



UNIVERSIDAD NACIONAL
AUTONOMA DE MEXICO

72
ZET
RECIBIDA EN LA DIVISION DE ESTUDIOS PROFESIONALES
MEXICO D.F. 1995

FACULTAD DE CIENCIAS

APLICACION DE METODOS DE INTERPOLACION BIVARIADOS
Y TRAZO DE CONTORNOS SUAVES PARA EL PRONOSTICO DE
PARAMETROS METEOROLOGICOS EN LA CUARTA REGION
METEOROLOGICA MUNDIAL.

T E S I S
QUE PARA OBTENER EL TITULO DE:
A C T U A R I O
P R E S E N T A:

EDGAR PEREZ PEREZ



MEXICO, D. F. FACULTAD DE CIENCIAS
SECCION ESCOLAR ENERO DE 1995

FALLA DE ORIGEN

TESIS CON
FALLA DE ORIGEN



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

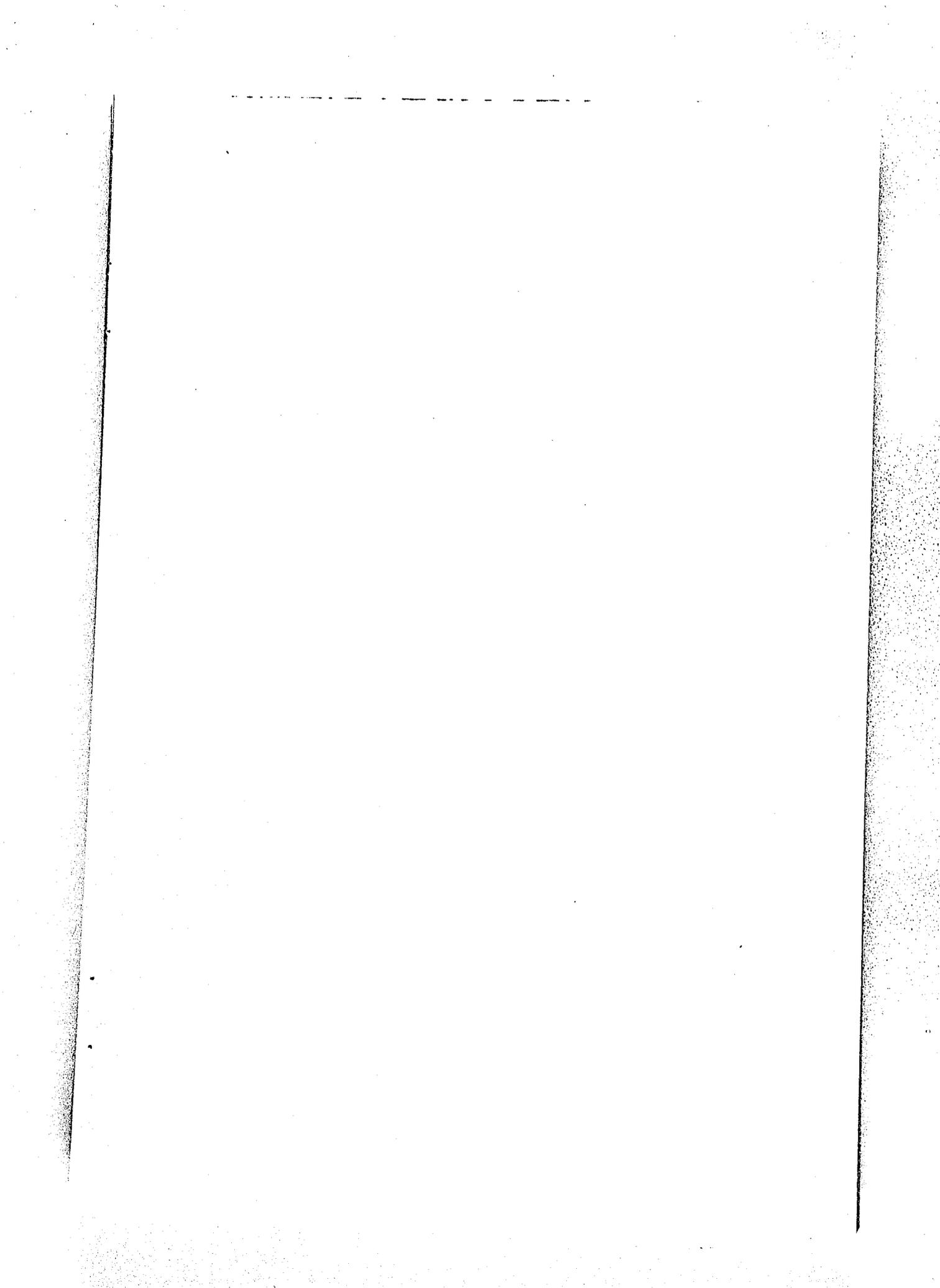
M. EN C. VIRGINIA ABRIN BATULE
Jefe de la División de Estudios Profesionales
Facultad de Ciencias
Presente

Los abajo firmantes, comunicamos a Usted, que habiendo revisado el trabajo de Tesis que realiz(ó)ron el pasante(s) **EDGAR PEREZ PEREZ**

con número de cuenta 8620746-1 con el Título: **APLICACION DE METODOS DE INTERPOLACION BIVARIADOS Y TRAZO DE CONTORNOS SUAVES PARA EL PRONOSTICO DE PARAMETROS METEOROLOGICOS EN LA CUARTA REGION METEOROLOGICA MUNDIAL.**

Otorgamos nuestro Voto Aprobatorio y consideramos que a la brevedad deberá presentar su Examen Profesional para obtener el título de **ACTUARIO**

GRADO	NOMBRE(S)	APELLIDOS COMPLETOS	FIRMA
DR.	PABLO	BARRERA SANCHEZ	
Director de Tesis LIC.	FRANCO	TOLEDO DE LA CRUZ	
M. en C.	MA. ELENA	GARCIA ALVAREZ	
M. en C.	MARIANO	LOZANO MARTINEZ	
Suplente M. en C.	JOSE	GUERRERO GRAJEDA	
Suplente			



INDICE

	PAGINA
I. MOTIVACIÓN	5
II. INTRODUCCIÓN	10
III. MÉTODO DE INVERSO DE LAS DISTANCIAS	14
IV. MÉTODO KRIGING	35
V. APLICACIÓN DE LOS MÉTODOS DESCRITOS	57
VI. ANÁLISIS DE LOS RESULTADOS OBTENIDOS	74
VII. DISCUSIÓN DE LA EFICIENCIA	77
VIII. CONCLUSION	88
IX. APENDICE (PROGRAMAS)	89
XI. BIBLIOGRAFIA.	134

MÉTODOS DE INTERPOLACION BIVARIADOS

MOTIVACION

Es en una gran diversidad de ramas de las ciencias naturales y sociales (meteorología, ecología, arqueología, geología, minerología, economía, demografía) en las que el problema de conocer información "puntual" o "zonal", sin tener específicamente una fuente generadora de datos en el área de interés, toma gran relevancia. Los medios de que se dispone, generalmente, en este tipo de problemas están limitados a proporcionar datos alrededor, cerca o lejos, del punto de interés y naturalmente, lo que se pretende es utilizar la información disponible en esta vecindad, tomando límites razonables para de alguna forma "adivinar", pronosticar o inferir la característica que nos interesa.

Los fenómenos meteorológicos, evidentemente, no pueden ser la excepción, todo lo contrario pues su estudio a un nivel generalmente de mesoescala se ajusta precisamente a las características antes descritas, es decir, estamos interesados, teóricamente, en conocer el estado de los parámetros meteorológicos (temperatura, presión, humedad, precipitación, etc.) En cualquier punto de, por ejemplo, la república mexicana y sin embargo no contamos con los datos pertinentes en cada uno de esos sitios, no obstante, sí disponemos de información en algunos "nodos" de la región: la proporcionada por la red de observatorios del Servicio Meteorológico Nacional (SMN) (quizás posteriormente, una vez que los datos que emiten sean lo suficientemente confiables, pudiera incluirse también la red de estaciones automáticas) y es de esta forma que a partir del comportamiento conocido de las variables meteorológicas en las inmediaciones del lugar de estudio se pretende clarificar el comportamiento desconocido que ahí se tiene.

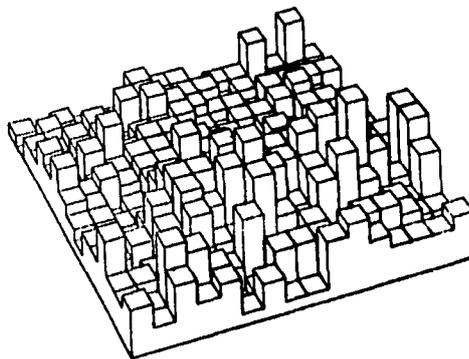
Este tipo de técnicas, matemáticamente, recibe el nombre de *interpolación*, y si bien su fundamentación teórica para funciones reales de variable real (interpolación con polinomios) es muy vieja, su generalización para funciones con dominio en \mathbb{R}^2 e imágenes en \mathbb{R} (*superficies*) y aplicación para inferir el comportamiento de fenómenos naturales, es relativamente nueva (aproximadamente en 1960 se comenzaron a publicar los primeros trabajos financiados por organizaciones militares).

Genéricamente, salvo algunas variaciones, este tipo de problemas han recibido el nombre de métodos de interpolación bivariados (multivariados) y ajuste de superficies suaves ($f: \mathbb{R}^2 \rightarrow \mathbb{R}$, tal que $f \in C^1$) para conjuntos de datos distribuidos irregularmente.

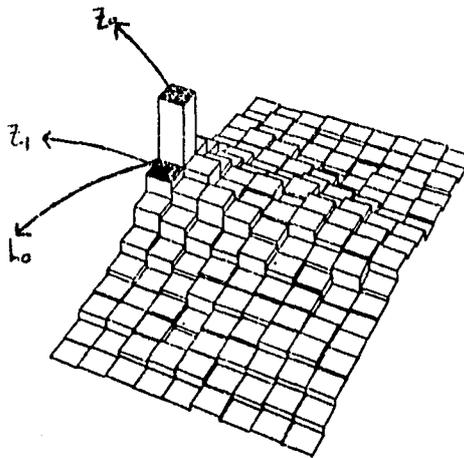
Tratando de justificar este nombre podemos decir lo siguiente:

Interpolación bivariada: sabemos, en términos básicos, que la interpolación consiste en *ajustarle* a un conjunto de datos una función que los caracterice o aproxime de la mejor manera posible de tal forma que pueda estimarse o predecirse, también, el valor de puntos "internos" o "externos" del conjunto original. Para la interpolación bivariada, la idea es exactamente la misma, sólo que en este caso la función que ajusta será una superficie.

Ajuste de superficies suaves: porque en realidad eso es lo que estamos haciendo. Recordemos que una superficie es una función con dominio en \mathbb{R}^2 e imágenes en \mathbb{R} y que el atributo de suave se les confiere cuando son funciones de clase C^1 , es decir, además de continuas derivables y tal que la derivada siga siendo una función continua. Esto es importante e indispensable, puesto que se pretende que el método de interpolación sea lo más exacto posible. Claramente una superficie discontinua o no suave (con muchos "picos") no es una buena herramienta de aproximación.



**EJEMPLO DE SUPERFICIE DISCONTINUA Y NO SUAVE
(MAL INTERPOLADOR)**



LIMITACIONES DE LAS SUPERFICIES DISCONTINUAS Y NO SUAVES

En las figuras anteriores observamos que se carece tanto de continuidad como de suavidad y por lo tanto no son superficies que nos den aproximaciones exactas en un problema de interpolación. Obsérvese, por ejemplo, (fig. B) que si la gráfica correspondiera a una superficie isotérmica (obtenida por algún método de estimación o aproximación) existiría una enorme incongruencia entre las aproximaciones referentes al valor máximo y a las que se encuentran alrededor de éste.

Supongamos que la región rectangular (R_0) alrededor del punto máximo (Z_0) tuviera un valor de 30°C , y la correspondiente (R_1) a un punto circundante (Z_1) un valor de 15°C ; ¿Qué pasa si nos preguntamos por el valor que dicha superficie interpolante le asigna al segmento límite (L_0) entre R_0 y R_1 ?

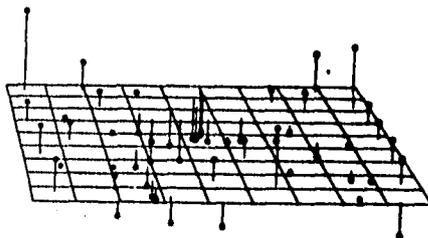
Resultaría, por un lado, que si nos aproximamos a L_0 por la región R_0 , obtendríamos 30°C , mientras que si lo hacemos por la región R_1 encontraríamos un valor de 15°C , es decir, dos magnitudes totalmente diferentes para el mismo segmento. De ahí la necesidad de la continuidad y suavidad de la superficie.

Conjunto de datos distribuidos irregularmente: en realidad el caso de una distribución irregular de los datos es la generalización para cualquier tipo de ubicación que pudieran tener los "nodos" generadores de información (observatorios en nuestro caso).

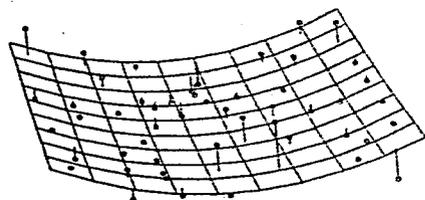
Teóricamente los datos conocidos (información emitida por los nodos) se encuentran distribuidos espacialmente con respecto al plano XY (o con respecto a la longitud y latitud es un sistema de coordenadas geográficas) con una altura o imagen en función de su localización en el plano base (en el caso de la temperatura, por ejemplo, dicha altura será igual al valor que se registre en el observatorio: 15 °c, 20 °c, 22 °c, 24 °c, etc.) Sin embargo, en términos prácticos no se trabaja tanto con las alturas o imágenes como con sus proyecciones sobre el plano XY o plano longitud-latitud.

Similarmente, no se trabaja tanto con la superficie interpolante como con las proyecciones de sus curvas de nivel en el plano base.

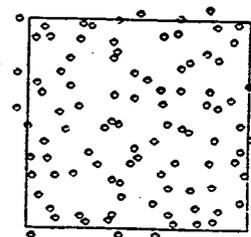
Así pues, diremos que una isolinia, o línea que une puntos de igual valor, es la proyección sobre el plano base de las curvas de nivel de una superficie, es decir :
ISOLINEA = $\{(X,Y) \in R^2 : Z(X,Y) = CTE.\}$



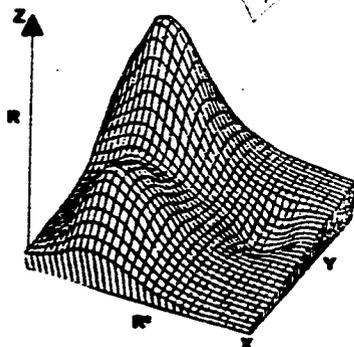
TEORICAMENTE LOS DATOS A INTERPOLAR SE ENCUENTRAN DISTRIBUIDOS ESPACIALMENTE CON RELACION A UN PLANO BASE. PRACTICAMENTE SE TRABAJA CON SUS PROYECCIONES SOBRE EL.



PROBLEMA ORIGINAL: DATOS A INTERPOLAR
DISTRIBUIDOS IRREGULARMENTE EN EL
ESPACIO

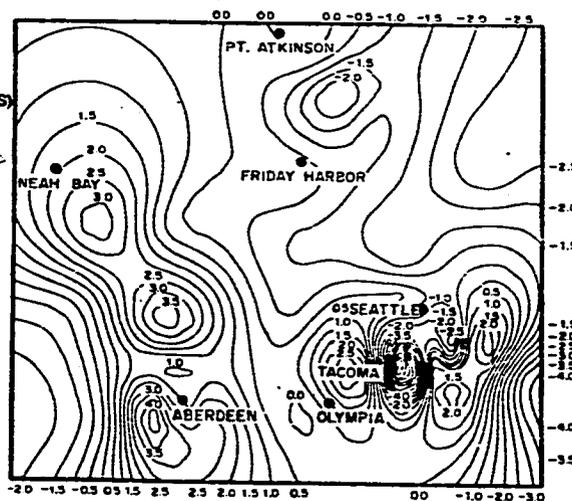


PROYECCION DE LOS VALORES
SOBRE EL PLANO BASE



SUPERFICIE INTERPOLANTE ($F: R^2 \rightarrow R$)
SUAVE GENERADA POR ALGUN METODO
DE INTERPOLACION

CURVAS DE NIVEL
DE LA SUPERFICIE
INTERPOLANTE PRO-
YECTADAS EN EL
PLANO BASE (ISOLINEAS)
SOLUCION DEL
PROBLEMA



INTRODUCCION

El problema de estimar, aproximar o inferir el comportamiento global de algún fenómeno mediante el conocimiento escueto de éste, es central en algunas ramas del quehacer humano (teórico y práctico). Tradicionalmente (en el caso de una sola variable dependiente y una explicatoria) se ha utilizado, por ejemplo, el análisis de regresión lineal, el no lineal y el ajuste con Splines; sin embargo, los problemas en la naturaleza o los que se crea el hombre no son tan simples como para modelarlos e intentar resolverlos utilizando únicamente un par de variables; es indispensable hacer un análisis multivariado. La regresión múltiple intenta aproximar en la medida de sus limitaciones soluciones óptimas para estos problemas; la exactitud, sin embargo, es fundamental, y, desafortunadamente, no puede ser proporcionada por un simple plano o un paraboloides, por ejemplo, es indispensable una función más versátil, más complaciente con los "caprichos" del fenómeno en estudio.

Así pues, formalmente, lo que se pretende es la construcción de funciones biviadas suaves (al menos con sus primeras derivadas parciales continuas) capaces de resolver "optimamente" el problema.

Actualmente, se dispone de un gran número de métodos generados desde muy diversos puntos de vista, los cuales pretenden solucionar el problema empleando herramientas matemáticas tan variadas como la formación de sus autores. En términos generales, sin embargo, podemos clasificar todas esas alternativas, tomando como referencia sus características de operación, fundamentación teórica e interdependencia de sus variables conocidas, a través de dos etapas; la primera etapa los divide en globales y locales:

METODOS GLOBALES Y METODOS LOCALES:

Por sus características de operación y fundamentación teórica, los métodos de interpolación pueden dividirse en dos grandes categorías: *métodos globales* y *métodos locales*. El concepto de método global es fácilmente comprensible: la función interpolante es dependiente de todos los datos puntuales de tal forma que la adición o supresión de uno de ellos o el cambio en una de sus coordenadas repercutirá a lo largo de todo el dominio de la función. En los métodos locales, la adición o supresión de un punto o el cambio en una de sus coordenadas afectará a la función interpolante sólo en los puntos cercanos, es decir, ésta permanecerá invariable en los puntos más retirados.

La segunda etapa, independientemente de su carácter de locales o globales, los clasifica en tres grandes grupos:

1.METODOS INTUITIVOS

2.METODOS ESTADISTICOS

3.METODOS POLINOMIALES

METODOS INTUITIVOS: Son los primeros que se comenzaron a utilizar, matemáticamente son sencillos y de muy fácil asimilación.

Sus representantes más conocidos son, por un lado, los denominados *métodos de inverso de las distancias ponderadas* y por el otro, *métodos basados en la teoría del elemento finito*. Un desglose más detallado de ellos incluye mínimamente a los siguientes:

- Inverso de las distancias
- Métodos basados en triangulaciones
- Métodos basados en cuadrículas o rejillas

METODOS ESTADISTICOS: De las tres clasificaciones ésta es la más reciente, su desarrollo presenta una historia larga e interesante en la teoría de la probabilidad. Su fundamentación teórica descansa, naturalmente, en la probabilidad y estadística clásicas. Su representante más notable es el método conocido como *Kriging*.

Kriging es un *método lineal insesgado de varianza mínima*, es decir un *estimador óptimo*, que aproxima el valor de una *variable aleatoria* en algún punto a partir de los valores disponibles a su alrededor; proporciona por lo tanto un mapa de isolíneas tan exacto como es posible.

En el área de Previsión del SMN se utilizan para el diseño de algunos productos (Cartas de Superficie y Altura) programas comerciales como SURFER el cual trabaja con dos de los métodos descritos anteriormente: Inverso de las distancias y Kriging.

METODOS POLINOMIALES: Sus antecesores directos son los típicos métodos de interpolación con *polinomios* y *Splines* para funciones con dominio y contradominio en los números reales, por ello se les considera como los algoritmos con el desarrollo teórico más antiguo.

Un posible desglose de ellos nos conduciría a lo siguiente:

Métodos que utilizan Splines

Métodos que utilizan Superficies Cuádricas

Ambas alternativas utilizan frecuentemente algunos resultados del álgebra lineal.

CARACTERISTICAS DESEABLES

Los 'atributos' mínimos deseables para cualquier método de interpolación son:

EXACTITUD: Dado que el objetivo fundamental es el "pronóstico" en base a los datos conocidos, es indispensable que el método sea capaz de producir superficies suaves que redunden finalmente en isollneas representativas y exáctas.

TIEMPO DE EJECUCION (COMPUTACIONAL): Algunos métodos son bastante eficientes en términos del tiempo requerido para efectuar los cálculos pero desafortunadamente son deficientes en otros aspectos; recíprocamente, hay métodos "lentos" con muy buenas otras características.

REQUERIMIENTOS DE ALMACENAJE (COMPUTACIONAL): Algunos autores consideran que este no es un aspecto crucial a menos que llegue a "grandes dimensiones". Los requerimientos de memoria se cuantifican sólo en términos de los arreglos adicionales necesarios para almacenar los valores entre los puntos (x_k, y_k, f_k) y no tanto por el número de 'variables simples' o la longitud del programa.

FACILIDAD DE IMPLEMENTACION (COMPUTACIONAL): Esta característica está muy vinculada con la naturaleza del algoritmo que se pretenda desarrollar ; como veremos hay algoritmos "matemáticamente muy pesados" con los que independientemente de la *habilidad del programador y de lo claro que pueda tener el problema* resultará un programa rebuscado y extenso en código .

Las cotas que nos hemos impuesto para el contenido del presente trabajo, nos permitirán:

Desarrollar dos métodos de acuerdo a la segunda clasificación mencionada, es decir, inverso de las distancias ponderadas y Kriging y hacer algunas comparaciones entre cada una de estas alternativas, mencionando sus bondades y limitaciones más severas así como los posibles vínculos teóricos que pudieran existir entre ellas (Splines es un caso particular de Kriging).

Proponer varios programas de cómputo:

Un programa decodificador de mensajes sinópticos meteorológicos que proporcionará las materias primas (datos) para la consecución del objetivo final.

Un programa que traslada la información meteorológica decodificada a un ambiente gráfico "real" (construcción de mapas y ubicación de los parámetros decodificados sobre ellos).

Un programa que a partir de los datos meteorológicos decodificados (76 a lo más) estimará el valor de 20X33 valores más correspondientes a las coordenadas geográficas (de 5 en 5 grados de latitud y longitud) de la república mexicana.

Utilizar, para la generación de los mapas de isolinias, un programa comercial auxiliar (*SURFER*), el cual podrá emplearse como un simple "trazador" dado un archivo de datos interpolados (archivo .GRD), o como una herramienta que resuelva el problema en paralelo dado el archivo original de datos meteorológicos decodificados. Y finalmente:

Aplicar todo lo anterior a la solución de un problema concreto: Explotar eficientemente la información meteorológica que se recibe de la red de observatorios de México este aprovechamiento óptimo de información abarca desde su decodificación, su "ploteo" en mapas, su interpolación y trazo de isolinias (todo en forma automática), hasta la generación de un pronóstico en tiempo real.

METODO DE INVERSO DE LAS DISTANCIAS PONDERADAS POR MINIMOS CUADRADOS.

ENTORNO

Este método utiliza y desarrolla al máximo la idea intuitiva de que del conjunto de puntos conocidos influyen de manera más representativa sobre un valor intermedio, desconocido que se pretenda estimar aquellos que se encuentran más cerca del punto en cuestión que los que se ubiquen más retirados de él.

Esa influencia en función de la distancia tiene que ser cuantificada de alguna forma y que mejor si esa forma es la "óptima", en el sentido de que tanto a los puntos cercanos como a los lejanos se les asigna un peso mayor o menor pero el mejor de entre todos los posibles.

PLANTEAMIENTO

Para este método es conveniente considerar que los datos de que se dispone provienen de alguna función "subyacente" f y visualizar el proceso de interpolación como un operador aplicado a esta función. Es decir, $f_i = f(x_i, y_i) : i = 1, 2, \dots, n$ es la función subyacente (no olvidemos que se parte del "supuesto" de que se cuenta con una muestra de n puntos distribuidos espacialmente) y $S(x, y) = P[f](x, y)$ es el operador.

Definamos $\rho_i(x, y)$ como una función con ciertas propiedades de una función distancia referida a un punto conocido (x_i, y_i) . En particular, asumiremos que $\rho_i(x_i, y_i) = 0$, es decir la distancia de un punto a él mismo es cero, y que $1/\rho_i(x, y)$ es una función decreciente no negativa cuando (x, y) se aleja de (x_i, y_i) , por ejemplo la distancia Euclidiana usual

$$d_i(x, y) = \sqrt{(x - x_i)^2 + (y - y_i)^2}$$

es una posibilidad para $\rho_i(x, y)$.

Denotaremos por $\Phi_j : j = 1, \dots, m$ al conjunto de funciones base que serán usadas en la aproximación por mínimos cuadrados.

De esta forma, la expresión general de la función interpolante será

$$P[f](x, y) = \sum_{j=1}^m \bar{a}_j(x, y) \phi_j(x, y) \quad (1.1)$$

Donde

$$\bar{a}_j(x,y) : j=1,\dots,m$$

representa la solución de

$$\text{Min}_{a_1,\dots,a_m} \sum_{i=1}^N \left[\frac{a_1\phi_1(x_i,y_i) + \dots + a_m\phi_m(x_i,y_i) - f_i}{\rho_i(x,y)} \right]^2 \quad (1.2)$$

para un (x,y) dado.

CASO TIPICO LINEAL (Regresión Lineal Simple)

Recordemos, con el propósito de lograr una mejor comprensión de lo anterior, las ideas básicas del problema de mínimos cuadrados aplicado al análisis de regresión:

Supongamos que como función interpolante queremos utilizar una recta

$$\rightarrow P[f](x,y) = ax + b$$

o bien

$$\bar{\beta}_0 x^0 + \bar{\beta}_1 x = \sum_{i=0}^1 \bar{\beta}_i(x,y) x^i = \sum_{i=0}^1 \bar{\beta}_i(x,y) \phi_i(x,y)$$

con

$$\phi_0(x,y) = 1 \quad \text{y} \quad \phi_1(x,y) = x$$

Donde las funciones beta i's son los parámetros que determinan en forma única la recta (ordenada al origen y pendiente).

Lo que se pretende es, entonces, minimizar la suma de los residuales al cuadrado (diferencia entre los valores observados y los estimados) para de esta forma tener la certeza de estar construyendo la recta óptima que ajusta al conjunto de puntos observados, es decir:

$$\frac{\text{Min}}{\beta_0, \beta_1} \sum_{i=1}^n e_i^2 = \frac{\text{Min}}{\beta_0, \beta_1} \sum_{i=1}^n (v_i - y_i)^2$$

donde los puntos estimados están dados por la ecuación

$$y_i^* = \bar{\beta}_0 + \bar{\beta}_1 x$$

$$\Rightarrow \frac{\text{Min}}{\beta_0, \beta_1} \sum_{i=1}^n (v_i^* - y_i)^2 = \frac{\text{Min}}{\beta_0, \beta_1} \sum_{i=1}^n (\bar{\beta}_0 x_i^0 + \bar{\beta}_1 x_i^1 - y_i)^2 \quad (1.3)$$

$$= \frac{\text{Min}}{\beta_0, \beta_1} \sum_{i=1}^n (\bar{\beta}_0 x_i^0 + \bar{\beta}_1 x_i^1 - f_i)^2 = \frac{\text{Min}}{\beta_0, \beta_1} \sum_{i=1}^n \frac{(\bar{\beta}_0 \phi_0(x_i, y_i) + \bar{\beta}_1 \phi_1(x_i, y_i) - f_i)^2}{1}$$

La necesidad de que la función distancia sea, en este caso, constante e igual a 1 queda de manifiesto más claramente si consideramos por un lado la solución de la ecuación (1.3) (parámetros de regresión óptimos) y por el otro la solución de la ecuación (1.2) con i variado de 0 a 1 y m igual a 1

Según sabemos, para el primer caso obtenemos:

$$\bar{\beta}_0 = \bar{y} - \bar{\beta}_1 \bar{x}$$

$$\bar{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i - (\sum x_i)^2}$$

Mientras que para el segundo

REDUCCION AL CASO LINEAL DEL METODO EN ESTUDIO

Problema

$$\text{Min}_{a_0, a_1} \sum_{i=1}^n \left[\frac{a_0 + a_1 x_i - f_i}{\rho_i(x, y)} \right]^2$$

→ Sea

$$K = \sum_{i=1}^n \left[\frac{a_0 + a_1 x_i - f_i}{\rho_i(x, y)} \right]^2$$

→

$$\frac{\partial K}{\partial a_0} = 2 \sum_{i=1}^n \left[\frac{a_0 + a_1 x_i - f_i}{\rho_i(x, y)} \right] * \frac{1}{\rho_i(x, y)}$$

→

$$\frac{\partial K}{\partial a_1} = 2 \sum_{i=1}^n \left[\frac{a_0 + a_1 x_i - f_i}{\rho_i(x, y)} \right] * \frac{x_i}{\rho_i(x, y)}$$

$$\rightarrow \frac{\partial K}{\partial a_0} = 0 \quad \rightarrow \quad \sum_{i=1}^n \left[\frac{a_0 + a_1 x_i - f_i}{\rho_i^2(x,y)} \right] = 0$$

$$\rightarrow \quad a_0 \sum_{i=1}^n \frac{1}{\rho_i^2} + a_1 \sum_{i=1}^n \frac{x_i}{\rho_i^2} - \sum_{i=1}^n \frac{f_i}{\rho_i^2} = 0$$

$$\rightarrow \quad a_0 = \frac{\sum_{i=1}^n \frac{f_i}{\rho_i^2} - a_1 \sum_{i=1}^n \frac{x_i}{\rho_i^2}}{\sum_{i=1}^n \frac{1}{\rho_i^2}} \quad (1.3.1)$$

A continuacion podemos observar que si en

$$\sum_{i=1}^n \frac{f_i}{\rho_i^2} \quad y \quad \sum_{i=1}^n \frac{x_i}{\rho_i^2} \quad , \quad \rho_i^2 \quad \text{fuera siempre } 1$$

$$\rightarrow \quad \sum_{i=1}^n \frac{f_i}{\rho_i^2} = \sum_{i=1}^n f_i \quad y \quad \sum_{i=1}^n \frac{x_i}{\rho_i^2} = \sum_{i=1}^n x_i$$

y además

$$\sum_{i=1}^n \frac{1}{\rho_i^2} = \sum_{i=1}^n \frac{1}{1} = \sum_{i=1}^n 1 = n$$

Con lo que (1.3.1) quedaría como

$$a_0 = \frac{\sum_{i=1}^n f_i - a_1 \sum_{i=1}^n x_i}{n} = \bar{y} - a_1 \bar{x}$$

$$\therefore \bar{\beta}_0 = a_0$$

Con respecto al otro parámetro observamos que

$$\rightarrow \frac{\partial K}{\partial a_1} = 0 \quad \Leftrightarrow \quad \sum_{i=1}^n \left[\frac{a_0 x_i + a_1 x_i^2 - f_i x_i}{\rho_i^2(x, y)} \right] = 0$$

$$\Leftrightarrow \quad a_0 \sum_{i=1}^n \frac{x_i}{\rho_i^2} + a_1 \sum_{i=1}^n \frac{x_i^2}{\rho_i^2} - \sum_{i=1}^n \frac{f_i x_i}{\rho_i^2} = 0$$

$$\Leftrightarrow \quad (\bar{y} - a_1 \bar{x}) \sum_{i=1}^n \frac{x_i}{\rho_i^2} + a_1 \sum_{i=1}^n \frac{x_i^2}{\rho_i^2} - \sum_{i=1}^n \frac{y_i x_i}{\rho_i^2} = 0$$

$$\Leftrightarrow \quad a_1 \left(\sum_{i=1}^n \frac{x_i^2}{\rho_i^2} - \frac{n \bar{x}^2}{\rho_i^2} \right) = \sum_{i=1}^n \frac{y_i x_i}{\rho_i^2} - \bar{y} \sum_{i=1}^n \frac{x_i}{\rho_i^2}$$

Si consideramos nuevamente a ρ_i^2 como 0 obtendremos:

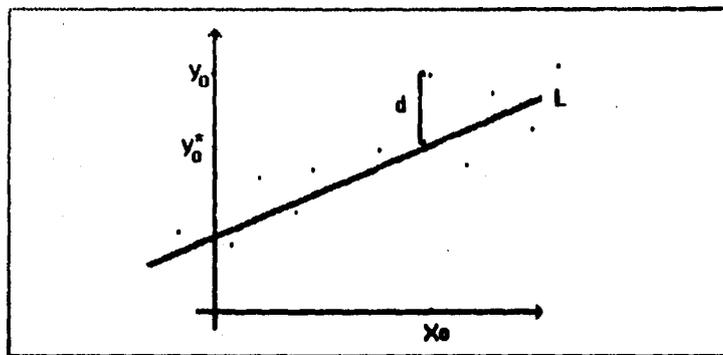
$$a_1 = \frac{\sum_{i=1}^n \frac{y x_i}{\rho_i^2} - y \sum_{i=1}^n x_i}{\left(\sum_{i=1}^n \frac{x_i^2}{\rho_i^2} - \frac{n \bar{x}^2}{\rho_i^2} \right)}$$

$$a_1 = \frac{\sum_{i=1}^n y x_i - y \sum_{i=1}^n x_i}{\sum_{i=1}^n x_i^2 - n \bar{x}^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

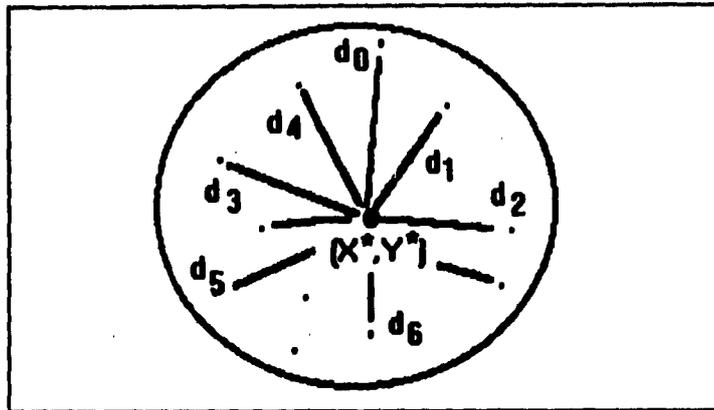
y $\therefore \bar{\beta}_1 = a_1$

INTERPRETACION GEOMETRICA

La distancia entre dos puntos, el estimado y el observado referida en unidades de ella misma siempre es igual a la unidad. En este caso, sobre el valor "intermedio, desconocido" que queremos estimar (y_0) solamente influirá el valor conocido y_0 .



Cuando la información disponible está constituida por más de un punto, digamos K valores en una vecindad de radio R de un punto (x^*, y^*) intermedio que queremos estimar, la distancia a cada uno de ellos no tiene por qué ser constante, es variable, es una función que depende de su localización con respecto al punto en cuestión.



Así pues, el enfoque sencillo y tradicional de los mínimos cuadrados aplicados al análisis de regresión lineal simple puede considerarse como un caso particular del método que estamos describiendo y esto a su vez tiene la ventaja de permitirnos "fijar ideas" y avanzar paralelamente con un modelo de fácil comprensión.

APLICACION DE LAS ECUACIONES NORMALES (Caso Lineal Típico)

Consideremos ahora las *ecuaciones normales* en el modelo de regresión lineal simple (1.3), obtenidas de la derivación de la función interpolante con respecto a sus parámetros, las cuales están dadas por lo siguiente

$$\begin{aligned} \sum y_i &= n\bar{\beta}_0 + \bar{\beta}_1 \sum x_i \\ \sum y_i x_i &= \bar{\beta}_0 \sum x_i + \bar{\beta}_1 \sum x_i^2 \end{aligned}$$

entonces el determinante de la matriz de coeficientes será

$$\begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{vmatrix} = n \sum x_i^2 - (\sum x_i)^2$$

Por otro lado, para poder hablar de la derivada de la función interpolante, dicha función derivada debe de existir, sin embargo en este caso encontramos que

$$P[f](x,y) = ax+b = \bar{\beta}_0 + \bar{\beta}_1 x$$

$$= \bar{y} - \bar{\beta}_1 \bar{x} + \bar{\beta}_1 x = \bar{y} + \bar{\beta}_1 (x - \bar{x}) = \bar{y} + \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} (x - \bar{x})$$

$$= \bar{y} + \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i - (\sum x_i)^2} (x - \bar{x})$$

Donde el denominador del segundo sumado es igual al determinante de la matriz de coeficientes de las ecuaciones normales. Por lo tanto, observamos que la existencia de las derivadas del interpolador está sujeta a la no singularidad de dicha matriz.

APORTACIONES Y SELECCION DE HERRAMIENTAS

McLain [7] aportó resultados de un número de pruebas realizadas para la elección de las funciones base, Φ_i , y funciones de distancia ρ_i . Sugirió que los Φ_i consistan de los monomios de "bajo orden" $x^j y^k$, y $\rho_i(x,y)$ sea tomada como

$$d_i, d_i^2, \text{ o } d_i e^{-ad_i^2}$$

La potencia más elevada para d_i y la exponencial se deben al deseo de darle menos peso a los puntos lejanos del que les asigna d_i por sí sola.

Notemos que la función (1.1) no estará definida en cualquier punto (x, y) si se elige alguna de las funciones anteriormente propuestas, puesto que el peso para la mejor aproximación por mínimos cuadrados

$$\frac{1}{\rho_i^2(x, y)} \rightarrow \infty$$

cuando $(x, y) \rightarrow (x_i, y_i)$, y por tanto es indispensable pedir que

$$p[f](x, y) \rightarrow f_i$$

cuando $(x, y) \rightarrow (x_i, y_i)$, pues de lo contrario la suma de los errores al cuadrado no estaría acotada. Además obtendremos una aproximación continua si definimos

$$p[f](x_i, y_i) \rightarrow f_i \quad ; \quad i=1, \dots, N$$

es decir, el interpolante evaluado en los n puntos disponibles debe de "regresarnos" la imagen conocida.

Mclain demostró que la función interpolante es infinitamente diferenciable lo cual es fácilmente comprobable, según hemos visto, bajo la suposición de no singularidad de la matriz de coeficientes del conjunto de ecuaciones normales obtenidas de (1.2)

En particular, Mclain propuso, retomando la idea de que las funciones base deben de ser igual a $x^i y^j$ y dadas las alternativas para la función distancia (de ponderación) la siguiente función cuadrática bivariable

$$P[f](x,y) = \bar{a}_1 + \bar{a}_2x + \bar{a}_3y + \bar{a}_4x^2 + \bar{a}_5xy + \bar{a}_6y^2$$

y la función de pesos

$$\rho_i = d_i e^{-ad_i^2}$$

y encontró que trabajaban bien para un gran número de casos.

Sin embargo, hasta este nivel y con las herramientas desarrolladas, este método, tiene 2 grandes desventajas:

1. Se requiere un gran esfuerzo computacional puesto que para cada evaluación se necesita resolver un problema de mínimos cuadrados.

2. El método es global, es decir, la función interpolante depende de todos los datos puntuales esparcidos sin importar que tan lejos estén del punto donde la evaluación se está llevando a cabo.

Es indispensable, por lo tanto, de alguna forma introducir modificaciones que redunden finalmente en un método eficiente y exacto

SUSTITUCION DE LOS PARAMETROS OPTIMOS

Modificaciones del método anterior (inverso de las distancias ponderadas por mínimos cuadrados) han reemplazado los "parámetros óptimos"

$$\bar{a}_j(x,y)$$

que dependen de toda la información disponible y que tienen que calcularse (con su correspondiente problema de mínimos cuadrados) para cada punto (x,y) del dominio de interés, con la aproximación

$$A[\bar{a}_j](x,y)$$

De tal forma que se cumpla que

$$A[\bar{a}_j](x_i, y_i) = \bar{a}_j(x_i, y_i) : i=1, \dots, N$$

Dicha sustitución tiene por objeto utilizar herramientas que no alteren los atributos que hasta ese momento se habían conseguido y que además sean computacionalmente más eficientes pues aunque los

$$\bar{a}_j(x, y)$$

son estimadores óptimos, tienen por otro lado la desventaja de proporcionar información puntual (son *estimadores puntuales*) es decir, para el punto para el cual fue resuelto el problema de mínimos cuadrados y por tanto, habría que resolver tantos problemas de este tipo como puntos quisieramos estimar.

La aproximación utilizada queda definida de la siguiente forma:

$$A[\bar{a}_j](x, y) = \sum_{i=1}^N \bar{a}_j(x_i, y_i) W_i(x, y) : j=1, \dots, m$$

Tal que $W_i(x_j, y_j) = \delta_{ij} : i, j=1, \dots, N$ (1.4)

(delta de Kronecker, definida como 1 si $i=j$ y como 0 si $i \neq j$)

Ahora observemos lo siguiente

$$A[\bar{a}_j](x_i, y_i) = \sum_{i=1}^N \bar{a}_j(x_i, y_i) W_i(x_i, y_i) = \bar{a}_j(x_i, y_i) \cdot 1 = \bar{a}_j(x_i, y_i)$$

*Con esta expresión encontramos que la aproximación $A[\bar{a}_j](x, y)$ se mantiene igual a $\bar{a}_j(x, y)$ en cualquiera de los n puntos conocidos y para cualquiera de los parámetros, y hasta ahí no existe ningún avance pues para cada punto conocido se tiene que calcular un problema de mínimos cuadrados, pero que pasa en algún punto intermedio (x, y) , ¿cómo aprovechamos el hecho de ya conocer nm parámetros óptimos para calcularle los que a él le corresponden? : Observemos para su primer parámetro

$$A[\bar{a}_1](x,y) = \bar{a}_1(x_1,y_1)W_1(x,y) + \bar{a}_1(x_2,y_2)W_2(x,y) + \bar{a}_1(x_3,y_3)W_3(x,y) + \dots + \bar{a}_1(x_n,y_n)W_n(x,y)$$

es decir, una "combinación lineal ponderada" del primero de los m parámetros en cada uno de los n puntos conocidos. Nótese que aquí W_i no es ni cero ni uno sino una cantidad que debe de estar en función de la proximidad del punto estimado (x,y) a cada uno de los (x_i,y_i) conocidos.

Con la aproximación anterior para $a_j(x,y)$, es claro, se logra una enorme reducción en el número de operaciones que se tienen que efectuar para estimar un punto intermedio, pues se sustituye el problema de mínimos cuadrados para estimar los m parámetros óptimos (sacar derivadas parciales, plantear un sistema de ecuaciones normalizadas, etc.) por un conjunto de m sumas de productos de dos funciones, una de las cuales (las a_j) siempre aparecerá con un comportamiento muy regular sin importar el punto (x,y) intermedio que se esté estimando, la única función que variará considerablemente es W_i y de ella, en realidad, dependerá la bondad (precisión, exactitud) de las aproximaciones $A[a_j](x,y)$.

*Nótese que con el enfoque anterior, sólo hay que resolver n problemas de mínimos cuadrados: los correspondientes a los n puntos conocidos.

Por lo tanto es muy importante hacer una adecuada elección de las funciones W_i , las cuales, según podemos ver intuitivamente dependerán o quedarán expresadas en términos de las distancias del punto intermedio estimado a cada uno de los n valores conocidos.

Así pues, el operador

$$Q[f](x,y) = \sum_{j=1}^m A[\bar{a}_j](x,y)\phi_j(x,y)$$

mantendrá las propiedades interpolatorias de $P[f]$.

Por lo tanto podemos reescribir la expresión anterior para Q como

$$Q(f)(x,y) = \sum_{i=1}^N W_i(x,y) Q_i(x,y)$$

Donde

$$Q_i(x,y) = \sum_{j=1}^m \bar{a}_j(x_i,y_i) \phi_j(x,y)$$

es la estimación por inverso de la distancia ponderada por mínimos cuadrados en el punto (x,y)

Es decir, el interpolante $Q(f)(x,y)$ queda expresado como una combinación lineal de funciones nodales ponderadas con una función que depende de la proximidad del punto estimado a cada uno de los puntos conocidos.

Nos referimos a las funciones $Q_i, i=1,\dots,N$ como las funciones nodales puesto que ellas están asociadas con los nodos $(x_i,y_i) : i=1,\dots,N$ respectivamente. Notemos que $Q_i(x_i,y_i) = f_i, i=1,\dots,N$ y que $Q_i(x,y)$ es una aproximación local (cerca de (x,y)) a $f(x,y)$ y como tal debe esperarse que "emule" el comportamiento y forma de $f(x,y)$ a menos que los puntos distantes le afectaran en forma considerable.

SELECCION DE LOS W_i

Con el propósito de que el interpolante $Q(f)$ mantenga las características de forma local de las funciones nodales requeriremos ciertas propiedades para los W_i en adición a (1.4). Específicamente, para conservar las siguientes ecuaciones

$$\frac{\partial Q(f)}{\partial x}(x_i,y_i) = \frac{\partial Q_i}{\partial x}(x_i,y_i)$$

y

$$\frac{\partial Q(f)}{\partial y}(x_i,y_i) = \frac{\partial Q_i}{\partial y}(x_i,y_i) : i=1,\dots,N$$

es indispensable que

$$\frac{\partial W_i(x_j, y_j)}{\partial x} = \frac{\partial W_i(x_j, y_j)}{\partial y} = 0 \quad : \quad i, j = 1, \dots, N$$

puesto que

$$\frac{\partial Q[f](x_i, y_i)}{\partial x} = \sum \frac{\partial W_i(x_i, y_i)}{\partial x} Q_i(x_i, y_i) + \sum \frac{\partial Q_i(x_i, y_i)}{\partial x} W_i(x_i, y_i)$$

$$\frac{\partial Q[f](x_i, y_i)}{\partial x} = 0 + \sum \frac{\partial Q_i(x_i, y_i)}{\partial x} W_i(x_i, y_i) = \frac{\partial Q_i(x_i, y_i)}{\partial x}$$

Similarmente para la otra ecuación.

Si proponemos el uso de funciones cuadráticas bivariadas para $Q_i(x, y) : i = 1, \dots, N$, entonces, si la propia función f fuera una cuadrática, la función Q , sería idéntica a f , es decir, $Q_i(x, y) = f(x, y) : i = 1, \dots, N$

Y, por lo tanto

$$Q[f](x, y) = \sum_{i=1}^N W_i(x, y) Q_i(x, y) = f(x, y) \sum_{i=1}^N W_i(x, y)$$

siempre que f sea una función cuadrática. Con el propósito de obtener una precisión cuadrática del interpolante modificado $Q[f]$, agregamos la siguiente condición

$$\sum_{i=1}^N W_i(x, y) = 1 \quad (1.5)$$

La elección de la función distancia que será utilizada en (1.2) cuando se calculen las funciones nodales fue hecha sobre la base de exhaustivas pruebas numéricas [4]. Franke and Little [2] propusieron

$$\frac{1}{\rho_i} = \frac{(R_i - d_i)_+}{R_i d_i} \quad ; \quad (R_i - d_i)_+ = \begin{cases} R_i - d_i & : R_i - d_i \geq 0 \\ 0 & : R_i - d_i < 0 \end{cases} \quad (1.7)$$

Lo cual trabaja bastante bien para elecciones adecuadas de los valores R_i .

Recordemos que Q_k es la solución del problema de inverso de distancias ponderadas por mínimos cuadrados en el punto $(x, y) = (x_k, y_k)$. La discusión acerca de la construcción de Q_k se simplifica si asumimos que

$$Q_k(x, y) = f_k + \bar{a}_{k2}(x - x_k) + \bar{a}_{k3}(y - y_k) + \bar{a}_{k4}(x - x_k)^2 + \bar{a}_{k5}(x - x_k)(y - y_k) + \bar{a}_{k6}(y - y_k)^2$$

Y así sólo buscaríamos la solución de

$$\text{Min}_{a_{k2}, \dots, a_{k6}} \sum_{\substack{i=1 \\ i \neq k}}^N \left[\frac{f_i + a_{k2}(x_i - x_k) + \dots + a_{k6}(y_i - y_k)^2 - f_i}{\rho_i(x_i, y_i)} \right]^2$$

Si $1/\rho_i$ está dado por (1.7), entonces siempre que $d_i(x_i, y_i) > R_i$, el punto (x_i, y_i, f_i) no tiene influencia puesto que su correspondiente término en la suma es cero. Así Q_k depende sólo de los puntos "cercaños" y es por lo tanto una aproximación local a f . Si ahora utilizamos funciones de peso W_i las cuales no se anulan sólo en alguna vecindad de (x_k, y_k) obtendremos un interpolante local.

CONSTRUCCION DE LOS R_i

La elección adecuada del "radio de influencia", R_i , como veremos a continuación, es un factor determinante para el buen desarrollo del método.

Con respecto a la selección del R_i involucrado en la definición de ρ , podemos decir lo siguiente:

Mientras que el uso de un radio variable, es decir $R_i \neq R_j$, proporciona flexibilidad al método, se ha encontrado que para propósitos de "interpolación general", la selección de esos valores puede muy frecuentemente ser un obstáculo que puede ser evitado utilizando un valor uniforme ($R_i = R$ para toda i).

Por otro lado, es importante notar que las funciones distancia ρ , se presentan en dos lugares:

1. En la definición de las funciones nodales Q_k
2. En la definición de las funciones de peso W_k (ver el *caso particular*, "Método de Shepard")

No es necesario utilizar el mismo "radio de influencia" en ambas situaciones, diversas pruebas han mostrado que es preferible utilizar diferentes valores, digamos $R = R_q$ en la definición de las funciones nodales y $R = R_w$ en la definición de las funciones de peso. Puesto que R_q denota el "radio de influencia" de los puntos en las funciones nodales y R_w el "radio de influencia" de las funciones nodales en la función interpolante $Q[f](x,y)$, resulta claro que debemos de elegir $R_w \leq R_q$. Con el propósito de hacer una elección razonable de ambos radios, se ha encontrado conveniente definir valores N_w y N_q , y calcular el radio de la región de influencia de acuerdo a las siguiente relaciones:

$$R_q = \frac{D}{2} \sqrt{\frac{N_q}{N}}$$

$$R_w = \frac{D}{2} \sqrt{\frac{N_w}{N}}$$

Donde

$$D = \max_{i,j} d_i(x_j, y_j)$$

Los valores de N_w y N_q representan el número de puntos que anticipadamente proponemos se encontrarán en círculos de radio R_w y R_q respectivamente. Para datos que de alguna forma están distribuidos uniformemente, se ha encontrado que un valor de $N_q = 18$ funciona bastante bien. Para conjuntos de datos que tienen algunas regiones en las que están relativamente separados y otras en las que están relativamente densos, o para conjuntos pequeños de datos ($N < 25$), es necesario incrementar el número de datos, puesto que la función interpolante está definida únicamente en la unión de los círculos de radio R_w centrados en los puntos conocidos $(x_i, y_i) : i = 1, \dots, N$. También se ha encontrado que el uso de la relación $N_q / N_w \approx 2$ es muy útil. Para evitar problemas cuando son menos de 5 datos los que se encuentran a una distancia R_q de un (x, y) dado, se ha incorporado un "fallback" automático a las funciones nodales lineales. En general se utiliza una *descomposición en valores singulares* para calcular los coeficientes de las funciones nodales, con lo cual se evita un posible problema de no unicidad.

SINTESIS

Finalmente como recapitulación, y a manera de "ejemplo", mencionaremos nuevamente los pasos esenciales de este método:

i) Seleccionar N_q y N_w con el propósito de definir:

$$\frac{1}{\rho_i} = \frac{(R_q - d_i)_+}{R_q d_i} : R_q = \frac{D}{2} \sqrt{\frac{N_q}{N}}$$

$$\frac{1}{\rho_i} = \frac{(R_w - d_i)}{R_w d_i} \quad ; \quad R_w = \frac{D}{2} \sqrt{\frac{N_w}{N}}$$

Los valores por *Default* para N_s y N_w son 18 y 9 respectivamente.

ii) Para $k = 1, \dots, N$, resolver el siguiente problema de mínimos cuadrados:

$$a_{kj} : j=2, \dots, 6 \quad \min \sum_{i=k}^N \frac{f_k + a_{k2}(x_i - x_k) + a_{k3}(y_i - y_k) + a_{k4}(x_i - x_k)^2 + a_{k5}(x_i - x_k)(y_i - y_k) + a_{k6}(y_i - y_k)^2 - f_i}{\rho_i(x_i, y_i)}$$

para obtener

$$\overline{a_{kj}} \quad ; \quad j=2, \dots, 6$$

iii) Definir

$$Q_k(x, y) = f_k + \overline{a_{k2}}(x - x_k) + \overline{a_{k3}}(y - y_k) + \overline{a_{k4}}(x - x_k)^2 + \overline{a_{k5}}(x - x_k)(y - y_k) + \overline{a_{k6}}(y - y_k)^2$$

y calcular

$$Q(x,y) = \frac{\sum_{k=1}^N \frac{Q_k(x,y)}{\rho_k^2(x,y)}}{\sum_{k=1}^N \frac{1}{\rho_k^2(x,y)}}$$

UN CASO PARTICULAR (Método de Shepard)

Otro caso particular, y más sencillo, para el método de inverso de las distancias ponderadas por mínimos cuadrados, (1.1)-(1.2), queda establecido cuando $m = 1$ y $\phi_i(x) = 1$. Entonces, al resolver (1.2), obtendremos que

$$\bar{a}_1(x,y) = \frac{\sum_{i=1}^N \frac{f_i}{\rho_i^2(x,y)}}{\sum_{i=1}^N \frac{1}{\rho_i^2(x,y)}}$$

Y así

$$P(x,y) = \frac{\sum_{i=1}^N \frac{f_i}{\rho_i^2}}{\sum_{i=1}^N \frac{1}{\rho_i^2}}$$

Inicialmente, este método fue propuesto por Shepard. Desafortunadamente, si se trabaja con él sin ninguna modificación no se tendrán buenas propiedades de ajuste. Claramente, el método de Shepard es un caso particular del método de inverso de las distancias ponderadas por mínimos cuadrados y ello ha sido puntualizado por varios autores [10]. Puede verificarse fácilmente que mientras $\rho(x, y) = 0$, las funciones

$$W_i = \frac{\frac{1}{\rho_i^2}}{\sum_{k=1}^N \frac{1}{\rho_k^2}}$$

cumplirán las condiciones impuestas por (1.4), (1.5) y (1.6). En esta última proposición, en realidad, estamos asumiendo que las funciones distancia ρ_i , $i = 1, \dots, N$ son lo suficientemente suaves como para que las derivadas (1.5) existen. Afortunadamente, este es el caso para la elección de las funciones ρ_i dadas por (1.7).

KRIGING.

PLANTEAMIENTO

Kriging es una técnica de estimación local que proporciona el mejor estimador lineal insesgado (MELI) de la característica desconocida en estudio. Esta limitación a la clase de estimadores lineales es bastante natural, puesto que ello significa que exclusivamente los momentos de segundo orden de la Función Aleatoria (es decir, la covarianza o variograma) son requeridos, y en general, es posible en la práctica *inferir* cualquiera de ellos.

Desde luego que cuando la información estructural (distribución de los puntos conocidos) disponible incluye más que el simple conocimiento de los momentos de segundo orden, entonces los estimadores no lineales, más precisos que Kriging simple, pueden ser definidos.

ECUACIONES DE KRIGING

Sea $Z(x)$ la función aleatoria bajo estudio. $Z(x)$ está definida en los puntos soporte y es estacionaria de segundo orden con :

Esperanza

$$E(Z(x)) = m$$

donde m es una constante generalmente desconocida.

Covarianza centrada

$$E(Z(x+h)Z(x)) - m^2 = C(h)$$

Variograma

$$E(Z(x+h) - Z(x))^2 = 2\gamma(h)$$

Cualquiera de esos dos momentos de segundo orden se supone conocido. Sin embargo, cuando únicamente existe el variograma, diremos entonces que la Función Aleatoria $Z(x)$ es intrínseca.

Es indispensable definir también la estimación del valor medio de $Z(x)$ sobre un dominio $V(x_0)$

$$Z_v(x_0) = \frac{1}{V} \int_{V(x_0)} Z(x) dx$$

Los datos experimentales que serán utilizados consisten de un conjunto de n valores discretos $\{Z_\alpha : \alpha = 1, \dots, n\}$

Notemos que bajo la hipótesis de estacionariedad (proceso débilmente estacionario) la esperanza de cada uno de esos datos es m , es decir $E(Z_\alpha) = m, \forall \alpha$

La estimación lineal Z_K^* considerada es una combinación lineal de los n valores conocidos, es decir,

$$Z_K^* = \sum_{\alpha=1}^n \lambda_\alpha Z_\alpha$$

Los n pesos λ_α son calculados de tal forma que se garantice que el estimador sea insesgado y que la estimación de la varianza sea mínima, y por lo tanto, diremos que el estimador Z_K^* es óptimo.

CONDICIONES DE INSESGAMIENTO:

Para obtener una *esperanza* de error igual a cero en $[Z_v - Z_K^*]$ (es decir, realizar un promedio sobre un gran número de situaciones similares) es suficiente con imponer la siguiente condición

$$\sum_{\alpha=1}^n \lambda_\alpha = 1$$

puesto que

$$E(Z_K^*) = \sum_{\alpha} E(\lambda_{\alpha} Z_{\alpha}) = \sum_{\alpha} \lambda_{\alpha} E(Z_{\alpha}) = m \sum_{\alpha} \lambda_{\alpha} = m = E(Z_V)$$

lo cual implica que

$$E(Z_V - Z_K^*) = 0$$

VARIANZA ESTIMADA MINIMA

Consideremos la siguiente relación $\text{Var}(Z_V - Z_K^*) = E\{(Z_V - Z_K^*)^2\} - E^2\{Z_V - Z_K^*\}$

$$= E(Z_V - Z_K^*)^2 \rightarrow Z_K^* = E(Z_V) = m \rightarrow \sum_{\alpha} \lambda_{\alpha} = 1$$

entonces $E\{(Z_V - Z_K^*)^2\}$ puede ser considerada como una estimación de la varianza, la cual puede, además, expresarse de la siguiente forma

$$E(Z_V - Z_K^*)^2 = E(Z_V^2) - 2E(Z_V Z_K^*) + E(Z_K^{*2})$$

con

$$E(Z_V^2) = \frac{1}{V^2} \int_V \int_V E(Z(x)Z(x')) dx dx' = \bar{C}(V, V) + m^2$$

$$E(Z_V Z_K^*) = \sum_{\alpha} \lambda_{\alpha} \frac{1}{V v_{\alpha}} \int_V \int_{v_{\alpha}} E(Z(x)Z(x')) dx dx' = \sum_{\alpha} \lambda_{\alpha} \bar{C}(V, v_{\alpha}) + m^2$$

$$E(Z_K'^2) = \sum_{\alpha} \sum_{\beta} \lambda_{\alpha} \lambda_{\beta} \frac{1}{v_{\alpha} v_{\beta}} \int dx \int E(Z(x)Z(x')) dx = \sum_{\alpha} \sum_{\beta} \lambda_{\alpha} \lambda_{\beta} \bar{C}(v_{\alpha}, v_{\beta}) + m^2$$

Como los términos m^2 se eliminan obtenemos:

$$E((Z_V - Z_K')^2) = \bar{C}(V, V) - 2 \sum_{\alpha} \lambda_{\alpha} \bar{C}(V, v_{\alpha}) + \sum_{\alpha} \sum_{\beta} \lambda_{\alpha} \lambda_{\beta} \bar{C}(v_{\alpha}, v_{\beta})$$

La notación estándar

$$\bar{C}(V, v_{\alpha})$$

designa el valor medio de la función de covarianza $C(h)$ cuando los dos extremos del vector h en forma independiente describen los dominios V y v_{α} respectivamente.

La varianza estimada puede, de esta forma, ser expresada como una forma cuadrática en λ_{α} y será minimizada sujeta a las condiciones de no insesgamiento

$$\sum_{\alpha} \lambda_{\alpha} = 1$$

Las ponderaciones o pesos óptimos se obtienen a través de las técnicas Lagrangianas estándar igualando a cero cada una de las n derivadas parciales

$$\frac{\partial [E(Z_V - Z_K)^2 - 2\mu \sum_{\beta} \lambda_{\beta}]}{\partial \lambda_{\alpha}}$$

$$\begin{aligned} \frac{\partial}{\partial \lambda_1} [2(\lambda_1 \bar{C}(V, v_1) + \dots + \lambda_n \bar{C}(V, v_n)) + \lambda_1 \lambda_1 \bar{C}(v_1, v_1) + \dots + \lambda_1 \lambda_n \bar{C}(v_1, v_n) + \dots + \lambda_n \lambda_1 \bar{C}(v_n, v_1) + \dots + \\ \lambda_n \lambda_n \bar{C}(v_n, v_n) - \mu(\lambda_1 + \dots + \lambda_n)] \\ \vdots \\ \frac{\partial}{\partial \lambda_n} [2(\lambda_1 \bar{C}(V, v_1) + \dots + \lambda_n \bar{C}(V, v_n)) + \lambda_1 \lambda_1 \bar{C}(v_1, v_1) + \dots + \lambda_1 \lambda_n \bar{C}(v_1, v_n) + \dots + \lambda_n \lambda_1 \bar{C}(v_n, v_1) + \dots + \\ \lambda_n \lambda_n \bar{C}(v_n, v_n) - \mu(\lambda_1 + \dots + \lambda_n)] \end{aligned}$$

Según vemos, este procedimiento proporciona un sistema de $(n + 1)$ ecuaciones lineales con $(n + 1)$ incógnitas (las n ponderaciones λ_{α} y el parámetro de Lagrange μ) el cual es conocido como el "Sistema Kriging".

$$\sum_{\beta=1}^n \lambda_{\beta} \bar{C}(v_{\alpha}, v_{\beta}) - \mu = \bar{C}(v_{\alpha}, V) \quad : \quad \forall_{\alpha} = 1, \dots, n \quad (2.1)$$

$$\sum_{\beta=1}^n \lambda_{\beta} = 1$$

La estimación mínima de la varianza o *varianza Kriging* puede entonces ser escrita de la siguiente forma

$$\sigma^2_K = E((Z_v - Z_K^*)^2) = \bar{C}(V, V) + \mu - \sum_{\alpha=1}^n \lambda_{\alpha} \bar{C}(v_{\alpha}, V) \quad (2.2)$$

Este sistema Kriging también puede ser expresado en términos de la función semivariograma $\gamma(h)$, particularmente cuando la función aleatoria $Z(x)$ es intrínseca solamente y la función de covarianza $C(h)$ no está definida.

$$\sum_{\beta=1}^n \lambda_{\beta} \bar{\gamma}(v_{\alpha}, v_{\beta}) - \mu = \bar{\gamma}(v_{\alpha}, V) : \quad \forall_{\alpha} = 1, \dots, n \quad (2.3)$$

$$\sum_{\beta=1}^n \lambda_{\beta} = 1$$

$$\gamma \quad \sigma^2_K = \sum_{\alpha} \lambda_{\alpha} \bar{\gamma}(v_{\alpha}, V) + \mu - \bar{\gamma}(V, V)$$

FORMA MATRICIAL

En forma matricial el sistema Kriging puede ser expresado como

$$[K] [\lambda] = [M2](V)$$

a través de lo cual y utilizando el sistema (2.1), obtenemos

$$[\lambda] = [K]^{-1} [M2](V)$$

que es la matriz de coeficientes óptimos que da solución al problema de minimización de la varianza sujeta a las condiciones de insesgamiento planteado inicialmente.

Y por lo tanto, llegamos a la siguiente expresión

$$\begin{aligned} Z_k^* &= \sum_{\alpha}^n \lambda_{\alpha} Z_{\alpha} \\ &= Z^T [K]^{-1} [M2] \end{aligned} \quad (2.4)$$

con $Z^T = (Z_1, \dots, Z_n)$

Observemos que si $\alpha = \alpha_i$, $[M2](V)$ es la i -ésima columna de $[K]$, entonces $[K]^{-1} [M2](\alpha_i) = (0, \dots, 0, 1, 0, \dots, 0)^T$ donde la unidad está en la i -ésima posición y por lo tanto, $Z_k^* = Z_i$ como se requería.

Lo anterior nos permite expresar la siguiente relación para la varianza

$$\sigma_k^2 = \bar{C}(V, V) - [\lambda]' [M2](V)$$

Donde la matriz de incógnitas $[\lambda]$ y el segundo miembro $[M2](V)$ pueden ser escritas como dos matrices columna

$$\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_\alpha \\ \cdot \\ \cdot \\ \cdot \\ \lambda_n \\ -\mu \end{bmatrix}$$

$$[M2](V) = \begin{bmatrix} \bar{C}(v_1, V) \\ \cdot \\ \cdot \\ \cdot \\ \bar{C}(v_\alpha, V) \\ \cdot \\ \cdot \\ \cdot \\ \bar{C}(v_n, V) \\ 1 \end{bmatrix}$$

Además el primer miembro $[K]$ o "matriz Kriging" se puede escribir como

$$[K] = \begin{bmatrix} \bar{C}(v_1, v_1) & \dots & \bar{C}(v_1, v_\beta) & \dots & \bar{C}(v_1, v_n) & 1 \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \bar{C}(v_\beta, v_1) & \dots & \bar{C}(v_\beta, v_\beta) & \dots & \bar{C}(v_\beta, v_n) & 1 \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \cdot & & \cdot & & \cdot & \cdot \\ \bar{C}(v_n, v_1) & \dots & \bar{C}(v_n, v_\beta) & \dots & \bar{C}(v_n, v_n) & 1 \\ 1 & \dots & 1 & \dots & 1 & 0 \end{bmatrix}$$

Nótese que la "matriz Kriging" es simétrica, es decir

$$\bar{C}(v_\alpha, v_\beta) = \bar{C}(v_\beta, v_\alpha) \quad \forall \alpha, \beta$$

Pasemos ahora a la situación en la que todas las mediciones son realizadas con errores σ^2 . Introduciremos, para ello, notación que nos permitirá comparar este método con el que ha sido desarrollado en varios escritos referidos en la bibliografía (Método de Splines).

Consideremos un conjunto de n observaciones y_i , cada una correspondiente a un punto x_i , $i = 1, \dots, n$; la función aleatoria en cuestión estará denotada por $f(x)$. Inicialmente, supusimos que dicha función aleatoria era observada sin error de tal forma que $y_i = f(x_i)$, pero para este caso tendremos que $y_i = f(x_i) + u_i$, donde los u_i tienen media cero y varianza σ^2 y no están correlacionados.

Asumiremos como algo *completamente natural*, el hecho de que esos términos de perturbación no están correlacionados con la función aleatoria f . El objetivo ahora es: *predecir $f(x)$* a través de

$$f^*(x) = \sum_1^n c_i y_i$$

donde las c_i 's serán elegidas de tal forma que minimicen

$$E\{f^*(x) - f(x)\}^2$$

que puede ser expresado como

$$\sum \sum c_i c_j [\sigma^2 \delta_{ij} + C(x_i, x_j)] - 2 \sum c_i C(x, x_i) + C(x, x)$$

Derivando con respecto a c_1, \dots, c_n encontramos que

$$D(x) = (\sigma^2 I + C)c$$

de lo que se deduce que

$$c = (\sigma^2 I + C)^{-1} D(x)$$

Donde $D(x) = [C(x_1, x), \dots, C(x_n, x)]^T$; $C(x, y)$ es la covarianza entre x, y ; C es la matriz cuadrada con elementos $C(x_i, x_j)$; $c = [c_1, \dots, c_n]^T$, e I la matriz identidad.

Por lo tanto, el mejor predictor lineal para y_1, \dots, y_n está dado por

$$f^*(x) = y^T (\sigma^2 I + C)^{-1} D(x) \quad (2.5)$$

claramente, esta ecuación se reduce a (2.4) cuando $\sigma \rightarrow 0$.

Al método anteriormente expuesto se le conoce como *Kriging Simple* cuando se tiene la certeza de que no existen errores en las mediciones efectuadas y como *Cokriging* cuando se sabe que sí existen tales errores, y en ambos casos, las ideas subyacentes básicas son mucho más antiguas en la estadística que en la literatura geoestadística.

ALGUNAS OBSERVACIONES PRACTICAS IMPORTANTES

OBSERVACION 1. *La existencia y unicidad de la solución.* El sistema Kriging (2.1) tiene una única solución si y sólo si la matriz de covarianza

$$[\bar{C}(v_\alpha, v_\beta)]$$

es definida estrictamente positiva y así necesariamente tiene un determinante estrictamente positivo. Por esta razón es suficiente que el modelo de covarianza puntual $C(h)$ utilizado sea definido positivo y que ninguno de los datos soporte v_α coincida completamente con otro. En realidad $v_\alpha = v_\beta$ significa que

$$\bar{C}(v_\alpha, v_\beta) = \bar{C}(v_\alpha, v_\beta) : \forall \beta$$

y que el determinante $|\bar{C}(v_\alpha, v_\beta)|$

sea, por lo tanto, cero.

OBSERVACION 2. Kriging que es un estimador insesgado, es también un interpolador exacto, es decir, si un punto (x, y) que va a ser estimado coincide con uno de los puntos soporte (dato conocido (x, y)), entonces el sistema Kriging proporciona:

- i) Un estimador Z_h^* idéntico al valor conocido
- ii) Una varianza igual a cero, es decir, $\sigma_h^2 = 0$.

En realidad, la "superficie Kriging" pasa a través de los datos conocidos, y esta es una propiedad que no todos los procedimientos de estimación satisfacen. Especialmente los métodos que utilizan polinomios generados por mínimos cuadrados.

OBSERVACION 3. El sistema Kriging, así como la varianza Kriging, dependen únicamente de las relaciones $C(h)$ o $\gamma(h)$ y de las "geometrías relativas" de los puntos soporte pero no de sus imágenes particulares, y por lo tanto, una vez que la configuración de los datos es conocida y antes de que se proceda a la obtención de las imágenes (decodificación de la información o consecución de los valores por métodos manuales), el sistema Kriging puede ser resuelto y su correspondiente estimación mínima de la varianza puede ser pronosticada.

OBSERVACION 4. La matriz Kriging $[K]$ depende únicamente de las geometrías relativas de los puntos soporte y no de la forma del dominio que será estimado, y por tanto, dos configuraciones idénticas de datos deben proporcionar la misma matriz Kriging y será suficiente calcular su matriz inversa $[K]^{-1}$ sólo una vez.

Las dos matrices solución (matrices columna) $[A]$ y $[A']$ se obtienen considerando los productos de $[K]^{-1}$ con las matrices del segundo miembro, es decir :

$$[A] = [K]^{-1} [M2] \quad \text{y} \quad [A'] = [K]^{-1} [M2']$$

Si adicionalmente a una idéntica configuración de los puntos soporte, es decir $[K] = [K']$, los dos dominios que serán estimados tienen la misma forma, entonces $[M2] = [M2']$, $[A] = [A']$, y por lo tanto es suficiente con resolver el sistema Kriging sólo una vez con el consecuente ahorro de tiempo computacional.

ENTORNO

VARIABLE ALEATORIA, FUNCION ALEATORIA.

Recordemos que una variable aleatoria es una función cuyo dominio es el espacio muestral Ω y su contradominio el conjunto de los números reales ($X : \Omega \rightarrow \mathbb{R}$), de tal forma que las imágenes que asume están determinadas por alguna función de distribución de probabilidad (*función de densidad*).

Ej: Se cuenta con un dado que en lugar de un número en cada cara tiene estampado un color diferente. El dado se lanza 5 veces.

Sea X = Color obtenido en cada lanzamiento

$$\Omega = \{\text{rojo, azul, verde, blanco, amarillo, violeta}\}$$

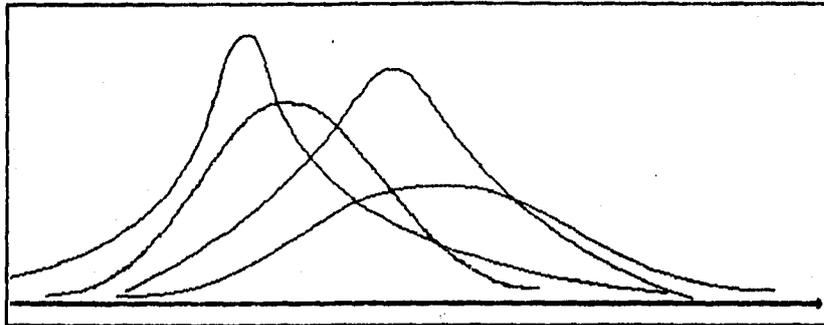
$$\Rightarrow \text{Supongamos que } X(\text{rojo}) = \pi, X(\text{azul}) = 7 + \sqrt{2}, X(\text{verde}) = 24, \\ X(\text{blanco}) = 13, X(\text{amarillo}) = 62.5$$

Notemos, sin embargo, que cada una de las 5 evaluaciones $X_i(\omega_i)$ $i = 1, 2, \dots, 5$ de la variable aleatoria puede ser pensada como la realización de una variable aleatoria que "cobra vida" por sí sola, es decir, $X(\text{rojo}) = \pi$ puede pensarse como una nueva variable aleatoria evaluada en uno de sus seis elementos del dominio y cuya imagen en este caso es π pero que claramente pudo haber sido cualquier otro real. Análogamente los cuatro restantes.

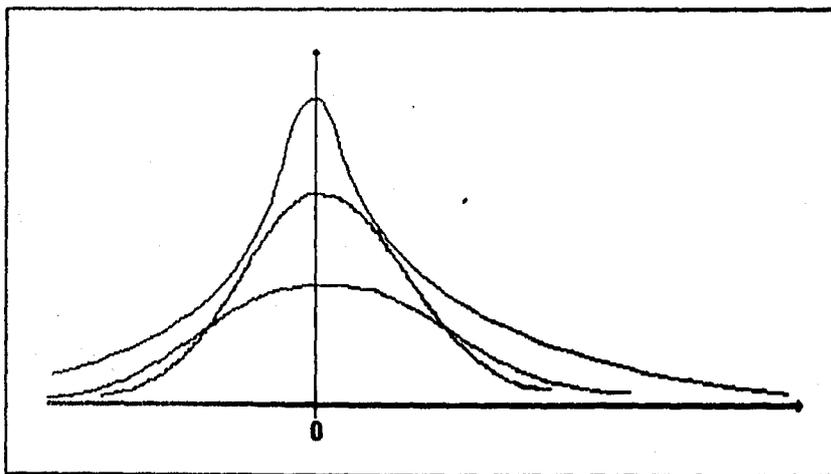
A este conjunto de nuevas variables aleatorias generadas a partir de una variable original se le conoce como *Función Aleatoria*.

PARAMETROS, ESTIMACION, ESTIMADORES MELI.

En el ambiente de la probabilidad y la estadística, un *parámetro* se define como una "variable" que determina en forma unívoca alguna función de distribución de probabilidades (proveniente del estudio de algún fenómeno de una población) por tanto, esa variable una vez determinada queda fija y caracterizando perfectamente el comportamiento de la población a través de su función de densidad (distribución de probabilidades).



Supongamos, por ejemplo, que tenemos una población cuya función de densidad (determinada por algún método) es una normal con media y varianzas no especificadas pero existentes. En este problema tenemos no uno sino dos parámetros que, en realidad, tratados de la forma adecuada pueden considerarse como uno solo (1 vector con dos componentes o un parámetro en \mathbb{R}^2). Claramente, mientras no conozcamos el valor del parámetro estamos limitados a saber sólo que la media y varianza de la función de densidad pueden ser cualquier par de números reales y que su representación gráfica, por ejemplo, puede ubicarse en cualquier punto del eje real y tener cualquier forma (ancha, angosta, "picuda")



Pero si ahora supieramos que la media de la función de densidad es cero (aunque siguiera sin especificarse la varianza) dicha función estaría ya más cerca de ser determinada puesto que de esta forma se tendría que de entre todos los números reales posibles es justamente alrededor el cero donde la función es simétrica y esto "reduce enormemente" las posibilidades para la forma de la función.

No obstante, son todavía una infinidad de funciones las que se ajustan a la pseudocarecterización impuesta por $E(X) = 0$.

Qué pasará ahora si especificamos también la varianza de la función. Observemos, dado que de alguna forma este número nos indica donde estarán los puntos de inflexión (en $E(X) - (Var(X))^{1/2}$ y $E(X) + (Var(X))^{1/2}$ o $\mu - \sigma$ y $\mu + \sigma$) que aparentemente el número de funciones candidatas se reduciría considerablemente a aquellas que siendo simétricas en $x = 0$ y con puntos de inflexión en $\mu - \sigma$ y $\mu + \sigma$ tuvieran cualquier imagen (máximo global) en $x = 0$ sin embargo no es así, solamente existe una función que debido a su naturaleza de función de densidad o de distribución de probabilidades

$$\int_{-\infty}^{\infty} f = 1$$

posee ambos atributos.

Así pues, vemos como la determinación de los parámetros de una función de densidad caracteriza en forma exclusiva su forma, comportamiento y peculiaridades y ello es valiosísimo para la *inferencia estadística*.

Para la estadística clásica, el problema de la *estimación de parámetros (θ)* a partir de una muestra presentó gran relevancia, se pretendía encontrar algún "indicador" que nos permitiera "inferir o pronósticar" el comportamiento de alguna característica desconocida o poco precisa de la población.

Evidentemente, se buscó que esta aproximación fuera lo más exacta posible, lo cual se consigue si minimizamos la diferencia entre el valor propuesto y el valor verdadero (parámetro poblacional) es decir:

$$\text{Minimizar} \quad E[(\theta^* - \theta)^2] \quad (2.1)$$

$$\text{y como (2.1) es igual a} \quad \text{Var}(\theta^*) + \text{Sesgo}^2(\theta^*) \quad (2.2)$$

donde θ = parámetro poblacional (valor verdadero desconocido), θ^* = estimador o predictor.

Entonces el problema se reduce a minimizar cada uno de los sumandos, es decir, a encontrar un estimador con la varianza y sesgo mínimos.

Así pues, como una primera aproximación, bastante natural, se limitó el problema de la "construcción" del estimador del parámetro poblacional de tal forma que únicamente se considerarían *predictores o estimadores "insesgados"*, es decir, aquellos para los que su esperanza es igual al parámetro mismo o con otras palabras, aquellos para los que su sesgo es igual a cero.

$$E(\theta^*) = \theta$$

o bien

$$E(\theta^*) - \theta = 0$$

pues de esta forma se garantiza que será un indicador preciso, exacto, fiel (medianamente pues falta ver qué pasa con su varianza).

Con respecto a la *minimización de la varianza*, se desarrollaron cotas superiores que permiten visualizar la bondad de los estimadores propuestos y se aportaron algunos teoremas con suficiente aplicación práctica que permiten decidir fácilmente si algún estimador tiene o no la menor varianza.

Ahora bien, sobra decir que en términos prácticos siempre es más codiciada una relación lineal, una ecuación lineal, o en términos más generales, la representación mediante un modelo lineal de algún fenómeno en estudio, que una relación no lineal (suponiendo que son igualmente buenos en sus otras características) pues la primera resulta ser de manipulación algebraica más sencilla y de interpretación geométrica más inmediata.

Los estimadores, evidentemente, tienen un comportamiento similar, es decir, los hay lineales y no lineales y los primeros son los más valorados.

A esta selecta clase de estimadores insesgados, lineales y que tienen varianza mínima se les conoce como estimadores *MELI*.

Ej: Supongamos tener una Muestra Aleatoria X_1, X_2, \dots, X_n , de una función de distribución de probabilidades $f(x, \theta) = \theta e^{-\theta x} : x \in (0, \infty)$ y que queremos estimar el parámetro $1/\theta$.

\bar{x}

es un estimador insesgado, lineal y se puede probar que tiene varianza mínima; por lo tanto es un estimador MELI

MOMENTOS, MOMENTOS CENTRALES, MOMENTOS DE 2º ORDEN.

Por otro lado, sabemos que si X es una Variable Aleatoria, el r -ésimo momento de X denotado por μ_r' se define como

$$\mu_r' = E(X^r)$$

si la esperanza existe. Notemos que $\mu_1' = E(X)$, es decir la media de X .

Además, si X es Variable Aleatoria, el r -ésimo momento central de X alrededor de a (a constante) se define como $E[(X-a)^r]$ y si $a = \mu_x$, tendremos el r -ésimo momento central alrededor de la media denotado por μ_r , y definido como

$$\mu_r = E[(X - \mu_x)^r] = \int_{\mathbb{R}} (X - \mu_x)^r f_X(x) dx$$

Notemos que $\mu_1 = E(X - \mu_x) = 0$ y que $\mu_2 = E(X - \mu_x)^2$ es igual a la varianza de X , es decir, el segundo momento central alrededor de la media es igual a la varianza.

Si ahora tenemos 2 Variables Aleatorias X, Y o X_t, X_{t+h} definidas en el mismo espacio de probabilidad, la covarianza de X y Y (o autocovarianza, pues se trata de la misma variable aleatoria recorrida h períodos en el tiempo, de X_t y X_{t+h}) se define como

$$\text{Cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)] = E(XY) - \mu_x \mu_y$$

Notemos que la covarianza (autocovarianza) de una Variable Aleatoria consigo misma es igual a su varianza. Así pues en el fondo, la covarianza también es un momento de segundo orden.

VARIOGRAMA

Si consideramos 2 Variables Aleatorias $Z(x)$ y $Z(x+h)$ en dos puntos x y $x+h$ separados por un vector h , la variabilidad entre esas dos cantidades está caracterizada por la función variograma $2\gamma(x, h)$ la cual está definida como la esperanza de de la también Variable Aleatoria $(Z(x) - Z(x+h))^2$, es decir,

$$2\gamma(x,h) = E(Z(x) - Z(x+h))^2$$

Esta función depende tanto del punto x como del vector h y por lo tanto su estimación requiere de varias realizaciones, $(z(x), z(x+h))$, del par de Variables Aleatorias $(Z(x), Z(x+h))$ pero en la práctica sólo una de tales realizaciones está disponible y es la que corresponde a las mediciones actuales de los valores en los puntos x y $x+h$.

Para evitar este problema, se introduce la hipótesis intrínseca, la cual consiste en considerar que el variograma $2\gamma(x,h)$ depende sólo del vector de separación h (módulo y dirección) y no de la localización del punto x . De esta forma, es posible estimar el variograma $2\gamma(h)$ a partir de los datos disponibles.

Un estimador $2\gamma'(h)$ es la media aritmética de las diferencias al cuadrado entre dos medidas experimentales $(Z(x_i), Z(x_i+h))$ en cualquier par de puntos separados por el vector h ; es decir,

$$2\gamma'(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i+h)]^2$$

Donde $N(h)$ es el número de pares experimentales $(Z(x_i), Z(x_i+h))$ de datos separados por el vector h .

Nótese que el variograma así definido puede ser visto como una estimación de la varianza. La varianza del error cometido cuando la imagen en el punto x es estimada por la imagen en el punto $x+h$

Por lo tanto, resulta que también el variograma puede considerarse como un momento de segundo orden

ESTACIONALIDAD DE ORDEN 2 (PROCESO DÉBILMENTE ESTACIONARIO)

Estacionalidad de orden 2 o Proceso débilmente estacionario: Se dice que una Función Aleatoria es estacionaria de orden 2 o débilmente estacionaria cuando:

i) La esperanza matemática, $E(Z(x))$, existe y no depende del punto soporte x , es decir:

$$E(Z(x)) = m : \forall x$$

ii) Para cada par de Variables Aleatorias $\{Z(x), Z(x+h)\}$ la covarianza existe y depende de la distancia de separación h

$$C(h) = E(Z(x+h) \cdot Z(x)) - m^2 : \forall x$$

donde h representa un vector de coordenadas (h_x, h_y, h_z) en el espacio tridimensional.

La estacionalidad de la covarianza implica la estacionalidad de la varianza y del variograma, según podemos ver en las siguientes ecuaciones:

$$Var(Z(x)) = E(Z(x) - m)^2 = C(0) : \forall x$$

$$\gamma(h) = \frac{1}{2} E(Z(x+h) - Z(x))^2 = C(0) - C(h) : \forall x$$

Esta última ecuación indica que bajo la hipótesis de estacionalidad de segundo orden, la covarianza y el variograma son dos herramientas equivalentes para caracterizar la autocorrelación entre dos variables $Z(x+h)$ y $Z(x)$ separadas por una distancia h .

HIPOTESIS INTRINSECA

Decimos que una función aleatoria es intrínseca cuando:

- i) La esperanza matemática existe y no depende del punto soporte x
- ii) Para todos los vectores h , el incremento $(Z(x+h) - Z(x))$ tiene una varianza finita la cual no depende de x

$$\text{Var}(Z(x+h) - Z(x)) = E(Z(x+h) - Z(x))^2 = 2\gamma(h) : \forall x$$

De esta forma, la estacionalidad de segundo orden implica la hipótesis intrínseca pero el inverso no es válido

ESTIMACION DE LA VARIANZA

Sea $Z(x)$ una Función Aleatoria estacionaria de segundo orden con Esperanza m , Covarianza $C(h)$, y semivariograma $\gamma(h)$. Consideremos primero el caso simple de la estimación de la media aritmética z_K de los K valores desconocidos $\{z(x_k) : k = 1 \text{ a } K\}$:

$$z_K = \frac{1}{K} \sum_{k=1}^K z(x_k)$$

Un estimador lineal z_K^* de z_K , dada la información de que se dispone, es la media aritmética de los n datos disponibles $\{z(x_i) : i = 1 \text{ a } n\}$:

$$z_K^* = \frac{1}{n} \sum_{i=1}^n z(x_i)$$

GENERALIZACION AL CASO CONTINUO

Supongamos ahora que los K puntos x_k están localizados en una región V centrada en el punto x y que los n puntos x_i dentro de otra región V' centrada en el punto x' , si K y n tienden a infinito las anteriores medias aritméticas z_K y z_K^* tienden a los valores medios en V y V' de la variable $z(y)$, es decir,

$$z_{k'} \rightarrow z_v(x) = \frac{1}{V} \int_{v(x)} z(y) dy \quad y \quad z_{k'} \rightarrow z_v(x') = \frac{1}{V} \int_{v(x')} z(y) dy$$

Esos valores medios $z_v(x)$ y $z_v(x')$ son interpretados como realizaciones particulares de dos variables aleatorias $Z_v(x)$ y $Z_v(x')$:

$$Z_v(x) = \frac{1}{V} \int_{v(x)} Z(y) dy \quad y \quad Z_v(x') = \frac{1}{V} \int_{v(x')} Z(y) dy$$

$$\rightarrow Z_v^2(x) = \frac{1}{V^2} \left[\int_{v(x)} Z(y) dy \cdot \int_{v(x)} Z(y) dy \right]$$

$$= \frac{1}{V^2} \left[\int_v dy \int_v Z(y) Z(y') dy' \right] \rightarrow E(Z_v^2(x)) = \frac{1}{V^2} \left[\int_v dy \int_v E(Z(y) Z(y')) dy' \right]$$

y como

$$\text{Cov}(X, Y) = E(X, Y) - E(X)E(Y)$$

$$\rightarrow E(Z_v^2(x)) = \frac{1}{V^2} \left[\int_v dy \int_v (C(y-y') + m^2) dy' \right]$$

donde $y-y'$ denota la distancia que existe entre el par de variables $Z(y)$ y $Z(y')$

$$= \frac{1}{V^2} \left[\int_V dy \left(\int_V C(y-y') dy' + \int_V m^2 dy' \right) \right]$$

$$= \frac{1}{V^2} \int_V \int_V C(y-y') dy dy' + \int_V m^2 dy' = \bar{C}(V,V) + m^2$$

Donde $\bar{C}(V,V)$

denota el valor medio de la función covarianza cuando ambos extremos del vector h describen el dominio $V(x)$

APLICACION

DESCRIPCION DEL PROBLEMA

Dependiendo del área en la que se esté inmerso y del tema que se desee abordar se clasifican sectores geográficos del planeta estableciendo como directrices básicas de clasificación algunas de las características propias de la región. Así por ejemplo, se habla de países tercermundistas, capitalistas y socialistas, de países miembros o no de tal o cual mercado o asociación comercial o financiera, de países de raza negra, nórdica o amarilla, de países con armamento nuclear o que carecen de él, etc.

"Meteorológicamente" hablando, también se establece una división del planeta, y en este caso se hace referencia a 4 regiones: *"Las cuatro regiones meteorológicas mundiales"*. México, Alaska, Canadá, EE.UU, Centro América y el Caribe conforman la denominada cuarta región. En ella individualmente cada país lleva un control más o menos serio, más o menos moderno, más o menos de calidad (según el país en cuestión), de sus condiciones meteorológicas, pero además existe el compromiso institucional entre los Servicios Meteorológicos de estos países por mantener un intercambio permanente de información.

De esta forma, México cuenta con información generada internamente a través de su red de observatorios y con datos obtenidos en los países restantes vía comunicación satelital con EE.UU. (pues este país es un centro regional de recolección) es decir, con información de toda la cuarta Región.

Todo este cúmulo de mensajes meteorológicos ayuda a comprender mejor y a pronosticar, que es de gran importancia, las condiciones meteorológicas a una escala incluso hemisférica, y, en realidad, así debe de ser, se debe de contar con una visión global, entera, real de la situación meteorológica, sin embargo debido a una gran diversidad de limitaciones de toda índole y para fines de nuestra aplicación, abusaremos de lo bondadoso en cuanto a "restricciones al abordar fenómenos de tipo natural" se refiere en los métodos de interpolación presentados y asumiremos que las regiones fuera de la república Mexicana en conjunción con los "nodos internos" generadores de información constituyen simplemente un conjunto de datos esparcidos irregularmente a través de los cuales queremos obtener información en las áreas intermedias. Notemos, sin embargo, que en la práctica este enfoque sencillo se aleja mucho de la realidad pues al estudiar fenómenos naturales hay que considerar "condiciones externas" que intervienen directamente en su comportamiento.

Téngase en cuenta, por ejemplo que los sistemas de altas y bajas presiones varían a lo largo del día y a través del año, que las masas de aire y los frentes se desplazan a lo largo de extensas porciones del globo, que las Ondas Tropicales y la posición de la Zona Intertropical de Convergencia está en función de las

temperaturas de los océanos y los continentes y que todo ello, en conjunción con muchos fenómenos meteorológicos más, afecta un comportamiento sencillo e ideal y regulan notablemente las condiciones atmosféricas.

Estas "variables extras", evidentemente, presentan una dificultad inata para "dejarse" modelar y conjuntarse de tal forma que el resultado sea un producto (modelo, sistema o fórmula mágica) receptor de información y emisor de, por ejemplo, pronósticos.

En las aplicaciones de la *Meteorología Clásica o Tradicional*, el producto que se desarrolla y posteriormente se explota con fines de pronosticar el estado del tiempo lo constituyen las denominadas cartas de superficie y cartas de altura, las cuales no son más que mapas de isolinias o curvas de nivel generadas a partir de un conjunto de datos meteorológicos de superficie o de algunos niveles de altura en la atmósfera, estos mapas o cartas en muchos casos son todavía elaborados a mano; se consideran, por ejemplo, para el caso de observaciones de altura, los datos que los centros de radiosondeo a través de las radiosondas registran y transmiten para un nivel de presión específico, 500 milibares entre otros, se conjunta toda la información y a partir de esta colección discreta e irregularmente distribuida de datos se "adivina" el comportamiento del parámetro en todos los puntos restantes; salta a la vista que con este método, incluso al especialista más experimentado se le presentan cotidianamente diversos tipos de impresiones o, incluso, errores severos.

Para nuestra aplicación, nosotros utilizaremos la información de superficie proporcionada por la red de Observatorios del Servicio Meteorológico Nacional (84 en total de los cuales sólo 79 transmiten regularmente)

NATURALEZA DE LA INFORMACION UTILIZADA

Diariamente el Servicio Meteorológico Nacional recibe una gran cantidad de información de carácter meteorológico proveniente directa o indirectamente de su red de observatorios distribuidos a lo largo y ancho de la república mexicana.

Los mensajes se reciben a través de teléfono (el observador dicta y el operador en el Centro Nacional de Captura, *Radio XBA*, captura en terminales), vía módem (el personal en los observatorios crea el archivo de cómputo y utilizando programas comerciales de comunicación como *Telix* o *Bitcom*, se enlaza con XBA y le transmite el mensaje), por radio de banda lateral (BLU) y vía satélite (GOES).

El contenido, como decíamos, es de naturaleza meteorológica, es decir, cada uno de los observatorios registra periódicamente las condiciones meteorológicas de su entorno (temperatura del aire, presión atmosférica, dirección e intensidad del viento, precipitación, nubosidad, temperatura de punto de rocío, etc.) las codifica

según formatos y convenciones reconocidos mundialmente y las transmite a un centro regional o directamente al centro nacional según su ubicación geográfica.

Existen diferentes tipos de mensajes según la naturaleza de las observaciones que contengan y el lugar desde el cual hayan sido efectuadas, éstas pueden ser de superficie y de altura, dentro de los de superficie se encuentran aquellos que son generados por los observatorios (puntos fijos y "corto alcance"), aquellos que son

generados desde diferentes puntos de los océanos por barcos equipados con equipo meteorológico (puntos móviles) y aquellos que son efectuados con radares (puntos fijos y "gran alcance"). La información de altura, por otro lado, se obtiene a través de globos piloto, globos equipados con *radiosondas* y con *Satélites Meteorológicos*, para dar origen a mensajes PILOT, RADIOSONDEOS e IMAGENES de la situación actual del territorio nacional respectivamente.

Para un buen análisis del estado del tiempo que redunde finalmente en un excelente pronóstico, se requiere, indiscutiblemente, establecer un adecuado balance entre el tipo de información que se explotará: de superficie o de altura, de puntos fijos o puntos móviles, de corto o gran alcance. Es evidente que a mayor información trabajada, corresponden mejores elementos de análisis y resultados de calidad más confiable.

Para nuestra aplicación, nos limitaremos exclusivamente a explotar información de superficie de puntos fijos y de corto alcance, dicha información se encuentra contenida en mensajes denominados *sinópticos* elaborados 8 veces al día (las 0,3,6,9,...,21 hrs, se conocen como horas sinópticas) por c/u de los observatorios de la red del SMN.

FORMATO QUE GUARDA LA INFORMACION UTILIZADA

El formato (una parte) que sirve de patrón para codificar y transmitir este tipo de mensajes se denomina *FM12-IX SYNOP* y está constituido como sigue:

Sección 0

M_iM_jM_kM_l YYGGI_w IIIII

Sección 1

I_hX_hhVV Nddff (00fff) 1S_nTTT 2S_nTdTd

3P_oP_oP_oP_o 4PPPP 5appp 6RRRt_R 7wwW₁W₂

8N_hC_LC_MC_H 9hh//

La codificación completa incluye, cuando menos, una sección más (sección 3) aunque eventualmente se incluye otra más (sección 5).

Cada uno de los grupos representa un elemento meteorológico: El grupo 1 se refiere a la temperatura del aire, el 3 a la presión atmosférica en el nivel de la estación, el 4 presión atmosférica al nivel medio del mar, el 8 a la cantidad y tipo de nubes, etc. Dentro de cada grupo, cada símbolo tiene un significado preciso y claro.

Con toda la información que hemos mostrado y la que se incluye en las secciones restantes se pueden hacer diferentes tipos de análisis, éstos pueden referirse al

comportamiento de sólo un elemento meteorológico o a la variación conjunta de varios de ellos. Nosotros abordaremos el primer caso, esto es, consideraremos a cada elemento en forma aislada y con él desarrollaremos, mediante la aplicación de

la herramienta teórica propuesta algunos programas de cómputo, puntos de vista y argumentos que nos permitan decidir sobre las condiciones inciertas de algún lugar.

SELECCION Y ACOTACION DEL PROBLEMA

Supongamos para fijar ideas, que nos enfrentamos a la siguiente situación que dicho sea de paso es bastante real. La Secretaría de Gobernación desea conocer el valor de precipitación que se registrará "en cada punto" de la República Mexicana el 21 de agosto para evitar cualquier tipo de contingencia y en su caso implementar los mecanismos que permitan que la afluencia de votantes no se vea disminuida, para ello encomienda al SMN que de alguna forma proporcione un análisis que permita tener un punto de vista claro y que contribuya a la vez a la consecución de alternativas de solución del problema.

PROCEDIMIENTO A SEGUIR

El SMN haciendo uso de la información histórica con que cuenta para cada observatorio, de alguna forma (mediante técnicas de estadística descriptiva y análisis de regresión) obtiene un número representativo que utilizará en combinación con la información actual (último mes por ejemplo) como información disponible en varios nodos (pues no en todos existía información y a partir de ese conjunto de datos distribuidos irregularmente emitirá un "pronóstico" para la precipitación.

Para explotar la información actual que se encuentra codificada en archivos de cómputo y con formato FM12-IX, los archivos son "pasados" por el programa DECODI.FOR el cual, a partir de estos "datos crudos" genera tablas de información decodificada (archivos) que servirán como insumos para otros programas o como material para usuarios finales. Los archivos generados son los siguientes: TAB_COM, DESCRIPT y DAT_MAP

de los cuales, el último, DAT_MAP, es el que contiene la información indispensable para nuestro trabajo: El valor de un parámetro decodificado, que el usuario selecciona, al momento de correr el programa y que puede ser cualquiera de los que contiene el mensaje sinóptico, sin embargo, con fines ilustrativos el código que se presenta únicamente está habilitado para crear DAT_MAP con información de temperatura ambiente, presión atmosférica y precipitación. Además DAT_MAP incluye las coordenadas geográficas del observatorio del cual proviene el valor registrado. Los otros dos archivos son creados para satisfacer necesidades de la sociedad civil y los medios de comunicación pues contienen *toda* la información decodificada de cada observatorio que envió sus mensajes (TAB_COM) así como *números descriptivos o medidas de tendencia central* de estos datos (DESCRIPTIVO).

De esta forma, la información actual se hace manejable y es posible vincularla a los datos históricos para crear finalmente el "numero mágico" que nos permita efectuar las actividades de "prognosis".

Asumamos pues que la siguiente es la lista de información con su representación geográfica:

LATITUD	LONGITUD	DATOS
33 0	101 00.	24 0
32 0	94 00.	10 0
30 4	107 92.	20 0
29 0	111 00.	18 0
29 2	118 32.	18 0
29 0	107 62.	17 0
28 6	106 07.	23 7
28 7	100 52.	26 0
28 0	88 00.	27 0
28 0	110 80.	30 0
26 9	101 42.	26 8
26 0	94 00.	25 0
25 5	103 45.	24 4
25 9	100 20.	24 0
25 0	91 00.	23 0
24 2	110 35.	25 7
24 0	104 67.	18 4
23 0	114 00.	26 0
23 2	106 42.	27 0
23 7	99 13.	23 0
23 8	98 20.	21 0
22 8	102 58.	11 8
22 2	100 98.	12 8
22 2	97 85.	27 2
21 5	104 88.	19 0
21 9	102 30.	12 0
21 0	101 25.	12 4
21 9	100 00.	20 0
20 7	103 68.	17 4
20 1	98 73.	11 2
20 1	98 37.	11 0
21 0	80 70.	23 8
10 0	115 00.	24 0
19 1	104 33.	25 1
19 7	103 47.	18 5
19 7	101 18.	13 2
19 4	99 13.	12 7
19 9	98 20.	11 5
19 5	96 02.	18 8
19 2	96 12.	25 2
19 9	90 53.	23 3
18 8	111 00.	26 3
18 2	94 42.	26 8
18 5	88 30.	27 8
17 0	117 00.	27 0
17 8	99 50.	18 0
16 6	99 92.	25 5
16 8	83 10.	22 9
16 3	92 13.	19 2
14	105 00.	26
14 92	92 27.	23 3

NOTA: Debemos hacer notar que en las coordenadas geográficas reales las longitudes son negativas, sin embargo, para simplificar les hemos invertido el signo.

Ya en este nivel, atravesaremos, cuando menos, por 3 etapas para la obtención del objetivo final:

PRIMERA FASE

El archivo con formato DAT_MAP le es pasado como archivo de entrada al programa que dibuja mapas DIBMAP.FOR (ver sección de programas) el cual, utilizando 3 archivos más con datos de coordenadas geográficas reales y algunas transformaciones en coordenadas polares, nos muestra en la pantalla un mapa de la cuarta región con sus divisiones de latitud y longitud cada 5 grados, incluyendo además la localización, en la república mexicana exclusivamente, de los observatorios que transmitieron información y que fue decodificada por DECODI.FOR (se marca dibujando una circunferencia).

es posible además, aunque esto aún no esté bien implementado, substituir la circunferencia que representa la localización del observatorio por el valor del parámetro decodificado y posteriormente utilizar los datos resultado de la interpolación para "plotearlos" sobre el mismo mapa y obtener así toda el área de interés rellena de datos estimados, lo cual nos sirve como información distribuida en forma "global y micro" y puede contribuir, indiscutiblemente, a nuestra actividades de pronóstico (ver diagrama de flujo).

SEGUNDA FASE

Para esta opción se desarrollan un par de programas de cómputo cuyo objetivo central es el de tratar de rebasar un marco exclusivamente teórico del tratamiento de los métodos de interpolación, es decir, de alguna forma se intenta implementar herramientas que contribuyan a que realmente sea manejable en la práctica la teoría propuesta.

Ambos programas codifican las ecuaciones del método de inverso de las distancias ponderadas por mínimos cuadrados, aunque esto no quiere decir que sean una aplicación fiel de la teoría pues en ellos se dejan de manifiesto muchas suposiciones y se hacen varias simplificaciones a las ideas teóricas presentadas, por ejemplo: Se utiliza el caso en que $m = 1$ y $\phi_i(x) = 1$ (Método de Shepard); No se definen "radios de influencia" de una manera formal (pues notemos que $m = 1$ y $\phi_i(x) = 1$) además al no considerar las cantidades N_w , tampoco se pueden definir radios R_w centrados en los puntos conocidos.

Básicamente, el desarrollo que siguen los programas propuestos está regido por las siguientes etapas:

Se lee el archivo DAT_MAP y su información (latitud, longitud, valor del parámetro decodificado) es trasladada a un arreglo bidimensional en cuyas variables horizontales (renglones) se representan las longitudes y en las verticales (columnas) las latitudes y en la localidad de memoria de cada dirección (referencia cruzada) se almacena el valor del parámetro.

Debido a que la ubicación real de los observatorios no corresponde a coordenadas enteras sino reales y a que un arreglo (en fortran) no puede referenciarse con índices que no sean enteros, los datos de coordenadas que son leídos de DAT_MAP son redondeados pues de otra forma no sería posible trasladarlos a la matriz; es claro que este traslado de información, que en términos reales puede traducirse en cientos de kilómetros, redundará finalmente en una pérdida de exactitud que puede incluso ser muy severa.

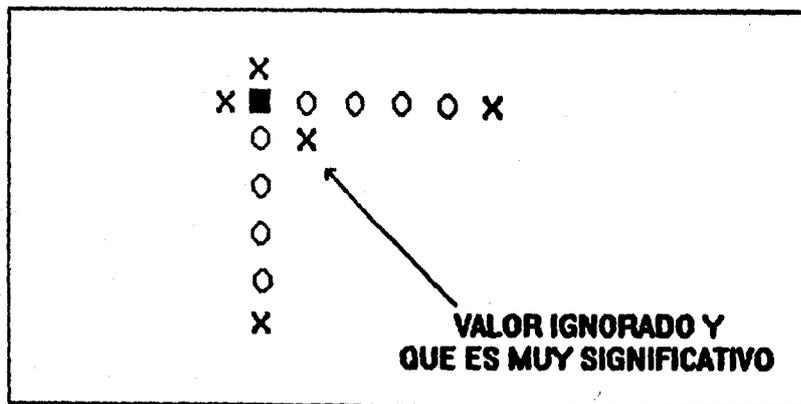
La matriz donde se almacena la información, cuenta con 20 renglones y 33 columnas que es el espacio suficiente para emular las intersecciones geográficas de 5 en 5 grados de latitudes y longitudes en la República Mexicana. Evidentemente, el archivo DAT_MAP no es suficiente para cubrir 20x33 variables pues este contiene a lo más 79 datos.

```
01 01 0101 01 010101 01 010101 01 01 01 01 01 01 01 01 01 01 01 0101 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 010101 01 01 01 01 01 01 240 01 01 01 01 0101 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 010101 01 01 01 01 01 01 01 01 01 01 01 0101 180 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 010101 01 010101 200 01 01 01 01 01 01 01 01 01 01 01 0101 01 01 01 01 01 010101
01 180 0101 01 010101 180 010101 01 01 01 01 01 01 01 01 01 01 01 0101 01 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 010101 170 237 01 01 01 01 01 01 260 01 01 01 0101 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 300 0101 01 01 01 01 01 01 01 01 01 01 01 0101 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 010101 01 01 01 01 01 01 268 01 01 01 01 0101 250 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 010101 01 01 01 01 244 01 01 240 01 01 01 0101 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 257 0101 01 01 01 184 01 01 01 01 01 01 0101 01 01 01 01 01 01 010101
01 01 0101 01 260 0101 01 01 0101 01 270 01 01 01 01 01 01 01 230 210 01 0101 01 01 01 01 01 010101
01 01 0101 01 010101 01 010101 01 01 01 01 118 01 128 01 01 272 0101 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 010101 01 01 198 01 120 124 200 01 01 01 0101 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 010101 01 01 01 174 01 01 01 01 110 01 01 01 01 01 01 01 01 238 01 01 0101
01 01 0101 240 010101 01 010101 01 01 251 168 01 132 01 127 118 01 262 01 01 01 01 233 01 01 010101
01 01 0101 01 010101 263 010101 01 01 01 01 01 01 01 01 01 01 01 01 265 01 01 01 01 01 278 01 01 0101
01 01 270 01 01 0101 01 010101 01 010101 01 01 01 01 01 01 180 01 01 0101 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 010101 01 01 01 01 01 01 268 01 01 0101 01 228 192 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 010101 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 010101 01 01 290 01 01 01 01 01 01 01 01 01 01 01 233 01 01 01 01 01 01 010101
01 01 0101 01 010101 01 010101 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 01 010101
```

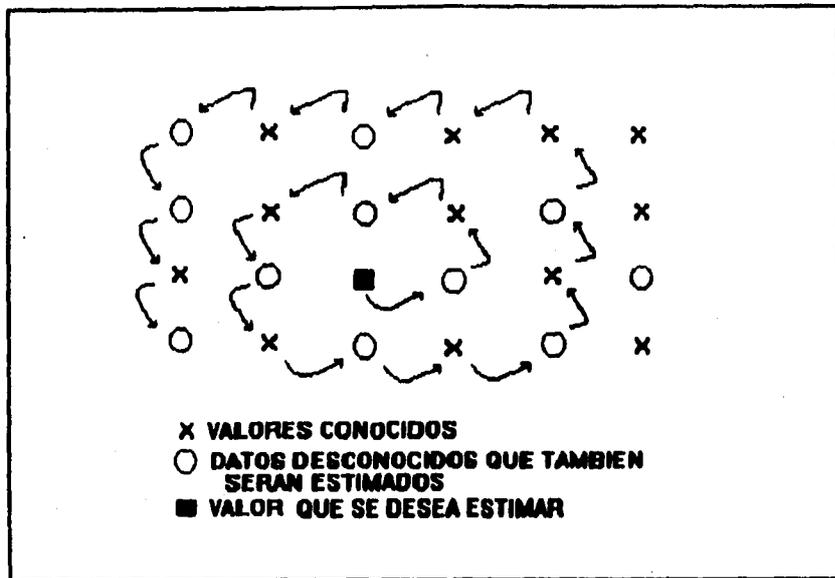
MATRIZ QUE MUESTRA LAS VARIABLES INICIALIZADAS (EL ARCHIVO DAT_MAP YA ESTA VACIADO SOBRE ELLA)

Con la información dispuesta de esta forma, el programa INTERP1.FOR se ubica en la posición (2,2) que corresponde al primer punto que potencialmente será estimado (la primera y última columna y el primero y último renglón no son tomadas en cuenta pues sirven sólo como frontera de la matriz) a partir de ahí se hace una búsqueda en forma de cruz, es decir, se busca, el primer valor disponible ($\neq 0$) hacia arriba, el primero hacia abajo, el primero hacia la izquierda y el primero hacia la derecha y se almacenan en variables auxiliares; simultáneamente a la identificación de un valor conocido se calcula la distancia a él. Estos números nos permiten poder calcular el inverso de las distancias al cuadrado, la suma de los cocientes entre los valores encontrados y las distancias al cuadrado y el cociente entre la suma anterior y la suma de los inversos de las distancias al cuadrado y todo ello nos permite estimar el valor para la variable (2,2). Posteriormente, pasamos a la posición (2,3) y repetimos el procedimiento y así sucesivamente hasta recorrer las 20×33 variables. Cuando en alguna "entrada" de la matriz existe un número $\neq 0$, esto es, un valor conocido, es ignorada y se pasa a la siguiente.

Esta forma de buscar información, puede ser, en muchos problemas generadora de resultados satisfactorios, sin embargo, para nuestros propósitos la exactitud es algo fundamental y es fácil descubrir configuraciones de datos en las que información relevante no es tomada en cuenta; necesitamos por lo tanto un algoritmo que, independientemente de la distribución de la información sea capaz de capturar todos los puntos que circunden a la variable que se desea estimar. El segundo programa de interpolación INTERP2.FOR resuelve este problema contundentemente pues utiliza para la búsqueda un algoritmo que recorre la matriz en forma de espiral alrededor de una casilla específica.



- VALOR QUE SE DESEA ESTIMAR
- X VALORES CONOCIDOS
- O DATOS DESCONOCIDOS QUE TAMBIEN SERAN ESTIMADOS



Con la búsqueda en espiral, encontramos nueve valores conocidos, mientras que con el algoritmo de cruz sólo dos.

Aparentemente, la diferencia entre ambos programas es mínima, sin embargo, es fundamental y muy significativa para la exactitud de las estimaciones; el primero, sabemos, realiza una búsqueda en forma de cruz, esto es, a partir de la localización del punto que se desea conocer se identifican a lo más 4 puntos, mientras que en el segundo, la búsqueda es más eficiente puesto que se "mira" en todas direcciones evitando así que información valiosa sea ignorada.

Además, con el algoritmo de búsqueda en espiral es posible establecer de alguna forma "radios de interés" que para nuestro caso quedarán definidos de la siguiente forma: cuando la localización del valor a estimar corresponde a latitudes y longitudes muy grandes o muy pequeñas (rangones y columnas ≤ 5 o rangones ≥ 15 y columnas ≥ 28) se propone un "radio grande" (valor mayor de un contador que se traduce en obtener aproximadamente 5 valores conocidos, alrededor del punto en cuestión) y cuando la localización nos conduce a una posición intermedia el radio de búsqueda se hace más pequeño sin que exista una disminución en el número de puntos circundantes encontrados. Naturalmente, se pueden definir más rangos para radios de búsqueda y precisar sus dimensiones con mayor claridad.

Es importante aclarar que para estimar una casilla de la matriz dentro del mismo algoritmo de búsqueda en espiral pueden establecerse diversas variantes, algunos son las siguientes:

- Utilizar toda la información cercana incluyendo también a los puntos que previamente hayan sido estimados.

- Omitir los valores que previamente fueron estimados obligando a que el algoritmo no tome en cuenta las casillas que se ubican en los renglones y columnas menores a los de la celda que en ese momento queremos estimar.

- Utilizar exclusivamente la información inicial.

Intuitivamente pareciera ser que la primera alternativa es la mejor y la segunda la menos favorable, no obstante numerosas pruebas nos han mostrado y tendremos oportunidad de descubrir que la mejor es la tercera alternativa.

Desde nuestro punto de vista, fundado no sólo en la intuición sino en los resultados obtenidos en varias de las pruebas y comparaciones realizadas, creemos que la búsqueda en espiral es mejor que la búsqueda en forma de cruz y por tanto es la que utilizaremos para el desarrollo ulterior de nuestra aplicación. Más adelante tendremos oportunidad de comparar resultados numéricos y gráficos obtenidos con ambos enfoques.

De esta forma, cada que el programa INTERP2.FOR estime un valor se lo asigna a su correspondiente entrada en la matriz y por tanto, después de 20x33 (menos a lo más 79 localidades correspondientes a los datos conocidos) repeticiones del código que selecciona variable, busca datos conocidos, realiza operaciones y asigna al arreglo bidimensional obtendremos una matriz completamente llena de valores estimados (menos a lo más 79 conocidos).

La representación de la información de esta manera compacta, puede desdoblarse y presentarse, por ejemplo, en forma de liste acompañada por datos adicionales que permitan identificar claramente la posición para cada valor estimado. Un resultado con estas características lo constituye el archivo PRO_INT el cual contiene 20x33 renglones y 3 columnas: ubicación geográfica (latitud y longitud) y valor estimado

Dentro del escenario del ejemplo planteado, estos números representan la precipitación pronosticada para los puntos ubicados en las intersecciones de 5 en 5 grados de latitud y longitud, sin embargo, aunque nos encontramos cerca, este no es el objetivo final al que pretendemos arribar, pues carecemos de información precisa en el interior de las regiones determinadas por la intersecciones, y no se olvide que lo que nuestras necesidades reclaman es la determinación, lo más exacta posible, de información en cualquier punto de la República Mexicana.

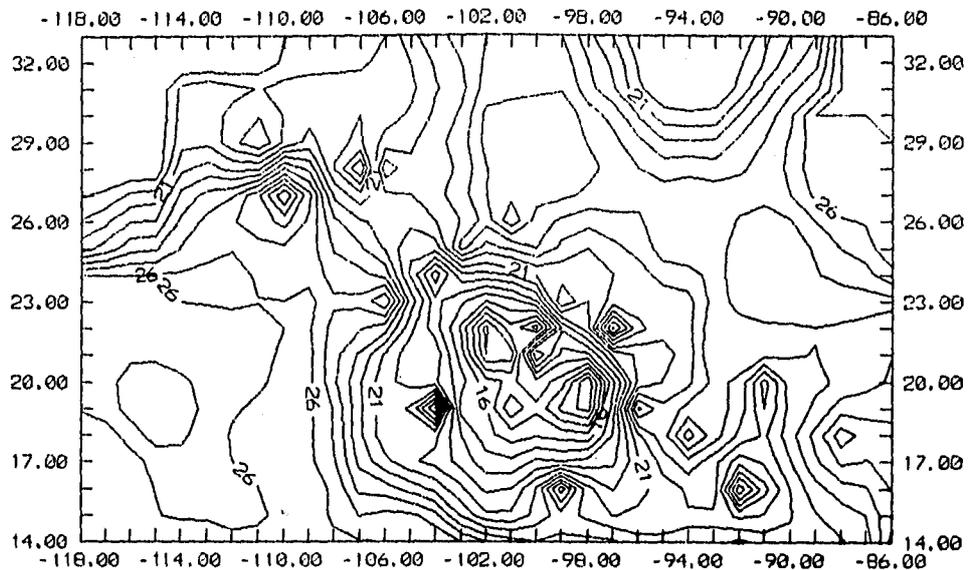
A partir de esta etapa, se puede avanzar por algunas líneas regidas por directrices de índole muy diverso, nosotros, con la intención de simplificar el problema, seguiremos una opción muy rápida y segura, aunque lo confiable debido a su exactitud es algo cuestionable que queda abierto y que se abordará posteriormente.

Así pues, la siguiente etapa a desarrollar, dentro de esta segunda alternativa global en la que estamos inmersos, consiste de la utilización de un programa comercial de uso específico (SURFER) el cual en términos generales está diseñado para lo siguiente:

A partir de un conjunto discreto de datos distribuidos irregularmente que pueden ser capturados directamente en él o ser importados (archivos con extensión .DAT) el programa construye una malla, con dimensiones establecidas por el usuario, en la que los datos desconocidos son estimados por alguno de 3 métodos con que cuenta: Inverso de las distancias, Kriging, Curvaturas Mínimas (no conocemos con precisión la fundamentación teórica subyacente a este último) como resultado de este proceso se obtiene un archivo con extensión .GRD que represente la malla completa con datos estimados.

El archivo GRD puede ser pasado, cuando menos, por dos procedimientos de gran importancia: el primero utilizando la información traza las curvas de nivel (isolinias) mientras que el segundo construye directamente la superficie.

Nosotros aquí utilizaremos Surfer solamente como un "trazador", es decir, no explotaremos la facilidad que tiene para interpolar datos, sino que le pasaremos ya la información interpolada que el programa INTERP2.FOR produce, para ello es indispensable que la salida del programa tenga un formato de matriz con algunos encabezados que le indiquen a SURFER que se trata de un archivo .GRD y que le indiquen además el tamaño de la malla y los mínimos y máximos en X, Y y Z, lo cual se consigue utilizando como archivo de entrada a PRO_INT para el programa DOGRD.FOR; todo lo anterior debido a que nuestro objetivo en este momento es la obtención de isolinias a partir de los resultados proporcionados por INTERP2.FOR, y el archivo de resultados, PRO_INT, es una colección discreta de 20x33 datos con la que, claramente, no es posible dibujarlas. No obstante, siguiendo este procedimiento, finalmente será posible entrar a las opciones de trazado de isolinias que nos permitan conocer información en cualquier punto que se desee, lo cual de alguna u otra forma puede pensarse como un pronóstico.



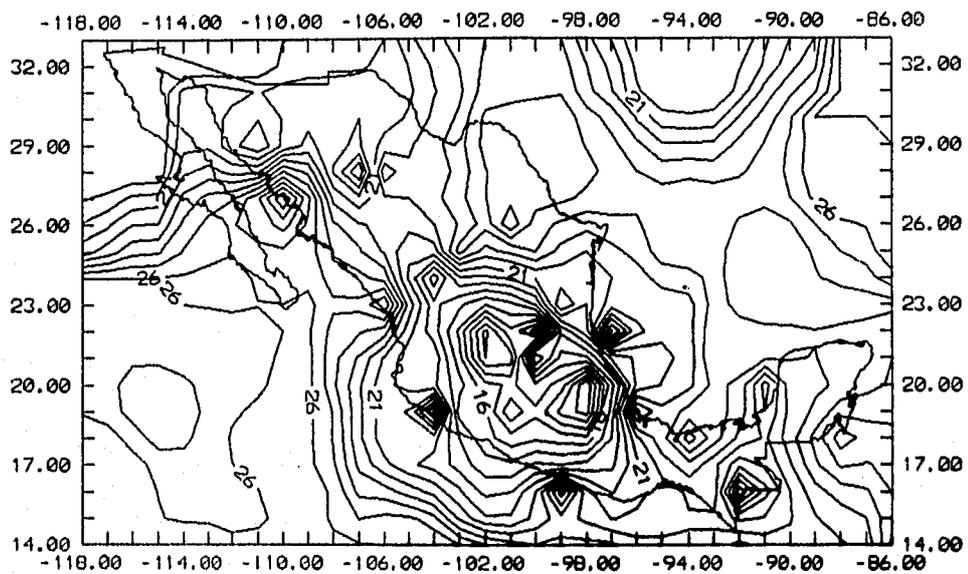
Lo más importante a resaltar desde nuestro punto de vista es que los valores interpolados (pag. 68) son "razonablemente congruentes" con la naturaleza de la información inicial, la cual, debido a que así se diseñó el programa, jamás se pierde y eso para la aplicación que estamos desarrollando es algo fundamental, en otras palabras, notemos que no sólo las estimaciones obtenidas son consistentes sino que además su emplazamiento o ubicación geográfica también es bastante convincente.

Su interpretación, por otro lado, es algo en lo que no haremos mucho énfasis, pues además de que ya se bosquejó una explicación, es algo mecánico y fácilmente comprensible. Observemos sólo lo siguiente:

La primera casilla que se estimó de la matriz (pag. 64) es la (2,2), que le llamamos (1,1) para simplificar pues los renglones y columnas de la frontera (renglones 1 y 22, columnas 1 y 35) son utilizados como puntos de referencia que nos indican hacia donde ya no existen datos, y para ella no existen datos disponibles en cientos de kilómetros, la estación más cercana (Isla Guadalupe) se encuentra a 4 grados de latitud sur (aproximadamente 450 Km.) y la siguiente más próxima es más de 1200 Km. de distancia, sin embargo, el algoritmo implementado en el programa de interpolación le asigna el mismo valor que reportó Isla Guadalupe y en este caso creemos, esto representa un acierto.

En alguna otra región, matriz (6,15) por ejemplo, donde tenemos cerca varias estaciones obtendremos también una estimación bastante convincente.

De esta forma, como respuesta al planteamiento original, proporcionamos dos herramientas (el mapa como interpretación gráfica de la matriz) mediante las cuales es posible dilucidar plenamente todas las interrogantes establecidas inicialmente.



ANALISIS

A simple vista, una comparación entre ambas matrices (la producida por INTERP2.FOR y la generada por SURFER) y ambos mapas de isolineas nos proporciona resultados muy satisfactorios pues existe una "compatibilidad inmediata" en los trazos de los mapas y en los valores de las matrices. Para hacer más evidente esta "semejanza" como primera instancia podríamos recurrir a una inspección visual más detallada a través de la cual nos fuera posible descubrir lo análogo en la forma de los contornos de subregiones delimitadas por varios pares de coordenadas.

Sin embargo, no hay que olvidar que un corrimiento de estos núcleos de isolineas en a los más un grado y hacia la izquierda se debe al truncamiento que en la SEGUNDA ALTERNATIVA se hace del valor de las coordenadas geográficas para poder referenciarlas en un arreglo bidimensional junto con la información conocida.

Es claro, según hemos aprendido, que una interpretación geométrica no es suficiente como argumento formal sobre el cual puedan erigirse nuevas ideas, es indispensable una justificación más seria en la que lo geométrico no juegue un papel central.

Para ello haremos un análisis de los datos estimados utilizando su representación matricial. Consideraremos cada una de las casillas en ambas matrices y calcularemos el valor absoluto de las diferencias de una con su correspondiente análoga y de esta forma obtendremos 20x33 nuevos numeros que nos indicarán las variaciones que cada una de las variables de una matriz guarda con sus equivalentes en la otra, y, posteriormente, promediaremos tales diferencias a fin de conocer la naturaleza de la tendencia central. Por ejemplo:

	CASILLAS		
	(2,2), ..., (12,17), ... (20,31)		
Matriz generada por INTERP2.FOR	18.0	17.3	22.9
Matriz generada por SURFER	18.0	11.8	22.8
DIFERENCIAS	0.0	5.5	0.1

lo cual nos genera la siguiente tabla:

Todo lo anterior, nos parece, constituye un resultado muy alentador pero que bien pudo haber sido producto de la casualidad; qué nos garantiza que nuestros resultados son consistentes con la realidad, cómo saber si nuestras estimaciones no presentan desviaciones bruscas de los valores verdaderos. ¿Bastará con formular un par de ejemplos similares, analizarlos y entonces concluir en función de lo obtenido de lo bondadoso o deficiente de nuestro método?, creemos que no, creemos que ni 5, ni 10, ni 20 ejemplos son suficientes para decidir si las ideas que desarrollamos son correctas o no, y aunque quizás para algún otro tipo de aplicación sí sería suficiente, el hecho es que aquí, debido a la distribución de los puntos conocidos (los observatorios no se mueven de lugar) no basta cualquier número de ejemplos pues los resultados serán siempre esencialmente los mismos que los que anteriormente hemos analizado, y por tanto, obtendríamos un punto de vista sesgado creyendo que nuestro algoritmo, programas, etc. invariablemente producen resultados con una confiabilidad superior al 89%.

Ahora bien, no debemos de pasar por alto que estamos utilizando como "información patrón" los resultados de un programa que nada nos garantiza que esté bien construido, y así, es posible que nuestros resultados sean más correctos que en un 89% o, por otro lado, que ambos, nosotros y SURFER, estemos proporcionando estimaciones erróneas. Más adelante abordaremos algunos puntos que cuestionan la confiabilidad del programa SURFER, por ahora baste lo anterior como motivación de la discusión referente a la confiabilidad de nuestro trabajo y a las limitaciones del programa SURFER.

DISCUSION **(calibración del algoritmo)**

En esta sección abordaremos, mediante el tratamiento de un par de argumentos, el problema de la confiabilidad no sólo de nuestro trabajo sino también del programa comercial SURFER.

creemos que las justificaciones que hemos esgrimido acerca de lo limitada que resulta ser la ayuda proporcionada por los ejemplos particulares en el intento de verificar la confiabilidad y calidad de nuestro trabajo han sido suficientes, por un lado, pero además, nos han incitado a la utilización de alternativas más formales.

Básicamente, la opción a desarrollar está constituida por 2 etapas en las que existe un incremento gradual en la complejidad del argumento empleado y por tanto una mayor seguridad para identificar los alcances de nuestro trabajo.

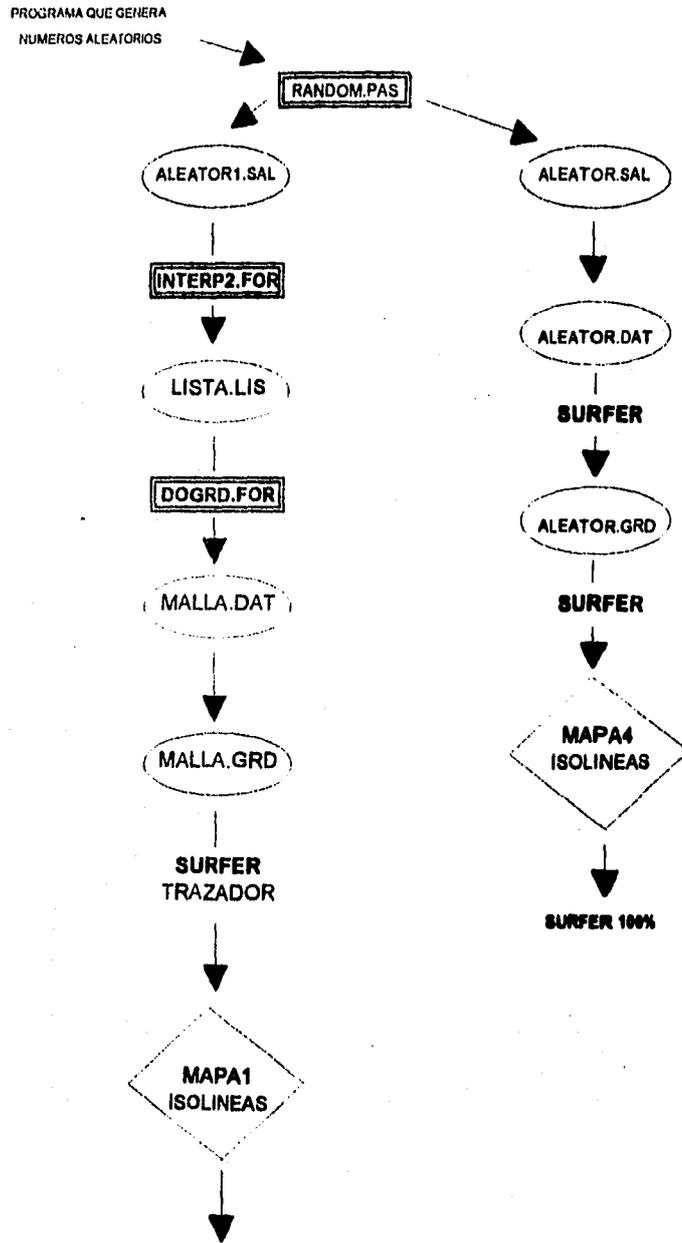
En la primera etapa tomaremos 3 funciones con un comportamiento perfectamente conocido (un paraboloides, un hiperboloides de dos ramas y un paraboloides hiperbólico) y utilizando los patrones base que hemos implementado, tamaño de la malla, algoritmo de interpolación, forma de búsqueda, etc., generaremos números aleatorios renglones ϵ [2,21], columnas ϵ [2,34], que nos permitan ubicar en forma imparcial, por ejemplo 80 puntos a partir de los cuales con ayuda del programa de interpolación INTERP2.FOR construiremos la malla de datos estimados con los que dibujaremos las curvas de nivel que nos ayudarán a decidir mediante la comparación con algo conocido y verdadero sobre lo bueno o erróneo de nuestro algoritmo y su ambiente.

Es decir, cuando trabajemos por ejemplo el conjunto de datos discretos de un paraboloides esperaremos, después del proceso de interpolación obtener circunferencias concéntricas, etc.

En caso de obtener resultados aceptables, para la segunda etapa emplearemos funciones cuyo comportamiento no es tan suave y regular como el de las 3 anteriores, dichas funciones serán extraídas del artículo "A Critical Comparison of Some Methods for Interpolation of Scattered Data" publicado por Richard Franke en 1979 y cuyo contenido se refiere precisamente a una serie de técnicas empleadas para poner a prueba la eficiencia de varios de los métodos de interpolación existentes hasta esa fecha, por tanto las funciones que se utilizarán serán funciones de un comportamiento complicado creadas justamente para "poner en jaque" a los métodos de interpolación y hacer evidentes sus limitaciones.

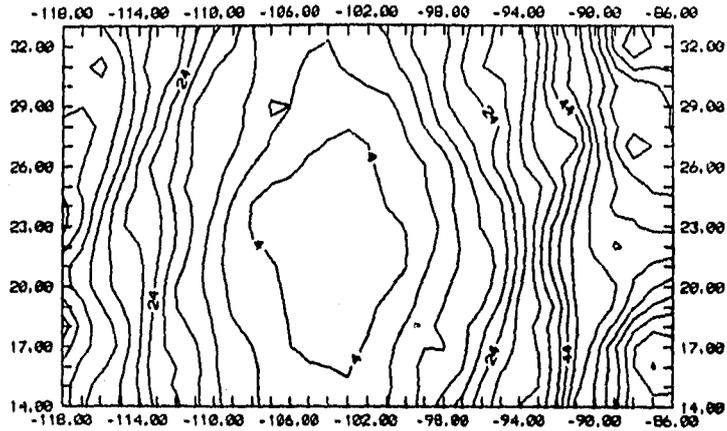
Sobra decir que, en caso de que nuestro algoritmo arroje buenos resultados en ambas etapas del proceso de prueba tendremos plena confianza para su uso y habremos sentado las bases para el desarrollo de aplicaciones que rebasarán por mucho la utilidad que a veces se encuentra en el empleo de los programas comerciales.

FLUJO DE LA INFORMACION UTILIZADA PARA LAS PRUEBAS DE EFICIENCIA DE NUESTRO ALGORITMO

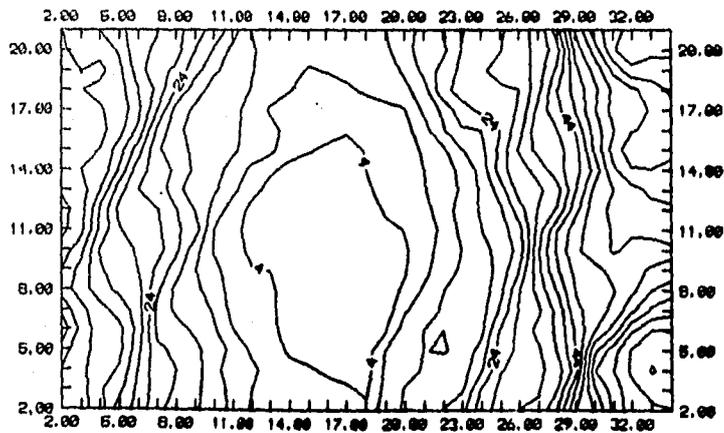


ETAPA 1

FUNCION 1: Parabolode Elíptico: $Z = (X-16)^2 / 4 + (Y-10)^2 / 9$



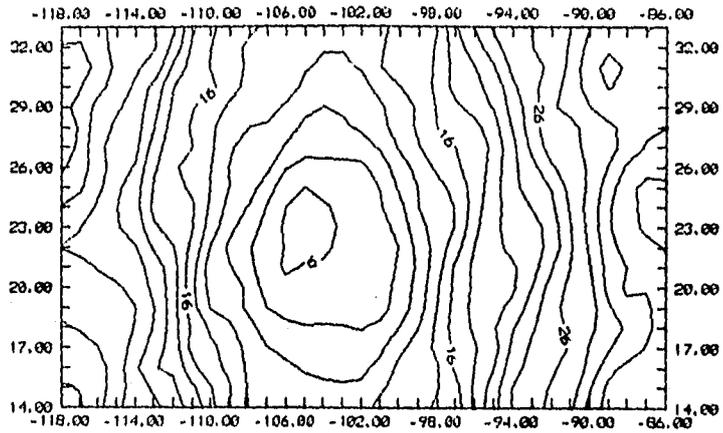
MAPA GENERADO A PARTIR DE LOS DATOS INTERPOLADOS POR NUESTRO PROGRAMA



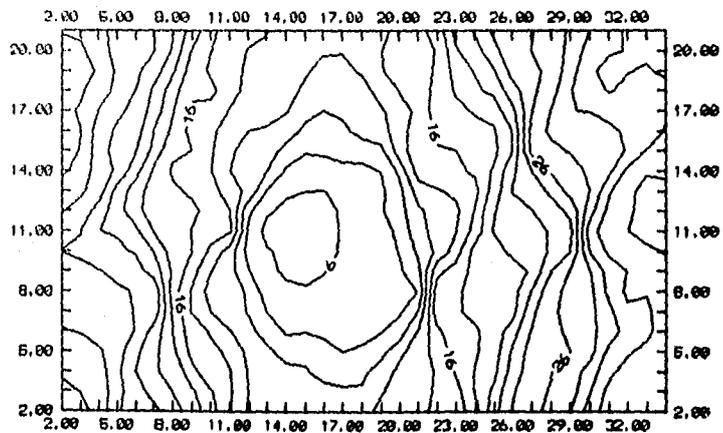
MAPA GENERADO A PARTIR DE LAS ESTIMACIONES HECHAS POR SURFER

ESTA TESIS NO DEBE SALIR DE LA BIBLIOTECA.

FUNCION 2: Elipsoide de 2 ramas: $Z = 4 \text{ SQRT}(1 + (X-10)^2 / 4 + (Y-10)^2 / 9)$

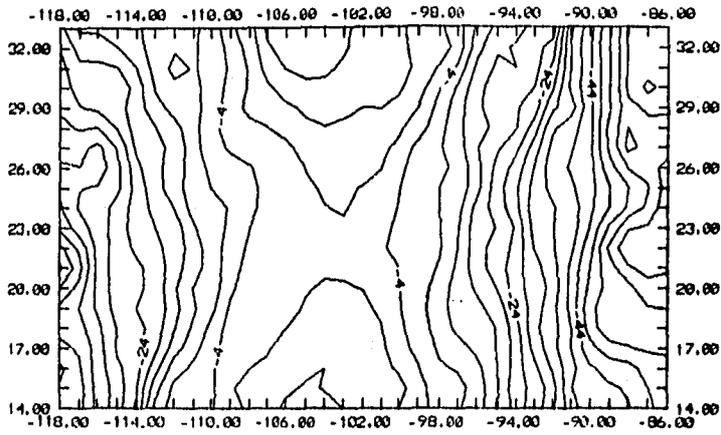


**MAPA GENERADO A PARTIR DE LOS
DATOS INTERPOLADOS POR NUESTRO PROGRAMA**

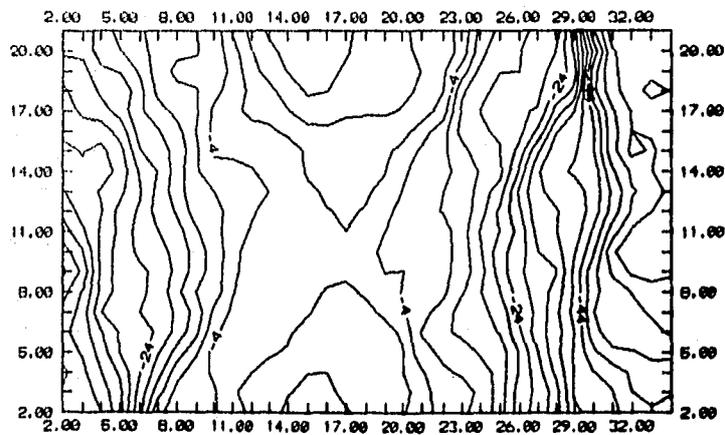


**MAPA GENERADO A PARTIR
DE LAS ESTIMACIONES HECHAS POR SURFER**

FUNCION 3: Paraboloide Hiperbólico : $Z = -(X-16)^2 / 4 + (Y-10)^2 / 9$



**MAPA GENERADO A PARTIR DE LOS
DATOS INTERPOLADOS POR NUESTRO PROGRAMA**



**MAPA GENERADO A PARTIR
DE LAS ESTIMACIONES HECHAS POR SURFER**

Si en los mapas anteriores nos limitamos a hacer un "análisis visual" concluiremos fácilmente que los resultados obtenidos en esta etapa de prueba son lo suficientemente adecuados como para, sin ninguna objeción, pasar al siguiente nivel en el proceso de calibración de nuestro algoritmo, sin embargo, pretendemos hacer más natural este avance y para ello observaremos también la naturaleza de las estimaciones numéricas que en cada uno de los 3 casos se han hecho. Consideremos por ejemplo la función número 2, con ella y a través de la metodología descrita en la sección denominada ANALISIS es posible obtener una tabla como la siguiente:

```

00 00 00 00 00 00 00 00 00 80 08 80 00 00 00 88 00 88 08 00 00 80 00 00 00 00 00 00 00 80 00 80 00 00
00 00 03 05 02 00 02 02 00 00 00 00 02 01 00 00 01 01 02 00 02 06 00 09 00 00 00 03 04 02 01 08 05 03 00
00 03 00 05 04 02 03 04 00 00 00 00 01 00 01 00 00 01 02 01 03 03 08 00 04 02 00 00 02 03 01 00 06 07 00
00 04 00 03 08 04 00 00 00 02 04 00 08 02 03 01 00 00 05 03 00 02 00 02 06 02 03 01 04 04 02 04 12 09 00
00 00 00 02 02 06 06 05 00 00 02 00 04 01 07 00 00 05 02 01 02 02 00 00 02 00 05 00 01 10 09 06 05 00 00
00 01 00 00 02 00 06 10 07 01 04 00 01 07 01 05 03 04 04 09 02 09 05 05 09 07 04 01 01 00 02 06 00 00 00
00 00 01 00 00 01 12 12 07 00 01 02 00 03 06 04 02 00 04 09 10 32 11 01 88 14 06 02 08 01 01 03 00 00 80
00 02 00 00 02 04 20 18 05 03 00 00 00 03 00 01 00 04 11 28 34 13 02 00 05 01 05 00 82 01 00 00 00 00
00 00 00 02 01 01 00 11 04 09 05 05 02 08 03 02 01 02 00 05 11 23 65 03 00 08 10 02 01 01 03 00 03 00 00
00 00 00 00 00 04 01 12 07 15 13 11 04 04 04 04 00 00 00 01 04 84 02 10 02 00 08 85 17 17 13 05 03 00 00
00 02 00 09 04 08 05 01 01 18 36 14 17 85 80 84 08 01 00 08 02 00 81 04 03 80 07 04 26 26 17 05 00 01 88
00 02 02 08 13 28 08 00 02 17 28 14 18 05 00 03 02 02 01 04 84 80 04 88 05 00 02 83 16 84 18 03 08 02 88
00 00 00 08 11 19 07 02 00 08 09 08 89 04 00 05 84 80 84 18 09 05 83 14 03 85 80 83 88 83 87 82 00 81 00
00 02 00 04 08 07 07 04 00 01 00 04 04 05 88 05 02 04 08 07 22 89 00 05 85 01 05 08 80 81 05 01 83 81 80
00 00 03 05 00 04 03 01 00 01 00 00 00 87 07 05 00 01 08 08 14 84 00 05 11 15 18 05 14 08 01 88 02 05 00
00 00 02 84 00 01 00 06 04 02 01 00 01 03 02 08 02 03 02 84 07 03 80 04 15 22 18 08 23 00 02 01 88 02 80
00 01 03 07 02 02 05 07 00 02 02 01 03 03 02 80 00 01 00 01 03 08 04 05 13 14 06 05 08 01 00 01 05 81 08
00 02 02 04 05 03 82 08 02 00 00 00 00 80 08 03 01 00 01 01 03 00 04 00 01 08 08 01 10 04 00 02 02 00 80
00 02 00 03 02 00 00 08 00 02 01 81 00 01 01 83 01 80 81 00 00 08 01 02 00 00 00 81 04 03 08 83 02 11 88
00 04 81 03 01 00 00 05 01 00 00 02 02 01 01 01 01 00 02 01 03 08 88 80 00 81 00 02 03 01 01 05 86 07 00
00 02 02 00 00 03 00 00 09 00 00 00 03 00 81 02 00 08 00 81 85 08 09 80 00 08 02 03 04 00 81 15 04 05 08
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 80 00 08 00 00 00 00 00 80 08 00 00 88 80 00 88 00 00 00

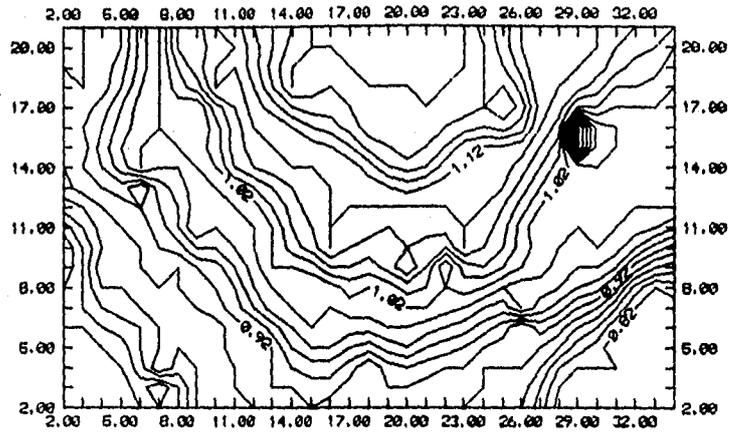
```

VALOR PROMEDIO DE LAS DIFERENCIAS	0.365
DIFERENCIA MAXIMA	3.648
DIFERENCIA MINIMA	0.0
NUMERO DE VECES QUE LA DIFERENCIA > 2	11
NUMERO DE VECES QUE LA DIFERENCIA > 3	3

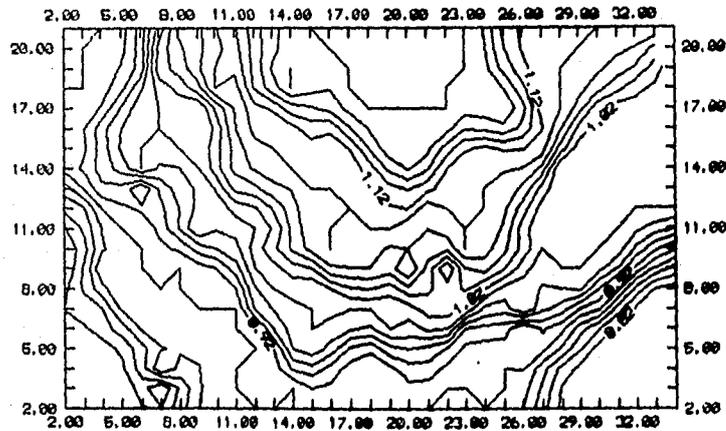
Y por tanto, con entera confianza podemos avanzar a la siguiente etapa.

ETAPA 2

FUNCION 1: $0.75 \exp(-((9x-2)^2 + (9y-2)^2)/4) +$
 $0.75 \exp(-((9x+1)^2/49 - (9y+1)/10)) +$
 $0.50 \exp(-((9x-7)^2 + (9y-3)^2)/4) -$
 $0.20 \exp(-((9x-4)^2 - (9y-7)^2)) = Z.$

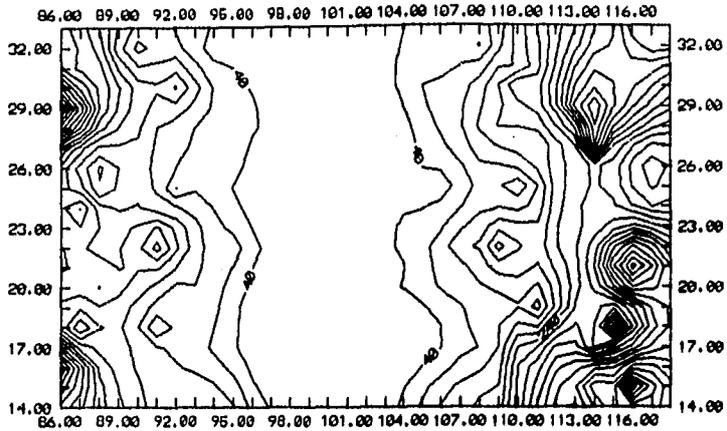


MAPA GENERADO A PARTIR DE LOS
DATOS INTERPOLADOS POR NUESTRO PROGRAMA

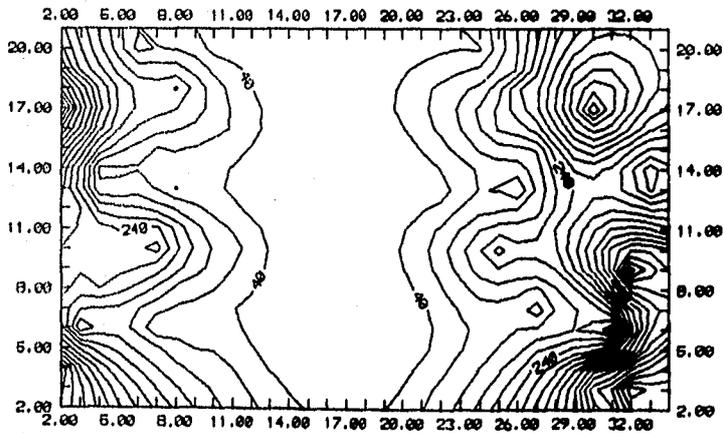


MAPA GENERADO A PARTIR
DE LAS ESTIMACIONES HECHAS POR SURFER

FUNCION 2: $Z = (1.25 + \text{Cos}(5.4y)) / (6(1+(3x-1)^2))$

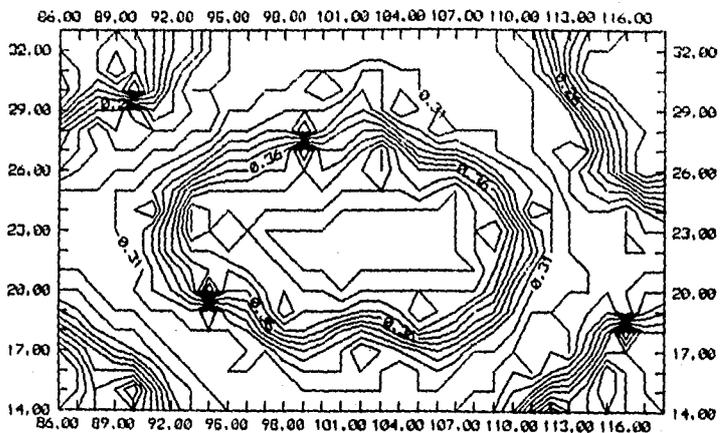


**MAPA GENERADO A PARTIR DE LOS
DATOS INTERPOLADOS POR NUESTRO PROGRAMA**

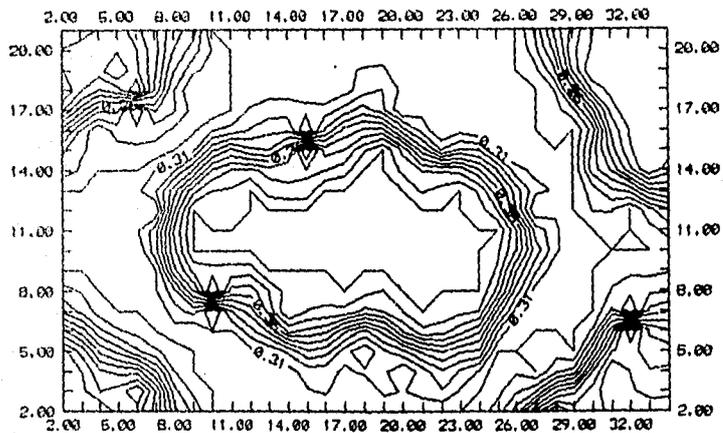


**MAPA GENERADO A PARTIR
DE LAS ESTIMACIONES HECHAS POR SURFER**

FUNCION 3: $Z = (1/9) \text{SQRT}(64 - 81((x-0.5)^2 + (y-0.5)^2)) - 0.5$



MAPA GENERADO A PARTIR DE LOS DATOS INTERPOLADOS POR NUESTRO PROGRAMA



MAPA GENERADO A PARTIR DE LAS ESTIMACIONES HECHAS POR SURFER

i) No se pierden jamás los datos conocidos, es decir, si en el campo inicial la casilla i,j tiene el valor Z_0 , éste prevalece en el campo final (matriz llena de valores estimados) mientras que Surfer no lo garantiza pues en algunas ocasiones los conserva y en otras, que son las más de los casos, los pierde.

Nótese que este argumento no fue utilizado en el momento de considerar las matrices de diferencias y que de haberlo hecho el porcentaje de "confiabilidad y eficiencia" de nuestro trabajo, claramente habría aumentado.

ii) Es relativamente sencillo implementar en nuestro programa las restricciones orográficas que, en la práctica, son de gran importancia; mientras que Surfer es un programa cerrado, obscuro, que nos impide conocer y mucho menos modificar, varios de los procesos que contiene.

Para aclarar más este punto, considerese lo siguiente: Si en la falda norte de una montaña existe una temperatura T_0 y en la falda sur (punto opuesto) existe el mismo valor, ¿Qué es lo que hace Surfer al trazar las isoclinas?, ¿Qué temperatura existe en la cima de la montaña?, ¿Es correcto que Surfer atraviese la montaña de un lugar a otro?

Debemos estar seguros, finalmente, de que los resultados obtenidos, tomando como punto de referencia a Surfer-Inverso de las Distancias que, a pesar de las limitaciones ya antes mencionadas, es un programa esencialmente correcto (pues tuvimos oportunidad de contrastarlo con opciones más exactas y confiables), no son producto de la casualidad sino de una síntesis de técnicas e ideas correctas.

CONCLUSION

En este apartado retomaremos un par de comentarios que a lo largo del contenido presentado quedaron incompletos, de igual forma, bosquejaremos algunas ideas generales que nos proporcionarán una visión global del ambiente en el que está inmerso nuestro trabajo y las expectativas que tiene considerado globalmente o en forma parcial a través de la selección y desarrollo "más a fondo" de algunos de los tópicos mostrados.

Es necesario recalcar, adulando a la insistencia, que los algoritmos, producto de la abstracción de la teoría, y los programas, resultado de la codificación de los primeros nos proporcionaron tanto para la distribución de los datos reales como para las funciones prueba en las etapas de calibración resultados muy satisfactorios no obstante que la teoría subyacente a las directrices constituye tan sólo un conjunto de suposiciones y simplificaciones para el caso más sencillo. Lo que pretendemos hacer notar es que las ecuaciones que intervienen en el programa se pueden mejorar notablemente con el consecuente beneficio en los resultados.

En este mismo sentido hay que agregar que los programas que hemos desarrollado tienen la ventaja adicional, sobre los productos comerciales, de ser susceptibles de modificación en un aspecto que para el caso que fueron desarrollados es fundamental: *El comportamiento frente a la Orografía*. Este problema en términos muy generales ya fue descrito en la sección denominada DISCUSION, pero brevemente se refiere a la imposibilidad que deben de tener las isolneas para atravesar longitudinalmente una montaña, un volcán, una depresión o cualquier factor orográfico considerable pues a medida que se va avanzando a través de ellas hasta alcanzar sus cimas o sus valles las condiciones meteorológicas van cambiando considerablemente, y por tanto, es incorrecto que a partir de dos datos iguales y opuestos en las faldas de una montaña se pretenda trazar mecánicamente una línea muy parecida a una recta; lo correcto sería trazar una línea que rodee la montaña.

Por otro lado, hay que reconocer que actualmente la teoría que hemos mostrado constituye únicamente los cimientos, o motivación quizás, de notables investigaciones que se han desarrollado con la intención de resolver el mismo problema pero de una manera más eficiente, donde la eficiencia radica en la exactitud que se pretende, en el número de variables y situaciones que se desea involucrar, en el alcance en unidades de tiempo de los resultados logrados, etc., y que en este momento en forma automatizada se explotan en los más grandes y poderosos centros de pronóstico meteorológico del mundo entero.

Es bueno mencionar en este momento, continuando con la intención de mostrar la situación real de nuestro trabajo, que de acuerdo a clasificaciones actuales las ideas que hemos desarrollado, quedan englobadas dentro de una etapa de lo que formalmente se

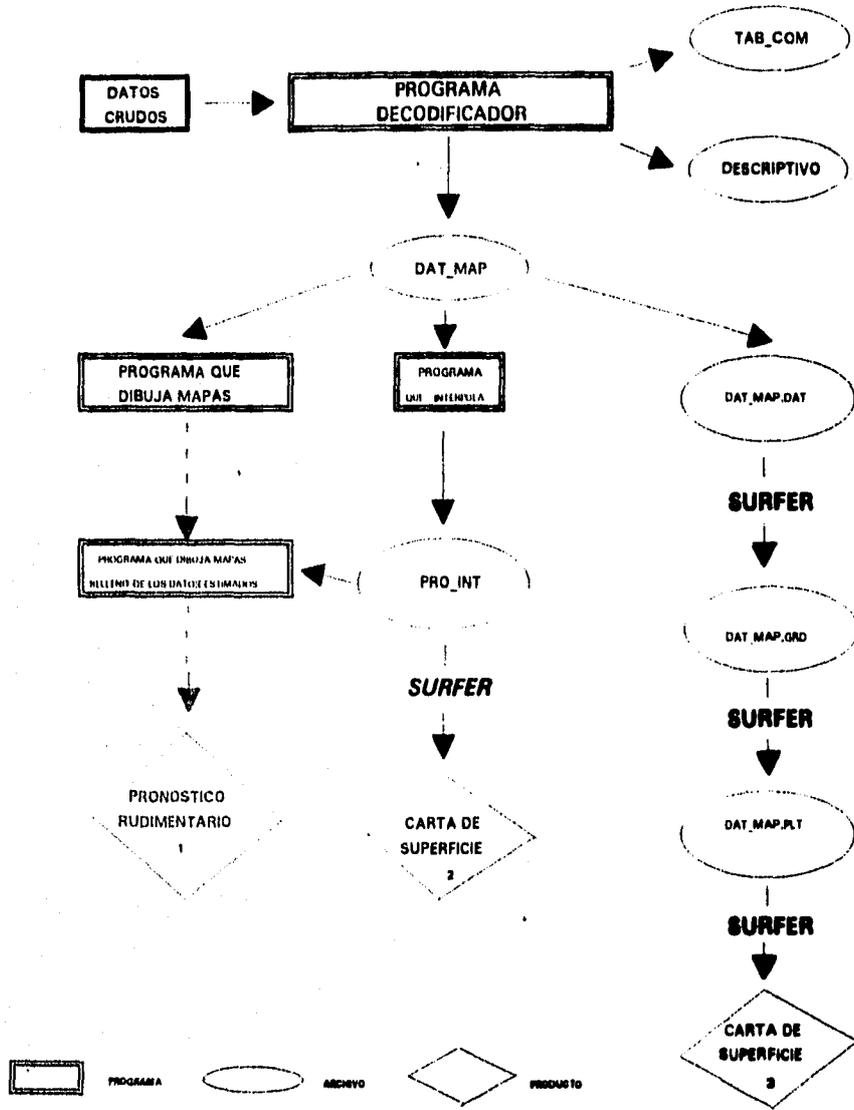
conoce como **ANALISIS OBJETIVO** que a su vez forma parte de las Técnicas de Análisis Meteorológico el cual tiene como tarea básica la reproducción a partir de un conjunto de observaciones de el estado verdadero de la atmósfera.

El propósito del Análisis Objetivo radica en obtener la descripción más exacta de un conjunto de datos mediante el uso de un conjunto de instrucciones que pueden ser igualmente aplicables a cualquier otro conjunto de datos. En la práctica, las instrucciones se encuentran en forma de programas de cómputo que corren en una computadora digital.

El uso de los procedimientos de análisis objetivo se ha convertido en algo ordinario en los Servicios Meteorológicos de los países desarrollados en estos últimos años, mientras que las técnicas manuales de análisis (**ANALISIS SUBJETIVO**) permanecen como las formas comunes en los Servicios Meteorológicos de los países pobres.

Así pues, ubicándonos en un escenario real que nos permita contar con un punto de vista global y objetivo, tendremos que darnos cuenta de que el trabajo que hemos desarrollado constituye tan sólo una síntesis de ideas muy elementales pero que a pesar de todo, o quizás precisamente por ello, según la posición en la que se esté, tiene su valor teórico y práctico.

ALTERNATIVAS PARA EL FLUJO DE LA INFORMACION HASTA LA OBTENCION DE UN PRODUCTO



NO DEBE SALIR DE LA BIBLIOTECA

 *PROGRAMA DE DECODIFICACION (DECODI.FOR): LEE DATOS DE LOS ARCHIVOS DE
 * INFORMACION METEOROLOGICA SINOPTICA ACTUAL O DE LOS ARCHIVOS DE
 * RESPALDO CON INFORMACION SINOPTICA HISTORICA. LA SALIDA CONSISTE DE TRES
 * ARCHIVOS 2 CON OBJETIVOS AJENOS A LA APLICACION QUE ESTAMOS DESARROLLANDO
 * Y OTRO, DAT MAP, CON LA INFORMACION INDISPENSABLE PARA LA CONSECUCION DE
 * DE NUESTRO OBJETIVO FINAL.
 *

* EL PROGRAMA DECOFICA TODOS LOS GRUPOS DE LOS MENSAJES CON FORMATO FM12-IX
 * PERO ESTA HABILITADO PARA CREAR DATA MAP UNICAMENTE CON INFORMACION DE
 * TEMPERATURA AMBIENTE, PRESION ATMOSFERICA O PRECIPITACION (LLUVIA)
 *

```

* INCLUDE 'FGRAPH.PI'
* INCLUDE 'FGRAPH.FD'
REAL*4      MEDI,MEDIA,VARIO,VARI,LAT,LON
REAL       VARIA10,SUBTOTAL1
INTEGER    ESTACION,Z,ELECCION
INTEGER*1  BANDERA,HORA,DIA
INTEGER*2  L,B,REGLON,CEN,DEC,UNI,DUMMY2,DUMMY
INTEGER*2  CENTENA,DECENA,UNIDAD,DMIL,DMILES,REN
INTEGER*2  CONTADOR,NOCERO,MIL,MILES,APUN,JDIA
INTEGER*4  COLUMNNA,ITAD5,ITRD5,ITXD5,MED10,ITND5
INTEGER*4  NMS,D1,NQ,SUBTOTAL,IPSD6,IPMD6,IPRD5
CHARACTER*1 SG,PUNTO,N,LIN,DR,DRS,RESP
CHARACTER*1 GUION,CERO,CREG
CHARACTER*2 ANIO,NMES,CHH,CDD,CHH1
CHARACTER*3 NSTN,ANIOB,NMESB,CDOB
CHARACTER*4 AA
CHARACTER*5 TAD5,TRD5,TKD5,TND5,PRD5
CHARACTER*6 PSD6,PMD6
CHARACTER*7 INICIA1,INICIA2,RESET
CHARACTER*8 DIRECCION,INTENSIDAD,DRCTOR
CHARACTER*10 INFORME
CHARACTER*(*) SP
PARAMETER (SP=' ')
CHARACTER*15 ARCH S,INICIAS
CHARACTER*31 NOMBRE
  
```

```

CHARACTER*35 ARCH_SAL
CHARACTER*30 ARCH_DATOS,MES
DIMENSION LIN(85)
DIMENSION LAT(80)
DIMENSION LON(80)
DIMENSION INFORME(80,10)
DIMENSION NOMBRE(79)
DIMENSION ESTACION(80,310)
DIMENSION MEDIA(79)
DIMENSION VARIANZA(79)
RECORD /xycoord/ xy

```

```

*
*
*****
* INICIALIZAMOS ALGUNAS VARIABLES QUE SE UTILIZARAN A LO LARGO DEL
* PROGRAMA
*****
*

```

```

L=-1
GUION='-'
CERO='0'
NP=0
NSTN='---'
SG=' '
PUNTO='.'
CREG=' '
BANDERA=0
DO 444,RENGLON=1,80
LAT(RENGLON)=0
LON(RENGLON)=0
444 CONTINUE
SUBTOTAL=0
SUBTOTAL1=0
NOCERO=0

```

```

*****
* SE LLENA UN ARREGLO UNIDIMENSIONAL DE 79 ELEMENTOS CON LOS NOMBRES DE
* LAS ESTACIONES, DEJANDO FUERA A LAS ESTACIONES:
*

```

```

*   Esmeralda, Coah.      280
*   Tamuin, S.L.P.       543
*   Islas Marias, Nay.   551
*   Isla Lobos, Ver.     570
*   Isla Contoy, Camp.   590
*

```

```

* DEBIDO A QUE NO TRANSMITEN REGULARMENTE INFORMACION.
*****

```

```

DATA NOMBRE(01) // 050 //
DATA NOMBRE(02) // 055 //
DATA NOMBRE(03) // 061 //
DATA NOMBRE(04) // 113 //
DATA NOMBRE(05) // 118 //
DATA NOMBRE(06) // 122 //
DATA NOMBRE(07) // 151 //
DATA NOMBRE(08) // 160 //
DATA NOMBRE(09) // 220 //
DATA NOMBRE(10) // 225 //
DATA NOMBRE(11) // 243 //
DATA NOMBRE(12) // 253 //
DATA NOMBRE(13) // 256 //

```

* DATA NOMBRE(14) // 258 //
 DATA NOMBRE(15) // 280 //
 DATA NOMBRE(15) // 305 //
 DATA NOMBRE(16) // 311 //
 DATA NOMBRE(17) // 323 //
 DATA NOMBRE(18) // 342 //
 DATA NOMBRE(19) // 373 //
 DATA NOMBRE(20) // 382 //
 DATA NOMBRE(21) // 390 //
 DATA NOMBRE(22) // 393 //
 DATA NOMBRE(23) // 401 //
 DATA NOMBRE(24) // 402 //
 DATA NOMBRE(25) // 405 //
 DATA NOMBRE(26) // 412 //
 DATA NOMBRE(27) // 423 //
 DATA NOMBRE(28) // 458 //
 DATA NOMBRE(29) // 471 //
 DATA NOMBRE(30) // 491 //
 DATA NOMBRE(31) // 499 //
 DATA NOMBRE(32) // 519 //
 DATA NOMBRE(33) // 525 //
 DATA NOMBRE(34) // 539 //
 * DATA NOMBRE(35) // 543 //
 DATA NOMBRE(35) // 548 //
 * DATA NOMBRE(38) // 551 //
 DATA NOMBRE(36) // 556 //
 * DATA NOMBRE(40) // 570 //
 DATA NOMBRE(37) // 571 //
 DATA NOMBRE(38) // 577 //
 DATA NOMBRE(39) // 581 //
 DATA NOMBRE(40) // 585 //
 * DATA NOMBRE(45) // 590 //
 DATA NOMBRE(41) // 593 //
 DATA NOMBRE(42) // 612 //
 DATA NOMBRE(43) // 625 //
 DATA NOMBRE(44) // 632 //
 DATA NOMBRE(45) // 634 //
 DATA NOMBRE(46) // 640 //
 DATA NOMBRE(47) // 644 //
 DATA NOMBRE(48) // 647 //
 DATA NOMBRE(49) // 648 //
 DATA NOMBRE(50) // 654 //
 DATA NOMBRE(51) // 656 //
 DATA NOMBRE(52) // 662 //
 DATA NOMBRE(53) // 665 //
 DATA NOMBRE(54) // 675 //
 DATA NOMBRE(55) // 679 //
 DATA NOMBRE(56) // 680 //
 DATA NOMBRE(57) // 683 //
 DATA NOMBRE(58) // 685 //
 DATA NOMBRE(59) // 687 //
 DATA NOMBRE(60) // 692 //
 DATA NOMBRE(61) // 695 //
 DATA NOMBRE(62) // 698 //
 DATA NOMBRE(63) // 723 //
 DATA NOMBRE(64) // 726 //
 DATA NOMBRE(65) // 737 //
 DATA NOMBRE(66) // 741 //
 DATA NOMBRE(67) // 743 //
 DATA NOMBRE(68) // 750 //

```

DATA NOMBRE(69) // 762 //
DATA NOMBRE(70) // 773 //
DATA NOMBRE(71) // 775 //
DATA NOMBRE(72) // 805 //
DATA NOMBRE(73) // 833 //
DATA NOMBRE(74) // 840 //
DATA NOMBRE(75) // 843 //
DATA NOMBRE(76) // 845 //
DATA NOMBRE(77) // 848 //
DATA NOMBRE(78) // 855 //
DATA NOMBRE(79) // 903 //
DO 1, I=1, 79

```

```

1 NOMBRE(I) (1:1) = CREG

```

```

*****
*
* EN FORMA INTERACTIVA CON EL USUARIO SE OBTIENEN EL VALOR DE LAS
* VARIABLES QUE DETERMINARAN EL NOMBRE DEL ARCHIVO DE ENTRADA, EL DE
* SALIDA, Y EL PARAMETRO CON EL QUE SE CREARA DAT_MAP

```

```

DO 399 B=1, 45
399 WRITE(*,*)
WRITE (*, '(A\)' ) ' PROPORCIONAR EL AÑO, MES Y DIA DE LOS '
WRITE (*, '(A24)' ) ' MENSAJES A DECODIFICAR'
*
WRITE(*,*)
WRITE (*, '(A\)' ) ' AÑO : (AAAA) ? '
READ (*, '(A4)' ) AA
WRITE(*,*)
WRITE (*, '(A\)' ) ' MES : (ENERO, ..., DICIEMBRE) ? '
READ (*, '(A12)' ) MES
WRITE(*,*)
WRITE (*, '(A\)' ) ' DIA : ( 01,02, ..., 30,31 ) ? '
READ (*, '(I2)' ) DIA
WRITE(*,*)
WRITE (*, '(A\)' ) ' HORA : ( 00,03, ..., 18,21 ) ? '
READ (*, '(I2)' ) HORA
*
IF (MES.EQ.'ENERO') THEN
NMES='01'
ENDIF
IF (MES.EQ.'FEBRERO') THEN
NMES='02'
ENDIF
IF (MES.EQ.'MARZO') THEN
NMES='03'
ENDIF
IF (MES.EQ.'ABRIL') THEN
NMES='04'
ENDIF
IF (MES.EQ.'MAYO') THEN
NMES='05'
ENDIF
IF (MES.EQ.'JUNIO') THEN
NMES='06'
ENDIF
IF (MES.EQ.'JULIO') THEN
NMES='07'
ENDIF
ENDIF

```

```

IF (MES.EQ.'AGOSTO') THEN
  NMES='08'
ENDIF
IF (MES.EQ.'SEPTIEMBRE') THEN
  NMES='09'
ENDIF
IF (MES.EQ.'OCTUBRE') THEN
  NMES='10'
ENDIF
IF (MES.EQ.'NOVIEMBRE') THEN
  NMES='11'
ENDIF
IF (MES.EQ.'DICIEMBRE') THEN
  NMES='12'
ENDIF
WRITE(*,*)
WRITE(*,*) ' CON QUE PARAMETRO SE TRAZARAN LOS MAPAS '
WRITE(*,*)
WRITE(*,*) '          2. TEMPERATURA AMBIENTE '
WRITE(*,*) '          7. PRESION ATMOSFERICA '
WRITE(*,*) '          1. PRECIPITACION '
WRITE(*,*(A\)) '          2,7,1 : ?
READ(*,*(I1)) ELECCION
*
WRITE(*,*)
ANIO(1:1)=AA(3:3)
ANIO(2:2)=AA(4:4)
WRITE (*,*(A\)) ' EN QUE DRIVE ESTA EL DISCO DE DATOS ?
READ (*,*(A1)) DR
*
WRITE(*,*)
WRITE (*,*(A\)) ' DRIVE DE SALIDA ?
READ (*,*(A1)) DRS
*
WRITE(*,*)
WRITE (*,*(A\)) ' EN EL DIRECTORIO RAIZ (S/N) ?
READ (*,*(A1)) RESP
*
IF ((RESP.EQ.'S').OR.(RESP.EQ.'s')) THEN
  WRITE (*,*)
  WRITE(*,*(A\)) ' NOMBRE DEL ARCHIVO DE SALIDA ?
  READ (*,*(A15)) ARCH_S
  ARCH_SAL= DRS // ':' // ARCH_S
*
  ARCH_SAL='LPT1'
ELSE
  IF ((RESP.EQ.'N').OR.(RESP.EQ.'n')) THEN
    WRITE(*,*)
    WRITE(*,*(A\)) ' NOMBRE DEL DIRECTORIO (DDDDDDDD) ?
    READ (*,*(A8)) DRCTOR
    WRITE(*,*)
    WRITE(*,*(A\)) ' NOMBRE DEL ARCHIVO DE SALIDA ?
    READ(*,*(A12)) ARCH_S
*
    ARCH_SAL= DRS // ':' // DRCTOR // '\' // ARCH_S
    ARCH_SAL='LPT1'
  ENDIF
ENDIF
PRINT 377, ARCH_SAL
*
* ABRIMOS ARCHIVOS DE SALIDA: DONDE QUEDARAN TODOS LOS DATOS

```

```

* DECODIFICADOS, TAB COM, Y DONDE QUEDARA LA LISTA DE DATOS QUE
* FUNGIRAN COMO MATERIAS PRIMAS PARA EL DESARROLLO ULTERIOR DE LA
* APLICACION
*

```

```

      IF (BANDERA.EQ.0) THEN
        BANDERA=1
        OPEN(3,FILE=ARCH_SAL,IOSTAT=K,ACCESS='SEQUENTIAL',
1FORM='FORMATTED')
        OPEN(5,FILE="C:DAT_MAP",IOSTAT=K,ACCESS='SEQUENTIAL',
1FORM='FORMATTED')
        IF (K.NE.0) GOTO 99
      ENDIF

```

```

*****

```

```

*
* INICIALIZACION DE LA IMPRESORA
*
* INICIA1=' ' // CHAR(27) // "&18D"
* INICIA2=INICIA1 // CHAR(27)
* INICIA3=INICIA2 // "(#Op10H"
* RESET =CHAR(27) // "E"
*
* WRITE(3,31) INICIA3
*
* DO 10, I=1,80
*   WRITE(7,20) I, CHAR(64+i)
*10 CONTINUE
*   WRITE(3,31) RESET
*
*20 FORMAT(I10, A10)
*31 FORMAT(80A)
*   END
*

```

```

*****

```

```

*
* LOS MENSAJES SINOPTICOS SE GUARDAN EN DIRECTORIOS CON NOMBRE
*
* SINOH
*
* LOS QUE A SU VEZ CONTIENEN ARCHIVOS DE ALMACENAMIENTO FINAL LLAMADOS
*
* SINOPDDH
*
* PARA LOS DIRECTORIOS Y ARCHIVOS HH PUEDE SER (00,03,06,09,...,18,21)
* Y PARA LOS ARCHIVOS DD PUEDE SER CUALQUIER NUMERO ENTERO MAYOR QUE
* 1 Y MENOR O IGUAL QUE 31. DE ESTA FORMA, LOS MENSAJES DE LAS 0 HRS.
* DEL DIA 19 SE ENCUENTRAN EN SIN000\SINOP1900.
*
* POR LO TANTO ES FACIL TENER ACCESO A LA INFORMACION DE, POR EJEMPLO,
* TODO UN MES CON TODAS LAS HORAS SINOPTICAS REFERENCIANDO LOS
* DIRECTORIOS Y ARCHIVOS CON CONTADORES ENTEROS
*

```

```

      DO 313, IHR=0,0
      DO 314, IDIA=1,1

      SELECT CASE (HORA)
      CASE(0)
        CHH='00'
      CASE(3)

```

```
CHH='03'  
CASE(6)  
CHH='06'  
CASE(9)  
CHH='09'  
CASE(12)  
CHH='12'  
CASE(15)  
CHH='15'  
CASE(18)  
CHH='18'  
CASE(21)  
CHH='21'  
END SELECT
```

```
SELECT CASE(DIA)
```

```
CASE(1)  
CDD='01'  
CASE(2)  
CDD='02'  
CASE(3)  
CDD='03'  
CASE(4)  
CDD='04'  
CASE(5)  
CDD='05'  
CASE(6)  
CDD='06'  
CASE(7)  
CDD='07'  
CASE(8)  
CDD='08'  
CASE(9)  
CDD='09'  
CASE(10)  
CDD='10'  
CASE(11)  
CDD='11'  
CASE(12)  
CDD='12'  
CASE(13)  
CDD='13'  
CASE(14)  
CDD='14'  
CASE(15)  
CDD='15'  
CASE(16)  
CDD='16'  
CASE(17)  
CDD='17'  
CASE(18)  
CDD='18'  
CASE(19)  
CDD='19'  
CASE(20)  
CDD='20'  
CASE(21)  
CDD='21'  
CASE(22)  
CDD='22'
```

```

CASE(23)
CDD='23'
CASE(24)
CDD='24'
CASE(25)
CDD='25'
CASE(26)
CDD='26'
CASE(27)
CDD='27'
CASE(28)
CDD='28'
CASE(29)
CDD='29'
CASE(30)
CDD='30'
CASE(31)
CDD='31'
END SELECT

*
WRITE(*,*)
NMS=0
*
* ARMAMOS EL NOMBRE DEL ARCHIVO DE DATOS Y POSTERIORMENTE LO ABRIMOS
*
ARCH DATOS=DR // ':SINOP' // CHH // '\SINO' // CDD // CHH
PRINT 377, ARCH_DATOS
377 FORMAT(T3,35A)
OPEN(1,FILE=ARCH_DATOS,IOSTAT=K)
IF (K.NE.0) GOTO 99
DO 310,I=1,80
DO 311,J=1,10
INFORME(I,J)='
311 CONTINUE
310 CONTINUE

* COMENZAMOS A LEER DEL ARCHIVO DE DATOS
51 READ (1,30,IOSTAT=K,END=92) (LIN(I),I=2,86)
30 FORMAT (85A1)
IF (K) 92,53,99
*
* Y SI ENCONTRAMOS EN EL REGISTRO LEIDO QUE SU ENCABEZADO ES UN NUMERO
* VALIDO ENTONCES PROCEDEMOS CON LA DECODIFICACION DE LOS GRUPOS
* SIGUIENTES, DE LO CONTRARIO LO ABANDONAMOS Y LEEMOS EL SIGUIENTE
* REGISTRO.
*
53 IF (LIN(2).EQ.'7'.AND.LIN(3).EQ.'6') GOTO 54
L=80
200 IF (L-10) 51,215,215
215 LB=L-9
*
*****
* INICIAMOS CON EL CODIGO DE DECODIFICACION

DO 201,IB=2,LB
IB1=IB+1
IB2=IB+2
IB3=IB+3
IB4=IB+4

```

```

                IF (LIN(IB).EQ.'3'.AND.LIN(IB1).EQ.'3'.AND.LIN(IB2).EQ.'3'.
1AND.LIN(IB3).EQ.' ') GOTO 250
201 CONTINUE
    TXD5=' '
    TND5=' '
    PRD5=' '
    GOTO 56
250 LB2=L-5
    ISB=IB3
    DO 202,ITX=ISB,LB2
        ITX1=ITX+1
        ITX2=ITX+2
        ITX3=ITX+3
        ITX4=ITX+4
        ITX5=ITX+5
        IF (LIN(ITX).EQ.' '.AND.LIN(ITX1).EQ.'1') GOTO 220
202 CONTINUE
    TXD5='-9999'
    GOTO 224
220 IF (LIN(ITX2).EQ.'0'.OR.LIN(ITX2).EQ.'1') GOTO 20
    TXD5='-9999'
    GOTO 224
20 IF (LIN(ITX2).NE.' '.AND.LIN(ITX3).NE.' '.AND.LIN(ITX4).NE.' '
1AND.LIN(ITX5).NE.' ') GOTO 222
    TXD5='-9999'
    GOTO 224
222 IF (LIN(ITX2).EQ.'0') SG=' '
    IF (LIN(ITX2).EQ.'1') SG='- '
    DO 207,INTX=ITX3,ITX4
        IMM=INTX-(ITX1)
*
* DECODIFICAMOS TEMPERATURA MAXIMA
*
207 TXD5(IMM:IMM)=LIN(INTX)
    TXD5(1:1)=SG
    TXD5(4:4)=PUNTO
    TXD5(5:5)=LIN(ITX5)
224 DO 203,ITN=ISB,LB2
    ITN1=ITN+1
    ITN2=ITN+2
    ITN3=ITN+3
    ITN4=ITN+4
    ITN5=ITN+5
    IF (LIN(ITN).EQ.' '.AND.LIN(ITN1).EQ.'2') GOTO 21
203 CONTINUE
    TND5='-9999'
    GOTO 230
21 IF (LIN(ITN2).NE.' '.AND.LIN(ITN3).NE.' '.AND.LIN(ITN4).NE.' '
1AND.LIN(ITN5).NE.' ') GOTO 225
    TND5='-9999'
    GOTO 230
225 IF (LIN(ITN2).EQ.'0'.OR.LIN(ITN2).EQ.'1') GOTO 226
    TND5='-9999'
    GOTO 230
226 IF (LIN(ITN2).EQ.'0') SG=' '
    IF (LIN(ITN2).EQ.'1') SG='- '
    DO 208,INTN=ITN3,ITN4
        IIM=INTN-(ITN1)
*
* DECODIFICAMOS TEMPERATURA MINIMA

```

```

*
208 TND5(IIM:IIM)=LIN(INTN)
    TND5(1:1)=SG
    TND5(4:4)=PUNTO
    TND5(5:5)=LIN(ITN5)
230 PRD5='
    DO 204,IPN=ISB,LB2
    IPN1=IPN+1
    IPN2=IPN+2
    IPN3=IPN+3
    IPN4=IPN+4
    IPN5=IPN+5
    IF (LIN(IPN).EQ.' '.AND.LIN(IPN1).EQ.'7') GOTO 217
204 CONTINUE
    PRD5='-9999'
    GOTO 56
217 IF (LIN(IPN2).EQ.'9'.AND.LIN(IPN3).EQ.'9'.AND.LIN(IPN4).EQ.
1'9'.AND.LIN(IPN5).EQ.'9') GOTO 237
    IF (LIN(IPN2).EQ.'9'.AND.(LIN(IPN3).EQ.'9'.AND.
1LIN(IPN4).EQ.'9')) GOTO 238
    GOTO 235
237 PRD5='-9999'
    GOTO 56
238 PRD5='000.'//(LIN(IPN5))
    GOTO 56
235 IF (LIN(IPN2).NE.'/'.AND.LIN(IPN3).NE.'/'.AND.LIN(IPN4).NE.
1'/''.AND.LIN(IPN5).NE.'/') GOTO 22
    PRD5='-9999'
    GOTO 56
22 IF (LIN(IPN2).NE.' '.AND.LIN(IPN3).NE.' '.AND.LIN(IPN4).NE.' '.
1AND.LIN(IPN5).NE.' ') GOTO 236
    PRD5='-9999'
    GOTO 56
236 DO 209,INPN=IPN2,IPN4
    IIM=INPN-(IPN2-1)
*
* DECODIFICAMOS PRECIPITACION
*
209 PRD5(IIM:IIM)=LIN(INPN)
    PRD5(4:4)=PUNTO
    PRD5(5:5)=LIN(IPN5)
    GOTO 56
54 NMS=NMS+1
    IF (LIN(4).NE.' '.AND.LIN(5).NE.' '.AND.LIN(6).NE.' ') GOTO 23
    GOTO 51
23 DO 2,I=4,6
    IM=I-3
2 NSTN(IM:IM)=LIN(I)
109 PRINT *, ' Mensaje : ',NMS, ' Estacion : ',NSTN, ' NP=',NP
*
*
    ID1=ICHAR(LIN(14))
    D1=ID1-48
    IF (LIN(13).EQ.' '.AND.(D1.GE.0.AND.D1.LE.9)) GOTO999
    GOTO 998
998 IF (LIN(13).EQ.' '.AND.LIN(14).EQ.'/') GOTO 997
997 N='0'
    GOTO 996
*
* DECODIFICAMOS DIRECCION E INTENSIDAD DEL VIENTO

```

```

*
999 N=LIN(14)
996 DIRECCION=LIN(15)//LIN(16)//CERO
    IF (DIRECCION.EQ.'99') THEN DIRECCION='999'
    INTENSIDAD=LIN(17)//LIN(18)
*
*
    IF (LIN(19).EQ.' ') GOTO 110
    TAD5='
    GOTO 113
110 IF (LIN(20).EQ.'1') GOTO 24
    TAD5='-9999'
    GOTO 113
24  IF (LIN(21).NE.' '.AND.LIN(22).NE.' '.AND.LIN(23).NE.' '.AND.
    1LIN(24).NE.' ') GOTO 111
    TAD5='-9999'
    GOTO 113
111 IF (LIN(21).EQ.'0'.OR.LIN(21).EQ.'1') GOTO 112
    TAD5='-9999'
    GOTO 113
112 IF (LIN(21).EQ.'0') SG=' '
    IF (LIN(21).EQ.'1') SG='- '
    DO 3,I=22,23
        IM=I-20
*
*   DECODIFICAMOS TEMPERATURA AMBIENTE
*
3   TAD5(IM:IM)=LIN(I)
    TAD5(1:1)=SG
    TAD5(4:4)=PUNTO
    TAD5(5:5)=LIN(24)
113 IF (LIN(25).EQ.' ') GOTO 114
    TRD5='-9999'
    GOTO 117
114 IF (LIN(26).EQ.'2') GOTO 25
    TRD5='-9999'
    GOTO 117
25  IF (LIN(27).NE.' '.AND.LIN(28).NE.' '.AND.LIN(29).NE.' '.AND.
    1LIN(30).NE.' ') GOTO 115
    TRD5='-9999'
    GOTO 117
115 IF (LIN(27).EQ.'0'.OR.LIN(27).EQ.'1') GOTO 116
    TRD5='-9999'
    GOTO 117
116 IF (LIN(27).EQ.'0') SG=' '
    IF (LIN(27).EQ.'1') SG='- '
    DO 4,I=28,29
        IM=I-26
*
*   DECODIFICAMOS TEMPERATURA DE PUNTO DE ROCIO
*
4   TRD5(IM:IM)=LIN(I)
    TRD5(1:1)=SG
    TRD5(4:4)=PUNTO
    TRD5(5:5)=LIN(30)
117 SG=' '
    IF (LIN(31).EQ.' ') GOTO 118
    PSD6='-99999'
    GOTO 121
118 IF (LIN(32).EQ.'3') GOTO 26

```

```

PSD6='-99999'
SG=' '
GOTO 121
26 IF (LIN(33).NE.' '.AND.LIN(34).NE.' '.AND.LIN(35).NE.' '.AND.
1LIN(36).NE.' ') GOTO 119
PSD6='-99999'
GOTO 121
119 IF (LIN(33).NE.'/'.AND.LIN(34).NE.'/'.AND.LIN(35).NE.'/'.AND.
1LIN(36).NE.'/') GOTO 120
PSD6='-99999'
GOTO 121
120 IF (LIN(33).EQ.'0') SG='1'
DO 5,I=33,35
IM=I-31
*
* DECODIFICAMOS PRESION ATMOSFERICA EN EL NIVEL DE LA ESTACION
*
5 PSD6(IM:IM)=LIN(I)
PSD6(1:1)=SG
PSD6(5:5)=PUNTO
PSD6(6:6)=LIN(36)
121 IF (LIN(37).EQ.' ') GOTO 122
PMD6='-99999'
GOTO 55
122 IF (LIN(38).EQ.'4') GOTO 27
PMD6='-99999'
SG=' '
GOTO 55
27 IF (LIN(39).NE.' '.AND.LIN(40).NE.' '.AND.LIN(41).NE.' '.AND.
1LIN(42).NE.' ') GOTO 123
PMD6='-99999'
GOTO 55
123 IF (LIN(39).NE.'/'.AND.LIN(40).NE.'/'.AND.LIN(41).NE.'/'.AND.
1LIN(42).NE.'/') GOTO 124
PMD6='-99999'
GOTO 55
124 IF (LIN(39).EQ.'0') SG='1'
DO 6,I=39,41
IM=I-37
*
* DECODIFICAMOS PRESION ATMOSFERICA REDUCIDA AL NIVEL MEDIO DEL MAR
*
6 PMD6(IM:IM)=LIN(I)
PMD6(1:1)=SG
PMD6(5:5)=PUNTO
PMD6(6:6)=LIN(42)
125 GOTO 55
*****
*
* ALMACENAMOS LA UBICACION GEOGRAFICA DEL OBSERVATORIO QUE ENVIO EL
* MENSAJE
55 NP=0
IF (NSTN.EQ.'050') THEN
NP=1
LAT(1)=31.85
LON(1)=-116.83
ENDIF
IF (NSTN.EQ.'055') THEN

```

```

      NP=2
*     LAT(2)=
*     LON(2)=
ENDIF
IF (NSTN.EQ.'061') THEN
  NP=3
  LAT(3)=31.35
  LON(3)=-113.67
ENDIF
IF (NSTN.EQ.'113') THEN
*   NP=4
*   LAT(NP)=
*   LON(NP)=
ENDIF
IF (NSTN.EQ.'118') THEN
*   NP=5
*   LAT(NP)=
*   LON(NP)=
ENDIF
IF (NSTN.EQ.'122') THEN
  NP=6
  LAT(6)=30.42
  LON(6)=-107.92
ENDIF
IF (NSTN.EQ.'151') THEN
  NP=7
  LAT(NP)=29.17
  LON(NP)=-118.32
ENDIF
IF (NSTN.EQ.'160') THEN
  NP=8
  LAT(NP)=29.08
  LON(NP)=-110.95
ENDIF
IF (NSTN.EQ.'220') THEN
  NP=9
  LAT(NP)=28.95
  LON(NP)=-107.82
ENDIF
IF (NSTN.EQ.'225') THEN
  NP=10
  LAT(NP)=28.63
  LON(NP)=-106.07
ENDIF
IF (NSTN.EQ.'243') THEN
  NP=11
  LAT(NP)=28.70
  LON(NP)=-100.52
ENDIF
IF (NSTN.EQ.'253') THEN
*   NP=12
*   LAT(NP)=
*   LON(NP)=
ENDIF
IF (NSTN.EQ.'256') THEN
  NP=13
  LAT(NP)=27.95
  LON(NP)=-110.80
ENDIF
IF (NSTN.EQ.'258') THEN

```

```

NP=26
LAT(NP)=24.82
LON(NP)=-107.40
ENDIF
IF (NSTN.EQ.'423') THEN
NP=27
LAT(NP)=24.03
LON(NP)=-104.67
ENDIF
IF (NSTN.EQ.'458') THEN
NP=28
LAT(NP)=23.20
LON(NP)=-106.42
ENDIF
IF (NSTN.EQ.'471') THEN
NP=29
*   LAT(NP)=
*   LON(NP)=
ENDIF
IF (NSTN.EQ.'491') THEN
NP=30
LAT(NP)=23.73
LON(NP)=-99.13
ENDIF
IF (NSTN.EQ.'499') THEN
NP=31
LAT(NP)=23.80
LON(NP)=-98.20
ENDIF
IF (NSTN.EQ.'519') THEN
NP=32
*   LAT(NP)=
*   LON(NP)=
ENDIF
IF (NSTN.EQ.'525') THEN
NP=33
LAT(NP)=22.78
LON(NP)=-102.58
ENDIF
IF (NSTN.EQ.'539') THEN
NP=34
LAT(NP)=22.15
LON(NP)=-100.98
ENDIF
IF (NSTN.EQ.'543') THEN
*
IF (NSTN.EQ.'548') THEN
NP=35
LAT(NP)=22.22
LON(NP)=-97.85
ENDIF
IF (NSTN.EQ.'551') THEN
*
IF (NSTN.EQ.'556') THEN
NP=36
LAT(NP)=21.50
LON(NP)=-104.88
ENDIF
IF (NSTN.EQ.'570') THEN
*
IF (NSTN.EQ.'571') THEN
NP=37
LAT(NP)=21.88

```

```

LON(NP)=-102.30
ENDIF
IF (NSTN.EQ.'577') THEN
NP=38
LAT(NP)=21.02
LON(NP)=-101.25
ENDIF
IF (NSTN.EQ.'581') THEN
NP=39
LAT(NP)=21.93
LON(NP)=-100
ENDIF
IF (NSTN.EQ.'585') THEN
NP=40
LAT(NP)=
LON(NP)=
ENDIF
IF (NSTN.EQ.'590') THEN
IF (NSTN.EQ.'593') THEN
NP=41
LAT(NP)=21.3
LON(NP)=-86.67
ENDIF
IF (NSTN.EQ.'612') THEN
NP=42
LAT(NP)=20.68
LON(NP)=-103.68
ENDIF
IF (NSTN.EQ.'625') THEN
NP=43
LAT(NP)=20.60
LON(NP)=-100.38
ENDIF
IF (NSTN.EQ.'632') THEN
NP=44
LAT(NP)=20.13
LON(NP)=-98.73
ENDIF
IF (NSTN.EQ.'634') THEN
NP=45
LAT(NP)=20.08
LON(NP)=-98.37
ENDIF
IF (NSTN.EQ.'640') THEN
NP=46
LAT(NP)=20.95
LON(NP)=-97.38
ENDIF
IF (NSTN.EQ.'644') THEN
NP=47
LAT(NP)=20.95
LON(NP)=-89.70
ENDIF
IF (NSTN.EQ.'647') THEN
NP=48
LAT(NP)=20.68
LON(NP)=-88.20
ENDIF
IF (NSTN.EQ.'648') THEN
NP=49

```

```
LAT(NP)=20.52
LON(NP)=-86.95
ENDIF
IF (NSTN.EQ.'654') THEN
  NP=50
  LAT(NP)=19.05
  LON(NP)=-104.33
ENDIF
IF (NSTN.EQ.'656') THEN
  NP=51
  LAT(NP)=19.70
  LON(NP)=-103.47
ENDIF
IF (NSTN.EQ.'662') THEN
  NP=52
  LAT(NP)=
  LON(NP)=
ENDIF
IF (NSTN.EQ.'665') THEN
  NP=53
  LAT(NP)=19.70
  LON(NP)=-101.18
ENDIF
IF (NSTN.EQ.'675') THEN
  NP=54
  LAT(NP)=19.30
  LON(NP)=-99.67
ENDIF
IF (NSTN.EQ.'679') THEN
  NP=55
  LAT(NP)=19.43
  LON(NP)=-99.08
ENDIF
IF (NSTN.EQ.'680') THEN
  NP=56
  LAT(NP)=19.43
  LON(NP)=-99.13
ENDIF
IF (NSTN.EQ.'683') THEN
  NP=57
  LAT(NP)=19.32
  LON(NP)=-98.23
ENDIF
IF (NSTN.EQ.'685') THEN
  NP=58
  LAT(NP)=19.85
  LON(NP)=-98.20
ENDIF
IF (NSTN.EQ.'687') THEN
  NP=59
  LAT(NP)=19.53
  LON(NP)=-96.92
ENDIF
IF (NSTN.EQ.'692') THEN
  NP=60
  LAT(NP)=19.15
  LON(NP)=-96.12
ENDIF
IF (NSTN.EQ.'695') THEN
  NP=61
```

```

LAT(NP)=19.85
LON(NP)=-90.53
ENDIF
IF (NSTN.EQ.'698') THEN
NP=62
*   LAT(NP)=
*   LON(NP)=
ENDIF
IF (NSTN.EQ.'723') THEN
NP=63
LAT(NP)=18.83
LON(NP)=-111.00
ENDIF
IF (NSTN.EQ.'726') THEN
NP=64
*   LAT(NP)=
*   LON(NP)=
ENDIF
IF (NSTN.EQ.'737') THEN
NP=65
LAT(NP)=18.85
LON(NP)=-97.10
ENDIF
IF (NSTN.EQ.'741') THEN
NP=66
LAT(NP)=18.15
LON(NP)=-94.42
ENDIF
IF (NSTN.EQ.'743') THEN
NP=67
*   LAT(NP)=
*   LON(NP)=
ENDIF
IF (NSTN.EQ.'750') THEN
NP=68
LAT(NP)=18.50
LON(NP)=-88.30
ENDIF
IF (NSTN.EQ.'762') THEN
NP=69
LAT(NP)=17.55
LON(NP)=-99.50
ENDIF
IF (NSTN.EQ.'773') THEN
NP=70
*   LAT(NP)=
*   LON(NP)=
ENDIF
IF (NSTN.EQ.'775') THEN
NP=71
LAT(NP)=17.07
LON(NP)=-96.72
ENDIF
IF (NSTN.EQ.'805') THEN
NP=72
LAT(NP)=16.83
LON(NP)=-99.92
ENDIF
IF (NSTN.EQ.'833') THEN
NP=73

```

```

LAT(NP)=16.17
LON(NP)=-95.20
ENDIF
IF (NSTN.EQ.'840') THEN
NP=74
LAT(NP)=
LON(NP)=
ENDIF
IF (NSTN.EQ.'843') THEN
NP=75
LAT(NP)=16.75
LON(NP)=-93.10
ENDIF
IF (NSTN.EQ.'845') THEN
NP=76
LAT(NP)=16.73
LON(NP)=-92.63
ENDIF
IF (NSTN.EQ.'848') THEN
NP=77
LAT(NP)=16.25
LON(NP)=-92.13
ENDIF
IF (NSTN.EQ.'855') THEN
NP=78
LAT(NP)=15.70
LON(NP)=-96.50
ENDIF
IF (NSTN.EQ.'903') THEN
NP=79
LAT(NP)=14.92
LON(NP)=-92.27
ENDIF

```

```

*
* LLENAMOS UNA MATRIZ CON INFORMACION DECODIFICADA PARA POSTERIORMENTE
* IMPRIMIRLA, ADEMAS OBTENEMOS EL VALOR NUMERICO DE LOS PARAMETROS
* PUESTO QUE TODO EL TIEMPO, EL PROGRAMA LOS TRABAJA COMO CADENAS DE
* CARACTERES

```

```

INFORME(NP,2)=TAD5//SP
IF (TAD5.NE.'-9999'.AND.TAD5.NE.' ') THEN
CEN=ICHAR(TAD5(2:2))
CENTENA=CEN-48
DEC=ICHAR(TAD5(3:3))
DECENA=DEC-48
UNI=ICHAR(TAD5(5:5))
UNIDAD=UNI-48
ITAD5=CENTENA*100+DECENA*10+UNIDAD
NQ=10*(IDIA-1)+2
ESTACION(NP,NQ)=ITAD5
ELSE
ESTACION(NP,NQ)=0
ENDIF
INFORME(NP,8)=N//SP
INFORME(NP,9)=DIRECCION//SP
INFORME(NP,10)=INTENSIDAD//SP
INFORME(NP,3)=TRD5//SP
IF (TRD5.NE.'-9999'.AND.TRD5.NE.' ') THEN

```

```

CEN=ICHAR(TRD5(2:2))
CENTENA=CEN-48
DEC=ICHAR(TRD5(3:3))
DECENA=DEC-48
UNI=ICHAR(TRD5(5:5))
UNIDAD=UNI-48
ITRD5=CENTENA*100+DECENA*10+UNIDAD
NQ=10*(IDIA-1)+3
ESTACION(NP,NQ)=ITRD5
ELSE
ESTACION(NP,NQ)=0
ENDIF
INFORME(NP,6)=PSD6//SP
IF (PSD6.NE.'-9999'.AND.PSD6.NE.' ') THEN
DMIL=ICHAR(PSD6(1:1))
DMILES=DMIL-48
MIL=ICHAR(PSD6(2:2))
MILES=MIL-48
CEN=ICHAR(PSD6(3:3))
CENTENA=CEN-48
DEC=ICHAR(PSD6(4:4))
DECENA=DEC-48
UNI=ICHAR(PSD6(6:6))
UNIDAD=UNI-48
IPSD6=DMILES*10000+MILES*1000+CENTENA*100+DECENA*10+UNIDAD
NQ=10*(IDIA-1)+6
ESTACION(NP,NQ)=IPSD6
ELSE
ESTACION(NP,NQ)=0
ENDIF
INFORME(NP,7)=PMD6//SP
IF (PMD6.NE.'-9999'.AND.PSD6.NE.' ') THEN
DMIL=ICHAR(PMD6(1:1))
DMILES=DMIL-48
MIL=ICHAR(PMD6(2:2))
MILES=MIL-48
CEN=ICHAR(PMD6(3:3))
CENTENA=CEN-48
DEC=ICHAR(PMD6(4:4))
DECENA=DEC-48
UNI=ICHAR(PMD6(6:6))
UNIDAD=UNI-48
IPMD6=DMILES*10000+MILES*1000+CENTENA*100+DECENA*10+UNIDAD
NQ=10*(IDIA-1)+7
ESTACION(NP,NQ)=IPMD6
ELSE
ESTACION(NP,NQ)=0
ENDIF
TAD5=' '
TRD5=' '
PSD6=' '
PMD6=' '
N=' '
DIRECCION=' '
INTENSIDAD=' '
GOTO 51
INFORME(NP,4)=TXD5//SP
IF (TXD5.NE.'-9999'.AND.TXD5.NE.' ') THEN
CEN=ICHAR(TXD5(2:2))
CENTENA=CEN-48

```

56

```

DEC=ICHAR(TXD5(3:3))
DECENA=DEC-48
UNI=ICHAR(TXD5(5:5))
UNIDAD=UNI-48
ITXD5=CENTENA*100+DECENA*10+UNIDAD
NQ=10*(IDIA-1)+4
ESTACION(NP,NQ)=ITXD5
ELSE
ESTACION(NP,NQ)=0
ENDIF
INFORME(NP,5)=TND5//SP
IF (TND5.NE.'-9999'.AND.TND5.NE.' ') THEN
CEN=ICHAR(TND5(2:2))
CENTENA=CEN-48
DEC=ICHAR(TND5(3:3))
DECENA=DEC-48
UNI=ICHAR(TND5(5:5))
UNIDAD=UNI-48
ITND5=CENTENA*100+DECENA*10+UNIDAD
NQ=10*(IDIA-1)+5
ESTACION(NP,NQ)=ITND5
ELSE
ESTACION(NP,NQ)=0
ENDIF
INFORME(NP,1)=PRD5//SP
IF (PRD5.NE.'-9999'.AND.PRD5.NE.' ') THEN
CEN=ICHAR(PRD5(2:2))
CENTENA=CEN-48
DEC=ICHAR(PRD5(3:3))
DECENA=DEC-48
UNI=ICHAR(PRD5(5:5))
UNIDAD=UNI-48
IPRD5=CENTENA*100+DECENA*10+UNIDAD
NQ=10*(IDIA-1)+1
ESTACION(NP,NQ)=IPRD5
ELSE
ESTACION(NP,NQ)=0
ENDIF
TXD5=' '
TND5=' '
PRD5=' '
NP=0
GOTO 51

```

```

*****
*
* MANDAMOS TODA LA INFORMACION DECODIFICADA AL ARCHIVO TAB_COM
*
92 PRINT 910,NMS
910 FORMAT (' Se decodificaron : ',I4,' mensajes ')
*
WRITE (3,34,IOSTAT=K) SP
34 FORMAT (' AA MM DD HR No. PCPN TA TR TX TN
1PSTN PNMN NUBES DIRECC INTEN')
*
WRITE(3,*)
*
ANIOB=ANIO// ' '
NMESE=NMESE// ' '
CDDB=CDD// ' '

```

```

DO 19,NR=1,79
WRITE (3,33,Iostat=K) ANIOB,NMESB,CDDB,CIH,NOMBRE(NR),
1(INFORME(NR,I),I=1,10)

```

```

*****

```

```

* LLENAMOS EL ARCHIVO DAT_MAP CON INFORMACION DEL PARAMETRO
* DECODIFICADO QUE SE SELECCIONO PARA EL TRAZADO DE LOS MAPAS DE
* ISOLINEAS

```

```

* IF (LAT(NR).NE.0.O.AND.INFORME(NR,ELECCION).NE.'-9999') THEN
* WRITE (5,36,Iostat=K) LAT(NR),' ',LON(NR),' ',
1(INFORME(NR,ELECCION))
* ENDF
19 CONTINUE
33 FORMAT (' ',4A,7A7,' ',10A,' ',10A,' ',10A)
36 FORMAT(F6.2,A1,F7.2,A1,10A)

```

```

IF (K.NE.0) GOTO 99

```

```

* 314 CONTINUE
313 CONTINUE

```

```

*****

```

```

* CREAMOS EL ARCHIVO EN EL QUE SE ALMACENARA INFORMACION RELACIONADA
* CON LAS TENDENCIAS Y "COMPORTAMIENTO ESTADISTICO" DE LOS DATOS
* DECODIFICADOS

```

```

* OPEN(5,FILE='C:DESCRIPT.SAL',Iostat=K,ACCESS='SEQUENTIAL',
* IFORM='FORMATTED')
* IF (K.NE.0) GOTO 99

```

```

* DO 315,RENGLON=1,79
* WRITE(5,36,Iostat=K) (ESTACION(RENGLON,COLUMNA),COLUMNA=1,25)
*36 FORMAT(25(I5))
*315 CONTINUE
* IF (K.NE.0) GOTO 99

```

```

* DO 517,REN=1,79
* NOCERO=0
* SUBTOTAL=0
* SUBTOTAL1=0
* DO 501,CONTADOR=1,31
* MEDIO=ESTACION(REN,10*CONTADOR-8)
* IF (MEDIO.NE.0) THEN
* NOCERO=NOCERO+1
* SUBTOTAL=SUBTOTAL+MEDIO
*501 ENDF

```

```

* IF (NOCERO.NE.0) THEN
* MEDI=LOG(SUBTOTAL)-LOG(10*NOCERO)
* MEDIA(REN)=EXP(MEDI)
* ELSE
* MEDIA(REN)=0
* ENDF

```

```

* DO 502,CONTADOR=1,31
* IF (ESTACION(REN,10*CONTADOR-8).NE.0) THEN
* VAR10=(ESTACION(REN,10*CONTADOR-8)-10*MEDIA(REN))**2

```

```

*      VARI10=LOG(VARI10)-LOG(100)
*      VARIA10=EXP(VARI10)
*      SUBTOTAL1=SUBTOTAL1+VARIA10
*502  ENDIF
*      IF (NOCERO.GT.1) THEN
*          VARI=LOG(SUBTOTAL1)-LOG(NOCERO-1)
*      ENDIF
*      IF (NOCERO.EQ.1) THEN
*          VARI=LOG(SUBTOTAL1)-LOG(NOCERO)
*      ENDIF
*      IF (NOCERO.NE.0) THEN
*          VARIANZA(REN)=EXP(VARI)
*      ENDIF
*      IF (NOCERO.EQ.0) THEN
*          VARIANZA(REN)=0
*      ENDIF
*517  CONTINUE
*
*      WRITE(5,513) ' ESTACION', '   HORA', '   MEDIA', '   VARIANZA '
*513  FORMAT(10A,10A,11A,16A)
*      DO 510,APUN=1,79
*          WRITE(5,519) NOMBRE(APUN)
*519  FORMAT(10A)
*      DO 512,JDIA=1,8
*          SELECT CASE(JDIA)
*              CASE(1)
*                  CHH1='00'
*              CASE(2)
*                  CHH1='03'
*              CASE(3)
*                  CHH1='06'
*              CASE(4)
*                  CHH1='09'
*              CASE(5)
*                  CHH1='12'
*              CASE(6)
*                  CHH1='15'
*              CASE(7)
*                  CHH1='18'
*              CASE(8)
*                  CHH1='21'
*          END SELECT
*          WRITE(5,'(A5,3X,F6.2,3X,F4.2)')
*          + CHH1,MEDIA(APUN),VARIANZA(APUN)
*514  FORMAT(2A,5X,2F7.2)
*512  CONTINUE
*510  CONTINUE
*
*      DUMMY2 = SETVIDEOMODE($ERESCOLOR)
*      CALL MOVETO(80,50,xy)
*      DUMMY = LINETO(80,200)
*      DUMMY = LINETO(240,200)
*      CALL MOVETO(79,50,xy)
*      DUMMY = LINETO(79,201)
*      DUMMY = LINETO(240,201)
*      DUMMY2 = SETCOLOR(4)
*      DO 516,I=1,NOCERO
*          J=30*I+60
*          K=30*I+80

```

```

* L=400-ESTACION(9,10*I-8)
*516 DUMMY2 = RECTANGLE($GFILLINTERIOR,J,L,K,200)
* READ(*,*)
* DUMMY = SETVIDEOMODE($DEFAULTMODE)
*
* WRITE(3,31) RESET
WRITE(*,*)
WRITE(*,*)
PRINT 370, '          PROGRAMA TERMINADO SIN ERROR '
370 FORMAT(60A)
GOTO 103
99  WRITE(*,*)
WRITE(*,*)
PRINT 371, '          !! ERROR !! , NO SE PUEDE ABRIR EL ARCHIVO '
PRINT 371, '          VERIFICAR QUE DIRECTORIOS O ARCHIVOS EXISTAN '
371 FORMAT (80A)
103 CLOSE(1)
CLOSE(3)
END

```



```

OPEN(1, FILE='C:DAT_MAP', IOSTAT=K, ACCESS='SEQUENTIAL',
1FORM='FORMATTED')
IF (K.NE.0) GOTO 99
100 READ(1, 20, IOSTAT=N, END=92) LATITUD, LONGITUD, COMMA, IMAGEN
20  FORMAT(F6.2, F7.2, A2, F5.2)
IF (N) 92, 53, 99
53  I=INT(LATITUD)-13
J=INT(LONGITUD)+85
MATRIZ(I, J)=IMAGEN
GOTO 100
92  DO 50, L=0, 20
WRITE(*, 70, IOSTAT=K) (MATRIZ(L, M), M=-34, 0)
70  FORMAT(35(F4.1))
50  CONTINUE
WRITE(*, *) ' '
WRITE(*, *) ' '

DO 200, RENG=1, 19
DO 300, COL=1, 33
CALL TOFIX(MATRIZ, RENG, COL, TEMP2)
CALL LOOKUP(TEMP1, MATRIZ, RENG, COL, LAT, LONG)
CALL LOOKDOWN(TEMP2, MATRIZ, RENG, COL, Z, LAT, LONG)
CALL LOOKLEFT(TEMP3, MATRIZ, RENG, COL, Z, LAT, LONG)
CALL LOOKRIGHT(TEMP4, MATRIZ, RENG, COL, Z, LAT, LONG)
CALL CALCULATE(LAT, LONG, Z, MATRIZ, RENG, COL, SUMDIST, INTERP)
300  CONTINUE
200  CONTINUE
*
DO 55, L=0, 20
55  WRITE(*, 70, IOSTAT=K) (MATRIZ(L, M), M=-34, 0)

STOP
END

```

SUBROUTINE TOFIX(MATRIZ, RENG, COL, TEMP2)

*
* SUBROUTINA QUE UBICA LAS CASILLAS VACIAS Y EN LAS QUE SE DEBE DE ESTIMAR
* EL VALOR DE LA VARIABLE. TAMBIEN SE OBTIEN SUS COORDENADAS GEOGRAFICAS
* (LATITUD Y LONGITUD)
*

```

INTEGER RENG, COL, L, M, TEMP1, TEMP2, TEMP3, TEMP4
INTEGER SUMDIST, LAT, LONG, D
REAL MATRIZ, INTERP, Z
DIMENSION MATRIZ(20, -36)
DIMENSION LAT(5), LONG(5)
DIMENSION Z(4)
DIMENSION D(8)

```

```

IF (MATRIZ(RENG, COL-34).EQ.0) THEN
DO 56, L=1, 5
LAT(L)=0
56 LONG(L)=0
DO 58, L=1, 4
58 Z(L)=0

```

```

57      DO 57,M=1,8
      *      D(M)=0

      TEMP1=RENG
      TEMP2=RENG
      TEMP3=COL-34
      TEMP4=COL-34
      LAT(1)=-RENG+33
      LONG(1)=COL-34-85
      SUMDIST=0
      INTERP=0.0
      ENDIF
      RETURN
      END

```

SUBROUTINE LOOKUP(TEMP1,MATRIZ,RENG,COL,LAT, LONG)

```

*
* SUBROUTINA QUE BUSCA EL PUNTO MAS CERCANO HACIA ARRIBA DE LA CABILLA
* QUE SE DESEA ESTIMAR, ESTE VALOR SERA GUARDADO EN UN ARREGLO DE
* LONGITUD 4. TAMBIEN SE OBTIENEN SUS CORDENADAS GEOGRAFICAS
* ( LATITUD Y LONGITUD ) PARA PODER CALCULAR, POSTERIORMENTE,
* LA DISTANCIA ENTRE AMBOS PUNTOS (EL ESTIMADO Y EL ENCONTRADO) Y
* UTILIZARLA COMO COCIENTE.
*

```

```

      INTEGER      TEMP1,RENG,COL,LAT, LONG
      REAL         MATRIZ,Z
      DIMENSION   LAT(5),LONG(5)
      DIMENSION   Z(4)
      DIMENSION   MATRIZ(20,-36)

```

```

      TEMP1=TEMP1-1
      IF (MATRIZ(TEMP1,COL-34).NE.0) THEN
        Z(1)=MATRIZ(TEMP1,COL-34)
        LAT(2)=-TEMP1+33
        LONG(2)=COL-34-85
      ENDIF
      DO WHILE(MATRIZ(TEMP1,COL-34).EQ.0.AND.TEMP1.GT.0)
        TEMP1=TEMP1-1
        IF (MATRIZ(TEMP1,COL-34).NE.0) THEN
          Z(1)=MATRIZ(TEMP1,COL-34)
          LAT(2)=-TEMP1+33
          LONG(2)=COL-34-85
          TEMP1=0
        ENDIF
      ENDDO
      RETURN
      END

```

SUBROUTINE LOOKDOWN(TEMP2,MATRIZ,RENG,COL,Z,LAT, LONG)

```

*
* SUBROUTINA QUE BUSCA EL PUNTO MAS CERCANO HACIA ABAJO DE LA CASILLA
* QUE SE DESEA ESTIMAR, ESTE VALOR SERA GUARDADO EN UN ARREGLO DE
* LONGITUD 4. TAMBIEN SE OBTIENEN SUS CORDENADAS GEOGRAFICAS
* ( LATITUD Y LONGITUD ) PARA PODER CALCULAR, POSTERIORMENTE,
* LA DISTANCIA ENTRE AMBOS PUNTOS (EL ESTIMADO Y EL ENCONTRADO) Y
* UTILIZARLA COMO COCIENTE.
*

```

```

INTEGER      TEMP2, RENG, COL, LAT, LONG
REAL         MATRIZ, Z
DIMENSION   Z(4)
DIMENSION   LAT(5), LONG(5)
DIMENSION   MATRIZ(20, -36)

```

```

TEMP2=TEMP2+1
IF (MATRIZ(TEMP2, COL-34).NE.0) THEN
  Z(2)=MATRIZ(TEMP2, COL-34)
  LAT(3)=-TEMP2+33
  LONG(3)=COL-34-85
ENDIF
DO WHILE(MATRIZ(TEMP2, COL-34).EQ.0.AND.TEMP2.LT.20)
  TEMP2=TEMP2+1
  IF (MATRIZ(TEMP2, COL-34).NE.0) THEN
    Z(2)=MATRIZ(TEMP2, COL-34)
    LAT(3)=-TEMP2+33
    LONG(3)=COL-34-85
    TEMP2=20
  ENDIF
ENDDO
RETURN
END

```

SUBROUTINE LOOKLEFT(TEMP3, MATRIZ, RENG, COL, Z, LAT, LONG)

```

*
* SUBROUTINA QUE BUSCA EL PUNTO MAS CERCANO HACIA LA IEQUIERDA DE LA
* CASILLA QUE SE DESEA ESTIMAR, ESTE VALOR SERA GUARDADO EN UN ARREGLO
* DE LONGITUD 4. TAMBIEN SE OBTIENEN SUS CORDENADAS GEOGRAFICAS
* ( LATITUD Y LONGITUD ) PARA PODER CALCULAR, POSTERIORMENTE, LA
* DISTANCIA ENTRE AMBOS PUNTOS (EL ESTIMADO Y EL ENCONTRADO) Y
* UTILIZARLA COMO COCIENTE.
*

```

```

INTEGER      TEMP3, RENG, COL, LAT, LONG
REAL         MATRIZ, Z
DIMENSION   MATRIZ(20, -36)
DIMENSION   LAT(5), LONG(5)
DIMENSION   Z(4)

```

```

101 TEMP3=TEMP3-1
    IF (MATRIZ(RENG, TEMP3).NE.0) THEN
      Z(3)=MATRIZ(RENG, TEMP3)
      LAT(4)=-RENG+33
    
```

```

LONG(4)=TEMP3-85
ELSE
IF (TEMP3.GT.-33) THEN
TEMP3=TEMP3-1
GOTO 101
ENDIF
ENDIF
RETURN
END

```

SUBROUTINE LOOKRIGHT(TEMP4, MATRIZ, RENG, COL, Z, LAT, LONG)

```

*
* SUBROUTINA QUE BUSCA EL PUNTO MAS CERCANO HACIA LA DERECHA DE LA
* CASILLA QUE SE DESEA ESTIMAR, ESTE VALOR SERA GUARDADO EN UN ARREGLO
* DE LONGITUD 4. TAMBIEN SE OBTIENEN SUS CORDENADAS GEOGRAFICAS
* ( LATITUD Y LONGITUD ) PARA PODER CALCULAR, POSTERIORMENTE, LA
* DISTANCIA ENTRE AMBOS PUNTOS (EL ESTIMADO Y EL ENCONTRADO) Y
* UTILIZARLA COMO COCIENTE.
*

```

```

INTEGER    TEMP4, RENG, COL, LAT, LONG
REAL       MATRIZ, Z
DIMENSION  MATRIZ(20, -36)
DIMENSION  LAT(5), LONG(5)
DIMENSION  Z(4)

```

```

TEMP4=TEMP4+1
IF (MATRIZ(RENG, TEMP4).NE.0) THEN
Z(4)=MATRIZ(RENG, TEMP4)
LAT(5)=-RENG+33
LONG(5)=TEMP4-85
ENDIF
DO WHILE(MATRIZ(RENG, TEMP4).EQ.0.AND.TEMP4.LE.-1)
TEMP4=TEMP4+1
IF (MATRIZ(RENG, TEMP4).NE.0) THEN
Z(4)=MATRIZ(RENG, TEMP4)
LAT(5)=-RENG+33
LONG(5)=TEMP4-85
TEMP4=0
ENDIF
ENDDO
RETURN
END

```

SUBROUTINE CALCULATE(LAT, LONG, Z, MATRIZ, RENG, COL, SUNDIST, INTERP)

```

*
* SUBROUTINA EN LA QUE SE HACEN LOS CALCULOS UTILIZANDO LA INFORMACION
* PROPORCIONADA POR LOS PROCEDIMIENTOS ANTERIORES
*

```



```
MATRIZ (RENG, COL-34) = INTERP/SUMDIST  
WRITE (2, *) -RENG+34, COL-119, MATRIZ (RENG, COL-34)  
ENDIF  
RETURN  
END
```

00

* PROGRAMA DE INTERPOLACION (INTER222.FOR): SINTESIS FINAL DE VARIAS IDEAS
 * SE SIGUE UTILIZANDO EL ALGORITMO DE BUSQUEDA EN ESPIRAL PERO AHORA SOLO
 * SE TOMAN EN CUENTA PARA LAS ESTIMACIONES A LOS DATOS INICIALES, ES DECIR
 * LA INFORMACION QUE PAULATINAMENTE SE VA GENERANDO EN CADA UNA DE LAS
 * CELDAS DE LA MATRIZ (MALLA) NO INTERVIENE EN EL PROCESO DE INTERPOLACION
 * DE LAS CASILLAS POSTERIORES

* UTILIZAMOS POR LO TANTO, DOS MATRICES: UNA DE TRABAJO Y LA OTRA DE
 * ESCRITURA EN LA PRIMERA UNICAMENTE SE ENCUENTRAN EMPLAZADOS LOS DATOS
 * INICIALES, MIENTRAS QUE EN LA SEGUNDA VAMOS INCORPORANDO TAMBIEN, EN
 * FORMA PAULATINA CADA UNA DE LAS ESTIMACIONES CALCULADAS.

```

CHARACTER  COMMA
REAL      SUMDIST,INTERP,D,SUMRECIPR
INTEGER   LAT, LONG, DLON, DLAT
INTEGER   I, J, K, L, M, N, COL, RENG, P, Q, ANT1, ANT2, STEEP
INTEGER   PASO, PASOL, PASA, PASA1, RENORI, COLORI, LIM ANT1, LIM ANT2
INTEGER   VECES, POT, POTEN, INCREM1, INCREM2, SUMANDO1, SUMANDO2, O, P1
INTEGER   LIMITE, INDICE1, INDICE2, INDICES, CONT, CONT1
REAL      IMAGEN, MATRIZ1, MATRIZ2, Z, LATITUD, LONGITUD
DIMENSION MATRIZ1(22,35), MATRIZ2(22,35)
DIMENSION Z(100), D(100), LAT(100), LONG(100), DLAT(100), DLON(100)
  
```

* INICIALIZAMOS LA MATRIZ

```

      STEEP=0
      DO 39, L=1, 22
DO 49, M=1, 35
      MATRIZ1(L, M)=0.1
      MATRIZ2(L, M)=0
49 CONTINUE
39 CONTINUE
  
```

* ABRIMOS ARCHIVOS DE ENTRADA Y SALIDA

```

      OPEN(4, FILE='C:TESS.SAL')
      OPEN(2, FILE='C:TESS.LIS')
      OPEN(1, FILE='C:DAT_MAP', IOSTAT=K, ACCESS='SEQUENTIAL',
      IFORM='FORMATTED')
      IF (K.NE.0) GOTO 99
  
```

* LEEMOS DEL ARCHIVO DE ENTRADA LOS DATOS NECESARIOS PARA LLENAR LA
 * MATRIZ

```

100 READ(1, 27, IOSTAT=N, END=92) LATITUD, LONGITUD, COMMA, IMAGEN
27  FORMAT(F7.2, F6.2, A3, F8.1)
      IF (N) 92, 53, 99
53  I=35-INT(LATITUD)
      J=120-INT(LONGITUD)
      WRITE(*, 21) ' ', LATITUD, ' ', LONGITUD, ' ', IMAGEN
21  FORMAT(A1, F5.2, A1, F7.2, A1, F8.2)
      MATRIZ1(I, J)=IMAGEN
      MATRIZ2(I, J)=IMAGEN
      GOTO 100

92  DO 50, L=1, 22
      WRITE(*, 70, IOSTAT=K) (MATRIZ1(L, M), M=1, 35)
  
```

```

70 FORMAT(35(F6.1))
50   CONTINUE

      WRITE(*,*)' '
      WRITE(*,*)' '

*   COMENZAMOS A RECORRER LAS ENTRADAS DE LA MATRIZ BUSCANDO LOCALIDADES
*   EN LAS QUE SI SEA NECESARIO ESTIMAR SU VALOR

      DO 200,RENG=2,21
      DO 300,COL=2,34

*
*   SI LA ENTRADA DE LA MATRIZ ESTA "VACIA", ES DECIR, SI AUN CONSERVA
*   SU VALOR DE INICIALIZACION, ENTONCES HAY QUE ESTIMAR SU VALOR, EN
*   CASO CONTRARIO NOS PASAMOS A LA SIGUIENTE ENTRADA.
*
      IF (MATRIZ1(RENG,COL).EQ.0.1) THEN

*
*   INICIALIZAMOS LAS VARIABLES AUXILIARES
*
      DO 56,L=1,100
      LAT(L)=0
56     LONG(L)=0
      DO 58,L=1,100
58     Z(L)=0
      DO 57,M=1,100
      D(M)=0
      DLAT(M)=0
57     DLON(M)=0
*
      SUMRECIPR=0
      INDICE1=1
      INDICE2=1
      LAT(INDICE2)=-RENG+35
      LONG(INDICE2)=-COL+120
      INDICE2=INDICE2+1
      SUMDIST=0.0
      INTERP=0.0

*****
*   INICIA EL CODIGO QUE PERMITE QUE LA BUSQUEDA SEA EN ESPIRAL

      PASO=1
      PASA=0
      CONT=0
      CONT1=1
      PASA1=0
      RENORI=RENG
      COLORI=COL
      LIM ANT1=RENORI
      LIM ANT2=COLORI-1
      VECES=0

*   SI EL PUNTO A ESTIMAR SE ENCUENTRA MUY A LA ORILLA, ENTONCES EL
*   RADIO DE BUSQUEDA REBASARA EN ALGUNAS DIRECCIONES LOS LIMITES DE
*   LA MATRIZ, ESTO SE TRADUCE EN QUE EL NUMERO DE PUNTOS QUE SEAN
*   CAPTURADOS POR EL RADIO SEA MENOR Y POR LO TANTO SE HACE
*   INDISPENSABLE DEFINIR UN RADIO MAYOR PARA INCLUIR MAS PUNTOS EN

```

* LA DIRECCION EN LA QUE LOS LIMITES DE LA MATRIZ NO SON REBASADOS
*

```
      IF (RENORI.LT.5.AND.COLORI.GT.28.OR.RENORI.LT.5.AND.COLORI
1.LT.5.OR.RENORI.GT.14.AND.COLORI.GT.28.OR.RENORI.GT.14.AND.
2.COLORI.LT.5) THEN
        LIMITE=10
      ELSE
        LIMITE=8
      ENDIF
DO 10,POT=1,LIMITE
  POTEN=(-1)**POT
  IF (PASA1.EQ.0) THEN
    INCREM1=0
    PASA1=1
  ELSE
    IF (PASA1.EQ.1) THEN
      INCREM1=1
      PASA1=2
    ENDIF
  ENDIF
  IF (INCREM1.LT.0) THEN
    INCREM1=-INCREM1
  ENDIF
  INCREM1=INCREM1+1
  IF (MOD(INCREM1,2).EQ.0) THEN
    INCREM1=INCREM1
  ELSE
    INCREM1=-INCREM1
  ENDIF
  IF (PASA.EQ.0) THEN
    INCREM2=-1
    PASA=1
  ELSE
    IF (PASA.EQ.1) THEN
      INCREM2=1
      PASA=2
    ENDIF
  ENDIF
  IF (INCREM2.LT.-1) THEN
    INCREM2=-INCREM2
  ENDIF
  INCREM2=INCREM2+1
  IF (MOD(INCREM2,2).EQ.0) THEN
    INCREM2=INCREM2
  ELSE
    INCREM2=-INCREM2
  ENDIF
```

*
* ESTABLECEMOS LOS LIMITES DE VARIACION DE LOS INDICES QUE
* RECORRERAN LA MATRIZ EN FORMA DE ESPIRAL. PRIMERA PARTE DE LA
* BUSQUEDA.
*

```
500      DO 20,I=LIM_ANT1,LIM_ANT1+INCREM1,POTEN
          DO 30,J=LIM_ANT2+INCREM2,LIM_ANT2+INCREM2,POTEN
```

* SI ESTAMOS EN UNA POSICION VALIDA DE LA MATRIZ ASIGNAMOS EL VALOR

```

* ENCONTRADO A UNA VARIABLE AUXILIAR

      IF (I.GE.1.AND.I.LE.21.AND.J.GE.1.AND.
1J.LE.34) THEN
      Z(INDICE1)=MATRIZ1(I,J)
*      write(2,'(f6.2)')z(indice1)
*
* CALCULAMOS LAS DISTANCIAS A LOS PUNTOS ENCONTRADOS
*
      INDICE1=INDICE1+1
      LAT(INDICE2)=35-I
      LONG(INDICE2)=-J+120
      CONT1=CONT1+1
      INDICE2=INDICE2+1
30      ENDIF
20      CONTINUE

      VECES=VECES+1
      IF (MOD(POT,2).EQ.1) THEN
      O=I+1
      P=J+1
      ENDIF
      IF (MOD(POT,2).EQ.0) THEN
      O=I-1
      P=J-1
      ENDIF
      PASO=PASO+1
      PASO1=1
      SUMANDO1=0

*
* ESTABLECEMOS LOS LIMITES DE VARIACION DE LOS INDICES QUE
* RECORRERAN LA MATRIZ EN FORMA DE ESPIRAL. ULTIMA PARTE DE LA
* BUSQUEDA.
*
      DO 40,M=1,VECES
      IF (PASO1.NE.0) THEN
      SUMANDO2=(-1)**PASO
      PASO1=0
      ENDIF
      SUMANDO1=SUMANDO1+SUMANDO2
      P1=P+SUMANDO1

      IF (O.GE.1.AND.O.LE.21.AND.P1.GE.1.AND.
1P1.LE.34) THEN
*      Z(INDICE1)=MATRIZ1(O,P1)
*      write(2,'(f6.2)')z(indice1)
*      INDICE1=INDICE1+1
*
* CALCULAMOS LAS DISTANCIAS A CADA UNO DE LOS PUNTOS ENCONTRADOS
*
      LAT(INDICE2)=35-O
      LONG(INDICE2)=-P1+120
      CONT1=CONT1+1
      INDICE2=INDICE2+1

      ENDIF

```

```

40    CONTINUE

*
*    ACTUALIZAMOS LAS VARIABLES QUE NOS PERMITIRAN CONTINUAR CON LA
*    BÚSQUEDA EN ESPIRAL
*
      LIM_ANT1=LIM_ANT1+INCREM1
      LIM_ANT2=LIM_ANT2+INCREM2
10    CONTINUE

*
*    TERMINA EL CÓDIGO QUE PERMITE QUE LA BÚSQUEDA SEA EN ESPIRAL
*****

      DO 122,INDICE3=2,CONT1+1

        IF (Z(INDICE3-1).NE.0.1) THEN
          DLAT(INDICE3-1)=ABS(LAT(1)-LAT(INDICE3))
122      ENDIF

        DO 123,INDICE3=2,CONT1+1
          IF (Z(INDICE3-1).NE.0.1) THEN

            DLON(INDICE3-1)=ABS(LONG(1)-LONG(INDICE3))
123      ENDIF

*
*    CALCULAMOS LAS DISTANCIAS QUE EXISTEN ENTRE CADA UNO DE LOS PUNTOS
*    ENCONTRADOS CON RESPECTO A LA POSICIÓN DE LA VARIABLE QUE SE DESEA
*    ESTIMAR
*
      DO 124,INDICE3=1,CONT1-1
        D(INDICE3)={(DLAT(INDICE3)**2+DLON(INDICE3)**2)**.5)**2
124      CONT=CONT+1
*
*    CALCULAMOS LOS INVERSOS DE LAS DISTANCIAS Y PAULATINAMENTE LOS
*    VAMOS SUMANDO
*
      DO 125,INDICE3=1,CONT1
        IF (D(INDICE3).NE.0) THEN
          D(INDICE3)=1.0/D(INDICE3)
          SUMRECIPR=SUMRECIPR+D(INDICE3)
        ENDIF
125      CONTINUE

*****

      DO 52,P=1,CONT1
52      SUMDIST=SUMDIST+D(P)
*
*    CALCULAMOS EL NUMERADOR DE LA EXPRESIÓN QUE NOS PERMITIRÁ ESTIMAR
*    LAS VARIABLES DESCONOCIDAS
*
      DO 54,Q=1,CONT1
        IF (D(Q).NE.0) THEN

```

```

        INTERP=INTERP+(Z(Q)*D(Q))
        ENDIF
54    CONTINUE
*
*   ASIGNAMOS A LA MATRIZ EN LA LOCALIDAD CORRESPONDIENTE EL VALOR
*   ESTIMADO, ESE MISMO NUMERO LO ESCRIBIMOS EN EL ARCHIVO PRO_INT
*   ACOMPAÑADO DE INFORMACION ADICIONAL QUE PERMITA SABER LA
*   LOCALIZACION GEOGRAFICA PARA ESA VARIABLE.
*
        IF (SUMDIST.NE.0) THEN
            MATRIZ2(RENG,COL)=(INTERP/SUMDIST)
        ENDIF
*
        ELSE
            WRITE(*,*) 'DIFERENTE DE CERO',RENG,COL,MATRIZ1(RENG,COL)
        ENDIF

300    CONTINUE
200    CONTINUE
*
*   CERRAMOS LOS CICLOS QUE PERMITEN RECORRER TODA LA MATRIZ
*
*****
*
*   ESCRIBIMOS LA NUEVA MATRIZ, CON TODAS SUS VARIABLES ESTIMADAS
*
        DO 55, L=1,22
55    WRITE(4,70,IOSTAT=K) (MATRIZ2(L,M),M=1,35)

        DO 61, L=1,22
61    WRITE(4,70,IOSTAT=K) (MATRIZ1(L,M),M=1,35)

        WRITE(*,*)
        DO 62,L=2,21
        DO 63,M=2,34
63    WRITE(2,*)MATRIZ2(L,M)
62    CONTINUE

        CLOSE(1)
        CLOSE(2)
        CLOSE(4)
        STOP
99    END

```

```

*****
*PROGRAMA QUE DIBUJA MAPAS (DIBMAP.FOR): LOS ARCHIVOS INDISPENSABLES PARA
* LA EJECUCION DEL PROGRAMA ESTAN CONSTITUIDOS POR PAREJAS DE NUMEROS
* (LONGITUDES Y LATITUDES REALES) Y ESTOS SON REPMEX.BLN, PARA DIBUJAR EL
* CONTORNO DE LA REPUBLICA MEXICANA; PNTEEUU1.POS, PARA DIBUJAR LA PARTE
* ESTE DE EE.UU, CANADA, ETC.; Y PUNTEEUU.BAK PARA LA PARTE OESTE DE ESTA
* ULTIMA REGION. OTRO ARCHIVO QUE TAMBIEN DESEMPEÑA UN PAPEL IMPORTANTE
* ES DAT MAP, PUES DE EL SE LEE LA INFORMACION NECESARIA PARA PODER
* DIBUJAR LA UBICACION DE LOS OBSERVATORIOS O BIEN, CON ALGUNAS PEQUEÑAS
* MODIFICACIONES AL CODIGO PRESENTADO, PLASMAR EN LA PANTALLA EL VALOR
* NUMERICO DEL PARAMETRO DECODIFICADO.
* EN REALIDAD SE PUEDEN IMPLEMENTAR DIVERSAS VARIACIONES, PUES
* ADEMAS DE LO YA MENCIONADO, ES POSIBLE MODIFICAR EL TIPO DE PROYECCION
* UTILIZANDO LAS TRANSFORMACIONES ADECUADAS (A COORDENADAS RECTANGULARES
* POR EJEMPLO) O MEDIANTE UNA ADECUADA LECTURA DEL ARCHIVO PRODUCTO DE
* DECODI.FOR, PRO_INT, CUBRIR LOS MAPAS CON TODOS LOS VALORES ESTIMADOS
*

```

```

INCLUDE 'FGRAPH.FI'
INCLUDE 'FGRAPH.FD'
INTEGER*2 DUMMY,DUMMY2,ELECCION
INTEGER*2 VECES,CONT,VENTANA

```

```

*****
* PROGRAMA PRINCIPAL
*****

```

```

CALL CLEARSCREEN($GCLEARSCREEN)
DUMMY2=SETVIDEOMODE($ERESCOLOR)
DUMMY=SETBKCOLOR($RED)
DO 1,CONT=1,13
  WRITE(*,*)
  VECES=1
  WRITE(*,'(A\)' )' MAPA DESEADO: 1) MEXICO 2) 4ª REGION ? '
  READ(*,'(I1)')ELECCION
  IF (ELECCION.EQ.1) THEN
    VENTANA=1
    CALL DIBUJA("C:REPMEX.BLN",12,VENTANA,VECES)
    CALL DIVISION()
    CALL OBSERVAT()
  ELSE
    VENTANA=2
    CALL DIBUJA("C:REPMEX.BLN",12,VENTANA,VECES)
    CALL DIBUJA("PNTEEUU1.POS",13,VENTANA,VECES)
    CALL DIBUJA("PUNTEEUU.BAK",13,VENTANA,VECES)
    CALL DIVISION()
    CALL OBSERVAT()
  ENDIF
  READ(*,*)
  DUMMY = SETVIDEOMODE ($DEFAULTMODE)
STOP
END

```

```

*****

```

```

SUBROUTINE DIBUJA(ARCHIVO,TONE,VENTANA,VECES)

```

```

* SUBROUTINA QUE: ESTABLECE LOS LIMITES DE LA VENTANA EN LA QUE SE HARAN

```

* LOS DIBUJOS EN FUNCION A LA INFORMACION PROPORCIONADA POR EL USUARIO,
 * LEE LOS DATOS DE LOS ARCHIVOS CORRESPONDIENTES Y EFECTUA, ES ESTE
 * CASO, UNA TRANSFORMACION QUE EMULE EL "COMPORTAMINETO REAL" DE LA
 * SITUACION GEOGRAFICA DE LAS REGIONES PARA, FINALMENTE, DIBUJAR LOS
 * MAPAS SELECCIONADOS
 *

```

INCLUDE      'FGRAPH.FD'
CHARACTER*1  COMMA
CHARACTER*12 ARCHIVO
REAL         LONG,LAT
INTEGER*2    DUMMY4,I,DUMMY,VENTANA,VECES
INTEGER*4    TONE
REAL         LONGI,LATI
RECORD /WXYCOORD/ XY
  
```

```

IF (VECES.EQ.1) THEN
  DUMMY2=SETVIDEOMODE($ERESCOLOR)
  DUMMY3=SETBKCOLOR($blue)
  VECEs=2
ENDIF
IF (VENTANA.EQ.1) THEN
  DUMMY=SETWINDOW(.TRUE.,-119.0,15.2,-82.0,35.2)
ELSE
  DUMMY=SETWINDOW(.TRUE.,-135.0,12.0,-50.0,60.0)
ENDIF
OPEN(1,FILE=ARCHIVO)
I=1
10  READ(1,11,END=99) LONG,COMMA,LAT
11  FORMAT(F8.3,A1,F8.4)
    LONGI=-100+(89-LAT)*(SIN((LONG+100)*3.14159/180))
    LATI=89-(89-LAT)*COS((LONG+100)*3.14159/180)
    DUMMY4=SETCOLOR(TONE)
    IF (I.EQ.1) THEN
      CALL MOVETO_W(LONGI,LATI,XY)
      I=I+1
    ELSE
      DUMMY = LINETO_W(LONGI,LATI)
    ENDIF
    GOTO 10
99  RETURN
    END
  
```

SUBROUTINE DIVISION()

* SUBROUTINA QUE LEE DEL ARCHIVO DIVIS.DAT, LA INFORMACION REFERENTE
 * A LA DIVISION GEOGRAFICA DEL LAS REGIONES CONSIDERADAS EN CUANTO
 * A LATITUDES Y LONGITUDES REALES SE REFIERE CON UN ESPACIAMIENTO
 * DE 5 GRADOS

```

INCLUDE      'FGRAPH.FD'
INTEGER      A,B
REAL*4      TR1,TR2
CHARACTER*1  COMMA

70  OPEN(4,FILE='C:DIVIS.DAT')
80  READ(4,80,END=98) A,COMMA,B
    FORMAT(I4,A1,I3)
  
```

```

DUMMY4=SETCOLOR(15)
TR1=-100+(90-B)*SIN((A+100)*3.14159/180)
TR2=90-(90-B)*COS((A+100)*3.14159/180)
DUMMY=SETPIXEL_W(TR1,TR2)
* DUMMY=SETPIXEL_W(A,B)
IF (MOD(B,5).EQ.0) THEN
DO 90, K=1,90
TR3=-100+(90-B)*SIN(((A+K)+100)*3.14159/180)
TR4=90-(90-B)*COS(((A+K)+100)*3.14159/180)
90 DUMMY = SETPIXEL_W(TR3,TR4)
*90 DUMMY = SETPIXEL_W(A+K,B)
ENDIF
GOTO 70
98 RETURN
END

```

SUBROUTINE OBSERVAT()

* SUBROUTINA QUE UTILIZA LA INFORMACION CONTENIDA EN EL ARCHIVO DAT MAP
* PARA DIBUJAR, EN UNA POSICION REAL, LOS OBSERVATORIOS QUE PARA EL
* DIA CONSIDERADO ENVIARON SUS MENSAJES SINOPTICOS
*

```

INCLUDE 'FGRAPH.FD'
REAL TRLO,TRLA,LA,LO,PARAM
RECORD /WXYCOORD/ XY

```

```

98 DUMMY2=SETCOLOR(5)
OPEN(6,FILE='C:DAT MAP')
71 READ(6,*,END=96) LA,LO,PARAM
*80 FORMAT(I4,A1,I3)
TRLO=-100+(90-LA)*SIN((LO+100)*3.14159/180)
TRLA=90-(90-LA)*COS((LO+100)*3.14159/180)
DUMMY=ELLIPSE_W($GFILLINTERIOR,TRLO-.1,TRLA-.1,TRLO+.1,
1TRLA+.1)
GOTO71
96 RETURN
END

```

(* PROGRAMA QUE GENERA LOS NUMEROS ALEATORIOS, DOMINIO, CONTRADOMINIO E IMAGEN, CON LOS QUE PROBAREMOS LA EFICIENCIA DEL METODO QUE IMPLEMENTAMOS *)

```

program numalea(input,output);
uses crt;
var
  x,y,x1,y1,iaux,yaux:real;
  z:double;
  i:integer;
  archivo,archivol:text;

begin
  assign(archivo,'c:\surfer\6fun.dat');
  rewrite(archivo);
  assign(archivol,'6fun1.sal');
  rewrite(archivol);
  clrscr;
  for i:=1 to 80 do
    begin
      randomize;
      x:=2+random(32);
      y:=2+random(19);

      {z:=0.75*exp(-((9*(x-16)-2)*(9*(x-16)-2)+(9*(y-10)-2)*(9*(y-10)-2)
      +0.75*exp(-((9*(x-16)+1)*(9*(x-16)+1)/49-(9*(y-10)+1)/10)
      +0.50*exp(-((9*(x-16)-7)*(9*(x-16)-7)+(9*(y-10)-3)*(9*(y-10)-3)
      -0.20*exp(-((9*(x-16)-4)*(9*(x-16)-4))-((9*(y-10)-7)*(9*(y-10)-
      )

    { FUNCION 1 O.k.
      XAUX:=X/100;
      YAUX:=Y/100;
      z:=0.75*exp(-((9*iaux-2)*(9*iaux-2)+(9*yaux-2)*(9*yaux-2))/4)
      +0.75*exp(-((9*iaux+1)*(9*iaux+1)/49-(9*yaux+1)/10)
      +0.50*exp(-((9*iaux-7)*(9*iaux-7)+(9*yaux-3)*(9*yaux-3))/4)
      -0.20*exp(-((9*iaux-4)*(9*iaux-4))-((9*yaux-7)*(9*yaux-7)));
    }

    { FUNCION 3 O.k.
      z:=(1.25+cos(5.4*(y-10)))/6*(1+(3*(x-16)-1)*(3*(x-16)-1));
    }

    { FUNCION 2 O.k.
      if (exp(9*y-9*x)-exp(9*x-9*y)) <> 0 then
        begin
          z:=(1/9)*((exp(9*y-9*x)+exp(9*x-9*y))/(exp(9*y-9*x)-exp(9*x-
          end;
        }

    { FUNCION 6 O.k.
      XAUX:=X/34;
      YAUX:=Y/21;
      z:=(1/9)*SQRT(64-81*((XAUX-0.5)*(XAUX-0.5)+(YAUX-0.5)*(YAUX-0.5))
    }
  
```

```

x1:=120-x;
y1:=35-y;
if z<0 then
begin
  if ((x<=9) and (y<=9)) then
    writeln(archivo,' ',y:3:2,' ',x:2:2,',',' ',z:4:1)
  else
    if ((x<=9) and (y>9)) then
      writeln(archivo,' ',y:3:2,' ',x:2:2,',',' ',z:4:1)
    else
      if ((x>9) and (y<=9)) then
        writeln(archivo,' ',y:3:2,' ',x:2:2,',',' ',z:4:1)
      else
        writeln(archivo,' ',y:3:2,' ',x:2:2,',',' ',z:4:1);
  end
else
begin
  if ((x<=9) and (y<=9)) then
    writeln(archivo,' ',y:3:2,' ',x:2:2,',',' ',z:4:1)
  else
    if ((x<=9) and (y>9)) then
      writeln(archivo,' ',y:3:2,' ',x:2:2,',',' ',z:4:1)
    else
      if ((x>9) and (y<=9)) then
        writeln(archivo,' ',y:3:2,' ',x:2:2,',',' ',z:4:1)
      else
        writeln(archivo,' ',y:3:2,' ',x:2:2,',',' ',z:4:1);
  end;
if z<0 then
begin
  if (x1<=99) then
    writeln(archivo1,' ',y1:3:2,' ',x1:2:2,',',' ',z:4:1)
  else
    if (x1>99) then
      writeln(archivo1,' ',y1:3:2,' ',x1:2:2,',',' ',z:4:1)
  end
else
begin
  if (x1<=99) then
    writeln(archivo1,' ',y1:3:2,' ',x1:2:2,',',' ',z:4:1)
  else
    if (x1>99) then
      writeln(archivo1,' ',y1:3:2,' ',x1:2:2,',',' ',z:4:1)
  end;
writeln(' ',y:2:2,' ',x:3:2,' ',z:5:2);
delay(100)
end;
close(archivo);
close(archivo1);
end.

```

* PROGRAMA QUE DA FORMATO A UNO DE LOS ARCHIVOS RESULTADO DE INTERP2.FOR
 * PARA QUE PUEDA SER LEIDO POR SURFER Y SE PUEDA GENERAR EL MAPA DE ISO-
 * LINEAS. (PROGRAMA QUE A PARTIR DE LOS DATOS QUE INTERPOLAMOS ARMA EL
 * "ARCHIVO-MALLA" (GREED) QUE SURFER RECONOCE PARA TRAZAR LAS ISOLINEAS.

```

integer k,j
real imagen,min,max
dimension imagen(800)

k=1
j=0
m=1
min=100.0
max=0.0
open(2,file='C:\SURFER\funcion.GRD')
open(1,file='funcion.LIS')
1 read(1,*,end=99)imagen(k)

10 format(f8.2)
WRITE(*,*)'VALORES DE IMAGEN K ',IMAGEN(K)
if (imagen(k).le.min) then
  min=imagen(k)
endif
if (imagen(k).ge.max) then
  max=imagen(k)
endif
k=k+1
goto 1

99 write(2,16)'DSAA'
write(2,*)' 33 20'
write(2,*)' -118 -86'
write(2,*)' 14 33'
write(2,19)' ',MIN,MAX
16 FORMAT(A4)
19 FORMAT(A3,F6.2,F6.2)

do while(m.le.66)
write(2,20)(imagen(k+j),k=1,10)
if (mod(m,3).eq.0) then
write(2,*)' '
endif
20 format(10f10.2)
j=j+10
m=m+1
enddo

write(*,*)'este es el minimo ',min
write(*,*)'este es el maximo ',max
stop
end

```

* PROGRAMA QUE LEE EL ARCHIVO DE DATOS INTERPOLADOS POR INTERP2.FOR Y
 * EL ARCHIVO DE DATOS INTERPOLADOS POR SURFER, RESTA CADA ESTIMACION CON
 * SU CORRESPONDIENTE EN EL OTRO ARCHIVO Y ARMA LA "MATRIZ DIFERENCIA"
 * QUE HEMOS UTILIZADO PARA PROBAR LA EFICIENCIA Y BUEN DESEMPEÑO DEL
 * PROGRAMA DE INTERPOLACION INTERP2.FOR.
 *

```

INTEGER    I,J,K,IN1,IN2
REAL      MATRIZ,SUMDIF,DIFMIN,VECES
REAL      VAL2,PATRON,DIF,DIFMAX,VECES1
DIMENSION MATRIZ(22,35)

OPEN(1,FILE='arcos.LIS')
OPEN(2,FILE='C:\SURFER\CONVarc.DAT')
OPEN(3,FILE='DIFerenc.ias')
OPEN(4,FILE='DIFarc.MAT')
SUMDIF=0
DIFMIN=10
DIFMAX=0
VECES=0
VECES1=0

DO 1,I=2,21
  DO 2,J=2,34

    READ(1,*)VAL2
    READ(2,*)IN1,IN2,PATRON
    DIF=ABS(PATRON-VAL2)
    IF (DIF.GT.2) THEN
      VECEs=VECEs+1
    ENDIF
    IF (DIF.GT.3) THEN
      VECEs1=VECEs1+1
    ENDIF
    IF (DIF.LE.DIFMIN) THEN
      DIFMIN=DIF
    ENDIF
    IF (DIF.GE.DIFMAX) THEN
      DIFMAX=DIF
    ENDIF
    SUMDIF=SUMDIF+DIF
    WRITE(3,*)IN1,IN2,DIF
    MATRIZ(I,J)=DIF
  2 CONTINUE
1 CONTINUE
DO 3,I=1,22
  WRITE(4,70,IOSTAT=K)(MATRIZ(I,J),J=1,35)
70 FORMAT(35(F4.1))
3 CONTINUE
WRITE(4,*)' '
WRITE(4,*)' '
WRITE(4,*)' '
WRITE(4,4)' VALOR PROMEDIO DE LAS DIFERENCIAS      ',SUMDIF/660
WRITE(4,4)' DIFERENCIA MAXIMA                      ',DIFMAX
WRITE(4,4)' DIFERENCIA MINIMA                      ',DIFMIN
WRITE(4,4)' NUMERO DE VECES QUE LA DIFERENCIA > 2  ',VECES
WRITE(4,4)' NUMERO DE VECES QUE LA DIFERENCIA > 3  ',VECES1

```

```
4  FORMAT(A43,F10.7)
   STOP
   END
```

* PROGRAMA QUE A PARTIR DEL ARCHIVO DE DATOS INTERPOLADOS POR SURFER,
* ARMA UNA MATRIZ PARA TENER UNA REPRESENTACION COMPACTA DE ELLOS Y
* PODER HACER COMPARACIONES CON LA MATRIZ DE INTERPOLACIONES GENERADA
* POR INTER2.FOR.

```
      INTEGER      LONG,LAT,CONT
      REAL         MATRIZ,IMAGEN
      DIMENSION   MATRIZ(22,35)

      DO 39,L=1,22
DO 49,M=1,35
      MATRIZ(L,M)=0
49      CONTINUE
39      CONTINUE
      CONT=0
      OPEN(1,FILE='C:\SURFER\FUNCONV.DAT')
      OPEN(2,FILE='FUNCONV.MAT')
1      READ(1,*,END=99)LONG,LAT,IMAGEN
      IF (LONG.GE.86.AND.LAT.GE.14) THEN
          LONG=120-LONG
          LAT=35-LAT
      ENDIF
      CONT=CONT+1
10     FORMAT(I2,I2,F7.1)
      WRITE(*,*)LONG,LAT,IMAGEN,CONT
      MATRIZ(LAT, LONG)=IMAGEN
      GOTO 1
99     DO 50,L=1,22
          WRITE(2,70,Iostat=k) (MATRIZ(L,M),M=1,35)
70     FORMAT(35(F5.1))
50     CONTINUE
      STOP
      END
```

BIBLIOGRAFIA

Dubrule O., 1983, Two Methods with different objectives: Splines and Kriging: *Journal Mathematical Geology* , V.15, pp.245-257.

Journel A., and Huijbregts, C.J., 1978, *Mining Geostatistics*: Academic Press, London.

G.S.Watson, 1983, Smoothing and Interpolation by Kriging and with Splines: *Journal Mathematical Geology*, V.16, pp.601-615.

Franke Richard and Greg Nielsen, 1979, Smooth Interpolation of Large Sets of Scattered Data, Naval Postgraduate School Technical Report, NPS-53-79-005

Franke Richard, 1979, A Critical Comparison of Some Methods for Interpolation of Scattered Data, Naval Postgraduate School Technical Report, NPS-53-79-005

Brian D. Ripley, 1981, *Spatial Statistics*, John Wiley and Sons.

Sabin, M.A., 1985, *Countouring*, NATO A.S.I. On Fundamental Algorithms for Computer Graphics, Ilkley U.K.

Dowd. P.A., 1985, A Review of Geostatistical Techniques for Contouring, NATO A.S.I. On Fundamental Algorithms for Computer Graphics, Ilkley U.K.