



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

---

# POSGRADO EN CIENCIAS BIOLÓGICAS

Facultad de Ciencias

**MODELACIÓN DE LA DINÁMICA POBLACIONAL DE  
CACTÁCEAS ENDÉMICAS DEL VALLE DE TEHUACÁN  
A LO LARGO DE UN GRADIENTE DE DISTURBIO  
ANTROPOGÉNICO CRÓNICO**

## **T E S I S**

QUE PARA OBTENER EL GRADO ACADÉMICO DE

**MAESTRO EN CIENCIAS BIOLÓGICAS  
(BIOLOGÍA AMBIENTAL)**

P R E S E N T A

**EDGAR JAVIER GONZÁLEZ LICEAGA**

**DIRECTOR DE TESIS: DR. CARLOS MARTORELL DELGADO**

MÉXICO, D. F.

MAYO, 2008



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

# Agradecimientos institucionales

Agradezco:

- Al Posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México.
- Al Consejo Nacional de Ciencia y Tecnología por la beca proporcionada durante el período 2006–2007.
- A los miembros de mi Comité Tutorial:
  - Al Dr. Martorell por el apoyo y supervisión constante durante la realización de este trabajo.
  - A la Dra. Valverde por sus aportaciones a este trabajo, sus múltiples correcciones y por las sesiones de examen tutorial que me forzaban a tener claras mis ideas para aclarárselas a ella.
  - Al Dr. Méndez por el constante entusiasmo que demostró en los exámenes tutorales.
- A los sinodales que revisaron mi tesis contribuyendo con sus comentarios y correcciones: el Dr. Luis Morales Mendoza y el Dr. Jordan Golubov Figueroa.
- A la comunidad de Concepción Buenavista, Oaxaca, por permitirnos el estudio de las cactáceas amenazadas que habitan su municipio.

# Agradecimientos personales

Agradezco:

- A mi madre por el apoyo incondicional que me ha ofrecido a lo largo de toda mi vida: nunca podré retribuirte todo lo que me has dado.
- A mi padre que, aun con sus altibajos, sigue apoyándome.
- A mi tutor, Carlos Martorell, que, desde el primer curso que tome con él, me despertó el interés por la ecología, motor que me ha movido estos últimos años.
- A todas las personas que se vieron involucrados en la toma de los datos demográficos de *Mammillaria dixanthocentron*; por experiencia reconozco el esfuerzo requerido para obtenerlos.
- Al grupo de trabajo de Ecología y Diversidad Vegetal; amigos y compañeros con los que he convivido estos tres últimos años.
- A mis amigos los exPAADOS: sigamos juntos durante mucho tiempo más.

# Índice

## Archivo

- A3    Introducción
  - 1.1.    Disturbio antropogénico crónico (DAC) y su efecto en la dinámica de poblaciones vegetales
  - 1.2.    Modelación de la dinámica poblacional afectada por el DAC
  - 1.3    Las cactáceas endémicas de México
- A4    Resumen
- A4    Abstract
- A5    Modelo
  - 2.1    Modelo Integral de Proyección
  - 2.2    Modelo Integral de Proyección modificado para incorporar el DAC
  - 2.3    Estimación de los parámetros a partir de las estructuras poblacionales observadas
- A5    Método
  - 3.1    Modelación de la dinámica poblacional de acuerdo al tamaño de los individuos y al DAC
  - 3.2    Estimación de los parámetros del modelo a partir de las estructuras poblacionales observadas
  - 3.3    Algoritmo
    - 3.3.1    Entrada de datos

- 3.3.2 Generación de la primera estructura observada y del gradiente de DAC
- 3.3.3 Generación de partículas, obtención de los valores que determinan al núcleo, generación de éstos e iteración del modelo
- 3.3.4 Cálculo y maximización de la verosimilitud: Optimación por enjambre de partículas
- 3.4 Validación del modelo con datos artificiales
- 3.5 Aplicación del modelo a *Mammillaria dixanthocentron*
  - 3.5.1 Especie bajo estudio
  - 3.5.2 Sitio de estudio
  - 3.5.3 Información demográfica disponible
  
- A5 Resultados
  - 4.1 Datos artificiales
  - 4.2 Datos de *Mammillaria dixanthocentron*
  
- A5 Discusión
  - 5.1 Datos artificiales
  - 5.2 Datos de *Mammillaria dixanthocentron*
  - 5.3 Supuestos subyacentes al modelo
  - 5.4 Consideraciones adicionales y mejoras posibles al modelo
  - 5.5 Perspectivas
  
- A6 Apéndice: Programa MIPDAC
  - A.1 Diagrama de flujo del algoritmo
  - A.2 Código del programa
    - A.2.1 Entrada de datos
    - A.2.2 Cuadratura Gauss-Legendre
    - A.2.3 Obtención de parámetros que integran el núcleo
    - A.2.4 Generación del núcleo
    - A.2.5 Generación de las estructuras y cálculo de la verosimilitud de la partícula
    - A.2.6 Maximización de la verosimilitud
  
- A7 Literatura citada

# Índice de Tablas

- Tabla 3.1 Valores que tomó el criterio de información de Akaike para diferentes funciones  $\mu(X)$  para la modelación de la supervivencia ( $s(x)$ ) de *M. dixanthocentron* en dos sitios de estudio durante cinco años. Se resaltan en negritas los mejores ajustes. Sólo se presentan aquellas funciones  $\mu(X)$  que mejor modelaron al menos un conjunto de datos.
- Tabla 3.2 Evaluación del desempeño de distintas funciones de distribución para modelar el crecimiento de los individuos usando información de *M. dixanthocentron* de La Cucharilla en 2001. Se resalta en negritas el modelo elegido.
- Tabla 3.3 Valores que tomó el criterio de información de Akaike para distintos modelos que describen el comportamiento de los parámetros de la distribución gamma en la modelación del crecimiento de los individuos usando información de *M. dixanthocentron* de dos sitios de estudio durante dos años. Se resaltan en negritas los mejores ajustes.
- Tabla 3.4 Valores que tomó el criterio de información de Akaike para diferentes funciones  $\mu(X)$  para la modelación de la fecundidad ( $f_1(x)$ ) de *M. dixanthocentron* en dos sitios de estudio durante cinco años. Se resaltan en negritas los mejores ajustes.
- Tabla 3.5 Valores del DAC para los 10 sitios bajo estudio.

# Índice de Figuras

- Figura 4.1 Verosimilitud de las mejores partículas obtenidas en cada corrida del programa para los datos artificiales. La línea horizontal indica la verosimilitud de la partícula generadora.
- Figura 4.2 Distancia Manhattan de las mejores partículas obtenidas en cada corrida del programa con respecto a la partícula generadora. La línea horizontal indica la distancia esperada entre dos partículas generadas al azar.
- Figura 4.3 Estructura de tamaños “observada” (histograma), generada por la partícula generadora (línea oscura) y obtenida por la mejor partícula encontrada (línea clara) para las nueve intensidades de DAC.
- Figura 4.4 Supervivencia (a y b), crecimiento (c y d) y fecundidad (e y f) de los individuos dependiendo del grado de DAC y del tamaño de los mismos

de acuerdo con la dinámica arrojada por la partícula generadora y la mejor partícula encontrada.

- Figura 4.5 Supervivencia, crecimiento y fecundidad de los individuos dependiendo del grado de DAC y del tamaño de los mismos de acuerdo a la dinámica arrojada por la segunda mejor partícula encontrada.
- Figura 4.6 Verosimilitud de las mejores partículas obtenidas en cada corrida del programa para los datos reales.
- Figura 4.7 Verosimilitud de las mejores partículas obtenidas en cada corrida del programa partiendo de las 25 mejores partículas obtenidas previamente. Se indica la verosimilitud de la mejor partícula obtenida después de 150 corridas del programa.
- Figura 4.8 Distancia Manhattan de las mejores partículas obtenidas en cada corrida del programa partiendo de las 25 mejores partículas obtenidas previamente con respecto a la mejor partícula original (la resultante de 150 corridas del programa).
- Figura 4.9 Estructura de tamaños observada (histograma) y obtenida por el modelo desarrollado (línea) para las seis intensidades de DAC observadas en el campo.
- Figura 4.10 Supervivencia (a), crecimiento (b) y fecundidad (c) de los individuos de *Mammillaria dixanthocentron* dependiendo del grado de DAC y del tamaño de los mismos de acuerdo con la dinámica arrojada por la mejor partícula encontrada.
- Figura 4.11 Supervivencia (a), crecimiento (b) y fecundidad (c) de los individuos de *Mammillaria dixanthocentron* obtenida a partir de los datos de dinámica demográfica para los sitios La Cucharilla y La Pedrera (línea clara y oscura, respectivamente) durante el periodo 2001-2002.
- Figura 5.1 *Splines* cúbicos naturales correspondientes a la primera estructura poblacional observada (Cerro Gordo) de *Mammillaria dixanthocentron* para diferentes tamaños de partición: 10 (línea clara), 50 (línea oscura) y 100 (línea punteada) subintervalos.
- Figura A.1 Diagrama de flujo del algoritmo MIPDAC.

# Resumen

El disturbio antropogénico crónico (DAC) es un agente que impacta negativamente a las poblaciones de muchas especies. Sin embargo, no todas responden a éste de la misma manera a nivel poblacional, ni ven afectados sus procesos demográficos de la misma forma. Con el objetivo de modelar el efecto del DAC sobre la dinámica poblacional de una cactácea vulnerable (*Mammillaria dixanthocentron*) cuyas poblaciones, ubicadas en el Valle Tehuacán-Cuicatlán, están sujetas a diferentes intensidades de DAC, se desarrolló un modelo integral de proyección que parte únicamente de las estructuras de tamaño obtenidas para las poblaciones bajo estudio (10 en este trabajo) y de una medición cuantitativa del DAC al que están sujetas. El modelo aquí desarrollado encontró, mediante máxima verosimilitud y optimización por enjambre de partículas, la dinámica poblacional sujeta a DAC que mejor se ajustó a las estructuras observadas. Esta dinámica sugiere que el DAC afecta de manera preponderantemente negativa a todos los procesos demográficos (supervivencia, crecimiento y reproducción), y que este efecto es dependiente del tamaño de los individuos. La dinámica anterior se comparó con las de dos poblaciones sujetas a diferente intensidad de DAC: la supervivencia y el crecimiento presentaron los mismos patrones en ambas dinámicas; la fecundidad, aunque se vio afectada negativamente por el DAC en el campo, dependió del tamaño de los adultos en las dinámicas observadas, pero no así en el modelo. Se discuten los supuestos subyacentes y mejoras posibles al modelo, así como perspectivas de trabajo posterior.

# Abstract

Chronic anthropogenic disturbance (CAD) is an agent that impacts negatively the populations of many species. However, not all of these respond to CAD in the same manner at the population level, nor see their demographic processes determined in the same way. With the aim of modeling the effect of CAD over the population dynamics of a vulnerable cactus (*Mammillaria dixonanthocentron*) whose populations, located in the Tehuacán-Cuicatlán Valley, Mexico, present different CAD intensities, an integral projection model was developed that depends solely on the size structures of the populations under study (10 in this work) and on the quantitative measurement of the CAD present in the sites. Through maximum likelihood and particle swarm optimization, the model proposed here found the CAD-determined demographic dynamics that best fitted the above information. This dynamics suggests that CAD modifies the demographic processes in a size-dependent, mainly negative way. A comparison of this dynamics with those found in two populations located in areas with different CAD intensities, showed that survival and growth conform to the same patterns in both dynamics; fecundity, though negatively modified by CAD in both dynamics, was size-independent in the modeled dynamics and size-dependent in the observed one. Underlying assumptions and possible improvements to the model are discussed, along with an outlook on subsequent work.

# Introducción

El hombre es y ha sido un factor de disturbio cuyo impacto en la biodiversidad de nuestro planeta ha ido incrementándose de manera palpable a lo largo de los últimos siglos. Dicho disturbio se ha reflejado en la extinción de especies (Kesseli, 1992; Vega y Peters, 2003), la reducción de la producción primaria y secundaria de los ecosistemas (Rapport y Whitford, 1999), la reducción y fragmentación del hábitat de muchas especies (Bradshaw y Marquet, 2003), y la consecuente disminución de los números poblacionales y las áreas de distribución de estas especies. Dado el acelerado crecimiento de la población humana y el consecuente aumento en el disturbio ejercido sobre las poblaciones de muchas especies vulnerables, es de vital importancia comprender el modo en el que dicho disturbio antropogénico afecta a las poblaciones de las especies amenazadas para así tomar las medidas que permitan su conservación. La construcción de modelos matemáticos que relacionen las respuestas demográficas de las especies de interés con este disturbio puede contribuir a tomar dichas medidas (Vega y Peters, 2003).

## **1.1 Disturbio antropogénico crónico (DAC) y su efecto en la dinámica poblacional**

El hombre perturba a las poblaciones de muy diversas formas, tanto directas como indirectas, variando el disturbio tanto en su escala de afectación, como en su frecuencia e intensidad en comparación con los disturbios naturales que se presentan en todo ecosistema (White y Pickett, 1985). Esto involucra desde episodios de disturbio esporádicos pero intensos, hasta eventos que aparentemente no tienen un gran impacto, pero que, debido a su constante incidencia en el tiempo, acaban teniendo, respecto de los primeros, efectos mucho más dañinos y a más largo plazo sobre las poblaciones, al no permitirles recuperarse adecuadamente de la perturbación generada por el disturbio (Singh, 1998; Vega y Peters, 2003).

Aunque la definición de lo que es un disturbio ha estado bajo debate durante las últimas décadas (ver *e. g.* Sousa, 1984; Rykiel, 1985; Pickett *et al.*, 1989), para este trabajo definiré *disturbio* como aquel evento relativamente discreto en el tiempo que altera la estructura de un ecosistema, comunidad o población y modifica la disponibilidad de recursos, de sustrato para colonización o el ambiente físico (White y Pickett, 1985). Partiendo de lo anterior, definiré el *disturbio antropogénico crónico* (DAC) como aquel disturbio de baja intensidad causado por el hombre de manera reiterada a lo largo de un intervalo amplio de tiempo (Singh, 1998; Vega y Peters, 2003; Martorell y Peters, 2005). Éste consiste generalmente en la remoción sistemática de pequeñas fracciones de biomasa (para leña, forraje y construcción y como productos no maderables) de la vegetación, así como el sobrepastoreo y la erosión generada por la ganadería extensiva (Vega y Peters, 2003).

A diferencia del disturbio agudo, el cual es fácilmente discernible en el campo y genera una dicotomía tajante entre sitios perturbados y no perturbados, el DAC genera un gradiente que va de un sitio prístino a uno severamente degradado, teniéndose entonces que tratar como una variable cuantitativa (Martorell y Peters, 2005). Esto aumenta la dificultad de estudiar el fenómeno del efecto del DAC sobre las poblaciones vegetales, pues no basta comparar dos sitios (con vs. sin disturbio) sino que hay que considerar sitios que sean representativos del gradiente de DAC presente en la zona. Además, es importante notar que la respuesta de cada especie al disturbio es particular y que diferentes tipos e intensidades de disturbio tienen diferentes efectos (McIntyre y Lavorel, 1994; Watt, 1998).

A nivel poblacional, el disturbio es un fenómeno que afecta la dinámica demográfica de las especies (Bazzaz, 1983), la cual se suele descomponer en tres procesos vitales: supervivencia,

crecimiento y reproducción de los individuos. Sin embargo, el hecho de que el disturbio afecte la dinámica poblacional, no implica que éste incida en dichos procesos de igual modo, ni que distintos individuos (*e. g.*, de diferente tamaño) respondan de igual forma al disturbio. Por lo tanto, es importante conocer cómo los distintos procesos e individuos se ven afectados por el DAC para de esta forma asignar los esfuerzos de conservación y manejo de manera más adecuada, enfocándolos en aquel proceso o en aquellos individuos que se ven afectados más negativamente por el DAC.

## **1.2 Modelación de la dinámica poblacional afectada por el DAC**

Una manera de evaluar las consecuencias que a nivel poblacional tiene el DAC consiste en modelar matemáticamente la dinámica poblacional e incorporar el disturbio en la modelación para así poder evaluar sus efectos en los distintos procesos demográficos. Por modelo matemático se entiende aquel modelo cuyas partes son conceptos matemáticos, tales como constantes, variables, funciones, ecuaciones, etc. que tienen como objetivo describir alguna parte del mundo real (Meyer, 1984). Dentro de los modelos enfocados al estudio de la dinámica poblacional se ubican los llamados modelos poblacionales estructurados, los cuales toman en cuenta el hecho de que los individuos dentro de una población difieren en ciertos atributos (edad, tamaño, estadio de desarrollo, localización espacial o genotipo), que afectan, de manera diferenciada, a las tasas vitales. Éstos incorporan dichas diferencias en la modelación de la dinámica numérica. De esta forma un modelo estructurado suele consistir de una serie de reglas matemáticas que especifican el cambio en el tiempo del número de individuos en cada categoría considerada (Tuljapurkar y Caswell, 1997). Para usar estos modelos usualmente se suelen estudiar poblaciones cuyos individuos se siguen a lo largo de varios años y se obtienen, de manera empírica, los datos de crecimiento, supervivencia y reproducción, los cuales se incorporan al modelo, analizándose posteriormente los resultados arrojados por éste.

Los modelos matriciales de proyección poblacional (Caswell, 2001) son un ejemplo de modelos estructurados que han sido utilizados ampliamente en las últimas dos décadas (*e. g.* Crouse *et al.*, 1987; Golubov *et al.*, 1999; Contreras y Valverde, 2002; Valverde y Zavala-Hurtado, 2006). Estos modelos operan en tiempo discreto, proyectando los cambios numéricos de una población de un tiempo  $t$  a uno  $t+1$ . En estos modelos se establece en primer lugar el intervalo de proyección (*i. e.*, la escala temporal) bajo el cual se va a trabajar. En segundo lugar se elige el número y características de las categorías de tamaño, edad o estadio del ciclo de vida que disgregarán a los individuos de la población, tratando de reflejar, en la medida de lo posible, la división natural del ciclo de vida que presenta la especie estudiada. De esta manera, cada individuo en dicha población pertenecerá a una y sólo una categoría. Posteriormente se calcula la probabilidad de transición de un individuo de una categoría a otra, así como una fecundidad promedio para los individuos en cada categoría. Esta información se resume en una matriz  $\mathbf{A}$ , la cual contiene por lo tanto la información demográfica de los cambios que sufrirá una población de un momento a otro en el tiempo. Si nombramos como  $\mathbf{n}_t$  al vector que contiene la cantidad de individuos de la población en cada categoría de tamaño en un tiempo  $t$  (*i. e.*, la estructura de la población en categorías de tamaño), se tiene que

$$\mathbf{n}_{t+1} = \mathbf{A} \cdot \mathbf{n}_t, \quad (1.1)$$

donde  $\cdot$  es la multiplicación matricial.

Debido a la forma como se construye la matriz  $\mathbf{A}$ , ésta posee la propiedad de ser ergódica (Cohen, 1979), lo que implica que la iteración de la ecuación (1.1) generará una estructura asintótica  $\mathbf{n}_\infty$  que será independiente de la estructura inicial  $\mathbf{n}_0$ . Dicha estructura asintótica  $\mathbf{n}_\infty$  se interpreta como la estructura a la que dicha población convergería de permanecer durante un largo tiempo sometida a la dinámica poblacional contenida en  $\mathbf{A}$ . Sin embargo, como se desprende de lo anterior, un supuesto importante de este modelo es que las condiciones ambientales que determinan la dinámica poblacional contenida en  $\mathbf{A}$  permanecen constantes a lo largo del tiempo, lo cual impide el modelación de la dinámica poblacional sujeta a un proceso como el DAC que la modifica año con año.

Los modelos matriciales estocásticos han sido usualmente vistos como una solución a lo anterior (Caswell, 2001). Dichos modelos intentan incorporar algo de la estocasticidad ambiental

(e. g., el disturbio en Silva *et al.*, 1991) a la que está sometida una población. Sin embargo, aún en estos modelos, lo que finalmente se analiza es la dinámica estocástica a largo plazo y no hay manera de evaluar la adecuación de dicha dinámica con respecto a la realidad (Fox y Gurevitch, 2000). Aún más, al ser el DAC un cambio determinístico, direccional y, por lo tanto, no estocástico, dichos modelos no permiten integrar el efecto del DAC.

Descartando los modelos matriciales estocásticos, hay que considerar la alternativa de hacer un seguimiento del devenir demográfico de una población inicialmente en estado prístino a medida que el DAC la fuese afectando cada vez con mayor intensidad. Considerando que éste es un proceso paulatino, requeriríamos contar con datos de dinámica poblacional a muy largo plazo (del orden de  $1 \times 10^2$  a  $1 \times 10^3$  años), lo cual hace inviable el seguimiento de una sola población durante la mayor parte del proceso de degradación generado por el DAC.

Una alternativa a lo anterior es recurrir al estudio de distintas poblaciones sometidas a distintos grados de DAC y concebir dichas poblaciones como representantes de una cronosecuencia a lo largo del gradiente de DAC. Una cronosecuencia resulta de una sustitución espacio por tiempo en la cual se extrapola una tendencia temporal de una serie de muestras de diferente edad (Pickett, 1989; este supuesto se discute en la sección 5.3). De esta manera, en una cronosecuencia de DAC, las poblaciones, ubicadas en sitios en los que se presenta una distinta intensidad de DAC, pueden representar el devenir potencial de una sola población ante un incremento en el DAC a lo largo del tiempo. La información idónea que se requeriría entonces para cada población integrante de la cronosecuencia sería la de su dinámica demográfica para la intensidad de DAC particular a la cual está sometida.

Dado que para modelar la dinámica de una población se requiere de al menos un año de seguimiento en el campo, estamos hablando de un esfuerzo de trabajo bastante elevado para seguir la dinámica de todas las poblaciones involucradas. Aun contando con esta información, se tiene que tomar en cuenta el factor azaroso involucrado en haber estudiado a una población en un año particular, el cual puede no ser representativo de las condiciones imperantes en un año promedio. Si consideramos la enorme biodiversidad de México y la gran cantidad de especies que pueden verse amenazadas por el DAC, este tipo de estudios pueden requerir de un esfuerzo prohibitivo, especialmente considerando la escasa disponibilidad presupuestal y de personal capacitado que existe en el país. Debido a esta gran cantidad de especies y la rapidez con que sus

poblaciones están siendo mermadas, el factor tiempo es preponderante, y se hace imperativo contar con un método rápido y sencillo que nos proporcione dicha respuesta.

Una alternativa al seguimiento de las poblaciones de la especie de interés a lo largo del tiempo que se consideró para este trabajo fue el uso de las estructuras de tamaños de dichas poblaciones. Una estructura poblacional en un momento dado en el tiempo es el resultado de la dinámica demográfica a la que estuvieron sujetos sus individuos en el pasado y de la estructura poblacional pretérita (ecuación (1.1)). Hablando de individuos longevos, como es el caso de las cactáceas, el intervalo de tiempo reflejado por la estructura puede comprender varios años. De esta manera, una estructura poblacional contiene información sobre la dinámica poblacional a la que estuvo sometida la población en los últimos años. Siendo el DAC uno de los factores que pueden afectar dicha dinámica, mi intención es utilizar una cronosecuencia de estructuras poblacionales para reconstruir la dinámica demográfica a lo largo del tiempo, así como los cambios que ésta ha sufrido como resultado del DAC.

El utilizar las estructuras poblacionales en lugar de las dinámicas demográficas completas tiene la gran ventaja de la facilidad y rapidez con la que las estructuras se pueden obtener en el campo, al requerir de una sola toma de datos para cada población, la cual se realiza en el término de horas y no de años, como es el caso de las dinámicas.

De esta forma, este trabajo tuvo la intención de desarrollar un modelo que partiera de información mínima, fácilmente obtenible, como lo son las estructuras poblacionales, y valorar la respuesta que una especie en particular tendrá ante el DAC. Este modelo se suscribe dentro de los llamados “métodos rápidos y sucios” (*quick-and-dirty*), los cuales tienen como objetivo proporcionar un estimado aproximado del efecto de un agente sobre una población o ecosistema de manera rápida, fácil y barata (U. S. EPA, 2002; Benton, 2007).

Así, el objetivo de mi trabajo fue desarrollar un modelo que pudiese utilizar estructuras de tamaños obtenidas de poblaciones representantes de una cronosecuencia de perturbación por DAC para reconstruir los cambios en la dinámica poblacional debidos al DAC. Esta información podría ser utilizada para evaluar el efecto que sobre los distintos procesos demográficos (supervivencia, crecimiento y fecundidad) tiene el DAC y plantear estrategias de manejo y conservación apropiadas para la especie en cuestión.

La idea conceptual de mi modelo consiste en lo siguiente:

- 1) Partir de una estructura poblacional inicial.

- 2) Plantear un modelo poblacional estructurado que proyecte las estructuras poblacionales de un tiempo  $t$  a un tiempo  $t+1$ .
- 3) Plantear un conjunto de reglas de cambio que describan la forma en la que la dinámica demográfica va cambiando a lo largo del gradiente de DAC.
- 4) Utilizar todo lo anterior para proyectar las estructuras poblacionales a lo largo de un periodo largo de tiempo en el cual el DAC va aumentando paulatinamente.
- 5) Utilizar las estructuras poblacionales obtenidas en el campo para evaluar qué tanto las reglas establecidas en el paso 3 pueden generar estructuras poblacionales similares a las observadas.
- 6) Siguiendo un procedimiento heurístico, generar diferentes reglas (paso 3) y evaluarlas (paso 5) hasta obtener aquéllas que produzcan un grado satisfactorio de similitud entre las estructuras poblacionales generadas y las observadas. Dichas reglas nos permitirían discernir la forma en la que la dinámica demográfica de la especie bajo estudio se ha visto afectada por el DAC.

La descripción detallada del modelo utilizado se encuentra en el siguiente capítulo.

### **1.3 Las cactáceas endémicas de México**

Como bien se sabe, México es un país con una gran diversidad de especies, muchas de las cuales se encuentran bajo peligro real de extinguirse. Las cactáceas, en particular, se encuentran enlistadas en el Apéndice II de CITES y muchas de ellas están catalogadas como “sujetas a protección especial”, “amenazadas” o “en peligro de extinción” por la Norma Oficial Mexicana (NOM-059-ECOL-2001; SEMARNAT, 2002). Esta condición de vulnerabilidad hace urgente conocer el estado actual de dichas especies.

Para el caso de muchas de las cactáceas de México, se ha sugerido que éstas son altamente vulnerables al disturbio, dada su baja tasa de crecimiento poblacional, su alta longevidad y su bajo reclutamiento (Hernández y Godínez, 1994; Contreras y Valverde, 2002) que determinan un restablecimiento demográfico extremadamente lento tras una perturbación. Estas plantas se enfrentan a fenómenos de DAC tales como la degradación y fragmentación del hábitat por

sobrepastoreo, erosión y extracción de recursos maderables y no maderables, así como al saqueo de ejemplares para su comercio ilegal nacional e internacional (Hernández y Godínez, 1994; Martorell y Peters, 2005). El hecho de que muchas cactáceas sean endémicas de México y presenten una reducida distribución geográfica (Arias, 1993) es un factor que hace más relevante el estudio del efecto que sobre las especies tiene el DAC, pues se espera que poblaciones reducidas corran un mayor riesgo de extinguirse por factores externos como el DAC (Kunin y Gaston, 1997; sin embargo ver Kessler, 2001).

Frente a esto, el conocimiento de la demografía de estas especies es el medio para entender y evitar la extinción de las especies, pues es la merma de las poblaciones la que genera el riesgo de extinción (Vega y Peters, 2003). Enfocar así los esfuerzos de investigación en la comprensión de la dinámica poblacional ofrecería una herramienta para conservar a estas especies (Godínez-Álvarez *et al.*, 2003). Además, dado que muchas de ellas coexisten con el hombre y se ven afectadas por él, es importante caracterizar la respuesta (idiosincrática) que cada especie tendrá ante el disturbio humano, si se tiene como objetivo el conservarlas *in situ*.

*Mammillaria dixanthocentron* es un ejemplo de una cactácea que se enfrenta a DAC debido a la degradación de su ambiente (Peters y Martorell, 2000) que se presenta en distintas intensidades debido a la mayor o menor presión humana sobre el mismo.. Esta especie está catalogada como sujeta a protección especial de acuerdo con la NOM-059-ECOL-2001 y enlistada en el Apéndice II de CITES.

# Modelo

Para evaluar las consecuencias que el DAC tiene sobre la dinámica poblacional, el enfoque que decidí utilizar fue el de la modelación de los distintos procesos demográficos (supervivencia, crecimiento y fecundidad), incorporando el efecto que tendría el DAC sobre ellos y utilizando las estructuras poblacionales observadas para encontrar aquel modelo de dinámica poblacional que mejor se ajustara a éstos.

El modelo desarrollado para modelar la dinámica demográfica de la especie bajo estudio parte del propuesto por Easterling *et al.* (2000), el cual permite modelar el cambio en el tiempo de las estructuras poblacionales suponiendo una dinámica poblacional que permanece constante a lo largo de éste. Para fines de comprensión presento en primer lugar dicho modelo y después la versión modificada que realicé de éste para incorporar el DAC.

## 2.1 Modelo Integral de Proyección

Easterling *et al.* (2000) desarrollaron el modelo integral de proyección (MIP), el cual conserva las propiedades de los modelos matriciales (Caswell, 2001), pero elimina la necesidad de agrupar a los individuos en clases discretas de estado al tratar al tamaño como una variable continua.

El MIP es un modelo que, al igual que el modelo matricial de proyección, estima diferentes atributos demográficos de importancia ecológica: la tasa de crecimiento asintótico, la distribución estable de tamaños, los valores reproductivos y las sensibilidades y elasticidades para los diferentes procesos demográficos y tamaños de los individuos (Easterling *et al.*, 2000). Este modelo, a diferencia del matricial, es más apropiado para el caso de especies que no presentan una división natural de su ciclo de vida en categorías (Tuljapurkar y Caswell, 1997); esto es importante, pues se sabe que la división artificial de un ciclo de vida continuo en categorías, como se realiza en los modelos matriciales, afecta los resultados (Enright *et al.*, 1995). En particular, las cactáceas constituyen un grupo de organismos que no presentan una división clara en los estadios que alcanzan durante su ciclo de vida, ni tampoco en sus tasas vitales a lo largo del tiempo, sino más bien un aumento paulatino en biomasa a lo largo de toda su vida.

Al igual que el modelo matricial de proyección, el MIP parte de la iteración de una función, correspondiente a la distribución por tamaños de los individuos  $n(x, t)$ , por otra que permanece constante, la función de proyección  $k(x, y)$ , también llamada el *núcleo*. Es decir, se tiene la iteración:

$$n(y, t+1) = \int_X k(x, y)n(x, t)dx; \quad x, y \in X, t \in \mathbb{N}, \quad (2.1)$$

donde  $X$  es el intervalo de posibles tamaños que puede tener un individuo y  $t$  es el tiempo (discreto). La descripción detalla del MIP original puede consultarse en su fuente original (Easterling *et al.*, 2000).

Como se puede ver, existe una correspondencia entre (1.1) y (2.1), pues la multiplicación matricial equivale a la multiplicación e integración de funciones.

Al igual que una función de distribución normal está determinada por dos parámetros ( $\mu$  y  $\sigma^2$ ), el núcleo es una función cuya forma está determinada por un conjunto de parámetros (12 en el modelo original de Easterling *et al.* (2000)). De esta forma, cuando contamos con datos de campo y buscamos ajustar el modelo a éstos, el ajuste se hace encontrando aquellos valores de los parámetros que mejor se ajusten a dichos datos.

Presento ahora la versión modificada del MIP en la que incorporé el DAC (MIPDAC).

## 2.2 Modelo Integral de Proyección modificado para incorporar el DAC

Para incorporar el DAC al MIP, modifiqué  $k(x, y)$  y  $n(x, t)$  en la ecuación (2.1) para que dependieran del DAC, de modo que la dinámica poblacional,  $k(x,y)$ , se viera afectada por el disturbio, haciendo por ende depender a las estructuras poblacionales,  $n(x, t)$ , del DAC.

Esta modificación tiene repercusiones sobre la ergodicidad del modelo (Ellner y Rees, 2007), que impiden hablar de la tasa de crecimiento asintótico, la distribución estable de tamaños, etc., atributos asociados comúnmente al MIP. Esta limitante no afecta el desarrollo de mi trabajo, pues el interés de este trabajo se enfoca en las consecuencias que tiene el DAC sobre la dinámica de una población sujeta al gradiente de DAC observado; y, en este contexto, no tiene sentido tratar de indagar el comportamiento asintótico del modelo.

En primer lugar, hay que dejar clara la relación que existe entre el gradiente de DAC, la cronosecuencia de sitios de estudio y el tiempo. Como se dijo en el capítulo anterior, el DAC es un proceso paulatino que hace inviable su estudio en el tiempo; por ello recurrí a una cronosecuencia de sitios de estudio sometidos a diferente intensidad de DAC para reconstruir la dinámica poblacional a lo largo de todo el gradiente de DAC. Así, el gradiente de DAC está en función del tiempo; *i. e.*, el gradiente  $\theta$  se puede expresar como  $\theta(t)$ . Para este trabajo, la función elegida para describir el gradiente fue la función identidad, ( $\theta(t) = t$ ), donde cada unidad de DAC

corresponde a un año. Esta función se agregó a  $k(x,y)$  como un nuevo eje  $\Theta \subset \aleph$  (subconjunto de los números naturales) y sustituyó a  $t$  en  $n(x, t)$ . Este nuevo eje  $\Theta$  consiste en los números naturales que van desde el valor (entero) de DAC asociado a la población menos perturbada,  $\theta(t_0)$ , hasta el valor  $\theta(t_k)$  que corresponde a la intensidad de DAC de la población más perturbada (suponiendo que se estudian  $k$  poblaciones). Dada la definición anterior de  $\theta(t)$ , el tiempo que se requeriría para que un sitio con un DAC de  $\theta(t_0)$  presentara un DAC de  $\theta(t_k)$  sería de  $t_k-t_0$  años.

La forma como se incorpora el DAC al MIP es la siguiente: una vez que se tenga determinada la forma funcional del núcleo en los ejes de tamaño ( $X^2$ ), se tendrá un conjunto de parámetros que determinan al núcleo en estos ejes; estos parámetros se convertirán en funciones dependientes del DAC, de forma que el núcleo, en última instancia, dependerá del DAC.

Dado lo anterior y partiendo de la ecuación (2.1), la versión modificada del MIP que para este trabajo realicé (MIPDAC) se expresa formalmente mediante la siguiente ecuación:

$$n(y, \theta(t+1)) = \int_X k(x, y, \theta(t)) n(x, \theta(t)) dx, \quad x, y \in X, \theta(t) \in \Theta, \quad (2.2)$$

donde  $X$  es el intervalo de posibles tamaños de los individuos,  $\Theta$  es el conjunto de valores que puede tomar el DAC,  $n(x, \theta(t))$  es la estructura poblacional correspondiente a una intensidad de DAC  $\theta(t)$ , y  $k(x, y, \theta(t))$  contiene la información sobre el cambio en los procesos demográficos de una intensidad de DAC,  $\theta(t)$ , a la siguiente. Hay que observar que  $X \subseteq \mathfrak{R}^+$  (*i. e.*, son números reales positivos) y  $\Theta \subset \aleph$  (*i. e.*, son números naturales). Como requiero que  $n(x, \theta(t))$  integre uno en el eje  $X$  para que sea así una función de densidad probabilística en este eje, después de cada iteración, la función se estandarizó (*i. e.*, se dividió entre la integral  $\int_X n(x, \theta(t)) dx$ ). Nótese

entonces que  $n(x, \theta(t))$  no nos dice la cantidad de individuos en la población que tienen un tamaño  $x$ , sino la probabilidad de encontrar un individuo con dicho tamaño.

El núcleo  $k(x, y, \theta(t))$  es una función que contiene la información sobre los distintos procesos dinámicos; por lo cual se descompone en dos funciones: la de la transición de los individuos de un tamaño  $x$  a otro  $y$ ,  $p(x, y, \theta(t))$ , y la de la fecundidad expresada como el número promedio de descendientes de tamaño  $y$  producidos por un individuo de tamaño  $x$ ,  $f(x, y, \theta(t))$ ; *i. e.*,

$$k(x, y, \theta(t)) = p(x, y, \theta(t)) + f(x, y, \theta(t)). \quad (2.3)$$

De igual manera,  $p(x, y, \theta(t))$  se descompone en dos funciones: la que corresponde a la probabilidad de supervivencia de un individuo de tamaño  $x$ ,  $s(x, \theta(t))$ , y la que describe la probabilidad de que un individuo de tamaño  $x$  crezca (decrezca o permanezca) a un tamaño  $y$ ,  $g(x, y, \theta(t))$ . Esta función, al ser una probabilidad, corresponde, para cada valor de  $x$ , a una función de distribución de probabilidad en  $X$ . Entonces la función de transición queda de la siguiente forma:

$$p(x, y, \theta(t)) = s(x, \theta(t))g(x, y, \theta(t)). \quad (2.4)$$

Similarmente, la función de fecundidad se puede descomponer en la función que describe el número promedio de descendientes de un individuo de tamaño  $x$ ,  $f_1(x, \theta(t))$ , y la que describe la probabilidad de que el descendiente de un adulto de tamaño  $x$  tenga un tamaño  $y$ ,  $f_2(x, y, \theta(t))$ . Al igual que  $g(x, y, \theta(t))$ ,  $f_2(x, y, \theta(t))$  corresponde, para cada valor de  $x$ , a una función de distribución de probabilidad en  $X$ . Así,

$$f(x, y, \theta(t)) = f_1(x, \theta(t))f_2(x, y, \theta(t)). \quad (2.5)$$

De esta forma tenemos que el núcleo está determinado por cuatro funciones:  $s(x, \theta(t))$ ,  $g(x, y, \theta(t))$ ,  $f_1(x, \theta(t))$  y  $f_2(x, y, \theta(t))$ .

La razón por la cual las funciones se suman o multiplican se debe a razones lógicas.  $p(x, y, \theta(t))$  y  $f(x, y, \theta(t))$  son funciones que se suman ya que ambas contribuyen al total de individuos de tamaño  $y$  en la estructura poblacional  $n(y, \theta(t+1))$ :  $p(x, y, \theta(t))$  con individuos existentes que transitaron de tamaño  $x$  y  $f(x, y, \theta(t))$  con individuos nuevos. Por su parte, las probabilidades  $s(x, \theta(t))$  y  $g(x, y, \theta(t))$  se multiplican pues sólo si un individuo sobrevive puede modificar su tamaño. Finalmente,  $f_1(x, \theta(t))$  y  $f_2(x, y, \theta(t))$  se multiplican porque sólo hasta que el individuo se

reproduce y se estima el número de descendientes que produjo es que se puede estimar la fracción de éstos que tendrán un tamaño  $y$ .

Al igual que el MIP propuesto originalmente por Easterling *et al.* (2000), la forma funcional que las cuatro funciones que integran al núcleo pueden tomar es bastante flexible: el modelo admite cualesquiera funciones para  $s(x, \theta(t))$  y  $f_1(x, \theta(t))$  y cualesquiera distribuciones para  $g(x, y, \theta(t))$  y  $f_2(x, y, \theta(t))$ . La elección de dichas funciones dependerá de los datos con los que se cuente para la especie bajo estudio.

## 2.3 Estimación de los parámetros a partir de las estructuras poblacionales observadas

Una vez que se tienen seleccionadas las funciones que determinan al núcleo, se procede a utilizar las estructuras poblacionales observadas para encontrar aquella dinámica poblacional que iría generando dichas estructuras a medida que aumentara la intensidad del DAC. Esto se hace comparando las estructuras poblacionales generadas por la iteración del modelo (ecuación (2.2)) con las estructuras poblacionales observadas en cada valor de intensidad de DAC observado. La forma como se comparan dichas estructuras es mediante la verosimilitud de las estructuras generadas con respecto a los tamaños de los individuos muestreados en cada sitio de estudio.

Así, dado un conjunto de valores de los parámetros que determinan al núcleo (y que determinan, en última instancia, una dinámica poblacional particular afectada por el DAC), se calcula la verosimilitud mediante la siguiente ecuación:

$$l(\rho_1, \dots, \rho_n) = \sum_{i=1}^k \sum_{j=1}^{m_i} \ln(n(x_{ij}, \theta(t_i) | \rho_1, \dots, \rho_n)), \quad (2.6)$$

donde  $n$  es el número de parámetros que determinan al núcleo,  $m_i$  es el número de individuos muestreados en el  $i$ -ésimo sitio,  $x_{ij}$  es el tamaño del  $j$ -ésimo individuo del  $i$ -ésimo sitio y

$n(x_{ij}, \theta(t_i) | \rho_1, \dots, \rho_n)$  es la estructura poblacional correspondiente a la intensidad observada de DAC  $\theta(t_i)$  generada a partir de los parámetros  $\rho_1, \dots, \rho_n$  y evaluada en  $x_{ij}$ .

De esta manera, aquel conjunto de parámetros que maximice el valor de la función de verosimilitud será aquél que determine una dinámica poblacional afectada por el DAC que genere las estructuras poblacionales más parecidas a las observadas en el campo. Esta dinámica poblacional nos dirá cómo el DAC va modificando los distintos procesos demográficos (supervivencia, crecimiento y fecundidad) a medida que éste aumente en intensidad.

La manera como dicho máximo se encuentre dependerá de la heurística de exploración del espacio paramétrico que utilicemos. Para este trabajo la heurística elegida fue la optimización por enjambre de partículas (sección 3.3.4).

# Método

Una vez establecido de manera estructural el MIPDAC, paso ahora a mostrar la aplicación del mismo a una especie particular: *Mammillaria dixanthocentron*. Para esta especie cuento con dos tipos de información: datos sobre la dinámica demográfica de dos poblaciones y datos sobre la estructura de tamaños de 10 poblaciones. El primer conjunto de datos se utilizará para establecer de manera explícita la forma funcional de las funciones que integran el núcleo. Una vez establecido el modelo, las estructuras poblacionales observadas se utilizarán para encontrar aquella dinámica poblacional afectada por el DAC que mejor se ajusta a ellas, obteniéndose así una dinámica poblacional que explica cómo el DAC ha afectado los procesos demográficos de *M. dixanthocentron* a medida que éste fue incrementando su intensidad en el tiempo.

### 3.1 Modelación de la dinámica poblacional de acuerdo al tamaño de los individuos y al DAC

Utilicé el conjunto de datos sobre la dinámica poblacional de *M. dixanthocentron* (10 dinámicas demográficas correspondientes al estudio de dos poblaciones durante cinco años) para determinar la estructura del núcleo en los ejes de tamaño  $X^2$  (*i. e.*, las funciones  $s(x)$ ,  $g(x, y)$ ,  $f_1(x)$  y  $f_2(x, y)$ ). Nótese que el eje  $\Theta$  no está involucrado (*i. e.*, el DAC no se considera en este momento).

El enfoque que utilicé para modelar la función  $s(x)$  fue el de los modelos lineales generalizados (MLGs; Dobson, 2002; McCullagh y Nelder, 1989). En estos modelos se tiene una función  $f$ , denominada de *enlace*, que cumple con la condición de que

$$f(Y) = \mu(X) + \varepsilon, \quad (3.1)$$

donde  $Y$  es el vector de datos de supervivencia de los individuos,  $X$  es el vector de datos de tamaño de éstos,  $\mu(X)$  es una función lineal de los datos  $X$  y  $\varepsilon$  es un error que sigue una función de distribución perteneciente a la familia exponencial. La elección de las funciones de enlace  $f$  y de la distribución del error  $\varepsilon$  depende del proceso demográfico particular a modelar. De lo anterior se desprende que, para que el MLG esté completamente definido, debo especificar las funciones que lo integran, *i. e.*, las funciones que integran la ecuación (3.1).

Hay que resaltar que para incorporar al núcleo la función  $s(x)$  sólo me interesa conocer la forma de las funciones  $f(Y)$  y  $\mu(X)$ , pues el error modela la distribución de los individuos alrededor de  $f(Y) = \mu(X)$  y aquí sólo se usa con fines de estimación estadística, mientras que lo que me interesa es conocer el comportamiento promedio de los individuos para cada tamaño particular  $x$ .

Para la función  $s(x)$  la función de enlace logit es la más adecuada, pues restringe los valores del modelo lineal a valores de la imagen entre 0 y 1. Al ser la supervivencia un fenómeno binario, la distribución  $\varepsilon$  más adecuada es la de Bernoulli, pues sólo toma valores de 0 y 1.

Una vez definidas las funciones de enlace y de distribución del error, postulé distintas funciones  $\mu(X)$  y verifiqué cómo se ajustaban a los datos, eligiendo aquella que tuviera el mejor

ajuste global para los 10 conjuntos de datos. Evalué el desempeño de estas funciones utilizando máxima verosimilitud vía el criterio de información de Akaike (1974), que permite elegir un modelo de entre varios que difieren en el número de parámetros que utilizan; aquel modelo que presenta un valor más bajo del criterio es aquel que se considera como el mejor. El programa que utilicé para realizar este procedimiento fue GLIM 4.0. Los resultados del mismo se resumen en la tabla 3.1.

Tabla 3.1. Valores que tomó el criterio de información de Akaike para diferentes funciones  $\mu(X)$  para la modelación de la supervivencia ( $s(x)$ ) de *M. dixanthocentron* en dos sitios de estudio durante cinco años. Se resaltan en negritas los mejores ajustes. Sólo se presentan aquellas funciones  $\mu(X)$  que mejor modelaron al menos un conjunto de datos.

Sitio		La Pedrera					La Cucharilla				
Año		2001	2002	2003	2004	2005	2001	2002	2003	2004	2005
$\mu(X)$	$a$	78.910	119.465	161.529	91.284	102.559	41.703	97.064	109.798	61.136	80.137
	$a+b/x$	80.295	119.682	119.413	91.121	99.901	39.748	95.991	102.587	61.958	<b>76.946</b>
	$a+bx+cx^2+dx^3$	81.164	123.552	152.92	<b>73.694</b>	102.331	39.714	<b>85.176</b>	107.996	63.793	84.41
	$a+b/x^3$	80.819	<b>118.367</b>	120.25	91.351	99.825	38.715	95.584	106.789	62.611	78.159
	$a+b\ln(x)$	<b>78.587</b>	121.27	138.536	86.756	<b>98.017</b>	37.214	85.859	102.827	61.134	78.885
	$a+bx^2$	79.219	120.930	162.753	93.129	102.208	<b>35.171</b>	93.931	111.773	59.823	81.374
	$a+b\ln(x)+c\ln^2(x)$	80.540	122.509	<b>118.72</b>	88.732	100.016	38.682	85.637	<b>99.936</b>	63.003	79.780
	$a+b\ln(x)+cx^3$	79.956	122.624	140.535	88.755	99.772	37.974	87.526	104.766	<b>55.439</b>	80.761

La función  $\mu(X)$  que mostró un mejor ajuste para todos los sitios y años fue  $a+b\ln(x)$ , por lo que tenemos que dos parámetros ( $a$  y  $b$ ) describen la supervivencia de los individuos.

De esta manera,  $s(x)$  es de la forma:

$$\log\left(\frac{s(x)}{1-s(x)}\right) = a + b\ln(x) \quad (3.2)$$

Para el caso de la funciones de crecimiento  $g(x, y)$ , al ser una función que se distribuye en  $X^2$ , se debe utilizar un enfoque diferente al de los MLG ya que lo que necesitamos ahora es una función de distribución que describa la probabilidad de un individuo de pasar de un tamaño  $x$  a

uno  $y$ . Esta función describe, en el caso de la función de crecimiento, cómo se distribuyen en el continuo de tamaños  $X$  los individuos que originalmente tenían un tamaño  $x$ .

Dada la flexibilidad del modelo en cuanto a la forma funcional del núcleo, que permite introducir cualesquiera funciones de distribución para describir a  $g(x, y)$  (Easterling *et al.*, 2000), evalué el desempeño que tenían distintas funciones de distribución que compartían el hecho de tener un dominio positivo (*i. e.*, la función tomaba valores  $x$  en  $\mathfrak{R}^+$ ), ya que el tamaño de los individuos sólo toma valores positivos.

Se evaluó el desempeño de diferentes funciones de distribución vía el coeficiente de determinación ( $r^2$ ), además de la fracción de los datos observados *vs.* la fracción correspondiente de datos esperados en varias bandas de confianza vía una prueba de  $\chi^2$  y una evaluación visual del desempeño general de la función y del desempeño particular para distintos tamaños de individuos. La función que mejor modeló el crecimiento fue la distribución gamma (Tabla 3.2), pues presentó valores buenos en todo lo anterior.

Sin embargo, para modelar la función  $\mu(X)$ , un modelo lineal no dio buenos ajustes, por lo que tuve que postular funciones no lineales. Dado que el GLIM sólo permite hacer modelos lineales y, para el caso de la gamma, el GLIM necesita mantener el valor del parámetro  $\alpha$  constante (lo cual limita la forma que dicha distribución puede tomar), utilicé el programa Excel para evaluar el desempeño de las funciones no lineales propuestas, vía la herramienta incorporada *Solver*. Esta herramienta permite encontrar el óptimo de una función objetivo, modificando los valores de ciertas variables. En mi caso se trató de encontrar los valores del modelo no lineal que dieran una máxima verosimilitud con respecto a los datos observados.

Como la función gamma es descrita por dos parámetros,  $\alpha$  y  $\beta$ , y dado que para cada tamaño de individuos ( $x$ ) deseo aproximar una función gamma particular, es necesario describir el comportamiento de estos dos parámetros a lo largo del continuo de tamaños  $X$ .

Evalué el desempeño de diferentes funciones no lineales para modelar  $\alpha(x)$  y  $\beta(x)$  vía el criterio de información de Akaike (1974). Adicionalmente, evalué la fracción de los datos observados *vs.* la fracción correspondiente de datos esperados en varias bandas de confianza, vía una prueba de  $\chi^2$ . El modelo que decidí elegir fue aquél que presentó valores moderadamente buenos en ambos objetivos para los dos sitios de estudio en dos años que fueron seleccionados por contener datos muy diferentes entre sí. Con esto se buscaba una función que pudiese ajustarse bien a los diferentes comportamientos de la especie en cualquier sitio y cualquier año. Todos los

modelos presentaron un valor de  $p > 0.05$  en la prueba  $\chi^2$  para los dos sitios y los dos años; sin embargo, el modelo  $\Gamma(x; \alpha(x), \beta(x))$ ,  $\alpha(x) = a+bx$ ,  $\beta(x) = (cx^d+e)/\alpha(x)$ , fue el que presentó mejores valores de  $\chi^2$  (Tabla 3.3), por lo que fue éste el que se eligió para modelar el crecimiento. Obsérvese entonces que para la función  $g(x, y)$  se requieren cinco parámetros ( $a, b, c, d$  y  $e$ ).

Tabla 3.2. Evaluación del desempeño de distintas funciones de distribución para modelar el crecimiento de los individuos usando información de *M. dixanthocentron* de La Cucharilla en 2001. Se resalta en negritas el modelo elegido.

Modelo*	$\chi^2(x; k)$ , $k$ variable	$\Gamma(x; \alpha, \beta)$ , $\alpha$ fija, $\beta$ variable	<b><math>\Gamma(x; \alpha, \beta)</math></b> , <b><math>\alpha</math> y <math>\beta</math></b> <b>variables</b>	$B(x; \alpha, \beta)$ , $\alpha$ fija, $\beta$ variable	$B(x; \alpha, \beta)$ , $\alpha$ y $\beta$ variables	$\log N(x; \mu, \sigma)$ , $\mu$ variable, $\sigma$ fija	$\log N(x; \mu, \sigma)$ , $\mu$ y $\sigma$ variables
$r^2$	0.7924	0.7676	0.7440	0.9279	0.9284	0.6840	0.7751
Valor $p$ de la prueba $\chi^2$	0.0001	0.3743	0.3488	0.0905	0.1584	0.4841	0.3601
Desempeño general	medio	malo	bueno	malo	bueno	malo	malo
Desempeño individuos chicos	bueno	bueno	bueno	bueno	bueno	bueno	bueno
Desempeño individuos grandes	bueno	malo	medio	malo	medio	malo	malo
Desempeño individuos medianos	medio	bueno	bueno	medio	bueno	bueno	bueno

\* El formato del modelo es  $f(x; \theta)$  donde  $f$  es una función de distribución,  $x$  es el tamaño del individuo y  $\theta$  es (son) el (los) parámetro(s) que determina(n) a la función de distribución;  $\chi^2 =$  ji-cuadrada,  $\Gamma =$  gamma,  $B =$  beta y  $\log N =$  log normal.

De esta manera la función de crecimiento  $g(x, y)$  es de la forma:

$$g(x, y) = y^{\alpha(x)-1} \frac{\beta(x)^{\alpha(x)} e^{-\beta(x)y}}{\Gamma(\alpha(x))}, \quad (3.3)$$

donde  $\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$ ,  $\alpha(x) = a+bx$ ,  $\beta(x) = (cx^d+e)/\alpha(x)$ .

Tabla 3.3. Valores que tomó el criterio de información de Akaike para distintos modelos que describen el comportamiento de los parámetros de la distribución gamma en la modelación del crecimiento de los individuos usando información de *M. dixanthocentron* de dos sitios de estudio durante dos años. Se resaltan en negritas los mejores ajustes.

Sitio		La Pedrera		La Cucharilla	
Año		2001-2002	2002-2003	2001-2002	2002-2003
Modelo	$\Gamma(x; \alpha(x), \beta(x)),$ $\alpha(x) = a+bx$ $\beta(x) = c+dx$	369.669	362.018	248.200	252.417
	$\Gamma(x; \alpha(x), \beta(x)),$ $\alpha(x) = a+bx$ $\beta(x) = x/(acx+\alpha d)$	381.585	371.014	252.248	269.743
	$\Gamma(x; \alpha(x), \beta(x)),$ $\alpha(x) = a+bx$ $\beta(x) = (cx^d+e)/\alpha(x)$	<b>365.023</b>	355.447	255.500	<b>242.835</b>
	$\Gamma(x; \alpha(x), \beta(x)),$ $\beta(x) = a+bx$ $\alpha(x) = (ax^b+c)/\beta(x)$	365.462	<b>352.251</b>	<b>241.296</b>	246.794

Para el caso de la función  $f_1(x)$ , nuevamente se utilizó un MLG. La función de enlace  $f(Y)$  utilizada en este caso fue la logarítmica, ya que restringe los valores del modelo lineal a valores de la imagen en los reales positivos; y la distribución  $\epsilon$  más adecuada fue la de Poisson, pues ésta describe el comportamiento del error en un conteo. Hay que considerar que no todo individuo presenta fecundidad, por lo que la función  $f_1(x)$  tiene un valor de cero antes de los  $7.5 \text{ cm}^3$ , tamaño aproximado al cual un individuo de *M. dixanthocentron* tiene la posibilidad de reproducirse.

Postulé luego distintas funciones  $\mu(X)$  y verifiqué la manera en la que se ajustaban los datos, eligiendo aquella que tuviera el mejor ajuste global para los 10 conjuntos de datos del mismo modo como se hizo con la función  $s(x)$ . Los resultados del ajuste se resumen en la tabla 3.4.

La función  $\mu(X)$  que mejor se ajustó a los datos fue  $a+b\ln(x)+c\ln^2(x)$ ; sin embargo, dado que esta función involucra tres parámetros y mi intención era utilizar el mínimo de parámetros en

el modelo, consideré utilizar en su lugar la función  $a+b\ln(x)$  que presentó, en general, ajustes mejores que  $a+b\ln^2(x)$ . Por lo tanto, la fecundidad se describe adecuadamente sólo con dos parámetros ( $a$  y  $b$ ).

Tabla 3.4. Valores que tomó el criterio de información de Akaike para diferentes funciones  $\mu(X)$  para la modelación de la fecundidad ( $f_1(x)$ ) de *M. dioxanthocentron* en dos sitios de estudio durante cinco años. Se resaltan en negritas los mejores ajustes.

Sitio		La Pedrera					La Cucharilla				
Año		2001	2002	2003	2004	2005	2001	2002	2003	2004	2005
$\mu(X)$	$a$	560.686	No hubo reproducción	474	309.84	380.3	376.33	642.73	1177.4	458.32	1246.8
	$a+bx$	232.699		264.857	156.141	168.942	338.818	543.474	963.618	317.21	897.119
	$a+b\ln(x)$	166.481		304.841	116.613	253.273	233.748	381.722	652.653	182.817	423.139
	$a+bx+cx^2$	200.34		216.052	147.61	159.314	<b>129.669</b>	460.744	879.545	<b>159.664</b>	461.969
	$a+bx+c\ln(x)$	167.281		169.052	117.312	129.836	159.047	364.995	642.544	165.777	<b>345.576</b>
	$a+bx^2+c\ln(x)$	167.281		171.057	<b>115.908</b>	129.836	146.514	370.704	652.252	160.466	347.077
	$a+b\ln(x)+c\ln^2(x)$	<b>165.878</b>		<b>165.044</b>	117.312	<b>128.031</b>	177.897	<b>359.586</b>	<b>614.22</b>	171.289	347.277
	$a+b\ln^2(x)$	172.6255		357.4256	143.4057	163.3672	258.226	454.2755	888.8788	297.1565	640.3243

De esta manera  $f_1(x)$  es de la forma:

$$\log(f_1(x)) = a + b\ln(x) . \quad (3.4)$$

Finalmente, para el caso de la función  $f_2(x, y)$  que describe el tamaño de los descendientes que tienen los individuos adultos, se usó la función gamma, la cual está definida sólo en el dominio de los reales positivos. Debido a la falta de información sobre el efecto del tamaño de la planta madre sobre el tamaño de su descendencia (Rojas-Aréchiga, com. pers.), se hizo el supuesto de que el tamaño de las plántulas no depende del de la planta madre. De esta manera, se consideró que los parámetros  $\alpha$  y  $\beta$  de la distribución no dependen de  $x$ . Por lo tanto, la función  $f_2(x, y)$  está determinada por dos parámetros ( $\alpha$  y  $\beta$ ) y es de la forma:

$$f_2(x, y) = y^{\alpha-1} \frac{\beta^\alpha e^{-\beta y}}{\Gamma(\alpha)}, \quad (3.5)$$

donde  $\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$ .

Obsérvese entonces que para que las funciones  $s(x)$ ,  $f_1(x)$  y  $f_2(x, y)$  queden determinadas se requieren seis parámetros por cada una ( $a$  y  $b$  en la funciones  $\mu(X)$  de  $s(x)$  y  $f_1(x)$ , y  $\alpha$  y  $\beta$  en la función  $f_2(x, y)$ ), que, sumandos a los cinco parámetros que determinan a  $g(x, y)$ , hacen un total de 11 parámetros que determinan el comportamiento del núcleo en los ejes X y Y.

El comportamiento de estos 11 parámetros a lo largo del eje  $\Theta$  se decidió, para mantener un modelo simple, que fuera lineal, *i. e.*, cada uno de los parámetros anteriores se convertirá en una función de la forma:

$$\rho(\theta(t)) = a + b\theta(t) \quad (3.6)$$

donde  $\rho$  es el parámetro en cuestión y  $a$  y  $b$  son los parámetros que determinan su comportamiento a lo largo del eje  $\Theta$ . Nótese que una pendiente  $b$  distinta de cero hace que el modelo pierda la propiedad de ser ergódico (Ellner y Rees, 2007).

Tenemos así que el núcleo está determinado en sus tres ejes (X, Y y  $\Theta$ ) por un total de 22 parámetros ( $a$  y  $b$  para cada parámetro en la ecuación (3.6)). Resta ahora encontrar aquellos 22 valores de los parámetros que mejor ajusten la dinámica del núcleo a la dinámica reflejada por las estructuras poblacionales observadas.

## 3.2 Estimación de los parámetros del modelo a partir de las estructuras poblacionales observadas

Como se recordará, mi objetivo es utilizar las estructuras poblacionales observadas en la reconstrucción de los cambios en la dinámica poblacional de *M. dixanthocentron* a lo largo del gradiente de DAC. Es en este momento donde utilizaré dicha información: para estimar los 22 valores de los parámetros que determinen una dinámica demográfica que se ajuste mejor a dichas estructuras, vía la ecuación iterativa (2.2). Nótese que hablo de estructuras poblacionales observadas para distinguirlas de las estructuras poblacionales generadas por el modelo.

Queremos entonces encontrar el conjunto de 22 valores de los parámetros (en adelante a un conjunto de este tipo lo denominaré una *partícula*) que maximice el ajuste de las estructuras poblacionales generadas por el modelo a las estructuras poblacionales observadas en el campo, utilizando máxima verosimilitud. Para ello fue necesario generar un programa que cumpliera con lo siguiente:

- 1) Tomara los datos de los tamaños de los individuos de las distintas poblaciones a lo largo de la cronosecuencia, las mediciones de DAC tomadas en cada sitio de estudio y los límites de los intervalos que determinan el espacio paramétrico.
- 2) Generara la primera estructura poblacional observada. Bajo el supuesto de que la población menos perturbada refleja el estado que las restantes tuvieron en el pasado, se utiliza su estructura de tamaños como la más temprana en el tiempo ( $t_0$ ). Como la primera estructura de tamaños en la ecuación (2.2),  $n(x, \theta(t_0))$ , es una función cuya forma se desconoce, el programa debe generar un *spline* cúbico natural que la ajuste. Además, las mediciones de DAC tomadas en el campo deberán usarse para generar el gradiente de DAC en el que se modelará la dinámica poblacional.
- 3) Una vez establecido el espacio paramétrico, generara un conjunto aleatorio de  $n$  partículas, las cuales explorarán el espacio paramétrico con el objetivo de encontrar la partícula que maximice el ajuste de las estructuras poblacionales generadas por el modelo a las estructuras poblacionales observadas en el campo. Para cada partícula, el

programa evaluaría en cada intensidad de DAC la ecuación (2.7) y obtendría los 11 parámetros del núcleo correspondiente, generando posteriormente dicho núcleo. A continuación, el modelo se iteraría, partiendo del *spline* de la primera estructura observada y de los núcleos generados en el paso anterior, obteniéndose así las estructuras poblacionales esperadas a lo largo del tiempo.

- 4) Calculara la verosimilitud de la partícula comparando las estructuras generadas por el modelo, para aquellos años en los cuales los  $\theta(t)$  fueran iguales a las intensidades de DAC observadas, con las estructuras poblacionales observadas correspondientes.
- 5) Repitiera el proceso hasta encontrar aquella partícula que maximizase la verosimilitud de las estructuras.

A continuación presento cómo se incorporaron dichos apartados en un programa que se diseñó para tal efecto.

### 3.3 Algoritmo

El programa que diseñé toma la información obtenida en el campo (las muestras de tamaño de los individuos de las poblaciones estudiadas y la medición del DAC del sitio de estudio para cada población asociada) y arroja aquella partícula cuyas estructuras generadas se ajustan mejor a las estructuras observadas (un diagrama de flujo del programa MIPDAC se presenta en el Apéndice, sección A.1). Dicho programa lo codifiqué y corrí en Matlab (el código se presenta en el Apéndice, sección A.2).

Presento así el procedimiento que se siguió para generar las subrutinas del programa que realizan cada uno de los incisos presentados en la sección anterior.

### 3.3.1 Entrada de datos

Los datos que el programa requiere son de tres tipos:

1. Las muestras del tamaño de los individuos que integran cada una de las poblaciones representantes de la cronosecuencia bajo estudio. Estos datos se introducen en el programa como una matriz de  $m \times k$  entradas, donde  $m$  es el tamaño de la muestra más grande y  $k$  es el número de muestras (*i. e.*, de poblaciones estudiadas).
2. La medición del DAC bajo una escala continua (*e. g.*, la dada por Martorell y Peters, 2005) para cada uno de los sitios de estudio asociados a cada población bajo estudio. Estos datos se introducen como un vector de  $k$  entradas.
3. Los límites de los intervalos continuos que determinan los valores que pueden tomar los parámetros del modelo. Estos datos se introducen como dos vectores de  $p$  entradas, donde  $p$  es el número de parámetros que determinan al núcleo (en mi caso  $p = 11$ ): uno para los valores máximos de los parámetros y otro para los valores mínimos.

### 3.3.2 Generación de la primera estructura observada y del gradiente de DAC

Para el caso de la primera estructura poblacional observada, como ésta se utilizará como la primera estructura poblacional en el proceso iterativo dado por la ecuación (2.2), necesitamos una función de densidad continua que la ajuste. Como estructura inicial utilicé un *spline* cúbico natural. Este tipo de *spline* tiene la ventaja de que es continuo en la vecindad de los extremos de cada subintervalo de la partición, obteniéndose así una función “suave” (con continuidad  $C^2$ ), lo cual es deseable para modelar la estructura poblacional de una especie que presenta individuos en todo el continuo de tamaños posibles (Knott, 2000). Adicional a lo anterior, y dado que este *spline* no tiene por qué ser una función de distribución, éste se estandarizó para que integrara uno

en el intervalo  $[0.0035, 1200]$  (el intervalo de tamaños que los individuos de *M. dixanthocentron* presentan en la naturaleza).

El gradiente de DAC que se utilizará en la modelación de la dinámica poblacional, como se dijo arriba, parte del valor mínimo y máximo de DAC observado en el campo y comprende todos los números naturales entre dichos valores.

### **3.3.3 Generación de partículas, obtención de los valores que determinan al núcleo, generación de éstos e iteración del modelo**

Para la generación de las  $n$  partículas (conocidas como *enjambre*) que explorarán el espacio paramétrico se eligieron  $n$  puntos al azar en cada intervalo del espacio paramétrico, generándose así los interceptos  $a$  de la recta (3.6). Las pendientes  $b$  se eligieron al azar de forma que el valor  $\rho(\theta(t_k))$  se mantuviera dentro del intervalo correspondiente para ese parámetro, lográndose así que todos los valores  $\rho(\theta(t))$  estuvieran dentro del intervalo paramétrico. Estos interceptos y pendientes constituyen las  $n$  partículas del enjambre.

Entonces, para cada partícula, se obtuvieron los 11 valores de los parámetros del núcleo correspondientes a cada intensidad de DAC  $\theta(t)$  usando la ecuación (3.6). Posteriormente, estos 11 valores se introdujeron en las funciones  $s(x, \theta(t))$ ,  $g(x, y, \theta(t))$ ,  $f_1(x, \theta(t))$  y  $f_2(x, y, \theta(t))$  generándose así un núcleo por cada valor de  $\theta(t)$ .

Además, para que  $g(x, y, \theta(t))$  y  $f_2(x, y, \theta(t))$  fueran funciones de distribución en el intervalo  $[0.0035, 1200]$ , éstas se estandarizaron para que integraran uno en dicho intervalo. Esto se hizo porque la función gamma (la usada para modelar a  $g(x, y, \theta(t))$  y  $f_2(x, y, \theta(t))$ ) está definida en el intervalo  $[0, \infty)$ , por lo que, de no restringirse al intervalo  $[0.0035, 1200]$ , se tendría que existirían individuos que de  $\theta(t)$  a  $\theta(t+1)$  quedarían fuera de dicho intervalo y, por lo tanto, se eliminarían de la población.

Ahora bien, hay que considerar el hecho de que, ya que el proceso de obtención de las estructuras poblacionales, dado por la ecuación (2.2), requiere de una integración en cada iteración, es necesario incorporar un método de integración numérica al programa; estos métodos discretizan y aproximan integrales. El método elegido para la discretización fue el de la cuadratura de Gauss-Legendre (ver Krommer y Ueberhuber, 1998), el cual proporciona un conjunto de valores de  $X$  (conocido como la *cuadratura*), en los cuales se va a evaluar la función a integrar para discretizarla, y un conjunto de pesos, los cuales ponderan la importancia de cada punto de la cuadratura en el valor final de la integral; de tal forma que, dada una función  $f(x)$ , su integral se aproxima por

$$\int f(x)dx \approx \sum_{i=1}^n w_i f(x_i), \quad (3.7)$$

donde  $n$  es el tamaño de la cuadratura, las  $x_i$  son los valores de  $X$  en los que se evalúa la función  $f$  y  $w_i$  es el peso correspondiente a cada  $x_i$ .

Se eligió esta cuadratura porque presenta un mayor refinamiento (*i.e.*, una mayor cantidad de intervalos) en el extremo derecho e izquierdo del continuo de tamaños. Así, la cuadratura hace una partición más fina para los individuos de tallas más pequeñas y más grandes. Esto es importante porque: 1) los individuos más pequeños sólo pueden pasar, de un año al siguiente, a tener una talla cercana a la que tenían anteriormente; si los intervalos son demasiado grandes las transiciones entre tamaños sólo podrían ocurrir en “saltos” demasiado grandes; y 2) los individuos más grandes son los que suelen presentar, en el caso de muchas cactáceas, los valores de sensibilidad y elasticidad más importantes (Godínez-Álvarez *et al.*, 2003), por lo que pequeños errores en el valor de la integral para dichos tamaños pueden tener mayor impacto en el resultado de las iteraciones del modelo.

Una vez resuelto esto y obtenidos los núcleos, utilicé el *spline* de la primera estructura poblacional y realicé la iteración dada por la ecuación (2.2) tomando, en cada iteración, el núcleo correspondiente a la intensidad de DAC correspondiente. Los núcleos se evaluaron en la cuadratura; ésta generada en el intervalo de  $X$  [0.0035, 1200] con un tamaño de la norma de la cuadratura de 150.

### 3.3.4 Cálculo y maximización de la verosimilitud: Optimación por enjambre de partículas

Para evaluar la similitud entre las estructuras generadas por el modelo y las observadas en el campo, se utilizó la verosimilitud  $l(\rho)$ , la cual, para una partícula  $\rho$  está dada por la ecuación (2.6).

Una vez calculado el valor de la verosimilitud de una partícula dada, nos enfrentamos ahora a un problema de optimización: encontrar el máximo global de una función (la de verosimilitud) que involucra 22 variables (los 22 parámetros que constituyen la partícula). Dada la complejidad de la función que quería optimar, no era factible obtener una solución analítica, por lo que tuve que recurrir a un método heurístico que explorara el espacio paramétrico y de ese modo se aproximara al máximo global.

Aunque para encontrar el máximo global se requeriría explorar todo el espacio paramétrico (Thompson, 1999), al enfrentarnos a un espacio de 22 dimensiones integradas por subconjuntos de los números reales, realizar esto implicaría ocupar muchísimo tiempo de computadora; con la adición de que dicha exploración sería muy ineficiente pues, sin el objetivo de encontrar el máximo, la computadora ocuparía mucho tiempo explorando zonas del espacio paramétrico con bajas verosimilitudes.

Por ello, hay métodos que intentan, de manera sistemática, explorar dicho espacio e intentar identificar al óptimo global. Entre éstos se encuentra la optimación por enjambre de partículas (OEP; Kennedy y Eberhart, 1995) que decidí utilizar incorporándole distintas mejoras que han sido desarrolladas para su uso (en Shi y Eberhart, 1998; Eberhart y Shi, 2001; Hu y Eberhart, 2002; Peram *et al.*, 2003; Hu *et al.*, 2004).

La OEP tiene como fundamento emular el proceso natural de la comunicación del conocimiento individual al grupo que se observa en los enjambres (o parvadas) de animales cuando se congregan, migran, forrajea, etc. con el objetivo de alcanzar un objetivo (Krusienski y Jenkins, 2005).

La heurística de la OEP es la siguiente: posicionando aleatoriamente un conjunto de partículas en el espacio paramétrico, se les calcula su verosimilitud y se les asigna

arbitrariamente una velocidad que las llevará a cambiar de posición. Posteriormente, la velocidad de cada partícula va cambiando dependiendo tanto de la ubicación previa de la partícula, como de la ubicación de la partícula líder (aquella con la mayor verosimilitud), de forma que cada partícula explora el área circundante a ella, pero con una cierta tendencia a acercarse a la líder. Si la nueva posición de una partícula es mejor (en verosimilitud) a la anterior, la partícula se queda ahí; de lo contrario, regresa a su posición original. Cada vez que se realiza esto se identifica a la nueva líder. Esto se realiza hasta encontrar un óptimo.

Formalmente, la OEP parte de dos ecuaciones y de un algoritmo de tres pasos. Las ecuaciones son:

$$v_f = v_i + c_1 \alpha_1 (p_h - p_i) + c_2 \alpha_2 (p_L - p_i) \quad (3.8)$$

$$p_f = p_i + v_f \quad (3.9)$$

donde  $v_f$  y  $v_i$  son las velocidades actual y anterior de la partícula,  $p_f$  y  $p_i$  es la posición actual y anterior de la partícula,  $c_1$  y  $c_2$  son pesos asignados,  $p_h$  es la mejor posición que ha tenido la partícula en su historia,  $p_L$  es la partícula líder y  $\alpha_1$  y  $\alpha_2$  son dos número aleatorios tomados de una distribución normal con media 0.9 y desviación estándar 0.2.

El algoritmo de la OEP, en su versión más simple, consiste en:

- I. Generar aleatoriamente un conjunto  $n$  de partículas.
- II. Calcular las verosimilitudes de dichas partículas.
- III. Mientras no se alcance un óptimo:
  - A. Para cada partícula:
    1. Calcular la velocidad de la partícula de acuerdo con la ecuación (3.8).
    2. Calcular la nueva posición de la partícula de acuerdo con la ecuación (3.9).
    3. Calcular la verosimilitud de las partículas en su nueva posición.
    4. Si la nueva verosimilitud es mejor que la anterior, reubicar definitivamente a la partícula a la nueva posición. Si es menor devolverla a su posición original.
  - B. Escoger la partícula con la mayor verosimilitud como la partícula líder.

En mi versión de la OEP la mejor partícula (la óptima) se considera como la partícula líder después de que se ha repetido el paso III 250 veces o la suma de las distancias euclidianas de todas las partículas a la líder es inferior a la unidad. El número de partículas fue de 25 y los valores de  $c_1$  y  $c_2$  se inicializaron en 2 (Hu *et al.*, 2004); para el caso de  $c_1$  su valor fue decreciendo en 0.5 cada 50 repeticiones del paso III con el objetivo de que, a medida que éstas pasaran, la partícula líder fuera teniendo mayor peso en la velocidad de las demás, obligando con ello a que las partículas se fueran reuniendo más rápidamente alrededor de la líder y exploraran zonas cercanas a la líder más factibles de tener verosimilitudes altas.

Partiendo de esta versión de la OEP, las mejoras incorporadas por otros autores son las siguientes:

- 1) Siempre que se genera o actualiza una partícula se verifica que genere núcleos viables siguiendo ciertas reglas de desigualdades (Hu y Eberhart, 2002).
- 2) Como se observa en la ecuación (3.8), la velocidad de la partícula se ve influida por  $p_n$  y por  $p_L$ . Adicional a estos dos factores, se agregó un tercero: la mejor partícula de entre una vecindad de partículas cercanas a la partícula en cuestión (Peram *et al.*, 2003). A esta partícula se le denomina  $p_n$  y tiene asociada en la ecuación (3.8) un peso  $c_3$ . En mi versión de la OEP, el valor de  $c_3$  tuvo el mismo comportamiento que el de  $c_1$ .
- 3)  $v_i$  en la ecuación (3.8) se multiplica por un peso  $w$  que agrega un factor de aleatoriedad a la velocidad, lo cual se ha demostrado que da mejores resultados en el proceso de optimación (Shi y Eberhart, 1998). En mi versión de la OEP, para cada partícula y cada parámetro,  $w$  es un número aleatorio tomado de una distribución normal con media 0.9 y desviación estándar 0.2.

Adicionalmente, incorporé mejoras *ad hoc*:

- 4) Cuando una partícula se genera o actualiza se verifica que ésta se encuentre dentro del espacio paramétrico, *i. e.*, que cada parámetro se ubique dentro del intervalo definido para él.

- 5) Dado que en la superficie de verosimilitud pueden existir óptimos locales, es posible que en una corrida del algoritmo no se logre encontrar el óptimo global, sino que la partícula líder permanezca en un óptimo local. Por ello el algoritmo se tiene que repetir varias veces, eligiéndose como óptimo global al mejor de entre los óptimos encontrados en cada corrida. A cada repetición del paso III se le llama un *ciclo*. En mi versión de la OEP el número máximo de ciclos por corrida fue de 250.
- 6) Para evitar que una partícula, al llegar a un extremo de alguno de los intervalos en los que estaban definidos los parámetros, se quedara estática en dicho parámetro, se incorporó una subrutina de rebote que cambia el signo de la velocidad de la partícula en ese parámetro de manera que ésta siga explorando dicho intervalo.
- 7) A partir del ciclo 100, para evitar que la partícula con la peor verosimilitud siga explorando una zona de baja verosimilitud, ésta se reubica en la posición de la partícula líder ( $p_L$ ). Con ello se reduce el número de partículas que exploran zonas diferentes a la de la partícula líder y se explora con mayor intensidad el área que circunda a ésta.

Los intervalos donde dichos parámetros pueden tomar valores se establecieron como  $\bar{\rho} \pm 4S_\rho$ , *i. e.*, cuatro desviaciones estándar alrededor de la media de los valores de los parámetros obtenidos en el procedimiento de la sección 3.1 a partir de los datos demográficos de los dos sitios de estudio durante cinco años. Para garantizar que las partículas siempre tuvieran valores dentro de dichos intervalos se restringieron los posibles valores de los parámetros  $a$  y  $b$  de tal suerte que el intercepto  $a$  de las rectas (3.6) se encontrara dentro del intervalo correspondiente; sin embargo, la pendiente  $b$  asociada debía ser tal que los valores que tomara el parámetro  $\rho$  a lo largo del eje  $\Theta$  se encontraran dentro del intervalo de valores posibles para dicho parámetro. Lo anterior ocurre cuando el valor de la función  $\rho(\theta(t))$  en el máximo de  $\Theta$  se encuentra dentro del intervalo de valores de  $a$ .

También hay que notar que el exponente  $d$  en la función  $g(x,y)$  fue un caso particular: a partir de los estimados de los parámetros  $c$  y  $d$  para los sitios La Pedrera y La Cucharilla para los cinco años de estudio, se observó que existe una fuerte correlación ( $r^2 = 0.9105$ ) entre dichos parámetros; esto quiere decir que el valor de un parámetro determina en gran medida el valor del otro, y la relación entre estos dos parámetros está dada por una línea recta. De esta manera,

explorar el espacio paramétrico más allá del segmento correspondiente a la recta significaría considerar combinaciones de valores de  $c$  y  $d$  que no serían biológicamente posibles. Así se decidió que el valor de  $d$  estuviera determinado por la recta dada por la regresión ( $d = 1.0671 - 0.0881c$ ). Por lo tanto, la función de crecimiento está definida sólo por cuatro parámetros. Lo anterior permitió reducir el espacio paramétrico a 20 dimensiones, lo cual permitió al programa encontrar mejores soluciones y de manera más rápida que al usar el espacio paramétrico completo.

### **3.4 Validación del modelo con datos artificiales**

Para evaluar el desempeño del programa en la detección de un máximo global usando OEP, elegí una partícula que generara núcleos viables a lo largo de un gradiente de DAC entre 0 y 100 y suponiendo un incremento anual de DAC de uno. Al azar, escogí 10 intensidades de DAC (1, 12, 18, 23, 26, 56, 68, 74, 85 y 91) y obtuve, por el proceso iterativo dado por la ecuación (2.2), las estructuras poblacionales  $n(x, \theta(t))$  correspondientes a cada intensidad. A partir de estas estructuras obtuve, a través de simulaciones Monte Carlo, 150 individuos para cada estructura que correspondieran a una muestra “observada” de una población sujeta a la intensidad de DAC correspondiente. Los datos de estos individuos conformaron una matriz de  $150 \times 10$  entradas que se introdujo en el programa junto con los 10 valores de DAC correspondientes. Se calculó la verosimilitud de la partícula generadora con respecto a estos datos y este valor, -4830.47, se consideró como la mejor verosimilitud a obtener mediante el programa.

## 3.5 Aplicación del modelo a *Mammillaria dixanthocentron*

### 3.5.1 Especie bajo estudio

*Mammillaria dixanthocentron* Backeberg (1963) es una cactácea que presenta tallos simples, de forma globosa a cortamente cilíndrica (5-30 cm de alto, 7-25 cm de ancho), con tubérculos cónicos (5-8 mm de largo, 3-6 mm de ancho), flores rosa-carmín (1.0-1.6 cm de largo, 0.9-1.2 cm de ancho), frutos rojo-verdosos (1.5-2.3 cm de largo, 0.4-0.7 cm de ancho) y semillas pardo-oscuros (0.8-1.0 cm de largo, 0.5-0.6 cm de ancho). Presenta floración entre los meses de octubre y febrero (Arias *et al.*, 1997), alcanzando su máximo en este último mes (Santini, 2007).

Esta especie es endémica del Valle de Tehuacán-Cuicatlán, en los estados de Oaxaca y Puebla, México, donde habita en bosque tropical caducifolio y bosque de *Quercus-Pinus* (Arias *et al.*, 1997) en localidades a elevaciones de 600 a 2500 m (Peters y Martorell, 2000). En estos tipos de vegetación se le encuentra usualmente como planta epífita o saxícola (Peters *et al.*, 2007). La especie se encuentra clasificada, de acuerdo con la Norma Oficial Mexicana (NOM-059-ECOL-2001; SEMARNAT, 2002), como sujeta a protección especial y se ubica en el Apéndice II de CITES. Es una especie muy susceptible ante la degradación del ambiente (Peters y Martorell, 2000) y sufre, además del DAC presente en los bosques que habita, de la extracción de ejemplares para su comercialización como planta de ornato ya que, debido a su lento crecimiento, no es de fácil cultivo (Oldfield, 1991).

## 3.5.2 Sitio de estudio

El valle de Tehuacán-Cuicatlán está situado en el sureste de Puebla y el noreste de Oaxaca, México (entre 17° 30' y 18° 55' N y entre 96° 42' y 97° 47' O). Las altitudes que abarca van de los 560 a los 2480 m s.n.m., la temperatura media es de 19.4 °C (con un intervalo de variación anual de los 10.8 a los 26.2 °C), la precipitación media anual es de 899 mm (368–1572 mm) y los climas presentes pertenecen a tres subgrupos áridos: semiárido (BS<sub>1</sub>), árido (BS<sub>0</sub>) y árido seco (BW), de acuerdo con la clasificación de Köppen modificada por García (1988).

Los datos de dinámica demográfica para *M. dixanthocentron* con los que cuento corresponden a dos poblaciones (La Pedrera y La Cucharilla) ubicadas en el municipio Concepción Buenavista, Oaxaca, México, a altitudes de 2,200 y 2,300 m s.n.m., respectivamente. En estos sitios el clima es semiárido (BS<sub>1</sub>), con una precipitación anual de 530.3 mm (temporada de lluvias de abril a octubre) y una temperatura media anual de 14.9 °C (Ureta, 2006). Los suelos son someros, alcanzando una profundidad de hasta 50 cm en las partes más altas y arboladas. La vegetación en el área donde crece *M. dixanthocentron* es, en ambas poblaciones, un bosque de *Quercus*. Este bosque, que se presenta a la manera de islas inmersas en una matriz de terreno destinado a la agricultura y/o ganadería, representa parte de la vegetación remanente en la región y se conserva principalmente en las partes altas de la zona; la población humana de la región utiliza, hasta la actualidad, estos bosques como fuentes de madera y forraje.

Los datos de estructura poblacional para *M. dixanthocentron* corresponden a 10 poblaciones ubicadas en diversos municipios del estado de Oaxaca, México. Las altitudes de los mismos varían de los 640 (Tecomavaca) hasta los 2500 m s.n.m. (Pala), lo que se refleja en la gran diversidad de hábitats en los que se encuentra la especie: los bosques tropicales caducifolios (Tecomavaca, Los Cues y Teotitlán), los bosques de *Quercus* (Teotitlán, La Pedrera, La Cucharilla, Cerro Gordo y Pala), los bosques de *Pinus* y *Quercus* (Pápalo) y los bosques de *Juniperus* (Miltepec y Tepelmeme).

La región ha tenido presencia humana desde tiempos prehispánicos (MacNeish, 1967), por lo que ha estado sometida a DAC durante un intervalo amplio de tiempo, ubicándose áreas con un alto grado de erosión por escorrentía en suelos desprovistos de vegetación. Las principales causas de DAC en el área son el pastoreo extensivo, la agricultura y la extracción de

productos no cultivados (Martorell y Peters, 2005), teniéndose en este último rubro la extracción ilegal de cactáceas globosas tales como la misma *M. dixanthocentron*.

### 3.5.3 Información demográfica disponible

La información con la que cuento para *M. dixanthocentron* es de dos tipos: datos de dinámica poblacional y datos de estructura poblacional. La siguiente es la descripción de ambos conjuntos de datos:

- Los datos de dinámica poblacional provienen de dos poblaciones (La Cucharilla y La Pedrera) seguidas a lo largo de cinco años (2001-2006). Estos datos contienen información sobre los cambios en el tamaño y la fecundidad de los individuos en dichas poblaciones. Los datos de tamaño se tomaron entre finales de julio y principios de agosto de cada año y los de fecundidad en los meses de enero y febrero. Se trazaron transectos de 50 m de largo y 4 m de ancho elegidos aleatoriamente en ambas áreas. Se marcaron aproximadamente 150 individuos en La Cucharilla y 250 en La Pedrera. Se midió la altura y el ancho de cada individuo con un vernier y el tamaño de cada individuo correspondió a la estimación de su volumen aproximado a un cilindro. La fecundidad, por su parte, se estimó a partir del número de flores observadas (Santini, 2007) y de un cálculo del número de plántulas por flor a partir de las plántulas observadas en el campo en el periodo julio-agosto del año correspondiente. Obsérvese que aquí se está haciendo el supuesto de que no existe banco de semillas, y que toda semilla presente en un año es resultado de la reproducción de ese año. Así, los datos de dinámica poblacional incluyen el incremento (o decremento) en el tamaño de los individuos de un año al siguiente, así como la fecundidad (número de plántulas) promedio por cada individuo en la población.
- Los datos de estructura poblacional provienen de 10 sitios de estudio (en orden de intensidad de DAC): Cerro Gordo, Teotitlán, Miltepec, Tecomavaca, Pápalo, La

Cucharilla, Pala, La Pedrera, Los Cues y Tepelmeme. Estos datos contienen únicamente el tamaño de los individuos en el año en que se tomaron (1999). El tamaño muestral para cada sitio fue, respectivamente: 281, 268, 55, 52, 100, 202, 125, 92, 40 y 42. La información detallada sobre la toma de estos datos se encuentra descrita en Peters y Martorell (2000).

Como el esfuerzo de muestreo que se realizó para la obtención de los datos de dinámica poblacional fue mayor que el realizado para los datos de estructura poblacional, se tiene más confianza sobre la cantidad de plántulas producidas por los individuos de las poblaciones de La Cucharilla y La Pedrera (6.56% de la población) que en las demás poblaciones. Por ello se ajustó la proporción de plántulas observadas en las estructuras poblacionales para que coincidiera con la observada en las poblaciones cuya dinámica se estudió. Para ello se agregaron plántulas (individuos de 0.0035 a 0.0898 cm<sup>3</sup>) a las demás poblaciones de forma tal que se tuviera una proporción de plántulas cercana a 6.56%.

Adicional a los datos anteriores, para cada sitio de estudio se contó con el índice de DAC para el año en el que se registró la estructura poblacional (1999). Dicho índice corresponde al propuesto por Martorell y Peters (2005), y modificado por los autores, el cual varía en una escala continua y positiva, incrementándose el índice a medida que aumenta la intensidad del DAC. Este índice está compuesto por 15 indicadores agrupados en tres categorías:

#### Ganadería

- Densidad de excretas de cabras (fracción de 10 cuadros de 1 m<sup>2</sup> donde se encontraron excretas sobre un transecto de 50 m).
- Densidad de excretas de ganado bovino y equino (medida de la misma forma que excretas de cabras pero para las excretas de cualquier otro tipo de ganado).
- Fracción de plantas ramoneadas (en un área de 50 m<sup>2</sup>).
- Caminos ganaderos (número promedio de caminos hechos por el ganado que cruzan un metro).

- Compactación del suelo por el ganado (tasa de infiltración de 250 ml de agua en un camino ganadero vs. uno no pisoteado utilizando cilindros de PVC de 10.4 cm de diámetro enterrados 4 cm en el suelo).

#### Actividades humanas

- Fracción de plantas macheteadas (medida de la misma forma que las plantas ramoneadas).
- Cobertura de caminos humanos (suma de los anchos de los caminos que cruzan un transecto de 50 m dividida entre la longitud de éste).
- Cercanía a poblaciones (inverso multiplicativo de la distancia en km a la población más cercana).
- Adyacencia a núcleos de actividad humana (presencia-ausencia).
- Cambio de uso del suelo (fracción de suelo cuyo uso se ha modificado en las zonas de estudio).
- Evidencia de incendio (presencia-ausencia).

#### Degradación del suelo

- Erosión (porcentaje de la superficie que presenta erosión grave, tales como cauces de arrastre de materiales, afloramientos de roca madre y caminos humanos profundos).
- Islas de suelo: montículos de suelo cubiertos de vegetación en una matriz de suelo fuertemente erosionado y desnudo (presencia-ausencia).
- Superficie totalmente modificada: superficie tan modificada que hace imposible o carente de significado realizar las mediciones anteriores (fracción de la superficie total).

De esta manera, utilizando análisis de componentes principales, se resumió la información de los índices previos en uno solo. Los valores del índice de DAC para los 10 sitios se reportan en la tabla 3.5.

Tabla 3.5. Valores del DAC para los 10 sitios bajo estudio.

<b>Sitio</b>	<b>Grado de DAC</b>
Cerro Gordo	28.09
Teotitlán	31.39
Miltepec	31.46
Tecomavaca	31.49
Pápalo	31.50
La Cucharilla	34.74
Pala	38.14
La Pedrera	51.58
Los Cues	56.08
Tepelmeme	89.55

Dado que los sitios Teotitlán, Miltepec, Tecomavaca y Pápalo coinciden en su valor entero, tomé a toda las poblaciones como representantes de una sola sometida a un grado de DAC de 31.

# Resultados

## 4.1 Datos artificiales

El programa, partiendo de datos de tamaño de 1500 individuos provenientes de 10 poblaciones hipotéticas representantes de 10 intensidades diferentes de DAC, reconstruyó una dinámica poblacional similar, en algunos aspectos, a la determinada por la partícula generadora, aunque también mostró diferencias sustanciales. Se realizaron 100 corridas del programa, las cuales dieron como resultado 100 partículas líderes cuyas verosimilitudes presentaron un aparente comportamiento asintótico, sobrepasando la verosimilitud de la partícula generadora (Fig. 4.1).

Las distancias (Manhattan) a las que las partículas obtenidas quedaron de la generadora fueron en general menores a las esperadas al azar (Gabourne *et al.*, 1993), aunque hubo algunas corridas en las que la partícula líder estuvo muy lejana de la generadora (Fig. 4.2).

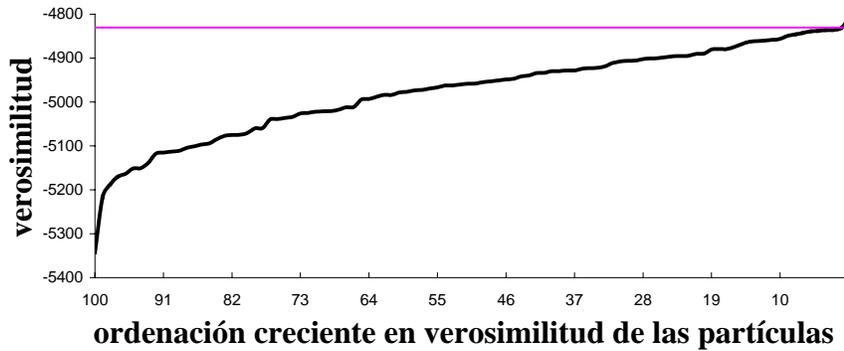


Figura 4.1. Verosimilitud de las mejores partículas obtenidas en cada corrida del programa para los datos artificiales. La línea horizontal indica la verosimilitud de la partícula generadora.

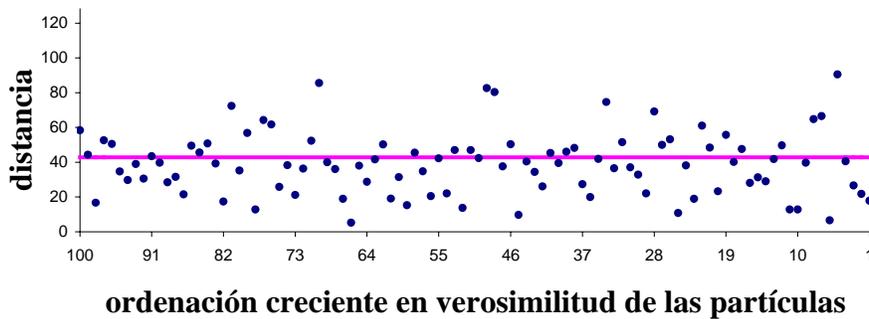


Figura 4.2. Distancia Manhattan de las mejores partículas obtenidas en cada corrida del programa con respecto a la partícula generadora. La línea horizontal indica la distancia esperada entre dos partículas generadas al azar.

El ajuste dado por la mejor partícula a las estructuras “observadas” fue muy bueno (Fig. 4.3). Sólo se presentan nueve intensidades de DAC en la figura 4.3 porque la primera estructura poblacional es con la que inicia el modelo y es igual para todas las partículas.

De acuerdo con la partícula generadora, la población sujeta a un DAC bajo posee una estructura dominada por plántulas e individuos de tallas grandes, que, al aumentar el DAC, lleva a la población hacia una estructura dominada por individuos de tallas medias y totalmente carente de individuos de tallas grandes. Este cambio en las estructuras lo logró reconstruir la mejor partícula encontrada. La prueba Kolmogorov-Smirnov de bondad de ajuste mostró que, para las intensidades de DAC 85 y 91, las estructuras generadas por la mejor partícula encontrada son estadísticamente iguales a las generadas por la partícula generadora ( $p > 0.1$ ); mientras que, para el resto de las intensidades de DAC “observadas” (12, 18, 23, 26, 56, 68 y 74), las estructuras generadas por la mejor partícula son estadísticamente diferentes a las generadas por la generadora ( $p < 0.05$ ).

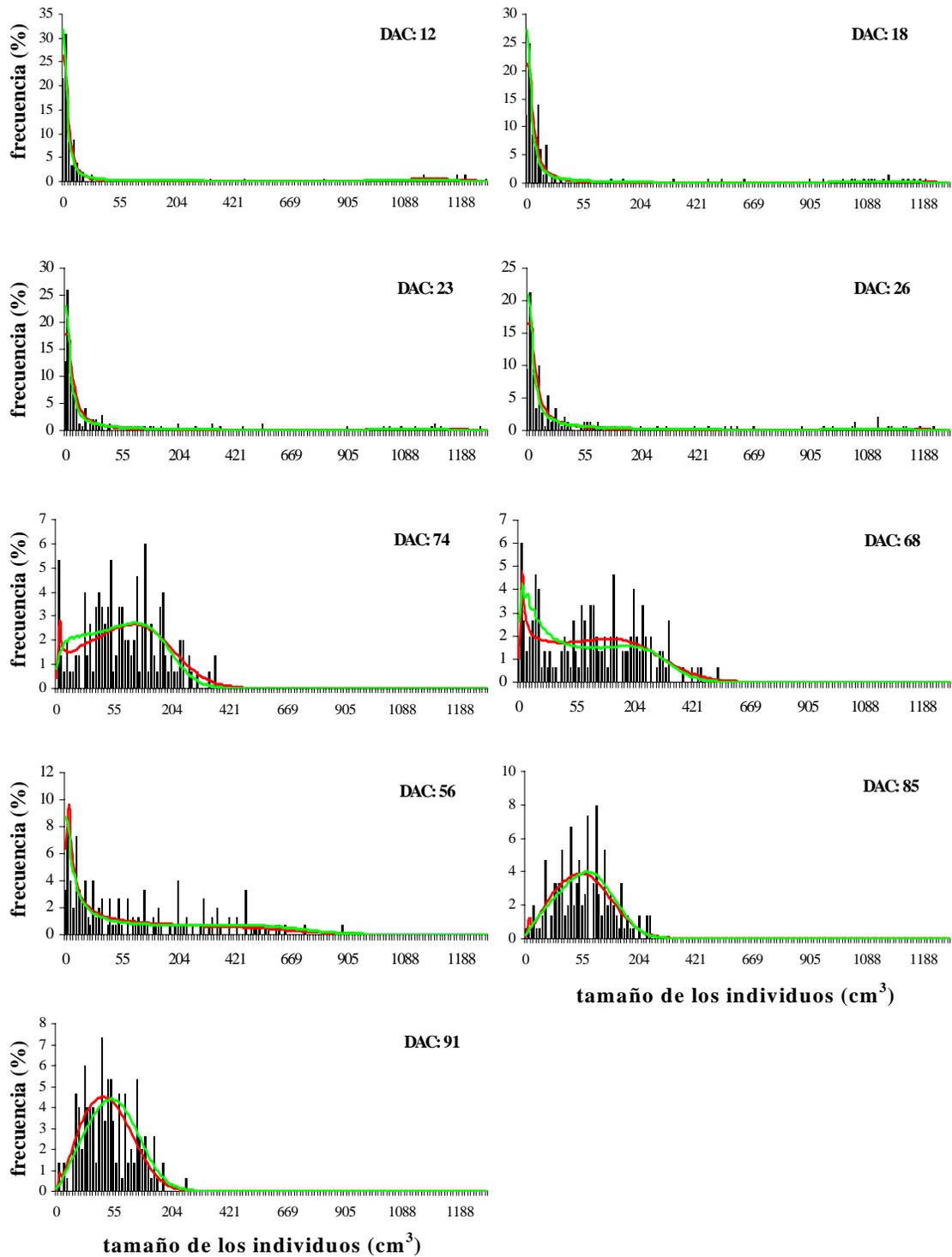


Figura 4.3. Estructura de tamaños “observada” (histograma), generada por la partícula generadora (línea oscura) y obtenida por la mejor partícula encontrada (línea clara) para las nueve intensidades de DAC.

En cuanto al ajuste de la dinámica demográfica de la mejor partícula encontrada con respecto a la dada por la partícula generadora, se tuvo un desempeño heterogéneo en los tres procesos demográficos de interés (Fig. 4.4), siendo la fecundidad la que presentó la mejor reconstrucción de la dinámica dada por la partícula generadora, y la función de supervivencia la que mostró diferencias más marcadas.

Primeramente, la supervivencia de los individuos en el caso de la partícula generadora depende del tamaño del individuo (Fig. 4.4 a): los individuos menores a  $120 \text{ cm}^3$  ven reducida su probabilidad de supervivencia a medida que aumenta el DAC; los individuos por encima de esta talla mantienen una probabilidad de supervivencia igual a uno a lo largo de todo el gradiente. Para el caso de la mejor partícula encontrada, la supervivencia, a intensidades bajas de DAC, es similar a la dada por la generadora (Fig. 4.4 b). Sin embargo, a intensidades altas, los individuos por encima de los  $80 \text{ cm}^3$  vieron su supervivencia reducida a cero a partir de una intensidad de DAC de 60.

Por su parte, el crecimiento de los individuos presentó un mejor ajuste entre la dinámica dada por la partícula generadora y la mejor partícula encontrada (Fig. 4.4 c y d), siendo éste mejor a intensidades altas de DAC, área donde la escala de crecimiento se mantuvo. Estas partículas comparten un comportamiento similar en el crecimiento frente al DAC de los individuos de tallas pequeñas: éste no afecta su crecimiento. Para los individuos de tallas medias y grandes, la partícula generadora predice un crecimiento a intensidades bajas de DAC y un decremento en el crecimiento a medida que el DAC alcanza valores más altos, siendo este patrón más pronunciado a mayor tamaño de los individuos. Por su parte, la mejor partícula encontrada muestra un patrón similar a intensidades altas de DAC, pero a intensidades bajas los individuos no crecen tanto como con la generadora.

Por último, la dinámica de la fecundidad de la mejor partícula fue, en general, similar a la dada por la partícula generadora, tanto en la forma de la superficie como en la cantidad de plántulas por individuo por año (Fig 4.4 e y f). Aunque el DAC no parece tener un efecto muy pronunciado sobre la fecundidad, un observación más detallada de las figuras 4.4 e y f muestra que, mientras que la partícula generadora arroja una fecundidad que aumenta para los individuos de talla mediana a medida que aumenta el DAC y disminuye para los de tallas grandes, la mejor partícula da lugar a una fecundidad que declina en los individuos medianos y aumenta en los individuos grandes al aumentar la intensidad del DAC.

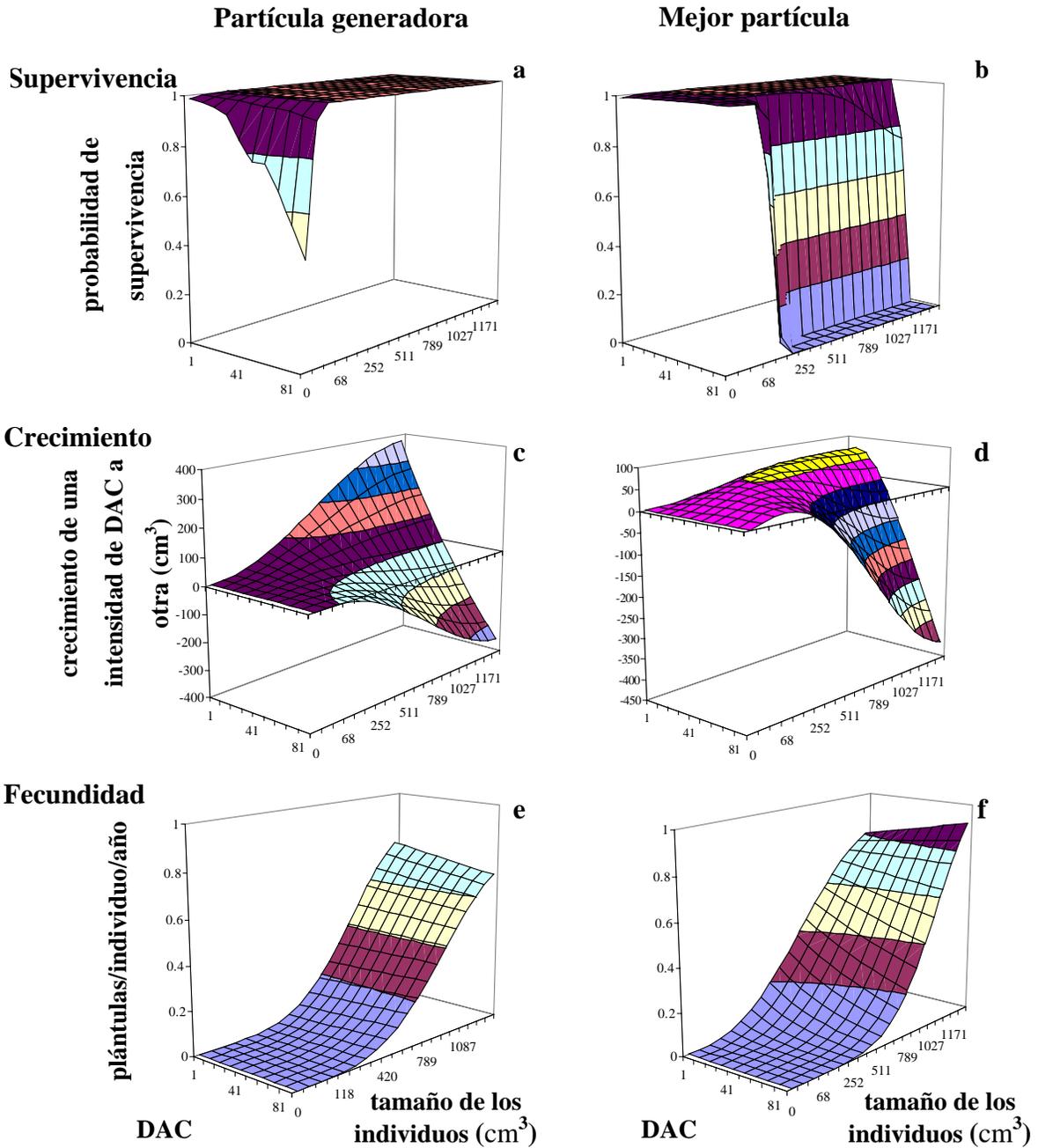


Figura 4.4. Supervivencia (a y b), crecimiento (c y d) y fecundidad (e y f) de los individuos dependiendo del grado de DAC y del tamaño de los mismos de acuerdo con la dinámica arrojada por la partícula generadora y la mejor partícula encontrada.

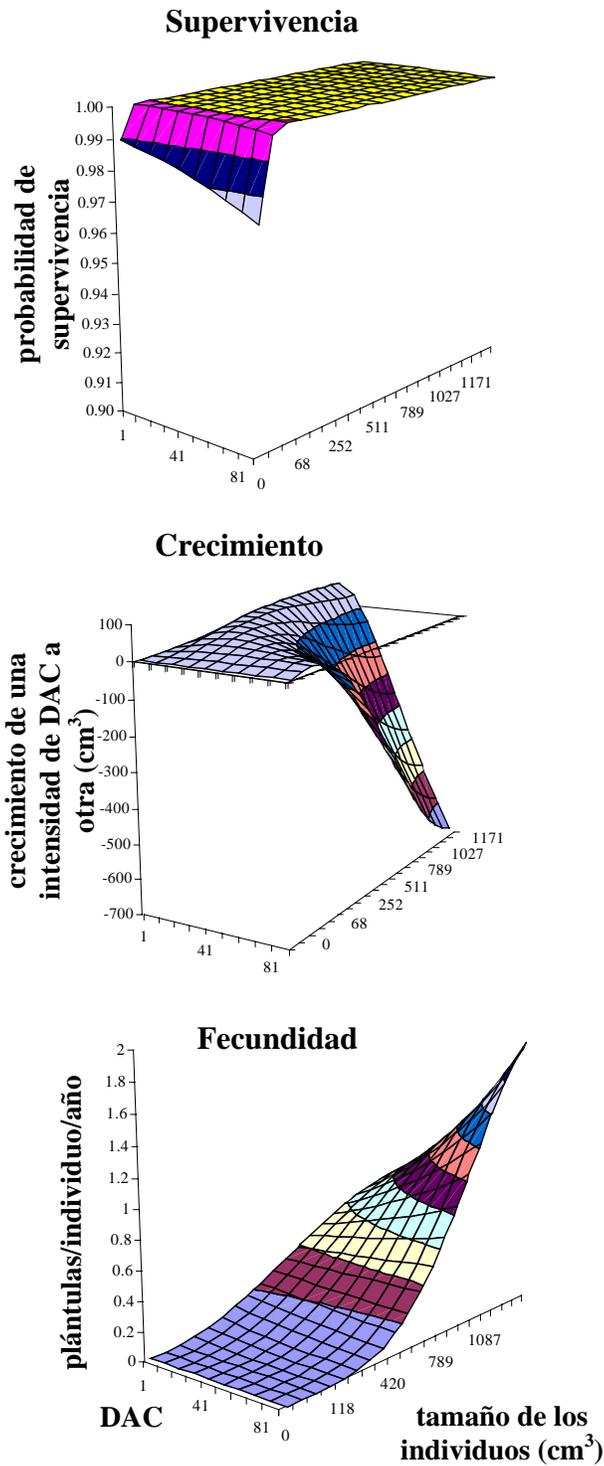


Figura 4.5. Supervivencia, crecimiento y fecundidad de los individuos dependiendo del grado de DAC y del tamaño de los mismos de acuerdo a la dinámica arrojada por la segunda mejor partícula encontrada.

Hay que recalcar que la segunda mejor partícula, que tuvo una verosimilitud 1.37 unidades por debajo de la generadora, no mantuvo la dinámica dada por la mejor partícula encontrada, y, a nivel de la dinámica poblacional asociada, fue más similar a la dada por la generadora (Fig. 4.5). A diferencia de la mejor partícula, la segunda mejor arrojó una supervivencia muy similar a la dada por la generadora aunque a una escala menor, mientras que el crecimiento se mantuvo igual al dado por la mejor partícula; la fecundidad por su parte, aunque similar a la mejor partícula, presentó un efecto del DAC mucho más pronunciado, y opuesto al determinado por la partícula generadora.

## 4.2 Datos de *Mammillaria dixanthocentron*

La aplicación del programa desarrollado a los datos de estructura de tamaños con los que se cuenta para las 10 poblaciones de *M. dixanthocentron* arrojó resultados consistentes entre corridas del programa y con los datos de dinámica demográfica con que se contaba para esta especie.

Partiendo de siete estructuras poblacionales con un tamaño muestral variable provenientes de 10 sitios de estudio con diferente grado de DAC, se realizaron 150 corridas del programa obteniéndose 150 partículas cuyas verosimilitudes presentaron un comportamiento asintótico (Fig. 4.6) mucho más patente que el dado por los datos artificiales (Fig. 4.1).

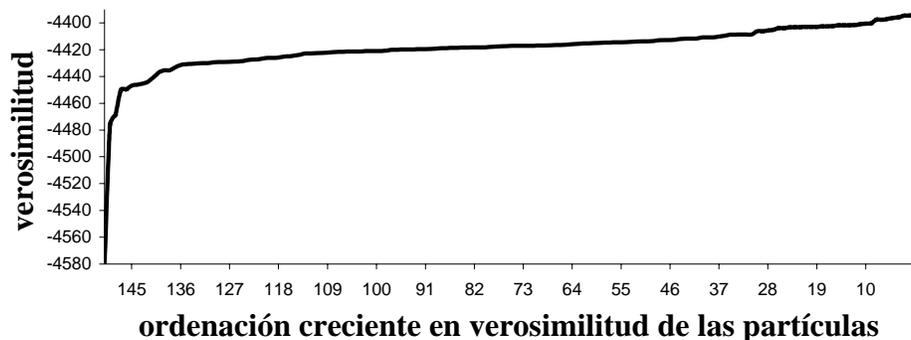


Figura 4.6. Verosimilitud de las mejores partículas obtenidas en cada corrida del programa para los datos reales.

Para verificar si corridas posteriores del programa arrojarían mejores partículas, realicé lo siguiente: tomé las mejores 25 partículas de las 150 obtenidas previamente en las correspondientes 150 corridas del programa con los datos reales y las incorporé al programa como las 25 partículas iniciales; así, corrí el programa y obtuve una mejor partícula que fue incorporada al conjunto de partículas iniciales; realicé lo anterior 25 veces y obtuve una nueva mejor partícula.

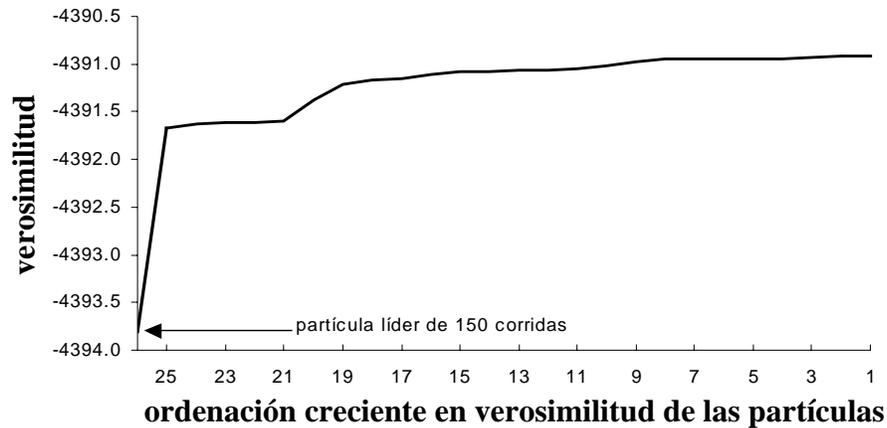


Figura 4.7. Verosimilitud de las mejores partículas obtenidas en cada corrida del programa partiendo de las 25 mejores partículas obtenidas previamente. Se indica la verosimilitud de la mejor partícula obtenida después de 150 corridas del programa.

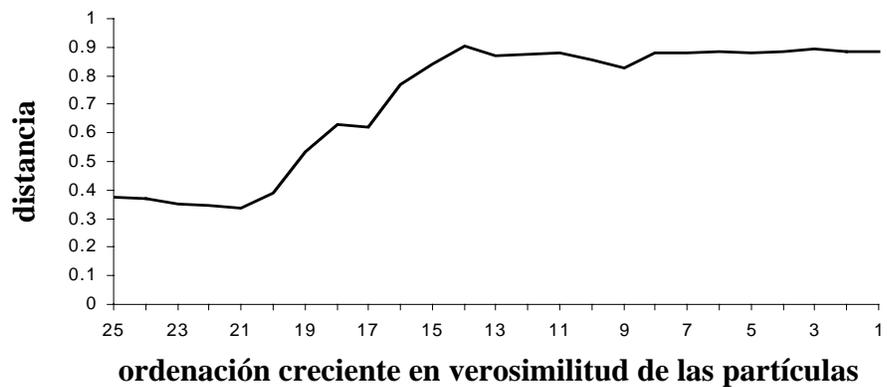


Figura 4.8. Distancia Manhattan de las mejores partículas obtenidas en cada corrida del programa partiendo de las 25 mejores partículas obtenidas previamente con respecto a la mejor partícula original (la resultante de 150 corridas del programa).

Las nuevas 25 partículas presentaron verosimilitudes bastante cercanas a la de la mejor partícula resultante de las 150 corridas del programa (Fig. 4.7). Además, estas partículas se ubicaron en un área cercana a la de la mejor partícula de la que partimos (Fig. 4.8). La mejor de estas 25 partículas arrojó una verosimilitud de -4390.91, siendo la diferencia entre las últimas dos partículas obtenidas por este proceso de 0.000338.

En cuanto a las estructuras poblacionales obtenidas por la mejor partícula para cada una de las intensidades de DAC observadas en el campo, éstas fueron más similares a las observadas a mayor tamaño muestral (Fig. 4.9).

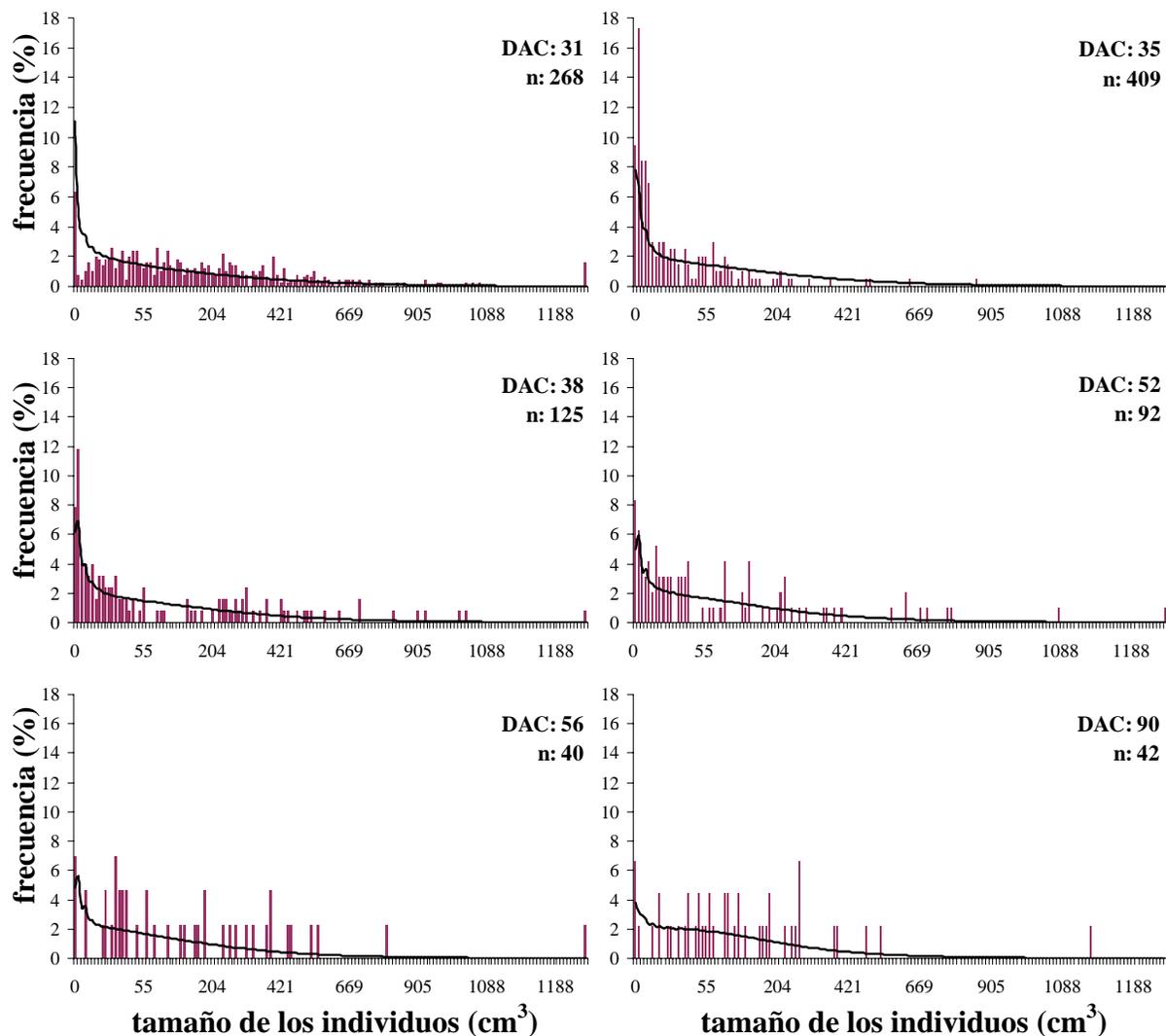


Figura 4.9. Estructura de tamaños observada (histograma) y obtenida por el modelo desarrollado (línea) para las seis intensidades de DAC observadas en el campo.

La mejor partícula arrojó una dinámica poblacional en donde el DAC afectó a los tres procesos demográficos (Fig. 4.10). En primer lugar, la supervivencia se vio afectada dependiendo del tamaño del individuo (Fig. 4.10 a). Para las plántulas, el DAC redujo su probabilidad de supervivencia. Por el contrario, para individuos más grandes, pero menores a  $250 \text{ cm}^3$ , el efecto del DAC fue inverso. Los individuos medianos y grandes no vieron afectada su supervivencia por el DAC.

De igual manera, el crecimiento de los individuos se vio afectado por el DAC dependiendo del tamaño de los individuos (Fig. 4.10 b). Los individuos pequeños se vieron favorecidos por el DAC, creciendo más a mayor intensidad de DAC. Lo contrario pasó para los individuos medianos y grandes, los cuales presentaron tasas de crecimiento menores a medida que el DAC aumenta, presentándose retrogresión en los individuos más grandes.

Finalmente, el DAC afectó negativamente a la fecundidad de todos los individuos, sin importar su tamaño (Fig. 4.10 c).

Estas tendencias en los cambios en los procesos demográficos son consistentes en las tres mejores partículas obtenidas, lo cual da mayor certidumbre de que el óptimo global presenta las mismas tendencias.

Comparando ahora los resultados anteriores para los datos reales con la información proporcionada por las dos dinámicas demográficas (La Cucharilla y La Pedrera, periodo 2001-2002), tenemos una correspondencia bastante buena (Fig. 4.11).

En cuanto a la supervivencia, la dinámica dada por los dos sitios de estudio es similar, en lo general, a la dada por la mejor partícula encontrada (Fig. 4.10 a). Sin embargo, tenemos que La Cucharilla ( $\text{DAC} = 34.74$ ) con respecto a La Pedrera ( $\text{DAC} = 51.58$ ) presenta una menor supervivencia de las plántulas (1.98%) y una mayor supervivencia en los individuos un poco mayores (Fig. 4.11 a). Este patrón es inverso del obtenido por la mejor partícula. No obstante, es importante recalcar que la prueba Kolmogorov-Smirnov no logró distinguir diferencia entre las supervivencias de ambos sitios ( $p > 0.1$ ).

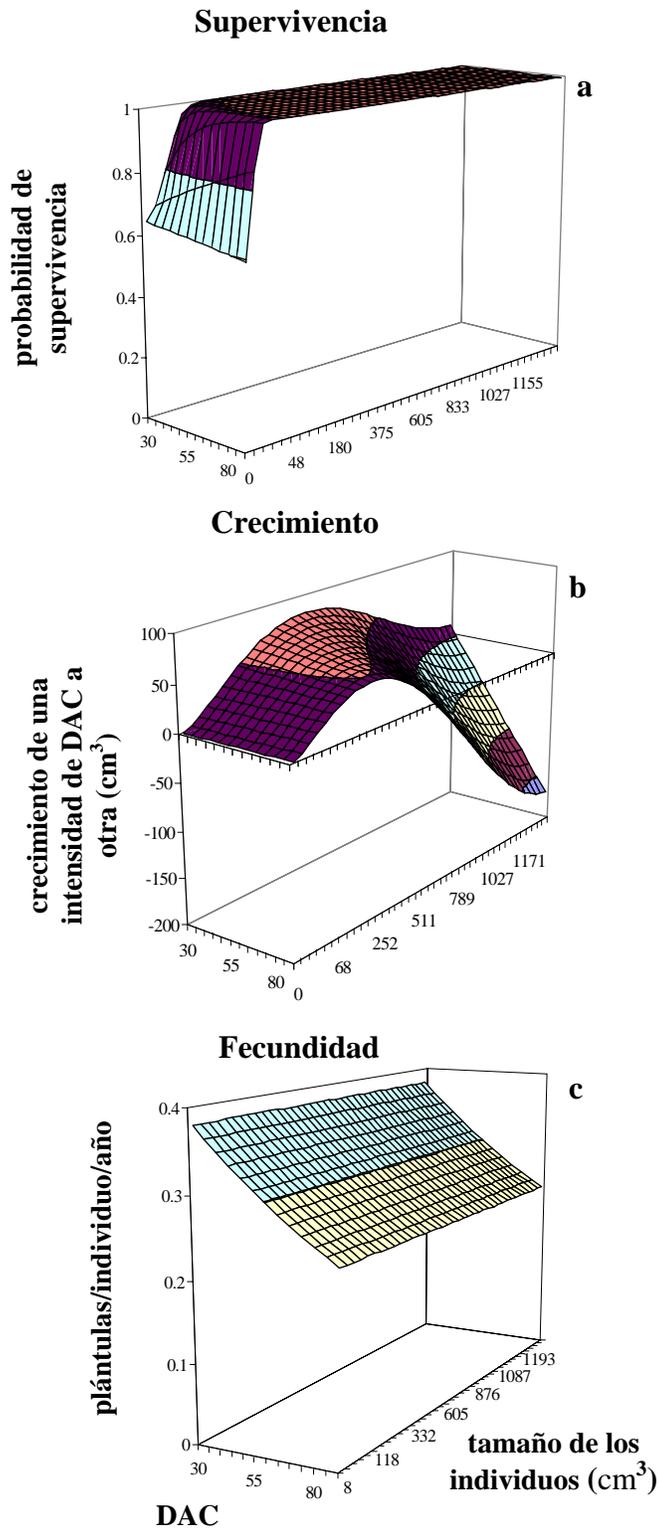


Figura 4.10. Supervivencia (a), crecimiento (b) y fecundidad (c) de los individuos de *Mammillaria dioxanthocentron* dependiendo del grado de DAC y del tamaño de los mismos de acuerdo con la dinámica arrojada por la mejor partícula encontrada.

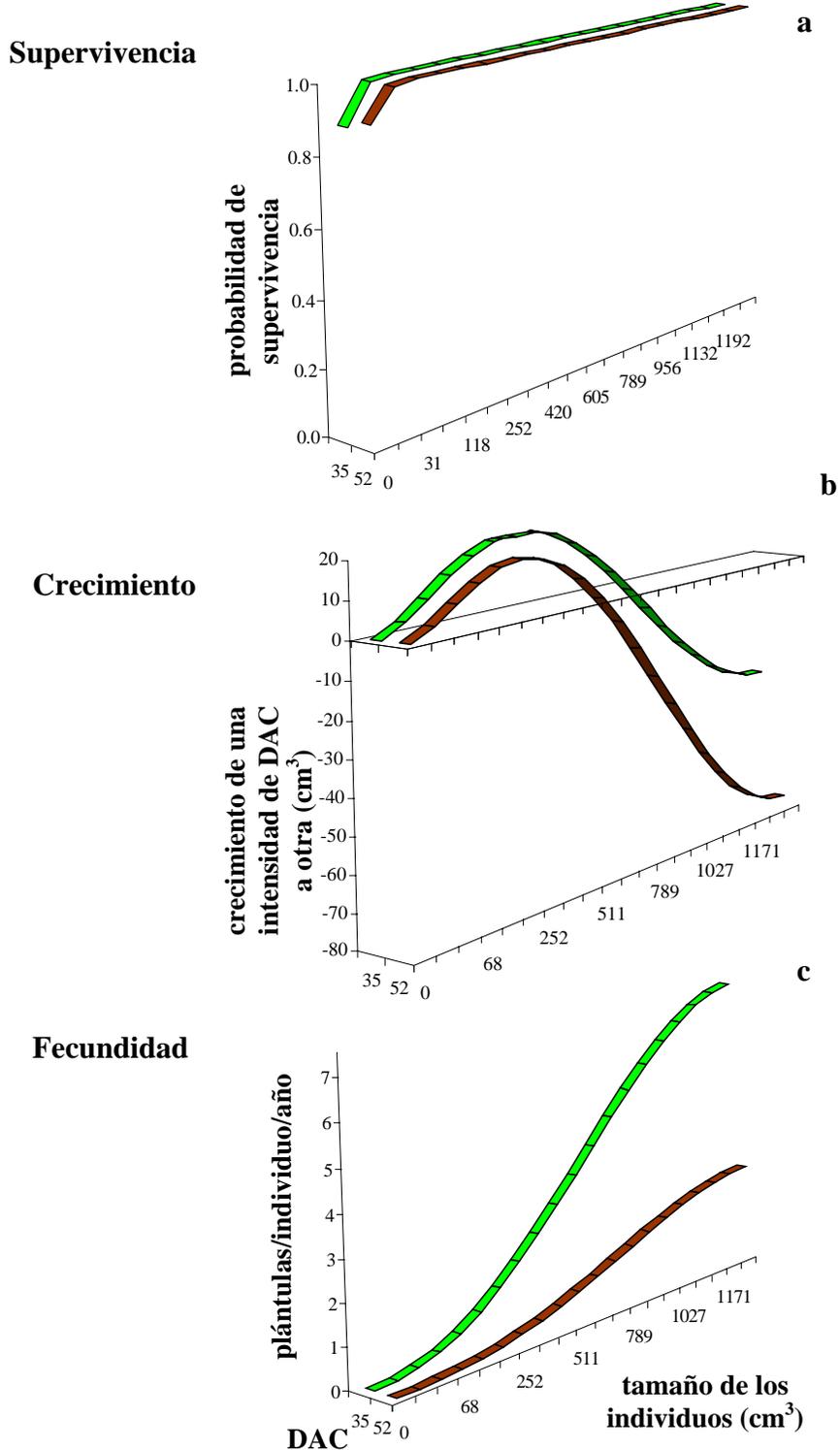


Figura 4.11. Supervivencia (a), crecimiento (b) y fecundidad (c) de los individuos de *Mammillaria dixanthocentron* obtenida a partir de los datos de dinámica demográfica para los sitios La Cucharilla y La Pedrera (línea clara y oscura, respectivamente) durante el periodo 2001-2002.

El crecimiento fue el proceso que mejor ajuste presentó. Tanto en la mejor partícula (Fig. 4.10 b) como en la dinámica demográfica (Fig 4.11 b), las plantas que crecen son las de tallas pequeñas y medianas, mientras que las de mayor talla decrecen en su tamaño; el DAC afecta negativamente este proceso, reduciendo el crecimiento de los individuos. Estos crecimientos difirieron únicamente en la escala, que en la mejor partícula varió en un intervalo de  $300 \text{ cm}^3$  de amplitud y en la dinámica demográfica varió en  $100 \text{ cm}^3$ .

La dinámica de la fecundidad dada por la mejor partícula, como se recordará, presenta una disminución en el número de plántulas generadas por año y no presenta diferente grado de fecundidad dependiendo del tamaño de los individuos (Fig. 4.10 c). Los dos sitios de estudio son consistentes con el hecho de que al haber mayor DAC disminuye el número de plántulas producidas; además en este caso, la fecundidad sí es dependiente del tamaño de los individuos, lo cual es más congruente con lo esperado para la especie bajo estudio (Fig. 4.11 c).



# Discusión

## 5.1 Datos artificiales

El proceso de validación del modelo arrojó resultados satisfactorios al poder reconstruir las estructuras poblacionales “observadas” arrojadas por la partícula generadora; sin embargo, a nivel de la dinámica asociada a la mejor partícula encontrada, los resultados son sólo parcialmente satisfactorios.

Como se observa en la figura 4.1, existe un aparente crecimiento asintótico en la verosimilitud, que, después de 100 corridas del programa, no permite afirmar que la asíntota se ha alcanzado. Dicha asíntota se esperaría alcanzar a medida que aumentara el número de corridas y cada vez más partículas líder llegasen al óptimo global.

Otro punto a notar es que, al no tender las partículas a estar cada vez más cerca de la generadora a medida que se incrementaba la verosimilitud (Fig. 4.2) queda claro que el espacio paramétrico presenta muchos máximos locales, siendo de hecho la partícula generadora sólo uno de ellos al haber quedado su verosimilitud por debajo de la de la mejor partícula (Fig. 4.1). Como

puntos distantes en el espacio paramétrico implican dinámicas demográficas asociadas diferentes, tenemos que distintas dinámicas producen estructuras poblacionales semejantes.

El hecho de que la mejor partícula encontrada presentara una verosimilitud 19.22 unidades por encima de la de la generadora (Fig. 4.1) se explica por el componente aleatorio que contienen las estructuras “observadas”: el método Monte Carlo parte de las estructuras generadas por la partícula generadora y obtiene de manera aleatoria la muestra “observada” que determina las estructuras “observadas”. Dicha aleatoriedad agregó un componente de error a las estructuras que fue mejor ajustado por la mejor partícula encontrada (Fig. 4.3).

Con respecto al ajuste que hizo la mejor partícula encontrada de la dinámica demográfica dada por la partícula generadora, podemos decir que éste fue bueno, especialmente para los individuos de menos de  $200 \text{ cm}^3$ : en general su supervivencia es alta, y crecen y se reproducen menos que los individuos más grandes, y el DAC afecta negativamente a ambos procesos.

Sin embargo, también se observaron diferencias muy importantes, atribuibles, en gran medida, a la muestra “observada” anómala de la que partimos para generar las estructuras poblacionales “observadas”: en ella, se tiene una ausencia de individuos grandes. Esta muestra no sería la esperada en un estudio real, pues los individuos grandes suelen ser los más conspicuos y, por lo tanto, los que, de existir, están siempre presentes en una muestra de la población.

Al no existir información sobre los individuos grandes, éstos no influyen en la ecuación (2.2). Con ello, la supervivencia, crecimiento y fecundidad para estos tamaños pueden ser modelados de cualquier forma, no teniendo efecto alguno sobre el valor de la verosimilitud de la partícula. Adicionalmente, esta muestra impide que exista consistencia entre las dinámicas asociadas a las mejores partículas encontradas, como puede apreciarse al comparar la mejor partícula (Fig. 4.4) y la segunda mejor partícula (Fig. 4.5). Ambas partículas estiman una dinámica semejante para los individuos pequeños (con excepción de la supervivencia) y difieren en la de los grandes cuando el DAC es elevado.

El caso más dramático de diferencias entre la partícula generadora y la mejor partícula ocurrió en la reconstrucción de la supervivencia, donde la dinámica dada por la mejor partícula fue, a intensidades altas de DAC, muy diferente a la dada por la generadora. Que esta dinámica de la supervivencia corresponda a la mejor partícula se debe a que las estructuras a ajustar, al no presentar individuos de tallas grandes a partir de la intensidad de DAC 56, permite que la supervivencia de los individuos que existían previamente a esta intensidad caiga a cero. Así, la

dinámica anómala obtenida para la supervivencia se debe en mayor medida a la singularidad de la muestra “observada” que a la identidad de la partícula generadora.

Estos resultados ilustran que, bajo ciertas condiciones, como puede ser la falta de representación de algunas categorías de tamaños en la muestra, se pueden obtener estimaciones inadecuadas de la dinámica poblacional, y que partículas con dinámicas demográficas diferentes pueden arrojar estructuras poblacionales similares. Esto destaca la importancia de evaluar cuáles son los requisitos que deben reunir los datos para que los resultados de un modelo de este tipo sean consistentes y más precisos.

## 5.2 Datos de *Mammillaria dixanthocentron*

Los resultados que obtuve para *Mammillaria dixanthocentron* mediante el MIPDAC nos dicen que el DAC afecta la dinámica poblacional de esta especie preponderantemente de forma negativa, lo que parece confirmarse por los datos de campo relacionados con la dinámica demográfica para sus poblaciones.

De nuevo en este caso se observa un crecimiento asintótico en la verosimilitud (Fig. 4.6), que, después de 150 corridas del programa, no permite afirmar que la asíntota se ha alcanzado. Sin embargo, hay que considerar que la diferencia en verosimilitud entre el primer y segundo lugar es menor a una unidad, y dado que el intervalo de variación en verosimilitud entre las 150 partículas líderes fue de 182 unidades, podemos considerar que el óptimo global no diferirá en gran medida del óptimo alcanzado.

El proceso que se realizó posteriormente, en el cual se tomaron las mejores 25 partículas de las 150 obtenidas previamente y se usaron como partículas iniciales en el programa, permitió alcanzar una solución con una mayor verosimilitud. Sin embargo, estamos hablando de tan sólo 2.8951 unidades de diferencia entre el óptimo alcanzado después de 150 corridas del programa y el óptimo alcanzado por este proceso, lo cual no representa una mejoría considerable. Asimismo, la dinámica de esta nueva partícula tampoco difirió sustancialmente de las anteriores, sugiriendo nuevamente que la solución proporcionada por el modelo es sólida.

Por ello, aunque este método sólo parte de las 25 mejores partículas obtenidas previamente y no de nuevas partículas aleatoriamente seleccionadas en el espacio paramétrico, sí apoya el hecho de que una cantidad mayor de corridas del programa incrementaría la verosimilitud del óptimo.

En cuanto al ajuste dado por la mejor partícula de las estructuras observadas, se tuvo que éste fue bueno. Dicho ajuste fue mejor para aquellas poblaciones con las que se contaba con un mayor tamaño de muestra (Fig. 4.9 DAC: 31 y 38) que para las que presentaban menor tamaño muestral (Fig. 4.9 DAC: 56 y 90). Esto se debe a que, a mayor tamaño muestral, una población tendrá mayor peso sobre la verosimilitud y, por lo tanto, mayor será la importancia que tendrá el buen ajuste de su estructura poblacional.

En cuanto a la dinámica demográfica asociada a la mejor partícula encontrada, tenemos una buena correspondencia entre ésta y la dada por los datos de dinámica demográfica para los dos sitios con los que se contaba información (Fig. 4.10 y 4.11).

Con relación a la supervivencia, el hecho de que la prueba Kolmogorov-Smirnov no lograra distinguir diferencia entre las supervivencias en los sitios La Pedrera y La Cucharilla, nos indica que dicha diferencia puede ser adjudicable más al azar que a un efecto diferenciado del DAC. Podemos afirmar así que el patrón general de la supervivencia dado por la dinámica demográfica fue reconstruido adecuadamente por el MIPDAC.

El crecimiento mostró el mejor ajuste de la dinámica dada por la mejor partícula encontrada, sólo presentándose una diferencia en la escala en que se dio este crecimiento.

El único problema importante fue que el modelo no pudo reconstruir el hecho de que la fecundidad es dependiente del tamaño de los individuos (Fig. 4.10 c), lo que esperaríamos que sucediera pues, en general, a mayor tamaño de un cacto, mayor el número de flores (Godínez-Álvarez *et al.*, 2003), y por ende, de plántulas que puede producir. Dado que la diferencia en fecundidad debida al tamaño no tiene un valor mínimo en la naturaleza, el intervalo que determina dicha diferencia en mi modelo incluye al cero (*i. e.*, una diferencia nula); por ello, establecer un intervalo que lo excluyera hubiera sido un acto arbitrario, dado que todos los intervalos se establecieron siguiendo un mismo procedimiento.

Hay que resaltar sin embargo el hecho de que existe un supuesto subyacente a la modelación de la fecundidad: las plántulas presentes en un año son el resultado de las flores producidas durante éste. Este supuesto implica la ausencia de un banco de semillas a largo plazo

(mayor a un año de edad). La existencia de dicho banco no ha sido aún demostrada para *M. dixanthocentron* (Rojas-Aréchiga y Vázquez-Yanes, 2000), aunque en otras especies del mismo género dicho banco sí existe (Bowers, 2005). Dado lo anterior, consideré no incluir un banco de este tipo en el modelo; la demostración de lo contrario sería fácilmente incorporable al modelo agregando a las plántulas producidas por flores las procedentes del banco de semillas.

De acuerdo con los resultados arrojados por el programa en los tres procesos demográficos de interés, podemos decir que, en términos generales, el desempeño que tuvo el programa al reconstruir la dinámica demográfica de *M. dixanthocentron* afectada por el gradiente de DAC fue bueno, revelando los mismos efectos negativos del DAC sobre el crecimiento y la fecundidad mostrados por la dinámica demográfica conocida para esta especie (Fig. 4.11).

Como se pudo ver con los datos artificiales (Figs. 4.4 y 4.5), el principal problema que presenta el MIPDAC es que, dado que distintas dinámicas poblacionales pueden generar estructuras poblacionales muy parecidas, la mejor partícula encontrada puede no reflejar la dinámica original que generó las estructuras observadas.

Al tomar las estructuras poblacionales, desearíamos que la dinámica poblacional que diera como resultado estas estructuras fuera única. Este parece ser el caso del conjunto de datos de *M. dixanthocentron*, pero el ejemplo con los datos artificiales demuestra que esto no necesariamente es lo que sucede. Dado que la estructura poblacional es resultado de la interacción de tres procesos demográficos básicos (supervivencia, crecimiento y fecundidad), éstos se conjugan para arrojar, en cada momento en el tiempo, una estructura poblacional diferente. Dichos procesos pudieran compensarse entre sí para que valores diferentes de los parámetros den lugar a dinámicas diferentes con estructuras asociadas similares. Así por ejemplo, una disminución en la supervivencia de los individuos de tallas grandes se puede ver compensada con un mayor crecimiento de los individuos de tallas menores que los sustituyan. Como ya se señaló, la ausencia de individuos de ciertas tallas puede agravar aún más este problema.

Esta compensación de los procesos se puede presentar tanto dentro de cada año observado como entre años. Por ejemplo, para un año en particular, una cantidad fija de plántulas producidas en ese año puede deberse a una fecundidad diferencial de acuerdo con el tamaño de los individuos o a una fecundidad homogénea de éstos. Entre años, una disminución, a medida que pasa el tiempo (y aumenta el DAC), en el número total de plántulas producidas en un año puede deberse a una de dos causas: a una disminución en el número de plántulas producidas por

individuo en una población con una cantidad constante de individuos adultos o a una disminución en el número de individuos adultos produciendo una cantidad constante de plántulas por individuo.

De esta forma, el modelo permite que distintas combinaciones de parámetros, asociados a distintas dinámicas demográficas, arrojen estructuras similares. Posiblemente, de entre las partículas con buena verosimilitud, a una le corresponderá una dinámica más congruente con la especie bajo estudio; sin embargo, la verosimilitud de esta partícula no tiene por qué ser la más grande. Esto lo demuestra la modelación que se hizo con los datos artificiales, donde la segunda mejor partícula presentó una mejor modelación de la supervivencia. Esto sugiere la necesidad de contar con otros elementos adicionales que permitan identificar cuál es la partícula apropiada cuando el modelo produzca varios resultados distintos. Algunos de estos elementos se discuten en la sección 5.4.

### **5.3 Supuestos subyacentes al modelo**

Aunque no se establecieron explícitamente en el modelo, existen supuestos que lo subyacen. El primer supuesto importante es que los sitios de estudio representan una cronosecuencia de perturbación, *i. e.*, que las poblaciones ubicadas en cada sitio son representantes del devenir de una sola población ante un incremento en el DAC a lo largo del tiempo. Al sustituir espacio por tiempo estamos suponiendo que la variación espacial y la temporal son equivalentes y que el destino de una población sujeta a DAC es único (Pickett, 1989). Aunque esta sustitución proporciona sólo una evaluación indirecta del efecto del DAC, este enfoque provee la oportunidad de aproximarse a un fenómeno relativamente complejo que por métodos directos sería difícil de estudiar.

Para que una cronosecuencia se considere válida es importante que el conjunto de condiciones e interacciones que determinan el comportamiento de cada población sean semejantes entre sí o sean controlados adecuadamente (Pickett, 1989), ya sea considerando sitios que difieran entre sí lo menos posible, de modo que éstos no afecten a las poblaciones al grado de

convertirse en un factor que introduzca un ruido considerable, o controlando estadísticamente dichos factores, lo cual no siempre es fácil de hacer.

Las condiciones que determinan el comportamiento poblacional incluyen factores abióticos como tipo de suelo, pendiente, temperatura, precipitación, etc.; así como factores bióticos como interacciones de depredación y competencia. Hay que tener también en cuenta el efecto sinérgico que pueden tener estos factores entre sí y con el factor bajo estudio.

En el caso de mi estudio, un elemento importante que puede ver afectado el supuesto de cronosecuencia fue la altitud a la que se ubicaron los sitios de estudio, ubicándose a altitudes que abarcaron un intervalo desde los 640 a los 2500 m s.n.m. La altitud juega un papel preponderante en la determinación de muchos factores abióticos (temperatura, precipitación, concentración de gases en la atmósfera, vientos, tipos de suelo, etc.), los cuales podrían confundir su efecto con el del DAC, al existir una posible sinergia entre estos factores y el DAC (Bühler y Schmid, 2001).

En cuanto a la identidad de la función  $\theta(t)$ , se hizo el supuesto de que ésta corresponde a la función identidad, *i. e.*, que el DAC aumenta un punto en la medición de su intensidad con cada año que pasa. Hay que recordar que el DAC es un fenómeno multifactorial que, para este estudio, tomó como agentes determinantes al ganado, las actividades humanas y la degradación del suelo, los cuales pueden darse con mayor o menor intensidad dependiendo del grado de DAC que presente el entorno en un momento dado. De acuerdo con Singh (1998), la degradación ambiental causada por el DAC suele ser un fenómeno discontinuo y no lineal, por lo que no se esperaría ver una afectación en la dinámica poblacional de *M. dixanthocentron* sino hasta que se alcanzaran valores altos de DAC. Sin embargo, debido en parte a la identidad que se le dio a la función  $\theta(t)$  el modelo no permite cambios no lineales en los procesos demográficos.

Otro supuesto que impide este tipo de cambios en los procesos demográficos es la función de cambio de los 11 parámetros a lo largo del gradiente de DAC: la recta (3.6). Este supuesto de linealidad se hizo para poder mostrar un efecto del DAC en los parámetros, pero manteniendo, en la medida de lo posible, un modelo relativamente simple.

Finalmente, existen otros supuestos implícitos en el modelo que vale la pena mencionar:

- Las funciones que determinan el núcleo ( $s(x, \theta(t))$ ,  $g(x, y, \theta(t))$ ,  $f_1(x, \theta(t))$  y  $f_2(x, y, \theta(t))$ ) son las mismas a lo largo del gradiente de DAC, sólo se modifican los 11 parámetros que las determinan.

- El único factor que afecta a la dinámica demográfica es el DAC; no se incorpora al modelo estocasticidad ambiental o demográfica.
- No existe migración en la población, de forma que los nuevos individuos que se ven representados en la estructura poblacional corresponden a plántulas producto de la reproducción de los adultos previamente presentes en la población; asimismo, los individuos salen de la población y, por lo tanto, ya no se ven representados en la estructura poblacional, cuando mueren.
- Existe densoindependencia en la modelación de la dinámica, *i. e.*, las tasas demográficas no se ven afectadas por el tamaño poblacional.

## 5.4 Consideraciones adicionales y mejoras posibles al modelo

Un elemento importante para elucidar la idoneidad de una partícula óptima, sin afectar las conclusiones a las que se lleguen en cuanto al efecto del DAC sobre la dinámica poblacional, podría ser el verificar que dicha partícula tenga un comportamiento biológicamente lógico para individuos de diferente tamaño. Aquí entraría en juego el conocimiento que sobre la dinámica demográfica de la especie bajo estudio tenga el investigador. Dicho conocimiento permitiría discernir, de entre diferentes dinámicas, aquella que tuviera propiedades concebibles para la especie particular bajo estudio.

Lo anterior, sin embargo, no garantiza un óptimo único e involucra un componente de subjetividad que no es deseable.

Alternativamente, una posible solución al problema sería incorporar información no sólo de las estructuras poblacionales, sino sobre la dinámica poblacional a corto plazo que permitiera reducir el espacio paramétrico a explorar y de este modo eliminar aquellas partículas con dinámicas no factibles para la especie. Dicha información podría consistir en el seguimiento de una o más de las poblaciones bajo estudio por un periodo de un año. Esta adición en la información tendría el gran inconveniente de requerir mucho más tiempo para llegar a

conclusiones sobre el efecto del DAC en la dinámica poblacional, pero proporcionaría más confianza sobre éstas.

Un punto a considerar en la ubicación del óptimo global es la amplitud de los intervalos que integran el espacio paramétrico. Para el caso de los datos reales, en cinco de los 11 parámetros que corresponden a los interceptos en la recta (3.6), el óptimo se encontró en alguno de los extremos del intervalo propuesto para ese parámetro. Dado que los intervalos fueron establecidos a partir de los datos de dinámica demográfica para *M. dixanthocentron* y se hicieron bastante amplios (cuatro desviaciones estándar alrededor de la media), no se esperaba que el óptimo se encontrara fuera de estos intervalos. El incrementar el ancho de los intervalos permitiría posiblemente encontrar una partícula con mayor verosimilitud pero que en su dinámica no fuera biológicamente factible para la especie. Además, se incorporarían al espacio paramétrico dinámicas que, biológicamente, podrían considerarse menos creíbles; de ahí que no haya considerado ampliar dichos intervalos para la especie estudiada.

Hay que resaltar el hecho de que los intervalos correspondientes a cada parámetro deberán ser establecidos por el usuario y que de la amplitud o estrechez de los mismos dependerá en gran medida la identidad y consistencia biológica del óptimo encontrado. Es en el establecimiento de los intervalos en donde entra en juego el conocimiento que se tenga sobre la especie a estudiar, pues dicho conocimiento permitirá establecer un conjunto de valores de los parámetros que sea lógico a nivel biológico. Así, una mayor cantidad de información, permitirá establecer intervalos más estrechos, reduciendo con ello el espacio paramétrico a explorar y, por ende, eliminando dinámicas no factibles para la especie, facilitando asimismo la identificación del óptimo global. Nótese también que el usuario aporta otros elementos específicos a la especie bajo estudio como son el intervalo de tamaños posibles y la edad/tamaño a la primera reproducción.

Otro punto a considerar es el tamaño muestral de las poblaciones estudiadas, el cual varió entre sitios de estudio. Así, la muestra por población, a diferencia de las “observadas” en los datos artificiales (que fue de 150 individuos por población), varió de 281 (Cerro Gordo) a 40 individuos (Los Cues). Esto implica un diferente grado de certidumbre sobre el efecto del DAC sobre las poblaciones observadas en el sentido de que, en una población pequeña, el azar pudiera jugar un papel más importante sobre la estructura poblacional de lo que podría afectar el DAC. Así, a mayor tamaño muestral, mayor confianza tenemos sobre el efecto que ha tenido el DAC sobre dicha estructura. De esta manera, el programa tiene que ajustar estructuras que, al diferir en

tamaño muestral, difieren en la confiabilidad de la información que nos proporcionan con respecto al efecto del DAC sobre ellas.

Una forma de mejorar las estimaciones, considerando este hecho, sería hacer que el MIPDAC tomara en cuenta el tamaño poblacional como un parámetro más del programa, estimando, para cada intensidad de DAC observada, el tamaño poblacional que mejor se ajuste a los tamaños observados. Una amplia gama de modelos incorporan el tamaño poblacional como un elemento preponderante en la dinámica demográfica de una especie (Finkelstein, 1985; Caswell, 2001), pues biológicamente se sabe que éste tiene un papel muy importante en el tipo de interacciones que se darán entre los individuos y, por ende, en la dinámica poblacional que de ellas resulte. El tamaño poblacional se incorporaría al modelo, no estandarizando la estructura poblacional resultado de la iteración dada por la ecuación (2.2). De esta manera, la iteración no daría como resultado una función de densidad, sino una estructura de tamaños real. Viendo la correlación existente entre los tamaños poblacionales obtenidos por el modelo (*i. e.*, los valores de  $\int_x n(x, \theta(t_i)) dx$  para cada intensidad de DAC  $\theta(t_i)$  observada) y los tamaños observados se determinaría la adecuación de una partícula en la estimación del tamaño poblacional. Dicha correlación, junto con la verosimilitud de la partícula, determinarían aquella considerada como óptima. Estos dos elementos podrían resumirse en una sola cantidad haciendo una suma ponderada de los mismos. Con lo anterior se obtendría un óptimo cuya dinámica asociada consideraría el hecho de que el tamaño poblacional puede verse afectado por el DAC, pudiéndose así considerar este óptimo como más confiable en cuanto a los resultados que arrojará.

Otro punto relevante es el tipo de función que cada uno de los 11 parámetros tuvo en el eje del DAC: la recta (3.6). Dado que el número de dimensiones del espacio paramétrico crece aritméticamente respecto del número de parámetros que esta función contenga, consideré la función más sencilla que permitiera mostrar un efecto del DAC sobre los parámetros: una recta. Sin embargo, esta función no tiene por qué ser la más idónea para modelar el efecto del DAC; una especie que se viera beneficiada por una intensidad media de DAC no sería modelada adecuadamente por una recta sino por una parábola (el cual, *e. g.*, es el caso de *Mammillaria pectinifera* [Martorell y Peters, 2005]). No obstante, hay que tener en cuenta que el considerar una parábola haría pasar la dimensión del espacio paramétrico de 22 a 33, lo cual redundaría en un mayor tiempo de computadora requerido para que el programa explorara dicho espacio. Como se ve, el considerar funciones más complejas presenta costos y beneficios que, dependiendo de

las capacidades computacionales y del conocimiento de la especie que tenga el usuario, serán ponderados por éste.

Como se dijo en la sección anterior, la identidad de la función  $\theta(t)$  también juega un papel importante en el modelo, la cual se consideró como la función identidad. Ésta, sin embargo, no tiene por qué ser la función más idónea para analizar el cambio en los parámetros que determinan la dinámica demográfica. Como se puede observar en el modelo desarrollado, la identidad de la función  $\theta(t)$  entra en juego en el momento de definir las funciones  $\rho(\theta(t))$ . De esta forma, un cambio en la identidad de la función  $\theta(t)$  afectará la forma como los parámetros cambien en el gradiente de DAC y viceversa. Al igual que en el caso de  $\rho(\theta(t))$ , un aumento en la complejidad del modelo afectará el tiempo que le tomará a éste encontrar el óptimo global.

Otra función importante es la asociada a la estructura con la que inicia la iteración descrita por la ecuación (2.2). Ésta se genera a partir de la estructura de la población sujeta al menor valor de DAC medido en el campo, ajustando un *spline* cúbico natural a la estructura poblacional y estandarizándolo para que integre uno en el intervalo de tamaños posibles (requisito fundamental en una función de distribución). Dado que el *spline* requiere como entrada una estructura discreta la cual ajustar, la forma que dicha estructura tiene es muy importante. Dicha forma, en última instancia, depende de la partición del intervalo de tamaños, pues a menor tamaño de la partición, menos nitidez tendrá la estructura; empero, a mayor tamaño, mayor el número de ceros que habrá en la estructura resultante (Fig. 5.1). Esto sugiere que una partición de tamaño medio es la más adecuada; así, para este modelo, el intervalo de tamaños se dividió en 50 subintervalos. Sin embargo, este valor es arbitrario, y un tamaño diferente de partición podría ser más adecuado para ajustar la primera estructura.

La forma que el *spline* tenga no es un aspecto menor en el modelo, pues el proceso iterativo de la ecuación (2.2), al partir de la primera estructura e ir multiplicándola por el núcleo bajo cada intensidad de DAC, haría que una pequeña diferencia entre dos primeras estructuras similares entre sí pudiera verse amplificada en cada iteración dando como resultado dinámicas demográficas muy diferentes.

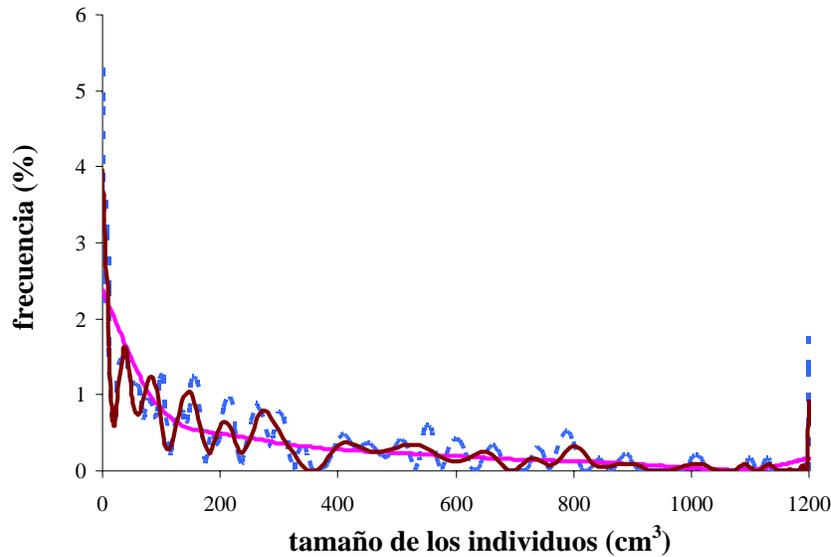


Figura 5.1. *Splines* cúbicos naturales correspondientes a la primera estructura poblacional observada (Cerro Gordo) de *Mammillaria dixanthocentron* para diferentes tamaños de partición: 10 (línea clara), 50 (línea oscura) y 100 (línea punteada) subintervalos.

Por lo anterior, sería deseable incorporar al modelo un parámetro que correspondiera al tamaño de la partición, que, al variar, permitiera explorar diferentes *splines* y ver el efecto que sobre la verosimilitud tendrían.

Una partición adicional que se encuentra presente en el modelo es la cuadratura usada en el método de integración de Gauss-Legendre. Como el tiempo requerido para el cálculo de la verosimilitud aumenta de forma exponencial con respecto al tamaño de la norma de la cuadratura, éste se estableció en 150 subintervalos con el objetivo de obtener un valor de la integral cercano al real pero que no consumiera demasiado tiempo en su cálculo. Además, más que el valor de verosimilitud arrojado por el programa, me interesó que la ordenación entre partículas por verosimilitud no se viera afectada por el tamaño de la cuadratura; lo cual se verificó que no ocurría.

## 5.5 Perspectivas

El modelo aquí presentado ha demostrado tener la capacidad de reconstruir con gran detalle cómo es que los procesos demográficos se ven afectados por el DAC (por ejemplo, en el caso de la función de crecimiento de *M. dixanthocentron*), pero también que puede proporcionar resultados ambiguos y poco satisfactorios bajo condiciones que aún no están claras.

Como se comentó en la introducción, el objetivo buscado por este modelo fue obtener un método “rápido y sucio” que partiera de información mínima, fácilmente obtenible, y valorara la respuesta que una especie en particular tendría ante el DAC. Para el caso de *M. dixanthocentron*, el modelo mostró un efecto negativo del DAC sobre los procesos demográficos y este efecto dependió del tamaño de los individuos. Esto nos habla de que el modelo logró su objetivo para el caso de esta especie. Sin embargo, como lo demostró la aplicación del modelo a datos artificiales, es posible que este modelo, aplicado a otras especies, dé resultados inadecuados al encontrar una dinámica poblacional, que, aunque capaz de conectar las estructuras poblacionales observadas, no se corresponda con lo que realmente aconteció.

Así, en este momento se puede considerar que se cuenta con un modelo prometedor pero que requiere de varias mejoras que permitan dar mayor confiabilidad a los resultados. Dichas mejoras ya han sido señaladas en la sección previa.

Además de la implementación de estas mejoras, el trabajo posterior se deberá enfocar en la evaluación de la sensibilidad y la robustez del modelo. Por sensibilidad se entiende el grado de variación en el valor de sus parámetros que admite un modelo sin ver afectados sus resultados (Murthy *et al.*, 1990); y por robustez, el grado de dependencia que tienen las conclusiones del modelo de que se cumplan sus supuestos (Giordano *et al.*, 1997). El análisis de sensibilidad se realizaría tanto a los datos que se introducen al modelo (número de poblaciones, tamaño de la muestra y espaciamiento entre valores de DAC), así como en las funciones que usamos (tamaño de la norma de la partición en el *spline* cúbico natural de la primera estructura poblacional, tamaño de la cuadratura, funciones que conforman al núcleo y la función de cambio de los parámetros a lo largo del DAC). Por su parte, el análisis de robustez evaluaría el efecto que tiene la violación de los supuestos arriba mencionados sobre los resultados arrojados por el modelo. De

no ser éstos graves, este análisis permitiría considerar la aplicación el modelo a un número mayor de especies: aquellas que no cumplieran con dichos supuestos.

También se deberá considerar traducir el programa a un lenguaje de programación de nivel bajo, pues el usado (Matlab), consume muchos recursos de la computadora al correr sobre la interfase de Matlab, que a su vez corre sobre Microsoft® Windows. Una alternativa sería pasar el programa a lenguaje C. Con ello el programa se ejecutaría en un tiempo menor al actual, que es de alrededor de 6 h/corrída en una máquina Pentium 4 a 2.40 GHz y 1.49 GB de RAM. Adicionalmente, el pasar el programa a C permitiría correr el programa sin la necesidad de instalar Matlab que es un programa muy grande.

Podemos concluir por tanto que el MIPDAC es un modelo con potencial de aplicación a otras especies de cactáceas. Una vez aplicado a un conjunto mayor de cactáceas y evaluada su sensibilidad y robustez, podrá evaluarse la aplicabilidad del modelo a otras especies vegetales. También sería importante considerar incorporar algunas de las mejoras arriba mencionadas para así permitir al programa ser más flexible en funciones y constantes que en esta versión fueron fijos.

# Literatura Citada

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **AC-19**: 716-723.
- Arias M., S. 1993. Cactáceas: conservación y diversidad en México. *Revista de la Sociedad Mexicana de Historia Natural* **XLIV**: 109-115.
- Arias M., S., U. Guzmán y S. Gama L. 1997. Cactaceae. Flora del Valle de Tehuacán-Cuicatlán, México. Fascículo **14**: 32-33.
- Bazzaz, F. A. 1983. Characteristics of populations in relation to disturbance in natural and man-modified ecosystems. En H. A. Mooney y M. Godron, eds. *Disturbance and Ecosystems*. Springer-Verlag: Berlin, West Germany. Pp. 259-275.
- Benton, T. G. 2007. Managing farming's footprint on biodiversity. *Science* **315**: 341-342.
- Bowers, J. E. 2005. New evidence for persistent or transient seed banks in three Sonoran desert cacti. *The Southwestern Naturalist* **50**: 482-487.
- Bradshaw, G. A. y P. A. Marquet. 2003. Introduction. En G. A. Bradshaw, P. A. Marquet y K. L. Ronnenberg (eds.). *How landscapes change: human disturbance and ecosystem fragmentation in the Americas*. Ecological Studies 162. Springer-Verlag: New York. Pp. 1-4.
- Bühler, Ch. y B. Schmid. 2001. The influence of management regime and altitude on the population structure of *Succisa pratensis*: implications for vegetation monitoring. *Journal of Applied Ecology* **38**: 689-698.

- Caswell, H. 2001. Matrix population models. Construction, analysis, and interpretation. 2<sup>nd</sup> edition. Sinauer: Sunderland, Massachusetts. 722 p.
- Cohen, J. E. 1979. Ergodic theorems in demography. Bulletin of the American Mathematical Society **1**: 275-295.
- Contreras, C. y T. Valverde. 2002. Evaluation of the conservation status of a rare cactus (*Mammillaria crucigera*) through the analysis of its population dynamics. Journal of Arid Environments **51**: 89-102.
- Crouse, D. T., L. B. Crowder y H. Caswell. 1987. A stage-based population model for loggerhead sea turtles and implications for conservation. Ecology **68**: 1412-1423.
- Dobson, A. J. 2002. An introduction to generalized linear models. 2<sup>nd</sup> edition. Chapman and Hall/CRC: Boca Raton. 225 p.
- Easterling, M. R., S. P. Ellner y P. M. Dixon. 2000. Size-specific sensitivity: applying a new structured population model. Ecology **81**: 694-708.
- Eberhart, R. C. y Y. Shi. 2001. Particle swarm optimization: developments, applications and resources. Proceedings of the Congress on Evolutionary Computation 2001. IEEE Press: Piscataway, New Jersey. Pp. 81-86.
- Ellner, S. P. y M. Rees. 2007. Stochastic stable population growth in integral projection models: theory and application. Journal of Mathematical Biology **54**: 227-256.
- Enright, N. J., M. Franco y J. Silvertown. 1995. Comparing plant life histories using elasticity analysis: the importance of life span and the number of life-cycle stages. Oecologia **104**: 79-84.
- Finkelstein, L. 1985. Mathematical modelling of dynamic biological systems. J. Wiley: New York. 355 p.
- Fox, G. A. y J. Gurevitch. 2000. Population numbers count: tools for near-term demographic analysis. The American Naturalist **156**: 242-256.
- Gaboune, B., G. Laporte y F. Soumis. 1993. Expected distances between two uniformly distributed random points in rectangles and rectangular parallelepipeds. Journal of the Operational Research Society **44**: 513-519.
- García, E. 1988. Modificaciones al sistema de clasificación climática de Köppen (para adaptarlo a las condiciones de la República Mexicana). Larios: México, D. F. 218 p.

- Giordano, F. R., M. D. Weir y W. P. Fox. 1997. A first course in mathematical modelling. 2<sup>nd</sup> ed. Brooks/Cole: Pacific Grove, California. 526 p.
- Godínez-Álvarez, H., T. Valverde y P. Ortega-Baes. 2003. Demographic trends in the Cactaceae. *The Botanical Review* **69**: 173-203.
- Golubov, J., M. C. Mandujano, M. Franco, C. Montaña, L. E. Eguiarte y J. Lopez-Portillo. 1999. Demography of the invasive woody perennial *Prosopis glandulosa* (honey mesquite). *Journal of Ecology* **87**: 955-962.
- Hernández, H. M. y H. Godínez A. 1994. Contribución al conocimiento de las cactáceas mexicanas amenazadas. *Acta Botánica Mexicana* **26**: 33-52.
- Hu, X. y R. C. Eberhart. 2002. Solving constrained nonlinear optimization problems with particle swarm optimization. Proceedings of the 6<sup>th</sup> World Multiconference on Systemics, Cybernetics and Informatics. Orlando, Florida.
- Hu, X., Y. Shi y R. C. Eberhart. 2004. Recent advances in particle swarm. Proceedings of the 2004 Congress on Evolutionary Computation. Portland, Oregon. Pp. 90-97.
- Kennedy, J. y R. C. Eberhart. 1995. Particle swarm optimization. Proceedings of the IEEE International Conference on Neural Networks. Perth, Australia. IEEE Service Center, Piscataway, New Jersey **IV**: 1942-1948.
- Kesseli, R. V. 1992. Population biology and conservation of rare plants. En S. K. Jain y L. W. Botsford (eds.). *Applied Population Biology*. Kluwer Academic Publishers: Dordrecht, Netherlands. Pp. 69-90.
- Kessler, M. 2001. Maximum plant-community endemism at intermediate intensities of anthropogenic disturbance in Bolivian montane forests. *Conservation Biology* **15**: 634-641.
- Knott, G. D. 2000. Interpolating cubic splines. Birkhauser: Boston, Massachusetts. 244 p.
- Krommer, A. R. y Ch. W. Ueberhuber. 1998. Computational integration. Society for Industrial and Applied Mathematics: Philadelphia, Pennsylvania. 445 p.
- Krusienski, D. J. y W. K. Jenkins. 2005. Design and performance of adaptive systems based on structured stochastic optimization strategies. *Circuits and Systems Magazine, IEEE* **5**: 8-20.
- Kunin, W. E. y K. J. Gaston. 1997. The biology of rarity: causes and consequences of rare-common. Chapman and Hall: London. 280 p.

- Martorell, C. y E. M. Peters. 2005. The measurement of chronic disturbance and its effects on the threatened cactus *Mammillaria pectinifera*. *Biological Conservation* **124**: 199-207.
- MacNeish, R. S. 1967. The prehistory of the Tehuacan Valley. Vol. 1. Pp. 290-331.
- McCullagh, P. y J. A. Nelder. 1989. *Generalized Linear Models*. Chapman and Hall: London. 511 p.
- McIntyre, S. y S. Lavorel. 1994. Predicting richness of native, rare, and exotic plants in response to habitat and disturbance variables across a variegated landscape. *Conservation Biology* **8**: 521-531.
- Meyer, W. J. 1984. *Concepts of mathematical modeling*. McGraw-Hill: New York. 440 p.
- Murthy, D. N. P., N. W. Page y E. Y. Rodin. 1990. *Mathematical modelling: a tool for problem solving in engineering, physical, biological and social science*. Pergamon: Oxford. 339 p.
- Oldfield, S. 1991. Review of significant trade in species of plants included in Appendix II of CITES 1983-1989. Report prepared for the 8<sup>th</sup> meeting of the Conference of the Parties. Royal Botanic Gardens Kew. 8-9 September 1991.
- Peram, T., K. Veeramachaneni y Ch. K. Mohan. 2003. Fitness-distance-ratio based particle swarm optimization. *Swarm Intelligence Symposium 2003. Proceedings of the 2003 IEEE*.
- Peters, E. M. y C. Martorell. 2000. Reporte final del proyecto: Conocimiento y conservación de las mamilarias endémicas del Valle de Tehuacán-Cuicatlán. Proyecto R166, CONABIO, 1998-2000. Instituto de Ecología, UNAM. 38 p.
- Peters, E. M., C. Martorell y E. Ezcurra. 2007. Nurse rocks are more important than nurse plants in determining the distribution and establishment of globose cacti (*Mammillaria*) in the Tehuacán Valley, Mexico. *Journal of Arid Environments*, doi:10.1016/j.jaridenv.2007.10.004.
- Pickett, S. T. A. 1989. Space-for-time substitution as an alternative to long-term studies. En G. E. Likens (ed.). *Long-term studies in ecology*. Wiley: Chichester. Pp. 71-88.
- Pickett, S. T. A., J. Kolasa, J. J. Armesto y S. L. Collins. 1989. The ecological concept of disturbance and its expresion at various hierarchical levels. *Oikos* **54**: 129-136.
- Rapport, D. J. y W. G. Whitford. 1999. How ecosystems respond to stress: common properties of arid and aquatic systems. *BioScience* **49**: 193-203.

- Rojas-Aréchiga, M. y C. Vázquez-Yanes. 2000. Cactus seed germination: a review. *Journal of Arid Environments* **44**: 85-104.
- Rykiel Jr., E. J. 1985. Towards a definition of ecological disturbance. *Australian Journal of Ecology* **10**: 361-365.
- Santini, B. 2007. Efecto de la heterogeneidad temporal y el disturbio sobre la dinámica poblacional de dos especies raras del género *Mammillaria* (Cactaceae). Tesis de licenciatura. Facultad de Ciencias. Universidad Nacional Autónoma de México.
- SEMARNAT, 2002. Norma Oficial Mexicana NOM-059-ECOL-2001. Protección ambiental, especies nativas de México de flora y fauna silvestres, categorías de riesgo y especificaciones para su inclusión, exclusión o cambio, lista de especies en riesgo. En: Diario Oficial de la Federación. 6 de marzo de 2002. Segunda Sección.
- Shi, Y. H. y R. C. Eberhart. 1998. A modified particle swarm optimizer. IEEE International Conference on Evolutionary Computation. Anchorage, Alaska.
- Silva, J. F., J. Raventos, H. Caswell y M. C. Trevisan. 1991. Population responses to fire in a tropical savanna grass, *Andropogon semiberbis*: a matrix model approach. *Journal of Ecology* **79**: 345-355.
- Singh, S. P. 1998. Chronic disturbance, a principal cause of environmental degradation in developing countries. *Environmental Conservation* **25**: 1-2.
- Sousa, W. P. 1984. The role of disturbance in natural communities. *Annual Review of Ecology & Systematics* **15**: 353-391.
- Thompson, J. R. 1999. *Simulation: a modeler's approach*. Wiley: New York. 297p.
- Tuljapurkar, S. y H. Caswell (eds.). 1997. *Structured-population models in marine, terrestrial, and freshwater systems*. Chapman and Hall: New York. 643 p.
- Ureta, C. 2006. Evaluación del disturbio antropogénico crónico sobre el comportamiento demográfico de dos especies de *Mammillaria* (Cactaceae). Tesis de licenciatura. Facultad de Ciencias. Universidad Nacional Autónoma de México.
- U. S. EPA. 2002. *Methods for evaluating wetland condition: land-use characterization for nutrient and sediment risk assessment*. Office of Water, U.S. Environmental Protection Agency: Washington, D.C. EPA-822-R-02-025.

- Valverde, P. L. y J. A. Zavala-Hurtado. 2006. Assessing the ecological status of *Mammillaria pectinifera* Weber (Cactaceae), a rare and threatened species endemic of the Tehuacán-Cuicatlán region in central Mexico. *Journal of Arid Environments* **64**: 193-208.
- Vega, E. y E. Peters. 2003. Conceptos generales sobre el disturbio y sus efectos en los ecosistemas. En: O. Sánchez, O. Monroy, E. Vega, E. Peters y C. Donovarros (comps.). *Conservación de ecosistemas templados de montaña en México*. Instituto Nacional de Ecología: México, D. F. Pp. 137-150.
- Watt, A. D. 1998. Measuring disturbance in tropical forests: a critique of the use of species-abundance models and indicator measures in general. *Journal of Applied Ecology* **35**: 467-469.
- White, P. S. y S. T. A. Pickett. 1985. Natural disturbance and patch dynamics: an introduction. En S. T. A. Pickett y P. S. White (eds.). *The ecology of natural disturbance and patch dynamics*. Academic Press, Inc: Burlington, Massachusetts. Pp. 3-16.

## Apéndice

# Programa MIPDAC

## A.1 Diagrama de flujo del algoritmo

En la figura A.1 presento el diagrama de flujo del programa que realicé. Como se observa, lo primero que se hace es la introducción de los datos, que incluye, además de la introducción de las muestras de tamaños de los individuos de las distintas poblaciones sujetas a DAC y el valor de la medición del DAC del sitio de estudio para cada población, la delimitación del espacio paramétrico, *i. e.*, establecer los límites del intervalo de valores que puede tomar cada parámetro. En mi caso, y como se describió en la sección 2.2.2, se tuvieron 11 intervalos que determinaron las intersecciones de las rectas dadas por la ecuación (3.2).

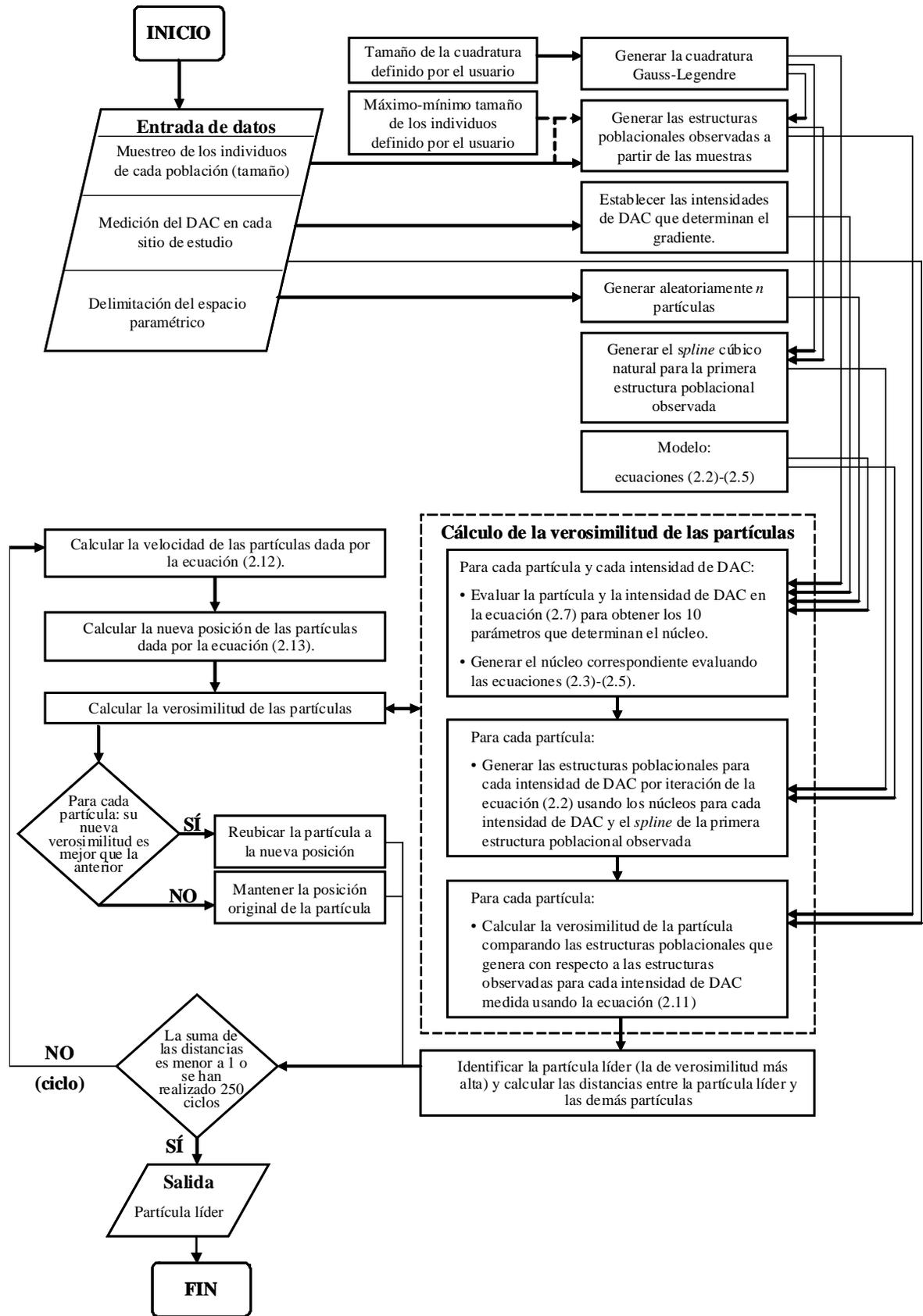


Figura A.1. Diagrama de flujo del algoritmo MIPDAC.

Posterior a esto, se genera la cuadratura Gauss-Legendre que se utilizará en el proceso de integración (la cuadratura puede tener como mínimo y máximo los tamaños mínimo y máximo observados en las muestras o tener valores preestablecidos por el usuario). Después, las muestras se utilizan para generar las estructuras poblacionales observadas, a la primera de las cuales se le ajusta un *spline* cúbico natural. Todas las estructuras se discretizan en la cuadratura.

El primer y el último valor de DAC observado se utilizan para establecer el conjunto de valores (de números naturales) que el programa considerará como el gradiente de DAC en el que se modelará la dinámica poblacional. Una vez que el programa tiene delimitado el espacio paramétrico a explorar, se genera un conjunto de  $n$  partículas, las cuales, en mi versión del MIPDAC fue de 25. A estas partículas se les calcula su verosimilitud; este cálculo involucra, para cada partícula:

1. Evaluar, para cada intensidad del gradiente de DAC, la partícula en la ecuación (3.6), obteniendo así los 11 parámetros que determinan el núcleo correspondiente a cada intensidad de DAC.
2. Generar el núcleo para cada intensidad de DAC usando las ecuaciones (2.3)-(2.5).
3. Generar las estructuras poblacionales para cada intensidad de DAC vía la iteración de la ecuación (2.2) usando los núcleos que generamos en el paso anterior y comenzando el proceso iterativo con el *spline* de la primera estructura poblacional observada.
4. Calcular la verosimilitud de la partícula comparando las estructuras poblacionales observadas con las estructuras poblacionales que se generaron en el paso anterior y que corresponden a las intensidades de DAC observadas.

Una vez que se tienen las verosimilitudes de todas las partículas se identifica la partícula líder (*i. e.*, la que tiene la mayor verosimilitud) y se calculan las distancias (euclidianas) a las que se ubican las demás partículas con respecto a la líder. Dichas distancias se suman y de ser esta suma menor a uno, la partícula líder es la óptima.

De no ser éste el caso, se entra a un ciclo que consiste en el paso III de la optimación por enjambre de partículas (OEP; sección 3.3.5), el cual se repite hasta que la distancia entre las partículas y la líder es menor a uno o hasta que se hayan realizado 250 ciclos.

## A.2 Código del programa

El programa MIPDAC se realizó dentro del editor de funciones de Matlab y consta de seis funciones. Estas funciones corresponden a:

- 1) Entrada de datos.
- 2) Generación de la cuadratura Gauss-Legendre.
- 3) Evaluación, a partir de una partícula e intensidad de DAC particulares, de la ecuación (2.7) para obtener los 11 parámetros que determinan el núcleo.
- 4) Generación, a partir de un conjunto de 11 parámetros, del núcleo asociado a estos parámetros evaluando las ecuaciones (2.3)-(2.5).
- 5) Generación de las estructuras poblacionales para cada intensidad del gradiente de DAC y cálculo de la verosimilitud de las estructuras generadas por la iteración de los núcleos con respecto a las estructuras poblacionales observadas.
- 6) Ubicación de aquella partícula que maximiza la verosimilitud.

Esta última función contiene los demás apartados del diagrama de flujo (Fig. A.1) que no se ubican en una función particular.

Presento a continuación las funciones desarrolladas, de acuerdo a la categoría a la que pertenecen. Las funciones están en el tipo de letra Courier para distinguirlas del resto del texto. El símbolo % es el usado por Matlab para hacer anotaciones dentro de la función y no forman parte del código de ésta.

## A.2.1 Entrada de datos

La función `datos.m` permite introducir la información básica que el programa requiere: las estructuras poblacionales observadas, los momentos de DAC observados y los intervalos que integran el espacio paramétrico:

```
function datos
global dac estructura minimo maximo nmm
dac = input('Dame el vector de valores de disturbio antropogenico
cronico:\n ');
estructura = input('Dame la matriz de tamaños de los individuos para los
distintos momentos de disturbio:\n ');
if length(dac) ~= length(estructura(1,:))
    error('El tamaño del vector de disturbios no coincide con el numero
de columnas de la matriz de tamaños');
end
minimo = zeros(1,nmm);
maximo = zeros(1,nmm);
minimo = input('Dame el vector de los valores minimos de los
parametros:\n');
maximo = input('Dame el vector de los valores maximos de los
parametros:\n');
if length(minimo) ~= length(maximo)
    error('El tamaño de los vectores no coincide');
end
for i = 1:nmm
    if maximo(i) <= minimo(i)
        error('El maximo es menor o igual al minimo');
    end
end
end
```

## A.2.2 Cuadratura Gauss-Legendre

Como se dijo en la sección 2.2, la generación de núcleos implica un proceso iterativo, dado por la ecuación (2.2), que involucra una integración de funciones. Para ello se usó el método de integración de Gauss-Legendre. Este método parte de un conjunto de valores de  $X$ , llamado cuadratura, en los cuales se evaluará la función a integrar para discretizarla, y un conjunto de pesos. `cuadrgl.m` es la función que obtiene la cuadratura y requiere como entrada los puntos extremos del intervalo continuo donde se va a hacer la integración ( $x_1$  y  $x_2$ ) y el tamaño de la cuadratura ( $n_i$ ); y arroja la cuadratura ( $c$ ) y los pesos asociados ( $w$ ).

```
function [c,w] = cuadrgl(x1,x2,ni)
eps = 3.0e-14;
m = round((ni+1)/2);
% como la cuadratura Gauss-Legendre es en el intervalo [-1,1]
% hay que transformar el intervalo [x1,x2] al [-1,1] via xm y x1
xm = 0.5*(x2+x1);
x1 = 0.5*(x2-x1);
for i = 1:m
    z = cos(pi*(i-.25)/(ni+.5));
    p1 = 1;
    p2 = 0;
    for j = 1:ni
        p3 = p2;
        p2 = p1;
        p1 = ((2*j-1)*z*p2-(j-1)*p3)/j;
    end
    pp = ni*(z*p1-p2)/(z*z-1);
    z1 = z;
    z = z1-p1/pp;
while abs(z-z1) > eps
    p1 = 1;
    p2 = 0;
    for j = 1:ni
        p3 = p2;
```

```

        p2 = p1;
        p1 = ((2*j-1)*z*p2-(j-1)*p3)/j;
    end
    pp = ni*(z*p1-p2)/(z*z-1);
    z1= z;
    z = z1-p1/pp;
end
c(i) = xm-xl*z; % se tiene una cuadratura simetrica
c(ni+1-i) = xm+xl*z;
w(i) = 2*xl/((1-z*z)*pp*pp); % se tienen pesos simetricos
w(ni+1-i) = w(i);
end

```

## A.2.3 Obtención de parámetros que integran el núcleo

Dada una partícula y una intensidad de DAC particulares, es necesaria una función que tome los valores de dicha partícula y determine los valores de los 11 parámetros correspondientes a cada momento de DAC  $\theta(t)$  (evaluando  $\theta(t)$  en la ecuación (3.6)). `parametros.m` es la función que toma la partícula (`p`) y el momento de DAC (`t`) y arroja los 11 valores que definen las funciones que integran al núcleo:

```

function [sat,sbt,cat,cbt,cct,cdt,cet,f1at,f1bt,f2at,f2bt] =
parametros(t,p)
global cda cdb
sat = p(1)+p(2)*t;
sbt = p(3)+p(4)*t;
cat = p(5)+p(6)*t;
cbt = p(7)+p(8)*t;
cct = p(9)+p(10)*t;
cdt = cda+cdb*cct+p(21);
cet = p(11)+p(12)*t;
f1at = p(13)+p(14)*t;

```

```

f1bt = p(15)+p(16)*t;
f2at = p(17)+p(18)*t;
f2bt = p(19)+p(20)*t;

```

## A.2.4 Generación del núcleo

Una vez que se tienen los 11 valores para una partícula e intensidad de DAC particulares, éstos se usan en la generación del núcleo. La función `nucleo.m` toma estos 11 valores, además de `x1`, `x2` y `ni`, y arroja el núcleo evaluado en la cuadratura `c`:

```

function nxy = nucleo(sa,sb,ca,cb,cc,cd,ce,f1a,f1b,f2a,f2b,x1,x2,ni)
global flmin wx x y
sx = exp(sa+sb.*x)./(1+exp(sa+sb.*x));
ax = ca+cb.*x;
bx = (cc.*x.^cd+ce)./ax;
cxy = wx*gampdf(y,ax,bx);
if sum(cxy) ~= 0
    cxy = cxy./(ones(ni,1)*sum(cxy));
end
flx = zeros(ni,ni);
flx(:,flmin:ni) = exp(f1a).*x(:,flmin:ni).^f1b;
f2y = wx*gampdf(y,f2a,f2b);
if sum(f2y) ~= 0
    f2y = f2y./((sum(f2y)'+ones(1,ni)));
end
nxy = sx.*cxy+flx.*f2y;

```

## A.2.5 Generación de las estructuras y cálculo de la verosimilitud de la partícula

Ya que se tienen los núcleos correspondientes a cada intensidad del gradiente de DAC evaluados en la cuadratura para una partícula ( $p$ ) particular, se tiene que calcular la verosimilitud que ésta tiene con respecto a las estructuras poblacionales observadas (frecuencias). La función `verosimilitud.m` toma lo anterior y arroja un valor de verosimilitud ( $v$ ):

```
function v = verosimilitud(p)
global t x1 x2 ni frecuencias ex1
estructuras = [];
inex = ex1';
for i = 1:(length(t)-1)
    for j = t(i):(t(i+1)-1)
        [sa,sb,ca,cb,cc,cd,ce,f1a,f1b,f2a,f2b] = parametros(j,p);
        nxy = nucleo(sa,sb,ca,cb,cc,cd,ce,f1a,f1b,f2a,f2b,x1,x2,ni);
        inex = nxy*inex;
        inex = inex./sum(inex);
    end
    inex = max(1e-323,inex);
    estructuras = [estructuras inex];
end
v = sum(sum(frecuencias.*log(estructuras)));
```

## A.2.6 Maximización de la verosimilitud

La función principal que incorpora a la OEP, así como al establecimiento de las intensidades de DAC que determinan el gradiente, la generación de las partículas y el *spline* cúbico natural de la

primera estructura observada, es `psoreal.m`. Esta función toma como entrada el número de partículas que va a tener nuestro enjambre (`np`), el número de parámetros que integran el núcleo (`nmm`) y el número de corridas que la OEP realizará (`ne`). El resultado que arroja esta función es aquella partícula que, después de `ne` corridas, maximiza la verosimilitud.

```
function psoreal(np,nmm,ne)
global dac t estructura minimo maximo x1 x2 ni cda cdb dat c x y w wx
flmin frecuencias exl
% contador aleatorio
for i = 1:sum(clock)
    p = random('unif',0,1);
end
format long;
% valores defaults de los parametros np, nmm y ne
if nargin < 1 | isempty(np)
    np = 25;
end
if nargin < 2 | isempty(nmm)
    nmm = 10;
end
if nargin < 3 | isempty(ne)
    ne = 20;
end
m = 2*nmm+1; % numero de parametros que integran una partícula
x1 = 0.0035; % tamaño minimo que puede tener un individuo (cm3)
x2 = 1200; % tamaño maximo que puede tener un individuo (cm3)
ni = 150; % numero de intervalos que tiene la cuadratura
cda = 1.06709675; % intercepto de la recta que describe la
                % correlación entre cc y cd
cdb = -0.08808192; % pendiente de la recta que describe la
                % correlación entre cc y cd
if dat ~= 1
    datos; % llama a la funcion datos.m
end
% redondea, ordena y verifica que los valores de DAC no
% coincidan
t = round(dac);
```

```

[t,r] = sort(t);
for i = 1:(length(t)-1)
    for j = (i+1):length(t)
        if t(i) == t(j)
            error('Dos o mas momentos de disturbio coinciden (en valor
redondeado)');
        end
    end
end
% ordena y acota los valores de las estructuras de tamaños
estructural = estructura;
for i = 1:length(t)
    estructura(:,i) = estructural(:,r(i));
    for j = 1:max(find(estructura(:,i)))
        estructura(j,i) = min(x2,max(x1,estructura(j,i)));
    end
end
% calcula los valores de la cuadratura y los ejes x e y
c = cuadrgr1(x1,x2,ni); % llama a la funcion cuadrgr1.m
z = zeros(2,ni);
w = zeros(ni,1);
z(1,1) = x1;
z(2,ni) = x2;
for i = 1:ni-1
    z(2,i) = 2*(c(i)-z(1,i))+z(1,i);
    z(1,i+1) = z(2,i);
    w(i) = z(2,i)-z(1,i);
end
w(ni) = z(2,ni)-z(1,ni);
wx = diag(w);
x = c;
y = x;
[x,y] = meshgrid(x,y);
% ubica al individuo de menor tamaño > 7.5cm3
flmin = min(find(max(0,c-7.5)));
% calcula las frecuencias observadas de tamaños a partir de las
% estructuras observadas
frecuencias = zeros(ni,length(t)-1);

```

```

for i = 1:length(t)-1
    for j = 1:max(find(estructura(:,i+1)))
        for k = 1:ni
            if estructura(j,i+1) > z(1,k) & estructura(j,i+1) <= z(2,k)
                frecuencias(k,i) = frecuencias(k,i)+1;
            end
        end
    end
end

end

% calcula la primera estructura de tamaños
q = 50; % tamaño de la particion de la primera estructura
        % poblacional
estruc1f = zeros(1,q);
estruc1x = cuadr1(x1,x2,q);
zex1 = zeros(2,q);
zex1(1,1) = x1;
zex1(2,q) = x2;
for i = 1:q-1
    zex1(2,i) = 2*(estruc1x(i)-zex1(1,i))+zex1(1,i);
    zex1(1,i+1) = zex1(2,i);
end

for i = 1:max(find(estructura(:,1)))
    for j = 1:q
        if estructura(i,1) > zex1(1,j) & estructura(i,1) <= zex1(2,j)
            estruc1f(j) = estruc1f(j)+1;
        end
    end
end

end

ex1 = csaps(estruc1x,estruc1f,1,c); % llama a la funcion
% generadora de splines cubicos naturales incorporada a Matlab
% csaps.m
ex1 = ex1./sum(ex1);
ex1 = max(1e-323,ex1);
gBests = zeros(ne,m);
vgBests = zeros(1,ne);
for z = 1:ne
    % inicia las particulas y verifica que sean "buenas"
    % particulas

```

```

z
p = zeros(np,m);
for i = 1:np
    contador = 0;
    while contador ~= length(t)
        for j = 1:nmm
            p(i,2*j-1) = random('unif',minimo(j),maximo(j));
            p(i,2*j) = random('unif',(minimo(j)-p(i,2*j-
1))/max(t),(maximo(j)-p(i,2*j-1))/max(t));
        end
        p(i,21) = random('unif', minimo(nmm+1),maximo(nmm+1));
        contador = 0;
        for j = 1:length(t)
            [sat,sbt,cat,cbt,cct,cdt,cet,flat,f1bt,f2at,f2bt] =
parametros(t(j),p(i,:));
            if ((cat > 0 & cbt > 0) | (cat > 0 & cbt < 0 & -cat/cbt
>= x2) | (cat < 0 & cbt > 0 & -cat/cbt <= x1)) & ((cct > 0 & cet >= 0) |
(cct > 0 & cet < 0 & cdt >= 0 & x1^cdt >= -cet/cct) | (cct > 0 & cet <
0 & cdt >= 0 & x2^cdt >= -cet/cct) | (cct < 0 & cet > 0 & cdt >= 0 &
x2^cdt <= -cet/cct) | (cct < 0 & cet > 0 & cdt < 0 & x1^cdt <= -
cet/cct)) & f2at > 0 & f2bt > 0
                contador = contador+1;
            end
        end
    end
end
end
% calcula la verosimilitud de las particulas
pBest = zeros(np,m);
vpBest = zeros(1,np);
disp('vpBests');
for i = 1:np
    pBest(i,:) = p(i,:);
    vpBest(i) = verosimilitud(pBest(i,:)); % llama a la
% funcion verosimilitud.m
    if isfinite(vpBest(i)) == 0 | vpBest(i) == 0
        while isfinite(vpBest(i)) == 0 | vpBest(i) == 0
            contador = 0;
            while contador ~= length(t)

```

```

        for j = 1:nmm
            p(i,2*j-1) = random('unif',minimo(j),maximo(j));
            p(i,2*j) = random('unif',(minimo(j)-p(i,2*j-1))/max(t),
1)/max(t),(maximo(j)-p(i,2*j-1))/max(t));
        end
        p(i,21) = random('unif',
minimo(nmm+1),maximo(nmm+1));
        contador = 0;
        for j = 1:length(t)
            [sat,sbt,cat,cbt,cct,cdt,cet,flat,f1bt,f2at,f2bt] =
parametros(t(j),p(i,:)); % llama a la funcion parametros.m
                if ((cat > 0 & cbt > 0) | (cat > 0 & cbt < 0 & -
cat/cbt >= x2) | (cat < 0 & cbt > 0 & -cat/cbt <= x1)) & ((cct > 0 & cet
>= 0) | (cct > 0 & cet < 0 & cdt >= 0 & x1^cdt >= -cet/cct) | (cct > 0
& cet < 0 & cdt >= 0 & x2^cdt >= -cet/cct) | (cct < 0 & cet > 0 & cdt >=
0 & x2^cdt <= -cet/cct) | (cct < 0 & cet > 0 & cdt < 0 & x1^cdt <= -
cet/cct)) & f2at > 0 & f2bt > 0
                    contador = contador+1;
                end
            end
        end
        pBest(i,:) = p(i,:);
        vpBest(i) = verosimilitud(pBest(i,:)); % llama a
% la funcion verosimilitud.m
        end
    end
    disp(vpBest(i));
end
% ubica los mejores vecinos
nBest = zeros(np,m);
for i = 1:np
    for j = 1:m
        [q,r] = max((vpBest-vpBest(i))*ones(1,np))./(abs(pBest(:,j)'-
pBest(i,j))*ones(1,np))+1));
        nBest(i,j) = pBest(r,j);
    end
end
% ubica la mejor particula de las encontradas en este ciclo

```

```

[vgBest,ngBest] = max(vpBest);
gBest = pBest(ngBest,:);
disp('particula lider');
disp(ngBest);
disp('gBest');
disp(gBest');
disp('vgBest');
disp(vgBest);
% establece las velocidades iniciales
velmax = (max(pBest)-min(pBest))/10; % velocidad maxima que
% puede tener la entrada de una particula
vel = zeros(np,m);
for i = 1:np
    for j = 1:m
        q = round(rand);
        vel(i,j) = q*velmax(j)-(1-q)*velmax(j);
    end
end
weightvel = zeros(np,m);
clrandom = zeros(np,m);
c2random = zeros(np,m);
c3random = zeros(np,m);
rebote = zeros(np,nmm+1);
c1 = 2; % ver ecuación (3.5)
c2 = 2;
c3 = 2;
u = 0;
dif = 1000;
histvpBest = [];
histgBest = [];
histvgBest = [];
hstdif = [];
histangulo = [];
histvelocidad = [];
while dif > 1 & u <= 250
    u = u+1
    if u == 50
        c1 = 1.5;

```

```

        c2 = 1.5;
    end
    if u == 100
        c1 = 1;
        c2 = 1;
    end
    if u == 150;
        c1 = 0.5;
        c2 = 0.5;
    end
    if u == 200;
        c1 = 0;
        c2 = 0;
    end
    end
    % calcula velocidades
    weight = random('norm',0.9,0.2,np,1);
    c1r = c1*rand(np,1);
    c2r = c2*rand(np,1);
    c3r = c3*rand(np,1);
    for i = 1:np
        weightvel(i,:) = weight(i)*vel(i,:);
        c1random(i,:) = c1r(i);
        c2random(i,:) = c2r(i);
        c3random(i,:) = c3r(i);
    end
    vel1 = vel(ngBest,:);
    vel = weightvel+c1random.*(pBest-p)+c2random.*(nBest-
pBest)+c3random.*(ones(np,1)*gBest-pBest); % ecuacion (3.9)
    vel = min(ones(np,1)*velmax,max(-ones(np,1)*velmax,vel));
    angulo = subspace(vel1',vel(ngBest,:));
    angulo = convang(angulo,'rad','deg')
    velocidad = sqrt(sum(vel(ngBest,:).^2))
    % rebote
    for i = 1:np
        for j = 1:nmm+1
            if (pBest(i,2*j-1) == maximo(j) | pBest(i,2*j-1) ==
minimo(j)) & rebote(i,j) == 0
                vel(i,2*j-1) = -vel(i,2*j-1);
            end
        end
    end
end

```

```

        rebote(i,j) = 1;
    end
end
end
% reubica las particulas: ecuacion (3.9)
for i = 1:nmm
    p(:,2*i-1) =
min(maximo(i)*ones(np,1),max(minimo(i)*ones(np,1),pBest(:,2*i-
1)+vel(:,2*i-1)));
        p(:,2*i) = min((maximo(i)*ones(np,1)-
p(:,2*i))/max(t),max((minimo(i)*ones(np,1)-
p(:,2*i))/max(t),pBest(:,2*i)+vel(:,2*i)));
    end
    p(:,21) =
min(maximo(nmm+1)*ones(np,1),max(minimo(nmm+1)*ones(np,1),pBest(:,21)+ve
l(:,21)));
    dif = sum(sqrt(sum(((ones(np,1)*gBest-pBest).^2)))
    v = vpBest;
    disp('vpBests');
    for i = 1:np
        contador = 0;
        for j = 1:length(t)
            [sat,sbt,cat,cbt,cct,cdt,cet,f1at,f1bt,f2at,f2bt] =
parametros(t(j),p(i,:));
            if ((cat > 0 & cbt > 0) | (cat > 0 & cbt < 0 & -cat/cbt
>= x2) | (cat < 0 & cbt > 0 & -cat/cbt <= x1)) & ((cct > 0 & cet >= 0) |
(cct > 0 & cet < 0 & cdt >= 0 & x1^cdt >= -cet/cct) | (cct > 0 & cet <
0 & cdt >= 0 & x2^cdt >= -cet/cct) | (cct < 0 & cet > 0 & cdt >= 0 &
x2^cdt <= -cet/cct) | (cct < 0 & cet > 0 & cdt < 0 & x1^cdt <= -
cet/cct)) & f2at > 0 & f2bt > 0
                contador = contador+1;
            end
        end
    if contador == length(t)
        v(i) = verosimilitud(p(i,:));
        if isfinite(v(i)) == 1 & v(i) >= vpBest(i)
            pBest(i,:) = p(i,:);
            vpBest(i) = v(i);
        end
    end
end
end

```

```

        disp(vpBest(i));
        rebote(i,:) = zeros(1,nmm+1);
    end
end
end
for i = 1:np
    for j = 1:m
        [q,r] = max((vpBest-
vpBest(i)*ones(1,np))./(abs(pBest(:,j))-pBest(i,j)*ones(1,np))+1));
        nBest(i,j) = pBest(r,j);
    end
end
if max(vpBest) > vgBest
    [vgBest,ngBest] = max(vpBest);
    gBest = pBest(ngBest,:);
    disp('particula lider');
    disp(ngBest);
    disp('gBest');
    disp(gBest');
end
% reubica las peores particulas
if u > 150
    [q,r] = min(vpBest);
    pBest(r,:) = gBest;
    vpBest(r) = vgBest;
    rebote(r,:) = zeros(1,nmm+1);
end
disp('vgBest');
disp(vgBest);
histvpBest = [histvpBest vpBest'];
histgBest = [histgBest gBest'];
histvgBest = [histvgBest vgBest];
hisdif = [hisdif dif];
histangulo = [histangulo angulo];
histvelocidad = [histvelocidad velocidad];
save(num2str(z),'histvpBest','histgBest','histvgBest','u','hisdif','his
tvelocidad');
gBests(z,:) = gBest;

```

```
        vgBests(z) = vgBest;
    end
end
% presenta la particula optima encontrada despues de ne corridas
[vgBest,r] = max(vgBests);
gBest = gBests(r,:);
disp('El conjunto de parametros que maximiza la verosimilitud es: ');
disp(gBest');
disp('con una verosimilitud de: ');
disp(vgBest);
```