

Precipitación

Temperatura

Temperatura máxima

Temperatura mínima

Evaporación

Cobertura de cielo

Tormentas eléctricas

Granizo

Evaporación

Cobertura de cielo

Tormentas eléctricas

Niebla

Granizo

Isabel Quintas

Eric II[©]

Documentación de la base de datos
climatológica y del programa extractor



IMTA
INSTITUTO MEXICANO
DE TECNOLOGÍA DEL AGUA

34

1

Precipitación

Temperatura

Temperatura máxima

Temperatura mínima

ERIC II

DOCUMENTACIÓN DE LA BASE DE DATOS
CLIMATOLÓGICA Y DEL PROGRAMA EXTRACTOR

Isabel Quintas

ERIC II

Documentación de la base
de datos climatológica y del
programa extractor

Manuales

IMTA

Coordinación de Desarrollo Profesional e Institucional
Subcoordinación de Desarrollo Institucional

Mayo, 2000

**INSTITUTO MEXICANO DE TECNOLOGÍA DEL AGUA
CENTRO DE CONSULTA DEL AGUA**

Clasif. G 551.634

Q47

21381

C.B. 19955

Proced. DONACIÓN

Fecha 27.07.05

551.634
Q47

Quintas, Isabel
ERIC II, Base de datos climatológica compactada, archivos y programa extractor / Isabel Quintas.- México: Instituto Mexicano de Tecnología del Agua, 2000.
54 p.; 17 x 21.5 cm (Manuales)

ISBN 968-7417-40-4

1. Informática. 2. Climatología-Base de datos-Programa extractor.

Revisión literaria:
Antonio Requejo.

Diseño de portada:
Gema Alín Martínez.

Diseño de interiores e ilustraciones:
Rosario Castro.

Formación:
Luisa Guadalupe Ramírez

Comercialización y gestión de tecnología:
Guillermo Laríos de Anda

Primera edición: 2000.

© Instituto Mexicano de Tecnología del Agua, 2000
Paseo Cuauhnáhuac 8532,
62550, Progreso, Morelos.

ISBN 968-7417-40-4

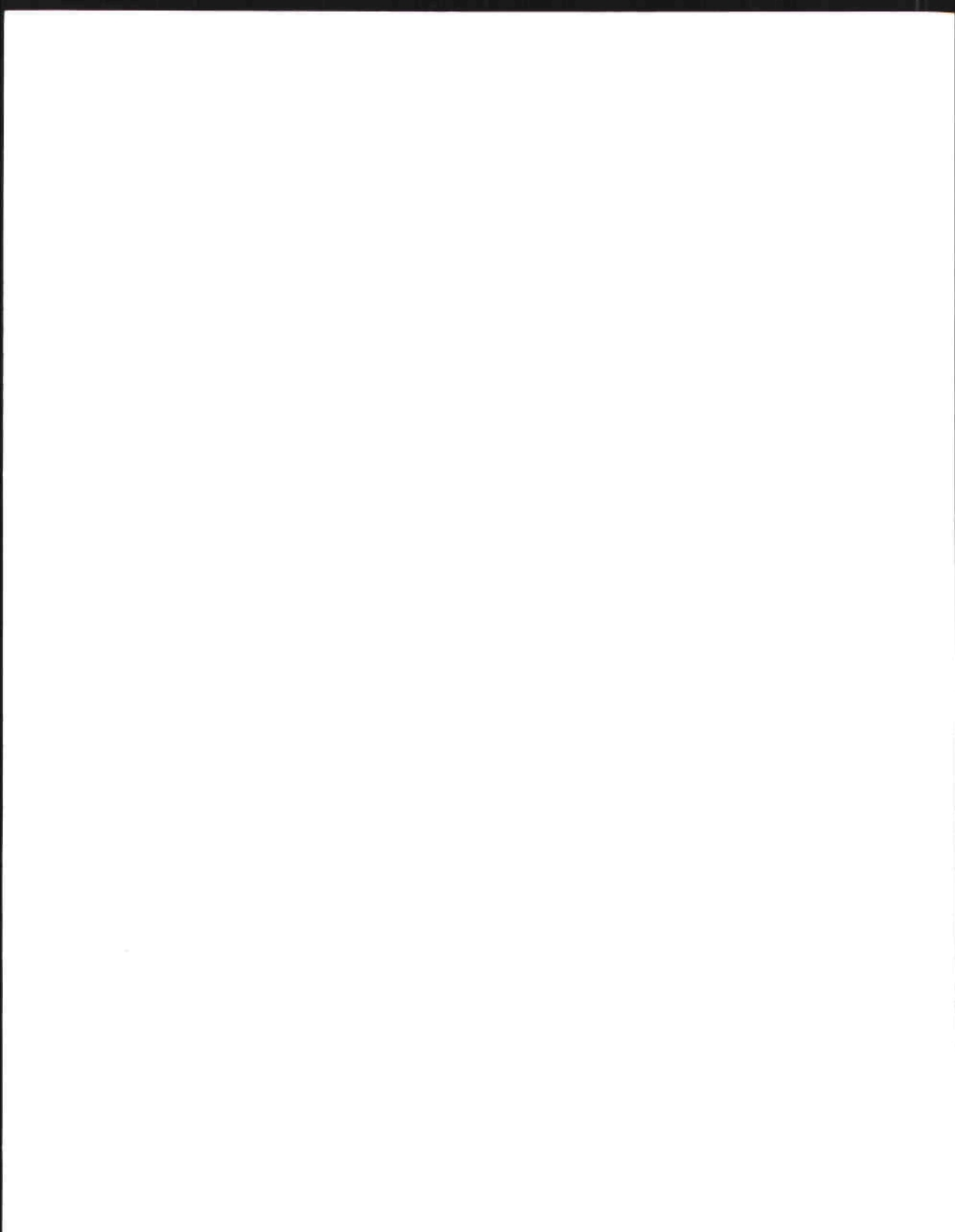
Hecho en México – *Printed in Mexico*

AGRADECIMIENTOS

El presente documento partió de la información existente en la base de datos *Clicom*, de la Gerencia del Servicio Meteorológico Nacional de la Comisión Nacional del Agua, instancia que facilitó los archivos originales. En particular, queremos agradecer la valiosa colaboración del gerente técnico del Servicio Meteorológico Nacional, ingeniero Carlos Espinosa González, quien con su apoyo posibilitó la labor.

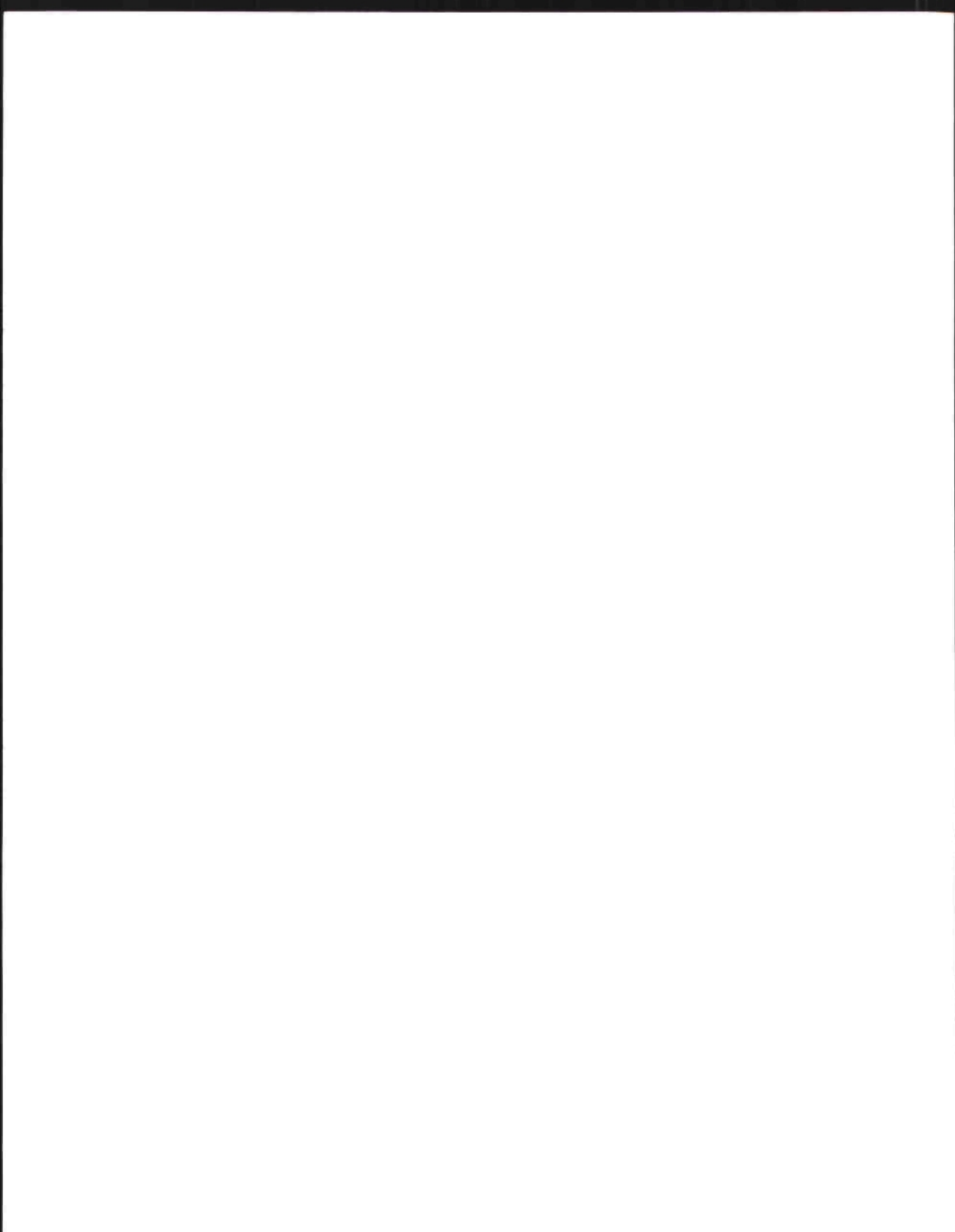
También, de manera especial, queremos agradecer a las gerencias técnicas de los estados de: Aguascalientes, Baja California Sur, Campeche, Colima, Durango, Guanajuato, Guerrero, Estado de México, Morelos, San Luis Potosí, Veracruz y Zacatecas por la seriedad y responsabilidad con que realizaron la recolección, mantenimiento y actualización de sus respectivas bases de datos. Asimismo, se incluyó la información parcial enviada desde los estados de: Chiapas, Puebla, Querétaro, Quintana Roo, Sinaloa y Yucatán.

En un trabajo de este tipo parece imposible terminar de encontrar errores. Algunas fallas de importancia fueron detectadas por el doctor Víctor Magaña y su grupo de alumnos del Instituto de Ciencias de la Atmósfera, de la Universidad Nacional Autónoma de México, las que a tiempo pudieron ser corregidas. Mención especial merece el doctor Lorrain Giddings, del Instituto de Ecología de Xalapa, Ver., quien ha efectuado y continúa un minucioso estudio de los datos, señalando varios problemas, ya aquí documentados.



ÍNDICE

INTRODUCCIÓN	11
1 DESCRIPCIÓN DE LA INFORMACIÓN CONTENIDA EN LA BASE DE DATOS <i>ERIC II</i>	13
2 TIPOS DE VARIABLES REGISTRADAS	19
3 SISTEMA DE CONSULTA A LA BASE DE DATOS	23
3.1 <i>Uso del sistema</i>	25
3.2 <i>Presentación de resultados</i>	27
4 ARCHIVOS ORIGINALES DEL <i>CLICOM</i>	31
4.1 <i>Archivo XXXXX.DBM</i>	33
4.1.1 Encabezado	34
4.1.2 Datos	35
4.1.3 Banderas	35
4.2 <i>Archivos de la base de datos ERIC</i>	35
4.2.1 Encabezado	36
4.2.2 Archivos de temperaturas: <i>TEMccc.BD</i>	37
4.2.2.1 Campo de valores del archivo: <i>TEMccc.BD</i>	37
4.2.3 Archivos de precipitaciones: <i>LLUccc.BD</i>	38
4.2.3.1 Campo de valores del archivo: <i>LLUccc.BD</i>	38
4.2.4 Archivos de variables cualitativas: <i>BOOccc.BD</i>	38
4.2.4.1 Campo de valores del archivo: <i>BOOccc.BD</i>	38
4.2.5 Archivos de evaporaciones: <i>EVAccc.BD</i>	39
4.2.5.1 Campo de valores del archivo: <i>EVAcccBD</i>	40
4.2.6 Archivos índices	40
5 ORGANIZACIÓN DE LA NUEVA BASE DE DATOS	43
5.1 <i>Directorios</i>	45
APÉNDICE 1 Procedimientos de detección de errores realizados por el SMN	47
APÉNDICE 2 Errores aún presentes en <i>ERIC II</i>	51
APÉNDICE 3 Estado de actualización en cada entidad federativa	55



ÍNDICE DE CUADROS

Cuadro 1	Clave asignada por <i>Clicom</i> a las variables meteorológicas.	29
Cuadro 2	Referencia de los estados.	29
Cuadro 3	Valores de precipitación y temperatura utilizados para la validación de la información.	50



INTRODUCCIÓN

En 1994 el Instituto Mexicano de Tecnología del Agua (IMTA) decidió crear una herramienta sustituta del sistema *Clicom*, que almacenaba la base de datos climatológica nacional, ya que éste no satisfacía las necesidades de información para los diferentes proyectos de investigación del Instituto. *Clicom* era un sistema que la Organización Meteorológica Mundial había diseñado para los países pequeños, cuando comenzaron a surgir las microcomputadoras.

El sistema realizado por el IMTA, conocido como *ERIC (Extractor Rápido de Información Climatológica)*, representó una herramienta mucho más adecuada a las necesidades de consulta, de un gran volumen y detalle de datos. Consistía en la base de datos organizados y ordenados de manera eficiente, conteniendo toda la información de la base de datos nacional, con un programa para búsquedas y para extraer la información de interés.

Como el número de usuarios iba en aumento, se hizo una edición de trescientos ejemplares en disco compacto. Los diferentes usuarios comenzaron a solicitar la actualización de los datos, debido a que la información recopilada en el *ERIC* finalizaba entre los años 1985-1990. Además se detectaron muchos errores en los datos, y con la evolución de las plataformas para PC, había la necesidad de emigrar a *Windows*.

En 1998 se decidió llevar a cabo dicha actualización, para lo cual se comenzaron a recopilar los datos en cada estado, ya que a partir de 1990 la Gerencia del Servicio Meteorológico Nacional (GSMN), había descentralizado tal actividad y no había integrado la información de los diferentes estados. Se encontró una gran disparidad en el avance de la actualización. Después de un año de efectuar esta tarea, se reunió la información que aquí se presenta, se hizo un diagnóstico de la situación y se ha propuesto una serie de tareas para capturar la información faltante a través de un proyecto conjunto con la Gerencia del Servicio Meteorológico Nacional.

El motivo de haber editado esta versión con la actualización inconclusa fue la demanda permanente de datos que nos hacen al IMTA los distintos usuarios, tanto de universidades como del sector público y privado. La cantidad de información agregada es relevante, ya que se sumaron 15,000 registros anuales de datos,¹ a los ochenta mil ya existentes. En algunos casos hay datos hasta 1998, en otro sólo llegan a 1995 o 1996. En algunos estados no existe información adicional. En la siguiente figura se muestra la información sobre cada estado, medida en registros anuales, existente en la versión anterior y los datos agregados.²

¹ Un registro anual corresponde a los datos de los 365 días del año de información de nueve variables en una estación; esto es más de 1,500 valores numéricos.

² El estado de Veracruz está actualizado hasta 1997; esto no se refleja en la gráfica porque en dicho estado no se ha terminado el proceso de validación de los datos y aún no habían sido entregados por el SMN al momento de la evaluación.

Número de años-estación en ERIC I y ERIC II

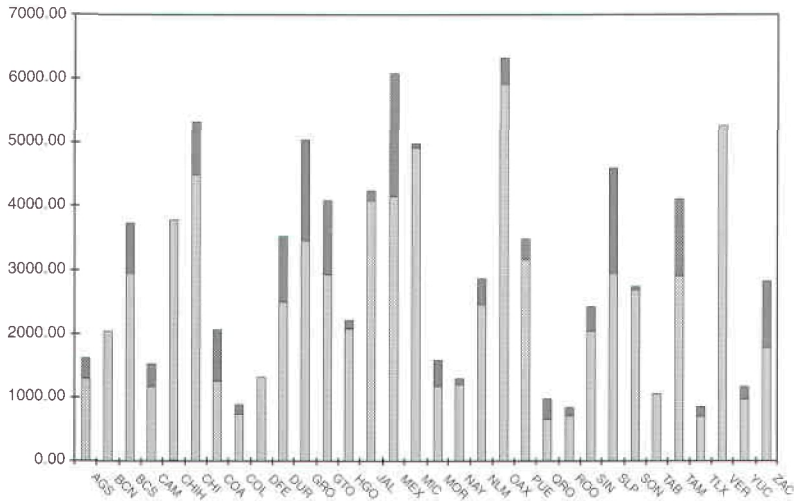
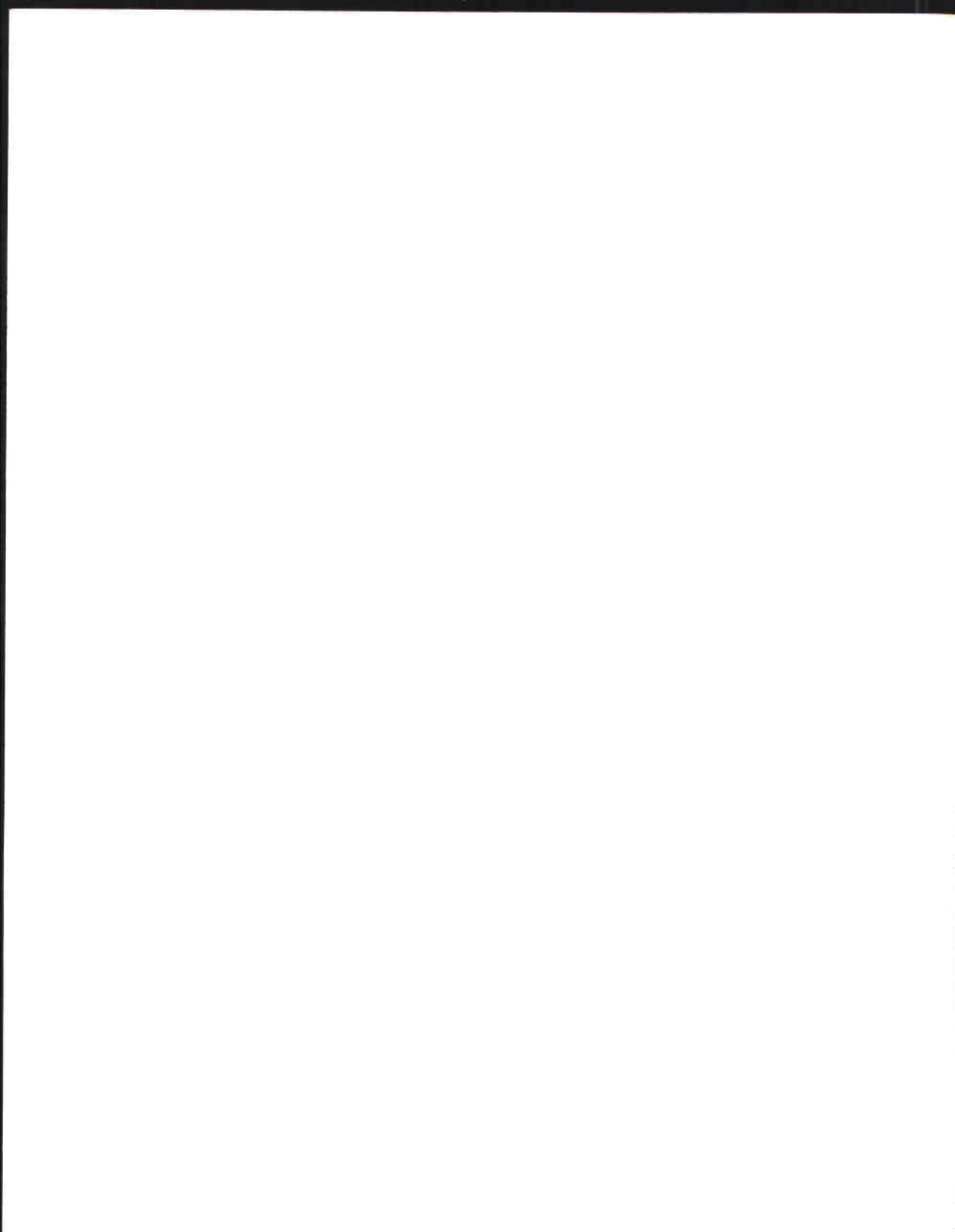


Figura 1. Número de años-estación en ERIC I y ERIC II.

Hubo necesidad de publicar la información y se presenta el estado de la base de datos nacional tal como se encontraba en enero del año 2000. En este momento, ya tenemos más datos que no han podido ser incluidos en la presente versión, y éstos irán aumentando en el transcurso del año. Toda base de datos es dinámica y una edición en disco compacto contendrá una imagen de dicha base congelada en un momento determinado. Los usuarios podrán consultar los datos más recientes en la base de datos del IMTA o en la del Servicio Meteorológico Nacional (SMN).

ERIC II es la nueva herramienta que trata de subsanar parte de las deficiencias de la anterior. Su presentación corresponde al estándar en plataformas de desarrollo: funciona bajo cualquier versión de *Windows*; no necesita ser instalado, por lo cual no requiere más que el espacio en disco para guardar los archivos de resultados de las búsquedas efectuadas. La salida respeta el formato de la versión anterior, ya que los usuarios han desarrollado herramientas para procesar dichos datos, pero agrega el cálculo de valores promedios mensuales en el caso de temperaturas, y los acumulados mensuales para las otras variables.

**1 DESCRIPCIÓN DE LA
INFORMACIÓN CONTENIDA
EN LA BASE DE DATOS
*ERIC II***



ERIC II contiene la información climatológica de la base de datos *Clicom* del SMN tal como se encontraba en diciembre de 1999. Se trata de la información medida por la red de estaciones climatológicas existentes en el país, que reportan el valor de nueve parámetros meteorológicos cada 24 horas. Esta red actualmente está a cargo de las gerencias estatales y regionales de la Comisión Nacional del Agua. Su mantenimiento, y especialmente la recopilación y digitalización de la información, no se realiza de la misma manera en cada una de las gerencias, por lo que esta base no es homogénea.

La red llegó a tener 5,000 estaciones, pero a partir de 1980 se produjo un deterioro importante debido a la falta de mantenimiento y al poco interés puesto por las instancias responsables, ya que se pretendió sustituir por una red automática de estaciones de monitoreo atmosférico, que nunca llegó a funcionar. En las figuras 2 y 3 se muestra el número de estaciones que se encontraban funcionando y de las que se conservan los datos, para cada estado, desde 1930 a 1999.

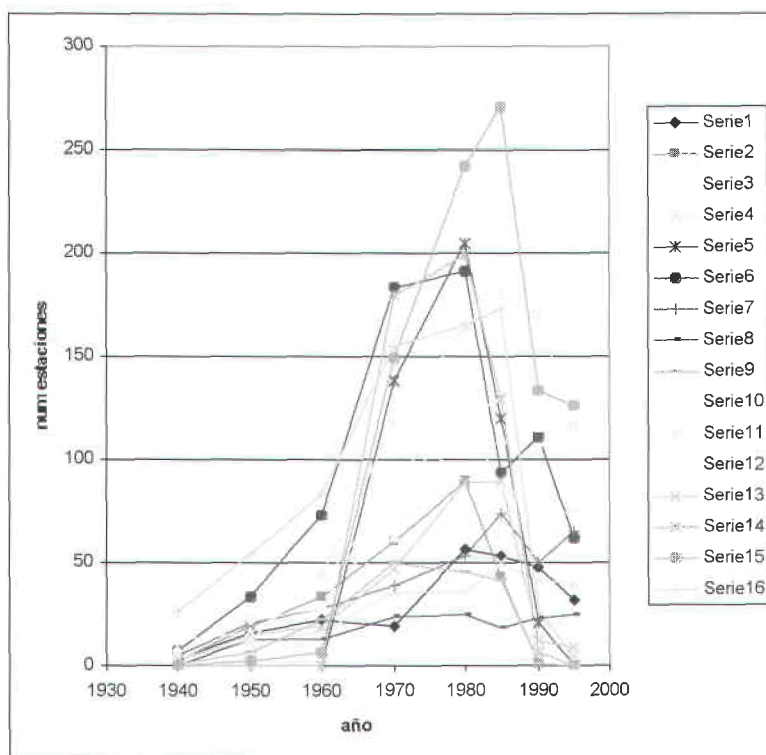


Figura 2. Estados del 1 (Aguascalientes), al 16 (Michoacán).

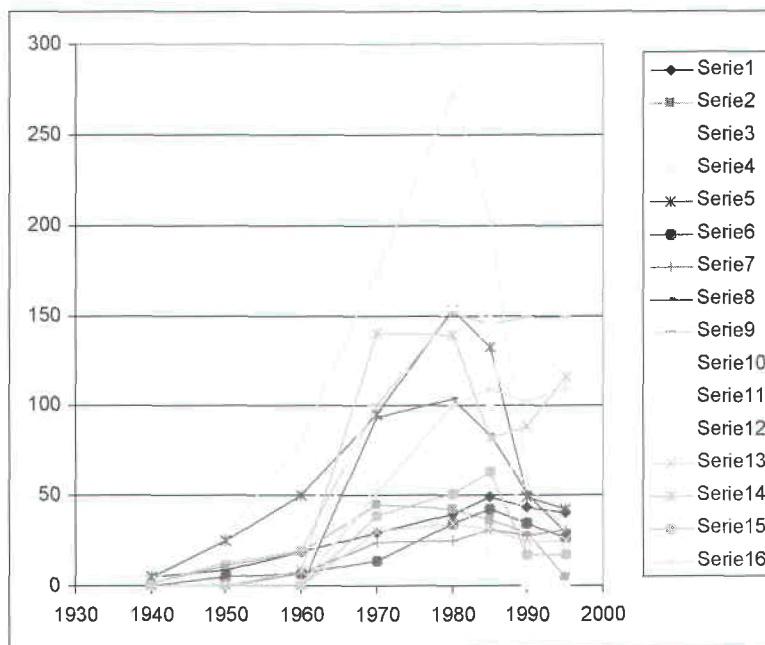


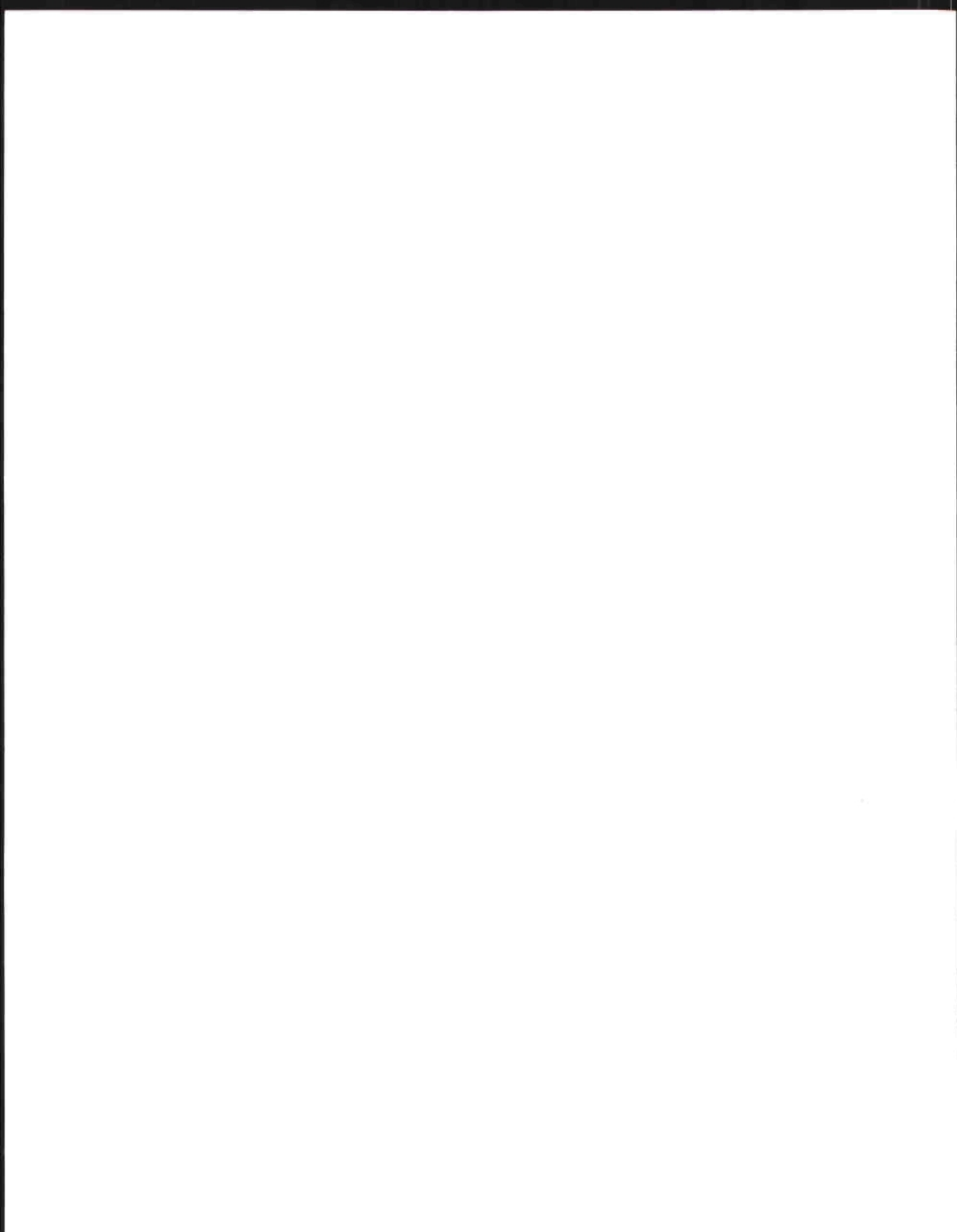
Figura 3. Estados del 17 (Morelos), al 32 (Zacatecas).

En 1990 la GSMN digitalizó la información existente para almacenarla y consultarla con el sistema *Clicom*. Se trataba de información de 5,000 estaciones, cuyos datos concluían entre 1985 y 1990. En algunos estados la información digitalizada comienza en 1960, aunque existan registros anteriores en papel. Esta información es la que se encuentra disponible en la versión anterior del *Extractor Rápido de Información Climatológica*.

A partir de 1990, se transfirió a las gerencias estatales y regionales el sistema *Clicom*, que desde ese momento debería ser actualizado en cada una de las regiones. Durante los últimos años el SMN realizó una tarea de depuración de los datos contenidos en *Clicom* hasta ese momento, con procedimientos muy sencillos que sólo permiten eliminar algunos errores.³ Esta tarea aún no ha finalizado.

³ Ver Apéndice 1: Procedimientos de detección de errores realizados por el SMN.

En el transcurso de 1999 el IMTA, con apoyo de la GSMN, comenzó a recabar la información que se había actualizado en cada estado. El IMTA también participó en la digitalización de datos de los estados de Coahuila y Nuevo León. Esta información se integró al sistema *Clicom* de la gerencia nacional y es la que se presenta en el disco compacto *ERIC II*. Se trata de información de 5,330 estaciones que, en algunos casos cubre hasta 1998, aunque en otros estados la situación no ha variado con respecto a la que se tenía en la versión anterior. También se actualiza el catálogo de estaciones. En el Apéndice 3 se muestra el estado de actualización en cada entidad federativa.



2 TIPOS DE VARIABLES REGISTRADAS



Las estaciones climatológicas que reportan datos cada 24 horas son asistidas por un observador que las visita alrededor de las 8.00 AM para realizar las mediciones y registrar las variables en un formato determinado. Se registran cinco variables cuantitativas: temperatura máxima y mínima de las 24 horas anteriores, temperatura observada al momento de la medición, lámina de lluvia acumulada y evaporación en 24 horas. Además, se registran cuatro variables cualitativas: estado del cielo y ocurrencia o no ocurrencia de tormenta, niebla y granizo.

La medición de las variables cuantitativas depende, básicamente, de la existencia en la estación del instrumento correspondiente; y la calidad de las mediciones, del estado de calibración del instrumento y la pericia del observador.

En el caso de las variables cualitativas, éstas son muy subjetivas dependiendo de la capacitación previa que se le haya dado al personal, influyendo también consideraciones locales.⁴ Por otro lado, los observadores codifican estos fenómenos por medio de símbolos que no siempre son suficientemente explícitos.

A continuación se especifica cada una de las variables:

- **Temperatura máxima:** se mide con un termómetro de máxima, en grados Celsius con un dígito decimal. Debido a que se lee alrededor de las 8.00 AM, la temperatura máxima registrada corresponderá a la temperatura máxima ocurrida posiblemente entre las 11.00 AM y las 5.00 PM del día anterior.
- **Temperatura mínima:** se mide con un termómetro de mínima, en grados Celsius con un dígito decimal. Debido a que se lee alrededor de las 8.00 AM, la temperatura mínima corresponde a la mínima ocurrida durante la madrugada del día de la lectura. Esta lectura siempre deberá ser menor o igual a la temperatura observada.
- **Temperatura observada:** se mide con un termómetro normal, en grados Celsius con un decimal. Aunque la observación se debe realizar a las 8.00 AM, esto puede variar ligeramente de una a otra estación de medición; sin embargo, este valor debe ser mayor o igual que la temperatura mínima y menor que la temperatura máxima.
- **Precipitación:** se trata de la lámina de lluvia acumulada en las 24 horas previas, esto es, de las 8.00 AM del día anterior, a las 8.00 AM de ese día. Dependiendo del régimen de lluvias de la región, esta medición puede corresponder a la lluvia caída la tarde anterior, o si se trata de lluvias en la madrugada, podría corresponder a la lámina caída durante la madrugada del día de la medición. Se mide en milímetros con un dígito decimal.

⁴ Lo que en una región puede ser considerado *totalmente nublado*, en otra puede ser visto como *medio nublado*.

- **Evaporación:** se mide la lámina de evaporación ocurrida en las 24 horas anteriores en un evaporímetro de tanque. La lámina se mide en milímetros. Debido a que ésta es una medición indirecta, ya que el observador debe calcular en función del nivel del día anterior, el nivel del día de la observación, y considerando la lámina de lluvia caída, suele tener mayor probabilidad de error. Es conveniente que el usuario sea precavido, observando los datos de la serie de tiempo para descubrir errores de cálculo.
- **Niebla:**⁵ esta variable indica presencia (1) o ausencia (0) de niebla durante las 24 horas previas.
Granizo: se indica la ocurrencia (1) o no (0) de granizo durante las 24 horas previas.
- **Tormenta eléctrica:** se indica la ocurrencia de tormenta en forma de truenos o relámpagos, durante las 24 horas previas. Ocurrencia (1) o ausencia (0).
- **Cobertura del cielo:** variable cualitativa para indicar cielo despejado (0), medio nublado (1) o nublado (2).

⁵ Debido a la escasa capacitación del personal a cargo de estas estaciones, la apreciación de las variables cualitativas es muy subjetiva, no pudiendo garantizarse la homogeneidad de criterios.

3 SISTEMA DE CONSULTA A LA BASE DE DATOS



ERIC II consiste en la base de datos que contiene la información que se describió en las secciones 1 y 2, compactada y ordenada⁶ de manera adecuada para realizar búsquedas de datos de manera rápida y eficiente. La información es exactamente la misma que se encontraba en el sistema *Clicom*, a principios de enero del año 2000.

Además, *ERIC II* cuenta con el programa administrador de la base de datos compactada, que ayuda a hacer las consultas al catálogo de estaciones y extraer la información climatológica, seleccionando la ventana espacio temporal que se requiera consultar.

El sistema *ERIC II* trabaja directamente desde el disco compacto, no siendo necesario instalarlo previamente, por lo que no ocupa lugar en el disco duro de la máquina. Debido a que su aplicación principal es la búsqueda de grandes cantidades de datos, éstos se depositan en archivos en el disco duro o algún otro medio que el usuario indique.

3.1 *Uso del sistema*

Para utilizar el sistema debe seleccionarse el icono de *ERIC II* del directorio raíz del disco compacto o una copia de éste para su acceso rápido.

ERIC II permite consultar las variables meteorológicas seleccionando las estaciones en una región geográfica, según los siguientes criterios:

- 1 Una estación.
- 2 Todas las estaciones.
- 3 Una región rectangular.
- 4 Una región poligonal.
- 5 Un estado.

Cualquiera que sea la ventana geográfica de la consulta, deberá indicar la ventana temporal y alguna de las nueve variables climatológicas:

- 1 Temperatura observada.
- 2 Temperatura máxima.
- 3 Temperatura mínima.
- 4 Precipitación.
- 5 Evaporación.

⁶ La estructura de la base de datos se explica en la sección 4.2: "Archivos de la base de datos *ERIC II*".

- 6 Tormenta.
- 7 Granizo.
- 8 Niebla.
- 9 Nubosidad.

Dependiendo del criterio de búsqueda seleccionado, se deberán dar datos adicionales según el caso:

1 Por estación

Hay que dar la clave de la estación en cinco caracteres. Los dos primeros para el estado y los tres últimos para la estación, por ejemplo:

04025 corresponde a la estación 25 de Campeche.⁷

Cuando el usuario no conozca la clave de la estación, puede consultar el catálogo de las 5,330 estaciones, ordenadas por estado, especificando el número de la clave, el nombre de la estación, su latitud, longitud y altura sobre el nivel del mar.⁸

Una alternativa de búsqueda en una localidad específica es con la opción de ventana rectangular alrededor de las coordenadas de latitud y longitud conocidas.

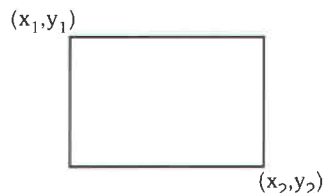
2 Todas las estaciones

No requiere información adicional, buscará la información solicitada para todas las estaciones del país. Esta opción tiene el inconveniente de que la salida puede ser un archivo de datos de varios megabytes, muy difícil de manejar para su posterior procesamiento.

3 Dentro de una ventana rectangular

Esta opción requiere la longitud y la latitud de la esquina superior izquierda (x_1, y_1) , y luego la longitud y latitud de la esquina inferior derecha (x_2, y_2) . Es conveniente conocer las coordenadas alrededor de donde se quiere tener información y elegir la ventana apropiada, según el área que se quiera estudiar. Hay que recordar que en México un grado equivale aproximadamente a 110 kilómetros.

(x_1, y_1)	102	24
(x_2, y_2)	101.5	23.7



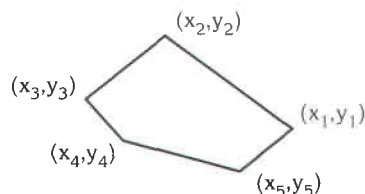
⁷ Es necesario agregar los ceros adelante para formar la clave de cinco caracteres.

⁸ Existe 1% de estaciones con coordenadas erróneas.

4 Dentro de una región poligonal

Esta opción permite introducir los vértices de un polígono convexo. Por cuestiones prácticas se limita el número de vértices a diez. En casos particulares como los contornos de una cuenca, por ejemplo, ésta se podrá descomponer en varios polígonos convexos. Las coordenadas x_i, y_i de cada vértice, deben introducirse en forma ordenada, en cualquier sentido de circulación. Cada uno de los vértices a considerar debe llevar la marca correspondiente.

(x_1, y_1)	99	22.1
(x_2, y_2)	102	24.5
(x_3, y_3)	103.7	22
(x_4, y_4)	102.5	21
(x_5, y_5)	100	20.8



5 Por estado

Debe seleccionarse de la lista de estados.

3.2 Presentación de resultados

Como resultado de la búsqueda se obtienen tres archivos *ASCII* observables con un editor de textos o con una hoja de cálculo; los archivos se encontrarán en el directorio de trabajo y tendrán el nombre asignado por el usuario, pero variando la extensión. La unidad de búsqueda del *ERIC* corresponde a los datos de un año, de una variable para una estación que se presentan en forma de tabla.

El archivo principal tendrá las tablas de datos diarios de la variable solicitada, para cada una de las estaciones que se encuentran en la ventana espacial indicada y para cada año de datos en que existan valores dentro del intervalo solicitado. Cada tabla tiene una línea de encabezado que proporciona los datos generales de la estación y el año de que se trata, seguidos de una matriz con las doce columnas correspondientes a los meses y 31 renglones correspondientes a los días.

Un segundo archivo contiene el resumen de los valores mensuales⁹ para las mismas estaciones y años que el archivo principal. En el caso de las temperaturas, el valor mensual es el valor promedio mensual. En el caso de las variables de lluvia y evaporación, se indica el valor acumulado mensual. En el caso de las variables cualitativas, se indica el número de veces que ocurrió el fenómeno durante el mes.

⁹ Ver Apéndice 2: "Errores aún presentes en *ERIC II*", porque hay errores en los acumulados.

Un tercer archivo contiene la lista de las estaciones sobre las que se hizo la búsqueda, indicando la clave de la estación, su latitud y longitud. Este archivo puede ser muy útil si se quiere localizar la información sobre un mapa geográfico.

Notas o recomendaciones

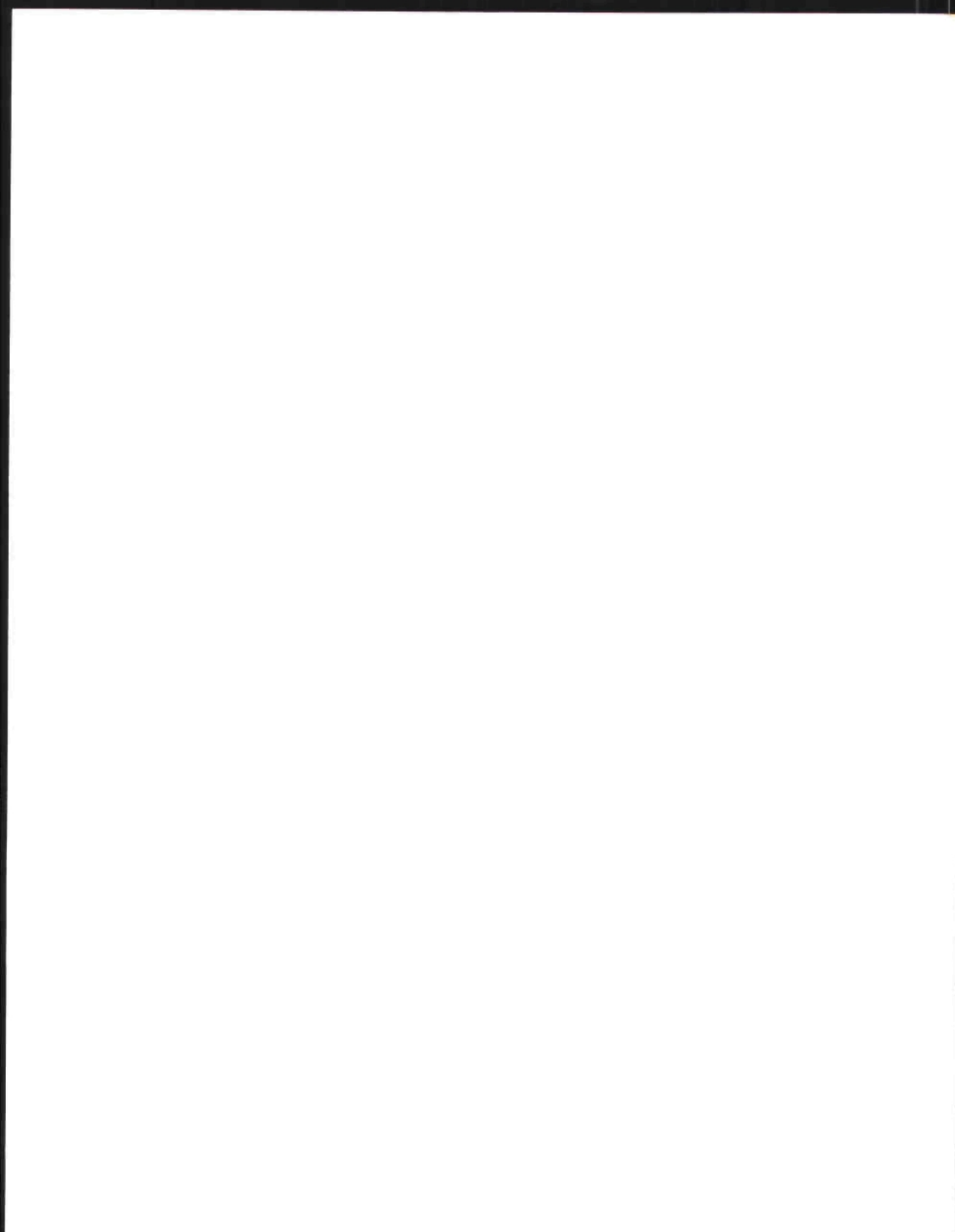
Se recomienda no hacer búsquedas demasiado amplias, ya que el resultado será un archivo muy grande; por ejemplo, los datos de evaporación de todo el país para el año 1980 ocuparon 5.4 Mbytes. Deberá existir espacio suficiente en el disco y contar con un editor que pueda procesar el archivo.

Cuadro 1: Clave asignada por *Clicom* a las variables meteorológicas

Núm.	Variable
1	Temperatura observada.
2	Temperatura máxima.
3	Temperatura mínima.
5	Precipitación.
18	Evaporación.
30	Tormenta.
31	Granizo.
32	Niebla.
43	Cobertura del cielo.

Cuadro 2: Referencia de los estados

Núm.	Estado	Abrev.
1	Aguascalientes	AGS
2	Baja California Norte	BCN
3	Baja California Sur	BCS
4	Campeche	CAM
5	Coahuila	COA
6	Colima	COL
7	Chiapas	CHP
8	Chihuahua	CHH
9	DF	DFE
10	Durango	DUR
11	Guanajuato	GTO
12	Guerrero	GRO
13	Hidalgo	HGO
14	Jalisco	JAL
15	México	MEX
16	Michoacán	MIC
17	Morelos	MOR
18	Nayarit	NAY
19	Nuevo León	NLN
20	Oaxaca	OAX
21	Puebla	PUE
22	Querétaro	QRO
23	Quintana Roo	ROO
24	San Luis Potosí	SLP
25	Sinaloa	SIN
26	Sonora	SON
27	Tabasco	TAB
28	Tamaulipas	TAM
29	Tlaxcala	TLX
30	Veracruz	VER
31	Yucatán	YUC
32	Zacatecas	ZAC



**4 ARCHIVOS ORIGINALES
DEL *CLICOM***



En esta sección se explica la estructura de datos del *Clicom* para el caso de los datos diarios. Se presentan los archivos que fueron leídos y compactados, con el detalle de sus registros. Esta información es útil para quienes quieran realizar aplicaciones leyendo los datos originales del *Clicom*.

Clicom está soportado por el manejador de base de datos *Dataease*. Se desarrolló a principios de los años ochenta para manejar y procesar información climatológica en microcomputadoras, en países pequeños. Como las capacidades de las máquinas eran aún muy limitadas, no se trató de tener una herramienta eficiente, sino de aprovechar las estructuras de las grandes bases de datos con las salidas interactivas de PC. Las nuevas versiones de *Dataease* y de *Clicom*, y los medios de almacenamiento actuales, dados por la posibilidad de contar con discos duros de varios gigabytes, discos ópticos, discos compactos y otros medio magnéticos de la última década, han permitido aumentar su capacidad, pero no se ha modernizado consecuentemente la estructura inicial.

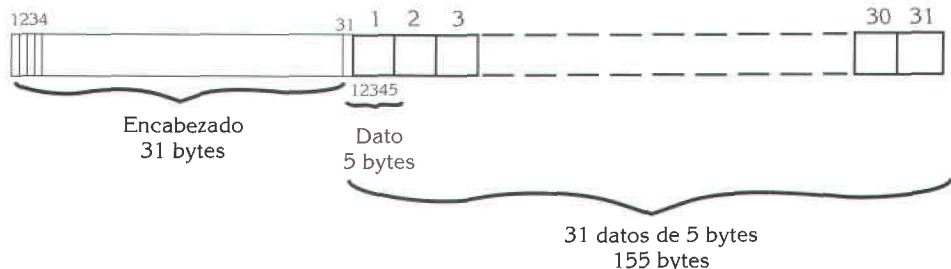
Los datos que se utilizaron se tomaron de los archivos que corresponden a la información diaria de la base *Clicom*, compuestos por registros de longitud fija. Cada registro contiene la información de un mes, para un elemento climatológico.

Los datos diarios están agrupados en 32 archivos, uno por cada estado. El nombre de estos archivos es, por ejemplo, *DLY020.DBM*, donde *DLY* indica que son datos diarios, *020* corresponde a la clave del estado y la extensión *DBM* es la que indica el formato de la base *Dataease*. Otra alternativa es obtener de *Clicom* unos archivos en formato de texto, tipo reporte, con los datos separados por comas. Los archivos generados de esta manera tiene por extensión *.scv*; es necesario escribir un programa para su interpretación y lectura. El problema de estos archivos es que no son un estándar de la base de datos, sino que los define el operador del *Clicom* cada vez que le solicita el reporte y, por lo tanto, puede variar su estructura y orden de archivo a archivo.

A continuación se indica la estructura de los archivos *.DBM*

4.1 Archivo *XXXXX.DBM*

El archivo *.DBM* es un archivo binario para acceso directo. Está compuesto por registros de 186 bytes. Estos 186 bytes están divididos en un encabezado de 31 bytes y 31 grupos de cinco bytes que guardan los datos diarios de todo un mes.



4.1.1 Encabezado

El encabezado está formado por 31 bytes que contienen la siguiente información:

Bytes	Contenido
1:1	Contiene un 12.
2:11	Contiene ceros.
12:14	Identificador del estado, tres caracteres. ¹⁰
15:22	Identificador de la estación, ocho caracteres.
23:25	Identificador del elemento climatológico. ¹¹
26:29	Año, cuatro caracteres.
30:31	Mes, dos caracteres numéricos.



El identificador de la estación está compuesto por el identificador del estado, antecedido de ceros, seguido por un número de tres dígitos asignado a cada estación en el estado.¹²

¹⁰ En el cuadro 2 se dan los identificadores de cada estado.

¹¹ En el cuadro 1 se dan los identificadores de los elementos o variables meteorológicas almacenadas en *Clicom*.

¹² En el archivo *posicion.idx* se encuentran los datos para localizar cada estación.

4.1.2 Datos

Los 155 bytes restantes están divididos en 31 grupos de cinco bytes cada uno. Se utilizan los cuatro primeros bytes para guardar el valor del dato diario y el quinto como bandera. Los datos son interpretados como un número real de cuatro bytes ($\text{real} \times 4$).

Clicom incluye un código fuente de procedimientos en *Fortran* para el desarrollo de aplicaciones propias. Se utilizó la rutina de lectura *READDLY* para leer los archivos originales.

4.1.3 Banderas

Cada dato tiene la posibilidad de tener una bandera para indicar características especiales del dato acompañante. *Clicom* define algunas de estas banderas y permite que el administrador defina otras. Se trata de un carácter y están definidos los siguientes: D, dudoso; E, valor estimado; G, valor generado; M, faltante; T, traza de lluvia.

Dentro de la base de datos, se encontró que la bandera asignada a cada dato está vacía en la mayoría de los casos. En aquellos lugares donde no hay un dato válido, como los lugares correspondientes al 29 y 30 de febrero, o como los días 31 de meses de treinta días, contiene el número -9999 en los bytes correspondientes al valor, y en el de bandera contiene una M. En estos casos la bandera es innecesaria, pues el valor ya indica que el dato no es válido. En el caso de los datos de precipitación aparece una T para indicar que hubo trazas de lluvia aunque no se registró nada; en *ERIC*, se reporta como 0 lluvia.

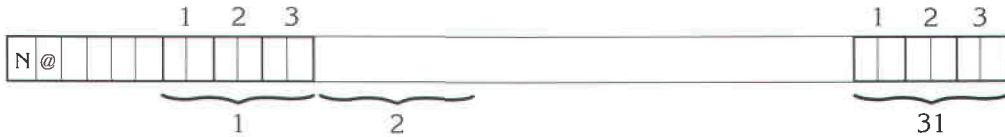
4.2 Archivos de la base de datos ERIC

En esta sección se describe la estructura de los registros de los archivos de la base de datos *ERIC*, según el tipo de datos, para los registros de temperaturas, precipitación, evaporación y variables cualitativas.

Aunque los datos originales son de distinta índole y sus rangos de variación son también diferentes, *Clicom* los considera a todos como variables reales para utilizar registros de longitud fija. Esta no es la mejor alternativa en lo que respecta al espacio utilizado, ya que para guardar una variable discreta, que puede tomar dos o tres valores, se utilizan cuatro bytes, más un quinto para indicar si el dato es válido o no.

4.2.2 Archivos de temperaturas: *TEMccc.BD*

El archivo de temperaturas de la base de datos es un archivo binario de acceso directo, formado por registros de 192 bytes que contienen un encabezado seguido de 31 campos con tres valores, cada uno: el primero, corresponde a la temperatura observada a las 8.00 AM; el siguiente, es la temperatura máxima diaria y, el tercero, la temperatura mínima diaria.



4.2.2.1 Campo de valores del archivo: *TEMccc.BD*

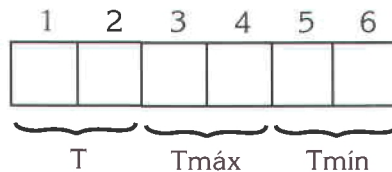
El valor de la temperatura está guardado en décimas de grado centígrado. El rango de valores posibles es:

$$-20^{\circ}\text{C} < T < 50^{\circ}\text{C}$$

que se convierte en:

$$-200 < T < 500 \quad \text{en décimas de grado}$$

El rango de -200 a 500 se codifican en enteros de dos bytes (de -32768 a 32767).¹³ Por lo tanto, cada campo de valores consiste en seis bytes en los que se almacenan la temperatura observada, la temperatura máxima y la temperatura mínima, en dos bytes cada una.



¹³ Los bytes que forman el entero se enumeran de derecha a izquierda, que corresponde a la notación de la familia *Intel*. En el caso de los enteros de dos bytes, la parte más significativa se encuentra en el byte de dirección menor, y la parte menos significativa, en la siguiente localidad.

4.2.3 Archivo de precipitaciones: *LLUccc.BD*

El archivo de precipitaciones de la base de datos es un archivo binario para acceso directo, formado por registros de 68 bytes que contienen un encabezado seguido de 31 campos para el valor de la precipitación diaria correspondiente al mes.



4.2.3.1 Campo de valores del archivo: *LLUccc.BD*

La precipitación se lee en milímetros con una resolución de 0.1mm. El rango de valores normales es de:

$$0 \text{ mm} < P < 500 \text{ mm}$$

que se convierte en décimas de milímetro para guardarlo en un número entero, quedando en el rango de :

$$0 < P < 5,000$$

codificándose en enteros de dos bytes (de -32768 a 32767)

4.2.4 Archivo de variables cualitativas: *BOOccc.BD*

El archivo de variables cualitativas de la base de datos es un archivo binario de acceso directo, formado por registros de 37 bytes, que tienen un encabezado, seguido de 31 campos que contienen, en forma codificada, la información correspondiente a la nubosidad, tormenta, granizo y niebla para cada día del mes.



4.2.4.1 Campo de valores del archivo: *BOOccc.BD*

Para codificar estas cuatro variables en un solo byte, se toman en cuenta sus valores relativos, se multiplican por un factor múltiplo de 2 y se suman:

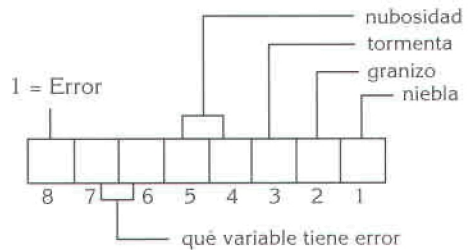
Variable	Valores posibles	Factor	Valores numéricos
Nubosidad	0,1,2	8	0,8,16
Tormenta	0,1	4	0,4
Granizo	0,1	2	0,2
Niebla	0,1	1	0,1

El factor sirve para colocar el valor en un bit determinado. Por ejemplo, si no hay niebla, se indica colocando un 0 en el bit 1 o el menos significativo; si hay niebla el valor de dicho bit será 1. Para el granizo se realiza de igual manera pero ocupando la posición del bit 2, o lo que es lo mismo, multiplicando el valor: 0×2 o 1×2 . En el caso de la variable tormenta los valores serán 0×4 o 1×4 correspondiendo entonces al bit número tres del byte; en el caso de nublado: 0×8 , 1×8 , 2×8 modificarán los bits 4 y 5 del byte. Al guardarlo en un número entero, queda en el rango de

o de 0 a 23 en decimal
 00000000 a 00010111 en binario

que se puede codificar en un entero de 1 byte. Los bits 6, 7 y 8 se utilizan para indicar si alguna variable tiene error o no está definida. En la siguiente tabla se indica cómo se codifican estos errores:

Bits 8, 7, 6	Tipo de error
0 0 0	No hay error
1 1 0	Error en tormenta
1 0 1	Error en granizo
1 0 0	Error en niebla
1 1 1	Error en nubosidad
0 1 1	Más de un error



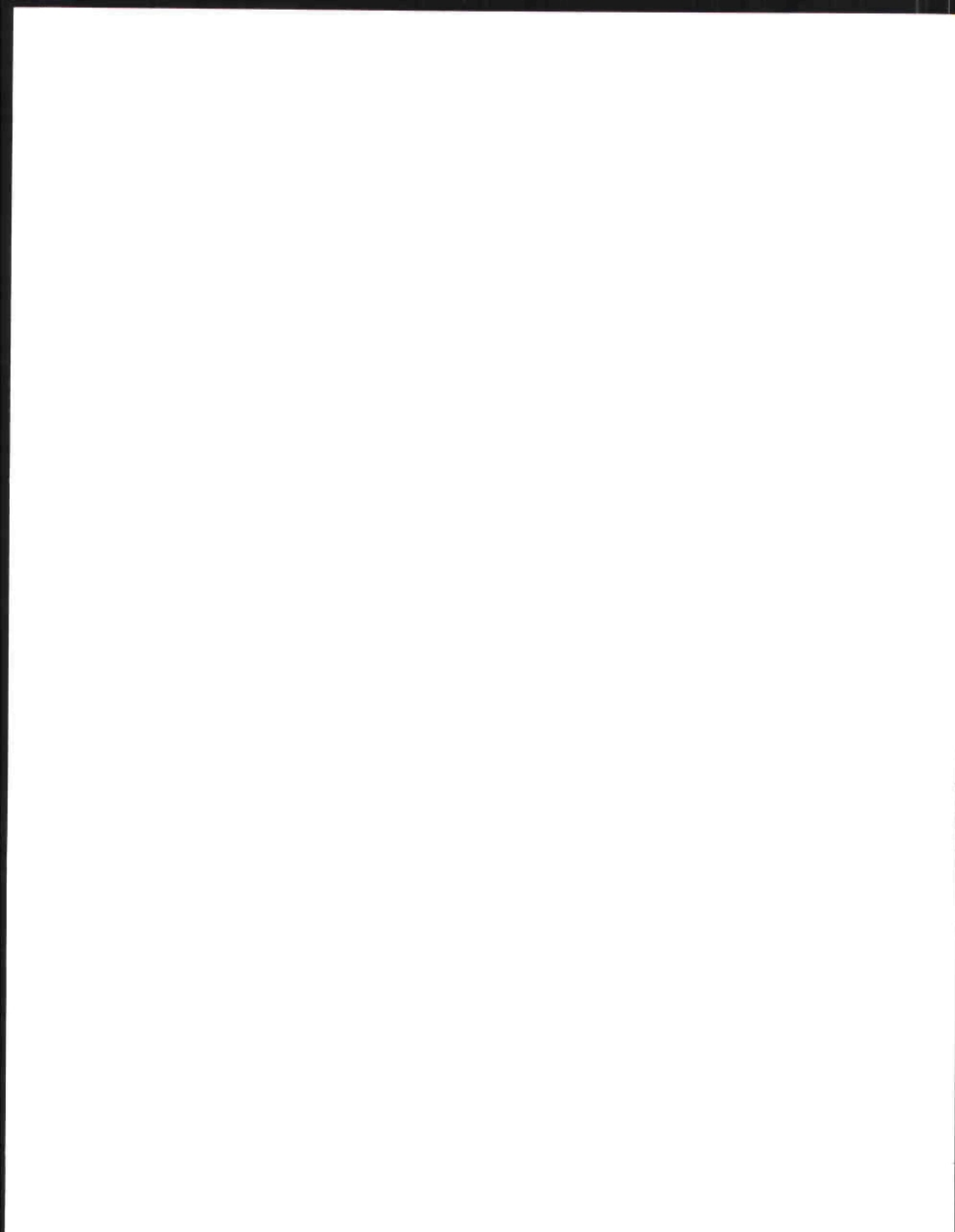
4.2.5 Archivo de evaporaciones: *EVAcc.BD*

El archivo de evaporación de la base de datos es un archivo binario para acceso directo, formado por registros de 37 bytes que tienen un encabezado seguido de 31 campos que contienen el valor de la evaporación diaria que corresponde al mes.

74	13	Año 74 empieza en el registro 103.
75	19	Año 75 empieza en el registro 109.
..
86	112	Año 86 empieza en el registro 202.
0	0	Fin de datos de la estación.
004	214	Otra estación.
..
..

En dicho ejemplo, la estación 001 tiene datos a partir de 1970. El primer registro de esta estación es el uno. El año 70 ocupa los registros del 1 al 6, por lo que sólo hay información de seis meses de ese año, ya que el registro 7 contiene información del año 71, para el que existen 12 registros; o sea, están todos los meses del 71. El año 72 no tiene datos y el último año de esta estación es el 88.

El par de ceros indica el fin de la estación. A continuación empieza la información de la estación 002, cuyo primer registro ocupa la posición 91. Para esta estación el año 73 comienza en el primer registro, o sea, en el 91; el año 74 empieza en el treceavo, o sea, en el 103, etcétera.



5 ORGANIZACIÓN DE LA NUEVA BASE DE DATOS



5.1 Directorios

Los archivos se encuentran agrupados en diferentes carpetas. En el directorio raíz de la base de datos se encuentran los programas de manejo de la base de datos *eric2.exe* con algunos archivos auxiliares. En la raíz también se encuentra un archivo de texto con información sobre las 5,360 estaciones existentes en la base de datos.

En este archivo se encuentran las estaciones ordenadas por estado, indicando su nombre, clave, latitud, longitud y altura sobre el nivel del mar.

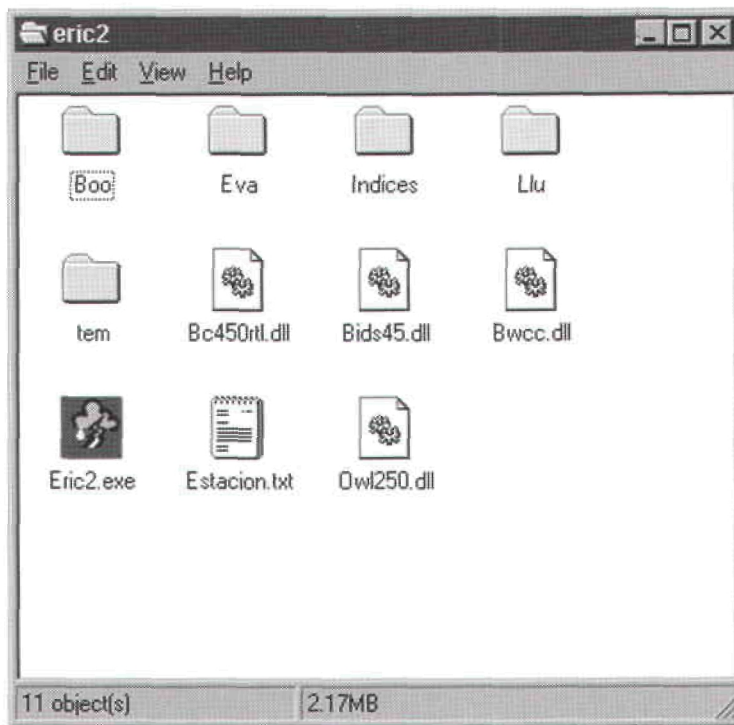


Figura 4. Organización de los archivos de la base de datos.

Además existen cinco carpetas en las que se encuentran, respectivamente, las bases de datos de temperatura, precipitación, evaporación y de variables cualitativas, y una conteniendo los archivos índices correspondientes.

El archivo *estaciones.txt* es un archivo de texto que contiene la clave de cada estación, su nombre, latitud, longitud y altura sobre el nivel del mar. Las estaciones que se han incorporado en esta segunda versión del *ERIC* aparecen con un tipo de letra diferente.

APÉNDICE 1
Procedimientos de
detección de errores
realizados por el SMN



Debido al gran número de errores encontrados en la base de datos *Clicom*, digitalizada en 1990, el SMN comenzó un trabajo de detección de errores de los datos que habían sido capturados hasta esa fecha. Dicho procedimiento se describe aquí, tal como lo reporta el personal a cargo de *Clicom* en la GSMN. Este procedimiento no fue aplicado a los datos agregados en esta segunda versión, donde cada estado realizó la captura de datos de los últimos años y, posteriormente, se incorporaron a la base de datos nacional. Es muy probable que, dado que en la mayoría de los estados, tal vez con la excepción de Guanajuato, hicieron la captura utilizando al mismo sistema *Clicom*, hayan utilizado los controles de calidad que trae predefinidos el sistema; sin embargo, esto no está documentado.

Para efectuar la validación de datos, el SMN creó un programa basado en los programas de utilerías en *Fortran* de *Clicom*, el cual lee todos los archivos en formato binario de las estaciones de un estado del país en una sola corrida y verifica los registros climatológicos contra límites preestablecidos, por elemento y tipo de validación. Los valores que no cumplen con los criterios establecidos se reportan como posiblemente erróneos, para cotejarlo posteriormente con los registros en papel de dichos datos, y finalmente, en su caso, corregir los registros erróneos con el dato real. El programa realiza las siguientes verificaciones:

- 1 De amplitud.
Verifica los valores máximos y mínimos de todos los elementos contra valores preestablecidos.
- 2 Coherencia interna entre elementos.
Revisa que haya coherencia entre las observaciones de diferentes elementos en un mismo día, por ejemplo, verifica que la temperatura mínima sea menor que la máxima.
- 3 Coherencia temporal con observaciones anteriores.
El programa revisa en las temperaturas que la diferencia absoluta entre un dato registrado y el del día anterior, no sobrepase de un cierto límite preestablecido.

La validación se efectuó y completó en todos los estados del país a excepción de Veracruz, que prácticamente se está recapturando y aún no se ha concluido, motivo por el cual *ERIC II* no tiene datos anteriores a 1990 para las estaciones cuya clave sea mayor a 157. Los interesados en estos datos pueden solicitarlos al Departamento de Climatología del SMN o al Instituto Mexicano de Tecnología del Agua.

En cuanto a los criterios de validación utilizados, fueron valores genéricos para todo el país. Un trabajo más cuidadoso exige determinar los valores para cada región y cada mes, lo que implica hacer una climatología previa. Aún así, se debe reconocer que se han depurado gran cantidad de datos inconsistentes.

Los valores de precipitación y temperatura utilizados por el SMN para todo el país, como valores límite preestablecidos para la validación, se muestran en el cuadro siguiente.

Cuadro 3. Valores de precipitación y temperatura utilizados para la validación de la información.

Concepto	Condición y valores
Diferencia en grados de temperaturas de un día con el anterior	≤ 15
Temperatura mínima	≥ -10.0 y ≤ 30.0
Temperatura máxima	≥ -10.0 y ≤ 48.0
Temperatura a la hora de la observación	≥ -5.0 y ≤ 30.0
Precipitación máxima diaria	≤ 100.0
Evaporación máxima diaria	≤ 16.0
Días con niebla	≥ 0 y ≤ 1
Días con granizo	≥ 0 y ≤ 1
Días con tormenta eléctrica	≥ 0 y ≤ 1
Nubosidad	≥ 0 y ≤ 2
Los valores antes enunciados se ajustaron a casos especiales de estaciones climatológicas que presentan valores de temperaturas y precipitación fuera de los rangos anotados en todo el país, para lo cual se hicieron corrimientos especiales.	

APÉNDICE 2
Errores aún presentes
en *ERIC II*



Se ha detectado una serie de errores que aún no han sido corregidos, pero ya se trabaja en ello. Los problemas encontrados y las medidas a tomar son:

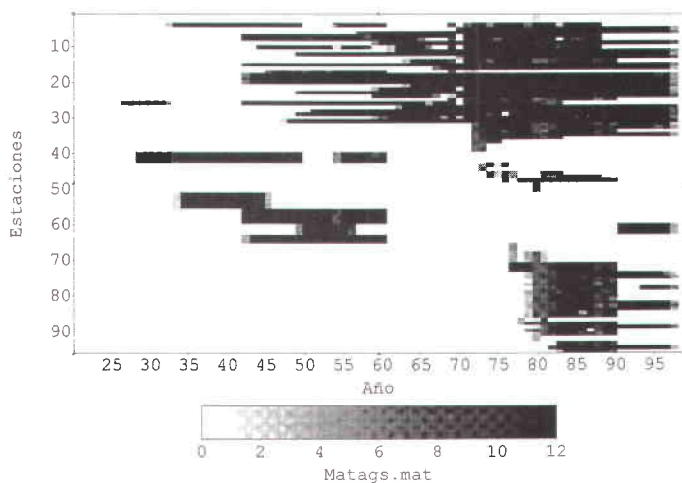
- 1 Las coordenadas de muchas estaciones están equivocadas. Se intentará, en todos los casos en que sea posible, tomar la posición de las estaciones utilizando posicionadores geográficos satelitales, *GPS*. La gran mayoría de las estaciones se corregirán utilizando la cartografía con mayor resolución.
- 2 En la base de datos aparecen algunos registros duplicados. Esto se debe a que se anexaron datos que ya existían en *Clicom*. Se trataron de eliminar, pero es posible que aún existan algunos.
- 3 Por el cambio de máquinas de 16 bits a las actuales de 32 bits, y las diferentes versiones de los compiladores, aparece como valor de dato no válido para la variable de evaporación el número 15. Este error se detectó cuando ya estaba armada la base de datos *ERIC II* y no se pudo corregir, pues 15 sí es un valor posible de evaporación. Queda a cargo del usuario decidir, observando la serie diaria, si el valor 15 corresponde a la evaporación o es un indicador de que no existe dato ese día. En el caso de los acumulados mensuales, siempre se considera como sin dato, por lo que la evaporación mensual será siempre igual o mayor a la calculada.
- 4 Para el caso del estado de Veracruz, los datos anteriores han sido revisados y en muchos casos tuvieron que ser recapturados, por lo que habrá diferencias importantes respecto a lo que tenía la versión anterior del *ERIC*. Faltan los datos anteriores a 1990 para todas las estaciones cuyo número de clave sea mayor a 157, porque aún no estaban corregidos y se prefirió no introducir parte de las series corregidas y otras no. El usuario puede recabar estos datos en el SMN o seguir utilizando los de la versión anterior.
- 5 En la variable *evaporación*, existe algún error en el estado de Baja California Sur que no permite realizar la búsqueda de todas las estaciones del país simultáneamente. Es necesario hacer las búsquedas estado por estado.
- 6 En el archivo de la variable *precipitación* del estado de Zacatecas quedó algún error que no permite obtener todos los datos de Zacatecas simultáneamente. Es necesario auxiliarse con la búsqueda por regiones o por estaciones. Los datos faltantes, después de recurrir a estas opciones, pueden solicitarlos en iquintas@tlaloc.imta.mx.
- 7 Después de editado el disco se detectó que los promedios de temperatura y los valores acumulados de lluvia que obtiene *ERIC II* están mal calculados en los siguientes estados: Aguascalientes, Baja California Norte y Sur, Campeche,

Guerrero, Hidalgo, Jalisco, Nayarit, Puebla, Querétaro, Tlaxcala, Yucatán y Zacatecas. Los datos diarios son correctos y los usuarios podrán calcular los promedios a partir de éstos. El problema se originó en que al procesar esos estados se utilizó como valor interno de dato no válido 99.9, mientras que en los demás estados se utilizó 999.9. La rutina que calcula los valores mensuales considera que si el valor es 999.9 es un dato faltante; esto hace que incorrectamente tome el valor 99.9 como valor posible y lo promedie obteniendo un resultado incorrecto.

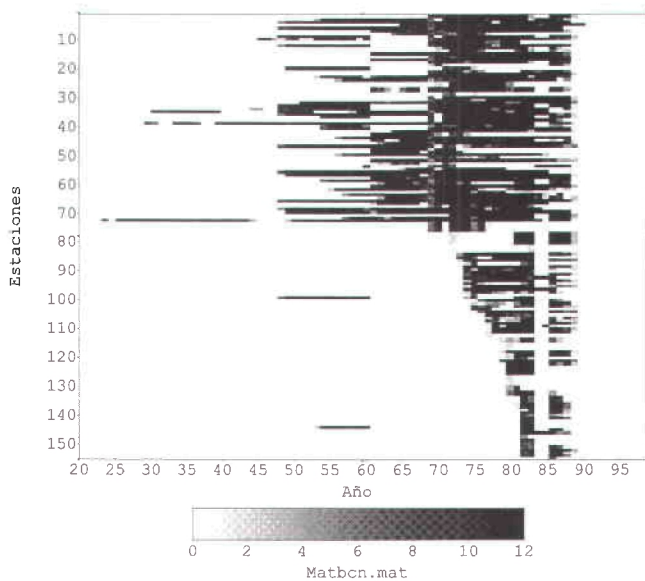
APÉNDICE 3
Estado de actualización
en cada entidad
federativa



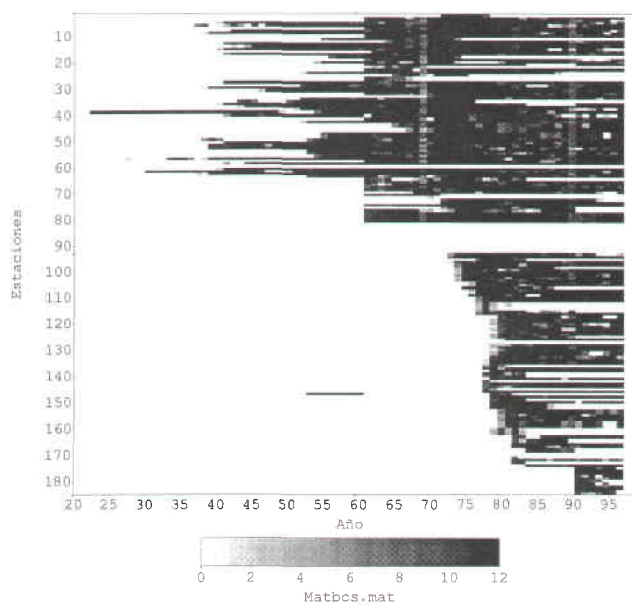
Aguascalientes



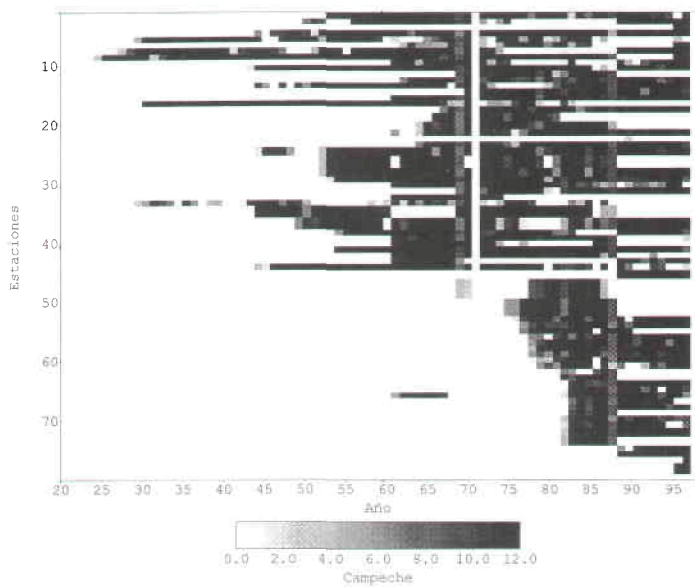
Baja California Norte



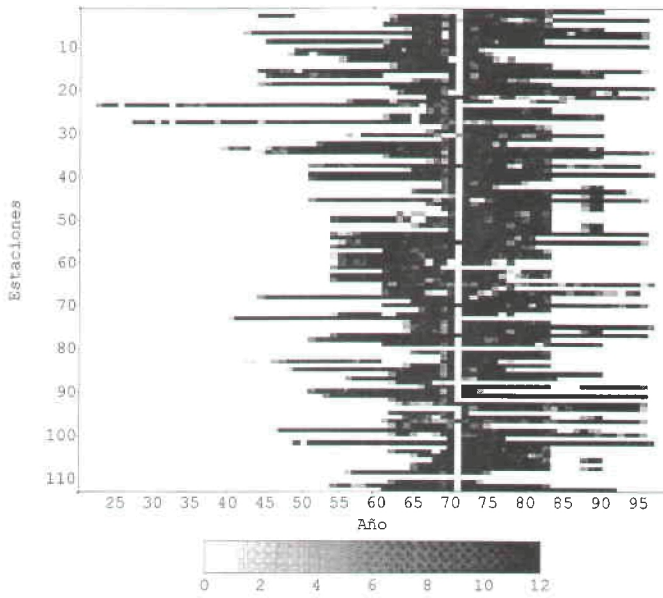
Baja California Sur



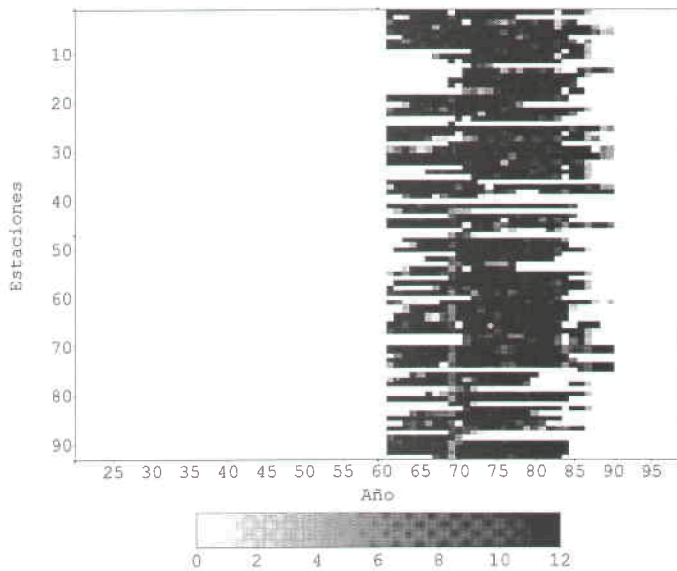
Campeche

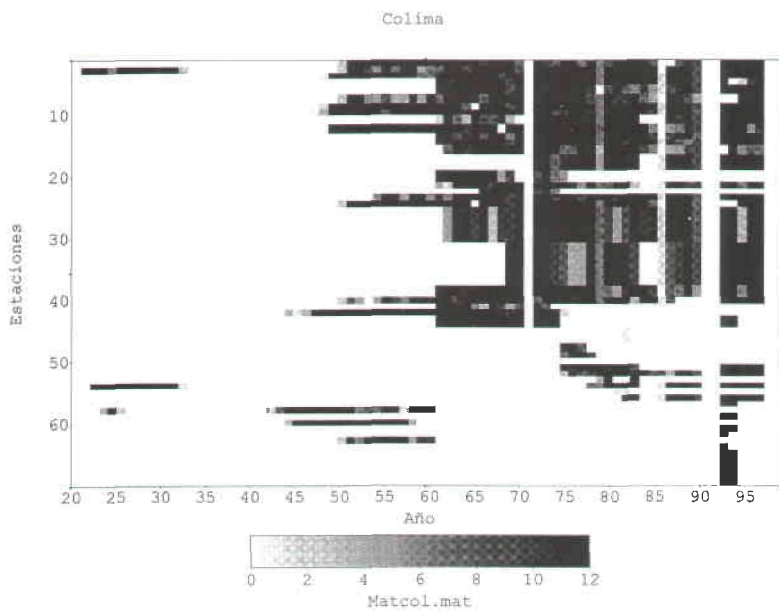
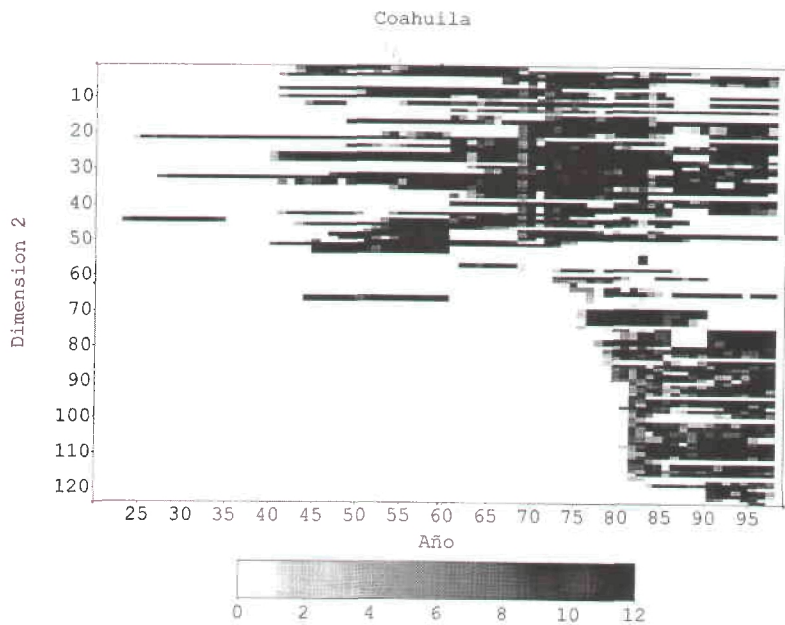


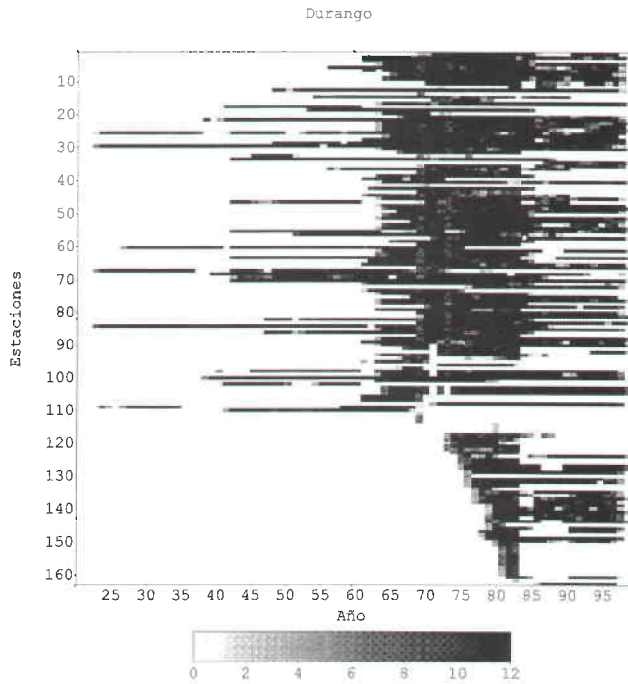
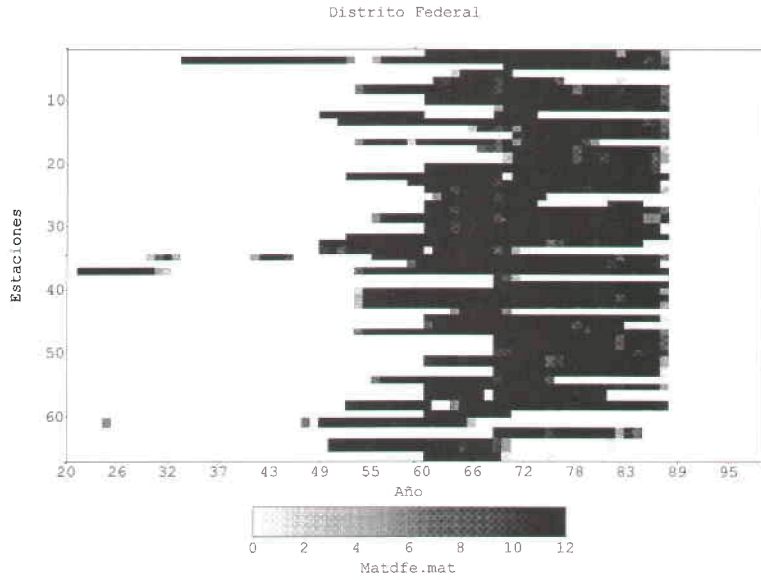
Chiapas



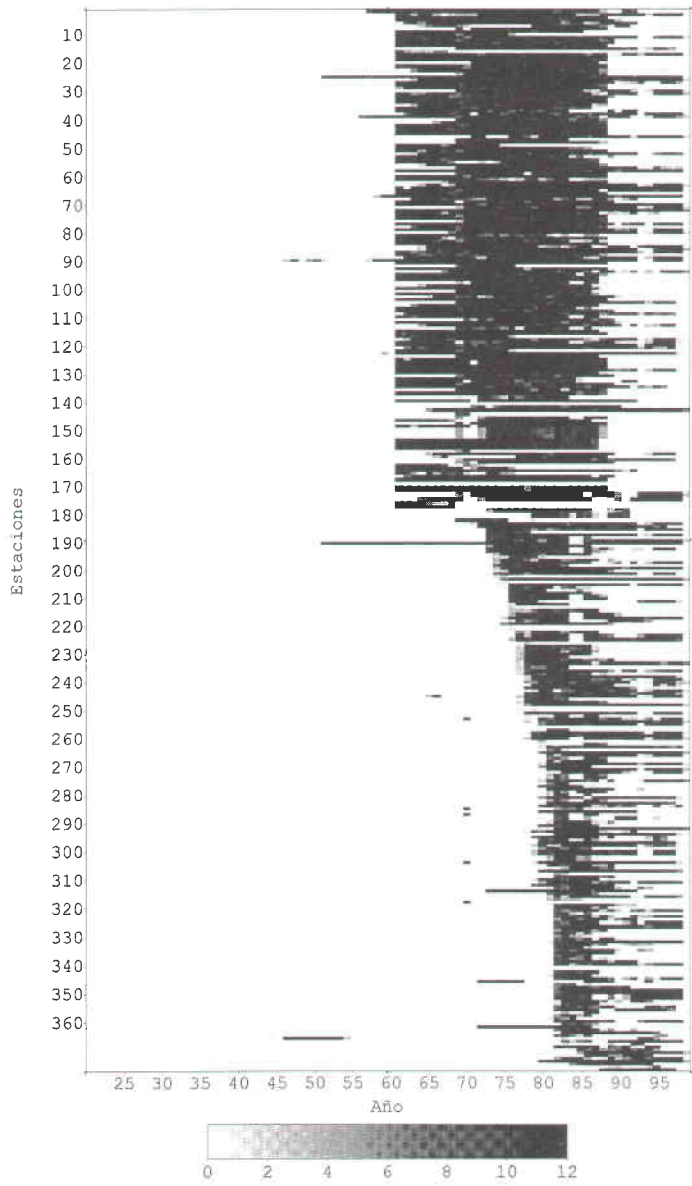
Chihuahua

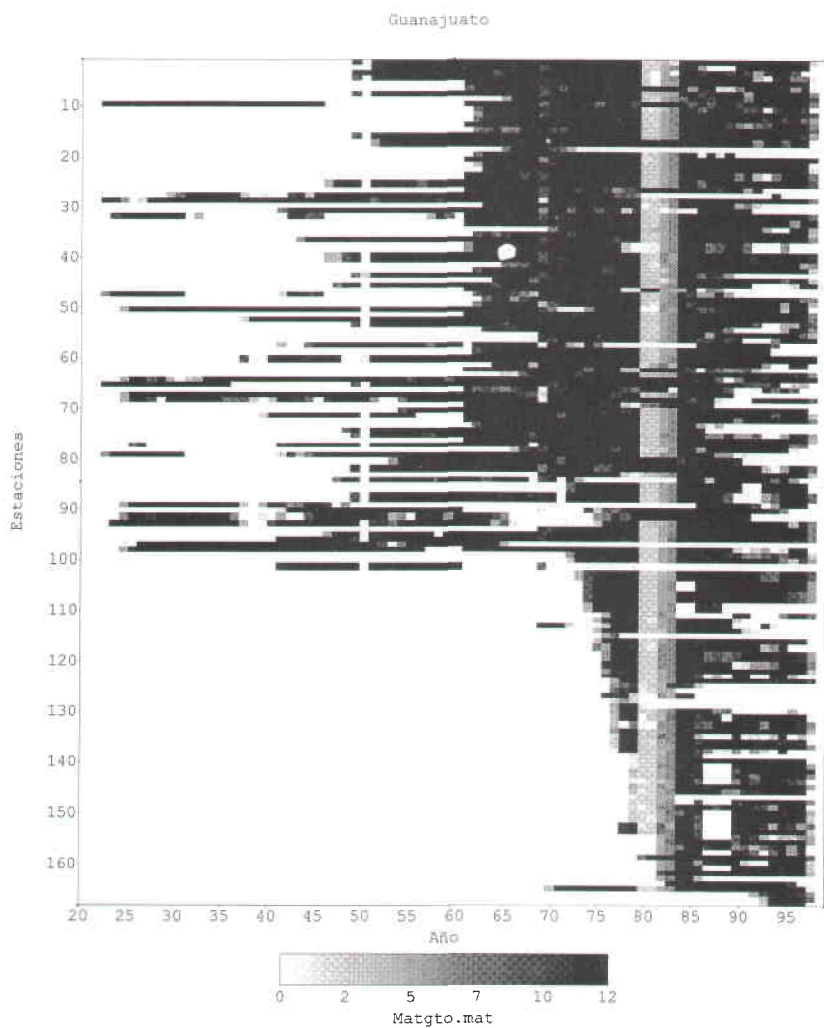


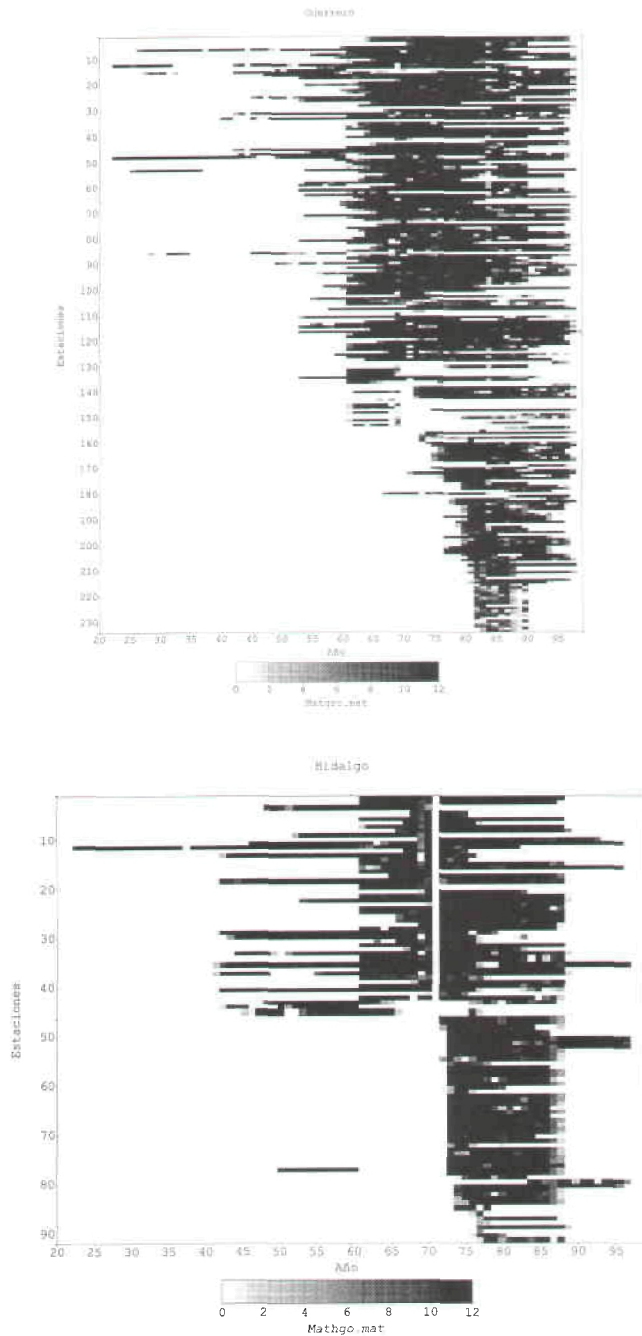


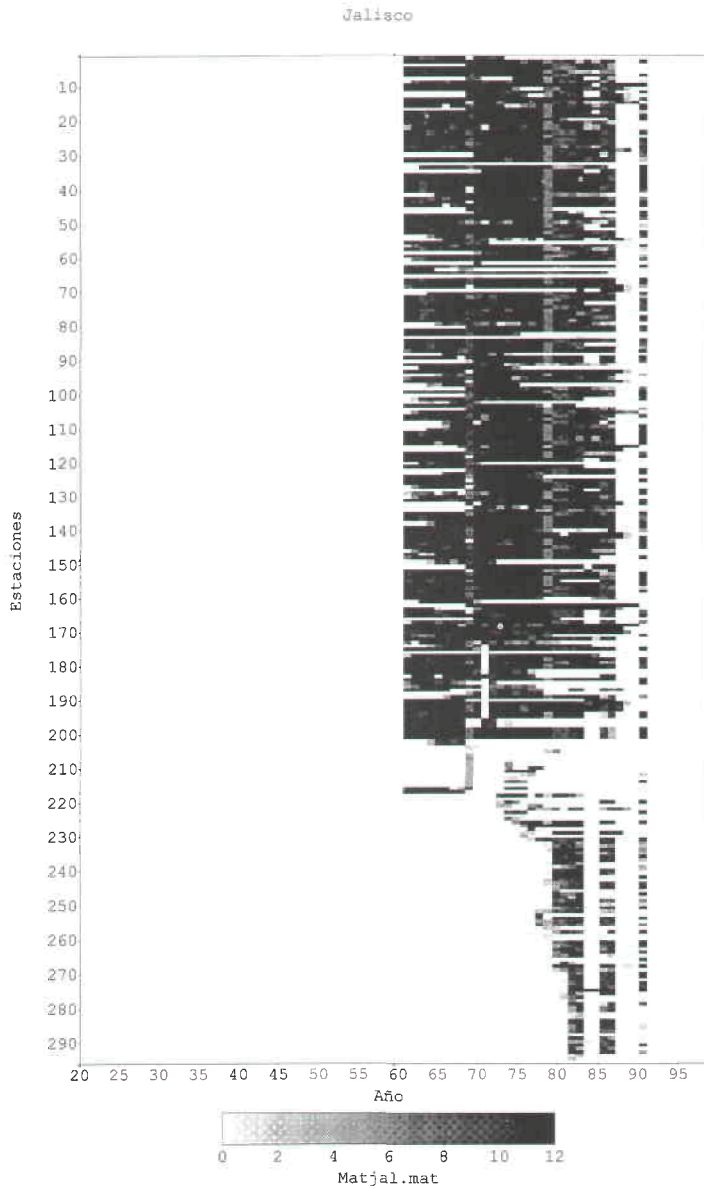


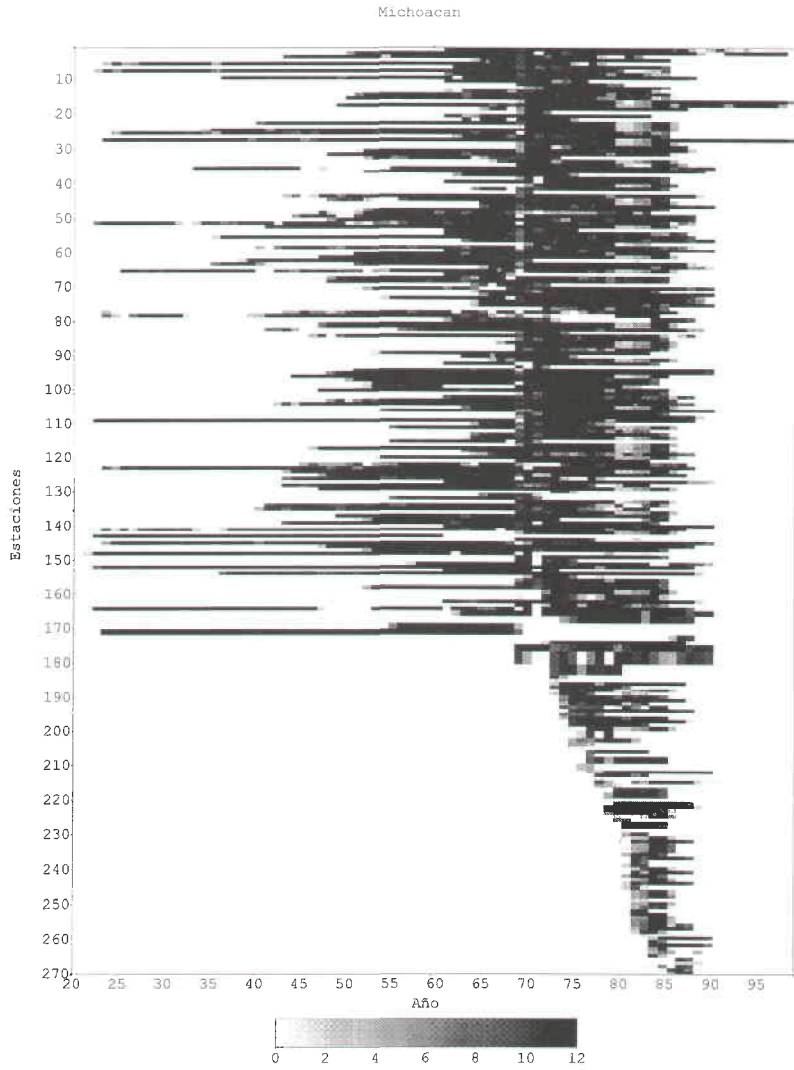
Estado de México



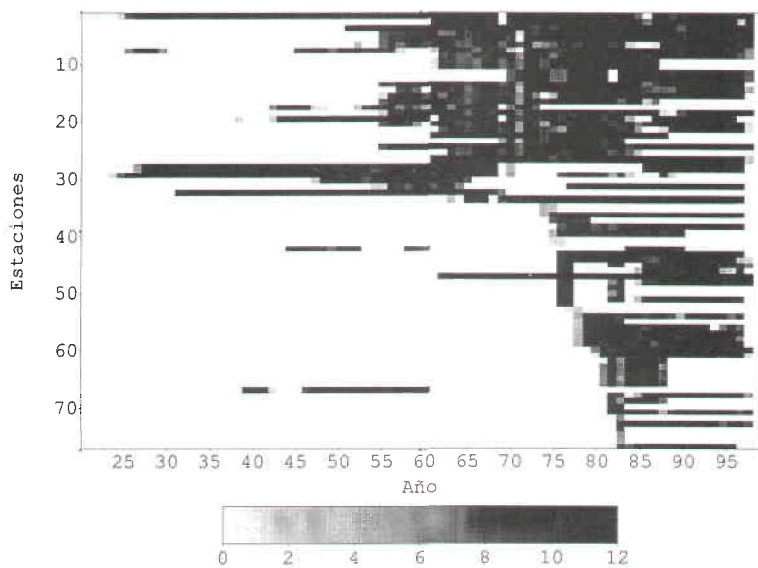




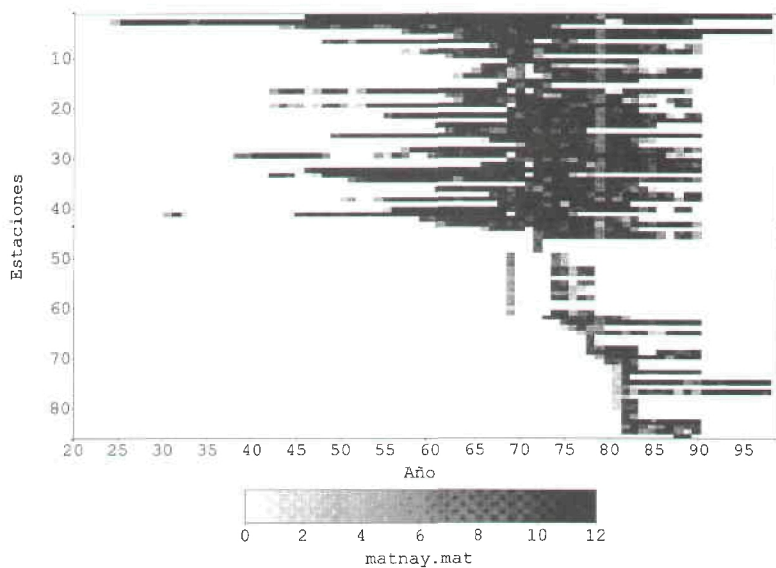




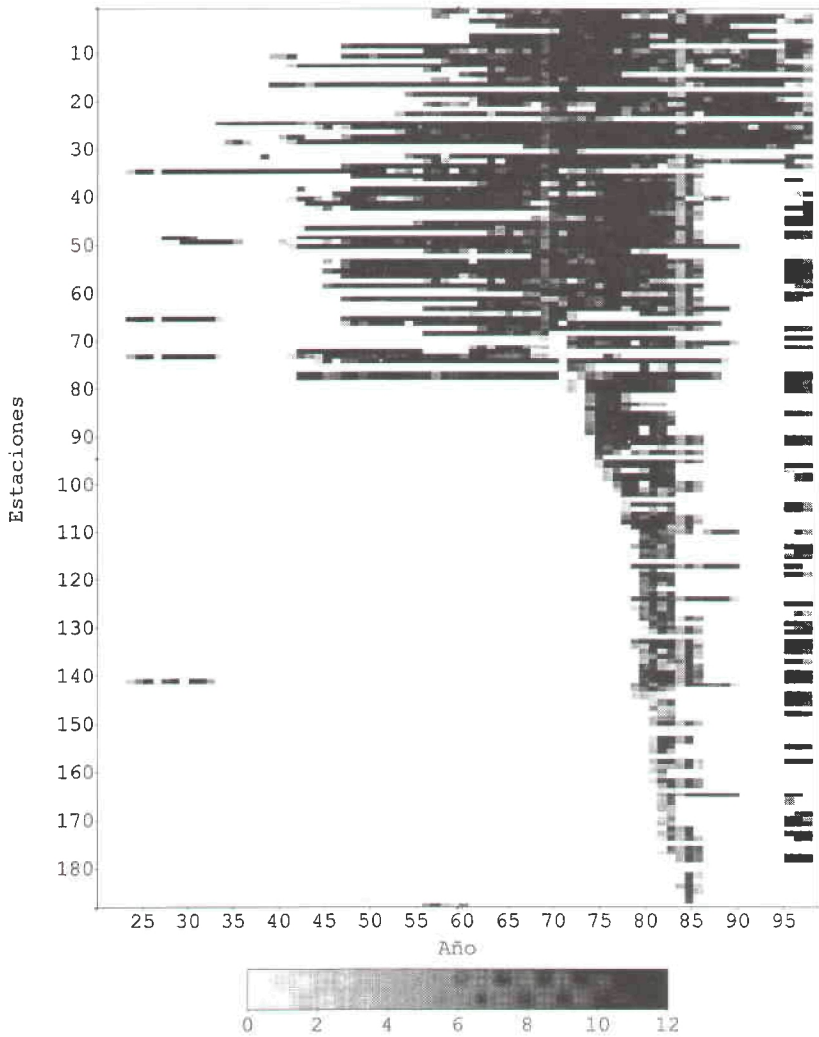
Morelos

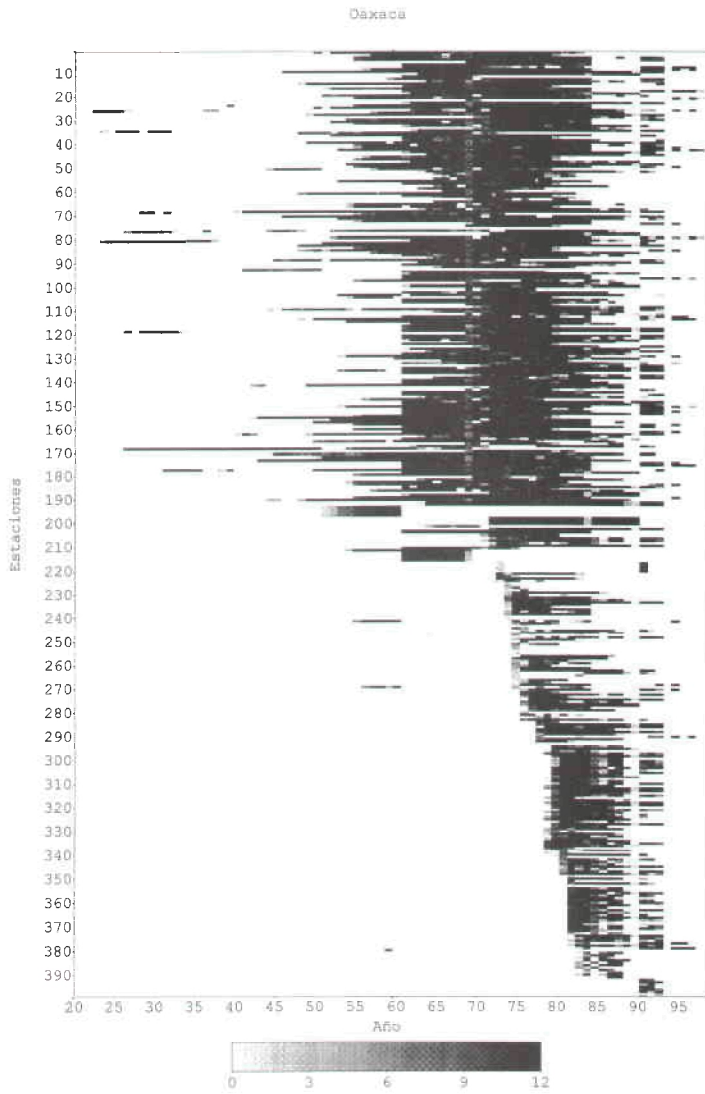


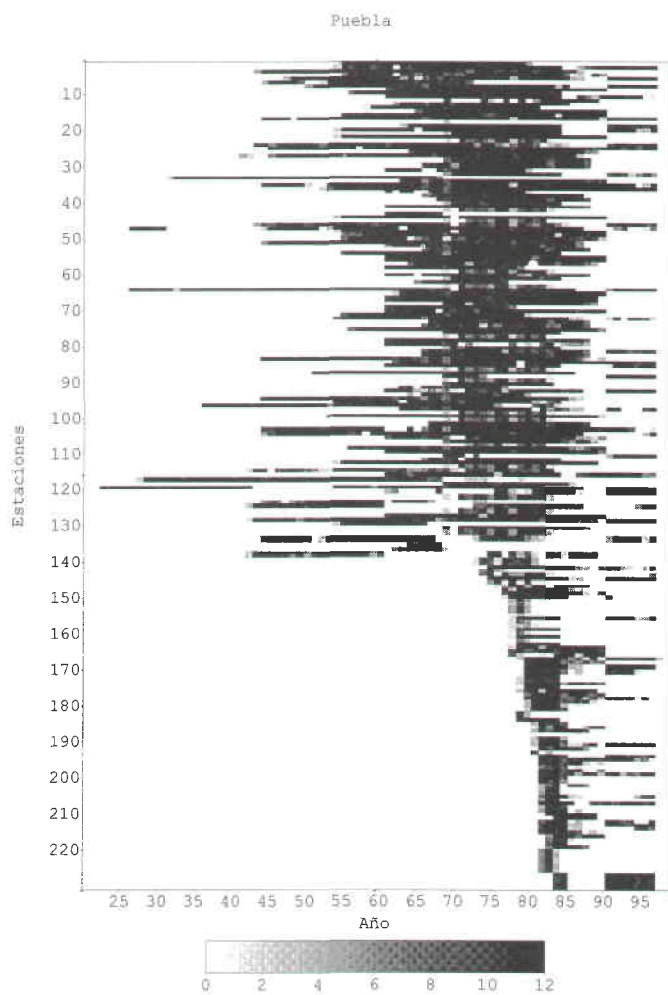
Nayarit



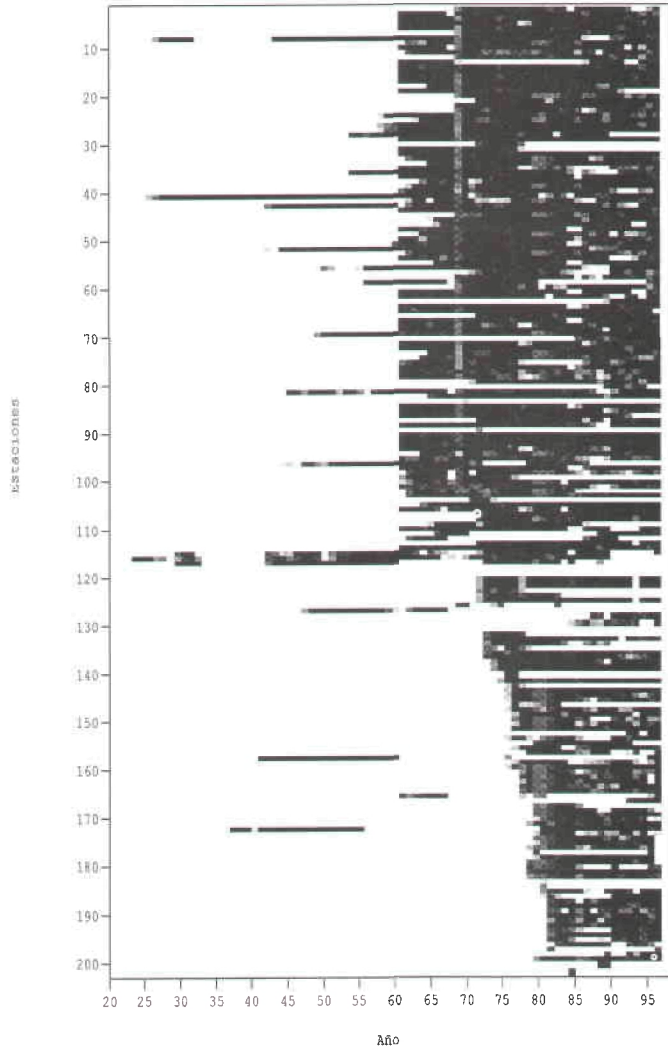
Nuevo León

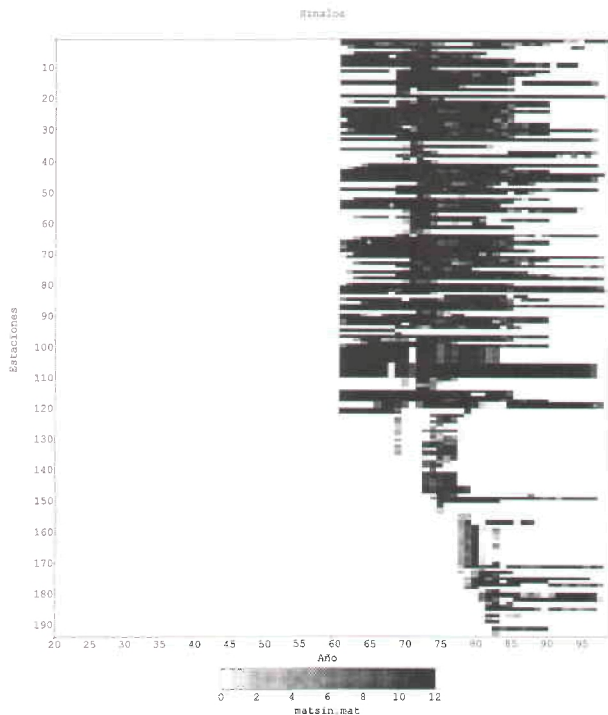




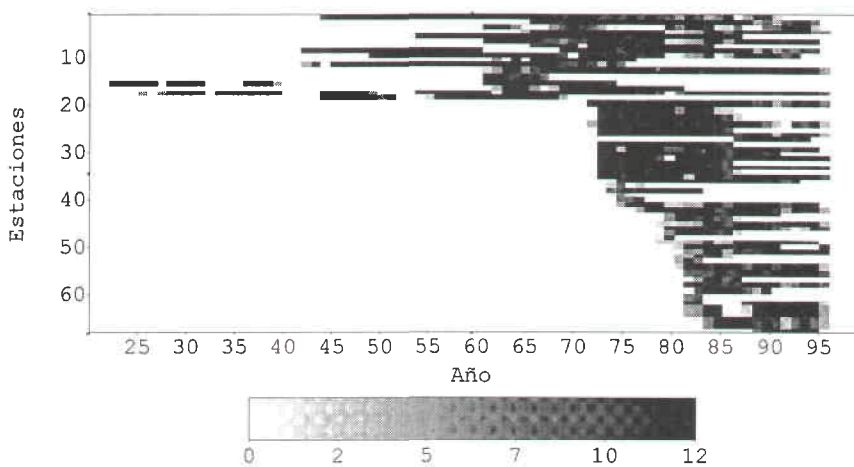


San Luis Potosí

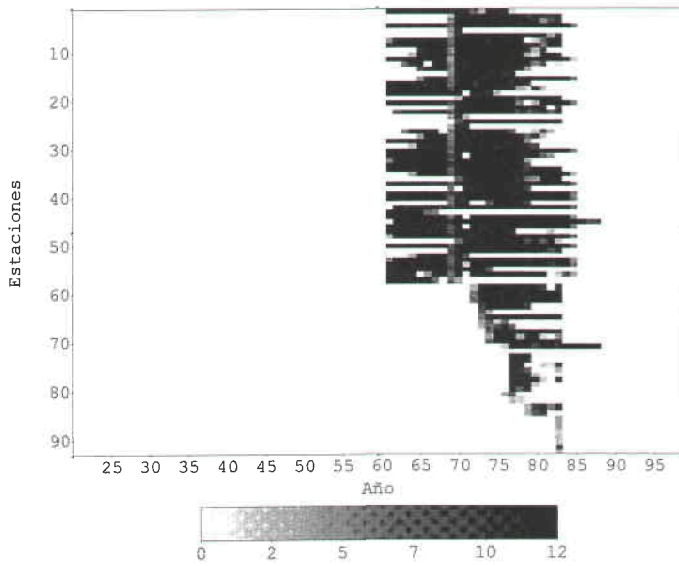




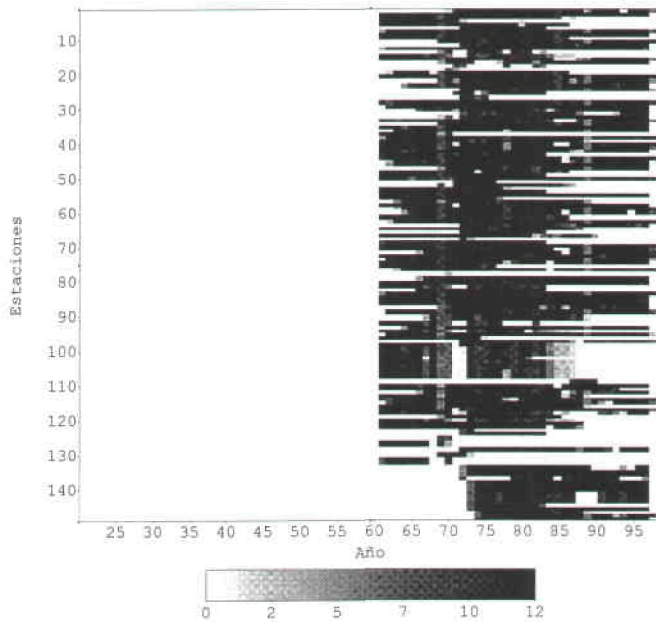
Querétaro

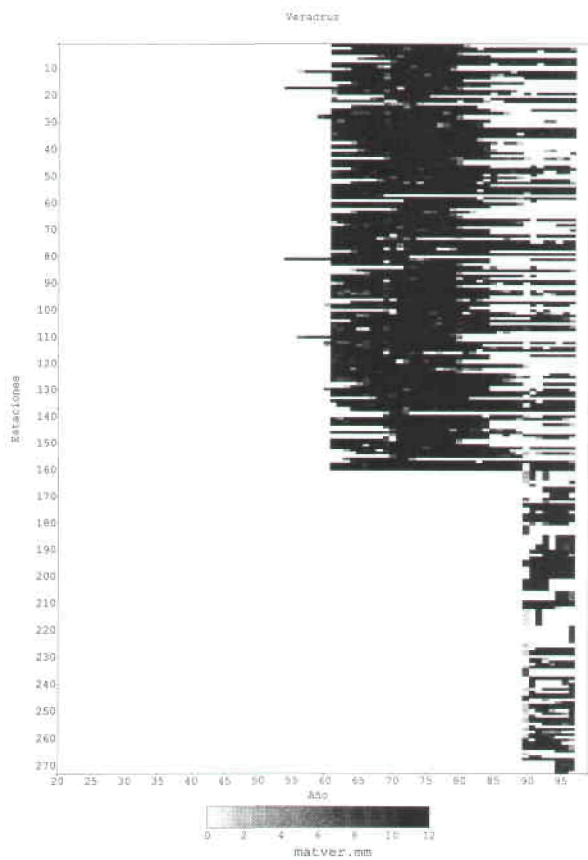
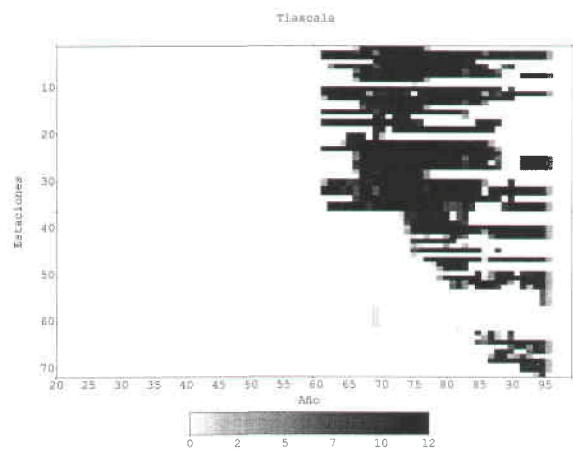


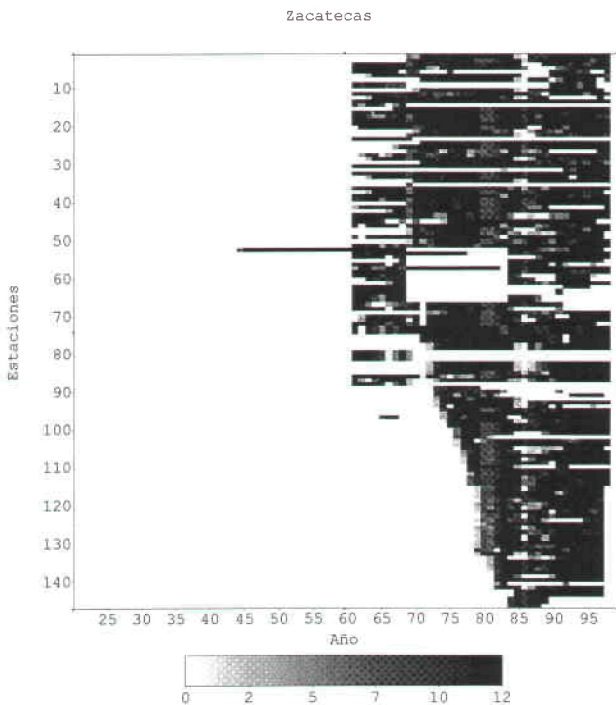
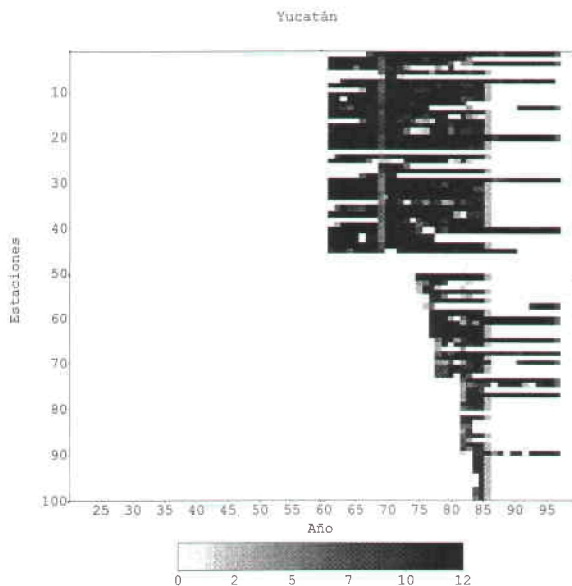
Tabasco



Tamaulipas









ERIC II se terminó de imprimir el mes de noviembre de 2000 en los talleres de Impresión y Diseño. Se utilizó tipografía Times New Roman de 10:12. Para los interiores se usó papel couché de 135 grs y para los forros cartulina couché de 169.5 kg. La edición consta de quinientos ejemplares.

CENTRO DE CONSULTA DEL AGUA

PAPELETA DE DEVOLUCIÓN

El lector se obliga a devolver este libro a la fecha del vencimiento del préstamo señalado en el último sello.

3- MAY - 11 <i>GR</i>		
-----------------------	--	--



19955
CENCA (Ej.1)

Eric II[®]

Documentación de la base de datos climatológica y del programa extractor

- *Eric II[®]* agrega 15 mil registros anuales a los ochenta mil existentes en la versión anterior, actualizándose así de acuerdo con la base de datos del Servicio Meteorológico Nacional.
- La actualización no es homogénea. La incorporación de información de la década de los años noventa relativa a los estados de Aguascalientes, Baja California Sur, Campeche, Coahuila, Durango, Guanajuato, Guerrero, Estado de México, Morelos, Querétaro, San Luis Potosí, Tamaulipas, Tlaxcala, Veracruz y Zacatecas significa un importante avance; mientras que en los estados de Colima, Chiapas, Nuevo León, Quintana Roo, Sinaloa y Yucatán se presentan pocas estaciones con datos actualizados. El resto aún no está disponible.
- Para la totalidad de los estados se depuraron los datos, corrigiéndose los erróneos de años anteriores y completando los faltantes.



Para mayor información relacionada con los productos y servicios del IMTA, comunicarse a:
Paseo Cuauhnáhuac 8532, Progreso, Jiutepec, Morelos, 62550, México.

Tel.: (73) 19-40-81

Fax: (73) 19-38-01

E-mail: cocom@tfaloc.imta.mx



19955

CENCA (Ej.1)