



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS
Y DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

UN NUEVO ENFOQUE PARA LA MODELACIÓN DE ESPECIES
BIOLÓGICAS BASADO EN LA MECÁNICA CUÁNTICA

TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS

PRESENTA:
GUILLERMO GARRO GÓMEZ

DIRECTOR DE LA TESIS
MIGUEL ARTURO BALLESTEROS MONTERO
IIMAS-UNAM

MIEMBROS DEL COMITÉ TUTOR
LUIS BERNARDO MORALES
IIMAS-UNAM

FRANCISCO TORRES AYALA
FACULTAD DE CIENCIAS-UNAM

CIUDAD UNIVERSITARIA, JULIO, 2021



Universidad Nacional
Autónoma de México



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.

Reconocimientos

Quiero agradecer en primer lugar al Dr. Miguel Ballesteros por la propuesta para realizar esta tesis, así como por el invaluable apoyo y constante motivación. Por su generosidad y amistad. Nada de este trabajo hubiera sido posible sin su participación. Quiero agradecer también a la Dra. Silvia Ruiz por la confianza depositada en este proyecto. Agradezco al Dr. Carlo Díaz-Ávalos por introducirnos en el tema del que se ocupa la primera parte de esta tesis y por sus valiosas aportaciones y conocimientos a la investigación desarrollada aquí. Agradezco también al Dr. Luis B. Morales por sus acertadas correcciones y múltiples sugerencias las cuales mejoraron en buena medida esta investigación. También quiero expresar mi agradecimiento al Dr. Francisco Torres Ayala y al Dr. Luis Silva por sus continuas observaciones y comentarios muchos de los cuales se integraron a la presente tesis. Quiero dar reconocimiento al Dr. Ómar Hernández y al Biol. Daniel Aguilar quienes nos proporcionaron una inestimable asesoría sobre aspectos de ecología y distribución de especies, a la vez que nos brindaron acceso a sus bases de datos recolectados a o largo de 1 año de trabajo de campo. Debo reconocer también el apoyo brindado por el Mat. Gerardo Franco.

Agradezco al Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas (IIMAS) de la UNAM, por proporcionarme una beca de lugar durante la realización de la presente tesis. En particular quiero agradecer al Laboratorio Universitario de Cómputo de Alto Rendimiento (LUCAR) y al Ing. Adrián Durán Chavesti por el apoyo y asesoría en cuanto al uso de los sistemas computacionales del IIMAS necesarios para la realización de este trabajo.

La investigación de esta tesis contó con diversos apoyos:

1. Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) de la UNAM IN108818.
2. Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) de la UNAM IN101621
3. Proyecto apoyado por el CONACYT, FORDECYT-PRONACES, PRONACES/429825/2020.

Mi profundo agradecimiento por los apoyos recibidos.

Índice general

1. Introducción	1
1.1. Objetivos	1
1.2. Metodología de la investigación aplicada a la ecología	2
1.2.1. Hipótesis ecológica	2
1.2.2. Motivación del Modelo Matemático	3
1.2.3. Comparación con los modelos de la Geoestadística	3
1.3. El modelo MaxEnt	5
1.3.1. Representación Matemática de la Situación Experimental	6
1.3.2. Objetivos del Método	8
1.3.2.1. Perspectiva Geográfica	8
1.3.2.2. Perspectiva de las Variables Ambientales	8
1.3.3. Procedimiento Matemático y Postulados	9
1.3.3.1. Perspectiva Geográfica	9
1.3.3.2. Perspectiva de las Variables Ambientales	10
1.3.4. Comparación Entre las Perspectivas Geográficas y de Variables Ambientales.	13
1.3.5. Solución Numérica a los Problemas de Minimización de los Postulados	14
1.4. Muestreo de Gibbs	14
1.4.1. El algoritmo <i>simulated annealing</i>	14
1.4.2. Muestreo de Gibbs	17
1.5. Breve descripción del Modelo Matemático	17
1.5.1. El espacio de configuraciones	17
1.5.2. Funciones de energía (Hamiltonianos)	18
1.5.3. Medida de Gibbs	19
1.5.4. Datos recolectados sobre el río	19
1.5.5. Generación del muestreo de Gibbs	20
1.6. Selección de Parámetros	23

1.6.1.	Temperaturas auxiliares	23
1.6.2.	Estimación de \mathbf{g} , \mathbf{T} y el estado inicial	24
1.7.	El problema de los mínimos locales y la intratabilidad	24
1.8.	Aplicaciones	25
1.9.	Teoría de operadores y de dispersión en la recta real discretizada	26
1.9.1.	Hamiltonianos	26
1.9.2.	Soluciones de Jost	28
1.9.3.	Resultados principales	29
2.	Un modelo numérico para la descripción de la distribución y abundancia de individuos	31
2.1.	Introducción	31
2.2.	Hipótesis ecológicas	31
2.3.	Antecedentes históricos	34
2.4.	Modelo Matemático	36
2.4.1.	La cuadrícula, la estructura de vecinos de primer orden y las configuraciones (o estados)	36
2.4.2.	Hamiltonianos (Funciones de energía)	38
2.4.2.1.	El Hamiltoniano Libre	38
2.4.2.2.	Energía Potencial	38
2.4.2.3.	Hamiltoniano Completo	39
2.4.3.	Medida de probabilidad (medida de Gibbs) y el campo aleatorio de Markov	39
2.4.3.1.	Medida de probabilidad (medida de Gibbs)	39
2.4.3.2.	Campo aleatorio de Markov	40
2.4.4.	Inclusión de los datos del río (el atractor) en el modelo: los hamiltonianos finales	40
2.4.4.1.	Retícula del río y componentes conexas: regiones	40
2.4.4.2.	Simulación Poisson	41
2.4.4.3.	Hamiltoniano con condiciones de contorno en el río	42
2.4.4.4.	Hamiltonianos sobre las Regiones 1,2,3 y 4	43
2.5.	Muestreo de Gibbs	44
2.5.1.	Ordenamiento de las celdas	44
2.5.2.	Probabilidades condicionales para celdas individuales	46
2.5.3.	Cadenas de Markov y distribuciones invariantes	47
2.5.4.	Algoritmo del Muestreo de Gibbs	48
2.5.5.	Selección de parámetros	51
2.5.5.1.	Temperaturas auxiliares y estados iniciales	51

2.5.5.2. Estado inicial, temperatura y constante de acoplamiento	52
2.6. Resultados	53
2.6.1. Resultados principales: distribución y abundancia de individuos	53
2.6.2. Resultados auxiliares: estimación de parámetros y caída de energía	57
2.7. Conclusiones	60
2.8. Un ejemplo simple de convergencia a mínimos locales	61
2.9. Algoritmos	63
2.9.1. Algoritmo para el orden de las celdas	63
2.9.2. Algoritmo Muestreo de Gibbs	64
3. Dinámica poblacional de <i>Plectrohyla Sagorum</i>	65
3.1. Introducción	65
3.1.1. Descripción geográfica de la región	66
3.1.1.1. Bioma	66
3.1.1.2. Geología y edafología	67
3.1.1.3. Climatología	67
3.1.2. Descripción de la especie <i>Plectrohyla Sagorum</i>	67
3.1.2.1. Importancia de los anfibios	68
3.2. Datos recopilados	69
3.3. Muestreo artificial en el río	70
3.4. Hamiltonianos para cada temporada	70
3.5. Muestreo de Gibbs	71
3.6. Selección de T y el Número de Iteraciones.	72
3.7. Estimación del número total de individuos y simulaciones en invierno	72
3.7.1. Selección de la constante de acoplamiento $g^{(i,W_i)}$	73
3.7.2. Número total de individuos	73
3.8. Dinámica de la población durante el año	73
3.8.1. Selección de las constantes de acoplamiento	74
3.9. Resultados	74
3.9.1. Calibración de la temperatura y el número de iteraciones (caso libre)	74
3.9.2. Las constantes de acoplamiento $g^{(i,W_i)}$ y el número total de individuos por región en invierno	77
3.9.3. las constantes de acoplamiento $g^{(i,x)}$, $x \in \{\text{Sp, Su, Fa}\}$	77
3.9.3.1. Número total de individuos	82
3.9.3.2. Resultados principales: distribución y abundancia de individuos	83
3.10. Conclusiones	86

4. Modelado de distribución de baja densidad y especies en peligro crítico: <i>Plectrohyla Guatemalensis</i>	89
4.1. Introducción	89
4.2. El modelo aplicado a especies de baja densidad	90
4.3. Descripción de la especie <i>Plectrohyla Guatemalensis</i>	91
4.4. Daros recolectados	92
4.5. Datos del río	93
4.6. Estados (configuraciones)	94
4.7. Medidas de probabilidad y Hamiltonianos condicionales (dados los datos del río)	95
4.7.1. Orden de las celdas y Muestreo de Gibbs	96
4.7.2. Estimación de parámetros	97
4.7.2.1. Temperaturas auxiliares: T_1 y un rango de T_2	97
4.7.2.2. Estimación de \mathbf{g} y \mathbf{T}	98
4.7.3. Densidades promedio anuales por celda y representación gráfica .	100
4.7.3.1. Individuos representados por puntos	100
4.7.3.2. Mapas de calor	101
4.8. Resultados	101
4.8.1. Temperatura e iteraciones (Caso Libre)	101
4.8.2. La constante de acoplamiento \mathbf{g}	102
4.8.3. Resultados Principales: Distribución y abundancia de individuos	103
4.9. Conclusiones	107
5. Análisis del límite a bajas energías de la matriz de dispersión asociada a la ecuación de Schrödinger discreta	109
5.1. Las soluciones de Jost	109
5.2. La matriz de dispersión	112
5.3. Prueba del resultado principal	114
Bibliografía	121

Introducción

1.1. Objetivos

En la primera parte de la investigación reportada en esta tesis, introducimos un modelo en el contexto de la ecología que puede usarse para describir la distribución y abundancia de individuos de una especie, cuando los datos recolectados de campo son extremadamente limitados (por ejemplo, en el caso de especies en peligro de extinción), y cuando no la información de covariables ambientales no están disponibles. Nuestro procedimiento se basa en una comprensión intuitiva de las propiedades físicas del fenómeno. La idea es que los individuos tienden a sentirse atraídos (o rechazados) por ciertas propiedades del entorno. Al mismo tiempo, se dispersan de tal manera que si no hay razón para que se encuentren en algunas localizaciones específicas, entonces se distribuyen uniformemente por toda la región. Nuestro modelo se basa en la mecánica cuántica, utilizando Hamiltonianos cuánticos en el contexto de la mecánica estadística clásica. El equilibrio entre las fuerzas esparcidoras y atractivas (o repulsivas) determina el comportamiento de las especies que modelamos, y esto se expresa en términos de un problema de control global de un operador energético que es la suma de un término cinético (esparcimiento) y un potencial (atracción o repulsión). Proponemos una solución numérica a este problema de control global que supera la conocida dificultad del muestreo de Gibbs (*annealing*) que es el hecho de que un control global es difícilmente accesible cuando el número de variables es grande (los algoritmos se atascan en un estado no óptimo). En [17, 19, 20] aplicamos este modelo al estudio de dos especies de anfibios en peligro crítico de extinción endémicas de la región del Volcán Tacaná (Soconusco), Chiapas, México. En las aplicaciones del modelo hacemos uso de los datos reportados en [4]. Esta investigación pretende contribuir a una mejor comprensión de las especies citadas y con ello contribuir a su preservación.

La segunda parte de la tesis trata sobre la teoría de operadores y de dispersión

en la recta real discretizada. Específicamente, hacemos un análisis del límite a bajas energías de la matriz de dispersión asociada a la ecuación de Schrödinger discreta con un potencial matricial que tiene un primer momento finito. Los coeficientes de transmisión se continúan analíticamente y se extienden a los bordes de la banda. Se proporciona una expresión explícita para estas extensiones. También se calculan los límites de los coeficientes de reflexión en los bordes de la banda. Esta parte de la investigación se ha realizado en colaboración con el matemático Gerardo Franco, quien realiza la mayor contribución en este tópico como parte de sus estudios de doctorado que están por iniciar bajo la dirección del Dr. Miguel Ballesteros.

1.2. Metodología de la investigación aplicada a la ecología

1.2.1. Hipótesis ecológica

La idea clave de nuestro modelo es notar que existen características geográficas y ambientales que atraen o repelen a los individuos de las especies biológicas. Nuestras hipótesis ecológicas principales están basadas en dos ideas intuitivamente claras: Primero, que los individuos de una especie biológica tienden a esparcirse uniformemente sobre la región de su hábitat en caso de ausencia de variables ambientales que actúen como atractores o repulsores; segundo, los individuos de una especie biológica son sensibles a la presencia de atractores o repulsores ambientales, esto es, o bien se concentran o se alejan de ciertas regiones de acuerdo a condiciones topográficas (presencia de cañones o zonas inaccesibles) o bien por razones relativas a presencia o ausencia de depredadores, y otros factores medio ambientales (ver Sección 2.2).

Estos dos comportamientos compiten entre sí y la distribución real de los individuos está determinada por un problema de minimización u optimización de una función de energía cuya inspiración radica en la mecánica cuántica. Por un lado, definimos una *energía cinética* que representa la capacidad de la especie objetivo a dispersarse en toda la región de su hábitat; y por otro lado, consideramos una *energía potencial* la cual controla la fuerza con que la especie es atraída o bien dispersada en caso de existir tales fuerzas. La energía del sistema, el Hamiltoniano, queda así definida por la suma de estas dos energías. Entonces suponemos que la tendencia natural de la especie es permanecer en un estado de mínima energía. Así que nuestra tarea consiste en encontrar dicha configuración de mínima energía, que además coincide con el estado de máxima probabilidad del sistema, porque la distribución de probabilidad será definida de manera natural como una medida espacial de Gibbs (o campo aleatorio de Markov) que tiene al Hamiltoniano del sistema como generador.

1.2.2. Motivación del Modelo Matemático

En nuestra investigación, es fundamental considerar el equilibrio entre las fuerzas de atracción y la tendencia de los individuos a dispersarse, para determinar el comportamiento de las especies que modelamos, y esto se expresa matemáticamente en términos de un problema de minimización de un operador de energía, es cual tiene una expresión matemática como la suma de un término cinético y un potencial.

Abordamos el problema utilizando la mecánica estadística clásica. A partir de las probabilidades condicionales, la medida de probabilidad de todo el sistema se obtiene con la ayuda de un hecho bien conocido en mecánica estadística que es que los campos aleatorios de Markov se distribuyen de acuerdo a medidas de Gibbs (teorema de Hammersley-Clifford). Por lo tanto, en lugar de centrarnos en probabilidades condicionales, abordamos directamente la medida conjunta, que es una medida de Gibbs. En particular, elegimos un Hamiltoniano que representa el comportamiento de los individuos en el espacio geográfico.

1.2.3. Comparación con los modelos de la Geoestadística

Algunos problemas relacionados a la reconstrucción de mapas de abundancia (o solo presencia) y distribución de especies han sido tratados por la Geoestadística. La base principal de la Geoestadística son los *campos aleatorios de Markov* (Besag [26]) los cuales se definen asumiendo distribuciones muy específicas (pertenecientes a la familia exponencial generalmente) para las probabilidades condicionales, tomando un criterio de máxima verosimilitud para estimar los parámetros. A partir de las probabilidades condicionales, la medida de probabilidad de todo el sistema se obtiene con la ayuda de un hecho bien conocido en mecánica estadística que es que los campos aleatorios de Markov se distribuyen de acuerdo con las medidas de Gibbs.

Los modelos de Besag introdujeron la base de la geoestadística en términos de campos aleatorios de Markov, y en la actualidad se utilizan con mucha frecuencia en el área de estadística aplicada en diversos contextos, como la ecología, y la medicina contra el cáncer (ver por ejemplo [28, 36, 37, 47, 71, 75, 88, 102, 150, 151, 152, 161, 162, 163, 164]). Como hemos mencionado, el enfoque fundamental de tales trabajos es que se asume que las probabilidades condicionales siguen distribuciones de probabilidad conocidas (generalmente pertenecientes a la familia de distribuciones exponencial). Desde este enfoque, el análisis de los datos se realiza entonces de acuerdo a los modelos jerárquicos de estimación paramétrica de la estadística bayesiana.

Es también conocido que si se aborda el problema desde la perspectiva usual mediante los estimadores de máxima verosimilitud, entonces nos enfrentamos a la intra-

tabilidad de los cálculos involucrados, por lo que generalmente se recurre a métodos numéricos de optimización aproximada como descenso del gradiente o, alternativamente, el método conocido como *simulated annealing*. Nosotros recurrimos a un método elaborado en el trabajo realizado por Geman-Geman [61], en donde se presenta por primera vez el algoritmo numérico conocido como *muestreo de Gibbs* o *Gibbs sampling*.

En Geman-Geman [61] se destaca que la distribución conjunta completa de un campo aleatorio de Markov toma una forma simple como una medida de Gibbs (también conocida como medida de Gibbs-Boltzmann), hecho que Besag [26] había ya demostrado en el teorema de Hammersley-Clifford, pero que no había sido explotado con mayor profundidad. Geman-Geman [61] además proporciona un algoritmo computacional que es matemáticamente robusto (el *muestreo de Gibbs* o *Gibbs sampling*). Esto permite implementar la maquinaria de la mecánica estadística como los métodos numéricos como Monte Carlo Markov Chain y Metropolis-Hastings. La estimación estadística equivale a encontrar las configuraciones más probables con respecto a las medidas de Gibbs del sistema. Adicionalmente, Geman-Geman [61] proporciona una de las primeras pruebas matemáticas de convergencia para una técnica de optimización de Monte Carlo llamada *simulated annealing* (después fue mejorado por Hayek [69] y otros autores). Una revisión completa sobre este asunto se encuentra en [129, 158]).

Motivados por el trabajo de Geman-Geman [61], nosotros planteamos el problema de la abundancia y distribución de especies desde la perspectiva de la física estadística clásica, y no asumimos ninguna distribución para los parámetros que intervienen en el problema de optimización (por lo que descartamos los métodos disponibles de la estadística bayesiana). Recurrimos para ello al algoritmo propuesto en [61], con una variante en la secuencia de ejecución del tal algoritmo, la cual explicamos con mayor detalle en el capítulo siguiente. En la siguiente sección también expondremos brevemente esta propuesta. En el presente trabajo también discutimos la deficiencia más importante y a menudo presente del algoritmo de muestreo de Gibbs, la cual consiste en que no proporciona aproximaciones cercanas a los estados óptimos en situaciones poco controladas (a diferencia del *simulated annealing*). Esto sucede debido fundamentalmente a que el algoritmo de muestreo de Gibbs considera las distribuciones condicionales locales (en cada sitio del sistema) para realizar las transiciones aleatorias a los estados más favorables (con energía más baja) en cada iteración. Esto imposibilita el control global del modelo y provoca que el algoritmo se atasque en estados que llamaremos mínimos locales en el sentido establecido en la Definición 1.5.1. Los resultados que obtenemos en las aplicaciones (Capítulos 3 y 4) parecen sugerir que nuestro enfoque permite superar esta deficiencia.

Debemos enfatizar que no estamos proponiendo una solución al problema analítico de minimización global (nótese que un mínimo local de Definición 1.5.1 no es un míni-

mo local de la función objetivo). Proponemos un modelo matemático y un esquema numérico con el objetivo de mantener un control global sobre el sistema, en el sentido de que los parámetros se estiman de acuerdo con el comportamiento global esperado de los fenómenos. En nuestra opinión, de acuerdo con los resultados obtenidos en las aplicaciones del modelo (Capítulos 3 y 4), alcanzamos este objetivo.

1.3. El modelo MaxEnt

En esta sección presentamos el modelo matemático conocido como MaxEnt, que se usa comúnmente para predecir presencia de especies a través de variables ambientales. Este método es substancialmente más sencillo del que se introduce en la tesis y no es compatible con las situaciones que se describen en la misma (y tampoco se puede aplicar para modelar los fenómenos que se presentan en la tesis). La principal razón por la que MaxEnt es mucho más sencillo es que considera celdas desacopladas, es decir, cada celda es independiente de las otras, de manera que se analiza cada una por separado. En contraste, en el modelo de la tesis el comportamiento de cada celda influye a todas las demás como en una membrana, es como una discretización de una ecuación de onda (o una ecuación de Schrödinger). De esta manera, nuestro modelo permite resolver el problema de abundancia y distribución usando muy pocos datos (como ocurre en las ecuaciones diferenciales parciales, en donde los datos iniciales, o datos a la frontera, determinan las soluciones). Contrariamente, el modelo MaxEnt requiere de información de variables ambientales en todos los píxeles de la región de estudio. Estos datos son imposibles de obtener en las situaciones que nosotros consideramos porque la escala en la que trabajamos es de 5 metros (la medida de los lados de las celdas), por lo que no existe ninguna base de datos apropiada (el muestreo de datos ambientes precisos a esta escala requeriría de una cantidad exorbitante de horas hombre). Además, nuestro modelo es aplicable para muestreos con muy pocos registros (como especies en peligro de extinción) y se utiliza también para estudiar regiones muy pequeñas (por ejemplo, de una hectárea).

MaxEnt es uno de los modelos para distribución de especies biológicas basados en ideas termodinámicas más usuales y actualmente en constante desarrollo (ver por ejemplo [12] y [87] para una revisión general del modelo). Fue introducido por Edwin T. Jaynes en [85] y [86] hacia 1957 en el contexto de la mecánica estadística y la teoría de la información, como un método eficiente y matemáticamente simple para estimar distribuciones de probabilidad. Este procedimiento fue adaptado rápidamente a los estudios de especies biológicas en el contexto de la ecología (para una visión general del desarrollo de MaxEnt en este sentido se puede consultar [72]). Es particularmente

usado en el problema de determinar la distribución geográfica de especies usando datos de solo presencia, ya que está basado únicamente en condiciones ambientales, lo que lo convierte en un método capaz de superar las deficiencias reativas a la calidad de los datos (generalmente los datos de ausencia no están disponibles o son deficientes). Es además un método muy flexible y cuenta con un creciente número de aplicación en diversas áreas como astronomía, optimización de portafolios financieros, reconstrucción de imágenes y procesamiento de señales, entre otras ([90, 121]).

En el contexto de la ecología, el objetivo de método MaxEnt es estimar la distribución potencial de una especie de acuerdo a la “idoneidad” del ambiente geográfico que constituye su hábitat natural (requisitos de supervivencia de la especie), bajo el principio termodinámico fundamental de que sin influencias externas, el sistema biológico tiende naturalmente a un estado de máxima entropía (principio de máxima entropía) [11, 54, 118, 119, 120, 133]. El modelo es fundamentalmente desacoplado, en el sentido de que no es necesario considerar correlaciones espaciales, y el ajuste depende exclusivamente los datos de solo presencia en una muestra aleatoria de sitios, y de los datos del ambiente ecológico en la totalidad del espacio geográfico de estudio dentro del cual se encuentra el hábitat de la especie (el ambiente geográfico está descrito mediante variables ambientales o covariables, tales como temperatura, precipitación media o humedad, entre otros - o bien por funciones las mismas, llamadas “características”). Esto difiere de una de las características fundamentales del modelo propuesto en esta tesis, en el cual las relaciones espaciales controlan las perturbaciones globales de la distribución de la especie a través de un sistema de vecinos local, como una membrana. Además, como se menciona arriba, nuestro modelo no requiere del conocimiento de las variables ambientales en la región completa de estudio.

1.3.1. Representación Matemática de la Situación Experimental

La región de estudio se divide en un conjunto finito de celdas. Denotamos por L a este conjunto. El modelo considera que los datos de las variables ambientales (temperatura presión, etc.) se conocen en cada celda. Estos datos se denotan con una función

$$Z : L \rightarrow \mathbb{R}^d$$

(en donde d es el número de variables ambientales). El modelo incluye un vector aleatorio

$$(X, Y) : S \rightarrow L \times \{0, 1\}.$$

En donde $S \equiv (S, \mathcal{F}, \mu)$ es un espacio de medida (S es un conjunto, \mathcal{F} una sigma álgebra y μ es una medida). Además,

$$X : S \rightarrow L, \quad Y : S \rightarrow \{0, 1\}$$

son funciones medibles (variables aleatorias) tales que:

1. $\mathbb{P}(X = x) = \mu(X^{-1}(\{x\})) =$ “Probabilidad de elegir x como punto de muestreo”.
- 2.

$$\mathbb{P}(X = x, Y = 0) = \mu((X, Y)^{-1}(\{(x, 0)\}))$$

= “Probabilidad de elegir x como punto de muestreo y que no se encuentre la especie en este punto”.

- 3.

$$\mathbb{P}(X = x, Y = 1) = \mu((X, Y)^{-1}(\{(x, 1)\}))$$

= “Probabilidad de elegir x como punto de muestreo y que se encuentre la especie en este punto”.

La densidad marginal

$$p(x) := \mathbb{P}(X = x)$$

se elige de acuerdo al criterio del ecólogo. La selección de puntos de muestreo (algunos elementos de L) se debería hacer con respecto a la distribución p (por medio de un generador de elementos aleatorios). El experimento consiste en visitar los sitios (celdas o pixeles) elegidos y verificar si se encuentra la especie. De esta forma, en cada celda visitada se obtendría la información de presencia o ausencia. Sin embargo, generalmente los datos de ausencia no se reportan o sus registros son muy deficientes. En consecuencia, normalmente solo se cuenta con datos de presencia. Estos datos consisten de un conjunto de celdas

$$x_1, x_2, \dots, x_m$$

en donde el muestreo reporta presencia.

1.3.2. Objetivos del Método

1.3.2.1. Perspectiva Geográfica

En este contexto, la meta del algoritmo MaxEnt es encontrar la probabilidad de presencia de una especie en cada punto x de L . De manera matemática, el objetivo es encontrar la siguiente distribución de probabilidad.

$$\pi(x) = \mathbb{P}(X = x \mid Y = 1), \quad x \in L. \quad (1.1)$$

Esto es, la probabilidad dada la presencia de la especie, de que ésta se encuentre en el sitio x (ver [118, 119, 120]).

1.3.2.2. Perspectiva de las Variables Ambientales

Definimos

$$\mathbf{Z} := Z \circ X$$

la variable aleatoria que representa la distribución de las variables ambientales: la probabilidad de que las variables ambientales de las celdas de muestreo estén en el conjunto K es

$$\mathbb{P}(\mathbf{Z} \in K).$$

En este contexto, la meta de MaxEnt es encontrar la probabilidad de presencia de la especie sujeta a los valores de las variables ambientales. Es decir, se busca reconstruir

$$\mathbb{P}(Y = 1 \mid \mathbf{Z} = \mathbf{z}). \quad (1.2)$$

En otras palabras, se quiere saber cuales son los valores de las variables ambientales que provocan la presencia de la especie. Esta probabilidad se puede encontrar con dato de presencia y ausencia. Sin embargo, en el caso de que únicamente contemos con datos presencia solo se puede obtener (1.2) multiplicada por una constante desconocida (ver [54] y [120]).

1.3.3. Procedimiento Matemático y Postulados

1.3.3.1. Perspectiva Geográfica

El modelo MaxEnt se usa para estimar la distribución π mediante la solución de un problema de minimización convexo del modo siguiente (ver [85] y [118]). Nos interesa encontrar una distribución $\hat{\pi}$ que se aproxime a π y que esté sujeta además a ciertas constricciones (que sea compatible con los dato de muestreo).

El modelo considera n funciones

$$h_i : \mathbb{R}^d \rightarrow \mathbb{R}, \quad i \in \{1, \dots, n\}$$

llamadas características ambientales que el ecólogo elige de manera arbitraria de acuerdo a su experiencia (representan la información relevante de las variables ambientales, de acuerdo a la opinión del ecólogo).

La medida empírica de π , que se obtiene de los datos experimentales, es

$$\tilde{\pi}(x) = \frac{1}{m} |\{1 \leq i \leq m : x_i = x\}|. \quad (1.3)$$

El postulado de MaxEnt asume que la mejor estimación, $\hat{\pi}$, para π es la función de densidad que satisface lo siguiente:

Postulado 1.3.1 (MaxEnt - Variables Ambientales 1). *La mejor aproximación para π es la función $\hat{\pi}$ que satisface:*

1. *Los valores de expectación de $h_i \circ Z$ con respecto a $\tilde{\pi}$ y $\hat{\pi}$ coinciden, para toda i :*

$$\mathbb{E}_{\tilde{\pi}}(h_i \circ Z) = \mathbb{E}_{\hat{\pi}}(h_i \circ Z). \quad (1.4)$$

2. *$\hat{\pi}$ es la distribución de menor entropía que satisface las constricciones (1.13):*

$$H(\hat{\pi}) = \min\{H(\rho) \mid \rho \text{ es una distribución que satisface (1.4)}\} \quad (1.5)$$

Nota 1.3.1. *En el postulado 1.3.1, $\mathbb{E}_\rho(g)$ representa el valor de expectación de g con*

respecto a la distribución ρ . Además

$$H(\rho) := \mathbb{E}_\rho(\ln \rho) \quad (1.6)$$

es la entropía de ρ . Esta entropía representa el desorden de la medida ρ . El postulado asevera que la mejor elección de la distribución $\hat{\pi}$ es aquella que tiene los mismos valores de expectación que la distribución empírica de los resultados experimentales y que presenta mayor desorden. En otras palabras, se elige la distribución que tiene el menor sesgo (que tiende a no favorecer ninguna tendencia particular).

Las constricciones 1.4 en el Postulado 1.3.1 resultan ser en muchas ocasiones muy restrictivas. Esto produce que la solución $\hat{\pi}$ al Postulado 1.3.1 pueda presentar una propiedad de sobre-ajuste (overfitting), ver [50, 54, 118]. Esto implica que el modelo resultante no siempre apropiado. Para relajar estas constricciones se se asume que los valores de expectación en (1.4) no sean iguales sino que se acercan lo suficiente, a una distancia λ_i . Los valores de $\lambda_i \geq 0$ se eligen de acuerdo al criterio del ecólogo:

Postulado 1.3.2 (MaxEnt - Variables Ambientales 2). *La mejor aproximación para π es la función $\hat{\pi}$ que satisface:*

1. Los valores de expectación de $h_i \circ Z$ con respecto a $\tilde{\pi}$ y $\hat{\pi}$ coinciden, para toda i :

$$|\mathbb{E}_{\tilde{\pi}}(h_i) - \mathbb{E}_{\hat{\pi}}(h_i)| \leq \lambda_i. \quad (1.7)$$

2. $\hat{\pi}$ es la distribución de menor entropía que satisface las constricciones (1.13):

$$H(\hat{\pi}) = \min\{H(\rho) \mid \rho \text{ es una distribución que satisface (1.7)}\} \quad (1.8)$$

1.3.3.2. Perspectiva de las Variables Ambientales

Como L es finito, entonces

$$Z(L) = D$$

es finito. Entonces consideramos

$$\mathbf{Z} : S \rightarrow D.$$

Denotamos por $f(\mathbf{z})$ la densidad marginal de \mathbf{Z} :

$$f(\mathbf{z}) = \frac{|\{x \in L : \mathbf{Z}(x) = \mathbf{z}\}|}{|L|}, \quad \mathbf{z} \in D. \quad (1.9)$$

Sea el vector aleatorio (Y, \mathbf{Z}) definido en L y con valores en el espacio producto $\{0, 1\} \times D$. Definimos

$$f_1(\mathbf{z}) = \mathbb{P}(\mathbf{Z} = \mathbf{z} \mid Y = 1), \quad \mathbf{z} \in D \quad (1.10)$$

y notamos que la regla de Bayes implica que

$$\mathbb{P}(Y = 1 \mid \mathbf{Z} = \mathbf{z}) = \frac{f_1(\mathbf{z})}{f(\mathbf{z})} \mathbb{P}(Y = 1), \quad \mathbf{z} \in D. \quad (1.11)$$

Como se explica arriba, la función de densidad f se supone conocida. De manera que la probabilidad objetivo $\mathbb{P}(Y = 1 \mid \mathbf{Z} = \mathbf{z})$ se puede encontrar conociendo f_1 y $\mathbb{P}(Y = 1)$. Este último número se puede encontrar únicamente si se cuenta con dato de presencia y ausencia. Con los de presencia solo podemos encontrar el ratio f_1/f la cual es la información objetivo para MaxEnt en este contexto. El problema se reduce a encontrar la distribución marginal f_1 .

El modelo MaxEnt se usa para estimar la distribución f_1 mediante la solución de un problema de minimización convexo del modo siguiente (ver [85] y [118]). Nos interesa encontrar una distribución \hat{f}_1 que se aproxime a f_1 y que esté sujeta además a ciertas constricciones (que sea compatible con los dato de muestreo).

Denotamos por

$$\mathbf{z}_i := Z(x_i), \quad i \in \{1, \dots, m\}.$$

El modelo considera n funciones

$$h_i : D \rightarrow \mathbb{R}, \quad i \in \{1, \dots, n\}.$$

llamadas características ambientales que el ecólogo elige de manera arbitraria de acuerdo a su experiencia (representan la información relevante de las variables ambientales, de acuerdo a la opinión del ecólogo).

La medida empírica de f_1 , que se obtiene de los datos experimentales, es

$$\tilde{f}_1(\mathbf{z}) = \frac{1}{m} |\{1 \leq i \leq m : \mathbf{z}_i = \mathbf{z}\}|, \quad \mathbf{z} \in D. \quad (1.12)$$

1. INTRODUCCIÓN

El postulado de MaxEnt asume que la mejor estimación, \hat{f}_1 , para f_1 es la función de densidad que satisface lo siguiente:

Postulado 1.3.3 (MaxEnt - Variables Ambientales 1). *La mejor aproximación para f_1 es la función \hat{f}_1 que satisface:*

1. *Los valores de expectación de h_i con respecto a f_1 y \hat{f}_1 coinciden, para toda i :*

$$\mathbb{E}_{f_1}(h_i) = \mathbb{E}_{\hat{f}_1}(h_i). \quad (1.13)$$

2. *\hat{f}_1 es la distribución de menor entropía relativa con respecto a f que satisface las constricciones (1.13):*

$$KL(f, \hat{f}_1) = \min\{KL(f, \rho) \mid \rho \text{ es una distribución que satisface (1.13)}\} \quad (1.14)$$

Nota 1.3.2. *En el postulado 1.3.3, $\mathbb{E}_\rho(g)$ representa el valor de expectación de g con respecto a la distribución ρ . Además*

$$KL(f, \rho) := \mathbb{E}_f \left(\ln \frac{f}{\rho} \right) \quad (1.15)$$

es la entropía relativa entre f y ρ , también conocida como divergencia de Kullback-Leibler. Esta entropía representa una noción de distancia entre f y ρ , puesto que es una función convexa cuyo mínimo se alcanza cuando $f = \rho$. El postulado asevera que la mejor elección de la distribución \hat{f}_1 es aquella que tiene los mismos valores de expectación que la distribución empírica de los resultados experimentales y que se acerca más a la distribución de las variables ambientales f . Esto último quiere decir, por ejemplo, que las variables ambientales con mayor presencia geográfica inducen que sea más factible encontrar la especie en alguna locación asociada a estas variables ambientales.

Las constricciones 1.13 en el Postulado 1.3.3 resultan ser en muchas ocasiones muy restrictivas. Esto produce que la solución \hat{f}_1 al Postulado 1.3.3 pueda presentar una propiedad de sobre-ajuste (overfitting), [50, 54, 118]. Esto implica que el modelo resultante no sea apropiado. Para relajar estas constricciones se asume que los valores de

expectación en (1.13) no sean iguales sino que se acercan lo suficiente, a una distancia λ_i . Los valores de $\lambda_i \geq 0$ se eligen de acuerdo al criterio del ecólogo:

Postulado 1.3.4 (MaxEnt - Variables Ambientales 2). *La mejor aproximación para f_1 es la función \hat{f}_1 que satisface:*

1. *Los valores de expectación de h_i con respecto a f_1 y \hat{f}_1 coinciden, para toda i :*

$$|\mathbb{E}_{f_1}(h_i) - \mathbb{E}_{\hat{f}_1}(h_i)| \leq \lambda_i. \quad (1.16)$$

2. *\hat{f}_1 es la distribución de menor entropía relativa con respecto a f que satisface las restricciones (1.13):*

$$KL(f, \hat{f}_1) = \min\{KL(f, \rho) \mid \rho \text{ es una distribución que satisface (1.16)}\} \quad (1.17)$$

1.3.4. Comparación Entre las Perspectivas Geográficas y de Variables Ambientales.

En [54] se explica que las perspectivas geográficas y de variables ambientales son equivalentes. El problema geográfico tiene una caracterización como maximización de cierta familia paramétrica de distribuciones de Boltzmann-Gibbs (ver [85] para una prueba detallada) de la forma

$$q_{\beta}(x) = \frac{e^{\beta \cdot \mathbf{h}(Z(x))}}{Q_{\beta}}, \quad (1.18)$$

donde $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{R}^n$ es un vector de parámetros (que interpretamos como los “pesos” que tiene cada una de las características h_1, \dots, h_n), $\mathbf{h} = (h_1, \dots, h_n)$ es el vector de funciones características, y Q_{β} es la constante de normalización.

De acuerdo a [54], el problema de variables ambientales tiene una caracterización similar:

$$f_1(\mathbf{z}) = f(\mathbf{z})e^{\alpha + \beta \cdot \mathbf{h}(\mathbf{z})}, \quad \mathbf{z} \in D, \quad (1.19)$$

en donde hemos mantenido la notación para el parámetro β , y α es un parámetro que asegura que $\sum_{\mathbf{z}} f_1(\mathbf{z}) = 1$. En [54], se argumenta que las ecuaciones (1.18) y (1.19) implican que las perspectivas de variables ambientales y geográficas son equivalentes.

1.3.5. Solución Numérica a los Problemas de Minimización de los Postulados

En [50], Figure 2 y Figure 3, se proponen dos algoritmos, llamados “actualización secuencial” y “actualización en paralelo”, respectivamente, que resuelven los problemas de minimización planteados en los Postulados 1.3.3, 1.3.4, 1.3.1 y 1.3.2 (ver [50], Theorem 2 y Theorem 3). El problema de minimización en [50] se plantea en la página en donde se presentan las Ecuaciones (1) y (2) del mismo texto.

1.4. Muestreo de Gibbs

1.4.1. El algoritmo *simulated annealing*

Uno de los métodos numéricos más usuales al que se recurre en los problemas de optimización es llamado *simulated annealing*, inspirado en lo procesos de templado de la física de materiales, fue introducido independientemente en [93] y en [146]. En su forma mas simple, el algoritmo simulated annealing está basado en la idea de que la distribución de probabilidades que determina el estado físico de una partícula en un determinado momento t que se mantiene bajo un baño de calor a temperatura $T(t)$, está dada por la bien conocida distribución de Gibbs-Boltzmann

$$G_{T(t)}(x) = \frac{1}{Z_{T(t)}} \exp \left[-\frac{H(x)}{T(t)} \right], \quad (1.20)$$

donde:

- $H(x)$ es el operador Hamiltoniano de la energía y es la función a minimizar (función objetivo), y x está en un dominio finito S (el espacio de *estados* o *configuraciones*),
- $T(t) > 0$ es un parámetro de *temperatura*, que es una función determinista que depende únicamente del tiempo t que satisface un *programa determinista de descenso*, es decir,

$$\begin{aligned} T(0) &\geq T(1) \geq \dots \geq 0, \\ \lim_{t \rightarrow \infty} T(t) &= 0. \end{aligned} \quad (1.21)$$

El esquema o programa de descenso de la temperatura es la cuestión más importante del algoritmo. Controla la aleatoriedad con la que ocurren los cambios a

estados más favorables durante la implementación del algoritmo.

- $Z_{T(t)}$ es la función partición de normalización.

Note que cuando la temperatura $T(t)$ es pequeña, la probabilidad de desplazamiento (transiciones) ascendentes es pequeña y, por lo tanto, el algoritmo tardará mucho en escapar de estados no óptimos (Hajek en [69] define una compleja estructura de mínimos locales en términos del comportamiento del algoritmo, nosotros adaptamos su definición en nuestro contexto en la Definición 1.5.1); y si la temperatura es grande, las transiciones descendentes suceden raramente y el algoritmo puede oscilar durante mucho tiempo entorno a un mínimo global antes de alcanzarlo. En [61, 69, 115, 130, 140, 141, 144] se estudia ampliamente la dependencia del comportamiento asintótico del algoritmo respecto al esquema (programa) de descenso de la temperatura y del “tamaño” del espacio de estados o configuraciones del sistema (“critical depth”). En la práctica se ha comprobado que el mejor programa de descenso de temperatura con garantía de convergencia es de orden logarítmico, lo que explica la principal desventaja teórica del método, la cual es que se necesita mucho tiempo para encontrar los óptimos globales de la función objetivo; y aunque en la práctica se han desarrollado con cierto éxito algoritmos con programas de enfriamiento más rápidos, no hay hasta ahora pruebas matemáticas de convergencia por lo que no hay garantía de convergencia a los óptimos globales.

Implementación del algoritmo *simulated annealing*

El algoritmo del *simulated annealing* se basa en un esquema usual tipo Metropolis Markov Chain que se describe como sigue (ver [46] para más detalles): El estado físico de la partícula (regido por (1.20)) pasa de un estado x al tiempo t , a un estado y al tiempo $t + 1$ con probabilidad

$$p_t = \min \left\{ 1, \frac{G_{T(t)}(y)}{G_{T(t)}(x)} \right\}. \quad (1.22)$$

El cociente

$$\frac{G_{T(t)}(y)}{G_{T(t)}(x)} = \exp \left(-\frac{1}{T(t)} [H(y) - H(x)] \right)$$

evita en la práctica computacional el cálculo de la función partición $Z_{T(t)}$, lo cual se convierte generalmente en un cálculo intratable según el tamaño del espacio de estados y la complejidad de H .

Ahora bien, en el contexto del *simulated annealing*, los resultados acerca de la convergencia del algoritmo a la solución óptima requieren además de ciertas restricciones

de comportamiento locales relacionadas a una estructura de vecinos sobre el espacio de estados S . Esta dependencia se introduce en el modelo interpretando al espacio de estados S como un gráfico no dirigido, o bien, interpretando a S como el conjunto imagen de un gráfico no dirigido E , el conjunto de *sitios*, con respecto a una función que asigna a cada sitio de E una configuración en S . Los conjuntos de estados “vecinos” son las componentes conexas de este gráfico. De esta manera, las probabilidades de transición p_t en la ecuación 1.22 son modificadas de la siguiente manera: si el sistema se encuentra en un estado x en un tiempo t , solo puede cambiar a un estado y al tiempo $t + 1$ con probabilidad positiva si y es vecino de x . Esto es,

$$p_t = \begin{cases} \min \left\{ 1, \frac{G_{T(t)}(y)}{G_{T(t)}(x)} \right\} & \text{si } x \text{ y } y \text{ son vecinos,} \\ 0, & \text{en otro caso.} \end{cases}$$

En Hajek [69] se prueba uno de los mejores resultados sobre convergencia (para otras pruebas ver también Tsitsiklis [141] y Nolte y Schrader [115]). Hajek [69] demuestra que el algoritmo converge a los óptimos globales en probabilidad si y sólo si la serie

$$\sum_{t=1}^{\infty} \exp(-d/T(t))$$

diverge, en donde d es una constante (“profundidad crítica”). Hajek [69] demuestra entonces que para alcanzar los mínimos globales es suficiente considerar un programa logarítmico propuesto anteriormente por Geman-Geman [61] de la forma

$$T(t) = \frac{c}{\log(t+1)}, \tag{1.23}$$

donde c es una constante tal que $c \geq d$. El problema radica en que la constante d es inalcanzable en la práctica (Nolte y Schrader [115]). Se han probado ([59, 113, 115]) algunos otros resultados de convergencia donde la constante c tiene cotas menores que d para el programa de enfriamiento logarítmico (1.23), pero requieren igualmente números grandes de iteraciones (de orden exponencial) para asegurar suficiente proximidad a los mínimos globales (ver Capítulo 3 en [97] para un análisis comparativo de estos algoritmos). No obstante, el algoritmo de *simulated annealing* ha sido extensamente usado (y algunas hibridaciones del mismo) en problemas de optimización en diversas áreas (ver [2, 23, 24, 34, 61, 69, 97, 140] para una extensa revisión del método, y para algunas variantes de las técnicas y aplicaciones modernas ver [33, 46, 76, 83, 108, 130, 130, 142]).

1.4.2. Muestreo de Gibbs

En Geman-Geman [61], el muestreo de Gibbs se introduce como una manera de simular una muestra aleatoria de un campo aleatorio de Markov definido sobre un gráfico (un *lattice*) por la distribución de probabilidad (1.20), en un esquema de temperatura constante, es decir, $T(t) = T > 0$ para toda t (Theorem A [61]), con la cual podemos aproximar el valor de expectación (relativo a la distribución del campo aleatorio de Markov) de cualquier otra variable (Theorem C [61]). En este trabajo, ambos resultados son de suma importancia para elaborar una simulación consistente del estado real de la distribución y abundancia de la especie biológica, bajo únicamente los supuestos ecológicos descritos antes en la Sección 1.2.1.

La cuestión importante del algoritmo es que las transiciones a estados más favorables son realizadas en un sólo sitio a la vez en cada iteración. La principal razón de ello es explotar la “observación pivotal” ([61], p. 722) de que las distribuciones condicionales de un campo aleatorio de Markov son nuevamente distribuciones de Gibbs, es decir, tienen la forma (1.20), y al mismo tiempo “explotar” la estructura de relaciones de vecindad del sistema: la distribución condicional sobre un solo sitio con respecto al resto del sistema, es equivalente a la distribución condicional únicamente con respecto a el conjunto de vecinos de dicho sitio. Esto ofrece una ventaja computacional al algoritmo (porque el conjunto de sitios vecinos es un conjunto pequeño). Pero este esquema de transiciones locales también engendra la principal deficiencia computacional del muestreo de Gibbs, pues sin una elección cuidadosa de los estados iniciales y la temperatura se corre el riesgo de quedar atascado en algún mínimo local (en el sentido de la Definición 1.5.1).

1.5. Breve descripción del Modelo Matemático

A continuación describimos en términos generales el modelo matemático que introducimos en la investigación en la parte aplicada a la ecología. La descripción más detallada está en el Capítulo 2.

1.5.1. El espacio de configuraciones

Por simplicidad, partimos de una situación en la que el campo de estudio es una región geográfica rectangular, aunque el modelo matemático que proponemos puede extenderse a regiones con cualquier otra forma. Nos enfrentamos al problema de determinar la distribución y abundancia de los individuos de alguna especie biológica dentro de dicha región. Suponemos además que dicha región cuenta con características geográficas que determinan la abundancia y la distribución de los individuos de la espe-

1. INTRODUCCIÓN

cie. Por motivos expositivos, explicamos nuestro método usando un ejemplo: elegimos una región específica (con una orografía complicada, un cañón, ubicado en un bosque nuboso de Chiapas, México, ver Figura 1.1). La región que estudiamos se ubica en los alrededores del Volcán Tacaná (Pueblo de Chiquihuites) en el municipio de Unión Juárez, Chiapas, México. Es un área rectangular con lados de longitud $960m$ y $1010m$. Las coordenadas UTM (usamos la proyección WGS 84 15N) de sus vértices suroeste y noreste son $(594378, 1668578)$ y $(595383, 1669533)$, respectivamente (ver Figura 1.1). Contiene un río que se bifurca en dos ramales que interpretamos como el único atractor sobre la región. Consideramos además especies de anfibios (ver Capítulos 3 y 4) como nuestra especie objetivo ya que se sienten atraídos naturalmente por los ríos y su acceso a lugares fuera de ellos depende de la orografía de los ríos de la región.

Dividimos la región geográfica en una área cuadrículada (un lattice)

$$\Lambda = \{1, \dots, M\} \times \{1, \dots, N\},$$

donde M y N son enteros positivos (en nuestro caso específico, $M = 192$ y $N = 202$). Geográficamente, todas las celdas (i, j) en Λ tienen las mismas dimensiones. En nuestro caso, cada celda cuadrada tiene $5m$ de lado. Una *configuración* (o un *estado*) del sistema es una especificación $\omega = (\omega_{i,j})_{(i,j) \in \Lambda}$ tal que $\omega_{(i,j)}$ denota el número de individuos en la celda (i, j) . El *espacio de configuraciones* es el espacio

$$\Omega = \Gamma^\Lambda,$$

donde $\Gamma = \{0, \dots, K\}$ y K es el número máximo de individuos que puede haber en cada celda.

Adicionalmente, definimos la *estructura de vecinos de primer orden* caracterizada por la siguiente propiedad: Dos celdas (r, s) y (i, j) en Λ son vecinos si se cumple la igualdad

$$|r - i| + |s - j| = 1.$$

Denotamos $(r, s) \sim (i, j)$ para las celdas vecinas.

1.5.2. Funciones de energía (Hamiltonianos)

Ahora describimos las funciones de energía que formalizarán matemáticamente las hipótesis que fundamentan la investigación (ver Sección 1.2.1). Recordemos que en la región existe un único atractor: el río.

Hamiltoniano Libre. Denotamos por H_0 el *Hamiltoniano libre*, el cual está descrito

como

$$H_0(\omega) := \sum_{(r,s) \sim (i,j)} (\omega_{r,s} - \omega_{i,j})^2, \quad \omega \in \Omega, \quad (1.24)$$

Observe que H_0 es no negativo y es cero si y sólo si ω es constante, lo que implica que las configuraciones ω que minimizan la energía H_0 son exactamente las que describen el mismo número de individuos en cada celda.

Energía Potencial. Nosotros tomamos al río como el único atractor (y no hay área de repulsión). Denotamos por V la energía potencial que controla la fuerza de atracción ejercida por el atractor, el cual está dado por

$$V_g(\omega) := g \sum_{i,j} d_{i,j}^2 \omega_{i,j}, \quad \omega \in \Omega, \quad (1.25)$$

donde $d_{i,j}$ es la distancia de la celda (i, j) al río.

Está claro que V_g es positivo y que se minimiza en los estados donde no hay individuos fuera del río. Incluimos una *constante de acoplamiento* g que mide la *intensidad* de la fuerza de atracción. Esta constante se seleccionará para que coincida con los datos experimentales (en general, podemos utilizar más variables, pero en nuestro caso solo una constante de acoplamiento es suficiente).

Hamiltoniano del Sistema. El hamiltoniano del sistema es la suma de la energía libre y la energía potencial, y es solo un oscilador armónico con constante de acoplamiento g dado por:

$$H(\omega) := H_0(\omega) + V_g(\omega), \quad \omega \in \Omega. \quad (1.26)$$

1.5.3. Medida de Gibbs

La medida de probabilidad del sistema está dada por la medida de Gibbs-Boltzmann definida por

$$\pi(\omega) := \frac{1}{Z} e^{-\frac{1}{T} H(\omega)}, \quad \omega \in \Omega, \quad (1.27)$$

donde Z es la función de partición (que es solo una constante de normalización) y $T > 0$ es la temperatura del sistema.

1.5.4. Datos recolectados sobre el río

En el caso de la región que estudiamos, el atractor es el río y es aproximadamente homogéneo con respecto a las variables ambientales y geográficas. Basta, por tanto,

realizar trabajo de campo en un pequeño tramo del río. Dado que el río está en promedio $5m$ de ancho, podemos suponer que consta de un conjunto de celdas alineadas. Creemos que, para el trabajo de campo, muestrear 15 celdas ($75m$) es suficiente para obtener la densidad por celda en el río. De hecho, en [4], se informa que se muestrearon 75 metros y se recolectaron diferentes especies de anfibios. Sin embargo, no se informan las densidades que necesitamos. En el trabajo futuro analizaremos diferentes especies de anfibios e informaremos sobre el trabajo de campo adicional. Por ahora es conveniente asumir una densidad de 5 individuos por celda para visualizar las características del modelo (y esto es realista en términos de [4]). Para situaciones reales, se debe realizar un análisis detallado de los aspectos biológicos y adaptaciones específicas del modelo y considerar otras estrategias numéricas para realizar simulaciones óptimas (ver Capítulos 3 y 4).

Estableciendo la densidad de individuos por celda igual a 5, definimos un proceso de Poisson sobre el río que simulará la distribución de la especie sobre el río (ver Sección 2.4.4.2). Este muestreo artificial a lo largo del río lo tomamos como dato inicial. El subconjunto de celdas que corresponden al río será denotado por R . Una vez que se fijan los datos en el río (es decir, sobre R), definimos la medida de probabilidad final del sistema, que es la probabilidad condicional dado que el número de individuos en las celdas que tocan el río es fijo. Denotamos esta medida de probabilidad por π^R . Es bien sabido que las probabilidades condicionales de las medidas de Gibbs también son medidas de Gibbs [91]. Denotamos por H^R el correspondiente Hamiltoniano:

$$\pi^R(\omega) = \frac{1}{Z^R} e^{-\frac{1}{T} H^R(\omega)},$$

donde Z^R es una constante de normalización (función de partición). Aquí la cuadrícula es la cuadrícula original menos todos los rectángulos que tocan el río (tiene 4 componentes conexas). A estos componentes conexas las llamamos Regiones 1, 2, 3 y 4 (ver Figura 1.1). Y las denotaremos por Λ^1 , Λ^2 , Λ^3 y Λ^4 , respectivamente.

1.5.5. Generación del muestreo de Gibbs

Dado que las Regiones 1, 2, 3 y 4 están desconectadas, podemos analizarlos por separado. Aquí nos enfocamos en la Región 4 (el análisis de las otras regiones es el mismo). Recordamos que nuestro objetivo final es encontrar estados cuya energía se acerque al mínimo (lo que equivale a un criterio de máxima verosimilitud). Hay pocos métodos que resuelvan rigurosamente este problema. El método preferido en este texto se llama *annealing* [60, 69], aunque también están disponibles estrategias de pendiente descendente con descensos de temperatura aleatorios [157]. En el caso del *annealing*, las

tasas de convergencia óptimas son accesibles [60, 69]. Desafortunadamente, tales tasas de convergencia son de orden logarítmico, comenzando con una temperatura grande, y la mayoría de las veces intratables. En nuestro caso, el *annealing* puede requerir cálculos de más de 10^{30000} . Las formulaciones abstractas en [60] y [69] son independientes de los estados (o configuraciones) iniciales y del orden de muestreo. Esta generalidad debe pagar el precio de bajas tasas de convergencia. No obstante, la selección inteligente de los estados iniciales, las temperaturas y los órdenes de muestreo pueden producir estados cuyas energías estén muy cerca del mínimo (pero las únicas tecnologías disponibles a este respecto son las numéricas).

Dado que un procedimiento de temperatura descendente no tiene sentido en nuestro caso, tenemos que elegir una temperatura adecuada (sin embargo, usamos un esquema de temperatura descendente para elegir el estado inicial y la temperatura del sistema). Las altas temperaturas implican una alta probabilidad para estados con energías que no se acercan al mínimo de energía. Las bajas temperaturas implican que los estados tendrán una alta probabilidad de quedarse estancados en los mínimos locales (en el sentido de la Definición 1.5.1). La única posibilidad es entonces elegir temperaturas bajas que promuevan estados cuyas energías estén cerca del mínimo global, pero no demasiado bajas para evitar el mínimo local. El muestreo de Gibbs es un método que produce una sucesión de configuraciones (o estados) que tienden (en probabilidad) a los estados más probables y, por tanto, en el régimen de baja temperatura tienden a estados cuyas energías están próximas al mínimo. Sin embargo, si la temperatura es demasiado baja o los estados iniciales no se eligen correctamente, es posible que nuestros cálculos numéricos solo produzcan mínimos locales. Si tenemos suerte, entonces las energías de los estados tienden a la energía mínima, pero la mayoría de las veces se acercan a los mínimos locales. La idea es elegir configuraciones iniciales, métodos de muestreo y temperaturas que promuevan el acercamiento al mínimo global de la energía (nótese que el método de *annealing* no considera estas características). El orden del muestreo es crucial y se define mediante un orden de las celdas de la Región 4. Esto se lleva a cabo de la siguiente manera: Denotamos por m_4 el número de celdas en la Región 4 y definimos una sucesión

$$(\ell_n)_{n \geq 1}$$

la cual satisface

- Para toda n , ℓ_n es una celda de la Región 4, y la sucesión cubre todas las células de esta región.
- Es periódica con periodo m_4 , donde m_4 es el número de celdas en la Región 4.
- Los primeros elementos de la sucesión cubren los vecinos del río, los siguientes

1. INTRODUCCIÓN

elementos cubren los vecinos de los primeros elementos de la sucesión (que no pertenecen a los primeros elementos), y así sucesivamente hasta cubrir todas las celdas de la Región 4.

El muestreo de Gibbs se basa fundamentalmente en una cadena de Markov \mathbf{X}^4 con valores en el espacio de configuraciones (o estados) de la Región 4, la cual está definida mediante las siguientes propiedades:

- $\mathbf{X}^4(0)_{i,j} = a^{(0)}$, para toda (i, j) en la Región 4, donde $a^{(0)}$ es una configuración inicial.
- $\mathbf{X}^4(n)_{i,j} = \mathbf{X}^4(n+1)_{i,j}$, para toda $(i, j) \neq \ell_{n+1}$ en la Región 4.
- Denotando por \mathbb{P}^4 la medida de probabilidad asociada a \mathbf{X}^4 , si $k \in \mathbb{N}$ y si a es una configuración en la Región 4, luego la probabilidad de transición de un paso en el momento n de a a la configuración que es igual a a excepto en la celda ℓ_{n+1} es dado por

$$\mathbb{P}^4(\mathbf{X}^4(n+1)_{\ell_{n+1}} = k | \mathbf{X}^4(n) = a) = \pi^4(\omega_{\ell_{n+1}} = k | \omega_{i,j} = a_{i,j}, \text{ for } (i, j) \neq \ell_{n+1}), \quad (1.28)$$

donde π^4 es la restricción de π^R a la Región 4.

Un muestreo de Gibbs es una realización

$$(\omega^{(n)})_{n \in \mathbb{N}} \quad (1.29)$$

del proceso estocástico \mathbf{X}^4 .

Según el Teorema 2.5.1 (Theorem A [60]), la distribución invariante de esta cadena de Markov es la distribución de π^4 , y el Teorema Ergódico 4.7.1 (Theorem C [61], 5.1.4 Theorem [158]) establece que podemos obtener una muestra aleatoria de la distribución π^4 a través de los promedios de muestreo de Gibbs. En el régimen de baja temperatura, esta muestra estará naturalmente más cerca del estado más probable de la distribución π^4 (el estado de energía mínima). Luego usamos el muestreo de Gibbs para estimar los parámetros del modelo que proponemos. Evitamos situaciones en las que el algoritmo se atasca en estados no óptimos (esta es una desventaja mayoritaria del muestreo de Gibbs). Dichos estados se denominarán mínimos locales de acuerdo con la siguiente definición:

Definición 1.5.1. Dada una configuración a , decimos que a es un **mínimo local** si para cada celda ℓ y cada $k \in \mathbb{N}$

$$H^4(a) \leq H^4(\tilde{a}),$$

donde \tilde{a} es otra configuración tal que $\tilde{a}_{i,j} = a_{i,j}$ for $(i,j) \neq \ell$ y $\tilde{a}_\ell = k$, and H^4 es el Hamiltoniano generador de la distribución π^4 sobre la Región 4.

Se puede encontrar una definición similar en [69], en el contexto del método de *simulated annealing* óptimo. En la Sección 2.8 mostramos en detalle un pequeño modelo en el que el algoritmo de muestreo de Gibbs aplicado ingenuamente queda atrapado en mínimos locales.

1.6. Selección de Parámetros

En la Sección 1.5.5 presentamos la principal herramienta matemática para nuestro modelo y simulaciones, la sucesión de muestreo de Gibbs $(\omega^{(n)})_{n \in \mathbb{N}}$. Esta sucesión es una realización de un proceso estocástico y, por tanto, cambia con cada simulación. Depende de la constante de acoplamiento g y la temperatura T . En esta sección, seleccionamos los parámetros que producen resultados que se asemejan mejor a los de la Sección 2.6. El procedimiento general se describe en la Sección 2.5.5.

1.6.1. Temperaturas auxiliares

Para calibrar nuestro modelo, tomamos un número constante de individuos (que denotamos por $K \in \{1, 2, \dots, 13\}$) en cada celda del río y $g = 0$. Específicamente, el número 13 es el número máximo de individuos por celda en el río (ver Sección 2.4.4.2). Realizamos simulaciones de muestreo de Gibbs con diferentes temperaturas y $m_4 \times 1000000$ iteraciones (y estado inicial cero). Seleccionamos una temperatura T_1 y eso produce un mínimo para la energía en nuestras simulaciones numéricas para las últimas iteraciones. Calculamos la media de los últimos $m_4 \times 10000$ iteraciones y lo utilizamos como estado inicial para nuevas simulaciones de muestreo de Gibbs, bajo diferentes temperaturas T_2 (más pequeña que T_1). Buscamos un intervalo para las temperaturas T_2 que produzca un descenso de energía significativo en las iteraciones de muestreo de Gibbs. Llamamos a este intervalo el *rango* de T_2 .

1.6.2. Estimación de g , T y el estado inicial

La estimación de los parámetros depende de los datos recolectados en el trabajo de campo en la Región 4. Una vez seleccionada la temperatura T_1 , que consideramos fija, aumentamos lentamente el valor de g (partiendo de $g = 0$) y simulamos muestreos de Gibbs para cada valor de g . Usamos estas simulaciones (y el Teorema Ergódico 2.5.2) para estimar los valores esperados del número de individuos en las celdas donde se realizó el trabajo de campo. Seleccionamos la constante de acoplamiento que mejor se aproxima a los datos recopilados. Esto produce una primera aproximación g_0 del parámetro g , que no es lo suficientemente precisa pero lo usamos como parámetro de referencia.

Por cada g en un vecindario de g_0 y cada T_2 en su rango, calculamos $m_4 \times 1000000$ iteraciones de muestreo de Gibbs (con estado inicial cero, temperatura T_1 y constante de acoplamiento g) y tomamos el promedio de los últimos $m_4 \times 10000$ estados para producir un estado inicial. Con este estado inicial, calculamos $m_4 \times 1000000$ iteraciones adicionales (con temperatura T_2 y constante de acoplamiento g). Calculamos el promedio de los últimos $m_4 \times 10000$ estados y comparamos el resultado con los datos del trabajo de campo. Elegimos el valor g_{T_2} que mejor coincida con los datos recopilados. A continuación, ejecutamos $m_4 \times 2000000$ iteraciones como describimos anteriormente con diferentes valores de T_2 (tomando g_{T_2} para la constante de acoplamiento) y elegimos el número T_2 que produce la mejor tasa de decaimiento para el energías. Llamamos T a este número y $g = g_T$, la temperatura y la constante de acoplamiento del sistema. El estado inicial del sistema es el promedio de las últimas $m_4 \times 10000$ iteraciones de $m_4 \times 1000000$ iteraciones de muestreo de Gibbs con estado inicial cero, temperatura T_1 y constante de acoplamiento g .

1.7. El problema de los mínimos locales y la intratabilidad

En el algoritmo *simulated annealing*, el programa de descenso (1.21) se elige de manera que dependa del número de iteraciones k en el esquema de muestreo de Gibbs, y en [60, 69, 157] está probado que el límite

$$T(k) \geq \frac{c}{\log(1+k)}, \quad k \geq 0,$$

debe cumplirse, para c constante lo suficientemente grande. Esta constante se estima en [60, 69, 157]), y en nuestro caso, conduce a un problema insoluble. Entonces, un descenso de la temperatura no tiene sentido y simplemente lo fijamos como un número

T pequeño por determinar. Esto conduce a un problema importante: la mayoría de las veces uno se queda atascado en los mínimos locales. Explicamos cómo abordamos este problema en las Secciones 2.2 y 1.5. En Sección 2.8 damos un ejemplo simple en el que el algoritmo se atasca en un mínimo local, y este mínimo es sustancialmente diferente del mínimo global.

1.8. Aplicaciones

Usamos el modelo anteriormente descrito para predecir la abundancia y distribución de dos especies de anfibios endémicas de la región de estudio: *Plectrohyla Sagorum* y *Plectrohyla Guatemalaensis*. Para la aplicación del modelos seleccionamos una subregión rectangular dentro de la Reserva natural del Volcán Tacaná (ver Figura 1.1). Para la primera especie, modelamos el comportamiento global de la población, en la segunda nos centramos más bien en la densidad poblacional.



(a) Reserva Natural del Volcán Tacaná (cuadro azul)



(b) Región rectangular de estudio, el río se muestra en azul.

Figura 1.1: Reserva Natural del Volcán Tacaná, cuadro azul en (a), y región de estudio (b). Nótese en (b) que la región rectangular de estudio está dividida en 4 subregiones geográficas disconexas por las corrientes del río (azul).

Caso *Plectrohyla Sagorum* (ver Capítulo 3).

Usamos nuestro modelo para predecir la dinámica poblacional de la especie durante cuatro estaciones distintas del año (tomamos el año 2018, en el cual se recopilaban los datos de campo [4]). Nuestro método explota la idea de que la población es atraída por los cuerpos de agua. Al mismo tiempo, la población tiende a esparcirse por todos lados si las condiciones de humedad son propicias. El equilibrio entre moverse y ser atraído determina su distribución y se estima a partir de los datos recopilados. Durante

invierno (la estación seca), la población en su totalidad está adherida a cuerpos de agua, y asumimos que todos los individuos están restringidos a un radio muy cercano a éstos. Fijamos entonces el número de individuos de la región a partir de los datos recopilados durante invierno, y reconstruimos la distribución y abundancia de individuos para el resto de las estaciones.

En la Figura 3.14 mostramos una representación gráfica de los valores esperados del número de individuos en cada celda, usando un punto para representar a cada individuo. Vemos claramente que los individuos se concentran en el río en invierno (la estación seca) y se dispersan a medida que llega la lluvia a la región (primavera). En verano, cuando la precipitación alcanza su máximo, la dispersión de los individuos es máxima. En otoño, los individuos vuelven a acercarse al río y su distribución es similar a la de primavera.

Caso *Plectrohyla Guatemalensis* (ver Capítulo 4).

Los datos recopilados para esta especie son sumamente escasos, por lo que optamos por un enfoque distinto. Aquí, nuestras variables aleatorias describen promedios anuales de valores esperados del número de individuos en lugar de individuos. Describimos los resultados obtenidos mediante un mapa de calor (ver Figura 4.5). Utilizamos los 7 colores del arco iris: rojo, naranja, amarillo, verde, cian, azul y violeta (este orden corresponde a una energía ascendente de los colores). Cada celda (i, j) está coloreada con uno de estos colores dependiendo de los valores de $\omega_{i,j}$ de tal manera que el color es constante en intervalos de la forma $\omega_{i,j} \in (\frac{s}{7}K, \frac{s+1}{7}K]$, $s \in \{0, \dots, 6\}$ y la energía del color aumenta a medida que aumenta $\omega_{i,j}$, donde K es la mayor densidad por celda registrada. Es posible observar que el modelo arroja una distribución de la especie “por bandas” concéntricas sobre la fuente de agua.

1.9. Teoría de operadores y de dispersión en la recta real discretizada

1.9.1. Hamiltonianos

En [15] se estudia la teoría de la dispersión para la matriz unidimensional de los Hamiltonianos de Schrödinger en \mathbb{Z} de la forma

$$H = H_0 + V, \tag{1.30}$$

donde H_0 es el operador discreto de Laplace definido sobre el espacio de Hilbert $\ell^2(\mathbb{Z}, \mathbb{C}^L)$ dado por

$$(H_0\phi)(n) := \phi(n+1) + \phi(n-1), \quad \phi \in \ell^2(\mathbb{Z}, \mathbb{C}^L), \quad (1.31)$$

y $V \in \mathcal{M}_{L \times L}(\mathbb{C})^{\mathbb{Z}}$ ($\mathcal{M}_{L \times L}(\mathbb{C})$ es el espacio de las matrices de $L \times L$ con coeficientes en \mathbb{C}) es un operador de multiplicación matricial con primer momento finito, definido por

$$(V\Psi)(n) := V(n)\Psi(n), \quad \Psi \in \ell^2(\mathbb{Z}, \mathbb{C}^L). \quad (1.32)$$

Suponemos además que $V(n) = V(n)^*$, $n \in \mathbb{Z}$ y

$$\sum_{n \in \mathbb{Z}} \|nV(n)\| < \infty. \quad (1.33)$$

En [16] se estudian los operadores de la forma (1.30), pero con V de soporte compacto. En [16] se sigue un enfoque analítico diferente, a saber, las soluciones de la ecuación de valor propio de H se calculan en términos de la matriz de transferencia, mientras que aquí se utiliza la ecuación de Volterra. Esto implica una gran diferencia y prácticamente no hay intersección entre las pruebas en esta investigación y las pruebas en [16]. Además, un potencial con soporte compacto admite continuaciones meromórficas de la matriz de dispersión en todo el plano complejo, algo que no es posible en el marco actual.

En [94] se aborda el caso escalar continuo y en [79] se ha tratado el caso escalar discreto. Para los operadores de Schrödinger con valores matriciales continuos, hay trabajos de Aktosun, Klaus, Van Der Mee y Weder [7, 8, 9]. Es importante enfatizar que nuestra prueba del límite de la matriz de dispersión es más corta y está estrechamente ligada a un tratamiento conceptual del Wronskiano. En realidad, no usamos la descomposición de Jordan como en [7]. Los estados ligados para la versión de medio espacio de la ecuación de Schrodinger de matriz discreta con potenciales soportados de forma compacta se han analizado en [6].

Este modelo se usa ampliamente en el contexto de los fenómenos de baja energía en la física del estado sólido. Además, H es un operador tridiagonal, también llamado operador Jacobi, que es el análogo discreto de un operador Sturm-Liouville. Su análisis está conectado a polinomios ortogonales (matriciales) y, a través de su teoría espectral, a medidas con valores matriciales en \mathbb{R} . Como es usual, estudiamos autovalores, además, no solo se abordan las energías reales, sino que también estudian los vectores propios generalizados correspondientes a las energías complejas. Usamos los mismos símbolos

1. INTRODUCCIÓN

H y H_0 para denotar los operadores definidos (con las mismas expresiones que arriba) ya sea sobre $(\mathcal{M}_{L \times L})^{\mathbb{Z}}$ o $(\mathbb{C}^L)^{\mathbb{Z}}$ y parametrizamos los autovalores en la forma

$$E = z + 1/z, \quad z \in \mathbb{C} \setminus \{0\}. \quad (1.34)$$

Entonces las ecuaciones de autovalores tienen la forma

$$Hu = Eu, \quad (1.35)$$

$$H_0u = Eu, \quad (1.36)$$

donde $u \in (\mathcal{M}_{L \times L})^{\mathbb{Z}}$. De particular importancia son las soluciones de Jost, las cuales son soluciones de (5.12) con a la frontera asintóticos en $-\infty$ o ∞ dadas por soluciones libres (esto es, soluciones a (5.13)).

1.9.2. Soluciones de Jost

Las soluciones de Jost, son soluciones del sistema (5.12) que se comportan como ondas libres (ondas planas) en el infinito.

Definición 1.9.1 (Soluciones de Jost). *Para toda $z \in \mathbb{D} \setminus \{0\}$, donde \mathbb{D} es el disco unitario abierto complejo, denotamos por u_+^z y $u_-^{1/z}$ la soluciones $\mathcal{M}_{L \times L}$ -valuadas de*

$$Hu_+^z = Eu_+^z, \quad Hu_-^{1/z} = Eu_-^{1/z}, \quad E = z + 1/z, \quad (1.37)$$

tales que, cuando $n \rightarrow +\infty$ y $n \rightarrow -\infty$, respectivamente,

$$u_+^z(n) = z^n(\mathbf{1} + o(1)), \quad u_-^{1/z}(n) = z^{-n}(\mathbf{1} + o(1)). \quad (1.38)$$

Adicionalmente, para $z = 1$, denotamos por v_{\pm}^z la soluciones $\mathcal{M}_{L \times L}$ -valuadas de

$$Hv_{\pm}^z = Ev_{\pm}^z, \quad E = 2, \quad (1.39)$$

tales que cuando $n \rightarrow \pm\infty$,

$$v_{\pm}^z(n) = n(\mathbf{1} + o(1)). \quad (1.40)$$

Debido al comportamiento asintótico de las soluciones de Jost (ver Ecuación (1.38)), las columnas de la matriz $(u_{\pm}^z, u_{\pm}^{1/z})$ son linealmente independientes para $z \in \mathbb{S}^1 \setminus \{-1, 1\}$

(donde \mathbb{S}^1 es la esfera unitaria compleja) y por tanto forman una base para el espacio de las soluciones. Lo mismo es válido para (u_{\pm}^1, v_{\pm}^1) . Esto implica que existen matrices $M_{\pm}^z, N_{\pm}^z \in \mathcal{M}_{L \times L}$ tales que

$$u_+^z = u_-^z M_+^z + u_-^{1/z} N_+^z, \quad u_-^{1/z} = u_+^z N_-^z + u_+^{1/z} M_-^z. \quad (1.41)$$

Además, probamos que las matrices tienen una continuación meromórfica a \mathbb{D} . Asumiendo que M_{\pm}^z son invertibles, podemos reescribir estas ecuaciones como

$$u_+^z T_+^z = u_-^z - u_-^{1/z} R_+^z, \quad u_-^{1/z} T_-^z = u_+^{1/z} - u_+^z R_-^z, \quad (1.42)$$

donde

$$T_{\pm}^z = (M_{\pm}^z)^{-1}, \quad R_{\pm}^z = -N_{\pm}^z (M_{\pm}^z)^{-1}, \quad (1.43)$$

son los coeficientes de transmisión y reflexión, respectivamente. La interpretación de (1.42) en el caso que $z = e^{ik} \notin \{-1, 1\}$, corresponde a una onda desplazándose a la derecha: la onda entrante u_-^z produce la onda saliente $u_+^z T_+^z$ viajando hacia la derecha (es decir, una onda transmitida) y la onda saliente $u_-^{1/z} R_+^z$ viajando a la izquierda (es decir, una onda reflejada).

1.9.3. Resultados principales

Teorema 1.9.1. *Existe una vecindad de 1 tal que para todo z en dicha vecindad con $z \in \overline{\mathbb{D}} \setminus \{1\}$, las matrices M_{\pm}^z son invertibles. Además, los límites*

$$T_{\pm}^1 := \lim_{z \rightarrow 1} T_{\pm}^z \quad (1.44)$$

existen, donde los límites son tomados en $\overline{\mathbb{D}}$. Los núcleos e imágenes de T_{\pm}^1 pueden calcularse explícitamente y están estrechamente conectados a estados semi-acotados.

Los límites

$$R_{\pm}^1 := \lim_{z \rightarrow 1} R_{\pm}^z \quad (1.45)$$

de los coeficientes de reflexión R_{\pm}^z existen, donde los límites son tomados en \mathbb{S}^1 , y tienen expresiones explícitas. Los núcleos e imágenes de $\mathbf{1} - R_{\pm}^1$ pueden calcularse explícitamente.

Un modelo numérico para la descripción de la distribución y abundancia de individuos

2.1. Introducción

En este capítulo explicamos con precisión el modelo matemático el método numérico que introducimos en la investigación. Como hemos explicado en la Sección 1.5.1, evitamos una exposición abstracta y en cambio, explicamos nuestro método usando un ejemplo: elegimos una región específica ubicada en un bosque nuboso de Chiapas, México, ver Figura 1.1). En los Capítulos 3 y 4 usaremos y adaptaremos nuestro método para estudiar especies específicas de anfibios del género *Plectrohyla* en la misma región que estudiamos aquí. Ambas especies requieren algunas características adicionales para su análisis, ya que se encuentran en peligro de extinción (los datos recolectados son extremadamente pobres). Los resultados en este capítulo están reportados en [19].

2.2. Hipótesis ecológicas

Elegimos una región específica (con una orografía complicada, un cañón) y una especie de anfibio endémica de la región. Esto es conveniente porque en este caso la interpretación de los conceptos matemáticos del presente trabajo es sencilla, sobre todo si tenemos en cuenta a los anfibios que son atraídos naturalmente por los ríos y su acceso a lugares fuera de ellos depende de la orografía de la región. En este capítulo, abordamos la parte matemática de nuestro programa y nos enfocamos en definiciones y análisis

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

matemáticos precisos y detallados, y todos los aspectos numéricos. En los Capítulos 3 y 4 usaremos y adaptaremos nuestro método para estudiar especies específicas de anfibios del género *Plectrohyla* en la región que aquí estudiamos ([17, 20]). El modelo que presentamos también será usado en una región diferente, que es más accesible para el trabajo de campo [18]. En concreto, en el Capítulo 3 (y en [20]) estudiamos la dinámica poblacional (cómo cambia la distribución en las diferentes estaciones del año) de *Plectrohyla Sagorum*; y en el Capítulo 4 (y en [17]) abordamos el problema de reconstruir la densidad poblacional una especie en peligro crítico llamada *Plectrohyla Guatemalensis* (esto es complicado porque los datos recolectados son muy pobres), y en [18] aprovecharemos los datos disponibles más precisos para usar inteligencia artificial (redes neuronales) para estimar los potenciales.

La idea clave de nuestro modelo es notar que hay variables ambientales que atraen o repelen a los individuos de una especie biológica. Por ejemplo, en el caso de los anfibios, los ríos juegan el papel de atractores. En ellos se alcanza la mayor densidad de individuos, y la densidad disminuye a medida que aumenta la distancia (y la elevación) a los ríos. Nuestro enfoque se basa en dos hipótesis básicas y razonables:

- H1** Si no existe una variable ambiental específica que atraiga o repele a los individuos, entonces se esparcen por todos lados, y es igualmente probable encontrarlos en cualquier lugar de la región que estudiamos.
- H2** En el caso de que podamos identificar atractores (o regiones de repulsión), entonces los individuos tienden a acercarse a ellos (o alejarse de las regiones de repulsión) pero al mismo tiempo, quieren moverse y extenderse. Estos dos comportamientos compiten entre sí y la distribución real de los individuos está determinada por la fuerza de la atracción (o la repulsión).

Las hipótesis H1 y H2 se encuentran entre los conceptos de la física que son los más utilizados desde los orígenes de la física moderna. La propagación descrita en H1 podría entenderse como energía cinética y las fuerzas de atracción o repulsión descritas en H2 se conocen como energía potencial.

Probablemente, la forma más natural de abordar el problema es mediante el uso de la mecánica estadística clásica, porque ha sido utilizada (conscientemente o no) por los estadísticos desde los orígenes de la geoestadística (ver Sección 2.3). Sin embargo, los conceptos de energías cinética y potencial no se han utilizado en geoestadística, porque sus modelos normalmente se limitan al alcance de las distribuciones habituales (más prominentemente de la familia exponencial). La idea clave es que la presencia de individuos en algún lugar depende de los individuos del entorno. Por lo tanto, las variables aleatorias asociadas a las posiciones de los individuos no pueden ser independientes y

generan un campo aleatorio de Markov o, de manera equivalente, una medida de Gibbs [60]. Esto establece las características generales del modelo matemático y los datos de muestreo se utilizan luego para estimar los parámetros libres del modelo.

Nuestro enfoque difiere significativamente de los métodos bayesianos habituales de geoestadística porque intentamos encontrar un proceso estocástico (un campo aleatorio de Markov) en lugar de ajustar un modelo estadístico. En primer lugar, estamos pensando en una situación en la que no tenemos una cantidad suficiente de datos muestrales lo cual es esencial en la inferencia estadística (específicamente, no tenemos covariables medibles *in situ* que brinden la posibilidad de calcular medidas de tendencia central y tasas de convergencia). Esta es una suposición razonable, ya que nuestro método se aplicará en trabajos futuros a modelos de especies en peligro de extinción para las que es difícil recopilar datos. Recordamos que las técnicas habituales de estimación bayesiana también utilizan una simulación muestral de distribuciones propuestas a priori para cada uno de los parámetros con el fin de obtener sus estimaciones. En cambio, nuestro modelo intenta reconstruir el mapa de abundancia y distribución de la especie sin la necesidad de tales suposiciones. En segundo lugar, nuestro modelo propone directamente una distribución global, a diferencia de los métodos habituales de geoestadística, en los que se proponen distribuciones condicionales locales para derivar (a través del teorema de Hammersley-Clifford) una distribución global (que es una medida de Gibbs).

Ciertamente, el problema de la estimación de la distribución real de las especies biológicas ha sido abordado desde la perspectiva de la geoestadística, en los que la información sobre covariables (cobertura vegetal, precipitación, isothermalidad) es fundamental. Sin embargo, esta información sobre covariables y el volumen de datos recolectados son prácticamente inexistentes en las regiones de nuestro interés, que son el bosque nuboso de Chiapas y el bosque lluvioso de Oaxaca en México. El trabajo de campo requiere mucho trabajo debido a que las condiciones geográficas son inaccesibles desde el entorno, por lo que la calibración de nuestro parámetro principal (la constante de acoplamiento, que mide la fuerza de atracción hacia las fuentes de agua) no se puede realizar con las técnicas propuestas por la geoestadística.

Las nuevas ideas introducidas en nuestro modelo computacional son las siguientes:

- (a) No estimamos localmente, sino globalmente: en lugar de enfocarnos en probabilidades condicionales (de un conjunto limitado de distribuciones), abordamos directamente la medida conjunta del sistema, la cual es una medida de Gibbs.
- (b) El modelo se concreta utilizando la intuición física de los fenómenos a describir. En particular, elegimos un Hamiltoniano extraído de la mecánica cuántica (que es el generador de una medida de Gibbs) que representa el comportamiento de los individuos.

- (c) Nuestro esquema numérico se basa en el algoritmo de muestreo de Gibbs [60]. Permite un control global del sistema que se identifica fácilmente a partir de nuestras simulaciones. En el caso libre, obtenemos una aproximación muy precisa del estimador de verosimilitud (mínimo global del Hamiltoniano libre). Es importante elegir el estado inicial con cuidado, ya que no siempre es posible alcanzar el mínimo global con estados iniciales arbitrarios, incluso en el caso libre (Sección 2.8).

2.3. Antecedentes históricos

La geoestadística es en la actualidad el principal enfoque, aunque no el único, de los métodos de investigación ecológica de la distribución espacial de las especies (para una descripción de los métodos clásicos y modernos de análisis espacial, véase [22, 29, 48, 126, 127]). La geoestadística se originó a partir del estudio de la distribución de ciertos minerales en los yacimientos mineros [13, 35, 42, 84, 103, 104, 105, 106, 107, 148]. La idea principal de los modelos geoestadísticos es que la densidad (o la simple presencia o ausencia) de minerales (o individuos de una población biológica, en el contexto de la ecología) sobre determinada posición geográfica, normalmente depende de dicha densidad en las posiciones geográficas circundantes. Lo que implica que los modelos estocásticos de la geoestadística no suponen independencia de las variables que intervienen en el sistema, como es el caso de los métodos típicos de los modelos jerárquicos generalizados del análisis de regresión lineal (ver [38, 39, 100, 128]). Tales modelos lineales son populares en la ecología porque utilizan covariables que describen las características espaciales de las regiones de estudio. Sin embargo, estos modelos no consideran la dependencia espacial intrínseca de todo el sistema, en contraste con los métodos de la geoestadística (ver [14, 74, 80, 81, 159]).

El primer paso importante en el desarrollo de la geoestadística en el contexto de la ecología es el trabajo de Besag [25, 26], donde el lema de Brook [31] es un ingrediente esencial. Besag utiliza campos aleatorios de Markov (introducidos originalmente por F. Spitzer [136] y C. Preston [123]), para caracterizar mediante el teorema de Hammersley-Clifford [26] (según Besag, demostrado en 1971 por Hammersley y Clifford en un trabajo inédito en 1971) los llamados automodelos, o modelos autorregresivos, diseñados específicamente para describir datos espacialmente autocorrelacionados, basados en relaciones del vecindad (ver [22, 62, 67, 135, 139, 154, 155]). La idea es que las propiedades estadísticas en ciertas posiciones geográficas se determinen una vez que se conocen estas mismas propiedades en las posiciones circundantes.

Esencialmente, el teorema de Hammersley-Clifford afirma que la distribución con-

junta de un campo aleatorio de Markov se puede factorizar utilizando distribuciones condicionales, lo que equivale al hecho de que existe una correspondencia uno a uno entre las medidas de Gibbs y los campos aleatorios de Markov [67, 92]. Desde el punto de vista de los automodelos de Besag, estas probabilidades condicionales están dadas por las distribuciones habituales de la estadística clásica, generalmente asociadas con la familia exponencial. La ventaja del enfoque de Besag es que en lugar de construir estimadores de máxima verosimilitud a partir de la distribución conjunta (que generalmente es muy difícil), propuso construirlos solo para las probabilidades condicionales (que normalmente es una tarea simple). Estos estimadores se denominan pseudo-verosimilitud, pero desafortunadamente, no coinciden con los estimadores de verosimilitud y pueden diferir sustancialmente. La hipótesis del enfoque de pseudo-verosimilitud es que los eventos estadísticamente significativos deben maximizar las probabilidades condicionales. Sin embargo, las complejas dependencias en el modelo de conjunto completo podrían jugar un papel importante (ver por ejemplo [162]). Además, la eficiencia de los métodos de máxima pseudo-verosimilitud depende de la fuerza de la dependencia espacial y es comparable a la eficiencia de las estimaciones de máxima verosimilitud solo si la dependencia espacial no es demasiado fuerte [68, 159].

Los métodos de restauración de imágenes (recuperar escenas reales de imágenes degradadas) se basan en la misma idea, utilizando campos aleatorios de Markov. Un trabajo fundamental sobre este tema fue el trabajo de Geman-Geman [61], quien remarcó que los campos aleatorios de Markov son equivalentes a las medidas de Gibbs, cuyo origen se remonta a los trabajos sobre termodinámica de L. Boltzmann [27] y J. W. Gibbs [65]. A diferencia de Besag y otros, Geman-Geman señaló que la distribución conjunta completa de un campo aleatorio de Markov toma una forma simple como una medida de Gibbs (teorema de Hammersley-Clifford). Proporcionaron un algoritmo computacional matemáticamente robusto, llamado muestreo de Gibbs. Esto permitió implementar la maquinaria de mecánica estadística y, en particular, métodos numéricos como Monte Carlo Markov Chain y Metropolis-Hastings. Geman-Geman entendió que la estimación estadística equivale a encontrar las configuraciones más probables con respecto a las medidas de Gibbs. También aplicaron el concepto de *simulated annealing* (introducido primero en [93]) para minimizar la energía utilizando un esquema donde la temperatura desciende lentamente. La estimación estadística se presenta también en términos de un problema de minimización. Geman-Geman proporcionó una de las primeras pruebas matemáticas de convergencia del método Metropolis que introducen (muestreo de Gibbs). Demostraron que el *simulated annealing* conduce a una energía mínima. Posteriormente, en [69], se demostraron las tasas óptimas de convergencia. En [129] y [158] se pueden encontrar trabajos más recientes sobre las tasas de convergencia de Monte Carlo.

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

Los trabajos de Besag y Geman-Geman introdujeron la base de la geoestadística en el contexto de la ecología, en términos de campos aleatorios de Markov y sus métodos, o modificaciones de los mismos, se utilizan con mucha frecuencia en el contexto de la ecología en la actualidad, en lugar de los métodos tradicionales de análisis de regresión generalizado (para ejemplos recientes, véase [28, 37, 71, 102, 163, 164]). Un enfoque bayesiano se utiliza en [47, 75, 162]. Sin embargo, en términos computacionales, la inferencia bayesiana puede llevar las simulaciones mucho más tiempo [162].

2.4. Modelo Matemático

2.4.1. La cuadrícula, la estructura de vecinos de primer orden y las configuraciones (o estados)

Como mencionamos anteriormente en la Sección 1.5.1, ofrecemos una explicación del modelo sobre un región específica comprendida en los alrededores del Volcán Tacaná, en el estado de Chiapas, México (ver Figura 1.1). Se trata de un área rectangular con lados de longitud $960m$ y $1010m$, cuya ubicación de sus vértices suroeste y noreste son, en coordenadas UTM bajo la proyección WGS 84 15N, $(594378, 1668578)$ y $(595383, 1669533)$, respectivamente. Está caracterizada por la presencia de un río que se bifurca en dos ramales principales, el cual suponemos como único atractor de la región, y la presencia de cañones (en los caudales de las zonas más altas del río). Dividimos la región en una cuadrícula (un *lattice*) compuesta de celdas cuadradales de $5m$ de longitud cada una (lo que resulta en un *lattice* rectangular de tamaño 192×202 celdas), ver Figura 2.1.

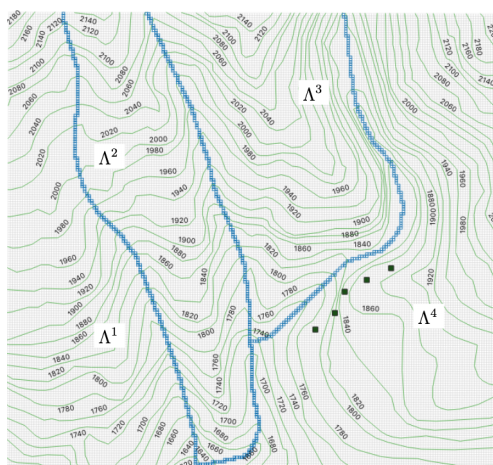


Figura 2.1: Regiones Λ^1 , Λ^2 , Λ^3 , Λ^4 y el Río R (azul)

Denotamos esta cuadrícula por

$$\Lambda := \{1, \dots, M\} \times \{1, \dots, N\},$$

donde M y N son números enteros positivos (en nuestro estudio, $M = 192$ y $N = 202$). Denotamos por K es el número máximo de individuos observados en cada celda de los datos recolectados (en nuestro caso asumimos que los datos fueron recolectados de todas las celdas que tocan el río) y obtenemos que

$$K = 13$$

Denotamos por Γ o bien el conjunto finito $\{0, 1, \dots, K\}$ o el intervalo $[0, K]$. El conjunto $\{0, 1, \dots, K\}$ representa el número posible de individuos en las celdas, y los elementos de $[0, K]$ son interpretados como los valores esperados del número de individuos en cada celda.

El espacio de estados (o configuraciones) se denota por

$$\Omega := \Gamma^\Lambda.$$

Los elementos de Ω se denominan estados o configuraciones y generalmente se denotan por

$$\Omega \ni \omega : \Lambda \rightarrow \Gamma,$$

e identificamos

$$\omega((i, j)) \equiv \omega_{i,j}$$

que es el número de individuos de la especie objetivo en la celda $(i, j) \in \Lambda$ (o el valor esperado del número de individuos en el caso de que $\Gamma = [0, K]$).

La *estructura de vecinos de primer orden* que usamos está dada por lo siguiente: Para cada celda $(i, j) \in \Lambda$, denotamos por

$$N(i, j) := \{(i-1, j), (i+1, j), (i, j-1), (i, j+1)\} \cap \Lambda$$

el conjunto vecinos de primer orden de (i, j) . Además, el símbolo

$$(i, j) \sim (l, m)$$

denota que las celdas (i, j) y (l, m) son vecinas, es decir, $(l, m) \in N(i, j)$, ver [Figura 2.2](#).

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

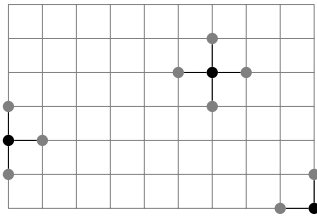


Figura 2.2: Las celdas vecinas (punto gris) de una celda (i, j) (punto negro).

En el contexto de la teoría de grafos, las estructuras de los vecinos se definen en términos de cliques. En el caso de los vecinos de primer orden que describimos anteriormente, el conjunto de cliques es

$$\mathcal{C} := \left\{ \{(r, s), (r, s + 1)\} \mid (r, s), (r, s + 1) \in \Lambda \right\} \cup \left\{ \{(i, j), (i + 1, j)\} \mid (i, j), (i + 1, j) \in \Lambda \right\}. \quad (2.1)$$

Por lo tanto, para cada $(i, j) \in \Lambda$,

$$N(i, j) = \{(l, m) \in \Lambda : \{(i, j), (l, m)\} \in \mathcal{C}\}. \quad (2.2)$$

2.4.2. Hamiltonianos (Funciones de energía)

2.4.2.1. El Hamiltoniano Libre

El Hamiltoniano Libre está dado por

$$H_0(\omega) = \sum_{\{\ell, j\} \in \mathcal{C}} (w_\ell - w_j)^2, \quad \omega \in \Omega,$$

donde \mathcal{C} es la clase de cliques (2.1). H_0 es positivo y se anula solo si ω es constante, de modo que las configuraciones ω que minimizan la energía H_0 son exactamente los que tienen el mismo número de individuos en cada celda. Entonces H_0 es una formalización matemática de la hipótesis ecológica H1 enunciada en la Sección 2.2.

2.4.2.2. Energía Potencial

La Energía Potencial está dada por

$$V_g(\omega) := g \sum_{\ell \in \Lambda} d_\ell^2 \omega_\ell, \quad \omega \in \Omega,$$

donde d_ℓ es la distancia de la celda ℓ al río definida como por

$$d_\ell = M_\ell + A_\ell$$

donde M_ℓ es la distancia horizontal desde la celda ℓ hasta el río, es decir, la distancia mínima desde el centro de la celda ℓ hasta los puntos del río, y A_ℓ es la distancia vertical desde el centro de la celda ℓ hasta el punto del río que minimiza la distancia horizontal. La información geográfica se obtiene a partir de datos proporcionados por INEGI [82].

En nuestro modelo de la región, el río es el único atractor (y no hay área de repulsión). Observe que V_g es mínima en las configuraciones ω en las que no hay individuos fuera del río. La energía potencial V_g describe qué tan rápido y de qué manera decae la presencia de individuos a medida que aumenta la distancia al río, por lo que es una formalización matemática de la hipótesis ecológica H2 enunciada en la Sección 2.2.

Esta es la parte del modelo que debe reconstruirse a partir de los datos del trabajo de campo. La precisión de la reconstrucción depende de la cantidad de datos disponibles. Para simplificar el modelo, e inspirados por el oscilador armónico, aquí asumimos una tasa cuadrática con fuerza determinada por una constante de acoplamiento g , el cual es el parámetro a estimar.

2.4.2.3. Hamiltoniano Completo

El Hamiltoniano del sistema es la suma de la energía libre y la energía potencial, y se asemeja a un oscilador armónico con constante de acoplamiento g :

$$H(\omega) := H_0(\omega) + V_g(\omega), \quad \omega \in \Omega. \quad (2.3)$$

2.4.3. Medida de probabilidad (medida de Gibbs) y el campo aleatorio de Markov

2.4.3.1. Medida de probabilidad (medida de Gibbs)

El siguiente paso es asignar una distribución de probabilidad en el espacio de configuraciones Ω . Esta distribución está dada por una distribución de Gibbs $\pi : \Omega \rightarrow [0, 1]$ que depende del hamiltoniano y de la temperatura T . Se define por

$$\pi_{T,g}(\omega) \equiv \pi(\omega) := \frac{1}{Z} \exp \left\{ -\frac{1}{T} H(\omega) \right\}, \quad \omega \in \Omega,$$

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

donde H es la energía (Hamiltoniano) definido previamente en (2.3). La constante Z es un factor de normalización, dado por

$$Z = \sum_{\omega \in \Omega} \exp \left\{ -\frac{1}{T} H(\omega) \right\}.$$

2.4.3.2. Campo aleatorio de Markov

Para cada $\nu \subset \Lambda$ y $p \in \nu$, definimos la variable aleatoria $W_p : \Gamma^\nu \rightarrow \mathbb{R}$ por

$$W_p(\omega) = \omega_p.$$

Para cada $\Theta \subset \nu$, establecemos

$$W_\Theta := \{W_p : p \in \Theta\}.$$

Omitimos el símbolo ν en nuestra notación. Observamos que las variables aleatorias W_p definen un campo aleatorio de Markov y W_p es una cantidad aleatoria que representa el número de individuos en la celda p .

2.4.4. Inclusión de los datos del río (el atractor) en el modelo: los hamiltonianos finales

2.4.4.1. Retícula del río y componentes conexas: regiones

Denotamos por $R \subset \Lambda$ la cuadrícula del río, que consiste en todas las celdas en Λ que tocan el río. La cuadrícula del río R divide la cuadrícula Λ en cuatro componentes conexas, que denotamos por $\Lambda^1, \Lambda^2, \Lambda^3$ y Λ^4 . A estos componentes los llamamos Regiones 1, 2, 3 y 4, respectivamente. Tenemos que

$$\Lambda = \bigcup_{i=1}^4 \Lambda^i \cup R.$$

Dado que las Regiones 1, ..., 4 son desconexas, los sistemas correspondientes están desacoplados y por esta razón cada región puede estudiarse por separado. Definimos

$$\Lambda^R := \Lambda \setminus R \quad \text{y} \quad \Omega^R := \Gamma^{\Lambda^R}.$$

Además

$$\Omega^i := \Gamma^{\Lambda^i}, \quad i \in \{1, 2, 3, 4\}.$$

Denotamos por

$$\omega^R : R \rightarrow \Gamma,$$

a las observaciones (simuladas) sobre el río. Entonces, ω^R es una muestra fija y lo tomamos como datos recopilados.

2.4.4.2. Simulación Poisson

Como ya se ha mencionado en la Sección 1.5.4, asumimos que la densidad de individuos por celda es 5. Por lo tanto, usamos un proceso de Poisson unidimensional homogéneo, con intensidad

$$\hat{\lambda} = 1,$$

que es el número promedio de individuos por metro. En este escenario, las observaciones de los individuos son los eventos del proceso y los tiempos corresponden a las distancias (a lo largo del río) donde ocurren los eventos. Simulamos el proceso de Poisson usando el teorema del intervalo [98], del cual obtenemos que la distribución conjunta condicional de los tiempos de espera dado un número fijo J de individuos (eventos) hasta el tiempo (distancia) t es igual a la distribución del estadístico de orden J -ésimo de J variables aleatorias independientes, distribuidas de manera idéntica con distribución uniforme sobre $(0, t)$. De esta forma, generamos artificialmente la abundancia de la especie en el río de la siguiente manera:

1. Elegimos una muestra aleatoria fija de una distribución unidimensional de Poisson con intensidad $\hat{\lambda}L$, donde L es la longitud total del río. Denotamos por J el número correspondiente de eventos, es decir, el número de individuos en el río.
2. Generamos una muestra aleatoria ordenada de tamaño J a partir de una distribución uniforme en $(0, L)$.

De nuestras simulaciones de Poisson, seleccionamos un número máximo de individuos en cada celda dado por

$$K = \max\{\omega_p^R : p \in R\} = 13.$$

Elegimos una simulación con este máximo después de haber realizado 1000 simulaciones Poisson y comprobar que esta era la situación más común (alrededor del 33% de las simulaciones realizadas).

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

2.4.4.3. Hamiltoniano con condiciones de contorno en el río

Para todo estado ω en Ω^R , sea ω^R el estado en Ω definido por

$$\omega_p^R := \begin{cases} \omega_p, & p \in \Lambda^R \\ \omega_p^R, & p \in R. \end{cases}$$

Denotamos por π^R la función de probabilidad condicional en Λ^R , dado que se conocen los datos del río ω^R . Entonces tenemos

$$\pi^R(\omega) := \mathbb{P}_\pi(W_{\Lambda^R} = \omega \mid W_R = \omega^R) := \frac{\pi(\omega^R)}{\sum_{a \in \Omega^R} \pi(a^R)}.$$

Es bien sabido que las medidas condicionales de las medidas de Gibbs son medidas de Gibbs en sí mismas. Un cálculo directo proporciona el Hamiltoniano H^R correspondiente a π^R :

Definición 2.4.1 (Hamiltoniano con condiciones de contorno en R). *Para $\omega \in \Omega^R$, definimos*

$$H_0^R(\omega) := \sum_{s_1 \sim s_2} (\omega_{s_1} - \omega_{s_2})^2 + \sum_{\substack{s \sim r \\ r \in R}} (\omega_s - \omega_r^R)^2, \quad (2.4)$$

$$V_g^R(\omega) = g \sum_{s \in \Lambda^R} d_s^2 \omega_s. \quad (2.5)$$

El Hamiltoniano que incluye los datos recolectados sobre el río es

$$H^R(\omega) = H_0^R(\omega) + V_g^R(\omega).$$

Observación. Después de un cálculo sencillo, obtenemos que

$$\pi^R(\omega) = \frac{1}{Z^R} \exp \left\{ -\frac{1}{T} H^R(\omega) \right\}.$$

y la función de partición Z^R es

$$Z^R := \sum_{\omega \in \Omega^R} \exp \left\{ -\frac{1}{T} H^R(\omega) \right\}.$$

Por lo tanto, H^R es el Hamiltoniano que corresponde a π^R .

2.4.4.4. Hamiltonianos sobre las Regiones 1,2,3 y 4

Para todo estado ω en Ω^i , sea ω^i el estado en $\Omega^i \cup R$ definido por

$$\omega_p^i := \begin{cases} \omega_p, & p \in \Lambda^i \\ \omega_p^R, & p \in R. \end{cases} \quad (2.6)$$

Denotamos por π^i la función de probabilidad condicional en Λ^i , dado que se conocen los datos del río ω^R :

$$\pi^i(\omega) := \mathbb{P}_\pi(W_{\Lambda^i} = \omega \mid W_R = \omega^R) := \frac{\pi(\omega^i)}{\sum_{a \in \Omega^i} \pi(a^i)}.$$

Definición 2.4.2 (Hamiltonianos sobre las Regiones 1,2,3 y 4). *Para cada $i = 1, 2, 3, 4$ y para todo $\omega \in \Omega^i$, definimos*

$$H_0^i(\omega) := \sum_{s_1 \sim s_2} (\omega_{s_1} - \omega_{s_2})^2 + \sum_{\substack{s \sim r \\ r \in R}} (\omega_s - \omega_r^R)^2,$$

$$V_g^i(\omega) = g \sum_{s \in \Lambda^i} d_s^2 \omega_s,$$

y definimos el Hamiltoniano sobre la Región i por

$$H^i(\omega) = H_0^i(\omega) + V_g^i(\omega). \quad (2.7)$$

Definición 2.4.3. *Para todo $\omega \in \Omega^R$ (o $\omega \in \Omega^i$) y $\nu \subset \Lambda^R$ ($\nu \subset \Lambda^i$) denotamos por*

$$\omega|_\nu$$

la restricción de ω a ν .

Observación. Un cálculo directo nos lleva a

$$\pi^i(\omega) = \frac{1}{Z^i} \exp \left\{ -\frac{1}{T} H^i(\omega) \right\},$$

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

donde la función de partición Z^i está dada por

$$Z^i := \sum_{\omega \in \Omega^i} \exp \left\{ -\frac{1}{T} H^i(\omega) \right\}.$$

De las definiciones anteriores se sigue que para todo $\omega \in \Omega^R$

$$H^R(\omega) = \sum_{i=1}^4 H^i(\omega|_{\Lambda^i})$$

y

$$\pi^R(\omega) = \prod_{i=1}^4 \pi^i(\omega|_{\Lambda^i}),$$

y esto implica que las variables aleatorias W_{Λ^i} son independientes de las variables W_{Λ^j} siempre que $i \neq j$. Consecuentemente, podemos estudiar las separadamente las regiones $\Lambda^1, \Lambda^2, \Lambda^3$ y Λ^4 , i.e. tales regiones están desacopladas.

2.5. Muestreo de Gibbs

2.5.1. Ordenamiento de las celdas

Primero clasificamos las celdas según la distancia a R (las celdas que tocan el río) en términos de vecinos. Establecemos la Clase 0, que denotamos por G_0 , las celdas que tocan el río: $G_0 := R$. La clase 1, denotada por G_1 , consta de las celdas en $\Lambda \setminus G_0$ que son vecinas de las celdas en G_0 . En general, definimos para cada $k \in \mathbb{N}$,

$$G_k := \left\{ l \in \Lambda \setminus \bigcup_{m=0}^{k-1} G_m : \exists t \in G_{k-1} \text{ tal que } l \sim t \right\}.$$

Dado que Λ es finito, existe \mathbf{k} tal que $G_{\mathbf{k}} \neq \emptyset$ y para todo $k > \mathbf{k}$, $G_k = \emptyset$. Se sigue entonces que

$$\Lambda = \bigcup_{m=0}^{\mathbf{k}} G_m, \quad \Lambda^R = \bigcup_{m=1}^{\mathbf{k}} G_m,$$

y las clases G_m son disjuntas.

En la Figura 2.3, representamos un ejemplo simplificado en una cuadrícula de 10×10 , donde las celdas azules representan el río y contienen el número 0, lo que muestra que pertenecen a G_0 . De manera similar, las celdas con el número i pertenecen a la

clase G_i , $i \in \{1, 2, 3, 4\}$. La intensidad del color gris decrece con la distancia al río.

2	1	0	0	0	1	1	0	1	2
3	2	1	1	1	0	1	1	0	1
4	3	2	1	0	0	1	2	1	0
3	2	1	1	0	0	1	2	2	1
2	1	0	0	1	0	0	1	2	2
1	0	1	1	2	1	1	0	1	2
0	1	2	2	1	0	0	0	1	2
1	2	2	1	0	1	1	1	2	3
2	3	3	2	1	0	1	2	3	4
3	4	3	2	1	0	1	2	3	4

Figura 2.3: Una cuadrícula de 10×10 con celdas de río representadas en azul. El número de cada celda corresponde a su clase.

Las regiones Λ^1 , Λ^2 , Λ^3 , Λ^4 se ordenan independientemente. La forma específica en que se ordenan varía de una región a otra, pero lo importante es que las primeras celdas corresponden a las celdas en G_1 y son seguidas por las celdas en G_2 . Procedemos de esta manera hasta llegar a las celdas en G_k . A continuación, solo describimos el orden en la Región Λ^4 (ver Figura 2.4, las celdas del río R son de color azul. La intensidad del color gris es uniforme en cada una de las clases G_i y disminuye a medida que aumenta i) y no abordamos la descripción del ordenamiento de las otras regiones, porque se analizan de manera muy similar:

En este caso $k = 108$, y el número de celdas en Λ^4 es

$$m_4 = 13715.$$

Proporcionamos un orden en Λ^4 , que denotamos por $<_{\Lambda^4}$, de la siguiente manera: Sea $l, \theta \in \Lambda^4$, luego

$$l <_{\Lambda^4} \theta$$

si hay algunos $i < j$ tales que $l \in G_i$ y $\theta \in G_j$ o

$$l <_{\text{lex}} \theta,$$

donde $<_{\text{lex}}$ es el orden lexicográfico. En la Sección 2.9.1 mostramos el algoritmo computacional implementado para realizar este orden.

Sea $\Lambda^4 = \{c_k : 0 \leq k \leq m_4 - 1\}$ una lista ordenada de Λ^4 , i.e. $c_k <_{\Lambda^4} c_{k+1}$, para toda $k \in \{0, \dots, m_4 - 2\}$. Ahora damos una definición precisa de la sucesión $(\ell_n)_{n \geq 1}$, introducida en la Sección 1.5.5: Definimos

$$\ell_n = c_k, \quad \text{donde } n - 1 \equiv k \pmod{m_4}. \quad (2.8)$$

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

La sucesión $(\ell_n)_{n \geq 1}$ es periódica con periodo $m_4 = 13715$ y satisface los ítems presentados en la Sección 1.5.5.

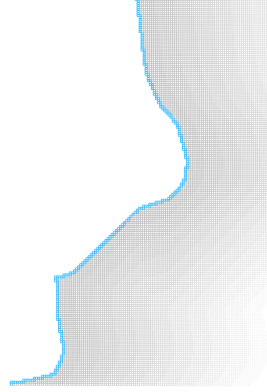


Figura 2.4: Orden de las celdas de la Región 4. Las celdas del río están en azul. El color gris de las clases de las celdas se degrada conforme la distancia al río. Puede observarse un patrón de “franjas” desde el río hasta el exterior.

2.5.2. Probabilidades condicionales para celdas individuales

En esta sección, damos una expresión explícita para las probabilidades condicionales en cada celda asumiendo que el número de individuos se conoce para todas las demás celdas. Estas probabilidades condicionales son ingredientes clave en el método de muestreo de Gibbs.

Definición 2.5.1. Sea $v \subsetneq \Lambda$, $q \in \Lambda \setminus v$ y $\omega \in \Gamma^v$. Para cada $x \in \Gamma$, denotamos $\iota_{q,x}(\omega) \in \Gamma^{\{q\} \cup v}$, el estado

$$\iota_{q,x}(\omega)(r) := \begin{cases} \omega(r), & \text{if } r \neq q, \\ x, & \text{if } r = q. \end{cases}$$

Para cada $p \in \Lambda^i$ y cada estado $\omega \in \Omega^i \setminus \{p\}$, definimos $H_p^{(\omega)} : \Gamma \rightarrow \mathbb{R}$ por (ver (2.6))

$$H_p^{(\omega)}(x) := \sum_{q \in N(p)} (x - \omega_q^i)^2 + g d_p^2 x,$$

donde $N(p)$ es el conjunto de celdas vecinas de p (ver (2.2)). Note que $H_p^{(\omega)}$ es el

Hamiltoniano asociado a la medida de Gibbs

$$\mathbb{P}_{\pi^i}(W_p = x | W_{\Lambda^i \setminus \{p\}} = \omega) := \frac{\pi^i(\iota_{p,x}(\omega))}{\sum_{y \in \Gamma} \pi^i(\iota_{p,y}(\omega))} = \frac{1}{Z^{(\omega)}} \exp \left\{ -\frac{1}{T} H_p^{(\omega)}(x) \right\}, \quad (2.9)$$

y la constante de normalización $Z^{(\omega)}$ es

$$Z^{(\omega)} := \sum_{y \in \Gamma} \exp \left\{ -\frac{1}{T} H_p^{(\omega)}(y) \right\}.$$

2.5.3. Cadenas de Markov y distribuciones invariantes

En esta sección, construimos cadenas de Markov que tienen a π^i como distribuciones invariantes. Lo hacemos solo para la región Λ^4 (es decir, $i = 4$), ya que la construcción de las otras regiones es análoga. Luego, construimos el proceso \mathbf{X}^4 introducido en la Sección 1.5.5 cuya distribución invariante es π^4 .

Dada la sucesión $(\ell_n)_{n \geq 1}$ definida anteriormente en la Sección 2.5.1, definimos una sucesión de matrices de transición $(\mathbf{M}_n^4)_{n \geq 1}$ con espacio de estados Ω^4 mediante la siguiente fórmula:

$$\mathbf{M}_n^4(\omega', \omega) = \begin{cases} \frac{1}{Z^{(\omega |_{\Lambda^4 \setminus \{\ell_n\}})}} e^{-\frac{1}{T} H_{\ell_n}^{(\omega |_{\Lambda^4 \setminus \{\ell_n\}})}(\omega_{\ell_n})}, & \text{si } \omega' |_{\Lambda^4 \setminus \{\ell_n\}} = \omega |_{\Lambda^4 \setminus \{\ell_n\}}, \\ 0, & \text{si } \omega' |_{\Lambda^4 \setminus \{\ell_n\}} \neq \omega |_{\Lambda^4 \setminus \{\ell_n\}}. \end{cases}$$

El estado inicial $\mathbf{X}^4(0)$ es elegido como

$$\mathbf{X}^4(0) = a^{(0)} \in \Omega^4,$$

con probabilidad $\pi^4(a^{(0)})$. Usando el Teorema de extensión de Kolmogorov, construimos un espacio de probabilidad $(A^4, \mathcal{F}^4, \mathbb{P}^4)$ y una cadena de Markov no homogénea $\mathbf{X}^4 = (\mathbf{X}(n))_{n \geq 0}$ con espacio de estados Ω^4 , tal que

$$\mathbb{P}^4(\mathbf{X}^4(n) = \omega \mid \mathbf{X}^4(n-1) = \omega') = \mathbf{M}_n^4(\omega', \omega), \quad \omega', \omega \in \Omega^4.$$

Observamos que \mathbf{X}^4 no es estacionario (las probabilidades de transición dependen de n), sin embargo, es aperiódico e irreducible (ver Geman-Geman [61]). Además, converge a la distribución invariante π^4 .

Teorema 2.5.1 (Relaxation. Theorem A [60]). *Para casi todas las configuraciones*

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

iniciales $\omega_0 \in \Omega^4$ y cada $\omega \in \Omega^4$,

$$\lim_{n \rightarrow \infty} \mathbb{P}^4(\mathbf{X}^4(n) = \omega \mid \mathbf{X}^4(0) = \omega_0) = \pi^4(\omega).$$

Este resultado de convergencia es fundamental porque nos coloca en la situación usual en la que queremos obtener una muestra cuya distribución sea lo suficientemente cercana a la distribución objetivo π^4 . La implementación del algoritmo se explicará más adelante en la Sección 2.5.4. A medida que aumenta T , el sistema tiende a ser puramente aleatorio, y cuando la temperatura tiende a cero, el sistema se concentra en los mínimos. Si la temperatura es lo suficientemente pequeña (pero no cero), los estados se concentran (en probabilidad) alrededor de los mínimos (pero generalmente se tarda demasiado en alcanzarlos y uno se atasca en los mínimos locales). Entonces, para fines prácticos, es suficiente tener una temperatura pequeña que dé una probabilidad baja de tener configuraciones fuera de los mínimos (esto se puede probar numéricamente). Es importante mantener la temperatura fija porque esto nos permite utilizar teoremas ergódicos [61]. El algoritmo de Metropolis-Hastings la mayoría de las veces produce mínimos locales (en el sentido de la Definición 1.5.1) que dependen en gran medida de las condiciones iniciales (ver Sección 2.8). El problema es grave, de hecho en [69], se demuestra que el *simulated annealing* requiere demasiadas iteraciones (tantas que las computadoras esencialmente nunca podrían realizar). Para resolver este problema, el orden de las celdas es crucial descrito en la Sección 2.5.1, porque promueve que la energía libre de los estados se acerque al mínimo global. Usamos la energía libre para probar el modelo y elegir la temperatura. Luego, ejecutamos el algoritmo Metropolis-Hastings e iteramos hasta que la energía se estabilice. Usamos un esquema de temperatura descendente (*simulated annealing*) solo para elegir estados iniciales.

2.5.4. Algoritmo del Muestreo de Gibbs

Como antes, solo nos enfocamos en la Región 4, porque el análisis de las otras regiones es muy similar. Dado que la temperatura es T es baja, la energía de los estados con mayor probabilidad está cerca del mínimo global. El principal interés de este trabajo es estimar el número de individuos en cada celda. Entonces tenemos que estimar

$$\mathbb{E}_{\pi^R}(W_p) = \mathbb{E}_{\pi^4}(W_p),$$

para cada $p \in \Lambda^4$. Esto se lleva a cabo utilizando un teorema ergódico (Teorema Ergódico 2.5.2 a continuación). La idea es encontrar una sucesión cuyas energías son descendentes y se acercan al mínimo cuando el parámetro de la sucesión tiende a infinito. La

cola de la sucesión contiene los estados más probables y, por lo tanto, podemos usarla para calcular los valores esperados. Esto solo es posible si elegimos estados iniciales y esquemas de muestreo apropiados; de lo contrario, la convergencia en el teorema ergódico podría hacer que el problema sea irresoluble (el Teorema Ergódico 2.5.2 se informan las tasas de convergencia).

A continuación, construimos la sucesión (aleatoria) de estados descrita en la Sección 1.5.5. Denotamos por $a^{(0)} \in \Omega^4$ al estado inicial:

$$\omega_p^{(0)} = a_p^{(0)}, \quad \forall p \in \Lambda^4.$$

Elegimos $\omega^{(n)}$ recursivamente de la siguiente manera. Suponemos que

$$\omega^{(0)}, \dots, \omega^{(n-1)}$$

están definidos. Elegimos aleatoriamente un elemento $x \in \Gamma$ y seguimos los siguientes pasos (recuerde (2.8) y la Sección 1.5.5)

- (1) Si (recordar Definición 2.5.1 y (2.7))

$$H^4\left(\iota_{\ell_n, x}(\omega^{(n-1)}|_{\Lambda^4 \setminus \{\ell_n\}})\right) < H^4(\omega^{(n-1)}),$$

entonces elegimos

$$\omega^{(n)} = \iota_{\ell_n, x}(\omega^{(n-1)}|_{\Lambda^4 \setminus \{\ell_n\}}).$$

- (2) Supongamos que

$$H^4\left(\iota_{\ell_n, x}(\omega^{(n-1)}|_{\Lambda^4 \setminus \{\ell_n\}})\right) \geq H^4(\omega^{(n-1)}).$$

Recordemos

$$\frac{\mathbb{P}_{\pi^i}\left(W_{\ell_n} = x \mid W_{\Lambda^4 \setminus \{\ell_n\}} = \omega^{(n-1)}|_{\Lambda^4 \setminus \{\ell_n\}}\right)}{\mathbb{P}_{\pi^i}\left(W_{\ell_n} = \omega_{\ell_n}^{(n-1)} \mid W_{\Lambda^4 \setminus \{\ell_n\}} = \omega^{(n-1)}|_{\Lambda^4 \setminus \{\ell_n\}}\right)} = e^{-\frac{1}{T} \left[H^i\left(\iota_{\ell_n, x}(\omega^{(n-1)}|_{\Lambda^4 \setminus \{\ell_n\}})\right) - H^i(\omega^{(n-1)}) \right]}.$$

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

Elegimos aleatoriamente $z \in (0, 1]$. If

$$z \leq e^{-\frac{1}{T} \left[H^i \left(\iota_{\ell_n, x} \left(\omega^{(n-1)} \Big|_{\Lambda^4 \setminus \{\ell_n\}} \right) \right) - H^i \left(\omega^{(n-1)} \right) \right]},$$

entonces definimos

$$\omega^{(n)} = \iota_{\ell_n, x} \left(\omega^{(n-1)} \Big|_{\Lambda^4 \setminus \{\ell_n\}} \right),$$

en otro caso, seleccionamos

$$\omega^{(n)} = \omega^{(n-1)}.$$

Definición 2.5.2 (Muestreo de Gibbs). *Un muestreo de Gibbs (con el estado inicial $a^{(0)}$) es una sucesión seleccionada al azar construida de forma recursiva como en los pasos (1) y (2) anteriores.*

La importancia del muestreo de Gibbs es que $(\omega^{(n)})_{n \in \mathbb{N}}$ se puede usar para calcular los valores esperados del número de individuos en cada celda, este es el contenido del siguiente teorema ergódico:

Teorema 2.5.2 (Teorema Ergódico - Theorem C [61], 5.1.4 Theorem [158]). *Para toda función $f : \Omega^4 \rightarrow \mathbb{R}$,*

$$\int f(\omega) d\pi^4(\omega) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{r=1}^n f(\omega^{(r)}),$$

en probabilidad y en L^2 . Para toda $\epsilon > 0$,

$$\pi^4 \left(\left| \frac{1}{n} \sum_{r=1}^n f(\omega^{(r)}) - \int f(\omega) d\pi^4(\omega) \right| > \epsilon \right) \leq \frac{c}{n\epsilon^2} e^{m_4 \Delta^4}, \quad (2.10)$$

donde $c = 13\|f\|^2$, m_4 es el número de celdas en Λ^4 y Δ^4 es la oscilación local máxima de H^4 dada por

$$\Delta^4 = \max\{\delta_s : s \in \Lambda^4\},$$

y δ_s es la oscilación de H^4 en la celda s dada por

$$\delta_s = \text{máx}\{|H^4(\omega) - H^4(\hat{\omega})| : \omega_t = \hat{\omega}_s, \text{ para toda celda } t \neq s\}.$$

En particular, para toda $p \in \Lambda^4$ estimamos el valor esperado del número de individuos en la celda p usando que

$$\mathbb{E}_{\pi^R}(W_p) = \mathbb{E}_{\pi^4}(W_p) = \int W_p d\pi^4(\omega) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{r=1}^n W_p(\omega^{(r)}),$$

en probabilidad y en L^2 .

El Teorema Ergódico 2.5.2 nos permite estimar los valores esperados del número de individuos en cada celda. Sin embargo, su aplicabilidad no es obvia porque los límites de error son enormes en nuestro caso y la convergencia es solo en probabilidad. Aquí nos enfrentamos nuevamente a los problemas de intratabilidad descritos en muchos lugares de este trabajo. En general, n tiene que ser muy gr y e (intratable) para que $\frac{1}{n} \sum_{m=1}^n W_p(\omega^{(m)})$ sea un aaproximación eficiente $\mathbb{E}_{\pi^R}(W_p)$ porque la sucesión $(\omega^{(n)})$ puede atascarse en los mínimos locales. Nuestro ordenamiento de celdas específicas produce estados con energías cercanas al mínimo para el caso libre (es decir, los estados más probables), y usamos perturbaciones en el caso de que $g \neq 0$. Esto se lleva a cabo con $m_4 \times 1000000$ iteraciones. Sin embargo, esta cantidad de iteraciones no proporciona buenas aproximaciones para el mínimo de energía libre para cada estado inicial. El estado inicial debe elegirse cuidadosamente. Una vez que se elige, $m_4 \times 1000000$ iteraciones son suficientes para obtener una aproximación bastante buena de $\mathbb{E}_{\pi^R}(W_p)$, el número esperado de individuos en la celda p , para cada p , en el caso libre. Para el caso perturbado, obtenemos un control global de tal manera que cuando $g = 0$ hay una distribución homogénea en toda la región y a medida que g aumenta, los individuos tienden a concentrarse en el río de manera continua y controlada. Esto es todo lo que necesitamos para modelar, sin importar si nos acercamos o no al mínimo global para el caso perturbado. Sin embargo, este comportamiento indica que nos acercamos a él.

2.5.5. Selección de parámetros

2.5.5.1. Temperaturas auxiliares y estados iniciales

Nos enfocamos en Región 4 para determinar la temperatura (y el estado inicial). En general, cada región debe analizarse por separado y los parámetros pueden elegirse

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

para cada región de forma independiente. Seleccionamos un conjunto de temperaturas T_1, \dots, T_n y un conjunto de estados iniciales $a^{(1)}, \dots, a^{(n)}$ de la siguiente manera: Tomamos $g = 0$ y $a^{(1)} = 0$ (y $\Gamma = \{0, \dots, K\}$). Ejecutamos muestreos de Gibbs con $m_4 \times 1000000$ iteraciones, donde $m_4 = 13715$ es el número de celdas en la Región 4, estado inicial $a^{(1)}$ y diferentes valores para la temperatura T . Tomamos los promedios de las últimas $m_4 \times 10000$ iteraciones y calculamos las energías (libres) de estos promedios. Elegimos el valor de T que (numéricamente) minimiza estas energías. Denotamos este valor por T_1 y el promedio correspondiente de los estados $a^{(2)}$. Ahora tomamos $a^{(2)}$ como estado inicial, $\Gamma = [0, K]$, y ejecutamos nuevamente $m_4 \times 1000000$ iteraciones de muestreo de Gibbs con diferentes valores de la temperatura T . Elegimos el valor de T que minimiza las energías libres como se indicó anteriormente. Llamamos a este valor T_2 y el promedio correspondiente de los estados (como arriba) $a^{(3)}$. Continuamos con este procedimiento hasta obtener una energía suficientemente pequeña (de acuerdo con nuestras expectativas con respecto a la aplicación que se pueda tener en mente). De esta forma, obtenemos T_1, \dots, T_n y $a^{(1)}, \dots, a^{(n)}$. En nuestro caso, tomamos $n = 2$.

2.5.5.2. Estado inicial, temperatura y constante de acoplamiento

El procedimiento de la Sección 2.5.5.1 produce temperaturas T_1, \dots, T_n que minimizan la energía libre. Sin embargo, esta minimización es solo numérica y, por lo tanto, son admisibles pequeñas variaciones de estos valores mínimos. Luego, el método de la Sección 2.5.5.1 proporciona, en realidad, pequeños intervalos para los valores de T_1, \dots, T_n . Tomamos ventaja de esta flexibilidad. Denotamos por

$$a_g^{(1)}, \dots, a_g^{(n)}$$

los estados iniciales correspondientes pero con $g > 0$ (las temperaturas T_1, \dots, T_n se calculan con $g = 0$). Definimos un conjunto de constantes de acoplamiento g_1, \dots, g_n (que dependen de las temperaturas T_1, \dots, T_n) de la siguiente manera: A partir de $g = 0$, aumentamos lentamente el valor de g y calculamos $m_4 \times 1000000$ iteraciones de muestreo de Gibbs (con estado inicial $a_g^{(1)} = 0$ y temperatura T_1). Tomamos el promedio de las últimas $m_4 \times 10000$ iteraciones, lo comparamos con los datos del trabajo de campo y tomamos el valor de g que mejor coincide con estos datos (numéricamente). A este valor lo llamamos g_1 . Repetimos el procedimiento, pero ahora con el estado inicial $a_g^{(2)}$ por g en un pequeño vecindario de g_1 y $\Gamma = [0, K]$. Denotamos por g_2 el valor que mejor coincide con los datos del trabajo de campo. Continuamos con este procedimiento hasta obtener el conjunto de constantes de acoplamiento g_1, \dots, g_n (dependen de T_1, \dots, T_n).

Finalmente, variamos los valores de T_1, \dots, T_n dentro de los intervalos descritos al

principio de esta sección para obtener el mejor decaimiento de energía (tomando $g = g_n$) para la sucesión que consta de $m_4 \times 1000000$ Iteraciones de muestreo de Gibbs con estado inicial $a_{g_n}^{(1)}$ y temperatura T_1 seguidas de $m_4 \times 1000000$ iteraciones con estado inicial $a_{g_n}^{(2)}$ y temperatura T_2 y así sucesivamente hasta llegar a $n \times (m_4 \times 1000000)$ iteraciones. Los valores resultantes para T_n , g_n y $a_{g_n}^{(n)}$ se llaman

$$\mathbf{T}, \quad \mathbf{g}, \quad \mathbf{a},$$

y son la temperatura real, la constante de acoplamiento y el estado inicial de nuestro sistema.

2.6. Resultados

2.6.1. Resultados principales: distribución y abundancia de individuos

El equilibrio entre atracción y repulsión se expresa mediante los estados de baja energía que son los estados más probables, en el régimen de baja temperatura, obtenidos por muestreo de Gibbs. La fuerza de atracción está determinada por la constante de acoplamiento \mathbf{g} , y se define para que coincida con los datos experimentales. Hay dos tipos de datos experimentales que se utilizarán en nuestros trabajos futuros: el primero es el conocimiento de que no hay individuos a distancias superiores a un cierto valor (esto ocurre cuando los individuos se concentran en el río y se toman muestras en las cercanías de es muy difícil); el segundo es la recopilación de datos sobre determinadas celdas fuera del río. En esta investigación, nos enfocamos en el primer tipo de información porque brinda una mejor comprensión del control global que logramos. Presentamos nuestros resultados en cuatro situaciones:






















- S1** No hay atracción hacia el río y tenemos un número constante de individuos en cada celda del río.
- S2** No se recolectaron individuos a distancias (horizontales) superiores a $300m$ hasta el río en la Región 4 (pero podría haber individuos a distancias más cortas).
- S3** No se recolectaron individuos a distancias (horizontales) superiores a $150m$ hasta el río en la Región 4 (pero podría haber individuos a distancias más cortas).
- S4** No se recolectaron individuos a distancias (horizontales) superiores a $15m$ hasta el río en la Región 4 (pero podría haber individuos a distancias más cortas).

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

Tomamos el estado inicial, la temperatura T y la constante de acoplamiento g como se describe en la Sección 2.5.5, para la Región 4, de acuerdo con las situaciones S1-S4 anteriores. Para las otras regiones también elegimos T y g , y el estado inicial se calcula de una manera análoga usando una temperatura auxiliar T_1 . Calculamos $m_i \times 1000000$ iteraciones de muestreo de Gibbs y calculamos el promedio de las últimas $m_i \times 10000$ iteraciones (donde m_i es el número de celdas en cada región i , con $i = 1, 2, 3, 4$). Este estado proporciona el valor esperado previsto de los individuos en cada celda y constituye nuestro principal resultado. Usamos representaciones gráficas para presentarlos (para cada situación S1-S4):

En las Figuras 2.5, 2.6, 2.7, y 2.8, mostramos una representación gráfica de los valores esperados del número de individuos en cada celda, utilizando mapas de calor, para cada situación descrita S1-S4. De las cifras se desprende que obtenemos un control muy preciso del comportamiento global, partiendo de una distribución homogénea (S1). A medida que aumenta la constante de acoplamiento, los individuos se retraen hacia el río de tal manera que ningún individuo a distancias del río superiores a $300m$, $150m$ y $15m$ se obtienen en las situaciones S2, S3 y S4, respectivamente. Los códigos de color de los mapas de calor siguen las siguientes convenciones: Usamos los 7 colores del arco iris y cada color se divide en tres tonos diferentes: El color más energético (violeta) representa la densidad más alta y el color menos energético (rojo) representa la densidad más baja. La escala de colores va acompañada de una degradación de la intensidad, que representa un número esperado de individuos según la Tabla 2.1.

Cuadro 2.1: Degradación de la intensidad de color y correspondiente número esperado de individuos

	(0.001, 0.620]		(5.572, 6.191]		(11.143, 11.762]
	(0.620, 1.239]		(6.191, 6.810]		(11.762, 12.381]
	(1.239, 1.858]		(6.810, 7.429]		(12.381, 13]
	(1.858, 2.477]		(7.429, 8.048]		
	(2.477, 3.096]		(8.048, 8.667]		
	(3.096, 3.715]		(8.667, 9.286]		
	(3.715, 4.334]		(9.286, 9.905]		
	(4.334, 4.953]		(9.905, 10.524]		
	(4.953, 5.572]		(10.524, 11.143]		

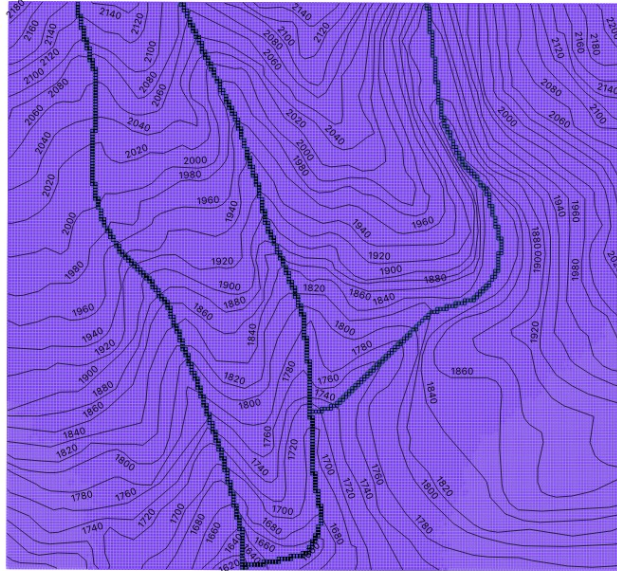


Figura 2.5: Abundancia y distribución de individuos en situación S1 con $K = 13$, $g_{S1} = 0$ y $T = 0.000027$

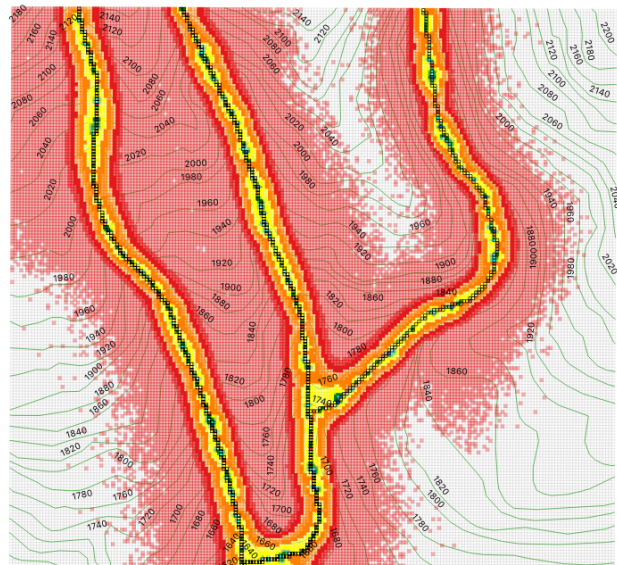


Figura 2.6: Abundancia y distribución de individuos en situación S2, con $g_{S2} = 0.0014$ y $T = 0.000027$

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

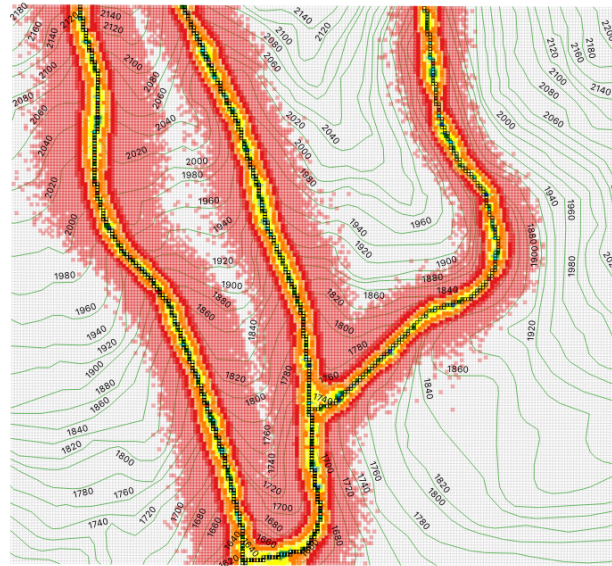


Figura 2.7: Abundancia y distribución de individuos en situación S3, con $g_{S3} = 0.005$ y $T = 0.000027$

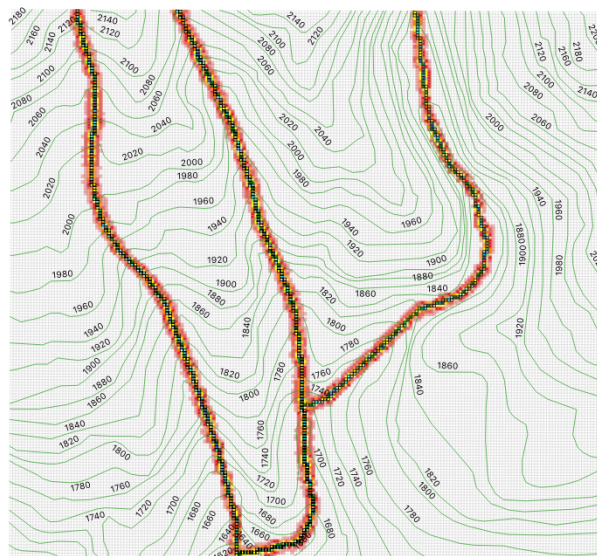


Figura 2.8: Abundancia y distribución de individuos en situación S4, con $g_{S4} = 0.5$ y $T = 0.000027$

2.6.2. Resultados auxiliares: estimación de parámetros y caída de energía

Temperatura auxiliar La temperatura auxiliar descrita en la Sección 1.6.1 es:

$$T_1 = 0.202083.$$

La constante de acoplamiento g y las temperaturas del sistema El parámetro de referencia g_0 descrito en la Sección 1.6.2 es

$$g_0 = 0.01.$$

Usamos los subíndices $S1, S2, S3, S4$, para describir las situaciones correspondientes presentadas en las situaciones descritas en la Sección 2.6.1. Obtenemos

$$\mathbf{T} = 0.000027$$

y los parámetros \mathbf{g} para cada situación $S1-S4$ están dados por

$$\mathbf{g}_{S1} = 0, \quad \mathbf{g}_{S2} = 0.0014, \quad \mathbf{g}_{S3} = 0.005, \quad \mathbf{g}_{S4} = 0.5.$$

En las Figuras 2.9, 2.10, 2.11, y 2.12, presentamos las energías de las simulaciones de muestreo de Gibbs con $m_i \times 2000000$ iteraciones (m_i es el número de celdas en la Región i , para $i = 1, 2, 3, 4$) como se describe en la Sección 1.6.2 para cada situación $S1-S4$: Para cada $X \in \{S1, S2, S3, S4\}$, las primeras $m_i \times 1000000$ las iteraciones se calculan con T_1, \mathbf{g}_X , y estado inicial cero. Las segundas $m_i \times 1000000$ las iteraciones se calculan con \mathbf{T}, \mathbf{g}_X , y estado inicial el promedio de los últimos $m_i \times 10000$ iteraciones de las primeras $m_i \times 1000000$ iteraciones. En estas figuras, observamos un descenso brusco de la energía y una estabilización en las últimas iteraciones. Esto indica que nuestro método produce estados con energías muy bajas.

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

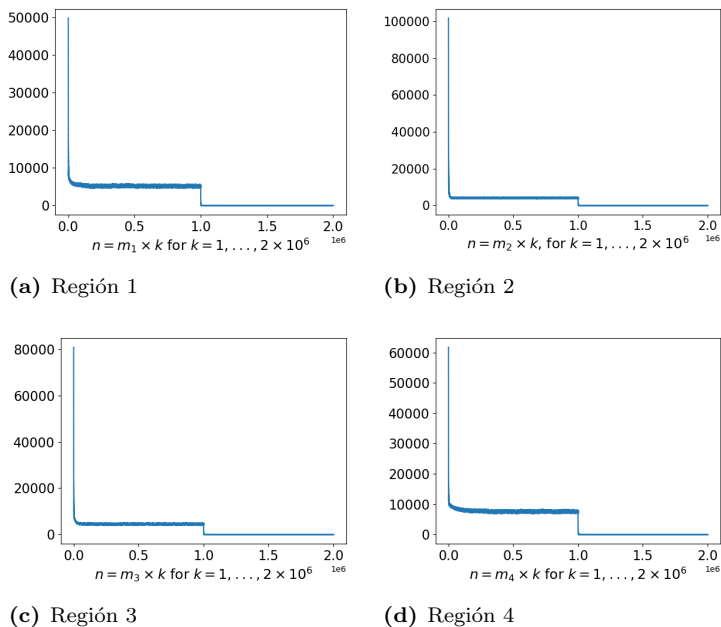


Figura 2.9: Decaimiento de la energía en la situación S1, con $g_{S1} = 0$ y $K = 13$

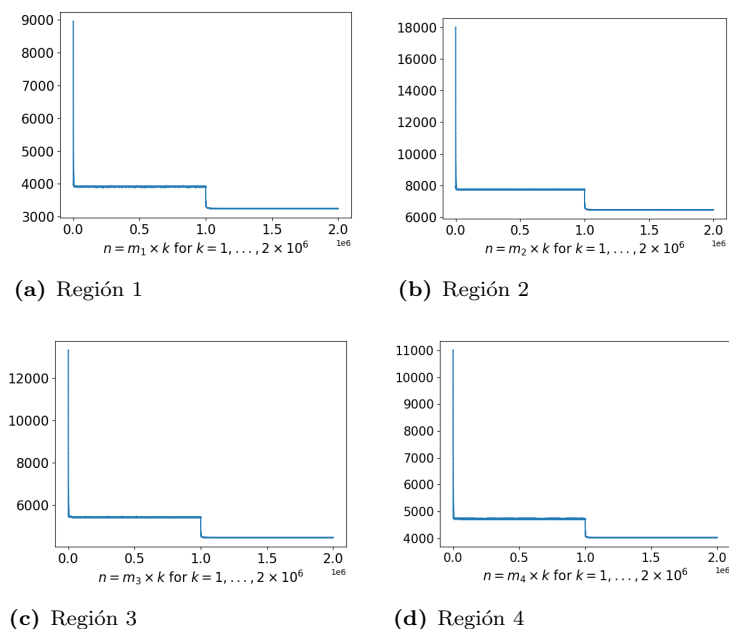


Figura 2.10: Decaimiento de la energía en la situación S2, con $g_{S2} = 0.0014$ y $T = 0.000027$

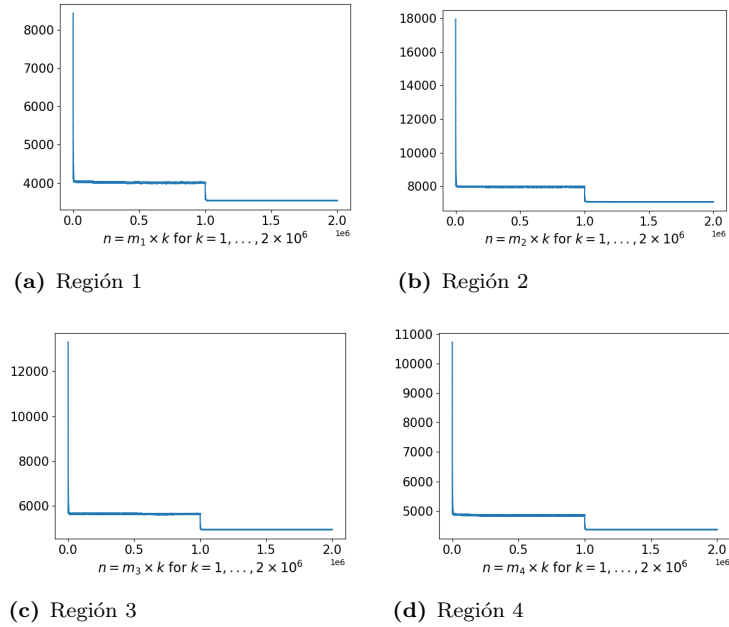


Figura 2.11: Decaimiento de la energía en la situación S3, con $g_{S3} = 0.005$ y $T = 0.000027$

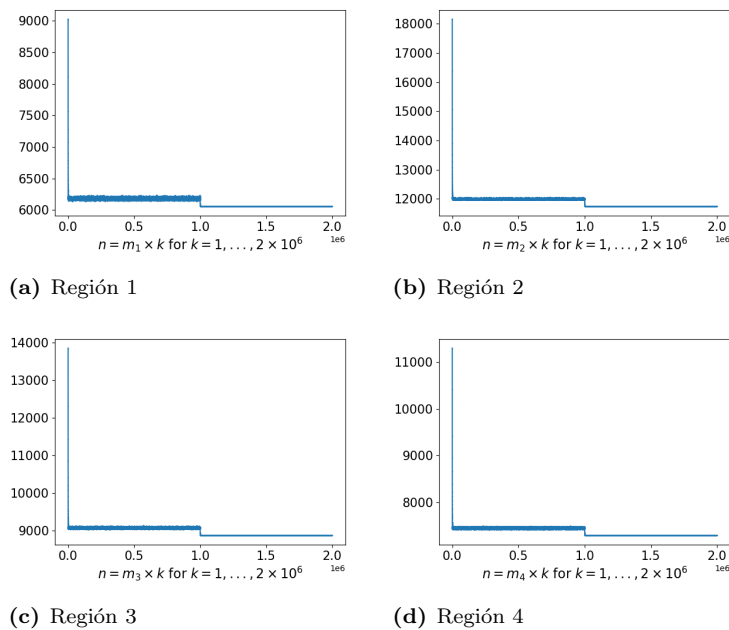


Figura 2.12: Decaimiento de la energía en la situación S4, con $g_{S4} = 0.5$ y $T = 0.000027$

2.7. Conclusiones

Introducimos un modelo y un método numérico en el contexto de geoestadística que se aparta sustancialmente de los métodos comunes de geoestadística. En lugar de centrarnos en las propiedades locales de los campos aleatorios de Markov, estudiamos la medida global y proponemos un enfoque numérico que resuelve el problema de la aproximación al mínimo global, que en la mayoría de los casos es intratable o la técnica de muestreo de Gibbs es deficiente y no produce aproximaciones consistentes al mínimo global. Nuestra propuesta se basa en un esquema de perturbación y un ordenamiento de muestreo inteligente.

El caso libre es importante porque en este caso se conoce la energía mínima, así como los estados fundamentales. Introducimos un orden de muestreo del muestreo de Gibbs que nos permite acercarnos a la energía mínima en el caso libre. Perturbamos el sistema y realizamos el mismo esquema de muestreo que en el caso libre. Esto nos permite controlar el comportamiento global del sistema y, por lo tanto, proporciona una armadura de herramientas que se pueden utilizar para comprender la distribución y abundancia de especies.

Las simulaciones numéricas se comportan exactamente como se esperaba: cuando no hay razón para que estén en un lugar específico (o lejos de algún lugar), entonces la presencia y abundancia de individuos es homogénea en toda la región (el caso libre). Adicionalmente, este esquema de ordenamiento en las celdas nos permite elegir estados iniciales que evita el problema persistente del método de muestreo de Gibbs, de no alcanzar los estados óptimos. Si hay algunas características específicas del entorno que atraen (o repelen) a los individuos, entonces podemos controlar nuestros parámetros de tal manera que produzcamos todos los escenarios posibles: se mueven continuamente del caso libre a la situación en la que cada individuo está en el atractor.

Nuestro enfoque se basa en una intuición física: comenzamos nuestro modelo definiendo la medida global en términos de un Hamiltoniano que contiene un término de energía cinética que promueve distribuciones homogéneas, una energía potencial que describe las fuerzas de atracción (o repulsión) y una constante de acoplamiento que mide la fuerza de las fuerzas. De esta manera, tenemos una imagen completa del problema y una visión global, en contraste con el enfoque habitual en geoestadística donde solo se consideran las probabilidades condicionales (el comportamiento local en una sola celda). Es importante tener en cuenta que nuestro método de muestreo es producto de muchos intentos fallidos anteriores.

En todos estos intentos anteriores nos quedamos atascados en los mínimos locales, y el problema de estos mínimos es que no pueden usarse para modelar la situación que

describimos: La dependencia de los mínimos locales de la constante de acoplamiento es altamente irrazonable, como incluso en el caso libre no hay esperanza de conseguir una distribución homogénea. La estructura de mínimos locales (en el sentido de Definición 1.5.1) es extremadamente complicada y errática con respecto a los parámetros del modelo y no hay control sobre qué mínimos locales (en el sentido de Definición 1.5.1) terminamos después de realizar simulaciones.

2.8. Un ejemplo simple de convergencia a mínimos locales

El principal problema con el algoritmo de muestreo de Gibbs es que las relaciones de vecindad local bloquean el algoritmo en los mínimos locales del Hamiltoniano (en el sentido de Definición 1.5.1), si las condiciones iniciales y la temperatura no se eligen cuidadosamente. En este apéndice proporcionamos solo unos pequeños ejemplos de cómo se observa esta situación, posicionándonos en una situación típica de estimación por máxima verosimilitud.

Entonces suponemos el caso de una cuadrícula Λ con 16 celdas, etiquetadas por (i, j) , con $i, j \in \{1, \dots, 4\}$. Tomamos solo un atractor, que es una celda fija ℓ . La situación es que queremos reconstruir la abundancia y distribución en toda la región dado que solo tenemos una observación y realizada en la celda ℓ . Para simplificar, consideramos solo el caso libre (ver Sección 2.4.2.1). En este caso, la función de verosimilitud, definida por la distribución de probabilidad condicional de la abundancia en todas las celdas de Λ excepto la celda ℓ dada la observación y (con $y \in \mathbb{N}$) realizada en la celda ℓ , viene dada por

$$L(\omega) := \pi^{\{\ell\}}(\omega) = \frac{1}{Z} e^{-H_0^{\{\ell\}}(\omega)}, \quad \omega \in \{1, \dots, y\}^{\Lambda \setminus \{\ell\}},$$

donde $H_0^{\{\ell\}}$ es el Hamiltoniano Libre (ver (2.4)) del sistema

$$H_0^{\{\ell\}}(\omega) = \sum_{(i,j) \sim (u,v)} (\omega_{i,j} - \omega_{u,v})^2 + \sum_{(i,j) \sim \ell} (\omega_{i,j} - y)^2,$$

y Z es la constante de normalización. El estado de probabilidad máxima global de L es equivalente al estado de energía mínima global de $H_0^{\{\ell\}}$. Este estado viene dado por

$$\omega_{i,j} = y$$

individuos en cada celda (i, j) .

Habitualmente, en sistemas mucho más gr y es imposible realizar directamente los cálculos para maximizar la probabilidad L , y se prefiere utilizar algún método como

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

	1	2	3	4
1	8	9	9	6
2	4	9	2	7
3	4	3	4	6
4	8	3	1	5

	1	2	3	4
1	1	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0

	1	2	3	4
1	4	2	1	1
2	2	2	1	1
3	1	1	1	1
4	1	1	1	1

Figura 2.13: Algunos ejemplos de cuadrícula 4×4 con un estado inicial que es un mínimo local en el sentido de Definición 1.5.1

el descenso de gradiente o el muestreo de Gibbs, que es lo que aquí nos interesa. Luego, queremos mostrar en la Figura 2.13 muestra tres ejemplos de estados iniciales para este sistema en el que, después de ejecutar el algoritmo de muestreo de Gibbs (con 100000 iteraciones) nos quedamos atascados precisamente en ese estado inicial (que es, por lo tanto, un mínimo local en el sentido de Definición 1.5.1. Este no es un resultado consistente, por supuesto, pero habla de la deficiencia fundamental del algoritmo de muestreo de Gibbs del que hemos hablado repetidamente en este manuscrito y que pretendemos Esta situación a menudo ocurre en nuestro modelo sin una elección cuidadosa de los estados iniciales. En nuestro caso, los métodos de probabilidad e incluso pseudo-verosimilitud son intratables debido al tamaño de la cuadrícula del sistema, por lo que nos propusimos lograr los estados iniciales que superan esta dificultad.

2.9. Algoritmos

2.9.1. Algoritmo para el orden de las celdas

A continuación presentamos el algoritmo implementado para ordenar las celdas de acuerdo al método descrito en la Sección 2.5.1

Algorithm 1 Algoritmo para el orden de las celdas

Require:

$\Lambda = R \cup S$, set of cells ordered con the lexicographic order, where

R , river cells,

S , cells that are not on the river.

```
1:  $k = 0$ 
2: while  $S \neq \emptyset$  do
3:    $R_1 = \emptyset$ 
4:   for  $i \in R$  do
5:      $g_i = k$  (class of cell  $i$ )
6:     for  $j \in S$  do
7:       if  $j \sim i$  then
8:          $R_1 = R_1 \cup \{j\}$ 
9:          $S = S \setminus \{j\}$ 
10:      end if
11:    end for
12:  end for
13:   $R = R_1$ 
14:   $k = k + 1$ 
15: end while
```

2. UN MODELO NUMÉRICO PARA LA DESCRIPCIÓN DE LA DISTRIBUCIÓN Y ABUNDANCIA DE INDIVIDUOS

2.9.2. Algoritmo Muestreo de Gibbs

Ahora se muestra el algoritmo implementado para las simulaciones del muestreo de Gibbs de la Sección 2.5.4.

Algorithm 2 Algoritmo Muestreo de Gibbs

```
1: global variables
2:    $T$ , temperature (global variable)
3: end global variables
4: while  $k \leq I$  do
5:   for  $i$  cell do
6:      $x_i^{(k)} \leftarrow$  number of individuals on  $i$  at step  $k$ 
7:     if  $i$  belongs to the river  $R$  then
8:        $x_i^{(k+1)} \leftarrow x_i^{(k)}$ 
9:     else
10:       $x \leftarrow$   $RyOM(0, 1, \dots, M)$ 
11:       $\Delta E \leftarrow$  energy difference on  $i$  con respect to the new value  $x$ 
12:      if  $\Delta E < 0$  then
13:         $x_i^{(k+1)} \leftarrow x$ 
14:      else
15:         $y \leftarrow$   $RyOM(0, 1)$ 
16:         $q \leftarrow \exp\left(-\frac{1}{T}\Delta E\right)$ 
17:        if  $y \leq q$  then
18:           $x_i^{(k+1)} \leftarrow x$ 
19:        else
20:          pass
21:        end if
22:      end if
23:    end if
24:  end for
25:   $k \leftarrow k + 1$ 
26: end while
```

Dinámica poblacional de *Plectrohyla* *Sagorum*

3.1. Introducción

En este capítulo presentamos nuevas características al modelo que presentamos en el capítulo anterior (ver también [19]), con el fin de reconstruir la dinámica poblacional de especies biológicas. Específicamente aplicamos nuestro método al estudio de una especie de anfibio en peligro de extinción llamada *Plectrohyla Sagorum* sobre la región de estudio del capítulo anterior (Sección 2.4.1). Los resultados de este capítulo están reportados en [20].

Como hemos explicado antes, la idea clave de nuestro método es utilizar los conceptos de energías cinética y potencial de la forma en que se entienden en la mecánica cuántica para construir un modelo usando la mecánica estadística clásica. El modelo se basa en un Hamiltoniano definido por la suma de un término cinético (energía o Hamiltoniano libre) y un término potencial. Entonces reducimos el problema a un problema de minimización: los estados con energías cinéticas cercanas al mínimo son casi constantes (describen probabilidades iguales para cualquier posición en el espacio) y representan individuos que se extienden por toda la región de estudio. Mientras que los estados con energías cercanas al mínimo del Hamiltoniano reflejan un equilibrio entre la fuerza de propagación y la de permanecer en los atractores. La intensidad las fuerzas de atracción está determinada por una constante de acoplamiento, que se selecciona a partir de los datos recopilados.

Suponemos que los individuos de *Plectrohyla Sagorum* se concentran esencialmente en las cercanías del río en la estación seca (invierno). Sin embargo, la temporada de lluvias proporciona suficiente suministro de agua para que los individuos puedan ubicarse

a mayor distancia del río. Esto se observa en el trabajo de campo de manera indirecta, ya que la densidad de individuos en el río disminuye significativamente en la temporada de lluvias. Nuestra hipótesis es que los individuos se concentran principalmente a una distancia de 5 metros del río (en invierno) y que el número de individuos en la región no varía drásticamente durante el año. Usamos nuestro método presentado en el capítulo anterior (y en [19]) para simular la distribución y abundancia de individuos en la estación seca (invierno). Esto nos permite estimar el número total de individuos y usamos esta información para reconstruir distribuciones de probabilidad (constantes de acoplamiento) las restantes temporadas del año (primavera, verano y otoño) haciendo coincidir tanto los datos recopilados en el río como el número total de individuos que calculamos.

3.1.1. Descripción geográfica de la región

La región de estudio está ubicada en el polígono de Chiquihuite al sureste de la reserva natural Volcán Tacaná de la UNESCO [134] y la posición exacta está descrita en la Sección 2.4.1 (ver Figuras 1.1 y 2.1). Contiene un río que se bifurca en dos ramales el cual pertenece a la cuenca del río Coatán (el nacimiento de este río se encuentra en el noroeste del Volcán Tacaná y su desembocadura es el Océano Pacífico). Las altitudes de la región de estudio están en un rango entre 1660 y 2030 metros de altura.

3.1.1.1. Bioma

El bioma específico de la región de estudio es el bosque nuboso. La precipitación es significativamente más alta que en otros tipos de bosque debido a la presencia permanente de nubes de agua que producen precipitación horizontal (también llamada precipitación oculta), que es la fuente de al menos un 5 % (hasta un 20 %) de la precipitación anual [70]. Esto hace que el bosque nuboso sea una de las principales fuentes de agua potable en la región [10]. Solo el 1 % del bosque global está cubierto por bosque nuboso, y el 90 % de él pertenece a las 200 ecorregiones prioritarias para la conservación global del Fondo Mundial para la Naturaleza (*World Wide Fund for Nature*, WWF), ver [116].

En áreas continentales, la presencia de bosque nuboso está restringida a altitudes de 1000 a 3000 metros de altura [10]. Están en grave peligro y la mayor parte del bosque nuboso ha sido destruido y está desapareciendo rápidamente. En México, solo el 1 % del territorio está cubierto por bosque nuboso [3, 134] y solo está presente en la Sierra Madre de Chiapas (que contiene la mayor parte), la meseta central de México y la región de Montañas en Guatemala. Más del 80 % pertenece a zonas perturbadas

[44, 112] y presenta una alta tasa de deforestación [10, 122]. Este es un problema muy importante por varias razones, una de ellas es que su biodiversidad es excepcionalmente alta y contiene un gran porcentaje de especies endémicas. Por ejemplo, contiene 12 % de la flora en México [3].

3.1.1.2. Geología y edafología

Los estratos rocosos son principalmente volcánicos, que consisten en hornblenda, tobas, arenas y brechas volcánicas [134]. El tipo de suelos de la región son yosoles, acrisoles, cambisoles y litosoles, y están permanentemente húmedos, anegados y altamente orgánicos, a menudo con mor humus y turba [70]. Los fluvisoles eutróficos de textura media predominan en la parte baja de la cuenca del río Coatán, sobre una pendiente plana o ligeramente ondulada [45].

3.1.1.3. Climatología

La temperatura media anual de la región de estudio es $20,7^{\circ}\text{C}$, y la precipitación anual es 3674.4 milímetros [134] (la mayor parte de abril a noviembre). El clima es húmedo con fuertes lluvias en verano (la clasificación corresponde a A(c)m(w)ig, [134]). La lluvia invernal representa menos de 5 % de la lluvia anual. La precipitación de los meses más secos (enero) ronda los 32,59 milímetros mientras que en el mes más húmedo (septiembre) alcanza los 687,26 milímetros. Las temperaturas medias mensuales mínima y máxima son $20,3^{\circ}\text{C}$ (diciembre) y $21,1^{\circ}\text{C}$ (abril), respectivamente [134]. En la temporada de invierno, la humedad es del 90 %, mientras que en el resto del año es 50 % [134]. La región está nublada todo el año, debido a los vientos del Golfo de México y los sistemas de nubes originados por los vientos polares.

3.1.2. Descripción de la especie *Plectrohyla Sagorum*

El polígono de Chiquihuite es el hábitat de 32 especies de anfibios de los tres géneros *Plectrohyla*, *Craugastor* y *Bolitoglossa* [4]. Las especies más abundante son *Plectrohyla Sagorum*, *Plectrohyla Matudai* y *Craugastor Pygmaeus* [4]. *Plectrohyla Sagorum* pertenece a la familia *Hylidae* del orden *Anura* [64]. Fue observada por primera vez por E. Hartweg en 1941 [73], quien dio la primera descripción taxonómica de la misma. Esta especie es pequeña (machos 45,5 mm, hembras 51,9 mm) y se caracteriza por las siguientes propiedades físicas: prepolítico espina puntiaguda; tubérculos en las superficies dorsales de la cabeza y extremidades; quilla vertical rostral, hendiduras vocales; dorso marrón opaco con manchas pequeñas e irregulares de color marrón oscuro; flancos bronceados con finas reticulaciones de color marrón oscuro; renacuajo con disco oral

3. DINÁMICA POBLACIONAL DE PLECTROHYLA SAGORUM

no agryado y vainas de mandíbula con vainas cortas estrías romas y procesos laterales cortos [21, 53, 95].

Plectrohyla Sagorum está en peligro porque su distribución está restringida a un área menor de $5,000 \text{ km}^2$, y está altamente fragmentado y constantemente reducido (en Guatemala, El Salvador y México [64]). Esta especie está presente en las laderas de la Sierra Madre del Sur (en la vecindad del Océano Pacífico) en el sur de Chiapas (México), en el sureste de Guatemala y en el norte de El Salvador. Su hábitat está restringido a altitudes entre 1000 y 2050 metros de altura [64] (ver Figura 3.1). Es abundante en dos localidades de los Cuchumatanes, Guatemala. Es muy raro en México, donde se conoce de una sola localidad. Solo se conoce un espécimen de El Salvador.



Figura 3.1: Rango geográfico de *Plectrohyla Sagorum* (fuente: [64])

La principal amenaza para *Plectrohyla Sagorum* es la pérdida de hábitat debido a la deforestación y la transformación de los hábitats forestales originales tanto en el sur de Chiapas como en Guatemala para la agricultura y la tala, plantaciones de madera y asentamientos humanos. Dado que su distribución se encuentra en elevaciones relativamente altas, la enfermedad que se reproduce en los arroyos, la quitridiomycosis, es una posible amenaza futura. Ya ha provocado descensos en otras especies de este género en Centroamérica.

3.1.2.1. Importancia de los anfibios

Los anfibios son un componente importante del bosque nuboso por varias razones. Una de ellas es que pueden modificar niveles tróficos de sus hábitats locales [40, 41, 124, 156, 160], de tal manera que otras especies de flora y fauna dependen de los recursos que generan.

Crisis de extinción de anfibios. En 2020, se estableció que los anfibios son los vertebrados terrestres más amenazados del mundo (el 41 % de las especies evaluadas están amenazadas [1]). De acuerdo a [1], 575 especies están en peligro crítico, 944 especies están en peligro, 638 especies son vulnerables y 130 especies probablemente están extintas (la mayoría de estas extinciones ocurrieron desde 1980 [138]). En 2015, se estimó que una cuarta parte de los anfibios del mundo estaba en peligro [137]. Las investigaciones de 2008 [125] señalan que los anfibios representan el 50 % de las especies amenazadas en las regiones donde viven especies en peligro crítico. El bosque nuboso es el hábitat de una parte importante de las especies de anfibios en el mundo, debido a las temperaturas estables, la humedad, la alta precipitación y las fuentes de agua que presentan. [52, 99, 143, 153]. México ocupa el quinto lugar en diversidad de anfibios en el mundo [56, 58, 77, 117, 131, 147] y el La reserva natural Volcán Tacaná contiene al menos 41 especies de los anfibios, el 55 % de ellos están amenazados [1, 44, 96, 114].

3.2. Datos recopilados

Usamos datos que se reportan en [4] y en [5]. En [4] se muestrearon tres transectos de 25 m en el río. La información precisa que utilizamos en este documento es que, en invierno, no se encontraron individuos a una distancia mayor de 5 metros del río y las siguientes densidades en cada una de las estaciones:

Primavera de 2018

Densidad de individuos sobre el río (Individuos / m)

$$\rho_{Sp} := 0.56.$$

Verano de 2018

Densidad de individuos sobre el río (individuos / m)

$$\rho_{Su} := 0.23.$$

Otoño de 2018

Densidad de individuos sobre el río (Individuos / m)

$$\rho_{Fa} := 0.31.$$

Invierno de 2018

Densidad de individuos sobre el río (Individuos / m)

$$\rho_{\text{Wi}} := 0.81.$$

3.3. Muestreo artificial en el río

En la Sección 3.1.2 dejamos establecidas las densidades de individuos en cada estación del año 2018. Se calcularon a partir de los datos recopilados en un transecto de 75 m. Suponemos que el río es homogéneo y, por lo tanto, los datos recopilados sobre estos 75 m son representativos. Esto nos permite recurrir a un proceso de Poisson para cada temporada con el fin de reconstruir el muestreo artificial que identificamos como datos recopilados (ver Sección 2.4.4.2). Representamos estos datos como configuraciones en el río dadas por:

$$\begin{aligned}\omega^{\text{Sp}} &: R \rightarrow \{0, 1, \dots, K^{\text{Sp}}\}, \\ \omega^{\text{Su}} &: R \rightarrow \{0, 1, \dots, K^{\text{Su}}\}, \\ \omega^{\text{Fa}} &: R \rightarrow \{0, 1, \dots, K^{\text{Fa}}\}, \\ \omega^{\text{Wi}} &: R \rightarrow \{0, 1, \dots, K^{\text{Wi}}\}.\end{aligned}\tag{3.1}$$

Donde $\omega^{\text{Sp}}(p) := \omega_p^{\text{Sp}}$ representa el número de individuos en la celda p del río en la primavera de 2018 y, de manera similar, los otros símbolos proporcionan el número de individuos en cada celda del río en verano (Su), otoño (Fa) e invierno (Wi) de 2018. Además, por cada $x = \text{Sp}, \text{Su}, \text{Fa}, \text{Wi}$, la constante K^x es el máximo número de individuos de la correspondiente simulación de Poisson por temporada x . En nuestras simulaciones obtuvimos lo siguiente: $K^{\text{Sp}} := \text{máx}(\omega^{\text{Sp}}) = 9$, $K^{\text{Su}} := \text{máx}(\omega^{\text{Su}}) = 7$, $K^{\text{Fa}} := \text{máx}(\omega^{\text{Fa}}) = 7$, y $K^{\text{Wi}} := \text{máx}(\omega^{\text{Wi}}) = 11$.

3.4. Hamiltonianos para cada temporada

Definición 3.4.1 (Hamiltonianos para cada temporada). *Para toda $\omega \in \Omega^i$ y toda $x \in \{\text{Sp}, \text{Su}, \text{Fa}, \text{Wi}\}$, definimos*

$$H_0^{(i,x)}(\omega) := \sum_{s_1 \sim s_2} (\omega_{s_1} - \omega_{s_2})^2 + \sum_{\substack{s \sim r \\ r \in R}} (\omega_s - \omega_r^x)^2,\tag{3.2}$$

$$V_{g^{(i,x)}}^{(i,x)}(\omega) = g^{(i,x)} \sum_{s \in \Lambda^i} d_s^2 \omega_s,\tag{3.3}$$

donde $g^{(i,x)}$ es la constante de acoplamiento. Y definimos los Hamiltonianos por temporada sobre las regiones 1, 2, 3 y 4 por

$$H^{(i,x)}(\omega) = H_0^{(i,x)}(\omega) + V_{g^{(i,x)}}^{(i,x)}(\omega). \quad (3.4)$$

Observe que en nuestro modelo, las constantes de acoplamiento $g^{(i,x)}$ dependen de la temporada $x \in \{\text{Sp}, \text{Su}, \text{Fa}, \text{Wi}\}$ y la Región $i \in \{1, 2, 3, 4\}$. Las constantes de acoplamiento se estiman usando datos recopilados en [4] y [5].

Definición 3.4.2. Para toda $\omega \in \Omega^i$ y toda $x \in \{\text{Sp}, \text{Su}, \text{Fa}, \text{Wi}\}$, definimos

$$\pi^{(i,x)}(\omega) := \frac{1}{Z^{(i,x)}} e^{-\frac{1}{T} H^{(i,x)}(\omega)}, \quad (3.5)$$

donde la función partición $Z^{(i,x)}$ está dada por

$$Z^{(i,x)} := \sum_{\omega \in \Omega^i} \exp \left\{ -\frac{1}{T} H^{(i,x)}(\omega) \right\}.$$

3.5. Muestreo de Gibbs

Para cada $x \in \{\text{Sp}, \text{Su}, \text{Fa}, \text{Wi}\}$ y cada $i \in \{1, 2, 3, 4\}$ consideramos la sucesión de muestreo de Gibbs sobre Ω^i , construida de acuerdo a la Sección 2.5.4, denotada por

$$(\omega^{(n;i,x)})_{n \in \mathbb{N} \cup \{0\}}.$$

Recordemos que la importancia del muestreo de Gibbs es que la sucesión $(\omega^{(n;i,x)})_{n \in \mathbb{N}}$, $i \in \{1, 2, 3, 4\}$, $x \in \{\text{Sp}, \text{Su}, \text{Fa}, \text{Wi}\}$, se pueden usar para calcular los valores esperados del número de individuos en cada celda, de acuerdo al Teorema Ergódico 2.5.2,

$$\int \omega_p d\pi^{(i,x)}(\omega) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n \omega_p^{(m;i,x)}, \quad (3.6)$$

para cada celda $p \in \Lambda^i$.

3.6. Selección de T y el Número de Iteraciones.

En la Sección 3.3, establecemos simulaciones de Poisson para el río y obtenemos que el número máximo de individuos para todas las estaciones es 11 (recordamos que nuestro potencial es atractivo y, por lo tanto, hay más individuos en las celdas del río que en otros lugares). La temperatura T y el número de iteraciones se seleccionan de tal manera que comenzando con un número constante de individuos en cada celda del río (y 0 individuos en otros lugares), la sucesión de muestreo de Gibbs correspondiente (para el caso libre: $g = 0$) termina en estados con un número aproximadamente constante de individuos en cada celda de la región, y esto debe suceder para cada constante $K \in \{6, \dots, 11\}$. Nótese que en este caso (caso libre) la energía mínima es 0 y se alcanza en estados que son constantes. En la Sección 3.9.1, informamos la temperatura que obtenemos. El número de iteraciones (para las sucesiones de muestreo de Gibbs) para cada Λ^i , para $i = 1, 2, 3, 4$, se fija en

$$\text{Número de Iteraciones para } \Lambda^i = m_i \times 1000000,$$

donde m_i es el número de celdas de la Región i . Remarcamos que acercarse al mínimo global es posible gracias al método de muestreo que proponemos utilizar, y que introducimos el capítulo anterior (y en [19]), con otros esquemas de muestreo es poco probable que uno se acerque a un mínimo (uno se atasca en los mínimos locales).

3.7. Estimación del número total de individuos y simulaciones en invierno

Los individuos de *Plectrohyla Sagorum* están adheridos a cuerpos de agua y se encuentran únicamente en las cercanías del río. Sin embargo, no son igualmente abundantes en el río durante todo el año. En invierno, la densidad de individuos en el río alcanza su máximo y disminuye hasta el verano, cuyo el número de individuos en el río alcanza su mínimo. Este comportamiento coincide con las características lluviosas de la región: el invierno es la estación seca y el verano es época de lluvias. La lluvia aumenta de invierno a verano y disminuye de verano a invierno. Nuestra hipótesis sobre este comportamiento es que los individuos permanecen en lugares alejados del río porque los suministros de agua de la lluvia satisfacen parcialmente sus necesidades. Por lo tanto, asumimos que el número total de individuos en cada región permanece esencialmente constante, y la variación de densidad implica que los individuos se dispersan lejos del

río. El invierno es la temporada en la que los datos recolectados son más confiables porque la ausencia de lluvias facilita considerablemente el trabajo de campo. De acuerdo a [4], en esta temporada, solo se observaron individuos raramente a 5 metros del río (en la Región 4, donde se realizaron las observaciones). Luego, construimos nuestras simulaciones con la restricción de que no hay individuos a distancias mayores de 5 metros del río, en invierno.

3.7.1. Selección de la constante de acoplamiento $g^{(i,Wi)}$

Las constantes de acoplamiento $g^{(i,Wi)}$ se eligen de la siguiente manera:

Generamos simulaciones de muestreo de Gibbs, con temperatura e iteraciones fijas de acuerdo con la Sección 3.6 para diferentes valores de $g^{(4,Wi)}$, comenzando con $g^{(4,Wi)} = 0$ y aumentando lentamente hasta alcanzar los promedios de las sucesiones de muestreo de Gibbs entre los últimos $m_4 \times 10000$ iteraciones ($m_4 = 13715$ es el número de celdas en la Región Λ^4) no producen individuos dentro de una distancia mayor que 5 metros hasta el río (en todas las simulaciones que realizamos). La constante de acoplamiento $g^{(4,Wi)}$ se selecciona como el valor más pequeño de $g^{(4,Wi)}$ con este requisito. Luego elegimos $g^{(i,Wi)} = g^{(4,Wi)}$, $i \in \{1, 2, 3\}$. Recordar que las observaciones se obtuvieron de la Región 4 ([4]).

3.7.2. Número total de individuos

Una vez que se determinan las constantes de acoplamiento $g^{(i,Wi)}$, junto con la temperatura T y el número de iteraciones, ejecutamos una simulación de muestreo de Gibbs y calculamos los promedios del número de individuos en cada celda. Esto y el teorema ergódico (Teorema 4.7.1) nos permiten estimar el número esperado de individuos en cada celda y, por lo tanto, el número total de individuos en cada región (usamos los últimos $m_i \times 10^4$ iteraciones en las sucesiones de muestreo de Gibbs - recordamos que siempre calculamos $m_i \times 10^6$ iteraciones -, donde m_i es el número de celdas en la Región Λ^i . Mantenemos estos números fijos y los usamos para predecir la distribución y abundancia de individuos en las otras temporadas.

3.8. Dinámica de la población durante el año

Aquí proporcionamos simulaciones de la distribución y abundancia de individuos en cada estación, excepto el invierno que se analiza anteriormente. las constantes de acoplamiento $g^{(i,x)}$, para $x \in \{\text{Sp}, \text{Su}, \text{Fa}\}$ (primavera, verano y otoño) se eligen para

ajustar el número total de individuos calculados en la Sección 3.7.2 y los datos recolectados sobre el río, según cada temporada, expresados en (3.1). Observe que los hamiltonianos $H^{(i,x)}$ ya contienen la información de los datos recopilados sobre el río.

3.8.1. Selección de las constantes de acoplamiento

Elegimos las constantes de acoplamiento $g^{(i,x)}$ de la siguiente manera: Ejecutamos simulaciones de muestreo de Gibbs $(\omega^{(n;i,x)})_{n=1}^{m_i \times 10^6}$, $i \in \{1, 2, 3, 4\}$, $x \in \{\text{Sp}, \text{Su}, \text{Fa}\}$, con la temperatura y las iteraciones fijadas de acuerdo con la Sección 3.6 para diferentes valores de $g^{(i,x)}$, comenzando con $g^{(i,x)} = 0$ y aumentando lentamente hasta que los promedios de las últimas $m_i \times 10000$ iteraciones de las sucesiones de muestreo de Gibbs producen el número de individuos, en cada región y cada temporada, establecido en la Sección 3.7.2.

3.9. Resultados

3.9.1. Calibración de la temperatura y el número de iteraciones (caso libre)

Aquí presentamos la temperatura que elegimos y el número de iteraciones que necesitamos para acercarnos al mínimo global, según el método descrito en el capítulo anterior (ver también [19]). Encontramos que la temperatura

$$T = 0.2 \tag{3.7}$$

funciona muy bien, y como se indica en la Sección 3.6, tomamos $m_i \times 1000000$ iteraciones para las sucesiones de muestreo de Gibbs, donde m_i es el número de celdas en Λ^i , para cada $i = 1, 2, 3, 4$. En las Figuras 3.2, 3.3, 4.3 y 3.5, mostramos el decaimiento de la energía para las situaciones descritas en la Sección 3.6, aquí solo reportamos las gráficas para $K = 6$ y $K = 11$. Las gráficas muestran los valores de energía del sistema completo para iteraciones de la forma $n = m_i k$ con $k = 1, \dots, 1000000$. Las Figuras 3.3 y 3.5 muestran los valores de energía para los primeros 2 millones de iteraciones del algoritmo. Estas cifras muestran una rápida disminución al mínimo de la función de energía. En las últimas iteraciones de $m_i \times 10000$ la energía está muy cerca de cero, con variaciones mínimas.

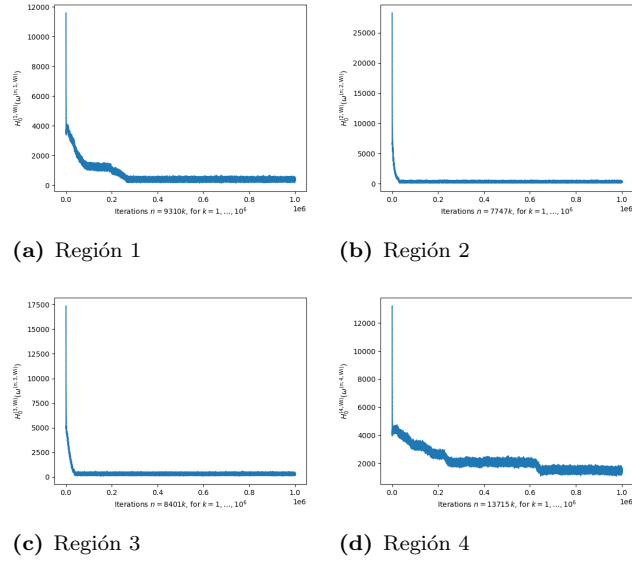


Figura 3.2: Selección de T . Las gráficas muestran cómo decae la energía con un número fijo de individuos $K = 6$ en las celdas del río R , temperatura $T = 0.2$ y $g = 0$. Los gráficos muestran los valores de la energía del sistema en las iteraciones de la forma $n = m_i k$, para $k = 1, \dots, 10^6$, donde m_i es el número de celdas para cada Región i , para $i = 1, 2, 3, 4$.

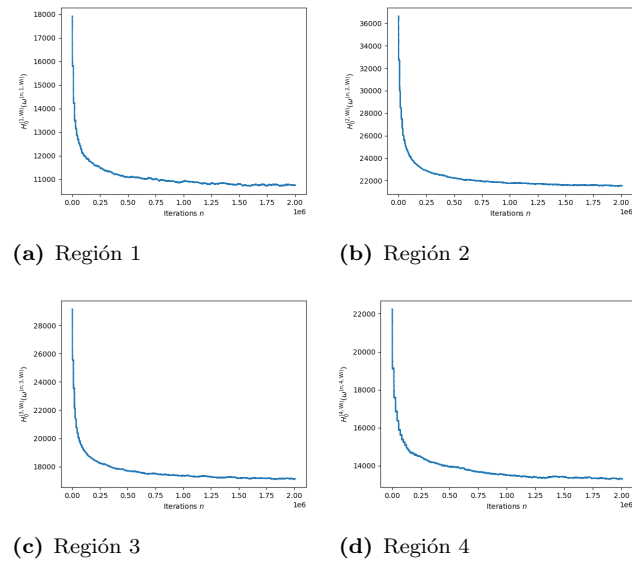


Figura 3.3: Selección de T . Los gráficos muestran cómo decae la energía con un número fijo de individuos $K = 6$ en las celdas del río R , temperatura $T = 0.2$ y $g = 0$, para el primeros 2 millones de iteraciones del muestreo de Gibbs.

3. DINÁMICA POBLACIONAL DE PLECTROHYLA SAGORUM

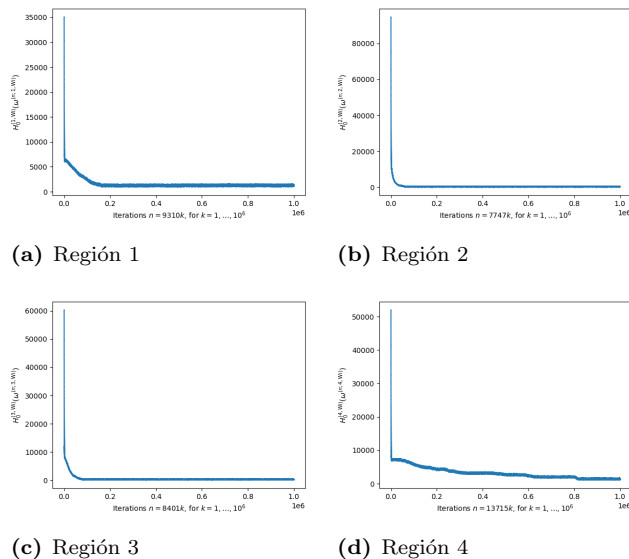


Figura 3.4: Selección de T . Los gráficos muestran cómo decae la energía con un número fijo de individuos $K = 11$ en las celdas del río R , temperatura $T = 0.2$ y $g = 0$. Los gráficos muestran los valores de la energía del sistema en las iteraciones de la forma $n = m_i k$, para $k = 1, \dots, 10^6$, donde m_i es el número de celdas para cada Región i , para $i = 1, 2, 3, 4$.

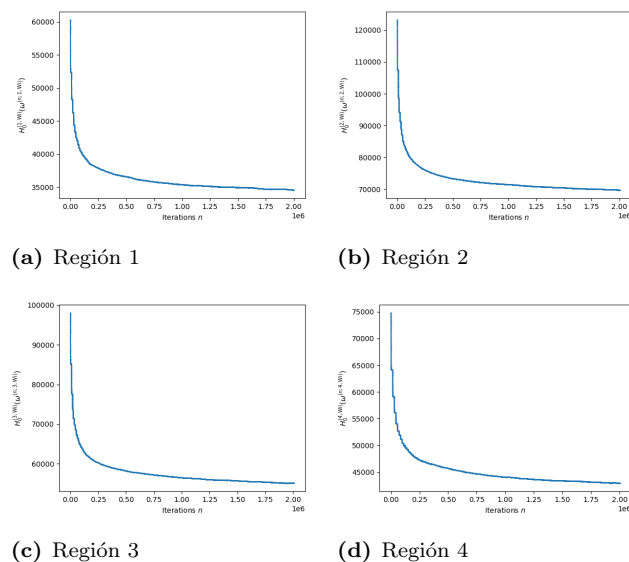


Figura 3.5: Selección de T . Los gráficos muestran cómo decae la energía con un número fijo de individuos $K = 11$ en las celdas del río R , temperatura $T = 0.2$ y $g = 0$, para los primeros 2 millones de iteraciones del muestreo de Gibbs.

3.9.2. Las constantes de acoplamiento $g^{(i,Wi)}$ y el número total de individuos por región en invierno

El invierno es la estación seca y, a partir del trabajo de campo [4, 5], creemos que los individuos permanecen dentro de una distancia horizontal de 5 metros del río y, por lo tanto, seleccionamos $g^{(4,Wi)}$ de tal manera que solo en las celdas del río R y sus vecinas contienen individuos. Luego elegimos $g^{(i,Wi)} = g^{(4,Wi)}$, $i \in \{1, 2, 3\}$ (ver Sección 3.7.1). Obtenemos

$$g^{(i,Wi)} = 2, \quad i \in \{1, 2, 3, 4\}.$$

En las Figuras 3.12 y 3.13, mostramos las gráficas de las energías de la sucesiones $(\omega^{(n;i,Wi)})_{n=1}^{m_i \times 10^6}$, para $i \in \{1, 2, 3, 4\}$. La Figura 3.13 muestra los valores de la energía de las primeras 2000000 iteraciones del algoritmo. Es claro a partir de estas figuras que la energía se estabiliza rápidamente. Usando estas constantes de acoplamiento, obtenemos los valores esperados para el número de individuos por región en invierno (ver Tabla 3.1).

Región	Número esperado de individuos (Wi)
Λ^1	1028.33
Λ^2	2198.16
Λ^3	1516.26
Λ^4	1208.29

Cuadro 3.1: Número esperado de individuos por región en invierno (Wi).

3.9.3. las constantes de acoplamiento $g^{(i,x)}$, $x \in \{\text{Sp}, \text{Su}, \text{Fa}\}$

Las constantes de acoplamiento $g^{(i,x)}$, $x \in \{\text{Sp}, \text{Su}, \text{Fa}\}$, $i \in \{1, 2, 3, 4\}$, están dadas en la Tabla 3.2. Las Figuras 3.6, 3.7, 3.8, 3.9, 3.10, 3.11, 3.12, y 3.13 muestran las gráficas de las energías para iteraciones de la forma $n = m_i k$, para $k = 1, \dots, 1000000$ (Figuras 3.6, 3.8, 3.10, 3.12) y para las primeras 2000000 de iteraciones (Figuras 3.7, 3.9, 3.11, 3.13) en regiones $\Lambda^1, \dots, \Lambda^4$ para cada temporada. Las energías se estabilizan rápidamente presentando variaciones mínimas.

Región	$g^{(i,\text{Sp})}$	$g^{(i,\text{Su})}$	$g^{(i,\text{Fa})}$
Λ^1	0.18	0.1	0.12
Λ^2	0.24	0.1	0.12
Λ^3	0.2	0.05	0.1
Λ^4	0.15	0.06	0.07

Cuadro 3.2: Las constantes de acoplamiento $g^{(i,x)}$, $x \in \{\text{Sp}, \text{Su}, \text{Fa}\}$ y $i = 1, 2, 3, 4$.

3. DINÁMICA POBLACIONAL DE PLECTROHYLA SAGORUM

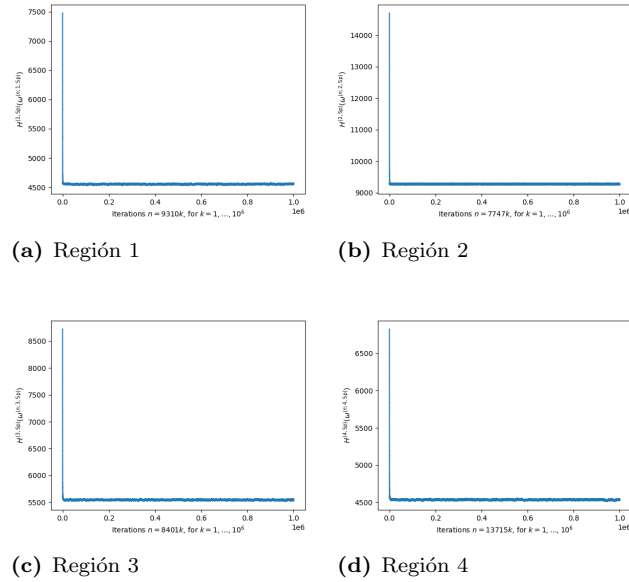


Figura 3.6: Estabilización de la energía sobre primavera (Sp). las gráficas muestran los valores de la energía del sistema en las iteraciones de la forma $n = m_i k$, para $k = 1, \dots, 10^6$, donde m_i es el número de celdas para cada región i , para $i = 1, 2, 3, 4$.

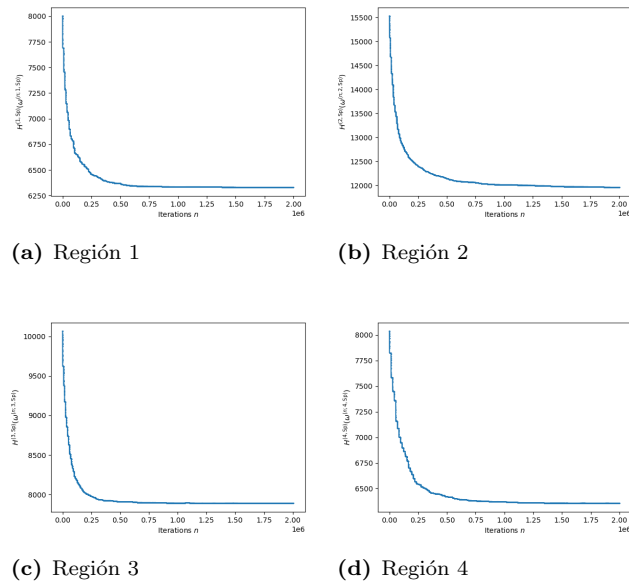


Figura 3.7: Estabilización de la energía en primavera (Sp). Los gráficos muestran los primeros dos millones de iteraciones del muestreo de Gibbs.

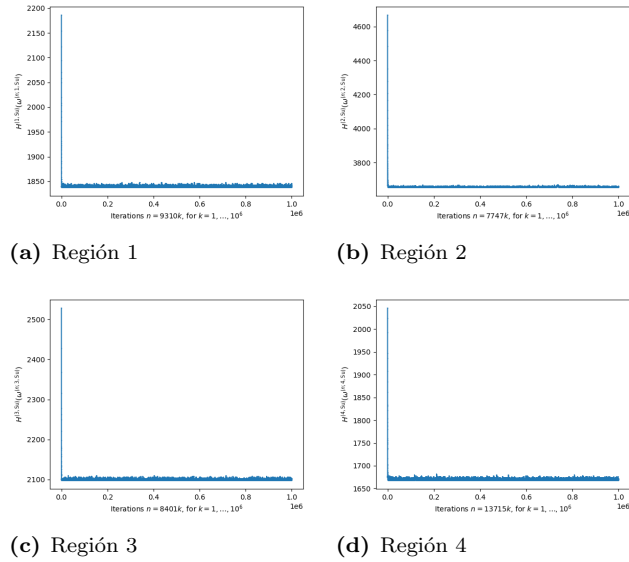


Figura 3.8: Estabilización de la energía sobre verano (Su). Las gráficas muestran los valores de la energía del sistema en las iteraciones de la forma $n = m_i k$, para $k = 1, \dots, 10^6$, donde m_i es el número de celdas para cada región i , para $i = 1, 2, 3, 4$.

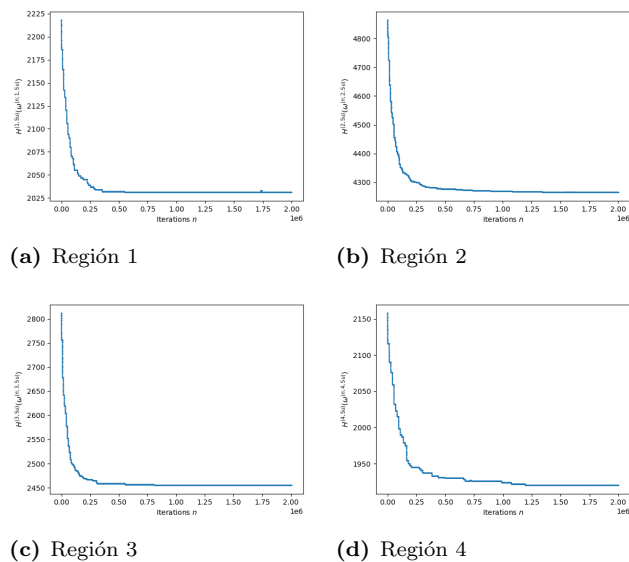


Figura 3.9: Estabilización de la energía en verano (Su). Los gráficos muestran los primeros dos millones de iteraciones del muestreo de Gibbs.

3. DINÁMICA POBLACIONAL DE PLECTROHYLA SAGORUM

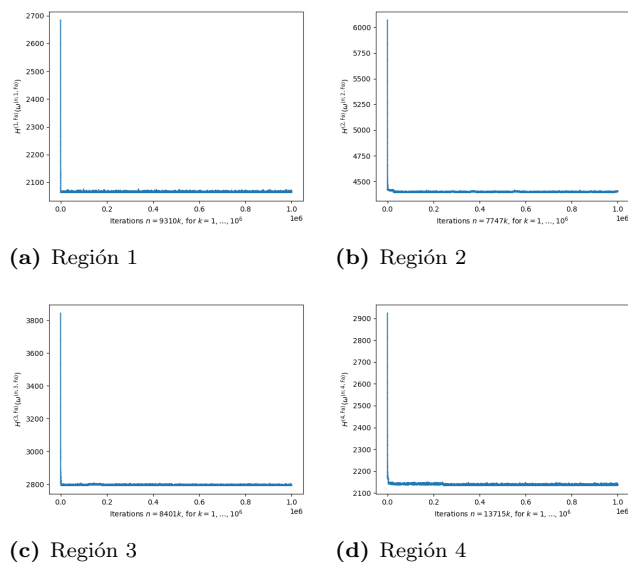


Figura 3.10: Estabilización de la energía en otoño (Fa). Las gráficas muestran los valores de la energía del sistema en las iteraciones de la forma $n = m_i k$, para $k = 1, \dots, 10^6$, donde m_i es el número de celdas para cada región i , para $i = 1, 2, 3, 4$.

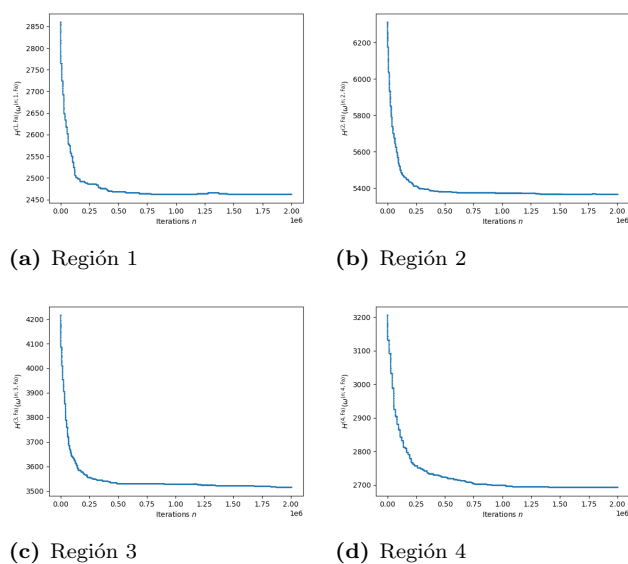


Figura 3.11: Estabilización de la energía en otoño (Fa). Los gráficos muestran los primeros dos millones de iteraciones del muestreo de Gibbs.

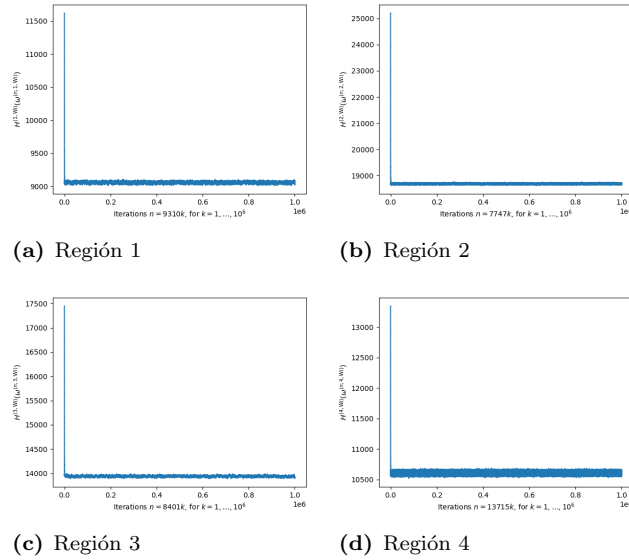


Figura 3.12: Estabilización de la energía en invierno (Wi). Las gráficas muestran los valores de la energía del sistema en las iteraciones de la forma $n = m_i k$, para $k = 1, \dots, 10^6$, donde m_i es el número de celdas para cada región i , para $i = 1, 2, 3, 4$.

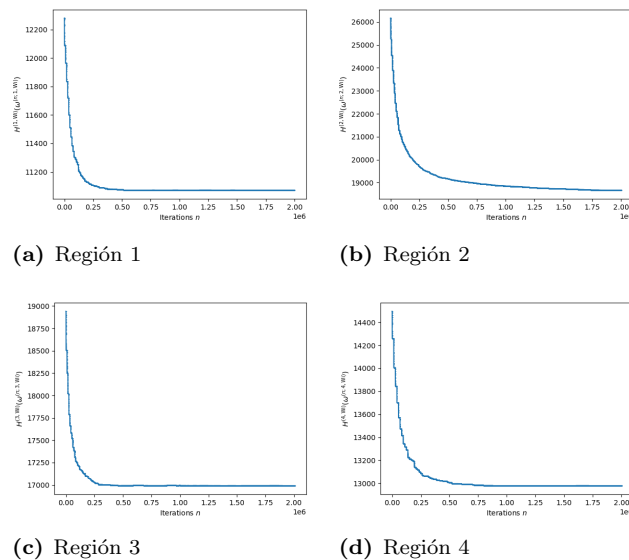


Figura 3.13: Estabilización de la energía en invierno (Wi). Los gráficos muestran los primeros dos millones de iteraciones del muestreo de Gibbs.

3.9.3.1. Número total de individuos

Para cada $i \in \{1, 2, 3, 4\}$ y cada temporada $x \in \{\text{Sp}, \text{Su}, \text{Fa}, \text{Wi}\}$ tomamos las últimas $m_i \times 10000$ iteraciones del muestreo de Gibbs

$$(\omega^{(n;i,x)})_{n=m_i \times 990000+1}^{m_i \times 10^6},$$

de la sucesión de $m_i \times 10^6$ iteraciones

$$(\omega^{(n;i,x)})_{n=0}^{m_i \times 10^6}.$$

Luego, estimamos el valor esperado del número de individuos promedio de las últimas $m_i \times 10000$ iteraciones usando el Teorema Ergódico 2.5.2 (Sección 3.7), ver Tabla 3.3. Se seleccionan las constantes de acoplamiento $g^{(i,x)}$ de tal manera que el número total de individuos se mantenga casi constante para cada temporada, ver Tabla 3.3. En la Tabla 3.4 se muestran las diferencias del valor esperado del número de individuos con respecto a los valores en invierno (que son los valores de referencia). Estas diferencias oscilan entre 9.8111 y 41.6925, lo que implica que el porcentaje de variación con respecto al invierno es menor al 3%, ver Tabla 3.5.

Región	Sp	Su	Fa	Wi
Λ^1	1056.95	1047.67	1047.22	1028.33
Λ^2	2188.35	2221.69	2234.57	2198.16
Λ^3	1551.78	1540.24	1557.96	1516.26
Λ^4	1229.23	1236.27	1222.42	1208.29

Cuadro 3.3: Número esperado de individuos por temporada y región.

Región	Sp	Su	Fa
Λ^1	28.61	19.34	18.88
Λ^2	9.81	23.53	36.41
Λ^3	35.51	23.98	41.69
Λ^4	20.93	27.98	14.12

Cuadro 3.4: Diferencias en el número esperado de individuos con respecto al invierno

Región	Sp	Su	Fa
Λ^1	2.78 %	1.88 %	1.84 %
Λ^2	0.45 %	1.07 %	1.66 %
Λ^3	2.34 %	1.58 %	2.75 %
Λ^4	1.73 %	2.32 %	1.17 %

Cuadro 3.5: Porcentaje de las diferencias en el número esperado de individuos con respecto al invierno

3.9.3.2. Resultados principales: distribución y abundancia de individuos

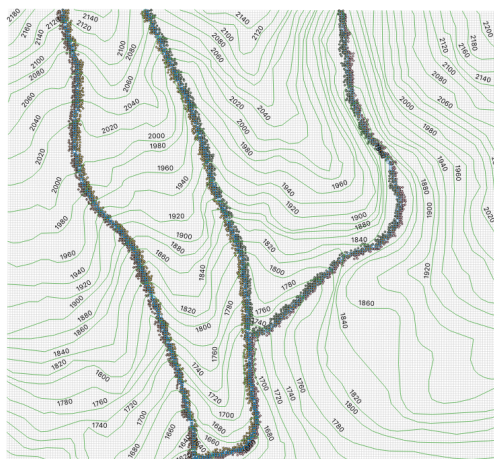
Como explicamos en la Sección 3.9.3.1, para la estación invierno (Wi) calculamos $m_4 \times 10^6$ iteraciones de nuestro muestreo de Gibbs y tomamos las últimas $m_4 \times 10^4$ iteraciones para estimar el valor esperado del número de individuos en cada celda, usando teoremas ergódicos. Nuestros resultados se presentan en varios mapas:

- En la Figura 3.14, mostramos una representación gráfica de los valores esperados del número de individuos en cada celda, usando un punto para representar a cada individuo. Vemos claramente que los individuos se concentran en el río en invierno (la estación seca) y se dispersan a medida que llega la lluvia a la región (primavera). En verano, cuyo la precipitación alcanza su máximo, la dispersión de los individuos es máxima. En otoño, los individuos vuelven a acercarse al río y su distribución es similar a la de la primavera.
- En la Figura 3.15 usamos un mapa de calor para representar la densidad y abundancia de individuos. Usamos los colores de 7 del arco iris: el color más energético (violeta) representa la densidad más alta y el color menos energético (rojo) representa la densidad más baja. En el medio, un valor ascendente de las densidades está representado (en orden) por los colores, naranja, amarillo, verde, cian y azul. Para tener una representación gráfica aún mejor, aumentamos adicionalmente la intensidad de los colores a medida que aumenta la densidad (de tal manera que el rojo corresponde a la menor intensidad y el azul a la mayor intensidad). Los colores y las correspondientes intensidades se muestran en la Tabla 3.6. En la Figura 3.16 proporcionamos un primer plano para facilitar la visibilidad. En las Figuras 3.15 y 3.16 (a diferencia de las otras figuras de este capítulo), los límites de las celdas del río son negros (de lo contrario, el río sería difícil distinguir). De nuestros mapas de calor, podemos ver claramente que la densidad de individuos es máxima en el río y disminuye a medida que el la distancia al río aumenta.

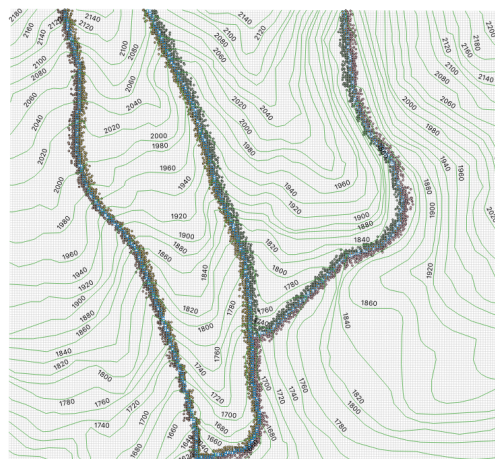
Cuadro 3.6: Degradación de la intensidad de color y correspondiente número esperado de individuos

■	(0, 1.5714]
■	(1.5714, 3.1429]
■	(3.1429, 4.7143]
■	(4.7143, 6.2857]
■	(6.2857, 7.8571]
■	(7.8571, 9.4286]
■	(9.4286, 11]

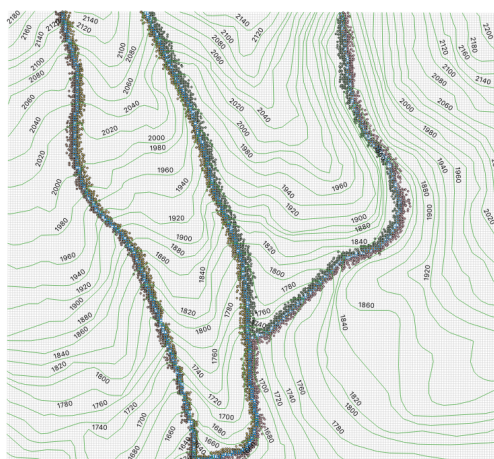
3. DINÁMICA POBLACIONAL DE PLECTROHYLA SAGORUM



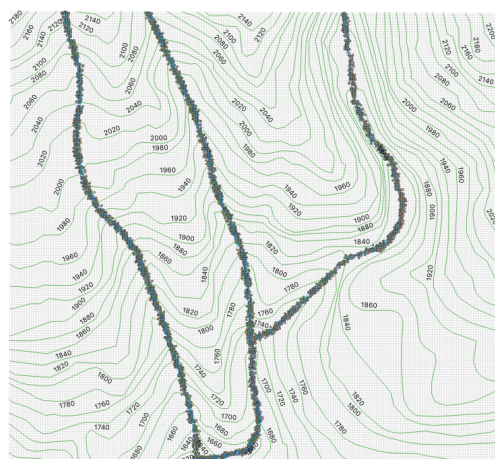
(a) Primavera 2018



(b) Verano 2018

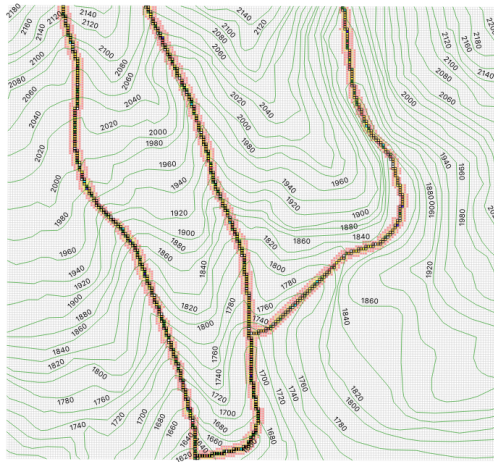


(c) Otoño 2018

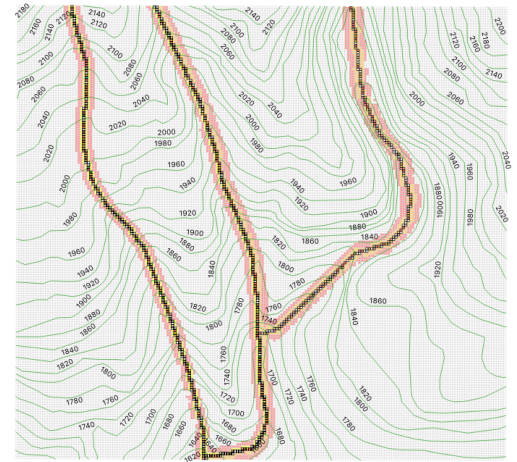


(d) Invierno 2018

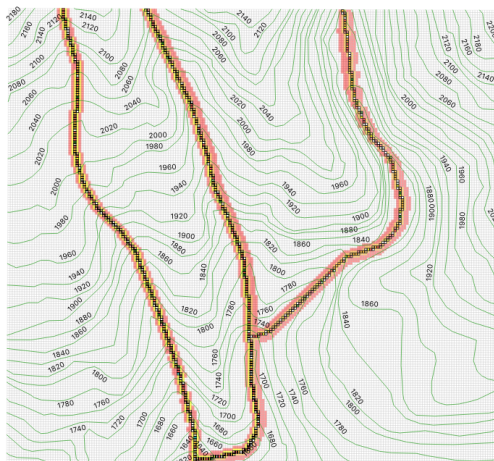
Figura 3.14: Distribución y abundancia por temporada de la especie *Plectrohyla sagorum*.



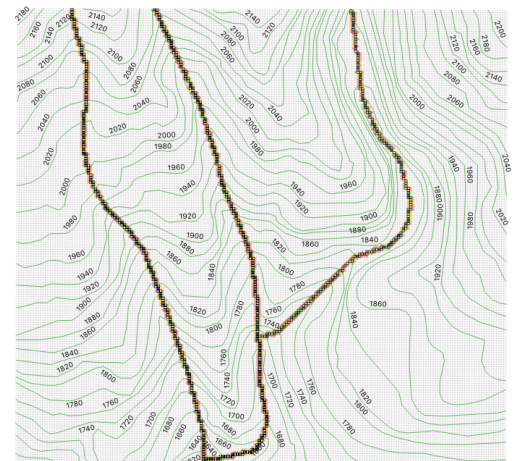
(a) Primavera 2018



(b) Verano 2018



(c) Otoño 2018



(d) Invierno 2018

Figura 3.15: Mapa de calor de la densidad de individuos por temporada de la especie *Plectrohyla sagorum*.

3. DINÁMICA POBLACIONAL DE PLECTROHYLA SAGORUM

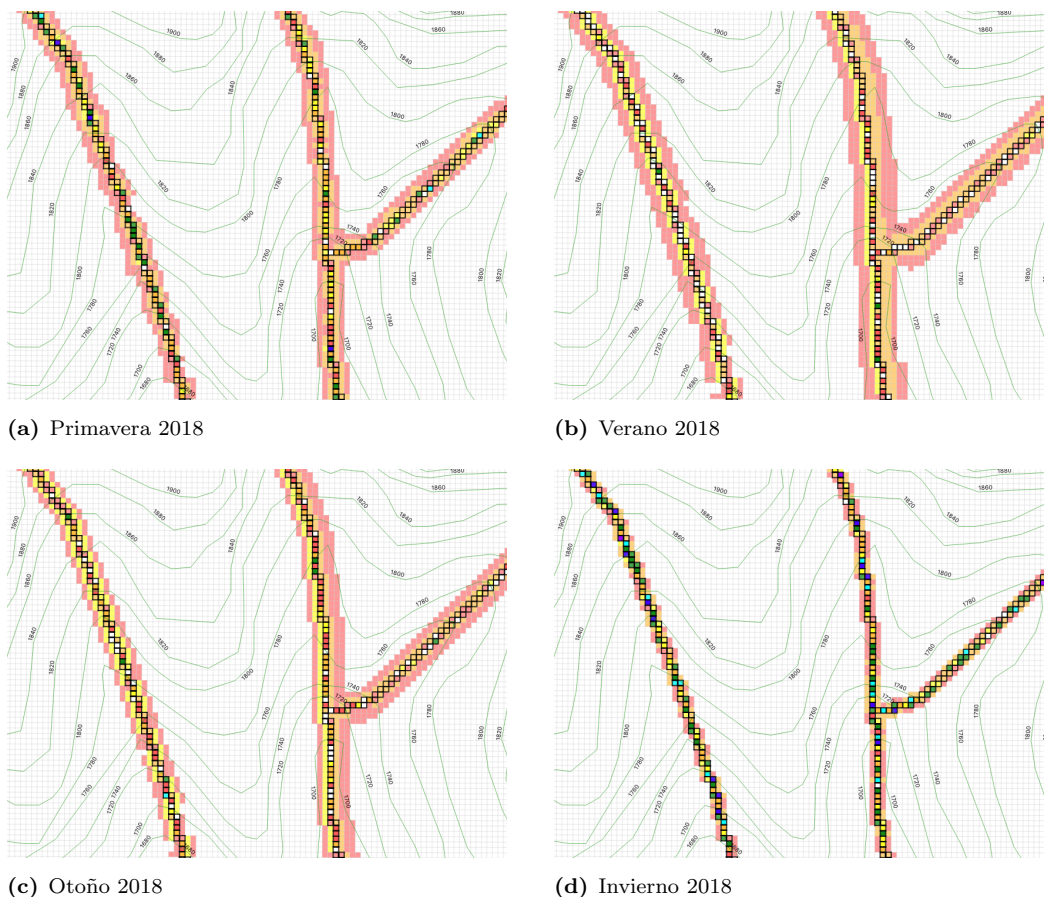


Figura 3.16: Detalle del mapa de calor de la densidad de individuos por temporada de la especie *Plectrohyla Sagorum*.

3.10. Conclusiones

Usamos el modelo que presentamos en el capítulo anterior (ver también [19]) para estimar la distribución y abundancia de individuos de *Plectrohyla Sagorum* en nuestra región de estudio. Agregamos nuevas características que nos permiten abordar la dinámica de la población. En particular, el trabajo de campo sugiere que en la estación seca los individuos se ubican en las cercanías del río. Hacemos uso de esta información junto con una estimación de la distribución y abundancia en el río (derivada de los datos recopilados en [4] y [5]), aplicamos nuestro método en [19] para estimar el número total de individuos en la región. Esto nos permite utilizar simulaciones de distribución y abundancia en el río (obtenidas de los datos recopilados en [4] y [5]) con el fin de reconstruir la distribución y abundancia de la región completa en cada temporada de el

año. Nuestro método difiere de trabajos anteriores porque controlamos las medidas de distribución conjunta y la aproximación a los mínimos globales (en lugar de las probabilidades condicionales locales y los mínimos locales, que es el enfoque habitual). Con este modelo somos capaces entonces de abordar el problema difícil (un control global, en lugar de un control local) porque utilizamos ideas de atracción y difusión inspiradas en la mecánica cuántica y un esquema de muestreo inteligente.

Modelado de distribución de baja densidad y especies en peligro crítico: Plectrohyla Guatemalensis

4.1. Introducción

El objetivo de este capítulo es doble, primero, contribuimos a la comprensión de la distribución y abundancia de una especie en peligro crítico, llamada *Plectrohyla Guatemalensis*, y segundo, presentamos algunas herramientas matemáticas que pueden usarse para estudiar la abundancia y distribución de especies que rara vez se observan en la naturaleza. La descripción geográfica de la región se detalla en la Sección 2.4.1. Nuestro estudio se centra en una cuadrícula rectangular con celdas cuadradas de $5m$ de longitud de lado (ver Figura 4.2). El trabajo de campo se reporta en [4], y consistió en 10 viajes de campo durante el año 2018, donde se muestreó una porción muy pequeña de la región de estudio (aquí proporcionamos muestreos simulados para toda la región). Nuestros principales resultados se resumen en las Figuras 4.5, 4.6 y 4.7. En la Figura 4.5, proporcionamos un mapa de calor donde se representan los promedios de un año de los valores esperados del número de individuos en cada celda de la cuadrícula. En la Figura 4.7, representamos mediante puntos las observaciones simuladas de individuos (cada punto es un individuo) considerando promedios durante el año. En la Figura 4.6, usamos puntos para describir simulaciones de los individuos que podrían observarse después de 10 viajes de campo durante el año (las fechas del trabajo de campo se informan en [4]). Los resultados de este capítulo están reportados en [17]

4.2. El modelo aplicado a especies de baja densidad

En el Capítulo 2 explicamos con detalle nuestro modelo matemático, el cual utiliza medidas de Gibbs (campos aleatorios de Markov) para definir un Hamiltoniano que está dado como la suma de un término de energía cinética y un término de energía potencial. El primero describe la expansión de los individuos y el segundo las fuerzas de atracción hacia el río. El equilibrio entre esparcirse y ser atraído por el río determina la distribución de los individuos y los datos recogidos fijan los parámetros de energía potencial. La idea original de usar campos de aleatorios de Markov [25, 26] es que los individuos tienden a comportarse como sus vecinos, de tal manera que una vez que se fija la distribución de las celdas vecinas de una celda específica en la cuadrícula, se determina la distribución de probabilidad completa para esa celda. En el caso de las especies en peligro crítico, las observaciones son extremadamente raras (generalmente hay menos de dos individuos para cada muestreo de campo) y, por lo tanto, los datos recopilados se limitan a unos pocos individuos. Los lugares donde se encuentran estos individuos no tienen características especiales que los atraigan, se encuentran por casualidad. A partir de un par de observaciones, es imposible reconstruir un mapa de distribución en la forma en que se lleva a cabo en la literatura de geoestadística (ver [13, 25, 26, 35, 42, 61, 84, 105, 148, 149]), y se deben introducir nuevas características al modelo detallado en el Capítulo 2 (ver también [19]).

El mayor problema es que solo a partir de las propiedades locales de los campos aleatorios de Markov (como en [25, 26, 43, 49, 66, 89, 92, 101, 135, 136, 163]), lo máximo que podemos obtener son agrupaciones alrededor de estas dos observaciones. Estos conglomerados no son razonables porque, como mencionamos anteriormente, no hay características en estos lugares que puedan atraer a los individuos y, obviamente, estas observaciones no se repetirían si se realizara un nuevo trabajo de campo. Nuestra solución al problema es la siguiente: utilizamos datos recolectados de [4], donde la recolección de individuos se realizó 10 veces durante el año, y el resultado fue 17 individuos en el río y 3 fuera del río. Estas observaciones se llevaron a cabo en 15 celdas de río de la cuadrícula y 5 celdas fuera del río.

Calculamos la densidad promedio (durante el año) de individuos en celdas de río R y la tomamos como datos fijos, que denotamos por $K > 0$. Las variables aleatorias que usamos no toman valores en el número de individuos (como lo hacemos en los capítulos 2 y 3 (ver también [19] y [20]), sino que toman valores del número promedio anual esperado de individuos por celda, que son números en el intervalo $[0, K]$. Asumimos que el río es homogéneo y, por lo tanto, los valores de las variables aleatorias asociadas con las celdas del río son constantes e iguales a K . Entonces, tomamos probabilidades con-

dicionales dados los datos del río. Utilizamos el modelo que presentamos en el Capítulo 2 (ver también [19]). Las variables aleatorias se distribuyen según una medida de Gibbs (ver Sección 2.4.3). El generador de esta medida es un Hamiltoniano del sistema.

En el régimen de baja temperatura, las configuraciones más probables son las menos energéticas. Como mencionamos anteriormente, nuestro hamiltoniano es la suma de un término cinético que promueve la propagación y un término de energía potencial que atrae a las personas al río. Los estados menos energéticos presentan un equilibrio entre extenderse y ser atraídos por el río y la energía potencial se fija utilizando datos de muestreo en las celdas que no pertenecen al río. En nuestro caso, solo se muestrearon 5 de tales celdas y, por lo tanto, usamos una forma simple del potencial inspirado en el oscilador armónico que depende solo de un parámetro al que llamamos constante de acoplamiento. Esta constante de acoplamiento se estima utilizando que el valor esperado en las celdas donde se recopilaron los datos debe coincidir con el valor esperado predicho por nuestro modelo, esta es una característica nueva en relación con el modelo del Capítulo 2 y la aplicación del Capítulo 3 (ver también [19] y [20]).

4.3. Descripción de la especie *Plectrohyla Guatemalensis*

La descripción del hábitat de la región de estudio se detalla en la Sección 3.1. En esa misma sección se explica la importancia de los anfibios en la naturaleza y el problema relevante que enfrentan en términos de extinción. La especie *Plectrohyla Guatemalensis* pertenece a la familia Hylidae del orden Anura [53, 57, 63]. El hábitat de esta especie es el bosque nuboso desde la Sierra Madre de Chiapas, hasta las regiones altas del sureste de Guatemala y las montañas del noreste de El Salvador y Honduras (y también se encuentra en el norte de Nicaragua [55]). Solo está presente en la franja altitudinal de 900 a 2800 metros de altura [63, 78, 95, 109, 156] - ver Figura 4.1). La primera observación de *Plectrohyla Guatemalensis* fue reportada por P. Brocchi en 1887, en la localidad de Chimaltenango, Guatemala [30]. En la reserva natural Volcán Tacaná (Soconusco, Chiapas), esta especie fue observada por primera vez en 1944 (ver [32])



Figura 4.1: Rango geográfico de *Plectrohyla Guatemalensis* (Fuente: [63]).

4. MODELADO DE DISTRIBUCIÓN DE BAJA DENSIDAD Y ESPECIES EN PELIGRO CRÍTICO: PLECTROHYLA GUATEMALENSIS

Plectrohyla Guatemalensis es una especie rara de anfibio que es pequeño pero con cuerpo robusto, con un rango de longitud hocico de 40,0 a 52,1 mm para machos adultos y un rango de 42,1 a 54,1 mm para hembras adultas. Las fosas nasales ligeramente protuberantes están colocadas dorsolateralmente y más cerca del ojo que del hocico. El canto es casi recto con una cresta angular. La región loreal es cóncava, pero las regiones internarial e interorbital son planas. La distancia entre el ojo y la fosa nasal es aproximadamente el 80 % del diámetro del ojo grande. El tímpano y el anillo timpánico son indistintos. Dedos largos y discos terminales grandes y redondos. El tercio superior del tímpano está cubierto por un pliegue supratimpánico pesado que se extiende desde la esquina posterior del ojo hasta la inserción del brazo. Los labios son redondeados y en los machos reproductores hinchados. Los individuos de esta especie no tienen hendiduras vocales ni sacos vocales. La espina prepolical bífida y los tubérculos redondos en el dorso distinguen a este género de ranas de otros *Plectrohyla* (ver [51, 53]).

El hábitat de *Plectrohyla Guatemalensis* incluye arroyos de montaña en ambientes arbóreos y acuáticos. Durante el día, se pueden encontrar adultos en grietas cerca de arroyos y bromelias arbóreas; de noche, se pueden encontrar adultos en las orillas de los arroyos y en las rocas cerca de los arroyos [53]. Esta especie está en “peligro crítico”, de la lista roja de la UICN [63]. Esto implica una disminución esperada de la población de al menos el 80 % en los próximos años debido a la fragmentación, la pérdida de hábitat, la hibridación, la competencia con especies introducidas, la contaminación, los parásitos y la quitridiomycosis especial producida por el hongo *batrachochytrium dendrobatidis* [110, 111, 114, 132, 145]. La población de esta especie disminuyó significativamente en México, Guatemala y El Salvador, y es más abundante en Honduras [111, 132].

4.4. Daros recolectados

Usamos datos de [4], donde se realizaron 10 viajes de campo durante el año 2018. Las fechas específicas en las que se recopilaban los datos son: 2 de febrero, 8 de febrero, 1 de marzo, 8 de marzo, 6 de junio, 5 de julio, 19 de agosto, 26 de agosto, 7 de noviembre y 18 de diciembre. Se estudió un transecto fijo de 75m en el río y 5 parcelas cuadradas sobre la Región Λ^4 (ver Figura 4.2, en donde las ubicaciones de las parcelas se muestran en color verde oscuro), con una longitud de lado de 5m, y se registraron las observaciones de los individuos durante cada viaje de campo. El número total de individuos registradas en el río es 17 y el número correspondiente a las parcelas es 3. Las coordenadas de las parcelas están Tabla 4.1.

Parcelas	UTM Coord.	Retícula Coord.
1	(595008.5, 1668868.5)	(133, 126)
2	(595048.5, 1668903.5)	(126, 134)
3	(595068.5, 1668948.5)	(117, 138)
4	(595113.5, 1668973.5)	(112, 147)
5	(595163.5, 1668998.5)	(107, 157)

Cuadro 4.1: Coordenadas de las parcelas

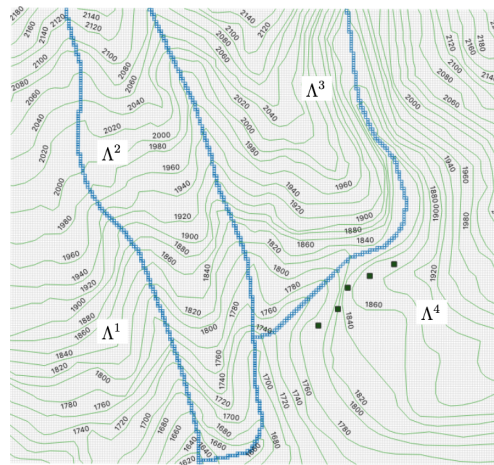


Figura 4.2: Regiones Λ^1 , Λ^2 , Λ^3 , Λ^4 y el Río R (azul). Las parcelas se muestran en color verde oscuro sobre la Región Λ^4

4.5. Datos del río

Denotamos por K la densidad promedio anual en el río por celda. Esto se calcula de la siguiente manera: En [4] se informa que se muestrearon $75m$ en el río y el ancho del río es de aproximadamente $5m$. Entonces, consideramos que esta parte del río contiene 15 celdas. La densidad promedio anual por celda es solo el número total de individuos recolectados durante 2018 en el río dividido por 15 (el número total de celdas muestreadas en el río) multiplicado por 10 (el número de fechas en las que se realizó el trabajo de campo). Denotamos este número por

$$K := \text{Densidad promedio anual de individuos por celda en el río} = 0.113$$

Este es un número muy pequeño porque la especie que modelamos está en peligro crítico de extinción. Esta es la razón por la que tenemos que tomar un promedio anual, y las simulaciones de Poisson son inútiles: una simulación de Poisson en el río produciría muy pocas ocurrencias y esto genera agrupaciones (artificiales) alrededor ellos (y esto no es

razonable, no hay características en la región que promuevan este comportamiento). La forma correcta de modelar esta situación es asumir que el río tiene una densidad homogénea de individuos por celda y no generar ninguna realización o simulación en él, como hacemos en los Capítulos 2 y 3 (ver también [19] y [20]). Observe que la realización de la distribución de probabilidad que proponemos es prácticamente lo mismo que la realización de un proceso de Poisson. Mantenemos las densidades como están (sin realizaciones) y procedemos con nuestro análisis (usando muestreo de Gibbs) con densidades promedio anuales por celda en lugar de individuos, como hacemos en los Capítulos 2 y 3 (ver también [19] y [20]), esto es una nueva característica del modelo que proponemos. Las realizaciones solo se simulan una vez que se obtiene la densidad promedio anual de individuos en cada celda de la región utilizando nuestro método (Figura 4.7). Esto implica una diferencia importante en nuestro modelo matemático, con respecto a [19] y [20], porque las densidades promedio anuales en las celdas pertenecen al intervalo

$$[0, K]$$

y no a un conjunto finito que describe individuos, como en los Capítulos 2 y 3 (ver también [19] y [20]).

4.6. Estados (configuraciones)

Denotamos por

$$\omega^R : R \mapsto [0, K], \quad \omega_p^R = K, \quad \forall p \in R,$$

el estado en el sistema, que describe las densidades promedio anuales en cada celda del río. Observe que esto es diferente del esquema del Capítulo 2 y la aplicación del Capítulo 3, donde se usa una simulación de Poisson y las celdas de río toman valores sobre el número de individuos, un conjunto finito, en lugar de densidades promedio anuales, que pertenecen al intervalo $[0, K]$.

Denotamos

$$\Omega^i := [0, K]^{\Lambda^i}$$

al conjunto de estados o configuraciones en la Región Λ^i . Para todo estado $\omega \in [0, K]^{\Lambda^i}$, denotamos

$$\omega_{i,j} \in [0, K],$$

la densidad promedio anual de individuos por celda en la celda (i, j) .

4.7. Medidas de probabilidad y Hamiltonianos condicionales (dados los datos del río)

Definición 4.7.1 (Hamiltonianos condicionales en las Regiones 1, 2, 3 y 4). *Para cada $\omega \in \Omega^i$, definimos*

$$H_0^i(\omega) := \sum_{s_1 \sim s_2} (\omega_{s_1} - \omega_{s_2})^2 + \sum_{\substack{s \sim r \\ r \in R}} (\omega_s - \omega_r^R), \quad (4.1)$$

$$V_g^i(\omega) = g \sum_{s \in \Lambda^i} d_s^2 \omega_s, \quad (4.2)$$

y los hamiltonianos condicionales en las Regiones 1, 2, 3 y 4 por

$$H^i(\omega) = H_0^i(\omega) + V_g^i(\omega). \quad (4.3)$$

El Hamiltoniano H_0^i (energía libre) se minimiza cuando las densidades promedio de todos los años, en cada celda, son iguales al valor máximo posible (K - la densidad en las celdas del río) - es decir, promueve la propagación. La energía potencial V_g^i se minimiza cuando no hay individuos fuera del río (desempeña el papel de una fuerza de atracción hacia el río). Minimizar H implica un equilibrio entre no tener individuos en el río y tener el número máximo. La función de acoplamiento g mide la fuerza de la atracción al río.

Definición 4.7.2 (Probabilidades condicionales). *Para todo $\omega \in \Omega^i$ denotamos por*

$$\pi_T^i(\omega) = \frac{1}{Z^i} e^{-\frac{1}{T} H^i(\omega)}, \quad (4.4)$$

donde la función partición Z^i está dada por

$$Z^i := \sum_{\omega \in \Omega^i} \exp \left\{ -\frac{1}{T} H^i(\omega) \right\}.$$

En el Capítulo 2 (ver también [19]) mostramos que π_T^i se puede considerar como una probabilidad condicional dados los datos del río ω^R . El parámetro T es la temperatura. Cuando T es pequeño, los estados más probables se concentran entre los estados menos energéticos. Sin embargo, es un problema difícil acceder a estados de energía

tan bajos (este es un problema de minimización). Usamos diferentes valores de T para minimizar este parámetro con la ayuda del muestreo de Gibbs. En el Capítulo 2 (ver también [19]), explicamos y justificamos matemáticamente nuestro procedimiento. Como especificamos en la Definición 3.1, los estados de baja energía son los que modelan el equilibrio entre esparcirse y ser atraídos por el río.

4.7.1. Orden de las celdas y Muestreo de Gibbs

Por cada $i \in \{1, 2, 3, 4\}$, definimos una sucesión $(\ell_n^i)_{n \in \mathbb{N}}$, con $\ell_n^i \in \Lambda^i$ y con las siguientes propiedades.

- $(\ell_n^i)_{n \in \mathbb{N}}$ es periódica, con un período igual al número total de celdas en Λ^i , que denotamos por m^i . Además, $\ell_1^i, \dots, \ell_{m^i}^i$ son todos diferentes y cubren la región completa Λ^i .
- Los primeros elementos de $(\ell_n^i)_{n \in \mathbb{N}}$ cubren a todos los vecinos del río, los segundos elementos cubren a todos los vecinos de los primeros elementos. Los terceros cubren a todos los vecinos del primer y segundo elementos. Procedemos de la misma manera hasta que cubramos todas las celdas en Λ^i .

Las propiedades anteriores no especifican con precisión una sucesión, pero son las cualidades importantes de nuestro esquema de muestreo. En el Capítulo 2 (ver también [19]) especificamos la sucesión que usamos, para la región Λ^4 (las otras se construyen de manera similar).

Definición 4.7.3 (Muestreo de Gibbs). *Un muestreo de Gibbs es una sucesión de estados $(\omega^{i,n})_{n \in \mathbb{N} \cup \{0\}}$ en Ω^i definido por lo siguiente: Elegimos un estado inicial $\omega^{i,0}$. Si $\omega^{i,0}, \dots, \omega^{i,n-1}$ están definidos, tomamos aleatoriamente $x \in [0, K]$ y establecemos $\tilde{\omega}$ en ser el estado que coincide con $\omega^{i,n-1}$ en todas las celdas con la excepción de ℓ_n . Elegimos $\omega^{i,n-1} = \tilde{\omega}$ si $H^i(\tilde{\omega}) < H^i(\omega^{i,n-1})$. De lo contrario, seleccionamos al azar $y \in (0, 1)$. Si $y \leq e^{-\frac{1}{T}(H^i(\tilde{\omega}) - H^i(\omega^{i,n-1}))}$, entonces $\omega^{i,n-1} = \tilde{\omega}$. Si este no es el caso, $\omega^{i,n} = \omega^{i,n-1}$.*

La importancia del muestreo de Gibbs es que $(\omega^{(i,n)})_{n \in \mathbb{N} \cup \{0\}}$ se puede usar para calcular los valores esperados del número de individuos en cada celda:

Teorema 4.7.1 (Teorema Ergódico - Theorem C [61]). *Supongamos que $(\theta^n)_{n \in \mathbb{N} \cup \{0\}}$ es una muestra de Gibbs en Λ^i . Entonces*

$$\text{Valor esperado de la densidad de individuos en } p = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n \theta_p^m,$$

con probabilidad 1.

El teorema ergódico también se aborda en [129, 158]. En estas fuentes, no está probado de la manera en que lo necesitamos porque aquí tenemos un conjunto finito Γ en lugar de $[0, K]$ no es finito. Sin embargo, en situaciones prácticas, el teorema ergódico es generalmente inútil porque las tasas de convergencia son demasiado lentas la mayoría de las veces ([69] y [158]). No pretendemos aplicar tal teorema, sino encontrar un método que represente la idea de un equilibrio entre atracción y propagación, de tal manera que las fuerzas de atracción puedan ser controladas con precisión para hacer coincidir los datos experimentales. Esto es exactamente lo que obtenemos en la Sección 4.8, y este es el núcleo de nuestro método introducido en el Capítulo 2 (ver también [19]): Controlamos globalmente nuestras medidas, de tal manera que para el caso libre nos acercamos al mínimo de energía y como el acoplamiento aumentos constantes nos movemos continuamente desde una situación homogénea hasta el punto en que todos los individuos están en el río. Esto sugiere que ya estamos en el régimen donde el teorema ergódico es válido, pero la prueba de esto no es útil en este manuscrito porque ya tenemos lo que queremos en términos de modelado (el teorema ergódico sirve de inspiración).

4.7.2. Estimación de parámetros

4.7.2.1. Temperaturas auxiliares: T_1 y un rango de T_2

Como explicamos anteriormente, nuestro modelo se basa en el uso de estados de baja energía de los hamiltonianos para estimar la distribución de probabilidad de los individuos. La técnica de minimización que utilizamos es encontrar los estados más probables para las medidas π_T^4 , utilizando muestreo de Gibbs en el régimen de baja temperatura. Aquí nos enfocamos en la Región Λ^4 porque el trabajo de campo se llevó a cabo en esta región. Como explicamos en el Capítulo 2 (y en [19]), el caso libre está bajo control porque conocemos el minimizador y usamos este caso para la calibración. En el Capítulo 2 (y en [19]), explicamos con todo detalle este procedimiento. Tomamos diferentes opciones de la temperatura T y consideramos las secuencias de muestreo de Gibbs (finitas) correspondientes con $1000000 \times m_4$ iteraciones, donde $m_4 = 13715$ es

el número de celdas en la Región Λ^4 , y un estado inicial igual a cero. Seleccionamos la temperatura T_1 que minimiza (numéricamente) las energías medias de las últimas $10000 \times m_4$ iteraciones. En nuestro caso, esto da ya una muy buena aproximación para la energía mínima. Sin embargo, queremos cubrir un rango mayor para los valores de K y, por lo tanto, reemplazamos K por $K' = 0.01$ y ejecutamos $1000000 \times m_4$ iteraciones de muestreo de Gibbs con T_1 y el estado inicial es igual a 0. Calculamos el promedio de las últimas $10000 \times m_4$ iteraciones. Esto define un estado cuya energía ya está cerca del mínimo (pero no lo suficientemente cerca), lo llamamos

$$\omega_0'. \tag{4.5}$$

Volvemos a ejecutar un muestreo de Gibbs con $1000000 \times m_4$ iteraciones, pero ahora con el estado inicial ω_0' (y K' en lugar de K) y descendemos la temperatura T hasta observar que las energías de las últimas $10000 \times m_4$ iteraciones están muy próximas al mínimo. Esto produce un intervalo de temperaturas que proporciona buenas aproximaciones al mínimo de energía. Denotamos por T_2 estas temperaturas y el intervalo correspondiente lo llamaremos el *rango* de T_2 . Las temperaturas T_1 y T_2 son importantes porque si sustituimos K por cualquier número en $[0.01, 0.113]$ y ejecutamos $2000000 \times m_4$ iteraciones de muestreo de Gibbs (las primeras $1000000 \times m_4$ con T_1 y las segundas $1000000 \times m_4$ con T_2) como se describió anteriormente, obtenemos estados cuya energía es muy cerca del mínimo. A continuación, mantenemos la temperatura T_1 fija, pero no se supone que T_2 sea fija: Necesitamos cierta flexibilidad para poder ajustar el parámetro g que coincida con los datos del trabajo de campo (en realidad, T_2 y g se estiman simultáneamente; el valor final de T_2 es la temperatura del sistema, y la llamamos T).

4.7.2.2. Estimación de g y T

Denotamos por g y T los valores “verdaderos” para los parámetros g y T . El parámetro g es el más importante en este manuscrito, porque fija la fuerza de atracción al río y para determinarla, usamos datos recolectados del parcelas.

Denotamos por $P \subset \Lambda^4$ el conjunto de 5 parcelas. Como hacemos con los datos de los ríos, el hecho de que las especies que modelamos estén en peligro crítico implica que no podemos considerar las parcelas individualmente: las observaciones son extremadamente improbables. Si se observa un individuo en una parcela, no significa que esta parcela sea una fuente de individuos, sino que por casualidad la ocurrencia ocurrió en esa parcela (aquí nuevamente tenemos el problema de los conglomerados descrito). Dado que todas las parcelas pertenecen a la Región 4, usamos esta región para fijar la

4.7 Medidas de probabilidad y Hamiltonianos condicionales (dados los datos del río)

constante de acoplamiento g . Hacemos explícita temporalmente la dependencia de g de las sucesiones de muestreo de Gibbs

$$(\omega^{i,n})_{n \in \mathbb{N} \cup \{0\}} \equiv (\omega^{i,n;g})_{n \in \mathbb{N} \cup \{0\}}. \quad (4.6)$$

Comenzamos con simulaciones de muestreo de Gibbs con temperatura $T = T_1$ y $g = 0$ y aumentamos lentamente g . Estimamos el valor esperado del número promedio anual de individuos en las parcelas P con los promedios de las últimas $10000 \times m_4$ iteraciones de simulaciones de muestreo de Gibbs (y el estado inicial cero):

$$\text{Número esperado de individuos en } P = \int_P \omega d\pi^4(\omega) \approx \frac{1}{10000 \times m_4} \sum_{n=990001 \times m_4}^{1000000 \times m_4} \sum_{p \in P} \omega_p^{i,n;g}. \quad (4.7)$$

Aumentamos g hasta que aproximadamente igualemos el promedio anual del número total de individuos recolectados en las parcelas (que es igual a 0.3), y este valor, que denotamos por g_0 , es nuestra estimación inicial de g :

$$\text{Número esperado de individuos en } P \approx \frac{1}{10000 \times m_4} \sum_{n=990001 \times m_4}^{1000000 \times m_4} \sum_{p \in P} \omega_p^{i,n;g}. \quad (4.8)$$

La estimación en (4.8) no es lo suficientemente precisa, pero es un punto de referencia útil para comenzar. Se obtiene más precisión con la ayuda de un esquema de temperatura descendente y 1000000 iteraciones adicionales. Tomamos diferentes valores de T_2 en el rango de T_2 y ejecutamos $1000000 \times m_4$ iteraciones con el estado inicial cero y T_1 . Tomamos el promedio de las últimas $10000 \times m_4$ iteraciones y lo tomamos como valor inicial para $1000000 \times m_4$ iteraciones adicionales con temperatura T_2 . Usamos las últimas $10000 \times m_4$ iteraciones (de las $2000000 \times m_4$ iteraciones) para calcular promedios y modificar los valores de g (en un vecindario de g_0) para que coincidan (3.13). Con este procedimiento, obtenemos un valor g_{T_2} de g para cada valor de T_2 en el rango de T_2 (esto es fácil de lograr numéricamente). A continuación, cambiamos los valores de T_2 y calculamos un nuevo muestreo de Gibbs con $2000000 \times m_4$ iteraciones como se describe anteriormente (tomando g_{T_2}), y creamos gráficos de las energías correspondientes. Seleccionamos el valor de T_2 que proporciona la mejor tasa de decaimiento para las energías de tal manera que observemos un decaimiento de Energía en la secuencia completa ($2000000 \times m_4$ iteraciones) y una estabilización en las últimas iteraciones (observe que hay valores de T_2 que producen comportamientos crecientes para la energía en las segundas $1000000 \times m_4$ iteraciones). Este valor de T_2 es la temperatura de nuestro

sistema, y lo llamamos

$$\mathbf{T}.$$

La constante de acoplamiento correspondiente $g_{\mathbf{T}}$ define la constante de acoplamiento del sistema, se denota por

$$\mathbf{g}.$$

4.7.3. Densidades promedio anuales por celda y representación gráfica

Una vez que se determinan \mathbf{T} y \mathbf{g} , calculamos $1000000 \times m_4$ iteraciones de muestreo de Gibbs con el estado inicial cero, acoplado la constante \mathbf{g} y la temperatura T_1 . Tomamos el promedio de los últimos 10000 estados. Elegimos este promedio como estado inicial para una nueva secuencia de muestreo de Gibbs con 1000000 iteraciones, utilizando \mathbf{T} y \mathbf{g} . Finalmente, denotamos por

$$\omega$$

el promedio de los últimos $10000 \times m_4$ estados (de las $2000000 \times m_4$ iteraciones).

El valor $\omega_{i,j}$ representa el valor esperado del número promedio anual de individuos en la celda (i, j) . Dado que $\omega_{i,j}$ no representa un valor esperado de los individuos, como es el caso de los modelos de los Capítulos 2 y 3 (ver también [19] y [20]), necesitamos definir visualizaciones gráficas más sofisticadas en comparación a las mostradas en los Capítulos 2 y 3 (ver también [19] y [20]). En las siguientes secciones explicamos esto con mayor detalle.

4.7.3.1. Individuos representados por puntos

Como mencionamos anteriormente, $\omega_{i,j}$ es el valor esperado del número promedio anual de individuos en la celda (i, j) . Usamos esta información para obtener representaciones gráficas de las personas que podríamos observar en la región.

Individuos simulados observados durante todo el año. Primero, denotamos por

$$\omega_Y := 10\omega,$$

donde

$$(10\omega)_{i,j} = 10\omega_{i,j}.$$

Entonces, $(\omega_Y)_{i,j}$ representa el valor esperado de los individuos observados durante todo el año (10 viajes de campo) en la celda (i, j) .

Para cada (i, j) , simulamos un número aleatorio $n_{i,j} \in \{0, 1, 2, 3, \dots\}$ según una distribución de Poisson con valor esperado $(\omega_Y)_{i,j}$. Entonces definimos

$$\mathbf{n} := (\mathbf{n}_{i,j})_{\text{para todas las celdas } (i,j)},$$

que es un estado construido aleatoriamente con valores en $\{0, 1, 2, 3, \dots\}$, y $n_{i,j}$ representa el número de individuos que podrían haber sido vistos en la celda (i, j) , en los 10 viajes de campo (Figura 4.6).

Observaciones promedio anuales simuladas. Para cada (i, j) , simulamos un número aleatorio $x \in [0, 1]$ (distribuido uniformemente). Si $x \leq \omega_{i,j}$, entonces elegimos $\mathbf{m}_{i,j} = 1$, de lo contrario $\mathbf{m}_{i,j} = 0$. Entonces definimos

$$\mathbf{m} := (\mathbf{m}_{i,j})_{\text{para todas las celdas } (i,j)},$$

que es un estado construido al azar con valores en $\{0, 1\}$, y $\mathbf{m}_{i,j}$ representan el número promedio anual de individuos que podrían haber sido vistos en la celda (i, j) , en un viaje de campo (Figura 4.7).

4.7.3.2. Mapas de calor

Usamos un mapa de calor (Figura 4.5) para representar los valores $\omega_{i,j}$, para cada i y cada j . Utilizamos los 7 colores del arco iris: rojo, naranja, amarillo, verde, cian, azul y violeta (este orden corresponde a la energía ascendente de los colores). Cada celda (i, j) está coloreada con uno de estos colores dependiendo de los valores de $\omega_{i,j}$ de tal manera que el color es constante en intervalos de la forma $\omega_{i,j} \in (\frac{s}{7}K, \frac{s+1}{7}K]$, $s \in \{0, \dots, 6\}$, y la energía del color aumenta a medida que aumenta $\omega_{i,j}$. Nosotros, además, aumentamos la intensidad del color a medida que aumenta la energía.

4.8. Resultados

4.8.1. Temperatura e iteraciones (Caso Libre)

Obtenemos que

$$T_1 = 0.00000475 \quad \text{and} \quad T = 0.00000002. \quad (4.9)$$

4. MODELADO DE DISTRIBUCIÓN DE BAJA DENSIDAD Y ESPECIES EN PELIGRO CRÍTICO: PLECTROHYLA GUATEMALENSIS

En la Figura 4.3 presentamos el gráfico de las energías libres de 2000000 de iteraciones de muestreo de Gibbs, en cada región. Se puede ver claramente que las energías descienden dramáticamente y se estabilizan cerca de la energía mínima (cero)

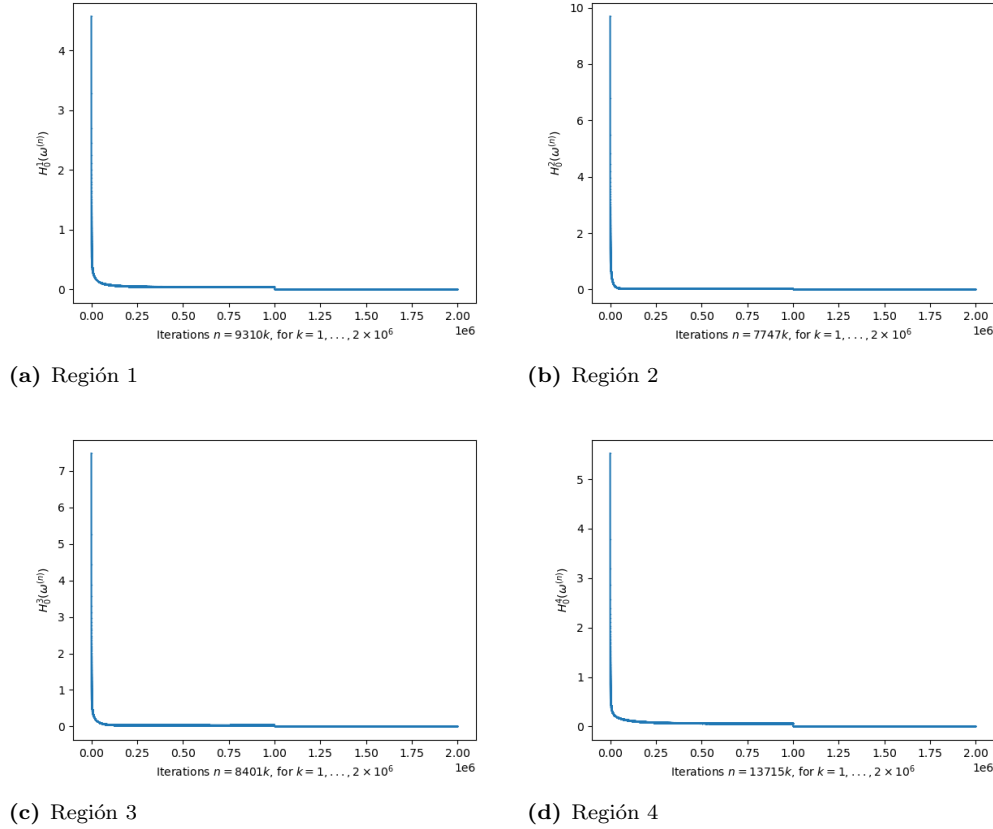


Figura 4.3: Energía libre con temperatura T , de acuerdo a la Sección 4.7.2, para 2000000 de iteraciones de un muestreo de Gibbs de la forma $k = m_i k$, para $k = 1, \dots, 2000000$ y donde m_i es el número de celdas en cada Región Λ^i , $i = 1, 2, 3, 4$.

4.8.2. La constante de acoplamiento g

Obtenemos que

$$g := 0.000000325. \quad (4.10)$$

En la Figura 4.4, presentamos el gráfico de las energías de las iteraciones de muestreo de 2000000 Gibbs en cada región (como se describe en la Sección 4.7.2.2). Se puede ver claramente que las energías descienden dramáticamente y se estabilizan cerca de cero.

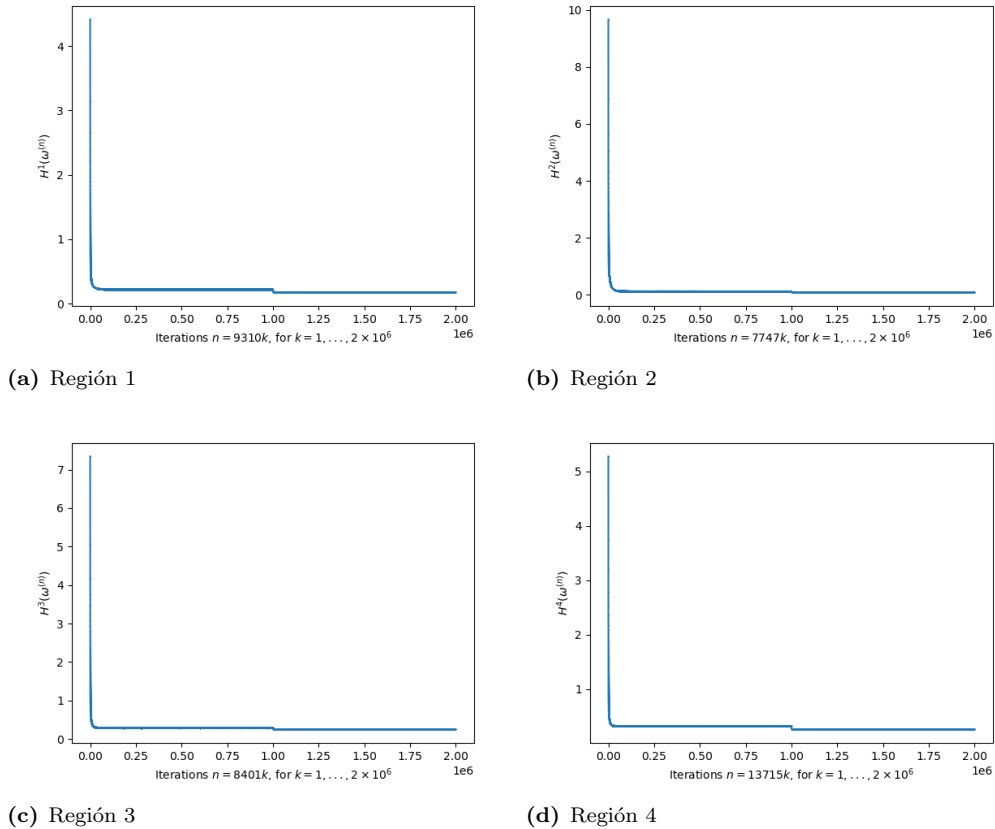


Figura 4.4: Estabilización de la energía en el caso perturbado: energía con temperatura T y constante de acoplamiento g , de acuerdo a la Sección 4.7.2.2, para 2000000 de iteraciones de un muestreo de Gibbs de la forma $k = m_i k$, para $k = 1, \dots, 2000000$ y donde m_i es el número de celdas de cada Región Λ^i , $i = 1, 2, 3, 4$.

4.8.3. Resultados Principales: Distribución y abundancia de individuos

Nuestros principales resultados son las representaciones gráficas de los individuos en cada celda de la región.

- En la Figura 4.6 representamos con puntos a cada individuo que se podría haber visto durante las fechas de 10 reportadas en la Sección 4.4, como se explica en la Sección 4.7.3.1. Presentamos 5 realizaciones n .
- En la Figura 4.7 describimos con puntos los individuos promedio del año que se podría haber visto en un viaje de campo en las fechas de 10 dadas en la Sección 4.4 (ver Sección 4.7.3.1). Presentamos 5 realizaciones de m .

4. MODELADO DE DISTRIBUCIÓN DE BAJA DENSIDAD Y ESPECIES EN PELIGRO CRÍTICO: PLECTROHYLA GUATEMALENSIS

La densidad de individuos es mayor en el río y disminuye a medida que aumenta la distancia al río. La especie *Plectrohyla Guatemalensis* se siente atraída por el río, pero no fuertemente. Esto difiere de *Plectrohyla Sagorum* (ver Capítulo 3 y [20]) que solo se puede encontrar en las proximidades del río.

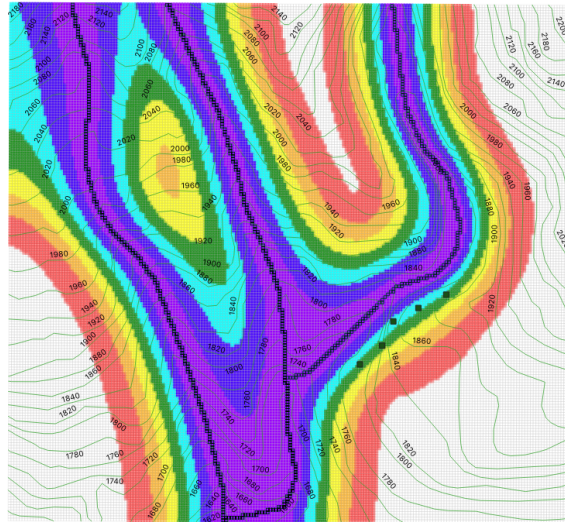
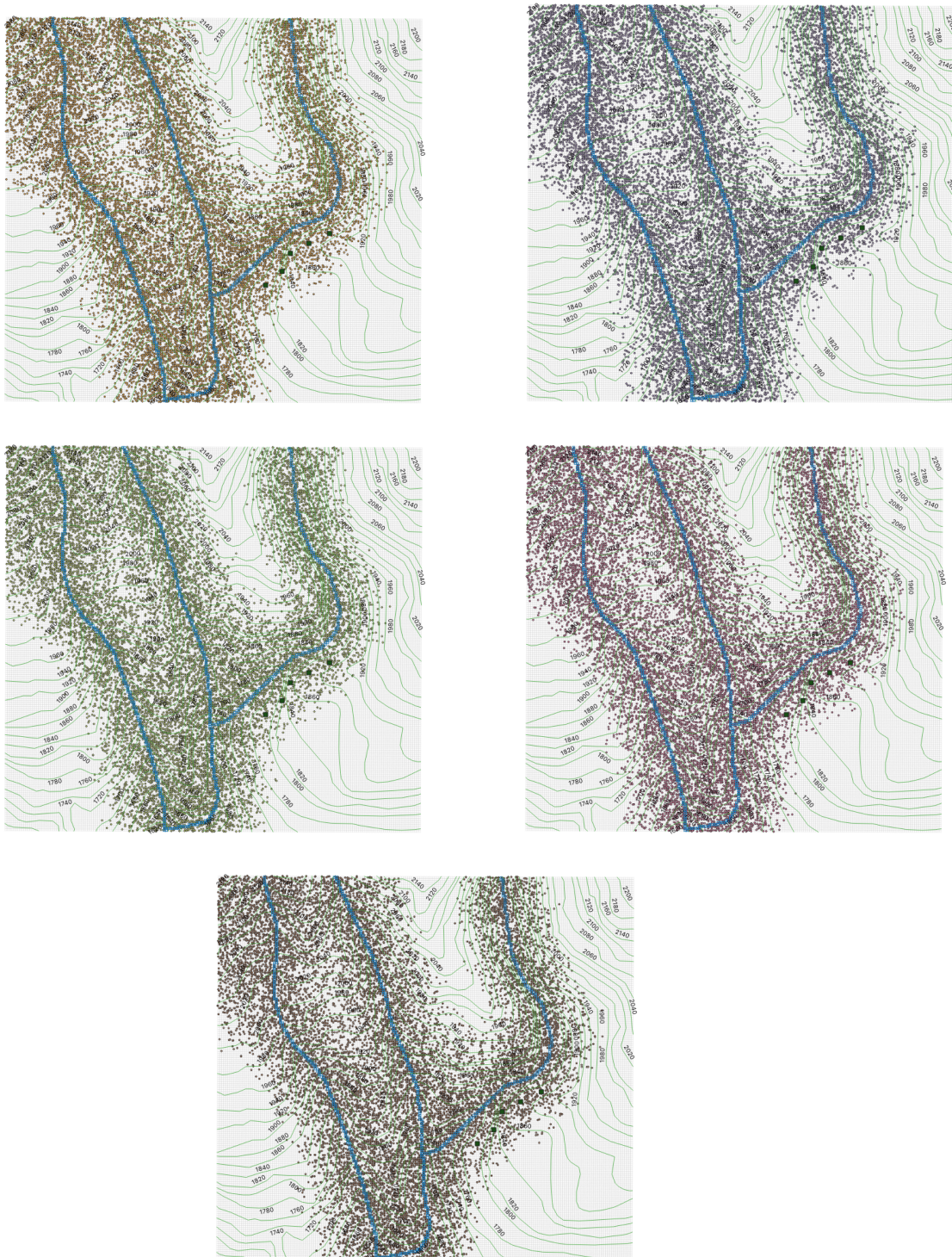


Figura 4.5: Mapa de calor de la densidad de individuos de la especie *Plectrohyla Guatemalensis*. Usamos negro para los bordes de los ríos con el fin de hacerlos distinguibles. Usamos los colores de 7 del arco iris: el color más energético (violeta) representa la densidad más alta y el color menos energético (rojo) representa la densidad más baja. La escala de colores va acompañada de una degradación de la intensidad. Rojo ■ representa un número esperado de individuos en el intervalo $(0, 0.016143]$; naranja ■ representa un número esperado de individuos en $(0.016143, 0.032286]$; amarillo ■ representa un número esperado de individuos en $(0.032286, 0.048429]$; verde ■ representa un número esperado de individuos en $(0.048429, 0.064572]$; cian ■ representa un número esperado de individuos en $(0.064572, 0.080715]$; azul ■ representa un número esperado de individuos en $(0.080715, 0.096858]$; violeta ■ representa un número esperado de individuos en $(0.096858, 0.113]$.

Figura 4.6: 5 realizaciones de n

4. MODELADO DE DISTRIBUCIÓN DE BAJA DENSIDAD Y ESPECIES EN PELIGRO CRÍTICO: PLECTROHYLA GUATEMALENSIS

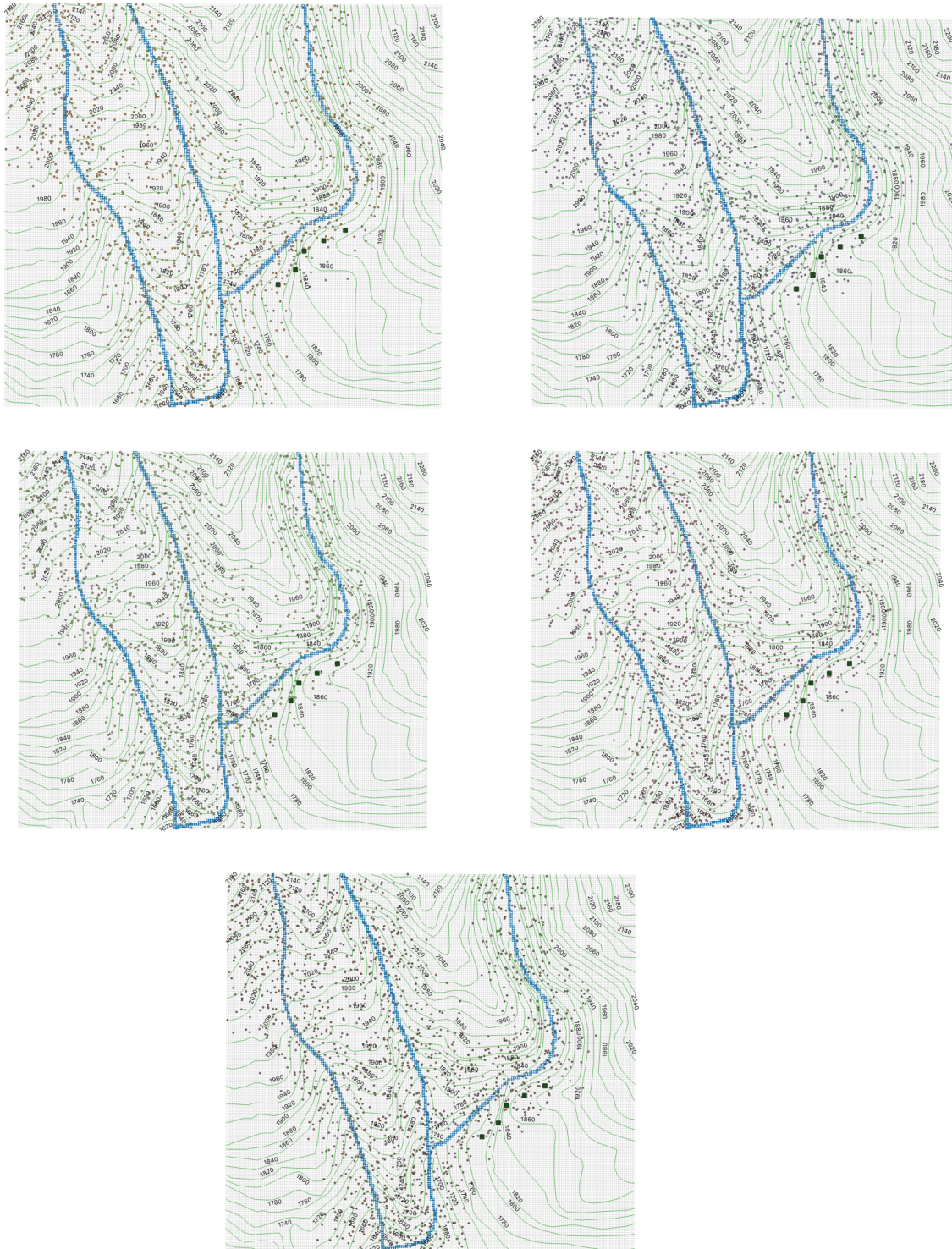


Figura 4.7: 5 realizaciones de m

4.9. Conclusiones

Estimamos la distribución y abundancia de la especie en peligro crítico de extinción *Plectrohyla Guatemalensis* en una región ubicada en la reserva natural Volcán Tacaná. Nuestro trabajo contribuye al conocimiento de esta especie y esperamos que pueda ayudar a su preservación. Obtenemos que los individuos se sienten atraídos por el río, pero la fuerza de atracción es mucho más débil que otras especies del mismo género como *Plectrohyla Sagorum* (ver Capítulo 3 y también [20]). Desde el punto de vista matemático, presentamos un método que se puede utilizar para describir especies en peligro crítico (basado en el modelo introducido en el Capítulo 2 y en [19]).

Este es un problema delicado porque la densidad extremadamente baja de individuos que ocurre en tales situaciones hace que sea muy difícil reconstruir una distribución de probabilidad. En nuestro caso, hay en promedio un poco más de 2 individuos recolectados en cada viaje de campo (y los métodos estándar frecuentemente usan datos de un solo muestreo de campo). Nuestro modelo asume que la presencia de individuos decae a medida que aumenta la distancia al río, y se supone que la tasa de decaimiento se ajusta a la energía potencial cuadrática (inspirada en el oscilador armónico de la mecánica cuántica). Entonces, la estimación de parámetros se reduce a una estimación de la constante de acoplamiento g .

Intencionalmente dejamos solo un parámetro libre (la constante de acoplamiento g), porque los datos del trabajo de campo que usamos son pobres y no son suficientes para estimar más de un parámetro. Sin embargo, nuestro método tiene la flexibilidad de obtener tasas de descomposición más precisas, pero esto requeriría trabajo de campo adicional (esto no es trivial, porque la región que estudiamos es un cañón con una orografía muy complicada). En este trabajo, las tasas de decaimiento precisas son más relevantes que en el Capítulo 3 (ver también [20]), porque en este último caso, los individuos permanecen en las cercanías del río y aquí todavía están presentes a largas distancias.

Análisis del límite a bajas energías de la matriz de dispersión asociada a la ecuación de Schrödinger discreta

En este capítulo estudiamos la teoría de dispersión para la ecuación matricial de Schrödinger discreta, en particular, definimos la matriz de dispersión a través de las soluciones de Jost, probamos que existe el límite de la matriz de dispersión en el límite y encontramos una expresión para el límite relacionándolo con la existencia o no de soluciones acotadas para los umbrales (half-bounded states). Esta parte de la investigación se ha realizado en colaboración con Gerardo Franco, cuya contribución es mayoritaria y forma parte de sus estudios de doctorado que están por iniciar. Los resultados más importantes sobre este tema junto con las demostraciones están [15].

5.1. Las soluciones de Jost

La ecuación matricial de Schrödinger unidimensional está dada por

$$H\phi = E\phi \tag{5.1}$$

en donde

$$H = -\frac{d^2}{dx^2} + V \tag{5.2}$$

es un operador autoadjunto densamente definido sobre $L^2(\mathbb{R}, \mathbb{C}^L)$. La ecuación (5.1) modela a una partícula (cuántica) que se desplaza en el eje real y el hecho de que es

5. ANÁLISIS DEL LÍMITE A BAJAS ENERGÍAS DE LA MATRIZ DE DISPERSIÓN ASOCIADA A LA ECUACIÓN DE SCHRÖDINGER DISCRETA

matricial se puede asociar a grados de libertad del sistema como por ejemplo el spin. Si imponemos además la restricción

$$V \in L^1_1(\mathbb{R}; \mathbb{C}^{L \times L}), \quad (5.3)$$

o equivalentemente, $(1 + |x|)\|V(x)\|$ es integrable sobre \mathbb{R} , con E en \mathbb{R} , entonces podemos encontrar soluciones fundamentales acotadas (en cada dirección), que no están en $L^2(\mathbb{R}, \mathbb{C}^L)$, las cuales son llamadas eigenvalores generalizados.

La versión discreta de (5.1) es una ecuación en diferencias simétrica de segundo orden de la forma

$$Hu = Eu, \quad (5.4)$$

donde $H : \ell^2(\mathbb{Z}, \mathbb{C}^L) \rightarrow \ell^2(\mathbb{Z}, \mathbb{C}^L)$ es un operador de segundo orden tal que

$$(Hu)(n) = -u(n-1) + 2u(n) - u(n+1) + V(n)u(n) \quad (5.5)$$

y las matrices $V(n) \in \mathcal{M}_{L \times L}(\mathbb{C})$ son autoadjuntas.

Luego, si consideramos el operador derivada discreta (Laplaciano discreto)

$$\frac{d}{dn}u(n) = u(n+1) - u(n),$$

entonces con un cálculo sencillo podemos ver que

$$-\frac{d^2}{dn^2}u(n-1) = -u(n-1) + 2u(n) - u(n+1) \quad (5.6)$$

de donde se desprende la analogía con el caso continuo expresado en la (5.1).

En este capítulo estudiamos la matriz unidimensional de Schrödinger discreta de la forma

$$H = H = H_0 + V, \quad (5.7)$$

donde H_0 (el Hamiltoniano libre) está dado por el Laplaciano discreto sobre el espacio de Hilbert $\ell^2(\mathbb{Z}, \mathbb{C}^L)$ definido como

$$(H_0\phi)(n) := \phi(n+1) + \phi(n-1), \quad \phi \in \ell^2(\mathbb{Z}, \mathbb{C}^L). \quad (5.8)$$

Consideramos además un operador $V \in \mathcal{M}_{L \times L}(\mathbb{C})^{\mathbb{Z}}$ de multiplicación matricial con

primer momento finito, definido por

$$(V\Psi)(n) := V(n)\Psi(n), \quad \Psi \in \ell^2(\mathbb{Z}, \mathbb{C}^L). \quad (5.9)$$

Suponemos además que $V(n) = V(n)^*$, $n \in \mathbb{Z}$ y

$$\sum_{n \in \mathbb{Z}} \|nV(n)\| < \infty. \quad (5.10)$$

y parametrizamos los autovalores en la forma

$$E = z + 1/z, \quad z \in \mathbb{C} \setminus \{0\}. \quad (5.11)$$

Entonces las ecuaciones de autovalores tienen la forma

$$Hu = Eu, \quad (5.12)$$

$$H_0u = Eu, \quad (5.13)$$

donde $u \in (\mathcal{M}_{L \times L})^{\mathbb{Z}}$.

Definición 5.1.1 (Soluciones de Jost). *Para toda $z \in \overline{\mathbb{D}} \setminus \{0\}$, donde \mathbb{D} es el disco unitario abierto complejo, denotamos por u_+^z y $u_-^{1/z}$ las soluciones $\mathcal{M}_{L \times L}$ -valuadas de*

$$Hu_+^z = Eu_+^z, \quad Hu_-^{1/z} = Eu_-^{1/z}, \quad E = z + 1/z, \quad (5.14)$$

tales que, cuando $n \rightarrow +\infty$ y $n \rightarrow -\infty$, respectivamente,

$$u_+^z(n) = z^n(\mathbf{1} + o(1)), \quad u_-^{1/z}(n) = z^{-n}(\mathbf{1} + o(1)), \quad (5.15)$$

donde $\mathbf{1}$ es la matriz identidad. Adicionalmente, para $z = 1$, denotamos por v_{\pm}^z las soluciones $\mathcal{M}_{L \times L}$ -valuadas de

$$Hv_{\pm}^z = Ev_{\pm}^z, \quad E = 2, \quad (5.16)$$

tales que cuando $n \rightarrow \pm\infty$,

$$v_{\pm}^z(n) = n(\mathbf{1} + o(1)). \quad (5.17)$$

5.2. La matriz de dispersión

Definición 5.2.1 (Wronskiano). *Para dos funciones $u, v : \mathbb{Z} \rightarrow \mathcal{M}_{L \times L}$ y $n \in \mathbb{Z}$ el Wronskiano se define por*

$$W(u, v)(n) = i(u(n+1)^*v(n) - u(n)^*v(n+1)) \in \mathcal{M}_{L \times L}. \quad (5.18)$$

Es fácil ver que $W(u, v)^* = W(v, u)$ y que para soluciones matriciales que satisfagan $Hu = Eu$ and $Hv = \bar{E}v$, el Wronskiano $W(u, v)$ es independiente de n . En estos casos omitimos el argumento n . Para $z \in \mathbb{D}$, los Wronskianos de las soluciones Jost pueden evaluarse usando el hecho de que son independientes de n : tomando los límites, o bien $n \rightarrow \infty$ o $n \rightarrow -\infty$, encontramos que

$$W(u_+^{\bar{z}}, u_+^z) = 0 = W(u_-^{1/\bar{z}}, u_-^{1/z}), \quad (5.19)$$

y si además $z \in \mathbb{S}^1 \setminus \{-1, 1\}$

$$W(u_\pm^z, u_\pm^z) = (\nu^z)^{-1} \mathbf{1}, \quad (5.20)$$

donde $\mathbf{1}$ es la matriz identidad y

$$\nu^z = \frac{i}{z - z^{-1}}. \quad (5.21)$$

Ahora, recuerde que para $z \in \mathbb{S}^1 \setminus \{-1, 1\}$, es posible descomponer los estados u_+^z y $u_-^{1/z}$ sobre las bases $(u_-^z, u_-^{1/z})$ y $(u_+^z, u_+^{1/z})$, respectivamente, y que las matrices M_\pm^z y N_\pm^z están definidas por

$$u_+^z = u_-^z M_+^z + u_-^{1/z} N_+^z, \quad u_-^{1/z} = u_+^z N_-^z + u_+^{1/z} M_-^z. \quad (5.22)$$

La ecuación (5.20) da lugar a

$$\begin{aligned} M_+^z &= \nu^z W(u_-^{1/z}, u_+^z), \\ N_+^z &= -\nu^z W(u_-^{1/z}, u_+^z), \\ N_-^z &= \nu^z W(u_+^z, u_-^{1/z}), \\ M_-^z &= -\nu^z W(u_+^z, u_-^{1/z}). \end{aligned} \quad (5.23)$$

Esto muestra que M_{\pm}^z puede extenderse a funciones analíticas en $\mathbb{D} \setminus \{0\}$ que son continuos en $\overline{\mathbb{D}} \setminus \{-1, 0, 1\}$. Las Ecuaciones (5.23) implican que

$$(N_+^z)^* = -N_-^z, \quad (M_+^z)^* = M_-^{\bar{z}}, \quad (5.24)$$

donde la primera ecuación es válida para $z \in \mathbb{S}^1 \setminus \{-1, 1\}$ y el segunda se puede extender a $\overline{\mathbb{D}} \setminus \{-1, 0, 1\}$.

Tenemos entonces (ver Lema 15 [15]):

$$(M_-^z)^* M_-^z = \mathbf{1} + (N_-^z)^* N_-^z, \quad (5.25)$$

$$M_+^z N_-^z = -N_+^{1/z} M_-^z, \quad (5.26)$$

$$(M_+^z)^* M_+^z = \mathbf{1} + (N_+^z)^* N_+^z, \quad (5.27)$$

$$M_-^z N_+^z = -N_-^{1/z} M_+^z. \quad (5.28)$$

Proposición 5.2.1. *Para $z \in \mathbb{S}^1 \setminus \{-1, 1\}$, M_{\pm}^z es invertible y \mathcal{S}^z es unitaria.*

Demostración. Que M_{\pm}^z es invertible se sigue de las ecuaciones (5.25) y (5.27). Ahora probaremos que es unitaria. Los términos fuera de la diagonal de $(\mathcal{S}^z)^* \mathcal{S}^z$ son

$$-((M_+^z)^{-1})^* N_-^z (M_-^z)^{-1} - ((M_+^z)^{-1})^* (N_+^z)^* (M_-^z)^{-1}, \quad (5.29)$$

$$-((M_-^z)^{-1})^* (N_-^z)^* (M_+^z)^{-1} - ((M_-^z)^{-1})^* N_+^z (M_+^z)^{-1} \quad (5.30)$$

los cuales son nulos por (5.24). Los términos en la diagonal son

$$((M_+^z)^{-1})^* (1 + (N_+^z)^* N_+^z) (M_+^z)^{-1}, \quad (5.31)$$

$$((M_-^z)^{-1})^* (1 + (N_-^z)^* N_-^z) (M_-^z)^{-1},$$

los cuales son ambos iguales a $\mathbf{1}$, ver (5.27) y (5.25). Esto prueba que \mathcal{S}^z es unitaria. \square

Las columnas de la matriz $(u_{\pm}^z, u_{\pm}^{1/z})$ son linealmente independientes para $z \in \mathbb{S}^1 \setminus \{-1, 1\}$ y por tanto forman una base para las soluciones. Lo mismo es cierto para (u_{\pm}^1, v_{\pm}^1) . Se sigue que existen matrices $M_{\pm}^z, N_{\pm}^z \in \mathcal{M}_{L \times L}$ tales que

$$u_+^z = u_-^z M_+^z + u_-^{1/z} N_+^z, \quad u_-^{1/z} = u_+^z N_-^z + u_+^{1/z} M_-^z. \quad (5.32)$$

5. ANÁLISIS DEL LÍMITE A BAJAS ENERGÍAS DE LA MATRIZ DE DISPERSIÓN ASOCIADA A LA ECUACIÓN DE SCHRÖDINGER DISCRETA

Nosotros probamos además que las matrices M_{\pm}^z tienen una continuación meromórfica en \mathbb{D} (el disco unitario complejo). Suponiendo que M_{\pm}^z son invertibles, podemos reescribir estas ecuaciones como

$$u_+^z T_+^z = u_-^z - u_-^{1/z} R_+^z, \quad u_-^{1/z} T_-^z = u_+^{1/z} - u_+^z R_-^z, \quad (5.33)$$

donde

$$T_{\pm}^z = (M_{\pm}^z)^{-1}, \quad R_{\pm}^z = -N_{\pm}^z (M_{\pm}^z)^{-1}, \quad (5.34)$$

son los coeficientes de transmisión y reflexión, respectivamente.

Definición 5.2.2. *Para toda $z \in \mathbb{S}^1 \setminus \{-1, 1\}$, la matriz de dispersión $\mathcal{S}^z \in \mathcal{M}_{2L \times 2L}$ is defined by*

$$\mathcal{S}^z = \begin{pmatrix} T_+^z & R_-^z \\ R_+^z & T_-^z \end{pmatrix} = \begin{pmatrix} (M_+^z)^{-1} & -N_-^z (M_-^z)^{-1} \\ -N_+^z (M_+^z)^{-1} & (M_-^z)^{-1} \end{pmatrix}.$$

Note que las matrices M_{\pm}^z son invertibles (Proposición 5.2.1). El caso $z = e^{ik}$ representa una ola que viaja hacia la derecha. Entonces u_-^z y $u_+^{1/z}$ están entrando y u_+^z y $u_-^{1/z}$ están saliendo. En este caso, la matriz de dispersión expresa las soluciones de Jost entrantes u_-^z y $u_+^{1/z}$ en términos de los salientes u_+^z y $u_-^{1/z}$:

$$\begin{pmatrix} u_-^z & u_+^{1/z} \end{pmatrix} = \begin{pmatrix} u_+^z & u_-^{1/z} \end{pmatrix} \mathcal{S}^z. \quad (5.35)$$

5.3. Prueba del resultado principal

Lema 5.3.1. *Sea $u \in (\mathbb{C}^L)^{\mathbb{Z}}$ una solución de (5.12), para $z = 1$. Son equivalentes:*

- (i) u es $o(n)$ para $n \rightarrow +\infty$.
- (ii) u es acotada cuando $n \rightarrow +\infty$.
- (iii) u converge cuando $n \rightarrow +\infty$.

Además, u es $o(1)$ para $n \rightarrow +\infty$, si y sólo si $u = 0$. Lo anterior se mantiene válido si reemplazamos $+\infty$ por $-\infty$.

Demostración. De las Ecuaciones (5.15) y (5.17) se sigue que el número de columnas de u_+^1 y v_+^1 forman una base para todas las soluciones. Entonces existen $\alpha, \beta \in \mathbb{C}^L$ tales que

$$u = v_+^1 \alpha + u_+^1 \beta.$$

El comportamiento asintótico de u_+^1 y v_+^1 en ∞ produce el resultado deseado. \square

Definición 5.3.1. *Introducimos la notación:*

$$\mathcal{N} := \text{Ker}(W(u_-^1, u_+^1)), \quad \mathcal{L} := \text{Ker}(W(u_+^1, u_-^1)). \quad (5.36)$$

El caso genérico se conoce como $\mathcal{N} = \{0\}$, de lo contrario, se habla del caso excepcional. Para $\xi \in \mathcal{N}$ definimos

$$\Gamma \xi := \left(\xi - \sum_{j=-\infty}^{\infty} j V(j) u_+^1(j) \xi \right). \quad (5.37)$$

Definición 5.3.2. *Denotamos*

$$\Omega := u_+^1(1)^{-1} u_-^1(1) \in \mathcal{M}_{L \times L}. \quad (5.38)$$

Lema 5.3.2. *Las siguientes ecuaciones son verdaderas:*

$$u_+^1(n) = n(iW(u_-^1, u_+^1) + o(1)), \quad n \rightarrow -\infty.$$

$$u_-^1(n) = n(-iW(u_+^1, u_-^1) + o(1)), \quad n \rightarrow +\infty.$$

Demostración. Del Lema 20 [15], tenemos

$$W(u_-^1, u_+^1) = -i \sum_{j=-\infty}^{\infty} V(j) u_+^1(j), \quad (5.39)$$

5. ANÁLISIS DEL LÍMITE A BAJAS ENERGÍAS DE LA MATRIZ DE
DISPERSIÓN ASOCIADA A LA ECUACIÓN DE SCHRÖDINGER DISCRETA

y

$$W(u_+^1, u_-^1) = i \sum_{j=-\infty}^{\infty} V(j)u_-^1(j). \quad (5.40)$$

Ahora, del Lema 7 [15] se sigue que

$$\begin{aligned} u_+^1(n+1) - u_+^1(n) &= - \sum_{j=n+2}^{\infty} (j-n-1)V(j)u_+^1(j) + \sum_{j=n+1}^{\infty} (j-n)V(j)u_+^1(j) \\ &= \sum_{j=n+1}^{\infty} V(j)u_+^1(j) \\ &= iW(u_-^1, u_+^1) - \sum_{j=-\infty}^n V(j)u_+^1(j). \end{aligned}$$

Por lo tanto, $u_+^1(n+1) - u_+^1(n) \rightarrow iW(u_-^1, u_+^1)$, como $n \rightarrow -\infty$ (recordar $|u_+^1(n)| \leq C|n|$, Lema 10 de [15]). Esto implica que

$$\frac{1}{n}(u_+^1(n+1) - u_+^1(1)) = \frac{1}{n} \sum_{j=1}^n u_+^1(j+1) - u_+^1(j) \rightarrow iW(u_-^1, u_+^1)$$

y por tanto

$$\frac{u_+^1(n)}{n} \rightarrow iW(u_-^1, u_+^1), \quad n \rightarrow \infty.$$

Esto prueba la primera igualdad, la segunda igualdad se prueba de forma similar. \square

Lema 5.3.3. *Las siguientes ecuaciones son verdaderas:*

$$\mathcal{N} = \{\xi \in \mathbb{C}^L : u_+^1 \xi \text{ es acotada}\}, \quad \mathcal{L} = \{\chi \in \mathbb{C}^L : u_-^1 \chi \text{ es acotada}\}. \quad (5.41)$$

Además, dado que $W(u_-^1, u_+^1)^* = W(u_+^1, u_-^1)$ por definición, se sigue que

$$\mathcal{N} = (W(u_+^1, u_-^1)\mathbb{C}^L)^\perp, \quad \mathcal{L} = (W(u_-^1, u_+^1)\mathbb{C}^L)^\perp, \quad (5.42)$$

y por tanto, $\dim(\mathcal{N}) = \dim(\mathcal{L})$.

Sea $\{e_1, \dots, e_L\}$ una base ortonormal de \mathbb{C}^L tal que los primeros d vectores formen una base de \mathcal{L} y los últimos vectores $L - d$ a partir de $\mathcal{L}^\perp = W(u_-^1, u_+^1)\mathbb{C}^L$, ver Lema

5.3.3. Tomamos otra base ortonormal $\{v_1, \dots, v_L\}$ de \mathbb{C}^L tal que los primeros d vectores formen una base de \mathcal{N} y Los últimos $L - d$ vectores forman una base de \mathcal{N}^\perp . Entonces definimos

$$P := \begin{pmatrix} e_1 & e_2 & \cdots & e_L \end{pmatrix}^*, \quad Q := \begin{pmatrix} v_1 & v_2 & \cdots & v_L \end{pmatrix}. \quad (5.43)$$

Recordamos que $P_{\mathcal{L}}$ y $P_{\mathcal{L}^\perp}$ son las proyecciones sobre \mathcal{L} y \mathcal{L}^\perp , respectivamente. Entonces

$$P_{\mathcal{L}} \left(i(1-z)(\Omega^* + \Gamma)P_{\mathcal{N}} + W(u_-^1, u_+^1)P_{\mathcal{N}^\perp} \right) P_{\mathcal{N}} = i(1-z)P_{\mathcal{L}}(\Omega^* + \Gamma)P_{\mathcal{N}} \quad (5.44)$$

y $P_{\mathcal{L}}(\Omega^* + \Gamma)P_{\mathcal{N}}$ define una biyección entre \mathcal{N} y \mathcal{L} (Lema 5.3.3). Basta entonces con demostrar que es inyectiva. Esto es cierto porque $\Gamma : \mathcal{N} \rightarrow \mathcal{L}$ es una biyección y, para toda $\xi \in \mathcal{N}$,

$$\begin{aligned} \langle \Gamma\xi, P_{\mathcal{L}}(\Omega^* + \Gamma)\xi \rangle &= \langle P_{\mathcal{L}}\Gamma\xi, (\Omega^* + \Gamma)\xi \rangle = \langle \Gamma\xi, (\Omega^* + \Gamma)\xi \rangle \\ &= \langle \Omega\Gamma\xi, \xi \rangle + \|\Gamma\xi\|^2 = \|\xi\|^2 + \|\Gamma\xi\|^2. \end{aligned}$$

Además,

$$P_{\mathcal{L}^\perp} \left(i(1-z)(\Omega^* + \Gamma)P_{\mathcal{N}} + W(u_-^1, u_+^1)P_{\mathcal{N}^\perp} \right) P_{\mathcal{N}^\perp} = P_{\mathcal{L}^\perp} W(u_-^1, u_+^1)P_{\mathcal{N}^\perp} \quad (5.45)$$

define una biyección entre \mathcal{N}^\perp y \mathcal{L}^\perp (ver Lema 5.3.3 y Definición 5.3.1). De ello se deduce que hay funciones con valores matriciales $A(z), B(z), C(z), D(z)$ tales que

$$P \left(i(1-z)(\Omega^* + \Gamma)P_{\mathcal{N}} + W(u_-^1, u_+^1)P_{\mathcal{N}^\perp} + X(z) + Y(z)P_{\mathcal{N}^\perp} \right) Q = \begin{pmatrix} A(z) & B(z) \\ C(z) & D(z) \end{pmatrix} \quad (5.46)$$

y

$$\begin{aligned} A(z) &= i(1-z)\mathbf{A} + o(|1-z|), & B(z) &= o(1), \\ C(z) &= O(|1-z|), & D(z) &= \mathbf{D} + o(1), \end{aligned} \quad (5.47)$$

donde

$$\mathbf{A} := [P_{\mathcal{L}}(\Omega^* + \Gamma)P_{\mathcal{N}}]_{\alpha}^{\beta}, \quad \mathbf{D} := [P_{\mathcal{L}^\perp} W(u_-^1, u_+^1)P_{\mathcal{N}^\perp}]_{\gamma}^{\delta}, \quad (5.48)$$

son invertibles donde $[T]_{\theta}^{\eta}$ denota la representación matricial de una transformación

5. ANÁLISIS DEL LÍMITE A BAJAS ENERGÍAS DE LA MATRIZ DE
DISPERSIÓN ASOCIADA A LA ECUACIÓN DE SCHRÖDINGER DISCRETA

lineal T en términos de las bases θ, η . Aquí $\alpha = \{v_1, \dots, v_d\}$ es una base de \mathcal{N} , $\beta = \{e_1, \dots, e_L\}$ es una base de \mathcal{L} , y $\gamma = \{v_{d+1}, \dots, v_L\}$ y $\delta = \{e_{d+1}, \dots, e_d\}$, son bases de \mathcal{N}^\perp y \mathcal{L}^\perp , respectivamente.

Teorema 5.3.1. *Los coeficientes de transmisión satisfacen las siguientes propiedades:*

$$T_+^z = T_+^1 + o(1), \quad T_-^z = T_-^1 + o(1), \quad (5.49)$$

cuando z tiende a 1 en $\overline{\mathbb{D}}$, donde

$$T_+^1 := Q \begin{pmatrix} 2\mathbf{A}^{-1} & 0 \\ 0 & 0 \end{pmatrix} P, \quad T_-^1 := \left(Q \begin{pmatrix} 2\mathbf{A}^{-1} & 0 \\ 0 & 0 \end{pmatrix} P \right)^*. \quad (5.50)$$

Además

$$\begin{aligned} T_+^1 \mathbb{C}^L &= \mathcal{N}, & \text{Ker}(T_+^1) &= W(u_-^1, u_+^1) \mathbb{C}^L, \\ T_-^1 \mathbb{C}^L &= \mathcal{L}, & \text{Ker}(T_-^1) &= W(u_+^1, u_-^1) \mathbb{C}^L. \end{aligned} \quad (5.51)$$

Los coeficientes de reflexión satisfacen las siguientes propiedades:

$$R_+^z = R_+^1 + o(1), \quad R_-^z = R_-^1 + o(1), \quad (5.52)$$

cuando z tiende a 1 en \mathbb{S}^1 , donde

$$R_+^1 := \mathbf{1} - \Gamma T_+^1, \quad R_-^1 := \mathbf{1} - \Gamma^{-1} T_-^1 + o(1). \quad (5.53)$$

Además

$$\begin{aligned} \mathcal{L} &= (\mathbf{1} - R_+^1) \mathbb{C}^L, & \text{Ker}(T_+^1) &= \text{Ker}(\mathbf{1} - R_+^1), \\ \mathcal{N} &= (\mathbf{1} - R_-^1) \mathbb{C}^L, & \text{Ker}(T_-^1) &= \text{Ker}(\mathbf{1} - R_-^1). \end{aligned} \quad (5.54)$$

Demostración. La ecuación (5.47) implica que las matrices $D(z)$ y $A(z) - B(z)D(z)^{-1}C(z)$ son invertibles para $z \in \overline{\mathbb{D}}$ en una vecindad de 1, entonces usando la fórmula del com-

plemento de Schur, se sigue que (recuerde (5.21))

$$\begin{aligned}
 & (\nu^z)^{-1} \begin{pmatrix} A(z) & B(z) \\ C(z) & D(z) \end{pmatrix}^{-1} \\
 &= (\nu^z)^{-1} \begin{pmatrix} \mathbf{1} & 0 \\ -D(z)^{-1}C(z) & \mathbf{1} \end{pmatrix} \begin{pmatrix} (A(z) - B(z)D(z)^{-1}C(z))^{-1} & 0 \\ 0 & D(z)^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{1} & -B(z)D(z)^{-1} \\ 0 & \mathbf{1} \end{pmatrix} \\
 &= \begin{pmatrix} 2A^{-1} & 0 \\ 0 & 0 \end{pmatrix} + o(1), \tag{5.55}
 \end{aligned}$$

donde se usaron (5.21) y (5.47).

Debido a (5.43) de P , el núcleo de T_+^1 es generado por $\{e_{d+1}, \dots, e_L\}$, y son una base de $W(u_-^1, u_+^1)\mathbb{C}^L$: dado que P es unitario, $P^* = P^{-1} = (e_1 \dots e_L)$. Por lo tanto, el núcleo en juego es igual al núcleo de $\begin{pmatrix} \mathbf{1} & 0 \\ 0 & 0 \end{pmatrix} (e_1 \dots e_L)^{-1}$. Debido a la definición (5.43) de Q , la imagen de T_+^1 es generada por $\{v_1, \dots, v_d\}$, y estos vectores forman una base de \mathcal{N} . Esto prueba la primera línea en (5.51). El segundo sigue del hecho de que $T_-^1 = (T_+^1)^*$ (que se puede deducir de (5.24) y (5.34)).

A continuación, tomemos n lo suficientemente pequeños como para que $(u_-^1(n))^{-1}$ existe. Entonces, por continuidad, $(u_-^{1/z}(n))^{-1}$ existe, por z en una vecindad de 1. Utilizando (5.33) conduce a

$$R_+^z = (u_-^{1/z}(n))^{-1}u_-^z(n) - (u_-^{1/z}(n))^{-1}u_+^z(n)T_+^z \rightarrow (u_-^1(n))^{-1}u_-^1(n) - (u_-^1(n))^{-1}u_+^1(n)T_+^1, \tag{5.56}$$

como z tiende a 1. Tomando el límite $n \rightarrow -\infty$ en el lado derecho de (5.56), llegamos a la primera ecuación en (5.53) (ver también Lema 29 [15], en donde se prueba que $\Gamma\xi = \lim_{n \rightarrow -\infty} u_+^1(n)\xi$; ver también (5.15)). La segunda ecuación se obtiene de manera similar. Las ecuaciones (5.54) se siguen de (5.51), (5.53) y el hecho de que Γ es una biyección de \mathcal{N} en \mathcal{L} . \square

Bibliografía

- [1] «IUCN Red List of Threatened Species». <https://www.iucnredlist.org/>, 2020.
- [2] Aarts, Emile y Korst, Jan *Simulated annealing and Boltzmann machines*. Wiley-Intersci. Ser. Discrete Math. John Wiley & Sons, 1989.
- [3] Aceves, Toledo «El bosque mesófilo de montaña en México: amenazas y oportunidades para su conservación y manejo sostenible». *Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO)*, 2010.
- [4] Aguilar, Juan D. *Estructura y Composición de la Comunidad de Anfibios en Sitios Conservados, de Zonas altas del Volcán Tacaná, Chiapas, México*. Tesis de Licenciatura, Facultad de Ciencias, UNAM, 2019.
- [5] Aguilar, Juan D.; Ballesteros, Miguel y Hernández, Omar «Structure and composition of the amphibian community in conserved sites, from high areas of the Tacana Volcano, Chiapas, Mexico». *Preprint in preparation*.
- [6] Aktosun, Tuncay; Choque-Rivero, Abdon E. y Papanicolaou, Vassilis G. «On the bound states of the discrete Schrödinger equation with compactly supported potentials». *arXiv preprint arXiv:1809.08150*, 2018.
- [7] Aktosun, Tuncay; Klaus, Martin y van der Mee, Cornelis «Small-energy asymptotics of the scattering matrix for the matrix Schrödinger equation on the line». *Journal of Mathematical Physics*, 2001, **42(10)**, pp. 4627–4652.
- [8] Aktosun, Tuncay; Klaus, Martin y Weder, Ricardo «Small-energy analysis for the selfadjoint matrix Schrödinger operator on the half line. II». *Journal of Mathematical Physics*, 2014, **55(3)**, p. 032103.
- [9] Aktosun, Tuncay y Weder, Ricardo «Direct and Inverse Scattering for the Matrix Schrödinger Equation». *Applied Mathematical Sciences*, 2020, **203**.

- [10] Aldrich, M; Bubb, P; Hostettler, S y Van de Wiel, H «Bosques nublados tropicales montanos, Tiempo para la acción». *WWF International/IUCN*, 2000.
- [11] Aoki, Ichiro «Holological study of lakes from an entropy viewpoint-lake Mendota». *Ecological Modelling*, 1989, **45(2)**, pp. 81–93.
- [12] ——— *Entropy principle for the development of complex biotic systems: organisms, ecosystems, the Earth*. Elsevier, 2012.
- [13] Armstrong, Margaret *Basic Linear Geostatistics*. Springer, 1998.
- [14] Augustin, N. H.; Muggleston, M. A. y Buckland, S. T. «An Autologistic Model for the Spatial Distribution of Wildlife». *Journal of Applied Ecology*, 1996, pp. 339–347.
- [15] Ballesteros, Miguel; Córdova, Gerardo Franco; Garro, Guillermo y Schulz-Baldes, Hermann «Band edge limit of the scattering matrix for quasi-one-dimensional discrete Schrödinger operators». *arXiv preprint arXiv:2008.02177*, 2020.
- [16] Ballesteros, Miguel; Córdova, Gerardo Franco y Schulz-Baldes, Hermann «Analyticity properties of the scattering matrix for matrix Schrödinger operators on the discrete line». *arXiv preprint arXiv:2004.13099*, 2020.
- [17] Ballesteros, Miguel; Díaz-Ávalos, Carlos; Garro, Guillermo y Hernández, Omar «Low Density Distribution Modeling and Critically Endangered Species: Plectrohyla Guatemalensis». *Preprint in preparation*, 2020.
- [18] Ballesteros, Miguel y Falcón, José «Artificial Intelligence and Spatial Statistics: On the Distributions and Abundance of Species Subject to Repulsion and Rejection Areas». *Preprint in preparation*, 2020.
- [19] Ballesteros, Miguel y Garro, Guillermo «A Model and a Numerical Scheme for the Description of Distribution and Abundance of Individuals». *Preprint*, 2020.
- [20] Ballesteros, Miguel; Morales, Luis B. y Garro, Guillermo «Population Dynamics of Plectrohyla Sagorum». *Preprint in preparation*, 2020.
- [21] Barrio-Amorós, Cesar L; Grünwald, CI; Franz-Chavez, Hector; Mendoza, AM y La Forest, Brandon Thomas «Notes on natural history and call description of the critically endangered Plectrohyla avia (Anura: Hylidae) from Chiapas, Mexico». *Amphibian & Reptile Conservation*, 2016, **10(2)**, pp. 11–17.
- [22] Bartlett, Maurice Stevenson *The statistical analysis of spatial pattern*, tomo 15. Springer Science & Business Media, 2013.

-
- [23] Bertsimas, Dimitris y Nohadani, Omid «Robust optimization with simulated annealing». *J. Global Optim.*, 2010, **48(2)**, pp. 323–334.
- [24] Bertsimas, Dimitris; Tsitsiklis, John et al. «Simulated annealing». *Statistical science*, 1993, **8(1)**, pp. 10–15.
- [25] Besag, Julian E. «Nearest-neighbour Systems and the Auto-Logistic Model for Binary Data». *Journal of the Royal Statistical Society. Series B (Methodological)*, 1972, pp. 75–83.
- [26] ——— «Spatial Interaction and Statistical Analysis of Lattice Systems». *Stat. Soc. B*, 1974, **36**, pp. 721–741.
- [27] Boltzmann, Ludwig *Leçons sur la théorie des gaz*, tomo 2. Gauthier-Villars, Réédition jean gabay, paris, 1987.^a edición, 1902.
- [28] Bonneau, Mathieu; Gaba, Sabrina; Peyrard, Nathalie y Sabbadin, Regis «An adaptive sampling method using Hidden Markov Random Fields for the mapping of weed at the field scale». En: *EWRS 2011-4th Workshop of the EWRS Working Group Weeds and Biodiversity, Dijon, FRA., , 2011*.
- [29] Briassoulis, Helen; Kavrouidakis, Dimitris y Soulakellis, Nikolaos (Eds.) *The Practice of Spatial Analysis*. Springer, 2019.
- [30] Brocchi, P «Description d'un nouveau genre de phaneroglosse hylaeforme (*Plectrohyla guatemalensis*)». *Bulletin de la Société Philomathique de Paris, Series*, 1877, **7(1)**, pp. 92–93.
- [31] Brook, D «On the distinction between the conditional probability and the joint probability approaches in the specification of nearest-neighbour systems». *Biometrika*, 1964, **51(3/4)**, pp. 481–483.
- [32] Bumzahem, Carlos B y Smith, Hobart M «Additional records and descriptions of Mexican frogs of the genus *Plectrohyla*». *Herpetologica*, 1954, **10(1)**, pp. 61–66.
- [33] Catoni, Olivier «Rough large deviation estimates for simulated annealing: application to exponential schedules». *Ann. Probab.*, 1992, **20(3)**, pp. 1109–1146.
- [34] Cercignani, Carlo *The Boltzmann equation and its applications*, tomo 67 de *Applied Mathematical Sciences*. Springer, New York, 1988.
- [35] Chilès, Jean-Paul y Delfiner, Pierre *Geostatistics*. John Wiley & Sons, 2.^a edición, 2012. ISBN 978-0-470-18315-1. [2850475].

- [36] Clark, Nicholas J; Wells, Konstans; Dimitrov, Dimitar y Clegg, Sonya M «Co-infections and environmental conditions drive the distributions of blood parasites in wild birds». *Journal of Animal Ecology*, 2016, **85(6)**, pp. 1461–1470.
- [37] Clark, Nicholas J; Wells, Konstans y Lindberg, Oscar «Unravelling changing interspecific interactions across environmental gradients using Markov Random Fields». *Ecology*, 2018, **99(6)**, pp. 1277–1283.
- [38] Cliff, A. D. y Ord, J. K. *Spatial autocorrelation*. Pion, London, 1973.
- [39] ——— *Spatial processes: models and applications*. Pion, London, 1981.
- [40] Connelly, Scott; Pringle, Catherine M; Whiles, Matt R; Lips, Karen R; Kilham, Susan y Brenes, Roberto «Do tadpoles affect leaf decomposition in neotropical streams?» *Freshwater Biology*, 2011, **56(9)**, pp. 1863–1875.
- [41] Cortés-Gomez, AM; Ruiz-Agudelo, César Augusto; Valencia-Aguilar, Anyelet y Ladle, Richard J «Ecological functions of neotropical amphibians and reptiles: a review». *Universitas Scientiarum*, 2015, **20(2)**, pp. 229–245.
- [42] Cressie, Noel *Statistics for spatial data*. John Wiley & Sons, 2015.
- [43] Cressie, Noel y Lele, Subhash «New models for Markov Random Fields». *Journal of Applied probability*, 1992, **29(4)**, pp. 877–884.
- [44] Cruz-Parra, C. A. *Diversidad de anfibios y reptiles en Bosque Mesófilo y Cafetal, Reserva de la Biósfera Volcán Tacaná*. Universidad de Ciencias y Artes de Chiapas, 2012.
- [45] De La Cruz Martínez, V y Hernández Zuniga, R «Geología del Volcán Tacaná». *Geotermia: Revista Mexicana de Geoenergía*, 1986, **2(1)**, pp. 5–21.
- [46] Delahaye, Daniel; Chaimatanan, Supatcha y Mongeau, Marcel *Simulated Annealing: From Basics to Applications*, tomo 272 de *International Series in Operations Research & Management Science*, capítulo 1, pp. 1–35. Springer, Cham, 3.^a edición, 2019.
- [47] Díaz-Ávalos, Carlos «Spatial Modeling of Habitat Preferences of Biological Species using Markov Random Fields». *Journal of Applied Statistics*, 2007, **34(7)**, pp. 807–821.
- [48] Diggle, Peter J *Statistical analysis of spatial and spatio-temporal point patterns*. Chapman and Hall/CRC, 2013.

-
- [49] Dobruschin, PL «The Description of a Random Field by Means of Conditional Probabilities and Conditions of Its Regularity». *Theory of Probability & Its Applications*, 1968, **13**(2), pp. 197–224. doi: 10.1137/1113026.
- [50] Dudík, Miroslav; Phillips, Steven J y Schapire, Robert E «Performance guarantees for regularized maximum entropy density estimation». En: *International Conference on Computational Learning Theory*, pp. 472–486. Springer, 2004.
- [51] Duellman, W. E. «The Hylid Frogs of Middle America.» *Study Amph. Rept., Univ. Kansas*, 2001, (1), p. 694.
- [52] Duellman, William E *Patterns of distribution of amphibians: a global perspective*. JHU Press, 1999.
- [53] Duellman, William Edward y Campbell, Jonathan A *Hylid frogs of the genus Plectrohyla: systematics and phylogenetic relationships*. University of Michigan Museum of Zoology, 1992.
- [54] Elith, Jane; Phillips, Steven J; Hastie, Trevor; Dudík, Miroslav; Chee, Yung En y Yates, Colin J «A statistical explanation of MaxEnt for ecologists (with Appendices)». *Diversity and distributions*, 2011, **17**(1), pp. 43–57.
- [55] Faivovich, Julián; Haddad, Célio FB; Garcia, Paulo CA; Frost, Darrel R; Campbell, Jonathan A y Wheeler, Ward C «Systematic review of the frog family Hylidae, with special reference to Hylinae: Phylogenetic analysis and taxonomic revision». *Bulletin of the American Museum of natural History*, 2005, (294), pp. 1–240.
- [56] Frías-Alvarez, Patricia; Zúniga-Vega, J Jaime y Flores-Villela, Oscar «A general assessment of the conservation status and decline trends of Mexican amphibians». *Biodiversity and Conservation*, 2010, **19**(13), pp. 3699–3742.
- [57] Frost, Darrel R «*Plectrohyla guatemalensis* Brocchi, 1877. *Amphibian Species of the World: an Online Reference. Version 6.0*». <https://amphibiansoftheworld.amnh.org/Amphibia/Anura/Hylidae/Hylinae/Plectrohyla/Plectrohyla-guatemalensis>, Retrieved 7 July 2020.
- [58] García-Castillo, Mirna G; Rovito, Sean M; Wake, David B y Parra-Olea, Gabriela «A new terrestrial species of Chiropterotriton (Caudata: Plethodontidae) from Central Mexico.» *Zootaxa*, 2017, **4363**(4), pp. 489–505.
- [59] Gelfand, S. B. y Mitter, S. K. «Analysis of simulated annealing for optimization». En: *1985 24th IEEE Conference on Decision and Control*, pp. 779–786, 1985.

- [60] Geman, Stuart y Geman, Donald «Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1984, **PAMI-6(6)**, pp. 721–741. doi: 10.1109/TPAMI.1984.4767596.
- [61] — «Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images». *Journal of Applied Statistics*, 1993, **20(5-6)**, pp. 25–62.
- [62] Georgii, Hans-Otto *Gibbs Measures and Phase Transitions*. De Gruyter, 2.^a edición, 2011.
- [63] Georgina Santos-Barrera, Manuel Acevedo, Luis Canseco-Márquez «*Plectrohyla guatemalensis*. *The IUCN Red List of Threatened Species 2010*: e.T55885A11371639.» <http://dx.doi.org/10.2305/IUCN.UK.2010-2.RLTS.T55876A11367513.en>, 2010.
- [64] — «*Plectrohyla sagorum*. *The IUCN Red List of Threatened Species 2010*: e.T55885A11371639.» <http://dx.doi.org/10.2305/IUCN.UK.2010-2.RLTS.T55885A11371639.en>, 2010.
- [65] Gibbs, Josiah Willard *Elementary principles in statistical mechanics: developed with especial reference to the rational foundation of thermodynamics*. Yale University Press, New Haven, 1902.
- [66] Grimmett, Geoffrey y Welsh, Dominic (Eds.) *Disorder in physical systems: A volume in honour of John M. Hammersley*. Oxford University Press, 1990.
- [67] Grimmett, Geoffrey R. «A theorem about Random Fields». *Bulletin of the London Mathematical Society*, 1973, **5(1)**, pp. 81–84.
- [68] Gumpertz, Marcia L. «Autologistic Model of Spatial Pattern of Phytophthora Epidemic in Bell Pepper». *J Agric Biol Environ S*, 1997, **2**, pp. 131–156.
- [69] Hajek, Bruce «Cooling Schedules for Optimal Annealing». *Math. Oper. Res.*, 1988, **13(2)**, pp. 311–329.
- [70] Hamilton, Lawrence S «Mountain Cloud Forest Conservation and Research: A Synopsis». *Mountain Research and Development*, 1995, pp. 259–266.
- [71] Harris, David J «Estimating species interactions from observational data with Markov Networks». *bioRxiv*, 2015, p. 018861.
- [72] Harte, John *Maximum entropy and ecology: a theory of abundance, distribution, and energetics*. OUP Oxford, 2011.

-
- [73] Hartweg, Norman «Notes on the genus *Plectrohyla*, with descriptions of new species». *Occasional Papers of the Museum of Zoology*, 1941, **(437)**, pp. 1–31.
- [74] He, Fangliang; Zhou, Julie y Zhu, Hongtu «Autologistic Regression Model for the Distribution of Vegetation». *Journal of Agricultural, Biological, and Environmental Statistics*, 2003, **8(2)**, p. 205.
- [75] Heikkinen, Juha y Hogmander, Harri «Fully Bayesian Approach to Image Restoration with an Application in Biogeography». *Applied Statistics*, 1994, pp. 569–582.
- [76] Henderson, Darrall; Jacobson, Sheldon H. y Johnson, Alan W. «The theory and practice of simulated annealing». En: Fred Glover y Gary A. Kochenberger (Eds.), *Handbook of metaheuristics*, tomo 57 de *Internat. Ser. Oper. Res. Management Sci.*, pp. 287–319. Kluwer Acad. Publ., Boston, MA, 2003.
- [77] Hernández-Ordóñez, Omar; de Jesús Cervantes-López, Martín; González-Hernández, Adriana; Andresen, Ellen y Reynoso, Víctor Hugo «First record of the limestone rainfrog *Craugastor Psephosypharus* (Amphibia: Anura: Craugastoridae) in Mexico». *Revista mexicana de biodiversidad*, 2017, **88(1)**, pp. 260–264.
- [78] Hidalgo, Emanuel Stanley Morán y Ruballo-Marroquín, Néstor Esaú «Nueva localidad de *Plectrohyla Guatemalensis* (BROCCHI, 1987) (Anura: Hylidae) en el departamento de Chalatenango, El Salvador». *Revista Biodiversidad Neotropical*, 2012, **2(2)**, pp. 126–130.
- [79] Hinton, Don B; Klaus, Martin y Shaw, J Ken «Half-bound states and Levinson's theorem for discrete systems». *SIAM journal on mathematical analysis*, 1991, **22(3)**, pp. 754–768.
- [80] Hoef, Jay M. Ver; Peterson, Erin E.; Hooten, Mevin B.; Hanks, Ephraim Mont y Fortin, Marie José «Spatial Autoregressive Models for Statistical Inference from Ecological Data». *Ecological Monographs*, 2018, **88(1)**, pp. 36–59.
- [81] Hoeting, Jennifer A.; Leecaster, Molly y Bowden, David «An Improved Model for Spatially Correlated Binary Responses». *Journal of Agricultural, Biological, and Environmental Statistics*, 2000, pp. 102–114.
- [82] INEGI «Continuo de Elevaciones Mexicano (CEM)». <https://www.inegi.org.mx/app/geo2/elevacionesmex/>.
- [83] Ingber, L. «Simulated annealing: practice versus theory». *Math. Comput. Modelling*, 1993, **18(11)**, pp. 29–57.
-

BIBLIOGRAFÍA

- [84] Isaaks, Edward H y Srivastava, Mohan R *Applied Geostatistics*. Oxford University Pres, 1989.
- [85] Jaynes, Edwin T «Information theory and statistical mechanics». *Physical review*, 1957, **106(4)**, p. 620.
- [86] — «Information theory and statistical mechanics. II». *Physical review*, 1957, **108(2)**, p. 171.
- [87] Jorgensen, Sven E *Thermodynamics and ecological modelling*. CRC press, 2018.
- [88] Juan, Pablo; Díaz-Ávalos, Carlos; Mejía-Domínguez, Nancy R. y Mateu, Jorge «Hierarchical spatial modeling of the presence of Chagas disease insect vectors in Argentina. A comparative approach». *Stoch. Environ. Res. Risk Assess.*, 2017, **31(2)**, pp. 461–479.
- [89] Kaiser, Mark S y Cressie, Noel «The Construction of Multivariate Distributions from Markov Random Fields». *Journal of Multivariate Analysis*, 2000, **73(2)**, pp. 199–220.
- [90] Kapur, Jagat Narain y Kesavan, Hiremaglur K «Entropy optimization principles and their applications». En: *Entropy and energy dissipation in water resources*, pp. 3–20. Springer, 1992.
- [91] Kindermann, Ross y Snell, J. Laurie *Markov Random Fields and their applications*, tomo 1 de *Contemporary Mathematics*. American Mathematical Society, Providence, R.I., 1980. [620955].
- [92] — *Markov Random Fields and their applications*. AMS, 1980.
- [93] Kirkpatrick, S.; Gelatt, C. D. y Vecchi, M. P. «Optimization by Simulated Annealing», 1983, **220(4598)**, pp. 671–680. ISSN 0036-8075.
- [94] Klaus, Martin «Low-energy behaviour of the scattering matrix for the Schrodinger equation on the line». *Inverse Problems*, 1988, **4(2)**, p. 505.
- [95] Köhler, Gunther *Amphibians of Central America*. Herpeton, Verlag Elke Kohler, 2011.
- [96] Koller, Daphne y Friedman, Nir *Probabilistic graphical models: principles and techniques*. MIT press, 2009.
- [97] Laarhoven, P. J. M. y Aarts, E. H. L. *Simulated annealing: theory and applications*, tomo 37, 1987.

- [98] Last, Günter y Penrose, Mathew *Lectures on the Poisson Process*. Cambridge University Press, 2017.
- [99] Lee, Julian C y JC, LEE «An ecogeographic analysis of the herpetofauna of the Yucatan Peninsula», 1980.
- [100] Legendre, Pierre «Spatial Autocorrelation: Trouble or New Paradigm?» *Ecology*, 1993, **74(6)**, pp. 1659–1673.
- [101] Li, Stan Z *Markov Random Field Modeling in Image Analysis*. Springer, 2009.
- [102] Lindberg, Oscar et al. *Markov Random Fields in Cancer Mutation Dependencies*. Department of Mathematics and Statistics, University of Turku, 2016.
- [103] Matheron, Georges «Krigage d'un Panneau Rectangulaire par sa Périphérie». *Note géostatistique*, 1960, **28**.
- [104] — «Exemple de krigage d'un panneau riche». *Note géostatistique*, 1961, **28**.
- [105] — *Traité de géostatistique appliquée*, tomo 1 and 2. Editions Technip, 1962.
- [106] — *Le krigage universel*. École nationale supérieure des mines de Paris, Paris, 1969.
- [107] — *Matheron's theory of regionalized variables*. International Association for Mathematical Geosciences. Studies in Mathematical Geosciences. Oxford University Press, Oxford, 2019. Edited by Vera Pawlowsky-Glahn and Jean Serra.
- [108] Maxwell, Brian *Advanced Simulated Annealing*. Clanrye International, 2015.
- [109] McCranie, JR «Specific status of the Montaña de Celaque Honduran frogs previously referred to as *Plectrohyla Guatemalensis* (Anura: Hylidae: Hylinae)». *Mesoamerican Herpetology*, 2017, **4**, pp. 390–401.
- [110] Mendelson, JR; Brodie Jr, ED; Malone, JH; Acevedo, ME; Baker, MA; Smatresk, NJ y Campbell, JA «Factors associated with the catastrophic decline of a cloud-forest frog fauna in Guatemala». *Revista de Biología Tropical*, 2004, **52(4)**, pp. 991–1000.
- [111] Mendoza-Almeralla, Cinthya; Burrowes, Patricia y Parra-Olea, Gabriela «La quítridiomicosis en los anfibios de México: una revisión». *Revista Mexicana de Biodiversidad*, 2015, **86(1)**, pp. 238–248.

BIBLIOGRAFÍA

- [112] Meza-Parral, Yocoyani y Pineda, Eduardo «Amphibian diversity and threatened species in a severely transformed Neotropical region in Mexico». *PloS ONE*, 2015, **10(3)**, p. e0121652.
- [113] Mitra, Debasis; Romeo, Fabio y Sangiovanni-Vincentelli, Alberto «Convergence and Finite-Time Behavior of Simulated aAnnealing». *Adv. in Appl. Probab.*, 1986, **18(3)**, pp. 747–771.
- [114] Muñoz Alonso, Luís Antonio «Riqueza, diversidad y estatus de los anfibios amenazados en el Sureste de México; una evaluación para determinar las posibles causas de la declinación de sus poblaciones». *El Colegio de la Frontera Sur*, 2010.
- [115] Nolte, Andreas y Schrader, Rainer «A Note on the Finite Time Behavior of Simulated Annealing». *Math. Oper. Res.*, 2000, **25(3)**, pp. 476–484.
- [116] Olson, David M y Dinerstein, Eric «The Global 200: Priority ecoregions for global conservation». *Annals of the Missouri Botanical garden*, 2002, pp. 199–224.
- [117] Parra-Olea, Gabriela; Flores-Villela, Oscar y Mendoza-Almeralla, Cinthya «Biodiversidad de anfibios en México». *Revista mexicana de biodiversidad*, 2014, **85**, pp. 460–466.
- [118] Phillips, Steven J; Anderson, Robert P y Schapire, Robert E «Maximum entropy modeling of species geographic distributions». *Ecological modelling*, 2006, **190(3-4)**, pp. 231–259.
- [119] Phillips, Steven J y Dudík, Miroslav «Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation». *Ecography*, 2008, **31(2)**, pp. 161–175.
- [120] Phillips, Steven J; Dudík, Miroslav y Schapire, Robert E «A maximum entropy approach to species distribution modeling». En: *Proceedings of the twenty-first international conference on Machine learning*, p. 83, 2004.
- [121] Polpo, Adriano; Stern, Julio; Louzada, Francisco; Izbicki, Rafael y Takada, Hellinton «Bayesian Inference and Maximum Entropy Methods in Science and Engineering». *Springer Proceedings in Mathematics & Statistics*, 2018.
- [122] Ponce-Reyes, Rocio; Nicholson, Emily; Baxter, Peter WJ; Fuller, Richard A y Possingham, Hugh «Extinction risk in cloud forest fragments under climate change and habitat loss». *Diversity and Distributions*, 2013, **19(5-6)**, pp. 518–529.

-
- [123] Preston, Christopher J *Gibbs states on countable sets*. Cambridge University Press, 1974.
- [124] Ranvestel, Anthony W; Lips, Karen R; Pringle, Catherine M; Whiles, Matt R y Bixby, Rebecca J «Neotropical tadpoles influence stream benthos: evidence for the ecological consequences of decline in amphibian populations». *Freshwater Biology*, 2004, **49(3)**, pp. 274–285.
- [125] Ricketts, Taylor H; Dinerstein, Eric; Boucher, Tim; Brooks, Thomas M; Butchart, Stuart HM; Hoffmann, Michael; Lamoreux, John F; Morrison, John; Parr, Mike; Pilgrim, John D et al. «Pinpointing and preventing imminent extinctions». *Proceedings of the National Academy of Sciences*, 2005, **102(51)**, pp. 18497–18501.
- [126] Ripley, B. D. «Modelling spatial patterns». *Journal of the Royal Statistical Society: Series B (Methodological)*, 1977, **39(2)**, pp. 172–192.
- [127] — «Spatial point pattern analysis in ecology». En: *Develoments in Numerical Ecology*, pp. 407–429. Springer, 1987.
- [128] — *Statistical inference for spatial processes*. Cambridge university press, 1991.
- [129] Robert, Christian y Casella, George *Monte Carlo statistical methods*. Springer, 2013.
- [130] Robini, Marc C. «Theoretically Grounded Acceleration Techniques for Simulated Annealing». En: Ivan Zelinka; Václav Snášel y Ajith Abraham (Eds.), *Handbook of Optimization*, tomo 38 de *Intelligent Systems Reference Library*, pp. 311–335. Springer, Berlin Heidelberg, 2013.
- [131] Sandoval-Comte, Adriana; Pineda, Eduardo; Rovito, Sean M y Luria-Manzano, Ricardo «A new species of *Isthmura* (Caudata: Plethodontidae) from the montane cloud forest of central Veracruz, Mexico». *Zootaxa*, 2017, **4277(4)**, pp. 573–582.
- [132] Santos-Barrera, G «Enfermedades infecciosas en poblaciones de anfibios». *Biodiversitas*, 2004, **56**, pp. 1–6.
- [133] Schneider, Eric D y Kay, James J «Life as a manifestation of the second law of thermodynamics». *Mathematical and computer modelling*, 1994, **19(6-8)**, pp. 25–48.
- [134] SEMARNAT y CONANP «Programa de Manejo Reserva de la Biosfera Volcán Tacaná». *Secretaría de Medio Ambiente y Recursos Naturales, Comisión Nacional de Áreas Naturales Protegidas*, 2013.

- [135] Sherman, S «Markov Random Fields and Gibbs Random Fields». *Israel Journal of Mathematics*, 1973, **14(1)**, pp. 92–103.
- [136] Spitzer, Frank «Markov Random Fields and Gibbs Ensembles». *The American Mathematical Monthly*, 1971, **78(2)**, pp. 142–154.
- [137] Stuart, S. *Threatened amphibians of the world*. Lynx Edicions, 2008.
- [138] Stuart, Simon N; Chanson, Janice S; Cox, Neil A; Young, Bruce E; Rodrigues, Ana SL; Fischman, Debra L y Waller, Robert W «Status and trends of amphibian declines and extinctions worldwide». *Science*, 2004, **306(5702)**, pp. 1783–1786.
- [139] Sullivan, Wayne G «Finite range random fields and energy fields». *Journal of Mathematical Analysis and Applications*, 1973, **44(3)**, pp. 710–724.
- [140] Tsitsiklis, John N. «A Survey of Large Time Asymptotics of Simulated Annealing Algorithms». En: Wendell Fleming y Pierre-Louis Lions (Eds.), *Stochastic Differential Systems, Stochastic Control Theory and Applications*, tomo 10 de *IMA Vol. Math. Appl.*, pp. 583–599. Springer, New York, 1988.
- [141] ——— «Markov chains with rare transitions and simulated annealing». *Math. Oper. Res.*, 1989, **14(1)**, pp. 70–90.
- [142] Tsuzuki, Marcos de Sales Guerra y de Castro Martins, Thiago (Eds.) *Simulated annealing: strategies, potential uses and advantages*. Nova Publishers, New York, 2014.
- [143] Tyler, M «Distribution patterns of amphibians in the Australo-Papuan region». John Hopkins University Press, 1999.
- [144] Urakawa, Hajime «Convergence Rates to Optimal Distribution of the Boltzmann Machine on Simulated Annealing». *Interdisciplinary Information Sciences*, 2009, **15(2)**, pp. 291–299.
- [145] Urbina-Cardona, J Nicolás y Loyola, Rafael D «Applying Niche-Based Models to Predict Endangered-Hyloid Potential Distributions: Are Neotropical Protected Areas Effective Enough?» *Tropical Conservation Science*, 2008, **1(4)**, pp. 417–445.
- [146] Černý, V. «Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm». *J. Optim. Theory Appl.*, 1985, **45(1)**, pp. 41–51.

- [147] Vié, Jean-Christophe; Hilton-Taylor, Craig y Stuart, Simon N *Wildlife in a changing world: an analysis of the 2008 IUCN Red List of threatened species*. IUCN, 2009.
- [148] Wackernagel, Hans *Multivariate Geostatistics: An Introduction with Applications*. Springer Science & Business Media, 2013.
- [149] Waller, Lance A y Gotway, Carol A *Applied Spatial Statistics for Public Health Data*, tomo 368. John Wiley & Sons, 2004.
- [150] Wei, Peng y Pan, Wei «Network-based genomic discovery: application and comparison of Markov random-field models». *J. R. Stat. Soc. Ser. C*, 2010, **59(1)**, pp. 105–125.
- [151] Wei, Zhi y Li, Hongzhe «A Markov random field model for network-based analysis of genomic data». *Bioinformatics*, 2007, **23(12)**, pp. 1537–1544.
- [152] — «A hidden spatial-temporal Markov random field model for network-based analysis of time course gene expression data». *Ann. Appl. Stat.*, 2008, **2(1)**, pp. 408 – 429.
- [153] Wells, Kentwood D *The ecology and behavior of amphibians*. University of Chicago Press, 2010.
- [154] Whittle, Peter «On stationary processes in the plane». *Biometrika*, 1954, pp. 434–449.
- [155] — «Stochastic-processes in several dimensions». *Bulletin of the International Statistical Institute*, 1963, **40(2)**, pp. 974–994.
- [156] Wilson, Larry David y McCranie, James R «The herpetofauna of the cloud forests of Honduras». *Amphibian and Reptile Conservation*, 2004, **3(1)**, p. 34.
- [157] Winkler, Gerhard *Image analysis, random fields and Markov chain Monte Carlo methods*, tomo 27 de *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 2.^a edición, 2003. A mathematical introduction, With 1 CD-ROM (Windows), Stochastic Modelling and Applied Probability.
- [158] — *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: a Mathematical Introduction*. Springer Science & Business Media, 2nd.^a edición, 2012.

BIBLIOGRAFÍA

- [159] Wu, Hulin y Huffer, F Red W «Modelling the distribution of plant species using the autologistic regression model». *Environmental and ecological Statistics*, 1997, **4(1)**, pp. 31–48.
- [160] Young, Bruce E; Stuart, Simon N y Chanson, JS «Joyas que Están Desapareciendo: El Estado de los Anfibios en el Nuevo Mundo». *NatureServe*, 2004.
- [161] Yuan, Ming y Kendzierski, Christina «Hidden Markov Models for Microarray Time Course Data in Multiple Biological Conditions». *J. Amer. Statist. Assoc.*, 2006, **101(476)**, pp. 1323–1332.
- [162] Zheng, Yanbing y Zhu, Jun «Markov Chain Monte Carlo for a Spatial-Temporal Autologistic Regression Model». *Journal of Computational and Graphical Statistics*, 2008, **17(1)**, pp. 123–137.
- [163] Zhu, Jun; Huang, Hsin-Cheng y Wu, Jungpin «Modeling Spatial-Temporal Binary Data using Markov Random Fields». *Journal of Agricultural, Biological, and Environmental Statistics*, 2005, **10(2)**, p. 212.
- [164] Zhu, Jun; Rasmussen, Jakob G; Møller, Jesper; Aukema, Brian H y Raffa, Kenneth F «Spatial-temporal modeling of forest gaps generated by colonization from below-and above-ground bark beetle species». *Journal of the American Statistical Association*, 2008, **103(481)**, pp. 162–177.