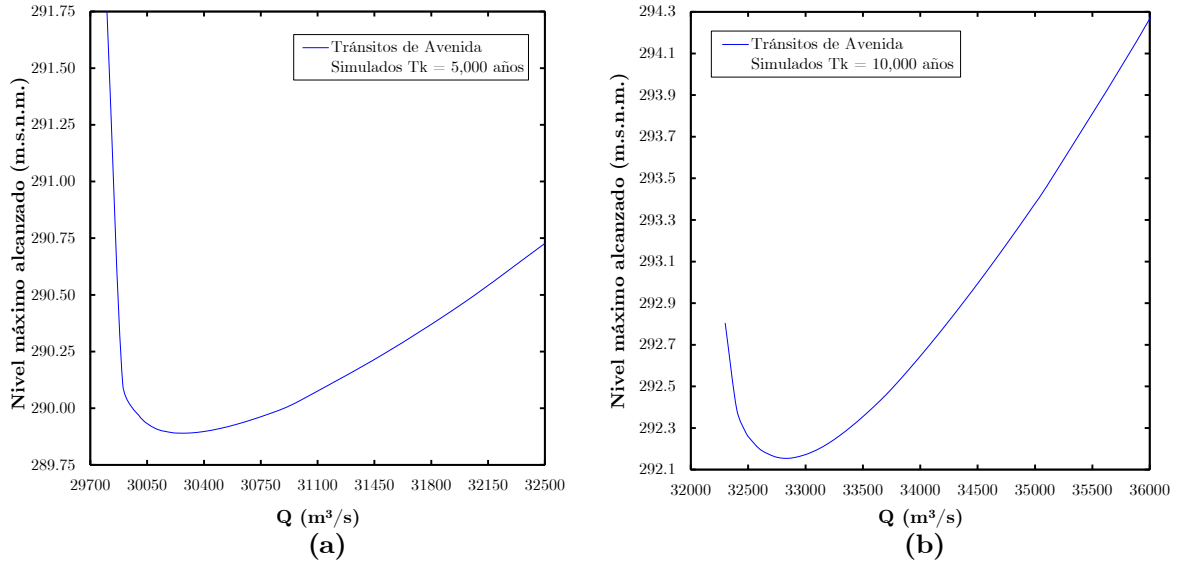
Por medio de los hidrogramas sintéticos generados (10,000), es posible obtener avenidas de diseño para diferentes escenarios, en un principio se señaló que se obtendrán avenidas de diseño únicamente para la estación Huites, motivo por el cual se sitúa la presa Luis Donaldo Colosio "Huites". Las avenidas de diseño se obtienen para diferentes periodos de retorno, de modo que, se ha comentado que las curvas de periodo de retorno bajo un enfoque bivariado, proporcionan infinitas soluciones de parejas  $(Q, V)$ , bajo este contexto, entonces se tendría infinitas avenidas de diseño. *Para averiguar cuál es la avenida más desfavorable de todo el conjunto, se tiene que simular tránsitos de avenida en función de las características físicas del vaso y vertedor de la presa Huites.*

Las estaciones de la Cuenca del Río Coatzacoalcos sirvieron exclusivamente para ejemplificar la metodología de selección del mejor modelo cópula y la generación de curvas de periodo de retorno conjunto, por lo que no se generarán avenidas de diseño.

### **7.2.9. Simulación de Tránsitos de avenidas de los hidrogramas sintéticos de la presa Huites**

La simulación de los Tránsitos de Avenida, toma en cuenta las características del vaso de almacenamiento y la obra de excedencias de la presa Huites; se utilizó la curva E-V-Qs del vaso de Huites, elaborada por la Comisión Federal de Electricidad CFE y proporcionada por el Instituto de Ingeniería UNAM, Tabla 7.5; la elevación y longitud de la cresta del vertedor son de 258 m.s.n.m. y 62 m respectivamente; las elevaciones al NAMO es de 270 m.s.n.m., NAME 290 m.s.n.m., corona de la cortina 290.50 m.s.n.m., y al parapeto de la cortina 292 m.s.n.m.; El tiempo pico calculado es  $tp = 25.68280$  (h), la forma de los hidrogramas sintéticos son del tipo Función Gamma 2P (Métodos 1 y 3, cap. 5). Los hidrogramas con diferentes combinaciones  $(Q_p, V)$  pueden alcanzar el mismo nivel máximo de agua en la presa, lo que implica un riesgo similar de sobrevertido en la presa; en consecuencia, se puede considerar que tienen el mismo período de retorno. Las simulaciones de los tránsitos de avenida se ejecutaron con el programa **TRATE.bas**, proporcionado por IIUNAM.

En la Figura 7.31 (a) y (b), se muestra un conjunto de los 10,000 hidrogramas que fueron transitados, cuyas parejas de puntos ( $Q_p$ ,  $V$ ) están asociados a periodos de retorno  $T_k = 5,000$  y  $T_k = 10,000$  años.



**Figura 7.31.** Elevaciones máximas presentadas en el embalse de la presa Huites para un determinado periodo de retorno Kendall y cópula de Joe/B5, obtenidas por medio de la simulación del Tránsito de Avenidas a los hidrogramas sintéticos generados. Estación Huites.

Para la elección de la avenida de diseño más desfavorable asociada con distintos periodos de retorno conjuntos, basta con identificar el conjunto de valores ( $Q_p$ ,  $V$ ), que hagan máximo el nivel alcanzado en el embalse de la presa Huites, de acuerdo a las curvas de la Figura 7.31 (a) y (b).

Se hace la observación de que, matemáticamente por medio de la Figura 7.31, la elevación óptima es el punto mínimo de ambas curvas, mientras que desde el criterio físico, la avenida de diseño más desfavorable es aquella que posee el nivel máximo sobre el embalse de la presa.

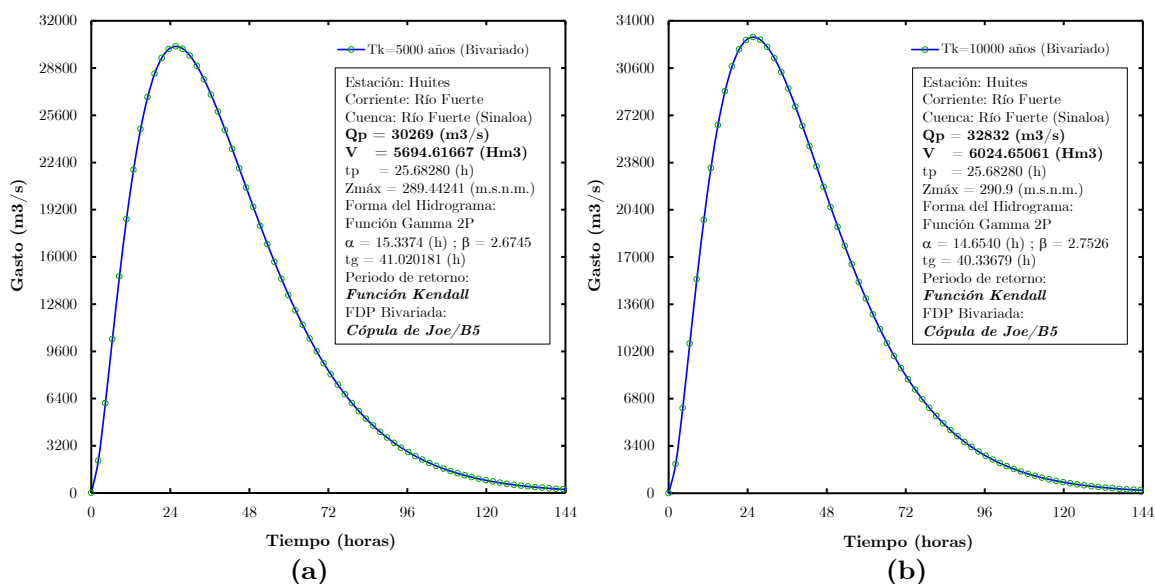
Lo anterior tiene cierto sentido geométrico; en primer lugar, se utilizó las ecuaciones de periodo de retorno bivariado por medio de la función Kendall y al cópula de Joe/B5, las curvas que gobiernan este comportamiento son de tipo convexas, por lo tanto las curvas de la Figura 7.31 también lo son, debido a que los volúmenes de las muestras sintéticas fueron generadas con la función Kendall y por ende los hidrogramas.



### 7.2.10. Avenidas de Diseño más desfavorables en el embalse de la presa Huites.

La construcción de la avenida de diseño más desfavorable está caracterizada por tres parámetros de gran importancia ( $Q_p$ ,  $V$ ,  $t_p$ ), otros dos parámetros de forma indirecta, como el periodo de retorno conjunto y el nivel más crítico en la superficie libre del embalse ( $T_k$ ,  $Z_{m\acute{a}x}$ ). Para  $T_k=10,000$  años, en la Tabla 7.28a se muestra el hidrograma de entrada a la presa Huites con base en la Función Gamma 2P y  $t_p = 25.68$  (h), en la Tabla 7.28b se muestran los resultados de la simulación del tránsito de la avenida a través del embalse de la presa Huites y en la Figura 7.33 se presentan sus correspondientes gráficas. En la Tabla 7.28c, se muestra la relación de las avenidas de diseño más desfavorables y asociadas a distintos periodos de retorno respectivamente.

La avenida de diseño más desfavorable para un  $T_k = 5,000$  años, resultó un  $Q_p = 30,269$  ( $m^3/s$ ) y  $V = 5,695$  ( $Hm^3$ ), condiciones que alcanzan una  $Z_{m\acute{a}x} = 289.4$  (m.s.n.m). Para el escenario con un  $T_k = 10,000$  años, resultó el conjunto  $Q_p = 32,832$  ( $m^3/s$ ) y  $V = 6,025$  ( $Hm^3$ ) como las condiciones que alcanzan una  $Z_{m\acute{a}x} = 290.9$  (m.s.n.m), Figura 7.32.



**Figura 7.32.** Avenidas de diseño con periodos de retorno  $T_k = 5,000$  años y  $T_k = 10,000$  años respectivamente, para la presa Huites; la forma de los hidrogramas se hicieron con base en la Función Gamma 2P, con  $t_p = 25.68$  (h) para ambos casos,  $t_g = 41.02$  (h) y  $t_g = 40.34$  (h) respectivamente.

**Tabla 7.28a.** Hidrograma de entrada al embalse de la presa Huites, se construyó con base en la forma de la Función Gamma de dos parámetros.

i	t (h)	Q (m3/s)	i	t (h)	Q (m3/s)	i	t (h)	Q (m3/s)	i	t (h)	Q (m3/s)	i	t (h)	Q (m3/s)
1	0	0	32	66.34723	10802.77162	63	132.69447	393.38527	94	199.04170	8.65230	125	265.38893	0.15481
2	2.14023	2101.98634	33	68.48747	9868.96248	64	134.83470	349.59686	95	201.18193	7.61804	126	267.52917	0.13567
3	4.28047	6120.54811	34	70.62770	9000.44798	65	136.97493	310.54543	96	203.32217	6.70607	127	269.66940	0.11888
4	6.42070	10764.18721	35	72.76793	8195.16096	66	139.11517	275.73819	97	205.46240	5.90214	128	271.80963	0.10416
5	8.56093	15400.02734	36	74.90817	7450.61513	67	141.25540	244.73074	98	207.60263	5.19359	129	273.94987	0.09125
6	10.70117	19676.06193	37	77.04840	6764.02448	68	143.39563	217.12277	99	209.74287	4.56925	130	276.09010	0.07993
7	12.84140	23403.62697	38	79.18863	6132.40285	69	145.53587	192.55404	100	211.88310	4.01924	131	278.23033	0.07001
8	14.98163	26496.33842	39	81.32887	5552.64603	70	147.67610	170.70069	101	214.02333	3.53479	132	280.37057	0.06132
9	17.12187	28933.04655	40	83.46910	5021.59873	71	149.81633	151.27182	102	216.16357	3.10820	133	282.51080	0.05370
10	19.26210	30733.75264	41	85.60933	4536.10846	72	151.95657	134.00636	103	218.30380	2.73261	134	284.65103	0.04702
11	21.40233	31943.43412	42	87.74957	4093.06825	73	154.09680	118.67024	104	220.44403	2.40201	135	286.79127	0.04117
12	23.54257	32621.03922	43	89.88980	3689.44984	74	156.23703	105.05370	105	222.58427	2.11106	136	288.93150	0.03604
13	<b>25.68280</b>	<b>32832</b>	44	92.03003	3322.32888	75	158.37727	92.96898	106	224.72450	1.85504	137	291.07173	0.03155
14	27.82303	32643.19625	45	94.17027	2988.90331	76	160.51750	82.24808	107	226.86473	1.62982	138	293.21197	0.02761
15	29.96327	32119.64843	46	96.31050	2686.50623	77	162.65773	72.74081	108	229.00497	1.43172	139	295.35220	0.02417
16	32.10350	31322.43567	47	98.45073	2412.61413	78	164.79797	64.31299	109	231.14520	1.25750	140	297.49243	0.02115
17	34.24373	30307.48059	48	100.59097	2164.85138	79	166.93820	56.84482	110	233.28543	1.10432	141	299.63267	0.01851
18	36.38397	29124.94252	49	102.73120	1940.99162	80	169.07843	50.22940	111	235.42567	0.96965	142	301.77290	0.01619
19	38.52420	27819.03177	50	104.87143	1738.95683	81	171.21867	44.37139	112	237.56590	0.85129	143	303.91313	0.01417
20	40.66443	26428.10841	51	107.01167	1556.81447	82	173.35890	39.18585	113	239.70613	0.74726	144	306.05337	0.01239
21	42.80467	24984.96651	52	109.15190	1392.77319	83	175.49913	34.59707	114	241.84637	0.65586	145	308.19360	0.01084
22	44.94490	23517.23200	53	111.29213	1245.17744	84	177.63937	30.53770	115	243.98660	0.57555	146	310.33383	0.00948
23	47.08513	22047.82278	54	113.43237	1112.50135	85	179.77960	26.94776	116	246.12683	0.50502	147	312.47407	0.00829
24	49.22537	20595.43468	55	115.57260	993.34207	86	181.91983	23.77395	117	248.26707	0.44306	148	314.61430	0.00725
25	51.36560	19175.02793	56	117.71283	886.41285	87	184.06007	20.96884	118	250.40730	0.38866	149	316.75453	0.00634
26	53.50583	17798.29729	57	119.85307	790.53590	88	186.20030	18.49034	119	252.54753	0.34089	150	318.89477	0.00554
27	55.64607	16474.11492	58	121.99330	704.63530	89	188.34053	16.30101	120	254.68777	0.29896	151	321.03500	0.00485
28	57.78630	15208.93938	59	124.13353	627.73001	90	190.48077	14.36766	121	256.82800	0.26215	152	323.17523	0.00424
29	59.92653	14007.18766	60	126.27377	558.92699	91	192.62100	12.66080	122	258.96823	0.22985	153	325.31547	0.00370
30	62.06677	12871.56902	61	128.41400	497.41467	92	194.76123	11.15431	123	261.10847	0.20150	154	327.45570	0.00324
31	64.20700	11803.38136	62	130.55423	442.45655	93	196.90147	9.82499	124	263.24870	0.17663	155	329.59593	0.00283

Para la realización del tránsito de avenida se consideraron las siguientes características y combinación de  $Q_p$  y  $V$ , para un  $T_k=10000$  años. En la Tabla 7.28b, se muestran los resultados de la simulación del tránsito de la avenida y en la Figura 7.33 se muestran las gráficas.

Número de gastos de entrada	155
Número de puntos de la curva elevaciones	221
Intervalo de tiempo	2.13 h
Gasto de salida por la obra de toma	0 m3/s
Factor para el intervalo de cálculo	1
Tipo de restricción en el gasto de salida	0
Sin restricciones en la descarga	0
Condiciones iniciales	
Volumen (NAMO)	2409.15 Hm3
Elevación inicial (NAMO)	270 m
Gasto pico	32832 m3/s
Volumen asociado al pico	6024.65 Hm3
Tiempo pico	25.68280 h

Tabla 7.28b. Resultados del Tránsito de Avenida en la presa Huites

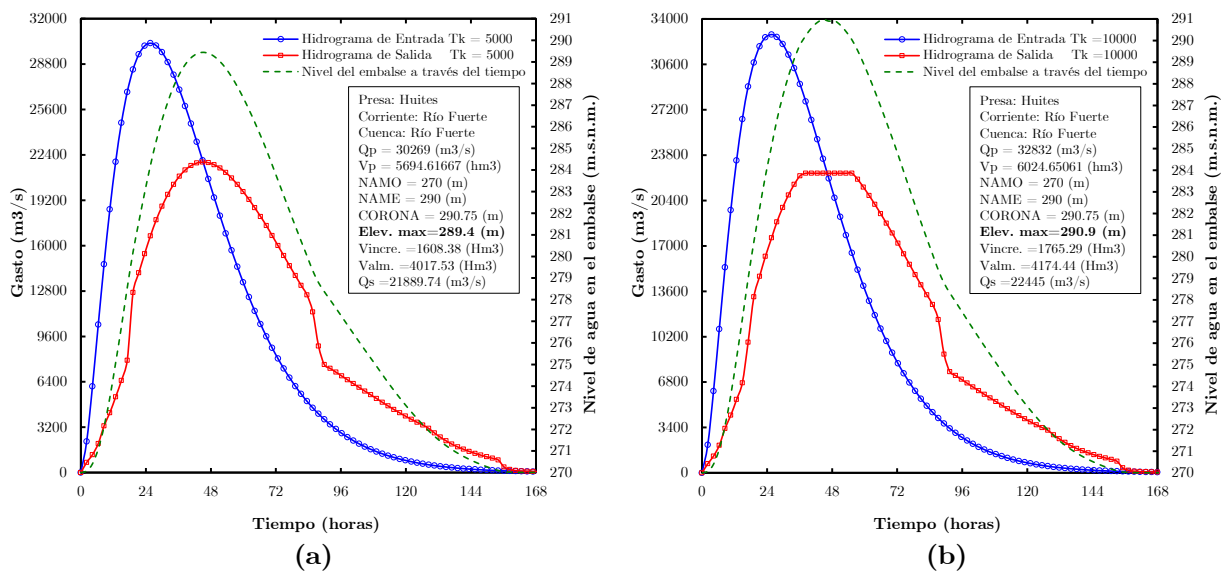
i	t	Volumen almacenado	Elevación	Q entrada	Q salida	i	t	Volumen almacenado	Elevación	Q entrada	Q salida
	hr	hm3	m	m3/s	m3/s		hr	hm3	m	m3/s	m3/s
0	0	2409.14990	270.00000	0.00000	0.00000	68	144.59	2445.56982	270.49408	192.55405	1359.97229
1	2.13	2414.58154	270.07382	2101.98633	682.81201	69	146.71	2437.04199	270.37839	170.70068	1231.43445
2	4.25	2438.63354	270.39999	6120.54834	1255.43994	70	148.84	2429.29517	270.27328	151.27182	1114.65320
3	6.38	2490.57324	271.10464	10764.18750	2058.41406	71	150.97	2422.25806	270.17786	134.00636	1009.24445
4	8.51	2570.09863	272.18356	15400.02734	3327.39478	72	153.09	2416.05005	270.09357	118.67024	865.44098
5	10.63	2675.10107	273.60809	19676.06250	4313.61475	73	155.22	2412.15405	270.04068	105.05370	376.26639
6	12.76	2802.45093	275.31024	23403.62695	5492.04150	74	157.34	2410.71899	270.02127	92.96898	196.74245
7	14.88	2946.60571	277.11731	26496.33789	6743.07959	75	159.47	2410.15479	270.01361	82.24808	125.89259
8	17.01	3095.50928	278.98389	28933.04688	9780.71680	76	161.60	2409.90405	270.01022	72.74081	94.56058
9	19.14	3235.97998	280.69061	30733.75195	13183.78223	77	163.72	2409.76978	270.00839	64.31299	77.62435
10	21.26	3369.05835	282.23755	31943.43359	14722.62109	78	165.85	2409.68140	270.00720	56.84482	66.61581
11	23.39	3497.77344	283.73380	32621.03906	16211.02637	79	167.98	2409.61377	270.00629	50.22940	58.14770
12	25.68	3618.87134	285.13162	32832.00000	17601.52539	80	170.10	2409.55786	270.00552	44.37139	51.09094
13	27.64	3730.15063	286.33490	32643.19531	18798.51367	81	172.23	2409.50903	270.00488	39.18585	45.16326
14	29.77	3830.01147	287.41473	32119.64844	19872.68555	82	174.35	2409.46631	270.00430	34.59707	39.80012
15	31.89	3917.11963	288.35669	31322.43555	20809.70898	83	176.48	2409.42920	270.00378	30.53770	35.00153
16	34.02	3990.67773	289.15210	30307.48047	21600.93359	84	178.61	2409.39648	270.00336	26.94776	31.04974
17	36.15	4050.34082	289.79724	29124.94336	22242.71484	85	180.73	2409.36694	270.00296	23.77395	27.38023
18	38.27	4097.24951	290.22354	27819.03125	22445	86	182.86	2409.34180	270.00259	20.96884	23.99298
19	40.40	4133.06201	290.58542	26428.10938	22445	87	184.99	2409.31982	270.00229	18.49034	21.17028
20	42.53	4158.02783	290.83597	24984.96680	22445	88	187.11	2409.29980	270.00204	16.30101	18.91212
21	44.65	4171.85303	290.97415	23517.23242	22445	89	189.24	2409.28198	270.00177	14.36766	16.37168
22	46.78	4174.43652	290.98000	22047.82227	22445	90	191.37	2409.26660	270.00159	12.66080	14.67806
23	48.90	4165.83740	290.91406	20595.43555	22445	91	193.49	2409.25317	270.00137	11.15431	12.70717
24	51.03	4146.24316	290.71790	19175.02734	22445	92	195.62	2409.24170	270.00122	9.82499	11.29081
25	53.16	4115.94385	290.41275	17798.29688	22445	93	197.74	2409.23022	270.00110	8.65230	10.16173
26	55.28	4075.30688	290.03888	16474.11523	22445	94	199.87	2409.22046	270.00095	7.61804	8.75038
27	57.41	4026.51514	289.53961	15208.93945	21986.42188	95	202.00	2409.21240	270.00082	6.70607	7.62130
28	59.54	3972.27002	288.95303	14007.18750	21402.92188	96	204.12	2409.20557	270.00073	5.90214	6.77449
29	61.66	3913.72266	288.31995	12871.56934	20773.16016	97	206.25	2409.19824	270.00067	5.19359	6.20995
30	63.79	3851.70410	287.64932	11803.38184	20106.03125	98	208.38	2409.19238	270.00055	4.56925	5.08087
31	65.91	3786.98511	286.94949	10802.77148	19409.88086	99	210.50	2409.18750	270.00052	4.01924	4.79860
32	68.04	3720.27368	286.22812	9868.96289	18692.27930	100	212.63	2409.18286	270.00043	3.53479	3.95179
33	70.17	3652.21240	285.49216	9000.44824	17960.17969	101	214.75	2409.17920	270.00040	3.10820	3.66951
34	72.29	3583.44995	284.72974	8195.16113	17201.74805	102	216.88	2409.17554	270.00034	2.73261	3.10497
35	74.42	3514.70068	283.93057	7450.61523	16406.78125	103	219.01	2409.17236	270.00031	2.40201	2.82270
36	76.55	3446.53394	283.13818	6764.02441	15618.53125	104	221.13	2409.16919	270.00027	2.11106	2.54043
37	78.67	3379.31372	282.35675	6132.40283	14841.20996	105	223.26	2409.16699	270.00021	1.85504	1.97589
38	80.80	3313.35132	281.59000	5552.64600	14078.45898	106	225.39	2409.16528	270.00021	1.62982	1.97589
39	82.92	3248.90942	280.84088	5021.59863	13333.28516	107	227.51	2409.16309	270.00018	1.43172	1.69362
40	85.05	3186.20361	280.11200	4536.10840	12608.20605	108	229.64	2409.16138	270.00015	1.25750	1.41135
41	87.18	3127.03711	279.37906	4093.06836	11479.98438	109	231.76	2409.16064	270.00012	1.10432	1.12908
42	89.30	3078.88232	278.77545	3689.44995	8884.46875	110	233.89	2409.15991	270.00012	0.96965	1.12908
43	91.43	3042.71484	278.32205	3322.32886	7577.14746	111	236.02	2409.15820	270.00012	0.85129	1.12908
44	93.56	3009.95776	277.91144	2988.90332	7292.87158	112	238.14	2409.15649	270.00009	0.74726	0.84681
45	95.68	2976.95166	277.49771	2686.50635	7006.44043	113	240.27	2409.15552	270.00009	0.65586	0.84681
46	97.81	2943.93262	277.08380	2412.61401	6719.88281	114	242.40	2409.15479	270.00006	0.57555	0.56454
47	99.94	2911.10425	276.67230	2164.85132	6434.99414	115	244.52	2409.15454	270.00006	0.50502	0.56454
48	102.06	2878.63965	276.26532	1940.99158	6153.24609	116	246.65	2409.15381	270.00006	0.44306	0.56454
49	104.19	2846.68457	275.86475	1738.95679	5875.92822	117	248.77	2409.15381	270.00003	0.38866	0.28227
50	106.31	2815.36108	275.47211	1556.81445	5604.09521	118	250.90	2409.15332	270.00006	0.34089	0.56454
51	108.44	2784.76904	275.08862	1392.77319	5338.60107	119	253.03	2409.15259	270.00003	0.29896	0.28227
52	110.57	2755.04956	274.69272	1245.17749	5064.50781	120	255.15	2409.15259	270.00003	0.26215	0.28227
53	112.69	2726.33838	274.30319	1112.50134	4794.85107	121	257.28	2409.15234	270.00003	0.22985	0.28227
54	114.82	2698.68921	273.92807	993.34204	4535.14502	122	259.41	2409.15186	270.00003	0.20150	0.28227
55	116.95	2672.12378	273.56769	886.41284	4285.64355	123	261.53	2409.15210	270.00000	0.17663	0.00000
56	119.07	2646.65259	273.22214	790.53589	4046.41333	124	263.66	2409.15234	270.00003	0.15481	0.28227
57	121.20	2622.27759	272.89142	704.63531	3817.46240	125	265.79	2409.15112	270.00003	0.13567	0.28227
58	123.32	2598.99292	272.57553	627.72998	3598.76270	126	267.91	2409.15112	270.00000	0.11888	0.00000
59	125.45	2576.78564	272.27426	558.92700	3390.19800	127	270.04	2409.15088	270.00003	0.10416	0.28227
60	127.58	2555.66577	271.98773	497.41467	3184.35913	128	272.16	2409.15063	270.00000	0.09125	0.00000
61	129.70	2536.17798	271.72336	442.45654	2847.27344	129	274.29	2409.15015	270.00003	0.07993	0.28227
62	131.83	2518.73682	271.48672	393.38528	2545.57520	130	276.42	2409.14966	270.00000	0.07001	0.00000
63	133.96	2503.12866	271.27496	349.59686	2275.58618	131	278.54	2409.15015	270.00000	0.06132	0.00000
64	136.08	2489.16113	271.08548	310.54544	2033.97998	132	280.67	2409.15063	270.00000	0.05370	0.00000
65	138.21	2476.62207	270.91537	275.73819	1828.52466	133	282.80	2409.14990	270.00003	0.04702	0.28227
66	140.33	2465.27393	270.76141	244.73074	1657.00977	134	284.92	2409.14917	270.00000	0.04117	0
67	142.46	2454.95313	270.62140	217.12277	1501.42810	135	287.05	2409.14941	270.00000	0.03604	0

**Tabla 7.28b.** Continua

i	t	Volumen almacenado	Elevación	Q entrada	Q salida	i	t	Volumen almacenado	Elevación	Q entrada	Q salida
	hr	hm3	m	m3/s	m3/s		hr	hm3	m	m3/s	m3/s
136	289.17	2409.14966	270.00000	0.03155	0	146	310.44	2409.15088	270.00000	0.00829	0
137	291.30	2409.14990	270.00000	0.02761	0	147	312.56	2409.15088	270.00000	0.00725	0
138	293.43	2409.15015	270.00000	0.02417	0	148	314.69	2409.15088	270.00000	0.00634	0
139	295.55	2409.15039	270.00000	0.02115	0	149	316.82	2409.15088	270.00000	0.00554	0
140	297.68	2409.15063	270.00000	0.01851	0	150	318.94	2409.15088	270.00000	0.00485	0
141	299.81	2409.15088	270.00000	0.01619	0	151	321.07	2409.15088	270.00000	0.00424	0
142	301.93	2409.15088	270.00000	0.01417	0	152	323.19	2409.15088	270.00000	0.00370	0
143	304.06	2409.15088	270.00000	0.01239	0	153	325.32	2409.15088	270.00000	0.00324	0
144	306.18	2409.15088	270.00000	0.01084	0	154	327.45	2409.15088	270.00000	0.00283	0
145	308.31	2409.15088	270.00000	0.00948	0						

Conforme a la Tabla 7.28b y con base en un análisis de  $T_k=10,000$  años, el máximo incremento de volumen almacenado en el vaso de la presa Huites es de 1,765.28638 (hm3), y la máxima elevación del nivel del agua en el embalse de la presa Huites es de 290.9 (m.s.n.m.).

En la Figura 7.33, se muestran los tránsitos de avenida simulados para los casos de análisis de probabilidad de retorno, en las comentadas gráficas se ha incluido para cada caso, el comportamiento de la curva de niveles de agua que se presentan en la superficie libre del embalse de la presa Huites, en función del tiempo.

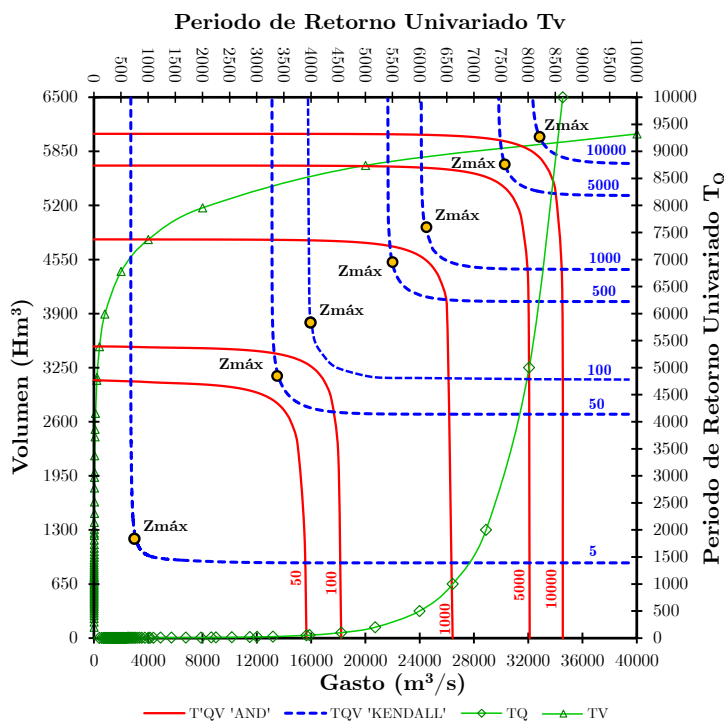


**Figura 7.33.** Tránsitos de avenidas simulados sobre el embalse de la presa Huites, Sinaloa, México, que produjeron los efectos más críticos en relación con el nivel máximo en el embalse; para periodos de  $T_k = 5,000$  años (a) y  $T_k = 10,000$  años (b) respectivamente.

**Tabla 7.28c.** Eventos de diseño que se pueden presentar en la presa Huities; análisis mediante un enfoque bivariado con base en cópulas.

EVENTOS DE DISEÑO BIVARIADOS			
RETORNO	ELEV.	CÓPULA DE JOE/B5	
$T_{kendall}$	$Z_{m\acute{a}x}$	$Q_p$	$V$
(años)	(m.s.n.m.)	( $m^3/s$ )	( $Hm^3$ )
50	279.61	13,500	3,150.00
100	280.09	15,895	3,875.00
500	284.75	22,000	4,520.00
1000	286.18	24,500	4,940.00
5000	289.44	30,269	5,694.62
10000	290.90	32,832	6,024.65

En la Figura 7.34, se muestran las curvas de periodo de retorno mediante los enfoques bivariados de "Kendall" (líneas discontinuas) y "AND" (líneas continuas), asimismo se presentan las curvas de periodo de retorno con enfoque univariado considerando a los gastos pico y volúmenes de la avenida. En este mismo gráfico, se ubican los conjuntos ( $Q_p$ ,  $V$ ) que generan el máximo nivel en el embalse de la presa Huities, por lo tanto, en esos puntos críticos se localizan las condiciones más desfavorables para caracterizar las distintas avenidas de diseño bajo un enfoque bivariado por medio de cópulas.



**Figura 7.34.** Comparación entre periodos de retorno bivariados y univariados

### **7.3. Avenidas de diseño de la presa Huites obtenidas en el libro de Ramírez y Aldama (2000)**

Para fines de comparación entre el método que se ha propuesto en esta tesis con el método presentado en el libro de Ramírez y Aldama (2000), se resumirán los hallazgos que consiguieron los autores mencionados. Cabe señalar que se utilizó idéntica información hidrométrica de Huites.

El análisis de frecuencia bivariado de avenidas que utilizaron los autores se basa en la Función de distribución de probabilidad logística bivariada de valores extremos (ec. 4.8), emplearon como marginales por default para  $Q_p$  y  $V$ , la FDP Doble Gumbel, el parámetro de asociación  $m$  empleado está en función del coeficiente de correlación lineal de Pearson  $\rho$  (ec. 4.11), en el ámbito de estadística avanzada este coeficiente se considera de baja potencia.

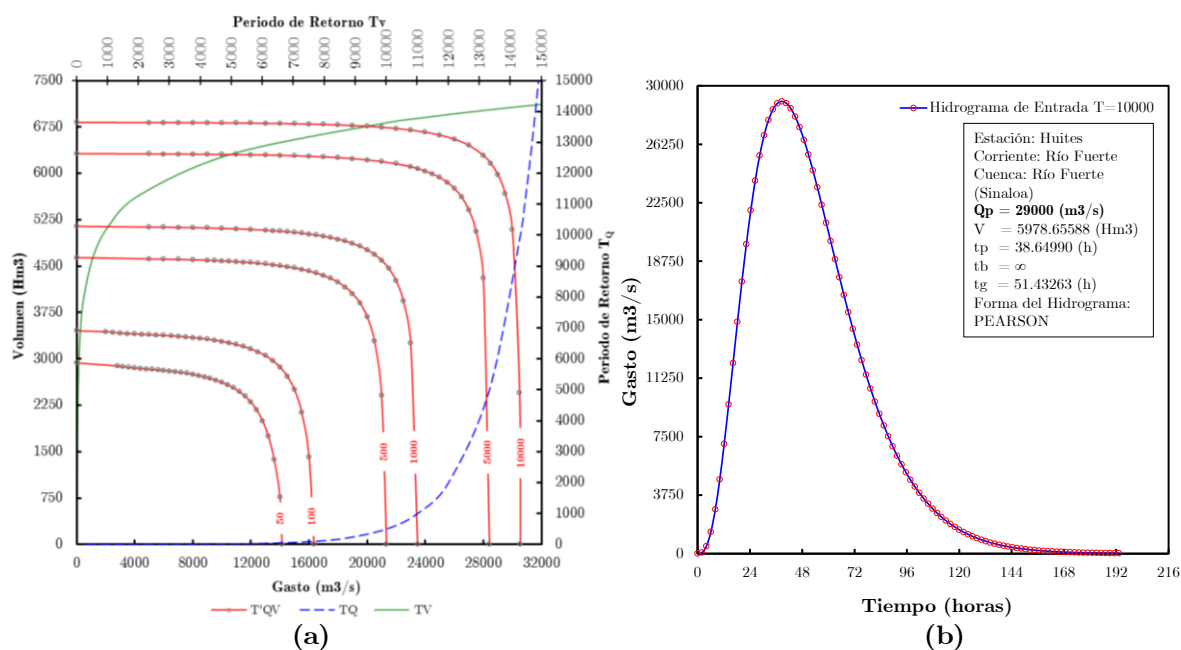
Los parámetros de las distribuciones de  $Q$  y  $V$  resultaron:  $\alpha_{q1}=740.660$ ,  $\beta_{q1}=1604.570$ ,  $\alpha_{q2}=3071.530$ ,  $\beta_{q2}=6669.270$ ,  $\alpha_{v3}=304.020$ ,  $\beta_{v3}=531.940$ ,  $\alpha_{v4}=728.610$ ,  $\beta_{v4}=1324.470$ ; los parámetros de segregación son  $P_q = 0.7618$  y  $P_v = 0.8101$ ; el parámetro de asociación es  $m = 1.6021$ , la optimización de parámetros se realizó con el método de Rosenbrock; la FDP bivariada es el modelo logístico bivariado; para estimar el periodo de retorno conjunto emplearon un tipo del modelo "AND" (ec. 4.19), la diferencia es que utilizaron una FDP en vez de una cópula; utilizaron la parametrización de Pearson III para la construcción de los hidrogramas completos de entrada; generaron combinaciones de gasto pico y volumen de escurrimiento (generación de muestras cortas para cada  $T_{q,v}$ ).

La combinación de gasto pico y volumen de escurrimiento que produce los efectos más desfavorables para un periodo de retorno  $T_{q,v} = 10,000$  años, es de  $Q_p = 29,000$  ( $m^3/s$ ) y  $V = 5,979$  ( $Hm^3$ ) (Figura 7.35). Calcularon el un  $t_p = 38.65$  (h), con base en las características de la cuenca del río Fuerte. Ésta combinación genera una elevación máxima en el vaso de la presa Huites de  $Z_{m\acute{a}x} = 290.6$  (m.s.n.m.).

Las elevaciones máximas para diferentes periodos de retorno y sus correspondientes conjuntos de valores ( $Q$ ,  $V$ ) que obtuvieron Ramírez y Aldama (2000), se presentan en la Tabla 7.29.

**Tabla 7.29.** Elevaciones máximas para diferentes periodos de retorno.

EVENTOS DE DISEÑO BIVARIADOS			
RETORNO	ELEV.	FDP LOGÍSTICO	2D
$T_{AND}$ (años)	$Z_{m\acute{a}x}$ (m.s.n.m.)	$Q_p$ ( $m^3/s$ )	$V$ ( $Hm^3$ )
50	279.44	12,400	2,174
100	282.05	15,000	2,500
500	285.38	20,000	3,681
1000	286.42	22,000	4,261
5000	289.33	27,000	5,410
10000	290.58	29,000	5,979



**Figura 7.35.** Curvas de periodo de retorno "AND" y Avenida de diseño.

## 7.4. Avenidas de diseño de la presa Huites obtenidas por el IIUNAM

Investigadores del Instituto de Ingeniería UNAM determinaron eventos de diseño ( $Q$ ,  $V$ ) con base en la Función de distribución logística bivariada de Gumbel (ec. 4.8), para la estación hidrométrica Huites, Sinaloa, México [Arganis *et al.* (2013)]. Utilizaron similar longitud de muestra de datos observados de gastos instantáneos que se enuncian en este trabajo, Tabla 7.8. Emplearon como mejores marginales que caracterizan el comportamiento de las variables  $Q$  y  $V$ , a las FDP Doble Gumbel (ec. 2.10). Emplearon al parámetro de asociación  $m$  para medir la dependencia de las variables  $Q$  y  $V$ , el cual está en función del

coeficiente de correlación lineal de Pearson  $\rho$  (ec. 4.11). Los once parámetros de la FDP logística bivariada de Gumbel se optimizaron utilizando el método de máxima verosimilitud, con base en esta técnica maximizaron el logaritmo de la función de verosimilitud  $L$ , usando un algoritmo genético simple [Arganis *et al.* (2013)].

El procedimiento que siguieron para la optimización de la FDP logística bivariada de Gumbel es el siguiente: Los parámetros estadísticos de ambas marginales (cinco parámetros  $c/u$ ) se ajustaron por separado y junto con el parámetro de asociación  $m$ , se sustituyeron en la FDP logística bivariada de Gumbel y por medio de la técnica de algoritmos genéticos se optimizó nuevamente de manera global los once parámetros que involucra la FDP bivariada, llegando así a lo siguiente:  $\alpha q_1=680.94$ ,  $\beta q_1=1516.39$ ,  $\alpha q_2=3140.60$  y  $\beta q_2=5729.79$ ;  $\alpha v_3=314.77$ ,  $\beta v_3=560.35$ ,  $\alpha v_4=686.07$  y  $\beta v_4=2000.00$ ; los parámetros de segregación son  $P_q = 0.7383$  y  $P_v = 0.9056$ ; el parámetro de asociación es  $m = 1.6668$  [Arganis *et al.* (2013)] y [Domínguez *et al.* (2009)].

El periodo de retorno conjunto se estimó con el tipo de modelo "AND" (ec. 4.19), para resolver dicha ecuación y obtener eventos de diseño, probaron diversos métodos numéricos, siendo el más eficiente en cuanto a estabilidad y convergencia el método de bisección [Arganis *et al.* (2013)], en el Apéndice D se puede consultar un código programado en Matlab.

Los investigadores del IIUNAM, los alcances que lograron en su investigación, están basados únicamente en determinar combinaciones de parejas  $Q$  y  $V$ , considerando periodos de retorno conjuntos  $T_{AND}= 1,000$  y  $T_{AND}= 5,000$  años, para cada caso, propusieron una serie de gastos y con la expresión 4.8 calcularon los volúmenes asociados, en primera instancia el proceso de cálculo fue fijar al gasto pico. Arganis *et al.* (2013), presentaron una tabla con resultados importantes al incluir diferentes combinaciones de eventos de diseño  $Q$  y  $V$ , exclusivamente se muestran los casos cuando  $T_{AND}= 1,000$  años para un  $Q_p= 22,000$  ( $m^3/s$ ) y alcanzar un  $V= 4,242$  ( $Hm^3$ );  $T_{AND}= 5,000$  años para un  $Q_p= 27,000$  ( $m^3/s$ ) y alcanzar un  $V= 5,390$  ( $Hm^3$ ).

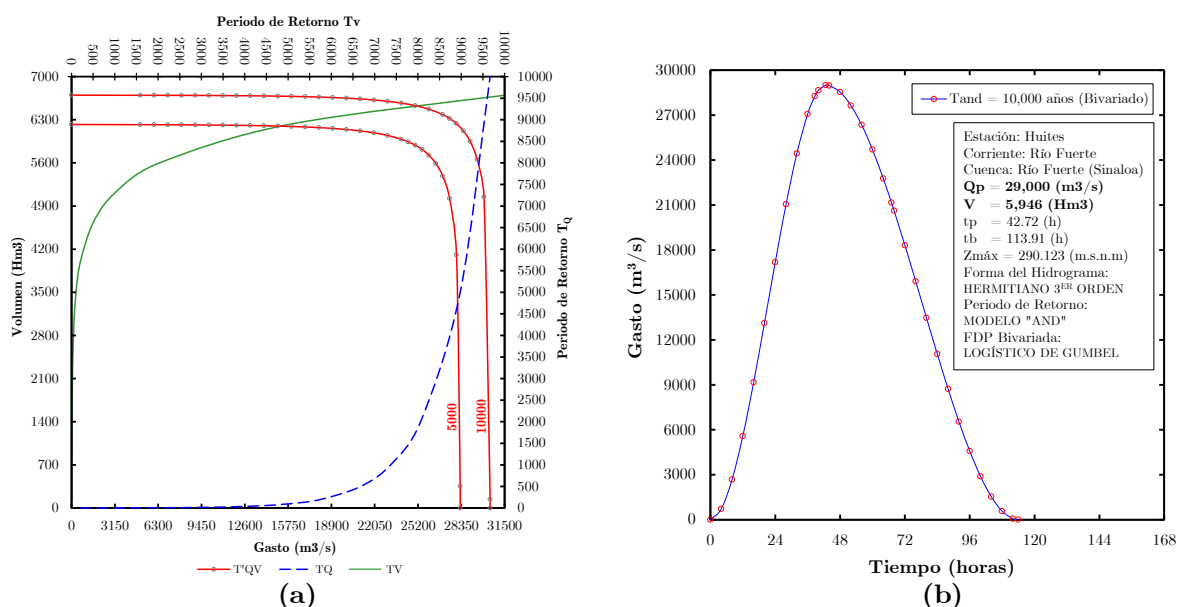
Para fines de esta tesis, el estudio elaborado por Arganis *et al.* (2013), se ha complementado, con el objetivo de realizar comparaciones. El procedimiento es similar como el expuesto en el Capítulo 6. Conforme a ello, la combinación de gasto pico y volumen de escurrimiento que



produce los efectos más desfavorables para un periodo de retorno  $T_{AND} = 10,000$  años, es de  $Q_p = 29,000$  ( $m^3/s$ ) y  $V = 5,946$  ( $Hm^3$ ), Figura 7.36. Ésta combinación genera una elevación máxima en el vaso de la presa Huites de  $Z_{m\acute{a}x} = 290.12$  m.s.n.m. Las elevaciones máximas para diferentes periodos de retorno y sus correspondientes conjuntos de valores  $(Q, V)$ , se presentan en la Tabla 7.30.

La citada metodología implementada en Arganis *et al.* (2013), no considera en construir avenidas de diseño ni mucho menos transitarlas. Con base en la investigación de Domínguez y Arganis (2012), ambos investigadores del IIUNAM, han implementado en dar forma al hidrograma de entrada a la presa, citan algunas metodologías, entre las que destacan son la Función Gamma de dos parámetros con tiempo pico al centro, tratado en el subtema 5.3 de este trabajo y el hidrograma hermitiano de tercer orden, bajo este último enfoque se calculó un  $t_p = 42.72$  (h) y  $t_b = 113.91$  (h), con base en  $Q_p$  y  $V$ .

Para complementar los resultados obtenidos por Arganis *et al.* (2013), se le ha dado forma al hidrograma de entrada a la presa Huites, empleando el hidrograma hermitiano de tercer orden, para el caso de  $T_{AND} = 10,000$  años, generando muestras cortas de  $Q$  y  $V$ , Figura 7.36.



**Figura 7.36.** Curvas de periodo de retorno "AND" (a) y Avenida de diseño con periodo de retorno  $T_{AND} = 10,000$  años (b), para la presa Huites; la forma de los hidrogramas son de tipo Hermitiano de tercer orden.

**Tabla 7.30.** Elevaciones máximas para diferentes periodos de retorno.

EVENTOS DE DISEÑO BIVARIADOS			
RETORNO	ELEV.	FDP LOGÍSTICO 2D	
$T_{AND}$ (años)	$Z_{m\acute{a}x}$ (m.s.n.m.)	$Q_p$ (m <sup>3</sup> /s)	$V$ (Hm <sup>3</sup> )
1,000	286.001	22,000	4,242
<b>5,000</b>	289.094	<b>27,000</b>	<b>5,390</b>
<b>10,000</b>	290.123	<b>29,000</b>	<b>5,946</b>

Para fines de comparación de los resultados obtenidos con la FDP Logística Bivariada analizados por IIUAM con otros modelos cópula, se realizó la prueba de bondad de ajuste por medio del método del mínimo error estándar, en la Tabla 7.31 se muestra la relación de los modelos evaluados, cuyo parámetro de descarte es el EEA, conforme a esta relación, el modelo cópula Joe/B5 es la mejor opción, mientras que la FDP Logística Bivariada obtuvo la quinta posición de mejor ajuste.

No se optó por realizar la prueba de bondad de ajuste con el método de González-Hoyos, por el hecho de que la FDP Logística Bivariada no es un modelo cópula.

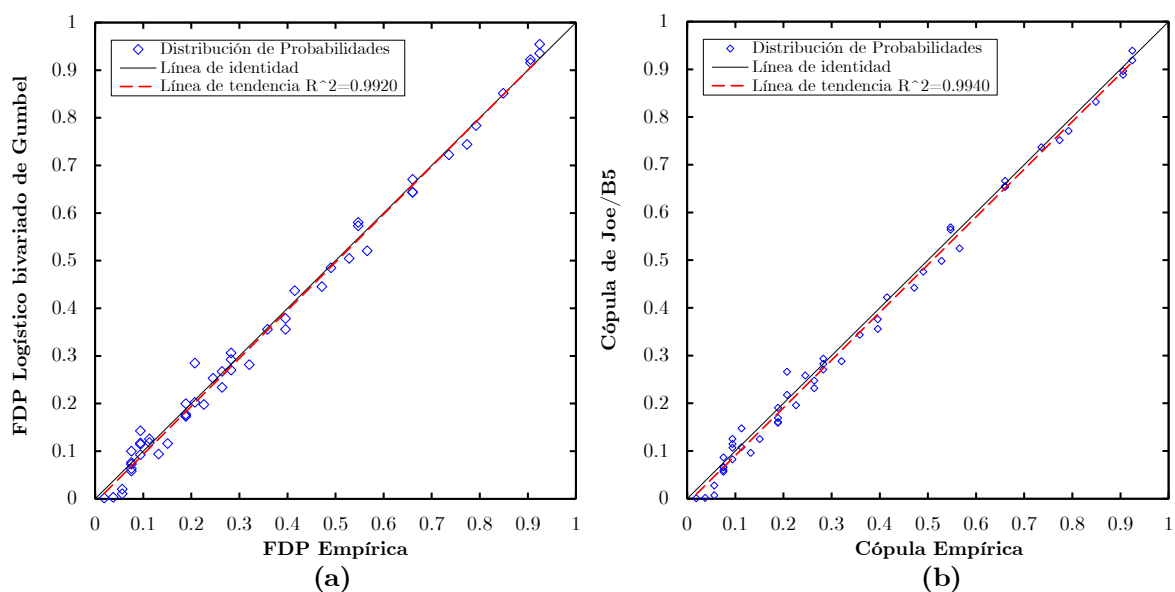
**Tabla 7.31.** Resumen de la prueba de bondad de ajuste mediante el EEA

Posición	Modelo	EEA
1	Joe/B5	0.0238
2	Gumbel	0.0250
3	Galambos	0.0251
4	Husler-R	0.0252
<b>5</b>	<b>FDP Log. Biv.</b>	<b>0.0258</b>
6	Gaussiana	0.0284
7	Frank	0.0293
8	Plackett	0.0295
9	Clayton	0.0366

En la Figura 7.37, se muestran dos gráficos que representan la correlación que existe entre la FDP Logística Bivariada con la FDP Empírica y la Cópula de Joe/B5 con la Cópula Empírica respectivamente.

En la Figura 7.37 (b), se observa que la línea de tendencia es prácticamente paralela a la línea de identidad, por lo que existe una mayor concentración de puntos sobre la línea de tendencia

En la Figura 7.37 (a) se observa una mayor dispersión de puntos, asimismo la línea de tendencia cruza en la parte superior derecha a la línea de identidad.



**Figura 7.37.** Distribución de probabilidades con base en la FDP Log. Biv. vs FDP Empírica (a) y la Cópula de Joe/B5 vs Cópula Empírica (b).

- **Estudio de Avenidas de diseño de la presa Huites: un enfoque univariado**

El Instituto de Ingeniería UNAM realizó un estudio relativo a la determinación y actualización de las avenidas de diseño de la presa Huites sobre el río Fuerte, a partir de registros de gastos medios diarios e instantáneos, con el fin de revisar la operación de su obra de excedencias. En el estudio elaborado por Domínguez *et al.* (2019) presentaron un análisis de frecuencia univariado de avenidas, para las épocas de invierno (meses de octubre a marzo) y de verano (meses de abril a septiembre).

Domínguez *et al.* (2019), realizaron el proyecto para extender el registro histórico de gastos medios diarios presentados en la presa Huites, que incluye al periodo de 1942 al 2018. Con el propósito de extender la longitud de la muestra citada, utilizaron datos de gastos medios diarios y gastos máximos instantáneos, extraídos del Banco Nacional de Aguas Superficiales (BANDAS) de la Comisión Nacional del Agua (CONAGUA) durante el periodo (1942-1992) con referencia a la estación Huites; de 1993-1995, se consideró porcentajes a partir de la estación Tubares; de 1995-2006, se seleccionó los meses con los mayores datos reportados del funcionamiento diario proporcionados por la Comisión

Federal de Electricidad (CFE); para 2007-2016 se obtuvieron los gastos medios diarios deducidos con base en los volúmenes de ingreso a partir del funcionamiento diario que reporta la CONAGUA; y la información de 2017 a 2018 se obtuvo de archivos del funcionamiento diario que proporcionó la CFE [Domínguez *et al.* (2019)].

Para la generación de avenidas de diseño, se consideró la época de verano e invierno por separado, y se empleó el método del Instituto de Ingeniería [Domínguez *et al.* (2006)]. ***A partir del registro histórico de gastos medios diarios de 1942 al 2018***, se obtuvo las series de gastos máximos anuales para duraciones de 1 a 40 días con el fin de obtener curvas gasto-duración-periodo de retorno (Q-d-T). Al aplicar el método a los gastos medios máximos de verano, el análisis de frecuencias dio como funciones de distribución de mejor ajuste las de tipo Doble Gumbel y Gumbel, para el caso de los datos de invierno se seleccionó a las funciones Doble Gumbel como la mejor opción. Con las funciones de distribución obtenidas se estimaron eventos para distintos periodos de retorno y se construyeron las curvas Q-d-T.

La avenida de diseño de verano a nivel diario, para un periodo de retorno de 10,000 años se estimó un gasto máximo de 7562.02 m<sup>3</sup>/s. Asimismo, para generar el hidrograma se tomó en cuenta una variación horaria y se consideró un incremento en el tiempo de 3 horas. Para analizar la capacidad de regulación del embalse se optó por darle una forma al pico de la avenida, para ello revisaron avenidas históricas durante el verano en la estación y se localizó al gasto medio diario más desfavorable de 3977 m<sup>3</sup>/s y un gasto máximo instantáneo de 4790 m<sup>3</sup>/s. Conforme a lo anterior, se realizó el cociente entre del gasto máximo instantáneo histórico y el gasto medio diario, resultando un factor de 1.2044, en este sentido, al gasto máximo de diseño diario (7562.02 m<sup>3</sup>/s), se multiplicó por el factor antes obtenido y se obtuvo un gasto pico de 9107.89 m<sup>3</sup>/s.

La avenida de diseño de invierno a nivel diario, correspondiente a un periodo de retorno de 10,000 años se estimó un gasto máximo de 18791.71 m<sup>3</sup>/s, para darle una forma al pico de la avenida se revisaron avenidas históricas, reportándose un gasto medio diario de 8992 m<sup>3</sup>/s y un gasto instantáneo de 15,000 m<sup>3</sup>/s, posteriormente se aplicó un procedimiento similar al de verano y se obtuvo un factor de 1.6681, para finamente calcular un gasto pico de 31,347.38 m<sup>3</sup>/s.

Se transitaron de las avenidas de verano y de invierno, sin considerar la forma en el pico de las avenidas, se trabajó con tres escenarios de elevación inicial en el embalse [Domínguez *et al.* (2019)]: primero considerando el NAMO que actualmente maneja el organismo operador del embalse CFE (270 m.s.n.m.) y la política de operación actual; Para las avenidas de invierno, se propuso la elevación inicial en el embalse de 276 m.s.n.m., se logró sobreelevar el NAMO original en 6 m, se planteó una nueva política de las descargas del vertedor, con lo que se conseguiría un mayor almacenamiento en el embalse, sin riesgo de rebasar el NAME (290 m.s.n.m.), ya que la elevación máxima que se podría alcanzar en la época de invierno ante una avenida con periodo de retorno de 10,000 años sería de 285.68 m.s.n.m.; Para el caso del verano se propuso un NAMO a la elevación 280 m.s.n.m., asimismo se planteó una nueva política de extracción de verano y no habría peligro de rebasar el NAME, aun considerando un pico de avenida de 9107.89 m<sup>3</sup>/s, por lo tanto, la elevación máxima que podría alcanzarse al transitar esta avenida por el embalse es 282.40 m.s.n.m., para un periodo de retorno de 10,000 años.

## **7.5. Comparación de resultados con las tres metodologías del análisis de frecuencias bivariado de avenidas.**

En este apartado se hace un análisis de resultados sobre los tres enfoques para la determinación del análisis de frecuencia conjunto y las avenidas de diseño más desfavorables, estos son: método propuesto en esta tesis (M1), basado en cópulas y forma de hidrogramas con la Función Gamma 2p; método propuesto por Ramírez y Aldama (2000) (M2), basado en la FDP Logística Bivariada de Gumbel (marginales asignadas por conveniencia Doble Gumbel) y forma de hidrogramas con la Función Pearson; método propuesto por el IIUNAM con enfoque bivariado (M3), basado en la FDP Logística Bivariada de Gumbel (marginales asignadas por conveniencia Doble Gumbel) y forma de hidrogramas triparamétricos hermitianos, finalmente el método (M4) con enfoque univariado.

Cabe destacar que, la selección de las marginales más apropiadas que se ajustaron a los datos observados de la estación Huites obtenidas con los tres métodos corresponden a la FDP Doble Gumbel, en lo que respecta a la selección del mejor modelo bivariado que caracterice a las marginales,

en los métodos M2 y M3 existe la restricción de que la FDP Logística Bivariada, única y exclusivamente acepta marginales de valores extremos, lo cual es una limitante, mientras que el método M1 se tiene plena libertad de elegir una amplia gama de modelos cópula y utilizando las marginales que mejor se ajustan a los datos observados, por este motivo la metodología M1 que se ha propuesto tiene mayor sustento matemático para un adecuado análisis bivariado o n-variado.

Otro aspecto a notar, es la forma de estimar el periodo de retorno conjunto  $T_{Q,V}$ , en los métodos M2 y M3 utilizaron el modelo "AND" adaptado a funciones de distribución, mientras que en el método M1 se probaron tres modelos, "Kendall", "AND" y "OR", siendo el modelo "Kendall" el que ofrece mejores resultados. El modelo "AND" obtiene los valores de  $T_{Q,V}$  más pequeños en relación a los otros dos, el modelo "OR" obtiene los valores de  $T_{Q,V}$  más grandes en relación a los otros dos, y el modelo "Kendall" sus curvas de  $T_{Q,V}$  se posicionan entre las curvas de "AND" y "OR". Las curvas de  $T_{Q,V}$  de los modelos "Kendall" y "OR" son convexas, mientras que las curvas de  $T_{Q,V}$  de "AND" son cóncavas.

La ventaja de emplear el método de cópulas para la mejor elección de una distribución bivariada, radica en que no existe restricciones para utilizar marginales, asimismo es posible extender muestras con longitudes muy grandes, extraídas directamente del modelo cópula seleccionado, lo cual hace que se tenga una muestra bivariada con infinidad de probabilidades de ocurrencia. En relación a la obtención de las avenidas de diseño, primeramente se hace un análisis gráfico de las avenidas históricas más representativas del registro en la estación, posteriormente con la forma de la avenida se busca una función de distribución que asemeje su forma, conservando en todo momento el volumen de la avenida y el gasto pico.

En los métodos M1 y M2 se empleó la Función Gamma 2P para generar los hidrogramas, mientras que el método M3 se empleó los hidrogramas hermitianos 3<sup>o</sup> orden. La simulación del tránsito de avenidas varía en sus resultados entre un método y otro, esto se debe a la curva E-V-Qs que se utilice, asimismo al tipo de hidrograma que se emplee y por supuesto la elección del tiempo pico, todas estas variables influenciarán al resultado de  $Z_{m\acute{a}x}$ , para M1, M3 y M4 se utilizó el programa TRATE.bas (proporcionado por IIUNAM) para transitar las avenidas en embalses.

En la Tabla 7.32, se muestra la comparación de parámetros estadísticos de las marginales que representan a  $Q$  y  $V$ , de los tres métodos que se han comentado. En el método M1 se emplearon los parámetros estadísticos que originalmente resultaron al realizar el análisis de ajuste de los datos observados con base en el programa AX.exe [Jiménez, (1997)], tales parámetros estadísticos se utilizaron sin alteraciones para generar la cópula y extender las muestras sintéticas.

*En los métodos M2 (algoritmo de Rosenbrock) y M3 (algoritmo genético) se hace un comentario de gran interés, en primera instancia lo que hicieron los autores es calcular los parámetros estadísticos de la FDP Doble Gumbel para las variables  $Q$  (5 parámetros) y  $V$  (5 parámetros), asimismo calcularon el parámetro de asociación  $m$  de la FDP bivariada, luego conformaron la FDP Logística Bivariada a través de las marginales  $Q$  y  $V$  ajustadas con la FDP Doble Gumbel en ambos casos y el parámetro  $m$ , nuevamente los autores optimizaron los once parámetros estadísticos que se incluyen en la distribución bivariada.*

*Este último proceso es de tomarse en cuenta, ya que al optimizar por segunda vez los parámetros de la marginales ya incluidas en la FDP Logística Bivariada, automáticamente dichos parámetros ya no generan el mínimo error estándar de ajuste en las variables  $Q$  y  $V$ , consulte la Tabla 7.32 para más detalles, por este hecho, las muestras sintéticas que se llegaron a generar con base en estas metodologías, no se acercarán a la realidad del comportamiento de los datos observados, ya que se tendría una imprecisión en los parámetros.*

Con el método M1, no existen estas imprecisiones, ya que la metodología de funciones cópula está gobernada por las medidas de dependencia  $\rho$  y  $\tau$  y a su vez por el parámetro de la cópula  $\theta$ , asimismo la cópula acepta cualquier marginal y sin necesidad de ajustar los parámetros nuevamente.

**Tabla 7.32.** Comparación de la medida EEA que resultaron al optimizar por segunda vez los parámetros estadísticos en M2 y M3 con M1.

Método	$\alpha_{q1}$	$\beta_{q1}$	$\alpha_{q2}$	$\beta_{q2}$	$\alpha_{v3}$	$\beta_{v3}$	$\alpha_{v4}$	$\beta_{v4}$	$P_q$	$P_v$	$\theta$ $m$	EEA; Q	EEA; V
M1	888.10	1581.37	3521.13	8569.59	315.96	558.43	550.36	2207.01	0.84	0.89	2.0837	286.37	84.33
M2	740.66	1604.57	3071.53	6669.27	304.02	531.94	728.61	1324.47	0.76	0.81	1.6021	507.43	121.51
M3	680.94	1516.39	3140.60	5729.79	314.77	560.35	686.07	2000.00	0.74	0.91	1.6668	543.25	122.90

En lo que respecta al método M4, no se podría realizar alguna comparación de manera detallada, puesto que el enfoque que se utiliza en M1, M2 y M3 es bivariado, mientras que M4 es un análisis univariado con gran detalle, la metodología toma en cuenta una variable estacional para su estudio (verano e invierno).

Emplearon la FDP Doble Gumbel para ajustar los datos observados de gastos medios diarios, para la obtención de los parámetros estadísticos utilizaron el programa AX.exe. Para construir el hidrograma de diseño para un periodo de retorno dado, se seleccionaron los gastos medios máximos diarios de 1 a 40 días (dado que para duraciones menores los hidrogramas varían en tiempo, por ejemplo véase la Figura 7.6, en la cual se representa la avenida máxima observada sobre el río Fuerte, asimismo se perciben dos picos), y con estos valores se estimaron los gastos diarios o individuales y se le da forma al hidrograma con el método de bloques alternos [Domínguez *et al.* (2006)].

## **7.6. Conclusiones sobre los resultados obtenidos**

La Función de distribución de probabilidad bivariada que mejor representa a la muestra de datos observados de la estación Huites, es el modelo con la Función Cópula Joe/B5 de un parámetro. Con base en esta cópula fue posible generar muestras sintéticas para cada periodo de retorno conjunto.

El conjunto de variables de gasto pico y volumen de escurrimiento que produce los efectos más desfavorables en la avenida, con la metodología propuesta M1 y para un periodo de retorno  $T_k = 5,000$  años, es de  $Q_p = 30,269$  (m<sup>3</sup>/s) y  $V=5,695$  (Hm<sup>3</sup>), condiciones que alcanzan un nivel máximo en el vaso de la presa Huites de  $Z_{m\acute{a}x} = 289.4$  (m.s.n.m.), parámetro que se acepta entre los límites permisibles de la infraestructura de la presa Huites en función del nivel del NAME y la corona.

Para el escenario con un  $T_k = 10,000$  años, se produjo la avenida más desfavorable con las variables  $Q_p = 32,832$  (m<sup>3</sup>/s) y  $V= 6,025$  (Hm<sup>3</sup>), esta avenida, al ser transitada por el vaso de la presa Huites alcanzó un nivel máximo en la superficie libre del agua de  $Z_{m\acute{a}x} = 290.9$  (m.s.n.m.),



---

parámetro que sobrepasa el límite del nivel del NAME por 90 (cm), asimismo la corona de la cortina es superada por 15 (cm).

Derivado de los resultados obtenidos y una vez aplicado métodos bivariados, la presa hidroeléctrica Luis Donaldo Colosio "Huites" no resulta hidrológicamente segura para el periodo de retorno de 10,000 años. El riesgo hidrológico de falla de la presa Huites, considerando una vida útil de cincuenta años y un periodo de retorno original de  $T=10,000$  años es del 0.499%, mientras que el riesgo de falla actual de la presa con base en los resultados con  $T_{NAME} = 6875$  años, es del 0.725%. Por esta razón, se puede afirmar que la presa Huites es 1.45 veces más insegura de lo que se tenía presupuestado.

Por otro lado, lo que se percibe en los resultados del método M4, es por ejemplo el valor del gasto pico obtenido durante el período de invierno, cuyo gasto se estimó en  $Q_p = 31,347$  ( $m^3/s$ ), resultado que no se obtuvo directamente del análisis, mejor dicho, se le dio forma al pico con base en la avenida histórica de 1960 utilizando gastos medios máximos diarios, este valor de gasto se aproxima al obtenido en M1, que es de  $Q_p = 32,832$  ( $m^3/s$ ), para ambos casos se analiza la avenida de diseño para un  $T=10,000$  años, y condición inicial de 270 (m.s.n.m.) al iniciar el tránsito de avenida.

Algunas de las diferencias sustanciales que impiden hacer diversas comparaciones con la metodología planteada en este trabajo, radica en lo siguiente: no se utiliza la misma longitud de muestra de datos observados; se difiere en la curva E-V- $Q_s$  debido al criterio de estacionalidad; la duración que se propone en M4 es muy grande en comparación con M1; en M4 no se define claramente la manera de estimar el tiempo pico  $t_p$ , sin embargo en M1 se establecieron diversos criterios; en M1 se utilizaron gastos máximos instantáneos, no así en M4, optaron por gastos medios diarios máximos.

Conforme a los resultados de M4, al aplicar el método del IIUNAM sí resulta hidrológicamente segura la presa Huites, hasta con un periodo de retorno de 10,000 años.

Uno de los inconvenientes de aplicar el enfoque bivariado en el análisis de avenidas máximas anuales, radica en la obtención de las series de datos de las variables  $Q_p$  y  $V$ , por un lado, la variable  $Q_p$  es

relativamente sencilla de obtener porque es el máximo valor del gasto medio diario en el año, sin embargo la estimación del volumen  $V$  generado durante esa avenida, tiene cierto grado de subjetividad, ya que depende de la identificación del punto de recesión del escurrimiento directo del hidrograma, para corregir este problema se tiene que implementar métodos de filtros digitales para su obtención, que son capaces de separar el flujo base del directo. Lo ideal es disponer de datos de volúmenes medidos en el embalse de la presa, a través su nivel.

El análisis del hidrograma por medio del método IIUNAM toma en cuenta un número suficiente de días de duración, de modo que no es necesario definir fechas de inicio y fin de las avenidas históricas, existiendo mayor versatilidad de formas en el hidrograma y logrando representar ampliamente la naturaleza del hidrograma, motivo suficiente de que otros métodos opten por utilizar criterios subjetivos ante este hecho.

En contraste, los métodos bivariados ofrecen un mayor sustento estadístico para la dependencia de las variables  $Q$  y  $V$ , se debe poner suma atención en obtener el volumen de la avenida ocasionada, optando por métodos más precisos que ofrece la Física de la ondas, como es el diseño de los filtros digitales, cursos que usualmente se enseñan en Geofísica y Eléctrica-Electrónica.

En el trabajo de Aranda D. (2014), muestra una aplicación muy interesante sobre un análisis hidrológico, tomando en cuenta la estacionalidad de las muestras de invierno y verano, el análisis comprende la aplicación de funciones cópula.

Durante el proceso para asignar la forma que tendría el hidrograma de entrada a la presa y por ello la estimación del tiempo pico, se identificó que en el momento de realizar simulaciones del tránsito de la avenida, los resultados del nivel del agua en el embalse varían considerablemente al emplear el tiempo pico calculado con diversas metodologías.

**Tabla 7.33.** Cálculo del tiempo pico con  $Q_p=32823$  (m<sup>3</sup>/s) y  $T=10000$  años

Método	tp (h)	Zmáx (m.s.n.m.)
SCS-CHOW (Características cuenca)	25.6828000	290.9
Hienstra y Francis (1981)	25.6988194	291.0
SCS-MRA (Características cuenca)	36.5213709	291.7
Ramírez y Aldama (2000)	38.2290271	292.2
Croley II (1980) (tp al centro)	59.8521895	292.8



# Capítulo 8

## Conclusiones, recomendaciones y futuras líneas de investigación

### 8.1. Conclusiones

La presente investigación proporciona metodologías para conducir a los profesionales especializados en temas hidrológicos, a realizar análisis de frecuencia bivariados de avenidas basados en Funciones Cópula, con el propósito de lograr obtener cuantiles  $q$  y  $v$  que caractericen a una avenida más eficientemente y que su vez sean adecuados para el diseño de presas, la gestión de presas y la evaluación del riesgo ante inundaciones.

La presente investigación desarrollada, ha estado encaminada a la justificación de la necesidad de un modelo multivariado que fuera capaz de integrar todas las variables que intervienen en el análisis de seguridad de presas, gasto pico y volumen de la avenida. Una vez justificado lo anterior, se ha propuesto un modelo multivariado (con enfoque bivariado) que fuera capaz de abordar la problemática de la seguridad de presas.

Los pilares básicos sobre los que se apoya esta tesis, son las Funciones Cópula, lo que permite desarrollar un análisis de frecuencias de avenida con enfoque bivariado y por lo tanto generar la avenida de diseño más desfavorable en el embalse de una presa. Se propuso una metodología y se aplicó a un caso particular.

Se ha llevado a cabo una exhaustiva revisión del estado del conocimiento actual, en especial la teoría de cópulas. En nuestros días, esta teoría comienza a tener importantes avances en cuanto a la aplicación en la ingeniería hidrológica. Conforme a la revisión bibliográfica, en Latinoamérica a la fecha no se han elaborado trabajos bajo este enfoque de manera formal.

El método para la estimación de avenidas para el diseño o revisión de presas se basa en una caracterización probabilista bivariada de los hidrogramas. El gasto pico y el volumen de escurrimiento de la avenida son determinados mediante el análisis de frecuencias bivariado, por lo cual se asigna un periodo de retorno a toda la avenida y no sólo al gasto máximo de la misma. Este proceso evita la arbitrariedad en la asignación del volumen, el cual tiene grandes efectos en la respuesta del vaso.

La dependencia entre las variables de gasto pico ( $Q$ ) y volumen de la avenida ( $V$ ), queda reflejado con mayor claridad desde un punto de vista gráfico (K-plot y Chi-plot), así como el estadístico, justificado por los indicadores Tau de Kendall y Rho de Sperman. Una vez vista la dependencia entre las variables  $Q$  y  $V$ , se procede al ajuste de las marginales y una vez ajustadas se procede a la elección de un modelo cópula. De esta forma se obtiene la función de distribución bivariada que es capaz de representar a la muestra inicial, procedimiento que se hace para cada una de las estaciones por analizar. Se obtiene de esta forma una herramienta (mejor cópula) capaz de generar muestras aleatorias, el tamaño de la misma se ha recomendado entre 10,000 y 100,00 pares ( $Q$ ,  $V$ ), que para este caso ha sido de 10,000 conjuntos por cada estación.

Con base en lo anterior, se genera un conjunto de hidrogramas sintéticos de entrada con determinada forma y asociados a diferentes periodos de retorno. Se transitan los hidrogramas sintéticos y se analiza el nivel

máximo alzado en el embalse de la presa, el hidrograma que produzca el nivel máximo será la avenida de diseño más desfavorable.

En este sentido, se encontró que la cópulas Joe/B5 es la mejor cópula que representa la estructura de dependencia entre el gasto pico y el volumen de la avenida en el estudio de caso considerado en la investigación, la presa Huites. En resumen, la mejor cópula debe representar adecuadamente los siguientes casos, la Estructura de dependencia de los pares observados  $(Q, V)$  y el comportamiento en la cola superior.

Se ha propuesto un análisis de frecuencia bivariado de avenidas, contemplando un enfoque de períodos de retorno bivariados teóricos basado en cópulas para obtener estimaciones de eventos de diseño (cuantiles) en la presa Huites. Las curvas teóricas del período de retorno conjunto representan la probabilidad de ocurrencia de un evento de avenida (el riesgo teórico sin tener en cuenta la estructura). Las curvas del período de retorno de Kendall (llamadas secundarias), se utilizaron para representar el riesgo de desbordamiento de la presa Huites. Además, el período de retorno de Kendall será útil en las evaluaciones de riesgo de avenidas.

El uso de hidrogramas obtenidos por medio de las Función Gamma 2P, descritos por dos parámetros evita la arbitrariedad en la selección del volumen de escurrimiento de la avenida además del gasto pico. El volumen de la avenida se elige de acuerdo con el criterio del análisis de frecuencias bivariado mediante la teoría de cópulas y para un determinado periodo de retorno de tipo Kendall. El evento de diseño resultante para un periodo de retorno dado se obtiene como la solución de un problema de optimización al transitar los hidrogramas sintéticos, de tal forma que se asegura que se está obteniendo la condición más desfavorable para el nivel de riesgo aceptado. El proceso de optimización permite además que las características del vaso y vertedor de la presa participen en la estimación del evento de diseño.

Derivado de los resultados obtenidos, la presa hidroeléctrica Luis Donaldo Colosio "Huites", para el escenario con un  $T_k = 10,000$  años, se produjo la avenida más desfavorable con las variables  $Q_p = 32,832$  (m<sup>3</sup>/s) y  $V = 6,025$  (Hm<sup>3</sup>), esta avenida al ser transitada por el vaso de

la presa alcanzó un nivel máximo en la superficie libre del agua de  $Z_{\text{máx}} = 290.9$  m.s.n.m., parámetro que sobrepasa el límite del nivel del NAME y la corona de la cortina en 15 (cm), asimismo  $Z_{\text{máx}}$  está al límite del parapeto de la cortina. El diseño actual de la presa Huites está sujeta a un riesgo de falla en su vida útil (50 años) ligeramente mayor que el correspondiente a un periodo de retorno de 10,000 años. De cualquier forma, la presa Huites es 1.45 más insegura de lo que se cree.

Esta metodología nos permite entender mejor la seguridad hidrológica de presas, permitiéndonos definir con mayor detalle la seguridad de las mismas así como fijar normas de explotación o si existiera algún déficit de seguridad establecer las medidas necesarias, lo cual influye en una mejora en cuanto a temas de inversión.

La metodología se aplica a un caso real de revisión, y muestra que algunas presas podrían ser más inseguras de lo que parecen (de acuerdo con el enfoque tradicional). El enfoque conjunto aplicado al caso de presas parece indicar que es necesario cambiar la filosofía en cuanto a la asignación del riesgo aceptado (periodo de retorno de diseño).

## **8.2. Aportaciones relevantes**

Análisis multivariado de frecuencias de avenidas con enfoque a Funciones Cópula.

Mediante las Funciones Cópula, siendo estas además de una función de distribución de probabilidad multivariada, son una medida de la relación que existe entre las variables que intervienen en el modelo.

Aplicación del método Cópula muestral de orden  $m$ , para evaluar la prueba de bondad de ajuste para seleccionar el mejor modelo cópula.

Un nuevo enfoque de estimar periodos de retorno conjuntos con base en la Función de Kendall y aun modelo cópula.

Un nuevo enfoque en la estimación de avenidas para el diseño de obras hidráulicas específicas, como las presas.

### **8.3. Recomendaciones para próximas investigaciones**

Se recomienda emplear otros modelos cópula de dos parámetros, enfocados al tratamiento de valores extremos: Cópula t-VE, Cópula Tawn y la Cópula Joe/BB5. Las cópulas de dos parámetros permiten una mayor flexibilidad para modelar la dependencia de la cola inferior y superior de la cópula (Joe, 1997).

Asimismo, en futuros análisis se tiene que realizar la evaluación de la dependencia en la cola superior de los diferentes modelos cópula a estudiar, con esto se garantiza un mejor análisis del comportamiento de valores extremos de las muestras sintéticas extendidas y por ello los periodos de retorno de magnitud grande [Joe (2015); Requena (2015)].

Es de suma importancia que se incluyan otros métodos para evaluar la Bondad de Ajuate para Cópulas, por ejemplo: la prueba estadística de Cramér-Von Mises adaptado a cópulas, la prueba estadística por medio de la cópula empírica propuesta por Genest et al. (2009), la prueba de Vuong (1989) y la de Clarke (2007) para cópulas empleadas en Guarín (2016). Como criterio de selección del mejor modelo, por ejemplo, el empleo de las técnicas de AIC y BIC mediante el procedimiento paramétrico de Bootstrap [Genest y Rémillard (2008)].

Proponer un procedimiento de análisis de sensibilidad para determinar la longitud mínima de la muestra sintética extendida.

La aplicación de las metodologías propuestas en la presente tesis a otros estudios de casos, con el fin de obtener más información sobre los resultados proporcionados.

Incorporar el análisis de filtros digitales para la estimación del volumen de la avenida, ya que los datos de volumen medidos en el embalse tienen un error considerable de medición, debido al oleaje, cubicación, etc.



## 8.4. Futuras líneas de investigación

De la presente investigación con enfoque bivariado  $(Q, V)$ , realizar la extensión formal al contexto trivariado, por ejemplo, incluyendo en el análisis la duración variable  $(d)$  del hidrograma, o el tiempo pico  $(t_p)$ .

Es importante resaltar que se han llevado a cabo desarrollos en la literatura en relación con el enfoque multivariado, aunque la aplicación de cópulas multivariadas todavía es compleja en la práctica. A este respecto, sería atractiva la aplicación de las prometedoras Cópulas Vine, que pueden proporcionar estructuras de dependencia complejas entre variables.

La adaptación y aplicación de las metodologías propuestas a escala nacional para proporcionar recomendaciones a los profesionales. Por ejemplo, proporcionar la distribución de probabilidad basada en la cópula bivariada para cada región del contexto nacional, para realizar un análisis de frecuencia bivariada de avenidas local o regional.

El estudio sobre la extensión de las variables de avenida observadas, podría aplicarse a otras cuencas para comparar el número de hidrogramas de avenidas necesarios, así como la distribución bivariada seleccionada en cada caso.

Quedan estos temas como líneas de investigación abiertos para futuros desarrollos en el sector hidrológico y la planificación hídrica.

# Apéndice A

## Coeficiente de Correlación de Spearman basado en Orden-Rango

### A.1. Introducción

En ocasiones se cuenta con datos que no siguen ningún tipo de distribución como la normal, binomial, etc., sin embargo es posible realizar pruebas y verificar hipótesis para formular conclusiones. Estos tipos de pruebas se denominan no paramétricas y no se consideran tan robustos como las pruebas paramétricas (*media, mediana o desviación estándar*). De hecho, en algunas pruebas no paramétricas, se ignoran los valores numéricos y simplemente se trabaja con el rango del valor de los datos (Steven *et al.*, 2016).

El coeficiente de Rho de Spearman es en realidad un procedimiento estadístico muy conocido, de hecho, es probablemente uno de los procedimientos no paramétricos más conocidos que existen. Este coeficiente, denotado como  $\rho_s$ , es esencialmente el equivalente no paramétrico de la  $\rho$  de Pearson, tal coeficiente es capaz de determinar si existe o no una correlación entre dos variables (Jared Linebach *et al.*, 2014).

La Rho de Spearman es extremadamente versátil, y la prueba se puede utilizar cuando ambas variables son ordinales o categóricas, asimismo, se puede usar datos donde una variable esté clasificada y la otra sea escala de intervalo o incluso cuando ambas variables son de escala de intervalo. Sin embargo, si se tienen datos que son escalas de intervalo, previamente deben convertirse en datos ordenados por categorías para poder emplearlos en el cálculo de Rho de Spearman (Jared Linebach *et al.*, 2014). Los datos pueden consistir en observaciones numéricas a las que se aplican rangos, o en observaciones no numéricas que únicamente se pueden clasificar. En el caso de empates en los valores de  $X_i$  o  $Y_i$ , se asigna un rango promedio.

El coeficiente de correlación de Spearman no es más que el coeficiente de correlación de Pearson aplicado a los rangos  $(R(X_i), R(Y_i))$  de las observaciones, de modo que no utiliza directamente las observaciones originales  $(X_i, Y_i), i = 1, \dots, n$  en el procedimiento (Stefano Bonnini *et al.*, 2014).

## A.2. Hipótesis

Sea  $(X_1, Y_1), \dots, (X_n, Y_n)$  una muestra aleatoria de una población bivariada continua con función de distribución conjunta  $H(X, Y)$  y funciones de distribución marginal  $F(X)$  y  $G(Y)$ . Sea  $R(X_i)$  el rango de  $X_i$  y de manera similar  $R(Y_i)$  el rango de  $Y_i$ . Es importante tener en cuenta que  $X$  e  $Y$  pueden ser incluso variables no numéricas, siempre y cuando puedan ser clasificadas, es decir, que sea factible ordenarlas categóricamente (Myles Hollander *et al.*, 2014). En esta sección, se probará la independencia entre las variables  $X$  e  $Y$  correspondientes a la hipótesis nula  $H_0$ .

## A.3. Procedimiento

Para calcular el coeficiente de correlación de Spearman  $\rho_s$  basado en rangos, existen dos escenarios, *una primera forma consiste en el análisis de una muestra cuyos valores no deben haber repeticiones, en el segundo*

---

*escenario la muestra presenta diversas repeticiones en sus valores, las cuales se identifican como "empates".* Para ambos casos, conforme a la literatura en estadística no paramétrica, existen dos expresiones para calcular el coeficiente, según la naturaleza de las muestras. Para fines de este trabajo, se presenta la metodología para ambos casos, además de que se ilustra la metodología con un ejemplo, y se comparan los resultados obtenidos con la implementación del *Software R* (de dominio público).

### A.3.1. Muestra con ausencia de empates

En ausencia de empates de una muestra, el cálculo del coeficiente de correlación de Spearman basado en rangos, computacionalmente es más accesible obtenerlo.

Se comienza por clasificar las  $n$  observaciones de la variable  $X$ , analizando el tamaño de observación (Rango) en magnitud ascendente, de modo que, se identifica la puntuación más baja y se etiqueta con el número "1", la segunda puntuación más baja se etiqueta como "2", y así sucesivamente, de manera similar, y por separado se clasifican las  $n$  observaciones de la variable  $Y$  de forma ascendente y análogamente se etiquetan las puntuaciones (Jared Linebach *et al.*, 2014).

Ahora bien, se obtiene la diferencia de los rangos de las variables  $X$  e  $Y$ , mejor dicho, se restan los rangos de las puntuaciones asociadas a cada variable, dicha diferencia se identifica como  $D_i$ .

Conforme a las diferencias obtenidas  $D_i$ , es posible que resulten números negativos, para eliminar los términos negativos se procede a elevar al cuadrado todas las diferencias resultantes asociadas a cada variable. El siguiente paso, consiste en calcular la suma de todas las diferencias al cuadrado  $D_i^2$ .

Cuando no hay empates dentro de una muestra, según Stefano Bonnini *et al.*, (2014), la  $\rho_s$  de Spearman es:

$$\rho_s = 1 - \frac{6 \sum_{i=1}^n [R(X_i) - R(Y_i)]^2}{n(n^2 - 1)} \quad (\text{A.1})$$

$$\rho_s = 1 - \frac{6 \sum_{i=1}^n D_i^2}{n(n^2 - 1)} \quad (\text{A.2})$$

donde

$n$  es el tamaño de la muestra bivariada

$D_i$  es la diferencia en las clasificación de rangos de cada variable.

La versión del coeficiente  $\rho_s$  de Spearman expresada en la ecuación A.1, es la más usada en el análisis hidrológico, por su versatilidad y facilidad de empleo y que no considera repeticiones en sus observaciones, la expresión A.1 se basa en que los rangos  $R(X_i)$  y  $R(Y_i)$  no se repiten.

Cabe señalar que, los registros históricos que concentran datos de escurrimientos y precipitaciones es muy poco probable que existan registros repetidos, debido a la aleatoriedad en que se presentan los eventos de lluvias y ciclones.

### A.3.2. Muestra con vínculos de empates

En el caso de observaciones repetidas en una muestra bivariada con determinada ordenación  $k$ , las cuales se identifican como "empates", el cálculo del coeficiente de correlación de Spearman basado en rangos, primeramente *se clasifican las  $n$  observaciones de la variable  $X$ , analizando su tamaño (columna "Ranking") en magnitud ascendente*, de modo que, se identifica la puntuación más baja y se etiqueta con el número "1", la segunda puntuación más baja se etiqueta como con el número "2", y así sucesivamente, de manera similar, y por separado se clasifican las  $n$  observaciones de la variable  $Y$  de forma ascendente y análogamente se etiquetan las puntuaciones conforme se hizo con  $X$  (Jared Linebach *et al.*, 2014).

Enseguida se **ordenan** las  $n$  observaciones de la variable  $X$  en **magnitud ascendente (menor a mayor)**, al tomar como pivote los puntajes de la variable  $X$ , simultáneamente también se modifica la clasificación del Ranking en forma ascendente además del orden  $k$  (tiempo en años, hablando hidrológicamente) que originalmente asocia a los puntajes de dicha variable, lo propio se realiza para la variable  $Y$ . Acorde a esta nueva organización de datos, el Ranking numérico más bajo es 1, mientras que el Ranking numérico más alto es  $n$ . Durante este desarrollo lógicamente se presentarán puntuaciones con empates.

Ahora bien, toda vez que se han ordenado las  $n$  observaciones de las variables  $X$  e  $Y$  con magnitud ascendente, se identifica a  $R(X_i)$  como el rango de  $X_i$ , de manera similar,  $R(Y_i)$  denota el rango  $Y_i$ . Conforme a este último arreglo de información, los puntajes de cada variable  $X$  e  $Y$  se han ordenado por categorías debido al surgimiento de empates, lo que ha originado que se conformen conjuntos de datos con igual puntuación, de modo que, es factible obtener rangos en función del Ranking asignado.

En otras palabras, supongamos que las variables  $X$  e  $Y$  presentan puntuaciones con empates, para obtener los rangos asociados a dichas variables y que previamente sus puntuaciones fueron ordenadas por categorías, **es factible que a todos los conjuntos de empates se les deba asignar un mismo rango promedio, calculado en función de sus posiciones (en orden ascendente) vinculadas al Ranking.**

Por ejemplo, suponga que en la variable  $X$  los cuatro primeros Rankings  $(1, 2, 3, 4)$  comparten la misma puntuación  $(a, a, a, a)$ , entonces se necesita promediar las posiciones de los Rankings  $(1, 2, 3, 4)$  asignados a las puntuaciones  $(a, a, a, a)$ , por lo tanto, los primeros cuatro Rankings comparten un rango de 2.5 asociado a las puntuaciones de la variable  $X$ . El siguiente rango asignado a la variable  $X$ , en principio de cuentas, tendría un rango de 5, a menos que existiesen más conjuntos de empates posteriores a la cuarta posición del Ranking.

En seguida de haber obtenido los rangos promedio, por separado, de las variables  $X$  e  $Y$ , **la organización actual de datos sufre nuevamente una **ordenación** simultánea con magnitud ascendente, tomando**

como pivote el **número de orden  $k$** , que en este caso, se refiere al tiempo en años en que se presentó un evento de gasto o volumen.

Ahora bien, se obtiene la diferencia de rangos de las variables  $X$  e  $Y$ , mejor dicho, se restan los rangos de las puntuaciones asociadas a cada variable, dicha diferencia es conocida como  $D_i$ . Conforme a las diferencias obtenidas  $D_i$ , es posible que resulten números negativos, para eliminarlos se procede a elevar al cuadrado todas las diferencias resultantes, el procedimiento continúa mediante el cálculo de la suma de todas las diferencias al cuadrado  $D_i^2$ . Inmediatamente después, se determina el conteo de las clasificaciones de empates vinculados a determinados conjuntos de empates asociados a las variables  $X$  e  $Y$ .

Por ejemplo, suponga que en la variable  $X$ , un conjunto de empates contiene la puntuación  $(a, a, a, a)$ , el conteo se entendería como; el tamaño de empates con puntuación  $a$  es de 4 elementos. Se aconseja construir una tabla que contenga todos los conjuntos de empates observados en ambas variables.

Según Myles Hollander *et al.*, (2014), la ecuación A.3 toma en cuenta el criterio de empates en un conjunto dado, conforme a este escenario el coeficiente de correlación de Spearman basado en rangos se calcula como:

$$\rho_s = \frac{n(n^2 - 1) - 6 \sum_{k=1}^n D_k^2 - \frac{1}{2} \left\{ \sum_{i=1}^g [t_i (t_i^2 - 1)] + \sum_{j=1}^h [u_j (u_j^2 - 1)] \right\}}{\sqrt{\left[ n(n^2 - 1) - \sum_{i=1}^g [t_i (t_i^2 - 1)] \right] \cdot \left[ n(n^2 - 1) - \sum_{j=1}^h [u_j (u_j^2 - 1)] \right]}}, \quad (\text{A.3})$$

donde

- $g$  número de conjuntos de vinculados a la variable  $X$
- $t_i$  tamaño de la clasificación de un conjunto empates de la variable  $X_i$
- $h$  número de conjuntos de vinculados a la variable  $Y$
- $u_j$  tamaño de la clasificación de un conjunto empates de la variable  $Y_j$
- $n$  número de observaciones.

La expresión A.3, es de gran utilidad cuando existen repeticiones de magnitud importante en las observaciones de una determinada muestra, por ejemplo cuando los datos son discretos, en el análisis hidrológico se puede emplear cuando se está trabajando con datos de gastos horarios, en este caso, es muy probable que se presenten mediciones de igual magnitud entre cada hora (*medidores ultrasónicos o tiempo en tránsito*).

*En el caso de que se presentasen algunas observaciones repetidas en una muestra (Gasto, Volumen, Precipitación), con precaución se puede suponer sin ningún problema, que no hay empates, dicho de otra manera, los empates se pueden evitar modificando ligeramente los valores repetidos.*

### A.3.3. Significado estadístico de la prueba

Sea  $H(X, Y)$  la función de distribución conjunta de la variable aleatoria bivariada  $(X, Y)$ ,  $F(X)$  y  $G(Y)$  sus funciones de distribución marginal, respectivamente.

*La hipótesis nula de interés es que  $X$  e  $Y$  son mutuamente independientes.*

$$H_0 = \{H(X, Y) = F(X) \cdot G(Y), \text{ para todo } (x, y) \in \mathbb{R}^2\}. \quad (\text{A.4})$$

*Las hipótesis alternativas de interés corresponde a los conceptos de correlación positiva y negativa entre  $X$  e  $Y$ .*

- i) Prueba unilateral (correlación positiva). Para probar la hipótesis nula de independencia entre  $X$  e  $Y$  con la hipótesis alternativa unilateral*

$$H_1 : \{X \text{ e } Y \text{ están correlacionados positivamente}\}, \quad (\text{A.5})$$

*es decir, los valores más grandes de  $X$  tienden a emparejarse con los valores más grandes de  $Y$ , y los valores más pequeños de  $X$  tienden a emparejarse con los valores más pequeños de  $Y$*



- ii) *Prueba unilateral (correlación negativa). Para probar la hipótesis nula de independencia entre X e Y con la hipótesis alternativa unilateral*

$$H_1 : \{X \text{ e } Y \text{ están correlacionados negativamente}\}, \quad (\text{A.6})$$

*es decir, los valores más grandes de X tienden a emparejarse con los valores más pequeños de Y, y los valores más pequeños de X tienden a emparejarse con los valores más grandes de Y.*

- iii) *Prueba bilateral. Para probar la hipótesis nula de independencia entre X e Y con la hipótesis alternativa bilateral*

$$H_1 : \{(X \text{ e } Y \text{ están correlacionadas positivamente}) \text{ o } (X \text{ e } Y \text{ están correlacionados negativamente})\} \quad (\text{A.7})$$

En el caso de trabajar con muestras pequeñas ( $n \leq 10$ ), para determinar la importancia que comprende el coeficiente de Rho de Spearman, se realiza una comparación entre el cuantil t obtenido con la ecuación A.8 con un valor crítico del Coeficiente de Correlación de Spearman basado en Rangos, dichos valores críticos son consultados en la tabla B de la sección de apéndices de Jared Linebach *et al.*, 2014.

Si se trabaja con una muestra de tamaño ( $n \geq 10$ ), se considera una muestra grande de acuerdo a la literatura, se calcula el valor de t mediante la ecuación A.8, donde t es la estadística t de Student, bajo la hipótesis nula de independencia, t es un cuantil aleatorio de la distribución t con  $df = n - 2$  grados de libertad. El valor obtenido t se compara con un valor crítico de la distribución t, los valores críticos se consultan en la tabla K de los apéndices de Jared Linebach *et al.*, 2014.

El cuantil aleatorio t de la distribución t con  $df = n - 2$  grados de libertad, se obtiene como

$$t = \rho_s \sqrt{\frac{n-2}{1-\rho_s^2}} \quad (\text{A.8})$$

donde

- $t$  es la estadística  $t$  de Student
- $\rho_s$  es el coeficiente de correlación de Spearman,
- $n$  es el número de observaciones bivariadas

Sí, el valor calculado de la estadística  $t$  es mayor que el valor crítico (tabla K de los apéndices de Jared Linebach *et al.*, (2014) se rechaza la hipótesis nula.

### A.3.4. Código en lenguaje R para calcular el Coeficiente de Correlación de Spearman.

El uso del Software R, para el cálculo del coeficiente de correlación de Spearman, es sumamente versátil y sencillo. En primera instancia, se debe de contar con la muestra bivariada capturada en un archivo de valores separado por comas con extensión .scv. A continuación se ejemplifica mediante un código en R para obtener la Rho de Spearman.

```
#####
#CÁLCULO DEL COEFICIENTE DE CORRELACIÓN RHO DE SPEARMAN
inpath <- 'C:\\NOMBRE DE LA CARPETA\\nombre del archivo.csv'
Alaux <- read.csv(inpath)
A1 <- data.matrix(Alaux)
n <- nrow(A1)
print("n")
print(n)
print("A1")
print(A1)
rho0 <- cor(A1,use="complete.obs",method="spearman")
rho0 <- rho[1,2]
print(rho)
#####
```

**Ejemplo 1 (Muestra con ausencia de empates):** Se cuenta con una muestra de gastos pico  $Q$  en,  $m^3/s$ , derivada de ella, se dedujo una nueva muestra referida a sus volúmenes de escurrimiento  $V$  en,  $Hm^3$ . Obtener el coeficiente de correlación de Spearman (por rangos) que asocie a las marginales  $Q$  y  $V$ . La información de los registros históricos de gastos pico, se recopiló del Banco Nacional de Datos de Aguas Superficiales (BANDAS), base de datos de la Comisión Nacional del Agua (2014), órgano desconcentrado del Gobierno Federal de México.

Los registros históricos provienen de la estación hidrométrica Las Perlas, con longitud de muestra de 57 años, ubicada en la cuenca del Río Coatzacoalcos en el Estado de Veracruz (*BANDAS, 2014*), vea tabla A.1.

**Tabla A.1.** Generalidades de la Estación Hidrométrica Las Perlas.

<i>Clave</i>	<i>Estación</i>	<i>Corriente</i>	<i>Cuenca</i>	<i>Estado</i>	<i>Coordenadas Geográficas</i>
29005	<i>Las Perlas</i>	<i>Río Coatzacoalcos</i>	<i>Río Coatzacoalcos</i>	<i>Veracruz</i>	17° 26' 17"N, -94° 52' 00"W

La muestra de gastos pico, con longitud de 57 años de mediciones, comprende un registro histórico de 1953 a 2014, y la correspondiente muestra de volúmenes de escurrimiento con idéntica longitud de 57 años, se exhiben en la tabla A.2. Primeramente se escanea visualmente a las muestras, y se identifica que no existan repeticiones en sus puntajes, esta observación previa permitirá al usuario decidir cuál metodología emplear de los dos escenarios que anteriormente se abordaron.

**Tabla A.2.** Registros Históricos de la Estación Hidrométrica Las Perlas.

<i>Tiempo</i> ( <i>años</i> )	$Q_p$ ( $m^3/s$ )	$V$ ( $Hm^3$ )	<i>Tiempo</i> ( <i>años</i> )	$Q_p$ ( $m^3/s$ )	$V$ ( $Hm^3$ )
1953	1947.374	4581.725	1987	1928.050	7430.401
1954	2181.895	5857.019	1988	2916.014	8913.535
1955	3098.750	6115.796	1989	4274.984	7113.683
1956	2200.750	7542.663	1990	1542.900	3398.817
1957	1738.249	5220.917	1991	4073.784	7127.026
1958	2406.000	7247.795	1992	2644.343	8705.360
1959	1815.312	5219.094	1993	2582.860	9383.771
1960	3439.749	8390.354	1994	958.511	3627.363
1961	4101.375	7332.841	1995	3649.163	9419.645
1962	2235.041	6651.430	1996	2168.700	6504.538
1963	6497.666	9504.804	1997	2586.347	7046.819
1964	2239.448	8055.928	1998	2198.073	5771.812
1965	2197.403	5597.068	1999	3353.591	8526.117
1966	3566.507	8313.473	2000	2258.943	10230.013
1967	2555.979	5094.033	2001	3122.859	7845.769
1968	6082.677	14061.219	2002	3389.663	7605.741
1969	5769.958	12055.792	2003	4038.793	10144.786
1970	4417.460	11217.620	2004	3516.746	12115.124
1971	5798.591	8372.399	2005	3638.544	14949.500
1972	3862.104	9225.439	2006	2845.812	11959.499
1973	5099.650	16356.012	2007	1261.411	3786.419
1975	2986.931	7298.026	2008	3842.252	7599.782
1976	2226.037	7156.266	2009	2470.312	5922.941
1978	6336.781	10730.080	2010	6399.943	12521.900
1981	2929.631	9253.425	2011	1098.361	5392.886
1982	4245.799	14272.462	2012	1228.784	5157.358
1983	3048.729	13621.620	2013	1172.337	5786.891
1984	3260.750	11251.216	2014	1065.888	5613.939
1985	4319.889	10437.004			

El análisis comienza por clasificar las  $n$  observaciones de las muestras  $Q$  y  $V$ , la clasificación se realiza en magnitud ascendente, Rangos  $R(Q_i)$  y  $R(V_i)$ , de modo que, se identifica la puntuación más baja y se etiqueta con el número "1", la segunda puntuación más baja se etiqueta como "2", y así sucesivamente, de manera similar y por separado, análogamente se etiquetan las puntuaciones de  $V$ , consulte la tabla A.3.

En seguida se obtiene la diferencia  $D_i$  entre los rangos  $R(Q_i)$  y  $R(V_i)$ , las diferencias resultantes se elevan al cuadrado ( $D_i^2$ ). Finalmente el coeficiente de correlación de Spearman es obtenido con la ecuación A.2.

$$\rho_s = 1 - \frac{6 \sum_{i=1}^n D_i^2}{n(n^2 - 1)} = 1 - \frac{6 \cdot (7292)^2}{57(57^2 - 1)} = 0.7636764324604615$$

Si se utiliza el Software R, el coeficiente de correlación de Spearman se estima con una aproximación de hasta siete decimales,  $\rho_s = 0.7636764$ .

**Ejemplo 2 (Muestra con vínculos de empates):** Se cuenta con una muestra de gastos pico  $Q$  en,  $m^3/s$ , derivada de ella, se dedujo una nueva muestra referida a sus volúmenes de escurrimiento  $V$  en,  $Hm^3$ . Obtener el coeficiente de correlación de Spearman (por rangos) que asocie a las marginales  $Q$  y  $V$ .

La información de los registros históricos de gastos pico, se recopiló del Banco Nacional de Datos de Aguas Superficiales (BANDAS), base de datos de la Comisión Nacional del Agua (2014), órgano desconcentrado del Gobierno Federal de México.

Los registros históricos que se analizarán provienen de la estación hidrométrica Minzapan, con longitud de muestra de 39 años, ubicada en la cuenca del Río Coatzacoalcos en el Estado de Veracruz (BANDAS, 2014), sus generalidades se resumen en la tabla A.4.

**Tabla A.3.** Rangos y diferencias de las muestras  $Q$  y  $V$

<i>Tiempo</i> (años)	$Q_p$ ( $m^3/s$ )	$V$ ( $Hm^3$ )	$R(Q_i)$	$R(V_i)$	$D_i$	$D_i^2$
1953	1947.374	4581.725	11	4	7	49
1954	2181.895	5857.019	13	14	-1	1
1955	3098.750	6115.796	32	16	16	256
1956	2200.750	7542.663	16	27	-11	121
1957	1738.249	5220.917	8	8	0	0
1958	2406.000	7247.795	21	23	-2	4
1959	1815.312	5219.094	9	7	2	4
1960	3439.749	8390.354	37	34	3	9
1961	4101.375	7332.841	46	25	21	441
1962	2235.041	6651.430	18	18	0	0
1963	6497.666	9504.804	57	42	15	225
1964	2239.448	8055.928	19	31	-12	144
1965	2197.403	5597.068	14	10	4	16
1966	3566.507	8313.473	39	32	7	49
1967	2555.979	5094.033	23	5	18	324
1968	6082.677	14061.219	54	54	0	0
1969	5769.958	12055.792	52	50	2	4
1970	4417.460	11217.620	50	47	3	9
1971	5798.591	8372.399	53	33	20	400
1972	3862.104	9225.439	43	38	5	25
1973	5099.650	16356.012	51	57	-6	36
1975	2986.931	7298.026	30	24	6	36
1976	2226.037	7156.266	17	22	-5	25
1978	6336.781	10730.080	55	46	9	81
1981	2929.631	9253.425	29	39	-10	100
1982	4245.799	14272.462	47	55	-8	64
1983	3048.729	13621.620	31	53	-22	484
1984	3260.750	11251.216	34	48	-14	196
1985	4319.889	10437.004	49	45	4	16
1987	1928.050	7430.401	10	26	-16	256
1988	2916.014	8913.535	28	37	-9	81
1989	4274.984	7113.683	48	20	28	784
1990	1542.900	3398.817	7	<b>1</b>	6	36
1991	4073.784	7127.026	45	21	24	576
1992	2644.343	8705.360	26	36	-10	100
1993	2582.860	9383.771	24	40	-16	256
1994	958.511	3627.363	<b>1</b>	<b>2</b>	-1	1
1995	3649.163	9419.645	41	41	0	0
1996	2168.700	6504.538	12	17	-5	25
1997	2586.347	7046.819	25	19	6	36
1998	2198.073	5771.812	15	12	3	9
1999	3353.591	8526.117	35	35	0	0
2000	2258.943	10230.013	20	44	-24	576
2001	3122.859	7845.769	33	30	3	9
2002	3389.663	7605.741	36	29	7	49
2003	4038.793	10144.786	44	43	1	1
2004	3516.746	12115.124	38	51	-13	169
2005	3638.544	14949.500	40	56	-16	256
2006	2845.812	11959.499	27	49	-22	484
2007	1261.411	3786.419	6	<b>3</b>	3	9
2008	3842.252	7599.782	42	28	14	196
2009	2470.312	5922.941	22	15	7	49
2010	6399.943	12521.900	56	52	4	16
2011	1098.361	5392.886	<b>3</b>	9	-6	36
2012	1228.784	5157.358	5	6	-1	1
2013	1172.337	5786.891	4	13	-9	81
2014	1065.888	5613.939	<b>2</b>	11	-9	81
$\Sigma$						<b>7292</b>

**Tabla A.4.** Generalidades de la Estación Hidrométrica Minzapan.

<i>Clave</i>	<i>Estación</i>	<i>Corriente</i>	<i>Cuenca</i>	<i>Estado</i>	<i>Coordenadas Geográficas</i>
29009	Minzapan	Río Huazuntlán	Río Coatzacoalcos	Veracruz	18° 09' 36"N, -94° 47' 20"W

La muestra de gastos pico, con longitud de 39 años de mediciones, comprende un registro histórico de 1974 a 2014, y la correspondiente muestra de volúmenes de escurrimiento con idéntica longitud, se exhiben en la tabla A.5. Primeramente se escanea visualmente a las muestras, y se identifica que si existen repeticiones en sus puntajes, esta observación previa conduce a emplear la metodología del segundo escenario, el caso donde las algunas observaciones se repiten.

El análisis comienza por *clasificar* las  $n$  observaciones de las muestras  $Q$  y  $V$ , la clasificación se realiza por el tamaño de la observación y en magnitud ascendente, (columna "Ranking"), de modo que, se identifica la puntuación más baja y se etiqueta con el número "1", la segunda puntuación más baja se etiqueta como "2", y así sucesivamente, de manera similar y por separado, análogamente se etiquetan las puntuaciones de  $V$ , consulte las tablas A.6 y A.7.

Enseguida se *ordenan* las  $n$  observaciones de las muestras  $Q$  y  $V$  por separado en magnitud ascendente (menor a mayor), al tomar como pivote los puntajes en las muestras  $Q$  y  $V$ , simultáneamente también se modifican las clasificaciones del Ranking en forma ascendente además del orden  $k$ , que en este caso, se refiere al tiempo en años, acorde a esta nueva organización de datos, el Ranking más bajo es 1, mientras que el Ranking más alto es 39, apóyese en las tablas A.8 y A.9.

En este momento es factible obtener los rangos, para ello se identifica a  $R(Q_i)$  como el rango de  $Q_i$ , de manera similar,  $R(V_i)$  denota el rango  $V_i$

Conforme al último arreglo de información, los puntajes de cada muestra  $Q$  y  $V$  se han ordenado por categorías debido al surgimiento de empates, lo que ha originado que se conformen conjuntos de datos con igual puntuación, vea las tablas A.8 y A.9.

**Tabla A.5.** Registros Históricos de la Estación Hidrométrica Las Perlas.

<i>Tiempo</i> (años)	$Q_p$ ( $m^3/s$ )	$V$ ( $Hm^3$ )
1974	164.66	61.72
1975	119.64	76.03
1976	336.93	57.39
1977	94.21	14.84
1978	156.56	39.99
1979	91.19	28.78
1980	309.79	83.33
1981	197.51	39.36
1982	206.14	61.72
1983	195.03	38.27
1986	175.55	46.71
1987	178.91	51.65
1988	91.19	35.17
1989	64.64	16.40
1990	63.02	18.87
1991	120.85	64.33
1992	131.17	27.69
1993	152.58	65.44
1994	73.30	18.23
1995	92.77	12.64
1996	172.44	51.65
1997	236.15	41.32
1998	87.74	27.69
1999	119.64	54.66
2000	82.24	26.23
2001	126.19	35.93
2002	80.42	14.84
2003	180.76	45.89
2004	52.74	13.35
2005	186.56	47.73
2006	141.90	26.23
2007	283.42	48.08
2008	99.01	25.71
2009	116.80	37.46
2010	145.52	50.60
2011	109.51	20.50
2012	129.29	31.23
2013	141.90	31.23
2014	70.54	37.46

**Tabla A.6.** Clasificación por tamaño de  $Q_p$  ascendentemente

<i>Tiempo</i> (años)	$Q_p$ ( $m^3/s$ )	<i>Ranking</i> $Q_p$
1974	164.66	27
1975	119.64	16
1976	336.93	39
1977	94.21	12
1978	156.56	26
1979	91.19	9
1980	309.79	38
1981	197.51	34
1982	206.14	35
1983	195.03	33
1986	175.55	29
1987	178.91	30
1988	91.19	10
1989	64.64	3
1990	63.02	<b>2</b>
1991	120.85	18
1992	131.17	21
1993	152.58	25
1994	73.30	5
1995	92.77	11
1996	172.44	28
1997	236.15	36
1998	87.74	8
1999	119.64	17
2000	82.24	7
2001	126.19	19
2002	80.42	6
2003	180.76	31
2004	52.74	<b>1</b>
2005	186.56	32
2006	141.90	22
2007	283.42	37
2008	99.01	13
2009	116.80	15
2010	145.52	24
2011	109.51	14
2012	129.29	20
2013	141.90	23
2014	70.54	4

**Tabla A.7.** Clasificación por tamaño de  $V$  ascendentemente

<i>Tiempo</i> (años)	$V$ ( $Hm^3$ )	<i>Ranking</i> $V$
1974	61.72	34
1975	76.03	38
1976	57.39	33
1977	14.84	3
1978	39.99	23
1979	28.78	14
1980	83.33	39
1981	39.36	22
1982	61.72	35
1983	38.27	21
1986	46.71	26
1987	51.65	30
1988	35.17	17
1989	16.40	5
1990	18.87	7
1991	64.33	36
1992	27.69	12
1993	65.44	37
1994	18.23	6
1995	12.64	<b>1</b>
1996	51.65	31
1997	41.32	24
1998	27.69	13
1999	54.66	32
2000	26.23	10
2001	35.93	18
2002	14.84	4
2003	45.89	25
2004	13.35	<b>2</b>
2005	47.73	27
2006	26.23	11
2007	48.08	28
2008	25.71	9
2009	37.46	19
2010	50.60	29
2011	20.50	8
2012	31.23	15
2013	31.23	16
2014	37.46	20

Las observaciones de las muestras  $Q$  y  $V$  presentan puntuaciones con empates (tablas A.8 y A.9), para obtener los rangos asociados a dichas muestras y que previamente sus puntuaciones fueron ordenadas por categorías, a todos los conjuntos de empates se les asigna un mismo rango promedio calculado en función de sus posiciones del Ranking.

**Tabla A.8.** Empates y Rangos  $Q_p$

<i>Tiempo</i> (años)	$Q_p$ ( $m^3/s$ )	<i>Ranking</i> $Q_p$	$R(Q_p)$
2004	52.74	1	1
1990	63.02	2	2
1989	64.64	3	3
2014	70.54	4	4
1994	73.30	5	5
2002	80.42	6	6
2000	82.24	7	7
1998	87.74	8	8
1979	<b>91.19</b>	9	<b>9.5</b>
1988	<b>91.19</b>	10	<b>9.5</b>
1995	92.77	11	11
1977	94.21	12	12
2008	99.01	13	13
2011	109.51	14	14
2009	116.80	15	15
1975	119.64	16	16.5
1999	119.64	17	16.5
1991	120.85	18	18
2001	126.19	19	19
2012	129.29	20	20
1992	131.17	21	21
2006	141.90	22	22.5
2013	141.90	23	22.5
2010	145.52	24	24
1993	152.58	25	25
1978	156.56	26	26
1974	164.66	27	27
1996	172.44	28	28
1986	175.55	29	29
1987	178.91	30	30
2003	180.76	31	31
2005	186.56	32	32
1983	195.03	33	33
1981	197.51	34	34
1982	206.14	35	35
1997	236.15	36	36
2007	283.42	37	37
1980	309.79	38	38
1976	336.93	39	39

**Tabla A.9.** Empates y Rangos  $V$

<i>Tiempo</i> (años)	$V$ ( $Hm^3$ )	<i>Ranking</i> $V$	$R(V)$
1995	12.64	1	1
2004	13.35	2	2
1977	14.84	3	3.5
2002	14.84	4	3.5
1989	16.40	5	5
1994	18.23	6	6
1990	18.87	7	7
2011	20.50	8	8
2008	25.71	9	9
2000	26.23	10	10.5
2006	26.23	11	10.5
1992	27.69	12	12.5
1998	27.69	13	12.5
1979	28.78	14	14
2012	31.23	15	15.5
2013	31.23	16	15.5
1988	35.17	17	17
2001	35.93	18	18
2009	37.46	19	19.5
2014	37.46	20	19.5
1983	38.27	21	21
1981	39.36	22	22
1978	39.99	23	23
1997	41.32	24	24
2003	45.89	25	25
1986	46.71	26	26
2005	47.73	27	27
2007	48.08	28	28
2010	50.60	29	29
1987	51.65	30	30.5
1996	51.65	31	30.5
1999	54.66	32	32
1976	57.39	33	33
1974	61.72	34	34.5
1982	61.72	35	34.5
1991	64.33	36	36
1993	65.44	37	37
1975	76.03	38	38
1980	83.33	39	39

Por ejemplo, conforme a la tabla A.8, los Rankings (9,10) comparten la misma puntuación (91.19, 91.19), entonces se necesita promediar las posiciones de los Rankings (9,10) asignados a las puntuaciones (91.19, 91.19), por lo tanto, los Rankings noveno y décimo comparten un rango de 9.5 asociado a las puntuaciones de la muestras  $Q$ .



El siguiente rango asignado a la muestras  $Q$ , en principio de cuentas, tendría un rango de 11, a menos que existiesen más conjuntos de empates posteriores a la décima posición del Ranking.

En seguida de haber obtenido los rangos promedio, por separado, de las muestras  $Q$  y  $V$ , la organización actual de datos sufre nuevamente una ordenación simultánea con magnitud ascendente, tomando como pivote el número de orden  $k$  (*tiempo en años*) que originalmente clasifica a las muestras  $Q$  y  $V$ , consulte la tabla A.10.

Ahora bien, se obtiene la diferencia de rangos entre las muestras  $Q$  y  $V$  denominada como  $D_i$ . Conforme a las diferencias obtenidas, se procede a elevar al cuadrado todas las diferencias resultantes, el procedimiento continúa mediante el cálculo de la suma total de las diferencias al cuadrado  $D_i^2$ , como en la tabla A.10.

Inmediatamente después, se determina el *conteo de las clasificaciones de empates* vinculados a determinados conjuntos de empates asociados a las muestras  $Q$  y  $V$ , los resultados se exhiben en la tabla A.11.

Por ejemplo, conforme a la muestra  $V$  de la tabla A.10, un conjunto de empates contiene la puntuación (27.69, 27.69), el conteo se entendería como; en **2 años** del registro histórico (1974-2014) se presentó un volumen de escurrimiento de 27.69 Hm<sup>3</sup>, vea la tabla A.11.

La ecuación A.3, toma en cuenta el criterio de empates en un conjunto dado por las muestras  $Q$  y  $V$ , conforme a este escenario, el coeficiente de correlación de Spearman basado en rangos se calcula como:

$$\rho_s = \frac{n(n^2 - 1) - 6 \sum_{k=1}^n D_k^2 - \frac{1}{2} \left\{ \sum_{i=1}^g [t_i(t_i^2 - 1)] + \sum_{j=1}^h [u_j(u_j^2 - 1)] \right\}}{\sqrt{\left[ n(n^2 - 1) - \sum_{i=1}^g [t_i(t_i^2 - 1)] \right] \cdot \left[ n(n^2 - 1) - \sum_{j=1}^h [u_j(u_j^2 - 1)] \right]}}$$

**Tabla A.10.** Estadística de Rangos y Diferencias al cuadrado

<i>Tiempo</i> (años)	$Q_p$ ( $m^3/s$ )	$V$ ( $Hm^3$ )	<i>Ranking</i> $Q_p$	<i>Ranking</i> $V$	$R(Q_p i)$	$R(V i)$	$D_i$	$D_i^2$
1974	164.66	61.72	27	34	27	34.5	-7.5	56.25
1975	119.64	76.03	16	38	16.5	38	-21.5	462.25
1976	336.93	57.39	39	33	39	33	6	36
1977	94.21	14.84	12	3	12	3.5	8.5	72.25
1978	156.56	39.99	26	23	26	23	3	9
1979	91.19	28.78	9	14	9.5	14	-4.5	20.25
1980	309.79	83.33	38	39	38	39	-1	1
1981	197.51	39.36	34	22	34	22	12	144
1982	206.14	61.72	35	35	35	34.5	0.5	0.25
1983	195.03	38.27	33	21	33	21	12	144
1986	175.55	46.71	29	26	29	26	3	9
1987	178.91	51.65	30	30	30	30.5	-0.5	0.25
1988	91.19	35.17	10	17	9.5	17	-7.5	56.25
1989	64.64	16.40	3	5	3	5	-2	4
1990	63.02	18.87	2	7	2	7	-5	25
1991	120.85	64.33	18	36	18	36	-18	324
1992	131.17	27.69	21	12	21	12.5	8.5	72.25
1993	152.58	65.44	25	37	25	37	-12	144
1994	73.30	18.23	5	6	5	6	-1	1
1995	92.77	12.64	11	1	11	1	10	100
1996	172.44	51.65	28	31	28	30.5	-2.5	6.25
1997	236.15	41.32	36	24	36	24	12	144
1998	87.74	27.69	8	13	8	12.5	-4.5	20.25
1999	119.64	54.66	17	32	16.5	32	-15.5	240.25
2000	82.24	26.23	7	10	7	10.5	-3.5	12.25
2001	126.19	35.93	19	18	19	18	1	1
2002	80.42	14.84	6	4	6	3.5	2.5	6.25
2003	180.76	45.89	31	25	31	25	6	36
2004	52.74	13.35	1	2	1	2	-1	1
2005	186.56	47.73	32	27	32	27	5	25
2006	141.90	26.23	22	11	22.5	10.5	12	144
2007	283.42	48.08	37	28	37	28	9	81
2008	99.01	25.71	13	9	13	9	4	16
2009	116.80	37.46	15	19	15	19.5	-4.5	20.25
2010	145.52	50.60	24	29	24	29	-5	25
2011	109.51	20.50	14	8	14	8	6	36
2012	129.29	31.23	20	15	20	15.5	4.5	20.25
2013	141.90	31.23	23	16	22.5	15.5	7	49
2014	70.54	37.46	4	20	4	19.5	-15.5	240.25
$\Sigma D_i^2$								<b>2805</b>

**Tabla A.11.** Conteo de observaciones con empates en un conjunto dado

<i>Conjunto</i>	$Q_p$ ( $m^3/s$ )			$V$ ( $Hm^3$ )						
	I	II	III	I	II	III	IV	V	VI	VII
<i>Empates</i> (muestras)	91.19	119.64	141.90	14.84	26.23	27.69	31.23	37.46	51.65	61.72
<i>Conteo <math>t_i</math></i> ( $t$ , años)	2	2	2	2	2	2	2	2	2	2

Se comienza por obtener los tamaños totales de la clasificación de empates para un conjunto dado de las muestras  $Q$  y  $V$ ,

para la muestra  $Q$ , con  $g = 3$

$$T_Q = \sum_{i=1}^g [t_i(t_i^2 - 1)]$$

$$T_Q = (2^3 - 2) + (2^3 - 2) + (2^3 - 2)$$

$$T_Q = 18$$

para la muestra  $V$ , con  $h = 7$

$$T_V = \sum_{j=1}^h [u_j(u_j^2 - 1)]$$

$$T_V = (2^3 - 2) + (2^3 - 2) + (2^3 - 2) + (2^3 - 2) + (2^3 - 2) + (2^3 - 2) + (2^3 - 2)$$

$$T_V = 42$$

y la longitud de las muestras es 39

$$n(n^2 - 1) = 39(39^2 - 1) = 59280$$

Sustituyendo las variables  $T_Q$ ,  $T_V$  y  $n(n^2 - 1)$  en la ecuación A.3

$$\rho_s = \frac{59280 - (6 \cdot 2805) - \frac{(18 + 42)}{2}}{\sqrt{[59280 - 18] \cdot [59280 - 42]}}$$

$$\rho_s = 0.715949381772396$$

El software R, ofrece un excelente resultado,  $\rho_s = 0.7159494$ .

Haciendo un balance de los dos escenarios para obtener el coeficiente de correlación de Spearman basado en rangos, es muy poco probable que en

variables hidrológicas como gasto o volumen existan cantidades repetidas, conforme a ello, el primer escenario es el que generalmente se aplica para obtener tal coeficiente. En este apartado se presentó un ejemplo con el segundo escenario, habiendo empates o repeticiones en las observaciones de una muestra bivariada, como se mostró, es bastante engorroso obtenerlo, por lo que en proyectos de ingeniería recomendaría utilizar programación mediante lenguaje R para obtener éste y otros procedimientos estadísticos para el análisis de frecuencias bivariado mediante cópulas.



# Apéndice B

## Métodos gráficos para detectar dependencia

### B.1. Introducción

Las herramientas gráficas que se presentan en esta sección, ayudan a realizar un primer estudio sobre la posible dependencia funcional de dos variables aleatorias. Estas herramientas para detectar dependencia son recientes en la literatura y se denominan **Chi-plot** y **K-plot**. Estas herramientas también se pueden aplicar para observar la dependencia que existe entre diferentes modelos de Funciones Cópula.

### B.2. Chi-plot

El gráfico Chi-Plot fue propuesto inicialmente por Fisher (1985). Su construcción está basada en el estadístico Chi-cuadrado para independencia. Sea  $(X_1, Y_1), \dots, (X_n, Y_n)$  una muestra aleatoria bivariada de una función de distribución conjunta y continua,  $H(X, Y)$ , y sea  $I(A)$  la función indicadora del evento A. Para cada observación  $(x_i, y_i)$  se desarrolla el siguiente procedimiento:

$$H_i = \frac{1}{n-1} \sum_{j \neq i} I(X_j \leq X_i, Y_j \leq Y_i) \quad (\text{B.1})$$

$$F_i = \frac{1}{n-1} \sum_{j \neq i} I(X_j \leq X_i)$$

$$G_i = \frac{1}{n-1} \sum_{j \neq i} I(Y_j \leq Y_i)$$

Ninguna de estas cantidades dependen exclusivamente de los rangos de las observaciones, Fisher (1985) propone graficar los pares  $(\lambda_i, \chi_i)$ , donde:

$$\lambda_i = 4 \text{sign}(\tilde{F}_i \tilde{G}_i) \text{máx}(\tilde{F}_i^2, \tilde{G}_i^2),$$

$$\chi_i = \frac{H_i - F_i G_i}{\sqrt{F_i(1-F_i)G_i(1-G_i)}},$$

donde

$\tilde{F}_i = F_i - 1/2$ ,  $\tilde{G}_i = G_i - 1/2$  para  $i \in \{1, \dots, n\}$ ,  $\lambda_i$  es una medida de distancia de la observación  $(X_i, Y_i)$  al centro de los datos. Todos los valores de  $\lambda_i$  deben estar en el intervalo  $[-1, 1]$ .

De modo que, el Chi-plot es un gráfico de dispersión de los pares  $(\lambda_i, \chi_i)$ , para todo  $i = 1, \dots, n$ . Al observar el gráfico Chi-plot, si los datos constituyen una muestra bivariada con marginales continuas independientes, los valores de  $\lambda_i$  estarán distribuidos uniformemente en el gráfico. Sin embargo, si X e Y están asociados, los valores de  $\lambda_i$  se mostrarán formando grupos; en particular, valores positivos de  $\lambda_i$  indican que  $X_i$  e  $Y_i$  son relativamente grandes o pequeños (al mismo tiempo) en relación a sus medianas, mientras que si los valores de  $\lambda_i$  son negativos, corresponde a que  $X_i$  e  $Y_i$  están ubicados en lados opuestos con respecto a sus medianas.

Las líneas horizontales del gráfico están dadas por  $\chi = -c_p/n^{1/2}$  y  $\chi = c_p/n^{1/2}$ , donde  $c_p$  se selecciona de manera que aproximadamente el 100% de los pares  $(\lambda_i, \chi_i)$  estén entre las dos líneas horizontales. Para  $p = 0.90, 0.95, 0.99$  los valores de  $c_p$  son 1.54, 1.78 y 2.18 respectivamente [Fisher (2001)]. Usando el método Monte Carlo se pueden calcular otros valores. También se recomienda representar solamente los pares para los

que  $|\lambda_i| < 4 \left( \frac{1}{n-1} - \frac{1}{2} \right)^2$ , con el fin de evitar observaciones engañosas. Con el propósito de mostrar un ejemplo de este tipo de gráfico, Guarín *et al.* (2018), hacen un análisis por métodos gráficos para detectar dependencia. En la Figura B.1 (b) se muestra un gráfico Chi-plot, para una muestra de 200 datos, y se logra detectar la dependencia entre las variables u1 y u2, pues la mayoría de los puntos caen por fuera de las bandas, asimismo la concentración de puntos se ubica en el primer cuadrante del gráfico.

Genest y Boies (2007), introdujeron la siguiente expresión:

$$H_i = \frac{1}{n-1} \# \{j \neq i : X_j \leq X_i, Y_j \leq Y\} = \frac{n W_i^* - 1}{n-1}, \quad (\text{B.2})$$

donde  $W_i^*$  es igual a la cópula empírica  $C_n$

$$W_i^* = C_n \left( \frac{R_i}{n+1}, \frac{S_i}{n+1} \right). \quad (\text{B.3})$$

En la expresión anterior, se observa que los eventos de las marginales u1 y u2 siguen una probabilidad de no excedencia con clasificación ascendente en sus eventos de u1 y u2, o sea  $k_i/(n+1) = 1 - (1/T)$ , de otra manera, u1 y u2 pueden verse como la fórmula de Ley empírica de Weibull con clasificación ascendente en sus eventos.

### B.3. K-plot

El K-plot (o Kendall-plot) fue creado por Genest y Boies (2003). Esta herramienta se construye sobre los rangos de las observaciones usando la transformación integral de probabilidades multivariadas, produciendo un gráfico similar al Q-Q plot convencional. Sea  $(X_1, Y_1), \dots, (X_n, Y_n)$  una muestra aleatoria de una función de distribución conjunta y continua  $H(X, Y)$ . Para construir el gráfico K-Plot se procede de la siguiente manera:

1. Para cada  $1 \leq i \leq n$  calcule  $H_i$  (como en el gráfico Chi-plot).
2. Ordene los valores  $H_i$ , de manera que  $H_{(1)} \leq \dots \leq H_{(n)}$ .
3. Grafique los pares  $(W_{i:n}, H_{(i)})$ , donde  $W_{i:n}$  es la esperanza del



$i$ -ésimo estadístico de orden en una muestra de tamaño  $n$ , las cuales se calculan como:

$$W_{i:n} = n \binom{n-1}{i-1} \int_0^1 w \cdot k_0(w) \cdot [K_0(w)]^{i-1} \cdot [1 - K_0(w)]^{n-i} dw \quad (\text{B.4})$$

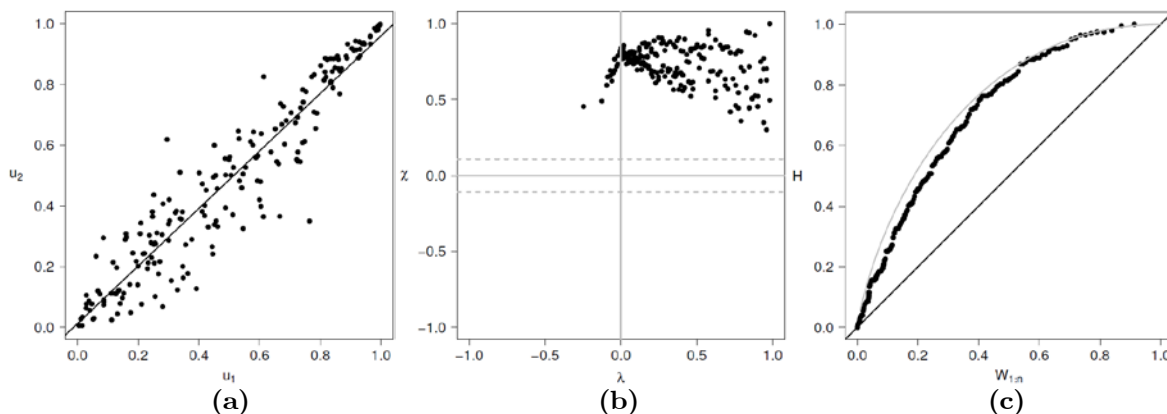
donde

$$k_0(w) = -\ln(w),$$

$$K_0(w) = w - w \ln(w) \quad 0 \leq w \leq 1, \text{ (línea curva),}$$

$n \binom{n-1}{i-1} = \frac{n!}{(i-1)! (n-1)!} = C_{n-1}^{i-1}$ , es el número de combinaciones de los  $(i - 1)$  elementos elegidos entre las  $(n - 1)$  observaciones.

Finalmente se realiza un gráfico de puntos de  $H_i$  contra  $W_{i:n}$ , y a medida que los puntos se alejan de la diagonal, se asume dependencia funcional entre las dos variables aleatorias. En la Figura B.1 (c) se muestra un gráfico K-plot, para la misma muestra de 200 datos, y se logra detectar la dependencia entre las variables  $u_1$  y  $u_2$ .



**Figura B.1.** Gráfico de dispersión (a), Chi-plot (b) y K-plot (c), para un tamaño de muestra  $n = 200$ .

# Apéndice C

## Determinación de la distancia supremo $D$ , rho de Spearman $\rho_s$ y parámetro de asociación $\theta$ en lenguaje R

```
#####  
Este programa esta codificado en lenguaje R, su operación consiste en lo  
siguiente:
```

- 1.- Ingresar la muestra bivariada (modificada en rangos [0-1]) en un archivo de valores separado por comas con extensión .scv.
- 2.- Ingresar el tamaño de la muestra  $n$  (primera entrada).
- 3.- Ingresar el tamaño del Grid,  $m \geq 2$  (segunda entrada).
- 4.- Conforme a la estructura y nomenclatura de programación, ingresar el número de modelo que se requiera probar (quinta entrada).
- 5.- Las demás entradas se rellenan con el número 1.

Ejemplo:

```
#system.time(TestCmvsC(sample.size,order.part,itors,itors2,model0,model1,rho0,  
rho1,model3))  
#system.time(TestCmvsC(57,3,1,1,3,1,1,1,1))
```

Resultados esperados:

Al término del proceso del programa, los resultados que se arrojan son:

- 1.- Impresión de la muestra transformada [0-1], correspondiente a los datos observados.

- 
- 2.- Impresión del parámetro Rho de Spearman, conforme a la naturaleza de la muestra bivariada.
  - 3.- Impresión del parámetro  $\theta$  del modelo en prueba.
  - 4.- Matriz de conteos  $S$ , para cada una de las  $d$ -cajas.
  - 5.- Matriz de la cópula  $C_m$  obtenida con la aproximación checkerboard.
  - 6.- Matriz cuyo arreglo corresponde a la distancia  $D$ , entre la cópula muestral  $C_m$  y la propuesta  $C$  (check).
  - 7.- La distancia de variación total  $D$ .

```
#####
```

```
TestCmvsC<-
function(sample.size,order.part,itors,itors2,model0,model1,rho0,rho1,model3)

{
d<-2
n<-sample.size
m<-order.part

realC0 <- array(rep(0,times=m^2),dim=c(m,m))
auxC0 <- array(rep(0,times=(m+1)^2), dim=c(m+1,m+1))
checkC0 <- array(rep(0,times=m^2),dim=c(m,m))

inpath <- 'C:\\RWORK\\minatitlan33.csv'
Alaux <- read.csv(inpath)
A1 <- data.matrix(Alaux)
n <- nrow(A1)
print("n")
print(n)
print("A1")
print(A1)
plot(A1[,1],A1[,2])

rho0 <- cor(A1,use="complete.obs",method="spearman")
rho0 <- rho0[1,2]
print("RHO DE SPEARMAN")
print(rho0)

# WE OBTAIN THE SAMPLE COPULA OF ORDER $m$ AND THE TOTAL VARIATION DISTANCE
FOR $C_{m}^{n}$ AND $CheckC0$

AA<-A1[do.call(order, lapply(1:NCOL(A1), function(i) A1[,i])), ]
B=c(AA,1:n)
dim(B)=c(n,3)
#print(B)

BB<-B[do.call(order, lapply(2:NCOL(B), function(i) B[,i])), ]
C=c(BB,1:n)
dim(C)=c(n,4)
```

## Distancia supremo $D$ , $\rho_s$ de Spearman y parámetro $\theta$ de las familias cópula

---

```
#print(C)

#CC<-C[do.call(order, lapply(3:NCOL(C), function(i) C[,i])), ]
#D=c(CC,1:n)
#dim(D)=c(n,6)
#print(D)

Z=C[,3:4]/n
#print(Z)
#return(Z)

u1<-Z[,1]
v1<-Z[,2]
Z<-c(u1,v1)
dim(Z)<-c(n,2)
#print(Z)

# for (m in 2:n){ #OJO

s=rep(0,times=m^2)

# s= matrix of frequencies for the sample 2-copula of order m

dim(s)<-c(m,m)

for (i in 1:n)
{
  for(j in 1:m)
  {
    for(k in 1:m)
    {
      if( u1[i]>(j-1)/m & u1[i]<=j/m & v1[i]>(k-1)/m & v1[i]<=k/m)
        s[j,k]=s[j,k]+1
    }
  }
}

print("s enteros => CONTEOS EN CADA CAJA")
print(s)

s <- s/n

print("s")
print(s)
print("SUMA s")
print(sum(s))

# PART one          WE FIND THE CHECKERBOARD COPULAS FOR BOTH DISTRIB. UNDER
# $rho0$ AND $rho1$,
#                   WE ALSO GIVE THE THEORETICAL TOTAL VARIATION DISTANCE
```

---

```
# WE FIND THE CORRESPONDING CHECKERBOARD COPULAS OF ORDER $m$ FOR MODEL 0 AND
MODEL 1 USING THE
# VALUES OF $rho0$ and $rho1$ AND EVALUATE THE TOTAL VARIATION DISTANCE
BETWEEN THEM.
```

```
if(model0==1){ #Clayton
theta0 <- iRho(claytonCopula(),rho0)
for(ij in 1:m){
for(jk in 1:m){
f <- (ij/m)^(-theta0)+(jk/m)^(-theta0)-1
if(f>=0){
g <- max(f^(-1/theta0),0)}
if(f<0){g<-0}
realC0[ij,jk] <- g}}
#print("real checkerboard copula of order m")
#print(realC0)
}

if(model0==3){ #Gumbel
if(rho0<0){print("Gumbel model does not apply")}
stop()}
theta0 <- iRho(gumbelCopula(),rho0)
for(ij in 1:m){
for(jk in 1:m){
f <- ((-log(ij/m))^(theta0)+(-log(jk/m))^(theta0))^(1/theta0)
g <- exp(-f)
realC0[ij,jk] <- g}}
#print("real checkerboard copula of order m")
#print(realC0)
}

if(model0==4){ #Frank
theta0 <- iRho(francCopula(),rho0)
for(ij in 1:m){
for(jk in 1:m){
e <- (exp(-theta0*(ij/m))-1)*(exp(-theta0*(jk/m))-1)
f <- 1+ (e/(exp(-theta0)-1))
g <- -(1/theta0)*log(f)
realC0[ij,jk] <- g}}
#print("real checkerboard copula of order m")
#print(realC0)
}

if(model0==5){ #Joe
if(rho0 >= 0 && rho0< 0.03){theta0 <- 1.012}
if(rho0 >=0.03 && rho0< 0.08){theta0 <- 1.0597}
if(rho0 >=0.08 && rho0< 0.2){theta0 <- 1.13}
if(rho0 >=0.2 && rho0< 0.4){theta0 <- 1.36}
if(rho0 >=0.4 && rho0< 0.65){theta0 <- 1.99}
if(rho0 >=0.65 && rho0< 0.85){theta0 <- 3.45}
```

```

if(rho0 >=0.85 && rho0< 0.93){theta0 <- 6.4}
if(rho0 >=0.93 && rho0< 0.98){theta0 <- 9.6}
if(rho0 >=0.98 && rho0<= 1){theta0 <- 23}
if(rho0<0){print("Joe Model does not apply")}
for(ij in 1:m){
for(jk in 1:m){
f <- (1-(ij/m))^(theta0)+(1-(jk/m))^(theta0)- ((1-(ij/m))^(theta0))*((1-
(jk/m))^(theta0))
g <- 1 - (f^(1/theta0))
realC0[ij,jk] <- g}}
#print("real checkerboard copula of order m")
#print(realC0)
}

```

```

if(model0==6){      #Plackett
if(rho0< -0.98 || rho0== -0.98){theta0 <- 0.0008168}
if(rho0 >-0.98 && rho0<= -0.93){theta0 <- 0.005990}
if(rho0 >-0.93 && rho0<= -0.85){theta0 <- 0.015125}
if(rho0 >-0.85 && rho0<= -0.65){theta0 <- 0.0583448}
if(rho0 >-0.65 && rho0<= -0.4){theta0 <- 0.195478}
if(rho0 >-0.4 && rho0<= -0.2){theta0 <- 0.465523}
if(rho0 >-0.2 && rho0<= -0.08){theta0 <- 0.740148}
if(rho0 >-0.08 && rho0<= -0.03){theta0 <- 0.860611}
if(rho0 >-0.03 && rho0<= 0){theta0 <- 0.970445}
if(rho0 >0 && rho0<= 0.03){theta0 <- 1.03046}
if(rho0 >0.03 && rho0<= 0.08){theta0 <- 1.16197}
if(rho0 >0.08 && rho0<= 0.2){theta0 <- 1.35108}
if(rho0 >0.2 && rho0<= 0.4){theta0 <- 2.14812}
if(rho0 >0.4 && rho0<= 0.65){theta0 <- 5.11566}
if(rho0 >0.65 && rho0<= 0.85){theta0 <- 17.1395}
if(rho0 >0.85 && rho0<= 0.93){theta0 <- 66.1154}
if(rho0 >0.93 && rho0<= 0.98){theta0 <- 166.939}
if(rho0 >0.98 && rho0<= 1){theta0 <- 1224.17}
for(ij in 1:m){
for(jk in 1:m){
e <- (1+(theta0-1)*((ij+jk)/m))^2 -4*(ij/m)*(jk/m)*theta0*(theta0-1)
f <- (1+(theta0-1)*((ij+jk)/m))-e^(1/2)
g <- f/(2*(theta0-1))
realC0[ij,jk] <- g}}
#print("real checkerboard copula of order m")
#print(realC0)
}

```

```

if(model0==9){      #T-Student df=4
#theta0 <- iRho(tCopula(df=4),rho0)
theta0 <- rho0
vvv <- tCopula(theta0, df=4)
for(ij in 1:m){
for(jk in 1:m){
realC0[ij,jk] <- pCopula(c(ij/m,jk/m),vvv)}}

```

---

```

#print("real checkerboard copula of order m")
#print(realC0)
}

if(model0==10){      #Normal dim=2
theta0 <- iRho(normalCopula(),rho0)
vvv <- normalCopula(theta0, dim=2)
for(ij in 1:m){
for(jk in 1:m){
realC0[ij,jk] <- pCopula(c(ij/m,jk/m),vvv)}}
#print("real checkerboard copula of order m")
#print(realC0)
}

if(model0==22){      #Gumbel Decreasing-Decreasing
if(rho0>=0){theta0 <- iRho(gumbelCopula(),rho0)}
if(rho0<0){print("The GumbelDD does not apply")}
for(ij in 1:m){
for(jk in 1:m){
f <- ((-log(1-ij/m))^(theta0)+(-log(1-jk/m))^(theta0))^(1/theta0)
g <- exp(-f)
realC0[ij,jk] <- (ij/m)+(jk/m)-1+g}}
#print("real checkerboard copula of order m")
#print(realC0)
#print("theta Gumbel D-D")
#print(theta0)
}

if(model0==26){      #Galambos
if(rho0<0){print("Galambos model does not apply")
stop()}
theta0 <- iRho(galambosCopula(),rho0)
for(ij in 1:m){
for(jk in 1:m){
f <- ((-log(ij/m))^(-theta0)+(-log(jk/m))^(-theta0))^(-1/theta0)
g <- (ij/m)*(jk/m)*exp(f)
realC0[ij,jk] <- g}}
#print("real checkerboard copula of order m")
#print(realC0)
}

if(model0==27){      #Galambos Decreasing-Decreasing
if(rho0>=0){theta0 <- iRho(galambosCopula(),rho0)}
if(rho0<0){print("The Galambos DD model does not apply")}
#stop()}
theta0 <- iRho(galambosCopula(),rho0)
for(ij in 1:m){
for(jk in 1:m){
f <- ((-log(1-ij/m))^(-theta0)+(-log(1-jk/m))^(-theta0))^(-1/theta0)
g <- (1-ij/m)*(1-jk/m)*exp(f)

```

---

## Distancia supremo $D$ , $\rho_s$ de Spearman y parámetro $\theta$ de las familias cópula

---

```
    realC0[ij,jk] <- (ij/m)+(jk/m)-1+g
    if(ij==m && jk==m){
      realC0[ij,jk] <- 1}
  }}
  #print("real checkerboard copula of order m")
  #print(realC0)
}

if(model0==28){    #Husler-Reiss
  if(rho0<0){print("Husler-Reiss model does not apply")
    stop()}
  theta0 <- iRho(huslerReissCopula(),rho0)
  for(ij in 1:m-1){
    for(jk in 1:m-1){
      z1 <- (1/theta0)+((theta0/2)*log(log(ij/m)/log(jk/m)))
      z2 <- (1/theta0)+((theta0/2)*log(log(jk/m)/log(ij/m)))
      p1 <- pnorm(z1,0,1)
      p2 <- pnorm(z2,0,1)
      f <- (log(ij/m)*p1)+(log(jk/m)*p2)
      g <- exp(f)
      realC0[ij,jk] <- g}}

  for(ij in 1:m){
    realC0[ij,m] <- ij/m}
  for (jk in 1:m){
    realC0[m,jk] <- jk/m}
  print("real checkerboard copula of order m")
  print(realC0)
}

#PRINT checkC0

for(ijk in 1:m){
  for(jkl in 1:m){
    auxC0[ijk+1,jkl+1]<-realC0[ijk,jkl]}}

for(ik in 1:m){
  for(jl in 1:m){
    checkC0[ik,jl]<- auxC0[ik+1,jl+1]-auxC0[ik+1,jl]-auxC0[ik,jl+1]+auxC0[ik,jl]}}

print("checkC0")
print(checkC0)

#DISTANCIA (D) ENTRE LA CÓPULA MUESTRAL Y LA PROPUESTA (CHECK)

D <- abs(checkC0-s)
print("DISTANCIA (D) ENTRE LA CÓPULA MUESTRAL Y LA PROPUESTA (CHECK)")
print(D)

print("SUMA DE D")
```

---



---

```
print(sum(D)/2)
#distVTH0[ki,1] <- sum(D)/2

print("PARÁMETRO TETA")
print(theta0)

}

#COPULA MODELS

#model 1=Clayton (theta\in [-1,\infty) non zero IN PRACTISE theta<=125, rho in
[-1,1])
#model 3=Gumbel (theta in [1,\infty) IN PRACTISE theta<=115, rho in [0,1])
#model 4=Frank (theta in (-\infty,\infty) non zero
#      IN PRACTISE |theta|<=709.78 and THETA<=24, rho in [-1,1])
#model 5=Joe (theta in [1,\infty), rho in [0,1])
#model 6=Plackett (theta in (0,\infty) not one rho in [-1,1])
#model 9=T-Student df=4 (theta in [-1,1], rho in [-1,1])
#model 10=Normal dim=2 (theta in [-1,1], rho in [-1,1])
#model 22=Gumbel Decreasing-Decreasing(theta in [1,\infty) IN PRACTISE
theta<=115), rho in [0,1]
#model 26=Galambos
#model 27=Galambos Decreasing-Decreasing
#model 28=Husler and Reiss
```

# Apéndice D

## Compendio de códigos fuente utilizados en la investigación

### D.1. Códigos fuente programados en Matlab

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
ESTE PROGRAMA OBTIENE LA PSEUDOMUESTRA, CALCULADAS A PARTIR DE LOS DATOS
OBSERVADOS DE GASTOS Y VOLÚMENES DE ESCURRIMIENTO DE LAS ESTACIONES OBJETO DE
ESTUDIO, ES DECIR, LAS SERIES SON TRANSFORMADAS EN RANGOS (Ri/n, Si/n).
By Luis Ramírez 2019
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clc;clear;
fprintf('Gastos Pico Q y Volúmenes V \n\n')

load nombredearchivo.txt
Q=reg5_2(:,1);V=reg5_2(:,2);
tam=length(Q);N=(1:tam).';
[Q2,IQ]=sort(Q);[Q3,IIQ]=sort(IQ);[V2,IV]=sort(V);[V3,IIV]=sort(IV);
QQ=IIQ;VV=IIV;
QU=QQ/tam;VU=VV/tam;

fileID = fopen('Resultados.txt','w');
%8s\r\n', 'N', 'Q', 'V', 'Qi', 'Vi', 'Q"', 'V"');
fprintf(fileID,'%2.0f %10.3f %10.3f %3.0f %3.0f %8.5f
%8.5f\r\n',[N,Q,V,QQ,VV,QU,VU]);
fclose(fileID);
%END PROGRAM
```

---

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
CALCULA LOS VALORES DEL VOLUMEN DE ESCURRIMIENTO, PARA UNA DETERMINADA
COMBINACION DE Q-T, Y UN DETERMINADO PERIODO DE RETORNO APLICANDO LA CÓPULA DE
JOE/B5. CADA VEZ QUE SE MODIFIQUE EL PERIODO DE RETORNO T, SE TENDRÁ QUE
PROPONER UN NUEVO VECTOR v. By Luis Ramírez 2019
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

load nombredearchivo.txt
Gast = nombredearchivo(:,1); %Gastos Propuestos
tamQ = length(Gast); %tamaño del vector de gastos propuestos

V1 = 6200:-100:0; %Ingresar el Vector de volúmenes, modificar si cambia el
Periodo de Retorno T
T = 10000; %Ingresar Periodo de Retorno en, años
tamV1 = length(V1); %Tamaño del vector de volúmenes propuestos
PA = [888.0995 1581.3713 3521.1268 8569.5929 315.9558 558.4338 550.3577
2207.0087 0.840 0.890 2.083687];

for i= 1:tamQ
    Gasto = Gast(i);
    for j = 1:tamV1
        aa = exp(-exp(-((Gasto-PA(2))./PA(1))));
        bb = exp(-exp(-((Gasto-PA(4))./PA(3))));
        cc = exp(-exp(-((V1-PA(6))./PA(5))));
        dd = exp(-exp(-((V1-PA(8))./PA(7))));
        FQ = (PA(9).*aa) + ((1-PA(9)).*bb); %Marginal 1, tipo Doble Gumbel
        FV = (PA(10).*cc) + ((1-PA(10)).*dd); %Marginal 2, tipo Doble Gumbel
        A1 = (1-FQ).^PA(11);
        B1 = (1-FV).^PA(11);
        CQV = 1-(A1+B1-(A1.*B1)).^(1./PA(11)); %CÓPULA DE JOE
        gl = 1 - FQ - FV + CQV - (1/T); %Periodo de Retorno Conjunto
    Tqv
    end
    t=gl; %Vector para analizar cambio de signo
    for k = 1:tamV1
        if t(k)*t(k+1) < 0
            vi=k;
            break
        end
    end
    %Función que calcula los volúmenes de escurrimiento, ecuación del Periodo
de Retorno Conjunto
    %y la CÓPULA DE JOE, resuelta por el Método Numérico de Bisección

    a=V1(vi); %Primer límite para el Método de Bisección
    b=V1(vi+1); %Segundo límite para el Método de Bisección
    for jj=1:70
        a=a;
        b=b;

```

## Compendio de códigos fuente utilizados en la investigación

---

```
fa = double(1 - FQ - ((PA(10).*(exp(-exp(-(a-PA(6))./PA(5)))))) + ((1-
PA(10)).*(exp(-exp(-(a-PA(8))...
    ./PA(7)))))) + (1-(A1+((1-((PA(10).*(exp(-exp(-(a-
PA(6))./PA(5)))))) + ((1-PA(10)).*(exp...
    (-exp(-(a-PA(8))./PA(7)))))))).^PA(11)-(A1.*((1-((PA(10).*(exp(-
exp(-(a-PA(6))./PA(5))))))...
    + ((1-PA(10)).*(exp(-exp(-(a-
PA(8))./PA(7)))))))).^PA(11))).^(1./PA(11))) - (1./T));

fb = double(1 - FQ - ((PA(10).*(exp(-exp(-(b-PA(6))./PA(5)))))) + ((1-
PA(10)).*(exp(-exp(-(b-PA(8))...
    ./PA(7)))))) + (1-(A1+((1-((PA(10).*(exp(-exp(-(b-
PA(6))./PA(5)))))) + ((1-PA(10)).*(exp...
    (-exp(-(b-PA(8))./PA(7)))))))).^PA(11)-(A1.*((1-((PA(10).*(exp(-
exp(-(b-PA(6))./PA(5))))))...
    + ((1-PA(10)).*(exp(-exp(-(b-
PA(8))./PA(7)))))))).^PA(11))).^(1./PA(11))) - (1./T));

if fa*fb > 0
    disp('MODIFICAR LÍMITES a Y b')
else
    q = (a + b)/2;
    fq = double(1 - FQ - ((PA(10).*(exp(-exp(-(q-PA(6))./PA(5)))))) + ((1-
PA(10)).*(exp(-exp(-(q-PA(8))...
    ./PA(7)))))) + (1-(A1+((1-((PA(10).*(exp(-exp(-(q-
PA(6))./PA(5)))))) + ((1-PA(10)).*(exp...
    (-exp(-(q-PA(8))./PA(7)))))))).^PA(11)-(A1.*((1-((PA(10).*(exp(-
exp(-(q-PA(6))./PA(5))))))...
    + ((1-PA(10)).*(exp(-exp(-(q-
PA(8))./PA(7)))))))).^PA(11))).^(1./PA(11))) - (1./T));
    err = abs(fq);
    if fa*fq < 0
        a = a;
    else
        a = q;
    end
    if fb*fq < 0
        b = b;
    else
        b = q;
    end
end
end
end
```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%  
 CALCULA EL TIEMPO AL CENTROIDE (tg), POR MEDIO DE LA FUNCIÓN GAMMA 2P. PARA  
 CONSTRUIR EL DEL HIDROGRAMA DE AVENIDA DE ENTRADA A LA PRESA. By Luis Ramírez 2019  
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

load nombredearchivo.txt
Gast = nombredearchivo(:,1); %Gastos Pico Propuestos en m3/s
Vol = nombredearchivo(:,2); %Volúmenes asociados en, Hm3
tamQ = length(Gast); %tamaño del vector de gastos propuestos

V1 = 305:-5:40; %V1=Vl, Ingresar el Vector de volúmenes, modificar si
cambia el Periodo de Retorno T
tamV1 = length(V1); %Tamaño del vector de tiempos al centroide tg del
Hidrograma Gamma 2p
tp= 38.6498982116263; %Ingresar el tiempo pico, en horas
for i= 1:tamQ
    Gasto = Gast(i);
    Volu = Vol(i);
    for j = 1:tamV1
        A=(tp./(V1-tp));
        B=(A.^-1);
        C=1+A;
        gama=gamma(C);
        g1 = ((Gasto.*0.0036).*exp(A).*(B.^A).*(V1-tp).*gama)-Volu; %Gasto en
Hm3/h, Volumen en Hm3
    end
    t=g1; %Vector para analizar cambio de signo
    for k = 1:tamV1
        if t(k)*t(k+1) < 0
            vi=k;
            break
        end
    end
    end
%Función que calcula el tiempo al centroide tg por el Método Numérico de
Bisección
    a1=V1(vi); %Primer límite para el Método de Bisección
    b1=V1(vi+1); %Segundo límite para el Método de Bisección

    g2 = @(tg) ((Gasto.*0.0036).*exp(tp./(tg-tp)).*(((tp./(tg-tp)).^-
1).^(tp./(tg-tp))).*(tg-tp).*gamma...
        (1+(tp./(tg-tp))))-Volu;
    T2(i) = bisection(g2,a1,b1);
    GAM(i)= gamma(1+(tp./(T2(i)-tp)));
    VP(i) = (((Gasto.*0.0036).*exp(tp./(T2(i)-tp)).*(((tp./(T2(i)-tp)).^-
1).^(tp./(T2(i)-tp))).*(T2(i)...
        -tp).*gamma(1+(tp./(T2(i)-tp)))));
    VPC(i)= (((Gasto.*0.0036).*exp(tp./(T2(i)-tp)).*(((tp./(T2(i)-tp)).^-
1).^(tp./(T2(i)-tp))).*(T2(i)...
        -tp).*gamma(1+(tp./(T2(i)-tp)))))-Volu;
end

```

## Compendio de códigos fuente utilizados en la investigación

---

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
OBTIENE LA CÓPULA GAUSSIANA (NORMAL)  $C(u_1, u_2)$ , LA CÓPULA ES EVALUADA CON
MARGINALES  $F(Q) = u_1$ ,  $F(V) = u_2$ . Se requiere ingresar el parámetro  $\theta$ .
```

```
By Luis Ramírez 2019
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
load nombredearchivo.txt
FQ = nombredearchivo(:,1);
FV = nombredearchivo(:,2);
tam=length(FQ);
%N=(1:tam).'; %N => es un contador

for i=1:tam                                %INGRESAR EL PARÁMETRO THETA (OJO)
C = copulacdf('Gaussian',[FQ,FV],0.6179674);
end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
OBTIENE LA FUNCIÓN INVERSA DE LA FDP GAMMA 3P, SE RESUELVE NUMÉRICAMENTE CON
EL MÉTODO NUMÉRICO DE BISECCIÓN. CÁLCULO DE CUANTILES. By Luis Ramírez 2019
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
load nombredearchivo.txt
Proba = nombredearchivo(:,1); %Vector de Probabilidades originales ORDENADOS
DE MAYOR A MENOR
tamQ = length(Proba); %tamaño del vector de Probabilidades

Q1 = 600:2000:13000; %Ingresar el Vector de gastos/volúmenes para análisis de
signos (met. bisección)
tamQ1 = length(Q1);

PA = [806.3649 3.2779 494.2317]; %INGRESAR LOS PARÁMETROS ESTADÍSTICOS
gg = gamma(PA(2));

for i= 1:tamQ                                %tamaño del vector de Probabilidades
    P1 = Proba(i);
    for j = 1:tamQ1 %Tamaño del Vector de gastos/volúmenes para análisis de
signos (met. bisección)
        RQ(j) = integral(@(x) (1./(PA(1).*gg)).*(((x-PA(3))./PA(1)).^(PA(2)-
1)).*exp(-(x-PA(3))./...
            PA(1)), PA(3), Q1(j)) - P1;
    end
    t = RQ.'; %Vector para analizar cambio de signo
    for k = 1:tamQ1
        if t(k)*t(k+1) < 0
            vi=k;
            break
        end
    end
    a=Q1(vi); %Primer límite para el Método de Bisección
    b=Q1(vi+1);
```

---

```

%SE APLICA EL MÉTODO DE BISECCIÓN PARA OBTENER EVENTOS DE DISEÑO CON ADF
fa = integral(@(x1) (1./(PA(1).*gg)).*(((x1-PA(3))./PA(1)).^(PA(2)-
1)).*exp(-(x1-PA(3))./PA(1))), PA(3), a) - P1;
fb = integral(@(x2) (1./(PA(1).*gg)).*(((x2-PA(3))./PA(1)).^(PA(2)-
1)).*exp(-(x2-PA(3))./PA(1))), PA(3), b) - P1;

if fa*fb > 0
    disp('MODIFICAR LÍMITES a Y b')
else
    q = (a + b)/2;
    fq = integral(@(x3) (1./(PA(1).*gg)).*(((x3-PA(3))./PA(1)).^(PA(2)-
1)).*exp(-(x3-PA(3))./PA(1))), PA(3), q) - P1;
    err = abs(fq);
    while err > 1e-15
        if fa*fq < 0
            a = q;
        else
            b = q;
        end
        q = (a + b)/2;
        fq = integral(@(x3) (1./(PA(1).*gg)).*(((x3-
PA(3))./PA(1)).^(PA(2)-1)).*exp(-(x3-PA(3))./PA(1))), PA(3), q) - P1;
        err = abs(fq);
    end
end
end
end

```

## D.2. Códigos fuente programados en Fortran

```

!*****
CALCULA EL ANÁLISIS DE FRECUENCIA UNIVARIADO, EL PROGRAMA TAMBIEN AJUSTA UNA
SERIE DE GASTOS MÁXIMOS ANUALES MEDIANTE LA FUNCION DE DISTRIBUCION DE
PROBABILIDAD DOBLE GUMBEL "Q2G". ASIMISMO SE ESTIMAN EVENTOS DE DISÑO POR
MEDIO DE LA LEY EMPIRICA DE WEIBULL,  $1-(1/T) = k/(n+1)$ , (datos ascendentes)
By Luis Ramírez 2010
!*****
PROGRAM EVENTOS_Q
    IMPLICIT NONE

    INTEGER, PARAMETER:: dimqf = 52 !INGRESAR el tamaño de la muestra
    INTEGER, PARAMETER:: dimp = 5 !INGRESAR el número de parámetros estadísticos
    INTEGER, PARAMETER:: dTr = 17 !INGRESAR el Tamaño del vector de Periodos de Retorno
    REAL, DIMENSION(1:dimqf):: QA,Qi,Q2,P,Tre,Qc,fQc
    REAL, DIMENSION(1:dimp):: pi !Vector de parámetros estadísticos

```

## Compendio de códigos fuente utilizados en la investigación

---

```
REAL, DIMENSION(1:dTr):: Tr,QTr,fQTr
INTEGER:: i,j,k,Salida
REAL:: temp          !DATOS OBSERVADOS (ORIGINALES)

OPEN(UNIT=100,FILE='Registros.dat',STATUS='UNKNOWN',ACTION='READ')
OPEN(UNIT=200,FILE='Resultados.dat',STATUS='UNKNOWN',ACTION='WRITE')

!INGRESAR parámetros estadísticos YA optimizados (alfal,beta1,alfa2,beta2,P) de la
!FDP PDOBLE GUMBEL, EL VECTOR DE PARÁMETROS DE GASTOS ACEPTA SOLO NÚMEROS REALES

pi = (/680.94,1516.39,3140.60,5729.79,0.7383/) !Vector de parámetros de gastos

!INICIA EL ANÁLISIS DE FRECUENCIA UNIVARIADO
!=====

!Lectura de datos observados
DO i=1,dimqf
  READ(100,*) QA(i)
END DO
CLOSE(UNIT=100)
!Vector de gastos máximos en su forma original (datos observados)
DO k=1,dimqf
  Qi(k)=QA(k)
END DO
!SORT, datos observados(gastos máx) para organizar ascendentemente
DO k=1,dimqf-1
  DO j=k+1,dimqf
    IF (Qi(k)>Qi(j)) THEN
      temp=Qi(k)
      Qi(k)=Qi(j)
      Qi(j)=temp
    END IF
  END DO
END DO

!Vector de gastos máximos en forma ascendente (menor a mayor)
DO k=1,dimqf
  Q2(k)=Qi(k)
END DO

!Periodo de retorno Tr, empleando fórmula de Weibull (para datos ascendentes),
;T = (n+1)/[(n+1)-k]
DO k=1,dimqf
  Tre(k)= (dimqf+1)/((dimqf+1)-REAL(k))
END DO

!Probabilidad de NO excedencia  $P(Q \leq q) = 1 - (1/T) = k/(n+1)$ , empleando la
! fórmula de la Ley empírica de Weibull (para datos ascendentes)
DO j=1,dimqf
  P(j)=1-(1/Tre(j))
END DO

CONTAINS
```

---



```

!*****
!Subrutina Ajuste_G2P ( )
!Esta subrutina calcula los Gastos Máximos de diseño Qc(m3/s) para diferentes
!periodos de retorno Tr(años), utilizando el método de Newton Raphson.
!*****
      SUBROUTINE Ajuste_G2P (dp,dqf,p,Q,PQ,Qcal,fQcal)
            IMPLICIT NONE

            INTEGER, INTENT(IN):: dp,dqf
            REAL, DIMENSION(1:dp):: p
            REAL, DIMENSION(1:dqf):: Q,PQ,Qcal,fQcal
            INTEGER:: i
            REAL:: tol,Gast,Prob

            tol = 1.0E-07

            DO i=1,dqf
                  Gast = Q(i)
                  Prob = PQ(i)
                  DO WHILE (ABS(fQ_NR(dp,p,Gast,Prob)) >= tol)
                        Gast = Gast - fQ_NR(dp,p,Gast,Prob)/dfQ(dp,p,Gast)
                  END DO
                  Qcal(i) = Gast
                  fQcal(i) = fQ_NR(dp,p,Gast,Prob)
            END DO

      END SUBROUTINE Ajuste_G2P

!*****
!Función fQ_NR( )
!Ésta función calcula el valor parcial de la función objetivo f(Qc)
!*****

      REAL FUNCTION fQ_NR (dp,p,Gast,Prob)
            IMPLICIT NONE

            REAL, INTENT(IN):: Gast,Prob
            INTEGER:: dp
            REAL:: a,b,p(dp)

            a = EXP(-EXP(-((Gast-p(2))/p(1))))
            b = EXP(-EXP(-((Gast-p(4))/p(3))))

            fQ_NR = (p(5)*a) + ((1-p(5))*b) - Prob

      END FUNCTION fQ_NR

!*****
!Función dfQ( )
!Ésta función calcula el valor la derivada de la función f(Qc)
!*****

      REAL FUNCTION dfQ (dp,p,Gast)

```

## Compendio de códigos fuente utilizados en la investigación

---

```
IMPLICIT NONE

REAL, INTENT(IN):: Gast
INTEGER:: dp
REAL:: a,b,c,d,p(dp)

a = EXP(-EXP(-((Gast-p(2))/p(1))))
b = EXP(-EXP(-((Gast-p(4))/p(3))))
c = EXP(-((Gast-p(2))/p(1)))
d = EXP(-((Gast-p(4))/p(3)))

dfQ = ((p(5)/p(1))*a*c) + (((1-p(5))/p(3))*b*d)

END FUNCTION dfQ

!*****
!Subrutina Extr_GastDis ()
!Esta subrutina calcula los eventos de diseño para diferentes Periodos de Retorno
!Tr = 2,5,10,20,50,100,200,500,1000,2000,5000,10000-15000 (años)
!*****
SUBROUTINE Extr_GastDis (dp,dqf,dT,p,Qcal,T,QT,fQT)
IMPLICIT NONE

INTEGER, INTENT(IN)      :: dp,dqf,dT
REAL, DIMENSION(1:dp)   :: p
REAL, DIMENSION(1:dT)   :: T,QT,fQT,PQ
REAL, DIMENSION(1:dqf)  :: Qcal
INTEGER:: i,j,k
REAL:: tol,Gast,Prob
tol = 1.0E-07

IF (MOD(dqf,2)==0) THEN
!Elige gasto con Tr=2 años; cuando el tamaño de la muestra es número par
k=dqf/2
ELSE
!Elige gasto con Tr=2 años; cuando el tamaño de la muestra es número impar
k=(dqf+1)/2
END IF

Gast = Qcal(k)
T = (/2,5,10,20,50,100,200,500,1000,2000,5000,10000,11000,12000,13000,14000,15000/)
DO i=1,dT
PQ(i)=1-(1/REAL(T(i)))
END DO
DO j=1,dT
Prob = PQ(j)
DO WHILE (ABS(fQ_NR(dp,p,Gast,Prob)) >= tol)
Gast = Gast - fQ_NR(dp,p,Gast,Prob)/dfQ(dp,p,Gast)
END DO
QT(j) = Gast
fQT(j) = fQ_NR(dp,p,Gast,Prob)
END DO
END SUBROUTINE Extr_GastDis

END PROGRAM EVENTOS_Q
```

### D.3. Códigos fuente programados en Macros Visual Basic

```
'////////////////////////////////////
Ubica los gastos máximos contenidos en una matriz de gastos medios diarios
incluida en una hoja de un determinado libro de Excel. Realiza un resumen con los
gastos máximos ocurridos en cada año del registro histórico comprendido. Dibuja el
hidrograma de la avenida ocurrida para cada año, posicionando el gasto máximo al
centro y tomando diez gastos medios diarios a la izquierda del gasto máximo y
quince gastos medios diarios a la derecha del gasto máximo.
```

By Luis Ramírez 2018

```
'////////////////////////////////////
```

```
Sub QMAX() ' UBICA LOS GASTOS MÁXIMOS y MUESTRA UN RESUMEN EN LISTA (AÑO,QMÁX)
```

```
'QMAX Macro
```

```
'DECLARACIÓN DE VARIABLES
```

```
Dim a As Integer
```

```
Dim b As Integer
```

```
Dim i As Integer
```

```
Dim tam As Integer
```

```
Dim ubica As String
```

```
Dim valor As Long
```

```
Dim celdaDestino As String
```

```
Dim rango As Range
```

```
Dim busca As Range
```

```
a = 2
```

```
b = 13
```

```
'CALCULA EL TAMAÑO DE REGISTRO DE GASTOS MEDIOS DIARIOS
```

```
tam = ActiveSheet.Cells(2, 35).Value
```

```
'UBICA LOS GASTOS MÁXIMOS EN LA MATRIZ DE GASTOS MEDIOS DIARIOS, Y LOS SEÑALA
!SOMBREÁNDOS EN COLOR AZUL.
```

```
For i = 1 To tam
```

```
    Set rango = ActiveSheet.Range(Cells(a, 3), Cells(a + 11, 33))
```

```
    Set busca = rango.Find(Application.WorksheetFunction.Max(rango),
```

```
    LookIn:=xlValues, lookat:=xlWhole)
```

```
    If Not busca Is Nothing Then
```

```
        ubica = busca.Address
```

```
    End If
```

```
    Range(ubica).Select
```

```
    With Selection.Interior
```

```
        .Pattern = xlSolid
```

```
        .PatternColorIndex = xlAutomatic
```

```
        .Color = 16711680
```

```
        .TintAndShade = 0
```

```
        .PatternTintAndShade = 0
```

```
    End With
```

## Compendio de códigos fuente utilizados en la investigación

---

```
With Selection.Font
    .ThemeColor = xlThemeColorDark1
    .TintAndShade = 0
End With
Selection.Font.Bold = True

'MUESTRA UN RESUMEN EN FORMA DE LISTA, SOBRE LA RELACIÓN DE GASTOS MÁXIMOS
'OCURRIDOS ANUALMENTE.

celdaDestino = "AL" & i + 1
Range(celdaDestino).Value = "=" & ubica

Range(Cells(b, 1), Cells(b, 1)).Select
Selection.Copy
celdaDestino = "AK" & i + 1
valor = Cells(b, 1)
Range(celdaDestino).Value = valor

a = a + 12
b = b + 12
Next i
Range("AI2").Select
Calculate

End Sub

'////////////////////////////////////

Sub HIDRO_RANG() 'DIBUJA EL HIDROGRAMA DE LA AVENIDA OCURRIDA PARA CADA AÑO

Dim rngTable As Range
Dim rngDest As Range
Dim intIndexCount As Integer, iX As Integer
Dim a As Integer
Dim b As Integer
Dim e As Integer
Dim f As Integer
Dim i As Integer
Dim tam As Integer
Dim valArray()
Dim ul As Long, rango1 As Range
Dim x As Integer
Dim rango As Range
Dim busca As Range
Dim ubica As String

a = 2
b = 4
e = 2
f = 2

tam = ActiveSheet.Previous.Cells(2, 35).Value
```

---

```
For i = 1 To tam

'TRANSFORMA LOS ELEMNTOS DE UNA MATRIZ A COMPONENTES DE UN VECTOR VERTICAL

Set rngTable = ActiveSheet.Previous.Range(ActiveSheet.Previous.Cells(a, 3),
ActiveSheet.Previous.Cells(a + 11, 33))
intIndexCount = rngTable.Count
ReDim valArray(intIndexCount)

For iX = 1 To intIndexCount
    valArray(iX - 1) = rngTable(iX)
Next iX

Set rngDest = ActiveSheet.Range("N1").Resize(UBound(valArray), 1)
rngDest = Application.Transpose(valArray)

'ELIMINA CELDAS VACIAS DE UNA COLUMNA Y RECORRE CELDAS DE UN RANGO

ul = Range("N" & Rows.Count).End(xlUp).Row
Set rangol = Range("N1:N" & ul)
x = 2
For Each celda In rangol
    If celda <> "" Then
        Range("O" & x) = celda
        x = x + 1
    End If
Next

Range("N1:N" & ul).Select
Selection.ClearContents

'ENCUENTRA EL VALOR MÁXIMO DE UN VECTOR Y UBICA SU POSICIÓN

Set rango = ActiveSheet.Range("O1:O" & ul)
Set busca = rango.Find(Application.WorksheetFunction.Max(rango),
LookIn:=xlValues, lookat:=xlWhole)
If Not busca Is Nothing Then
    ubica = busca.Address
End If

'SELECCIONA EL RANGO DEL HIDROGRAMA DE LA AVENIDA

Range(ubica).Offset(-10, 0).Select
Selection.Resize(Selection.Rows.Count + 25).Select
Selection.Copy
Range(Cells(b, 2), Cells(b, 2)).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
:=False, Transpose:=False

'IMPRESIONA LOS PUNTOS DEL HIDROGRAMA DE AVENIDA

ActiveSheet.Previous.Select
```

---

```
Cells(e, 1).Select
Selection.Copy
ActiveSheet.Next.Select
Cells(f, 1).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=False

a = a + 12
b = b + 30
e = e + 12
f = f + 30

Next i

Range("O1:O" & ul).Select
Selection.ClearContents
Range("A1").Select
Calculate

End Sub
```

### D.4. Códigos fuente programados en lenguaje R

```
#####
Este programa calcula un vector de medidas de dependencia de Rho de Spearman,
a través de simular una muestra aleatoria de 10,000 puntos generada con la
Cópula de Joe/B5, este procedimiento se realiza para cada estimación de Rho de
Spearman. Para ello se propone un vector de parámetros theta en el intervalo
[1.01-4] con espaciamentos de  $\Delta=0.1$ .
#####
```

```
vector1<-array(rep(0,times=39),dim=c(39,1))
v <- c(1.01,1.02,1.03,1.04,1.05,1.06,1.07,1.08,1.09,1.1,...,4)
for(i in 1:39){
  uu <- joeCopula(v[i])
  B1 <- rCopula(10000,uu)
  a <- cor(B1,use="complete.obs",method="spearman")
# print(a[1,2])
  vector1[i] <- a[1,2]}
print(vector1)
```

```
##### Para la cópula de t-student se utiliza la siguiente instrucción
##### uu <- tCopula(v[i])
```

```
#####
ELIMINA ALEATORIAMENTE UN PAR DE DATOS OBSERVADOS (Q,V) DE UNA MUESTRA
ORIGINAL. LO ANTERIOR SE HACE PARA AJUSTAR EL TAMAÑO DE LA MUESTRA DE DSATOS
OBSERVADOS PARA APLICAR EL MÉTODO DE LA PRUEBA DE BONDAD DE AJUSTE PROPUESTO
POR GONZÁLEZ-HOYOS (2019), ESPECÍFICAMENTE LA GENERACIÓN DEL GRID.
#####
```

---

```
ELIMINA_UNO<- function(sample.size,order.part)

{
d<-2
n<-sample.size
m<-order.part

inpath <- 'C:\\Users\\LUIS\\Desktop\\CARPETA\\morada37.csv'

Alaux <- read.csv(inpath)
A1 <- data.matrix(Aiaux)
n <- nrow(A1)
print("A1")
print(A1)
print(n)

du <- n/m
fdu <- floor(du)

if(du-fdu>0){
q <- runif(1,min=0,max=1)
print("q")
print(q)
for(k in 1:n){
if((k-1)/n<q && q <k/n)
{quitar <- k
print("quitar")
print(quitar)}}
B1 <- array(rep(0,times=2*(n-1)), dim=c((n-1),2))
if(quitar>1){
qq <- quitar-1
qqq <- quitar+1
B1[1:qq,1] <- A1[1:qq,1]
B1[1:qq,2] <- A1[1:qq,2]
B1[quitar:(n-1),1] <- A1[qqq:n,1]
B1[quitar:(n-1),2] <- A1[qqq:n,2]}
A1 <- B1
print("B1")
print(B1)
n <- n-1}

print("n")
print(n)
}
#system.time(ELIMINA_UNO(sample.size,order.part))
```

## Compendio de códigos fuente utilizados en la investigación

---

```
#####  
GENERACIÓN DE UNA MUESTRA DE 10,000 PARES DE PUNTOS DE LAS MARGINALES U1 Y U2  
OBTENIDAS POR MEDIO DE LA CÓPULA DE CLAYTON. TAMBIÉN OBTIENE LA PROBABILIDAD  
DE LA CÓPULA C(U1,U2). DATOS DE LA ESTACIÓN LAS PERLAS, CUENCA DEL RÍO  
COATZACOALCOS, VERACRUZ.  
#####
```

```
Cu <- claytonCopula(2.723024)  
U <- rCopula(10000,Cu)  
CclaytonU1U2 <- pCopula(U,copula=Cu)  
setwd("C:/NOMBREDCARPETA")  
Datos = data.frame(CclaytonU1U2, U[,1], U[,2])
```

```
#####  
GENERACIÓN DEL GRÁFICO K-PLOT, UTILIZANDO EL PAQUETE DE "ICOPULA" DEL LENGUAJE  
R. SE UTILIZA LA CÓPULA DE GALAMBOS. ESTE PAQUETE TOMA EN CUENTA A LAS CÓPULAS  
DE VALORES EXTREMOS, ELÍPTICAS, ARQUIMEDIANAS Y OTRAS. DATOS DE LA ESTACIÓN  
LAS PERLAS, CUENCA DEL RÍO COATZACOALCOS, VERACRUZ.  
#####
```

```
K.plot(rCopula(n=52,galambosCopula(param=iTau(galambosCopula(0.8992416),0.3831  
071))),add=TRUE,col=2)
```

```
#####  
GENERACIÓN DEL GRÁFICO CHI-PLOT, UTILIZANDO EL PAQUETE "ASBIO" DEL LENGUAJE R.  
SE UTILIZA LA CÓPULA DE HÜSLER-REISS. ESTE PAQUETE TOMA EN CUENTA A LAS  
CÓPULAS DE VALORES EXTREMOS, ELÍPTICAS, ARQUIMEDIANAS Y OTRAS. DATOS DE LA  
ESTACIÓN LAS PERLAS, CUENCA DEL RÍO COATZACOALCOS, VERACRUZ.  
#####
```

```
Cu <- huslerReissCopula(1.350887)  
U <- rCopula(52,Cu)  
chi.plot(U[,1], U[,2])
```





# Apéndice E

## Registros de Gastos medios diarios de la Estación Huites, Sinaloa

### E.1. Introducción

La información hidrométrica utilizada se recopiló del Banco Nacional de Datos de Aguas Superficiales (BANDAS), que es la base de datos de la Comisión Nacional del Agua (2014a), constituida como un órgano desconcentrado de SEMARNAT y máxima autoridad en el sector hídrico del Gobierno Federal de México.

El periodo de registro histórico de datos para la estación hidrométrica Huites ubicada en la cuenca del Río Fuerte, en el Estado de Sinaloa, México, abarca de 1941 a 1992.

Se estima un gasto medio de  $132.963 \text{ m}^3/\text{s}$  en dicho periodo, y un volumen de escurrimiento medio de  $4,273.544 \text{ Hm}^3$ . El 29 de diciembre de 1990 se presentó el Gasto máximo registrado en el periodo comprendido,  $Q_{\text{máx}} = 10,129.264 \text{ (m}^3/\text{s)}$ . En la Tabla E.1, se puede consultar la información de los registros de gastos medios diarios durante el periodo.

Tabla E.1. Gastos Medios Diarios de la Estación Huites, Sinaloa. Cuenca del Río Fuerte.

Año Mes		ESTACION HUITES, SINALOA																																								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31										
1941	AGO	607.327	551.013	543.184	498.663	569.212	620.958	522.567	448.665	380.804	481.609	280.283	363.270	282.236	215.144	215.004	260.075	305.064	333.851	434.328	605.002	755.235	1423.514	911.218	532.031	352.888	257.152	218.108	222.925	199.354												
	SEPT	165.085	163.859	161.746	145.804	116.699	98.737	84.123	74.193	65.922	263.490	655.531	431.414	306.016	208.190	154.728	123.768	101.181	87.278	75.000	66.591	60.375	55.061	49.561	46.250	43.098	41.830	39.250	37.849	34.961												
	OCT	33.013	31.280	30.248	29.217	27.764	27.109	25.138	24.935	23.516	22.750	22.280	21.989	21.590	21.220	20.936	20.699	20.493	20.316	20.150	20.000	19.875	19.765	19.665	19.575	19.495	19.425	19.365	19.315	19.275												
	NOV	33.653	33.449	32.861	31.490	30.288	29.078	28.313	27.553	26.798	26.038	25.278	24.519	23.759	23.001	22.245	21.491	20.739	20.000	19.250	18.500	17.750	17.000	16.250	15.500	14.750	14.000	13.250	12.500	11.750												
	DIC	18.331	17.967	17.633	17.318	17.024	16.749	16.484	16.229	15.984	15.749	15.524	15.309	15.104	14.909	14.724	14.549	14.384	14.229	14.084	13.949	13.824	13.709	13.604	13.509	13.424	13.349	13.284	13.229	13.184	13.149											
	ENE	34.653	33.449	32.861	31.490	30.288	29.078	28.313	27.553	26.798	26.038	25.278	24.519	23.759	23.001	22.245	21.491	20.739	20.000	19.250	18.500	17.750	17.000	16.250	15.500	14.750	14.000	13.250	12.500	11.750												
	FEB	12.300	12.000	11.750	11.500	11.250	11.000	10.750	10.500	10.250	10.000	9.750	9.500	9.250	9.000	8.750	8.500	8.250	8.000	7.750	7.500	7.250	7.000	6.750	6.500	6.250	6.000	5.750	5.500	5.250	5.000											
	MAR	12.300	12.000	11.750	11.500	11.250	11.000	10.750	10.500	10.250	10.000	9.750	9.500	9.250	9.000	8.750	8.500	8.250	8.000	7.750	7.500	7.250	7.000	6.750	6.500	6.250	6.000	5.750	5.500	5.250	5.000											
	ABR	12.300	12.000	11.750	11.500	11.250	11.000	10.750	10.500	10.250	10.000	9.750	9.500	9.250	9.000	8.750	8.500	8.250	8.000	7.750	7.500	7.250	7.000	6.750	6.500	6.250	6.000	5.750	5.500	5.250	5.000											
	MAY	7.885	7.695	7.519	7.354	7.199	7.054	6.919	6.794	6.679	6.574	6.479	6.394	6.319	6.254	6.199	6.154	6.109	6.074	6.039	6.004	5.969	5.934	5.904	5.874	5.844	5.814	5.784	5.754	5.724	5.694	5.664										
	JUN	5.881	5.699	5.524	5.359	5.204	5.059	4.924	4.799	4.684	4.579	4.484	4.409	4.344	4.289	4.244	4.209	4.184	4.169	4.154	4.149	4.144	4.139	4.134	4.129	4.124	4.119	4.114	4.109	4.104	4.099	4.094	4.089									
	JUL	49.403	51.365	49.777	47.403	45.133	42.969	40.912	38.969	37.133	35.403	33.777	32.251	30.825	29.500	28.275	27.150	26.125	25.200	24.375	23.650	23.025	22.500	22.075	21.750	21.525	21.400	21.375	21.350	21.325	21.300	21.275	21.250									
	AGO	136.261	96.125	193.676	297.932	273.096	254.046	231.610	206.676	179.513	152.933	128.141	104.114	80.844	59.244	40.414	194.912	337.477	483.277	487.936	381.117	590.227	274.695	186.331	146.439	106.728	66.331	32.831	14.639	106.728	113.237	99.244										
	SEPT	493.815	497.833	615.137	1306.135	1210.981	1775.511	2306.573	331.198	361.239	300.184	320.236	428.716	381.439	292.027	274.431	279.942	333.951	788.211	676.130	676.302	676.492	540.590	785.439	717.297	598.261	370.007	263.809	148.809	156.599	125.199	109.169										
	OCT	107.565	150.682	149.865	442.404	317.515	245.967	133.069	179.033	200.862	110.213	202.611	188.868	132.237	110.398	94.438	84.164	78.661	69.237	62.842	57.855	54.000	52.250	49.398	47.200	45.288	44.000	42.500	40.968	39.353	37.668	35.933	34.163									
	NOV	72.840	59.149	49.000	38.362	34.981	242.637	144.679	157.841	106.973	77.658	61.570	51.348	44.848	40.400	38.949	38.749	38.549	38.349	38.149	37.949	37.749	37.549	37.349	37.149	36.949	36.749	36.549	36.349	36.149	35.949	35.749	35.549									
	DIC	32.450	29.149	27.600	26.051	24.502	22.953	21.404	19.855	18.306	16.757	15.208	13.659	12.110	10.561	9.012	7.463	5.914	4.365	2.816	1.267	0.718	0.169	0.620	1.071	1.522	1.973	2.424	2.875	3.326	3.777	4.228	4.679									
	ENE	21.524	21.408	21.292	21.176	21.060	20.944	20.828	20.712	20.596	20.480	20.364	20.248	20.132	20.016	19.900	19.784	19.668	19.552	19.436	19.320	19.204	19.088	18.972	18.856	18.740	18.624	18.508	18.392	18.276	18.160	18.044	17.928	17.812								
	FEB	17.955	17.912	17.869	17.826	17.783	17.740	17.697	17.654	17.611	17.568	17.525	17.482	17.439	17.396	17.353	17.310	17.267	17.224	17.181	17.138	17.095	17.052	17.009	16.966	16.923	16.880	16.837	16.794	16.751	16.708	16.665	16.622	16.579	16.536							
	MAR	11.456	10.959	10.463	9.967	9.471	8.975	8.479	7.983	7.487	6.991	6.495	5.999	5.503	5.007	4.511	4.015	3.519	3.023	2.527	2.031	1.535	1.039	0.543	0.047	0.551	1.055	1.559	2.063	2.567	3.071	3.575	4.079	4.583								
	ABR	6.748	6.639	6.530	6.421	6.312	6.203	6.094	5.985	5.876	5.767	5.658	5.549	5.440	5.331	5.222	5.113	5.004	4.895	4.786	4.677	4.568	4.459	4.350	4.241	4.132	4.023	3.914	3.805	3.696	3.587	3.478	3.369	3.260	3.151							
	MAY	4.699	4.629	4.560	4.491	4.422	4.353	4.284	4.215	4.146	4.077	4.008	3.939	3.870	3.801	3.732	3.663	3.594	3.525	3.456	3.387	3.318	3.249	3.180	3.111	3.042	2.973	2.904	2.835	2.766	2.697	2.628	2.559	2.490	2.421							
	JUN	4.116	4.038	3.960	3.882	3.804	3.726	3.648	3.570	3.492	3.414	3.336	3.258	3.180	3.102	3.024	2.946	2.868	2.790	2.712	2.634	2.556	2.478	2.400	2.322	2.244	2.166	2.088	2.010	1.932	1.854	1.776	1.698	1.620	1.542							
	JUL	4.116	4.038	3.960	3.882	3.804	3.726	3.648	3.570	3.492	3.414	3.336	3.258	3.180	3.102	3.024	2.946	2.868	2.790	2.712	2.634	2.556	2.478	2.400	2.322	2.244	2.166	2.088	2.010	1.932	1.854	1.776	1.698	1.620	1.542							
	AGO	395.220	482.852	348.124	575.940	457.940	627.940	489.940	659.940	521.940	691.940	553.940	723.940	585.940	755.940	617.940	787.940	649.940	819.940	681.940	851.940	713.940	883.940	745.940	915.940	777.940	947.940	809.940	979.940	841.940	1011.940	873.940	1043.940	905.940	1115.940							
	SEPT	497.630	482.852	348.124	575.940	457.940	627.940	489.940	659.940	521.940	691.940	553.940	723.940	585.940	755.940	617.940	787.940	649.940	819.940	681.940	851.940	713.940	883.940	745.940	915.940	777.940	947.940	809.940	979.940	841.940	1011.940	873.940	1043.940	905.940	1115.940							
	OCT	43.397	41.249	39.101	36.953	34.805	32.657	30.509	28.361	26.213	24.065	21.917	19.769	17.621	15.473	13.325	11.177	9.029	6.881	4.733	2.585	0.437	17.862	127.488	225.386	178.272	148.882	127.948	110.018	97.185	78.651	64.623	51.595	39.567	28.539	18.511						
	NOV	22.144	21.908	21.672	21.436	21.200	20.964	20.728	20.492	20.256	20.020	19.784	19.548	19.312	19.076	18.840	18.604	18.368	18.132	17.896	17.660	17.424	17.188	16.952	16.716	16.480	16.244	16.008	15.772	15.536	15.300	15.064	14.828	14.592	14.356	14.120						
	DIC	65.245	66.871	108.004	97.762	85.159	75.998	70.117	65.518	60.919	56.320	51.721	47.122	42.523	37.924	33.325	28.726	24.127	19.528	14.929	10.330	5.731	1.132	16.362	106.458	197.900	147.108	104.375	62.642	21.909	1.132	16.362	106.458	197.900	147.108	104.375	62.642					
	ENE	149.305	161.139	180.024	153.958	135.962	153.959	268.873	222.792	183.977	171.961	141.479	117.706	91.900	79.375	73.500	68.625	63.750	58.875	53.999	49.125	44.250	39.375	34.500	29.625	24.750	19.875	15.000	10.125	5.250	0.375	15.000	10.125	5.250	0.375	15.000	10.125	5.250	0.375			
	FEB	70.200	530.957	333.003	322.988	252.101	199.304	166.924	142.483	126.073	112.268	101.427	91.949	85.614	80.279	75.944	71.609	67.274	62.939	58.604	54.269	49.934	45.599	41.264	36.929	32.594	28.259	23.924	19.589	15.254	10.919	6.584	2.249	1.132	16.362	106.458	197.900	147.108	104.375	62.642		
	MAR	52.805	49.221	46.549	44.034	41.519	39.004	36.489	33.974	31.459	28.944																															



Tabla E.1. Continuación.

Table with columns for month/year (Año/Mes), day (Día), and 31 columns of meteorological data (GASTOS MEDIOS DIARIOS EN m³/h). The table is organized into sections for the years 1953, 1954, 1955, 1956, 1957, 1958, and 1959, with each section containing monthly data from January to December.

# Registros de Gastos medios diarios de la Estación Huites, Sinaloa

**Tabla E.1. Continuación.**

Año		ESTACION HUITES, SINALOA																																
		GASTOS MEDIOS DIARIOS EN m³																																
Año	Mes	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31		
1959	ENE	25.077	23.682	21.549	19.625	18.370	17.586	16.778	16.143	15.535	15.033	14.468	13.971	13.478	12.978	12.478	12.023	11.619	11.216	11.183	11.183	11.183	11.183	11.183	11.183	11.183	11.183	11.183	11.183	11.183	11.183	11.183		
	FEB	10.026	11.218	12.660	12.180	12.159	12.439	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459	12.459		
	MAR	23.687	22.363	21.195	20.056	19.823	19.599	19.375	19.151	18.927	18.703	18.479	18.255	18.031	17.807	17.583	17.359	17.135	16.911	16.687	16.463	16.239	16.015	15.791	15.567	15.343	15.119	14.895	14.671	14.447	14.223	14.000		
	ABR	6.817	6.770	6.338	5.807	5.850	5.498	4.937	4.561	4.562	4.185	3.722	3.346	2.970	2.594	2.218	1.842	1.466	1.090	0.714	0.338	0.338	0.338	0.338	0.338	0.338	0.338	0.338	0.338	0.338	0.338	0.338		
	MAY	5.650	4.729	4.418	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396	4.396		
	JUN	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150	4.150		
	JUL	217.064	242.132	270.227	228.023	188.994	211.630	165.765	167.333	133.750	193.291	167.833	136.955	224.457	179.589	171.269	159.033	136.955	136.955	136.955	136.955	136.955	136.955	136.955	136.955	136.955	136.955	136.955	136.955	136.955	136.955	136.955		
	AGO	104.992	124.297	189.853	216.199	208.190	411.472	374.465	470.170	346.875	305.065	413.118	167.580	172.580	169.580	169.580	169.580	169.580	169.580	169.580	169.580	169.580	169.580	169.580	169.580	169.580	169.580	169.580	169.580	169.580	169.580	169.580	169.580	
	SEPT	37.186	28.219	296.872	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	186.737	
	OCT	37.123	35.268	1827.479	877.089	391.588	242.755	174.480	130.935	107.708	82.824	69.338	60.434	54.038	49.037	49.019	45.998	41.988	37.217	34.321	32.550	30.789	27.482	23.104	21.950	20.049	18.789	18.221	17.411	16.875	16.384	15.894	15.404	
	NOV	50.093	46.890	41.459	32.742	27.174	22.092	19.862	18.739	18.092	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	18.043	
	DIC	15.620	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	15.159	
1960	ENE	218.537	169.138	128.355	745.935	600.643	173.933	172.518	524.194	430.363	309.622	228.262	869.811	183.728	245.914	136.000	670.088	670.088	670.088	670.088	670.088	670.088	670.088	670.088	670.088	670.088	670.088	670.088	670.088	670.088	670.088	670.088		
	FEB	38.012	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	37.456	
	MAR	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	
	ABR	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985	13.985
	MAY	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	5.987	
	JUN	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	3.290	
	JUL	15.095	15.475	11.022	36.775	48.492	60.032	113.954	121.038	258.718	181.034	202.911	210.037	280.084	237.341	289.088	230.540	338.928	219.502	137.350	178.252	337.977	204.279	243.965	394.645	460.844	521.666	366.436	269.531	234.242	200.340	169.531	139.380	
	AGO	225.614	184.771	211.511	143.119	144.239	228.300	236.791	359.462	426.583	595.018	847.914	481.388	611.954	481.388	611.954	481.388	611.954	481.388	611.954	481.388	611.954	481.388	611.954	481.388	611.954	481.388	611.954	481.388	611.954	481.388	611.954	481.388	
	SEPT	65.614	398.843	286.966	234.452	213.641	283.931	441.408	881.735	660.980	589.389	408.935	296.326	220.520	162.872	127.822	104.853	86.262	66.262	46.262	26.262	6.262	26.262	26.262	26.262	26.262	26.262	26.262	26.262	26.262	26.262	26.262	26.262	
	OCT	32.605	33.349	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	34.398	
	NOV	22.181	20.312	18.654	18.158	17.306	16.854	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	15.988	
	DIC	15.517	14.269	13.314	12.865	11.134	10.089	9.519	9.769	9.672	15.637	16.169	19.439	30.210	31.370	25.859	23.188	19.928	16.351	16.249	16.300	16.850	16.850	16.850	16.850	16.850	16.850	16.850	16.850	16.850	16.850	16.850	16.850	
1961	ENE	435.988	458.963	195.027	95.023	64.775	57.575	45.164	44.897	54.903	51.454	47.337	42.790	39.150	34.650	31.519	28.235	26.426	32.606	185.301	631.869	872.025	878.314	894.657	1314.925	1044.741	664.765	430.202	305.718	230.483	183.219	150.778		
	FEB	122.999	103.222	89.128	73.161	65.217	59.615	52.226	45.887	42.612	40.615	38.964	36.029	33.400	32.875	30.925	27.768	27.886	28.018	27.136	26.489	25.821	24.665	24.018	23.366	22.798	22.227	21.661	21.095	20.529	19.963	19.397		
	MAR	20.638	20.668	20.230	20.168	19.590	19.057	17.938	17.587	17.884	18.598	15.742	15.699	15.699	15.265	14.187	14.132	13.551	13.266	13.211	12.451	12.389	12.389	12.389	12.389	12.389	12.389	12.389	12.389	12.389	12.389	12.389	12.389	
	ABR	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	11.600	
	MAY	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	6.684	
	JUN	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	
	JUL	18.618	24.618	34.618	44.618	54.618	64.618	74.618	84.618	94.618	104.618	114.618	124.618	134.618	144.618	154.618	164.618	174.618	184.618	194.618	204.618	214.618	224.618	234.618	244.618	254.618	264.618	274.618	284.618	294.618	304.618	314.618	324.618	
	AGO	349.551	313.034	244.634	203.863	165.348	129.827	95.308	60.787	26.266	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	3.189	
	SEPT	315.531	282.384	205.760	251.488	415.348	615.348	815.348	1015.348	1215.348	1415.348	1615.348	1815.348	2015.348	2215.348	2415.348	2615																	





# Registros de Gastos medios diarios de la Estación Huites, Sinaloa

**Tabla E.1. Continuación.**

ESTACION HUITES, SINALOA																																														
GASTOS MEDIOS DIARIOS EN m <sup>3</sup> /E																																														
Año	Mes	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31														
1971	ENE	25.692	25.227	24.399	24.533	24.283	23.527	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930	24.930												
	FEB	9.149	9.077	8.886	8.630	8.432	8.374	7.763	7.597	7.406	7.207	7.027	6.857	6.687	6.517	6.347	6.177	6.007	5.837	5.667	5.497	5.327	5.157	4.987	4.817	4.647	4.477	4.307	4.137	3.967	3.797	3.627	3.457	3.287												
	MAR	5.137	5.039	4.914	4.873	4.729	4.741	4.705	4.651	4.611	4.537	4.471	4.405	4.341	4.275	4.210	4.144	4.078	4.012	3.946	3.880	3.814	3.748	3.682	3.616	3.550	3.484	3.418	3.352	3.286	3.220	3.154	3.088	3.022												
	ABR	3.697	3.698	3.614	3.572	3.473	3.453	3.396	3.456	3.473	3.406	3.347	3.289	3.231	3.173	3.115	3.057	3.000	2.942	2.884	2.826	2.768	2.710	2.652	2.594	2.536	2.478	2.420	2.362	2.304	2.246	2.188	2.130	2.072	2.014											
	MAY	2.627	2.453	2.408	2.227	2.210	2.208	2.222	2.210	2.208	2.206	2.204	2.202	2.200	2.198	2.196	2.194	2.192	2.190	2.188	2.186	2.184	2.182	2.180	2.178	2.176	2.174	2.172	2.170	2.168	2.166	2.164	2.162	2.160	2.158	2.156										
	JUN	2.956	3.056	3.046	3.054	3.055	3.056	3.057	3.058	3.059	3.060	3.061	3.062	3.063	3.064	3.065	3.066	3.067	3.068	3.069	3.070	3.071	3.072	3.073	3.074	3.075	3.076	3.077	3.078	3.079	3.080	3.081	3.082	3.083	3.084	3.085	3.086									
	JUL	4.780	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682	4.682									
	AGO	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193	1.193								
	SEPT	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015	2.015								
	OCT	5.771	5.996	6.021	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978	5.978								
	NOV	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183	3.183							
	DIC	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261	4.261							
1972	ENE	45.549	37.611	39.760	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716	38.716							
FEB	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061						
MAR	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061	42.061					
ABR	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901	10.901					
MAY	5.084	4.987	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924	4.924					
JUN	11.438	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545	11.545					
JUL	10.170	9.242	10.737	9.531	9.472	7.503	63.709	62.498	70.110	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450	127.450					
AGO	87.268	103.963	336.240	464.404	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724	450.724					
SEPT	535.689	413.003	391.130	424.936	450.430	420.795	446.866	404.489	500.424	566.376	522.273	719.251	830.052	810.735	943.836	804.813	589.916	650.778	622.532	547.844	598.577	571.057	598.060	281.256	95.641	150.453	136.148	116.150	136.148	116.150	136.148	116.150	136.148	116.150	136.148	116.150	136.148	116.150	136.148	116.150	136.148	116.150				
OCT	102.278	89.204	78.992	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340	80.340				
NOV	961.932	615.043	412.610	294.844	228.537	183.318	148.819	123.333	107.873	101.322	93.749	86.821	80.477	74.211	68.141	62.065	55.989	50.913	45.837	40.761	35.685	30.609	25.533	20.457	15.381	10.305	5.229	0.153	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621				
DIC	167.243	140.966	122.750	109.780	104.010	100.780	99.135	95.806	90.694	77.713	70.731	68.141	62.065	55.989	50.913	45.837	40.761	35.685	30.609	25.533	20.457	15.381	10.305	5.229	0.153	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621	159.621			
1973	ENE	73.041	90.362	142.460	167.362	161.369	488.206	1985.415	1571.302	841.473	569.362	418.106	324.252	254.852	207.739	172.296	145.812	121.617	107.186	94.688	85.637	80.250	76.105	72.626	68.012	63.921	60.932	58.348	56.359	54.048	51.723	49.401	47.079	44.757	42.435	40.113	37.791	35.469	33.147	30.825	28.503	26.181	23.859			
FEB	54.657	60.700	58.988	54.924	52.698	50.074	49.250	55.194	344.377	390.851	292.738	214.906	190.581	128.318	104.414	87.547	75.769	68.012	60.950	70.148	518.957	5501.070	2407.948	1305.277	675.513	530.789	423.034	316.282	209.449	152.262	105.492	58.742	12.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422
MAR	345.159	312.280	286.745	261.819	245.737	223.746	208.478	192.285	176.367	169.882	173.759	174.208	171.420	163.688	208.286	194.251	164.589	146.378	129.588	116.378	103.168	90.448	77.726	65.004	52.282	39.560	26.838	14.116	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422	1.422
ABR	74.530	72.804	71.588	69.588	67.590	65.592	63.594	61.596	59.598																																					





# Registros de Gastos medios diarios de la Estación Huites, Sinaloa

## Tabla E.1. Continuación.

ESTACION HUITES, SINALOA																																			
GASTOS MEDIOS DIARIOS EN m³																																			
Año	Mes	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31			
1983	ENE	590.066	458.723	313.980	236.137	187.135	168.553	148.442	133.954	120.380	112.266	98.211	89.117	80.428	86.103	95.489	102.479	92.827	84.710	80.428	96.103	140.152	188.884	170.550	152.275	123.016	116.254	105.989	97.857	90.126	434.818				
	FEB	49.434	314.700	385.942	1038.411	1133.377	907.667	582.498	378.528	276.755	216.859	188.222	155.729	136.213	114.387	103.388	95.273	85.372	78.284	70.284	65.302	59.211	56.000	53.617	51.609	53.623	53.617	56.000	59.211	63.965	59.241	53.623	46.485	43.782	
	MAR	49.874	706.377	6478.583	1058.333	2066.184	1206.181	936.951	773.830	620.249	489.420	383.520	274.915	222.436	162.033	110.393	66.315	43.602	30.263	14.034	12.924	10.923	9.635	8.320	7.263	6.395	5.924	5.303	4.853	4.446	4.085	3.742	3.419	3.107	
	ABR	41.650	20.571	37.871	35.430	34.091	41.197	36.433	34.091	31.106	28.850	26.433	24.016	21.600	19.183	16.767	14.350	11.934	9.517	7.100	4.683	2.266	1.851	1.436	1.021	0.606	0.191	0.777	1.162	1.547	1.932	2.317	2.702	3.087	
	MAY	20.851	20.639	21.871	18.956	18.630	14.018	15.533	15.368	13.743	13.008	12.262	11.527	10.792	10.057	9.322	8.587	7.852	7.117	6.382	5.647	4.912	4.177	3.442	2.707	1.972	1.237	0.502	0.765	1.028	1.291	1.554	1.817	2.080	
	JUN	26.058	24.826	16.378	15.532	13.952	13.952	17.811	16.479	15.487	14.500	13.513	12.527	11.541	10.555	9.569	8.583	7.597	6.611	5.625	4.639	3.653	2.667	1.681	0.695	0.709	0.723	0.737	0.751	0.765	0.779	0.793	0.807	0.821	
	JUL	11.054	10.290	9.711	8.875	8.856	9.189	12.538	13.968	13.594	12.706	11.819	10.932	10.045	9.158	8.271	7.384	6.497	5.610	4.723	3.836	2.949	2.062	1.175	0.288	0.401	0.514	0.627	0.740	0.853	0.966	1.079	1.192	1.305	
	AGO	160.088	143.413	115.201	89.166	110.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008
	SEPT	177.198	163.475	125.202	118.562	110.504	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008	104.026	105.665	121.008
	OCT	154.108	169.834	127.628	101.024	83.987	70.280	62.738	68.388	59.282	50.885	42.732	34.580	26.428	18.276	10.124	1.972	0.820	0.668	0.516	0.364	0.212	0.060	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	NOV	104.130	87.676	71.278	75.549	213.944	238.382	136.827	148.348	114.982	98.073	86.592	75.025	63.460	51.899	40.338	28.778	17.220	5.660	3.100	0.540	0.080	0.020	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	DIC	97.684	379.550	1280.640	770.108	425.223	259.720	193.945	152.771	123.334	105.555	97.154	88.521	79.887	71.253	62.620	54.000	45.380	36.760	28.140	19.520	10.900	2.280	0.660	0.140	0.030	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1984	ENE	22.689	22.682	28.863	31.201	30.721	29.870	28.863	26.832	24.803	22.774	20.745	18.716	16.687	14.658	12.629	10.600	8.571	6.542	4.513	2.484	0.455	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	FEB	19.301	14.898	15.392	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898
	MAR	17.310	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898	14.898
	ABR	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	7.142	
	MAY	10.972	6.155	6.002	6.212	6.881	7.817	6.881	6.155	6.002	6.212	6.881	7.817	6.881	6.155	6.002	6.212	6.881	7.817	6.881	6.155	6.002	6.212	6.881	7.817	6.881	6.155	6.002	6.212	6.881	7.817	6.881	6.155	6.002	6.212
	JUN	10.770	11.252	12.456	36.779	25.391	13.856	15.855	17.854	18.853	16.212	14.151	12.526	10.946	9.325	7.704	6.083	4.462	2.841	1.220	0.600	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	JUL	844.407	395.244	633.664	404.263	460.335	644.934	677.890	651.485	624.124	685.812	687.346	594.900	447.359	283.656	163.422	63.422	574.162	646.163	763.541	722.905	655.663	550.776	399.410	434.035	516.856	567.488	522.488	468.957	424.702	379.792	334.421	288.847	243.476	
	AGO	328.306	237.215	216.843	275.721	297.063	321.414	460.239	381.668	157.355	165.753	195.110	1068.308	1433.185	1250.246	965.400	799.893	639.473	578.437	551.649	658.275	720.156	785.801	713.228	500.243	645.270	892.999	850.700	441.714	344.421	278.813	211.621	145.922	80.845	
	SEP	415.986	423.051	502.905	472.243	332.291	257.800	203.314	204.344	416.814	533.052	382.265	334.312	271.378	217.147	158.016	105.065	48.369	32.307	10.542	4.300	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	OCT	102.580	100.096	100.096	167.222	164.062	89.864	83.964	83.964	101.717	60.058	45.346	34.300	24.422	16.858	10.360	5.406	2.442	1.104	0.488	0.200	0.088	0.042	0.020	0.010	0.005	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	NOV	31.420	29.615	26.947	23.061	21.749	20.330	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899
	DIC	31.420	29.615	26.947	23.061	21.749	20.330	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899	19.899
1985	ENE	691.102	462.817	350.948	289.000	217.135	186.213	160.767	142.917	127.828	115.866	102.738	89.483	76.227	63.000	49.732	36.465	22.900	11.300	4.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	FEB	374.302	381.325	343.422	330.964	321.708	302.988	274.312	244.351	218.479	201.940	181.902	165.181	150.247	138.791	128.275	116.380	105.645	91.215	85.385	82.407	80.936	78.860	74.054	69.248	64.442	59.636	54.830	49.924	45.118	40.312	35.506	30.700	25.894	
	MAR	53.957	54.954	54.683	51.661	46.689	43.746	42.337	41.854	40.583	38.779	37.443	35.788	34.283	32.778	31.273	29.768	28.263	26.758	25.253	23.748	22.243	20.738	19.233	17.728	16.223	14.718	13.213	11.708	10.203	8.698	7.193	5.688	4.183	
	ABR	92.704	71.127	55.376	47.507	41.828	37.412	33.882	30.483	27.128	23.772	20.416	17.060	13.704	10.348	6.992	3.636	0.280	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	MAY	21.415	19.484	18.817	18.330	18.854	19.377	19.900	20.423	20.946	21.469	21.992	22.515	23.038	23.561	24.084	24.607	25.130	25.653	26.176	26.700	27.223	27.746	28.269											



# Apéndice F

## Análisis de Frecuencias Univariado de Avenidas

### F.1. Introducción

Las distribuciones más apropiadas que representan a las variables Q y V, se han elegido con base en un análisis de ajuste de funciones de distribución de probabilidad (FDP). El ajuste se realizó con el programa AX.exe. [Jiménez (1997)], y cuenta con una paquetería de hasta seis FDP: Normal, Log Normal 2-3P, Gamma 2-3P, Exponencial, Gumbel y Doble Gumbel 5p. Se calcula el Error Estándar de Ajuste (EEA) [Kite (1988)] para comparar el ajuste realizado a las series con otros modelos de FDP. En las Figuras F.1 a F.12, se presentan las curvas de mejor ajuste a los datos observados de las estaciones de la cuenca del Rio Coatzacoalcos (Las Perlas, Jesús Carranza II, Tierra Morada y Minzapan) y la cuenca del Río Fuerte (Huites).

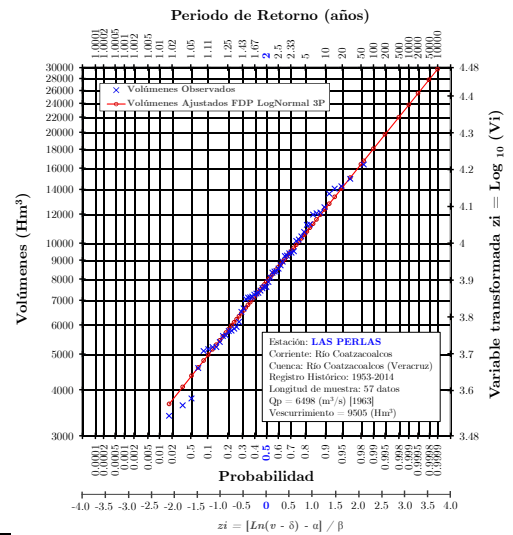
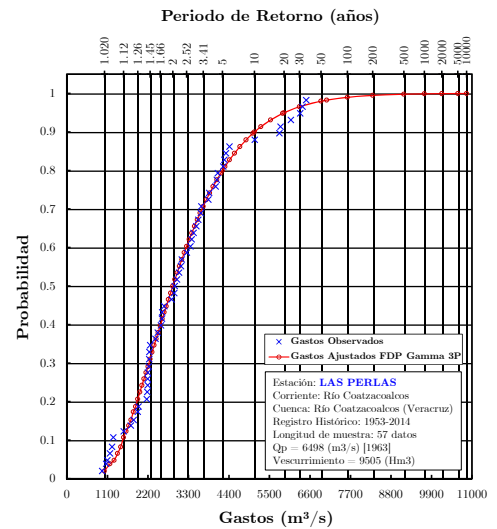
El procedimiento del Análisis de frecuencias univariado de avenidas consiste para las variables Q y V en: organizar los datos de forma ascendente; asignar un periodo de retorno (en años) por medio de la Ley empírica de Weibull con clasificación ascendente en sus datos,  $T = (n + 1) / [(n + 1) - k_i]$ ; calcular la probabilidad de no excedencia,  $P = 1 - (1/T) = k_i / (n + 1)$ ; ajustar diferentes FDP a las series y obtener los cuantiles q y v; obtener eventos de diseño para diferentes T.

Tabla F.1. Análisis de Frecuencia Univariado

ORD k	ESTACIÓN LAS PERLAS (VERACRUZ)					
	DATOS OBSERVADOS		PER. RET.	PROB.	GAMMA 3P	LOG - NORMAL 3P
	Q <sub>ASC.</sub> m <sup>3</sup> /s	V <sub>ASC.</sub> Hm <sup>3</sup>	T años	P(X ≤ x) 1 - 1/T	Q <sub>CUANTIL</sub> m <sup>3</sup> /s	V <sub>CUANTIL</sub> Hm <sup>3</sup>
1	958.511	3398.817	1.0175	0.0172	1017.160	3657.590
2	1065.888	3627.363	1.0357	0.0345	1168.190	4072.370
3	1098.361	3786.419	1.0545	0.0517	1281.240	4362.740
4	1172.337	4581.725	1.0741	0.0690	1376.130	4597.540
5	1228.784	5094.033	1.0943	0.0862	1460.260	4799.890
6	1261.411	5157.358	1.1154	0.1034	1537.160	4980.610
7	1542.900	5219.094	1.1373	0.1207	1608.760	5146.210
8	1738.249	5220.917	1.1600	0.1379	1676.510	5300.630
9	1815.312	5392.886	1.1837	0.1552	1741.100	5446.290
10	1928.050	5597.068	1.2083	0.1724	1803.470	5585.190
11	1947.374	5613.939	1.2340	0.1897	1863.840	5718.650
12	2168.700	5771.812	1.2609	0.2069	1922.820	5847.660
13	2181.895	5786.891	1.2889	0.2241	1980.500	5973.240
14	2197.403	5857.019	1.3182	0.2414	2037.290	6095.840
15	2198.073	5922.941	1.3488	0.2586	2093.310	6216.150
16	2200.750	6115.796	1.3810	0.2759	2148.760	6334.630
17	2226.037	6504.538	1.4146	0.2931	2203.660	6451.650
18	2235.041	6651.430	1.4500	0.3103	2258.420	6567.560
19	2239.448	7046.819	1.4872	0.3276	2313.020	6682.750
20	2258.943	7113.683	1.5263	0.3448	2367.540	6797.310
21	2406.000	7127.026	1.5676	0.3621	2422.140	6911.700
22	2470.312	7156.266	1.6111	0.3793	2476.970	7026.130
23	2555.979	7247.795	1.6571	0.3966	2532.030	7140.880
24	2582.860	7298.026	1.7059	0.4138	2587.550	7256.050
25	2586.347	7332.841	1.7576	0.4310	2643.310	7371.920
26	2644.343	7430.401	1.8125	0.4483	2699.750	7488.640
27	2845.812	7542.663	1.8710	0.4655	2756.930	7606.600
28	2916.014	7599.782	1.9333	0.4828	2814.870	7725.900
29	2929.631	7605.741	2	0.5000	2873.700	7846.830
30	2986.931	7845.769	2.0714	0.5172	2933.490	7969.550
31	3048.729	8055.928	2.1481	0.5345	2994.440	8094.480
32	3098.750	8313.473	2.2308	0.5517	3056.420	8221.900
33	3122.859	8372.399	2.3200	0.5690	3119.940	8351.940
34	3260.750	8390.354	2.4167	0.5862	3185.070	8485.140
35	3353.591	8526.117	2.5217	0.6034	3251.900	8621.760
36	3389.663	8705.360	2.6364	0.6207	3320.570	8762.290
37	3439.749	8913.535	2.7619	0.6379	3391.470	8907.050
38	3516.746	9225.439	2.9000	0.6552	3464.380	9056.550
39	3566.507	9253.425	3.0526	0.6724	3540.200	9211.370
40	3638.544	9383.771	3.2222	0.6897	3618.950	9372.410
41	3649.163	9419.645	3.4118	0.7069	3701.000	9540.200
42	3842.252	9504.804	3.6250	0.7241	3786.820	9715.760
43	3862.104	10144.786	3.8667	0.7414	3876.570	9900.160
44	4038.793	10230.013	4.1429	0.7586	3971.380	10094.650
45	4073.784	10437.004	4.4615	0.7759	4071.820	10300.820
46	4101.375	10730.080	4.8333	0.7931	4178.710	10520.840
47	4245.799	11217.620	5.2727	0.8103	4292.980	10756.810
48	4274.984	11251.216	5.8000	0.8276	4416.490	11012.230
49	4319.889	11959.499	6.4444	0.8448	4551.220	11291.200
50	4417.460	12055.792	7.2500	0.8621	4698.870	11599.220
51	5099.650	12115.124	8.2857	0.8793	4864.050	11944.540
52	5769.958	12521.900	9.6667	0.8966	5051.230	12338.310
53	5798.591	13621.620	11.6000	0.9138	5269.010	12798.550
54	6082.677	14061.219	14.5000	0.9310	5530.010	13356.140
55	6336.781	14272.462	19.3333	0.9483	5858.840	14066.930
56	6399.943	14949.500	29	0.9655	6311.020	15057.290
57	6497.666	16356.012	58	0.9828	7054.810	16739.270

Tabla F.2. Eventos de diseño

ESTACIÓN LAS PERLAS			
PER. RET.	PROB.	GAMMA 3P	LOG - NORMAL 3P
T	P(X ≤ x)	Q	V
años	1 - 1/T	m <sup>3</sup> /s	Hm <sup>3</sup>
2	0.5	2873.70	7846.83
5	0.8	4223.62	10613.07
10	0.9	5091.99	12424.35
20	0.95	5897.29	14149.50
50	0.98	6898.54	16379.39
100	0.99	7617.10	18056.46
200	0.995	8302.45	19740.22
500	0.998	9146.51	21994.47
1000	0.999	9707.58	23748.20
2000	0.9995	10180.06	25534.26
5000	0.9998	10613.17	27762.13
10000	0.9999	10849.41	29683.18



**Tabla F.3.** Análisis de Frecuencia Univariado

ESTACIÓN JESÚS CARRANZA II (VERACRUZ)						
ORD	DATOS OBSERVADOS		PER. RET.	PROB.	LOG - NORMAL 3P	GUMBEL
k	Q <sub>ASC.</sub> m <sup>3</sup> /s	V <sub>ASC.</sub> Hm <sup>3</sup>	T años	P(X ≤ x) 1 - 1/T	Q <sub>CUANTIL</sub> m <sup>3</sup> /s	V <sub>CUANTIL</sub> Hm <sup>3</sup>
1	228.800	136.322	1.0192	0.0189	265.200	119.950
2	239.622	141.249	1.0392	0.0377	315.650	163.410
3	316.693	144.191	1.0600	0.0566	352.730	193.330
4	331.694	201.572	1.0816	0.0755	383.740	217.240
5	371.301	219.770	1.1042	0.0943	411.190	237.690
6	411.467	230.696	1.1277	0.1132	436.320	255.900
7	463.000	235.820	1.1522	0.1321	459.820	272.520
8	504.806	262.742	1.1778	0.1509	482.140	287.980
9	513.018	264.317	1.2045	0.1698	503.570	302.550
10	529.254	290.698	1.2326	0.1887	524.340	316.420
11	549.551	292.760	1.2619	0.2075	544.620	329.750
12	563.437	295.910	1.2927	0.2264	564.510	342.650
13	609.083	316.397	1.3250	0.2453	584.140	355.200
14	656.626	363.241	1.3590	0.2642	603.610	367.470
15	662.346	364.398	1.3947	0.2830	622.960	379.520
16	679.601	399.939	1.4324	0.3019	642.260	391.410
17	685.447	416.056	1.4722	0.3208	661.580	403.170
18	709.586	421.695	1.5143	0.3396	680.990	414.860
19	742.389	459.628	1.5588	0.3585	700.520	426.490
20	757.476	478.026	1.6061	0.3774	720.210	438.110
21	791.234	478.387	1.6563	0.3962	740.130	449.740
22	843.185	479.946	1.7097	0.4151	760.310	461.420
23	844.249	528.853	1.7667	0.4340	780.820	473.170
24	844.267	535.701	1.8276	0.4528	801.680	485.030
25	863.267	537.973	1.8929	0.4717	822.960	497.010
26	919.583	540.811	1.9630	0.4906	844.720	509.150
27	929.625	551.801	2.0385	0.5094	866.980	521.480
28	954.854	556.790	2.1200	0.5283	889.880	534.040
29	968.670	563.598	2.2083	0.5472	913.420	546.850
30	1021.627	566.359	2.3043	0.5660	937.730	559.960
31	1032.827	572.835	2.4091	0.5849	962.880	573.400
32	1034.651	628.348	2.5238	0.6038	988.950	587.220
33	1035.372	630.104	2.6500	0.6226	1016.070	601.480
34	1043.875	650.001	2.7895	0.6415	1044.330	616.220
35	1065.129	662.981	2.9444	0.6604	1073.940	631.510
36	1074.025	666.248	3.1176	0.6792	1105.030	647.440
37	1091.737	668.909	3.3125	0.6981	1137.780	664.080
38	1099.945	703.318	3.5333	0.7170	1172.460	681.550
39	1158.125	710.303	3.7857	0.7358	1209.370	699.960
40	1165.443	720.520	4.0769	0.7547	1248.870	719.470
41	1178.874	733.054	4.4167	0.7736	1291.350	740.280
42	1295.875	737.115	4.8182	0.7925	1337.390	762.600
43	1353.315	797.909	5.3	0.8113	1387.760	786.760
44	1387.100	812.417	5.8889	0.8302	1443.400	813.140
45	1387.380	833.070	6.6250	0.8491	1505.580	842.290
46	1464.520	885.963	7.5714	0.8679	1576.190	874.960
47	1610.187	911.077	8.8333	0.8868	1657.970	912.260
48	1903.280	969.758	10.6	0.9057	1755.180	955.900
49	1924.668	1027.261	13.2500	0.9245	1875.080	1008.750
50	1967.920	1124.572	17.6667	0.9434	2031.630	1076.170
51	2041.277	1223.053	26.5	0.9623	2256.350	1170.240
52	2908.776	1231.761	53	0.9811	2654.850	1329.440

**Tabla F.4.** Eventos de diseño

ESTACIÓN JESÚS CARRANZA II			
PER. RET.	PROB.	LOG - NORMAL 3P	GUMBEL
T años	P(X ≤ x) 1 - 1/T	Q m <sup>3</sup> /s	V Hm <sup>3</sup>
2	0.5	855.79	515.29
5	0.8	1356.98	772.03
10	0.9	1724.03	942.01
20	0.95	2099.70	1105.06
50	0.98	2620.37	1316.11
100	0.99	3036.50	1474.26
500	0.998	4090.62	1839.73
1000	0.999	4592.76	1996.84
2000	0.9995	5123.74	2153.91
5000	0.9998	5812.76	2361.49
10000	0.9999	6429.81	2518.51

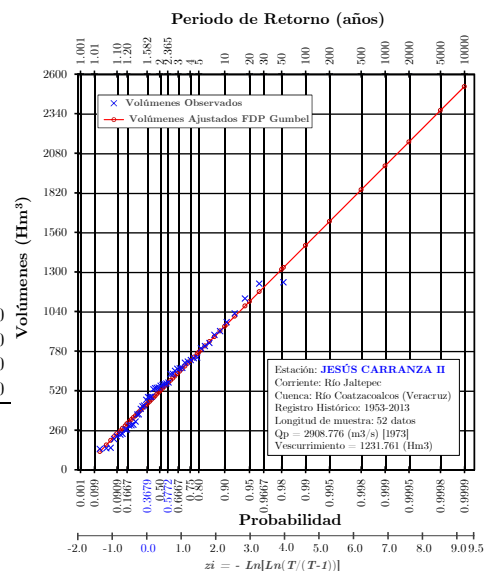
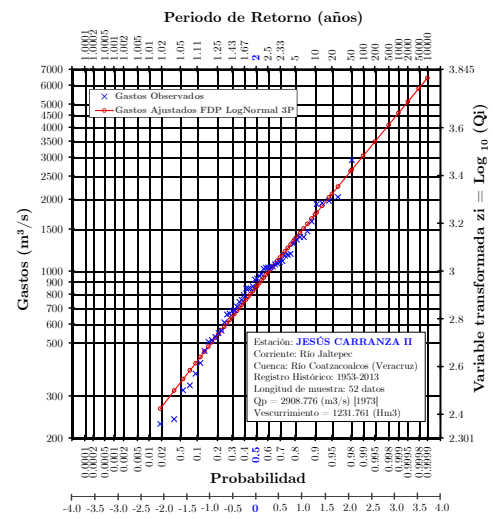


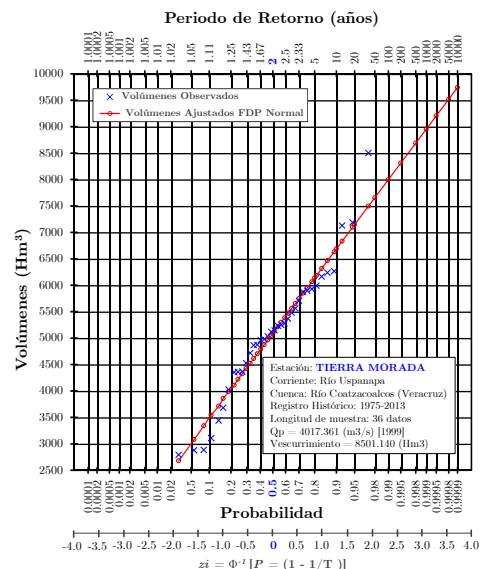
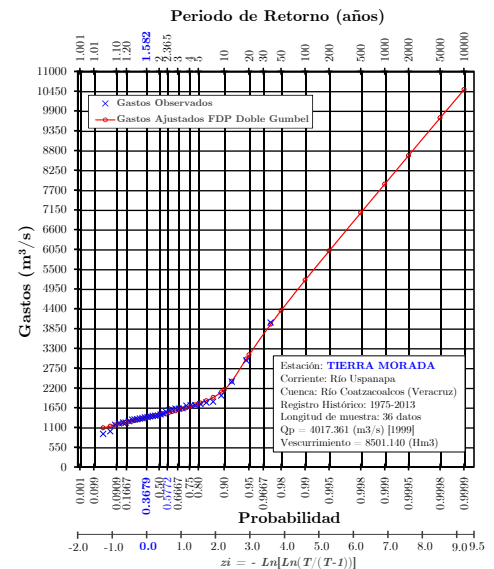


Tabla F.5. Análisis de Frecuencia Univariado

ESTACIÓN TIERRA MORADA (VERACRUZ)						
ORD	DATOS OBSERVADOS		PER. RET.	PROB.	DOBLE GUMBEL	NORMAL
k	Q <sub>ASC.</sub> m <sup>3</sup> /s	V <sub>ASC.</sub> Hm <sup>3</sup>	T años	P(X ≤ x) 1 - 1/T	Q <sub>CUANTIL</sub> m <sup>3</sup> /s	V <sub>CUANTIL</sub> Hm <sup>3</sup>
1	919.033	2790.896	1.0278	0.0270	1090.56	2678.12
2	993.897	2878.562	1.0571	0.0541	1137.13	3078.90
3	1171.189	2883.500	1.0882	0.0811	1168.86	3340.90
4	1213.470	3107.467	1.1212	0.1081	1194.44	3543.04
5	1234.824	3441.189	1.1563	0.1351	1216.63	3711.33
6	1235.790	3680.570	1.1935	0.1622	1236.70	3857.79
7	1271.945	4028.189	1.2333	0.1892	1255.35	3989.03
8	1317.131	4354.193	1.2759	0.2162	1273.02	4109.10
9	1322.793	4359.807	1.3214	0.2432	1290.01	4220.68
10	1329.181	4366.273	1.3704	0.2703	1306.53	4325.66
11	1347.556	4527.039	1.4231	0.2973	1322.76	4425.44
12	1356.309	4716.113	1.4800	0.3243	1338.83	4521.05
13	1379.116	4863.636	1.5417	0.3514	1354.86	4613.35
14	1387.614	4872.658	1.6087	0.3784	1370.95	4703.00
15	1409.545	4960.950	1.6818	0.4054	1387.22	4790.58
16	1419.625	4977.438	1.7619	0.4324	1403.75	4876.58
17	1427.064	5033.078	1.8500	0.4595	1420.64	4961.44
18	1429.465	5116.982	1.9474	0.4865	1438.02	5045.54
19	1458.875	5146.921	2.0556	0.5135	1456.00	5129.32
20	1485.977	5231.154	2.1765	0.5405	1474.73	5213.42
21	1527.332	5250.013	2.3125	0.5676	1494.35	5298.27
22	1574.334	5276.912	2.4667	0.5946	1515.07	5384.27
23	1600.253	5365.303	2.6429	0.6216	1537.14	5471.85
24	1624.675	5478.731	2.8462	0.6486	1560.85	5561.51
25	1625.988	5563.175	3.0833	0.6757	1586.60	5653.80
26	1633.918	5706.877	3.3636	0.7027	1614.95	5749.42
27	1704.797	5874.413	3.7	0.7297	1646.65	5849.19
28	1715.450	5893.518	4.1111	0.7568	1682.80	5954.18
29	1716.212	5928.174	4.6250	0.7838	1725.15	6065.76
30	1748.050	5989.554	5.2857	0.8108	1776.55	6185.83
31	1792.200	6164.973	6.1667	0.8378	1842.31	6317.06
32	1815.872	6237.837	7.4	0.8649	1933.80	6463.52
33	1987.238	6264.537	9.2500	0.8919	2081.21	6631.82
34	2374.517	7129.018	12.3333	0.9189	2384.92	6833.95
35	2961.060	7186.606	18.5	0.9459	3002.38	7095.96
36	4017.361	8501.140	37	0.9730	3969.09	7496.74

Tabla F.6. Eventos de diseño

ESTACIÓN TIERRA MORADA			
PER. RET.	PROB.	DOBLE GUMBEL	NORMAL
T años	P(X ≤ x) 1 - 1/T	Q m <sup>3</sup> /s	V Hm <sup>3</sup>
2	0.5	1446.93	5087.43
5	0.8	1754.63	6136.63
10	0.9	2147.64	6688.12
20	0.95	3119.94	7143.73
50	0.98	4351.80	7656.36
100	0.99	5195.21	7997.96
200	0.995	6010.33	8310.48
500	0.998	7068.61	8689.04
1000	0.999	7863.85	8954.50
2000	0.9995	8659.08	9205.16
5000	0.9998	9711.25	9517.38
10000	0.9999	10494.25	9741.21



**Tabla F.7.** Análisis de Frecuencia Univariado

ESTACIÓN MINZAPAN (VERACRUZ)						
ORD	DATOS OBSERVADOS		PER. RET.	PROB.	DOBLE GUMBEL	DOBLE GUMBEL
k	Q <sub>ASC.</sub> m <sup>3</sup> /s	V <sub>ASC.</sub> Hm <sup>3</sup>	T años	P(X ≤ x) 1 - 1/T	Q <sub>CUANTIL</sub> m <sup>3</sup> /s	V <sub>CUANTIL</sub> Hm <sup>3</sup>
1	52.738	12.643	1.0256	0.0250	56.61	12.07
2	63.021	13.351	1.0526	0.0500	64.70	14.53
3	64.639	14.761	1.0811	0.0750	70.41	16.29
4	70.540	14.912	1.1111	0.1000	75.06	17.74
5	73.302	16.405	1.1429	0.1250	79.11	19.02
6	80.424	18.229	1.1765	0.1500	82.79	20.20
7	82.235	18.874	1.2121	0.1750	86.20	21.31
8	87.736	20.501	1.2500	0.2000	89.44	22.38
9	91.131	25.710	1.2903	0.2250	92.55	23.42
10	91.247	26.142	1.3333	0.2500	95.57	24.45
11	92.768	26.328	1.3793	0.2750	98.53	25.47
12	94.206	27.607	1.4286	0.3000	101.46	26.50
13	99.005	27.781	1.4815	0.3250	104.37	27.55
14	109.505	28.777	1.5385	0.3500	107.28	28.62
15	116.796	31.176	1.6	0.3750	110.22	29.72
16	119.539	31.287	1.6667	0.4000	113.19	30.87
17	119.742	35.167	1.7391	0.4250	116.21	32.08
18	120.847	35.934	1.8182	0.4500	119.31	33.35
19	126.194	37.452	1.9048	0.4750	122.49	34.72
20	129.288	37.464	2	0.5000	125.79	36.18
21	131.166	38.272	2.1053	0.5250	129.22	37.75
22	141.514	39.357	2.2222	0.5500	132.81	39.43
23	142.282	39.990	2.3529	0.5750	136.59	41.19
24	145.523	41.322	2.5	0.6000	140.61	42.98
25	152.583	45.892	2.6667	0.6250	144.91	44.77
26	156.555	46.713	2.8571	0.6500	149.56	46.52
27	164.662	47.734	3.0769	0.6750	154.62	48.23
28	172.440	48.083	3.3333	0.7000	160.21	49.92
29	175.553	50.598	3.6364	0.7250	166.46	51.60
30	178.910	51.616	4	0.7500	173.54	53.30
31	180.764	51.685	4.4444	0.7750	181.69	55.04
32	186.558	54.658	5	0.8000	191.17	56.86
33	195.031	57.392	5.7143	0.8250	202.34	58.79
34	197.514	61.518	6.6667	0.8500	215.53	60.91
35	206.141	61.925	8	0.8750	231.13	63.28
36	236.153	64.333	10	0.9000	249.64	66.06
37	283.417	65.436	13.3333	0.9250	272.22	69.48
38	309.789	76.034	20	0.9500	301.88	74.11
39	336.925	83.327	40	0.9750	348.85	81.73

**Tabla F.8.** Eventos de diseño

ESTACIÓN MINZAPAN			
PER. RET.	PROB.	DOBLE GUMBEL	DOBLE GUMBEL
T	P(X ≤ x)	Q	V
años	1 - 1/T	m <sup>3</sup> /s	Hm <sup>3</sup>
2	0.5	125.79	36.18
5	0.8	191.17	56.86
10	0.9	249.64	66.06
20	0.95	301.88	74.11
50	0.98	363.36	84.14
100	0.99	407.45	91.53
200	0.995	450.72	98.85
500	0.998	507.27	108.47
1000	0.999	550.00	115.75
2000	0.9995	592.73	123.03
5000	0.9998	649.70	132.74
10000	0.9999	689.44	139.97

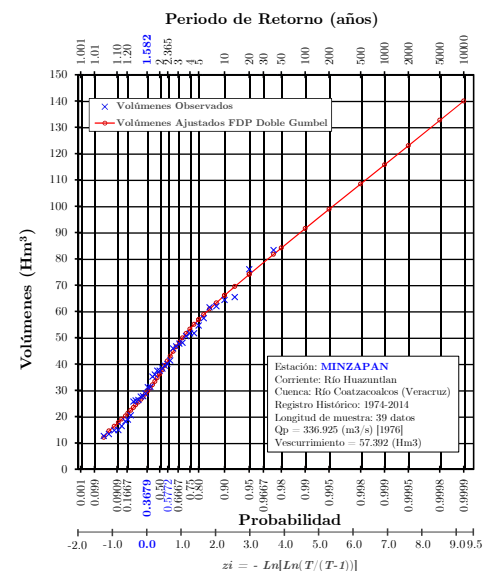
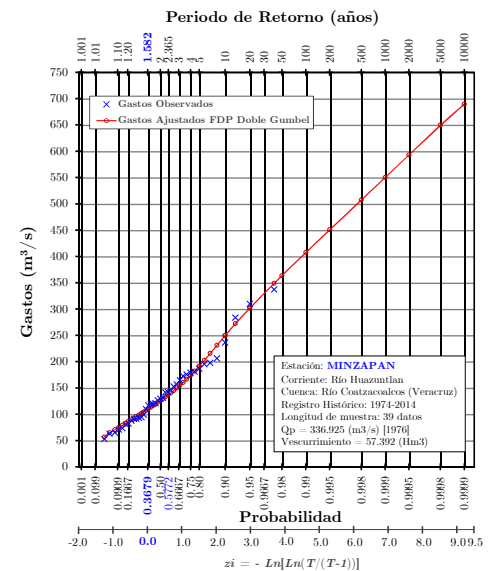




Tabla F.9. Análisis de Frecuencia Univariado

ESTACIÓN MINATITLAN (VERACRUZ)						
ORD	DATOS OBSERVADOS		PER. RET.	PROB.	GUMBEL	LOG - NORMAL 3P
k	Q <sub>ASC.</sub> m <sup>3</sup> /s	V <sub>ASC.</sub> Hm <sup>3</sup>	T años	P(X ≤ x) 1 - 1/T	Q <sub>CUANTIL</sub> m <sup>3</sup> /s	V <sub>CUANTIL</sub> Hm <sup>3</sup>
1	1675.660	6704.137	1.0303	0.0294	1985.540	7678.720
2	2014.331	7893.388	1.0625	0.0588	2257.180	8661.050
3	2032.395	8157.220	1.0968	0.0882	2448.870	9316.190
4	2627.356	10557.501	1.1333	0.1176	2605.420	9829.580
5	2628.518	10998.466	1.1724	0.1471	2742.100	10262.420
6	2731.955	11271.250	1.2143	0.1765	2866.150	10643.320
7	3024.572	11272.745	1.2593	0.2059	2981.660	10988.440
8	3145.874	11453.083	1.3077	0.2353	3091.220	11307.410
9	3157.922	11719.436	1.3600	0.2647	3196.610	11607.040
10	3650.161	11807.150	1.4167	0.2941	3299.120	11891.520
11	3650.578	12025.152	1.4783	0.3235	3399.760	12164.800
12	3750.334	12093.419	1.5455	0.3529	3499.350	12429.380
13	3832.553	12313.410	1.6190	0.3824	3598.610	12687.440
14	3924.047	12792.656	1.7000	0.4118	3698.170	12940.910
15	3926.905	12986.947	1.7895	0.4412	3798.640	13191.330
16	3960.814	13368.214	1.8889	0.4706	3900.610	13440.090
17	4082.258	13431.360	2	0.5000	4004.680	13688.640
18	4097.641	13551.066	2.1250	0.5294	4111.500	13938.270
19	4153.879	13633.401	2.2667	0.5588	4221.770	14190.590
20	4193.450	14644.227	2.4286	0.5882	4336.300	14447.040
21	4642.538	14736.010	2.6154	0.6176	4456.020	14709.140
22	4729.402	15148.540	2.8333	0.6471	4582.040	14978.650
23	4941.280	15261.417	3.0909	0.6765	4715.730	15257.790
24	5110.423	15771.178	3.4	0.7059	4858.810	15549.140
25	5156.578	16455.574	3.7778	0.7353	5013.540	15855.800
26	5275.151	16606.251	4.2500	0.7647	5182.900	16182.530
27	5285.819	16613.017	4.8571	0.7941	5371.190	16534.660
28	5375.723	17211.417	5.6667	0.8235	5584.190	16920.750
29	5632.659	17238.013	6.8	0.8529	5831.620	17353.160
30	5982.003	17271.696	8.5	0.8824	6129.130	17852.710
31	6503.965	18129.470	11.3333	0.9118	6506.170	18456.860
32	7037.829	20107.457	17	0.9412	7028.890	19246.710
33	7613.413	22075.669	34	0.9706	7908.160	20472.600

Tabla F.10. Eventos de diseño

ESTACIÓN MINATITLAN			
PER. RET.	PROB.	GUMBEL	LOG - NORMAL 3P
T	P(X ≤ x)	Q	V
años	1 - 1/T	m <sup>3</sup> /s	Hm <sup>3</sup>
2	0.5	4004.68	13688.64
5	0.8	5411.47	16608.79
10	0.9	6342.90	18198.95
20	0.95	7236.34	19546.44
50	0.98	8392.81	21102.25
100	0.99	9259.42	22162.04
200	0.995	10122.87	23148.07
500	0.998	11262.03	24365.80
1000	0.999	12122.97	25244.46
2000	0.9995	12983.61	26086.39
5000	0.9998	14121.09	27071.57
10000	0.9999	14981.48	27870.45

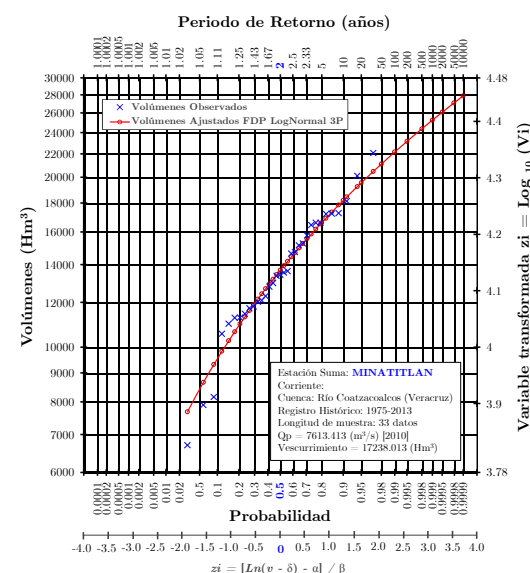
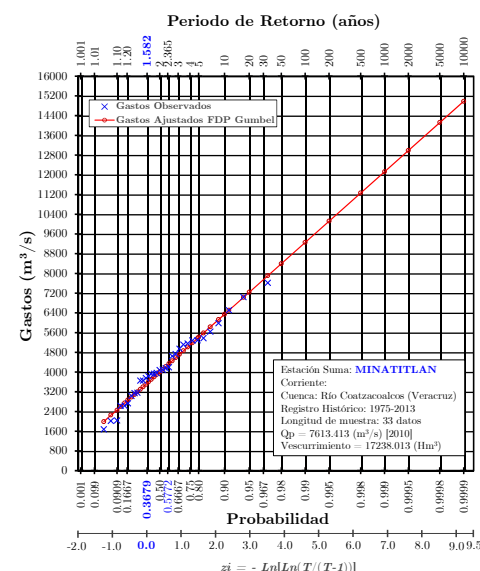
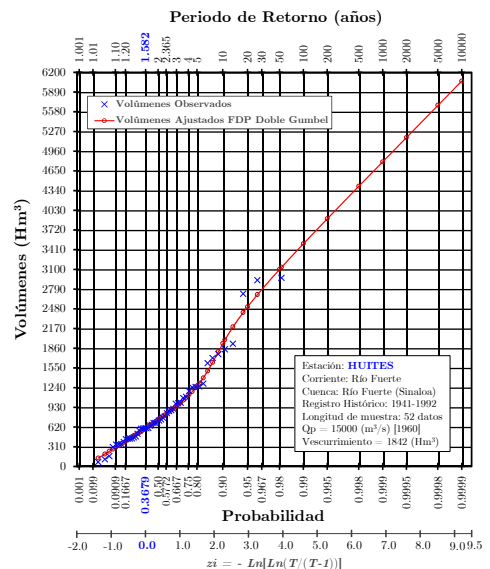
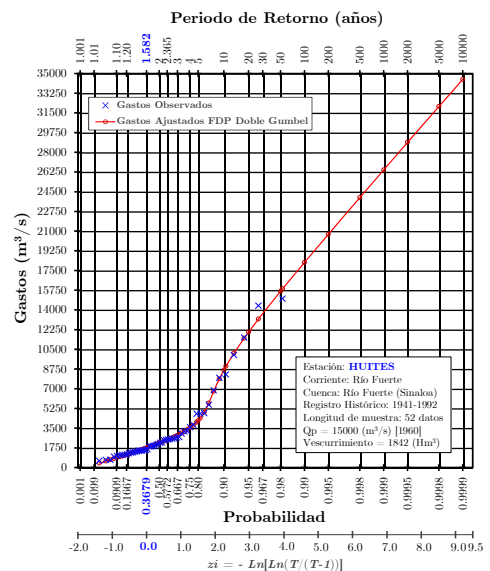


Tabla. F.11. Análisis de Frecuencia Univariado

ESTACIÓN HUITES (SINALOA)						
ORD	DATOS OBSERVADOS		PER. RET.	PROB.	DOBLE GUMBEL	DOBLE GUMBEL
k	Q <sub>ASC.</sub> m <sup>3</sup> /s	V <sub>ASC.</sub> Hm <sup>3</sup>	T años	P(X ≤ x) 1 - 1/T	Q <sub>CUANTIL</sub> m <sup>3</sup> /s	V <sub>CUANTIL</sub> Hm <sup>3</sup>
1	593	69	1.0192	0.0189	396.980	132.190
2	677	111	1.0392	0.0377	576.000	194.840
3	696	163	1.0600	0.0566	700.330	238.220
4	955	305	1.0816	0.0755	800.450	273.070
5	1025	344	1.1042	0.0943	886.720	303.010
6	1095	351	1.1277	0.1132	964.010	329.780
7	1127	352	1.1522	0.1321	1035.050	354.330
8	1135	384	1.1778	0.1509	1101.550	377.260
9	1165	428	1.2045	0.1698	1164.650	398.960
10	1266	435	1.2326	0.1887	1225.170	419.730
11	1319	437	1.2619	0.2075	1283.700	439.760
12	1349	442	1.2927	0.2264	1340.730	459.230
13	1396	458	1.3250	0.2453	1396.640	478.270
14	1429	474	1.3590	0.2642	1451.740	496.980
15	1496	513	1.3947	0.2830	1506.310	515.460
16	1499	565	1.4324	0.3019	1560.570	533.780
17	1508	593	1.4722	0.3208	1614.750	552.010
18	1534	596	1.5143	0.3396	1669.050	570.220
19	1558	601	1.5588	0.3585	1723.670	588.470
20	1620	601	1.6061	0.3774	1778.770	606.810
21	1866	607	1.6563	0.3962	1834.560	625.310
22	1868	644	1.7097	0.4151	1891.230	644.020
23	1908	653	1.7667	0.4340	1948.970	662.990
24	1944	683	1.8276	0.4528	2008.020	682.290
25	2025	684	1.8929	0.4717	2068.570	701.980
26	2085	689	1.9630	0.4906	2130.920	722.130
27	2200	720	2.0385	0.5094	2195.350	742.810
28	2225	740	2.1200	0.5283	2262.170	764.100
29	2420	766	2.2083	0.5472	2331.780	786.110
30	2450	787	2.3043	0.5660	2404.620	808.930
31	2506	837	2.4091	0.5849	2481.210	832.700
32	2531	871	2.5238	0.6038	2562.230	857.550
33	2563	885	2.6500	0.6226	2648.430	883.650
34	2580	905	2.7895	0.6415	2740.840	911.220
35	2677	986	2.9444	0.6604	2840.690	940.500
36	2702	992	3.1176	0.6792	2949.690	971.830
37	3010	1001	3.3125	0.6981	3070.040	1005.600
38	3215	1076	3.5333	0.7170	3204.880	1042.320
39	3229	1092	3.7857	0.7358	3358.690	1082.710
40	3585	1185	4.0769	0.7547	3538.290	1127.720
41	3790	1245	4.4167	0.7736	3754.690	1178.720
42	4780	1250	4.8182	0.7925	4026.950	1237.690
43	4790	1258	5.3	0.8113	4390.960	1307.630
44	4828	1302	5.8889	0.8302	4918.220	1393.170
45	5580	1625	6.6250	0.8491	5720.320	1501.030
46	6860	1706	7.5714	0.8679	6777.500	1638.330
47	7960	1768	8.8333	0.8868	7884.270	1805.390
48	8275	1842	10.6	0.9057	8989.560	1992.410
49	10000	1928	13.2500	0.9245	10156.260	2194.360
50	11558	2716	17.6667	0.9434	11490.650	2420.990
51	14376	2930	26.5	0.9623	13197.000	2702.710
52	15000	2966	53	0.9811	15883.480	3134.800

Tabla. F.12. Eventos de diseño

ESTACIÓN HUITES			
PER. RET.	PROB.	DOBLE GUMBEL	DOBLE GUMBEL
T	P(X ≤ x)	Q	V
años	1 - 1/T	m <sup>3</sup> /s	Hm <sup>3</sup>
2	0.5	2162.86	732.39
5	0.8	4158.62	1264.15
10	0.9	8655.38	1934.79
20	0.95	12029.12	2510.74
50	0.98	15665.02	3099.86
100	0.99	18227.51	3505.42
200	0.995	20726.23	3898.42
500	0.998	23990.16	4409.35
1000	0.999	26436.92	4792.36
2000	0.9995	28893.14	5173.89
5000	0.9998	32067.32	5682.61
10000	0.9999	34485.74	6061.18





# Bibliografía

Abramowitz, M. y Stegun, I. (1965). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. National Bureau of standards applied Mathematics, Series **55**.

Aldama, A., Ramírez, A. I., Aparicio, F., Mejía, R., Ortega, G. (2006). *Hydrologic Safety of Dams in Mexico*, Instituto Mexicano de Tecnología del Agua, México.

Alexander, C. (2008). *Market Risk Analysis* Volume. II, John Wiley & Sons Ltd, United Kingdom.

Aparicio, F. J. (2000). *Fundamentos de hidrología de superficie*, Ed. Limusa, México.

Aranda, D. J. (2014). *Estimation of the probability of oververting and maximum flows downstream of reservoir dams*. Tesis de Doctorado, Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos, Universidad Politécnica de Valencia, España.

Arganis, J.M.L., Domínguez, M.R. y Herrera, A.J., (2013). Getting design events of bivariate functions by means of the bisection method, *Revista Chilena de Ingeniería*, **21** (2), 293-300.

Arganis, J.M.L. (1998). *Tránsito de Avenidas en Embalses y Cauces*. Tesis de Maestría, DEPFI. Universidad Nacional Autónoma de México.

Arreguín, F., Alcocer, V. (2011). *Diseño Hidráulico de Vertedores*, Instituto Mexicano de Tecnología del Agua, México.

Bender, J., Wahl, T. y Jensen, J., (2014). Multivariate design in the presence of non-stationarity, *Journal of Hydrology*, **514**, 123-130.

Berezowsky, V.M., Fuentes, M.O. (1983). *Métodos Numéricos*. Manual de Diseño de Obras Civiles. CFE-México.

Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G. and Roncalli, T. (2000). Copulas for finance A reading guide and some applications. Working Paper, Groupe de Recherche Opérationnelle, Crédit Lyonnais.

Candela, A., Brigandí, G. y Aronica, G., (2014). Estimation of synthetic flood design hydrographs using a distributed rainfall-runoff model coupled with a copula-based single storm rainfall generator, *Natural Hazards and Earth System Science*, **14**, 1819-1833.

Chebana, F., Ouarda, TBMJ. (2011). Multivariate quantiles in hydrological frequency analysis. *Environmetrics*, Vol. **22**.

Chen, L. y Guo, S. (2019). *Copulas and Its Application in Hydrology and Water Resources*, Springer.

Chowdhary, H., Escobar, L.A. and Singh, V.P., (2011). Identification of suitable copulas for bivariate frequency analysis of flood peak and flood volume data, *Hydrology Research*, **42**, 193-216.

Clarke, Kevin A. (2007). A simple distribution-free test for nonnested model selection. *Political Analysis*, **15** (3), 347-363.

Comisión Nacional del Agua (CNA) (1997). *Presa y Central Hidroeléctrica Luis Donald Colosio Murrieta (Huites) Memoria Técnica (1992-1995)*, T I, México.

Comisión Nacional del Agua (CONAGUA) (2014a). Banco Nacional de Datos de Aguas Superficiales (BANDAS). GASIR, Subdirección General Técnica. <https://app.conagua.gob.mx/bandas/>

Comisión Nacional del Agua (CONAGUA) (2014b). Estudio de Alternativas de Solución para Suministrar Agua Potable a las Ciudades

de Minatitlán y Coatzacoalcos en el Estado de Veracruz. Dirección General del Organismo de Cuenca Golfo Centro y la Dirección Local Veracruz, México.

Comisión Nacional del Agua (CONAGUA) (2014c). *Programa de medidas preventivas y de mitigación de la sequía. Consejo de Cuenca Río Coatzacoalcos*, México.

Conway, D. (1979). *Multivariate Distributions with Specified Marginals*. (Technical Report 145, Stanford University).

Croley II, T.E., (1977). *Hydrologic and Hydraulic Computations on Small Programmable Calculators*. Iowa Inst. Hydraul. Res., Univ. Iowa, Iowa City, Iowa, 837 pp.

Croley II, T.E., (1980). Gamma Synthetic Hydrographs, *Journal of Hydrology*, **47**, 41-52 pp.

Cuberos, A., Masiello, E., y Maume-Deschamps, V. (2019). Copulas checker-type approximations: application to quantiles estimation of aggregated variables. *Comm. Statist - Theory and Methods*.

Daniels, H. E. (1950). Rank correlation and population models. *J. Roy. Statist. Soc. Ser. B***12**, 171-181.

De Michele, C., Salvadori, G., Canossi, M., Petaccia, A. and Rosso, R., (2005). Bivariate statistical approach to check adequacy of dam spillway, *Journal of Hydrologic Engineering*, **10**, 50-57.

Demarta, S. y McNeil, A. (2004). The t copula and related copulas. *International Statistical Review*, **73**(1), 111–129.

Devroye, (1986). L. *Non-Uniform Random Variate Generation*. Springer-Verlag, New York.

Domínguez, M.R., Arganis, J.M.L., (2012). Validation of methods to estimate design discharge flow rates for dam spillways with large regulating capacity, *Hydrological Sciences Journal*, **57** (3), 460-478.

Domínguez, R., Arganis, J.M.L., Carrizosa, E.E., Fuentes, M.G.E., Echeverri, C.A. (2006). *Determinación de Avenidas de Diseño y Ajuste de los Parámetros del Modelo de Optimización de las Políticas de Operación del Sistema de Presas del Río Grijalva*. Para CFE. Informe Final.

Domínguez, R., Arganis, J.M.L., Carrizosa, E.E., Osnaya, R.J., Esquivel, G.G. (2019). *Actualización de avenidas de diseño de la presa Huites, Sin., México*. Para CFE. Informe Final.

Domínguez, R., J.M.L., Carrizosa, E.E., Fuentes, M.G.E., Berezowsky, V.M. (2003). *Normatividad interna para determinar de forma racional y congruente los periodos de retorno que deben cumplirse en el diseño de obras de almacenamiento, control y protección*. Para CNA. Informe Final.

Domínguez, R.M., Fuentes, M.G.E., Arganis, J.M.L. (2004). *Optimización de los parámetros de la función de distribución doble Gumbel usando algoritmos genéticos en una serie de gastos máximos anuales*. XXI Congreso Latinoamericano de Hidráulica, Sao Paulo, Brasil.

Durante, F. y Sempi, C. (2010). Copula theory: an introduction, In: Jaworski, P., Durante, F., Härdle W., Rychlik, T. *Proceedings of the Copula Theory and its Applications, Lecture Notes in Statistics*, vol. 198, Springer, Berlin, Heidelberg, pp. 3–31.

Eckhardt K., (2005). How to construct recursive digital filters for baseflow separation. *Hydrol Process*, **19**, 507–515.

Embrechts, P., McNeil, A., Straumann, D. (2001). *Correlation and dependency in risk management: properties and pitfalls*. In: Dempster, e.M., Moffatt, H. (Eds.), *Risk Management: Value at Risk and Beyond*. Cambridge University Press.

Erdely, A. (2009). *Cóputas y dependencia de variables aleatorias: Una introducción*. *Miscelánea Matemática*, **48**, 7-28. Universidad Autónoma Metropolitana-Cuajimalpa, México.

Escalante, S.C. (1998). Multivariate extreme value distributions with mixed Gumbel marginals, *American Water Resources Association*, Vol. **34**, No. 2.

Escalante, S.C., Reyes, C.L. (2005). *Técnicas Estadísticas en Hidrología*. Facultad de Ingeniería. Universidad Nacional Autónoma de México.

Eschenburg, P. (2013). *Properties of extreme-value copulas*, Tesis de Diploma, Technische Universität München Fakultät für Mathematik, Munich, Alemania.

Fang, H. & Fang, K. (2002). The meta-elliptical distributions with given marginals. *J. Multivariate Anal.* **82**, 1–16.

Favre, A.C., Brunner, M. I., Seibert, J., (2016). Bivariate return periods and their importance for flood peak and volume estimation, *Wires Water*, Vol. **03**, 819-833 pp.

Favre, A.C., Meylan, P. y Musy, A., (2012). *Predictive hydrology: A frequency analysis approach*. CRC Pres.

Favre, A.C. Adlouni, S., Perreault, L., Thiémondge, N., and Bobée, B., (2004). Multivariate hydrological frequency analysis using copulas, *Water Resour.*

Finkelstein, B.V., (1953). On the limiting distributions of the extreme terms of a variation a series of a two dimensional random quantity. *Doklad Akad, S.S.S.R.*, **91**(2), 209.

Fisher, N. I., Switzer, P. (2001). Graphical assessment of dependence: Is a picture worth 100 tests?, *Am. Stat.*, **55**(3), 233-239.

Fisher, N., Switzer, P. (1985). Chiplots for asses-sing dependence, *Biometrika*, **72**(2), 253-265.

Fréchet, M., (1951). Sur les tableaux de corrélation dont les marges sont données. *Ann Univ Lyon Sect*, **A9**, 53-77.



Fréchet, M., (1958). Remarques au sujet de la note précédente. *C. R. Acad Sci. Paris. Sér I Math* **246**, 2719-2720.

Fuentes, M.O., Arganis, J.M.L., Domínguez, M.R., Fuentes, M.G., Rodríguez, V.K. (2015). Maximization of the Likelihood Function of Probability Distributions using Genetic Algorithms. *Ingeniería del Agua*. Editorial UPV.

Fuentes, M.O.A., Fuentes, M.G.E., Domínguez, M.R. (2005). Optimización de los parámetros de algunas funciones de distribución de probabilidad de gastos máximos anuales usando un algoritmo genético simple. *4a. Conferencia Iberoamericana en Sistemas Cibernética e Informática, Cicsi, Orlando, Flo., Usa*, Vol. **2**, 156-159.

Galambos, J. (1975). Order Statistics of Samples from Multivariate Distributions. *J. Am. Stat. Assoc.* Vol. **70**(351), 674-680.

Galambos, J. (1978). *The asymptotic theory of extreme order statistics*, Wiley & Sons.

Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika*.

Genest, C. and Boies, J.C., (2003). Detecting dependence with Kendall plots, *The American Statistician*, **57**, 275-284.

Genest, C. and Rémillard, B., (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales d'Institut Henri Poincaré, Probabilités et Statistiques*, **44**, 1096-1127.

Genest, C., Boies, J. (2003). Detecting dependence with Kendall plots, *The American Statistician*, **57**(4), 275-284.

Genest, C., Favre, A. (2007). Everything you always wanted to know about copula modeling but were afraid to ask, *Journal of Hydrologic Engineering*, **12**(4), 347-368.

Genest, C., Rémillard, B., Beaudoin, D. (2009). Goodness-of-fit tests for copulas: a review and a power study. *Insurance: Mathematics and economics*, **44** (2), 199-213.

Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, USA.

González, F. J. (1970). *Contribución al análisis de frecuencias de valores extremos de los gastos máximos en un río*, Pub. **277**, Instituto de Ingeniería, UNAM, México.

González-Barrios, J. y Hernández, M. (2013). Sample d-copula of order m. *Kybernetika*, **49**(5), 663–693.

González-Barrios, J. y Hoyos, R. (2019). Estimating checkerboard approximations with sample d-copulas. *Comm. Stat. – Simulation and computation*.

Guarín, J., Jaramillo, M., Lopera, C. (2018). A comparison of two graphical methods for detecting dependence, *Ciencia en Desarrollo*, **9**(1), 71-88.

Gudendorf, G. y Segers, J. (2010). Copula theory: an introduction, In: Jaworski, P., Durante, F., Härdle W., Rychlik, T. *Proceedings of the Copula Theory and its Applications, Lecture Notes in Statistics*, vol. **198**, Springer, Berlin, Heidelberg, pp. 127–145.

Gumbel E.J., (1958). *Statistics of extremes*, Columbia University Press, USA.

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

Gumbel, E.J. (1962). *Statistical theory of extreme values*, Chap. 6, Contributions to order statistics, Sarhan & Greenberg Ed., Wiley & Sons.

Gumbel, E.J., Mustafi, C.K. (1967). Some analytical properties of bivariate extremal distributions, *American Statistical Association*, Vol. **62**.

Hiemstra, A. V., Francis, D.M. (1981). Run hydrographs for prediction of flood hydrographs, *Journal of the Hydraulics Division*, ASCE Vol. **107**, No. HY6, 759-775 pp.

Holland, J.H. (1975). *Adaptation in Natural and Artificial Systems*. The University of Michigan Press.

Hosking, J.R.M, and Wallis, J.R., (1997). *Regional frequency analysis: an approach based on L-moments*, Cambridge University Press.

Hüsler, J. y Reiss, R. (1989). Maxima of normal random vectors: Between independence and complete dependence. *Statist. Probab. Lett.*, **7**(4):283-286.

Instituto Geológico y Minero de España (IGME) (2008). Guía metodología para la determinación de los caudales ambientales. Capítulo de aguas subterráneas, Tomo I. Métodos de estimación de las aportaciones subterráneas a los caudales ambientales, España.

Jared A. Linebach et al., (2014). *Nonparametric Statistics for Applied Research*, Springer, New York USA.

Jiménez, E.M. (1997). Manual de operación de los programas AX.exe y CARACHID.exe (para hidrología), CENAPRED, México.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman-Hall.

Joe, H. (2015). *Dependence Modeling with Copulas*. CRC Press Taylor & Francis Group.

Johnson, M.E. (1987). *Multivariate Statistical Simulation*. J. Wiley Sons, New York.

Kite, G.W. (1988). *Frequency and risk analyses in hydrology*, Water Resources Publication, Littleton, Colorado.

Klein, B., Pahlow, M., Hundecha, Y. and Schumann, A., (2010). Probability analysis of hydrological loads for the design of flood control systems using copulas, *Journal of Hydrologic Engineering*, **15**, 360-369.

Kotz, S., Nadarajah, S., (2004). *Multivariate t-Distributions and their Applications*. Cambridge University Press.

Kuester, J. L. y J. H. Mize. (1973). *Optimization techniques with FORTRAN*. McGraw-Hill. USA.

Ladson, A. R., Brown, R., Neal, B., Nathan, R. (2013). A standard approach to baseflow separation using the Lyne and Hollick Filter, *Australian Journal of Water Resources*, **17**, 25-34.

Laio, F., (2004). Cramer-von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters, *Water Resources Research*, **40**.

Mächler, M (2014). Spearman's Rho for the AMH Copula: a Beautiful Formula. URL <https://cran.r-project.org/web/packages/copula/vignettes/rhoAMH-dilog.pdf>

Mendoza, G. R., Zavala, A. J., Villa, I. S. (2014). Revisión de gastos de diseño de la presa Huites mediante relaciones lluvia-escorrentamiento. *Ingeniería Hidráulica y Ambiental*, **35**, 77-89.

Meyer, C. (2009). The Bivariate Normal Copula. *Commun. Stat. – Theory Methods*. **42**, 2402-2422.

Mockus, V. (1957). *Use of Storm and Watershed Characteristics in Synthetic Unit Hydrograph Analysis and Application*, U.S. Soil Conservation Service.

Mood A., Graybill F. y Boes D. (1974). *Introduction to the Theory of Statistics*. Tercera edición., McGraw-Hill.

Myles Hollander et al., (2014). *Nonparametric Statistical Methods*, John Wiley, New Jersey USA.

- 
- Nash, J.E. (1957). The Form of the Instantaneous Unit Hydrograph. *Publications International Association of Hydrological Sciences*, **45**, 114-121 pp.
- Natural Environment Research Council., (1975). *Estimation of flood volumes over different durations*. Flood Studies Report, Vol. **1**, United Kingdom.
- Nelsen, R. B. (1986). Properties of a one-parameter family of bivariate distributions with specified marginals. *Commun. Stat. - Theory Methods*.
- Nelsen, R.B. (2006). *An Introduction to Copulas*, 2nd edition. Springer, New York.
- Pickands, J. (1981). Multivariate extreme value distributions (with a discussion). In Proceedings of the 43rd Session of the International Statistical Institute. *Bull. Inst. Internat. Statist.* **49**, 859–878.
- Ponce, V. M. (1989). *Engineering hydrology, Principles and Practices*. Prentice Hall, Englewood Cliffs, New Jersey.
- Ramírez, A. I., Aldama, A. (2000). *Análisis de frecuencias conjunto para la estimación de avenidas de diseño*. Asociación Mexicana de Hidráulica e Instituto Mexicano de Tecnología del Agua, Avances en hidráulica **7**.
- Ramírez, S. L. (2011). *Actualización de la Avenidas de Diseño de las Presas del Río Grijalva*. Tesis de Licenciatura, Facultad de Ingeniería, Universidad Nacional Autónoma de México.
- Rao, A.R. and Hamed, K.H., (1999). *Flood frequency analysis*. CRC Press.
- Raynal J., Guevara J. (1997). Maximum Likelihood Estimators for the two Populations Gumbel distribution. *Hydrological Science and Technology*, **13**(1-4), 47-56.
- Raynal, V.J.A. (1985). Bivariate extreme value distributions applied to flood frequency analysis Ph. D. Dissertation. Colorado State University, Ft. Collins, Colorado, U.S.A.

Reich, B. M. (1962). Design hydrographs for very small watersheds from rainfall, Report No. CER 62 MNR 4150, Civil Engineering Section, Colorado State University.

Requena, R.A.I. (2015). *Flood frequency analysis by a bivariate model based on copulas*. Tesis Doctoral. Universidad Politécnica de Madrid, Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos, España.

Rodríguez, R. (2003). *Cóputas Arquimedeanas*, Tesis de Maestría en Ciencias, Posgrado en Ciencias Matemáticas UNAM, México.

Rosenbrock, H.H. (1960). An automatic method for finding the greatest or least value of a function, *Computer Journal*, Vol. **3**.

Salvadori, G., De Michele, C. (2004). Frequency analysis via copulas: Theoretical aspects and applications to hydrological events, *Water Resources Research*, Vol. **40**.

Salvadori, G., De Michele, C., Kottegoda, N. T. y Rosso, R. (2007). *Extremes in nature. An approach using copulas*, Springer.

Secretaría de Agricultura y Recursos Hidráulicos (SARH) (1987). Instructivo de Hidrología para determinar la Avenida Máxima Ordinaria, México.

Shiau, J.T. (2003). Return period of bivariate distributed extreme hydrological events, *Stoch. Environ. Res. Risk Assess.*, **17**.

Singh, Vijay P. (1988). *Hydrologic Systems: Rainfall-runoff modeling, Vol. 1 Hydrologic Systems*, Prentice Hall, Englewood Cliffs, New Jersey, 480 pp.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Inst. Statist. Univ. Paris, Publ.* **8**, 229-231.

Soranzo, A., Epure, E. (2014). Very Simply Explicitly Invertible Approximations of Normal Cumulative and Normal Quantile Function. *Applied Mathematical Sciences*, Vol. **8**.

Stedinger, J.R., Vogel, R.M., Foufoula, G.E. (1993). *Frequency analysis of extreme events*. D.R. Maidment (Ed.), Handbook of Hydrology, McGraw-Hill, New York.

Stefano Bonnini et al., (2014). *Nonparametric Hypothesis Testing Rank and Permutation Methods with Applications in R*, John Wiley, United Kingdom.

Steven Davis and Evelyn R. Davis (2016). *Data Analysis with SPSS Software Linear Regression and Non-Parametric Tests*, Cognella Academic Publishing.

Tawn, J. (1988). Bivariate extreme value theory: models and estimation. *Biometrika*, **75**(3):397-415.

The Mathworks, Inc. (1992). The Mathworks Matlab Reference Guide.

Tiago de Oliveira, J., (1958). Extremal distributions. *Revista Faculdade de Ciencias*. Universidad de Lisboa. Portugal. Serie **A7**, 219-228.

Vuong, Quang H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Journal of the Econometric Society*, 307-333.

Yue, S., Ouarda, T., Bobée, B., Legendre, P., y Bruneau, P. (1999). The Gumbel mixed model for flood frequency analysis, *J. Hydrol.*, **226**.

Zhang, L., Singh, V. P. (2019). *Copulas and their Applications in Water Resources Engineering*. Cambridge University Press, London.