

A web-based GIS tool for exploring the world's biodiversity: The Global Biodiversity Information Facility Mapping and Analysis Portal Application (GBIF-MAPA)

Group 8: Ewelina Kwiatkowska, Katharina Kunstman,
Jan Schwalb and Malte Simon

Keywords:

Biodiversity mapping, Online Geographic Information Systems, Species richness, Survey gap analysis, Global Biodiversity Information Facility

1. Introduction

„Overall, we are locked into a race. We must hurry to acquire the knowledge on which a wise policy of conservation and development can be based for centuries to come“. E.O. Wilson (1988)

In the quote, Wilson (1988) urges the global biodiversity community to develop a wise set of policies for biological conservation.

- Wisdom: is the top portion of a hierarchy, includes data, information and knowledge
- In order to achieve wisdom, one must first have data — raw observations and measurements.

The raw data formuch of biodiversity research are specimen occurrence records — when and where species are found.

- GBIF (Global Biodiversity Information Facility)- has developed an infrastructure by which museums and herbaria can publish their databases to a global network of biodiversity data.

As of March 2007, provides GBIF access over the Internet to approximately 120 million species occurrence records from over 1000 separate collections.

- EstimateS provide estimations of species richness

- DIVA-GIS provide GIS capabilities and a suite of biodiversity analysis modules

- GBIF-MAPA- a web-based biodiversity workflow tool that provides users the means to semi-automate raw biodiversity data acquisition, geospatial visualization and deployment of core biodiversity analyses

2. Case studies

- The point of developing GBIF-MAPA was to provide a means for users to select any region in the world and ask core biodiversity research questions like:
 - Is there enough existing biodiversity data to determine accurate measures of species richness?
 - “Where is the most likely spot to survey for more biodiversity given the current species occurrences and environmental conditions?”
- Case study 1: **Survey Gap Analysis tool** (SGA) = Best location for field survey depending on previous study sites.
- Case study 2: **Species Richness Analysis tool** (SRA) = Find the best sites with high species richness.
- Case study criteria:
 - Criteria 1: The regions and taxa were not ones with which we were intimately familiar.
 - Criteria 2: The regions are known to be hotspots for biodiversity.

(FLEMONS et al. 2007)

Case study 1: „Where to look for new frog and toad species on Madagascar?”

- The amphibian Order Anura (frogs and toads) was chosen for this case study because there are only a few Anura records from the GBIF data cache on mainland Madagascar.
- This distribution suggests that there may be undiscovered anuran species on Madagascar (Case study criteria 1: Taxa not well researched)
- We used the Survey Gap Analysis tool to determine best survey locations on Madagascar for anurans

General workflow

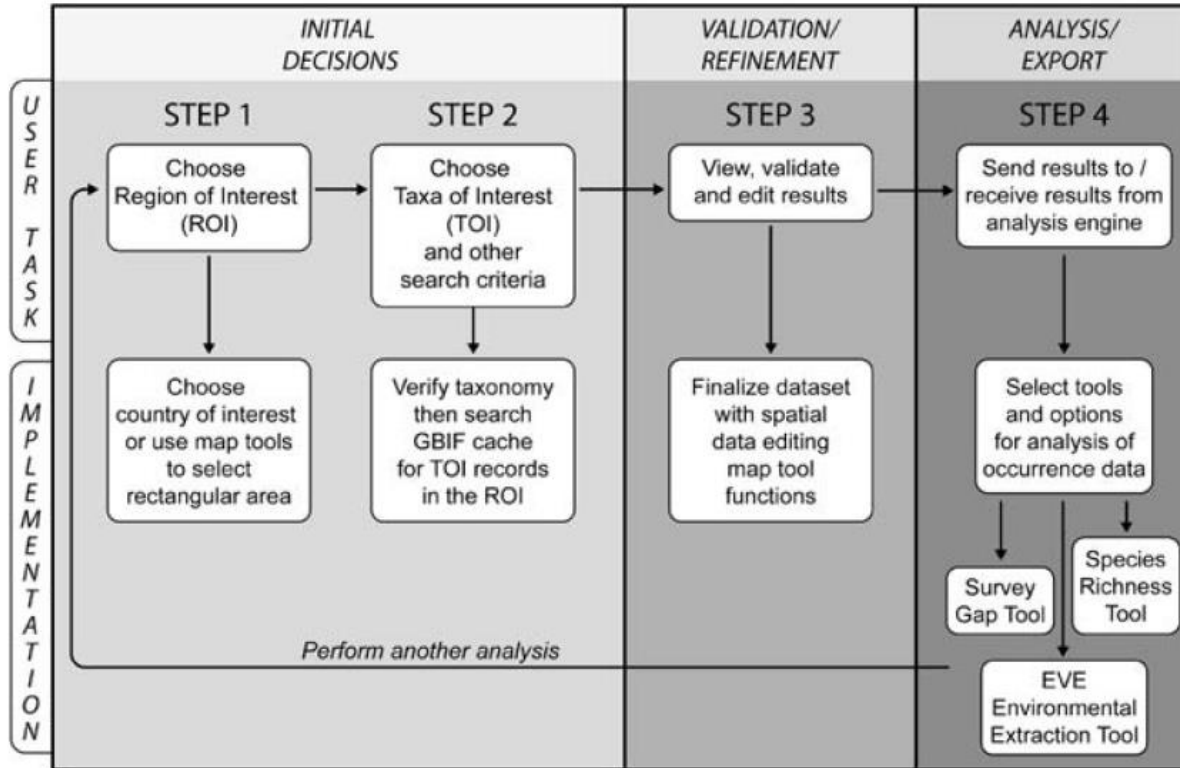
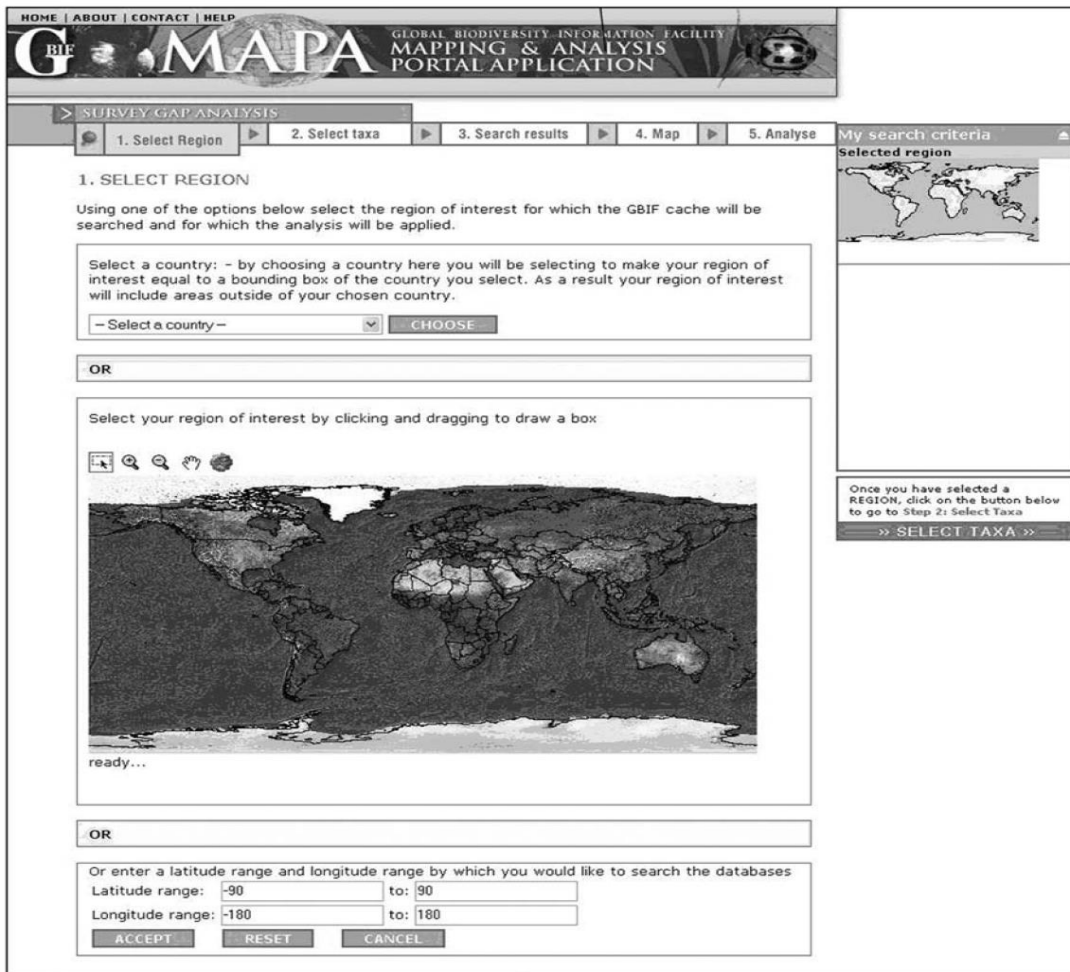


Fig. 1 – Diagram showing the user tasks and how they are implemented in the Global Biodiversity Information Facility Mapping and Analysis Portal. The tasks include: making initial decisions about the region of interest and taxon of interest; validating and refining the record set by mapping and checking records; performing one of three biodiversity analyses (species richness assessment, survey gap analysis and environmental extraction) and exporting results.



Application Interface

Fig. 2–GBIF-MAPA interface showing GUI design methods for maximizing user certainty. At the top of the screen under the GBIF-MAPA banner is the “process flow bar” and on the right of the page is the “search criteria summary box”.

(FLEMONS et al. 2007)

Workflow for case study 1

I. Initial decisions:

Step 1: ROI = Madagascar

Step 2: TOI = Anura

II. Validation/Refinement:

Step 3: View and validate results

→ Delete anura taxa occurring in the sea

III. Analysis/Export

Step 4: Tool selection = Survey Gap Analysis

→ Data extraction (Best study sites to search for anura)

Results of Survey Gap Analysis

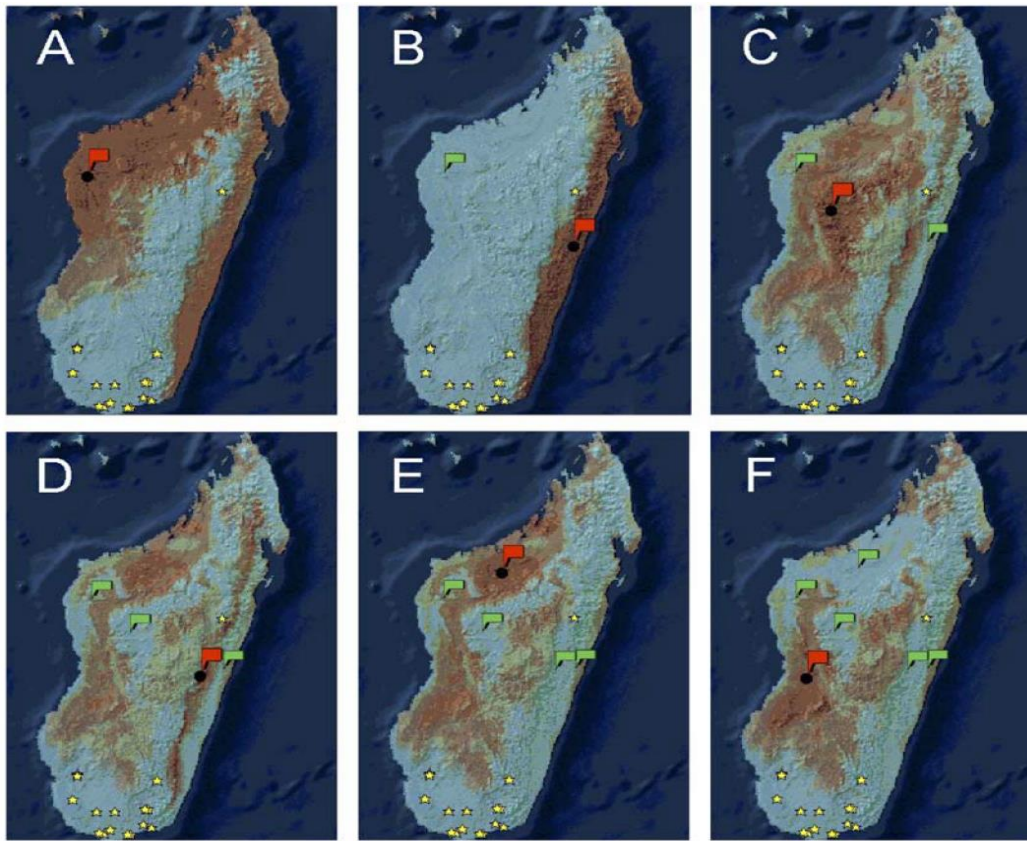


Fig. 4—Panel of maps showing the results of the iterative selection of points by the Survey Gap Analysis algorithm. The colour gradation from blue to brown indicates areas of low complementarity to high complementarity. The red flags represent the optimal survey location in terms of most complementing existing survey effort (yellow stars) B to F — addition of one new site per iteration, each iteration identifies the next highest site of complementarity. Each green flag represents a previously chosen new survey location. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(FLEMONS et al. 2007)

Case study 2: “Can Ethiopian rodent species richness be estimated using GBIF data?”

- The case study for species richness assessment is rodent richness in the country Ethiopia.
- The country comprises the Ethiopian Highlands, which is one of thirty four global biodiversity hotspots. (Case study criteria 2: High biodiversity)
- It is also an area of global conservation concern given high species richness and number of endemics, and given increasing and dramatic human impact on the area.

Workflow for case study 2

I. Initial decisions:

Step 1: ROI = Ethiopia

Step 2: TOI = Rodentia

II. Validation/Refinement:

Step 3: View and validate results (delete unfitting entries)

→ 1053 specimen occurrence records

→ delete uncompleted entries

III. Analysis/Export

Step 4: Tool selection = Species Richness Analysis

→ Data extraction

> SPECIES RICHNESS ASSESSMENT

☒ 1. Select Region
 ☒ 2. Select taxa
 ☒ 3. Search results
 ☒ 4. Map
  5. Analyse

Summary data

NAME	VALUE
Total number of species observed in all samples pooled:	63.
value of rare species (10 or fewer individuals):	44.
value of abundant species (more than 10 individuals):	19.
value of infrequent species (found in 10 or fewer samples):	61.
value of frequent species (found in more than 10 samples):	2.
Total number of samples:	100
value of samples with at least one infrequent species:	27.
value of singletons:	9.
value of doubletons:	12.
value of uniques:	30.
value of duplicates:	10.
Total number of individuals in rare species:	157.
Total number of occurrences of infrequent species:	156.
Sample abundance coverage estimator:	0.942675159235667
Sample incidence coverage estimator:	0.8076923076923069

Tests of assumption

No statistically significant trend was detected in the number of records per unit effort (percent change = 142.373%, $p = 0.344981$). This is good!

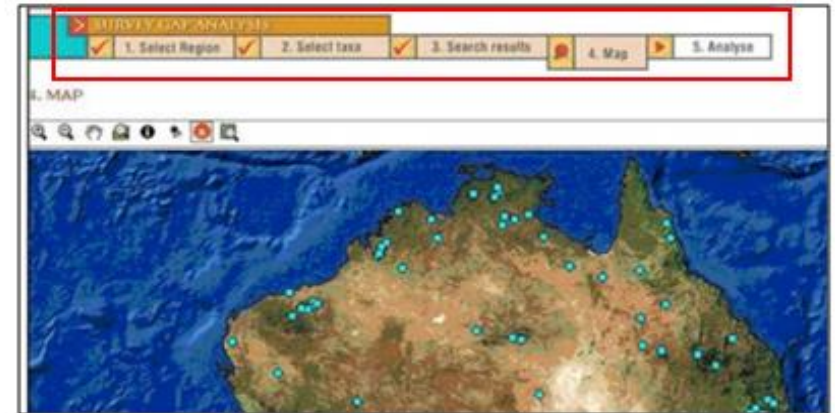
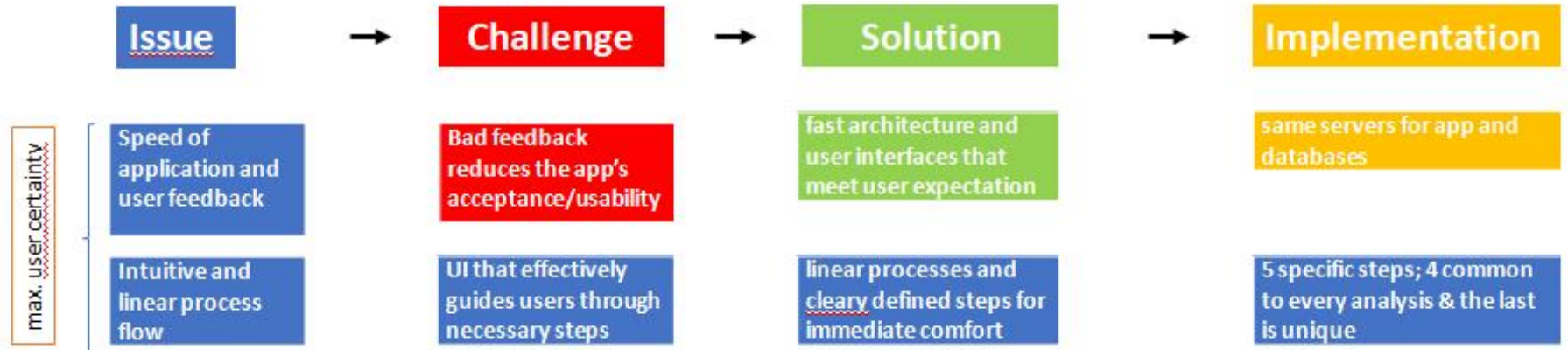
Canonical correspondence analysis of the data indicates no substantial, significant change in community composition over the sequence of samples (percent variance explained = 1.95388%, $p = 0.59$). This is good!

Species Richness Analysis

Fig. 5 – Tabular summary output from the species richness analysis (SRA) of rodents in Ethiopia as generated from GBIF-MAPA. After validation, there were 998 usable species occurrence data points from the GBIF data cache. The summary data is part of a much larger set of output from SRA.

(FLEMONS et al. 2007)

GENERAL TECHNICAL ISSUES



TECHNICAL ISSUES OF THE WORKFLOW

Step	Challenge	Solution	Implementation
1 ROI	ability to select a meaningful ROI	3 approaches: 1. Database field 2. Bounding box 3. Polygonal bounding area	-
2 TOI	Large amount of data, exponentially increasing loading times	Local JDBC (Java Database Connectivity)	MAPA app is installed directly into GBIF portals
3 <u>map/validate species occurrences</u>	effectively view multiple taxa on multiple base layers	allow user selection of symbology view or hide different <u>basemap</u> layers	e. g. drop-down list of symbols for each taxon
4 <u>Data analysis & results</u>	Select tools that are needed & simple to implement	select tools that are already in use & well known among the biodiversity researchers	implement two main tools (SRA & SGA)

(FLEMONS et al. 2007)

a. What were the main questions (or objectives) addressed?

- Supporting research, conservation management, and education with high quality global diversity data
- How to build tools which use the GBIF data portal, giving access to researchers to allow to use the whole spectrum of available data

b. How did the authors tackle them? (what did they do?)

Building a tool to perform biodiversity analysis with a end-to-end workflow

- Three types of analyses:
 - environmental extraction
 - survey gap analysis
 - species richness assessment
- Focusing on:
 - user-friendly interface
 - speed
 - flexibility in designing output

c. What aspects could be added/done differently/problems?

- Paper was published in 2007, outdated
- MAPA webpage is not available, only reference screenshots in the paper
- No possibility to try it out, hard to visualize technical issues and service
- No information how gbif.org looked like back then
- In second case study, incomprehensible what quality the output has

Get data How-to Tools Community About

GBIF | Global Biodiversity Information Facility

Free and open access to biodiversity data


OCCURRENCES SPECIES DATASETS PUBLISHERS RESOURCES

Search

WHAT IS GBIF? ABOUT GBIF GERMANY


Link to GBIF records observed in Saint-Benoît, Réunion by Vincent Poirier, Photo via Naturalist (CC BY-NC)

Occurrence records 1.410.166.702	Datasets 52.491	Publishing institutions 1.596	Peer-reviewed papers using data 4.429
-------------------------------------	--------------------	----------------------------------	--




News

Call for data papers from European Russia




News

Virtual workshop: Advancing the Catalogue of the World's Natural History Collections




News

2020 Ebbe Nielsen Challenge seeks open-data innovations for biodiversity




News


Call for nominations to the 2020 GBIF Young Researchers Award



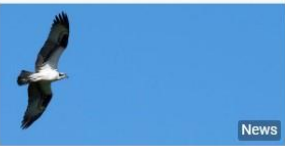
Taxonomy



Data



Data use



News

- Able to search after species, occurrences, data sets
- Tables, pictures of occurrences

(GBIF.ORG 2020)

Get data How-to Tools Community About

Classification

Select a species

Kingdom Animalia

Phylum Chordata

Class Mammalia

Order Carnivora

Family Felidae

Genus *Felis* Linnaeus, 1758

Species *Felis silvestris* Schreber, 1777

Immediate children

Subspecies *Felis silvestris* subsp. *cafra* Desmarest, 1822

Subspecies *Felis silvestris* subsp. *caucasica* Satunin, 1905

Subspecies *Felis silvestris* subsp. *caudata* Gray, 1874

Subspecies *Felis silvestris* subsp. *chutuchta* Birula, 1916

Subspecies *Felis silvestris* subsp. *cretensis* Haltenorth, 1953

Subspecies *Felis silvestris* subsp. *euxina* Pocock, 1943

Subspecies *Felis silvestris* subsp. *foxi* Pocock, 1944

SPECIES | ACCEPTED


Felis silvestris Schreber, 1777

Published in: Schreber, Johann C. D. von. 1777. Die Säugethiere in Abbildungen nach der Natur, mit Beschreibungen. Wolfgang Walther, Erlangen. Vol. 3. source: Catalogue of Life


Wildkatze In German

OVERVIEW 1 TREATMENT METRICS REFERENCE TAXON ⇌ 23,501 OCCURRENCES 26 INFRASPECIES

929 OCCURRENCES WITH IMAGES



21,366 GEOREFERENCED RECORDS



Generated an hour ago © OpenStreetMap contributors, © OpenMapTiles, GBIF

Any year 1620 - 2020 EXPLORE


RECORDED AS INTRODUCED IN 7 COUNTRIES AND ISLANDS

- Gbif has an interactive map
- Able to look into particular regions

Community About

SEARCH OCCURRENCES | 2,268 WITH COORDINATES

TABLE GALLERY MAP TAXONOMY METRICS DOWNLOAD



France

Andorra

Portugal

Spain

© OpenStreetMap contributors, © OpenMapTiles, GBIF

SEARCH RESOURCES | 110 RESULTS

ALL NEWS DATA USE EVENTS PROJECTS PROGRAMMES TOOLS DOCUMENTS LITERATURE

Georeferencing Calculator Tool

The Georeferencing Calculator is a tool created to aid in the georeferencing of descriptive localities such as those found in museum-based natural history collections. It was originally designed for t...

Frictionless Darwin Core Tool

DarwinCore Archive as Frictionless Data Package ... A tool converting Darwin Core Archive into Frictionless Data Package. DarwinCore standard, created and maintained by Biodiversity Informatics Standard...

BRAHMS – Management of Natural History Tool

Flexible database management software for preserved and living natural history collections, taxonomic research and biogeographic study, providing wide-ranging and innovative functionality to assemble...

Online capacity self-assessment tool for national biodiversity information facilities Tool

Capacity self-assessment is a tool to help countries with planning at any stage in the development of their national biodiversity information facility. This online survey is a tool to assess capacities...

Sequence ID Tool

This tool will query taxonomy reference databases for Fungi or Animals ITS sequences for Fungi will be queried against the UNITE Species Hypothesis General FASTA release, a reference database for DNA ...

Relative observation trends Tool

A pilot that explores whether relative observation trends for species within GBIF provide useful information for researchers ... This pilot tool provides users with a rapid assessment of whether the tr...

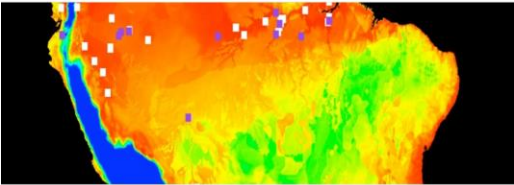
FILTERS

TOOL | 27 MARCH 2017

Maxent

Software for modeling species niches and distributions

[HTTP://BIODIVERSITYINFORMATICS.AMNH.ORG/OPEN_SOURCE/MAXENT/](http://biodiversityinformatics.amnh.org/open_source/maxent/)



The Maxent software is based on the maximum-entropy approach for modeling species niches and distributions. From a set of environmental (e.g., climatic) grids and georeferenced occurrence localities (e.g. mediated by GBIF), the model expresses a probability distribution where each grid cell has a predicted suitability of conditions for the species. Under particular assumptions about the input data and biological sampling efforts that led to occurrence records, the output can be interpreted as predicted probability of presence (cloglog transform), or as predicted local abundance (raw exponential output).


Species models are determined from a set of environmental or climate layers (or "coverages") for a set of grid cells in a landscape, together with a set of sample locations where the species has been observed. The model expresses the suitability of each grid cell as a function of the environmental variables at that grid cell. A high value of the function at a particular grid cell indicates that the grid cell

TOOL | 10 FEBRUARY 2015

Biodiverse tool

Biodiverse is a tool for the spatial analysis of diversity using indices based on taxonomic, phylogenetic and matrix-based (e.g. genetic distances) relationships, as well as related environmental and temporal variations. Biodiverse supports four processes: (1) linked visualization of data distributions in geographic, taxonomic, phylogenetic and matrix space; (2) spatial moving window analysis including richness, endemism, phylogenetic diversity and beta diversity; (3) spatially constrained agglomerative cluster analysis; and (4) randomizations for hypothesis testing. Biodiverse is open-source and supports user designed extensions. It can be used both through a graphical user interface (GUI) and through user-written scripts.

[HTTP://WWW.PURL.ORG/BIODIVERSE](http://www.purl.org/biodiverse)



Thumbnail image for tool

Citation

Laffan, S.W., Loder, E. & Rosauer, D.F. (2015) Biodiverse, a tool for the spatial analysis of biological and related diversity. *Ecography* Vol. 33, 643-647 (version number).

- A lot of tools for cleaning and validating data
- Now more tools for SDM and predicting survey locations: MaxEnt, Biodiverse, Rpackages, OpenModeller, etc.

References

FLEMONS, P., GURALNICK, R., KRIEGER, J., RANIPETA, A. & NEUFELD, D. (2007): A web-based GIS tool for exploring the world's biodiversity: the global biodiversity information facility mapping and analysis portal application (GBIF-MAPA). *Ecological Informatics*, **2**, 49– 60.