

UNIVERSIDAD DE COSTA RICA
SISTEMA DE ESTUDIOS DE POSGRADO

**APLICACIÓN DE TECNOLOGÍAS DE INFORMACIÓN
PARA MODELAR LA DISTRIBUCIÓN DE PLAGAS**

Trabajo final de investigación aplicada sometido a la consideración de la Comisión del Programa de Estudios de Posgrado en Computación e Informática para optar al grado y título de Maestría Profesional en Computación e Informática

ADRIAN ALONSO HERRERA CERVANTES

Ciudad Universitaria Rodrigo Facio, Costa Rica

2012

A mi familia, que siempre ha estado dispuesta
a apoyarme en cada una de mis aventuras
y planes en esta vida

“Este trabajo final de investigación aplicada fue aceptado por la Comisión del Programa de Estudio de Posgrado en Computación e Informática de la Universidad de Costa Rica, como requisito parcial para optar al grado y título en Maestría Profesional en Computación e Informática”

Dr. Carlos Rojas Alvarado

Representante de la Decana Sistema de Estudios de Posgrado

MSc. Manuel Vargas del Valle

Profesor Guía

Dra. Gabriela Marín Raventós

Representante del Director Programa de Posgrado en Informática y Computación

Adrián Alonso Herrera Cervantes

Sustentante

Tabla de contenido

Introducción	1
Descripción del Problema	3
Justificación.....	4
Marco Teórico	6
Concepto de Plaga.....	6
Modelado de Distribución de Especies	6
Datos Biológicos	9
Datos Ambientales	10
Algoritmos de Modelado.....	12
Evaluación de los Resultados de un Modelo	14
Teoría de Dinámica de Poblaciones.....	15
Ley de Malthus	16
La Ecuación de Verhulst	17
Las Ecuaciones de Lotka-Volterra	18
Antecedentes	20
Investigación Relacionada	20
Proyecto del Curso PF3342 – Laboratorio de Informática Aplicada a la Ecología y a la Conservación de la Biodiversidad.....	22
Objetivos	25
General.....	25
Específicos.....	25
Delimitación	26
Metodología.....	27
Análisis de Resultados	28
Implementación de la Aplicación	28
Pruebas de la Aplicación con Datos Reales de Plagas.....	42
Pruebas de la Proyección de Crecimiento Poblacional.....	47
Conclusiones	51
Bibliografía	52

Resumen

Actualmente, uno de las grandes amenazas para la seguridad alimentaria de un país es la proliferación de poblaciones de plagas que puedan atacar los cultivos destinados a la satisfacción de las necesidades alimenticias de su población. Una forma de atacar este problema es mediante el uso de tecnologías que utilicen toda la información disponible para tomar las decisiones que permitan planificar las acciones contra las especies consideradas plagas.

Para ayudar a generar información para la toma de decisiones contra las plagas se propone una aplicación que reúne varias tecnologías de información *open source* de diversos campos como los sistemas de información geográfica, las bases de datos, el modelado de distribuciones espaciales y el desarrollo de software. El objetivo de esta aplicación es proporcionar una herramienta de fácil utilización, con una arquitectura escalable y que permita obtener tanto los modelos de distribución como las estadísticas de evaluación de los mismos.

La aplicación permite tanto la manipulación de los datos utilizados en los modelos, como la generación de los mismos y la posterior consulta de los resultados obtenidos en el pasado. Así mismo, proporciona una herramienta básica para generar modelos de dinámica población utilizando como base modelos generales desarrollados por diversos estudiosos desde el siglo XIX.

La aplicación fue puesta a prueba, para comprobar los resultados para una especie de la mosca de la fruta (*Ceratitis capitata*) que se considera como plaga cuarentenaria en Costa Rica. Se ejecutaron los modelos utilizando dos algoritmos distintos y dos conjuntos de variables distintos. Luego de las pruebas el algoritmo MAXENT pareció no ser muy efectivo para el problema planteado en la prueba. Sin embargo, el algoritmo GARP proporcionó resultados bastante buenos en uno de los modelos que se aproxima bastante a la realidad de esta plaga en nuestro país.

La conclusión más relevante de este trabajo es que es posible conjuntar tecnologías de información de diferentes ámbitos para crear una aplicación más robusta y escalable que ofrezca una alternativa al problema planteado, y adicionalmente, deje abiertas las opciones para mejorarla y ampliarla.

Lista de Cuadros

Cuadro 1. Modelos de distribución de especies, año de aparición, método utilizado y referencias.	12
Cuadro 2. Tabla comparativa entre herramientas.	42

Lista de Figuras

Figura 1. Pantalla inicial de la aplicación desarrollada en el curso PF3342.	23
Figura 2. Pantalla para ingresar los parámetros de un modelo de distribución en la aplicación desarrollada en el curso PF3342.	24
Figura 3. Pantalla que muestra los resultados de un modelo generado por la aplicación desarrollada en el curso PF3342.	24
Figura 4. Arquitectura de la aplicación.	31
Figura 5. Pantalla para administrar las variables ambientales de los modelos.	33
Figura 6. Pantalla para administrar la inclusión de observaciones de una plaga en un modelo.....	33
Figura 7. Formulario para registrar una nueva observación de una especie.	34
Figura 8. Pantalla para consultar la lista de modelos generados por la aplicación.....	35
Figura 9. Pantalla para parametrizar la generación de un modelo de distribución.	36
Figura 10. Pantalla principal de la aplicación Maxent	37
Figura 11. Pantalla principal de la aplicación Desktop GARP.....	38
Figura 12. Pantalla de generación de un modelo en OpenModeller Desktop.	39
Figura 13. Pantalla para visualizar el resultado de un modelo de distribución.	40
Figura 14. Pantalla para parametrizar un modelo de comportamiento poblacional.	41
Figura 15. Resultado usando el algoritmo MAXENT con 19 variables climáticas.....	44
Figura 16. Resultado usando el algoritmo GARP con 19 variables climáticas.....	44
Figura 17. Resultado usando el algoritmo MAXENT con 8 variables climáticas.....	45
Figura 18. Resultado usando el algoritmo GARP con 8 variables climáticas.....	46
Figura 19. Resultado de un modelo de crecimiento poblacional utilizando el modelo de Lotka-Volterra.	48
Figura 20. Resultado de un modelo de crecimiento poblacional utilizando el modelo de Verhulst.....	49

Introducción

La población mundial ha llegado a alcanzar más de 7000 millones de habitantes el año pasado y sigue en constante crecimiento. Esta situación hace que el ser humano busque la manera de satisfacer las necesidades alimenticias de esta población. La situación se complica aún más si se tiene en cuenta que el calentamiento global está afectando las condiciones climáticas en las que se cultiva. Sin embargo, nuestros agricultores no deben lidiar solo con este cambio, sino también con los efectos de este calentamiento en las poblaciones de plagas que atacan los cultivos y ponen en riesgo el alimento del ser humano (CNN, 2011).

Este Trabajo Final de Investigación Aplicada (TFIA) busca plantear y demostrar que es posible conjuntar diferentes herramientas del campo de las tecnologías de información para crear una herramienta que coadyuve en el control y prevención de plagas en los cultivos de nuestro país.

En la Descripción del Problema se plantean las amenazas que tienen que afrontar las naciones del planeta para establecer planes de seguridad alimentaria que les proporcionen los requerimientos alimenticios a sus poblaciones. En la Justificación se plantea como diversas instituciones recomiendan la inversión en investigación y desarrollo para atacar las amenazas a la seguridad alimentaria y la relevancia de un proyecto como este para nuestro país.

En el Marco Teórico se detallan los estudios y la teoría detrás de la generación de los modelos de distribución de especies, los datos a utilizar, los algoritmos, las consideraciones a la hora de trabajar con este tipo de modelos así como la evaluación que se debe hacer de los mismos si se quieren tomar decisiones. Así mismo, esta sección proporciona una introducción y explicación a varios modelos de dinámica y crecimiento poblacional que se han desarrollado a lo largo de la historia.

Los Antecedentes hacen una breve descripción del trabajo previo realizado en el curso PF3342 – Laboratorio de Informática Aplicada a la Ecología y a la Conservación de la Biodiversidad y que sirve de base a este TFIA, así como la descripción de otros trabajos de investigación realizados alrededor del tema de las plagas y los modelos de distribución de especies.

La sección de Objetivos detalla que la meta de este TFIA es la creación de una aplicación para crear modelos de distribuciones de plagas a partir de tecnologías de información *open source*. La sección de Delimitación presenta los casos de uso que limitarán el alcance de la herramienta. La Metodología describe los pasos que se siguieron para trabajar en este TFIA.

Finalmente, el Análisis de Resultados describe los resultados obtenidos a nivel de la arquitectura de implementación de la herramienta, los resultados en la generación de modelos de distribución de especies y los resultados en la generación de modelos de dinámica poblaciones; y la sección de Conclusiones describe las principales conclusiones alcanzadas gracias al trabajo realizado en este TFIA.

Descripción del Problema

En diversos medios de comunicación es posible observar noticias relacionadas con el tema de la seguridad alimentaria para un país. El Instituto Interamericano de Cooperación para la Agricultura (IICA) define la seguridad alimentaria como “la existencia de condiciones que posibilitan a los seres humanos tener acceso físico, económico y de manera socialmente aceptable a una dieta segura, nutritiva y acorde con sus preferencias culturales, que les permita satisfacer sus necesidades alimentarias y vivir de una manera productiva y saludable” (IICA, 2012).

Las condiciones a las que hace referencia el IICA son la disponibilidad física de los alimentos en cantidades y calidad suficientes, el acceso de las personas a los alimentos ya sea por medio de recursos económico o de otra índole, el logro de un bienestar nutricional que permita satisfacer las necesidades de las personas y la estabilidad en acceso a los alimentos de forma que no se presente escasez o carencia de los mismos debido a circunstancias ajenas a las personas (IICA, 2012).

En un reporte de una conferencia sobre la seguridad alimentaria mundial organizada por la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO por sus siglas en inglés) en Roma en el 2008, se establece dentro de las bases de conocimiento, que el movimiento de plagas de plantas y enfermedades animales a través de fronteras físicas y políticas amenazan la seguridad alimentaria y generan una preocupación pública en el mundo; así mismo, se estableció que existen evidencias de que el cambio climático está alterando la distribución, la incidencia y la intensidad de las plagas y enfermedades animales (FAO, 2008).

Es así como el control de las plagas que atacan a los cultivos en el territorio nacional, se convierte en un tema importante si se quiere asegurar que el país tendrá las condiciones apropiadas, definidas por la IICA, para alcanzar un nivel de seguridad alimentaria que permita el bienestar de la población costarricense.

Justificación

Debido a los retos que se enfrentan con respecto al mejoramiento de los procesos agrícolas con el fin de evitar una crisis a nivel alimenticio, el IICA propone que uno de los bastiones para hacer frente a los cambios sea retomar la inversión en investigación y desarrollo. Menciona que el aprovechamiento del potencial de las tecnologías de información debe beneficiar a los productores para alcanzar no solo mayores rendimientos económicos, sino una dieta saludable que sea capaz de satisfacer las necesidades nutricionales de la población (Brathwaite, 2009).

La FAO, adicionalmente, nos habla que, dentro de las respuestas que debería de tomarse en cuenta para atender los riesgos de las plagas para la seguridad alimentaria, hay que trabajar con información que permita prever la distribución, gravedad e incidencia de las plagas usando métodos e instrumentos de análisis de riesgo. Adicionalmente, dice la misma FAO, es importante una alerta temprana y medidas de prevención mediante mecanismos de control y detección tempranos para evitar los mayores costos generados por la erradicación y el control (FAO, 2008).

El presente proyecto surge al buscar desarrollar una herramienta, usando tecnologías de información, que permita generar información para comprender cómo se pueden distribuir las poblaciones de una plaga. Cuando los recursos para la atención de una plaga pueden llegar a ser escasos y es necesario asegurar que los cultivos no se verán afectados, el contar con la mayor cantidad de información posible cobra importancia.

Actualmente el país se encuentra en un período de ajustes a nivel presupuestario en todos los niveles de la sociedad, esto puede afectar la atención de una plaga en los cultivos. La importancia de este proyecto radica en que información de alerta temprana, como lo puede ser una proyección de distribución poblacional, puede contribuir a la planificación de los recursos para controlar una plaga.

Finalmente, es del interés del autor de este TFIA el realizar un proyecto que proponga una idea que ayude a encontrar soluciones a un problema presente en nuestra realidad nacional y que de una u otra forma concierne a toda la población costarricense. Creo que el seguimiento y

tratamiento de la plagas en nuestro país es un tema que cae dentro de esta descripción, pues como se expuso anteriormente, la presencia, ausencia y control de las mismas tiene repercusiones en la oferta de productos alimenticios para consumo humano.

Marco Teórico

Concepto de Plaga

La FAO (1995) define plaga como "cualquier especie, raza o biotipo vegetal o animal o agente patógeno dañino para las plantas o productos vegetales".

Otras definiciones (Universidad Nacional de Colombia, 2013) nos indican que se puede considerar como plaga:

- Un animal o planta cuyas actividades interfieren con la salud humana o su bienestar y que afecta sus ingresos económicos.
- Una criatura que reduce la disponibilidad, calidad o valor de un recurso importante para la humanidad
- Un animal o planta cuya densidad de población excede un nivel arbitrario no aceptable para la humanidad, el cual resulta en un daño económico.

De todas estas acepciones, se puede sintetizar una definición que las incluya a todas y considerar que una plaga es cualquier especie, raza o biotipo vegetal o animal o agente patógeno dañino cuya densidad de población excede un nivel no aceptable para la humanidad, lo cual reduce la disponibilidad, calidad o valor de un recurso importante y con ello afecta la calidad de vida del ser humano, usualmente, en el aspecto económico.

Modelado de Distribución de Especies

Guisan y Thuiller (2005) por un lado, y Pearson (2003) por otro, establecen que la forma usual de modelar y predecir las posibles ubicaciones de una especie consiste en cuantificar de alguna forma la relación de la especie con las condiciones ambientales de los espacios donde habitan. Los primeros autores antes mencionados resaltan que los primeros esfuerzos, primero en 1924 para predecir el comportamiento de un cactus invasivo en Australia y luego en 1963 para predecir la distribución de distintas especies en Europa, buscaron correlacionar la distribución de las especies y el clima.

Basado en lo anterior y de acuerdo con lo expuesto por Guisan y Thuiller, y Franklin (2010), es que podemos definir el modelado de distribución de especies (o SDM por sus siglas en inglés) como modelos empíricos que relacionan observaciones de especies en el campo con variables ambientales de predicción basados, usualmente en modelos estadísticos, con lo cual es posible

extrapolar la distribución de especies en el tiempo y en el espacio. Para Peterson (2003), el hecho de poder realizar esta relación permite, principalmente por el uso de información de observaciones de una especie, el poder aplicar estos modelos a casi cualquier especie sin importar si la misma ha sido bien estudiada o no.

Los usos que se le puede dar a la información generada a partir de un SDM son múltiples, Franklin destaca los siguientes:

- **Diseño de reservas y planes de conservación:** el realizar mapas de distribución de una especie sirve de herramienta para definir nuevas zonas de protección y definir planes de conservación para las reservas existentes o futuras.
- **Análisis de impacto y manejo de recursos:** si se cuenta con información sobre la distribución o potencial distribución de una especie, se puede medir el impacto que una actividad puede tener sobre el hábitat de la misma, así mismo es posible definir mejor el destino de los recursos con los que se cuenta. Por ejemplo económicos y humanos.
- **Restauración ecológica:** la información de los SDM puede ser utilizada para buscar ubicaciones apropiadas para la reintroducción de especies en nuevos hábitats que cumplan con las necesidades de la especie para sobrevivir.
- **Riesgo e impacto de especies invasivas:** las especies invasivas pueden tener un impacto económico y ecológico muy grande, es por ellos que es uno de los principales usos de los SDM. Usualmente se utilizan con dos fines: predecir potenciales nuevas zonas geográficas donde la especie se podría establecer con base en las condiciones del ambiente nativo y utilizar la información de nuevos asentamientos para ubicar potenciales nuevos asentamientos. Este uso de los SDM es el objetivo del proyecto que se detalla en este documento y que busca proporcionar una herramienta que facilite su uso para la toma de decisiones en este sentido.
- **Estudio del efecto del cambio climático en los ecosistemas:** se sabe que existe una relación entre las especies y el clima. Se asume que éstas buscan estar en equilibrio con

las condiciones climáticas de sus hábitats. Los SDM puede ayudar a modelar los cambios en las distribuciones geográficas de las especies debidas a los cambios ocurridos en las distintas variables climáticas que conforman el entorno en el que las especies se desarrollan.

Los modelos de distribución de especies se basan en dos asunciones o principios (Guisan y Thuiller, 2005):

- **Postulado de equilibrio:** La información de condiciones climáticas y de las observaciones de las especies se limitan a un tiempo y espacio específico, por lo que la relación que se pueda establecer entre estos datos proporciona una fotografía de una situación en un momento dado. Un postulado conveniente es asumir que las especies están en pseudo-equilibrio con su ambiente. No existe mucha documentación al respecto de que tan aproximado es un modelo a un estado de equilibrio o cuánto tiempo requiere una especie para alcanzar un nuevo equilibrio luego de un cambio. Es posible que algunas especies no estén en equilibrio con el ambiente y otros factores hagan que sus puntos de observación estén muy dispersos, lo que puede afectar el resultado de un modelo pues no es posible saber a ciencia cierta si hay equilibrio o no.
- **Concepto de nicho:** Los SDM están estrechamente relacionados con el concepto de nicho. El concepto puede ser visto desde el punto de vista de las necesidades ambientales de una especie para sobrevivir en un lugar o desde el punto de vista del rol e impacto de la especie en el lugar que habita. Sin embargo, la definición más aceptada por su cercana relación con los SDM es la de Hutchinson de 1957 que define el nicho de una especie como el conjunto n-dimensional de variables ambientales que permite que la especie pueda subsistir de forma indefinida. Usualmente esta definición es usada para definir el nicho fundamental de una especie, el nicho realizado de una especie hace referencia al subconjunto del nicho fundamental donde las especies han sido realmente observadas.

Los modelos de distribución de poblaciones requieren de dos tipos de datos como insumo para poder realizar las proyecciones: los datos biológicos que describen la distribución conocida de una especie y los datos ambientales que definen las características del hábitat de una especie.

Datos Biológicos

Debido a que los datos de observaciones de una especie están asociados a mapas digitales y las mismas variables ambientales se almacenan como mapas, es que los conceptos de sistemas de información geográficos son aplicables cuando se analiza con detalle cada uno de estos tipos de datos. Usualmente los datos se almacenan como puntos, como polígonos o en formato raster¹, sin embargo, este último es el elegido para proveer de la información necesaria a los algoritmos utilizados en los SDM.

Los datos biológicos de una especie pueden ser solamente de presencia o de presencia/ausencia. Algunos algoritmos solo pueden utilizar los primeros y existen algoritmos que pueden utilizar ambos. Los datos pueden provenir de información de colecciones personales, de datos recopilados por museos, de investigaciones realizadas independientemente o de datos disponibles en la red.

Muchas veces se tiene la duda de cuántas observaciones son necesarias para que los resultados de un modelo sean confiables. Algunos autores hablan de 50 a 100 observaciones, otros hablan de al menos 30 y algunos hablan de 100 a 500 observaciones (Franklin, 2010). Sin embargo, más allá del número de observaciones, lo importante es contar con ocurrencias de la especie que se encuentren bien distribuidas a lo largo del espacio que se sabe ocupa la misma.

Adicionalmente, la ubicación temporal de los datos que se utilicen para un modelo es importante, pues los resultados de un modelo que tome en cuenta información de los últimos meses o par de años, por ejemplo, no serán los mismos que si se incluyen dentro de los datos

¹ El formato raster se refiere a mapas o imágenes formadas por una cuadrícula donde cada celda tiene asignado un número de acuerdo a la condición o atributo de la superficie que representa, por ejemplo la altura, temperatura promedio, promedio de lluvias entre otros (FAO, 2001).

observaciones de hace 10, 20 o 30 años, pues la distribución en amplios rangos de tiempo se pudo ver afectada por cambios en las condiciones ambientales de la especie.

Datos Ambientales

Estos son los datos que proveen de la información necesaria sobre las condiciones ambientales en las que habita o puede habitar una especie. La idea es utilizar esta información de variables que se sabe tienen un efecto en el comportamiento y la fisiología de una especie, y por ende en su distribución. Es recomendable dejar de lado variables que tienen un impacto indirecto sobre la distribución de una especie, pues pueden incluir datos erróneos en los modelos (Franklin, 2010).

Los datos ambientales se pueden caracterizar como datos continuos o datos categóricos. Los datos continuos son aquellos que pueden tomar un valor dentro de una escala, como la temperatura o las precipitaciones. Los datos categóricos son aquellos donde la información está dividida en categorías, como el tipo de suelo o el tipo de uso de la tierra. La información categorizada no se puede utilizar en la gran mayoría de algoritmos utilizados en los SDM, es por ello que la misma debe ser manipulada en muchos casos para generar datos continuos que sí se puedan utilizar como variables en un modelo.

Dentro de los datos usualmente utilizados por un SDM están:

- Mapas del clima: como se mencionó anteriormente, se sabe que las especies reaccionan ante las condiciones climáticas para definir su localización, es por ello que muchas de las variables que definen el clima se utilizan en los SDM. Usualmente se trabaja con los promedios y los extremos superior e inferior de variables como la temperatura, precipitación, radiación, humedad y cobertura nubosa entre otras. Adicionalmente, en muchos casos se realizan proyecciones de estas variables a futuro con el fin de utilizar un modelo para definir la posible distribución de una especie a futuro.
- Mapas digitales del terreno: En algunas oportunidades se utilizan variables topográficas. Dentro de las variables utilizadas usualmente están: la elevación, el

ángulo de inclinación de las pendientes, la dirección de las pendientes, curvatura de las laderas, posición de las laderas. Algunas veces también se incluyen otros elementos topográficos como los cuerpos de agua y los caminos.

- Mapas geológicos y con factores del suelo: aspectos como la composición química y el pH del suelo, o la textura del suelo pueden ser factores que afecten la distribución de una especie.

También es posible tomar en cuenta variables menos comunes, pero igualmente válidas por la información que pueden aportar a los algoritmos de los SDM:

- Vegetación: ciertas categorías de vegetación o ecosistemas pueden ser de preferencia para algunas especies por lo que una variable de este tipo puede ser representativa para el hábitat de las mismas. El uso que haga el ser humano de la tierra también tiene influencia en las distribuciones.
- Disturbios e historia de disturbios: Los disturbios naturales son importantes en la dinámica poblacional de una especie. Los ecosistemas presentan disturbios como inundaciones, incendios, avalanchas o tornados que se caracterizan por su frecuencia y magnitud.
- Distribución de otras especies: Algunas veces se considera que las distribuciones observadas de una especie tienen intrínseco el efecto que otras especies pueden tener sobre la misma; sin embargo, es posible incorporar en los modelos otras especies que tienen una relación de mutualismo, competencia o depredación con la especie que se estudia y con ello incorporar una variable más que afecta el hábitat de una especie en particular.

La calidad de los resultados de un modelo dependerá de la calidad de los resultados que se le proporcionen al algoritmo. Es por ello que hay que seleccionar con cuidado tanto los datos de ocurrencia que se proveen al algoritmo, como los datos de las variables que se quieren utilizar, e inclusive, las variables que se requieren; el objetivo es siempre utilizar información relevante para el estudio que se quiere hacer con el fin de obtener resultados relevantes.

Algoritmos de Modelado

A lo largo de los años se han desarrollado diferentes modelos matemáticos que permiten relacionar los distintos tipos de datos mencionados anteriormente y con ello poder generar los modelos de distribución de las especies.

Se han aplicados diferentes técnicas y metodologías para desarrollar cada uno de ellos. Algunos modelos aplican métodos estadísticos como modelos lineales generalizados o modelos lineales aditivos, otros aplican técnicas de aprendizaje como máxima entropía o inteligencia artificial, algunos incluso hacen una mezcla de algoritmos. Algunos de los algoritmos desarrollados a la fecha de describen brevemente en el Cuadro 1.

Cuadro 1. Modelos de distribución de especies, año de aparición, método utilizado y referencias.

Modelo	Año	Método	Referencia
BIOCLIM	1991	“Sobre climático”, se puede considerar como el pionero.	No disponible
BIOMAPPER	2002	Análisis del factor del nicho ecológico.	http://www2.unil.ch/biomapper/
BIOMOD	2003	Modelo lineal generalizado, modelo aditivo generalizado, árboles de clasificación y regresión, redes neuronales artificiales.	http://www.will.chez-alice.fr/Software.html
SPECIES	2002	Redes neuronales artificiales.	No disponible
GARP	1999	Algoritmos genéticos (con una mezcla de “sobre climático”, modelos generalizados lineales y redes neuronales artificiales).	http://www.nhm.ku.edu/desktopgarp/
MAXENT	2005	Máxima entropía.	http://www.cs.princeton.edu/~schapire/maxent/

Para el propósito del presente trabajo se van a tomar en cuenta los algoritmos GARP y MAXENT, que son los que proporcionan mejores modelos de acuerdo con los resultados de trabajos previos realizados en este campo (Peterson, 2001; Wiens y otros, 2009). A continuación se detalla en qué consiste cada uno de estos algoritmos:

- GARP: es el acrónimo de Genetic Algorithm for Rule-set Production. Fue desarrollado por David Stockwell. Se basa en datos de presencia de una especie y produce una clasificación binaria: presencia o ausencia. El algoritmo genera una población de reglas para clasificar la ausencia o presencia de una especie basado en varios mecanismos:

generación de reglas a partir de los datos obtenidos de las variables ambientales, creación de reglas basadas en rangos usando un principio similar al de BIOCLIM (Franklin, 2010), se generan reglas utilizando una regresión logística donde los coeficientes son estimados para favorecer algunas variables buscando predecir la ausencia o presencia de una especie. Las reglas luego de que son generadas por muestreo, se entrenan dividiendo los datos de ocurrencias en muestras de entrenamiento y muestras testigo. El proceso de descarte y generación de nuevas reglas se va a repetir hasta una cantidad determinada de iteraciones o hasta que se llega a un punto de convergencia en el proceso. Existe una implementación del algoritmo conocida como Desktop GARP.

- MAXENT: es el acrónimo de Maximum Entropy, algoritmo base de este tipo de modelado. Algunas veces se conoce como un modelo de aprendizaje estadístico. La proyección de la distribución de hábitats apropiados para una especie se hace basado en el principio de máxima entropía que establece que la mejor aproximación de una distribución desconocida es aquella con entropía máxima (la más dispersa) sujeta a restricciones conocidas. Las restricciones están establecidas por el valor esperado de la distribución, el cuál es estimado a partir de las observaciones de presencia de una especie. Los modelos de distribución se obtienen por la estimación de la densidad de las variables ambientales condicionada por la presencia de las especies. Al compararlo con otros métodos, MAXENT ha demostrado tener un mejor rendimiento en cuanto a la eficiencia en la generación de modelos y la precisión de los resultados obtenidos. Existe una implementación del algoritmo Maxent, cuyo nombre es el mismo algoritmo.

Al igual que ocurre con los datos, es importante seleccionar el modelo adecuado de acuerdo a los objetivos que se planteen para el uso de los resultados obtenidos a partir de un algoritmo. En general, lo mejor es seleccionar los algoritmos que identifiquen de una forma más acertada la relación entre los datos de observación de una especie y las variables que se proporcionan al modelo. Adicionalmente es importante seleccionar el modelo de acuerdo con el tipo de datos de ocurrencia que se tenga (presencia o ausencia) y el tipo de variables con que se cuenta, pues

no todos los algoritmos trabajan con los mismos tipos (por ejemplo BIOCLIM se enfoca en variables bioclimáticas, GARP o MAXENT son más apropiados para incorporar otro tipo de variables).

Evaluación de los Resultados de un Modelo

Evaluar los resultados obtenidos de un modelo es de gran importancia debido a que si no se tiene una idea de que tan buenos son esos resultados, se podrían cometer errores a la hora de ejecutar cualquier tipo de plan que tenga como sustento dicho modelo. Para llevar a cabo la evaluación existen algunas métricas que muchas veces son provistas por los mismos algoritmos luego de que realizan el procesamiento de los datos para generar los modelos. Entre estas métricas tenemos:

- **Matriz de confusión:** también es conocida como matriz de error o tabla de contingencia. Esta matriz tabula las 4 posibilidades de predicción que se pueden generar en un modelo de distribución: verdadero positivo (el modelo dice que la especie se encuentra en un lugar y efectivamente está ahí), falso positivo (el modelo predice que la especie se encuentra en un lugar, pero está ausente del mismo), falso negativo (el modelo predice ausencia de la especie pero la misma se encuentra en el lugar) y verdadero negativo (el modelo dice que existe ausencia de la especie y efectivamente la hay). Lo ideal es que la cantidad de valores de falso positivo y falso negativo sean lo menor posible, pues esto indica que las predicciones del modelo generado son bastante acertadas.
- **Errores de comisión y omisión:** el error de comisión se define como el error de predecir la presencia de la especie cuando esta se encuentra ausente, mientras que el error de omisión se define como predecir la ausencia de una especie cuando en realidad está presente o como predecir que el hábitat es inhabitable para la especie cuando en realidad si lo es. La idea es que ambos errores sean lo más bajos posibles.
- **La curva ROC y la prueba AUC:** la curva ROC describe la relación entre la proporción de observaciones de presencia correctamente predichas y la proporción de observaciones de ausencia incorrectamente predichas. La prueba AUC hace referencia al área bajo la

curva AUC y su valor oscila entre 0,5 y 1. Si el valor es cercano a 0,5 significa que los resultados del modelo fueron generados de forma aleatoria y por ende no son confiables, si el valor es cercano a 1 significa que el modelo fue capaz de discriminar de una forma más precisa las regiones en las que la especie se encuentra presente o ausente.

Teoría de Dinámica de Poblaciones

Las plagas, a pesar de ser especies que producen un daño a nivel de los cultivos utilizados por el ser humano para alimentarse, forman parte de complejos ecosistemas, por lo que es posible poder aplicarles ciertos principios que han sido planteados a lo largo de la historia para tratar de comprender la dinámica de poblaciones de los seres vivos.

Alan Berryman (Berryman, 2003) nos plantea la discusión que existe sobre si es posible aplicar leyes y principios de otras ciencias a la ecología para poder explicar los fenómenos que se observan en las poblaciones. Así es como a partir de su análisis, plantea cinco principios que deberían ser tomados en cuenta si se desea un modelo que permita explicar los cambios en la población de una especie. Dichos principios se listan a continuación:

- Crecimiento geométrico: el principio de crecimiento geométrico está presente en muchos sistemas tanto naturales como de la vida cotidiana del ser humano, desde el decaimiento de muestras radioactivas hasta el cálculo de intereses compuestos en cuentas de ahorro (Berryman, 2003). Desde este punto de vista puede considerarse que los sistemas ecológicos van a estar acordes con esta ley. Adicionalmente, como se ha podido observar en la vida real, parece describir el comportamiento de una población bajo ciertas condiciones, por lo que se puede considerar el primer principio fundamental en la dinámica de poblaciones. Este principio ha sido discutido por el economista inglés Thomas Robert Malthus en su libro “An Essay on the Principle of Population” publicado en 1798 (Bacaër, 2011).

- Cooperación: el segundo principio se basa en la premisa de que, a nivel ecológico, un individuo va a ver incrementados sus beneficios si la densidad de población aumenta. Ejemplos de estos beneficios son el incremento en las posibilidades de encontrar una pareja, obtener más alimento o escapar de enemigos. Al mismo tiempo, este principio permite establecer que hay un límite en la capacidad de crecimiento, pues todos los organismos tienen una capacidad límite de reproducción (Berryman, 2003).
- Competencia: este principio se basa en otra premisa ecológica que establece que los individuos de una población van a tener problemas para adquirir los recursos que necesitan o se van a volver más vulnerables conforme la población aumenta. Esto va a reducir la capacidad de reproducción o supervivencia de los individuos y por ende va a afectar la población de una especie (Berryman, 2003).
- Interacción entre especies: los tres principios anteriores se refieren a situaciones dentro de una población de una especie, sin embargo, a nivel ecológico ninguna población está aislada y más bien interactúan de una u otra forma con otras especies. Entre las interacciones que se pueden tomar en cuenta para crear un modelo están las relaciones entre especies como el mutualismo y la depredación (Berryman, 2003).
- Factores limitantes: este principio reconoce que el control sobre la dinámica de una población va a cambiar con el tiempo y en el espacio como resultado de una retroalimentación interna para responder a los cambios en el ambiente o la densidad de la población (Berryman, 2003).

Ley de Malthus

Thomas Robert Malthus nació en Londres en 1766. En 1784 empezó sus estudios en la Universidad de Cambridge y obtuvo su diploma en 1791. En 1798 escribe un libro titulado "An Essay on the Principle of Population" el cual se basa en las ideas de que debido a que la población tiende a crecer más rápido que la producción de alimento y que parte de la sociedad está condenada a la miseria, hambruna y epidemias, la población no puede crecer y eso hace que la sociedad no pueda progresar (Bacaër, 2011).

“Podemos, pues, sentar como cierto que cuando no lo impide ningún obstáculo, la población va doblando cada 25 años, creciendo de período en período en una progresión geométrica”

(Malthus, 1846. Pág. 4)

Para establecer la afirmación anterior, Malthus se valió de los datos de población de los Estados Unidos. De dicha frase, el concepto que interesa es que la población crece en forma geométrica, si se representa esta idea en forma de una ecuación diferencial, la expresión que permite representar el cambio de la población en el tiempo es:

$$\frac{dP}{dt} = rP \quad (\text{Pianka, 1999. Pág. 107})$$

Donde P es la población en un momento dado y r es una constante que representa la proporción en la que la población crece a lo largo del tiempo.

La Ecuación de Verhulst

Pierre-François Verhulst nació en Bruselas en 1804. En 1825 obtuvo un PhD en matemáticas en la Universidad de Ghent. En 1835 se convirtió en profesor de la Universidad Libre de Bruselas. En 1838 publica “Note on the Law of Population Growth”, donde expone que, basado en la idea de Malthus sobre el crecimiento geométrico, el tamaño de una población depende de la fertilidad y tamaño del territorio que habita, por lo que con el pasar del tiempo se va a acercar más y más a un punto límite (Bacaër, 2011).

Bajo esta idea es que propone la siguiente ecuación diferencial para representar el cambio de la población en el tiempo:

$$\frac{dP}{dt} = rP \left(1 - \frac{P}{K}\right) \quad (\text{Bacaër, 2011. Pág. 36})$$

Donde r es la constante de crecimiento proporcional de la población, P es la población en un momento dado y la nueva constante K se le conoce como la capacidad de carga de una población y se refiere a la cantidad máxima de individuos que puede tener una población cuando llega a su punto máximo de crecimiento.

Para valores muy pequeños de P :

$$\frac{dP}{dt} \simeq rP \quad (\text{Bacaër, 2011. Pág. 36})$$

Es posible observar que en este planteamiento se produce un crecimiento exponencial hasta que se empieza a llegar a un punto donde dicho crecimiento se detiene y se empieza a acercar a un máximo donde se estabiliza a lo largo del tiempo.

Las Ecuaciones de Lotka-Volterra

Las ecuaciones de Lotka-Volterra fueron propuestas de forma independiente por el italiano Vito Volterra y el estadounidense Alfred Lotka en la década de 1920 (Berryman, 2003). Volterra las planteó al estudiar la variación de las poblaciones de peces en el Mar Adriático. Lotka las usó para tratar de explicar el comportamiento de reacciones químicas donde las concentraciones de los compuestos oscilan en el tiempo.

El modelo de Lotka-Volterra supone un área geográfica cerrada que contiene una cantidad P de presas en un tiempo t . En ausencia de depredación, la población de presas crece de forma exponencial con una tasa de crecimiento r (Padín y otros, 2003).

Para estimar la tasa de mortalidad impuesta por los depredadores sobre las presas, se supone que las presas están distribuidas de forma aleatoria en un área geográfica A y que los depredadores buscan en un área A_s por unidad de tiempo. Si en una fracción σ de los encuentros entre ambos la presa muere, entonces cada depredador consume un promedio de $\sigma (A_s / A) P$ presas por unidad de tiempo. Se define como tasa de ataque $\alpha = \sigma(A_s / A)$ y en consecuencia la tasa de actualización sobre la población que establece de un depredador se define como αP .

Si el área geográfica de interés contiene Q depredadores en un tiempo t , entonces la tasa de mortalidad por depredación experimentada por una presa es αQ . De esta forma, la tasa de

crecimiento de una población de presas explotada por Q depredadores es $r - \alpha Q$. Así llegamos a la primera de las ecuaciones:

$$\frac{dP}{dt} = (r - \alpha Q)P \quad (\text{Padín y otros, 2003. Pág. 2370})$$

En la ecuación anterior se consideró que cada predador consume αP presas por unidad de tiempo. Suponiendo que cada presa consumida de lugar a ϵ nacimientos, entonces la fecundidad de los depredadores en el tiempo t estará dada por $\epsilon \alpha P$. Si la tasa de mortalidad de los depredadores es una constante δ , entonces la dinámica poblacional de los depredadores puede ser descrita por medio de la segunda ecuación:

$$\frac{dQ}{dt} = (\epsilon \alpha P - \delta)Q \quad (\text{Padín y otros, 2003. Pág. 2371})$$

Estas ecuaciones dos ecuaciones forman lo que se conoce como un sistema de ecuaciones acoplado, pues las variaciones de uno de los componentes del sistema van a afectar al otro. Así mismo, es posible inferir que el crecimiento de una población de presas será exponencial en ausencia de depredadores y que el decrecimiento de la población de depredadores será exponencial en ausencia de presas.

Este conjunto de ecuaciones ha llegado a ser generalizado para lograr aplicarlas en otras situaciones de interacción de especies, como relaciones simbióticas de competencia o mutualismo. Así mismo, es posible relacionar n cantidad de especies con las generalizaciones de este modelo.

Antecedentes

Dentro de los antecedentes del presente Trabajo Final de Investigación Aplicada se tienen dos tipos: una serie de investigaciones realizadas para el tema de modelado de distribución de poblaciones de plagas y el proyecto elaborado en el curso PF3342 – Laboratorio de Informática Aplicada a la Ecología y a la Conservación de la Biodiversidad.

Investigación Relacionada

En la bibliografía consultada se encontró que existen algunos trabajos de investigación realizados para el tema del uso de modelos de distribución de poblaciones para el estudio de las poblaciones de plagas.

- En el 2000, Sánchez-Cordero y Martínez-Meyer tomaron como base los datos de observación, recolectados para las colecciones de varios museos, de 17 tipos de roedores que atacan los cultivos del estado de Veracruz, México. Los autores tomaron estos datos, distintos arreglos de variables ambientales y elevación y el algoritmo GARP para llevar a cabo diferentes proyecciones, con lo cual pudieron realizar un estudio de las potenciales distribuciones para cada una de las especies considerada como una plaga (Sánchez-Cordero y Martínez-Meyer, 2000).
- Vera, Rodríguez y otros, en el 2002, desarrollaron un estudio sobre una especie de mosca de la fruta (la *Ceratitis capitata*) donde estudiaron su comportamiento en Europa y el norte de África, para luego estudiar el potencial comportamiento de dicha especie en Argentina y Australia. Para dicho estudio se utilizó un algoritmo comercial conocido como CLIMEX para obtener los modelos de distribución requeridos (Segura, Rodríguez, Vera, Sutherst y Cladera; 2002).
- Un estudio de dos moscas de la fruta, la *Ceratitis capitata* y la *Ceratitis rosa*, fue realizado en el 2008 por DeMeyer, Robertson y otros. En este caso se utilizaron el algoritmo GARP y una técnica conocida como Análisis de Componentes Principales para generar los modelos de distribución de ambas especies. Al final se contrastan los resultados donde la *Ceratitis capitata* es una especie bastante extendida a nivel mundial y se considera que la *Ceratitis rosa*, a pesar de no ser tan extendida, puede

llegar a serlo gracias a las similitudes con la primera (De Meyer, Robertson, Peterson y Mansell; 2008).

- Li, Ma, Hu y otros realizaron un estudio en el 2009 para analizar el posible comportamiento de tres moscas de la fruta (*Ceratitis capitata*, *Ceratitis rosa* y *Ceratitis cosyra*) en China. El análisis de ambas especies se realizó utilizando los algoritmos GARP y MAXENT sobre los datos de observaciones de estas plagas y 19 variables de condiciones ambientales. El estudio fue motivado debido al impacto económico de estas especies en los cultivos y el movimiento de estas especies desde el continente africano a Asia (Li, Ma, Hu, Liu y Zhang; 2009).
- Más recientemente, en el 2011, Jaramillo y otros desarrollaron una investigación para estudiar el comportamiento de la Broca del Café en relación al cambio climático que se está experimentando en el planeta, motivado por el impacto económico que dicha plaga tiene sobre los procesos productivos en los cultivos de café. En el estudio se utilizó el algoritmo CLIMEX para obtener modelos de distribución de esta plaga en África para las condiciones ambientales actuales, y proyecciones a futuro en diferentes escenarios provocados por el cambio climático (Jaramillo, Muchugu, Vega, Davis, Borgemeister y otros; 2011).

De estas investigaciones se puede valorar el hecho de que se ha realizado un esfuerzo por demostrar que este tipo de modelado puede ser utilizado para la proyección de distribución de plagas, sin embargo, no es el objetivo de este TFIA demostrar la validez o la precisión de los modelos generados o de los algoritmos que se utilizaron, si no elaborar una herramienta que permita configurar de una forma sencilla la realización de dichos modelos con el fin de pasar de la teoría y las posibilidades; a una herramienta real y más sencilla de utilizar por los usuarios finales, sin requerir de un extenso conocimiento para entender como se ha atacado el problema de realizar este tipo de modelados.

Proyecto del Curso PF3342 – Laboratorio de Informática Aplicada a la Ecología y a la Conservación de la Biodiversidad

El proyecto del curso consistió en la elaboración de una herramienta en plataforma web. Para el desarrollo de la aplicación se utilizaron herramientas que permitieron tanto el tratamiento de los datos para ser publicados en la aplicación, como las herramientas para la plataforma web propiamente dichas.

Para el tratamiento de los mapas y preparación de los mismos se utilizaron herramientas GIS, en este caso Diva-GIS y Quantum GIS. El primero se seleccionó por su capacidad para realizar transformaciones entre distintos formatos, lo cual ayudó a procesar los datos las variables climáticas. El segundo se seleccionó por su versatilidad para procesar archivos raster y manipularlos para generar mapas con escala de colores.

A nivel de plataforma web, como servidor se eligió Glassfish en su versión 2.1 para hospedar la aplicación que se desarrolló. Para hospedar y manipular los mapas utilizados en la aplicación se escogió GeoServer, versión 2.0.2. Para la generación de modelos se escogió OpenModeller. Con el fin de facilitar la integración de los mapas proporcionados por GeoServer y los resultados generados por OpenModeller es que se escogió OpenLayers en su versión 2.0.

La implementación de la solución se dividió en tres partes:

- Una aplicación web que sirve de cliente al publicar un sitio que permite generar las peticiones la generación de nichos. Es la cara de la aplicación y se conecta a GeoServer para obtener los mapas y a un webservice para enviar las peticiones. Aquí es donde el usuario puede seleccionar las variables con las que desea crear el modelo y los parámetros del algoritmo que se utiliza.
- Un servidor que publica el servicio web que se encarga de generar los nichos que se le soliciten. Sirve de puente entre los datos que provee y solicita el sitio cliente, y OpenModeller. Publica los servicios para: generar un modelo, verificar si un modelo ha sido generado, obtener los resultados de un modelo, obtener las variables disponibles para generar los nichos y obtener los parámetros de un algoritmo. Cada vez que el servicio web recibe una petición para crear un nuevo modelo, se crea un hilo que

atiende dicha petición, con lo que se logra la atención de peticiones múltiples en un momento dado.

- Un componente que interactúa con OpenModeller de forma directa. Es el componente que se encarga de generar los archivos de petición, invocar la aplicación para generar el nicho y procesar el resultado para crear un mapa agradable a la vista para el usuario.

Los resultados finales del proyecto del curso se muestran en la Figura 1, Figura 2 y Figura 3.



Figura 1. Pantalla inicial de la aplicación desarrollada en el curso PF3342.

PARAMETRIZAR MODELO

Mapas Modelar Nicho

Seleccione las capas que desea tomar en cuenta en el modelo

- Todas / Ninguna
- Temperatura Media Anual
- Rango Medio Diurno (Promedio Mensual de (Temperatura Máxima - Temperatura Mínima))
- Isotermalidad (P2/P7) * 100
- Temperatura Estacional (Desviación estándar * 100)
- Temperatura Máxima del Mes mas Calido
- Temperatura Mínima del Mes mas Frio
- Rango de Temperatura Anual (P5-P6)
- Temperatura Media del Trimestre mas Lluvioso
- Temperatura Media del Trimestre mas Seco
- Temperatura Media del Trimestre mas Calido
- Temperatura Media del Trimestre mas Frio
- Precipitacion Anual
- Precipitacion del Mes mas Lluvioso
- Precipitacion del Mes mas Seco
- Precipitacion Estacional (Coeficiente de Variacion)
- Precipitacion del Trimestre mas Lluvioso
- Precipitacion del Trimestre mas Seco
- Precipitacion del Trimestre mas Calido
- Precipitacion del Trimestre mas Frio
- Altura

Seleccione la especie

- Pharomachus mocinno
- Trogon violaceus

Seleccione el algoritmo

GARP

Indique los parametros del algoritmo

Generaciones Máximas: 400

Límite de Convergencia: 0.01

Tamaño de la Población: 50

Remuestreo: 2500

Modelar

Figura 2. Pantalla para ingresar los parámetros de un modelo de distribución en la aplicación desarrollada en el curso PF3342.

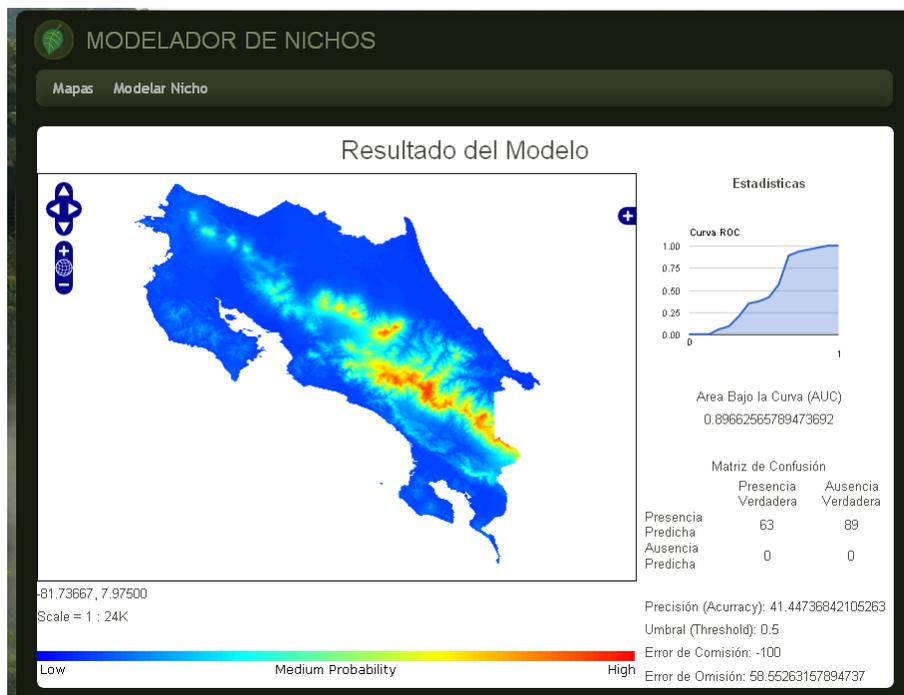


Figura 3. Pantalla que muestra los resultados de un modelo generado por la aplicación desarrollada en el curso PF3342.

Objetivos

Como se aprecia en las distintas investigaciones expuestas en la sección de Antecedentes, el tema del modelado de distribución de especies está cobrando relevancia en la actualidad por la información que pueda aportar para la toma de decisiones. Uno de los campos donde estas herramientas puede llegar a tener mucha importancia es en el campo de la seguridad alimentaria, pues de ella depende el bienestar de una población.

Es así como, dado que instituciones como el IICA y la FAO recomiendan invertir en tecnologías de información para salvaguardar la seguridad alimentaria y que se cuenta con una aplicación básica para el modelado de distribución de especies, se decide crear una aplicación más robusta que permita modelar la distribución de especies consideradas plagas para los cultivos y coadyuve al control y prevención de las mismas. Para lograr esto, se plantean los siguientes objetivos.

General

Desarrollar una aplicación, utilizando herramientas *open source*, que permita modelar la posible distribución de especies consideradas como plagas, que pueda servir como una base para crear una aplicación más robusta en el futuro, para apoyar la optimización de los recursos en el seguimiento y control de plagas.

Específicos

- Elaborar un módulo que permita la administración de los datos (datos de observación y variables ambientales) que se utilizan en la generación de los modelos de distribución.
- Elaborar un módulo para la generación y consulta de los resultados de los modelos de distribución de plagas.
- Implementar un método de proyección de poblaciones que permita modelar el comportamiento poblacional de una especie de una plaga.

Delimitación

El alcance de la aplicación que se propone desarrollar se encuentra delimitado por la siguiente lista de casos de uso:

- Administrar la lista de variables disponibles para generar un modelo.
- Administrar la información de ocurrencia de una especie.
- Configurar la generación de un modelo.
- Consultar los resultados de un modelo.
- Descargar un reporte con los resultados de un modelo.
- Configurar y generar la proyección del comportamiento poblacional.

Elementos como los siguientes quedan fuera de los objetivos y funcionalidad de la aplicación propuesta:

- Comparación en cuanto al rendimiento de los distintos algoritmos de modelado.
- Definición de una estrategia que permita mejorar el rendimiento en cuanto al tiempo de generación de los modelos.
- La incorporación de herramientas de modelado distintas a las provistas por OpenModeller en la versión de la aplicación a desarrollar en este TFIA.

Metodología

Básicamente la metodología fue sencilla. Se tomó como base el proyecto realizado en el laboratorio del curso PF3342 – Laboratorio de Informática Aplicada a la Ecología y a la Conservación de la Biodiversidad, mencionado en la sección de antecedentes. Al mismo se le hicieron las mejoras correspondientes para mejorar su funcionalidad e interfaz de usuario con el fin de facilitar el uso de la aplicación.

Se implementó una base de datos para guardar la información utilizada por la aplicación, con el fin de mejorar los diversos procesos que poseía el proyecto base y dar soporte a las nuevas funcionalidades implementadas en esta nueva versión de la aplicación.

Luego de que la aplicación fue desarrollada, la misma fue sometida a pruebas utilizando información de una especie que se consideran plagas y afectan o pueden afectar cultivos de nuestro país.

Análisis de Resultados

El presente análisis de resultados se divide en dos secciones. La primera presenta los resultados correspondientes a la implementación de la aplicación que fue desarrollada. La segunda hace referencia a los resultados obtenidos al utilizar la aplicación con información real para dos plagas distintas.

Implementación de la Aplicación

Dentro del objetivo principal de este TFIA se establece que se busca utilizar tecnologías y herramientas *open source* para el desarrollo de una aplicación que permita realizar proyecciones para la distribución de plagas. Son varias las herramientas de este tipo que se integraron, ya sea de una u otra forma, para lograr crear la aplicación planteada. Estas herramientas intervienen en las distintas etapas del ciclo de vida de la aplicación. Las herramientas utilizadas, tanto durante el desarrollo como para proporcionar la funcionalidad propiamente dicha de la aplicación, fueron:

- Netbeans, version 7.0.1 como ambiente de desarrollo.
- Glassfish Community, version 3.1.1 como servidor de aplicaciones para hospedar tanto la aplicación como un servidor de mapas.
- GeoServer, versión 2.1.3 como servidor de mapas.
- OpenModeller, version 1.2.0 como herramienta para generar modelos de distribución de especies (también existe una versión conocida como OpenModeller Desktop que proporciona una interfaz gráfica para la herramienta).
- PostgreSQL, version 9.0.4 como motor de base de datos.
- Java Server Faces (JSF), versión 2.0 como framework de desarrollo web.
- PrimeFaces, version 2.2.1 como librería de componentes web.

- PostGIS, versión 1.5.2 como soporte en PostgreSQL para administrar información geográfica.
- OpenLayer, versión 2.11 como soporte para desplegar mapas en la aplicación.
- JasperReports, versión 4.5.0 como *framework* para la generación de reportes.

Donato (2009) menciona que la información para la toma de decisiones en cuanto a los recursos para conservación del medio ambiente tiene que estar disponible de una forma clara y sencilla para las personas que al final toman las mismas. A pesar de que en este caso, el objetivo de la aplicación no es la conservación del medio ambiente, sino la lucha contra las plagas que afectan cultivos, el principio expuesto es igualmente aplicable. Por ello se decide continuar con la línea de desarrollar la aplicación en una plataforma web, de esta forma la información que pueda proporcionar el sistema puede ser accedida en cualquier momento y lugar.

La aplicación sigue una arquitectura de dos niveles principales en una plataforma web, estos dos niveles se interconectan a diferentes piezas de software *open source* con el fin de proporcionar una aplicación completa.

El primer nivel es el de presentación, este se encarga de definir la forma en la que el usuario interactúa con la aplicación, así como la interconexión con las fuentes de información. La implementación se hizo utilizando JSF como *framework* de desarrollo web, JSP para la capa de presentación y Managed Beans para la capa de controladores. A nivel de presentación, se complementa la funcionalidad por defecto de JSF con los componentes de PrimeFaces para proporcionar una interfaz sencilla, pero que a la vez sea poderosa en la experiencia de usuario. Adicionalmente, se incorpora la conexión a otras herramientas para adicionar un valor agregado a la funcionalidad de la aplicación, por ejemplo, se tiene el uso de OpenLayers para desplegar en pantalla los resultados del modelo y poder visualizar su relación con otros mapas provistos por GeoServer.

El segundo nivel es el de lógica y acceso a base de datos, aquí se proporciona los servicios de generación de modelos propiamente dichos, así como el acceso a la información que administra el sistema. Este nivel publica los servicios por medio de servicios web basados en SOAP. Utiliza un componente desarrollado desde el curso PF3342 – Laboratorio de Informática Aplicada a la Ecología y a la Conservación de la Biodiversidad, que sirve de interfaz entre el servicio web respectivo y la herramienta OpenModeller para generar los distintos modelos de distribución de las plagas. Adicionalmente, se conecta a una base de datos en PostgreSQL para guardar, actualizar o recuperar información de los modelos de acuerdo a las acciones que el usuario realice en la aplicación. PostGIS se utilizó como complemento para almacenar la información geográfica de las ocurrencias de una plaga en un formato que pudiera ser consultado y utilizado por GeoServer para proveer mapas de los registros de ocurrencia de una plaga.

Ambos niveles se hospedan en un servidor Glassfish, el cuál fue seleccionado por ser el servidor de referencia para la implementación de la plataforma JEE, que es sobre la cual se apoyan las aplicaciones web desarrolladas en Java. Un diagrama de la arquitectura completa de la aplicación se observa en la Figura 4.

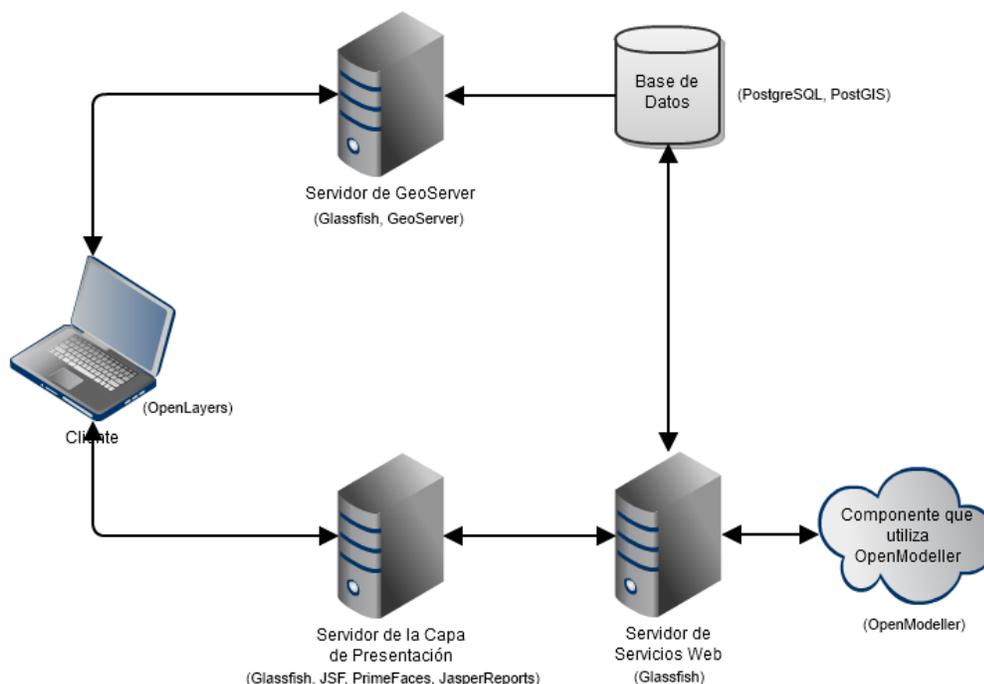


Figura 4. Arquitectura de la aplicación (Entre paréntesis las tecnologías en cada nivel).

Esta separación de niveles se hizo por dos motivos principales:

- La generación de modelos es una tarea que puede llegar a consumir bastantes recursos, tanto de procesamiento como de memoria, debido a esto es posible tener un servidor con mayores recursos para soportar la generación de varios modelos de forma concurrente y que sea distinto del servidor que publica la interfaz de usuario, que requiere de menos recurso.
- Proporcionar la libertad o facilidad a otros desarrolladores de crear una interfaz gráfica en otra plataforma o lenguaje sin tener que reescribir toda la lógica de generación de modelos y administración de información.

Una vez explicada la arquitectura seleccionada para el desarrollo de la aplicación objetivo de este TFIA (a la cuál denominaremos Versión TFIA de aquí en adelante), se detallan y analizan a continuación las mejoras realizadas a la misma con respecto a la aplicación desarrollada en el laboratorio del curso PF3342 – Laboratorio de Informática Aplicada a la Ecología y a la

Conservación de la Biodiversidad (a la cuál denominaremos Versión Laboratorio de aquí en adelante).

La principal de las mejoras implementadas se dio en la forma en la que se maneja la información que utiliza y genera la aplicación. En la Versión TFIA se hace uso de una base de datos relacional para almacenar la información que utiliza el sistema para generar los modelos y donde almacena los resultados de los mismos. En la Versión Laboratorio, toda la información se encontraba almacenada en archivos de texto plano (los datos de ocurrencia de una especie), en formato XML (los resultados de los modelos generados por OpenModeller) o en el código fuente (los algoritmos disponibles y sus parámetros, así como la lista de variables disponibles). En la Versión TFIA, al utilizar un motor de base de datos como PostgreSQL, se obtiene una mayor flexibilidad en cuanto al manejo de las opciones de configuración (los algoritmos y sus parámetros se pueden modificar de forma dinámica y en cualquier momento sin requerir una recompilación del código fuente pues ya no están incluidos en este), manejo de la información utilizada como insumo para los modelos (se puede agregar o quitar variables para la generación de modelos sin requerir una recompilación del código fuente, además de que el usuario puede mantener en la aplicación las variables que considera relevantes) y actualización de la misma (es posible actualizar ciertos atributos de los datos de ocurrencias de una forma más sencilla y dinámica).

Al analizar la teoría sobre los modelos de distribución, un aspecto que sale a relucir es el hecho de la importancia que tiene que la información de las variables y los datos de ocurrencia posean un nivel de calidad que permita confiar en que el modelo resultante también sea de calidad. Es por ese motivo que la aplicación proporciona un mecanismo para el registro de nuevas variables, así como la modificación de las mismas si es requerido, o la eliminación en los casos que se considere que las variables ya no son significativas para modelos futuros. Adicionalmente, posee una herramienta que permite indicar si un dato de ocurrencia debe ser incluido en los modelos o no, esto con el fin de depurar dicha información para descartar en los modelos observaciones de una especie que pueden ser consideradas erróneas o poco significativas; y también es posible agregar datos de una nueva observación directamente desde la aplicación usando el formulario que se muestra en la Figura 7.

Sistema de Proyección de Plagas Inicio Datos Modelo Población



Administrar Variables

Variables para Proyectar

(1 of 1) ← → 1 → → 10

Id	Nombre	Descripción	🔍	🗺️	⬇️
1	Temperatura Media Anual	Temperatura Media Anual	🔍	🗺️	⬇️

(1 of 1) ← → 1 → → 10

AGREGAR VARIABLE

Figura 5. Pantalla para administrar las variables ambientales de los modelos.

Sistema de Proyección de Plagas Inicio Datos Modelo Población



Depurar Datos de Observación

Pharomachus mocinno

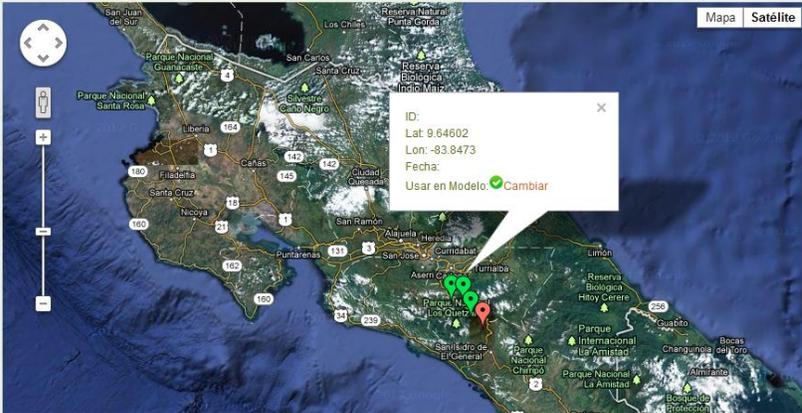


Figura 6. Pantalla para administrar la inclusión de observaciones de una plaga en un modelo.



El formulario, titulado "Agregar Observación", contiene dos campos de entrada de texto. El primer campo, etiquetado "Latitud:", tiene un cursor de texto visible. El segundo campo, etiquetado "Longitud:", está vacío. En la parte inferior del formulario, hay dos botones: "Registrar" y "Cancelar", ambos con un fondo verde.

Figura 7. Formulario para registrar una nueva observación de una especie.

La Versión Laboratorio de la aplicación no posee una forma de acceder a una lista de modelos generados previamente. En este punto, el hecho de que la información de los resultados de los modelos se guarde en un motor de bases de datos relacional proporciona una segunda mejora en la Versión TFIA, pues existe una pantalla que permite consultar la lista de modelos generados previamente y en la cuál es posible filtrar por algoritmo utilizado, especie analizada en el modelo y fecha de generación; adicionalmente se tiene un indicador del estado en el que se encuentra la generación del modelo así como un enlace para consultar los detalles del mismo.

Sistema de Proyección de Plagas Inicio Datos Modelo Población

⚙️ 📄

Resultados de Modelos

(1 of 1) << << 1 >> >> 10 ▾

Nombre Comun	Nombre Científico	Fecha	Algoritmo	Estado	Consultar
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>		
Quetzal	Pharomachus mocinno	2012-03-20 22:59:46.067-05	GARP	✔	🔍
Quetzal	Pharomachus mocinno	2012-03-04 13:45:42.845-06	MAXENT	✔	🔍
Quetzal	Pharomachus mocinno	2012-03-23 18:09:30.928-05	MAXENT	✔	🔍
Quetzal	Pharomachus mocinno	2012-03-23 11:29:20.981-05	MAXENT	✔	🔍

(1 of 1) << << 1 >> >> 10 ▾

Figura 8. Pantalla para consultar la lista de modelos generados por la aplicación.

A nivel de la pantalla que permite configurar los parámetros que se proporcionan al modelo, las mejoras que se aplicaron consisten en mejorar el orden de la información en la pantalla al separa de una forma más clara los parámetros del algoritmo de las variables del modelo (comparar la Figura 2 con la Figura 9). Los parámetros del algoritmo se cargan de forma automática de acuerdo al algoritmo seleccionado. La lista de variables se carga de acuerdo a las variables habilitadas en la aplicación. Un cambio significativo es el permitir al usuario seleccionar la variable que va a funcionar como máscara a la hora de hacer el recorte del raster con el resultado final del modelo; en la Versión Laboratorio esto no era posible, sino que dicha máscara se encontrada definida en el código fuente.

Configurar Modelo

Algoritmo:

Especie:

Parámetros del Algoritmo	
Parámetro	Valor
Number Of Background Points	<input type="text" value="10000"/>
Number Of Iterations	<input type="text" value="500"/>
Terminate Tolerance	<input type="text" value="0.00001"/>
Output Format	<input type="text" value="2"/>

Variables del Modelo				
Máscara Entr.	Entrenamiento	Máscara Proy.	Proyección	Variable
<input checked="" type="radio"/>	<input type="checkbox"/>	<input checked="" type="radio"/>	<input type="checkbox"/>	Temp. Media Anual Afr 2.5 min
<input checked="" type="radio"/>	<input type="checkbox"/>	<input checked="" type="radio"/>	<input type="checkbox"/>	Rango Diurno Medio Afr 2.5 min
<input checked="" type="radio"/>	<input type="checkbox"/>	<input checked="" type="radio"/>	<input type="checkbox"/>	Isotermalidad Afr 2.5 min
<input checked="" type="radio"/>	<input type="checkbox"/>	<input checked="" type="radio"/>	<input type="checkbox"/>	...

Figura 9. Pantalla para parametrizar la generación de un modelo de distribución.

Si se compara la pantalla de configuración de un modelo de la Figura 9 con la pantalla de la aplicación Maxent de la Figura 10 o la pantalla de Desktop GARP de la Figura 11, se aprecia que la primera es más sencilla que las otras dos en cuanto a una disminución sustancial de las opciones de configuración del algoritmo a utilizar. La razón de esta sencillez se debe a que se busca que la Versión TFIA pueda ser utilizada por usuarios finales que no son expertos en cada algoritmo y por lo tanto no se requiera de una configuración sumamente detallada de todas y cada una de las variantes que pueden afectar el modelo en cuanto a su configuración.

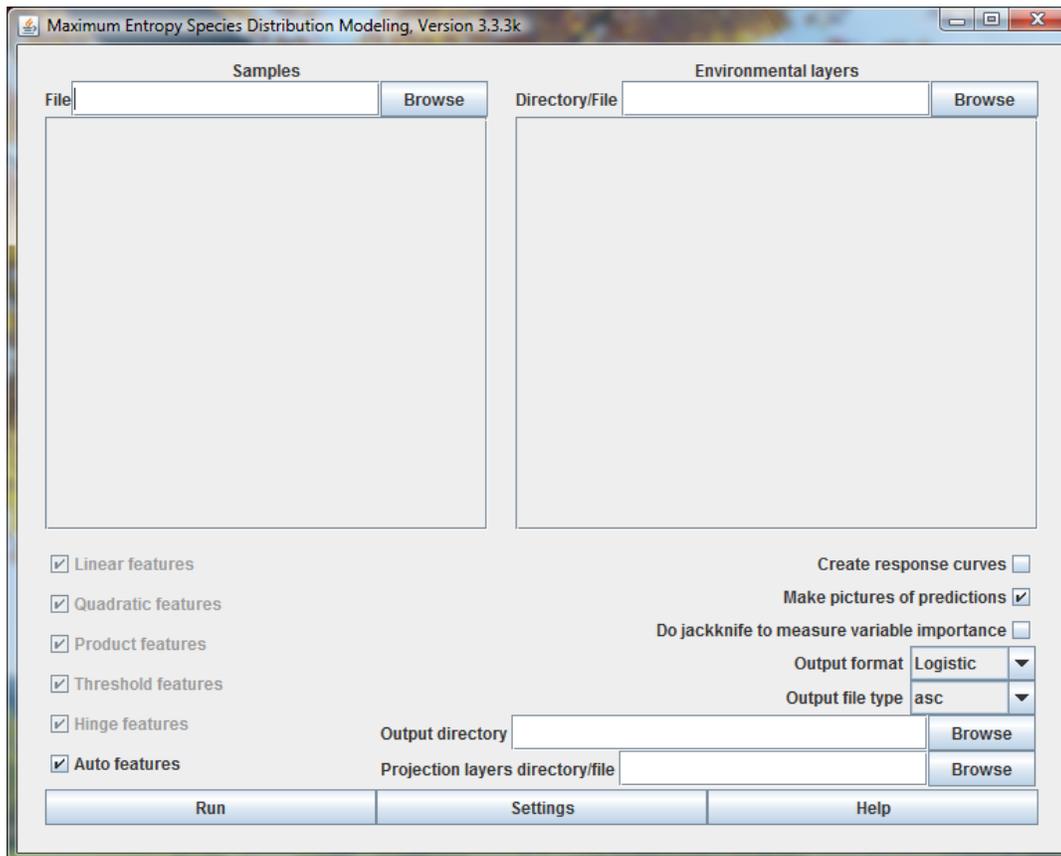


Figura 10. Pantalla principal de la aplicación Maxent

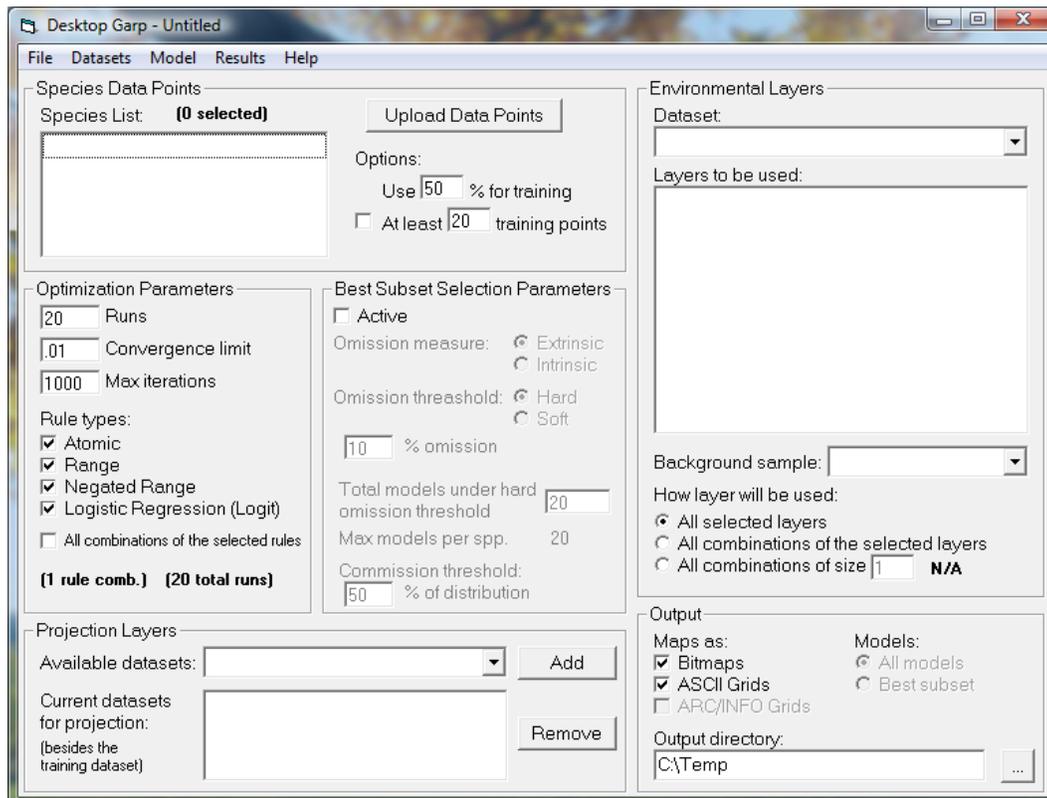


Figura 11. Pantalla principal de la aplicación Desktop GARP.

Con respecto a la pantalla de OpenModeller Desktop que se muestra en la Figura 12, tiene ciertas similitudes en cuanto a la configuración de un algoritmo al permitir la selección de la especie, el algoritmo y los conjuntos de variables que se utilizarán en los modelos. Una diferencia importante es que en la Versión TFIA no se requiere de pasos adicionales para modificar los valores de los parámetros de los algoritmos, contrario a lo que se necesita en OpenModeller Desktop, donde se debe crear una copia del algoritmo con los nuevos valores para cada parámetro. Adicionalmente, al estar la Versión TFIA desarrollada para una plataforma web, permite tener acceso a la aplicación en cualquier lugar y momento; y debido a su arquitectura que separa la presentación de la lógica de generación de modelos, la misma posee alta escalabilidad pues se puede hospedar la lógica de generación en un servidor robusto con grandes recursos de procesamiento y tener uno o más servidores web que permiten el acceso a la capa de presentación y a la funcionalidad.

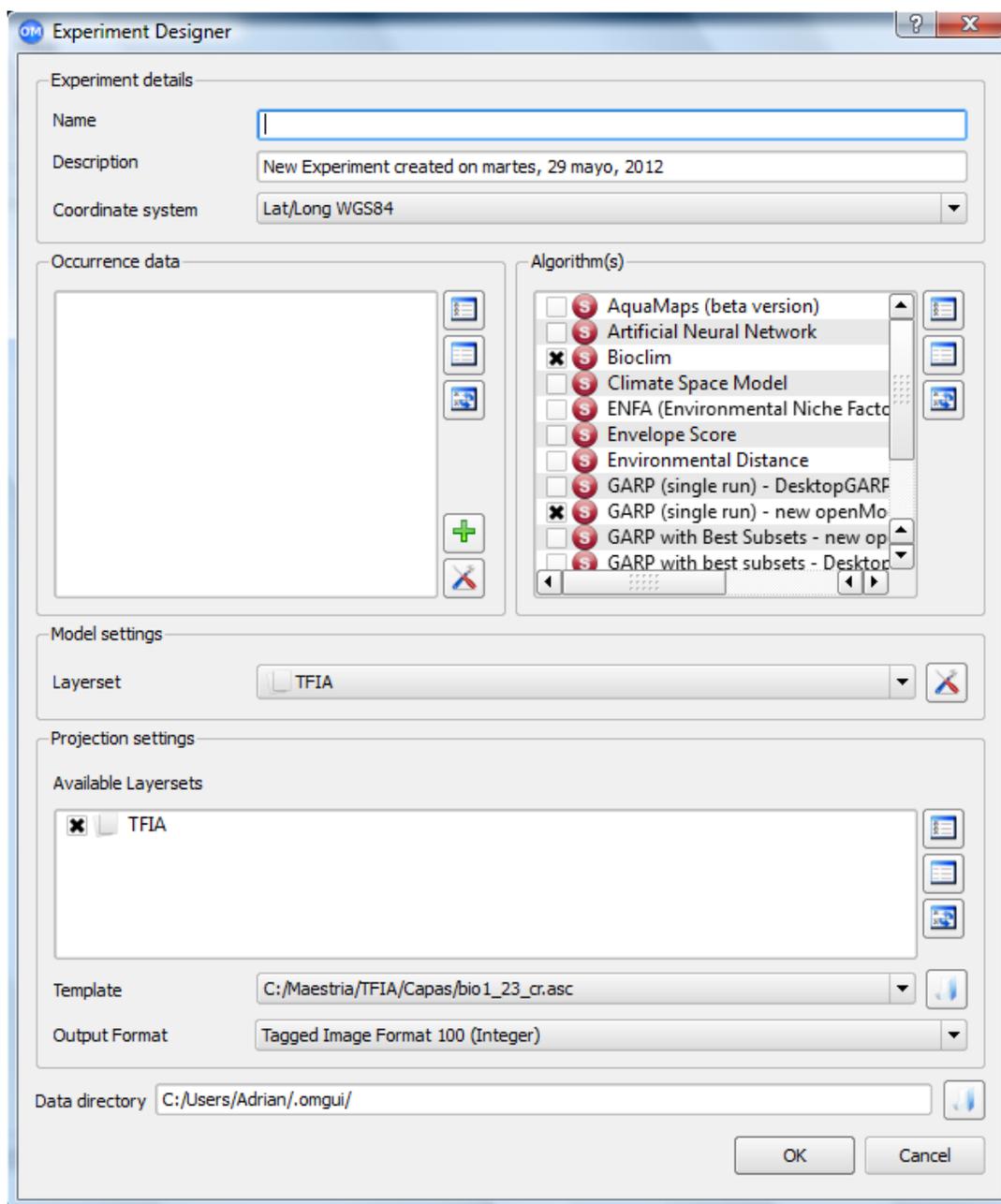


Figura 12. Pantalla de generación de un modelo en OpenModeller Desktop.

Otra diferencia significativa de la Versión TFIA con respecto a las otras tres herramientas, es que las mismas no proporcionan una opción de manipulación de los datos de observación, cosa que si posee la herramienta desarrollada para este TFIA.

Con respecto a la comparación entre la Versión Laboratorio y la Versión TFIA, la pantalla que muestra la información con los datos resultados del modelo de distribución potencial de una plaga en la Versión TFIA es prácticamente la misma que en la Versión Laboratorio, pues se conserva el despliegue del mapa resultado utilizando OpenLayers, es posible sobreponer este resultado una serie de mapas adicionales, como por ejemplo el de las ocurrencias de la especie estudiada en el modelo, se tienen los datos que permiten evaluar la calidad del modelo y un gráfico que muestra la tendencia de la curva ROC, también utilizada para evaluar la validez del modelo. Un mejora es la posibilidad de generar un reporte detallado del modelo y descargarlo en formato PDF, así mismo es posible descargar la imagen del modelo o el archivo raster con el resultado del modelo para poderlo manipular en cualquier herramienta GIS que tenga la capacidad de leer raster en formato ASCII.

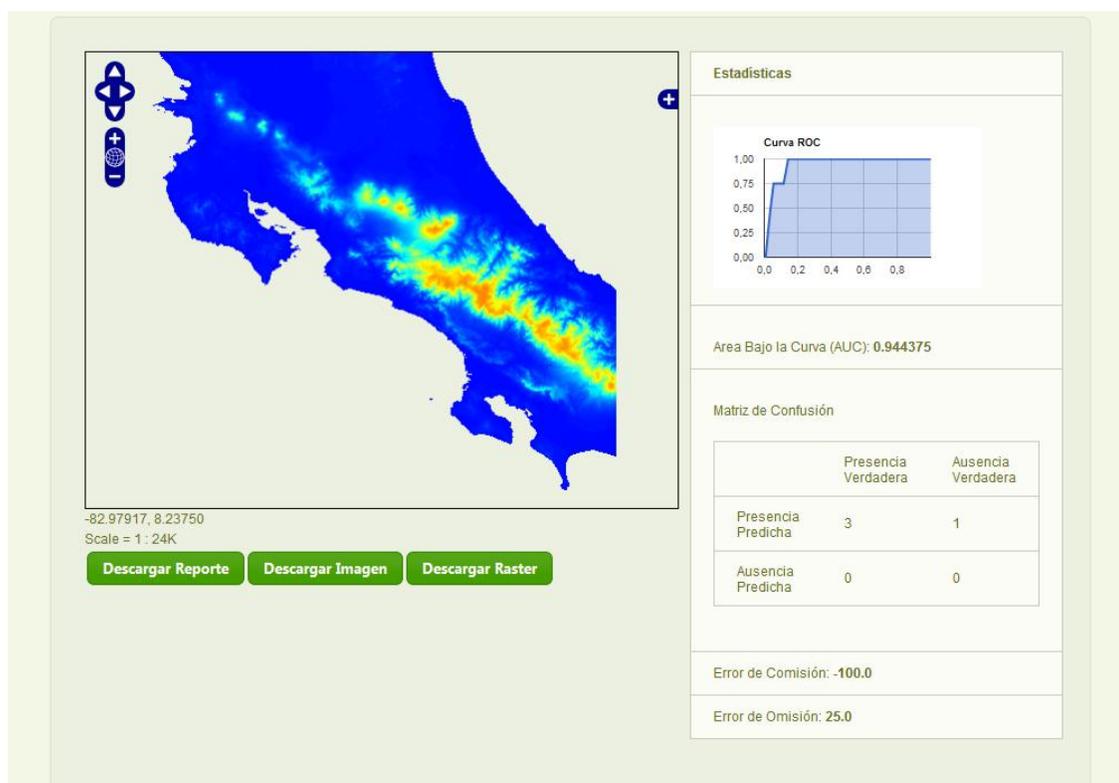


Figura 13. Pantalla para visualizar el resultado de un modelo de distribución.

Finalmente, un aporte completamente nuevo de la Versión TFIA es la herramienta para predecir la distribución de una especie con base en ciertas características de la misma y algunas

variables externas que pueden afectarla, como lo son la presencia de depredadores o el uso de plaguicidas. El modelo utilizado se basa en los distintos modelos de crecimiento poblacional que se exponen en la sección Teoría de Dinámica de Poblaciones del Marco Teórico. El modelo básico utilizado es el de Verhulst, para el caso de la influencia de un depredador sobre la plaga se utiliza una adaptación del modelo de Verhulst incluyendo las ecuaciones de Lotka-Volterra. Finalmente, si se quiere incluir el efecto de un pesticida, se hace un ajuste sobre cualquiera de las dos ecuaciones resultantes anteriores para quitar de la proyección los individuos que se verían afectados por los efectos del pesticida.

Sistema de Proyección de Plagas Inicio Datos Modelo Población

Configurar Modelo de Población

Datos de la Población

Tamaño Inicial de la Población	<input type="text"/>
Factor de Crecimiento	<input type="text"/>
Equilibrio Poblacional	<input type="text"/>

+ Efecto de Depredador

+ Efecto de Pesticida

CALCULAR POBLACIÓN

Figura 14. Pantalla para parametrizar un modelo de comportamiento poblacional.

El Cuadro 2 presenta un resumen de la comparación entre la Versión Laboratorio y la Versión TFIA, así como la comparación con otras herramientas para la generación de modelos de distribución.

Cuadro 2. Tabla comparativa entre herramientas.

Criterio	Maxent	Desktop GARP	OpenModeller Desktop	Versión Laboratorio	Versión TFIA
Requiere de un conocimiento profundo del algoritmo utilizado para configurar el modelo (sencillez de configuración)	Sí	Sí	No	No	No
Posee funcionalidad orientada a la realización de tareas colaborativas	No	No	No	Parcial*	Sí
Permite manipular los datos de ocurrencia de una especie	No	No	No	No	Sí
Permite relacionar resultados de modelos con otros mapas	No	No	No	Sí	Sí
Posee modelos básicos de dinámica poblacional	No	No	No	No	Sí

*La aplicación está en plataforma web, pero no cuenta con una consulta para ver modelos generados anteriormente

Pruebas de la Aplicación con Datos Reales de Plagas

Como ha sido mencionado a lo largo de este documento, la idea de este proyecto es proporcionar una herramienta para modelar la posibilidad de que una plaga se asiente o se disperse en el país, si éste ofrece las condiciones ideales para el mismo. Es por eso que es importante hacer pruebas a la aplicación que permitan demostrar su utilidad.

A la hora de definir las pruebas a realizar se pensó en una especie que no se encuentre en el país o que la misma se encuentre localizada de forma restringida. Al revisar las investigaciones previas descritas en la sección de Antecedentes de este documento y la lista de plagas cuarentenarias definida por el Servicio Fitosanitario del Estado (SFE), se encuentra la *Ceratitis capitata*, una especie de la mosca de la fruta que es originaria de África, pero que se ha extendido alrededor del mundo y en nuestro país se ha localizado en La Cruz, Guanacaste.

Las razones para escoger esta plaga son que cumplen con el requisito de presencia/ausencia en el país descrito anteriormente y que gracias a los estudios previos de esta especie, se cuenta con información detallada para realizar distintos modelos de prueba.

Para realizar las pruebas se obtuvieron datos de ocurrencia para ambas especies en el sitio web data.gbif.org y las variables ambientales se obtuvieron del sitio web www.worldclim.org. Para los datos de ocurrencia se filtraron los resultados de la búsqueda en el sitio web mencionado

con el fin de recolectar datos que tuvieran información sobre la latitud y longitud de la observación de la especie en el campo o del sitio donde se recolectó un espécimen. Con respecto a los datos de variables ambientales, la resolución de las mismas es de 2,5 minutos de coordenadas geográficas, de acuerdo al sistema de referencia WGS84.

Las pruebas se realizaron aplicando los algoritmos MAXENT y GARP que, como se mencionó en el Marco Teórico de este TFIA, son los que presentan los mejores resultados en la actualidad a la hora de generar los modelos de distribuciones potenciales de especies.

Los datos ocurrencia se limpiaron para utilizar solamente los puntos de ocurrencia observados en el continente africano y europeo, sin tomar en cuenta los puntos observados en las islas o el océano. Esta depuración de datos se hizo debido a que en el sitio web de WorldClim, en la metodología de muestran mapas donde se observa que la cobertura de los datos de variables climáticas parecen ser mejores en África continental que en las pequeñas islas a su alrededor.

Se generaron 4 modelos en las pruebas. El primer juego de modelos utilizó las 19 variables ambientales en los algoritmos GARP y MAXENT. El segundo juego de modelos solo utilizó 8 variables de las 19; esta selección se hizo tomando como base el estudio de De Meyer, Robertson y otros en el 2008. No está dentro de los objetivos de este TFIA el análisis de cómo seleccionar las variables para generar los modelos, es por ello que se utilizan estudios previos como base para sustentar este tipo de decisiones.

La Figura 15 muestra los resultados de la primera prueba, utilizando las 19 variables ambientales con el algoritmo MAXENT. El valor del área bajo la curva (AUC) es de 0,94 y el error de omisión de 37%.

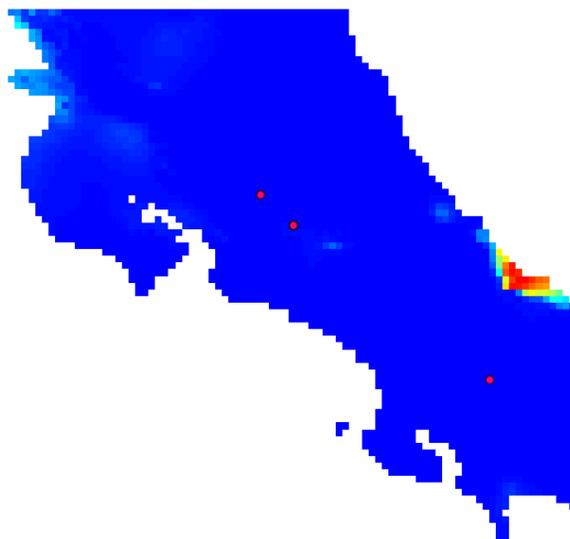


Figura 15. Resultado usando el algoritmo MAXENT con 19 variables climáticas.

Utilizando las mismas 19 variables, pero esta vez con el algoritmo GARP, se obtiene el resultado que se aprecia en la Figura 16. Esta vez el valor de AUC es de 0,87 y el error de omisión es de un 24%.

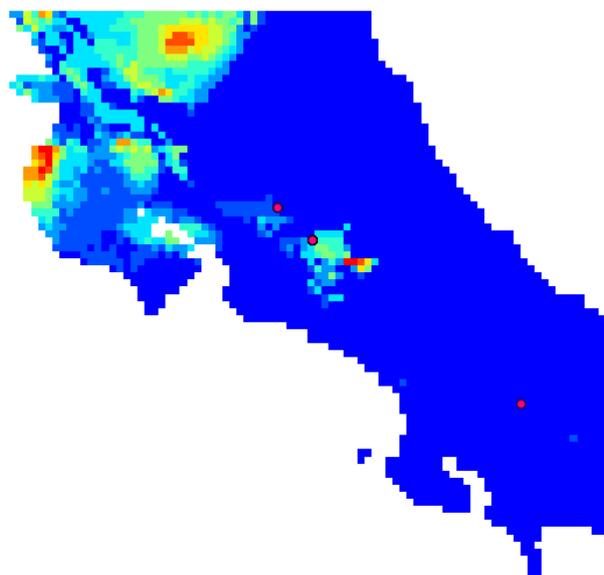


Figura 16. Resultado usando el algoritmo GARP con 19 variables climáticas.

Si se hace la prueba, pero esta vez usando solamente 8 variables, aplicando el algoritmo MAXENT, se obtiene un valor AUC de 0,92 y un error de omisión del 39%. El resultado se puede observar en la Figura 17.

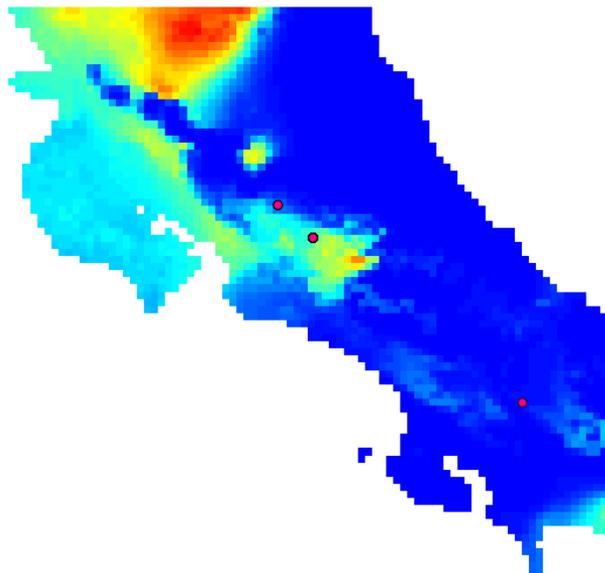


Figura 17. Resultado usando el algoritmo MAXENT con 8 variables climáticas.

Finalmente, el modelo obtenido para las 8 variables con GARP arroja un valor AUC de 0,83 con un error de omisión de solamente 8%. La Figura 18 muestra ese resultado.

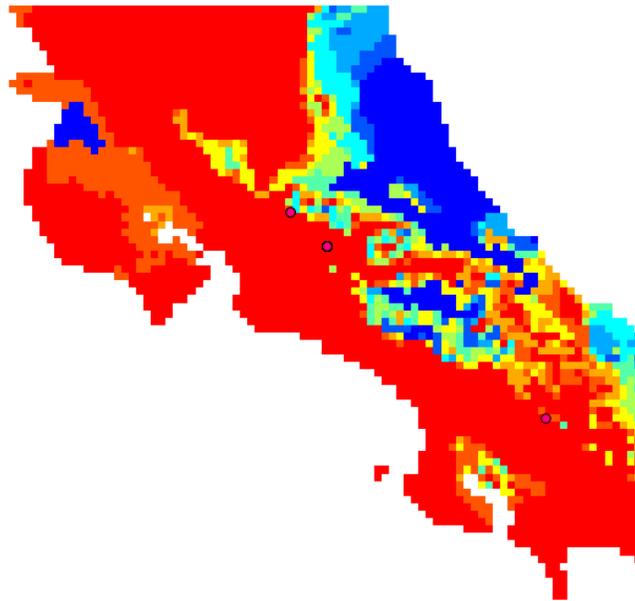


Figura 18. Resultado usando el algoritmo GARP con 8 variables climáticas.

De acuerdo con los datos de ocurrencia obtenidos para la *Ceratitidis capitata*, se tienen registros de ocurrencia en el país para esta especie. Los mismos no fueron utilizados a la hora de generar los modelos, pues se buscó comparar los resultados de los modelos con estas observaciones para comprobar la precisión de los modelos. En las figuras con los resultados de los modelos, dichas ocurrencias se muestran como círculos de color fusia.

Al analizar los dos primeros modelos, el valor de AUC sugiere que el nivel de aleatoriedad a la hora de generar los modelos es menor usando el algoritmo MAXENT que usando el algoritmo GARP, sin embargo, si se utiliza la métrica del error de omisión, el algoritmo GARP obtiene un mejor resultado. A nivel de las observaciones, el resultado del modelo usando GARP se aproximó de forma leve a dichos puntos de ocurrencia.

Con respecto a los resultados de los otros dos modelos, nuevamente el valor AUC se mantuvo relativamente alto lo que indica un nivel de aleatoriedad aceptablemente bajo a la hora de generar los resultados. Sin embargo, existe una diferencia muy marcada en relación a los valores de los errores de omisión, pues para el caso de MAXENT fue de 39%, pero para GARP fue de solamente un 8%. Al tomar en cuenta ambas métricas, se puede sugerir que el modelo generado por GARP es el mejor de los dos, dado que el valor AUC es relativamente alto y el

valor del error de omisión es muy bajo; y según el Marco Teórico de este TFIA, ambas condiciones definen un buen modelo. Adicionalmente, es posible apreciar como las observaciones de *Ceratitis capitata* se ubican en áreas de alta probabilidad de acuerdo con el modelo, situación que no ocurren en el modelo generado con MAXENT.

Ahora, si se comparan los 4 modelos, es posible observar que la reducción de variables de un conjunto de pruebas a otra si tuvo un efecto considerable en algunos resultados. Para el caso de los modelos con MAXENT, el efecto fue mínimo, pues los valores de AUC y del error de omisión presentaron una variación ínfima y negativa, pues el valor AUC se encareció en 0.02 unidades y el error de omisión aumentó en un 2%. En contraste, dicha reducción tuvo un efecto muy positivo en los resultados finales, pues aunque el valor de AUC decreció en 0.04 unidades, el error de omisión pasó de 24% a un 8%, una reducción sustancial e importante. De estos resultados se puede sospechar que el algoritmo MAXENT tal vez no es el mejor para el estudio de esta especie en específico y se puede confirmar una variación de las variables utilizadas en un modelo tienen un efecto en el resultado final según se estudió en el Marco Teórico de este TFIA.

Es importante resaltar nuevamente que el tener datos de ocurrencias de la *Ceratitis capitata* en Costa Rica ayuda a realizar una evaluación final de los modelos para determinar que tan buenos pueden ser o no para tomar decisiones.

Pruebas de la Proyección de Crecimiento Poblacional

En relación a los modelos de crecimiento poblacional, en el Marco Teórico se desarrolla una explicación sobre diferentes modelos de crecimiento poblacional que han sido planteados a lo largo de la historia, algunos son modificaciones o mejoras de otros y funcionan para una sola población. Otros son más complejos y relacionan dos o más poblaciones, buscando modelar la interacción entre ellas.

En la aplicación Versión TFIA, se desarrolló una interfaz que permite, con unos pocos parámetros, trabajar con un modelo sencillo el de Verhulst. Es posible proporcionar más parámetros para aplicar un modelo como el definido por Lotka-Volterra. Finalmente es posible

proporcionar parámetros adicionales para incluir en estos modelos el efecto que un plaguicida puede llegar a tener sobre la población de una plaga.

La Figura 19 permite observar el resultado para un modelo de crecimiento poblacional que aplica el modelo de Lotka-Volterra. La Figura 20 muestra el resultado de un modelo de crecimiento poblacional utilizando el modelo de Verhulst. Se aclara que los resultados que se presentan en esta figuras no tienen relación con el depredador de la *Ceratitis capitata*, sino que buscan mostrar el tipo de resultados que se pueden visualizar con la herramienta.

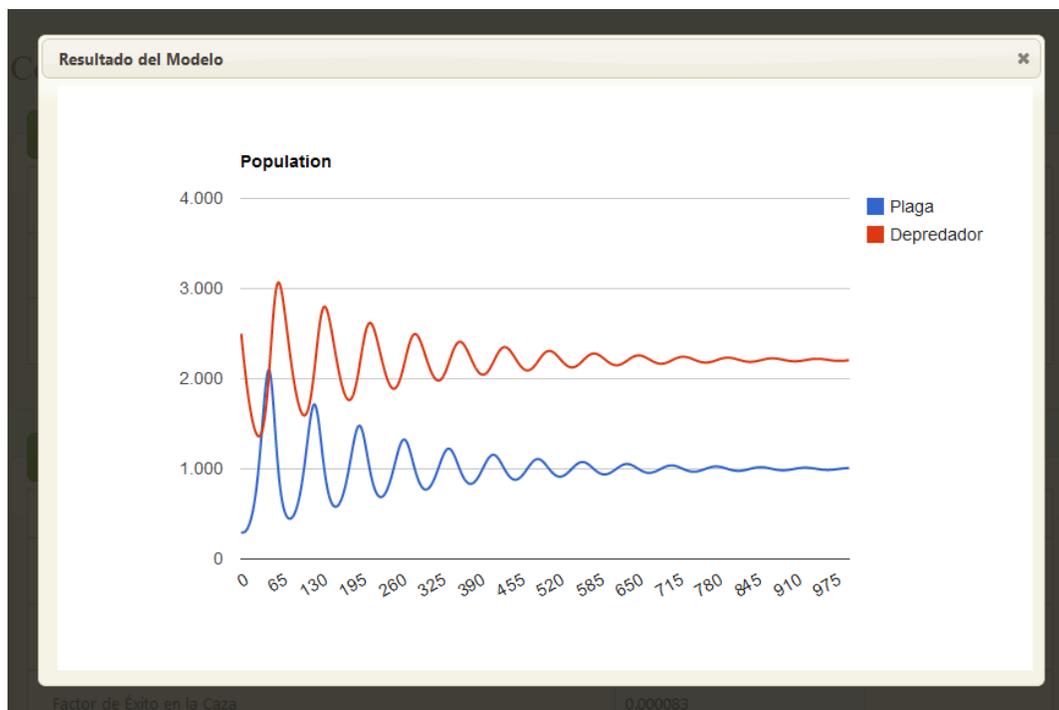


Figura 19. Resultado de un modelo de crecimiento poblacional utilizando el modelo de Lotka-Volterra.

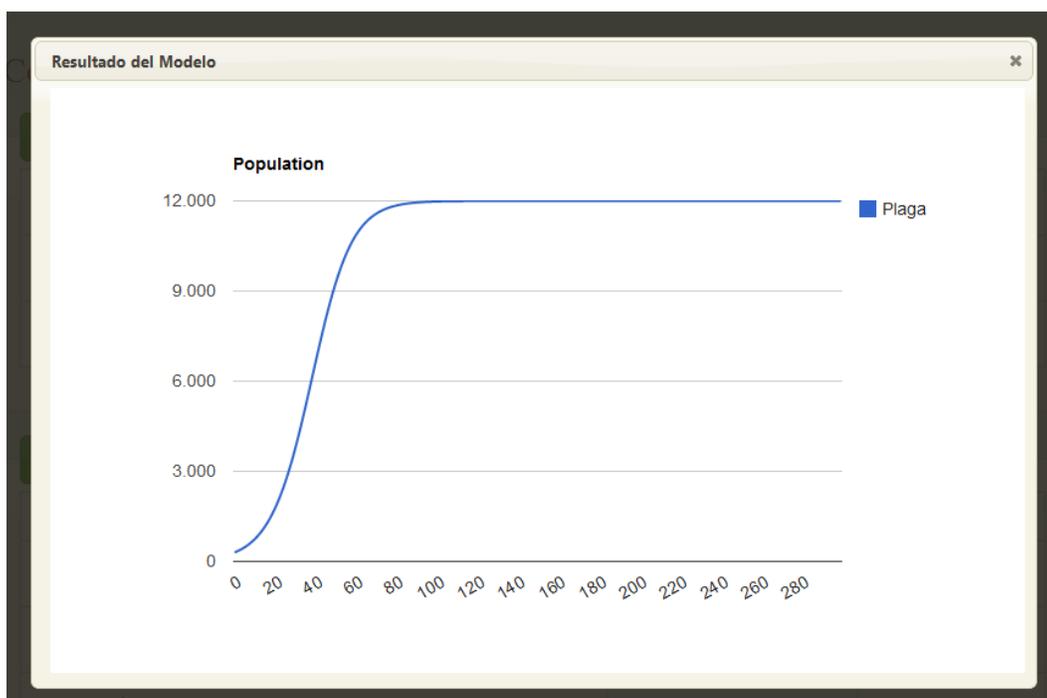


Figura 20. Resultado de un modelo de crecimiento poblacional utilizando el modelo de Verhulst.

Luego de la implementación y pruebas realizadas en estos modelos, es posible establecer dos hechos interesantes:

- Para que los resultados que genera el modelo sean lo más útiles y representativos posibles, es necesario proporcionar al modelo información que sólo es posible determinar con la observación del comportamiento de una especie. Entre ellos se tienen los valores de porcentaje de crecimiento y el valor de carga máxima, los cuales solo se pueden llegar a estimar si se cuenta con información sobre el desarrollo de una población de una plaga a lo largo del tiempo.
- Los modelos de crecimiento poblacional que se presentan en el Marco Teórico y que fueron implementados en la Versión TFIA de la aplicación, describen modelos generales que pueden proporcionar una tendencia, pero es posible establecer a partir de trabajos como “Efecto de la variabilidad climática sobre la fluctuación poblacional de la rata cañera (*Sigmodon hispidus*) en Cañas, Guanacaste” y “Relación entre la temperatura superficial del mar TSM y anomalías de temperatura del aire en el Pacífico

norte de Costa Rica durante años ENOS como posible predictor de potencialidad de plaga de langosta.”, que cada población de una plaga se ve afectada por una serie de factores que alteran los resultados de estos modelos generales. En el caso de la Rata de Campo, de acuerdo con Retana y otros (Retana, Solera, Solano y Álvarez; 1999), se observa el impacto de los períodos de lluvia y sequía en la fluctuación de las poblaciones de esta especie; un factor como este no está incluido en los modelos de la Versión TFIA, pero sería interesante incluirlo en un trabajo futuro para mejorar las estimaciones de crecimiento poblacional generadas por la aplicación desarrollada hasta el momento.

Conclusiones

- Es posible reunir varias herramientas *open source*, de diferentes áreas, para crear una aplicación que permita generar modelos de potenciales distribuciones de poblaciones de plagas.
- La decisión tomada para desarrollar la aplicación, al utilizar una arquitectura en plataforma web que separa la presentación de la lógica de generación de los modelos, le proporciona una escalabilidad y una serie de ventajas de acceso a la herramienta, que no tendría si se hubiera utilizado una estrategia de desarrollar una aplicación de escritorio.
- El desarrollo de una aplicación que permita manipular la información necesaria para generar modelos de distribución de plagas (como las variables y los datos de ocurrencia de la especie) y que permita la visualización de los resultados, proporciona al usuario final una herramienta de fácil manejo que le ayude a trabajar con los distintos algoritmos sin necesidad de conocer a profundidad toda la teoría asociada a la generación de modelos.
- La creación de modelos de crecimiento poblacional es un área que requiere bastante estudio si se pretende lograr un modelo que puede representar lo que ocurre en la realidad con la población de una especie, pues los modelos teóricos se pueden ajustar y mejorar si se toman en condiciones adicionales características de cada especie.

Bibliografía

- Bacaër, N. 2011. A Short History of Mathematical Population Dynamics. Pág 31-38. Londres, Reino Unido: Springer-Verlag.
- Berryman, A. 2003. On principles, laws and theory in population ecology. OIKOS 103: 3.
- Brathwaite, C. 2009. Seguridad alimentaria en las Américas: se necesita un nuevo modelo de desarrollo. San José, Costa Rica: IICA.
- CNN. 2011. ¿Quién alimenta a 7,000 millones? Disponible en <http://www.cnnexpansion.com/economia/2011/10/31/quien-alimenta-a-7000-millones>.
- De Meyer, M., Robertson, M. P., Peterson, A. T., Mansell, M. W. 2008. Ecological niches and potential geographical distributions of Mediterranean fruit fly (*Ceratitis capitata*) and Natal fruit fly (*Ceratitis rosa*). Journal of Biogeography 35: 270-281.
- Donato, K., Amaral, S., Vieira, A., Câmara, G., Carvalho, A., Satiko, L. 2009. Making Species Distribution Models Available on the Web for Reuse in Biodiversity Experiments: *Euterpe edulis* Case Study.
- FAO. 1995. Glosario de términos fitosanitarios. Disponible en <http://www.fao.org/docrep/W3587E/w3587e03.htm>.
- FAO. 2001. El Uso de Computadoras, Programas e Instrumentos Electrónicos en la Planificación y Seguimiento de Planes de Manejo del Bosque Húmedo Tropical - Un Caso en Costa Rica. Disponible en <http://www.fao.org/docrep/003/x8763s/x8763s00.htm>.
- FAO. 2008. Plagas y Enfermedades Transfronterizas Relacionadas con el Cambio Climático. Pág 1-13. Conferencia de Alto Nivel sobre la Seguridad Alimentaria Mundial: Los Desafíos del Cambio Climático y la Bioenergía, Roma.
- Franklin, J. 2010. Mapping Species Distributions: Spatial inference and distribution. Cambridge, Reino Unido: Cambridge University Press.

- GBIF. Species: *Ceratitis capitata* (Wiedemann, 1824). Disponible en <http://data.gbif.org/species/1626096/>.
- Guisan, A y Thuiller, W. 2005. Predicting Species Distribution: offering more than simple habitats model. *Ecology Letters* 8: 993-1009.
- IICA. 2012. Preguntas Frecuentes Sobre Seguridad Alimentaria. Disponible en <http://www.iica.int/Esp/conocimiento/actualidad/Documentos%20Seguridad%20Alimentaria/Preguntas%20frecuentes%20sobre%20la%20seguridad%20alimentaria.pdf>
- Jaramillo, J., Muchugu, E., Vega, F. E., Davis, A., Borgemeister, C. y otros. 2011. Some Like It Hot: The Influence and Implications of Climate Change on Coffee Berry Borer (*Hypothenemus hampei*) and Coffee Production in East Africa. *PLoS ONE* 6(9): e24528. doi:10.1371/journal.pone.0024528.
- Li, B., Ma, J., Hu, X., Liu, H., Zhang, R. 2009. Potential Geographical Distributions of the Fruit Flies *Ceratitis capitata*, *Ceratitis cosyra*, and *Ceratitis rosa* in China. *Journal of Economic Entomology* 102: 1781-1790.
- Malthus, T. 1846. *Ensayo Sobre el Principio de la Población*. Pág 4-6, traducido por Noguera J y Miquel J. Madrid, España.
- Padín, M. y otros. 2003. Visualización de Modelos Poblacionales. *Mecánica Computacional* 22: 2369-2382.
- Pearson, R.G. 2007. *Species' Distribution Modeling for Conservation Educators and Practitioners*. Synthesis. American Museum of Natural History. Disponible en <http://ncep.amnh.org>.
- Peterson, A. T. y Vieglais, D. 2001. Predicting Species Invasions Using Ecological Niche Modeling: New Approaches from Bioinformatics Attack and Pressing Problem. *BioScience* 51: 363-371.
- Peterson, A. T. 2003. Predicting the Geography of Species' Invasion Via Ecological Niche Modeling. *The Quarterly Review of Biology* 78: 419-433.

- Pianka, E. 1999. *Evolutionary Ecology*. Pág 105-107. Pearson International.
- Retana, J. 2003. Relación entre la temperatura superficial del mar TSM y anomalías de temperatura del aire en el Pacífico norte de Costa Rica durante años ENOS como posible predictor de potencialidad de plaga de langosta. *Tópicos Meteorológicos y Oceanográficos* 10: 1.
- Retana, J., Solera, M., Solano, J. y Álvarez, H. 1999. Efecto de la variabilidad climática sobre la fluctuación poblacional de la rata cañera (*Sigmodon hispidus*) en Cañas, Guanacaste. Ministerio de Ambiente y Energía. Instituto Meteorológico Nacional. Gestión de Desarrollo. San José, Costa Rica.
- Sánchez-Cordero, V y Martínez-Meyer, E. 2000. Museum specimen data predict crop damage by tropical rodents. *PNAS* 97: 7074-7077.
- Segura, D. F., Rodríguez, R., Vera, M. T., Sutherst, R. W., Cladera, J. L. 2002. Potential Geographical Distribution of the Mediterranean Fruit Fly, *Ceratitis capitata* (Diptera: Tephritidae), with Emphasis on Argentina and Australia. *Environmental entomology* 31: 1009-1022.
- Universidad Nacional de Colombia. 2013. Concepto de Plaga. Dirección Nacional de Servicios Académicos Virtuales. Disponible en <http://www.virtual.unal.edu.co/cursos/agronomia/2006631/lecciones/cap01/lec01a.html>.
- Wiens, J y otros. 2009. Niches, models and climate change: Assessing the assumptions and uncertainties. *PNAS* 106: 19729-19736.
- WorldClim. Data for current conditions (~1950-2000). Disponible en <http://www.worldclim.org/current>.