

FloraMap™

**Una Herramienta para Computadora que Permite Predecir la
Distribución de Plantas y de Otros Organismos
en Estado Silvestre**

Versión 1.01 2002

**P.G. Jones y A. Gladkov
Editada por Annie L. Jones**

La misión del Centro Internacional de Agricultura Tropical (CIAT) es reducir el hambre y la pobreza en los trópicos mediante una investigación colaborativa que mejore la productividad agrícola y el manejo de los recursos naturales.

El CIAT es uno de los 16 centros internacionales de investigación agropecuaria, conocidos ya como los Centros de la Cosecha del Futuro, que son auspiciados por el Grupo Consultivo para la Investigación Agrícola Internacional (GICIAI).

El trabajo del CIAT es financiado por un gran número de países, organizaciones para el desarrollo regional e internacional y fundaciones privadas. En el 2001, los siguientes países son donantes del CIAT: Alemania, Australia, Bélgica, Brasil, Canadá, Colombia, Dinamarca, España, Estados Unidos de América, Francia, Holanda, Italia, Japón, México, Noruega, Nueva Zelanda, Perú, Reino Unido, Sudáfrica, Suecia, Suiza y Tailandia. Entre las entidades donantes están el Banco Asiático de Desarrollo (ADB), el Banco Interamericano de Desarrollo (BID), el Banco Mundial, el Centro Internacional de Investigaciones para el Desarrollo (CIID), el Fondo Internacional para el Desarrollo Agrícola (IFAD), la Fundación Ford, la Fundación Kellogg, la Fundación Nippon, la Fundación Rockefeller, la Fundación Wallace y la Unión Europea (UE).

La información y las conclusiones contenidas en esta publicación no reflejan necesariamente los puntos de vista de los donantes.

FloraMap™

**Una Herramienta para Computadora que Permite Predecir la
Distribución de Plantas y de Otros Organismos
en Estado Silvestre**

Versión 1.01 2002

**P.G. Jones y A. Gladkov
Editada por Annie L. Jones**

Derechos de Autor del Manual © CIAT 1998-1999
Derechos de Autor del Software © CIAT 1978-1999

FloraMap esta escrito usando MapObjects LT de ESRI.
MapObjects LT es una marca registrada del Environmental
Systems Research Institute, Inc. Derechos de Autor 1999.
Environmental Systems Research Institute, Inc. Todos
Derechos Reservados.

Centro Internacional de Agricultura Tropical
International Center for Tropical Agriculture
Apartado Aéreo 6713
Cali, Colombia

Fax: (57-2) 4450-073
Correo electrónico: p.jones@cgiar.org

Impreso en Colombia
xxxx de 2002

Jones, P.G.; Gladkov, A.

FloraMap: una herramienta para computadora que permite predecir
la distribución de plantas y de otros organismos en estado silvestre;
versión 1.01 en español, 2002. --Editada por Annie L. Jones. -- Serie CD-
ROM del CIAT. -- Cali, Colombia : Centro Internacional de Agricultura
Tropical, 2002.

1 CD-ROM + guía (XX p.)

I. Gladkov, A. II. Jones, Annie L. III. Centro Internacional de
Agricultura Tropical.

Contenido

Introducción	1
Cómo Usar este Manual	2
1. Ponerse en Marcha	3
Requerimientos Mínimos de Equipo y de Software	3
Instalación	3
2. Tutorial	5
Configuración del Mapa	6
Verificación de los Datos	10
Análisis de Componentes Principales	22
Análisis de Probabilidad y Mapeo	25
Análisis de Agrupación	29
Conclusión	38
3. Sección de Referencia para el Usuario	40
La Ventana del Mapa	40
Iconos de Servicio del Menú Principal	46
La Ventana de Control de Capas	57
La Ventana de Análisis de Componentes Principales	62
La Ventana de Agrupación	65
4. Teoría	69
Superficies Climáticas	69
Estandarización de Fechas de Clima (Rotación)	70
Cálculos del Modelo	76
Cálculos de Probabilidad	79
Probabilidades Divergentes	82
Análisis de Grupos	84
Bibliografía	91
Apéndice A	93
Tipos de Archivos FloraMap	93
Creación de Archivos de Puntos de Acceso	94
Glosario	96
Índice	101

Introducción

FloraMap es un sistema para generar la distribución pronosticada o las posibles zonas de adaptación, para organismos naturales cuando se conoce poco o nada sobre la fisiología detallada del organismo. Se asume que el clima en el punto de recolección de un conjunto de individuos es representativo del rango de ambientes en que habita el organismo. En el caso de las plantas, éstas son, generalmente, accesiones de las colecciones de germoplasma o especímenes de herbarios.

El clima en estos puntos de recolección se usa para orientar la calibración cuando se computa un modelo de probabilidad del clima. El método usa un resultado de transformación matemática Fourier para estandarizar la sincronización climática y un análisis de los componentes principales (PCA, el acrónimo inglés) para generar una distribución de probabilidad en múltiples dimensiones. Los puntajes de componente principal se usan para mapear la superficie de probabilidad a partir de un conjunto de superficies climáticas interpoladas. El sistema se ha usado para guiar la recolección de plantas, para investigar la variación taxonómica y genética, y para mapear las plagas de cultivos y sus depredadores potenciales.

El sistema es el resultado de más de 20 años de trabajo en el CIAT y reúne, en un interfaz fácil de usar (o al menos no agresivo), algunas de las técnicas que se han desarrollado para hacer frente a los requerimientos de los científicos del CIAT para hacer estos análisis. Peter Jones desarrolló la base de datos climáticos, las superficies interpoladas y algunas de las funciones de manejo del clima. Nick Galwey las reunió con el PCA en un paquete Genstat durante una licencia de estudio en el CIAT, que fue registrado en Jones et al. (1997). Alexander Gladkov puso en marcha el actual sistema para ejecutar bajo Windows. Nos gustaría agradecer a los científicos del CIAT que, en el transcurso de los años, nos han impulsado a extender las capacidades del sistema, en especial David Wood, Anthony Bellotti, Steve Beebe y Joe Tohme. También extendemos un agradecimiento especial a Luigi Guarino por sus sugerencias detalladas, muchas de las cuales se usaron para mejorar el manual.

Los menús y ventanas de FloraMap están todavía en inglés en la pantalla. Por eso son explicados en inglés y español donde se encuentran por primera vez en el texto, luego se usa las palabras en inglés para corresponder con la pantalla. Al final del manual el lector puede encontrar un glosario para su ayuda.

Cómo Usar Este Manual

La primera regla es **¡LEERLO!** La mayoría de los usuarios de computadora recurren al manual de usuarios cuando todo lo demás fracasa; los autores son no diferentes a los demás usuarios. Sin embargo, el análisis que usted hará en FloraMap corresponde a un tipo altamente específico de mapeo. Es fácil aplicar mal los algoritmos y aún generar un mapa de aspecto bonito.

De incluir Microsoft una sección sobre la composición de sonetos shakesperianos en el manual de Word, el usuario se sorprendería, con toda razón. FloraMap no pretende ser un sistema generalizado de análisis de datos y la correcta interpretación del análisis especializado requiere de una buena comprensión de las bases de datos y las opciones usadas. Esperamos que usted encuentre tiempo para examinar el TUTORIAL con nosotros, paso a paso.

En primer lugar, vaya al TUTORIAL detallado (Capítulo 2) y comience a trabajar con el ejemplo. Usted puede verificar las funciones de FloraMap en la sección de REFERENCIA PARA EL USUARIO (Capítulo 3) o con las ayudas en línea. Si usted desea leer los antecedentes teóricos en la medida que vaya avanzando, diríjese, por supuesto, a la TEORIA del proceso (Capítulo 4). El ejemplo se extrae de una base de datos real que ha sido analizada en el CIAT. Hemos alterado algunas de las coordenadas de la accesión para simular el tipo de errores que usted puede encontrar en otras colecciones de datos.

Una vez que usted conoce lo que puede hacer FloraMap y lo que no puede hacer, usted estará listo para aplicarlo a sus propios datos. El Apéndice A le ayudará a crear sus archivos de insumo. **¡De allí en adelante, está por cuenta propia!** Pero no vacile en llamarnos si usted considera que FloraMap no está haciendo lo que debería hacer.

1. Ponerse en Marcha

Requerimientos Mínimos de Equipo y de Software

- **UCP 486 DS, 66 MHz** o más
Nota: Mientras se producía el manual, FloraMap estaba siendo ensayado en un Pentium G6 233 MHz. Su ejecución en equipos más lentos es factible, pero muchas veces penosamente lento.
- **32 Mb de RAM**
- **Drive de CD-ROM**
- **Al menos un espacio libre en la unidad de disco duro de 200 Mb**
*Nota: Esta aplicación de evaluación utilizará cerca de **64 Mb**, el resto se necesita para espacio de trabajo del mapa.*
- **Monitor de 15 pulgadas que corre 256 colores. 1024 x 768 pixeles.**
Nota: Se prefiere color de 16 ó 32 bit. Con menos de 256 colores, la gama de color para la superficie de probabilidad es limitada y puede aparecer como textura. Si su pantalla usa menos de 768 pixeles verticalmente, usted perderá partes de algunas ventanas.
- **Windows NT o Windows 95**
Nota: FloraMap no ha sido ensayada bajo Windows 3.1.
- **Impresora o graficadora a color – preferiblemente Postscript**

Instalación

FloraMap instalará c:\Program Files\CIAT\FloraMap en su disco duro C: en el directorio a menos que usted especifique una vía diferente durante la instalación. Puede que usted desee preparar un directorio alternativo ahora, antes de comenzar, o puede dejar que los procedimientos de instalación de FloraMap creen un directorio para usted.

- ✓ Coloque el CD-ROM de FloraMap en la unidad de CD.
- ✓ Cuando aparece el mensaje de petición de ejecutar (run), entre x:\set up, donde x es la letra del dispositivo de la unidad de CD-ROM.
- ✓ Conteste las preguntas que aparecen en el mensaje de petición en la medida que corre el programa de instalación.

El programa creará un directorio de trabajo c:\Program Files\CIAT\FloraMap\demo. Este directorio contiene dos archivos de puntos de acceso para el TUTORIAL en el Capítulo 2 de este manual. Usted puede mover estos archivos a otro directorio de su elección.

Los archivos de rejilla climática y de cobertura asociada permanecen en el CD-ROM para evitar llenar su disco duro con estos archivos grandes. Esto significa que usted tendrá que dejar el CD-ROM en la unidad para utilizar FloraMap.

Algunos usuarios pueden tener problemas en instalar FloraMap bajo Windows 2000. Esto es porque Windows 2000 se comporta un poco diferente a NT. Bajo NT cualquiera puede acceder una máquina y, si ellos no son un usuario registrado, Windows crea un perfil para ellos. Bajo Windows 2000 es un poco diferente. Se crea un perfil pero es uno virtual que desaparecerá cuando ellos salen del sistema. Esto afecta los privilegios que el usuario tiene. Si usted no puede cargar ni usar FloraMap significa probablemente que usted no es un usuario registrado en esa máquina. Haga lo siguiente:

- ✓ Entra al sistema como administrador
- ✓ Vaya a Comienzo–Configuración–Tablero de Control
- ✓ Haga clic en Usuarios y Contraseña de Acceso (entra)
- ✓ Haga clic en Agregar
- ✓ (Curiosear el nombre del usuario en la red)
- ✓ Haga clic en Siguiente. El nivel del acceso TIENE QUE ser “Standard User” [Power Use Group]
- ✓ Haga clic en Terminar
- ✓ Luego sale del sistema como administrador
- ✓ Accede el sistema como usuario
- ✓ Instale FloraMap

Únase al listado de servidor del grupo de usuarios de FloraMap para obtener información actualizada e indicaciones para usuarios.

Envíe un mensaje a: listserv@cgiar.org con el texto:
suscribir a FloraMap “su dirección de correo electrónico”.

NOTICIA IMPORTANTE

Si usted ya tiene instalado el “Borland Database Engine” puede obtener el siguiente mensaje de error. Si BDE no es presente, FloraMap pondrá el valor correctamente durante instalación.

“Error en acceso de rejillas de clima. Demasiados archivos abiertos, puede necesitar aumentar el límite de MAXFILEHANDLE en la configuración de IDAPI.”

En este caso, abra el Control Panel, abra el Administrador de BDE y seguir – SYSTEM y INIT a MAXFILEHANDLES. Aumente el número a acerca de 200.

2. Tutorial

Este tutorial busca familiarizar al usuario con el tipo de problemas que pueden presentarse al usar FloraMap con datos globales reales. Usted puede encontrar detalles sobre el manejo de las ventanas en el Capítulo 3 de este manual, pero también incluiremos recordatorios en esta sección. (Nadie lee los manuales hasta que es absolutamente necesario hacerlo, ¿o no?) Asimismo los detalles teóricos se encuentran en el Capítulo 4, Teoría.

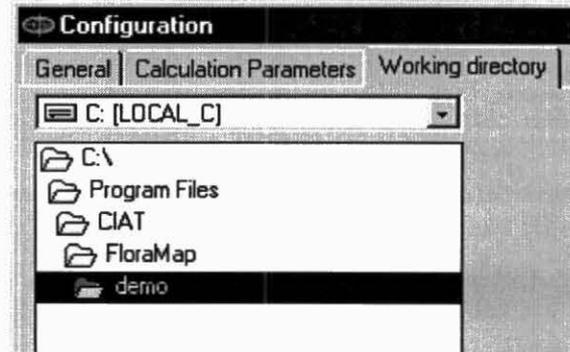
Trabajaremos, del principio al fin del tutorial, con un conjunto de datos que han sido preparados a partir de un conjunto real de accesiones de germoplasma de *Stylosanthes guianensis* de la colección de germoplasma del CIAT. *Stylosanthes guianensis* es una leguminosa potencialmente útil para uso en pasturas tropicales. Se encuentra ampliamente distribuido en América Latina y frecuentemente se desarrolla en zonas perturbadas por intervención humana. La especie tiene un número de formas diferentes, taxonómicamente diferenciadas. Adjunto a cada accesión, la base de datos tiene un identificador que denota la agrupación asignada en un análisis realizado en CIAT de un isoenzima, $\alpha\beta$ fosfatasa ácida. El identificador indica si la accesión provino del grupo 1 o del grupo 12 en ese análisis. Las accesiones del grupo 1 fueron predominantemente *S. guianensis* var. *vulgaris*, mientras que el grupo 12 incluyó muchas accesiones de *S. guianensis* var. *pauciflora* (para detalles adicionales, consulten a Jones et al. 1996).

En ciertos casos se han alterado los datos sobre las accesiones para permitirnos mostrar algunas de las características del sistema FloraMap y para ofrecer al usuario cierta práctica en buscar errores que probablemente pueden encontrarse en series equivalentes de datos de germoplasma. Se han cambiado todos los números de pasaporte de las accesiones para eliminar confusión con la base de datos original. En algunos casos, se han adicionado una serie de notas de campo imaginarias y datos de pasaporte para hacer que el ejercicio sea más verosímil.

Configuración del Mapa

Se ha establecido un directorio de trabajo para este tutorial. Despliegue hacia abajo marcos de trabajo (settings) y revise la ventana de configuración (configuration). El directorio de trabajo (working directory) debe verse como aparece a la derecha.

Este directorio contiene dos archivos. El archivo `stylo_guianensis.dbf` contiene nuestros archivos de puntos de acceso para el tutorial. El otro archivo, `stylo_secondfile.dbf`, se usará posteriormente en el tutorial para ahorrarle tiempo. Los archivos de puntos de acceso también pueden ser archivos ASCII delimitados por espacios (ver el APENDICE A para preparar los archivos de puntos de acceso).



Haga clic en el icono atajo para FloraMap y podemos comenzar. Aparecerá un mapa en blanco con un menú de capas en la esquina derecha superior. Como primera medida, asignaremos la configuración que utilizaremos para el tutorial.

Haga clic en **settings** y en **configuration**.

- ✓ En la ventana de **configuration**, haga clic en **working directory**.
- ✓ Verifique que el directorio de trabajo es el correcto para este tutorial. Generalmente debe ser `c:\Program Files\CIAT\FloraMap\demo\`.
- ✓ Haga clic en **general**.
- ✓ Defina las opciones de autoguardar configuración (**autosave configuration**), autoguardar mapa (**autosave map**) y agregue los símbolos de la capa a la leyenda (**add layer symbols to legend**).
- ✓ Verifique que el mar es azul en el color de fondo del mapa (**map background color**).
- ✓ Escoge rejillas climáticas incorporadas (**built-in climate grids**) y seleccione América Latina (**Latin America**).
- ✓ Haga clic en **parámetros de cálculo (calculation parameters)**.
- ✓ Active la opción **muestre el clima promedio para los puntos seleccionados (show average climate for selected points)**.

Aparecerá en pantalla el diagrama del clima para una serie seleccionada de puntos en cualquier momento durante el análisis.

- ✓ Fije **corregir temperatura (correct temperature)** en **apagado**.

La corrección para la elevación de los puntos de acceso permite estimar el clima con mayor precisión. Vamos a comenzar el tutorial como si no estuvieran disponibles datos de elevación.

- ✓ Active trate a las accesiones con coordenadas idénticas como una sola observación (*treat accessions with identical coordinates as a single observation*).

Esta opción cambia la forma en que se interpreta la base de datos de accesiones. La forma en que usted desea que se interprete depende de su comprensión de lo que significa accesiones múltiples para el análisis. Las accesiones múltiples que provienen del mismo par de coordenadas pueden surgir de diversas maneras. Debemos aclarar que aquí estamos discutiendo puntos múltiples que provienen exactamente de las mismas coordenadas, no solo del mismo pixel climático. Esto significa que las coordenadas se entraron como idénticas **a propósito**. O, el recolector tomó varias accesiones de lo que él o ella consideraba como el mismo sitio (la misma parada en el camino o en el transecto) o en el mismo campo o en la misma finca; o la muestra recolectada se ha subdividido durante su procesamiento en el banco de germoplasma. Esto a menudo sucede cuando se siembra la muestra para determinar sus características fenotípicas o agronómicas, o ambas cosas. En el banco de germoplasma del CIAT, a menudo se puede identificar estas últimas porque el número de accesoión es seguido de una serie de letras (a, b, c, etc.) para indicar la subdivisión de la muestra.

¿Significa esto que una muestra/ un sitio es más importante porque allí se recolecta un rango de diversidad genética o se debe considerar la muestra/ sitio como un punto climático? Cualquiera de estos puntos de vista es completamente válido y usted, el analista, debe decidir cuál usar. En este análisis tutorial, hemos decidido desactivar esta opción, lo que significa que muestreo /sitio con entradas múltiples tendrá una ponderación más alta que muestreo /sitio única.

- ✓ Ajuste la opción desajustes (*mismatches*) en movido manualmente (*moved manually*).

Esta opción es muy importante. La alternativa es mover automáticamente los puntos desiguales al punto más cercano en la superficie climática. Más adelante en el tutorial, usted verá que esto puede producir graves errores si no se utilice correctamente.

- ✓ Haga clic en OK para aceptar las configuraciones y estamos listos para comenzar.

La primera cosa que debemos hacer es instalar el mapa. Vaya al

menú de capas (layers menu) y haga clic en el icono cargue capas. Aparecerá en pantalla el menú abrir capas (open layers). Elija los archivos de tipo shapefile. Usted ahora tiene que encontrar el shapefile correcto para proporcionar los antecedentes al análisis. Hemos colocado una selección de shapefiles en el CD-ROM, y se encuentran bajo \COVERAGES\america.

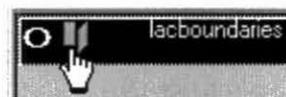


SAMCOUNTRIES.shp es el shapefile que nos dará los antecedentes a América Latina con los límites de los países. También está SAMMUNICIP.shp, que consiste en los límites municipales (condado) y los caminos, los ríos y las aldeas, que se explica por sí mismo. Las cuatro coberturas se explican en mucho detalle y no deben desplegarse en el mapa de todo el continente. Son, sin embargo, útiles cuando el mapa está aumentado para mirar en detalle una región.



Aun cuando usted ha ampliado una región pequeña, estas coberturas tomarán tiempo para usarlas. La ventana completa de cobertura se calcula aunque no muestre en la pantalla. Esperamos solucionar este problema en versiones posteriores de FloraMap.

Resalte a SAMCOUNTRIES.shp y haga clic en Abrir (Open). El mapa de América Latina y el Caribe aparecerá en la ventana del mapa. Si llega negro, usted deberá cambiar las configuraciones de color y rellenar (fill) para generar una base útil para su mapa.



- ✓ Haga clic en control de capa (layer control). Esto abrirá el menú de control de capa.

- ✓ Cambie el color de fill a un buen color de fondo. Un verde opaco es muy efectivo.
- ✓ Seleccione relleno sólido (solid fill).
- ✓ Active muestre en la leyenda (show in legend), usted no necesitará el fondo nombrado.
- ✓ Active con contorno (with outline).
- ✓ Coloque el color del contorno (outline color) en un color oscuro, el negro generalmente sirve.
- ✓ Haga clic en aplicar (apply) para verificar lo que usted ha hecho. Si le parece adecuado, haga clic en OK para fijar los cambios.

Usted ahora está listo para cargar el conjunto de datos de las accesiones (accessions dataset).

- ✓ Haga clic en el icono cargar capa, al igual que antes.
- ✓ Elija puntos de accesión (accession points) como tipo de archivo.



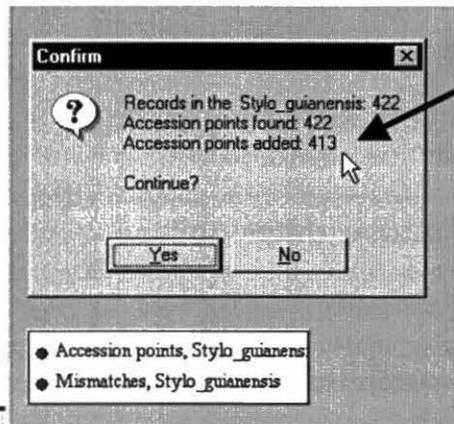
NOTA: Dado que la extensión de archivos .dbf se usa también para los shapefiles, estos también aparecerán como puntos de accesión. Lamentablemente, debido a sus contenidos geográficos, FloraMap puede aceptarlos como archivos de puntos de accesión válidos. Tenga cuidado de abrirlos como archivos de puntos de accesión porque son, a menudo, muy grandes y tomarán una cantidad desmesurada de tiempo para cargar. Estamos trabajando para que este aspecto sea más seguro.

- ✓ Cambie nuevamente el directorio al directorio de trabajo c:\Program Files\CIAT\FloraMap\demo.
- ✓ Seleccione el archivo stylo_guianensis.dbf y haga clic en open.

FloraMap abrirá el archivo y procederá a verificar el contenido. FloraMap hace esta verificación asegurando que cada punto en el archivo pueda asignarse a un pixel de clima válido de la base de datos climáticos. Esta operación toma mucho tiempo ya que se debe buscar en toda la base de datos para determinar la validez de los puntos. Las palabras verificando accesiones (checking accessions) aparecerán en la parte inferior de la ventana del mapa (map window) en el indicador del proceso (process indicator) y aparecerá una escala de tiempo. Debido a que esta operación es una actividad de alto nivel, la ventana de capas (layers window) no desaparecerá de la pantalla si usted trata de cambiar a otra aplicación mientras la verificación de las accesiones está en proceso. Para evitar que esto interfiera con su otro trabajo, mueve layers window a una parte discreta de la pantalla antes de abrir el archivo de puntos de accesión (accession points file).

Verificación de los Datos

Se ha terminado la verificación de los puntos de acceso y usted puede ver la siguiente ventana.



No se encontraron todos los puntos de acceso durante el proceso de verificación. Afortunadamente, usted seleccionó el movimiento manual de los puntos de acceso faltantes. Ahora usted debe revisar el mapa y determinar dónde están. Son $422 - 413 = 9$ puntos que deben encontrarse.

✓ Haga clic en no en la ventana confirmar (confirm). Libere las ventanas confirm y capas (layers) para liberar espacio

para manipular el mapa y vaya en busca de los puntos faltantes. La leyenda del mapa indica que ellos se visualizarán en rojo.

Aquí usted ha encontrado un grupo de accesiones que está completamente solo en el Océano Atlántico.

✓ Use zoom y pan hasta que usted pueda ver las accesiones claramente.



✓ Haga un círculo alrededor de los puntos con la herramienta seleccionar área. Asegúrese de que el conjunto de datos de desiertos está resaltado en la ventana de control de capas (layers control window).



✓ O señale las accesiones con el cursor hasta que aparezca la información.

Este es un caso típico de mala ubicación. Es casi siempre corregible por la más leve de las pistas. En este caso, mire la identidad del recolector y la fecha en que se recolectaron las accesiones.

W.B.S. fue quien hizo la recolección y ella recolectó estas accesiones el 12 de agosto de 1978. Usted puede estar muy seguro de que ella no se encontraba en el Atlántico en esa fecha, ¿luego dónde estaba? Usted tiene varias opciones.

- ✓ Examine el conjunto de accesiones en dBase, Excel u otra hoja electrónica en busca de la fecha y el recolector.
- ✓ Devuélvase a la base de datos de germoplasma de la cual se tomo el conjunto de accesiones y mire si hay otras pistas.
- ✓ Verifique si hay un catálogo publicado de la colección.
- ✓ Verifique con el libro de campo original de recolección si usted lo tiene a mano.

En este caso, tenemos acceso a la información en dos formas. La lista del conjunto de accesiones nos muestra un grupo de accesiones recolectado en Brasil por W.B.S., más o menos en esa fecha.

	A	B	C	D	E	F
1	LONGITUDE	LATITUDE	ISOCCLASS	ACCESSION_	COLL_N_DAT	COLLECTOR
10	-74.3170	-15.0500	1.0000	Sgui2624	14-Aug-78	W.B.S
11	-45.0830	-12.4830	1.0000	Sgui1931	10-Aug-78	W.B.S
12	-41.8670	-12.4000	1.0000	Sgui415	12-Aug-78	W.B.S
13	-41.8670	-12.4000	1.0000	Sgui95	12-Aug-78	W.B.S
14	-38.7170	-12.4000	1.0000	Sgui731	12-Aug-78	W.B.S
15	-38.3500	-12.3830	1.0000	Sgui1006	12-Aug-78	W.B.S
16	-38.6670	-12.3330	1.0000	Sgui1367	12-Aug-78	W.B.S
17	-44.8830	-12.0830	1.0000	Sgui2027	14-Aug-78	W.B.S
18	-44.6670	-12.0170	1.0000	Sgui3292	14-Aug-78	W.B.S
19	-44.6670	-12.0170	1.0000	Sgui2970	14-Aug-78	W.B.S

Ahora usted sabe que W.B.S. estuvo en Brasil en 12.4 Sur, 41.867 Oeste en el 12 de agosto de 1978. Por consiguiente, es muy probable que las accesiones provinieron de este sitio. El segundo fragmento de información que confirma lo anterior proviene de las anotaciones de campo de los datos de pasaporte en la base de datos de germoplasma. Dicen así:

12 de agosto, entre Itaberaba y Seabra, valle de Paraguacu, estribaciones de Serra do Sincora. La evidencia, por consiguiente, apunta hacia el caso del signo faltante. En este caso, en la latitud. Dado que las coordenadas corregidas ubican a las accesiones “desacertadas” directamente en el camino de Itaberaba a Seabra, podemos suponer que este fue el único error (lamentablemente común) en los datos.

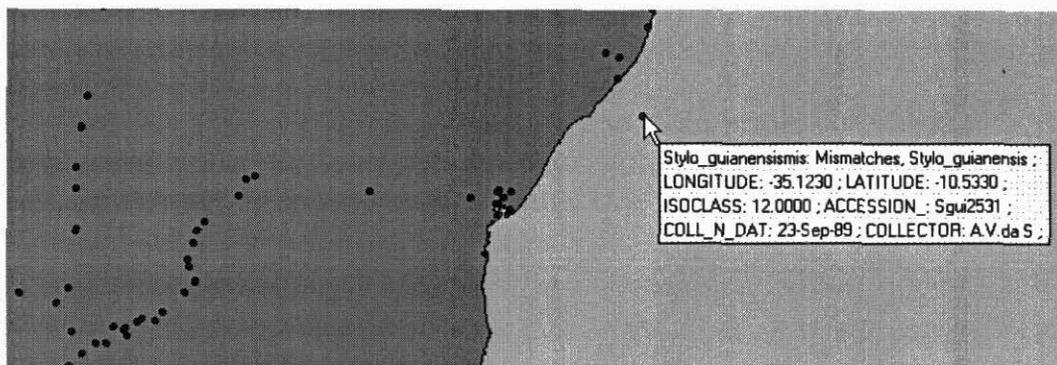
- ✓ En su base de datos o programa de hoja electrónica, corrija las latitudes correspondientes a estos puntos en la base de datos de puntos de accesión (accesión points dataset).



NOTA: Recuerde cerrar el cuadro de accesiones después de efectuar los cambios. FloraMap no podrá accederlo si usted deja la hoja electrónica en forma de ventana minimizada.

Hay otro punto único mar dentro de la costa de Brasil.

- ✓ Elija la herramienta zoom y logre un acercamiento a la costa este de Brasil.
- ✓ Usando el icono de pan, mueve el mapa hacia arriba y hacia la izquierda para ubicar la zona en un espacio de trabajo fácil.
- ✓ Señale el punto rojo de desierto hasta que aparezca la información. O haga un círculo alrededor de éste con la herramienta seleccione área.



Este vez usted tiene una accesión recolectada por el famoso (y novelesco) botánico brasileño A. V. da Silva, según lo indicado por la información sobre el punto de accesión. Usted debe determinar adónde estaba el Dr. da Silva el 23 de septiembre de 1989. Al revisar la lista de accesiones, usted encuentra, afortunadamente, que en septiembre de 1989 él se encontraba haciendo recolecciones en Mato Grosso do Norte donde los *cerrados* se unen con la selva amazónica.

LONGITUDE	LATITUDE	ISOCLASS	ACCESSION_	COLL_N_DAT	COLLECTOR
-52.1	-12.05	12	Sgui2529	26-Sep-89	A.V.da S
-55.083	-10.583	12	Sgui1561	23-Sep-89	A.V.da S

La base de datos de germoplasma no contiene más información sobre estas accesiones. Afortunadamente, el Dr. da Silva publicó el diario de sus viajes de recolección desde 1956 hasta 1991, año en que se retiró. En estos diarios, usted puede leer las siguientes descripciones:

Setem 26 Varzias do rio Xingu, Floresta de varzia muito denso, poucos leguminos baixos. Uma S. gui.

Setem 23 Noreste de Telez Pirez. Floresta de cerradao denso. Mata de varzia em baixos, dois S.gui em clareiras dos acampamientos.

Caminho do varios kilometros.

Esto pone la accesión en la zona correcta, al nordeste de Telez Pirez.

¿Pero cuáles son las coordenadas correctas? Ni la latitud ni la longitud coinciden con el otro sitio de accesión. Sin embargo, observamos que el recolector se estaba trasladando entre los lugares de campamento (¿mineros quizás?) de la misma localidad, así que el otro sitio no debía quedar muy lejos. ¿Podría un sencillo error de entrada de datos cambiar estas coordenadas a un lugar ubicado unos cuantos kilómetros del otro? Sí, en este caso un error común de digitación ha reemplazado un 5 con un 3. En efecto, la ubicación 55.083 Oeste, 10.583 Sur es exactamente 5 km 460 metros de 55.123 Oeste, 10.533 Sur. 55.123 Oeste se había convertido en 35.123 debido a un error de digitación.

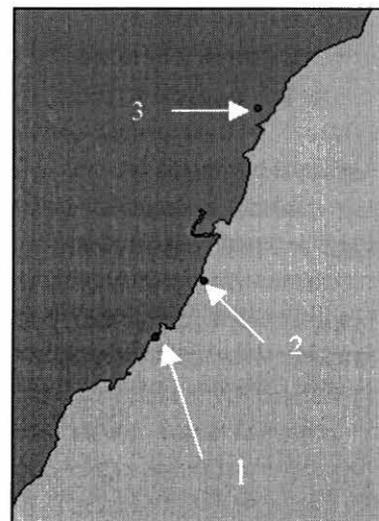
- ✓ Si usted está de acuerdo que estas pruebas son convincentes, prosiga y corrija el punto en la lista de accesiones (accessions list).

El próximo problema es uno para el cual FloraMap está diseñado para afrontar. Cuando se computan las superficies de la base de datos climáticos, dependemos de un modelo de elevación digital (MED) del área. El modelo es una representación de la topografía que es generada por computadora, en este caso una sencilla rejilla cuadrada que da las alturas modales de la tierra en cada celda de la rejilla. Esto depende, a su vez, de los mapas a diversas escalas y proyecciones.

El fondo y las capas geográficas clave incorporadas en FloraMap también provienen de mapas en diversas proyecciones. Los mapas usados por los botánicos y los recolectores de plantas tienen las mismas características.

Todo lo anterior hace que sea relativamente probable que, a las escalas continentales a las cuales estamos trabajando, los bordes de las coberturas no coincidan exactamente, lo cual generalmente ocasiona que algunos puntos caigan fuera del área de cobertura climática.

Observe, ahora, la línea costera al norte de Río de Janeiro. El punto 1 parece estar tan cerca de la costa que casi podría estar en la playa. (¡Las preguntas sobre por qué los recolectores de plantas están descansando en las playas brasileñas deben dirigirse a las autoridades pertinentes y no a los autores de FloraMap!) El punto 2 también parece estar ubicado cerca de la orilla y un pequeño toque lo pondría en un píxel válido en la base de datos climáticos. El punto 3 parece ser cuestión de adivinanza. Está en tierra, pero sigue siendo una accesión desacertada. En este punto, usted debe recordar que los píxeles de la base de



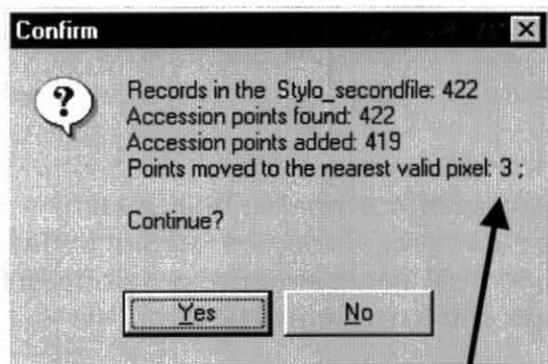
datos climáticos tienen cerca de 18 km en cada lado. El punto 3 está en la esquina entre dos píxeles que se están aproximando a la costa en este punto. Generalmente es buena idea revisar las coordenadas de los puntos que se encuentran en una situación similar a la de estos tres, en caso de que usted observe errores obvios. Sin embargo, todo lo que usted tiene que hacer después es comenzar nuevamente el análisis después de cambiar la configuración.

- ✓ Despeje el mapa.
- ✓ Haga clic en settings.
- ✓ Seleccione calculation parameters.
- ✓ Bajo mismatches, elija movido automáticamente (moved automatically).

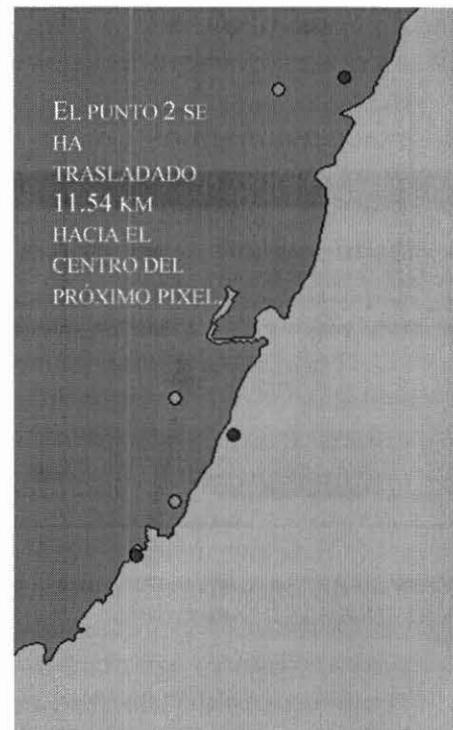
NOTA: No lo hace hasta que haya terminado de verificar todos los puntos de desierto. ¡Si usted deja puntos sin verificar, FloraMap los ubicará, muy gentilmente, en el punto más cercano en la línea costera, independientemente de si es prudente hacerlo!

En este caso lo ha hecho, así que haga clic en OK y proceda con el tutorial.

Ahora comience nuevamente el programa.



Tres píxeles ya no son desiertos. Se han trasladado al píxel válido más cercano.



Estos puntos han sido adicionados al final de un nuevo archivo de 422 puntos de acceso que FloraMap usará de ahora en adelante. Si usted desea, puede suprimir el conjunto de puntos de desierto del mapa. No se tomarán en cuenta estos puntos en cualquier análisis subsiguiente.

Puede habersele ocurrido que si usted puede detectar errores obvios, por ejemplo los puntos de desacierto, puede haber errores que no han traducido los puntos de la capa de trabajo del clima, sino que solamente los mueven a otro pixel climático válido, pero erróneo. Este, lamentablemente, es generalmente el caso.

Un control obvio es verificar que los puntos se encuentren al menos en los países y/o las regiones donde usted sabe que las accesiones se recolectaron. En el caso de estas colecciones de *S. guianensis*, podríamos decir, sin lugar a dudas, que cualquier punto que se encuentre en Argentina o Chile está fuera de límite. Pero esta observación dependería del conocimiento de que ninguno de los recolectores visitó esos países, pues no necesariamente es imposible encontrar estas plantas allí.

Otro control, si las accesiones tienen la fecha correcta, consiste en hacer efectivamente un seguimiento al viaje de recolección en el mapa. La cobertura de caminos de FloraMap, que usted encontrará en el CD-ROM, puede ayudarlo en esta tarea. Los cambios abruptos o desvíos en el itinerario deben investigarse. La herramienta de distancia (p. 51) puede indicarle rápidamente si es factible haber hecho un desvío en función de tiempo.

Cuando se hayan agotado todos estos métodos obvios, FloraMap aún tiene algunas herramientas que lo pueden ayudar. Continúe con el análisis y usted tendrá la oportunidad de aplicarlas.

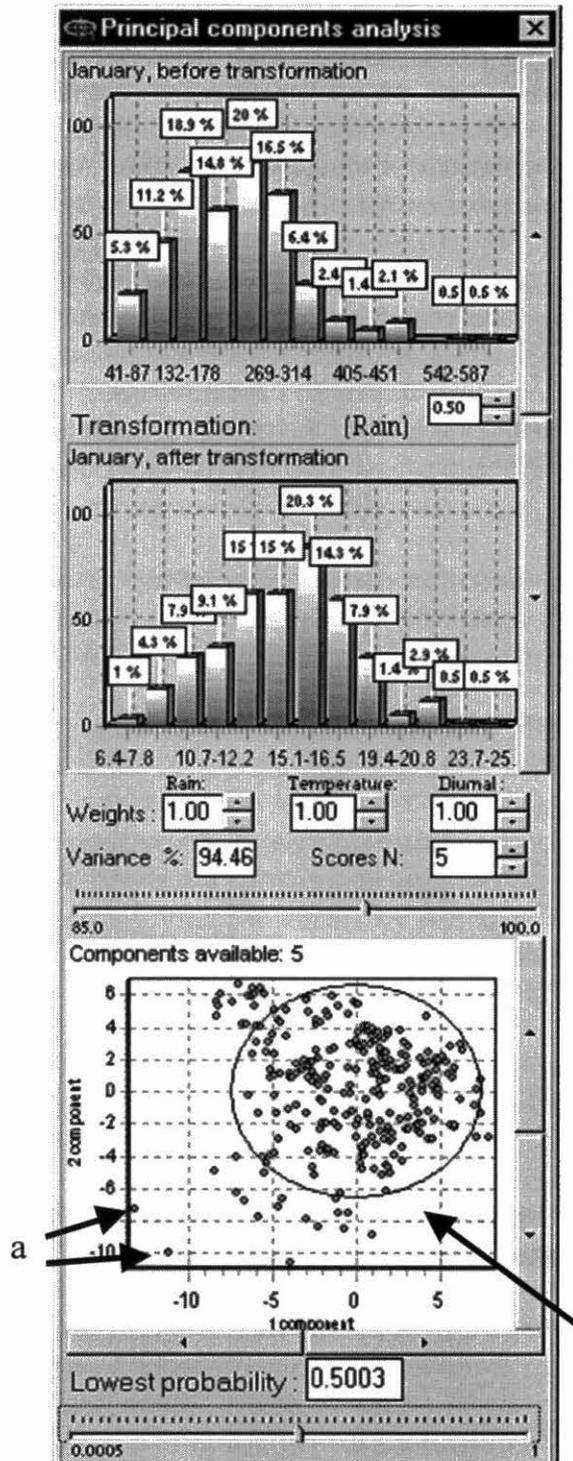
- ✓ Haga clic en sí (yes) en el menú de confirmación (confirm menu) a la izquierda.

El mensaje *creating climate file (creando el archivo de clima)* aparecerá en el *process indicator*, junto con su barra de escala de tiempo. Una vez creado el archivo, aparecerá la ventana PCA. Esta ventana es muy compleja. Por favor tome el tiempo para leer la descripción que se hace de esta ventana en el capítulo 3.

Todavía estamos, ante todo, interesados en verificar el conjunto de datos. En un análisis real, esto puede tomar días o hasta semanas para hacer. Las personas tienden a pensar que los datos son mejores porque han sido apuntados y que una vez dentro de una computadora son 'verdaderos'. Los datos de pasaporte de germoplasma pueden haber pasado por una docena o más de etapas durante su procesamiento. Las coordenadas van del mapa al cuaderno, del cuaderno al informe. A menudo el informe pasa por las manos de un editor. Esto implica edición y composición. Los datos luego se codifican para su insumo en la base de datos, y, debido a que las accesiones se pasan de colección en colección, los datos codificados se transfieren de un sistema de computación a otro.

Es un milagro que logremos siquiera algo con que valga la pena trabajar, pero debido a que todas las personas que participan del proceso son sumamente dedicadas, generalmente sí lo hacemos. Sin embargo, los errores son inevitables.

Avancemos directamente al próximo nivel de la verificación de datos.



Observe la ventana de PCA a la izquierda.

- ✓ Ajuste ponderaciones (weights) en 1.00.
- ✓ Ajuste transformación (transformation) en raíz cuadrada-lluvia (rain) a la potencia de 0.5.
- ✓ Ajuste puntuación N (scores N) en 5.

Usted debe tener el análisis como aparece en la imagen. De no ser así, asegúrese de que las correcciones que hizo en la base de datos de accesiones fueron debidamente aceptadas.

Observe el diagrama de dispersión del PCA en la parte inferior de la ventana de PCA. Este diagrama muestra la representación de las accesiones en el espacio bidimensional definido por el primer y el segundo componente principal. Si el suyo no lo hace, mueva las barras laterales hasta que obtenga el diagrama apropiado en pantalla. Este plano de componente-1-por-componente-2 es casi siempre el plano en el espacio del PCA que muestra la mayor parte de la variación en los datos. De hecho, en este análisis representa cerca del 70% de la variación.

La elipse muestra el límite en 2 desviaciones estándares. En dos dimensiones esperamos que cerca del 14% de los puntos caigan fuera de

esta elipse. Con 422 puntos de acceso, tendríamos, entonces, 49, de manera que la población no es seriamente anormal. Sin embargo, algunos puntos deben ser estudiados.

Los dos puntos en la esquina inferior izquierda (a) obviamente son resultados aislados. Verifiquen dónde están en el mapa utilizando la herramienta **seleccionar área** para dibujar alrededor de ellos. Las accesiones deben parecerse a las siguientes dos: Sgui2624 y Sgui2163.

¡Ahora espere un minuto! Ya hemos visto que W.B.S. estaba recolectando accesiones en Brasil el 14 de agosto de 1978, entonces ¿cómo podría estar también en Perú? Lamentablemente, en este caso no tenemos acceso a sus libros de campo y los datos de pasaporte no contienen notas útiles. Sabemos aproximadamente dónde se recolectaron las otras accesiones ese día, pero las coordenadas mal registradas no aportan una pista útil. No hay ningún error de digitación obvio. Alguien debe haber saltado un renglón o revuelto las hojas cuando se registraron estos datos. Este caso nos ha vencido. Al menos sabemos que está equivocado, pero nuestro único recurso consiste en eliminar la accesión.

- ✓ Sálgase de FloraMap, vaya a la hoja electrónica de simulación y suprima la hilera para la accesión Sgui2624.

Y entonces hay 421.

LONGITUDE	LATITUDE	ISOCCLASS	ACCESSION_	COLL_N_DAT	COLLECTOR
-74.317	-15.05	1	Sgui2624	14-Aug-78	W.B.S
-71.033	8.683	12	Sgui2163		

Observe el siguiente punto, Sgui2163. Se encuentra en Venezuela. No hay ninguna nota, ningún libro de campo, ninguna fecha propicia ni recolector. (Esto no es realmente el caso pues se trata, en la vida real, de una accesión recolectada por el CIAT y conoceríamos su procedencia exacta si tuviera su número de accesión correcta. En efecto, este es el primer caso de un problema que no se ha inventado, pues lo hemos descubierto en el transcurso de este ejercicio.)

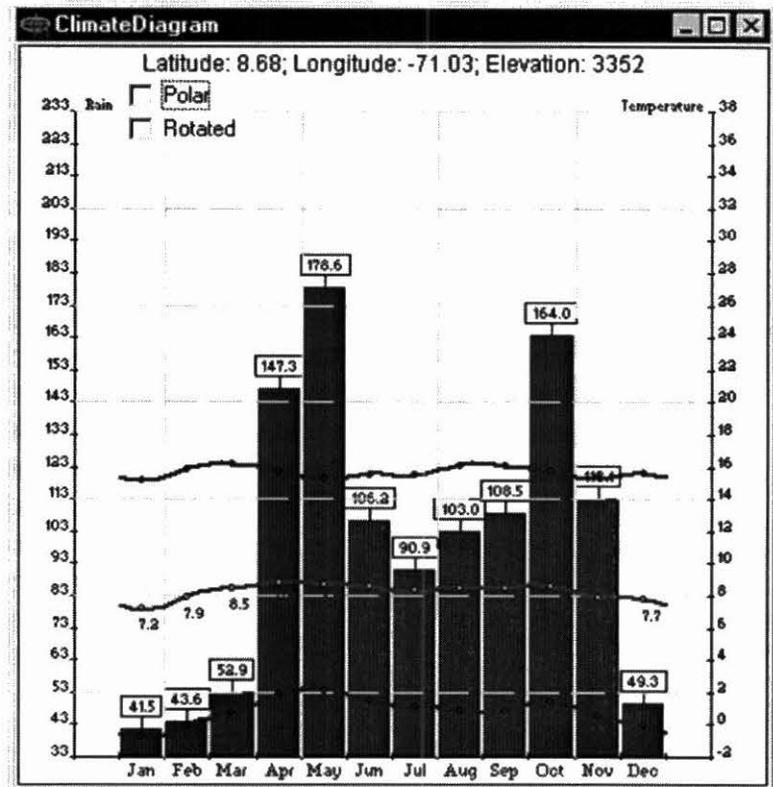
Entonces, ¿qué hace que el clima de esta accesión sea tan diferente? Observe el registro de clima que se generó para esta accesión a partir de la base de datos.

- ✓ Dibuje alrededor del punto en el diagrama de dispersión con la herramienta **seleccionar área**. Usted tiene activado la opción **calcular promedio (calculate average)**, entonces un diagrama del clima aparecerá. Ignórelo por el momento. Observe que el punto en

Venezuela empieza a titilar en el mapa. ¡Usted ha identificado dónde está!

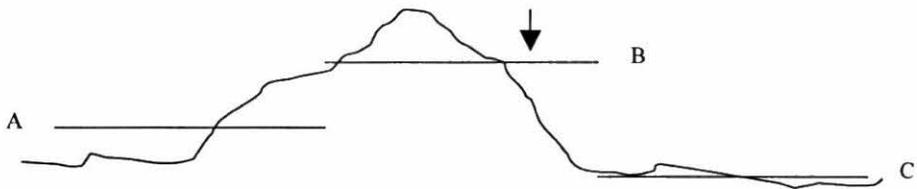
- ✓ Escoja zoom para agrandar bien la zona alrededor del punto. Usted puede ampliarla hasta que no se observe ningún otro punto.
- ✓ Utilice la herramienta para elaborar el diagrama del clima para señalar la acesión. El diagrama del clima aparecerá como al derecho.

Sería una planta de *S. guianensis* muy extraña la que creciera en un ambiente con temperaturas diarias promedio de cerca de 8 grados centígrados y con temperaturas nocturnas por debajo de cero durante al menos dos meses al año. ¿Qué ha sucedido? Observe la elevación del pixel en la rejilla climática. Es de 3352 metros, o sea está bien arriba en la cordillera.



Este problema se presenta debido al tamaño del píxel, 18 km, utilizado por la base de datos climáticos. Sin embargo, no se resolverá completamente con una rejilla climática más estrecha. ¿Recuerde los problemas de los desaciertos que se presentaron a lo largo de la costa? Exactamente lo mismo puede suceder, por todas las mismas razones, a lo largo de una discontinuidad topográfica —que podría ser una hilera de cerros, una cordillera andina o el contorno del valle Rift en Kenya. El clima es generalmente insensible a unos cientos de metros, o hasta unos cuantos kilómetros, de desplazamiento lateral. Los leves desaciertos de las capas del mapa son generalmente poco importantes para la precipitación. Pero, si el desacierto cambia la elevación, entonces la temperatura puede ser bastante diferente a lo esperado.

Cada uno de estos píxeles — A, B y C — representa la elevación modal del terreno que cubre. En la base de datos

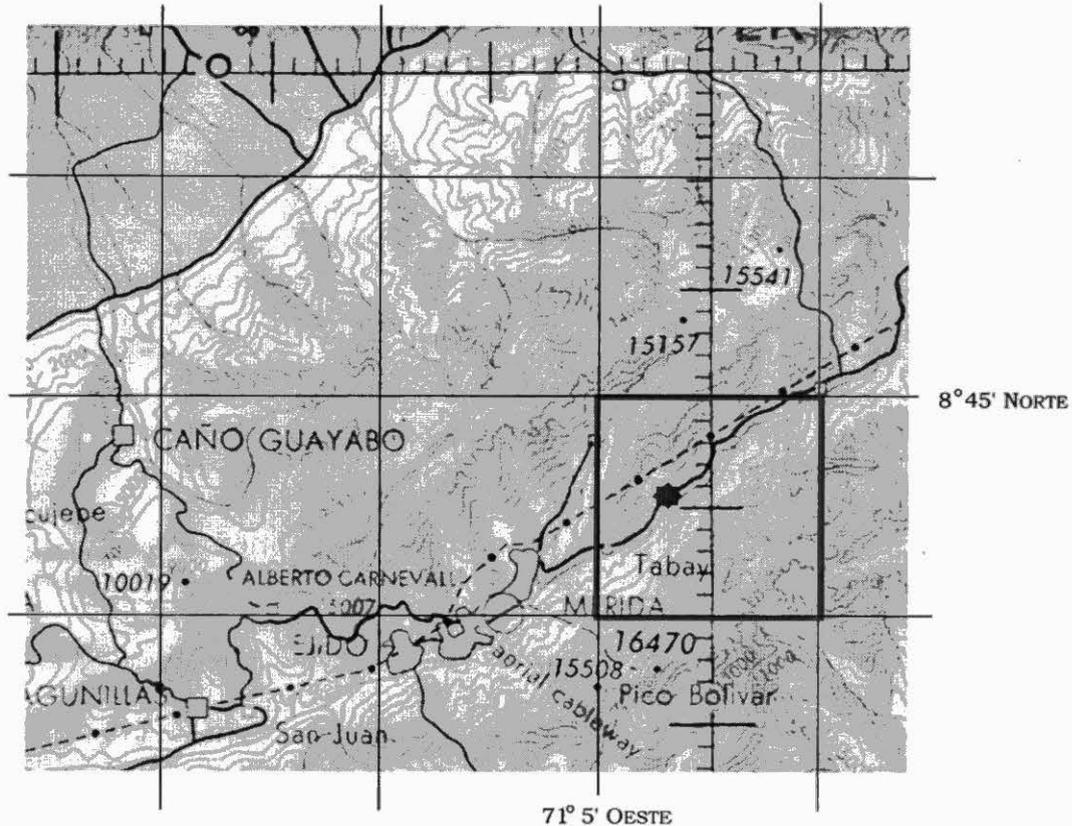


que estamos utilizando, los píxeles tienen una longitud de 18 km. A esta escala no es sorprendente encontrar un cerro importante que corte los píxeles A y B. Si un punto de acceso es representado por la flecha en el píxel B, ¿cuál es la solución? Podríamos moverlo al píxel C, tal como lo hubiéramos hecho de haber faltado una línea costera. Esta es una solución viable si usted conoce el terreno y sabe que es válido mover el punto de acceso al píxel C. El problema sería resuelto con un pequeño cambio de coordenadas.

Sin embargo, supongamos que el recolector realmente recogió el espécimen en el punto en el píxel B y no se ha presentado ningún desplazamiento. La elevación en ese punto está, entonces, bien representada por la flecha en el píxel C, pero no queremos desplazar el punto. Podemos corregir el registro del clima correspondiente a la elevación en el punto de recolección. Hacer la corrección de las tasas de lapso para temperatura es sencillo y fiable. Las temperaturas descienden aproximadamente 6 grados cada 1000 metros de elevación (FloraMap usa un modelo más preciso pero la idea es la misma).

Si FloraMap va hacer esta corrección, necesita saber la elevación real en el punto de recolección. Hemos diseñado a FloraMap para que sea flexible en este aspecto porque (especialmente en lo que se refiere a las colecciones viejas) no siempre tenemos la elevación de origen según lo midió el recolector. Si no está disponible ningún dato de elevación, usted comenzará el análisis con un conjunto de datos de accesiones que no tendrá una columna para elevación. FloraMap aceptará esta situación y completará las elevaciones a partir de la base de datos climáticos. De hecho, puesto que las temperaturas en la base de datos son ajustadas a la elevación del píxel, no hay nada que hacer. Sin embargo, si usted agrega una columna denominada elevación (elevation), FloraMap tomará esta columna y, dondequiera que se presente un campo válido que no está en blanco, sustituirá esa elevación con la que se encuentre en la base de datos y automáticamente corregirá la temperatura con el modelo de la tasa de lapso.

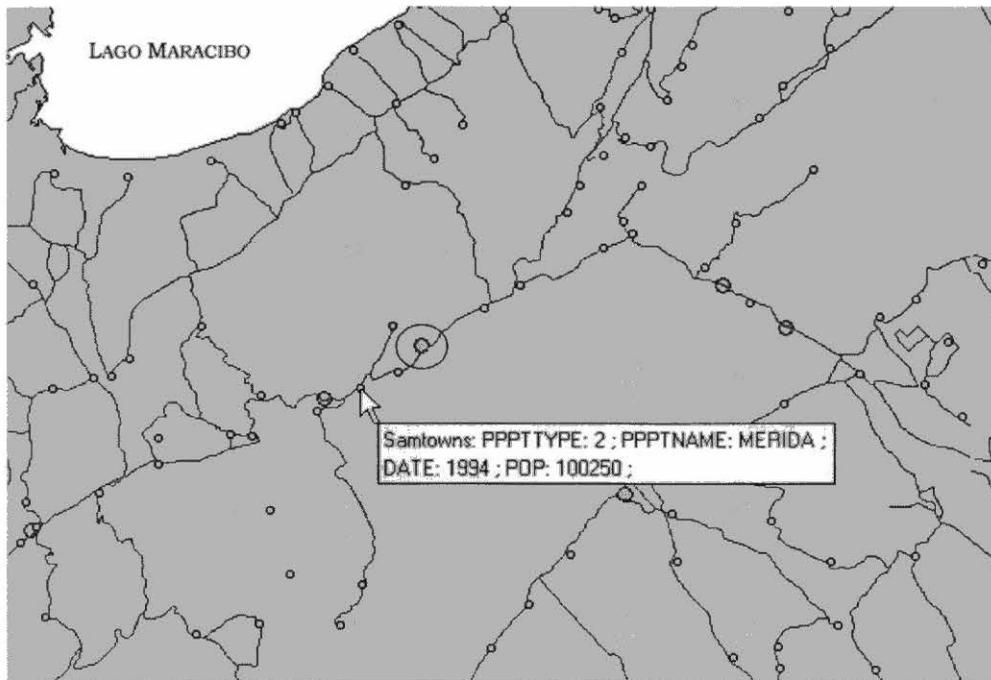
A continuación presentamos una parte del Mapa de Navegación Operacional (ONC, su acrónimo en inglés) K 26, a escala 1:1 000 000, para la región alrededor de Mérida en Venezuela.



Estas graficas están disponibles para todo el mundo. Publicadas por el Defense Mapping Agency Aerospace Center, localizado en St Louis AFS Missouri 63118, Estados Unidos, son razonablemente fiables y no son costosas. Los pilotos de avión las usan para trazar los cursos de vuelo, de manera que tienen que ser lo suficientemente económicas para ser desechables. Constituyen una ayuda imprescindible para el trabajo geográfico investigativo que usted debe realizar para validar sus series de datos de accesiones.

La rejilla de pixel climática se muestra como la red geográfica en negro. El píxel escogido para la accesión Sgui2163 aparece demarcado en púrpura. La coordenada del píxel indica su esquina noroccidental y el píxel cubre una extensión de 10 minutos como lo muestra la red geográfica del mapa. La estrella roja indica que la planta se recolectó (como lo son muchas accesiones de ciertas especies) al lado del camino. Este es un control útil respecto a la validez de las coordenadas. Usted

puede ver por qué la elevación del píxel es superior a 3300 metros. Toda la zona color amarillo oscuro está por encima de los 9000 pies (lamentablemente las gráficas del ONC no están en el sistema métrico), dominando el paisaje en el píxel con áreas relativamente grandes que sobrepasan los 12 000 pies. Por consiguiente, el píxel es una generalización fiel del paisaje, pero Sgui2163 provino del fondo del valle, cerca al río, a poco menos de 7000 pies. Esto se traduce en cerca de 2100 metros, justo dentro del límite para esta especie.



Hay una manera alternativa de visualizar el área desde FloraMap. Si usted quiere intentarlo, siga los pasos a continuación:

- ✓ Haga zoom cerca en la zona indicada.
- ✓ Seleccione **samroads** en el directorio de coberturas (**coverages**) en el CD-ROM.
- ✓ Abra el icono de control de capas para **samroads** en la ventana de **layer control**.
- ✓ Fije el tamaño en 1 y el color en rojo.
- ✓ Seleccione **samtowns** del directorio de **coverages**.
- ✓ Fije el tamaño en 3 y el color en naranja.
- ✓ Cambia el tamaño de los puntos de acceso a 6 y el color a azul.

Ahora debemos hacer la corrección respecto a la elevación.

- ✓ Vaya a su hoja electrónica.
- ✓ Abre el archivo **S-guianensis.dbf** e inserte una columna para elevación.

- ✓ Usted ahora puede introducir 2100 para Sgui2163.

En este momento, usted puede ejecutar nuevamente el análisis y observará que el punto para esta accesión desaparecería de los valores alejados de lo normal en el diagrama de dispersión y se uniría a los otros que se encuentran dentro de la elipse. Sin embargo, algunas accesiones provienen de situaciones andinas similares. Valdría la pena revisar estos otros puntos por el mismo tipo de problema.

- ✓ Utilice la herramienta de diagrama del clima para señalar un candidato factible. Sabemos que los Andes son elevados cerca de la frontera de Colombia con Ecuador. Pruebe los puntos ubicados en 77.5 grados oeste, 0.9 grados norte. Resulta ser la accesión Sgui1285.

El diagrama del clima indica que la elevación es de 2743– claramente no una elevación en la cual se encontraría *S. guianensis*. Sería mejor, entonces, proporcionar a FloraMap la elevación de todos los puntos de accesión para que no tengamos que dar caza a todos los posibles errores, uno por uno.

Para ahorrarle tiempo y esfuerzo (¡aunque no existe atajo cuando se trata de datos reales!) hemos preparado una nueva lista de accesiones para usted, la cual incorpora los datos de elevación para todas las accesiones andinas. Recomience el análisis con el archivo `stylo_second` y usted tendrá los cálculos de elevación de los recolectores, además de todas las correcciones que usted haya incorporado hasta el momento.

Análisis de Componentes Principales

- ✓ Borre del mapa todas las capas de puntos de accesiones (`accession points layers`), o reinicie FloraMap.
- ✓ Active `configuration` del menú de `settings`.
- ✓ Seleccione `calculation parameters`.
- ✓ Seleccione `correct temperatures`.
- ✓ Bajo `mismatches`, seleccione `moved automatically`.



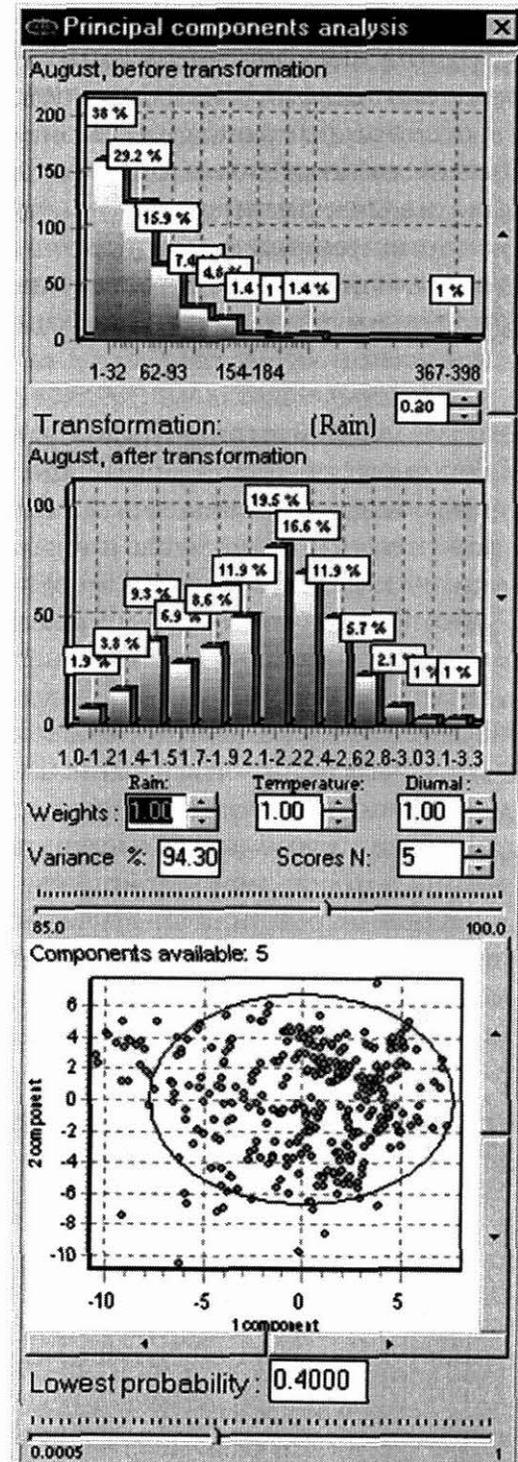
Es importante que usted utilice esta última opción solamente cuando está seguro de que los puntos a mover son los que están rozando con el límite de la base de datos climáticos. Si usted utilice esta opción cuando los puntos caen muy lejos del límite, serán automáticamente movidos al límite más cercano, lo cual, generalmente, es un error.

- ✓ Haga clic en OK para fijar nuevamente la configuración.
- ✓ De `layers menu`, seleccione `accession points` y luego `Stylo_secondfile.dbf`.

- ✓ En esta etapa, usted puede suprimir la capa mismatched de layers control window. Estos puntos ahora han sido corregidos y agregados al archivo de puntos de acceso. Para asegurar que están registrados, guarde el archivo de puntos de acceso de layers control window.
- ✓ Haga clic en el icono PCA para dar comienzo al nuevo análisis.

La ventana PCA aparecerá como a la derecha. Vamos a utilizar algunas de sus características para ajustar el siguiente análisis. En primer lugar, fije la transformación para datos de precipitación; estos datos son notoriamente anormales, a diferencia de los datos de temperatura, que casi siempre funcionan bien. A menudo se modelan como distribución gamma (ver Jones y Thornton 1993; 1997). Dado que FloraMap utiliza las características especiales de la distribución normal para determinar probabilidades, usted deberá seleccionar una transformación que cambie la distribución de los datos a una distribución que se aproxime tan estrechamente como sea posible a una distribución normal.

FloraMap actualmente ofrece dos transformaciones que son apropiadas para los datos de precipitación—el logaritmo natural, $\ln(\text{rain})$, y el exponencial, rain^x . En el primer caso, solamente existe una opción. En el segundo, los exponentes varían ampliamente, desde -3.0 hasta -0.1 y desde $+0.1$ hasta $+3.0$. No se admite el intervalo del medio por razones computacionales (exponentes pequeños positivos o negativos transforman todos los valores muy cercanos a 1.0 para ser de uso práctico). $\text{Rain}^{0.5}$ es, desde luego, la transformación de la raíz cuadrada. Rain^{-1} es el número recíproco. El cambio del exponente en la ventana proporcionada puede especificar muchas otras transformaciones. Las transformaciones comunes, como la transformación logit, la probit y la



trigonométrica, no están incluidos en FloraMap porque son de uso dudoso con datos de precipitación.

- ✓ Haga clic en **transformation** para pasar alternadamente entre la transformación del logaritmo y la transformación exponencial.

Observe cómo la elección de la transformación cambia la distribución en la ventana del diagrama de dispersión. En la medida que usted cambia el exponente, el número de puntuaciones seleccionados cambiará ocasionalmente. Esto sucede porque el algoritmo está tomando el porcentaje de variación explicado como constante y trata de ajustar el número de puntuaciones para compensar los datos cambiantes. Algunas de las transformaciones que usted puede lograr con la transformación de exponentes son inusuales, para decir lo menos, pero, aunque nunca sean usados para elaborar un mapa, pueden brindar una perspectiva interesante de las agrupaciones de los datos en el diagrama de dispersión.

La transformación debe producir, en la medida que sea posible, una típica curva normal, en forma de campana, para cada mes (desplácese a través de los meses con las barras de desplazamiento ubicadas al lado de los histogramas). Actualmente, no podemos ajustar la transformación individual a cada mes; debe bastar una sola transformación para todo el año. Generalmente no se pueden satisfacer simultáneamente los requisitos de una buena transformación para todos los meses. Aquellos meses con menor precipitación son especialmente difíciles de normalizar. En este caso, son julio, agosto y septiembre. Utilice las herramientas de transformación en forma interactiva para familiarizarse con lo que está haciendo y elija la transformación que le dé los mejores resultados para el año entero. Lamentablemente, quizás no sea la transformación que se adapte mejor a ciertos meses. La herramienta que se presenta aquí es lluvia a la potencia 0.2 o sea la quinta raíz de precipitación. Comprueba que ésta ofrezca el mejor ajuste para todos los meses, que junio sea el mes que peor ajuste presente, y que cualquier otra potencia introduzca más distorsión.



Esperamos que las versiones futuras de FloraMap incorporarán las ayudas estadísticas para seleccionar transformaciones y transformaciones más generalizadas.

Ahora que usted ha seleccionado su transformación, puede proceder al siguiente paso. Por ahora, dejaremos las ponderaciones en 1 y observaremos el efecto de escoger diferentes números de componentes principales.

Análisis de Probabilidad y Mapeo

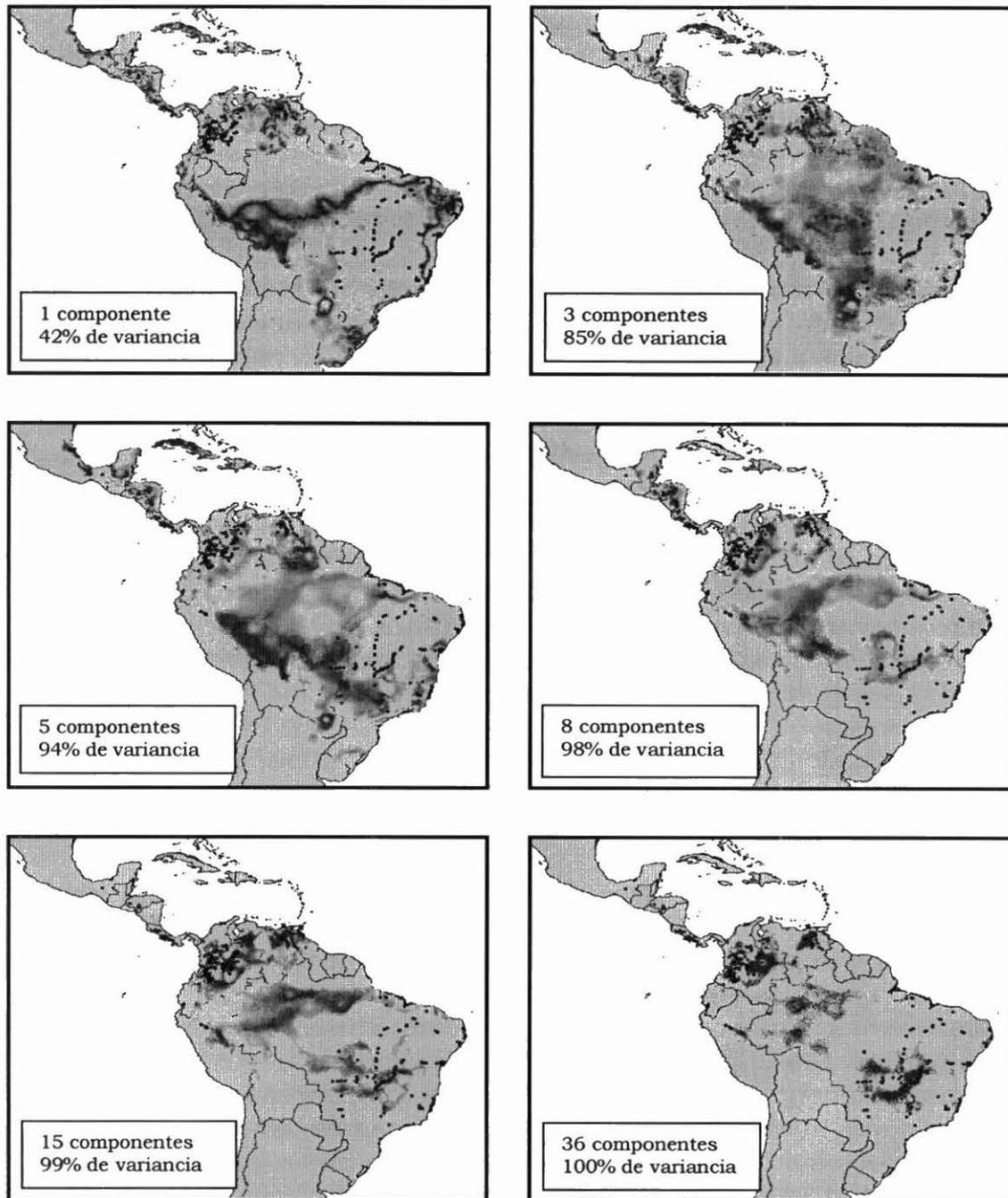
Estamos ahora preparados para comenzar a mapear las probabilidades. Lo que mapearemos es la probabilidad de que cada uno de los píxeles en el mapa y, en consecuencia, en la rejilla de datos climática, pertenezca a la población climática descrita por el análisis de PCA, que ahora aparece en la ventana de PCA (PCA window). Consulte la sección de TEORÍA, página 69, para los detalles.

- ✓ Verifique el recuadro de scores N y pase alternadamente entre los valores hasta lograr el número de componentes que le gustaría usar. Haga varios ensayos con diferentes números de componentes para ver que sucede. Usted puede realizar la misma operación al mover la palanca de variación (variance lever) o al indicar la variación deseada en el recuadro de variación (variance box). Observe que un rango de variaciones se aplica a cada serie de componentes (scores N) seleccionada.
- ✓ Siempre utilice el icono de visualizar mapa para producir la capa de probabilidad. Fije la probabilidad mínima en 0.4 para asegurar una cobertura total de los puntos de alta probabilidad. Usted puede suprimir las capas anteriores del layers menu si desea mantener un mapa menos confuso.
- ✓ Usted puede guardar las capas de probabilidad que parecen interesantes, dándoles un nuevo nombre y llamándolos como shapefiles generados (rendered shapefiles) del layers menu.



Las imágenes a continuación ilustran el efecto de utilizar diferentes números de componentes principales en el presente análisis con weights fijadas en 1 y la probabilidad más baja (lowest probability) en 0.4.

Los mapas con pocos componentes se ajustan a áreas grandes que están por fuera del rango de puntos de accesión y muestran un ajuste sorprendentemente inadecuado en las principales áreas de accesión. En la medida que aumenta el número de componentes, las probabilidades se ajustan más a los puntos de accesión y pierden importancia las áreas que están por fuera de las principales áreas de recolección. Esto no debe sorprenderlo. Los primeros dos o tres componentes principales miden características climáticas generales, de amplio espectro. El primero es generalmente un componente de tipo **tamaño**. La variación a lo largo de su eje a menudo se asociará con, por ejemplo, más lluvia o menos lluvia, temperaturas más altas o temperaturas más bajas. Es, por lo tanto, una medición aproximada del clima, pero explica gran parte de la variación (42%). En la medida que se vayan adicionado otros componentes al análisis, la proporción de la variación explicada aumenta rápidamente.



Mapas de probabilidad derivados de números diferentes de componentes principales.

Cuando ya se hayan incluido cinco componentes, estos acomodan un 94% de la variación. Solo queda escaso 6% para todos los otros componentes. Esto normalmente se consideraría como interferencia, que queda del registro y procesamiento de los datos climáticos. Podríamos considerar la posibilidad de ajustar hasta ocho componentes, lo cual

llevaría el ajuste a un nivel tan alto que sólo quedaría el 2% de la variación sin explicarse. Esto definitivamente debe ser interferencia. Las temperaturas solamente se miden a 0.1 °C y, sin duda, no con tanta precisión a largo plazo.

Entonces, ¿por qué parece aumentar la precisión del ajuste del área de probabilidad hasta el ajuste con 36 componentes? Tiene una doble respuesta. En primer lugar, la interferencia tiene suficiente coherencia aleatoria para que el ajuste parezca estar refinado, bastante más allá del punto en que debe sobreponerse la interferencia aleatoria. En segundo lugar, cada componente adicional trae consigo unos 13/12 grados de libertad respecto a los cálculos del parámetro de población.

Por tanto, si fuéramos a ajustar una serie de calibraciones de 36 observaciones, el mapa de la superficie de probabilidad para 33 componentes, por definición, se ajustaría exactamente a los puntos de calibración. En efecto, esto es lo que sucede—cada punto de accesión está cerca de un punto máximo local en la superficie de probabilidad. Pero no hay datos libres para producir una superficie de interpolación que sea válida. Por consiguiente, no queremos ajustar demasiados componentes principales, especialmente cuando se trata de pequeñas series de calibración. Saber exactamente cuántos componentes escoger no es una ciencia exacta, entre menos componentes mejor, pero deben arrojar una superficie de probabilidad que sea realista. Normalmente, preferiríamos usar entre cuatro y ocho componentes principales, según el tamaño de la serie de calibración.



FloraMap no le permitirá ajustar un mayor número de componentes que los datos teóricamente permiten. Sin embargo, usted todavía puede ajustar más componentes si lo justifica la información contenida en los datos.

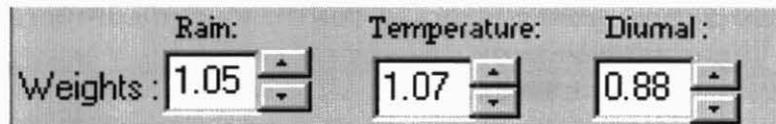
Lamentablemente, cuando ajustamos hasta ocho componentes principales en este conjunto de datos, quedan grandes áreas de alta probabilidad donde ningún punto de calibración cae, y grandes áreas donde los puntos de accesión presentan una probabilidad muy baja. Tener áreas de probabilidad positiva, sin puntos de calibración, quizás no sea una situación desfavorable. Bien puedan estar prediciendo correctamente la posible existencia de germoplasma que aún no se ha recolectado. Lo inverso, cuando el mapa muestra baja probabilidad para áreas con puntos de accesión, obviamente está equivocado. Para una posible explicación de lo que está equivocado, consulte la sección de TEORÍA, página 79.

Si aceptamos los argumentos expuestos en la sección de TEORÍA,

entonces estamos tratando con más de una distribución de población climática. En este momento, no tenemos pruebas aparte del ajuste de probabilidad en el mapa. No tenemos idea de cuántas poblaciones climáticas pueden estar ocultos en los datos de la serie de calibración. Mucho menos tenemos evidencia de que las diferentes poblaciones climáticas reflejan la variación genética en la especie que estamos trazando. FloraMap incorpora una herramienta que permite investigar estas posibilidades a partir de los datos climáticos. Le recomendamos enérgicamente regresar ahora a los datos de germoplasma y ver si puede recoger allí cualquier información que puede ser útil para separar posibles poblaciones. Más adelante, en esta misma sección, usted encontrará un sencillo ejemplo ilustrativo. En la práctica, es prudente empezar a pensar cuanto antes en este problema, desde ambos puntos de vista.

Una última opción que debemos cubrir en el análisis de probabilidad es la ponderación de las variables. Hasta ahora, hemos dejado los valores relativos en 1. Las 36 variables se agrupan en tres grupos de 12. Usted ya ha visto cómo se aplica la transformación a uno de los grupos de variables (precipitación). Las ponderaciones se aplican a todo el conjunto de datos en grupos de 12. Por consiguiente, hay tres valores relativos: uno para precipitación, uno para temperatura y otro para rango de temperatura diurna. Suman 3.

Para cambiar los valores relativos, observe la ventana de PCA donde aparecen los



recuadros de valores relativos y palancas. Aunque aún no hemos mirado la herramienta de agrupación, la siguiente es una recomendación útil. Es mejor desactivar la agrupación cuando usted cambia los valores relativos. Cada vez que usted cambia un valor relativo, se recalcula la agrupación, lo cual es un proceso engorrosamente demorado. Desde luego, si usted quiere ver el efecto del cambio del valor relativo en las agrupaciones, entonces debe proceder a la reagrupación después de que haya cambiado los valores relativos. En este caso, es preferible entrar al recuadro de valores relativos y digitar directamente el nuevo valor relativo. Si usted trata de pasar alternadamente entre los valores relativos utilizando las palancas, se recalculan tanto el PCA como las agrupaciones para cada punto que usted pasa.

- ✓ Mueve la palanca de valor relativo de la lluvia hasta llegar al valor más alto posible.

Observe cómo el diagrama de dispersión responde—las agrupaciones se concentran y luego desaparecen. Los puntos se dispersan fuera de la

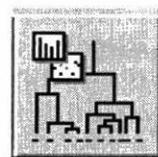
superficie de la elipse de distribución. Usted observará un comportamiento similar si usted pasa alternadamente a través de los valores relativos correspondientes al rango de temperatura o al rango de temperatura diurna. Es probable que nunca usará una ponderación de 3 para la precipitación, pero la opción es sumamente útil mientras usted se desliza, aumentando interactivamente los valores relativos. A menudo, la acción dinámica de los puntos puede indicarle cuáles puntos son especiales. Para un punto de datos, ser especial muchas veces es, lamentablemente, indicación de error. El comportamiento extraño debe investigarse. Al menos, esto lo llevará a identificar entornos inusuales en los datos de accesión. Haga un círculo alrededor de éstos en el diagrama de dispersión utilizando la herramienta de selección de áreas. Observe dónde están ubicados estos entornos.

Generalmente los cambios en la ponderación para uso real en el análisis de PCA y agrupación no se pasaran a los extremos. A diferencia de la transformación, no podremos comprobar inmediatamente qué tan bueno se ve el resultado. En el caso de la transformación, este es la ventana). El diagrama de dispersión es la única retroinformación inmediata de la cual usted dispone cuando altera los valores relativos. Use el diagrama para buscar modelos interesantes que luego puede utilizar para alimentar la herramienta de agrupación. Si usted tiene conocimientos previos acerca del germoplasma en que está trabajando, entonces use esa información para fijar los valores relativos. Si sabe que algunas de las especies son especialmente sensibles a la temperatura, entonces se justificaría aumentar el valor relativo de la temperatura. Sin embargo, recuerde que cuando fija uno o dos valores relativos muy altos, está perdiendo la información de la serie de calibración.

Cuando se desplazó a través de los valores relativos de precipitación, ¿observó algo interesante? Al pasar usted por el valor relativo de 1.5, el diagrama de dispersión se unió en grupos compactos, sólo para rápidamente dispersarse nuevamente en la medida que usted aumentaba el valor relativo. Esta es el tipo de indicación de estructura que estamos buscando en los datos. Fijemos el valor relativo de la precipitación en 1.5 y los otros valores relativos de la temperatura en 0.75 cada uno.

Análisis de Agrupación

La herramienta de función del análisis de agrupación se incluye en FloraMap para investigar la posibilidad de poblaciones múltiples.



- ✓ Cierre el layers menu para liberar espacio en la pantalla.
- ✓ Haga clic en el icono de agrupación en map window.

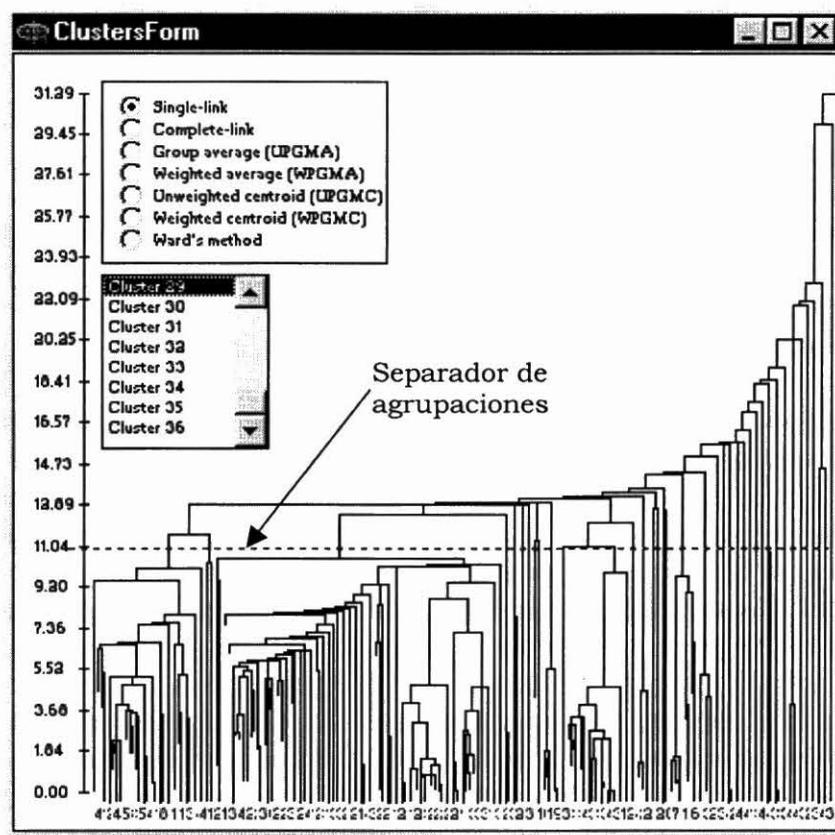
Se abrirá la ventana de agrupación (*cluster window*) y el *process indicator* mostrará cálculo de distancias (*calculating distances*) y luego agrupación (*clustering*). Hemos incorporado siete técnicas de agrupación, todas de tipo aglomeración; es decir, las agrupaciones aumentan mediante la adición de miembros o la fusión de agrupaciones. Las siete técnicas usan una medida de distancia euclidiana, son jerárquicas y producen un dendrograma de agrupaciones. Aplicamos estas técnicas al conjunto de datos climáticos original después de la transformación y la ponderación, pero antes del PCA. Por consiguiente, la elección de número de componentes principales no afecta a la agrupación. Para más detalles, consulte la sección de TEORÍA, página 81.

No existe un algoritmo **correcto** para el proceso de agrupación, y no existe resultado **correcto** para un análisis de agrupación. Aunque los métodos se basan en gran medida de las matemáticas de la teoría de gráficas y el álgebra de matriz, son métodos esencialmente subjetivos de búsqueda de maneras de simplificar las complejas estructuras de datos. El resultado final depende de lo que el usuario desea obtener del análisis. En nuestro caso, necesitamos subdividir la matriz de datos en series que cubren los diversos tipos de clima que encontramos en la serie de calibración.

- ✓ Ensaye las técnicas de agrupación (*cluster techniques*). Seleccione un método del menú en *cluster window*. El proceso de agrupación empezará automáticamente.

El método que usted eventualmente seleccione dependerá de qué está buscando exactamente. Si desea sencillamente dividir la serie de accesiones en dos o tres agrupaciones, entonces seleccionará un método que arroje agrupaciones compactas, limpias, bien divididas, a un alto nivel. Los métodos de enlace completo (*complete-link*) y Ward generalmente hacen esto bien. El método de Ward es particularmente apropiado porque intenta minimizar el término de error al cuadrado dentro de las agrupaciones.

Si usted está buscando estructura detallada en el conjunto de datos, entonces a veces es útil el algoritmo de enlace único (*single-link*) que une a las agrupaciones y agrega los elementos por la distancia más corta. Está estrechamente relacionado con el árbol mínimo con conectividad completo (*minimum spanning tree*) de la teoría de gráficas. A menudo produce un dendrograma de apariencia desordenada debido a su tendencia de vincular los miembros uno a la vez, en agrupaciones extendidas en largas hileras. Rara vez sirve para proporcionar cortes limpios en las agrupaciones principales, pero el detalle puede ser ilustrativo.



- ✓ Hale hacia abajo el separador de agrupaciones haciendo clic en el eje *y* donde cruza la línea horizontal punteada. Hálolo hacia abajo hasta que tenga 36 agrupaciones en pantalla.
- ✓ Seleccione cada agrupación que comienza con el número 36. Haga clic a la derecha en la superficie del mapa y seleccione visualizar conjunto de datos (view dataset) del menú pequeño.

Usted observará que las últimas siete agrupaciones contienen una sola accesión o dos accesiones muy similares. Muchas veces se considera como un punto débil la manera irregular en que los algoritmos de enlace único siguen agregando el próximo punto más distante. Ciertamente lo es si lo que usted desea es una serie depurada de unas cuantas agrupaciones de tipo globular.

Es excelente como herramienta para buscar resultados aislados en los datos. Mire la lista de accesiones de la agrupación 36 hasta la 30. Contiene:

- (a) El único punto que cae en México;
- (b) Tres puntos sin datos de elevación;
- (c) Nuestro viejo amigo Sgui2163 – domesticado, pero no verdaderamente alineado a una elevación de 2150 metros; y
- (d) Dos pares de puntos de la región de Chocó, Colombia, donde la

precipitación es excepcionalmente intensa.

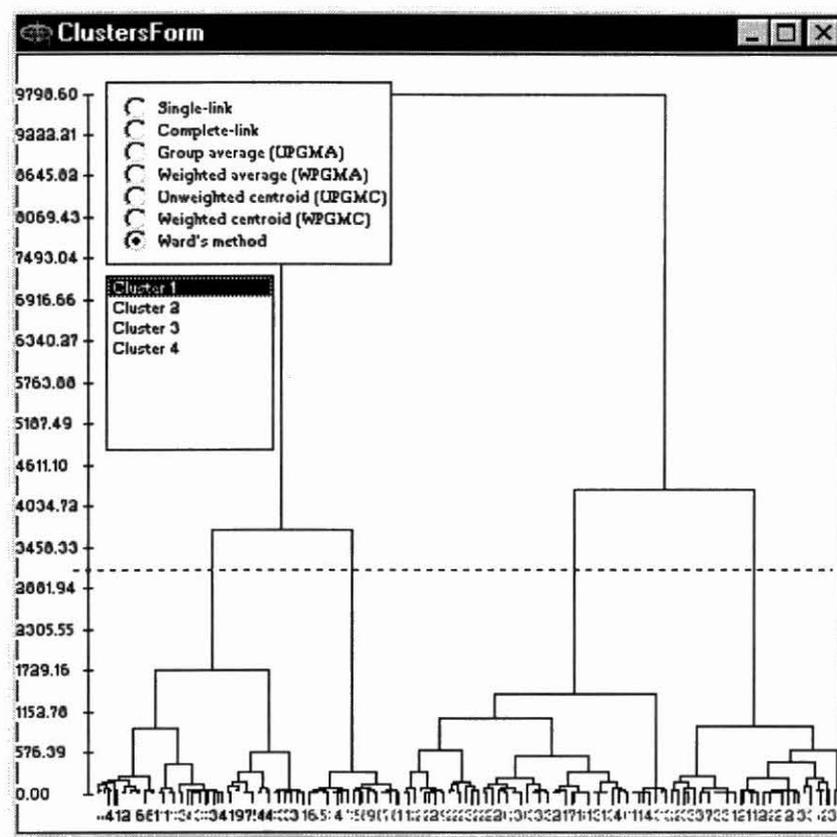
No vamos a agobiarle con el ajuste de estos datos. Tomaría mucho tiempo y hasta puede significar tratar de completar algunas de las colecciones. Sin embargo, es el tipo de depuración que se debe hacer en un conjunto de datos real y puede conducir a nuevas apreciaciones acerca del germoplasma en estudio. Por ejemplo, ¿por qué hay un solo punto en México y uno por encima de los 2000 metros? ¿Es una característica del germoplasma o del recolector?

Así que procederemos con el análisis como si no se hubieran presentando estos resultados aislados. Miremos nuevamente lo que usted puede obtener del algoritmo de enlace único. En la parte inferior izquierdo del diagrama, usted puede observar lo que parecen ser grupos de agrupaciones óptimas. Para llegar a ellos, deberá que halar hacia abajo el **separador de agrupaciones** hasta que se delinee cerca de 55 agrupaciones.

- ✓ Seleccione las agrupaciones desde el 1 hasta, por ejemplo, el 16, haciendo clic en el número de la agrupación en la ventana. Ignore las pequeñas lonjas de agrupaciones y observe las agrupaciones principales. Igual que antes, llame la hoja electrónica con **view dataset**.

Observe dónde aparecen las agrupaciones en el diagrama de dispersión. Los puntos de la agrupación seleccionada aparecerán de color azul. Note que algunos comprenden casi todas las accesiones con valor 12 bajo ISO, otros tienen predominantemente el valor 1. Estos valores son las agrupaciones de un estudio que se realizó con isoenzimas. El clima y la agrupación de isoenzimas parecen estar relacionados, pero este análisis de agrupaciones no ilustra claramente esta relación.

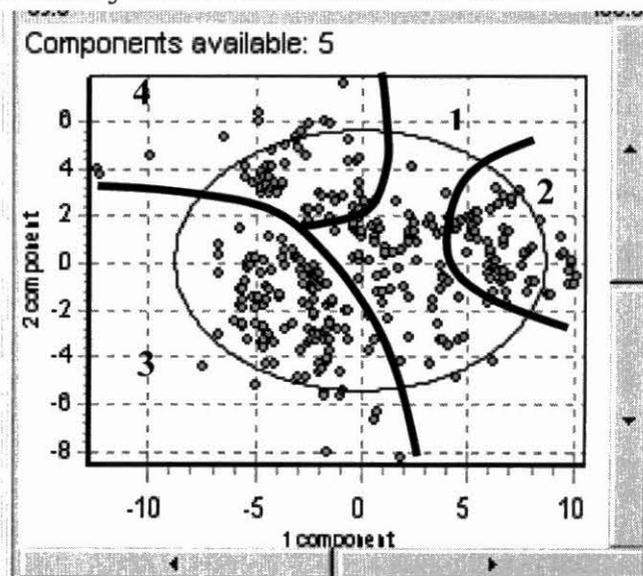
Mire ahora las agrupaciones obtenidas con el **método de Ward**. Son mucho más definidas y compactas. El **separador de agrupaciones** fácilmente puede cortar el conjunto de datos en cuatro agrupaciones de casi el mismo tamaño.



- ✓ Ubique el cursor en una línea de dendrograma para visualizar las características de la agrupación por debajo de ese punto.
- ✓ Ubique el cursor en el dendrograma, haga clic y hale hacia abajo hacia la derecha para abrir una ventana de zoom. Ubique el cursor en la superficie del dendrograma, mantenga suprimida la tecla de cambio y, con el ratón, empuje el dendrograma para girar y lograr una visión panorámica de nuevas áreas.

Mire ahora el diagrama de dispersión y seleccione cada uno de las cuatro agrupaciones, una a la vez, con el diagrama de dispersión sobre los primeros dos componentes. Usted puede ver la manera en que las agrupaciones dividen nitidamente el diagrama de dispersión en cuatro secciones. Hasta se puede apreciar un reducido espacio entre las secciones.

- ✓ Seleccione cada agrupación en forma sucesiva y llame la



hoja electrónica con **view dataset**.

- ✓ Desplácese a través de las hojas electrónicas observando la incidencia de 12 ó 1 para el carácter de ISO.

Las agrupaciones 3 y 4 son, con solo dos excepciones, todas de la clase 12 de ISO. La agrupación 2 es mixta, más o menos mitad-y-mitad, pero la agrupación 1 es predominantemente de clase 1 de ISO. La indicación de que el clima y la clasificación de isoenzimas puede estar relacionado es potencialmente importante, pero retomaremos este punto más adelante.

Tratemos ahora de ajustar las agrupaciones climáticas independientes como conjuntos de datos individuales. Usted puede hacer estos directamente desde la ventana del mapa, sin tener que salir de FloraMap.

- ✓ Despeje cualquier capa de probabilidad del mapa con el menú de capas.
- ✓ Seleccione la agrupación 1 (**cluster 1**) en **cluster window**.
- ✓ Verifique que la agrupación se seleccione correctamente en el diagrama de dispersión.

Ahora usted puede seguir una de dos rutas para elaborar el mapa de la agrupación. La primera ruta sirve como verificar rápidamente cómo se verán las agrupaciones en el mapa. La segunda le dará completo control sobre el nuevo PCA y le dejará verificar la coherencia de los datos.

Ruta 1

- ✓ Sencillamente haga clic en el icono del mapa una vez haya seleccionado la agrupación. El PCA se recalcula para la agrupación seleccionada utilizando los valores fijados en la **ventana de PCA**. Usted no tiene la posibilidad de cambiar las opciones, y nada cambia en la ventana de PCA. Para visualizar los mapas de las otras agrupaciones, sencillamente seleccione la agrupación y luego elija nuevamente el icono del mapa.



Ruta 2

- ✓ Haga clic en el icono PCA. Se calculará el PCA de la agrupación seleccionada. Verifique que todo esté bien en el diagrama de dispersión. A veces aparecen valores alejados de lo normal que no eran perceptibles en etapas anteriores del análisis. Aprieta los dientes y espere que esto no suceda, porque si pasa, deberá buscar el error y comenzar de



nuevo. Usted ahora puede fijar valores relativos, cambiar la transformación y seleccionar el número de componentes en el PCA como si fuera un análisis completamente independiente (lo cual es). Si todo sale como usted espera, seleccione la probabilidad mínima para elaboración de mapas. Dado que usted eventualmente estará sobreponiendo tres superficies de probabilidad, le sugerimos fijar esta probabilidad en un valor razonablemente alto—en 0.5 o hasta 0.6.

Si usted encuentra puntos dudosos, querrá ver cual información tiene acerca de los mismos.

- ✓ Haga clic derecho en cualquier parte de la pantalla y seleccione **view dataset**. Aparecerá una hoja electrónica con las accesiones de la agrupación seleccionada.



La funcionalidad de las hojas electrónicas mostradas en FloraMap es mínima. Estas hojas sólo existen para mostrar los datos. Usted no puede cambiar, agregar, suprimir o clasificar los datos en estas hojas electrónicas. Para hacerlo, deberá salir de FloraMap y usar su propio programa de hoja electrónica para hacer los cambios.

- ✓ Haga clic en el icono **visualizar mapa**.
- ✓ Haga clic en el icono **control de capas** y cambie el color a uno distintivo.
- ✓ Despeje la ventana de PCA y la hoja electrónica cerrando las ventanas. Seleccione otra agrupación y haga clic en el icono PCA para seguir reevaluando el PCA y elaborando mapas de la agrupación.

Usted puede guardar las capas de probabilidad de la agrupación para uso futuro utilizando el icono **guardar capa** en **layers menu**. Le recomendamos dar un nuevo nombre a la capa cuando la vaya a archivar. Si le deja los nombres por defecto, el archivo puede sobrescribirse en un análisis posterior. Para guardar el conjunto de accesiones seleccionado, elija **archivar (file)**, luego **guardar (save)** y asigne un nombre.

El mapa que usted ha creado con las cuatro agrupaciones, a partir de la agrupación del método de Ward, presenta un ajuste mucho mejor respecto a las accesiones que el ajuste de población única, aunque no cubre todos los puntos de accesión. Usted no debe esperar que el modelo de probabilidad alguna vez represente todos los puntos en el conjunto de accesiones. Damos por sentado que éstos se extraen de una población normal y, por consiguiente, algunos siempre se ubicarán por fuera de

cierto nivel de probabilidad.

Las agrupaciones tienen una realidad geográfica y climática obvia. La agrupación 1 refleja los climas moderados de la región costera de Brasil, Paraguay y el Caribe. La segunda agrupación abarca principalmente los climas secos cálidos de los *cerrados* brasileños, con sólo una zona muy pequeña en El Salvador que queda por fuera de la región. La tercera incluye las regiones más frescas de la región andina, el sur de Brasil y los altiplanos de América Central. Finalmente, la cuarta agrupación representa los cálidos climas húmedos de la región central del Amazonas, los *llanos* colombianos y la Costa de Mosquitos de América Central. Esta clasificación y el mapeo son lógicos de modo tranquilizador. Moviendo hacia abajo en el dendrograma a más agrupaciones, nos da una clasificación aún más ajustada, con algunas zonas que pueden ser útiles para nuevas exploraciones. Sin embargo, existe el mismo peligro de ir demasiado lejos con la técnica de agrupación como encontramos con el número de componentes. Entre más agrupaciones usted separa, más grados de libertad usará de los datos. Una multitud de agrupaciones pequeñas pueden ajustarse extremadamente bien a los puntos, pero casi no le darán capacidad de predicción.



Planeamos incorporar algunas estadísticas de bondad de ajuste y de validez del modelo tanto en la selección de componentes principales como en la herramienta de agrupación, las cuales estarán disponibles en la próxima versión completa de FloraMap.

Ahora hemos llegado al final de esta sección de análisis en FloraMap, usando solamente datos climáticos. Usted ya ha aprendido qué debe buscar en los datos en función de uniformidad y error. Ha visto algunas técnicas para corregir errores al igual que la forma cómo usar el diagrama de dispersión y cómo aplicar las transformaciones y los valores relativos. Usted ha realizado pruebas de análisis de agrupaciones para determinar si más de una población está mostrando una adaptación del clima diferente en la serie de calibración. Usted ha visto cómo se puede mirar datos externos sobre el germoplasma, desde adentro de las agrupaciones, para tener una idea de cómo estos factores están conectados con la variabilidad del clima. Ahora consideraremos la aplicación directa de la información sobre el germoplasma en la elaboración de mapas.

- ✓ Salga de FloraMap y lleve *stylo-secondfile.dbf* a su programa de hoja electrónica.
- ✓ Clasifique el archivo por la columna ISO y crea dos nuevos conjuntos de datos de accesiones. Nómbralos Grupo_1 y Grupo_12, o algo parecido.

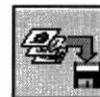
La agrupación aplicada a los datos de la isoenzima $\alpha\beta$ fosfatasa ácida ahora separa estos conjuntos de datos de accesión. Los grupos también se correlacionan con las variedades *Stylosanthes guianensis* var. *vulgaris* (grupo 1) y *S. guianensis* var. *pauciflora* (grupo 12). En la actualidad, FloraMap solamente puede manejar un archivo de clima generado a la vez. Los siguientes procedimientos, sin embargo, le permiten construir un mapa con tanto las probabilidades del clima como los puntos de accesión de dos conjuntos de accesiones.

- ✓ Ingrese nuevamente a FloraMap y cargue las accesiones para el primer grupo.
- ✓ Proceda con las operaciones de PCA y MAP para este conjunto de datos.
- ✓ Archive los puntos de accesión y la superficie de probabilidad como archivos de capas del mapa (map layer files), dándoles nombres diferentes.



Estos shapefiles están de hecho disponibles en el directorio de trabajo, pero esta característica puede cambiar en versiones futuras.

- ✓ Elimine los accession points y la capa de probabilidad (probability layer) del mapa. En la mayoría de las circunstancias, usted puede dejar la capa de probabilidad en el mapa, pero como a veces se presentan conflictos con el uso de ciertos shapefiles es mejor despejar el mapa.
- ✓ Repita el procedimiento con el segundo conjunto de accesiones.
- ✓ Cargue nuevamente el mapa con los guardados map layer files. Se cargarán los puntos de accesión como un shapefile, y de esta manera no tendrán que pasar por el proceso normal prolongado de verificación y creación del archivo del clima.
- ✓ Asigne colores nuevamente al mapa, a su gusto, y archive el mapa completo utilizando el icono guardar mapa.



Vimos indicaciones en algunos de las agrupaciones de clima que la evidencia genética y la adaptación climática pueden asociarse. El mapa que usted acaba de crear claramente corrobora este hecho. Cada grupo, definido según los datos isoenzimáticos, se adapta a los regímenes diferenciados de clima con poca superposición geográfica excepto en el sudoeste de Amazonia.

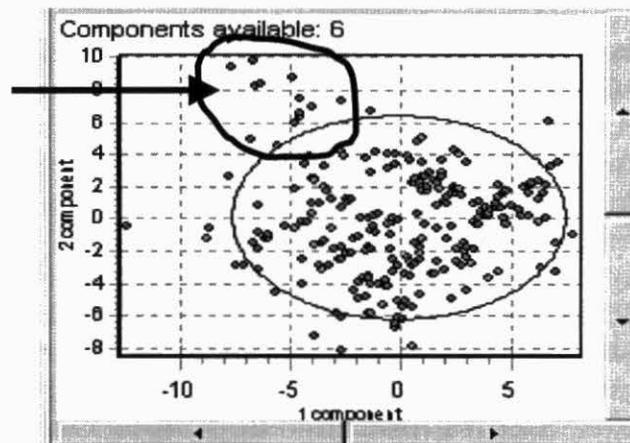
Hay una última cosa que podemos hacer con estos datos. Supongamos que usted es responsable de enviar germoplasma promisorio de leguminosas a África. Acabamos de mostrar que los materiales del grupo 12 tendrán un rango de adaptación diferente a los materiales del grupo 1. Entonces, ¿qué materiales enviaras adónde?

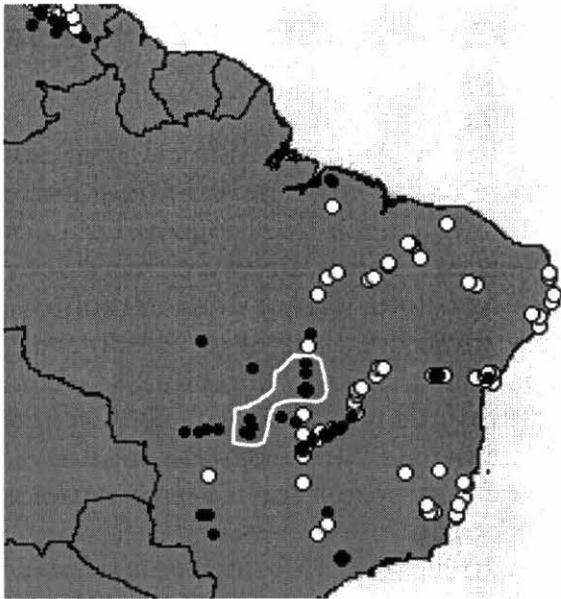
- ✓ Seleccione el archivo de rejilla climática correspondiente a África bajo la configuración general de marcos de trabajo (settings configuration general).
- ✓ Vaya carga una capa (load a layer) en layers control window.
- ✓ Abra el shapefile AFRCOUNTRIES del directorio \COVERAGES\ en el CD-ROM.
- ✓ Nada aparecerá en la pantalla hasta que usted seleccione zoom lejos (zoom out); la cobertura adicional para África se ha creado en el espacio a la derecha de su pantalla.
- ✓ Una vez que usted puede ver ambos continentes, puede cargar los dos archivos de accesiones agrupados para elaborar el mapa de la distribución potencial tanto en África como en América Latina.

Conclusión

Si usted ha logrado terminar este tutorial con nosotros, ha visto prácticamente todo lo que FloraMap puede hacer. Por si mismo, el programa rara vez puede probar o refutar hipótesis por la naturaleza subjetiva de muchos de los procedimientos que empleamos. FloraMap puede servir para guiar el pensamiento y ayudar a formular hipótesis. Usted ha visto que FloraMap puede separar poblaciones al agrupar los climas en que se encontraron accesiones. Algunos de estas agrupaciones pueden asociarse con variaciones taxonómicas y genéticas. Cuando trazamos los diferentes grupos genéticos, se destacan diferentes adaptaciones climáticas. El sistema de FloraMap es particularmente eficaz para la investigación interactiva de conjuntos de datos.

Tratando este ultimo punto, miremos el análisis final que usted acaba de terminar. El diagrama de dispersión del grupo 12 muestra un grupo compacto de puntos al lado superior izquierdo del diagrama que no parecen ser un resultado casual.





En el mapa a la izquierda, los puntos negros conforman el grupo 12 y los puntos blancos, el grupo 1. La línea blanca encierra los puntos que se encuentran en el círculo en el anterior diagrama de dispersión. Son los puntos del grupo 12 que están mostrando un clima atípico para ese grupo y se recopilaron en una zona donde el tipo del clima es predominantemente del tipo del grupo 1.

Hasta aquí puede ayudarle FloraMap en este caso específico.

El próximo paso es regresar al germoplasma y averiguar si este es un acontecimiento aleatorio o si tiene alguna importancia para estas accesiones específicas de germoplasma.

3. Sección de Referencia para el Usuario

FloraMap usa una variedad de ventanas diferentes. Algunas, como la pantalla de mapas, presentan una barra de títulos en la parte superior que muestra el icono de control de la aplicación FloraMap al lado izquierdo, luego el título de la ventana y, a la derecha, los iconos de minimizar, maximizar y cerrar. Se puede cambiar el tamaño de estas ventanas tirando de la esquina derecha inferior. La ventana de PCA contiene un icono de control de la aplicación, el título de la ventana, y solamente un icono para cerrar, a la derecha. No se puede cambiar el tamaño de esta ventana o minimizarla. La ventana de capas tiene una barra de títulos más sencilla y no se puede cambiar su tamaño. Cuando se activa, siempre aparecerá en la parte superior de la ventana y debe moverse a una parte de la pantalla poco importante para que no oculte las otras pantallas. Puede apagarse cuando no se necesite. La barra de títulos se ve azul oscuro cuando la ventana está activa.

Debido a la alta demanda de la unidad central de procesamiento (CPU, el acrónimo inglés) que la mayoría de los cálculos hechos en FloraMap exigen, hemos diseñado para que gran parte de estos cálculos se hagan a un alto nivel de procesamiento. Esto significa que a veces las ventanas de FloraMap flotarán en frente de otra aplicación en uso. Si esto le molesta, salga de FloraMap o reduzca la ventana del mapa a un icono mientras esté usando una aplicación alternativa.

La Ventana del Mapa

La pantalla del mapa (mapscreen) es la principal pantalla de elaboración de mapas y es la primera que aparece al entrar al programa. El ejemplo muestra un mapa de América Latina y consta de dos capas, el mapa básico con los límites de los países y los puntos de acesión para *Stylosanthes guianensis*. Este mapa preparado con antelación ha sido cargado de un MAP file.

Los menús desplegables presentan la barra de menú debajo de la barra de títulos. Las funciones disponibles se describen a continuación y también están en gran parte disponibles a través de los iconos de función. Los números que aparecen al lado inferior derecho de la ventana corresponden a la longitud y la latitud del cursor en grados decimales; la latitud tiene un valor negativo hacia el sur y la longitud un valor negativo hacia el oeste.

La ventana de información (information window) aparecerá en cualquier momento en que el cursor queda sobre la superficie del mapa sin moverse durante unos segundos. Describirá las características de la capa superior del mapa señalado por el cursor.

El menú pequeño está disponible en cualquier momento haciendo clic derecho con el cursor en cualquier parte de la superficie del mapa. Las opciones que muestra son:

View dataset	muestra una página electrónica con el conjunto actual de accesiones.
Load	carga un mapa de un MAP file; es equivalente al icono de cargar mapa.
Save	guarda el mapa actual con todas sus capas en un MAP file, y
Configuration	muestra el menú de configuración.

En la barra inferior de la ventana se ha incorporado un indicador de proceso que nombra el proceso en operación actual y el intervalo transcurrido.



Menús desplegables

Los menús desplegables incluyen operaciones que están disponibles mediante la selección de iconos. Más adelante se dan detalles adicionales bajo la explicación de cada icono.

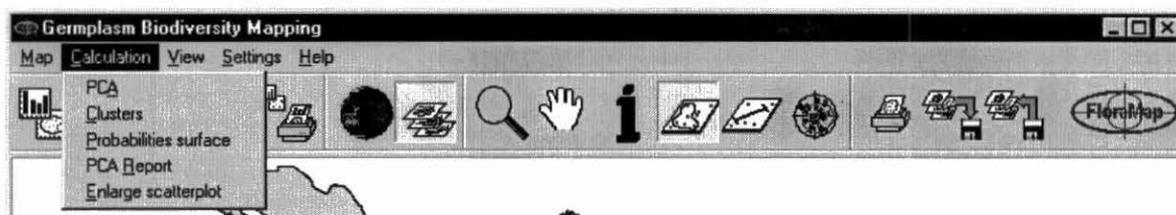


Haga desplegar la opción **map** y aparece como arriba

1. Layers – continúa con el menú de capas.
2. Load – carga un mapa anteriormente elaborado.
3. Save – guarda el mapa que usted está elaborando.
4. Print map – imprime el mapa.
5. Print diagrams – imprime los histogramas de precipitación antes de la transformación y después de ella, además de todos los diagramas de dispersión posibles.
6. Exit – sale de la sesión de FloraMap.



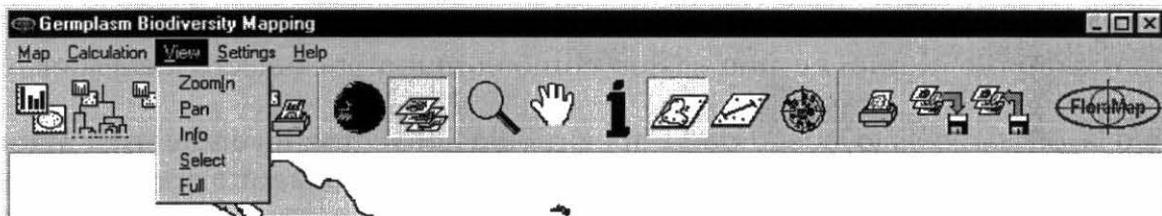
Nota que cuando se selecciona todas las diagramas de dispersión posibles puede existir un enorme número de componentes.



Haga desplegar cálculo (**calculation**) y aparece como arriba:

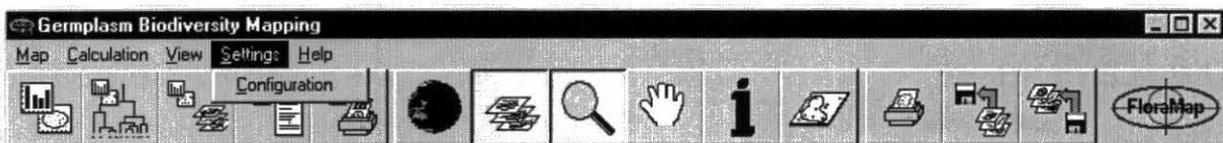
1. PCA – ejecuta el análisis de componentes principales.
2. Clusters – hace el análisis de conglomerados.
3. Probabilities surface – calcula la probabilidad y elabora el mapa correspondiente.
4. PCA report – produce el archivo de salida del análisis de PCA.

- Enlarge scatterplot – copia la ventana del diagrama de dispersión en una ventana más grande donde se pueden separar las accesiones individuales.



Haga desplegar visualizar (**view**) y aparece como arriba.

- ZoomIn – hace un acercamiento para magnificar las áreas del mapa.
- Pan – Mueve el mapa magnificado para permitir ver detalles adicionales.
- Info – Muestra la información sobre las capas del mapa.
- Select – Selecciona un área del mapa y los puntos que abarca.
- Full – Hace un alejamiento para que se pueda visualizar el mapa completo.



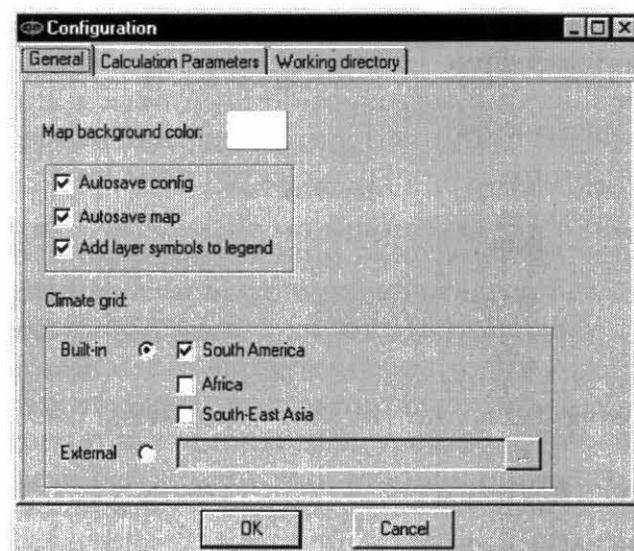
Haga desplegar **settings** y aparece como arriba.

Solo aparece configuration y no está en la lista de iconos; por tanto, se explica aquí. Haga clic configuration en el menú desplegable o en el menú de clic derecho. Aparecerá la ventana de configuración que es una ventana ajustable con tres ventanas de funciones.

Las siguientes opciones aparecen bajo **general**.

Autosave configuration – Si se selecciona (activa), la configuración que usted eligió

mientras trabajaba se mantiene en la configuración por defecto cuando usted sale de FloraMap. Siempre usted ingrese al sistema, aparecerá la última



en

configuración que usted usó. Si usted no hizo selección alguna, entonces aparecerá la configuración por defecto.

Autosave map ofrece la misma opción que la anterior, pero respecto al mapa de trabajo. Todas las capas con sus formatos y colores se guardan automáticamente al final de una sesión como el mapa por defecto (default.map).

Add layers symbols to legend da la opción de colocar una leyenda en el mapa. La opción también se puede activar bajo la ventana de capas, que se discutirá más adelante.

Map background color solamente se puede fijar aquí. Normalmente será un azul claro para el mar, pero se puede cambiar a otros tonos según la impresora a color utilizada.



Su impresora a color quizás no reproduzca fielmente los colores que usted ve en pantalla. Las impresoras varían en las tonalidades que pueden producir. Haga varias impresiones a manera de ensayo para determinar cuál es el mejor color para usar de fondo para la impresión.

Rejillas climáticas

Rejillas incorporadas

Las rejillas climáticas son cruciales para el trabajo de FloraMap. Usted debe especificar las rejillas que usted necesitará para su análisis. Actualmente existen tres rejillas climáticas internas (ya incorporadas) que cubren la mayor parte del trópico. Usted puede seleccionar cualquiera combinación de éstas. Las rejillas que corresponden a América Latina y África están a un tamaño de pixel de 10 minutos del arco (cerca de 18 km); la rejilla correspondiente a Asia está a 5 minutos del arco (cerca de 9 km).



*El motor de la base de datos utilizado en esta versión de FloraMap es lento con estos archivos grandes. **No seleccione las rejillas climáticas a menos que definitivamente las vaya a usar.** Es poco probable que usted tenga un conjunto de calibración que abarque los continentes. Solamente seleccionará dos o más rejillas cuando vaya trazar la superficie de probabilidad sobre dos continentes juntos en el mismo mapa.*

Rejillas externas

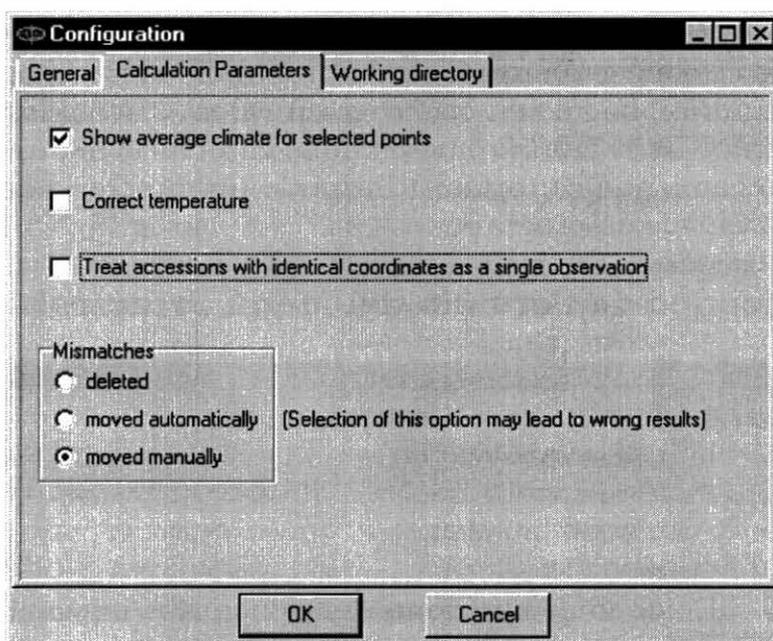
Estos son archivos de rejilla similares a los de las rejillas internas, pero no apoyan las múltiples funciones de trazado de mapas que tienen las rejillas

internas. Para usarlas, todas las demás rejillas deben desactivarse, lo cual se hace automáticamente al seleccionar usted una rejilla externa. La superficie de probabilidad aún puede trazarse sobre otra rejilla, pero no puede hacerse sobre dos o más rejillas al mismo tiempo. Las rejillas externas pueden ubicarse en cualquier directorio que usted desea usar.

La única rejilla externa que se incluye en FloraMap versión 1 es la rejilla de 5 minutos del arco, correspondiente a los 48 estados inferiores de Estados Unidos continental. Este es un nuevo muestreo de las superficies climáticas de 2.5 minutos del arco gentilmente proporcionado por Chris Daly de la Universidad de Oregon (ver Daly y Taylor 1998a y b). Usted encontrará esta rejilla en el subdirectorío \external\ bajo climate grids en el CD-ROM.

Bajo **calculation parameters** usted encontrará diversas opciones que son sumamente importantes.

Show average climate for selected points es una nueva e importante característica de esta versión. FloraMap calculará el clima promedio para un grupo seleccionado de puntos. Esto funcionará para los puntos seleccionados como un conglomerado en el dendrograma [ver cluster window, página 65], los puntos seleccionados en el diagrama de dispersión o el mapa [ver icono seleccionar área), página 49].



El clima promedio se indicará en el diagrama del clima (ver página 53).

Solamente puede usarse correct temperature si usted tiene datos de elevación en su conjunto de datos de accesión. Recomendamos enfáticamente que use esta opción de ser posible. No es necesario tener una serie completa de elevaciones para todas las accesiones (ver sección TUTORIAL, página 18).

Treat accessions with identical coordinates as a single observation controla la forma en que FloraMap interpreta un conjunto de accesiones. Determina si la unidad del análisis es una accesión o un punto de recolección. Para mayor información, ver la sección TUTORIAL, página 7.

Mismatches establece la forma en que se manejan los desaciertos. Los

desaciertos son los puntos de acceso que no coinciden con un pixel válido en el archivo de la rejilla climática. (Ver la sección TUTORIAL, página 11). Establecer que estos desaciertos sean suprimidos (deleted) por el valor preasignado es más bien drástico, pero a veces se necesitan medidas drásticas. Escoger la opción de **moved automatically** ubicará los desaciertos en el pixel más cercano en el archivo de la rejilla climática. Esta opción es apropiada para aquellos puntos que caen unos cuantos kilómetros de la costa, pero los puntos que están más lejos deberán verificarse manualmente, usando **moved manually**.



*Recomendamos enfáticamente que la opción **moved automatically** no se active hasta que se expliquen todos los puntos distantes de desaciertos. Los puntos distantes de desaciertos **casi nunca** son correctos cuando se mueven al punto más cercano en la superficie climática.*

Working directory es el directorio de trabajo para sus conjuntos de datos de accesiones, la superficie de probabilidad y los archivos MAP. Para trabajar el tutorial, hágalo en **c:\Program Files\CIAT\FloraMap\ demo**. Después de trabajar el Tutorial, usted debe cambiarlo al directorio de su elección. No se recomienda continuar utilizando un directorio de datos bajo su directorio de archivos de programa.

Iconos de Servicio del Menú Principal

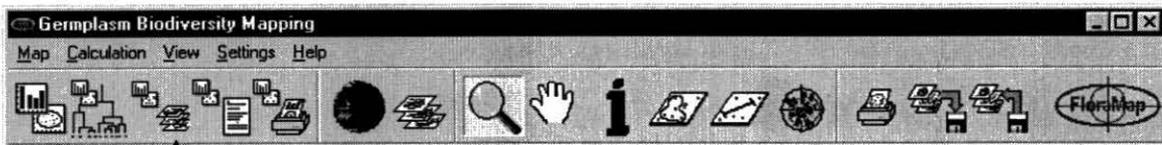


El **icono de componentes principales** inicia el proceso de PCA con el conjunto de datos de accesiones que ha sido cargado en el mapa, un conjunto definido por la *herramienta para seleccionar área*, o un conglomerado seleccionado de cluster window. Para detalles adicionales del análisis, consulte bajo ventana de PCA, página 62.

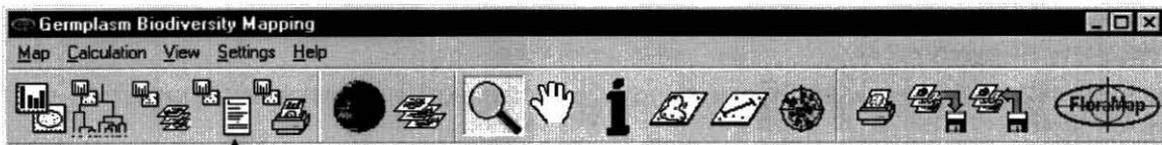


El **icono de agrupación** inicia la agrupación con todo el conjunto de accesiones, un conjunto definido por la herramienta para seleccionar área, o cualquier conjunto definido en la ventana de PCA. Este último puede usarse

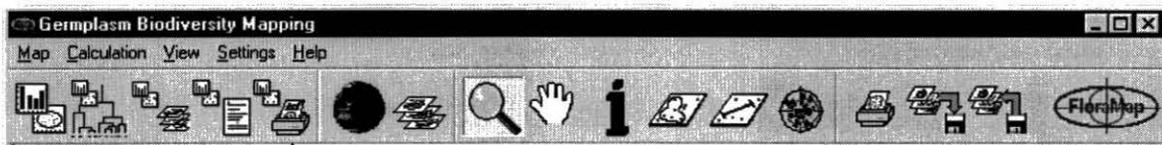
para producir una forma de agrupación repetitiva. El icono de conglomerados no puede iniciar el proceso de agrupación en un conglomerado seleccionado, pero si se introduce el conglomerado seleccionado en el PCA por medio del icono de componentes principales, entonces el icono de conglomerados usará ese conglomerado como el conjunto para agrupaciones subsiguientes. Es importante no activar la opción de agrupación hasta no necesitarla, porque hacer operaciones en la ventana de PCA genera agrupación automática, lo cual puede hacer que la operación sea lenta si los conjuntos de accesiones son grandes.



El **icono para visualizar mapa** toma el modelo de probabilidad calculado actualmente en la ventana de PCA y mapea la capa de probabilidad en el mapa que se encuentra en la ventana del mapa. Calcular la superficie de probabilidad puede tomar algún tiempo. Si usted inició el proceso y cambia de opinión, puede detenerlo moviendo el icono nuevamente.



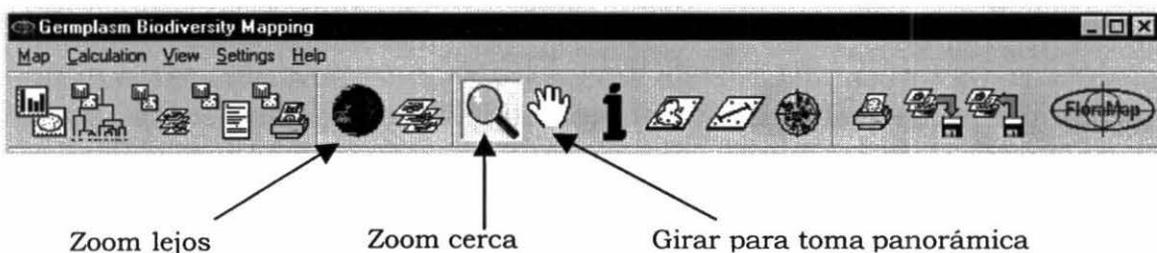
El **icono del informe del modelo** produce un archivo de informe que incluye los datos crudos, los detalles de la transformación, las medias y las variaciones, y los datos transformados, seguidos por valores específicos (*eigenvalues*) y coeficientes de componentes (vectores específicos, *eigenvectors*). Finalmente, se presentan las puntuaciones de cada componente para el número de componentes seleccionados. Cuando se trata de una serie de calibración grande y se selecciona muchos componentes, el archivo resultante puede ser muy grande. Recomendamos que lo inspeccione antes de imprimir.



El **icono para imprimir diagramas** imprime los diagramas de la ventana de PCA. Imprimirá los 12 histogramas de precipitación mensual antes de la transformación y después de la misma, además de todos los diagramas de dispersión posibles.



Nota que cuando se seleccionan muchos componentes, el número de diagramas de dispersión posibles puede ser grande.



Los **iconos de zoom** consisten en tres iconos relacionados, los cuales se tratarán en conjunto a continuación.

Ensaye el icono zoom cerca.

- ✓ Coloque el cursor en el mapa y tire hacia fuera un rectángulo sobre el área que usted desea ampliar. FloraMap ajustará el eje más largo de este rectángulo para llenar la ventana del mapa.

Una vez que usted tenga un acercamiento del área, puede ser necesario circular.

- ✓ Pase alternadamente el icono de giro para toma panorámica y coloque el cursor en el mapa.
- ✓ Haga clic izquierdo sostenido y mueve el mapa con el cursor. Aparecerá un

espacio en blanco en el área donde antes estaba el mapa, pero ésta se llenará rápidamente con la nueva cobertura del mapa en la medida que avanza el giro.

- ✓ Suelte el ratón y coja nuevamente el mapa soltando el botón izquierdo del ratón y haciendo clic.



Usted puede liberar y mover el mapa a una velocidad más rápida que la de la pantalla en llenarse. Tenga cuidado entonces de no perderse. FloraMap recordará todos los movimientos que usted ha hecho y seguirá moviendo el mapa hasta que todos sus giros se hayan cumplido.

Es sencillo hacer zoom lejos.

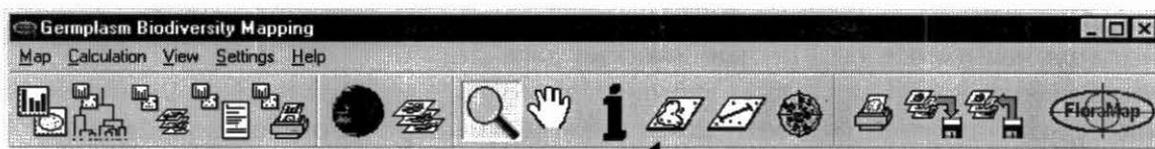
- ✓ Active el icono para zoom lejos. Se dibujará nuevamente el mapa completo. Las versiones futuras pueden tener una herramienta más sofisticado para visualizar de lejos, pero, por el momento, debe regresar al comienzo, no pasar adelante y no recolectar \$200.



El **icono de capas** controla las capas del mapa. Utilice este icono para activar y desactivar la ventana de control de capas. Con la ventana activada usted controla, en forma individual, todas las capas en el mapa. Ver la página 57 para una descripción detallada de la ventana de control de capas.



El **icono de información** es un método activo para encontrar información sobre cualquier parte del mapa. Seleccione el icono de información y haga clic con el cursor en la posición necesaria. El elemento del mapa se iluminará momentáneamente y aparecerá una pequeña hoja electrónica, la cual puede desplazarse a otra parte de la pantalla y permanecerá abierta hasta que usted lo cierra, aún si usted suprime la capa del mapa a la cual se refiere. La selección de otro icono del menú desactiva el icono de información.



El **icono para seleccionar el área** se usa para identificar los conjuntos de puntos de accesión según aparecen en la ventana del mapa y en el diagrama de dispersión en la ventana de PCA. Ver página 62 para la descripción de esta ventana.

A diferencia de algunos de los otros iconos que abren ventanas, el icono para seleccionar el área opera directamente tanto en la ventana PCA como en la ventana MAP. Es una de las herramientas más útiles que se proporcionan como icono. Como usted podrá apreciar en las secciones de TUTORIAL y de TEORÍA, muchas veces es necesario considerar la serie de calibración de accesiones como un número de poblaciones diferentes. En muchos casos, el icono para seleccionar el área es una forma sumamente flexible de hacer esta consideración. Usted puede usar este icono para identificar subconjuntos dentro del conjunto de accesiones, separarlos y archivarlos como conjuntos para analizarlos en forma independiente, o hasta proceder directamente a un nuevo análisis.

Para usar la herramienta, dibuje alrededor de los puntos en la ventana map o en la del scatterplot. Una línea roja, la cual se iluminará en forma intermitente, seguirá el cursor cuando se hace clic sostenido con el botón izquierdo del ratón. El polígono cerrará automáticamente tan pronto se libera el botón. Si usted está dibujando alrededor de puntos en el mapa y tiene más de un conjunto de puntos de accesión cargado, deberá seleccionar el conjunto pertinente en la ventana de control de capas. Por tanto, si usted quiere identificar un conjunto de puntos de desaciertos, debe resaltar el conjunto de desaciertos.

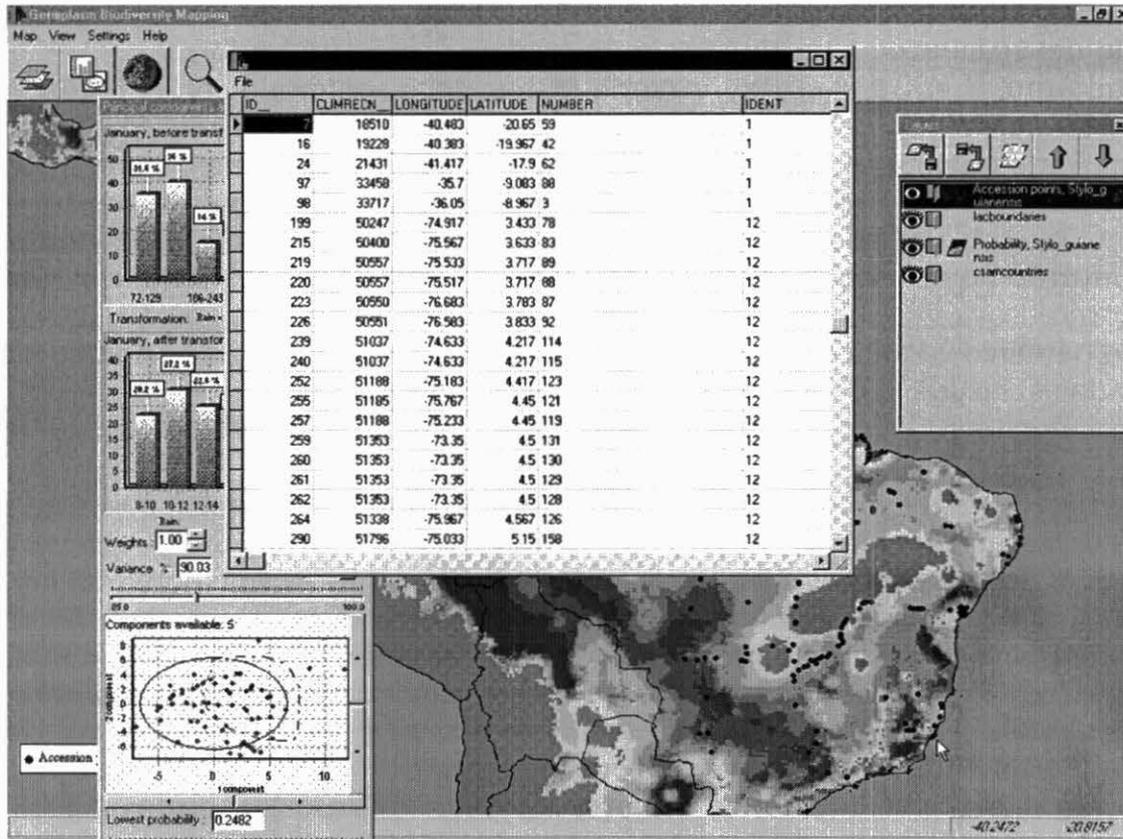
La herramienta para seleccionar el área siempre está disponible en la ventana ampliada del diagrama de dispersión, independientemente de que haya usted movido el icono.

Si usted ha activado la opción **show average climate for selected points** en el menú de configuración, el programa calculará el clima promedio y lo mostrará en el DIAGRAMA DEL CLIMA (ver página 53). El promedio se calcula en los datos climáticos que han sido rotados para asegurar que los registros del clima coinciden para el cálculo. Una vez el promedio se muestra en el diagrama del clima, usted observará que no existe la opción de **rotar/deshacer rotación** que aparece cuando se mira el clima de una sola accesión o punto del mapa. Esto sucede porque la rotación no tiene sentido cuando se aplica a un clima promedio. Los registros pueden ser procedentes de puntos geográficamente diversos y no existe, por tanto, ningún ángulo de fase correcto para rotar el promedio hacia atrás.



Aún si usted dibuja alrededor de un punto único, no tendrá oportunidad de deshacer la rotación de los datos. En este caso, simplemente señale la accesión con la herramienta del diagrama del clima.

A continuación, hemos movido alternamente el icono para seleccionar el área y hemos dibujado a pulso un polígono que encierra un conjunto de accesiones según aparecen en el diagrama de dispersión.



El polígono dibujado a pulso para encerrar el conjunto aparece como una línea roja punteada que gira alrededor de los puntos para llamar la atención. Los puntos seleccionados han cambiado de color verde a color azul oscuro, y la ventana de la hoja electrónica ha aparecido en pantalla. La hoja electrónica contiene ahora el conjunto de datos correspondiente a los puntos seleccionados. Se selecciona el punto de acesión con ID 7, y el punto en la costa brasileña, justo norte de Río de Janeiro hacia donde apunta el cursor, se ilumina en forma intermitente. Puesto que en papel no hay tiempo de respuesta, usted no puede ver esta iluminación intermitente aquí en la página.

Desplácese a través de los scatterplots y continúe por donde caen los puntos seleccionados. Es la manera ideal de lograr una idea exacta de la distribución de los puntos de acesión en un espacio multidimensional.

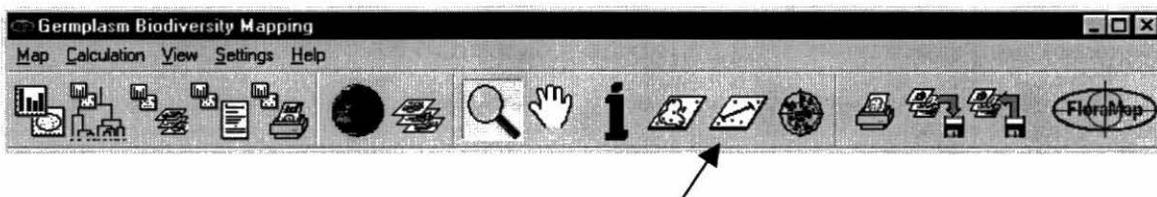
Con mover alternadamente el icono para seleccionar el área, usted puede hacer exactamente lo mismo para cualquier grupo de puntos en la ventana del mapa. Haga clic con el cursor en el mapa y dibuje alrededor de un conjunto de puntos. Observe que el análisis de PCA (PCA analysis) cambia constantemente. Los puntos en el scatterplot cambian de color en la medida que usted cambia el conjunto de datos. Continuamente se está actualizando y es completamente interactivo. Si le gusta la manera que se ve un subconjunto que ha elegido, puede usted regresar a la ventana de PCA y cambiar la transformación o las

ponderaciones, el marco de probabilidad, o el número de componentes.

Archivar la hoja electrónica es procedimiento estándar. Seleccione **file**, luego **save**, y asigne un nombre. Luego usted puede salir del análisis actual y recomenzar, utilizando el nuevo subconjunto. O, si no tiene tiempo para hacer este procedimiento, haga clic en el icono de PCA y el análisis de PCA se hará en el subconjunto que usted ha designado. Posteriormente usted puede alterar los parámetros del modelo (las transformaciones, las ponderaciones y el número de componentes). Si usted tiene aún menos tiempo y es muy impaciente, seleccione el icono de **visualizar mapas** y se calculará la superficie de probabilidad para el modelo del subconjunto escogido y se mostrará en el mapa. En este caso, usted tendrá que aceptar los parámetros del modelo que se fijaron en el análisis anterior.

Es fácil perderse entre las complejidades de las diferentes modalidades de selección. Usted puede escoger un subconjunto en el mapa, regresar al diagrama de dispersión, que ahora contiene solamente ese subconjunto, y escoger un nuevo subconjunto. Recomendamos que, a menos que usted solamente se está divirtiendo para ver que tan lejos puede llegar nuestro producto sin invalidarlo, debe hacer anotaciones minuciosas de lo que está haciendo y archivar con cierta regularidad la hoja electrónica con nombres que le permitan hacerle seguimiento durante el análisis. Si usted escoge las formas más rápidas de análisis y visualización, el mismo programa asignará un nombre a cada hoja electrónica, agregando un número de secuencia al final del nombre del archivo.

Es posible que usted quiera dibujar alrededor de un número de subconjuntos que le gustaría analizar en conjunto. Para hacerlo, seleccione y archive las hojas electrónicas por separado y luego haga una concatenación de éstas en su programa de hoja electrónica.

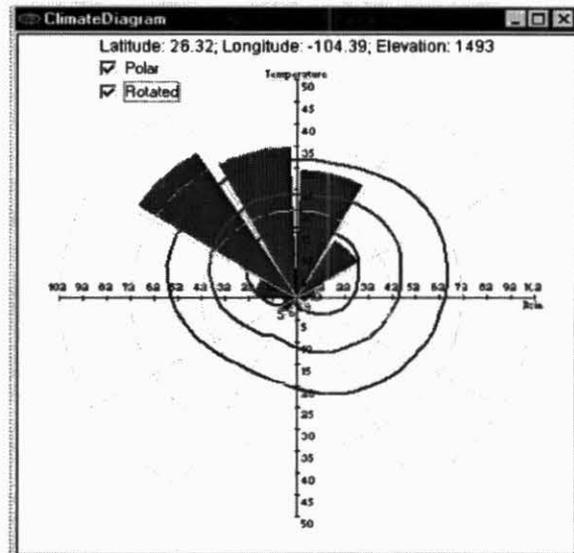
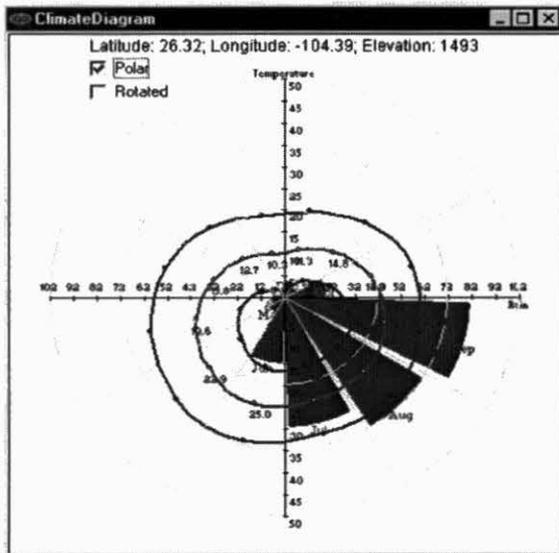
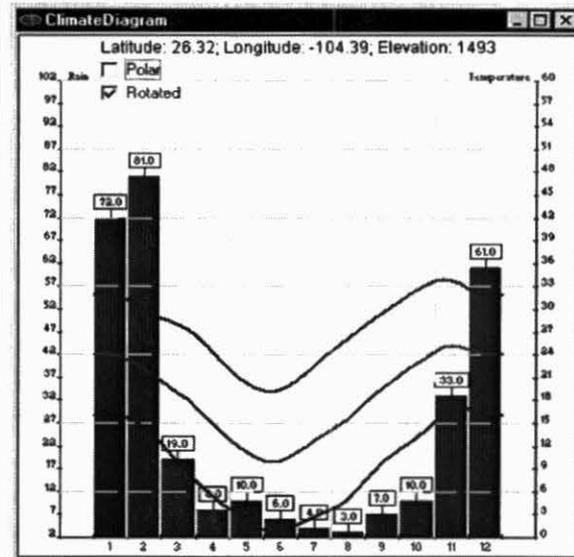
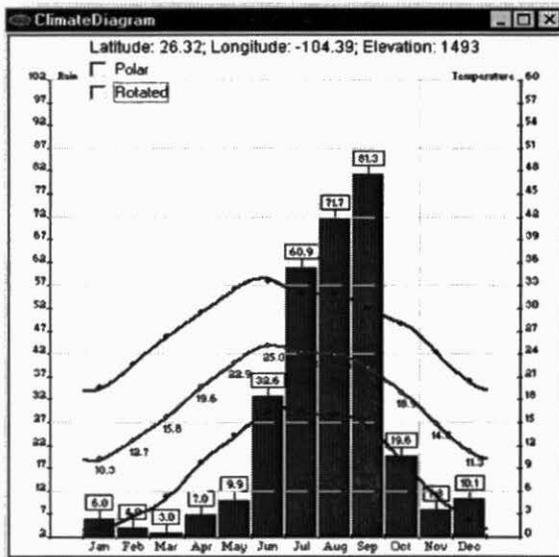


El icono para medir distancia en el mapa es especialmente útil cuando se está eliminando errores de un conjunto de datos de accesión. Mueva alternadamente el icono, posicione el cursor en un punto en el mapa, haga clic izquierdo y prolongue la línea para medir la distancia en el mapa en kilómetros. El dato de la distancia se indica en la esquina inferior izquierda de la ventana del mapa. Con este icono usted puede medir la distancia entre puntos imprecisos o la distancia entre una accesión y un elemento geográfico como un camino (ver página 21).



El icono del diagrama del clima es una herramienta útil para estudiar la estructura de un conjunto de accesiones, para eliminar posibles errores en las accesiones y hasta para explorar la base de datos climáticos. Mueve alternadamente el icono y señale cualquier punto de accesión o cualquier punto en el mapa. FloraMap buscará el registro de clima pertinente (es un poco lento haciendo esta operación con la presente versión de FloraMap—pero prometemos agilizar el proceso en la próxima versión).

Usted puede mover el diagrama del clima alternadamente entre cualquiera de las cuatro representaciones que se presentan a continuación.



Los dos diagramas a la izquierda muestran la precipitación mensual total y las temperaturas media, máxima y mínima en las coordenadas cartesianas y polares. El trazado polar comienza en la parte superior y corre en sentido contrario al de las manecillas del reloj, desde enero hasta diciembre. Los totales de precipitación corresponden cada una a una columna del histograma, y las temperaturas medias mensuales se muestran contra el trazado de temperatura media.

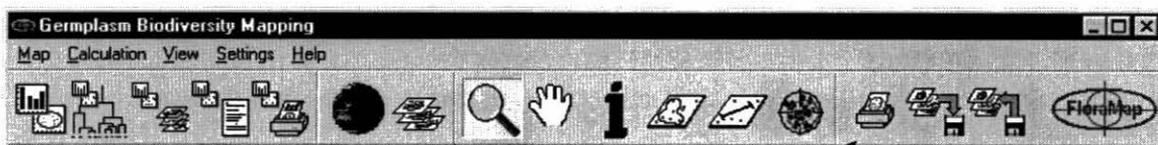
Los dos diagramas a la derecha muestran los datos que han sido rotados. No consiste en un sencillo cambio mensual (ver sección de TEORÍA, página 70), sino que es producido por una transformación Fourier especial que conserva los totales de cada período de 30 días y el promedio anual global. Dado que el ángulo de rotación puede ser de fracción de mes, los valores mensuales trazados cambian, como se observa en este ejemplo de las tierras altas de México central. Usted puede ver fácilmente la rotación que desplaza la época lluviosa entre junio y octubre hasta el comienzo del año virtual que se ha rotado.

No hay meses marcados en el año que se ha rotado porque el punto cero de la escala de tiempo es arbitrario.

Si usted ha activado la opción *show average climate for selected points*, un diagrama de clima aparecerá automáticamente cuando usted selecciona los puntos con la herramienta para seleccionar el área o en el diagrama de agrupación (*cluster diagram*). En ambos casos, la opción de rotación no aparecerá porque los climas promedios se han calculado utilizando los puntos cero virtuales de los datos que han sido rotados. El ángulo de rotación promedio no tiene significado alguno; por tanto, no hay ningún ángulo por el cual se pueda deshacer la rotación del registro. Existe perpetuamente en un espacio virtual que ha sido rotado.



No existe la opción de imprimir el diagrama del clima, pero si usted presiona alt-impr pant (alt-print screen) mientras la ventana está activada (es decir, cuando la barra de títulos está de color azul), la ventana se copiará en el portapapeles de edición. Usted luego puede pegar la ventana en un documento en Word, en una hoja electrónica, etcétera.



El icono para imprimir mapa envía el mapa, tal como aparece en la ventana del mapa, a una impresora que usted haya elegido entre los que están

disponibles. No imprimirá la pantalla que lo rodea o cualquier ventana superpuesta. El mapa consta de las capas activas en la ventana de capas, más cualquier información de leyenda que usted seleccione.



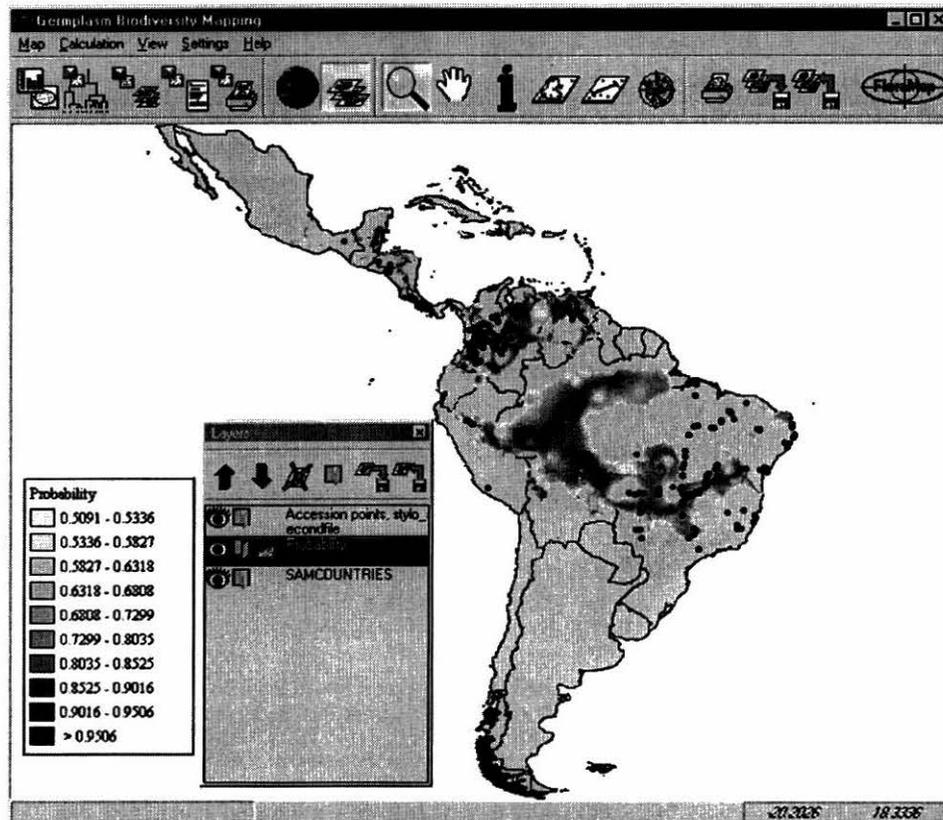
Diferentes impresoras manejan los colores de manera diferente. Casi nunca son como aparecen en pantalla. Prepárese para

experimentar y llegar a conocer los colores del menú, o de su menú de colores personalizado, que se imprimen bien en las impresoras que está usando.



El icono para guardar el mapa hace lo que indica la imagen—toma todas las capas que usted tiene activadas en layers window y las refiere a un MAP file en el directorio de trabajo. El archivo de MAP es un meta-archivo (ver APÉNDICE A). No contiene datos sobre la capa en sí, pero sí contiene información sobre cuáles capas se muestran con qué opciones de visualización para recrear el mapa. Cada capa del mapa se guarda de manera independiente como un shapefile en el directorio de su elección. Cada shapefile consta de tres partes: el archivo **DBF**, el archivo **SHP** y el archivo índice **SHX**.

A continuación presentamos un ejemplo. El mapa tiene las tres capas



—`stylo_secondfile`, probabilidad y `SAMCOUNTRIES`. El mapa se ha guardado en el directorio `\transfer\`. El resultado es el siguiente conjunto de 10 archivos.

Contents of 'transfer'					
Name	Size	Type	Modified	Attributes	
SAMCOUNTRIES	110KB	DBF File	8/23/99 9:24 AM	AC	
SAMCOUNTRIES.shp	3,648KB	SHP File	8/23/99 9:24 AM	AC	
SAMCOUNTRIES.shx	29KB	SHX File	8/23/99 9:24 AM	AC	
stylo_secondfile_01acp	74KB	DBF File	8/23/99 9:24 AM	AC	
stylo_secondfile_01acp.shp	12KB	SHP File	8/23/99 9:24 AM	AC	
stylo_secondfile_01acp.shx	4KB	SHX File	8/23/99 9:24 AM	AC	
stylo_secondfile_01prb	1,616KB	DBF File	8/23/99 9:24 AM	AC	
stylo_secondfile_01prb.shp	2,176KB	SHP File	8/23/99 9:24 AM	AC	
stylo_secondfile_01prb.shx	129KB	SHX File	8/23/99 9:24 AM	AC	
trial 23 aug	2KB	MAP File	8/23/99 9:24 AM	AC	

El archivo MAP 'trial 23 aug.map' (*mapa de prueba del 23 de agosto*) esta acompañado por los tres shapefiles, cada uno con tres componentes.

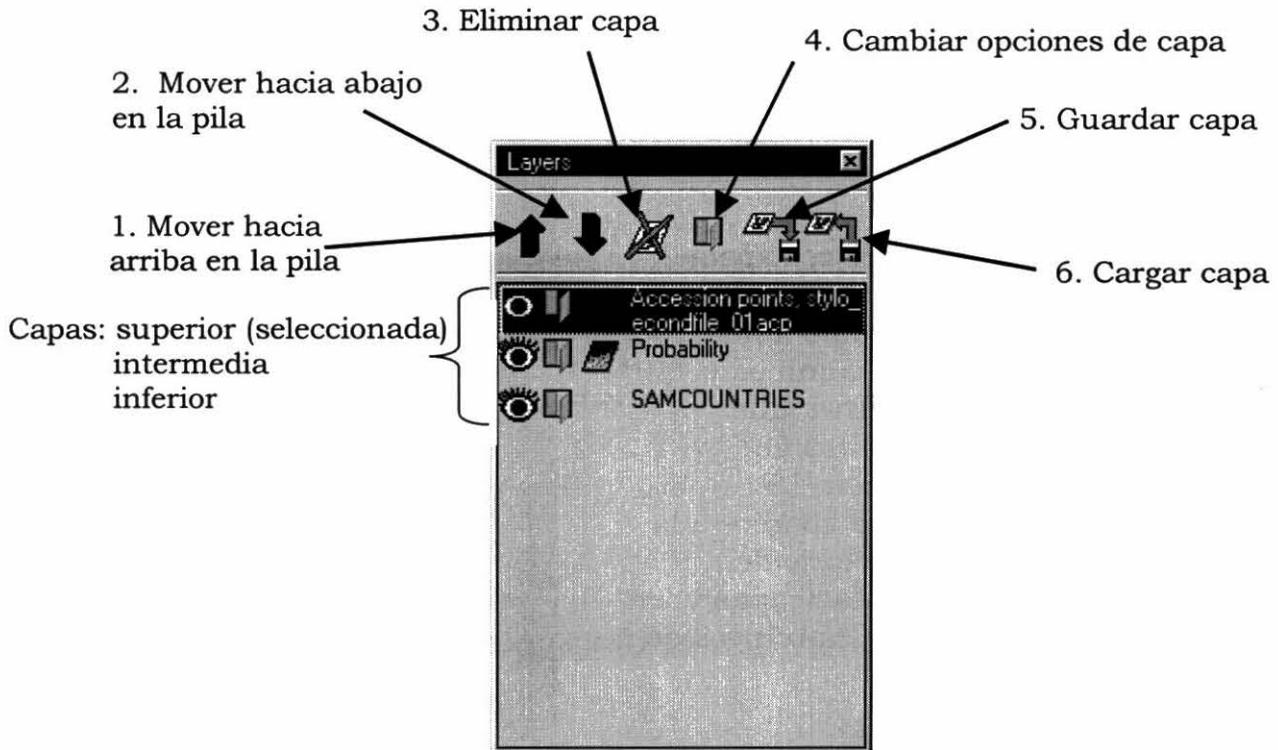
Usted no puede utilizar el **icono para cargar mapa** hasta que se haya creado un .MAP file. Un .MAP file viene con el sistema, es el `default.map`. Si usted ha activado la opción `autosave map` en la ventana de configuración, el mapa que usted tiene en el momento de salirse de FloraMap será guardado en este MAP file.



¡IMPORTANTE! El mapa que usted carga a partir de un MAP file consta de solamente los shapefiles que crean la imagen. El conjunto de accesiones que usted vea en el mapa (si tiene uno) es solamente la representación gráfica. Si usted trata de calcular un **PCA** a partir de este mapa, el sistema le pedirá cargar un conjunto de accesiones. Esto ocurre porque no hay un archivo de clima asociado con el mapa. Para proceder con el análisis, elimine la capa de puntos de accesión del mapa y seleccione nuevamente el archivo de puntos de accesión. FloraMap entonces procederá a construir el respectivo archivo de clima.

La Ventana de Control de Capas

Este es un menú de control de capas (layers control menu). El menú que aparece a continuación especifica tres capas. Puede haber las capas que se consideren necesarias, pero la sobresaturación crea confusión en el mapa.

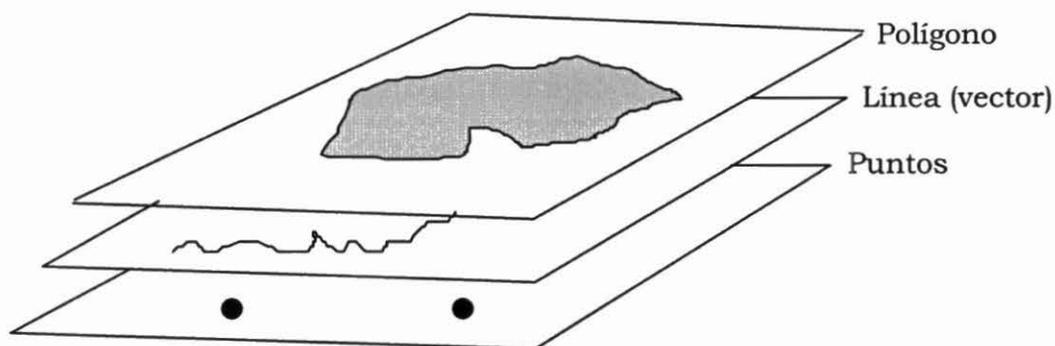


- Icono 1 mueve la capa seleccionada hacia arriba en la pila.
- Icono 2 mueve la capa seleccionada hacia abajo en la pila.
- Icono 3 elimina la capa seleccionada del mapa, pero no del directorio.
- Icono 4 cambia las opciones para la capa seleccionada.
- Icono 5 guarda la capa.
- Icono 6 carga una capa.

La pila controla lo que se ve en el mapa y también su visualización. Las capas se colocan en el mapa de abajo hacia arriba. Las capas superiores oscurecen las capas inferiores. Se puede cambiar el orden de las capas utilizando los iconos de flecha o haciendo clic y halando la capa hasta ubicarla en el lugar indicado dentro de la pila.

Las capas del polígono que tiene relleno de color oscurecerán los datos de línea o de punto por debajo de ellas (ver diagrama a continuación). Asegúrese de que su mapa está correcto al ordenar las capas. Una nueva capa de probabilidad siempre aparece como la capa superior. Si usted no tiene en pantalla la ventana de control de capas, haga clic en el icono de capas y reordene

los puntos de acceso de tal manera que usted los pueda visualizar.



Cargar una capa con el icono para cargar capas abarca cargar tanto las capas pasivas del mapa (como fondo, límites de los países, ríos, caminos y ciudades) como la capa activa, que consta del conjunto de puntos de acceso que se va a analizar. La visualización del menú que aparece a la derecha le indicará el directorio de trabajo donde se archivan las capas y, en la parte inferior, los tipos de archivos que usted puede cargar.



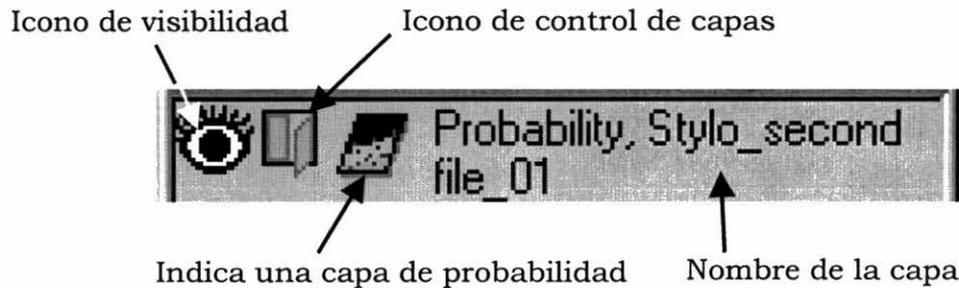
Un accession points file puede tener un extensión .dbf o .acp. Los archivos DBF están en formato dBASE 4. Los archivos ACP son archivos ASCII delimitados por espacios con encabezamientos de columna. Para información adicional sobre archivos de FloraMap, ver el APÉNDICE A.

Los shapefiles están compuestos de tres archivos con extensiones .shp, .shx, y .dbf. Son archivos estándares ESRI y son compatibles con ArcView.



¡SUMAMENTE IMPORTANTE! FloraMap no tiene manera de diferenciar un archivo de puntos de acceso en formato DBF de un archivo DBF de un shapefile o de un archivo del clima. Si el archivo contiene columnas para latitud y longitud, el archivo puede cargarse como un archivo de puntos de acceso. Dado que muchos shapefiles son muy extensos, esto puede constituirse en un grave error.

Personalización de las capas del mapa

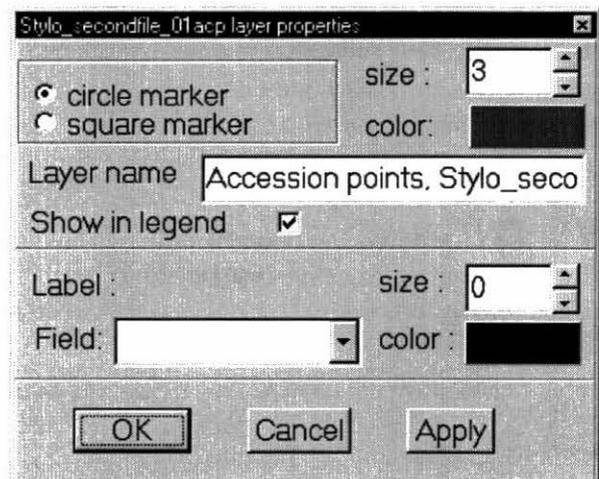


El **icono de visibilidad** es útil cuando se elabora un mapa. Puede suprimir o mostrar una capa en respuesta a un doble clic. La capa lógicamente permanece conectada al mapa y, por tanto, no necesita ser borrada y recargada.

Haga doble clic en la ventana abierta en el **icono de control de capas** en la capa seleccionada y usted abre una ventana para seleccionar las características de esa capa en el mapa. El primer ejemplo es **accession points layer**. Usted puede elegir el tipo de marcador (**marker**) y su color y tamaño (**size**). Haga clic en el color mostrado y usted abrirá el menú de colores (**color menu**). O puede elegir de la paleta de colores que se ofrece, o puede usted mismo combinar para lograr un color personalizado. Recuerde que los colores de una impresora de color no serán iguales a los que usted ve en pantalla, de manera que usted necesitará algo de práctica con esto.

El nombre de la capa (**layer name**) es el que aparece en la leyenda. El nombre es asignado por defecto a partir del nombre de archivo de esa capa, pero en el mapa se puede cambiar las palabras a las que usted desea para la leyenda. Si usted desactiva la opción **show in legend**, entonces no se hará referencia a la capa en la leyenda.

El uso de etiqueta: campo: (**label: field:**) a veces puede ser útil para verificar elementos individuales en el mapa, pero el mapa generalmente



contiene demasiada información para poder rotular un mapa de salida. El sistema asignará una etiqueta a cada característica individual en el shapefile. Esto puede funcionar para puntos de acceso trazados dispersamente. Sin embargo, en el caso de etiquetas de un shapefile de país, se asigna etiquetas no solamente al país, como se puede pensar, sino también a cada isla pequeña, ensenada y lago.



Desafortunadamente, los campos de etiqueta aumentan con el mapa en la medida que éste se amplíe; en la medida en que usted hace un acercamiento, los campos no se vuelven menos llenos, solo más grandes. Trataremos de resolver este problema en las versiones futuras.

Una vez que usted haya producido una superficie de probabilidad, FloraMap lo trazará sobre cualquier mapa que usted tenga en la ventana del mapa. Mueve alternadamente las capas hacia arriba y hacia abajo utilizando las flechas en la ventana de capas, o colóquelos en su lugar con el icono de mano en el cursor.

Una superficie de probabilidad es diferente a la superficie de fondo que hemos visto anteriormente. Tiene muchos valores. En efecto, cada pixel puede tomar cualquier valor del mínimo que usted especificó en la pantalla de PCA hasta el valor 1.0. Por consiguiente, el color de la superficie será una gama de colores para denotar el intervalo de probabilidades. Usted puede cambiar los colores y el número de niveles en el ámbito.

La opción **size** no tiene significado para una superficie de probabilidad. Si el relleno se cambia a transparente (**transparent**) no se verá. **Label: field:** no debe activarse ya que tiene un efecto desastroso sobre el mapa.

Los campos de valor (**value**) y pintar desde y hasta (**render from and to**), se fijan automáticamente en la probabilidad más baja que usted solicitó con el control de probabilidad (**probability control**) en la ventana de PCA. Muy raramente será necesario reasignar valores a estos campos.

Style_secondfile_01prb layer properties

solid fill size : 3
 transparent fill color: [black]

Layer name: Newname
Show in legend:

Label: size : 1
Field: <none> color: [black]

Render	from	to
Value	0.5003	1
Color	[white]	[black]

Quantity of breaks: 20

without outline outline color: [black]
 with outline

OK Cancel Apply

El hardware que usted está usando puede restringir el número de niveles que usted elige. Un controlador de pantalla de 256 colores limita el número de niveles que son únicos. En la actualidad, el sistema presenta, desafortunadamente, una limitación artística. No se puede definir una

secuencia de color que pase de un color por un segundo a un tercero, lo cual sería una técnica sumamente útil para mostrar en detalle el rango de probabilidad. Por ejemplo, la secuencia negro, café, rojo, amarillo, blanco no puede definirse todavía, aunque es una secuencia natural llamativa para ver. Estaremos trabajando para incluir esta flexibilidad en futuras versiones. Actualmente, usted puede especificar los puntos finales y la secuencia de color asignada será la manera más sencilla de llegar desde el primero color hasta el segundo.



NOTA: Los contornos de niveles en el ámbito que usted ve en el mapa en pantalla no son líneas reales. Los píxeles en la superficie siguen manteniendo sus valores originales, a diferencia de algunos sistemas de información geográfica que requieren la reclasificación de la imagen para fijar el rango de clases. Una consecuencia de esta situación es que los límites del ámbito no se comportan como polígonos. Si usted active outline en este menú, usted obtendrá el marco de cada pixel, no los polígonos del nivel de probabilidad. Si usted desea convertirlos a vectores, entonces tendrá que exportar la capa a otro sistema que puede manejar la conversión trama-a-vector.

La Ventana de Análisis de Componentes Principales

No se puede cambiar el tamaño de esta ventana. Si falta una parte en la parte superior o inferior, entonces su pantalla no reúne la configuración mínima para esta aplicación (ver PONERSE EN MARCHA).

Principal components analysis

January, before transformation

Bin Range	Percentage
41-87	5.7 %
132-178	10.5 %
269-314	14.5 %
405-451	19 %
542-587	20.9 %
	16.2 %
	6.4 %
	2.1 %
	1.7 %
	2.1 %
	0.5 %
	0.5 %

Transformation: (RainValue) 0.50

January, after transformation

Bin Range	Percentage
6.4-7.8	0.7 %
10.7-12.2	5 %
15.1-16.5	7.4 %
19.4-20.8	8.6 %
23.7-25	15.4 %
	21.1 %
	14.7 %
	14 %
	7.8 %
	1.4 %
	2.9 %
	0.5 %
	0.5 %

Weights: Rain: 1.00 Temperature: 1.00 Diurnal: 1.00

Variance: 95.2 Scores N: 6

Components available: 6

2 component

1 component

Lowest probability: 0.5003

0.0005

1

Barra de títulos - ventana de tamaño no variable

Barra de desplazamiento del mes del histograma

Control de la transformación

Control de las ponderaciones

Control de la variancia

Control de los componentes

Barra de desplazamiento del diagrama de dispersión

Control de la probabilidad

Histograma antes de la transformación

Histograma después de la transformación

Diagrama de dispersión

Los dos histogramas nos muestran los datos de precipitación para el conjunto de datos de los puntos de acceso. Estos son los diagramas de frecuencia para cada mes antes de la transformación y después de ella. Utilice la barra de desplazamiento del mes del histograma a la derecha para seleccionar el mes. Haga clic en el histograma y hale hacia abajo y a la derecha para crear una ventana de ampliación con el cursor.

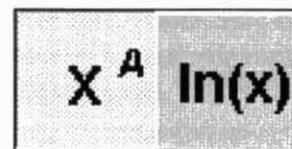


Haga clic en cualquier parte de la imagen del histograma y asegure el cursor sobre ella haciendo clic derecho. Usted ahora puede girar con el ratón para una toma panorámica y desplazarse hasta el detalle de su elección, en la dirección que desea. Para regresar a la imagen original, haga clic en el histograma con el botón izquierdo del ratón, creando una ventana de zoom, y hale hacia arriba y a la izquierda, al contrario de lo que usted hizo al comienzo.



La transformación es crítica para el análisis. Todo el análisis depende de que las variables se distribuyen como variables NORMALES. Las temperaturas generalmente se distribuyen normalmente con una curva bellamente simétrica en forma de campana...pero la precipitación casi nunca. En efecto, la mayoría de los estudios lo consideran como una distribución gamma. Usted puede ver su forma característica en las cifras de enero en este ejemplo. Se sesga en gran medida hacia los valores más pequeños. No podemos usar una distribución gamma en FloraMap debido a los cálculos de probabilidad que estaremos haciendo más adelante. Por consiguiente, debemos transformar la distribución para que sea lo más normal posible.

Haga clic en cualquier parte del control de la transformación entre los dos histogramas y la aplicación le dará la opción de transformaciones que aparece a la derecha. Las dos transformaciones básicas son logaritmo natural y transformación exponencial. Escoja entre las dos haciendo clic en el menú. Si usted eligió logaritmo natural, no hay que hacer nada más.



Si usted eligió la transformación exponencial, entonces tendrá que especificar el exponente en la pequeña ventana facilitada. El intervalo permitido es de +3.0 hasta +0.1, luego de -0.1 hasta -3.0. Esto le da un rango amplio de exponentes de los cuales elegir. Va desde x al cubo hasta el recíproco de x al cubo. Pasa por alto el intervalo +0.1 hasta -0.1 porque en ese intervalo los datos se transforman casi al valor fijo 1.0. Gran parte de este intervalo produce transformaciones bastante excepcionales, pero es útil tener la opción. Para comenzar, una buena transformación para ensayar es la de raíz cuadrada, o sea x a la potencia 0.5. Elija la que genera la curva con la campana mejor formada en la ventana después de la transformación (after-transformation). Tenemos que usar la misma transformación para los 12 meses. Por tanto, haga este ejercicio cuidadosamente con cada mes y escoja la transformación que hace el mejor trabajo en los meses con las peores distribuciones. Rara vez

encontrará una transformación que sea óptima para todos los meses; por tanto, debe usarse la que comporta mejor.

Usted puede alterar las ponderaciones que se aplican a las variables climáticas usando el control de ponderaciones. Mueve alternadamente los botones y observe cómo cambian las ponderaciones. Todos las ponderaciones deben sumar 3. Por tanto, si usted modifica una de ellas, las demás cambiarán. El algoritmo para asignar el cambio a las ponderaciones restantes no es siempre predecible. Si usted tiene dificultad en lograr la ponderación que necesita, introduzca los números directamente en las ventanas.

Observe cómo el *scatterplot*, en la parte inferior de la pantalla, cambia en la medida que usted cambia las ponderaciones. Más adelante discutiremos las implicaciones estadísticas de este cambio. Por ahora, ¿con qué velocidad se mueve? Está recalculando el análisis de los componentes principales en la medida que usted cambia las ponderaciones. La velocidad con que el programa hace este nuevo cálculo depende de su computadora (ver PONERSE EN MARCHA).

El número de puntuaciones de los componentes (*component scores*) es una consideración importante para su análisis (ver sección de TEORÍA, páginas 73-74). Los componentes se encuentran organizados en orden descendente de variancia y generalmente los dos o tres primeros explican una parte relativamente grande de la variación general, pero ¿**cúantos y en qué** proporción?

Recoja la palanca de control de la variancia (*variance control*) con el cursor y muévela. La variancia cambia en la ventana ubicada encima, al igual que el número en control de los componentes (*components control*), lo que indica la variancia que es explicada por el número de puntuaciones de los componentes que usted ha seleccionado. Una manera alterna es mover hacia arriba o hacia abajo los botones que se encuentran al lado de la ventana de *components control*, o fijar usted mismo los números en las ventanas.

El *scatterplot* constituye una parte sumamente importante del análisis (ver la sección de TUTORIAL y también la de TEORÍA). Hay diagramas $N.(N-1)/2$ cuando usted tiene puntuaciones N . Por tanto, hay tres con 3 puntuaciones, seis con 4 puntuaciones y 45 con 10 puntuaciones. Todas le suministran información. La elipse se fija en dos desviaciones estándares del origen de los dos componentes mostrados. Usted puede mover alternadamente entre los diagramas de dispersión utilizando las cuatro barras de control del diagrama de dispersión. Las dos barras inferiores lo llevarán hacia arriba y hacia abajo en el componente mostrado en la escala x ; las dos a la derecha lo llevarán entre los componentes en la escala y .

Agrandamiento del diagrama de dispersión

El diagrama de dispersión es una herramienta tan útil para investigar las propiedades de los conjuntos de puntos de accesión que hemos proporcionado la opción de agrandar la ventana para ver los puntos más claramente. Haga desplegar el menú de cálculo (calculation menu) y luego haga clic en agrandar diagrama de dispersión (enlarge scatterplot). Una vez que el diagrama de dispersión se haya duplicado en la ventana grande, la herramienta para seleccionar el área es seleccionada automáticamente, aunque la ventana del mapa detrás de ésta queda bloqueada. Usted puede escoger los conjuntos de puntos y desplazarse a través de los componentes principales en la ventana agrandada sin afectar el conjunto que se encuentra en la ventana de PCA. Al seleccionar un conjunto de puntos, se muestra la hoja electrónica de estos puntos. Si usted tiene activada la opción de diagrama del clima, verá en pantalla el clima promedio para el conjunto seleccionado.

Cuando usted ha tomado su decisión respecto a un conjunto de puntos para estudio posterior, haga clic en **OK** y ese conjunto se convertirá en el conjunto seleccionado en la ventana de PCA y usted puede proceder a entrarlo en otra pasada del análisis. Usted notará que los puntos se tornan de color azul al seleccionarlos, pero si usted ha desplazado la ventana agrandada a una combinación diferente de componentes, entonces igualmente tendrá que desplazar la ventana pequeña del diagrama de dispersión. Si usted cierre el diagrama de dispersión grande con el botón **X** de Windows o con el botón de cancelar (cancel), nada cambiará.

El **probability control** fija la probabilidad más baja que usted quiere que se trace. Esto es importante cuando usted está construyendo un mapa, especialmente si quiere mostrar más de una superficie de probabilidad, porque la superficie de probabilidad es una capa muy colorida y no se pueden visualizar otras características a través de ella. Entre más baja sea la probabilidad que usted escoge para mapear, más completo será el embaldosado del mapa con la superficie. Escoja un valor ya sea digitándolo en el recuadro o moviendo el botón de control hacia la izquierda o hacia la derecha.

La Ventana de Agrupación

Anteriormente vimos cómo podíamos fijar subconjuntos de datos en forma manual, ya sea en el mapa mismo o en la ventana del diagrama de dispersión. A menudo existen buenas razones para pensar que los climas del conjunto de accesiones no son homogéneos (ver secciones de TEORÍA y de TUTORIAL). El dibujar alrededor de grupos en el scatterplot, utilizando el icono para seleccionar el área, es, a menudo, una manera tan buena como cualquier otra para fijar subconjuntos de datos. Si usted puede ver las agrupaciones, puede delinear conglomerados relativamente complejos donde un algoritmo numérico podría fallar. Sin embargo, usted está restringido a considerar dos dimensiones a la

vez. De hecho, aunque podemos visualizar el espacio tridimensional, pocas personas pueden visualizar cuatro dimensiones.

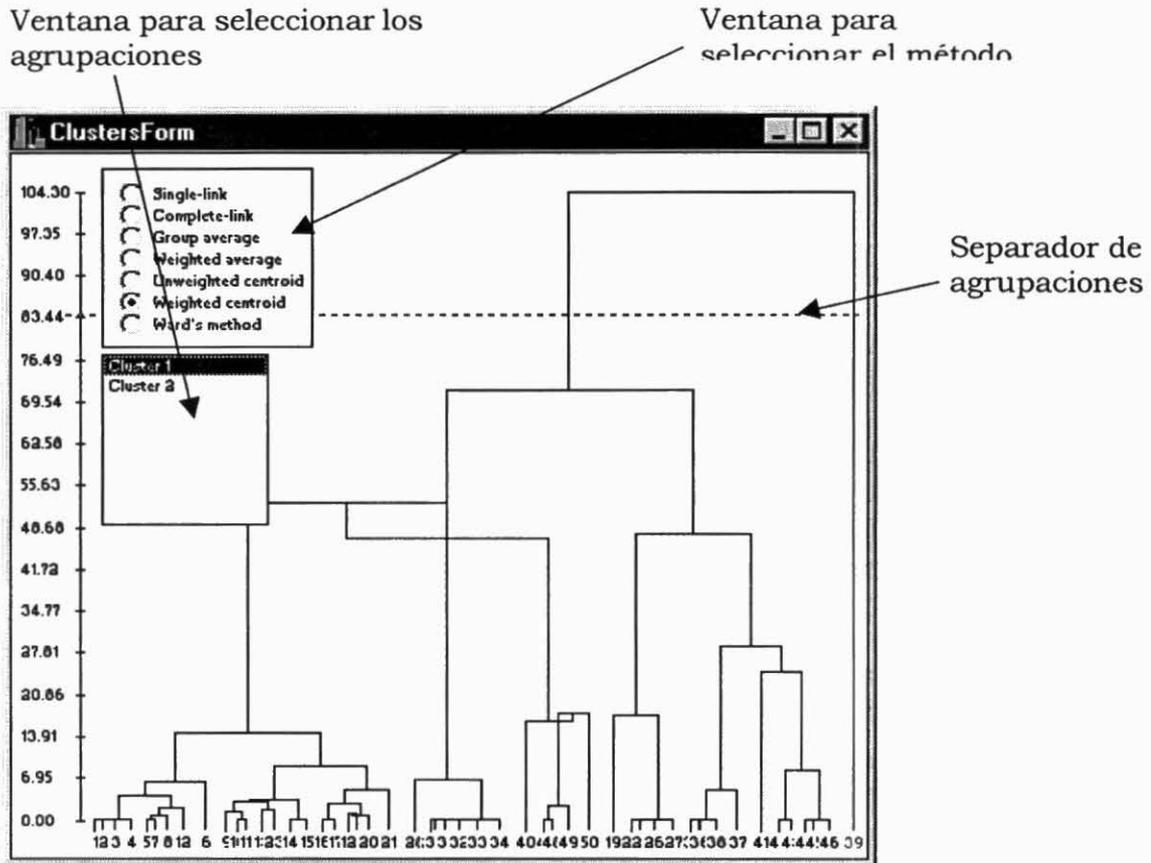
El análisis numérico de conglomerados es una herramienta potente (pero altamente subjetiva). Hemos incorporado un rango de métodos de análisis de conglomerados para ayudar con este problema (ver Análisis de Grupos, sección de TEORÍA).

Para activar la ventana de agrupación, selecciónala del menú desplegable de calculaciones (calculations), o mueva el icono de agrupación.

La presentación de los conglomerados es completamente interactiva con el PCA. En la medida que usted ajusta el PCA, la ventana de agrupación cambia hacia la agrupación.

Lo más importante de la presentación es un dendrograma que muestra las relaciones entre las accesiones. La escala a la izquierda es la distancia del conglomerado. La línea inferior no es una escala en el sentido usual de la palabra —muestra todo lo que podemos mostrar de las identidades de las accesiones. En la mayoría de los casos, el conjunto de accesiones será tan grande que estos identificadores son demasiados para mostrarlos en el diagrama completo. Por consiguiente, hemos facilitado una opción de cambio de visualización en el diagrama.

Coloque el cursor en el diagrama y hale hacia abajo y hacia la derecha mientras hace clic sostenido con el botón izquierdo del ratón. La parte encerrada del dendrograma se ampliará hasta llenar el marco. Normalmente sólo será útil de hacer zoom cerca para un conglomerado completo o parcial que tenga los identificadores de accesión visibles en la parte inferior del dendrograma. Sin embargo, usted puede hacer zoom cerca en cualquier área para mirar los vínculos del conglomerado si así lo desea. Para hacer zoom lejos, coloque el cursor en el diagrama y hale hacia la izquierda y hacia arriba mientras hace clic sostenido con el botón izquierdo. El diagrama volverá a su tamaño completo. La función de zoom es netamente para manipular la presentación. La función no selecciona accesiones para procesamiento adicional. Con zoom cerca de la presentación, usted puede girar para una toma panorámica de la zona del dendrograma al presionar sostenidamente la tecla de mayúsculas mientras mueve el cursor a través del área del dendrograma.



En el área de presentación hay dos ventanas del menú cuyo tamaño no se puede variar. Una de éstas es la ventana para seleccionar el método (method-select window), que está en la parte superior y sirve para seleccionar el método de agrupación. Actualmente hay siete métodos disponibles (ver sección de TEORÍA, página 85), pero solamente se puede seleccionar un método a la vez. Cada vez que usted seleccione un método diferente se hace un nuevo cálculo de la presentación. Ensaye los métodos diferentes y observe los resultados para su conjunto de datos. No existe un método correcto; cada uno tiene sus ventajas para diferentes conjuntos de datos.

Selección, análisis y archivo de las agrupaciones

La ventana para seleccionar las agrupaciones (cluster select window) se encuentra debajo de la ventana para seleccionar el método e indica el número de grupos que usted ha seleccionado. Para seleccionar los niveles de agrupación, haga clic izquierdo en el eje de distancia de la agrupación a la izquierda. Esto traerá la línea punteada horizontalmente, es decir la línea de selección del conglomerado (cluster-select line), hasta la punta de la flecha del cursor. Mueva el cursor hacia arriba y hacia abajo en la pantalla para seleccionar el nivel de agrupación que desea. El número de conexiones verticales cortadas por la línea punteada determina el número de agrupaciones

que se muestren en la ventana. Haga clic en uno de las agrupaciones dentro de la ventana para resaltarlo y seleccionarlo. Solamente puede seleccionar una agrupación a la vez. La agrupación seleccionada se resaltará en azul en el dendrograma. Todas las otras conexiones de la agrupación aparecerán en negro y los identificadores de accesión en rojo. Al mismo tiempo, observe el diagrama de dispersión. Los puntos para la agrupación seleccionada cambiarán a color azul. Usted puede desplazarse a través de los diagramas de dispersión para determinar cuál de los componentes está afectando a la agrupación.



NOTA: *La agrupación se calcula a partir de las variables transformadas escaladas. Los componentes principales son otra manera de visualizar las variables, pero cada diagrama de dispersión representa solamente la variación explicada por los dos componentes mostrados.*

Usted ahora tiene tres opciones para hacer análisis adicionales de la agrupación seleccionada.

- ✓ Haga clic en el icono para visualizar mapa y la superficie de probabilidad calculada a partir de los puntos de la agrupación se mostrará en el mapa. Aunque se calcula un nuevo PCA para la agrupación de puntos, la ventana de presentación del PCA no cambia.
- ✓ Haga clic en el icono de los componentes principales y se recalculará el PCA. Los histogramas y el diagrama de dispersión cambiarán para reflejar el nuevo PCA de los puntos de la agrupación. La superficie de probabilidad no se visualizará hasta que usted use nuevamente el icono para visualizar mapa.
- ✓ Haga clic derecho en el ratón con el cursor ubicado en cualquier parte del mapa y aparecerá un menú pequeño. Haga clic en **view dataset** y aparecerá una hoja electrónica con los datos de las accesiones de la agrupación seleccionada. Usted puede asignar un nombre a la agrupación y guardarla normalmente. Si más adelante desea proceder con el PCA y la elaboración del mapa, puede cargar la agrupación desde la ventana de control de capas. Esta opción es especialmente útil si desea analizar varias agrupaciones que están agrupados juntas. Guárdelos uno a la vez bajo diferentes nombres, utilice su programa de hoja electrónica para incorporarlos en un cuadro y cargue nuevamente este cuadro con el menú de control de capas.

4. Teoría

El sistema de FloraMap se basa en el cálculo de la probabilidad de que un registro de clima pertenezca a una distribución normal de múltiples variables descrita por los climas en los puntos de recolección de una serie de calibración de organismos. Fue diseñado para especies de plantas que se presentan naturalmente; su uso puede extenderse para cubrir el acontecimiento natural de cualquier organismo cuya distribución es determinada en gran parte por el clima. FloraMap utiliza una serie de superficies climáticas interpoladas, un método para calcular el modelo de probabilidad y un método para mapear las probabilidades de clima en la superficie climática.

Superficies Climáticas

Ahora están disponibles superficies climáticas interpoladas espacialmente para muchas áreas. Estas superficies generalmente manejan normales climáticas a largo plazo interpoladas sobre un modelo digital de elevación (MDE) mediante diferentes métodos (Jones 1991; Hutchinson 1997). El tamaño de pixel depende del modelo de elevación subyacente. Puede ser tan pequeño como 90

m (Jones 1996), lo cual da lugar a un extenso conjunto de datos, ó 10 minutos de arco (cerca de 18 km), que es tan grande como práctico en muchos casos. En este último, el modelo de elevación normal es el TGPO006 de la Administración Nacional Oceanográfica y Atmosférica (NOAA, su acrónimo en inglés) (NOAA 1984). En el CIAT hemos producido conjuntos de datos interpolados para América Latina y África, utilizando los datos de cerca de 10 000 estaciones en el caso de América Latina y 7000 en el caso de África. Cada serie de superficies consta de los totales de precipitación mensual, las temperaturas mensuales promedio y el intervalo de temperatura diurna promedio mensual, generando 36 variables de clima en tres grupos de 12.

Utilizamos un algoritmo sencillo de interpolación basado en el cuadrado inverso de la distancia entre la estación y el punto interpolado. Para cada pixel interpolado encontramos las cinco estaciones más cercanas. Luego se calculan las ponderaciones de distancia inversa y se aplican a cada valor mensual del tipo de datos que está siendo interpolado. Así que, para cinco estaciones con valores de datos x y distancias de la distancia de pixel d :

$$x_{pixel} = \sum_{i=1}^5 d_i^{-2} \times \sum_{i=1}^5 \frac{x_i}{d_i^{-2}} \quad (1)$$

Los datos de temperatura se estandarizan a la elevación del pixel en el MDE que usa un modelo *lapse rate* (Jones 1991).

El uso de esta interpolación sencilla conlleva varias ventajas. En primer lugar, es el más rápido de todos los métodos comunes. Segundo, la superficie interpolada atraviesa exactamente cada punto de la estación, porque la ponderación $1/(d(l)**2)$ se convierte en infinita en la medida que d se aproxima a cero. Tercero, la interpolación es sumamente estable en las áreas de datos dispersos. Se acerca al medio de las estaciones más cercanas en la medida que todas ellas se tornan igualmente distantes. Cuarto, es relativamente estable frente a errores en la elevación de las estaciones; solamente se afecta la región local de esa estación. Por otro lado, tanto la técnica spline laplaciana como la co-Kriging propagan estos errores más ampliamente. Esta es una de las ventajas de usar un modelo comprobado de *lapse rate* en vez de ajustar un modelo local, como lo hacen ambas de las últimas técnicas mencionadas.

El método de interpolación sencilla tiene dos pequeñas desventajas. En primer lugar, el derivado de la superficie se convierte en cero en la medida que pasa a través del punto de la estación. En otras palabras, cada estación se encuentra en una pequeña meseta o escalón de la superficie interpolada, lo cual es generalmente mucho más pequeño que el tamaño del pixel y, en consecuencia, no se puede apreciar. Segundo, se presenta un escalón (generalmente pequeño) en la superficie ajustada en la medida que las estaciones entran o salen de la ventana de ajuste. En aquellos casos en que la densidad de la estación es alta respecto al tamaño de pixel, es casi imposible visualizar. En aquellos casos en que estaciones no son tan densas, puede producir líneas rectas o arcos fluidos de aspecto desagradable en los datos ajustados de precipitación, los cuales no están vinculados con la elevación. La inspección del perfil de la superficie indica, generalmente, que éstos son unidades mínimas, pero tienen un aspecto desagradable y pueden minar la confianza en los mapas superficiales.

Estandarización de Fechas de Clima (Rotación)

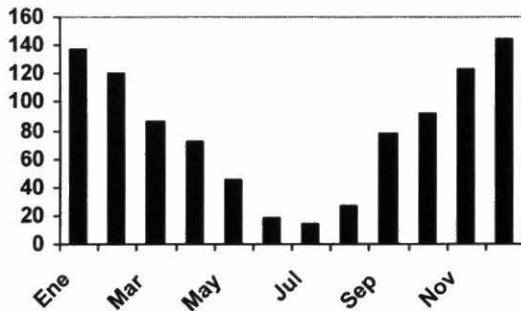
Los sucesos climáticos que ocurren a lo largo del año, como verano /invierno y comienzo /fin de la época lluviosa, son de importancia fundamental al comparar un clima con otro. Lamentablemente, en muchos de los tipos de clima ocurren en fechas diferentes. El caso más obvio es cuando se comparan los climas entre puntos en los hemisferios norte y sur, pero pueden verse diferencias más sutiles en la distribución temporal de sucesos climáticos en todo el trópico. Lo que necesitamos es un método para eliminar estas diferencias que nos permita hacer comparaciones libres de estos efectos anuales de época.

Miremos dos estaciones de clima hipotéticas. Se encuentran en un clima mediterráneo característico —inviernos húmedos calientes, veranos secos calientes. Northville podría estar en alguna parte de California, y Southville en Chile. La precipitación de agosto en Southville cae en enero en Northville. Si

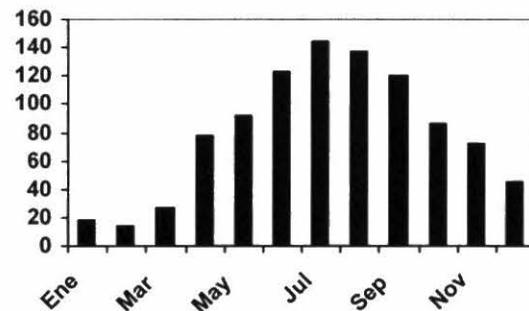
trazamos estas precipitaciones en las coordenadas polares, fácilmente podemos ver que necesitamos rotarlas a un tiempo estándar para poder compararlas.

Precipitación mensual para Northville y Southville

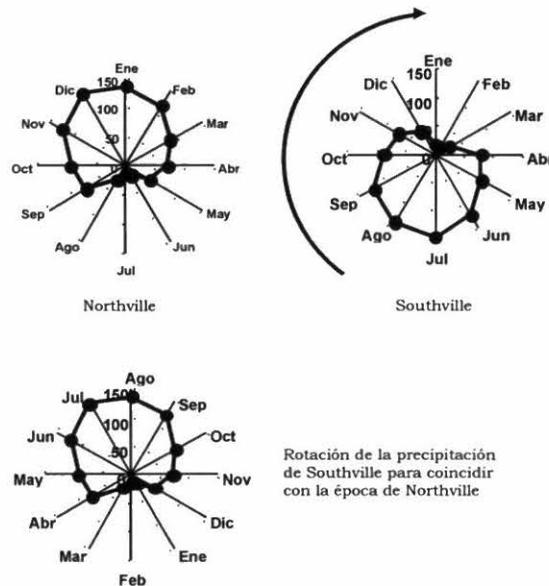
	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Northville	137	120	87	72	46	18	14	27	78	92	123	145
Southville	18	14	27	78	92	123	145	137	120	87	72	46



Mensual Northville



Mensual Southville



¿Cómo hacemos esta rotación en forma automática? La respuesta es la transformación Fourier de 12 puntos. Afortunadamente es el más sencillo de todos los algoritmos posibles de transformación Fourier. Es sumamente eficaz y rápido en términos computacionales. De hecho, es la base de casi todos los algoritmos rápidos de transformación Fourier que descomponen el problema de manera secuencial hasta llegar al caso sencillo de 12 puntos. Toma los 12 valores mensuales y los convierte en una serie de funciones seno y coseno. La transformación utilizada en FloraMap ha sufrido una modificación para que se

conservan los valores mensualmente totales (Jones 1987). La ecuación generada es:

$$r = a_0 + \sum_{i=1}^6 a_i \sin(ix) + b_i \cos(ix) \quad (2)$$

Puede reescribirse como una serie de vectores de frecuencia, cada uno con una amplitud α_i y un ángulo de fase, θ_i :

$$\alpha_i = \sqrt{a_i^2 + b_i^2} \quad \theta_i = \sin\left(\frac{b_i}{\alpha_i}\right) = \cos\left(\frac{a_i}{\alpha_i}\right) \quad (3)$$

Si restamos el primer ángulo de fase de todos los demás vectores en el conjunto, entonces hemos producido una rotación rígida de los vectores. Esta es la rotación que estamos buscando. La rotación ubica el valor máximo de la primera frecuencia en un ángulo de fase cero y los demás valores en posiciones equivalentes a su separación angular en los datos originales. Usamos, entonces, el primer ángulo de fase correspondiente a la precipitación para rotar los datos correspondientes al rango de temperatura y de temperatura diurna, y estas variables se rotan rígidamente junto con la precipitación.

Es obvio cómo funciona este algoritmo para los registros de clima en los

casos de precipitación unimodal. Pueden existir climas que son ambivalentes respecto a una primera rotación de frecuencias. En la práctica, casi nunca se encuentran; el único caso grave es cuando no se presenta precipitación en todo el año.

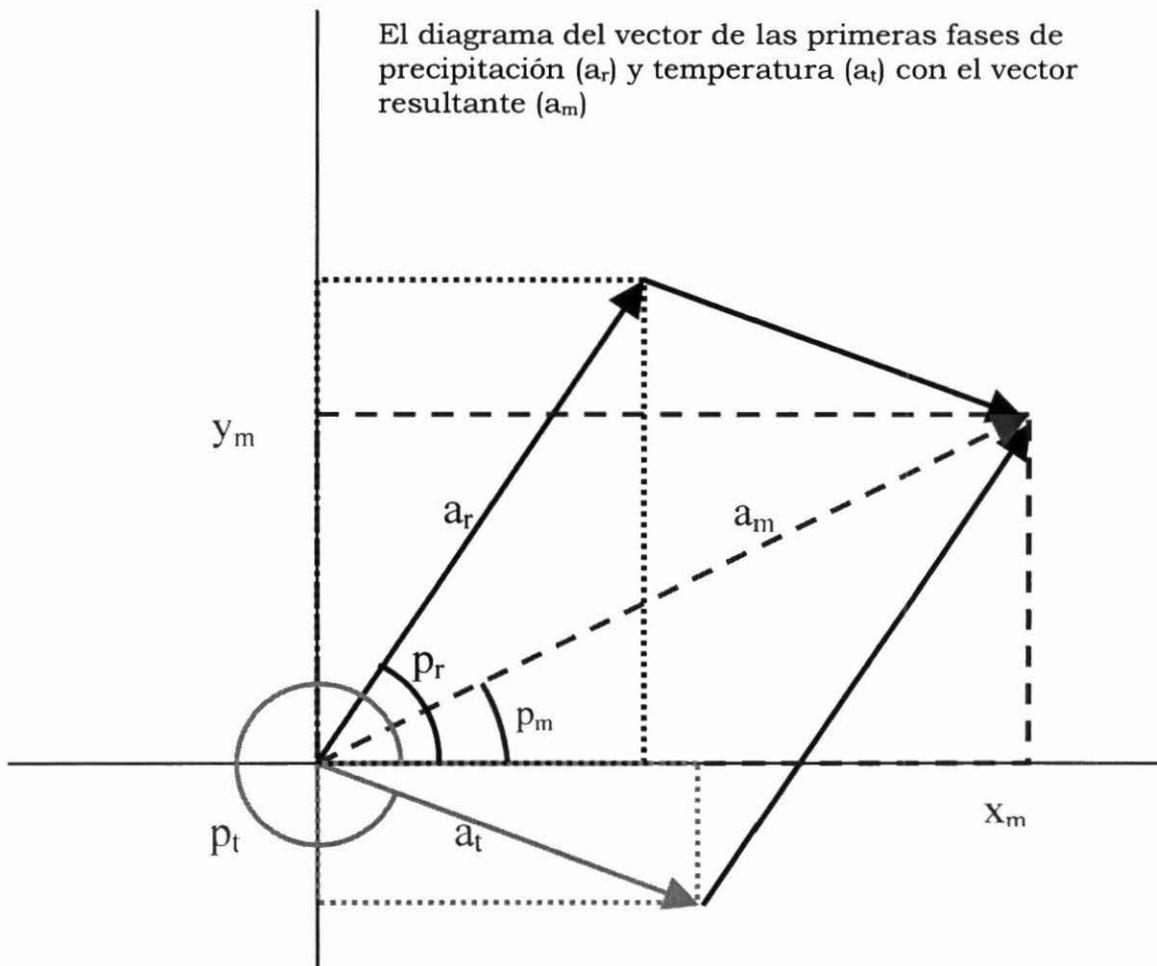
Esta explicación funciona bien para zonas tropicales y para otras partes de casi todo el mundo, según lo indica el manual de la versión 1.0 de FloraMap. Existía una leve probabilidad de que el procedimiento no funcionara si se ajustaba una accesión a un modelo con latitudes suficientemente altas como para presentar un clima tipo mediterráneo (como el que se utilizó en el ejemplo anterior). En caso de que algunas de las accesiones caigan en áreas de lluvias durante el invierno y algunas en áreas de lluvias intensas durante el verano (no Mediterráneo), el modelo resultante podría tener un ajuste bastante deficiente. Dado que esta situación es improbable en términos botánicos, es posible que aún no se haya observado en la práctica, aunque sí se presentó el caso al hacer ejecutar por el programa un conjunto simulado a manera de ensayo a través de los Andes en Chile / Argentina.

Dr. Ian Makin del Instituto Internacional para el Manejo del Agua (IWMI, su acrónimo en inglés) gentilmente nos permitió acceder al Atlas Mundial de Agua y Clima de ese Instituto para hacer las rejillas climáticas que permitieran ampliar el rango de FloraMap. Escogimos ensayar la rejilla correspondiente a

Europa porque tenemos usuarios potenciales que quieren examinar esta zona. Fue entonces que surgió el problema. La temperatura es, por lejos, el determinante de clima que predomina en Europa occidental. Los modelos de precipitación pueden ser invernal, estival, o indeterminado sobre distancias relativamente cortas.

Por consiguiente, tenemos la posibilidad de rotar la precipitación o la temperatura, pero ¿cuando se decide cuál es el factor dominante? Tratamos muchas combinaciones de reglas, pero lamentablemente llegamos a la conclusión de que ninguna era aceptable. Todas dieron como resultado una línea destacada que atraviesa el mapa en algún punto donde se cambió la base de rotación, lo que dio lugar a que climas, que debieron haber ido cambiando de un tipo a otro de manera gradual e imperceptible, de repente cambiaran frente a una discontinuidad. Los usuarios hubieran tenido serios problemas al ajustar modelos en estas zonas.

Se encontró que la mejor solución es utilizar TANTO la precipitación COMO la temperatura para calcular el ángulo de fase de la rotación. Por tanto:



El ángulo y la amplitud de fase resultantes son, entonces:

$$\begin{aligned}y_m &= a_r \cos p_r + a_t \cos p_t \\x_m &= a_r \sin p_r + a_t \sin p_t \\a_m &= \sqrt{y_m^2 + x_m^2} \\p_m &= \text{angle}\left(\frac{x_m}{a_m}, \frac{y_m}{a_m}\right)\end{aligned}$$

Lamentablemente, esto no resuelve completamente el problema de ajustar un modelo a climas con diferentes factores determinantes del tiempo. Sin embargo, la gran mayoría de los climas a nivel mundial son:

- (1) Determinados por la precipitación, donde la temperatura no constituye un efecto estacional importante (grandes áreas del trópico y del subtropico),
- (2) Determinados por la temperatura, donde la precipitación es uniforme a lo largo del año (la mayor parte del resto del trópico y algunos climas templados), o
- (3) Determinados por la precipitación y la temperatura, cuando los dos variables presentan una alta correlación (lluvias estivales, en la mayor parte del resto del mundo).

El caso diferente es:

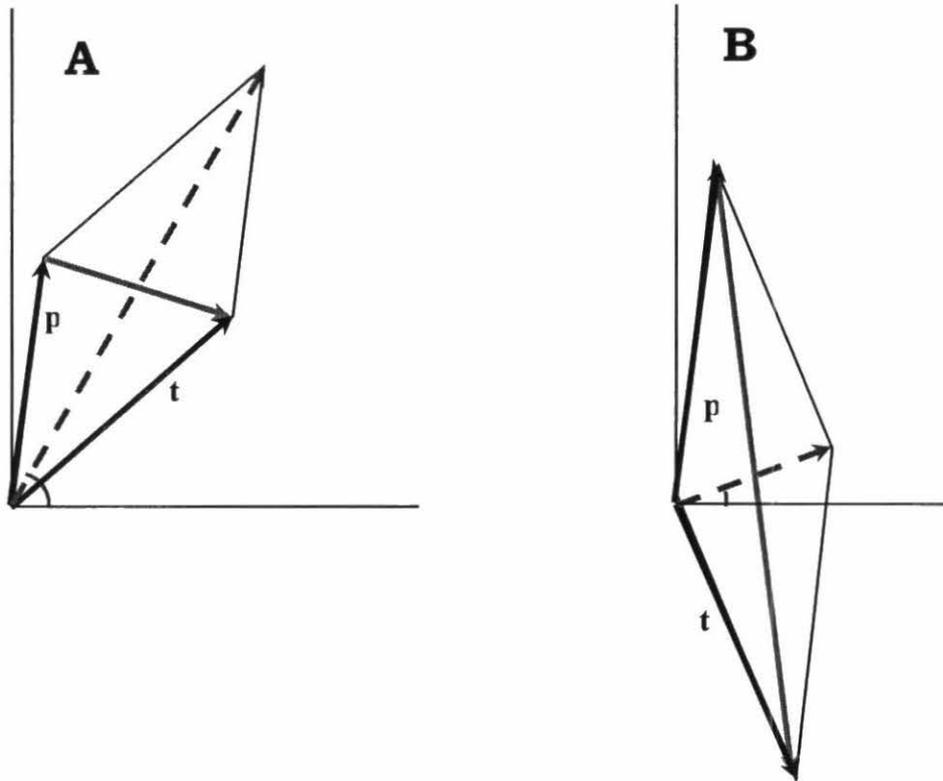
- (4) Lluvias invernales y veranos secos calientes (casi exclusivamente climas tipo mediterráneo).

Afortunadamente, los climas mediterráneos se encuentran en las latitudes moderadamente altas y podemos permitir que la rotación sea dominada por la temperatura, sin perder el carácter general de las rotaciones y las comparaciones. Por consiguiente, debemos aumentar paulatinamente la ponderación para el vector de temperatura en la medida que nos acercamos a los climas mediterráneos (para evitar una oscilación súbita).

Se encontró que las siguientes ponderaciones funcionan bien:

- p** = mm de precipitación
- t** = temperatura x 2 x abs(latitud)





Se puede presentar un obstáculo cuando dos vectores casi se cancelan entre sí, lo cual resultaría en oscilaciones fuertes del ángulo de rotación en respuesta

a cambios leves en los vectores de precipitación y de temperatura. Esta situación es más factible en la medida que la situación pasa de A (vea figura arriba) a B y más allá. Como antes, las flechas punteadas constituyen los vectores de rotación, pero calculados a partir de los vectores ponderados de precipitación y temperatura.

Cuando el vector de rotación es el vector suma de $\mathbf{r} + \mathbf{t}$, el vector contra-diagonal es la diferencia $\mathbf{r} - \mathbf{t}$. Es fácil detectar que las áreas peligrosas se presentarán cuando $\mathbf{r} - \mathbf{t}$ es mucho mayor que $\mathbf{r} + \mathbf{t}$. Por consiguiente, podemos utilizar un práctico índice de estabilidad, s .

$$s = \arctan\left(\frac{|\mathbf{r} - \mathbf{t}|}{|\mathbf{r} + \mathbf{t}|}\right)$$

Este índice sería igual a cero en el caso de estados estables, donde el ángulo de rotación es dominado por la precipitación, la temperatura o por ambos obrando en concierto. El índice se acercará a $\pi/2$ en la medida que los vectores tienden a cancelar sus efectos. Dado que podemos trazar este índice, podemos verificar la presencia de áreas donde puede presentarse esta rotación

indeterminada. Las zonas con un s relativamente alto (inestabilidad potencial) se presentan a lo largo de la Costa Pacífica estadounidense, en Chile, el nordeste de Brasil, Sri Lanka y en algunas zonas de África central. Sin embargo, el índice no llega a alcanzar los 80 grados en ninguna área. Aunque este índice parece ser alto, los ángulos de fase se rotan correctamente y, de hecho, es poco probable que haya una rotación falsa.

Si no está seguro de los ajustes del modelo al incluir accesiones de estas zonas, sírvase usar la herramienta *ClimateDiagram* para investigar la situación. En el caso de las rejillas de alta precisión, puede haber un pixel ocasional que rote de manera anormal y examinaremos esta posibilidad cuando creamos las rejillas nuevas. Sin embargo, para las rejillas actuales de *FloraMap* no habrá ningún problema.

Para ahorrar tiempo de computación, toda la superficie climática se rota según estas reglas y todas las operaciones en *FloraMap* se hacen dentro del espacio de la fase que ha sido rotada.



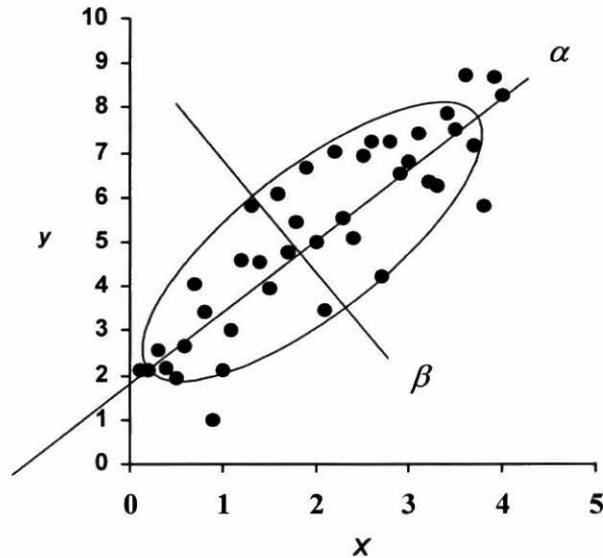
La única excepción es cuando el usuario solicita un diagrama de clima para un punto de acesión o un punto de la superficie climática.

Cálculos del Modelo

Una vez que hayamos rotado cada registro en la serie de calibración, estaremos listos para construir el modelo de probabilidad. Los datos con los cuales trabajamos tienen 36 variables de clima. Si usáramos todas estas variables en la forma en que se presentan, nos enfrentaríamos al reto de construir un modelo de probabilidad en un espacio que contiene 36 dimensiones. Aunque no es demasiado difícil para una computadora moderna, sí presenta problemas para el usuario cuando éste trate de visualizar lo que está sucediendo. La otra consideración es que pueden estar altamente correlacionadas todas las variables del clima, haciendo que el modelo de probabilidad sea aun más difícil de entender. Una manera de superar este problema es usar un análisis de componentes principales (PCA, el acrónimo inglés). Un PCA construye series de combinaciones lineales de las variables para maximizar la variancia en cada uno desde los datos originales. Estas combinaciones lineales tienen otra particularidad que es sumamente útil. Son ortogonales a sí mismos, completamente sin correlación y, por tanto, pueden ser manejadas independientemente o en conjuntos sin interacciones inesperadas.

La operación puede ilustrarse en dos dimensiones del siguiente modo. La figura a continuación muestra un diagrama de dispersión de dos variables x y y , altamente correlacionadas, y, por consiguiente, nada independientes. Para

cualquier cambio en x , esperaríamos un cambio en y . Sin embargo, podemos encontrar dos ejes nuevos, α y β , que no están correlacionadas, y se maximiza la variancia representada en el primero de los ejes nuevos. Observe que α no es la línea de regresión de y en x , y, por tanto, pasa directamente por el grupo de puntos.



En este caso, $\alpha = 0.454x + 0.891y$ y $\beta = 0.891x - 0.454y$. Estos ejes nuevos son ortogonales y no correlacionados. El movimiento a lo largo del eje α no implica movimiento alguno a lo largo del eje β . El componente α explica el 95.6% de la variancia original, β solamente el 4.4%. El truco de esta transformación lineal es calcular los valores específicos y los vectores específicos de la matriz de variancia-covariancia del sistema de variables. En el caso de FloraMap, es una matriz 36 x 36 de variables climáticas.

En la anotación de la matriz necesitamos encontrar una matriz \mathbf{Q} y una matriz diagonal Λ tal que:

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \text{diag } \lambda = \Lambda \quad (4)$$

donde \mathbf{A} es nuestra matriz de variancia-covariancia.

La matriz Λ , compuesto de los elementos λ , es la matriz diagonal de los valores específicos, que en nuestro caso contienen la variancia de los vectores específicos. La matriz \mathbf{Q} es una matriz simétrica, que contiene vectores específicos como filas y como columnas. Los vectores específicos tienen dos propiedades sumamente útiles, una de las cuales ya se mencionó —son

linealmente independientes los unos de los otros. La segunda propiedad útil es que un vector específico multiplicado por cualquiera escala es todavía un vector específico.

Para esta operación, la matriz de variación-covariancia no tiene que ser un rango completo. FloraMap se ajusta a un número tan reducido como tres puntos de calibración. Sin embargo, tal vez los resultados no sean fiables con conjuntos de accesiones de menos de 15 puntos.

El PCA puede realizarse a partir de la matriz de las sumas de los cuadrados y los productos cruzados (SSCP, su acrónimo en inglés), la matriz de variancia-covariancia, o la matriz de correlación de un grupo de variables. En FloraMap, usamos la matriz de variación-covariancia mediante la estandarización de las variables antes de calcular la SSCP. Pero diferimos de muchos análisis estándares en que nuestros datos tienen una estructura que deseamos preservar en vez de estandarizar completamente. Los datos son, en realidad, tres grupos de 12 valores para las diferentes variables climáticas — precipitación, temperatura e intervalo de temperatura diurna. Queremos conservar esta diferencia para permitir al usuario aplicar la ponderación en

todo el cuadro para variables climáticas, aumentando, por ejemplo, la importancia de la precipitación respecto a la de temperatura. Además, la información correspondiente a los 12 valores mensuales es de interés crucial y no queremos estandarizarlo de tal manera que desaparezca. Por consiguiente, estandarizamos todos los valores de precipitación por variancia común para precipitación y así sucesivamente.



Al momento de escribir la presente versión, estandarizamos cada variable mensual a un valor medio de cero. Estamos estudiando la posibilidad de también ofrecer la opción de estandarización utilizando solamente el medio del grupo, lo cual daría un ajuste más determinante a los efectos a lo largo del año. Sin embargo, en la actualidad, los efectos no deseados se trasladan a los valores del componente cuando se cambian las ponderaciones.

Una vez que hayamos encontrado Λ y Q , podemos describir el sistema de variables climáticas en función de los componentes principales y sus variancias (vectores específicos y valores específicos). Podemos escoger un subconjunto de los componentes (porque los vectores específicos son independientes), y podemos escalarlos individualmente (porque la multiplicación o la división por una constante no cambia las propiedades del vector específico). Este último punto es importante porque es exactamente lo que queremos hacer para calcular las probabilidades.

Cálculos de Probabilidad

La función de densidad de probabilidad normal para una variable única es representada por:

$$z = \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \quad (5)$$

A partir de la integral de esta función podemos calcular la probabilidad de observar un punto extraído de esta población. Tradicionalmente, miramos la probabilidad de que un punto pueda estar ubicado más allá del origen que el punto en cuestión. En general, también calculamos los parámetros de distribución de la muestra que estamos investigando. Debido a esto, usamos otras estadísticas tales como **t de Student** para calcular la probabilidad.



En FloraMap, hacemos una suposición simplificadora de que la serie de calibración de accesiones contendrá puntos suficientes para que el cálculo de la muestra sea equivalente a conocer los parámetros de población. Esto no será cierto para series pequeñas de calibración y, por tanto, las probabilidades calculadas no serán estrictamente exactas. Sin embargo, siempre y cuando el usuario reconozca este hecho, las probabilidades aún pueden usarse como un índice para el mapeo.

Para múltiple dimensiones con n variables independientes (ortogonal), la función de densidad de probabilidad se convierte en:

$$z = \frac{1}{\sqrt{2\pi}} e^{-(t_1^2 + t_2^2 + \dots + t_n^2)/2} \quad (6)$$

La integral de esta función puede obtenerse mediante la integración repetida, pero la especificación de los límites de integración para cada integración subsiguiente en términos de las funciones anteriores es desordenada y tediosa. A continuación presentamos una manera más fácil de considerarlo. Queremos la probabilidad de que cualquier punto en una distribución caiga dentro de un radio de:

$$r = \sqrt{(t_1^2 + t_2^2 + \dots + t_n^2)} \quad (7)$$

El volumen de una esfera de la dimensión n es:

$$\frac{r^n \sqrt{\pi^n}}{\Gamma\left(\frac{n+2}{2}\right)} \quad (8)$$

Observe que en la medida que aumenta n , el volumen de la esfera tiende ajustarse a cero. Por tanto, la integral de probabilidad construida en el espacio con n grande será pequeña contra-intuitivamente.

El volumen de una capa incalculablemente delgada de esta esfera en el radio r es:

$$\frac{nr^{n-1} \sqrt{\pi^n}}{\Gamma\left(\frac{n+2}{2}\right)} \quad (9)$$

El derivado de la integral de probabilidad en esta capa es:

$$z = \frac{nr^{n-1} \sqrt{\pi^n}}{\Gamma\left(\frac{n+2}{2}\right)} e^{-r^2/2} \quad (10)$$

Por consiguiente, la integral desde 0 hasta r es:

$$\frac{n\sqrt{\pi^n}}{\Gamma\left(\frac{n+2}{2}\right)} \cdot \int_0^r r^{n-1} e^{-r^2/2} dr \quad (11)$$

Si tomamos solamente la porción a la derecha del signo integral, y dividimos por el límite en la medida que r pasa a la infinidad desde la izquierda, tenemos para dimensiones pares:

$$\lim_{r \rightarrow \infty} \int_0^r r^{n-1} e^{-r^2/2} = \Gamma\left(\frac{n}{2}\right) 2^{(n-2)/2} \quad (12)$$

$$p = 1 - \frac{e^{-r^2/2} \left(r^{n-2} + (n-2)r^{n-4} + (n-2)(n-4)r^{n-6} \dots \Gamma\left(\frac{n}{2}\right) 2^{(n-2)/2} \right)}{\Gamma\left(\frac{n}{2}\right) 2^{(n-2)/2}} \quad (13)$$

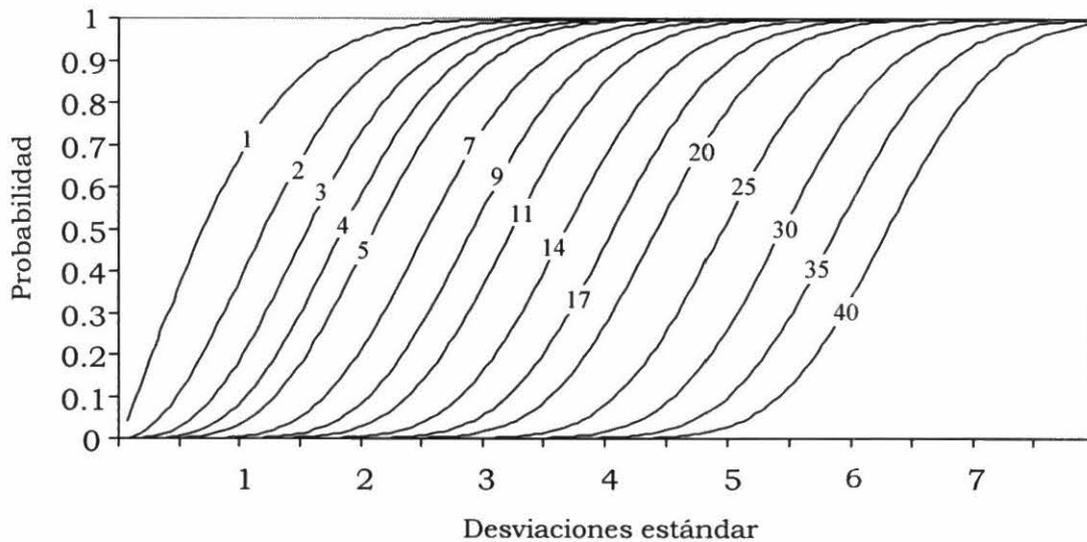
Al descomponer en factores, esto se convierte en:

$$p = 1 - e^{-r^2/2} \left(\frac{1}{(n-2)} r^2 + 1 \right) \left(\frac{1}{(n-4)} r^2 + 1 \right) \left(\frac{1}{(n-6)} r^2 + 1 \right) \cdots \left(\frac{1}{(n-(n-2))} r^2 + 1 \right) \quad (14)$$

Y, para las dimensiones impares, vamos descomponiendo en factores en la medida que vamos avanzando:

$$\lim_{r \rightarrow \infty} = \frac{\sqrt{2} \cdot \sqrt{\pi} \cdot (3)(5)(7) \cdots (n-2)}{2} \quad (15)$$

$$p = \operatorname{erf}\left(\frac{\sqrt{2}r}{2}\right) - \frac{\sqrt{2}r e^{1r^2/2}}{\sqrt{\pi}} \cdot \left(\frac{1}{(n-2)} r^2 + 1 \right) \left(\frac{1}{(n-4)} r^2 + 1 \right) \left(\frac{1}{(n-6)} r^2 + 1 \right) \cdots \left(\frac{1}{(n-(n-3))} r^2 + 1 \right) \quad (16)$$

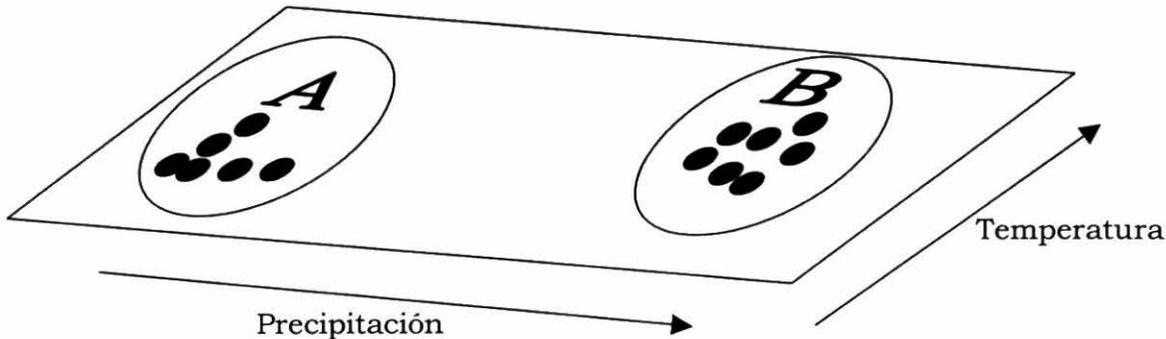


La integral de probabilidad en dimensiones múltiples: La probabilidad de encontrar un punto entre el origen y el radio r para las poblaciones $N(0,1)$ en las dimensiones seleccionadas de 1 a 40.

Este resultado es importante. Si no lo tuviéramos, no podríamos mantener el nivel correcto de probabilidad en la medida que pasamos de una serie de dimensiones a otra. Esto es efectivamente lo que hacemos cuando escogemos diferentes conjuntos de componentes principales.

Probabilidades Divergentes

Imagínese un plano con la precipitación que varía de la izquierda hacia la derecha y la temperatura que varía de adelante hacia atrás.

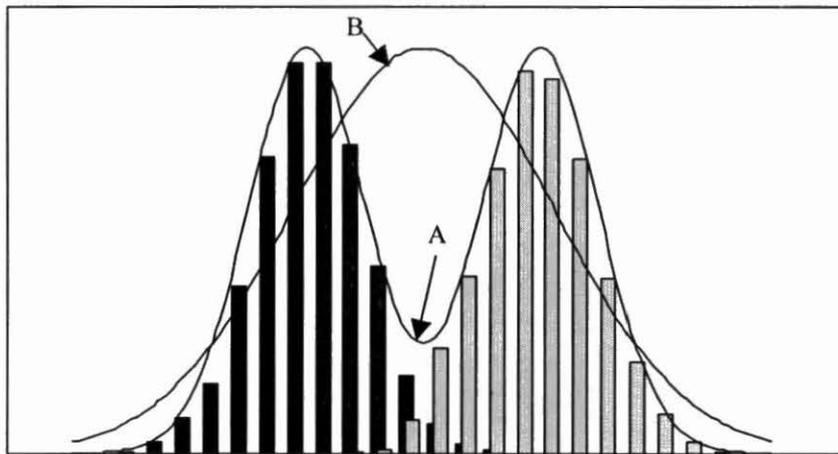


Los puntos en (A) experimentan un régimen fresco seco, mientras que los puntos en (B) experimentan un clima más húmedo y más caliente. Si suponemos que todos son puntos de accesión de la misma especie, no hay razón para no tratar de ajustar para ellos un modelo de probabilidad único. Si el caso es que la brecha entre ellos es netamente fortuita y solamente existe porque no se ha realizado recolecciones allí, estaríamos haciendo lo correcto al ajustar un solo modelo sencillo. Por otro lado, si los recolectores han explorado la brecha y no han encontrado accesiones, entonces estaríamos equivocados al ajustar un modelo único. Los recolectores individuales pueden conocer cuál es el caso pertinente pero, lamentablemente, los resultados negativos rara vez pasan a los bancos de germoplasma. No existe una entrada para “estuve allí y no encontré X”, aunque quizás esa información esté registrada en los libros de campo del recolector.

FloraMap no puede dar respuestas específicas a problemas de este tipo, pero sí puede indicar que se está presentando un problema de este tipo. Hemos incluido herramientas que permite al usuario investigar esa posibilidad.

Un indicador útil es el mismo mapa de probabilidad. Si las áreas de alta probabilidad caen sistemáticamente en áreas con pocos puntos de accesión, mientras que muchos de los puntos de accesión caen en las áreas de probabilidad mediana a baja, entonces puede existir un problema del tipo descrito anteriormente. Este resultado aparentemente improbable es, para nada, inusual.

La figura que se presenta a continuación indica cómo esto puede ocurrir. Las dos distribuciones normales se muestran en los histogramas de color negro y gris. La una se encuentra compensada por la otra, con solamente una pequeña superposición en el medio. La línea bimodal (A) es la suma de sus curvas de distribución normal. Se ajusta bien a los datos del histograma y muestra claramente la caída en el medio de la distribución donde hay pocas observaciones. Si fuéramos a ajustar el total de las observaciones como si fueran una población continua, obtendríamos la curva de distribución (B). Observe que el punto máximo de la función de densidad de probabilidad en este caso se presenta donde ocurre menor número de observaciones. Es un caso claro de un medio sin sentido. Trate de decir a alguien con una mano en agua hirviendo y la otra en hielo que la temperatura promedio está buena. Una vez que haya curado sus quemaduras por calor y por frío, lo más seguro es que no te agradecerá por tus observaciones.



Dos poblaciones normales como histogramas, que indican la suma de sus dos curvas de distribución (A), y la curva de distribución ajustada al conjunto completo de puntos (B).

Este tipo de problema puede deberse a una sola razón o, más frecuentemente, a una combinación de múltiples razones.

1. La especie existe en las áreas indicadas, pero nunca ha sido recolectada allí.
2. Las barreras geográficas o ecológicas han evitado la propagación de la especie a estas regiones.
3. Lo que se tomó como grupo homogéneo de germoplasma al comienzo del análisis está demostrando ser, en realidad, grupos diversos de adaptación al clima, es decir, diferenciación ecotípica.

4. Las especies que han emergido recientemente han presentado una dispersión inadecuada.
5. Ha habido interferencia humana.

En muchas situaciones, la elucidación de las razones tendrá un impacto importante en el estudio y en el uso del germoplasma.

Una primera verificación de la posibilidad de una distribución discontinua en el espacio climático es el diagrama de dispersión provisto en la ventana de PCA (ver Capítulo 3). El diagrama le permite visualizar las accesiones en cualquiera de los planos definidos por los componentes principales. Dado que los componentes son ortogonales, con cada transecto del espacio del componente, está logrando una imagen bidimensional en ángulo recto a todos los demás componentes. (No trate visualizar hacia donde se dirigen los otros ángulos rectos pues le puede dar dolor de cabeza.)

La variancia representada por cada componente disminuye rápidamente en la medida que baje la lista. Por tanto, primero debe mirar el transecto del componente 1 frente al componente 2, lo cual muchas veces representa más de tres cuartos de toda la variancia. El primer componente a menudo se considera un componente de 'tamaño'. El 'tamaño' de un clima es un concepto algo difícil de manejar, pero lo que realmente queremos decir es que es un componente que muestra una tendencia básica. Es decir, los climas en una dirección pueden ser generalmente más húmedos y más calientes, mientras que los de la otra dirección son más secos y más frescos. En la medida que usted baja por los componentes, se describen combinaciones progresivamente más complicadas y esotéricas de los datos. Sin embargo, a veces es posible visualizar la estructura general de un componente hasta el cuarto o quinto, y le puede dar un significado descriptivo. A veces surgen conceptos de la 'forma' o la 'relación' de las variables. Rara vez se puede hacer una interpretación después del quinto componente, pero los componentes subsiguientes representan, entonces, tan poquita de la variancia que generalmente no se tienen en cuenta.

Observe cada transecto del diagrama de dispersión respecto a cualquier agrupación o discontinuidad en los datos. Si se presentan agrupaciones obvias, entonces esto puede apoyar las pruebas de un mapa que muestra el tipo de distribución de probabilidad descrito anteriormente. En este momento es bueno volver a mirar los datos y determinar si se presentan cualquier agrupación genética o morfológica conocida dentro de la muestra de accesiones que puede explicar el comportamiento del modelo. FloraMap tiene una herramienta para ayudar con este análisis desde el punto de vista de los datos climáticos.

Análisis de Grupos

Hemos incorporado diversos métodos de agrupación en FloraMap para ayudar a desarrollar un modelo de probabilidad climático que puede manejar poblaciones múltiples. Si desea leer más acerca del análisis de grupos, recomendamos consultar a Jain y Dubes (1988) para un tratado mucho más detallado de las siguientes descripciones. Aunque Everitt (1974) es un tratado más antiguo, es aún muy comprensible. Hartigan (1975) da tanto una buena descripción práctica de las técnicas como el código fuente de Fortran de algunas aplicaciones interesantes.

Los métodos que hemos incorporado constituyen un subconjunto pequeño, pero ampliamente usado, de métodos de agrupación. Estos son enlace único (*single-link*), enlace completo (*complete-link*), promedio del grupo (*group average*), promedio ponderado del grupo (*weighted group average*), centroide sin ponderar (*unweighted centroid*), centroide ponderado (*weighted centroid*) y el método de Ward (*Ward's method*). Jain y Dubes clasifican todos estos métodos como **SAHN** (de tipo secuencial, aglomerativo, jerárquico y sin superposición, por las palabras en inglés). Son secuenciales porque los elementos operan uno a la vez, en contraposición con todos juntos. Son aglomerativos en que los grupos se van formando, etapa por etapa, al adicionar miembros o al fusionar grupos. Son jerárquicos, y en todos ellos se puede construir un árbol de dendrograma que muestre la relación entre los grupos a cada nivel de agrupación y entre niveles. Los grupos resultantes no se superponen en el espacio n en el cual están dibujados (en nuestro caso 36 dimensiones). La medida de la distancia es siempre una distancia euclidiana calculada a partir de los datos climáticos después de su transformación y ponderación, pero antes del PCA. Una distancia euclidiana se calcula como la raíz cuadrada de una suma de los cuadrados. En un caso, el método de Ward, ésta es la distancia euclidiana al cuadrado; no se toma la raíz cuadrada.

Un conjunto adicional de acrónimos puede aplicarse a algunos de los métodos (ver Jain y Dubes). El núcleo de las siglas, **PGM**, representa métodos de grupos de pares (**paired group methods**), los prefijos **U** y **W** significan sin ponderar (**unweighted**) y ponderado (**weighted**) y los sufijos **A** y **C**, media aritmética (**arithmetic mean**) y centroide (**centroid**). Por tanto, el método del promedio del grupo (*group average method*) a menudo se conoce como UPGMA, el promedio del grupo ponderado (*weighted group average*) como WPGMA, el centroide sin ponderar (*unweighted centroid*) como UPGMC y el centroide ponderado (*weighted centroid*) como WPGMC. El uso de estas nomenclaturas es lo suficientemente generalizado para que se incluyan en los nombres del método en la ventana de análisis de grupos.

Lance y Williams (1967) sugirieron que una fórmula generalizada para la actualización de matrices de distancia podría cubrir el método de agrupación

SAHN más común. Siguiendo el método de Jain y Dubes, esto sería para los grupos k , r y s :

$$d[k, (r, s)] = \alpha_r d[k, r] + \alpha_s d[k, s] + \beta d[r, s] + \gamma |d[k, r] - d[k, s]| \quad (17)$$

Donde $d[]$ es la función de distancia y $d[k, (r, s)]$ es la distancia del grupo recién formado (r, s) y el grupo existente k , que tiene n_k elementos.

El siguiente cuadro muestra los coeficientes de las medidas de distancia según se utilizan durante la implementación de FloraMap, donde n_r es el número de puntos en la agrupación r , n_s , en la agrupación s y n_k en k .

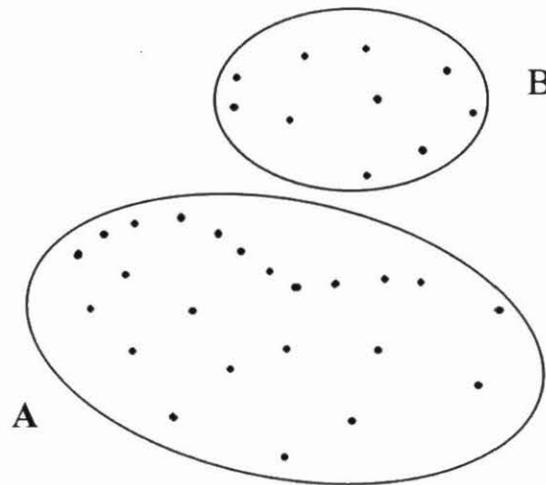
Valores del coeficiente para algoritmos de actualización de matrices tipo SAHN (secuencial, aglomerativo, jerárquico y sin superposición) (según Jain y Dubes 1988).

Método de agrupación	α_r	α_s	β	γ
Enlace único	1/2	1/2	0	-1/2
Enlace completo	1/2	1/2	0	1/2
UPGMA (promedio del grupo)	$\frac{n_r}{n_r + n_s}$	$\frac{n_s}{n_r + n_s}$	0	0
WPGMA (promedio ponderado)	1/2	1/2	0	0
UPGMC (centroide sin ponderar)	$\frac{n_r}{n_r + n_s}$	$\frac{n_s}{n_r + n_s}$	$\frac{-n_r n_s}{(n_r + n_s)^2}$	0
WPGMC (centroide ponderado)	1/2	1/2	-1/4	0
Método de Ward (variancia mínima)	$\frac{n_r + n_k}{n_r + n_s + n_k}$	$\frac{n_s + n_k}{n_r + n_s + n_k}$	$\frac{-n_k}{n_r + n_s + n_k}$	0

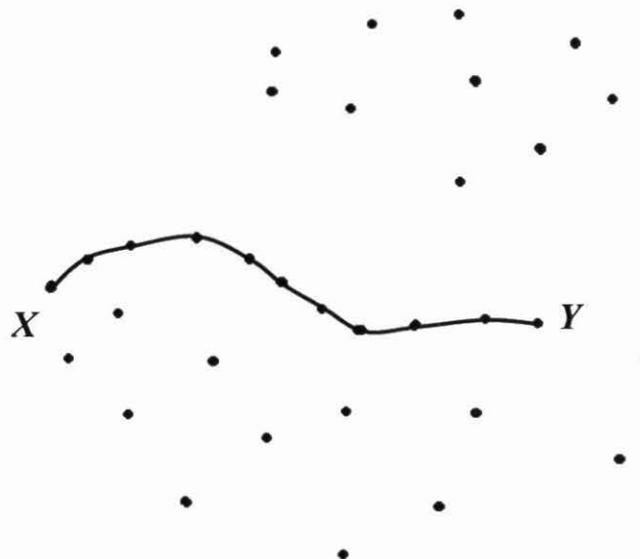
El método de enlace único se relaciona estrechamente con el árbol mínimo con conectividad completo (**minimum spanning tree**) de la teoría de gráficos. El mismo dendrograma puede generarse a partir de un algoritmo aglomerativo de enlace único o al eliminar progresivamente el enlace más grande del árbol mínimo. Los grupos, por consiguiente, dependen completamente de la distancia entre los puntos individuales y no de las propiedades de los grupos emergentes. En la práctica, el método es rápido y

útil para extraer numerosos grupos pequeños de una población donde se espera una agrupación altamente local. Tiene una característica que, dependiendo de su aplicación, puede considerarse como una fortaleza o como una debilidad.

Con base en la colección de puntos presentados a continuación, es posible que queramos resaltar los dos grupos más obvios, (A) y (B).



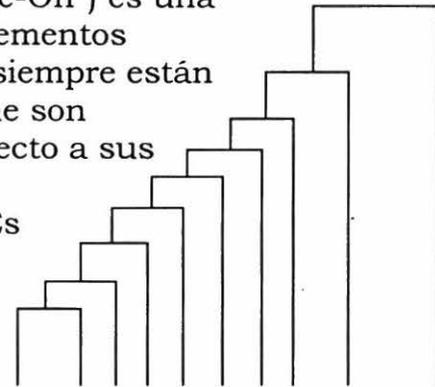
Sin embargo, se presenta otra agrupación evidente que puede tener un significado físico.



El grupo de puntos unidos por la línea X a Y bien puede ser un transecto en medio de muchos puntos aleatorios. Se pueden presentar casos en que la capacidad para separar estos datos va a ser importante. El algoritmo de enlace único es ideal para este tipo de grupo, aunque a menudo no tiene éxito con los más comunes. Tiende a formar grupos adicionando elementos individuales y

produce dendrogramas cuya forma característica así lo indica. Puede ser frustrante si usted está buscando grupos diferenciados.

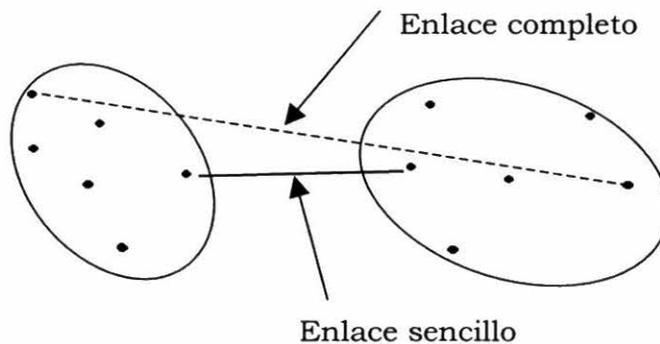
En algunos casos, “Adicionar Uno” (“Add-One-On”) es una herramienta muy útil. Si miramos los últimos elementos incorporados en el dendrograma completo, casi siempre están en la forma de “Add-One-On”, lo que significa que son puntos que presentan una distancia mayor respecto a sus vecinos más cercanos, es decir son lo que comúnmente llamaríamos resultados aislados. Es casi seguro de que corresponden a puntos con errores de ubicación en sus pasaportes o a accesiones interesantes encontrados en ambientes atípicos (ver la sección de TUTORIAL, página 31).



El temido dendrograma “adicionar uno”

La mayoría de los otros métodos que hemos ejecutado corresponden a intentos de producir grupos más compactos y evitar el temido dendrograma “Add-One-On”.

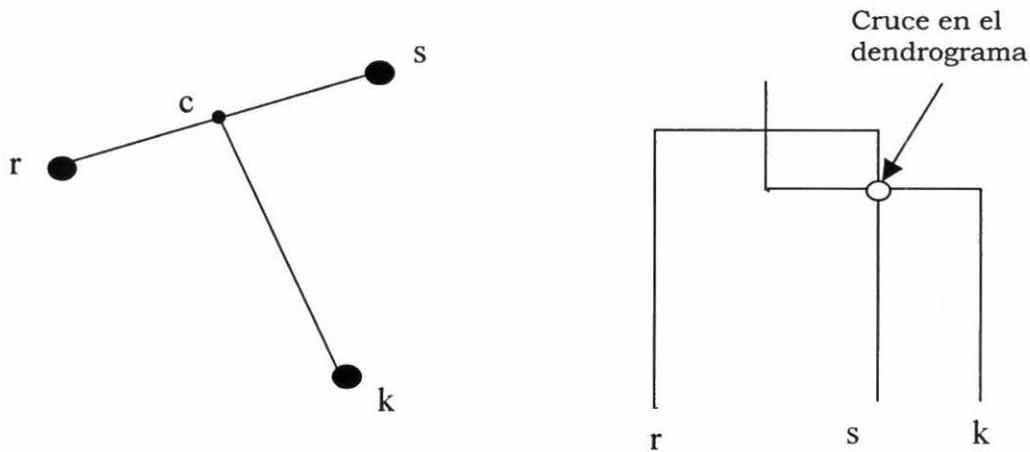
El método de enlace completo evita este problema al medir entre los componentes más lejanos de los dos grupos a unirse.



Los métodos que se basan en promedios buscan mantener las mejores características de los dos métodos anteriores, y utilizan la media aritmética de las dos líneas (punteada y llena) que aparecen en el diagrama anterior. Los métodos UPGMA y WPGMA difieren en el sentido de que la primera opción (sin ponderar) considera a todos los individuos de un grupo; el segundo método (ponderado) le asigna la misma ponderación a cada grupo, independientemente del número de individuos que contiene. Tanto UPGMA como WPGMA funcionan razonablemente bien con grupos compactos, pero en algunas circunstancias pueden generar grupos cuya forma es anormal. Esto les da cierta ventaja sobre UPGMC y WPGMC, pero el usuario debe considerar los datos cuidadosamente para determinar si los grupos son reales. No se puede dar una representación

geométrica a los métodos que se basan en promedios, ya que no hay localidad en la distancia promediada.

Sin embargo, UPGMC y WPGMC tienen una representación gráfica, que puede ser utilizada para ilustrar una consecuencia desafortunada de los métodos, como se ilustra en el siguiente diagrama.



Aquí los grupos r y s se han unido para formar un nuevo grupo (r, s) con su centroide en c . Las líneas r , k y s , k son más largas que la línea r , s . Por tanto, el grupo (r, s) se forma antes de considerar k . Sin embargo, cuando el grupo k

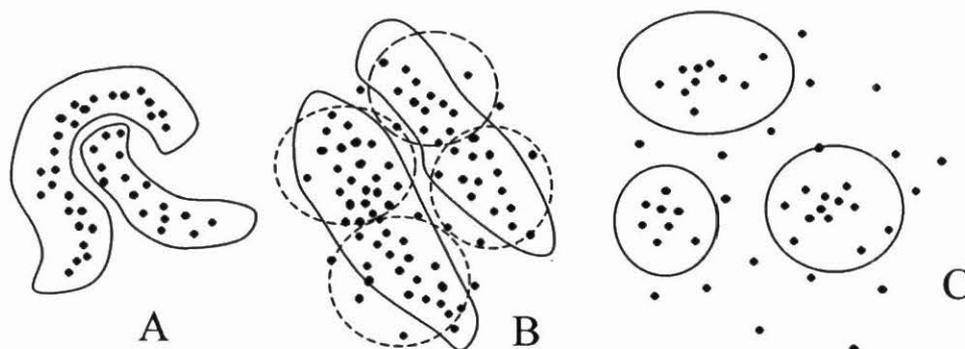
se une al centroide de (r, s) , la línea c, k es más corta que la línea (r, s) , lo cual significa que los grupos no se forman en una escala monótono y el dendrograma así lo refleja al tener un cruce como se ilustra en el diagrama arriba a la derecha.

Aunque los dendrogramas producidos por los algoritmos UPGMC y WPGMC son difíciles de interpretar debido a este efecto, estos métodos generalmente producen grupos compactos y bien formados. Si los puntos son completamente aleatorios, cerca del 13% de las fusiones de grupo producen cruces.

El método de Ward produce los grupos más compactos y el dendrograma que se ve más depurado. Este procedimiento busca minimizar la variancia dentro de los grupos y maximizar la que existe entre las distancias de los grupos.

Hay que tener cuidado en la interpretación de cualquiera de los métodos anteriores. Los métodos de agrupación compacta pueden generar grupos compactos a partir de datos completamente aleatorios, mientras que el algoritmo de enlace único tiene más probabilidad de producir un dendrograma "Add-One-On". Se debe aprovechar toda oportunidad para estudiar la distribución de los puntos de datos. El diagrama de flujo del PCA es una oportunidad para hacerlo.

El siguiente diagrama muestra algunas de las posibles configuraciones e indica las técnicas de agrupación que pueden producir los mejores resultados.



Solamente el algoritmo de enlace único encontrará los grupos en (A). Si son realmente definitivos, con una brecha entre ellos al menos tan ancha como el enlace único más grande dentro de ellos, la complejidad de la forma del grupo no será demasiado importante. En el caso de esta distribución demostrativa, el dendrograma probablemente parecería como dos o más cascadas “Add-One-On” enlazadas.

A simple vista, parece que existen dos grupos elípticos alargados en el ejemplo (B) (ver las curvas sólidas dibujadas a mano). Los algoritmos de agrupación compacta, como el método de Ward y el método centroide, probablemente definirán los cuatro grupos denotados por las elipses punteadas.

Los métodos *single-link* y *group average* tienen mayor probabilidad de encontrar grupos alargados. Si los datos se reescalaban para depurar los grupos, entonces los métodos de agrupación compacta funcionarían bien.

El método de Ward es el método que mejor encuentra los grupos esféricos densos en un fondo de puntos aleatorios —ejemplo (C). Hay que tener mucho cuidado en asegurar que los grupos producidos son reales y no un mecanismo de distribución aleatoria de puntos.

Bibliography

- Carnahan, B.; Luther, H.A.; Wilkes, J.O. 1969. Applied numerical methods. John Wiley, NY, US. 604 p.
- Cooley, W.W.; Lohnes, P.R. 1971. Multivariate data analysis. John Wiley, NY, US. 364 p.
- Daly, C.; Taylor, G. 1998a. Annual maximum, minimum, and mean temperatures of the coterminous United States, 1961-1990. Water and Climate Center of the Natural Resources Conservation Service, April 1998. 12 GIS grids, GRASS format. Portland, OR, US.
- Daly, C.; Taylor, G. 1998b. United States average monthly precipitation, 1961-1990. Water and Climate Center of the Natural Resources Conservation Service, April 1998. 12 GIS grids, GRASS format. Portland, OR, US.
- Digby, P.; Galwey, N.; Lane, P. 1989. Genstat 5. A second course. Clarendon Press, GB. 233 p.
- Everitt, B. 1974. Cluster analysis. Reviews of current research, no. 11, Social Science Research Council. Heinemann, GB. 122 p.
- Hartigan, J.A. 1975. Clustering algorithms. John Wiley, NY, US. 351 p.
- Hutchinson, M.F. 1997. ANUSPLIN Version 3.2 Users guide. The Australian National University. Centre for Resource and Environmental Studies. Canberra, AU. 39 p.
- Jain, A.K.; Dubes, R.C. 1988. Algorithms for clustering data. Prentice Hall, New Jersey, US. 320 p.
- Jones, P.G. 1987. Current availability and deficiencies data relevant to agroecological studies in the geographical area covered in IARCS. *In*: Bunting, A.H. (ed.). Agricultural environments: Characterization, classification and mapping. CAB International, GB. p. 69-82.
- Jones, P.G. 1991. The CIAT Climate Database Version 3.41. Machine readable dataset. Centro Internacional de Agricultura Tropical (CIAT), Cali, CO.
- Jones, P.G. 1996. Climate Database for Haiti. Machine readable dataset. Centro Internacional de Agricultura Tropical (CIAT), Cali, CO.

- Jones, P.G.; Thornton, P.K. 1993. A rainfall generator for agricultural applications in the tropics. *Agric. For. Meteor.* 63:1-19.
- Jones, P.G.; Thornton, P.K. 1997. Spatial and temporal variability of rainfall related to a third order Markov model. *Agric. Forest Meteor.* 86:127-138.
- Jones, P.G.; Rebgetz, R.; Maas, B.L.; Kerridge, P.C. 1996. Genetic diversity in *Stylosanthes* species: A GIS mapping approach. Paper and poster presented at the First International Symposium on Tropical Savannas, 24-29 Mar, Brasilia, Brazil. 14 p. Typescript copies available from CIAT, Cali, CO.
- Jones, P.G.; Galwey, N.; Beebe, S.E.; Tohme, J. 1997. The use of geographical information systems in biodiversity exploration and conservation. *Biodivers. Conserv.* 6:947-958.
- Lance, G.N.; Williams, W.T. 1967. A general theory of classificatory sorting strategies: II. Clustering systems. *Computer J.* 10:271-277.
- Morrison, D.F. 1967. *Multivariate statistical methods*. McGraw-Hill series in probability and statistics. McGraw-Hill, NY, US. 338 p.
- NOAA (National Oceanographic and Atmospheric Administration). 1984. TGP-006 D. Computer compatible tape. NOAA, Boulder, CO, US.
- Press, W.H.; Flannery, B.P.; Teukolsky, S.A.; Vetterling, W.T. 1986. *Numerical recipes: The art of scientific computing*. Cambridge University Press, GB. 818 p.

Apéndice A

Tipos de Archivo FloraMap

Archivos ACP son archivos de puntos de acceso compuestos como archivos ASCII delimitados por espacios con encabezamientos de columna. Son el equivalente ASCII de los archivos de puntos de acceso en DBF. FloraMap puede leerlos directamente.

Archivos DBF se utilizan para diversos tipos de archivo asociados con una cobertura del mapa. Pueden contener datos de puntos de acceso, datos climáticos o información acerca de los polígonos en los **shapefiles**.

Archivos MAP son los archivos de FloraMap que contienen la información que permite relacionar archivos DBF, archivos de puntos de acceso y **shapefiles** en una cobertura del mapa. Un **MAP** file contiene referencias a las diversas capas del mapa combinadas juntas para representar un mapa. FloraMap construirá un conjunto de todos los archivos necesarios para el mapa y les asignará un nombre a partir del nombre que usted haya asignado al archivo **MAP**. También contiene los atributos visuales para cada capa tales como colores, tamaños de tipo de letra, y así sucesivamente.

Archivos SHP son los **shapefiles** que delimitan los datos correspondientes a puntos, líneas o polígonos. Son también compatibles con ArcView.

Archivos SHX son archivos índice especializados que dan sentido a los **shapefiles**.

Archivos TXT se producen cuando se guarda un archivo de informe (**report file**). Estos son archivos de datos ASCII delimitados por espacios.

Un **shapefile ESRI** consta de un archivo principal (**main file**), un archivo índice y un cuadro en dBASE. El archivo principal es un archivo de acceso directo, con una longitud variable de registro, en el cual cada registro describe una forma ("shape") con una lista de sus vértices. En el archivo índice, cada registro contiene la dirección del registro correspondiente del archivo principal, desde el comienzo del registro principal. El cuadro en dBASE contiene los atributos característicos con un registro por característica. La relación una-a-una entre geometría y atributos se basa en el número del registro. Los registros de atributos en el archivo en dBASE deben estar en el mismo orden como los registros en el archivo principal.

Ejemplo

Archivo principal: counties.shp
 Archivo índice: counties.shx
 Cuadro en dBASE: counties.dbf

Los tres archivos anteriores describen una capa del mapa.

Creación de Archivos de Puntos de Acceso

Los archivos de puntos de acceso pueden construirse como archivos **ACP** o **DBF**. Deben contener por lo menos dos columnas con el encabezamiento **latitud longitud**. No es necesario tener ninguna otra columna de datos.

```

latitud longitud
"
5.6 -76.3
"
4.3 -74.3
10.4 -73.0

```

Este es un archivo ACP mínimo con cuatro puntos de acceso.

Si usted presenta a FloraMap un archivo como este, sin una columna para **elevation**, entonces el análisis trabajará utilizando la elevación tomada de la rejilla de la Base de Datos Climáticos, aunque usted haya especificado la temperatura correcta en **configuration window**. Para permitir la corrección de la temperatura mediante la elevación, usted debe proporcionar una columna para los datos de elevación. Usted también puede adicionar el número de columnas de datos que desea. Estas serán leídas junto con los puntos de acceso y se mantendrán como identificación junto con los puntos de acceso a través de todo el análisis.

Latitud	longitud	elevación	id	info	
5.6	-76.3	1200	1	x	Usted puede adicionar el número de columnas que desea siempre y cuando estén completas.
4.3	-74.3	600	2	v	
10.4	-73.0	200	3	bbb	
-20.3	-44.3	120	4	uuvv	

Si usted tiene columnas de datos que no están completas para todas las accesiones, entonces no podrá utilizar este formato delimitado por espacios. Debe preparar el archivo en un formato delimitado por comas que permite reconocer los valores faltantes. Por tanto:

ce A

```
LATITUD, longitud, elevación, id, info
5.6, -76.3, 1200, 1, x
4.3, -74.3,      , 2, v
10.4, -73.0, 200, 3, bb
-20.3, -44.3, 120, 4, uuvv
```

La elevación para el punto 2 ahora tiene un valor faltante para la elevación.

FloraMap no puede leer este archivo delimitado por comas. Léalo en su programa de hoja electrónica, verifique que el formateo esté correcto y guárdelo como un archivo en dBASE 4, con la extensión de archivo .dbf.

E18					
	A	B	C	D	E
1	LATITUDE	longitude	elevation	id	info
2	5.6	-76.3	1200		1 x
3	4.3	-74.3			2 v
4	10.4	-73	200		3 bb
5	-20.3	-44.3	120		4 uuw
6					

Glosario

Accession points	Puntos de acceso
Accession points added	Puntos de acceso agregados
Accession points dataset	Base de datos de puntos de acceso
Accession points file	Archivo de puntos de acceso
Accession points found	Puntos de acceso encontrados
Accession points layers	Capas de puntos de accesiones
Accessions dataset	Conjunto de datos de las accesiones
Accessions list	Lista de accesiones
Add layer symbols to legend	Agregue los símbolos de la capa a la leyenda
Add-One-On tool	Herramienta de Adicionar Uno
After	Después
After transformation window	Ventana de datos transformados
Alt-print screen	Alt-impr pant
Analysis	Análisis
Apply	Aplicar
Arithmetic mean	Media aritmética
Attributes	Atributos
August	Agosto
Autosave	Autoguardar
Before	Antes
Boundaries	Limites
Built-in climate grids	Rejillas climáticas incorporadas
Calculate average	Calcular promedio
Calculating distances	Calculo de distancias
Calculation	Cálculo
Calculation menu	Menú de cálculo
Calculation parameters	Parámetros de cálculo
Cancel	Anular
Centroid	Centroide
Checking accessions	Verificando accesiones
Circle	Circulo
Climate diagram	Diagrama del clima
Climate grids	Rejillas climáticas
Cluster	Grupo / agrupación
Clustering	Agrupación
Cluster select line	Línea de selección de agrupación
Cluster select window	Ventana para seleccionar las agrupaciones
Clusters form	Ventana de agrupación
Cluster techniques	Técnicas de agrupación
Cluster window	Ventana de agrupación
Collector	Coleccionista
Complete-link	Enlace completo
Component	Componente

Components available	Componentes disponibles
Components control	Control de los componentes
Component scores	Puntuaciones de los componentes
Configuration	Configuración
Confirm	Confirmar
Confirm menu	Menú de confirmación
Contents of transfer	Contenidos de la transferencia
Continue	Seguir
Control Panel	Tablero de Control
Correct temperature	Corregir temperatura
Countries	Países
Coverages	Coberturas
CPU	Unidad central de procesamiento
Creating climate file	Creando el archivo de clima
Database	Base de datos
Default.map	Mapa por defecto
Deleted	Suprimido
Diurnal	Diurno
Eigenvalues	Valores específicos
Eigenvectors	Vectores específicos
Elevation	Elevación
Enlarge scatterplot	Agrandar diagrama de dispersión
Exit	Salida
External	Externo
Field	Campo
File	Archivar
File name	Nombre del archivo
File of type	Archivo de tipo
Fill	Rellenar
From	Desde
Full	Completo
Germplasm	Mapeo de biodiversidad de germoplasmo
biodiversity mapping	
Group average	Promedio del grupo
Help	Ayuda
Information window	Ventana de información
January	Enero
Label	Etiqueta
Lapse rate	Variación de la temperatura en razón de la altura
Latin America	América Latina
Latitude	Latitud
Layer	Capa
Layer control	Control de capa
Layers control menu	Menú de control de capas
Layers control window	Ventana de control de capas

Layers menu	Menú de capas
Layers window	Ventana de capas
Layer symbols	Símbolos de la capa
Load	Cargar
Load a layer	Cargar una capa
Longitude	Longitud
Look in	Buscar en
Lowest probability	La probabilidad más baja
Main file	Archivo principal
Map	Mapa
Map background color	Color de fondo del mapa
Map layer files	Archivos de capas del mapa
Mapscreen	Pantalla del mapa
Map window	Ventana del mapa
Marker	Marcador
Menu	Menú
Method	Método
Method-select window	Ventana para seleccionar el método
Minimum spanning tree	Árbol mínimo con conectividad completo
Mismatches	Desaciertos
Modified	Modificado
Moved automatically	Movido automáticamente
Moved manually	Movido manualmente
Name	Nombre
None	Ningún
Number	Número
Open	Abrir
Open layers	Abrir capas
Outline color	Color del contorno
Paired group methods	Métodos de grupos de pares
Pan	Girar para toma panorámica
PCA window	Ventana de análisis de componentes principales
Points moved to the nearest valid pixel	Puntos trasladados al pixel válido mas cercano
Power Use Group	Grupo de Usuarios de la Potencia
Principal components analysis	Análisis de componentes principales
Print diagrams	Imprimir diagramas
Print map	Imprimir mapa
Probabilities surface	Superficie de probabilidades
Probability	Probabilidad
Probability control	Control de probabilidad
Probability layer	Capa de probabilidad
Process indicator	Indicador del proceso
Program files	Archivos de programa

Quantity of breaks	Cantidad de niveles
Rain	Lluvia
Record	Registro
Render	Pintar
Rendered shapefiles	Shapefiles generados
Report	Informe
Report file	Archivo de informe
River	Río
Road	Camino
Rotated	Rotado
Run	Ejecutar
Save	Guardar
Scatterplot	Diagrama de dispersión
Scores N	Puntuación N
Select	Seleccione
Selection of this option may lead to wrong results	Selección de esta opción podría llevar a resultados incorrectos
Settings	Marcos de trabajo
Settings configuration general	Configuración general de marcos de trabajo
Show average climate for selected points	Muestre el clima promedio para los puntos seleccionados
Show in legend	Muestre en la leyenda
Single-link	Enlace único
Size	Tamaño
Solid fill	Relleno sólido
South America	América del Sur
Southeast Asia	El sudeste de Asia
Square	Cuadrado
Standard user	Usuario típico
Start	Comienzo
System	Sistema
Temperature	Temperatura
To	Hasta
Town	Aldea
Transformation	Transformación
Transparent	Transparente
Treat accessions with identical coordinates as a single observation	Trate a las accesiones con coordenadas idénticas como una sola observación
Type	Tipo
Unweighted centroid	Centroide sin ponderar
Value	Valor
Variance	Variación
Variance box	Recuadro de variación
Variance control	Control de la varianza
Variance lever	Palanca de variación

View	Visualizar
View accessions	Visualizar accesiones
View dataset	Visualizar conjunto de datos
Ward's method	Método de Ward
Weighted	Ponderado
Weighted average	Promedio ponderado
Weighted centroid	Centroide ponderado
Weighted group average	Promedio del grupo ponderado
Weights	Ponderaciones
With / without outline	Con / sin contorno
Working directory	Directorio de trabajo
Yes	Sí
Zoom in	Zoom cerca
Zoom out	Zoom lejos