

## FICHA METODOLÓGICA

### Autores

María Cecilia Londoño  
Instituto Humboldt  
[mlondono@humboldt.org.co](mailto:mlondono@humboldt.org.co)

Laura Carolina Bello  
Instituto Humboldt  
[lbello@humboldt.org.co](mailto:lbello@humboldt.org.co)

Iván González  
Instituto Humboldt  
[igonzaez@humboldt.org.co](mailto:igonzaez@humboldt.org.co)

### Enlaces a recursos en línea

- GBIF  
<http://www.gbif.org/occurrence>
- SiB Colombia  
<http://data.sibcolombia.net/inicio.htm>

### Fuentes de datos utilizadas

#### *Registros biológicos*

La base de datos de registros biológicos de especies incorporada en el análisis fue obtenida a partir de consultas al Sistema de información sobre Biodiversidad de Colombia (SIB) a *Global Biodiversity Information Facility* (GBIF), descargando todos los registros que se presentaban para el área geográfica de Colombia durante el 20 y 29 de septiembre de 2013. Se obtuvo un total de 2'273.830 de registros para Colombia de los cuales 1'238.590 registros tenían coordenadas geográficas y 1'035.240 estaban sin coordenadas.

#### *Cartografía*

Para realizar los análisis de verificación y vacíos geográficos dentro del territorio colombiano se utilizaron los siguientes insumos cartográficos. Capas de departamentos, municipios, cascos urbanos y altitud de la cartografía oficial generada a escala 1:100000 (1). Para los municipios se cuenta con un conjunto de 6 shapefiles históricos que reconstruyen los cambios en la delimitación de los municipios desde 1964 hasta el 2011 (años 1964, 1973, 1985, 1993, 2003, 2011) (2). Áreas protegidas (3), Corporaciones autónomas regionales (4), Parques Nacionales Naturales (3), Biomas (5), vías (1) y ríos principales (1).

Por último, el análisis de vacíos incorporó la representatividad ambiental para lo cual se requirió de capas de variables bioclimáticas, y variables topográficas en formato raster. Las variables climáticas fueron obtenidas de WorldClim versión 1.4 (6) a resolución de

1 km<sup>2</sup>. Estas métricas son derivadas de series mensuales de temperatura y precipitación obtenidas a partir de una red mundial de estaciones meteorológicas de diversas fuentes. Estas variables climáticas son consideradas variables biológicamente significativas para la distribución de las especies. En total se contó con 19 variables bioclimáticas de las cuales 11 estaban relacionadas con la temperatura y 8 con la precipitación (6). Las 2 variables topográficas (altura y pendiente) fueron derivadas del modelo de elevación digital (DEM) a resolución de 1 km<sup>2</sup> obtenido de (6).

### Fuentes bibliográficas cifras de biodiversidad

| Grupo            | Reportado en Bibliografía | Observado en bases de datos | % conocimiento | Fuente Bibliográfica cifras   |
|------------------|---------------------------|-----------------------------|----------------|---|
| <b>Aves</b>      | <b>1889</b>               | 1837                        | 97%            | Donegan T., Quevedo A., s McMullan M., & Salaman P. 2011. Revision of the status of bird species occurring or reported in Colombia 2011 Revisión del estatus de las especies de aves que existen o han sido reportadas en Colombia 2011. <i>Conservacion Colombiana</i> 15: 4-21  |
| <b>Mamíferos</b> | <b>492</b>                | 405                         | 82%            | Solari S., Muñoz-Saba Y., Rodríguez-Mahecha, J. V., Defler T.R., Ramírez-Chaves H. y Trujillo, F. 2014. Riqueza, endemismo y conservación de los mamíferos de Colombia. <i>Mastozoología Neotropical</i> , 20(2):301-365.   |
| <b>Anfibios</b>  | 784                       | 556                         | 71%            | Frost, Darrel R. 2013. Amphibian Species of the World: an Online Reference. Version 5.6 (9 January 2013). Electronic Database accessible at <a href="http://research.amnh.org/herpetology/amphibia/index.html">http://research.amnh.org/herpetology/amphibia/index.html</a> . American Museum of Natural History, New York, |
| <b>Reptiles</b>  | <b>593</b>                | 284                         | 48%            | Castaño-Mora, O.V. (Ed.). 2002. Libro Rojo de Reptiles de Colombia. Serie Libros Rojos de Especies Amenazadas de Colombia. Bogotá, Colombia. Instituto de Ciencias Naturales Universidad Nacional de Colombia, Ministerio de Ambiente Vivienda y Desarrollo Territorial. 160 p.   |
| <b>Peces</b>     | 1796                      | 1643                        | 91%            | Mojica, J.I., C. Castellanos, S. Usma y R. Álvarez (eds.). 2002. Libro Rojo de Peces dulceacuícolas de Colombia. Serie Libros Rojos de Especies Amenazadas de Colombia. Bogotá, Colombia. Instituto de Ciencias Naturales Universidad Nacional  |

|                          |             |       |     |  |
|--------------------------|-------------|-------|-----|--|
|                          |             |       |     | de Colombia, Ministerio de Ambiente Vivienda y Desarrollo Territorial. 288 p.  |
| <b>Mariposas</b>         | <b>3274</b> | 349   | 11% | Mojica, J.I., C. Castellanos, S. Usma y R. Álvarez (eds.). 2002. Libro Rojo de Peces dulceacuícolas de Colombia. Serie Libros Rojos de Especies Amenazadas de Colombia. Bogotá, Colombia. Instituto de Ciencias Naturales Universidad Nacional de Colombia, Ministerio de Ambiente Vivienda y Desarrollo Territorial. 288 p.   |
| <b>Hormigas</b>          | <b>900</b>  | 307   | 34% | Chaves, M.E. y Santamaría, M. (eds). 2006. Informe Nacional sobre el Avance en el Conocimiento y la Información de la Biodiversidad 1998-2004. Instituto de Investigación en Recursos Biológicos Alexander von Humboldt. Bogotá D.C. Colombia. 2 Tomos.  |
| <b>Arácnidos</b>         | 1089        | 56    | 5%  | Flórez G. & Sánchez-C. H. 1997. La diversidad de los arácnidos de en Colombia, aproximación inicial en: Rangel-Ch. O. Colombia Diversidad Biótica Vol I. pp 327<br><a href="http://issuu.com/diversidadbiotica/docs/dbi.-cap19.diversidad-aracnidos?e=2165212/2622944">http://issuu.com/diversidadbiotica/docs/dbi.-cap19.diversidad-aracnidos?e=2165212/2622944</a> |
| <b>Decápodos</b>         | <b>688</b>  | 189   | 27% | Andrade-C., M. G. 2011. Estado de conocimiento de la biodiversidad en Colombia y sus amenazas. Consideraciones para fortalecer la interacción ciencia-política. Rev. Acad. Colomb. Cienc., 35(137), 491-507.   |
| <b>Abejas</b>            | <b>398</b>  | 66    | 17% | Melo G., 2007. Introductory Remarks. In Moure, J. S., Urban, D. & Melo, G. A. R. (Orgs). Catalogue of Bees (Hymenoptera, Apoidea) in the Neotropical Region. Sociedade Brasileira de Entomologia. Xiv, 1058p.  |
| <b>Plantas</b>           | 27881       | 13263 | 48% | Andrade-C., M. G. 2011. Estado de conocimiento de la biodiversidad en Colombia y sus amenazas. Consideraciones para fortalecer la interacción ciencia-política. Rev. Acad. Colomb. Cienc., 35(137), 491-507.   |
| <b>Helechos y afines</b> | <b>1641</b> | 290   | 18% | Bernal et al 2007 en: Stuessy T. F. 2007. PLANT SYSTEMATICS WORLD. TAXON 56 (1) February 2007: 273–274   |
| <b>Orquídeas</b>         | <b>4010</b> | 1286  | 32% | Sarmiento, J. 2007. La Familia Orchidaceae en Colombia. Actual Biol (Supl. 1): 84.   |

|                       |             |     |     |  |
|-----------------------|-------------|-----|-----|--|
| <b>Musgo y afines</b> | <b>1721</b> | 837 | 49% | Bernal et al 2007 en: Stuessy T. F. 2007. PLANT SYSTEMATICS WORLD. TAXON 56 (1) February 2007: 273–274           |
| <b>Palmas</b>         | <b>262</b>  | 229 | 87% | Bernal, R., Celis, M. & Gradstein, S.R. 2007. <i>Plant diversity of Colombia catalogued</i> . Taxon 56 (1): 273. |
| <b>Líquenes</b>       | <b>1377</b> | 284 | 21% | Bernal et al 2007 en: Stuessy T. F. 2007. PLANT SYSTEMATICS WORLD. TAXON 56 (1) February 2007: 273–274           |

## Descripción metodológica

### *Verificación de los registros biológicos*

Los registros biológicos fueron analizados para verificar: a) la existencia de coordenadas geográficas dentro de Colombia; b) la consistencia de la ubicación de los registros con los departamentos y municipios reportados en la base de datos, para este paso se tiene en cuenta los cambios de los límites históricos de los municipios a través de los años 1964, 1973, 1985, 1993, 2003 y 2011; c) los registros que caen en los cascos urbanos; d) la consistencia altitudinal eliminando los registros que presentan valores extremos según el método propuesto por (7); y e) eliminar los duplicados de registros que presenten la misma especie, coordenadas, fecha de registro y localidad. Por cada uno de estos pasos se generan subsets donde se depositan los registros que presentan inconsistencias. Los códigos para realizar este análisis se encuentran disponibles en: <http://code.google.com/p/verificacion-registros-biologicos/> junto con una descripción detallada. El resultado de esta etapa es una base de datos depurada con los registros que presentan consistencia geográfica.

### *Índice de vacíos geográficos*

Para el Índice de vacíos geográficos (IVG) se siguió la metodología propuesta por (8). El IVG identifica áreas carentes de información y donde el muestreo adicional mejoraría la representatividad de la base de datos, la representatividad ambiental y la complementariedad biológica. Para desarrollar este índice la unidad mínima de análisis es 10 km<sup>2</sup> y se desarrolla en 3 etapas. La primera estima la densidad de registros por celdas de 10 km<sup>2</sup>, la segunda analiza la representatividad ambiental y la tercera la representatividad taxonómica con base en la completitud de la riqueza de especies.

Para el análisis de representatividad ambiental se realizó un muestreo de puntos aleatorios de las variables ambientales y un muestreo de las variables solo para los puntos de los registros. A partir de este muestreo se construyeron las distribución de densidad para cada variables y se comparó las dos distribuciones de probabilidad con una prueba de Kolmogorov-Smirnov con la hipótesis nula que las dos distribuciones de probabilidad son iguales (9). Posteriormente cada variable se dividió en cuatro intervalos y el sesgo se calculó para cada intervalo según el índice de (10):

$$\text{SESGO} = \frac{n_d - p_d N}{\sqrt{p_d * (1 - p_d) * N}}$$

Donde  $n_d$  es en número de registros en el intervalo (d), N es el número de total de registros en la base de datos, y  $p_d$  es la probabilidad que un registro caiga en el intervalo (d).

Para la tercera etapa se usó el índice de completitud de (11). Este índice se basa en la comparación del total (estimado) de especies presentes en un área geográfica ( $S^*$ ) con el número de especies observadas en esa área, de tal forma que el índice de complementariedad C se define como:

$$C = S_{\text{obs}} - S^*$$

El cálculo de C se realiza para cada pixel de 10 km<sup>2</sup>. El número de especies estimadas se calculó con el método de Jackknife que tiene en cuenta una corrección por el sesgo de muestreo de tal forma que:

$$S^* = S_{\text{obs}} + L \frac{n - 1}{n}$$

Donde n es el número de muestras y L el número de especies que ocurren solo en una muestra (12).

A partir de los insumos generados en los pasos previos se calculó en IVG propuesto por García Márquez et al. (2012), que toma la capa de densidad estimada de registros (d), la capa de representatividad ambiental (b) y la capa de representatividad taxonómica (C). Cada uno de estas capa fueron estandarizadas de 0 a 1 siguiendo la ecuación de (13).

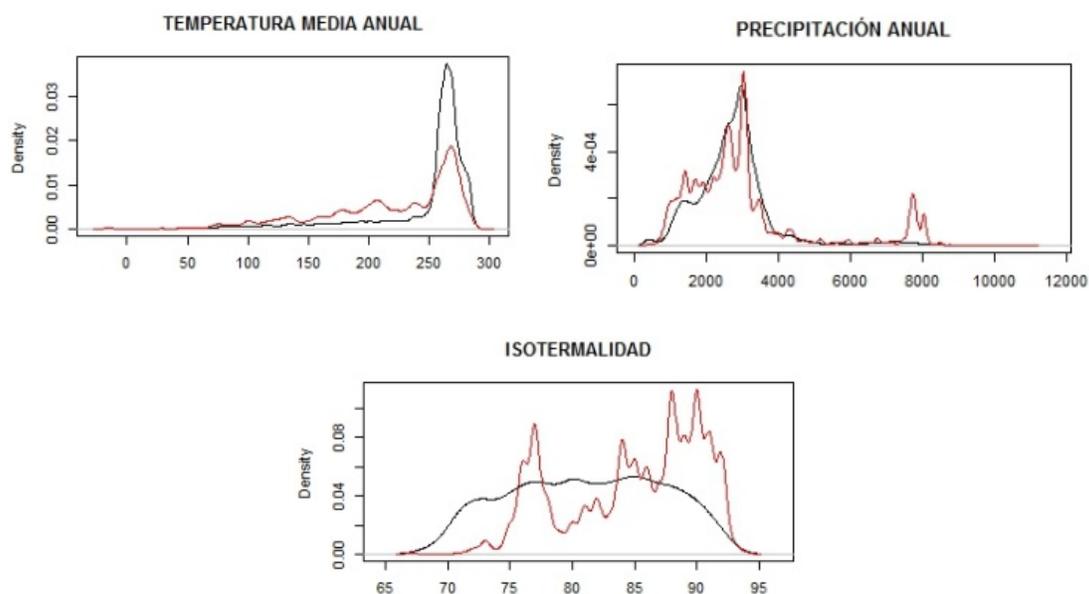
$$y' = \frac{y_i - y_{\text{min}}}{y_{\text{max}} - y_{\text{min}}}$$

Una vez las capas fueron estandarizadas se procedió a calcular el IVG de la siguiente forma:

$$\text{IVG} = \frac{3 - d' - b' - C'}{F}$$

Dónde: d' es la capa densidad estandarizada, b' es la capa de sesgo ambiental estandarizada, C' la capa de complementariedad estandarizada y F es el número de factores introducidos en el índice. Los resultados del índice están en el rango de 0 -1 donde 0 representa áreas que han sido correctamente representadas mientras que 1 representa áreas que no están bien representadas en la base de datos (8). Para este

estudio el índice IVG será modificado de la forma  $1 - IVG$  para que los valores del índice indiquen lo contrario: 1 para zonas muy bien representadas y 0 para las que no.

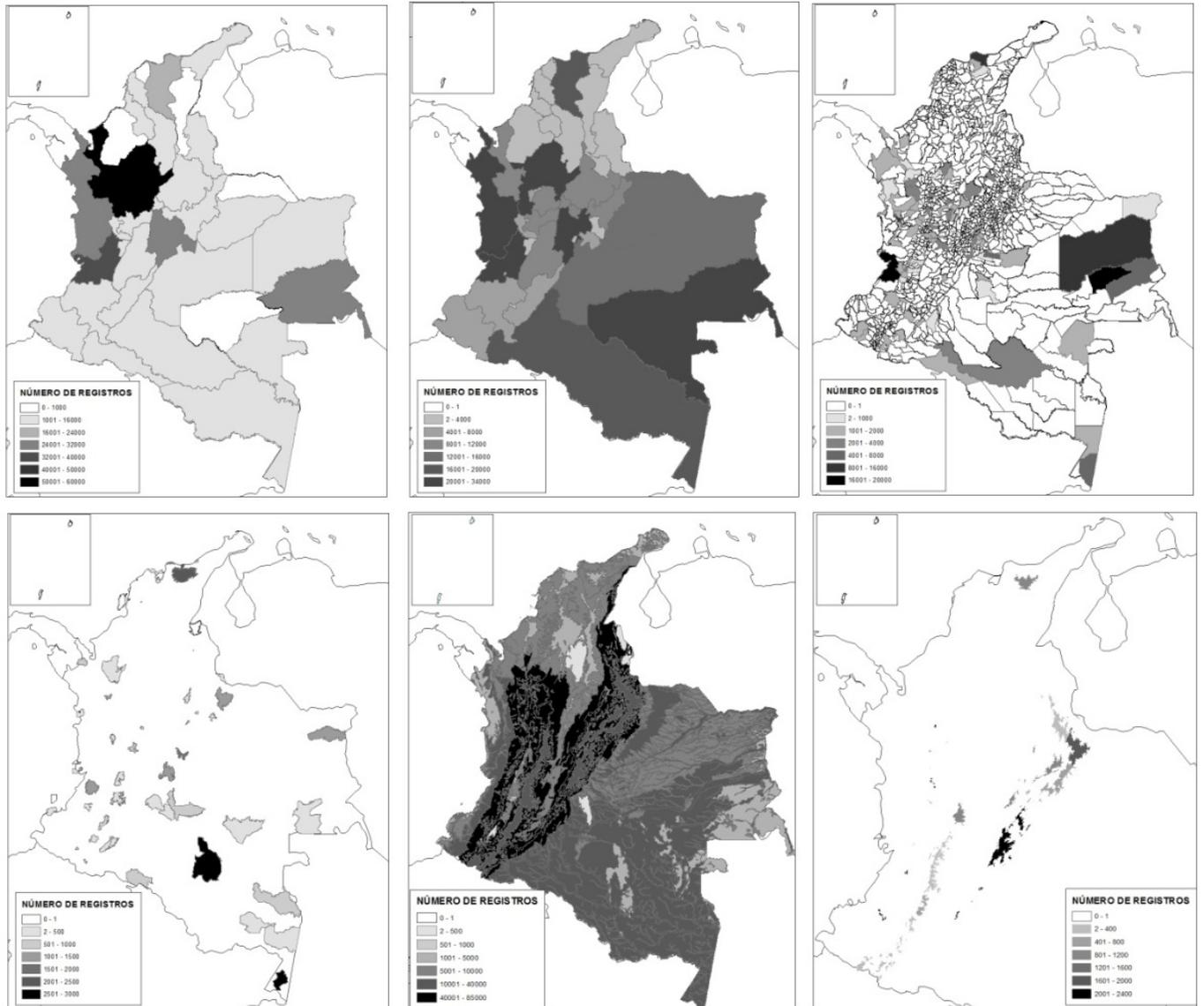


**Figura 1. Distribución de probabilidad empírica de las variables ambientales en Colombia (negro) y distribución de probabilidad de las variables ambientales representadas por los registros (rojo).** Sub-representado cuando la línea negra está por encima de la roja. Sobre-representado cuando la roja está por encima de la línea negra.

### Resultados detallados

- [Si bien Antioquia, Valle del Cauca, Cundinamarca, Chocó y Guainía son los departamentos con mayor número de registros](#), no significa que al interior de los departamentos se tenga una buena representatividad [geográfica](#), con excepción de Antioquia en donde los registros se encuentran mejor distribuidos en los municipios,; en el departamento del Valle del Cauca los registros se concentran en Buenaventura, en el Guainía se concentran en Barranco Minas e Inírida, en el Magdalena los registros provienen en su mayoría de los Parques Nacionales Naturales del Tayrona y Sierra Nevada de Santa Marta, en el Chocó y en el Amazonas más de la mitad de los registros se encuentran concentrados en los municipios del Cantón del San Pablo (Managru) y en Leticia respectivamente.
- Dentro de los grupos taxonómicos evaluados los invertebrados, con excepción de las abejas, los líquenes y helechos son los menos representados en las plataformas en línea. Mayores esfuerzos se deben realizar para incluir un mayor número de datos sobre estos grupos taxonómicos.
- El 46% de registros no puede utilizarse porque no cuenta con georreferencias, este porcentaje puede reducirse para potenciar el uso de la información ya obtenida, en análisis espaciales.

- Del total de datos con georreferencias, el 22,6% pasan todos los filtros de verificación (con información y consistencia geográfica y altitudinal). Las principales limitaciones se encuentran al contrastar la ubicación de las coordenadas geográficas con los polígonos municipales, bien sea porque son inconsistentes o porque no se reportan los datos de municipio en el registro biológico disponible en línea. La información general asociada a la localidad del registro es valiosa para los procesos de verificación de la información, así como en posteriores actividades de complementación de información geográfica y georreferenciación.
- Los registros no están distribuidos homogéneamente en el territorio Colombiano. Los departamentos de Atlántico, Arauca, Guaviare, Cesar y Córdoba cuentan con el menor número de datos que resultan verificados. Del total de 1.123 municipios de Colombia tan solo el 9% presentan registros biológicos verificados, el restante 91% de los municipios no están representados en los resultados depurados (Figura 2c), evidenciando los serios problemas de agregación de los registros a localidades puntuales. Incrementar la información por unidades político-administrativas es importante para poder usar información de biodiversidad en los procesos de ordenamiento territorial.
- Tres biomas no cuentan con información en las bases de datos en línea: Heliobioma del río Zulia, Bioma insular del Caribe y Bioma insular del Pacífico (Figura 2e). A pesar de que los biomas de los Andes están bien representados, no se cuenta con información disponible en línea para los páramos de la Cordillera Occidental (Cerro plateado, Citará, Farallones de Cali y Paramillo), y algunos de la Cordillera Oriental (Miraflores, Los picachos y Yariguíes) (Figura 2f). Esta situación se prevé que mejore, especialmente dados los esfuerzos que realiza el país para obtener información que le permita abordar la delimitación de ecosistemas estratégicos.
- Los Parques Nacionales Naturales, por tratarse de las áreas destinadas para la conservación de los recursos naturales, suelen ser lugares estratégicos para el estudio y registro de información biológica. Sin embargo 7 de las 56 áreas protegidas nacionales no están representadas en las bases de datos en línea de registros biológicos (PNN Corales del Rosario, SFF Malpelo, PNN Gorgona, PNN Old Providence Mc Bean Lagoon, SFF El Chorrón, PNN Catatumbo Bari, SFF Isla de la Corota). La distribución agregada de los registros en el territorio colombiano tiene efectos en la representación del gradiente ambiental del territorio. En todas las variables se observa que la distribución de probabilidades de las variables representadas no corresponde a la distribución de probabilidad real de las variables en Colombia ( $D > 0.12$  y  $p < 0.05$  para todos los casos). Las zonas de mayor altitud están sobrerrepresentadas y las zonas bajas subrepresentadas. Esto ocasiona el patrón que se observa para el gradiente de temperatura e isothermalidad, ya que las altas temperaturas de las zonas bajas donde la temperatura es más constante a lo largo del día se encuentran subrepresentadas. Tener una buena representatividad del gradiente ambiental del territorio colombiano permite fortalecer el análisis de la biodiversidad a partir de información climática y topográfica.



**Figura 2. Número de registros por unidad geográfica de análisis. a. Departamentos, b. Corporaciones Autónomas regionales, c. Municipios, d. Áreas protegidas de carácter nacional, e. Biomas, f. Complejos de páramo de Colombia**

### Literatura asociada

- (1) IGAC .2012. Base cartográfica integrada a escala 1:100000. Bogotá. Colombia.
- (2) IGAC (s.f.).Limite municipal de los años 1964, 1973, 1985, 1993, 2003, 2011. Bogotá. Colombia.
- (3) Parques Nacionales Naturales. 2012. Registro único nacional de áreas protegidas. Bogotá D.C. Colombia.

- (4) IGAC (Ajustado por IAVH). 2010. Corporaciones Autónomas Regionales y Autoridades Ambientales de Grandes Centros Urbanos. Ajuste realizado con base a información tomada del IGAC, Asocars, Corporaciones Autónomas Regionales y Autoridades Ambientales de Grandes Centros Urbanos. Bogotá D.C. Colombia.
- (5) IDEAM, IGAC, IAvH, Invemar, I. Sinchi e IIAP. 2007. Ecosistemas continentales, costeros y marinos de Colombia. Instituto de Hidrología, Meteorología y Estudios Ambientales, Instituto Geográfico Agustín Codazzi, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Instituto de Investigaciones Ambientales del Pacífico Jhon von Neumann, Instituto de Investigaciones Marinas y Costeras José Benito Vives De Andrés e Instituto Amazónico de Investigaciones Científicas Sinchi. Bogotá, D. C, 276 p. + 37 hojas cartográficas.
- (6) Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25:1965-1978.
- (7) Iglewicz, B. H., D. 1993. How to Detect and Handle Outliers.
- (8) Garcia Márquez, J. R. G., C. F. Dormann, J. H. Sommer, M. Schmidt, A. Thiombiano, S. S. Da, C. Chatelain, S. Dressler, and W. Barthlott. 2012. A methodological framework to quantify the spatial quality of biological databases. *Biodiversity and Ecology* 4:25-39.
- (9) Abadie, Alberto. 2002. "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models." *Journal of the American Statistical Association*, 97:457 (March) 284-292.
- (10) Kadmon, R., O. Farber, and A. Danin. 2004. Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications* 14:401-413.
- (11) Soberón, J.M., Llorente, J.B., Oñate, L. 2000. The use of specimen-label databases for conservation purposes: an example using Mexican papilionid and pierid butterflies. – *Biodiversity and Conservation* 9: 1441–1466.
- (12) Burnham, K.P., Overton, W.S. 1979. Robust estimation of population size when capture probabilities vary among animals. – *Ecology* 60: 927–936.
- (13) Legendre, P., Legendre, L. 1998. *Numerical Ecology*. 2nd ed. – Amsterdam: Elsevier Science.

### **Para tener en cuenta**

Este análisis está basado solo en los registros biológicos que están disponibles en línea, por lo que sus conclusiones solo son aplicables para este tipo de información y no refleja el conocimiento general de todo el país. En Colombia se han realizado grandes esfuerzos en el conocimiento de la diversidad biológica que aún no están disponibles para su consulta en línea, por lo cual no fueron incluidos en este análisis.

Los estimadores de riqueza aumentan su robustez si se tiene un esfuerzo de muestreo documentado que se incorpore en el cálculo, o si el número de registros para un área es suficiente. Dado que no se cuenta con información sobre el esfuerzo de muestreo y para muchas áreas el número de registros es muy bajo, los estimadores de riqueza deben ser interpretados con cautela entendiendo que es posible que el IVG cambie ligeramente al incorporarse una mejor técnica para calcular la representatividad de la riqueza, o cuando esta se pueda hacer con base en listados taxonómicos compilados.

### **Vacios de información**

El vacío más relevante es la dificultad para encontrar la información sobre listados de especies para Colombia por unidades administrativas, políticas, o biomas, pues está dispersa y requiere un gran esfuerzo de organización y síntesis de información a partir de referencias bibliográficas. Esta información resulta crucial para poder aproximarse de una manera más robusta a la representatividad de los datos con respecto a la riqueza taxonómica del país.

Otro tema relevante es reconocer que las bases de datos en línea aun no recopilan todo el conocimiento de la diversidad. En especial se requiere de esfuerzos interinstitucionales para visibilizar la información existente de regiones como la Amazonia, ya que esta zona ha sido ampliamente estudiada pero la información georreferenciada asociada a los registros biológicos no se encuentra disponible en línea. Además, es necesario realizar un esfuerzo de georreferenciación de los registros disponibles ya que de la información consultada para Colombia solo el 54% presenta coordenadas.