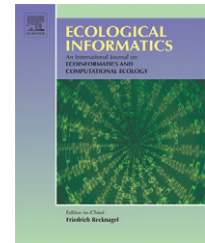


available at www.sciencedirect.comwww.elsevier.com/locate/ecolinf

GBD-Explorer: Extending open source java GIS for exploring ecoregion-based biodiversity data

Jianting Zhang^{a,*}, Deana D. Pennington^a, Xianhua Liu^b

^aLTER Network Office, the University of New Mexico, Albuquerque, NM, 87131, United States

^bDepartment of Biology, the University of North Carolina, Chapel Hill, NC, 27599, United States

ARTICLE INFO

Article history:

Received 10 October 2006

Received in revised form 9 May 2007

Accepted 9 May 2007

Keywords:

Species distribution

Ecoregion

Taxonomic tree

Open source GIS

ABSTRACT

Biodiversity and ecosystem data are both geo-referenced and “species-referenced”. Ecoregion classification systems are relevant to basic ecological research and have been increasingly used for making policy and management decisions. There are practical needs to integrate taxonomic data with ecoregion data in a GIS to visualize and explore species distribution conveniently. In this study, we represent the species distributed in an ecoregion as a taxonomic tree and extend the classic GIS data model to incorporate operations on taxonomic trees. A prototype called GBD-Explorer was developed on top of the open source JUMP GIS. We use the World Wildlife Fund (WWF) terrestrial ecoregion and WildFinder species databases as an example to demonstrate the rich capabilities implemented in the prototype.

© 2007 Elsevier B.V. All rights reserved.

1. Introduction

Biodiversity and ecosystem data are not only geographically referenced but also “species-referenced” (BDEI, 2001). Several large scale species distribution datasets associate species with ecoregions (Loveland and Merchant, 2004). The ecoregions and the species associated with them have been widely used in basic ecological research and for making policy and management decisions as well (Thompson et al., 2004; McDonald et al., 2005; Lamoreux et al., 2006). Geographical Information System (GIS) has been used to visualize and analyze species distribution data. Desktop GIS systems, such as ESRI ArcGIS (ESRI, 2006), have been used to generate distribution maps for single species and species richness maps for quite some time. More recently Web-based GIS (Peng and Tsou, 2003) systems have been routinely used for exploring species distribution data. For example, the World Wildlife Fund (WWF) uses ESRI ArcIMS to map the ecoregions in which a selected species is distributed (WWF, 2006). Unfortunately, none of the existing commercial or open source GIS software support visualizing and exploring large number of species distributions simultaneously based on their taxonomic relationships. This is mostly due to the

relational data model used by most of the current leading GIS systems to manage non-geometric data. The data model requires all the information associated with a geometric object (such as a polygon) be fields with primitive data types (such as integer, real or string). On the other hand, species distributed in an ecoregion can be represented as a taxonomic tree which provides more information than simply a list of species. We are not aware of existing GIS packages that support user-defined data types declaratively. Existing open source GIS systems must be programmatically extended to support taxonomic trees and the operations on them.

In this study, we adopt Darwin Core (TDWG, 2006) and use the following eight levels of taxonomy: *Kingdom/Phylum/Class/Order/Family/Genus/Species/SubSpecies*. Hereafter we will refer to these eight levels of taxonomy as taxonomic ranks and taxon names at all taxonomic ranks as taxonomic data. The benefits of supporting a taxonomic tree in an extended geographical information system are three-folds: 1) From a data modeling perspective, the taxonomic data are treated similar to the geometric data — both are the extensions of primitive data types. Several new operations can be defined systematically based on the newly introduced taxonomic data type. 2) From a

* Corresponding author. Department of Computer Science, the University of California, Davis, USA. Tel.: +1 530 752 5764.
E-mail address: jianting@cs.ucdavis.edu (J. Zhang).

From/To	Tabular	Spatial	Taxonomic
Tabular	Relational (RDBMS)	GIS	(4)
Spatial			(1)
Taxonomic		(2)	(3)

Fig. 1 – Supported high-level operations in the prototype. Numbers label operation types that will be discussed in the text.

system perspective, the linkage between geometry and taxonomy is changed from external to internal and programming work to link the two types of data can be greatly reduced. 3) From a user perspective, the taxonomic information is displayed in a tree structure which is natural and familiar to ecologists. Queries can be performed conveniently by making use of the structures of taxonomic trees.

A prototype system called GBD-Explorer was developed by extending the open source Java Unified Mapping Platform (JUMP) GIS package (Vivid Solutions, 2004) to incorporate taxonomic data into GIS. The prototype system aims at supporting users to explore ecoregion-based biodiversity data visually and interactively, stimulating hypotheses and seeking possible explanations. The current implementation focuses on interlinking geographical data and taxonomic data of ecore-

gions for exploratory species distribution analysis. Supporting additional types of geographical data (such as grids rasterized from species range maps), incorporating environmental data (such as topographical, bioclimate and satellite data), allowing more complex user interactions (such as defining gradients interactively) and linking with analytical models (such as multivariate regression techniques and machine learning algorithms), have been planned for future developments.

In this paper, WWF’s terrestrial ecoregion data and Wild-Finder species data (WWF, 2006) are used as an example to demonstrate the prototype’s rich capabilities, such as mapping multiple species distributions based on complex spatial/taxonomic queries, comparison of taxonomic trees of interactively selected region groups and navigating among taxonomic trees and their associated geographical regions. The rest of this paper is arranged as follows. Section 2 introduces the prototype’s data model and supported high-level operations. Section 3 presents an overview of the prototype system. Sections 4-7 discuss the design and implementation for each of the four typical application scenarios using examples. Finally Section 8 is the summary and future work directions.

2. Data model and supported operations

The prototype was developed by adopting and extending JUMP GIS (Vivid Solutions, 2004). One of the fundamental data structures in JUMP is called *Feature*. Geometric data are

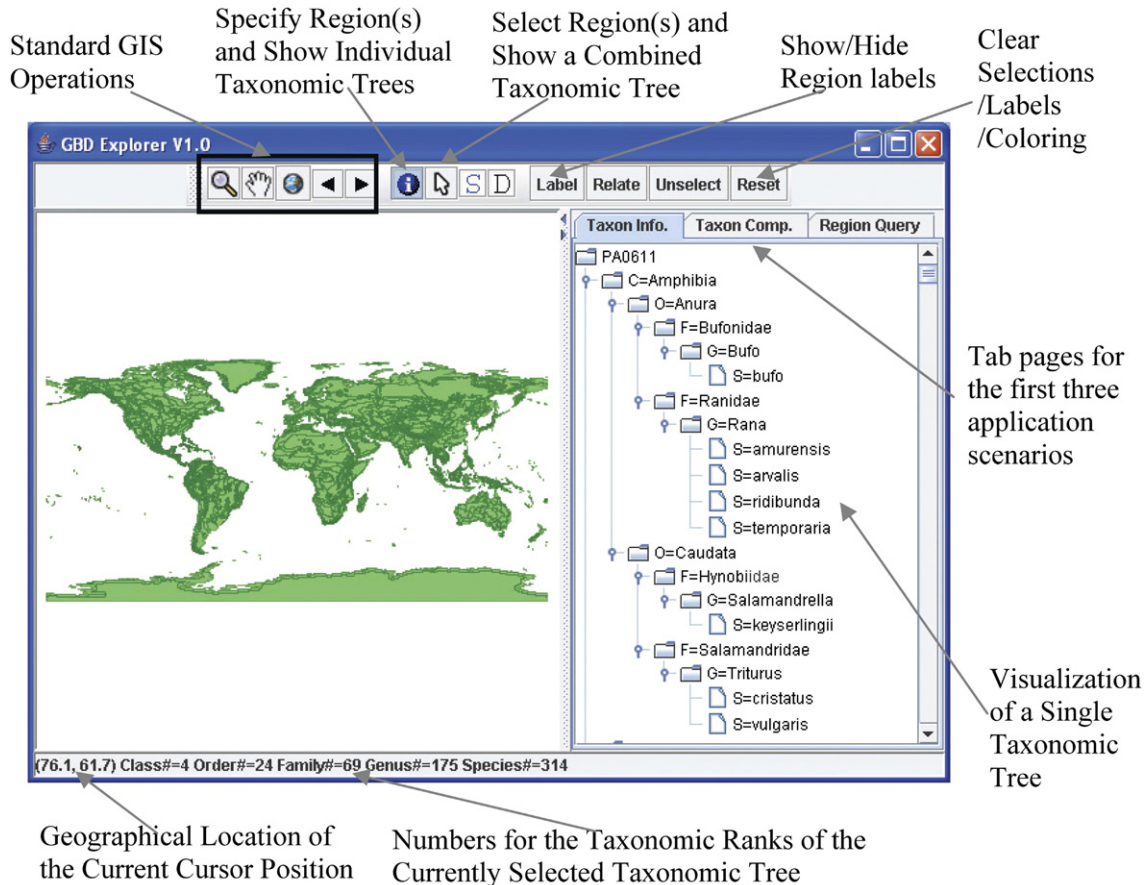


Fig. 2 – Overview of GBD-Explorer.

abstracted and encapsulated as a predefined field called *Geometry* and associated with the tabular fields. We define a new field for taxonomic data and add it to the data model of the prototype. The layout of the tabular, spatial and taxonomic fields of a Feature is shown in Fig. 1.

The prototype currently supports four types of high-level operations as shown in Fig. 1. (1) From spatial to taxonomic. Users can select regions in maps to obtain the corresponding taxonomic trees. (2) From taxonomic to spatial. Users can select multiple paths in a taxonomic tree as the query criteria and regions associated with the selected taxonomic data will be selected and mapped. (3) From taxonomic to taxonomic. Several operations on two taxonomic trees, including Union, Intersection and Differences are supported. (4) From tabular to taxonomic. Since a taxonomic tree shares the same key with the tabular fields of a Feature, any query result set on the tabular data can be mapped to a taxonomic tree set, similar to finding spatial data from a tabular query result set. We have not defined operations from taxonomic to tabular since we have not identified practical needs for that functionality.

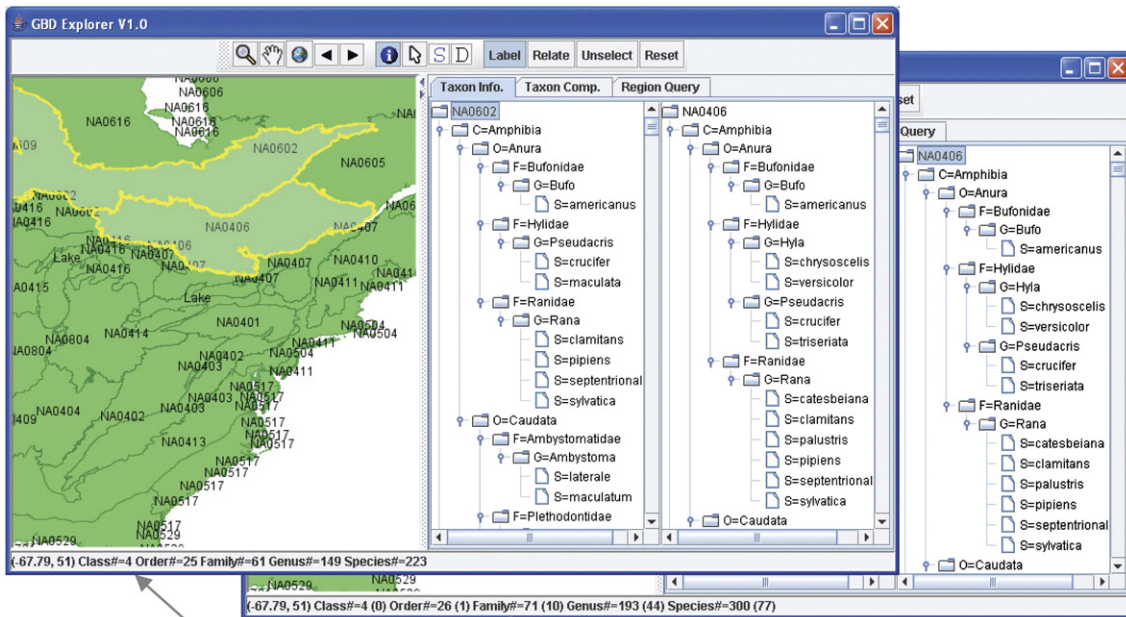
Multiple types of high-level operations can be combined in several different ways to reflect different application scenarios. Four typical application scenarios are currently identified and supported in the prototype system. The supported scenarios are:

- Select one or more regions and get their taxonomic trees. This is the simplest application scenario and involves type 1 operations only. Details are provided in Section 4.
- Select source and destination regions or region groups, compute the union, intersection, and differences of the taxonomic trees of the two region groups. This application scenario is the combination of type 1 and type 3 operations. Details are provided in Section 5.

- Map to regions by querying taxonomic trees. For a given taxonomic tree, the prototype allows users to select multiple tree nodes and the paths from the taxonomic tree root to the nodes will be used as the conjunctive criteria for querying and mapping regions. This scenario requires type 2 operations and more details will be provided in Section 6.
- Navigate among geospatial regions and taxonomic trees. Users can start with a particular region or region groups and get the associated taxonomic tree; from the taxonomic tree, users can select one or more species and map to their distribution regions. This application scenario actually is a combination of the first and the third scenarios and thus can potentially involve types 1 and 2 operations. The exploration process can be conducted in an iterative manner to explore the species–region relationships of interests. More details are provided in Section 7.

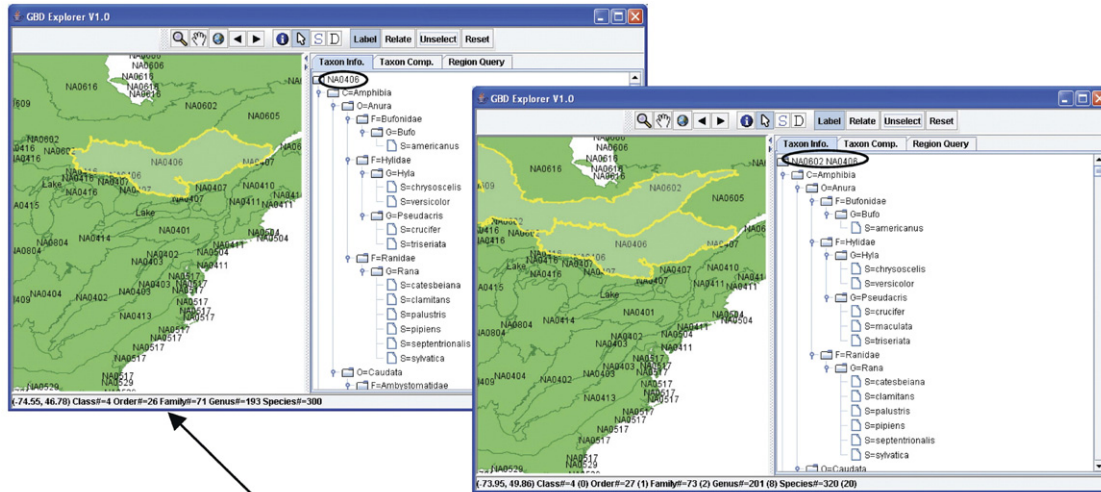
3. System overview

Fig. 2 provides an overview of the prototype’s interface. The prototype does not have menus but instead provides a toolbar at the top. There is also a status bar at the bottom to show feedbacks during operations, such as system information, current cursor locations and numbers of species at different taxonomic ranks associated with a taxonomic tree. The major canvas is divided into two parts. Geographical (ecoregion) data are displayed in the left part. The right part has three tabbed pages, namely *Taxon Info* (display the taxonomic trees corresponding to selected regions), *Taxon Comp[arison]* (display/compare original and derived taxonomic trees of selected source and destination region groups), and *Region Query* (query regions that satisfy taxonomic criteria). These three tabs correspond to the first three application scenarios. The canvas



Numbers for the Taxonomic Ranks of the Currently Selected Taxonomic Tree and Differences Compared to the Previous Selected Taxonomic Tree

Fig. 3 – Display taxonomic trees and species richness among specified regions.



Numbers for the Taxonomic Ranks of the Currently Selected Taxonomic Tree and Their Differences while New Regions are Selected and Their Taxonomic Trees are Combined

Fig. 4– The combined taxonomic trees and their differences of species richness of selected regions.

is implemented as a split pane to allow users to set the proportion of space used to display geographical data and taxonomic data. At one extreme, the geographical data can take all of the canvas space and the prototype serves as a simple GIS to display geographical data. Similarly, the taxonomic data can take all of the canvas space and the prototype can be used to browse and compare taxonomic trees.

The far left part of the toolbar is inherited from JUMP GIS with some modifications to use the related functions as a toolkit rather than as a workbench for which JUMP was originally designed (Vivid Solutions, 2004). While we still use the *Specify* and *Select* icons (represented by the letter *i* and an arrow, respectively) as in JUMP and some other GIS systems, their functions are rewritten to show taxonomic trees rather than tabular records. The difference between *Specify* and *Select* for taxonomic data in the prototype is that *Specify* shows individual taxonomic trees that correspond to the selected regions while *Select* combines the taxonomic trees using a Union operation (see Sections 4 and 5 for more details).

Since JUMP allows only one type of selection (which is tied to the *Select* tool), we added two more types of selections and tied them to “S” (*Source*) and “D” (*Destination*) tools for comparison of taxonomic trees (see Section 5 for more details). The boundaries of selected regions are colored differently to reflect their types: yellow for *Select*, red for *Source* and blue for *Destination*. Selected regions are highlighted using a lighter fill color than non-selected regions. By default, new selections (including regions selected by *Select*, *Source*, *Destination*) replace previous selections. However users can keep previous selections by holding the shift key, functionality provided by JUMP. The prototype also keeps data availability in mind when constructing user interfaces. If only geographical data are available, the selection tools and the related tab pages will be disabled so that the system can still function as a simple GIS.

Throughout the rest of the paper we will use the WWF’s terrestrial ecoregion data and WildFinder species data (WWF, 2006) to demonstrate the functionality of the prototype system. WWF ecoregion data were provided in ESRI Shapefile format and have 14458 polygons representing the 825 ecoregions in 8 realms and 26 biomes. The WWF WildFinder species database was provided in Microsoft Access database format which has 29,112 species, 4815 genera, 445 families and 69 orders in 4 classes (amphibians, reptiles, birds, and mammals). There are 350,045 species-ecoregion records, i.e., about 400 species per ecoregion on average. The sizes of the GIS data and the species data are about 70 megabytes and 80 megabytes, respectively. All the experiments are performed on a Dell OptiPlex GX270 desktop machine with a Pentium 4 3.2 GHZ processor and 256 megabytes Java Virtual Machine (VM) memory are used. While the hardware configuration is below a typical desktop machine at present time, the response time and the general performance are satisfactory.

4. Application scenario 1: region to species

The simplest application scenario is to visualize a taxonomic tree after selecting one or more regions (termed as a region group). As introduced in Section 3, the prototype allows users to visualize each individual taxonomic tree by using the *Specify* tool and display a combined taxonomic tree by using the *Select* tool. These two tools help users to identify species distributed in a particular region or region groups visually and interactively.

In Fig. 3, when the cursor is placed near the border of Ecoregions NA0602 and NA0406, the two regions are specified and the corresponding taxonomic trees are displayed in the *Taxon Info.* tab page. When users select the taxonomic tree corresponding to region NA0602, the numbers of species at the

five taxonomic ranks (Class, Order, Family, Genus and Species) will be displayed in the status bar. When users also select the tree corresponding to region NA0407, the differences at the five ranks are also displayed in the status bar. From the results, users can easily see that ecoregion NA0406 has greater species richness than ecoregion NA0602 (1 more order, 10 more families, 44 more genera and 77 more species). When users want to work on two taxonomic trees that are not spatially adjacent (and hence can not be selected by using the Specify tool, they can follow the taxonomic tree comparison scenario by explicitly identify the source and destination regions for comparison as detailed in Section 5.

While the Specify tool selects regions that intersect with the current cursor position within a certain distance threshold and displays the corresponding taxonomic trees individually, the Select tool allows the user to select regions that intersect with an interactively drawn rectangle or to add multiple regions resulting from using the Specify tool multiple times. In contrast to the Specify tool, the Select tool displays a combined taxonomic tree from all the selected regions. Since visualizing a combined taxonomic tree takes much less screen space than visualizing multiple taxonomic trees, it is easier for users to browse through the combined taxonomic tree. As shown in Fig. 4, the status bar shows the numbers for each

taxonomic rank of the combined taxonomic tree. It also shows the differences compared with previously combined taxonomic tree as the regions are being added and their taxonomic trees are being combined. The Select tool allows selection of any number of regions in an arbitrary order, showing the changes of species richness among the selected regions in a manner similar to the Specify tool.

5. Application scenario 2: taxonomic comparisons

Defining the differences in species compositions among regions and/or region groups is essential to exploring species distributions, stimulating hypotheses and seeking possible explanations. We have implemented tools in GBD-Explorer that allow users to select arbitrary source and destination regions or region groups for visualization and comparison. This tool is useful especially when users choose regions or region groups along the environmental gradients for comparisons.

We extend operations on sets to taxonomic trees, including *Union*, *Intersect* and *Differences*. The taxonomic trees of regions are the sub-trees of the taxonomic tree of the whole datasets (the union of all the corresponding taxonomic trees) and thus they have the same maximum possible heights. The Union operation

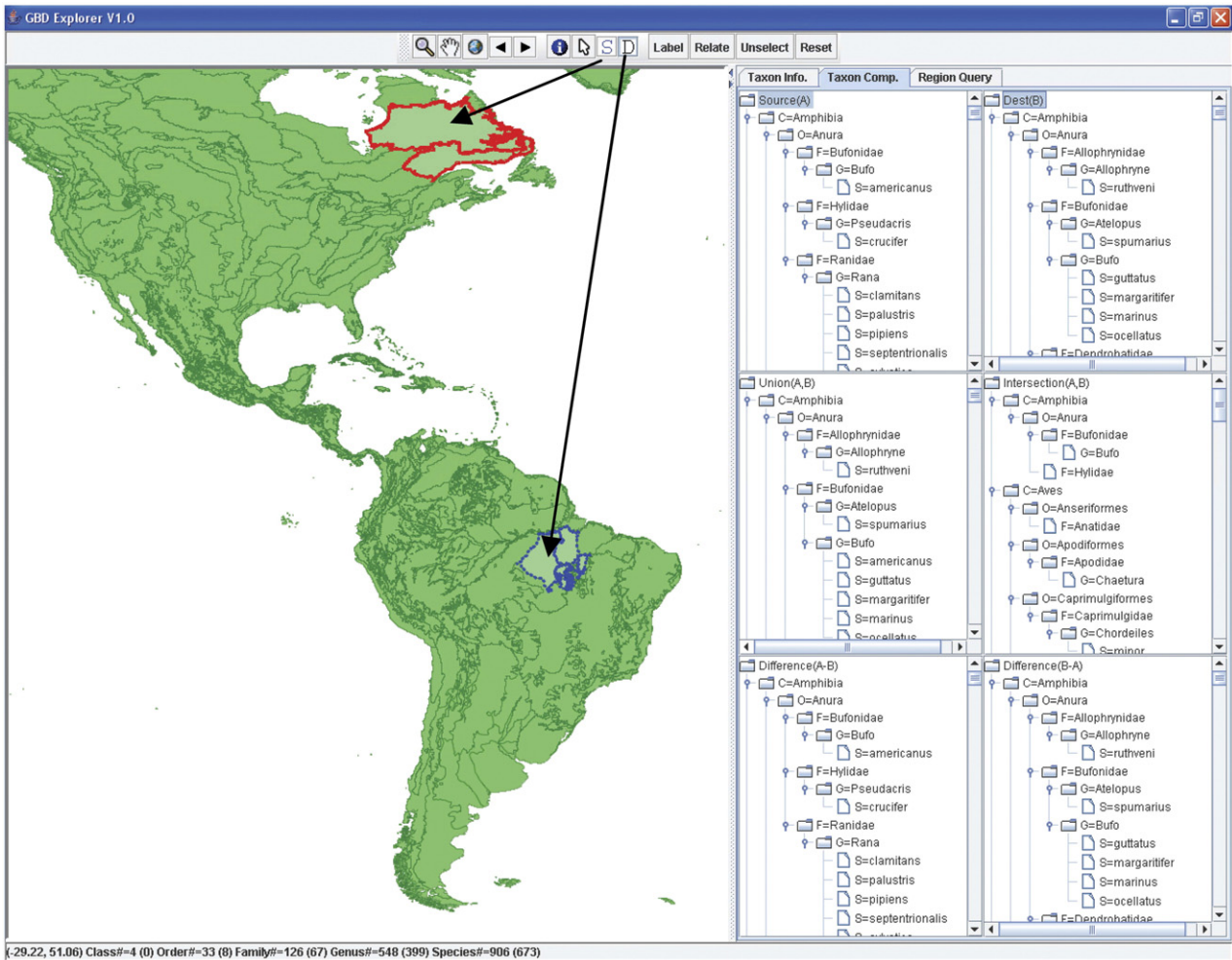


Fig. 5 – Comparing taxonomic trees of the source and destination region groups.

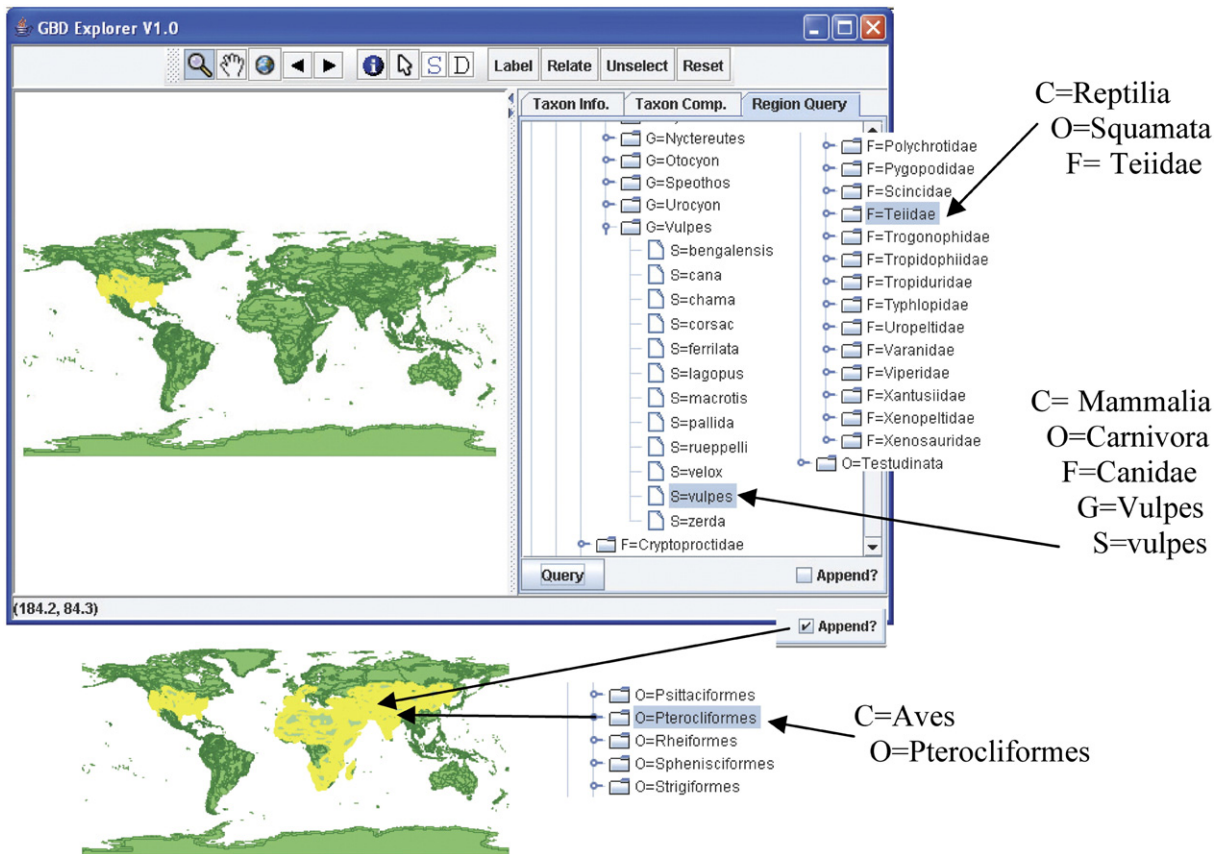


Fig. 6 – Query interface for mapping to regions based on taxonomic queries. Top: using query criteria of both family and species, bottom: appending query results using query criteria at the taxonomic rank order.

combines the nodes in the source and destination taxonomic trees (including both leaf and non-leaf nodes). The Intersect operation only keeps the nodes that are in both source and destination taxonomic trees. The Difference operation removes all the leaf nodes in the source tree that are also in the destination tree, i.e. *src-dest*. The taxonomic tree operations are implemented in Java using set operations at each level of the taxonomic trees. Tree structures are created for the resulting taxonomic tree based on the results of the set operations during the processes. Note that while Union and Intersection operations are order-invariant, the Difference operation relies on the order of the two involving taxonomic trees. To remove all the leaf nodes in *dest* that are also in *src*, i.e. *dest-src*, we only need to switch the order of *src* and *dest* when applying the Difference operation.

Fig. 5 shows the results of comparing two ecoregions (NA0605 and NA0606) in North America and two ecoregions in South America (NT0168 and NT0140). The numbers shown in the status bar indicate that the two ecoregions of the South America have significantly higher species richness, i.e., 8 more orders, 67 more families, 399 more genera and 673 more species. By further looking into their intersected taxonomic tree, we can see that they have few in common at all taxonomic ranks, i.e., the intersected taxonomic tree has only 13 species, 36 genera, 38 families, 22 orders. The reason that the number of genera is larger than the number of species (and similarly the numbers between the families and genera) in the intersected taxonomic tree is that the two taxonomic

trees (Source and Dest) have some common genera but do not have common species under the genera. The prototype system allows user to explore the other three derived taxonomic trees in a similar manner.

6. Application scenario 3: mapping to regions

GBD-Explorer has the capability of mapping species to their geographical distribution regions based on complex taxonomic criteria. This capacity distinguishes it from most existing species mapping systems that can only map a single species at a time. The functionality is also useful to explore species collocation patterns interactively.

The mapping to regions function by querying taxonomic trees (“Region Query” for short) is provided as a tab page in GBD-Explorer (Fig. 6). The taxonomic tree for the whole dataset is displayed in the tab page and users can select one or more nodes from the taxonomic tree. The “Append” checkbox at the very right-bottom allows the user to choose whether a new query result will replace (when unchecked) or append (when checked) to previously selected regions. For a single selected node in the taxonomic tree, the path from the root all the way to the node (which we call a query path) is used as a sub-query as follows. Suppose the query path is *a.b.c* and the corresponding taxonomic ranks of the labels along the path are A, B and C (such as Family, Genus and Species), then the sub-query will be

Select Geometry From Data where $A = a$ and $B = b$ and $C = c$. For example, the “where” condition for the first sub-query shown in Fig. 6 will be Class = Reptilia and Order = Squamata and Family = Teiidae. Similarly the “where” condition for the second query path in Fig. 6 is Family = Mammalia and Order = Carnivora and Family = Canidae and Genus = *Vulpes* and Species = *vulpes*. Note that all the involved taxonomic ranks appearing in the “where” condition are conjunctive and in the order of their appearances in the query path. Also according to the translation rule, the taxonomic ranks below the rank corresponding to the selected node are not involved in the query, which essentially chooses all the species under the node (taxonomic rank).

The policy to combine multiple sub-queries into a complete query to be submitted to the query engine is also conjunctive by design. The top-left side of Fig. 6 highlights the regions resulting from the query shown at the top-right side. The query includes two conjunctive sub-queries that are corresponding to the two query paths, i.e., regions have species in Class Reptilia Order

Squamata Family Teiidae and Class Mammalia Order Carnivora Family Canidae Genus *Vulpes* Species *vulpes*. The reason is primarily that conjunctive query of multiple species is frequently used. In addition, since the sub-query for a single selected node is conjunctive, a disjunctive query can be promoted to the top level and use the “Append” option to combine the query results. This can be shown using the following example. Suppose we have a disjunctive condition somewhere along a query path like ($A = a$ and ($B = b1$ or $B = b2$ or ($B = b3$ and ($C = c1$ or $C = c2$))). The query can be easily decomposed to ($(A = a$ and $B = b1)$ or ($A = a$ and $B = b2)$ or ($A = a$ and $B = b3$ and $C = c1)$ or ($A = a$ and $B = b3$ and $C = c2$)). The bottom part of Fig. 6 shows the result after appending the query result of $C = Aves$ and $O = Pterociformes$. Allowing appending new query results not only supports more complex taxonomic queries but also allows users to compare the spatial relationships among multiple species groups (as expressed in their query criteria), such as contain, disjoint and degrees of overlaps.

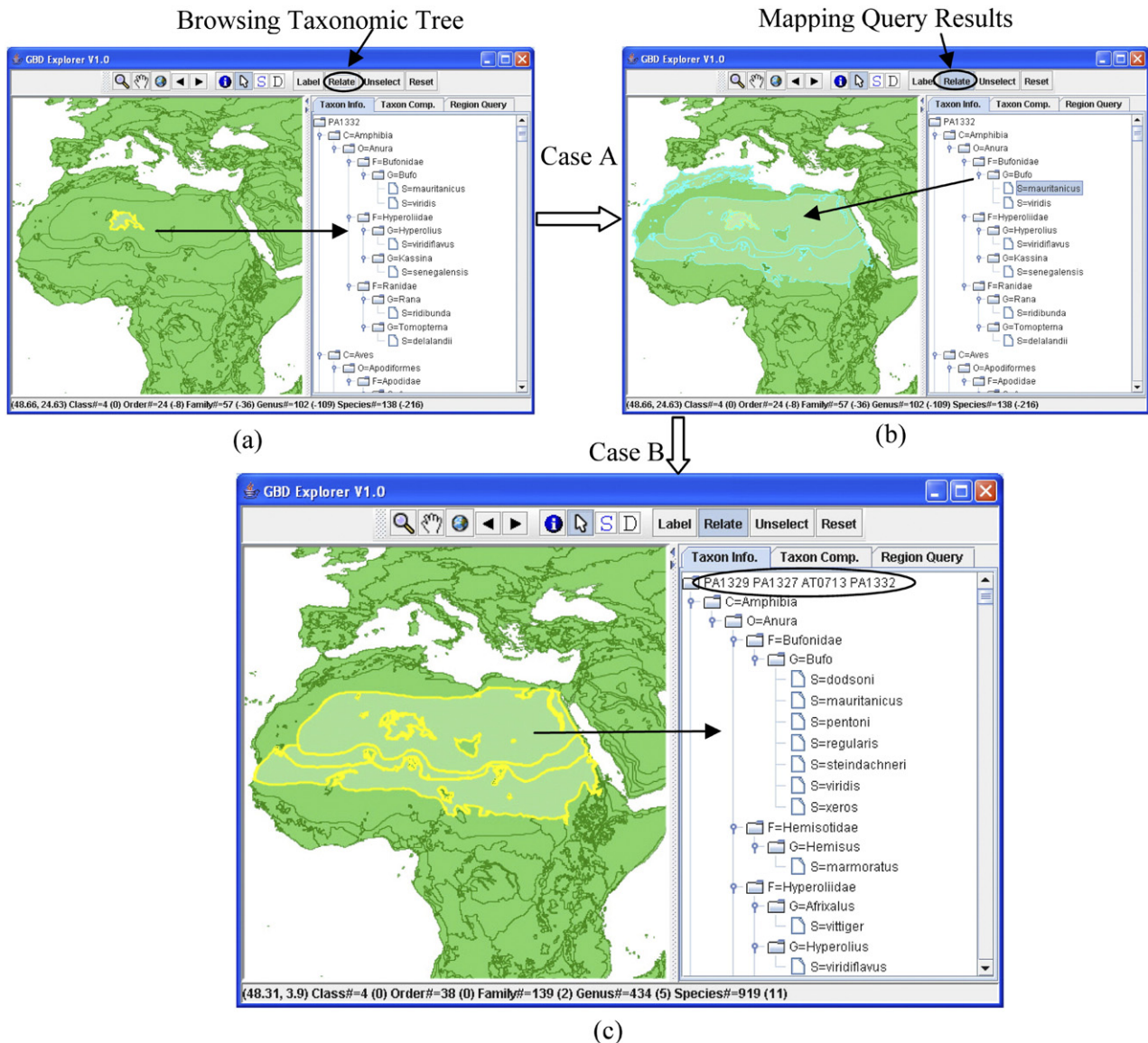


Fig. 7 – Scenario 4 example to demonstrate geospatial-taxonomic navigation. (a) Region to Species, (b) Species to Regions, and (c) Regions to Species Again.

7. Application scenario 4: geospatial-taxonomic navigations

The above three application scenarios involve data retrievals either from geospatial to taxonomic or from taxonomic to geospatial. Moreover, the following two cases may happen during the exploration process: A) After selecting a region or region group and obtaining its taxonomic tree (as in scenarios 1 and 2), users may want to know the other regions that also have a common subset of species. B) After querying regions based on a certain taxonomic criteria on a taxonomic tree (including the one of the whole dataset), users may want to see the individual or combined taxonomic trees of all or some of the resulting regions. The prototype's capabilities in supporting the iterative geospatial-taxonomic navigations, as demonstrated in the example below, can be useful in exploring species-area relationships (Scheiner, 2003) visually and interactively.

For case B, users can perform the operations involved in application scenario 3 and scenarios 1 and 2 sequentially and achieve the goal. On the other hand, support for case A essentially requires implementing type 2 operation (as in scenario 3) for each of the resulting taxonomic trees in scenarios 1 and 2. The difference between type 2 operations in the application scenario 3 and the application scenario 4 (case A) is that the former operates on the taxonomic tree for the whole dataset while the later operates on the taxonomic tree for a region or a region group, which is a subset of the taxonomic tree for the whole dataset. GBD-Explorer implements the type 2 operation for each taxonomic tree resulting from either scenario 1 or scenario 2 operations by reusing the program codes for the type 2 operation in scenario 3 through modular design.

Users can switch between browsing a taxonomic tree and mapping the query result with the toggle button *Relate* (Fig. 7). While all tree nodes are selectable in either browsing or querying mode, when the button is toggled, the selected tree nodes and their corresponding tree paths will be used to query against the geometries of the whole datasets and map the resulting regions. The obvious reason to distinguish the two modes is the computation cost. While browsing a taxonomic tree incurs little computation cost once the tree is rendered, the computation cost for query and mapping is proportional to the multiplication of the number of regions in the whole dataset and the number of query paths, which could be quite computationally expensive. Our experimental results show that GBD-Explorer can respond within a fraction of a second for datasets with tens of thousands of geometric objects and up to 10 query paths after query optimization. The navigation performance is satisfactory.

The two cases in scenario 4 can also be combined in arbitrary orders that allow users to navigate among taxonomic trees and their distributed geographical regions. For the example shown in Fig. 7, the ecoregion PA1332 was first selected using the *Select* tool and the corresponding taxonomic tree is visualized (Fig. 7(a)). When users toggle down the *Relate* button and clicked the taxonomic tree node *S* = *mauritanicus*, the prototype perform the query *C* = Amphibia and *O* = Anura and *F* = Bufonidae and *G* = *Bufo* and *S* = *mauritanicus* against its databases and map the resulting regions (Fig. 7(b)).

When users select four of the resulting ecoregions (PA1329, PA1327, PA0713 and PA1332), the combined taxonomic tree is then displayed (Fig. 7(c)). Comparing the combined taxonomic tree of the four ecoregions (Fig. 7(c)) and the taxonomic tree of ecoregion PA1332 alone (Fig. 7(a)), we can see that the number of species under Genus *Bufo* is increased from 2 to 7.

8. Summary and future work

In this paper, we report the design, development and application of the GBD-Explorer prototype system that extends GIS functionality and incorporates taxonomic data from species distributions. We identify several basic operations and typical application scenarios that are built on top of the basic operations. Examples have been provided for each of the application scenarios to demonstrate the prototype's capabilities using WWF terrestrial ecoregions GIS and Wildfinder species data. While the prototype was primarily designed for ecoregion-based biodiversity data, it can be used for generic region-based (such as administrative) geographical data and other domain-specific data that can be represented as rooted trees (such as certain types of evolution trees).

For future work, we plan to support more application scenarios by combining the basic operations and designing new user interfaces. From an ecological research perspective, we plan to link bioclimate variables (Hijmans et al., 2005) and remote sensing data (Pettoirelli et al., 2005) with the ecoregions and explore species-energy (Hawkins et al., 2003) and species-productivity (Waide et al., 1999) relationships at the ecoregion level, in addition to exploring species distribution patterns. Finally, following our previous practices in scientific workflow based ecological modeling and biodiversity studies in distributed computing environments (Zhang et al., 2005), we also plan to use Kepler scientific workflow system (Ludäscher et al., 2006) to access distributed online environmental and species distribution data and build reusable workflows to use the prototype system for similar purposes.

9. Software access

The prototype system, including source codes, binary distributions, third-party libraries and data, are publically available at <http://cvs.ecoinformatics.org/cvs/cvsweb.cgi/seek/projects/beam/GBDExplorer/>.

Acknowledgement

This work is supported in part by DARPA grant # N00014-03-1-0900 and NSF grant ITR #0225665 SEEK.

REFERENCES

- BDEI, 2001. Executive Summary, Report of NSF/USGS/NASA Workshop on Biodiversity and Ecosystem Informatics. <http://www.evergreen.edu/bdei/2001/bdes1.pdf>.
 ESRI, 2006. ArcGIS. <http://www.esri.com/software/arcgis/>.

- Hawkins, B.A., Field, R., Cornell, H.V., Currie, D.J., Guegan, J.F., Kaufman, D.M., Kerr, J.T., Mittelbach, G.G., Oberdorff, T., O'Brien, E.M., Porter, E.E., Turner, J.R.G., 2003. Energy, water, and broad-scale geographic patterns of species richness. *Ecology* 84 (12), 3105–3117.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25 (15), 1965–1978.
- Lamoreux, J.F., Morrison, J.C., Ricketts, T.H., Olson, D.M., Dinerstein, E., McKnight, M.W., Shugart, H.H., 2006. Global tests of biodiversity concordance and the importance of endemism. *Nature* 440 (7081), 212–214.
- Loveland, T.R., Merchant, J.M., 2004. Ecoregions and ecoregionalization: geographical and ecological perspectives. *Environmental Management* 34, S1–S13.
- Ludäscher, B., Altintas, I., Berkley, C., Higgins, D., Jaeger, E., Jones, M., Lee, E.A., Tao, J., Zhao, Y., 2006. Scientific workflow management and the Kepler system. *Concurrency and Computation: Practice and Experience* 18 (10), 1039–1065.
- McDonald, R., McKnight, M., Weiss, D., Selig, E., O'Connor, M., Violin, C., Moody, A., 2005. Species compositional similarity and ecoregions: do ecoregion boundaries represent zones of high species turnover? *Biological Conservation* 126 (1), 24–40.
- Peng, Z.-R., Tsou, M.-H., 2003. *Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks*. John Wiley & Sons, Hoboken, New Jersey.
- Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.M., Tucker, C.J., Stenseth, N.C., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution* 20 (9), 503–510.
- Scheiner, S.M., 2003. Six types of species–area curves. *Global Ecology and Biogeography* 12 (6), 441–447.
- TDWG, 2006. Taxonomic Databases Working Group Darwin Core 2. <http://darwincore.calacademy.org/>.
- Thompson, R.S., Shafer, S.L., Anderson, K.H., Strickland, L.E., Pellier, R.T., Bartlein, P.J., Kerwin, M.W., 2004. Topographic, bioclimatic, and vegetation characteristics of three ecoregion classification systems in North America: comparisons along continent-wide transects. *Environmental Management* 34, S125–S148.
- Vivid Solutions, 2004. Unified Mapping Platform (JUMP). <http://www.vividsolutions.com/jump/>.
- Waide, R.B., Willig, M.R., Steiner, C.F., Mittelbach, G., Gough, L., Dodson, S.I., Juday, G.P., Parmenter, R., 1999. The relationship between productivity and species richness. *Annual Review of Ecology and Systematics* 30, 257–300.
- WWF, 2006. World Wildlife Fund WildFinder: Online database of species distributions, ver. Jan-06. <http://www.worldwildlife.org/WildFinder>.
- Zhang, J., Pennington, D.D., Michener, W.K., 2005. Using web services and scientific workflow for species distribution prediction modeling. *Lecture Notes in Computer Science* 3739, 610–617.