



6
UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

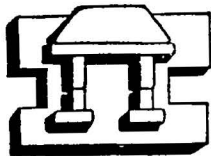
ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES
CAMPUS IZTACALA

LA ENTROPIA COMO MEDIDA DE INCERTIDUMBRE DE
MODELOS DE DISTRIBUCION DE ESPECIES.

T E S I S

QUE PARA OBTENER EL TITULO DE
B I O L O G A
P R E S E N T A
GEORGINA ALCANTAR LOPEZ

ASESOR: DR. LUIS ANTONIO BOJORQUEZ TAPIA



IZTACALA

LOS REYES IZTACALA, EDO. DE MEX.

2007

TESIS CON
FALLA DE ORIGEN



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

PAGINACION DISCONTINUA

A

Gabriela, Raúl, Alonso y Miguel.

Josefina y Raúl

Vladimir

A mi familia, sin ella no hubiese llegado a donde ahora estoy.

A Gabriela, mi hermana, por su ávida curiosidad sobre las cosas que no entiende y para las que siempre espera una explicación de mi parte (que puedo asegurar no siempre es la mejor).

A Vladimir por ser el compañero perfecto de muchos de mis momentos. Por ser uno de los motores que impulsan mi vida.

A todos los amigos que a lo largo de mi formación como bióloga encontré:

A los Omahos compañeros de enseñanzas y maestros de muchas otras: Nohemi, compañera y contrincante de tantas y tantas discusiones; Penelope con su constante mal humor, dispuesta siempre a verle el lado pútrido a las cosas y cómplice perfecta; César célebre antagonista de Penelope y mio, siempre amable, dispuesto y cariñoso para ser la cálida pared en donde terminaban mis ansiedades; Raymundo por su amistad breve, pero sincera y Fabricio, como él mismo lo dice *alter ego*, no más.

A mis amigos y compañeros del ex-laboratorio de Análisis Ambientales. Si me hubieran dicho las cosas que me pasarían en ese lugar no las hubiera creído. A Majose por enseñarme que la paciencia es una actitud que hay que cultivar todos los días; Gerardo, amigo y cómplice reciente; Daniela por su confianza y compañía en el proceso de crecimiento; Salomón porque con sus negativas me hizo independiente. A los internacionales. Liliana por su excelente humor y entrañable amistad; Luli por su amistad que a pesar de los kilómetros se mantiene; Rodrigo por ser quién hasta en mis peores momentos consiguió hacerme reír a carcajadas, excelente amigo, compañero y cómplice pero sobre todo ponchin bag. A Camilo por su paciencia.

A Luis Bojórquez por su manera tan peculiar de enseñar y por su paciencia infinita a pesar de mi insistencia en llevarlo a sus límites. Estoy segura que en más de una ocasión se arrepintió de haberme aceptado en su equipo.

A Mariano Hernández por su apoyo en lo profesional y en lo personal. Sin él, este trabajo no hubiera sido posible aún con toda la paciencia de Luis y toda mi disposición.

A los profesores que me sufrieron y de quienes, las más de las veces, obtuve paciencia.

A quienes revisaron este trabajo, por ayudarme a aterrizar las ideas.

A todos, *gracias*.

INDICE GENERAL

1. RESUMEN	1
2. INTRODUCCIÓN	2
3. MARCO TEÓRICO	4
3.1. LA INCERTIDUMBRE EN LOS MODELOS ESPACIALES	5
3.2. INTERPOLACIÓN	7
3.3. ENTROPÍA Y TEORÍA DE LA INFORMACIÓN	8
4. MÉTODO	10
4.1. MODELO DE DISTRIBUCIÓN	14
4.2. BASE DE DATOS	14
4.3. INTERPOLACIÓN	15
4.4. SIMULACIÓN ESTOCÁSTICA	16
4.5. CÁLCULO DE ENTROPÍA	17
5. RESULTADOS	19
5.1. MODELO DE DISTRIBUCIÓN	19
5.2. ANÁLISIS CLIMATOLÓGICO	21
5.3. INTERPOLACIÓN	21
5.4. SIMULACIÓN ESTOCÁSTICA	25
5.5. ENTROPÍA	25
6. DISCUSIÓN	32
7. CONCLUSIONES	37
8. BIBLIOGRAFÍA	38

INDICE DE FIGURAS

Figura 1. Ubicación geográfica del área de estudio.....	11
Figura 2. Características del área de estudio (escala 1 000 000, INEGI 1994). a)- Tipos de vegetación y uso del suelo y b)- Altura sobre el nivel del mar.....	12
Figura 3. Mapa con el número de registros de colectas para las 9 especies de aves (a) y el modelo ajustado en GLIM (b).	20
Figura 4. Mapas de temperatura (a) y precipitación (b) interpolados con <i>kriging</i>	22
Figura 5. Errores de <i>kriging</i> en la interpolación de la temperatura (a) y la precipitación (b).....	23
Figura 6. Mapa de entropía en los resultados simulados para el modelo de distribución de especies (a) y área seleccionada (b) con base en la eficiencia del modelo según la teoría de la información.26	

EL S CON
FALLA DE ORIGEN

INDICE DE TABLAS Y GRÁFICAS

Tabla 1. Extensión en hectáreas y porcentaje de cada tipo de vegetación (escala 1:1 000 000, INEGI 1994) en el área de estudio.....	13
Tabla 2. Prueba de T-Student de diferencia entre medias para los datos de precipitación y temperatura observados y simulados (con g.l. de 150 y 100 para temperatura y de 175 y 152 para precipitación.....	24
Gráfica 1. Valores de la entropía (H) calculada por pixel con base en el número de respuestas diferentes en cada uno; la línea superior indica la H máxima del sistema en 5.09 nats.	27
Gráfica 2. Predicción del número de individuos con los valores interpolados promedio (P) de precipitación y temperatura. Además, se muestra la predicción mínima (Min) y máxima simulada (Max) por pixel.	27
Gráfica 3. Valores de redundancia relativa (R) de las predicciones del modelo (información en el sistema) y de eficiencia (E) del modelo en transmitir esa información.	29
Gráfica 4. Valores de entropía (H) y coeficiente de variación (CV) de los resultados simulados por el modelo de distribución de especies.	29
Gráfica 5. Valores del coeficiente de variación (CV), eficiencia del modelo (E) y redundancia relativa en las predicciones del modelo por pixel. El recuadro muestra el área seleccionada en función de la exactitud y precisión de las predicciones.	31

1. RESUMEN

El aumento en la explotación de los recursos naturales como consecuencia del crecimiento poblacional provoca la pérdida de la diversidad biológica. Para limitar los efectos de estos procesos es necesario identificar las áreas naturales que deban ser conservadas por la diversidad biológica que albergan. Sin embargo, la información sobre la distribución geográfica de las especies presentes en una región pocas veces existe. En efecto, los modelos de distribución de especies proporcionan la información que, como primer acercamiento, permite reconocer las áreas en que potencialmente se encuentran dichas especies. No obstante, el principal problema que se enfrenta en la construcción de este tipo de modelos es que pocas veces se evalúan los niveles de incertidumbre de sus predicciones.

El propósito final de este trabajo es presentar el cálculo de la entropía como una alternativa en la evaluación de la incertidumbre que afecta a los modelos de distribución de especies. Para ello se utilizó la información de colecta de 9 especies de aves en la Sierra Madre Occidental y se construyó un modelo *ad hoc* para predecir su distribución. Además, se recurrió a la geoestadística y la simulación estocástica para generar valores aleatorios de precipitación y temperatura en la región de estudio, a partir de la información de las estaciones climatológicas dentro del área. Con los valores de precipitación y temperatura y el modelo construido, se simuló la distribución de las especies cuyos resultados se utilizaron en el cálculo de la entropía en el índice de Shannon.

La entropía se utilizó como una medida de precisión en las predicciones del modelo, a diferencia del coeficiente de variación que fue una medida de la exactitud de las mismas. En función de lo anterior, se determinaron las áreas en donde las predicciones pueden tener mayor exactitud sin perder precisión. Con los niveles de incertidumbre caracterizados se podría apoyar la toma de decisiones para la protección de la biodiversidad.

2. INTRODUCCIÓN

Ante la pérdida de la diversidad biológica, como consecuencia del crecimiento poblacional y la extracción desmedida de los recursos naturales, es necesario proteger las zonas con mayor diversidad a través de programas que permitan la conservación de su patrimonio biológico (Margules *et al.* 1989). Sin embargo, la información sobre la distribución de las especies que permita la selección de las áreas a proteger en una región pocas veces está disponible. Pese a esto, no es posible aplazar las decisiones de conservación hasta que se obtengan suficientes datos de campo a causa de las altas tasas de transformación de la cubierta natural que existen en todo el mundo (Davis *et al.* 1990).

El conocimiento sobre la distribución de las especies es incompleto e impreciso pues los sitios de colecta están generalmente cerca de las vías de comunicación o de las principales poblaciones (Bojórquez-Tapia *et al.* 1994). En consecuencia, existen huecos en el resto de la región que pueden ser compensados con el uso de técnicas de modelación espacial que combinen sistemas de información geográfica (SIG) y métodos estadísticos multivariados. En México, estas técnicas se han empleado para predecir la riqueza de especies e identificar áreas críticas para la conservación, a partir de registros de colecta de especímenes con referencia geográfica y capas cartográficas digitales (Bojórquez-Tapia *et al.* 1994, Cruz 1996). En particular, Bojórquez-Tapia *et al.* 1994 y Cruz 1996 combinan los Modelos Lineales Generalizados, GLIM (Nicholls 1989), y el Análisis de Correspondencias de Especies por Variables Ambientales, CASEV (Montaña y Greig-Smith 1990) para tal fin. Estos procedimientos multivariados permiten encontrar las variables ambientales con mayor influencia en la presencia de las distintas especies y que, en la práctica, resultan más útiles para modelar su distribución (Ezcurra *et al.* 1987, Margules y Stein 1989 y Nicholls 1989).

Procedimientos como los ya mencionados carecen de una evaluación formal de los niveles de incertidumbre asociados a los modelos estadísticos. Según Borhr y Heisenberg (citados por Brillouin 1969), la incertidumbre es el límite en la exactitud del mecanismo de medición. En los modelos de distribución de especies, los registros de especies geográficamente referenciados y los mapas temáticos equivalen a tales mecanismos. El fundamento de esta equivalencia se arraiga en las fuentes de incertidumbre que afectan los

modelos espaciales como la variación natural de los sistemas físicos y biológicos (Haefner 1996), el cálculo de parámetros estadísticos (Summers *et al.* 1993, Monte *et al.* 1996) y la interpolación espacial de los datos (Phillips *et al.* 1996).

En el momento que se determina estadísticamente el nivel de incertidumbre de los modelos espaciales es posible conocer los límites de confianza de las predicciones (Summers *et al.* 1993). Así entonces, es posible evaluar la protección que ofrece cada área a las diferentes especies cuando se han identificado los niveles de confianza de los modelos utilizados en su diseño. Uno de los métodos que se utilizan para analizar los niveles de incertidumbre en los modelos espaciales está basado en el concepto de entropía.

El concepto de entropía proviene de la Teoría de la información (Shannon 1949; citado por George 1965), cuyo planteamiento establece que la información en un sistema no es única y estable, sino una amplia gama de posibilidades que puede generar una respuesta particular. Así, la incertidumbre es la falta de conocimiento *a priori* sobre la respuesta generada por el sistema. Entonces, un sistema con alta incertidumbre tiene muchas posibilidades de respuesta, mucha información y, por lo tanto, alto contenido de entropía (Raisbeck 1963, George 1965, Wicken 1987).

Para evaluar la entropía del modelo de distribución de especies es necesario evaluar todos los resultados posibles sobre la ocurrencia de las especies que modela, a través de métodos de simulación estocástica. Con este propósito se utilizó la geoestadística y la simulación estocástica para generar valores aleatorios de precipitación y temperatura, que a su vez fueron utilizados en las simulaciones de la distribución de 9 especies de aves en un área de la Sierra Madre Occidental.

Con base en la eficiencia del modelo y al coeficiente de variación se determinaron las zonas de predicción con mayor certidumbre. Son estas áreas donde se puede tener mayor confianza en la predicción sobre la presencia de las especies modeladas y, en consecuencia, donde se pueden proteger de manera más eficiente.

3. MARCO TEÓRICO

La pérdida de la diversidad biológica puede ser limitada con la protección de áreas naturales (Terborgh *et al.* 1983, Miller *et al.* 1986, Flores-Villela *et al.* 1988, Margules *et al.* 1988. Margules *et al.* 1989, Scott 1990). Para ubicar una reserva natural es necesario conocer la distribución y abundancia de las especies en una región. Sin embargo, la información disponible se refiere casi siempre a la presencia y abundancia de las especies más evidentes. Además, los registros de colecta se concentran en las cercanías de las principales poblaciones y vías de comunicación produciendo huecos de información en el resto de la región (Williams *et al.* 1992, Margules *et al.* 1987, Bojórquez-Tapia *et al.* 1994).

Frente a esta escasez de información, los modelos estadísticos espaciales han demostrado ser una alternativa en la descripción y simulación para la toma de decisiones en el manejo de recursos bióticos (Bojórquez-Tapia *et al.* 1990, Phillips *et al.* 1996, Corsi *et al.* 2000). Los modelos estadísticos permiten obtener la distribución y abundancia potencial de las especies a partir de los registros de colecta existentes y su relación con las variables ambientales de cada sitio (Ezcurra *et al.* 1987, Margules *et al.* 1988, Margules *et al.* 1989, Montaña *et al.* 1990, Hone 1991, Miller *et al.* 1994, Bojórquez-Tapia *et al.* 1995).

Los modelos de distribución se han favorecido con el uso de los sistemas de información geográfica (SIG) ya que incrementan la capacidad de análisis y permiten el manejo de grandes cantidades de información. Un SIG es un sistema especializado de base de datos que se distingue por su capacidad para almacenar, consultar, manejar y analizar datos con referencia geográfica (Azuaa *et al.* 1994, Burroughs 1996, Agumya *et al.* 1999, Elston *et al.* 1993). Los SIG facilitan el estudio de dinámicas espaciales complejas como el comportamiento de las especies en espacio y tiempo (Bracken *et al.* 1990, Miller 1994, Griffith *et al.* 1999) por lo que su uso es conveniente en proyectos de conservación y manejo de los recursos bióticos. No obstante, aún con la ayuda de los SIG, la eficacia de los modelos espaciales depende de la cantidad y calidad de las observaciones en campo (Braithwaite 1991).

TESIS CON
FALLA DE ORIGEN

3.1. LA INCERTIDUMBRE EN LOS MODELOS ESPACIALES

Hasta ahora, la construcción de los modelos espaciales carece de una evaluación formal del error que afecta sus resultados y las decisiones que se toman con ellos (Monte *et al.* 1996, Griffith *et al.* 1999, Heuvelink 1999). El error en el modelado espacial tiene dos componentes, la exactitud y la precisión. La exactitud se refiere a la cercanía entre el dato original y el promedio de la predicción del modelo (Summers *et al.* 1993, Mowrer 1999). La precisión, por su parte, se refiere a la variación inherente en las respuestas del modelo (Summers *et al.* 1993, Phillips *et al.* 1996). Un modelo tiene mayor precisión cuando obtiene el mismo resultado con los datos de diferentes repeticiones. Un modelo tiene poca precisión cuando produce resultados diferentes en cada repetición.

Desde luego, los errores en los modelos espaciales provocan incertidumbre en los resultados obtenidos. La incertidumbre es el efecto de los mecanismos de medición en la exactitud y precisión de las mediciones (Brillouin, 1969, Davis *et al.* 1997). En el caso de los modelos de distribución de especies estos mecanismos están representados en los registros de campo con referencia geográfica y los mapas temáticos.

En los modelos estadísticos, la incertidumbre puede existir por diferentes razones: la variación natural de los fenómenos medidos, la falta de precisión en la medición de las variables, la alteración de los valores durante el proceso de construcción, el error en la estimación de los parámetros y el error en la calibración del modelo (Gardner *et al.* 1990, Haefner 1996, Davis *et al.* 1997). No obstante, la información incompleta es quizá la causa más frecuente de esta incertidumbre (Zimmerman 2000). En consecuencia, las alternativas de solución para éstos problemas pueden ser: la medición de las variables en períodos de tiempo cortos o en escalas pequeñas que permitan la extrapolación; la evaluación del error en la estimación de los parámetros con base en análisis de frecuencias o la estimación de los intervalos de confianza (Gardner *et al.* 1990, Summers *et al.* 1993, Haefner 1996, Monte *et al.* 1996).

Klepper (1997) identifica las etapas en el modelado de los procesos ambientales que son afectadas por la incertidumbre: *i)* en los datos fuente, *ii)* en los valores iniciales y *iii)* en la estructura del modelo. Básicamente, existen dos puntos de vista diferentes en los métodos

utilizados para analizar el error en éstos modelos. La principal distinción entre ellos es el objetivo del análisis. Cuando se pretende conocer el error en la predicción como consecuencia de los cambios en los parámetros del modelo, se utiliza el análisis de sensibilidad (Hwang *et al.* 1998). Por el contrario, el análisis de incertidumbre se utiliza cuando el interés está en las causas del error en la predicción del modelo. En el análisis de incertidumbre se considera la variación inherente en los datos de entrada y sus efectos en los resultados del modelo. Mientras que el análisis de sensibilidad no pone atención en los efectos de la fuentes de información imprecisas o en la variación de los datos de entrada al modelo (Summers *et al.* 1993).

Con frecuencia se utilizan métodos de simulación para estimar la incertidumbre en la predicción de los modelos (Gardner *et al.* 1983, Summers *et al.* 1993, Phillips *et al.* 1996, Davis *et al.* 1997, Hwang *et al.* 1998, Gascoigne *et al.* 1999, Griffith *et al.* 1999). Estos métodos se basan en el muestreo aleatorio y repetitivo de los parámetros, las variables principales, las condiciones limitantes y las condiciones iniciales de las variables en el modelo. En dichos análisis, se han utilizado como medidas de incertidumbre el cálculo del coeficiente de determinación (r^2 , Gardner *et al.* 1990); la varianza; la desviación estándar (CCDF, Summers *et al.* 1996) o el cálculo de los límites de confianza (Monte *et al.* 1996, Hakanson 1996), sin considerar la propagación espacial de la incertidumbre en las predicciones.

El desarrollo de los sistemas de información geográfica hace posible el análisis espacial de la incertidumbre en las predicciones de los modelos. Esta incertidumbre se reconoce cuando los resultados obtenidos tienen diferente ubicación geográfica durante el proceso de simulación (Hwang *et al.* 1998, Gascoigne *et al.* 1999). Así, por ejemplo, Phillips *et al.* (1996) evalúan los errores espaciales de la interpolación de las variables ambientales y su efecto en el cálculo del potencial de evapotranspiración (PET, por sus siglas en inglés) en la Cuenca del Río Columbia en E.U.A. con el coeficiente de variación como el estimador para identificar las zonas de mayor incertidumbre en su modelo.

El conocimiento limitado que se tiene sobre la dinámica de los procesos ambientales hace conveniente la evaluación de la incertidumbre en los modelos que se utilizan para conocerlos (Klepper 1997, Hwang *et al.* 1998). Idealmente, debe ser posible la

TESIS CON
FALLA DE ORIGEN

identificación, descripción y formalización de la incertidumbre que afecta los modelos espaciales utilizados (Hwang *et al.* 1998). Así entonces, se podría producir no sólo el mapa con las predicciones del modelo, sino también un mapa con la incertidumbre asociada a ellas. Con el análisis de incertidumbre en los modelos se estiman los límites de confianza (o incertidumbre) en los valores predichos (Monte *et al.* 1996, Hakánson 1996).

3.2. INTERPOLACIÓN

Cuando no existe la información suficiente para la construcción de modelos estadísticos se recurre a métodos de interpolación para generarla. Sin embargo, el error de interpolación de las variables utilizadas en los modelos espaciales es tomado pocas veces en cuenta. El error de interpolación puede no ser homogéneo en el espacio pues depende del número y cercanía de los datos de muestreo. La evaluación de este error permite caracterizar la incertidumbre en los resultados cuando la densidad de los datos de campo es baja y existe variación entre datos contiguos para la interpolación (Phillips *et al.* 1996, Hwang *et al.* 1998).

Actualmente, con la geostatística es posible estimar el error en los modelos espaciales de manera puntual. A diferencia de los métodos estadísticos clásicos, la geostatística supone que cada medición depende de su ubicación espacial. Para ello, utiliza métodos clásicos de regresión modificados *ad hoc* con el fin de describir la complejidad de los fenómenos naturales en una región determinada. El variograma y el semivariograma son los estimadores utilizados en la interpolación espacial por el método de *kriging* de la variable con base en la varianza promedio entre pares de observaciones en función de la distancia (h) que las separa (Isaaks *et al.* 1989, Funk *et al.* 1999):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{(i,j)h_{ij}=h} (v_i - v_j)^2$$

Donde v_i es el valor en el punto i , v_j es el valor en el punto j y $N(h)$ es el número de pares ij separados por la distancia h .

En la interpolación por *kriging* se estima, a través del semivariograma, el promedio de la variable y el error asociado a éste utilizando los valores observados y la distancia entre ellos, la dirección de la variación y redundancia de los datos. De esta manera, el error

consiste en la dispersión de las mediciones (desviación estándar) alrededor del valor estimado con base en la variación observada en la muestra. Finalmente, éste método de interpolación supone que la variable de interés esta regionalizada, es decir, resulta de procesos que operan a escala regional (Isaaks *et al.* 1989, Fortin *et al.* 1999, Moeur *et al.* 1999). Por lo tanto, ésta estimación del error puede ser una herramienta eficiente en el cálculo de la incertidumbre de los modelos que la utilizan

3.3. ENTROPÍA Y TEORÍA DE LA INFORMACIÓN

El concepto de entropía establece que la información disponible en un sistema no es única y estable, sino una amplia gama de posibilidades que pueden generar una respuesta en particular. Así, la incertidumbre es la ignorancia *a priori* sobre la respuesta generada por el sistema en un momento dado. Un sistema con alta incertidumbre tiene diferentes posibilidades de respuesta, alto contenido de información y, por lo tanto, alta entropía. Cuando todas las respuestas tienen la misma posibilidad se dice que el sistema tiene su contenido máximo de información, máxima entropía (Raisbeck 1963, George 1965, Wicken 1987, Bracken *et al.* 1990, Singh 2000).

Con el desarrollo de la teoría de probabilidad en 1949 aparece el planteamiento de Shannon sobre la Teoría de la información que actualmente es utilizada en las áreas cuyo objeto de estudio presente comportamiento estocástico. De manera general, la Teoría de la información se refiere al flujo de información entre el emisor y el receptor (Bracken *et al.* 1990). Originalmente, la Teoría de la información propone que la información en cualquier sistema con comportamiento aleatorio es su capacidad en generar diferentes respuestas. (George 1965). Sin embargo, los mecanismos que se utilizan para conocer estos sistemas suelen ser imprecisos y es necesario evaluar la capacidad que tienen de transmitir la información que el sistema genera (transinformación, Reza 1961).

En general, para Shannon y Weaver el contenido de información o entropía en un sistema se mide con el índice (Reza 1961, Wicken 1987, Singh 2000):

$$H = \sum_{i=1}^n p_i \ln \frac{1}{p_i}$$

Donde p_i es la probabilidad de transmisión de la respuesta i y H es la entropía.

En efecto, si se parte de un conjunto de respuestas que pueden definir la cualidad de interés en el sistema, la cantidad de información del sistema se obtiene a partir de la probabilidad de expresión (p_i) que tiene cada una de esas respuestas. La unidad que se utiliza para medir la cantidad de la información en un sistema aún no ha sido estandarizada pero se determina en función de la base de los logaritmos que se utiliza en la fórmula para calcularla. Así, cuando se utiliza logaritmo base 2 la unidad es el *bit* (por la contracción del término en inglés '*binary digit*'); cuando es logaritmo base 10 se habla de *decit* ('*bel*' o '*decimal digit*') y cuando se trata logaritmos naturales (base e) se habla de *nats* (o '*natural bel*', Pielou 1977).

Un mecanismo capaz de transmitir todas las respuestas generadas por el sistema es eficiente. Por otro lado, la redundancia en la información está dada por la cantidad de estados diferentes en el sistema. Cuando hay menor cantidad de estados en el sistema hay mayor redundancia, pues la información generada es repetitiva, es decir, precisa (Reza 1961, George 1965).

En este trabajo, el sistema a conocer es la distribución de las especies en una región a través de un modelo construido *ad hoc*. La simulación estocástica y el índice de Shannon-Weaver se utilizan, entonces, para evaluar la incertidumbre en las predicciones del modelo de distribución de especies, como consecuencia del error de estimación en las variables ambientales que éste emplea. El propósito final es establecer los niveles de incertidumbre en la predicción del modelo, que permita elegir entre zonas con mayor o menor confiabilidad para la protección de las especies cuya distribución se modela.

4. MÉTODO

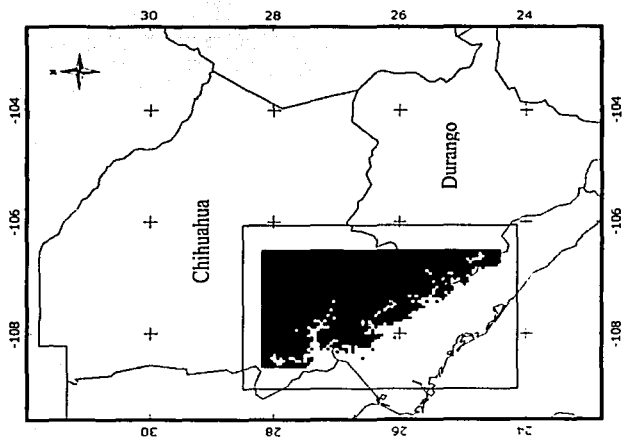
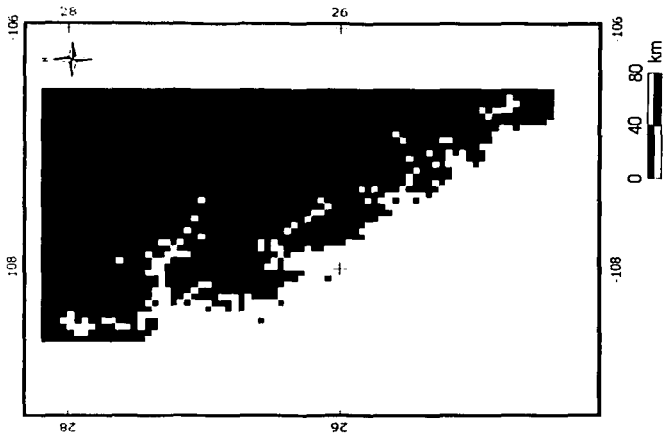
En este trabajo se evalúa la incertidumbre en las predicciones de un modelo de distribución de especies a partir del concepto de entropía. Para ilustrar este método se eligió uno de los grupos de especies de vertebrados terrestres reportados por Solares (2001). En este grupo se encuentran 9 especies de aves (*Amazilia violiceps*, *Amazilia beryllina* y *Cyananthus latirostris*, Familia Trochilidae; *Basileuterus rufifrons* y *Parula pitiayumi*, Familia Parulinae; *Myiodynastes luteiventris* y *Sayornis nigricans*, Familia Tyrannidae; *Molothru aeneus*, Familia Icteridae y *Vireo hypochryseus*, Familia Vireonidae) que ocupan generalmente zonas de matorrales, los límites de éstas con los bosques, cultivos y zonas abiertas (Howell *et al.* 1995).

El área de estudio se definió con base en la ubicación de los 20 registros de colecta de las especies utilizadas. Esta área ocupa 5 194 000 ha al Suroeste del estado de Chihuahua. Los límites Norte y Este del área de estudio fueron marcados con base en la cercanía de las estaciones climatológicas. El límite al Suroeste, que marca una diagonal en la región, es la cota de 600 msnm de la Sierra Madre Occidental (Figura. 1).

Los tipos de vegetación más abundantes son el Bosque de pino-encino (en las partes altas) y la Selva mediana y Selva baja caducifolia (en las laderas del lado oeste de la Sierra) pues ocupan en conjunto el 65 % de su extensión (40 % y 25 %, respectivamente Tabla 1). El resto del área está cubierto por Bosque de pino, Bosque de encino y Pastizal en un 15 %, 14 % y 5 % respectivamente. El 2 % restante está ocupado con Agricultura y Chaparral (Figura 2a).

En términos generales, el área de estudio está en las una de las zonas más altas de la sierra pues el 43 % de su extensión presenta altitudes mayores a los 2 000 msnm y un 40 % más esta entre los 1 000 y 2 000 msnm (Figura 2b). Presenta un clima templado y seco pues la temperatura media esta entre los 12 °C y 16 °C (en el 50 % del área) mientras que sólo el 11 % presenta la mayor precipitación total anual durante los meses de junio a septiembre (1 000 a 1 200 mm.).

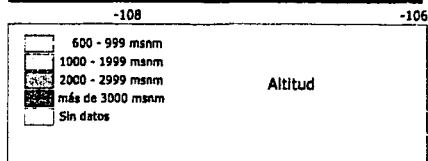
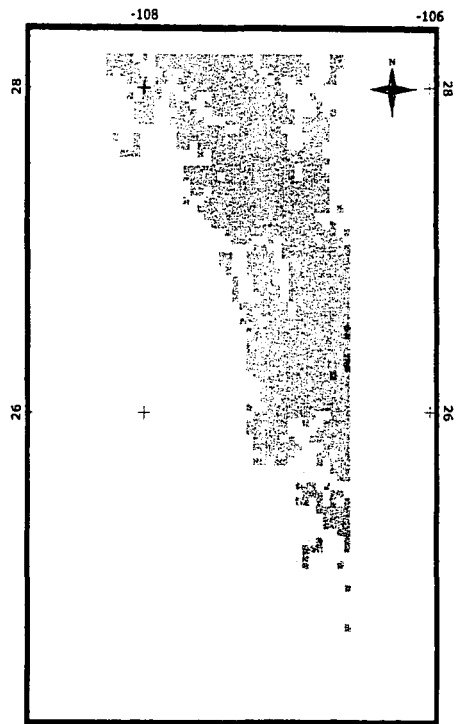
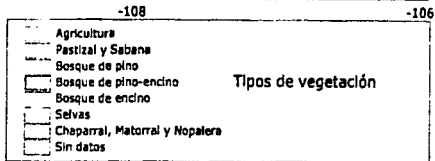
Figura 1. Ubicación geográfica del área de estudio.



TESIS CON
FALLA DE ORIGEN

Figura 2. Características del área de estudio (escala 1 000 000, INEGI 1994). a)- Tipos de vegetación y uso del suelo y
b)- Altura sobre el nivel del mar

TESIS CON
FALTA DE ORIGEN



0 20 40
km

Tabla 1. Extensión en hectáreas y porcentaje de cada tipo de vegetación (escala 1:1 000 000, INEGI 1994) en el área de estudio.

Tipo de Vegetación	Extensión	
	(ha)	(%)
Agricultura	55 650	1
Pastizal y Sabana	259 700	5
Bosque de pino	720 800	14
Bosque de pino-encino	2 090 850	40
Bosque de encino	760 550	15
Selvas	1 279 950	25
Chaparral, Matorral y Nopalera	26 500	1
Total	5 194 000	100

4.1. MODELO DE DISTRIBUCIÓN

El modelo de distribución de especies se construyó en con base en los modelos lineales generalizados GLIM (por sus siglas en inglés, Crawley 1993, Cruz 1996, Margules *et al.* 1989, Nicholls 1989) con base en los 17 registros de colecta reportados en los estados de Chihuahua y Durango (León 2000, Navarro 2000) para las 9 especies de aves del grupo. Además, se obtuvieron los valores de las variables ambientales en cada sitio de colecta a través de la sobreposición de éstos con los mapas disponibles en la base de datos. Con este

$$y = e^{(a+vt+bp)}$$

conjunto de datos se construyó el modelo con base en la forma:

donde y : es la frecuencia esperada del conjunto de especies; a , b y c : son los coeficientes de regresión para cada variable según el modelo ajustado; v : es el coeficiente de regresión para cada tipo de vegetación y t y p : son los valores de temperatura y precipitación en cada unidad. Este modelo se calculó por unidad a través de un archivo ejecutable en UNIX.

De las variables utilizadas, el tipo de vegetación y la altitud en cada unidad permanecieron constantes. Por el contrario, los datos de precipitación y temperatura se modificaron con base en las medias estimadas en la interpolación por *kriging*. Además, con base en los resultados de la simulación estocástica se calculó el modelo por pixel para obtener el conjunto de predicciones a partir del cual se calculó la entropía del modelo.

4.2. BASE DE DATOS

La base de datos utilizada incluyó los mapas digitales de altitud y vegetación disponibles el Laboratorio de Análisis Ambientales del Instituto de Ecología, UNAM; los registros de precipitación y temperatura de las estaciones climatológicas y los registros de colecta de las 9 especies de aves agrupadas. Para el manejo de esta información, se utilizó el SIG Geographic Resource Analysis Support System, GRASS versión 4.3 (USA-CERL 1998), instalado en una estación de trabajo Sun Sparc Ultra10, con el sistema operativo Solaris 5.6.

Los mapas de precipitación y temperatura se obtuvieron por interpolación a partir de la consulta, por medio del Extractor Rápido de Información Climatológica (ERIC; IMTA 1996), de los registros de las estaciones climatológicas del Sistema Meteorológico Nacional con mayor cantidad de años registrados. Primero, se identificaron las estaciones climatológicas en el área de estudio a fin de obtener los datos disponibles de precipitación y temperatura en cada una. Con éstos registros, se calculó la temperatura promedio anual y la precipitación total anual promedio en cada una de las estaciones en la región. El manejo de la base de datos climatológicos se realizó a través de programas ejecutables en UNIX generados especialmente para cada caso.

Con el propósito de representar el comportamiento espacial de la temperatura y la precipitación, se ajustaron las distribuciones de frecuencias correspondientes al conjunto de datos definido para cada variable. Así, se ajustó una distribución Normal para los datos de temperatura y una LogNormal para los datos de precipitación. Además, se calculó el coeficiente de correlación entre ambas variables, que después se incorporó en la simulación estocástica.

Por otro lado, los registros de las especies de aves se obtuvieron de la base de datos proporcionada por León y Navarro (2000) para los vertebrados terrestres en Chihuahua y Durango con base en los reportes de los proyectos que esta institución ha apoyado en la región. El archivo de registros se depuró de coordenadas apócrifas y se exportó al SIG para su manejo.

4.3. INTERPOLACIÓN

Los mapas de precipitación y temperatura se obtuvieron a través de la interpolación de los datos de las estaciones climatológicas en la región. Esta interpolación permite estimar los valores de estas variables de manera puntual, cuando, como en este caso, no se cuenta con la información de campo completa. En primer lugar, se obtuvieron los promedios anuales de temperatura media y precipitación total de cada una de las estaciones seleccionadas. Enseguida, se realizaron las interpolaciones correspondientes a fin de caracterizar la variación espacial en la región a través del análisis geoestadístico de los datos. Para tal efecto, se utilizó el programa *GS+ Geostatistics for the Environmental Sciences*.

version 5.0.3 beta para el cálculo del semivariograma e interpolación por *krigging* de las variables en el resto de la región. Así, se obtuvo la media y la desviación estándar por pixel para cada variable.

Por otro lado, con base en la variación estimada por la interpolación para precipitación y temperatura, se determinó el tamaño mínimo del pixel que permitiera el manejo de la información evitando los datos redundantes, pero conservando la variación en la región. Así, la región se dividió en 1 960 unidades (pixeles) de 2 650 ha cada una. Por medio del SIG, los mapas de altitud y vegetación se hicieron compatibles a éste tamaño de unidad. En ambos casos la categoría asignada por unidad fue la moda del conjunto de pixeles agrupados.

4.4. SIMULACIÓN ESTOCÁSTICA

La simulación estocástica se realizó por unidad con base en la rutina PRISM (Gardner *et al.* 1983). Los archivos de entrada se construyeron con la media y la desviación estándar de precipitación y temperatura en cada unidad, producto de la interpolación, los mínimos y máximos de cada variable por pixel y el grado de correlación de las variables. Con esto, se definió el dominio de cada distribución para el ajuste de la función bivariada en la que PRISM hace los muestreos correspondientes. PRISM esta programado en FORTRAN 77 y se compiló en la estación de trabajo. En términos generales, el proceso de simulación en PRISM tiene tres fases:

- Cálculo de las funciones de probabilidad para cada variable. PRISM calculó la función de distribución de probabilidad a partir del mínimo y máximo observado, media y desviación estándar determinados anteriormente y señalados en el archivo de entrada.
- Cálculo de la distribución bivariada. Para esto, PRISM utilizó el coeficiente de correlación señalado en el archivo de entrada y con base en él determina la distribución bivariada entre las variables.
- Muestreo aleatorio. A partir de la distribución de frecuencias, PRISM calculó los intervalos de clase y con éstos generó los números al azar dentro del dominio de la función bivariada hasta completar las simulaciones requeridas. La principal ventaja del Hipercubo Latino sobre la simulación Monte Carlo es que asegura que los números

generados son el resultado del muestreo de la distribución completa; esto es, se hace el muestreo al menos una vez en cada intervalo de clase que define la función correspondiente.

El número de simulaciones se determinó por medio de la prueba T para diferencias de medias entre los datos observados y los simulados en toda la región (50 y 100 simulaciones). Como resultado de este proceso se obtuvieron 163 valores aleatorios para precipitación y temperatura por archivo de entrada, esto es, por unidad (pixel) en función del máximo número de individuos predicho por el modelo.

4.5. CALCULO DE ENTROPÍA

Con los 163 mapas de distribución simulados se calculó el histograma de frecuencias por pixel. El histograma se definió con base en el número de repeticiones de cada predicción (0-162 individuos). A continuación, se calcularon las probabilidades de cada una con base en:

$$p_i = \frac{n_i}{163} \quad i=0,1,2,3,\dots,162$$

Donde p_i es la probabilidad y n_i es la frecuencia de la predicción.

Una vez calculadas las probabilidades, se determinó la entropía de la predicción en cada pixel utilizando el índice de Shannon-Weaver. El valor de entropía por pixel se llevó al SIG para obtener el mapa de incertidumbre asociada en la predicción del modelo de distribución de especies.

La eficiencia del modelo en transmitir la información generada a partir de las simulaciones se evaluó con base en el cálculo de la redundancia relativa y la entropía máxima en cada pixel. Así, la entropía máxima es:

$$H_{max} = \ln n$$

Donde n es el número total de respuestas posibles en el modelo, esto es, el número máximo de individuos predichos en el área. La redundancia relativa está definida por:

$$\frac{\ln n - H_x}{H_{max}} \quad x=0,1,2,3,\dots,1960$$

y la eficiencia:

$$\frac{H_x}{H_{max}} \quad x=0,1,2,3,\dots,1960$$

En ambos casos, H_x es la entropía en el pixel x y n el número total de respuestas posibles en el modelo.

Para finalizar el análisis de los resultados se compararon los resultados de entropía, redundancia relativa y eficiencia con el coeficiente de variación de las predicciones del modelo y de las simulaciones de precipitación y temperatura por pixel.

5. RESULTADOS

5.1. MODELO DE DISTRIBUCIÓN

Los 17 sitios de colecta se ubican sólo en Selva, Bosque de pino y Bosque de pino-encino (9, 7 y 4, respectivamente, Figura 3a). Estos sitios se ubican principalmente en zonas bajas (<2 000 msnm) con precipitaciones entre 600 y 700 mm, sólo un sitio se localiza en áreas con 900 mm; y temperatura entre 8 °C y 14 °C. El sitio de colecta con mayor número de registros se encuentra por encima de los 2 000 msnm.

El modelo GLIM final para predecir la distribución y abundancia del conjunto de especies utilizado es:

$$y = e^{(-1.681 - 0.00515p + 0.3882t + 0.001535a + v)}$$

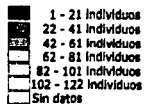
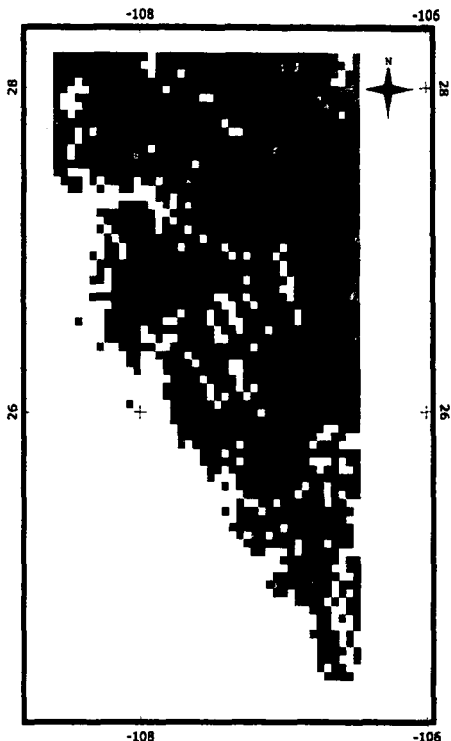
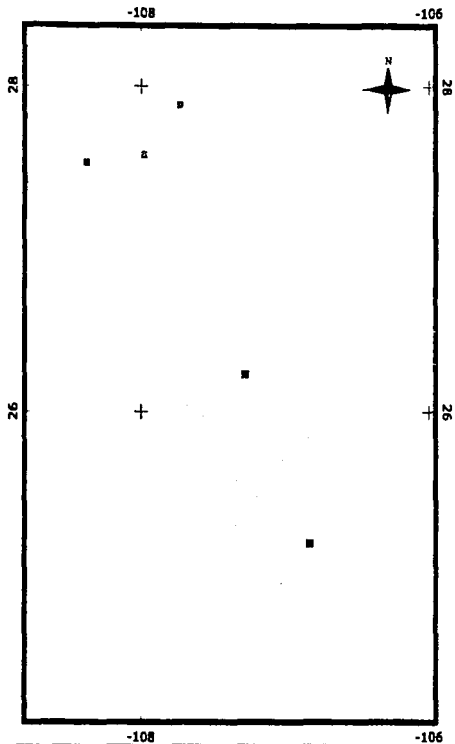
Donde p , t , a y v son los valores de precipitación, temperatura, altitud y vegetación. La prueba de bondad de ajuste (χ^2) demostró que el cambio de la devianza entre el modelo nulo y el modelo propuesto es significativo (0.03945, $p \leq 0.05$). Además, el modelo explica en un 100 % el cambio en la devianza.

Este modelo se realizó por pixel con los datos de las variables en cada uno. Los coeficientes de regresión asociados a cada una de las variables corresponden a la influencia que tiene cada una en el resultado final. En contraste, la vegetación tiene un coeficiente por cada tipo de cobertura, -1.256 para el Bosque de pino-encino y -0.3695 para las Selvas. El resultado es el mapa de distribución del conjunto de especies de aves (Figura 3b).

En el mapa de distribución ajustado, el máximo predicho es 122 individuos (Figura 3b). Sólo en el 16 % (752 760 ha) del área se predicen más de 20 individuos, de éstas en tan sólo 26 560 ha se predicen más de 50 individuos y en únicamente 15 920 ha más de 100 individuos. Estas áreas se encuentran en Bosques de encino y Pastizales entre 2 000 msnm y 3 000 msnm, con precipitación entre 300 mm y 400 mm y con temperaturas de 10 °C y 11 °C. En el resto del área las predicciones no sobrepasan los 40 individuos por pixel. Siguen siendo las zonas con mayor abundancia aquellas cubiertas de bosques y pastizales, aunque en menores altitudes.

Figura 3. Mapa con el número de registros de colectas para las 9 especies de aves (a) y el modelo ajustado en GLIM (b).

TESIS CON
FALLA DE ORIGEN



**TESIS CON
FALLA DE ORIGEN****5.2. ANÁLISIS CLIMATOLÓGICO**

En la región de estudio se localizan 59 estaciones climatológicas, 48 en Chihuahua y 11 en Durango. Los registros utilizados en este análisis abarcan de 1960 a 1990 pues es el periodo de tiempo que contiene datos para un número mayor de estaciones. Aunque ninguna estación cubre los 30 años de registro, se encontraron datos de temperatura en 33 estaciones y de precipitación en 53 estaciones.

El 36 % de las estaciones que registraron la temperatura tienen datos de 2 a 5 años. En el caso de la precipitación, sólo el 17 % de las estaciones registró ese mismo número de años. Pese a que sólo una estación registró la precipitación durante 29 años y la temperatura se registró en 9 estaciones durante un máximo de 16 años, en general, existe mayor cantidad de registros de precipitación que de temperatura (14.7 y 9.3 años en promedio, respectivamente).

La temperatura media anual observada presentó un mínimo de 2.1 °C y 22.8 °C como máximo. La distribución de frecuencias de tipo Normal se ajustó con base en 11.6 °C de media y 6 °C de desviación estándar. La prueba de ajuste K-S fue de 0.098 en comparación con el valor de tablas $\alpha_{0.1}=0.1794$ y $\alpha_{0.05}=0.1542$.

El rango de precipitación total anual registrada en las estaciones fue entre 141 mm y 1 283 mm. La media observada de 709.5 mm y la desviación estándar de 246.6 mm describen una distribución de frecuencias del tipo LogNormal. El valor de Kolmogorov-Sminorv para bondad de ajuste fue de 0.090 en comparación con los valores de tablas $\alpha_{0.1}=0.1416$ y $\alpha_{0.05}=0.1217$.

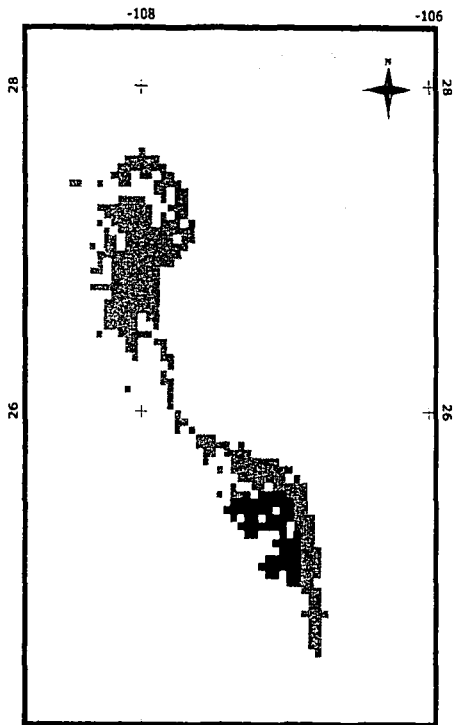
5.3. INTERPOLACIÓN

Los valores de temperatura estimados en la interpolación por krigging varían entre 4 °C y 19 °C (Figura 4a). El 70 % del área (3 379 810 ha) mostró temperaturas menores a los 12 °C en zonas de más de 1 000 msnm. Del resto del área (1 446 970 ha), poco menos del 45 % se encuentra a menos de 1 000 msnm.

Por su parte, la estimación de la precipitación presentó valores entre 310 mm y 990 mm (Figura 4b). Del área total, 1 752 800 ha (33 %) presentaron valores promedio de

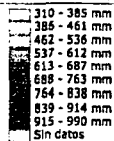
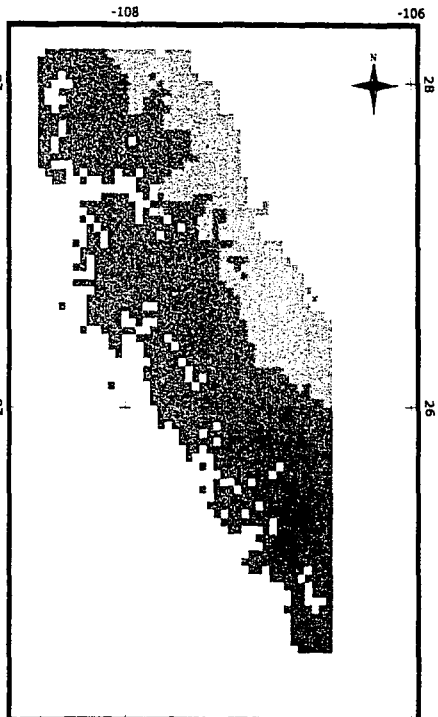
Figura 4. Mapas de temperatura (a) y precipitación (b) interpolados con *kriging*.

TESIS CON
FALLA DE ORIGEN



Interpolación de temperatura

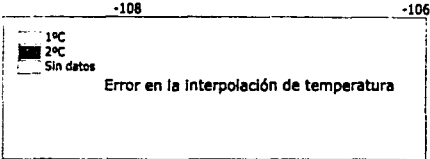
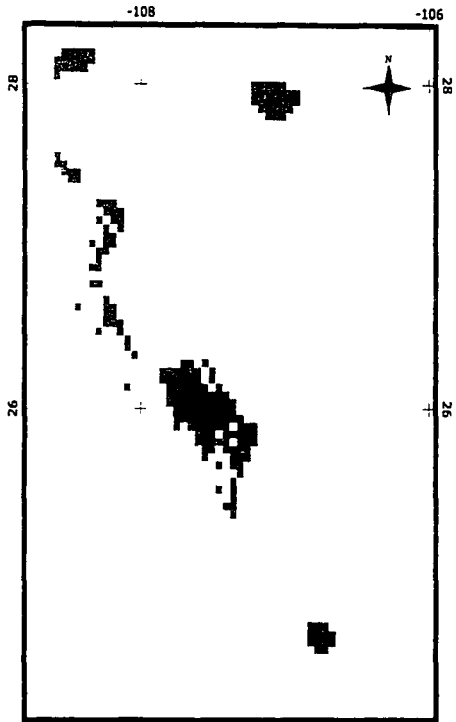
0 20 40
km



Interpolación de precipitación

Figura 5. Errores de *kriging* en la interpolación de la temperatura (a) y la precipitación (b).

TESIS CON
FALTA DE ORIGEN



0 20 40 km

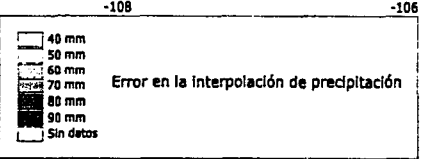


Tabla 2. Prueba de T-Student de diferencia entre medias para los datos de precipitación y temperatura observados y simulados (con g.l. de 150 y 100 para temperatura y de 175 y 152 para precipitación).

Estadístico	Observados	Simulados		
		Temperatura		
η	33	50	100	500
media	11.62	11.79	11.79	11.78
varianza	36.96	0.25	0.23	0.23
desviación estándar	6.08	0.49	0.48	0.47
prueba de T		-0.122588	-0.122728	-0.124027
		Precipitación		
η	53	50	100	500
media	709.48	709.56	709.56	709.56
varianza	60814.28	0.75	0.75	0.75
desviación estándar	246.60	0.86	0.87	0.87
prueba de T		0.001762	0.001745	0.001797

precipitación menores a los 600 mm en zonas de más de 1 000 msnm. El resto de la región (3 073 980 ha) son zonas con precipitaciones mayores a los 600 mm.

En sólo el 5 % del área (200 460 ha) el error de estimación de la temperatura fue de 2 °C. Estas áreas se encuentran en zonas de más de 1 000 msnm (Figura 5a). En el caso de la precipitación, el error de estimación presentó valores entre 50 mm y 80 mm. En el 85 % de la región se obtuvieron errores mayores a los 60 mm, de ésta área sólo 627 800 ha corresponden a zonas con menos de 1 000 msnm (Figura 5b).

5.4. SIMULACIÓN ESTOCÁSTICA

Con la simulación estocástica se obtuvieron 163 valores por pixel. El número de simulaciones fue establecido con base en el número máximo de individuos predicho por el modelo de distribución de especies utilizado por un lado, y por otro, con base en las pruebas de T-Student para diferencias de medias entre los valores de las estaciones climatológicas y los simulados (Tabla 2). Este resultado muestra que la media de los datos observados y la media de los datos simulados no son significativamente diferentes, además, es evidencia del desempeño de la rutina PRISM en cuanto a la simulación a partir de los datos observados.

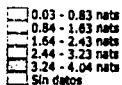
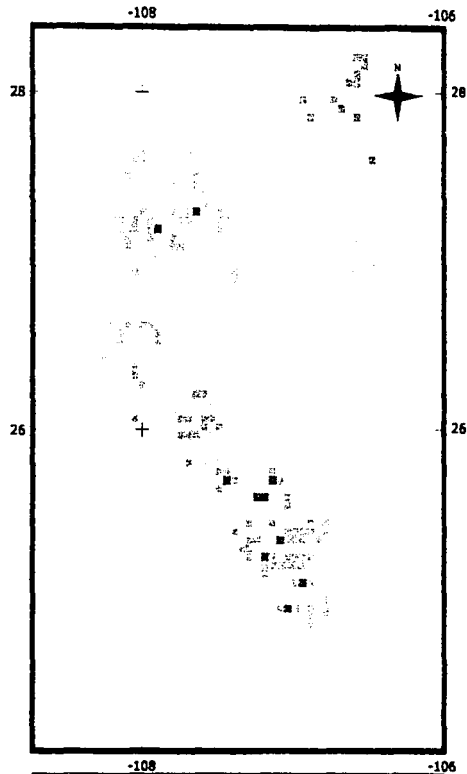
En términos generales, en los 163 mapas simulados el mayor número de individuos predicho es de 162. Únicamente en el 1 % del área se predicen valores mayores a los 50 individuos. Sólo en 6 unidades del área, que equivale a 15 900 ha, se predicen valores mayores a los 80 individuos. Estas zonas están cubiertas principalmente de Bosque de encino, aunque existen áreas de Bosque de pino, Pastizales y Chaparrales; a más de 2 000 msnm y con temperaturas entre los 9 °C y 16 °C y precipitaciones entre 300 mm y 600 mm.

5.5. ENTROPÍA

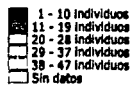
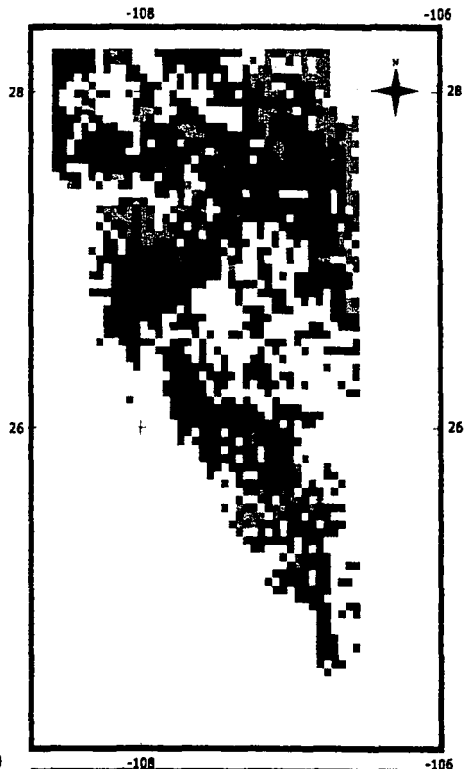
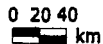
La entropía depende de la probabilidad de predicción que tienen los estados del sistema. Cuando el número de predicciones posibles aumenta, la probabilidad de presencia de cada una de ellas disminuye. El valor de $H_{\text{máx}}$ fue de 5.09 nats puesto que el modelo predice desde 0 individuos y, por lo tanto, se consideraron 163 respuestas por pixel. Así pues, la entropía del modelo de distribución varía entre 0 nats y 3.89 nats (Figura 6a y Gráfica 1). Hay que destacar que la entropía del modelo aumenta en función, primero, del número de

Figura 6. Mapa de entropía en los resultados simulados para el modelo de distribución de especies (a) y área seleccionada (b) con base en la eficiencia del modelo según la teoría de la información.

TESIS CON
FALLA DE ORIGEN

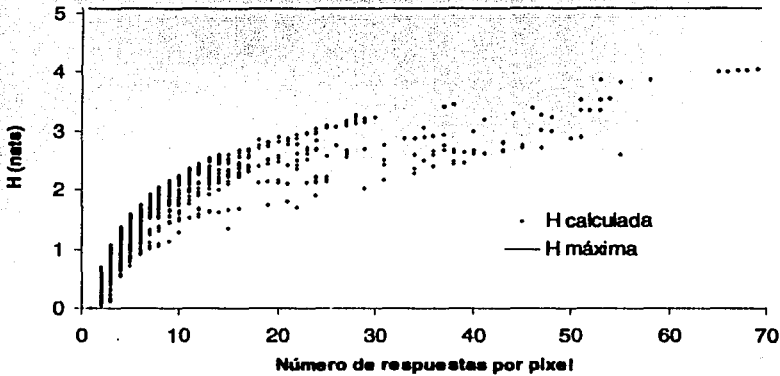


Entropía en la predicción del modelo de distribución

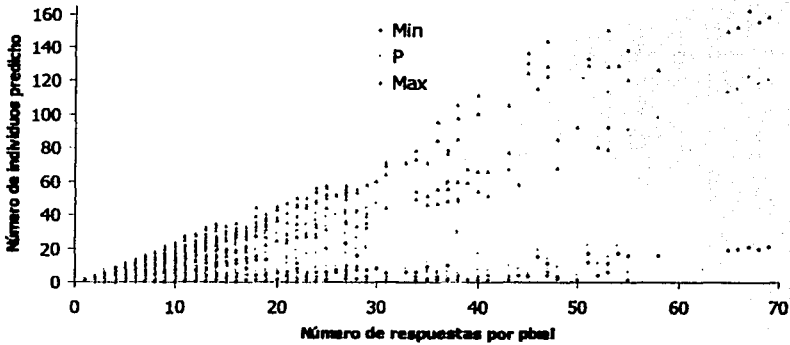


Predicción dentro de los límites de entropía

Gráfica 1. Valores de la entropía (H) calculada por pixel con base en el número de respuestas diferentes en cada uno; la línea superior indica la H máxima del sistema en 5.09 nats.



Gráfica 2. Predicción del número de individuos con los valores interpolados promedio (P) de precipitación y temperatura. Además, se muestra la predicción mínima (Min) y máxima simulada (Max) por pixel.



respuestas simuladas y, segundo, de la frecuencia de cada respuesta por pixel. Sólo en 160 000 ha, aproximadamente, se presentan valores de entropía de más del 50 %.

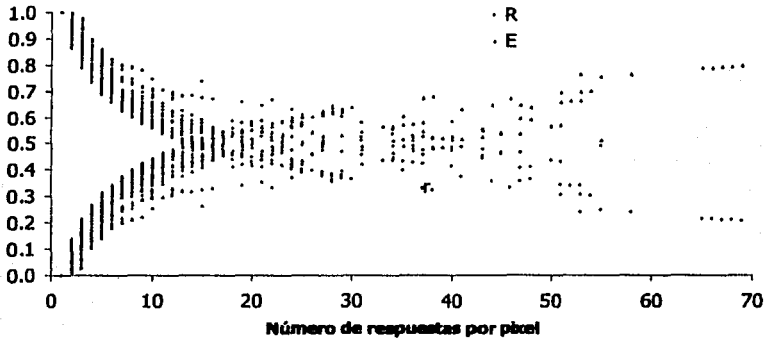
Las áreas con mayor entropía son aquellas en donde el modelo de distribución, con los promedios interpolados de precipitación y temperatura, predice un mayor número de individuos (Figura 3b y Figura 6a). Al término de la simulación, estas áreas tienen un alto grado de incertidumbre (>70 %) a causa del número de respuestas por pixel (Gráfica 2), que corresponden, en su mayoría, a zonas cubiertas con Bosque de encino, Bosque de pino y Selvas (43 %, 25 % y 22 %, respectivamente); a más de 1 000 msnm; con precipitaciones entre los 300 mm y 50 mm y temperatura entre los 9 °C y 12 °C. Además, parece relevante que en ellas el error de interpolación de la precipitación es entre 60 mm y 80 mm.

A partir del cálculo de la entropía en los resultados por pixel es posible estimar la redundancia relativa de la información producida por el sistema y la eficiencia del mecanismo que se utiliza para su transmisión. Así, los valores de redundancia relativa disminuyen al aumentar el número de predicciones en el pixel. Por el contrario, la eficiencia del modelo aumenta en tanto aumenta el número de predicciones por pixel (Gráfica 3).

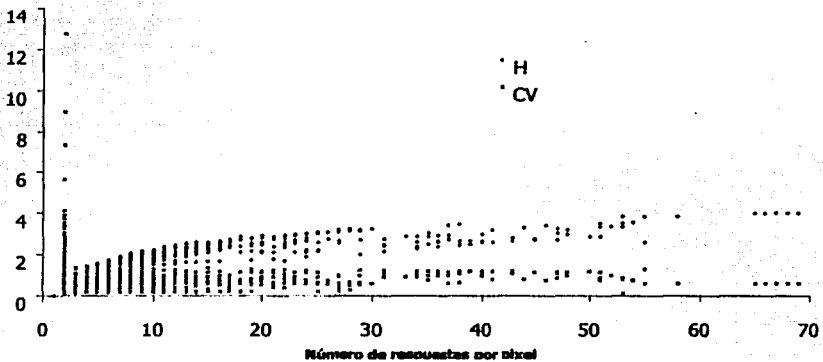
De la misma manera, la redundancia relativa de los resultados del modelo presenta una relación inversa con la entropía. Al aumentar la entropía disminuye la redundancia en las predicciones. En el cálculo de la eficiencia del modelo, al aumentar la entropía, aumenta también la eficiencia del modelo. En los pixeles en donde la entropía es mayor, la eficiencia del modelo también lo es. En cambio, los pixeles con mayor entropía tienen mayor información y por lo tanto menor redundancia en los resultados producidos.

Por otro lado, al comparar los resultados de la entropía y el coeficiente de variación propuesto por Phillips (1996), como medidas de incertidumbre en las predicciones por pixel, se encontró que a diferencia de la entropía, la gráfica del coeficiente de variación tiene un comportamiento más errático (Gráfica 4). Las variaciones más evidentes se hallaron principalmente en los pixeles con menor número de predicciones. Esta tendencia está dada por los valores de las predicciones. En los pixeles donde la mayoría de las simulaciones predicen 0 individuos —y en sólo algunos casos se predicen valores mayores— el conjunto de

Gráfica 3. Valores de redundancia relativa (R) de las predicciones del modelo (información en el sistema) y de eficiencia (E) del modelo en transmitir esa información.



Gráfica 4. Valores de entropía (H) y coeficiente de variación (CV) de los resultados simulados por el modelo de distribución de especies.



predicciones es menos homogéneo y, por lo tanto, el coeficiente es mayor. Contrariamente, cuando los valores predichos son diferentes de 0, la desviación estándar es menor y en consecuencia el coeficiente de variación es menor.

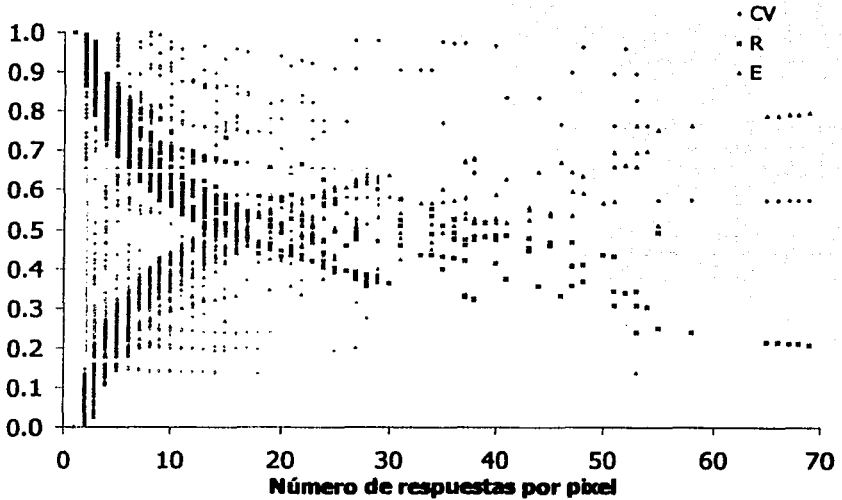
Sin embargo, en la gráfica de entropía es fácil distinguir la tendencia en función del incremento en x . En tanto que la entropía se encuentra en sus niveles más bajos, incluso en algunos casos con entropía nula, el coeficiente de variación presenta valores por encima de la unidad. La diferencia observada se debe a que la entropía es una medida de la ignorancia *a priori* en los resultados producidos por el modelo, que depende del número de respuestas posibles por sitios y de la probabilidad de ocurrencia de cada una. Mientras que el coeficiente de variación es una medida de la homogeneidad entre los valores predichos en cada pixel.

Aunque ambos métodos miden diferentes características en los resultados, ésta comparación permitió la selección de las áreas con mayor cantidad predicha de individuos y menor entropía. Lo anterior se consiguió al graficar la eficiencia del modelo, la redundancia relativa y el coeficiente de variación de las predicciones contra el número de respuestas por pixel y el número de individuos predicho, respectivamente. Los puntos en donde las líneas se cruzan mostraron el intervalo de eficiencia del modelo en que la predicción del número de individuos es confiable (Gráfica 5 y Figura 6b). Así entonces, el intervalo de eficiencia definido fue $0.14 \leq E \leq 0.65$ (Figura 6b).

La zona definida por el intervalo de eficiencia ocupa 3 858 400 ha, esto es, el 95 % del área de estudio con valores de entropía entre 0.69 nats y 3.29 nats. De esta área, el 69 % tiene precipitaciones entre los 500 mm y 800 mm; el 92 %, temperaturas entre 5 °C y 15 °C y el 79 %, altitudes entre 1 000 msnm y 3 000 msnm. Los tipos de vegetación predominantes son Bosque de pino-encino (32 %), Selvas (27 %), Bosque de pino (17 %) y Bosque de encino (15 %).

Finalmente, en sólo 219 950 ha se predicen más de 20 individuos por pixel. Estas áreas ocupan Bosques de encino, Pastizales y Bosques de pino; en altitudes entre 1 000 msnm y 2 000 msnm, precipitaciones entre 300 mm y 500 mm y temperatura de 12 °C, principalmente. Con base en la selección señalada, el máximo número de individuos predicho en esta área fue de 47 individuos, que se presentó en un único pixel.

Gráfica 5. Valores del coeficiente de variación (CV), eficiencia del modelo (E) y redundancia relativa en las predicciones del modelo por pixel. El recuadro muestra el área seleccionada en función de la exactitud y precisión de las predicciones.



g

TESIS CON
FALLA DE ORIGEN

6. DISCUSIÓN

El conocimiento sobre la distribución de las especies es un requisito importante en las diferentes políticas de conservación. Sin embargo, descifrar las condiciones que rigen la distribución de las especies en el espacio no es sencillo, de ahí la importancia de la inferencia estadística como herramienta en la toma de decisiones. La relación estadística entre las colectas en campo y las variables ambientales que pueden influir en la presencia de cualquier especie hace posible la obtención de resultados en poco tiempo.

Los modelos estadísticos utilizados en la predicción de la distribución espacial de las especies están sujetos a varios tipos de errores que generan incertidumbre en los resultados. Esta incertidumbre puede existir por: *i*) la falta de precisión en la medición de las variables, *ii*) la alteración de los valores durante el proceso de construcción del modelo, *iii*) la estimación de los parámetros y la calibración del modelo o *iv*) por la propia variación natural de los fenómenos medidos (Gardner *et al.* 1990, Haefner 1996, Davis *et al.* 1997). Las tres primeras causas son producto del error humano y pueden disminuirse o eliminarse con sólo tomar precauciones durante la captura de la información y durante la construcción, calibración y ejecución del modelo. Sin embargo, las consecuencias de la variación natural en los fenómenos medidos son pocas veces tomadas en cuenta.

La estimación de la variación en espacio y tiempo es difícil dado que la información con la que se cuenta es escasa y de mala calidad. Por ejemplo, la información climatológica en las estaciones del Servicio Meteorológico Nacional reporta actividad desde 1940, aunque en realidad los años en que se tienen datos son relativamente pocos. Además, el trabajo de interpolación se dificulta cuando se trabaja en grandes regiones y, en consecuencia, con un número mayor de estaciones, pues los datos disponibles generalmente no son comparables. A pesar de todo, la región que se utilizó en este trabajo fue pequeña y con pocas estaciones, lo que facilitó la comparación entre la información disponible que, además, tuvo una alta correlación entre estaciones y que permitió mejores resultados en la simulación estocástica.

Las variables ambientales que determinan el hábitat de las especies cambian en espacio y tiempo. Sin duda, ésta variación afecta los métodos utilizados para conocer la dinámica de los sistemas naturales. Cuando es posible estimar el comportamiento de las

variables de interés y su incertidumbre, es posible estimar el efecto de esta incertidumbre en los modelos que las utilizan. El punto de partida es determinar la variación de estos elementos en el espacio para poder estimar sus efectos en las predicciones del modelo que las ocupa.

El uso de la geostatística en la estimación del comportamiento de las variables ambientales que utiliza el modelo, permitió determinar la incertidumbre con que éstas influyen en las predicciones. El modelo utilizado como ejemplo de distribución de especies, responde a la variación de la temperatura, precipitación, tipo de vegetación y altitud, respectivamente. Esto hizo posible estimar la incertidumbre en las predicciones del modelo con base en la simulación de sus resultados a partir de la variación de la temperatura y precipitación.

En este ejemplo, el modelo construido permite hacer una clara distinción entre las zonas con más o menos individuos aunque, a partir de la estimación de la incertidumbre, las zonas en donde se predice el mayor número de individuos fueron también las zonas con mayor incertidumbre (Figura 5b y Figura 6a). La entropía del modelo de distribución de especies proporcionó una medida de la incertidumbre en sus resultados. Sin duda, las zonas con nula o menor entropía son las más confiables, sin embargo, corresponden a las predicciones más bajas. Incluso con la simulación se hace evidente que en estas zonas no es posible que, bajo las condiciones de éste trabajo, se registren las especies utilizadas como ejemplo.

En el contexto de la Teoría de la información, el desempeño del modelo se evalúa a través de la redundancia relativa de la información producida por el sistema y, en función de ésta, la eficiencia de su transmisión a través del modelo. La información producida por un sistema puede medirse en función de cualquier variable que tenga influencia sobre el sistema y sobre todo si presenta un comportamiento aleatorio. En este caso, la información es el número de individuos predicho cuyo factor aleatorio es la simulación de los valores de precipitación y temperatura que se utilizan para su predicción. A su vez, el modelo de distribución es el canal por el cual se transmite ésta información.

Si existe poca redundancia en el sistema entonces hay mayor variación, mayor información y, por consiguiente, mayor incertidumbre o entropía en las respuestas

variables de interés y su incertidumbre, es posible estimar el efecto de esta incertidumbre en los modelos que las utilizan. El punto de partida es determinar la variación de estos elementos en el espacio para poder estimar sus efectos en las predicciones del modelo que las ocupa.

El uso de la geoestadística en la estimación del comportamiento de las variables ambientales que utiliza el modelo, permitió determinar la incertidumbre con que éstas influyen en las predicciones. El modelo utilizado como ejemplo de distribución de especies, responde a la variación de la temperatura, precipitación, tipo de vegetación y altitud, respectivamente. Esto hizo posible estimar la incertidumbre en las predicciones del modelo con base en la simulación de sus resultados a partir de la variación de la temperatura y precipitación.

En este ejemplo, el modelo construido permite hacer una clara distinción entre las zonas con más o menos individuos aunque, a partir de la estimación de la incertidumbre, las zonas en donde se predice el mayor número de individuos fueron también las zonas con mayor incertidumbre (Figura 5b y Figura 6a). La entropía del modelo de distribución de especies proporcionó una medida de la incertidumbre en sus resultados. Sin duda, las zonas con nula o menor entropía son las más confiables, sin embargo, corresponden a las predicciones más bajas. Incluso con la simulación se hace evidente que en estas zonas no es posible que, bajo las condiciones de éste trabajo, se registren las especies utilizadas como ejemplo.

En el contexto de la Teoría de la información, el desempeño del modelo se evalúa a través de la redundancia relativa de la información producida por el sistema y, en función de ésta, la eficiencia de su transmisión a través del modelo. La información producida por un sistema puede medirse en función de cualquier variable que tenga influencia sobre el sistema y sobre todo si presenta un comportamiento aleatorio. En este caso, la información es el número de individuos predicho cuyo factor aleatorio es la simulación de los valores de precipitación y temperatura que se utilizan para su predicción. A su vez, el modelo de distribución es el canal por el cual se transmite ésta información.

Si existe poca redundancia en el sistema entonces hay mayor variación, mayor información y, por consiguiente, mayor incertidumbre o entropía en las respuestas

producidas por él mismo. Los valores predichos varían en función de las simulaciones de la precipitación y la temperatura, por lo tanto hay mayor información. Al aumentar la entropía en las respuestas del sistema disminuye la redundancia. Por el contrario, el modelo es más eficiente en tanto hay más información por transmitirse en el sistema.

A diferencia del enfoque clásico en biología sobre el uso del índice de Shannon como medida de diversidad, en este trabajo, la entropía se utilizó como medida de la incertidumbre en la predicción del modelo de distribución, con el propósito de establecer áreas donde las predicciones son más confiables. Con estas áreas identificadas es posible entonces optimizar los esfuerzos y recursos en función de la conservación de las especies modeladas. Bajo el enfoque tradicional del índice de Shannon, se preferirían aquellos sitios con mayor entropía (diversidad). Bajo la perspectiva de la incertidumbre, los sitios con mayor entropía (incertidumbre) son aquellos en donde los resultados del modelo de distribución tienen menor confiabilidad.

Sin embargo, el principal problema fue identificar las zonas con las mayores predicciones a pesar de su incertidumbre. Para esto se utilizó la comparación entre los resultados de la eficiencia y la redundancia relativa del modelo, a partir del concepto de entropía, con los resultados del cálculo del coeficiente de variación de las predicciones en cada pixel (Gráfica 5).

La tendencia del coeficiente de variación está dada por los valores de las predicciones. El comportamiento errático de la gráfica del coeficiente de variación es consecuencia de la diferencia entre los valores predichos en cada pixel, esto es, son valores poco homogéneos, principalmente en los pixeles donde se predicen menos de 30 individuos. En estos pixeles, la mayoría de las simulaciones predicen 0 individuos, y sólo en algunas ocasiones se predicen valores mayores, en consecuencia, tanto la desviación estándar como el coeficiente de variación tienen valores altos. Por el contrario, cuando los valores son diferentes de 0 en la mayoría de las predicciones, la desviación estándar y el coeficiente de variación tienen valores menores.

En general, al aumentar la entropía disminuye el coeficiente de variación. Mientras la entropía es una medida del número de resultados diferentes y la proporción de cada uno de

éstos por pixel, el coeficiente de variación es una medida de la diferencia entre los resultados obtenidos.

Por otro lado, la eficiencia y la redundancia relativa del modelo siguen una relación inversa. La eficiencia aumenta en función de la cantidad de información disponible en el sistema para ser transmitida. En este ejemplo, el modelo es eficiente en la transmisión de la información que se está generando en el sistema. Por el contrario, la redundancia señala la proporción de resultados semejantes en el sistema. Así mientras la eficiencia aumenta, la redundancia relativa disminuye.

Consecuentemente, los puntos de intersección entre las gráficas definen los límites de confianza en las predicciones del modelo. El límite inferior está dado por el cruce entre la gráfica de eficiencia y la gráfica del coeficiente de variación. Este punto señala donde la eficiencia del modelo en transmitir la información y la homogeneidad de los valores predichos en cada pixel aumentan. Por su parte, el límite superior es el punto donde se predice el número de individuos más alto sin perder la redundancia que produzca tal cantidad de información que haga a los resultados perder confiabilidad.

En esta comparación, el coeficiente de variación es una medida de la exactitud de los resultados del modelo. La exactitud se refiere a la cercanía entre el dato original y la predicción del modelo (Summers *et al.* 1993, Mowrer 1999). Si se considera al promedio de las predicciones como el dato original, entonces, la desviación estándar muestra qué tanto varían los resultados al cabo de las simulaciones. Así, el cociente de la desviación estándar y el promedio de las predicciones muestran la dispersión en los resultados simulados, es decir, la exactitud de la predicción en el pixel.

Asimismo, la redundancia es una medida de precisión en las predicciones simuladas (George 1965). La precisión se refiere a la variación inherente en las respuestas del modelo (Summers *et al.* 1993, Phillips *et al.* 1996). Cuando se obtiene el mismo resultado con los datos procedentes de las diferentes repeticiones del modelo entonces, el modelo tiene un alto nivel de precisión. Por el contrario, cuando los resultados del modelo a partir de la simulación de la incertidumbre varían ampliamente, significa que el modelo tiene un bajo nivel de precisión. Esto sucede en el modelo analizado en tanto aumenta el número de

respuestas posibles por pixel, hay más información, más entropía y en consecuencia menos precisión.

El intervalo definido por los puntos de intersección en la gráfica presentada (Gráfica 5) es, entonces, el rango de eficiencia del modelo en donde las predicciones ganan exactitud pues tienen menor coeficiente de variación sin perder precisión, pues conservan altos niveles de redundancia. En cambio, en los extremos de las gráficas se encuentran, por un lado, los píxeles que son precisos pero poco exactos pues tienen altos valores de coeficiente de variación y, por otro, los píxeles que son exactos pero poco precisos pues la redundancia tiene sus valores más bajos.

En resumen, con este método se pueden establecer los niveles de confianza en las predicciones de los modelos que se utilizan en la identificación de áreas importantes para la conservación. De esta manera, es posible conocer cómo se propaga la incertidumbre de la información que se usa en la construcción de los modelos espaciales y sus efectos en los resultados. La intención final es ofrecer un apoyo en la toma de decisiones en torno a los lugares susceptibles de protección, pues permite identificar las zonas en donde las predicciones tienen mayor o menor incertidumbre. Por lo tanto, permitirá la optimización de los esfuerzos humanos y económicos en términos de conservación de la biodiversidad en el planeta.

7. CONCLUSIONES

El enfoque de la teoría de la información, a través de la utilización del índice de Shannon-Weaver, permitió evaluar las predicciones del modelo de distribución de especies en función su la variación durante la simulación estocástica.

La entropía es una medida sintética y eficiente del desempeño de los modelos de distribución de especies. Los resultados arrojados por ésta son sencillos de interpretar en el contexto de la teoría de la información. Sin embargo, para que esta función se cumpla se debe verificar la existencia de un valor máximo de entropía, de lo contrario, el índice pierde validez.

La entropía toma en cuenta el número de predicciones del modelo y su probabilidad al cabo de las simulaciones en cada sitio. Por lo tanto, la redundancia relativa es una medida de la precisión en los resultados simulados. En cambio, el coeficiente de variación es una medida de la exactitud de las predicciones, pues evalúa la diferencia en las predicciones del modelo simuladas en cada sitio. Consiguientemente, la combinación de los resultados de ambos índices permitió establecer el intervalo de confianza en las predicciones del modelo utilizado.

Con este método es posible determinar los intervalos de confianza de los modelos utilizados en la predicción de la distribución de las especies en términos de conservación. En consecuencia, se propone como método de apoyo en la toma de decisiones para la ubicación y selección de las áreas importantes para la conservación de la diversidad biológica.

El factor aleatorio en la simulación estocástica de la precipitación y la temperatura en el modelo de distribución de especies, permitió la evaluación de sus predicciones con base en el concepto de entropía. La simulación estocástica hizo posible la utilización del concepto de entropía como medida de incertidumbre pues suministró la variación necesaria para la definición del sistema a evaluar.

El uso de la geoestadística como herramienta en la interpolación de variables ambientales facilitó la simulación del modelo de distribución de especies, pues funcionó como fuente de incertidumbre en los valores de precipitación y temperatura que se utilizaron en el modelo.

8. BIBLIOGRAFÍA

- Agumya, A. y G.J. Hunter. 1999. Assessing "Fitness for use" of geographic information: What risk are we prepared to accept in our decisions?. En K. Lowell y A. Jaton (editores) *Spatial Accuracy Assessment: Land information uncertainty in natural resources*. Ann Arbor Press. Chelsea, Michigan.
- Azuara, I. Y A. Ramírez. 1994. Tecnologías y manejo de información geográfica en bioconservación. *Ciencia y Desarrollo*. 118:58-65.
- Bojórquez-Tapia, L. A., P. Ffolliot y D.P. Guertin. 1990. Multiple -Resource Modeling as a tool for conservation: Its Applicability in Mexico. *Environmental Management*.14: 3 317-324.
- Bojórquez-Tapia. L.A., I. Azuara, E. Ezcurra. O. Flores-Villela. 1995. Identifying conservation priorities in Mexico through Geographic Information systems and modeling, *Ecological Applications*. 5 (1): 215-231.
- Bojórquez-Tapia, L.A., P. Balvanera y A.D. Cuarón. 1994. Biological Inventories and Computer Data Bases: Their Role in Environmental Assessment. *Environmental Management*. 18 (5): 775-785.
- Bracken, I. y C. Wester. 1990. *Information Technology-Geography and Planning. Including Principles of GIS*. Routledge. New York.
- Braithwaite, R.W. 1991. Fauna y habitats surveys as ecological pathfinders. En: C.R. Margules y M.P. Austin (editores).1991 *Nature conservation: Cost effective biological survey and data analysis*. CSIRO. Australia.
- Brillouin, L. 1969. *La información y la incertidumbre en la ciencia*. UNAM. México.
- Brown, D.E. 1982. Madrean Evergreen Woodland. En: Brown D.E. (ed.) *Biotic Communities of the American Southwest-United States and Mexico*. Desert Plants. 4 (1-4): 1-10; 59-65.
- Brown, D.E.y C.H.L Lowe. 1982. Introduction. En: Brown D.E. (ed.) *Biotic Communities of the American Southwest-United States and Mexico*. Desert Plants. 4 (1-4): 10.

- Corsi, F. J. de Leeuw y A. Skidmore. 2000. Modeling Species Distribution with GIS. En: Boitani L y T.K. Fuller (editores) *Research Techniques in Animal Ecology*. Columbia University Press .NY. 442 pp.
- Crawley, M.J. 1993. *Methods in ecology: GLIM for ecologist*. Blackwell scientific publications.massachussetts. U.S.A.
- Cruz, G.M. 1996. *Modelaje espacial de la pérdida de habitat y áreas ricas en especies en Durango y Chihuahua, México*. Tesis de Maestría. Centro de Ecología. UNAM. México.
- Davis, F.K., D.M. Stoms, J.E. Esthes, J. Scean y J. M. Scott. 1990. An information systems approach to the preservation of biological diversity. *Geographical Information Systems*. 4 (1): 55-78.
- Davis, T.J. y C.P. Keller. 1997. Modelling and visualizing multiple spatial uncertainties. *Computers & Geosciences* 23(4):397-408.
- Ezcurra, E., M. Equihua y J. López-Portillo. 1987. The desert vegetation of pinacate, Sonora, México. *Vegetatio* 71: 49-60.
- Flores-Mata, G., L. Jiménez, S. Madrigal X., R. Moncayo y F. Takaki. 1971. Memoria del mapa de tipos de vegetación de la República mexicana. Secretaría de Recursos Hidráulicos. México, D.F. 59 pp.
- Flores-Villela, O. y P. Gerez. 1988. *Conservación en México: Síntesis sobre Vertebrados Terrestres, Vegetación y uso del Suelo*. Instituto Nacional de Investigaciones sobre Recursos Bióticos. Conservación Internacional. México D.F.
- Gardner, R., B. Rödger y U. Bergström. 1983. PRISM. A systematic Method for determining the effect of Parameter Uncertainties on Model Predictions. Technical report Studsvik Energiteknik AB report NW-83-/555. Nykoping, Suecia.
- Gardner, R., V. Dale y R. O'Neill. 1990. Error propagation and uncertainty in process modeling. En: R. Dixon, R. Meldahl, G. Ruark, G. Warren. (editores). 1990. *Process modeling of forest Growth responses to enviromental stress*. Timber Press. Portland Oregon, 208-219 p.
- Gascoigne, J. Y R. Wadsworth. 1999. Mapping misgivings: Monte Carlo Modeling of Uncertainty and the provision of spatial information for international policy. En K.

- Lowell y A. Jaton (editores) *Spatial Accuracy Assessment: Land information uncertainty in natural resources*. Ann Arbor Press. Chelsea, Michigan.
- George, F.H. 1965. *Cybernetics and Biology*. Oliver & Boyd Editions. Great Britain.
- Griiffith, D.A., R.P. Haining y G. Arbia. 1999. Uncertainty and error propagation in map analyses involving arithmetic and overlay operations: Inventory and prospects. En K. Lowell y A. Jaton (editores) *Spatial Accuracy Assessment: Land information uncertainty in natural resources*. Ann Arbor Press. Chelsea, Michigan.
- Hacfnr, J.W. 1996. *Modeling Biological Systems. Principles and Aplications*. Chapman & Hall.
- Hakånson, L. 1996. A general method to define confidence limits for model predictions based on validation. *Ecological Modelling*. 91:153-168.
- Heuvelink, G y M.F. Goodchild. 1998. *Error Propagation in Envirometal Modelling with GIS*. Taylor & Francis. 127 p.
- Howell, S.N.G. y S. Webb. 1995. *A Guide to the Birds of Mexico and Northern Central America*. Oxford University Press. New York. 851 pp.
- Hone, J. 1991. Analysis of animal survey data . En: C.R. Margules y M.P. Austin (editores).1991 *Nature conservation: Cost effective biological survey and data analysis*. CSIRO. Australia.
- Hwang, D., H. A. Karimi y D.W. Byun. 1998. Uncertainty analysis of enviromental models within GIS enviroments. . *Computers & Geosciences* 24 (2):119-130.
- IMTA. 1996. *Extractor Rápido de Información Climatológica (ERIC)*. Colección de Proyectos IMTA. Serie Programas.
- Klepper, O. 1997. Multivariate aspects of model incertainty analysis: tools for sensitivity analysis and calibration. *Ecological Modelling*. 101:1-13.
- León Paniagua, Livia. 2000. *Computarización de las colecciones del Museo de Zoología "Alfonso L. Herrera" para su incorporación a la REMB: Fase I*. Museo de Zoología "Alfonso L Herrera". Facultad de Ciencias. UNAM. Proyecto J123. Base de datos SNIB-CONABIO. México.

- List, R., O. Moctezuma y P. Manzano-Fischer. 1999. Informe final del proyecto: Identificación de áreas prioritarias para la conservación, corredores y zonas de amortiguamiento en el norte de la Sierra Madre Occidental (Convenio No. B-4-97-7). Naturalia A.C. y The Wildlands Project. Toluca, México. 131 pp.
- Margules, C.R. y A.O. Nicholls. 1987. Assessing the Conservation Value of Remnant Habitat 'Islands': Mallee Patches on the Western Eyre Peninsula, South Australia. p 89-102. En: D.A. Saunders, G.W. Arnold, A. Burbidge y A. Hopkins (editors) *Nature Conservation: The Role of Remnants of Native Vegetation*. Surrey Beatty and Sons. CSIRO. Australia.
- Margules, C.R. y J.L. Stein. 1989. Patterns in the Distributions of Species and the Selection of Nature Reserves: An Example from Eucalyptus Forests in South-eastern New South Wales. *Biological Conservation*. 50: 219-238.
- Margules, C.R., A.O. Nicholls y R.L. Pressey. 1988. Selecting Networks of reserves to maximise biological diversity. *Biological Conservation*. 43: 63-76.
- Miller, R. I. y P.S. White. 1986. Considerations for preserve design based on the distribution of rare plants in Great Smoky Mountains National Park. USA. *Environmental Management* 10:1. 119-124 p.
- Miller, R.I. 1994. *Mapping the Diversity of Nature*. Chapman & Hall. London. England.
- Mittermier, R. A. y C. Goetsch De Mittermier. La importancia de la Diversidad Biológica en México. 63-73 p. En: Sarukhán J. y Dirzo R. 1992. *México ante los retos de la Biodiversidad*. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. México.
- Montaña, C. y P. Greig-Smith. 1990. Correspondence Analysis of Species by Environmental Variable Matrices. *Journal of Vegetation Science*. 1:453-460.
- Monte, L., L. Hakanson, U. Bergstrom, J. Brittain y R. Helting. 1996. Uncertainty Analysis and Validation of Environmental Models: The Empirically Based Uncertainty Analysis. *Ecological Modelling*. 91: 139-152.
- Mowrer, H.T. 1999. Accuracy (re)assurance: Sellin uncertainty assessment to the Uncertain. En: K. Lowell y A. Jaton (editores) *Spatial Accuracy Assessment: Land information uncertainty in natural resources*. Ann Arbor Press. Chelsea, Michigan.

- Navarro Singtienza, Adolfo. 2000. *Atlas de las Aves de México. Fase II*. Museo de Zoología "Alfonso L. Herrera". Facultad de Ciencias. UNAM. Proyecto E18. Base de datos SNIB-CONABIO. México.
- Nicholls, A. O. 1991. Examples of the use of Generalised Linear Models in analysis of survey data for conservation evaluation. En: C. R. Margules y M. P. Austin (editores). 1991 *Nature conservation: Cost effective biological survey and data analysis*. CSIRO. Australia.
- Nicholls, A. O. 1989. How to Make Biological Surveys Go Further with Generalised Linear Models. *Biological Conservation*. 50: 51-75.
- Pendergast, J.R., R.M. Quinn, J.H. Lawton, B.C. Eversham y D.W. Gibbons. 1993. Rare species, the coincidence of diversity hotspots and conservation strategies. *NATURE*. 365(23):335-337
- Phillips, D.L., D.G. Marks. 1996. Spatial Uncertainty Analysis: Propagation of Interpolation Errors in Spatially distributed Models. *Ecological Modelling*. 91:213-229.
- Pielou, E.C. 1977. *Mathematical Ecology*. John Wiley and Sons. 385 p.
- Pressey, R.L., C.J. Humphries, C.R. Margules, R.I. Vane-Wright y P.H. Williams. 1993. Beyond Opportunism: Key principles for systematic Reserve Selection. *TREE*. 8(4):124-128
- Raisbeck, G. 1963. *Information Theory. An Introduction for Scientists and Engineers*. M.I.T. Massachusetts. U.S.A.
- Reza, F.M. 1961. *An introduction to information theory*. McGraw-Hill. New York.
- Rzedowski, J. 1978. *Vegetación de México*. Ed. Limusa. México, D.F. 432 pp.
- Rzedowski, J. 1981. Diversidad y orígenes de la flora fanerógama de México. *Acta Botánica Mexicana*. 14:3-21.
- Scott, J.M., C. Blair, J.D. Jacobi, J.E. Estes. 1987. Species richness. *Bioscience* 37(11):782-788.
- Scott, J.M., F. Davis, B. Csuti, R. Noss, B. Butterfield, C. Groves, H. Anderson, S. Caicco, F. D'Erchia, T.C. Edwards, J. Ulliman, R.G. Wright. 1993. Gap analysis: a geographic approach to protection of biological diversity. *Wildlife Monographs*. 123:1-41.

- Singh, V.P. 2000. The entropy theory as a tool for modelling and decision-making in environmental and water resources. *Water SA*. 26(1):1-11.
- Solares, M. M.J. 2001. *Aplicación de modelos de optimización en la identificación de áreas prioritarias para la conservación: El caso de la Sierra Madre Occidental*. Tesis de Maestría Instituto de Ecología. UNAM. México.
- Summers, J.K., H.T. Wilson, y J. Kou. 1993. A method for quantifying the prediction uncertainties associated with water quality models. *Ecological Modelling*. 65:161-176.
- Terborgh, J. y B. Winter. 1983. A method for siting parks and reserves with special reference to Colombia and Ecuador. *Biological Conservation*. 27:45-58
- USA-CERL. 1998. GRASS 4.2 User's Manual. United States Army Corps of Engineers Construction. Engineering Research Laboratory. Champaign, Illinois.
- Wicken, J.S. 1987. *Evolution, Thermodynamics and Information. Extending the darwinian program*. Oxford University Press. 17-29 p.
- Williams, P.H., Vane-Wright R.I. y C.J. Humphries. 1992. Measuring Biodiversity for Choosing Conservation Areas. In: LaSalle y Gauld (Eds). *Hymenoptera and Biodiversity*. CAB International. Wallingford, UK.
- Zimmerman, H.J. 2000. An application-oriented view of modeling uncertainty. *European Journal of Operational Research*. 122: 190-198.