



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

ANÁLISIS REGIONAL PARA LA ESTIMACIÓN DE GASTOS DE DISEÑO. APLICACIÓN
A LAS REGIONES HIDROLÓGICAS 18, 22 Y 23.

TESIS
QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN INGENIERÍA

PRESENTA:
ERICK QUENDER PÉREZ ESPINOSA

TUTOR
DR. RAMÓN DOMÍNGUEZ MORA
INSTITUTO DE INGENIERÍA

CIUDAD DE MÉXICO, JULIO 2017

JURADO ASIGNADO:

Presidente: Dr. Jesús Gracia Sánchez

Secretario: Dr. Oscar Arturo Fuentes Mariles

Vocal: Dr. Ramón Domínguez Mora

1^{er}. Suplente: M.I. Víctor Franco

2^{d o}. Suplente: Dra. Maritza Liliana Arganis Juárez

Facultad de Ingeniería, Ciudad Universitaria. Ciudad de México.

TUTOR DE TESIS:

Dr. Ramón Domínguez Mora

FIRMA

Agradecimientos

Agradezco al CONACYT por el apoyo económico brindado para la realización de los estudios de posgrado.

Al Dr. Ramón Domínguez y a mis sinodales Dr. Jesus Gracia, Dr. Oscar Fuentes, Dra. Maritza Arganis y al M.I. Víctor Franco por su apoyo y guía en la realización de este trabajo.

Resumen

En la República Mexicana existen regiones con estaciones hidrométricas que proporcionan información valiosa para la estimación de avenidas de diseño. Sin embargo, en tiempos recientes, se ha observado una baja en el número de sitios que cuentan con mediciones, lo que ha acrecentado la dificultad de disponer de registros que nos permitan realizar estimaciones de gastos para distintos periodos de diseño.

En este trabajo se presenta una metodología de obtención de avenidas de diseño, a partir de sitios con o sin información, mediante factores regionales y ecuaciones del tipo potencial. La aplicación de los métodos se llevó a cabo a en las regiones hidrológicas 18, 22 y 23 de la República Mexicana.

Mediante el uso de técnicas regionales, se obtuvieron factores que permiten calcular gastos de diseño en sitios con poca información a partir de la media de gasto máximo instantáneo anual registrado. Además, se obtuvieron ecuaciones regionales de tipo potencial que permiten la estimación de gastos en sitios donde únicamente se cuente con información de las características propias de la cuenca y de precipitación.

Con la realización de este estudio, se aportan nuevas herramientas para la estimación de avenidas de diseño, además se hace ver la importancia de la calidad de la información que se recaba. Para dar una estimación con la menor incertidumbre posible se deben verificar los registros, validarlos y someterlos a análisis mediante pruebas estadísticas. Además de verificar la congruencia de las ecuaciones potenciales en cuanto al sentido y naturaleza de las variables que utilicen.

Índice

Agradecimientos	ii
Resumen.....	iii
Índice de tablas	v
1. Introducción	1
2. Antecedentes	3
3. Conceptos teóricos sobre regionalización	6
3.1 Pruebas de homogeneidad e independencia de eventos	6
3.1.1 Prueba estadística de Helmert	6
3.1.2 Prueba t de Student	6
3.1.3 Prueba estadística de Cramer	7
3.1.4 Prueba de independencia de eventos.....	8
3.2 Funciones de distribución de probabilidad	9
3.2.1 Distribución Gumbel.....	9
3.2.2 Doble Gumbel.....	11
3.3 Regionalización.....	11
3.3.1 Delimitación de regiones homogéneas.....	11
3.3.2 Técnicas regionales para la estimación de gastos de diseño	15
4. Metodología	18
4.1 Selección de estaciones y extracción de datos	18
4.2 Cálculo de los estadísticos.....	18
4.3 Pruebas de homogeneidad e independencia.....	18
4.4 Obtención de rasgos fisiográficos de la cuenca	19
4.5 Agrupación de estaciones hidrométricas.....	20
4.6 Análisis estadístico regional	21
5. Aplicación y resultados.....	23
5.1 Descripción de las zonas de estudio.....	23
5.1.1 Región hidrológica 18.....	23
5.1.2 Región hidrológica 22.....	23
5.1.3 Región hidrológica 23.....	23
5.2 Recopilación y análisis de la información.....	25
5.2.1 Región hidrológica 18.....	25
5.2.2 Región hidrológica 22	28
5.2.3 Región hidrológica 23.....	29
5.3 Pruebas de homogeneidad e independencia.....	31
5.3.1 Región hidrológica 18.....	32
5.3.2 Región Hidrológica 22	33
5.3.3 Región Hidrológica 23	34
5.4 Identificación de grupos homogéneos.....	35
5.4.1 Región hidrológica 18.....	35
5.4.2 Región hidrológica 22.....	40
5.4.3 Región hidrológica 23.....	43
5.5 Análisis estadístico regional y factores regionales.....	45
5.6 Ecuaciones regionales	50
5.7 Generación de muestras sintéticas	56
6. Conclusiones y recomendaciones	61
Bibliografía	62

Índice de tablas

TABLA 1. ESTACIONES SELECCIONADAS DE LA REGIÓN HIDROLÓGICA 18.	25
TABLA 2. RESUMEN DE ESTADÍSTICOS DE LA RH 18.....	27
TABLA 3. LISTA DE ESTACIONES DE LA RH-22.....	28
TABLA 4. ESTADÍSTICOS PARA LA RH-22.....	29
TABLA 5. RELACIÓN DE ESTACIONES RH-23.	29
TABLA 6. ESTADÍSTICOS PARA LA RH-23.....	31
TABLA 7. RESUMEN DE RESULTADOS DE LAS PRUEBAS DE HOMOGENEIDAD E INDEPENDENCIA, RH-18.	32
TABLA 8. RESUMEN DE RESULTADOS DE LAS PRUEBAS DE HOMOGENEIDAD E INDEPENDENCIA, RH-18.	33
TABLA 9. RESUMEN DE RESULTADOS DE LAS PRUEBAS DE HOMOGENEIDAD E INDEPENDENCIA, RH-23.	34
TABLA 10. ESTACIONES SOMBREADAS LOCALIZADAS DEBAJO DE LA CURVA DE AJUSTE.....	36
TABLA 11. RESULTADOS OBTENIDOS POR EL MÉTODO DEL COEFICIENTE DE VARIACIÓN.....	36
TABLA 12. AGRUPACIÓN POR MÉTODO DE FISHER.	38
TABLA 13. ESTACIONES DE CADA UNA DE LAS REGIONES HOMOGÉNEAS DE LA RH-18.....	39
TABLA 14. COEFICIENTES DE VARIACIÓN Y LOGARITMOS DEL ÁREA Y LOS QMIA, RH-22.	40
TABLA 15. COEFICIENTES DE VARIACIÓN Y LOGARITMOS DEL ÁREA Y LOS QMIA, RH-23.	44
TABLA 16. FACTORES REGIONALES OBTENIDOS POR EL MÉTODO DE ESTACIONES - AÑO.....	50
TABLA 17. ECUACIONES POTENCIALES DE LA RH-18.....	52
TABLA 18. ECUACIONES POTENCIALES DE LA RH-22.....	53
TABLA 19. ECUACIONES POTENCIALES DE LA RH-23.....	54
TABLA 20. ECUACIONES REGIONALES PARA ESTIMAR LA MQMIA.....	55
TABLA 21. PARÁMETROS DE LA DISTRIBUCIÓN DOBLE GUMBEL.....	57

1. Introducción

En hidrología los métodos regionales para la estimación de avenidas son cada vez más usados debido a la falta de registros, mala calidad de los datos e inclusive nula medición. Los pocos registros históricos impiden un análisis confiable, ya que es común encontrar más estaciones con registros de precipitaciones que estaciones con datos de escurrimiento. Ante esta situación se requiere el uso de métodos que nos ayuden a estimar con la menor incertidumbre posible escurrimientos en sitios con escasa o nula información.

En ocasiones se plantea el desarrollo de infraestructura hidráulica en sitios donde no se cuenta con información, por lo tanto, es imprescindible contar con métodos que ayuden a enfrentar dicha problemática. El análisis aquí presentado contempla aplicar modelos regionales y verificar sus resultados. Recientemente se han realizado trabajos de este tipo, por parte del Instituto de Ingeniería de la UNAM, en varias regiones hidrológicas del país, sin embargo, aún restan varias por analizar. Las regiones hidrológicas 18, 22 y 23 forman parte de las zonas en el país en las cuales no se han realizado, hasta el momento, estudios de este tipo.

El presente trabajo gira entorno a poder estimar avenidas de diseño empleando la media de los gastos máximos instantáneos anuales, de acuerdo con esto último se plantea una metodología con dos enfoques.

La primera consiste en estimar avenidas de diseño a partir de factores regionales obtenidos por el método de estaciones-año en sitios con información disponible.

En tanto, la segunda va enfocada a obtener gastos de diseño en zonas con nula información. Para ello se emplearán ecuaciones regionales en función de las características fisiográficas e hidrológicas de la cuenca, que tendrán como fin estimar la media del gasto máximo instantáneo anual correspondiente a dicha zona y, con este último, obtener los eventos de diseño deseados con ayuda de los factores regionales.

En distintas regiones del mundo se han llevado a cabo estudios regionales con diversos objetivos, en el capítulo de Antecedentes se pone en contexto el planteamiento del problema con lo antes mencionado.

En el capítulo *Conceptos teóricos* se explica en que consiste cada uno de los métodos y pruebas estadísticas que se aplicaran durante el análisis.

La estructura general de la metodología, abarcando los dos enfoques antes descritas se expone en el capítulo *Metodología*. En este apartado se hace una explicación paso a paso sobre cómo obtener eventos de diseño en cuencas aforadas y no aforadas.

En *Aplicación y resultados* se describen las regiones hidrológicas en estudio, además, se aplican los conceptos teóricos siguiendo la metodología expuesta en el capítulo anterior. Para cada resultado, se hace un breve análisis y comentarios sobre lo obtenido. Hasta este punto ya se tendrán los factores y ecuaciones regionales que nos permitirán estimar eventos de diseño.

Finalmente, en *Conclusiones y recomendaciones* como su nombre lo indica se exponen las principales conclusiones de la metodología empleada así como comentarios sobre el uso y aplicabilidad de los resultados obtenidos.

2. Antecedentes

El problema de determinar gastos de diseño en sitios con poca información no es exclusivo de nuestro país, en varias partes del mundo se han aplicado distintos métodos de regionalización con resultados satisfactorios y que, además, muestran ser una buena herramienta para afrontar dichos problemas. A continuación se citan algunos trabajos realizados recientemente en algunos países.

Allasia & Avruch Goldenfun (2000) llevaron a cabo un proceso de regionalización en la región del alto Uruguay, en ese trabajo se emplearon ecuaciones de tipo potencial que relacionaban el área de la cuenca, precipitación y número de orden de corrientes. El modelo que emplearon fue como el presentado a continuación:

$$Q = aA^b P^c D^d \quad (2.1)$$

Siendo A , P , D y Q el área, precipitación, número de orden de corriente y gasto respectivamente; a , b , c y d son los parámetros por determinar. La ecuación anterior se fue linealizando y añadiendo parámetros para cada proceso de regresión lineal múltiple, para encontrar un mejor ajuste. Un ajuste de este tipo tiene la limitante que por ser para una única zona homogénea la aplicabilidad de la función encontrada se limita a sitios con área menor o igual a la zona de estudio original.

Por otro lado, en la República de El Salvador, se aplicaron técnicas de delimitación de regiones homogéneas y de estimación regional de avenidas de diseño para periodos de retorno de 5, 10, 15, 20, 25, 50 y 100 años (Erazo Ch., 2004). Primero realizaron un análisis de los datos a la que denominaron "Outliners", eliminando datos dudosos que pudieran meter algún tipo de ruido al análisis estadístico que se realizaría posteriormente.

Al tratarse de una regionalización que comprendía todo el país, delimitaron regiones homogéneas empleando el *Test de homogeneidad de Gumbel* el cual consiste en evaluar si la curva de ajuste para un análisis puntual no se aleja de la curva de frecuencia regional compuesta por n estaciones, tomando como fronteras los límites de confianza en función de la variable reducida Z .

$$Z = -\ln\left(-\ln\left(1 - \left(\frac{1}{Tr}\right)\right)\right) \quad (2.2)$$

Una vez definidas las regiones homogéneas aplicaron la técnica regional denominada "Avenida índice" con la hipótesis de "estaciones-año", es decir, una mezcla de dos métodos de regionalización. Después de obtener el registro completo para cada región homogénea, se ajustaron a funciones de distribución de probabilidades del tipo Gumbel, Log Pearson III, Log Normal II y Log Normal III, para

determinar la función del mejor ajuste se aplicaron las pruebas de Chi cuadrada, Kolmogorov Smirnov y ajuste gráfico (Erazo Ch., 2004).

Erazo Ch., (2004), en la aplicación a la república de El Salvador, propone ecuaciones regionales que involucra al área de las cuencas, la ecuación sigue la forma siguiente:

$$Q = aA + b \quad (2.3)$$

Donde Q es el gasto estimado, A el área de la cuenca, a y b parámetros obtenidos mediante regresión lineal. El estudio abarca estimación de gastos medios y gastos para distintos periodos de retorno, los límites de aplicabilidad son las áreas de las cuencas tomadas en consideración en el proceso de delimitación de regiones homogéneas.

En Turquía, hace varios años se llevaron a cabo grandes proyectos de infraestructura hidroeléctrica. En tiempos recientes se han visto en la necesidad de crear infraestructura menor para la generación de energía eléctrica que abastezca a empresas privadas de reciente creación. Sin embargo, no se cuentan con datos para hacer los estudios pertinentes, es por eso que Isik & Singh, (2008) propusieron modelos regionales para estimar gastos de diseño. Delimitaron regiones homogéneas combinando tres métodos denominados:

- Algoritmo de agrupamiento jerárquico de aglomeración
- Método de partición K- Medias
- Método de curvas de caudales

Una vez que se delimitaron regiones homogéneas, hicieron ajustes para estimar gastos en sitios sin información, la ecuación propuesta tiene la forma:

$$Q = kW_1^{a1}W_2^{a2} \dots W_n^{an} \quad (2.4)$$

Donde Q es el gasto estimado, W son las características fisiográficas o climáticas de la cuenca y an son coeficientes empíricos. Cuando no se tienen datos y únicamente se cuenta con el área, la ecuación se reduce en la siguiente expresión.

$$Q = kA^a \quad (2.5)$$

La aplicación de estos métodos se hizo a 26 cuencas hidrológicas con 1410 sitios de aforo. Finalmente, en dicho estudio recomiendan realizar una delimitación de regiones homogéneas para abarcar un área de análisis menor y de esta manera tener una estimación de gastos con la menor incertidumbre posible.

Por otro lado, en Irán también se han realizado análisis regionales de gastos, principalmente, el realizado por Byzedi, Saghafian, Mohammadi, & Siosemarde, (2013) el suroeste de dicho país, con el fin de crear acciones o planes de trabajo que ayuden a afrontar las consecuencias de la sequía. Ellos toman un enfoque, más que de diseño, de aprovechamiento y manejo de los recursos hídricos. Como bien sabemos Irán es un país de oriente medio, el cual se encuentra en condiciones áridas y con recursos hídricos muy limitados.

Ante esa situación, han emprendido una política de trasvase de cuencas, mediante el cual se pretende abastecer zonas donde el agua es escasa o prácticamente no existe por las condiciones climáticas. Debido a las condiciones climatológicas extremas de la región se hace casi imposible contar con un monitoreo que lleve a cabo un registro de los escurrimientos. Es por eso que algunos investigadores de dicho país han puesto en práctica métodos de regionalización para estimar gastos mínimos y poder definir políticas de trasvase de cuencas en zonas donde fuese necesario.

Byzedi, et al, (2013) aplicaron técnicas para delimitar regiones homogéneas para luego usar modelos multivariados para ello usaron variables propias de cada cuenca como el área, la precipitación total, densidad de drenaje y factores climáticos. Finalmente, concluyeron que las condiciones de déficit por sequías en los escurrimientos era menor en cuencas con altitudes elevadas, caso contrario en zonas bajas donde las condiciones áridas son más severas.

En México se han llevado a cabo varias investigaciones en cuanto al tema de regionalización. Existen modelos complejos como los multivariados para regionalizar gastos, sin embargo, no resultan del todo práctico y en ocasiones debido a la complejidad de los modelos hacen que se lleguen a resultados pocos congruentes sin mencionar la poca practicidad. Es por lo anterior que Domínguez, Arganis, Gonzáles, Carriosa, & Guzmán, (2013) propusieron un nuevo procedimiento para la regionalización de escurrimientos máximos registrados en estaciones hidrométricas de las 37 regiones hidrológicas que comprenden la República Mexicana. La importancia de esta propuesta radica en la simplicidad del método que permite ser aplicado en cuencas con escasa información hidrométrica pero que cuenten con datos de las características fisiográficas y cinemáticas medibles, tales como el área de la cuenca, tiempo de concentración, precipitación media, entre otros.

Es a raíz de investigaciones hechas por Domínguez, et al, (2013), que en fechas recientes se han realizado regionalizaciones en varias zonas del país, parte de este trabajo comprende, de igual manera, aplicar dicha metodología, pero ahora considerando diferentes técnicas de agrupación, con el fin de comprobar que los resultados de dicho método muestre resultados confiables y congruentes.

3. Conceptos teóricos sobre regionalización

3.1 Pruebas de homogeneidad e independencia de eventos

A menudo se requiere que los datos hidrológicos tengan estadísticos homogéneos para tener una mayor garantía y disminuir la incertidumbre en los métodos a aplicarse posteriormente. Generalmente, la falta de homogeneidad de los datos hidrológicos es ocasionada por actividades humanas tales como la deforestación, cambios de uso de suelo, construcción y operación de obras hidráulicas, mediciones erróneas, etc.

3.1.1 Prueba estadística de Helmer

La prueba consiste en analizar el cambio entre cada valor de la serie de datos respecto a su media, si el cambio entre un valor de la serie respecto a la media es menor se considera un signo negativo si fuese mayor se toma signo positivo.

De esta manera se otorga un signo a la desviación que tiene un dato respecto a media, si una desviación de cierto signo es seguida por otro diferente se considera un cambio "C" si no ocurre el cambio se estaría hablando de una secuencia "S".

De esta manera una muestra se considera homogénea si cumple la siguiente condición:

$$-\sqrt{n-1} \leq (S - C) \leq \sqrt{n-1} \quad (3.1)$$

Donde n es el tamaño de la muestra, C el número de cambios y S el número de secuencias.

3.1.2 Prueba t de Student

Otra de las hipótesis que llevan a que una muestra no sea homogénea es un cambio abrupto en la media, la prueba t de Student es de gran ayuda para el análisis de dichos cambios.

Si se considera que una serie Q_i para $i=1,2,\dots, n$, se divide en dos conjuntos n_1 y n_2 ambos de tamaño $n_1 = n_2 = n/2$, entonces el estadístico t para la prueba de homogeneidad se calcula con la ecuación 3.2:

$$t_d = \frac{\bar{x}_1 - \bar{x}_2}{\left[\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \right]^{1/2}} \quad (3.2)$$

Donde:

\bar{x}_1 y S_1^2 son la media y la varianza del primer conjunto.

\bar{x}_2 y S_2^2 son la media y la varianza del segundo conjunto.

El valor absoluto de t_d se compara con el valor de la distribución t de Student de dos colas, y con $\nu = n_1 + n_2 - 2$ grados de libertad para un nivel de significancia de $\alpha = 0.005$. Si y solo si el valor absoluto de t_d es mayor que el de la distribución t de Student, se concluye que la diferencia entre las medias muestra inconsistencias y por lo tanto la serie Q_i se considera no homogénea.

3.1.3 Prueba estadística de Cramer

La prueba estadística de Cramer tiene el propósito de verificar la homogeneidad de una serie de datos Q_i , mediante la determinación de la variación del valor medio en distintos periodos de tiempo dentro de la misma serie. Este método de verificación plantea la creación de tres bloques, el primer bloque del tamaño de la muestra original, el segundo con el 60 % del total de la muestra (Q_{60}), el tercero con 30 % (Q_{30}).

Una serie se considera homogénea cuando no existe variación entre las medias de cada bloque, para realizar la comparación se utiliza la prueba estadística t de Student basadas en las siguientes ecuaciones:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (3.3)$$

$$S_x = \left[\frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{1/2} \quad (3.4)$$

$$\bar{x}_{60} = \sum_{i=1}^{n_{60}} \frac{x_i}{n_{60}} \quad (3.5)$$

$$\bar{x}_{30} = \sum_{i=1}^{n_{30}} \frac{x_i}{n_{30}} \quad (3.6)$$

$$\tau_{60} = \frac{\bar{x}_{60} - \bar{x}}{S_x} \quad (3.7)$$

$$\tau_{30} = \frac{\bar{x}_{30} - \bar{x}}{S_x} \quad (3.8)$$

$$t_w = \left\{ \frac{n_w (n - 2)}{n - n_w [1 + (\tau_w)^2]} \right\}^{1/2} |\tau_w|, \text{ para } w = 60 \text{ y } 30 \quad (3.9)$$

El estadístico t_w tiene una distribución t de Student de dos colas con $\nu = n_1 + n_2 - 2$ grados de libertad y para un nivel de significancia de 0.05.

Si y solo si el valor absoluto de t_w , para $w=60$ y $w=30$, es mayor que el de la distribución t de Student, se puede concluir que la diferencia entre las medias es significativa, por lo tanto, se considera no homogénea.

3.1.4 Prueba de independencia de eventos

Se dice que los datos individuales de una muestra son independientes entre sí cuando no guardan relación alguna entre ellos, es decir, ninguno depende de otro. En resumen, la prueba de independencia de evento comprueba que cada dato de la muestra sea producto de un proceso estocástico.

Para comprobar que los datos pertenecen a una serie de variables aleatorias se aplica la prueba de independencia de Anderson. Esta prueba hace uso del coeficiente de correlación serial r_k para distintos tiempos de retraso " k ".

El coeficiente de correlación serial se calcula con la ecuación 3.10.

$$r_k = \frac{\sum_{i=1}^n (X_i - \bar{x})(X_{i+k} - \bar{x})}{\sum_{i=1}^n (X_i - \bar{x})^2} \quad (3.10)$$

$$\bar{x} = \sum_{i=1}^n \frac{X_i}{n} \quad (3.11)$$

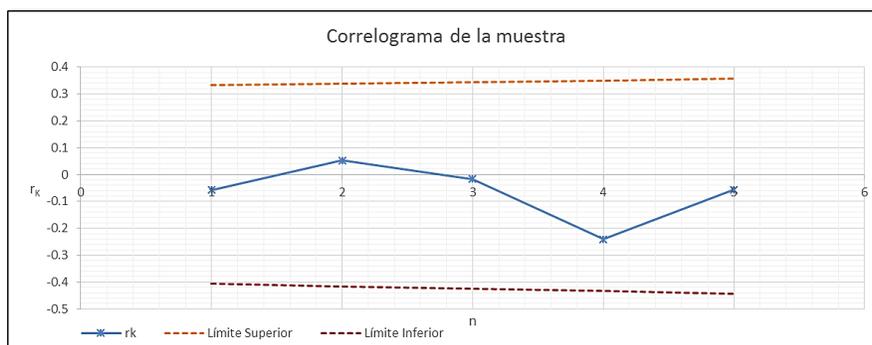
para $k = 1, 2, \dots, n/3$

Se calculan los límites de confianza del 95% de r_k .

$$r_k(95\%) = \frac{-1 \pm 1.9 b \sqrt{(n-k-1)}}{n-k} \quad (3.12)$$

Se grafican los valores estimados para r_k contra los tiempos de retraso k junto con sus correspondientes límites de confianza, a esta gráfica se le denomina Correlograma de la muestra.

Si solo el 10% de los valores r_k sobrepasan los límites de confianza se dice que la muestra es independiente, por lo tanto, los datos son producto de un proceso estocástico, es decir, sigue las leyes de la probabilidad.



Gráfica 1. Ejemplo de un correlograma.

3.2 Funciones de distribución de probabilidad

3.2.1 Distribución Gumbel

La función de distribución de probabilidad se expresa con la siguiente ecuación:

$$F(x) = e^{-e^{\frac{x-v}{\alpha}}} \quad (3.13)$$

La función de densidad de probabilidad es:

$$f(x) = \frac{1}{n} e^{\left(-\frac{x-v}{\alpha}\right)} e^{-e^{\left(-\frac{x-v}{\alpha}\right)}} \quad (3.14)$$

donde: v = parámetro de ubicación y α = parámetro de escala.

La curtosis y el coeficiente de asimetría, típicos para una distribución de este tipo, son de 5.4 y 1.14 respectivamente. La variable Z estandarizada se calcula con:

$$Z_i = \frac{x_i - v}{\alpha} \quad (3.15)$$

Los parámetros de la distribución Gumbel pueden ser calculados por el método de momentos o máxima verosimilitud.

Momentos:

$$\hat{v} = \hat{x} - 0.45S \quad (3.16)$$

$$\hat{\alpha} = 0.78S \quad (3.17)$$

Máxima verosimilitud:

$$P = n - \sum_{i=1}^n e^{-z_i} \quad (3.18)$$

$$R = n - \sum_{i=1}^n z_i + \sum_{i=1}^n z_i e^{-z_i} \quad (3.19)$$

Las ecuaciones, para una estimación de parámetros por máxima verosimilitud, deben converger bajo los siguientes criterios.

$$\frac{P}{\hat{\alpha}} \approx 0 \quad (3.20)$$

$$\frac{-R}{\hat{\alpha}} \approx 0 \quad (3.21)$$

Si no cumplen con los criterios de convergencia se emplean incrementos calculados con:

$$\delta_{vi} = (1.11P_j - 0.26R_j) \frac{\hat{\alpha}_i}{n} \quad (3.22)$$

$$\hat{v}_{j+1} = \hat{v}_j + \delta_{vj} \quad (3.23)$$

$$\delta_{\alpha i} = (0.26P_j - 0.61R_j) \frac{\hat{\alpha}_i}{n} \quad (3.24)$$

$$\hat{\alpha}_{j+1} = \hat{\alpha}_j + \delta_{\alpha j} \quad (3.25)$$

3.2.2 Doble Gumbel

Cuando existe un comportamiento en el cual se observa que los datos, al ser ordenados de mayor a menor, tiendan a formar dos poblaciones, la función Doble Gumbel generalmente es la que mejor se ajusta.

La función de distribución de probabilidades está dada por la ec. 3.26, mientras que la de densidad por la ec. 3.27.

$$F(x) = e^{-e^{-\alpha_1(x-\beta_1)}} p + (1-p)e^{-e^{-\alpha_2(x-\beta_2)}} \quad (3.26)$$

$$f(x) = p\alpha_1 e^{-\alpha_1(x-\beta_1)} + (1-p)\alpha_2 e^{-\alpha_2(x-\beta_2)} \quad (3.27)$$

Donde: α_1 y β_1 = parámetros de la primera población, α_2 y β_2 = parámetro de la segunda población, y p = es la relación que existe entre el número de datos de la segundo y primera población. La variable estandarizada Z se calcula de igual manera que la distribución Gumbel.

3.3 Regionalización

El concepto de regionalización en hidrología surge de la necesidad de contar con muestras más completas que ayuden a la estimación de eventos de diseño. Este proceso consiste en estimar lluvias o escurrimientos en sitios donde no se cuenta con información a partir de uno o varios sitios donde si cuentan con ella.

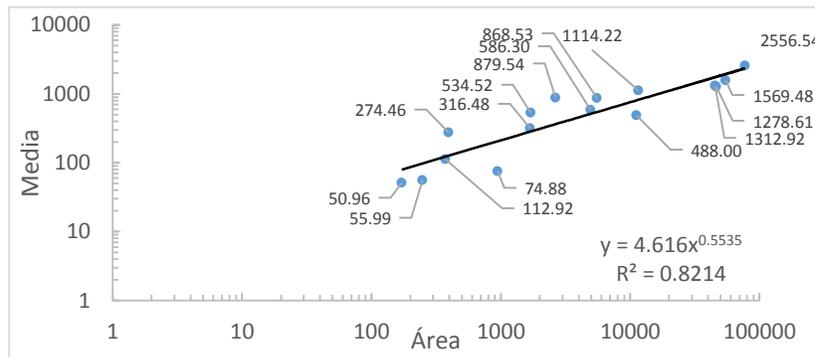
El principal problema que se tiene al aplicar el concepto de regionalización es la delimitación de grupos que cumplan con homogeneidad, lo que lleva a un nuevo concepto al que se denomina *regiones homogéneas*. Para poder crear regiones homogéneas se debe contar con varios métodos que permitan hacer una comparación y ayude tomar la mejor decisión en cuanto a dicha delimitación.

3.3.1 Delimitación de regiones homogéneas

3.3.1.1 Ajuste potencial de áreas drenadas

Este método consiste en hacer dos agrupaciones de regiones homogéneas en toda la zona en estudio. Se grafican las áreas drenadas de cada cuenca y las medias de los gastos máximos instantáneos anuales. Posteriormente, se ajustan a una ecuación de tipo potencial.

La separación entre ambos grupos homogéneos estará dada por la curva de ajuste de la ecuación potencial encontrada. Un grupo homogéneo será aquel que se ubique por encima o debajo de la curva de ajuste, el eje donde se ubican las áreas se grafican en escala logarítmica para facilitar la identificación de los grupos, como se muestra en la Gráfica 2.



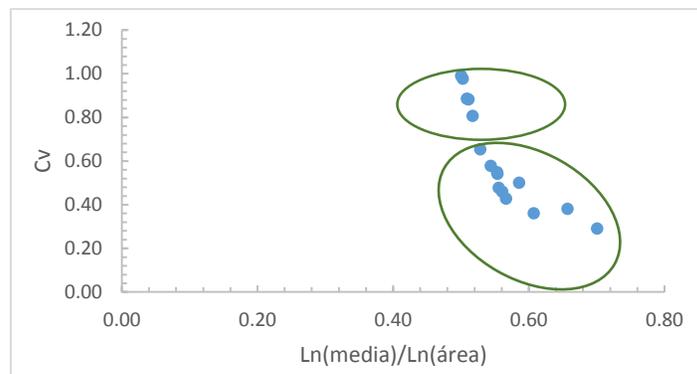
Gráfica 2. Ejemplo de ajuste potencial de áreas drenadas y gastos.

3.3.1.2 Coeficiente de variación

Como su nombre lo indica este método hace una comparación gráfica del coeficiente de variación y las áreas drenadas de cada sitio. Para cada serie de registros se calcula el coeficiente de variación C_v y su media. La media se utiliza para realizar una estandarización con las áreas y poder hacer una comparación con cada una de ellas.

Se grafican los valores de C_v contra la relación entre áreas drenadas y la media de cada sitio, se verifican si tienden a formar grupos alrededor de algún valor de C_v .

Este método requiere que los coeficientes de variación sean calculados de muestras con distribución normal y que además compartan el mismo periodo común, si esto último no ocurre la comparación de los C_v puede tener cierta desproporción, y con ello no lograrse formar los grupos. Por otra parte, si no se



Gráfica 3. Ejemplo de agrupación de los coeficientes de variación.

verifica que las series sean normales, se debe recurrir a algún proceso de normalización. Un ejemplo de la formación de grupos homogéneos aplicando el coeficiente de variación se muestra en la variación. Gráfica.

3.3.1.3 Trazos multidimensionales o Curvas de Andrews

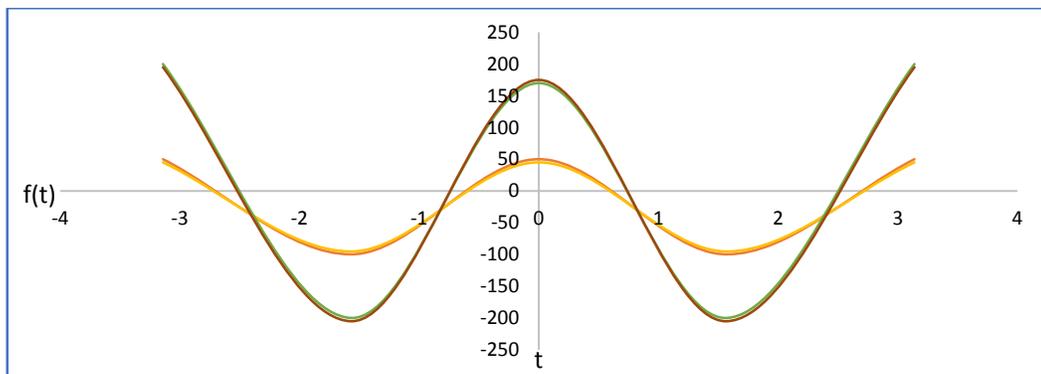
El grado de dispersión de grupos homogéneos se evalúan mediante un proceso de posicionamiento propuesto por Andrews (Andrews, 1972), en él se expone que un punto en el espacio multidimensional se representa por una curva en dos dimensiones de acuerdo con la siguiente función:

$$f(x) = \frac{x_1}{\sqrt{2}} + x_2 \sin(t) + x_3 \cos(t) + x_4 \sin(2t) + x_5 \cos(2t) + x_6 \sin(3t) + \dots \quad (3.28)$$

Donde $x_1, x_2, x_3, \dots, x_n$ son los atributos más significativos de cada sitio y la función se evalúa en el rango $-\pi \leq t \leq \pi$.

Como la función preserva las distancias de comparación esto la hace una técnica ideal de comparación visual para la localización de grupos homogéneos. Cuando varias curvas se agrupan, compartiendo una amplitud y periodo similar, se dice que forman parte de un grupo homogéneo.

Los resultados dependen en gran medida del orden de aparición de los atributos en la función. Se recomienda, para obtener mejores resultados realizar un análisis del estadístico t entre los atributos de cada sitio y los gastos de diseño estimados de manera puntual. Los atributos en la función multidimensional representan las componentes de baja frecuencia que generalmente son las más fáciles de observar.



Gráfica 4. Ejemplo de curvas de Andrews o Trazos multidimensionales.

3.3.1.4 Prueba F de Fisher

Para realizar la prueba se emplea la distribución F , con ella se prueba la hipótesis de que la varianza de una muestra pertenezca a una cierta población y de esta manera aceptar que dicha muestra pertenece a un grupo homogéneo.

Al considerar dos muestras aleatorias independientes de una población con distribución normal, la relación de las varianzas S_1^2/S_2^2 tiene una distribución F . Es decir, el estadístico F se puede calcular como el cociente de las varianzas de cada muestra.

$$F = \frac{S_1^2}{S_2^2} \quad (3.29)$$

Los grados de libertad gl_1 y gl_2 se calculan de la siguiente manera:

$$gl_1 = n_1 - 1 \quad (3.30)$$

$$gl_2 = n_2 - 1 \quad (3.31)$$

donde n_1 y n_2 son el tamaño de las muestras 1 y 2, respectivamente.

Cuando se comparan las varianzas el estadístico F tiene los mismos grados de libertad que su recíproco $1/F$. Tanto F como $1/F$ tienen el mismo tipo de distribución, es común tomar como numerador la varianza mayor en el cociente S_1^2/S_2^2 .

La distribución F de Fisher no es simétrica con respecto a la media, es decir, tiene una asimetría positiva (Ver Figura 1). Para facilitar los cálculos se considera que la región donde se rechaza la hipótesis es la parte asimétrica más alejada de la media.

Para la evaluación de los límites de aceptación y rechazo se consultan tablas para niveles de significancia del 5 y 1%. Los grados de libertad para el numerador se localizan en la columna izquierda, en tanto que para el denominador en la parte superior de dichas tablas.

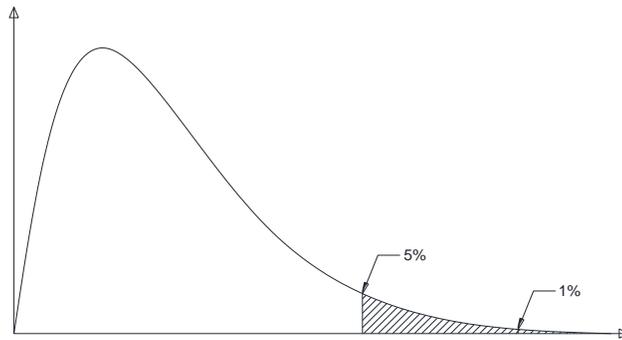


Figura 1. Distribución F de Fisher.

3.3.2 Técnicas regionales para la estimación de gastos de diseño

3.3.2.1 Estaciones-Año

Esta técnica considera el modelado de una sola serie de datos “Estaciones-año”, no requiere que las muestras analizadas tengan la misma longitud de registro.

Para un sitio con escasa información $n < 30$ años se debe contar con información de cuencas vecinas que pertenezcan a la misma región homogénea.

Para cada una de estas muestras se debe obtener su valor medio \bar{x}_j para $j=1, \dots, m$, donde m = número de estaciones involucradas.

1.- Se modulan las muestras de la forma:

Para los sitios A, B, C y D con registros de los años 1980 a 1994.

$$\begin{array}{cccc} \frac{Q_{1980}^A}{\bar{X}_A} & \frac{Q_{1980}^B}{\bar{X}_B} & \frac{Q_{1980}^C}{\bar{X}_C} & \frac{Q_{1980}^D}{\bar{X}_D} \\ & \cdot & & \\ & \cdot & & \\ & \cdot & & \\ \frac{Q_{1994}^A}{\bar{X}_A} & \frac{Q_{1994}^B}{\bar{X}_B} & \frac{Q_{1994}^C}{\bar{X}_C} & \frac{Q_{1994}^D}{\bar{X}_D} \end{array}$$

*Muestras moduladas.

2.- Se unen las muestras moduladas de cada sitio para crear una serie con n datos igual a la suma de n_A , n_B , n_C y n_D .

$$\frac{Q_{1980}^A}{\bar{X}_A}$$

$$\vdots$$

$$\frac{Q_{1994}^D}{\bar{X}_D}$$

3.- La nueva serie se ordena de mayor a menor.

4.- Se le asigna un periodo de retorno con la fórmula siguiente:

$$T_r = \frac{n + 1}{m} \tag{3.32}$$

Donde: $n = n_A + n_B + n_C + n_D$, T_r el periodo de retorno en años y m es el m -ésimo evento de los n registrados.

5.- Se obtiene una probabilidad de no excedencia.

$$P(X \leq x) = 1 - \frac{1}{T_r} \tag{3.33}$$

6.- Se realiza el análisis de frecuencias y se tabulan los periodos de retorno, donde la primera fila suele llamarse Regional del sitio.

	2	5	10	...	10000
$\left[\frac{Q}{\bar{X}}\right]_T$	a	b	c	...	d
$Q_T^A = \bar{X}_A \left[\frac{Q}{\bar{X}}\right]_T$	e	f	g	...	h
$Q_T^B = \bar{X}_B \left[\frac{Q}{\bar{X}}\right]_T$	i	j	k	...	l
$Q_T^C = \bar{X}_C \left[\frac{Q}{\bar{X}}\right]_T$	m	n	o	...	p
$Q_T^D = \bar{X}_D \left[\frac{Q}{\bar{X}}\right]_T$	a	r	s	...	t

7.- Si se cuenta con datos en el sitio en estudio, basta con multiplicar la regional del sitio por la media del lugar.

8.- Cuando no se cuenta con información del sitio, se puede estimar la media del gasto mediante el empleo de información de estaciones vecinas que pertenezcan a un mismo grupo homogéneo.

Observaciones:

1.- El procedimiento cuenta con un problema de escala, si se lleva la información de cuencas pequeñas a una cuenca grande los valores no serán realistas y viceversa.

2.- Se requiere al menos tres sitios, además del que está en estudio, para crear un modelo regional.

4. Metodología

4.1 Selección de estaciones y extracción de datos

Una vez definida la región hidrológica a estudiar, se seleccionan las estaciones que tengan por lo menos 20 años de registro y que además no se encuentren en zonas con obras de regulación que puedan afectar a la medición de gastos

Los datos se obtienen del Banco Nacional de Datos de Aguas Superficiales (BANDAS) de la Comisión Nacional del Agua y del Instituto Mexicano de Tecnología del Agua (IMTA); de este banco de información se extraen los datos de gastos máximos instantáneos ($Q_{\text{máx.inst.}}$) y los gastos medios diarios ($Q_{\text{med.diario}}$).

Antes de realizar el análisis estadístico y aplicar distintos conceptos teóricos, se debe efectuar una depuración de datos buscando anomalías o inconsistencias. Los criterios para la detección de datos dudosos en una estación pueden ser varios, por ejemplo, la comparación de la media con el máximo de la serie de datos que por lo general el máximo llega a ser hasta 5 veces la media, si ese no fuera el caso se debe indagar si el dato es válido o no.

Otro criterio a considerar en la eliminación de datos erróneos, es la comparación del valor mínimo con el máximo, que generalmente llega a ser el valor mínimo el 10 por ciento del valor máximo. También se puede hacer una comparación entre los valores de los gastos máximos instantáneos con respecto a los gastos medios máximos de un día. En los casos en que se observen mayores los medios diarios que los instantáneos se opta por eliminar el dato o bien se revisa la información de gastos máximos mensuales. Finalmente, cuando subsiste la duda y el valor correspondiente puede ser importante, se recurre a conseguir la información original (Domínguez Mora, *et al*, 2016).

4.2 Cálculo de los estadísticos

Para cada una de las estaciones hidrométricas se determinan los estadísticos, es decir, la media, la varianza, desviación estándar, coeficiente de variación, así como el valor máximo y mínimo.

4.3 Pruebas de homogeneidad e independencia

Las características estadísticas de las series hidrológicas como la media, la desviación estándar y los coeficientes de correlación serial, se afectan cuando la serie presenta cierta tendencia en la media o en la varianza. Tales anomalías son producidas por la pérdida de homogeneidad y la inconsistencia.

En general, la falta de homogeneidad de los datos es producto de las actividades humanas como la deforestación, aperturas de nuevas áreas al cultivo, rectificación de cauces, construcción de embalse y reforestación. También es consecuencia de

los procesos naturales súbitos tales como incendios forestales, terremotos, deslizamientos de laderas y erupciones volcánicas.

Ante la necesidad de corroborar la independencia y homogeneidad se aplicaran las siguientes pruebas:

- Prueba Estadística de Helmer
- Prueba *t* de Student
- Prueba estadística de Cramer
- Prueba de Independencia de Eventos

4.4 Obtención de rasgos fisiográficos de la cuenca

Para la realización de un análisis regional es imprescindible contar con información sobre las características propias de la cuenca, la obtención de las mismas se explica a continuación.

Área. En el BANDAS de la CONAGUA se puede consultar el área drenada de la cuenca. En ocasiones existen estaciones que no cuentan con dicho dato, en esos casos se calcula con un SIG (Sistema de Información Geográfica) como por ejemplo Arcmap© o Arcview©.

Tiempo de concentración. El tiempo de concentración (*tc*) se determina con la longitud (*L*) y pendiente del cauce principal (*S*) con la ecuación experimental de Kirpich (Aparicio, 2010).

$$tc = 0.000325 \frac{L^{0.77}}{S^{0.385}} \quad (4.1)$$

Donde: *tc* en horas., *L* en m y *S* en m/m, determinada con método de la pendiente media.

Precipitación media hp. Para el cálculo de este parámetro primero se identifican las estaciones climatológicas que se encuentren dentro o cerca de la cuenca asociada a cada estación hidrométrica analizada; se calcula la media de las precipitaciones diarias máximas anuales para cada estación climatológica y se obtiene su promedio espacial.

Número de escurrimiento N y retención potencial máxima S. El numero *N* se determina con el tipo y uso de suelo a partir de las tablas del U.S. Soil Conservation Service. A partir del número de escurrimiento se obtiene el espesor del suelo *S* (infiltración potencial máxima), la deducción detallada de este parámetro se puede consultar en Domínguez, *et al*, 2008, *S* se determina mediante la siguiente ecuación:

$$S = \frac{2540 - (25.4N)}{N} \quad (4.2)$$

donde: N es el número de curva y S en cm.

4.5 Agrupación de estaciones hidrométricas

Se consideran los grupos formados por cada una de las 37 regiones hidrológicas delimitadas por CONAGUA (Figura 2).



Figura 2. División de las 37 regiones hidrológicas del país (CONAGUA, 2016).

Posteriormente se realiza una subagrupación para la estaciones de cada región hidrológica, a esta agrupación se le denomina *Delimitación de regiones homogéneas*. Para disminuir la incertidumbre al emplear un solo método, se propone la utilización de algunos otros para tener una delimitación más precisa y fundamentada. Los métodos a emplear son los siguientes:

- Ajuste potencial
- Coeficiente de variación
- Trazos multidimensionales
- Prueba F de Fisher

4.6 Análisis estadístico regional

Normalización del registro para cada grupo de estaciones hidrométricas. Los datos anuales del gasto máximo instantáneo de cada estación hidrométrica se dividen entre su media y por cada grupo de estaciones hidrométricas se forma un solo registro extendido con todos los datos normalizados, el cual se trabaja como una serie anual.

La serie anual formada con los datos normalizados de cada grupo de hidrométricas se analiza estadísticamente ordenándolos de mayor a menor, se les asigna un periodo de retorno con la ecuación de Weibull (Aparicio, 2010) y se realiza el ajuste con distintas funciones de distribución de probabilidades, de estas se selecciona la que nos proporcione un mejor ajuste.

Con la función de distribución seleccionada se calculan los $Q_{m\acute{a}x.inst.}$ normalizados para distintos periodos de retorno. Los valores obtenidos para cada grupo de estaciones hidrométricas constituyen factores normalizados, de tal manera que al multiplicarlos por la media de los $Q_{m\acute{a}x.inst.}$ de cada estación hidrométrica del grupo correspondiente se obtiene una estimación regional de gastos máximos para distintos periodos de retorno (Domínguez Mora, *et al*, 2016).

Curvas regionales para estimar la media de los $Q_{m\acute{a}x.inst.}$ Se proponen modelos no lineales de tipo potencial como las siguientes:

$$\bar{Q}_{max.inst.} = a_1 A^{a_2} \quad (4.3)$$

$$\bar{Q}_{max.inst.} = a_1 A^{a_2} t a^{a_3} \quad (4.4)$$

$$\bar{Q}_{max.inst.} = a_1 V^{a_2} t a^{a_3} S^{a_4} \quad (4.5)$$

Donde: $\bar{Q}_{max.inst.}$ es la media de los gastos máximos instantáneos obtenida para cada estación hidrométrica del grupo de regiones consideradas en m^3/s , A es el área de la cuenca en km^2 , tc es el tiempo de concentración en horas., V es el volumen de escurrimiento en la cuenca, obtenido del producto entre el área de la cuenca y la precipitación, y S es la retención potencial máxima del suelo en cm. Los valores de precipitación asociados a distintos periodos de retorno y duración pueden consultarse en Domínguez, Carriosa, Fuentes, & Galván, 2010, en la que se usó la información de poco más de dos mil estaciones pluviométricas con más de 20 años de registro.

Los parámetros a de las ecuaciones potenciales se calculan minimizando el error cuadrático mediante algún algoritmo de optimización.

Selección del modelo. La selección del modelo debe hacerse tomando en cuenta la cantidad y calidad de la información de cada sitio de estudio. Para cada modelo se construyen gráficas que nos permitan hacer una comparación de los coeficientes de correlación entre los datos medidos, los datos calculados, se selecciona aquél que tenga una mejor correlación y en última instancia el modelo más simple.

Todo lo anterior se resume en lo siguiente:

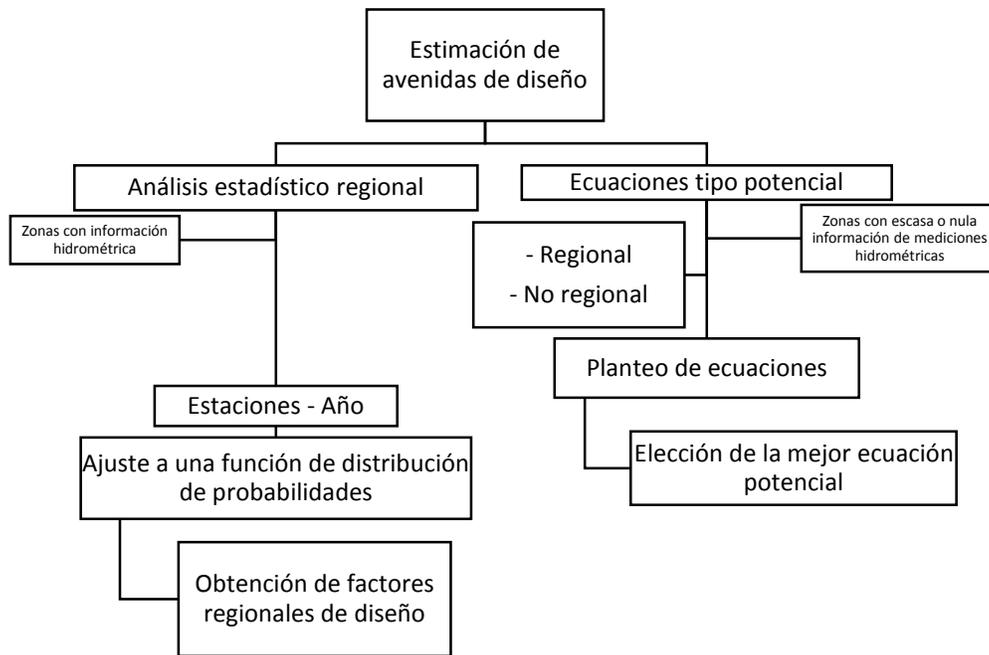


Figura 3. Resumen de la metodología para estimar avenidas de diseño.

5. Aplicación y resultados

5.1 Descripción de las zonas de estudio

5.1.1 Región hidrológica 18

La región hidrológica 18 se encuentra ubicada en el centro del país y tiene salida en la costa del pacífico, con un área de aproximadamente 117 405.60 km² comprende los estados de Guerrero, Michoacán, Jalisco y el estado de México.

Las zonas bajas y medias de la región están comprendidas hacia el norte por la cordillera neo-volcánica transversal, hacia el este con la sierra de Taxco y Chichihualco y hacia el suroeste por la sierra Madre del Sur.

La principal corriente de la región hidrológica es el río Balsas, este se origina en la unión del río Atoyac con el río Mixteco. Los principales afluentes son: río Tepecoacuilco, Huacapa, Cocula o Iguala, Sabinos, Poliutla, Ajuchitlan, Anuco, Cuiro, Cutzamala, entre otros.

5.1.2 Región hidrológica 22

Se encuentra en la región del istmo de Oaxaca, está limitada al norte con las cuencas de los ríos Papaloapan y Coatzacoalcos, al oriente con la región hidrológica 30 y 23, por el sur con el golfo de Tehuantepec. En la parte occidental se encuentra la cuenca del río Tehuantepec que escurre paralela al litoral del pacífico hasta encontrar salida cerca de Salina Cruz, Oaxaca.

El resto de corrientes siguen el comportamiento típico de los ríos de la vertiente del pacífico, lo que significa que fluyen de manera casi perpendicular a la costa.

Existen dentro de la región tres importantes accidentes hidrográficos que son: la Laguna superior, la Laguna Inferior y la Laguna Oriental, a través de las cuales, con excepción del río Tehuantepec, las corrientes de esta región drenan sus aguas al mar. Estas lagunas se ubican al final de una llanura aluvial sujeta a continua emersión debido al depósito de sedimentos.

El río de Tehuantepec es el más importante de la región debido a que su área drenada es considerablemente mayor al de otras cuencas de la misma zona, este río nace a una altitud de 3 200 m.s.n.m. y recorre aproximadamente 240 km hasta su desembocadura. Otros ríos de la región son: Los Perros, Chicapa, Niltepec, Ostuta, entre otros.

5.1.3 Región hidrológica 23

La región hidrológica 23 se localiza en la costa de Chiapas y parte de Oaxaca en la región del Istmo. Al noroeste limita con la región 22, al noreste con varios afluentes del río Grijalva de la región hidrológica 30, al sureste y suroeste con

Guatemala y el océano pacífico respectivamente. Tiene una superficie de aproximadamente 13 591 km².

Esta región está delimitada por la sierra Madre de Chiapas, misma que da origen a su línea de parteaguas. Tiene altitudes máximas de hasta 2900 m.s.n.m. en la sierra del Soconusco y 2500 m.s.n.m. en el Cerro de Tres Picos, al noreste de Tonalá, Chiapas.

Las corrientes de la región hidrológica 23 siguen el mismo patrón de comportamiento de los ríos de las vertientes del pacífico, esto es, con nacimiento en el parteaguas de la Sierra Madre de Chiapas y recorrido hacia el océano pacífico. De esta manera los escurrimientos forman un peine que se originan en la Sierra Madre de Chiapas y desembocan en el litoral del pacífico.

Los principales ríos de esta región son: Tapanatepec o Lagartero, Zanatenco, San Diego, De Jesús, Pijjiapan, Margaritas, Novillero, Sesecapa, Huixtla, Huehuetán, Coatán, Cahuacán, Suchiate, entre otros.



Figura 4. Localización de las regiones hidrológicas no. 18, 22 y 23.

5.2 Recopilación y análisis de la información

Se realizó una revisión previamente sobre las estaciones, cuidando que no se encontraran en obras o estructuras de regulación que pudieran provocar que los registros no se comporten de manera homogénea. Se descartaron aquellas estaciones con registros menores a 20 años por ser poco representativas para la finalidad de este estudio.

Para cada región hidrológica se hace una descripción de algunas anomalías encontradas en las MQMIA en cada una de las tres regiones hidrológicas. Las estaciones seleccionadas y sus respectivos estadísticos se muestran a continuación.

5.2.1 Región hidrológica 18

En la Tabla 1 se muestra la lista de las 16 estaciones seleccionadas en la región 18 mismas que serán empleadas para aplicarse la metodología descrita en el Capítulo 4.

Estaciones	Nombre
18129	Los Granjenos
18195	Ziritzicuaro
18201	El Cajon
18230	Santo Tomás
18281	La Limonera
18432	Ixcamilpa
18434	El Fraile
18439	San Juan Tetelcingo
18459	Rio Grande
18460	Rio Chiquito
18461	Zitácuaro
18481	La Caimanera
18487	Los Pinzanes
18494	Los Panches
18495	La Pastoria
18518	Mezcala

Tabla 1. Estaciones seleccionadas de la región hidrológica 18.

Inicialmente el Instituto de Ingeniería de la UNAM llevó a cabo un depuración de datos en estaciones hidrométricas de todo el país (Domínguez M. , Arganis, González, Carroza, & Guzmán, 2013), sin embargo, por tratarse de demasiadas estaciones se optó por realizar una nueva revisión y corroboración de los registros para las estaciones a estudiar. En el nuevo estudio de los datos se encontró lo siguiente:

Estación 18195.

En esta estación se halló un dato poco coherente, ocurrido en el mes de febrero del año 1974, de 13140 m³/s. La media, al ser una medida estadística estable, nos indicó que ese valor es exageradamente grande respecto a la media y no correspondía con la capacidad de drenaje de la cuenca. Se consultaron los gastos máximos instantáneos por mes del BANDAS donde se encontró que el valor de dicho gasto no pudo haber ocurrido para ese mes, ya que los gastos para otros años van de 10 a 13 m³/s. Las mediciones de los gastos mensuales son respaldadas con los valores de escala consultados en el mismo banco de datos, donde el valor de escala no concuerda con el dato de 13140 m³/s. Por lo tanto se tomó la decisión de tomar ese valor como un gasto con una magnitud de 13.14 m³/s, y el máximo para ese año sería sustituido por 555.80 m³/s congruente con los registros de otros años.

Estación 18201.

Se eliminó el valor del año 1961 por ser un valor muy pequeño para ese año. Se notó claramente, en los registros de estaciones cercanas, la observación de grandes escurrimientos.

Estación 18432.

En el año 1989 se localizó un máximo de 3210 m³/s que comparado con la media parecía ser un dato incierto, sin embargo, consultando los valores de escala y comparándolos con los ocurridos en el mismo mes, se encontró que ese dato si ocurrió para ese año. Se hicieron comparaciones con otras estaciones en cuanto a la magnitud de los escurrimientos registrados para el mismo periodo, resultando para todas con escurrimientos altos.

Estación 18460.

Se encontró un dato anómalo en el año 1974 de 875 m³/s que comparado con la media de 71.90 m³/s resulta ser un dato que difícilmente pudo haber ocurrido, por lo que se examinaron los datos de gastos máximos instantáneos mensuales. En ellos se encontró que ese registro ocurría en el mes de diciembre y que no corresponde con la media de los gastos para los meses de diciembre donde los gastos son menores a 10 m³/s, esto fue igualmente corroborado por los registros de escala. El dato fue descartado y el máximo para ese año cambió a 47.70 m³/s congruente para los datos de otros años de las demás estaciones cercanas.

Estación 18481.

En el año de 1998 se detectó un gasto muy grande de 9350 m³/s, se verificó con las lecturas de escala y se revisaron los registros de ese año en estaciones vecinas. Se comprobó que en el año 1967 se tiene registrado un evento similar, por lo que se opta por dejar dicho gasto.

Estación 18494.

Se encontró un gasto muy elevado de 4778 m³/s respecto a la media, ocurrido en 1966, al hacer la verificación correspondiente se halló que el dato es válido y se optó por dejarlo.

Estación 18518.

En esta estación, se encontró que los años 1978 y 1980 cuentan con el mismo registro, tanto en los datos anuales como en los mensuales. Se hizo una comparación con los datos de la estación vecina 18439 que cuenta con escurrimientos similares, y que además aforan la misma corriente, por lo que el dato correcto debió ser el de 1978, por lo tanto se eliminó el de 1980.

Además, se agregaron datos que no se encontraban en el registro cuando se hizo el primer análisis. Los datos agregados son del periodo del año 2004 a 2007 para las estaciones 18434 y 18439; para el periodo 2004 a 2009 en las estaciones 18481, 18487, 18494 y 18495.

Finalmente los estadísticos de las estaciones con los datos corregidos se presentan en la Tabla 2.

Estación	Nombre	No. De datos	Máx.	Media	Mín.	Varianza	SD	Cv
18129	Los Granjenos	42	258.00	50.96	11.00	2480.31	49.80	0.98
18195	Ziritzécuaro	21	1230.20	534.52	155.00	83553.76	289.06	0.54
18201	El Cajón	26	1116.00	264.98	28.12	58603.73	242.08	0.91
18230	Santo Tomás	41	3881.00	1569.48	405.21	560212.61	748.47	0.48
18281	La Limonera	31	673.56	316.48	173.00	13018.37	114.10	0.36
18432	Ixcamilpa	49	3210.00	586.30	159.00	223380.97	472.63	0.81
18434	El Fraile	42	1569.74	488.00	184.00	71848.76	268.05	0.55
18439	San Juan Tetelcingo	57	3264.00	1312.92	382.00	316448.50	562.54	0.43
18459	Rio Grande	54	136.00	74.88	40.50	471.59	21.72	0.29
18460	Rio Chiquito	52	107.75	55.99	12.00	457.01	21.38	0.38
18461	Zitácuaro	30	297.47	112.92	40.90	3188.32	56.47	0.50
18481	La Caimanera	51	9906.00	2556.54	826.00	2798393.85	1672.84	0.65
18487	Los Pinzanes	44	5422.00	868.53	146.00	739975.67	860.22	0.99
18494	Los Panches	44	4778.00	1114.22	128.00	974196.89	987.01	0.89
18495	La Pastoria	45	2873.00	879.54	335.00	258562.25	508.49	0.58
18518	Mezcala	46	2771.46	1278.61	325.88	345181.38	587.52	0.46

Tabla 2. Resumen de estadísticos de la RH 18.

5.2.2 Región hidrología 22

Las regiones seleccionadas para este estudio son las siguientes:

Estación	Nombre
22008	Boquilla Número 1
22015	Tequisistlán
22016	Chicapa
22017	Ixtepec
22018	Ostuta
22026	Zanatepec
22028	Niltepec

Tabla 3. Lista de estaciones de la RH-22.

Al igual que en la región hidrológica anterior se detectaron registros demasiado pequeños en relación con la media mismos que fueron descartados. En la revisión de los datos se encontró lo siguiente:

Estación 22008

En esta estación no se detectaron registros que generaran duda, únicamente se agregaron nuevos datos con registros mensuales completos en los años 2012 y 2014.

Estación 22015

No se encontró dato alguno que pudiera ser erróneo, sin embargo, se agregaron cuatro nuevos registros para los años 2010, 2011, 2012 y 2014.

Estación 22016 y 22018

Se agregaron dos nuevos datos en los años 2010 y 2011.

Estación 22017

Se agregaron nuevos registros para los años 2012, 2013 y 2014.

Estación 22026

No se encontraron datos que generaran incertidumbre al compararlos con la media, sin embargo, se agregó un nuevo valor para el año 2011.

Los estadísticos de las estaciones con los datos verificados y agregados se presentan en la Tabla 4.

Estación	Nombre	No. De datos	Máx.	Media	Mín.	Varianza	SD	Cv
22008	Boquilla Número 1	73	1100.00	215.18	18.43	32075.19	179.10	0.83
22015	Tequisistlán	60	1440.00	394.08	22.30	152749.13	390.83	0.99
22016	Chicapa	58	2455.43	366.38	20.61	210292.54	458.58	1.25
22017	Ixtepec	60	1625.00	226.23	7.32	79909.60	282.68	1.25
22018	Ostuta	55	3171.00	548.37	62.34	391867.91	625.99	1.14
22026	Zanatepec	51	2400.00	244.17	29.72	119635.06	345.88	1.42
22028	Niltepec	41	479.00	89.12	5.26	12708.46	112.73	1.26

Tabla 4. Estadísticos para la RH-22.

5.2.3 Región hidrológica 23

Las estaciones seleccionadas son las siguientes:

Estación	Clave
23003	Cahuacán
23006	Malpaso
23007	Suchiate II
23008	Huixtla
23009	Pijjiapan
23011	Tonalá
23012	Novillero
23013	Jesús
23014	San Diego
23015	Cacaluta
23016	Cintalapa
23017	Coapa
23018	Margaritas
23019	Despoblado
23022	San Nicolás
23023	Talismán II

Tabla 5. Relación de estaciones RH-23.

Para esta región los resultados del análisis de los registros se presentan a continuación.

Estación 23003.

Se optó por descartar registros de gastos máximos instantáneos anuales por no contar con el registro mensual completo que incluya a los meses donde se presentan dichos máximos, como ocurre en estaciones cercanas. Se eliminaron los

años 1993 a 1997, 2001 y de 2003 a 2005, que presentaban dicho problema. No se agregaron nuevos datos.

Estación 23008 y 23015.

Se agregaron datos para el periodo comprendido de 2011 a 2014.

Estación 23009.

Se descartó el registro de 1995 dado que solo se tenía registro de un mes, igualmente el de 1998 por solo contar con registro del mes de enero. Los años 2002 y 2008 se eliminaron debido a que eran menores al gasto medio diario de un día. Se agregaron nuevos datos para los años 2011, 2012 y 2013.

Estación 23012.

Se suprimieron los datos de los años 2004 y 2005 por que no tenían registro de datos mensuales y diarios. Se agregaron nuevos datos para los años de 2010 a 2014.

Estación 23013.

Se eliminó el dato del año 1976 ya que solo contaba con un mes de registro y además era un valor muy pequeño respecto a la media.

Estación 23016.

De igual manera se agregaron 2 años al registro correspondientes al 2010 y 2011.

Estación 23018.

Para esta estación se eliminó el registro del año 2005, ya que únicamente cuenta con datos de los meses de julio, agosto y septiembre y se tienen antecedentes de inundaciones provocadas por el huracán Stan en la zona en el mes de octubre de ese mismo año.

Estación 23019.

Se añadieron nuevos datos al registro de los años 2011 y 2012.

Estación 23022.

Se agregaron los años 2010 al 2014 a la serie de datos.

Los estadísticos de las estaciones con los datos verificados y agregados, se presentan en la Tabla 6.

Estación	Nombre	No. De datos	Máx.	Media	Mín.	Varianza	SD	Cv
23003	Cahuacán	48	498.92	173.93	62.60	7178.92	84.73	0.49
23006	Malpaso	35	1327.00	289.28	55.41	69936.39	264.45	0.91
23007	Suchiate II	38	2200.00	949.72	205.10	176354.82	419.95	0.44
23008	Huixtla	45	702.00	260.76	23.09	25851.97	160.79	0.62
23009	Pijjiapan	45	1004.80	239.99	27.05	32252.39	179.59	0.75
23011	Tonalá	28	741.16	169.41	26.36	19616.78	140.06	0.83
23012	Novillero	34	450.00	226.57	41.82	13486.58	116.13	0.51
23013	Jesús	36	140.00	57.97	11.20	1018.16	31.91	0.55
23014	San Diego	33	480.70	204.46	18.65	14943.58	122.24	0.60
23015	Cacaluta	45	479.75	167.32	35.06	9949.66	99.75	0.60
23016	Cintalapa	36	738.93	154.42	24.55	18231.32	135.02	0.87
23017	Coapa	34	481.41	178.61	13.22	17136.20	130.91	0.73
23018	Margaritas	33	703.14	227.35	24.81	35921.75	189.53	0.83
23019	Despoblado	43	591.20	226.69	20.68	25480.67	159.63	0.70
23022	San Nicolás	32	339.73	112.80	35.95	5100.47	71.42	0.63
23023	Talismán II	24	1374.65	387.97	100.20	143984.93	379.45	0.98

Tabla 6. Estadísticos para la RH-23.

Los datos de las estaciones de las regiones hidrológicas fueron obtenidos del BANDAS (Banco Nacional de Datos de Aguas Superficiales) de la Comisión Nacional del Agua (CONAGUA). En general, la revisión de los datos contenidos en dichos registros consistió en descartar valores que en ocasiones no eran congruentes a la capacidad de la misma cuenca, valores muy elevados o mediciones muy pequeñas.

5.3 Pruebas de homogeneidad e independencia

Se aplicaron tres pruebas de homogeneidad (Helmert, t de Student y Cramer), a cada una de las estaciones de las tres regiones hidrológicas en estudio. Estas pruebas nos ayudarían a comprobar si la serie de registros es estacionaria o no, es decir, si existe cierta tendencia en el comportamiento de los registros provocada por distintos factores naturales o humanos.

Además, se aplicó la prueba de independencia para comprobar que en realidad pertenezcan a una serie de variables aleatorias.

5.3.1 Región hidrológica 18

Estación	Nombre	No. De datos	Pruebas de homogeneidad			Independencia
			Helmert	t de Student	Cramer	Anderson
18129	Los Granjenos	42	X	✓	✓	✓
18195	Ziritzcuaro	21	✓	✓	✓	✓
18201	El Cajón	25	✓	✓	X	✓
18230	Santo Tomás	41	✓	✓	✓	✓
18281	La Limonera	31	✓	✓	X	✓
18432	Ixcamilpa	49	X	✓	✓	✓
18434	El Fraile	42	✓	✓	✓	✓
18439	San Juan Tetelcingo	57	✓	✓	✓	✓
18459	Río Grande	54	✓	✓	X	✓
18460	Río Chiquito	52	✓	X	✓	✓
18461	Zitacuaro	30	✓	✓	✓	✓
18481	La Caimanera	51	✓	✓	✓	✓
18487	Los Pinzanes	44	✓	✓	✓	✓
18494	Los Panches	44	X	✓	✓	✓
18495	La Pastoria	45	✓	✓	✓	✓
18518	Mezcala	46	X	✓	X	✓

Tabla 7. Resumen de resultados de las pruebas de homogeneidad e independencia, RH-18.

Para la región hidrología 18, de acuerdo a los resultados mostrados en la Tabla 7 (el símbolo “✓” significa que pasaron la prueba, mientras que las marcadas con “X” no), se comprueba que las muestras son homogéneas puesto que pasaron dos o más pruebas. La estación 18518 únicamente pasa la prueba de t Student, esto puede deberse a cambios de uso de suelo o deforestación a lo largo de tiempo, sin embargo, al no contar con un registro mayor a 60 años la prueba no es concluyente (Dalrymple, 1960). Además, es importante señalar que cuando se considera a una estación carente de homogeneidad es común caer en una subestimación de gastos para periodos de retorno relativamente grandes.

En cuanto a la prueba de independencia se confirma que hay poca correlación entre los datos, ya que todas las estaciones pasaron las pruebas, por lo tanto, corresponde a una serie de variables aleatorias.

5.3.2 Región Hidrológica 22

Estación	Nombre	No. De datos	Pruebas de homogeneidad			Independencia
			Helmert	t de Student	Cramer	Anderson
22008	Boquilla Número 1	73	X	✓	X	✓
22015	Tequisistlán	60	X	✓	✓	✓
22016	Chicapa	58	X	✓	✓	✓
22017	Ixtepec	60	X	✓	✓	✓
22018	Ostuta	55	X	✓	✓	✓
22026	Zanatepec	51	X	✓	✓	✓
22028	Niltepec	41	X	✓	X	X

Tabla 8. Resumen de resultados de las pruebas de homogeneidad e independencia, RH-18.

De acuerdo a la Tabla 8, la mayoría cumple con las pruebas de homogeneidad, salvo las estaciones 22008 y 22028, mismos que solo pasan un prueba. Esto se debe posiblemente, y de acuerdo a imágenes satelitales tomadas de Google Earth®, al cambio de uso de suelo aguas arriba de las estaciones hidrométricas. Por ejemplo, en la zona aguas arriba de la estación 22028-Niltepec se encuentran zonas agrícolas y sitios con una marcada deforestación tal como se puede apreciar en la Figura 5 . Se deberá tomar en cuenta el no subestimar gastos de diseño para sitios cercanos a esas dos estaciones.

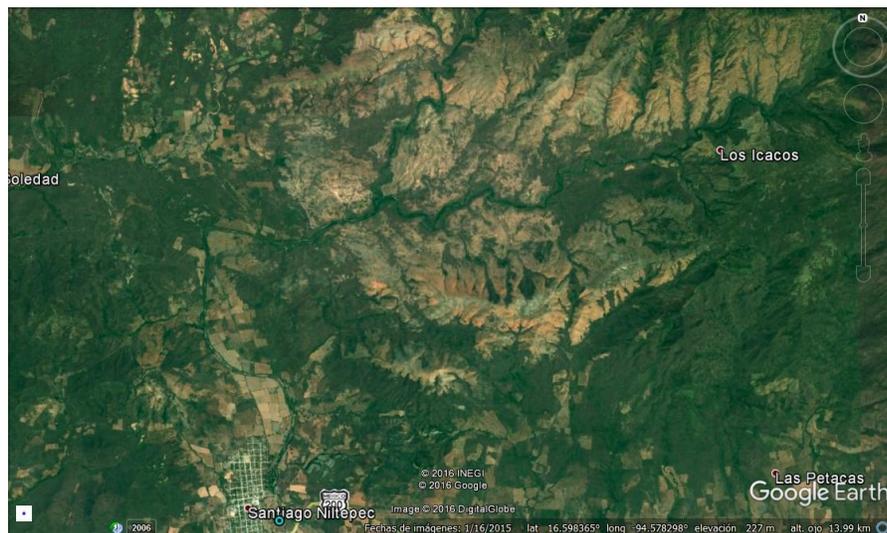
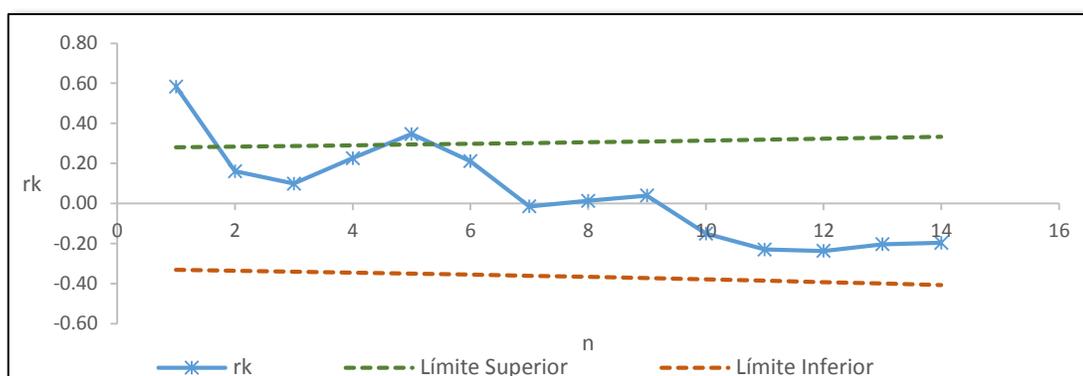


Figura 5. Imagen satelital de la zona de aguas arriba de la estación 22018-Niltepec (Google Earth®).

En cuanto a la prueba de independencia de Anderson, la estación 22028 no pasó la prueba, sin embargo, los datos que quedaron fuera del correlograma se encuentran muy cercano a los límites marcados considerando un 95% de límite de confianza. La hipótesis de aceptación de la prueba consiste en que al menos el 10% de un tercio de los valores de la muestra deben quedar fuera de los límites, para esta estación 1.4 valores deben quedar fuera.

Como se aprecia en la Gráfica 5, siendo muy rigurosos la prueba no se acepta porque dos valores se encuentran fuera, sin embargo si se tomase un límite de confianza menor, por ejemplo del 90% la prueba sería aceptada. Considerando lo anterior y sabiendo que no contamos con una muestra lo suficientemente grande como para aplicar una rigurosa estadística, para fines de este trabajo, se puede considerar como una muestra de variables aleatorias.



Gráfica 5. Correlograma de la muestra para la estación 22028.

5.3.3 Región Hidrológica 23

Región 23	Nombre	No. De datos	Pruebas de homogeneidad			Independencia
			Helmert	t de Student	Cramer	Anderson
23003	Cahuacán	48	X	✓	✓	✓
23006	Malpaso	35	X	✓	✓	✓
23007	Suchiate ii	38	X	✓	X	✓
23008	Huixtla	45	X	✓	X	✓
23009	Pijijiapan	45	✓	✓	X	✓
23011	Tonalá	28	✓	✓	✓	✓
23012	Novillero	34	X	✓	X	X
23013	Jesús	36	✓	✓	✓	✓
23014	San diego	33	✓	X	✓	X
23015	Cacaluta	45	X	✓	X	✓
23016	Cintalapa	36	X	✓	X	✓
23017	Coapa	34	X	✓	X	X
23018	Margaritas	33	X	✓	✓	✓
23019	Despoblado	43	✓	✓	X	X
23022	San Nicolás	32	X	✓	X	✓
23023	Talismán ii	24	X	✓	✓	✓

Tabla 9. Resumen de resultados de las pruebas de homogeneidad e independencia, RH-23.

En este caso se observa que en todas las estaciones pasan por lo menos una de las pruebas de homogeneidad, el hecho de que no pasen las demás puede ser a los cambios físicos en las cuencas. Se vuelve a recalcar en la no subestimación de los gastos de diseño estimados, para un periodo de retorno elevado, en las zonas donde no pasan más de una prueba.

En la prueba de independencia se observa que las estaciones 23012, 23014, 23017 y 23019 no cumplen la prueba, sin embargo, esto no es concluyente debido a los pocos datos ya que al menos se debe contar con un registro de 60 datos (Dalrymple, 1960) ocurriendo algo semejante a lo explicado en la región 22. Por otra parte, las estaciones consideradas para las tres regiones hidrológicas fueron seleccionadas de tal manera que no tuvieran influencia ajena a la naturaleza propia del cauce, por lo tanto, para fines de este trabajo todos los registros se tomaran como series de variables aleatorias.

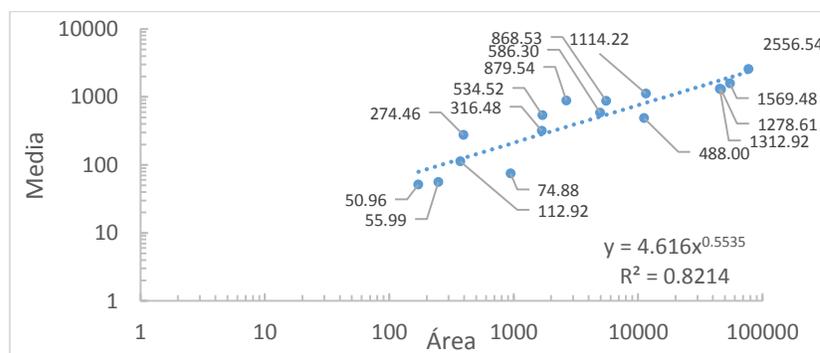
5.4 Identificación de grupos homogéneos

Una vez revisados los datos de cada estación se procedió a identificar grupos homogéneos de cada región hidrológica, para esto se emplearon cuatro técnicas de agrupación.

5.4.1 Región hidrológica 18

Al usar más de un método para la delimitación de las zonas homogéneas, se tiene un panorama más amplio sobre las estaciones que, en ocasiones, resulten difíciles de agrupar.

Ajuste potencial. Como se ha descrito anteriormente el método consiste en relacionar el área drenada de cada cuenca con la media de los gastos máximos instantáneos, se grafican y se hace el ajuste a una curva de tipo potencial. Los grupos de estaciones se forman según si los valores de los gastos quedan por arriba o por debajo de dicha curva de ajuste.



Gráfica 6. Curva de ajuste del tipo potencial, RH-18.

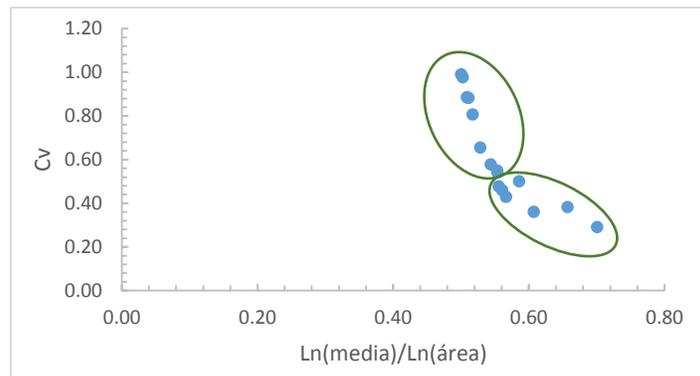
Media	Área	Estación
50.96	171.40	18129
55.99	247.80	18460
74.88	943.50	18459
112.92	372.50	18461
274.46	395.00	18201
316.48	1681.30	18281
488.00	11175.50	18434
534.52	1701.40	18195
586.30	4946.60	18432
868.53	5525.10	18487
879.54	2655.90	18495
1114.22	11596.40	18494
1278.61	46530.00	18518
1312.92	45361.70	18439
1569.48	54697.00	18230
2556.54	77312.00	18481

Tabla 10. Estaciones sombreadas localizadas debajo de la curva de ajuste.

Con el método del coeficiente de variación se pudieron identificar dos regiones integradas por las estaciones de la Tabla 10, tomando como punto de separación un Cv (coeficiente de variación) igual a 0.55 y la agrupación que resulta de graficar con el cociente de los logaritmos naturales del área y la media.

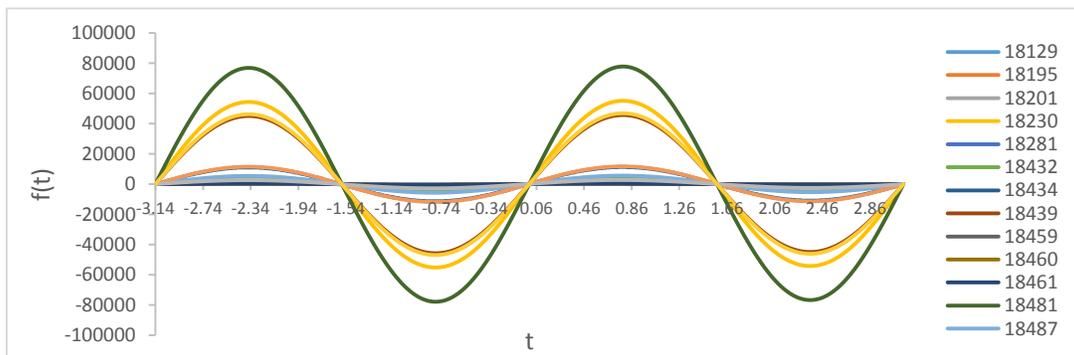
Clave	Cv	Ln(Media)	Ln(Área)	Ln(media)/Ln(Área)
18459	0.29	4.32	6.16	0.70
18281	0.36	5.76	9.47	0.61
18460	0.38	4.03	6.12	0.66
18439	0.43	7.18	12.66	0.57
18518	0.46	7.15	12.75	0.56
18230	0.48	7.36	13.24	0.56
18461	0.50	4.73	8.07	0.59
18195	0.54	6.28	11.33	0.55
18434	0.55	6.19	11.18	0.55
18495	0.58	6.78	12.46	0.54
18481	0.65	7.85	14.84	0.53
18432	0.81	6.37	12.32	0.52
18201	0.88	5.61	10.98	0.51
18494	0.89	7.02	13.79	0.51
18129	0.98	3.93	7.82	0.50
18487	0.99	6.77	13.51	0.50

Tabla 11. Resultados obtenidos por el método del coeficiente de variación.



Gráfica 7. Grupos formados de coeficientes de variación.

En tanto que con el método de las curvas de Andrews se obtuvieron los resultados mostrados en la Gráfica 8.



Gráfica 8. Curvas de Andrews para la región hidrológica no. 18.

Para la construcción de las curvas se emplearon las características fisiográficas en el orden siguiente:

1. Área (km²)
2. Pendiente media de la cuenca (m/m)
3. Longitud del cauce principal (km)
4. Tiempo de concentración (horas)
5. Nc promedio
6. S espesor del suelo (cm)
7. Hp media (mm)

El orden se obtuvo al comparar el estadístico t después de realizar un proceso de regresión entre las características fisiográficas y los eventos de diseño calculados de manera puntual.

Las Curvas de Andrews indican que aquellas estaciones que pertenezcan a una región homogénea tenderán a formar grupos de curvas en una especie de racimos. Como se muestra en la Gráfica 8, las curvas forman claramente dos grupos homogéneos, el primero conformado por las estaciones 18518, 18481, 18230 y 18439, mientras el segundo grupo lo forman las estaciones restantes que se agrupan entorno al eje de las abscisas.

Finalmente, se emplea el método de la prueba F de Fisher, para esto se usaron los datos registrados de los gastos máximos instantáneos anuales (QMIA) en cada estación y se dividieron entre la media de cada una de ellas, para el eliminar las diferencias de escala producidas por características particulares de cada estación. Posteriormente se calcularon los cocientes de las varianzas entre estaciones y se agruparon de acuerdo al límite dado por la distribución F de Fisher.

Región 1	Región 2
18201	18129
18494	18281
18487	18495
18481	18195
	18459
	18461
	18460
	18230
	18518
	18439
	18432
	18434

Tabla 12. Agrupación por método de Fisher.

Tomando en cuenta la agrupación de los cuatro métodos, se identificaron dos regiones homogéneas conformadas por las estaciones que se muestran en la Tabla 13.

De los resultados se pudo apreciar que los métodos de ajuste exponencial y la prueba F de Fisher dieron resultados similares. Los resultados de las curvas de Andrews se usaron para confirmar o rechazar la permanencia de algunas estaciones dentro de alguna de las dos regiones encontradas, por ejemplo las estaciones 18495 y 18195 quedaban en la región 1, pero revisando las agrupaciones por curvas de Andrews se confirmó que correspondían al grupo 2, lo que resulta congruente cuando se localizan las estaciones geográficamente (Ver Figura 6).

Región 1	Región 2
18201	18129
18494	18281
18487	18495
18481	18195
	18459
	18461
	18460
	18230
	18518
	18439
	18432
	18434

Tabla 13. Estaciones de cada una de las regiones homogéneas de la RH-18.

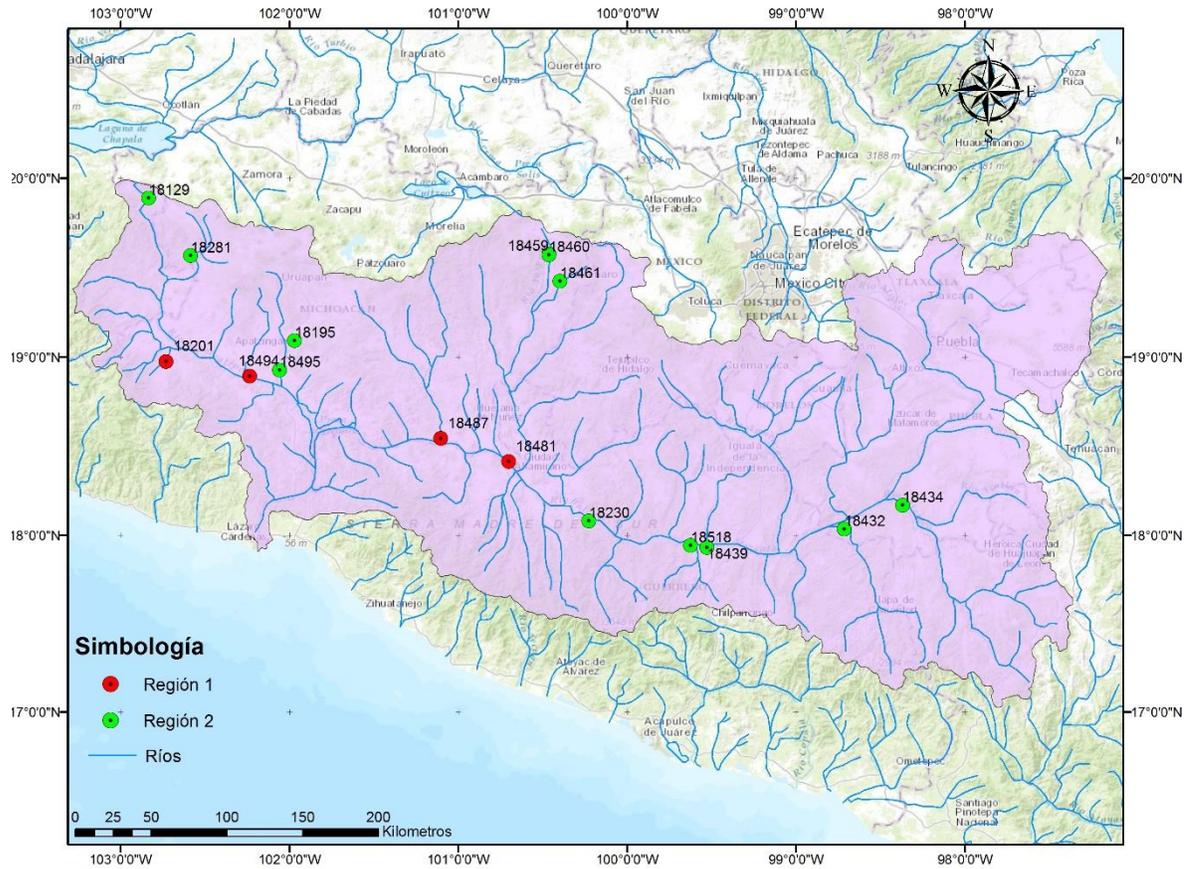


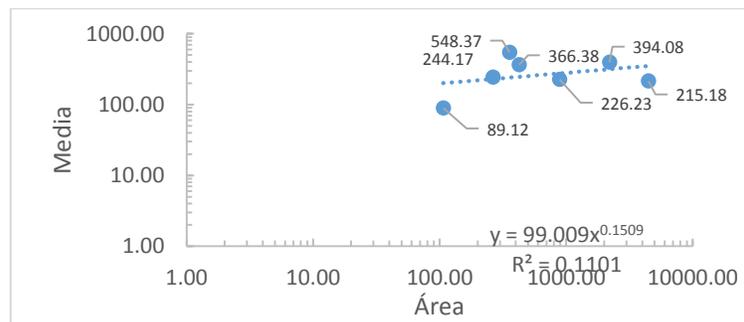
Figura 6. Regiones homogéneas para la Región hidrológica no. 18.

Una vez obtenida la agrupación de regiones homogéneas se procedió a verificar, mediante la prueba de Fisher, si correspondían a una misma población. Al aplicar la prueba se encontró que no pasan dicha prueba, sin embargo, esto no es concluyente. Al tomar en cuenta únicamente la delimitación de regiones homogéneas únicamente con el método de Fisher las estaciones quedaban dispersas, por lo que hablar de una región homogénea, basándose únicamente en este método, no sería posible.

Garantizar que las regiones homogéneas pasen la prueba de Fisher, nos da una mayor certeza si queremos reproducir estadísticos similares al momento de generar muestras sintéticas, estas últimas se explicarán más adelante.

5.4.2 Región hidrológica 22

En la región 22, se identificaron dos regiones homogéneas, el primer método fue el potencial, donde se apreció la formación de dos grupos homogéneos. El primer grupo (puntos por encima de la curva de ajuste potencial), están formados por las estaciones 22015, 22016, 22018 y 22026; el segundo grupo (puntos debajo de la curva de ajuste potencial) integrado por las estaciones 22008, 22017 y 22028. Cabe mencionar que para una mejor visualización de las agrupaciones el eje de las abscisas (área) fue puesto en escala logarítmica.

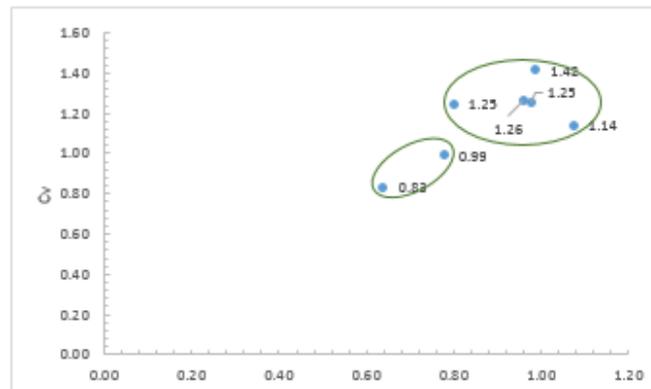


Gráfica 9. Ajuste a una curva potencial de la región hidrológica no. 22.

El segundo método de identificación, aplicado a esta región, fue el del coeficiente de variación. Al graficar los coeficientes de variación Cv con el cociente de los logaritmos de la media los gastos máximos instantáneos anuales y del área, se identificaron dos grupos homogéneos.

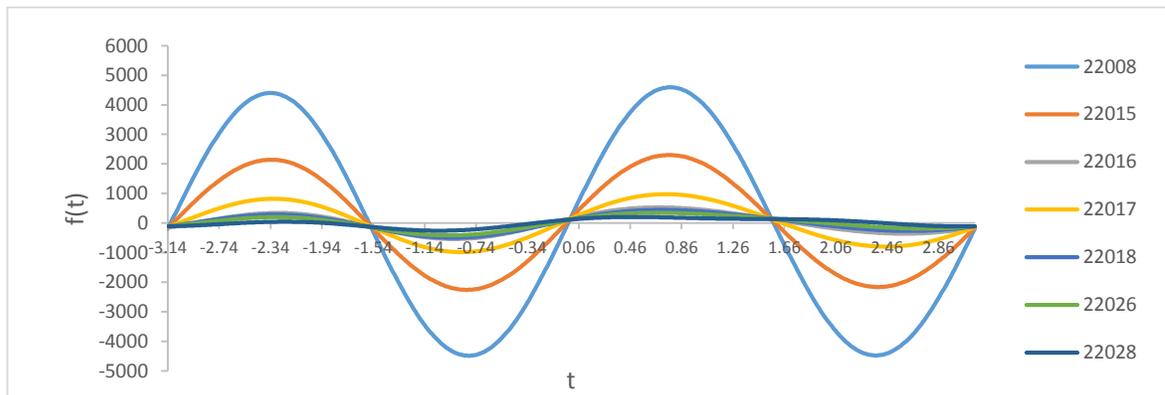
Clave	Cv	Ln(Media)	Ln(Área)	Ln(media)/Ln(Área)
22008	0.83	5.37	8.41	0.64
22015	0.99	5.98	7.70	0.78
22016	1.25	5.90	6.05	0.98
22017	1.25	5.42	6.79	0.80
22018	1.14	6.31	5.88	1.07
22026	1.42	5.50	5.58	0.99
22028	1.26	4.49	4.67	0.96

Tabla 14. Coeficientes de variación y logaritmos del área y los QMIA, Rh-22.



Gráfica 10. Grupos formados por el método del coeficiente de variación, Rh-22.

Al aplicar el método de las curvas de Andrews se identificaron dos grupos, el formado por las estaciones 22008, 22015 y 22017 resulta estar un poco disperso (Ver Gráfica 11), pero en base a los resultados mostrados en el método del coeficiente de variación, el grupo homogéneo de las curvas dispersas estaría formado únicamente por las estaciones 22008 y 22015, mientras que el resto integra la segunda región homogénea.



Gráfica 11. Curvas de Andrews, RH-22.

Por último, para esta región hidrológica el método de la prueba de Fisher arrojó para un nivel de significancia de 0.05 una $F = 1.51$, siendo que las únicas cocientes de varianzas que dieron menor al valor de la prueba de Fisher fueron las estaciones 22008 y 22015, se asume que estas forman parte de una región homogénea, quedando las restantes en otro grupo homogéneo.

Analizando las agrupaciones de cada uno de los métodos, se encontró que el método del ajuste potencial difiere de los otros, donde la estación 22017 parece pertenecer a la región comprendida por las estaciones 22008 y 22015, sin embargo, las curvas de Andrews, los coeficientes de variación y la prueba F de Fisher muestran que en realidad corresponde al segundo grupo homogéneo. En la Figura 7 se pueden observar las estaciones agrupados en dos regiones homogéneas.

Se puede afirmar que las estaciones 22008 y 22015 forman parte de una misma región por ubicarse en una zona geográficamente distinta a las demás estaciones, por ejemplo, estas dos estaciones se encuentran sobre afluentes del río Tehuantepec, en cambio las de la otra región homogénea corresponden a corrientes independientes que tienen un comportamiento similar a los ríos de la región hidrológica 23 como se verá más adelante.

Revisando nuevamente a los grupos de estaciones de cada región homogénea se encontró que pasan la prueba de Fisher, en ambos grupos se tiene el mismo resultado. Esto ocurre debido a que se obtuvieron resultados similares con el método de Fisher y las otras de técnicas. En consecuencia, se puede esperar que al generar muestras sintéticas, generadas a partir de alguna de las dos regiones homogéneas, se puedan reproducir estadísticos similares a los de la muestra original.

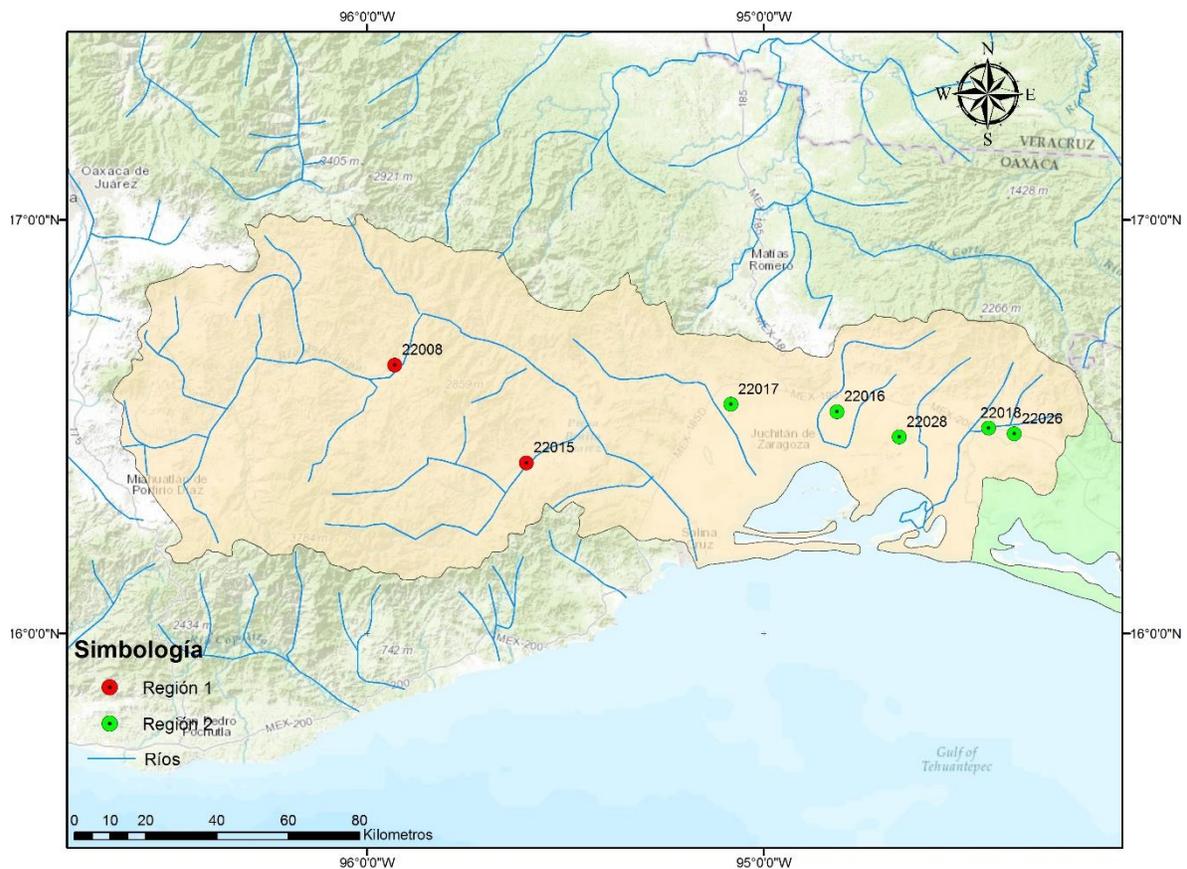


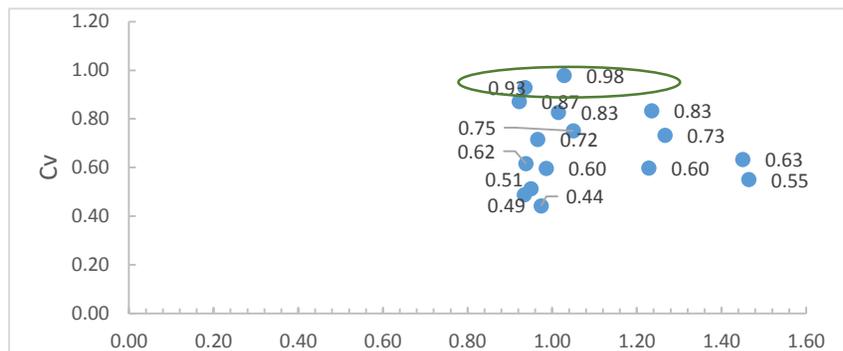
Figura 7. Regiones homogéneas para la región hidrológica no. 22.

5.4.3 Región hidrológica 23

En esta región la delimitación de las regiones homogéneas no fue tan clara, ya que los métodos anteriormente descritos no dan resultados coherentes al localizar los grupos geográficamente, es decir, estaciones pertenecientes a un mismo grupo quedan muy separados uno de otros. Las estaciones de la región se ubican sobre escurrimientos que se desarrollan de forma casi paralelas una de otras, por lo que no son tributarias a un cauce principal, posiblemente se una de las causas por las que no se pudo reconocer fácilmente una región homogénea. Por otro lado, las estaciones se encuentran a una distancia promedio de la costa del pacifico de 20 km, siendo la causa principal de los escurrimientos las lluvias de tipo orográfica.

Las técnicas de delimitación de regiones homogéneas no mostraron resultados consistentes, salvo el del coeficiente de variación y las curvas de Andrews mismo que a continuación se presentan.

La separación de dos zonas homogéneas se hizo agrupando estaciones con un coeficiente de variación cercano a uno. En este caso las estaciones 23006 y 23023, que tienen un coeficiente de variación de 0.93 y 0.98 respectivamente, formaran parte de una región homogénea de acuerdo al método del coeficiente de variación.



Gráfica 12. Grupos formados por el método del coeficiente de variación, Rh-23.

Cv	Clave	Ln(Media)	Ln(Área)	Ln(media)/Ln(Área)
0.44	23007	6.86	7.04	0.97
0.49	23003	5.16	5.52	0.93
0.51	23012	5.42	5.71	0.95
0.55	23013	4.06	2.77	1.46
0.60	23015	5.10	5.17	0.99
0.60	23014	5.32	4.33	1.23
0.62	23008	5.56	5.93	0.94
0.63	23022	4.73	3.26	1.45
0.72	23019	5.42	5.61	0.97
0.73	23017	5.19	4.09	1.27
0.75	23009	5.49	5.23	1.05
0.83	23011	5.13	5.06	1.02
0.83	23018	5.43	4.39	1.23
0.87	23016	5.04	5.46	0.92
0.93	23006	5.67	6.05	0.94
0.98	23023	5.96	5.80	1.03

Tabla 15. Coeficientes de variación y logaritmos del área y los QMIA, Rh-23.

En la Gráfica 12 se puede observar un grupo formado por las estaciones 23006 y 23023, se tomó ese criterio de agrupación ya que al agregar más estaciones estas quedan alejadas de la zona homogénea geográfica, lo cual no tiene sentido dado los objetivos de este estudio. En los demás métodos no se pudo observar una agrupación clara y coherente.

La delimitación de la región homogénea por el método del coeficiente de variación muestra que dos de las estaciones se separan de las demás. Este resultado es coherente dado que dichas estaciones se encuentran tierra adentro y a una elevación mayor que el resto de las demás, lo que explicaría que tengan un coeficiente de variación alto.

Por otra parte, la estación 23006 y 23023, que se ubican sobre el río Coatán y el río Suchiate respectivamente, tienen las mayores áreas drenadas salvo la estación 23007 que se encuentra sobre el río Suchiate, este último no entra dentro de esta región puesto que tiene el coeficiente de variación más bajo con 0.44 debido a la zona de planicie donde se localiza.

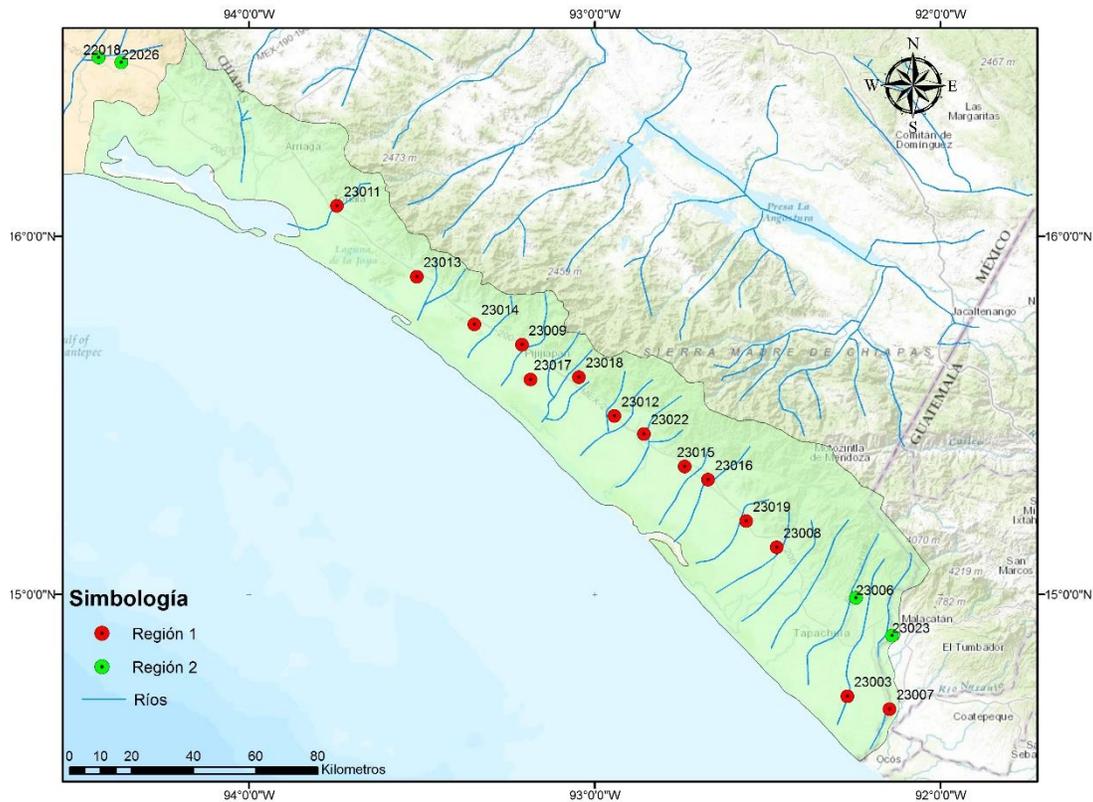


Figura 8. Regiones homogéneas para la RH-23.

Al aplicar la prueba F de Fisher para la comprobación de los grupos homogéneos, los resultados indican que la región 1 es la que puede ser fácilmente reproducible sus estadísticos al generar una muestra sintética. En tanto, la región 2 no pasa dicha prueba, por lo tanto, puede haber cierta incertidumbre en las muestras sintéticas que se llegasen a generar a partir de ella.

5.5 Análisis estadístico regional y factores regionales

Para efectuar el análisis estadístico regional, se aplicó la técnica regional *Estaciones-Año*. Como primer paso se modularon los datos de cada estación, dividiendo la media de cada uno entre cada uno de los valores registrados.

Una vez que se tienen los datos modulados se forma un único registro extendido formado por los datos modulados de cada estación, por ejemplo, para región homogénea número uno de la Región hidrológica 18 (Rh-18) se tendría un registro extendido de 164 valores modulados. Cabe mencionar que el análisis estadístico regional se hace únicamente con las estaciones de cada una de las regiones homogéneas delimitadas anteriormente.

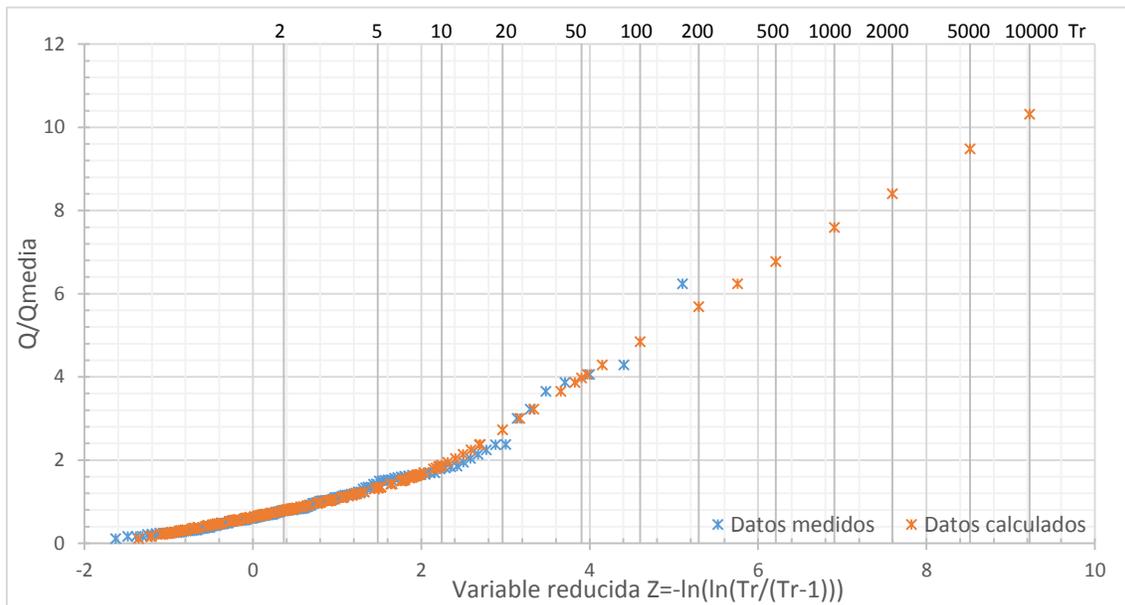
Posteriormente, se ajustó el registro modulado a una función de distribución de probabilidades usando el software AX.exe desarrollado por el Centro Nacional de

Prevención de Desastres (CENAPRED). Del análisis se encontró que la función Doble Gumbel fue la que mejor se ajustó a los datos modulados de las regiones homogéneas en estudio.

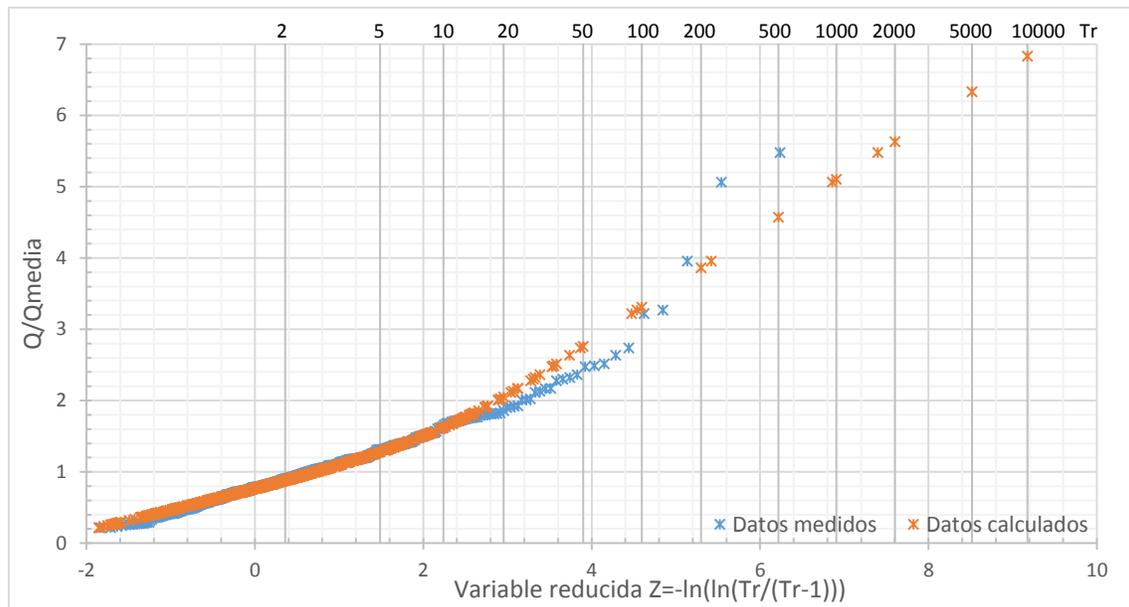
Con la función de distribución Doble Gumbel se obtuvieron factores regionales asociados a distintos periodos de retorno, estos factores regionales ayudan a obtener gastos de diseño al ser multiplicados por la media del gasto máximo instantáneo anual de un sitio que se encuentre dentro de la misma región homogénea.

Los resultados de los ajustes se muestran a continuación.

Región hidrológica 18.



Gráfica 13. Ajuste regional con la distribución Doble Gumbel de la región homogénea 1.



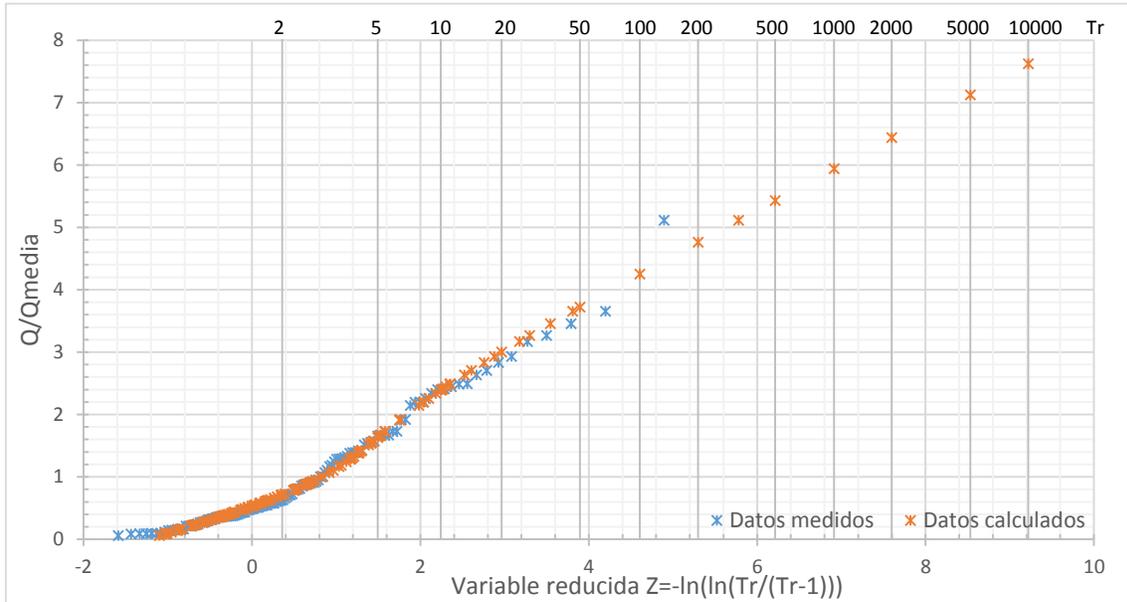
Gráfica 14. Ajuste regional con la distribución Doble Gumbel de la región homogénea 2.

El ajuste regional para región homogénea 1 es mejor si se compara con el de la región 2, esto se debe a valores extremos ocurridos en las estaciones 18129 y 18432, que corresponde a corrientes de primer orden y por tanto susceptibles a picos de gastos extremos (Ver Figura 6). Generalmente, los altos gastos son registrados cuando no existe gasto base suficiente que regule el tránsito a lo largo del cauce, otros factores como la pendiente, tiempo de concentración y número de curva N suelen ser importantes.

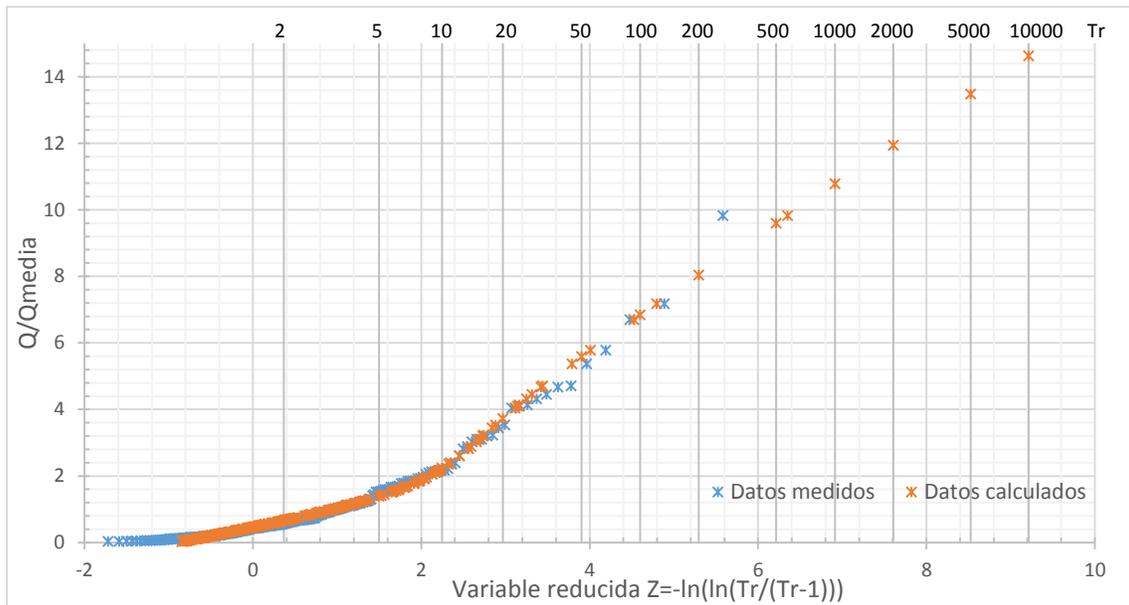
Aquí es importante señalar que los picos registrados en la región homogénea 2 afecta a un grupo de estaciones que teóricamente comparten características similares. La importancia de un análisis regional radica en poder hacer una extrapolación de esos eventos a estaciones donde no fueron registrados.

Un análisis puntual en estaciones vecinas nos llevaría a subestimación o sobreestimación, según sea el caso, de cualquier evento de diseño superior a 20 años de periodo de retorno. Se recalca que el propósito de este estudio es estimar avenidas de diseño a partir de modelos regionales con la menor incertidumbre posible al ampliar los registros mediante la técnica de estaciones-año para luego extrapolarlos.

Región hidrológica 22.



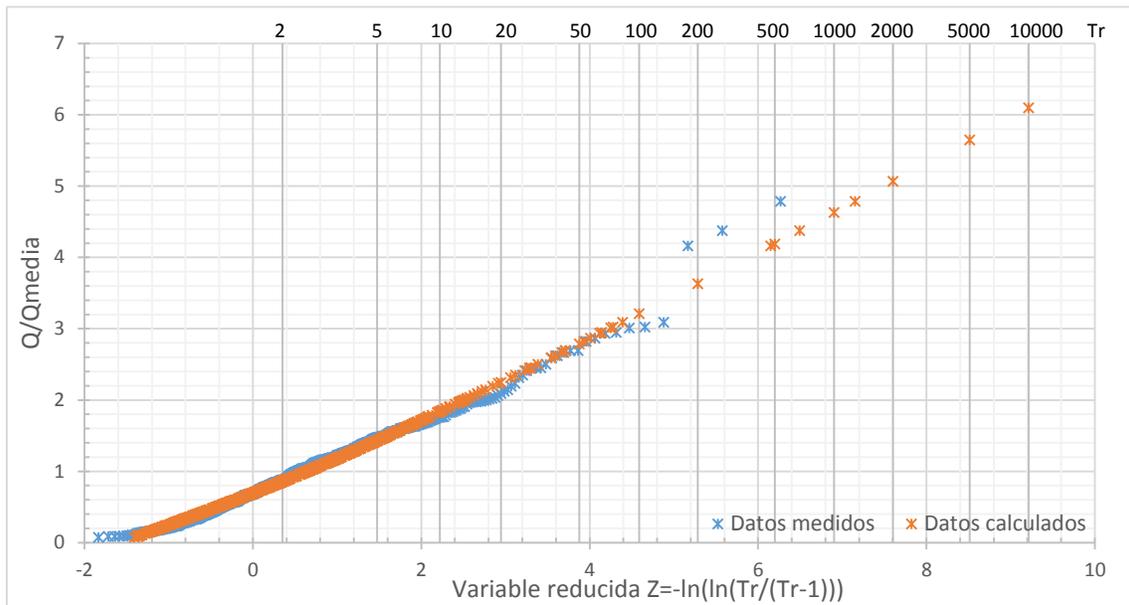
Gráfica 15. Ajuste regional con la distribución Doble Gumbel de la región homogénea 1.



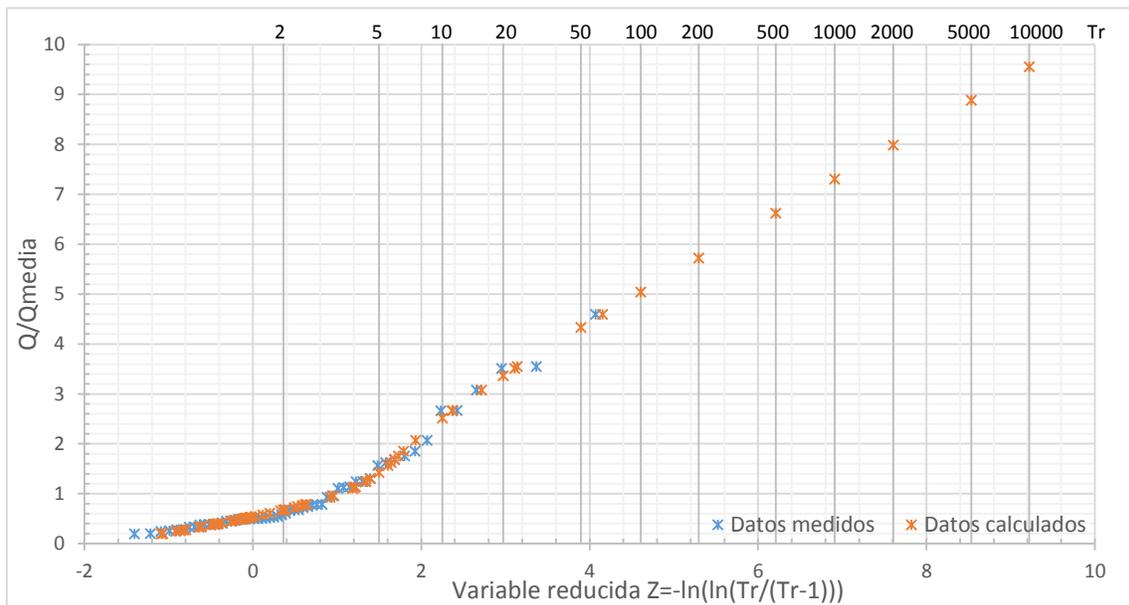
Gráfica 16. Ajuste regional con la distribución Doble Gumbel de la región homogénea 2.

En este caso el ajuste de la distribución Doble Gumbel con los datos medidos modulados es muy bueno, por lo tanto la incertidumbre es baja al momento de estimar gastos para periodos de retorno de hasta 100 años periodo de retorno.

Región hidrológica 23.



Gráfica 17. Ajuste regional con la distribución Doble Gumbel de la región homogénea 1.



Gráfica 18. Ajuste regional con la distribución Doble Gumbel de la región homogénea 2.

El ajuste para la región homogénea 1 no es tan bueno debido a tres eventos máximos que provocan un salto respecto al comportamiento de los demás valores. Cabe recordar que la región hidrológica 23, ubicada en la costa de Chiapas, es vulnerable al impacto de fenómenos meteorológicos. Los eventos extremos que se aprecian en la Gráfica 17 corresponden a eventos ciclónicos de los años 1973 y 1963 (Hernández Unzón, 2012; Rosengaus Moshinsky, Jiménez Espinosa, &

Vázquez Conde, 2014). Estos últimos no fueron registrados en estaciones vecinas, y debido a esto el ajuste para esos datos no es bueno.

A continuación se presentan en la Tabla 16 los factores regionales, para cada una de las tres regiones hidrológicas en estudio, obtenidos del análisis regional.

Tr	RH - 18		RH - 22		RH - 23	
	Región 1	Región 2	Región 1	Región 2	Región 1	Región 2
2	0.79	0.88	0.71	0.66	0.86	0.67
5	1.33	1.28	1.63	1.39	1.43	1.42
10	1.88	1.62	2.39	2.23	1.84	2.51
20	2.73	2.05	3.00	3.73	2.25	3.36
50	3.98	2.76	3.72	5.59	2.79	4.33
100	4.85	3.31	4.25	6.84	3.21	5.04
200	5.69	3.86	4.76	8.04	3.63	5.72
500	6.78	4.57	5.43	9.60	4.19	6.62
1000	7.60	5.10	5.94	10.78	4.63	7.30
2000	8.41	5.63	6.44	11.94	5.07	7.98
5000	9.49	6.33	7.12	13.48	5.65	8.88
10000	10.32	6.83	7.62	14.63	6.10	9.55

Tabla 16. Factores regionales obtenidos por el método de estaciones - año.

Con estos factores se pueden estimar gastos de diseño para distintos periodos de retorno en cuencas aforadas con escasa información. El proceso para la estimación de dichos gastos de diseño es el siguiente:

1. Ubicar la estación en estudio.
2. Identificar a que región homogénea pertenece.
3. Calcular la media de los gastos máximos instantáneos anuales.
4. Multiplicar la media con su respectivo factor regional obtenido de la Tabla 16.
5. Obtener los gastos de diseño para cada periodo de retorno requerido.

Cabe recordar que todo el análisis realizado en el presente trabajo gira en torno a la media de los gastos máximos instantáneos anuales, es por eso que se emplea este último para obtener los gastos de diseño a partir de factores regionales.

5.6 Ecuaciones regionales

Cuando en una zona, donde se requiera realizar un estudio hidrológico para estimar avenidas de diseño, no se cuenten con datos de aforo o los registros sean muy escasos el uso de ecuaciones regionales puede ser una buena opción.

Las ecuaciones regionales tiene la particularidad de relacionar la media del gasto máximo instantáneo anual, con las características fisiográficas y de precipitación de

las cuencas que involucren. Estas ecuaciones se formularon en primera instancia para cada una de las regiones homogéneas, sin embargo, en los casos donde las regiones homogéneas tenían pocas estaciones el ajuste era muy bueno pero los coeficientes de los exponentes determinados no eran congruentes.

Con la premisa de buscar una función potencial que involucre el mayor número de estaciones y que, además, tenga un buen ajuste, se planteó usar todas las estaciones para cada región hidrológica, de esta manera contar con más datos que den un mayor peso al ajuste.

Se concibieron dos formas de plantear las ecuaciones potenciales, la primera fue ponderando el área y la precipitación media.

$$Q_{estim} = a_1 A^{a_2} H p^{a_3} K^{a_4} \quad (5.1)$$

Donde A = área drenada, $H p$ = la altura de precipitación media y K = alguna otra característica de la cuenca, por ejemplo, tiempo de concentración (t_c), pendiente media (S_m), número de curva (N_c) y longitud del cauce principal (L_c).

El segundo empleando el volumen drenado que está en función del área drenada y la precipitación media. Donde K puede ser cualquier otra característica de la cuenca, exceptuando al área y la precipitación media, ya que se encuentran implícitas en el volumen escurrido (V).

$$Q_{estim} = a_1 V^{a_2} K^{a_3} \quad (5.2)$$

Los ajustes se realizaron empleando la herramienta Solver de Excel® y Algoritmos Genéticos en Matlab®, fijando como función objetivo aquella que minimizara el error cuadrático. A continuación se presentan los siguientes resultados.

Región hidrológica 18

Modelo	Solver			Algoritmo genético			Coeficientes (Solver)			
	Error cuadrático	R ²	R	Error cuadrático	R ²	R	a1	a2	a3	a4
a_1A^{a2}	1022.04	0.85	0.92	1022.06	0.85	0.92	1.32E+01	0.45		
$a_1A^{a2}Hp^{a3}$	616.10	0.95	0.97	681.60	0.93	0.97	1.68E-07	0.62	4.01	
$a_1A^{a2}Hp^{a3}Sm^{a4}$	477.37	0.97	0.98	2907.06	0.93	0.97	6.56E-07	1.02	3.85	0.98
$a_1A^{a2}Hp^{a3}tc^{a4}$	616.21	0.95	0.97	2709.72	0.94	0.97	1.50E-02	0.61	3.98	3.74E-03
$a_1A^{a2}Hp^{a3}Nc^{a4}$	509.05	0.96	0.98	3129.34	0.96	0.98	1.50E-02	0.66	4.55	-3.29
$a_1A^{a2}Hp^{a3}Lc^{a4}$	498.18	0.97	0.98	868.38	0.96	0.98	4.72E-08	-0.20	4.11	1.55
$a_1A^{a2}tc^{a3}$	1021.80	0.85	0.92	1021.80	0.85	0.92	1.60E+01	0.39	0.10	
$a_1A^{a2}tc^{a3}Sm^{a4}$	722.41	0.93	0.96	1091.54	0.90	0.95	1.96E+05	-0.45	3.66	2.91
$a_1A^{a2}tc^{a3}Nc^{a4}$	1019.61	0.85	0.92	1244.28	0.85	0.92	1.80E+02	0.41	0.08	-0.59
$a_1A^{a2}tc^{a3}Lc^{a4}$	722.54	0.93	0.96	809.30	0.91	0.95	2.70E-04	-0.43	-3.87	5.75
$a_1A^{a2}Sm^{a3}$	921.49	0.88	0.94	1022.04	0.85	0.92	3.28E+01	0.93	1.19	
$a_1A^{a2}Sm^{a3}Nc^{a4}$	921.49	0.88	0.94	947.30	0.88	0.94	3.29E+01	0.93	1.19	-1.68E-03
$a_1A^{a2}Sm^{a3}Lc^{a4}$	722.41	0.93	0.96	815.51	0.92	0.96	9.50E+00	-0.45	1.50	2.82

Tabla 17. Ecuaciones potenciales de la RH-18.

De la Tabla 17 se puede observar que el modelo potencial con un mayor coeficiente de determinación es el que involucra el área de la cuenca (A), altura de precipitación media (Hp) y la pendiente media (Sm) y además presenta el menor error cuadrático, por lo tanto, ésta será la propuesta para la región hidrológica 18.

Región hidrológica 22

Modelo	SOLVER			Algoritmo genético			Coeficientes (Solver)			
	Error cuadrático	R ²	R	Error cuadrático	R ²	R	a1	a2	a3	a4
a_1A^{a2}	364.04	0.02	0.14	405.31	0.03	-0.17	218.54	0.05		
$a_1A^{a2}Hp^{a3}$	277.03	0.43	0.66	279.38	0.44	0.67	3.73E-05	0.67	2.46	
$a_1A^{a2}Hp^{a3}Sm^{a4}$	222.69	0.63	0.80	229.31	0.63	0.79	3.97E-05	0.86	2.59	0.55
$a_1A^{a2}Hp^{a3}tc^{a4}$	259.49	0.50	0.71	552.90	0.16	0.40	2.33E+8	0.85	2.31	-0.48
$a_1A^{a2}Hp^{a3}Nc^{a4}$	186.29	0.74	0.86	869.05	0.40	0.63	2.33E+8	1.22	5.02	-10.37
$a_1A^{a2}Hp^{a3}Lc^{a4}$	276.98	0.43	0.66	93055.89	0.42	0.65	4.38E-05	0.68	2.44	-0.03
$a_1A^{a2}tc^{a3}$	317.05	0.26	0.51	317.05	0.26	0.51	110.55	0.39	-0.85	
$a_1A^{a2}tc^{a3}Sm^{a4}$	286.43	0.39	0.63	292.34	0.38	0.62	1866.69	0.14	0.42	0.99
$a_1A^{a2}tc^{a3}Nc^{a4}$	307.89	0.30	0.55	342.45	0.29	0.54	1.28E-04	0.45	-0.87	3.07
$a_1A^{a2}tc^{a3}Lc^{a4}$	286.43	0.39	0.63	308.92	0.32	0.56	1.73	0.14	-2.15	1.98
$a_1A^{a2}Sm^{a3}$	289.39	0.38	0.62	289.76	0.38	0.62	811.23	0.24	0.72	
$a_1A^{a2}Sm^{a3}Nc^{a4}$	270.39	0.46	0.68	319.81	0.44	0.66	3.80E-06	0.32	0.79	4.35
$a_1A^{a2}Sm^{a3}Lc^{a4}$	286.43	0.39	0.63	116393.94	0.05	0.23	590.68	0.14	0.83	0.33

Tabla 18. Ecuaciones potenciales de la RH-22.

Revisando los resultados de los modelos de la región hidrológica 22, se aprecia que los coeficientes de determinación son bajos. Las ecuaciones planteadas le dan poco peso al área como se puede observar en la ecuación potencial, sin embargo, al agregar más variables los ajustes mejoran.

De las medias de los gastos máximos instantáneos anuales (MQMIA) se observa que todas las estaciones tienen un valor que ronda los 300 m³/s, que se mantiene constante para todas las estaciones excepto en la estación 22028. Además, para ese mismo gasto el área varía en demasía de una estación a otra, por lo tanto, la ecuación potencial que solo involucra al área no puede presentar un mejor ajuste debido a que, al parecer, no guarda relación alguna con los escurrimientos medidos. Esto puede deberse a la propia orografía de la región, donde las estaciones con mayor área están protegidas por cadenas montañosas y tiene un tránsito lento de sus avenidas, a diferencia del resto de estaciones donde sucede todo lo contrario.

Por otra parte, se obtuvo que la ecuación potencial que mejor se ajusta a los datos medidos de la RH-22 es la que contiene las variables del área, altura de precipitación media, y Número de curva, sin embargo al revisar el signo de los exponentes, se encontró que para el Nc es negativo cosa que no es congruente. Al tener un Nc negativo implicaría que al tener un mayor valor de esta variable el gasto disminuiría, algo que no es posible pues el significado de Nc implica todo lo contrario. Expuestas las razones anteriores, la ecuación que mejor se ajusta con exponentes de signo congruentes es la que relaciona el área, la precipitación media y la pendiente media, con un coeficiente de correlación de 0.8.

Región hidrológica 23

MODELO	SOLVER			Algoritmo Genético			Coeficientes (Solver)			
	Error c.	R ²	R	Error c.	R ²	R	a1	a2	a3	a4
$a1A^{a2}$	333.67	0.85	0.92	533.05	0.69	0.83	1.73	0.89		
$a1A^{a2}Hp^{a3}$	304.60	0.87	0.93	304.60	0.87	0.93	0.15	0.80	0.58	
$a1A^{a2}Hp^{a3}Sm^{a4}$	302.33	0.88	0.94	404.49	0.88	0.94	0.99	1.07	0.46	0.99
$a1A^{a2}Hp^{a3}tc^{a4}$	304.29	0.87	0.93	372.82	0.78	0.88	0.02	0.84	0.56	-0.09
$a1A^{a2}Hp^{a3}Nc^{a4}$	726.38	0.13	0.36	736.38	0.13	0.35	0.02	0.30	0.83	0.89
$a1A^{a2}Hp^{a3}Lc^{a4}$	303.76	0.87	0.93	345.90	0.82	0.90	0.19	0.86	0.56	-0.12
$a1A^{a2}tc^{a3}$	327.36	0.86	0.93	327.36	0.86	0.93	0.95	1.10	-0.44	
$a1A^{a2}tc^{a3}Sm^{a4}$	316.06	0.87	0.93	328.76	0.89	0.94	50.20	1.27	0.30	2.09
$a1A^{a2}tc^{a3}Nc^{a4}$	279.74	0.90	0.95	297.58	0.92	0.96	17019.78	0.89	-0.35	-2.04
$a1A^{a2}tc^{a3}Lc^{a4}$	295.68	0.87	0.93	307.93	0.88	0.94	280.58	1.02	2.06	-2.24
$a1A^{a2}Sm^{a3}$	317.46	0.87	0.93	317.85	0.87	0.93	16.71	1.29	1.61	
$a1A^{a2}Sm^{a3}Nc^{a4}$	239.88	0.91	0.96	367.22	0.93	0.96	54114.62	0.56	-0.62	-2.39
$a1A^{a2}Sm^{a3}Lc^{a4}$	316.06	0.87	0.93	318.52	0.86	0.93	22.59	1.27	1.99	0.23

Tabla 19. Ecuaciones potenciales de la RH-23.

Para la región hidrológica 23 se encontró, una vez descartando las ecuaciones con exponentes no coherentes, que la mejor ecuación potencial es la que tiene por variables a la altura de precipitación media y el área de la cuenca. De esta forma, se observa que añadiendo la pendiente media a dicha ecuación existe una mejora en el coeficiente de determinación. Sin embargo, la mejora no es muy significativa y, partiendo del principio de parsimonia, se optó por usar el modelo potencial menos complejo, es decir, con menos variables.

Cabe mencionar que también se probaron ecuaciones potenciales que involucraban al volúmenes escurridos, no obstante, estas no presentaron un buen ajuste por lo tanto fueron descartados y no se presentan en este trabajo.

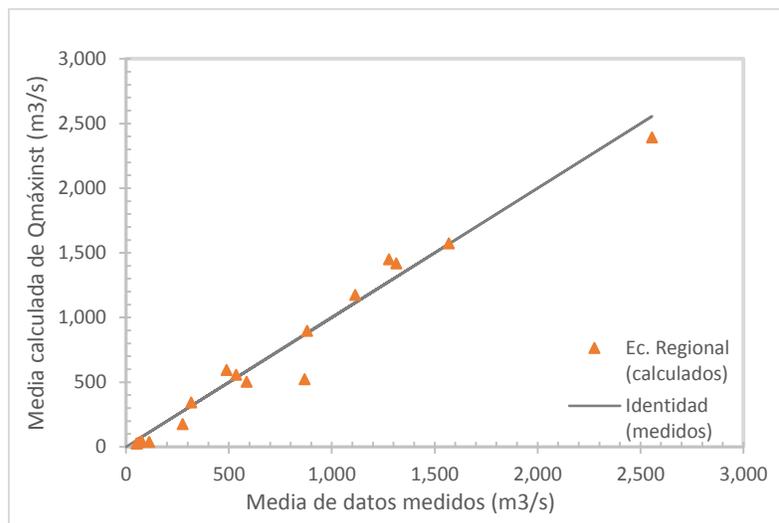
En resumen, las ecuaciones regionales que servirán para estimar la *media de los gastos máximos instantáneos anuales* con ayuda de las características fisiográficas se muestran en la Tabla 20. Finalmente, con el gasto calculado se pueden determinar los gastos de diseño con ayuda de los factores regionales obtenidos anteriormente y que se muestran en la Tabla 16.

Región hidrológica	Ecuación potencial
RH - 18	$\bar{Q} = 6.5629 \times 10^{-7} A^{1.0226} Hp^{3.8502} Sm^{0.9799}$
RH - 22	$\bar{Q} = 3.9715 \times 10^{-5} A^{0.8586} Hp^{2.5941} Sm^{0.5478}$
RH - 23	$\bar{Q} = 0.1504 A^{0.7986} Hp^{0.5840}$

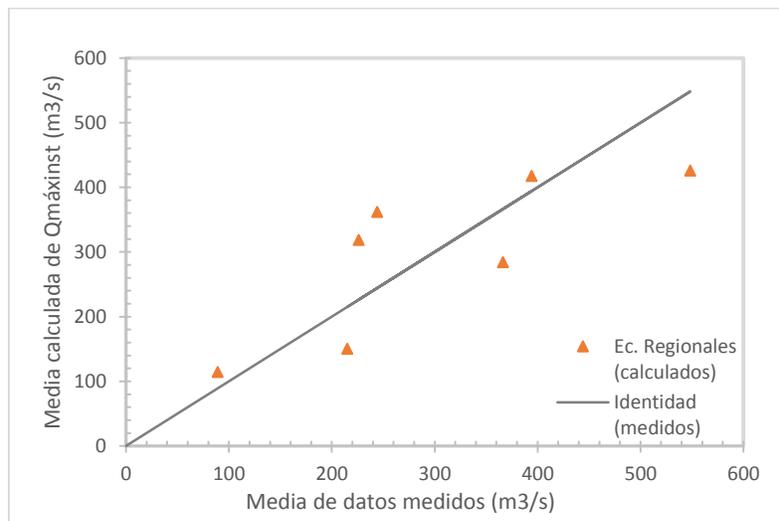
Unidades: \bar{Q} en $\frac{m^3}{s}$, A en Km^2 , Hp en mm y Sm en $\frac{m}{m}$

Tabla 20. Ecuaciones regionales para estimar la MQMIA.

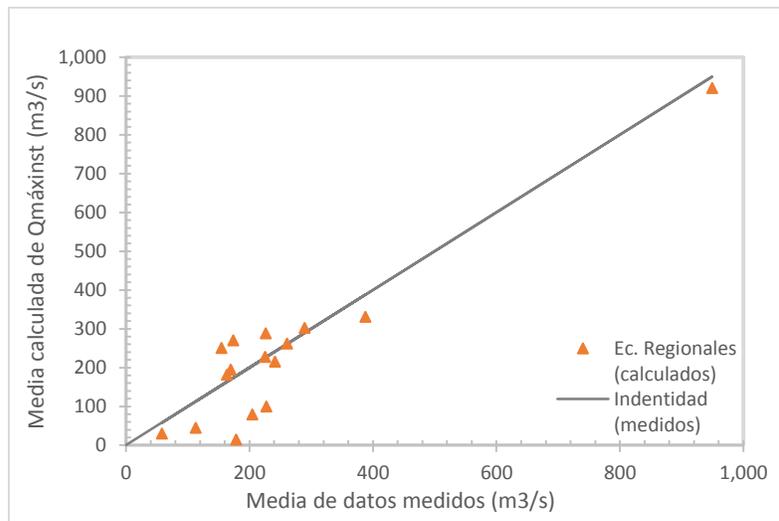
En las gráficas siguientes se muestran la comparación de las MQMIA medidos con los calculados por las ecuaciones regionales de la Tabla 20.



Gráfica 19. Comparación de MQMIA medidos y calculados de la RH-18.



Gráfica 20. Comparación de MQMIA medidos y calculados de la RH-22.



Gráfica 21. Comparación de MQMIA medidos y calculados de la RH-23.

5.7 Generación de muestras sintéticas

Los registros hidrológicos, en general, se describen a partir de las propiedades estadísticas de un registro original tales como la media, desviación estándar, coeficiente de asimetría, coeficiente de curtosis y distribución de probabilidad (Escalante Sandoval & Reyes Chavéz, 2002).

En el presente trabajo se expone la generación de muestras, a partir de los parámetros de distribución de probabilidades, de esta forma empleando la función de distribución de probabilidades se pueden generar muestras que comúnmente se denominan sintéticas.

Como se ha explicado anteriormente una distribución de F de Fisher comprueba que ciertos registros formen parte de una misma población manteniendo, en teoría, el comportamiento de estadísticos al generar muestras sintéticas.

El proceso para la generación de muestras sintéticas consiste en generar números aleatorios con distribución uniforme y con valores entre 0 y 1. Cada valor de número aleatorio representa la probabilidad de ocurrencia del evento que queremos generar. Mediante la función de distribución que mejor se ajuste a la serie de datos original procedemos a calcular los eventos para cada probabilidad aleatoria.

Generalmente, las muestras sintéticas para un sitio son generadas de manera puntual, es decir, sin tomar en consideración estaciones el comportamiento hidrológico en otros sitios. En este trabajo se plantea la generación de muestras sintéticas a partir de la serie resultante de la técnica estaciones año, es decir, se generaran muestras sintéticas de tipo regional. Para las regiones en estudio se

empleó la función de distribución Doble Gumbel por ser la que mejor se ajusta a la serie regional, los parámetros se muestran en la Tabla 21.

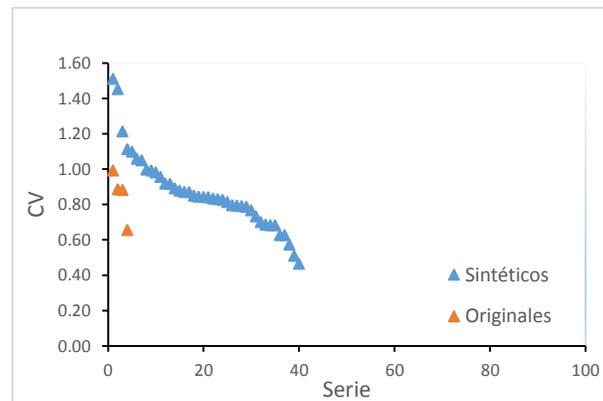
Se generaron un total de 10 muestras sintéticas por cada estación que abarcara el método de estaciones-año, para cada una de ellas se calculó el coeficiente de variación. Una vez que se calcularon los coeficientes de variación se procedió a compararlos con los coeficientes de variación originales. La comparación de Cv es importante para verificar que las muestras sintéticas reproducen eventos con el mismo comportamiento estadístico que las muestras originales.

	RH - 18		RH - 22		RH - 23	
	R1	R2	R1	R2	R1	R2
$\alpha_1 =$	2.7445	3.5059	2.7019	2.1150	2.4000	3.657
$\beta_1 =$	0.5990	0.7390	0.4350	0.4090	0.6490	0.467
$\alpha_2 =$	0.8540	1.3125	1.3745	0.5980	1.5401	1.028
$\beta_2 =$	2.2140	1.6750	2.0830	3.0760	1.5980	2.145
$P =$	0.90	0.91	0.80	0.90	0.90	0.80

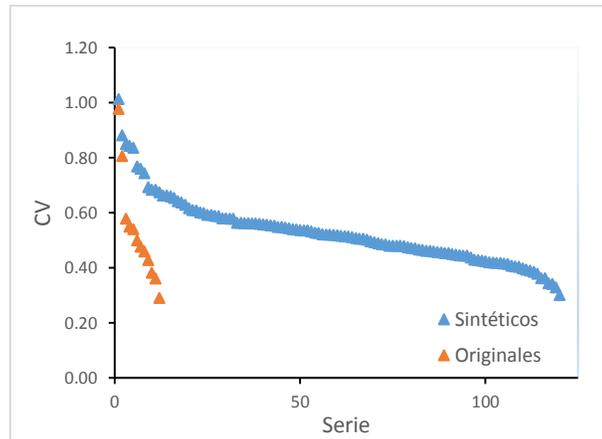
Tabla 21. Parámetros de la distribución Doble Gumbel.

Resultados de las muestras sintéticas.

Región hidrológica 18.



Gráfica 22. Comparación de Cv, región homogénea 1 (RH-18).

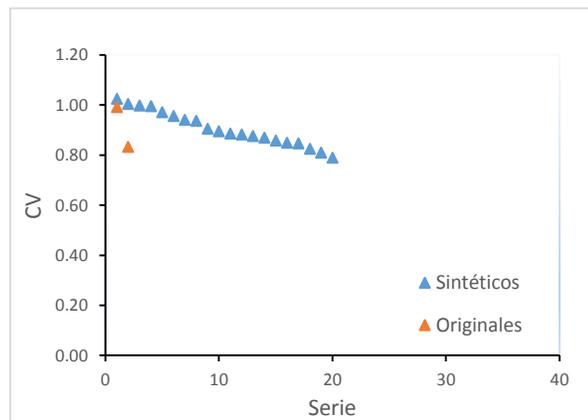


Gráfica 23. Comparación de Cv, región homogénea 2 (RH-18).

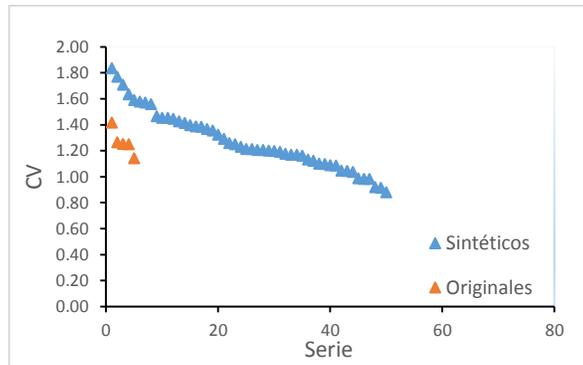
Al graficar los coeficientes de variación de las muestras sintéticas se observa que estas varían con cierto margen de holgura respecto de los originales, por lo que se puede concluir que las muestras sintéticas si conservan los estadísticos de las muestras originales.

Únicamente en la región homogénea dos parece estar en un margen muy ajustado, sin embargo, el margen es pequeño y por lo tanto se considera que si se reproducen los estadísticos de la muestra original.

Regio hidrológica 22.



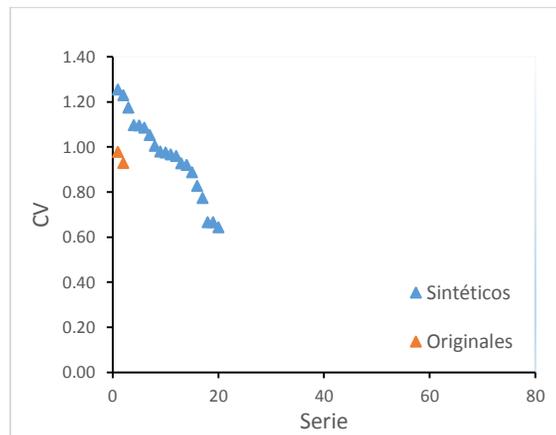
Gráfica 24. Comparación de Cv, región homogénea 1 (RH-22).



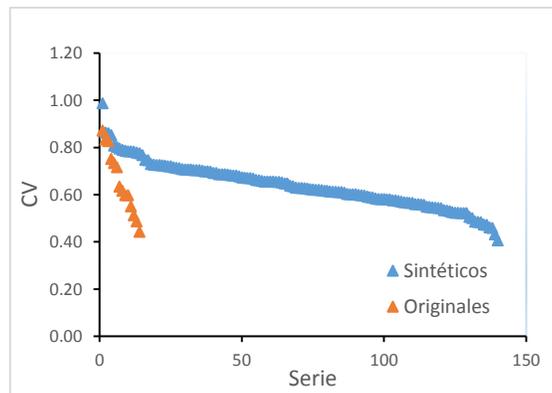
Gráfica 25. Comparación de Cv, región homogénea 2 (RH-22).

En el caso de la región hidrológica 22 se observa que los coeficientes de variación sintéticos si reproducen a los originales, es decir que se obtuvieron valores en un rango que envuelve al Cv de las muestras originales.

Región hidrológica 23.



Gráfica 26. Comparación de Cv, región homogénea 1 (RH-23).



Gráfica 27. Comparación de Cv, región homogénea 2 (RH-23).

En cuanto a esta región, se tiene un comportamiento similar a la región 18, ya que en una zona homogénea se logran reproducir los valores de Cv holgadamente, mientras que en la segunda zona no es así, sin embargo, el margen es muy pequeño y por lo tanto es aceptable.

Las muestras sintéticas, presentadas en este trabajo sirvieron para comprobar que las series regionales, además de reproducir gastos para distintos periodos de retorno, también pueden generar series sintéticas que conserven sus variables estadísticas, como lo es el coeficiente de variación. La generación de series sintéticas es muy importante cuando se realizan estudios de aprovechamientos hidráulicos, estudios de factibilidad hídrica o manejo integral de cuencas.

6. Conclusiones y recomendaciones

Los métodos regionales son muy útiles, ya que estos procedimientos facilitan la estimación de escurrimientos con ayuda de zonas que si cuentan con información hidrométrica.

La realización de este análisis regional trajo como resultados la obtención de factores regionales por medio del método de estaciones año, que pueden ser usados para estimar gastos de diseño en cuencas con poca información hidrométrica, únicamente conociendo la media del gasto medio instantáneo anual (MQMIA) de la zona de estudio.

Por otra parte, en el caso de que no se cuenten con datos en el sitio donde se requiera determinar gastos de diseño, se proponen ecuaciones del tipo potencial que involucran variables fácilmente medibles. Cabe mencionar que las ecuaciones potenciales tienen como límite de aplicabilidad el área máxima y mínima de las cuencas consideradas en la delimitación de las regiones homogéneas. La elegibilidad está en función de los coeficientes de determinación, correlación y error cuadrático, además de la verificación del exponente de cada una de las ecuaciones ya que hubo casos donde se obtuvieron exponentes no congruentes con las variables que se estaban empleando.

La revisión de los exponentes en las ecuaciones potenciales es algo que se debió hacer a detalle, por ejemplo en la Región 18 se obtuvieron valores de exponente negativo para el número de curva (N_c) lo cual no es congruente ya que a mayor número de curva existe un incremento en el gasto por lo que un exponente negativo significaría lo contrario. Algo similar ocurre con el exponente del tiempo de concentración ya que debe poseer un exponente negativo, por ejemplo, al referirnos a un hidrograma nos daremos cuenta que para un determinado volumen escurrido y para un tiempo de concentración relativamente pequeño, tendremos un gasto pico más elevado mientras que para un tiempo mayor se tendrá un gasto pico menor.

Finalmente, se debe tener especial cuidado al momento de estimar avenidas de diseño de más de 100 años de periodo de retorno, pues a partir de dicho periodo el ajuste no es muy bueno por lo que puede llevarse a una subestimación. Se recomienda optar por revisar los resultados de las pruebas de homogeneidad para verificar que efectivamente los registros no sufren una variación abrupta, ya sea con una tendencia a la alza o a la baja. Si se comprueba que no existe homogeneidad se deberá tomar con reserva periodos de retorno mayores. En general, para periodos de retorno pequeños los ajustes son muy buenos y por ende podemos decir que la estimación de los gastos adquiere una menor incertidumbre.

Bibliografía

- Allasia, D., & Avruch Goldenfun, J. (2000). Regionalización de caudales en la Región del Alto Uruguay-Río grande do sul Brasil. *Universidad Nacional del Nordeste Comunicaciones Científicas y Tecnológicas*.
- Andrews, D. (1972). Plots of high-dimensional data. *Biometrics*, 28, 125-136.
- Aparicio, F. (2010). *Fundamentos de hidrología de superficie*. México: Limusa.
- AX.exe. (1997). CENAPRED, Ciudad de México.
- BANCO NACIONAL DE DATOS DE AGUAS SUPERFICIALES (BANDAS). (Agosto de 2015). Obtenido de <http://www.conagua.gob.mx/CONAGUA07/Contenido/Documentos/Portada%20BANDAS.htm>
- Byzedi, M., Saghafian, B., Mohammadi, K., & Siosemarde, M. (2013). Regional analysis of streamflow drought: a case study in southwestern Iran. *Environmental Earth Sciences*.
- CONAGUA, C. N. (Mayo de 2016). *Sistema Nacional de Información del Agua, Atlas digital del agua México*. Obtenido de <http://www.conagua.gob.mx/atlas/ciclo10.html>
- Dalrymple, T. (1960). Flood-Frequency analyses. Manual of hidrology: Part 3. *Geological Survey Water-Supply*(1513-A).
- Domínguez Mora, R., Arganis Juárez, M., Guzmán García, H., Carriosa Elizondo, E., Esquivel Garduño, G., & Fuentes Mariles, G. (2016). *Modelos regionales de escurrimientos máximos instantáneos en la República Mexicana*. México: Instituto de Ingeniería, UNAM.
- Domínguez, M., Arganis, J., Capella, V., Luna, V., Fuentes, M., Carriosa, E., . . . Reyes, M. (2012). *Análisis Estadístico. Manual de Diseño de Obras Civiles, Comisión Federal de Electricidad. Hidrotécnica, Hidrología, (A.1.6)*.
- Domínguez, M., Arganis, J., Gónzales, V., Carriosa, E., & Guzmán, G. (2013). *Regionalización de datos de escurrimientos de las regiones hidrológicas de México para la estimación de avenidas de diseño*. México: Instituto de Ingeniería, UNAM.
- Domínguez, M., Carriosa, E., Fuentes, M., & Galván, T. (2010). *Actualización de mapas de precipitaciones para diferentes periodos de retorno y diferentes duraciones*. México: Elaborado por el Centro Nacional de Prevención de Desastres.
- Domínguez, M., Esquivel, G., Baldemar, M., Méndez, A., Mendoza, R., Arganis, J., & Carriosa, E. (2008). *Modelo de Pronóstico de Escurrimiento MPE*. México: Instituto de Ingeniería.
- Erazo Ch., A. M. (2004). *Regionalización de caudales máximos y medios*. San Salvador: Servicio Hidrológico Nacional de El Salvador.
- Escalante Sandoval, C., & Reyes Chavéz, L. (2002). *Técnicas Estadísticas en Hidrología*. México, D.F.: Facultad de Ingeniería UNAM.
- Hernández Unzón, A. (2012). *Base de datos de ciclones tropicales que afectaron a México durante el período de 1970 a 2011*. Comisión Nacional del Agua.
- Isik, S., & Singh, V. P. (2008). Hydrologic Regionalization of Watersheds in Turkey. *JOURNAL OF HYDROLOGIC ENGINEERING © ASCE*, 824-834.

- Nathan, R., & McMahon, T. (1990). Identification of homogeneous regions for the purposes of regionalisation. *Journal of Hydrology*, 121, 217-238.
- Rosengaus Moshinsky, M., Jiménez Espinosa, M., & Vázquez Conde, M. (2014). *Atlas climatológico de ciclones tropicales en México*. México: Centro Nacional de Prevención de Desastres.
- Secretaría de recursos hidráulicos. (1969). *Boletín hidrológico num. 37*. México.
- Secretaría de recursos hidráulicos. (1971). *Boletín hidrológico num. 49. Tomo I*. México.
- Secretaría de recursos hidráulicos. (1974). *Boletín hidrológico num. 31. Tomo V*. México.
- Sutcliffe, J. (1978). *Methods of flood estimation: A guide to the flood studies report*. England: Institute Hydrology.