## APENDICE 6.1.

#### **Metodología detallada del Paso 1. “Fundamento Hidrológico”**

#### Análisis de consistencia, homogeneidad y complementación de series hidrometeorológicas.

##### Complementación de series hidrometeorológicas

Con respecto a las estaciones hidrométricas el rellenado de datos se realiza tomando un periodo común de todas las estaciones y se eliminan los huecos que se tuvieran, al determinar una la matriz de coeficiente de correlación entre las estaciones. Se eligió la relación más alta y se graficaron ambas series normalizadas, una vez teniendo el gráfico se agregó la línea de tendencia con mayor R2 y esa ecuación se aplicó para rellenar la estación con menor número de datos (y) empleando los datos de la otra estación (x). Así se hizo sucesivamente hasta tener cada una de las estaciones completas

##### Análisis de independencia

Para que se pueda llevar a cabo el análisis de frecuencias se requiere que la muestra ***x****,* **para *j* = *1, 2, 3,…. n***esté compuesta por las variables aleatorias. En hidrología, independencia significa que un evento que ocurre en un tiempo dado no depende de eventos anteriores.

Una vez rellenados las series de tiempo de la información hidrométrica, se procede a verificar independencia en cada serie 𝑗 mediante la prueba de Anderson la independencia de eventos.

###### Prueba de Anderson

La prueba independencia de Anderson hace uso del coeficiente de correlación serial **ϒk** para distintos tiempos de retraso de ***´k´****.*

La expresión para el coeficiente de correlación serial de retraso ***´k´***es:

**ϒk =**

Ecuación 6.2.1.1

Para  **= K = 1, 2,… =**

Además, los límites a 95 % de confianza para **ϒk** se obtienen como:

**ϒk (95 %) =**

Ecuación 6.2.1.2

La gráfica de los valores estimados para ´**ϒk´** (ordenadas) contra los tiempos de retraso ´***k****´* (abcisas) junto con sus correspondientes límites de confianza, se llama correlograma de la muestra, Fig. 6.2.1.1.

Si solo el 10 % de los valores de ´**ϒk´** sobrepasan los límites de confianza se dice que la serie es independiente y: es una variable que sigue las leyes de la probabilidad.



Fig. 6.2.1.1. Precipitación media anual en la Cuenca Hidrológica Río La Antigua.

##### Análisis de homogeneidad

Las características estadísticas de las series hidrológicas, como la media, desviación estándar y los coeficientes de correlación serial, se afectan cuando las series presentan tendencia en la media o en la varianza; o cuando ocurren saltos negativos o positivos, tales anomalías son producidas por la pérdida de homogeneidad y la inconsistencia.

En general, la falta de homogeneidad de los datos es inducida por las actividades humanas como la deforestación, apertura de nuevas áreas al cultivo, rectificación de cauces, construcción de embalses, reforestación, entre otras. También es producto de los procesos naturales súbitos, como incendios forestales, huracanes, terremotos, deslizamiento de laderas, erupciones volcánicas anomalías climáticas que evidencian la ocurrencia del fenómeno de El Niño y otros fenómenos macroclimáticos.

La naturaleza homogénea de los procesos hidrológicos puede examinarse estadísticamente con respecto al tiempo y al espacio. Las pruebas estadísticas que miden la homogeneidad de una serie de datos presentan una hipótesis nula y una regla para aceptarla o rechazarla. Las pruebas de homogeneidad utilizadas y las características que analizan se describen a continuación:

* Prueba estadística de Helrmert, analiza los cambios de las desviaciones estándar con respecto a su media de toda la serie.
* Prueba estadística de “t” de Student, analiza los cambios abruptos de la media dividiendo a la serie en dos bloques.
* Regla de Cramer, analiza los cambios de la media, diferenciando tres bloques de la serie histórica.

El análisis de homogeneidad se realiza para evaluar la calidad de la información hidrológica, no con la finalidad de rechazarla sino de seleccionar los registros que representen mejor las condiciones hidrológicas. No se realizó ningún tratamiento de las series, debido a que no se tienen los criterios suficientes para corregirlas, es difícil detectar el origen de estos cambios.

Los resultados de las pruebas estadísticas proporcionan información sobre el estado de las series de precipitación y caudal, sin embargo, se toman con precaución en su interpretación, ya que ciertos comportamientos hidrológicos no son observados cuando las series son muy cortas o incompletas, y pueden ser detectados por las pruebas estadísticas como un cambio o una tendencia.

A continuación, se explican las pruebas de homogeneidad utilizadas.

###### Prueba estadística de Helmert

Esta prueba es sencilla y consiste en analizar el signo de las desviaciones de cada evento en estudio con respecto a su media. Si una desviación de un cierto signo es seguida de otra del mismo signo, entonces se dice que se forma una secuencia (s), de lo contrario se considera como un cambio (c)

La serie se considera homogénea si cumple:

Donde “n” es el tamaño de la muestra.

###### Prueba estadística de “t” de Student

Cuando la causa probable de la pérdida de homogeneidad de la serie sea un cambio abrupto en la media, la prueba del estadístico *´t´* es muy útil.

Si se considera una serie de tamaño *´n´* la cual se divide en dos conjuntos de tamaños

Entonces, el estadístico de prueba se define como:

Ecuación 6.2.1.3

Donde

 Media y varianza de la segunda parte del registro de tamaño *´n´.*

El valor absoluto de ´*td´* se compara con el valor de la distribución *´t´* de student de dos colas, y con **Ѵ = + – 2** grados de libertad para un nivel de significancia de **ɚ = 0.005.** Los valores de t se muestran en la Tabla 6.2.1.1:

Tabla 6.2.1.1. Valores "t" de Student.

| Grados de libertad | Una Cola 5% | Dos Colas 5% |
| --- | --- | --- |
| 1 | 6.314 | 12.706 |
| 2 | 2.920 | 4.303 |
| 3 | 2.353 | 3.182 |
| 4 | 2.132 | 2.776 |
| 5 | 2.015 | 2.571 |
| 6 | 1.943 | 2.447 |
| 7 | 1.895 | 2.365 |
| 8 | 1.860 | 2.306 |
| 9 | 1.833 | 2.262 |
| 10 | 1.812 | 2.228 |
| 11 | 1.796 | 2.201 |
| 12 | 1.782 | 2.179 |
| 13 | 1.771 | 2.160 |
| 14 | 1.761 | 2.145 |
| 15 | 1.753 | 2.131 |
| 16 | 1.746 | 2.120 |
| 17 | 1.740 | 2.110 |
| 18 | 1.734 | 2.101 |
| 19 | 1.729 | 2.093 |
| 20 | 1.725 | 2.086 |
| 21 | 1.721 | 2.080 |
| 22 | 1.717 | 2.074 |
| 23 | 1.714 | 2.069 |
| 24 | 1.711 | 2.064 |
| 25 | 1.718 | 2.060 |
| 26 | 1.706 | 2.056 |
| 27 | 1.703 | 2.052 |
| 28 | 1.701 | 2.048 |
| 29 | 1.699 | 2.045 |
| 30 | 1.697 | 2.040 |
| 40 | 1.684 | 2.021 |
| 60 | 1.671 | 2.000 |
| 120 | 1.658 | 1.980 |
| ∞ | 1.645 | 1.960 |

###### Regla de Cramer

Esta prueba se utiliza con el propósito de verificar la homogeneidad en el registro en estudio, y también para determinar si el valor medio no varía significativamente de un periodo de tiempo corto. Con este propósito se consideran 3 bloques, el primero, del tamaño total de la muestra n; el segundo de tamaño n60 (60% de los últimos valores de la muestra n); y el tercero de tamaño n30 (30% de los últimos valores de la muestra n).

La prueba compara el valor de **ẍ** del registro total con cada una de las medias de los bloques elegidos Para que se considere la serie analizada como estacionaria en la media, se deberá cumplir que no exista diferencia significativa entre las medias de los bloques.

**ẍ =**

Ecuación 6.2.1.4

Ecuación 6.2.1.5

Ecuación 6.2.1.6

 **;**

Ecuación 6.2.1.7

*tw* = w = 60 y 30

Ecuación 6.2.1.8

El estadístico ***´tw´***tiene distribución ***´t´*** de Student de dos colas con una **Ѵ = + – 2** grados de libertad y para un nivel de significancia **ɚ = 0.005.**

Si y sólo si el valor absoluto de *´tw´* para **w = 60 y 30** es mayor que el de la distribución **´t´** de Student se concluye que la diferencia entre las medias es evidencia de inconsistencia y: la serie ***´x´****,* se considera no homogénea.

En caso de que no sean homogéneas se disminuye un T calculado a utilizar por fines de seguridad. En el caso de mínimos es para asegurar el mínimo requerimiento del ecosistema haciendo mayor la magnitud propuesta.

#### Técnicas de estimación de parámetros

Las técnicas de estimación de parámetros se utilizan en cada una de las Funciones de Distribución de Probabilidades (FDP), para mínimos utilizadas para la definición del Periodo de retorno con el que se define la región homogénea.

El análisis de frecuencias de eventos extremos, consiste en obtener una distribución de probabilidad que mejor describe el comportamiento de la variable aleatoria.

Estas distribuciones de probabilidad están definidas por el valor que toman sus parámetros. Estos parámetros pueden “estimarse” por diferentes técnicas las cuales cumplan las siguientes propiedades:

a) Sesgo nulo: Un estimador de un parámetro se dice que es no sesgado, si el valor esperado esto es, el sesgo.

Sesgo=

Ecuación 6.2.1.9

b) Consistencia: Un estimador de un parámetro se dice es consistente si la probabilidad que difiere de por más de una constante arbitraria , se aproxima a cero a medida que el tamaño de la muestra tiende al infinito.

c) Eficiencia: Un estimador se dice que es un estimador más eficiente de si no es sesgado y si su varianza es por lo menos menor que la de cualquier otro estimador de .

d) Suficiencia: Un estimador se dice que es suficiente para , si usa la información de relevancia para que está contenida en la muestra.

###### Estimación por momentos

El método de los momentos es un procedimiento muy sencillo para encontrar un estimador de uno o más parámetros poblacionales. Consiste básicamente en plantear un sistema de ecuaciones, cuyo tamaño depende del número de parámetros a estimar. Esto se hace al igualar los momentos poblacionales con los muestrales.

Los momentos muestrales, también conocidos como estadísticos muestrales, se obtienen con las siguientes expresiones.

Media

Ecuación 6.2.1.10

Varianza

Sesgada:

Ecuación 6.2.1.11

No sesgada:

Ecuación 6.2.1.12

Cabe indicar que la denominación “no sesgado” hace referencia al hecho de considerarse que en la serie restan n-1 valores independientes luego de haberse calculado la media de la muestra (al compararse los valores de la serie con su propia media se pierde un grado de libertad).

Coeficiente de asimetría

Sesgado:

Ecuación 6.2.1.13

No sesgado:

Ecuación 6.2.1.14

Coeficiente de curtosis

Sesgado:

Ecuación 6.2.1.15

No sesgado:

Ecuación 6.2.1.16

Coeficiente de variación

Ecuación 6.2.1.17

En el análisis hidrológico se recomienda el uso de los estadísticos no sesgados, ya que generalmente se trabaja con muestras relativamente pequeñas. Cuando n es relativamente grande (al menos mayor de 30), los valores de los momentos muestrales sesgados y no sesgados son prácticamente similares.

###### Método de máxima verosimilitud

Sea una función de densidad f(x) de “n” parámetros, si existe una muestra aleatoria (θ) de esta función de densidad, entonces se puede encontrar una función de densidad conjunta que se puede escribir de la siguiente manera.



Ecuación 6.2.1.18

Esta es la función de verosimilitud, el método estima los parámetros maximizando L, es decir maximizando la probabilidad de que ocurran conjuntamente las variables aleatorias, en otras palabras asegurando que la muestra es la única que puede obtenerse al seleccionar n observaciones aleatorias a partir de la función de densidad. La función logarítmica está definida por:



Ecuación 6.2.1.19

De igual manera esto lo realizamos diferenciando la función parcialmente respecto de sus parámetros e igualando a cero, obteniendo de esta manera un sistema de ecuaciones que al resolverlo nos da como resultado los parámetros de la función.

###### Estimación por momentos L

Estos momentos son capaces de caracterizar a un mayor número de distribuciones, además de estar virtualmente libres de sesgo aún para muestras pequeñas. Las muestras deben estar ordenadas de mayor a menor.

 Ecuación 6.2.1.20

  Ecuación 6.2.1.21

 Ecuación 6.2.1.22

 Ecuación 6.2.1.23

Donde:



Coeficiente de variación – L



Ecuación 6.2.1.24

Coeficiente de sesgo – L



Ecuación 6.2.1.25

Coeficiente de curtosis – L



Ecuación 6.2.1.26

#### Distribuciones de probabilidad. Teoría distribucional

Se obtienen los estadísticos muéstrales de las series : para i= 1, 2,...,n años y para j=1, 2, 3, …, s sitios, en la región a analizar.

Para cada serie 𝑗 se genera una muestra modulada de la forma:

Ecuación 6.2.1.27

El análisis de frecuencia de los caudales de una muestra 𝑄𝑖, con i =1,2,..., n, se emplea para proveer la magnitud de un evento 𝑄𝑇 de cierto período de retorno T, por medio del ajuste de una distribución de probabilidad, la cual se selecciona como la mejor de un grupo de ellas por el menor EEA tenga. Específicamente, se calcula los periodos de retorno de 2 y 10 años.

Las distribuciones de probabilidad utilizadas para el análisis de mínimos son las siguientes:

* Distribución Weibull de 2 parámetros, recomendado por Manciola *et al*., 2004.
* Distribución Log-Normal de 2 parámetros, recomendado por Ashkar, J. et al., 2009.
* Distribución Log-Normal de 3 parámetros, , recomendado por Osorio-González M. *et al*., 2014.
* Distribución Gamma de 2 parámetros, recomendado por Ashkar, J. et al., 2009.
* Distribución Gamma de 3 parámetros, recomendado por Osorio-González M. *et al*., 2014.
* Distribución Gumbel, recomendado por Huang, W. *et al.*, 2014.
* General de Valores Extremos, recomendado por Huang, W. *et al.*, 2014.

Para cada una de las series determinadas para las estaciones seleccionadas, se hizo lo siguiente:

1. Se ordenaron los gastos 7Qmin de cada una de las series de menor a mayor.

2. Se asignó a cada serie una distribución de probabilidad empírica, siguiendo la ley de Weibull.

3. Para cada una de las series se determinaron los parámetros de ajuste de las distribuciones de probabilidad.

###### Distribución Weibull

 Ecuación 6.2.1.28

 Ecuación 6.2.1.29

Donde y k son los parámetros de escala y forma.

###### Distribución Log Normal con dos parámetros:

Ecuación 6.2.1.30

 Parámetro de ubicación

 Parámetro de escala

γ > 0

Estimadores por momentos y máxima verosimilitud:

Ecuación 6.2.1.31

Ecuación 6.2.1.32

Evento de diseño:

Ecuación 6.2.1.33

###### Distribución lognormal con 3 parámetros

Ecuación 6.2.1.34

Donde:

x0 Parámetro de ubicación

µy Parámetro de forma

σy Parámetro de escala

Estimadores por momentos:

Ecuación 6.2.1.35

Donde:

Ecuación 6.2.1.36

Ecuación 6.2.1.37

Ecuación 6.2.1.38

Ecuación 6.2.1.39

Ecuación 6.2.1.40

Estimadores por máxima verosimilitud:

Ecuación 6.2.1.41

Ecuación 6.2.1.42

Y el estimador se obtiene al resolver:

Ecuación 6.2.1.43

Evento de diseño:

Ecuación 6.2.1.44

###### Distribución gamma con 2 parámetros

Ecuación 6.2.1.45

Ecuación 6.2.1.46

Si

Donde:

α Parámetro de escala

β Parámetro de forma

 Función Gamma completa

Además:

Ecuación 6.2.1.47

Ecuación 6.2.1.48

Ecuación 6.2.1.49

Estimadores por momentos:

Ecuación 6.2.1.50

Ecuación 6.2.1.51

Estimadores por máxima verosimilitud:

Ecuación 6.2.1.52

El estimador β se obtiene al resolver:

Ecuación 6.2.1.53

Empleando como valor inicial de parámetro :

Ecuación 6.2.1.54

Ecuación 6.2.1.55

La derivada de la función es:

Ecuación 6.2.1.56

La aproximación de Thon de la función digamma de es:

Ecuación 6.2.1.57

La aproximación de Thon de la función trigamma de es:

Ecuación 6.2.1.58

Evento de diseño:

Ecuación 6.2.1.59

###### Distribución gamma con 3 parámetros

Ecuación 6.2.1.60

Donde:

x0 Parámetro de ubicación

α Parámetro de escala

β Parámetro de forma

Además:

Ecuación 6.2.1.61

Ecuación 6.2.1.62

Ecuación 6.2.1.63

Ecuación 6.2.1.64

Estimadores por momentos:

Ecuación 6.2.1.65

Ecuación 6.2.1.66

Ecuación 6.2.1.67

Estimadores por máxima verosimilitud:

Ecuación 6.2.1.68

Ecuación 6.2.1.69

El estimador se obtiene al resolver:

Ecuación 6.2.1.70

Evento de diseño:

Ecuación 6.2.1.71

###### Distribución Gumbel

Ecuación 6.2.1.72

Donde:

 Parámetro de ubicación

 α Parámetro de escala

Estimadores por momentos:

Ecuación 6.2.1.73

Ecuación 6.2.1.74

Estimadores por máxima verosimilitud

La variable reducida Gumbel es:

Ecuación 6.2.1.75

Considerando la variable reducida Gumbel se tiene el siguiente proceso iterativo:

Ecuación 6.2.1.76

Ecuación 6.2.1.77

El criterio de convergencia es:

Incrementos:

Ecuación 6.2.1.78

Ecuación 6.2.1.79

Nuevos valores:

Ecuación 6.2.1.80

Ecuación 6.2.1.81

Eventos de diseño:

Ecuación 6.2.1.82

###### Distribución log pearson tipo 3

Ecuación 6.2.1.83

Donde:

y0 Parámetro de ubicación

α Parámetro de escala

β Parámetro de forma

Estimadores por momentos:

Ecuación 6.2.1.84

Ecuación 6.2.1.85

Ecuación 6.2.1.86

Donde son los estadísticos de la serie:

Estimadores por máxima verosimilitud:

Ecuación 6.2.1.87

Ecuación 6.2.1.88

El estimador se obtiene al resolver:

Ecuación 6.2.1.89

Evento de diseño:

Ecuación 6.2.1.90

###### Distribución General de Valores Extremos (GVE)

Ecuación 6.2.1.91

Ecuación 6.2.1.92

Si ** = 0, , entonces la distribución es tipo I (Gumbel)

Si ** < 0, , entonces la distribución es tipo II (Fréchet)

Si ** > 0, , entonces la distribución es tipo III (Weibull)

Donde *υ* es el parámetro de ubicación, ** es el parámetro de escala, ** > 0, y ** es el parámetro de forma.

La variable reducida GVE es:

Ecuación 6.2.1.93

Estimadores por momentos

Para

Ecuación 6.2.1.94

Para

Ecuación 6.2.1.95

Ecuación 6.2.1.96

Ecuación 6.2.1.97

Ecuación 6.2.1.98

Ecuación 6.2.1.99

Ecuación 6.2.1.100

Para (Distribución tipo II)

Ecuación 6.2.1.101

Ecuación 6.2.1.102

Para (Distribución tipo III)

Ecuación 6.2.1.103

Ecuación 6.2.1.104

Para (Distribución tipo I)

Ecuación 6.2.1.105

Ecuación 6.2.1.106

Estimadores por máxima verosimilitud

Se tiene el siguiente proceso iterativo.

Ecuación 6.2.1.107

Ecuación 6.2.1.108

Ecuación 6.2.1.109

El criterio de convergencia es:

Ecuación 6.2.1.110

Ecuación 6.2.1.111

Ecuación 6.2.1.112

Incrementos:

Ecuación 6.2.1.113

Ecuación 6.2.1.114

Ecuación 6.2.1.115

Ecuación 6.2.1.116

Ecuación 6.2.1.117

Ecuación 6.2.1.118

Ecuación 6.2.1.119

Ecuación 6.2.1.120

Ecuación 6.2.1.121

Nuevos valores

Ecuación 6.2.1.122

Ecuación 6.2.1.123

Ecuación 6.2.1.124

Eventos de diseño

Ecuación 6.2.1.125

#### Método de selección de distribuciones

La calidad de los valores de caudal estimados para un cierto período de retorno, con distribuciones de probabilidad teórica, está dado principalmente por la comparación de dichos valores estimados con los valores realmente observados o medidos. Para ello es posible utilizar diferentes técnicas denominadas Técnicas de Evaluación o Métodos de Bondad de Ajuste. Entre los diferentes métodos más difundidos se encuentran los de:

• Chi Cuadrado

• Kolmogorov – Smirnov

• Papeles probabilísticos

• Error Estándar de Ajuste

En el presente trabajo sólo se utilizará la Técnica del Error Estándar de Ajuste, se obtiene con la ecuación:

Ecuación 6.2.1.126

Donde:

 Son los eventos ordenados de mayor a menor con un período de retorno asignado:

 y una probabilidad de no excedencia 𝑃=1−1/T

𝑛𝑗: Longitud en años del registro analizado

 : Número de orden del registro.

: Eventos estimados por cierta distribución de probabilidad para cada período de retorno T asignado a la muestra ordenada .

𝑚𝑝: Número de parámetros de la distribución ajustada.

La distribución de mejor ajuste será aquella que proporcione el mínimo valor del estadístico EEA. Si una o más distribuciones tienen valores similares de EEA, entonces se deberá optar por aquella distribución que tenga el menor número de parámetros.

#### Análisis de Precipitación

El proceso que se describe a continuación se realiza debido a que se utilizan parámetros de precipitación tal y como lo propone Reydi (2012) y Orsini (2011), en la aplicación de técnicas regionales para cuencas aforadas y no aforadas.

##### Análisis previo de las series de datos de precipitación

Con información consultada en la Base de Datos Climatológica Nacional (CLICOM) del Servicio Meteorológico Nacional se determinó que en la cuenca existen 94 estaciones climatológicas, cuyos registros comprenden diferentes periodos de tiempo. Además, las series están incompletas por uno o varios días, inclusive por años.

Para poder utilizar esta información y posteriormente realizar el rellenado de datos faltantes se realizaron pruebas de homogeneidad (Helmert, T de Student y Cramer) e independencia (método de Anderson) para descartar estaciones que no cuenten con datos independientes o no homogéneos.

Inicialmente se obtuvieron los datos anuales de la precipitación para cada una de las estaciones con años completos de datos. Posterior a ello y con ayuda del software Análisis de Frecuencias (AFA v.1.1) creado por el Instituto Mexicano de Tecnología del Agua (IMTA) en el 2010 se realizaron dichas pruebas.

Se muestra el proceso para la estación 30008 Altotonga, Veracruz en el Anexo 6.1.1. Se basa en el proceso mostrado en el “Manual de análisis de frecuencias en hidrología” (Aparicio *et al.*, 2010) para el uso del software AFA.

##### Generación de datos de precipitación faltantes

Con el análisis previo de las series de datos no fue descartado el uso de ninguna estación climatológica debido a que todas las muestras resultaron ser independientes con la aplicación del método de Anderson.

Para poder hacer uso de dicha información es necesario que todas las series cuenten con el mismo periodo de registros, además el periodo debe de ser continuo. Con esta situación es necesaria la aplicación de un método para el rellenado de datos faltantes para cumplir con estos parámetros.

Dentro de los métodos de interpolación mediante técnicas geoestadísticas se seleccionó el método de peso del inverso de la distancia (IDW) presentado por Real, 2015. Dicho método se muestra a continuación:

Ecuación 6.2.1.127

Donde:

Ecuación 6.2.1.128

Donde:

Para aplicar este método se seleccionaron las 5 estaciones más cercanas a la estación del dato a interpolar, teniendo en cuenta que por lo menos 3 de ellas contaran con el registro del valor faltante. Además, con ayuda de SIG se obtuvieron las distancias entre las estaciones.

Finalmente, con el Software de creación de hojas de cálculo, se programaron las operaciones necesarias con las fórmulas del método IDW para realizar la generación de los valores de los registros faltantes.

Observando el registro de las series de datos de todas las estaciones se determinó trabajar con el periodo comprendido de 1960-2009 debido a que la mayoría de las estaciones mostraba registros en ese periodo. De las 94 estaciones existentes, se eliminaron 30 debido a que en el periodo de análisis no contaban con un registro continuo de datos y las estaciones circundantes a ellas también presentaban esa falta de datos, por lo que no se realizó la generación datos faltantes de precipitación para dichas estaciones.

##### Pruebas de homogeneidad e independencia

Terminada la generación de los datos faltantes en cada una de las 64 estaciones en el periodo 1960-2009 se realizó nuevamente una serie de pruebas estadísticas de homogeneidad e independencia para verificar que no existan problemas en el registro de la información, además permitieron corroborar que la generación de los valores faltantes de las series climatológicas no introduzca mayor incertidumbre a los registros.

Para la independencia se aplicó la prueba de Anderson y en homogeneidad se aplicaron las pruebas de Helmert, T de Student y Cramer. De manera similar al análisis previo se usó AFA v.1.1.

El proceso que se llevó a cabo es idéntico al mostrado en el análisis previo, la única variante fue el periodo, ahora de 50 años homogéneo para todas las estaciones.

##### Parámetros de lluvia

Concluido el rellenado de datos y con de software de creación de hojas de cálculo se determinaron las características de “Días con precipitación apreciable” y “Precipitación antecedente” para cada una de las 64 estaciones climatológicas determinadas en la sección anterior, se procedió a generar rasters de cada una de estas variables.

Una vez calculada la precipitación promedio mensual, con datos diarios, se procedió a interpolar para poder lograr un resultado por área de estudio. La técnica de interpolación que se utilizó fue Kriging bayesiano empírico (EBK). Una vez teniendo los rasters de la media diaria. Se usó álgebra de mapas con SIG para determinar el valor medio de X cantidad de días sumados consecutivamente. Posteriormente, se adjudicó el valor medio a las poligonales representativas de las áreas drenadas por las estaciones hidrométricas independientes ya analizadas para cada región homogénea. Se trabajó de la misma forma el parámetro de “Días con precipitación apreciable”, se interpoló con el mismo método los valores promedios de las estaciones meteorológicas creando un raster y adjudicando el valor medio a las áreas drenadas de las estaciones depuradas dependiendo la región homogénea en el área de estudio.

#### Delimitación de Regiones homogéneas

##### Delimitación Geográfica

En esta técnica se tiene que ubicar el centroide (estación base) de la cuenca que se va a estudiar y a partir de este se traza un círculo de 80 km de radio, donde las estaciones hidrométricas que se encuentran dentro de esta circunferencia son una misma región. Todas las estaciones ubicadas dentro de esta circunferencia se considera que presentan el mismo patrón meteorológico y que cualquier predicción de tipo hidrológico será válida en la región delimitada (Escalante, [sin fecha]; Cruz-Miguel, 2012).



Fig. 6.2.1.2.Ejemplo de aplicación de técnica geográfica de delimitación de regiones homogéneas.

Fuente: Cruz-Miguel, 2012

##### Método región influencia

En esta técnica cada sitio se considera el centro de su propia región. La identificación de una región de influencia se basa en la medición de la distancia Euclidiana en un espacio de atributos multidimensionales. Se propone esta técnica para definir las regiones homogéneas de las cuencas no aforadas.

El conjunto de atributos o características fisiográficas y climatológicas se relaciona a las características de los eventos extremos. También se define una función de peso que refleja la importancia relativa de cada sitio dentro de la región.

La distancia Euclidiana entre una estación cualquiera ***j***  y la estación base ***b*** para lo cual se genera la región de influencia, se expresa como:



Ecuación 6.2.1.129





El proceso estandarizado elimina las unidades de cada atributo y reduce cualquier diferencia de escala dentro del rango de valores. El grupo de atributos seleccionado es en función de la información disponible en la red de estaciones. La elección de los atributos requiere cierto juicio ingenieril. Sin embargo, pueden obtenerse como la técnica de delimitación multidimensional al examinar la correlación entre las características fisiográficas y las climatológicas.

##### Método de coeficiente de variación

Esta técnica hace uso de una comparación gráfica entre atributos estadísticos de los registros analizados.

Para cada una de las series j, se obtienen los estadísticos muestrales: media y coeficiente de variación (Cvj). Posteriormente se construye una gráfica situando en el eje de las ordenadas los coeficientes Cvj y en las abscisas las relaciones donde Aj son las áreas drenadas de cada uno de los sitios j. Los grupos se forman por la concentración de puntos alrededor de determinado valor del coeficiente de variación. Es importante mencionar que estos estadísticos deben estimarse de registros con distribución Normal.



Fig. 6.2.1.8. Esquema de la aplicación de la técnica de CV.

De las muestras recortadas, periodo común, se obtienen los estadísticos, donde:

A= Área de la cuenca



Ecuación 6.2.1.130

Donde:

S= Desviación estándar muestral

Xmed= media de la muestra

##### Parámetro β de la Distribución General de Valores Extremos (GVE)

Está técnica se basa en la FDP de GVE. El parámetro β es el parámetro que determina el tipo de distribución de valores extremos al que converge una serie de datos. Por esto, tener varias series de datos con valores similares del parámetro β es otro criterio para identificar una región homogénea.



Ecuación 6.2.1.131

Donde representa el parámetro forma, el cual puede tener valores negativos, positivos o iguales a cero, de lo cual se puede dar una primera clasificación y agrupación. Para estimar una región homogénea las muestras deben comenzar y terminar en un mismo periodo y poder estimar sus parámetros. Es común obtener valores para el parámetro entre -0.25 y 0.25.

##### Trazos multidimensionales. Curvas de Andrews

La heterogeneidad de los grupos que se forman de manera preliminar, se evalúa mediante una técnica de posicionamiento propuesta por Andrews en la cual un punto en el espacio multidimensional se representa por una curva de dos dimensiones de acuerdo con la función. Se propone esta técnica para definir las regiones homogéneas de las cuencas no aforadas.

Ecuación 6.2.1.132

Donde:

x1, x2, x3, x4, … son las características fisiográficas y/o meteorológicas de la cuenca obtenidas en el análisis de regresión y la función se evalúa en el rango

El hecho de que esta función preserva las distancias la hace una técnica ideal de comparación visual para la formación de los grupos homogéneos. Los racimos de cuencas con comportamientos similares aparecen como una banda de curvas muy próximas unas de otras.

Una característica del método que no se identifica de forma inmediata, consiste en que los resultados que se obtienen dependen del tipo y del ordenamiento de las variables seleccionadas, Las primeras variables son asociadas con componentes cíclicas de baja frecuencia y las últimas de alta frecuencia. Las bajas frecuencias son más fáciles de observar, de esta manera x1 representará aquella que en el análisis de frecuencias resulto más significativa desde el punto de vista estadístico, x2 a la segunda y así sucesivamente.

Una vez se han desarrollado las funciones del tipo *f(t)* para cada uno de los sitios involucrados, se generan envolventes que formen conjuntos de curvas de comportamiento similar y se deriva una curva media *(t)* que represente los atributos físicos de cada grupo formado.

#### Técnicas Regionales

Posteriormente se realizaron tres métodos de análisis regional para cuencas aforadas y no aforadas: el Método Estaciones-Año, el Método de Regresión Lineal Múltiple y Avenida índice para localizar un valor de caudal ecológico (7Q10). El parámetro de selección entre estos métodos es el mínimo EEA.

##### Método Estaciones-Año

Este método considera el tratamiento de una sola muestra de datos conformada por un registro estandarizado de eventos, el cual una vez que se construye se ajusta a un conjunto de distribuciones de probabilidad. Esta técnica regional a diferencia de otras no requiere que las muestras tengan una longitud de registro común. Con referencia al tamaño de muestra aceptable en cada muestra participante en el análisis, se recomienda que al menos sea de 10 años. Esta recomendación no es del todo estricta, ya que, si se tiene un sitio con, por ejemplo, 9 años, pero dentro de este están eventos extremos que puedan afectar el comportamiento del fenómeno en la región, entonces se deberán incluir en el estudio.

Existe otra recomendación en cuanto el número de estaciones-año que se requieren para una correcta estimación del evento , y es que donde 𝑇 es el periodo de retorno requerido en años.

El Método Estaciones-Año descrito por Escalante *et al.* (2008) se describe de la siguiente manera:

Su característica más relevante es el registro llamado “estaciones-año” de tamaño con todos los eventos modulados

Una vez obtenida la distribución de mejor ajuste, es posible estimar los eventos regionales para diferentes periodos de retorno, los cuales serían válidos para cualquier punto dentro de la zona analizada.

Ecuación 6.2.1.133

T = 2, 5, 10, 20, 50, 100, 500, 1000, 5000 y 10000 años

Si se quiere obtener un evento en un sitio j que tiene escasa información, basta multiplicar la relación regional obtenida en el paso anterior por el correspondiente valor de .

##### Avenida índice

Usar mínimo tres estaciones para este método. Se requiere un periodo común entre estaciones con 10 años mínimo.



Fig. 6.2.1.9. Mismo periodo de registro en todas las estaciones.

Si el máximo no está en el periodo común, se prefiere quitar una estación, si se incluye el máximo en las restantes



Fig. 6.2.1.10. Ejemplo de definición de registro a utilizar en técnica de avenida índice.

Se obtienen los eventos de diseño para todas las estaciones usando los parámetros de la distribución Gumbel que hayan dado el menor EEA y adicionalmente se obtiene el evento de diseño para 2.33 años.



Fig. 6.2.1.11. Obtención de Q con Tr de 2.33 años en cada estación.

Tr=2.33 años

Calcular μ y α por la técnica de momentos y máxima verosimilitud

Ecuación 6.2.1.134

Se modulan los gastos, se ordenan por bloques de periodo de retorno y se obtiene la mediana de cada bloque,

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Estación |  |  | … |  |
| A |  |  | … |  |
| B | . | . |  |  |
| C | . | . |  |  |
| D | . | . |  |  |
| E |  |  | … |  |

La mediana de cada bloque se considera la ecuación regional, para obtener los eventos de diseño de las demás estaciones sólo se multiplica la ecuación regional por el Q 2.33 de cada estación.

|  |  |  |  |
| --- | --- | --- | --- |
| A |  | … |  |
| B | . | . | . |
| C | . | . | . |
| D | . | . | . |
| E |  | … |  |

Se busca una relación funcional entre los q 2.33 y la característica fisiográfica más significativa. Con ello se establece la ecuación para el Qx, con el Tr requerido.

Gráfica 6.2.1.1. Ejemplo de definición de ecuación para Qx.

##### Método de correlación y regresión Lineal Múltiple

Las características en la frecuencia de los caudales mínimos anuales pueden expresarse en términos de algunos índices geomorfológicos y factores climatológicos de la cuenca. La dependencia de las variables se estima para cada sitio mediante un proceso de correlación y regresión discriminante, donde la aportación de cada variable independiente es probada hasta lograr un modelo de regresión en el cual todas las variables que en él intervengan sean estadísticamente significativas. (Escalante *et al*. 2008).

Ahora bien, este método se describe en lo siguiente:

Recabar la información sobre los eventos por ser analizados en cada una de las estaciones de la región, así como la determinación de sus características fisiográficas, partiendo de las propuestas por Reidy *et al.* (2012) y complementadas con Orsini (2011).

Para cada periodo de retorno analizado se plantea un sistema de ecuaciones:

Ecuación 6.2.1.135

Donde

Si se quiere obtener un evento en un sitio j no aforado o con escasa información, basta sustituir las características fisiográficas de j en la ecuación regional obtenida en el paso anterior.

Hasta este punto se cuenta con las diferentes regiones con mismo funcionamiento hidrológico, además de la metodología para definir el 7Q10 en cuencas no aforadas.

Debido que se utilizó características geomorfológicas para regionalizar, es decir, se hizo una caracterización que permitió clasificar diversas cuencas; se pretende partir de esta clasificación para detallar en el paso 2 del marco ELOHA.