

01162  
6  
2ej



**DIVISION DE ESTUDIOS DE POSGRADO**  
**Facultad de Ingeniería**

**MODELOS DE TRANSFERENCIA  
DE INFORMACION HIDROLOGICA**

**MARTIN ALFONSO GUTIERREZ LOPEZ**

**TESIS**

**PRESENTADA A LA DIVISION DE ESTUDIOS DE  
POSGRADO DE LA**

**FACULTAD DE INGENIERIA**

**DE LA**

**UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO**

**COMO REQUISITO PARA OBTENER  
EL GRADO DE**

**MAESTRO EN INGENIERIA**

**( HIDRAULICA )**

**CIUDAD UNIVERSITARIA**

**NOVIEMBRE, 1995**

**FALLA DE ORIGEN**



Universidad Nacional  
Autónoma de México



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

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

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

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

A



**ROCIO PAREDES LOPEZ**

*" Dios bendiga a los ausentes..."*

## INDICE

	PAGINA
<b>CAPITULO 1</b>	
<b>1. INTRODUCCION</b>	<b>1</b>
<b>CAPITULO 2</b>	
<b>2. IDENTIFICACION DE REGIONES HOMOGENEAS</b>	<b>4</b>
2.1 Características fisiográficas de la cuenca	4
2.1.1 Area y perímetro del parteaguas de la cuenca	5
2.1.2 Precipitación media anual	6
2.1.3 Pendiente media de la cuenca	7
2.1.4 Elevación media de la cuenca	9
2.1.5 Orden y longitud de la corriente principal	9
2.1.6 Pendiente de la corriente principal	10
2.1.7 Parámetro de forma	13
2.1.8 Frecuencia de la corriente	13
2.1.9 Densidad de drenaje	14
2.1.10 Ubicación del centroide de la cuenca	14
2.1.11 Coeficiente de escurrimiento	15
2.2 Regionalización de cuencas de acuerdo con sus características fisiográficas y climatológicas.	18
2.2.1 Trazo multidimensional	20
2.2.2 Ventajas y desventajas	25
2.3 Regionalización de cuencas de acuerdo con el coeficiente de variación de los gastos máximos anuales.	26
2.3.1 Técnicas de agrupación por racimos	26
2.4 Regionalización de cuencas con base en los gastos máximos anuales	31
2.5 Selección del análisis de grupo más apropiado	36

<b>CAPITULO 3</b>	
<b>3. ANALISIS DE VARIANZA</b>	<b>38</b>
3.1 Distribución F	38
3.1.1 Comparación de dos varianzas poblacionales	40
3.2 Noción general del análisis de varianza	40
3.3 Análisis de varianza en dos sentidos	43
<b>CAPITULO 4</b>	
<b>4. ANALISIS DE CORRELACION Y REGRESION</b>	<b>47</b>
4.1 Procesos de normalización	48
4.1.1 Transformación de Box-Cox	50
4.1.2 Transformación Box-Cox de doble potencia	53
4.2 Correlación simple	54
4.3 Correlación múltiple	58
4.3.1 Intervalos de confianza y pruebas de hipótesis	64
4.3.2 Criterio de selección	68
4.4 Extensión de registros	72
4.4.1 Criterios de mejora en la estimación de la media y varianza	74
<b>CAPITULO 5</b>	
<b>5. DISTRIBUCIONES DE PROBABILIDAD</b>	<b>79</b>
5.1 Distribuciones univariadas	79
5.1.1 Estimación de parámetros	80
5.1.2 Distribución Normal	84
5.1.3 Distribución Lognormal	85
5.1.4 Distribución General de Valores Extremos Tipo I (Gumbel)	85
5.1.5 Distribución General de Valores Extremos	86
5.1.6 Distribución Exponencial	86
5.1.7 Distribución Gamma	87
5.1.8 Distribución Gumbel Mixta	88
5.2 Distribuciones bivariadas	88

## **CAPITULO 6**

<b>6. APLICACIONES</b>	<b>93</b>
6.1 Descripción del área de estudio	93
6.2 Metodología	94
6.3 Normalización de muestras	100
6.4 Discriminación de variables	100
6.5 Identificación de regiones homogéneas	112
6.5.1 Criterio tradicional	112
6.5.2 Trazo multidimensional	112
6.5.3 Regiones homogéneas con base en los coeficientes de variación	115
6.5.4 Regiones homogéneas con base en la prueba de Langbein	137
6.6 Estimación en cuencas aforadas	143
6.6.1 Extensión de registros	148
6.7 Estimación en cuencas no aforadas	161
6.7.1 Ecuaciones regionales con base en parámetros estimados en forma univariada	163
6.7.2 Ecuaciones regionales con base en parámetros estimados en forma bivariada	173
6.7.3 Ecuaciones regionales con base en registros extendidos	173

## **CAPITULO 7**

<b>7. CONCLUSIONES Y RECOMENDACIONES</b>	<b>190</b>
7.1 Conclusiones	190
7.2 Recomendaciones	192

## **BIBLIOGRAFIA**

## **AGRADECIMIENTOS**

## **ANEXOS**

ANEXO 1 Programas en Quick Basic 4.5

ANEXO 2 Planos de la Región Hidrológica 10

## INDICE DE TABLAS

		PAGINA
2.1	Características fisiográficas empleadas en el estudio	5
2.2	Resumen de las características fisiográficas de la región hidrológica 10	17
3.1	Valores críticos de la distribución $F$ al nivel de significancia de 5%	45
3.2	Valores críticos de la distribución $F$ al nivel de significancia de 1%	46
4.1	Coefficientes de A, B y C para estimar la varianza y el coeficiente de correlación simple crítico de la varianza para $\theta = 1$	77
4.2	Coefficiente de correlación simple crítico de la varianza para $\theta = 0$ y $\theta = 1$	78
6.1	Principales estaciones de la región hidrológica 10	96
6.2	Gastos máximos instantáneos región hidrológica 10	97
6.3	Estadísticos de las muestras originales	99
6.4	Resumen de las características fisiográficas de la región hidrológica 10 (muestras normalizadas)	101
6.5	Estadísticos de las muestras normalizadas	102
6.6	Regresión de 15 características fisiográficas con gastos de 100 años de período de retorno	105
6.7	Regresión de 10 características fisiográficas con gastos de 100 años de período de retorno	108
6.8	Regresión de 6 características fisiográficas con gastos de 100 años de período de retorno	111
6.9	Distancias entre estaciones para círculos de 80 km	113
6.10	Análisis de varianza para los coeficientes de variación (1a. etapa)	130
6.11	Análisis de varianza para los coeficientes de variación (2a. etapa)	131
6.12	Análisis de varianza para los coeficientes de variación (3a. etapa)	133
6.13	Análisis de varianza para los coeficientes de variación (4a. etapa)	134
6.14	Análisis de varianza para los coeficientes de variación (etapa de calibración)	136
6.15	Prueba de homogeneidad de Langbein, ajuste a todas las estaciones	139
6.16	Prueba de homogeneidad de Langbein, ajuste a la región "A" (23 estaciones)	140

6.17	Prueba de homogeneidad de Langbein, ajuste a la región "B" (19 estaciones)	140
6.18	Mínimos errores estándar de ajuste utilizando distribuciones univariadas	149
6.19	Parámetros de la distribución Gumbel Mixta para las diferentes estaciones	150
6.20	Gastos de diseño para diferentes períodos de retorno ajustados con la distribución Gumbel Mixta, utilizando distribuciones univariadas	151
6.21	Parámetros de la distribución Gumbel Mixta, obtenidos por distribuciones bivariadas	152
6.22	Gastos máximos instantáneos, región "A" (23 estaciones) muestras normalizadas	156
6.23	Coefficientes de correlación para la región "A"	157
6.24	Período efectivo de registro para correlaciones >0.45	158
6.25	Gastos máximos instantáneos, región "A" (23 estaciones) registros extendidos por correlación lineal simple	159
6.26	Parámetros de la distribución Gumbel Mixta, obtenidos de las muestras extendidas	160
6.27	Características fisiográficas normalizadas con Box Cox de doble potencia	165
6.28	Parámetros de la distribución Gumbel Mixta, normalizados con Box Cox de doble potencia	166
6.29	Regresión de las características fisiográficas con el parámetro alfa 1	167
6.30	Parámetros de la distribución Gumbel Mixta, ajustados con la ecuación regional	181
6.31	Parámetros de la distribución Gumbel Mixta, obtenidos con la ecuación regional de parámetros bivariados	182
6.32	Parámetros de la distribución Gumbel Mixta, obtenidos con la ecuación regional de muestras extendidas	183



## INDICE DE FIGURAS

		PAGINA
1	Curvas de Andrews para 15 características fisiográficas	22
2	Curvas de Andrews para 6 características fisiográficas	23
3	Curvas de Andrews para las estaciones Huites y Piaxtla	24
4	Límites de confianza para la prueba de homogeneidad de Langbein	36
5	Distribución $F$ para diferentes grados de libertad	39
6	Transformación de Box-Cox	52
7	Curvas de Andrews para la región "A"	116
8	Curvas de Andrews para la región "B"	116
9	Curvas de Andrews para la región "A" y Choix	117
10	Curvas de Andrews para la región "B" y Choix	117
11	Curvas de Andrews para las regiones "A" y "B" de las estaciones: Jaina, Tecusiapa y Guamuchil	118
12	Curvas de Andrews para las regiones "A" y "B" de las estaciones: Badiraguato, El Quelite y Picachos	119
13	Curvas de Andrews para las estaciones: San M. Zapotitlán, San Blas, La Tina, Bamicori, y Las Cañas (región "A")	120
14	Curvas de Andrews para las estaciones: El Mahone, Choix, Huites, Palo Dulce, Chinipas y San Francisco (región "A")	121
15	Curvas de Andrews para las estaciones: San Ignacio, La Veranera Urique, Guerachic, Alamos y Cazanate (región "A")	122
16	Curvas de Andrews para las estaciones: Jaina, Naranjo, Zopilote Guamuchil, Pericos y Tierra Blanca (región "A")	123
17	Curvas de Andrews para las estaciones: Pte. Sud Pacífico, Palos Blancos, El Varejonal y Badiraguato (región "A")	124
18	Curvas de Andrews para las estaciones: Toahayana, Tecusiapa, Los Molinos, Guatenipa y La Huerta (región "B")	125
19	Curvas de Andrews para las estaciones: Pte. Cañedo, Sanalona, Picachos, Tamazula, El Bledal y Santa Cruz (región "B")	126
20	Curvas de Andrews para las estaciones: Acatitán, Ixpalino, Piaxtla y El Quelite (región "B")	127

21	Coefficientes de variación vs QBAR/AREA	128
22	Prueba de homogeneidad de Langbein para las 42 estaciones	141
23	Prueba de homogeneidad de Langbein para las 23 estaciones de la región "A"	142
24	Prueba de homogeneidad de Langbein para las 19 estaciones de la región "B"	142
25	Distribución Gumbel Mixta univariada, estación Huites	145
26	Distribución Gumbel Mixta bivariada, estación Huites	146
27	Hidrograma triangular y unitario adimensional para el gasto de 10,000 años de período de retorno de la estación Huites	148
28	Distribución Gumbel Mixta bivariada, estación La Veranera	175

## **I. INTRODUCCION**

De todas las sustancias necesarias para la vida, el agua es sin duda la más importante. El agua sigue sustentando todas las formas de vida: algunos organismos de gran simplicidad pueden existir sin aire, pero ninguno puede desarrollarse sin agua. El agua ha dado origen a grandes civilizaciones, y en el transcurso de millones de años ha sido uno de los principales agentes en la estructuración de la superficie terrestre. Del volumen total de agua en el planeta, un poco más del 97% es de mar; del restante 3%, el agua freática representa el 22% y el 77% es hielo. Esto deja apenas un 1% del total de agua dulce en el ciclo hidrológico, del que aproximadamente la mitad se encuentra en ríos, lagos y marismas. Bien administrada, habría suficiente agua para cubrir las necesidades humanas, presentes y futuras pero su irregular distribución, la contaminación y sobre todo el mal uso están poniendo en peligro su abastecimiento.

Esta irregular distribución espacial y temporal del agua hace necesaria la construcción de obras de aprovechamientos hidráulicos. El buen diseño de estos proyectos radica en maximizar la cantidad de agua disponible. El fundamento para estimarlo correctamente, es la construcción de obras que puedan controlar y administrar el exceso de agua. El análisis detallado de los eventos extremos proporcionará sin duda, la mejor base para llevar a cabo estos diseños. Cabe señalar que el diseño final debe evitar derrames catastróficos y permitir una adecuada operación, además de garantizar un costo económicamente factible.

El ingeniero hidrólogo desconoce la magnitud de los escurrimientos que afectarán al proyecto en el futuro, es por eso que se debe calcular la probabilidad de ocurrencia de cierto evento, asociándolo a un período de retorno. La técnica estadística del análisis de frecuencias, permiten analizar los escurrimientos para luego poder realizar extrapolaciones de carácter probabilista. Sin embargo, uno de los problemas más comunes a los que se enfrenta el ingeniero es la escasa o nula información en lugares en donde se requiere diseñar y construir alguna obra para formar un aprovechamiento hidráulico.

La regionalización es empleada en la Hidrología para facilitar la transferencia de información desde lugares con una mayor longitud de registro, a sitios cuyos registros son escasos o nulos. El proceso de regionalizar abarca una serie de detalles que deben cuidarse para lograr una confiable y existosa estimación de los eventos a predecir. Una de las partes más importantes en el análisis regional es el lograr una correcta discriminación de las variables que intervendrán en el modelo, así como las cuencas que serán empleadas y de las cuales se generará toda la información; he aquí la importancia de contar con un buen procedimiento, ya que una mala metodología puede ocasionar problemas tan graves como alterar la muestra o perjudicar registros con el afán de querer mejorar la calidad de la información.

La inherente dificultad que entraña regionalizar no depende de una simple comparación geográfica o del empleo de los registros de una sola estación de aforos, sino el lograr construir un modelo con el mayor número de variables significativas y lograr, en algunos casos, que la información de pocas estaciones sirva para mejorar o completar la de otra y enriquecer de esta manera y cuando sea posible los eventos a predecir.

En este trabajo se desarrolla un procedimiento para inferir los eventos de diseño tanto en cuencas aforadas como en no aforadas, mediante el empleo de distribuciones multivariadas de valores extremos y de las características fisiográficas y meteorológicas de la región en estudio.

En el capítulo 2 serán presentadas las características fisiográficas y meteorológicas seleccionadas para el análisis regional, así como los criterios para subdividir regiones, llamadas "*técnicas de homogeneidad de cuencas hidrológicas*", que comprenden de la selección del grupo más apropiado para el análisis, pruebas de significancia de grupo y la definición de regiones homogéneas con base en cinco conceptos fundamentales: primero, el criterio tradicional de homogeneidad con base en regiones geográficamente contiguas, segundo, empleando las características fisiográficas de la región (curvas de Andrews) en donde la definición de regiones se realiza a través de un criterio gráfico. El tercer y cuarto método consideran determinar la homogeneidad de los coeficientes de

variación de los gastos máximos anuales, tanto entre grupos (criterio de Wiltshire), así como derivar una metodología de agrupación denominada por racimos.

Finalmente, la subdivisión de cuencas se realiza de acuerdo con pruebas que requieren un análisis previo de frecuencias (prueba de Langbein).

Algunos de estos análisis de homogeneidad de cuencas y sobre todo, la discriminación de las variables en competencia, requieren conceptos de "*análisis de varianza*", que se presentan en el capítulo 3.

En el capítulo 4 se desarrolla la metodología de "*correlación y regresión*". En él se señalan las consideraciones más importantes para crear un modelo de transferencia de información, tanto para cuencas aforadas como no aforadas. Una parte muy importante dentro del análisis de correlación y regresión es garantizar que las muestras cumplan con el requisito de normalidad. Es por esto que se presentan algunos algoritmos de transformación de muestras.

En lo que se refiere a las cuencas en donde sí se tenga una serie de registros de escurrimientos máximos (cuencas aforadas), se busca una extrapolación confiable de los eventos de diseño, para lo cual en el capítulo 5 se analizan las "*distribuciones univariadas de probabilidad*" de mayor uso en la predicción de fenómenos hidrológicos. En la parte final del capítulo se presenta la distribución logística multivariada de valores extremos como una alternativa para el análisis de frecuencias de eventos extremos.

En el capítulo 6 se lleva a cabo la aplicación de las técnicas expuestas, a la región noroeste de la República Mexicana (región hidrológica 10). En donde se inicia con el cálculo de las características fisiográficas de la región, seguido de un análisis univariado de frecuencias, de un proceso de normalización y discriminación de variables, así como de las correspondientes pruebas de homogeneidad. Además, se comprueba que existe una mejora en la estimación de los parámetros a través de distribuciones bivariadas, logrando disminuir significativamente el error estándar de ajuste en el análisis de frecuencias. Finalmente, al relacionar las características fisiográficas de las cuencas con los parámetros de sus distribuciones de probabilidad, se desarrollan las correspondientes ecuaciones regionales para determinar eventos de diseño en cuencas no aforadas.

Las conclusiones y principales contribuciones de este trabajo se detallan en el capítulo 7

## **2. IDENTIFICACION DE REGIONES HOMOGENEAS**

### **2.1 Características fisiográficas de la cuenca**

Los procesos hidrológicos que ocurren dentro de una cuenca drenada, son determinados principalmente por sus propiedades físicas. Para definirse la relación entre estas propiedades y su respuesta hidrológica, el comportamiento hidrológico de las cuencas puede predecirse sin la necesidad de mediciones directas de los procesos. Sin embargo, esto puede no ser práctico o posible; si se desea tomar todas las características físicas que pueden considerarse como determinantes, Bolton en 1965 propone trece índices morfométricos de estas características.

Aparte de éstos existen índices de clima, tipo de suelo y uso de la tierra. El área de una cuenca, en particular, tiene una fuerte correlación con los índices de escurrimiento, (Anderson, 1957). Otras variables que pueden usarse para estimar el gasto son, la pendiente (Benson, 1962; Nash y Shaw, 1966; Newson, 1976), la capacidad del almacenamiento (Acreman, 1985); la longitud de la corriente principal ha probado ser un valioso estimador del tiempo de retraso en hidrogramas, (Bell y Om Kar, 1969; Kennedy y Watt, 1969) y al gasto base se le ha relacionado con la geología de la cuenca (Institute of Hidrology, 1980).

En la tabla 2.1 se presentan las principales variables asociadas a las características de una cuenca, que se emplearán en este estudio como indicativas de su comportamiento hidrológico.

**TABLA 2.1. CARACTERISTICAS FISIOGRAFICAS EMPLEADAS EN EL ESTUDIO**

VARIABLE	DESCRIPCION	UNIDADES
AREA	Area drenada hasta la estación	km <sup>2</sup>
PRECMEDA	Precipitación media anual	mm
PENDCUENC	Pendiente media de la cuenca	m km <sup>-1</sup>
ELEVMEDE	Elevación media de la cuenca	msnm
LONGITUD	Longitud de la corriente principal	km
PENDCAUC	Pendiente de la corriente principal	m km <sup>-1</sup>
PARMFORM	Parámetro de forma	km <sup>-1</sup>
FRECUENC	Frecuencia de corriente	km <sup>-2</sup>
DENSIDAD	Densidad de drenaje	km <sup>-1</sup>
CENTLAT	Latitud del centroide de la cuenca	grados
CENTLON	Longitud del centroide de la cuenca	grados
COEFESCR	Coefficiente de escurrimiento	
ORDEN	Orden de la corriente principal	
PERIMET	Perímetro del parteaguas	km
TEMP	Temperatura media anual	°C

A continuación se presenta una breve descripción de cómo fueron obtenidas estas características. La región hidrológica 10 ubicada al noroeste de la República Mexicana, y dada su importancia desde el punto de vista agrícola, y por sus caudalosas corrientes, es una excelente región para llevar a cabo un estudio hidrológico. Además la estación hidrométrica Huites se encuentra muy bien instrumentada, por lo cual será empleada como ejemplo en el desarrollo de todas estas características y de los conceptos que se estudiarán posteriormente.

### **2.1.1 Area y perímetro del parteaguas de la cuenca**

Se entiende como área drenada de una cuenca, al área en proyección horizontal encerrada por el parteaguas. El perímetro de éste se obtuvo con un curvímeter, midiendo su longitud total en km; mientras que el área se determinó con un planímetro y fue expresada en kilómetros cuadrados. Estas mediciones se efectuaron en un plano editado por la Coordinación General de Servicios Nacionales de Estadística, Geografía e Informática (INEGI), a escala 1:1 000 000

## 2.1.2 Precipitación media anual

La precipitación media anual se calculó por el método de las isoyetas, se empleó para ello un plano de isoyetas medias anuales a la misma escala del que fue empleado en el cálculo de las áreas. La altura de la precipitación media se calculó con la siguiente expresión:

$$hp_m = \sum_{i=1}^n hp_i \frac{A_i}{A} \quad (2.1)$$

donde

- A      área total de la cuenca, en km<sup>2</sup>
- A<sub>i</sub>    área entre isoyetas, en km<sup>2</sup>
- hp<sub>i</sub>    altura de precipitación entre isoyetas, en mm
- hp<sub>m</sub>    altura de precipitación media, en mm
- n      número de tramos entre isoyetas

Con ayuda del programa ISOYETAS.BAS se calcularon las precipitaciones medias para las 42 cuencas. Para la estación Huites, por ejemplo, se obtuvieron los siguientes resultados:

ISOYETAS	PRECIPITACION hp <sub>i</sub> ( mm )	AREA ENTRE ISOYETAS ( km <sup>2</sup> )	hpi*Ai ( mm*km <sup>2</sup> )
1500-1200	1350	0.0	0
1200-1000	1100	2375.0	2612500
1000-800	900	9025.0	8122500
800-700	750	11962.5	8971875
700-600	650	2407.5	1564875
600-500	550	250.0	137500
500-400	450	0.0	0
400-300	350	0.0	0
		26020.0	21409252

Sustituyendo valores en la ecuación 2.1, se tiene:

$$hp_m = \frac{21409252}{26020} = 822.8 \text{ mm}$$



### 2.1.3 Pendiente media de la cuenca

Para aplicar el criterio de Horton se traza una malla cuyo eje principal (x) sigue aproximadamente el sentido del cauce principal. Como ejemplo, en la cuenca de la estación hidrométrica Huites, se llevaron 19 divisiones sobre el eje x y 18 sobre el eje y, obteniéndose 342 cuadros de 13 km por lado, (ver plano1).

Una vez hecho lo anterior, se mide la longitud de cada línea de la malla comprendida dentro de la cuenca y se cuentan las intersecciones y tangencias de cada línea con las curvas de nivel. La pendiente de la cuenca en cada dirección de la malla se valúa como:

$$S_x = \frac{N_x D}{L_x} \quad (2.2)$$

$$S_y = \frac{N_y D}{L_y} \quad (2.3)$$

donde

D	desnivel constante entre curvas de nivel
$L_x$	longitud total de las líneas de la malla en la dirección x, comprendidas dentro de la cuenca
$L_y$	longitud total de las líneas de la malla en la dirección y, comprendidas dentro de la cuenca
$N_x$	número total de intersecciones y tangencias de las líneas de la malla con las curvas de nivel en la dirección x
$N_y$	número total de intersecciones y tangencias de las líneas de la malla con las curvas de nivel en la dirección y,
$S_x$	pendiente de la cuenca en la dirección x
$S_y$	pendiente de la cuenca en la dirección y

Finalmente, Horton considera que la pendiente media de la cuenca puede determinarse como

$$S_c = \frac{N D \sec \theta}{L} \quad (2.4)$$

$$L = L_x + L_y \quad (2.4 a)$$

$$N = N_x + N_y$$

(2.4 b)

donde

$\theta$  ángulo entre las líneas de la malla y las curvas de nivel

Como resulta demasiado laborioso determinar la  $\sec \theta$  de cada intersección, Horton sugiere usar un valor promedio de 1.57. En la práctica, y para propósitos de comparación, es igualmente eficaz ignorar el término  $\sec \theta$ , o bien considerar el promedio aritmético o geométrico de las pendientes  $S_x$  y  $S_y$  como pendiente de la cuenca. El programa PENDMED.BAS que se muestra en el apéndice, se desarrolló para optimar los cálculos de la pendiente media en cada una de las cuencas en estudio.

A continuación se contaron las intersecciones con las curvas de nivel de cada recta paralela a los ejes, y sus longitudes correspondientes limitadas por el parteaguas. Los resultados se muestran en la siguiente tabla

LINEA EN LA MALLA	INTERSECCIONES		LONGITUDES ( km )	
	$N_x$	$N_y$	$L_x$	$L_y$
1	0	0	20.0	0.0
2	1	0	37.0	0.0
3	12	0	58.0	0.0
4	13	0	90.0	0.0
5	8	8	112.0	90.0
6	13	9	133.0	94.0
7	18	18	160.0	120.0
8	11	13	162.0	142.0
9	16	20	165.0	162.0
10	22	15	160.0	167.0
11	10	17	137.0	166.0
12	9	19	124.0	150.0
13	17	20	117.0	142.0
14	19	13	126.0	131.0
15	8	13	117.0	117.0
16	7	16	72.0	144.0
17	1	8	80.0	125.0
18	1	3	41.0	88.0
19	0	3	0.0	74.0
20	0	10	0.0	75.0
21	0	2	0.0	67.0
22	0	0	0	20.0
	186	207	1911.0	2074.0

Como el desnivel entre curvas de nivel es de  $D = 0.500$  km, la pendiente de la cuenca, según la ecuación 2.4 vale, considerando  $\sec \theta = 1$

$$S_c = \frac{393 \times 0.500}{3985} = 0.04931$$

y la pendiente en cada dirección será:

$$S_x = \frac{186 \times 0.500}{1911} = 0.04867$$

$$S_y = \frac{207 \times 0.500}{2074} = 0.04990$$

### 2.1.4 Elevación media de la cuenca

La variación en elevación de una cuenca, así como su elevación media, puede obtenerse fácilmente con el método de las intersecciones. El mapa topográfico de la cuenca se divide en cuadrados de igual tamaño, considerando que por lo menos 100 intersecciones estén comprendidas dentro de la cuenca. La elevación media de la cuenca se calcula como el promedio de las elevaciones de todas las intersecciones. El programa ELEV MED.BAS que se muestra en el anexo de programas, ayudó a sistematizar el cálculo.

Para aplicar el método de las intersecciones, se utilizó la malla que se trazó para calcular la pendiente media. Así, 149 puntos de intersección quedaron dentro de la cuenca, y se tomó la elevación de cada punto. Finalmente la elevación media es igual a la suma de todas las elevaciones entre el número total de intersecciones, o sea:

$$E_m = \frac{285029.55}{149} = 1912.95 \text{ msnm}$$

### 2.1.5 Orden y longitud de la corriente principal

Otras características importantes de cualquier cuenca son las trayectorias o el arreglo de los cauces de las corrientes naturales dentro de ella. La razón de su importancia se manifiesta en la eficiencia del sistema de drenaje en el escurrimiento resultante. Por otra parte, la forma de drenaje proporciona indicios de las condiciones del suelo y de la superficie de la cuenca.

Las características de una red de drenaje pueden describirse, principalmente, de acuerdo con el orden de las corrientes, longitud de tributarios y densidad de corriente.

Antes de hablar del orden de las corrientes, conviene ver su clasificación. Todas las corrientes pueden dividirse en tres clases generales, dependiendo del tipo de escurrimiento, el cual está relacionado con las características físicas y condiciones climáticas de la cuenca. Así una corriente puede ser efímera, intermitente o perenne. Una corriente efímera es aquella que sólo lleva agua cuando llueve e inmediatamente después. Una corriente intermitente lleva agua la mayor parte del tiempo pero, generalmente, en época de lluvias su aporte cesa cuando el nivel freático desciende por debajo del fondo del cauce. La corriente perenne contiene agua todo el tiempo, ya que aún en época de sequía es abastecida continuamente, pues el nivel freático siempre permanece por arriba del fondo del cauce.

El orden de las corrientes es una clasificación que proporciona el grado de bifurcación dentro de la cuenca. El procedimiento más común para esta clasificación es considerar como corrientes de orden uno, aquellas que no tienen ningún tributario; de orden dos a las que sólo tienen tributarios de orden uno; de orden tres aquellas corrientes con dos o más tributarios de orden dos, etc. Así el orden de la corriente principal indicará la extensión de la red de corrientes dentro de la cuenca. Para hacer esta clasificación se requiere de un plano de la cuenca que incluya tanto corrientes perennes como intermitentes. Una vez que se sabe cual es el orden de la red de drenaje, es muy fácil determinar cual es la corriente principal; se siguen las corrientes de mayor orden partiendo de la estación hasta llegar a la parte más alta de la cuenca (en un tributario de orden 1). Sin embargo, no se debe perder de vista que la escala del plano base es determinante en este concepto; la *Ley de Horton de números de ríos* (Horton, 1945) determina que la relación de bifurcaciones es relativamente constante de un orden a otro, ésta, es una buena alternativa cuando estos problemas son significativos.

#### **2.1.6 Pendiente de la corriente principal**

La pendiente de un tramo de río es el desnivel que existe entre los extremos inicial y final, y la distancia horizontal de dicho tramo. El concepto de pendiente de la corriente principal, representa un valor medio; ya que cada tramo de río tiene una pendiente propia. De esta manera, dividiendo el cauce en "n" número de tramos; el perfil del río se aproximará más al real mientras mayor sea este número y la pendiente del cauce será la real conforme disminuya la distancia horizontal de cada tramo analizado.

La fórmula que proponen A.B. Taylor y H.E. Schwarz, se basa en considerar que el río está formado por una serie de canales con pendiente uniforme, cuyo tiempo de recorrido " $t_i$ " es igual al del río. Si se representa a la distancia horizontal entre cada curva de nivel como " $X_i$ " para cada tramo, entonces

$$t_i = \frac{X_i}{V_i} \quad (2.5)$$

Según la ley de Chezy, la velocidad de cada tramo, es función de la pendiente del tramo y de un coeficiente "k" que depende de la forma de la sección transversal y por consiguiente de la rugosidad

$$V_i = C\sqrt{Rh_i S_i} \quad (2.6)$$

Por otro lado si consideramos a la distancia total "L" del cauce como la suma de las "n" distancias por los "n" tramos

$$L = \sum_{i=1}^n X_i \quad (2.7)$$

y el tiempo total de recorrido "T" como la suma de los tiempos parciales "t<sub>i</sub>", tenemos:

$$T = \sum_{i=1}^n t_i \quad (2.8)$$

$$t_i = \frac{X_i}{k\sqrt{S_i}} \quad (2.9)$$

$$T = \frac{L}{k\sqrt{S}} \quad (2.10)$$

sustituyendo (2.7) y (2.8) e igualando (2.9) y (2.10) obtenemos:

$$\frac{L}{k\sqrt{S}} = \sum_{i=1}^n \frac{X_i}{k\sqrt{S_i}} \quad (2.11)$$

Si de (2.11) despejamos "S" obtenemos la ecuación propuesta por A.B. Taylor y H.E. Schwarz.

$$S = \left( \frac{n}{\frac{1}{\sqrt{S_1}} + \frac{1}{\sqrt{S_2}} + \dots + \frac{1}{\sqrt{S_n}}} \right)^2 \quad (2.12)$$

donde

- n número de tramos en que se divide el río
- $S_1, S_2, \dots, S_n$  pendiente de cada tramo
- S pendiente media del río

A esta fórmula se le conoce como de Taylor-Schwarz y es la que ayudó a calcular las 42 pendientes de la región en estudio. Para lograr una mayor eficiencia y rapidez en éste cálculo, se realizó un programa para computadora TAYLOR.BAS al que se le proporcionan los datos del cadenamamiento del río, y el número de tramos en los que se quiere dividir la longitud total.

Para aplicar el criterio de Taylor y Schwarz, al cauce principal (río San Miguel) de la estación Huites, se dividió la corriente en estudio en 30 tramos iguales de 8.9 km cada uno. Mediante interpolación de valores se tiene la elevación de cada punto y por consecuencia, su desnivel; con este valor se puede calcular la pendiente de cada tramo, con solo dividir este desnivel entre su distancia horizontal. En la siguiente tabla se proporciona el desnivel de cada tramo, su correspondiente pendiente y el factor  $1/\sqrt{S}$ , que requiere la fórmula.

CALCULO DE LA PENDIENTE POR TAYLOR-SCHWARZ, ESTACION HUITES

TRAMO	$S_i$	$1/\sqrt{S}$	TRAMO	$S_i$	$1/\sqrt{S}$
1	0.011	9.393	16	0.004	16.125
2	0.011	9.393	17	0.004	16.125
3	0.011	9.393	18	0.004	16.125
4	0.011	9.393	19	0.004	16.125
5	0.011	9.393	20	0.004	16.125
6	0.011	9.393	21	0.004	16.125
7	0.012	8.965	22	0.004	16.125
8	0.016	8.000	23	0.004	16.125
9	0.016	8.000	24	0.004	16.125
10	0.016	8.000	25	0.004	16.125
11	0.016	8.000	26	0.004	16.125
12	0.016	8.000	27	0.004	16.125
13	0.016	8.000	28	0.004	16.125
14	0.015	8.211	29	0.013	8.751
15	0.004	16.125	30	0.024	6.476

$$S = \left( \frac{n}{9.393+9.393+\dots+6.476} \right)^2 = 0.00685$$

### 2.1.7 Parámetro de forma

Esta característica indica de una manera muy precisa, la forma que tiene cada cuenca expresándola como una relación entre el perímetro de ésta (P) y su área (A) se expresa en 1/km, y es un buen indicador de como será la respuesta al escurrimiento a partir de la forma de la cuenca. Sin embargo, también en este concepto esta muy ligada la escala del plano que se emplea para calcular (P) y (A).

$$P_f = \frac{P}{A} \quad (2.13)$$

El parámetro de forma para la cuenca de la estación Huites será:

$$P_f = \frac{843}{26020} = 0.0324$$

Esto significa que por cada kilómetro cuadrado se tienen 0.0324 km del perímetro de la cuenca, esto influye en el escurrimiento, ya que entre mayor sea este valor, mayor será el perímetro por km<sup>2</sup> de área, que se encuentre limitando la cuenca.

### 2.1.8 Frecuencia de la corriente

La frecuencia junto con la densidad de drenaje, son dos de las características que intentan describir la relación lluvia escurrimiento de una cuenca, la primera se expresa como la relación entre el número total de corrientes (N<sub>s</sub>) y el área drenada (A).

$$F_c = \frac{N_s}{A} \quad (2.14)$$

Para determinar el número de corrientes (N<sub>s</sub>) se consideran solo las corrientes perennes e intermitentes. La corriente principal se cuenta como una desde su nacimiento hasta su desembocadura. Después se tendrán todos los tributarios de orden inferior, desde su nacimiento hasta la unión con la corriente principal, así sucesivamente hasta llegar a los tributarios de orden uno.

Esta relación no siempre proporciona una medida real de la eficiencia de drenaje, pues puede suceder que se tengan dos cuencas con la misma frecuencia de corriente y estén drenadas en muy diferente forma, dependiendo de la longitud de la corriente principal y de sus tributarios.

Para la cuenca en la estación Huites se tiene:

$$F_c = \frac{28}{26020} = 0.0011$$

Es decir que por cada kilómetro cuadrado de área se tienen 0.0011 cauces capaces de drenar los escurrimientos.

### 2.1.9 Densidad de drenaje

Esta característica es la que proporciona una información más real que la anterior, ya que se expresa como la longitud de las corrientes (L) por unidad de área drenada (A).

$$D_d = \frac{L}{A} \quad (2.15)$$

Para la cuenca en la estación Huites se tiene:

$$D_d = \frac{267}{26020} = 0.0103$$

Esto quiere decir que el 1% del área total de la cuenca se encuentra formada por cauces naturales.

### 2.1.10 Ubicación del centroide de la cuenca

El centroide o centro de gravedad, es el lugar geométrico en donde se supone concentrada toda la superficie drenada por la cuenca. Este punto se determina para efectos de diseño y se designa como el centro de la tormenta que es empleada en un modelo lluvia-escurrimiento.

Los ejes coordenados empleados para este cálculo, generalmente, son los paralelos de latitud y meridianos de longitud. Para una exacta determinación de este punto, se sugiere descomponer el área total de la cuenca ( $A_t$ ) en "n" figuras regulares, cuyos centroides ( $X_1, Y_1$ ) y áreas ( $A_i$ ) sean conocidos; hasta que, de una manera aproximada se logre reproducir la forma de la cuenca. El siguiente paso es calcular las coordenadas del centro de gravedad, para esto se hace uso de la fórmula para superficies geométricas compuestas, que está dada por:



$$X_c = \frac{X_1 A_1 + X_2 A_2 + \dots + X_n A_n}{A_t} \quad (2.16)$$

$$Y_c = \frac{Y_1 A_1 + Y_2 A_2 + \dots + Y_n A_n}{A_t} \quad (2.17)$$

donde

- $X_i$  distancia para cada figura de su centroide al eje vertical de referencia
- $Y_i$  distancia para cada figura de su centroide al eje horizontal de referencia
- $A_i$  área para cada "i-ésima" figura en que se divide la cuenca.
- $A_t$  área total de la cuenca
- $X_c$  distancia del centro de gravedad de la cuenca al eje vertical de referencia
- $Y_c$  distancia del centro de gravedad de la cuenca al eje horizontal de referencia.

Los centros de gravedad para cada una de las cuencas, obtenidos mediante el procedimiento anterior, se muestran en la tabla resumen 2.2.

### 2.1.11 Coeficiente de escurrimiento

Se define como coeficiente de escurrimiento anual a la relación entre el volumen del escurrimiento directo y el volumen total de lluvia.

$$C = \frac{V_e}{V_{ll}} \quad (2.18)$$

donde

- $C$  coeficiente de escurrimiento, adimensional
- $V_e$  volumen de escurrimiento directo, en  $m^3$
- $V_{ll}$  volumen total de lluvia, en  $m^3$

Este coeficiente, también puede obtenerse a través de tablas en función de usos indicativos de la tierra y tipos de superficie; ellas aparecen en libros o manuales relacionados con Hidrología. Sin embargo, para lograr la selección correcta del coeficiente de escurrimiento, no basta con escogerlo de alguna tabla o multiplicarlo por valores arbitrarios. Se recomienda que cuando existan áreas tributarias en zonas con diferentes valores de "C" se utilice un coeficiente ponderado, según sea la magnitud de cada área.

El coeficiente de escurrimiento *real* para el área drenada de la cuenca de Huites (26020 km<sup>2</sup>) se puede obtener al conocer el volumen medio anual de escurrimiento (3871x10<sup>6</sup> de m<sup>3</sup>) y la precipitación media que se calculó en el punto 2.1.2 (822.8 mm), así se tiene que:

$$C = \frac{3871 * 1000}{26020 * 822.8} = \frac{3871000}{21409256} = 0.1808$$

Mientras que para obtener un coeficiente ponderado es necesario conocer los diferentes usos del suelo. Los coeficientes de escurrimiento asociados a los distintos tipos de suelos, pueden obtenerse de tablas (Campos, 1987). Se emplearon para el estudio cartas geológicas, fisiográficas, de vegetación y de uso del suelo a escala 1:1 000 000 de planos editados por la Coordinación General de Servicios Nacionales de Estadística, Geografía e Informática (INEGI), de los cuales se obtuvo la siguiente información:

Uso del suelo	Area (km <sup>2</sup> )	Coefficiente
Selva Baja	7098	0.17
Bosque de Encino/Pino	15791	0.22
Matorral	1204	0.20
Agrícola de temporal	1926	0.10

El coeficiente ponderado será entonces:

$$C_p = \frac{(7098 * 0.17) + (15791 * 0.22) + (1204 * 0.20) + (1926 * 0.10)}{(7098 + 15791 + 1204 + 1926)} = 0.197$$

Lo anterior indica que un coeficiente ponderado es un valor muy cercano al real, por tanto puede considerarse confiable, ya que su cálculo implica varias áreas y coeficientes de escurrimiento. Esta misma metodología fue la que se empleó al calcular todos los coeficientes de escurrimiento ponderados para todas las cuencas. Para comparar los valores ponderados con los reales se calcularon éstos para toda la zona en estudio; obteniéndose valores más exactos. Por ejemplo, el valor real para la cuenca de la estación Toahayana resultó 0.196 y el ponderado de 0.193, para Acatitán el real fue 0.204 y el ponderado de 0.200. Dado que estas cuencas se encuentran en lugares significativos de la zona, puede considerarse que el emplear un coeficiente ponderado de escurrimiento es perfectamente válido.

La tabla 2.2 muestra el resumen de las características fisiográficas de la región hidrológica 10.

TABLA 2.2 RESUMEN DE LAS CARACTERISTICAS FISIOGRAFICAS DE LA REGION HIDROLOGICA No.10

MUESTRAS ORIGINALES

ESTACION / CARACTERISTICA	AREA (km <sup>2</sup> )	PRECMEDA (mm)	PENDCUENC (m km-1)	ELEVEMD (mnm)	LONGITUD (km)	PENDCAUC (m km-1)	PARAFORM (km-1)	FRECUENC (km-2)	DENSIDAD (km-1)	CENTLAT (grados)	CENTLON (grados)	COEFESCK	ORDEN	PERIMET (km)	TEMP (°c)
1 SM ZAPOTITLAN	34450	789.42	0.0406	1650.70	440	0.00342	0.0336	0.0046	0.0128	26.93	107.80	0.188	5	1158	23
2 SAN BLAS	33590	799.52	0.0414	1682.76	402	0.00353	0.0324	0.0046	0.0120	26.95	107.82	0.188	5	1088	23
3 LA TINA	275	654.73	0.0167	405.00	30.3	0.00413	0.2109	0.0036	0.0738	26.17	108.52	0.166	2	58	23
4 BANCORI	223	800.45	0.0061	415.00	20.2	0.00490	0.3363	0.0045	0.0906	26.37	108.40	0.162	1	75	23
5 LAS CAÑAS	29529	825.14	0.0445	1759.82	342	0.00538	0.0328	0.0049	0.0118	26.98	107.68	0.192	5	968	23
6 EL MAHONE	29428	826.08	0.0445	1759.82	337	0.00548	0.0326	0.0049	0.0117	26.98	107.68	0.192	5	959	23
7 CHOLX	1403	972.06	0.0221	1050.00	82	0.00234	0.1447	0.0050	0.0513	26.57	108.07	0.172	3	203	21
8 HUITES	26020	822.80	0.0493	1912.95	267	0.00685	0.0324	0.0051	0.0097	27.03	107.58	0.197	5	843	15
9 PALO DULCE	6323	911.37	0.0569	1742.05	184	0.01028	0.0614	0.0044	0.0269	27.60	108.20	0.210	4	388	15
10 CHINPAS	5098	909.69	0.0600	1940.00	131	0.01449	0.0651	0.0047	0.0235	27.70	108.15	0.211	4	332	15
11 SAN FRANCISCO	17531	773.47	0.0504	2015.74	246	0.00720	0.0379	0.0058	0.0131	26.80	107.80	0.191	4	665	25
12 SAN IGNACIO	10920	802.23	0.0536	2002.43	209	0.01041	0.0497	0.0067	0.0183	26.57	107.17	0.183	4	543	21
13 LA VERANERA	4106	739.77	0.0363	2354.62	137	0.01565	0.0852	0.0049	0.0312	27.30	107.40	0.215	3	350	11
14 URIQUE	4000	738.56	0.0358	2408.80	127	0.01537	0.0843	0.0050	0.0300	27.30	107.37	0.215	3	337	11
15 GUERACHIC	6262	818.71	0.0423	2322.16	145	0.01249	0.0564	0.0053	0.0216	26.37	106.80	0.177	4	353	17
16 ALAMOS	2270	634.25	0.0158	505.00	124	0.00184	0.1093	0.0035	0.0520	26.77	108.80	0.156	3	248	23
17 CAZANATE	1813	654.11	0.0156	556.36	93	0.00156	0.1120	0.0033	0.0458	26.85	108.82	0.161	3	203	23
18 JAINA	8179	982.60	0.0518	1279.18	177	0.01014	0.0560	0.0056	0.0241	26.00	107.43	0.187	4	458	21
19 TOAHAYANA	5281	1003.39	0.0579	1602.67	131	0.01130	0.0583	0.0064	0.0235	25.97	107.27	0.193	4	308	19
20 TECUSIAPA	3773	1040.37	0.0666	1893.68	90	0.01504	0.0861	0.0053	0.0220	26.00	107.12	0.197	4	325	15
21 LOS MOLINOS	501	919.96	0.0417	1300.00	35	0.02159	0.1936	0.0020	0.0639	25.73	107.27	0.182	2	97	23
22 NARANJO	2064	809.04	0.0209	586.36	107	0.00760	0.1177	0.0029	0.0484	26.20	108.30	0.158	3	243	23
23 ZOPILOTE	666	743.32	0.0283	437.50	68	0.00690	0.1922	0.0030	0.0871	26.00	108.23	0.158	2	128	25
24 GUAMUCHIL	1645	767.17	0.0096	201.11	65	0.00698	0.1173	0.0030	0.0383	25.53	107.83	0.145	3	193	25
25 PERICOS	270	734.26	0.0130	150.00	31	0.00613	0.2778	0.0037	0.1111	25.20	107.62	0.153	1	75	25
26 TIERRA BLANCA	11614	977.22	0.0434	1541.27	276	0.00603	0.0527	0.0030	0.0215	25.45	106.85	0.180	4	612	13
27 PTE. S PACIFICO	11434	978.50	0.0434	1541.27	268	0.00632	0.0525	0.0031	0.0214	25.45	106.85	0.180	4	600	13
28 PALOS BLANCOS	11409	982.29	0.0434	1563.29	256	0.00649	0.0508	0.0031	0.0210	25.45	106.85	0.181	4	580	13
29 EL VAREJONAL	10987	990.33	0.0470	1653.33	233	0.00680	0.0519	0.0029	0.0191	25.48	106.80	0.182	4	570	13
30 BADIRAGUATO	1018	983.50	0.0520	595.00	59	0.02118	0.1473	0.0029	0.0570	25.60	107.50	0.187	2	150	23
31 GUATEMPA	8252	1006.34	0.0491	1994.90	178	0.01202	0.0594	0.0032	0.0194	25.50	106.58	0.185	4	490	17
32 LA HUERTA	6149	1019.94	0.0412	2217.57	117	0.01569	0.0631	0.0036	0.0171	25.53	106.42	0.182	4	388	13
33 PTE. CAÑEDO	4086	1045.49	0.0433	774.80	117	0.00470	0.0710	0.0037	0.0289	25.00	106.90	0.185	3	290	25
34 SANALONA	3657	1080.66	0.0518	895.24	85	0.00438	0.0662	0.0036	0.0202	25.02	106.83	0.191	3	242	25
35 PICACHO	3280	1106.39	0.0518	895.24	85	0.00438	0.0729	0.0040	0.0213	25.02	106.83	0.191	3	239	25
36 TAMAZULA	2241	1171.95	0.0684	1146.43	48	0.01024	0.0848	0.0036	0.0192	25.05	106.77	0.196	3	199	23
37 EL BLEDAL	371	856.13	0.0206	267.50	31	0.00799	0.2372	0.0027	0.0809	24.70	107.00	0.185	1	88	25
38 SANTA CRUZ	8919	1068.89	0.0532	1689.60	180	0.01267	0.0561	0.0085	0.0205	24.70	106.18	0.199	4	500	13
39 ACAITAN	1884	1046.59	0.0497	726.92	115	0.01131	0.1131	0.0080	0.0515	24.32	106.55	0.200	3	213	25
40 DXPALINO	6166	1006.73	0.0430	1393.68	174	0.01310	0.0592	0.0070	0.0264	24.13	106.03	0.205	4	365	11
41 PIAXTLA	5307	1044.76	0.0496	1644.52	142	0.01689	0.0660	0.0066	0.0239	24.15	105.98	0.208	4	350	11
42 EL QUELITE	835	720.06	0.0084	211.67	42	0.00579	0.1653	0.0096	0.0515	23.65	106.37	0.182	2	138	25

## **2.2 Regionalización de cuencas de acuerdo con sus características fisiográficas y climatológicas**

La heterogeneidad de regiones, es un gran inconveniente cuando se trata de regionalizar cuencas; la magnitud de los errores en los que se puede incurrir pueden ser muy significativos afectando tanto a los eventos de diseño como a la ampliación de registros.

Para dividir un área de estudio en subregiones homogéneas es necesario considerar que tendrán un comportamiento similar, de esta manera, los datos requeridos serán inferidos con gran precisión, ya que las ecuaciones de regresión se basan en las características de la cuenca, logrando así gran confiabilidad para predecir cualquier variable hidrológica. Mosley (1981) estudió los problemas relacionados con la regionalización y notó que tanto la teoría como la práctica de esta técnica han recibido la atención de muchos investigadores de las ciencias geográficas, pero no se ha logrado obtener una metodología general para identificar regiones que pueden ser utilizadas en los análisis hidrológicos.

La mayor parte de las veces se emplea la técnica de ubicación geográfica, que se basa simplemente en trazar radios entre 70 y 80 km en cada estación y considerar como cuencas homogéneas, a las estaciones que se ubiquen dentro de estos círculos.

Muchos investigadores han definido a las subregiones homogéneas a partir del error cuadrático de los residuales de ecuaciones extremas de regresión. (Natural Environment Research Council, 1975, Tasker, 1982). Estos residuales son marcados en un mapa del área en estudio; se dibujan líneas que subjetivamente abarquen subregiones geográficamente contiguas formando de esta manera, grupos con valores residuales similares en tipo y medida.

Recientemente, técnicas multivariadas utilizan las características fisiográficas y meteorológicas de cada cuenca como un índice para identificar regiones hidrológicamente homogéneas (Tasker, 1982; Acreman y Sinclair, 1986; Hughes, 1987; Haines, 1988). Si el deseo es extrapolar los resultados a cuencas no aforadas, obviamente es preferible realizarlo con base en una clasificación de las características medias de la cuenca, que a registros hidrológicos que no existen en el sitio. Si se considera que los escurrimientos de la cuenca siguen una distribución multivariada normal, una región no aforada puede ser colocada dentro de un grupo en particular empleando para ello, un análisis de discriminación (Rao, 1965).

Tradicionalmente, y como lo propone el método de ubicación geográfica, los grupos de regionalización abarcan áreas geográficamente contiguas, generalmente, para un círculo de 80 km de radio. Sin embargo, las subregiones deben dividirse con base en similitudes hidrológicas o a las características de la cuenca, las cuales no siempre tienen un significado geográfico.

Con las técnicas de clasificación que se presentan a continuación, las cuencas se van ubicando en función de la similitud que existe entre las características que previamente han sido seleccionadas como significativas, y las de los elementos del grupo en el que se desea ubicar la estación.

A las técnicas que ubican estaciones en forma iterativa, con base en un algoritmo puramente matemático se les denomina: *técnicas de racimos*. Con esta metodología no es necesario que una cuenca se encuentre dentro de un grupo que sea geográficamente contiguo. Estas regiones podrá decirse, serán comunes respecto a sus series de datos en un espacio variable multidimensional antes que en el espacio geográfico. Sin embargo, existen numerosos problemas asociados a la regionalización en general y en particular al análisis de grupos. El primer problema común en las técnicas de regionalización, es el relacionado con la selección de las variables empleadas para valuar el grado de similitud entre las diferentes cuencas.

Algunas cuencas quizá son similares con respecto a ciertas variables (como el grado de la cubierta forestal o la lluvia anual) pero no a otras (por ejemplo, la fisiografía). Además, con el análisis por racimos, literalmente cualquier conjunto de variables es capaz de generar grupos. Lo cual es muy útil ya que de esta manera se podrán seleccionar las variables de acuerdo con la relevancia del problema. Si dichas variables son medidas a partir de longitudes de registros grandes, ésto contribuirá a obtener una mayor similitud que aquellas características que se obtienen a partir de registros más cortos. La solución general de este problema es un ordenamiento de todas las variables, así como una segura medición o estimación de cada una de ellas.

Otro problema encontrado cuando se usa el análisis por racimos, es la variedad de los diferentes algoritmos que pueden emplearse para formar grupos. Desafortunadamente, las diferentes técnicas de agrupación aplicadas a un mismo registro, a menudo producen estructuras que son sustancialmente diferentes. Esto es porque el escoger un método de agrupación implica imponer cierta estructura a la muestra de datos, lo cual a menudo implica definir una nueva y completa estructura. Desde luego que aún cuando de manera particular algún elemento no caiga dentro de ningún grupo natural, el análisis por racimos todavía detectará la existencia de nuevos grupos (o la no existencia).

Si existe una estructura significativa, entonces se puede esperar que una variedad de diferentes algoritmos, produzcan el mismo número de grupos, de tal manera que las cuencas serán clasificadas dentro de grupos similares. Sin embargo, aunque existen algunos lineamientos generales en la clasificación de la mayoría de las cuencas, existen algunas que deben clasificarse bajo ciertas bases arbitrarias, por ejemplo, cuencas que se encuentren dentro de una misma zona de influencia ciclónica que desea ser estudiada o que por cuestiones económicas, en donde necesariamente alguna región debe formar parte de un grupo de cuencas que abastecen o delimitan una zona de aprovechamientos hidráulicos.

La correcta delimitación de cada una de las regiones homogéneas sentará las bases de una técnica con la cual las cuencas que no hayan sido incluidas en el proceso de regionalización puedan ser correctamente colocadas en uno u otro de los grupos homogéneos. La solución de este problema es de crucial importancia, ya que permite realizar uno de los principales objetivos de la regionalización: la estimación de las características hidrológicas en cuencas no aforadas.

Un procedimiento alternativo fue adoptado por Hasines (1988). En él se asienta un algoritmo en forma de un árbol de decisiones, en donde las estaciones se agrupan con base en la característica que mejor define a esa cuenca. Esta técnica fue desarrollada para imitar la estructura de la clasificación derivada de un análisis de grupos. Desafortunadamente, hay un problema asociado con la aplicación de la técnica; ya que una nueva cuenca siempre será colocada en uno u otro de los grupos disponibles, pero como su clasificación depende solo de la variable escogida, entonces no puede determinarse qué tan indiferentes es al grado (o carencia) de similitud de las otras variables. Tales técnicas no son indicativas de si la predicción del criterio derivador de grupos es aplicable a cuencas no aforadas y bajo qué consideraciones.

En este capítulo se presenta una regionalización metodológica dirigida a problemas asociados con la selección de un grupo de técnicas apropiadas, de una definición de regiones homogéneas y de la predicción de los miembros de un grupo de cuencas nuevas, que de otro modo eran desconocidos.

### **2.2.1 Trazo multidimensional**

Si se considera que un conjunto de variables son capaces de especificar la semejanza hidrológica entre diferentes cuencas, entonces, cada cuenca y sus correspondientes características fisiográficas pueden ser dibujadas en un diagrama a escala, que revelará las similitudes (o agrupaciones) entre cada una de ellas, además, como el método preveé una discriminación previa de variables, éste

servirá para resaltar la importancia de cada una de ellas. A esta técnica se le conoce con el nombre de *trazo multidimensional en dos dimensiones* (Andrews, 1992).

A continuación se presenta el procedimiento y las recomendaciones para predecir el peso de cada una de las variables que se emplean:

- a) Obtener una serie de variables que se consideren significativas para la predicción de diferentes grupos (características fisiográficas).
- b) Identificar las variables más relevantes del problema a través de un análisis de varianza, técnica que será desarrollada en el capítulo 3.
- c) Seleccionar un peso para cada variable, acorde a su relativa importancia. Este procedimiento se lleva a cabo probando el estadístico "t" para cada característica, técnica que será desarrollada en el capítulo 4.
- d) Obtener las curvas propuestas por Andrews para cada una de las cuencas, con base en las características seleccionadas, representándolas por una curva en dos dimensiones a través de la siguiente función:

$$f(x) = \frac{x_1}{\sqrt{2}} + x_2 \sin(t) + x_3 \cos(t) + x_4 \sin(2t) + x_5 \sin(2t) + \dots \quad (2.19)$$

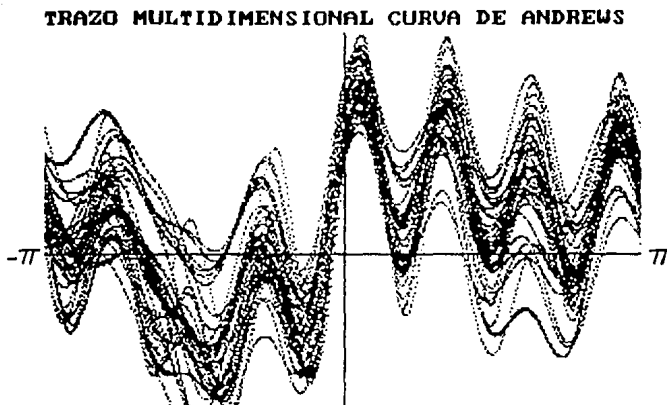
donde

$X_1, X_2, \dots$  representan a cada una de las variables fisiográficas de la cuenca.

t se dibuja en un rango entre  $-\pi$  a  $+\pi$ .

Una característica del método, que no es inmediatamente identificada, consiste en que los resultados que se obtengan dependen de la clase de las variables que son seleccionadas. Las primeras variables son asociadas con componentes cíclicas de bajas frecuencias y las últimas con componentes de frecuencias altas; esto puede ayudar a asociar a las variables con su importancia. Las bajas frecuencias son más fáciles de observar, de esta manera  $X_1$  representará la variable más importante,  $X_2$  a la segunda, y así sucesivamente.

La completa homogeneidad entre cuencas está ilustrada en las curvas de Andrews que se muestran en la figura 1. Estas curvas en particular fueron obtenidas para las cuencas de la región hidrológica 10, y fueron construidas con 15 características fisiográficas. Cada curva corresponde a una cuenca en particular, cabe señalar que se obtuvieron patrones similares de curvas para estas características, ya que en la elaboración de esta gráfica no se ha efectuado ninguna discriminación de variables.



**FIGURA 1 CURVAS DE ANDREWS PARA 15 CARACTERISTICAS FISIOGRAFICAS**

Esto puede interpretarse como que la mayoría de las cuencas describen un trazo regular, lo cual indica una total homogeneidad de la región. Sin embargo, la identificación de grupos homogéneos en términos visuales, implica la separación e identificación de grupos de curvas que formen franjas similares.

Cabe señalar la gran importancia que representa el buen uso y selección de las variables, ya que de alguna manera, alguien que no esté familiarizado con la técnica, podría suponer al observar estas curvas, que no existen subregiones homogéneas dentro de la región hidrológica 10.

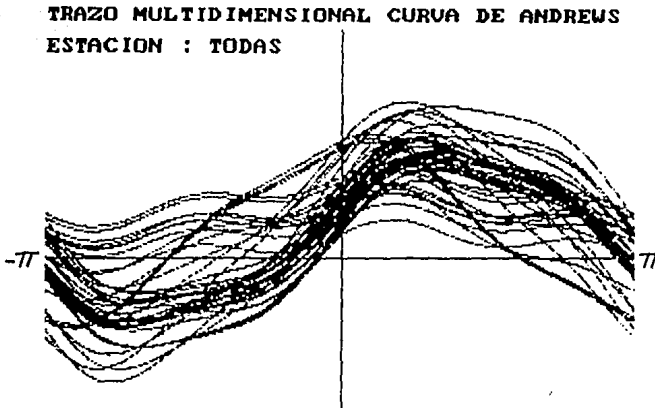
La figura 2 muestra las curvas, una vez que se discriminaron y jerarquizaron las características fisiográficas. Si se observa con cuidado esta figura, se concluye que el comportamiento de cada



curva depende de las variables, y del peso que se le dé a cada una de ellas. Aparentemente la figura 1 no marca ninguna diferencia entre ninguna de las cuencas, mientras que al seleccionar las variables, pueden distinguirse claramente dos grupos de cuencas.

- e) A partir de estas curvas de Andrews, discriminar envolventes intentando formar conjuntos de curvas similares y derivar de ésto una curva media que represente los atributos físicos de cada grupo.

Esta curva media del grupo, o asignada al grupo, es empleada como la curva característica de la región. Grupos de cuencas similares formarán una sola franja; mientras que las cuencas que caigan fuera de dicha franja de identificación, podrá decirse que pertenecen a otro grupo diferente.



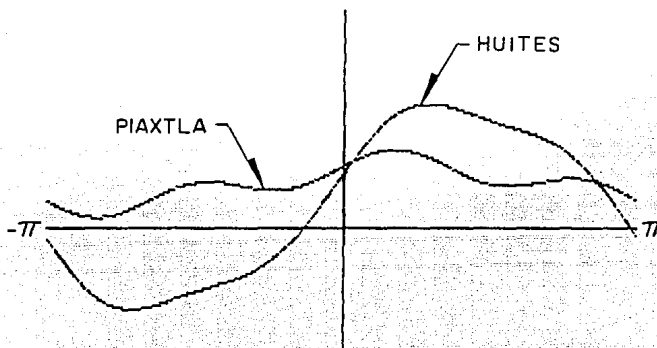
**FIGURA 2 CURVAS DE ANDREWS PARA 6 CARACTERISTICAS FISIOGRAFICAS**

Andrews demostró que las "n" variables analizadas, producirán curvas similares para cuencas homogéneas y además al extraer las envolventes, el análisis gráfico habrá preservado las distancias entre estaciones.

- f) Refinar la óptima agrupación obtenida en el punto anterior, a través de una comparación de la curva de cada cuenca con los diferentes grupos de asignación.

Una cuenca será subjetivamente identificada dentro de un grupo, si su curva sigue el mismo o la mayor parte del comportamiento de la envolvente que identifica al grupo. La figura 3 muestra dos ejemplos de curvas características que se generaron de la fisiografía de las cuencas Huites y Piaxtla, las cuales presentan claras diferencias en su comportamiento y por lo tanto pertenecen a dos regiones diferentes.

**TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : HUITES Y PIAXTLA**



**FIGURA 3 CURVAS DE ANDREWS PARA LAS ESTACIONES HUITES Y PIAXTLA**

g) Finalmente se prueban los criterios SEUCLID (SS) y COSINE (CS) en las curvas en donde se tenga la incertidumbre de a qué grupo pertenece, tal y como se indica a continuación.

Un de los criterios objetivos para asignar los miembros de cada grupo se obtiene por la suma del cuadrado de las diferencias (SS) entre los puntos (i) que forman la curva característica del j-ésimo grupo (S<sub>j</sub>) y la curva de la k-ésima cuenca (T<sub>k</sub>) en consideración.

$$SS = \sum (S_{ji} - T_{ki})^2 \tag{2.20}$$

La cuenca es localizada en el grupo con el menor valor de SS. El segundo criterio (CS) es la medida de unión entre estaciones, y aunque este procedimiento se basa en un análisis gráfico, puede

considerarse que el criterio (CS) sirve como un patrón similar al empleado en un análisis por racimos.

$$CS = \frac{\sum(S_{ji} * T_{ki})}{\sqrt{\sum S_{ji}^2 \sum T_{ki}^2}} \quad (2.21)$$

El criterio CS varía entre -1 y +1 en donde la perfecta correlación entre dos curvas se indica con un valor cercano a +1.

El hecho de que estos estadísticos consideren las diferencias entre la curva de una cuenca y la curva media de un grupo, los hace conceptualmente similares a un técnica de unión entre estaciones, denominados en la literatura como *métodos de liga*.

Con el propósito de hacer menos pesado el trabajo de dibujar a mano las cuencas en estudio, se desarrolló un programa para computadora que permite calcular, dibujar e imprimir de una manera muy rápida, las correspondientes curvas de Andrews, así como evaluar los criterios SS y CS. El listado del programa ANDREWS.BAS se muestra en el anexo de programas.

### 2.2.2 Ventajas y desventajas

Puede concluirse que la derivación de un criterio para la homogeneidad de cuencas a partir de las curvas de Andrews es recomendable por tres razones principales:

1. Permite que un gran número de cuencas puedan ser procesadas rápidamente.
2. Brinda una guía práctica para quien desee localizar una cuenca no aforada previamente.
3. Esta técnica no considera que los grupos de cuencas hidrológicamente similares sean geográficamente contiguos.

Gran importancia tiene la función de Andrews, ya que permite determinar los miembros de cada grupo para cuencas nuevas, los cuales no se conocían hasta entonces. La curva asociada con la nueva cuenca, es comparada con las curvas de los diferentes grupos, logrando así determinar cuál de ellas es la más parecida o apropiada para la región. Esta aproximación tienen además una considerable ventaja sobre la mayoría de los análisis de discriminación, ya que en éste es posible identificar a una cuenca que no pertenezca a ninguno de los grupos ya existentes.

La debilidad de este procedimiento y de cualquiera que se base solamente en el uso de las características de la cuenca, es que la agrupación obtenida es muy sensitiva a la selección inicial de las variables predictivas; las variables deben ser cuidadosamente escogidas y ponderadas, para asegurar que el resultado de similitud hidrológica entre grupos no sea tan solo un artefacto para seleccionar características de cuencas. Sin embargo, este método puede extenderse si se emplean en lugar de las características fisiográficas, por ejemplo, los estadísticos de la muestra, o los parámetros de alguna distribución de probabilidad asociados a cada cuenca en particular.

## **2.3 Regionalización de cuencas de acuerdo con el coeficiente de variación de los gastos máximos anuales**

### **2.3.1 Técnicas de agrupación por racimos**

Como ya se comentó, un algoritmo por racimos, es el que automáticamente agrupa cuencas en grupos acordes a un criterio específico.

Un algoritmo propuesto por Gordon (1981), llamado de *relocalización geográfica*, requiere para iniciar el proceso, de un número arbitrario  $M$  y la colocación de las cuencas se realiza aleatoriamente subdividiendo la región en  $M$  grupos, con igual número de cuencas cada uno. Para la localización exitosa de una cuenca dentro de un grupo, se debe obtener que el total de la suma del cuadrado de las distancias entre los centroides de las cuencas del grupo, sea mínimo, esto deberá ser probado para los  $M$  grupos que se formen.

En Nueva Zelanda, Mosley (1981) fue el primero en realizar un análisis por racimos para formar grupos de cuencas similares con base no solo en características fisiográficas o a su ubicación geográfica; sino que empleó específicamente las características de los gastos medios anuales y su correspondiente coeficiente de variación.

Siguiendo la iniciativa de Mosley, se presenta un procedimiento de ubicación iterativa de estaciones, empleando un algoritmo con base en pruebas estadísticas.

- a) Localizar en una gráfica a escala, una cuenca en forma individual a través de su área (AREA) y de dos de los estadísticos de su serie de escurrimientos: el escurrimiento medio anual (QBAR) y su coeficiente de variación (CV).

Se considera que estos dos estadísticos por sí solos, son capaces de describir mucho de lo que representa para el ingeniero una curva de frecuencias. QBAR describe la intensidad espacial de un gasto máximo promedio; mientras que CV caracteriza su máxima variación año por año, es decir, la variabilidad de una serie de escurrimientos máximos anuales. El coeficiente de variación (CV) está definido como:

$$CV = \frac{S_x}{\bar{X}} \quad (2.22)$$

donde:

$S_x$  desviación estandar de la muestra (insesgada)

$\bar{X}$  media aritmética de la serie

Una interpretación de los objetivos de esta técnica es que, las diferencias de los coeficientes de variación *dentro* de un grupo de cuencas, debe ser minimizadas; mientras que las variaciones *entre* grupos, deben ser maximizadas.

La gráfica se construye colocando los CV en el eje de las ordenadas y la relación QBAR / AREA en las abscisas; las cuencas que se localizan cerca unas de otras pueden ser reunidas formando grupos (racimos).

Esta técnica con dos estadísticos, prevé que los valores de QBAR y CV para cada cuenca sean obtenidos de registros con distribución normal, de tal manera que ninguna de las dos variables tenga una influencia que pueda desproporcionar a la otra.

Este procedimiento de dibujar los estadísticos es preferible sobre aquellos que solo jerarquizan el agrupamiento inicial a través de un número arbitrario; tendiendo a producir un gran y dominante grupo o un conjunto de pequeños satélites localizados en los límites del espacio de datos. Estos procedimientos que parten de una división arbitraria, Gordon (1981) deben posteriormente ser ajustados por un modelo, en donde intervengan el total de las características fisiográficas de áreas contiguas. Para esto debe considerarse que cada subregión producto de un análisis de este tipo es homogénea en términos de su respuesta hidrológica, pero esto no podía ser garantizado para cuencas vecinas con características fisiográficas muy diferentes.

b) En esta segunda etapa es necesario probar la significancia de cada grupo formado. Para ello debe aceptarse, que un grupo de cuencas presenta dos propiedades principales:

1. *Que pueden separarse o distinguirse de otros grupos de cuencas.*
2. *La homogeneidad en las características de sus escurrimientos, permite definir curvas de frecuencias similares.*

La prueba "F", una prueba del ratio de la varianza, es un método clásico para recalcar diferencias significantes entre grupos. En esta aplicación el estadístico "F" está definido como la varianza estimada según variación de los CV grupales

$$F = \frac{\text{Varianza estimada según variación entre CV grupales}}{\text{Varianza estimada según la variación de CV dentro del grupo}} \quad (2.23)$$

Para una eficiente subdivisión de cuencas, en grupos significativamente diferentes, el estadístico "F" se puede formular como se presenta a continuación:

Si "k" denota el grupo y "j" la estación, "m<sub>k</sub>" será el número de estaciones en el grupo "k". CV<sub>jk</sub> es el coeficiente de variación de la muestra en la estación "j" del grupo "k"

La variación total "T" del coeficiente de variación muestral, consiste en determinar la variación "A" dentro del grupo y la variación "B" entre grupos.

$$T = A + B \quad (2.24)$$

donde:

$$T = \left\{ \sum_k \sum_j CV_{jk}^2 - \left( \sum_k \sum_j CV_{jk} \right)^2 \right\} / \sum_k m_k \quad (2.25)$$

$$A = \left( \sum_k \sum_j CV_{jk}^2 \right) - \sum_k \left[ \left( \sum_j CV_{jk} \right)^2 / m_k \right] \quad (2.26)$$

$$B = \sum_k \left[ \left( \sum_j CV_{jk} \right)^2 / m_k \right] - \left[ \left( \sum_k \sum_j CV_{jk} \right)^2 / \sum_k m_k \right] \quad (2.27)$$

Si se toman "P" grupos se tendrá:

$$F = \frac{B/(p-1)}{A/(\sum m_k - p)} \quad (2.28)$$

Este desarrollo es el que se conoce como análisis de varianza (ANOVA) el cual se explicará a detalle en el capítulo 3, y es una prueba muy confiable de la regionalización de los coeficientes de variación.

En el anexo de programas se presenta el listado del programa ANOVA.BAS que se empleó para la clasificación de la región hidrológica 10. Wiltshire S.E. (1985) propuso modificar las expresiones anteriores empleando la varianza muestral de los coeficientes de variación  $U_{jk}$  calculada como:

$$U_{jk} = \frac{V_k}{n_{jk}} \quad (2.29)$$

donde:

$V_k$  varianza de los  $CV_{jk}$  del grupo  $k$   
 $n_{jk}$  longitud del registro en la estación "j" dentro del grupo "k"

Esta cualidad depende de las propiedades estadísticas de los escurrimientos en los sitios del grupo  $k$ , y se observa claramente que la varianza de la muestra es inversamente proporcional a su tamaño. Idealizando,  $V_k$  puede ser obtenida independientemente de los datos de escurrimiento, pero esto no es una manera práctica de realizarlo, ya que se obtendría un estimador sin sentido alguno de los registros empleados. (ver ecuación 2.38).

La variación total modificada "T" del coeficiente de variación muestral, se transforma en:

$$T = A + B \quad (2.30)$$

donde

$$T = \sum_k \sum_j \frac{(CV_{jk} - CV_{\bullet\bullet})^2}{U_{jk}} \quad (2.31)$$

$$A = \sum_k \sum_j \frac{(CV_{jk} - CV_{\bullet k})^2}{U_{jk}} \quad (2.32)$$

$$B = \sum_k \left( \sum_j \frac{1}{U_{jk}} \right) * (CV_{\bullet k} - CV_{\bullet\bullet})^2 \quad (2.33)$$

y

$$CV_{\bullet k} = \frac{\sum_j CV_{jk}/U_{jk}}{\sum_j 1/U_{jk}} \quad \text{es la media de los CV's en el grupo } k \quad (2.34)$$

$$CV_{\bullet\bullet} = \frac{\sum_k (CV_{\bullet k} \sum_j 1/U_{jk})}{\sum_k \sum_j 1/U_{jk}} \quad \text{es la media de los CV's en los grupos} \quad (2.35)$$

Si se toman los "P" grupos, se tendrá:

$$F = \frac{B/(p-1)}{A/(\sum m_k - p)} \quad (2.36)$$

La contribución del coeficiente A en cada grupo K, está denotada como  $S_k$ , donde:

$$A = S_k = \sum_j \frac{(CV_{jk} - CV_{\bullet k})^2}{U_{jk}} \quad (2.37)$$

El grupo de varianzas  $V_k$  de los coeficientes  $CV_{jk}$  es el promedio de las varianzas calculadas dentro del grupo en cada sitio. Esta exacta definición de la estimación de la varianza fue definida por Parr (1983) como la varianza extendida, la cual se describe a continuación:

Se considera una serie de elementos independientes, siguiendo una misma distribución  $Q_1, \dots, Q_n$

De esta manera la varianza extendida de los coeficientes de variación en cada sitio estará dada por:

$$V_{jk} = (n-1) \sum_i^n \left[ CV_{n-1}^i - \left( \frac{\sum_i CV_{n-1}^i}{n} \right) \right]^2 / n \quad (2.38)$$

donde

$CV_n$  denota el coeficiente de variación calculado para esa serie

$CV_{n-1}$  denota el coeficiente de variación calculado para los n-1 elementos de la serie y la varianza por grupo para los m sitios será igual a:



$$V_{jk} = \left( \sum_j n_{jk} V_{jk} \right) / m_k \quad (2.39)$$

El estadístico  $F$  distribuido con  $(p-1)$ ,  $(\sum m_k - p)$  grados de libertad, es la prueba básica que denota la diferencia entre los grupos. El parámetro  $S_k$  sigue la distribución  $X^2$  con  $(m_k - 1)$  grados de libertad y puede ser empleada como una prueba de la homogeneidad de los grupos, bajo la hipótesis nula, de que no existen diferencias entre el sitio,  $CV_n$  dentro de su grupo.

- c) Si se observa que el valor de  $F$  excede el valor crítico de  $(m_k - 1)$  grados de libertad, la hipótesis es rechazada y por lo tanto debe aceptarse que existen diferencias significativas entre el sitio (CV) y la región en cuestión. Deberán entonces, reacomodarse los grupos para buscar disminuir las diferencias significativas entre cada estación, y sobre todo lograr un incremento en el valor de  $F$ , lo que garantizará que un grupo en particular es homogéneo. Esto es importante, ya que aceptando esta alternativa de la hipótesis, se disminuye el poder de la prueba de  $F$ , así como los grados de libertad asociados con el denominador de la ecuación (2.36), ya que no serán exactamente  $(\sum m_k - p)$ .
- d) Repetir los pasos b) y c) para una agrupación diferente.
- e) Examinar la variación de los resultados de las pruebas estadísticas con las diferentes particiones, de tal manera que se encuentre una eficiente partición dentro de grupos homogéneos y distintos. Ello se hace donde sea necesario o se requiera, repetir los pasos b) y c) para converger a la solución óptima.

## 2.4 Regionalización de cuencas con base en los gastos máximos anuales

El uso de curvas adimensionales o de curvas de gastos crecientes, en las cuales se relaciona un valor estimado de la medida de los gastos ( $Q$ ) con gastos de menores frecuencias de ocurrencia, es uno de los procedimientos recomendados en el "Flood Studies Report" (NERC, 1975) para estimar gastos de diseño en sitios no aforados. El obtener una curva regional, es admitir, que existe homogeneidad en toda la región, en la cual, la distribución de gastos máximos anuales en todas las estaciones, es la misma cuando dichos gastos se dividen entre su respectiva media en cada estación. De este modo, si  $G(q/Q_i)$  es la distribución de los gastos adimensionales en el sitio "i", se tiene:

$$G_n(q/Q) = G(q/Q) \text{ para todos los sitios "i"}$$

donde

$G_n(q/Q)$  es la distribución adimensional de los gastos, para todos los sitios en la región "n".

Es necesario mencionar que dentro de la región, los datos en más de una estación pueden ser estimados por los parámetros regionales. Cualquiera de estos estimadores regionales, puede ser calculado para las cuencas de la región, especialmente para aquellas cuencas no aforadas, o en las que se tienen datos dudosos. Estos parámetros pueden ser reajustados para el sitio de interés ya sea a partir de los datos de la serie de escurrimientos o de alguna característica fisiográfica de la cuenca. (Acreman 1985).

La mayoría de los estudios regionales (NERC, 1975; Beable y Mckerchar 1982) consideran que las regiones deben ser geográficamente homogéneas. Sin embargo, las cuencas vecinas dentro de una región pueden ser física e hidrológicamente muy diferentes, por ejemplo, dos cuencas con similares curvas de distribución, pueden ser asignadas en distintas regiones. Además, de que los registros de escurrimientos es poco probable que sean independientes dentro de una cierta región geográfica.

Una alternativa de regionalización geográfica podría ser el escoger la curva de frecuencias para cada sitio en particular dentro de una región. Sin embargo, la gran variabilidad inherente en la estimación de parámetros en series de registros cortos (Benson 1960) implica que las curvas calculadas para cada estación pueden variar aún dentro de regiones homogéneas.

Estas diferencias en las muestras impedirá entonces la completa identificación de regiones con base en sus registros.

La regionalización con base en las características físicas de las cuencas, es un método útil de identificar grupos, los cuales pueden ser internamente homogéneos con respecto a sus distribuciones de  $q/Q$ .

Cada sitio puede ser ubicado en una región con base en sus características fisiográficas, independientemente de sus registros de escurrimientos, tal y como lo contempla el propuesto trazo multidimensional.

La mayoría de los métodos regionales que se basan en la distribución de los gastos máximos anuales, se ajustan por una de las tres distribuciones de valores extremos (Fisher y Tippett, 1928).

El procedimiento regional basado en el uso de la distribución de valores extremos tipo I (EVI) se hizo popular por el trabajo de Gumbel (1958). Jenkinson (1955) mostró que las tres distribuciones de Fisher y Tippett son un caso especial de la Distribución General de Valores Extremos (GEV).

Cabe señalar que la distribución GEV fue ampliamente recomendada para usarla en el Reino Unido por NERC (1975). Dalrymple (1960) encontró una prueba de homogeneidad investigando curvas de distribución en estaciones individuales, basada en la distribución (EVI), en esta prueba, hubo diferencias significativas para cada curva en particular, debido en parte a considerar regiones hipotéticas. Al emplear esta prueba, Cole (1966) fue capaz de identificar seis tipos de regiones homogéneas en Inglaterra y Gales.

La escasez de datos provocó que Biswas y Fleming (1966) encontraran diferencias significativas en las curvas de distribución estudiadas en Escocia empleando la metodología GVE. NERC (1975) consideró que la regionalización de Escocia podía realizarse sólo con dos curvas del tipo GEV, una para las partes bajas en el centro del país y las partes altas del sur. Estas regiones fueron formadas con base en la distribución geográfica de sus residuales (diferencia entre el gasto calculado y el medido). Estos residuales se obtuvieron a partir del modelo de regresión empleado para estimar gastos a partir de las características de todas las cuencas de Inglaterra. Presumiblemente, el comportamiento de los residuales se debió a las pobres especificaciones del modelo, y/o a las variaciones espaciales de los factores que influyen en el cálculo de los gastos y que no fueron tomados en cuenta por el modelo.

Stevens y Lynn (1978) examinaron si realmente estas regiones tenían diferencias significativas y encontraron que sólo una de las diez regiones no tenía mucha similitud con las otras.

Las curvas de NERC (1975) fueron obtenidas por un procedimiento gráfico para estimar los parámetros de la distribución GEV, pero Hosking (1985) recomendó el uso de un algoritmo alternativo, basado en el método de momentos pesados (PWM) introducido por Greenwood (1979) y seguido como análisis de frecuencias por Wallis (1980), Hosking demostró que para *muestras pequeñas*, la técnica de (PWM) es preferible a la estimación por máxima verosimilitud (MLE). Esto es evidente, ya que para series de muestras grandes (MLE) es ampliamente más eficiente.

A continuación se presenta la metodología a seguir cuando se desea identificar regiones homogéneas con base en la distribución de sus correspondientes gastos máximos anuales. La idea básica de esta prueba (Langbein, 1981) consiste en considerar, que dentro de una área hidrológicamente

homogénea, un número de estaciones hidrométricas han sido operadas y han registrado los efectos de los mismos fenómenos meteorológicos, entonces, se podrá probar la posible existencia de uno o varios grupos de cuencas, agrupados en regiones homogéneas, no hay que perder de vista que una correcta clasificación de este tipo de regiones producirá una confiable curva de frecuencias que será empleada principalmente, en cuencas no aforadas. Esta prueba de homogeneidad, ha sido utilizada en numerosos procedimientos de regionalización especialmente por el de la avenida índice (*index flood*), el cual intenta incrementar la seguridad al determinar las características de la curva de frecuencias dentro de una región.

a) Se propone un ajuste previo de frecuencias, con base en la distribución (EV1), la cual tiene una función de distribución del tipo:

$$F(x) = \begin{cases} \exp\{-[1 - \beta(x - \mu) / \alpha]^{1/\beta}\} & \beta \neq 0 \\ \exp\{-\exp[-(x - \mu) / \alpha]\} & \beta = 0 \end{cases} \quad (2.40)$$

donde  $x$  esta limitada por  $\mu + \alpha / \beta$ , para  $0 < \beta < \infty$

El programa GUMBEL.BAS que se muestra en el anexo de programas ajusta una muestra de datos a una distribución (EV1) estimando los parámetros por el método de momentos y de máxima verosimilitud. Sin embargo no hay que perder de vista que muchas veces resulta necesario probar cuál de las dos distribuciones (EV1) o (GVE) es más conveniente ajustar al registro; esto es, verificar si desde un punto de vista estadístico esta última resulta más conveniente que el modelo (EV1). La forma de llevar a cabo esto se realiza a través de diversas pruebas tal y como lo proponen Martínez y Raynal (1988), o al evaluar dos pruebas estadísticas; la primera desarrollada conceptualmente por Van Montfort (1970) y expuesto operativamente por Servruk y Geiger (1981), la segunda prueba corresponde a la propuesta por Otten y Van Montfort (1978).

b) A continuación se calculan los gastos de 2.33 (gasto medio) y de 10 años de periodo de retorno (gasto de creciente) con la forma inversa de la distribución (EV1).

$$X = \mu - \alpha \text{Ln}\{-\text{Ln}[F(x)]\} = \mu + \alpha(y) \quad (2.41)$$

$$y = -\text{Ln}\{-\text{Ln}[F(x)]\} \quad (2.42)$$

donde

$\mu$       parámetro de ubicación

$\alpha$       parámetro de escala

La variable reducida "y" tiene una desviación estandar igual a (Darlymple, 1960):

$$\sigma_y = \frac{e^y}{\sqrt{n}} \sqrt{\frac{1}{T-1}} \quad (2.43)$$

donde

- n      número de años del registro  
T      período de retorno en años

- c) Se obtiene el promedio de la relación  $q_{2.33}/q_{10}$  para toda la región. Además se vuelve a calcular el gasto de 10 años de período de retorno despejándolo a partir de la relación promedio. Entonces se asume que el 95% de los valores se comportan como una distribución normal (Spiegel, 1970) y por lo tanto caerán en el intervalo  $\pm 2\sigma_y$  del valor de T.
- d) Dado que la prueba se basa en la creciente con un período de retorno de 10 años, se tiene:

$$2\sigma_y = \frac{0.666e^y}{\sqrt{n}} \quad (2.44)$$

De esta manera, el valor de la variable reducida tipo I para T=10 años es 2.25 y los límites de confianza estarán dados por:

$$2.25 \pm 6.33/\sqrt{n} \quad (2.45)$$

En la siguiente tabla se muestran los límites de confianza superior e inferior con sus correspondientes periodos de retorno, para diferentes tamaños de muestras.

**Intervalos de confianza para la prueba de homogeneidad de Langbein**

No. de años (n)	Límite inferior		Límite superior	
	$y - 2\sigma_y$	$T_L$	$y + 2\sigma_y$	$T_U$
5	-0.59	1.2	5.09	160
10	0.25	1.8	4.25	70
20	0.83	2.8	3.67	40
50	1.35	4.4	3.15	24
100	1.62	5.6	2.88	18
200	1.80	6.5	2.70	15
500	1.97	7.7	2.53	13
1000	2.05	8.3	2.45	12

e) Finalmente, cada estación que va a ser probada puede ser localizada en una gráfica como la que se muestra en la figura 4 a través de sus coordenadas (T, n). Si alguna estación queda fuera de los límites de confianza es excluida de la región homogénea.

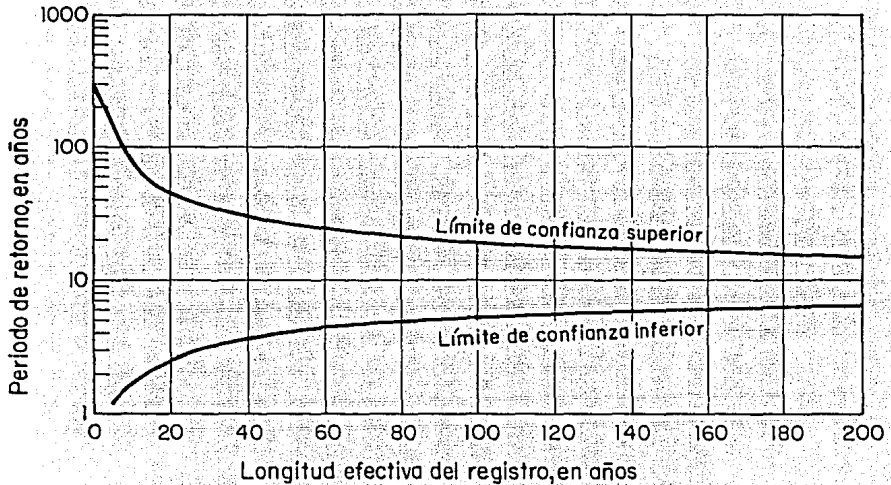


FIGURA 4 LÍMITES DE CONFIANZA PARA LA PRUEBA DE HOMOGENEIDAD DE LANGBEIN

El programa LANGBEIN.BAS desarrollado para esta prueba de homogeneidad, se presenta en el anexo 1.

## 2.5 Selección del análisis de grupo más apropiado

La primera etapa en la identificación de regiones homogéneas es intentar un análisis por grupos, empleando para ello, diferentes medidas de similitud y métodos que puedan ir *ligando* cada una de las cuencas en estudio. Muchas de las medidas de similitud y métodos de liga, han sido estudiados por numerosos investigadores para cada una de las variables de agrupación que ellos consideran como significativas. Una evaluación de cada uno de los procesos de agrupación puede producir tres o más diferentes pero uniformes grupos.

a) Se propone en primer lugar, el recurso gráfico como una técnica de selección preliminar de la más apropiada partición. Para ésto pueden dibujarse los CV contra (QBAR/AREA) tal y como se expuso en el apartado 2.3. Sin embargo, al final de este proceso algunas cuencas pueden quedar fuera de los grupos, lo que originaría una colocación arbitraria.

Otra alternativa es emplear las curvas de Andrews. Es claro que como esta prueba se basa en la investigación de cuáles cuencas son localizadas más frecuentemente y dentro de qué grupos, se espera que éste pueda ser un ejercicio relativamente directo (pero tedioso). Sin embargo, la ventaja de este método, es que prevé el emplear los criterios (SS) y (CS) como métodos de liga para agrupar estaciones cuya localización sea dudosa, a manera de un algoritmo por racimos.

b) Con los grupos que se obtienen de éste primer análisis gráfico, debe probarse la homogeneidad de cada uno *de sus miembros*, ésto, a través de un análisis de varianza. El estadístico  $F$  debe evaluarse a través de las expresiones propuestas por Wiltshire. Y reubicar las estaciones que sean necesarias, hasta lograr una completa homogeneidad dentro de los grupos.

c) A continuación deben identificarse las diferencias (si las hay) *entre los grupos* formados. Para lo cual, se propone un análisis de varianza en dos direcciones (capítulo 3), en donde se podrán detectar y reubicar estaciones que si bien son homogéneas, el colocarlas en otro grupo, ayudarán a que éste sea completamente homogéneo y a su vez sea significativamente diferente a los demás grupos. En esta etapa se prevé garantizar, que los elementos removidos del grupo, no eran significativos.

d) Finalmente, el mejor método debe ser el que mida las propiedades de una serie de gastos dentro de un grupo variable de cuencas, así como la completa selección de las características fisiográficas más relevantes. Es por eso que se propone una metodología con base en la combinación de varios criterios de homogeneidad, sin embargo, en un principio y en forma individual cada método deberá arrojar un número similar de grupos.

### 3. ANALISIS DE VARIANZA

Como ya se ha comentado, el análisis de varianza (ANOVA) es la prueba estadística que sirve para demostrar si dos variancias muestrales provienen de la misma población. Este concepto es muy útil en la regionalización, ya que permite discriminar cuáles variables son significativas en un modelo de regresión, también puede determinarse, si dos o más poblaciones son iguales, una vez que son comparadas simultáneamente sus medias poblacionales. En estos dos casos, las poblaciones analizadas deben tener una distribución normal. Esta prueba se basa en que el estadístico de prueba sigue la distribución  $F$ .

#### 3.1 Distribución $F$

El estadístico  $F$  se utiliza como una prueba en varias situaciones. Las principales características de esta distribución son:

1. Existe una familia de distribuciones  $F$ . Un elemento específico de la familia está determinado por los dos parámetros que caracterizan a la distribución, los grados de libertad (g.l.) en el numerador y en el denominador (figura 5). Obsérvese que la forma de las curvas varía a medida que cambian los grados de libertad.



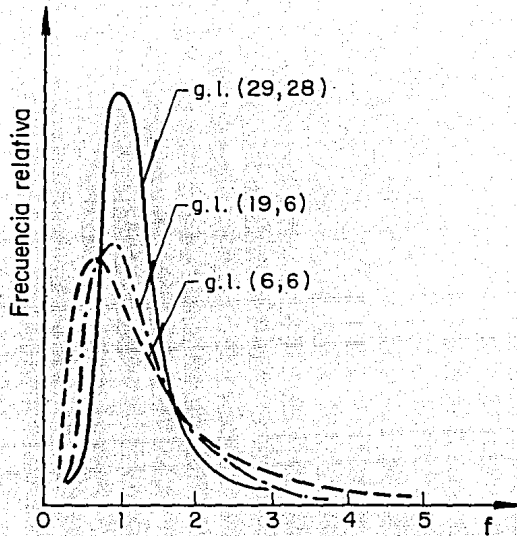


FIGURA 5 DISTRIBUCIÓN  $F$  PARA DIFERENTES GRADOS DE LIBERTAD

2. El valor de  $F$  no puede ser negativo.
3. La distribución  $F$  es una distribución continua.
4. La curva que representa una distribución  $F$  tiene un sesgo positivo.
5. Sus valores varían de 0 a  $\infty$  y tiene un comportamiento asintótico sobre el eje  $F$ .

La función  $F$  con  $\gamma_1 = m$  y  $\gamma_2 = n$  grados de libertad ( $m$  y  $n$  son conocidos como numerador y denominador, respectivamente) está dada por:

$$P_F(f) = \frac{\Gamma\left(\frac{\gamma_1 + \gamma_2}{2}\right) \gamma_1^{\gamma_1/2} \gamma_2^{\gamma_2/2} f^{\gamma_1/2} (\gamma_2 + \gamma_1 f)^{-(\gamma_1 + \gamma_2)/2}}{\Gamma(\gamma_1/2) \Gamma(\gamma_2/2)} \quad (3.1)$$

La media y la varianza de  $F$  serán:

$$E(F) = \frac{\gamma_1}{\gamma_2 - 2} \quad (3.2)$$

$$\text{Var}(F) = \gamma_2^2(\gamma_1 + 2) / [\gamma_1(\gamma_2 - 2)(\gamma_2 - 4)] \quad (3.3)$$

Las tablas 3.1 y 3.2 muestran los valores de  $F$  en función de los diferentes grados de libertad del numerador ( $m$ ) y denominador ( $n$ ) para valores de  $P_f(f)=0.95$  y  $0.99$

### 3.1.1 Comparación de dos varianzas poblacionales

La distribución  $F$  se utiliza para demostrar la hipótesis de que la varianza de una población normal es igual a la varianza de otra población normal. Así, la prueba es útil para determinar si una población tiene o no más variación que otra. La prueba  $F$  también puede usarse para validar supuestos con respecto a ciertas pruebas estadísticas, cabe recordar que independientemente de esto, primero se plantea la hipótesis que se quiere rechazar ( $H_0$ : hipótesis nula).

Para comparar varianzas poblacionales, la hipótesis nula es aquella en la que la varianza de una población normal,  $\sigma_1^2$  es igual a la varianza de otra población normal,  $\sigma_2^2$ . Para realizar la prueba, en una muestra aleatoria de  $N_1$  observaciones, y una muestra de  $N_2$  de una segunda población. El estadístico de prueba es  $S_1^2/S_2^2$ , en donde  $S_1^2$  y  $S_2^2$  son las varianzas muestrales respectivas.

Si la hipótesis nula es verdadera ( $H_0: \sigma_1^2 = \sigma_2^2$ ), el estadístico de prueba sigue la distribución  $F$  con  $N_1-1$  y  $N_2-1$  grados de libertad. La varianza muestral más grande se coloca en el numerador; en consecuencia,  $F$  es mayor que 1.0 y el valor crítico del estadístico  $F$  se obtiene con el nivel de significancia ( $\alpha$ ), buscando después el número adecuado de grados de libertad, en una tabla de valores críticos de  $F$  para diferentes niveles de significación y diferentes grados de libertad del numerador.

### 3.2 Noción general del análisis de varianza

El segundo uso de la distribución  $F$  comprende la técnica de análisis de varianza. Básicamente, en este análisis se emplea información muestral para determinar si tres o más causas de variación en un conjunto de datos, producen o no resultados diferentes. El uso de la palabra causa de variación es comúnmente llamado "tratamiento" y estará representado por las muestras que se estén analizando. Antes de realizar una prueba utilizando las técnicas ANOVA se examinarán las consideraciones en que se basa la prueba.

1. Las poblaciones de interés deben estar distribuidas normalmente
2. Tales poblaciones tiene desviaciones estandar iguales:  $S_1 = S_2$
3. Las muestras que se seleccionan de cada una de las poblaciones son aleatorias e independientes, es decir, no estan relacionadas entre si.

El procedimiento ANOVA, puede ilustrarse de manera más adecuada mediante los pasos siguientes:

- **HIPOTESIS NULA.**  $H_0$  expresa que no hay diferencias significativas entre las medias poblacionales; es decir  $\mu_1 = \mu_2 = \dots$ . La hipótesis alternativa denotada como  $H_1$  plantea que al menos una media es diferente. En la prueba, si  $H_0$  se rechaza,  $H_1$  se aceptará.
- **NIVEL DE SIGNIFICANCIA.** Se selecciona el nivel, por ejemplo al 5% ( $\alpha=0.05$ )
- **ESTADISTICO DE PRUEBA.** El estadístico de prueba adecuado es la distribución  $F$ . Este procedimiento se basa en las consideraciones antes mencionadas. De esta manera,  $F$  es la razón de dos varianzas.

$$F = \frac{\text{Varianza poblacional estimada según variación entre medias muestrales}}{\text{Varianza poblacional estimada según la variación en las muestras}}$$

La terminología común para el numerador es "varianza entre muestras", e "inter o intravarianza". Para el denominador es "varianza en las muestras". El numerador tiene  $k-1$  grados de libertad (g.l.). El denominador tiene  $N-k-1$ , donde  $k$  es el número de tratamientos o columnas y  $N$  es el número de observaciones.

- **REGLA DE DECISION.** Como se observó con anterioridad, la distribución  $F$  y la curva correspondiente tienen sesgo positivo y dependen, del número de tratamientos  $k$  y del número total de observaciones  $N$ .

El valor crítico, esto es, el punto divisorio entre la región de aceptación y la de rechazo, se puede obtener consultando las tablas 3.1 y 3.2 para los niveles 0.05 y 0.01. Los grados de libertad para el denominador se encuentran en la columna de la izquierda. Por ejemplo, para 2 g.l. en el numerador y 12 g.l. en el denominador, el valor crítico para el nivel de significancia 0.05 se obtendrá recorriendo horizontalmente a 2 grados de libertad en el numerador, después hacia abajo en las columnas hasta llegar al número opuesto a 12 g.l. en la columna de la izquierda. Este número es 3.89 y es el valor crítico de  $F$  para el nivel 0.05.

- **CALCULAR F Y TOMAR UNA DECISION.** Este paso consiste en construir una tabla ANOVA. Esta es sólo una forma conveniente de registrar la suma de cuadrados y otros cálculos. El formato general para un problema de análisis de varianza en un sentido se muestra en la tabla siguiente.

FORMATO GENERAL DEL ANALISIS DE VARIANZA

FUENTE DE VARIACION	SUMA DE CUADRADOS	GRADOS DE LIBERTAD	CUADRADO MEDIO
Entre tratamientos	SST	k-1	SST/k-1
Error (en los tratamientos)	SSE	N-k	SSE/N-k
Total	Total SS		

donde

$$MSTR = SST/k-1$$

$$MSE = SSE/N-k$$

SS

cuadrado medio entre tratamientos

cuadrado medio debido al error, también se denomina cuadrado medio dentro de tratamientos

abreviatura de tratamiento de suma de cuadrados, y se obtiene mediante:

$$SST = \sum \left( \frac{X^2}{n} \right) - \frac{(\sum X)^2}{N} \tag{3.4}$$

$$SSE = \sum (X^2) - \sum \left( \frac{X^2}{n} \right) \tag{3.5}$$

$$F = \frac{SST/k-1}{SSE/N-k} = \frac{MSTR}{MSE} \tag{3.6}$$

donde

- $X^2$  indica elevar al cuadrado el total de cada columna
- $n$  número de observaciones para cada tratamiento (columna)
- $\sum(X)$  suma de todas las observaciones
- $k$  número de tratamientos (columnas)
- $N$  número total de observaciones

La variación total (Total SS) es la suma de la variación entre columnas y entre renglones; es decir, Total SS = SST + SSE

La regla de decisión indica que si el valor calculado de  $F$  es menor o igual al valor crítico, la hipótesis nula se acepta. Si el valor de  $F$  es mayor que el crítico,  $H_0$  se rechaza y  $H_1$  se acepta. Con base en este esquema se desarrolló el programa ANOVA.BAS con el fin de emplearlo más adelante en la discriminación de variables, el listado puede consultarse en el anexo 1.

### 3.3 Análisis de varianza en dos sentidos

El análisis de varianza en dos sentidos permite considerar no solo la posibilidad de que exista variación en las medias de los tratamientos, (columnas) sino tomar en cuenta una posible variación dentro de los elementos de cada uno de ellos (bloques). De esta manera, existirán dos conjuntos de hipótesis que se prueban:

1.  $H_0$ : Las medias de tratamientos son iguales  
 $H_1$ : Las medias de tratamientos no son iguales

2.  $H_0$ : Las medias de bloques son iguales  
 $H_1$ : Las medias de bloques no son iguales

El mismo formato que sirve para el caso de ANOVA en un sentido se utiliza para la tabla ANOVA en dos sentidos. Análogo a la tabla ANOVA, el formato general en dos sentidos es:

#### FORMATO GENERAL DEL ANALISIS DE VARIANZA EN DOS SENTIDOS

FUENTE DE VARIACION	SUMA DE CUADRADOS	GRADOS DE LIBERTAD	CUADRADO MEDIO
Tratamientos	SST	k-1	SST/k-1
Bloques	SSB	n-1	SSB/n-1
Error	SSE	(k-1)(n-1)	SSE/(k-1)(n-1)
Total	SS		

donde

$$SS=SST+SSB+SSE$$

$$MSTR=SST/k-1 \quad \text{cuadrado medio entre tratamientos}$$

$$MSE=SSE/(k-1)(n-1) \quad \text{cuadrado medio debido al error, en los bloques}$$

$$MSB=SSB/n-1 \quad \text{cuadrado medio según bloques y se obtiene mediante:}$$

$$SSB = \sum \left( \frac{X_B^2}{k} \right) - \left[ \frac{(\sum X)^2}{N} \right] \quad (3.7)$$

donde

$X_B$  se refiere al total del bloque, es decir, al total de cada renglón  
 $k$  número de elementos de cada bloque

Así, el estadístico  $F$  entre tratamientos estará dado como:

$$F = \frac{SST/k-1}{SSE/(k-1)(n-1)} = \frac{MSTR}{MSE} \quad (3.8)$$

Para  $F$  entre bloques, se tiene:

$$F = \frac{SSB/n-1}{SSE/(k-1)(n-1)} = \frac{MSB}{MSE} \quad (3.9)$$

Finalmente, al igual que el análisis en un sentido la regla de decisión indica que si el valor calculado de  $F$  es menor o igual al valor crítico, la hipótesis nula se acepta. Si el valor de  $F$  es mayor que el crítico,  $H_0$  se rechaza y  $H_1$  se acepta. El programa ANOVA2.BAS desarrollado para este caso puede consultarse en el anexo 1.

**TABLA 3.1**  
**VALORES CRITICOS DE LA DISTRIBUCION F AL NIVEL DE SIGNIFICANCIA DE 5%**

↓	Grados de libertad del numerador																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	10.1	9.6	9.3	9.1	9.0	8.9	8.9	8.9	8.8	8.8	8.7	8.7	8.7	8.6	8.6	8.6	8.6	8.6	8.5
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.81	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.81	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Fuente: Esta tabla se ha tomado de M. Merrington y C.M. Thompson, "Tables of Percentage Points of the Inverted Beta (F) Distribution", Biometrika, Vol. 33

**TABLA 3.2**  
**VALORES CRITICOS DE LA DISTRIBUCION F AL NIVEL DE SIGNIFICANCIA DE 1%**

Grados de libertad del denominador	Grados de libertad del numerador																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4052	5000	5403	5625	5764	5859	5928	5982	6023	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.5	99.0	99.2	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	26.9	26.7	26.6	26.5	26.4	26.3	26.2	26.1
4	21.20	18.00	16.70	16.00	15.50	15.20	15.00	14.80	14.70	14.50	14.40	14.20	14.00	13.90	13.80	13.70	13.70	13.60	13.50
5	16.30	13.30	12.10	11.40	11.00	10.70	10.50	10.30	10.20	10.10	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.70	10.90	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.20	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.30	8.65	7.59	7.01	6.83	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.60	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.00	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.50	5.04	4.70	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.19	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.53	2.45	2.36	2.27	2.17
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Fuente: Esta tabla se ha tomado de M. Merrington y C.M. Thompson, "Tables of Percentage Points of the Inverted Beta (F) Distribution", Biometrika, Vol. 33



#### **4. ANALISIS DE CORRELACION Y REGRESION**

El análisis de correlación es un grupo de técnicas estadísticas empleadas para medir la intensidad de la relación (correlación) entre dos o más variables. A la expresión matemática que define esta relación se le llama ecuación de regresión. El estudio de la correlación es muy empleado en la Hidrología, ya que se puede determinar una ecuación que relacione una variable por ejemplo, una característica fisiográfica; con otra variable que se quiere predecir por ejemplo, los gastos máximos, o los parámetros de una distribución de probabilidad. También, de esta manera, podrá realizarse una transferencia de información entre estaciones, ya cuando éstas presenten una buena correlación; una ecuación de regresión, podrá inferir los valores en la estación de menor registro.

Las principales consideraciones para llevar a cabo una regresión lineal, pueden resumirse en:

1. Para cada valor de  $X$  existe un grupo de valores  $Y$ .  $X$  y  $Y$  se distribuyen en forma normal.
2. Las medias de estas distribuciones normales de valores  $Y$  se encuentran todas en la recta ajustada por la regresión.
3. Las desviaciones estándares de dichas distribuciones normales son iguales.

4. Los valores  $Y$  son estadísticamente independientes. Esto significa que al seleccionar una muestra, los valores  $Y$  seleccionados para un valor  $X$  específico no dependen de los valores  $Y$  para cualquier otro valor  $X$ . Las observaciones sucesivas de la variable dependiente no deben estar correlacionadas. Si tal consideración no se cumple, la situación se denomina autocorrelación. La autocorrelación ocurre con frecuencia cuando se recopilan datos durante periodos o intervalos de tiempo.
5. La variación en la diferencia entre los valores real y pronosticado debe ser la misma para todos los valores ajustados de  $Y$ . Los residuos calculados deben estar distribuidos en forma normal con media cero.
6. Las variables independientes no deben estar muy correlacionadas. Cuando las variables independientes sí lo están, la situación se conoce como multicolinealidad.

La consideración de que las muestras deben estar distribuidas en forma normal, a menudo se pasa por alto, sin embargo, es de gran importancia ya que pueden afectarse en forma significativa los resultados del análisis de correlación y regresión. Por eso, a continuación se presenta una serie de técnicas encaminadas a normalizar muestras.

#### 4.1 Procesos de normalización

Dado que la teoría de correlación y regresión simple o múltiple, esta basada en que la distribución de las variables analizadas es normal, se deberá garantizar esta condición antes de realizar cualquier proceso de transferencia de información mediante la aplicación de esta teoría.

Al emplear procesos de regresión y correlación, debe asegurarse que los registros del fenómeno sigan una distribución normal para lo cual deben emplearse procesos denominados de normalización, los cuales transforman una serie de registros no normales a una serie normal. Así, los valores de la serie " $y$ " (nombrados 1,2,3...), corresponderán a " $n$ " valores en donde el valor más pequeño corresponde a  $y=1$ , el siguiente es  $y=2$ , etc. Por lo tanto, si reemplazamos  $y=1$  por el valor de la observación más pequeña de una muestra con " $n$ " valores de una población normal, entonces estaremos más cerca de una distribución normal.

El valor de la  $i$ -ésima observación más pequeña de la muestra de " $n$ " valores de la población normal será:

$$Z_i = \frac{n!}{(i-1)!(n-i)!} \int_{-\infty}^{\infty} X \frac{1}{\sqrt{2\pi}} e^{-x^2/2} [\Phi(X)]^{i-1} [1-\Phi(X)]^{n-i} dX \quad (4.1)$$

donde

$$\Phi(X) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

Por lo tanto, la serie será normalizada al remplazar  $y=i$  por  $Z_i$ . Ya existen tablas calculadas de  $Z_i$ , (Pearson y Hartley, 1972) presentaron una tabla con base en la transformación  $Z_i$ .

A continuación se muestra una tabla con las transformaciones sugeridas, según sea la distribución de la muestra original  $Y_i$

DISTRIBUCION ORIGINAL DE Y	VARIANZA EN TERMINOS DE SU MEDIA $\mu$	TRANSFORMACION	VALOR DE LA VARIANZA TRANSFORMADA
POISSON	$\mu$	$\sqrt{y}$	0.25
BINOMIAL	$\mu(1-\mu)/n$	$\text{sen}^{-1} \sqrt{y}$	1/4n
BINOMIAL (-)	$\mu + a\mu^2$	$\frac{1}{\sqrt{a}} \text{senh}^{-1} \sqrt{ay}$	0.25
EMPIRICA	$a\mu$	$\sqrt{y}$	1.4a
EMPIRICA	$a\mu^2$	$\log_e y$	a
		$\log_{10} y$	0.189a
EMPIRICA	$b\mu + a\mu^2$	$\frac{1}{\sqrt{a}} \text{senh}^{-1} \sqrt{ay}$	0.25b

Otra formulación de las transformaciones normalizantes, fue sugerida por (Johnson, 1949). El uso de estas transformaciones por lo general requiere un conocimiento de la media y varianza de "y", además de los parámetros  $\sqrt{b_1}$ , y  $b_2$ .

Así, si se tienen "n" valores de "y", denominados como  $y_1, y_2, \dots, y_n$  entonces, la media, la varianza y los parámetros mencionados, podrán calcularse con las expresiones siguientes:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (4.2)$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 \quad (4.3)$$

$$\sqrt{b_1} = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^3}{\left[ \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \right]^{3/2}} \quad (4.4)$$

$$b_2 = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^4}{\left[ \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \right]^2} \quad (4.5)$$

Se definen en estos términos las 3 transformaciones siguientes:

$$\text{SB: } Z = \gamma + \delta \log \left[ \frac{x}{1-x} \right] \quad (0 < x < 1) \quad (4.6)$$

$$\text{SU: } Z = \gamma + \delta \sinh^{-1}(x) \quad (-\infty < x < \infty) \quad (4.7)$$

$$\text{SL: } Z = \gamma + \delta \log(x) \quad (0 < x < \infty) \quad (4.8)$$

donde  $x = (y - \zeta)/\lambda$  y  $\gamma$ ,  $\delta$ ,  $\zeta$ , y  $\lambda$  son valores desconocidos

Estos valores desconocidos se estiman usando  $\bar{y}$ ,  $S^2$ ,  $\sqrt{b_1}$ , y  $b_2$ . la transformación particular que mejor resuelve el problema se escoge en base de los valores de  $\sqrt{b_1}$ , y  $b_2$ .

Los metodos de estimación de  $\gamma$ ,  $\delta$ ,  $\zeta$ , y  $\lambda$  fueron desarrollados por Pearson y Hartley en (1971), en su trabajo pueden consultarse las tablas 34 y 35 para el cálculo de los parámetros mencionados.

#### 4.1.1 Transformación de Box-Cox

Box y Cox (1964) sugirieron las siguientes transformaciones, para ayudar a satisfacer los requerimientos de normalidad, varianza cosntante y aditividad. Así una serie de datos " $y_i$ " podrá ser transformada a una serie  $Z_i$  a través de las expresiones siguientes:

$$Z = \frac{y^\lambda - 1}{\lambda} \quad (\lambda \neq 0) \quad (4.9)$$

$$Z = \log_e y \quad (\lambda = 0) \quad (4.10)$$

El procedimiento sugerido, es escoger un valor particular para  $\lambda$  (parámetro de transformación) y calcular la varianza  $\hat{\sigma}^2$  usando la variable transformada "Z" como variable dependiente. Así, la función logarítmica de  $\lambda$  se calcula como:

$$L_b(\lambda) = (n - k - 1) \left[ \left( \frac{\lambda - 1}{n} \right) \sum_{i=1}^n \log_e y_i - \frac{1}{2} \log_e \hat{\sigma}^2 \right] \quad (4.11)$$

Variando  $\lambda$  y repitiendo el procedimiento arriba mencionado, puede trazarse una gráfica  $L_b(\lambda)$  contra  $\lambda$  tal y como se muestra en la figura 6.

La mejor elección para  $\lambda$  es tomar el valor que maximice  $L_b(\lambda)$ , sin embargo, si  $\lambda=0$  es el valor que la maximiza, entonces  $\log_e y$  se toma como la mejor transformación. Pueden existir algunos problemas con estas transformaciones si "y" es negativa, una transformación alternativa puede ser:

$$Z = \frac{(y + \lambda_2)^{\lambda_1} - 1}{\lambda_1} \quad (\lambda_1 \neq 0) \quad (4.12)$$

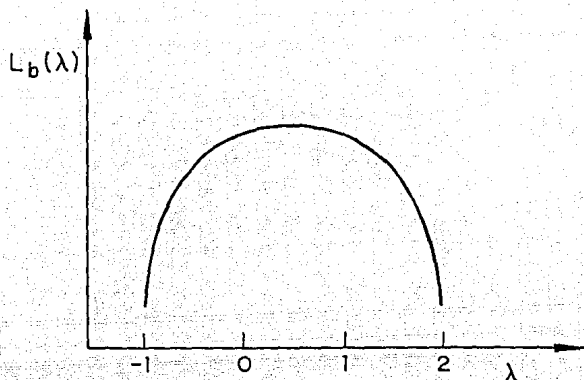
$$Z = \log_e(y + \lambda_2) \quad (\lambda_1 = 0) \quad (4.13)$$

Para resolver este problema, con este procedimiento se calcula una vez más la varianza de los residuales  $\hat{\sigma}^2$  con los valores específicos de  $\lambda_1$  y  $\lambda_2$  y entonces se evalúa:

$$L_b(\lambda_1, \lambda_2) = (n - k - 1) \left[ \left( \frac{\lambda_1 - 1}{n} \right) \sum_{i=1}^n \log_e (y_i + \lambda_2) - \frac{1}{2} \log_e \hat{\sigma}^2 \right] \quad (4.14)$$

El problema ahora es escoger  $\lambda_1$  y  $\lambda_2$  para maximizar  $L_b(\lambda_1, \lambda_2)$  esto se puede realizar tanto gráficamente como usando algunos logaritmos numéricos para maximizar.

Las transformaciones no solo se usan para las variables dependientes. Por ejemplo, será apropiado transformar  $x_1, x_2, \dots, x_k$  a una serie del tipo  $x_1^1, x_2^2, \dots, x_k^k$ .



**FIGURA 6 TRANSFORMACION DE BOX-COX**

Supongamos que la transformación correcta para  $x_1$ , pudiera aproximarse a  $x_1^{r_1^*}$  ( $r_1^*$  podría ser 1 como una suposición preliminar). Entonces una expansión a través de las series Taylor sería:

$$x_1^{r_1} = x_1^{r_1^*} + (r_1 - r_1^*) x_1^{r_1^*} \log_e x_1 + \dots \quad (4.15)$$

Así si las variables  $x_1^{r_1^*}$  y  $x_1^{r_1^*} \log_e x_1$  pueden incluirse en el modelo de regresión, entonces los coeficientes estimados de regresión de esas variables, se usan para calcular y estimar el valor correcto de  $x_1$ . Esta estimación de  $r_1$ , podrá estar dada por:

$$r_1^* = \frac{\text{coeficientes estimados por la regresión de } x_1^{r_1^*} \log_e x_1}{\text{coeficientes estimados por la regresión de } x_1^{r_1^*}}$$

Esta estimación para  $r_1$ , puede ser sustituida por  $x_1^{r_1^*}$  la regresión puede entonces repetirse para obtener una nueva estimación para  $r_1$ , y así sucesivamente incluyendo los términos  $x_1^{r_1^1}, x_2^{r_2^2}, \dots, x_k^{r_k^k}$  y  $x_1^{r_1^1} \log_e x_1, x_2^{r_2^2} \log_e x_2, \dots, x_k^{r_k^k} \log_e x_k$

#### 4.1.2 Transformación Box-Cox de doble potencia

La fuerza de la transformación original de Box-Cox, considerada de 1ª potencia para normalizar muestras según la ecuación (4.9) ha sido cuestionada, respecto a que, si bien produce una serie  $Z_i$  cuyo histograma llega a ser simétrico en el eje vertical pasando por la media, no produce una serie completamente normal. Este histograma transformado será normal solo cuando el coeficiente de asimetría se reduce a un valor de cero y el valor de la curtosis sea 3.

La corrección de la curtosis puede lograrse a través de otra transformación denominada: *transformación de doble potencia* (Two Step Power Transformation, TSPT)

$$t_i = \left( |y_i - \bar{y}| \right)^\gamma \quad (4.16)$$

donde

$\gamma$  parámetro de transformación para TSPT de signo positivo

$t_i$  i-ésimo elemento de la serie transformada TSPT y tiene el mismo signo de  $(y_i - \bar{y})$

$\bar{y}$  media de la serie transformada por Box-Cox (ecuación 4.9)

Es evidente que cuando  $\gamma \rightarrow 0$  la curtosis de la serie transformada tiende a uno y cuando  $\gamma \rightarrow \infty$  también lo hace la curtosis. Por lo tanto, si la curtosis de la serie normalizada por 1ª la potencia es mayor que 3, el parámetro de transformación estará entre 0 y 1, y si la curtosis es menor que 3,  $\gamma$  será mayor que uno.

Cuando una serie es transformada por TSPT a una serie normal, el valor correspondiente para un período de retorno específico puede obtenerse con la siguiente ecuación:

$$t_p = \bar{t} + k_p \sigma_t \quad (4.17)$$

donde

$t_p$  corresponde al "p" período de retorno en la serie transformada

$k_p$  desviación normal estándar correspondiente al mismo período de retorno

$\sigma_t$  desviación estándar de la serie TSPT transformada

El valor correspondiente a  $t_p$  en la serie original  $x_p$ , se puede obtener usando las expresiones siguientes (desnormalización):

$$x_p = (y_p^{\lambda+1})^{1/\lambda} \quad (4.18)$$

donde

$$y_p = y'_p + \bar{y} \quad (4.19)$$

y

$$y'_p = (|t_p|)^{1/\gamma} \quad (4.20)$$

donde

$y'_p$  tiene el mismo signo que  $t_p$

$x_p$  corresponde al "p" periodo de retorno de la serie original

$y_p$  corresponde al "p" periodo de retorno de la serie transformada

Finalmente cabe mencionar que este método de transformación, es empleado simultáneamente con el de Box-Cox. Con dicho método el coeficiente de asimetría se reduce a cero o cercano a cero, obteniendo una serie transformada. Esta serie se sustituye en la expresión de TSPT y con valores adecuados de  $\gamma$  se consigue un coeficiente de curtosis igual a 3. Esta transformación es la mejor y más nueva que se conoce, por lo cual es ampliamente recomendada cuando las muestras deben ser 100% normales, sin embargo la transformación original de Box-Cox es muy fácil de programar y produce valores de la curtosis en la mayoría de los casos entre 2.0 y 4.5, y de cero invariablemente para el coeficiente de asimetría.

## 4.2 Correlación simple

Posiblemente el más común de los modelos empleados en la Hidrología se basa en considerar una relación lineal entre dos variables. Generalmente, el objetivo de tal modelo es proporcionar una forma de predecir o estimar una variable dependiente a partir de la segunda variable (conocida).

$$Y = a + bX \quad (4.21)$$

La línea  $Y = a + bX$  es llamada comúnmente como recta de regresión de  $Y$  en  $X$ . La palabra regresión la usó por primera vez Sir Francis Galton en 1877 en su estudio de los factores hereditarios. Descubrió que las estaturas de los descendientes de padres altos tendían a una *regresión* (es decir, a volver o retornar) hacia la estatura promedio de la población. A la recta matemática desarrollada se le denominó línea o recta de regresión. El término *recta de regresión* es de uso común aunque sería más adecuado llamarla *ecuación de pronóstico* o *ecuación de estimación*.



Al procedimiento para determinar a y b se le llama *regresión simple*. El término "simple" hace referencia cuando solo existe una variable independiente. Cuando existe más de una, se le llama *regresión múltiple* y los parámetros estimados son conocidos como *coeficientes de regresión*.

Así, podemos decir que  $e_i$  representa la desviación entre los valores  $Y_i$  observados y los inferidos  $\hat{Y}_i$ , basados en la ecuación de regresión, pero sin conocer qué tanto se relacionan Y y X.

Considerando ahora la pregunta: ¿puede, la ecuación de regresión representar adecuadamente los datos?. Naturalmente, la respuesta debe ser atendida al definir el término "adecuado". Una posibilidad es determinar qué tanto de la variable dependiente, puede ser explicada por la regresión. Esto puede medirse de la forma siguiente:

$$Y_i - \hat{Y}_i = (Y_i - \bar{Y}) - (\hat{Y}_i - \bar{Y})$$

Elevando al cuadrado ambos miembros y sumando todas las observaciones se tiene:

$$\sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - \bar{Y})^2 - 2\sum (Y_i - \bar{Y})(\hat{Y}_i - \bar{Y}) + \sum (\hat{Y}_i - \bar{Y})^2 \quad (4.22)$$

De la ecuación;  $a = \bar{Y} - b\bar{X}$  y como  $(\hat{Y}_i - \bar{Y})$  puede escribirse como:

$$(\hat{Y}_i - \bar{Y}) = b(X_i - \bar{X})$$

sustituyendo se tiene:

$$-2\sum (Y_i - \bar{Y})(\hat{Y}_i - \bar{Y}) = -2b\sum (Y_i - \bar{Y})(X_i - \bar{X}) \quad (4.23)$$

$$\sum (Y_i - \bar{Y})(X_i - \bar{X}) = b\sum (X_i - \bar{X})^2$$

sustituyendo en (4.23)

$$-2\sum (Y_i - \bar{Y})(\hat{Y}_i - \bar{Y}) = -2b \left[ b\sum (X_i - \bar{X})^2 \right]$$

o

$$-2\sum (Y_i - \bar{Y})(\hat{Y}_i - \bar{Y}) = -2b^2\sum (X_i - \bar{X})^2 \quad (4.24)$$

que sustituido en la ecuación (4.22) se tiene:

$$\sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - \bar{Y})^2 - \sum (\hat{Y}_i - \bar{Y})^2$$

reagrupando términos

$$\sum (Y_i - \bar{Y})^2 = \sum (Y_i - \hat{Y}_i)^2 + \sum (\hat{Y}_i - \bar{Y})^2$$

Sin embargo,  $\sum (Y_i - \bar{Y})^2$  es igual a  $\sum Y_i^2 - n\bar{Y}^2$  y se tiene:

$$\sum Y_i^2 = n\bar{Y}^2 + \sum (Y_i - \hat{Y}_i)^2 + \sum (\hat{Y}_i - \bar{Y})^2 \quad (4.25)$$

De esta manera la suma de los cuadrados de  $\sum Y_i^2$  puede dividirse en tres componentes:

1.  $n\bar{Y}^2$  la suma de los cuadrados debidos a la media ( variación total)
2.  $\sum (Y_i - \hat{Y}_i)^2 = \sum e_i^2$  la suma del cuadrado de las desviaciones de la regresión o la suma del cuadrado de los residuales
3.  $\sum (\hat{Y}_i - \bar{Y})^2$  la suma de los cuadrados de la regresión (variación explicada por la regresión)

Así, la suma de cuadrados estará dada por:

$$\sum (Y_i - \bar{Y})^2 = \sum y_i^2 = \sum Y_i^2 - n\bar{Y}^2 = \sum (Y_i - \hat{Y}_i)^2 + \sum (\hat{Y}_i - \bar{Y})^2 \quad (4.26)$$

Empleando las ecuaciones (4.22) y (4.23)

$$\sum y_i^2 = \sum e_i^2 + b\sum x_i y_i$$

Esto quiere decir, que la suma total de los cuadrados esta formada por dos componentes: el error o variación no explicada y la variación explicada por la regresión. De esta manera, el llamado *coeficiente de determinación*  $r^2$  mide el porcentaje de la variación en Y que se explica por la variación de X. ( Coeficiente de no determinación,  $1 - r^2$  )

$$r^2 = \frac{\text{suma de cuadrados en la regresión}}{\text{suma de cuadrados debidos a la media}}$$

$$r^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum y_i^2} \quad (4.27)$$

de las ecuaciones (4.22) y (4.23) puede escribirse

$$r^2 = b \sum x_i y_i / \sum y_i^2 = (\sum x_i y_i)^2 / (\sum x_i^2 \sum y_i^2) \quad (4.28)$$

Si la ecuación de regresión predice perfectamente cada valor de  $Y_i$ ,  $e_i$  tenderá a ser cero; bajo esta condición  $\sum y_i^2 = \sum (\hat{Y}_i - \bar{Y})^2$ , el valor de  $r^2$  será uno. Si la ecuación no puede explicar las variaciones de  $Y$ ,  $\sum e_i^2 = \sum y_i^2$  esto quiere decir que  $\sum (\hat{Y}_i - \bar{Y})^2$  tomará el valor de cero. Bajo esta consideración el valor de  $r^2$  será cero.

Es decir, el rango posible de valores de  $r^2$  será entre cero y uno. Así, un valor de  $r^2$  igual a uno, indicará la mejor correlación entre  $Y$  y  $X$  y por lo tanto la ecuación de regresión más adecuada. Este coeficiente puede escribirse como:

$$r^2 = \frac{b \sum x_i y_i}{\sum y_i^2} = \frac{b^2 \sum x_i^2}{\sum y_i^2} = b^2 \frac{S_x^2}{S_y^2} \quad 0 \leq r^2 \leq 1 \quad (4.29)$$

$$r = b \frac{S_x}{S_y} \quad -1 \leq r \leq 1$$

El signo de "r" es el mismo que el de "b" siendo  $S_x$  y  $S_y$  siempre positivas.

*El coeficiente r es denominado de correlación* y mide la fuerza de asociación entre las dos variables. Comúnmente se escribe como:

$$r = \frac{\sum x_i y_i}{(\sum x_i^2 \sum y_i^2)^{1/2}} = \frac{S_{x,y}}{S_x S_y} \quad (4.30)$$

Cabe recordar que es necesario identificar las consideraciones para realizar una regresión lineal, porque si no se cumplen, los resultados podrían presentar un sesgo. Finalmente, a la raíz cuadrada positiva de  $\text{Var}(\varepsilon)$  se le conoce como *error estándar* de la regresión y puede calcularse a partir de los conceptos mencionados. Bajo el supuesto de que el valor esperado  $E(\varepsilon) = 0$  y así la  $\text{Var}(\varepsilon)$  estará dada por:

$$\text{Var}(\varepsilon) = E(\varepsilon^2) - E^2(\varepsilon) = E(\varepsilon^2) = \sigma^2$$

Según Graybill, (1961) determinó que  $\text{Var}(\varepsilon)$  puede calcularse como:

$$S^2 = \frac{\sum e_i^2}{(n-2)} = \frac{\sum (Y_i - \hat{Y}_i)^2}{(n-2)} \tag{4.31}$$

Si al estimar el error estándar de la regresión resulta un valor muy cercano o igual a la desviación estándar de  $Y$ ,  $r^2$  estará muy cerca de cero y por lo tanto, la ecuación de regresión no tendrá mucha capacidad de explicar el comportamiento de  $Y$ .

### 4.3 Correlación múltiple

Frecuentemente una variable denominada como dependiente, obtiene sus características de parámetros cuantitativos. Por ejemplo, el pico máximo en un escurrimiento, en cierta cuenca dada, estará relacionado con el área de la cuenca, la pendiente del cauce, la lluvia, etc. A continuación se extenderá el modelo visto en el primer punto de este capítulo, para que pueda incluir varias variables independientes, creando así, por ejemplo, un modelo que pueda predecir el pico máximo de un escurrimiento, y que contenga todas las variables antes mencionadas.

La forma general de un modelo lineal es:

$$Y = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \tag{4.32}$$

donde

- Y variable dependiente
- $X_1, X_2, \dots, X_p$  variables independientes
- $\beta_1, \beta_2, \dots, \beta_p$  parámetros no conocidos.

Este modelo es lineal en los parámetros  $\beta_1$ , un modelo del tipo:

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3^2 + \beta_4 \ln(X_4) \quad (4.33)$$

es también lineal respecto a los parámetros  $\beta_j$

Sin embargo los modelos:

$$Y = \beta_1 + \beta_2 X^{\beta_3}$$

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 \exp(\beta_4 X_3)$$

$$Y = \beta_1 + \beta_2 X_2 + \beta_2^2 X_2$$

son de parámetros no lineales

En la práctica, pueden obtenerse las  $n$  observaciones de  $Y$  con sus correspondientes  $n$  observaciones de cada una de las  $p$  variables independientes. De este modo, se pueden escribir  $n$  ecuaciones como la (4.32) una por cada observación. Esencialmente, se van a resolver  $n$  ecuaciones para los  $p$  parámetros desconocidos. Así que  $n$  podrá ser igual o mayor que  $p$ . En la práctica  $n$  debe ser entre 3 y 4 veces mayor que  $p$ . Las  $n$  ecuaciones, llamadas normales, son:

$$\begin{aligned} Y_1 &= \beta_1 X_{1,1} + \beta_2 X_{1,2} + \dots + \beta_p X_{1,p} \\ Y_2 &= \beta_1 X_{2,1} + \beta_2 X_{2,2} + \dots + \beta_p X_{2,p} \\ &\dots\dots\dots \\ Y_n &= \beta_1 X_{n,1} + \beta_2 X_{n,2} + \dots + \beta_p X_{n,p} \end{aligned} \quad (4.34)$$

donde

- $Y_j$        $i$ -ésima observación de  $Y$
- $X_{i,j}$      $i$ -ésima observación de la  $j$ -ésima variable independiente.

Esta ecuación también puede ser escrita como:

$$Y_i = \sum_{j=1}^p \beta_j X_{i,j} \quad (4.35)$$

desde  $i=1$  a  $n$

En notación matricial la ecuación (4.35) se transforma en:

$$\underline{Y} = \underline{X} \underline{\beta} \quad (4.36)$$

$\begin{matrix} \underline{Y} & \text{nx1} \\ \underline{X} & \text{nxp} \\ \underline{\beta} & \text{px1} \end{matrix}$

donde

$\underline{Y}$  vector de observaciones (n x 1)

$\underline{X}$  matriz de (nxp) creada con las n observaciones de cada uno de las p variables independientes

$\underline{\beta}$  vector de parámetros no conocidos de (px1)

De la forma matricial se obtiene

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} X_{1,1} & X_{1,2} & X_{1,3} & \cdots & X_{1,p} \\ X_{2,1} & X_{2,2} & X_{2,3} & \cdots & X_{2,p} \\ X_{3,1} & X_{3,2} & X_{3,3} & \cdots & X_{3,p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X_{n,1} & X_{n,2} & X_{n,3} & \cdots & X_{n,p} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_p \end{bmatrix} \quad (4.37)$$

En el siguiente desarrollo se considera que  $X_{i,1} = 1$  para  $i=1$  hasta n. El modelo que se discutió en el punto anterior

$$Y = \alpha + \beta X$$

es un caso especial de la ecuación (4.32) en donde  $X_{i,1} = 1$ ,  $X_{i,2} = X$ ,  $\beta_1 = \alpha$  y  $\beta_2 = \beta$ , siguiendo con este esquema, los parámetros no conocidos,  $\underline{\beta}$  pueden ser estimados por la minimización de:  $\sum e_i^2$  donde el término ( )' denota la matriz transpuesta.

$$\sum e_i^2 = \underline{e}' \underline{e} = (\underline{Y} - \underline{X}\hat{\underline{\beta}})' (\underline{Y} - \underline{X}\hat{\underline{\beta}}) \quad (4.38)$$

donde

$$e_i = Y_i - \hat{Y}_i = Y_i - \sum_{j=1}^p \hat{\beta}_j X_{i,j}$$

derivando la expresión (4.38) con respecto a  $\hat{\underline{\beta}}$  y haciendo la derivada parcial igual a cero se tiene:

$$0 = -2\underline{X}'(\underline{Y} - \underline{X}\hat{\underline{\beta}})$$

o

$$\underline{X}'\underline{Y} = \underline{X}'\underline{X}\hat{\underline{\beta}} \quad (4.39)$$

La solución de la ecuación (4.39) se obtiene por la premultiplicación de  $(\underline{X}'\underline{X})^{-1}$ :

$$\begin{aligned} (\underline{X}'\underline{X})^{-1} \underline{X}'\underline{Y} &= (\underline{X}'\underline{X})^{-1} (\underline{X}'\underline{X})\hat{\underline{\beta}} = \hat{\underline{\beta}} \\ \hat{\underline{\beta}} &= (\underline{X}'\underline{X})^{-1} \underline{X}'\underline{Y} \end{aligned} \quad (4.40)$$

La matriz  $\underline{X}'\underline{X}$  desempeña un papel importante en la estimación de  $\underline{\beta}$  y en las varianzas de los  $\hat{\beta}_i$ . La matriz  $\underline{X}'\underline{X}$  se construye con la suma de los cuadrados y el producto cruzado de las variables independientes. Para que la matriz  $\underline{X}'\underline{X}$  pueda ser invertida debe ser cuadrada y de orden  $p$  ( $p \times p$ )

Si definimos una matriz  $\underline{Z} = [Z_{i,j}]$  como:

$$Z_{i,j} = \frac{X_{i,j} - \bar{X}_j}{S_j}$$

entonces;  $\underline{Z}'\underline{Z} / (n-1)$  es una matriz de correlación de  $p \times p$ .  $\underline{R} = [R_{i,j}]$  donde  $R_{i,j}$  es el coeficiente de correlación entre la  $i$ -ésima y  $j$ -ésima variable independiente.

Por definición  $R_{i,j} = 1$  para  $i = j$ . Si  $|R_{i,j}| = 1$  para algunos " $i$ " diferentes de " $j$ " quiere decir que la  $i$ -ésima variable independiente es una función lineal de la  $j$ -ésima variable independiente y que el orden de  $\underline{X}'\underline{X}$  será menor que  $p$ .

Por lo tanto si la ecuación (4.39) se usa para estimar los " $p$ " parámetros no conocidos  $\underline{\beta}$ , podrá haber dependencia no lineal en  $\underline{X}$ . Esta medida de la variable independiente puede no ser una (perfecta) función lineal de alguna de las otras variables independientes. Además, el rango de  $\underline{X}'\underline{X}$  será  $p$ , una variable independiente podrá no tener dependencia lineal en alguna de las funciones lineales de las restantes variables independientes.

Por ejemplo, si  $p$  es 4 y:

$$X_2 = a X_1 + b X_3 + c$$

$X_2$  es una función lineal de  $X_1$  y  $X_3$  de modo que el rango de  $X'X$  puede ser a lo más igual a 3. Si esto se acerca a una dependencia lineal en  $X$  la solución de  $(X'X)^{-1}$  puede implicar errores y pérdida de significancia dando origen a una absurda estimación de  $\underline{\beta}$ , según fue comprobado por Draper y Smith, (1966).

Como en el caso de la regresión simple, en donde el total de la suma de los cuadrados puede ser desagregado en 3 partes

$$\sum Y_i^2 = n\bar{Y}^2 + \sum (Y_i - \hat{Y}_i)^2 + \sum (\hat{Y}_i - \bar{Y})^2$$

Draper y Smith, (1966) demostraron que esta ecuación puede denotarse también en forma matricial como:

$$\underline{Y}'\underline{Y} = n\bar{Y}^2 + \left( \hat{\underline{\beta}}'\underline{X}'\underline{Y} - n\bar{Y}^2 \right) + \left( \underline{Y}'\underline{Y} - \hat{\underline{\beta}}'\underline{X}'\underline{Y} \right) \quad (4.41)$$

De modo que los tres componentes de la suma total de los cuadrados  $\underline{Y}'\underline{Y}$  ó  $\sum Y_i^2$  son:

1.  $n\bar{Y}^2$  es la suma de los cuadrados debidos a la media
2.  $\underline{Y}'\underline{Y} - \hat{\underline{\beta}}'\underline{X}'\underline{Y} = (\underline{Y} - \underline{X}\hat{\underline{\beta}})'\underline{e} = \underline{e}'\underline{e} = \sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2$  es la suma de los cuadrados de las desviaciones de la regresión, es decir, la sumas de los residuales al cuadrado.
3.  $\hat{\underline{\beta}}'\underline{X}'\underline{Y} - n\bar{Y}^2 = \sum (\hat{Y}_i - \bar{Y})^2$  es la suma de los cuadrados debido a la regresión.

El coeficiente de determinación múltiple,  $R^2$ , puede ser definido ahora como:

$$R^2 = \left( \hat{\underline{\beta}}'\underline{X}'\underline{Y} - n\bar{Y}^2 \right) / \left( \underline{Y}'\underline{Y} - n\bar{Y}^2 \right) \quad (4.42)$$

como en el caso de  $r^2$ , el intervalo de  $R^2$  es de 0 a 1. El coeficiente de correlación múltiple es definido como la raíz cuadrada positiva de  $R^2$ . Una vez más  $R^2$  es una fracción de la suma total del cuadrado de los valores corregidos por la media, que se obtiene de la ecuación de regresión:  $\hat{\underline{Y}} = \underline{X}\hat{\underline{\beta}}$  Frecuentemente, la división a la suma de los cuadrados, se presenta en forma de una tabla de análisis de varianza (ANOVA). Continuando la analogía con la regresión simple, se define:  $\underline{e} = \underline{Y} - \underline{X}\hat{\underline{\beta}}$



De esta manera, nuestro proceso de estimación garantiza que  $E(\epsilon) = 0$ . Un estimador neutral para  $\text{Var}(\epsilon)$  ó  $\sigma^2$  es  $S^2$  donde:

$$S^2 = \sum e_i^2 / (n-p) = \sum (Y_i - \hat{Y}_i)^2 / (n-p) \quad (4.43)$$

$$S^2 = \underline{e}'\underline{e} / (n-p) = (\underline{Y} - \underline{X}\hat{\beta})'(\underline{Y} - \underline{X}\hat{\beta}) / (n-p) = (\underline{Y}'\underline{Y} - \hat{\beta}'\underline{X}'\underline{Y}) / (n-p)$$

El error estándar en la ecuación de regresión es estimado por  $S$ . Y una expresión para  $R^2$  análoga a  $r^2$  es:

$$R^2 = 1 - (n-p)S^2 / (n-1)S_Y^2 \quad (4.44)$$

TABLA ANOVA PARA REGRESION MULTIPLE

CONCEPTO	GRADOS DE LIBERTAD	SUMA DE CUADRADOS	MEDIA CUADRATICA ESPERADA
Media	1	$\frac{\overline{Y^2}}{n}$	
Regresión	p-1	$\hat{\beta}'\underline{X}'\underline{Y} - n\overline{Y^2}$	
Residual	n-p	$\underline{Y}'\underline{Y} - \hat{\beta}'\underline{X}'\underline{Y}$	$\sigma^2$
Total	n	$\underline{Y}'\underline{Y}$	

n número de observaciones

p número de parámetros

Una vez más se muestra que si la ecuación de regresión está explicando una gran parte de la variación en  $Y$ , el error estándar en la ecuación puede ser significativamente menor que la desviación estándar de  $Y$ .

### 4.3.1 Intervalos de confianza y pruebas de hipótesis

Como fue el caso en la regresión simple, y a fin de emplear de manera correcta, alguno de los teoremas desarrollados para intervalos de confianza y pruebas de hipótesis, en regresión múltiple, deben hacerse algunas consideraciones. Se supone que  $\varepsilon_i$  es idéntico para todos los casos, independiente y normalmente distribuido con media cero y varianza  $\sigma^2$ .

Los intervalos de confianza para  $\sigma^2$  podemos denotarlos en función de su distribución chi-cuadrada,  $(n-p)S^2/\sigma^2$ . De esta manera, los intervalos de confianza son:

$$L = \frac{(n-p)S^2}{\chi^2_{1-\alpha/2, n-p}} \qquad U = \frac{(n-p)S^2}{\chi^2_{\alpha/2, n-p}} \qquad (4.45)$$

Para hacer inferencias referentes al parámetro  $\underline{\beta}$ , la varianza del estimador  $\hat{\underline{\beta}}$  debe ser calculada. La matriz de varianza-covarianza de  $\underline{\beta}$  estará dada por:

$$\text{Cov}(\hat{\underline{\beta}}) = E\left\{ \left[ \hat{\underline{\beta}} - E(\hat{\underline{\beta}}) \right] \left[ \hat{\underline{\beta}} - E(\hat{\underline{\beta}}) \right]' \right\}$$

$$E(\hat{\underline{\beta}}) = (\underline{X}'\underline{X})^{-1} \underline{X}'\underline{Y} \quad \text{y} \quad \underline{Y} = \underline{X}\underline{\beta} + \underline{\varepsilon}$$

así es que se tiene:

$$\hat{\underline{\beta}} = (\underline{X}'\underline{X})^{-1} \underline{X}'(\underline{X}\underline{\beta} + \underline{\varepsilon})$$

$$\hat{\underline{\beta}} = (\underline{X}'\underline{X})^{-1} (\underline{X}'\underline{X})\underline{\beta} + (\underline{X}'\underline{X})^{-1} \underline{X}'\underline{\varepsilon}$$

$$\hat{\underline{\beta}} = \underline{\beta} + (\underline{X}'\underline{X})^{-1} \underline{X}'\underline{\varepsilon}$$

ó

$$\hat{\underline{\beta}}\underline{\beta} = (\underline{X}'\underline{X})^{-1} \underline{X}'\underline{\varepsilon}$$

y

$$\text{Cov}(\hat{\underline{\beta}}) = E\left[ (\hat{\underline{\beta}}\underline{\beta})(\hat{\underline{\beta}}\underline{\beta})' \right]$$

$$\text{Cov}(\hat{\beta}) = E\left[\left(\underline{X}'\underline{X}\right)^{-1} \underline{X}' \underline{\varepsilon} \underline{\varepsilon}' \underline{X} \left(\underline{X}'\underline{X}\right)^{-1}\right]$$

Una vez más, tratando a  $\underline{X}$  como fijo, se tiene:

$$\text{Cov}(\hat{\beta}) = \left(\underline{X}'\underline{X}\right)^{-1} \underline{X}' E(\underline{\varepsilon} \underline{\varepsilon}') \underline{X} \left(\underline{X}'\underline{X}\right)^{-1}$$

como  $\underline{\varepsilon}$  es una matriz de  $(n \times 1)$ ,  $\underline{\varepsilon} \underline{\varepsilon}'$  será una matriz de  $(n \times n)$ .

$$\underline{\varepsilon} \underline{\varepsilon}' = [\varepsilon_i \varepsilon_j]$$

$$E(\underline{\varepsilon} \underline{\varepsilon}') = [E(\varepsilon_i \varepsilon_j)]$$

$$E(\varepsilon_i \varepsilon_j) = \sigma^2 \quad \text{para } i = j$$

$$E(\varepsilon_i \varepsilon_j) = 0 \quad \text{para } i \neq j$$

Por lo tanto,  $E(\underline{\varepsilon} \underline{\varepsilon}')$  es una matriz simétrica con elementos en su diagonal principal ( $i = j$ ) iguales a  $\sigma^2$ , y el resto igual a cero.

$$E(\underline{\varepsilon} \underline{\varepsilon}') = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \sigma^2 & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix} = \sigma^2 \underline{I}$$

en donde  $\underline{I}$  es la matriz identidad de  $(n \times n)$

$$\text{Cov}(\hat{\beta}) = \left(\underline{X}'\underline{X}\right)^{-1} \underline{X}' \sigma^2 \underline{I} \underline{X} \left(\underline{X}'\underline{X}\right)^{-1}$$

o también

$$\text{Cov}(\hat{\beta}) = \sigma^2 \left(\underline{X}'\underline{X}\right)^{-1} \quad (4.46)$$

La varianza de  $\hat{\beta}_i$  es igual a la covarianza para esta misma  $\hat{\beta}_i$  y es  $\sigma^2$  veces los  $i$ -ésimos elementos de la diagonal principal de  $(\underline{X}'\underline{X})^{-1}$

La covarianza de  $\hat{\beta}_i$  con  $\hat{\beta}_j$  es  $\sigma^2$  veces los  $i$ -ésimos y  $j$ -ésimos elementos de la matriz  $(X'X)^{-1}$ . Si hacemos  $C^{-1} = (X'X)^{-1}$

$$\text{Var}(\hat{\beta}_i) = \sigma_{\beta_j}^2 = C_{ii}^2 \sigma^2 \quad (4.47)$$

donde

$C_{ii}^{-1}$   $j$ -ésimos términos de  $(X'X)^{-1}$

Si el modelo es correcto, y el valor de  $\hat{\beta}_i/S\hat{\beta}_i$  sigue la distribución "t" con  $(n-p)$  grados de libertad.  $S\hat{\beta}_i$  es un estimador de  $\sigma\hat{\beta}_i$  y es igual al valor positivo de la raíz cuadrada de  $C_{ii}^2 S^2$ .

Finalmente, los intervalos de confianza para  $\beta_i$  están dados por:

$$L\beta_i = \hat{\beta}_i - t_{1-\alpha/2, n-p} S\hat{\beta}_i \quad U\beta_i = \hat{\beta}_i + t_{1-\alpha/2, n-p} S\hat{\beta}_i \quad (4.48)$$

La hipótesis de prueba  $\beta_i = \beta_0$  en donde  $\beta_0$  es una constante conocida, puede hacerse bajo la hipótesis nula de que  $(\hat{\beta}_i - \beta_0)/S\hat{\beta}_i$  sigue la distribución "t". Esto es:

$$H_0: \beta_i = \beta_0 \quad \text{y} \quad H_1: \beta_i \neq \beta_0$$

En donde el estadístico 't' es calculado como"

$$t = \frac{\hat{\beta}_i - \beta_0}{S\hat{\beta}_i} \quad (4.49)$$

La hipótesis nula  $H_0$ , se rechaza si  $|t| > t_{1-\alpha/2, n-p}$

Dado que, en general el parámetro  $\hat{\beta}_i$  no es independiente de  $\hat{\beta}_j$  ( su covarianza esta dada por  $C_{ij}^{-1} \sigma^2$  ) se repite la ecuación (4.48).

$$H_0: \beta_i = \beta_{0,i} \quad \text{y} \quad H_0: \beta_j = \beta_{0,j}$$

La prueba de  $H_0: \beta_i = 0$  contra  $H_1: \beta_i \neq 0$  es equivalente a probar la hipótesis, de que la  $i$ -ésima variable independiente no contribuye significativamente a explicar la variación de la variable dependiente.

Si  $H_0: \beta_i = 0$  es aceptada, en muchas ocasiones es conveniente eliminar la  $i$ -ésima variable independiente del modelo.

Una prueba de hipótesis para probar que toda la ecuación de regresión no explica el significado total de la variación de  $Y$ , puede ser:

$$H_0: \beta_2 = \beta_3 = \dots = \beta_p = 0$$

$H_1: \beta_i \neq \beta_0$  por lo menos uno de los  $\beta$  no es cero

Dado que  $\beta_i$  no es independiente de  $\beta_j$ , repetir la aplicación de la ecuación (4.48) no es una manera válida para probar esta hipótesis. Debe emplearse el valor que aporta el radio de la media cuadrada, debido a la regresión de los residuales.  $Y$  que además, sigue la distribución "F" con  $(p-1)$  y  $(n-p)$  grados de libertad.

Para probar  $H_0: \beta_2 = \beta_3 = \dots = \beta_p = 0$ , calculamos la prueba estadística siguiente:

$$F = \frac{(\hat{\beta}' \underline{X}' \underline{Y} - n \overline{Y}^2) / (p-1)}{(\underline{Y}' \underline{Y} - \hat{\beta}' \underline{X}' \underline{Y}) / (n-p)} \quad (4.50)$$

y la hipótesis nula  $H_0$  se rechaza si  $F > F_{1-\alpha, p-1, n-p}$

Una prueba de hipótesis en que " $k$ " variables independientes no contribuyen significativamente a explicar la variación lineal de la variable dependiente, puede realizarse con este estadístico.

Para lograr esto, se sugiere arreglar el modelo de tal manera, que las últimas " $k$ " variables sean las que se van a probar. La hipótesis es tal, que se propone que las últimas " $k$ " variables independientes son las que no contribuyen significativamente a explicar la variación lineal en  $Y$ . Esto es equivalente a escribir.

$$H_0: \beta_{p-k+1} = \beta_{p-k+2} = \dots = \beta_p = 0$$

$H_1$ : por lo menos uno de los  $\beta$  no es cero

La prueba  $H_0$ , denota que la totalidad del modelo, contiene " $p$ " variables independientes. También denota un modelo reducido que se obtiene al eliminar las últimas " $k$ " variables independientes con  $(p-k)$  variables independientes. Al modificar la notación se obtiene:

$$F = \frac{(Q_2 - Q_2^*)/k}{Q_1/(n-p)} \quad (4.51)$$

donde

- $Q_2$  suma de los cuadrados, debida a la regresión en el modelo completo con  $(p-1)$  grados de libertad.
- $Q_1$  suma del cuadrado de los residuales en el modelo completo con  $(n-p)$  grados de libertad.
- $Q_2^*$  suma de los cuadrados, debida a la regresión en el modelo reducido con  $(p-k-1)$  grados de libertad.

Se tiene una distribución con " $k$ " y  $(n-p)$  grados de libertad.  $H_0$  es rechazada si:

$$F > F_{1-\alpha, k, n-p}$$

Nótese que  $Q_2 - Q_2^*$  es la suma reducida de los cuadrados en la regresión, ocasionada por eliminar " $k$ " variables independientes. Si el valor de  $Q_2^*$  es muy cercano o igual a  $Q_2$ , quiere decir que el haber eliminado las " $k$ " variables, no ha producido grandes cambios en la capacidad del modelo para explicar la variación lineal en  $Y$ , bajo estas condiciones  $F$  será pequeño y  $H_0$  no será rechazada esto a su vez indicará que deben eliminarse las " $k$ " variables que se han considerado. El rechazo de  $H_0$  no implica que todas las últimas " $k$ " son importantes; esto solo implica que, al menos una de estas variables es significativa en toda la variación de  $Y$ .

### 4.3.2 Criterio de selección

Una situación común en la cual se emplea la regresión múltiple, es cuando se dispone de una variable dependiente y varias independientes, y es necesario encontrar un modelo lineal para predecir valores no observados en la variable dependiente. El modelo que debe desarrollarse, no necesariamente debe contener a la totalidad de las variables independientes. Por lo tanto, los puntos de interés son:

- 1) ¿Puede ser empleado un modelo lineal?
- 2) ¿Cuál (es) variable (s) debe (n) ser incluida (s) en el modelo?

Un punto que complica la selección del modelo es que en la mayoría de los casos las variables independientes no son del todo estadísticamente independientes, sin embargo están correlacionadas.

Uno de los primeros pasos que debe hacerse en un análisis de regresión es calcular la matriz de correlación  $\underline{R}$  para las variables independientes:

$$Z_{i,j} = (X_{i,j} - \bar{X}_j) / S_j \quad (4.52)$$

donde

- $\bar{X}_j$  media de las  $j$ -ésimas variables independientes
- $S_j$  desviación estándar de las  $j$ -ésimas variables independientes

Si definimos  $\underline{Z} = [Z_{i,j}]$

$$\underline{R} = \underline{Z}' \underline{Z} / (n - 1) = [R_{i,j}] \quad (4.53)$$

donde

- $R_{i,j}$  correlación entre la  $i$ -ésima y la  $j$ -ésima variable independiente

Ahora, se debe analizar si  $R_{i,j} = 1$  para  $i \neq j$ , ya que alguna de las variables  $i$  o  $j$  pueden ser omitidas del modelo, de otro modo la matriz  $\underline{X}' \underline{X}$  no podrá ser invertida. Si  $R_{i,j}$  es cercano a la unidad (pero no igual),  $\underline{X}' \underline{X}$  podrá ser invertida, y podrá estimarse  $\underline{\beta}$ . Si  $R_{i,j}$  es cercano a la unidad, entonces  $\text{Var}(\hat{\beta}_i)$  o  $\text{Var}(\hat{\beta}_j)$  podrán tener valores grandes.

Las pruebas de hipótesis en  $\beta_i$  y  $\beta_j$  pueden indicar que ninguno es significativamente diferente de cero, cuando en realidad alguno de ellos es tratado aisladamente, puede resultar significativamente diferente de cero.

El problema aquí es que, dado que  $X_i$  y  $X_j$  están estrechamente relacionadas, ambas intentan dar explicación a la misma causa dentro del modelo lineal. Al considerar ambas variables  $X_i$  y  $X_j$  dentro del modelo, la parte de la variación en  $Y$  que uno u otro pudieran explicar si fueran empleadas de manera aislada, se divide entre ellos de tal modo que ninguno es significativo. En otras palabras, el único efecto causado por un factor, (el cual quizá se refleje en  $X_i$  o  $X_j$ ) está siendo dividido entre dos variables correlacionadas.

El emplear variables en la ecuación de regresión, que están altamente correlacionadas, hace de gran dificultad la interpretación de los coeficientes de regresión. Algunas veces el signo del coeficiente de regresión puede ser el contrario del que se espera, si dicha variable está fuertemente correlacionada con otra variable independiente dentro de la ecuación.

Una práctica común al seleccionar un modelo de regresión múltiple (que no necesariamente es la que se debe adoptar) es llevar a cabo algunas regresiones, en los datos dados, usando diferentes combinaciones en las variables independientes. La regresión que "mejor" ajuste los datos, es la que se selecciona. Un criterio empleado comúnmente para obtener el "mejor" ajuste, es seleccionar la ecuación que produce el mayor valor de  $R^2$ .

Cabe recordar que no siempre el empleo del criterio del mayor valor de  $R^2$  y el menor de  $S$  dan los mismos resultados. Además, aunque se le agregen más variables a la ecuación de regresión, el valor de  $R^2$  nunca aumentará. Así desde el punto de vista del criterio de  $R$  se deben emplear todas las variables que se tengan a la mano. Sin embargo, esto genera una ecuación difícil de manejar y en la cual, es extremadamente difícil dar una interpretación adecuada a los coeficientes. Además mientras más variables sean añadidas a la ecuación de regresión, aumentará el valor del error estándar, esto se aprecia claramente en la ecuación (4.42), en donde cada vez que se agrega una variable, se reduce el denominador  $(n-p)$  y aumenta el de  $\frac{Y'Y - \hat{\beta}'X'Y}{n-p}$ . Sin embargo, el numerador muchas veces puede o no aumentar proporcionalmente a  $(n-p)$ . Esto es muy importante de tomar en cuenta, ya que no siempre el adicionar variables contribuye significativamente a la regresión, y tal vez convenga dejarlas fuera.

Todas las variables que permanezcan en la regresión deben de contribuir con la regresión a menos que exista una razón de peso (teórica o intuitiva) para incluir alguna variable no significativa, además las variables incluidas deben de tener un adecuado significado físico. En el caso de que dos variables describan el mismo efecto cuando se usan por separado, y ambas sean significativas, se sugiere emplear la más fácil de obtener.

El número de coeficientes estimados no debe exceder del 25 al 35% del número de observaciones. Esto es una regla, para evitar un "sobre-ajuste" que provoque oscilaciones en la ecuación repercutiendo en la relación entre observaciones y variables independientes.

Debido al gran desarrollo de la computación es posible calcular un gran número de ecuaciones de regresión, cada una de ellas con una gran serie de registros. Esto ha conducido a varios procedimientos formales, dos de ellos serán discutidos. El método de las "regresiones de gran posibilidad" y las "regresiones de paso inmediato".

El método de las regresiones de gran posibilidad implica calcular ecuaciones de regresión teniendo cada una de ellas, cada posible combinación de las variables  $X$ . Si se requieren todas las ecuaciones para encontrar el término de intersección,  $2^p - 1$  ecuaciones de regresión pueden ser calculadas; (en



donde  $p$  es el número de variables independientes) una de las cuales es siempre igual a uno y se le denomina de intersección.

De este modo si  $p=4$ , 8 ecuaciones pueden obtenerse. Si  $p=11$ , se deberán calcular y examinar, 1024 ecuaciones; de este modo siempre que  $p$  resulte moderadamente grande, será prohibitivo encontrar el número de ecuaciones requerido, y la probabilidad de encontrar una regresión significativa, se dará solo por casualidad.

Uno de los más cómodos procedimientos para seleccionar la "mejor" ecuación de regresión es el método de las regresiones de paso inmediato. Este procedimiento consiste en construir una ecuación de regresión, adicionando en cada paso una variable a la vez, la cual irá explicando poco a poco la cantidad que queda de variaciones no explicadas. Después de cada paso, todas las variables dentro de la ecuación son examinadas en su significado y discutido su aporte al significado total de la variación.

Así, la primera variable adicionada, es una que tiene gran correlación simple con la variable dependiente. La segunda variable que se agrega es aquella que explica la variación en la variable dependiente, de lo que no ha podido explicar la primera variable adicionada. Hasta este punto las variables son probadas en su significancia, y según sea el resultado, serán o no aceptadas. La tercera variable que se incorpora es la que explica la variación, que no ha podido ser explicada por las variables ya existentes en la ecuación. Este procedimiento continúa, hasta haber encontrado las variables que hayan sido (o no hayan sido) significantes, más se debe tener cuidado en asegurarse de que se obtiene una ecuación racional.

Por supuesto, que la verdadera prueba de qué tan bueno resulta ser el modelo de regresión, depende de la habilidad que tenga éste para predecir la variable dependiente a partir de observaciones de las variables independientes que no hayan sido empleadas para estimar los parámetros de regresión. Para lograr hacer una comparación de esta naturaleza, es necesario dividir la muestra aleatoriamente en dos partes. La primera parte de los datos será empleada para desarrollar el modelo y la otra parte para probarlo. Desafortunadamente, la mayoría de las veces en las aplicaciones hidrológicas, no se tienen disponibles las suficientes observaciones para llevar a cabo este procedimiento.

#### 4.4 Extensión de registros

La disponibilidad de registros simultáneos en dos o más estaciones de medición, puede ser muy útil para extender el registro correspondiente a una estación con menos longitud de registro. Entonces el nuevo registro extendido puede usarse para obtener nuevos estimadores de sus estadísticos, logrando con esto la transferencia de información de un sitio a otro.

Sin embargo, no siempre los resultados son satisfactorios al extender una muestra con periodo de registro corto, ya que es necesario que los estimadores obtenidos a través de la muestra extendida deben ser estadísticamente superiores a los producidos por la muestra original.

Las técnicas de transferencia de información pueden usarse para:

1. Mejorar los estimadores de los parámetros de una muestra pequeña
2. Extender los valores observados de una muestra con periodo de registro corto
3. Completar la información faltante

Considerese un par de secuencias, una corta y otra larga de variables hidrológicas aleatorias  $X, Y$

$$\begin{aligned} & Y_1, Y_2, \dots, Y_n \\ & X_1, X_2, \dots, X_n, X_{n+1}, X_{n+2}, \dots \end{aligned}$$

Un modelo de regresión lineal simple puede ser expresado como:

$$Y_t = \bar{Y}_1 + \bar{r} \frac{S_1(Y)}{S_2(X)} (X_t - \bar{X}) + \alpha \theta \sqrt{1 - \bar{r}^2} S_1(Y) \varepsilon_t \quad (4.54)$$

donde

$\bar{Y}_1$       valor de regresión para un valor dado de  $X_1$   
 $\bar{r}$         coeficiente de correlación simple obtenido de las muestras como:

$$\bar{r} = \frac{1}{N_1} \sum_{i=1}^{N_1} \frac{(X_i - \bar{X}_1)(Y_i - \bar{Y}_1)}{S_1(X) S_1(Y)} \quad (4.55)$$

donde

$\bar{Y}_1$  y  $\bar{X}_1$       medias estimadas del registro simultáneo,  $N_1$  de las variables  $Y$  y  $X$ , respectivamente, y son obtenidas como:

$$\bar{X}_1 = \frac{1}{N_1} \sum_{t=1}^{N_1} X_t \quad (4.56)$$

$$\bar{Y}_1 = \frac{1}{N_1} \sum_{t=1}^{N_1} Y_t \quad (4.57)$$

$S_1(X)$  y  $S_1(Y)$  son las desviaciones estándar insesgadas y que son estimadas como:

$$S_1(X) = \left[ \sum_{t=1}^{N_1} \left( \frac{X_t - \bar{X}_1}{N_1 - 1} \right)^2 \right]^{1/2} \quad (4.58)$$

$$S_1(Y) = \left[ \sum_{t=1}^{N_1} \left( \frac{Y_t - \bar{Y}_1}{N_1 - 1} \right)^2 \right]^{1/2} \quad (4.59)$$

$\alpha$  es un parámetro usado para remover el sesgo en el estimador de la varianza de "Y" y es obtenido por:

$$\alpha^2 = \frac{N_2(N_1 - 4)(N_1 - 1)}{(N_2 - 1)(N_1 - 3)(N_1 - 2)} \quad (4.60)$$

La media y la varianza de la muestra extendida son:

$$\bar{Y} = \bar{Y}_1 + \frac{N_2}{(N_1 + N_2)} \hat{b}(\bar{X}_2 - \bar{X}_1) \quad (4.61)$$

$$S^2(Y) = \frac{1}{(N_1 + N_2 - 1)} \left[ (N_1 - 1)S_1^2(Y) + (N_2 - 1)\hat{b}^2 S_2^2(X) + \frac{N_1 N_2}{(N_1 + N_2)} (\bar{X}_2 - \bar{X}_1)^2 + \dots \right. \\ \left. \dots + (N_2 - 1)\theta^2 \alpha^2 (1 - \bar{r}^2) S_1^2(Y) \right] \quad (4.62)$$

Donde  $\theta = 1$  si va a usarse la componente aleatoria  $\varepsilon_t$  y  $\theta = 0$  cuando no se usa esta opción.  $\varepsilon_t$  es una variable aleatoria normal con media cero y varianza igual a uno (puede ser obtenida de una tabla de números aleatorios [0,1])

$$\hat{b} = \frac{\bar{r} S_1(Y)}{S_1(X)} \quad (4.63)$$

$\bar{X}_2$  y  $S_2(X)$  son estimadores de la media y la desviación estándar para la secuencia.

$X_{N_1+1}, \dots, X_{N_1+N_2}$  de la variable "X"

$$\bar{X}_2 = \frac{1}{N_2} \sum_{t=N_1+1}^{N_1+N_2} X_t \quad (4.64)$$

$$S_2(X) = \left[ \sum_{t=N_1+1}^{N_1+N_2} \frac{(X_t - \bar{X}_2)^2}{N_2 - 1} \right]^{1/2} \quad (4.65)$$

#### 4.4.1 Criterios de mejora en la estimación de la media y varianza

La mejora de la estimación de la media y varianza en una muestra extendida, es medida a través del *contenido de información relativa*.

$$I = \frac{\text{Var}(\hat{\gamma}_1)}{\text{Var}(\hat{\gamma})} \quad (4.66)$$

donde

$\text{Var}(\hat{\gamma}_1)$       varianza del parámetro  $\gamma$  estimado a través del registro original con longitud  $N_1$

$\text{Var}(\hat{\gamma})$       varianza del parámetro  $\gamma$  estimado a través del registro combinado  $N_1+N_2$

Para que exista superioridad al usar las dos estaciones, "I" tiene que ser mayor a la unidad. Por lo tanto, en el caso de la media, se tiene que la varianza de ésta cuando se usan ambas estaciones, es el siguiente para el registro extendido:

$$\text{Var}(\hat{\gamma}) = \frac{\hat{\sigma}_y^2}{N_1} \left\{ 1 - \frac{N_2}{N_1+N_2} \left[ \bar{r}^2 - \frac{(1-\bar{r}^2)}{(N_1-3)} \right] \right\} \quad (4.67)$$

La varianza de la media, tomando solo el registro original  $N_1$  es:

$$\text{Var}(\hat{y}_1) = \frac{\hat{\sigma}_y^2}{N_1} \quad (4.68)$$

Y dado que  $I > 1$  o lo que es lo mismo  $\text{Var}(\bar{Y}) < \text{Var}(\bar{Y}_1)$  se tiene que esto ocurre cuando:

$$\bar{r} > \left( \frac{1}{N_1 - 2} \right)^{1/2} < \hat{r}_{cm1} \quad (4.69)$$

$\hat{r}_{cm1}$  es función de  $N_1$  y se obtiene por la interpolación de los valores que se muestran en la tabla siguiente:

$N_1$	5	10	15	20	25	30	40	50
$\hat{r}_{cm1}$	0.58	0.35	0.28	0.24	0.21	0.19	0.16	0.14

El estimado de la varianza puede ser mejorado por análisis de correlación, si  $\text{Var}[S^2(Y)]$  la varianza del estimado después de que el registro es extendido por correlación, es menor que  $\text{Var}[S_1^2(Y)]$  la varianza del estimado de la muestra original es:

$$\text{Var}[S^2(Y)] = \frac{2\sigma_y^4}{(N_1 - 1)} + \frac{N_2\sigma_y^4}{(N_1 + N_2 - 1)^2} [A\bar{r}^4 + B\bar{r}^2 + C] \quad (4.70)$$

Para que  $S^2(Y)$  (registro extendido) sea un mejor estimador de  $\sigma_y^2$  ( $S_1^2(Y)$  registro corto)

$$\text{Var}[S^2(Y)] < \left\{ \text{Var}[S_1^2(Y)] = \frac{2\sigma_y^4}{(N_1 - 1)} \right\} \quad (4.71)$$

Esto ocurre si:

$$|\bar{r}| > \left( \frac{-B + \sqrt{B^2 - 4AC}}{2A} \right)^{1/2} \quad (4.72)$$

Si esta desigualdad no se satisface, entonces  $S_1^2(Y)$  es mejor estimador de  $\sigma_y^2$  que  $S^2(Y)$

Los valores de A, B y C se muestran en la Tabla 4.1. Si se prefiere, los valores del coeficiente de correlación simple crítico de la varianza para  $\theta = 0$  y  $\theta = 1$  se muestran en la Tabla 4.2.

No es suficiente el determinar que existirá mejora en la media y la varianza al extender un registro, además es necesario considerar el número de años que podrán extenderse. Para esto Langbein propone lo que se denomina *el período efectivo de registro*  $N_e$  el cual indica, con base en los registros combinados de las dos estaciones, cuántos años como máximo podrán extenderse sin que se perjudiquen las características estadísticas del registro. Este criterio relaciona el número de años de la estación base, los años de la estación a la que se le va ampliar el registro y su respectivo coeficiente de determinación  $R^2$  obtenido de la regresión lineal.

$$N_e = \frac{N_1 + N_2}{1 + \frac{N_2}{N_1 - 2}(1 - R^2)} \quad (4.73)$$

donde

- $N_e$  período efectivo de registro, en años
- $N_1$  número de años del registro común (registro corto)
- $N_2$  número de años del registro largo
- $R^2$  coeficiente de determinación

**TABLA 4.1**  
**COEFICIENTES DE A, B Y C PARA ESTIMAR LA VARIANZA Y EL COEFICIENTE**  
**DE CORRELACION SIMPLE CRITICO DE LA VARIANZA PARA  $\theta=1$**

N2	COEFICIENT	N1								
		10	15	20	25	30	50	75	100	
10	<b>A</b> <hr/> $r^4$	11.83	15.38	15.22	14.82	14.46	13.58	13.08	12.82	
15		20.99	26.35	26.26	25.76	25.29	24.13	23.46	23.11	
20		30.16	37.33	37.30	36.70	36.13	34.68	33.84	33.40	
25		39.32	48.30	48.34	47.64	46.96	45.22	44.22	43.68	
30		48.48	59.27	59.38	58.58	57.79	55.77	54.60	53.97	
50		85.13	103.16	103.53	102.34	101.13	97.97	96.11	95.13	
75		130.95	158.03	158.73	157.05	155.30	150.72	148.01	146.57	
100		176.77	212.89	213.92	211.75	209.48	203.47	199.90	198.02	
10		<b>B</b> <hr/> $r^2$	1.53	0.95	0.64	0.47	0.37	0.20	0.12	0.09
15			1.90	1.15	0.75	0.54	0.42	0.21	0.13	0.09
20	2.27		1.34	0.86	0.61	0.47	0.23	0.14	0.10	
25	2.64		1.53	0.97	0.68	0.52	0.25	0.15	0.10	
30	3.00		1.73	1.08	0.75	0.56	0.27	0.15	0.11	
50	4.47		2.50	1.52	1.03	0.76	0.33	0.18	0.12	
75	6.31		3.47	2.06	1.38	0.99	0.42	0.22	0.14	
100	8.15		4.43	2.61	1.72	1.23	0.50	0.26	0.16	
10	<b>C</b>		0.44	0.09	0.04	0.02	0.01	0.00	0.00	0.00
15			0.57	0.12	0.05	0.02	0.01	0.00	0.00	0.00
20		0.71	0.14	0.05	0.03	0.02	0.00	0.00	0.00	
25		0.84	0.16	0.06	0.03	0.02	0.00	0.00	0.00	
30		0.97	0.18	0.07	0.03	0.02	0.00	0.00	0.00	
50		1.51	0.27	0.10	0.05	0.03	0.01	0.00	0.00	
75		2.18	0.39	0.13	0.06	0.03	0.01	0.00	0.00	
100		2.85	0.50	0.17	0.08	0.04	0.01	0.00	0.00	

TABLA 4.2

COEFICIENTE DE CORRELACION SIMPLE CRITICO DE LA VARIANZA PARA  $\theta=0$

N2 \ NI	6	8	10	12	14	16	18	20	25	30	35	40	45	50	55	60
6	0.81	0.69	0.63	0.63	0.63	0.64	0.65	0.65	0.66	0.67	0.68	0.68	0.68	0.69	0.69	0.69
8	0.82	0.72	0.70	0.70	0.70	0.71	0.72	0.72	0.73	0.73	0.74	0.74	0.74	0.75	0.75	0.75
10	0.82	0.74	0.73	0.73	0.73	0.74	0.75	0.75	0.76	0.76	0.77	0.77	0.77	0.78	0.78	0.78
12	0.82	0.75	0.74	0.75	0.75	0.76	0.76	0.77	0.78	0.78	0.79	0.79	0.79	0.80	0.80	0.80
14	0.82	0.76	0.75	0.76	0.77	0.77	0.78	0.78	0.79	0.80	0.81	0.81	0.82	0.82	0.82	0.82
16	0.82	0.76	0.76	0.77	0.78	0.78	0.79	0.79	0.80	0.81	0.81	0.82	0.82	0.83	0.83	0.83
18	0.83	0.77	0.77	0.77	0.79	0.79	0.79	0.80	0.81	0.81	0.82	0.82	0.83	0.83	0.83	0.83
20	0.83	0.77	0.77	0.78	0.79	0.79	0.80	0.80	0.81	0.82	0.83	0.83	0.83	0.84	0.84	0.84
25	0.83	0.78	0.78	0.79	0.80	0.80	0.81	0.82	0.83	0.83	0.84	0.84	0.85	0.85	0.85	0.85
30	0.83	0.78	0.78	0.79	0.81	0.81	0.82	0.82	0.83	0.84	0.85	0.85	0.86	0.86	0.86	0.86
35	0.83	0.78	0.78	0.80	0.82	0.82	0.82	0.83	0.84	0.85	0.86	0.86	0.86	0.86	0.87	0.87
40	0.83	0.79	0.79	0.80	0.82	0.82	0.83	0.83	0.84	0.85	0.86	0.86	0.86	0.87	0.87	0.87
45	0.83	0.79	0.79	0.60	0.82	0.82	0.83	0.83	0.85	0.85	0.86	0.86	0.87	0.87	0.87	0.88
50	0.83	0.79	0.79	0.61	0.82	0.82	0.83	0.84	0.85	0.86	0.87	0.87	0.87	0.87	0.88	0.88
55	0.83	0.79	0.79	0.81	0.83	0.83	0.83	0.84	0.85	0.86	0.87	0.87	0.87	0.88	0.88	0.88
60	0.83	0.79	0.79	0.81	0.83	0.83	0.83	0.84	0.85	0.86	0.87	0.87	0.88	0.88	0.88	0.89

COEFICIENTE DE CORRELACION SIMPLE CRITICO DE LA VARIANZA PARA  $\theta=1$

N2 \ NI	6	8	10	12	14	16	18	20	25	30	35	40	45	50	55	60
6	0.82	0.71	0.64	0.59	0.55	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
8	0.82	0.71	0.65	0.59	0.55	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
10	0.82	0.72	0.65	0.59	0.55	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
12	0.82	0.72	0.65	0.59	0.55	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
14	0.82	0.72	0.65	0.59	0.55	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
16	0.82	0.72	0.65	0.60	0.55	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
18	0.82	0.72	0.65	0.60	0.55	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
20	0.82	0.72	0.65	0.60	0.55	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
25	0.82	0.72	0.65	0.60	0.55	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
30	0.82	0.72	0.65	0.60	0.55	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
35	0.82	0.72	0.65	0.60	0.56	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
40	0.82	0.72	0.65	0.60	0.56	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
45	0.82	0.72	0.65	0.60	0.56	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
50	0.83	0.72	0.65	0.60	0.56	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
55	0.83	0.72	0.65	0.60	0.56	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28
60	0.83	0.72	0.65	0.60	0.56	0.52	0.49	0.47	0.42	0.39	0.36	0.34	0.32	0.30	0.29	0.28



ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA

## 5. DISTRIBUCIONES DE PROBABILIDAD

Las distribuciones univariadas de probabilidad pueden emplearse como un medio para estimar eventos de diseño en cuencas no aforadas. Existen numerosas técnicas regionales que emplean este tipo de distribuciones, por ejemplo el método de la Avenida-Índice, de Estaciones-año y Box-Cox regional. Por otra parte las distribuciones bivariadas sirven para mejorar la calidad de la información cuando se tiene un registro grande en una cuenca vecina (Escalante, 1991).

### 5.1 Distribuciones univariadas

Definamos una función de densidad univariada para una función realmente variada, la cual esta definida en el eje real  $\Re(-\infty, \infty)$  y satisface:

$$f(x) \geq 0 \quad \text{para} \quad -\infty < x < \infty$$

Función de densidad de probabilidad

$$F(X) = \int_{-\infty}^x f(x) dx = 1$$

Distribución de probabilidad

(5.1)

Una variable aleatoria  $X$  es absolutamente continua si y solo si existe una función de densidad  $f$  para la cual:

$$\Pr(a < x \leq b) = \int_a^b f(x) dx \quad (5.2)$$

para cualquier  $a < b$

Dicha función de distribución asociada a la variable aleatoria  $X$  se define como la probabilidad de que dicha variable tome valores iguales o menores que un valor fijo  $X$  para toda  $x$  perteneciente al conjunto de los números reales.

### 5.1.1 Estimación de parámetros

Para que se puedan realizar inferencias estadísticas utilizando las distribuciones de probabilidad, se requiere de la estimación de sus parámetros.

Existen en la literatura una gran variedad de métodos para estimar parámetros de distribuciones, entre los que se pueden mencionar, entre otros, por mínimos cuadrados, promedios de probabilidad pesada, por momentos y por máxima verosimilitud.

El método de momentos es uno de los más empleados en la Hidrología y consiste en igualar los momentos poblacionales (características estadísticas de la muestra) con los muestrales:

Considerando que la esperanza de una variable aleatoria  $x$  ( $E(x)$ ) es conocida como la media de  $x$  denotada por  $\mu$ , los parámetros que caracterizan a una población se definen como:

*1er. Momento (media)*

Poblacional	$\mu = E(x)$	(5.3)
-------------	--------------	-------

Muestral	$\bar{x} = \frac{1}{N} \sum_{i=1}^n x_i$	(5.4)
----------	--	-------

*2o. Momento (Varianza)* Es una medida de dispersión de  $x$  y una medida de tendencia central para las desviaciones  $(x - \mu)^2$  al cuadrado.

$$\text{Poblacional} \quad \sigma^2 = E[(x - \mu)^2] \quad (5.5)$$

$$\text{Muestral} \quad S^2 = \frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{scsgada} \quad (S_{\text{scsg}}) \quad (5.6)$$

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N-1} \quad \text{inscsgada} \quad (S_{\text{inscsg}}) \quad (5.7)$$

**3er. Momento (Coeficiente de Asimetría).** Con frecuencia una función de densidad no es simétrica con respecto al máximo, sino que tiene una cola más grande que otras, por lo que se dice que la distribución esta sesgada a la derecha o a la izquierda. El coeficiente de sesgo o asimetría se obtiene por:

$$\text{Poblacional} \quad \alpha_3 = \frac{E[(x - \mu)^3]}{\sigma^3} \quad (5.8)$$

$$\text{Muestral} \quad g = \frac{\frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^3}{(S_{\text{scsg}}^2)^{3/2}} \quad \text{scsgado} \quad (5.9)$$

$$g = \frac{N^2}{(N-1)(N-2)} \frac{\frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^3}{(S_{\text{scsg}}^2)^{3/2}} \quad \text{inscsgado} \quad (5.10)$$

**4o. Momento (Curtosis).** En algunos casos una función de densidad puede tener sus valores concentrados cerca de la media, así que la función de densidad tiene un pico grande, en otros casos la densidad de probabilidad puede ser relativamente plana. A la medida del grado de apuntamiento o aplanamiento se llama coeficiente de Curtosis y se determina por:

$$\text{Poblacional} \quad \alpha_4 = \frac{E[(x - \mu)^4]}{\sigma^4} \quad (5.11)$$

$$\text{Muestral} \quad k = \frac{\frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^4}{(S_{\text{sesg}}^2)^2} \quad \text{sesgado} \quad (5.12)$$

$$k = \frac{N^3}{(N-1)(N-2)(N-3)} \frac{\frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^4}{(S_{\text{sesg}}^2)^2} \quad \text{insesgado} \quad (5.13)$$

En todas las ecuaciones anteriores las letras griegas identifican a estimadores poblacionales.

Así, por ejemplo, si se desea obtener el parámetro de escala  $\beta$  de la distribución exponencial por el método de momentos, se tendrá:

$$f(x) = \beta e^{-\beta x} \quad x > 0$$

Momento poblacional:

$$\mu = E(x) = \int_{-\infty}^{\infty} x f(x) dx \quad \mu = \frac{1}{\beta}$$

Momento muestral:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^n x_i \quad \mu = \bar{x}$$

$$\frac{1}{\beta} = \frac{1}{N} \sum_{i=1}^n x_i \quad \beta = \frac{1}{\bar{x}}$$

El método de Máxima Verosimilitud consiste en una función de densidad de probabilidad de  $x$  con parámetros  $a_i$ ,  $i=1, \dots, m$  que desean ser estimados, esto es:  $f(x, a_1, a_2, \dots, a_m)$  habrá una muestra aleatoria  $x_1, x_2, \dots, x_n$  de esta densidad, entonces su densidad conjunta es:

$$f(x_1, x_2, x_3, \dots, x_n ; a_1, a_2, a_3, \dots, a_m)$$

puesto que la muestra es aleatoria, la densidad conjunta puede ser escrita como:

$$\prod_{i=1}^m f(x_i ; a_1, a_2, a_3, \dots, a_m) \quad (5.14)$$

La probabilidad de obtener un valor dado de  $x$ , por ejemplo  $x_1$ , es proporcional a  $f(x_1, a_1, a_2, \dots, a_m)$  y la probabilidad de obtener la muestra aleatoria  $x_1, x_2, \dots, x_n$  de la población  $X$  es proporcional al producto de sus densidades individuales. La función de densidad conjunta es llamada la función de Verosimilitud y es denotada como:

$$L = \prod_{i=1}^n f(x_i ; \theta) \quad (5.15)$$

Para poder resolver esta expresión, es necesario hacer máxima la función  $L$ , y dado que una función alcanza su máximo en el mismo lugar que lo alcanza su logaritmo, por facilidad se maximiza la función logarítmica de verosimilitud.

$$LL = L_n \prod_{i=1}^n f(x_i ; \theta) \quad (5.16)$$

Finalmente, al derivar esta expresión con respecto a cada uno de los parámetros; se obtendrán tantas ecuaciones como parámetros tenga la función. Y al igualar las ecuaciones a cero, se despejan los parámetros.

Así por ejemplo, si se desean calcular los parámetros de la distribución normal, por el método de máxima verosimilitud se tiene:

$$L(x, \mu, \sigma^2) = L(x_i, \theta) = \prod_{i=1}^n f(x_i)$$

$$L(x, \mu, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma}} e^{-1/2 \left( \frac{x-\mu}{\sigma} \right)^2}$$

$$L(x, \mu, \sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-1/2 \sum \left( \frac{x-\mu}{\sigma} \right)^2}$$

tomando  $\frac{\partial L}{\partial \mu} = 0$  ;  $\frac{\partial L}{\partial \sigma^2} = 0$  y resolviendo conjuntamente

$$\hat{\mu} = \frac{1}{n} \sum x$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$$

En la Hidrología estadística, existen una gran variedad de funciones de distribución de probabilidad y cada una de ellas, presenta parámetros particulares. El término denominado *ajustar* una función de distribución a una serie de datos (por ejemplo escurrimientos), identifica al procedimiento de intentar que alguna de las funciones represente de la mejor manera a la población de esos eventos. En este renglón, el lograr determinar cuál de las distribuciones es la mejor se logra al aplicar lo que se conocen con el nombre de pruebas de bondad de ajuste, entre las que se pueden mencionar: prueba Chi-cuadrada, de Kolmogorov-Smirnoff y el Error Estándar de ajuste (Kite, 1977).

Por la facilidad y confiabilidad de la prueba, en este trabajo se utilizará el criterio del Error Estándar de ajuste, el cual se calcula como:

$$EE = \left[ \frac{\sum_{i=1}^n (x_i - y_i)^2}{n - m} \right]^{1/2} \quad (5.17)$$

donde

- $x_i$  eventos registrados
- $y_i$  eventos calculados
- $n$  tamaño de la muestra
- $m$  número de parámetros de la distribución de ajuste

A continuación se presentan las principales funciones de distribución de probabilidad  $F(x)$  o de densidad  $f(x)$  empleadas en la Hidrología, así como los parámetros que las caracterizan.

### 5.1.2 Distribución Normal

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2} \quad (5.18)$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty \quad (5.19)$$

donde

- $\mu$       parámetro de ubicación (media)
- $\sigma$       parámetro de escala (desviación estándar)

### 5.1.3 Distribución Lognormal

De dos parámetros

$$f(x) = \frac{1}{\sqrt{2\pi} x \sigma_y} e^{-1/2 \left[ \frac{\ln(x) - \mu_y}{\sigma_y} \right]^2} \quad (5.20)$$

donde

- $\mu_y$       parámetro de ubicación
- $\sigma_y$       parámetro de escala
- coeficiente de asimetría  $> 0$  (Kite, 1977)

De tres parámetros

$$f(x) = \frac{1}{\sqrt{2\pi} (x - x_0) \sigma_y} e^{-1/2 \left[ \frac{\ln(x - x_0) - \mu_y}{\sigma_y} \right]^2} \quad (5.21)$$

donde

- $x_0$       parámetro de ubicación
- $\mu_y$       parámetro de forma
- $\sigma_y$       parámetro de escala
- coeficiente de asimetría  $> 0$  y  $x > x_0$  (Kite, 1977)

### 5.1.4 Distribución General de Valores Extremos Tipo I (Gumbel)

$$F(x) = e^{-e^{-\left(\frac{x-\mu}{\alpha}\right)}} ; \alpha > 0 \quad x > 0 \quad (5.22)$$

- $\mu$       parámetro de ubicación
- $\alpha$       parámetro de escala

### 5.1.5 Distribución General de Valores Extremos

$$F(x) = \exp \left[ - \left[ 1 - \left( \frac{x-\mu}{\alpha} \right)^\beta \right]^{1/\beta} \right] \quad \alpha > 0 \quad x > 0 \quad (5.23)$$

$$f(x) = \frac{1}{\alpha} \left[ 1 - \left( \frac{x-\mu}{\alpha} \right)^\beta \right]^{1/\beta-1} e^{- \left[ 1 - \left( \frac{x-\mu}{\alpha} \right)^\beta \right]^{1/\beta}} \quad -3 \leq \beta \leq 4 \quad (5.24)$$

donde

$\mu$       parámetro de ubicación  
 $\alpha$       parámetro de escala  
 $\beta$       parámetro de forma

Tipo I      Gumbel  
 $\beta \rightarrow 0$  y coeficiente de asimetría = 1.14

Tipo II      Fréchet  
 $\beta < 0$  ;  $\mu + \frac{\alpha}{\beta} \leq x \leq \infty$  y coeficiente de asimetría > 1.14

Tipo III      Weibull  
 $\beta > 0$  ;  $-\infty < x < \mu + \frac{\alpha}{\beta}$  y el coeficiente de asimetría < 1.14

### 5.1.6 Distribución Exponencial

De un parámetro

$$F(x) = 1 - \beta e^{-\beta x} \quad (5.25)$$

$$f(x) = \beta e^{-\beta x} \quad (5.26)$$

donde

$\beta$       parámetro de escala

De dos parámetros

$$F(x) = 1 - e^{- \left( \frac{x-x_0}{\beta} \right)} \quad (5.27)$$



$$f(x) = \frac{1}{\beta} e^{-\left(\frac{x-x_0}{\beta}\right)} \quad (5.28)$$

donde

$x_0$     parámetro de ubicación  
 $\beta$       parámetro de escala

### 5.1.7 Distribución Gamma

De dos parámetros

$$F(x) = \int_0^x \frac{x^{\beta-1} e^{-x/\alpha}}{\alpha^\beta \Gamma(\beta)} \quad \alpha > 0 \quad \hat{\gamma}_x > 0 \quad (5.29)$$

$$f(x) = \frac{x^{\beta-1} e^{-x/\alpha}}{\alpha^\beta \Gamma(\beta)} \quad (5.30)$$

donde

$\alpha$       parámetro de escala  
 $\beta$       parámetro de forma  
 $\hat{\gamma}_x$     coeficiente de asimetría poblacional  
 coeficiente de asimetría muestral  $< 2$  (Kite, 1977)

De tres parámetros (Pearson tipo III)

$$f(x) = \frac{1}{\alpha^\beta \Gamma(\beta)} \left(\frac{x-x_0}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x-x_0}{\alpha}\right)} \quad \alpha > 0 \quad \hat{\gamma}_x > 0 \quad x_0 \leq x < \infty \quad (5.31)$$

donde

$x_0$     parámetro de ubicación  
 $\alpha$       parámetro de escala  
 $\beta$       parámetro de forma  
 $\hat{\gamma}_x$     coeficiente de asimetría poblacional  
 coeficiente de asimetría muestral  $< 2$  (Kite, 1977)

### 5.1.8 Distribución Gumbel Mixta

$$F(x) = p \exp^{-\exp\left(-\frac{x-\mu_1}{\alpha_1}\right)} + (1-p) \exp^{-\exp\left(-\frac{x-\mu_2}{\alpha_2}\right)} \quad (5.32)$$

$$f(x) = \frac{p}{\alpha_1} \exp^{-\exp\left(-\frac{x-\mu_1}{\alpha_1}\right)} \exp\left(-\frac{x-\mu_1}{\alpha_1}\right) + \left(\frac{1-p}{\alpha_2}\right) \exp^{-\exp\left(-\frac{x-\mu_2}{\alpha_2}\right)} \exp\left(-\frac{x-\mu_2}{\alpha_2}\right) \quad (5.33)$$

para  $\alpha_1 > 0$  ;  $x > 0$  ;  $0 < p < 1$

donde

- $\mu_1$  parámetro de ubicación de la 1a. población
- $\alpha_1$  parámetro de escala de la 1a. población
- $\mu_2$  parámetro de ubicación de la 2a. población
- $\alpha_2$  parámetro de escala de la 2a. población
- $p$  relaciona al tamaño de la 1a. población con el total de la muestra ( $N1/N$ )

La función logarítmica de verosimilitud de esta distribución esta dada por:

$$LL = \sum \text{Ln} \left[ \frac{p}{\alpha_1} e^{-e^{-\left(\frac{x-\mu_1}{\alpha_1}\right)}} e^{-\left(\frac{x-\mu_1}{\alpha_1}\right)} + \left(\frac{1-p}{\alpha_2}\right) e^{-e^{-\left(\frac{x-\mu_2}{\alpha_2}\right)}} e^{-\left(\frac{x-\mu_2}{\alpha_2}\right)} \right] \quad (5.34)$$

Esta expresión para resolverse, debe maximizarse. El algoritmo utilizado será el de Rosenbrock que es un modelo de optimización no lineal, multivariado y restringido. El ajuste de una muestra a esta distribución y el cálculo de los eventos de diseño puede realizarse con los programas GMBMIXT1.BAS y GMBMIXT2.BAS que se muestra en el anexo de programas.

## 5.2 Distribuciones bivariadas

El empleo de la estadística en la Hidrología es una herramienta en términos probabilísticos, para determinar confiablemente los escurrimientos que desean inferirse. El análisis de frecuencias a través de distribuciones univariadas han sido muy estudiado y empleado en la Hidrología moderna,

sin embargo, el estudio de las distribuciones multivariadas es un campo nuevo y por tanto poco estudiado. Los pioneros en los fundamentos teóricos de las distribuciones de valores extremos bivariadas, fueron: Finkelstein en 1953, Gumbel y Tiago de Oliveira en 1958.

Una distribución asintótica bivariada tiene cuatro características principales. (Escalante, 1991).

- Las distribuciones marginales son asintóticas extremas.
- Cumple el postulado de estabilidad (distribución estable para los valores más grandes de la muestra).
- Posee una función de densidad.
- Se elimina el caso trivial donde la distribución multivariada es el producto de sus distribuciones marginales extremas.

Si se toman dos variables aleatorias continuas e independientes  $X$ ,  $Y$ , con la misma distribución conjunta, perteneciente a alguno de los casos de la distribución General de Valores Extremos (apartado 5.1.4), es posible asignarles transformaciones lineales de la forma:

$$X_n = a_n X_{\max} + b_n \quad (5.35)$$

$$Y_n = c_n Y_{\max} + d_n \quad (5.36)$$

La distribución conjunta de los valores máximos de  $X$  y  $Y$  puede expresarse por:

$$F^n(x, y) = \Pr(\max X_i \leq x, \max Y_i \leq y) \quad (5.37)$$

Para cuando  $n \rightarrow \infty$  la función acumulada de esta distribución bivariada de valores extremos será:

$$F(x, y) = \lim_{n \rightarrow \infty} (a_n x + b_n, c_n y + d_n) \quad (5.38)$$

Dentro de las distribuciones bivariadas existen dos tipos: los modelos que cuentan con una función de densidad (diferenciables) y los no diferenciables. La estimación de parámetros se realiza de una manera más fácil en el caso de los modelos diferenciables, entre los que se encuentran los modelos logístico y el mezclado.

El modelo logístico considera que la diferencia de las variables reducidas tienen una distribución logística estándar, siempre y cuando sus marginales sean distribuciones Gumbel. Debido a que pueden manejarse amplios rangos de coeficientes de correlación con este modelo, es ampliamente aconsejable en la solución de problemas de análisis de frecuencias en Hidrología (Escalante, 1991).

La forma general de este modelo es:

$$F(x, y, z, \dots, \theta) = \exp\left\{-\left[(-\ln F(x))^m + (-\ln F(y))^m + (-\ln F(z))^m + \dots\right]^{1/m}\right\} \quad (5.35)$$

donde

$F(x)$ ,  $F(y)$  y  $F(z)$       distribuciones marginales.  
 $m$                                       parámetro de asociación.

Para  $m=1$  la función de distribución bivariada se reduce al producto de sus marginales:

$$F(x, y, m) = F(x)F(y)$$

La expresión (5.35) debe satisfacer las siguientes desigualdades:

$$F(x_1), \dots, F(x_n) \leq F(x_1, \dots, x_n) \leq \min[F(x_1), \dots, F(x_n)]$$

$$\left[ \prod_{i \neq j} F(x_i, x_j) \right]^{1/2(n-1)} \leq F(x_1, \dots, x_n) \leq \frac{\left[ \prod_{i,j} F(x_i, x_j) \right]^{1/2}}{\left[ \prod_i F(x_i) \right]^{n-2}}$$

Las marginales pueden ser:

Gumbel

$$F(s) = \exp^{-\exp\left(-\frac{s-\mu}{\alpha}\right)}$$

General de Valores Extremos

$$F(s) = \exp\left[-\left(1 - \left(\frac{s-\mu}{\alpha}\right)^\beta\right)^{1/\beta}\right]$$

Gumbel Mixta

$$F(s) = p \exp^{-\exp\left(-\frac{s-\mu_1}{\alpha_1}\right)} + (1-p) \exp^{-\exp\left(-\frac{s-\mu_2}{\alpha_2}\right)}$$

Finalmente, si se desea el caso bivariado "Bigumbel" por ejemplo, se tendría:

$$F(x, y, \mu_1, \alpha_1, \mu_2, \alpha_2, m) = \exp\left[-\left[e^{-m\left(\frac{x-\mu_1}{\alpha_1}\right)} + e^{-m\left(\frac{y-\mu_2}{\alpha_2}\right)}\right]^{1/m}\right]$$

$$f(x, y, \underline{\theta}) = \frac{1}{\alpha_1 \alpha_2} e^{-m\left(\frac{x-\mu_1}{\alpha_1}\right)} e^{-m\left(\frac{y-\mu_2}{\alpha_2}\right)} \left[e^{-m\left(\frac{x-\mu_1}{\alpha_1}\right)} + e^{-m\left(\frac{y-\mu_2}{\alpha_2}\right)}\right]^{1/m-2} \dots$$

$$\dots \left\{ (m-1) + \left[e^{-m\left(\frac{x-\mu_1}{\alpha_1}\right)} + e^{-m\left(\frac{y-\mu_2}{\alpha_2}\right)}\right]^{1/m} \right\} \exp\left\{-\left[e^{-m\left(\frac{x-\mu_1}{\alpha_1}\right)} + e^{-m\left(\frac{y-\mu_2}{\alpha_2}\right)}\right]^{1/m}\right\}$$

donde el parámetro de asociación  $m=2$

La correspondiente función de verosimilitud de esta expresión estará dada por:

$$L(x, y, \underline{\theta}) = \left[\prod_{i=1}^{n_1} f(p_i, \underline{\theta}_1)\right]^{I_1} \left[\prod_{i=1}^{n_2} f(p_i, q_i, \underline{\theta}_2)\right]^{I_2} \dots$$

$$\dots \left[\prod_{i=1}^{n_3} f(x, y, \underline{\theta}_3)\right]^{I_3} \left[\prod_{i=1}^{n_4} f(r_i, s_i, \underline{\theta}_4)\right]^{I_4} \left[\prod_{i=1}^{n_5} f(r_i, \underline{\theta}_1)\right]^{I_5}$$

donde

- L( ) función de verosimilitud
- $n_1$  longitud del registro univariado antes del periodo común  $n_3$
- $n_2$  longitud del registro con relación bivariada antes del periodo común  $n_3$
- $n_3$  longitud del registro durante el periodo común en las dos estaciones
- $n_4$  longitud del registro con relación bivariada después del periodo común  $n_3$
- $n_5$  longitud del registro univariado después del periodo común  $n_3$
- p variable del registro univariado antes del registro común
- (p,q) variables relacionadas en forma bivariada antes del registro común
- (x,y) variables relacionadas en forma bivariada durante el periodo común
- r variable del registro univariado después del registro común
- (r,s) variables relacionadas en forma bivariada después del registro común
- $I_i$  número indicador con valor:  
 $I_i=1$  si  $n_i>0$  o  $I_i=0$  si  $n_i=0$
- $\underline{\theta}_i$  vector de parámetros

Recordando la propiedad logaritmica de la función de verosimilitud, se tiene:

$$LL(x, y, \underline{\theta}) = I_1 \left[ \sum_{i=1}^{n_1} \text{Ln } f(p_i, \underline{\theta}_1) \right] + I_2 \left[ \sum_{i=1}^{n_2} \text{Ln } f(p_i, q_i, \underline{\theta}_2) \right] + \dots$$

$$\dots + I_3 \left[ \sum_{i=1}^{n_3} \text{Ln } f(x, y, \underline{\theta}_3) \right] + I_4 \left[ \sum_{i=1}^{n_4} \text{Ln } f(r_i, s_i, \underline{\theta}_4) \right] + I_5 \left[ \sum_{i=1}^{n_5} \text{Ln } f(r_i, \underline{\theta}_5) \right]$$

Los estimadores de máxima verosimilitud de los parámetros de distribuciones bivariadas son aquellos para los cuales esta ecuación es maximizada. Para obtener dichos estimadores se emplea el algoritmo de optimización no lineal multivariado restringido de Rosenbrock (Rosenbrock, 1960).

## **6. APLICACIONES**

La metodología que se propone en este capítulo esta encaminada a obtener eventos de diseño a partir de modelos de transferencia de información en dos tipos de cuencas. Las cuencas aforadas, en donde la transferencia se realiza a través de la estimación de parámetros en forma bivariada, y ampliando los registros cortos con un modelo de correlación lineal simple. En lo que se refiere a las cuencas no aforadas, se obtienen tres tipos de ecuaciones regionales. Estas ecuaciones relacionan las características fisiográficas de la región, con los parámetros de la distribución de probabilidad obtenidos en forma univariada, bivariada y de los registros extendidos. Por su importancia a nivel nacional se escogió la región hidrológica 10 para llevar a cabo esta aplicación.

### **6.1 Descripción del área de estudio**

La región hidrológica 10 se encuentra ubicada al noroeste de la República Mexicana, entre los  $105^{\circ}30'$  y  $109^{\circ}$  de longitud oeste, y los  $23^{\circ}30'$  y  $28^{\circ}30'$  de latitud norte. Es una de las regiones más importantes en el país, tanto por lo que se refiere al grupo de corrientes caudalosas que reúne, como también en lo referente a los usos que se le da a tales caudales; lo que ha dado lugar a la creación de una zona que figura entre las más productoras del país desde el punto de vista agrícola.

Para facilitar su estudio la región se divide en 9 grandes cuencas, las cuales son identificadas por el nombre del río principal al que drenan Cuencas de los ríos Fuerte, Sinaloa, Mocorito, de la Laguna de Caimanero, Culiacán, San Lorenzo, Elota, Piaxtla y Quelite. La tabla 6.1 muestra las 42 estaciones que son utilizadas en este trabajo, así como la cuenca a la que pertenecen, las coordenadas de ubicación, su localización por estado. Los gastos máximos anuales disponibles se muestran en la tabla 6.2. El plano 2 muestra la ubicación de la región así como la localización exacta de cada una de las estaciones.

## 6.2 Metodología

A continuación se presenta el detalle de la metodología a seguir y un cuadro resumen en donde se enumeran los principales puntos del desarrollo.

- a) Como primer paso se obtienen los estadísticos muestrales de cada uno de los registros, los que se pueden ver en la tabla 6.3.
- b) Determinar las características fisiográficas de la región: AREA, PRECMEDA, LONGITUD, PENDCAUC, FRECUENC Y CENTLON (ver capítulo 2).
- c) A continuación se lleva a cabo un análisis de frecuencias en todas las estaciones para determinar la distribución de probabilidad que identifique a la región (Gumbel Mixta, ver subcapítulo 6.6).
- d) Se obtienen los eventos de diseño en cada sitio, para los periodos de retorno de 2, 5, 10, 20, 50, 100, 500, 1000, 5000, y 10000 años.
- e) Normalizar las muestras de las 15 características fisiográficas y los gastos máximos para los periodos de retorno mencionados en el inciso d.
- f) Discriminar las características fisiográficas que deben intervenir en el modelo de correlación múltiple. Este procedimiento se lleva a cabo a través de un análisis de varianza, para lo cual se correlacionan estas características con los gastos máximos para los periodos de retorno mencionados.
- g) Aplicar las técnicas de homogeneidad de cuencas (ver capítulo 2).
- h) Llevar a cabo una transferencia de información para cuencas aforadas. Este procedimiento puede realizarse si se emplean las distribuciones univariadas de probabilidad, por diferentes métodos regionales, por ejemplo: la avenida índice, estaciones año, o Box Cox regional. Sin embargo, el desarrollo de estos métodos por su gran variedad, debe ser motivo de un estudio de regionalización. En este trabajo la transferencia de información para cuencas aforadas se lleva a cabo por dos caminos:
  - Estimar parámetros en forma bivariada.



- Extender registros en las estaciones en donde se mejore el contenido de información.

i) Llevar a cabo una transferencia de información para cuencas no aforadas, desarrollando ecuaciones regionales con un modelo de correlación lineal múltiple, en donde se relacionen las características fisiográficas significativas de la región con los parámetros de su distribución de probabilidad obtenidos en forma:

- Univariada
- Bivariada
- De registros extendidos

j) Estimar él o los eventos de diseño para el período de retorno requerido, ya sea para cuencas aforadas o no.

## REGION HIDROLOGICA 10

### CARACTERISTICAS FISIOGRAFICAS

#### SELECCION DE VARIABLES SIGNIFICATIVAS

- Eventos de diseño para diferentes periodos de retorno
- Normalización de muestras
- Discriminar variables

#### DELIMITAR LAS REGIONES HIDROLOGICAMENTE HOMOGENEAS

- Con base en las características fisiográficas
- Con base en los coeficientes de variación de los gastos máximos anuales
- Con base en la frecuencia de los gastos máximos anuales

### TRANSFERENCIA DE INFORMACION

#### CUENCAS AFORADAS

- Métodos regionales
- Estimación bivariada de parámetros
- Extensión de registros

#### CUENCAS NO AFORADAS (Ecuaciones regionales)

Correlacionando fisiografía con:

- Parámetros univariados
- Parámetros bivariados
- Parámetros de registros extendidos

### HIDROGRAMA DEL EVENTO DE DISEÑO

**TABLA 6.1**  
**PRINCIPALES ESTACIONES DE LA REGION**  
**HIDROLOGICA 10**

**CUENCA DEL RIO FUERTE**

ESTACION	ESTADO	LATIT	LONG
SM ZAPOTTLAN	Sin.	25°57'	109°03'
SAN BLAS	Sin.	26°06'	108°45'
LA TINA	Sin.	26°12'	108°37'
BAMICORI	Sin.	26°23'	108°30'
LAS CAÑAS	Sin.	26°28'	108°35'
EL MAHONE	Sin.	26°30'	108°35'
CHOIX	Sin.	26°44'	108°20'
HUITES	Sin.	26°54'	108°21'
PALO DULCE	Chih.	27°01'	108°24'
CHINIPAS	Chih.	27°25'	108°33'
SAN FRANCISCO	Sin.	26°54'	108°08'
SAN IGNACIO	Chih.	26°52'	107°53'
LA VERANERA	Chih.	27°14'	107°53'
URIQUE	Chih.	27°18'	107°50'
GUERACHIC	Chih.	26°38'	107°15'
ALAMOS	Chih.	27°42'	107°45'
CAZANATE	Son.	26°36'	108°46'

**CUENCA DEL RIO CULIACAN**

ESTACION	ESTADO	LATIT	LONG
TIERRA BLANCA	Sin.	24°50'	107°24'
PTE. S PACIFICO	Sin.	24°48'	107°24'
PALOS BLANCOS	Sin.	24°56'	107°23'
EL VAREJONAL	Sin.	25°05'	107°24'
BADIRAGUATO	Sin.	25°20'	107°32'
GUATENIPA	Sin.	25°21'	107°14'
LA HUERTA	Sin.	25°22'	106°42'
PTE. CAÑEDO	Sin.	24°49'	107°23'
SANALONA	Sin.	25°20'	107°14'
PICACHOS	Sin.	24°49'	107°09'
TAMAZULA	Dgo.	24°56'	106°59'
EL BLEDAL	Sin.	24°48'	107°09'

**CUENCA DEL RIO SAN LORENZO**

ESTACION	ESTADO	LATIT	LONG
SANTA CRUZ	Sin.	24°29'	106°57'

**CUENCA DEL RIO SINALOA**

ESTACION	ESTADO	LATIT	LONG
JAINA	Sin.	25°54'	108°00'
TOAHAYANA	Sin.	26°10'	107°42'
TECUSIAPA	Sin.	25°51'	107°23'
LOS MOLINOS	Sin.	25°51'	107°23'
NARANJO	Sin.	25°48'	108°28'
ZOPILOTE	Sin.	25°45'	108°22'

**CUENCA DEL RIO ELOTA**

ESTACION	ESTADO	LATIT	LONG
ACATITAN	Sin.	24°03'	106°39'

**CUENCA DEL RIO PIAXTLA**

ESTACION	ESTADO	LATIT	LONG
IXPALINO	Sin.	23°57'	106°36'
PIAXTLA	Sin.	23°56'	106°26'

**CUENCA DEL RIO MOCORITO**

ESTACION	ESTADO	LATIT	LONG
GUAMUCHIL	Sin.	25°28'	108°06'

**CUENCA DE LA LAGUNA DE CAIMANERO**

ESTACION	ESTADO	LATIT	LONG
PERICOS	Sin.	25°06'	107°42'

**CUENCA DEL RIO QUELITE**

ESTACION	ESTADO	LATIT	LONG
EL QUELITE	Sin.	23°31'	106°29'





**TABLA 6.3**  
**ESTADISTICOS DE LAS MUESTRAS**  
**ORIGINALES**

**CUENCA DEL RIO FUERTE**

ESTACION	DATOS	MEDIA	DESV.EST.	CV
SM ZAPOTITLAN	22	625	645	1.033
SAN BLAS	10	4018	3857	0.960
LA TINA	24	107	152	1.429
BAMICORI	32	195	177	0.908
LAS CAÑAS	24	2122	2776	1.308
EL MAHONE	20	198	26	0.132
CHOIX	29	393	337	0.856
HUITES	53	2942	3124	1.062
PALO DULCE	29	976	1196	1.225
CHINIPAS	21	1061	524	0.494
SAN FRANCISO	33	1725	1450	0.841
SAN IGNACIO	19	1622	813	0.501
LA VERANERA	14	413	341	0.826
URIQUE	19	303	148	0.489
GUERACHIC	13	890	1391	1.564
ALAMOS	22	248	179	0.722
CAZANATE	19	555	728	1.312

**CUENCA DEL RIO CULIACAN**

ESTACION	DATOS	MEDIA	DESV.EST.	CV
TIERRA BLANCA	7	1544	554	0.359
PTE. S PACIFICO	35	2960	2205	0.745
PALOS BLANCOS	47	1482	1726	1.165
EL VAREJONAL	7	674	626	0.929
BADIRAGUATO	27	958	1854	1.935
GUATENIPA	21	1889	1393	0.738
LA HUERTA	17	934	574	0.615
PTE. CAÑEDO	22	1116	933	0.836
SANALONA	42	447	505	1.130
PICACHOS	7	1128	419	0.372
TAMAZULA	23	584	278	0.476
EL BLEDAL	48	286	273	0.956

**CUENCA DEL RIO SAN LORENZO**

ESTACION	DATOS	MEDIA	DESV.EST.	CV
SANTA CRUZ	43	1268	1216	0.959

**CUENCA DEL RIO SINALOA**

ESTACION	DATOS	MEDIA	DESV.EST.	CV
JAINA	46	1197	1190	0.994
TOAHAYANA	29	1049	630	0.600
TECUSIAPA	17	976	793	0.813
LOS MOLINOS	13	159	73	0.461
NARANJO	47	622	656	1.054
ZOPILOTE	47	363	276	0.760

**CUENCA DEL RIO ELOTA**

ESTACION	DATOS	MEDIA	DESV.EST.	CV
ACATITAN	31	1032	865	0.838

**CUENCA DEL RIO PIAXTLA**

ESTACION	DATOS	MEDIA	DESV.EST.	CV
IXPALINO	31	1318	1218	0.925
PIAXTLA	16	1420	1588	1.118

**CUENCA DEL RIO MOCORITO**

ESTACION	DATOS	MEDIA	DESV.EST.	CV
GUAMUCHIL	36	606	630	1.040

**CUENCA DE LA LAGUNA DE CAIMANERO**

ESTACION	DATOS	MEDIA	DESV.EST.	CV
PERICOS	26	202	95	0.471

**CUENCA DEL RIO QUELITE**

ESTACION	DATOS	MEDIA	DESV.EST.	CV
EL QUELITE	26	469	445	0.950

### 6.3 Normalización de muestras

Uno de los problemas más importantes en el análisis de regresión regional es, sin duda, encontrar que las muestras de los escurrimientos no se encuentran distribuidas en forma normal. De esta forma no podrá aplicarse ningún procedimiento de regresión a estas series de eventos extremos.

Tal y como se comentó en el punto 4.1; debe realizarse una normalización de muestras a través de algún algoritmo de transformación. Empleando la transformación normalizante de Box-Cox (ecuaciones 4.9 y 4.10) se desarrolló el programa BOXCOX.BAS. Primeramente, fueron normalizados los escurrimientos máximos instantáneos ajustados a una distribución del tipo Doble Gumbel para los períodos de retorno de 2, 5, 10, 20, 50, 100, 500, 1 000, 5 000 y 10 000 años calculados para cada estación a través de un análisis de frecuencias (ver apartado 6.6), logrando un *coeficiente de asimetría* promedio de -0.03 y un valor de 3.64 para la *curtosis* (muestras estadísticamente normales). El resumen de los estadísticos de las muestras normalizadas, así como su respectivo parámetro  $\lambda$  de transformación se muestran en la tabla 6.5. También el programa fue aplicado en las series de 15 características fisiográficas para las 42 estaciones de la región, ver Tabla 6.4.

### 6.4 Discriminación de variables

Una vez que las muestras y características se encuentran normalizadas, es posible realizar una selección de las variables que deben emplearse. Un análisis de correlación múltiple y el estadístico "t" tal y como se desarrolló en el apartado 4.3, darán el criterio para saber cuáles variables contribuyen significativamente en la ecuación de regresión.

El primer análisis consistió en correlacionar las muestras normalizadas de las 15 características fisiográficas con cada uno de los gastos; para los períodos de retorno antes mencionados. Por facilidad, de aquí en adelante se empleará como ejemplo los resultados de la correlación con el gasto de 100 años de período de retorno. El primer paso consiste en proponer una ecuación del tipo  $Y = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$  en donde "Y" será la variable dependiente, en este caso el gasto de 100 años ( $Q_{100}$ ) y las características fisiográficas serán las variables independientes. La tabla 6.6 muestra los resultados de la regresión, para este gasto específico.

El siguiente paso es probar que la ecuación de regresión explica completamente el significado de la variable "Y"; esto deberá probarse al proponer la hipótesis nula  $H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$  equivalente a decir que la ecuación de regresión *no explica el significado de la variable "Y"*.

TABLA 6.4 RESUMEN DE LAS CARACTERÍSTICAS FISIOGRAFICAS DE LA REGION HIDROLOGICA 10

MUESTRAS NORMALIZADAS

ESTACION / CARACTERISTICA	λ Box Cox =															
	0.14	2.40	1.72	1.27	0.27	0.33	-0.51	-0.41	-0.43	3.24	1.00	2.47	1.70	4.14		
AREA	PHECMECA	PENDCUEN	ELEVEMED	LONGITUD	PENDCAUC	PARAFORM	FRECUENC	DENSIDAD	LATIT	LONG	COEFESCK	ORDEN	PERIMET	TEMP		
(km2)	(mm)	(m km-1)	(msnm)	(km)	(m km-1)	(km-1)	(km-2)	(km-1)	(grados)	(grados)			(km)	(°C)		
1	SM ZAPOTITLAN	23.69449	3743878	-0.5790496	9607.153	15.455060	-2.563026	-9.105381	-19.71671	-12.82509	13286.370	107.80	-0.3983353	8.486	13.63319	104845.7
2	SAN BLAS	23.58554	3859917	-0.5789664	9844.762	14.599340	-2.560140	-9.312546	-19.71671	-13.25143	13318.370	107.82	-0.3983353	8.486	13.42742	104845.7
3	LA TINA	8.53830	2389610	-0.5808876	1612.310	5.599348	-2.535142	-2.375831	-22.05911	-4.80734	12109.430	108.52	-0.4000617	1.323	5.84861	104845.7
4	BAMICORI	8.08484	3870632	-0.5813068	1663.061	4.634660	-2.506403	-1.457429	-19.91727	-4.20522	12411.860	108.40	-0.4003422	1.323	6.37259	104845.7
5	LAS CAÑAS	23.03618	4163366	-0.5786440	10420.860	14.195040	-2.489959	-9.242221	-19.15017	-13.36442	13366.460	107.68	-0.3979871	8.486	13.04758	104845.7
6	EL MAHONÉ	23.02171	4174810	-0.5786440	10420.860	14.124000	-2.486701	-9.277221	-19.15017	-13.42194	13366.460	107.68	-0.3979871	8.486	13.01754	104845.7
7	CHOIX	12.55683	6169428	-0.5805729	5408.032	8.468395	-2.619793	-3.294437	-18.97209	-6.01471	12719.460	108.07	-0.3996220	3.219	8.63291	71942.4
8	HUITES	22.50638	4135119	-0.5781126	11585.790	13.037760	-2.445611	-9.312546	-18.79955	-14.74390	13446.890	107.58	-0.3975366	8.486	12.60782	17865.3
9	PALO DULCE	17.17919	5285010	-0.5771933	10287.400	11.436660	-2.361279	-4.176214	-20.12421	-8.68318	14387.550	108.20	-0.3962847	5.621	10.32275	17865.3
10	CHINIPAS	16.45687	5261631	-0.5767946	11794.260	10.109640	-2.281035	-5.936973	-19.52221	-9.34178	14557.140	108.15	-0.3961835	5.621	9.98999	17865.3
11	SAN FRANCISCO	20.91171	5564936	-0.5779903	12382.140	12.671550	-2.435458	-8.446182	-17.70802	-12.67491	13079.680	107.80	-0.3980752	5.621	11.87712	148070.7
12	SAN IGNACIO	15.17274	3891391	-0.5776036	12278.400	11.966510	-2.358499	-7.102558	-16.55103	-10.66648	12719.460	107.17	-0.3987555	5.621	11.27563	71942.4
13	LA VERANERA	19.21261	3203433	-0.5794555	15083.720	10.277680	-2.261749	-4.924246	-19.15017	-8.00311	13886.980	107.40	-0.3957716	3.219	10.04128	4947.0
14	URIQUE	15.66893	3190886	-0.5795057	15525.890	9.994469	-2.266314	-4.961637	-18.97209	-8.17878	13886.980	107.37	-0.3957716	3.219	9.93914	4947.0
15	GUERACHIC	17.14620	4085985	-0.5788718	14820.110	10.493570	-2.316875	-6.536453	-18.46663	-9.77251	12411.860	106.80	-0.3992379	5.621	10.06440	29995.3
16	ALAMOS	13.92958	2214160	-0.5809344	2134.088	9.986340	-2.651100	-4.102860	-22.34370	-5.96625	13032.300	108.80	-0.4007442	3.219	9.13548	104845.7
17	CAZANATE	13.27672	2384200	-0.5809410	2413.531	8.889206	-2.671206	-4.027864	-22.94885	-6.43151	13158.910	108.82	-0.4004107	3.219	8.63291	104845.7
18	JAINA	18.07155	6331225	-0.5778176	6949.378	11.278940	-2.364299	-6.567353	-17.95998	-9.21598	11856.410	107.43	-0.3984207	5.621	10.78619	71942.4
19	TOAHAYANA	16.57367	6657468	-0.5770642	9253.518	10.109640	-2.340063	-6.394075	-16.91107	-9.34178	11812.140	107.27	-0.3978984	5.621	9.69896	47537.5
20	TECUSIAPAN	15.48311	7261533	-0.5758885	11437.760	8.778211	-2.271767	-4.887447	-18.46663	-9.67743	11856.410	107.12	-0.3975366	5.621	9.84192	17865.3
21	LOS MOLINOS	9.91202	5405375	-0.5789381	7093.356	5.968697	-2.175633	-2.569320	-28.73514	-5.26310	11462.090	107.27	-0.3988375	1.323	6.92036	104845.7
22	NARANJO	13.65078	3971042	-0.5806459	2580.053	9.375138	-2.424750	-3.878155	-24.33008	-6.22604	12154.470	108.30	-0.4006127	3.219	9.08357	104845.7
23	ZOMLOTE	10.60548	3240421	-0.5801326	1778.454	7.868423	-2.443754	-5.866119	-23.96057	-4.31680	11856.410	108.23	-0.4006127	1.323	7.53838	148070.7
24	GUAMUCHIL	13.00062	3495636	-0.5811979	662.274	7.728301	-2.441519	-3.888302	-23.96057	-7.13149	11175.920	107.63	-0.4014241	3.219	8.50879	148070.7
25	PERICOS	8.49807	3146456	-0.5810628	456.121	5.626894	-2.466217	-1.807351	-21.78544	-3.65681	10714.600	107.62	-0.4009368	1.323	6.37259	148070.7
26	TERRA BLANCA	19.34020	6248375	-0.5787568	8805.601	13.188290	-2.469270	-6.835650	-23.96057	-9.79668	11062.850	106.85	-0.3989997	5.621	11.62813	9879.0
27	PTE. S PACIFICO	19.28235	6267925	-0.5787568	8805.601	13.054670	-2.460506	-6.852725	-23.60804	-9.82100	11062.850	106.85	-0.3989997	5.621	11.56928	9879.0
28	PALIS BLANCOS	19.27426	6326443	-0.5787568	8965.694	12.848670	-2.455493	-7.001932	-23.60804	-9.91995	11062.850	106.85	-0.3989189	5.621	11.46899	9879.0
29	EL VAREJONAL	19.13523	6451455	-0.5783705	9626.599	12.433250	-2.446573	-6.901543	-24.33008	-10.42963	11105.160	106.80	-0.3988375	5.621	11.41777	9879.0
30	BADRAGUAT	11.69172	6345079	-0.5777985	2628.445	7.433235	-2.181043	-3.246923	-24.33008	-5.64529	11275.510	107.50	-0.3984207	1.323	7.90517	104845.7
31	GUATEMPA	18.10294	6704508	-0.5781412	12219.780	11.301740	-2.325849	-6.314806	-23.27118	-10.34444	11133.430	106.58	-0.3985894	5.621	10.97866	29995.3
32	LA HUERTA	17.08436	6924025	-0.5789866	13977.560	9.694477	-2.261101	-6.063661	-22.05911	-11.05096	11175.920	106.42	-0.3988375	5.621	10.32275	9879.0
33	PTE. CAÑEJO	15.73697	7347580	-0.5787756	3675.954	6.994477	-2.513559	-5.595152	-21.78544	-3.84888	10441.520	106.90	-0.3985894	3.219	9.54027	148070.7
34	SANALONA	15.38441	7954912	-0.5778270	4416.497	8.587059	-2.525445	-5.869770	-22.05911	-10.12618	10468.610	106.83	-0.3980752	3.219	9.07309	148070.7
35	PICACHO	15.04387	8417021	-0.5778270	4416.497	8.587059	-2.525445	-5.494068	-21.02337	-9.84549	10468.610	106.83	-0.3980752	3.219	9.04140	148070.7
36	TAMAZULU	13.89168	9664219	-0.5756256	6046.539	6.829754	-2.362139	-4.940790	-22.05911	-10.40102	10509.330	106.77	-0.3976281	3.219	8.47051	104845.7
37	EL BLEDAL	9.20963	4548620	-0.5806649	951.777	5.656894	-2.414667	-2.123551	-25.12597	-4.53110	10040.970	107.00	-0.3985894	1.323	6.71017	148070.7
38	SANTA CRUZ	18.37916	7748503	-0.5776582	9895.614	11.347080	-2.313498	-6.559597	-14.78581	-10.04750	10040.970	106.18	-0.3973516	5.621	11.03667	9879.0
39	ACATIPAN	13.38684	7366231	-0.5780702	3389.842	9.632250	-2.339861	-3.998088	-15.21932	-6.00077	9549.019	106.55	-0.3972581	1.323	8.75205	4947.0
40	IXPALINO	17.09372	6710842	-0.5788016	7748.808	11.209250	-2.305560	-6.329053	-16.21303	-8.77236	9309.409	106.03	-0.3967801	5.621	10.15528	148070.7
41	PIAXTLA	16.58999	7335306	-0.5780759	9561.494	10.413660	-2.242165	-5.881863	-16.68847	-9.25741	9334.434	105.98	-0.3964849	5.621	10.04128	4947.0
42	EL QUELITE	11.17638	3002387	-0.5812399	706.801	6.456751	-2.476740	-2.949547	-13.94745	-6.00077	8722.648	106.37	-0.3988375	1.323	7.71111	148070.7

**TABLA 6.5**  
**ESTADISTICOS DE LAS MUESTRAS**  
**NORMALIZADAS**

**CUENCA DEL RIO FUERTE**

ESTACION	$\lambda$	MEDIA	COEF.ASIM	CURTOSIS
SM ZAPOTITLAN	0.18	10.82	-0.04	2.94
SAN BLAS	-0.67	1.48	-0.04	3.52
LA TINA	0.20	6.53	0.00	5.41
BAMICORI	-0.13	3.60	-0.03	2.89
LAS CAÑAS	-0.11	4.84	-0.02	3.28
EL MAHONE	-0.33	2.49	-0.04	4.78
CHOIX	-0.07	4.70	-0.04	4.01
HUITES	-0.38	2.48	-0.04	3.26
PALO DULCE	-0.33	2.67	-0.04	4.17
CHINIPAS	0.52	67.82	-0.04	2.51
SAN FRANCISO	-0.48	2.01	-0.04	2.57
SAN IGNACIO	-0.78	1.27	-0.04	2.67
LA VERANERA	-0.10	4.35	-0.04	3.53
URIQUE	0.43	24.02	-0.04	5.06
GUERACHIC	-0.71	1.38	-0.03	4.37
ALAMOS	0.46	23.30	-0.03	3.26
CAZANATE	0.06	6.95	-0.01	5.15

**CUENCA DEL RIO CULIACAN**

ESTACION	$\lambda$	MEDIA	COEF.ASIM	CURTOSIS
TIERRA BLANCA	0.58	118.46	-0.04	7.18
PTE. S PACIFICO	0.36	43.29	-0.04	2.97
PALOS BLANCOS	0.08	8.71	-0.04	2.07
EL VAREJONAL	0.15	9.93	-0.04	4.18
BADIRAGUATO	-0.27	2.98	-0.03	6.41
GUATENIPA	-0.04	6.29	-0.04	2.29
LA HUERTA	0.49	53.27	-0.04	2.36
PTE. CANEDO	0.44	42.50	-0.04	2.61
SANALONA	-0.05	4.95	-0.03	3.88
PICACHO	0.87	515.34	-0.04	5.74
TAMAZULA	-0.30	2.82	-0.04	3.55
EL BLEDAL	-0.01	5.18	-0.04	3.43

**CUENCA DEL RIO SAN LORENZO**

ESTACION	$\lambda$	MEDIA	COEF.ASIM	CURTOSIS
SANTA CRUZ	0.12	10.63	-0.03	4.56

**CUENCA DEL RIO SINALOA**

ESTACION	$\lambda$	MEDIA	COEF.ASIM	CURTOSIS
JAINA	-0.27	3.10	-0.03	3.88
TOAHAYANA	0.25	17.96	-0.04	2.33
TECUSIAPA	-0.15	4.18	-0.04	3.41
LOS MOLINOS	0.98	145.41	-0.04	3.37
NARANJO	0.05	6.93	-0.04	2.82
ZOPILOTE	0.50	32.89	-0.03	2.41

**CUENCA DEL RIO ELOTA**

ESTACION	$\lambda$	MEDIA	COEF.ASIM	CURTOSIS
ACATITAN	-0.60	1.63	-0.04	2.54

**CUENCA DEL RIO PIAXTLA**

ESTACION	$\lambda$	MEDIA	COEF.ASIM	CURTOSIS
IXPALINO	-0.21	3.63	-0.02	4.03
PIAXTLA	-0.99	1.00	0.28	3.13

**CUENCA DEL RIO MOCORITO**

ESTACION	$\lambda$	MEDIA	COEF.ASIM	CURTOSIS
GUAMUCHIL	0.22	12.80	0.00	4.83

**CUENCA DE LA LAGUNA DE CAIMANERO**

ESTACION	$\lambda$	MEDIA	COEF.ASIM	CURTOSIS
PERICOS	0.82	91.92	-0.04	2.57

**CUENCA DEL RIO QUELITE**

ESTACION	$\lambda$	MEDIA	COEF.ASIM	CURTOSIS
EL QUELITE	-0.10	4.35	-0.03	2.84



Esto conduce a calcular la prueba  $F$  de la ecuación (3.56):

$$F = \frac{21.639}{12.671} = 1.7078$$

El valor crítico de  $F_{0.95,15,26} = 2.0711$  y como  $1.7078 < 2.0711$  **SE ACEPTA  $H_0$**  es decir *la ecuación de regresión no explica el significado de la variable "Y"*.

El siguiente paso consiste en determinar cuál(es) de las variables contribuyen en la ecuación. Esto es lo mismo que proponer las hipótesis  $H_0: \beta_i = 0$  y  $H_1: \beta_i \neq 0$

Cabe recordar que si el modelo es correcto, y el valor de  $\hat{\beta}_i / S\hat{\beta}_i$  sigue la distribución "t" con  $(n-p)$  grados de libertad.  $S\hat{\beta}_i$  es un estimador de  $\sigma\hat{\beta}_i$  y es igual al valor positivo de la raíz cuadrada de  $C_{ii}^2 S^2$ . En donde este término  $(i,i)$  se toma de la matriz varianza covarianza; así:

$$t = \frac{\hat{\beta}_i - \beta_0}{S\hat{\beta}_i}$$

para la primera variable, se tiene:

$$t = \frac{-1.565 - 0}{(13.479 * 12.671)^{1/2}} = -0.11974$$

El valor crítico de  $t_{0.975,15,26} = 2.0555$  y como  $-0.11974 < 0 < 2.0555$  **SE ACEPTA  $H_0$** . Por lo tanto, *ésta variable no contribuye a explicar el gasto de 100 años de periodo de retorno*. Sin embargo, al observar la tabla 6.6 se concluye que en un principio ninguna de las 15 variables contribuye en la ecuación de regresión, esto es evidente ya que en un principio se probó que la ecuación que se propone no explica la variación de la variable "Y". Este mismo procedimiento se realiza con los 9 gastos restantes, con el objeto de determinar cuáles variables deben ser removidas del modelo; de aquí en adelante se buscará en un principio, que el estadístico  $F$ , sea mayor que el valor crítico de  $F$  para rechazar la hipótesis  $H_0$  esto es, que la ecuación explica correctamente la variable dependiente  $Y$ . Y después garantizar que las variables que permanezcan en el modelo contribuyan con éste. Esta prueba se realizará al considerar que su estadístico "t" debe ser mayor que cero y lo más cercano posible al valor crítico de "t". Esto siempre y cuando se haya rechazado  $H_0$  con base en la prueba  $F$ . Se muestran a continuación las 15 características en competencia.

VARIABLE	DESCRIPCION	UNIDADES
1 AREA	Area drenada hasta la estación	km <sup>2</sup>
2 PRECMEDA	Precipitación media anual	mm
3 PENDCUENC	Pendiente media de la cuenca	m km <sup>-1</sup>
4 ELEV MED	Elevación media de la cuenca	m
5 LONGITUD	Longitud de la corriente principal	km
6 PENDCAUC	Pendiente de la corriente principal	m km <sup>-1</sup>
7 PARMFORM	Parámetro de forma	km <sup>-1</sup>
8 FRECUENC	Frecuencia de corriente	km <sup>-2</sup>
9 DENSIDAD	Densidad de drenaje	km <sup>-1</sup>
10 CENTLAT	Latitud del centroide de la cuenca	grados
11 CENTLON	Longitud del centroide de la cuenca	grados
12 COEFESCR	Coefficiente de escurrimiento	
13 ORDEN	Orden de la corriente principal	
14 PERIMET	Perímetro del parteaguas	km
15 TEMP	Temperatura media anual	°C

La tabla siguiente muestra el resumen de los resultados de la regresión aplicados a las 10 series de gastos máximos instantáneos, asociados con las 15 características fisiográficas.

Lugar	PERÍODO DE RETORNO ( años)									
	2	5	10	20	50	100	500	1000	5000	10000
1	3	14	6	6	6	6	6	6	6	6
2	8	3	8	8	8	8	8	8	8	8
3	5	2	14	2	2	2	2	2	2	2
4	6	11	2	14	5	5	5	5	5	5
5	2	6	5	5	10	10	10	10	10	10
6	10	8	11	11	14	11	11	11	11	11
7	14	5	3	3	11	14	14	1	1	1
8	7	4	10	10	1	1	1	14	14	14
9	1	9	9	9	3	7	7	7	7	7
10	11	12	4	1	7	3	3	3	9	9
11	9	15	13	7	9	9	9	9	3	3
12	13	13	7	4	12	12	12	12	12	12
13	4	10	1	13	13	13	13	13	13	13
14	15	7	15	12	4	4	4	4	4	4
15	12	1	12	15	15	15	15	15	15	15

Por ejemplo, para el gasto de 100 años de período de retorno las características más importantes son: la número 6 (PENDCAUC), la 8 (FRECUENC) y la 2 (PRECMEDA) mientras que las menos significativas son la 15 (TEMP), la 4 (ELEV MED) y la 13 (ORDEN) esto es evidente, ya que al tratarse de un modelo lluvia-escurrimiento, la temperatura, por ejemplo, no es una variable importante mientras que la precipitación media si lo es.

TABLA 6.6 REGRESION DE 15 CARACTERISTICAS FISIOGRAFICAS  
CON GASTOS DE 100 AÑOS DE PERIODO DE RETORNO

MUESTRAS NORMALIZADAS

ESTACION	GASTO	AREA	PRECMEDA	PENDCUENC	ELEVMEGD	LONGITUD	PENDCAUC	PARMFORM	FRECUENC	DENSIDAD	CENTLAT	CENTLON	COEFESCR	ORDEN	PERIMET	TEMP	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
1	SM ZAPOTITLAN	17.9912	23.69449	3743878	-0.5790496	9607.153	15.455060	-2.565026	-9.105381	-19.71671	-12.82509	13286.370	107.80	-0.3983353	8.486	13.63319	104845.7
2	SAN BLAS	25.5719	23.58554	3859917	-0.5789664	9844.762	14.993490	-2.560140	-9.312546	-19.71671	-13.25143	13318.370	107.82	-0.3983353	8.486	13.42742	104845.7
3	LA TINA	14.2347	8.53830	2389610	-0.5808876	1612.310	5.599348	-2.535142	-2.375831	-22.05911	-4.80734	12109.430	108.52	-0.4006171	1.323	5.84861	104845.7
4	BAMICORI	12.3471	8.08484	3870632	-0.5813068	1663.061	4.634660	-2.506403	-1.457429	-19.91727	-4.20522	12411.860	108.40	-0.4003422	1.323	6.37529	104845.7
5	LAS CAÑAS	24.7555	23.03618	4163366	-0.5786440	10420.860	14.195040	-2.489995	-9.242221	-19.15017	-13.36442	13366.460	107.68	-0.3979871	8.486	13.04758	104845.7
6	EL MAHONE	9.7156	23.02171	4174810	-0.5786440	10420.860	14.124000	-2.486701	-9.277221	-19.15017	-13.42194	13366.460	107.68	-0.3979871	8.486	13.01754	104845.7
7	CHOIX	15.6953	12.55683	6169428	-0.5805729	5408.332	8.468395	-2.619793	-3.294437	-18.97209	-6.01471	12719.460	108.07	-0.3996220	3.219	8.63291	71942.4
8	HUITES	26.0704	22.50638	4135119	-0.5781126	11585.790	13.037760	-2.445161	-9.312546	-18.79895	-14.74390	13446.890	107.58	-0.3975366	8.486	12.60782	17865.3
9	PALO DULCE	21.3247	17.17919	5285010	-0.5771933	10287.400	11.436660	-2.361279	-6.176214	-20.12421	-8.68318	14387.550	108.20	-0.3962847	5.621	10.32275	17865.3
10	CHINIPAS	17.2831	16.45687	5261631	-0.5767946	11794.260	10.109640	-2.281035	-5.936973	-19.52221	-9.34178	14557.140	108.15	-0.3961835	5.621	9.89899	17865.3
11	SAN FRANCISCO	21.5343	20.91171	3564936	-0.5779903	12382.140	12.671550	-2.435458	-8.446182	-17.70802	-12.67491	13079.680	107.80	-0.3980752	5.621	11.87712	148070.7
12	SAN IGNACIO	18.8584	19.11274	3891391	-0.5776036	12278.400	11.966510	-2.358499	-7.102558	-16.55103	-10.66648	12719.460	107.17	-0.3987555	5.621	11.27563	71942.4
13	LA VERANERA	15.4869	15.75261	3203433	-0.5794555	15083.720	10.277680	-2.261749	-4.924246	-19.15017	-8.00311	13886.980	107.40	-0.3957716	3.219	10.04128	4947.0
14	URIQUE	12.9657	15.66893	3190886	-0.5795057	15255.890	9.994469	-2.266314	-4.961637	-18.97209	-8.17878	13886.980	107.37	-0.3957716	3.219	9.39194	4947.0
15	GUERACHIC	21.5989	17.14620	4085985	-0.5788718	14820.110	10.493570	-2.311675	-6.536453	-18.46663	-9.77251	12411.860	106.80	-0.3992379	5.621	10.06440	29995.3
16	ALAMOS	12.6269	13.92958	2214160	-0.5809344	2134.088	9.906340	-2.651100	-4.102860	-22.34370	-5.96625	13032.300	108.80	-0.4007442	3.219	9.13548	104845.7
17	CAZANATE	18.4591	13.27672	2384200	-0.5809410	2413.531	8.889206	-2.671206	-4.027864	-22.94885	-6.43151	13158.910	108.82	-0.4004107	3.219	8.63291	104845.7
18	JAINA	20.4937	18.07155	6331225	-0.5778176	6949.378	11.278940	-2.364299	-6.567353	-17.99998	-9.21598	11856.410	107.43	-0.3984207	5.621	10.78619	71942.4
19	TOAHAYANA	17.3263	16.57367	6657468	-0.5770642	9253.518	10.109640	-2.340063	-6.394075	-16.91107	-9.34178	11812.140	107.27	-0.3978984	5.621	9.69896	47537.5
20	TECUSIAPA	18.9670	15.48311	7261533	-0.5758885	11437.760	8.778211	-2.271767	-4.887447	-18.46663	-9.67743	11856.410	107.12	-0.3975366	5.621	9.84192	17865.3
21	LOS MOLINOS	10.5746	9.91202	549.3475	-0.5789381	7093.356	5.968697	-2.175635	-2.569320	-28.73514	-5.26310	11462.090	107.27	-0.3988375	1.323	6.92036	104845.7
22	NARANJO	17.4953	13.65078	3971042	-0.5806459	2580.053	9.375138	-2.424750	-3.878155	-24.33008	-6.22604	12154.470	108.30	-0.4006127	3.219	9.08357	104845.7
23	PILOTE	13.8019	10.60548	3240421	-0.5801326	1778.454	7.868423	-2.443754	-2.586119	-23.96057	-4.31680	11856.410	108.23	-0.4006127	1.323	7.53838	148070.7
24	GUAMICHIL	17.9358	13.00062	3495636	-0.5811979	662.274	7.728301	-2.441519	-3.888302	-23.96057	-7.13149	11175.920	107.83	-0.4014241	3.219	8.50879	148070.7
25	PERICOS	11.1260	8.49807	3146456	-0.5810628	456.121	5.656894	-2.466217	-1.807351	-21.78544	-3.65681	10714.600	106.62	-0.4009368	1.323	6.32759	148070.7
26	TIERRA BLANCA	18.8853	19.34020	6248375	-0.5787568	8805.601	13.188290	-2.469270	-6.835650	-23.96057	-9.79668	11062.850	106.85	-0.3989997	5.621	11.62813	9879.0
27	PTE. S PACIFICO	23.0651	19.28235	6267925	-0.5787568	8805.601	13.054670	-2.460506	-6.827253	-23.60804	-9.82100	11062.850	106.85	-0.3989997	5.621	11.56928	9879.0
28	PALOS BLANCOS	21.1553	19.27426	6326443	-0.5787568	8965.694	12.848670	-2.455493	-7.001932	-23.60804	-9.91995	11062.850	106.85	-0.3989997	5.621	11.46899	9879.0
29	EL VAREJONAL	16.1341	19.13523	6415455	-0.5783705	9626.599	12.433250	-2.446573	-6.904543	-24.33008	-10.42963	11105.160	106.80	-0.3988375	5.621	11.41777	9879.0
30	BADIRAGUATO	23.2064	11.69172	6345079	-0.5777985	2628.445	7.433235	-2.181043	-3.246923	-24.33008	-5.64529	11275.510	107.50	-0.3984207	1.323	7.90517	104845.7
31	GUATENIPA	20.6650	18.10294	6704508	-0.5781412	12219.780	11.301740	-2.325849	-6.314806	-23.27118	-10.34444	11133.430	106.58	-0.3985894	5.621	10.97866	29995.3
32	LA HUERTA	16.3935	17.08436	6924025	-0.5789866	13977.560	9.694477	-2.261101	-6.063663	-22.05911	-11.05096	11175.920	106.42	-0.3988375	5.621	10.32275	9879.0
33	PTE. CAÑEDO	18.6307	15.73697	7347580	-0.5787756	3675.954	6.694477	-2.513559	-5.595152	-21.78544	-8.34888	10441.520	106.90	-0.3985894	3.219	9.54027	148070.7
34	SANALONA	17.1323	15.38441	7954912	-0.5778270	4416.497	8.587059	-2.525445	-5.869770	-22.05911	-10.12618	10468.610	106.83	-0.3980752	3.219	9.07309	148070.7
35	PICACHO	16.7893	15.04387	8417021	-0.5778270	4416.497	8.587059	-2.525445	-5.494068	-21.02337	-9.84549	10468.610	106.83	-0.3980752	3.219	9.04140	148070.7
36	TAMAZULA	14.6879	13.89168	9664219	-0.5756256	6046.539	6.829754	-2.362139	-4.940790	-22.05911	-10.40102	10509.330	106.77	-0.3976281	3.219	8.47051	104845.7
37	EL BLEDAL	14.8523	9.20963	4548620	-0.5806649	951.777	5.656894	-2.414667	-2.123551	-25.12597	-4.53110	10040.970	107.00	-0.3985894	1.323	6.71017	148070.7
38	SANTA CRUZ	20.9340	18.37916	7748503	-0.5776582	9895.614	11.347080	-2.313498	-6.559597	-14.78581	-10.04750	10040.970	106.18	-0.3973516	5.621	11.03667	9879.0
39	ACAITAN	18.7696	13.38684	7366231	-0.5780702	3389.842	9.632250	-2.339861	-3.998088	-15.21932	-6.00077	9549.019	106.55	-0.3972581	3.219	8.75205	148070.7
40	IPALONJA	21.6671	17.09372	6710842	-0.5788016	7748.808	11.209940	-2.305560	-6.329053	-16.21303	-8.77236	9309.409	106.03	-0.3967801	5.621	10.15528	4947.0
41	PIAXTLA	21.6649	16.58999	7335306	-0.5780759	9561.494	10.413660	-2.242165	-5.881863	-16.66847	-9.25741	9334.434	105.98	-0.3964849	5.621	10.04128	4947.0
42	EL QUELITE	16.1946	11.17638	3002387	-0.5812399	706.801	6.456751	-2.476740	-2.949547	-13.94745	-6.00077	8722.648	106.37	-0.3988375	1.323	7.71111	148070.7

**TABLA 6.6 REGRESION DE 15 CARACTERISTICAS FISIOGRAFICAS  
CON GASTOS DE 100 AÑOS DE PERIODO DE RETORNO**

...Continuación

*Estadísticos de la regresión*

Coef. correlac.	0.7045
Coef. determ.	0.4963
Error estándar	3.5596
Observaciones	42

*Análisis de Varianza*

	<i>gl</i>	<i>Suma cuadrados</i>	<i>Cuadrado medio</i>	<i>F</i>	<i>F crítico</i>
Regresión	15	324.5878	21.6392	1.7078	2.0711
Residuales	26	329.4373	12.6707		
Total	41	654.0250			

	<i>Coefficientes</i>	<i>Error Estándar</i>	<i>Estadístico t</i>	<i>Límites</i>		<i>t crítico</i>
				<i>Inferior 95%</i>	<i>Superior 95%</i>	
Independiente	-353.9082	736.3187	-0.4806	-1867.4339	1159.6174	
x1	-1.5649	13.0690	-0.1197	-28.4287	25.2988	2.0555
x2	0.0000	0.0000	0.6443	0.0000	0.0000	
x3	-298.5888	975.1405	-0.3062	-2303.0201	1703.8426	
x4	-0.0007	0.0007	-1.0045	-0.0022	0.0007	
x5	1.5308	2.8574	0.5357	-4.3426	7.4042	
x6	16.9413	10.7477	1.5763	-5.1510	39.0335	
x7	-2.0942	11.3281	-0.1849	-25.3794	21.1910	
x8	0.2855	0.2559	1.1158	-0.2404	0.8114	
x9	-1.0377	2.2646	-0.4582	-5.6926	3.6172	
x10	0.0004	0.0028	0.1250	-0.0055	0.0062	
x11	0.2822	5.3956	0.0523	-10.8086	11.3730	
x12	-513.4175	891.0047	-0.5762	-2344.9050	1318.0699	
x13	-0.6967	1.0851	-0.6420	-2.9272	1.5339	
x14	0.4332	15.4308	0.0281	-31.2853	32.1517	
x15	0.0000	0.0000	-1.2282	-0.0001	0.0000	

Un detalle interesante de esta tabla es observar, como la característica 3 (PENDCUENC) es *significativa* para gastos de período de retorno pequeños, mientras que para períodos de retorno de 50 años en adelante esta variable *no es significativa*. Sucede lo contrario con el AREA, mientras es significativa para períodos de retorno grandes, no lo es para períodos entre 2 y 20 años.

Un análisis detallado de la tabla anterior, nos lleva a eliminar las últimas 5 variables que consistentemente no son significativas para ninguno de los 10 períodos de retorno analizados. Así, las variables 4, 9, 12, 13 y 15 son eliminadas. Las restantes 10 variables en competencia serán:

VARIABLE	DESCRIPCION	UNIDADES
1 AREA	Area drenada hasta la estación	km <sup>2</sup>
2 PRECMEDA	Precipitación media anual	mm
3 PENDCUENC	Pendiente media de la cuenca	m km <sup>-1</sup>
4 LONGITUD	Longitud de la corriente principal	km
5 PENDCAUC	Pendiente de la corriente principal	m km <sup>-1</sup>
6 PARMFORM	Parámetro de forma	km <sup>-1</sup>
7 FRECUENC	Frecuencia de corriente	km <sup>-2</sup>
8 CENTLAT	Latitud del centroide de la cuenca	grados
9 CENTLON	Longitud del centroide de la cuenca	grados
10 PERIMET	Perímetro del parteaguas	km

No es suficiente eliminar las variables, es necesario probar a través de la ecuación (4.51) previamente analizada, que las variables eliminadas *no eran significativas*. Así, se tiene que:

$$n = 42 \quad p = 16 \quad k = 5 \quad n-p = 26$$

$$Q_2 = 324.5877 \quad (\text{en el modelo completo, con } p-1 \text{ grados de libertad})$$

$$Q_2^* = 301.9398 \quad (\text{en el modelo reducido, con } p-k-1 \text{ grados de libertad})$$

$$Q_1 = 329.4373 \quad (\text{en el modelo completo, con } n-p \text{ grados de libertad})$$

$$F = \frac{(Q_2 - Q_2^*)/k}{Q_1/(n-p)} = \frac{4.5296}{12.6707} = 0.35749$$

El valor crítico de  $F$  al 95% para  $k=5$  y  $(n-p)=26$  es 2.586, y como  $0.35749 < 2.586$  la hipótesis nula que propone que las últimas " $k$ " variables independientes son las que *no contribuyen significativamente* a explicar la variación lineal en  $Y$ . Lo que es equivalente a escribir.

$$H_0: \beta_{p-k+1} = \beta_{p-k+2} = \dots = \beta_p = 0$$

$H_1$ : por lo menos uno de los  $\beta$  no es cero

TABLA 6.7 REGRESION DE 10 CARACTERISTICAS FISIOGRAFICAS  
CON GASTOS DE 100 AÑOS DE PERIODO DE RETORNO

*Estadísticos de la regresión*

Coef. correlac.	0.6795
Coef. determ.	0.4617
Error estándar	3.3701
Observaciones	42

*Análisis de Varianza*

	gl	Suma cuadrados	Cuadrado medio	F	F crítico
Regresión	10	301.9398	30.1940	2.6585	2.1512
Residuales	31	352.0852	11.3576		
Total	41	654.0250			

	Coeficientes	Error Estándar	Estadístico t	Límites		t crítico
				Inferior 95%	Superior 95%	
Independiente	-649.5850	587.5430	-1.1056	-1847.8874	548.7174	
x1	0.7705	11.6845	0.0659	-23.0603	24.6013	2.0395
x2	0.0000	0.0000	1.1864	0.0000	0.0000	
x3	-555.8776	864.4609	-0.6430	-2318.9583	1207.2031	
x4	0.6475	1.0737	0.6030	-1.5423	2.8372	
x5	14.2350	8.4654	1.6815	-3.0303	31.5003	
x6	-0.6578	10.0285	-0.0656	-21.1110	19.7954	
x7	0.1464	0.1846	0.7930	-0.2301	0.5229	
x8	-0.0015	0.0010	-1.5337	-0.0036	0.0005	
x9	3.6291	2.3102	1.5709	-1.0825	8.3407	
x10	-1.4007	13.7858	-0.1016	-29.5171	26.7157	

**H<sub>0</sub> DEBE ACEPTARSE**, y por lo tanto se dice que las 5 variables eliminadas *no eran significativas en el modelo*. De esta manera al realizar una vez más la regresión para el modelo reducido a 10 variables, (Tabla 6.7) se tiene que:

$$F = \frac{30.1940}{11.3576} = 2.6585$$

El valor crítico de  $F_{0.95,10,31} = 2.151$  y como  $2.659 > 2.151$  **SE RECHAZA H<sub>0</sub>** es decir *la ecuación de regresión si explica el significado de la variable Q<sub>100</sub>*.

Cabe recordar que solo se toma el ejemplo de 100 años, pero el procedimiento se desarrolló para todos los eventos de diseño. Siguiendo el procedimiento anterior debe evaluarse ahora cuáles son las variables significativas, en este modelo reducido de 10 variables; esto con ayuda del estadístico "t". Para la primera variable AREA se tiene que:

$$t = \frac{0.7705}{11.6845} = 0.06594$$

El valor crítico de  $t_{0.975,31} = 2.0395$  y como  $0.0659 > 0$  podemos decir que, *ésta variable si contribuye a explicar el gasto de 100 años de periodo de retorno*.

Mientras que para la variable 8 (CENTLAT) por ejemplo:

$$t = \frac{-0.001538}{0.001003} = -1.533$$

y como  $-1.533 < 0$  podemos decir que, *ésta variable no contribuye a explicar el gasto de 100 años de periodo de retorno*. De esta forma, al analizar con cuidado los resultados de la tabla 6.7, puede crearse una tabla similar a la que se empleó en el modelo completo en donde se jerarquicen las características en este modelo. Así pues se observa que 4 de las 10 características son las que predominantemente se encuentran en los últimos lugares de significancia; o lo que es lo mismo decir que tienen estadísticos "t" negativos, lo cual indica que deben eliminarse del modelo.

Lugar	PERÍODO DE RETORNO (años)									
	2	5	10	20	50	100	500	1000	5000	10000
1	2	9	9	5	5	5	5	5	5	5
2	3	10	5	9	9	9	9	9	9	9
3	4	2	2	2	2	2	2	2	2	2
4	5	3	10	10	7	7	7	7	7	7
5	10	5	7	7	4	4	4	4	4	4
6	9	4	4	4	1	1	1	1	1	1
7	7	7	3	3	10	6	6	6	6	6
8	6	6	1	1	6	10	10	10	10	10
9	1	1	6	6	3	3	3	3	3	3
10	8	8	8	8	8	8	8	8	8	8

Las últimas 4 variables que consistentemente no son significativas para ninguno de los 10 periodos de retorno analizados, serían la 3, 6, 8, y 10. Así, que son eliminadas; finalmente las restantes 6 variables en competencia serán:

VARIABLE	DESCRIPCION	UNIDADES
1 AREA	Area drenada hasta la estación	km <sup>2</sup>
2 PRECMEDA	Precipitación media anual	mm
3 LONGITUD	Longitud de la corriente principal	km
4 PENDCAUC	Pendiente de la corriente principal	m km <sup>-1</sup>
5 FRECUENC	Frecuencia de corriente	km <sup>-2</sup>
6 CENTLON	Longitud del centroide de la cuenca	grados

Una vez más, es necesario probar a través de la ecuación (4.51) previamente analizada, que las variables eliminadas *no eran significativas*. Así, se tiene que:

$$n = 42 \quad p = 11 \quad k = 4 \quad (n-p) = 31$$

$$Q_2 = 301.9398 \text{ (en el modelo completo, con } p-1 \text{ grados de libertad)}$$

$$Q_2^* = 260.9290 \text{ (en el modelo reducido, con } p-k-1 \text{ grados de libertad)}$$

$$Q_1 = 352.0852 \text{ (en el modelo completo, con } n-p \text{ grados de libertad)}$$

$$F = \frac{10.2527}{11.3576} = 0.9027$$

El valor crítico de *F* al 95% para *k*=4 y *(n-p)*=31 es 2.685, y como 0.9027 < 2.685 la hipótesis nula que propone que las últimas "*k*" variables independientes son las que *no contribuyen significativamente* a explicar la variación lineal en *Y*. **DEBE ACEPTARSE**, y por lo tanto se dice que las 4 variables eliminadas *no eran significativas en el modelo*.



TABLA 6.8 REGRESION DE 6 CARACTERISTICAS FISIOGRAFICAS  
CON GASTOS DE 100 AÑOS DE PERIODO DE RETORNO

*Estadísticos de la regresión*

Coef. correlac.	0.6316
Coef. determ.	0.3990
Error estándar	3.3513
Observaciones	42

*Análisis de Varianza*

	<i>gl</i>	<i>Suma cuadrados</i>	<i>Cuadrado medio</i>	<i>F</i>	<i>F crítico</i>
Regresión	6	260.9290	43.4882	3.8720	2.3805
Residuales	35	393.0960	11.2313		
Total	41	654.0250			

	<i>Coefficientes</i>	<i>Error Estándar</i>	<i>Estadístico t</i>	<i>Limites</i>		<i>t crítico</i>
				<i>Inferior 95%</i>	<i>Superior 95%</i>	
Independiente	-3.9803	107.5930	-0.0370	-222.4061	214.4454	
x1	0.0842	0.5253	0.1603	-0.9822	1.1506	2.0301
x2	0.0000	0.0000	1.3390	0.0000	0.0000	
x3	0.6529	0.8032	0.8129	-0.9777	2.2836	
x4	3.7578	5.0952	0.7375	-6.5860	14.1016	
x5	0.1056	0.1760	0.5998	-0.2518	0.4630	
x6	0.2145	1.0448	0.2053	-1.9066	2.3356	

Finalmente, al realizar la regresión para el modelo reducido a 6 variables (Tabla 6.8) se tiene que:

$$F = \frac{43.4882}{11.2313} = 3.8720$$

El valor crítico de  $F_{0.95,6,35} = 2.380$  y como  $3.8720 > 2.380$  **SE RECHAZA  $H_0$**  es decir *la ecuación de regresión si explica el significado de la variable  $Q_{100}$ .*

Y dado que el valor de estadístico "t" es positivo para las 6 variables independientes, éste se considera como el mejor modelo de regresión, y no deben de eliminarse más variables de él; además esto se valida al observar de las Tablas 6.7 y 6.8 que el error de la regresión disminuyó de 3.560 en el modelo con 15 características fisiográficas a 3.351 para el modelo final de 6 variables independientes. De aquí en adelante se manejarán solo estas seis características fisiográficas, tanto para el análisis de homogeneidad de cuencas como para determinar las ecuaciones regionales.

## **6.5 Identificación de regiones homogéneas**

### **6.5.1 Criterio tradicional**

Frecuentemente, el considerar como homogéneas a cuencas geográficamente contiguas, es una de las suposiciones iniciales en un análisis regional. Como se comentó en el apartado 2.2, la distancia que se maneja para hacer tal consideración es de 70 a 80 km. Para el caso de la región hidrológica 10 se efectuó este criterio tradicional, tomando, en un plano a escala, radios de 80 km en cada estación y agrupando así las estaciones que quedarán dentro, procediendo de esta manera se obtuvieron 3 regiones, tal y como se muestra en el plano 3 y en la tabla 6.9. La primera (Región A) con 24 estaciones, la región B con 14 y la C con 4 estaciones.

### **6.5.2 Trazo multidimensional**

La metodología expuesta en el apartado 2.2.1 señala el empleo de la curva de Andrews como una técnica para regionalizar y agrupar cuencas con características fisiográficas similares. Una vez que se efectuó un análisis de correlación y regresión, pudo determinarse el número de variables significativas para la región hidrológica 10, las cuales resultaron ser: AREA, PRECMEDA, PENDCUENC, LONGITUD, FRECUENC Y CENTLON.

TABLA 6.9 DISTANCIAS ENTRE ESTACIONES PARA CIRCULOS DE 80 km

Distancias en km	
1	SM ZAPOTITLAN
2	SAN BLAS
3	LA TINA
4	BAMICORI
5	LAS CAÑAS
6	EL MAHONE
7	CHOIX
8	HUITES
9	PALO DULCE
10	CHINIPAS
11	SAN FRANCISCO
12	SAN IGNACIO
13	LA VERANERA
14	URIQUE
15	GUERACHIC
16	ALAMOS
17	CAZANATE
18	JAINA
19	TOAHAYANA
20	TECUSIAPA
21	LOS MOLINOS
22	NARANJO
23	ZOPILOTE
24	GUAMUCHIL
25	PERICOS
26	TIERRA BLANCA
27	PUENTE S.PACIF.
28	PALOS BLANCOS
29	EL VAREJONAL
30	BADIRAGUATO
31	GUATENIPA
32	LA HUERTA
33	PUENTE CAÑEDO
34	SANALONA
35	PICACHO
36	TAMAZULA
37	EL BLEDAL
38	SANTA CRUZ
39	ACATTITAN
40	IXPALINO
41	PIANTLA
42	EL QUELITE

LAS CAÑAS	73	45	30	15	LAS CAÑAS											
EL MAHONE	76	49	34	12	3	EL MAHONE										
CHOIX		66	44	40	32	CHOIX										
HUITES			60	53	50	20	HUITES									
PALO DULCE				73	69	61	34	15	PALO DULCE							
CHINIPAS								78	59	45	CHINIPAS					
SAN FRANCISCO									69	66	64	27	25	32	70	SAN FRANCISCO

EJEMPLO: Distancia entre la estación Huites y la estación Las Cañas, es de 53 km

Se obtuvo el trazo multidimensional de las 42 cuencas de la región, tanto individualmente como en conjunto, cabe recordar que el orden que ocupan las variables dentro de la ecuación es muy importante. Para este caso, y por tratarse de un modelo de escurrimientos, el área de la cuenca, la precipitación media y la red de drenaje, (representada por la frecuencia, longitud y la pendiente del cauce) fueron las características principales.

La figura 2 que se presentó en el segundo capítulo, muestra el trazo de todas las curvas de la región, de ésta se pudieron extraer en primera instancia dos regiones: la denominada región "A"; ver figura 7, la cual agrupó inicialmente a 15 cuencas que, inequívocamente tienen un comportamiento similar; mientras que 7 estaciones conformaron la región "B"; ver figura 8. Las restantes 20 cuencas, ofrecían una incertidumbre acerca de a qué grupo pertenecían.

A través del programa ANDREWS.BAS, fueron dibujadas, las curvas características de cada región; los criterios SEUCLID y COSINE fueron probados como un método de liga entre regiones, (tal y como se explicó en el apartado 2.2.1.). Para la cuenca en la estación Choix, por ejemplo, se obtuvieron los resultados siguientes:

REGION A		REGION B	
SS	CS	SS	CS
21,652.70	0.9150	87,327.34	0.4152

Estos resultados indican que al comparar a Choix con las curvas significativas de las regiones A y B, ver figuras 9 y 10, el criterio SEUCLID es menor en la región A, y por lo tanto las dos curvas tienen una asociación similar en cuanto al peso que se les dio a las variables significativas, mientras que el criterio COSINE prueba que mientras más se acerque al valor de +1 ( $1 > 0.915 > 0.415$ ) existe una mejor correlación entre las dos curvas, y por lo tanto CS es un criterio válido de liga entre estas dos cuencas.

Siguiendo el mismo procedimiento se ubicaron las 20 estaciones restantes, siempre siguiendo el criterio del menor valor de SEUCLID y el valor más cercano a +1 de COSINE. Las figuras 11 y 12 muestran algunas de estas pruebas; las cuales permitieron definir con certeza el grupo al que pertenecían las cuencas. El caso más cerrado, fue el de la cuenca de Picachos en donde:

REGION A		REGION B	
SS	CS	SS	CS
54,040.17	0.4974	34,662.43	0.6043

Para este caso, ambos valores del COSINE se aproximan a 0.50, sin embargo, tanto su ubicación geográfica como el criterio SEUCLID, fácilmente ubican a la cuenca de Picachos dentro de la región B.

Finalmente, la región A quedó formada por 27 cuencas homogéneas (figuras 13 a 17) esta agrupación, arrojó un conjunto de cuencas con características muy similares, podemos mencionar como la principal de ellas a la altura de la precipitación media, la cual tiene un promedio de 800 mm en la región B (ver figuras 18 a 20), que agrupa a 15 cuencas, el promedio de la precipitación es de 1000 mm. Esto puede explicarse claramente si se observa el plano de isoyetas de la región 10, el cual tiene una forma similar a la agrupación que se realizó a través del trazo multidimensional. El plano 4 muestra las regiones A y B en el cual se propone la subdivisión de la región hidrológica 10.

### **6.5.3 Regiones homogéneas con base en los coeficientes de variación.**

Tal y como se explicó en el apartado 2.3, Mosley en 1981 propone localizar una cuenca a través de las coordenadas (QBAR/AREA) y el coeficiente de variación de sus gastos máximos instantáneos (CV) dentro de una gráfica a escala.

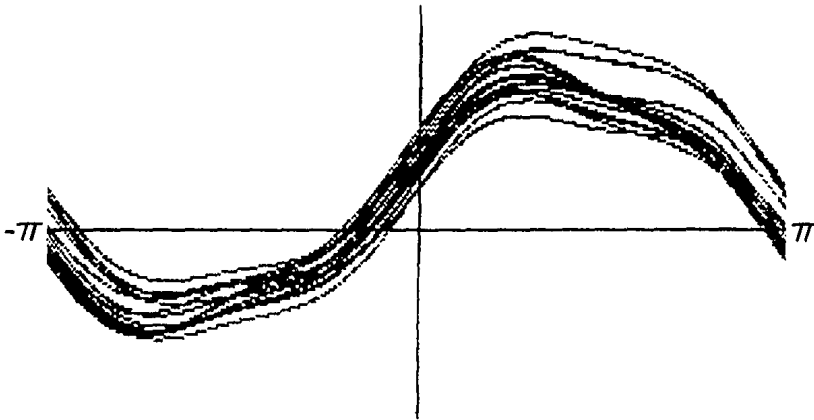
De tal manera, como se muestra en la figura 21, al dibujar los valores para las 42 cuencas, se forman 3 grupos: La región B agrupa a 14 cuencas, que son las que tienen un coeficiente de variación entre 0.0007 y 0.0663. La siguiente región agrupa a 9 cuencas en el intervalo de 0.0942 a 0.1798. Finalmente, la región A agrupó a las 19 estaciones restantes en el intervalo de 0.1989 a 0.4723. En la regionalización que resultó de este técnica, demuestra que solo algunas de las cuencas que pertenecían al mismo grupo en el análisis del trazo multidimensional, quedaron agrupadas juntas.

Aún así, ésta fué la agrupación inicial para dar comienzo a los ciclos de liga.

En la primera interacción, se aplicaron las expresiones que se estudiaron en el punto 2.3.2.; de ésta manera, se calculó el estadístico  $F$  tomando un valor de 91.688 mientras que su valor crítico (ver apartado 4.1.) tiene un valor de 3.238, esto indica que:

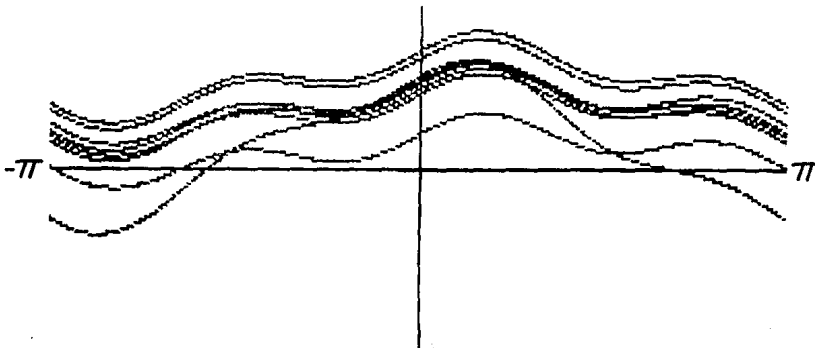
$$F(91.688) > F \text{ crítico } (3.238)$$

**TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : REGION A (15 estaciones)**



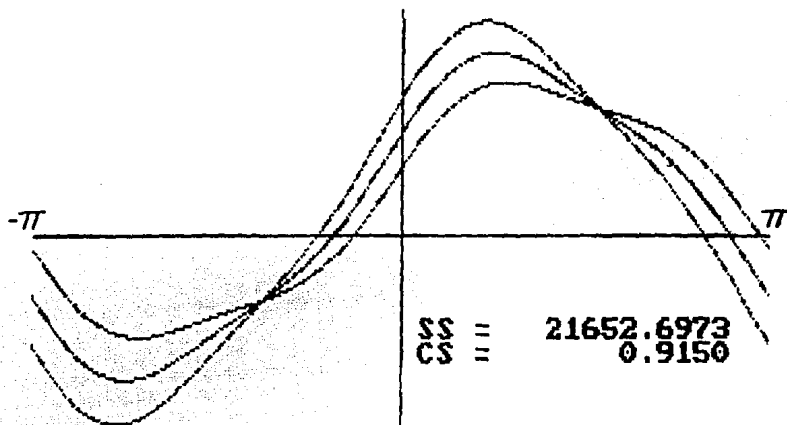
**FIGURA 7 CURVAS DE ANDREWS PARA LA REGION "A"**

**TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : REGION B (7 estaciones)**



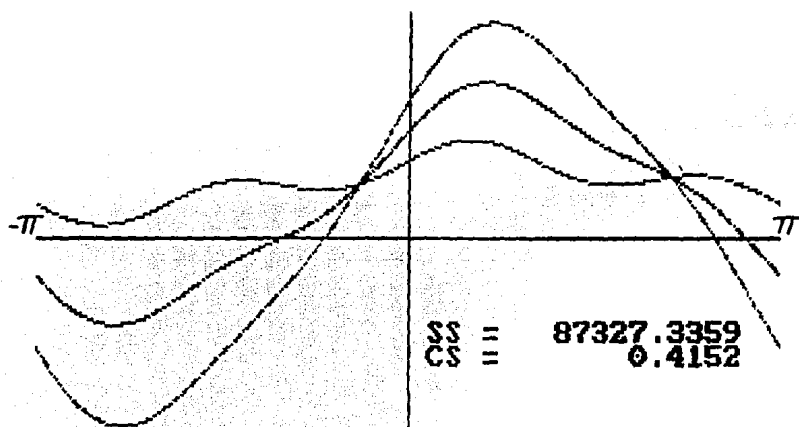
**FIGURA 8 CURVAS DE ANDREWS PARA LA REGION "B"**

**TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : REGION A Y CHOIX**



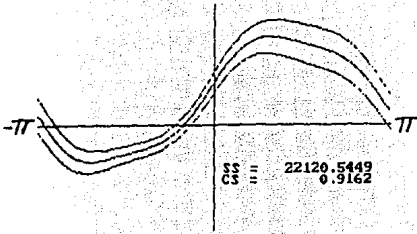
**FIGURA 9 CURVAS DE ANDREWS PARA LA REGION "A" Y CHOIX**

**TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : REGION B Y CHOIX**

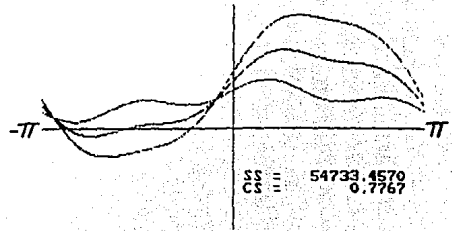


**FIGURA 10 CURVAS DE ANDREWS PARA LA REGION "B" Y CHOIX**

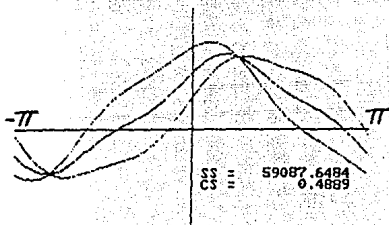
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION A Y JAINA



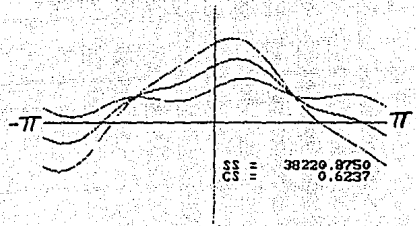
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION B Y JAINA



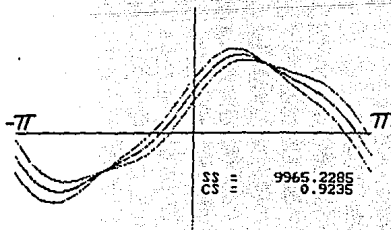
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION A Y TECUSIAPA



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION B Y TECUSIAPA



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION A Y GUAMUCHIL



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION B Y GUAMUCHIL

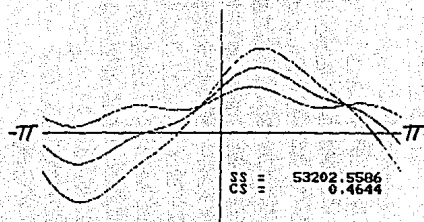
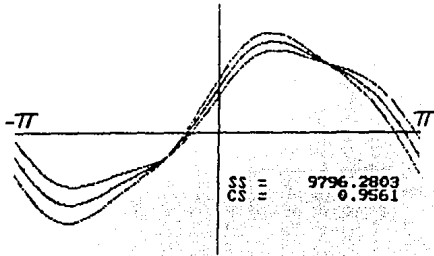


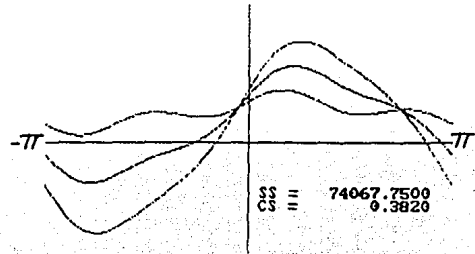
FIGURA 11 - CURVAS DE ANDREWS PARA LAS REGIONES "A" Y "B"  
 DE LAS ESTACIONES: JAINA, TECUSIAPA Y GUAMUCHIL.



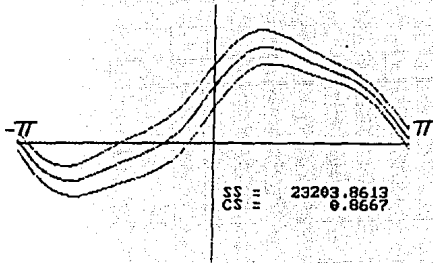
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION A Y BADIAGUATO



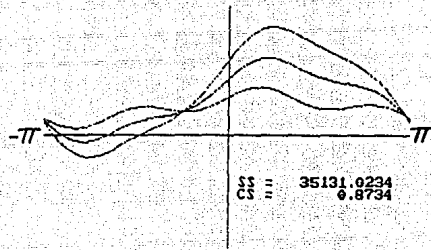
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION B Y BADIAGUATO



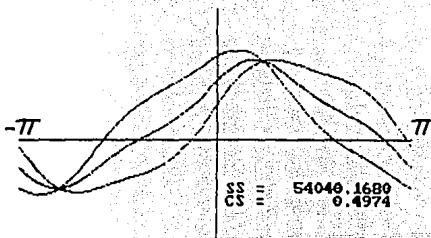
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION A Y EL QUELITE



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION B Y EL QUELITE



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION A Y PICACHOS



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION B Y PICACHOS

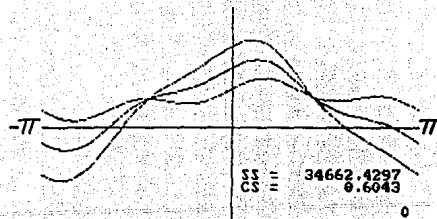
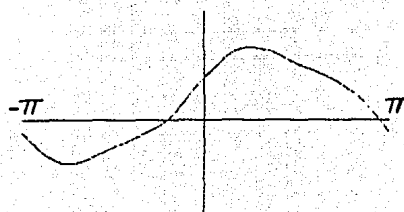


FIGURA 12 CURVAS DE ANDREWS PARA LAS REGIONES "A" Y "B"  
 DE LAS ESTACIONES: BADIAGUATO, EL QUELITE Y PICACHOS.

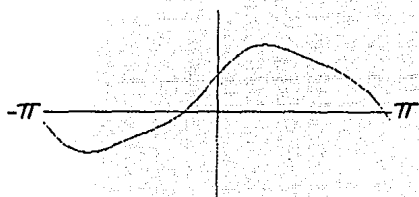
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION "A" (27 estaciones)



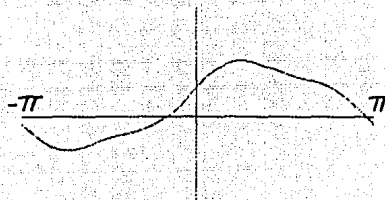
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 SM ZAPOTITLAN



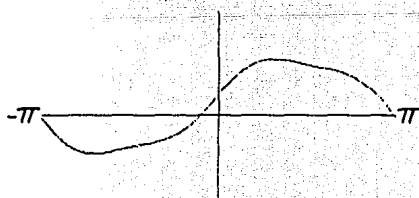
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : SAN BLAS



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : LA TINA



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : BAMICORI



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : LAS CAÑAS

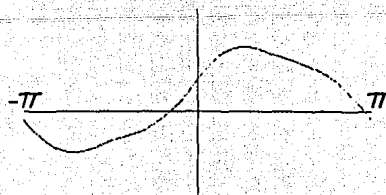
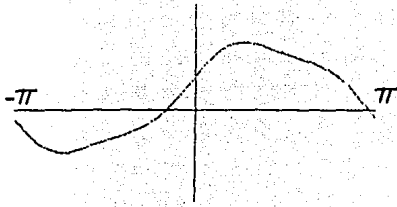
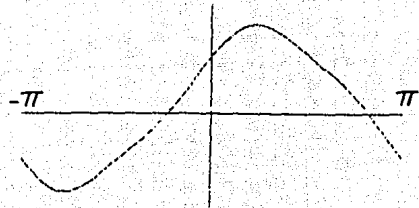


FIGURA 13 CURVAS DE ANDREWS PARA LAS ESTACIONES: SAN M. ZAPOTITLAN, SAN BLAS,  
 LA TINA, BAMICORI Y LAS CAÑAS ( REGION "A" )

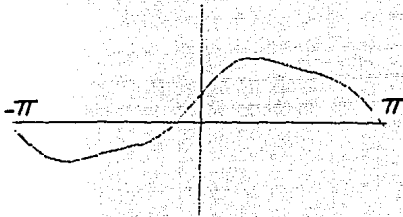
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : EL MAHONE



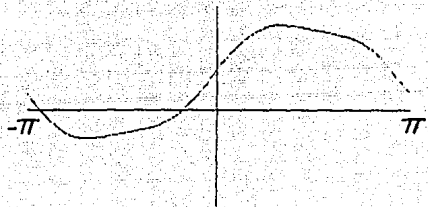
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : CHOIX



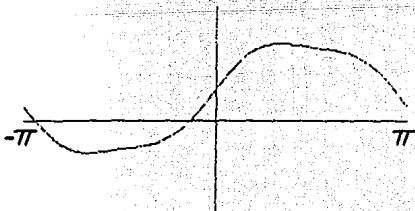
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : HUITES



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : PALO DULCE



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : CHINIPAS



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : SAN FRANCISCO

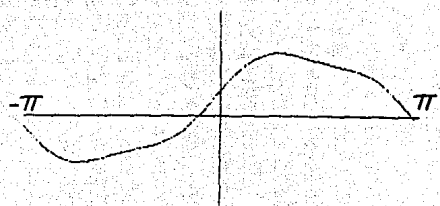
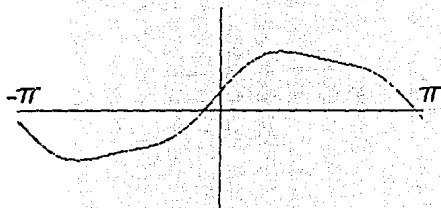
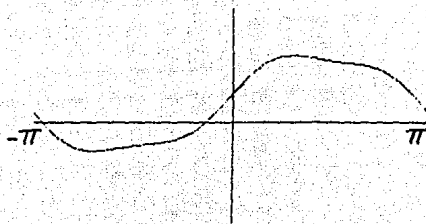


FIGURA 14 CURVAS DE ANDREWS PARA LAS ESTACIONES: EL MAHONE, CHOIX,  
HUITES, PALO DULCE, CHINIPAS Y SAN FRANCISCO ( REGION "A" )

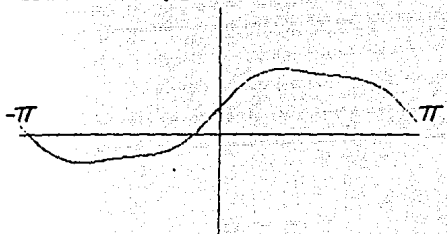
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : SAN IGNACIO



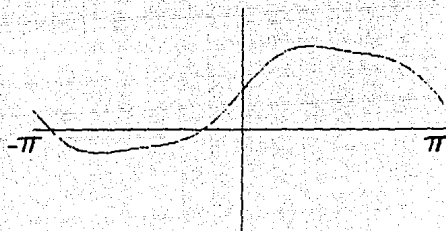
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : LA VERANERA



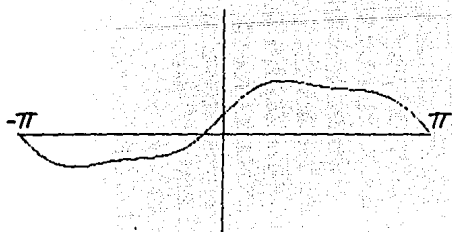
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : URIQUE



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : GUERACHIC



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : ALAMOS



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : CAZANATE

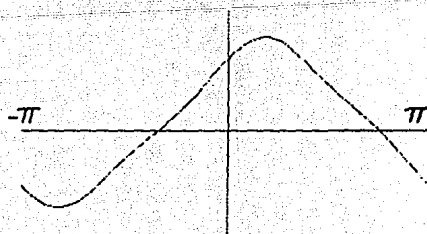
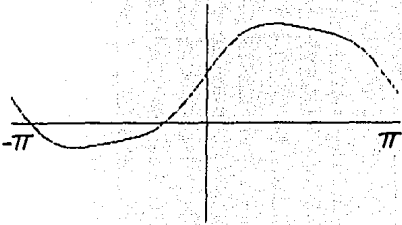
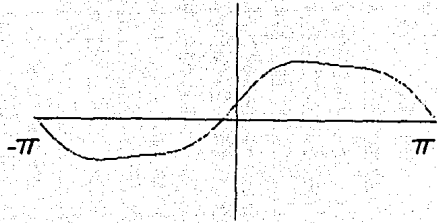


FIGURA 15 CURVAS DE ANDREWS PARA LAS ESTACIONES: SAN IGNACIO, LA VERANERA, URIQUE, GUERACHIC, ALAMOS Y CAZANATE ( REGION "A" )

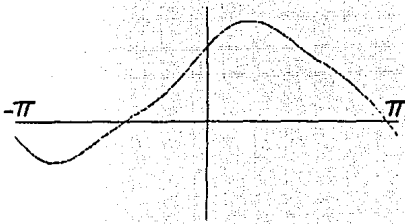
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : JAINA



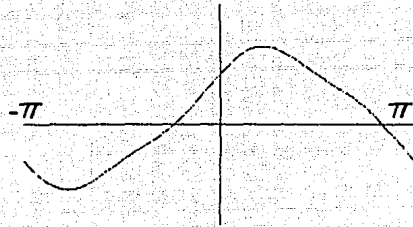
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : NARANJO



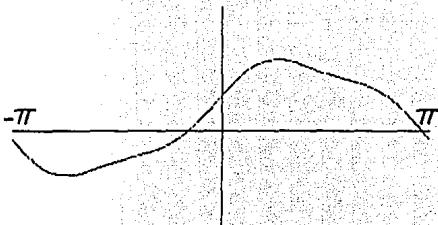
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : ZOPILOTE



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : GUAMUCHIL



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : PERICOS



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : TIERRA BLANCA

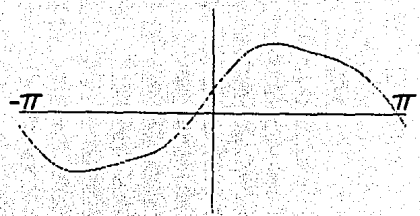
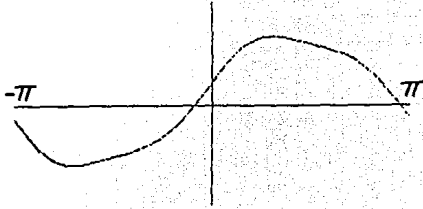
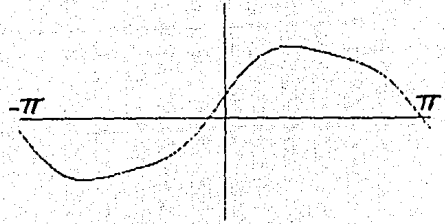


FIGURA 16 CURVAS DE ANDREWS PARA LAS ESTACIONES: JAINA, NARANJO, ZOPILOTE,  
GUAMUCHIL, PERICOS Y TIERRA BLANCA ( REGION "A" )

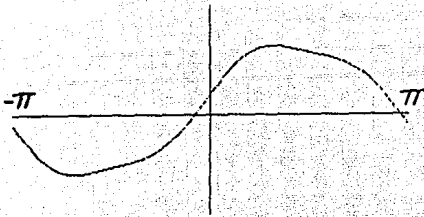
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : PUENTE SUD PACIFICO



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : PALOS BLANCOS



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : EL VAREJONAL

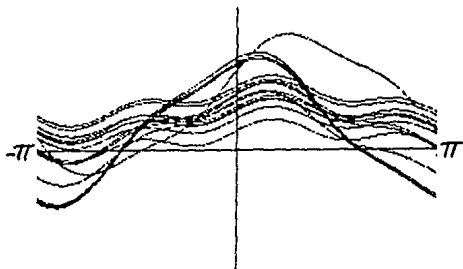


TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : BADI RAGUATO

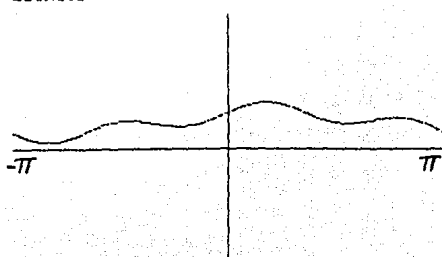


FIGURA 17 CURVAS DE ANDREWS PARA LAS ESTACIONES: PTE. SUD PACIFICO,  
PALOS BLANCOS, EL VAREJONAL Y BADI RAGUATO ( REGION "A" )

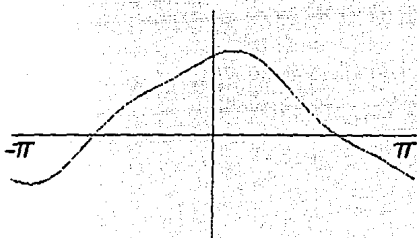
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : REGION "B" (15 estaciones)



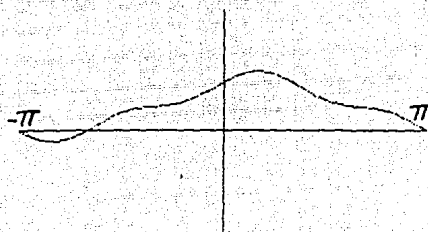
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : TOAHAYANA



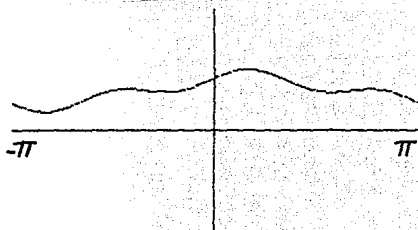
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : TECUSIAPA



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : LOS MOLINOS



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : GUATENIPA



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : LA HUERTA

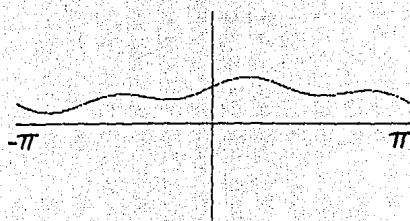
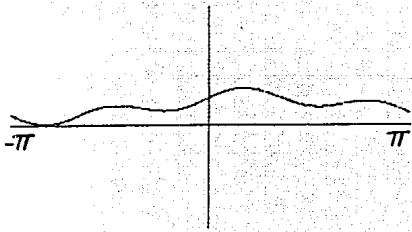
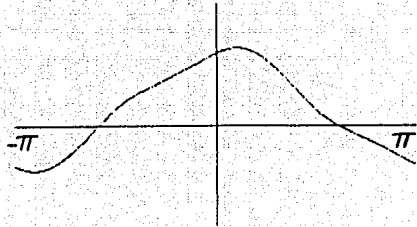


FIGURA 18. CURVAS DE ANDREWS PARA LAS ESTACIONES: TOAHAYANA, TECUSIAPA, LOS MOLINOS, GUATENIPA Y LA HUERTA ( REGION "B" )

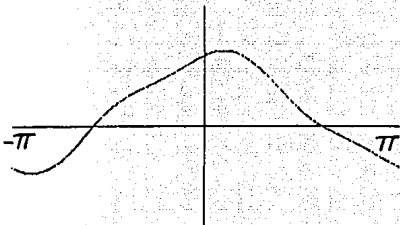
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : PUENTE CAÑEDO



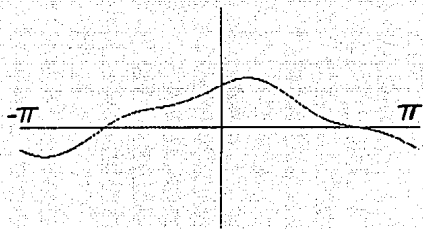
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : SANALONA



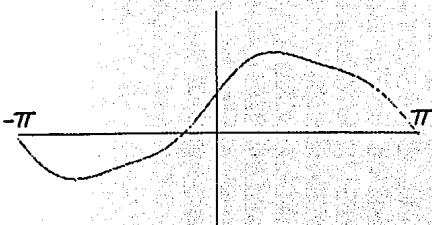
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : PICACHOS



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : TAMAZULA



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : EL BLEDAL



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
 ESTACION : SANTA CRUZ

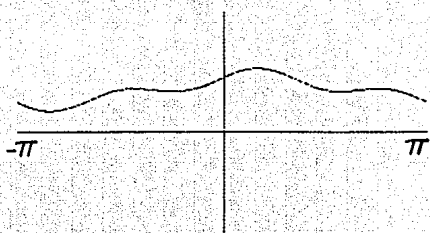
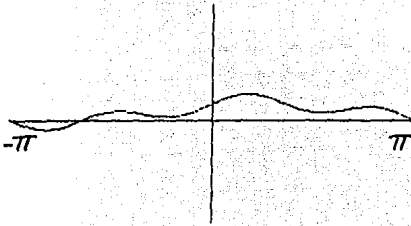


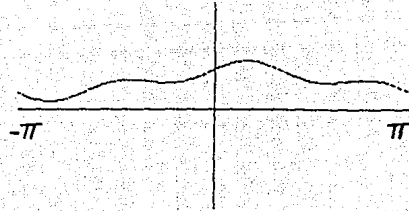
FIGURA 19 CURVAS DE ANDREWS PARA LAS ESTACIONES: PTE. CAÑEDO, SANALONA, PICACHOS, TAMAZULA, EL BLEDAL Y SANTA CRUZ (REGION "B" )



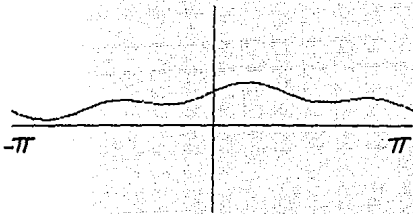
TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : ACATITAN



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : IXPALINO



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : PIAXTLA



TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS  
ESTACION : EL QUELITE



FIGURA 20 CURVAS DE ANDREWS PARA LAS ESTACIONES: ACATITAN, IXPALINO,  
PIAXTLA Y EL QUELITE ( REGION "B" )

La hipótesis nula debe rechazarse ( $H_0$ : no existen diferencias entre el sitio,  $CV_i$  dentro de su grupo), por lo tanto se dice que *para esta agrupación inicial, sí existen diferencias entre los CV de cada grupo*. ver Tabla 6.10

La segunda fase de fusión, consistió en trasladar estaciones del grupo A al grupo B y viceversa. Cabe recordar, que el procedimiento de relocalizar estaciones de un grupo a otro, puede realizarse a través de la metodología propuesta en el apartado 2.3.1. (método propuesto para formar grupos), pero dado que se tiene la agrupación previa del método del trazo multidimensional se prefirió ésta como indicativo para reubicar estaciones.

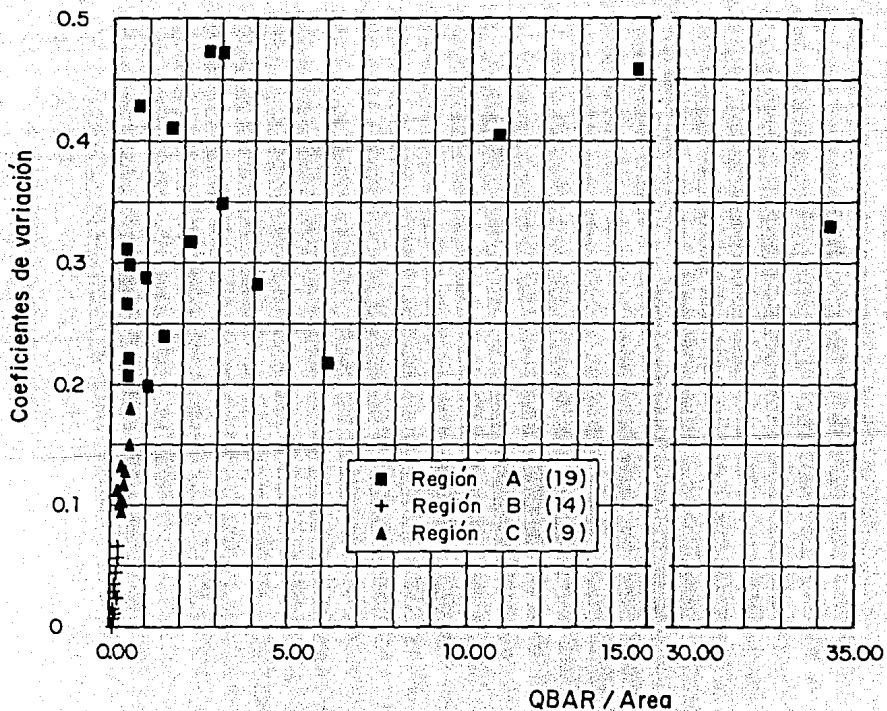


FIGURA 21 COEFICIENTES DE VARIACION vs QBAR/AREA

Así, las estaciones: San Blas, El Malone, Huites, Palo Dulce, San Francisco, San Ignacio y Querachic, fueron ubicadas en el grupo A, en donde se encuentran la mayoría de las estaciones, que según las curvas de Andrews, son homogéneas entre sí, tal y como lo muestra la tabla 6.11. Al probar el estadístico  $F$  en esta segunda fase, se obtuvo que:

$$F(1.002) < F_{\text{crítico}}(3.238)$$

Lo cual nos indica que debe aceptarse la hipótesis nula y considerar que no existen diferencias entre el sitio  $CV_i$  dentro de su grupo. Esto en sí, es ya una aceptable regionalización. Sin embargo, cabe hacer la pregunta: ¿Puede, con ésta agrupación considerarse que cada grupo es homogéneo?

La respuesta es sencilla, el análisis de varianza en una dirección, el cual se ha venido probando hasta esta segunda interacción, no podría responder a ésta pregunta. Un análisis de varianza en dos direcciones (apartado 3.3.) podría indicarnos qué tan homogéneo es cada grupo, y también si existen o no diferencias entre los coeficientes de variación dentro de su grupo. El programa ANOVA2.BAS desarrollado en el apartado 3.3., y que se muestra en el anexo 1, permite estimar el estadístico  $F$  en ambas direcciones (renglones y columnas).

Al aplicar este criterio en esta segunda interacción se obtienen resultados muy interesantes:

$$\text{Entre estaciones } F(3.143) > F_{\text{crítico}}(1.899)$$

$$\text{Entre grupos } F(9.352) > F_{\text{crítico}}(3.259)$$

Es fácil observar como no puede aceptarse ninguna de las dos hipótesis nulas  $H_0$  y debe concluirse inequívocamente, que existen diferencias entre los sitios  $CV_i$  dentro de su grupo y por consiguiente ninguno de los 3 grupos puede considerarse homogéneo.

El tercer ciclo de fusión, y de aquí en adelante el análisis comprenderá no solo el cálculo en una dirección sino un análisis de varianza en 2 direcciones.

La siguiente ubicación de estaciones, consistió en trasladar 4 estaciones del grupo C: (que hasta ahora no se había trabajado) Bamicori, Las Cañas, Choix y La Veranera al grupo A en donde, según la prueba del trazo multidimensional de Andrews, deben ubicarse; los resultados fueron los siguientes:

**TABLA 6.10 ANALISIS DE VARIANZA  
PARA LOS COEFICIENTES DE VARIACION  
( 1a Etapa )**

GRUPO A		GRUPO B		GRUPO C	
ESTACION	CV	ESTACION	CV	ESTACION	CV
ZAPOTTILAN	0.3108	SAN BLAS	0.0026	BAMICORI	0.1274
LA TINA	0.4282	EL MAHONE	0.0091	LAS CAÑAS	0.1121
CHINIPAS	0.2812	HUITES	0.0161	CHOIX	0.1015
URIQUE	0.2392	PALO DULCE	0.0288	LA VERANERA	0.1005
ALAMOS	0.4099	SN FCO.	0.0108	GUATENIPA	0.0942
CAZANATE	0.2219	SN IGNACIO	0.0011	SANALONA	0.1323
TOAHAYANA	0.1989	GUERACHIC	0.0064	EL BLEDAL	0.1492
LOS MOLINOS	0.4564	JAINA	0.0356	SANTA CRUZ	0.1798
NARANJO	0.2062	TECUSIAPA	0.0663	EL QUELITE	0.1162
ZOPILOTE	0.4714	BADIRAGUAT	0.0563		
GUAMUCHIL	0.2878	TAMAZULA	0.0235		
PERICOS	0.4042	ACATITAN	0.0068		
TIERRA BLANC	0.2166	XPALINO	0.0442		
PTE S PACIFICO	0.3167	PIAXTLA	0.0007		
PALOS BLANCO	0.2661				
EL VAREJONAL	0.2978				
LA HUERTA	0.3465				
PTE CAÑEDO	0.4723				
PICACHOS	0.3283				

ANOVA EN UNA DIRECCION

Grupos	Elementos	Suma	Promedio	Varianza
Columna 1	19	6.1604	0.3242	0.0085
Columna 2	14	0.3083	0.0220	0.0005
Columna 3	9	1.1132	0.1237	0.0007

Fuente Variación	SC	gl	CM	F	F critico
Entre Tratam.	0.7732	2	0.3866	91.6883	3.2381
Error	0.1644	39	0.0042		
Total	0.9376	41			

**TABLA 6.11 ANALISIS DE VARIANZA  
PARA LOS COEFICIENTES DE VARIACION  
( 2a. Etapa )**

GRUPO A		GRUPO B		GRUPO C	
ESTACION	CV	ESTACION	CV	ESTACION	CV
ZAPOTTILAN	0.3108	TIERRA BLAN	0.2166	BAMICORI	0.1274
LA TINA	0.4282	PTE S PACIFIC	0.3167	LAS CAÑAS	0.1121
CHINIPAS	0.2812	PALOS BLANC	0.2661	CHOIX	0.1015
URIQUE	0.2392	EL VAREJONA	0.2978	LA VERANERA	0.1005
ALAMOS	0.4099	LA HUERTA	0.3465	GUATENIPA	0.0942
CAZANATE	0.2219	PTE CAÑEDO	0.4723	SANALONA	0.1323
TOAHAYANA	0.1989	PICACHOS	0.3283	EL BLEDAL	0.1492
LOS MOLINOS	0.4564	JAINA	0.0356	SANTA CRUZ	0.1798
NARANJO	0.2062	TECUSIAPA	0.0663	EL QUELITE	0.1162
ZOPILOTE	0.4714	BADIRAGUAT	0.0563		
GUAMUCHIL	0.2878	TAMAZULA	0.0235		
PERICOS	0.4042	ACATITAN	0.0068		
SAN BLAS	0.0026	IXPALINO	0.0442		
EL MAHONE	0.0091	PIAXTLA	0.0007		
HUTTES	0.0161				
PALO DULCE	0.0288				
SN FCO.	0.0108				
SN IGNACIO	0.0011				
GUERACHIC	0.0064				

**ANOVA EN UNA DIRECCION**

Grupos	Elementos	Suma	Promedio	Varianza
Columna 1	19	3.991	0.2101	0.0309
Columna 2	14	2.4777	0.1770	0.0254
Columna 3	9	1.1132	0.1237	0.0007

Fuente Variación	SC	gl	CM	F	F critico
Entre Tratam.	0.0458	2	0.0229	1.0018	3.2381
Error	0.8918	39	0.0229		
Total	0.9376	41			

**ANOVA EN DOS DIRECCIONES**

Fuente Variación	SC	gl	CM	F	F critico
Filas	0.6598	18	0.0367	3.1429	1.8986
Columnas	0.2181	2	0.1091	9.3516	3.2594
Error	0.4199	36	0.0117		
Total	1.2978	56			

Anova en una dirección, ver Tabla 6.12

$$F(0.301) < F_{\text{crítico}}(3.238)$$

*(Se acepta  $H_0$ ) no hay diferencias*

Anova en dos direcciones:

$$\text{Entre estaciones: } F(2.971) > F_{\text{crítico}}(1.789)$$

$$\text{Entre grupos: } F(14.434) > F_{\text{crítico}}(3.209)$$

*(Se rechaza  $H_0$ ) "Hay diferencias de los sitios  $CV_i$  entre los grupos"*

La cuarta etapa de liga consiste en ubicar las estaciones del grupo C dentro del grupo B, esto previendo no dejar grupos con 23 estaciones (región A) y un grupo único de 5 estaciones, como es el caso del grupo C (atendiendo sobre todo, una vez más al estudio anterior del trazo multidimensional).

De ésta manera, la cuarta etapa comprende a la región A formada por 23 estaciones y la región B por 19. Los resultados después de aplicar los programas ANOVA.BAS y ANOVA2.BAS fueron:

ANOVA en una dirección, ver Tabla 6.13

$$F(0.325) < F_{\text{crítico}}(4.085)$$

*(Se acepta  $H_0$ )*

ANOVA en dos direcciones:

$$\text{Entre las estaciones } F(1.671) < F_{\text{crítico}}(2.048)$$

$$\text{Entre grupos: } F(2.060) < F_{\text{crítico}}(4.301)$$

*(Se acepta  $H_0$ )*

De estos valores puede concluirse que, al aceptar las dos hipótesis nulas  $H_0$  efectivamente es más correcto tomar solo dos regiones y no tres como se hizo en los primeros 3 pasos.

Además, al evaluar el estadístico  $F_i$  y aceptar  $H_0$  sabemos que para cada sitio  $CV_i$  no hay diferencias entre su grupo, ni entre los demás grupos (en éste caso solo entre el grupo B); es decir, una estación  $CV_i$  está correctamente localizada dentro de su grupo, y el grupo de esta manera se considera homogéneo.

**TABLA 6.12 ANALISIS DE VARIANZA  
PARA LOS COEFICIENTES DE VARIACION  
( 3a. Etapa )**

GRUPO A		GRUPO B		GRUPO C	
ESTACION	CV	ESTACION	CV	ESTACION	CV
ZAPOTITLAN	0.3108	TIERRA BLAN	0.2166	GUATENIPA	0.0942
LA TINA	0.4282	PTE S PACIFIC	0.3167	SANALONA	0.1323
CHINIPAS	0.2812	PALOS BLANC	0.2661	EL BLEDAL	0.1492
URIQUE	0.2392	EL VAREJONA	0.2978	SANTA CRUZ	0.1798
ALAMOS	0.4099	LA HUERTA	0.3465	EL QUELITE	0.1162
CAZANATE	0.2219	PTE CAÑEDO	0.4723		
TOAHAYANA	0.1989	PICACHOS	0.3283		
LOS MOLINOS	0.4564	JAINA	0.0356		
NARANJO	0.2062	TECUSIAPA	0.0663		
ZOPILOTE	0.4714	BADIRAGUAT	0.0563		
GUAMUCHIL	0.2878	TAMAZULA	0.0235		
PERICOS	0.4042	ACATITAN	0.0068		
SAN BLAS	0.0026	EXPALINO	0.0442		
EL MAHONE	0.0091	PIAXTLA	0.0007		
HUITES	0.0161				
PALO DULCE	0.0288				
SN FCO.	0.0108				
SN IGNACIO	0.0011				
GUERACHIC	0.0064				
BAMICORI	0.1274				
LAS CAÑAS	0.1121				
CHOIX	0.1015				
LA VERANERA	0.1005				

**ANOVA EN UNA DIRECCION**

Grupos	Elementos	Suma	Promedio	Varianza
Columna 1	23	4.4325	0.1927	0.0268
Columna 2	14	2.4777	0.1770	0.0254
Columna 3	5	0.6717	0.1343	0.0011

Fuente Variación	SC	gl	CM	F	F crítico
Entre Tratam.	0.0143	2	0.0071	0.3012	3.2381
Error	0.9234	39	0.0237		
Total	0.9376	41			

**ANOVA EN DOS DIRECCIONES**

Fuente Variación	SC	gl	CM	F	F crítico
Filas	0.6967	22	0.0317	2.9717	1.7889
Columnas	0.3076	2	0.1538	14.4342	3.2093
Error	0.4689	44	0.0107		
Total	1.4732	68			

**TABLA 6.13 ANALISIS DE VARIANZA  
PARA LOS COEFICIENTES DE VARIACION  
( 1a. Etapa )**

GRUPO A		GRUPO B	
ESTACION	CV	ESTACION	CV
ZAPOTTILAN	0.3108	TIERRA BLANC	0.2166
LA TINA	0.4282	PTE S PACIFICO	0.3167
CHINPAS	0.2812	PALOS BLANC	0.2661
URIQUE	0.2392	EL VAREJONAL	0.2978
ALAMOS	0.4099	LA HUERTA	0.3465
CAZANATE	0.2219	PTE CAÑEDO	0.4723
TOAHAYANA	0.1989	PICACHIOS	0.3283
LOS MOLINOS	0.4564	JAINA	0.0356
NARANJO	0.2062	TECUSIAPA	0.0663
ZOPILOTE	0.4714	BADIRAGUATO	0.0563
GUAMUCHIL	0.2878	TAMAZULA	0.0235
PERICOS	0.4042	ACATITAN	0.0068
SAN BLAS	0.0026	IXPALINO	0.0442
EL MAHONE	0.0091	PIAXTLA	0.0007
HÜTTES	0.0161	GUATENIPA	0.0942
PALO DULCE	0.0288	SANALONA	0.1323
SN FCO.	0.0108	EL BLEDAL	0.1492
SN IGNACIO	0.0011	SANTA CRUZ	0.1798
GUERACHIC	0.0064	EL QUELITE	0.1162
BAMICORI	0.1274		
LAS CAÑAS	0.1121		
CHOIX	0.1015		
LA VERANERA	0.1005		

**ANOVA EN UNA DIRECCION**

Grupos	Elementos	Suma	Promedio	Varianza
Columna 1	23	4.4325	0.1927	0.0268
Columna 2	19	3.1494	0.1658	0.0190

Fuente Variación	SC	gl	CM	F	F critico
Entre Tratam.	0.0076	1	0.0076	0.3252	4.0847
Error	0.9301	40	0.0233		
Total	0.9376	41			

**ANOVA EN DOS DIRECCIONES**

Fuente Variación	SC	gl	CM	F	F critico
Filas	0.6386	22	0.0290	1.6705	2.0478
Columnas	0.0358	1	0.0358	2.0597	4.3009
Error	0.3823	22	0.0174		
Total	1.0566	45			



Al observar la agrupación como resultado de esta etapa las estaciones Jaina y Tecusiapa que pertenecen al grupo B, en el análisis de Andrews son parte del grupo homogéneo A y lo mismo para las estaciones Guamuchil y Pericos, que con el criterio de Andrews se ubicaron en el grupo B y que en esta etapa están en el A. Cabe recordar que para la estación Guamuchil, tuvieron que probarse los criterios SECLUID y COSINE en el trazo multidimensional, ya que existía duda de a qué región pertenecía. De esta manera se incluye una etapa denominada de calibración en la cual se intercambian de grupo las estaciones mencionadas, obteniéndose los siguientes resultados:

ANOVA en una dirección Tabla 6.14

$$F(0.397) < F_{\text{crítico}}(4.085)$$

(Se acepta  $H_0$ )

ANOVA en dos direcciones:

$$\text{Entre estaciones: } F(1.870) < F_{\text{crítico}}(2.048)$$

$$\text{Entre grupos: } F(0.014) < F_{\text{crítico}}(4.301)$$

(Se acepta  $H_0$ )

La etapa de calibración carecería de sentido alguno, si no se tuviera un criterio para decidir sobre la mejora de la prueba, ya que tal vez se podría ensayar un par de reclasificaciones más, obteniéndose para ambos casos en la prueba ANOVA que las hipótesis  $H_0$  son aceptadas. Por tal, la mejor forma de probar que en la etapa de calibración se han mejorado los grupos, consiste en comparar los errores (SS) para ambos casos:

$$SS \text{ etapa 4 } (0.382) > SS \text{ etapa 5 } (0.368) \text{ calibración}$$

De tal forma, que al lograr disminuir el error en la etapa de calibración, se garantiza una mejor reubicación de las estaciones: Jaina, Tecusiapa, Guamuchil y Pericos. Es importante no perder de vista que en la etapa de calibración (si existe) se deben mover el mismo número de estaciones entre cada uno de los grupos. Esto con el objeto de no alterar el tamaño de las regiones ya clasificadas y por tanto pueda compararse dentro de un mismo criterio y tamaño de muestras, el error (SS) que servirá como guía de la correcta o incorrecta reubicación de las estaciones.

Finalmente, y dado que se trata de una prueba más rigurosa, debe probarse el estadístico  $F$  tal y como lo propone WILTSHIRE evaluado con el programa WILTSHIRE.BAS (ver apartado 2.3.1.). De esta manera se tiene que:

**TABLA 6.14 ANALISIS DE VARIANZA  
PARA LOS COEFICIENTES DE VARIACION  
( Etapa de calibración )**

GRUPO A		GRUPO B	
ESTACION	CV	ESTACION	CV
ZAPOTITLAN	0.3108	TIERRA BLANC	0.2166
LA TINA	0.4282	PTE S PACIFICO	0.3167
CHINIPAS	0.2812	PALOS BLANC	0.2661
URIQUE	0.2392	EL VAREJONAL	0.2978
ALAMOS	0.4099	LA HUERTA	0.3465
CAZANATE	0.2219	PTE CAÑEDO	0.4723
TOAHAYANA	0.1989	PICACHOS	0.3283
LOS MOLINOS	0.4564	BADIRAGUATO	0.0563
NARANJO	0.2062	TAMAZULA	0.0235
ZOPILOTE	0.4714	ACATITAN	0.0068
SAN BLAS	0.0026	IXPALINO	0.0442
EL MAHONE	0.0091	PIAXTLA	0.0007
HUITES	0.0161	GUATENIPA	0.0942
PALO DULCE	0.0288	SANALONA	0.1323
SN FCO.	0.0108	EL BLEDAL	0.1492
SN IGNACIO	0.0011	SANTA CRUZ	0.1798
GUERACHIC	0.0064	EL QUELITE	0.1162
BAMICORI	0.1274	GUAMUCHIL	0.2878
LAS CAÑAS	0.1121	PERICOS	0.4042
CHOIX	0.1015		
LA VERANERA	0.1005		
JAINA	0.0356		
TECUSIAPA	0.0663		

**ANOVA EN UNA DIRECCION**

Grupos	Elementos	Suma	Promedio	Varianza
Columna 1	23	3.8424	0.1671	0.0255
Columna 2	19	3.7395	0.1968	0.0204

Fuente Variación	SC	gl	CM	F	F critico
Entre Tratam.	0.0092	1	0.0092	0.3969	4.0847
Error	0.9284	40	0.0232		
Total	0.9376	41			

**ANOVA EN DOS DIRECCIONES**

Fuente Variación	SC	gl	CM	F	F critico
Filas	0.6883	22	0.0313	1.8701	2.0478
Columnas	0.0002	1	0.0002	0.0138	4.3009
Error	0.3681	22	0.0167		
Total	1.0566	45			

$$F_{\text{Wiltshire}}(0.191) > F_{\text{crítico}}(4.085)$$

(Se acepta  $H_0$ )

Y por tanto la región hidrológica 10 quedará subdividida en las siguientes regiones:

REGION A (23 estaciones)		REGION B (19 estaciones)	
Sn M. Zapotitlan	Palo Dulce	Tierra Blanca	Sanalona
La Tina	San Francisco	Puente Sud Pacífico	El Bledal
Chinipas	San Ignacio	Palos Blancos	Santa Cruz
Urique	Guerachic	El Varejonal	El Quelite
Alamos	Bamicori	La Huerta	Guamuchil
Cazamate	Las Cañas	Puente Cañedo	Pericos
Toahayana	Choix	Picachos	
Los Molinos	La Veranera	Badiraguato	
Naranja	Jaina	Tamazula	
Zopilote	Tecusiapa	Acatitan	
San Blas		Ixpálino	
El Mahone		Pixtla	
Huites		Guatenipa	

#### 6.5.4 Regiones homogéneas con base en la prueba de Langbein

La prueba de homogeneidad de Langbein, tal y como se comentó en el apartado 2.4. exige un análisis previo de frecuencias a través de la distribución general de valores extremos tipo 1 (ver subcapítulo 5.1.4); al aplicar este procedimiento a los registros de las 42 estaciones de la región hidrológica 10, se obtuvo que la mayoría de los ajustes se lograba al estimar los parámetros de la distribución, a través del método de momentos; siempre atendiendo al criterio del mínimo error cuadrático.

Una vez que se tienen los parámetros, se calculan los gastos medio ( $Q_{2.33}$ ) y de creciente ( $Q_{10}$ ) para las 42 estaciones de la región. A continuación se obtuvo la relación  $Q_{10}/Q_{2.33}$  tal como se muestra en la tabla 6.15; arrojando un promedio regional de 2.151. El siguiente paso es obtener los valores de ( $Q_{10}$ ) regionales. Así por ejemplo para el caso de la estación de Huites, se tiene que: para su gasto de  $Q_{2.33} = 2946.10 \text{ m}^3/\text{s}$ , el  $Q_{10}$  regional, es:

$$Q_{10} = 2.151 \cdot (2946.10); \quad Q_{10} = 6337.06 \text{ m}^3/\text{s}$$

Con sus respectivos parámetros:  $\mu = 1536.12$  y  $\alpha = 2436.92$  se sustituyen en la ecuación (2.22), haciendo  $k=0$  y  $F(x)=1-1/T$ . Despejando el valor de T (período de retorno) se tiene:

$$1 - \frac{1}{T} = \exp\left\{-\exp\left[-\frac{(6337.06 - 1536.12)}{2436.92}\right]\right\}$$

$$1 - \frac{1}{T} = \exp(-0.13945)$$

$$1 - \frac{1}{T} = 0.8698$$

$$\frac{1}{T} = 0.1302 ; \quad T = 7.683$$

que con los 53 años de registro de la estación Huites se obtienen las coordenadas ( 7.7 , 53 ) las cuales se ubican en la gráfica que propone Langbein con sus respectivos límites de confianza. La figura 22 muestra el trazo de los 42 pares de coordenadas, que se calcularon con ayuda del programa INDEX.BAS

De aquí puede concluirse que al caer todas las estaciones dentro de los límites de confianza, *toda la región 10 se considera hidrológicamente homogénea* (bajo la prueba de Langbein).

Esto como se demostró en los anteriores métodos de homogeneidad, no es cierto. Sin embargo, a manera de comprobación y aprovechando el fácil manejo del programa INDEX.BAS, se decidió probar la homogeneidad de las regiones A y B descritas en el punto anterior, así para la región A, que comprende 23 estaciones (tabla 6.16 y figura 23) se considera bajo esta prueba como una región homogénea.

Y para la región B que comprende las 19 estaciones restantes (tabla 6.17 y figura 24), según puede observarse en la figura, también es una región hidrológicamente homogénea.

Esta prueba lejos de considerarse intrascendente, reafirma una vez más la existencia y homogeneidad de los dos grupos considerados. Sin embargo cabe señalar, que el hecho de obtener en un principio homogeneidad en las 42 estaciones lleva a la idea de que esta prueba no es muy poderosa para discriminar o reubicar estaciones; pero, logra reafirmar la idea de probar la homogeneidad en los registros de los escurrimientos máximos, idea principal de Wiltshire en 1985 al analizar los coeficientes de variación de éstos.

**TABLA 6.15**  
**PRUEBA DE HOMOGENEIDAD DE LANGBEIN**  
**AJUSTE A TODAS LAS ESTACIONES**

	Q2.33	Q 10	Q10 / Q2.33	$\mu$	$\alpha$	Tr	N
<b>CUENCA DEL RIO FUERTE</b>							
SM ZAPOTTILAN	625.47	1467.12	2.35	334.19	503.45	7.97	22
SAN BLAS	4023.01	9052.74	2.25	2282.26	3008.61	8.83	10
LA TINA	106.74	305.28	2.86	38.03	118.76	5.54	24
BAMICORI	194.73	425.09	2.18	115.00	13.80	3.69	32
LAS CAÑAS	2125.65	5745.41	2.70	872.88	2165.22	6.04	24
EL MAHONE	199.51	237.93	1.19	186.21	22.99	3.89	20
CHOIX	393.37	832.18	2.12	241.50	262.48	10.52	29
HUITES	2946.10	7020.09	2.38	1536.12	2436.92	7.68	53
PALO DULCE	977.41	2536.54	2.60	437.81	932.62	6.47	29
CHINIPAS	1064.31	1790.20	1.68	813.08	434.20	30.49	21
SAN FRANCISCO	1726.52	3617.37	2.10	1072.11	1131.04	10.85	33
SAN IGNACIO	1623.48	2648.13	1.63	1236.40	634.45	34.45	19
LA VERANERA	413.87	859.18	2.08	259.75	266.37	11.18	14
URIQUE	305.92	504.53	1.65	237.18	118.81	35.08	19
GUERACHIC	891.35	2705.75	3.04	263.40	1085.31	5.11	13
ALAMOS	247.78	480.90	1.94	167.09	139.45	14.30	22
CAZANATE	555.95	1505.07	2.71	227.46	567.74	6.02	19
<b>CUENCA DEL RIO SINALOA</b>							
JAINA	1198.72	2750.10	2.29	661.80	927.98	8.40	46
TOAHAYANA	1049.68	1870.77	1.78	765.51	491.15	21.39	29
TECUSIAPA	976.68	2010.33	2.06	618.94	618.29	11.50	17
LOS MOLINOS	161.86	271.35	1.68	123.97	65.49	31.20	13
NARANJO	622.75	1477.56	2.37	36.90	511.32	7.76	47
ZOPILOTE	363.55	723.31	1.99	239.04	215.20	12.98	47
<b>CUENCA DEL RIO MOCORITO</b>							
GUAMUCHIL	606.74	1428.79	2.35	322.23	491.73	7.89	36
<b>CUENCA LAGUNA CAIMANERO</b>							
PERICOS	204.50	344.56	1.68	156.03	83.78	30.14	26
<b>CUENCA DEL RIO CULLIACAN</b>							
TIERRA BLANCA	1560.06	2325.81	1.49	1296.05	458.04	90.52	7
PTE. S PACIFICO	2963.35	5838.66	1.97	1968.23	1719.91	13.47	35
PALOS BLANCOS	1484.01	3735.23	2.52	704.88	1346.60	6.86	47
EL VAREJONAL	674.96	1491.70	2.21	392.29	488.55	9.26	7
BADIRAGUATO	960.30	3377.31	3.52	123.79	1445.78	4.35	27
GUATENIPA	1890.57	3707.28	1.96	1261.83	1086.69	13.73	21
LA HUERTA	930.96	1699.56	1.83	664.95	459.75	18.86	17
PTE. CAÑEDO	1117.21	2333.48	2.09	696.27	727.53	10.96	22
SANALONA	447.97	1106.98	2.47	219.89	394.20	7.11	42
PICACHOS	1141.49	1735.85	1.52	935.79	355.53	72.40	7
TAMAZULA	583.97	946.47	1.62	458.51	216.83	40.12	23
EL BLEDAL	286.31	642.79	2.25	162.94	213.23	8.88	48
<b>CUENCA DEL RIO SAN LORENZO</b>							
SANTA CRUZ	1269.88	2855.84	2.25	721.00	948.66	8.84	43
<b>CUENCA DEL RIO ELOTA</b>							
ACATITAN	1032.70	2160.03	2.09	642.55	674.33	10.91	31
<b>CUENCA DEL RIO PIAXTLA</b>							
IXPALINO	1319.36	2907.93	2.20	769.57	950.23	9.33	31
PIAXTLA	1421.88	3492.28	2.46	705.32	1238.45	7.20	16
<b>CUENCA DEL RIO QUELITE</b>							
EL QUELITE	469.12	1049.64	2.24	268.00	347.25	8.96	26

Promedio = 2.151

**TABLA 6.16**  
**PRUEBA DE HOMOGENEIDAD DE LANGBEIN**  
**AJUSTE A LA REGION "A" ( 23 ESTACIONES )**

	Q 2.33	Q 10	Q10 / Q2.33	mu	alfa	Tr	N
<b>CUENCA DEL RIO FUERTE</b>							
SM ZAPOTITLAN	625.47	1467.12	2.35	334.19	503.45	8.03	22
SAN BLAS	4023.01	9052.74	2.25	2282.26	3008.61	8.90	10
LA TINA	106.74	305.28	2.86	38.03	118.76	5.57	24
BAMICORI	194.73	425.09	2.18	115.00	13.80	4.07	32
LAS CAÑAS	2125.65	5745.41	2.70	872.88	2165.22	6.08	24
EL MAHONE	199.51	237.93	1.19	186.21	22.99	4.14	20
CHOIX	393.37	832.18	2.12	241.50	262.48	10.63	29
HUTES	2946.10	7020.09	2.38	1536.12	2436.92	7.75	53
PALO DULCE	977.41	2536.54	2.60	437.81	932.62	6.52	29
CHINIPAS	1064.31	1790.20	1.68	813.08	434.20	31.00	21
SAN FRANCISCO	1726.52	3617.37	2.10	1072.11	1131.04	10.96	33
SAN IGNACIO	1623.48	2648.13	1.63	1256.40	634.45	35.06	19
LA VERANERA	413.87	859.18	2.08	259.75	266.37	11.29	14
URIQUE	305.92	504.53	1.65	237.18	118.81	35.71	19
GUERACHIC	891.35	2705.75	3.04	263.40	1085.31	5.14	13
ALAMOS	247.78	480.90	1.94	167.09	139.45	14.47	22
CAZANATE	555.95	1505.07	2.71	227.46	567.74	6.06	19
<b>CUENCA DEL RIO SINALOA</b>							
JAINA	1198.72	2750.10	2.29	661.80	927.98	8.47	46
TOAHAYANA	1049.68	1870.77	1.78	765.51	491.15	21.71	29
TECUSIAPA	976.68	2010.33	2.06	618.94	618.29	11.62	17
LOS MOLINOS	161.86	271.35	1.68	123.97	65.49	31.73	13
NARANJO	622.75	1477.56	2.37	36.90	511.32	7.82	47
ZOPILOTE	363.55	723.31	1.99	239.04	215.20	13.13	47

Promedio = 2.158

**TABLA 6.17**  
**PRUEBA DE HOMOGENEIDAD DE LANGBEIN**  
**AJUSTE A LA REGION "B" ( 19 ESTACIONES )**

	Q 2.33	Q 10	Q10 / Q2.33	mu	alfa	Tr	N
<b>CUENCA DEL RIO MOCORITO</b>							
GUAMUCHIL	606.74	1428.79	2.35	322.23	491.73	7.82	36
<b>CUENCA LAGUNA CAJMANERO</b>							
PERICOS	204.50	344.56	1.68	156.03	83.78	29.54	26
<b>CUENCA DEL RIO CULIACAN</b>							
TIERRA BLANCA	1560.06	2325.81	1.49	1296.05	458.04	87.98	7
PTE. S PACIFICO	2963.35	5838.66	1.97	1968.23	1719.91	13.29	35
PALOS BLANCOS	1484.01	3735.23	2.52	704.88	1346.60	6.80	47
EL VAREJONAL	674.96	1491.70	2.21	392.29	488.55	9.16	7
BADRAGUATO	960.30	3377.31	3.52	123.79	1445.78	4.33	27
GUATENIPA	1890.57	3707.28	1.96	1261.83	1086.69	13.53	21
LA HUERTA	930.96	1699.56	1.83	664.95	459.75	18.55	17
PTE. CAÑEDO	1117.21	2333.48	2.09	696.27	727.53	10.82	22
SANALONA	447.97	1106.98	2.47	219.89	394.20	7.05	42
PICACHOS	1141.49	1735.85	1.52	935.79	355.53	70.48	7
TAMAZULA	583.97	946.47	1.62	458.51	216.83	39.24	23
EL BLEDAL	286.31	642.79	2.25	162.94	213.23	8.79	48
<b>CUENCA DEL RIO SAN LORENZO</b>							
SANTA CRUZ	1269.88	2855.84	2.25	721.00	948.66	8.75	43
<b>CUENCA DEL RIO ELOTA</b>							
ACAITTAN	1032.70	2160.03	2.09	642.55	674.33	10.78	31
<b>CUENCA DEL RIO PIAXTLA</b>							
KXPALINO	1319.36	2907.93	2.20	769.57	950.23	9.23	31
PIAXTLA	1421.88	3492.28	2.46	705.32	1238.45	7.14	16
<b>CUENCA DEL RIO QUELITE</b>							
EL QUELITE	469.12	1049.64	2.24	268.00	347.25	8.86	26

Promedio = 2.143

Por lo resultados obtenidos, debe descartarse la prueba de Langbein como un confiable criterio para subdividir cuencas, sin embargo, es ya una prueba clásica de homogeneidad y puede servir como revisión de las regiones previamente formadas, además, esta prueba es parte del método de la avenida índice (*index flood*) (Kite, 1978) y debe realizarse incuestionablemente para garantizar, sobre todo, la homogeneidad del análisis previo de frecuencias que esta técnica requiere.

Finalmente, al evaluar los resultados de las siete pruebas de homogeneidad:

- Distancia entre estaciones
- Trazo multidimensional (curvas de Andrews)
- Gráfica CV vs QBAR/AREA
- Análisis de varianza de los CV, en una dirección
- Análisis de varianza de los CV, en dos direcciones
- Estadístico "F" según Wiltshire
- Prueba de Langbein

Se concluye que la regionalización más conveniente, es la que se obtiene de análisis de varianza en dos direcciones; sin perder de vista que para llegar a ésta fue necesario ir tomando los resultados que arrojaban las pruebas anteriores, sobre todo para ahorrar tiempo en la reubicación de estaciones.

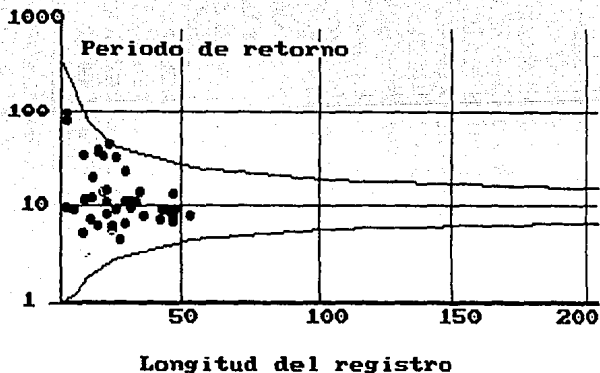


FIGURA 22 PRUEBA DE HOMOGENEIDAD DE LANGBEIN PARA LAS 42 ESTACIONES

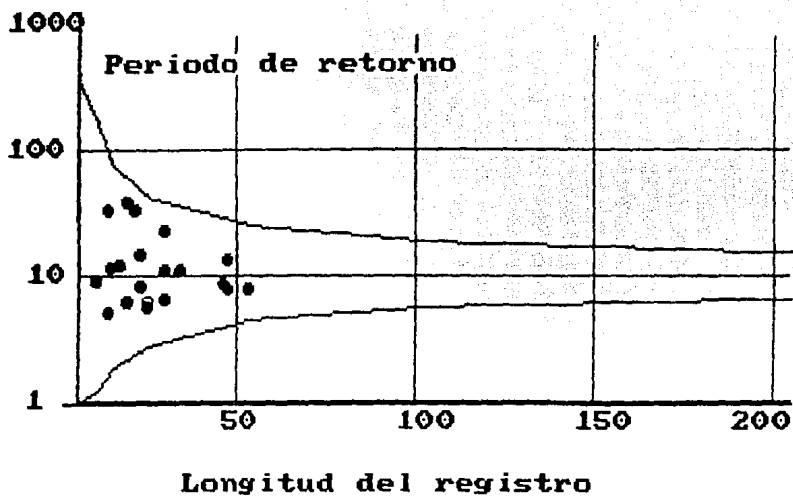


FIGURA 23 PRUEBA DE HOMOGENEIDAD DE LANGBEIN PARA LAS 23 ESTACIONES DE LA REGION "A"

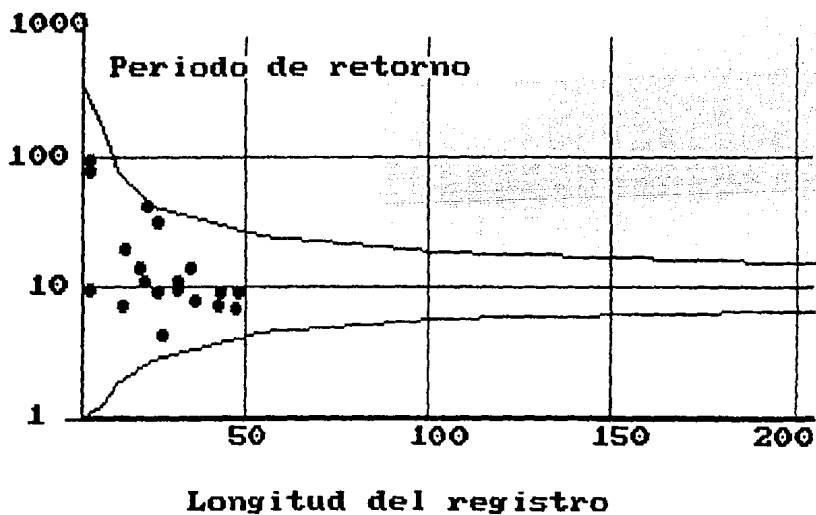


FIGURA 24 PRUEBA DE HOMOGENEIDAD DE LANGBEIN PARA LAS 19 ESTACIONES DE LA REGION "B"



## 6.6 Estimación en cuencas aforadas

El problema de la estimación de eventos de diseño en cuencas aforadas, se desarrolla a través de un análisis de frecuencias, en el cual, la estimación de los parámetros representa la parte básica y en si, el fundamento para lograr disminuir la incertidumbre en los escurrimientos máximos a inferir.

Esta estimación de parámetros se realiza siguiendo tres lineamientos fundamentales:

- Con base en un análisis univariado de frecuencias.
- Estimar los parámetros a través de un análisis bivariado.
- Extendiendo los registros de las muestras originales.

Primeramente se debe mencionar que el objetivo de un análisis de frecuencias es determinar la distribución de probabilidad que mejor se ajuste a los eventos en estudio, para lo cual se propone el procedimiento siguiente:

1. Selección de las muestras. Para este caso lo constituyen las series de escurrimientos máximos para las 23 estaciones que forman la región A.
2. Determinar los estadísticos muestrales de cada uno de los registros.
3. Ordenar cada una de las muestras de mayor a menor, asignándole un periodo de retorno (  $T$  ) y una probabilidad (  $P$  ) a cada uno de los escurrimientos

$$T = \frac{n+1}{m}$$

$$P = 1 - \frac{1}{T}$$

donde

T	período de retorno, en años
P	probabilidad de no excedencia
n	tamaño de la muestra
m	número de orden

4. Dibujar la muestra en papel de probabilidad. En el eje de las abscisas se miden los periodos de retorno (  $T$  ) mientras que los gastos máximos (  $Q$  ) se dibujan en el eje de las ordenadas.
5. Seleccionar las posibles distribuciones de probabilidad en competencia, en función de los estadísticos de la muestra.

6. Obtener los estimadores de los parámetros de las funciones seleccionadas, a través de los métodos propuestos en el apartado 5.1.1
7. Calcular a través de las pruebas de bondad de ajuste, la distribución de probabilidad que minimice el Error Estándar. La tabla 6.18 muestra los valores de estos errores para las diferentes distribuciones en competencia. Para este caso en 28 de las 42 estaciones la mejor combinación de los parámetros se logró ajustar a través de la distribución Gumbel Mixta. Estos resultados demuestran que los escurrimientos de la región presentan una fuerte componente ciclónica, y como los errores para las 14 estaciones restantes no difieren mucho, se toma a la distribución Gumbel Mixta, como la característica de esta región. Para el caso de la región A, en solo 7 estaciones no fue el Error Estándar de ajuste de la Gumbel Mixta la mejor distribución. La tabla 6.19 muestra los parámetros univariados de esta distribución para toda la región.
8. Calcular los eventos de diseño para los periodos de retorno de 2, 5, 10, 20, 50, 100, 500, 1 000, 5 000 y 10 000 años ( $F(x) = 1 - (1/T)$ ). Los valores de  $F(x)$  y los parámetros se sustituyen en la ecuación de la Gumbel Mixta, en donde solo la variable  $x = X(T)$  es incógnita. Con ayuda de un método numérico por ejemplo, el de bisección, se encuentran los valores de  $X(T)$ . La tabla 6.20 muestra los escurrimientos máximos asociados a esta distribución.

A continuación, se tomaron los parámetros bivariados obtenidos por Raynal, (1985) y Escalante, (1991) siguiendo el procedimiento del apartado 5.2 y se calcularon los errores, ver Tabla 6.21. En esa tabla se observa claramente que el error estándar disminuye significativamente en la mayoría de las estaciones de la región A.

Los resultados de estos dos procedimientos pueden resumirse si se observan y analizan las figuras 25 y 26 en donde claramente se ve que la estimación de parámetros a través de una distribución bivariada, por ejemplo, en la estación Huites produce el mejor ajuste; lo que también ocurre para la mayoría de las estaciones de la región A.

Tomando esta opción como el "mejor ajuste", y si se desea obtener el hidrograma de diseño para la estación Huites, para un periodo de retorno, por ejemplo de 10 000 años; basta tomar los parámetros para el "mejor ajuste" sustituirlos en la expresión de la distribución Gumbel Mixta (tal como se indica en el paso 8 mencionado anteriormente), y obtener el gasto para el periodo de retorno seleccionado. Finalmente con ayuda de algún método para construcción de hidrogramas, como se conoce la fisiografía de la región puede emplearse, por ejemplo, el hidrograma triangular o el adimensional (Dominguez, 1979) para dar la forma y distribución al escurrimiento.

DISTRIBUCION GUMBEL MIXTA UNIVARIADA  
METODO: MAXIMA VEROSIMILITUD  
ESTACION HUITES  
ERROR DE AJUSTE= 968.69

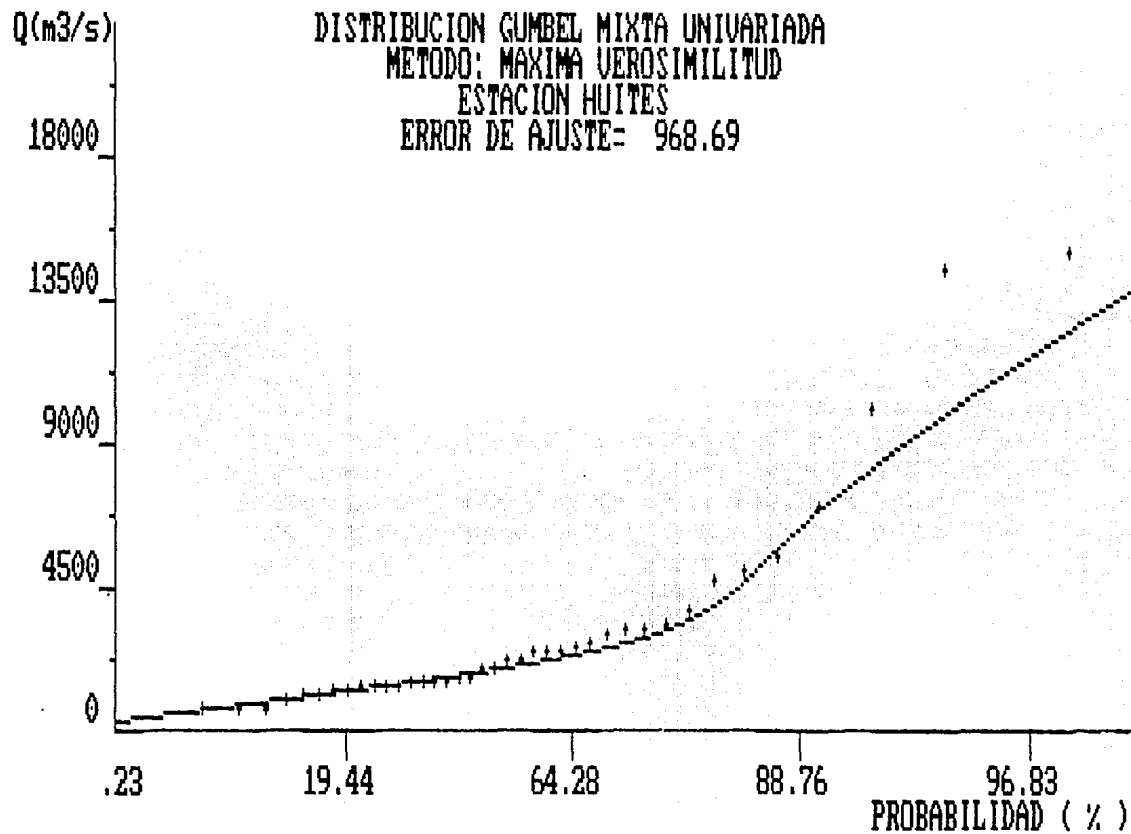


FIGURA 25 DISTRIBUCION GUMBEL MIXTA UNIVARIADA, ESTACION HUITES

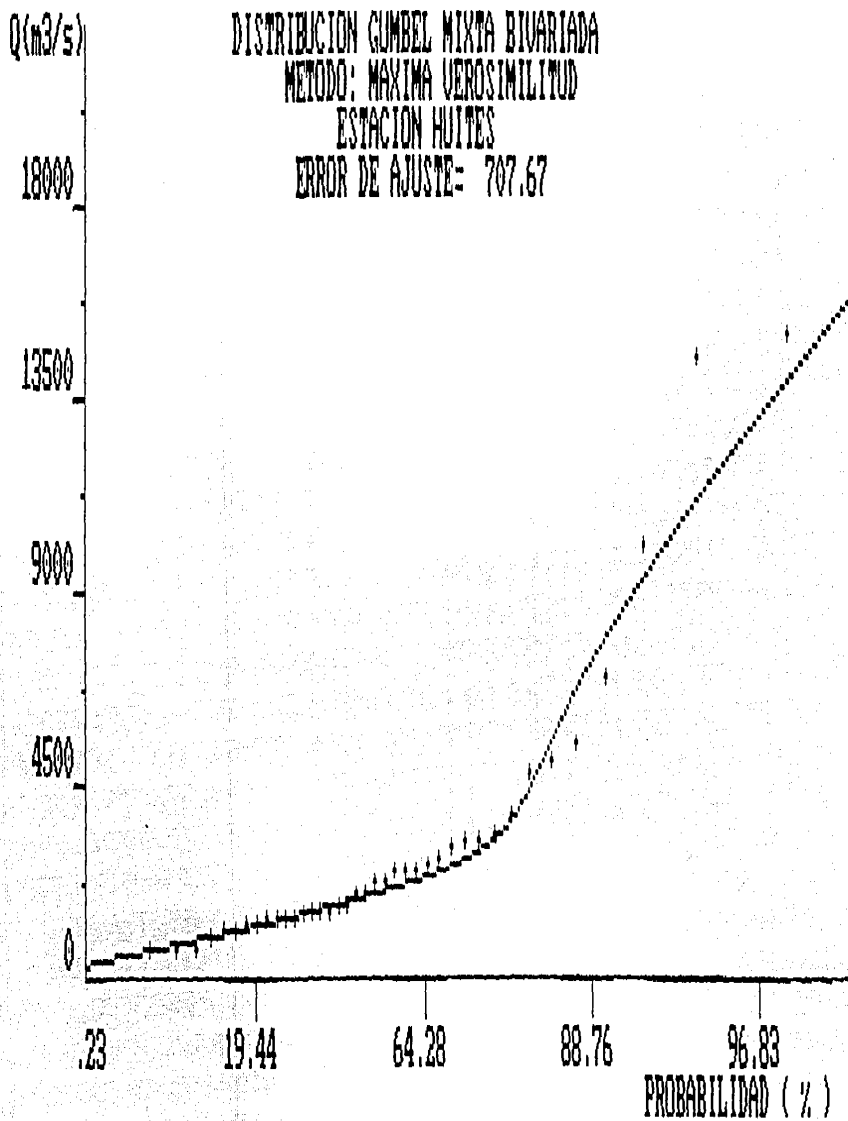


FIGURA 26 DISTRIBUCION GUMBEL MIXTA BIVARIADA, ESTACION HUITES

Siguiendo el ejemplo de Huites El gasto para 10,000 años de periodo de retorno resultó de: 34,376 m<sup>3</sup>/s. El tiempo base (en horas) puede calcularse con la ecuación de Mockus.

$$t_b = 2.67 t_p$$

El tiempo pico (en horas) puede expresarse como:

$$t_p = \sqrt{t_c} + t_r$$

$$d_e = 2\sqrt{t_c}$$

donde

$t_c$  tiempo de concentración, en h y puede calcularse con la expresión propuesta por Kirpich.

$d_e$  duración en exceso, en días

$$t_c = 0.000325 \frac{L^{0.77}}{S^{0.385}}$$

$t_r$  tiempo de retraso, en h y puede calcularse según Chow.

$$t_r = 0.005 \left( \frac{L}{\sqrt{S}} \right)^{0.64}$$

L es la longitud del cauce principal, en metros y S es su pendiente. Sustituyendo en estas expresiones para las características fisiográficas de Huites, se tiene:

$$L=267 \text{ km} \quad S=0.00685 \text{ m/km-1}$$

$$t_c = 0.000325 \frac{267000^{0.77}}{[0.00685(1/0.1)]^{0.385}} = 13.76 \text{ horas} = 0.57 \text{ día}$$

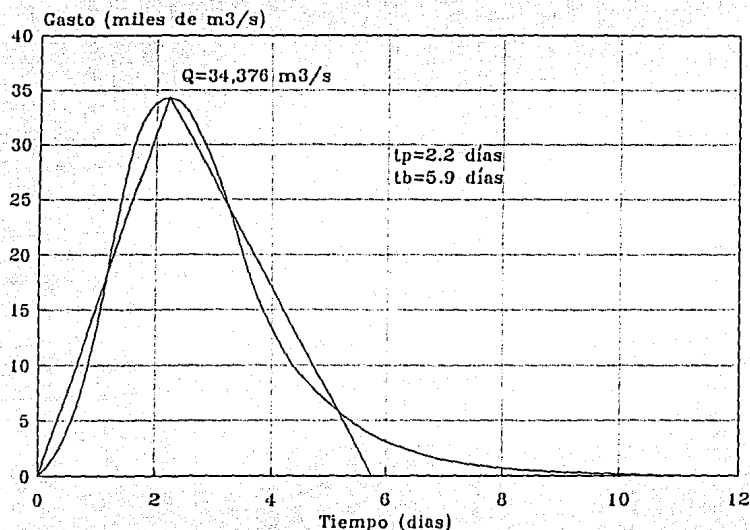
La duración en exceso es:  $d_e = 2\sqrt{0.57} = 1.51 \text{ día}$

$$t_r = 0.005 \left( \frac{267000}{\sqrt{[0.00685(1/0.1)]}} \right)^{0.64} = 35.04 \text{ horas} = 1.46 \text{ día}$$

$$t_p = 0.76 + 1.46 = 2.22 \text{ día}$$

$$t_b = 2.67(2.22) = 5.93 \text{ día}$$

La figura 27 muestra los hidrogramas unitario adimensional y triangular para el gasto de 10,000 años de período de retorno en la estación Huites.



**FIGURA 27 HIDROGRAMA TRIANGULAR Y UNITARIO ADIMENSIONAL PARA EL GASTO DE 10,000 AÑOS DE PERÍODO DE RETORNO DE LA ESTACION HUITES**

### 6.6.1 Extensión de registros

Con el objeto de mejorar la estimación de los parámetros en alguna de las estaciones que dentro de la subregión A registren periodos muy cortos de escurrimientos, se lleva a cabo una extensión de registros. Es decir, se crean muestras extendidas de las estaciones escasamente aforadas, en función de estaciones (estación base) con las cuales presentan una buena correlación de datos y además garantizan una mejora en la estimación de la media y la varianza extendidas.

Esta transferencia de información debe realizarse siguiendo la metodología propuesta en el apartado 4.4, en donde primeramente deben calcularse los coeficientes de correlación para todos los años comunes, en las 23 estaciones que componen la región A.

**TABLA 6.18**  
**MINIMOS ERRORES ESTANDAR**  
**DE AJUSTE UTILIZANDO DISTRIBUCIONES UNIVARIADAS**

**CUENCA DEL RIO FUERTE**

ESTACION	DISTRIBUCION	EE
SM ZAPOTTILAN	GAMA 2P	283.7
SAN BLAS	EXPONENCIAL	425.7
LA TINA	GUMBEL MIXTA	10.4
BAMICORI	GUMBEL MIXTA	31.4
LAS CAÑAS	GUMBEL MIXTA	523.5
EL MAHONE	GUMBEL	7.3
CHOIX	GUMBEL MIXTA	75.6
HUITES	GUMBEL MIXTA	612.1
PALO DULCE	GUMBEL MIXTA	409.3
CHINIPAS	GUMBEL MIXTA	81.8
SAN FRANCISCO	GUMBEL MIXTA	324.6
SAN IGNACIO	GUMBEL MIXTA	348.0
LA VERANERA	GUMBEL MIXTA	171.7
URIQUE	LOGNORMAL	40.8
GUERACHIC	GUMBEL MIXTA	701.1
ALAMOS	GUMBEL MIXTA	28.3
CAZANATE	EXPONENCIAL	211.2

**CUENCA DEL RIO SINALOA**

ESTACION	DISTRIBUCION	EE
JAINA	GUMBEL MIXTA	415.3
TOAHAYANA	GUMBEL MIXTA	110.6
TECUSIAPA	GUMBEL MIXTA	206.2
LOS MOLINOS	NORMAL	9.3
NARANJO	GUMBEL MIXTA	112.9
ZOPILOTE	GUMBEL MIXTA	41.5

**CUENCA DEL RIO MOCORITO**

ESTACION	DISTRIBUCION	EE
GUAMUCHIL	GUMBEL	195.6

**CUENCA DE LA LAGUNA DE CAIMANERO**

ESTACION	DISTRIBUCION	EE
PERICOS	GUMBEL	8.0

**CUENCA DEL RIO CULIACAN**

ESTACION	DISTRIBUCION	EE
TIERRA BLANCA	GUMBEL	293.0
PTE. S PACIFICO	GAMA 3P	528.4
PALOS BLANCOS	GAMA 2P	269.3
EL VAREJONAL	GAMA 3P	109.6
BADIRAGUATO	GUMBEL MIXTA	721.5
GUATENIPA	GUMBEL MIXTA	210.3
LA HUERTA	GUMBEL MIXTA	98.9
PTE. CAÑEDO	GAMA 2P	74.5
SANALONA	GUMBEL MIXTA	97.1
PICACHOS	LOGNORMAL	191.3
TAMAZULA	GUMBEL MIXTA	40.4
EL BLEDAL	GUMBEL MIXTA	68.9

**CUENCA DEL RIO SAN LORENZO**

ESTACION	DISTRIBUCION	EE
SANTA CRUZ	GUMBEL MIXTA	368.7

**CUENCA DEL RIO ELOTA**

ESTACION	DISTRIBUCION	EE
ACATITAN	GUMBEL MIXTA	336.8

**CUENCA DEL RIO PIAXTLA**

ESTACION	DISTRIBUCION	EE
IXPALINO	GUMBEL MIXTA	342.3
PIAXTLA	GUMBEL MIXTA	556.2

**CUENCA DEL RIO QUELITE**

ESTACION	DISTRIBUCION	EE
EL QUELITE	GUMBEL MIXTA	94.9

**TABLA 6.19 PARAMETROS DE LA DISTRIBUCION GUMBEL MIXTA  
PARA LAS DIFERENTES ESTACIONES**

	MU 1	ALFA 1	MU 2	ALFA 2	P
<b>CUENCA DEL RIO FUERTE</b>					
SM ZAPOTITLAN	259.140	245.682	1,117.320	648.486	0.798
SAN BLAS	1,803.834	810.222	10,080.029	1,449.381	0.800
LA TINA	52.069	43.588	787.196	289.502	0.958
BAMICORI	79.216	41.844	420.511	76.862	0.752
LAS CAÑAS	606.832	424.383	5,521.198	2,264.759	0.786
EL MAHONE	178.690	24.053	199.216	9.958	0.546
CHOIX	241.344	134.241	1,300.143	235.943	0.934
HUITES	1,408.928	656.131	5,994.867	3,355.593	0.814
PALO DULCE	578.278	311.451	5,652.882	664.912	0.965
CHINIPAS	394.360	26.315	1,064.175	349.040	0.236
SAN FRANCISCO	776.890	288.363	2,291.792	1,206.159	0.625
SAN IGNACIO	865.156	8.018	1,488.047	512.682	0.202
LA VERANERA	201.827	114.683	593.699	312.387	0.716
URIQUE	215.551	147.624	290.740	15.470	0.543
GUERACHIC	330.472	128.852	2,628.486	1,507.200	0.842
ALAMOS	128.918	93.121	464.771	84.461	0.805
CAZANATE	219.204	156.745	1,120.338	788.475	0.814
<b>CUENCA DEL RIO SINALOA</b>					
JAINA	597.549	258.171	1,838.403	1,094.330	0.748
TOAHAYANA	554.848	290.447	1,630.532	282.000	0.700
TECUSIAPA	515.652	275.814	2,001.848	638.010	0.821
LOS MOLINOS	112.785	58.222	250.352	17.424	0.868
NARANJO	244.704	186.116	1,581.303	392.168	0.818
ZOPILOTE	188.134	161.163	785.373	73.440	0.853
<b>CUENCA DEL RIO MOCORITO</b>					
GUAMUCHIL	304.321	216.629	1,093.256	700.825	0.841
<b>CUENCA LAGUNA CAIMANERO</b>					
PERICOS	137.252	70.301	323.314	19.491	0.834
<b>CUENCA DEL RIO CULIACAN</b>					
TIERRA BLANCA	1,327.293	588.576	1,333.422	75.698	0.621
PTE. S PACIFICO	559.468	279.463	2,975.849	1,413.800	0.275
PALOS BLANCOS	118.582	55.450	1,398.712	1,130.060	0.327
EL VAREGONAL	145.610	88.681	1,182.695	200.298	0.572
BADIRAGUATO	325.908	158.710	3,659.212	2,327.191	0.883
GUATENIPA	836.955	459.159	3,353.286	650.254	0.704
LA HUERTA	365.247	159.898	1,368.151	181.867	0.534
PTE. CAÑEDO	322.052	293.162	1,807.969	474.855	0.610
SANALONA	234.218	152.429	1,844.643	338.000	0.929
PICACHOS	635.315	208.723	1,184.022	259.992	0.324
TAMAZULA	435.517	148.854	1,260.517	24.098	0.916
EL BLEDAL	151.588	95.475	551.042	310.887	0.852
<b>CUENCA DEL RIO SAN LORENZO</b>					
SANTA CRUZ	680.838	411.470	2,944.569	1,157.654	0.870
<b>CUENCA DEL RIO ELOTA</b>					
ACATITAN	491.063	149.233	1,372.646	629.048	0.618
<b>CUENCA DEL RIO PIAXTLA</b>					
IXPALINO	765.098	388.441	3,198.840	1,578.029	0.895
PIAXTLA	645.505	188.806	3,530.821	1,133.011	0.811
<b>CUENCA DEL RIO QUELITE</b>					
EL QUELITE	222.505	155.443	1,224.341	283.318	0.859



**TABLA 6.20 GASTOS DE DISEÑO (en m<sup>3</sup>/s) PARA DIFERENTES PERIODOS DE RETORNO AJUSTADOS CON LA DISTRIBUCION GUMBEL MIXTA UTILIZANDO DISTRIBUCIONES UNIVARIADAS**

	PERIODO DE RETORNO EN AÑOS									
	2	5	10	20	50	100	500	1,000	5,000	10,000
<b>CUENCA DEL RIO FUERTE</b>										
SM ZAPOTITLAN	521	1,003	1,299	1,579	1,939	2,207	2,827	3,095	3,712	4,005
SAN BLAS	2,415	7,466	10,609	11,884	13,340	14,384	16,749	17,750	20,118	21,095
LA TINA	71	128	174	263	776	779	785	787	793	796
BAMICORI	117	383	471	535	611	666	791	845	965	1,026
LAS CAÑAS	988	4,939	7,711	9,745	12,168	13,929	17,945	19,630	23,634	25,391
EL MAHONE	198	214	226	239	259	275	314	330	369	386
CHOIX	304	491	683	1,222	1,540	1,726	2,120	2,284	2,662	2,833
HUITES	1,864	3,442	6,875	9,892	13,288	15,711	21,192	23,536	28,907	31,251
PALO DULCE	709	1,099	1,406	1,864	5,750	6,367	7,527	7,997	9,083	9,571
CHINIPAS	1,043	1,481	1,750	2,004	2,331	2,575	3,138	3,382	3,956	4,200
SAN FRANCISCO	1,143	2,628	3,704	4,637	5,795	6,648	8,607	9,449	11,329	12,110
SAN IGNACIO	1,496	2,125	2,519	2,892	3,372	3,730	4,560	4,914	5,763	6,056
LA VERANERA	355	616	776	927	1,122	1,267	1,602	1,747	2,076	2,223
URIQUE	294	344	451	561	700	804	1,043	1,145	1,380	1,490
GUERACHIC	413	700	2,620	4,081	5,639	6,736	9,205	10,255	12,696	13,673
ALAMOS	204	396	496	585	698	781	971	1,054	1,246	1,319
CAZANATE	324	667	1,334	2,037	2,835	3,404	4,688	5,238	6,544	7,032
<b>CUENCA DEL RIO SINALOA</b>										
JAINA	808	1,563	2,591	3,491	4,567	5,348	7,130	7,887	9,669	10,352
TOAHAYANA	1,005	1,529	1,846	2,145	2,529	2,816	3,477	3,761	4,420	4,688
TECUSIAPA	709	1,387	2,151	2,717	3,363	3,823	4,865	5,311	6,349	6,837
LOS MOLINOS	169	218	244	269	2,99.7	323	377	400	453	477
NARANJO	377	942	1,673	2,027	2,424	2,711	3,346	3,614	4,298	6,251
ZOPILOTE	326	573	727	872	1,060	1,200	1,522	1,661	1,979	2,125
<b>CUENCA DEL RIO MOCORITO</b>										
GUAMUCHIL	436	813	1,237	1,792	2,503	3,012	4,157	4,640	5,763	6,251
<b>CUENCA LAGUNA CAIMANERO</b>										
PERICOS	184	313	341	365	407	450	561	610	721	770
<b>CUENCA DEL RIO CULIACAN</b>										
TIERRA BLANCA	1,396	1,885	2,351	2,786	3,340	3,753	4,707	5,116	6,056	6,446
PTE. S PACIFICO	2,754	4,574	5,673	6,706	8,032	9,022	11,305	12,281	14,552	15,626
PALOS BLANCOS	1,052	2,576	3,463	4,293	5,355	6,147	7,972	8,766	10,548	11,329
EL VAREJONAL	324	1,276	1,448	1,601	1,792	1,933	2,258	2,397	2,711	2,857
BADIRAGUATO	414	686	2,111	5,010	7,548	9,272	13,111	14,747	18,556	20,313
GUATENIPA	1,329	3,295	3,935	4,453	5,084	5,546	6,599	7,057	8,106	8,595
LA HUERTA	802	1,474	1,627	1,764	1,937	2,065	2,360	2,485	2,784	2,906
PTE. CAÑEDO	794	1,971	2,388	2,752	3,206	3,541	4,310	4,640	5,421	5,763
SANALONA	307	524	761	1,779	2,219	2,482	3,047	3,285	3,834	4,054
PICACHO	1,137	1,466	1,667	1,856	2,099	2,281	2,699	2,882	3,297	3,468
TAMAZULA	510	734	1,040	1,266	1,296	1,318	1,386	1,454	1,686	1,808
EL BLEDAL	209	376	573	832	1,151	1,378	1,887	2,101	2,613	2,833
<b>CUENCA DEL RIO SAN LORENZO</b>										
SANTA CRUZ	923	1,664	2,677	3,796	5,019	5,869	7,765	8,570	10,450	11,329
<b>CUENCA DEL RIO ELOTA</b>										
ACATITAN	696	1,562	2,122	2,608	3,211	3,655	4,676	5,116	6,153	6,544
<b>CUENCA DEL RIO PIAXTLA</b>										
IXPALINO	973	1,587	2,366	3,897	5,650	6,831	9,425	10,548	13,087	14,063
PIAXTLA	782	1,442	3,848	4,864	6,010	6,828	8,680	9,474	11,329	12,110
<b>CUENCA DEL RIO QUELITE</b>										
EL QUELITE	318	632	1,174	1,460	1,756	1,963	2,427	2,625	3,077	3,272

**TABLA 6.21  
PARAMETROS DE LA DISTRIBUCION GUMBEL MIXTA  
OBTENIDOS POR DISTRIBUCIONES BIVARIADAS**

	MU 1	ALFA 1	MU 2	ALFA 2	P	ERROR
<b>CUENCA DEL RIO FUERTE</b>						
SM ZAPOTITLAN	61.06	17.88	559.37	406.2	0.238	287.21
SAN BLAS	1759.95	799.92	10201.59	1751.02	0.779	743.98
LA TINA	52.246	43.992	774.13	3.587	0.958	10.38
BAMICORI	79.32	41.63	422.76	79.86	0.753	30.17
LAS CAÑAS	608.7	435.08	6097.66	2550.87	0.729	529.75
EL MAHONE	178.69	24.05	199.22	9.96	0.546	8.73
CHOIX	242.13	133.04	1322.07	256.28	0.931	63.62
HUTES	1375.64	611.04	6490.04	3637.32	0.797	458.89
PALO DULCE	548.89	320	6682.75	129.85	0.962	102.42
CHINIPAS	394.36	26.31	1064.17	349.03	0.236	81.75
SAN FRANCISCO	804.84	284.65	2594.95	1341.72	0.638	228.57
SAN IGNACIO	865.24	7.849	1472.65	534.36	0.222	339.22
LA VERANERA	113.83	24.36	389.15	227.13	0.197	155.22
URIQUE	228.1	152.91	289.13	14.394	0.576	46.18
GUERACHIC	330.47	128.85	2628.73	1507.21	0.842	700.99
ALAMOS	99.67	68.68	352.81	124.77	0.609	23.91
CAZANATE	219.21	156.74	1120.32	788.62	0.814	321.84

**CUENCA DEL RIO SINALOA**

JAINA	588.58	240.79	2018.89	1158.74	0.741	371.53
TOAHAYANA	331.92	31.27	1026.15	458.22	0.26	87.4
TECUSIAPA	562.37	282.31	2890.9	374.33	0.862	126.61
LOS MOLINOS	77.95	38.78	190.48	33.11	0.456	17.56
NARANJO	244.57	182.07	1513.92	493.88	0.806	92.11
ZOPILOTE	41.46	32.08	330.54	204.58	0.212	38.56

MEDIA = 211.591

Con ayuda de la ecuación (4.55) se formó la matriz de correlación que se presenta en la tabla 6.23.

Cabe recordar que para poder correlacionar dos muestras, es necesario que se garantice la normalidad de los registros. Como ya se ha desarrollado anteriormente se toma la transformación de Box-Cox para normalizar las 23 muestras (tabla 6.22) El criterio de selección a seguir, consiste en analizar todas las posibles parejas de estaciones que presenten un coeficiente superior a 0.45; de esta forma del total de los 253 coeficientes de correlación, solo 83 cumplen con esta condición. Además para todas estas parejas debe probarse que existe una mejora en la media y la varianza de la muestra extendida.

Por ejemplo, para la pareja de estaciones: Huites (1941-1993) - San Francisco (1941-1973), cuyas muestras presentan 33 años en común y un coeficiente de correlación de 0.728, se desea extender la muestra de la estación San Francisco, tomando a Huites como estación base.

$$\begin{array}{lll} N1 = 33 & N2 = 20 & \hat{r} = 0.728 \\ \overline{Y}_1 = 2.014 & S_1(y) = 4.17 \times 10^{-2} & \overline{X}_2 = 2.493 \\ \overline{X}_1 = 2.475 & S_1(x) = 2.18 \times 10^{-2} & S_2(x) = 0.0377 \end{array}$$

Primeramente debe verificarse si habrá mejora en la estimación de la media. De la ecuación (4.69) se tiene:

$$\hat{r} > \left( \frac{1}{33-2} \right)^{1/2} < \hat{r}_{cm_t}$$

en donde  $\hat{r}_{cm_t}$  es función de N1 y se obtiene de la tabla propuesta en el apartado 4.4.1.

$$0.728 > 0.179 < 0.181$$

por lo tanto, habrá mejora en la estimación de la media.

Para la mejora en la varianza se toma la ecuación (4.72) y si se considera para el modelo la componente aleatoria  $\varepsilon_t$  ( $\theta = 1$ ) y con ayuda de la tabla 4.2, para N1=33 y N2=20 se tiene:

$$|\hat{r}| > 0.372$$

por lo tanto, habrá mejora en la estimación de la varianza.

Antes de seguir adelante, es necesario determinar bajo estas condiciones, cuántos años podrán extenderse sin que se perjudiquen las características estadísticas del registro de San Francisco. Es decir, el período efectivo de registro  $N_e$  (ecuación 4.73)

$$N_e = \frac{33+20}{1 + \frac{20}{33-2}(1-0.5293)} = 40.65$$

Por lo tanto si el registro original de San Francisco tiene una longitud de 33 años; con base en el registro de la estación Huites, podrán extenderse a lo más 7 años (33+7). La tabla 6.24 muestra los valores de las longitudes efectivas de registro para las cuencas en donde existió una mejora en la media y varianza, y que además tienen coeficientes de correlación mayor o igual a 0.45. Una vez verificado lo anterior, se procede a calcular la media de la muestra extendida, ver ecuación (4.67)

$$\bar{Y} = 2.014 + \frac{20}{(33+20)} \left[ 0.728 \left( \frac{4.172}{2.182} \right) \right] (2.493 - 2.475) = 2.01618$$

y la varianza de la muestra extendida, ver ecuación (4.62)

$$S^2(y) = \frac{1}{(33+20-1)} \left[ (33-1)(4.172 \times 10^{-2})^2 + (20-1)(1.391)^2 (2.182 \times 10^{-2})^2 + \dots \right. \\ \left. \dots + \frac{660}{(33+20)} (2.493 - 2.475)^2 + (20-1)(1)^2 (1.05037)^2 (1-0.5293)(4.172 \times 10^{-2})^2 \right] \\ S^2(y) = 0.00046$$

Finalmente, el modelo de regresión lineal simple de transferencia de información entre los gastos máximos anuales de las estaciones Huites y San Francisco, es el siguiente:

$$Y_t = 2.014 + (0.728) \left( \frac{4.172}{2.182} \right) (X_t - 2.475) + (1.05037)(1-0.5293)^{1/2} (4.172 \times 10^{-2}) \varepsilon_t$$

$$Y_t = 1.0720 + 0.3804 X_t + 0.0157 \varepsilon_t$$

Con esta expresión y tomando los 7 años a inferir de la estación Huites (1974 a 1980) se obtienen los datos pronosticados pero normalizados. La expresión inversa de la transformación Box-Cox ayuda a obtener valores reales de los escurrimientos inferidos:

DATO BASE	DATO INFERIDO	INFERIDO DESNORMALIZADO
2.519	2.03496	2538
2.446	2.00854	1024
2.510	2.03161	2208
2.449	2.01859	1383
2.536	2.05218	6347
2.523	2.03814	2924
2.467	2.01490	1232

A continuación se presenta una lista de las 13 combinaciones de estaciones de la región A en donde se llevó a cabo la extensión de registros, siguiendo el procedimiento antes mencionado, y garantizando en todos los casos, que existe mejora en la estimación de la media y la varianza de las muestras extendidas, además de lograr siempre el mayor periodo efectivo de registro.

COMPLETAR	CON EL REGISTRO DE	Ne	AÑOS QUE SE AGREGAN
S.M.Zapotitlán	Jaina	28	5
San Blas	Naranjo	15	5
Las Cañas	Jaina	29	5
Choix	Huites	32	3
Palo Dulce	Huites	32	3
Chinipas	Huites	26	5
San Francisco	Huites	40	7
San Ignacio	Huites	26	7
La Veranera	San Francisco	19	5
Urique	Huites	22	3
Guerachic	Huites	19	6
Toahayana	Jaina	33	4
Tecusiapa	Jaina	23	6

Finalmente la tabla 6.25 muestra los resultados ya desnormalizados de las 23 estaciones de la región A. Y aplicando el mismo procedimiento del análisis de frecuencias se obtiene un tercer juego de parámetros obtenido al ajustar las muestras extendidas a la distribución Gumbel Mixta (tabla 6.26).



TABLA 6.23 COEFICIENTES DE CORRELACION PARA LA REGION "A"

1	SM ZAPOTITLAN	SM ZAPOTITLAN																						
2	SAN BLAS	0.00	SAN BLAS																					
3	LA TINA	0.16	0.00	LA TINA																				
4	BAMICORI	0.08	0.00	0.20	BAMICORI																			
5	LAS CAÑAS	0.69	0.00	0.33	0.17	LAS CAÑAS																		
6	EL MAHONE	-0.08	0.00	0.25	-0.03	-0.13	EL MAHONE																	
7	CHOIX	0.46	0.00	0.08	-0.23	0.26	0.12	CHOIX																
8	HUITES	0.46	0.53	-0.23	-0.08	0.45	0.19	0.60	HUITES															
9	PALO DULCE	0.35	0.00	-0.19	-0.10	0.44	-0.25	0.70	0.55	PALO DULCE														
10	CHINIPAS	0.38	0.00	-0.27	-0.16	-0.41	-0.25	0.29	0.66	0.44	CHINIPAS													
11	SAN FRANCISCO	0.56	0.00	-0.02	0.06	0.69	-0.28	0.56	0.72	0.56	0.25	SAN FRANCISCO												
12	SAN IGNACIO	0.48	0.00	-0.21	0.19	0.49	0.01	0.51	0.73	0.39	0.47	0.86	SAN IGNACIO											
13	LA VERANERA	0.66	0.00	0.38	0.00	0.47	0.00	0.35	0.59	0.43	0.00	0.79	0.00	LA VERANERA										
14	URIQUE	0.58	0.00	-0.14	0.01	0.70	0.07	0.45	0.58	0.18	0.61	0.61	0.63	0.00	URIQUE									
15	GUERACHIC	0.48	0.00	-0.11	-0.26	0.51	0.00	0.76	0.75	0.80	0.00	0.80	0.00	0.47	0.00	GUERACHIC								
16	ALAMOS	0.15	0.94	-0.45	0.28	0.39	0.82	-0.03	0.26	0.19	-0.02	0.47	-0.67	0.53	-0.88	0.41	ALAMOS							
17	CAZANATE	0.31	0.00	-0.08	-0.12	0.07	-0.52	0.41	0.05	0.29	0.38	0.03	-0.01	0.00	0.19	0.00	0.99	CAZANATE						
18	JAINA	0.71	0.69	0.16	-0.06	0.70	0.27	0.39	0.37	0.30	0.11	0.76	0.32	0.64	0.27	0.72	0.32	0.30	JAINA					
19	TOAHAYANA	0.64	0.00	0.13	-0.15	0.68	0.00	0.51	0.54	0.41	0.54	0.71	0.56	0.51	0.43	0.68	0.05	0.12	0.63	TOAHAYANA				
20	TECUSIAPA	0.57	0.00	0.14	0.07	0.44	-0.17	0.49	0.65	0.37	0.57	0.69	0.87	0.01	0.86	0.77	0.04	0.12	0.74	0.80	TECUSIAPA			
21	LOS MOLINOS	0.41	0.00	0.55	-0.13	-0.23	-0.25	-0.13	-0.07	-0.37	-0.86	-0.21	0.55	-0.10	0.35	-0.06	-0.55	-0.03	-0.05	0.08	0.03	LOS MOLINOS		
22	NARANJO	0.18	0.71	0.34	0.16	0.23	0.26	0.49	0.40	0.24	-0.17	0.62	0.34	0.72	0.01	0.27	0.26	0.03	0.44	0.27	0.26	-0.06	NARANJO	
23	ZOPILETE	0.11	0.74	0.19	0.15	0.25	0.29	0.17	0.06	-0.11	-0.45	0.44	-0.09	0.04	-0.20	0.31	0.17	-0.10	0.42	0.21	0.19	-0.01	0.64	ZOPILETE

LAS CAÑAS	0.69	0.00	0.33	0.17	LAS CAÑAS										
EL MAHONE	-0.08	0.00	0.25	-0.03	-0.13	EL MAHONE									
CHOIX	0.46	0.00	0.08	-0.23	0.26	0.12	CHOIX								
HUITES	0.46	0.53	-0.23	-0.08	0.45	0.19	0.60	HUITES							
PALO DULCE	0.35	0.00	-0.19	-0.10	0.44	-0.25	0.70	0.55	PALO DULCE						
CHINIPAS	0.38	0.00	-0.27	-0.16	-0.41	-0.25	0.29	0.66	0.44	CHINIPAS					
SAN FRANCISCO	0.56	0.00	-0.02	0.06	0.69	-0.28	0.56	0.72	0.56	0.25	SAN FRANCISCO				

EJEMPLO: El coeficiente de correlación entre la estación Las Cañas y la estación Huites es de 0.45





TABLA 6.25 GASTOS MAXIMOS INSTANTANEOS, REGION "A" ( 23 ESTACIONES )  
REGISTROS EXTENDIDOS POR CORRELACION LINEAL SIMPLE

AÑO	S M ZAPOTITLAN	SAN BLAS	LA TINA	BAMICORI	LAS CAÑAS	EL MAHONE	CHOIX	HUITES	PALO DULCE	CHINIPAS	SAN FRANCISCO	SAN IGNACIO	LA VERANERA	URIQUE	QUERACHIC	ALAMOS	CAZANATE	JAINA	TOAHAYANA	TECUSIAPA	LOS MOLINOS	NARANJO	ZOPILOTE	
1938																								
1939																							277	162
1940																							170	35
1941		1562						2085*		1326							759						90	10
1942		2284						2531		1558							2065						1405	400
1943		12675			18019			14376		2292							6091						1776	819
1944		3447			893			2580		1376							580						221	110
1945		1636			1117			1499		894							714						403	116
1946		1154			2178			1165		559							746						304	194
1947		1142			2347			1127		1137			379				771						61	23
1948		3280			842			623		1668			539			137	693						659	690
1949		9200			10000			2298		4710			946			395	2614						1320	776
1950		3800			4579			3229		2920			1095			133	2336						374	182
1951		1139		58	946			677		493			234			77	437						31	75
1952		1683		153	2364			1266		695			88		248	594							96	386
1953		1526		124	1210			1025		678			120		309	42	546		828				90	487
1954		2619		405	1619			955	474	683			371		260	177	516		837				185	54
1955		4548		379	7477		112	4780	1284	3101		1343	595	542	1600	1483							475	206
1956				379	659		74	532	313	583		138	496	297	639	1283							57	377
1957				170	129		73	593	248	434		143	181	246	362	420							190	70
1958				192	3888		220	1045	455	3750		861	354	445	2232	1560	601	85	3093	1030				
1959				67	1624		299	1908	743	947	1572	384	525	669	616	343	387	33	593	334				
1960	563	60	89	8562	1150	15000	6800	2174	6640	3819	561	5300	376	2003	2239	1525	99	334	177					
1961	488	73	155	697	159	1396	472	749	8801	1147	250	320	114	795	793	581	116	119	230					
1962	213	75	46	329	376	1620	347	1098	9971	1211	285	422	115	1137	888	816	231	434	331					
1963	648	34	110	824	407	2702	5841	1493	2134	1969	407	1860	521	1226	1814	1547	161	433	806					
1964	609	120	94	607	250	1319	481	742	735	1296	565	224	547	175	454	392	271	206	223	82				
1965	127	17	650	172	464	1944	1360	1178	952	1359	272	292	495	258	650	329	496	82	534	293				
1966	1480	71	93	1507	197	275	2420	674	406	2144	1470	316	695	292	958	812	620	262	330	611				
1967	512	113	49	595	271	300	2506	635	843	1175	1341	282	758	7	21*	900	1251	1010	172	684	710			
1968	1131	778	438	10000	194	336	1534	530	363	1308	1164	216	504	74	90	1338	1298	767	273	1720	452			
1969	41	30	22	213	203	248	1508	1100	721	605	960	115	708	106	199	340	295	287	182	255	10			
1970	82	2	138	420	176	136	1558	390	417	980	860	88	745		218	357	311	212	167	435	838			
1971	498	115	44	766	150	360	2300	1283	1420	950	884	308	691		3223	1109	1060	894	600	318				
1972	720	10	70		164	273	2225	688	1152	2129	1740	274			715	932	908	2521	405	275				
1973	711	47	141		180	351	7960	1370	1516	4855	3075	319			185	1349	1570	3075	1150	518				
1974	479	110*	650		192*	194	4001	1588	1742	2538	1897	354			424	680	963	976*	800	298				
1975	69	110*	100		215	192	1067	383	430	1024	1002	308			111	488	368	466	349	561				
1976	69	29	495		205*	580	3233	740	420	2208	4020	293			198	900	809	942	1995	508				
1977	60	160	214		207	189	1119	333	752	1383	861	79			259	791	672	787	73	370				
1978	729	58	94		189	348	6178	535	1446	6347	1696	300			475	989	1500	1442	230	93				
1979	1529	2	54		196*	576	4443	1825	2158	2924	2210	716			509*	1620	1930	2467	104	43				
1980	198	156	363		189	417	1474	562	720	1232	876	338			685	400	331	395	755	320				
1981	2786	122	134		216	1700	2508	1180	980*	1765	285*				1365	2832	952*		2040	924				
1982	2633	193	54		217	748	1530	951	1001	1369	243				988	4440	2580		1630	796				
1983	14	72			165*	588	8000	1010*	1740	1795	528*				255	179	1713		889	189				
1984	390				218	607	5496	190*	1271	1413	563				375	694	736*		296	334				
1985	1313				224	512	3385	843	1605	1898	387				250	1213*	1580*		542*	447				
1986	397					284	1374										494*							
1987							1245																	
1988							2299																	
1989							1345																	
1990							11350																	
1991							2509																	
1992							2006																	
1993							1180																	
1994																								

Data inferido

**TABLA 6.26**  
**PARAMETROS DE LA DISTRIBUCION GUMBEL MIXTA**  
**OBTENIDOS DE LAS MUESTRAS EXTENDIDAS**

	MU 1	ALFA 1	MU 2	ALFA 2	P	ERROR
<b>CUENCA DEL RIO FUERTE</b>						
SM ZAPOTITLAN	259.722	247.604	1,514.773	603.357	0.808	175.673
SAN BLAS	1,782.267	811.696	10,081.250	1,450.044	0.867	753.319
LA TINA	52.246	43.992	774.130	3.587	0.958	10.388
BAMICORI	79.216	41.844	420.511	76.862	0.752	31.386
LAS CAÑAS	693.207	489.409	5,801.190	3,718.911	0.771	1,141.262
EL MAHONE	178.690	24.053	199.216	9.958	0.546	8.734
CHOIX	254.235	141.194	1,308.034	239.327	0.942	71.378
HUTES	1,408.928	656.131	5,994.867	3,355.593	0.814	612.050
PALO DULCE	569.506	312.160	6,815.405	664.549	0.969	185.431
CHINIPAS	394.339	26.298	1,035.739	353.185	0.181	94.255
SAN FRANCISCO	821.572	315.185	2,367.663	1,236.476	0.594	333.685
SAN IGNACIO	865.037	7.927	1,459.169	497.762	0.144	373.446
LA VERANERA	247.139	141.277	986.072	143.656	0.807	63.432
URIQUE	213.805	138.043	291.975	15.180	0.525	51.808
GUERACHIC	409.254	176.927	2,580.150	1,561.710	0.892	550.995
ALAMOS	97.641	67.846	346.625	117.898	0.599	28.342
CAZANATE	219.204	156.745	1,120.338	788.475	0.814	321.871
<b>CUENCA DEL RIO SINALOA</b>						
JAINA	597.549	258.171	1,838.403	1,094.330	0.748	415.319
TOAHAYANA	594.137	311.380	1,573.848	275.727	0.711	108.105
TECUSIAPA	528.160	268.808	1,913.288	596.498	0.795	160.437
LOS MOLNOS	78.271	38.985	189.831	32.712	0.461	17.942
NARANJO	244.704	186.116	1,581.303	392.168	0.818	112.993
ZOPILOTE	41.123	32.222	327.215	200.007	0.211	41.524

MEDIA = 246.251

## 6.7 Estimación en cuencas no aforadas

Una de las partes fundamentales en la metodología que se desarrolla en este trabajo, es lograr una serie de ecuaciones regionales que puedan ser empleadas, principalmente, en la estimación de eventos de diseño en lugares no aforados o con poca información. Estas ecuaciones deben contener el mayor número posible de características significativas; esto es, deben estar en función de la fisiografía de la región y del análisis de frecuencia de los gastos máximos instantáneos. La metodología plantea correlacionar las seis características fisiográficas de la región, con la serie de parámetros de la distribución de probabilidad obtenidos al analizar el comportamiento de los escurrimientos; (ver capítulo 5).

Cabe mencionar que también se correlacionaron las características fisiográficas de la región, con cada uno de los gastos de diseño para períodos de retorno de 2, 5, 10, 20, 50, 100, 500, 1 000, 5 000 y 10 000 años, obtenidos en el análisis de frecuencias univariado. Pero dado que no se obtuvieron los resultados esperados, se prefirió correlacionar la fisiografía con los cinco parámetros de la distribución Gumbel Mixta.

Una vez que se han discriminado las características fisiográficas que deben emplearse, se realiza la subdivisión de regiones homogéneas y se especifica la variable dependiente (parámetros o gastos), esto se efectúa con el fin de obtener ecuaciones regionales confiables, es decir, un análisis de regresión múltiple con estricto apego a sus consideraciones estadísticas básicas.

Cabe recordar que una de estas consideraciones, es garantizar que las muestras estén distribuidas en forma normal. Así, antes de pretender hacer un análisis de correlación y regresión, las muestras de las seis características fisiográficas, de los parámetros univariados, bivariados, extendidos y si se desea de los gastos, deberán cumplir con este requisito. De aquí en adelante y a manera de ejemplo, se desarrollarán las ecuaciones para la región A que comprende las 23 estaciones ya enumeradas anteriormente.

El primer paso consiste en normalizar las muestras, pero esta vez solo para los 23 registros de la región A. Esta vez la normalización se hace (para obtener una mayor precisión en las ecuaciones) a través de la transformación normalizante de doble potencia (ecuación 4.16) la cual garantiza, como ya se comentó una asimetría igual a cero y curtosis 3.0. Este proceso se desarrolló con la ayuda del programa BOXTWO.BAS que se muestra en el anexo 1.

A continuación se presenta el resumen de los resultados de esta normalización, y los valores de  $\lambda$  y  $\gamma$  que normalizaron las muestras tanto en el primer paso de Box-Cox, como en el de doble potencia, respectivamente. Para las características fisiográficas se tiene:

CARACTERISTICA	BOX-COX		DOBLE POTENCIA		
	$\lambda$	MEDIA	$\gamma$	ASIMETRIA	CURTOSIS
AREA	0.1499	17.3713	1.2590	-0.0727	3.0003
PRECMEDA	0.2399	16.6797	0.8940	0.0409	3.0005
LONGITUD	0.3399	13.0317	0.8980	0.0326	3.0003
PENDCAUC	0.4199	-2.0734	1.1020	0.0017	3.0000
FRECUENC	0.0000	-5.4139	0.5024	-0.8810	3.0001
CENTLON	0.0000	4.6804	1.0690	0.1960	3.0022

Los estadísticos para los cinco parámetros de la distribución Gumbel Mixta univariada son:

CARACTERISTICA	BOX-COX		DOBLE POTENCIA		
	$\lambda$	MEDIA	$\gamma$	ASIMETRIA	CURTOSIS
$\mu 1$	0.0299	6.0686	1.1360	0.0341	3.0000
$\alpha 1$	0.0099	4.6187	1.1920	0.0029	3.0017
$\mu 2$	-0.1100	4.8611	1.2170	0.0497	3.0015
$\alpha 2$	0.2699	14.8775	0.9269	-0.0193	3.0002
p	0.0000	-0.6071	1.5330	-0.9805	3.0001

Para los parámetros de la distribución Gumbel Mixta estimados en forma bivariada:

CARACTERISTICA	BOX-COX		DOBLE POTENCIA		
	$\lambda$	MEDIA	$\gamma$	ASIMETRIA	CURTOSIS
$\mu 1$	0.0399	6.2611	1.1520	0.0273	3.0011
$\alpha 1$	0.0199	4.7329	1.1840	0.0029	3.0001
$\mu 2$	-0.1100	4.8774	1.2910	0.0336	3.0007
$\alpha 2$	0.1999	11.1307	0.9299	-0.0846	3.0017
p	0.0000	-0.6116	1.5620	-0.9920	3.0006

Así, por ejemplo, el área de la cuenca en la estación Huites es 26020 km<sup>2</sup> su valor normalizado con el primer paso de Box-Cox será, tomando todos los decimales, igual a:

$$y = \frac{(26020^{0.1499993205} - 1)}{0.1499993205} = 23.9673$$

El valor normalizado con la transformación de doble potencia es:

$$t = (|23.9673 - 17.3713|)^{1.2590121031} = 10.7519$$

Al normalizar las áreas para las 23 estaciones, y calcular sus estadísticos se tiene un coeficiente de asimetría de -0.0727 y una curtosis de 3.0003. De esta forma, la muestra de las 23 áreas se considera completamente normal. Igualmente, se realizó para las otras cinco características fisiográficas y para los cinco parámetros. Las tablas 6.27 y 6.28 muestran respectivamente las características fisiográficas y los parámetros normalizados.

### 6.7.1 Ecuaciones regionales con base en parámetros estimados en forma univariada

Una vez que se tienen las muestras normalizadas, se realiza la regresión de las características fisiográficas con los parámetros; para ello se emplean las mismas expresiones de regresión que ayudaron a discriminar las variables. Esta vez la matriz estará formada por seis variables independientes (características fisiográficas) y la variable dependiente irá cambiando dependiendo del parámetro analizado.

En la tabla 6.29 se muestran los resultados de la regresión múltiple para el parámetro  $\alpha_1$ . A continuación se presenta a manera de ejemplo el cálculo de este parámetro (regional) para la estación Huites a partir de los resultados de la regresión. Procedimiento que servirá como base para construir las restantes ecuaciones de los otros cuatro parámetros.

Se toman los respectivos coeficientes y el término independiente, que relacionarán las características particulares de cada estación, con el parámetro en análisis a través de una ecuación lineal de regresión. Por ejemplo, al inferir el parámetro  $\alpha_1$  en la estación Huites se tiene: Si los coeficientes de la ecuación de regresión son:

VARIABLE	COEFICIENTE
Término Independiente	0.291770140
AREA	0.378314381
PRECMEDA	0.468930092
LONGITUD	-0.787009347
PENDCAUC	-4.369646183
FRECUENC	-0.811886890
CENTLON	59.652652500

La inferencia del parámetro univariado  $\alpha_1$  puede realizarse con la ecuación:

$$t_{\alpha_1} = 0.0292 + (0.3783\text{AREA}) + (0.4689\text{PRECMEDA}) - (0.787\text{LONGITUD}) - \dots \\ \dots - (4.3696\text{PENDCAUC}) - (0.8119\text{FRECUENC}) + (59.6527\text{CENTLON}) \quad (6.1)$$

Sustituyendo los valores *normalizados* de las características fisiográficas de la estación Huites se obtiene:

$$t_{\alpha_1} = 0.0292 + 4.0676 + 0.0146 - 2.5388 + 0.0394 - 0.2973 - 0.0835$$

$$t_{\alpha_1} = 0.0292 + 1.2018$$

$$t_{\alpha_1} = 1.2310$$

Este valor es el parámetro  $\alpha_1$  normalizado, por lo que para conocer su valor real es necesario desnormalizarlo. Empleando las ecuaciones inversas de 4.9 y 4.16 para desnormalizar este valor tanto de la transformación de Doble Potencia como de Box-Cox se tiene finalmente que el valor de  $\alpha_1$  es igual a:

$$y_{\alpha_1} = |1.2310|^{1/1.192} + 4.6187; \quad y_{\alpha_1} = 5.8092$$

$$\alpha_1 = [(5.8092 * 0.009999329) + 1]^{1/0.0009999329}$$

$$\alpha_1 = 283.352$$

De la misma manera se calculan y desnormalizan los otros cuatro parámetros. Cada ecuación regional, (una por cada parámetro) tiene la forma de la ecuación (6.1) sin embargo, por facilidad, cada una de las características se expresan a través de sus respectivos parámetros de desnormalización, esto con el fin de introducir directamente en la ecuación los valores originales de las características fisiográficas. Así, no será necesario desnormalizar, y sin ningún problema los parámetros que se obtengan de las ecuaciones regionales serán sustituidos en la expresión de la distribución de probabilidad (para este caso fue la Gumbel Mixta) para finalmente, inferir los eventos de diseño (escurrimientos máximos) para los periodos de retorno seleccionados.

TABLA 6.27 CARACTERISTICAS FISIOGRAFICAS  
NORMALIZADAS CON BOX COX DE DOBLE POTENCIA

	$\lambda =$	0.149999321	0.239999369	0.339999288	0.419999212	Logaritmo natural	Logaritmo natural
	$\gamma =$	1.259012103	0.893992007	0.897991955	1.102004767	0.502369881	1.069003224
	AREA	PRECMEDA	LONGITUD	PENDCAUC	FRECUENC	CENTLON	
<b>CUENCA DEL RIO FUERTE</b>							
SM ZAPOTITLAN	13.521680	-0.222084	5.977627	-0.068871	0.177866	-0.000053	
SAN BLAS	13.262030	-0.153217	5.458606	-0.066350	0.177866	0.000044	
LA TINA	-14.920420	-1.082652	-5.438878	-0.053490	-0.459795	0.004634	
BAMICORI	-15.979880	-0.146743	-6.325925	-0.038847	0.099791	0.003803	
LAS CAÑAS	11.969880	0.049697	4.555013	-0.030605	0.307049	-0.000764	
EL MAHONE	11.936230	0.056920	4.474191	-0.028965	0.307049	-0.000764	
CHOIX	-6.218555	0.885094	-2.532152	-0.097137	0.338174	0.001583	
HUITES	10.751900	0.031070	3.225883	-0.009013	0.366147	-0.001403	
PALO DULCE	0.683181	0.575421	1.307196	0.029314	-0.109736	0.002444	
CHINIPAS	-0.022241	0.566375	-0.576832	0.074405	0.230051	0.002110	
SAN FRANCISO	7.270523	-0.327357	2.798059	-0.004705	0.512149	-0.000053	
SAN IGNACIO	3.746618	-0.134270	1.958841	0.030782	0.637556	-0.004187	
LA VERANERA	-0.772757	-0.543312	-0.344404	0.085938	0.307049	-0.002600	
URIQUE	-0.882855	-0.550984	-0.729405	0.083187	0.338174	-0.002804	
GUERACHIC	0.641540	-0.007749	0.000327	0.053633	0.415176	-0.006828	
ALAMOS	-3.639781	-1.214651	-0.843479	-0.113201	-0.489405	0.006605	
CAZANATE	-4.832522	-1.086629	-2.061932	-0.123441	-0.546142	0.006748	
<b>CUENCA DEL RIO SINALOA</b>							
JAINA	1.970898	0.935978	1.107728	0.027729	0.476731	-0.002397	
TOAHAYANA	0.040661	1.034224	-0.576832	0.040712	0.600548	-0.003491	
TECUSIAPA	-1.137326	1.202670	-2.187449	0.079918	0.415176	-0.004538	
LOS MOLINOS	-11.794430	0.621221	-5.088375	0.140014	-0.894377	-0.003491	
NARANJO	-4.140938	-0.085297	-1.497603	-0.000409	-0.653814	0.003119	
ZOPILOTE	-10.270060	-0.520782	-3.179116	-0.008375	-0.627338	0.002645	

TABLA 6.28 PARAMETROS DE LA DISTRIBUCION GUMBEL MIXTA  
NORMALIZADOS CON BOX-COX DE DOBLE POTENCIA

	$\lambda =$				
	0.029999329	0.009999329	-0.110000655	0.269999355	Logaritmo natural
	1.136006355	1.192008972	1.21701014	0.926991582	1.533024907
	MU 1	ALFA 1	MU 2	ALFA 2	P
<b>CUENCA DEL RIO FUERTE</b>					
SM ZAPOTITLAN	-1.8312520	-1.9005200	-0.2449241	-0.1352599	-0.7510636
SAN BLAS	2.6258910	2.7100480	0.9178676	6.7554240	0.2308080
LA TINA	-2.0340420	-0.7234352	-0.0943731	-11.0504600	0.4158227
BAMICORI	-1.4616030	-0.7826219	-0.3757833	-5.7666620	0.1759962
LAS CAÑAS	0.9683410	1.7869070	0.6903282	9.9576700	0.1679386
EL MAHONE	-0.4116348	-1.4773040	-0.8182458	-9.7709540	0.0000823
CHOIX	-0.0773099	0.3384445	0.0597127	-2.2422580	0.3878171
HUITES	2.2364610	2.3977100	0.6909155	11.9870800	0.2466805
PALO DULCE	0.9307101	1.3553530	0.6652131	2.6289650	0.4245921
CHINIPAS	0.4324819	-1.3602480	0.0022527	-0.6071355	-0.7629133
SAN FRANCISO	1.3444490	1.2539050	0.2771675	5.7289170	0.0475658
SAN IGNACIO	1.5003830	-3.0024700	0.1071625	1.3523890	-0.9892532
LA VERANERA	-1.0149200	-1.5611570	-0.4425567	-2.7229810	-0.7166344
URIQUE	-0.1958402	0.4406431	-0.5822670	-9.0951230	-0.0001321
GUERACHIC	0.2233358	0.2958300	0.3343764	6.9789630	0.2795803
ALAMOS	-1.1807650	-0.2487582	-0.4808534	-4.5736020	0.0269613
CAZANATE	-0.1774925	0.5072626	0.0143035	3.4854480	0.2465038
<b>CUENCA DEL RIO SINALOA</b>					
JAINA	0.9755420	1.1106710	0.1878987	5.1999930	0.1714038
TOAHAYANA	0.2183709	-1.1519680	0.0015542	0.2344363	-0.5978528
TECUSIAPA	0.7763780	1.1959280	0.2219086	2.4242840	0.2549787
LOS MOLINOS	-1.4778950	-0.8674834	-0.8500915	-7.7322330	-0.0648392
NARANJO	-0.0638288	0.7059470	0.1297205	-0.0121830	0.2512230
ZOPILOTE	-2.3693650	-1.1021590	-0.5134946	-2.8508010	-0.9200394



**TABLA 6.29**  
**REGRESION DE LAS CARACTERISTICAS FISOGRAFICAS**  
**CON EL PARAMETRO ALFA 1**

*Regression Statistics*

Multiple R	0.4953
R Square	0.2453
Adjusted R Square	-0.0377
Standard Error	1.4905
Observations	23

*Analysis of Variance*

	df	Sum Squares	Mean Square	F	Significance F
Regression	6	11.5538	1.9256	0.8668	0.5397
Residual	16	35.5461	2.2216		
Total	22	47.0999			

	Coefficients	Stand Error	t Statistic	P-value	Lower 95%	Upper 95%	t critic
Intercept	0.0292	0.3240	0.0901	0.9291	-0.6576	0.7160	
x1	0.3783	0.2455	1.5410	0.1376	-0.1421	0.8988	2.1199
x2	0.4689	0.5902	0.7945	0.4354	-0.7822	1.7201	
x3	-0.7870	0.6018	-1.3078	0.2044	-2.0627	0.4887	
x4	-4.3696	7.9388	-0.5504	0.5876	-21.1991	12.4598	
x5	-0.8119	1.0652	-0.7622	0.4540	-3.0699	1.4462	
x6	59.6527	179.5928	0.3322	0.7429	-321.0670	440.3723	

**PARAMETRO ALFA 1**

**RESULTADOS DE LA ECUACION REGIONAL**

$\lambda = 0.009999329$   
 $\gamma = 1.192008972$   
 Media = 4.6187

**ESTACION**

	Observation	Predicted Y	Residuals	Real	Y'	Y'+Ymed	X real
SM ZAPOTITLAN	1	0.4894	-2.3899	-1.9005	0.5491	5.1678	154.2690
SAN BLAS	2	0.8267	1.8833	2.7100	0.8525	5.4712	205.7698
LA TINA	3	-0.9592	0.2358	-0.7234	-0.9657	3.6530	36.1585
BAMICORI	4	-0.7909	0.0083	-0.7826	-0.8213	3.7974	41.5560
LAS CAÑAS	5	0.8349	0.9520	1.7869	0.8595	5.4782	207.1460
EL MAHONE	6	0.8820	-2.3593	-1.4773	0.9000	5.5187	215.2522
CHOIX	7	0.3288	0.0096	0.3384	0.3934	5.0121	133.0204
HUITES	8	1.2310	1.1667	2.3977	1.1904	5.8091	283.3516
PALO DULCE	9	-0.3645	1.7199	1.3554	-0.4289	4.1898	60.6087
CHINIPAS	10	0.6283	-1.7146	-1.3602	0.6772	5.2959	174.2340
SAN FRANCISO	11	0.0257	1.2282	1.2539	0.0463	4.6650	95.5354
SAN IGNACIO	12	-1.0599	-1.9426	-3.0025	-1.0500	3.5687	33.3312
LA VERANERA	13	-1.0268	-0.5344	-1.5612	-1.0224	3.5963	34.2303
URIQUE	14	-0.7945	1.2351	0.4406	-0.8245	3.7942	41.4311
GUERACHIC	15	-0.7107	1.0066	0.2958	-0.7509	3.8678	44.4721
ALAMOS	16	0.0324	-0.2812	-0.2488	0.0563	4.6750	96.4539
CAZANATE	17	0.6995	-0.1922	0.5073	0.7409	5.3596	185.1094
JAINA	18	-0.3093	1.4200	1.1107	-0.3737	4.2450	63.9052
TOAHAYANA	19	0.1098	-1.2618	-1.1520	0.1567	4.7754	106.1558
TECUSIAPA	20	0.9274	0.2685	1.1959	0.9387	5.5574	223.3008
LOS MOLINOS	21	-0.2308	-0.6366	-0.8675	-0.2923	4.3264	69.0899
NARANJO	22	0.3199	0.3861	0.7059	0.3844	5.0031	131.8863
ZOPILOTE	23	-0.8946	-0.2075	-1.1022	-0.9108	3.7079	38.1219

## ECUACION REGIONAL PARA EL PARAMETRO $\mu_1$

$$\mu_1 = (0.02999933\mu_1' + 1)^{1/0.02999933} \quad (6.2)$$

donde

$$\mu_1' = (-0.08601 + A' + PR' + L' + PD' + F' + C')^{1/1.136} + 6.0686$$

$$A' = 0.2178 \left[ \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right]^{1.2590}$$

A      área de la cuenca, en  $\text{km}^2$

$$PR' = 0.4006 \left[ \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right]^{-0.8940}$$

PR      precipitación media, en mm

$$L' = -0.3427 \left[ \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right]^{0.8980}$$

L      longitud del cauce principal, en km

$$PD' = 1.8553 \left[ \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right]^{-1.1020}$$

PD      pendiente del cauce principal, en  $\text{m}/\text{km}^{-1}$

$$F' = 0.7729 \left[ |\text{Ln}(F) + 5.4139| \right]^{-0.5024}$$

F      frecuencia del cauce principal, en  $\text{km}^{-2}$

$$C' = 91.6066 \left[ |\text{Ln}(C) - 4.6804| \right]^{1.0690}$$

C      coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

## ECUACION REGIONAL PARA EL PARAMETRO $\alpha_1$

$$\alpha_1 = (0.00999933\alpha'_1 + 1)^{1/0.00999933} \quad (6.3)$$

donde

$$\alpha'_1 = (0.0292 + A' + PR' + L' + PD' + F' + C')^{1/1.192} + 4.6187$$

$$A' = 0.3783 \left[ \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right]^{1.2590}$$

A      área de la cuenca, en km<sup>2</sup>

$$PR' = 0.4689 \left[ \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right]^{0.8940}$$

PR     precipitación media, en mm

$$L' = -0.7870 \left[ \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right]^{0.8980}$$

L      longitud del cauce principal, en km

$$PD' = -4.3696 \left[ \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right]^{1.1020}$$

PD     pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = -0.8119 \left[ |\ln(F) + 5.4139| \right]^{0.5024}$$

F      frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = 59.6527 \left[ |\ln(C) - 4.6804| \right]^{1.0690}$$

C      coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

## ECUACION REGIONAL PARA EL PARAMETRO $\mu_2$

$$\mu_1 = (-0.110000655 \mu_1' + 1)^{1/-0.110000655} \quad (6.4)$$

donde

$$\mu_1' = (-0.017299 + A' + PR' + L' + PD' + F' + C')^{1/1.217} + 4.8611$$

$$A' = 0.018498 \left[ \left( \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right) \right]^{1.2590}$$

A área de la cuenca, en km<sup>2</sup>

$$PR' = 0.215843 \left[ \left( \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right) \right]^{0.8940}$$

PR precipitación media, en mm

$$L' = 0.006871 \left[ \left( \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right) \right]^{0.8980}$$

L longitud del cauce principal, en km

$$PD' = -1.25291 \left[ \left( \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right) \right]^{1.1020}$$

PD pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = 0.154483 \left[ \left| \ln(F) + 5.4139 \right| \right]^{0.5024}$$

F frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = 21.73113 \left[ \left| \ln(C) - 4.6804 \right| \right]^{1.0690}$$

C coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

## ECUACION REGIONAL PARA EL PARAMETRO $\alpha_2$

$$\alpha_2 = (0.26999936 \alpha_2' + 1)^{1/0.26999936} \quad (6.5)$$

donde

$$\alpha_2' = (-0.01057 + A' + PR' + L' + PD' + F' + C')^{1/0.92699} + 14.8775$$

$$A' = 0.73717 \left[ \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right]^{1.2590}$$

A      área de la cuenca, en km<sup>2</sup>

$$PR' = 1.84302 \left[ \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right]^{-0.8940}$$

PR    precipitación media, en mm

$$L' = -0.95154 \left[ \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right]^{0.8980}$$

L      longitud del cauce principal, en km

$$PD' = -19.76871 \left[ \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right]^{1.1020}$$

PD    pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = -0.32968 \left[ |\ln(F) + 5.4139| \right]^{0.5024}$$

F      frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = -150.01223 \left[ |\ln(C) - 4.6804| \right]^{1.0690}$$

C      coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

## ECUACION REGIONAL PARA EL PARAMETRO P

$$p = \exp(p') \quad (6.6)$$

donde

$$p' = (-0.051506696 + A' + PR' + L' + PD' + F' + C')^{1/1.533025} - 0.6071$$

$$A' = 0.09372 \left[ \left( \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right) \right]^{1.2590}$$

A área de la cuenca, en km<sup>2</sup>

$$PR' = 0.14627 \left[ \left( \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right) \right]^{0.8940}$$

PR precipitación media, en mm

$$L' = -0.25022 \left[ \left( \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right) \right]^{0.8980}$$

L longitud del cauce principal, en km

$$PD' = -3.64459 \left[ \left( \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right) \right]^{1.1020}$$

PD pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = -0.25988 \left[ |\ln(F) + 5.4139| \right]^{0.5024}$$

F frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = -14.12994 \left[ |\ln(C) - 4.6804| \right]^{1.0690}$$

C coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

Los valores de los parámetros obtenidos de estas ecuaciones regionales, para la región homogénea A se muestran en la tabla 6.30.

### **6.7.2 Ecuaciones regionales con base en parámetros estimados en forma bivariada**

De la misma forma se desarrollaron las ecuaciones (6.7 a 6.11) que se originaron al correlacionar los parámetros obtenidos a través de distribuciones bivariadas. (que fue la que arrojó el mejor ajuste para el caso de cuencas aforadas). Los parámetros que se obtuvieron al aplicar estas ecuaciones se presentan en la tabla 6.31.

Si se observan con cuidado las Tablas 6.30 y 6.31 y se compara el error medio entre ambas, se concluye que al disminuirlo de 980 para el caso de las ecuaciones obtenidas de los parámetros originales a 975 para el ajuste de los parámetros con distribuciones bivariadas, se está logrando una mejor y sobre todo, más confiable estimación de parámetros para cuencas no aforadas. Este razonamiento es lógico, ya que las ecuaciones (6.7 a 6.11) se obtuvieron con base en una mejor estimación de parámetros.

### **6.7.3 Ecuaciones regionales con base en registros extendidos.**

Finalmente se obtiene un tercer juego de ecuaciones regionales. Aprovechando los registros extendidos, se repite el procedimiento antes descrito para obtener las ecuaciones que relacionen las características fisiográficas de la región, con los cinco parámetros de la distribución Gumbel Mixta así, estas ecuaciones (6.12 a 6.16) podrán ser empleadas en estaciones en donde se haya realizado una extensión del registro de sus escurrimientos máximos anuales.

El error medio que se obtiene con estas ecuaciones es de 981, sin embargo no está muy lejos de la estimación con las ecuaciones que se generaron inicialmente, y si se considera que algunas de las muestras fueron aumentadas en 5 años en promedio, y que además no se alteró el contenido de información de éstas; las ecuaciones (6.12 a 6.16) constituyen una tercera opción que en algún momento puede arrojar valores muy confiables al estimar eventos de diseño; recordando que se trata de disminuir la incertidumbre en cuencas escasamente aforadas o con registros incompletos.

Por ejemplo para la estación La Veranera (figura 28), en donde su registro original es de solo 14 años, el mejor ajuste considerándola como una cuenca bien aforada, se logró con la distribución

Gumbel Mixta estimando sus parámetros con una distribución bivariada, arrojando un error estándar de ajuste de 155.2. Mientras que, si se considera a ésta como una estación pobremente aforada y se obtienen los parámetros de su distribución empleando los tres juegos de ecuaciones desarrolladas, se logran resultados muy confiables, pues los errores que se obtienen para el caso de las ecuaciones regionales relacionando parámetros originales, en forma bivariada y parámetros obtenidos de muestras extendidas son: 196.1, 212.6 y 228.5, respectivamente. Este resultado se observa más claramente si se comparan los eventos de diseño para los periodos de retorno seleccionados en esta misma estación.



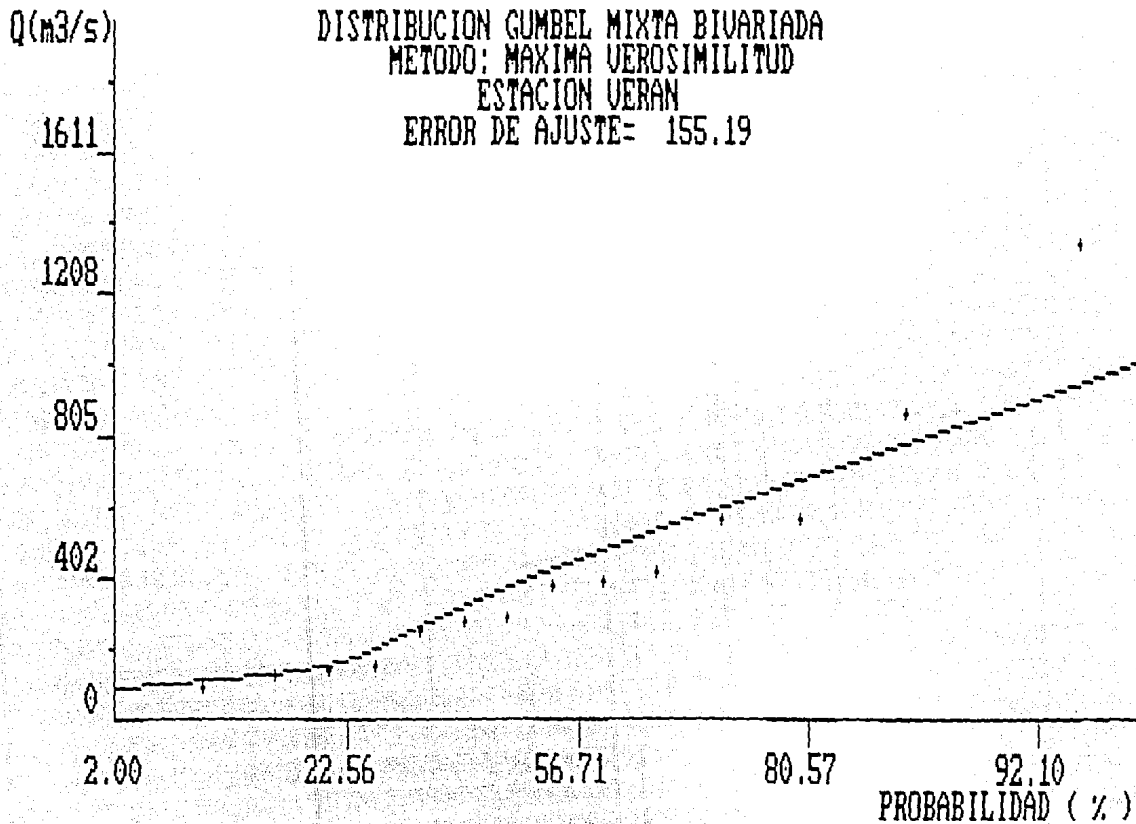


FIGURA 28: DISTRIBUCION GUMBEL MIXTA BIVARIADA, ESTACION LA VERANERA

**ESTIMANDO PARAMETROS EN FORMA BIVARIADA  
ECUACION REGIONAL PARA EL PARAMETRO  $\mu_1$**

$$\mu_1 = (0.03999933 \mu_1' + 1)^{1/0.03999933} \quad (6.7)$$

donde

$$\mu_1' = (-0.09518 + A' + PR' + L' + PD' + F' + C')^{1/1.152} + 6.2611$$

$$A' = 0.2481 \left[ \left( \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right) \right]^{1.2590}$$

A área de la cuenca, en km<sup>2</sup>

$$PR' = 0.4002 \left[ \left( \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right) \right]^{0.8940}$$

PR precipitación media, en mm

$$L' = -0.40583 \left[ \left( \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right) \right]^{0.8980}$$

L longitud del cauce principal, en km

$$PD' = 1.93863 \left[ \left( \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right) \right]^{1.1020}$$

PD pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = 0.8402 \left[ \left| \ln(F) + 5.4139 \right| \right]^{0.5024}$$

F frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = 95.72389 \left[ \left| \ln(C) - 4.6804 \right| \right]^{1.0690}$$

C coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

**ESTIMANDO PARAMETROS EN FORMA BIVARIADA  
ECUACION REGIONAL PARA EL PARAMETRO  $\alpha_1$**

$$\alpha_1 = (0.01999933\alpha'_1 + 1)^{1/0.01999933} \quad (6.8)$$

donde

$$\alpha'_1 = (0.0295 + A' + PR' + L' + PD' + F' + C')^{1/1.184} + 4.7329$$

$$A' = 0.3942 \left[ \left| \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right| \right]^{1.2590}$$

A      área de la cuenca, en km<sup>2</sup>

$$PR' = 0.47092 \left[ \left| \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right| \right]^{0.8940}$$

PR      precipitación media, en mm

$$L' = -0.81782 \left[ \left| \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right| \right]^{0.8980}$$

L      longitud del cauce principal, en km

$$PD' = -4.3144 \left[ \left| \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right| \right]^{1.1020}$$

PD      pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = -0.83193 \left[ \left| \ln(F) + 5.4139 \right| \right]^{0.5024}$$

F      frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = 64.14283 \left[ \left| \ln(C) - 4.6804 \right| \right]^{1.0690}$$

C      coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

**ESTIMANDO PARAMETROS EN FORMA BIVARIADA  
ECUACION REGIONAL PARA EL PARAMETRO  $\mu_2$**

$$\mu_1 = (-0.110000655\mu_1' + 1)^{1/-0.110000655} \quad (6.9)$$

donde

$$\mu_1' = (-0.017868 + A' + PR' + L' + PD' + F' + C')^{1/1.29101} + 4.8774$$

$$A' = 0.02343 \left[ \left( \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right) \right]^{1.2590}$$

A área de la cuenca, en km<sup>2</sup>

$$PR' = 0.22193 \left[ \left( \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right) \right]^{0.8940}$$

PR precipitación media, en mm

$$L' = -0.005859 \left[ \left( \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right) \right]^{0.8980}$$

L longitud del cauce principal, en km

$$PD' = -1.09009 \left[ \left( \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right) \right]^{1.1020}$$

PD pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = 0.16917 \left[ |\ln(F) + 5.4139| \right]^{0.5024}$$

F frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = 26.14231 \left[ |\ln(C) - 4.6804| \right]^{1.0690}$$

C coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

**ESTIMANDO PARAMETROS EN FORMA BIVARIADA  
ECUACION REGIONAL PARA EL PARAMETRO  $\alpha_2$**

$$\alpha_2 = (0.199999347 \alpha_2' + 1)^{1/0.199999347} \quad (6.10)$$

donde

$$\alpha_2' = (-0.012568 + A' + PR' + L' + PD' + F' + C')^{1/0.929991543} + 11.1307$$

$$A' = 0.64368 \left[ \left( \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right) \right]^{1.2590}$$

A      área de la cuenca, en km<sup>2</sup>

$$PR' = 0.77412 \left[ \left( \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right) \right]^{0.8940}$$

PR    precipitación media, en mm

$$L' = -1.04716 \left[ \left( \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right) \right]^{0.8980}$$

L      longitud del cauce principal, en km

$$PD' = -21.59728 \left[ \left( \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right) \right]^{1.1020}$$

PD    pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = -0.30239 \left[ \left| \ln(F) + 5.4139 \right| \right]^{0.5024}$$

F      frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = -258.75409 \left[ \left| \ln(C) - 4.6804 \right| \right]^{1.0690}$$

C      coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

**ESTIMANDO PARAMETROS EN FORMA BIVARIADA  
ECUACION REGIONAL PARA EL PARAMETRO P**

$$p = \exp(p') \quad (6.11)$$

donde

$$p' = (-0.0570463 + A' + PR' + L' + PD' + F' + C')^{1/1.562026262} - 0.6116$$

$$A' = 0.09034 \left[ \left[ \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right] \right]^{1.2590}$$

A área de la cuenca, en km<sup>2</sup>

$$PR' = 0.17593 \left[ \left[ \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right] \right]^{0.8940}$$

PR precipitación media, en mm

$$L' = -0.24346 \left[ \left[ \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right] \right]^{0.8980}$$

L longitud del cauce principal, en km

$$PD' = -4.07609 \left[ \left[ \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right] \right]^{1.1020}$$

PD pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = -0.27623 \left[ \left| \ln(F) + 5.4139 \right| \right]^{0.5024}$$

F frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = -19.06499 \left[ \left| \ln(C) - 4.6804 \right| \right]^{1.0690}$$

C coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

**TABLA 6.30**  
**PARAMETROS DE LA DISTRIBUCION GUMBEL MIXTA**  
**AJUSTADOS CON LA ECUACION REGIONAL**

	MU 1	ALFA 1	MU 2	ALFA 2	P	ERROR
<b>CUENCA DEL RIO FUERTE</b>						
SM ZAPOTITLAN	496.506	154.269	2,654.973	1,088.413	0.432	1,953.149
SAN BLAS	560.739	205.770	2,689.573	1,164.285	0.536	4,699.861
LA TINA	56.085	36.158	354.915	39.951	0.553	188.469
BAMICORI	103.484	41.556	550.485	73.655	0.655	135.166
LAS CAÑAS	649.329	207.146	2,512.003	1,092.101	0.505	1,766.528
EL MAHONE	661.678	215.252	2,501.983	1,098.811	0.523	2,393.530
CHOX	267.077	133.020	2,143.603	478.695	0.921	209.614
HUITES	753.430	283.352	2,134.009	1,088.459	0.692	2,704.398
PALO DULCE	276.877	60.609	1,573.941	349.555	0.331	1,069.723
CHINIPAS	570.969	174.234	2,230.664	303.938	0.511	676.263
SAN FRANCISCO	520.854	95.535	1,666.425	587.169	0.370	844.984
SAN IGNACIO	315.360	33.331	1,080.581	426.902	0.306	827.723
LA VERANERA	215.655	34.230	520.593	210.652	0.295	196.115
URIQUE	238.249	41.431	517.896	231.260	0.319	291.700
GUERACHIC	223.930	44.472	661.391	412.637	0.384	1,369.424
ALAMOS	107.017	96.454	686.326	207.341	0.668	207.456
CAZANATE	120.899	185.109	693.518	250.901	0.901	672.179
<b>CUENCA DEL RIO SINALOA</b>						
JAINA	429.477	63.905	1,970.827	558.192	0.385	854.655
TOAHAYANA	507.344	106.156	1,750.390	584.756	0.524	365.736
TECUSIAPA	587.052	223.301	1,346.569	636.557	0.800	434.576
LOS MOLINOS	82.090	69.090	320.279	114.214	0.502	149.455
NARANJO	138.148	131.886	666.338	248.715	0.628	410.856
ZOPILOTE	66.517	38.122	424.811	83.728	0.414	122.160

MEDIA = 980.162

**TABLA 6.31**  
**PARAMETROS DE LA DISTRIBUCION GUMBEL MIXTA**  
**OBTENIDOS CON LA ECUACION REGIONAL**  
**DE PARAMETROS BIVARIADOS**

	MU 1	ALFA 1	MU 2	ALFA 2	P	ERROR
<b>CUENCA DEL RIO FUERTE</b>						
SM ZAPOTITLAN	495.911	154.274	2769.485	1082.139	0.433	2039.823
SAN BLAS	564.656	204.741	2850.444	1198.329	0.563	4650.429
LA TINA	59.881	33.701	360.81	43.717	0.563	186.9206
BAMICORI	104.386	41.673	559.223	71.201	0.653	141.1456
LAS CAÑAS	657.311	206.593	2728.894	1088.397	0.498	1719.297
EL MAHONE	670.574	214.58	2725.62	1099.021	0.514	2583.321
CHOIX	263.674	128.531	2356.819	458.567	0.968	159.8556
HUITES	777.368	281.809	2385.777	1179.106	0.679	2517.004
PALO DULCE	257.83	61.504	1720.784	193.337	0.326	1188.959
CHINIPAS	548.171	137.81	1683.302	235.523	0.369	407.4519
SAN FRANCISCO	532.683	98.225	1828.831	542.817	0.356	839.8867
SAN IGNACIO	319.703	34.104	1040.155	400.258	0.3	869.5135
LA VERANERA	221.091	35.409	519.051	184.59	0.28	212.5923
URIQUE	248.666	42.881	520.032	217.876	0.302	291.5363
GUERACHIC	230.877	45.228	636.946	454.215	0.378	1344.677
ALAMOS	107.209	96.7	662.887	208.84	0.658	200.9126
CAZANATE	123.315	182.837	684.065	279.731	0.898	670.387

**CUENCA DEL RIO SINALOA**

JAINA	422.684	65.125	2168.825	433.719	0.476	871.8243
TOAHAYANA	516.133	105.448	1970.866	512.378	0.558	434.2823
TECUIAPA	611.299	219.927	1571.563	606.646	0.802	366.8304
LOS MOLINOS	82.184	69.275	321.397	83.431	0.486	137.7182
NARANJO	136.938	131.195	664.211	189.742	0.618	448.2481
ZOPILOTE	65.625	38.482	410.82	69.481	0.405	140.428

MEDIA = 974.915



**TABLA 6.32**  
**PARAMETROS DE LA DISTRIBUCION GUMBEL MIXTA**  
**OBTENIDOS CON LA ECUACION REGIONAL**  
**DE MUESTRAS EXTENDIDAS**

	MU 1	ALFA 1	MU 2	ALFA 2	P	ERROR
<b>CUENCA DEL RIO FUERTE</b>						
SM ZAPOTITLAN	677.609	325.205	3.267.739	1.312.961	0.664	1.757.162
SAN BLAS	710.766	365.705	3.130.369	1.369.420	0.687	4.762.519
LA TINA	49.081	33.276	360.631	36.548	0.619	170.410
BAMICORJ	102.964	40.937	558.704	57.575	0.648	142.565
LAS CAÑAS	773.697	346.240	2.847.110	1.198.958	0.655	1.712.719
EL MAHONE	779.633	351.195	2.814.255	1.199.899	0.655	2.313.230
CHOIX	307.224	176.858	1.832.773	465.247	1.553	299.961
HUITES	798.807	378.647	2.212.796	1.088.323	0.661	2.587.853
PALO DULCE	348.427	110.909	1.878.404	372.695	0.609	948.734
CHINIPAS	490.055	138.766	1.451.079	328.530	0.598	786.745
SAN FRANCISCO	603.959	186.394	1.885.958	559.154	0.617	918.247
SAN IGNACIO	463.769	108.467	1.508.512	400.794	0.617	867.608
LA VERANERA	296.419	68.170	740.343	173.105	0.558	228.499
URIQUE	307.156	78.389	706.367	183.544	0.582	306.309
GUERACHIC	346.766	126.225	960.555	370.568	0.628	1.366.090
ALAMOS	123.745	103.575	787.872	218.397	0.635	305.018
CAZANATE	123.228	144.396	710.187	247.929	0.712	577.724
<b>CUENCA DEL RIO SINALOA</b>						
JAINA	539.215	160.078	2.271.720	555.929	0.633	912.683
TOAHAYANA	561.916	186.492	1.819.290	525.488	0.672	213.479
TECUSIAPA	549.780	250.677	1.298.668	525.343	0.687	362.914
LOS MOLINOS	86.156	73.382	371.103	101.315	0.610	450.767
NARANJO	159.162	124.286	811.363	254.631	0.617	441.938
ZOPILOTE	73.839	47.855	517.603	87.774	0.617	125.257

MEDIA = 980.801

La tabla 6.32 muestra los parámetros obtenidos de las ecuaciones regionales (6.12 a 6.16) relacionando la fisiografía con los parámetros de las muestras extendidas.

Para redondear el ejemplo de la estación La Veranera a continuación se presentan los eventos de diseño para los periodos de retorno seleccionados, ajustados con los parámetros obtenidos en forma bivariada (mejor ajuste) y con los tres grupos de parámetros que se originan de las ecuaciones regionales:

**1er. grupo.** Parámetros obtenidos al relacionar la fisiografía con los parámetros originales univariados.

**2o. grupo.** Parámetros obtenidos al relacionar la fisiografía con los parámetros obtenidos en forma bivariada.

**3er. grupo.** Parámetros obtenidos al relacionar la fisiografía con los parámetros de las muestras extendidas.

GASTOS DE DISEÑO EN ( $m^3/s$ ) PARA LA ESTACION LA VERANERA

Periodo de retorno en años	En forma bivariada	ECUACIONES REGIONALES		
		Del 1er. grupo de parámetros	Del 2o. grupo de parámetros	Del 3er. grupo de parámetros
2	395	476	488	445
5	673	752	726	828
10	847	916	870	976
20	1013	1070	1005	1107
50	1225	1268	1178	1272
100	1384	1416	1307	1394
500	1750	1756	1606	1675
1000	1908	1902	1734	1795
5000	2272	2247	2027	2076
10000	2442	2394	2150	2198

Finalmente cabe mencionar que para el caso de las cuencas no aforadas las ecuaciones regionales demostraron ser un excelente recurso en la estimación de los parámetros, tal es el caso de la estación La Veranera, en donde el gasto para 10 000 años estimado con base en el primer grupo de ecuaciones regionales apenas difiere en  $48 m^3/s$  del que se considera como mejor ajuste de los escurrimientos máximos registrados

**ESTIMANDO PARAMETROS DE REGISTROS EXTENDIDOS**  
**ECUACION REGIONAL PARA EL PARAMETRO  $\mu_1$**

$$\mu_1 = (0.149999321 \mu_1' + 1)^{1/0.149999321} \quad (6.12)$$

donde

$$\mu_1' = (-0.1362339 + A' + PR' + L' + PD' + F' + C')^{1/0.97399098} + 9.189$$

$$A' = 0.26052 \left[ \left( \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right) \right]^{1.2590}$$

A área de la cuenca, en km<sup>2</sup>

$$PR' = 0.665567 \left[ \left( \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right) \right]^{0.8940}$$

PR precipitación media, en mm

$$L' = -0.252258 \left[ \left( \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right) \right]^{0.8980}$$

L longitud del cauce principal, en km

$$PD' = 2.043744 \left[ \left( \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right) \right]^{1.1020}$$

PD pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = 1.418273 \left[ |\ln(F) + 5.4139| \right]^{0.5024}$$

F frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = 79.442819 \left[ |\ln(C) - 4.6804| \right]^{1.0690}$$

C coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

**ESTIMANDO PARAMETROS DE REGISTROS EXTENDIDOS  
ECUACION REGIONAL PARA EL PARAMETRO  $\alpha_1$**

$$\alpha_1 = (0.23999937 \alpha_1' + 1)^{1/0.23999937} \quad (6.13)$$

donde

$$\alpha_1' = (0.0536678 + A' + PR' + L' + PD' + F' + C')^{1/1.032} + 9.5965$$

$$A' = 0.747687 \left[ \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right]^{1.2590}$$

A área de la cuenca, en km<sup>2</sup>

$$PR' = 1.2344 \left[ \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right]^{0.8940}$$

PR precipitación media, en mm

$$L' = -1.28211 \left[ \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right]^{0.8980}$$

L longitud del cauce principal, en km

$$PD' = -15.45046 \left[ \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right]^{1.1020}$$

PD pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = -1.42675 \left[ |\ln(F) + 5.4139| \right]^{0.5024}$$

F frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = -66.2401 \left[ |\ln(C) - 4.6804| \right]^{1.0690}$$

C coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

**ESTIMANDO PARAMETROS DE REGISTROS EXTENDIDOS  
ECUACION REGIONAL PARA EL PARAMETRO  $\mu_2$**

$$\mu_1 = (-0.04000067 \mu_1' + 1)^{1/-0.04000067} \quad (6.14)$$

donde

$$\mu_1' = (-0.018895 + A' + PR' + L' + PD' + F' + C')^{1/1.007} + 6.1867$$

$$A' = 0.003657 \left[ \left[ \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right] \right]^{1.2590}$$

A área de la cuenca, en km<sup>2</sup>

$$PR' = 0.36107 \left[ \left[ \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right] \right]^{0.8940}$$

PR precipitación media, en mm

$$L' = 0.103782 \left[ \left[ \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right] \right]^{0.8980}$$

L longitud del cauce principal, en km

$$PD' = -1.56956 \left[ \left[ \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right] \right]^{1.1020}$$

PD pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = 0.264957 \left[ \left| \ln(F) + 5.4139 \right| \right]^{0.5024}$$

F frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = 27.250884 \left[ \left| \ln(C) - 4.6804 \right| \right]^{1.0690}$$

C coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

**ESTIMANDO PARAMETROS DE REGISTROS EXTENDIDOS**  
**ECUACION REGIONAL PARA EL PARAMETRO  $\alpha_2$**

$$\alpha_2 = (0.20999935 \alpha_2' + 1)^{1/0.20999935} \quad (6.15)$$

donde

$$\alpha_2' = (0.04833 + A' + PR' + L' + PD' + F' + C')^{1/0.91099} + 11.7523$$

$$A' = 0.45093 \left[ \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right]^{1.2590}$$

A      área de la cuenca, en km<sup>2</sup>

$$PR' = 1.50415 \left[ \left( \frac{PR^{0.2399} - 1}{0.2399} \right) - 16.6797 \right]^{0.8940}$$

PR      precipitación media, en mm

$$L' = -0.41639 \left[ \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right]^{0.8980}$$

L      longitud del cauce principal, en km

$$PD' = -16.2087 \left[ \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right]^{1.1020}$$

PD      pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = -0.65601 \left[ |\ln(F) + 5.4139| \right]^{0.5024}$$

F      frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = -114.70818 \left[ |\ln(C) - 4.6804| \right]^{1.0690}$$

C      coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

**ESTIMANDO PARAMETROS DE REGISTROS EXTENDIDOS  
ECUACION REGIONAL PARA EL PARAMETRO P**

$$p = \exp(p') \quad (6.16)$$

donde

$$p' = (0.169079 + A' + PR' + L' + PD' + F' + C')^{1/0.3979984} - 0.4826$$

$$A' = 0.0500829 \left[ \left[ \left( \frac{A^{0.1499} - 1}{0.1499} \right) - 17.3713 \right] \right]^{1.2590}$$

A área de la cuenca, en km<sup>2</sup>

$$PR' = 0.23795 \left[ \left[ \left( \frac{PR^{0.2399} - 1}{-0.2399} \right) - 16.6797 \right] \right]^{0.8940}$$

PR precipitación media, en mm

$$L' = -0.14822 \left[ \left[ \left( \frac{L^{0.3399} - 1}{0.3399} \right) - 13.0317 \right] \right]^{0.8980}$$

L longitud del cauce principal, en km

$$PD' = -6.6873 \left[ \left[ \left( \frac{PD^{0.4199} - 1}{0.4199} \right) + 2.0734 \right] \right]^{1.1020}$$

PD pendiente del cauce principal, en m/km<sup>-1</sup>

$$F' = -0.09668 \left[ \left| \ln(F) + 5.4139 \right| \right]^{0.5024}$$

F frecuencia del cauce principal, en km<sup>-2</sup>

$$C' = -58.09761 \left[ \left| \ln(C) - 4.6804 \right| \right]^{1.0690}$$

C coordenada de longitud del centroide de la cuenca, en grados

Es necesario señalar en las expresiones anteriores, que al elevar a la potencia se tome en cuenta el valor absoluto; sin embargo, el valor que resulte de esto conservará el signo de la operación.

## **7. CONCLUSIONES Y RECOMENDACIONES**

### **7.1 Conclusiones**

La incertidumbre al estimar la magnitud y el período de ocurrencia de eventos hidrológicos extremos, así como la variabilidad en espacio y tiempo de las diferentes componentes del ciclo hidrológico, han hecho necesario desarrollar procedimientos regionales para analizar y explicar el comportamiento de estas variables. Es por eso que como primer modelo de transferencia de información, el análisis de frecuencias de la información hidrológica en cuencas aforadas, relaciona la magnitud de los eventos extremos con su frecuencia de ocurrencia mediante el uso de distribuciones de probabilidad univariadas y bivariadas.

En este análisis debe garantizarse que se ha seleccionado la mejor distribución de probabilidad, además de estimar los parámetros para el caso de las distribuciones con una población, con el método de máxima verosimilitud en forma bivariada. Para el caso de las distribuciones de dos poblaciones (Gumbel Mixta) los parámetros se deben obtener también en forma bivariada, iniciando el ajuste con los parámetros que se obtienen de un algoritmo de optimación. Esta estimación bivariada de los parámetros garantiza una significativa disminución en el error estándar de ajuste.



El término modelo es hoy en día de uso frecuente para describir las ecuaciones de sistemas físicos, y de técnicas para hallar soluciones óptimas. Un modelo matemático es una ecuación o conjunto de ecuaciones (o relaciones matemáticas) que describen en forma aproximada un proceso físico.

Las ecuaciones regionales que han sido desarrolladas en este trabajo para sitios no aforados, a partir de estimar parámetros en forma bivariada representan un procedimiento para el modelado de cuencas. Las características fisiográficas y el análisis de frecuencias de sus correspondientes gastos máximos instantáneos, representan la base fundamental para facilitar la primera etapa de la planeación, análisis y diseño de los fenómenos hidrológicos relacionados con proyectos de recursos hidráulicos.

Gran parte de la efectividad de estas ecuaciones nace del conocimiento de las funciones de distribución de probabilidad de valores extremos, específicamente de las distribuciones bivariadas. Estas distribuciones arrojan una ganancia en la precisión al obtener los parámetros de la distribución de probabilidad. De acuerdo a los resultados obtenidos en esta estimación de parámetros se permite garantizar al ingeniero hidrólogo una disminución en la incertidumbre de la frecuencia y niveles de las crecientes, a fin de predecir y controlar avenidas máximas que permitirán delimitar el plano de inundación, y que a su vez ayudarán para dimensionar los embalses utilizados en el control de avenidas.

Todo este desarrollo plantea tres consideraciones necesarias para garantizar su efectividad: Primero una correcta estimación y discriminación de las características fisiográficas de la región constituye una base sólida para iniciar el proceso de delimitación de regiones hidrológicamente homogéneas. Segundo un método de regionalización que parte de este estudio de homogeneidad y de las características seleccionadas, es el origen del análisis de correlación y regresión que dará la forma final a las ecuaciones regionales. Tercero una confiable estimación de parámetros a través de una relación bivariada, permitirá disminuir, en forma significativa la incertidumbre en el cálculo de los eventos de diseño en las cuencas aforadas. Ya que estos parámetros parten del mejor ajuste, las ecuaciones regionales que resulten serán ampliamente recomendadas en cuencas con registros cortos o escasamente aforadas.

En lo que se refiere a la primera consideración cabe mencionar que debe existir la mayor cantidad posible de variables fisiográficas en competencia y que la discriminación de éstas debe realizarse con pruebas estadísticas como el análisis de varianza, y que a su vez el estadístico  $F$ , debe probarse para garantizar que las características que son removidas no eran significativas en el modelo. La delimitación de cuencas homogéneas debe realizarse en un principio, con base en un método gráfico

en donde pueda reflejarse la variabilidad de la fisiografía, por ejemplo, las curvas de Andrews. Esta primera agrupación debe servir como el inicio para llevar a cabo un análisis de varianza en dos direcciones, de los coeficientes de variación de los gastos máximos anuales. Este desarrollo es el que dará la ubicación final de las estaciones.

Para el caso de las cuencas no aforadas se debe en un principio, extender registros a través de una correlación lineal simple y posteriormente construir una ecuación regional con base en una correlación múltiple, que relacione los parámetros de su distribución con sus correspondientes características fisiográficas. Una vez que se conoce la mejor combinación de los parámetros, el gasto de diseño debe calcularse para los periodos de retorno seleccionados. Finalmente, aplicando algún procedimiento para el cálculo de hidrogramas se le dará la forma al escurrimiento máximo.

Esta serie de modelos de transferencia de información hidrológica han sido desarrollados como una alternativa de regionalización, además estos modelos incorporan técnicas numéricas de manera que puedan ser adaptados a soluciones por computadora, dando eficiencia y seguridad en la estimación de los eventos de diseño.

## **7.2 Recomendaciones**

Es recomendable, si se escoge la regionalización como un recurso para estimar eventos de diseño, que previamente se realice una discriminación de variables fisiográficas, a fin de modelar correctamente el proceso lluvia-escurrimiento. Así mismo, que se desarrolle detalladamente un análisis de homogeneidad de cuencas basado en la metodología propuesta y que además se lleve a cabo un análisis de frecuencias con base en la estimación bivariada de parámetros. Finalmente las ecuaciones regionales deben ser obtenidas de algún modelo (para este caso se sugiere lineal) de correlación múltiple.

En lo que se refiere a las pruebas de homogeneidad, la prueba de Lagbein solo se recomienda como un método para revisar la homogeneidad de los grupos, ya que se ha demostrado que no es una prueba lo suficientemente confiable para delimitar regiones.

Otra recomendación importante se relaciona con el hecho de que al efectuar un análisis de correlación y regresión, debe garantizarse la total normalidad de las muestras, pues los errores que se cometen por pasar inadvertida esta consideración, son significativamente importantes al grado de alterar los resultados inferidos. Por tal razón se recomienda llevar a cabo una normalización por la

transformación de Box-Cox de doble potencia, para garantizar un coeficiente de asimetría igual a cero y una curtosis igual a tres.

Finalmente cabe aclarar, que este trabajo es solo parte de la modelación de cuencas, ya que no sólo son los escurrimientos los que influyen en un proceso lluvia-escurrimiento. Deben estudiarse de la misma manera los procesos que abarcan el total de los componentes del ciclo hidrológico; estudiando cada una de sus características y consideraciones fundamentales, tal y como se efectuó en este análisis de escurrimientos. Así como se modelaron escurrimientos, se puede mediante estas técnicas analizar gastos mínimos, análisis de lluvia, etc,... Dichos procesos deben abarcar modelos con patrones para modelar la circulación en la atmósfera, la densidad y variabilidad espacial de la lluvia en la zona de captación, la forma y las dimensiones del hidrograma de diseño, transitar la avenida por el cauce y terminar con un funcionamiento de vaso.

## **BIBLIOGRAFIA**

- ACREMAN, C. AND SINCLAIR C., "Classification of drainage basins according to their physical characteristics; an application for flood frequency analysis in Scotland", Journal of Hydrology. Number 84, 1986, 365-380 pp.
- ACREMAN, C., "Predicting the mean annual flood from basin characteristics in Scotland", Hydrological Sciences Journal: 1985. Number 30 (1), 37-49 pp.
- ALVARADO C. J., "Cálculo de avenidas de diseño para vertedores de presas de almacenamiento. Aplicación a la presa Peñitas", Tesis de Maestría, División de Estudios de Posgrado, Facultad de Ingeniería, UNAM., 1993: 179 pp.
- ANDERSON, W., "Relating sediment yield to watershed variables", Trans. Am. Geophys Union, 1957, Number 38, 921-924 pp.
- ANDREWS, C., "Plots of high dimensional data. Biometrics", Journal of Hydrology, Number 28, 1972, 125-136 pp.

- BEABLE, E., AND McHERCHAR, I., "Regional flood estimation in New Zealand",  
Natl. Water and Soil Conserv. Org., Water and Soil Taech, Publ. 1982, Number 20.
- BEDIENT, P., HUBER, W., AND ROSS, P., Hidrology and flood plain analysis.  
U.S.A. Addison-Wesley Publishing Company, 1988. 142 p.
- BELL, F. C., AND OM KAR, S., "Characteristics response times in design flood estimation",  
Journal of Hidrology, Number 8, 173-196 pp.
- BENSON, M. A., "Characteristics of frequency curves based on a theoretical 1,000 year record",  
U.S. Geol. Surv. Water Suppl. Pap., 1960, 1543-A.
- BENSON, M. A., "Factors influencing the ocurrence of floods in a humid region of divers terrain",  
U.S. Geol. Surv. Water Suppl. Pap., 1962, 1580-B.
- BEVEN J, KEITH, F WOOD, ERIC AND SIVAPLAN, MURUGESU. On hidrologycal  
heterogeneity-catchment morphology and catchment response" in Journal of  
Hidrology. Number 100, 1988. 353-375 pp.
- BISWAS A, K. AND FLEMING, G. "Floods in Scotland: Magnitude and frequency" in Water  
Water Eng. 1966 246-252 pp.
- BOULTON A, G. " Morphometric analysis of river basin characteristics". in U.K. Water Resour.  
Board, TN2, 10. 1965.
- BOX G, E.P. AND COX D,R. J.R. Stat. Soc., Series 1964.B, 26, 211-46 pp.
- COLE, G. "An application of the regional analysis of flood flows." in River Flood Hidrology. Inst.  
Civ. Eng., London, 1966. 39-57 pp.
- CAMPOS ARANDA, D.F. Procesos del ciclo hidrológico. México, Universidad  
Autónoma de San Luis Potosí , volúmen 20, tomo 1/2 , 1987.
- CAMPOS ARANDA, DANIEL..F. " Ajuste regionalizado de las distribuciones Gumbel y GVE en  
la región hidrológica No. 10, Sinaloa. I. Algoritmo similar al índice de crecientes. "en  
Memorias Técnicas, XIII Congreso Nacional de Hidráulica. Tomo II

- CAMPOS ARANDA, DANIEL..F. " Ajuste regionalizado de las distribuciones Gumbel y GVE en la región hidrológica No. 10, Sinaloa. 2. Método basado en los MPP". en Memorias Técnicas, XIII Congreso Nacional de Hidráulica. Tomo II
- DALRYMPLE, T. "Flood frequency analysis." in U.S. Geol. Surv., Water Suppl. Pap. 1960. 1543-A.
- ESCALANTE SANDOVAL, CARLOS A. " Análisis regional de eventos extremos" en Memorias técnicas, XIII Congreso Nacional de Hidráulica. Tomo I.
- ESCALANTE SANDOVAL, CARLOS A. Distribuciones trivariadas de valores extremos y sus aplicaciones en hidrología. Tesis Doctoral, División de Estudios de Posgrado, Facultad de Ingeniería, UNAM., 1991. 315 p.
- EVERITT, B.S. Graphical techniques for multivariate data. Great Britain, North-Holland, 1978. 117 p.
- FISHER R, A. AND TIPPETT L.H.C. "Limiting forms of the frequency distribution of the largest or smallest of a sample". in Proc. Cambr. Phil. Soc., 1928. 180-190 pp.
- GINGRAS, DENIS AND ADAMOWSKI, KAZ. "Homogeneous region delineation based on annual flood generation mechanisms", in Hydrological Sciences Journal. Vol. 38; number 2, 1993. 103-121 pp
- GUMBEL E, J. Statistics of Extremes. Columbia Univ Press, New York, NY, 1958a.
- GREENWOOD J, A., LANDWEHR J, M, MATALAS N , C AND WALLIS J, R. . "Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form." in Water Resour. Res. 1979. Number 15. 1049-1054 pp.
- H. BURN, DONALD. "Delineation of groups for regional flood frequency analysis". in Journal of Hidrology. Number 104, 1988. 345-361 pp.
- HAAN T, CHARLES. Statiscal methods in hidrology. U.S.A. , Iowa State University. 378 P.

- HAINES A, T., FINLAYSON, B, L. AND McMAHON, T, A. "A global classification of river regimes. Appl. Geog" in Journal of Hidrology.. Number 8, 1988. 255-272 pp.
- HISDAL, H AND TVEITO E. O. "Extension of runoff series using empirical orthogonal functions" in Hydrological Sciences Journal. Vol 38, number 2, 1993. 33-49 pp.
- HOLDER, L. R. Multiple regression in hidrology. Great Britain, Institute of Hidrology Wallingford, Oxfordshire, 1985. 147 p.
- HORTON, R, E., "Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology.  
Bull. Geol. Soc. Am., 1945, vol. 56, 275-370 pp.
- HOSKING J.R.M., WALLIS, J.R. AND WOOD E, F. "An appraisal of the regional flood frequency procedure " in The U.K. Flood Studies Rep., Hydrol. Sci. J., 1985a. Number 30 (1) 85-89 pp.
- HOSKING J.R.M., WALLIS, J.R. AND WOOD E, F. "Estimation of the generalised extreme value distribution by the method of probability weighted moments" in Technometrics. 1985b. Number 27. 251-261 pp.
- HUGHES J.M.R. "Hydrological characteristics and classification of Tasmanian rivers" in Journal of Hidrology. Aust. Geog, Studies. Number 25 (1). 61-82 pp.
- INSTITUTE OF HYDROLOGY. Low flows study report. Inst of Hidrology, Wallingford.
- JARQUE M, CARLOS. "A solution to the problem of optimum stratification in multivariate sampling" in Applied Statistics. Vol 30, number 2, 1981. 163-169 pp.
- JENKINSON A,F. "The frequency distribution of the annual maximum (or minimum) of meteorological elements".in Q.J.R. Meteorol. Soc., 1955. 158-171 pp.
- JOHNSON N,L. Biometrika. 1949. 149-76.
- KENDALL M.G. AND STUART, A. The advanced theory of statistics. Griffin, High Wycombe, Second Ed. 1961, vol 2.

- KENNEDY R, J AND WATT W, E. "The relationship between lag time and the physical characteristics of drainage basins" in Southern Ontario, I.A.S.H. Symp. Floods and their computations, Leningrad. 1969. Number 85. 866-874.
- KITE W, G. Frequency and risk analyses in hidrology. U.S.A. , Water Resources Publications, 1988. 257 p.
- KLEIN, M. " Effect of catchment size on runoff relationships a comment" in Journal of Hidrology. Number 71, 1984. 191-195 pp.
- LU HSIUNG, LI AND STEDINGER R, JERY. "Sampling variance of normalized GEV/PWM quantile estimators and a regional homogeneity test." in Journal of Hidrology. Number 138, 1992. 223-245 pp.
- McCUEN H, RICHARD. Microcumputer Applications in Statistical Hydrology. New Jersey, Prentice Hall.
- MARTINEZ, A ., H.R. AND RAYNAL V., J.A. Identificación de distribuciones de valores extremos con registros históricos. 10o Congreso Nacional de Hidráulica Morelia Michoacán, Tomo II, Subtema 6, 1978. 356-365 pp.
- MASON D, ROBERT Y LIND A, DOUGLAS. Estadística. U.S.A. , Alfaomega.
- MOSLEY P, M. "Delimination of New Zealand hidrologic regions" in Journal of Hydrology. Number 49, 1981. 173-192 pp.
- NASH J,E. AND SHAW B.L. "Flood frquency as a function of catchment characteristics" in River Flood Hydrology. Inst. Civ. Eng., London. 1966.
- NATHAN, J AND McMAHON A, T. "Identification of homogeneous regions for the purposes of regionalisation" in Journal of Hidrology. Number 121, 1990, 217-238 pp.
- NATURAL ENVIRONMENT RESEARCH COUNCIL. Flood Studies Report. NERC London. 1975. 5 vols.



- NEWSON M, D. Mapwork for flood studies. Part II: Analysis of indices and remapping. Inst. Hydrol., Wallingford, Rep. 1976. Number 25.
- OTTEN, A. AND VAN MONTFORT, M.A.J. "The power of two Test on the type of distributions of extremes. in Journal of Hydrology. 1978 Vol 37, 195-199 pp.
- PEARSON E,S. AND HARTLEY H.O. Biometrika Tables for Statisticians. Cambridge University Press, 1971, vol 2.
- PILGRIM H, DAVID, CORDERY, IAN AND BARON C, BRUCE. "Effects of catchment size on runoff relationships" in Journal of Hidrology. Number, 58, 1982 205-221 pp.
- RAO C, R. Linear statiscal inference and its applications. Wiley, New York, NY., 1965.
- RAYNAL, J. A. "Parameter Estimation Efficiency in Extreme Value Distributions type I and II", Report of Special Study, Civil Engineering Department. Colorado Sate University, 1982
- RAYNAL, J. A. "Bivariate Extreme Value Distributions Applied to Flood Frecuency Analysis", PhD dissertation Colorado Sate University, 1985
- RAYNAL, J. A. y SALAS J. D. "Multivaiate Extreme Value Distributions in Hydrological Analyses", Proceedings of the Rome Symposium, April 1987, IAHS Publications No. 164 pp 11-119, 1987
- RAYNAL J. A. y ESCALANTE S. C. "FLODRO" A User Friendly Personal Computer Package for Flood and Drought Frecuency Analyses Education and Training", Proc. International Symposium on Hydrology and Watwer Resources, 1991
- SEVRUK, B AND GEIGER, H. "Selection of distributions types for extremes of precipitation". in Operational Hydrology Report. 1981.Number 15. WMO-No 560. World Meteorological Organization. Geneva, Switzerland.
- SPRINGALL G, ROLANDO. Hidrologia. México, Instituto de Ingenieria., Facultad de Ingenieria, UNAM, 1970. 210 p.

- STEVENS, M.J AND LYNN P,P. Regional growth curves. Inst Hydrol., Wallingford, Rep. 1978. Number 52.
- TASKER G, D. "Simplified testing of hidrologic regression regions" in Journal Hydraul. Div., Proc ASCE; 1982a Number 108 (HY10), 128-1222 pp.
- TASKER G, D. "Comparing methods of hydrological regionalisation" in Journal of Hydraul. Water REs. Bull., 1982b Number 18(6). 965-970 pp.
- VAN DER SCHAAF, S. "Errors in level recorder data: prevention and detection" in Journal of Hydrology. Number 73, 1984. 373-382 pp.
- VAN MONTFORT, M.A.J: On testing that the distribution of extremes is of type I when type II is the alternative. in Journal of Hydrology. 1970 Vol 11: 421-427 pp.
- VIESSMAN, W, LEWIS, G AND KNAPP, J. Introduction to Hidrology. New York, Harper and Row, 1989. 499 p.
- WALLIS J.R.. "Risk and uncertainties in the evaluation of flood events for the design of hidrologic structures." Paper presented at Seminar on Extreme Hydrological Events Floods and Droughts. Centro di Cult. Sci. "E. Majorana", Erice. 1980.
- WILTSHIRE E , S. "Identification of homogeneous regions for flood frequency analysis" in Journal of Hidrology. Number, 84, 1986. 287-302 pp.
- WILTSHIRE, S AND BERAN, M. "Multivariate techniques for the identification of homegeneous flood frequency regions" in Regional flood frequency analysis, 1987. 133-145 pp.
- WILTSHIRE, S AND BERAN, M. "A significance test for homogeneity of flood frequency regions" in Regional flood frequency analysis., 1987.. 147-158 pp.
- ZOLT, ZRINJI AND BURN H, DONALD. " Flood frequency analysis for ungauged sites using a region of influence approach" in Journal of Hidrology. Number, 153, 1993. 1-21 pp.

## **AGRADECIMIENTOS**

Quiero agradecer el apoyo, comprensión y sobre todo el amor de mis padres Rosita y Alfonso y de mi hermana Rosa María.

A Angélica, por su amor y sus horas de trabajo.

Al Dr. Carlos Escalante, por su ayuda, paciencia y generosidad en la preparación de este trabajo.

A los doctores, Gabriel Echávez, Rolando Springall, Ramón Domínguez y al M.I. Víctor Franco, por sus comentarios y sugerencias.

A mi amigo M.I. Abel Herrera y a su familia, por la hospitalidad que me brindaron en Davis, Ca.

Al Sr. Abel Padilla, quien inició y apoyó mi formación como ingeniero.

A mi abuelo, mis tías, mis tíos y mis primos por su apoyo y compañía.

Al ingeniero Benjamín Domínguez, del Servicio Meteorológico Nacional, por haberme facilitado la información requerida en este estudio.

A Lupita Tejeida, y al Arq. René Olvera.

## **ANEXO I**

### **PROGRAMAS EN QUICK BASIC 4.5**

PROGRAMA: ANDREWS.BAS  
PROGRAMA PARA GRAFICAR Y DELIMITAR REGIONES HIDROLOGICAMENTE HOMOGENEAS EN BASE A  
SUS CARACTERISTICAS FIBIOGRAFICAS

```

# C I 8 : C I F A P
SCREEN 1, 0, 0
COLOR 1, 15
DIM C(50, 50), Y(50), Z(50), F#(50)
LOCATE 10, 7: PRINT "DEFINICION DE REGIONES"
LOCATE 12, 7: PRINT "HIDROLOGICAMENTE HOMOGENEAS"
LOCATE 14, 9: PRINT "TRAZO MULTIDIMENSIONAL"
LOCATE 16, 12: PRINT "CURVAS DE ANDREWS"
LOCATE 22, 11: PRINT " - ANDREWS BAS - "
3 IF INKEY$ = " " THEN LOCATE 25, 33: PRINT TIME$: GOTO 3

```

```

SCREEN 2: SCREEN 0
COLOR 15, 1: C I 5
LOCATE 1, 1: PRINT "TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS"
LOCATE 3, 1: INPUT "SE DESEA COMPARAR DOS REGIONES ( S / N ) = ? : G#
IF G# = "S" OR P# = "S" THEN 250
LOCATE 4, 1: INPUT "SE DESEA CREAR UN ARCHIVO DE DATOS ( S / N ) = ? : R#
IF G# = "S" OR R# = "S" THEN 27

```

```

CLS
LOCATE 3, 1: INPUT "No de estaciones por estudiar = ? : A: PRINT
INPUT " TITULO DE LA GRAFICA : $#
FOR J = 1 TO M
INPUT " NOMBRE DEL ARCHIVO DE DATOS : F#(J)
PRINT " Caracteristicas de la estacion : PRINT
OPEN F#(J) FOR INPUT AS #J
INPUT #J, N
FOR I = 1 TO N
INPUT #J, C(J, I)
PRINT USING "#####.##": C(J, I)
NEXT I: CLOSE #J: PRINT : PRINT : NEXT J
31 IF INKEY$ = " " THEN LOCATE 25, 5: PRINT " Cualquier tecla para continuar ... ": GOTO 31
GOTO 35

```

```

27 CLS
LOCATE 3, 1: INPUT "No de estaciones por crear archivo = ? : M
FOR J = 1 TO M
C I 5: LOCATE 5, 1: PRINT " NOMBRE DE LA " : J : " ESTACION = ? : INPUT: E#(J)
OPEN E#(J) FOR OUTPUT AS #J
LOCATE 7, 1: INPUT "No de caracteristicas de la estacion = ? : N
PRINT #J, N
FOR I = 1 TO N
PRINT " DATO No. = ? : I
INPUT C(J, I)
PRINT #J, C(J, I)
NEXT I: CLOSE #J: NEXT J

```

```

33 FOR J = 1 TO M: CLS
FOR I = 1 TO N
IF C(J, I) > 10 AND C(J, I) < 99 THEN C(J, I) = C(J, I) / 10
IF C(J, I) > 100 AND C(J, I) < 999 THEN C(J, I) = C(J, I) / 100
IF C(J, I) > 1000 AND C(J, I) < 9999 THEN C(J, I) = C(J, I) / 1000
IF C(J, I) > 10000 AND C(J, I) < 99999 THEN C(J, I) = C(J, I) / 10000
NEXT I
NEXT J

```

```

SCREEN 1, 0, 0: COLOR 1, 15: C I 5
LOCATE 1, 1: PRINT "TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS"

```

```

LOCATE 3, 1: PRINT "ESTACION : ", S#

```

```

XMIN = -3.14: XMAX = 3.14
YMIN = -5: YMAX = 25

```

```

WINDOW (XMIN, YMIN) - (XMAX, YMAX)
LINE (XMIN, 0) - (XMAX, 0): LINE (0, YMIN) - (0, YMAX)

```

```

IF N = 0 THEN 6
IF N = 1 THEN 10
IF N = 2 THEN 20
IF N = 3 THEN 30
IF N = 4 THEN 40
IF N = 5 THEN 50
IF N = 6 THEN 60
IF N = 7 THEN 70
IF N = 8 THEN 80
IF N = 9 THEN 90
IF N = 10 THEN 100
IF N = 11 THEN 110
IF N = 12 THEN 120
IF N = 13 THEN 130
IF N = 14 THEN 140
IF N = 15 THEN 150
IF N > 15 THEN 6

```

```

10
REM ----- UN ELEMENTO -----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###.100" * (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J, 1) / 14142)
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = " ": LOOP: END
REM -----

```

```

20
REM ----- DOS ELEMENTOS -----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###.100" * (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J, 1) / 14142) + (C(J, 2) * SIN(X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = " ": LOOP: END
REM -----

```

```

30
REM ----- TRES ELEMENTOS -----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###.100" * (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J, 1) / 14142) + (C(J, 2) * SIN(X)) + (C(J, 3) * COS(X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = " ": LOOP: END

```

```

NEXT J: NEXT X: DO WHILE INKEY$ = "": LOOP: END
REM

```

```

40
REM -----CUATRO ELEMENTOS-----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###", 100 * (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J, 1) / 14142) + (C(J, 2) * SIN(X)) + (C(J, 3) * COS(X)) + (C(J, 4) * SIN(2 * X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = "": LOOP: END
REM

```

```

50
REM -----CINCO ELEMENTOS-----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###", 100 * (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J, 1) / 14142) + (C(J, 2) * SIN(X)) + (C(J, 3) * COS(X)) + (C(J, 4) * SIN(2 * X)) + (C(J, 5) * COS(2 * X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = "": LOOP: END
REM

```

```

60
REM -----SEIS ELEMENTOS-----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###", 100 * (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J, 1) / 14142) + (C(J, 2) * SIN(X)) + (C(J, 3) * COS(X)) + (C(J, 4) * SIN(2 * X)) + (C(J, 5) * COS(2 * X)) + (C(J, 6) * SIN(3 * X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = "": LOOP: END
REM

```

```

70
REM -----SIETE ELEMENTOS-----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###", 100 * (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J, 1) / 14142) + (C(J, 2) * SIN(X)) + (C(J, 3) * COS(X)) + (C(J, 4) * SIN(2 * X)) + (C(J, 5) * COS(2 * X)) + (C(J, 6) * SIN(3 * X)) + (C(J, 7) * COS(3 * X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = "": LOOP: END
REM

```

```

80
REM -----OCHO ELEMENTOS-----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO

```

```

LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###", 100 * (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J, 1) / 14142) + (C(J, 2) * SIN(X)) + (C(J, 3) * COS(X)) + (C(J, 4) * SIN(2 * X)) + (C(J, 5) * COS(2 * X)) + (C(J, 6) * SIN(3 * X)) + (C(J, 7) * COS(3 * X)) + (C(J, 8) * SIN(4 * X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = "": LOOP: END
REM

```

```

90
REM -----NUEVE ELEMENTOS-----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###", 100 * (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J, 1) / 14142) + (C(J, 2) * SIN(X)) + (C(J, 3) * COS(X)) + (C(J, 4) * SIN(2 * X)) + (C(J, 5) * COS(2 * X)) + (C(J, 6) * SIN(3 * X)) + (C(J, 7) * COS(3 * X)) + (C(J, 8) * SIN(4 * X)) + (C(J, 9) * COS(4 * X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = "": LOOP: END
REM

```

```

100
REM -----DIEZ ELEMENTOS-----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###", 100 * (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J, 1) / 14142) + (C(J, 2) * SIN(X)) + (C(J, 3) * COS(X)) + (C(J, 4) * SIN(2 * X)) + (C(J, 5) * COS(2 * X)) + (C(J, 6) * SIN(3 * X)) + (C(J, 7) * COS(3 * X)) + (C(J, 8) * SIN(4 * X)) + (C(J, 9) * COS(4 * X)) + (C(J, 10) * SIN(5 * X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = "": LOOP: END
REM

```

```

110
REM -----ONCE ELEMENTOS-----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###", 100 * (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J, 1) / 14142) + (C(J, 2) * SIN(X)) + (C(J, 3) * COS(X)) + (C(J, 4) * SIN(2 * X)) + (C(J, 5) * COS(2 * X)) + (C(J, 6) * SIN(3 * X)) + (C(J, 7) * COS(3 * X)) + (C(J, 8) * SIN(4 * X)) + (C(J, 9) * COS(4 * X)) + (C(J, 10) * SIN(5 * X)) + (C(J, 11) * COS(5 * X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = "": LOOP: END
REM

```

```

120
REM -----DOCE ELEMENTOS-----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###", 100 * (XMAX - X)
FOR J = 1 TO M

```

```

Y(J) = (C(J,1) / 14142) + (C(J,2) * SIN(X)) + (C(J,3) * COS(X)) + (C(J,4) * SIN(2 * X)) + (C(J,5) * COS(2 * X)) +
(C(J,6) * SIN(3 * X)) + (C(J,7) * COS(3 * X)) + (C(J,8) * SIN(4 * X)) + (C(J,9) * COS(4 * X)) + (C(J,10) * SIN(5 *
X)) + (C
J,11) * COS(5 * X)) + (C(J,12) * SIN(6 * X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = "": LOOP: END
REM -----

```

```

130 -----TRICE ELEMENTOS-----
REM -----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###, 100" (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J,1) / 14142) + (C(J,2) * SIN(X)) + (C(J,3) * COS(X)) + (C(J,4) * SIN(2 * X)) + (C(J,5) * COS(2 * X)) +
(C(J,6) * SIN(3 * X)) + (C(J,7) * COS(3 * X)) + (C(J,8) * SIN(4 * X)) + (C(J,9) * COS(4 * X)) + (C(J,10) * SIN(5 *
X)) + (C
J,11) * COS(5 * X)) + (C(J,12) * SIN(6 * X)) + (C(J,13) * COS(6 * X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = "": LOOP: END
REM -----

```

```

140 -----CATORCE ELEMENTOS-----
REM -----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###, 100" (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J,1) / 14142) + (C(J,2) * SIN(X)) + (C(J,3) * COS(X)) + (C(J,4) * SIN(2 * X)) + (C(J,5) * COS(2 * X)) +
(C(J,6) * SIN(3 * X)) + (C(J,7) * COS(3 * X)) + (C(J,8) * SIN(4 * X)) + (C(J,9) * COS(4 * X)) + (C(J,10) * SIN(5 *
X)) + (C
J,11) * COS(5 * X)) + (C(J,12) * SIN(6 * X)) + (C(J,13) * COS(6 * X)) + (C(J,14) * SIN(7 * X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = "": LOOP: END
REM -----

```

```

150 -----QUINCE ELEMENTOS-----
REM -----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###, 100" (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J,1) / 14142) + (C(J,2) * SIN(X)) + (C(J,3) * COS(X)) + (C(J,4) * SIN(2 * X)) + (C(J,5) * COS(2 * X)) +
(C(J,6) * SIN(3 * X)) + (C(J,7) * COS(3 * X)) + (C(J,8) * SIN(4 * X)) + (C(J,9) * COS(4 * X)) + (C(J,10) * SIN(5 *
X)) + (C
J,11) * COS(5 * X)) + (C(J,12) * SIN(6 * X)) + (C(J,13) * COS(6 * X)) + (C(J,14) * SIN(7 * X)) + (C(J,15) * COS(7 *
X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X: DO WHILE INKEY$ = "": LOOP: END
REM -----

```

```

280 CLS
PRINT: M = 2
INPUT "TITULO DE LA GRAFICA "; S$

```

```

FOR J = 1 TO M
INPUT "NOMBRE DEL ARCHIVO DE DATOS "; E$(J)
PRINT "Características de la estación": PRINT
OPEN E$(J) FOR INPUT AS #J
INPUT #J, H
FOR I = 1 TO H
INPUT #J, C(J,I)
PRINT USING "#####": C(J,I)
NEXT I: CLOSE #J: PRINT: PRINT: NEXT J
30 IF INKEY$ = " " THEN LOCATE 25, 5: PRINT "Cualquier tecla para continuar...": GOTO 28

```

```

FOR J = 1 TO M: CLS
FOR I = 1 TO N
IF C(J,I) > 10 AND C(J,I) < 99 THEN C(J,I) = (C(J,I) / 10)
IF C(J,I) > 100 AND C(J,I) < 999 THEN C(J,I) = (C(J,I) / 100)
IF C(J,I) > 1000 AND C(J,I) < 9999 THEN C(J,I) = (C(J,I) / 1000)
IF C(J,I) > 10000 AND C(J,I) < 99999 THEN C(J,I) = (C(J,I) / 10000)
NEXT I
NEXT J

```

```

SCREEN 1, 0, 0: COLOR 1, 15: CLS
LOCATE 1, 1: PRINT "TRAZO MULTIDIMENSIONAL CURVA DE ANDREWS"
LOCATE 3, 1: PRINT "ESFACON": S$

```

```

XMIN = -3.14: XMAX = 3.14
YMIN = -5: YMAX = 25

```

```

WINDOW (XMIN, YMIN)-(XMAX, YMAX)
LINE (XMIN, 0)-(XMAX, 0): LINE (0, YMIN)-(0, YMAX)

```

```

S5 = 0: C5 = 0: C1 = 0: S1 = 0: S = 0: I = 0: S1 = 0

```

```

REM -----SEIS ELEMENTOS-----
INCREMENTO = (XMAX - XMIN) / 639
FOR X = XMIN TO XMAX STEP INCREMENTO
LOCATE 23, 22: PRINT "Calculando...": PRINT USING "###, 100" (XMAX - X)
FOR J = 1 TO M
Y(J) = (C(J,1) / 14142) + (C(J,2) * SIN(X)) + (C(J,3) * COS(X)) + (C(J,4) * SIN(2 * X)) + (C(J,5) * COS(2 * X)) +
(C(J,6) * SIN(3 * X))
PSET (X, Y(J)), 1
Z(J) = Z(J - 1) + Y(J)
S5 = S5 + ((Y(1) - Y(2)) ^ 2)
S1 = S1 + (Y(1) * Y(2))
S = S + (Y(1) ^ 2)
I = I + (Y(2) ^ 2)
IF J = M THEN PSET (X, Z(J) / M), 2
NEXT J: NEXT X
S1 = (S1 ^ 1) ^ (1 / 2)
CS = (S1 / S1)
LOCATE 20, 22: PRINT "S5 =": PRINT USING "#####": S5
LOCATE 21, 22: PRINT "CS =": PRINT USING "#####": CS
DO WHILE INKEY$ = "": LOOP: END
REM -----

```

PROGRAMA ANOVA.BAS  
PROGRAMA PARA REALIZAR UN ANALISIS DE VARIANZA EN UNA DIRECCION

```

CLS
CLFAR
SCREEN 1,0,0
COLOR 15
LOCATE 6,3:PRINT "MODELOS DE TRANSFERENCIA"
LOCATE 12,8:PRINT "DE INFORMACION HIDROLOGICA"
LOCATE 14,6:PRINT "LA PRUEBA DE ANALISIS DE"
LOCATE 16,8:PRINT "VARIANZA DE UN GRUPO DE DATOS"
LOCATE 22,10:PRINT " A N O V A . B A S "
IF INKEY$="" THEN 1
SCREEN 2:SCREEN 0
COLOR 15,1:CLS

DIM O(50,50),M(50),A(50,50),SA(50)

INPUT "NUMERO DE GRUPOS POR ANALIZAR: ";NK
FOR K=1 TO NK
PRINT "GRUPO No. =";K
INPUT " NOMBRE DEL ARCHIVO DE DATOS = ";A$(K)
OPEN A$(K) FOR INPUT AS #K
INPUT #K,M(K)

FOR J=1 TO M(K)
INPUT #K,O(J,K):NEXT J

FOR J=1 TO M(K)
SM(K)=O(J,K)+SM(K)
A(J,K)=O(J,K)^2
SA(K)=A(J,K)+SA(K)
NEXT J

T=SM(K)+T
W=M(K)+W
Q=SA(K)+Q
NEXT K:CLOSE #K

FOR K=1 TO NK
TN=(SM(K)^2)/M(K)+TN
SST=TN-(T^2)/W
SSE=Q-TN
SS=SST+SSE
NEXT K

CLS:PRINT
PRINT "SST =";SST
PRINT "SSE =";SSE
PRINT "SS =";SS
PRINT
PRINT "GRADOS DE LIBERTAD DEL NUMERADOR =";NK-1
PRINT "GRADOS DE LIBERTAD DEL DENOMINADOR =";W-NK
PRINT
INPUT "VALOR CRITICO DEL ESTADISTICO F =";FCR
MSTR=SST/(NK-1)
PRINT:PRINT "MSTR =";MSTR
MSE=SSE/(W-NK)
PRINT "MSE =";MSE

F=MSTR/MSE
PRINT:COLOR 15,4

```

```

PRINT "VALOR DEL ESTADISTICO F =";:PRINT USING "###.###":F:COLOR 15,1:PRINT
IF F<=FCR THEN:COLOR 30,1:PRINT "SE ACEPTA LA HIPOTESIS NULA";:COLOR 14,1:PRINT "NO HAY
DIFERENCIAS -";:COLOR 15,1:GOTO 66
COLOR 15,1:COLOR 30,1:PRINT "SE RECHAZA LA HIPOTESIS NULA";:COLOR 14,1:PRINT "AL MENOS UNA ES
DIFERENTE -";:COLOR 15,1
PRINT
66:LOCATE 23,2:INPUT "IMPRESION DE RESULTADOS (S/N) =";R$
IF R$="S" OR R$="s" THEN 66
END

```

```

66:COLOR 15,1
CLS:1:PRINT
FOR K=1 TO NK
PRINT "GRUPO No. =";K
PRINT " NOMBRE DEL ARCHIVO DE DATOS = ";A$(K)
NEXT K:1:PRINT
LPRINT "SST =";SST
LPRINT "SSE =";SSE
LPRINT "SS =";SS
LPRINT
LPRINT "GRADOS DE LIBERTAD DEL NUMERADOR =";NK-1
LPRINT "GRADOS DE LIBERTAD DEL DENOMINADOR =";W-NK
LPRINT
LPRINT "VALOR CRITICO DEL ESTADISTICO F =";FCR
MSTR=SST/(NK-1)
LPRINT:LPRINT "MSTR =";MSTR
MSE=SSE/(W-NK)
LPRINT "MSE =";MSE
F=MSTR/MSE
LPRINT
LPRINT "VALOR DEL ESTADISTICO F =";:LPRINT USING "###.###":F:LPRINT
IF F<=FCR THEN:LPRINT "SE ACEPTA LA HIPOTESIS NULA";:LPRINT "NO HAY DIFERENCIAS -":END
LPRINT "SE RECHAZA LA HIPOTESIS NULA";:LPRINT "AL MENOS UNA ES DIFERENTE -"
END

```



PROGRAMA ANOVA2.BAS  
PROGRAMA PARA REALIZAR UN ANALISIS DE VARIANZA EN DOS DIRECCIONES

```
CLS:CLFAR
SCREEN 1,0,0
COLOR 1,5
LOCATE 6,3:PRINT "MODELOS DE TRANSFERENCIA"
LOCATE 12,6:PRINT "DE INFORMACION HIDROLOGICA"
LOCATE 14,6:PRINT "LA PRUEBA DE ANALISIS DE"
LOCATE 16,6:PRINT "VARIANZA EN DOS DIRECCIONES"
LOCATE 22,10:PRINT " A N O V A 2 . B A S "
IF INKEY="" THEN 1
SCREEN 2:SCREEN 0
COLOR 15,1:CLS
```

DATA(50,50),M(50),A(50,50),SA(50),B(50),V(50,50)

```
INPUT "NUMERO DE GRUPOS POR ANALIZAR = ";NK:PRINT
INPUT "NUMERO MAXIMO DE DATOS EN UN GRUPO = ";MAX
FOR K = 1 TO NK
PRINT :PRINT "GRUPO No. = ";K
INPUT " NOMBRE DEL ARCHIVO DE DATOS = ";A$(K)
OPEN A$(K) FOR INPUT AS #K
INPUT #K,M(K)
```

```
FOR J = 1 TO M(K)
INPUT #K,Q(J,K):NEXT J
```

```
FOR J = 1 TO M(K)
SA(K) = Q(J,K) + SA(K)
NEXT J
```

```
T = SM(F) + T
W = M(F) + W
Q = SA(K) + Q
NEXT K:CLOSE #K
```

```
REM CALCULO DE LA SUMA POR RENGLONES
FOR J = 1 TO MAX
FOR K = 1 TO NK
V = Q(J,K) + V
NEXT K:B(J) = V
V = 0:NEXT J
```

```
FOR K = 1 TO NK
TN = (SM(K) ^ 2) / M(K) + TN
SST = TN - (T ^ 2 / W)
NEXT K
```

```
FOR J = 1 TO MAX
TB = (B(J) ^ 2) / NK + TB
SSB = TB - (T ^ 2 / W)
NEXT J
```

```
SS = Q - (T ^ 2 / W)
SSE = SS - SST - SSB
```

```
PRINT
PRINT "SST = ";SST
PRINT "SSB = ";SSB
PRINT "SSE = ";SSE
PRINT "SS = ";SS
PRINT
PRINT " GRADOS DE LIBERTAD ENTRE COLUMNAS = ";NK - 1
PRINT " GRADOS DE LIBERTAD ENTRE LAS FILAS = ";MAX - 1
PRINT " GRADOS DE LIBERTAD PARA EL ERROR = ";(NK - 1) * (MAX - 1)
PRINT
```

```
MSTR = SST / (NK - 1)
PRINT "MSTR = ";MSTR
MSB = SSB / (MAX - 1)
REM MSB CUADRADO MEDIO SEGUN FILAS
PRINT "MSB = ";MSB
MSE = SSE / ((NK - 1) * (MAX - 1))
PRINT "MSE = ";MSE
```

```
REM F PARA COLUMNAS
INPUT "VALOR CRITICO DE F ESTADISTICO F (Para columnas) = ";FCC
F = MSTR / MSE
PRINT :COLOR 15,4
PRINT "VALOR DE F ESTADISTICO F = ";PRINT USING "###.###":F:COLOR 15,1:PRINT
IF F <= FCC THEN :COLOR 30,1:PRINT "SE ACEPTA LA HIPOTESIS NULA";:COLOR 14,1:PRINT " - NO HAY
DIFERENCIAS - ";:COLOR 15,1:GOTO 64
COLOR 15,1:COLOR 30,1:PRINT "SE RECHAZA LA HIPOTESIS NULA";:COLOR 14,1:PRINT " - AL MENOS UNA ES
DIFERENTE - ";:COLOR 15,1
PRINT
```

```
REM F PARA FILAS
64 INPUT "VALOR CRITICO DE F ESTADISTICO F (Para filas) = ";FC1
F = MSB / MSE
PRINT :COLOR 15,4
PRINT "VALOR DE F ESTADISTICO F = ";PRINT USING "###.###":F:COLOR 15,1:PRINT
IF F <= FC1 THEN :COLOR 30,1:PRINT "SE ACEPTA LA HIPOTESIS NULA";:COLOR 14,1:PRINT " - NO HAY
DIFERENCIAS - ";:COLOR 15,1:GOTO 65
COLOR 15,1:COLOR 30,1:PRINT "SE RECHAZA LA HIPOTESIS NULA";:COLOR 14,1:PRINT " - AL MENOS UNA F5
DIFERENTE - ";:COLOR 15,1
PRINT
```

```
65 INPUT "IMPRESION DE RESULTADOS ( S / N ) = ";R$
IF R$ = "S" OR R$ = "N" THEN 66
END
```

```
66 COLOR 15,1
CLS:1:PRINT
PRINT "SST = ";SST
PRINT "SSB = ";SSB
PRINT "SSE = ";SSE
PRINT "SS = ";SS
PRINT
PRINT " GRADOS DE LIBERTAD ENTRE COLUMNAS = ";NK - 1
PRINT " GRADOS DE LIBERTAD ENTRE LAS FILAS = ";MAX - 1
PRINT " GRADOS DE LIBERTAD PARA EL ERROR = ";(NK - 1) * (MAX - 1)
PRINT
MSTR = SST / (NK - 1)
```

```

I PRINT "MSTR = "; MSTR
MSB = SSB / (MAX - 1)
REM MSB CUADRADO MEDIO SEGUN FILAS
I PRINT "MSB = "; MSB
MSE = SSE / ((NK - 1) * (MAX - 1))
I PRINT "MSE = "; MSE

```

```

REM F PARA COLUMNAS
I PRINT "VALOR CRITICO DEL ESTADISTICO F (Para columnas) ="; FCC
F = MSTR / MSE
I PRINT : COLOR 15, 4
I PRINT "VALOR DEL ESTADISTICO F = "; I PRINT USING "###.###"; F; COLOR 15, 1; I PRINT
IF F <= FCC THEN : COLOR 30, 1; I PRINT "SE ACEPTA LA HIPOTESIS NULA"; : COLOR 14, 1; I PRINT " - NO HAY
DIFERENCIAS - "; : COLOR 15, 1; GOTO 130
: COLOR 15, 1; : COLOR 30, 1; I PRINT "SE RECHAZA LA HIPOTESIS NULA"; : COLOR 14, 1; I PRINT " - AL MENOS UNA
ES DIFERENTE - "; : COLOR 15, 1
I PRINT

```

```

REM F PARA FILAS
130 I PRINT "VALOR CRITICO DEL ESTADISTICO F (Para filas) ="; FCC
F = MSB / MSE
I PRINT : COLOR 15, 4
I PRINT "VALOR DEL ESTADISTICO F = "; I PRINT USING "###.###"; F; COLOR 15, 1; I PRINT
IF F <= FCC THEN : COLOR 30, 1; I PRINT "SE ACEPTA LA HIPOTESIS NULA"; : COLOR 14, 1; I PRINT " - NO HAY
DIFERENCIAS - "; : COLOR 15, 1; END
: COLOR 15, 1; : COLOR 30, 1; I PRINT "SE RECHAZA LA HIPOTESIS NULA"; : COLOR 14, 1; I PRINT " - AL MENOS UNA
ES DIFERENTE - "; : COLOR 15, 1
I PRINT

```

PROGRAMA: COEFCOR.BAS  
PROGRAMA PARA CALCULAR EL COEFICIENTE DE CORRELACION ENTRE N ESTACIONES

```

CLS : CLEAR
SCREEN 1, 0, 0
COLOR 1, 5
LOCATE 6, 3; PRINT "MODELOS DE TRANSFERENCIA"
LOCATE 12, 8; PRINT "DE INFORMACION HIDROLOGICA"
LOCATE 14, 11; PRINT "EL COEFICIENTE DE "
LOCATE 16, 7; PRINT "CORRELACION UNIFILAS SU(1)"
LOCATE 18, 11; PRINT "TRIPLE ESTACIONES"
IF INKEY$ = "" THEN I
SCREEN 2; SCREEN 0
COLOR 15, 1; CLS

```

```

DIM X(120, 120), SUMX(120), MED(120), M(120, 120), R(100)
DIM C(100), SUMC(120), SUMM(120), Z(100), B(100), A9(100)

```

```

INPUT "No. DE ESTACIONES POR ANALIZAR = "; M
FOR I = 1 TO M
PRINT : PRINT I; "ESTACION"
INPUT "NOMBRE DEL ARCHIVO DE DATOS = "; A$(I)
OPEN A$(I) FOR INPUT AS #1
INPUT #1, N
FOR K = 1 TO N
INPUT #1, X(I, K)
NEXT K
CLOSE #1
NEXT I

```

```

REM CALCULO DE LAS MEDIAS Y CUADRADOS DE CADA ESTACION
FOR I = 1 TO M
FOR K = 1 TO N
SUMX(I) = SUMX(I) + X(I, K)
M(I, K) = (X(I, K)) ^ 2
SUMM(I) = SUMM(I) + M(I, K)
NEXT K
MED(I) = SUMX(I) / N
NEXT I

```

```

CLS
FOR I = 1 TO M
PRINT "DATOS DE LA ESTACION #"; A$(I); PRINT
FOR K = 1 TO N
PRINT USING "#####"; X(I, K)
NEXT K
PRINT "SUMA X = "; SUMX(I)
PRINT "SUMA MEDIANA ESTACION"; I = "; MED(I)
PRINT "SUMA X^2 = "; SUMM(I)
PRINT "CALCULO DE LA DIFERENCIA PARA DATOS DE SIGUIENTE ESTACION"
Z IF INKEY$ = "" THEN 2
CLS : NEXT I

```

```

PRINT "IMPRESION DE DATOS INICIALES ( 5 / N ) = "; JB
IF JB = "5" THEN 55
GOTO 100

```

```

55 CLS

```

```

FOR I = 1 TO M
  IPRINT " DATOS DE LA ESTACION : A$(I)
  FOR K = 1 TO H
  LPRINT USING " ***** "; X(I, K)
  NEXT K
NEXT I

180 C15 = 1 = 0
4011 = 1
50 J = H + 1

60 PRINT : PRINT " COEFICIENTE DE CORRELACION ENTRE LAS "
PRINT " ESTACIONES : A$(I); " Y "; A$(J); PRINT
  I = I + 1
  FOR K = 1 TO H
    C(I) = X(I, K) * X(J, K)
    SUMC(I) = SUMC(I) + C(I)
    PRINT C(I)
  NEXT K

Z(I) = (N * SUMC(I)) - (SUMX(I) * SUMX(J))
B(I) = (N * SUMX(I) - SUMX(I) ^ 2) * (N * SUMX(J) - SUMX(J) ^ 2)
R(I) = Z(I) / SQRT(B(I))
PRINT : PRINT " COEFICIENTE Rxy = "; PRINT USING "*****"; R(I)

5 IF INKEY$ = "" THEN 5
CLS
J = J + 1
IF J = M + 1 THEN 70
GOTO 60
70 H = H + 1
IF H = M THEN FIND
J = 0
GOTO 50
END

```

**PROGRAMA: DOBLEBOXBAS**  
**PROGRAMA PARA NORMALIZAR MUESTRAS CON LA TRANSFORMACION DE DOBLE POTENCIA**

```

CLS : CLEAR
SCREEN 1, 0, 0
COLOR 1, 15
LOCATE 6, 3: PRINT "MODELOS DE TRANSFERENCIA"
LOCATE 12, 8: PRINT "DE INFORMACION HIDROLOGICA"
LOCATE 14, 8: PRINT "NORMALIZACION DE BOXCOX"
LOCATE 16, 7: PRINT "DE DOBLE POTENCIA"
LOCATE 22, 12: PRINT " D O B L E B O X . B A S "
IF INKEY$ = "" THEN 1
SCREEN 2: SCREEN 0
COLOR 15, 1: C15

INPUT " NOMBRE DEL ARCHIVO DE DATOS = "; A$
OPEN A$ FOR INPUT AS #1
INPUT #1, N
PRINT
INPUT " NOMBRE ARCHIVO RESULTADOS NORMALIZADOS = "; B$
OPEN B$ FOR OUTPUT AS #2

DIM X(N), R(N), I(H)
DIM D(H), A(N), AA(N), AAA(N), AAAA(N), R(N)
DIM Y(H), Z(H), C(N, N), CR(N), SC(N, N)

CLS : REM TOMA DE DATOS
COLOR 30, 0
LOCATE 14, 13: PRINT "-- CALCULANDO ESTADISTICOS --"
COLOR 15, 0
FOR I = 1 TO N
  INPUT #1, D(I)
NEXT I
CLOSE #1

REM SUMA Y MEDIA
FOR I = 1 TO N
  S1 = D(I) + S1
NEXT I
M1 = S1 / N

REM VARIANZA Y DESVIACION ESTANDAR
FOR I = 1 TO N
  A(I) = D(I) - M1
  AA(I) = A(I) ^ 2
  AAA(I) = A(I) ^ 3
  AAAA(I) = A(I) ^ 4
  S1 = A(I) + S1
  SAA = AA(I) + SAA
  SAAA = AAA(I) + SAAA
  SAAAA = AAAA(I) + SAAAA
NEXT I

CLS
VAR = SAA / N
VNG = SAAA / (N - 1)
DSV = SQRT(VNS)

```

```

94 REM PRESENTANDO ESTADÍSTICOS EN PANTALLA
CLS:PRINT "ESTADÍSTICOS DE LA MUESTRA ORIGINAL ESTACION: ",A$:PRINT
PRINT
PRINT "*****:PRINT
PRINT " NÚMERO DE DATOS = ":PRINT USING "#####";N
PRINT " MEDIA = ":PRINT USING "#####";MD
PRINT " VARIANCIA NO SESGADA = ":PRINT USING "#####";VNS
PRINT " VARIANCIA SESGADA = ":PRINT USING "#####";VAR
PRINT " DESVIACION ESTANDAR = ":PRINT USING "#####";DSV

```

REM COEFICIENTES DE ASIMETRÍA, CURTOSIS Y DE VARIACIÓN

```

GS = (SAAA / N) / (DSV * 3)
GNS = GS * ((N - 2) / ((N - 1) * (N - 2)))
KS = (SAAAA / N) / (VAR * 2)
KNS = KS * ((N * 3) / ((N - 1) * (N - 2) * (N - 3)))
CDV = DSV / MD

```

REM PRESENTANDO COEFICIENTES EN PANTALLA

```

PRINT
PRINT "*****:PRINT
PRINT " COEF. SESGADO DE ASIMETRÍA = ":PRINT USING "#####";GS
PRINT " COEF. NO SESGADO DE ASIMETRÍA = ":PRINT USING "#####";GNS
PRINT
PRINT " COEF. SESGADO DE CURTOSIS = ":PRINT USING "#####";KS
PRINT " COEF. NO SESGADO DE CURTOSIS = ":PRINT USING "#####";KNS
PRINT
PRINT " COEFICIENTE DE VARIACIÓN = ":PRINT USING "#####";CDV
PRINT "*****:PRINT
PRINT
INPUT " IMPRESION DE ESTADÍSTICOS (S/N) = ";D$:PRINT
IF D$ = "S" THEN 690
GOTO 100

```

690 LPRINT "ESTADÍSTICOS DE LA MUESTRA ORIGINAL ESTACION: ",A\$:LPRINT

```

LPRINT
LPRINT "*****:LPRINT
LPRINT " NÚMERO DE DATOS = ":LPRINT USING "#####";N
LPRINT " MEDIA = ":LPRINT USING "#####";MD
LPRINT " VARIANCIA NO SESGADA = ":LPRINT USING "#####";VNS
LPRINT " VARIANCIA SESGADA = ":LPRINT USING "#####";VAR
LPRINT " DESVIACION ESTANDAR = ":LPRINT USING "#####";DSV

```

```

LPRINT
LPRINT "*****:LPRINT
LPRINT " COEF. SESGADO DE ASIMETRÍA = ":LPRINT USING "#####";GS
LPRINT " COEF. NO SESGADO DE ASIMETRÍA = ":LPRINT USING "#####";GNS
LPRINT
LPRINT " COEF. SESGADO DE CURTOSIS = ":LPRINT USING "#####";KS
LPRINT " COEF. NO SESGADO DE CURTOSIS = ":LPRINT USING "#####";KNS
LPRINT
LPRINT " COEFICIENTE DE VARIACIÓN = ":LPRINT USING "#####";CDV
LPRINT "*****:LPRINT

```

```

100
INPUT " QUIERES NORMALIZAR LA MUESTRA CON BOX-COX (S/N) = ";D$:PRINT
IF D$ = "S" OR D$ = "s" THEN 10
INPUT " QUIERES NORMALIZAR LA MUESTRA CON LOGARITMOS (S/N) = ";F$:PRINT
IF F$ = "S" OR F$ = "s" THEN 237

```

END

REM NORMALIZACION

```

10 CLS
I = 1
34
I = I + 01
MD = 0; SM = 0; GS = 0; GNS = 0
SA = 0
SAA = 0
SAAA = 0
SAAAA = 0
FOR I = 1 TO N
A(I) = 0
AA(I) = 0
AAA(I) = 0
AAAA(I) = 0
NEXT I

```

```

FOR I = 1 TO N
Y(I) = ((D(I) * T) - 1) / T
NEXT I

```

REM SUMA Y MEDIA

```

FOR I = 1 TO N
SM = Y(I) + SM
NEXT I
MD = SM / N

```

REM VARIANCIA Y DESVIACION ESTANDAR

```

FOR I = 1 TO N
A(I) = Y(I) - MD
AA(I) = A(I) ^ 2
AAA(I) = A(I) ^ 3
AAAA(I) = A(I) ^ 4
SA = A(I) + SA
SAA = AA(I) + SAA
SAAA = AAA(I) + SAAA
SAAAA = AAAA(I) + SAAAA
NEXT I

```

```

VAR = SAA / N
VNS = SAA / (N - 1)
DSV = SQRT(VNS)

```

REM COEFICIENTES DE ASIMETRÍA, CURTOSIS Y DE VARIACIÓN

```

GS = (SAAA / N) / (DSV * 3)
GNS = GS * ((N - 2) / ((N - 1) * (N - 2)))
KS = (SAAAA / N) / (VAR * 2)
KNS = KS * ((N * 3) / ((N - 1) * (N - 2) * (N - 3)))
CDV = DSV / MD

```

```

IF ABS(GNS) < .05 THEN GOSUB SAIR
IF I >= 39 THEN GOSUB SAIR
COR OR 30, 0
LOCAL I, IS:PRINT "-- NORMALIZANDO LA MUESTRA --"
COR OR 15, 0

```

```

LOCATE 17, 20: PRINT " VALOR NORMALIZADOR =: COLOR 10, 0: PRINT " *** ": COLOR 15, 0: PRINT USING
"######": I
LOCATE 19, 20: PRINT "COEF ASIMETRIA NO SESGADO =: COLOR 11, 0: PRINT " *** ": COLOR 15, 0: PRINT USING
"######": GHS
LOCATE 21, 20: PRINT "COEF. CURTOSIS NO SESGADO =: COLOR 12, 0: PRINT " *** ": COLOR 15, 0: PRINT USING
"######": KNS

```

GOTO 34

SALIR:

BEEP: BEEP

PRINT: PRINT " VALOR DE < I > PARA NORMALIZAR LA MUESTRA =: I

CLS: PRINT

PRINT " RESULTADOS DEL PROCESO DE NORMALIZACION ": PRINT

PRINT " VALOR DE < I > PARA NORMALIZAR LA MUESTRA =: I: PRINT USING "###": I: PRINT

PRINT " VALOR DE < GHS > AL NORMALIZAR LA MUESTRA =: I: PRINT USING "###": GHS: PRINT

PRINT " DATOS ORIGINALES DATOS NORMALIZADOS ": PRINT

FOR I = 1 TO N

PRINT USING "#####": D(I); Y(I)

NEXT I

PRINT: INPUT "IMPRESION DE RESULTADOS ( S / N ) =: F9

IF F9 = "S" THEN 195

GOTO 200

195 CLS

PRINT " RESULTADOS DEL PROCESO DE NORMALIZACION - ESTACION -: AS: PRINT

PRINT: PRINT " VALOR DE < T > PARA NORMALIZAR LA MUESTRA =: I: PRINT USING "#####":

T

PRINT: PRINT " VALOR DE < GHS > AL NORMALIZAR LA MUESTRA =: I: PRINT USING "###": GHS:

PRINT

PRINT " DATOS ORIGINALES DATOS NORMALIZADOS ": PRINT

FOR I = 1 TO N

PRINT USING "#####": D(I); Y(I)

NEXT I

200 CLS

PRINT "ESTADISTICOS DE LA MUESTRA NORMALIZADA ESTACION -: AS: PRINT

PRINT

PRINT " \*\*\*\*\*: PRINT

PRINT " NUMERO DE DATOS =: I: PRINT USING "#####": N

PRINT " MEDIA =: I: PRINT USING "#####": MD

PRINT " VARIANCA NO SESGADA =: I: PRINT USING "#####": VHS

PRINT " VARIANCA SESGADA =: I: PRINT USING "#####": VAR

PRINT " DESVIACION ESTANDAR =: I: PRINT USING "#####": DSV

PRINT

PRINT " \*\*\*\*\*: PRINT

PRINT " COEF. SESGADO DE ASIMETRIA =: I: PRINT USING "#####": GS

PRINT " COEF. NO SESGADO DE ASIMETRIA =: I: PRINT USING "#####": GHS

PRINT

PRINT " COEF. SESGADO DE CURTOSIS =: I: PRINT USING "#####": KS

PRINT " COEF. NO SESGADO DE CURTOSIS =: I: PRINT USING "#####": KNS

PRINT

PRINT " COEFICIENTE DE VARIACION =: I: PRINT USING "#####": CV

PRINT: PRINT

PRINT: INPUT "IMPRESION DE ESTADISTICOS ( S / N ) =: D9

IF D9 = "S" THEN 215

GOTO 347

215 CLS

PRINT "ESTADISTICOS DE LA MUESTRA NORMALIZADA ESTACION -: AS: PRINT

PRINT

PRINT " \*\*\*\*\*: PRINT

PRINT " NUMERO DE DATOS =: I: PRINT USING "#####": N

PRINT " MEDIA =: I: PRINT USING "#####": MD

PRINT " VARIANCA NO SESGADA =: I: PRINT USING "#####": VHS

PRINT " VARIANCA SESGADA =: I: PRINT USING "#####": VAR

PRINT " DESVIACION ESTANDAR =: I: PRINT USING "#####": DSV

PRINT

PRINT " \*\*\*\*\*: PRINT

PRINT " COEF. SESGADO DE ASIMETRIA =: I: PRINT USING "#####": GS

PRINT " COEF. NO SESGADO DE ASIMETRIA =: I: PRINT USING "#####": GHS

PRINT

PRINT " COEF. SESGADO DE CURTOSIS =: I: PRINT USING "#####": KS

PRINT " COEF. NO SESGADO DE CURTOSIS =: I: PRINT USING "#####": KNS

PRINT

PRINT " COEFICIENTE DE VARIACION =: I: PRINT USING "#####": CV

PRINT: PRINT

GOTO 347

237

MD = 0: DS = 0: GS = 0: GHS = 0

SA = 0

SAA = 0

SAAA = 0

SAAAA = 0

FOR I = 1 TO N

A(I) = 0

AA(I) = 0

AAA(I) = 0

AAAA(I) = 0

NEXT I

FOR I = 1 TO N

Y(I) = LOG(D(I))

NEXT I

REM SUMA Y MEDIA

FOR I = 1 TO N

SM = Y(I) + SM

NEXT I

MD = SM / N

REM VARIANCA Y DESVIACION ESTANDAR

FOR I = 1 TO N

A(I) = Y(I) - MD

AA(I) = A(I) ^ 2

AAA(I) = A(I) ^ 3

AAAA(I) = A(I) ^ 4

SA = A(I) + SA

SAA = AA(I) + SAA

SAAA = AAA(I) + SAAA

SAAAA = AAAA(I) + SAAAA

NEXT I

```

VAR = SAA / N
VNS = SAA / (N - 1)
DSV = SOR(VNS)

```

REM COEFICIENTES DE ASIMETRIA, CURTOSIS Y DE VARIACION

```

GS = (SAAA / N) / (DSV ** (3))
GHS = GS * ((N ** 2) / ((N - 1) * (N - 2)))
KS = (SAAAA / N) / (VAR ** 2)
KHS = KS * ((N ** 3) / ((N - 1) * (N - 2) * (N - 3)))
CDV = DSV / MD

```

CLS: PRINT

```

PRINT * RESULTADOS DEL PROCESO DE NORMALIZACION *: PRINT
PRINT * DATOS ORIGINALES DATOS NORMALIZADOS *: PRINT
FOR I = 1 TO N
PRINT USING * ***** * : D(I); Y(I)
NEXT I

```

```

PRINT * INPUT * IMPRESION DE RESULTADOS ( S / N ) = * : F$
IF F$ = "S" THEN 1350
GOTO 2000

```

1350 CLS

```

PRINT * RESULTADOS DEL PROCESO DE NORMALIZACION - ESTACION - *: A$: PRINT
PRINT * DATOS ORIGINALES DATOS NORMALIZADOS *: PRINT
FOR I = 1 TO N
PRINT USING * ***** * : D(I); Y(I)
NEXT I

```

2000 CLS

PRINT \* ESTADISTICOS DE LA MUESTRA NORMALIZADA ESTACION: \* : A\$: PRINT

```

PRINT * *****: PRINT
PRINT * NUMERO DE DATOS = *: PRINT USING "*****": N
PRINT * MEDIA = *: PRINT USING "*****": MD
PRINT * VARIANCIA NO SEGGADA = *: PRINT USING "*****": VNS
PRINT * VARIANCIA SEGGADA = *: PRINT USING "*****": VAR
PRINT * DESVIACION ESTANDAR = *: PRINT USING "*****": DSV
PRINT * *****: PRINT
PRINT * COEF. SEGGADO DE ASIMETRIA = *: PRINT USING "*****": GS
PRINT * COEF. NO SEGGADO DE ASIMETRIA = *: PRINT USING "*****": GHS
PRINT *
PRINT * COEF. SEGGADO DE CURTOSIS = *: PRINT USING "*****": KS
PRINT * COEF. NO SEGGADO DE CURTOSIS = *: PRINT USING "*****": KHS
PRINT *
PRINT * COEFICIENTE DE VARIACION = *: PRINT USING "*****": CDV
PRINT * *****: PRINT
PRINT * INPUT * IMPRESION DE ESTADISTICOS ( S / N ) = * : D$
IF D$ = "S" THEN 2150
GOTO 2340

```

2150 CLS

PRINT \* ESTADISTICOS DE LA MUESTRA NORMALIZADA ESTACION: \* : A\$: PRINT

```

PRINT * *****: PRINT
PRINT * NUMERO DE DATOS = *: PRINT USING "*****": N
PRINT * MEDIA = *: PRINT USING "*****": MD

```

```

PRINT * VARIANCIA NO SEGGADA = *: PRINT USING "*****": VNS
PRINT * VARIANCIA SEGGADA = *: PRINT USING "*****": VAR
PRINT * DESVIACION ESTANDAR = *: PRINT USING "*****": DSV
PRINT * *****: PRINT
PRINT * COEF. SEGGADO DE ASIMETRIA = *: PRINT USING "*****": GS
PRINT * COEF. NO SEGGADO DE ASIMETRIA = *: PRINT USING "*****": GHS
PRINT *
PRINT * COEF. SEGGADO DE CURTOSIS = *: PRINT USING "*****": KS
PRINT * COEF. NO SEGGADO DE CURTOSIS = *: PRINT USING "*****": KHS
PRINT *
PRINT * COEFICIENTE DE VARIACION = *: PRINT USING "*****": CDV
PRINT * *****: PRINT

```

347 CLS

MDT = MD

REM INCIAN DOBLE POTENCIA

```

IF KHS > 3 THEN GT = 1
IF KHS < 3 THEN GT = 1
340
G1 = G1 + DO1
MD = 0: SM = 0: GS = 0: GHS = 0
SA = 0
SAA = D
SAAA = 0
SAAAA = 0

```

FOR I = 1 TO N

```

A(I) = 0
AA(I) = 0
AAA(I) = 0
AAAA(I) = 0
NEXT I

```

FOR I = 1 TO N

```

T(I) = ABS(Y(I) - MDT) * GT
IF (Y(I) - MDT) > 0 THEN GOTO 777
T(I) = T(I) * -1
777 NEXT I

```

REM SUMA Y MEDIA

```

FOR I = 1 TO N
SM = T(I) + SM
NEXT I
MD = SM / N

```

REM VARIANCIA Y DESVIACION ESTANDAR

```

FOR I = 1 TO N
A(I) = T(I) * MD
AA(I) = A(I) * 2
AAA(I) = A(I) * 3
AAAA(I) = A(I) * 4
SA = A(I) + SA
SAA = AA(I) + SAA
SAAA = AAA(I) + SAAA
SAAAA = AAAA(I) + SAAAA
NEXT I

```

```
VAR = SAA / N
VNS = SAA / (N - 1)
DSV = SQR(VNS)
```

REM COEFICIENTES DE ASIMETRIA, CURTOSIS Y DE VARIACION

```
GS = (SAAA / N) / (DSV ** (3))
GNS = GS * ((N ** 2) / ((N - 1) * (N - 2)))
KS = (SAAAA / N) / (VAR ** 2)
KNS = KS * ((N ** 3) / ((N - 1) * (N - 2) * (N - 3)))
CDV = DSV / MD
```

IF ABS(GNS) < .0001 OR KNS >= 3 THEN GOSUB SALO

COLOR 30, 0

LOCATE 14, 1: PRINT "-- NORMALIZANDO DOBLE POTENCIA --"

COLOR 15, 0

LOCATE 17, 20: PRINT " VALOR NORMALIZADOR "; COLOR 10, 0: PRINT " \*\*\* "; COLOR 15, 0: PRINT USING

"###.###"; GT

LOCATE 19, 20: PRINT "COEF. ASIMETRIA NO SESGADO "; COLOR 11, 0: PRINT " \*\*\* "; COLOR 15, 0: PRINT USING

"###.###"; GHS

LOCATE 21, 20: PRINT "COEF. CURTOSIS NO SESGADO "; COLOR 12, 0: PRINT " \*\*\* "; COLOR 15, 0: PRINT USING

"###.###"; KNS

GOTO 340

SALO:

BEFP: BEFP

PRINT: PRINT " VALOR DE < GAMA > PARA NORMALIZAR LA MUESTRA = "; G

CLS: PRINT

PRINT " RESULTADOS DEL PROCESO DE NORMALIZACION DE DOBLE POTENCIA "; PRINT

PRINT " VALOR DE < GAMA > PARA NORMALIZAR LA MUESTRA = "; PRINT USING "###.###"; G: PRINT

PRINT " VALOR DE < GHS > AL NORMALIZAR LA MUESTRA = "; PRINT USING "###.###"; GHS: PRINT

PRINT " DATOS NORMALES BOX-COX DATOS DOBLE POTENCIA "; PRINT

FOR I = 1 TO N

PRINT USING "#####" " Y(I); T(I)

NEXT I

PRINT: INPUT " IMPRESION DE RESULTADOS ( S / N ) = "; I\$

IF I\$ = "S" OR I\$ = "n" THEN GHS

GOTO 2005

CLS: CLS

PRINT " RESULTADOS DEL PROCESO DE DOBLE POTENCIA - ESTACION - "; A\$: PRINT

PRINT: PRINT " VALOR DE < GAMA > PARA NORMALIZAR LA MUESTRA = "; PRINT USING "

"###.###"; GT

PRINT: PRINT " VALOR DE < GHS > AL NORMALIZAR LA MUESTRA = "; PRINT USING "###.###"; GHS:

PRINT

PRINT " DATOS NORMALES BOX-COX DATOS DOBLE POTENCIA "; PRINT

FOR I = 1 TO N

PRINT USING "#####" " Y(I); T(I)

NEXT I

2005 CLS

PRINT " ESTADISTICOS DE LA MUESTRA NORMALIZADA DOBLE POTENCIA ESTACION - "; A\$: PRINT

PRINT

PRINT "....."; PRINT

PRINT " NUMERO DE DATOS = "; PRINT USING "#####"; N

PRINT " MEDIA = "; PRINT USING "#####"; MD

```
PRINT " VARIANCIAS NO SESGADA = "; PRINT USING "#####"; VNS
PRINT " VARIANCIAS SESGADA = "; PRINT USING "#####"; VAR
PRINT " DESVIACION ESTANDAR = "; PRINT USING "#####"; DSV
```

PRINT "....."; PRINT

PRINT " COEF. SESGADO DE ASIMETRIA = "; PRINT USING "#####"; GS

PRINT " COEF. NO SESGADO DE ASIMETRIA = "; PRINT USING "#####"; GHS

PRINT

PRINT " COEF. SESGADO DE CURTOSIS = "; PRINT USING "#####"; KS

PRINT " COEF. NO SESGADO DE CURTOSIS = "; PRINT USING "#####"; KNS

PRINT

PRINT " COEFICIENTE DE VARIACION = "; PRINT USING "#####"; CDV

PRINT: PRINT

PRINT: INPUT " IMPRESION DE ESTADISTICOS ( S / N ) = "; D\$

IF D\$ = "S" OR D\$ = "n" THEN 2153

GOTO 2340

2153 CLS

PRINT " ESTADISTICOS DE LA MUESTRA NORMALIZADA ESTACION - "; A\$: PRINT

PRINT

PRINT "....."; PRINT

PRINT " NUMERO DE DATOS = "; PRINT USING "#####"; N

PRINT " MEDIA = "; PRINT USING "#####"; MD

PRINT " VARIANCIAS NO SESGADA = "; PRINT USING "#####"; VNS

PRINT " VARIANCIAS SESGADA = "; PRINT USING "#####"; VAR

PRINT " DESVIACION ESTANDAR = "; PRINT USING "#####"; DSV

PRINT

PRINT "....."; PRINT

PRINT " COEF. SESGADO DE ASIMETRIA = "; PRINT USING "#####"; GS

PRINT " COEF. NO SESGADO DE ASIMETRIA = "; PRINT USING "#####"; GHS

PRINT

PRINT " COEF. SESGADO DE CURTOSIS = "; PRINT USING "#####"; KS

PRINT " COEF. NO SESGADO DE CURTOSIS = "; PRINT USING "#####"; KNS

PRINT

PRINT " COEFICIENTE DE VARIACION = "; PRINT USING "#####"; CDV

PRINT: PRINT

2340 PRINT #2, N

FOR I = 1 TO N

PRINT #2, T(I)

NEXT I

CLOSE #2

PROGRAMA: ELEV.MED.BAS  
PROGRAMA PARA CALCULAR LA ELEVACION MEDIA DE UNA CUENCA

```
CLS:CL:EAR  
SCREEN 1,0,0  
COLOR 1,15  
LOCATE 6,4:PRINT "MODELOS DE TRANSFERENCIA"  
LOCATE 8,9:PRINT "DE INFORMACION HIDROLOGICA"  
LOCATE 14,6:PRINT "ELEVACION MEDIA DE UNA CUENCA"  
LOCATE 16,9:PRINT "A FORMULA DE HORTON"  
LOCATE 22,10:PRINT " - ELEV.MED.BAS - "  
IF BKEYS="" THEN LOCATE 25,33:PRINT TIME$:GOTO 11
```

```
SCREEN 2:SCREEN 0  
COLOR 15,1  
CLS
```

```
INPUT "NOMBRE DE LA ESTACION ":PS:PRINT  
INPUT "NOMBRE DE LA CUENCA PRINCIPAL ":AS:PRINT  
INPUT "CUANTOS PUNTOS SE TIENEN ",N:PRINT:CLS  
DIM F(N)()  
K=0  
FOR I=1 TO N  
PRINT "PARA EL PUNTO =":I  
INPUT "LA ELEVACION ES =":E(I):PRINT  
ET=EI+E(I)  
NEXT I  
PM=ET/N  
PRINT:PRINT
```

```
CLS  
PRINT " ELEVACION MEDIA DE UNA CUENCA"  
PRINT " FORMULA DE HORTON":PRINT  
PRINT " NOMBRE DE LA ESTACION":PS  
PRINT " NOMBRE DE LA CUENCA":AS  
PRINT:PRINT  
PRINT " PUNTO DE ELEVACIONES"  
PRINT " INTERSECCION (m.s.n.m)":PRINT  
FOR I=1 TO N  
PRINT USING " #### ";I:PRINT USING " #####";E(I)  
NEXT I  
PRINT:PRINT  
PRINT " ELEVACION MEDIA =":PRINT USING "####.##",ENL:PRINT "m.s.n.m"  
PRINT "#####":PRINT  
INPUT "DESEAS IMPRIMIR RESULTADOS (SI) =":RS  
IF RS="S" OR RS="5" THEN 770
```

END

770 CLS

```
PRINT " ELEVACION MEDIA DE UNA CUENCA"  
PRINT " FORMULA DE HORTON":PRINT  
PRINT " NOMBRE DE LA ESTACION":PS  
PRINT " NOMBRE DE LA CUENCA":AS  
PRINT:PRINT  
PRINT " PUNTO DE ELEVACIONES"  
PRINT " INTERSECCION (m.s.n.m)":PRINT  
FOR I=1 TO N  
PRINT USING " #### ";I:PRINT USING " #####";E(I)  
IF I=40 OR I=100 OR I=155 OR I=215 THEN PRINT:PRINT:PRINT:PRINT:PRINT:PRINT:PRINT  
PRINT:PRINT:PRINT
```

```
NEXT I  
LPRINT  
LPRINT " NOMBRE DE LA ESTACION":PS:LPRINT  
LPRINT:PRINT  
LPRINT " ELEVACION MEDIA =":PRINT USING "####.##",ENL:LPRINT "m.s.n.m"  
LPRINT:PRINT  
END
```



PROGRAMA: EXTREG.BAS  
PROGRAMA PARA EXTENDER REGISTROS A TRAVES DE CORRELACION LINEAL SIMPLE

```
CLS: CLEAR
SCREEN 1, 0, 0
COLOR 1, 15
LOCATE 6, 3: PRINT "MODULOS DE TRANSFERENCIA"
LOCATE 12, 8: PRINT "DE INFORMACION HIDROLOGICA"
LOCATE 14, 11: PRINT "EXTENSION DE REGISTROS"
LOCATE 16, 7: PRINT "CORRELACION LINEAL SIMPLE"
LOCATE 18, 11: PRINT "EXTREG.BAS"
IF INKEY$ = "" THEN 1
SCREEN 2: SCREEN 0
COLOR 15, 1: CLS
```

```
DATA X(120, 120), SUMX(120), MFD(120), M(120, 120), R(100)
DATA C(100), SUMC(120), SUMH(120), Z(100), B(100), AS(100)
DATA DSV(50, 50), SUMDSV(50), DS1(50), Y(50), DSV(50), Y(50)
```

```
REM INICIO DE ESTACIONES
M = 2
FOR I = 1 TO M
PRINT : PRINT I: "ESTACION"
INPUT "NOMBRE DEL ARCHIVO DE DATOS = "; AS(I)
OPEN AS(I) FOR INPUT AS #1
INPUT #1, N
FOR K = 1 TO N
INPUT #1, X(L, K)
NEXT K
CLOSE #1
NEXT I
```

```
REM CALCULO DE LAS MEDIDAS Y CUADRADOS DE CADA ESTACION
FOR I = 1 TO M
FOR K = 1 TO N
SUMX(I) = SUMX(I) + X(I, K)
M(L, K) = (X(L, K)) ^ 2
SUMM(I) = SUMM(I) + M(L, K)
NEXT K
MFD(I) = SUMX(I) / N
NEXT I
```

```
REM CALCULO DESVIACION ESTANDAR
FOR I = 1 TO M
FOR K = 1 TO N
DSV(I, K) = (X(L, K) - MFD(I)) ^ 2
SUMDSV(I) = SUMDSV(I) + DSV(I, K)
NEXT K
DS1(I) = SQR(SUMDSV(I) / (N - 1))
NEXT I
```

```
CLS
FOR I = 1 TO M
PRINT : PRINT : PRINT "DATOS DE LA ESTACION "; AS(I): PRINT
FOR K = 1 TO N
REM PRINT USING " ***** "; X(L, K)
NEXT K
PRINT "SUMA X = "; SUMX(I)
```

```
PRINT "MEDIA ESTACION"; I: "; MED(I)
PRINT "SUMA X^2 = "; SUMM(I)
PRINT "DESVIACION EST. = "; DS1(I)
NEXT I
```

```
I = 0
40 H = 1
50 J = H + 1
```

```
REM 60 PRINT: PRINT "COEFICIENTE DE CORRELACION ENTRE LAS"
REM PRINT "ESTACIONES "; AS(H); "Y "; AS(J): PRINT
L = L + 1
FOR K = 1 TO N
C(L) = X(H, K) * X(J, K)
SUMC(L) = SUMC(L) + C(L)
REM PRINT C(L)
NEXT K
```

```
Z(L) = ((N * SUMC(L)) - (SUMX(H) * SUMX(J)))
B(L) = ((N * SUMM(H) - SUMM(H) ^ 2) * (N * SUMM(J) - SUMM(J) ^ 2))
R(L) = Z(L) / SQR(B(L))
PRINT: PRINT "COEFICIENTE Rxy = "; PRINT USING "*****"; R(L)
```

```
B = R(L) * (DS1(2) / DS1(1))
PRINT "VALOR DEL COEF B = "; B
```

```
PRINT: INPUT "IMPRESION DE DATOS INICIALES (S/N) = "; JS
IF JS = "S" OR JS = "s" THEN 55
GOTO 100
```

```
55 CLS
FOR I = 1 TO M
PRINT "DATOS DE LA ESTACION "; AS(I)
FOR K = 1 TO N
PRINT USING " ***** "; X(L, K)
PRINT "SUMA X = "; SUMX(I)
PRINT "MEDIA ESTACION; B = "; MFD(I)
PRINT "SUMA X^2 = "; SUMM(I)
PRINT "DESVIACION EST. = "; DS1(I)
NEXT K
NEXT I
```

```
100 CLS
INPUT "AYUDA DE REGISTRO DE LA ESTACION BASE = "; H1
INPUT "AYUDA DE REGISTRO DE LA ESTACION A MEDIR = "; H2: PRINT
INPUT "VALOR CRITICO DE CORRELACION RCV = "; RCMC
```

```
RCMC = SQR(1 / (H1 - 2))
PRINT "VALOR CALCULADO DE R PARA LA MEDIDA = "; RCMC
```

```
IF ABS(R(L)) > RCMC AND RCMC < RCV1 THEN 200
GOTO 220
200 R# = "HABRA MEJORA EN LA ESTIMACION DE LA MEDIDA": GOTO 230
220 R# = "NO HABRA MEJORA EN LA ESTIMACION DE LA MEDIDA"
```

```
230 PRINT "RESULTADO = "; R#
```

```
A1 = ((H2 + 2) * (H1 + 6) * (H1 - 8)) / ((H1 - 3) * (H1 - 5))
```

```

A2 = (B * (N1 - 4)) / (N1 - 3)
A3 = (2 * N2 * (N1 - 4) * 2) / ((N1 - 3) ^ 2)
A4 = ((N1 * N2 * (N1 - 4) * 2) / ((N1 - 3) ^ 2) * (N1 - 2))
A = R(1) * 4 * (A1 - A2 - A3 + A4)
PRINT "A = "; A

```

```

B1 = (6 * (N2 + 2) * (N1 + 6)) / ((N1 - 3) * (N1 - 5))
B2 = (2 * (N1 * 2 - N1 - N1)) / (N1 - 3)
B3 = (2 * N2 * (N1 - 4) * (N1 - 5)) / ((N1 - 3) ^ 2)
B4 = (2 * (N1 - 4) * (N1 + 3)) / (N1 - 3)
B5 = (2 * N1 * N2 * (N1 - 4) * 2) / ((N1 - 3) ^ 2 * (N1 - 2))
B = R(1) * 2 * (B1 + B2 + B3 - B4 - B5)
PRINT "B = "; B

```

```

C1 = (2 * (N1 + 1)) / (N1 - 3)
C2 = (5 * (N2 + 2)) / ((N1 - 3) * (N1 - 5))
C3 = ((N1 + 1) * (2 * N1 + N2 - 2)) / (N1 - 1)
C4 = (2 * N2 * (N1 - 4)) / (N1 - 3) ^ 2
C5 = (2 * (N1 - 4) * (N1 + 1)) / (N1 - 3)
C6 = (N1 * N2 * (N1 - 4) * 2) / ((N1 - 3) ^ 2 * (N1 - 2))
C = C1 - C2 - C3 + C4 + C5 + C6
PRINT "C = "; C

```

```

INPUT D
RCV = SQRT((B * (SQRT(B * 2) - (4 * A * C)))) / (2 * A)

```

```

300 PRINT : PRINT
INPUT "VALOR MINIMO DE RCV PARA LA VARIANZA" = ; RCV
IF ABS(R(1)) > RCV THEN 500
GOTO 550
500 5# = "HABRA MEJORA EN LA ESTIMACION DE LA VARIANZA"; GOTO 600
550 5# = "NO HABRA MEJORA EN LA ESTIMACION DE LA VARIANZA"
600 PRINT "RESULTADO = "; 5#

```

```

NE = (N1 + N2) / (1 + (N2 / (N1 - 2)) * (1 - R(1) ^ 2))
PRINT : PRINT "AVOS MAXIMOS DE REGISTRO QUE PUEDEN SER AMPLIADOS = "; NE

```

```

INPUT "NUMERO DE AVOS A INFERIR" = ; W
PRINT "DATOS DE LA ESTACION BASE"
FOR I = 1 TO W
PRINT "DATO"
INPUT Y(I)
SUMY = Y(I) + SUMY
NEXT I
MEDY = SUMY / W
FOR I = 1 TO W
DSY(I) = (Y(I) - MEDY) ^ 2
SDSY = SDSY + DSY(I)
NEXT I

```

```

CLS
DSY = SQRT(SDSY / (W - 1))
PRINT "MEDIA DE LA MUESTRA BASE PARA INFERIR = "; MEDY
PRINT "DESVIACION ESTANDAR DE LA MUESTRA BASE PARA INFERIR = "; DSY

```

```

REM MEDIA DE LA MUESTRA EXTENDIDA
YMED = MED(2) + (N2 / (N2 + N1)) * B * (MEDY - MED(1))
PRINT "MEDIA DE LA MUESTRA EXTENDIDA = "; YMED

```

```

REM CALCULO DE ALFA CUADRADA
ALFA = (N2 * (N1 - 4) * (N1 - 1)) / ((N2 - 1) * (N1 - 3) * (N1 - 2))
PRINT "ALFA^2 = "; ALFA

```

```

REM VARIANZA DE LA MUESTRA EXTENDIDA
S1 = (N1 - 1) * (DS(2) ^ 2)
REM PRN1 "S1 = "; S1
S2 = (N2 - 1) * (B ^ 2) * (DSY ^ 2)
REM PRN1 "S2 = "; S2
S3 = (((N1 * N2 * B ^ 2)) / (N1 + N2)) * (MEDY - MED(1)) ^ 2
REM PRN1 "S3 = "; S3
S4 = ((N2 - 1) * ALFA) * (1 - R(1) ^ 2) * DS(2) ^ 2
REM PRN1 "S4 = "; S4
S = (1 / (N1 + N2 - 1)) * (S1 + S2 + S3 + S4)

```

```

PRN1 "VARIANZA DE LA MUESTRA EXTENDIDA = "; S

```

```

AB = R(1) * DS(2) / DS(1)
REM PRN1 "AB = "; AB
AC = (SQRT(1 - R(1) ^ 2)) * ALFA * DS(2)
REM PRN1 "AC = "; AC
AA = MED(2) - (AB * MED(1))
REM PRN1 "AA = "; AA
PRINT : PRINT
PRINT "VALOR MINIMO DEL MODULO DE TRANSFERENCIA DE INFORMACION ES:"
PRINT "YL = "; : PRINT USING "###.###"; AA; : PRINT " "; : PRINT USING "###.###"; AB; : PRINT "AL = "; : PRINT
USING "###.###"; AC; : PRINT " "; : PRINT " ";
3 IF INKEY# = " " THEN 3

```

```

REM DATOS INFERIDOS
CLS : PRINT " DATO BASE DATO INFERIDO"
FOR I = 1 TO W
Y(I) = AA + (AB * Y(I)) + (AC * RND)
PRINT USING "#####.#####"; Y(I); Y(I)
NEXT I

```

```

PRINT : INPUT "IMPRESION DE DATOS INICIALES (S / N) = "; S
IF S = "S" OR S = "s" THEN 123
END
123 CLS
PRINT "MODULOS DE TRANSFERENCIA DE INFORMACION HIDROLOGICA"
PRINT "EXTENSION DE REGISTROS"
PRINT "VALOR CRITICO DE CORRELACION RCV = "; RCV
PRINT "VALOR CALCULADO DE R PARA LA MEDIA = "; RCMC

```

```

PRINT "RESULTADO = "; R#
PRINT : PRINT
PRINT "VALOR MINIMO DE RCV PARA LA VARIANZA = "; RCV
PRINT "RESULTADO = "; S#
PRINT "AVOS MAXIMOS DE REGISTRO QUE PUEDEN SER AMPLIADOS = "; NE
PRINT "MEDIA DE LA MUESTRA BASE PARA INFERIR = "; MEDY
PRINT "DESVIACION ESTANDAR DE LA MUESTRA BASE PARA INFERIR = "; DSY
PRINT : PRINT "MEDIA DE LA MUESTRA EXTENDIDA = "; YMED
PRINT "ALFA^2 = "; ALFA
PRINT "VARIANZA DE LA MUESTRA EXTENDIDA = "; S

```

```

I PRINT "Yc = "; I PRINT USING "###.###"; AA; I PRINT "x"; I PRINT USING "###.###"; AB; I PRINT "Xc x";
I PRINT USING "###.###"; AC; I PRINT "et"
I PRINT "      DATOS INFERIDOS "
C15: I PRINT "      DATO BASE      DATO INFERIDO "
FOR I = 1 TO W
I PRINT USING "#####" " Y(I); Y(I)
NEXT I
END

```

PROGRAMA: GMBMIXT2.BAS  
PROGRAMA PARA EL CALCULO DE LOS EVENTOS DE DISEÑO DE LA DISTRIBUCION GUMBEL MIXTA UNA VEZ QUE SE CONOCEN LOS PARAMETROS DE LA DISTRIBUCION.

```

I C15: C1 FEAR
SCREEN 1,0,0
COLOUR 15
LOCATE 8,9: PRINT "MODULOS DE TRANSFERENCIA"
LOCATE 10,9: PRINT "DE INFORMACION HIDROLOGICA"
LOCATE 11,9: PRINT "DISTRIBUCION GUMBEL MIXTA"
LOCATE 16,9: PRINT "CALCULO DE EVENTOS DE DISEÑO"
LOCATE 20,11: PRINT "G M B M I X T 2 . B A S "
DO WHILE KEY$ = "": LOOP
SCREEN 2: SCREEN 0
COLOUR 15,1: C15

```

REM CALCULO DE LA MISMA CON GUMBEL MIXTA  
REM PARAMETROS OBTENIDOS CON EL ALGORITMO DE OPTIMIZACION

```

PRINT "PARAMETROS DE LA ESTACION ": PRINT
INPUT "Beta 1 = "; MU1
INPUT "Alpha 1 = "; A1 F A 1
INPUT "Beta 2 = "; MU2
INPUT "Alpha 2 = "; A1 F A 2
INPUT " P = "; P

```

TOL = .00001

DIM X(60), Y(60), TR(60), FR(60), ER(60), TD(60), O(60)

REM LECTURA DE DATOS

```

INPUT " NOMBRE DEL ARCHIVO DE DATOS = "; A#
OPEN A# FOR INPUT AS #1
INPUT #1, N

```

```

FOR I = 1 TO N
INPUT #1, X(I)
NEXT I
CLOSE #1

```

125 REM ORDENAMIENTO DE DATOS - MAYOR A MENOR  
REM PRINT " TABLA DE DATOS -- MAYOR A MENOR -- "; PRINT

```

I = 1
133 O(I) = X(I)
I = I + 1
IF I > N THEN 137
GOTO 133
137 I = 1
138 J = I + 1
139 IF O(I) > O(J) THEN 143
O = O(I)
O(I) = O(J)
O(J) = O
143 J = J + 1
IF J > N THEN 146
GOTO 139
146 I = I + 1

```

```

IF I = N THEN 150
GOTO 138
150 FOR I = 1 TO N
REM PRINT "      ", O(I)
NEXT I: PRINT

REM PRINT "EL VALOR MINIMO ES: O(N)
REM PRINT "EL VALOR MAXIMO ES: O(I)

CLS : FOR I = 1 TO N
TR(I) = (N + 1) / I
FX = 1 - 1 / TR(I)
PR(I) = 1 - 1 / TR(I)

X1 = O(I) - ABS(O(I)): X = X1: X$ = "X1: GOSUB 1000VALOR MINIMO DE X
X2 = 50 * O(I) + ABS(50 * O(I)): X = X2: X$ = "X2: GOSUB 1000VALOR MAXIMO DE X

500 FROM = (X1 + X2) / 2: X = FROM: X$ = "FROM: GOSUB 1000VALOR PROMEDIO DE X
IF ABS(GP) < TOL THEN GOTO 600
SE = GX * GP
IF SE > 0 THEN X1 = FROM ELSE X2 = FROM
GOTO 500

600 XT(I) = FROM
PRINT "X Ajustado ("; I; ") = "; XT(I)
NEXT I

GOTO IMPRIMIR ' SALE DE LA RUTINA DE BISECCION PARA IMPRIMIR

1000
A = EXP(-EXP(-(X - MU1) / AI FA1))
B = EXP(-EXP(-(X - MU2) / AI FA2))
GX = P * A + (1 - P) * B - FX

IF X$ = "FROM" THEN GOTO 1500
IF X$ = "X2" THEN GOTO 1200

GX1 = GX: RETURN
1200 GX2 = GX: RETURN
1500 GP = GX: RETURN

IMPRIMIR:
PRINT "-----"
PRINT TAB(10); "AJUSTE REALIZADO PARA LA ESTACION "; A$
PRINT TAB(10); "DISTRIBUCION: GUMBEL MIXTA"
PRINT "-----"
PRINT
PRINT TAB(10); "I   IR   DATO X   AJUSTADO XT   ERROR" Z"
PRINT
SUM = 0
FOR I = 1 TO N

ER(I) = (O(I) - XT(I)) ^ 2
SUM = SUM + ER(I)

PRINT TAB(10); USING "###"; I; PRINT USING "#####"; TR(I); PRINT USING "#####"; O(I);
PRINT USING "#####"; XT(I); PRINT USING "#####"; ER(I)
NEXT I

```

```

NEXT I

EE = SQR(SUM / (N - 5)) ' 5 ES EL # DE PARAMETROS

PRINT TAB(10); "===== METODO DE OPTIMIZACION DE LA FUNCION DE MAXIMA VEROSIMILITUD =====";
PRINT
PRINT TAB(10); "Error estandar de ajuste = "; EE: PRINT
PRINT TAB(10); "Beta 1 = "; MU1
PRINT TAB(10); "Alfa 1 = "; AI FA1: PRINT
PRINT TAB(10); "Beta 2 = "; MU2
PRINT TAB(10); "Alfa 2 = "; AI FA2: PRINT
PRINT TAB(10); "P = "; P: PRINT: PRINT

INPUT "DESEA IMPRIMIR (S/N) = "; I$
IF I$ = "S" OR I$ = "s" THEN 90
GOTO 145

90
COLOR 15, I: CLS
LOCATE 10, 12: PRINT "----IMPRI MIENDO----"

LPRINT "-----"
LPRINT TAB(10); "AJUSTE REALIZADO PARA LA ESTACION "; A$
LPRINT TAB(10); "DISTRIBUCION: GUMBEL MIXTA"
LPRINT "-----"
LPRINT
LPRINT TAB(10); "I   IR   DATO X   AJUSTADO XT   ERROR" Z"
LPRINT
SUM = 0
FOR I = 1 TO N

ER(I) = (O(I) - XT(I)) ^ 2
SUM = SUM + ER(I)

LPRINT TAB(10); USING "###"; I; PRINT USING "#####"; TR(I); PRINT USING "#####"; O(I);
LPRINT USING "#####"; XT(I); PRINT USING "#####"; ER(I)
NEXT I

EE = SQR(SUM / (N - 5)) ' 5 ES EL # DE PARAMETROS

PRINT CHR$(27); "G"
LPRINT TAB(10); "===== METODO DE OPTIMIZACION DE LA FUNCION DE MAXIMA VEROSIMILITUD =====";
LPRINT
LPRINT TAB(10); "Error estandar de ajuste = "; EE: LPRINT
LPRINT TAB(10); "Beta 1 = "; MU1
LPRINT TAB(10); "Alfa 1 = "; AI FA1: LPRINT
LPRINT TAB(10); "Beta 2 = "; MU2
LPRINT TAB(10); "Alfa 2 = "; AI FA2: LPRINT
LPRINT TAB(10); "P = "; P: LPRINT: LPRINT

N5 CLS
TD(1) = 2: TD(2) = 5: TD(3) = 10: TD(4) = 20: TD(5) = 50
TD(6) = 100: TD(7) = 500: TD(8) = 1000: TD(9) = 5000: TD(10) = 10000

FOR I = 1 TO 10
FX = 1 - 1 / TD(I)
X1 = I: X$ = "X1: GOSUB 7000VALOR MINIMO DE X
X2 = 10000: X$ = "X2: GOSUB 7000VALOR MAXIMO DE X

```

```

7500 PROM = (X1 + X2) / 2: X = PROM: X$ = "PROM": GOSUB 71000 VAI OR PROMEDIO DE X
IF ABS(GP) < 101 THEN GOTO 7600
SE = GX * GP
IF SE > 0 THEN X1 = PROM ELSE X2 = PROM
GOTO 7500

```

```

7600 XT(I) = PROM
PRINT "X Discreto [":t "] = ": XT(I)
NEXT I

```

```

GOTO IMPRESION ' SALE DE LA RUTINA DE BISECCION PARA IMPRIMIR
71000
A = EXP(-EXP(-(X - MU) / ALFA1))
B = EXP(-EXP(-(X - MU2) / ALFA2))
GX = P * A + (1 - P) * B - FX

```

```

IF X$ = "PROM" THEN GOTO 71300
IF X$ = "X2" THEN GOTO 71200

```

```

GX = GX: RETURN
71200 GX2 = GX: RETURN
71300 GP = GX: RETURN

```

IMPRESION:

```

CLS
PRINT TAB(10); " EVENTOS DE DISEÑO": PRINT
PRINT TAB(10); " ESTACION: "; A$
PRINT TAB(10); "DISTRIBUCION: GUMBEL MIXTA": PRINT
PRINT TAB(10); STRING$(60, "+")
PRINT TAB(10); " PERIODO DE RETORNO GASTO DE DISEÑO "
PRINT TAB(10); STRING$(60, "+")

```

```

FOR I = 1 TO 10
PRINT TAB(10); I: PRINT USING " *****"; TD(I);: PRINT USING " *****"; XT(I)
NEXT I: PRINT

```

```

INPUT "DESEA IMPRESION (S/N) = "; D$
IF D$ = "S" OR D$ = "s" THEN B5
END

```

```

B5 COLOR 15, 1: CLS
LOCATE 10, 12: PRINT " --- IMPRIMIENDO ---"

```

```

I: PRINT: I: PRINT
I: PRINT TAB(10); " EVENTOS DE DISEÑO": I: PRINT
I: PRINT TAB(10); " ESTACION: "; A$
I: PRINT TAB(10); "DISTRIBUCION: GUMBEL MIXTA": I: PRINT: I: PRINT
I: PRINT TAB(10); " PERIODO DE RETORNO GASTO DE DISEÑO ": I: PRINT

```

```

FOR I = 1 TO 10
I: PRINT TAB(10); I: PRINT USING " *****"; TD(I);: I: PRINT USING " *****"; XT(I)
NEXT I
END

```

PROGRAMA: GUMBEL.BAS  
PROGRAMA PARA OBTENER LOS PARAMETROS DE LA DISTRIBUCION GENERAL DE VALORES EXTREMOS TIPO I POR MOMENTOS Y MAXIMA VEROSIMILITUD

```

1:CLS: CLEAR
SCREEN 1, 0, 0
COLOR 1, 15

```

```

LOCATE 8, 9: PRINT "MODELOS DE TRANSFERENCIA"
LOCATE 10, 8: PRINT "DE INFORMACION HIDROLOGICA"
LOCATE 14, 9: PRINT "DISTRIBUCION GENERAL DE"
LOCATE 16, 9: PRINT "VALORES EXTREMOS TIPO I"
LOCATE 20, 11: PRINT "G U M B E L . B A S"
DO WHILE INKEY$ = "": LOOP
SCREEN 2: SCREEN 0
COLOR 15, 1: CLS

```

```

PRINT
E = 271221828#
INPUT " NOMBRE DEL ARCHIVO DE DATOS = "; A$
OPEN A$ FOR INPUT AS #1
INPUT #1, N

```

```

DIM SA(50), SAA(50), SAAA(50), SAAAA(50), T(50)
DIM D(50), D1(50), A(50), AA(50), AA1(50), AAA(50), AAA1(50)
DIM Y(50), YTE(50), FY(50), MUN(300), ALFA(300), P(300)
DIM K(300), DAIU(300), DA(300), ALFA1(300), X(50)

```

```

REM TOMA DE DATOS
FOR I = 1 TO N
INPUT #1, O(I)
NEXT I
CLOSE #1

```

```

REM ESTADISTICOS
CLS: REM TOMA DE DATOS
COLOR 15, 1
LOCATE 5, 26: PRINT " - CALCULANDO PARAMETROS -"
COLOR 14, 1
LOCATE 7, 15: PRINT " - POR EL METODO DE MOMENTOS -"
COLOR 30, 1: LOCATE 23, 15: PRINT " F e r r e . . . . .": COLOR 15, 1

```

```

REM SUMA Y MEDIA
FOR I = 1 TO N
SM = O(I) + SM
NEXT I
MD = SM / N

```

```

REM VARIANCIA Y DESVIACION ESTANDAR
WW = 100 - N
FOR I = 1 TO N

```

```

A(I) = O(I) - MD
AA(I) = A(I) ^ 2
AAA(I) = A(I) ^ 3
AAAA(I) = A(I) ^ 4
SA = A(I) + SA
SAA = AA(I) + SAA
SAAA = AAA(I) + SAAA
SAAAA = AAAA(I) + SAAAA

```

```

COLOR 15, 4: LOCATE 7, 48: PRINT "AVANCE": PRINT USING "####": (WV + I): PRINT "Z": COLOR 15, 1
NEXT I

```

```

VAR = SAA / N
VNS = SAA / (N - 1)
DSV = SQR(VNS)

```

```

REM COEFICIENTES DE ASIMETRIA, CURTOSIS Y DE VARIACION
GS = (SAAA / N) / (VAR * (3 / 2))
GNS = GS * ((N * 2) / ((N - 1) * (N - 2)))
KS = (SAAAA / N) / (VAR * 2)
KNS = KS * ((N * 3) / ((N - 1) * (N - 2) * (N - 3)))
CDV = DSV / MD

```

```

REM DETERMINACION DE LOS PARAMETROS

```

```

REM MOMENTOS
MU = MD - (.45 * DSV)
ALFA = (.76 * DSV)
REM PRINT "MU = "; MU
REM PRINT "ALFA = "; ALFA

```

```

REM MAXIMA VEROSIMILITUD
GOSUB MAXIMA
CLS

```

```

PRINT "PRINT " PARAMETROS DISTRIBUCION GUMBEL ( GVE T-1) ESTACION: "; AS: PRINT "PRINT
REM PRESENTACION DE PARAMETROS EN LA PANTALLA

```

```

PRINT "PRINT " ***** PRINT
PRINT " PARAMETROS POR EL METODO DE MOMENTOS "

```

```

PRINT "
PRINT " ALFA = "; PRINT USING "#####": ALFA
PRINT "
PRINT " MU = "; PRINT USING "#####": MU

```

```

PRINT " ***** PRINT
PRINT " PARAMETROS POR MAXIMA VEROSIMILITUD "

```

```

PRINT "
PRINT " ALFA = "; PRINT USING "#####": ALFA(J)
ALFA = ALFA(J)

```

```

PRINT "
PRINT " MU = "; PRINT USING "#####": MU(J)
MUM = MU(J)

```

```

PRINT " ***** PRINT
PRINT " INPUT " IMPRESION DE PARAMETROS ( S / N) = "; N$

```

```

IF N$ = "S" OR N$ = "s" THEN MO
GOTO 230

```

```

MO CLS: LPRINT "PARAMETROS DISTRIBUCION GUMBEL ( GVE T-1) ESTACION: "; AS: LPRINT: LPRINT
REM PRESENTACION DE PARAMETROS EN LA PANTALLA

```

```

LPRINT " ***** LPRINT
LPRINT " PARAMETROS POR EL METODO DE MOMENTOS "

```

```

LPRINT "
LPRINT " ALFA = "; PRINT USING "#####": ALFA
LPRINT "
LPRINT " MU = "; PRINT USING "#####": MU

```

```

LPRINT " ***** LPRINT
LPRINT " INPUT " IMPRESION DE PARAMETROS ( S / N) = "; N$

```

```

IF N$ = "S" OR N$ = "s" THEN MO
GOTO 230

```

```

LPRINT ***** LPRINT
LPRINT " PARAMETROS POR MAXIMA VEROSIMILITUD "

```

```

LPRINT "
LPRINT " ALFA = "; PRINT USING "#####": ALFA(J)
LPRINT "
LPRINT " MU = "; PRINT USING "#####": MU(J)
LPRINT "

```

```

GOTO 230

```

```

MAXIMA:

```

```

COLOR 14, 1: LOCATE 9, 18: PRINT " - CALCULANDO POR MAXIMA VEROSIMILITUD - "; COLOR 15, 1
COLOR 30, 1
LOCATE 23, 55: PRINT " F = perc. .... "; COLOR 15, 1
MUM() = MU
ALFA() = ALFA

```

```

FOR J = 1 TO 200

```

```

LOCATE 13, 27: PRINT " No. de Iteraciones ... "; J
SY = 0
SEY = 0
SYE = 0

```

```

FOR I = 1 TO N

```

```

Y(I) = (O(I) - MUM(J)) / ALFA(J)
SY = Y(I) + SY
EY(I) = E * (-Y(I))
SEY = EY(I) + SEY
YFE(I) = Y(I) * EY(I)
SYFE = YFE(I) + SYFE
NEXT I

```

```

P(J) = (N) - SEY
R(J) = (N - (SY) + SYFE

```

```

IF (P(J) / ALFA(J)) <= 1E-08 AND (R(J) / ALFA(J)) <= 1E-08 THEN 600

```

```

DMU(J) = ((1) * P(J) - (.26 * R(J))) * (ALFA(J) / N)
DA(J) = ((.26 * P(J)) - (.61 * R(J))) * (ALFA(J) / N)
MUM(J + 1) = MUM(J) + DMU(J)
ALFA(J + 1) = ALFA(J) + DA(J)
IF J = 199 THEN 224
NEXT J
600 RETURN

```

```

224 COLOR 30, 1

```

```

LOCATE 23, 40: PRINT " NO CONVERGE A MENOS DE 50 LUGAR "
DO WHILE INKEY$ = "": LOOP
RETURN

```

```

230 CLS

```

```

REM LIMPIANDO VARIABLES
FOR I = 1 TO N
A(I) = 0
AA(I) = 0
NEXT I

```

REM EVENTOS DE DISEÑO POR MOMENTOS Y MÁXIMA VEROSIMILITUD

PRINT \* PRINT \* ORDEN Tr VALOR MOMENTOS MAX.VEROSIMILITUD \* PRINT

```
  I = 1
  I33 O(I) = O(I)
  I = I + 1
  IF I > N THEN I37
  GOTO I33
  I37 I = I
  I36 J = I + 1
  I39 IF O(I) > O(J) THEN M3
    B = O(I)
    O(I) = O(J)
    O(J) = B
  I43 J = J + 1
  IF J > N THEN I46
  GOTO I39
  I46 I = I + 1
  IF I = N THEN I50
  GOTO I36
```

I50

```
FOR I = 1 TO N
  T(I) = (N + 1) / I
NEXT I
```

EM = 0; EV = 0

FOR I = 1 TO N

AA(I) = MUX - (ALFA \* LOG(-LOG(1 - (1 / T(I)))))

A(I) = MU - (ALFA \* LOG(-LOG(1 - (1 / T(I)))))

IF A(I) < 0 THEN A(I) = 0

IF AA(I) < 0 THEN AA(I) = 0

EM = ((O(I) - A(I)) ^ 2) + EM

EV = ((O(I) - AA(I)) ^ 2) + EV

PRINT USING "###"; I; PRINT USING "#####"; T(I); O(I); A(I); AA(I)

IF I = 20 OR I = 40 OR I = 60 OR I = 80 OR I = 100 OR I = 120 OR I = 140 THEN DO WHILE INKEY\$ = "": LOOP

NEXT I

PRINT

PRINT "EL VALOR MINIMO ES"; O(N)

PRINT "EL VALOR MAXIMO ES"; O(1)

REM ERRORES CUADRATICOS

EMT = SQRT(EM / N - 2)

EVT = SQRT(EV / N - 2)

PRINT "ERROR EN MOMENTOS="; PRINT USING "#####"; EMT

PRINT "ERROR EN MAXIMAV="; PRINT USING "#####"; EVT

DO WHILE INKEY\$ = "": LOOP

CLS

PRINT \* ESTACION : \* AS

PRINT \* EVENTOS DE DISEÑO ESTIMADOS POR MOMENTOS Y MÁXIMA VEROSIMILITUD \* PRINT

PRINT \* PERIODO DE RETORNO DISEÑO POR MOMENTOS MÁXIMA VEROSIMILITUD \*

PRINT \* [Años] [m3/s] [m3/s] \* PRINT

FOR I = 1 TO 11

G = 15

READ I

IF I = 233 THEN G = 14

COLOR G, 1

AA(I) = MUX - (ALFA \* LOG(-LOG(1 - (1 / T))))

X(I) = MU - (ALFA \* LOG(-LOG(1 - (1 / T))))

PRINT USING "#####"; T; PRINT USING "#####"; X(I); AA(I)

NEXT I

DO WHILE INKEY\$ = "": LOOP

DATA 2,2,3,5,10,20,50,100,500,1000,5000,10000

LOCATE 22, 8; INPUT "IMPRESION DE RESULTADOS ( S / N ) ="; W\$

IF W\$ = "S" OR W\$ = "s" THEN 240

END

240

PRINT \* PRINT \* ORDEN Tr VALOR MOMENTOS MAX.VEROSIMILITUD \* PRINT

FOR I = 1 TO 11

PRINT USING "###"; I; PRINT USING "#####"; T(I); O(I); A(I); AA(I)

NEXT I

PRINT

PRINT "EL VALOR MINIMO ES"; O(N)

PRINT "EL VALOR MAXIMO ES"; O(1)

PRINT

PRINT "ERROR EN MOMENTOS="; PRINT USING "#####"; EMT

PRINT "ERROR EN MAXIMAV="; PRINT USING "#####"; EVT

PRINT \* PRINT

PRINT \* ESTACION : \* AS

PRINT \* EVENTOS DE DISEÑO ESTIMADOS POR MOMENTOS Y MÁXIMA VEROSIMILITUD \* PRINT

PRINT \* PERIODO DE RETORNO DISEÑO POR MOMENTOS MÁXIMA VEROSIMILITUD \*

PRINT \* [Años] [m3/s] [m3/s] \* PRINT

FOR I = 1 TO 11

READ T

PRINT USING "#####"; T; PRINT USING "#####"; X(I); AA(I)

NEXT I

PRINT \* PRINT \* PRINT \* PRINT \* PRINT \* PRINT

DATA 2,2,3,5,10,20,50,100,500,1000,5000,10000

PROGRAMA: ISOYETAS BAG  
 PROGRAMA PARA CALCULAR LA PRECIPITACION MEDIA ANUAL EMPLEANDO EL METODO DE LAS ISOYETAS

```
CLS: CLEAR
SCREEN 1,0,0
COLOR 1,15
LOCATE 6,4:PRINT "MODOS DE TRANSFERENCIA"
LOCATE 8,9:PRINT "DE INFORMACION INFORMACION"
LOCATE 14,3:PRINT "CALCULO DE LA PRECIPITACION MEDIA"
LOCATE 16,4:PRINT "UTILIZANDO EL METODO DE LAS ISOYETAS"
LOCATE 22,10:PRINT "ISOYETAS BAG - "
IF BIT(0) = 1 THEN LOCATE 21,33:PRINT "MAYOR COLOR"
```

```
SCREEN 2, SCREEN 0
COLOR 15,1
CLS
```

```
INPUT "NOMBRE DE LA ESTACION "; P$: PRINT
INPUT "NOMBRE DE LA CUENCA PRINCIPAL "; A$: PRINT
```

```
DIM S(10), B(10), D(10), C(10), A(10)
K = D
```

```
FOR I = 1 TO B
  READ S(I): NEXT I
```

```
FOR I = 1 TO B
  READ B(I): NEXT I
```

```
FOR I = 1 TO B
  READ C(I): NEXT I
```

```
FOR I = 1 TO B
  PRINT "ENTRE LA ISOYETA - ", S(I)
  PRINT "Y LA ISOYETA - ", B(I): PRINT
  INPUT "SE TIENE UN AREA EN km2 DE = ", D(I): PRINT
  T = T + D(I)
  A(I) = C(I) * D(I)
  TA = TA + A(I)
NEXT I
```

```
PM = TA / T
```

```
CLS
PRINT "PRECIPITACION MEDIA ANUAL, METODO DE LAS ISOYETAS"; PRINT
PRINT " ESTACION: "; P$
PRINT " EN LA CUENCA: "; A$: PRINT
PRINT " ISOYETAS PRECIPITACION AREA ENTRE "
PRINT " TACION ISOYETAS hpt A1 "
PRINT " mm km2 mm*km2: PRINT
FOR I = 1 TO B
  PRINT USING "####"; S(I);: PRINT "-";: PRINT USING "####"; B(I);: PRINT USING "#####.##";
  C(I): D(I): A(I)
NEXT I
PRINT: PRINT " ";: PRINT USING "#####.##"; T;: TA
PRINT " Precipitacion media = "; PM;: PRINT "mm"
PRINT: PRINT
```

```
INPUT "DESEAS IMPRIMIR RESULTADOS (S/N) = "; R$
IF R$ = "S" OR R$ = "s" THEN 770
END
```

```
770 CLS
```

```
PRINT "PRECIPITACION MEDIA ANUAL, METODO DE LAS ISOYETAS";: PRINT
PRINT " ESTACION: "; P$
PRINT " EN LA CUENCA: "; A$:: PRINT
PRINT " ISOYETAS PRECIPITACION AREA ENTRE "
PRINT " TACION ISOYETAS hpt A1 "
PRINT " mm km2 mm*km2: PRINT
FOR I = 1 TO B
  PRINT USING "####"; S(I);: PRINT "-";: PRINT USING "####"; B(I);: PRINT USING "#####.##";
  C(I): D(I): A(I)
NEXT I
PRINT: PRINT " ";: PRINT USING "#####.##"; T;: TA
PRINT " Precipitacion media = "; PM;: PRINT "mm"
PRINT
```

```
DATA 1500,1200,1000,800,700,600,500,400
DATA 1200,1000,800,700,600,500,400,300
DATA 1260,1100,900,750,650,550,450,350
```



PROGRAMA: LANGBEIN.BAS  
 PROGRAMA PARA CALCULAR LA PRUEBA DE HOMOGENEIDAD DE CUENCAS EN BASE AL METODO DE  
 LANGBEIN

CLS: C1 FAR  
 SCREEN 1, 0, 0: COLOR 1, 15  
 LOCATE 10, 7: PRINT "MODELOS DE TRANSFERENCIA DE"  
 LOCATE 12, 9: PRINT "INFORMACION HIDROLOGICA"  
 LOCATE 16, 8: PRINT "PRUEBA DE HOMOGENEIDAD DE"  
 LOCATE 18, 12: PRINT "LANGBEIN"  
 3 IF INKEY\$ = "" THEN LOCATE 25, 33: PRINT TIME\$: GOTO 3

DIM R(50), T3(50), T10(50), T(50), DD(50), MU(50), ALFA(50)  
 DIM A\$(50), TR(50)

SCREEN 2: SCREEN 0  
 COLOR 15, 1: CLS  
 INPUT "NUMERO DE ESTACIONES POR ANALIZAR = "; N

FOR I = 1 TO N: PRINT: PRINT  
 INPUT "NOMBRE DEL ARCHIVO DE DATOS = "; A\$(I)  
 OPEN A\$(I) FOR INPUT AS #1  
 PRINT "ESTACION", I  
 INPUT #1, R(I)  
 PRINT "Longitud del registro = "; R(I)  
 INPUT #1, T3(I)  
 PRINT "Gasto de Tr = 2.33 = "; T3(I)  
 INPUT #1, T10(I)  
 PRINT "Gasto de Tr = 10.0 = "; T10(I)  
 INPUT #1, MU(I)  
 PRINT " Parametro MU = "; MU(I)  
 INPUT #1, ALFA(I)  
 PRINT " Parametro ALFA = "; ALFA(I)  
 CLOSE #1  
 NEXT I

REM Q10/Q2.33  
 FOR I = 1 TO N  
 T(I) = T10(I) / T3(I)  
 ST = T(I) + ST  
 NEXT I

REM PROMEDIO DE Q10/Q2.33  
 PST = ST / N

REM PST/Q2.33  
 FOR I = 1 TO N  
 DD(I) = T3(I) \* PST  
 NEXT I

REM DESPEJANDO EL Tr  
 FOR I = 1 TO N  
 TR(I) = 1 / (1 - (EXP(-EXP(-(DD(I) - MU(I)) / ALFA(I))))))  
 NEXT I

REM RESULTADOS  
 CLS

PRINT " PRUEBA DE HOMOGENEIDAD METODO DE LA AVENIDA INDICE "; PRINT  
 PRINT " Q2.33 Q10/Q2.33 MU ALFA Tr N ESTACION "; PRINT  
 FOR I = 1 TO N  
 PRINT USING "#####.##", T3(I); T10(I);: PRINT USING "###.##"; T(I); MU(I); ALFA(I); TR(I);: PRINT USING "  
 ## "; R(I);: PRINT A\$(I)  
 NEXT I  
 PRINT: PRINT " Promedio Q10/Q2.33 = ";: PRINT USING "###.##"; PST  
 LOCATE 22, 8: INPUT " SE DESEA IMPRIMIR RESULTADOS ( S / N ) = "; S\$  
 IF S\$ = "S" OR S\$ = "s" THEN 65  
 GOTO 70

REM IMPRESION  
 65 CLS  
 IPRINT " PRUEBA DE HOMOGENEIDAD METODO DE LA AVENIDA INDICE "; IPRINT  
 IPRINT " Q2.33 Q10/Q2.33 MU ALFA Tr N ESTACION "; IPRINT  
 FOR I = 1 TO N  
 IPRINT USING "#####.##", T3(I); T10(I);: IPRINT USING "###.##"; T(I); MU(I); ALFA(I); TR(I);: IPRINT  
 USING "## "; R(I);: IPRINT A\$(I)  
 NEXT I  
 IPRINT: IPRINT " Promedio Q10/Q2.33 = ";: IPRINT USING "###.##"; PST

REM LIMITE DE X y Y  
 70 CLS  
 XMIN = 20: XMAX = 200: YMIN = 1: YMAX = 3

SCREEN 1, 0, 0: COLOR 0, 1

REM DIBUJA LINEAS DE NUMEROS  
 FOR J = 0 TO 2  
 FOR I = 1 TO 10 STEP 10  
 Y = 185 - (J \* 60 + 60 \* LOG(I) / LOG(10))  
 LINE (28, 8 \* Y) (318, 8 \* Y), 2  
 NEXT I: NEXT J

FOR I = 9 TO 149: PSET (94, I), 2  
 PSET (167, I), 2: PSET (239, I), 2  
 PSET (312, I), 2: NEXT I

REM DIBUJA NUMEROS  
 LOCATE 1, 1: PRINT "000": LOCATE 7, 1: PRINT "100": LOCATE 13, 2: PRINT "10"  
 LOCATE 19, 2: PRINT "1": LOCATE 20, 12: PRINT "50"  
 LOCATE 20, 21: PRINT "100": LOCATE 20, 30: PRINT "150"  
 LOCATE 20, 38: PRINT "200"

REM DIBUJA EJE S  
 SCREEN 1, 0, 0: WINDOW (PRINT, YMIN) (XMAX, YMAX)  
 LINE (XMIN + 16, 0) (XMAX, 0): LINE (0, 1 + YMIN) (0, YMAX)

REM LIMITE INFERIOR  
 LINE (0, 0) (5, 05): LINE (5, 08) (10, 26)  
 LINE (10, 26) (20, 45): LINE (20, 45) (50, 64)  
 LINE (50, 64) (100, 75): LINE (100, 75) (200, 81)

REM LIMITE SUPERIOR  
 LINE (0, 248) (5, 22): LINE (5, 22) (10, 185)  
 LINE (10, 185) (20, 16): LINE (20, 16) (50, 136)  
 LINE (50, 136) (100, 126): LINE (100, 126) (200, 118)

```

REM I ETIQUETAS
LOCATE 1, 14: PRINT "TRIENSA DE I ANG BWH"
LOCATE 3, 6: PRINT "Período de retorno"
LOCATE 23, 10: PRINT "Longitud del registro"

```

```

REM DIBUJA ESTACIONES EN LA GRAFICA
FOR I = 1 TO N
CIRCLE (R(I) - 5, LOG(TR(I)) / LOG(10)), (RMAX / Y50), 1
PRINT (R(I) - 5, LOG(TR(I)) / LOG(10)), 1
NEXT I

```

```

4 IF INKEY$ = "" THEN LOCATE 25, 33: GOTO 4

```

PROGRAMA: MULTIPLE.BAS  
PROGRAMA DE CORRELACION LINEAL MULTIPLE

```

CLS: CLEAR
SCREEN 1, 0, 0
COLOR 1, 15
LOCATE 10, 7: PRINT "MODELOS DE TRANSFERENCIA DE"
LOCATE 12, 9: PRINT "INFORMACION HIDROLOGICA"
LOCATE 18, 6: PRINT "CORRELACION LINEAL MULTIPLE"
LOCATE 22, 10: PRINT " - ANALISIS BAS - "
IF INKEY$ = "" THEN LOCATE 25, 33: PRINT TIME$: GOTO 1

```

```

DIM X(50, 50), Y(50, 50), M(50, 50), SUMX(50), W(50, 50), E(50)
DIM SUMX(50), SUMW(50), MED(50), S(50), C(50), BB(50), B(50), SUME(50)
SCREEN 2: SCREEN 0
COLOR 15, 1: CLS

```

```

PRINT : INPUT "NUMERO DE TERMINOS DEPENDIENTES ( M ) = "; M: PRINT

```

```

REM NUMERO DE ECUACIONES Y ORDEN DE LA MATRIZ (P*P)
P = M + 1

```

```

FOR I = 1 TO M
PRINT : PRINT I "Termino dependiente"
INPUT "NOMBRE DEL ARCHIVO DE DATOS = "; A$(I)
OPEN A$(I) FOR INPUT AS #I
INPUT #I, N
FOR K = 1 TO N
INPUT #I, X(I, K)
X(I, K) = LOG(X(I, K)) / LOG(10)
NEXT K: CLOSE #I: NEXT I

```

```

REM : PRINT
INPUT "NOMBRE DEL ARCHIVO DE LOS TERMINOS INDEPENDIENTES = "; Y$
OPEN Y$ FOR INPUT AS #M + 1
INPUT #M + 1, N
FOR K = 1 TO N
INPUT #M + 1, Y(L, K)
Y(L, K) = LOG(Y(L, K)) / LOG(10)
NEXT K: CLOSE #M + 1

```

```

REM CALCULO DE LAS MEDIAS Y CUADRADOS DE CADA ESTACION DEPENDIENTE
FOR I = 1 TO M
FOR K = 1 TO N
SUMX(I) = SUMX(I) + X(I, K)
M(I, K) = (X(I, K)) ^ 2
SUMW(I) = SUMW(I) + W(I, K)
NEXT K
MED(I) = SUMX(I) / N
NEXT I

```

```

REM CALCULO DE LAS MEDIAS Y CUADRADOS DE LA ESTACION INDEPENDIENTE
FOR K = 1 TO N
SUMY = SUMY + Y(L, K)
NEXT K
MEDY = SUMY / N

```

```

CLS
FOR I = 1 TO M
PRINT "TERMINOS DEPENDIENTES DEL ARCHIVO "; A$(I): PRINT
FOR K = 1 TO N
PRINT USING " ***** "; X(I, K)

```

```

NEXT K
PRINT : PRINT " SUMA X = "; SUMX(I)
PRINT : PRINT " MEDIA ESTACION: t = "; MED(I)
PRINT : PRINT " SUMA X^2 = "; SUMX(I)
PRINT : PRINT " Cualquier tecla para siguientes datos "
2 IF INKEY$ = "" THEN 2
CLS : NEXT I

CLS
PRINT " TERMINOS INDEPENDIENTES DEL ARCHIVO : Y$ PRINT
FOR K = 1 TO N
PRINT USING " ***** " Y(L, K)
NEXT K
PRINT : PRINT " SUMA Y = "; SUMY
PRINT : PRINT " MEDIA ESTACION INDEPENDIENTE = "; MEDY
PRINT : PRINT " Cualquier tecla continuar "
3 IF INKEY$ = "" THEN 3
CLS

PRINT : INPUT " IMPRESION DE DATOS INICIALES ( S / N ) = "; J9
IF J9 = "S" THEN 55
GOTO 180

55 CLS
FOR I = 1 TO M
I PRINT " DATOS DE LA ESTACION DEPENDIENTE : A$(I)
FOR K = 1 TO N
I PRINT USING " ***** " X(L, K)
NEXT K
NEXT I
I PRINT : I PRINT : I PRINT
I PRINT " DATOS DE LA ESTACION INDEPENDIENTE : Y$
FOR K = 1 TO H
I PRINT USING " ***** " Y(L, K)
NEXT K

180 CLS : L = 0
40 H = 1
50 J = H + 1

60 PRINT : PRINT " MULTIPLICACION DE DATOS ENTRE LOS "
PRINT " ARCHIVOS : A$(I); Y ; A$(J); PRINT
L = L + 1
FOR K = 1 TO H
C(I) = X(L, K) * Y(L, K)
SUMC(I) = SUMC(I) + C(I)
PRINT C(I)
NEXT K
PRINT : PRINT " SUMA = "; SUMC(L)
5 IF INKEY$ = "" THEN 5
CLS
J = J + 1
IF J = M + 1 THEN 70
GOTO 60
70 H = H + 1
IF H = M THEN 120
J = 0
GOTO 50

120
REM TERMONOS I)
W(L, I) = N

```

```

PRINT : PRINT "W(L,I) ="; W(L, I)
FOR I = 2 TO M + 1
W(L, I) = SUMA(I - 1)
PRINT "W"; I; ": I "; I; " ="; W(L, I)
NEXT I

FOR I = 1 TO M
W(L + 1, I) = SUMX(I)
W(L + 1, I) = SUMD(I)
PRINT "W"; I; ": I "; I; " ="; W(L + 1, I)
PRINT "W"; I + 1; ": I "; I; " ="; W(L + 1, I)
NEXT I

FOR I = 1 TO M - 1
W(I + 1, I + 2) = SUMC(I)
W(I + 2, I + 1) = SUMC(I)
PRINT "W"; I + 1; ": I "; I + 2; " ="; W(I + 1, I + 2)
PRINT "W"; I + 2; ": I "; I; " ="; W(I + 2, I + 1)
NEXT I

REM TERMINOS INDEPENDIENTES
BB(I) = SUMY
PRINT "B"; I; "; BB(I)

FOR I = 2 TO M + 1
FOR K = 1 TO H
E(I) = X(I - 1, K) * Y(L, K)
SUME(I) = SUME(I) + E(I)
NEXT K
NEXT I

FOR I = 2 TO M + 1
BB(I) = SUME(I)
PRINT "B"; I; "; BB(I)
NEXT I

REM

PRINT : PRINT " Cualquier tecla continuar "
6 IF INKEY$ = "" THEN 6

H = M + 1
REM TERMINO INDEPENDIENTE
FOR K = 1 TO N
B(K) = BB(K)
FOR J = 1 TO H
REM PRINT "ecuacion :"; K; " variable: "; J
A(K, J) = W(K, J)
NEXT J
NEXT K

CLS
I9 = "*****"
REM RESOLVIENDO SISTEMA DE ECUACIONES
LOCATE 2, 25 : PRINT "Calculando ....."
FOR K = 1 TO H : PIVOT = A(K, K)
FOR J = 1 TO H
A(K, J) = A(K, J) / PIVOT
NEXT J
A(K, K) = 1 / PIVOT
FOR J = 1 TO N
IF A(J, K) = 0 OR J = K THEN 370
MU = A(J, K)

```

```

FDR1 = 1 TO N
  A(J,1) = A(J,1) - A(K,1) * MU
NEXT L
  A(L,K) = A(L,K) - MU
NEXT K
S(1) = S(1) + A(L,K) * B(K)
NEXT K
CLS
PRINT "SOLUCION:" PRINT: PRINT

S(I) = 10 * S(I)

FOR I = 1 TO N
  PRINT "A(:" I ":" ) = "; PRINT USING UR; S(I)
NEXT I

PRINT: PRINT
PRINT "LA ECUACION QUEDA DE: TPO": PRINT
PRINT "Y = A1 * X1 - (A2) * X2 - (A3) * X3 - (A4) * ...": PRINT
END

```

PROGRAMA: PENDMED.BAS  
PROGRAMA PARA CALCULAR LA PENDIENTE MEDIA DE UNA CUENCA CON EL METODO DE HORTON

```

CLS: CLEAR
SCREEN 1,0,0
COLOR 1,15
LOCATE 6,4: PRINT "MODELOS DE TRANSFERENCIA"
LOCATE 8,9: PRINT "DE INFORMACION HIDROLOGICA"
LOCATE 14,6: PRINT "CALCULO DE LA PENDIENTE MEDIA"
LOCATE 16,6: PRINT "EMPLEANDO EL METODO DE HORTON"
LOCATE 22,10: PRINT " - PENDMED.BAS - "
IF INKEY$ = "" THEN LOCATE 25,33: PRINT TIME$: GOTO 11

```

```

SCREEN 2: SCREEN 0
COLOR 15,1
CLS

```

```

INPUT "NOMBRE DE LA ESTACION = "; P$: PRINT
INPUT "NOMBRE DE LA CUENCA PRINCIPAL = "; A$: PRINT
INPUT "NUMERO MAXIMO DE LINEAS EN UNA MALLA = "; N

```

```

DIM H$(100), N$(100), L$(100), Y$(100)
CLS

```

```

FOR I = 1 TO N
  PRINT "PARA LA LINEA - ", I: PRINT
  PRINT "SE TIENEN LAS SIGUIENTES INTERSECCIONES": PRINT
  INPUT "EN LA DIRECCION X = "; N$(I)
  INPUT "EN LA DIRECCION Y = "; N1$(I): PRINT
  S1X = S1X + N$(I)
  S1Y = S1Y + N1$(I)

```

```

PRINT "SE TIENEN LAS SIGUIENTES LONGITUDES (km)": PRINT
INPUT "EN LA DIRECCION X = "; L$(I)
INPUT "EN LA DIRECCION Y = "; L1$(I): PRINT
S2X = S2X + L$(I)
S2Y = S2Y + L1$(I)

```

```

NEXT I

```

```

REM PENDIENTE EN LA DIRECCION X
Px = (S1X * S) / S2X
REM PENDIENTE EN LA DIRECCION Y
Py = (S1Y * S) / S2Y
REM PENDIENTE DE LA CUENCA
PC = ((S1X + S1Y) * S) / (S2X + S2Y)

```

```

CLS
PRINT " PENDIENTE MEDIA DE UNA CUENCA, CRITERIO DE HORTON": PRINT
PRINT " ESTACION: "; P$
PRINT " EN LA CUENCA: "; A$: PRINT
PRINT " NUMERO DE INTERSECCIONES          LONGITUDES "
PRINT " LA LINEA   Nk   Ny   Lk   Ly "
PRINT " DE LA MALLA": PRINT
FOR I = 1 TO N
  PRINT USING "### " I ":" PRINT USING "##### " N$(I); N1$(I): PRINT USING "#### " L$(I); L1$(I)
NEXT I
PRINT: PRINT " "; PRINT USING "#####"; S1X; S1Y; S2X; S2Y: PRINT: PRINT
PRINT " Pendiente en la direccion X = "; PRINT USING "## ####"; Px

```

```

PRINT "Pendiente en la dirección Y = "; PRINT USING "#####"; PY: PRINT
PRINT "Pendiente de la cuenca = "; PRINT USING "#####"; PC: PRINT
PRINT: PRINT

```

```

INPUT "Deseará imprimir resultados (S/N) = ", K9
IF K9 = "S" OR K9 = "s" THEN 770
END

```

770 CLS

```

PRINT "PENDIENTE MEDIA DE UNA CUENCA, CRITERIO DE HORTON"; LPRINT
LPRINT " ESTACION "; P9
LPRINT " EN LA CUENCA "; A9: LPRINT
LPRINT " NUMERO DE INTERSECCIONES LONGITUDES "
LPRINT " LA LINEA Nk Ny Lx ly "
LPRINT " DE LA MALLA: LPRINT
FOR I = 1 TO N
LPRINT USING "### "; I: LPRINT USING "##### "; NK(I); NY(I): LPRINT USING "##### "; LX(I);
LY(I)
NEXT I
LPRINT: LPRINT " ";: LPRINT USING "#####"; SINK; SLY: LPRINT: LPRINT
LPRINT "Pendiente en la dirección K = ";: LPRINT USING "#####"; PK
LPRINT "Pendiente en la dirección Y = ";: LPRINT USING "#####"; PY: LPRINT
LPRINT "Pendiente de la cuenca = ";: LPRINT USING "#####"; PC: LPRINT
LPRINT: LPRINT

```

PROGRAMA: TAYLOR.BAS  
PROGRAMA PARA CALCULAR LA PENDIENTE MEDIA DE UN RIO CON LA FORMULA DE TAYLOR SCHWARZ

```

CLS: CLEAR
SCREEN 1, 0, 0
COLOR 1, 15
LOCATE 6, 4: PRINT "MODELOS DE TRANSFERENCIA"
LOCATE 8, 9: PRINT "DE INFORMACION HIDROLOGICA"
LOCATE 14, 3: PRINT "CALCULO DE LA PENDIENTE"
LOCATE 16, 4: PRINT "CON LA FORMULA DE TAYLOR-SCHWARZ"
LOCATE 22, 10: PRINT " TAYLOR.BAS - "
IF INKEY$ = "" THEN LOCATE 25, 33: PRINT TIME$: GOTO 11

```

```

SCREEN 2: SCREEN 0
COLOR 15, 1
CLS

```

```

INPUT "NOMBRE DE LA ESTACION "; P9: PRINT
INPUT "NOMBRE DE LA CORRIENTE PRINCIPAL "; A9: PRINT
INPUT "LONGITUD DEL CAJUT EN Km "; L: PRINT
INPUT "TU CUANTAS DISTANCIAS PARCIALES SE TIENEN "; H: PRINT: CLS
DIM S(H * 4), B(H * 4), D(H * 4), C(H * 4), A(5000)
DIM X(5000), Y(1000), H(1000), I(1000), F(1000), T(1000)
DIM Q(1000), G(1000), E(1000), K(1000), O(1000), U(1000), Z(1000)
K = 0: FOR I = 1 TO H
INPUT "DE LA COTA - ", S(I)
INPUT "A LA COTA - ", B(I): PRINT
INPUT "SE TIENE UNA DISTANCIA EN KM DE ", D(I): PRINT
T = T + D(I)
NEXT I
T = L / T
IF L < K THEN 270
C = T: GOTO 280
270 C = T
280 A(0) = 0
FOR I = 1 TO H
C(I) = C * D(I)
A(I) = C(I) + A(I - 1)
NEXT I
CLS
PRINT "ELEVACIONES DISTANCIA DISTANCIA DISTANCIA"
PRINT " PARCIAL COMPENSADA ACUMULADA"
PRINT " m a n m. Km Km Km"; PRINT
FOR I = 1 TO H
PRINT S(I); "; "; B(I);: PRINT USING "##### "; D(I); C(I); A(I)
NEXT I
PRINT: PRINT: PRINT
INPUT "EN CUANTOS TRAMOS SE DIVIDIRA EL PERFIL "; M: CLS
LOCATE 10, 10: PRINT "CALCULANDO ..."
W = L / M
FOR I = 1 TO M
X(I) = W * I: NEXT I
J = 0
FOR I = 1 TO M + M
J = I: A(0) = 0
-80 IF X(I) >= A(J - 1) AND A(J) >= X(I) THEN 500
J = J + 1: GOTO 480
500 Q(I) = A(J - 1)

```

```

G(I) = A(J): F(I) = G(I) - Q(I)
K(I) = S(J): O(I) = B(J)
U(I) = (M(I) - Q(I)) / E(I) * O(I)
Z(I) = ((G(I) - X(I)) / E(I)) * K(I)
Y(I) = U(I) + Z(I)
NEXT I
V = D
FOR I = 1 TO M
  Y(I) = S(I): X(0) = D
  H(I) = Y(I) - Y(I)
  L(I) = X(I) - X(I) - I
  P(I) = (H(I) / L(I)) / 1000
  T(I) = (1 / P(I)) * (1 / 2))
  V = V + T(I)
NEXT I
F = (M / V) * 2: PRINT
CLS
PRINT " I " " S(1) " " US1 U2: PRINT
FOR I = 1 TO M
  PRINT USING " ## "; I: PRINT USING " ### ## " F(I): T(I)
NEXT I
PRINT: PRINT
PRINT " Sca ";: PRINT USING "#####"; F
PRINT: PRINT
51 IF INKEY$ = "" THEN LOCATE 25, 35: GOTO 51
CLS: PRINT " CALCULO DE PENDIENTE POR EL"
PRINT " METODO DE TAYLOR-SCHWARZ: PRINT
PRINT " NOMBRE DE LA ESTACION: F$
PRINT " NOMBRE DE LA CORRIENTE: A$
PRINT " LONGITUD DEL RIO:;: PRINT USING "### ##"; L:; PRINT " km"
PRINT " AREA DE LA CUENCA:;: PRINT USING "### ##"; A:; PRINT " km2"
PRINT " FACTOR DE AJUSTE:;: PRINT USING "### ##"; C
PRINT " FECHA ACTUALIZACION: B$: PRINT
PRINT " ELEVACIONES DISTANCIA DISTANCIA DISTANCIA"
PRINT " PARCIAL COMPENSADA ACUMULADA"
PRINT " { m a n m } (km) (km) (km): PRINT
FOR I = 1 TO N
  PRINT USING " ### "; S(I):; PRINT:;:; PRINT USING "### "; B(I):; PRINT USING "#####"
O(I): C(I): A(I)
NEXT I
PRINT: PRINT "#####"
PRINT " PENDIENTE: S = ";: PRINT USING "#####"; F
PRINT "#####";: PRINT
INPUT "DESEAS IMPRIMIR RESULTADOS (S/N) = "; R$
IF R$ = "S" OR R$ = "s" THEN 770
END
770 CLS: LPRINT " CALCULO DE PENDIENTE POR EL"
LPRINT " METODO DE TAYLOR-SCHWARZ: LPRINT
LPRINT " NOMBRE DE LA ESTACION: F$
LPRINT " NOMBRE DE LA CORRIENTE: A$
LPRINT " LONGITUD DEL RIO:;: LPRINT USING "### ##"; L:; LPRINT " km"
LPRINT " ELEVACIONES DISTANCIA DISTANCIA DISTANCIA"
LPRINT " PARCIAL COMPENSADA ACUMULADA"
LPRINT " { m a n m } (km) (km) (km): LPRINT
FOR I = 1 TO N
  LPRINT USING " ### "; B(I):; LPRINT:;:; LPRINT USING "### "; B(I):; LPRINT USING "#####"
O(I): C(I): A(I)
NEXT I

```

```

LPRINT: LPRINT "#####"
LPRINT " PENDIENTE: S = ";: LPRINT USING "#####"; F
LPRINT "#####";: PRINT
END

```

PROGRAMA: WILTSHIRE.BAS  
 PROGRAMA QUE DEFINE EL ESTADISTICO F SEGUN EL CRITERIO DE WILTSHIRE PARA DEFINIR REGIONES  
 HOMOGENEAS

```
CLS
CLEAR
SCREEN 1,0,0
COLOR 1,5
LOCATE 6,3:PRINT "MODILOS DE TRANSFERENCIA"
LOCATE 12,8:PRINT "DE INFORMACION HIDROLOGICA"
LOCATE 14,8:PRINT "ANALISIS DE VARIANZA"
LOCATE 16,7:PRINT "SEGUN WILTSHIRE"
LOCATE 22,12:PRINT "W.I.L.T.S.H.I.R.E. B.A.S"
IF INKEY$="" THEN 1
SCREEN 2:SCREEN 0
COLOR 15,1:CLS
```

```
DIM A$(50,50), N$(50,50), SM$(50), MD$(50), SA$(50), SAA$(50), SUMZ$(50)
DIM CV$(50,50), A$(50,50), A1$(50,50), VNS$(50), DSV$(50), Z$(50)
DIM CD$(50), V$(50,50), S$(50), P$(50,50), SP$(50), CV$(50)
```

```
INPUT "NUMERO DE GRUPOS POR ANALIZAR = "; NK
FOR K = 1 TO NK:CLS
PRINT "GRUPO No. = "; K
INPUT "No. DE ESTACIONES DEL GRUPO = "; NJ
W = NJ * W
FOR J = 1 TO NJ
PRINT
REN PRINT " NOMBRE DE LA. J. ESTACION = "; INPUT AS$(J, K)
INPUT " NÚMERO DE DATOS = "; N$(J, K)
INPUT " COEFICIENTE DE VARIACION = "; CV$(J, K)
SM$(K) = CV$(J, K) * SM$(K)
NEXT J
```

```
MD$(K) = SM$(K) / NJ
FOR J = 1 TO NJ
A$(J, K) = CV$(J, K) * MD$(K)
AA$(J, K) = A$(J, K) ^ 2
SA$(K) = A$(J, K) * SA$(K)
SAA$(K) = AA$(J, K) * SAA$(K)
NEXT J
```

```
VNS$(K) = SAA$(K) / (NJ - 1)
DSV$(K) = SQRT(VNS$(K))
CPV$(K) = DSV$(K) / MD$(K)
```

```
FOR J = 1 TO NJ
U$(J, K) = VNS$(K) / N$(J, K)
SU$(K) = (1 / U$(J, K)) * SU$(K)
P$(J, K) = CV$(J, K) / U$(J, K)
SP$(K) = P$(J, K) * SP$(K)
CV$(K) = SP$(K) / SU$(K)
NEXT J
```

```
FOR J = 1 TO NJ
Z$(J) = ((CV$(J, K) * CV$(K)) ^ 2 / U$(J, K))
SUMZ$(K) = Z$(J) + SUMZ$(K)
```

NEXT J

```
CLS:PRINT " ESTADISTICOS DE LOS COEF. DE VARIACION DEL GRUPO: K: "; PRINT
PRINT " ***** PRINT
PRINT " NÚMERO DE DATOS = "; PRINT USING "#####"; NJ
PRINT " MEDIA = "; PRINT USING "#####"; MD$(K)
PRINT " VARIANZA NO BISEGADA = "; PRINT USING "#####"; VNS$(K)
PRINT " DESVIACION ESTANDAR = "; PRINT USING "#####"; DSV$(K)
PRINT " COEFICIENTE DE VARIACION = "; PRINT USING "#####"; CV$(K)
PRINT
PRINT
PRINT " ESTADISTICOS DE LA PRUEBA F PARA EL GRUPO: K: PRINT
PRINT " ***** PRINT
FOR J = 1 TO NJ
PRINT " VARIANZA MUESTRAL DE "; A$(J, K); " CV = "; PRINT USING "#####"; U$(J, K)
NEXT J
PRINT " MEDIA DE LOS CV = "; PRINT USING "#####"; CV$(K)
PRINT " COEFICIENTE A. DE LOS CV = "; PRINT USING "#####"; SUMZ$(K)
```

```
LOCATE 22,12:PRINT "Cualquier tecla para continuar..."
2 IF INKEY$="" THEN 2
Z1 = SUMZ$(K) * Z1
CVP = (CV$(K) * SU$(K)) * CVP
SUP = SU$(K) * SUP
NEXT K
CLS
LOCATE 8,12:PRINT " VALOR TOTAL DEL COEFICIENTE A = "; PRINT USING "#####"; Z1
CV1 = CVP / SUP
```

```
FOR K = 1 TO NK
B$(K) = SU$(K) * (CV$(K) - CV1) ^ 2
SB = B$(K) * SB
LOCATE 10,12:PRINT " VALOR TOTAL DEL COEFICIENTE B = "; PRINT USING "#####"; SB
NEXT K
```

```
LOCATE 3,9:PRINT " ESTADISTICO DE PRUEBA F SEGUN WILTSHIRE"
F = (SB / (NK - 1)) / (W - NK)
LOCATE 12,12:PRINT " VALOR PARAMETRO ESTADISTICO F = "; PRINT USING "#####"; F
LOCATE 14,12:PRINT " GRADOS DE LIBERTAD FI NUMERADOR = "; PRINT USING "###"; (NK - 1)
LOCATE 16,12:PRINT " GRADOS DE LIBERTAD DENOMINADOR = "; PRINT USING "###"; (W - NK)
END
```

**ANEXO 2**

**PLANOS DE LA REGION HIDROLOGICA 10**



