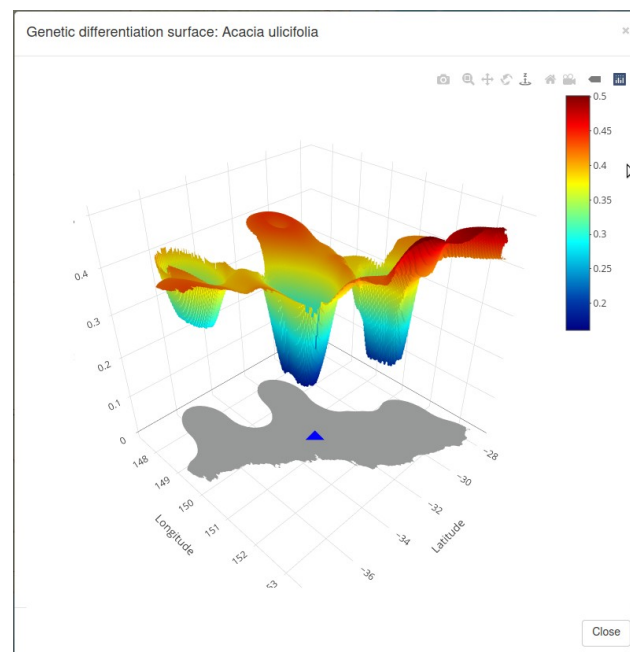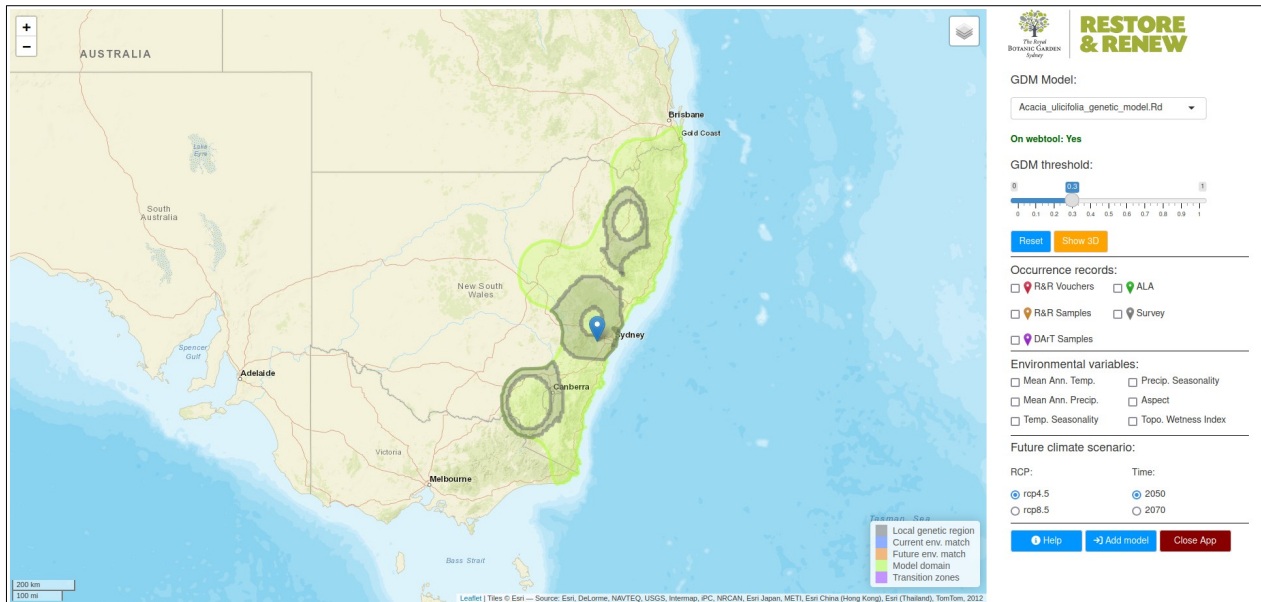# Restore & Renew Model Review Shiny App

Peter D. Wilson
Visiting Scientist
Research Centre for Ecological Resilience (ReCER)
Royal Botanic Garden, Sydney

2022-12-16

Introduction

In 2016, as the first experiments by Jason Bragg with fitting GDMs got underway, I began tinkering with an *R* Shiny app which would assist with the review of GDM models and also allow exploration of aspects of environmental site matching. These are two important features of the Webtool and I reasoned that a tool which allowed some degree of investigation of GDM performance and exploration of environmental data types, and their value for site matching, would be useful.

With the impending launch of the R&R webtool in mid-2018, I began an extensive revision of the first crude version. Although there seems to have been little (i.e. no known!) use of the app since I first made it available to the Evolutionary Ecology (now ReCER) team, I have pursued intermittent development of the app. The current version has advanced greatly since that first version.

Key attributes of the app include:

- Allows easy addition of new models
- Makes it possible to have multiple model version for a single taxon
- Site matching and genetic information in a single interface
- Switch seamlessly to pure site matching by selecting "no genetic model"

Steps for adding new models are explained in a following section.

Since it is written as an *R* application, it should be possible to run the app on any major computing platform including Windows (if you must), Linux and Mac (yeah for POSIX-based operating systems!). In addition to *R*, it needs to launch the web interface component of the app in a modern browser. I have successfully run it in Firefox and Chrome. I have not tested it in Microsoft's Edge and Apple's Safari browsers, but it should work OK in those environments.

My hope in expending effort on maintaining and expanding this tool is that, as we move towards a greatly increased rate of production for R&R, this tool may prove to be useful after all.

Installation

The app currently can be found on the ReCER team Sharepoint *Documents* folder within the *Restore and Renew* sub-folder. Navigate to the *Webtool maintenance* folder, and then drop into the *R-stuff* folder. You should download the *RandR_modelReview_app* folder.

Place the downloaded zip archive in a convenient folder on your PC and extract all contents.

Several steps must now be executed to enable the app to run in its new location:

1. Open the app script *app_Nov2022g.R* in *R*-Studio

2. Edit the base path at line 70 to show the folder into which you have extracted the application but ending with the "www" sub-folder. NOTE: the path should end in a final path separator.

3. Save the script

4. Open the *R*-script *set_covars.R* in *R*-Studio

5. Edit the base_path value to match the one you entered above, and save the script

6. Run the script

7. Open the *R*-script *clipDomainPolygon_v3.R* in *R*-Studio

8. Edit the base_path value to match the one you entered above, and save the script

9. Run the script

You should now be OK to start to app (see the next section).

Quick Use Guide

Start-up

- Open R-Studio

- Load the app file

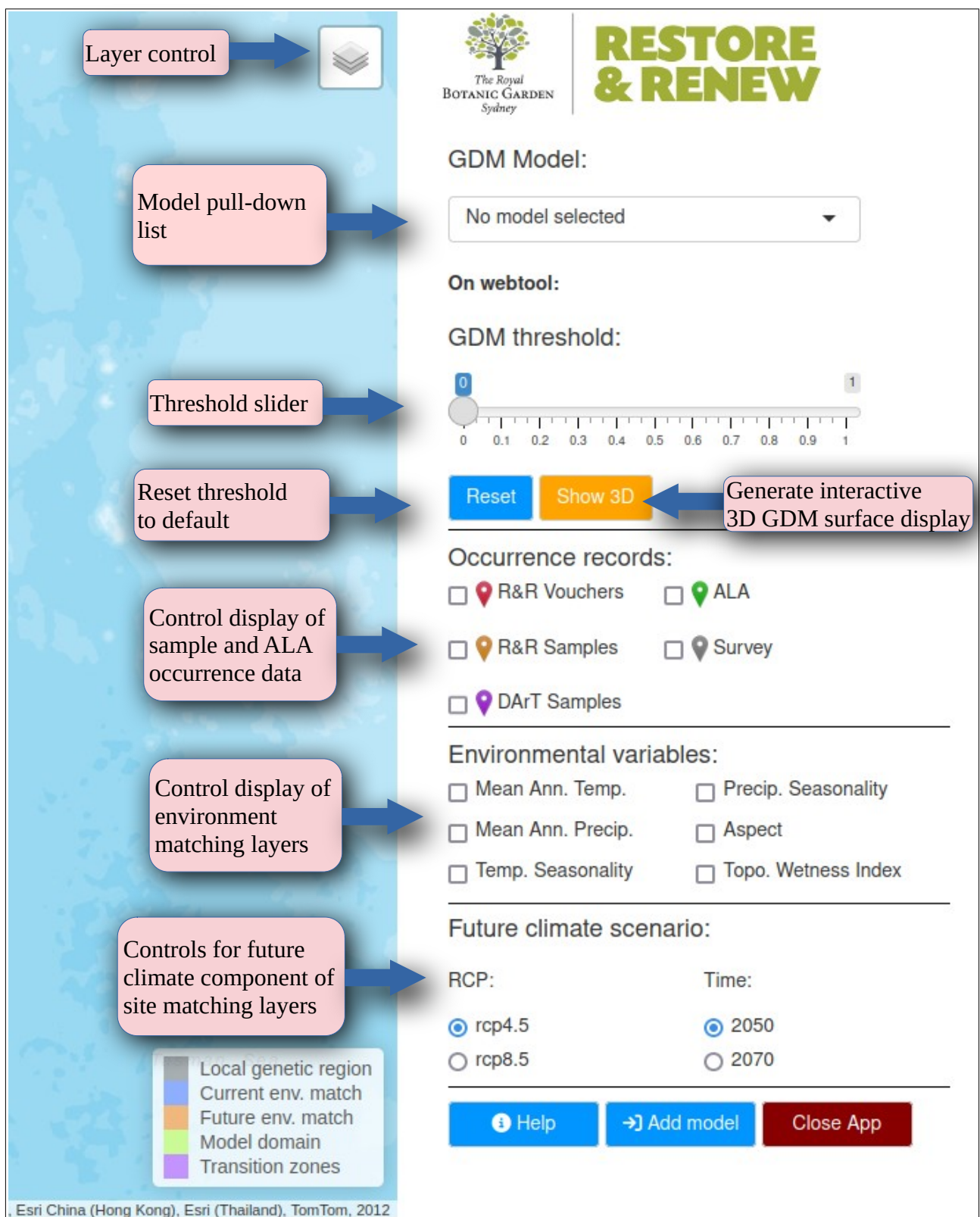- Click on the "Run app" link in the top right corner of the editor panel
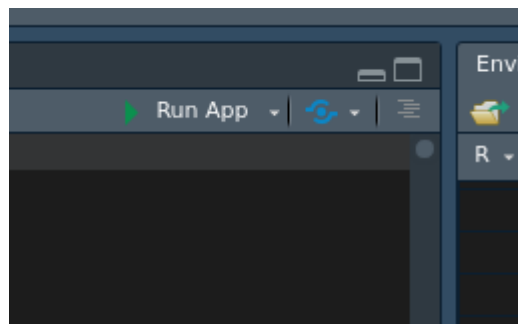
Controls and features

The table below summarises the major features of each control provided by the app. Please check the image on page 6 to help you relate the descriptive material to the app as seen on screen.

Element	Description
GDM Model	Model selection. At start-up "No model selected" is shown. The selected model may changed at anytime and displays are updated in response.
GDM Threshold	When a model is selected, the threshold value saved in the model object is displayed on the slider. After you have clicked on the map to select a location and the default thresholded genetic differentiation surface has been computed and displayed, you can move the slider to explore other threshold values. If you wish to return to the default threshold, click the "Reset" button. The "Show 3D" button will generate an interactive 3D surface representation of the un-thresholded surface for the selected location. Details of the controls available in the 3D display are provided in a separate section. **NOTE:** There is a lot of computing to generate the 3D graphic display and there will be a noticeable delay before the image appears in the popup window.

Element	Description
Occurrence records	Display markers for the indicated occurrence record types. A model must be selected for this to work.
Environmental variables	No site matching layers are computed and displayed until at least one variable environmental variable.
Future climate scenario	Allows selection of the future climate scenario and time.
Help	Opens a small help dialogue.
Add model	Specialised help on general steps to adding a new model. More of a reminder as the many steps involved are better described in this document.
Close App	Closes the app. **PLEASE USE THIS BUTTON TO CLOSE THE APP**.
Layer control	A checklist controlling the layers to be displayed. This control turns the selected layer on or off but does not generate re-computation of any layers. This allows your, for example to turn off site matching layers without having to laboriously uncheck all checked environmental variables.

Layer control

Model pull-down list

Threshold slider

Reset threshold to default

Generate interactive 3D GDM surface display

Control display of sample and ALA occurrence data

Control display of environment matching layers

Controls for future climate component of site matching layers

The Royal
BOTANIC GARDEN
Sydney

RESTORE & RENEW

GDM Model:

No model selected

On webtool:

GDM threshold:

| 0 | | | | | | | | | | 1 |
| 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |

Reset   Show 3D

Occurrence records:

☐ 📍 R&R Vouchers   ☐ 📍 ALA

☐ 📍 R&R Samples   ☐ 📍 Survey

☐ 📍 DArT Samples

Environmental variables:

☐ Mean Ann. Temp.   ☐ Precip. Seasonality

☐ Mean Ann. Precip.   ☐ Aspect

☐ Temp. Seasonality   ☐ Topo. Wetness Index

Future climate scenario:

RCP: Time:

◉ rcp4.5   ◉ 2050

◯ rcp8.5   ◯ 2070

ⓘ Help   →] Add model   Close App

Local genetic region
Current env. match
Future env. match
Model domain
Transition zones

, Esri China (Hong Kong), Esri (Thailand), TomTom, 2012

File storage

All files used by the app are found in the home folder you created when downloading and installing the app. The main *R*-script (*app.R*) and a few utility *R*-scripts (described in sections below) are found in the home folder. Data and other resource files (including GDM model files) are found in folders below the "www" folder which you will see in the home folder.

```
www/resources

www/models
      www/models/domain
      www/models/gdm
      www/models/zones

www/images

www/qData/eastOZ

www/envData/eastOZ
```

Installing a new model

Installing a new model to use in the app requires a number of steps to ensure that paths to covariate layers are correctly referenced in the model object, and that the file name used fits the requirements of the app. Updates to two ancillary files are also required, The first, updating *www/resources/onWebtool_lookup.csv*, is required for the app the function correctly. The second, updating the SQLite database which supplies occurrence information, is not critical. However, if these data are not updated then you may not be able to assess if the GDM and tissue sampling have been adequate.

Model object name

Change the name of the GDM object file to begin with the taxon name with spaces replaced by underscores. For example, the original GDM model file might be something like:

```
BankSerr_genetic_model.Rd
```

This should be changed to:

```
Banksia_serrata_genetic_model.Rd
```

Please be sure to check the spelling of the taxon name when renaming this file because the taxon name component of the file name is used to align the model with other important data components.

> The part of the file before "`_genetic_model.Rd`" must be unique.

If you wish to load alternate models for the same taxon, you may add elements to the file name after the taxon name. For example, if you wish to compare alternate models based on 2 or 3 ancestral groups, then use something like the following naming pattern:

```
Acacia_linifolia_K2_genetic_model.Rd

Acacia_linifolia_K3_genetic_model.Rd
```

The file name will appear in the pull-down list just as you have entered it here so you can display results for each variant.

Model object files should be copied to the gdm folder (i.e. /www/models/gdm).

Changing covariate path names

Paths to covariate files needed to evaluate a model are stored as fields in the model object. The names of any Q-covariate files must be changed to match the name used for the GDM model object file. For example,

```
ActiHeli_Qprops1.tif
```

should be changed to:

```
Actinotus_helianthi_Qprops1.tif
```

Q-covariate files for a new model should be copied to the folder *www/qFiles/eastOZ*.

Then, run the helper *R*-script *set_covars* to update the paths to fit the folder structure used by the app.

You can, if you wish, simply run the script as you did at installation. This will regenerate ALL Q-covariate entries for all models found in the */www/models/gdm* folder. This can be time consuming, particularly as the number of models increases.

Alternatively, you can edit the line which lists all files in the gdm folder with the FULL path to just the file you wish to update. For example,

```
model_files ← "/path_to_files/www/models/gdm/Greenus_plantus_genetic_model.Rd"
```

Generating a web-friendly domain polygon

The mapping component of the app uses an Industry-standard mapping engine called Leaflet. Like all Javascript applications which supply a WMS (Web Mapping Service) it needs polygons for display to be available in geoJSON file format. The R-script *clipDomainPolygon* which you meet when yo installed the app, is used to do several useful things:

- Look at the GDM model object and load the embedded model domain raster (each model domain is unique)
- Convert the raster into a polygon (i.e. trace the boundary of non-NA grid cells)
- Clip the polygon where it extends beyond the coastline
- Save the clipped polygon in geoJSON format

You can, if you wish, simply run the script as you did at installation. This will regenerate ALL polygons for all models found in the */www/models/gdm* folder. This can be time consuming, particularly as the number of models increases. You can alter the script to parse a single model as shown above to covariate file updates.

Update onWebtool.csv table

For each new model you add to the app you will need to add an line in the file
www/resources/onWebtool_lookup.csv. This can be done in a text editor or using a spreadsheet
application. If using the later, please be sure to save as a comma separated value (csv) file. Here are
a few lines from valid entries to give guidance on formatting new entries:

taxon	onWebtool
Banksia serrata	No
Banksia serrata K3	Yes

The name in the *taxon* column matches the name part of the GDM file name which has been
described in an earlier section (without the underscores in place of spaces). The entries in the
onWebtool column indicate if the model is currently posted to the webtool. Values here control the
display of the advisory prompt in the app control panel just below the model selection control.

Updating occurrence data

Data in the SQLite database (located in *www/resources*) uses the base taxonomic name. So, for
example, GDM models *Banksia_serrata_genetic_model.Rd* and
Banksia_serrata_K3_genetic_model.Rd will be associated with the same set of occurrence data
linked to *Banksia serrata*.

This implies, of course, that you will only need to add to the database when an entirely new taxon is
added.

To ease the burden as much as possible, the database has been pre-loaded with data for voucher and
tissue samples for all 1,323 named forms present in the ReCER database (exported data on 15
November 2022).

If you want to add ALA-derived data for a new taxon, then the process is more complex. The
procedures and supporting *R*-packages are used in preparing species entries for posting to the
webtool, and are described in the Webtool Maintenance Manual. If you wish to work through that
process, that is OK. If not, please contact me if you would like supporting obtaining these data.

Using the 3D model viewer

The 3D model view activated by clicking on the *Show 3D* button allows you to assess the form taken by the surface representing genetic differentiation across the model domain. The graphical display uses the *plotly* Javascript library and is a highly flexible, interactive display. This section gives you a quick overview of the most useful manipulations which can be made to the display.

What is displayed?

There are three key components to the display:
- 3D genetic differentiation surface to the selected species at the selected location
- The model domain as a light grey area on the XY-plane (below the surface)
- The selected location shown as a blue diamond embedded in the XY-plane
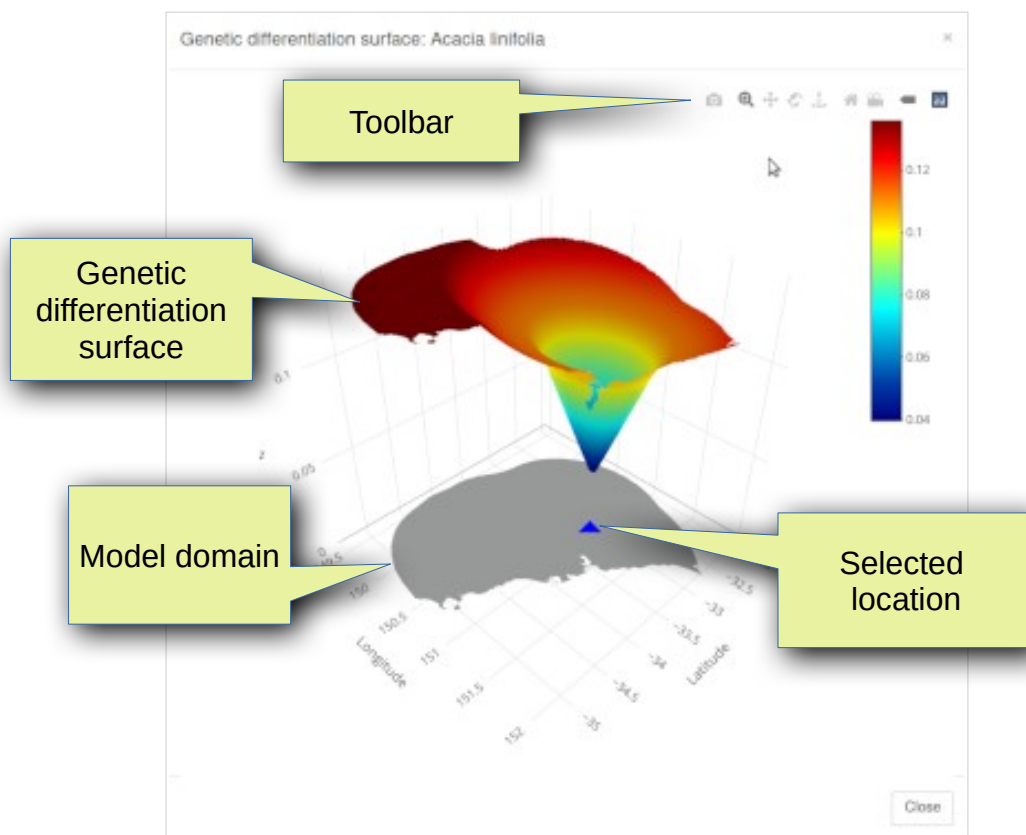
Toolbar

When you move the mouse pointer over the 3D display, you will notice a toolbar appear in the top right corner of the plot. A few tools are very useful in the context of reviewing the 3D surface.

Home

There are many manipulations you can make to the display so that returning to the original view to "unscramble" the display it tedious and rather inaccurate. Fortunately, clicking on the "Home" button returns all settings to the default state and returns the display to the initial or default view.

Zoom

Zooming in or out can be accomplished by two methods:
1. Rotating the mouse scroll wheel. This is active in other modes such as rotation; or,
2. Clicking on the zoom option on the toolbar. Now, holding the left mouse button zooms in as you drag the mouse "up", and zooms out as you move the mouse "down". This zoom mode remains active until to click on another tool on the toolbar.

Rotation

The 3D surface can be freely and dynamically rotated around all axes. When the display is first created, the rotation tool is active. You can switch back to rotation mode at any time by clicking on the Rotation tool on the toolbar.

To freely rotate the image around all three axes, ensure that the rotation mode is selected, and hold the left mouse button while dragging the mouse.

In this mode, you can zoom in and out at the last orientation by releasing the mouse button and moving the mouse scroll wheel.

Save a snapshot image

The camera icon (first icon on the toolbar) lets you save the current view as a PNG image.

Relationship with Restore & Renew Webtool

A quick outline about how the model review app relates to the R&R Webtool. Hopefully, you will be able to confirm that the way the model review app works mirrors the essential aspects of the R&R Webtool. And you will be able to map your discoveries using the app to motivate appropriate adjustments to species settings on the Webtool, or improvements to upstream processing such as generating GDMs and other resources for posting on the Webtool.

What is the same?

Some fundamental components of the code are identical to the code used in the webtool. This includes:

- GDM models and associated map layers (e.g. model domain polygon)
- Code to compute GDM surface and polygons
- Code to compute environmental match and the environmental data used for that feature

There is therefore absolute certainty that the results generated in the app are exactly as they would be seen on the webtool.

What differs?

The app differs in some important respects to the webtool:

- Ability to dynamically adjust the threshold used to truncate the GDM surface
- 3D model surface display
- Display the full range of site matching layers on the same display as the GDM results (the Webtool separates them at the moment into two distinct modules and only displays a very simple 2-variable climate match on the webtool provenancing module)
- "Special Zone" information is not displayed (but will be added in a future update).

Future developments

Some ideas come to mind for further development of this tool:

- Ability to extract sample data directly from the ReCER database if there is ever a legitimate API created
- Ability to show tissue samples cleaned/filtered and used in GDM model creation by loading metadata tables used in GDM fitting adding an appropriate category to the displayable occurrence records
- Adding additional or updated environmental data layers to enhance the ability to explore their utility and potential for inclusion in the webtool (e.g. ability to compare and contrast NARCLiM 2.0 data against other CMIP6 downscaled future climate data such as the widely-used CHELSA data)