Niche modelling practical with openModeller

openModeller is an open source project to create a library of niche modelling algorithms on a single platform (http://openmodeller.sourceforge.net/). The latest version has more than a dozen algorithms, including implementations of Maxent, GARP, ENFA and BIOCLIM. We will be using two versions of this software, a command-line driven library of algorithms (openModeller) and the version with a graphical user interface (openModeller Desktop).

Input data

Niche modelling requires two data types as input i) distribution data for your taxa, ii) environmental/climate layers for the area surrounding your taxa.

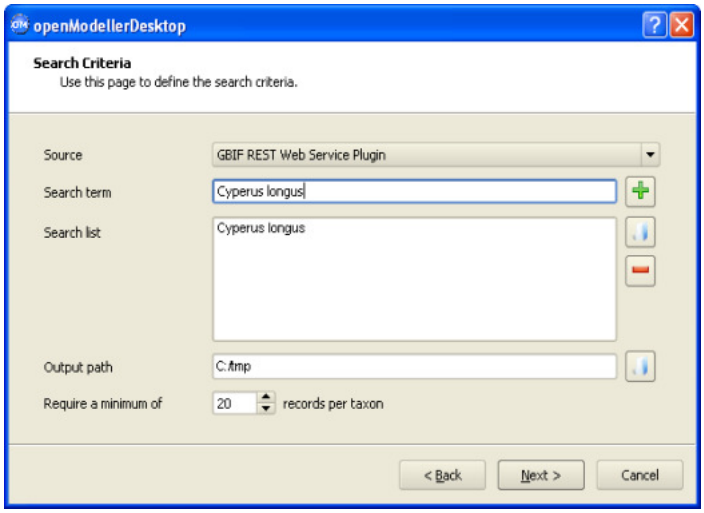
Distribution data

Distribution data can be found from a variety of sources such as museum specimen labels or literature searches. We are going to use the Global Biodiversity Information Facility (www.gbif.org) portal to access specimen data from Museums and Herbaria from around the world.

You can get data from GBIF either by direct download from www.gbif.org or using the openModeller search tool

Using the GBIF search tool in openModeller

- Click the "search for locality data" button.
- GBIF requires that you read and accept the terms and conditions of data use, click the tick box and OK.
- Type the name of your favourite taxon into the search term box and click the plus symbol.
- You also need to provide the path where you want to store the output.
- Click next to run the search.

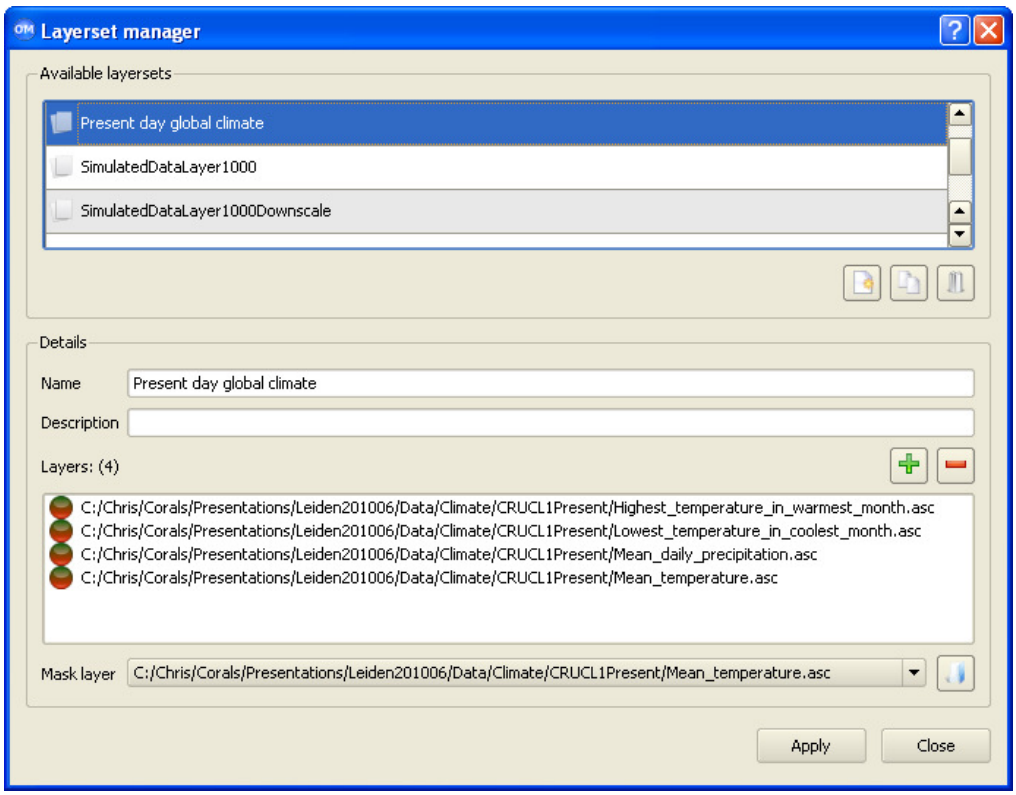


You should see a report of what has been found. You will get a text output and a GIS shapefile with the points. Note: this tool can break down if your species is very common (>10,000 points).

Environmental data

We need to prepare a collection of environmental layers which openModeller calls a 'Layerset'. When modelling lots of different species we often use the same set of environmental layers. The Layerset manager in openModeller lets you define this group of layers and also to give it a simple name like 'Present day global climate'.

- Click the layersets button to enter the layerset manager.
- Type in a name for your layerset "Present day global climate".
- Click the green plus to add layers to this layerset. Add all the layers (.asc files) in the "Data / climate / CRUCL1Present" directory. Once you click apply you should see your layerset appear in the "available layersets" box.
- Click close to finish.

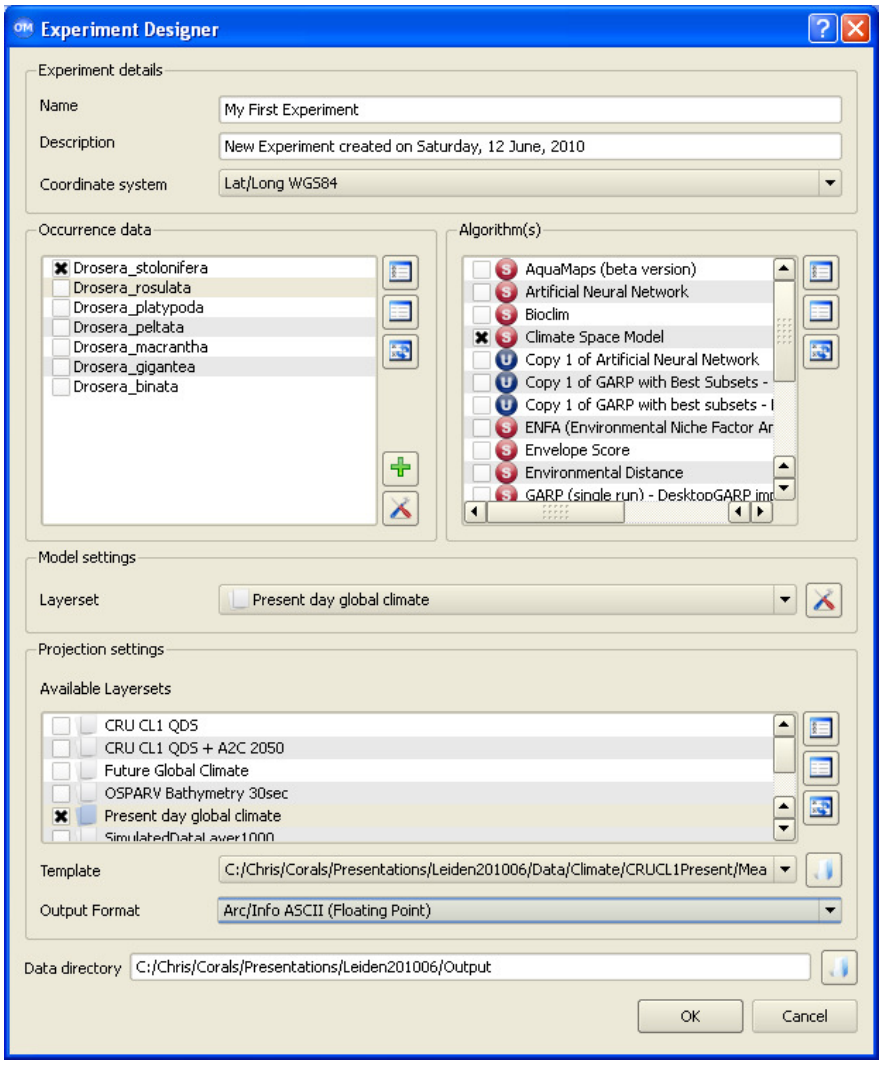


Where do I find environmental data?

There are lots of datasets to be found online. Here are a few to get you started. Temperature and precipitation can be found from the worldclim site (http://www.worldclim.org/). The highest resolution global elevation data is available at (http://topex.ucsd.edu/WWW_html/srtm30_plus.html). A limited number of future climate scenarios are available from the worldclim site, or you can get the original data from the IPCC data download website (http://www.ipcc-data.org/), but this data requires some processing to get into a format usable for niche modelling.

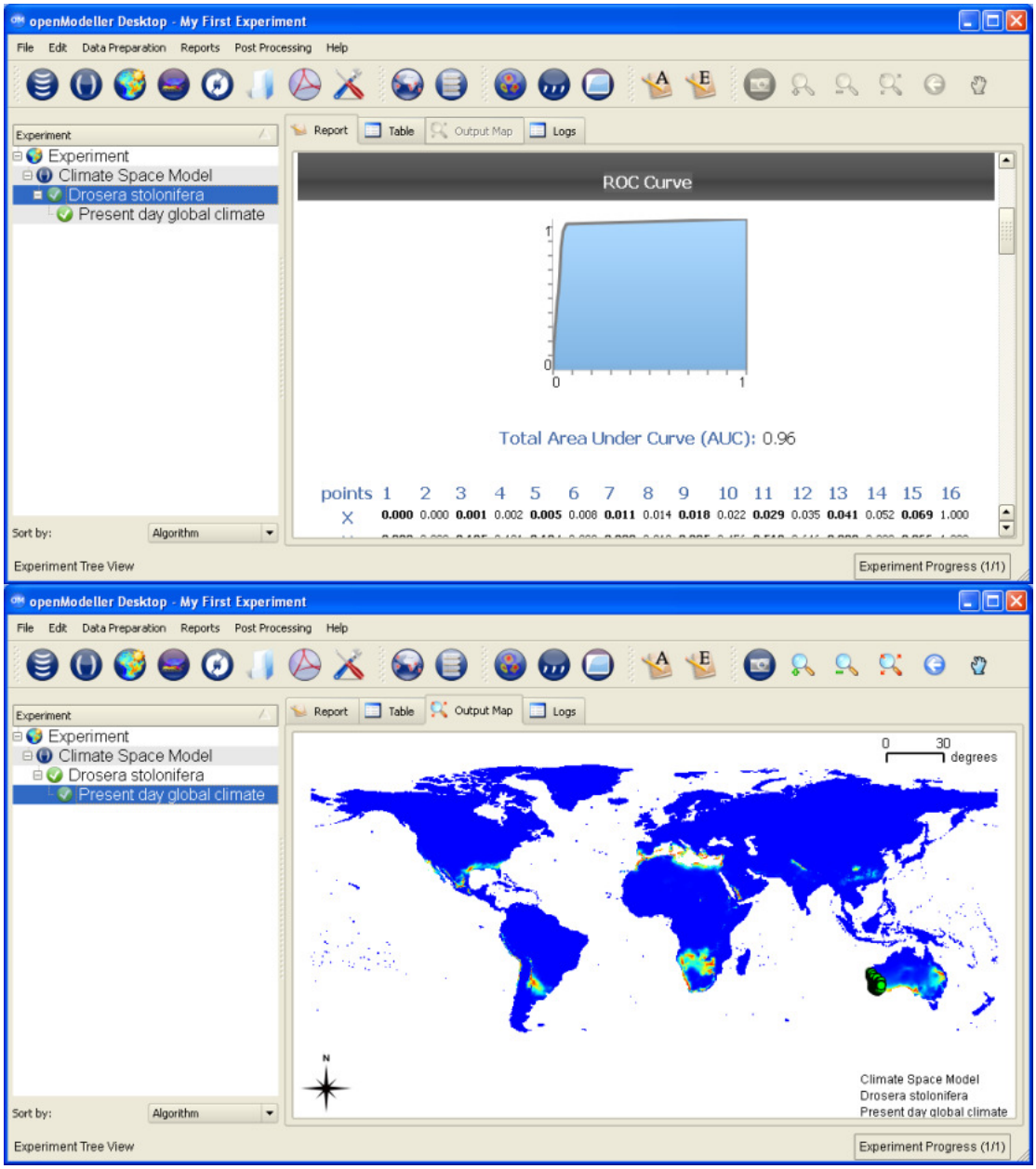
Running a model

- Click on the 'new experiment' button.
- In the "experiment designer" give your experiment a name (e.g. "my first experiment".
- In the 'occurrence data' section click the green + button to select your specimen localities file (the data you downloaded, or use the file "YessonCulham2006Localities.txt"). Your species name should appear in the occurrence data list. If you were to have multiple species in your localities file then these would all show up here.
- Check the box of species you want to model.
- In the Algorithms box you can tick one or more algorithms. Some algorithms have quite long compute times, so start with a fast algorithm - 'Climate Space Model'.
- The model settings section is used to specify which layersets are used to build and project. You only have one layerset defined at the moment so this should already be selected.
- Choose the Arc Info ASCII (floating point) output format
- Finally, in the data directory box set the place to store the output files
- Click OK to run the experiment



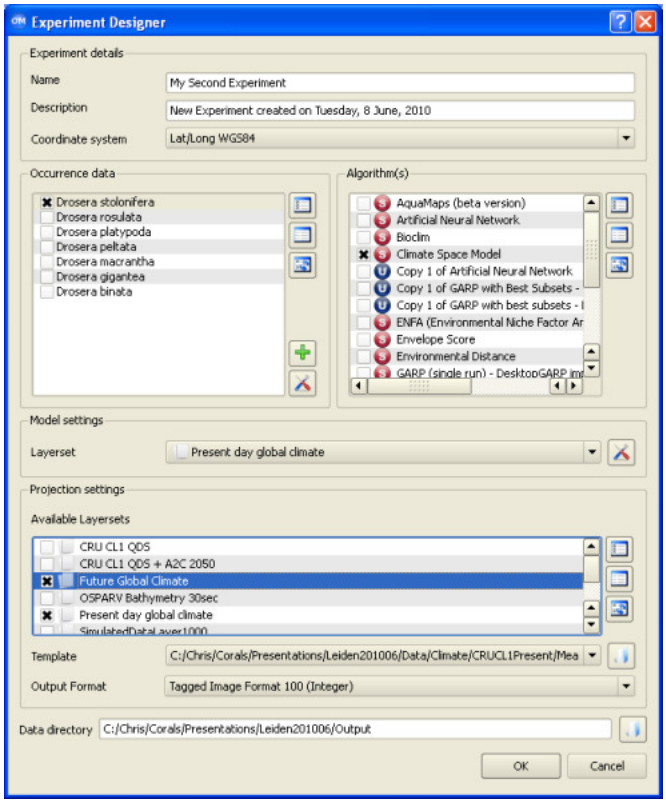
Your experiment queue should now open with a list of the models you've defined. If you've just selected one species and one algorithm there will be just one model listed. If you have five species and five algorithms this list would include all 25 model combinations.

The progress bar first shows the model building progress (in red) then the model projection progress (in green). Models with green ticks are complete. Click the taxon name under the model name on the left to view the model details including AUC validation. Click the layer name under this to view the output map.



GIS layers of your projected model output can be found in your chosen data directory.

Projecting a model into a different scenario



Often we want to take one model and project the model into different environmental layer, say a future climate scenarios to assess the impact of climate change on potential distribution. To do this we need the same set of environmental layers used to build the model, but for a different area/time frame. In this example we will use one of the IPCC future climate scenarios.

- Create a new layerset (as above), call the layerset "Future Global Climate" and add all files from the directory "Data \ Climate \ CRUCL12050A2c"
- Create a new experiment as before (call this one "My Second Experiment")
- Choose "Present Day Global Climate" for the model settings layerset
- In the projection layer box check both "Present Day Global Climate" and "Future Global Climate" in order to project the model into the present and future to compare the differences.

You should see two layerset names under your species name, click each one to see the model projected into that scenario.