

## ENSAYO DE PROCESAMIENTO DE DATOS PARA LA FLORA DE VERACRUZ \*

S. OLVERA FONSECA \*\*

A. GÓMEZ-POMPA \*\*

### RESUMEN

Este trabajo es un ensayo de un sistema codificado para capturar datos para el procesamiento de datos para el programa "Flora de Veracruz". Se utilizaron los datos de las colecciones de Gimnospermas y Monocotiledóneas del Estado de Veracruz existentes en el Herbario Nacional del Instituto de Biología (MEXU). Se obtuvieron respuestas para una serie de preguntas. Las computadoras utilizadas fueron una Burroughs 5500 y una IBM 1130-3D. Se utilizaron los datos geográficos de las colecciones para elaborar automáticamente mapas de distribución de diversos taxa, con el auxilio de un graficador IBM-1627 Calcomp. Se discuten los problemas y las ventajas de este sistema en contraste con el que se sigue como rutina en la Flora de Veracruz.

### ABSTRACT

This paper is an assay of a coded system for data capture used in the processing of data in the "Flora de Veracruz" program. Data were obtained from the Gymnosperm and Monocotyledon herbarium specimens from Veracruz which are deposited in the Herbario Nacional del Instituto de Biología (MEXU). Answers were obtained for a series of questions. The computers used were a Burroughs 5500 and an IBM 1130-3D. The geographic data from the collections were used to elaborate distribution maps of various taxa automatically. A graphic plotter, IBM-1627 Calcomp, was used in this aspect of the process. The advantages and problems of this system are discussed in relation to that being used routinely in the "Flora de Veracruz".

### ANTECEDENTES

La aplicación de las computadoras electrónicas al campo de las ciencias biológicas ha tenido gran auge en la última década. La literatura de trabajos científicos en los que se ha puesto en práctica la versatilidad de las computadoras aumenta cada día. Los informes de trabajos biológicos publicados en este campo abarcan casi todas las ramas de la biología; así, los hay de tipo fisiológico, genético, zoológico, bioquímico, ecológico, botánico, paleontológico, etcétera.

\* Flora de Veracruz. Contribución núm. 9. Proyecto conjunto del Instituto de Biología de la UNAM, y del Arnold Arboretum y el Gray Herbarium de la Universidad de Harvard, para llevar al cabo un estudio ecológico-florístico del Estado de Veracruz, México. Una información detallada del proyecto puede verse en: An. Inst. Biol. Univ. Nal. Autón. México 41, Ser. Bot. (1): 1-2, 1970.

\*\* Instituto de Biología, UNAM.

Por otra parte, la gran flexibilidad del procesamiento electrónico de datos, ha dado por resultado las más ingeniosas aplicaciones de las máquinas a la biología.

La utilización de las computadoras en botánica es muy reciente y ha alcanzado diversos campos, así vemos que en el campo de la bibliografía botánica existen varios trabajos en los que se usan para resolver distintos problemas bibliográficos (Biological Abstracts, 1970; Lawrence *et al.*, 1968; Cowan, 1970; Gómez-Pompa, *et al.*, 1972).

Otro uso sumamente interesante de las computadoras en botánica es el realizado por Morse, Beaman, y Shetler (1968 y 1971) sobre la elaboración de claves de identificación que puede ser de gran utilidad para los botánicos, ya que se combina la elaboración de claves de distintos tipos, con la identificación de ejemplares desconocidos.

Introduciéndonos ya en el problema de la recuperación de información en botánica, encontramos que prácticamente no existen trabajos publicados y que la mayoría de ellos son proyectos en desarrollo o planes futuros (Batra, 1970; Bell, 1970; Crovello, 1970; Gould, 1970; MacDonald, 1970; Sokal *et al.*, 1970; Throckmorton, 1970). Los trabajos más importantes realizados son los efectuados por Perring (1963), Perring y Walters (1962) Soper (1964, 1966) y los de Rogers (1970).

Últimamente algunos autores han hecho una recopilación de los trabajos biológicos que implican el uso de computadoras, los cuales dan un panorama general de los proyectos en proceso

(Beschel y Soper, 1970; Crovello y MacDonald, 1970).

Trabajos que pueden considerarse pioneros de la computación en botánica, prometen enorme utilidad en su aplicación, no sólo para esta ciencia, sino que pueden ser la base para trabajos biológicos en otros campos, pues como menciona Ledley (1960) ... "con una coordinación efectiva tales programas de computadora pueden ser generales y suficientemente flexibles para admitir datos de diferentes fuentes, previendo así duplicación en el esfuerzo de programación" ... y pudiéramos agregar que también en la planeación general del problema.

El uso de computadoras electrónicas en México para trabajos botánicos, es también muy reciente y los primeros informes aparecen en el Simposio sobre Problemas de Información de Ciencias Naturales celebrado en 1967 en México (Gómez-Pompa y Olvera F., 1969, Sheinvar y Gómez-Pompa, 1969). Otros usos en México son los trabajos realizados en el Inventario Nacional Forestal (1970) y en el CIMMYT (Centro Internacional de Mejoramiento de Maíz y Trigo).

En 1967 aparece la primera publicación completa de un proyecto piloto de recuperación de información para el Herbario Nacional (Scheinvar, Gómez-Pompa y Alonso: 1967), en el cual se hace un ensayo con la colección de pteridofitas. En 1969, Sheinvar y Gómez-Pompa publican un trabajo concerniente a la elaboración automática de etiquetas de herbario, en donde, por otra parte se discute algo acerca del uso de esta metodología y sus aplicaciones en botánica.

## OBJETIVOS

Durante la elaboración del proyecto Flora de Veracruz, se vio que era muy importante organizar la información existente referente a las plantas del

Estado de Veracruz (Gómez-Pompa y Nevling; 1973).

Estos datos pueden provenir tanto de la bibliografía, como de los herbarios:

esta última fuente de datos probablemente es la más importante, pues es aquí donde se obtendrán tanto datos del ejemplar propiamente dicho, como la información registrada en las etiquetas correspondientes.

Los datos así obtenidos podrán ser la base de una serie de estudios morfológicos, taxonómicos, ecológicos y evolutivos, que darán cuerpo a la flora. Ellos podrán servir para hacer la descripción de las familias, manejar los datos, elaborar claves, calcular coeficientes de correlación, etcétera.

Pero, ¿cómo obtener esta información tan importante sin el enorme consumo de tiempo que presupone la revisión de todo un herbario?

La solución no era sencilla, la revisión del herbario era inevitable, pero lo que sí se podía solucionar era el tener que procesar manualmente la gran cantidad de datos.

Como mencionamos anteriormente, la computadora electrónica es de utilidad en todos los campos en donde va a ser manejado un gran volumen de datos o

información y es aquí donde se encontró una más de las aplicaciones de las computadoras al campo de la biología.

Uno de los objetivos del programa Flora de Veracruz (Gómez-Pompa, 1969), es la creación de un banco general de datos de donde se pueda hacer la extracción de dicha información en el momento deseado (Fig. 1).

Debido a la falta de antecedentes en este campo, se decidió probar diferentes métodos de preparación de la información para la entrada a la computadora. En este trabajo se presentan los resultados de un sistema sencillo de recuperación de información con datos codificados que nos permitió al mismo tiempo, evaluar el número de colecciones actuales de Veracruz en el Herbario Nacional, usando en este caso las colecciones de Gimnospermas y Monocotiledóneas.

Un segundo objetivo que nos marcamos en este trabajo, fue el de ensayar el uso de computadoras para elaborar automáticamente mapas de distribución de especies.

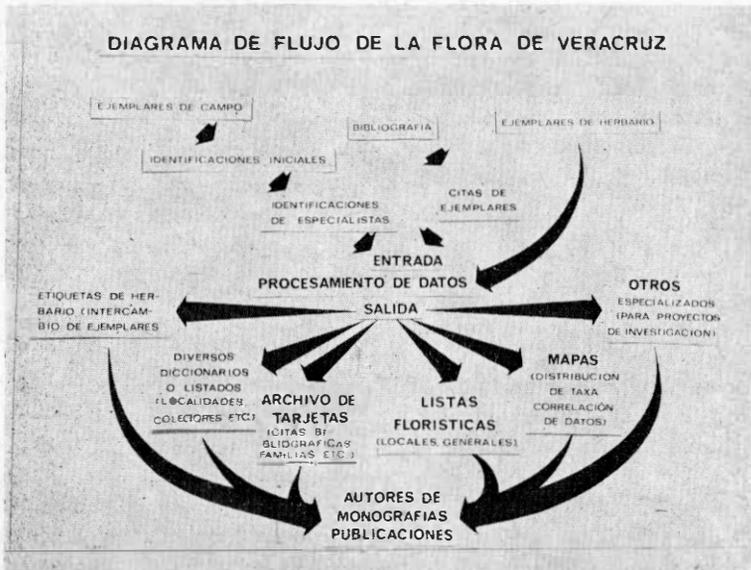


Fig. 1. Diagrama de flujo que muestra la creación del banco de información (tomado de Gómez-Pompa y Nevling, 1973).

## MATERIALES Y MÉTODOS

Para el desarrollo de este trabajo se utilizó parte de las colecciones del Estado de Veracruz, depositadas en el Herbario Nacional de México (MEXU) hasta el año de 1968.

De los ejemplares, se han obtenido hasta ahora 39 tipos diferentes de información, los cuales son codificados y puestos en 3 tarjetas IBM.

## FORMATO Y CODIFICACIÓN

## TARJETA 1

## Columna 1

Esta columna está reservada para el nivel taxonómico más alto en que se quiera subdividir cualquier programa.

Usamos esta información pensando que posiblemente en el futuro se podrían abarcar otros grupos de plantas del herbario.

Ejemplo: T = traqueofita

## Columnas 2-6

## Familia:

Codificamos tomando las primeras cinco letras del nombre de la familia.

Ejemplo: PODOC = Podocarpaceae

## Columnas 7-10

## Género:

Tomado del libro de Dalla Torre et Harms (1900-1907), se anota el número dado dentro de cada familia; cuando en algún caso no aparece anotado cierto género, se continúa la numeración después del último número anotado.

Ejemplo: 001 = *Podocarpus*

## Columnas 12-14

## Especie:

Asignado arbitrariamente en orden as-

cedente de acuerdo a las especies encontradas.

Ejemplo: 001 001 = *P. Matudai*

## Columnas 15-23

## Autor de la especie:

Codificado en nueve columnas de manera arbitraria,

Ejemplo: LUND = Lundell

## Columnas 25-50

## Localidad:

Puesta en 26 columnas se trató, hasta donde fue posible, de poner completos los nombres de los poblados, abreviando sólo en los casos necesarios. Asimismo, se procuró empezar con la entidad más importante anotada para determinada localidad.

Ejemplo: GUTIÉRREZ ZAMORA SN. MIGUEL  
CERC = Sn. Miguel, cerca de  
Gutiérrez Zamora.

## Columnas 51-53

## Estado:

En este trabajo, por ser exclusivo para Veracruz, sólo aparece la clave para este Estado, pero en tres columnas es posible codificar los nombres de todos los Estados de la República.

Ejemplo: VER = Veracruz

## Columnas 55-57

## País:

Aunque por ahora sólo estamos usando la clave para México, creemos que lo mismo que en el caso anterior, tres columnas parecen suficientes para abreviar los nombres de los países de los cuales existen recolecciones en el herbario.

Ejemplo: MEX = México

Columna 59

Nº de colectores:

Se anota exclusivamente el número de colectores que se citan en el ejemplar, que puede ser uno, dos o tres, o en ocasiones más.

Columnas 60-68

Nombre del colector:

El criterio usado para codificar esta información es el siguiente:

Para nombres latinos:

Se anotarán, las dos primeras letras del primer apellido, la primera letra del segundo apellido y la primera letra del nombre.

Ejemplo: BRHH = Bravo H., Helia

Para nombres no latinos:

Las dos primeras letras del primer apellido, la primera letra del primer nombre y la primera letra del segundo nombre.

Ejemplo: BEJH = Beaman John H.

Cuando los colectores son tres o más, sólo se anotarán los dos que aparezcan primero.

Cuando la recolección haya sido hecha por una institución, se usarán abreviaturas especiales.

Ejemplo: COBECOYOL = Corporación Benéfica El Coyol.

Columna 70

Nº de ejemplares:

Se anota el número de ejemplares de la misma recolección que existe en el herbario.

Columnas 72-79

Nº de registro:

Se ha iniciado un registro de las recolecciones del herbario el cual quedará dividido por grupos taxonómicos. Así, tenemos ahora las Pteridofitas, Gimnospermas, Monocotiledóneas y Dicotiledóneas, cada uno con su registro correspondiente.

Esta información sirve para evaluar el número de recolecciones que existen en el herbario.

Columna 80

Nº de tarjeta:

Indica si la tarjeta es la primera, segunda o tercera, sirve para tener un control de la información que se ha puesto en cada tarjeta.

## TARJETA 2

Columnas 1-4

Siglas del herbario:

En este caso las siglas son MEXU (Herbario Nacional, Instituto de Biología, U.N.A.M.)

Columnas 6-11

Fecha de recolección:

En esta información se codifica en forma numérica el mes, en las dos primeras columnas y el año se pone completo.

Ejemplo: 061954 = Junio de 1954

Columnas 13-19

Nº de recolección:

Corresponde al número de recolección de cada colector o colectores o bien de alguna institución.

Ejemplo: JB00325 = Jardín Botánico  
Nº 325.

Columnas 24-31

Altitud:

La altitud, dada en metros, se codifica en cinco columnas; pero, en ocasiones,

se anotan los límites entre los cuales se encuentra la planta, por ejemplo de 500 a 700 m, entonces se codifica 0500M-0700M.

#### Columnas 33-39

Latitud:

Usamos una G para abreviar grados, una M para los minutos y una N o una S para abreviar Norte o Sur.

Ejemplo: 25G 57MN = 25°57' Norte

#### Columnas 41-48

Longitud:

Codificada como en el caso anterior; se utiliza una E o una O para indicar Este u Oeste.

Ejemplo: 035G 25M0 = 035°25' Oeste

#### Columnas 51-59

Clave mapa eje X, clave mapa eje Y:  
Para cada localidad se anota tanto sus coordenadas, como una clave especial que corresponde a dichas coordenadas.

Ejemplo: 18G 19MN 095G 05M0  
20.0/56.0

#### Columna 60

Información de abundancia:

En los ejemplares aparece anotada la siguiente información:

M = muy abundante    A = abundante  
R = regular            E = escaso

#### Columna 62

Forma de vida:

Se codifica la información que aparece anotada en la etiqueta.

Ejemplo:            A = árbol

#### Columnas 63-64

Información ambiental:

Se registran datos muy variados, por ejemplo algunos referentes a climas:

H = húmedo; otros a relieve: E = acantilado, y otros más se refieren propiamente al hábitat: A = acuático.

#### Columnas 65-66

Tipo de vegetación:

Esta información es muy variada. Se toma el dato tal como aparece en la etiqueta y se hacen diccionarios para codificar esta información.

Ejemplo:            M = manglar

#### Columna 67

Textura y profundidad del suelo:

Información poco frecuente y muy variada. Se toma como aparece en el ejemplar, se codifica y se hace el diccionario correspondiente.

Ejemplo:            A = arenoso

#### Columna 68

Origen y composición química del suelo:  
Lo mismo que en el caso anterior, se codifica y se elabora el diccionario respectivo.

Ejemplo:            V = volcánico

#### Columnas 69-70

Usos:

Aunque es bastante raro encontrar esta información, creemos que en algunos casos es muy importante el tener este dato.

Ejemplo:            07 = forraje

#### TARJETA 3

##### Columnas 1-6

Tamaño de la planta:

Codificada en metros. Cuando aparecen dos datos, por ejemplo de 3 a 5 metros, se saca el promedio y se pone una P en el lugar de la M.

Ejemplos:            03. 0m, 04. 0p

Columnas 8-12

Tronco circunferencia (a.d.p.)

Se refiere al tamaño del tronco a la altura del pecho. Esta información es poco frecuente, pero creemos que puede ser de interés en algunos casos.

Ejemplo: A = ámbar

Columna 24

Tipo de asociación con animal.

Ejemplo: Hormigas = H

Columna 14

Presencia de flor u órgano correspondiente:

Esta información consideramos que es muy importante, pues ella dará idea de la época de floración de la planta, o bien, servirá para correlación en la identificación de material. Se anota solamente una S si el ejemplar posee la flor o una N en el caso contrario.

Columna 26

Polinización.

Ejemplo: Abejas = A

Columnas 56-59

Número cromosómico.

Ejemplo: n10, 2n20

Columna 16

Color de la flor:

Aunque esta información se puede considerar subjetiva y está sujeta a errores, creemos que es importante. Se codifica y se hace diccionario.

Estas 3 últimas informaciones, hasta el momento, no han aparecido en los ejemplares revisados. Siendo información que se considera tan importante se ha aumentado en el nuevo formato de la etiqueta de Veracruz, con el objeto de tener la información de la planta lo más completa posible.

Ejemplo: V = verde

Columnas 28-49

Nombre vulgar

Información muy importante que, aunque no es muy frecuente, creemos que debe anotarse. Se anota el nombre tal como aparece en el ejemplar.

Columna 18

Presencia de fruto u órgano correspondiente:

Tomada con el mismo propósito que en el caso de la flor. Esta información es menos frecuente. Se anota igualmente una S si el ejemplar posee el fruto o una N si no lo posee.

Ejemplo: Cabeza de negro

Columna 20

Látex (color)

Se elabora un diccionario para codificar el color del látex, que se encuentra anotado en la etiqueta.

Columna 51

Duración de la hoja

Anotamos una P en el caso de que sea perennifolia o una C cuando es caducifolia.

Ejemplo: R = rojo

Columna 22

Resina (color)

Se procede de la misma forma que en el caso anterior. Ambas informaciones son poco frecuentes.

Columnas 53-54

Información de Tipo (holotipo, isotipo, etcétera).

Con la implicación de esta información creemos que en un futuro podríamos tener una lista de los Tipos existentes en el Herbario Nacional. Se codifica y se elabora diccionario.

Ejemplo: TH = holotipo

Columnas 61-62

Tipo de material

Ésta es una información adicionada últi-

mamente, con el objeto de anotar si el ejemplar es sólo una fotografía de algún ejemplar o si es la planta más una fotografía de la misma.

Ejemplo: F = foto PF = planta y foto

## EQUIPO ELECTRÓNICO UTILIZADO

El procesamiento electrónico de los datos se hizo en el Centro de Cálculo Electrónico de la UNAM y en el Instituto Mexicano del Petróleo. La razón de hacerlo es dos lugares diferentes es que fue necesario hacer dos programas, uno exclusivamente para la recuperación general de la información y el otro para la elaboración de los mapas de distribución. Lo anterior se debe a que en la época en que fue hecho el programa no existía ningún graficador disponible en el Centro de Cálculo Electrónico.

El programa para la recuperación de la información está hecho en lenguaje Algol y fue procesado en la computadora B-5500.

Graficación:

Teniendo la experiencia del consumo de tiempo y de los problemas que representa la elaboración manual de mapas de distribución y, conociendo de antemano algunos trabajos publicados acerca del uso de computadoras para la resolución de estos problemas (Perring, 1963; Soper, 1964; Soper, 1966; Hawkes, Kershaw y Readett, 1968), se decidió probar el sistema.

La selección del equipo adecuado y el procesamiento que se le daría a los datos, no se decidió sino después de probar varios métodos.

Primeramente señalaremos que el mapa del Estado de Veracruz fue dividido en 93 pequeños cuadros de 20 minutos por lado (Fig. 2) (Gómez-Pompa y Nevling, 1967). A cada cuadro se le

asignó un número y se le subdividió a su vez en 100 pequeños cuadros (Fig. 3). Estos pequeños cuadros son utilizados en el campo para localizar los puntos de recolección, de acuerdo con una numeración progresiva que presenta todo el mapa. Esta numeración es codificada en una clave especial que se corresponde con las coordenadas cartesianas, con el objeto de tener acceso a la información de dos maneras.

Poco después fue hecho un pequeño programa para utilizar la computadora CDCG-15, la cual tenía entre otros accesorios, el graficador digital PA-3 (Gómez-Pompa y Olvera F., 1969). De aquí se obtuvieron las primeras pruebas del uso de computadoras en la graficación de mapas de especies; en ello también había algunos problemas tales como el que los símbolos usados no salían con una claridad óptima, lo que representaría un problema para la impresión; el mapa tenía que ser ajustado al papel graficador al iniciar la graficación, o bien se hubiese tenido que imprimir el mapa en el papel graficador. Por otra parte, el programa sólo comprendía la parte correspondiente a graficación.

Finalmente se probó el graficador IBM-1627 CalComp, modelo 2, que opera automáticamente bajo control de la computadora IBM 1130-3D (Fig. 4), la cual tiene una capacidad de memoria de 32K bytes con palabras de 16 bits. Asimismo, se hace uso de la lectora de tarjetas IBM-2501.

El uso de este tipo de graficador tiene la ventaja de poder graficar conti-

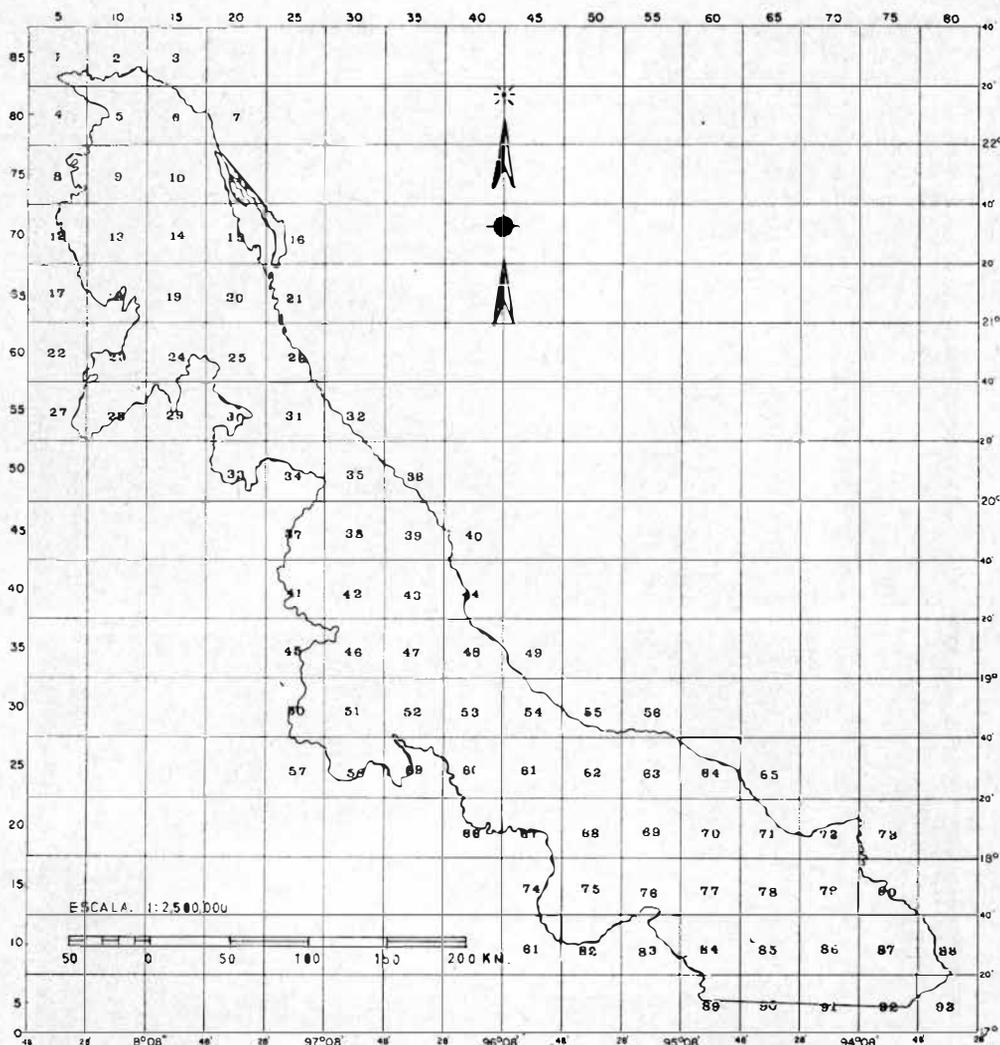


Fig. 2. Mapa general del Estado de Veracruz, que muestra los 93 cuadros en que fue dividido.

nuamente bajo control de programa, a velocidad adecuada y evitar estar cambiando de papel, la pluma puede despegarse y empezar a graficar en cualquier punto. El proceso de graficación va acompañado de una rotación mecánica del tambor del graficador, sincronizada con el movimiento transversal de la pluma a través del papel, lo que da una gran exactitud.

El sistema contiene un grupo de sub-

rutinas propias para el graficador que el operador puede aplicar de acuerdo a sus necesidades o bien desarrollar una subrutina especial para satisfacer un propósito particular.

Por otra parte, este graficador tiene una pluma con tinta negra, la cual es bastante visible y da por resultado una graficación óptima desde el punto de vista de visibilidad.

●tra innovación hecha al sistema es

el hacer el contorno del mapa por medio del aparato OSCAR F (Fig. 5). Tal aparato está diseñado para acelerar el análisis de varios registros de trazos continuos. El sistema convierte información analógica a valores digitales y da una lectura automática en una perforadora de tarjetas. La secuencia de las lecturas de salida es controlada por medio de un tablero.

Este aparato tiene varias aplicaciones en Ingeniería. El uso que nosotros le dimos es sencillo, y consiste en lo siguiente: en la pantalla fluorescente se coloca el mapa, se van señalando los puntos del contorno y se marcan. Estos puntos aparecen automáticamente en una perforadora de tarjetas que se encuentra conectada al sistema. Los datos que aparecen en las tarjetas, correspon-

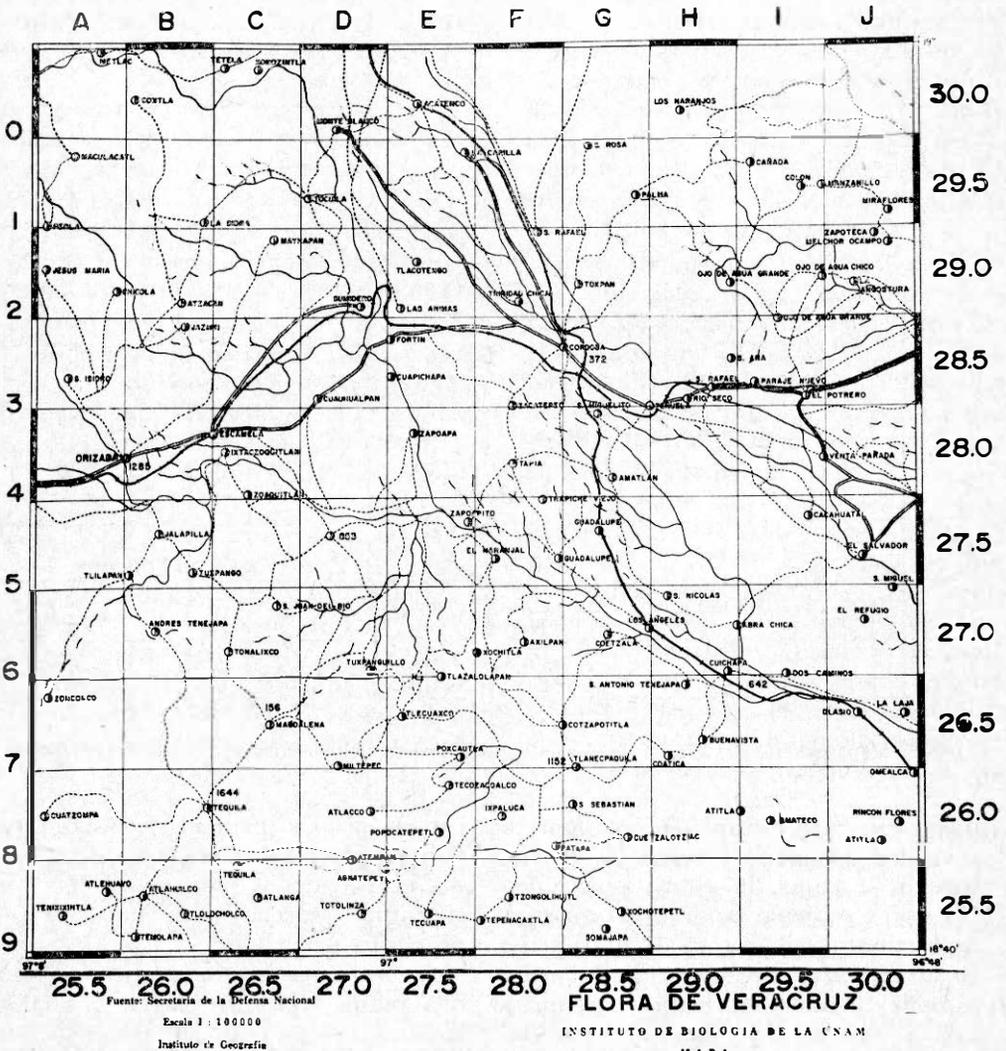


Fig. 3. Uno de los 93 cuadros mostrando la subdivisión en 100 cuadros y la numeración (a la derecha y abajo) que sirve para tomar la clave de localidad.

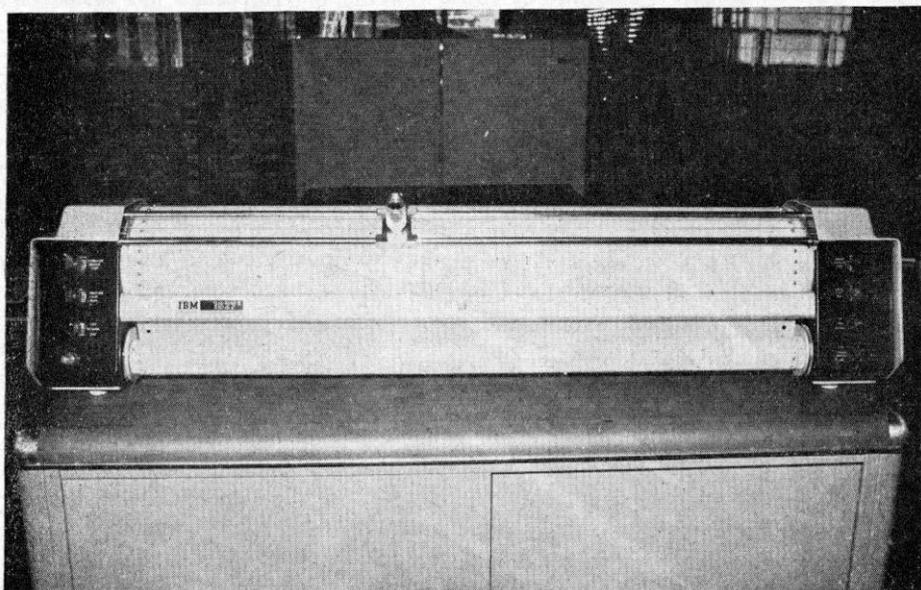


Fig. 4. Graficador IBM-1627 (A) que opera bajo control de la computadora IBM-1130-3D (B).

den a las coordenadas de los ejes X e Y de los puntos tomados y estas tarjetas puestas en el programa de graficación, servirán para delinear el contorno del mapa.

Es conveniente señalar también que

todas las leyendas que aparecen en los mapas están hechas bajo control de programa.

El programa está hecho en lenguaje FORTRAN.

Para la elaboración de los mapas se

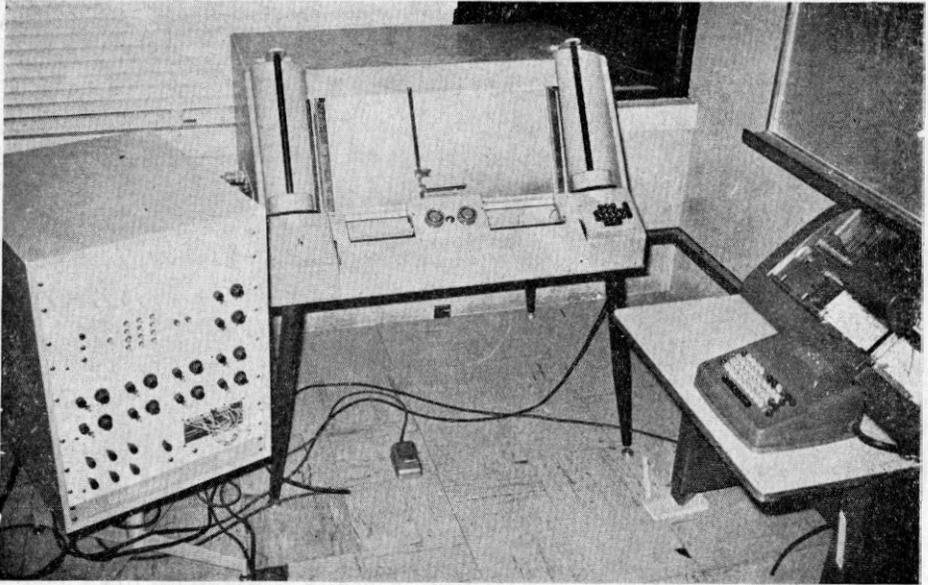


Fig. 5. Aparato OSCAR F, del Centro de Cálculo Electrónico, conectado a una perforadora de tarjetas.

usa, de la información contenida en las tres tarjetas IBM:

- a) La información de la familia.
- b) El número del género.
- c) El número de la especie.
- d) Los datos de las coordenadas X e Y, los cuales son tomados de la clave del mapa por mayor comodidad del manejo de esta información.

En resumen, las pruebas hechas con el graficador conectado a la computadora (Gómez-Pompa y Olvera Fonseca, en prensa), dieron la pauta para llegar a los resultados actuales. En ellos se delinea el contorno directamente por medio de la computadora, se grafican las localidades por medio de las letras del alfabeto, o con números o signos adi-

cionales cuando éstas no son suficientes, y se enlistan en el mismo mapa las especies o géneros que se encuentran en cada localidad.

En el trabajo de graficación hecho por Soper (1969), se presentan unos mapas muy similares a los nuestros, pero en ellos se usan sólo tres o cuatro signos diferentes para graficar.

El sistema usado para la Flora Británica (Perring, 1963), es un sistema donde se grafican las localidades por medio de una tabuladora. El sistema no es conveniente para nuestro estudio pues, como menciona Soper (1964), es propio para graficar zonas pequeñas y donde además se tiene una recolección masiva, debido a que la graficación se hace por columnas, lo que hace al sistema poco flexible.

## RESULTADOS

Los resultados son una serie de listados generales (Fig. 6) que se obtuvieron como respuesta a un grupo de pre-

guntas tomadas al azar, que consideramos pudieran tener importancia para el programa Flora de Veracruz, como

FAMILIAS DE VERACRUZ

GENEROS POR FAMILIA

AMARILLYDACEAE  
ARACEAE  
BROMELIACEAE  
BURMANNIACEAE  
CANNACEAE  
COMMELINACEAE  
CYCADACEAE  
CYCLANTACEAE  
CYPERACEAE  
DIOSCOREACEAE  
GRAMINEAE  
HAEMODORACEAE  
HYDROCHARITACEAE  
IRIDACEAE  
JUNCACEAE  
LILIACEAE  
MARANTACEAE  
MAYACACEAE  
MUSACEAE  
ORCHIDACEAE  
PALMAE  
PINACEAE  
PODOCARPACEAE

AGAVE  
BOMAREA  
CRINUM  
HYMENOCALYX  
HYPOXIS  
SPREKELIA  
ZEPHYRANTHES

ARACEAE

ANTHURIUM  
ARISAEMA  
MONSTERA  
PHILODENDRON  
SPATHIPHYLLUM  
SYNGONIUM  
XANTHOSOMA

BROMELIACEAE

AECHMEA  
CATOPSIS

## ESPECIES POR FAMILIA

## AMARILLYDACEAE

AGAVE RUBESCENS

AGAVE XALAPENSIS

BOMAREA EDULIS

BOMAREA SP.

CRINUM ERUBESCENS

CRINUM SP.

HYMENOCLALYX SP.

HYPOXIS DECUMBENS

HYPOXIS MEXICANA

SPREKELIA FORMOSISSIMA

ZEPHYRANTHES LINLEYANA

## ARACEAE

ANTHURIUM CRASSINERVIUM

ANTHURIUM MIRANDAE

ANTHURIUM MONTANUM

ANTHURIUM MYOSURROIDES

ANTHURIUM SCANDENS

ANTHURIUM SCHLECHTENDALLII

ANTHURIUM SILVIGAUDENS

por ejemplo: ¿qué familias, géneros y especies han sido recolectadas en Veracruz?: pero que de ninguna manera representa la totalidad de la información deseada. En términos generales se pueden considerar solamente como un ejemplo de la versatilidad de los métodos automáticos de recuperación de información.

Por otra parte, se obtuvo una serie de mapas elaborados automáticamente

(Figs. 7, 8 y 11), que creemos tienen una gran exactitud con respecto a la graficación. Estos mapas representan la distribución de géneros y especies por familias, en Veracruz, que se encontraron en el Herbario Nacional de la UNAM, hasta el año de 1968, fecha en que se empezó formalmente el programa intensivo de recolección.

Como se puede apreciar, la graficación está hecha de acuerdo a localidades

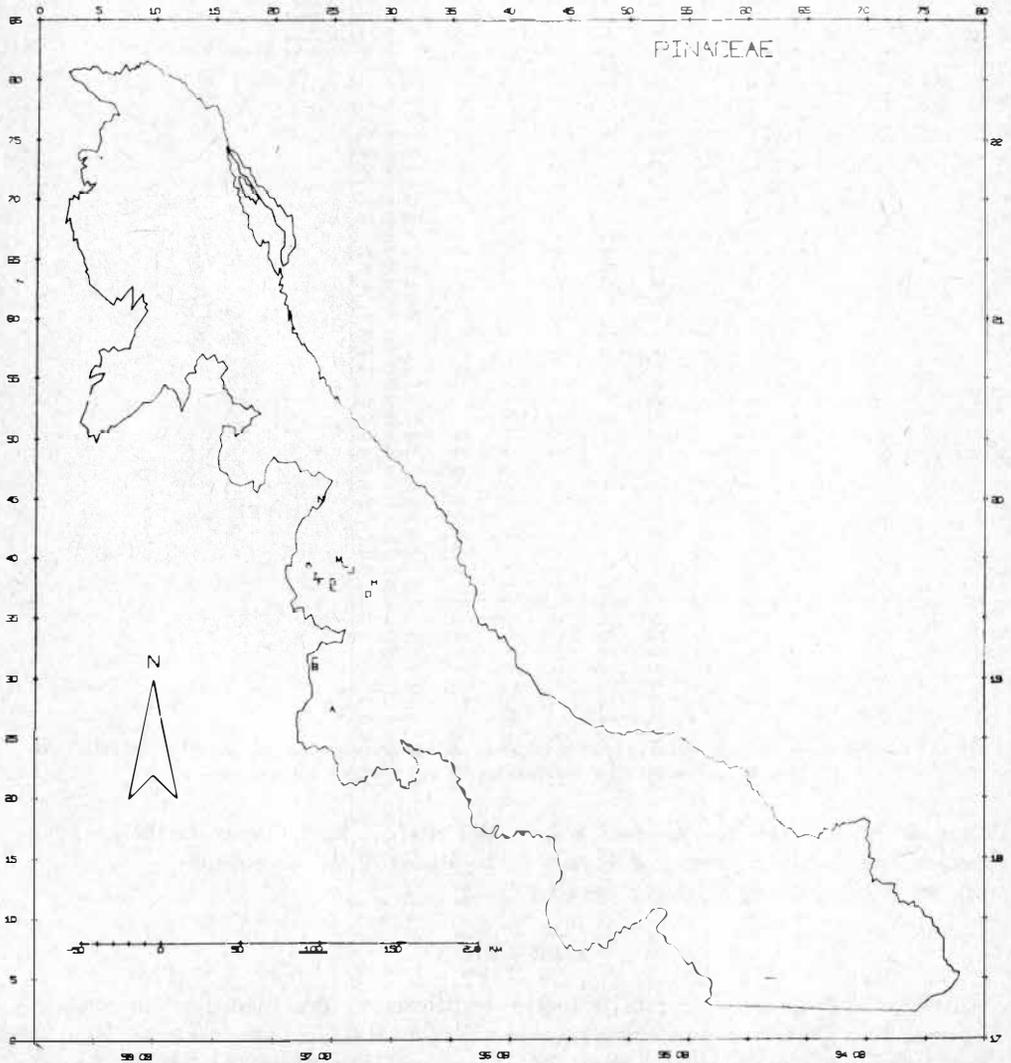


Fig. 7. Distribución de la familia Pinaceae en Veracruz.

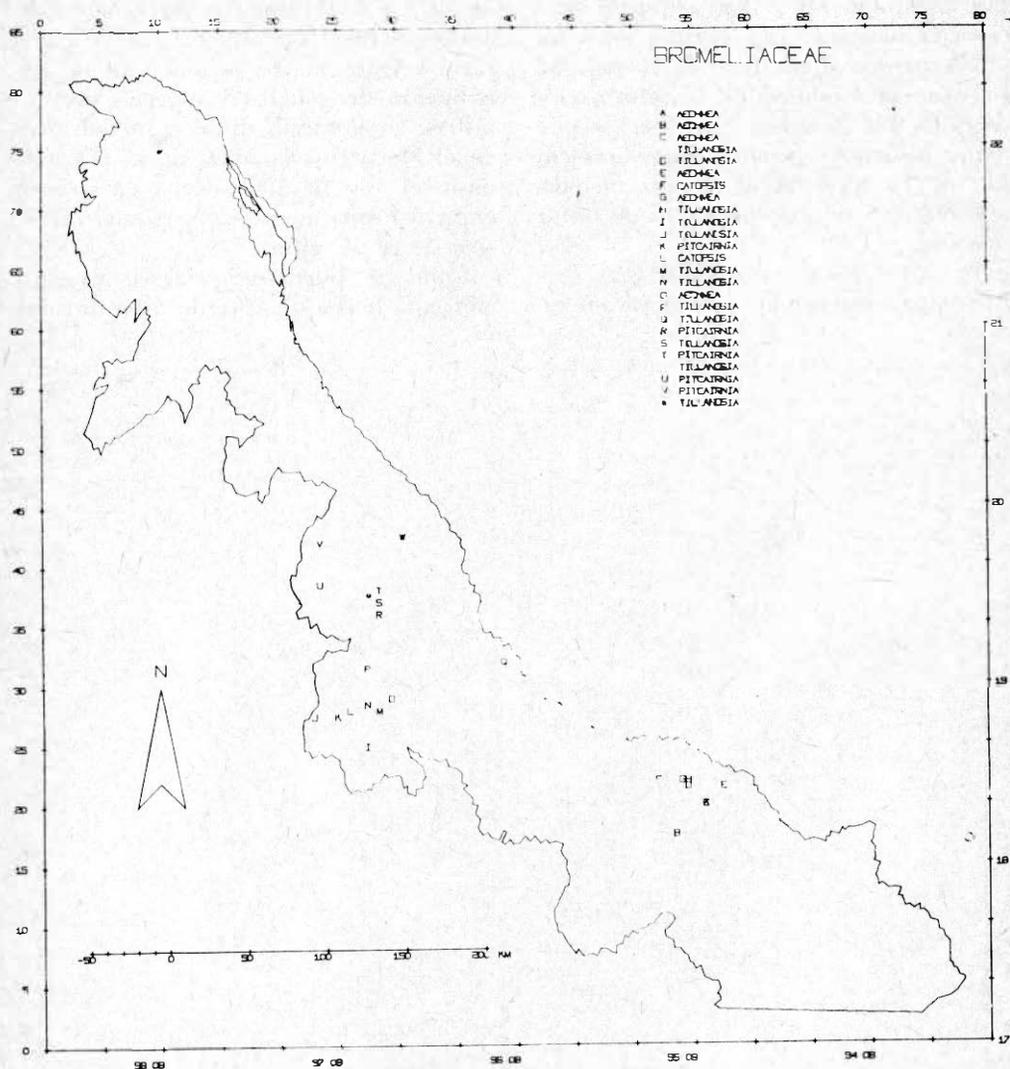


Fig. 8. Distribución de los géneros de la familia Bromeliaceae. En el margen derecho del mapa se anota en qué localidades se encuentran los géneros.

donde se encuentran los géneros o las especies, que por otra parte, se encuentran mencionados en la parte derecha

del mapa, para mayor facilidad en la localización de las mismas.

### DISCUSIÓN

Durante el desarrollo de este trabajo se presentaron diversos problemas derivados del sistema usado. Entre ellos destacan los problemas concernientes a la

codificación. Por ejemplo, con respecto a la localidades encontramos que, en ocasiones, el espacio para su codificación no era suficiente para anotar la locali-



dad completa, por lo que, a criterio personal se tenían que omitir algunos datos o bien abreviar. Por el contrario, en otros casos, se menciona solamente el nombre de algún poblado, el cual es difícil de localizar en el Estado, por no tener el municipio o cabecera a que pertenece. De todo esto resulta que la información en algunos casos es muy subjetiva y que está completamente sujeta al criterio o decisión del codificador.

Otro ejemplo sería que al codificar los nombres de los autores de especies, se toman las letras a codificar a criterio muy personal; asimismo, las letras o números usados en los diccionarios no siguen una regla especial para su selección, sino que se van dando a criterio propio o bien al azar.

En contraste con estas desventajas, tenemos que con la información codificada, se usa un mínimo de tarjetas IBM, en este caso tres (Fig. 9). En el primer trabajo de este tipo (Scheinvar y Gómez-Pompa, 1969), se usaron dos tarjetas para codificar la información; sin embargo, en este trabajo se pensó que no eran suficientes para un trabajo más específico de investigación como es el de la Flora de Veracruz, para lo cual, además, fue diseñada una nueva etiqueta de recolección que aportaría más información, con lo cual se hacía indispensable el uso de otra tarjeta más.

El formato general es casi el mismo al utilizado por Scheinvar y Gómez-Pompa (1967), salvo, desde luego, la información adicional en este trabajo; no así los resultados, que aunque tienen el mismo objetivo, o sea la recuperación de la información, no se presentan de la misma manera. En el primer caso, se presentan las listas de respuestas a las preguntas sin decodificar, mientras que, en el segundo, se pretende obtener la información completamente decodificada.

En general, se ha visto que, en ambos

casos, hubiese sido más conveniente usar codificación numérica en lugar de alfabética o alfanumérica como fue hecho, pues esto tiene restricciones con respecto al lenguaje de máquina que se va a emplear.

En el sistema utilizado para la Flora de Veracruz (Gómez-Pompa y Nevling, en prensa), se ha abandonado la idea de decodificar y se está pasando la información que estaba codificada en dos tarjetas, a un nuevo formato de diecinueve tarjetas sin codificar (Fig. 10).

Con este nuevo sistema se resuelven algunos problemas presentes en los dos sistemas anteriores. Por ejemplo, se ahorra tiempo al no pasar a codificación, sino directamente a perforación, lo cual implica también la disminución de posibles errores hechos al codificar. La información puesta tal cual en las cintas magnéticas, no necesita un programa especial de decodificación, sino simplemente el de recuperación de información.

Otros problemas encontrados, pero que son de carácter más específico, son los referentes a la elaboración de los mapas de distribución de especies y los de programación.

El primer problema, desde luego, fue el encontrar el equipo adecuado para este propósito y durante el proceso del mismo fueron surgiendo problemas colaterales. Por ejemplo, el mapa de Veracruz usado, aun cuando es el más completo disponible no está actualizado, por lo que es frecuente que diversas localidades no se encuentren en el mapa o que éstas estén localizadas en sitios erróneos.

Un problema evidente en algunos mapas fue que ciertas localidades aparecen graficadas en el mar, lo cual puede deberse a tres factores:

- 1º Que el contorno de los mapas ampliados no coincida con el del mapa general.

TRACH.	NO. REP. 03 MEXICO	MEXM	COMPOS65	000149019H
COL. L. NEVLING Y A. GOMEZ-POMPA		NO. 255	COMPOS65	000149018H
		DET. R. McVAUGH	COMPOS65	000149017H
		03/09/1967	COMPOS65	000149016H
		FL. MORADA	COMPOS65	000149015H
ABUND. REGULAR	HERBACEA	TAN. 50CM	COMPOS65	000149014H
			COMPOS65	000149013H
			COMPOS65	000149012H
			COMPOS65	000149011H
SUELO ARENOSO DE REAC. POSITIVA AL HCL.			COMPOS65	000149010H
		INCL. PLANO-ALUVIAL	COMPOS65	000149009H
SELVA BAJA		VEG. SECUNDARIA	COMPOS65	000149008H
		ALT. 25M	COMPOS65	000149007H
TAMAULIPAS, 10KM SA TAMPICO-PANUCCO		MAPA 75.5/05.5	COMPOS65	000149006H
4KM AL SUR DEL LIMITE DE LOS ESTADOS DE VERACRUZ-			COMPOS65	000149005H
MEXICO		EDI. VERACRUZ	COMPOS65	000149004H
PALAFOXIA ROSEA VAR. ROBUSTANRYDBK			COMPOS65	000149003H
			COMPOS65	000149002H
NO. REC. INF	0001490	COMPOSITE	COMPOS65	000149001H

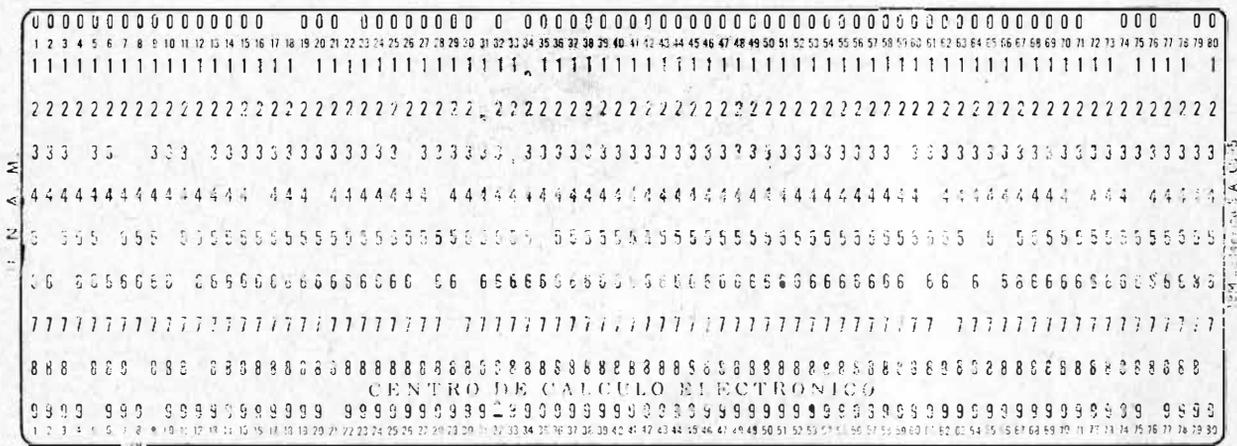


Fig. 10. El nuevo formato de diecinueve tarjetas sin codificar.

2º Que cuando se delincó el contorno del mapa en el OSCAR F, los puntos tomados en algunas pequeñas curvas hayan quedado ligeramente fuera del límite de tierra y como el mapa fue hecho al doble del tamaño delineado, la pequeña distancia de posible error, aparezca aumentada al doble, lo que haría que los puntos ciertamente se localicen en el mar.

3º Que la numeración dada para tomar nuestras claves resulta un tanto amplia y que debió hacerse más pequeña o minuciosa para que así los puntos graficados pudiesen aparecer con un error mínimo cercano a cero, en lugar de 3.5 km como sucede con la numeración actual tomada.

Por otra parte, viendo hacia el futuro, nuestro sistema de graficación tiene

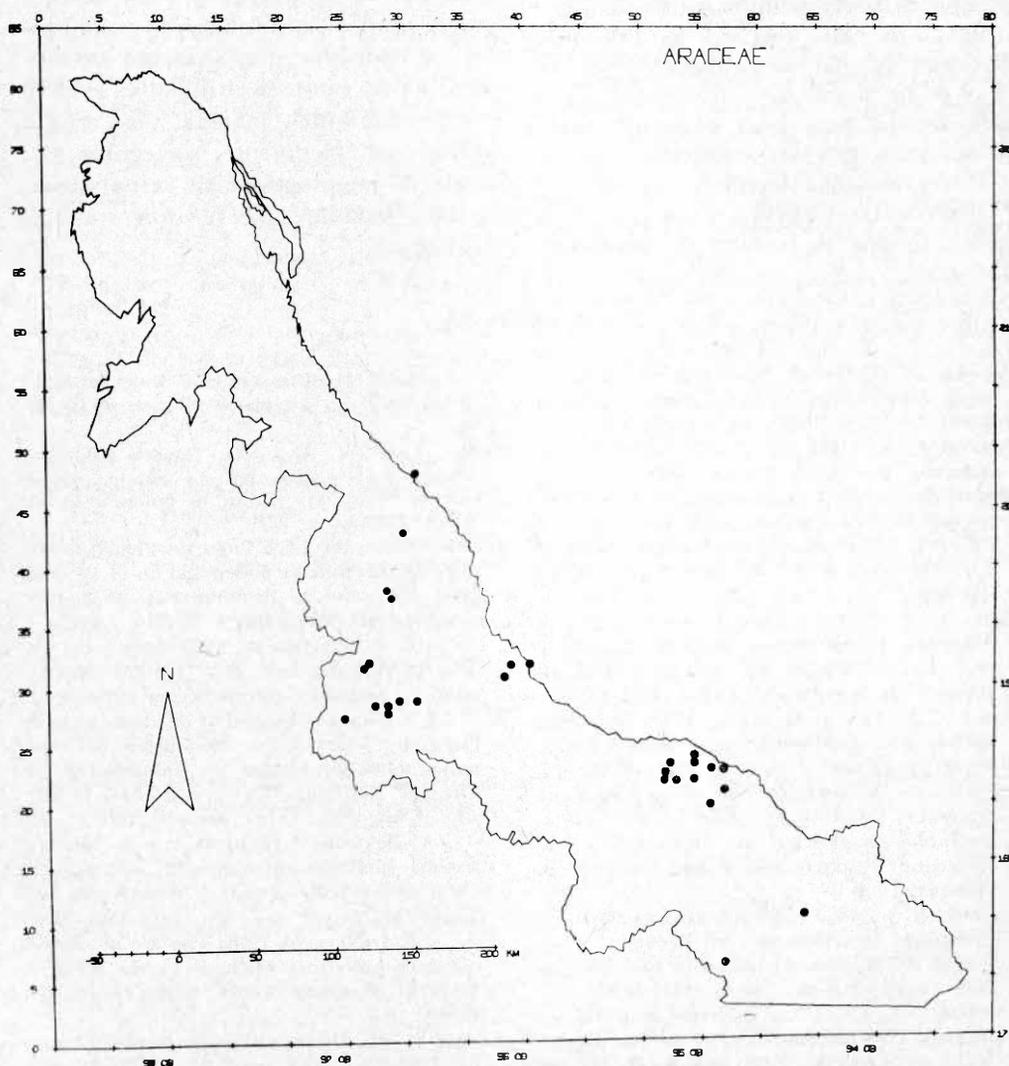


Fig. 11. Este mapa muestra el tipo de graficación que probablemente se use en la Flora de Veracruz.

una posible limitación, y es la restricción con respecto al número de signos disponibles y la enorme cantidad de posibles localidades. De cualquier manera, por ahora el sistema funcionó perfectamente, y si en el futuro se necesitase de la graficación de un número elevado de localidades, se tendrá que sacrificar un poco de la exactitud al poner dos signos combinados para cada localidad, o bien, quizá se encuentre un sistema de graficación más adecuado.

Por otro lado, ésta no ha sido más que una prueba de graficación, en la cual se ha usado la metodología antes mencionada, pero para nuestros propósitos finales, la graficación sería similar a la representada en la figura 13.

Una de las objeciones que se hacen a este sistema es la falta de confiabili-

dad de datos (Thomas, 1969). Debido a que el presente trabajo es una reproducción fiel de los datos que aparecen en las etiquetas de los ejemplares, es factible que pudieran aparecer errores en los datos referentes a familia, género y especie.

Tales errores se deben a una identificación errónea del material por parte de los recolectores o de la persona que identificó el ejemplar, si consideramos que se usan 39 informaciones y sólo las tres mencionadas anteriormente son susceptibles de error, la utilidad es evidente. De cualquier manera, los errores podrán ser fácilmente corregidos por medio de un programa de actualización de información, una vez que éstos se localicen.

#### LITERATURA

- ARNOLD, R. R., H. C. HILL y A. V. NICHOLS, 1966. *Introduction to data processing*. John Wiley and Sons., Inc., New York. 326 p.
- AWARD, F. M., 1970. *Proceso de datos en los negocios*. Ed. Diana, México. 510 p.
- BATRA, L. R., 1970. Automation of procedures of the mycology investigation unit. In: Crovello, T. J. and R. D. MacDonald. Index of EDP-IR projects in systematics. *Taxon* 19 (1):64.
- BELL, C. R., 1970. Inventory of the Flora of Southeast United States. In: Crovello, T. J. and R. D. MacDonald. Index of EDP-IR projects in systematics. *Taxon* 19 (1):64.
- BESCHEL, R. E. y J. H. SOPER, 1970. The automation and standardization of certain herbarium procedures. *Canad. J. Bot.* 48:547-554.
- BIOSCIENCIAS INFORMATION SERVICE OF BIOLOGICAL ABSTRACTS (Ed.), 1970. *Guide to the indexes for biological abstracts and biosearch index*. Biological Abstracts, Inc. Philadelphia, Pennsylvania, 20 p.
- COWAN, R. S., 1970. Computerization of Index Nominum Genericorum. In: Crovello, T. J. and R. D. MacDonald. Index of EDP-IR projects in systematics. *Taxon* 19 (1):66.
- CROVELLO, T. J., 1970. Automated retrieval information in the Edward Lee Greene Herbarium. In: Crovello, T. J. and R. D. MacDonald. Index of EDP-IR projects in systematics. *Taxon* 19 (1):67.
- , y R. D. MACDONALD, 1970. Index of EDP-IR projects in systematics. *Taxon* 19 (1):63-76.
- DALLA TORRE, C. G., y H. HARMS, 1900-1907. *Genera Siphonogamarum ad systema Englerianum*, conscripta, sumtibus guilelmi Eugelmani, Lipsiae, p. 921.
- GÓMEZ-POMPA, A., 1969. Flora de Veracruz. Manuscrito del trabajo presentado en el IV Congreso Mexicano de Botánica, 8 al 11 de septiembre 1969. Monterrey y Saltillo. Mérito.
- , y L. I. NEVLING JR., 1967. Mapas del Estado de Veracruz. *Inst. Biol., UNAM, México*, núm. 1. Enero 15 (reproducción en Xerox).
- , y S. OLVERA. Procesado de datos para la Flora de Veracruz. In: Resúmenes del Simposio sobre Problemas de Información en Ciencias Naturales. *Publ. Esp. No. 1 Inst. Biol. Univ. Nat. Autón. México* 1969.
- , A. BUTANDA, L. SCHEINVAR y A. MUHLIA. Sistema bibliográfico computarizado para la Flora de Veracruz. *An. Inst. Biol. Univ. Nat. Autón. México* 43, Ser. Bot. (1): 1-10. 1972.
- , y L. I. NEVLING, 1973. The use of electronic data processing methods in the Flora of Veracruz Program. *Contr. Gray Herb.* 203: 4964.
- GOULD, S. W., 1970. International plant index. In: Crovello, T. J. and R. D. MacDonald. Index of EDP-IR projects in systematics. *Taxon* 19 (1):68.

- HAWKES, G. J., B. L. KERSHAW y R. C. READETT, 1968. Computer mapping of species distribution in a county flora. *Watsonia* 6(6):350-364.
- LAWRENCE, G. H. M., A. F. G. BUCHHEIM, G. S. DANIELS y H. DOLEZAL (Eds.), 1968. B-P-II, *Botanicum-Periodicum-Huntianum*. Hunt Botanical Library. Pittsburgh, Pennsylvania. 1063 p.
- LEDLEY, R. S., 1960. *Digital computer and control engineering*. McGraw Hill Book Co., Inc. New York. 853 p.
- MACDONALD, R. D., 1970. Plant records center. In: Crovello, T. J. and R. D. MacDonald. Index of EDP-IR projects in systematics. *Taxon* 19(1):71.
- MORSE, L. E., 1971. Specimen identification and Key construction with time-sharing computers. *Taxon* 20(2/3):269-282.
- BEAMAN, J. H. y S. G. SUETLER, 1968. A computer system for editing diagnostic keys for Flora North America. *Taxon* 17:479-483.
- PERRING, F. H. y S. M. WALTERS, 1962. *Atlas of the British Flora*. Botanical Society of the British Isles. Thomas Nelson and Son., Ltd., London. 432 p.
- PERRING, F. H., 1953. Data processing for the Atlas of the British Flora. *Taxon* 12(5):183-190.
- ROGERS, D. J., 1970. Development of a prototype system for the storage and retrieval of biological data. In: Crovello, T. J. and R. D. MacDonald. Index of EDP-IR projects in systematics. *Taxon* 19(1):74.
- SCHILNVAR, L., A. GÓMEZ-POMPA y L. ALONSO. 1967. Sistema automático de recuperación de información para el Herbario Nacional del Instituto de Biología, UNAM. *An. Inst. Biol. Univ. Nat. Autón. México* 38 Ser. Bot. (1): 203-250.
- , y A. GÓMEZ-POMPA, 1969. Algunos métodos automáticos para la elaboración de etiquetas de herbario. *Bol. Soc. Bot. México* 39: 73-93.
- , y A. GÓMEZ-POMPA. 1969. Proyecto piloto de recuperación automática de información del Herbario Nacional, UNAM. In: Resúmenes del Simposio sobre Problemas de Información en Ciencias Naturales. *Publ. Esp. No. 1 Inst. Biol. Univ. Nat. Autón. México*: 30.
- SOKAL, R., F. J. ROHLF y T. J. CROVELLO, 1970. Automatic retrieval of collection data and character information in the genera *Pemphigus* and *Populus*. In: Crovello, T. J. and R. D. MacDonald. Index of EDP-IR projects in systematics. *Taxon* 19(1):75.
- SOPER, J. H., 1964. Mapping the distribution of plants by machine. *Canad. J. Bot.* 42:1087-1100.
- , 1966. Machine-plotting of phytogeographic data. *Canad. Geogr.* 10:15-26.
- , 1969. The use of data processing methods in the herbarium. *An. Inst. Biol. Univ. Nat. Autón. México* 40 Ser. Bot. (2) (en prensa).
- THROCKMORTON, T. D., 1970. Daffodil data bank. In: Crovello, T. J. and R. D. MacDonald. Index of EDP-IR projects in systematics. *Taxon* 19(1):76.
- THOMAS, J. H., 1969. Correct determination and the use of automated data processing in systematic collections, *Publ. Esp. No 1. An. Inst. Biol. Univ. Nat. Autón. México*: 36.

Nota: Debido a que este trabajo fue elaborado en 1967-1968 se conservó el contenido original, actualizando solamente algunas citas bibliográficas.