



Project no. 018340

Project acronym: EDIT

**Project title: Toward the European Distributed Institute of
Taxonomy**

Instrument: Network of Excellence

Thematic Priority: Sub-Priority 1.1.6.3: "Global Change and Ecosystems"

**C5.110 Design document for a prototype
software for data completeness
assessment (gap analysis) for selected
test data (e.g. ATBI and other datasets)**

Due date of Component: Month 49

Actual submission date: Month 49

Start date of project: 01/03/2006

Duration: 5 years

Organization name of lead contractor for this component: CSIC

Revision final

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

This report describes the methods of inventory completeness assessment developed in the EDIT geoplatform tools and presents a spatial analysis based on ATBI data. It has been compiled by Marcin Gašior (MIIZPAN) with a generous input from Pere Roca Ristol (CSIC), Alexander Kroupa (MFN), Patricia Mergen (RMCA) and the colleagues from the EDIT developer mailing list.

Introduction

The incompleteness of species inventories due to limited field sampling is considered as an important issue in many biodiversity studies. Biodiversity measures are often biased by undersampling and cannot provide a correct assessment of biodiversity patterns. Moreover, modelling methods cannot supply reliable and useful distribution predictions if they are based on biased and insufficient data.

As species distribution information may be insufficient for some taxonomic groups, GIS tools should provide additional functions to facilitate assessing inventory completeness. Those tools allow to:

1. examine the degree of completeness of distributional information,
2. discriminate well surveyed localities from those that do not have reliable inventories,
3. identify sets of areas where to carry out additional surveys, in order to increase the level of coverage of the environmental and spatial variation of a given region.

More info about requirements to assess inventory completeness can be found at: http://wp5.e-taxonomy.eu/blog/files_edit_wp5/2007-07-26_D5.35_&_D5.38.doc.

Above mentioned requirements were the basis to develop an essential gap analysis tools in the mapViewer web application on the EDIT geoplatform. This report presents how those tools work with a sample ATBI+M (All Taxa Biodiversity Inventory and Monitoring) data from Mercantour/Alpi Maritime area (<http://www.atbi.eu/mercantour-maritime/?q=node/481>).

Data completeness assessment tools developed on the EDIT geoplatform

A set of additional functions (*Spatial analysis*) can be enabled in the *Add/Remove modules* section after user data has been uploaded. This data must be in simple .csv (comma separated values) format and have at least latitude and longitude in WGS84 datum and a field (genus, species).

The datasets collected during ATBI+M activities contain diverse formats of geographic information although it is stored in latitude/longitude notation. That is why the data usually requires some effort to convert it to a generic form (decimal notation), the only format accepted by the mapViewer (the sample data already had a proper format).

Spatial analysis can be performed for different selectable GIS layers of surface units (see *Fig. 1.*):

- *Europe level 1* (administrative units) (*see Map 1.*);
- *0.5-, 1-, 2-, 5-, 10- degrees* quadricules (*see Maps 2 - 5. for 0.5 degrees grid*);
- *UTM World/ Europe grids* (too wide range with reference to the sample data).

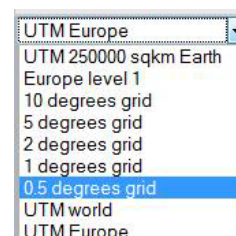


Fig. 1. Surface units

Launching an analysis results in calculation of 3 area statistics. This figures are visualized as thematic maps based on predefined aerial units – in this case the polygonal layer selected by the user. The corresponding legend of interval classes for each statistic is also created. Screenshots of the maps are shown in an annex. Those maps are as follow:

1. Map of sampling efforts (see Map 2.):

Number of records in each spatial unit – gives an idea of the collecting efforts.

2. Map of taxonomic richness (see Map 3.):

Number of genera in each spatial unit – gives some idea of the biodiversity.

3. Map of inventory uncertainty (see Map 4.):

Ratio richness/effort (number of genera/number of records) per spatial unit – gives a more accurate idea of the biodiversity than number of genera in spatial unit. This ratio allows to distinguish the polygons with the same number of genera but different number of collected records. A high ratio genera/records indicates a higher possibility to find new genera in case of collecting continuation. A low ratio signifies low biodiversity because the collecting effort has been high but only a few different genus have been found.

The map of inventory uncertainty illustrates the surface units where it is desirable to carry on additional surveys in order to collect specimens of new genera and recover the spatial variation of the area, or where data on absences should be recorded.

Interactive analysis tool (see Map 5.)

When the analysis has been performed, the *Interactive analysis tools* section is activated. It enables to display of all area statistics after highlighting some area from selected layer.

Other methods for inventory completeness assessment

Besides the methods developed in the mapViewer, there are number of other, more sophisticated methods. The most popular are a rarefaction curve generated from long-term sampling and an asymptotic fit of the cumulative number of recorded species which is used to estimate the approximate number of species not yet described. Due to their complexity and sometimes extensive computational needs, it would be very complicated to implement them in a webGIS application. Moreover, there are specialized desktop applications, developed for years, implementing those methods (e.g. EstimateS for the first method).

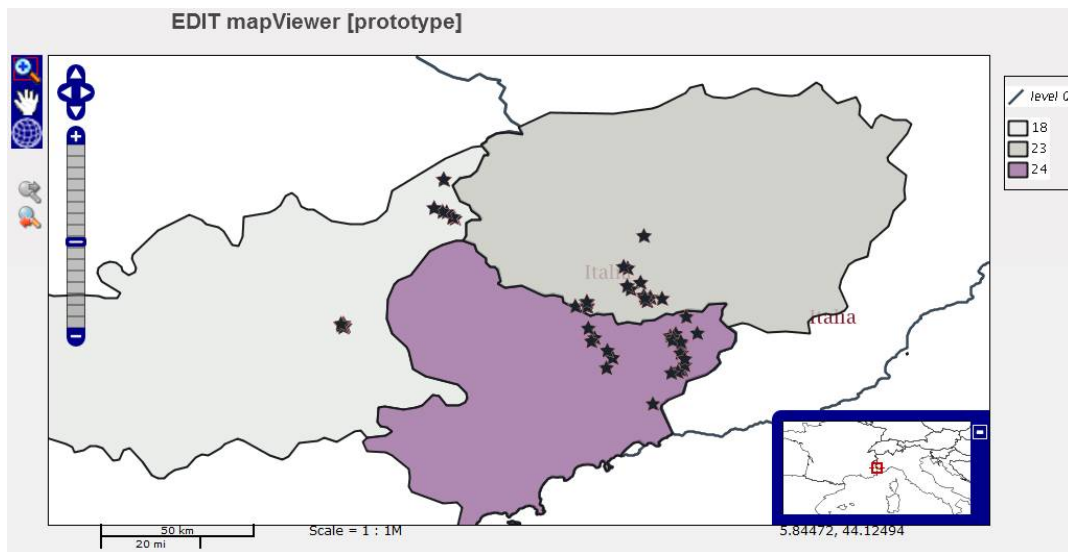
Conclusions and recommendations

The spatial analysis functions of the EDIT mapViewer allow to perform basic spatial investigations such as survey effort or species distribution and biodiversity measure in relation to spatial units. This information facilitates assessment of occurrences distribution and inventory completeness. It can also indicate the spatial units where additional surveys are desirable to carry on in order to obtain a spatial variation of the area.

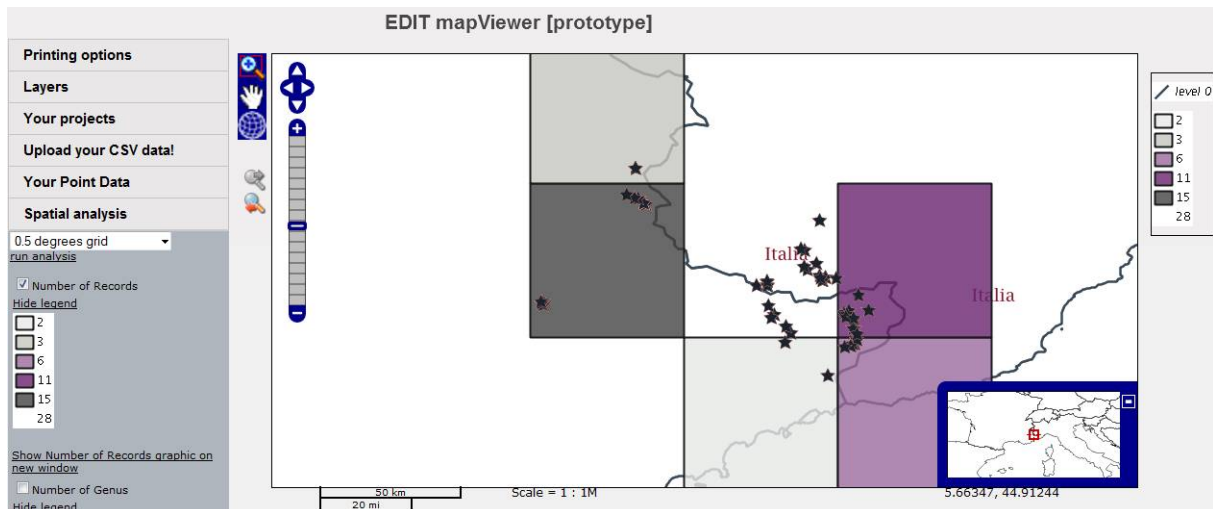
Due to the ATBI+M research sites cover relatively small areas, it was recommended to introduce additional GIS layers with smaller grid e.g. 25km or even 10km UTM grids. It should allow to assess the inventory characteristics more precisely. It was also pointed out that an accuracy of creation and display of interval classes should be enhanced (*compare Map 2.*).

Moreover, it was suggested that further functions to analyze which taxa groups are not well represented in the ATBI sites should be added. This could be done by comparing the species number from the ATBI sites for example with checklist data from the corresponding country.

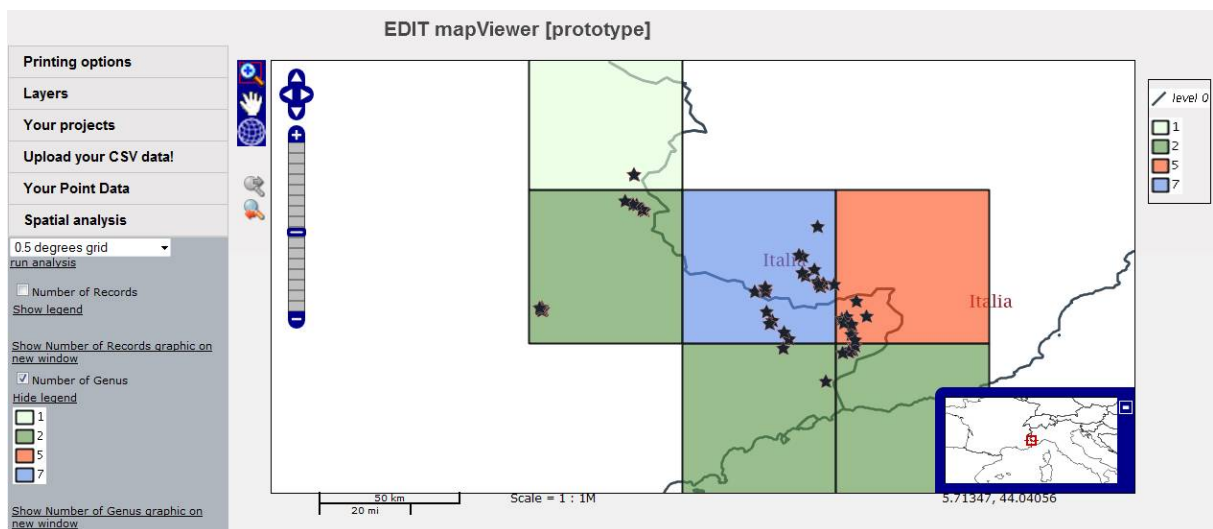
Annex: Maps.



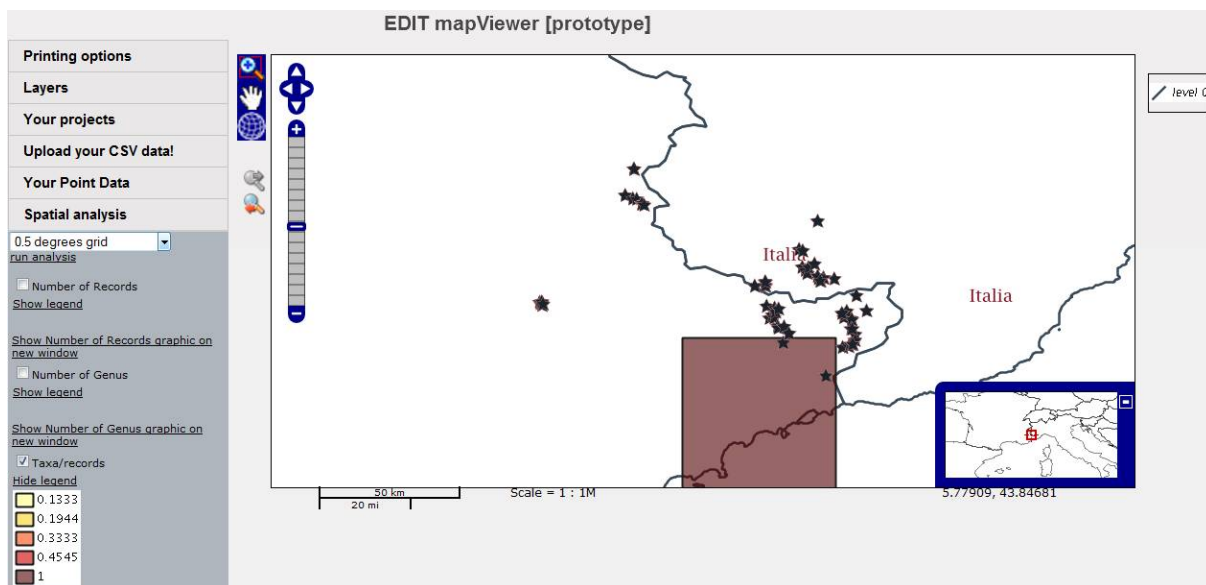
Map 1. Analysis for Europe level 1 layer.



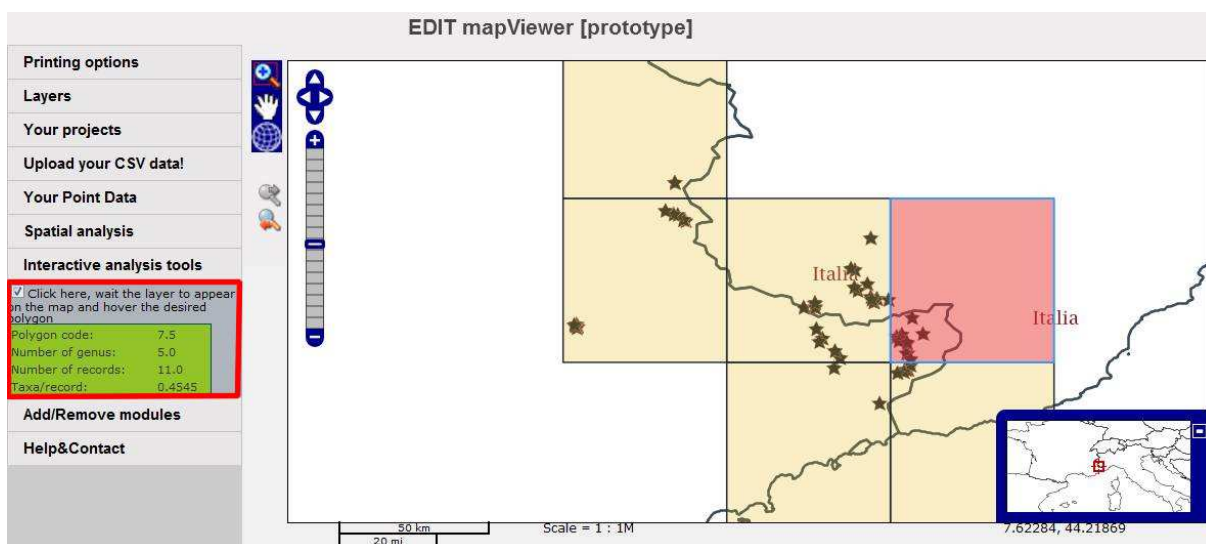
Map 2. Map of sampling effort.



Map 3. Map of taxonomic richness.



Map 4. Map of inventory uncertainty.



Map 5. Map with 'Interactive analysis tools' activated.