
OSGeo Journal

The Journal of the Open Source Geospatial Foundation

Volume 1 / May 2007

In This Volume

Real World Implementations of Open Source software

Introducing Mapbender, deegree, openModeller ...

Understanding Spatial Relationships

Examining the Web Processing Server (WPS) Specification

Package Interaction - GRASS-GMT, Tikiwiki, PyWPS, GRASS-R ...

Software Updates

News, and more...



Project Spotlights

Introducing openModeller

A fundamental niche modelling framework

by Tim Sutton, Renato de Giovanni and Marinez Ferreira de Siqueira

Introduction

In 1957 Hutchinson (4) formalised the fundamental niche principle. In essence, he proposed that if all the environmental conditions that allow a species to exist indefinitely could be tabulated as a multidimensional hypercube, then the resulting hypercube could be considered to be the organism's fundamental ecological niche. Naturally, fully understanding all the ecological conditions for any given species is a monumental task, limited primarily by the lack of rich data that could drive such an analysis. By extrapolation of the association between species occurrence localities and a set of GIS raster layers representing environmental conditions such as rainfall, temperature, solar radiation etc., a correlative approach can be taken towards describing an organism's ecological niche (10). Using this technique, a raster layer can be produced with areas that most resemble the environmental conditions at the original known sites of occurrence for the organism. Arguably, this approach first garnered widespread in-

terest with the publication of 'A biogeographic analysis of Australian Elapid snakes.' (5), where the authors produced predicted distribution maps for Elapid snakes based on simple models of climatic preference. As the world starts to engage more fully in the discussion of the potential impacts of global climate change, the ability to predict the impact of these changes on the distribution of organisms has become more topical. Fundamental niche modelling has been used to predict species loss in the face of future climate changes (12; 2). It has also been used to model invasive species, spread of disease vectors (6) and in phylogenetic modelling to produce distribution maps for hypothetical ancestors (13; 14).

A number of software applications exist for carrying out fundamental niche analysis. Popular choices include Bioclim (5), DesktopGARP (9), MaxEnt (7) etc. Elith *et al.* (3) provide an excellent overview and comparison of some of the more commonly used applications. These applications are usually each created to address a specific algorithmic approach to fundamental niche modelling. Typically these tools are not free (as in freedom), open source software. In addition, each application has different requirements in terms of data preparation, system requirements (Java, MS Windows etc.) and the learning curve required from the user to gain the necessary

skills to use the application. In other cases, algorithms for fundamental niche modelling were developed within scripting toolkits, such as the Climate Space Model (8) which was developed using MatLab. These are difficult to redistribute as user friendly applications.

To address these (and other) concerns, the *openModeller* project was initiated in 2003 by CRIA¹. The *openModeller* project is free and open source, and all code is available on SourceForge. *openModeller* development is supported by FAPESP², and by contributions from the open source developer community. The principle aim of the *openModeller* project is to provide tools for researchers interested in fundamental niche modelling. It is a fresh approach to the world of fundamental niche modelling since it provides a plugin architecture for new algorithms to be incorporated easily. Providing all algorithms under a common architecture allows for easier comparison of results between different algorithms since the all models can be prepared using the same independent data and can be output into a common format. Implementations of GARP, Bioclim, CSM, Environmental Distance and others have already been developed as plugins for *openModeller*.

openModeller capitalises on existing open source software, harnessing libraries such as GDAL³, GSL⁴, QGIS⁵ etc. *openModeller* is written in C++ and is cross-platform, running on MS Windows, Mac OS X and GNU/Linux.

openModeller sub projects

There are a number of development areas within the *openModeller* project. First and foremost is the *openModeller* library. The library provides a common platform for all other tools mentioned below. It includes the ability to read environmental and species occurrence data, setting model parameters, loading and running algorithm plugins and writing the results of an analysis to a geospatial dataset. C++ and Python bindings (generated using SWIG) for the library are available.

A Web Services API based on SOAP (using the document / literal style) is available for remote execution of *openModeller* jobs. A CGI application is

available, allowing this service to be deployed inside of an Apache web server instance.

A number of console tools have been created on top of this library. *om_console* provides a way to run a model using simple text configuration files that specify which input datasets to use. *om_create* and *om_project* provide a means to use *openModeller* using XML configuration files. Other helper tools are also available for the command line to provide simple data extraction and metadata query facilities. For example *om_sampledump* extracts the environmental values at each occurrence point and provides the result in a delimited format.

For end users there is a cross platform graphical user interface 'desktop' application (written in C++ and Qt4). *openModeller Desktop* provides a user friendly environment for preparing, running and visualising the results of fundamental niche models. *openModeller Desktop* will be described in more detail in the section that follows.

openModeller Desktop

openModeller Desktop is a solution for users wishing to use the *openModeller* library without programming or needing to be familiar with the various command line tools that the *openModeller* project provides. Within *openModeller Desktop*, users can carry out the following core activities:

- automatically retrieve occurrence data from online databases such as GBIF⁶ and species-Link⁷;
- manage the environmental raster layers that will be used for modelling;
- manage the parameters for the various algorithm plugins included with *openModeller*;
- design an experiment which will run multiple models for one or more species using one or more algorithms;
- visualise the results of an experiment by viewing individual models using a tabular and a map view.

¹Centro de Referência em Informação Ambiental: <http://www.cria.org.br>

²<http://www.fapesp.br>

³GDAL - Geospatial Data Abstraction Library: <http://www.gdal.org>

⁴GSL - GNU Scientific Library: <http://www.gnu.org/software/gsl/>

⁵QGIS - Quantum GIS: <http://qgis.org>

⁶GBIF - Global Biodiversity Information Facility: <http://gbif.org>

⁷<http://smlink.cria.org.br>

Modelling *Acacia cyclops* - a simple case study

One typical use of fundamental niche modelling is to predict the areas that may be vulnerable to the spread of invasive species. In this example *openModeller* will be used to find areas of potential distribution for *Acacia cyclops*⁸. *Acacia cyclops* is a native plant of south western Australia, but is known to be an invasive pest plant in other parts of the world. Thus the aim of this case study is to ascertain areas potentially vulnerable to invasion from this species. To start, the *openModeller Desktop* localities search tool was used to search for locality records (see Figure 1).

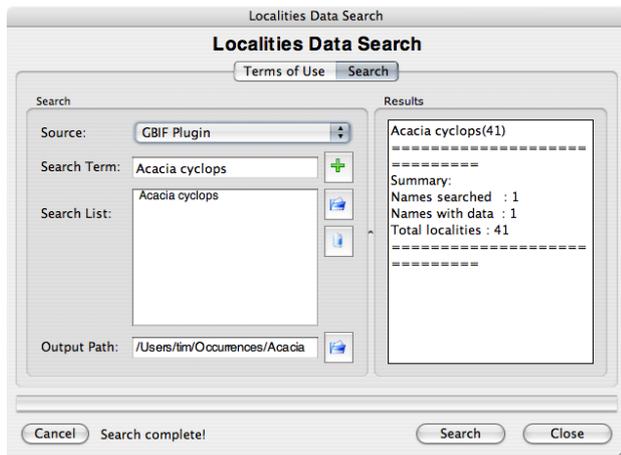


Figure 1: Searching GBIF for locality data.

Since the search results returned records outside of Australia, a text editor was used to delete all occurrence records that were not within Australia. This was done in order to model a fundamental niche based solely on the species' known native distribution. This reduced the initial 41 records retrieved from GBIF to 31 records, all of which originate from the Australian Virtual Herbarium (1). After this, the layerset manager in *openModeller Desktop* was used to define a layerset. 23 layers derived from CRU⁹ were selected for this purpose (see Figure 2.). These layers represent climate related indicators (based on recorded climate measurements) related to temperature, rainfall etc.

⁸*openModeller Desktop* is capable of modelling many species in a single experiment, but for the purposes of a simple case study only one species was used

⁹CRU - Climate Research Unit: <http://www.cru.uea.ac.uk>

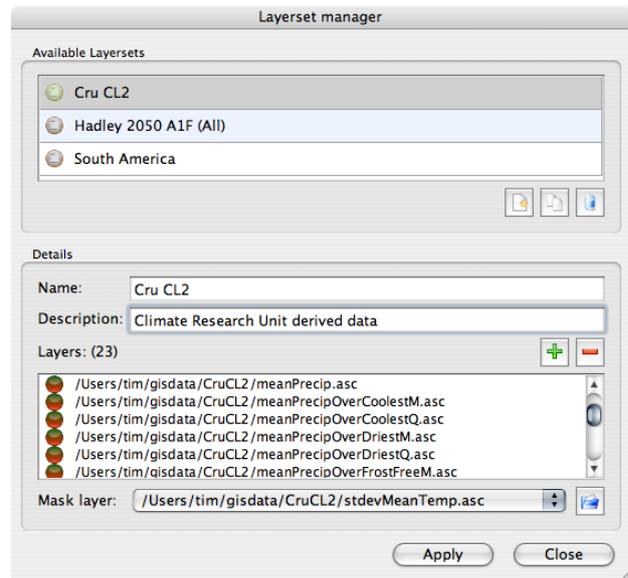


Figure 2: The layerset manager.

Creating named layersets provides great benefit to the user since they can easily refer to multiple layers as one conceptual entity and do not need to spend a lot of time selecting files from the file system each time they run an experiment. Once a layerset has been defined it can be used repeatedly in different experiments.

openModeller Desktop includes an algorithm manager (fig. 3). This tool lets the user view all the available modelling algorithm plugins that have been installed along with the *openModeller* library. All algorithms ship with sensible defaults (as determined by algorithm authors) so that they can be used 'out the box'. Within the algorithm manager these default parameters are presented as immutable 'system profiles'. System profiles can be cloned and modified to create 'user profiles'. Modification of these parameters requires some knowledge of the principles behind the underlying algorithm, but provides advanced users with a great deal of flexibility and control. Novice users can eschew the algorithm manager completely and simply work with the default 'system profiles'.

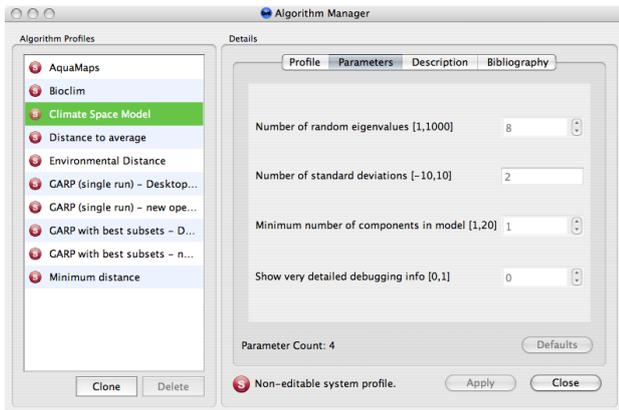


Figure 3: The algorithm manager.

The occurrence data, layerset and algorithm profile are the three essential elements for designing a modelling experiment. The experiment designer (see Figure 4.) is used to define how these data should be combined to form an experiment. Each created experiment is given a unique name and a description. The experiment designer provides a simple user interface with powerful functionality. With it, large experiments can be created that model multiple species using multiple algorithms. In the future, support will be added for modelling with multiple climate scenarios. In *openModeller Desktop*, the actual invocation of modelling processing, happens via a 'modelling plugin'. This means that the interaction between the Desktop interface and the *openModeller* library is very well separated. It also means that new modelling plugins can be written to support other modelling toolkits in the future. *openModeller Desktop* currently has two such plugins available: a 'local' plugin which uses libopenModeller directly, and a 'web services' plugin that can invoke modelling jobs on a remote server using the *openModeller* Web Services API. The latter is still somewhat experimental, but the aim is to allow the user to design large, CPU intensive experiments within the user friendly *openModeller Desktop* interface. The experiment will then be submitted via the Web Service API to run on a large cluster of powerful servers, with only the modelling outputs being returned to the client for local visualisation.

Once the experiment is completed, the main application window provides a treeview of the experiment. The top level node of the tree is the experiment itself. Beneath this, mid level nodes represent each algorithm profile used for the experiment. The terminal nodes of the tree represent the individual models. Clicking on a model node provides a de-

tailed report for the model. The model can also be visualised spatially using the map tab. The map tab uses an embedded GIS browser (based on Quantum GIS libraries) to allow the user to interactively pan and zoom around the geographic region of the output model.

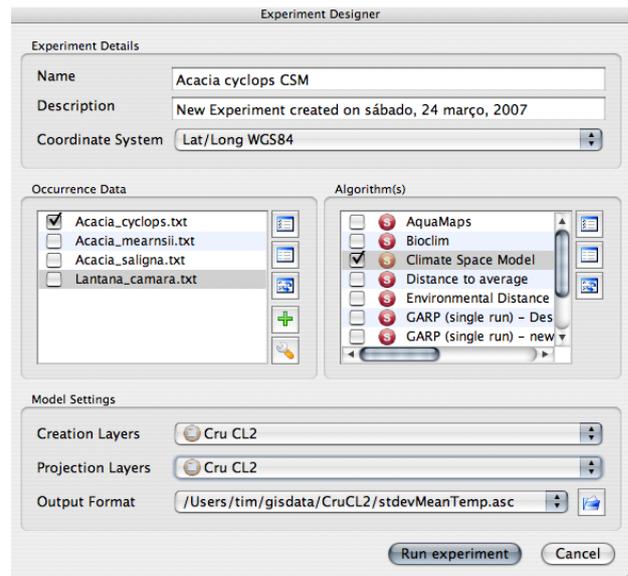


Figure 4: The experiment designer.

In the *Acacia cyclops* case study, Australia-only occurrence data was used to generate an fundamental niche model (Figure 4.) which was projected into a global climate dataset. In Figure 5, the results of this model are shown for south western Australia. The original occurrence points used to create the model are overlaid as black dots in the image. Areas in dark blue represent a poor match to the species' modelled fundamental niche. As the colours range through light blue, yellow and then to red, an increasingly good match to the species' modelled fundamental niche is shown. By inference, if the species is planted in areas demarcated in red, then it probably has a good chance of survival. In Figure 6, the same model output is shown, but the area of interest is set to southern Africa. Based on the model, we can postulate that the Western Cape region of South Africa represents a similarly optimal habitat for *Acacia cyclops*. *Acacia cyclops* is indeed a problematic invasive plant species in the Western Cape, having been introduced there to stabilise sand dunes (11).

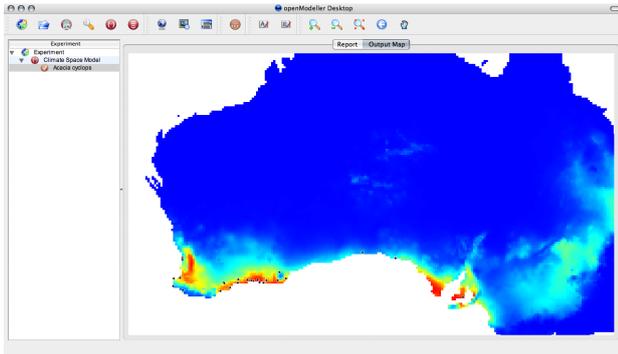


Figure 5: The model results for south western Australia.

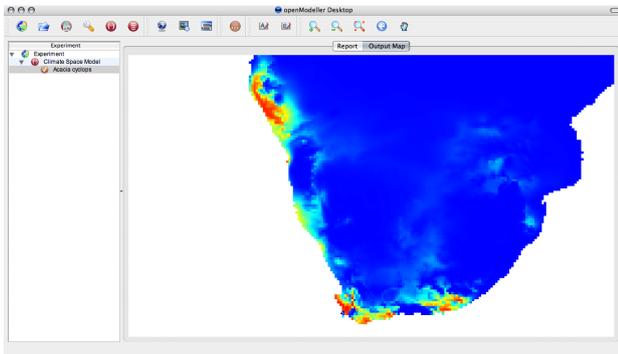


Figure 6: The model results for southern Africa.

Conclusion

openModeller provides solutions for users and developers interested in fundamental niche modelling. The *openModeller* library enables programmatic access to its capabilities. The console tools allow for scripting and interactive command line usage. The desktop application provides a user friendly graphical environment. The *openModeller* Web Services API provides a way to deploy *openModeller* on powerful servers (and in the near future, clusters). The plugin approach to implement different modelling algorithms allows the user to experiment with different modelling techniques without having to learn a new application each time, or having to prepare input data in a different way for each application. The algorithm plugin framework allows algorithm developers to focus exclusively on the core implementation of the algorithm without needing to 'reinvent the wheel' by writing data input / output routines and user interface logic.

The *openModeller Desktop* application provides an environment where all aspects of data preparation, experiment design, running an experiment and vi-

ualising the results of an experiment, can be carried out. *openModeller Desktop* can be compiled on all three major desktop operating systems (MS Windows, Mac OS X, GNU/Linux), and is freely available under the GPL.

The *openModeller* team is always eager to collaborate with other projects, users and developers. We encourage you to download and use the software and we welcome any feedback. We provide support via mailing lists and IRC - visit our home page for more details.

Bibliography

- [1] Council of Heads of Australian Herbaria. 2007. **Australia's Virtual Herbarium via Centre for Plant Biodiversity Research** In <http://www.cpbr.gov.au/cgi-bin/avh.cgi> (Accessed 25 March, 2007)
- [2] Brewer, P.W. 2003. **Modelling the global distribution patterns of Leguminosae species in past, present and future climates** PhD Thesis, University of Reading, UK.
- [3] Elith, J., Graham C.H., Anderson R.P., Dudik M., Ferrier S., Guisan A., Hijmans R. J., Huettmann F., Leathwick J.R., Lehmann A., Li J., Lohmann L.G., Loiselle B.A., Manion G., Moritz C., Nakamura M., Nakazawa Y., Overton J.McC., Peterson A.T., Phillips S.J., Richardson K.S., Scachetti-Pereira R., Schapire R.E., Soberon J., Williams S., Wisz M.S., Zimmermann N. E. 2006. **Novel methods improve prediction of species' distributions from occurrence data.** In *ECOGRAPHY* 29(2):129-151
- [4] Hutchinson, G.E. (1957). **Concluding Remarks.** In *Cold Spring Harbor Symposia on Quantitative Biology*. 22: 415-42.
- [5] Nix, H.A. (1986). **A biogeographic analysis of Australian Elapid snakes.** In Longmore, R. (ed.) *Atlas of Australian Elapid Snakes*. Australian Flora and Fauna Series 8: 4-15.
- [6] Peterson, A.T., Bauer, J.T., Mills, J.N. 2004. **Ecologic and Geographic Distribution of Filovirus Disease.** In *Emerging Infectious Diseases* 10(1):40-47

- [7] Phillips, S. J., Anderson, R. P., and Schapire, R. E., (2006). **Maximum entropy modeling of species geographic distributions.** In *Ecological Modelling*, 190, 231-259.
- [8] Robertson, M. P., Caithness, N., and Villet, M. H. 2001. **A PCAbased modelling technique for predicting environmental suitability for organisms from presence records.** In *Diversity and Distributions* vol.7: 15-27.
- [9] Scachetti-Pereira, R. (2002). **Desktop GARP.** In <http://lifemapper.org.br/desktopgarp> (Accessed on January 10th 2007).
- [10] Soberon, J., and Peterson, A. T. 2005. **Interpretation of Models of Fundamental Ecological Niches and Species' Distributional Areas.** In *Biodiversity Informatics*.
- [11] Stirton, C.H. (ed). 1983. **Plant invaders - beautiful but dangerous.** Department of Nature and Environmental Conservation, Cape Town.
- [12] Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B. F. N., de Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S., and Midgley, G.F., and Miles, L., Ortega-Huerta, M.A., Townsend, A., Phillips, O.L., and Williams, S.E. 2004. **Extinction risk from climate change** In *Nature*:427
- [13] Yesson, C., Culham, A. 2006a. **Phyloclimatic Modelling: Combining Phylogenetics and Bioclimatic Modelling.** In *Systematic Biology*: in press.
- [14] Yesson, C., Culham, A. 2006b. **A Phyloclimatic Study of Cyclamen.** In *BMC Evolutionary Biology*. 6:72

Tim Sutton

CRIA, Campinas, Brazil

<http://openmodeller.sf.net>

[tim AT linfiniti.com](mailto:tim@linfiniti.com)

Editor in Chief:Tyler Mitchell - [tmitchell AT osgeo.org](mailto:tmitchell@osgeo.org)**Editor, News:**

Jason Fournier

Editor, Case Studies:

Micha Silver

Editor, Project Spotlights:

Martin Wegmann

Editor, Integration Studies:

Martin Wegmann

Editor, Programming Tutorials:

Landon Blake

Acknowledgements

Various reviewers & the GRASS News Project

The *OSGeo Journal* is a publication of the *OSGeo Foundation*. The base of this newsletter, the $\text{\LaTeX}2^{\text{e}}$ style source has been kindly provided by the GRASS and R News editorial board. All articles are copyrighted by the respective authors. Please use the OSGeo Journal url for submitting articles, more details concerning submission instructions can be found on the OSGeo homepage.

Newsletter online: <http://www.osgeo.org/journal>OSGeo Homepage: <http://www.osgeo.org>

Mail contact through OSGeo, PO Box 4844, Williams Lake, British Columbia, Canada, V2G 2V8

ISSN 1994-1897