GBIF Data Modeling Workshop

Extrapolation of points to areas based on similarities

Course syllabus

4-8 April 2005,
CONABIO, Mexico City

1. Biodiversity and Biogeography

Here, the general scenario of biodiversity is reviewed. We will cover topics such as the dimensions of biological diversity (how many species?), and its general distribution. We will also discuss some of the studies that have been carried out to date, speaking to what has been learned and what remains *to be learned.*

- Dimensions of biological diversity
- How many species are there?
- Where are species distributed?
- Which general factors stand out in the distribution of species and the distribution of biological diversity?

2. Informatics of biodiversity

In order to learn more about biodiversity, we must marshal the information resources that we have available. This task leads not only to improving the storehouse of information available to us, but also to more educated efforts to improve and build further that storehouse without duplication of effort. As such, an emerging field is biodiversity informatics is introduced here.

- What is this new field about?
- What are its principal components and advances?
  - -Computerization of data
  - -Data integration
  - -Data analysis

3. Biological Problem

Here, we begin to treat ideas related to the intersection of biodiversity science with geography. That is, we speak to issues of why species are found where they are found, and why they are not found where they are not found. The answer to these questions is key in understanding global patterns of biodiversity.

- Matters related to biodiversity and its distribution in space and time
- Theory of the ecological niche
- Aerography (includes scales)
- Niche applied to the areas of distribution

4. Sources of Data

Here, a brief review is provided of existing biodiversity data resources. This will include demos of most of the existing distributed biodiversity information networks, as well as review of sources of geospatial data and resources for geo-referencing biodiversity data.

Sources of data (GBIF, Remib, TSA, Manis, Herpnet, Fishnet, BBS, E-bird...)
Sources of environmental data (Worldclim, Ipcc, hidro1K, CONABIO, AVHRR and Modis...)
Preparation of biological data 1.5 hours (includes practice),
Quality Control:

- Biological: Geographical and taxonomic bias in sampling mammals
- Gazeteers (participants may bring their own gazeteers)

5. Inferential Modeling

Here, we proceed to the question of how one can fill in the gaps in existing sampling of biodiversity. The idea is the use of what we term ecological niche modeling, and can be carried out via numerous different inferential algorithms.

- Basics of inferential or predictive modeling
- Brief discussion of genetic algorithms
- Prediction models, uses and applications (GARP, Domain, Floramap...)
- Statistics and selection of models

6. What is GARP and what does it do?

We present one inferential algorithm, GARP, in greater detail…

The focus of a GARP prediction:

7. Case studies:

Here, we present a series of applications to which ecological niche modeling has been used. The idea is to illustrate a fairly broad swath of potential applications … from terrestrial to aquatic, across space and time, etc.

** Predicting geographic distributions in space:*
Planning how to sample (use of databases)
Biogeography: (sister species)
Macroecological patterns
Prediction of new species

** Evaluating distributions in ecological space and ecological characteristics of species*
Ecology of niche evolution

** Predicting in space and time*
Dengue fever disease.

** Predicting potential distributions*
Invasive species

** Predicting scenarios of change*
Climate change

** Detecting species interactions*
Various possibilities

** Applications in conservation*
Deforestation and distribution of species
The design of ANP networks based on the distribution of species

** Applications in matters of economic importance*
Health: Chagas in Mexico- chagas clan
Agriculture: Impact of rodent pests on Mexican crops.
Genetically modified organisms
Sustainable use (certification of the origin of agave cacti)

8. Desktop GARP: GARP interface for PC
Review of the use of the GARP interface, which will be the principal (but not exclusive) focus of class exercises.

- Installation requirements
- Installation
- Using Desktop Garp
- Entering points
- Optimization parameters
- Environmental data layers
    - Selection of environmental data layers
- Exit parameters
- Projections
- Creation of data layers set with Dataset Manager
- Results
    - Prediction maps
    - Results table
- Errors in the algorithm
    - How can one test if a model is good and what is known about this/these tests?

Note: If Garp3 is completely developed by these dates, that version will be used. Otherwise, there will be an introduction to the version.

9. Your own cases: 8 hours (you can bring your data and layers or download them)

Participants will be encouraged to bring their own 'challenges' and develop their own products based on skills learned in class. By consulting in advance with participants, we hope to make these analyses possible (see relevant document).