**Sobre el intercambio de datos de vegetación: el estándar ‘Veg-X’ y el paquete de R ‘VegX’**

Miquel De Cáceres1,2,\*

1 Centro Tecnológico Forestal de Cataluña, Ctra. antiga St. Llorenç de Morunys km. 2; 25280 Solsona, Catalunya, España;

2 CREAF, Cerdanyola del Vallès 08193, España;

\* Autor de correspondencia: [miquelcaceres@gmail.com](mailto:miquelcaceres@gmail.com)

Actualmente existe una cantidad ingente de datos sobre la vegetación de los ecosistemas terrestres del mundo. En muchos casos se trata de datos de parcelas de muestreo, establecidas para estudios concretos o que forman parte de esfuerzos colectivos de observación a gran escala, como son los inventarios florísticos o los inventarios forestales. A menudo dichos datos se encuentran archivados en bases de datos relacionales, con una estructura de datos conocida. El *Global Index of Vegetation-Plot Databases* (GIVD; [www.givd.info](http://www.givd.info)) (Dengler et al. 2011) es una meta-base de datos que en marzo de 2018 contiene información sobre 260 bases de datos de vegetación. Sin embargo, los datos de la mayoría de las bases de datos listadas en GIVD a menudo están almacenados en bases de datos con estructuras únicas. Esto no supone un problema para estudios que utilicen bases de datos grandes, tales como los inventarios forestales, dado que solo es necesario comprender una sola estructura de datos (Wiser 2016). Sin embargo, cuando uno desea integrar distintas fuentes de datos para realizar nuevos análisis es necesario establecer mapas de campos y conceptos, cosa que enlentece en gran medida el proceso y puede ser fuente de errores. En el caso de redes de parcelas a escala mundial que comparten los mismos protocolos de muestreo de campo, el problema se puede aliviar estableciendo estándares de modelos de datos para bases de datos (Condit et al. 2014). Sin embargo, esta solución no es aplicable para armonizar e integrar datos provenientes de distintos estudios (e.g. distintas formas de las parcelas y/o subparcelas, distinta definición de estratos, etc.). Así pues, los análisis requiriendo la integración de datos de vegetación de varias fuentes se ven dificultados por la no existencia de estándares de intercambio de datos de vegetación reconocidos internacionalmente.

En 2003, el grupo de trabajo de ecoinformática de la *International Association for Vegetation Science* (IAVS; [www.iavs.org](http://www.iavs.org)) decidió promover el desarrollo de un estándar de intercambio de datos de parcelas de vegetación. A partir del trabajo en distintos talleres internacionales se llegó a la publicación de *Veg-X*, un estándar de intercambio de datos implementado en un esquema XML (extensible markup language) y diseñado para ser compatible con un gran número estructuras de datos de vegetación existentes (Wiser et al. 2011). El estándar distingue entre la entidad observada (p.ej. un árbol) y el acto de observación (p.e. las mediciones) y permite almacenar tanto datos de plantas individuales (p.ej. diámetros, alturas, …) como datos de plantas ‘agregados’ (p.ej. porcentaje de cobertura de una especie), incluyendo su posición en estratos verticales. El estándar también soporta el muestreo repetido de organismos y parcelas, así como la agrupación de observaciones por criterios no temporales (p.ej. a partir de definir sub-parcelas). Todos los elementos del esquema XML están claramente definidos para facilitar el mapeo de los campos de cualquier base de datos de vegetación.

Desde su publicación, una de las barreras a la adopción de *Veg-X* para el intercambio de datos ha sido su complejidad, por otro lado necesaria para acomodarse a la gran variedad de estructuras de datos existentes. Para conseguir que el formato de intercambio sea adoptado por un conjunto amplio de usuarios es importante desarrollar herramientas ecoinformáticas para facilitar el mapeo desde distintas estructuras de datos a *Veg-X*, así como para la integración de documentos en dicho formato, incluyendo la armonización de unidades, etc. Desde 2017, desde el grupo de trabajo de ecoinformática de la IAVS nos hemos puesto a desarrollar un paquete de R que sirva para realizar las tareas mencionadas. El desarrollo del paquete, llamado VegX, se ha realizado en paralelo a una revisión importante del estándar mismo. Tanto el esquema en XML del estándar (ver. 2.0) como el paquete de R y su documentación (descripción detallada del esquema, manual de usuario del paquete) se encuentran disponibles en un repositorio de desarrollo y distribución de GitHub (<https://github.com/miquelcaceres/VegX>).

Aunque se puede considerar funcional, el paquete VegX se encuentra actualmente en un estado de desarrollo en el que es necesario testear su usabilidad para importar e integrar distintos tipos de datos fuente. Además, uno de nuestros objetivos es el de promover que los datos de vegetación asociados a publicaciones científicas que sean archivados en repositorios públicos (e.g. DRYAD; <https://datadryad.org/>) estén almacenados mediante estándar tales como *Veg-X*, para facilitar su reutilización en estudios posteriores. Para ello, sería interesante disponer de varios ejemplos de importación de datos (escritos p.ej. en Rmarkdown). Es por ese motivo que quisiera aprovechar esta Nota Ecoinformática para hacer un llamamiento a investigadores (ecólogos terrestres, ingenieros forestales, botánicos, etc.) con datos de vegetación propios (p.ej. en tablas de Excel o ficheros ASCII de texto) interesados en probar la importación de sus datos y reportarnos los problemas o dudas que hayan surgido, así como el código usado en el proceso. No es necesario que dichos datos sean públicos, dado que solamente nos interesa su estructura, no su contenido. Esta información nos ayudará enormemente a mejorar la herramienta que proponemos.

**Agradecimientos**

Gracias al grupo de ecoinformática de la AEET por la permisión de escribir esta nota. Miquel De Cáceres cuenta con una ayuda Ramón y Cajal del MINECO (RYC-2012-11109).

**Referencias**

Condit, R., Lao, S., Singh, A., Esufali, S., & Dolins, S. 2014. Data and database standards for permanent forest plots in a global network. *Forest Ecology and Management* 316: 21–31.

Dengler, J., Jansen, F., Glöckler, F., Peet, R.K., De Cáceres, M., Chytrý, M., Ewald, J., Oldeland, J., Lopez-Gonzalez, G., Finckh, M., Mucina, L., Rodwell, J.S., Schaminée, J.H.J., & Spencer, N. 2011. The Global Index of Vegetation-Plot Databases (GIVD): a new resource for vegetation science. *Journal of Vegetation Science* 22: 582–597.

Wiser, S.K. 2016. Achievements and challenges in the integration, reuse and synthesis of vegetation plot data. *Journal of Vegetation Science* 27: 868–879.

Wiser, S.K., Spencer, N., De Cáceres, M., Kleikamp, M., Boyle, B., & Peet, R.K. 2011. Veg-X - an exchange standard for plot-based vegetation data. *Journal of Vegetation Science* 22: 598–609.