

**ANALISIS ESPACIAL DE LA PRECIPITACION  
EN LA CUENCA DEL RIO YACAMBU**

**MARIA ISABEL ROJAS POLANCO**

TRABAJO DE GRADO

Presentado ante la

Universidad de Los Andes

En cumplimiento parcial de los requisitos

para optar al título de

**MAGISTER SCIENTIARUM**

Instituto de Estadística Aplicada y Computación.

Mérida, Febrero 1996

**DONACION**

**SERBIULA**  
Tulio Febres Cordero

*A Verónica y Héctor,  
la luz y el amor de mi vida.*

## AGRADECIMIENTO

A la Dirección de Hidrología y Meteorología del Ministerio del Ambiente y de los Recursos Naturales Renovables, quien a través de la Unidad de Estudio del Clima Tropical del Centro Interamericano de Desarrollo e Investigación Ambiental y Territorial brindó el apoyo financiero y de equipos necesarios para llevar a cabo este trabajo.

A la Dra. Lelys Bravo de Guenni, tutora de la tesis, por su valiosa amistad, interés, asesoramiento, colaboración, revisión y comentarios en el presente trabajo de grado. La distancia no fue motivo de ausencia .

Al Dr. M.F, Hutchinson, por permitir el uso de los softwares ANUDEM y ANUSPLIN, y brindar gentilmente su asesoramiento.

Al Ing. Abraham Salcedo, por su gran apoyo y confianza en la realización de este proyecto.

A los profesores Elizabeth Torres y Arnaldo Goitia, por su orientación y amistad durante el curso de post-grado

A mis compañeros del post- grado por su amistad

A todas aquellas personas que de alguna u otra manera contribuyeron a la realización de este Trabajo Especial de Grado.

## INDICE

AGRADECIMIENTO	.....	iii
LISTA DE TABLAS	.....	vi
LISTA DE FIGURAS	.....	vi
LISTA DE SIMBOLOS	.....	ix
<b>CAPITULO I :</b>	<b>INTRODUCCION</b> .....	<b>1</b>
	<b>OBJETIVOS</b> .....	<b>3</b>
<b>CAPITULO II:</b>	<b>ANTECEDENTES</b> .....	<b>4</b>
<b>CAPITULO III:</b>	<b>DESCRIPCION DEL MODELO</b> .....	<b>8</b>
3.1	Definición del modelo.....	8
3.2	Descripción de $e(x_i)$ .....	9
3.3	Estimación de la función de suavizamiento por Spline Laplaciano.....	10
3.4	Análisis estadístico.....	12
3.5	Validación Cruzada Generalizada.....	13
3.6	Determinación del ruido.....	13
3.7	Estimación de la varianza.....	14
3.8	Dócima de la bondad de ajuste.....	15
3.9	Explicación aproximada de la GCV.....	15
3.10	Spline Parcial Laplaciano.....	16
<b>CAPITULO IV :</b>	<b>CARACTERÍSTICAS GENERALES DE LA CUENCA DEL RIO YACAMBU</b>	
4.1	Cuenca del Río Yacambú.....	18
4.2	Red de estaciones pluviométricas.....	20
4.3	Controles climáticos.....	20
4.4	Importancia socio-económica.....	20

<b>CAPITULO V:</b>	<b>METODOLOGIA</b>	
5.1	Definición y estructura del Sistema de Información Geográfico.....	23
5.2	Aplicación y evaluación del modelo Spline Laplaciano.....	25
5.3	Producción de la superficie ajustada de precipitación....	27
5.4	Generación y corrección del Modelo Digital de Terreno.....	27
5.5	Selección del modelo.....	28
5.6	Optimización del número de estaciones.....	28
	5.6.1 Selección del número óptimo de estaciones.....	29
<b>CAPITULO VI:</b>	<b>ANALISIS DE RESULTADOS</b>	
6.1	Aplicación y evaluación del modelo Spline Laplaciano...	32
6.2	Producción de la superficie ajustada de precipitación....	41
6.3	Selección del modelo.....	49
6.4	Optimización del número de estaciones.....	49
	6.4.1 Muestreo con reemplazo.....	49
	6.4.2 Muestreo sin reemplazo.....	53
	6.4.2.1 Análisis gráfico.....	58
	6.4.2.2 Análisis de varianza.....	66
<b>CONCLUSIONES</b>	.....	71
<b>RECOMENDACIONES</b>	.....	74
<b>BIBLIOGRAFIA</b>	.....	75
<b>ANEXO 1:</b>	<b>PROGRAMAS UTILIZADOS, CARACTERISTICAS, ARCHIVOS DE ENTRADA Y SALIDAS.....</b>	<b>78</b>

## LISTA DE TABLAS

TABLA		pag.
1	Lista de estaciones pluviométricas.....	21
2	Análisis del orden de derivada.....	33
3	Análisis de bondad de ajuste del modelo.....	34
4	Evaluación del modelo spline laplaciano.....	35
5	Muestreo con reemplazo.....	52
6	Rango intercuartílico.....	56
7	Muestreo sin reemplazo.....	59
8	Bondad de ajuste para n subconjuntos.....	60
9	Cálculo de la prueba F aproximada para la comparación de modelos.....	66
10	Lista de estaciones seleccionadas.....	68

## LISTA DE FIGURAS

FIG		pag.
1	Cuenca del Rio Yacambú.....	19
2	Análisis Anual.....	37
2.1	Parámetro de suavizamiento.....	37
2.2	Grados de libertad.....	37
2.3	Error predictivo.....	37
2.4	% Precipitación pronosticada.....	37
2.5	Relación RGCV - Sigma.....	37
2.6	Varianza Estimada.....	37
2.7	Cuadrado Medio Residual.....	37
2.8	Verdadero error Cuadrático.....	37
3	Análisis Mensual.....	39
3.1	Precipitación media mensual.....	39
3.2	Parámetro de suavizamiento.....	39
3.3	Grados de libertad del error .....	39
3.4	Grados de libertad de la señal.....	39
3.5	Varianza Estimada.....	40
3.6	Cuadrado Medio Residual.....	40
3.7	Error predictivo.....	40
3.8	Verdadero error Cuadrático.....	40
4.1	Modelo Digital de Terreno ( proyección).....	42
4.2	Modelo Digital de Terreno ( raster ) .....	43
5.1	Precipitación Anual (mod. Bidimensional - Perp.).....	45
5.2	Precipitación Anual (mod. Bidimensional .) .....	46
6.1	Precipitación Anual (mod. Tridimensional - Perp.) .....	47
6.2	Precipitación Anual (mod. Tridimensional .) .....	48
7.1	Precipitación Anual (mod. Parcial - Perp.) .....	50
7.2	Precipitación Anual (mod. Parcial ) .....	51
8	Muestreo con Reemplazo.....	54
8.1	Precipitación media.....	54
8.2	Grados de libertad del error.....	54

8.3	Parámetro de suavizamiento.....	54
8.3	Grados de libertad de la señal.....	54
8.4	Varianza Estimada.....	54
8.5	Cuadrado Medio Residual.....	55
8.6	Error predictivo.....	55
8.7	Verdadero error Cuadrático.....	55
9	Clasificación de Estaciones en Muestreo con Reemplazo.....	57
10	Muestreo sin reemplazo.....	62
10.1	Variación de las medias.....	62
10.2	Variación del RHO.....	62
10.3	Varianza Estimada.....	62
10.4	Variación del error predictivo.....	62
10.5	Variación del cuadrado medio residual.....	62
10.6	‡ Error cuadrático verdadero.....	62
10.7	‡ Variación del RHO.....	63
10.8	Variación de GLE - GLS.....	63
10.9	Análisis de los grados de libertad.....	63
11	Precipitación anual (33 estaciones) .....	64
12	Precipitación anual (46 estaciones) .....	65
13	Estaciones Seleccionadas.....	69
14	Estaciones Seleccionadas (crem u srem) .....	70

## LISTA DE SIMBOLOS

### Símbolo

%PREC.	:	Porcentaje de precipitación predictiva
AVR	:	Promedio de las varianzas locales.
GCV	:	Error predictivo
gle	:	Grados de libertad del error
gls	:	Grados de libertad de la señal
mm / Kms	:	Milímetros de precipitación por altura, expresada en kilómetros.
MSE	:	Error cuadrático verdadero.
MSR	:	Cuadrado medio residual
n	:	Número de estaciones
RAVR	:	Raíz de la varianza relativa media
RTGCV	† :	Raíz del error predictivo
RTMSE	† :	Raíz error cuadrático verdadero
RTMSR	:	Raíz Cuadrado medio residual
RTVAR	:	Raíz Varianza estimada
SCR	:	Suma de cuadrados residuales
VAR	:	Varianza estimada
X <sup>2</sup>	:	Variable chi- cuadrado

QC 925.3  
R64

**SUMARIO**

**ANALISIS ESPACIAL DE LA PRECIPITACION**

**EN LA CUENCA DEL RIO YACAMBU**

por

**MARIA ISABEL ROJAS POLANCO**

En el presente trabajo se aplicó el modelo de interpolación por spline multidimensional con suavizamiento Laplaciano, para estimar la precipitación sobre la cuenca y optimizar la red de estaciones pluviométricas.

A través del programa ANUSPLIN, se aplicó y evaluó el modelo Spline Laplaciano, en forma bidimensional, tridimensional y parcial, tanto en datos de precipitación mensual como anual, luego se obtuvo en cada dimensión una superficie de precipitación anual ajustada al modelo digital de terreno que fue generado en el área de estudio utilizando el programa ANUDEM.

Después de seleccionar el modelo de mejor ajuste en la Cuenca, se diseñó una metodología con la que se logró reducir en un 45% la Red de estaciones pluviométricas fundamentada en el análisis de la variación de los grados de libertad de la señal y del ruido, del parámetro de suavizamiento y de los estadísticos que caracterizaban la superficie ajustada obtenida a través del modelo Spline Laplaciano en cada muestra de estaciones pluviométricas seleccionada y aplicando un procedimiento similar de eliminación hacia atrás que se realiza en una regresión lineal múltiple, usando una prueba F aproximada.

**SERBIULA - TULIO FEBRES CORDERO**



QC925.3 R64

# *CAPITULO I*

## *INTRODUCCION*

Venezuela es un país cuya economía, aparte de las industrias extractivas (mineras y petroleras), es altamente dependiente de la agricultura y por lo tanto de los Recursos Hídricos. De esta manera los aspectos meteorológicos que inciden en un proyecto hidráulico o agrícola, son quizás los más importantes, ya que de ellos dependen todos los demás factores.

La precipitación, caída durante un cierto intervalo de tiempo, es un parámetro de primera importancia para la Hidrometeorología, debido a su relación básica y fundamental en la planificación, diseño, operación de los recursos hídricos, tanto desde el punto de vista de diseños hidráulicos como del manejo del uso del agua: urbano, industrial, agrícola, generación eléctrica, navegación fluvial y recreación. Por lo tanto, es esencial, contar con buenos datos de precipitación, de manera de poder estimar o predecir su magnitud en un área determinada. Las mediciones de la precipitación se llevan a cabo en estaciones pluviométricas, las cuales están ubicadas en puntos fijos del espacio que se va a estudiar. La estimación de la lluvia media en el área, a partir de medidas puntuales implica un cierto grado de incertidumbre, que dependerá de la densidad y distribución de la red de estaciones; tamaño, naturaleza y topografía de la cuenca y del método adoptado para evaluar los datos.

Por consiguiente, para estimar la precipitación sobre una cuenca o de alguna parte significativa de ésta, se requiere realizar una interpolación e integración de los valores puntuales de lluvia sobre esa área.

En el presente trabajo se aplicó el modelo de interpolación por Spline Laplaciano para estimar la precipitación media en la Cuenca Yacambú, Edo. Lara, con un área de 396 Km<sup>2</sup>,

cuya red de puntos de observación está formada por 83 estaciones pluviométricas. El presente modelo, es un método de interpolación por spline multidimensional con suavizamiento Laplaciano cuyas soluciones fueron encontradas por Wahba (1979 ) y modificadas por Hutchinson (1987) con el fin de aplicarlas a las variables climáticas, minimizando el error de estimación de las mismas.

Para la aplicación del modelo, Hutchinson (1990) desarrolló un software denominado ANUSPLIN, que por primera vez es usado en Venezuela en condiciones de relieve irregular, y cuyos datos de entrada además de la precipitación utiliza información sobre el relieve de la Cuenca, el cual es generado en forma digital mediante un modelo digital de terreno que utiliza la información hidrográfica del área en estudio.

El modelo se analizó primero, en forma bidimensional considerando tan sólo la posición espacial de los datos, dada por su latitud y longitud; segundo, de manera tridimensional, incluyendo la altitud de los mismos y por último como un spline multidimensional parcial en el cual se incorpora la elevación como un sub-modelo lineal paramétrico.

Una vez obtenido el modelo que presentó el mejor ajuste en el área de estudio, se diseñó un método para optimizar el número de estaciones con la finalidad de encontrar una red básica mínima de puntos de observación, considerando en forma prioritaria, la necesidad de reducir los costos de operación y mantenimiento de la Red de estaciones pluviométricas, sin desmejorar la información necesaria para estudiar la variabilidad espacial de la precipitación en la Cuenca.

## ***OBJETIVOS***

1. Estudiar el modelo Spline Laplaciano en el análisis espacial de la precipitación en la Cuenca del Río Yacambú , mediante el software ANUSPLIN (Hutchinson 1990).
2. Optimizar el número de estaciones en la Red Pluviométrica del Sistema Yacambú - Quibor.
3. Analizar las ventajas y desventajas de utilizar un modelo estadístico de interpolación para la estimación de la precipitación en una cuenca.

## *CAPITULO II*

### *ANTECEDENTES*

Históricamente en la estimación de la precipitación media en una cuenca, se han aplicado tres métodos: gráficos, topográficos y numéricos.

#### A. Métodos gráficos.

Los métodos gráficos, usados tradicionalmente en Venezuela, que calculan la precipitación media de la cuenca, como un promedio ponderado cuyo peso está dado por el área de influencia que corresponde a cada estación pluviométrica, en el área total de la cuenca. Entre ellos se encuentran:

- a. Promedio Aritmético, el cual supone una lluvia uniformemente distribuida, considerando que todas las estaciones tienen peso igual a uno.
- b. Thiessen (1911), donde el peso de cada estación está dado por su área de influencia, determinada por el área de polígonos calculados a partir de la ubicación de las mismas. Este supone que las estaciones están bien distribuidas y en zonas planas.
- c. Isoyético que consiste en trazar líneas de puntos de igual precipitación (isoyetas) y luego ponderar en base al área entre isoyetas. Este método interpola las isolíneas de manera subjetiva, dependiendo por tanto de la experiencia del hombre para el trazado de las mismas.

#### B. Métodos topográficos.

Los métodos topográficos (Solomon S. (1986); Daly C.(1994)), correlacionan los datos puntuales de precipitación con un conjunto de variables tales como los parámetros

topográficos (altura, pendiente, etc. ) y los climáticos tales como la exposición al sol, velocidad y dirección del viento, . . , etc.

### C. Métodos numéricos.

A partir de la década de los 60, con el avance de los medios computacionales, se empiezan a usar los métodos numéricos, basados en el desarrollo de modelos de interpolación con el fin de resolver el problema de la estimación de variables climáticas a partir de puntos de observación distribuidos irregularmente en el espacio. Estos métodos pueden clasificarse en locales y globales.

Entre los métodos locales se destacan los trabajos realizados por: Coons(1969); Nielson(1974), McLain(1976); Powel (1976), Barnhill (1977); Lawson (1977); Akima (1978); McCullagh (1981); Sibson (1981), los cuales se basaban en particionar una superficie en elementos de dimensiones diferenciales y luego ajustar una función a cada elemento de forma tal que la función generada sobre toda la región analizada sea continuamente diferenciable. Las funciones definidas en cada elemento tendrán sus propias restricciones que dependerán de la forma misma del elemento, es por ello que la superficie ajustada resultante puede ser sensitiva a la posición de cada punto de observación, especialmente cuando éstos estén muy dispersos y la densidad de puntos sea muy baja. La desventaja de utilizar este tipo de modelo es lo difícil que resulta extenderlos a más de dos dimensiones y determinar el ruido en los datos.

Otros métodos de interpolación local fueron los presentados por: Cressman (1959); Shepard (1968); Goodin (1979); Bourke (1982), en donde se selecciona de manera subjetiva una función de ponderaje, definida en términos de un radio de influencia,

descartando aquellos puntos que se encuentren fuera del alcance de éste. En este tipo de técnica los grados de suavizamiento del error dependen de la función de ponderaje y no del ruido existente en los datos, por otra parte, es complicado encontrar un radio óptimo en regiones donde los puntos de observación estén muy dispersos. En 1979, Gandin desarrolla una técnica denominada ‘análisis objetivo óptimo’, la cual ha sido ampliamente usada en el campo de la Meteorología y consiste básicamente en analizar la varianza de la diferencia entre el estimador y el valor verdadero de cualquier variable meteorológica, suponiendo que el campo es estacionario, superando de esta forma la precisión de los métodos mencionados anteriormente.

Paralelo a las técnicas antes contempladas se desarrollaron los métodos globales, en los cuales la función de interpolación debe satisfacer, en general, los criterios de rotación invariante en curvatura mínima total. Bajo este criterio los métodos globales más conocidos son el ‘Kriging’ propuesto por Matheron en 1965 y el Spline Laplaciano que nace bajo la concepción del modelo Spline univariado para suavizar un conjunto de datos determinado, formulándose luego soluciones para el caso bivariado ( Duchon (1976) ), hasta definirse como un modelo de interpolación por spline multidimensional con suavizamiento Laplaciano, cuyas soluciones fueron encontradas por Wahba y Wendelberger (1980) y modificadas por Bates y Wahba (1982), Elden (1984) y Hutchinson (1987) para trabajar con conjuntos compuesto de grandes cantidades de datos espaciales.

Tanto el Kriging como el Spline Laplaciano están fundamentados en la Teoría de las Variables Regionalizadas, mediante las cuales se puede describir un fenómeno que se distribuye en el espacio y que muestra una cierta estructura espacial . Ambos métodos permiten solucionar y desarrollar los aspectos siguientes:

1. Elegir un estimador insesgado que considere la estructura espacial real del fenómeno.
2. Dan la posibilidad de evaluar la varianza del error de estimación en todo punto de interés.
3. Se arreglan de tal forma que el estimador sea óptimo.

Además, con la aplicación de ellos, se logra minimizar el error de interpolación y es fácil extenderse a dimensiones de orden mayor a tres. No obstante, la principal desventaja del Kriging según estudios realizados por Hutchinson M.F (1993) es que éste depende de la estimación previa de una función espacial de covarianza o variograma que es muy imprecisa en áreas con baja densidad de información. En el caso de que exista un buen variograma, los resultados serán similares al obtenido al realizar interpolación por Spline Laplaciano.

## CAPITULO III

### DESCRIPCION DEL MODELO DE SPLINE LAPLACIANO

El modelo de Spline Laplaciano es un método de interpolación multidimensional mediante el cual se puede estimar la cantidad de precipitación, en un punto particular o como el valor promedio de lluvia en un área determinada, en base a la información aportada por un conjunto de pluviómetros ubicados irregularmente en el espacio.

El objetivo del método es encontrar una función  $f$  continua de suavizamiento que se ajuste a los puntos de observación, una vez que se ha seleccionado tanto el orden de la derivada que define la rugosidad de la superficie como el parámetro de suavizamiento que es calculado al minimizar la validación cruzada generalizada (GCV) evaluada en el área de estudio.

La superficie estimada, no depende de un supuesto de normalidad en los valores observados, lo que representa una gran ventaja sobre otros métodos usados en la interpolación espacial de datos.

#### 3.1 DEFINICION DEL MODELO

Dadas las mediciones  $y(x_i)$  de precipitación en  $n$  puntos de observación ubicados irregularmente en un área geográfica, el modelo Spline Laplaciano se define como:

$$y(x_i) = f(x_i) + \varepsilon(x_i) \quad (i=1,2,\dots,n) \quad (3.1)$$

donde:

$f(x_i)$  : Son los valores de una función continua de suavizamiento que va a ser estimada a partir de las observaciones  $y(x_i)$ .

$x_i$  : Representan las coordenadas geográficas, latitud y longitud del punto de observación proyectadas en un plano euclidiano bidimensional, incluyendo la altitud del lugar para el caso trimensional

$\varepsilon(x_i)$  : Son variables aleatorias independientes, que se distribuyen normal con media cero y varianza  $\sigma^2$ .

### 3.2 DESCRIPCION DE $\varepsilon(x_i)$

Los errores  $\varepsilon(x_i)$  miden la desviación de la función de suavizamiento estimada, incluyendo tanto el error de medición aleatorio que no está correlacionado espacialmente como la variación a microescala que solamente está correlacionada en rangos muy pequeños por debajo de los límites de una observación normal.

Los  $\varepsilon(x_i)$  se definen como variables aleatorias independientes que se distribuyen normal con media cero y varianza  $\sigma^2$ . Se asume además que tienen una estructura de covarianza dada por :

$$E(\varepsilon \varepsilon^T) = \sigma^2 V \quad (3.2)$$

donde  $\varepsilon$  es el vector de los errores,  $V$  es una matriz de covarianzas, de orden  $n \times n$ , diagonal, positiva definida.

En el caso del análisis de datos de precipitación la varianza de cada observación  $y(x_i)$  estará dada por  $d_i \sigma^2$ , donde  $d_i$  es una constante de proporcionalidad que define la variabilidad de la lluvia caída en cada punto de observación y  $\sigma^2$  es generalmente desconocida.

### 3.3 ESTIMACION DE LA FUNCION $f$ DE SUAVIZAMIENTO POR SPLINE LAPLACIANO

Una estimación de  $f$  para resolver el modelo 3.1 se obtiene encontrando una función que minimice:

$$\sum_{i=1}^n \left[ \frac{y_i - f(x_i)}{d_i} \right]^2 + \rho J_m(f) \quad (3.3)$$

donde :

- \*  $d_i$  es un estimado local de la varianza de  $\epsilon_i$ . Para el conjunto de datos de precipitación promedio  $d_i = \text{var}_i / n_i$ , donde  $\text{var}_i$  es la varianza y  $n_i$  es el número de años de registro.
- \*  $J_m$ : es una medida de la rugosidad de la función  $f$  definida en términos de las derivadas parciales de  $m$ -ésimo orden. Esta medida de rugosidad (Hutchinson M.F, 1988) define la función de interpolación y está generalmente definidas en términos del primer y segundo orden de las derivadas parciales de  $f$ , dadas por:

$$J = \int (f_{xx}^2 + f_{xy}^2 + f_{yy}^2) dx dy \quad (3.4)$$

$$f_x = \frac{\partial f}{\partial x}; f_y = \frac{\partial f}{\partial y}$$

cuyo rango de integración está definido por la región ocupada por la grilla ajustada.

- \*  $\rho$ : es un parámetro de suavizamiento, definido como un número positivo que determina la falta de exactitud de los datos y la rugosidad de la superficie; generalmente es calculado minimizando la validación cruzada generalizada, la cual será definida en la sección 3.5

La función  $f$  debe ser una función 'suave' y continua, lo que significa que  $f, f', f'', \dots, f^{m-1}, f^m$ , existen. Entonces, recordando la formula de Taylor, tenemos:

$$f(t) = \left\{ \sum_{v=0}^{m-1} \frac{t^v}{v!} f^{(v)}(0) \right\} + \left\{ \int_0^t \frac{(t-u)^{m-1}}{(m-1)!} f^{(m)}(u) du \right\} \quad (3.5)$$

$$(x)_+ = \begin{cases} x & \text{si } x \geq 0 \\ 0 & \text{en otro caso} \end{cases}$$

Es decir que la función  $f$  se puede expresar como la suma de un polinomio y una función de cuadrado integrable :

$$f = f_0 + f_1 \quad (3.6)$$

\*  $f_0$ , es una función generada por polinomios de la forma

$$\phi_v(t) = \frac{t^{v-1}}{(v-1)!} \quad v=1,2,\dots,m \quad (3.7)$$

\*  $f_1$  es una función con derivadas hasta el orden  $m-1$ , absolutamente continuas.

Por lo que la función  $f$  se puede definir explícitamente como:

$$f(x) = \sum_{j=1}^M a_j \phi_j(x) + \sum_{i=1}^n b_i \psi(r_i) \quad (3.8)$$

donde  $\phi_j$  son un conjunto de  $M$  monomios de bajo orden y  $\psi$  es función escalar de la distancia euclidiana  $r_i$  entre  $x$  y  $x_i$ . Tanto  $M$  como la función escalar  $\psi$  dependen de la dimensión de  $x$  y del orden de la derivada  $m$ .

Los coeficientes  $b_i$  están restringidos a satisfacer las condiciones límites dadas por:

$$\sum_{i=1}^n b_i \phi_j(x) = 0 \quad (j = 1, 2, \dots, M) \quad (3.9)$$

Los coeficientes  $a_j$  y  $b_i$ , en 3.8 pueden ser calculados en orden de  $n^3$  operaciones, resolviendo el sistema lineal de orden  $n$ , dado por :

$$\begin{aligned} (K + \rho V)b + Ta &= y \\ T^T b &= 0 \end{aligned} \quad (3.10)$$

donde :

- $K$  es una matriz simétrica, positiva definida, la cual se genera a partir del efecto por rugosidad de la superficie en (3.3) y está dada por :

$$K_{ij} = \psi(\|x_i - x_j\|) \quad (3.11)$$

- $T$  es una matriz  $n \times M$ , definida por :

$$T_{ij} = \phi_j(x_i) \quad (3.12)$$

### 3.4 ANALISIS ESTADISTICO

La interpretación estadística del modelo de interpolación es parcialmente análoga al enfoque de mínimos cuadrados aplicado al análisis de regresión lineal. El Spline Laplaciano puede ser visto como una generalización no-paramétrica del análisis de regresión lineal.

Sea  $A$  una matriz de influencia de orden  $n \times n$ , mediante la cual se pueden expresar los valores ajustados  $f(x_i)$ , dados por 3.8, como una función lineal de los valores de  $y(x_i)$ , definida como:

$$f = Ay$$

donde:  $y = (y(x_1), y(x_2), y(x_3), \dots, y(x_n))^T$  y  $f = (f_1, f_2, f_3, \dots, f_n)^T$ , siendo

$$f_i = f(x_i, y_i, h_i) \quad (i = 1, 2, \dots, n)$$

### 3.5 VALIDACION CRUZADA GENERALIZADA ( GCV )

La validación cruzada generalizada es una medida del error predictivo dado en una superficie ajustada y puede ser calculada omitiendo cada dato puntual en turno y estimando el cuadrado medio residual entre los valores observados y los valores ajustados al resto de los puntos de observación.

La GCV se determina en forma implícita debido a que los valores ajustados dependen linealmente de los datos puntuales, resultando por ello una medida buena de la potencia predictiva de la superficie spline ajustada (Wahba 1990, Hutchinson y Gessler, 1994). Ella se puede definir como:

$$GCV = \frac{\left(\frac{1}{n}\right) \sum_{i=1}^n [(y_i - f_i)/d_i]^2}{\left[(1/n) \text{traza}(I - A)\right]^2} = \frac{(y - Ay)^T V^{-1} (y - Ay)/n}{[\text{traza}(I - A)/n]^2} \quad (3.13)$$

donde  $A$  es la matriz de influencia.

### 3.6 DETERMINACION DEL RUIDO

La traza de  $(I - A)$ , donde  $I$  es la matriz identidad y  $A$  la matriz de influencia, define los grados de libertad del ruido existente en los datos y puede interpretarse como los grados de libertad del cuadrado de la suma residual entre los valores observados y los ajustados.

Hutchinson y Gessler (1994), demostraron que la traza  $(I - A)$  es una herramienta útil de diagnóstico del modelo. En la mayoría de las aplicaciones de suavizamiento es

razonable esperar que los grados de libertad sean mayor a  $n/2$ . Cuando este no es el caso, es probable que los puntos de observación no sean lo suficientemente densos como para permitir una interpolación por Spline Laplaciano o que los errores  $\varepsilon(x_i)$  estén correlacionados positivamente. Otro camino eficiente para detectar errores en los datos es la presencia de residuos grandes en los datos puntuales calculados a partir del Spline Laplaciano ajustado

Una interpolación por spline, exacta, corresponde a tener cero grados de libertad en el ruido, lo cual es físicamente imposible encontrarlo. En el caso de que la traza  $(I - A)$  alcance el máximo valor  $n-M$ , el spline ajustado es simplemente una regresión de mínimos cuadrados de los datos en los  $M$  monomios  $\phi_j$ .

### 3.7 ESTIMACION DE LA VARIANZA.

Por analogía a una regresión lineal, la varianza del error puede ser estimada como:

$$\hat{\sigma}_\varepsilon^2 = \frac{(y - Ay)^T V^{-1} (y - Ay)}{\text{traza}(I - A)} \quad (3.14)$$

La varianza estimada de una función lineal de spline ajustado, incluye la estimación de la varianza en cualquiera de sus puntos. En particular Wahba (1983) demostró que la matriz de covarianza resultante de los valores ajustados en los datos puntuales está dada por  $A \sigma_\varepsilon^2$  donde  $A$  es la matriz de influencia. Usando datos simulados, Wahba encontró que el valor estimado de  $\sigma_\varepsilon^2$  dado en la ec.(3.14) aportó estimados aceptables para intervalos de confianza calculados a partir de la matriz de covarianza estimada.

El parámetro  $\sigma_\varepsilon^2$  se estima después que el splines ha sido ajustado, minimizando todo tipo de error de predicción, a través del GCV. Si la traza  $(I - A)$  es menor a  $n/2$  la GCV deberá ser mayor a  $2\hat{\sigma}_\varepsilon^2$ .

### 3.8 DOCIMA DE LA BONDAD DE AJUSTE DEL MODELO

El cuadrado de la suma residual :

$$\sum_{i=1}^n [(y_i - f_i) / d_i]^2 \quad (3.15)$$

es una muestra de  $\sigma^2 \chi^2_m$  variables, donde  $\sigma^2$  es la varianza,  $m = \text{traza}(I - A)$  y  $\chi^2$  es una variable chi-cuadrado con  $m$  grados de libertad.

Si tenemos disponible un valor de  $\sigma^2$  estimado, *a priori*, entonces los puntos críticos de  $\chi^2$  pueden ser usados para evaluar la bondad de ajuste del spline estimado.

### 3.9 EXPLICACION APROXIMADA DE LA GCV

Debido a que cada dato puntual tiene un ruido específico, cuando se utiliza para evaluar el error predictivo se genera una sobreestimación del error verdadero de la superficie ajustada. O sea que :

$$\text{GCV} = \text{TMSE} + \sigma^2 \quad (3.16)$$

donde TMSE es el verdadero error cuadrático medio en los datos puntuales.

Si  $\sigma^2$  es conocida, entonces TMSE es un estimador insesgado y estará dado por :

$$(1/n) \sum_{i=1}^n [(z_i - f_i) / d_i]^2 - (2/n) \text{traza}(I - A) \sigma^2 + \sigma^2 \quad (3.17)$$

Este podrá ser usado para determinar el parámetro de suavizamiento, pero es sensible a los errores en los valores de  $\sigma^2$ . Es preferible determinar el parámetro de suavizamiento minimizando GCV, la cual no requiere de ningún estimado de varianza, como fue descrito anteriormente

### 3.10 SPLINE PARCIAL LAPLACIANO

El Spline Parcial ha sido descrito detalladamente por Wahba (1990), como un caso especial del Spline Laplaciano. Un resumen de la metodología básica con la interpolación de variables climáticas, principalmente, puede ser hallado en Hutchinson (1993).

El modelo observacional de spline parcial para  $n$  valores  $z_j$  en la posición  $x_j$  está dado por:

$$z_i = f(x_i) + \sum_{j=1}^p \beta_j \psi_j(x_i) + \varepsilon_i \quad (3.18)$$

$i = 1, \dots, n ; j = 1, \dots, p$

donde  $f$  es una función de suavizamiento desconocida que va a ser estimada, los  $\psi_j$  son un conjunto de  $p$  funciones conocidas y  $\beta_j$  son un conjunto de parámetros que también tienen que ser estimados. Los  $x_j$  representan comúnmente las coordenadas en un espacio Euclidiano de dos o tres dimensiones. Los  $\varepsilon_i$  son errores aleatorios con media 0 y estructura de covarianza dada por:

$$E(\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^T) = V\sigma^2 \quad (3.19)$$

donde  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$ , y  $V$  es una matriz  $n \times n$  positiva definida y  $\sigma^2$  puede ser conocida o desconocida. Los errores  $\varepsilon_j$  están no correlacionados si  $V$  es diagonal y correlacionados para  $V$  no diagonal. La función  $f$  y los parámetros  $\beta_j$  son estimados minimizando :

$$(z - g)^T V^{-1} (z - g) + \rho J_m(f) \quad (3.20)$$

donde :

$z = (z_1, \dots, z_n)^T$  y  $g = (g_1, \dots, g_n)^T$ , siendo

$$g_i = g(y_i) = f(x_i) + \sum_{j=1}^p \beta_j \psi_j(y_i)$$

Las variables independientes  $\psi_{ij}$  son denominadas generalmente como covariables, estableciendo que el submodelo lineal está dado en forma paramétrica.

$J_m$  es una medida de la rugosidad de la función spline  $f$  definida en la misma forma del spline Laplaciano y  $\rho$  es un número positivo denominado parámetro de suavizamiento.

La solución del problema de minimización puede ser resuelto explícitamente con el estimado de  $f$  que tenga una expansión en términos de una función escalar de distancia, cuya forma depende de la dimensión de  $x_i$  y del orden  $m$  de la derivada, que defina la rugosidad (Wahba, 1990). Es de hacer notar que cuando no existe un submodelo paramétrico, la solución se reduce a un Spline Laplaciano.

## *CAPITULO IV*

### **CARACTERISTICAS GENERALES DE LA CUENCA YACAMBU**

#### **4.1 CUENCA DEL RIO YACAMBU**

La Cuenca del Río Yacambú (Fig.1) está ubicada en el estado Lara, municipio Andrés Eloy Blanco y se localiza entre las coordenadas 69°39' y 69°43' de longitud oeste y entre los 09°33' y los 09°38' de latitud norte. Físicamente está situada en la vertiente Sur de la Sierra de Portuguesa, tiene una extensión de 396 Km<sup>2</sup> y limita al Norte con la cuenca del río Turbio, al Noroeste con la cuenca del río Tocuyo, al Sur con el río Morador y al Sudeste con el río Guache. La Cuenca forma parte del sistema montañoso de la Cordillera de Los Andes, su relieve se caracteriza por ser predominantemente montañoso con alturas hasta de 2160 metros sobre el nivel del mar en contraste con áreas de pendiente suaves y valles estrechos, existiendo un gradiente altitudinal significativo entre la zona llana y la montañosa. Las principales filas montañosas presentan crestas agudas y fuertemente inclinadas, cubiertas de una densa vegetación. El drenaje de la zona lo forma un sistema irregular compuesto por el río Yacambú como colector principal y al cual vierten sus aguas otros ríos y quebradas, por la margen izquierda: El Barranco, San Carlos, El Puente, La Negra, La Escalera y Qda. Honda, y por la margen derecha: Los Palmares, Pedro Jiménez, Agua Blanca, Los Villegas y Cerro Blanco. Todas estas quebradas tienen pendientes muy pronunciadas, descendiendo su longitud relativamente corta desde su nacimiento, 2160 m.s.n.m hasta el cauce, cuya elevación promedio es 800 m.s.n.m y con un desarrollo general de un patrón de drenaje predominantemente del tipo dentrítico.



#### ***4.2 Red de Estaciones Pluviométricas***

Los datos de precipitación para esta cuenca se obtuvieron de la División de Hidrología y Meteorología del Ministerio del Ambiente y de los Recursos Naturales Renovables (MARNR). En la tabla 1 se incluye la totalidad de las estaciones utilizadas y la localización de las mismas se muestra en la figura 1. Ellas forman un conjunto de 83 estaciones, de las cuales 10 se encuentran dentro de la cuenca, 15 en su área inmediatamente adyacente y el resto en el área de influencia sobre el comportamiento hidrometeorológico de la misma .

#### ***4.3 Controles Climáticos***

La ubicación de la cuenca del río Yacambú dentro del sistema montañoso de la Cordillera Andina, hace que el clima regional sufra las influencias de la altitud, de la circulación de Los Alisios y de las masas de aire del sur. La Cordillera de Los Andes actúa como una barrera climática sobre la circulación atmosférica del occidente del país, produciendo importantes consecuencias en el clima. Durante el verano astronómico, estación lluviosa en Venezuela de abril/mayo a septiembre /octubre, la alta precipitación en la cuenca se explica por la incidencia de los vientos dominantes de la región, los Alisios del noreste, los cuales cargados de humedad, chocan en un ángulo favorable contra el sistema montañoso. Allí los vientos se ven obligados a ascender y como consecuencia a enfriarse, con lo cual se producen lluvias orográficas, las cuales se acentúan por el caldeamiento excesivo de la llanura de Barlovento, originando un tipo de lluvia orográfico convectivo. También durante el verano astronómico el área se encuentra bajo la influencia de las masas de aire del sur relacionadas con la Zona de Convergencia Intertropical y durante el invierno

TABLA 1

LISTADO DE ESTACIONES											
Serial	Estación	Tipo	Alt. msnm	Latitud	Longitud	Serial	Estación	Tipo	Alt.	Latitud	Longitud
1174	QUEBRADA ARRIBA	PR	615	101439	703153	2448	CENDE	PC	3000	93149	701005
1177	ALTAGRACIA	PR	479	102056	701214	2151	SABANA GRANDE	PR	200	84225	703156
1186	BURERE	PR	430	100522	701439	2152	CENIZO-SECUNDARIO D3	PR	62	83300	704530
1187	CARORA-GRANJA	C3	413	100946	700455	2201	CURARIGUA	PR	627	95957	695507
1202	BARQUISIMETO-FUNDACION-B22	PC	528	100453	691740	2204	QUIBOR	C2	682	95511	693739
1203	CAMBURAL	PR	550	102541	693731	2208	MAPORAL	PR	415	95435	691132
1204	CANAPE	PR	658	100111	693167	2215	SAN MIGUEL	PR	1087	95240	693103
1208	BARQUISIMETO-SE	PR	582	100512	691812	2216	SAN MIGUEL	C1	1005	95129	693043
1210	BARQUISIMETO-AL	PR	598	100523	692208	2219	MIRACUY	PR	1161	93744	693350
1220	AGUA NEGRA	PR	1560	94843	693036	2221	CRUZ MACHADERA	PR	1845	95134	692623
1221	LAS TABLAS	PR	1330	94949	692853	2225	QUIBOR	C3	1502	84728	693502
1222	EL ZANCUDDO	PR	1355	94700	693220	2226	CAPILLA BUCARAL	PR	1050	84844	692228
1265	BOBARE	PR	660	101725	692748	2228	AURO	PR	280	94239	691421
1288	DUACA	PR	750	101659	690849	2231	CASPIYO	PR	1300	93934	693943
1271	ARENALES	PR	428	100924	695423	2232	PARQUE YACAMBU	PR	1484	94230	693436
1272	BANCO DE BARAGUA	PR	787	100849	693531	2233	LA CRUZ	PR	1200	84500	692947
1273	TURTURA-SAN PABLO	PR	720	100833	694221	2234	SAANRE	PR	1330	84433	693824
1274	GUADALUPE	PR	582	100229	694043	2235	SAANRE	C2	1350	84400	693900
1277	HACIENDA CUJI-CARACARAS	PR	486	100513	702258	2236	LA ESCALERA	PR	1340	84300	693100
1278	EL CUJI	C1	620	101000	691900	2237	RIEOTO	PR	1280	95148	692147
1282	BARQUISIMETO-BA	SB	613	100400	691900	2239	AGUA BLANCA	PR	175	94001	690553
1295	YARITAGUA-EXPERIMENTAL	C1	375	100400	690700	2241	HUMOCARO BAJO	PR	841	94128	695718
1321	CAYURE	PR	370	102543	665934	2246	HACIENDA CAMBURIITO	PR	285	93603	691910
1322	ARDA DISPENSARIO	PR	278	102608	665333	2248	ARADRE	C1	200	93600	691300
1342	GUAMA	PR	340	101707	664901	2251	HUMOCARO ALTO-VIVERO	PR	950	93732	695917
1354	GUARABAO	C1	260	101454	684811	2252	LA MESA	PR	1195	93353	695953
1362	CAMPO ELIAS	PR	480	101317	665434	2253	HACIENDA GUACHE	PR	255	92537	692153
1371	URACHICHE	PR	435	100848	690039	2255	ACARIGUA	PR	188	83408	691313
1381	CENTRAL MATILDE	C3	235	100839	685211	2259	PAYARA	PR	152	87843	690443
2106	EL JARDIN	PR	1499	94845	700711	2264	ACARIGUA-AEROPUERTO	SP	226	93300	691400
2118	EL EMPEDRADO	PR	844	95205	701624	2287	LAS CUMBRES	PR	1588	95212	691925
2119	TRENTINO-LA PASTORA	PR	559	84844	701041	2290	LAS TROJAS	PR	1409	95156	691656
2126	WALLE HORNO	PR	800	84007	701657	2331	SAN RAFAEL DE ONOTO	PR	175	84052	685815
2127	PUENTE VILLEGAS	PR	621	94548	701435	2333	LA CEIBA-CAMORICO	PR	150	83903	684423
2128	AGUA DE OBISPOS	PC	2114	84133	700553	2334	LAS MAJAGUAS	C3	182	93913	685937
2129	SAN ANTONIO	PR	237	84338	702651	2340	COJEDITOS	PR	160	93847	685510
2132	AGUA VIVA	PR	110	93335	703731	2342	SANTA TERESA	PR	150	93123	685137
2133	MONAY	PR	292	93300	702732	2351	TOCUTANO	PR	140	93132	685737
2138	LA CONCEPCION	PC	1125	93300	701508	2357	GUADINAS	PR	188	94127	685255
2139	SABANA GRANDE-GUAGO	PR	1388	93514	695136	2362	GUAYABITO	PR	140	93021	684831
2141	CENIZO-CAMPAMENTO VIVIAN	PR	124	93207	704026	6332	EL TRIGAL	PR	1680	102140	684838
2146	GUAMAS DE MONAY	PR	272	93607	702500						

astronómico, período de sequía en Venezuela (noviembre/diciembre a abril), bajo la influencia de incursiones de frentes fríos modificados que invaden a Venezuela produciendo lluvias de extraordinaria magnitud fuera de la época. En la mayor parte de la cuenca, la estación lluviosa se observa entre los meses de abril a noviembre, presentando el máximo de precipitación en el mes de junio, y la estación seca de diciembre a marzo, registrándose el mínimo en el mes de enero. Su régimen pluviométrico es del tipo bimodal observándose un máximo principal en el mes de junio y un máximo secundario en el mes de octubre. La temperatura media anual en la cuenca es de 21.3° C y en lo referente a la evaporación, los máximos corresponden al mes de marzo y los mínimos a junio, siendo el valor promedio anual de 1371.8 mm

#### ***4.4 IMPORTANCIA SOCIO-ECONOMICA DE LA CUENCA YACAMBU.***

En la Cuenca del río Yacambú se está desarrollando el sistema hidráulico Yacambú-Quibor, el cual constituye una de las obras de infraestructura más importantes del estado Lara, ya que permitirá abastecer de agua la zona agrícola del Valle de Quibor y el área metropolitana e industrial de Barquisimeto, generando 33640 millones de bolívares. En el Valle de Quibor se produce pimentón, tomate, frutales, uva, caña de azúcar y muy especialmente cebolla. En este rubro casi el 60% de la producción nacional proviene del Edo. Lara, y el 90% de la cebolla sembrada en el estado se cosecha en Quibor, lo que dice bastante de la relevancia económica que representa para el país esta región. Se estima que el sistema Yacambú producirá 330 millones de metros cúbicos de agua por año, pero su buen funcionamiento y operación dependerá de la disponibilidad de agua.

## *CAPITULO VI*

### *METODOLOGIA*

En este capítulo se describen las diferentes etapas que fueron necesarias para poder lograr una metodología que permitiera los objetivos propuestos. Para ello en principio, se utilizó la estructura de un Sistema de Información Geográfico, con la finalidad de organizar y realizar todo lo inherente al procesamiento de datos; y luego, en base a una revisión bibliográfica, se investigó la forma de aplicar y evaluar el modelo Spline Laplaciano, tanto a nivel anual como mensual.

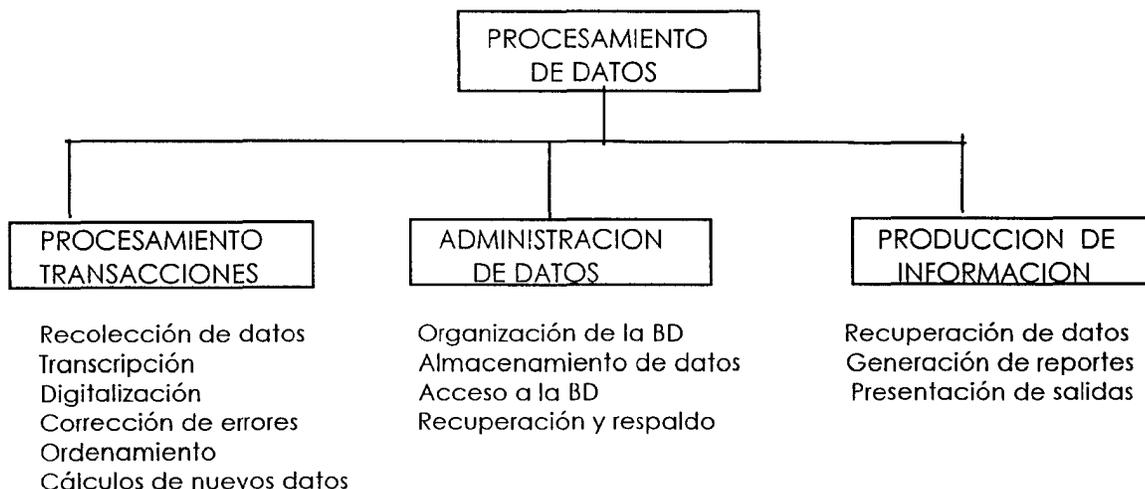
Por otra parte, se utilizó el programa ANUSPLIN para obtener el modelo digital de terreno del área en estudio sobre cual se evaluaron las superficies de precipitación obtenidas al aplicar el modelo Spline Laplaciano en forma bidimensional, tridimensional y parcial, para luego seleccionar el modelo que mejor se ajustase a la Cuenca del Río Yacambú

Por último, se diseñó una metodología gráfica y analítica para la optimización del número de estaciones pluviométricas, a fin de seleccionar un número mínimo de estaciones con el cual se pudiese estimar con un error mínimo la precipitación en el área Yacambú- Quibor, y reducir los costos por mantenimiento y operación de la red de puntos de observación.

#### 6.1 DEFINICIÓN Y ESTRUCTURA DEL SISTEMA DE INFORMACIÓN GEOGRAFICO

Un Sistema de Información Geográfico (SIG) es un sistema de hardware, software y procedimientos diseñados para apoyar la recolección, manejo, análisis y despliegue visual de datos referenciados espacialmente con el fin de resolver problemas de planificación .

Como en todo sistema de información su actividad principal fue el procesamiento de datos que fue definida según el diagrama siguiente:



La información básica estuvo conformada por registros de precipitación y la cartografía del área en estudio. La información pluviométrica consistió en datos mensuales y anuales de precipitación registrados entre 1970 y 1983 (14 años), por una red de 83 puntos de observación (estaciones pluviométricas), ubicados irregularmente en el espacio, según sus coordenadas geográficas y altura sobre el nivel del mar (cotas), dadas en la tabla 1.

La información cartográfica se obtuvo de las cartas N°6244-6245-6344-6345, elaboradas en 1976 por la Dirección de Cartografía Nacional, en Proyección Mercator Transversal a escala 1:100.000. A partir de esta información se crearon los mapas de relieve, ríos y perímetro de la Cuenca, los cuales fueron posteriormente digitalizados a través del ARC/Inform, usando para ello una tabla digitalizadora. Los ríos fueron digitalizados de aguas arriba hacia aguas abajo y las curvas de nivel cada 200 m.s.n.m, expresando toda la información en kilómetros

Una vez corregidos los errores por transcripción y digitalización, y calculados los valores puntuales promedios y varianzas locales de la precipitación, se conformó la base de datos con los archivos de entrada necesarios para la corrida del programa ANUSPLIN y ANUDEM.

El programa ANUDEM (anexo 1) requiere que los archivos de las curvas de nivel, la red hidrográfica y los puntos seleccionados con cotas conocidas sean exportados al Sistema de Información Geográfico IDRISI. En este sistema los archivos están en formato ASCII, del tipo vectorial y fue necesario pasarlos a la estructura de formato de los archivos de entrada requerida para la generación del modelo digital de terreno.

Las imágenes generadas por el ANUDEM y el ANUSPLIN, a su vez, formaron parte de la base de datos, usando el IDRISI para su manipulación y representación visual.

## 6.2 APLICACION Y EVALUACION DEL MODELO SPLINE LAPLACIANO

En la aplicación y evaluación del modelo se utilizó el programa SPLINA, incluido en el paquete ANUSPLIN (Anexo 1), el número de estaciones era moderado. El programa se aplicó a nivel mensual y anual; primero en forma bidimensional, considerando tan sólo la posición espacial de los datos, dada por su latitud y longitud; segundo, de manera tridimensional, incluyendo la altitud de los mismos y por último como un spline parcial, usando, en principio, la altura como una única covariable y luego incorporando otra covariable al modelo, utilizada para medir el efecto sobre la superficie estimada, de la posición de cada estación pluviométrica con respecto a la cuenca. Esta covariable fue definida como una variable dicotómica donde el valor (-1) indicaba que la estación se encontraba fuera de la cuenca y (1) en caso contrario.

A fin de evaluar el modelo, primero se aplicó una derivada de segundo y tercer orden en la definición de la rugosidad con el fin de analizar el efecto de ésta en la estimación de la función  $f$  referida en la sección 3.3, luego se probó la bondad de ajuste del modelo comparando la suma de cuadrados de residuales ( $n \cdot MSR$ , donde  $n$  es el número de datos puntuales) con los puntos críticos de una  $\hat{\sigma}_E^2 \cdot \chi^2$  con  $gle$  grados de libertad, siendo  $gle$  los grados de libertad del error. Por último, se analizaron los grados de libertad del

error y la señal, el parámetro de suavizamiento ( $RHO = n * \lambda$ ) y los estadísticos dados en los archivos de salida del SPLINA (Anexo 1), considerando en su análisis las proposiciones siguientes:

1. - En cada superficie, la suma de los grados de libertad del error y de la señal deberá ser igual al número de puntos de observación y su ajuste se considera incorrecto cuando el parámetro RHO sea muy pequeño y la señal muy grande (igual al número de puntos de observación) o un RHO muy grande y la señal muy pequeña (un número que depende del número de variables independientes y del orden de la derivada).
2. Después de un procedimiento de ajuste el valor de la varianza estimada (VAR) debe ser menor que uno, lo que significa que la varianza del error en la superficie ajustada no es mayor que la varianza promedio en los valores de los datos.
3. Si los grados de libertad del error es menor que  $n/2$  entonces la  $GCV > 2 * \hat{\sigma}_E^2$ .
4. El cuadrado medio residual (MSR) estará ponderado de acuerdo a las varianzas locales  $d_i$ , dadas en los archivos de entrada.
5. Las raíces cuadradas de VAR, GCV, MSR y MSE proveen una medida adecuada en milímetros de los valores calculados en estos estadísticos.
6. El porcentaje del error predictivo con respecto al valor promedio de precipitación medido en el conjunto total de estaciones fuese menor al 20% ( $\pm 3$ ).
7. La diferencia entre el valor observado y el estimado no fuese mayor al 20% del valor observado.

A nivel mensual, además, se caracterizó el comportamiento del modelo en relación a la cantidad de precipitación, comparando los valores de los estadísticos y de los otros parámetros en los meses secos y en los meses húmedos.

### **6.3 PRODUCCIÓN DE LA SUPERFICIE DE PRECIPITACIÓN.**

La superficie de interpolación generada a partir del modelo bidimensional, tridimensional y parcial fueron procesadas por el programa LAPGRD (Anexo 1), incluido dentro del ANUSPLIN, para lo cual se debió definir los archivos de entrada según su formato respectivo y luego procesar las imágenes en IDRISI, tanto en formato raster como en perspectiva sobre el modelo digital de terreno.

### **6.4 MODELO DIGITAL DE TERRENO.**

La generación del modelo digital de la cuenca fue realizada utilizando el programa ANUDEM (Hutchinson, 1990), para un área de 3275 Kms<sup>2</sup>, ubicada entre los 09°26' y 09°55' de latitud norte y los 69° 11' y 69° 44' de longitud oeste, cubriendo no sólo la cuenca sino también sus parteaguas oeste y este hasta alcanzar los piedemontes, con alturas comprendidas entre los 62 y 2200 metros sobre el nivel del mar

Los datos de elevación fueron dados en función de las curvas de nivel definidas como líneas de contornos, a las cuales se les aplicó el algoritmo de forzamiento de drenaje ya que la red de drenaje en el área es muy extensa y no fueron medidos los puntos de sumidero. Las líneas de contorno de las cimas fueron condicionadas solamente si sus longitudes eran menor que la cima máxima.

Por otra parte, el valor de la raíz del cuadrado medio residual de la grilla ajustada se condicionó a un 0.05% del intervalo entre las curvas de nivel, permitiendo que el valor máximo en la grilla ajustada necesariamente no estuviese restringido a caer en los puntos de cotas conocidas, y aplicando una penalidad por rugosidad de 0.0 para curvaturas mínimas. En cuanto a las tres tolerancias, permitidas en las diferentes altitudes (Anexo 1), se usaron

los valores según las recomendaciones ( Hutchinson M. y Guenni L., 1990) contempladas para mapas a escala 1:100.000, los cuales fueron : 20 para medir la exactitud en los datos de altura, 40 en la relación entre la primera tolerancia y el forzamiento de drenaje y 100 para el ajuste de los puntos de mínimo local que no son resueltos por el acondicionamiento de drenaje. El máximo número de iteraciones en los datos fue de 30 y el espaciamiento por retícula de 500 metros.

En total se digitalizaron 142980 puntos en curvas de nivel, de los cuales fueron aceptados el 87% ; y 29529 puntos en la red hidrográfica, de los que se reconocieron el 90 %. En relación a los puntos de cotas conocidos, los cuales fueron definidos en base a la altitud de cada estación pluviométrica, se determinó que el 36.1% presenta un residual en la altura superior a los 4 metros.

## **6.5 SELECCIÓN DEL MODELO**

La selección del modelo se realizó en base al mejor ajuste, basado en el menor error predictivo de la superficie generada y valores de los estadísticos utilizados en la evaluación del modelo.

## **6.6 OPTIMIZACIÓN DEL NÚMERO DE ESTACIONES:**

La optimización del número de estaciones en el presente estudio, significa conseguir un número mínimo de estaciones pluviométricas con las cuales en el Sistema Hidráulico Yacambú - Quibor se pueda estimar la precipitación con un mínimo de error, minimizando, de esta manera los costos por mantenimiento y operación de la red .

Para alcanzar este objetivo, en primer lugar se hizo un muestreo con reemplazo, seleccionando cada estación, una a una, manteniendo una muestra de 82 puntos de

observación, con el fin de analizar la variación de los grados de libertad del error y la señal, el parámetro de suavizamiento y los estadísticos que caracterizaban cada superficie interpolada generada por la muestra una vez que había sido extraída la estación pluviométrica; luego se trabajó con un muestreo sin reemplazo hasta llegar al mínimo de puntos de observaciones necesarios para que el error cuadrático medio no se incrementase en una cantidad estadísticamente significativa..

En este caso la extracción de cada estación, no se realizó en forma aleatoria sino que dependió del menor valor que presentase cada punto de observación en la raíz del cuadrado medio residual ponderado por su error estándar dentro de cada muestra, es decir, se eliminaba aquella estación cuyo error en la estimación fuese mínimo.

#### **6.6.1 SELECCION DEL NUMERO OPTIMO DE ESTACIONES**

La selección del “subconjunto óptimo “de estaciones se realizó primero en base a un análisis gráfico, donde se examinó la variación de los estadísticos, el parámetro de suavizamiento y los grados de libertad del error y de la señal que evalúan la superficie de interpolación, con respecto al tamaño de muestra seleccionado y luego se hizo un análisis de varianza del error en cada uno de los subconjuntos de estaciones .

Para el análisis de varianza se supuso que las  $k$  muestras son independientes y normalmente distribuidas, y se les aplicó una técnica paso a paso, fundamentada en el procedimiento de extraer secuencialmente las estaciones, considerándolas como variables independientes, de manera similar al procedimiento de eliminación hacia atrás que se realiza en una regresión lineal múltiple (Drapper y Smith, 1971). Luego se estudió la variación de la suma de cuadrados residuales de cada modelo ajustado, según una prueba F (Guenni (et al.), 1995). Cabe descartar que esta prueba F se considera aproximada

dado que se está trabajando en un espacio de Hilbert y además los sucesivos modelos ajustados no están anidados.

Para el cálculo de la razón de varianzas  $F = \frac{s_1^2}{s_2^2}$ , el procedimiento se basó en los fundamentos teóricos siguientes:

Sea  $S = \{X_1, \dots, X_n\}$ ,  $n$  puntos de observación denominados como localidades existentes. Ajustamos una superficie spline para  $S$ . La varianza de los residuales para la función ajustada a  $S$  es  $\sigma_s^2$ . El estimador de  $\sigma_s^2$  es estimado como:

$$\hat{\sigma}_s^2 = \frac{SSR_S}{g.l.error}$$

Bajo la suposición de normalidad  $SSR \sim \sigma_s^2 \chi^2(g.l.e)$

Sea  $d_i$  igual al residual escalado para la localidad  $i$ , consideramos

$$d_1^* = \min\{d_i\}, i = 1, \dots, n \quad \text{y el subconjunto } S^{(1)} = S - \{X_{(1)}\} \text{ donde } X_{(1)}$$

corresponde a la localidad donde  $d_i$  es mínimo

Ahora consideramos que  $SSR_S^{(1)}$  es igual a la suma de cuadrados de residuales, para todos los  $X_i$  cuando el punto  $X_{(1)}$  no ha sido incluido en el proceso de ajuste.

Se deseaba observar cuál era el incremento relativo de la suma de cuadrados residuales cuando se eliminan estaciones sucesivamente. Se esperaba que  $SSR_S < SSR_S^{(1)}$ , lo cual no es rigurosamente cierto ya que los modelos sucesivos no son estrictamente anidados.

En general, para dos conjuntos  $S^{(i)}$  y  $S^{(j)}$ , se comparó las sumas de cuadrados  $SSR_S^{(i)}$  y  $SSR_S^{(j)}$  tal que  $SSR_S^{(i)}$  es la suma de cuadrados de residuales con menor número de estaciones y  $SSR_S^{(j)}$  es la suma de cuadrados de residuales con mayor número de estaciones y además se cumpliera que :

$$SSR_S^{(i)} > SSR_S^{(j)}$$

Entonces se definió el estadístico F como:

$$F = \frac{(SSR(N) / g.l.n)}{(SSR(D) / g.l.d)} \sim F_{0.05 (g.l.n, g.l.d)}$$

∴

- N = número total de estaciones
- $n_i$  = número de estaciones ( i ) del conjunto menor de estaciones
- $n_j$  = número de estaciones ( j ) del conjunto mayor de estaciones
- g.l.n = grados de libertad del numerador =  $n_j - n_i$
- g.l.d = grados de libertad del denominador =  $N - n_i$
- $SS(N) = SSR(i) - SSR(j)$  = Suma de cuadrados residuales del numerador
- $SS(D) = SSR(j)$  = Suma de cuadrados residuales del denominador.

*Teniendo que :*

- Si  $F > F_{0.05 (g.l.n, g.l.d)}$ , no se podía extraer la estación  $X_{(1)}$ . El incremento en SSR es grande y la diferencia es significativa.

- Si  $F < F_{0.05 (g.l.n, g.l.d)}$ , se podía extraer la estación  $X_{(1)}$  y la diferencia no es significativa.

Si al comparar dos modelos consecutivos, la diferencia de su sumas de cuadrados residuales es negativa, se omitirá la comparación y se tomará el próximo siguiente.

En cada paso, la hipótesis nula ( $H_0$ ) que se prueba es que el estimador  $\hat{f}$  del spline teórico  $f$  definido en 3.1 es función de los puntos  $X_i$  que pertenecen al conjunto  $S^{(i)}$  con menor número de estaciones.

## *CAPITULO VII*

### *ANALISIS DE RESULTADOS*

En este capítulo se presenta en tres secciones los resultados obtenidos a partir de la aplicación de la metodología antes expuesta. En la primera sección se presentan la evaluación y aplicación del modelo Spline Laplaciano, a nivel anual y mensual. En la segunda, se obtuvieron las superficies de precipitación, para cada tipo del modelo (bidimensional, tridimensional y parcial) evaluadas sobre el modelo digital de terreno generado sobre la Cuenca analizada, seleccionando el modelo Spline Parcial como el modelo de mejor ajuste en el área de estudio y en la tercera sección se logró seleccionar una muestra óptima de estaciones, minimizando la red de estaciones pluviométricas de 83 a 46 puntos de observación, con lo que se reducirá en un 45% los costos por mantenimiento y operación de estaciones pluviométricas en el Sistema Hidráulico Yacambú - Quibor.

#### **7.1 APLICACION Y EVALUACION DEL MODELO SPLINE LAPLACIANO**

En general al aplicar el modelo Spline Laplaciano, según lo expuesto en la metodología, se encontró que el orden de la derivada que mejor define la rugosidad de la superficie es igual a dos, alcanzando de esta manera el menor error predictivo (tabla 2) en cada tipo del modelo ajustado (bidimensional, tridimensional y parcial).

Por otra parte, al examinar la validez del mismo, suponiendo que los residuos  $\epsilon$  se distribuyen como una Normal y comparando la suma de cuadrados residuales ( $n \cdot MSR$ ) con el valor crítico  $\sigma^2 \chi^2_m$ , donde  $\sigma^2$  es la varianza estimada,  $m = \text{traza} (I - A)$  y  $\chi^2$  es una chi-cuadrado con  $m$  grados de libertad, el modelo mostró siempre tener un buen ajuste (tabla 3), con un nivel de confianza del 95 % .

Además de estas propiedades, en la tabla 4, se muestra en cada modelo ajustado, la variación a nivel anual y mensual de los estadísticos, parámetro de suavizamiento, grados de

ANALISIS ANUAL CON RESPECTO AL ORDEN DE LAS DERIVADAS

TABLA 2

MODELO	ORDEN	MEDIA	Desv.std	VAR	RHO	ρ <sub>1</sub>	ρ <sub>2</sub>	ρ <sub>3</sub>	GCV	MNR	MSE	RTGCY	RTMSR	RTVAR	RTMSE	ρPREC.	mm/Km	AVR	RAVR
BIDIMENS.	2	1169.5	509.6	4.2	4.7	31.4	51.6	53400.0	7670.0	12600.0	231.0	87.6	2.1	112.0	19.8	4840.0	69.6	4840.0	69.6
	3	1169.5	509.6	5.6	264.0	37.2	45.8	60600.0	12200.0	15000.0	246.0	110.0	2.4	122.0	21.0	4840.0	69.6	4840.0	69.6
TRIDIMENS.	2	1169.5	509.6	1.8	0.1	15.1	67.9	48800.0	1610.0	7250.0	221.0	40.1	1.4	85.2	18.9	4840.0	69.6	4840.0	69.6
	3	1169.5	509.6	3.7	5.8	32.1	50.9	46300.0	6930.0	11000.0	215.0	83.2	1.9	105.0	18.4	4840.0	69.6	4840.0	69.6
PARCIAL	2	1169.5	509.6	3.1	3.6	28.1	54.9	44300.0	5080.0	9930.0	211.0	71.2	1.8	99.6	18.0	4840.0	69.6	4840.0	69.6
	3	1169.5	509.6	4.1	187.0	33.8	49.2	48900.0	8120.0	11800.0	221.0	90.1	2.0	109.0	18.9	4840.0	69.6	4840.0	69.6

ANALISIS DE LA BONDAD DE AJUSTE

MODELO BIDIMENSIONAL

TABLA 3

MES	$\sigma^2$	MSR	gle	$\chi^2_{0,95}$	$\sigma^2\chi^2$	SCR
ENE	36.3	22.2	50.7	71.4	2590.1	1842.6
FEB	16.4	4.0	20.3	34.2	559.7	332.8
MAR	8.4	0.8	8.4	17.5	147.0	70.3
ABR	284.2	140.0	40.8	59.3	16867.3	11620.0
MAY	614.4	240.0	32.4	49.5	30400.3	19920.0
JUN	1122.2	498.0	36.9	54.4	61088.4	41334.0
JUL	575.7	214.0	30.9	47.0	27044.1	17762.0
AGO	494.1	216.0	36.1	54.4	26896.4	17928.0
SEP	154.4	40.9	22.0	36.8	5678.9	3394.7
OCT	26.6	1.6	4.9	11.1	296.1	132.0
NOV	207.9	94.7	38.0	56.9	11828.6	7860.1
DIC	190.8	99.1	43.2	63.0	12015.4	8225.3

MODELO TRIDIMENSIONAL

MES	$\sigma^2$	MSR	gle	$\chi^2_{0,95}$	$\sigma^2\chi^2$	SCR
ENE	33.1	17.1	42.8	61.8	2047.8	1419.3
FEB	0.0	0.0	0.0	-	-	0.0
MAR	0.0	0.0	0.0	-	-	0.0
ABR	284.2	140.0	40.8	59.3	16867.3	11620.0
MAY	136.5	13.7	8.3	17.5	2393.1	1137.1
JUN	985.9	372.0	31.4	48.2	47550.4	30876.0
JUL	446.5	124.0	22.9	36.8	16424.0	10292.0
AGO	0.0	0.0	0.0	-	-	0.0
SEP	71.0	9.2	10.7	20.5	1454.8	759.5
OCT	0.0	0.0	0.0	-	-	0.0
NOV	63.6	10.9	14.2	26.1	1660.1	904.7
DIC	73.0	14.5	16.4	28.8	2106.6	1203.5

MODELO PARCIAL

MES	$\sigma^2$	MSR	gle	$\chi^2_{0,95}$	$\sigma^2\chi^2$	SCR
ENE	32.8	18.6	47.1	67.8	2225.9	1543.8
FEB	11.0	2.0	15.3	27.5	301.8	167.7
MAR	10.8	1.5	11.2	21.9	237.1	121.2
ABR	194.4	80.0	34.2	52.0	10104.6	6640.0
MAY	558.1	207.0	30.9	47.0	26219.6	17181.0
JUN	960.5	395.0	34.1	52.0	49913.9	32785.0
JUL	476.1	158.0	27.6	43.2	20566.2	13114.0
AGO	362.8	135.0	30.9	47.0	17045.9	11205.0
SEP	114.1	26.0	19.0	32.9	3747.2	2158.0
OCT	12.9	0.5	3.0	9.3	120.6	38.9
NOV	108.3	33.6	25.8	40.6	4402.3	2788.8
DIC	60.6	11.5	15.8	27.5	1665.9	954.5

NIVEL ANUAL

MOD	$\sigma^2$	MSR	gle	$\chi^2_{0,95}$	$\sigma^2\chi^2$	SCR
BID	20231.2	7670.0	31.4	48.2	975789.7	636610.0
TRID	8857.2	1610.0	15.1	27.5	243469.9	133630.0
PAR	15004.0	5080.0	28.1	44.5	667089.7	421640.0
PAR2	15100.8	5050.0	27.8	43.2	652271.8	419150.0

EVALUACION DEL MODELO SPLINE LAPLACIANO

TABLA 4

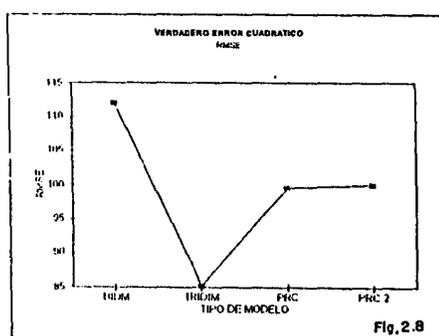
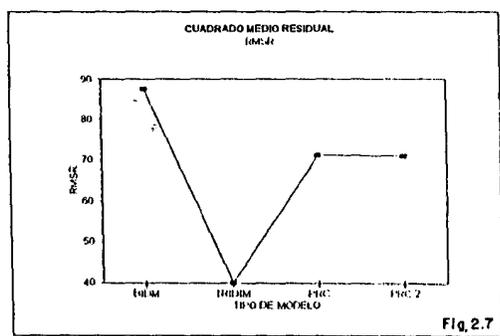
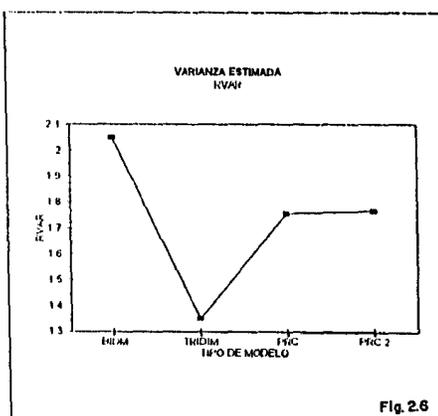
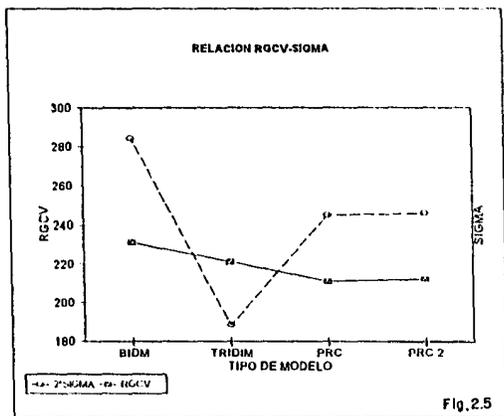
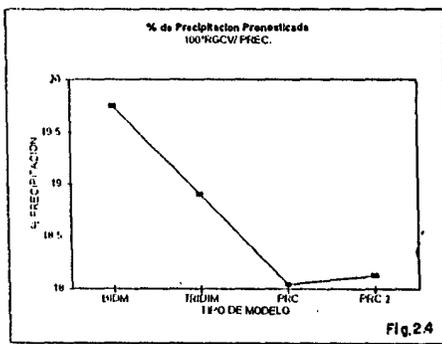
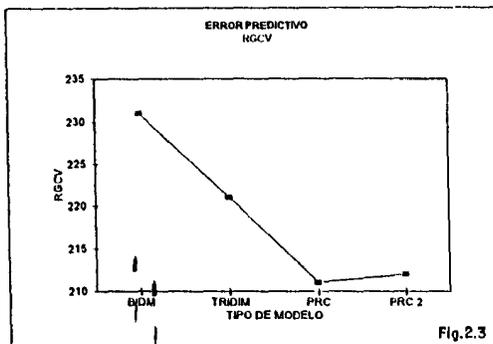
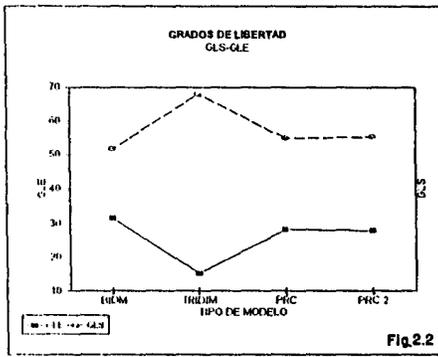
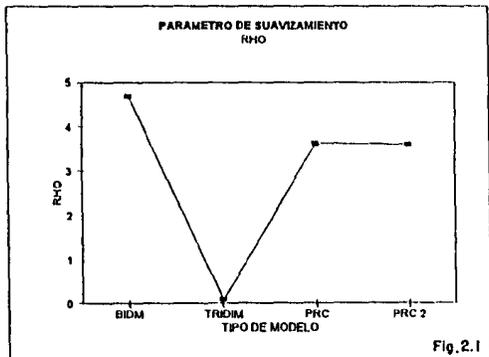
ANALISIS ANUAL																		
MODELO	MEDIA	Dev_std	s2	RHO	VAR	gle	glr	GCV	MSR	MSE	RTGCV	RTMSR	RTVAR	RTMSE	%PREC.	mm/Km	AVR	RAVR
BIDM	1169.5	509.6	20231.2	4.5	4.18	31.4	51.6	53400.0	7670.0	12600	231.0	87.6	2.1	112.0	19.8		4840.0	69.6
TRIDIM	1169.5	509.6	8857.2	0.1	1.83	15.1	67.9	48800.0	1610.0	7250	221.0	40.1	1.4	85.2	18.9		4840.0	69.6
PRC	1169.5	509.6	15004.0	3.6	3.10	28.1	54.9	44300.0	5080.0	9930	211.0	71.2	1.8	99.6	18.0	232.0	4840.0	69.6
PRC2	1169.5	509.6	15100.8	3.5	3.12	27.8	55.2	45000.0	5050.0	10000	212.0	71.0	1.8	100.0	18.1	232/40.6	4840.0	69.6
ANALISIS MENSUAL																		
MODELO BIDIMENSIONAL																		
MES	MEDIA	Dev_std	s2	RHO	VAR	gle	glr	GCV	MSR	MSE	RTGCV	RTMSR	RTVAR	RTMSE	%PREC.	mm/Km	AVR	RAVR
ENE	21.0	12.1	36.3	21.7	0.663	50.7	32.3	59.4	22.2	14.1	7.7	4.7	0.8	3.8	36.7		54.7	7.4
FEB	24.3	16.2	16.4	3.8	0.182	20.3	62.7	66.8	4.0	12.4	8.2	2.0	0.4	3.5	33.6		90.0	9.5
MAR	36.2	20.3	8.4	0.4	0.081	8.4	74.6	82.5	0.8	7.5	9.1	0.9	0.3	2.7	25.1		103.0	10.1
ABR	123.9	46.4	284.2	7.5	0.440	40.8	42.2	578.0	140.0	144.0	24.0	11.8	0.7	12.0	19.4		646.0	25.4
MAY	153.6	82.9	614.4	0.0	1.310	32.4	50.6	1560.0	240.0	375.0	39.7	15.5	1.2	19.4	25.8		469.0	21.7
JUN	156.9	111.0	1122.2	10.0	3.540	36.9	46.1	2520.0	498.0	623.0	50.2	22.3	1.9	25.0	32.0		317.0	21.7
JUL	140.3	95.0	575.7	7.0	2.140	30.9	52.1	1550.0	214.0	362.0	39.3	14.6	1.5	19.0	28.0		269.0	16.4
AGO	119.2	70.2	494.1	10.1	2.560	36.1	46.9	1140.0	216.0	280.0	33.7	14.7	1.6	16.7	28.3		193.0	13.9
SEP	118.1	58.4	154.4	2.1	0.800	22.0	61.0	580.0	40.9	113.0	24.1	6.4	0.9	10.6	20.4		193.0	13.9
OCT	132.2	49.8	26.6	0.7	0.103	4.9	78.1	448.0	1.6	25.1	21.2	1.3	0.3	5.0	16.0		258.0	16.1
NOV	90.7	32.7	207.9	16.9	1.050	38.0	45.0	451.0	94.7	112.0	21.2	9.7	1.0	10.6	23.4		198.0	14.1
DIC	59.7	27.3	190.8	0.7	0.875	43.2	39.8	366.0	99.1	91.5	19.1	10.0	0.9	9.6	32.0		218.0	14.8
MODELO TRIDIMENSIONAL																		
MES	MEDIA	Dev_std	s2	RHO	VAR	gle	glr	GCV	MSR	MSE	RTGCV	RTMSR	RTVAR	RTMSE	%PREC.	mm/Km	AVR	RAVR
ENE	21.0	12.1	33.1	1.1	0.606	42.8	40.2	64.3	17.1	16.1	8.0	4.1	0.8	4.0	38.2		54.7	7.4
FEB	24.3	16.2	0.0	0.0	0.000	0.0	83.0	60.8	0.0	0.0	7.8	0.0	0.0	0.0	32.1		90.0	9.5
MAR	36.2	20.3	0.0	0.0	0.000	0.0	83.0	75.1	0.0	0.0	8.7	0.0	0.0	0.0	23.9		103.0	10.1
ABR	123.9	46.4	284.2	0.1	0.440	40.8	42.2	578.0	140.0	144.0	24.0	11.8	0.7	12.0	19.4		646.0	25.4
MAY	153.6	82.9	136.5	0.0	0.291	8.3	74.7	1360.0	13.7	123.0	36.8	3.7	0.5	11.1	24.0		469.0	21.7
JUN	156.9	111.0	985.9	0.1	3.110	31.4	51.6	2600.0	372.0	611.0	51.0	19.3	1.8	24.7	32.5		317.0	21.7
JUL	140.3	95.0	446.5	0.0	1.660	22.9	60.1	1620.0	124.0	324.0	40.2	11.1	1.3	18.0	28.6		269.0	16.4
AGO	119.2	70.2	0.0	0.0	0.000	0.0	83.0	681.0	0.0	0.0	26.1	0.0	0.0	0.0	21.9		193.0	13.9
SEP	118.1	58.4	71.0	0.0	0.368	10.7	72.3	549.0	9.2	61.7	23.4	3.0	0.6	7.9	19.8		193.0	13.9
OCT	132.2	49.8	0.0	0.0	0.000	0.0	83.0	445.0	0.0	0.0	21.1	0.0	0.0	0.0	16.0		258.0	16.1
NOV	90.7	32.7	63.6	0.0	0.321	14.2	68.8	370.0	10.9	52.5	19.2	3.3	0.6	7.3	21.2		198.0	14.1
DIC	59.7	27.3	73.0	0.1	0.335	16.4	66.6	369.0	14.5	58.6	19.2	3.8	0.6	7.7	32.1		218.0	14.8
MODELO PARCIAL																		
MES	MEDIA	Dev_std	s2	RHO	VAR	gle	glr	GCV	MSR	MSE	RTGCV	RTMSR	RTVAR	RTMSE	%PREC.	mm/Km	AVR	RAVR
ENE	21.0	12.1	32.8	37.5	0.600	47.1	35.6	57.9	18.6	14.2	7.6	4.3	0.8	3.8	36.3	3.6	54.7	7.4
FEB	24.3	16.2	11.0	4.6	0.122	15.3	67.7	59.6	2.0	9.0	7.7	1.4	0.3	3.0	31.7	5.3	90.0	9.5
MAR	36.2	20.3	10.8	0.6	0.105	11.2	71.8	79.8	1.5	9.3	8.9	1.2	0.3	3.1	24.6	6.5	103.0	10.1
ABR	123.9	46.4	194.4	7.5	0.301	34.2	48.8	472.0	80.0	114.0	21.7	8.9	0.5	10.7	17.5	26.8	646.0	25.4
MAY	153.6	82.9	558.1	0.0	1.190	30.9	52.1	1490.0	207.0	349.0	38.6	14.4	1.1	18.7	25.1	25.5	469.0	21.7
JUN	156.9	111.0	960.5	7.7	3.030	34.1	48.9	2330.0	395.0	565.0	48.3	19.9	1.7	23.8	30.8		29.7	21.7
JUL	140.3	95.0	476.1	3.2	1.770	27.6	55.4	1430.0	158.0	317.0	37.8	12.6	1.3	17.8	26.9		20.9	16.4
AGO	119.2	70.2	362.8	6.2	1.880	30.9	52.1	975.0	135.0	228.0	31.2	11.6	1.4	15.1	26.2	26.3	193.0	13.9
SEP	118.1	58.4	114.1	2.1	0.591	19.0	64.0	497.0	26.0	87.7	22.3	5.1	0.8	9.4	18.9	21.1	193.0	13.9
OCT	132.2	49.8	12.9	0.4	0.050	3.0	80.0	355.0	0.5	12.4	18.8	0.7	0.2	3.5	14.2	24.0	258.0	16.1
NOV	90.7	32.7	108.3	3.8	0.547	25.8	57.2	347.0	33.6	74.5	18.6	5.8	0.7	8.6	20.5	23.7	198.0	14.1
DIC	59.7	27.3	60.6	2.1	0.278	15.8	67.2	318.0	11.5	49.0	17.8	3.4	0.5	7.0	29.8	15.7	218.0	14.8

libertad del error y la señal, que evalúan la superficie interpolada, se pudo observar lo siguiente:

**A .- Nivel Anual :**

- a.- El parámetro de suavizamiento (fig.2.1) en el modelo tridimensional es aproximadamente igual a cero y el porcentaje de estaciones para identificar la señal (fig.2.2) es muy alto (82%) por lo que se deduce que el modelo no es lo suficientemente preciso en la representación de las características de la precipitación en la cuenca analizada o que el número de estaciones por área no es el adecuado.
- b.- El menor error predictivo tuvo lugar en la superficie generada por el modelo spline parcial, (figs. 2.3 y 2.4) donde se incluyó la altura como un sub-modelo lineal paramétrico.
- c.- En las distintas formas en las que se evaluó el modelo, los grados de libertad del error son menor a la mitad del número de puntos de observación y la validación cruzada generalizada (GCV) es menor a  $2\hat{\sigma}_\varepsilon^2$  (fig.2.5 ), por lo que se deduce a partir de lo demostrado por Hutchinson y Gessler (1994) podría existir alguna deficiencia en la superficie interpolada, causada por la dispersión de los puntos de observación en algunas regiones de la cuenca analizada.
- d.- Con respecto de la varianza estimada (RVAR), el cuadrado medio residual (RMSR) y verdadero error cuadrático (RMTSE), medidos todos en base a la raíz cuadrada de sus valores, se encontró que éstos disminuyeron significativamente en el modelo tridimensional (figs. 2.6 a 2.8) sin que ello representase más exactitud en la superficie de interpolación generada, por lo expuesto en ( a ).

ANALISIS A NIVEL ANUAL



TIPO DE MODELO:  
 BIDM= BIDIMENSIONAL  
 TRIDIM= TRIDIMENSIONAL  
 PRC= PARCIAL ( 1 COVARIABLE)  
 PRC2= PARCIAL ( 2 COVARIABLES)

- e.-El hecho de incluir dentro del modelo Spline parcial una covariable adicional a la altura que midiera el efecto de la estación al encontrarse, dentro o fuera del área digitalizada no aportó mejoras importantes en la precisión de la superficie generada
- f.- El gradiente de precipitación en el área analizada fue de 232mm /km.
- g.-La precipitación media en toda el área es de 1169.5 mm, con una desviación estándar de los valores de precipitación de 69.6 mm y un coeficiente de variación del 43%.

#### **B.- Nivel Mensual :**

Según la cantidad mensual de precipitación, se pudo apreciar que:

- a.-El régimen de precipitación mensual (fig. 3.1) es del tipo bimodal, presentando un máximo principal en el mes de junio y otro secundario en el mes de octubre . El período seco se observa entre el mes de diciembre hasta marzo, ocurriendo el mínimo en el mes de enero
- b.-A excepción del mes más seco, el parámetro de suavizamiento (fig 3.2) tiende a ser mayor en el período seco que en el húmedo.
- c.-Durante los meses secos predomina el ruido en los datos y la señal es más débil, presentándose lo contrario en el período húmedo (fig . 3.3 y 3.4 )
- d.- La varianza estimada (fig. 3.5) es menor que uno en los meses secos, lo que significa que durante estos meses la varianza del error de la superficie ajustada es menor que la varianza promedio en los datos observados .
- e.- El valor del cuadrado medio residual, error predictivo y error cuadrático verdadero (figs.3.6 a 3.8) en los meses secos es mayor que en los meses húmedos, lo que implica que la precisión de la superficie interpolada depende de la cantidad de precipitación registrada en cada uno de los meses.

ANÁLISIS MENSUAL

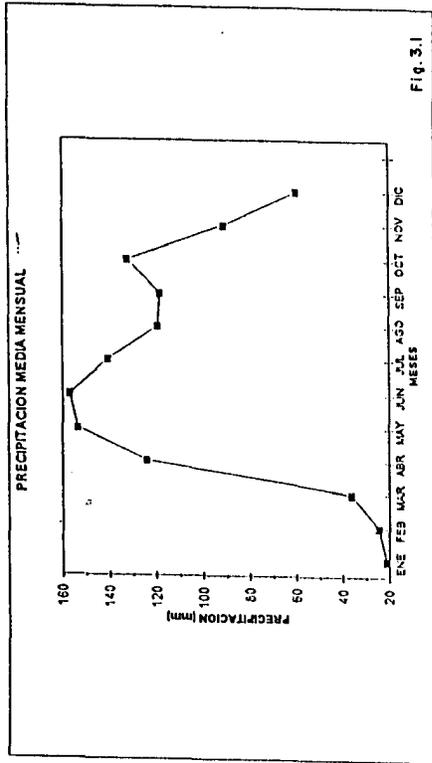


Fig. 3.1

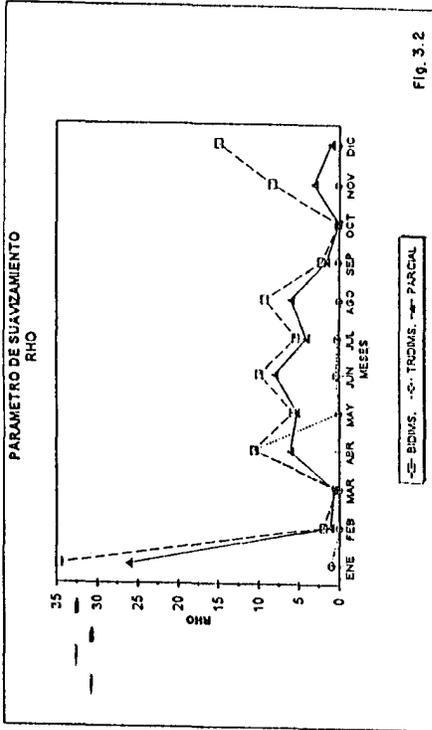


Fig. 3.2

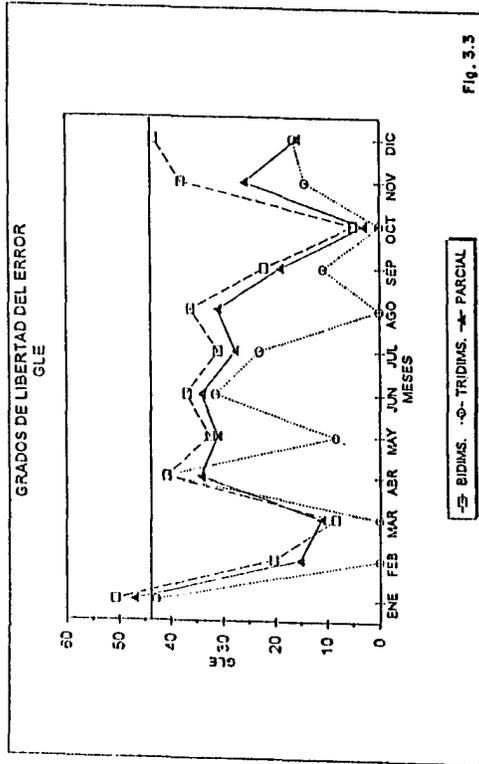


Fig. 3.3

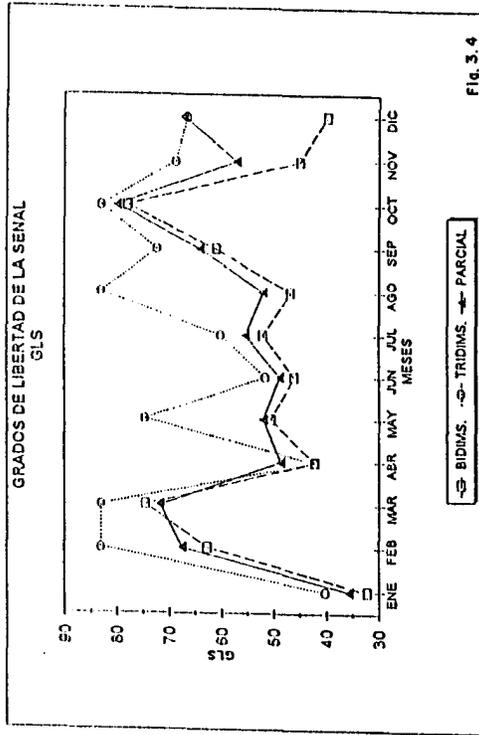
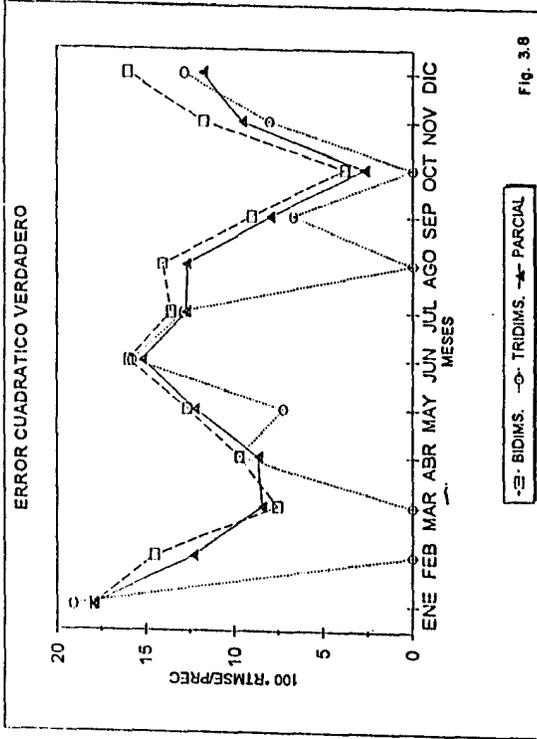
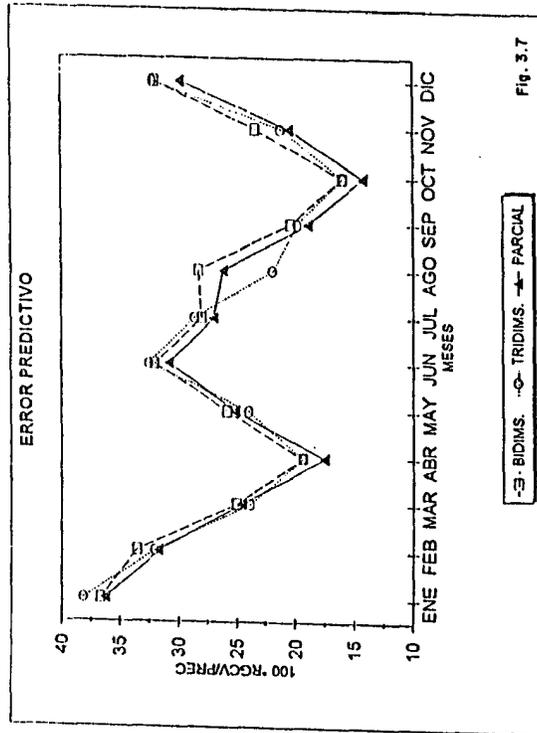
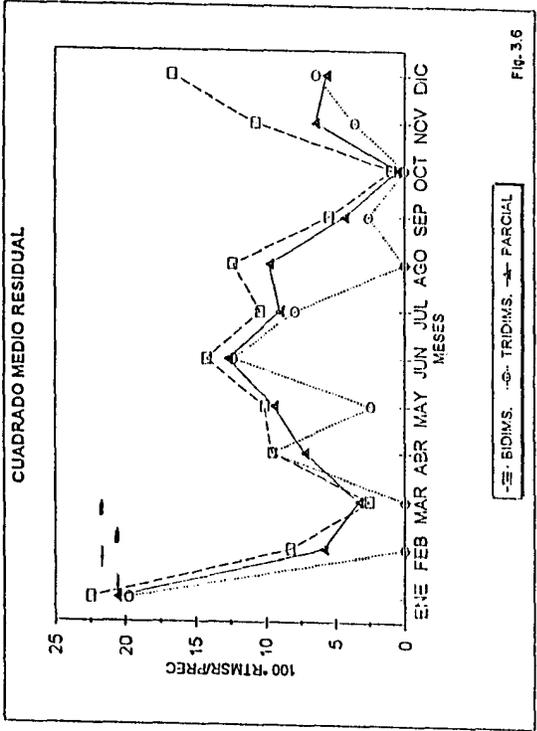
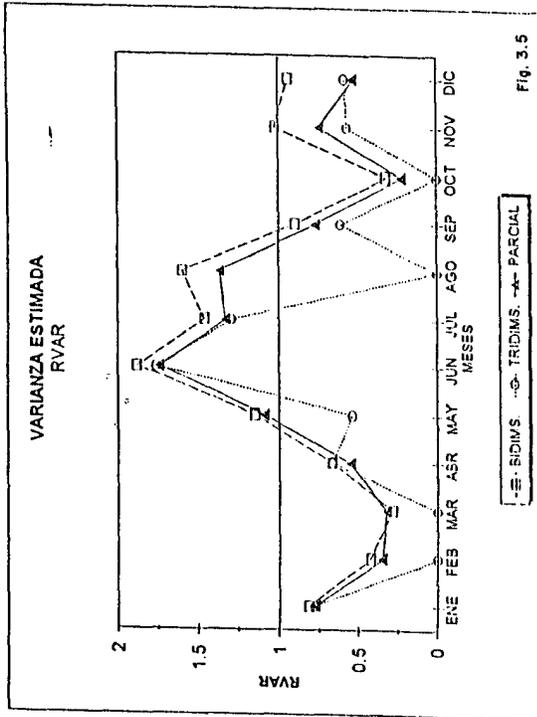


Fig. 3.4

ANALISIS MENSUAL



Al comparar, el modelo con respecto a la variación de los estadísticos, parámetro de suavizamiento, la señal y el ruido, se observó que los valores calculados con el modelo en forma bidimensional tienen igual curva de variación mensual que en el modelo spline parcial, en el que se logran magnitudes menores y por tanto menor error en la superficie de precipitación ajustada.

Con respecto al modelo aplicado en forma tridimensional, se observó que en la mayoría de los meses, el parámetro de suavizamiento es muy bajo y la señal es muy alta, y los valores de VAR, MSR, GCV y MTSE, llegan a ser cero, lo que indica que el modelo falla cuando es utilizado de esta forma y por tanto pierde significado físico.

## **7.2 PRODUCCION DE LA SUPERFICIE DE PRECIPITACION AJUSTADA**

La superficie de interpolación producida a partir del modelo Spline Laplaciano en forma bidimensional, tridimensional y parcial se evaluaron sobre el modelo digital de terreno, el cual se representó primero, en un diagrama en perspectiva (fig.4.1 ) con visión Noreste y un ángulo de inclinación sobre el plano de 30°, donde se pueden observar claramente el límite de la cuenca, los valles intramontanos, la forma y dirección del sistema orográfico, cimas y vertientes, siendo todas estas características importantes para el análisis del comportamiento espacial de la precipitación en el área en estudio. En la figura 4.2 se presenta el modelo del terreno como una imagen raster en la cual se clasificaron pisos altitudinales cada 150 m.s.n.m y se le acoplaron la red hidrográfica y la red de estaciones pluviométricas.

De esta manera, las características en las superficies de precipitación, fueron las siguientes:

**Modelo digital de terreno  
Cuenca del río Yacambu**

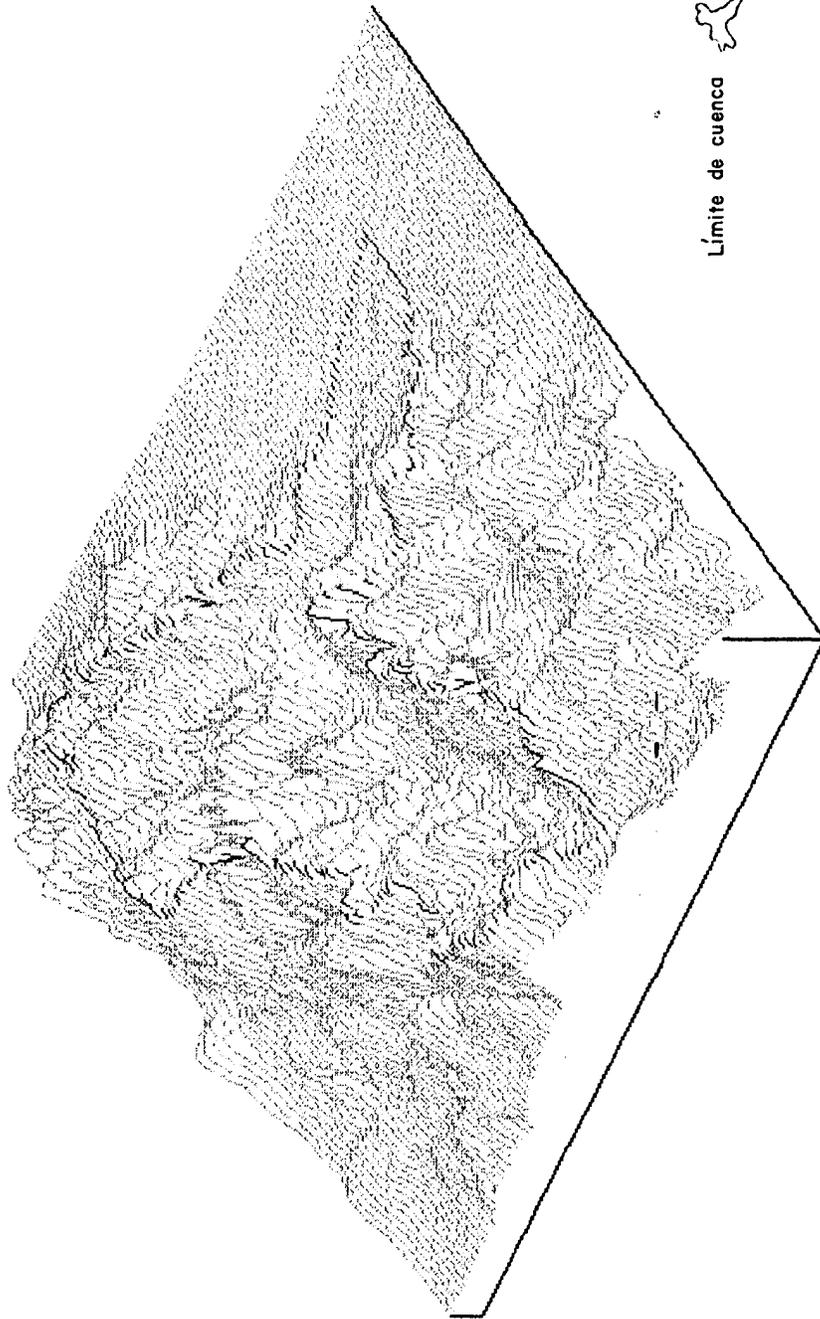


Fig. 4.1

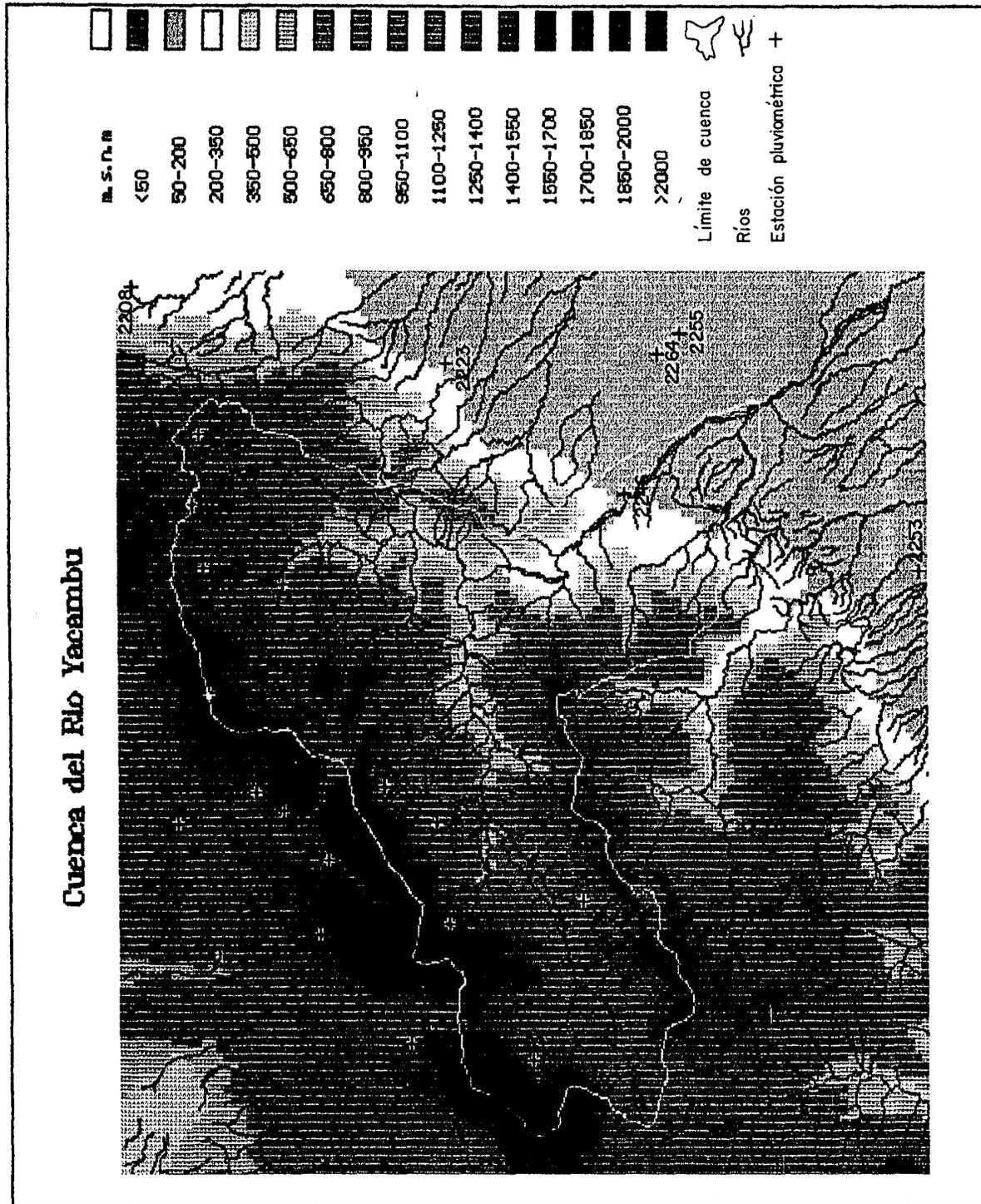


Fig. 4.2

### **A.- Modelo Bidimensional :**

La superficie de interpolación generada a partir de la aplicación del modelo spline laplaciano de forma bidimensional, en formato raster (fig.5.1) y acoplada al terreno (fig.5.2), se caracterizó principalmente por presentar un trazado regular en la demarcación de las áreas con igual precipitación .

La precipitación estimada por este modelo, varió entre un mínimo de 420 mm hasta un máximo de 2410.6 mm, ubicándose el área de mayor precipitación en un gran núcleo que se extiende desde el área central hasta el sur de la cuenca y la de menor precipitación hacia el lado oeste.

### **B.- Modelo Tridimensional :**

Al incluir la altura como una variable independiente en el modelo spline, la superficie de precipitación ajustada (fig.6.1 y 6.2 ) presentó un trazado irregular en la definición de los límites de aquellas áreas con igual precipitación, indicando lo complejo de la topografía . La precipitación estimada estuvo entre un mínimo de 467.21 mm. y un máximo de 2475.6 mm, localizando las áreas con mayor precipitación en los sistemas orográficos ubicados en la zona sur de la cuenca y las zonas de menor precipitación en lado oeste y este de la misma.

### **C. Modelo Spline Parcial .**

En este caso, el modelo pudo estimar los valores máximos de precipitación, en las zonas desprovistas de estaciones pluviométricas, según el cálculo del gradiente de precipitación estimado a partir de la inclusión de la altura como única covariable dentro

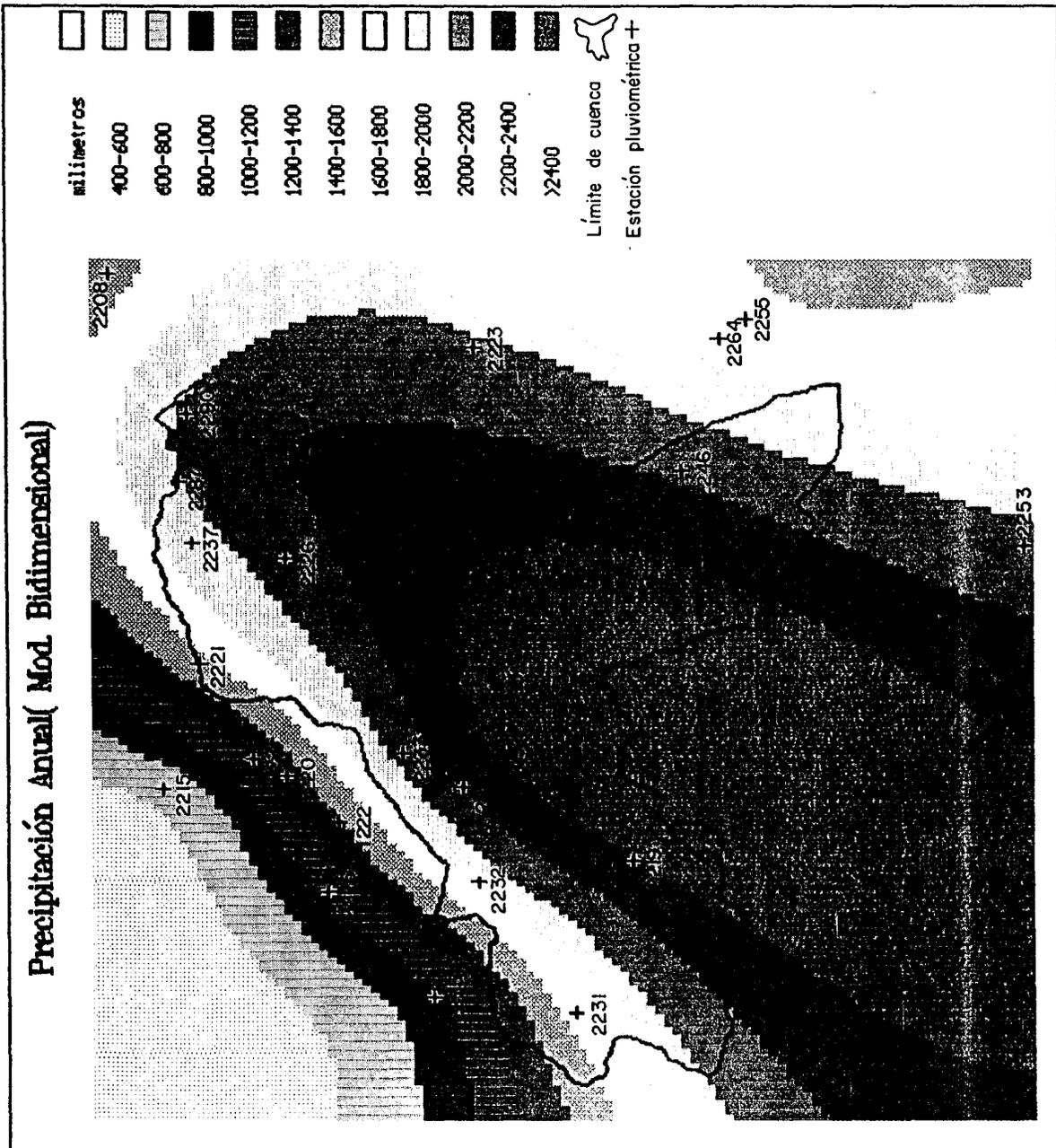


Fig. 5.1

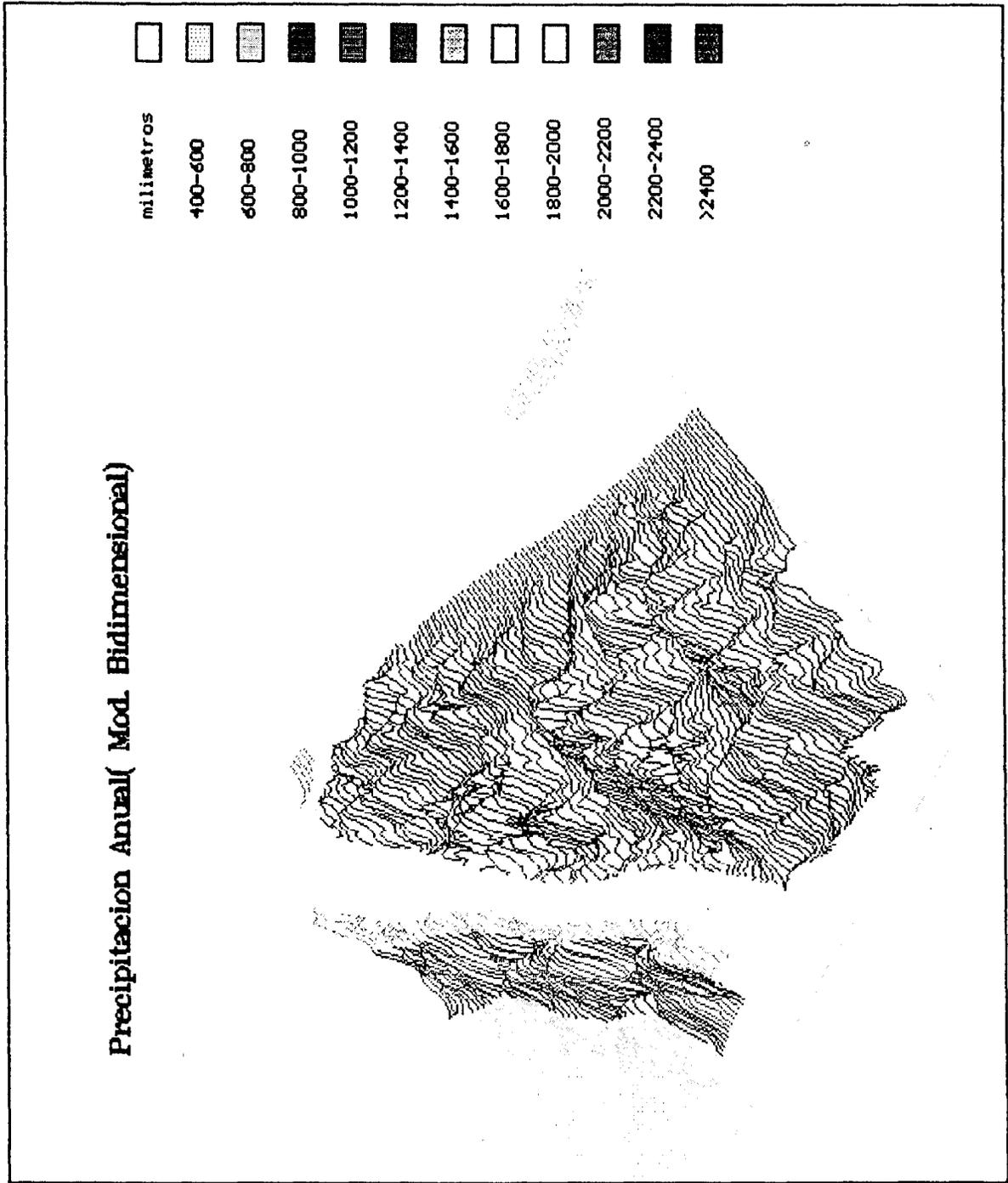


Fig. 5.2

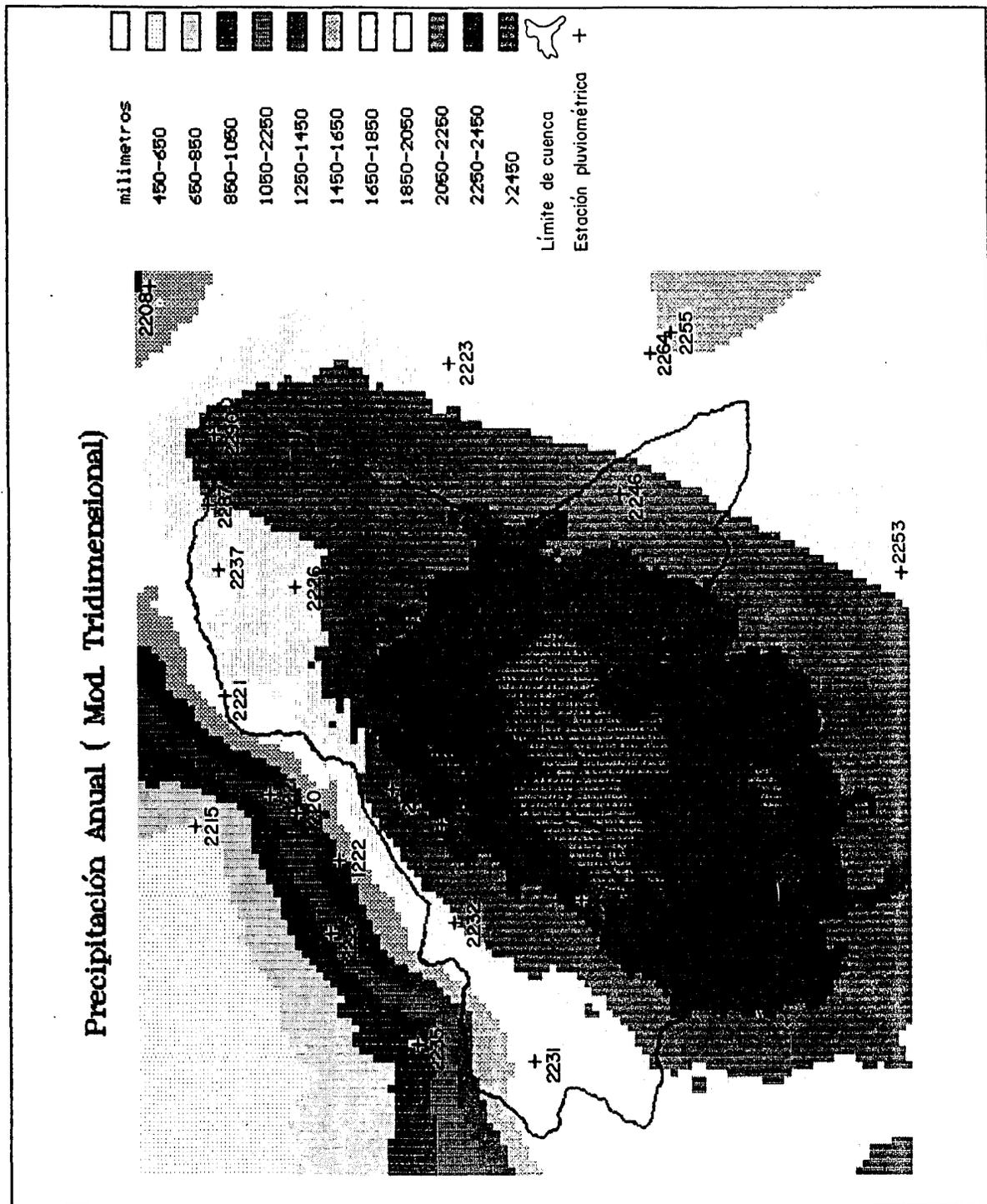


Fig. 6.1

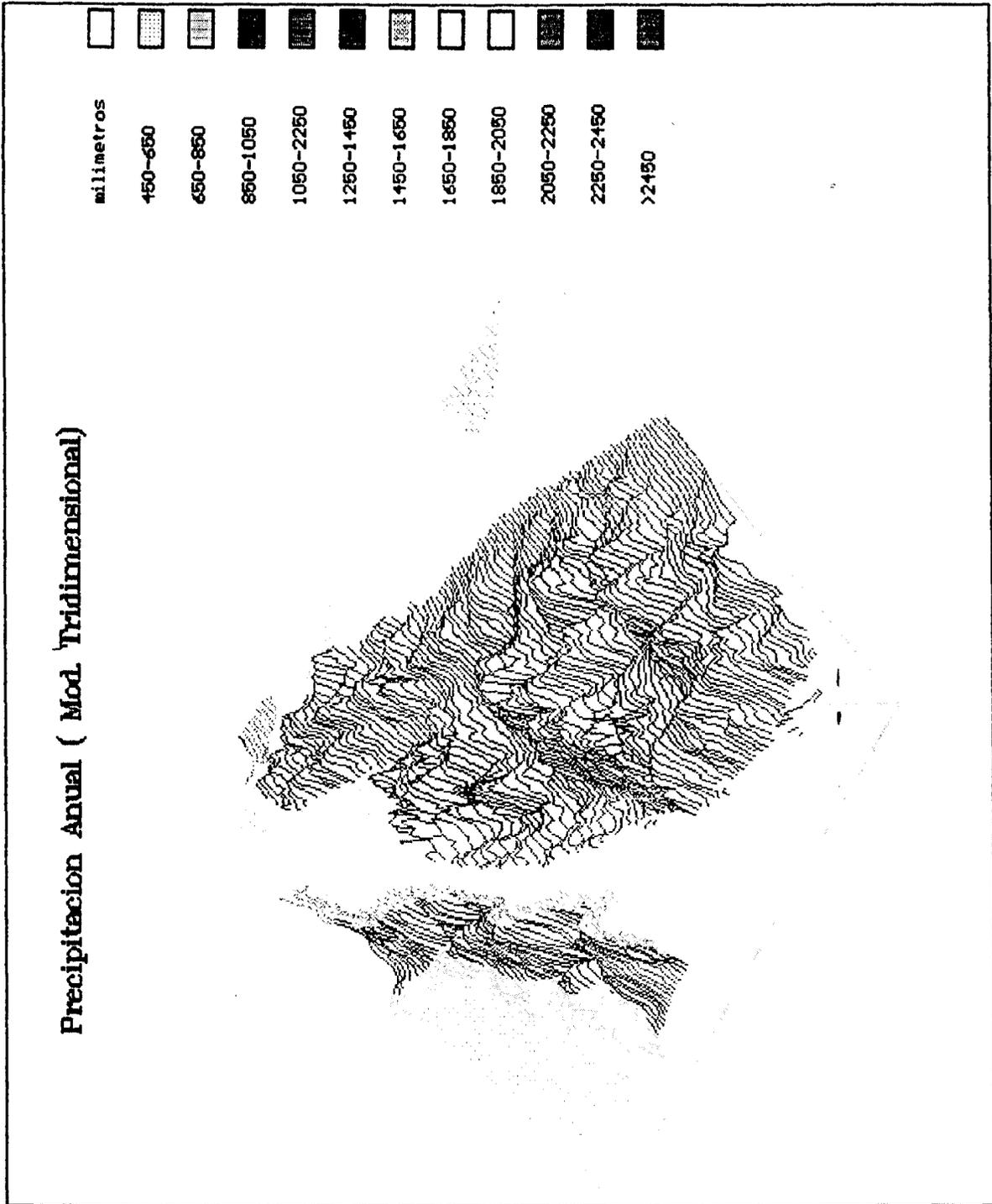


Fig. 6.2

de un sub-modelo lineal paramétrico. De esta manera, en la superficie de precipitación ( fig. 7.1) ajustada por el modelo spline parcial, se observaron núcleos de poca extensión con precipitación alta, en la zona sur de la cuenca donde no existen puntos de observación (fig.7.2) y las zonas de menor precipitación en lado oeste y este de la cuenca.

En forma similar al modelo tridimensional, el trazado en la demarcación de las áreas con igual precipitación fue muy irregular, debido a la incidencia del relieve en la estimación de la precipitación.

## **7.5 SELECCION DEL MODELO.**

El modelo Spline Laplaciano parcial fue el modelo que mejor se ajustó a las condiciones fisiograficas de la Cuenca, obteniéndose a través de él un error predictivo del 18% sobre la cantidad de precipitación anual, lo cual es completamente válido en la operaciones de usos del agua.

## **7.6 OPTIMIZACION DEL NUMERO DE ESTACIONES**

### **7.2. 1 MUESTREO CON REEMPLAZO**

El muestreo con reemplazo (tabla 5 ) se realizó con la finalidad de estudiar el aporte de información de cada estación en la superficie interpolada, en base a la variación que presentaban los grados de libertad del error y la señal, el parámetro de suavizamiento y los estadísticos calculados, una vez que era extraída la estación pluviométrica del conjunto total de puntos de observación.

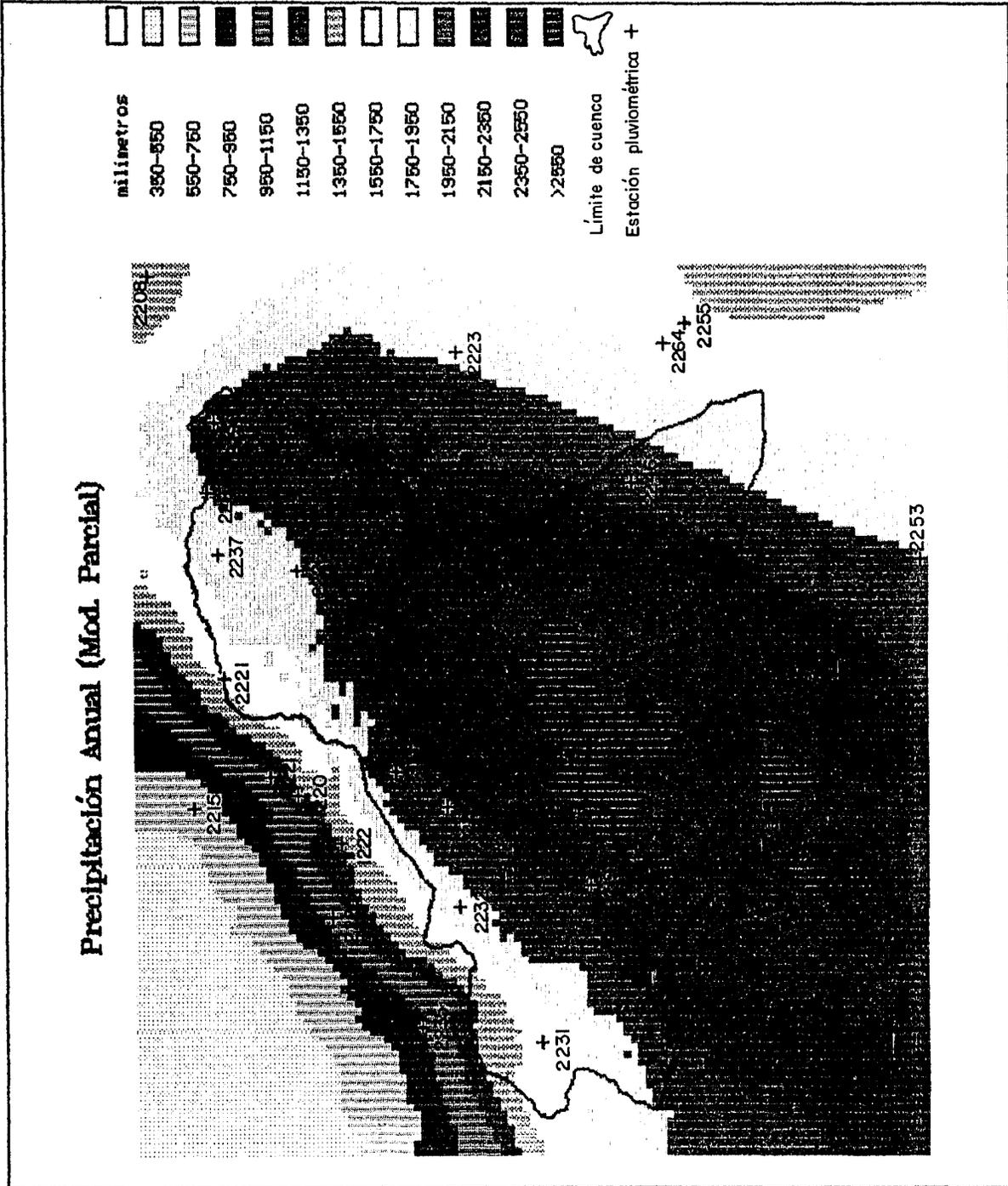


Fig. 7.1

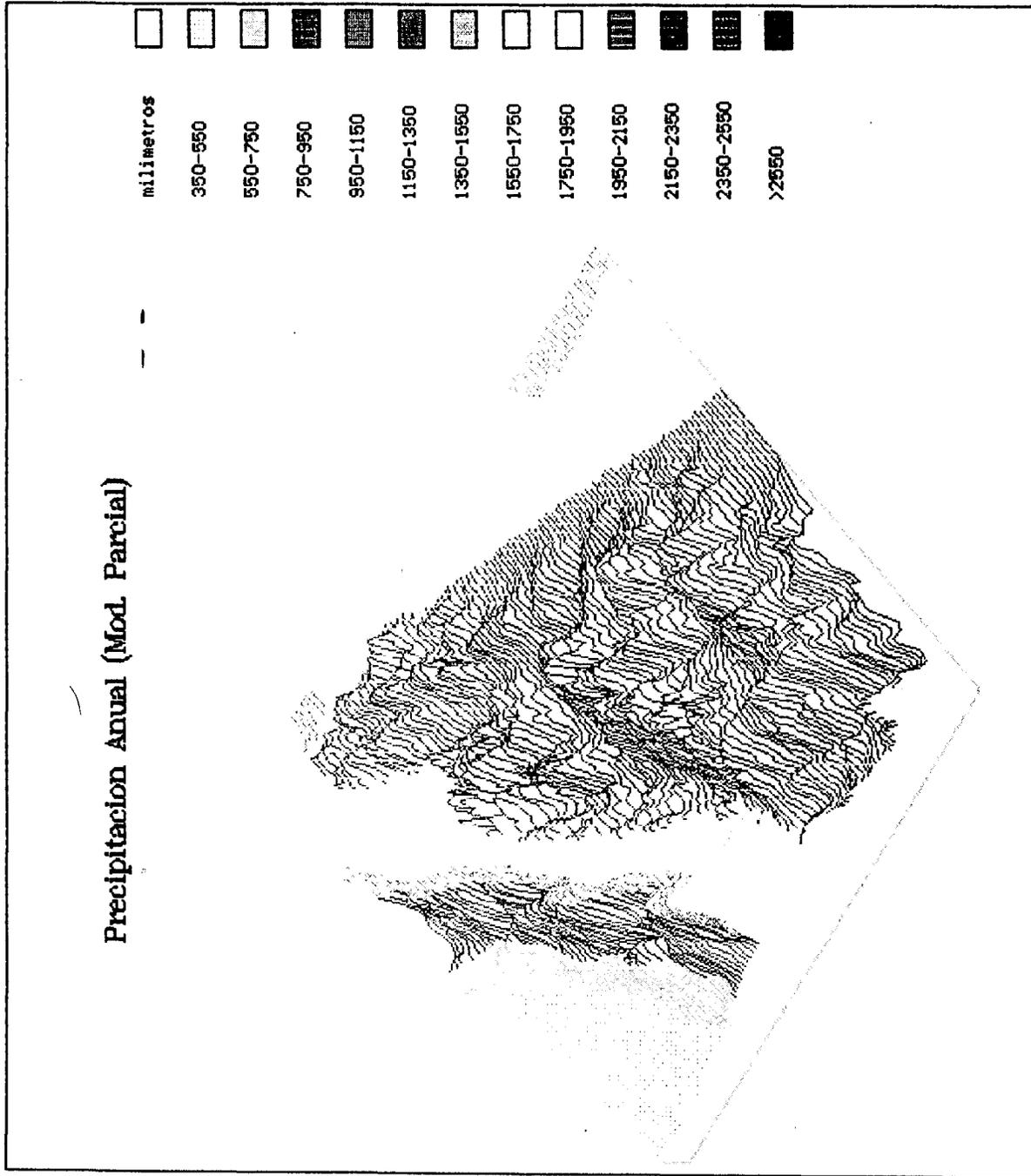


Fig. 7.2

TABLA 5

ESTACION	MUESTREO CON REEMPLAZO															
	MEDIA	SE	VAR	RHO	g <sub>1e</sub>	g <sub>1s</sub>	GCV	MSR	MSE	RTGCV	RTMSR	RTVAR	RTMSE	mm/ Km	mVr	RvR
1174	1175.0	511.89	3.10	3.57	27.8	54.2	44300	5090	9920	210.0	71.3	1.76	99.60	233	4850	69.6
1177	1175.9	509.29	3.09	3.55	27.8	54.2	44100	5060	9980	210.0	71.1	1.76	99.40	233	4830	69.5
1180	1175.1	510.08	3.10	3.5	27.4	54.8	44700	5000	9950	211.0	70.7	1.76	99.70	233	4810	69.4
1187	1175.9	508.20	3.10	3.52	27.5	54.5	44700	5030	9960	211.0	70.9	1.76	99.80	233	4840	69.6
1202	1175.4	509.77	3.15	3.71	27.5	54.5	45600	5130	10200	214.0	71.8	1.77	101.00	234	4870	69.9
1203	1178.6	505.78	3.10	3.61	27.9	54.1	44300	5140	9960	211.0	71.7	1.76	99.80	233	4860	69.9
1204	1170.0	505.12	3.10	3.58	27.7	54.3	44700	5110	10000	211.0	71.5	1.76	100.00	233	4860	69.9
1208	1176.4	508.69	3.24	3.84	27.9	54.1	46500	5370	10400	216.0	73.3	1.80	102.00	231	4860	69.9
1210	1176.9	508.14	3.09	3.53	27.3	54.7	45400	5020	10100	213.0	70.9	1.76	100.00	234	4890	69.9
1220	1188.9	512.67	3.21	3.81	27.8	54.2	45700	5260	10200	214.0	72.5	1.76	101.00	237	4830	69.5
1221	1170.4	512.62	3.30	3.96	28.3	53.7	46700	5400	10400	215.0	74.1	1.82	102.00	230	4820	69.4
1222	1187.3	512.30	3.10	3.7	27.0	54.4	46100	5210	10300	215.0	72.2	1.76	101.00	231	4860	69.7
1265	1178.9	505.29	3.09	3.6	27.8	54.2	44500	5130	9970	211.0	71.0	1.76	99.80	233	4880	69.9
1268	1174.3	510.81	3.06	3.7	28.1	53.9	43400	5100	9780	208.0	71.4	1.75	99.80	237	4870	69.8
1271	1177.4	507.51	3.10	3.54	27.6	54.4	44700	5050	9970	211.0	71.1	1.76	99.80	233	4840	69.6
1272	1178.1	505.01	3.10	3.58	27.7	54.3	44700	5110	10000	212.0	71.5	1.76	100.00	233	4860	69.9
1273	1177.9	506.87	3.10	3.51	27.4	54.6	44900	5010	9990	212.0	70.8	1.76	100.00	233	4840	69.6
1274	1178.2	508.36	3.10	3.53	27.5	54.5	45000	5060	10000	212.0	71.1	1.76	100.00	234	4870	69.8
1277	1173.7	511.24	3.10	3.53	27.5	54.5	44600	5040	9950	211.0	71.0	1.76	99.80	233	4840	69.6
1278	1176.1	509.09	3.14	3.7	28	54.0	44900	5240	10100	212.0	72.4	1.77	100.00	231	4860	69.9
1282	1178.0	508.09	3.09	3.6	27.3	54.7	45300	5030	10100	213.0	70.9	1.76	100.00	232	4860	69.9
1295	1171.2	517.44	3.07	3.63	27.8	54.2	44200	5080	9900	210.0	71.2	1.75	99.50	232	4880	69.9
1321	1169.8	512.68	3.11	3.67	27.9	54.1	44100	5110	9900	210.0	71.5	1.76	99.50	231	4830	69.5
1322	1171.4	512.38	3.18	3.82	28.3	53.7	44800	5330	10100	212.0	73.0	1.78	101.00	209	4860	69.7
1342	1168.9	512.68	3.19	3.68	27.8	54.4	46000	5210	10300	214.0	72.2	1.79	101.00	235	4850	69.6
1354	1169.6	512.89	3.18	3.69	27.7	54.3	45900	5230	10300	214.0	72.3	1.78	101.00	231	4870	69.8
1362	1168.9	512.68	3.08	3.53	27.5	54.5	44800	5000	9930	211.0	70.7	1.75	99.70	235	4870	69.8
1371	1184.2	510.39	3.07	3.88	28.8	53.2	42100	5190	9600	205.0	72.0	1.75	98.00	228	4820	69.4
1381	1168.4	512.69	3.10	3.58	27.8	54.4	44600	5070	9970	211.0	71.2	1.76	99.80	234	4850	69.6
2108	1168.1	512.54	3.18	3.84	28.5	53.5	44500	5380	10100	211.0	73.4	1.78	100.00	200	4870	69.8
2118	1172.7	511.85	3.09	3.58	27.6	54.4	44500	5050	9940	211.0	71.1	1.76	99.70	214	4850	69.8
2119	1173.1	511.62	3.15	3.69	27.7	54.3	45100	5130	10100	212.0	71.7	1.77	100.00	217	4840	69.6
2126	1173.3	511.47	3.09	3.53	27.4	54.8	44900	5010	9990	212.0	70.8	1.76	100.00	234	4860	69.7
2127	1171.4	512.37	3.06	3.48	27.1	54.9	44600	4880	9870	211.0	69.9	1.75	99.40	237	4830	69.5
2128	1172.0	512.18	3.09	3.49	27.3	54.7	44800	4870	9950	212.0	70.5	1.76	99.70	239	4820	69.4
2129	1168.7	512.65	3.15	3.50	27.5	54.5	45200	5090	10100	212.0	71.3	1.77	100.00	232	4810	69.4
2132	1171.5	512.35	3.18	3.71	27.8	54.2	45400	5220	10200	213.0	72.2	1.78	101.00	231	4840	69.8
2133	1166.8	512.02	3.16	3.77	28.2	53.8	44500	5250	10800	211.0	72.5	1.78	100.00	233	4830	69.5
2138	1175.3	509.93	2.98	3.35	26.8	55.2	44200	4730	9730	210.0	68.8	1.73	98.60	251	4850	69.6
2139	1163.1	509.38	3.22	3.78	28.3	53.7	45000	5370	10270	212.0	73.3	1.79	101.00	217	4830	69.5
2141	1171.2	512.43	3.17	3.63	27.5	54.5	45800	5160	10200	214.0	71.8	1.78	101.00	232	4850	69.6
2148	1168.9	512.67	3.14	3.54	27.3	54.7	45400	5020	10100	213.0	71.0	1.77	100.00	233	4810	69.3
2148	1175.6	509.62	2.93	3.72	28	54.0	41700	4850	9380	204.0	69.8	1.71	98.80	347	4860	69.7
2151	1168.2	512.89	3.14	3.61	27.6	54.4	45000	5123	10190	212.0	71.6	1.77	100.00	233	4830	69.5
2152	1171.1	512.48	3.13	3.62	27.8	54.2	44800	5140	10000	212.0	71.6	1.77	100.00	232	4840	69.6
2201	1176.1	509.04	3.09	3.48	27.5	54.5	44300	4980	9880	211.0	70.5	1.76	99.40	233	4800	69.3
2204	1177.8	508.93	3.06	3.75	28.1	53.9	43300	5110	9770	208.0	71.5	1.75	99.80	237	4860	69.7
2208	1187.3	512.30	3.07	3.57	27.8	54.4	44300	5010	9890	210.0	70.8	1.75	99.40	225	4840	69.6
2215	1178.4	508.74	3.25	3.8	27.9	54.1	46600	5380	10500	216.0	73.3	1.80	102.00	233	4870	69.8
2216	1175.2	509.99	3.40	4.14	28.7	53.3	47300	5800	10800	217.0	76.2	1.84	104.00	216	4870	69.8
2216	1158.0	499.93	2.86	2.87	26.2	56.8	44200	4170	9410	210.0	64.6	1.68	97.00	239	4740	69.8
2221	1158.8	504.84	2.29	2.7	24.4	57.8	36600	3240	7650	191.0	56.9	1.51	87.50	197	4760	69.9
2225	1171.2	512.44	3.20	3.7	27.7	54.3	45900	5230	10300	214.0	72.3	1.79	101.00	232	4850	69.6
2226	1160.5	509.58	3.10	3.82	28.1	54.9	44300	5080	9930	211.0	71.2	1.76	99.80	232	4840	69.8
2228	1169.5	508.58	3.10	3.62	28.1	54.9	44300	5080	9930	211.0	71.2	1.76	99.80	232	4840	69.8
2231	1162.4	508.58	2.97	3.27	26.5	55.5	44000	4600	9620	210.0	67.8	1.72	98.10	236	4790	69.2
2232	1157.4	508.68	2.93	3.37	26.7	55.3	42600	4530	9360	209.0	67.3	1.71	98.70	225	4750	69.9
2233	1154.5	494.08	2.86	3.47	27	55.0	41100	4460	9090	203.0	66.8	1.69	95.30	236	4740	69.8
2234	1173.0	511.65	2.52	2.27	22.2	58.8	45300	3330	8960	213.0	57.7	1.59	94.80	284	4870	69.8
2235	1174.0	510.98	2.31	2.38	22.6	58.4	40800	3090	8130	202.0	55.6	1.52	90.20	246	4870	69.8
2236	1160.5	498.85	3.48	4.2	29.2	52.8	46200	5860	10600	215.0	76.5	1.87	103.00	235	4730	69.8
2237	1163.1	509.37	3.23	3.67	27.7	54.3	44800	5100	10000	212.0	71.5	1.80	100.00	230	4690	69.5
2239	1164.4	510.57	3.13	3.6	27.6	54.4	45100	5110	10100	212.0	71.5	1.77	100.00	232	4850	69.6
2241	1174.5	510.60	3.15	3.73	28.1	53.9	44900	5260	10100	212.0	72.5	1.77	101.00	227	4870	69.8
2246	1158.7	503.04	3.13	3.53	27.4	54.6	44800	5020	9990	212.0	70.9	1.77	100.00	234	4790	69.2
2248	1163.0	509.86	3.05	3.47	27	55.0	45200	4890	9970	213.0	69.9	1.75	99.80	234	4860	69.7
2251	1172.0	511.74	3.15	3.81	27.5	54.5	45700	4140	10200	214.0	71.7	1.77	101.00	232	4860	69.7
2252	1170.3	512.63	3.12	3.82	27.6	54.4	44900	5090	10000	212.0	71.4	1.77	100.00	242	4850	69.6
2253	1160.7	508.30	3.12	3.59	27.9	54.1	44200	5100	9920	210.0	71.4	1.77	99.60	232	4810	69.4
2255	1168.2	511.80	3.14	3.75	27.6	54.4	45300	5130	10100	213.0	71.8	1.77	101.00	231	4860	69.7
2259	1168.1	511.74	3.11	3.58	27.6	54.4	44800	5070	9960	211.0	71.2	1.76	99.80	232	4840	69.6
2264	1164.8	510.89	3.23	3.75	27.7	54.3	46300	5290	10400	215.0	72.8	1.80	102.00	231	4840	69.6
2287	1159.8	504.97	3.29	3.82	28	54.0	40300	5300	10400	215.0	73.4	1.81	102.00	227	4800	69.3
2290	1157.4	500.70	3.07	3.82	27.5	54.5	43800	4830	97							

Graficamente (fig. 8.1 a 8.7 ) se determinó que existía un rango entre el 25 y 75% de los valores (tabla 6), que definía dos grupos de estaciones (fig.9) los cuales presentaban el mismo comportamiento en todos los términos considerados:

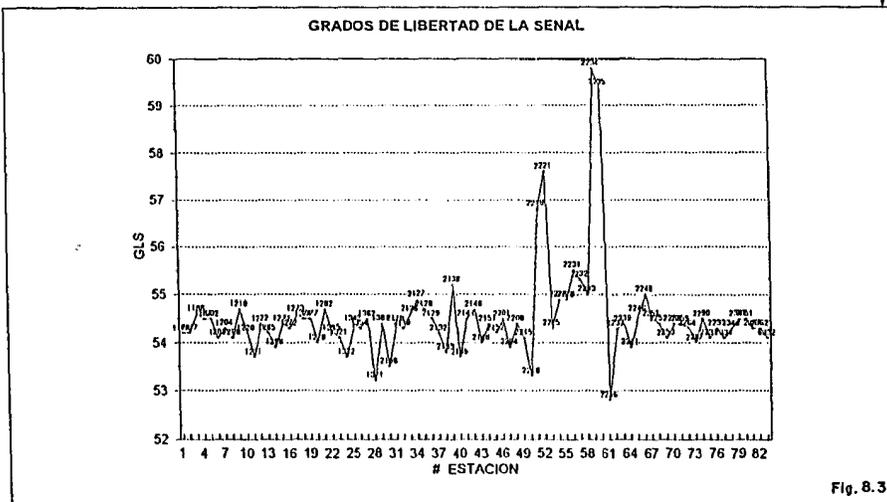
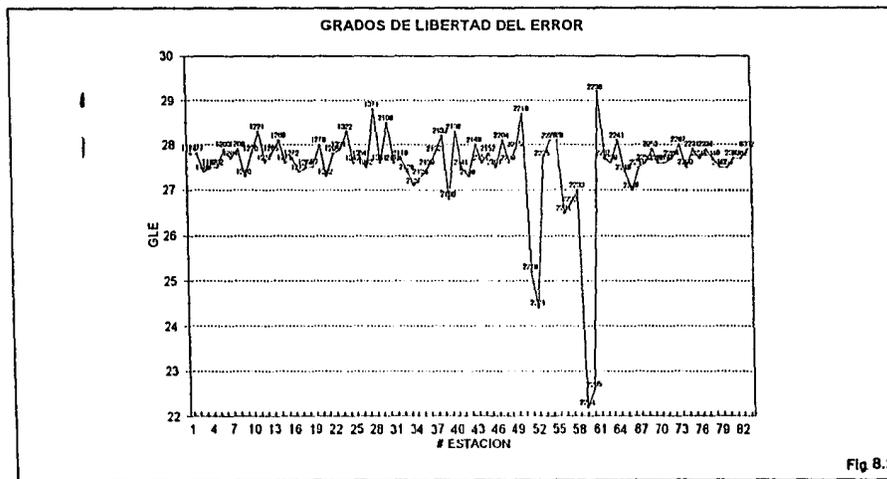
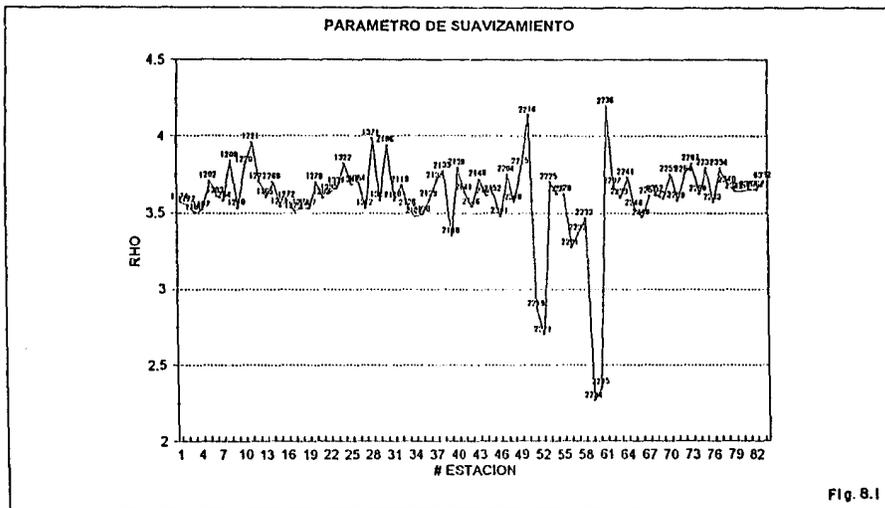
- Grupo A, cuyos valores se ubicaban dentro del rango y donde la característica principal es que se encontraban en las zonas más alejadas de la cuenca., a excepción de las estaciones 2233, 2255 y 2253 que se localizan dentro del área digitalizada.
- Grupo B, donde los valores estaban fuera del rango . En este grupo las estaciones extraídas se ubicaban dentro del área digitalizada a excepción de las estaciones: 2127, 2139, 2138, las cuales son necesarias para evitar el problema de extrapolación.

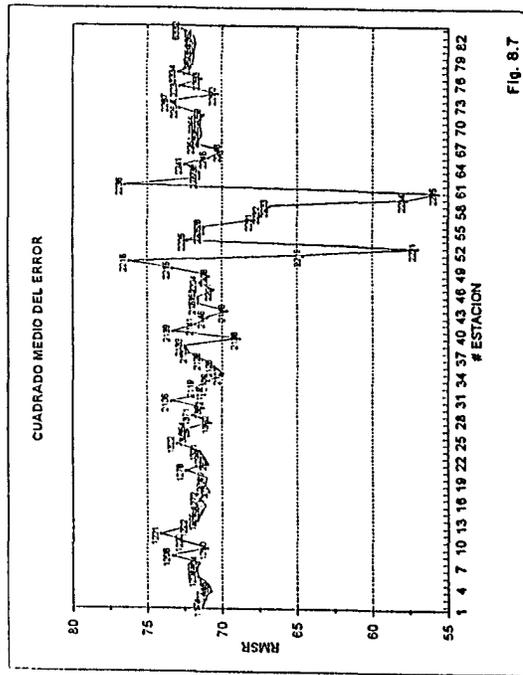
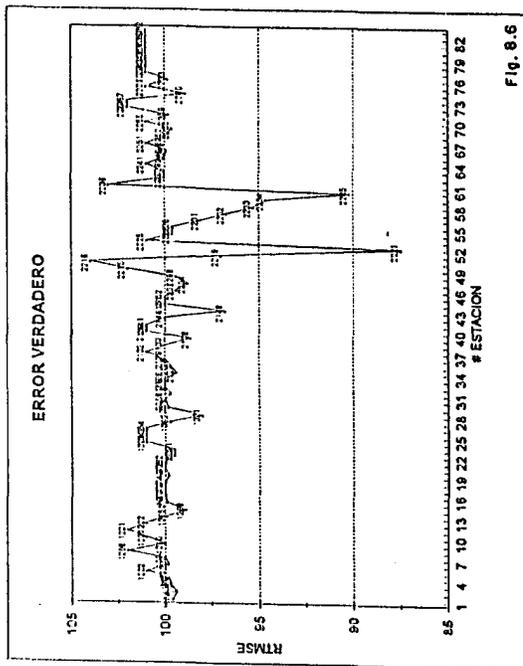
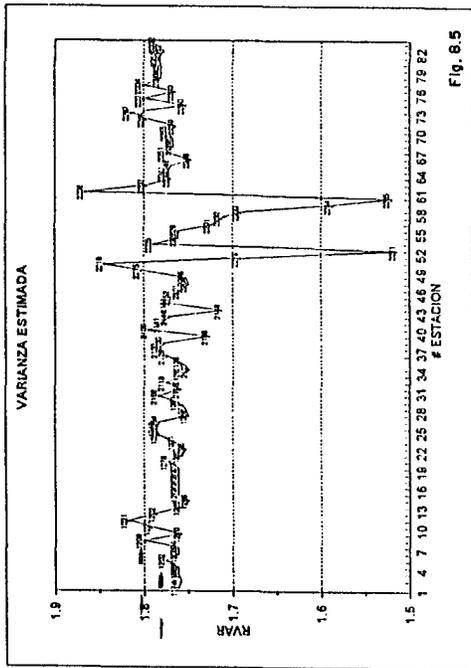
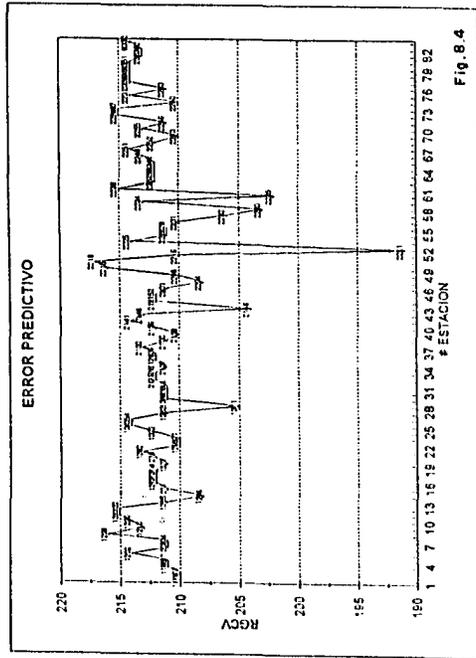
Además de ello, se determinó que aquellas estaciones cuya variación estuviese muy por encima o por debajo del rango definido, como es el caso de 2219, 2221,2216,2236, representaban puntos de observación que al ser eliminados podían originar un cambio significativo en el análisis espacial de la precipitación sobre el área de estudio , debido a que estas estaciones se ubicaban en localidades singulares de información, valles intramontanos o en lugares donde podía existir problemas de extrapolación.

### **7.2.2 MUESTREO SIN REEMPLAZO.**

El muestreo sin reemplazo, se realizó eliminando a partir del conjunto total de puntos de observación , aquella estación pluviométrica cuyo error en la estimación fuese mínimo, con la finalidad de alcanzar el número mínimo de puntos de

## MUESTREO CON REEMPLAZO

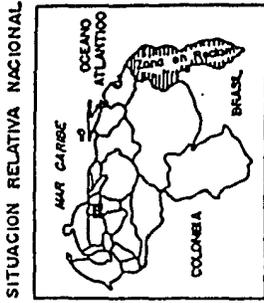




RANGO INTERCUATILICO  
(MUESTREO CON REEMPLAZO)

	CUARTIL INFERIOR	CUARTIL SUPERIOR	RANGO INTERCUATILICO
PRECIPITACION	1166.0	1175.1	9.1
PARAMETRO DE SUAVIZAMIENTO	3.5	3.7	0.2
GRADOS DE LIBERTAD DEL ERROR	27.5	27.9	0.4
GRADOS DE LIBERTAD DE LA SENAL	54.1	54.5	0.4
ERROR PREDICTIVO	211.0	214.0	3.0
ERROR MEDIO RESIDUAL	71.1	74.2	3.1
VARIANZA ESTIMADA	1.8	1.8	0.0
ERROR VERDADERO	99.6	101.0	1.4

TABLA 6

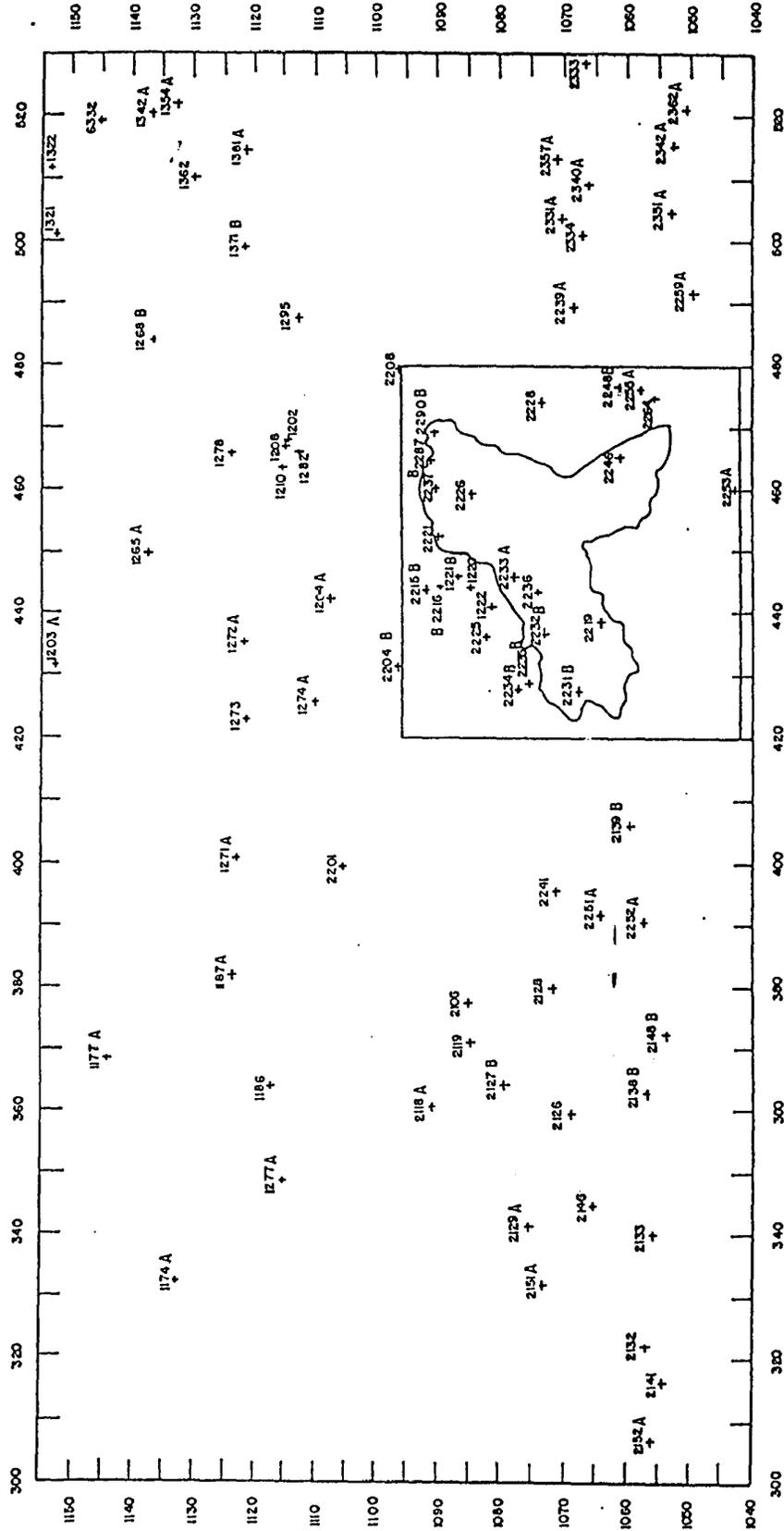


**LEYENDA**

(Serial)  
2232  
Estación Pluviométrica  
Área Digitalizada

# CUENCA DEL RIO YACAMBU

## CLASIFICACION DE ESTACIONES EN MUESTREO CON REEMPLAZO



A : ESTACION CON  $\Delta\theta < r_1$   
 B : ESTACION CON  $\Delta\theta > r_1$   
 r1 : RANGO INTERCUARTILICO  
 $\Delta\theta$ : VARIACION DEL ESTADISTICO

Fig. 9

observaciones necesarios para que el error cuadrático medio (MSR) no se incrementase en una cantidad estadísticamente significativa.

De esta manera, el tamaño de la muestra se varió desde  $n=83$  correspondiente al conjunto total de estaciones hasta  $n=8$  considerado por el modelo como el tamaño mínimo necesario para su desarrollo matemático .

En la tabla 7 se muestra la variación de los estadísticos, parámetro de suavizamiento y los grados de libertad del error y de la señal que caracterizaron a cada una de las superficies de precipitación generadas a partir del subconjunto de  $n$  estaciones seleccionadas.. Por otra parte se probó la bondad de ajuste del modelo en cada muestra (tabla 8) a un nivel de confianza del 95%, demostrándose la validez de cada una de las superficies de precipitación obtenidas y además se determinó que en cada muestra la GCV siempre es mayor a dos veces la varianza del error, aún cuando el ruido es siempre menor a  $n/2$

### 7.2.2.1 ANALISIS GRAFICO

El análisis gráfico se utilizó como una guía para seleccionar el número de estaciones a ser eliminadas a partir de la curva representativa de la variación de los estadísticos, parámetro de suavizamiento y los grados de libertad del error y de la señal, con respecto al tamaño de la muestra, al observar que el valor de cada término considerado tiende a ser constante hasta llegar a un tamaño de muestra en el cual el valor se dispara y comienza a presentarse un comportamiento inestable con valores altos, indicando que la variabilidad del campo observado pudiese estar explicado por este conjunto de estaciones. En este sentido, se pudieron notar las características siguientes :

MUESTREO SIN REEMPLAZO

TABLA 7

ESTACION	MUESTRA	MEDIA	SE	VAR	RHO	g1	g2	GCV	MSR	MSE	RTGCV	RTMSR	RTVAR	RTMSE	COV	mVr	Rv
2201	82	1176.2	509.0	3.09	3.5	27.5	54.5	44300	4953.1	9880	211	71.2	1.76	99.6	232	4840	69.6
2264	81	1171.5	510.5	3.22	3.6	27.1	53.9	46300	5094.9	10300	215	72.0	1.79	102.0	232	4810	69.4
1177	80	1178.1	510.2	3.21	3.5	26.8	53.2	46000	5020.9	10300	215	71.9	1.79	101.0	233	4800	69.3
1174	79	1181.8	512.3	3.21	3.5	26.6	52.4	45900	4997.5	10200	214	72.0	1.79	101.0	233	4810	69.4
1204	78	1192.1	507.4	3.20	3.5	26.2	51.8	46300	4924.8	10300	215	72.3	1.79	102.0	233	4860	69.7
2141	77	1194.3	510.4	3.27	3.5	25.6	51.4	47900	4924.2	10600	219	72.9	1.81	103.0	233	4880	69.9
1203	76	1204.5	505.8	3.26	3.5	25.5	50.5	47800	5067.8	10700	219	73.4	1.81	103.0	234	4920	70.1
1265	75	1215.3	500.3	3.26	3.5	25.3	49.7	47900	4952.4	10700	219	73.8	1.81	104.0	234	4970	70.5
2342	74	1212.7	503.2	3.33	3.5	24.7	49.3	49600	4964.1	11000	223	74.5	1.82	105.0	234	4970	70.5
2362	73	1209.4	505.9	3.33	3.4	24.3	48.7	50000	4872.6	11100	224	74.4	1.82	105.0	235	5000	70.7
2251	72	1213.9	508.0	3.38	3.4	23.7	48.3	51700	4828.8	11400	227	74.8	1.84	107.0	235	5030	70.9
1271	71	1223.6	504.8	3.37	3.3	23.2	47.8	52000	4708.1	11400	228	74.5	1.84	107.0	236	5040	71.0
1187	70	1232.0	503.5	3.36	3.2	22.8	47.2	52200	4639.6	11500	228	74.3	1.83	107.0	236	5050	71.1
2146	69	1232.2	507.2	3.40	3.1	22.0	47.0	53600	4544.6	11600	231	73.8	1.84	108.0	237	5020	70.9
2225	68	1235.3	510.3	3.53	3.2	21.6	46.4	55800	4621.7	12100	236	75.1	1.88	110.0	237	5030	70.9
1277	67	1241.4	511.6	3.50	3.1	21.1	45.9	56000	4695.7	12100	237	74.4	1.87	110.0	238	5030	70.9
1186	66	1249.5	511.2	3.48	3.0	20.8	45.2	55400	4400.6	12000	235	74.1	1.87	109.0	238	5010	70.8
1273	65	1261.3	510.0	3.45	2.8	20.1	44.9	55800	4396.7	11900	236	73.1	1.86	109.0	240	5010	70.8
1272	64	1275.1	497.5	3.43	2.8	19.9	44.1	56000	4589.8	12000	237	73.5	1.85	109.0	241	5070	71.2
1274	63	1288.2	490.3	3.38	2.7	19.3	43.7	56400	4829.4	12000	238	72.9	1.84	110.0	243	5120	71.6
2253	62	1278.5	488.2	3.40	2.6	19.1	42.9	56200	4900.4	12000	237	73.0	1.84	109.0	243	5090	71.3
2333	61	1274.2	491.0	3.41	2.6	18.8	42.2	56300	5013.7	12000	237	73.1	1.85	110.0	237	5090	71.3
2351	60	1272.7	495.0	3.41	2.5	18.3	41.7	57200	4793.4	12100	239	72.8	1.85	110.0	244	5100	71.4
2151	59	1274.2	499.3	3.42	2.4	17.7	41.3	58300	4477.8	12200	241	72.2	1.85	111.0	246	5100	71.4
1354	58	1276.2	503.3	3.51	2.4	17.2	40.8	61000	4466.3	12700	247	73.1	1.87	113.0	246	5140	71.7
1342	57	1277.2	507.7	3.41	2.2	16.2	40.8	62100	4124.3	12600	249	70.8	1.85	112.0	255	5170	71.9
2152	56	1281.5	511.2	3.77	2.1	15.8	40.2	61900	5087.7	12500	249	70.3	1.94	112.0	256	5190	72.0
2287	55	1269.1	507.3	3.78	2.5	16.1	38.9	66400	5516.4	13800	258	75.5	1.94	117.0	248	5140	71.7
2239	54	1263.3	510.3	3.80	2.4	15.6	38.4	68000	5470.0	14000	261	75.3	1.95	118.0	249	5160	71.8
2228	53	1252.5	508.9	3.77	2.3	14.9	38.1	68900	5278.9	13900	263	73.8	1.94	118.0	250	5140	71.7
1210	52	1265.8	504.5	3.72	2.2	14.2	37.8	70300	5444.4	14100	267	72.8	1.93	119.0	254	5220	72.2
2129	51	1266.6	509.5	3.66	2.0	13.6	37.4	70900	6798.6	13900	266	71.2	1.91	118.0	256	5180	72.0
2259	50	1262.9	514.0	3.59	1.9	13.3	36.7	70100	8474.1	13700	265	70.2	1.89	117.0	258	5180	72.0
2334	49	1259.3	518.6	3.96	2.2	13.4	35.6	74600	8378.6	14800	273	74.7	1.99	122.0	253	5160	71.8
2331	48	1253.5	522.5	3.93	2.0	12.7	35.3	76300	7787.1	14800	276	72.9	1.98	122.0	255	5110	71.5
2340	47	1250.2	527.7	3.89	1.9	11.9	35.1	79200	6759.6	15000	281	71.2	1.97	122.0	259	5150	71.8
2357	46	1243.5	531.5	3.88	1.8	11.8	34.2	77200	29207.5	14700	278	71.2	1.97	121.0	258	5100	71.4
2246	45	1225.5	523.1	3.73	1.6	11.0	34.0	77500	25641.6	1430	278	67.9	1.93	120.0	264	5030	70.9
1381	44	1226.6	529.0	3.68	1.5	10.5	33.5	77900	15349.6	14100	279	66.5	1.92	119.0	264	5050	71.1
6332	43	1219.6	533.2	3.69	1.5	10.2	32.8	79700	21758.4	14400	282	66.8	1.92	120.0	244	5100	71.4
1322	42	1224.6	538.7	3.59	1.4	9.7	32.3	80400	19194.2	14200	284	65.2	1.89	119.0	249	5150	71.8
1321	41	1226.6	545.2	3.57	1.4	9.6	31.4	78300	29798.0	14100	280	65.6	1.89	119.0	254	5150	71.8
1268	40	1237.8	547.3	3.53	1.4	9.5	30.5	77300	29951.0	14000	278	66.2	1.88	118.0	257	5210	72.2
1278	39	1253.5	545.3	3.54	1.4	9.4	29.6	78400	31739.6	14300	280	67.2	1.88	120.0	257	5320	72.9
2132	38	1260.1	551.0	3.56	1.4	9.3	28.7	77300	65426.3	14300	278	68.3	1.89	120.0	257	5340	73.1
2133	37	1256.2	558.1	3.50	1.3	9.1	27.9	76400	48323.2	14100	276	67.6	1.87	119.0	259	5340	73.1
2126	36	1267.4	561.8	3.34	1.2	8.3	27.7	78000	45171.8	13900	279	64.6	1.83	118.0	267	5410	73.6
2118	35	1277.7	566.5	3.25	1.1	8.0	27.0	77100	57321.6	13600	278	63.7	1.80	117.0	289	5450	73.8
1362	34	1279.4	574.9	3.30	1.2	8.1	25.9	76800	138506.3	13900	277	66.0	1.82	118.0	288	5550	74.5
1295	33	1287.1	582.3	3.22	1.1	7.8	25.2	77300	103373.7	13900	278	65.5	1.79	118.0	293	5660	75.2
1371	32	1277.2	588.6	3.24	1.1	7.7	24.3	76000	176419.5	1390	276	66.3	1.80	118.0	284	5650	75.2
2236	31	1246.3	571.3	3.82	1.2	7.5	23.5	84800	174431.8	15600	291	70.4	1.95	125.0	281	5380	73.3
1208	30	1267.9	568.0	4.75	1.7	8.1	21.9	98500	164400.4	19300	314	84.3	2.18	139.0	268	5520	74.3
1222	29	1265.1	577.8	5.12	1.8	7.7	21.3	10800	164623.9	21100	328	87.3	2.26	145.0	265	5600	74.8
2128	28	1275.9	585.5	4.73	1.5	6.8	21.2	109000	146433.6	20000	330	80.2	2.17	141.0	286	5590	74.8
2127	27	1285.8	594.3	4.65	1.3	6.3	20.7	102000	145868.0	19900	334	77.6	2.16	141.0	282	5580	74.7
2241	26	1306.2	596.2	4.45	1.3	6.0	20.0	111000	162481.3	19600	334	76.4	2.11	140.0	262	5720	75.6
2252	25	1314.5	607.1	4.11	1.1	5.4	19.6	111000	134763.5	18700	333	71.5	2.03	137.0	303	5790	76.1
2106	24	1315.9	620.1	3.66	0.8	4.4	19.6	117000	148835.0	17700	342	63.4	1.91	133.0	284	5930	77.0
2119	23	1335.2	626.6	3.38	0.7	4.0	19.0	115000	150536.0	16600	339	59.5	1.84	129.0	289	5970	77.3
2216	22	1364.1	625.4	4.94	1.1	4.5	17.5	148000	203600.7	24100	384	79.0	2.22	155.0	242	6140	78.4
2208	21	1364.9	640.9	5.20	1.2	4.6	16.4	149000	155577.2	25300	386	83.9	2.28	159.0	187	6230	78.9
2139	20	1348.6	653.1	6.21	1.8	5.4	14.6	145000	181008.4	28400	380	102.0	2.49	169.0	229	6260	79.1
2219	19	1302.5	636.6	4.41	0.8	3.5	15.5	142000	133009.1	21300	377	69.2	2.10	146.0	273	5920	76.9
2138	18	1336.3	637.3	5.21	1.3	4.1	13.9	137000	2061000.9	24200	370	85.2	2.28	156.0	606	6050	77.8
2148	17	1375.6	634.0	5.06	1.3	4.0	13.0	134000	425369.6	24100	366	86.1	2.25	155.0	636	6230	78.9
2226	16	1344.5	641.3	6.07	1.6	4.0	12.0	153000	444271.2	28800	391	98.1	2.46	170.0	600	6330	79.6
2290	15	1290.4	624.9	6.11	1.4	3.5	11.5	162000	354977.3	29000	403	93.9	2.47	170.0	541	6200	78.7
2237	14	1261.9	638.2	4.99	0.8	2.5	11.5	154000	323033.4	22300	392	68.9	2.23	149.0	487	5420	73.6
1221	13	1275.0	662.3	9.61	1.7	3.0	10.0	224000	434923.7	39800	474	109.0	3.10	199.0	488	5380	73.3
2248	12	1243.7	681.6	10.20	1.6	2.5	9.5	272000	406228.5	45100	521	110.0	3.19	212.0	529	5600	74.

# BONDAD DE AJUSTE MUESTREO SIN REEMPLAZO

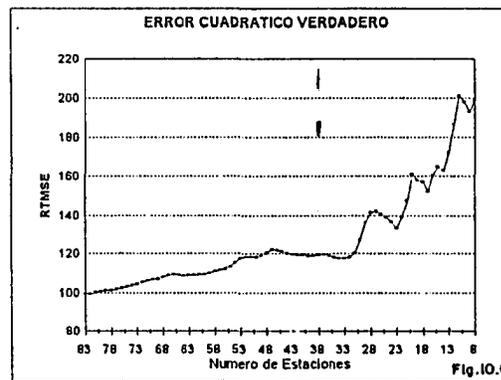
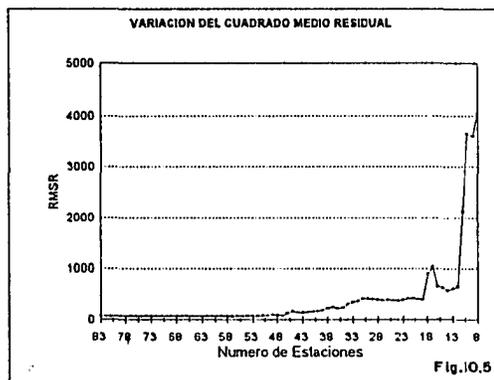
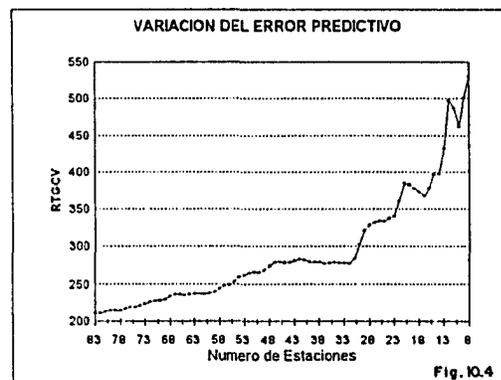
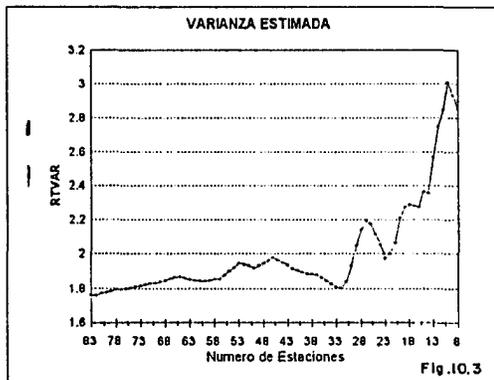
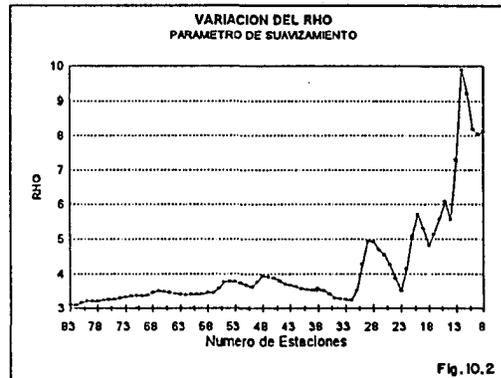
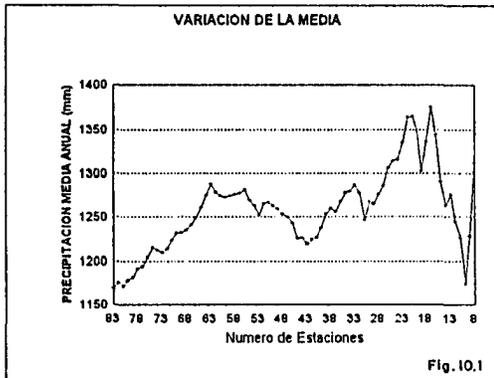
TABLA 8

ESTACION	MUESTRA	n/2	g <sub>0</sub>	MSR	SCR	σ <sup>2</sup>	χ <sup>2</sup> <sub>0,95</sub>	σ <sup>2</sup> χ <sup>2</sup>
1044	83	41.5	28.1	5080.0	421640.0	15004.0	41.3	620222.6
2201	82	41.0	27.5	4953.1	406151.8	14832.0	40.1	594860.0
2284	81	40.5	27.1	5094.9	412690.9	15488.2	40.1	621282.3
1177	80	40.0	26.8	5020.9	401870.8	15408.0	38.9	599142.1
1174	79	39.5	26.6	4997.5	394800.2	15440.1	38.9	600390.3
1204	78	39.0	26.2	4924.8	384134.4	15552.0	38.9	604741.5
2141	77	38.5	25.8	4924.2	379163.9	15957.6	37.7	600843.4
1203	76	38.0	25.5	5007.8	385149.5	16039.2	37.7	603915.8
1285	75	37.5	25.3	4952.4	371430.9	16202.2	37.7	610053.2
2342	74	37.0	24.7	4964.1	367342.3	16550.1	36.4	602672.3
2382	73	36.5	24.3	4872.8	355702.8	16650.0	36.4	608310.2
2251	72	36.0	23.7	4828.8	347673.8	17001.4	35.2	597981.1
1271	71	35.5	23.2	4708.1	334273.1	16984.8	35.2	597307.2
1187	70	35.0	22.8	4639.6	324771.2	16968.0	33.9	575630.2
2146	69	34.5	22.0	4544.6	313577.7	17068.0	33.9	579022.7
2225	68	34.0	21.6	4621.7	314273.6	17755.9	32.7	580095.2
1277	67	33.5	21.1	4695.7	314608.8	17605.0	32.7	575165.2
1186	66	33.0	20.8	4400.6	290441.0	17434.8	31.4	547634.4
1273	65	32.5	20.1	4306.7	285785.0	17284.5	31.4	542013.4
1272	64	32.0	19.9	4588.8	293745.4	17390.1	30.1	524188.6
1274	63	31.5	19.3	4820.4	304249.9	17305.8	30.1	521851.4
2253	62	31.0	19.1	4900.4	303826.2	17306.0	30.1	521663.5
2333	61	30.5	18.8	5013.7	305833.1	17356.9	28.9	501081.9
2351	60	30.0	18.3	4793.4	287601.3	17391.0	28.9	502066.4
2151	59	29.5	17.7	4477.8	284191.0	17442.0	27.6	481174.2
1354	58	29.0	17.2	4468.3	259045.8	18041.4	27.6	497709.0
1342	57	28.5	16.2	4124.3	235082.6	17029.7	26.3	463594.5
2152	56	28.0	15.8	5087.7	284911.9	18568.3	25.0	489075.3
2287	55	27.5	16.1	5516.4	303402.6	19420.2	26.3	510014.5
2239	54	27.0	15.6	5470.0	295379.2	19608.0	25.0	480117.8
2228	53	26.5	14.8	5278.9	279783.7	19377.8	23.7	458959.0
1210	52	26.0	14.2	5444.4	283110.4	19418.4	23.7	459920.6
2129	51	25.5	13.6	6798.0	346726.8	18958.8	22.4	423957.2
2250	50	25.0	13.3	8474.1	423707.4	18966.2	22.4	415848.7
2334	49	24.5	13.4	8378.6	410551.4	20433.6	22.4	450936.7
2331	48	24.0	12.7	7787.1	373782.3	20082.3	21.0	422251.8
2340	47	23.5	11.9	6759.8	317701.3	20033.5	19.7	394162.2
2257	46	23.0	11.8	29207.5	1343545.7	19788.0	19.7	389331.9
2248	45	22.5	11.0	25641.6	1153871.5	18741.8	19.7	368747.4
1381	44	22.0	10.5	16340.6	675384.1	18584.0	18.3	340217.8
6032	43	21.5	10.2	21758.4	935600.7	18810.0	18.3	344520.0
1022	42	21.0	9.7	19104.2	806154.8	18488.5	16.9	312808.2
1321	41	20.5	9.6	29798.0	1221718.6	18385.5	16.9	311083.5
1268	40	20.0	9.5	20951.0	1198030.5	18301.3	16.9	311161.7
1278	39	19.5	9.4	31730.8	1237846.3	18832.8	16.9	318631.4
2132	38	19.0	9.3	65426.3	2486200.8	19010.4	16.9	321636.2
2133	37	18.5	9.1	48323.2	1787958.7	18660.0	16.9	316215.4
2128	36	18.0	8.3	45171.8	1626186.5	18068.4	15.5	280207.8
2118	35	17.5	8.0	57321.8	2008256.0	17712.5	15.5	274673.3
1362	34	17.0	8.1	138508.3	4709214.2	18315.0	15.5	284018.4
1205	33	16.5	7.8	103373.7	3411331.1	18225.2	14.1	258378.2
1371	32	16.0	7.7	176419.5	5645422.8	18306.0	14.1	257512.8
2238	31	15.5	7.5	174431.8	5407385.1	20551.6	14.1	289102.0
1208	30	15.0	8.1	184400.4	4932011.3	26220.0	15.5	408601.7
1222	29	14.5	7.7	184623.0	4774093.7	28672.0	14.1	403332.7
2128	28	14.0	6.8	146433.6	4100141.4	26440.7	12.6	332930.1
2127	27	13.5	6.3	145868.0	3938437.1	25047.0	12.6	326713.7
2241	26	13.0	6.0	162481.3	4224512.7	25454.0	12.6	320560.0
2252	25	12.5	5.4	134763.5	3369086.5	23796.9	11.1	283443.2
2108	24	12.0	4.4	148835.0	3572039.2	21703.8	9.5	205918.8
2119	23	11.5	4.0	150536.0	3462328.2	20178.8	9.5	191448.1
2216	22	11.0	4.5	203600.7	4479216.0	30331.6	9.5	287778.0
2208	21	10.5	4.6	155577.2	3267120.2	32306.0	9.5	307384.5
2139	20	10.0	5.4	181008.4	3620168.4	38874.6	11.1	430360.6
2219	19	9.5	3.5	133009.1	2527173.7	26107.2	7.8	204020.8
2138	18	9.0	4.1	2061000.9	37098018.8	31520.5	9.5	280057.9
2148	17	8.5	4.0	425369.6	7231284.0	31523.8	9.5	298989.3
2226	16	8.0	4.0	444271.2	7108338.7	38423.1	9.5	364547.9
2280	15	7.5	3.5	354077.3	5324659.8	37882.0	7.8	286037.4
2237	14	7.0	2.5	323033.4	4522467.0	27045.8	6.0	162044.3
1221	13	6.5	3.0	438923.7	5654008.4	51701.8	7.8	404035.3
2248	12	6.0	2.5	406228.5	4874742.6	57120.0	6.0	342233.1
2265	11	5.5	2.5	1312479.7	14432789.9	47804.3	6.0	286777.8
2231	10	5.0	2.1	13386420.1	133864201.4	45399.0	6.0	272007.0
1202	9	4.5	1.5	12300505.4	111514548.5	46748.1	3.9	176580.7
1282	8	4.0	1.5	21090466.5	168795756.2	53747.2	3.8	208467.5

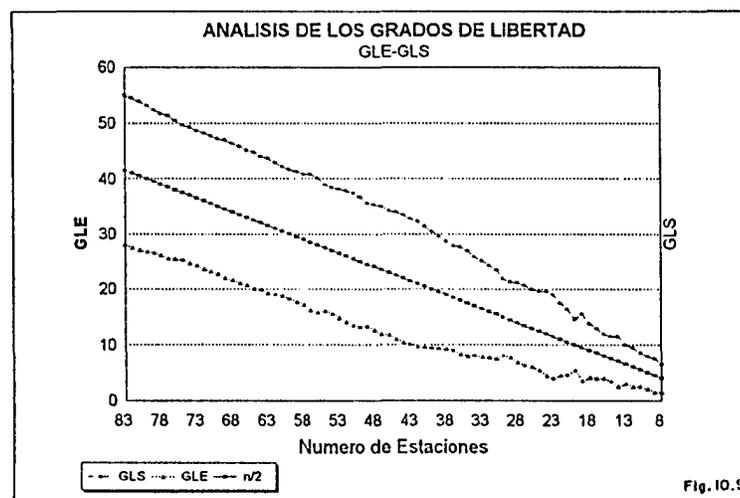
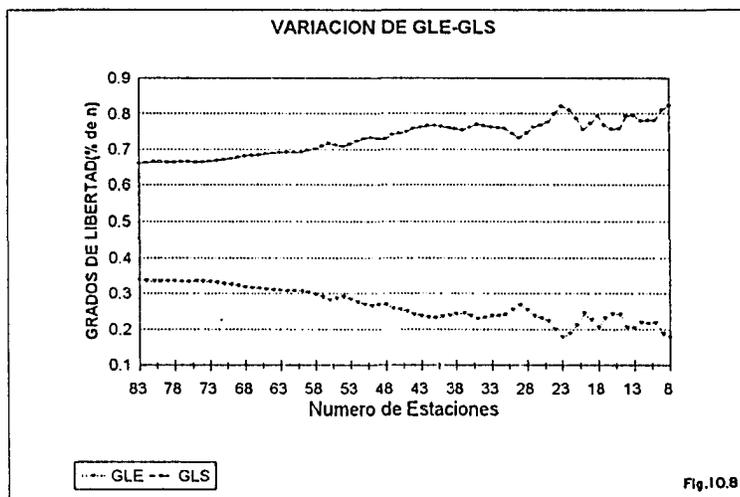
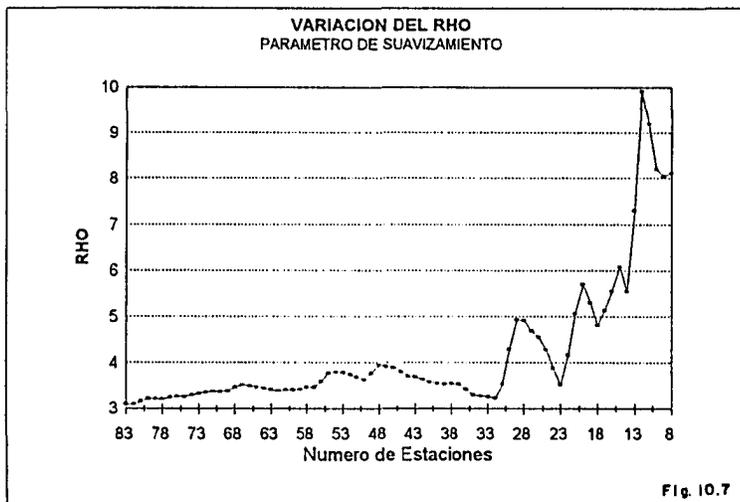
- a.- En las gráficas 10.1 a 10.6 representativas de la variación del parámetro de suavizamiento, la varianza estimada, error predictivo, cuadrado medio residual y del error cuadrático verdadero, se presenta en general un primer cambio de pendiente con un tamaño de muestra igual a 46, manteniéndose aproximadamente constante hasta llegar a una muestra de 33 donde se observa una fuerte variación en los valores.
- b.- Con respecto al porcentaje del número total de estaciones que se necesitan para identificar la señal y el ruido (fig.10.8) se pudo observar que a medida que el número de estaciones disminuye el porcentaje de estaciones para identificar la señal aumenta y a partir de cierto número de estaciones este porcentaje varía abruptamente. De esta manera se nota que existe un primer cambio de pendiente en la curva para un tamaño de muestra igual  $n= 55$  estaciones y luego a partir de una muestra de 33 puntos de observación se presenta un cambio brusco.
- c.- En la fig. 10.9 observamos que siempre los grados de libertad del error son menores a la mitad del número total de puntos de observación y tanto el ruido como la señal presentan el mismo comportamiento señalado en la fig.10.8 .

A partir de estas características, se infirió que el número de estaciones podría estar entre 33 ó 46 puntos de observación, cuyas superficies se muestran en las figs.11 y 12, respectivamente, con lo cual se verificó en cada caso el sentido físico del aporte de información de las estaciones seleccionadas.

MUESTREO SIN REEMPLAZO



## MUESTREO SIN REEMPLAZO



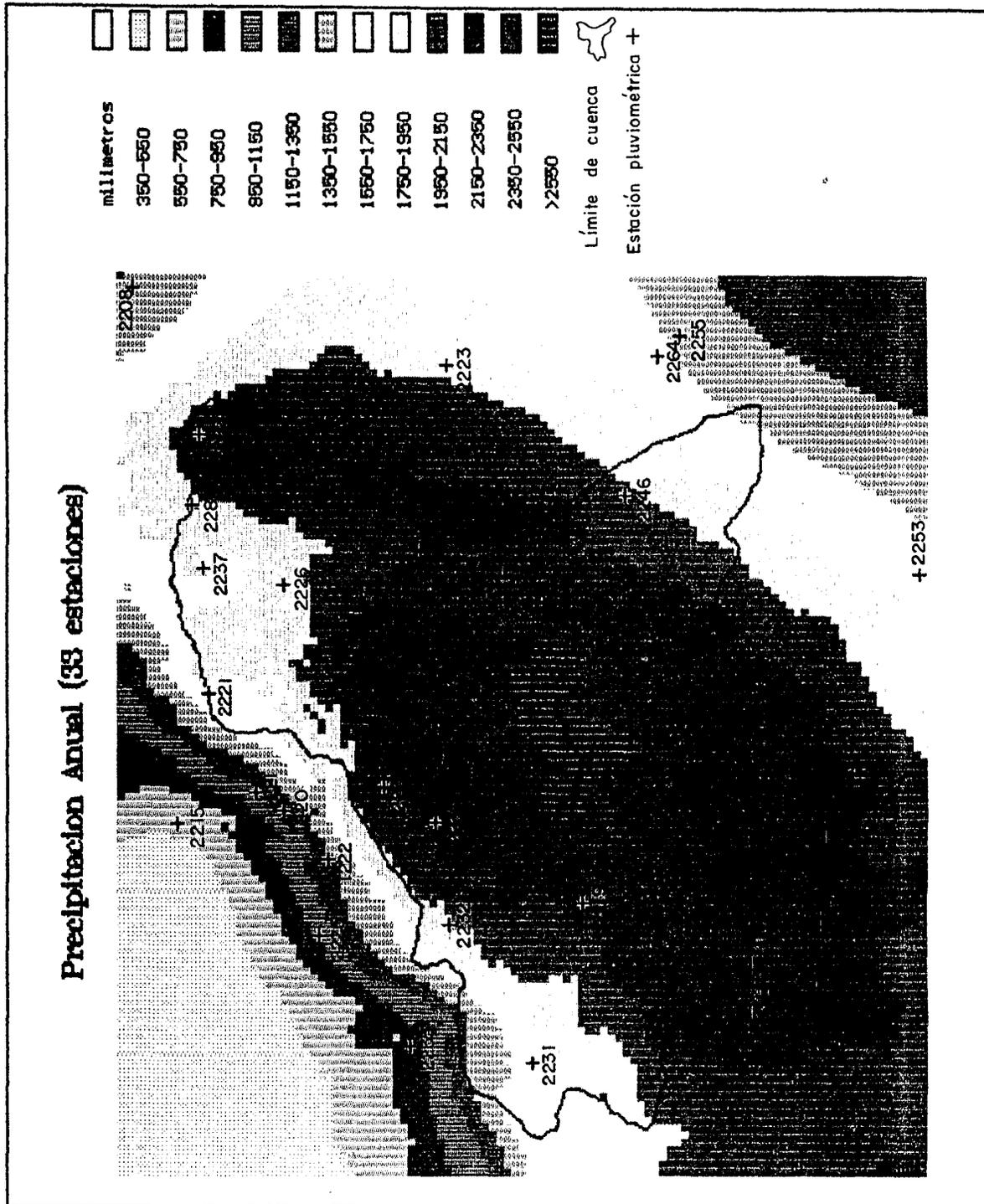


Fig. II

# Precipitación Anual (46 estaciones)



### 7.2.2.2 Análisis de varianza

A fin de precisar el número de estaciones, se aplicó la prueba F según lo expuesto en metodología y se plantearon las hipótesis siguientes:

$$H_0: \hat{f}_{S(i)} = \hat{f}_{S(i)};$$

$$H_1: \hat{f}_{S(i)} \neq \hat{f}_{S(i)};$$

donde la hipótesis nula significa que no existe diferencia en la superficie de precipitación estimada a partir de la información aportada por las muestra de estaciones de tamaño  $i$  y  $j$ , y la hipótesis alternativa, caso contrario.

En la tabla 7 observamos que existe un aumento de los valores del cuadrado medio residual ajustado (MSR), al extraer cada estación. Comenzamos el análisis con 81 estaciones y se comparó con el próximo mayor que él, si se encontraba una comparación no significativa se desechaban los valores previos no significativos y se comparaba con el valor MSR actual, en este caso el modelo fue de 55 estaciones, donde no se rechazó la hipótesis nula, luego no se consideró si no hasta el modelo con 46 estaciones, rechazando en este caso la hipótesis nula. En la tabla 9 se expresa de manera resumida la comparación de los modelos con su respectiva prueba F :

CALCULO DE LA PRUEBA F APROXIMADA PARA LA COMPARACION DE MODELOS

CONJUNTO DE ESTACIONES	n	MSR	g.l.n	SS( N )	g.l.d	SS ( D )	F	p
2264	81	5094.9						
2264 2287	55	5516.4	26	421.5	2	5094.9	0.006	NS
2287 2357	46	29207.5	9	23691.1	28	5516.4	13.36	**
2357 1321	41	29798	5	590.5	37	29207.5	0.1496	NS
1321 1268	40	29951	4	153	42	29798	0.054	NS
1268 1278	39	31739.6	1	1788.6	43	29951	2.57	NS
1278 2132	38	65422.3	1	33682.7	44	31739.6	42.7	**

TABLA 9

De esta manera, resultó que el subconjunto seleccionado estuvo formado por 46 estaciones, lo cual coincide con el resultado en el análisis gráfico, al presentarse un cambio de pendiente en la curva representativa de la variación de cada uno de los estadísticos considerados, en la muestra del mismo número de estaciones.

En la tabla 10, se presenta el listado de la muestra seleccionada, cuya ubicación geográfica se muestra en fig.13, denotando que ellas están localizadas dentro y fuera del área digitalizada, manteniéndose aquellas estaciones que estuviesen ubicadas en valles intramontanos, localidades singulares de información y como apoyo para evitar el problema de extrapolación. En general, se observa que estas estaciones forman en una franja en dirección suroeste -noreste representando, posiblemente, el efecto del viento en la zona analizada. Cabe destacar que el 37 % de estas estaciones al ser extraídas en el muestreo con reemplazamiento, originaron una variación fuera del rango intercuartílico de cada uno de los parámetros analizados. Las estaciones que no fueron seleccionadas se encuentran ubicadas en áreas donde la densidad de estaciones es muy alta o bien en zonas muy alejadas de la Cuenca, incidiendo poco sobre la variabilidad de precipitación dentro del sistema. Este grupo de estaciones coincidió en gran proporción el grupo B definido en el muestreo con reemplazo (fig.14).

## LISTADO DE ESTACIONES SELECCIONADAS

TABLA 10

Serial	Estación	Tipo	Alt. msnm	Latitud	Longitud
1202	BARQUISIMETO-FUNDACION	PC	528	100453	691740
1208	BARQUISIMETO-SE	PR	592	100512	691812
1220	AGUA NEGRA	PR	1560	94843	693036
1221	LAS TABLAS	PR	1330	94949	692953
1222	EL ZANCUDO	PR	1355	94700	693220
1268	DUACA	PR	750	101658	690849
1278	EL CUJI	C1	620	101000	691900
1282	BARQUISIMETO/B.A	SB	613	100400	691900
1295	YARITAGUA-EXPERIMENTAL	C1	375	100400	690700
1321	CAYURE	PR	320	102543	685934
1322	AROA DISPENSARIO	PR	278	102608	685333
1362	CAMPO ELIAS	PR	460	101317	685434
1371	URACHICHE	PR	435	100848	690039
1381	CENTRAL MATILDE	C3	235	100839	685211
2106	EL JABON	PR	1499	94845	700711
2118	EL EMPEDRADO	PR	844	95205	701624
2119	TRENTINO-LA PASTORA	PR	559	94844	701041
2126	VALLE HONDO	PR	800	94007	701657
2127	PUENTE VILLEGAS	PR	621	94548	701435
2128	AGUA DE OBISPOS	PC	2114	94133	700553
2132	AGUA VIVA	PR	110	93335	703731
2133	MONAY	PR	292	93300	702732
2138	LA CONCEPCION	PC	1125	93330	701508
2139	SABANA GRANDE-GUAGO	PR	1388	93514	695136
2148	CENDE	PC	3000	93149	701005
2204	QUIBOR	C2	682	95511	693739
2208	MAPORAL	PR	415	95435	691132
2215	SAN MIGUEL	PR	1097	95240	693103
2216	SAN MIGUEL	C1	1005	95129	693043
2219	MIRACUY	PR	1161	93744	693350
2221	CRUZ MACHADERA	PR	1845	95134	692623
2226	CAPILLA BUCARAL	PR	1050	94844	692229
2231	CASPITO	PR	1300	93934	693943
2232	PARQUE YACAMBU	PR	1484	94230	693436
2233	LA CRUZ	PR	1200	94500	692947
2234	SANARE	PR	1330	94433	693924
2235	SANARE	C2	1350	94400	693900
2236	LA ESCALERA	PR	1340	94300	693100
2237	RIECITO	PR	1280	95148	692147
2241	HUMOCARO BAJO	PR	941	94128	695719
2246	HACIENDA CAMBURITO	PR	285	93603	691910
2248	ARAURE	C1	200	93600	691300
2252	LA MESA	PR	1195	93353	695953
2255	ACARIGUA	PR	198	93408	691313
2290	LAS TROJAS	PR	1409	95156	691656
6332	EL TRIGAL	PR	1680	102140	684938

**SITUACION RELATIVA NACIONAL**

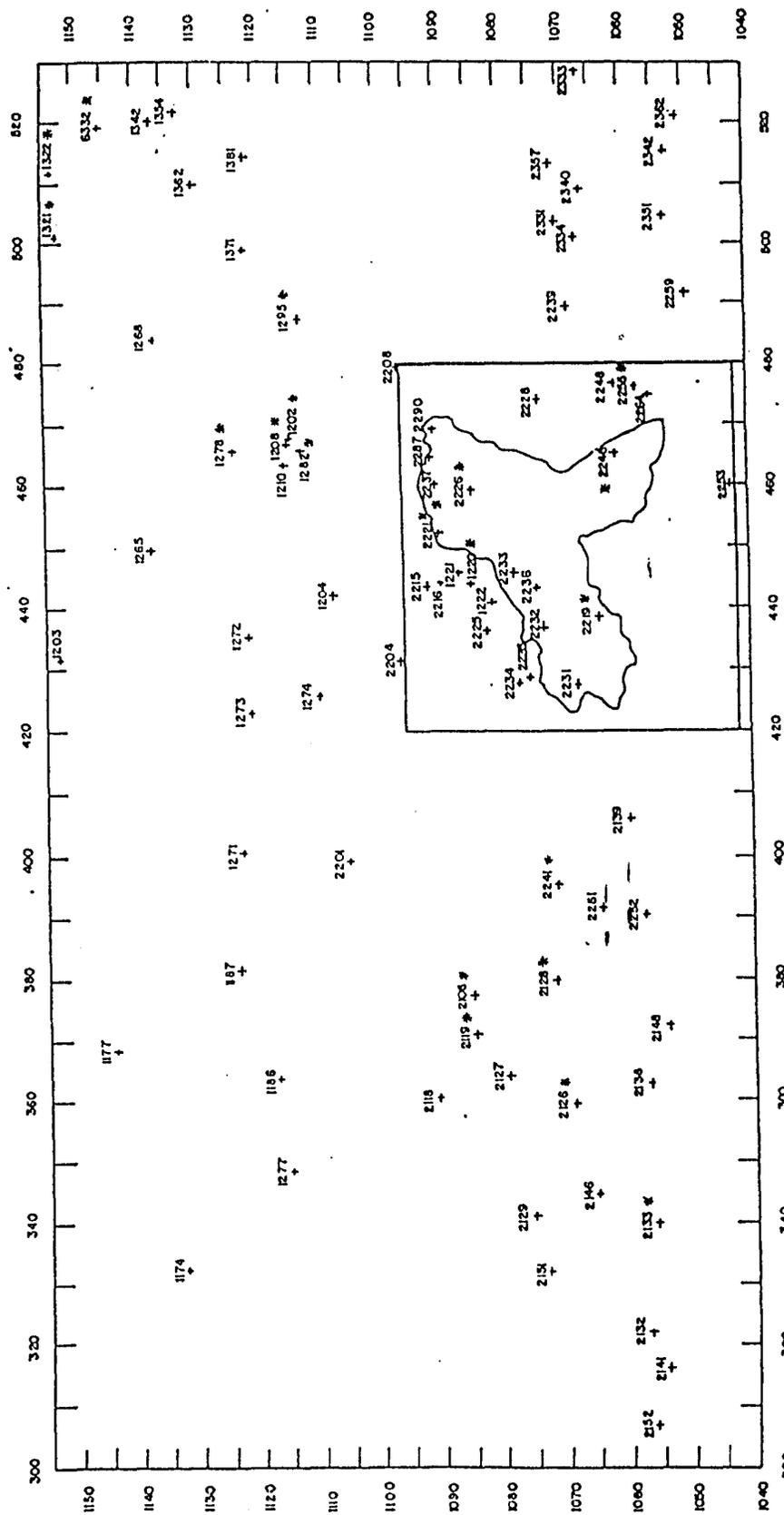
**LEYENDA**

(SeaHa)  
2232  
Estación Pluviométrica

Ante Digitalizada

# CUENCA DEL RIO YACAMBU

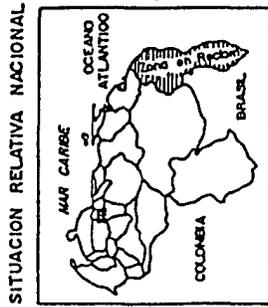
## ESTACIONES SELECCIONADAS



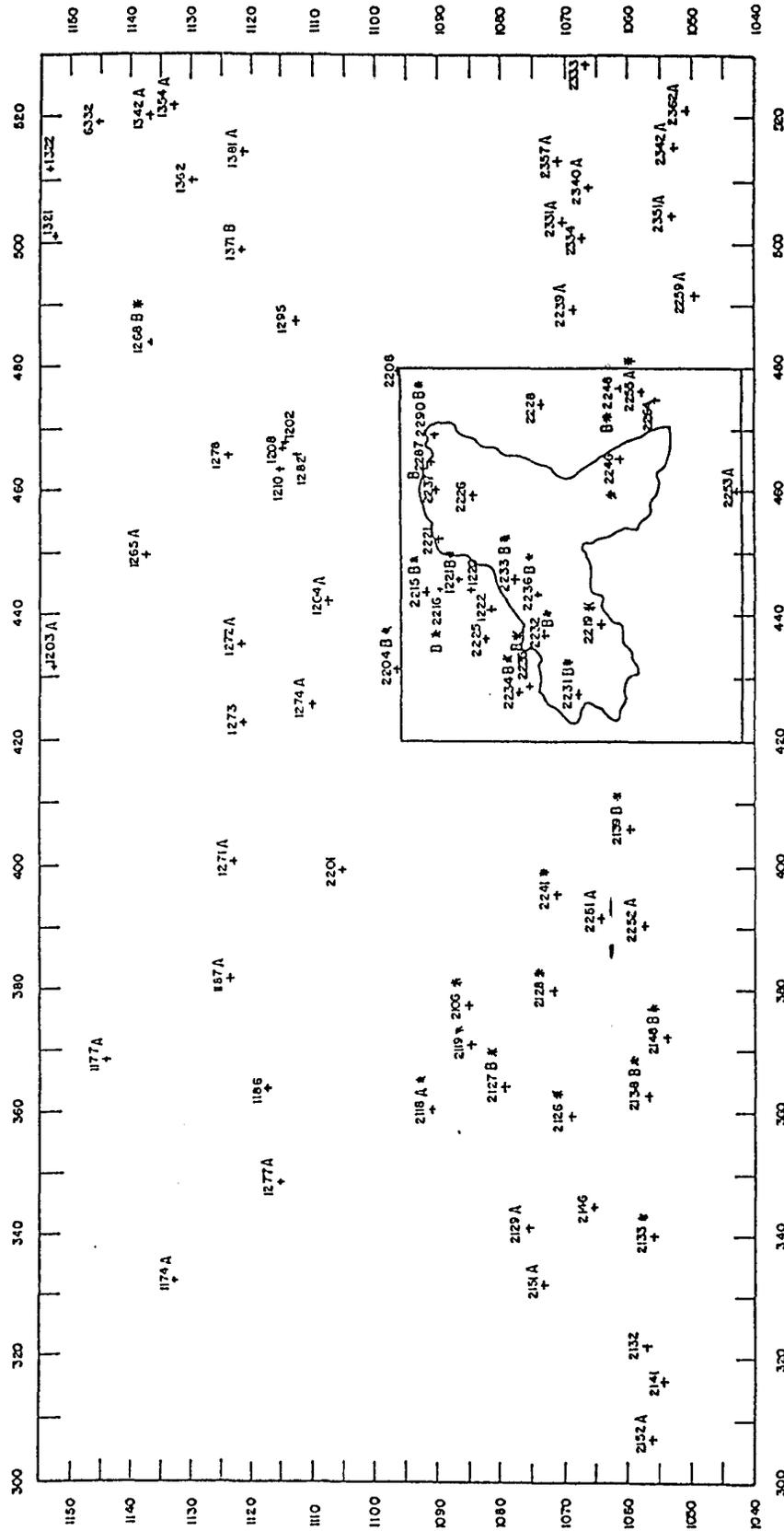
# CUENCA DEL RIO YACAMBU

ESTACIONES (CREM U SREM)

## LEYENDA



(6x101)  
 Estación Pluviométrica  
 Área Digitalizada



CREM = Muestreo con reemplazo  
 SREM = Muestreo sin reemplazo

Fig. 14

## CONCLUSIONES

Las conclusiones a las cuales se llegaron en el presente trabajo, fueron las siguientes:

- El modelo Spline Laplaciano, aplicado a través del programa ANUSPLIN, resultó ser una técnica eficaz para estimar la precipitación en el área de estudio seleccionada, proporcionando un análisis completamente objetivo y explícito.
- La superficie generada es válida en toda el área de estudio y no está sujeta al comportamiento singular en algunos puntos de observación.
- La varianza estimada en la superficie ajustada representa una medida de la validez del modelo.
- En general al aplicar el modelo Spline Laplaciano, se encontró que el orden de la derivada que mejor define la rugosidad de la superficie es igual a dos, alcanzando de esta manera el menor error predictivo en cada una de los distintos modelos analizados. Así mismo al examinar la validez del modelo, comparando la suma de cuadrados Residuales con el valor crítico de una  $\hat{\sigma}_E^2 \chi^2_m$  éste siempre mostró tener un buen ajuste, con un nivel de confianza del 95 %.
- La forma de aplicación del modelo Spline Laplaciano que mejor se ajustó al área de estudio, fue el modelo Spline parcial, en el cual se usó la elevación como única covariable en el sub-modelo lineal paramétrico. Se estimaron los valores máximos de precipitación en base al gradiente de precipitación estimado, alcanzándose un error predictivo del 18 % en la cantidad anual de precipitación, siendo más preciso que el modelo spline laplaciano bidimensional y tridimensional.

- El porcentaje de error de predicción estimado en la superficie ajustada permite que los valores precipitación sean usados en aplicaciones posteriores con un grado de confianza definido.
- El hecho de que los grados de libertad del error siempre fuesen menor a la mitad del número de puntos de observación y que la validación cruzada fuese menor a  $2\sigma_e^2$ , no resultó un factor relevante en el ajuste de la superficie interpolada, ya que se obtuvo un porcentaje de error de predicción de precipitación anual aceptable en los estudios hidrometeorológicos para fines de operación de embalses y manejo del recurso agua.
- La varianza del error de la superficie ajustada a nivel anual fue mayor que la varianza promedio en los datos observados. Esto se debió a la gran dispersión de los puntos de observación en algunas zonas del área de estudio y a la gran complejidad del relieve.
- A nivel mensual, se pudo verificar que durante los meses secos, la diferencia entre los valores estimados y los observados es menor que en los meses húmedos.
- Durante los meses secos, el parámetro de suavizamiento es mayor, presentándose además mayor ruido y menor señal en los datos.
- El error predictivo, medido por la  $\sqrt{GCV}$ , fue menor a nivel anual que a nivel mensual.
- En el caso de que se desee realizar inferencia sobre los valores estimados, se deberán transformar los valores observados por su raíz cuadrada, para que se cumplan los supuestos de normalidad y homoscedasticidad en los residuos.
- Con la realización de este trabajo se desarrolló una metodología la cual se basó en técnicas estadísticas, que de acuerdo a la bibliografía consultada, no habían sido utilizadas en el análisis espacial de datos, con fin de seleccionar el conjunto de puntos

de observación con el cual se pudiese representar la variabilidad de la precipitación en un área determinada. A través de ella, se pudo determinar que:

- El rango intercuartilico sirvió para indicar aquellos puntos de observación con mayor ponderación dentro de la variabilidad espacial de la precipitación.
- El error estimado en cada superficie ajustada permitió evaluar la densidad de la red con el fin de obtener redes más pequeñas, lo suficientemente representativas.
- El análisis grafico, aunque fue utilizado de manera empírica, resultó ser una buena guía para determinar el número de estaciones seleccionadas al inferirse que el tamaño de la muestra estaba determinado al presentarse un cambio de pendiente en la curva de variación de la raíz del cuadrado medio residual (RMSR) con respecto al número de estaciones, coincidiendo a su vez, este punto en las la curvas de variación de los estadísticos restantes (RGCV, RTMSE, RVAR ), parámetro de suavizamiento (RHO) y en el aporte de señal y ruido de los datos en cada una de las muestras.
- El uso de la la prueba F aproximada, se ajustó bien al proceso de selección del número mínimo de estaciones, considerando a las estaciones pluviométricas como variables independientes y aplicando un proceso de ‘eliminación hacia atrás’ similar al que se utiliza en un modelo de regresión.
- A partir de la metodología aplicada, se logró seleccionar una muestra óptima de estaciones, minimizando la red de estaciones pluviométricas de 83 a 46 puntos de observación, con lo que reducirá en un 45% los costos por mantenimiento y operación de estaciones pluviométricas en el Sistema Hidráulico Yacambú -Quibor.

## ***RECOMENDACIONES***

A partir de los resultados obtenidos en el presente trabajo de investigación se señalan las recomendaciones siguientes:

1. Extender el estudio a nivel nacional, a fin de generar un mapa digital de precipitación, a partir de la información puntual de la red nacional de estaciones pluviométricas, a nivel mensual y anual.
2. Generar un Modelo Digital de Terreno que permita representar la información topográfica de todo el país.
3. Rediseñar la red de estaciones pluviométricas tanto en la cuenca del Río Yacambú como en su área de influencia, modificando el número y la ubicación de los puntos de observación.
4. Aplicar a nivel nacional los métodos y procedimientos estadísticos utilizados en este trabajo, con el fin de obtener una la red mínima de estaciones pluviométricas, que se ajuste a la situación económica del país, pero que con ella se pueda tener una buena estimación de precipitación en las diferentes regiones del país.
5. Extender el estudio a otras variables climatológicas e incluir otros factores fisiográficos.

## BIBLIOGRAFIA

- Bosque S. J (1992). Sistema de Información Geográfica. Ed. Rialp
- Daly C, Neilson R. and Phillips (1994) A Statistical-Topographic for Mapping climatological precipitation over mountainous terrains. Journal of Applied Metyeorology, Vol 33, No. 2, Feb.1994,140-158
- Davis J. 1986. Statistics and Data Analysis in Geology. Jhon Wiley and Sons, 1986.
- Draper and Smith, Applied Regression Analysis. Wiley Series in Probability and Mathematical Statistics, 1966
- Graybill, F. Theory and Appication of the Linear Model. Duxbury Press,1976
- Mood, Graybill and Boes. Introduction to the Theory of Statistics. McGraw -Hill,1974
- Guenni, L, 1995. A model for seasonal variation of rainfall at Adelaaide and Turen. Ecological Modeling,1474 (1195) xxx C.
- Guenni, L,1995. Spial interpolation of rainfall statistics al different times scales. Prociding of six International of Statistical Climatology. Irlanda.
- Hutchinson, M.F (1984 ). A summary of some surface fitting and contouring programs for noisy data. Consulting Report ACT 84/6. CSIRO Division of Mathematics and Statistics, Canberra, Australia.
- Hutchinson, M.F, (1989a). A new objetive method for spatial interpolation of meteorological vartiables from irregular networks applied to the estimation of monthly mean solar radiation, temperature, precipitation and windrum. CSIRO Division of Water Resourse Tech. Memo. 89/5.
- Hutchinson, M.F, (1989b). A new prtocedure for grdding elevation and stream line data with automatic removal of spurious pits. Journal of Hydrology 106, 211-232.
- Hutchinson, M.F, (1991a). The application of thin plate smoothing splines to continent-wide data assimilation. In: J.D. Jaspewr (de), BMRC Research Report N° 27, "Data Assimilatibn Systems ", Bureau of Meteorology, Melbourne,104-113.
- Hutchinson, M.F, (1991b). Climatic analyses in data sparse regions. In: R. C. Muchow and J.A Bellamy(editor), "Climatic Risk in Crop Production", CAB International, 55-71
- Hutchinson, M.F Spatio-temporal analysis of monthly rainfall ( borrador no publicado). Centre for Resource and Environmental Studies. Australian National University, 319-322
- Hutchinson, M.F ( 1994 ) Interpolating mean rainfall using thin plate smoothing splines. ( borrador no publicado). Centre for Resource and Environmental Studies. Australian National University
- Hutchinson, M.F and R.J. Bishof (1983) A new method for estimating the spatial distribution of mean seasonal and annual rainfall applied to the Hunter Valley, New South Wales. Australian Meteorological Magazine 31, 179-184.
- Hutchinson, M.F and Johnson C.M. Accounting for Srong spatial correlations in the interpolation of mean rainfall for a standard period. Centre for Resource and Environmental Studies. Australian National University, 315-318.

- Hutchinson, M.F and Gessler, P.E (1994 ) Splines -more than just a smooth interpolator. *Geoderma* 62: 45-67.
- Hutchinson, M.F and Guenni L., (1992 )Documentation on how to use ANUSPLIN. Reporte no publicado
- Hutchinson, M.F and Guenni L , (1992). Procedures for spial interpolation of long term mean monthly rainfall. Reporte no publicado
- Hutchinson, M.F and Guenni L Interpolation of monthly climate means (1992). Reporte no publicado.
- Jones D.A, Gurney R.J and O'Connel (1979): Network Design using optimal estimation procedures. *Water Resources Researh*, Vol. 15, N<sub>o</sub> 6, Dic,1979.
- Montilva J. (1995). Apuntes de clase. Curso : Diseño y manejo de base de datos para GIS. CIDIAT
- Montgomery D. Introduction to linear regression analysis. Second Edition. Jhon Wiley & Sons, INC
- Orlandoni G.. Ajuste de Funciones Segmentadas (funciones Spline) por mínimos cuadrados. Publicación 83-06.Instituto de Estadística Aplicada y Computación. Universidad de Los Andes.
- Ponte R. (1983). Determinación de la precipitación media y análisis de profundidad - área - duración de tormentas por el método multicuadrático en Venezuela. Trabajo de Ascenso, Universidad de Los Andes, Dic. 1983.
- Seaman R. S. and Hutchinson, M.F (1985). Comaparative real data tests of some objetive analysis methods by whitholding observations, *Australian Meteorological Magazine* 33, 1985, 37-46.
- Thieboux H. and Pedder M.(1987). Spatial Objective Analysis whit applications in Atmospheric Science. Academic Press,1987
- Zuo H.,Hutchinson, M.F, McMahon J.P. Estimating the spatial distribution of monthly mean and extreme temperatures across China. Centrrc for Resource and Environmental Studies. Australian National University, 319-322
- Wahba, G. (1990). Spline Models for Observational Data. CBMS- NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia.
- Wahba G., (1980) Splines based, regularization, and generalized cross validation for solving appoximation problems wiht large quantities of noisy data. *Aproximation Theory III*, W. Cheney, de. Academic Press, 905-912
- Watson G.S (1984 ) Smoothing and Interpolation by krigging and Splines. *Mathematical Geology*, Vol,16, N<sub>o</sub> 6, 1984

**PROGRAMAS UTILIZADOS, CARACTERISTICAS,  
DATOS DE ENTRADA Y SALIDAS**

**ANEXO 1**

# **PROGRAMAS UTILIZADOS, CARACTERISTICAS, DATOS DE ENTRADA Y SALIDAS**

## **1.1 INTRODUCCION**

Los programas utilizados en el presente trabajo forman parte del software realizado por el Dr. Hutchinson (1990 ), en la Universidad Nacional de Australia, con la finalidad de facilitar la aplicación del modelo Spline Laplaciano en campos relacionados con la hidrometeorología.

En el desarrollo de este software se contemplan dos programas principales: el ANUDEM con el cual se genera un modelo digital de terreno y el ANUSPLIN en el que se ajusta una superficie spline en una región de datos puntuales, distribuidos de manera irregular, en donde cada valor es asociado con un número de variables geográficas independientes, tales como la latitud, longitud y elevación.

Los programas fueron realizados en FORTRAN 77 estandar y requieren el apoyo de un sistema de información geográfico ( SIG ) para la captura, manejo y visualización de la representación espacial de los datos.

## **1.2 PROGRAMA ANUDEM**

### **1.2.1 CARACTERISTICAS**

El programa ANUDEM ha sido diseñado con la finalidad de generar modelos digitales de terreno, obtenidos al calcular sobre una grilla regular, los valores de superficies suavemente discretizadas ajustadas a un gran número de datos puntuales de altura (metros sobre el nivel del mar) espaciados irregularmente, curvas de nivel y red hidrográfica de una determinada región geográfica. Normalmente esto se realiza

imponiendo dos criterios opcionales, que se asumen en los datos de entrada : el primero que los datos puntuales de altura incluyan todos los máximos locales importantes y el segundo que en los datos se especifiquen los mínimos locales significativos, denominados puntos sink.

El algoritmo implementado por el programa interpola los datos de elevación en una grilla regular minimizando la penalidad por rugosidad, convenientemente débil, en los valores de la grilla ajustada e imponiendo simultáneamente las restricciones siguientes:

- \* La primera que es opcional, se incorpora simplemente permitiendo que los máximos locales se ubiquen solamente en los puntos de la grilla en los cuales se ha localizado los datos puntuales, con lo cual se remueven los máximos falsos generados por las técnicas de interpolación para aplicaciones generales. Esta proposición puede ser omitida cuando los datos de entrada de altura están dados en curvas de nivel (líneas de puntos con igual altura o cota) en donde no se incluyen los máximos locales, por lo que en este caso el programa interpola las alturas de los máximos locales en base a la información de las curvas de nivel circundantes.
- \* La segunda fija la estructura de drenaje. Esto se realiza imponiendo una condición global de drenaje en los valores de la grilla ajustada, que automáticamente remueve los puntos sink falsos al calcular las restricciones impuestas por el drenaje en el modelo digital de terreno (DEM) calculado directamente a partir de los datos de la hidrografía en el área de estudio.

Estas son las innovaciones principales del programa con las cuales se logra la eliminación de las principales fallas de los modelos de terreno producidas por las técnicas de interpolación de propósito general que limitan su utilización en aplicaciones hidrológicas. Se ha demostrado que la condición global de drenaje impuesta, se convierte en la práctica, en una condición muy poderosa que puede incrementar significativamente

la exactitud, principalmente en las características del drenaje en modelos digitales de terreno interpolados a partir de conjuntos dispersos de datos específicos de superficies. El tamaño de tales conjuntos de datos puede ser al menos de un orden de magnitud más pequeño que el número de datos requerido para describir adecuadamente la altura, usando curvas de nivel. Esto minimiza el costo de la obtención de modelos de terreno confiables en términos de recolección, corrección y almacenaje de datos primarios de elevación. También elimina la necesidad de editar manuales detallados de grillas de elevación interpoladas para eliminar rasgos falsos de drenaje.

Por otra parte, el programa actúa en forma conservadora cuando se intenta eliminar sinks y no impone condiciones de drenaje que podrían contradecir los datos de altura. Una consecuencia de esto es que los errores tanto en altura y posición de los datos de entrada pueden a menudo estar indicados por sinks en la grilla ajustada al final, especialmente cuando los datos de entrada incluyen al menos la red del cauce principal. Esto genera muy buenos resultados al procesar conjuntos de datos muy grandes en donde el programa puede ocasionalmente no tomar en cuenta información de diagnóstico para cada cauce lo que sería muy útil en la corrección de datos. El número de tales sinks es usualmente muy pequeño .

### **Algoritmo de forzamiento de drenaje**

La condición global de drenaje es impuesta en un intento de remover todos los mínimos locales que no han sido identificados como tales en los archivos de entrada de puntos sink .

El algoritmo es controlado por la aplicación sistemática de tres tolerancias introducidas por el usuario, las cuales permiten que éste ajuste la potencia del forzamiento de drenaje en relación tanto a la exactitud como a la densidad de los datos de altura.

La acción de las tolerancias se hace menos crítica a medida que se mejora la exactitud y densidad de los datos. Cuando la tolerancia ha sido ajustada apropiadamente, los puntos sink no muy claros en el programa son aquellos asociados, bien con errores de elevación significantes en los datos de entrada o a áreas donde los datos de entrada no tienen densidad suficiente para resolver confiadamente las características de drenaje en la grilla ajustada.

La primera tolerancia dada por el usuario debe reflejar la exactitud de la altura de los datos de entrada pero puede también reflejar la densidad de los datos de altura . Las diferencias de altura entre los puntos de datos que no exceden estas tolerancias se toman como insignificantes con respecto al drenaje. Así los datos puntuales que bloquean el drenaje hasta un máximo de estas tolerancias son eliminados. Cuando los datos puntuales no son lo suficientemente densos para resolver con exactitud el drenaje esta tolerancia puede ser aumentada, lo cual lleva a un patrón de drenaje casi generalizado a expensas de la fidelidad de los datos de altura. Esto es especialmente útil cuando se trabaja a escalas más pequeñas (digamos menos 1:100.000) .

La segunda tolerancia limita la altura de cada dato puntual de profundidad que exceda al dato puntual que pueda considerarse como un mínimo local significativo. Esto permite eliminar un dato puntual, aún cuando halla sido asociado con un punto sink, para obtener una imagen clara del drenaje . Esta tolerancia normalmente tiene un efecto significativo, solo cuando los datos puntuales están relativamente espaciados . Cuando la densidad de los datos es suficiente como para resolver las características del drenaje, la segunda tolerancia, se podrá ajustar a su mínimo valor permisible, el cual es dos veces la primera tolerancia.

Si se ajusta la segunda tolerancia a un valor mayor se asegura que el programa actúe conservadoramente cuando se desee remover puntos sink y consecuentemente quedaran más de ellos en la grilla final. Esto es apropiado cuando la densidad de los datos no es

suficiente para resolver las características de la red hidrográfica en la grilla ajustada. En este caso, el usuario puede remover más puntos sink, agregando más información tanto puntual como de drenaje en aquellas áreas consideradas como críticas.

El forzamiento de drenaje también puede ser obtenido incorporando los datos de la red hidrográfica. Esto es útil, cuando se requiere un emplazamiento más exacto de los ríos del que puede ser calculado automáticamente por el programa. También se utiliza para remover sink que no son eliminados por el algoritmo de forzamiento de drenaje automático. Este es de hecho la forma recomendada de corregir anomalías en la línea de descenso del drenaje presentes en las grillas de altura calculadas por el programa. Todos los datos de altura que no concuerdan con la pendiente de cada línea de la red de drenaje son removidos y aquellos puntos que excedan la tercera tolerancia de altura especificada por el usuario serán salidas para posibles correcciones de errores.

Las condiciones laterales son dadas por las líneas que representan a los ríos. Esto asegura que el cauce actúa como una línea divisoria para la interpolación y simultáneamente asegura que cada río llegue al fondo de su valle respectivo. Las condiciones laterales no están ajustadas a datos puntuales que se encuentren fuera de la línea de cauce y cuyas alturas sean mayor a la tercera tolerancia bajo la altura del cauce. Estas discrepancias se indican en los archivos de salidas para posibles correcciones de errores. La tercera tolerancia también se usa para impedir la remoción automática de

sinks, cuyas elevaciones exceden la altura de los sink por más de dos veces el valor de la tercera tolerancia. Estas situaciones son normalmente causadas por errores en las cotas, pero también pueden originarse debido al espaciamiento entre los datos puntuales.

### **1.1.2 DATOS DE ENTRADA Y SALIDAS**

El programa procesa archivos diferentes de entrada de datos, los cuales están dados como:

1. Datos de altura
2. Datos sink
3. Datos sobre el drenaje
4. Curvas de nivel

El programa lee los datos de cada archivo de entrada y los clasifica en ventanas según los límites del mapa especificados generalizándolos luego en una la resolución de la grilla impuesta por el usuario . Los datos puntuales son generalizados aceptando un máximo por celda de grilla, generalmente igual a cuatro y descartando cualquier otro punto de datos restantes . Los datos de la red hidrográfica son generalizados aceptando a lo máximo una línea de datos por celda de grilla, eliminando los rizados innecesarios.

Los datos puntuales seleccionados son escritos en archivos no formateados. Se dispone de un archivo para cada conjunto de datos, lo cual facilita la recuperación rápida durante la ejecución. El programa luego emplea un método de multigrilla simple que calcula las grillas a resoluciones más finas sucesivamente, comenzando de la grilla gruesa inicial hasta la final, según la resolución dada por el usuario y simultáneamente impone condiciones de drenaje para remover depresiones donde sea posible al algoritmo descrito anteriormente.

Para cada resolución de grilla los datos puntuales aceptados son leídos de los cinco archivos improvisados y los valores en los puntos de grilla no ocupados por los datos

puntuales son calculados por la iteración GAUSS-Seidel, con método SORT sujeto a una penalidad de rugosidad apropiada y limitada a una cadena ordenada, cuyas restricciones son obtenidas de los datos drenaje dados por el usuario. Los datos de curvas de nivel y de depresiones (sink) y el forzamiento de drenaje se ejercen de una forma automática establecida por el programa. La iteración se termina cuando el número máximo de iteraciones dadas por el usuario, (entre 30 y 35 ) han sido alcanzadas. Los valores de inicio para cada grilla mas fina sucesiva son interpolados linealmente de la grilla mas gruesa precedente.

Una vez terminadas las iteraciones de multigrilla, el programa calcula todos los puntos sink remanentes en la grilla ajustada y opcionalmente escribe un resumen detallado a la salida. Estos puntos sink son también opcionalmente escritos al archivo especificado por el usuario, para propósitos de ploteo. El programa opcionalmente calcula la raíz media cuadrada residual verdadera de la grilla ajustada a los puntos de datos en sus posiciones originales, usando interpolación bilineal en las celdas de la grilla. La información de cauces, tal como esta incorporada en la grilla es también opcionalmente escrita en archivo en un formato conveniente para el ploteo.

### **Especificaciones de los archivos de datos de entrada:**

#### **1.-Datos puntuales de altura:**

Cada archivo contiene triples ordenadas X,Y, Z que denotan la posición y altura de cada punto de observación, con al menos, un dato puntual por registro. Para cada resolución de grilla estos datos puntuales son ubicados en el punto de grilla más cercano. Si se ubica más de un punto en el mismo punto de grilla, se toman los valores promedio. con no más que cuatro puntos de datos considerados para cada punto de grilla. No existe un límite en la entrada de datos puntuales en altura leídos de los archivos de datos suplidos por el usuario.

## 2. -Datos sink

Cada archivo de datos contiene triples ordenadas X,Y,Z que denotan la posición y altura de los puntos sink nominados. Este programa no intenta remover sinks a tales puntos. El punto sink mas alto puede dejarse no especificado, ajustando el valor z, a un valor que caiga fuera de los límites de altura especificados por el usuario.

## 3.- Ríos:

Cada archivo contiene listas de pares de coordenadas x,y, en un orden descendente de altura. Los pares de coordenadas en cada corriente de agua deben ser ordenados desde aguas arriba hacia hasta el punto mas bajo ya que los valores en la grilla están dispuestos en forma descendente para cada cauce. Los datos del cauce tienen prioridad sobre los puntos de altura y los datos sobre curvas de nivel. Así los datos puntuales de altura que entran en conflicto con la forma descendente de los cauces, son ignorados.

## 5. Curvas de nivel

Cada archivo contiene pares de coordenadas x,y precedidos por una línea que indica el número de pares de coordenadas y la cota de las curvas de nivel.

## PROGRAMA ANUSPLIN

### 1 CARACTERISTICAS

El programa ANUSPLIN fue diseñado con la finalidad de ajustar una superficie spline multidimensional con suavizamiento Laplaciano a un conjunto de valores puntuales distribuidos irregularmente en una determinada región geográfica. Cada valor puntual está asociado con un número definido de variables tales como : latitud, longitud y elevación.

ANUSPLIN está formado por cinco programas : Splina, Splinb, Selnot, Lapgrd y Lappnt, cuyo uso dependerá de los datos de entrada y del problema que se desee resolver .

Ellos se aplican de manera secuencial ya que los archivos de salida de uno son parte de los archivos de entrada del siguiente. pero su ejecución es independiente . Las características principales de cada uno de ellos son las siguientes:

a.- Programas Splina y Splinb:

Splina y Splinb son los programas principales de ANUSPLIN con los cuales se ajusta una función de spline laplaciano de  $i$  variables independientes al ruido de un conjunto de datos  $i$  dimensional . La diferencia entre ambos es que el programa Splinb está diseñado para operar en un espacio que está definido en términos de un conjunto de ‘nodos’ seleccionados en la misma región como datos puntuales, permitiendo ajustar hasta 12 superficies en más de 3500 localidades usando hasta 1000 nodos; en el caso de Splina se pueden ajustar hasta 12 superficies pero en no más de 1000 datos puntuales.

Los requerimientos de almacenaje en Splina son proporcionales al cuadrado del número de puntos y el tiempo de procesamiento es proporcional al cubo del número de datos puntuales. En el Splinb se mantienen las mismas proporciones pero con respecto al número de nodos. Es de hacer notar que en el programa Splina no se permite la coincidencia de puntos mientras que en el Splinb si pero no la de nodos, cuya ubicación es inicialmente seleccionada por el programa SELNOT. En el caso de que la ubicación de los nodos coincida con cada dato puntual, las operaciones Splinb serán idénticas a las de Splina.

b.- Programa SELNOT :

El programa tiene como función seleccionar un conjunto inicial de nodos a partir de datos puntuales. Este elige un número específico de datos puntuales como

nodos intentando un equi-muestreo en el espacio de la variable independiente spline. Teóricamente es posible seleccionar nodos que no se encuentren en la posición de los datos puntuales, no obstante esto no es permitido por el programa Splinb .

c.- Programa LAPGRD:

El programa Lapgrd es usado para crear archivos ASCII o archivos imagen (raster) de la superficies generadas por los programas Splina y Splinb.

Si la altura ha sido usada como variable independiente, el programa Lapgrd, necesitará un archivo separado con el modelo digital de terreno generado en por el ANUDEM.

d.- Programa LAPPNT :

El programa Lappnt evalúa los coeficientes de la superficies generadas por el Splina y Splinb, en los puntos que se especifiquen en los archivos de entrada

### 1.3.2 ARCHIVOS DE ENTRADA Y SALIDAS.

Los archivos de entradas y las salidas de cada uno de los programas descritos anteriormente son los siguientes:

a.- Programa Splina y Splinb:

Las entradas a ambos programas son muy similares. El archivo de entrada incluye: el número de variables independientes, el número de covariables independientes, los límites inferior y superior de las variables independientes y las covariables, el orden de la derivada en la función que representa el grado de rugosidad de f, el número de superficies que van ser ajustadas, un estimado de la desviación estándar de los

residuales para cada superficie (si se conoce) y las especificaciones de los archivos de datos y de las de los archivos donde se almacenaran los resultados . En el caso de Splinb, también se requiere el nombre del archivo que contiene el índice de los puntos de nodo .

Los programas chequean los errores en los datos de entrada, eliminando los valores que caen fuera de los límites de las variables independiente dada por el usuario. Esto es una forma muy útil de ajustar una superficie a un sub-conjunto de datos, sin tener que crear un archivo nuevo de datos.

En cuanto a los resultados de los programas Splina y Splinb estos son dispuestos en archivos de salidas que incluyen información estadística esencial de la superficie ajustada incluyendo residuales, estimados de error estandar y dependiendo de los valores asignados por el usuario, se genera una lista de los datos y valores ajustados con errores standard, archivos de salida( xxxxx.sur) que contienen los coeficientes de las superficies ajustadas y los parámetros usados para calcular el parámetro de suavizamiento óptimo. Los coeficientes de superfice son usados para calcular los valores de las superficies ajustadas por otros programas en ANUSPLIN Incluyendo LAPGRD y LAPPNT. La lista de los datos y valores ajustados también es útil para lo mismo, especialmente cuando se ajusta una superficie a un nuevo conjunto de datos.

La validación cruzada (GCV) y el error cuadrático verdadero (MTSE), explicados en el capítulo III forman parte de los estadísticos de salida, y adicionalmente a ellos se incluyen los siguientes:

- Varianza relativa media : es el promedio de las varianzas locales

$$AVR = \frac{\sum d_i}{n}$$

$d_i$ = varianza local;  $n$ = número de puntos de observación

- Raíz de la varianza relativa media

$$RAVR = \sqrt{AVR}$$

- MEDIA: promedio de los valores de precipitación
- STD DEV: desviación estándar en los valores de precipitación.
- MSR : Cuadrado medio residual, es cual es calculado ponderando el cuadrado de cada residual por su varianza local realtiva.

$$MSR = \frac{\sum_{i=1}^n \frac{(z_i - \hat{z}_i)^2}{d_i^2}}{n}$$

- VAR : varianza estimada, definida como la razón entre entre la varianza del error y la varianza relativa media

$$VAR = \frac{\hat{\sigma}_e^2}{AVR}$$

- COVARIABLE :gradiente(s) de la variable(s) independiente considerada en el sub-modelo paramétrico

También se incluyen las raíces cuadradas de : GCV, MSR, MTSE y VAR las cuales expresan a sus estadísticos respectivos en las unidades de los datos; los grados de libertad del error (gle ); los grados de libertad de la señal (gls ) y el paámetro de suavizamiento multiplicado por el número de datos ( RHO)

b.- Programa SELNOT :

Las entradas a selnot son como sigue:

1. -Numero de variables independientes (NDIM)
2. Para  $K= 1$  a NDIM, lee los límites superiores e inferiores de las variables independientes
3. Nombre del archivo de datos.
4. Formato del archivo de datos
5. Nombre del archivo de nodos de salida.
6. Nombre del archivo de nodos rechazados. ( Blanco si no se requiere)
7. Número de nodos requeridos

El archivo de salida de este programa, es una entrada al programa Splinb.

c.- Programa LAPGDR:

Las entradas o directivas del usuario deben ser preparadas en un archivo de entrada que va a ser leído por el programa al igual que de una unidad de entrada standard.

Las entradas son leídas de la manera siguiente:

1. Nombre del archivo de los coeficientes spline
2. Número de superficies que van a ser evaluadas
3. Nombre de la grilla de salida
4. Modelo de la grilla de salida
5. Opción de centraje.
6. Índice de la variable independiente.

7. Límite inferior, límite superior y espaciamiento de las variables independiente
8. Modo de la K ésima variable independiente
9. Constante suplida por el usuario
- 10.El nombre del archivo de grilla
11. Valor especial de la grilla de entrada
- 12.Factor de escala de la grilla de entrada
- 13.Formato de grilla de entrada (blanco para no formateada)

El archivo de salida de este programa, es del tipo imagen y debe ser manejado en IDRISI, para su representación visual

d.- Programa LAPPNT:

Los archivos de entrada de este programa estan dados por los resultados generados en los programas Splina y Splinb. Los valores de todas las variables independientes y las covariables deben ser dados para cada punto y estos pueden ser etiquetados ( por ejemplo código o nombre de la estación).

En el archivo de salida, se disponen las coordenadas de cada estación con su respectivo valor estimado una vez que ha sido ajustada la supercie interpolada.