

## Desarrollo de Herramientas de Exploración de Datos Geospaciales Aplicadas a una Base de Datos Fitosanitaria

Fernando R. A. BORDIGNON <sup>1,5</sup>, Walter F. SIONE <sup>1,4</sup>, Pablo G. ACEÑOLAZA <sup>1,3</sup>, Lisandra P. ZAMBONI <sup>1</sup>, Guillermo HEIT <sup>2</sup>, Javier ORZUZA <sup>2</sup>, Pablo HORAK <sup>2</sup> y Pablo CORTESE <sup>2</sup>.

1. Centro Regional de Geomática - Facultad de Ciencia y Tecnología. Universidad Autónoma de Entre Ríos. CeReGeo-UADER. España y Matteri S/N. CP:3105. Diamante. Entre Ríos. Argentina. [sione@ceregeo.org.ar](mailto:sione@ceregeo.org.ar)

2. Dirección de Vigilancia y Monitoreo. SENASA. Av. Paseo Colón 315. 4º Piso Ofic. 15. Ciudad Autónoma de Buenos Aires. C1063ACD. Bs. As. Argentina. [gheit@senasa.gov.ar](mailto:gheit@senasa.gov.ar)

3. Centro de Investigaciones Científicas y de Transferencias de Tecnologías a la Producción CICyTTP-CONICET. [acenolaza@gmail.com](mailto:acenolaza@gmail.com)

4. PRODITEL – Cs.Bs. UNLu.

5. Laboratorio de Redes de Datos Cs. Bs. UNLu. [bordi@unlu.edu.ar](mailto:bordi@unlu.edu.ar)

### Resumen

El presente trabajo corresponde al “Estudio sobre el desarrollo de Sistemas de Información asociados a aplicaciones SIG con base de datos fitosanitaria” desarrollado de forma conjunta por el Centro Regional de Geomática de la FCyT- UADER, y la DNPV-SINAVIMO dependiente del SENASA. El objetivo específico es generar modelos cartográficos de dispersión distribución de plagas a través de la implementación de aplicaciones SIG funcionando bajo entornos web.

Como resultado, se ha construido un Módulo de Exploración de Datos Geospaciales (MEDG) denominado “Sistema de Información Geográfica Fitosanitaria” (SIGAFIT), que actúa como herramienta central del sistema de información fitosanitaria de riesgo bioclimático, dado que permite que un usuario técnico pueda armar realizar consultas espaciales complejas a través de un formulario simple, sobre una base de datos climáticos, fitogeográficos y/o administrativos y derivados de

teledetección, para el desarrollo de modelos de distribución espacial de especies. El mismo puede ser consultado a través de la página web de Sinavimo o desde el sitio <http://sigafit.senasa.gov.ar>, y permite obtener, en formato mapa o tabular, información sobre aquellas áreas del territorio nacional que cumplen con las condiciones expresadas en la mencionada consulta. En cuanto a la aplicación propiamente dicha, cabe destacar que se encuentra desarrollada con software libre y puede ser modificada agregando nuevas funcionalidades.

**Palabras clave:** Base de datos geospacial. GIS. Vigilancia fitosanitaria, modelo de distribución de especies.

### Contexto

Esta presentación corresponde al proyecto “Estudio sobre el desarrollo de Sistemas de Información asociados a aplicaciones SIG con base de datos fitosanitaria” desarrollado por el Centro Regional de Geomática de la Facultad de Ciencia y

Tecnología, perteneciente a la Universidad Autónoma de Entre Ríos en el marco del Programa de apoyo al fortalecimiento Institucional del SENASA. (EuropeAid/123383/C/SER/AR), mediante la Licitación 6/2008 L4. Este proyecto se vincula – además – con líneas de investigación y desarrollo de sus integrantes y colaboradores en otras instituciones públicas.

## **Introducción**

Los modelos de distribución de especies (SDMs) han tomado un rol cada vez más destacado en la comprensión de las actuales y potenciales distribuciones de especies. Los SDMs permiten establecer relaciones de distribución las especies vegetales y/o animales con los patrones de distribución de variables ecológicas, dando como resultado un mapa de distribución potencial del hábitat.

Estudios actuales de biogeografía, biología de conservación, ecología, paleoecología, manejo de vida salvaje y control de plagas que afectan a cultivos, los utilizan. Estos modelos estudian las asociaciones entre la ocurrencia geográfica de una especie y un grupo de variables predictivas para establecer los mecanismos que gobiernan la distribución de las especies, apoyándose en un entorno SIG con soporte de almacenamiento de información en bases de datos geoespaciales. Así, estos modelos permiten la estimación de los requerimientos ecológicos de la especie. (Araújo y Guisan, 2006).

Los modelos de distribución son ampliamente utilizados para evaluar el potencial de propagación de especies invasoras (Peterson et al., 2003; Rouget et al., 2004; Thuiller et al., 2005b), identificar y gestionar especies

amenazadas (Engler et al., 2004; Norris, 2004), priorizar zonas de conservación de biodiversidad (Araujo et al., 2004; Ortega-Huerta et al., 2004; Sanchez-Cordero et al., 2005) y evaluar el impacto potencial del cambio climático en los patrones de distribución de las especies (Skov et al., 2004; Beaumont et al., 2005; Bomhard et al., 2005; Thuiller et al., 2005a; Thuiller, et al., 2005).

Los organismos responsables de la sanidad vegetal y animal de nuestro país, trabajan en la incorporación de nuevas tecnologías, estas se orientan a que expertos temáticos en el control y monitoreo de especies cuarentenarias, en general con menos experiencia de uso de SDMs, dispongan de una herramienta que, vía una interfaz web, permita el análisis de condiciones ambientales predisponentes para el establecimiento y desarrollo de estas plagas.

Siguiendo con lo anteriormente mencionado, se ha desarrollado una aplicación para la vigilancia fitosanitaria, la cual está compuesta de una base de datos bioclimáticos y territoriales, con sus correspondientes interfaces de consulta, exploración y visualización de datos. Por otro lado se ha comenzado a desarrollar una librería de modelos de plagas y enfermedades que se ejecutan en un ambiente web sobre datos actuales o proyectados.

En este contexto, el proyecto tiene entre sus principales objetivos la construcción de herramientas informáticas de asistencia al modelado de la distribución de plagas y enfermedades de interés fitosanitario. Se busca contar con herramientas que permitan a los organismos de control: a) disponer de información climática y geográfica relevante (en forma de mapa y tabular)

para la toma de decisiones en la temática fitosanitaria, b) diseñar, poner a punto y ejecutar modelos de distribución de plagas y c) identificar áreas potenciales de dispersión de plagas.

### **Características de la Base de Datos Aplicados a la Vigilancia Fitosanitaria**

A los efectos de relevar la mayor cantidad de información de interés fitosanitaria, se ha diseñado una estructura de datos que consiste en una grilla que recubre el territorio argentino. Cada elemento es de 0.1° lo que representa – aproximadamente – cuadrados de 10 km por 10 km. Definiendo en total 33.333 elementos o celdas, donde cada uno representa una zona geográfica determinada que tiene asociado numerosos atributos de interés. A este conjunto de celdas se lo denomina “grilla nacional”.

El sistema SIGAFIT posee una base de datos geoespacial fitosanitaria que almacena atributos de interés asociados a cada elemento de la grilla nacional. Las categorías o grupos de datos son las siguientes; Administrativos/políticos, biogeográficos/cobertura, Censo Agropecuario 2002, fotoperíodo/heliofanía, índices de vegetación, información de suelos, información topográfica, estadísticas climáticas mensuales, estadísticas bioclimáticas anuales, datos Worldclim<sup>1</sup> bioclimáticos y datos Worldclim climáticos.

La base de datos ha sido construida utilizando un software libre de gestión de datos (DBMS) denominado

---

<sup>1</sup> Worldclim, <http://www.worldclim.org>, Conjunto de datos climáticos de cobertura mundial.

PostgreSQL<sup>2</sup>. El manejo de los datos espaciales en la base de datos se realizó por medio de la extensión denominada PostGIS<sup>3</sup>, la cual brinda soporte de almacenamiento a datos geográficos y una serie de funciones para el tratamiento y consulta de los mismos.

La base de datos está compuesta por tablas que almacenan uno o varios atributos de interés a los objetivos del proyecto. El conjunto de todas las tablas forman la denominada grilla nacional. La relación entre cada una está dada por el atributo “gid”, el cual es identificador de celda e independientemente de la tabla que sea, un mismo número de gid siempre va a hacer referencia a un mismo espacio físico en el espacio argentino.

En la actualidad existen más de 120 tablas de datos temáticos que contienen en total más de 1.600 atributos, los cuales pueden ser consultados directamente por los operadores o los modelos de dispersión que se han implementado.

### **Módulo de Exploración de Datos Geospaciales (MEDG)**

El MEDG tiene por objetivo la permitir la formulación de consultas sobre la base de datos del Sistema de Información Geográfica Fitosanitaria (SIGAFIT) de forma simple, amigable y efectiva. Las consultas se diseñan sobre un formulario web, donde un técnico dispone de una serie de metadatos de asistencia a la tarea.

Así, se pueden realizar consultas que permitan delimitar el área potencial de

---

<sup>2</sup> PostgreSQL, <http://www.postgresql.org/>, Sistema de gestión de base de datos relacional orientada a objetos.

<sup>3</sup> PostGIS, <http://postgis.refractory.net/>, Extensión para soporte de objetos espaciales en PostgreSQL

presencia de una especie de acuerdo a criterios y umbrales de las variables consultados en la bibliografía de referencia. A modo de ejemplo, es posible estimar la distribución de una plaga con presencia en las siguientes condiciones: "Zonas geográficas en las provincias de K o L" donde "la temperatura promedio del mes de P esté entre X y Y °C", "se cultive A o B", "estén a más de N kilómetros de rutas provinciales o nacionales" y "no superen los Z metros de altura sobre el nivel del mar". Estos requerimientos se vuelcan sobre el MEDG mediante un formulario electrónico que acepta dos tipos de atributos, a saber: a) numéricos (por ejemplo valor de temperatura promedio o distancia a una ruta nacional) y b) alfanuméricos (por ejemplo datos territoriales y de ecoregiones).

Una vez que el usuario realiza una consulta al MEDG, la misma se convierte en instrucciones SQL que serán procesadas por el gestor de la base de datos y darán por resultado una tabla donde cada fila representa un elemento de la grilla nacional que satisface los requerimientos planteados. La información que conforma el conjunto "respuesta" puede ser visualizada de distintas formas (modos mapa y texto) y descargada en archivos en diferentes formatos que pueden ser exportados como capas temáticas para ser integradas a sistemas información geográfica, de acuerdo a lo que se describe a continuación:

El formato tabular ofrece la opción de visualización únicamente de atributos numéricos y alfanuméricos correspondientes a las celdas de la grilla nacional que componen el juego de resultados.

El formato mapa puede visualizarse gráficamente como una capa más de un SIG con interfaz web, mediante la aplicación P.mapper<sup>4</sup>, que permite al usuario una buena exploración. El sistema web aporta asimismo una serie de capas de información (en formato raster y vector) que optimizan la tarea del operador.

El formato Google Map/Earth muestra el conjunto de celdas de salida como una capa de Google Map o Google Earth (modo 3D) que se ejecuta sobre una página web normal (no requiere de la descarga de programas específicos en la computadora del operador). Cuando el operador selecciona una celda resultado se mostrarán los atributos intervinientes en la consulta con sus valores asociados.

Otros formatos de archivos de intercambio se han implementado, así se dispone de varias alternativas exportación de resultados. En particular, el sistema ofrece salidas para Google Earth en el formato KML, en el estándar GML (Geographic Markup Language) y en formato propietario de la empresa ESRI denominado Shapefile.

Estos diferentes formatos han sido incluidos en SIGAFIT como parte de una política de accesibilidad a la información espacial, los archivos tienen asociados metadatos con información de las fuentes y calidad de los datos.

A los efectos de validar el funcionamiento del módulo de exploración de datos geospaciales se han implementado -para su ejecución automática- dos modelos de dispersión de

---

<sup>4</sup> P.mapper, <http://www.pmapper.net/>, Framework de visualización de información .

potenciales enfermedades sobre cultivos vegetales. Los resultados (áreas de riesgo categorizadas) fueron contrastados con los obtenidos en forma manual, a partir de interactuar con el módulo de análisis de datos geoespaciales de un sistema de información geográfica. Del análisis se concluye que los modelos funcionan correctamente.

## Resultados

La base de datos geoespaciales y el MEDG se encuentra operando sobre la red intranet de la institución, y están siendo utilizados por profesionales y técnicos de SENASA de todo el país.. Actualmente se están realizando modificaciones a las interfaces de usuario, a fin de mejorar la operatividad por parte de usuarios no expertos en SIG.

Con respecto a los SDMs, un primer resultado obtenido fue el desarrollo de un mapa de distribución potencial de *Lobesia botrana* en Argentina, en base a parámetros bioclimáticos que regulan su distribución espacio-temporal en los países en los que se halla presente. Así como también, elaborar una herramienta para identificar a nivel nacional, las regiones donde se darían las condiciones ambientales apropiadas para que la plaga se instale y desarrolle su ciclo de vida. En el modelo intervienen como variables el NDVI, la precipitación, las temperaturas mínima, media y máxima. El mismo está basado en datos históricos de temperatura del Servicio Meteorológico Nacional.

El estudio de *L. botrana*, que al momento de realizarlo no se encontraba registrada en Argentina, resultó acertado en su pronóstico, acertando con gran precisión en las primeras apariciones en el país. En la figura 1 se puede observar una copia de

la imagen de pantalla correspondiente a las áreas de riesgo de *Lobesia botrana*.

Otro modelo desarrollado con buenos resultados es el que corresponde a *Lymantria dispar*. lepidóptero polífago que, en su etapa de larva, produce importante defoliaciones a especies de árboles en su actual rango de distribución (Liebhold, Elkinton, & Williams, 2000).

A futuro el desafío es el desarrollo de SDMs que además de incluir datos climáticos incluyan variables de tipo meteorológico (en tiempo real) para la evaluación de condiciones en terreno y toma de decisiones. Esto implica la necesidad de contar con datos suficientes por parte de los organismos encargados de la toma de estos datos (Servicio Meteorológico Nacional, INTA, entre otros) en tiempo y forma para su inclusión en el sistema. Asimismo, se está evaluando el desarrollo de herramientas que permitan el análisis conjunto de variables que expliquen la distribución de especies que afectan a cultivos con la finalidad de tener modelos para la toma de decisiones y planificación de las tareas de seguimiento, control y erradicación de plagas.

Por otro lado, a partir de la experiencia realizada obtenida, se está diseñando un sistema de información regional para la provincia de Entre Ríos que basado también en herramientas de tipo libre u *open source* pueda operar con mayor resolución espacial y permita asistir a organismos públicos y privados en las siguientes temáticas:

- Emergencias ambientales como por ejemplo incendios, y la inclusión de variables que intervienen en modelos de probabilidad de ignición.

- Problemas epidemiológicos y su relación con variables socioeconómicas y ambientales.

- Gestión de la producción agropecuaria y la interrelación entre datos productivos y variables ambientales.

### **Formación de Recursos Humanos**

En el marco del presente proyecto de investigación y desarrollo de herramientas de asistencia a la vigilancia de plagas en vegetales, se han organizado y dictado una serie de cursos de postgrado, talleres y jornadas sobre Geoestadística, Procesamiento Digital de Imágenes y análisis de series temporales entre otros, que han aportado nuevas miradas y potenciales soluciones a problemas derivados del proyecto. En la actualidad también se ha integrado al equipo de trabajo de CEREGeo/FCYT-UADER un becario de CONICET que tiene como línea de trabajo para su doctorado el desarrollo y aplicación de SDMs

### **Referencias**

Araújo, M.B., Cabeza, M., Thuiller, W., Hannah, L. y Williams, P.H. 2004. Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. *Global Change Biology*, 10, 1618–1626.

Araújo, M.B.; Guisan, A. 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33: 1677–1688.

Beaumont, L.J., Hughes, L. y Poulsen, M. 2005. Predicting species' distributions: use of climatic parameters in BIOCLIM and its impact on predictions of species'

current and future distributions. *Ecological Modelling*, 186, 250–269.

Bomhard, B., Richardson, D.M., Donaldson, J.S., Hughes, G.O., Midgley, G.F., Raimondo, D.C., Rebelo, A.G., Rouget, M. y Thuiller, W. 2005. Potential impacts of future land use and climate change on the Red List status of the Proteaceae in the Cape Floristic Region, South Africa. *Global Change Biology*, 11, 1452–1468.

Engler, R., Guisan, A. y Rechsteiner, L. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, 41, 263–274.

LeVeen, E. P. 1989. Economic evaluation of eradication programs. En D. L. Dahlsten & R. Garcia, (edts.), *Eradication of exotic pests*. Yale University Press, New Haven, Connecticut, USA. 41-56.

Liebhold, A., Elkinton, J., & Williams, D. (2000). What causes outbreaks of the gypsy moth in North America? *Population Ecology*, 257-266.

Norris, K. 2004. Managing threatened species: the ecological toolbox, evolutionary theory and declining-population paradigm. *Journal of Applied Ecology*, 41, 413–426.

Ortega-Huerta, M.A. y Peterson, A.T. 2004. Modelling spatial patterns of biodiversity for conservation prioritization in north-eastern Mexico. *Diversity and Distributions*, 10,39–54.

Peterson, A.T. y Robins, C.R. 2003. Using ecological-niche modelling to predict barred owl invasions with

implications for spotted owl conservation. *Conservation Biology*, 17, 1161–1165.

Rouget, M., Richardson, D.M., Nel, J.L., Le Maitre, D.C., Egoh, B. y Mgidi, T. 2004. Mapping the potential ranges of major plant invaders in South Africa, Lesotho and Swaziland using climatic suitability. *Diversity and Distributions*, 10, 475–484.

Sanchez-Cordero, V., Cirelli, V., Mungui, M. y Sarkar, S. 2005. Place prioritization for biodiversity content using species ecological niche modelling. *Biodiversity Informatics*, 2, 11–23.

Skov, F. y Svenning, J.C. 2004. Potential impact of climatic change on the distribution of forest herbs in Europe. *Ecography*, 27, 366–380.

Thuiller, W., Lavorel, S. y Araújo, M.B. 2005. Niche properties and geographical extent as predictors of species sensitivity to climate change. *Global Ecology and Biogeography*, 14, 347–357.

Thuiller, W., Lavorel, S., Araujo, M.B., Sykes, M.T. y Prentice, I.C. 2005a. Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences USA*, 102, 8245–8250.

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O. y Rouget, M. 2005b. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, 11, 2234–2250.

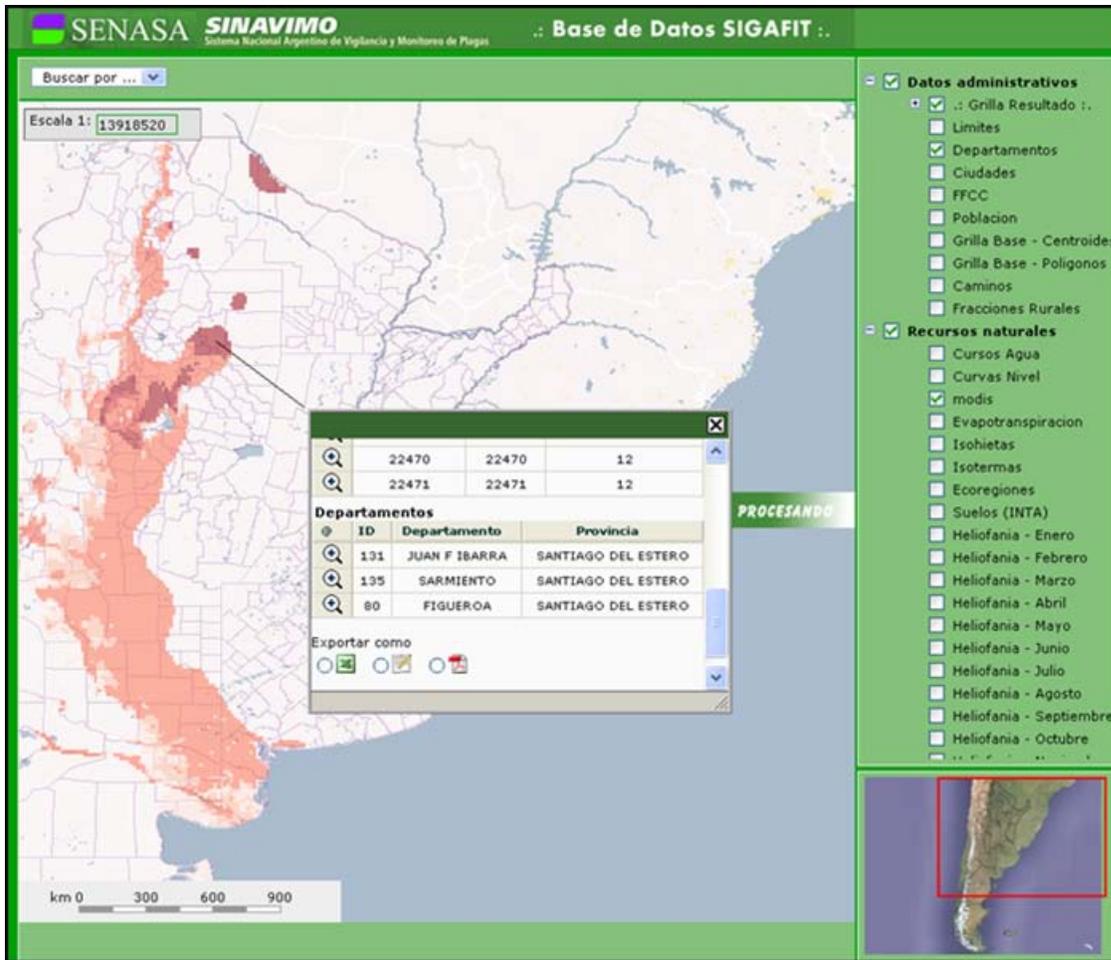


Figura 1 – Salida en formato mapa correspondiente a las áreas de riesgo de *Lobesia botrana*.