



01162
6
24

DIVISION DE ESTUDIOS DE POSGRADO
Facultad de Ingeniería

OPERACIONES ENTRE VARIABLES ALEATORIAS
HIDROLOGICAS

MARCO ANTONIO JACOBO VILLA

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)

DIRIGIDA POR:
DR. CARLOS ESCALANTE SANDOVAL

CIUDAD UNIVERSITARIA
1996

TESIS CON
FALLA DE ORIGEN

TESIS CON
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 esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

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

Dedicado a mis padres:

Carmen y Antonio.

Quiero hacer patente mi agradecimiento a las siguientes personas:
al Dr. Carlos Brebbia, al Dr. Rolando Springall, al Dr. Ramón
Domínguez, al M. en I. Victor Franco, y muy especialmente a los
Doctores Carlos Escalante y Gabriel Echávez, por el apoyo brindado
hasta ahora.

INDICE.

1. INTRODUCCION	1
2. TEORIA	5
2.1 Definiciones básicas de la probabilidad...	5
2.2 Convolución	10
2.3 Función logística bivariada y estadísticos básicos	12
2.4 Análisis de registros	16
3. SUMA DE VARIABLES ALEATORIAS:	
CONFLUENCIA DE DOS RIOS	19
3.1 Descripción general	19
3.2 Metodología propuesta	19
3.3 Aplicación de la metodología	22
4. RESTA DE VARIABLES ALEATORIAS:	
DEFICIT DE LAS SEQUIAS	28
4.1 Descripción general	28
4.2 Metodología propuesta	31
4.3 Aplicación de la metodología	37
5. DIVISION DE VARIABLES ALEATORIAS:	
ANALISIS DE HIDROGRAMAS DE ESCURRIMIENTOS EXTREMOS.....	48
5.1 Descripción general	48
5.2 Metodología propuesta	48
5.3 Aplicación de la metodología	51
6. COMBINACION LINEAL DE VARIABLES ALEATORIAS:	
RELACION LLUVIA-ESCURRIMIENTO.....	56
6.1 Descripción general	56
6.2 Metodología propuesta	57
6.3 Aplicación de la metodología	61
7. COMENTARIOS Y CONCLUSIONES	71
BIBLIOGRAFIA.....	75
APENDICE A. PROGRAMAS DE COMPUTADORA	78
A.1 Descripción y listados de los programas ..	79
A.2 Guía para el uso de los programas ejecutables	102
APENDICE B. HIDROGRAMAS DE GASTOS DE AVENIDAS DE LA ESTACION HUITES, SINALOA.....	110
B.1 Representación tabular de los gastos máximos mensuales para el periodo 1941-1993 (gastos en m ³ /s)	111
B.2 Representación esquemática de los gastos máximos mensuales en el periodo 1941-1993	115
B.3 Representación esquemática de los escurrimientos medios mensuales menores a 50 m ³ /s	121
APENDICE C. EJEMPLOS DE CALCULO DE LOS PARAMETROS ESTADISTICOS UTILIZADOS.....	127

1. INTRODUCCION

La diversidad de condiciones geográficas y climatológicas en el territorio nacional condicionan la disponibilidad del recurso agua, tanto en espacio como en tiempo. En términos globales se sabe que el 31 por ciento del territorio es árido, el 36 por ciento es semiárido y el 33 por ciento es húmedo o subhúmedo, CNA, 1994. Desde un punto de vista regional, la variación en la disposición del recurso se hace mas clara, como se muestra en el siguiente resumen de distribución de agua superficial por cuencas y vertientes:

Vertiente del Pacífico.- Constituida por la península de Baja California y la vertiente continental, la cual inicia desde la desembocadura del río Colorado, en el extremo norte del golfo de Cortés, y concluye en la desembocadura del río Suchiate, en la frontera de México con Guatemala. La península de Baja California tiene una extensión territorial equivalente al 7 por ciento del total nacional y sus aguas superficiales representan sólo el 0.2 por ciento de la cifra nacional. La vertiente continental ocupa más del 42 por ciento del territorio del país y el escurrimiento superficial sólo representa el 29 por ciento del total nacional (CNA, 1994).

Vertiente del Golfo.- Va desde el río Bravo, en la frontera de México con Estados Unidos, hasta el río Champotón, en Campeche. La superficie territorial de esta vertiente equivale a 33 por ciento de la extensión de todo el país y el agua que capta representa más del 67 por ciento del total nacional (CNA, 1994).

Vertiente interior.- Son todas aquellas cuencas que no descargan al mar, correspondiendo al 16 por ciento de la extensión territorial total y con un escurrimiento superficial del menos del 1 por ciento (CNA, 1994).

Vertiente del Caribe.- corresponde a la península de Yucatán y debido a las condiciones geológicas existentes en la mayor parte de la región, los escurrimientos superficiales son transformados en flujos subterráneos. La extensión territorial de la vertiente es del 2 por ciento del total nacional y la única descarga superficial significativa que se presenta, el río Hondo, es de alrededor del 0.2 por ciento (CNA, 1994).

Como puede apreciarse, es indispensable contar con herramientas estadísticas que permitan obtener la información necesaria tanto para el diseño y construcción de la infraestructura requerida para aumentar el control que se tiene en la actualidad sobre los recursos hidráulicos, como para continuar, y en caso de ser necesario, mejorar la operación y manejo de las obras existentes.

El objetivo del trabajo es presentar un conjunto de herramientas estadísticas con las cuales sea posible predecir, a partir de valores propuestos de periodo de retorno o de probabilidad de ocurrencia, el evento extremo probable que se espera de un fenómeno o proceso hidrológico, el cual puede ser representado a través de la suma, resta, división y combinación lineal de dos variables aleatorias; para desarrollar dichas herramientas se utiliza la convolución de funciones de densidad de probabilidad conjunta bivariada, empleando como modelo propuesto la función logística bivariada (distribución bivariada de valores extremos tipo II).

A continuación se exponen los motivos por los cuales se propone el uso de funciones multivariadas (bivariadas como caso particular) en el análisis estadístico de fenómenos hidrológicos:

- Los fenómenos y procesos hidrológicos son el resultado de la interacción de diferentes variables, en espacio y tiempo, que ocurren de manera simultánea. De modo que si se considera un modelo multivariado, se podría esperar una mejor representación del fenómeno en cuestión.
- Al utilizar funciones multivariadas se hace uso de mayor

información, la cual está contenida no sólo en los datos de cada una de las variables, sino también en las relaciones que existen entre ellas (transferencia de información).

- El análisis puede llevarse a cabo abarcando cualquier situación que se presente entre dependencia total e independencia total de las variables consideradas.

- La metodología propuesta puede ser aplicada a fenómenos con variables cuyos registros son de longitudes diferentes. Esto se lleva a cabo a través de la estimación de los parámetros de la función multivariada propuesta.

Para cada una de las operaciones consideradas se propuso una metodología a seguir, aplicándose a un fenómeno particular. Los casos considerados son los siguientes:

- Estimación del gasto máximo esperado del escurrimiento en la confluencia de dos corrientes aforadas aguas arriba, utilizando la suma de los gastos máximos anuales respectivos.

- Estimación de los valores de déficit y duración máximos esperados de las sequías, a partir de datos de escurrimiento en una corriente dada, restando las variables aleatorias volumen de referencia (obtenido al multiplicar el gasto umbral asociado al fenómeno de sequía por la duración de la sequía) menos el volumen de escurrimiento instantáneo (representado por el área abajo de la curva de los gastos instantáneos menores al gasto de referencia).

- Estimación del valor máximo esperado de la relación gasto pico/volumen total descargado de las avenidas máximas anuales registradas en una estación hidrométrica.

- Estimación del valor máximo esperado del gasto pico a partir de la combinación lineal de las variables aleatorias gasto antecedente (parámetro que indica el grado de humedad que tiene la cuenca antes del escurrimiento directo) y la precipitación efectiva, utilizando un modelo linealizado a través de una transformación logarítmica (modelo de Hewlett).

Para las cuatro aplicaciones se desarrollaron programas de

computadora que sirvieron de apoyo en la solución de los problemas tratados.

2. TEORIA

Las metodologías a utilizar en los capítulos tres, cuatro, cinco y seis, requieren del uso de algunos de los conceptos, definiciones y resultados de la teoría de la probabilidad, de la estadística, así como de la hidrología. En el subcapítulo 2.1 se presentan las definiciones y teoremas mas relevantes de la probabilidad; en el 2.2 se define la convolución para los casos de suma, diferencia, relación y combinación lineal de dos variables aleatorias; en el 2.3 se presentan los principales parámetros estadísticos a utilizar en los capítulos siguientes, así como la función logística bivariada; en el 2.4 se hace un resumen de los conceptos y resultados mas importantes para manejar los registros históricos de las variables hidrológicas de interés.

2.1 DEFINICIONES BASICAS DE LA PROBABILIDAD

Las siguientes definiciones son fundamentales para el resto del capítulo (Meyer, 1965):

Experimento .- Proceso por el cual se obtienen los puntos de interés para analizar algún fenómeno (determinístico o aleatorio).

Espacio muestral .- Es el conjunto de todos los valores posibles arrojados por el experimento.

Evento .- Es un conjunto de valores posibles resultado de un experimento, el cual está relacionado con un espacio muestral particular; es decir, es un subconjunto del espacio muestral.

Una función evaluada en los números reales y que está definida en

un espacio muestral S se denomina variable aleatoria (DeGroot, 1986). Se dice que una variable aleatoria X tiene una distribución discreta si X puede tomar solamente un número k de diferentes valores x_1, \dots, x_k o una secuencia infinita de diferentes valores x_1, x_2, \dots . Si una variable aleatoria X tiene una distribución discreta, la función de probabilidad de X es definida como la función f tal que para cualquier real x se cumpla que

$$f(x) = P_r(X = x) \quad (2.1)$$

donde P_r es la probabilidad. Así, la probabilidad de cualquier subconjunto A de los reales puede ser determinado por la relación

$$P_r(X \in A) = \sum_{x_1 \in A} f(x_1) \quad (2.2)$$

Se dice que una variable aleatoria X tiene una distribución continua si existe una función no negativa f , definida en los números reales, tal que para cualquier intervalo A

$$P_r(X \in A) = \int_A f(x) dx \quad (2.3)$$

La función f se denomina función de densidad de probabilidad de X . Si una variable aleatoria X tiene una distribución continua, la probabilidad de que X pertenezca a cualquier subconjunto de los reales puede ser encontrada integrando la función de densidad de X sobre aquel subconjunto.

Todas las funciones de distribución deben satisfacer los requerimientos siguientes:

$$i) \quad f(x) \geq 0 \quad (2.4)$$

$$ii) \quad \int_{-\infty}^{\infty} f(x) dx = 1 \quad (2.5)$$

Si una variable aleatoria X tiene una distribución continua, entonces $P_r(X=x)=0$ para cada valor individual x .

La función de distribución F de una variable aleatoria X es una función definida para cada número real x , esto es

$$F(x)=P_r(X \leq x) \quad \text{para } -\infty < x < \infty \quad (2.6)$$

Esto es válido para distribuciones de X discretas, continuas y mixtas (que son discretas y continuas a la vez). Esta función también recibe el nombre de función de distribución acumulada.

La función de distribución de cualquier variable aleatoria X es una función F definida en los reales. El valor de $F(x)$ en cualquier punto x debe ser un número en el intervalo $[0,1]$, debido a que $F(x)$ es la probabilidad del evento $\{X \leq x\}$. Esta función debe cumplir con las propiedades siguientes:

i) La función $F(x)$ es no decreciente conforme x crece; esto es, si $x_1 < x_2$ entonces $F(x_1) \leq F(x_2)$.

ii) $\lim_{x \rightarrow -\infty} F(x) = 0$ y $\lim_{x \rightarrow \infty} F(x) = 1$ (2.7)

iii) Una función de distribución es siempre continua por la derecha, es decir, $F(x) = F(x^+)$ en cada punto x .

Si la función de distribución de una variable X es conocida, entonces se puede determinar el valor de la probabilidad de que X se encuentre en cualquier intervalo específico perteneciente a los reales.

A continuación se presentan algunos teoremas importantes sobre la función de distribución y la probabilidad.

Teorema 2.1. Para cualquier valor x dado,

$$P_r(X > x) = 1 - F(x) \quad (2.8)$$

Teorema 2.2. Para cualesquiera valores dados x_1 y x_2 tales que $x_1 < x_2$,

$$P_r(x_1 < X < x_2) = F(x_2) - F(x_1) \quad (2.9)$$

Teorema 2.3. Para cualquier valor dado x ,

$$P_r(X < x) = F(x^-) \quad (2.10)$$

Teorema 2.4. Para cualquier valor dado x ,

$$P_r(X = x) = F(x^+) - F(x^-) \quad (2.11)$$

En el caso de las variables aleatorias discretas X , la distribución puede ser representada ya sea por la función de probabilidad como por la distribución.

En las variables aleatorias continuas, la relación entre la función de distribución y la función de densidad se expresa como

$$F(x) = P_r(X \leq x) = \int_{-\infty}^x f(t) dt \quad (2.12)$$

y si en cualquier punto se cumple que $f(x)$ es continua, entonces

$$F'(x) = dF(x)/dx = f(x) \quad (2.13)$$

La distribución de probabilidad conjunta de dos variables aleatorias es denominada la distribución bivariada. La función de probabilidad conjunta de dos variables aleatorias discretas X, Y se define como la función f tal que para cualquier punto (x, y) en el plano XY se cumple que

$$f(x, y) = P_r(X=x \text{ y } Y=y) \quad (2.14)$$

Si la secuencia $(x_1, y_1), (x_2, y_2), \dots$ incluye todos los valores posibles del par (X, Y) , entonces

$$\sum_{i=1}^{\infty} f(x_i, y_i) = 1 \quad (2.15)$$

Para cualquier subconjunto A del plano XY

$$P_r[(X, Y) \in A] = \sum_{(x_i, y_i) \in A} f(x_i, y_i) \quad (2.16)$$

Se dice que dos variables aleatorias continuas tienen una distribución conjunta, si existe una función no negativa f definida sobre el plano XY tal que para cualquier subconjunto A del plano

$$P_r[(X, Y) \in A] = \iint_A f(x, y) dx dy \quad (2.17)$$

donde f es la función de probabilidad conjunta de X y Y . Esta función debe satisfacer las siguientes propiedades:

i) $f(x, y) \geq 0$ para $-\infty < x < \infty$ y $-\infty < y < \infty$ (2.18)

ii) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$ (2.19)

La probabilidad de que un par ordenado (X, Y) pertenezca a cualquier región del plano XY se encuentra integrando f sobre dicha región.

Si X y Y tienen una distribución conjunta continua con una distribución de densidad conjunta, entonces la función de distribución conjunta para cualesquiera valores x, y es

$$F(x, y) = \int_{-\infty}^y \int_{-\infty}^x f(r, s) dr ds \quad (2.20)$$

La función de densidad conjunta puede ser derivada a partir de la función de distribución a través de la expresión

$$f(x,y) = \partial^2 F(x,y) / \partial x \partial y \quad (2.21)$$

esto siempre y cuando exista la derivada en el punto (x,y) .

Cuando la distribución bivariada conjunta depende solamente de una de las variables X o Y , entonces se obtiene la función F_1 , denominada distribución marginal.

Se dice que dos variables aleatorias X y Y son independientes si, para cualesquiera dos conjuntos A y B de números reales, se cumple que

$$P_r(X \in A \text{ y } Y \in B) = P_r(X \in A) P_r(Y \in B) \quad (2.22)$$

En particular, para cualesquiera números reales x,y , se tiene que

$$P_r(X \leq x \text{ y } Y \leq y) = P_r(X \leq x) P_r(Y \leq y) \quad (2.23)$$

Dos variables aleatorias X,Y son independientes si y sólo si, para todos los números reales x,y , se cumple que

$$F(x,y) = F_1(x) F_2(y) \quad (2.24)$$

donde $F_1(x)$ y $F_2(y)$ son las distribuciones marginales.

2.2 CONVOLUCION

En general, la convolución se define como se muestra a continuación:

Definición. La convolución de dos funciones de densidad f_1 y f_2 es la función F tal que

$$\forall x \in \mathbb{R}^1 : F(x) = \int_{-\infty}^{\infty} F_1(x-y) dF_2(y) \quad (2.25)$$

donde F es representada como $F = f_1 * f_2$ (Kai, 1974).

Teorema. Sean X_1 y X_2 variables reales independientes con funciones marginales F_1 y F_2 , respectivamente. Entonces $X_1 + X_2$ tiene la función $F_1 * F_2$ (Kai, 1974).

Corolario. La operación binaria de convolución, representada por el asterisco $*$, es conmutativa y asociativa.

A modo de resumen se presentan las definiciones de la convolución para la suma, resta, división y combinación lineal de dos variables aleatorias (Fraser, 1976).

Convolución de la suma .

$$F_{X+Y}(t) = \int_{-\infty}^{\infty} f(t-s, s) ds \quad (2.26)$$

donde f es la función de densidad en \mathbb{R}^2 y F_{X+Y} es la función marginal de la suma de las variables X, Y ($t=x+y$; $s=y$).

Convolución de la diferencia .

$$F_{X-Y}(t) = \int_{-\infty}^{\infty} f(t+s, s) ds \quad (2.27)$$

donde f es la función de densidad en \mathbb{R}^2 y F_{X-Y} es la función marginal de la diferencia de las variables X, Y ($t=x-y$; $s=y$).

Convolución de la relación.

$$F_{X/Y}(t) = \int_0^{\infty} f(ts, s) ds \quad (2.28)$$

donde f está definida sobre \mathbb{R}^2 con $f=0$ para $y \leq 0$ y $F_{X/Y}$ es la función marginal de la relación X/Y ($t=x/y$; $s=y$).

Convolución de la combinación lineal.

$$F_T(t) = |a|^{-1} \int_{-\infty}^{\infty} f((t-bs)/a, s) ds \quad (2.29)$$

donde f está definida en R^2 y F_T es la función marginal de la combinación lineal $aX + bY$, con a y b constantes y $a \neq 0$ ($t=ax+by$; $y=s$).

2.3 FUNCION LOGISTICA BIVARIADA Y ESTADISTICOS BASICOS

El interés de esta tesis está enfocado hacia los valores extremos de los diferentes fenómenos o procesos tratados en los capítulos tres, cuatro, cinco y seis; por esta razón se propone utilizar la función de densidad de probabilidad logística bivariada. A continuación se muestra dicha función, así como las expresiones para el cálculo de los parámetros asociados.

La forma general del modelo logístico de las distribuciones bivariadas de valores extremos para máximos es (Raynal-Villaseñor, 1990):

$$F(x, y, m) = \exp\{-[(-\ln F(x))^m + (-\ln F(y))^m]^{1/m}\} \quad (2.30)$$

donde $F(x, y, m)$ es la distribución bivariada de las variables aleatorias x , y , m es el parámetro de asociación y $F(x)$, $F(y)$, las distribuciones marginales. Cuando las marginales tienen una distribución de valores extremos tipo I, el modelo general proporciona la distribución bivariada de valores extremos tipo II:

$$F(x, y, m) = \exp\{-[\exp(-m(x-u_1)/\alpha_1) + \exp(-m(y-u_2)/\alpha_2)]^{1/m}\} \quad (2.31)$$

con función de densidad

$$\begin{aligned}
f(x,y,m) &= (1/\alpha_1\alpha_2) \exp(-m(x-u_1)/\alpha_1) \exp(-m(y-u_2)/\alpha_2) \\
&[\exp(-m(x-u_1)/\alpha_1) + \exp(-m(y-u_2)/\alpha_2)]^{(1/m)-2} \{m-1 + [\exp(-m(x-u_1)/\alpha_1) \\
&+ \exp(-m(y-u_2)/\alpha_2)]^{1/m}\} \exp\{-[\exp(-m(x-u_1)/\alpha_1) \\
&+ \exp(-m(y-u_2)/\alpha_2)]^{1/m}\} \quad (2.32)
\end{aligned}$$

donde u_1, u_2 son los parámetros de ubicación, α_1, α_2 son los parámetros de escala y $m \geq 1$, parámetro de asociación; los subíndices 1,2 representan las variables aleatorias x, y , respectivamente.

La estimación de los parámetros de la función logística bivariada conjunta se lleva a cabo a través del método de la máxima verosimilitud; la expresión general para el caso bivariado es de la forma siguiente (Escalante y Raynal, 1994):

$$L(x,y,\theta) = \prod_{i=1}^{n_1} f(p_i, \theta_1) \prod_{i=1}^{n_3} f(x,y, \theta_3) \prod_{i=1}^{n_5} f(r_i, \theta_5) \quad (2.33)$$

donde L es la función de máxima verosimilitud; x, y son variables aleatorias con registro bivariado coincidiendo en la longitud del registro (periodo común); θ representa a los parámetros de la función; \prod representa el producto de las funciones f ; θ_1 es el parámetro univariado que corresponde a la longitud del registro anterior al periodo común; θ_3 es el parámetro bivariado que corresponde al periodo común; θ_5 es el parámetro univariado que corresponde a la longitud posterior al periodo común; n_1, n_3, n_5 longitudes de los registros antes, durante y después del periodo común, respectivamente; p es la variable con registro univariado antes del periodo común; r es la variable con registro univariado después del periodo común. Ver figura 2.1.

El caso general contempla la posibilidad de utilizar dos registros con longitudes diferentes. Si el caso a estudiar tiene registros con la misma longitud y las mismas fechas de inicio y fin, entonces la expresión se reduce a

$$L(x, y, \theta) = \prod_{i=1}^{n_3} f(x, y, \theta_3) \quad (2.34)$$

Para hacer manipulable computacionalmente la función general de máxima verosimilitud, se aplica el logaritmo natural a los dos miembros de la ecuación, de modo que la expresión queda

$$LL(x, y, \theta) = \sum_{i=1}^{n_1} \ln\{f(p_i, \theta_1)\} + \sum_{i=1}^{n_3} \ln\{f(x, y, \theta_3)\} + \sum_{i=1}^{n_5} \ln\{f(r_i, \theta_5)\} \quad (2.35)$$

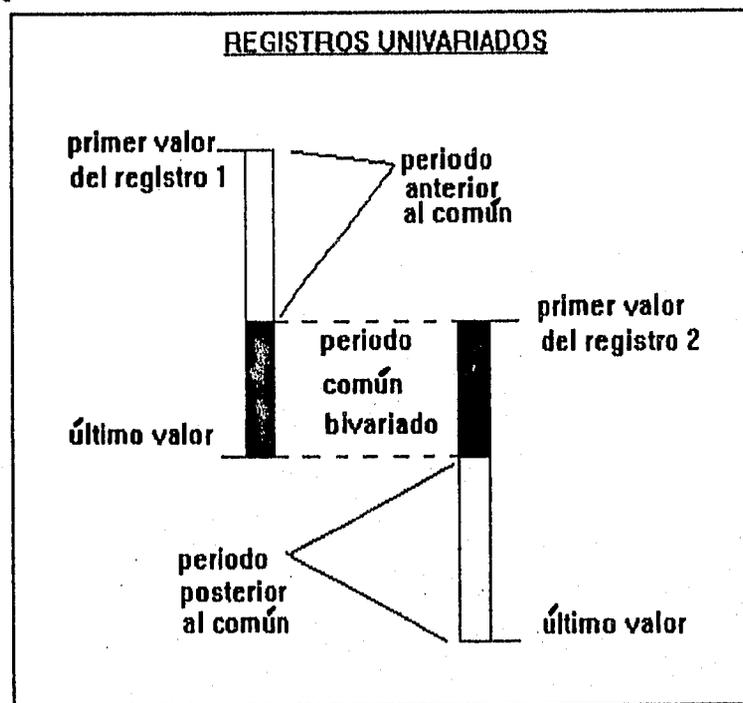


Figura 2.1 .Representación esquemática de los periodos anterior, común y posterior de dos registros de datos o muestras.

Sustituyendo la función logística en $f(p_i, \theta_1)$, $f(x, y, \theta_3)$ y $f(r_i, \theta_5)$, se llega a

$$\begin{aligned}
LL(x, y, u_1, \alpha_1, u_2, \alpha_2, m) = & -n_1 \ln\{\alpha_p\} + \sum_{i=1}^{n_1} [-(1-\beta_p(p_i-u_p)/\alpha_p)^{1/\beta_p} + \\
& \ln\{(1-\beta_p(p_i-u_p)/\alpha_p)^{1/\beta_p-1}\}] - n_3 (\ln\{\alpha_1\} + \ln\{\alpha_2\}) + \\
& \sum_{i=1}^{n_3} [\ln\{(1-\beta_1(x_i-u_1)/\alpha_1)^{m/\beta_1-1} + \ln\{(1-\beta_2(y_i-u_2)/\alpha_2)^{m/\beta_2-1}\}] + \\
& \ln\{(1-\beta_1(x_i-u_1)/\alpha_1)^{m/\beta_1} + (1-\beta_2(y_i-u_2)/\alpha_2)^{m/\beta_2}\}^{1/m-3} + \ln\{(1-m)(1-2m)\}] + \\
& ((1-\beta_1(x_i-u_1)/\alpha_1)^{m/\beta_1} + (1-\beta_2(y_i-u_2)/\alpha_2)^{m/\beta_2})^{2/m} + \\
& (3m-3) ((1-\beta_1(x_i-u_1)/\alpha_1)^{m/\beta_1} + (1-\beta_2(y_i-u_2)/\alpha_2)^{m/\beta_2})^{1/m} - n_5 \ln\{\alpha_r\} + \\
& \sum_{i=1}^{n_5} [-(1-\beta_r(r_i-u_r)/\alpha_r)^{1/\beta_r} + \ln\{(1-\beta_r(r_i-u_r)/\alpha_r)^{1/\beta_r-1}\}] \quad (2.36)
\end{aligned}$$

La solución de la expresión anterior se encuentra a través del esquema de optimación (maximación) denominado algoritmo de Rosenbrook, el cual ha sido implementado y adecuado computacionalmente por algunos autores (por ejemplo, Escalante, 1991).

A continuación se hace un resumen de los parámetros estadísticos básicos utilizados para caracterizar los registros o muestras de los datos a utilizar en las secciones siguientes (Ven Te Chow et. al., 1988):

Parámetros de tendencia central.

Media aritmética

Poblacional: $\mu = E(X) = \int_{-\infty}^{\infty} xf(x) dx \quad (2.37)$

Muestral: $x_m = (\sum_{i=1}^n x_i) (1/n) \quad (2.38)$

Mediana

Poblacional: x tal que $F(x) = 1/2 \quad (2.39)$

Muestral: el valor correspondiente al 50% de los datos
 (2.40)

Media Geométrica

Poblacional: antilogaritmo de $E(\log x) \quad (2.41)$

Muestral: $(\prod_{i=1}^n x_i)^{1/n} \quad (2.42)$

Parámetros de variación.

Varianza

$$\text{Poblacional: } \sigma^2 = E[(X-\mu)^2] \quad (2.43)$$

$$\text{Muestral: } s^2 = [1/(n-1)] \sum_{i=1}^n (x_i - x_m)^2 \quad (2.44)$$

Desviación estandar

$$\text{Poblacional: } \sigma = \{E[(X-\mu)^2]\}^{1/2} \quad (2.45)$$

$$\text{Muestral: } s = \{[1/(n-1)] \sum_{i=1}^n (x_i - x_m)^2\}^{1/2} \quad (2.46)$$

Coefficiente de variación

$$\text{Poblacional: } CV = \sigma/\mu \quad (2.47)$$

$$\text{Muestral: } CV = s/x_m \quad (2.48)$$

Parámetro de simetría.

Curtois

$$\text{Poblacional: } \gamma = E[(X - \mu)^3]/\sigma^3 \quad (2.49)$$

$$\text{Muestral: } C_s = n [\sum_{i=1}^n (x_i - x_m)^3 / \{(n-1)(n-2)s^3\}] \quad (2.50)$$

E representa la esperanza matemática o valor esperado, n es el número de datos de la muestra, X es la variable aleatoria, x es un elemento de dicha variable y F(x) es la función de distribución de probabilidad.

2.4 ANALISIS DE REGISTROS

El análisis de valores extremos de una variable hidrológica requiere del manejo de ciertos conceptos, los cuales se resumen a continuación (Ven Te Chow et. al., 1988).

Se dice que un evento extremo ha ocurrido cuando una variable aleatoria X es mayor o igual que algún valor base o de referencia definido x_T . Si en un registro cualquiera la variable aleatoria

rebasa el límite x_T un número de veces a lo largo del periodo de interés, entonces el intervalo de tiempo entre cada valor en el que se cumple que $X \geq x_T$ se llama intervalo de recurrencia y se representa con la letra τ . El valor promedio del intervalo de recurrencia entre eventos que igualan o exceden una magnitud específica es llamado el periodo de retorno del evento dado. En otras palabras, el periodo de retorno T del evento $X \geq x_T$ es el valor esperado del intervalo de recurrencia τ o el valor promedio de τ medido en un gran número de ocurrencias ($\tau_m = \text{número total de años/número de veces que } X \geq x_T \text{ rebasa el valor } x_T$).

La relación entre probabilidad y periodo de retorno está dada por la expresión

$$E(\tau) = T = 1/P_r \quad (2.51)$$

donde $E(\tau)$ es el valor esperado del intervalo de recurrencia, T es el periodo de retorno y P_r es la probabilidad. Esto puede ser escrito también como

$$P_r(X \geq x_T) = 1 - 1/T \quad (2.52)$$

Para los cálculos del periodo de retorno y de la probabilidad se utiliza la fórmula de Weibull, la cual se expresa como

$$P_r(X \geq x_m) = 1 - [m/(n+1)] \quad (2.53)$$

donde n es el número de datos de la muestra del registro, m es el rango y x_m el dato del registro asociado con el rango m .

Los métodos para analizar los registros de las variables hidrológicas se definen a continuación:

Serie de duración total. Es cuando se consideran todos los datos

disponibles en el registro.

Serie de duración parcial. Es cuando los datos de un registro se seleccionan de forma que su magnitud sea mayor que la de un valor base definido.

Serie de excedencia anual. Es una serie de duración parcial con el valor base seleccionado de forma que el número de años del registro es igual al número de valores en la serie.

Serie de valores extremos. Es cuando se seleccionan de un registro los valores máximos o mínimos en cada uno de los intervalos de tiempo igualmente espaciados.

Serie anual. Es una serie de valores extremos con un intervalo de tiempo de un año.

Cabe hacer mención que si se hace un análisis de serie de duración total, es conveniente utilizar la expresión

$$P_r(X \geq x_m) = 1 - [m/n] \quad (2.54)$$

con n refiriéndose al número de datos y no al número de años. Si se hace el análisis a través de una serie anual de máximos, se utiliza la expresión de Weibull.

3. SUMA DE VARIABLES ALEATORIAS:

CONFLUENCIA DE DOS RIOS

3.1 DESCRIPCION GENERAL

Conocer el gasto máximo que se puede presentar en un punto dado del sistema de drenaje en una cuenca es de vital importancia para el diseño y operación de obras hidráulicas como presas de almacenamiento, así como en el control de avenidas en zonas donde las inundaciones pueden provocar cuantiosas pérdidas humanas y materiales. Un caso particular que resulta común en la práctica es el de la confluencia de dos tributarios que están aforados aguas arriba en cualquier punto relativamente cercano; al no existir estaciones de medición en el punto donde se unen las descargas, es necesario contar con una herramienta estadística que permita predecir los gastos que se pueden esperar en dicho punto, para la condición de dependencia parcial entre las variables aleatorias.

Este problema fue resuelto por Raynal Villaseñor (Raynal, 1990) y es incluido en este trabajo para abarcar un mayor número de posibilidades de aplicación de las herramientas propuestas. Los datos utilizados (los mismos empleados por Raynal) son los gastos máximos mensuales aforados de los ríos "Bear " y "Dry creek", localizados en Estados Unidos. En la sección 3.3 se presentan los resultados.

3.2 METODOLOGIA PROPUESTA

Para obtener los valores de probabilidad conjunta bivariada en la confluencia de dos ríos, es necesario analizar por series anuales los registros de cada uno de los ríos. Una vez determinados los

valores máximos anuales se procede a sumarlos, obteniendo de esta forma la variable aleatoria T. La variable S corresponde a los valores de gasto de alguno de los dos ríos. Cada una de estas variables se arreglan de mayor a menor ya que en este caso particular se pretende trabajar con valores extremos máximos; el rango del valor máximo es uno y el mínimo corresponde al número total de valores del registro. Con las muestras arregladas y los valores de rango asignados, se procede a calcular la probabilidad de que las variables sean igualadas o excedidas para ciertos valores, utilizando la expresión de Weibull (Ven Te Chow et al, 1988 y Kite, 1988) como se muestra a continuación

$$\Pr(S \geq s) = 1 - [m/(n+1)] \quad (3.1)$$

y

$$\Pr(T \geq t) = 1 - [m/(n+1)] \quad (3.2)$$

donde Pr es la probabilidad de excedencia; m es el rango y n el número total de valores del registro o muestra (subcapítulos 2.1 y 2.4).

Una vez determinada la probabilidad de los valores de las muestras, se calculan los estadísticos muestrales media aritmética, media geométrica, varianza, desviación estándar, coeficiente de variación y coeficiente de curtosis; así mismo, se calculan los parámetros de la función conjunta bivariada, es decir, los parámetros de ubicación (u_1, u_2) , los parámetros de escala (α_1, α_2) , y el parámetro de asociación, m (subcapítulo 2.3).

Con los parámetros de ubicación, escala y asociación estimados e integrando la función conjunta bivariada, es posible calcular el valor de la probabilidad de ocurrencia de un gasto correspondiente a la confluencia de dos ríos. Las funciones de densidad y distribución para este caso particular se muestran a continuación

$$\begin{aligned}
f(t-s, s) &= (1/\alpha_1\alpha_2) \exp(-m(t-s-u_1)/\alpha_1) \exp(-m(s-u_2)/\alpha_2) \\
&\quad [\exp(-m(t-s-u_1)/\alpha_1) + \exp(-m(s-u_2)/\alpha_2)]^{(1/m)-2} \\
&\quad \{m-1 + [\exp(-m(t-s-u_1)/\alpha_1) + \exp(-m(s-u_2)/\alpha_2)]^{1/m}\} \\
&\quad \exp\{-[\exp(-m(t-s-u_1)/\alpha_1) + \exp(-m(s-u_2)/\alpha_2)]^{1/m}\} \quad (3.3)
\end{aligned}$$

$$F(t) = \int_{-\infty}^t \int_{-\infty}^{\infty} f(t-s, s) ds dt \quad (3.4)$$

donde $f(t-s, s)$ es la función de densidad, t el valor de la suma de los gastos de cada río ($t=x+y$), s el valor correspondiente a los gastos de uno de los ríos ($s=y$) y $F(t)$ es la función de distribución.

Los valores de probabilidad estimados por la expresión de Weibull sirven para estimar los valores de la variable t , suma de los gastos de los dos ríos, utilizando el modelo propuesto. El procedimiento para obtener t requiere de la integración de la función de densidad por algún método numérico (en este caso el método es el de la regla de Simpson simple bidimensional, ver Burden y Fairies, 1985) y el uso de un esquema de solución de ecuaciones por un método de bisección modificado (Nakamura, 1993) para obtener el valor de t que hace que la probabilidad calculada se aproxime a la probabilidad estimada hasta un cierto orden de magnitud que depende de la precisión deseada. Los valores resultantes se comparan con los datos y se determina que tan adecuado es el modelo para la predicción.

Finalmente, se utiliza el modelo para calcular los valores máximos esperados de la suma de los gastos máximos en la confluencia de dos ríos, definiendo primero los periodos de retorno de interés y los valores de probabilidad asociados.

3.3 APLICACION DE LA METODOLOGIA

Los valores máximos anuales para los ríos "Bear" y "Dry creek" se obtuvieron de una muestra mensual de veinte años (de 1951 a 1970); con esta información se forma la muestra que representa a la variable aleatoria T, donde cada valor de la variable representa la suma de los gastos máximos anuales. Los valores máximos anuales y la suma se presentan en la tabla 3.1.

AÑO	Gastos Río Bear (m ³ /s)	Gastos Río Dry creek (m ³ /s)	Suma de los gastos (m ³ /s)
1951	368.03	164.62	532.65
1952	325.57	164.14	489.71
1953	114.94	15.57	130.51
1954	123.43	28.59	152.02
1955	63.98	14.44	78.42
1956	302.92	170.99	473.91
1957	147.78	14.64	162.42
1958	164.48	35.1	199.58
1959	111.82	55.2	147.21
1960	532.23	77.29	609.52
1961	22.65	14.24	36.89
1962	44.16	19.82	63.98
1963	255.36	82.1	337.46
1964	29.73	5.38	35.11
1965	331.23	96.68	427.91
1966	43.03	15.85	58.88
1967	345.38	125.7	471.08
1968	11.92	3.34	15.26
1969	475.61	229.88	705.49
1970	413.33	114.66	527.99

Tabla 3.1. Gastos máximos anuales de los ríos Bear y Dry creek.

Los estadísticos de T y S se muestran a continuación:

	Variable T (gasto suma), en m ³ /s	Variable S (gasto Dry creek), en m ³ /s
Media aritmética:	283.7908	211.3790
Media geométrica:	179.4373	135.2167
Varianza:	50267.15	27020.11
Desviación estándar:	224.2034	164.3780
Coefficiente variación:	0.7900	0.7776
Curtosis:	0.0193	0.0223

Los valores de la suma y del río Bear se arreglan de mayor a menor y se les asigna el rango m a cada uno de ellos; con este valor sustituido en la expresión de Weibull y el número de datos, que en este caso es $n=20$, se calcula la probabilidad asignada a cada valor. Este valor de probabilidad es fundamental para encontrar el gasto máximo de la confluencia a través del modelo logístico bivariado; se utiliza un esquema de búsqueda de raíces basado en el método de bisección, de modo que se propone un intervalo bastante amplio para iniciar dicho proceso para encontrar el valor t que hace que el resultado de la integral doble (probabilidad estimada por el modelo) coincida hasta por un orden de magnitud de 10^{-3} con el valor de la probabilidad asignada. El intervalo de búsqueda va de 0 a 5000 m³/s, ya que así se asegura que los valores de probabilidad estarán comprendidos entre 0 y 1. Mas detalles del programa de computadora utilizado se encuentran en el apéndice A.

Antes de calcular los valores del gasto en la confluencia, utilizando el modelo logístico bivariado, se determinan los estimadores de los parámetros de asociación, escala y ubicación, haciendo uso del algoritmo desarrollado por Escalante, 1991. En este caso particular se utilizan los valores utilizado por Raynal, 1990, mostrados a continuación

Parámetro de asociación:

$$m = 2.6976$$

Parámetros de ubicación:

(río Bear)	(río Dry creek)
$u_1 = 134.32 \text{ m}^3/\text{s}$	$u_2 = 48.62 \text{ m}^3/\text{s}$

Parámetros de escala:

$$\alpha_1 = 116.71 \text{ m}^3/\text{s} \quad \alpha_2 = 50.58 \text{ m}^3/\text{s}$$

Con estos valores se procede a calcular el gasto máximo esperado en la confluencia y los resultados se comparan con los datos. Los resultados se muestran en la tabla 3.2 y en la figura 3.1.

Como se puede observar en la tabla 3.2, en la primera columna está la probabilidad asignada a cada uno de los datos de la suma (variable T), segunda columna, los cuales son comparados con los resultados obtenidos del modelo, tercera columna; la comparación se llevó a cabo a través del error relativo, cuarta columna, usando la expresión siguiente:

$$\text{error} = \frac{\|\text{dato} - \text{valor modelo}\|}{\|\text{dato}\|} \quad (3.5)$$

En la figura 3.1 se muestra la representación gráfica de los datos y de los resultados del modelo, es decir, los gastos máximos en la confluencia de dos ríos, contra la probabilidad asignada.

El siguiente paso fue calcular el gasto en la confluencia para diferentes valores de periodo de retorno. Los valores propuestos fueron 1.5, 2, 3, 5, 10, 50, 100, 1000 y 10000 años. Los resultados obtenidos se muestran en la tabla 3.3 y en la figura 3.2.

En la tabla 3.3 se muestra la probabilidad, primera columna, correspondiente al valor del periodo de retorno, segunda columna;

en la tercera columna se encuentran los valores de gastos máximos esperados en la confluencia. En la figura 3.2 se puede ver el comportamiento del modelo, a través de la representación gráfica del gasto máximo en la confluencia contra el periodo de retorno propuesto.

Probabili- dad	Suma de gastos, en m ³ /s		Error relativo
	datos	modelo	
0.048	15.26	7.324443	0.52
0.095	35.11	43.99795	0.2531
0.143	36.89	75.06	1
0.191	58.88	103.841	0.7636
0.238	63.98	130.21	1
0.286	78.42	154.0995	0.9651
0.333	130.51	176.1103	0.3494
0.381	152.02	196.7287	0.2941
0.429	162.42	216.4698	0.3328
0.476	167.02	236.0582	0.4134
0.524	199.58	255.9137	0.2823
0.571	337.46	277.5812	0.1774
0.619	427.91	302.2432	0.2937
0.667	471.08	330.2149	0.299
0.714	473.91	363.6216	0.2327
0.762	489.71	400.0519	0.1831
0.81	527.99	440.5639	0.1656
0.857	532.65	488.7435	0.0824
0.905	609.52	560.4598	0.0805
0.952	705.49	679.6308	0.0367

Tabla 3.2. Resultados del modelo logístico bivariado.

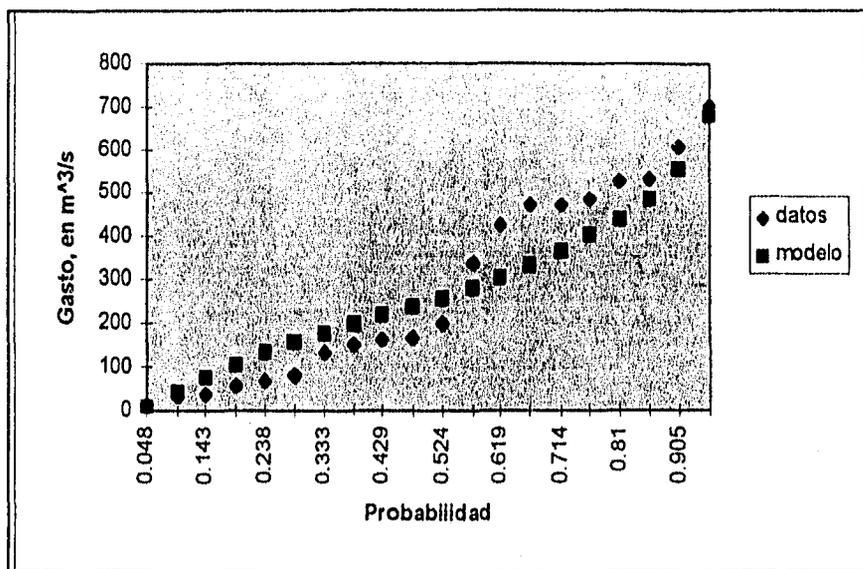


Figura 3.1. Representación gráfica de la suma de gastos (datos y valores estimados).

Probabili- dad	Periodo de retorno en años	Gasto estimado, en m ³ /s
0.333	1.5	176.1103
0.5	2	246.1103
0.667	3	330.2049
0.8	5	430.9128
0.9	10	551.6097
0.98	50	828.2512
0.99	100	921.6349
0.999	1000	1158.451
0.9999	10000	1164.555

Tabla 3.3. Valores estimados de gasto máximo en la confluencia de dos ríos a partir de valores de periodo de retorno propuestos.

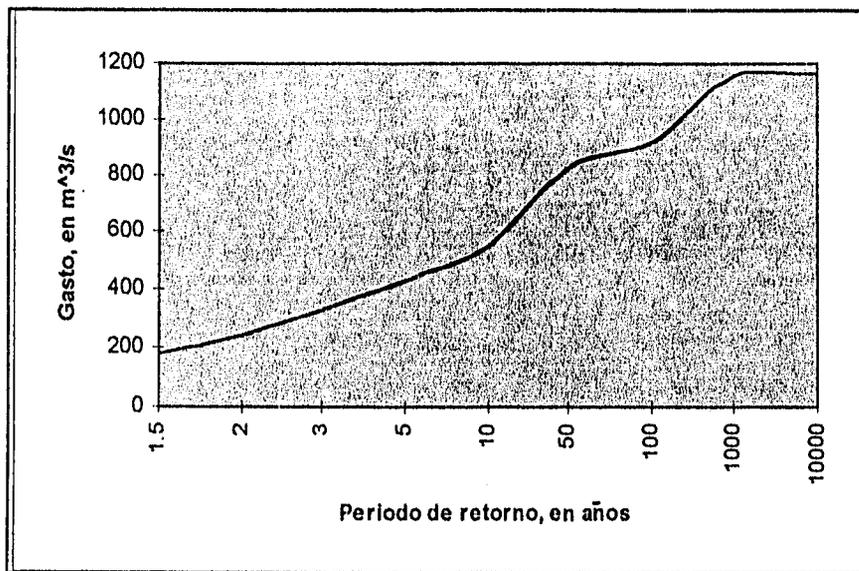


Figura 3.2. Representación gráfica gasto máximo estimado-periodo de retorno.

Como se puede observar, el modelo propuesto se ajusta satisfactoriamente a los datos, por lo que es altamente recomendable su uso para fines de predicción en este caso particular.

Se desarrolló el programa de computadora "suma.exe", escrito en el lenguaje de programación C++ que se ejecuta en un ambiente "windows" versión 3.1, el cual sirvió de apoyo en la mayor parte de los pasos seguidos de acuerdo a la metodología propuesta en el subcapítulo 3.2.

4. RESTA DE VARIABLES ALEATORIAS:

DEFICIT DE LAS SEQUIAS

4.1 DESCRIPCION GENERAL

En el norte del país grandes extensiones de terreno, en las cuales se llevan a cabo actividades económicas como agricultura y ganadería, se ven afectadas periódicamente por los efectos de las sequías. Es, por ello, importante contar con una herramienta predictiva que ofrezca estimaciones de la severidad o déficit de las sequías, así como de su duración. A continuación se hace un breve repaso del estado del conocimiento actual de este tipo de fenómenos. En el subcapítulo 4.2 se presenta la metodología propuesta para predecir los valores máximos esperados de déficit y duración de la sequía; en el subcapítulo 4.3 se aplica la metodología a los datos de la estación "Huites", ubicada en el río Fuerte, en el norte del estado de Sinaloa.

De acuerdo con los resultados obtenidos de modelos simplificados del sistema atmósfera-océano-continente aplicados en latitudes medias septentrionales, se desarrolla durante el año una alternancia de centros de baja y alta presión atmosférica que mantiene en balance la convergencia y la divergencia de humedad atmosférica en la superficie del terreno; cuando el balance es anulado por la prolongada duración de un centro de alta presión, se dan las condiciones para que se presenten las etapas iniciales del fenómeno de la sequía (Bravar y Kavvas, 1991).

Un centro de baja presión atmosférica provoca la convergencia horizontal del aire al nivel de terreno, de modo que el vapor de agua del océano y del sistema terreno-vegetación es acumulado en una zona particular y posteriormente forzado a moverse ascendentemente, abasteciendo de humedad a la atmósfera; esta

humedad retornará a la superficie y no necesariamente a la misma posición geográfica, debido básicamente a que los vientos en latitudes medias provienen del oeste. Si se centra la atención en el subsistema terreno-vegetación, la forma en la que la atmósfera cercana a la superficie obtiene la humedad es por la relación entre intensidad y duración de la evapotranspiración: puede ser que la intensidad sea baja pero con una duración prolongada, o por otro lado, que la intensidad sea alta pero la duración pequeña debido a la presencia de precipitación intermitente. Si se alcanza un balance entre intensidad y duración, se dice que se ha alcanzado la condición normal; por ejemplo, si no hay días con lluvia, la humedad de las partes someras de los suelos disminuye y la humedad relativa atmosférica se reduce, de modo que se establecen las condiciones para la sequía.

Un centro de alta presión atmosférica, también conocido como sistema anticiclónico, provoca la divergencia horizontal en la superficie del terreno, por lo que la humedad atmosférica recolectada en la superficie por la evapotranspiración es dispersada y alejada de la región geográfica por los vientos.

La presencia de las denominadas ondas de Rossby, es decir, aquellas ondas de perturbación atmosférica a nivel planetario en las cuales la velocidad longitudinal asociada al movimiento atmosférico es de 10 a 15 m/s con longitudes de onda de 3000 a 4000 km, extendiéndose desde las latitudes bajas hasta la región polar, es probablemente la causa por la cual el balance de la alternancia de centros de alta y baja presión es anulado y por lo tanto se da origen al fenómeno de la sequía (cabe mencionar que las ondas de Rossby parecen estar ligadas al fenómeno de "el Niño" y no se sabe exactamente cual es su origen), Bravar y Kavvas, 1991. Esto explica el origen de las sequías pero no la duración prolongada que se verifica mucho tiempo después de la presencia de los centros de alta presión; la razón para esto se encuentra básicamente en el subsistema terreno-vegetación. Después que se ha presentado un

centro de alta presión por un periodo de tiempo largo, la capacidad de la superficie del terreno y de la vegetación para abastecer de vapor de agua a la atmósfera, se ve reducida (también el posible aporte de los acuíferos a la atmósfera es reducido, debido al descenso del nivel freático). Aun cuando se presente un centro de baja presión, el abasto de vapor de agua a través de las corrientes de aire ascendentes será insuficiente para que se forme la precipitación; para que la precipitación ocurra, deben existir cuatro factores: 1) Enfriamiento del aire hasta un valor de temperatura cercano al punto de rocío; 2) condensación en núcleos para formar las gotas o los cristales de hielo; 3) crecimiento de las gotas y cristales hasta alcanzar las dimensiones propias para que se desplacen a la superficie en forma de gotas de lluvia, nieve o granizo; 4) Importación de suficiente vapor de agua para sostener el proceso (Dingman, 1994). La única forma en la que se puede terminar la sequía, es por el transporte de humedad atmosférica desde regiones cercanas a la zona afectada; sin embargo, la cantidad de humedad requerida para restablecer las condiciones normales debe ser considerable (Bravar y Kavvas, 1991).

Es importante distinguir el fenómeno de las sequías y el de la desertificación; las sequías son fenómenos reversibles, es decir, una vez que se cumplen ciertas condiciones, sus efectos terminan. La desertificación es un fenómeno no reversible que puede tener diversas causas: una de ellas está relacionada con la remoción de la cubierta vegetal en grandes extensiones de terreno. El agua de lluvia que era almacenada por el sistema vegetación-suelo, es transformada en escurrimiento directo y flujo subterráneo, provocando así una disminución en la disponibilidad de vapor de agua en la superficie del terreno y por lo tanto en la humedad relativa, lo que a su vez provoca una disminución en la evapotranspiración, que aunado a la remoción de vegetación, da como resultado un continuo abatimiento en la disponibilidad de humedad en la superficie. Otra de las causas tiene que ver con el incremento del albedo (porcentaje de la radiación incidente

recibida que es reflejada por una superficie); cuando se reduce la cubierta vegetal se aumenta el valor del albedo, presentándose una disminución en la radiación solar que llega a la superficie y hace que la evapotranspiración disminuya, implicando con ello la disminución del calor latente y del flujo de humedad a la atmósfera. Además, el aire tiende a ser enfriado por radiación y a mantener el equilibrio térmico, por lo tanto, descenderá y la precipitación será inhibida.

4.2 METODOLOGIA PROPUESTA

Los efectos de las sequías en una región determinada pueden ser estudiados a través de metodologías basadas en técnicas hidrológicas, como es, por ejemplo, el análisis de escurrimientos superficiales en ríos, propuesto por Zelenhasic y colaboradores; la técnica fue diseñada para ser empleada en el ámbito de la hidrología estocástica, sin embargo, en este trabajo se toma como base y se hacen las modificaciones pertinentes para aplicarla en la predicción.

El objetivo de la metodología es encontrar los valores máximos de déficit y duración de la sequía que se pueden esperar para diferentes niveles de probabilidad o periodos de retorno, utilizando el modelo logístico bivariado para valores extremos, por medio de la convolución de la resta de dos variables aleatorias.

La información requerida consiste únicamente de un registro completo de escurrimientos diarios en una sección de una corriente de importancia, siendo deseable que su longitud sea lo mas extensa posible.

El primer paso es determinar el valor del gasto de referencia o umbral, es decir, aquel gasto por medio del cual se determina que parte del hidrograma está asociado a las sequías. Para calcular

dicho gasto se hace uso del hidrograma (representación gráfica de gastos diarios contra tiempo) en toda su extensión temporal (todos los años del registro), estimándose el área bajo la curva. Una vez definido este valor de área, se propone un valor constante de gasto mayor a cero, de modo que la recta que lo representa divida al área bajo la curva del hidrograma en dos partes complementarias cuya suma sea el 100% del valor del área calculado inicialmente. De acuerdo con análisis estadísticos (Zelenhasic y Salvai, 1987), el gasto de referencia es aquel valor que divide el área bajo la curva de forma que la parte complementaria superior corresponda al 90 o 95% del valor de área calculada inicialmente. El valor del área total bajo la curva puede ser estimada en forma gráfica (subdividiendo el área total en áreas mas pequeñas o haciendo uso del planímetro) o utilizando un esquema numérico de integración (regla del trapecio, o fórmula de Simpson o cuadratura gaussiana). Para definir el 90 o 95% de esta área, sólo se estima el área abajo del valor propuesto de gasto de referencia y por prueba y error se encuentra el valor correspondiente al 10 o 5%, respectivamente.

Con el gasto de referencia estimado, se procede a construir el conjunto de variables aleatorias necesarias para el modelo: El volumen de referencia, el volumen de descarga y el déficit de la sequía. Las variables están contenidas en las expresiones siguientes (Zelenhasic y Salvai, 1987):

$$D_v = Q_r T_v - \int_{t_i}^{t_f} Q_t dt \quad \text{con } Q_t < Q_r \quad (4.1)$$

donde D_v es el déficit de la v -ésima sequía, Q_r es el gasto de referencia, T_v es la duración de la sequía (número de días consecutivos en el cual el gasto en cualquier instante está por debajo del gasto de referencia), Q_t es el gasto en cualquier instante o de descarga, t_i es el instante en el que inicia la sequía y t_f es el instante en el que finaliza.

El primer término del miembro de la derecha en la expresión para el déficit D_v (ecuación 4.1), es el volumen de referencia, definido como

$$V1(T_v) = Q_r * T_v \quad (4.2)$$

con Q_r constante. El segundo término del miembro de la derecha es el volumen de descarga, definido como

$$V2(Q_t) = \int_{t_i}^{t_f} Q_t dt \quad (4.3)$$

con la integral representando el área bajo la curva de los gastos menores al gasto de referencia. El déficit por lo tanto sólo es la resta de $V1-V2$.

La duración de las sequía y el área bajo la curva de los gastos menores al de referencia se obtienen del hidrograma, como se muestra en el esquema de la figura 4.1.

Existen dos casos en los cuales las variables $V1$ y $V2$ tienen que ser tratadas de manera especial. Uno de estos casos es cuando los volúmenes son irrelevantes comparados con los volúmenes medios o extremos; el otro caso es cuando la secuencia de dos o más eventos de sequía (gastos instantáneos por debajo del gasto de referencia Q_r) están separados por un evento de no sequía (valores de gasto por arriba de Q_r) o también denominado inter-evento. En ambos casos se requiere de un valor base por medio del cual sea posible decidir si el valor es irrelevante y se desecha del análisis, y si la secuencia de diferentes valores es mutuamente dependiente y puede ser considerada como un sólo valor. Para el primer caso, el valor base es el valor máximo del déficit, $D_{v \max}$, considerando todo el registro de escurrimientos; así, los valores de déficit en los que

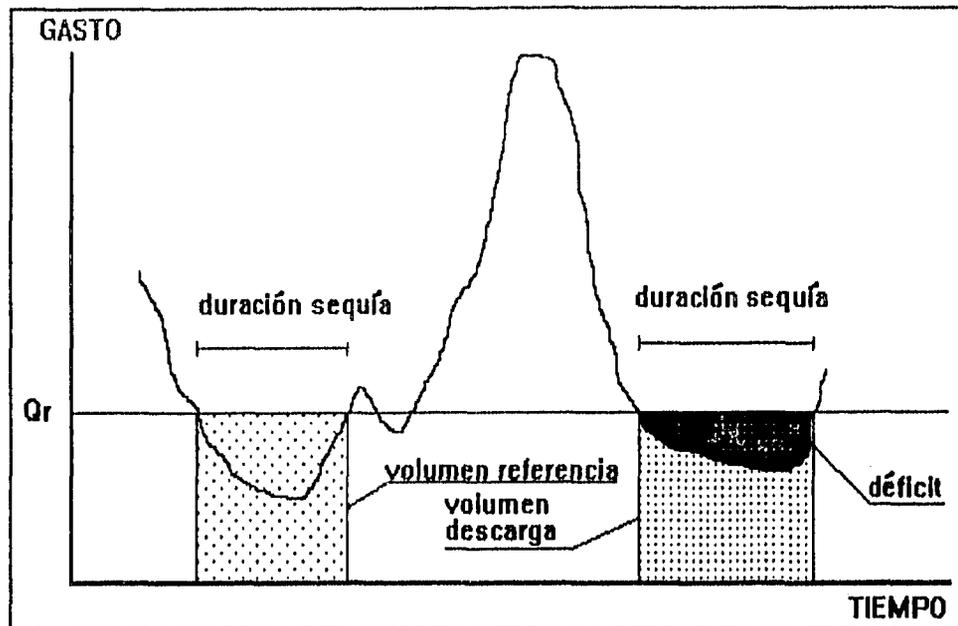


Figura 4.1. Representación esquemática de las variables en el análisis de sequías.

se cumpla que $D_i < 0.005 D_{v \max}$, serán desechados del análisis. Si se utilizan gastos diarios, el valor base para determinar la dependencia entre dos eventos es la duración del inter-evento: si la duración es menor o igual a 6 días, entonces los dos eventos de sequía son mutuamente dependientes y se consideran como un sólo valor de la variable aleatoria. Cabe hacer mención que en la metodología propuesta por Zelenhasic y Salvai, la forma de tratar la dependencia de eventos de sequía, requiere de un mayor refinamiento, como lo mencionan los autores.

En el caso que se presente una sequía que inicia al final del año y se prolonga hasta el año siguiente, el evento será asignado al año en el cual se presenta un mayor número de días; así por ejemplo, si al final de un año cualquiera se tiene que la sequía abarca 40 días y su continuación es de 60 días en el siguiente año,

entonces el evento será asignado a este último.

Una vez determinados los valores que corresponden a las variables aleatorias volumen de referencia, V_1 , y volumen de descarga, V_2 , se seleccionan los valores máximos por año en cada caso; con esta información se definen las variables aleatorias requeridas por el modelo logístico bivariado utilizando la convolución de la resta: la variable T será el déficit, es decir, volumen de referencia menos volumen de descarga y S será el volumen de descarga.

Con los valores de las variables aleatorias T y S definidas para cada año, se procede al arreglo de mayor a menor y a la asignación del valor de rango m , empezando con $m=1$ para el valor numérico máximo y $m=n$ para el mínimo de cada muestra, donde n representa el número total de datos. El paso siguiente es calcular la probabilidad de excedencia asociada a cada valor t y s (valores de las variables T y S , respectivamente) haciendo uso de la expresión de Weibull.

Para hacer uso del modelo propuesto y ver que tan adecuado es el ajuste a los datos disponibles, se calculan los estimadores de los parámetros de ubicación (u_1, u_2), escala (α_1, α_2) y asociación (m) del modelo logístico bivariado, como se mencionó en el capítulo 2. Los subíndices indican la muestra de la variable utilizada para el cálculo; el número 1 se refiere al volumen de referencia y el 2 al volumen de descarga. Se propone utilizar el programa de computadora desarrollado por Escalante, 1991, para encontrar los valores óptimos de los cinco estimadores. Al mismo tiempo, se calculan los estadísticos muestrales tales como la media aritmética, media geométrica, la varianza, la desviación estandar, el coeficiente de variación y el coeficiente de curtosis, de los cuales se puede obtener información complementaria importante sobre las variables T y S .

Para calcular el valor de T, déficit de sequía, a través del modelo logístico bivariado, se sustituyen los estimadores de los parámetros calculados y se utiliza el valor de probabilidad asignada como valor de referencia. La forma en la que se llevan a cabo los cálculos es la siguiente: se integra numéricamente la función de densidad logística bivariada

$$\begin{aligned}
 f(t+s,s) = & (1/\alpha_1\alpha_2) \exp(-m(t+s-u_1)/\alpha_1) \exp(-m(s-u_2)/\alpha_2) \\
 & [\exp(-m(t+s-u_1)/\alpha_1) + \exp(-m(s-u_2)/\alpha_2)]^{(1/m)-2} \\
 & \{m-1 + [\exp(-m(t+s-u_1)/\alpha_1) + \exp(-m(s-u_2)/\alpha_2)]^{1/m}\} \\
 & \exp\{-[\exp(-m(t+s-u_1)/\alpha_1) + \exp(-m(s-u_2)/\alpha_2)]^{1/m}\} \quad (4.4)
 \end{aligned}$$

obteniéndose así la función de distribución

$$F(t) = \int_{-\infty}^t \int_{-\infty}^{\infty} f(t+s,s) ds dt \quad (4.5)$$

la cual expresa el valor de probabilidad calculada. La integración se lleva a cabo usando el esquema de Simpson simple para dos variables (Burden y Faires, 1985), proponiendo un valor inicial de t (déficit de sequía), y una vez calculada la probabilidad, se compara con la probabilidad asignada considerada, de modo que en caso que sean diferentes se busque otro valor de t tal que ambas probabilidades tiendan a coincidir hasta un cierto orden de magnitud o tolerancia, que en este caso es de 10^{-3} . El esquema de búsqueda del valor t está basado en el método de bisección (Nakamura, 1993). Más detalles sobre ambos esquemas se presentan en el apéndice A.

Los valores calculados del déficit de sequía, se comparan con los datos a través del error relativo, expresado por

$$\text{error} = \|\text{dato} - \text{valor modelo}\| / \|\text{datos}\| \quad (4.6)$$

De esta comparación es posible definir que tan adecuado es el

ajuste del modelo a los datos, así como el intervalo de probabilidades donde puede aplicarse con certeza.

Con la información anterior, se proponen diferentes valores de periodo de retorno para los cuales se desea calcular el valor de déficit de sequía máximo esperado, así como la duración de la sequía máxima esperada en un año. En este caso, se requiere de una relación adecuada entre la duración y el déficit de la sequía. Se propone utilizar diferentes tipos de correlaciones y decidir en base a características estadísticas cual es la mejor opción.

Se desarrolló el programa de computadora "restavar.exe", el cual sirve de apoyo en la aplicación de la metodología propuesta. Este programa fue escrito en el lenguaje de programación C++ y se ejecuta en un ambiente "windows", versión 3.11. Comentarios sobre el programa se encuentran en el apéndice A.

4.3 APLICACION DE LA METODOLOGIA

Dado que en el norte del país se presentan periódicamente sequías que afectan en forma considerable las actividades económicas de extensas regiones, se propuso aplicar la metodología expuesta en el subcapítulo anterior a la estación hidrométrica "Huites", localizada en el tramo del río Fuerte, aguas abajo de la confluencia de los ríos Chinipa, Urique y Septentrión, en el norte del estado de Sinaloa; en la figura 4.2 se presenta un esquema de la ubicación de la estación.

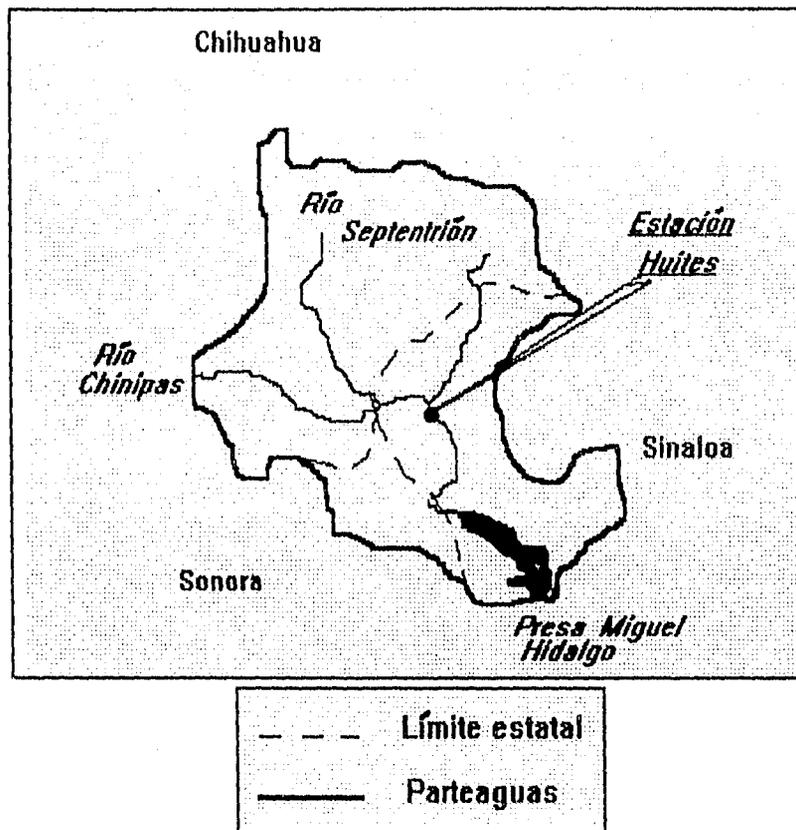


Figura 4.2. Localización de la zona de interés.

La región donde se localiza la estación hidrométrica presenta un tipo climatológico BW, que de acuerdo con la clasificación de Koeppen, es un clima seco desértico, caracterizado por intensidades de evaporación mayores a las de precipitación, régimen térmico extremo, gran oscilación térmica diaria y anual, lámina de precipitación anual menor a los 200 mm y en caso de existir vegetación, está compuesta básicamente por xerofitas (plantas rastreras en zonas arenosas), Ayllon y Chávez, 1992.

Como primer paso se determinó el gasto de referencia, utilizando el hidrograma de gastos diarios para el periodo 1942 - 1992 (en el

apéndice B se muestra el hidrograma de los datos máximos mensuales, incluyendo los datos representados en forma tabular); el valor obtenido para un porcentaje del 91% de la totalidad de los gastos diarios fue de 15.75 m³/s. Una vez trazado el gasto de referencia en el hidrograma (ver apéndice B.3, donde se representan los hidrogramas de gastos bajos y la recta correspondiente al gasto de referencia), se obtuvieron los valores de duración de la sequía, con lo cual fue posible determinar el volumen de referencia, que es una de las variables aleatorias a considerar; la otra variable aleatoria corresponde al volumen de escurrimiento, que a su vez es designado como la variable S para el modelo logístico bivariado. La variable T corresponde al déficit de la sequía, es decir, a la diferencia entre el volumen de referencia y el volumen de escurrimiento. En la Tabla 4.1 se muestran los datos de duración, volumen de referencia, volumen de escurrimiento y déficit de la sequía.

Los parámetros estadísticos para las variables utilizadas en el modelo logístico bivariado se resumen a continuación:

	Variable T (Déficit ×10 ⁶ , en m ³)	Variable S (Volumen descarga ×10 ⁶ , en m ³)
Media aritmética:	68.8776	63.0098
Media geométrica:	57.5041	53.9593
Varianza:	1596.89	890.33
Desviación estándar:	39.9611	29.8384
Coefficiente variación:	0.5803	0.4736
Curtosis:	0.0221	0.0171

Con los valores de las muestras de las variables T y S determinadas, se procede a estimar la probabilidad asignada, asociando el rango a las muestras arregladas de mayor a menor y

empleando la expresión de Weibull con $n=51$.

AÑO	Duración, en días	Duración, en M s	Volumen referencia, en M m3	Volumen descarga, en M m3	Déficit, en M m3
1942	94	8.1216	127.51	67.39	60.12
1943	117	10.1088	158.71	85.54	73.17
1944	39	3.3696	52.9	33.7	19.21
1945	129	11.1456	174.99	72.58	102.41
1946	139	12.0096	188.55	83.81	104.74
1947	72	6.2208	97.67	47.52	50.15
1948	70	6.048	94.95	56.16	38.79
1949	47	4.0608	63.75	40.61	23.15
1950	80	6.912	108.52	69.12	39.39
1951	159	13.74	215.68	103.68	112
1952	80	6.192	108.52	49.25	59.27
1953	148	12.7872	200.76	77.76	123
1954	131	11.3184	177.7	74.3	103.39
1955	113	9.7632	153.28	70.85	82.43
1956	126	10.89	170.92	81.22	89.7
1957	90	7.776	122.08	33.7	88.39
1958	46	3.9744	62.4	39.74	22.65
1959	103	8.8992	139.72	52.7	87.01
1960	92	7.9488	124.8	53.57	71.23
1961	91	7.8624	123.44	50.98	72.46
1962	98	8.4672	132.94	64.8	68.14
1963	112	9.6768	151.93	44.06	107.86
1964	70	6.048	94.95	60.48	34.47
1965	92	7.9488	124.8	37.15	87.64
1966	66	5.7024	89.53	57.02	32.5

Nota: M expresa millones

Tabla 4.1. Datos de duración, volumen de referencia, volumen de escurrimiento y déficit de la sequía.

AÑO	Duración, en días	Duración, en M s	Volumen referencia, en M m3	Volumen descarga, en M m3	Déficit, en M m3
1967	162	13.9968	219.75	123.55	96.2
1968	volumen	referencia	menor al	volumen	descarga
1969	97	8.3808	131.58	83.81	47.77
1970	68	5.8752	92.24	44.06	48.18
1971	155	13.392	210.25	71.71	138.54
1972	89	7.6896	120.73	59.62	61.11
1973	66	5.7024	89.53	57.02	32.5
1974	163	14.0832	221.11	3.11	218
1975	109	9.4176	147.86	94.18	53.68
1976	198	17.1024	268.58	171.07	97.51
1977	56	4.8384	75.96	38.02	37.95
1978	99	8.5536	134.29	54.43	79.86
1979	113	9.7632	153.29	97.63	55.65
1980	192	16.5888	260.44	114.05	146.4
1981	39	3.3696	52.9	33.7	19.21
1982	141	12.1824	191.26	94.18	97.09
1983	41	3.5424	55.62	35.42	20.19
1984	97	8.3808	131.58	83.81	47.77
1985	29	2.5056	39.34	25.06	14.28
1986	67	5.7888	90.88	57.88	33
1987	85	7.344	115.3	73.44	41.86
1988	139	12.0096	188.55	87.26	101.29
1989	82	7.0848	111.23	5.01	106.22
1990	64	5.5296	86.81	43.2	43.61
1991	85	7.344	115.3	73.44	41.86
1992	21	1.8144	28.49	18.14	10.34

Nota: M expresa millones.

Tabla 4.1. (Continuación)

El paso siguiente consiste en calcular los estimadores de los parámetros del modelo logístico bivariado, empleando el algoritmo desarrollado por Escalante, 1991. Los valores obtenidos son:

Parámetro de asociación:

$$m = 2.40587$$

Parámetros de ubicación:

(Volumen de referencia)	(Volumen de escurrimiento)
$u_1 = 102.859 \times 10^6 \text{ m}^3$	$u_2 = 49.966 \times 10^6 \text{ m}^3$

Parámetros de escala:

$$\alpha_1 = 50.292 \times 10^6 \text{ m}^3 \quad \alpha_2 = 26.391 \times 10^6 \text{ m}^3$$

Una vez conocidos los estimadores y los valores de probabilidad asignada, se procede al cálculo de los valores de déficit de la sequía a través del modelo propuesto. Los resultados son comparados con los datos por medio del error relativo, definido en el capítulo 3, y se resumen en la tabla 4.2 y figura 4.3.

Como se puede observar de la tabla 4.2 y de la figura 4.3, el modelo se ajusta muy bien en casi toda la extensión de la probabilidad asignada, excepto para el último valor de probabilidad (0.981), donde se tiene un error relativo muy grande (valor de 1 en la tabla 4.2). Esto puede deberse a que el modelo seleccionado no se ajusta adecuadamente en la parte extrema superior del rango de probabilidad, siendo necesario proponer otro modelo.

Los valores de periodo de retorno propuestos para predecir el déficit de sequía máximo esperado son 1.5, 2, 3, 5, 7, 10, 50, 100, 1000 y 10000 años. Por lo comentado en el párrafo anterior, la predicción para periodos de retorno mayores a 50 años no tiene sentido, ya que el modelo aquí propuesto no se ajusta adecuadamente para valores de probabilidad mayores o iguales a 0.981. Los resultados obtenidos se muestran en la tabla 4.3 y en la figura 4.4. El comportamiento de las curvas mostradas en la figura 4.4 se debe en gran parte al graficador utilizado.

Ya que se conocen los valores de déficit de sequía máximos esperados, es posible determinar la duración máxima esperada por año. Para llevar a cabo esto se requiere una relación entre las dos muestras, lo cual se consigue buscando una correlación adecuada.

Probabilidad	Déficit, en datos	Mm ³ modelo	Error relativo
0.019	10.35	0.005002	1
0.038	14.28	7.192807	0.4963
0.058	19.2	13.45365	0.2993
0.077	19.2	18.20483	0.0518
0.096	20.2	22.10725	0.0944
0.115	22.66	25.49087	0.1249
0.135	23.14	28.50637	0.2319
0.154	32.51	31.25103	0.0387
0.173	32.51	33.77253	0.0388
0.192	33	36.15289	0.0995
0.212	34.47	38.41118	0.1143
0.231	37.94	40.58555	0.0697
0.25	38.79	42.66073	0.0998
0.269	39.4	44.66153	0.1335
0.288	41.86	46.61274	0.1135
0.308	41.86	48.55822	0.16
0.327	43.61	50.40644	0.1558
0.346	47.77	52.26228	0.094
0.365	47.77	54.09332	0.1324
0.385	48.18	55.90911	0.1604
0.404	50.15	57.72871	0.1511
0.423	53.68	59.52542	0.1089
0.442	55.66	61.33358	0.1019
0.462	59.27	63.14937	0.0655
0.481	60.12	64.98041	0.0808
0.5	61.11	66.83435	0.0937

Nota: M expresa millones.

Tabla 4.2. Resultados del déficit de sequías máximo esperado.

Probabili- dad	Déficit, en datos	Mm ³ modelo	Error relativo
0.519	68.14	68.71498	0.0084
0.538	68.87	70.59944	0.0251
0.558	71.23	72.52966	0.0182
0.577	72.46	74.52092	0.0284
0.596	73.17	76.60374	0.0475
0.615	79.86	78.64078	0.0153
0.635	82.43	80.79988	0.0198
0.654	87.02	83.02003	0.046
0.673	87.65	85.33935	0.0264
0.692	88.38	87.83415	0.0062
0.712	89.7	90.58072	0.0098
0.731	96.2	93.03737	0.0329
0.75	97.08	95.93653	0.0118
0.769	97.52	99.06456	0.0158
0.788	101.29	102.3452	0.0104
0.808	102.41	105.7174	0.0323
0.827	103.4	109.9593	0.0634
0.846	104.73	114.3385	0.0917
0.865	106.22	119.3587	0.1237
0.885	107.87	125.0654	0.1594
0.904	112	132.0539	0.1791
0.923	123	140.9345	0.1458
0.942	138.54	153.9349	0.1111
0.962	146.39	176.5178	0.2058
0.981	218	568.1174	1

Tabla 4.2. (Continuación).

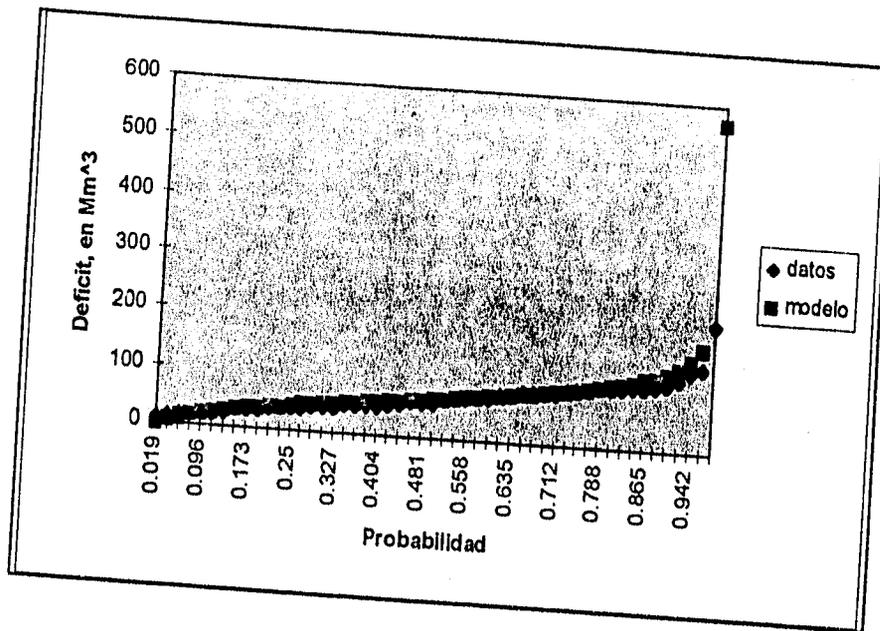


Figura 4.3. Representación de los datos y valores calculados para el déficit de sequías.

Los resultados para las diferentes relaciones propuestas se muestran a continuación:

Regresión lineal (duración = $a + b \cdot \text{déficit}$), con

$$a = -12.40364885,$$

$$b = 0.8359231521 \quad \text{y}$$

$$\text{coeficiente de correlación} = 0.8502$$

Regresión logarítmica (duración = $a + b \cdot \ln \text{déficit}$), con

$$a = -239.7552934,$$

$$b = 68.93406996 \quad \text{y}$$

$$\text{coeficiente de correlación} = 0.8169$$

Regresión exponencial (duración = $a \cdot b^{\text{déficit}}$), con

$$a = 14.77833617,$$

$$b = 1.014035306 \quad y$$

$$\text{coeficiente de correlación} = 0.8798$$

Regresión de potencia (duración = $a \cdot \text{déficit}^b$), con

$$a = 0.1968767904,$$

$$b = 1.267215411 \quad y$$

$$\text{coeficiente de correlación} = 0.9320$$

Notas: 1) En todos los casos, duración en días y déficit en 10^6 m^3 .

2) Coeficiente de correlación calculado por la expresión

$$r = \frac{E[xy] - E[x] E[y]}{(\sigma_x \sigma_y)}$$

donde r es el coeficiente de correlación; E es la esperanza matemática; x, y las variables aleatorias; σ_x, σ_y desviaciones estándar de las variables respectivas.

Si el coeficiente de correlación es un parámetro que brinda la información suficiente para decidir cual de las correlaciones es la más adecuada, entonces se selecciona la regresión de potencia.

Si cada valor de déficit estimado se sustituye en la regresión de potencia, entonces se obtienen las duraciones esperadas para cada periodo de retorno propuesto. Los resultados se muestran en la representación gráfica de la figura 4.4 y en la tabla 4.3.

Probabili- dad	Periodo de retorno, en años	Déficit, en Mm ³	Duración, en días
0.333	1.5	50.90807	29
0.5	2	66.83435	40
0.667	3	84.56116	54
0.8	5	104.4814	71
0.857	7	117.1004	82
0.9	10	130.5585	94
0.98	50	564.4553	604
0.99	100	736.5736	846
0.999	1000	749.2688	865
0.9999	10000	956.6653	1179

Tabla 4.3. Valores calculados de déficit y duración de sequía.

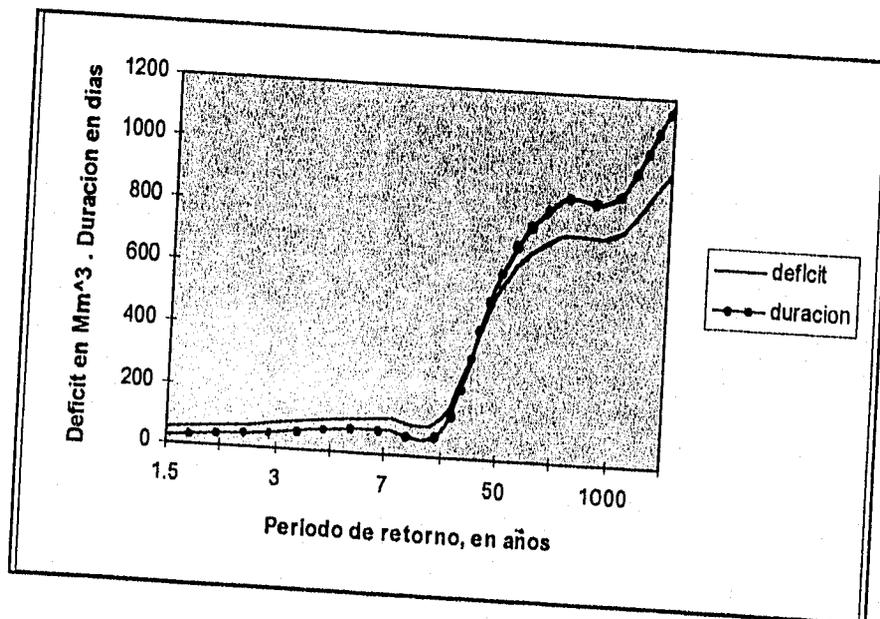


Figura 4.4. Representación gráfica de los valores estimados del déficit y duración de sequía contra periodo de retorno.

5. DIVISION DE VARIABLES ALEATORIAS:

ANALISIS DE HIDROGRAMAS DE ESCURRIMIENTOS EXTREMOS

5.1 DESCRIPCION GENERAL

Algunas de las aplicaciones mas relevantes del análisis de hidrogramas de escurrimientos extremos se encuentran en el control de avenidas e inundaciones (Bertoli et al, 1994), así como en la revisión y diseño de las obras hidráulicas utilizadas para el abastecimiento de agua potable, irrigación o generación de energía.

En este capítulo se propone el análisis de hidrogramas a través de la división del valor de gasto máximo entre el volumen de escurrimiento directo asociado, pretendiendo obtener una relación simple entre las dos variables, de modo que al proponer un valor de gasto se pueda estimar el volumen descargado de una avenida particular, siendo válido el caso inverso, es decir, proponiendo un valor de escurrimiento, encontrar el valor de gasto máximo.

La metodología propuesta para obtener el valor extremo de la relación gasto máximo-volumen escurrido, se muestra en la sección 5.2; la aplicación al caso de los escurrimientos registrados en la estación Huites, Sinaloa, en el periodo que va de 1941 a 1990 se encuentra en la sección 5.3.

5.2 METODOLOGIA PROPUESTA

Para calcular el gasto máximo que se puede esperar de un volumen escurrido, o al contrario, conocer el volumen máximo esperado a partir de un valor de gasto máximo definido, se requiere como paso

inicial contar con información de escurrimientos en una corriente de importancia; estos gastos pueden ser horarios o diarios. Una vez seleccionada la longitud del registro a estudiar, se procede a separar el escurrimiento directo del gasto base, para lo cual puede ser utilizada cualquiera de las técnicas propuestas en Viessman et al., 1989, Linsley et al., 1988, Linsley et al. 1972, Chow et al., 1988, Dingman, 1993 o Bren et al., 1987.

Una vez transformados los gastos totales en escurrimiento directo, se definen los valores de gasto máximo, Q_{max} , y del volumen escurrido V_{ed} , es decir, el área abajo de la curva del hidrograma, para cada avenida registrada.

Con la información disponible, se definen las variables que serán utilizadas en el modelo logístico bivariado: la variable aleatoria T corresponde a la relación Q_{max}/V_{ed} y la variable aleatoria S al volumen de escurrimiento V_{ed} .

Como información complementaria para las nuevas variables T y S , se calculan los estadísticos como la media aritmética, la media geométrica, la varianza, la desviación estandar, el coeficiente de variación y el coeficiente de curtosis (subcapítulo 2.3) para cada una de las muestras.

Los valores de cada una de las variables aleatorias, se arreglan de mayor a menor de forma que el valor máximo sea el primer valor y el valor mínimo el último; es decir, se asigna el valor de rango $m=1$ para el máximo valor de cada una de las variables, el dos al siguiente y así sucesivamente hasta asignar el máximo valor del rango al número de datos, $m=n$ (subcapítulo 2.4). Una vez arreglados los valores de T y S , se calcula para cada par de elementos la probabilidad asignada a través de la expresión de Weibull (ver ecuación 2.53).

Para utilizar el modelo logístico bivariado, se tienen que calcular

los estimadores de los parámetros de ubicación, escala y asociación, para cada una de las muestras originales; para los parámetros con el subíndice 1, se utilizan los valores del gasto máximo y para los que tienen asignado el 2, le corresponden los valores de volumen escurrido.

Con los valores conocidos de probabilidad asignada y de los estimadores de los parámetros del modelo, se calcula la probabilidad por medio de la integración numérica de la función de densidad del modelo logístico bivariado

$$\begin{aligned}
 f(t*s,s) &= (1/\alpha_1\alpha_2) \exp(-m(t*s-u_1)/\alpha_1) \exp(-m(s-u_2)/\alpha_2) \\
 &\quad [\exp(-m(t*s-u_1)/\alpha_1) + \exp(-m(s-u_2)/\alpha_2)]^{(1/m)-2} \\
 &\quad \{m-1 + [\exp(-m(t*s-u_1)/\alpha_1) + \exp(-m(s-u_2)/\alpha_2)]^{1/m}\} \\
 &\quad \exp\{-[\exp(-m(t*s-u_1)/\alpha_1) + \exp(-m(s-u_2)/\alpha_2)]^{1/m}\} \quad (5.1)
 \end{aligned}$$

obteniéndose así la función de distribución o probabilidad

$$F(t) = \int_{-\infty}^t \int_{-\infty}^{\infty} s \{f(t*s,s)\} ds dt \quad (5.2)$$

para el valor t . El valor de probabilidad calculado se compara con el de probabilidad asignada, y si son diferentes de acuerdo a un valor de tolerancia definida, se busca un nuevo valor de t , de modo que el próximo valor de probabilidad calculada tienda al de probabilidad asignada. La integración se lleva a cabo usando el esquema bidimensional de la regla de Simpson simple y la búsqueda del valor de t (cuyo valor inicial es asignado de forma que la probabilidad calculada sea 1) se lleva a cabo por un esquema modificado de bisección. Mas detalles sobre los esquemas se presentan en el apéndice A.

Con los valores calculados por el modelo, se define el ajuste que tiene sobre los datos y de esta forma se pueden determinar los valores de periodo de retorno para predecir la relación máxima

esperada. Esta relación servirá para encontrar el gasto máximo esperado a partir de un volumen de escurrimiento máximo, o el volumen de escurrimiento máximo esperado a partir del gasto máximo.

Fue desarrollado el programa de computadora "div_var.exe", que lleva a cabo la mayor parte de las tareas necesarias para aplicar la metodología. Este programa fue escrito en el lenguaje de programación C++ y funciona en ambiente Windows 3.11. Detalles del programa y del listado son mostrados en el apéndice A.

5.3 APLICACION DE LA METODOLOGIA

La estación hidrométrica Huites, como se menciona en el subcapítulo 4.3, está localizada en el estado de Sinaloa y cuenta con un extenso registro de escurrimientos. Estos escurrimientos fueron convertidos en directos, separando el flujo base del escurrimiento total, empleando las técnicas propuestas en Viessmann et al, 1989, Chow et al, 1988, Dingman, 1993 y Bren et al, 1987. Una vez restado el valor del flujo base a cada uno de los gastos diarios, se procede a calcular el área debajo de la curva trazada para los gastos de las avenidas mas importantes que se presentaron en el año; con esto se obtiene el volumen de escurrimiento y el gasto máximo de la avenida. Una vez determinados los valores máximos de cada una de las variables, se calcula la variable de la relación gasto máximo entre volumen escurrido, es decir, $T = Q_{max}/Volumen$. La variable simple S es el volumen de escurrimiento directo. En la tabla 5.1 se muestran los valores de cada una de las variables.

Los parámetros estadísticos de cada una de las variables se muestran a continuación:

	Variable T (Relación, en 10^{-6} l/s)	Variable S (Volumen, en 10^6 m ³)
Media aritmética:	4.9480	751.5965
Media geométrica:	5.6259	521.2656
Varianza:	3.5142	570743.53
Desviación estándar:	1.8746	755.4757
Coefficiente de variación:	0.3789	1.0052
Curtosis:	0.0318	0.0476

AÑO	Qmáx, Gasto máximo, en m ³ /s	Ved, Volumen esc. dir., en M m ³	R=Qmáx /Ved, en 10^{-6} l/s
1941	2084.8	455.779	4.574
1942	2531	666.895	3.795
1943	14376	2886.819	4.98
1944	2580	930.848	2.772
1945	1499	486.625	3.08
1946	1164.8	301.557	3.863
1947	1127.3	295.895	3.81
1948	3215	483.744	6.646
1949	10000	3843.441	2.582
1950	3229.2	535.079	6.15
1951	677	121.566	5.569
1952	1266	300.625	4.211
1953	1025	331.502	3.092
1954	954.8	77.237	12.362
1955	4780.3	837.579	5.707
1956	695.7	107.702	6.459
1957	593	245.627	2.414
1958	3010	550.028	5.472
1959	1908	337.053	5.661
1960	15000	2395.953	6.261
1961	1396.3	595.438	2.345
1962	1620	310.512	5.217
1963	2702	866.744	3.117
1964	1319	324.517	4.065
1965	1944	363.128	5.353

Tabla 5.1. Datos de gasto máximo, Qmáx, volumen de escurrimiento directo, Ved y relación Qmáx/Volumen.

AÑO	Qmáx, Gasto máximo, en m3/s	Ved, Volumen esc. dir., en M m3	R=Qmáx / Ved, en 10 ⁻⁶ l/s
1966	2420	494.948	4.889
1967	2505.8	606.884	4.129
1968	1534.3	206.798	7.419
1969	1508	312.877	4.82
1970	1558	380.084	4.099
1971	2200	788.517	2.79
1972	2225	460.274	4.834
1973	7960	1329.317	5.988
1974	3790	807.709	4.692
1975	1080	225.405	4.791
1976	3240	685.851	4.724
1977	1135	200.659	5.656
1978	4790	875.674	5.47
1979	6860	1130.679	6.067
1980	1496	344.059	4.348
1981	4828	874.841	5.519
1982	2450	448.419	5.464
1983	8279.2	1667.443	4.966
1984	5580	1429.244	3.904
1985	3585	817.297	4.386
1986	2081.5	1689.452	1.232
1987	1025.4	191.025	5.368
1988	3145	476.523	6.6
1989	1496.5	139.308	10.742
1990	11704	2314.64	5.057

Nota: M expresa millones.

Tabla 5.1. (Continuación).

Los estimadores para los parámetros de escala, ubicación y asociación se calcularon a través del algoritmo desarrollado por Escalante, 1991; los resultados son los siguientes:

Parámetro de asociación:

$$m = 2.6150$$

Parámetros de ubicación:

$$\begin{array}{ll} \text{(Qmáx)} & \text{(Ved)} \\ u_1 = 2217.17 \text{ m}^3/\text{s} & u_2 = 467.57 * 10^6 \text{ m}^3 \end{array}$$

Parámetros de escala:

$$\alpha_1 = 1724.16 \text{ m}^3/\text{s} \quad \alpha_2 = 347.56 * 10^6 \text{ m}^3$$

Después de arreglar las muestras de las variables T y S de mayor a menor y de asignarle el rango m a cada par de valores, se calcula la probabilidad asignada utilizando la expresión de Weibull, con $n=50$. Con esta probabilidad se calcula el valor de relación máxima para cada uno de los cincuenta datos, utilizando el modelo logístico bivariado; finalmente los resultados se comparan con los valores de la variable T y la representación gráfica se muestra en la figura 5.1.

El modelo propuesto no se ajusta al comportamiento de los datos para valores de probabilidad mayores de 90%, como puede observarse en la figura 5.1, por lo que no es posible llevar a cabo predicciones para valores de periodo de retorno mayores a 10 años. Esto se puede deber a que el modelo logístico no representa en forma adecuada la relación gasto máximo entre volumen para la estación "Huites". Para que el modelo pueda ser utilizado en la predicción de valores máximos de relación gasto máximo entre volumen escurrido, debe de proponerse algún otro modelo bivariado. En caso de que se desee probar otros modelos, es necesario cambiar la expresión del modelo seleccionado para ser remplazado por la

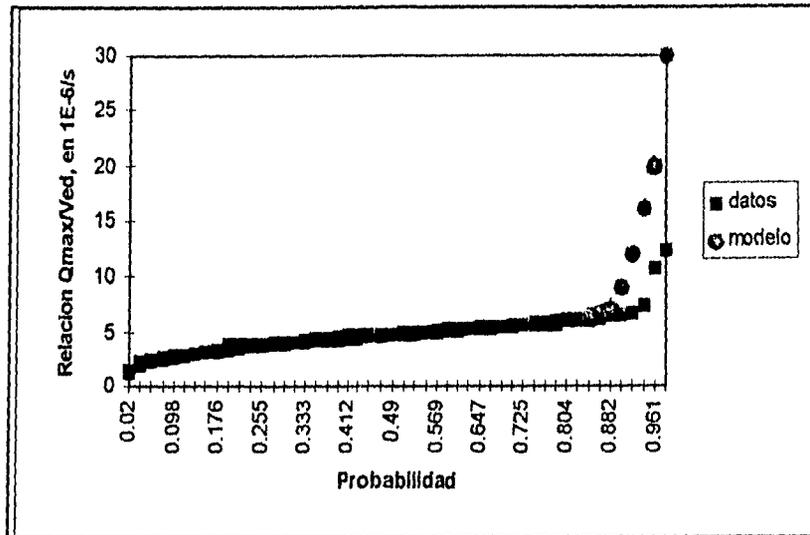


Figura 5.1. Representación gráfica de los valores de la relación Q_{\max}/V_{ed} y los calculados por el modelo.

expresión del modelo logístico bivariado. Esto se hace editando en el código fuente del programa "integ_d.cpp" la rutina "funcion_d()". El programa mencionado pertenece al proyecto "div_var.prj", comentado en el apéndice A.

6. COMBINACION LINEAL DE VARIABLES ALEATORIAS: RELACION LLUVIA ESCURRIMIENTO

6.1 DESCRIPCION GENERAL

El proceso lluvia-escorrimento es de los más importantes en la Hidrología, ya que tiene que ver con el manejo de presas de almacenamiento, control de avenidas y diseño de obras de protección en las márgenes de las secciones que presentan mayor posibilidad de daño en las zonas ribereñas, entre otras muchas aplicaciones. Hasta la fecha, se han desarrollado diversos modelos que están encaminados a predecir o pronosticar el gasto máximo que escurre a través de una corriente, tomando en cuenta que estos escurrimientos son generados por un conjunto de factores que varían en tiempo y espacio. Uno de estos modelos es el propuesto por Hewlett (Bren et al, 1987), el cual permite estimar el gasto máximo del escurrimiento directo, considerando variables independientes la lámina de precipitación y el gasto antecedente, el cual representa el grado de humedad del suelo que tiene la cuenca. Cabe hacer mención que este modelo fue propuesto para tormentas específicas de duraciones que van de los 15 a los 120 minutos. En este capítulo, la aplicación del modelo, que es una combinación lineal en el dominio de los logaritmos, se llevó a cabo en una subcuenca del estado de Guanajuato, donde la información disponible se encuentra en intervalos de tiempo de meses, es decir, lámina de precipitación acumulada en el mes y gastos mensuales (m^3/mes).

En el subcapítulo 6.2 se presenta la metodología general para encontrar el valor máximo esperado de la variable que representa la combinación lineal de dos variables aleatorias independientes, utilizando el modelo logístico bivariado. En el subcapítulo 6.3 se muestran los resultados de la aplicación de la metodología a la subcuenca del río De la Laja, en el tramo comprendido entre la

estación hidrométrica "Perico" y la estación "La Begoña"; además de la estación hidrométrica, se cuenta con las estaciones climatológicas "Mandujano", "Neutla", "Cortázar" y "Celaya".

6.2 METODOLOGIA PROPUESTA

Antes de describir la metodología, se presentan algunas definiciones importantes (Bren et al., 1987; Viessmann et al., 1989; Linsley et al., 1988; Chow et al. 1988; Linsley et al., 1972; Dingman, 1993):

- La duración del gasto de escurrimiento es el intervalo de tiempo asociado al escurrimiento directo.
- El escurrimiento directo es aquella parte del hidrograma que queda por arriba de la línea de separación con el gasto base (ver figura 6.1).
- El flujo antecedente es aquel gasto en el cual se percibe inicialmente el efecto del escurrimiento causado por la precipitación en las partes altas de la cuenca; este flujo está asociado al contenido de humedad en ese momento en la cuenca (ver figura 6.1).
- La lámina del escurrimiento directo es la representación del volumen escurrido dividido por el área de la cuenca.
- El gasto pico es el valor máximo del escurrimiento directo.
- La precipitación efectiva es aquella que se transforma en escurrimiento directo, después que ha sido eliminado el efecto de la evaporación, infiltración y pérdidas.

Dado que el modelo a utilizar (modelo de Hewlett) fue diseñado para

tormentas que van de 15 a 120 minutos de duración, es de esperarse que si se utiliza fuera de ese intervalo de duración (diario o mensual), la validez del modelo no se asegura y los resultados obtenidos pueden ser de poca utilidad para la predicción. Entonces, además de buscar un método por el cual sea posible predecir el valor de gasto máximo a partir de valores propuestos de precipitación y gasto antecedente, se verificará si el modelo puede representar alguna utilidad para la predicción, cuando la información disponible corresponde a un intervalo de tiempo mensual, situación muy común en nuestro país.

Independientemente del intervalo de tiempo de la información que se pretenda utilizar, la metodología general será válida en cualquier caso. Así, el primer paso es la adquisición de datos de escurrimiento y precipitación en una cuenca determinada. Una vez que se tiene la información, se procede a determinar cada una de las variables aleatorias implicadas en el modelo de Hewlett; esto es, el gasto máximo (o pico) y el gasto antecedente, los cuales serán obtenidos a partir del hidrograma de escurrimientos (ver figura 6.1), y la precipitación, obtenida de registros pluviométricos o pluviográficos, según sea el caso. Si el intervalo es horario o diario, la precipitación tiene que ser expresada como lámina de lluvia efectiva y los gastos máximo y antecedente en láminas de escurrimiento (valor de gasto dividido por el área de la cuenca que drena el río en cuestión en el intervalo de tiempo considerado); para el cálculo de estas variables referirse a cualquiera de las técnicas presentadas, por ejemplo, en Viessman et al., 1989, Linsley et al., 1988, Linsley et al. 1972, Chow et al., 1988, Dingman, 1993. Si el intervalo es mensual, entonces se propone utilizar la lámina de precipitación total y los gastos como unidad de volumen dividido por el tiempo; esto debido a la pérdida de información que se tiene al considerar valores acumulados mensuales, que no permiten establecer una relación causa-efecto clara, como en el caso horario o diario.

Como el modelo de Hewlett es una representación linealizada por medio de logaritmos, es necesario aplicar una transformación a los datos de las variables lámina de lluvia, gasto antecedente y gasto máximo. A continuación se muestra el modelo:

$$\ln(Q_{\max}) = a \cdot \ln(h_p) + b \cdot \ln(Q_{\text{ant}}) + c \quad (6.1)$$

donde \ln es el logaritmo natural, h_p es la precipitación efectiva, Q_{ant} es el flujo antecedente, Q_p es el gasto pico y a , b , c son constantes (Bren et al., 1987).

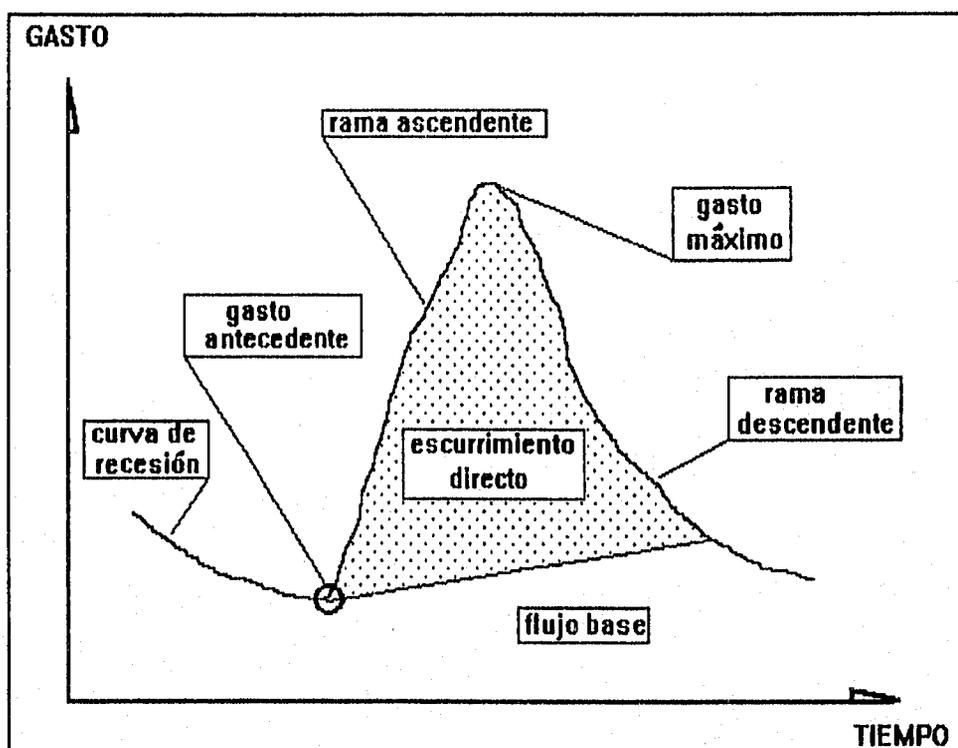


Figura 6.1. Representación esquemática de los gastos máximo o pico y antecedente.

Las constantes del modelo de Hewlett se determinan por medio de una regresión lineal múltiple. El programa de computadora "r_l_mult.exe" fue desarrollado para transformar las muestras al dominio del logaritmo y calcular las constantes. Para más detalles del programa, ver apéndice A.

Con las constantes conocidas, se determina la variable aleatoria T , la cual representa solo los términos $a \cdot \ln(h_p) + b \cdot \ln(Q_{ant})$, y la variable S , la cual es $\ln(Q_{ant})$. Como se puede observar, existirán dos valores de $\ln(Q_{máx})$, uno procede de los datos y el otro del modelo de Hewlett.

Como información complementaria, se calculan los parámetros estadísticos media aritmética, media geométrica, varianza, desviación estándar, coeficiente de variación y coeficiente de curtosis para las muestras de cada una de las variables aleatorias T y S .

Los valores de las variables aleatorias T y S se arreglan de mayor a menor, de modo que se asigne el rango $m=1$ al valor más grande y el rango $m=n$ al mínimo de las muestras (n es el número de datos). Utilizando la expresión de Weibull (ecuación 2.53), se determina la probabilidad asociada a cada par de elementos de las variables.

Para utilizar el modelo logístico bivariado en el cálculo de los valores máximos esperados de $\ln(Q_{máx})$, se requiere de la estimación de los parámetros de ubicación, escala y asociación, a través del algoritmo desarrollado por Escalante, 1991. Los parámetros con subíndice 1 se calculan de la muestra de $\ln(h_p)$ y los que tienen el subíndice 2 de la muestra $\ln(Q_{ant})$.

Con los estimadores de los parámetros y la probabilidad asignada conocidos, se calcula la probabilidad por medio de la integración numérica de la función de densidad logística bivariada

$$\begin{aligned}
f((t-b*s)/a, s) &= (1/\alpha_1\alpha_2) \exp(-m((t-b*s)/a-u_1)/\alpha_1) \\
&\exp(-m(s-u_2)/\alpha_2) [\exp(-m((t-b*s)/a-u_1)/\alpha_1) + \exp(-m(s-u_2)/\alpha_2)]^{(1/m)-2} \\
&\{m-1 + [\exp(-m((t-b*s)/a-u_1)/\alpha_1) + \exp(-m(s-u_2)/\alpha_2)]^{1/m}\} \\
&\exp\{-[\exp(-m((t-b*s)/a-u_1)/\alpha_1) + \exp(-m(s-u_2)/\alpha_2)]^{1/m}\} \quad (6.2)
\end{aligned}$$

obteniéndose así la función de distribución de probabilidad

$$F(t) = \int_{-\infty}^t \int_{-\infty}^{\infty} 1/|a| * f((t-b*s)/a, s) ds dt \quad (6.3)$$

Este valor calculado de probabilidad se compara con la probabilidad asignada, y si son diferentes, de acuerdo con una tolerancia definida de antemano, se busca un nuevo valor de t, de modo que la probabilidad calculada tienda al valor de probabilidad asignada. La integración del modelo logístico bivariado se lleva a cabo por el esquema bidimensional de la regla de Simpson simple y la búsqueda del valor de t ($\ln(Q_{max})$) se hace por medio de un esquema modificado de bisección, iniciando el proceso con un valor t que asegure una probabilidad de 1. Detalles de los dos esquemas se presentan en el apéndice A.

Los resultados del modelo logístico se comparan con los valores estimados por el modelo de Hewlett y con los datos, definiendo así que tan adecuado es el ajuste y los límites de validez. De esta forma se procede al paso final, el cual es la predicción de los gastos máximos esperados para los diferentes periodos de retorno propuestos. Lo adecuado del ajuste se puede caracterizar a través del error relativo definido en el capítulo 3.

6.3 APLICACION DE LA METODOLOGIA

La metodología expuesta en el subcapítulo anterior fue aplicada a la subcuenca del río "De La Laja", comprendida entre las estaciones "Perico" y "La Begoña", en el estado de Guanajuato; se cuenta con

información de precipitación acumulada mensual en las estaciones "Mandujano", "Neutla", "Cortázar" y "Celaya" y de volúmenes de escurrimiento mensuales en la estación "Perico". En la figura 6.2 se muestra el esquema de localización de la cuenca, incluyendo a las estaciones climatológicas e hidrométricas.

La información de precipitación en las estaciones "Mandujano", "Neutla" y "Cortázar", está limitada en longitud de registro a 9, 10 y 13 años, respectivamente. La estación "Celaya" tiene un periodo mayor, del cual sólo se emplearon 21 años, de 1946 a 1966, debido a que la información disponible en la estación "Perico" solo es útil en ese periodo (parece ser que hubo un cambio en el régimen del río aguas arriba de la estación, presentándose escurrimientos más uniformes durante todo el año a partir de 1967, aproximadamente). La relativa pequeña longitud del registro de precipitación de las estaciones mencionadas, aunado a la cantidad de datos estimados, obligó a considerar como estación representativa de la cuenca a la estación "Celaya". Para tener alguna certeza en cuanto la selección de la estación, se procedió a correlacionar estos registros con los de las tres estaciones restantes, suponiendo que una fuerte correlación asegure que la cuenca está sometida al mismo tipo de precipitaciones en toda su extensión; en caso contrario, sería de esperar que existen regimenes de precipitación muy diferentes en tiempo y espacio, y que la representatividad en una de las estaciones no sería una suposición mínimamente aceptable. Los resultados de la correlación son los siguientes:

Estación:	Celaya/Mandujano	Celaya/Neutla	Celaya/Cortázar
Periodo:	(1964-1972)	(1963-1972)	(1960-1972)
Media			
aritmética:	49.65mm 67.39mm	49.55mm 55.43mm	50.94mm 54.95mm
Desviación			
estándar:	60.52mm 83.57mm	61.57mm 68.33mm	62.87mm 64.23mm
Coefficiente			
correlación:	0.91	0.87	0.92

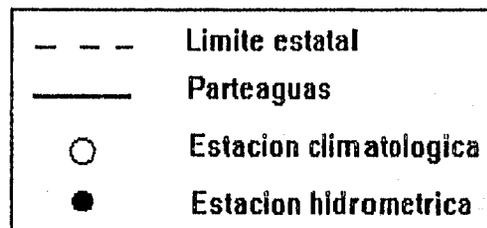
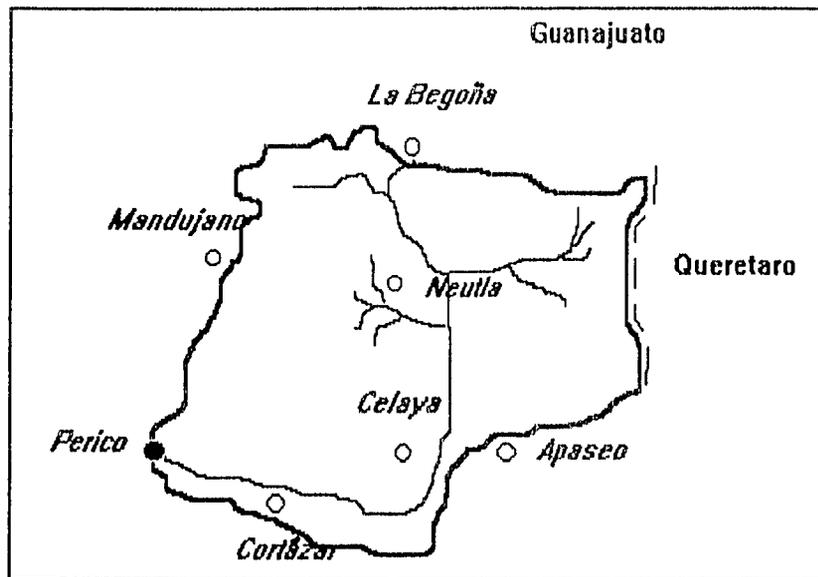


Figura 6.2. Localización de la zona de interés.

Como puede observarse, los parámetros estadísticos muestrales son muy similares y los datos en el periodo indicado muestran una correlación importante. A partir de esto se considera aceptable la representatividad, en cuanto a lámina precipitada, de la estación "Celaya" en toda la cuenca.

El gasto antecedente es obtenido del hidrograma de escurrimientos totales a partir de la curva de recesión del flujo base, es decir, aquel punto en el que inicia la curva ascendente del hidrograma (Ver figura 6.1, subcapítulo 6.1).

A partir de los datos disponibles, se seleccionan los máximos anuales de gasto antecedente, de lámina de precipitación y de gasto máximo de escurrimiento. Al gasto máximo se le resta el valor del gasto antecedente y se transforman los valores utilizando el logaritmo natural. En la Tabla 6.1 se muestran los datos originales y los transformados.

Con los datos en el dominio del logaritmo, se procede a calcular el conjunto de parámetros de la combinación lineal

$$\ln\{Q_{\max}'\} = a \ln\{h_p\} + b \ln\{Q_{\text{ant}}\} + c \quad (6.4)$$

utilizando el programa de computadora "r_l_mult.exe". Este programa hace una regresión lineal múltiple, que para este caso particular, dió los resultados siguientes:

$$a = 1.399637; \quad b = -0.044797; \quad c = 3.565873$$

$$\text{varianza} = 0.722048; \quad \text{desviación estandar} = 0.84934$$

$$R^2 = 0.187054; \quad R = 0.432497$$

Como puede apreciarse, el coeficiente de correlación múltiple es del orden del 40%, lo que hace suponer una mayor incertidumbre de las estimaciones del logaritmo del gasto máximo del escurrimiento directo utilizando la combinación lineal.

Con los valores de los coeficientes a, b y c conocidos, se calculan los valores de la variable aleatoria T, correspondientes a la combinación lineal $t = ax + by$. En este caso, "x" es el logaritmo de la lámina de precipitación y "y" es el logaritmo del gasto antecedente; esta última variable corresponde a la variable S, es decir, $s = \ln\{Q_{\text{ant}}\}$. Una vez definidas las variables para el modelo logístico bivariado, se calcula la probabilidad asociada, a partir de la cual se encuentran los valores obtenidos por el modelo propuesto. Los parámetros estadísticos de cada una de las

variables aleatorias se muestran a continuación:

Variable T Variable S

Media aritmética:	10.5230	5.1552
Media geométrica:	10.5160	5.1477
Varianza:	0.1495	0.0778
Desviación estándar:	0.3867	0.2789
Coefficiente de variación:	0.03675	0.0541
Coefficiente de curtosis:	-0.0632	-0.0628

Para encontrar los valores de los estimadores de los parámetros del modelo logístico bivariado, se utilizó el algoritmo desarrollado por Escalante, 1991. Los resultados obtenidos se presentan a continuación:

Parámetro de asociación:

$$m = 1.24643$$

Parámetros de ubicación:

$$\begin{array}{cc} (\ln\{h_p\}) & (\ln\{Q_{ant}\}) \\ u_1 = 4.9981 & u_2 = 5.2679 \end{array}$$

Parámetros de escala:

$$\alpha_1 = 0.3312 \quad \alpha_2 = 1.0228$$

Con los estimadores conocidos, se calcula el valor de $\ln\{Q_{max}\}$ para cada valor de probabilidad asignada. En la tabla 6.2 se muestran los resultados obtenidos por el modelo, y se comparan con los obtenidos de la combinación lineal y con los valores de la variable T.

AÑO	Qmáx, Gasto máximo, en 10 ³ m ³	hp, Lámina de lluvia, en mm	Qant, Gasto anteced., en 10 ³ m ³	Qmáx' = Qmáx - Qant, en 10 ³ m ³	ln Qmáx'	ln hp	ln Qant
1946	4480	117	644	3836	8.2522	4.762	6.468
1947	30558	91.2	89	30469	10.3245	4.513	4.489
1948	24507	173.9	429	24078	10.0891	5.158	6.061
1949	14799	176.4	96	14703	9.5958	5.173	4.564
1950	34435	230.6	70	34365	10.4448	5.441	4.248
1951	37549	164.5	62	37487	10.5317	5.103	4.127
1952	37535	206.5	81	37454	10.5309	5.33	4.394
1953	73681	188.8	84	73597	11.2064	5.241	4.431
1954	31098	171.9	122	30976	10.3409	5.147	4.804
1955	191899	157.6	562	191337	12.1618	5.06	6.332
1956	54380	202.1	441	53939	10.8956	5.309	6.089
1957	8469	85.4	414	8055	8.994	4.447	6.026
1958	121429	183.9	307	121122	11.7046	5.214	5.727
1959	33378	177.9	1048	32330	10.3838	5.181	6.955
1960	40233	280.7	1014	39219	10.5769	5.637	6.922
1961	22009	218.6	1302	20707	9.9382	5.387	7.172
1962	46638	187.4	656	45982	10.736	5.233	6.486
1963	29902	180.9	734	29168	10.2808	5.198	6.599
1964	85780	187.3	212	85568	11.3571	5.233	5.357
1965	58943	205.2	1001	57942	10.9672	5.324	6.909
1966	118133	175.5	1029	117104	11.6708	5.168	6.936

Todos los valores corresponden a un intervalo mensual.

Tabla 6.1. Valores de Gasto máximo, lámina de precipitación y gasto antecedente.

Probabilidad	Datos	Valores modelo Hewlett	Valores modelo logístico
0.046	8.2522	9.520602	10.93555
0.091	8.994	9.681434	11.13237
0.136	9.5958	9.941454	11.25405
0.182	9.9382	10.36449	11.35246
0.227	10.0891	10.48796	11.43599
0.273	10.2808	10.50616	11.50694
0.318	10.3245	10.51434	11.57102
0.364	10.3409	10.52321	11.629
0.409	10.3838	10.54551	11.68202
0.455	10.4448	10.55448	11.73275
0.5	10.5309	10.59996	11.78233
0.546	10.5317	10.60138	11.82353
0.591	10.5769	10.60758	11.86625
0.636	10.736	10.64981	11.9082
0.682	10.8956	10.70245	11.94787
0.727	10.9672	10.70803	11.98983
0.773	11.2064	10.72344	12.02378
0.818	11.3571	10.78479	12.06307
0.864	11.6708	10.8295	12.09358
0.909	11.7046	10.99054	12.13172
0.955	12.1618	11.14596	12.16453

Tabla 6.2. Resultados del modelo logístico bivariado para la combinación lineal.

De la tabla 6.2 y figura 6.3 se puede apreciar que el modelo logístico es una sobreestimación de los valores obtenidos por la combinación lineal y de los datos. Esto se puede deber, entre otras posibilidades, a la relativamente baja correlación lineal múltiple encontrada al calcular los coeficientes de la regresión y a que el modelo de Hewlett fue desarrollado para tormentas con duraciones de 15 a 120 minutos. Sin embargo, por los resultados obtenidos, cabe la posibilidad de que el modelo se pueda utilizar para intervalos de tiempo mensuales, siendo necesario aplicar esta metodología a otros casos para validarla.

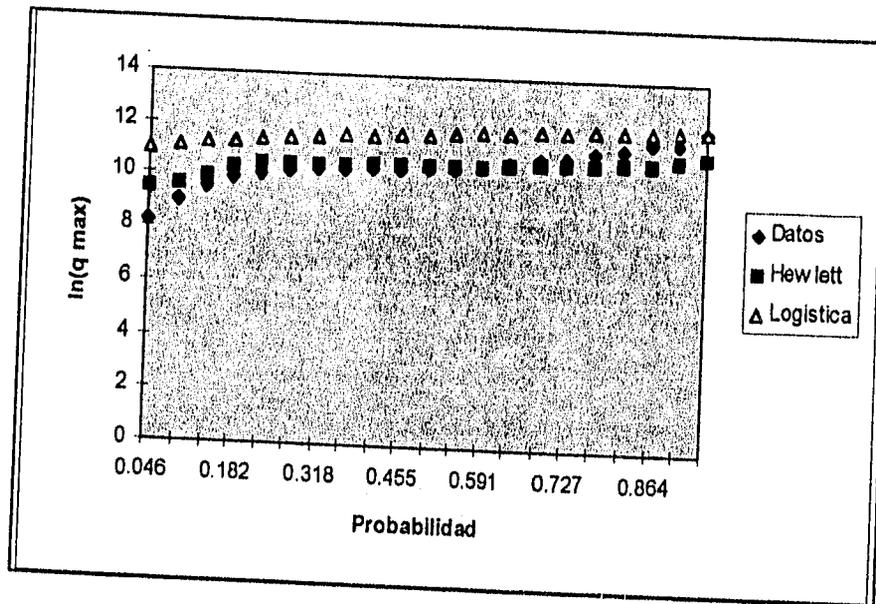


Figura 6.3. Representación de los valores calculados por el modelo, combinación lineal y datos.

Probabili- dad	Periodo de retorno en años	$\ln[Q_{\text{máx}}]$ estimado	$Q_{\text{máx}}$, en 10^3 m^3/mes
0.333	1.5	11.59619	108683
0.5	2	11.7823	130915
0.667	3	11.9364	152726
0.8	5	12.04552	170335
0.9	10	12.12257	183978
0.98	50	12.18436	195704
0.99	100	12.1897	196752
0.999	1000	12.19886	198562
0.9999	10000	12.19809	198410

Tabla 6.3. Resultados calculados por el modelo logístico bivariado para diferentes valores de periodo de retorno.

Aunque se sabe que el modelo es una sobreestimación, se hicieron los calculos para $\ln\{Q_{max}\}$ y Q_{max} para los periodos de retorno propuestos: 1.5, 2, 3, 5, 10, 50, 100, 1000 y 10000 años. Los resultados obtenidos se muestran en la tabla 6.3 y en las figuras 6.4 y 6.5.

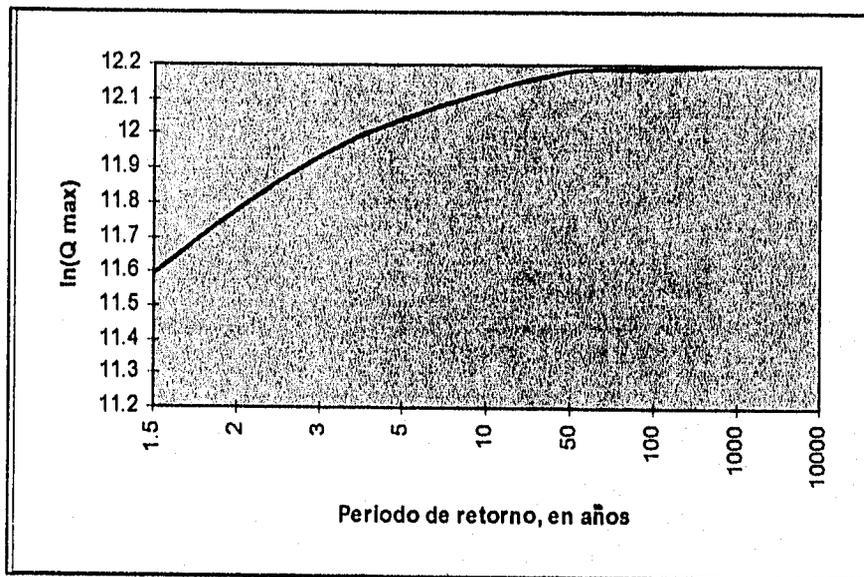


Figura 6.4. Relación $\ln\{Q_{max}\}$ contra periodo de retorno.

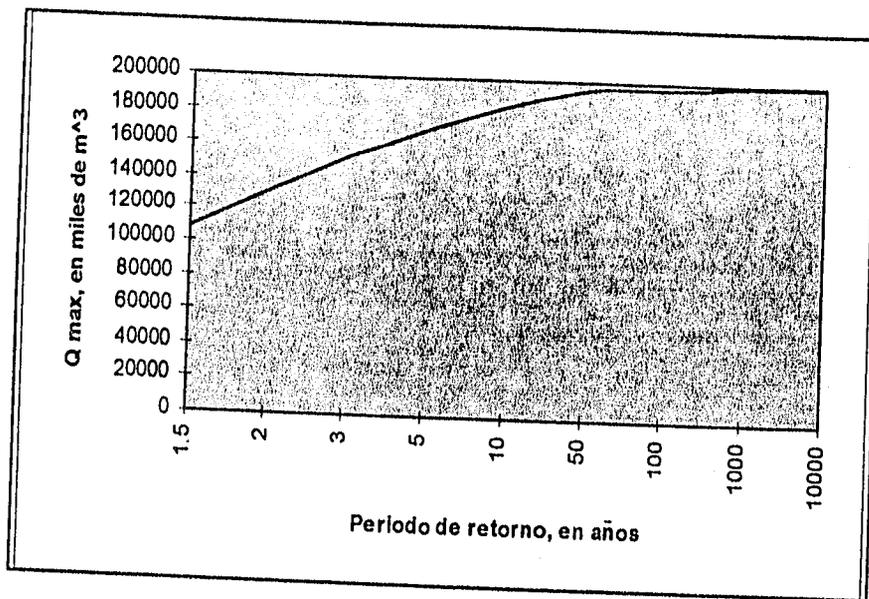


Figura 6.5. Relación Q_{max} mensual contra periodo de retorno.

7. COMENTARIOS Y CONCLUSIONES

La aplicación de las herramientas predictivas propuestas resultó en general satisfactoria, excepto para el caso de la división de las variables gasto máximo entre volumen escurrido, donde el modelo logístico presenta un comportamiento diferente al mostrado por los datos para los valores de probabilidad mayores al 90%.

En el caso de los gastos extremos esperados en la confluencia de dos ríos el modelo logístico bivariado presenta un ajuste adecuado, ya que representa la tendencia de los datos conforme aumenta el valor de la probabilidad asignada, como se observa en la figura 3.1 y en la tabla 3.2.

Para el análisis de las sequías en la estación hidrométrica "Huites", localizada en el norte de Sinaloa, se encontró que el modelo propuesto se ajustó muy bien a los datos de déficit de sequía para valores de probabilidad menores al 0.981; para valores mayores la discrepancia del modelo con los datos aumenta considerablemente, como se muestra en la figura 4.3 y en la tabla 4.2. Esto se puede deber a que el modelo propuesto no es completamente adecuado, siendo necesario buscar otro modelo que se ajuste mejor. A pesar de ello, el modelo puede ser utilizado en la predicción de valores de déficit y de duración de sequía para periodos de retorno de hasta 50 años.

Una cuestión interesante de la metodología propuesta es que a partir de los valores de déficit estimados, se puede llevar a cabo la predicción de la duración de la sequía usando una correlación de potencias, información muy importante en el manejo de los recursos hidráulicos y en la operación de presas de almacenamiento, sobre todo si se toma en cuenta que la zona de estudio es desértica. Los resultados de las predicciones para el déficit y la duración se

muestran en la figura 4.4 y en la tabla 4.3.

En cuanto al método de Zelenhasic se considera muy apropiado para cuestiones prácticas, ya que con un sólo registro se puede hacer el análisis de las sequías en una estación hidrométrica específica.

Aún cuando el modelo de Hewlett (combinación lineal para el logaritmo del gasto máximo) es un modelo diseñado para tormentas de duraciones de entre 5 y 120 minutos, los resultados obtenidos para la situación en la que se tiene información mensual, fueron satisfactorios ya que por lo observado en la tabla 6.2 y en la figura 6.3, el modelo logístico bivariado presentó un comportamiento similar, a pesar de ser una sobreestimación, al de los datos obtenidos de la estación hidrométrica y los estimados a partir de la información de precipitación y gasto antecedente. Algunas de las causas por las cuales el modelo logístico bivariado muestra esa sobreestimación son: 1) Que las variables gasto máximo mensual, lámina precipitada mensual y gasto antecedente mensual, presentan otro tipo de ajuste que no sea el de la regresión lineal múltiple; esto porque el coeficiente de correlación multivariado es de alrededor del 40%. 2) Que las suposiciones de pérdida de información, por tratarse de un intervalo de tiempo mayor para el cual fue diseñado el modelo de Hewlett, son incorrectas y por lo tanto los valores de lámina de precipitación tengan que ser transformados a valores de precipitación efectiva, siendo necesario transformar a la vez los gastos de escurrimiento directo en láminas de escurrimiento. 3) Que la suposición de representatividad de la estación climatológica "Celaya", en cuanto al comportamiento de la precipitación en toda la cuenca, es incorrecta. 4) Que existen factores externos que no son considerados en el análisis; esto es, un cambio en el régimen del río De La Laja, debido posiblemente a descargas al cauce desde algún punto aguas arriba o por la construcción de alguna obra reguladora, cuyo funcionamiento afecta a la estación "Perico".

El ajuste del modelo logístico bivariado a los datos de relación gasto máximo-volumen escurrido se verifica en la mayor parte del rango de probabilidad, excepto para los valores mayores al 90%, intervalo de mayor utilidad para la predicción. Esto se puede deber a que las variables consideradas para la estación hidrométrica Huites tiene un comportamiento diferente al del modelo logístico.

Las representaciones gráficas de las predicciones mostradas en las figuras 3.2, 4.4, 6.4 y 6.5 deben ser tomadas con precaución ya que el graficador empleado realiza el ajuste a los puntos de manera automática, no existiendo la posibilidad de representar en otro formato los resultados. Si se desea una mejor representación se recomienda utilizar otro graficador (las representaciones del presente trabajo se hicieron con el paquete de computadora "Excel" versión 5). Lo recomendable es utilizar los programas desarrollados en este trabajo y calcular los valores deseados.

En las cuatro aplicaciones se propuso el uso del error relativo para definir lo adecuado del ajuste entre datos y valores calculados por el modelo, debido a que sólo hay un modelo propuesto y por los alcances del trabajo, no se pretendió investigar si algún otro modelo presentaba un mejor ajuste (en ese caso se podrían usar parámetros estadísticos como el error estándar).

Los parámetros estadísticos pueden ser indicativos de lo adecuado de las muestras para ser utilizadas en alguno de los cuatro casos considerados en este trabajo. Así, por ejemplo, en el caso de la confluencia de dos ríos y del déficit de las sequías, que fueron los mejores resultados, los valores de media aritmética, media geométrica y coeficiente de variación presentan el mismo orden de magnitud entre las variables T y S. En el caso de la combinación lineal, las medias aritmética y geométrica presentan una diferencia de alrededor de un orden de magnitud entre T y S; en el caso del coeficiente de variación, las diferencias son mayores que en las otras dos operaciones. En el caso de la división de variables, que

fue el que arrojó los resultados más pobres, las diferencias en orden de magnitud son mayores a dos en los parámetros de tendencia central y alrededor de uno en el caso del coeficiente de variación. Lo importante de esto sería la posibilidad de mejorar los resultados pobres rescalando las variables. Si aún así los resultados no se mejoran, entonces la siguiente posibilidad es proponer otra función de densidad de probabilidad. En el caso de la división se trato de reescalar cada una de las muestras (gasto máximo y de volumen total de descarga) dividiendolas por el valor mas grande, de modo que los valores de ambas estarían acotadas entre 0 y 1; sin embargo, esto no funcionó ya que los resultados no mostraron mejora alguna, concluyendo que es necesario buscar otra función que se ajuste mejor a los datos.

Los programas de computadora desarrollados para cada herramienta predictiva presentan versatilidad en cuanto a la posibilidad de utilizar otros modelos diferentes al propuesto, escribiendo en el código fuente el nuevo modelo en la subrutina "funcion", que tiene incluida cada uno de los programas de integración.

La validación de estas herramientas dependerá del número y variedad de casos que se analicen en el futuro.

Se propone la utilización de modelos diferentes al logístico, de modo que existan amplias posibilidades en el análisis bivariado de valores extremo. Los modelos que podrían ser considerados a futuro son los de la familia de las funciones de valores extremos bivariadas, es decir, las BEV12, BEV21, BEV22 y BEV33 (la utilizada en este trabajo fue la BEV11).

BIBLIOGRAFIA.

1.- Ayllon Torres, T. y Chávez Flores, J.: México: Sus recursos naturales y su población; ed. LIMUSA, 2ª edición, México, 1992.

2.- Bertoli, L. and Moisello, U.: A regional model for estimating the probability distribution of routed peak discharges; Journal of Hydrology, 153, 1994, pp 103-138.

3.- Bravar L. and Kavvas, M.L.: On the physics of droughts. I. A conceptual framework; Journal of Hydrology, 129, 1991, pp 281-297.

4.- Bren, L.J., Farrel, P.W. and Leitch, C.J.: Use of weighted integral variables to determine the relation between rainfall intensity and storm flow and peak flow generation; Water Resources Research, vol. 23, no. 7, pp 1320-1326, july, 1987.

5.- Burden, R.I. y Faires, J.D.: Análisis numérico; Grupo Editorial Iberoamericana S.A. de C.V., 3ª edición, México, 1985.

6.- Chow V.T., Maidment, D.R. and Mays, L.W.: Applied Hydrology; ed. Mac Graw-Hill, International Edition, Singapur, 1988.

7.- Comisión Nacional del Agua: El agua y sus senderos: ríos, lagos y cascadas de México, México, 1994.

8.- DeGroot, M.H.: Probability and Statistics; Ed. Addison-Wesley Publishing Co., 2ª edición, USA, 1986.

9.- Deitel, H.M. and Deitel, P.J.: C++ How to program; Ed. Prentice-Hall International Inc., USA, 1994.

- 10.- Dingman, S.L.: Physical Hydrology; Ed. Prentice-Hall Inc., USA, 1993.
- 11.- Escalante Sandoval, C.: Distribuciones trivariadas de valores extremos y sus aplicaciones al análisis de frecuencias de avenidas; Tesis Doctoral, DEPFI, UNAM, México, 1991.
- 12.- Escalante Sandoval, C.A. and Raynal Villaseñor, J.A.: A trivariate extreme value distribution applied to flood frequency analysis; Journal of Research of the National Institute of Standards and Technology, vol. 99, num. 4, jul-aug, 1994.
- 13.- Fraser, D.A.S.: Probability and Statistics: Theory and Applications; Ed. Duxbury Press, USA, 1976.
- 14.- Johnson, R.A. and Wichern, D.W.: Applied Multivariate Statistical Analysis; Ed. Prentice-Hall International Inc., 3rd edition, USA, 1992.
- 15.- Kai Lai Chung.: A Course in Probability Theory; Ed. Academic Press, 2nd edition, USA, 1974.
- 16.- Kite, G.W.: Frequency and Risk Analyses in Hydrology; Water Resources Publications, USA, 1988.
- 17.- Linsley, R.K. and Franzini, J.B.: Water Resources Engineering, Ed. McGraw-Hill, Japan, 1972.
- 18.- Linsley, R.K., Kohler, M.A. and Paulhus, J.L.H.: Hydrology for Engineers; Ed. McGraw-Hill, 3rd edition, Singapur, 1988.
- 19.- Meyer. P.L.: Introductory Probability and Statistical Applications; Ed. Addison-Wesley Publishing Co., USA, 1965.

20.- Nakamura, S.: Applied Numerical Methods in C; Ed. Prentice-Hall International Inc., USA, 1993.

21.- Raynal-Villasenor, J.A.: Análisis de gastos máximos aguas abajo de confluencias de ríos; Revista Ingeniería Hidráulica en México, enero-abril 1990, pp 37-39, México.

22.- Viessman, W., Lewis, G.L. and Knapp, J.W.: Introduction to Hydrology; Ed. Harper and Row, 3rd edition, USA, 1989.

23.- Zelenhasic, E. and Salvai, A.: A Method of Streamflow Drought Analysis; Water Resources Research, vol. 23, no. 1, pp 156-168, january, 1987.

APENDICE A.

PROGRAMAS DE COMPUTADORA

A.1 DESCRIPCION Y LISTADOS DE LOS PROGRAMAS.

Los programas desarrollados son el apoyo para llevar a cabo las tareas necesarias en la metodología propuesta para cada una de las operaciones entre variables aleatorias hidrológicas.

Los programas fueron escritos en el lenguaje de programación C++, utilizando el compilador Turbo C++ versión 3.0, en un ambiente "windows", versión 3.11. Para cada una de las operaciones, se desarrollaron programas generales o proyectos, los cuales contienen a su vez programas que llevan a cabo tareas particulares. Los proyectos son: "suma.prj", aplicado en la confluencia de dos ríos; "resta.prj", en el déficit de sequías; "div_var.prj", en la relación de gastos-volumen de escurrimiento de avenidas; "r_l_mult.cpp" y "comb_lin.prj" en la regresión lineal (obtención de coeficientes) y en la combinación lineal del gasto máximo, lámina precipitada y gasto antecedente, respectivamente. A continuación se describen brevemente los proyectos, haciendo comentarios sobre cada uno de los programas que lo constituyen y mostrando el listado correspondiente.

Proyecto "suma.prj".

Esta constituido por los programas siguientes:

lectura.cpp . Captura el número de datos *n* y los valores de la muestra, los cuales se asignan a un arreglo unidimensional *x*, desde un archivo externo cuyo formato es definido en el programa *crea_arc.cpp*, comentado mas adelante.

```
#include <fstream.h>
#include <conio.h>
#include <iomanip.h>
#include <stdio.h>
#include <math.h>
#include "lect.h"
int lectura(int *n, double *x)
{
    int i;
    char arch_ent[30];
```

```

cout << "\nDar el nombre del archivo: ";
cin >> arch_ent;
ifstream entrada(arch_ent);
entrada >> *n;
cout << endl << "el numero de datos del archivo es "
    << *n << endl << endl;
for (i = 0; i < *n; i++) {
    entrada >> *x;
    cout << "    " << *x << endl;
    x++;
}
entrada.close();
return 0;
}

```

crea_arch.cpp . Crea los archivos de entrada para ser leídos por el programa "lectura.cpp". A través de la interacción con el usuario, se capturan cada uno de los datos, de forma que el archivo creado es una columna cuyo primer valor es el número de datos (número entero) y los renglones siguientes son los valores de la muestra (números en doble precisión).

```

#include <fstream.h>
#include <conio.h>
#include <iomanip.h>
#include <stdio.h>
#include <math.h>
#include "crea.h"

int crear()
{
    char nomb_arch[30], resp;
    int numtot, i;
    double valor;

    do {
        cout << "Dar el nombre del archivo a crear: ";
        cin >> nomb_arch;
        cout << "Dar el numero total de valores a capturar: ";
        cin >> numtot;
        ofstream archivo(nomb_arch);
        archivo << numtot << endl;
        for (i = 0; i < numtot; i++) {
            cout << "dar el valor " << (i+1) << ": ";
            cin >> valor;
            archivo << valor << endl;
        }
        cout << "repetir la creación del archivo? (s/n) : ";
        cin >> resp;
    }
    while ((resp == 's') || (resp == 'S'));

    return 0;
}

```

arregmax.cpp . Arregla los datos de mayor a menor utilizando el método de la "burbuja" (Deitel y Deitel, 1994); además se calcula la probabilidad asociada utilizando la expresión de Weibull

para valores de excedencia.

```
#include <fstream.h>

void arreglomap(double *val, int cont, double *prob)
{
    double temp;

    for (int p=1; p<cont; p++)
        for (int j=0; j<cont-1; j++)
            if (val[j] < val[j+1]) {
                temp = val[j];
                val[j] = val[j+1];
                val[j+1] = temp;
            }
    for (int i=0; i<cont; i++){
        prob[i]=1.0-((double)i+1)/((double)cont+1);
        cout << "valor[" << i+1 << "] = " << val[i] << " ; prob = "
            << prob[i] << endl;
    }
}
```

arregmin.cpp . Arregla la muestra en cuestión de menor a mayor. Esto se hace cuando se requiere un análisis de mínimos extremos. Es similar al programa anterior.

```
#include <fstream.h>

void arreglomin(double *val, int cont, double *prob)
{
    double temp;

    for (int p=0; p<cont; p++)
        for (int j=0; j<cont-1; j++)
            if (val[j] > val[j+1]) {
                temp = val[j];
                val[j] = val[j+1];
                val[j+1] = temp;
            }
    for (int i=0; i<cont; i++){
        prob[i]=1.0-((double)i+1)/((double)cont+1);
        cout << "valor[" << i+1 << "] = " << val[i] << " ; prob = "
            << prob[i] << endl;
    }
}
```

par_est.cpp . Calcula los estadísticos muestrales de los datos capturados; esto es, la media aritmética, la media geométrica, la varianza, la desviación estándar, el coeficiente de variación y el coeficiente de curtosis.

```
#include <fstream.h>
#include <stdio.h>
#include <math.h>
#include "par.h"

const int TAM = 51;
```

```

// constructor de la clase param
param::param()
{
    for (int j = 0; j <= TAM; j++)
        vx[j]=0.0;
}

param::~~param()
{
}

// funcion miembro para captura de datos
void param::datos(int nd, double *x)
{
    numdat=nd;
    for (int i=0; i<numdat; i++)
        vx[i]=x[i];
}

// funcion miembro para calculo de la media aritmetica
double param::media_arit()
{
    double suma=0.0;
    for (int i = 0; i < numdat; i++)
        suma+=vx[i];

    return (suma/numdat);
}

// funcion miembro para calcular la media geometrica
double param::media_geom()
{
    double prod=1.0;
    for (int i = 0; i<numdat; i++)
        prod *= vx[i];

    return exp(log(prod)/numdat);
}

// funcion miembro para calcular la varianza
double param::varianza(double xm)
{
    double dif2=0.0;
    for (int i = 0; i<numdat; i++)
        dif2 += pow((vx[i]-xm),2);

    return dif2/(numdat-1);
}

// funcion miembro para calcular la desviacion estandar
double param::desv_est(double v)
{
    return sqrt(v);
}

// funcion miembro para calcular el coeficiente de variacion
double param::coef_var(double xm, double d)
{
    return d/xm;
}

// funcion miembro para calcular el coeficiente de curtosis
double param::curtosis(double xm, double d)
{
    double suma3=0.0;

    for (int i = 0; i < numdat; i++)
        suma3 += pow((vx[i]-xm),3);

    return suma3/((numdat-1)*(numdat-2)*pow(d,3));
}

```

`integ_s.cpp` . Calcula la integral de la función de densidad logística bivariada para el caso de la suma, por medio de la regla de Simpson simple bidimensional; los límites de integración van de -500 a 500 para la variable simple S, y de -500 a t (valor de suma) para la variable compuesta T. El paso de la integración en el caso de S es de 66.67 con 30 puntos de cálculo en el intervalo. Para T es variable dependiendo de t. Para la búsqueda del valor t, se inicia el proceso de bisección con un valor de 5000, el cual asegura que la probabilidad calculada (`id[]`) esta alrededor de 1. Para detener el proceso de búsqueda, se propuso una tolerancia de 1/1000 o 30 iteraciones; el parámetro de comparación para la tolerancia es el error relativo cuadrado, considerando los valores calculados de la probabilidad (`id[]`), con respecto a la probabilidad asociada estimada por la expresión de Weibull.

```
#include <fstream.h>
#include <math.h>
#include <stdio.h>

const int TAM = 50;

int integ_s(int *ctr, double *pr, double *pa1, double *pa2,
            double *pul, double *pu2, double *pm, double *vlr)
{
    double funcion_s(double, double, double *, double *, double *, double *,
                    double *);

    int m=15, n=15, iter;
    double varx[TAM], vary[TAM], id[TAM];
    double a=-500.0, b, c=-500.0, d=500.0, h, k, term1, term2, term3,
            term4, term5, term6, term7, term8, term9, term10, term11,
            term12, term13, term14, term15, term16;
    double t2, t3, t5, t6, t7, t8, t9, t10, t11, t12, t14, t15, err=1.0, fp0,
            p0, pn, p, fp, fpn, cp;

    k = (d - c)/(2*n); // paso para y

    for (int ll = 0; ll<*ctr; ll++) {
        b=0.005;
        iter=1;
        do
        {
            for (int z = 0; z<2; z++) {
                t2=0.0; t3=0.0; t5=0.0; t6=0.0; t7=0.0; t8=0.0; t9=0.0;
                t10=0.0; t11=0.0; t12=0.0; t14=0.0; t15=0.0;

                h = (b - a)/(2*m); // paso para x

                for (int l = 0; l <= 2*m; l++)
                    varx[l] = a + l*h;

                for (int o = 0; o <= 2*n; o++)
                    vary[o] = c + o*k;
            }
        }
    }
}
```

```

term1 = funcion_s(varx[0], vary[0], pa1, pa2, pu1, pu2, pm);
term4 = funcion_s(varx[2*m], vary[0], pa1, pa2, pu1, pu2, pm);
term13 = funcion_s(varx[0], vary[2*n], pa1, pa2, pu1, pu2, pm);
term16 = funcion_s(varx[2*m], vary[2*n], pa1, pa2, pu1, pu2, pm);

for (int i = 1; i <= m-1; i++)
    t2 += funcion_s(varx[2*i], vary[0], pa1, pa2, pu1, pu2, pm);
term2 = 2*t2;

for (i = 1; i <= m; i++)
    t3 += funcion_s(varx[2*i-1], vary[0], pa1, pa2, pu1, pu2, pm);
term3 = 4*t3;

for (int j = 1; j <= n-1; j++)
    t5 += funcion_s(varx[0], vary[2*j], pa1, pa2, pu1, pu2, pm);
term5 = 2*t5;

for (j = 1; j <= n-1; j++)
    for (i = 1; i <= m-1; i++)
        t6 += funcion_s(varx[2*i], vary[2*j], pa1, pa2, pu1, pu2, pm);
term6 = 4*t6;

for (j = 1; j <= n-1; j++)
    for (i = 1; i <= m; i++)
        t7 += funcion_s(varx[2*i-1], vary[2*j], pa1, pa2, pu1, pu2, pm);
term7 = 8*t7;

for (j = 1; i <= n-1; i++)
    t8 += funcion_s(varx[2*m], vary[2*j], pa1, pa2, pu1, pu2, pm);
term8 = 2*t8;

for (j = 1; j <= n; j++)
    t9 += funcion_s(varx[0], vary[2*j-1], pa1, pa2, pu1, pu2, pm);
term9 = 4*t9;

for (j = 1; j <= n; j++)
    for (i = 1; i <= m-1; i++)
        t10 += funcion_s(varx[2*i], vary[2*j-1], pa1, pa2, pu1, pu2, pm);
term10 = 8*t10;

for (j = 1; j <= n; j++)
    for (i = 1; i <= m; i++)
        t11 += funcion_s(varx[2*i-1], vary[2*j-1], pa1, pa2, pu1, pu2, pm);
term11 = 16*t11;

for (j = 1; j <= n; j++)
    t12 += funcion_s(varx[2*m], vary[2*j-1], pa1, pa2, pu1, pu2, pm);
term12 = 4*t12;

for (i = 1; i <= m-1; i++)
    t14 += funcion_s(varx[2*i], vary[2*n], pa1, pa2, pu1, pu2, pm);
term14 = 2*t14;

for (i = 1; i <= m; i++)
    t15 += funcion_s(varx[2*i-1], vary[2*n], pa1, pa2, pu1, pu2, pm);
term15 = 4*t15;

id[11] = (h*k/9) * (term1+term2+term3+term4+term5
+term6+term7+term8+term9+term10+term11
+term12+term13+term14+term15+term16);
if ((z == 0) && (iter == 1)) {
    fp0=id[11];
    p0=b;
    pn=5000.0;
    p=p0+(pn-p0)/2;
    b=p;
}
if ((z == 0) && (iter > 1)) {
    fp0=id[11];
    p=p0+(pn-p0)/2;
    b=p;
}
]

```

```

        if (z == 1)
            fp=id[11];
    }
    cp=fp-pr[11];
    if (cp<0) {
        b=p;
        p0=b;
    }
    if (cp>0) {
        b=p;
        pn=b;
    }
    err=sqrt((pr[11]-fp0)*(pr[11]-fp0)/(fp0*fp0));
    iter+=1;
    } while ((err >= 0.0001) && (iter < 30));
    vlr[11]=b;
}
return 0;
}

double funcion_s(double x, double y, double *alfal, double *alfa2,
                double *u1, double *u2, double *m)
{
    double t1, t2, pot1, pot2;

    t1 = exp(- *m *(x-y- *u1)/ *alfal);
    t2 = exp(- *m *(y- *u2)/ *alfa2);
    pot1 = (1.0/ *m)-2.0;
    pot2 = 1.0/ *m;

    return 1.0/( *alfal * *alfa2)*t1*t2*pow((t1+t2), pot1)*( *m-1+
        pow((t1+t2), pot2))*exp(-(pow((t1+t2), pot2)));
}

```

suma_var.cpp . Este es el programa central que llama a los programas anteriores.

```

#include <fstream.h>
#include <conio.h>
#include <iomanip.h>
#include <stdio.h>
#include <math.h>
#include "par.h"
#include "lect.h"
#include "crea.h"
#include "arregmin.h"
#include "arregmax.h"
#include "integ_s.h"

const int TAM = 50;

main()
{
    /**** declaracion de variables ****/

    param P;
    char opc,nomb_arch[30],cte;
    int opcion, num, c;
    double valor1[TAM], valor2[TAM], med_ar1, med_geol, var1,
        desv_e1, coef_v1, curt1, med_ar2, med_geol2, var2, desv_e2,
        coef_v2, curt2, pral, pra2, prul, pru2,
        prm, pr_cal[TAM],v_cal[TAM],tr[15];

    /*
    variables en el programa principal:

```

```

P -variable del tipo param (calculo de parametros estadisticos)
opc,opcion -caracter y numero para direccionar control del programa
num -numero de valores
c
*/

cout << endl << endl
<< "      ssssss uu uu mmm mmm aaaaaa vv vv aaaaaa rrrrrr "
<< endl
<< "      sss uu uu mmm mmm aa aa vv vv aa aa rr rr "
<< endl
<< "      ssssss uu uu mmm mmm aaaaaa vv vv aaaaaa rrrrrr "
<< endl
<< "      sss uu uu mmm mmm aa aa vvvv aa aa rr rr"
<< endl
<< "      ssssss uuuuuu mm mm aa aa vv aa aa rr rr"
<< endl << endl << endl << endl << endl
<< "      PROGRAMA QUE CALCULA LA PROBABILIDAD CONJUNTA BIVARIADA "
<< endl
<< "      DE LA SUMA DE DOS VARIABLES ALEATORIAS UTILIZANDO EL MODELO "
<< endl
<< "      LOGISTICO (DISTRIBUCION DE VALORES EXTREMOS TIPO II). "
<< endl << endl << endl
<< "<oprimir tecla 'retorno'>" << endl;
getchar();
clrscr();

do {
    cout << "Seleccionar una de las opciones siguientes:"
        << endl << endl
        << "      (0)...terminar la sesion "
        << endl
        << "      (1)...captura de datos de un archivo externo "
        << endl
        << "      (2)...creacion de un archivo de datos "
        << endl
        << "      (3)...arreglo de muestras para analisis de maximos"
        << endl
        << "      (4)...arreglo de muestras para analisis de minimos"
        << endl
        << "      (5)...suma de variables aleatorias"
        << endl
        << "      (6)...calculo de parametros estadisticos "
        << endl
        << "      (7)...calculo de la probabilidad de los datos"
        << endl
        << "      (8)...calculo de la probabilidad de valores propuestos"
        << endl << endl
        << "====> ";
    cin >> opcion;
    clrscr();
    if (opcion != D) {
        switch (opcion){
            case 1:
                cout << "      ===== \n"
                    << "      lectura de datos desde un archivo externo \n"
                    << "      ===== \n"
                    << endl << endl
                    << "      Indicar el numero de archivos a leer: \n"
                    << "      (U)no" << endl
                    << "      (D)os" << endl << endl
                    << "      ===== ";
                cin >> opc;
                if ((opc == 'd') || (opc == 'D')) {
                    cout << endl << "      primer archivo a leer" << endl;
                    lectura(&num, valor1);
                    getchar();
                    clrscr();
                    cout << endl << "      segundo archivo a leer" << endl;
                    lectura(&num, valor2);
                }
            }
        }
    }
}

```

```

if ((opc == 'u') || (opc == 'U'))
    lectura(&num, valor1);

cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
break;

case 2:
cout << "          ===== \n"
<< "          creacion de archivo de entrada \n"
<< "          ===== \n"
<< endl << endl
<< "          Indicar el numero de archivos a crear: \n"
<< "          (U)no \n"
<< "          (D)os \n"
<< endl << "          ==> ";
cin >> opc;
if ((opc == 'u') || (opc == 'U'))
    crear();
if ((opc == 'd') || (opc == 'D')) {
    cout << "          primera muestra" << endl;
    crear();
    clrscr();
    cout << "          segunda muestra" << endl;
    crear();
}
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
break;

case 3:
cout << "          ===== \n"
<< "          arreglo de los datos de mayor a menor \n"
<< "          ===== \n"
<< endl << endl
<< "          Indicar el numero de muestras a ordenar: "
<< endl
<< "          (U)no" << endl
<< "          (D)os" << endl
<< endl << "          ==> ";
cin >> opc;
if ((opc == 'D') || (opc == 'd')) {
    cout << "          arreglo de la primera muestra : " << endl << endl;
    arreglomax(valor1,num,pr_cal);
    cout << endl << "para continuar oprimir <retorno>";
    getchar();
    cout << "          arreglo de la segunda muestra : " << endl << endl;
    arreglomax(valor2,num,pr_cal);
}
if ((opc == 'U') || (opc == 'u')){
    cout << "          arreglo de la muestra : " << endl << endl;
    arreglomax(valor1,num,pr_cal);
}
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
break;

case 4:
cout << "          ===== \n"
<< "          arreglo de los datos de menor a mayor \n"
<< "          ===== \n"
<< endl << endl
<< "          Indicar el numero de muestras a ordenar: "
<< endl
<< "          (U)no" << endl
<< "          (D)os" << endl
<< endl << "          ==> ";
cin >> opc;
if ((opc == 'D') || (opc == 'd')) {
    cout << "          arreglo de la primera muestra : " << endl << endl;
    arreglomin(valor1,num,pr_cal);
}

```

```

cout << endl << "para continuar oprimir <retorno>";
getchar();
cout << " arreglo de la segunda muestra :" << endl << endl;
arreglomin(valor2,num,pr_cal);
}
if ((opc == 'U') || (opc == 'u')){
cout << " arreglo de la muestra :" << endl << endl;
arreglomin(valor1,num,pr_cal);
}
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
break;
case 5:
cout << "          ===== \n"
<< "          suma de dos variables aleatorias \n"
<< "          ===== \n"
<< endl << endl
<< "variable " << "suma de " << endl
<< "simple " << "variables " << endl << endl;
for (c=0; c<num; c++) {
valor2[c]=valor2[c]+valor1[c];
cout << setw(10) << valor1[c] << setw(10) << valor2[c]
<< endl;
}
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
break;
case 6:
cout << "          ===== \n"
<< "          calculo de los parametros estadisticos \n"
<< "          ===== \n"
<< endl << endl
<< " Desea calcular los parametros de dos muestras? (s/n): ";
cin >> opc;
if ((opc == 's') || (opc == 'S')) {
P.datos(num,valor1);
med_ar1=P.media_arit();
med_geol=P.media_geom();
var1=P.varianza(med_ar1);
desv_e1=P.desv_est(var1);
coef_v1=P.coef_var(med_ar1,desv_e1);
curt1=P.curtosis(med_ar1,desv_e1);
cout << " valores de la muestra simple:" << endl << endl
<< " media aritmetica : " << med_ar1 << endl
<< " media geometrica : " << med_geol << endl
<< " varianza : " << var1 << endl
<< " desviacion estandar: " << desv_e1 << endl
<< " coef. variacion : " << coef_v1 << endl
<< " curtosis : " << curt1 << endl << endl;
P.datos(num,valor2);
med_ar2=P.media_arit();
med_geo2=P.media_geom();
var2=P.varianza(med_ar2);
desv_e2=P.desv_est(var2);
coef_v2=P.coef_var(med_ar2,desv_e2);
curt2=P.curtosis(med_ar2,desv_e2);
cout << " valores de la muestra compuesta:" << endl << endl
<< " media aritmetica : " << med_ar2 << endl
<< " media geometrica : " << med_geo2 << endl
<< " varianza : " << var2 << endl
<< " desviacion estandar: " << desv_e2 << endl
<< " coef. variacion : " << coef_v2 << endl
<< " curtosis : " << curt2 << endl;
}
else {
P.datos(num,valor1);
med_ar1=P.media_arit();
med_geol=P.media_geom();
var1=P.varianza(med_ar1);
desv_e1=P.desv_est(var1);

```

```

coef_v1=P.coef_var(med_ar1,desv_el);
curt1=P.curtosis(med_ar1,desv_el);
cout << " valores de la muestra:" << endl << endl
<< " media aritmetica : " << med_ar1 << endl
<< " media geometrica : " << med_geol << endl
<< " varianza : " << var1 << endl
<< " desviacion estandar: " << desv_el << endl
<< " coef. variacion : " << coef_v1 << endl
<< " curtosis : " << curt1 << endl;
}
cout << endl << " Dar el valor de alfa 1, alfa 2, u1, u2 y m: " << endl;
cin >> pral >> pra2 >> prul >> pru2 >> prm;

cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
break;
case 7:
cout << " ===== \n"
<< " calculo de la probabilidad de los datos \n"
<< " ===== \n"
<< endl << endl << endl << endl
<< " Esperar un momento. Calculando la probabilidad"
<< endl
<< " conjunta bivariada de los datos.";
integ_s(&num,&pr_cal,&pral,&pra2,&prul,&pru2,&prm,v_cal);
clrscr();

cout << " Desea almacenar los resultados en un archivo? (s/n): ";
cin >> cte;
if ((cte == 'S') || (cte == 's')) {0
cout << " Dar el nombre del archivo a crear: ";
cin >> nomb_arch;
ofstream archivo(nomb_arch);
archivo << " ARCHIVO: " << nomb_arch << endl << endl
<< "PROBABILIDAD VALOR DE LA VALOR OBTENIDO " << endl
<< "ASIGNADA VARIABLE (DATOS) DEL MODELO" << endl
<< endl;
for (c=0; c<num; c++) {
archivo << setw(15) << pr_cal[c] << setw(15) << valor2[c] << setw(15)
<< v_cal[c] << endl;
}
archivo << endl << endl << "Parametros: alfa 1= " << pral << " ; alfa 2= "
<< pra2 << " ; u1= " << prul << " ; u2= " << pru2 << endl
<< "m= " << prm << endl;
}

cout << endl
<< "probabilidad valor de la valor obtenido" << endl
<< "asignada variable (datos) del modelo" << endl << endl;

for (c=0; c<num; c++)
cout << setw(15) << pr_cal[c] << setw(15) << valor2[c] << setw(15)
<< v_cal[c]
<< endl;
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
break;
case 8:
cout << " ===== \n"
<< " calculo de la probabilidad de valores propuestos \n"
<< " ===== \n"
<< endl << endl;
cout << endl << " Dar el numero de valores de periodo de retorno: ";
cin >> num;
for (c=0; c<num; c++) {
cout << endl << " Dar el valor " << c+1 << " del periodo de retorno: ";
cin >> tr[c];
pr_cal[c]=1.0-1/tr[c];
}
cout << endl << endl << endl << endl

```

```

        << "          Esperar un momento. Calculando la probabilidad"
        << endl
        << "          conjunta bivariada de los valores propuestos." << endl;
    integ_s(&num,pr_cal,&pral,&pra2,&prul,&pru2,&prm,v_cal);
}

clrscr();

cout << "          Desea almacenar los resultados en un archivo? (s/n): ";
cin >> cte;
if ((cte == 's') || (cte == 'S')) {
    cout << "          Dar el nombre del archivo a crear: ";
    cin >> nomb_arch;
    ofstream archivo(nomb_arch);
    archivo << "          ARCHIVO: " << nomb_arch << endl << endl
        << "PERIODO DE          PROBABILIDAD          VALOR OBTENIDO " << endl
        << "RETORNO          ASOCIADA          DEL MODELO" << endl << endl;
    for (c=0; c<num; c++)
        archivo << setw(15) << tr[c] << setw(15) << pr_cal[c] << setw(15)
            << v_cal[c] << endl;
    archivo << endl << endl << "Parametros: alfa 1= " << pral << " ; alfa 2= "
        << pra2 << endl << "          u1= " << prul << " ; u2= "
        << prul << endl << "          m= " << prm << endl;
}
clrscr();
cout << "periodo de          probabilidad          valor obtenido" << endl
    << "retorno          asociada          del modelo" << endl << endl;

for (c=0; c<num; c++)
    cout << setw(15) << tr[c] << setw(15) << pr_cal[c] << setw(15) << v_cal[c]
        << endl;
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
break;

default:
    cout << "opcion incorrecta, intentar de nuevo" << endl
        << endl << "para continuar oprimir tecla <retorno>";
    getchar();
    clrscr();
    break;
}
}
}
while (opcion != 0);

return 0;
}

```

Proyecto "resta.prj".

Este proyecto contiene siete programas, de los cuales los primeros cinco ya han sido explicados en el proyecto anterior.

lectura.cpp .

crea_arc.cpp .

arregmax.cpp .

arregmin.cpp .

`par_est.cpp` .

`integ_r.cpp` . Calcula el valor de la diferencia de dos variables aleatorias a través de la integración del modelo logístico bivariado. Es el mismo esquema que en el caso de la suma, con la única diferencia en la función.

```
double funcion_r(double x, double y, double *alfa1, double *alfa2,
                 double *u1, double *u2, double *m)
{
    double t1, t2, pot1, pot2;

    t1 = exp(- *m *(x+y- *u1)/ *alfa1);
    t2 = exp(- *m *(y- *u2)/ *alfa2);
    pot1 = (1.0/ *m)-2.0;
    pot2 = 1.0/ *m;

    return 1.0/( *alfa1 * *alfa2)*t1*t2*pow((t1+t2), pot1)*( *m-1+
        pow((t1+t2), pot2))*exp(-(pow((t1+t2), pot2)));
}
```

`restavar.cpp` . Es el programa central que llama a los programas anteriores. Es muy similar al anterior, variando únicamente en la opción 5 y en el calculo de la variable T a partir de los valores propuestos de periodo de retorno (opción 8).

```
case 5:
    cout << "      ===== \n"
    << "      resta de dos variables aleatorias \n"
    << "      ===== \n"
    << endl << endl
    << "variable " << "resta de " << endl
    << "simple   " << "variables " << endl << endl;
    for (c=0; c<num; c++) {
        valor2[c]=fabs(valor2[c]-valor1[c]);
        cout << setw(10) << valor1[c] << setw(10) << valor2[c]
        << endl;
    }
    cout << endl << "para continuar oprimir tecla <retorno>";
    getchar();
    clrscr();
    break;
```

Proyecto "div_var.prj".

Contiene siete programas, de los cuales los primeros cinco ya fueron explicados en el caso de la suma.

`lectura.cpp` .

crea_arc.cpp .
 arregmax.cpp .
 arregmin.cpp .
 par_est.cpp .

integ_d.cpp . Calcula el valor de la división entre dos variables aleatorias a través de la integración del modelo logístico bivariado. Es el mismo esquema que en el caso de la suma y resta, con la diferencia en los límites de integración, que ahora van de un valor cercano a cero (0.0001) hasta 10000, en el caso de S y hasta t en el caso de T . Otros cambios son el producto del valor de la variable S en cada término de la integración, así como las variables de la función logística bivariada ($t=x/y$).

```

term1 = vary[0]*funcion_d(varx[0],vary[0],pa1,pa2,pu1,pu2,pm);
term4 = vary[0]*funcion_d(varx[2*m],vary[0],pa1,pa2,pu1,pu2,pm);
term13 = vary[2*n]*funcion_d(varx[0],vary[2*n],pa1,pa2,pu1,pu2,pm);
term16 = vary[2*n]*funcion_d(varx[2*m],vary[2*n],pa1,pa2,pu1,pu2,pm);
term2 = 2*vary[0]*t2;
term3 = 4*vary[0]*t3;

for (int j = 1; j <= n-1; j++) {
  t5 += vary[2*j]*funcion_d(varx[0],vary[2*j],pa1,pa2,pu1,pu2,pm);
}
for (j = 1; j <= n-1; j++)
  for (i = 1; i <= m-1; i++) {
    t6 += vary[2*j]*funcion_d(varx[2*i],vary[2*j],pa1,pa2,pu1,pu2,pm);
  }
term6 = 4*t6;

for (j = 1; j <= n-1; j++)
  for (i = 1; i <= m; i++) {
    t7 += vary[2*j]*funcion_d(varx[2*i-1],vary[2*j],pa1,pa2,pu1,pu2,pm);
  }
term7 = 8*t7;

for (j = 1; i <= n-1; i++) {
  t8 += vary[2*j]*funcion_d(varx[2*m],vary[2*j],pa1,pa2,pu1,pu2,pm);
}
term8 = 2*t8;

for (j = 1; j <= n; j++) {
  t9 += vary[2*j-1]*funcion_d(varx[0],vary[2*j-1],pa1,pa2,pu1,pu2,pm);
}
term9 = 4*t9;

for (j = 1; j <= n; j++)
  for (i = 1; i <= m-1; i++) {
    t10 += vary[2*j-1]*funcion_d(varx[2*i],vary[2*j-1],pa1,pa2,pu1,pu2,pm);
  }
term10 = 8*t10;

for (j = 1; j <= n; j++)
  for (i = 1; i <= m; i++) {
    t11 += vary[2*j-1]*funcion_d(varx[2*i-1],vary[2*j-1],pa1,pa2,pu1,pu2,pm);
  }
term11 = 16*t11;

for (j = 1; j <= n; j++) {

```

```

    t12 += vary[2*j-1]*funcion_d(varx[2*m], vary[2*j-1], pa1, pa2, pu1, pu2, pm);
}
term12 = 4*t12;

for (i = 1; i <= m-1; i++) {
    t14 += funcion_d(varx[2*i], vary[2*n], pa1, pa2, pu1, pu2, pm);
}
term14 = 2*vary[2*n]*t14;

for (i = 1; i <= m; i++) {
    t15 += funcion_d(varx[2*i-1], vary[2*n], pa1, pa2, pu1, pu2, pm);
}
term15 = 4*vary[2*n]*t15;

double funcion_d(double x, double y, double *alfa1, double *alfa2,
double *u1, double *u2, double *m)
{
    double t1, t2, pot1, pot2;

    t1 = exp(- *m *(x+y- *u1)/ *alfa1);
    t2 = exp(- *m *(y- *u2)/ *alfa2);
    pot1 = (1.0/ *m)-2.0;
    pot2 = 1.0/ *m;

    return 1.0/( *alfa1 * *alfa2)*t1*t2*pow((t1+t2), pot1)*(*m-1+
pow((t1+t2), pot2))*exp(-(pow((t1+t2), pot2)));
}

```

`div_var.cpp` . Es el programa central que llama a los programas anteriores. Es similar al programa para la suma; difiere en las opciones 5 y 8 (uso de "funcion_d()", en lugar de "funcion_s()").

```

case 5:
    cout << "          ===== \n"
    << "          division de dos variables aleatorias \n"
    << "          ===== \n"
    << endl << endl
    << "variable " << "division " << endl
    << "simple " << "variables " << endl << endl;
    for (c=0; c<num; c++) {
        valor2[c]=valor2[c]/valor1[c];
        cout << setw(10) << valor1[c] << setw(10) << valor2[c]
        << endl;
    }
    cout << endl << "para continuar oprimir tecla <retorno>";
    getch();
    clrscr();
    break;

```

Proyecto "r_1_mult.prj".

Contiene cinco programas, de los cuales los primeros dos ya fueron explicado en el caso de la suma.

lectura.cpp .
crea_arc.cpp .
translog.cpp . Transforma cada uno de los valores de la muestra en logaritmos naturales.

```
#include <fstream.h>
#include <math.h>

void transflog(int cont, double *v)
{
    for (int i=0; i<cont; i++) {
        v[i]=log(v[i]);
        cout << "valor[" << i << "] = " << v[i] << endl;
    }
}
```

r_l_mult.cpp . Calcula los valores de los coeficientes de la regresión lineal múltiple par tres términos (Johnson y Wichern, 1992). La solución de la matriz inversa y del sistema de ecuaciones para encontrar los valores deseados, se lleva a cabo por el método de Gauss, con sustitución hacia atrás y pivoteo máximo por columna (Nakamura, 1993).

```
#include <fstream.h>
#include <stdio.h>
#include <math.h>

// funcion para la regresion lineal multiple usando minimos cuadrados
int r_l_mult(int n, double *z1, double *z2, double *y, double *param1,
            double *param2, double *param0)
{
    const int TAM = 25;

    /**** prototipos de las funciones ****/

    // funcion para calcular los elementos de la matriz inversa
    int m_inv(int, double [3][3]);

    // funcion para calcular la media aritmetica de y
    double y_media(int, double *);

    /**** declaracion de variables ****/

    int i, j, k;
    double sumal, izt, izty, zb, ete=0.0, varianza, ym, y_ym=0.0,
        y1_ym=0.0, R2, z{TAM}{3}, zt{3}{TAM}, ztz{3}{3}, beta{3},
        yl{TAM}, e{TAM};

    /*
    calculo de los parametros de la regresion lineal multiple
    y = beta[0]*z0 + beta[1]*z1 + beta[2]*z2
    variables:

    n -numero de ternas z1,z2,y
    z1,z2,y -valores de las variables independientes
```

```

i,j,k -contadores
z -matriz asociada a z0=1, z1 y z2 (datos)
zt -transpuesta de la matriz z
ztz -producto zt*z
beta -coeficiente del modelo
sumal -valor usado para calcular los elementos de ztz
izt -valor usado para calcular los elementos de (ztz)^(-1)*zt
izty -valor usado para calcular los elementos de izt*y
yl -vector calculado a partir de los valores estimados de beta
zb -valor usado para calcular z*beta
e -error del valor y y el valor estimado yl (y-yl)
ete -producto del error e por su transpuesta et
ym -media aritmetica del vector y
y_ym -valor usado para calcular el cuadrado de y-ym
yl_ym -valor usado para calcular la diferencia yl-ym
varianza -valor estimado por ete/(n-r-1);
n numero de ternas
r rango de las matrices (en este caso r=3)
R2 -coeficiente de correlacion multiple calculado por
[sumatoria (yl-ym)^2]/[sumatoria (y-ym)^2]
*/

// transformacion de los vectores z0=1,z1,z2 a matrices
for (i=0; i<n; i++) {
  for (j=0; j<3; j++) {
    if (j == 0)
      z[i][0]=1.0;
    if (j == 1)
      z[i][1]=z1[i];
    if (j == 2)
      z[i][2]=z2[i];
  }
}

// construccion de la matriz transpuesta de z
for (i=0; i<n; i++)
  for (j=0; j<3; j++) {
    zt[j][i]=z[i][j];
  }

// construccion de la matriz zt*z
for (k=0; k<3; k++) {
  for (j=0; j<3; j++) {
    sumal=0.0;
    for (i=0; i<n; i++) {
      sumal+=z[i][k]*zt[j][i];
    }
    ztz[j][k]=sumal;
  }
}

// funcion para calcular la matriz inversa de ztz
m_inv(3,ztz);

cout << " =====>> endl
<< " = constantes calculadas para el modelo lineal =>> endl
<< " = >> endl
<< " = y = beta0 + beta1*z1 + beta2*z2 =>> endl
<< " =====>> endl
<< endl;

// construccion de los valores de los coeficientes beta
for (i = 0; i < 3; i++) {
  izty=0.0;
  for (j = 0; j < n; j++) {
    izt=0.0;
    for (k = 0; k < 3; k++){
      izt += ztz[i][k]*zt[k][j];
    }
    izty += izt*y[j];
  }
}

```

```

    beta[i] = izty;
}

*param0=beta[0];
*param1=beta[1];
*param2=beta[2];

cout << "  beta " << 0 << " = " << *param0 << endl
    << "  beta " << 1 << " = " << *param1 << endl
    << "  beta " << 2 << " = " << *param2 << endl;

// calculo del vector estimado y1 utilizando los valores beta
for (i = 0; i < n; i++) {
    zb=0.0;
    for (j = 0; j < 3; j++) {
        zb += z[i][j]*beta[j];
    }
    y1[i]= zb;
}

// funcion para calcular la media aritmetica de y
ym=y_media(n,y);

// construccion del vector error y calculo de las diferencias
// (y-ym)^2 y (y1-ym)^2
for (i = 0; i < n; i++) {
    e[i]=y[i]-y1[i];
    y_ym+=(y[i]-ym)*(y[i]-ym);
    y1_ym+=(y1[i]-ym)*(y1[i]-ym);
}

// calculo del producto et*e
for (i = 0; i<n; i++)
    ete+=e[i]*e[i];

// calculo de la varianza
varianza=ete/(n-3);
cout << endl << "  varianza (s^2)          = " << varianza << endl;
cout << "  desv. estandar (s)           = " << (sqrt(varianza)) << endl;

// calculo del coeficiente de correlacion multiple
R2=y1_ym/y_ym;

cout << "  coef. corr. multiple (R^2)= " << R2 << endl;

return 0;
}

// funcion para calcular la matriz inversa
int m_inv(int n, double a[3][3])
{
    /**** prototipo de la funcion ****/

    // funcion para resolver la matriz por el metodo de Gauss
    void gauss(int, double [3][6]);

    /**** declaracion de variables ****/
    int i, j, k;
    double a[3][6];

    /*
    calculo de la matriz inversa

    variables:

    n -numero de valores
    al -valores de la matriz a invertir
    i,j, i, l, k -contadores
    a -valores de la matriz a manipular

```

```

*/
// matriz identidad
a[0][3]=1.0; a[0][4]=0.0; a[0][5]=0.0;
a[1][3]=0.0; a[1][4]=1.0; a[1][5]=0.0;
a[2][3]=0.0; a[2][4]=0.0; a[2][5]=1.0;

// construccion de la matriz a
for (i = 0; i < n; i++) {
    for (j = 0; j < n; j++) {
        a[i][j]=a1[i][j];
    }
}

// funcion para resolver la matriz a
gauss(n, a);

// asignacion de los resultados a la matriz al
for (i = 0; i < n; i++) {
    k=0;
    for (j = n; j < (n*2); j++) {
        al[i][k]=a[i][j];
        ++k;
    }
}
return 0;
}

// funcion para resolver un sistema de ecuaciones por el metodo de Gauss
void gauss(int num, double mat[3][6])
{
    /**** declaracion de variables ****/
    int i, j, jc, jr, k, kc, m, nv, pv;
    double det, eps, eps1, eps2, r, temp, tm, va;

    /*
    resolucion de un sistema de ecuaciones lineales por el metodo de Gauss
    variables:

    num -numero de valores
    mat -valores de la matriz a resolver
    i,j,jc,jr,k,kc,m,nv,pv -contadores
    det -valor del determinante
    eps,eps1,eps2 -valores asociados con la epsilon de la maquina
    r,temp,tm,va -valores utilizados para resolver por renglon el sistema
    */

    while (eps1 > 0) {
        eps = eps/2.0;
        eps1 = eps*0.98+1.0;
        eps1+=-1;
    }
    eps *= 2;
    eps2 = eps*2;
    det = 1.0;
    for (i = 0; i < (num-1); i++) {
        pv = i;
        for (j = i+1; j < num; j++)
            if ((fabs(mat[j][i]) > (fabs(mat[pv][i])))
                pv = j;
        if (pv != i) {
            for (jc = 0; jc < (num*2); jc++) {
                tm = mat[i][jc];
                mat[i][jc] = mat[pv][jc];
                mat[pv][jc] = tm;
            }
        }
    }
}

```

```

    det = -det;
}
if (det == 0)
    cout << endl << "matriz singular" << endl;
for (jr = i+1; jr < num; jr++) {
    if (mat[jr][i] != 0) {
        r = mat[jr][i]/mat[i][i];
        for (kc = i+1; kc < (num*2); kc++) {
            temp = mat[jr][kc];
            mat[jr][kc] = mat[jr][kc] - r*mat[i][kc];
            if (fabs(mat[jr][kc]) < eps2*temp)
                mat[jr][kc] = 0.0;
        }
    }
}
}
for (i = 0; i < num; i++)
    det *= mat[i][i];

if (mat[num-1][num-1] != 0) {
    for (m = num; m < (num*2); m++) {
        mat[num-1][m] = mat[num-1][m]/mat[num-1][num-1];
        for (nv = num-2; nv >= 0; nv--) {
            va = mat[nv][m];
            for (k = nv+1; k < num; k++)
                va += -mat[nv][k]*mat[k][m];
            mat[nv][m] = va/mat[nv][nv];
        }
    }
}
}
}

```

```

// funcion para calcular la media aritmetica de y
double y_media(int numdat, double * vx)
{

```

```

/**** declaracion de variables ****/

```

```

double suma=0.0;

```

```

/*

```

```

    calculo de la media aritmetica de y

```

```

    variables:

```

```

    numdat -numero de valores

```

```

    vx -valor particular

```

```

    suma -suma de los valores vx

```

```

*/

```

```

for (int i = 0; i < numdat; i++)

```

```

    suma+=vx[i];

```

```

return (suma/numdat);
}

```

r_l_m.cpp . Programa central que llama a los programas anteriores.

```

#include <fstream.h>
#include <conio.h>
#include <iomanip.h>
#include <stdio.h>
#include <math.h>

```

```

#include "lect.h"
#include "crea.h"
#include "translog.h"
#include "rlm.h"

const int TAM = 25;

main()
{
    /*** declaracion de variables ***/

    char opc,o,nomb_arch[30],cte;
    int opcion, num, c;
    double valor1[TAM], valor2[TAM], valor3[TAM], coefa, coefb,
        coefc;

    /*
    variables en el programa principal:

    P -variable del tipo param (calculo de parametros estadisticos)
    opc,opcion -caracter y numero para direccionar control del programa
    num -numero de valores
    c
    */

    cout << endl << endl
    << "      rrrrrr      ll      mmm mmm uu  uu ll      tttttttt "
    << endl
    << "      rr rrr      ll      mmmmm mmmmm uu  uu ll      tt  "
    << endl
    << "      rrrrrr --- ll      --- mm mmm mm uu  uu ll      tt  "
    << endl
    << "      rr rr      ll      mm m mm uu  uu ll      tt  "
    << endl
    << "      rr rr      llll      mm      mm uuuuuuu llll      tt  "
    << endl << endl << endl << endl << endl
    << "      PROGRAMA QUE CALCULA LOS PARAMETROS DEL MODELO LINEAL : "
    << endl
    << "                      Z = a*X + b*Y + c                      "
    << endl
    << "      DONDE Z ES LA VARIABLE DEPENDIENTE, X-Y SON LAS VARIABLES "
    << endl
    << "      INDEPENDIENTES, a-b-c SON LOS VALORES CONSTANTES PARA LA "
    << endl
    << "      REGRESION LINEAL MULTIPLE."
    << endl << endl << endl
    << "<oprimir tecla 'retorno'>" << endl;
    getchar();
    clrscr();

    do {
        cout << "Seleccionar una de las opciones siguientes:"
        << endl << endl
        << "      (0)...terminar la sesion "
        << endl
        << "      (1)...captura de datos de un archivo externo "
        << endl
        << "      (2)...creacion de un archivo de datos "
        << endl
        << "      (3)...transformacion logaritmica de los valores"
        << endl
        << "      (4)...calculo de los coeficientes a, b, c"
        << endl << endl
        << "====> ";

        cin >> opcion;
        clrscr();
        if (opcion != 0) {
            switch (opcion){
                case 1:

```

```

cout << "      ===== \n"
  << "      lectura de datos desde un archivo externo \n"
  << "      ===== \n"
  << endl << endl
  << endl << "primer archivo a leer (variable X)" << endl;
lectura(&num, valor1);
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
cout << endl << "segundo archivo a leer (variable Y)" << endl;
lectura(&num, valor2);
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
cout << endl << "tercer archivo a leer (variable Z)" << endl;
lectura(&num, valor3);
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
break;
case 2:
cout << "      ===== \n"
  << "      creacion de archivo de entrada \n"
  << "      ===== \n"
  << endl << endl
  << "Si se pretende usar la combinacion lineal, se requieren tres \n"
  << "archivos de entrada; en los casos de suma, resta y division \n"
  << "se requieren dos. \n"
  << endl
  << "Indicar el numero de archivos a crear: \n"
  << "      (U)no \n"
  << "      (D)os \n"
  << "      (T)res \n"
  << endl << "====> ";
cin >> opc;
if ((opc == 'u') || (opc == 'U'))
  crear();
if ((opc == 'd') || (opc == 'D')) {
  cout << "primera muestra" << endl;
  crear();
  clrscr();
  cout << "segunda muestra" << endl;
  crear();
}
if ((opc == 't') || (opc == 'T')) {
  cout << "primera muestra" << endl;
  crear();
  clrscr();
  cout << "segunda muestra" << endl;
  crear();
  clrscr();
  cout << "tercera muestra" << endl;
  crear();
}
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
break;
case 3:
cout << "      ===== \n"
  << "      transformacion logaritmica de los valores \n"
  << "      ===== \n"
  << endl << endl;
transflog(num, valor1);
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
transflog(num, valor2);
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();

```

```

transflog(num, valor3);
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
break;
case 4:
cout << "      ===== \n"
<< "      calculo de los coeficientes a, b y c \n"
<< "      ===== \n"
<< endl << endl;

r_l_mult(num, valor1, valor2, valor3, &coefa, &coefb, &coefc);
cout << "coefa= " << coefa << " ,coefb= " << coefb << " , coefc= "
<< coefc << endl;
for (c=0; c<num; c++) {
    valor3[c]=coefa*valor1[c]+coefb*valor2[c]+coefc;
}
cout << "coefa= " << coefa << " ,coefb= " << coefb << " , coefc= "
<< coefc << endl;
cout << endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
break;
default:
cout << "opcion incorrecta, intentar de nuevo" << endl
<< endl << "para continuar oprimir tecla <retorno>";
getchar();
clrscr();
break;
}
}
while (opcion != 0);

return 0;
)

```

A.2 GUIA PARA EL USO DE LOS PROGRAMAS EJECUTABLES.

En esta sección se presentan los pasos que hay que seguir una vez que se tienen los datos listos para ser procesados y que se cuenta con la información de los estimadores de los parámetros de ubicación, escala y asociación para el modelo logístico bivariado. Todos los proyectos tienen una versión ejecutable que puede ser utilizada en el manejador de archivos de "windows"; también pueden ser ejecutados a través de alguno de los compiladores de C++ que sean compatibles. Al final del apéndice se encuentran los archivos de datos utilizados por cada uno de los programas.

Programa "suma.exe".

Si se utiliza el manejador de archivos en la ventana principal de "windows", sólo se selecciona el nombre del archivo con extensión ".exe" y se oprime la tecla <retorno>.

Aparece una primera pantalla de presentación y al oprimir la tecla 'retorno' se llega a la pantalla del menú general de opciones. Si se desea un análisis de máximos, el orden de las opciones es el siguiente: si se van a crear los archivos, entonces se empieza con la opción 2; una vez creados se pasa a la captura, opción 1; el paso que sigue es la suma de las dos muestras, opción 5; la segunda muestra capturada y la muestra formada por la suma se arreglan de mayor a menor, calculando la probabilidad asociada por la expresión de Weibull, opción 3; el paso siguiente es calcular los estadísticos muestrales y capturar los estimadores de los parámetros del modelo logístico bivariado, dejando sólo un espacio entre cada valor (no colocar ningún otro carácter, sólo dar el espacio con la barra espaciadora y al final del quinto valor oprimir la tecla 'retorno'), opción 6; con los pasos anteriores realizados, se procede al calculo de la probabilidad para los datos de la suma (debido a los esquemas de integración y búsqueda, esta parte es muy lenta), opción 7; finalmente, para estimar los valores de suma para diferentes periodos de retorno propuestos, se

selecciona la opción 8. En caso de análisis de mínimos, el orden de las opciones es el mismo sólo que se tiene que cambiar la opción 3 por la opción 4. En las opciones de 1 a 4 y la 6, se puede hacer uso de una sola muestra.

Seleccionar una de las opciones siguientes:

- (0)...terminar la sesion
- (1)...captura de datos de un archivo externo
- (2)...creacion de un archivo de datos
- (3)...arreglo de muestras para analisis de maximos
- (4)...arreglo de muestras para analisis de minimos
- (5)...suma de variables aleatorias
- (6)...calculo de parametros estadisticos
- (7)...calculo de la probabilidad de los datos
- (8)...calculo de la probabilidad de valores propuestos

====>

El valor máximo de datos es 51. Si se requiere un número mayor, se tendran que modificar los programas "integ_s.cpp", "suma_var.cpp" en la variable TAM, localizada en la parte superior del codigo, en ambos casos. Dependiendo de la computadora, se pueden presentar problemas para ejecutarlo debido básicamente al desborde de pila ("stack overflow"); esto limita fuertemente el número de datos a manipular.

Programa "resta.exe".

Es muy similar al programa anterior, con las variantes comentadas a continuación. La pantalla del menú general de opciones cambia únicamente en la opción 5. Si se desea un análisis de máximos, el orden de las opciones es el siguiente: si se van a crear los archivos, entonces se empieza con la opción 2; una vez creados se pasa a la captura, opción 1; el paso que sigue es la resta de las dos muestras, opción 5; la segunda muestra capturada y la muestra formada por la resta se arreglan de mayor a menor, calculando la probabilidad asociada por la expresión de Weibull, opción 3; el paso siguiente es calcular los parámetros muestrales estadísticos y capturar los estimadores de los parámetros del modelo logístico bivariado, dejando sólo un espacio entre cada valor (no colocar

ningun otro caracter, sólo dar el espacio con la barra espaciadora y al final del quinto valor oprimir la tecla 'retorno'), opción 6; con los pasos anteriores realizados, se procede al cálculo de la probabilidad para los datos de la suma (debido a los esquemas de integración y búsqueda, esta parte es muy lenta), opción 7; finalmente, para estimar los valores de suma para diferentes periodos de retorno propuestos, se selecciona la opción 8. En caso de análisis de mínimos, el orden de las opciones es el mismo sólo que se tiene que cambiar la opción 3 por la opción 4. En las opciones de 1 a 4 y la 6, se puede hacer uso de una sola muestra.

Seleccionar una de las opciones siguientes:

- (0)...terminar la sesion
- (1)...captura de datos de un archivo externo
- (2)...creacion de un archivo de datos
- (3)...arreglo de muestras para analisis de maximos
- (4)...arreglo de muestras para analisis de minimos
- (5)...resta de variables aleatorias
- (6)...calculo de parametros estadisticos
- (7)...calculo de la probabilidad de los datos
- (8)...calculo de la probabilidad de valores propuestos

====>

Programa "div_var.exe".

Como en el caso anterior, es muy similar al de "suma.exe", con las variantes siguientes. En el menú de opciones generales, la opción 5 lleva a cabo la división. En lo que respecta al orden de ejecución es el mismo: primero la opción 2, luego opción 1, opción 5, opción 3, opción 6, opción 7, opción 8. Para el caso de análisis de mínimos se cambia la opción 3 por la opción 4.

Seleccionar una de las opciones siguientes:

- (0)...terminar la sesion
- (1)...captura de datos de un archivo externo
- (2)...creacion de un archivo de datos
- (3)...arreglo de muestras para analisis de maximos
- (4)...arreglo de muestras para analisis de minimos
- (5)...division de variables aleatorias
- (6)...calculo de parametros estadisticos
- (7)...calculo de la probabilidad de los datos
- (8)...calculo de la probabilidad de valores propuestos

====>

Programa "r_l_mult.exe".

Este programa tiene que ser utilizado antes de hacer el análisis para la combinación lineal de variables aleatorias. Una vez activado el programa y ya situado en la pantalla del menú general, se procede a la creación de los archivos, opción 2; luego se pasa a la captura, opción 1; si se desean transformar las muestras, se selecciona la opción 3; finalmente, se calculan los coeficientes de la regresión lineal múltiple, opción 4.

Seleccionar una de las opciones siguientes:

- (0)...terminar la sesion
- (1)...captura de datos de un archivo externo
- (2)...creacion de un archivo de datos
- (3)...transformacion logaritmica de los valores
- (4)...calculo de los coeficientes a,b,c

=====>

Programa "comb_lin.exe".

Una vez determinados los coeficientes de la regresión lineal múltiple, se utiliza el programa "comb_lin.exe", que presenta la pantalla del menú general de opciones como el mostrado a continuación

Seleccionar una de las opciones siguientes:

- (0)...terminar la sesion
- (1)...captura de datos de un archivo externo
- (2)...arreglo de muestras para analisis de maximos
- (3)...arreglo de muestras para analisis de minimos
- (4)...combinacion lineal de variables aleatorias
- (5)...calculo de parametros estadisticos
- (6)...calculo de la probabilidad de los datos
- (7)...calculo de la probabilidad de valores propuestos

=====>

El orden de las opciones para el análisis de máximos es el siguiente: se capturan las tres muestras por medio de la opción 1; después se pasa a la combinación lineal, donde también se capturan los coeficientes calculados de la regresión lineal múltiple, opción 4; con esta información se pasa al arreglo de valores de mayor a

capturado
(volumen referencia)

capturado
(volumen descarga)

51
127.51
158.71
52.9
174.99
188.55
97.67
94.95
63.75
108.52
215.68
108.52
200.76
177.7
153.28
170.92
122.08
62.4
139.72
124.8
123.44
132.94
151.93
94.95
124.8
89.53
219.75
131.87
131.58
92.24
210.25
120.73
89.53
221.11
147.86
268.59
75.96
134.29
153.29
260.44
52.9
191.26
55.62
131.58
39.34
90.88
115.3
188.55
111.23
86.81
115.3
28.49

51
67.39
85.54
33.7
72.58
83.81
47.52
56.16
40.61
69.12
103.68
49.25
77.76
74.3
70.85
81.22
33.7
39.74
52.7
53.57
50.98
64.8
44.06
60.48
37.15
57.02
123.55
63.00
83.81
44.06
71.71
59.62
57.02
3.11
94.18
171.07
38.02
54.43
97.63
114.05
33.7
94.18
35.42
83.81
25.06
57.88
73.44
87.26
5.01
43.2
73.44
18.44

<=== número de datos

valores de volúmenes

valores de volúmenes

DIVISION.

Primer archivo
capturado
(Volumen escurrimiento)

Segundo archivo
capturado
(Gasto máximo)

50	50	<=== número de datos
455.779	2184.8	
666.895	2531.0	
2886.819	14376.0	
930.848	2580.0	
486.625	1499.0	
301.557	1164.8	
295.895	1127.3	
483.744	3215.0	
3873.441	10000.0	
535.079	3229.2	
121.566	677.0	
300.625	1266.0	valores de gastos y
331.502	1025.0	escurrimientos
77.237	954.8	
837.579	4780.3	
107.702	695.7	
245.627	593.0	
550.028	3010.0	
337.053	1908.0	
2395.953	15000.0	
595.438	1396.3	
310.512	1620.0	
866.747	2702.0	
324.517	1319.0	
363.128	1944.0	
494.948	2420.0	
606.884	2505.8	
206.798	1534.3	
312.877	1508.0	
380.084	1558.0	
788.517	2200.0	
460.274	2225.0	
1329.317	7960.0	
807.709	3790.0	
225.405	1080.0	
685.851	3240.0	
200.659	1135.0	valores de gastos y
875.674	4790.0	escurrimientos
1130.679	6860.0	
344.059	1496.0	
874.841	4828.0	
448.419	2450.0	
1667.443	8279.2	
1429.244	5580.0	
817.297	3585.0	
1689.452	2081.5	
191.025	1025.4	
476.523	3145.0	
139.308	1496.5	

2314.64

11704.0

COMBINACION LINEAL.

Primer archivo capturado (Lámina lluvia)	Segundo archivo capturado (Gasto antec.)	Tercer archivo capturado (Gasto máximo)	
21	21	21	<=== número de datos
117.0	644.0	4480.0	
91.2	89.0	30558.0	
173.9	429.0	24507.0	
176.4	96.0	14799.0	
230.6	70.0	34435.0	
164.5	62.0	37549.0	
206.5	81.0	37535.0	
188.8	84.0	73681.0	
171.9	122.0	31098.0	valores de lámina de
157.6	562.0	191899.0	lluvia, gasto antecedente
202.1	441.0	54380.0	y gasto máximo
85.4	414.0	8469.0	
183.9	307.0	121429.0	
177.9	1048.0	33378.0	
280.7	1014.0	40233.0	
218.6	1302.0	22009.0	
187.4	656.0	46638.0	
180.9	734.0	29902.0	
187.3	212.0	85780.0	
205.2	1001.0	58943.0	
175.5	1029.0	118133.0	

APENDICE B.

**HIDROGRAMAS DE GASTOS DE AVENIDAS DE LA ESTACION
HUITES, SINALOA.**

B.1 REPRESENTACIÓN TABULAR DE LOS GASTOS MÁXIMOS
MENSUALES PARA EL PERIODO 1941-1993 (GASTOS EN m3/s).

MES	AÑO						
	1941	1942	1943	1944	1945	1946	1947
Enero		37	66.6	768	197.41	62.2	185.9
Febrero		177.11	17.96	2580	24.3	14.4	82
Marzo		166.73	11.49	976	13.45	8.55	126.55
Abril		13.14	7.1	55	7.88	8.55	16.96
Mayo		8.6	6.07	17.36	5.05	4.63	16.2
Junio		68.73	414.2	89.05	5.45	151.4	4.05
Julio		780.3	416.4	274.82	1499.2	445	718.8
Agosto		2531	1085	474.6	1191.3	1164.8	1127.33
Septiembre	1668	1868.8	1396.9	1024.75	268.4	808.4	754.8
Octubre	756	2037.6	3283	52.96	1250	427.5	51.6
Noviembre	34	427.6	45	285.2	19.8	17.4	100.2
Diciembre	2084.8	226	14376	1262.5	11.24	19.41	634

MES	AÑO						
	1948	1949	1950	1951	1952	1953	1954
Enero	26.55	10000	3229.25	37.6	65.31	68.33	540.1
Febrero	3215	416.22	961	22.74	10	11.55	22.5
Marzo	104.05	47.45	28.91	13.46	355.8	39.02	9.24
Abril	13.62	24.11	12.6	11.59	11.81	7.06	6.19
Mayo	8.96	14.52	8.88	5.58	7.68	4.21	5.03
Junio	118.77	141	261.2	33.3	204.9	57.25	171.4
Julio	554.6	2297.5	1384	322.23	895	1025	481.24
Agosto	799	826.4	438.95	587.47	1266	885	954.8
Septiembre	244.2	942.4	438.95	234	187.55	587.2	715
Octubre	493.5	304.8	156	37.37	32.43	25.4	406.55
Noviembre	25.95	45.57	17	27	238.3	25.75	26.74
Diciembre	623.2	895.33	11.22	677	223.5	11.23	12.08

MES	AÑO						
	1955	1956	1957	1958	1959	1960	1961
Enero	4780.28	204.8	156.2	91.88	25.62	15000	1396.25
Febrero	35.8	21.7	79.9	114.2	652	102.2	136.65
Marzo	18.93	9.52	83.2	894	24.3	29.25	20.7
Abril	10	5.16	14.94	33.6	14.31	14.1	11.6
Mayo	4.92	17.25	4.99	92.94	5.42	6.02	7.02
Junio	23.23	278.6	26.64	307.05	108.75	15.38	208.9
Julio	662	493.2	330.9	608.5	544.3	721.4	558.6
Agosto	1069.88	695.7	380	1045.2	1345.5	1046	905.9
Septiembre	551.16	531.73	489	3010	397.9	790	771.2
Octubre	561.75	50.83	593	849	1908	140.5	682
Noviembre	21.09	10.38	22.73	286.85	51.8	24.05	831.6
Diciembre	31	11.44	362	136.7	1831	985.2	56.23

MES	AÑO						
	1962	1963	1964	1965	1966	1967	1968
Enero	912	27.28	50.05	30.93	188.8	15	401.6
Febrero	101.5	58.76	22.1	360	340	18.9	1118
Marzo	28.1	15.66	25.4	52.78	661.24	10.63	703.8
Abril	11.6	6.59	14.22	9.85	18.32	8.55	232.8
Mayo	8.43	4.43	34.38	5.35	8.76	9.3	38
Junio	64.74	32.9	105	32.1	688.89	272.6	100.51
Julio	501	980.13	912.4	491.2	892.24	1192.75	1478.25
Agosto	374.5	969.2	1319.13	663.12	2420	1310.5	1534.33
Septiembre	892.8	1054	938.5	1787.6	738	462.5	1019.6
Octubre	1620	323	145.2	147.6	131.2	113.6	106
Noviembre	30.12	47.2	19.7	18.88	33.9	147.5	385.5
Diciembre	32.4	2702	216	1944	17.64	2505.8	156.8

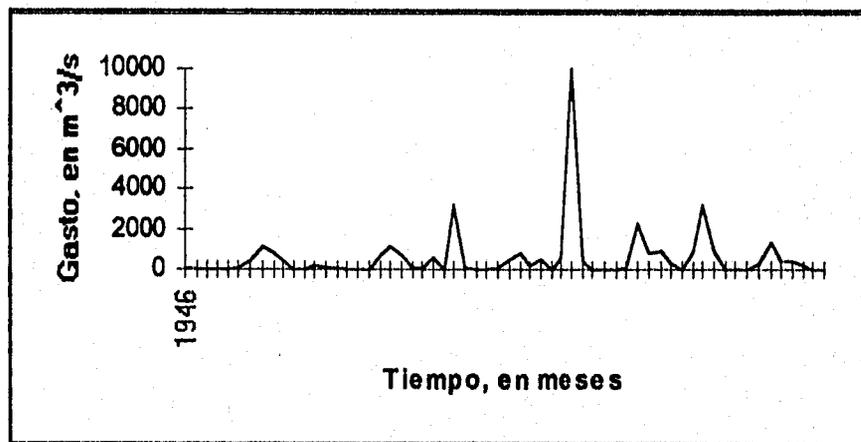
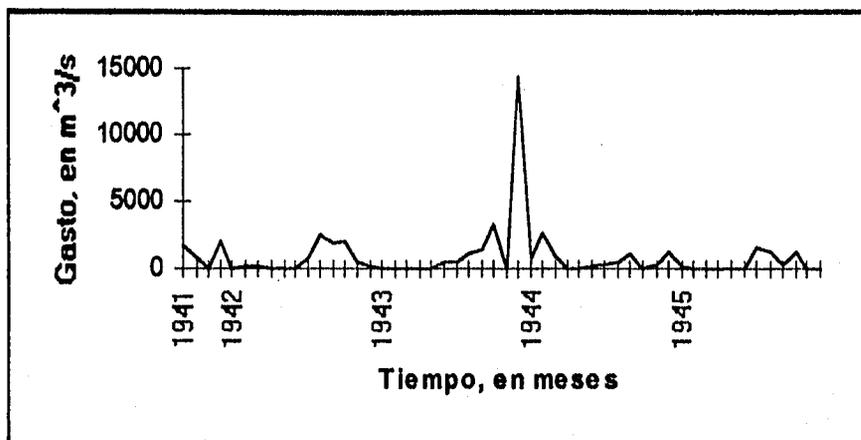
MES	AÑO						
	1969	1970	1971	1972	1973	1974	1975
Enero	68	136	25.97	732	2256.5	112.75	213.23
Febrero	300	37.94	9.39	23.96	7960	6.63	509.71
Marzo	39.4	206	9.2	16.82	372.5	6.26	15.92
Abril	13.65	29.5	5.19	10.43	77.25	4.11	13.18
Mayo	8.16	10.5	3.74	112.33	33.44	3.23	9.89
Junio	41.95	24	494.4	346	75.8	12.01	18.71
Julio	1508	1330	576.27	732.1	800	904.83	937.47
Agosto	417.6	1558	1176	1109	1546.1	988.73	960.89
Septiembre	504	970	591.2	1142	1380	3548.99	1066.9
Octubre	96	328.67	2200	2040	64.25	543.22	57.64
Noviembre	20.65	29.99	592	2225	13.38	4001	10.83
Diciembre	736	56.2	298	184.8	6.39	529.97	31.07

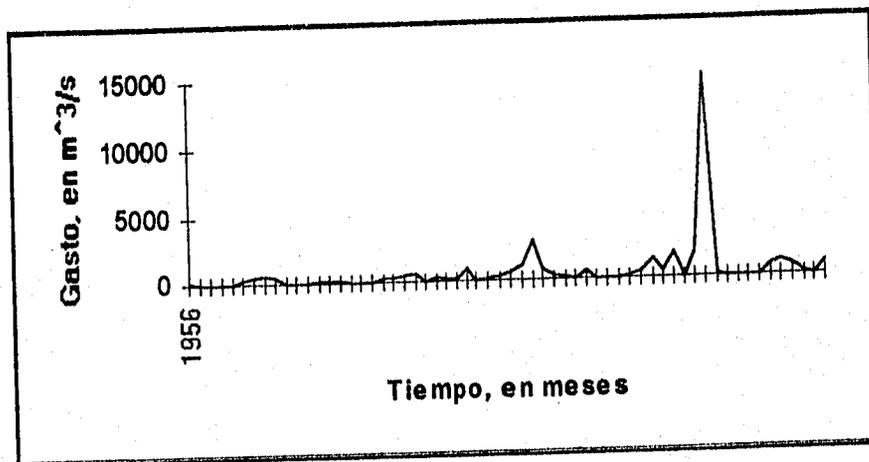
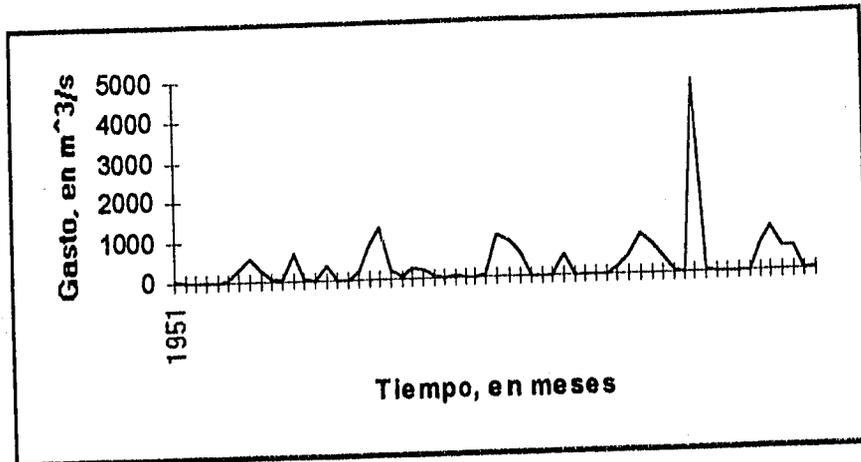
MES	AÑO						
	1976	1977	1978	1979	1980	1981	1982
Enero	15.69	92.58	15.35	3380.98	12.11	2039.73	251.58
Febrero	50.37	117.26	95.78	4443.03	14.9	67.48	21.51
Marzo	9.8	18.78	183.16	100.82	17.39	2433.74	8.72
Abril	8.26	25.96	13.85	31.35	14.17	71.6	6.83
Mayo	7.89	9.84	5.99	475.07	3.88	29.36	6.3
Junio	48.24	158.71	86.68	18.39	87.83	47	4.93
Julio	1070.71	551.27	477.35	789.64	450.73	783.39	301.69
Agosto	501.51	1118.88	714.45	1001.42	1473.54	934.45	407.21
Septiembre	1272.27	1080.66	6177.9	304	1451.87	2340.3	1264.52
Octubre	3233.41	264.17	1068.84	29.49	1051.37	2507.99	1530.25
Noviembre	259.05	25.6	120.2	10.14	65.83	68.58	797.68
Diciembre	66.18	15.62	1626.41	17.28	664.47	18.12	1498.54

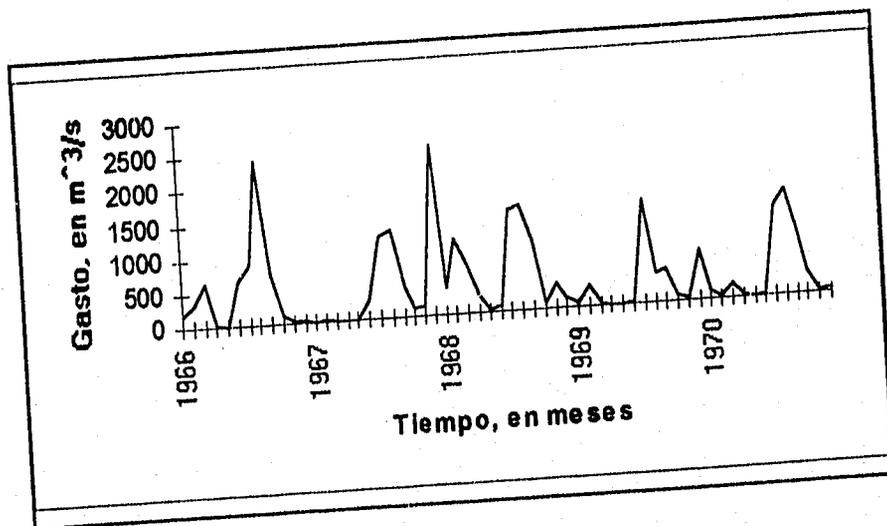
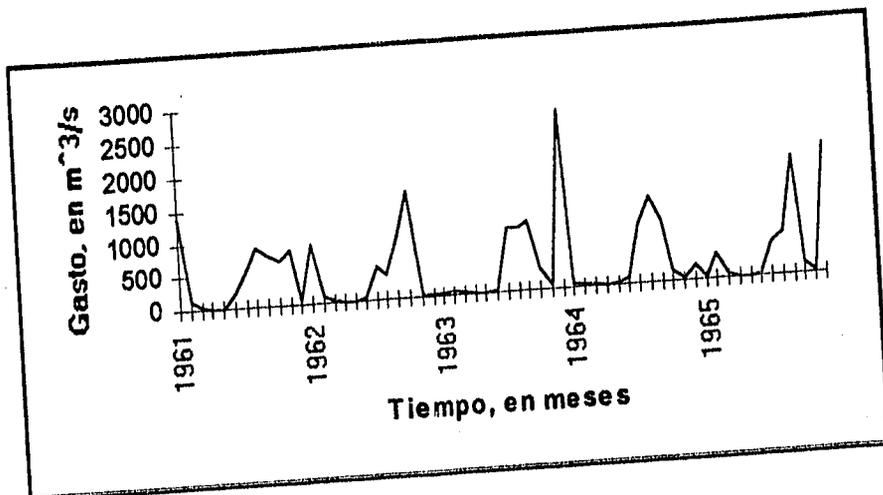
MES	AÑO						
	1983	1984	1985	1986	1987	1988	1989
Enero	797.92	138.75	3384.71	17.77	79	28.7	996
Febrero	1250.56	166.4	383.56	49.46	47	14.84	1216
Marzo	8000	11.9	149.33	19.14	31.68	10.55	38
Abril	206.17	7.67	129.72	12.7	14.19	6.67	22
Mayo	20.94	31.23	20.65	27.5	14.79	6.91	6.54
Junio	26.46	750.43	260.6	152.5	28.04	126.7	30.6
Julio	221.59	1004	1166.12	1374	1244.67	2299	236.87
Agosto	891.61	1623.29	1043.69	1348	1005.12	1753	1345.3
Septiembre	549.04	626.72	608.28	938	289.72	188.24	1332.38
Octubre	1098.02	348.63	652.23	366	76.01	119.2	71.88
Noviembre	296.33	81.48	167.27	65	10.58	17.32	102.68
Diciembre	1506.62	5496.41	65.27	323	154.67	20.62	200.82

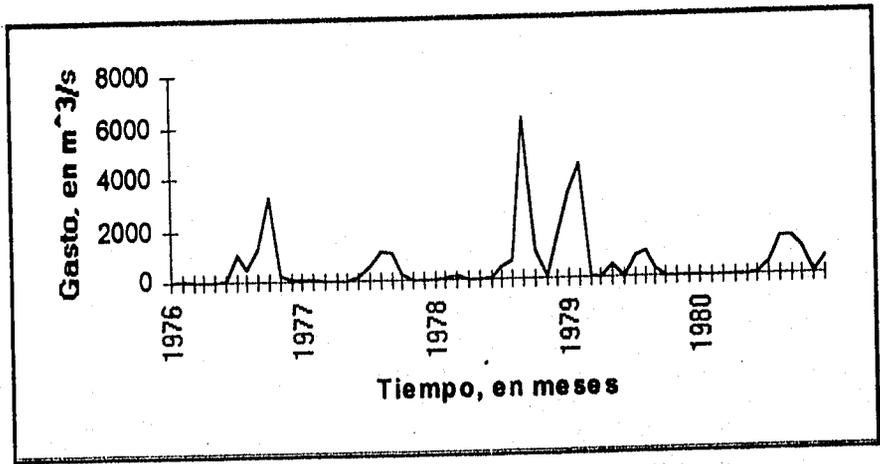
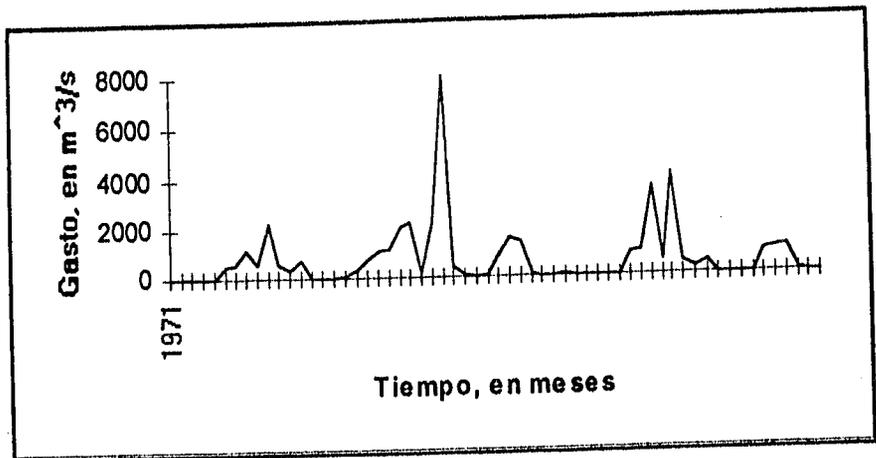
MES	AÑO			
	1990	1991	1992	1993
Enero	237.96	2509	2006	550
Febrero	425.26	196.59	1712	250
Marzo	287.15	504.27	489.38	1180
Abril	19.6	19.57	162	18
Mayo	7.62	12.36	185	12
Junio	102.56	7.6	19.9	55.4
Julio	939.57	1228.45	795.2	448
Agosto	682.51	1473.63	640	836
Septiembre	893.87	1467	400	
Octubre	3005	202	50	
Noviembre	165.32	456	20.5	
Diciembre	11350	1501	50	

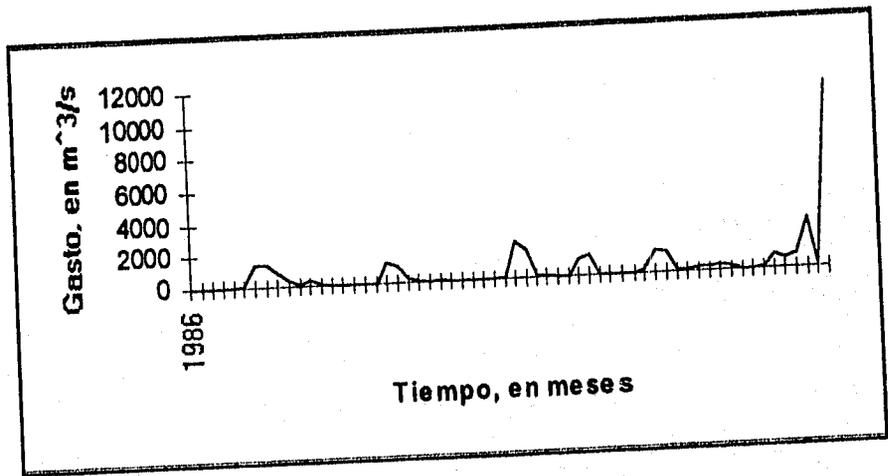
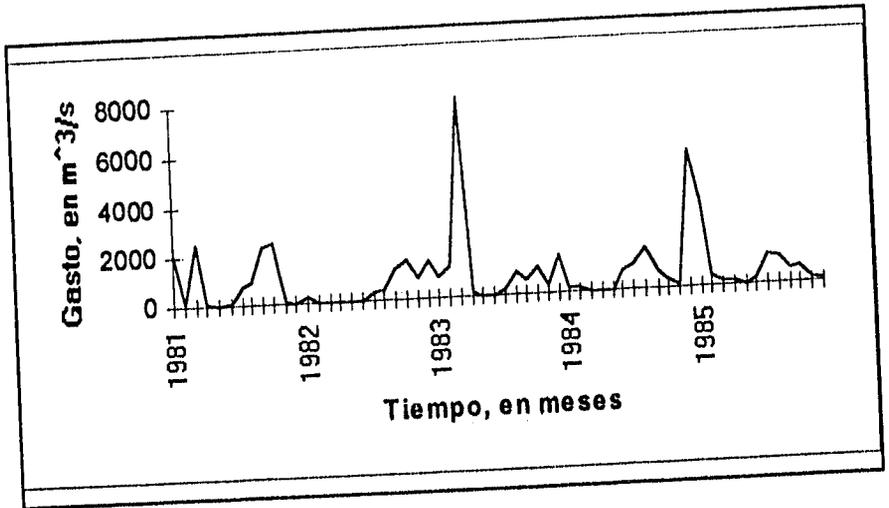
B.2 REPRESENTACIÓN ESQUEMÁTICA DE LOS GASTOS MÁXIMOS MENSUALES EN EL PERIODO 1941-1993. SE PUEDE APRECIAR EL COMPORTAMIENTO DEL GASTO MÁXIMO Y EL VOLUMEN ESCURRIDO ASOCIADO.

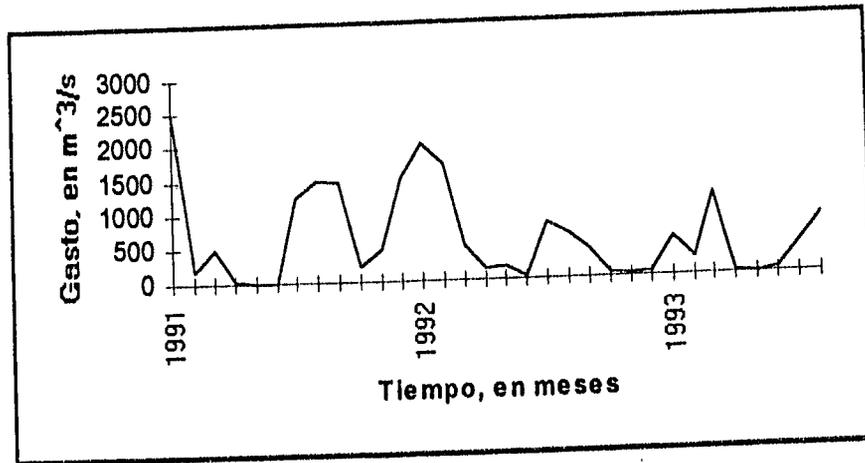




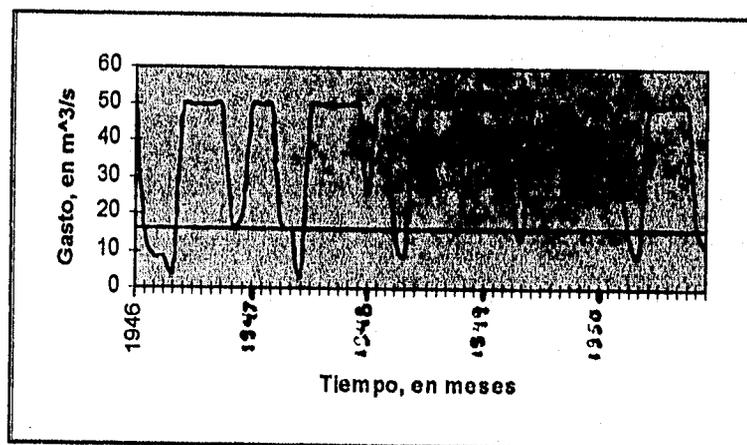
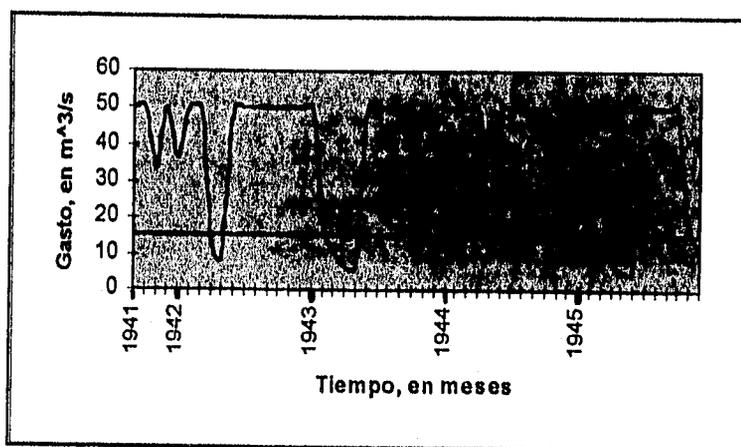


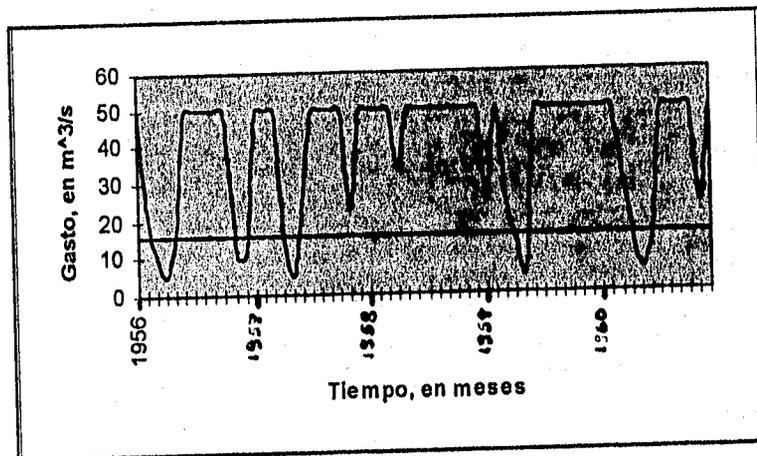
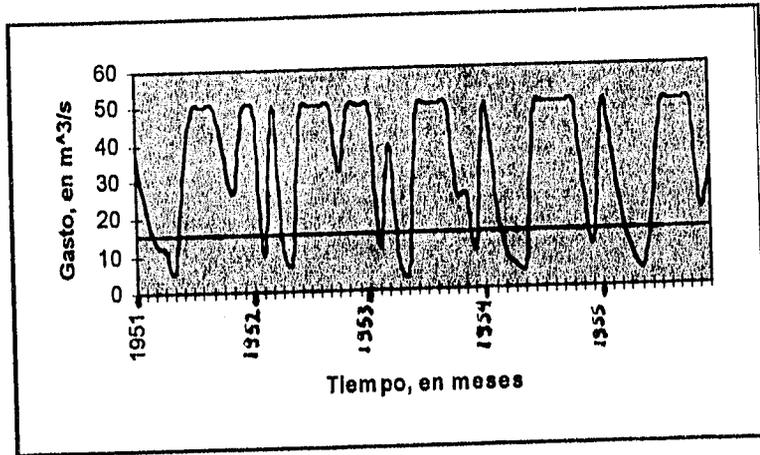


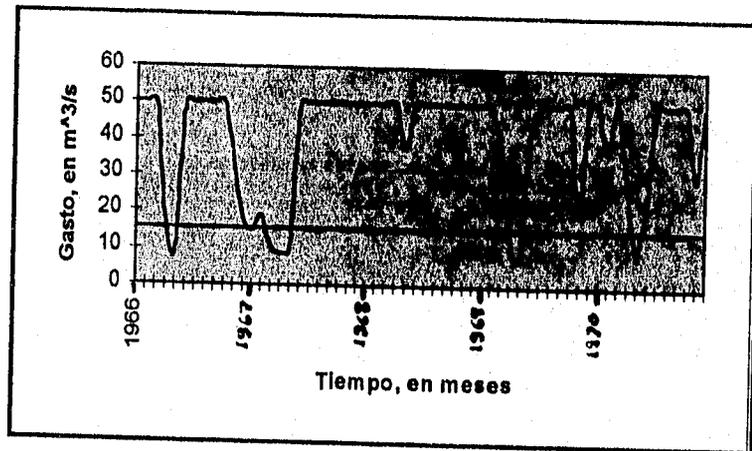
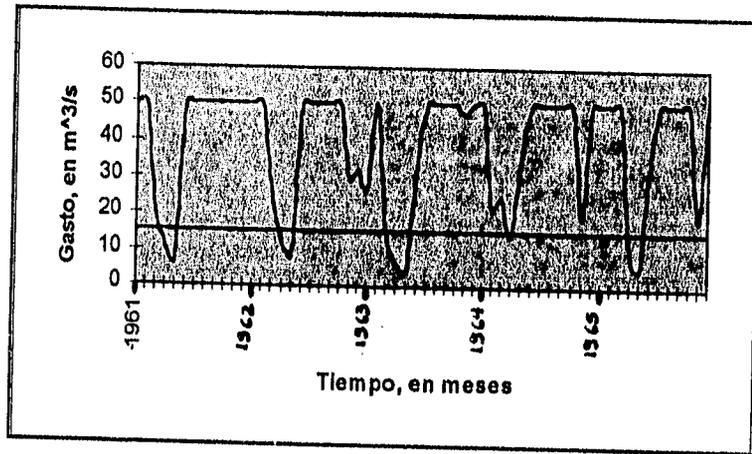


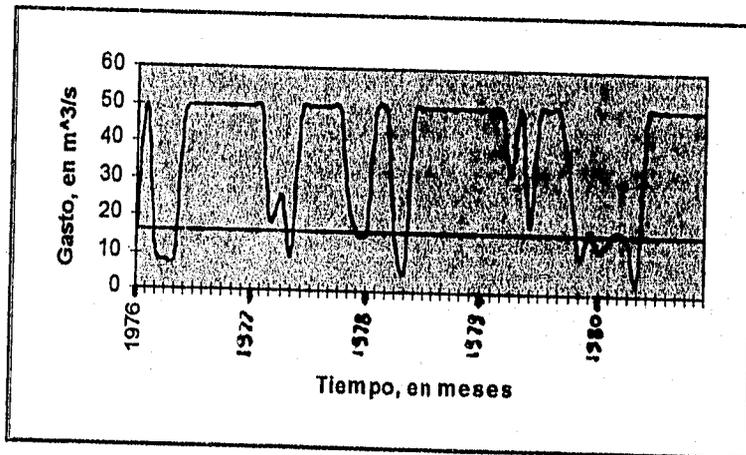
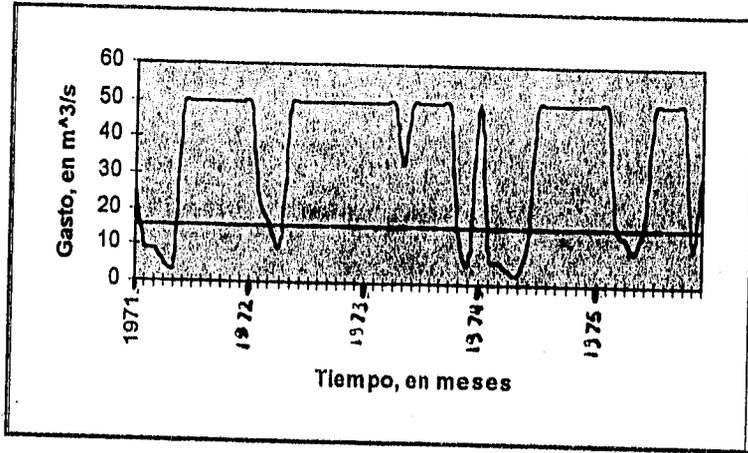


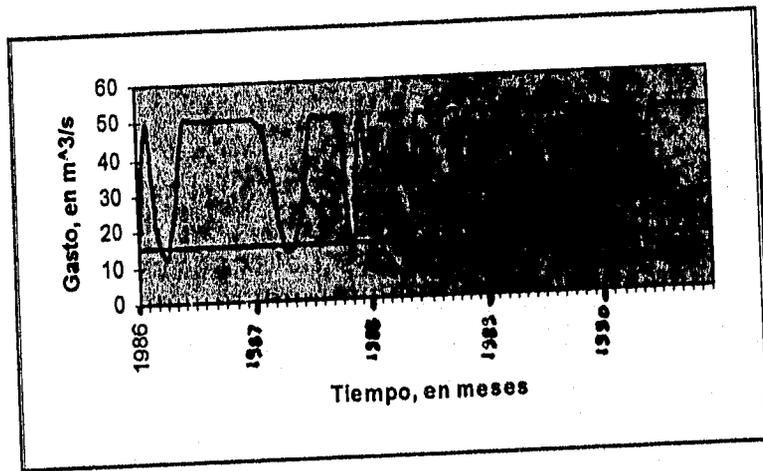
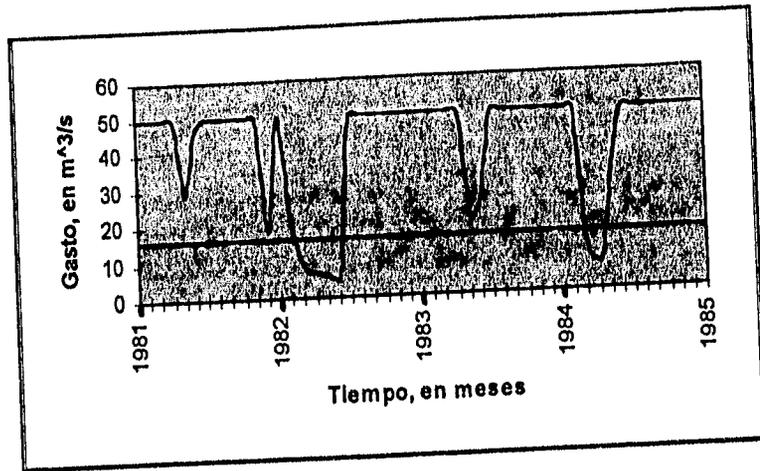
B.3 REPRESENTACIÓN ESQUEMÁTICA DE LOS ESCURRIMIENTOS MEDIOS MENSUALES MENORES A 50 m³/s. LA LÍNEA HORIZONTAL REPRESENTA EL GASTO DE REFERENCIA Y LA PARTE DE LA CURVA POR DEBAJO DE ÉSTA CORRESPONDE A LOS GASTOS ASOCIADOS CON EL FENÓMENO DE LA SEQUIA.

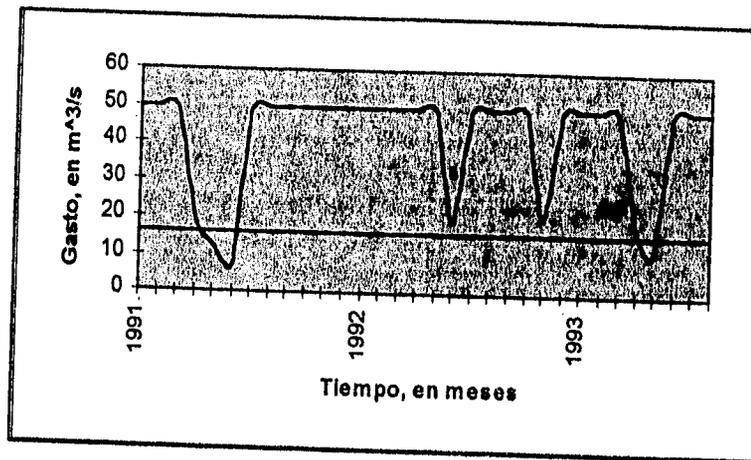












APENDICE C.

**EJEMPLOS DEL CALCULO DE LOS PARÁMETROS ESTADÍSTICOS
UTILIZADOS**

En este apéndice se muestra como se calculan los parámetros estadísticos utilizados en los capítulos 3, 4, 5 y 6. Los datos utilizados para calcular la media aritmética, la media geométrica, la varianza y la curtosis son los valores de gasto del río Dry Creek (20 valores expresados en m³/s, capítulo 3):

164.62

164.14

15.57

28.59

14.44

170.99

14.64

35.1

55.2

77.29

14.24

19.82

82.1

5.38

96.68

15.85

125.7

3.34

229.88

114.66

Cálculo de la media aritmética.

$$x_m = (164.62 + 164.14 + 15.57 + 28.59 + 14.44 + 170.99 + 14.64 + 35.1 + 55.2 + 77.29 + 14.24 + 19.82 + 82.1 + 5.38 + 96.68 + 15.85 + 125.7 + 3.34 + 229.88 + 114.66) / 20 = 211.379$$

Cálculo de la media geométrica.

$$x_G = (164.62 * 164.14 * 15.57 * 28.59 * 14.44 * 170.99 * 14.64 * 35.1 * 55.2 * 77.29 * 14.24 * 19.82 * 82.1 * 5.38 * 96.68 * 15.85 * 125.7 * 3.34 * 229.88 * 114.66)^{1/20} = 135.2167$$

Cálculo de la varianza.

$$s^2 = [(164.62-211.379)^2 + (164.14-211.379)^2 + (15.57-211.379)^2 + (28.59-211.379)^2 + (14.44-211.379)^2 + (170.99-211.379)^2 + (14.64-211.379)^2 + (35.1-211.379)^2 + (55.2-211.379)^2 + (77.29-211.379)^2 + (14.24-211.379)^2 + (19.82-211.379)^2 + (82.1-211.379)^2 + (5.38-211.379)^2 + (96.68-211.379)^2 + (15.85-211.379)^2 + (125.7-211.379)^2 + (3.34-211.379)^2 + (229.88-211.379)^2 + (114.66-211.379)^2] \{1/(20-1)\} = 27020.11$$

Cálculo de la curtosis.

$$C_s = [(164.62-211.379)^3 + (164.14-211.379)^3 + (15.57-211.379)^3 + (28.59-211.379)^3 + (14.44-211.379)^3 + (170.99-211.379)^3 + (14.64-211.379)^3 + (35.1-211.379)^3 + (55.2-211.379)^3 + (77.29-211.379)^3 + (14.24-211.379)^3 + (19.82-211.379)^3 + (82.1-211.379)^3 + (5.38-211.379)^3 + (96.68-211.379)^3 + (15.85-211.379)^3 + (125.7-211.379)^3 + (3.34-211.379)^3 + (229.88-211.379)^3 + (114.66-211.379)^3] \{20/[(20-1)(20-2)(164.378)^3]\} = 0.0223$$