

TALLER MODERADORES Y MODELADORES DE BIOMODELOS

biomodelos.humboldt.org.co

Medellín, Marzo 10 · 2016

Participantes:

Álvaro Cogollo
Cristina López Gallego
Danny Zurc
Diana Carolina Guzmán
Felipe Andrés Toro
Germán Forero
Hugo López
Hugo Mantilla Meluk
José Nicolás Urbina
Juan Manuel Daza
Norberto López
Sergio Chaparro Herrera
Tatiana Velásquez
William Agudelo
Yanira Cifuentes

Moderadores:

Jorge Velásquez Tibatá
María Helena Olaya Rodríguez

El alcance y relevancia de BioModelos

Los participantes consideraron que BioModelos debe ser tomada como una aplicación que provee información dinámica en el tiempo (como lo es la nomenclatura). Se debe hacer énfasis en que los modelos son hipótesis que están en constante cambio y que son objeto de críticas y modificaciones que buscan mejorar la información sobre la distribución real de las especies. En este sentido, se resalta que no se busca llegar a verdades absolutas con base en los mapas producidos, y que el gran valor de BioModelos radica en el la participación de los expertos en la construcción de la plataforma y en la validación de los modelos, pues son ellos quienes solicitan las herramientas para mejorar los modelos y así mismo dan el visto bueno o generan la confianza para la adopción de las hipótesis que se están publicando en un momento en el tiempo, que son las que podrían apoyar los procesos de toma de decisiones.

Se mencionó que aquellos grupos que ya tengan avanzada la curaduría de los datos (tanto en términos de taxonomía como de validación geográfica), como son el grupo de zamias y el de primates, deben ser priorizados para avanzar en el flujo de trabajo y publicar las hipótesis de distribuciones de especies lo antes posible durante este año. Se recomienda que los avances y procesos de los grupos de trabajo de BioModelos sean en lo posible documentados y compartidos con el resto de grupos (tanto con los moderadores como con las sociedades y asociaciones) para adaptar o seguir los flujos de trabajo que los demás hayan realizado, en especial aquellos que vayan más avanzados.

Pertinencia de la curaduría colaborativa de registros de BioModelos

Se considera que el trabajo de curaduría de registros biológicos es una tarea de vital importancia, tanto para BioModelos como para los investigadores y académicos del país. En este sentido, los participantes estuvieron de acuerdo con que la curaduría debe ser coordinada y elaborada por las sociedades o asociaciones. En términos de la actualización de la nomenclatura, se propuso que ésta se realizara como mínimo anualmente (idealmente semestralmente). Es importante conocer claramente el procedimiento que sigue el grupo de plantas, para poner en funcionamiento un sistema estable y continuo para la actualización de la nomenclatura, como lo es el Catálogo de Plantas de Colombia (<http://catalogoplantasdecolombia.unal.edu.co/>) el cual fue designado como la autoridad taxonómica para este grupo en particular. El catálogo es quien da la lista oficial de las especies presentes en Colombia y todos los botánicos se ciñen a ella. Para el caso de los grupos taxonómicos de vertebrados, se propone que sean las sociedades y asociaciones quienes se conviertan en las autoridades taxonómicas para las listas de especies de sus respectivos grupos. Así, los expertos no solo darán su visto bueno en términos del producto de BioModelos (validación de modelos) si no desde la misma base de su construcción (confiabilidad en los datos con los cuales se generan las hipótesis). Es importante destacar el proceso de curaduría que se hace desde las colecciones biológicas y dejar este trabajo plasmado en la memoria o metadatos de los modelos, porque los participantes son conscientes que si bien estos ejercicios deben ser reportados a las bases de datos como GIBIF o el SiB, la publicación de dichas curadurías en estas plataformas pueden tomar demasiado tiempo.

Sobre las actualizaciones de las bases de datos, los participantes recomiendan mantener un histórico de los cambios que se realicen para saber cuál es la versión más actual y monitorear qué tantos cambios se presentan y en qué grupos. El nuevo ingeniero del equipo del Laboratorio de Biogeografía Aplicada – Instituto Humboldt (LBA) contemplará esta tarea.

Publicación en línea de bases de datos de distribución curadas

La APC ha optado por seguir un flujo de trabajo en el que la asociación recibe los datos directamente de los investigadores, allí son depurados por los expertos y de haber correcciones se le devuelve la base de datos al publicador para que las realice y finalmente la base de datos sea entregada directamente al SiB. Para la APC, lo más importante de trabajar con el SiB es generar unos campos lo más sencillo posibles para que los publicadores no se sientan abrumados por el formato Darwin Core. Si bien este flujo se está implementando para la APC, los demás participantes consideran que una gran mayoría de publicadores no se pondrían en la tarea de corregir sus bases de datos, y en muchos casos este flujo de trabajo no aplicaría porque muchos registros son de publicadores que ya han fallecido. Si bien es importante que se tenga presente el asunto de los derechos de autor (como bien siempre lo ha mencionado el SiB), en tanto que la modificación de los datos solo lo puede hacer el publicador original, es importante que se haga el ejercicio de curaduría para la generación de los modelos y que dicha curaduría sea reconocida en la memoria o meta datos de los modelos publicados. Se mencionó también que la curaduría de datos a partir de fuentes secundarias tampoco se podría publicar en el SiB en la actualidad debido a los problemas de derechos de autor.

Las bases de datos curadas con las cuales se generen modelos podrían ser visualizadas en BioModelos en la medida que se hagan acuerdos de entendimiento, bajo restricciones de uso y visualización para proteger procesos de producción intelectual y/o grupos susceptibles a la extracción (especies traficadas). Dichas bases de datos, de ser visualizadas, deben reconocer tanto la fuente de los datos como la autoría de la curación.

Tareas BioModelos

- Definir con cada grupo la base de los registros biológicos y como se realizará la curaduría de los registros de las especies que se van a trabajar en el 2016. La fuente de los datos puede ser la compilada por el LBA, la compilada por los propios grupos, o una base de datos que integre ambas. Para curar los datos, cada grupo puede usar el software SIG y recursos cartográficos que estén a su alcance, o hacerlo a través de herramientas que se dispongan para ese fin en BioModelos.
- Identificar las herramientas de curación de registros biológicos que se necesitan implementar o mejorar en la aplicación BioModelos.
- Definir modelo de acuerdos de entendimiento para la visualización de registros en BioModelos, contemplando condiciones de uso (¿Qué usos son permitidos?) y de visualización (¿Quiénes pueden ver los registros en BioModelos?).

Desarrollo de modelos

Sin duda, el valor agregado de BioModelos es involucrar el “ojo del experto” en el proceso de validación de las hipótesis de distribución de las especies.

Para la identificación del método de modelamiento, se considera que es el equipo del LBA quien debe hacer la selección pues tienen mayor conocimiento sobre dichos algoritmos. Una vez se seleccione esto, los expertos entrarían a escoger el umbral y a realizar ediciones –delineación de la extensión de ocurrencia de las especies (adiciones o sustracciones) de los modelos propuestos para generar un modelo híbrido. En este sentido, los participantes consideran que no se debe hacer una elección de la mejor hipótesis como se presenta en este momento a los usuarios, si no que los expertos deben trabajar sobre los modelos estadísticos presentados por el equipo del LBA (los cuales deben ser construidos con datos curados), y este es el mapa consenso que se debe presentar a los usuarios tan pronto pongan una especie en el buscador del home.

Al hablar del módulo que le permita ver a los usuarios en tiempo real su mapa, los participantes opinaron que esta opción no es prioritaria y que más bien BioModelos podría enfocar más esfuerzos en crear espacios para hacer “Modelatones” (espacios para reunir a los expertos y generar los modelos consenso), que para este año se podrían hacer durante los congresos de herpetología y ornitología. En este sentido, los participantes aconsejan revisar si este módulo sería una prioridad para BioModelos considerando las opiniones de los ingenieros del equipo. Si bien reconocen que para los modeladores podría ser atractivo, consideran que la aplicación debe enfocarse en masificar la divulgación y apropiación de la información sobre distribuciones de especies.

Para los ejercicios de “modelatón”, los participantes recomiendan que BioModelos cuente con la visualización (y ojalá en un futuro con la posibilidad de hacer análisis espaciales de cruces de capas) de capas como ecoregiones o distritos biogeográficos para mejorar los procesos de edición y validación de las hipótesis de distribución. También se propuso desarrollar dentro de la aplicación una herramienta para que las ediciones o validaciones de los puntos de registros queden automáticamente con la autoría del experto.

Finalmente se propuso el flujo de trabajo de la figura 1. Allí, la base para la generación de los modelos deben ser las bases de datos curadas. La generación y selección de los modelos estadísticos o inductivos la realiza el LBA, como ya fue mencionado. Una vez se tenga un modelo inductivo, éste será evaluado por los expertos designados desde los grupos a través de los moderadores, mediante la elección del umbral para el modelo inductivo y la realización de ediciones espaciales que determinen el área de ocupación de la(s) especie(s) en cuestión. Si solamente participa un experto en este proceso, el mapa resultante se considerará el modelo híbrido o consenso y será el que se publica en BioModelos cuando un usuario busque una especie directamente en el home. Si hay más de un experto (y la opinión generalizada es que no habrá más de 4 expertos para una especie en particular o grupos de especies) y los modelos híbridos son similares, el LBA se encargará de integrar estos resultados en un solo modelo híbrido que será presentado a través de los moderadores a los miembros de los grupos para que avalen dicho mapa y posteriormente pueda ser presentado a los usuarios a través de la aplicación. Si llegasen a existir diferencias muy marcadas en los modelos híbridos de los expertos, ambos se mostrarían en BioModelos y la calificación de los usuarios podría servir como base para guiar a los usuarios acerca de cuál tiene más soporte.

Es importante tener en cuenta que ya hay varios investigadores que han desarrollado y publicado modelos con hipótesis de distribución de forma ajena a BioModelos. Estos mapas deben ser tenidos en cuenta antes de correr modelos, pues no se quiere duplicar esfuerzos. Si dichos mapas tienen robustez y buen soporte externo (algo que debe evaluar el moderador y los miembros del grupo que se consideren expertos en dichas especies, así como otros modeladores) esos modelos son los que se deberían publicar para las especies sin necesidad de la construcción de un modelo inductivo. Si los modelos externos no son avalados por los grupos, se desarrolla el flujo de trabajo de BioModelos, pero se publicarían tanto el mapa consenso resultante como el mapa externo y tendrán una votación dentro de los grupos que los usuarios podrán consultar para la descarga final de la información.

En cualquiera de estos casos donde se presenten más de un modelo para ofrecerle a los usuarios de BioModelos, se podrá adoptar una de las siguientes opciones: 1) los modelos son puestos a votación por los expertos de la especie y estas votaciones (o sus promedios) se publican junto a cada una de las hipótesis de distribución; en este sentido, el usuario visualizaría todas las hipótesis (modelos híbridos o modelos externos) y podría consultar cuál modelo tiene mayor apoyo por parte de los expertos para elegir cuál de todos descargar. 2) Los modelos son puestos a votación entre los expertos y se escoge y publica el que tenga mayor número de votos. El sistema de votación debe ser escogido: sistema de estrellas de 1 a 5 o sistema de pulgares arriba o abajo.

Para aquellas especies que no tengan modelo consenso, se propone que aparezcan con una advertencia sobre esta situación o que el modelo consenso se encuentra en proceso de ser construido. En este caso se podría presentar el modelo inductivo y el modelo deductivo de la IUCN si es el caso, para que los usuarios voten y descarguen el que les parezca mejor. Al final del proceso, lo que se espera es que se presente un solo mapa que contenga la mejor hipótesis de distribución de la especie de acuerdo a la opinión experta de los grupos y las sociedades/asociaciones.

Un aspecto importante que se mencionó fue el caso de aquellas especies que cuenten con menos de 10 registros. Los participantes señalaron que muchas de estas especies seguramente coinciden con especies endémicas, y por tanto, requieren el diseño de una metodología o módulo especial en BioModelos donde los expertos puedan generar los modelos a mano según sus conocimientos de distribución de dichas especies. Para esto sería necesario concertar una reunión previa dentro de los grupos para revisar e identificar estas especies. En BioModelos, es necesario que al consultar estas especies salga una advertencia que mencione que son especies con pocos registros y por tanto se utilizan mapas de experto.

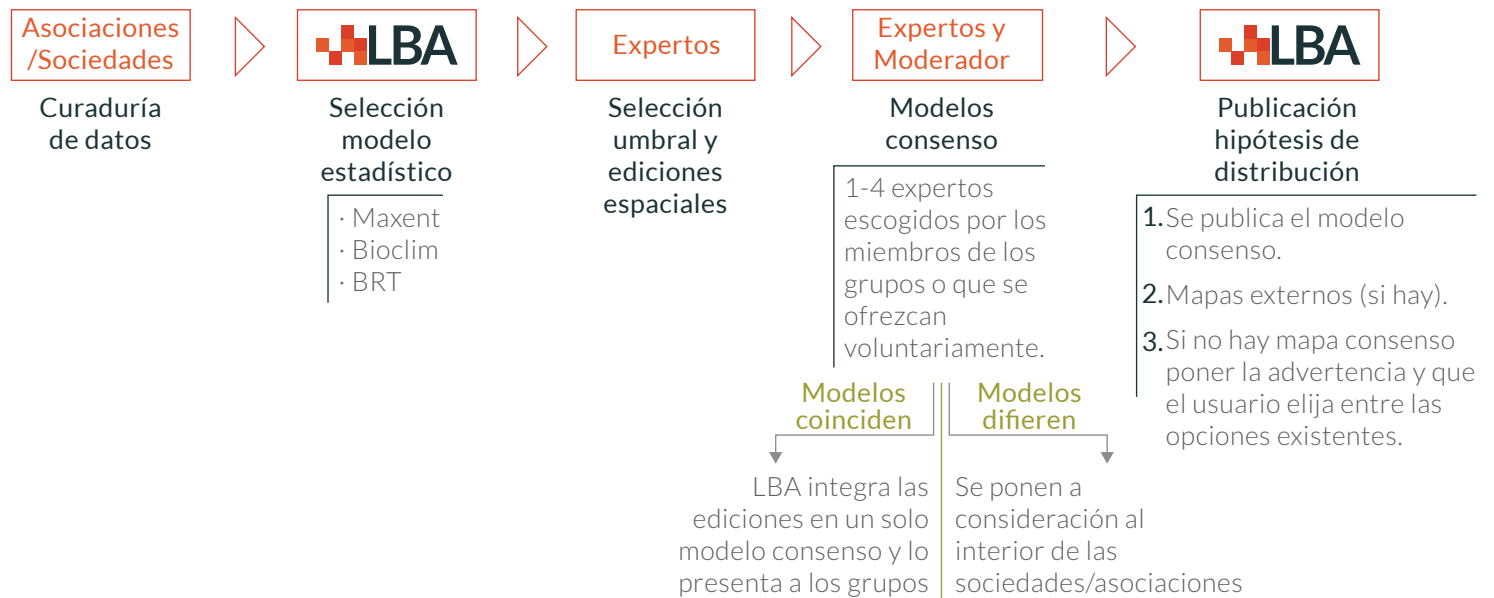


Figura 1. Flujo de trabajo para la publicación de hipótesis de distribución de especies, propuesto durante el taller de moderadores y modeladores de BioModelos.

Tareas BioModelos

- Modificar la lógica de navegación en el visor para mostrar apenas se realiza una consulta el modelo validado de la especie (ayuda muestra en la página donde está la información del modelo), mapa de IUCN consultado por servicios web en caso que el modelo no exista o registros (en caso que haya más de un modelo o no exista modelo; ayudas muestran que hacer en esos casos).
- Espacio “mis mapas” en el perfil de cada experto mostrando los resultados de sus ediciones a modelos. No se genera en tiempo real sino con cierta regularidad (y se notifica al usuario cuando está listo).
- Cargar capa de ecoregiones en Bio Modelos.
- Actualizar materiales informativos del proceso biomodelos para ajustarse a la nueva propuesta.
- Levantar inventario de especies con modelos.
- Modulo dentro de BioModelos que permita desarrollar mapa de experto para las especies con muy pocos registros para la elaboración de un modelo.

Agendas de trabajo por grupos

Pensando en la información que le puede ser útil a los tomadores de decisiones es importante tener en cuenta zonas de interés como pueden ser los ecosistemas estratégicos o amenazados, o priorizar especies de interés nacional como especies polinizadores o vectores de transmisión de enfermedades y en escenarios como cambio climático, cambios en el uso de la tierra, etc.

Sin embargo, también es importante considerar que el avance en el conocimiento de las especies de los diferentes grupos va a velocidades muy diferentes y están a diferentes niveles. En este sentido, solo cada grupo sabe con qué especies se puede comprometer y posiblemente no respondan o solo respondan en cierta medida a estas necesidades nacionales para los tomadores de decisiones. En este sentido se concluye que es mejor que cada grupo desarrolle sus agendas particulares, pero que se mantenga la comunicación entre grupos para retroalimentarse en el proceso de selección de las especies.

Mamíferos

Se lograron consolidar 17 especies de mamíferos entre los miembros del grupo, de las cuales 6 son objetos de conservación para PNN.

Hugo López invitó al equipo de BioModelos al taller que realizará en la Universidad Nacional sobre modelación de distribución de mamíferos.

Adicionalmente, la tesis de Tatiana Velásquez generará modelos de distribución para especies de delfines y nutrias que compartirá para su publicación en BioModelos.

Herpetos

El grupo decidió trabajar con las ranas de cristal, para las cuales ya se realizó la búsqueda de literatura y se cuenta con la curaduría de las bases de datos.

Adicionalmente el grupo está interesado en trabajar los anfibios andinos, priorizando los de páramo ya que cuentan con una base de datos curada inicial. También se mencionó que la WCS está trabajando una publicación sobre tortugas que sería interesante revisar.

Los miembros de grupo solicitaron a BioModelos las bases de datos para estas especies para su revisión y curaduría.

Primates

Se comentó que la APC ya se encuentra realizando un proceso con el SiB Colombia para curar las bases de datos directamente con los publicadores de los datos. En este sentido, se menciona que es necesario comunicarnos con el presidente de la APC para solicitar estas bases de datos.

Aves

Para las aves acuáticas se quiere retomar el trabajo de “Modelatón” que se había realizado previamente revisando cuidadosamente de nuevo la nomenclatura. Adicionalmente se quieren priorizar especies con pocos registros y especies de patos.

El grupo aves de Colombia está interesado en modelos de especies migratorias de importancia para la conservación (trabajadas para planes de manejo), que en total son 33 especies. Además se esperan modelos de trabajos de estudiantes y del libro de búhos de Colombia.

Zamias y magnolias

Estos dos grupos tienen bastante trabajo adelantado en la curaduría de datos y validación de modelos. En este sentido, consideran que este año habrá modelos consensados para publicar en BioModelos.

En el caso de magnolias también hay un gran avance en la curaduría de las bases de datos que se realizó para la publicación del libro rojo de magnolias de Colombia. En este sentido, se menciona que 15 especies ya cuentan con estas bases de datos y los esfuerzos se podrían enfocar en las 11 que se encuentran en estado crítico de amenaza.

Tareas pendientes

1. Contar con los listados de especies priorizadas por grupos.
2. Que los moderadores identifiquen los expertos para las especies priorizadas.
3. Plantear el cronograma de desarrollo de modelos por grupo con fecha límite Noviembre 2016.
4. Revisar para cuáles de estas especies ya existen hipótesis de distribución y que los grupos las evalúen para su publicación.