ENM Special Topics

Clayton J. Visger and Charlotte Germain-Aubrey
Florida Museum of Natural History
University of Florida

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. All images used with permission or are free from copyright.

Topics

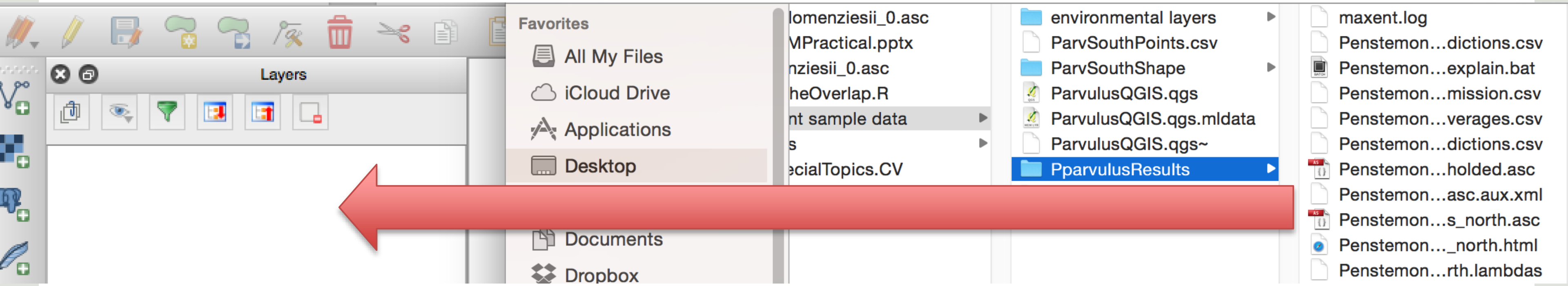
- QGIS
- Quantifying niche overlap
- PRISM
- Alternative Modeling Tools

QGIS

- Open-source alternative to ArcGIS
  - OSX compatible
  - Get it here http://www.qgis.org/en/site/
- We will cover
  - How to import MaxEnt output
  - Apply a basemap
  - Colorize and and alter transparency of the layers

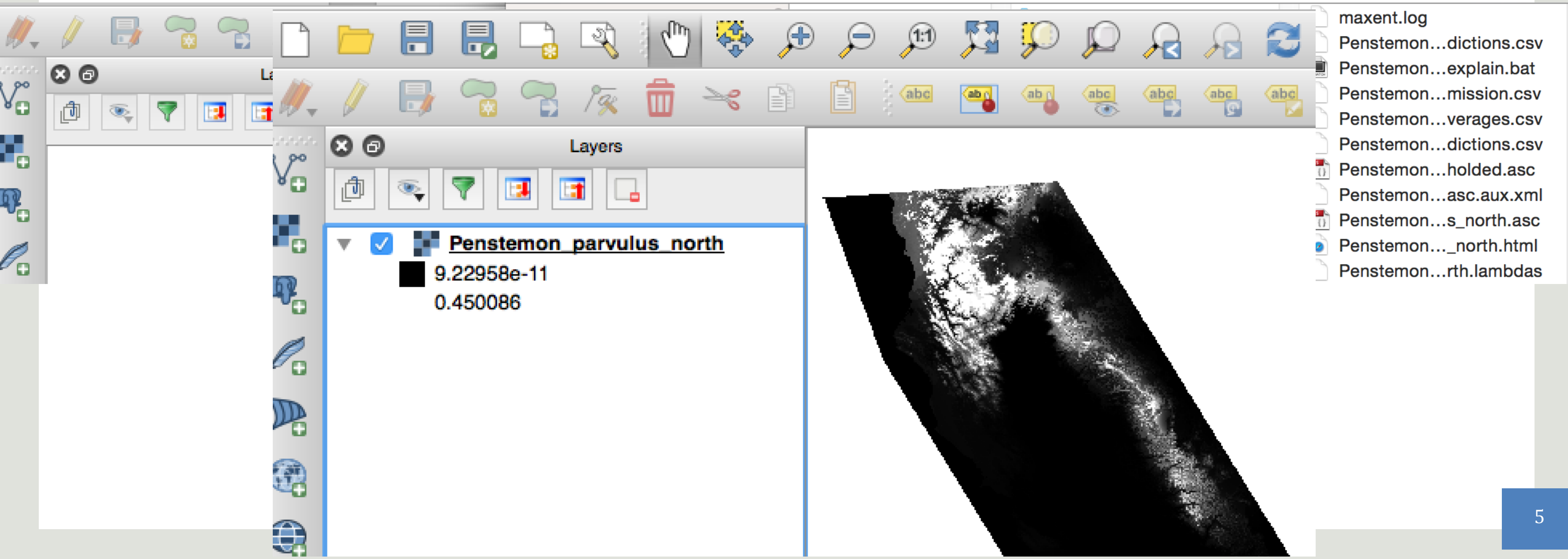
QGIS Map making

- Importing MaxEnt output
  - – Look for the .asc
    - There may be more than one
  - – Drag it in to the layer manager!



QGIS Map making

- Importing MaxEnt output
  - – Look for the .asc
    - There may be more than one
  - – Drag it in to the layer manager!



maxent.log
Penstemon…dictions.csv
Penstemon…explain.bat
Penstemon…mission.csv
Penstemon…verages.csv
Penstemon…dictions.csv
Penstemon…holded.asc
Penstemon…asc.aux.xml
Penstemon…s_north.asc
Penstemon…_north.html
Penstemon…rth.lambdas

☑ Penstemon_parvulus_north
■ 9.22958e-11
0.450086

QGIS Map making

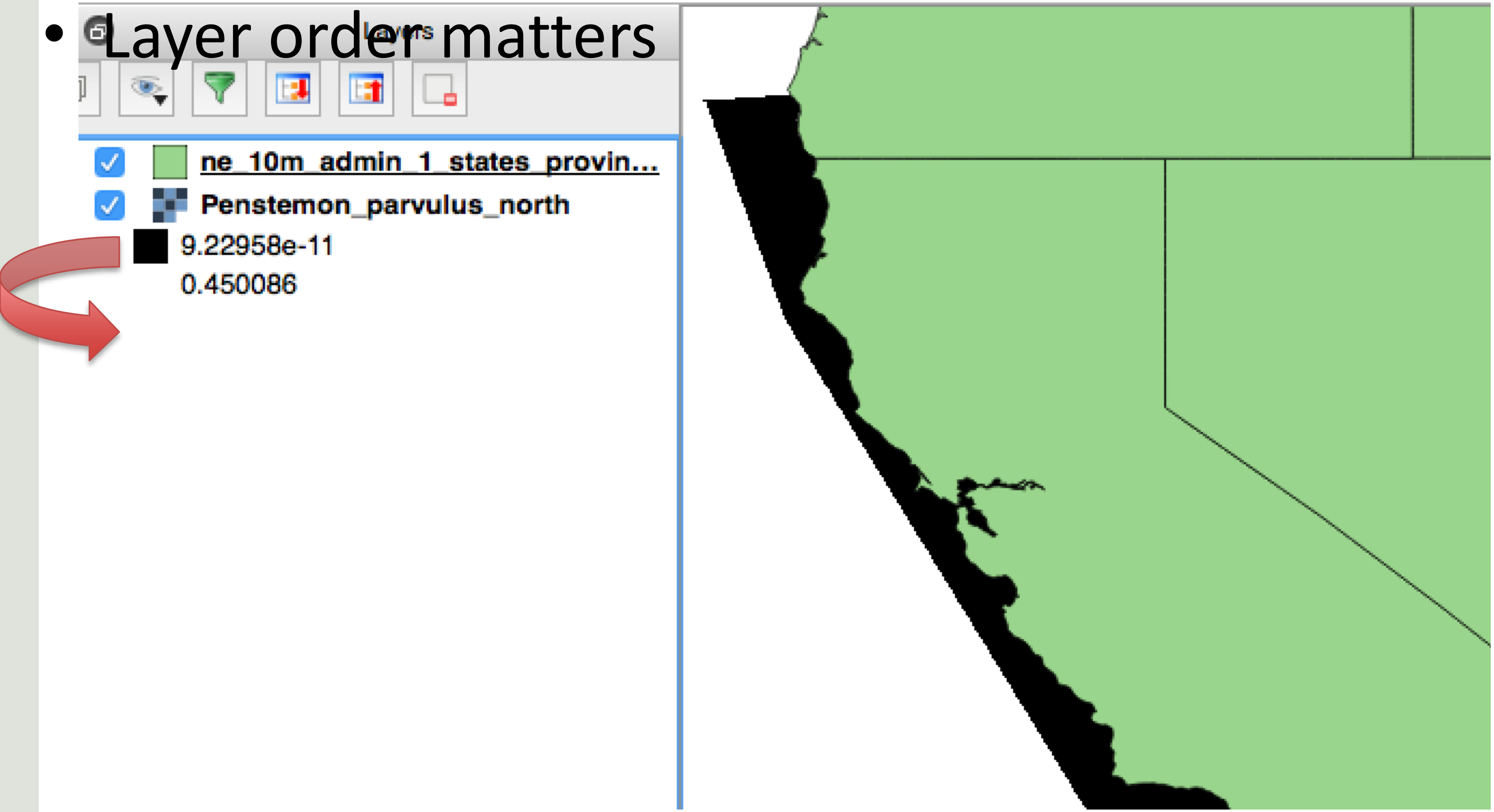
- Base Map
  - A few options here
    - Elevation layer from Worldclim
    - Natural Earth layer sets (What I'm using)

      www.naturalearthdata.com/

  - Drag it in!

QGIS Map making

- Layer order matters



☑ 🟩 ne_10m_admin_1_states_provin...

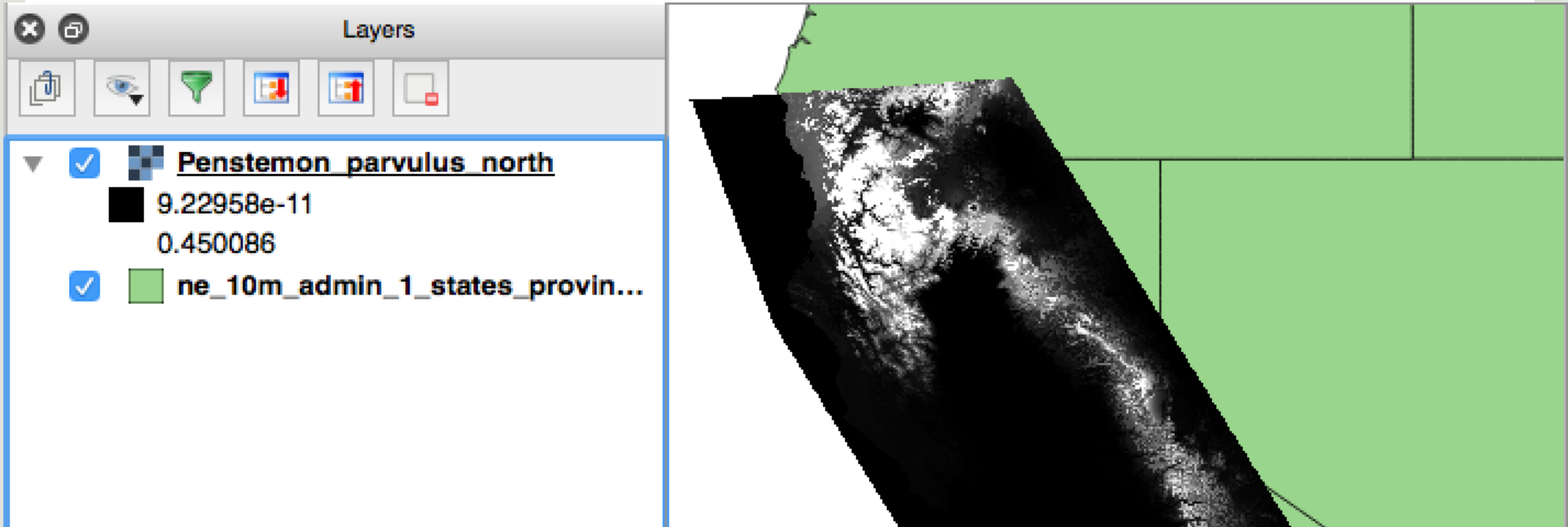
☑ Penstemon_parvulus_north

■ 9.22958e-11

0.450086

QGIS Map making

- Layer order matters
  - Still not looking great…



☑ Penstemon_parvulus_north

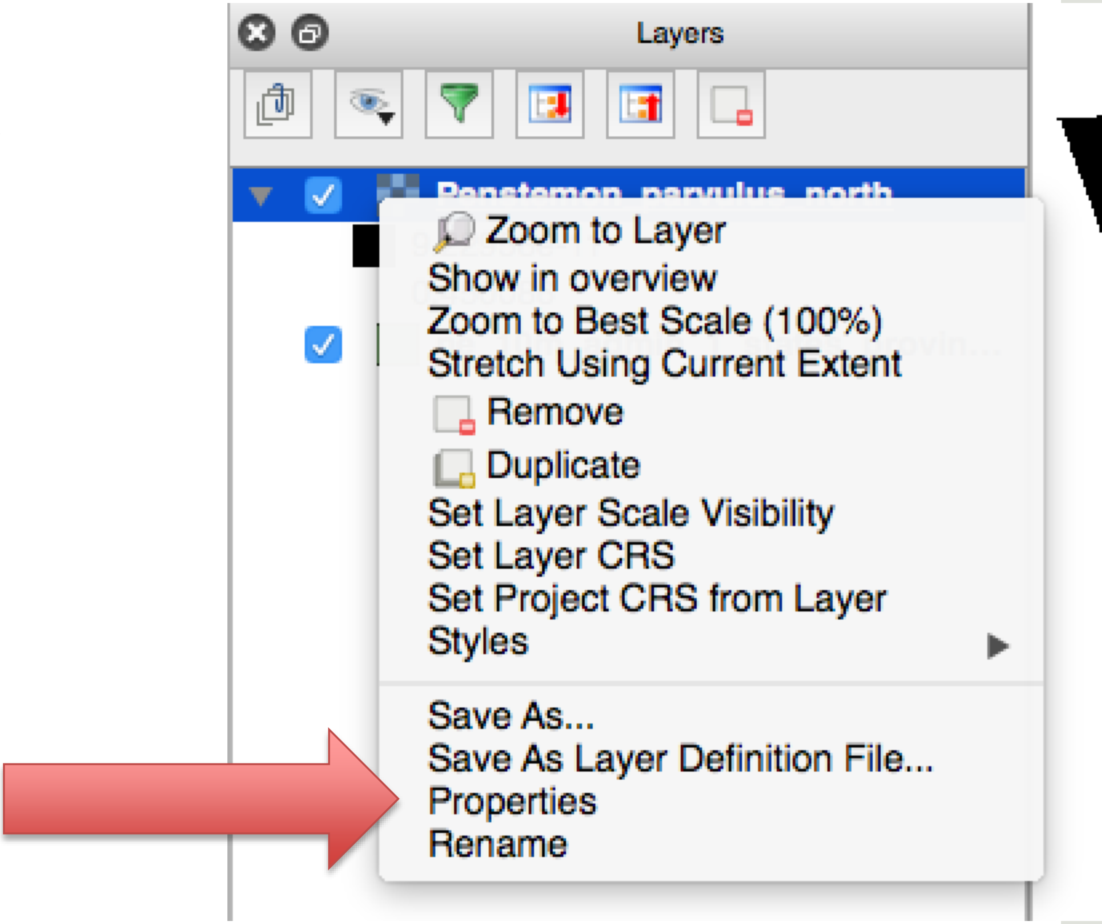
■ 9.22958e-11

0.450086

☑ 🟩 ne_10m_admin_1_states_provin...

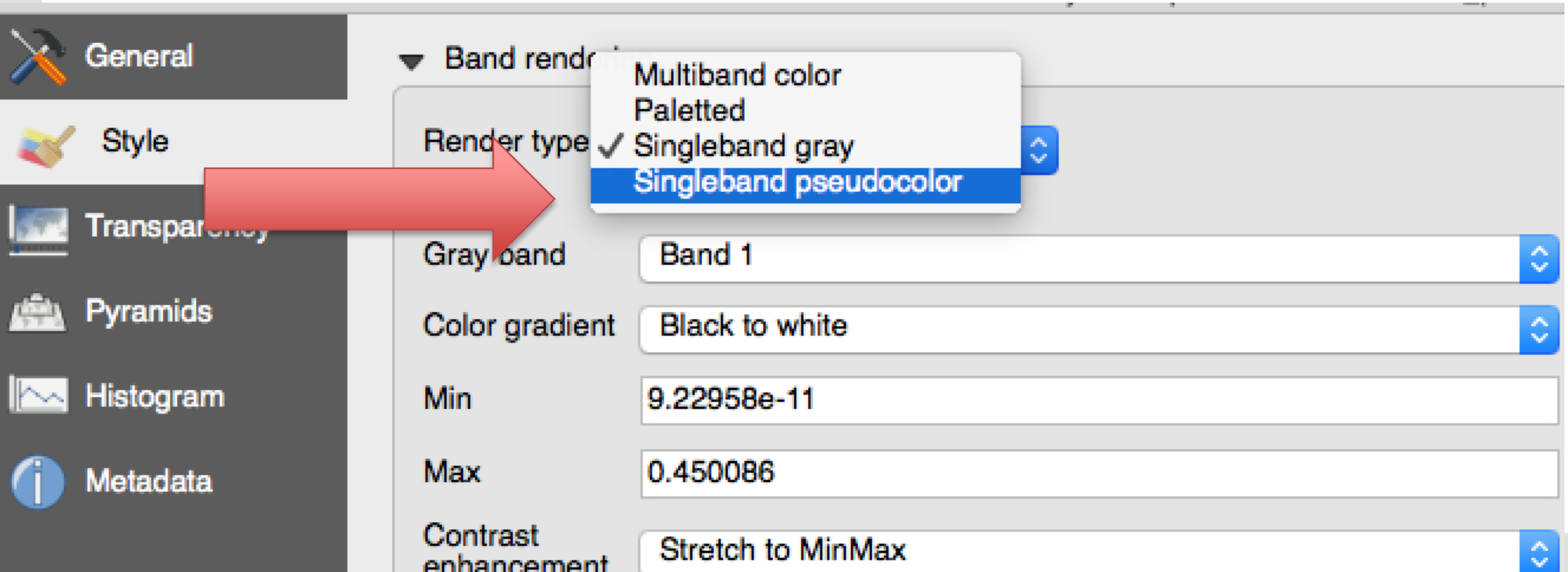
QGIS Map making

- Editing color and transparency
  - Rclick layer > properties



QGIS Map making

- Editing color and transparency
  - Style>Render type > Single band pseudocolor



Render type: Singleband pseudocolor

Gray band: Band 1

Color gradient: Black to white

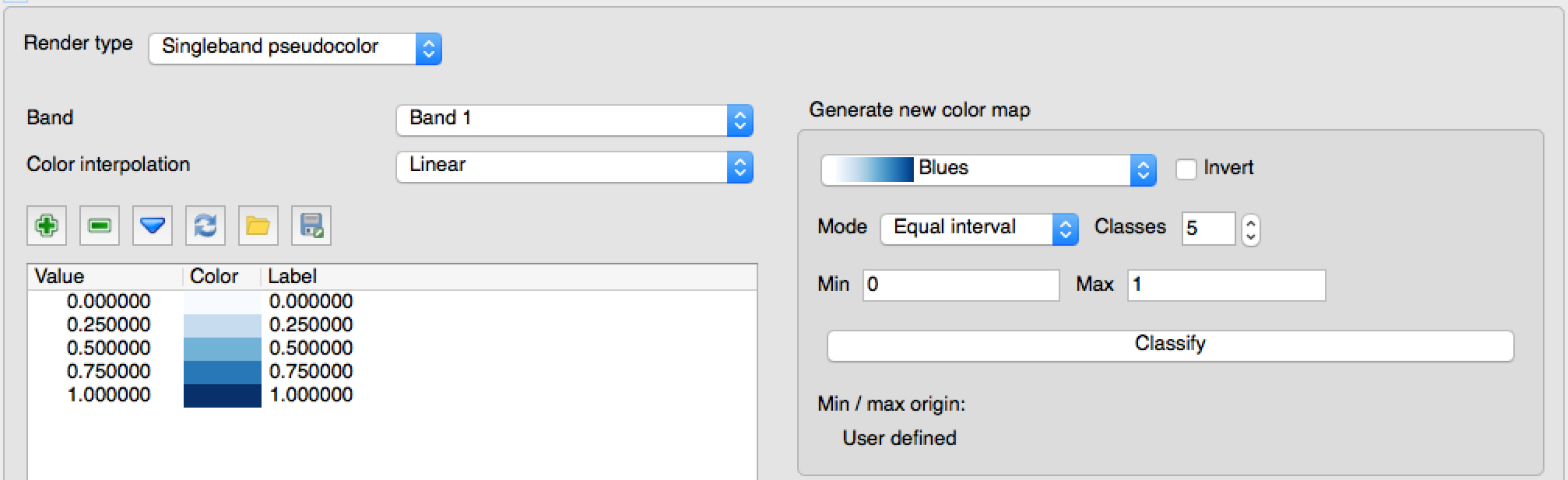
Min: 9.22958e-11

Max: 0.450086

Contrast enhancement: Stretch to MinMax

QGIS Map making

- Editing color and transparency
  - Pick a colorset
  - Equal interval, min =0 max =1, classes up to you
  - Push Classify



Render type: Singleband pseudocolor

Band: Band 1

Color interpolation: Linear

Value	Color	Label
0.000000	■	0.000000
0.250000	■	0.250000
0.500000	■	0.500000
0.750000	■	0.750000
1.000000	■	1.000000

Generate new color map

Blues ☐ Invert

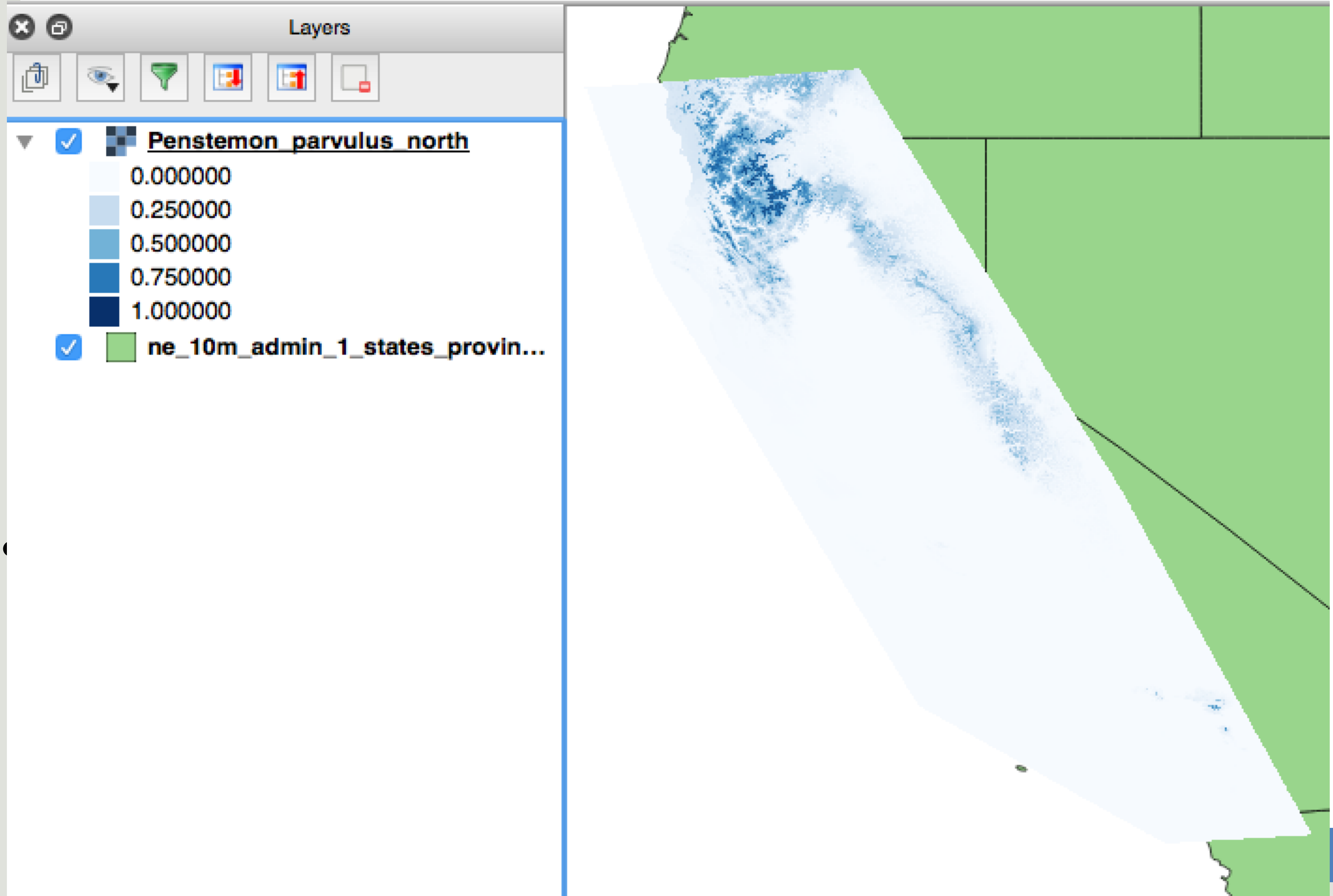
Mode: Equal interval Classes: 5

Min: 0 Max: 1

Min / max origin:

User defined

QGIS Map making

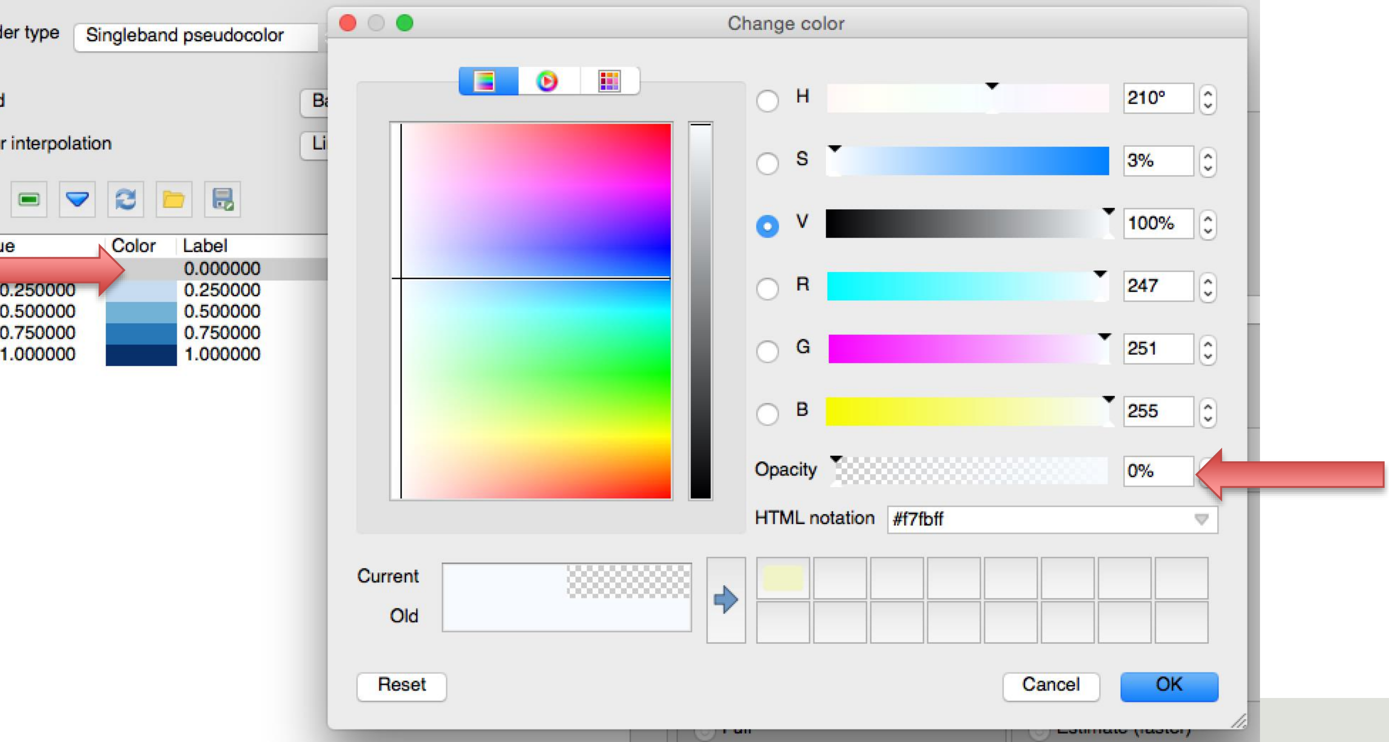


☑ Penstemon_parvulus_north
- 0.000000
- 0.250000
- 0.500000
- 0.750000
- 1.000000

☑ ne_10m_admin_1_states_provin...

QGIS Map making

- At the style menu double click the color and lower opacity to 0%



ler type: Singleband pseudocolor

r interpolation

ue	Color	Label
		0.000000
0.250000		0.250000
0.500000		0.500000
0.750000		0.750000
1.000000		1.000000

Change color

○ H 210°

○ S 3%

● V 100%

○ R 247

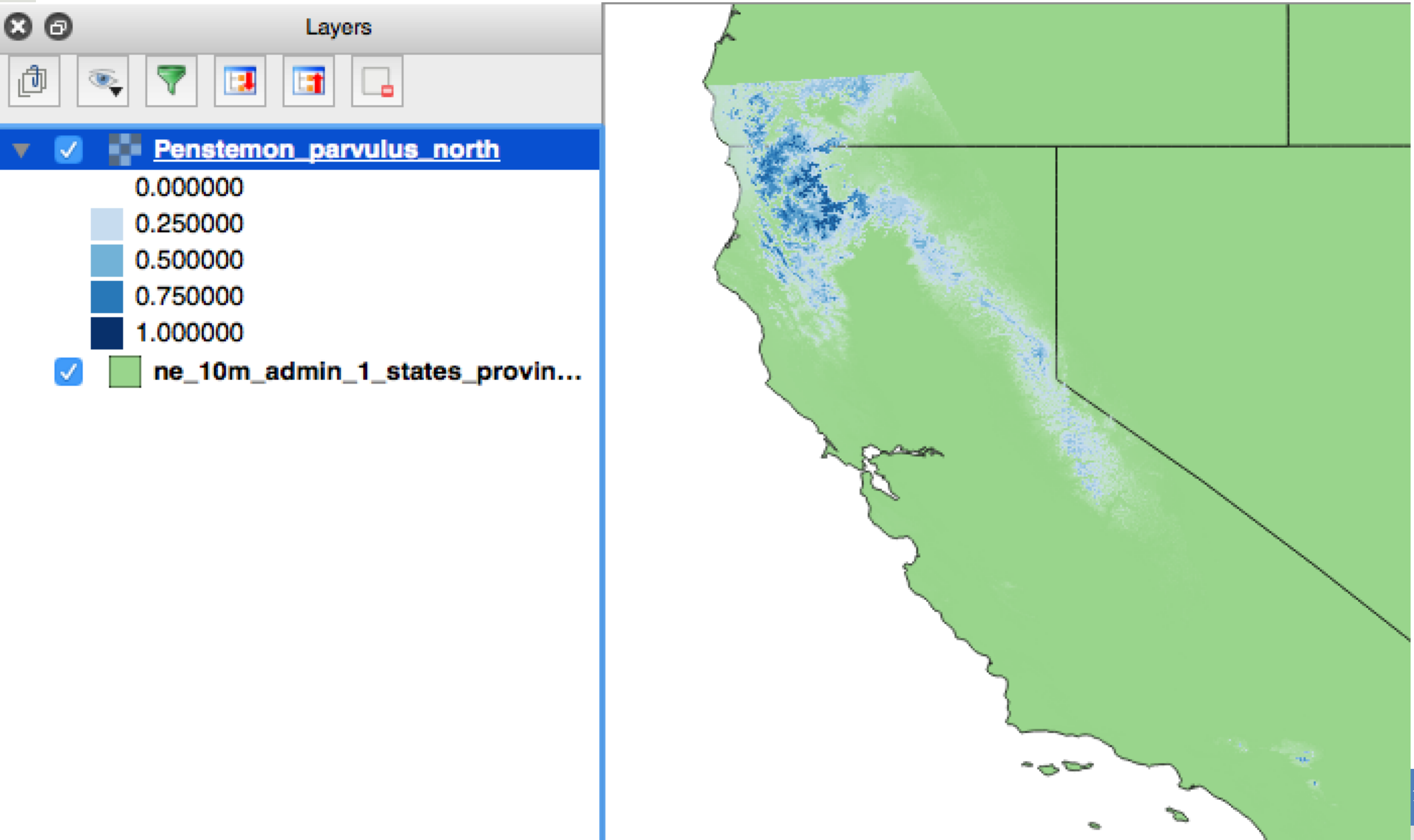
○ G 251

○ B 255

Opacity 0%

HTML notation #f7fbff

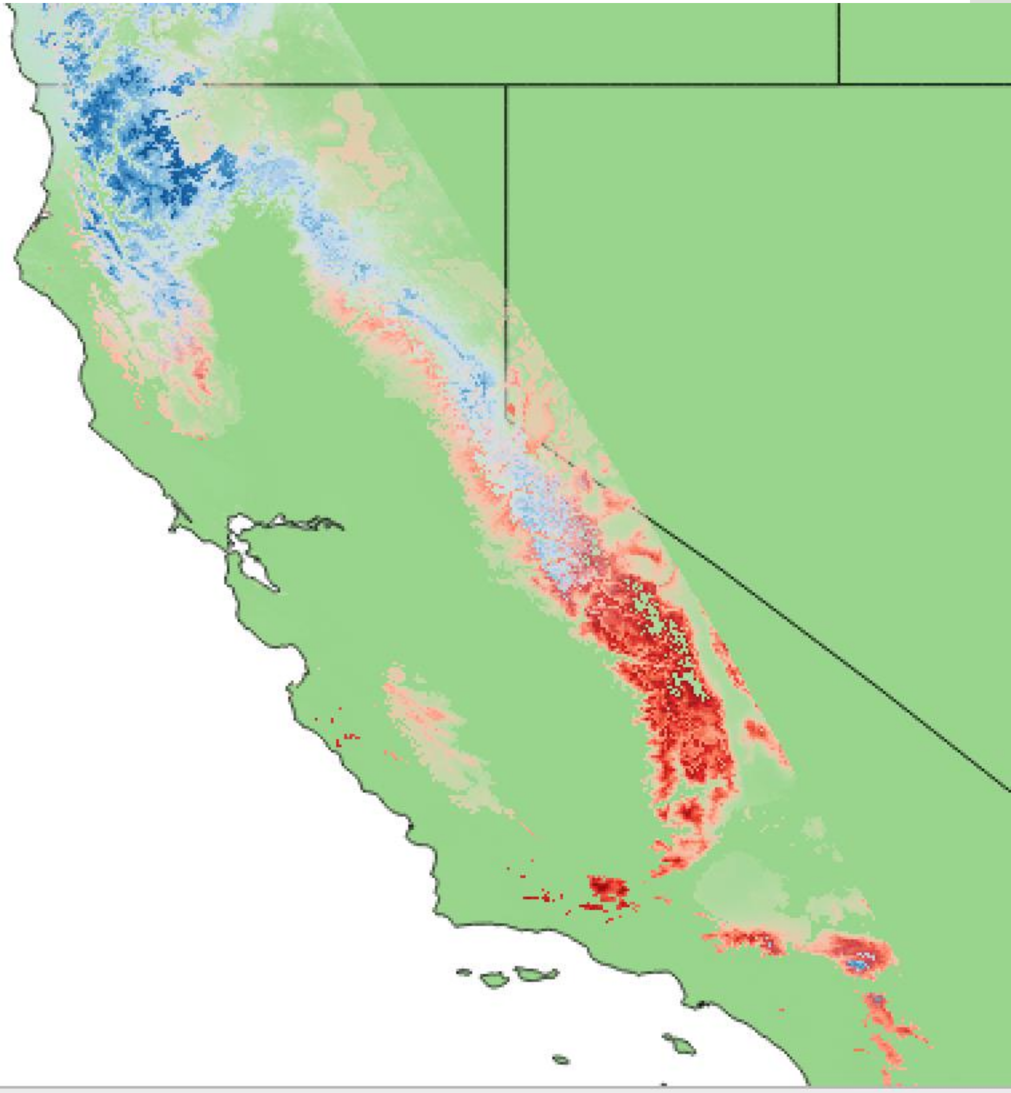
QGIS Map making



☑ Penstemon_parvulus_north
0.000000
0.250000
0.500000
0.750000
1.000000
☑ ne_10m_admin_1_states_provin...

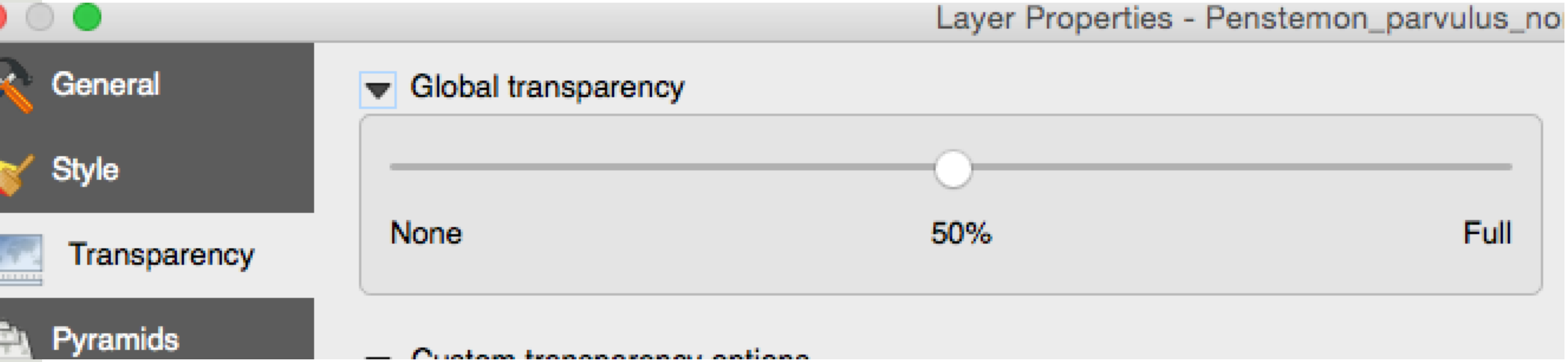
QGIS Map making

- Lets add in a second dataset
  - Hard to tell what is under the blue distribution



QGIS Map making

- Layer properties > Transparency =50%
  - Do this for both layers



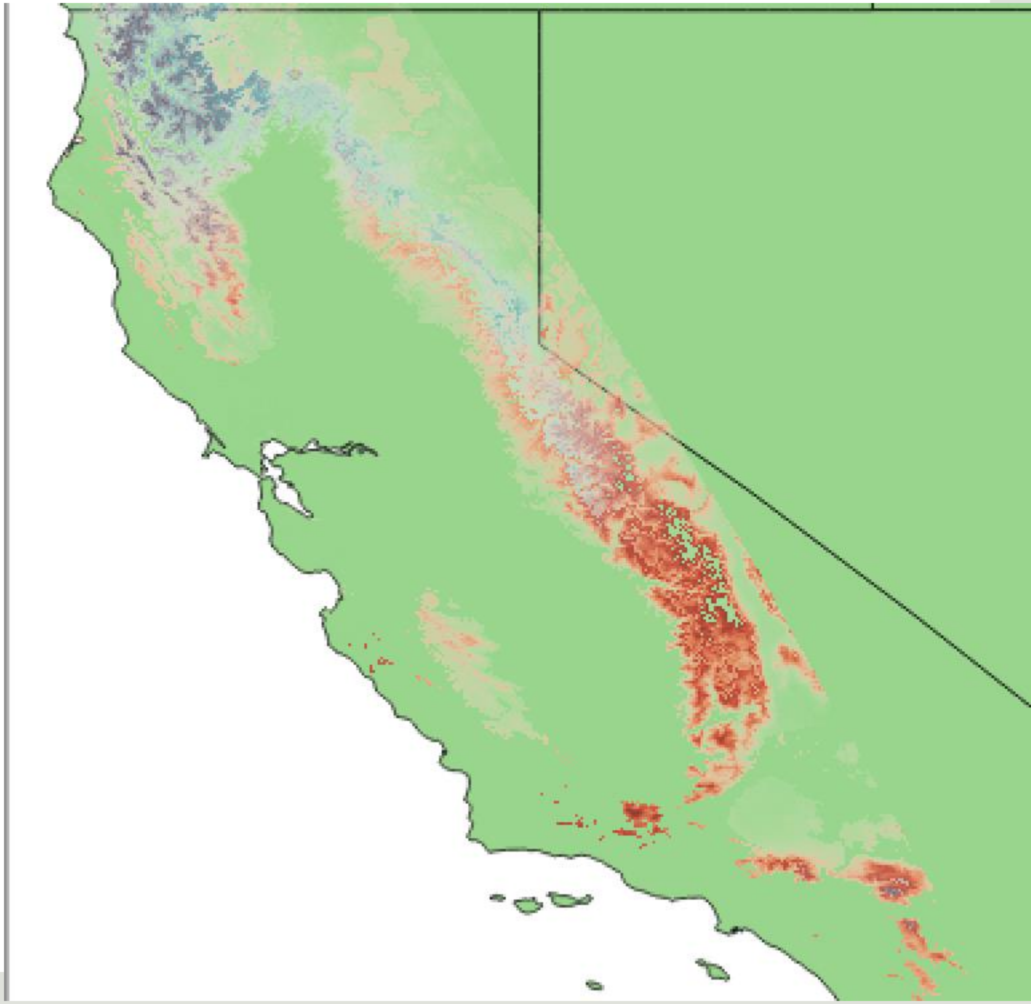
Layer Properties - Penstemon_parvulus_no

Global transparency

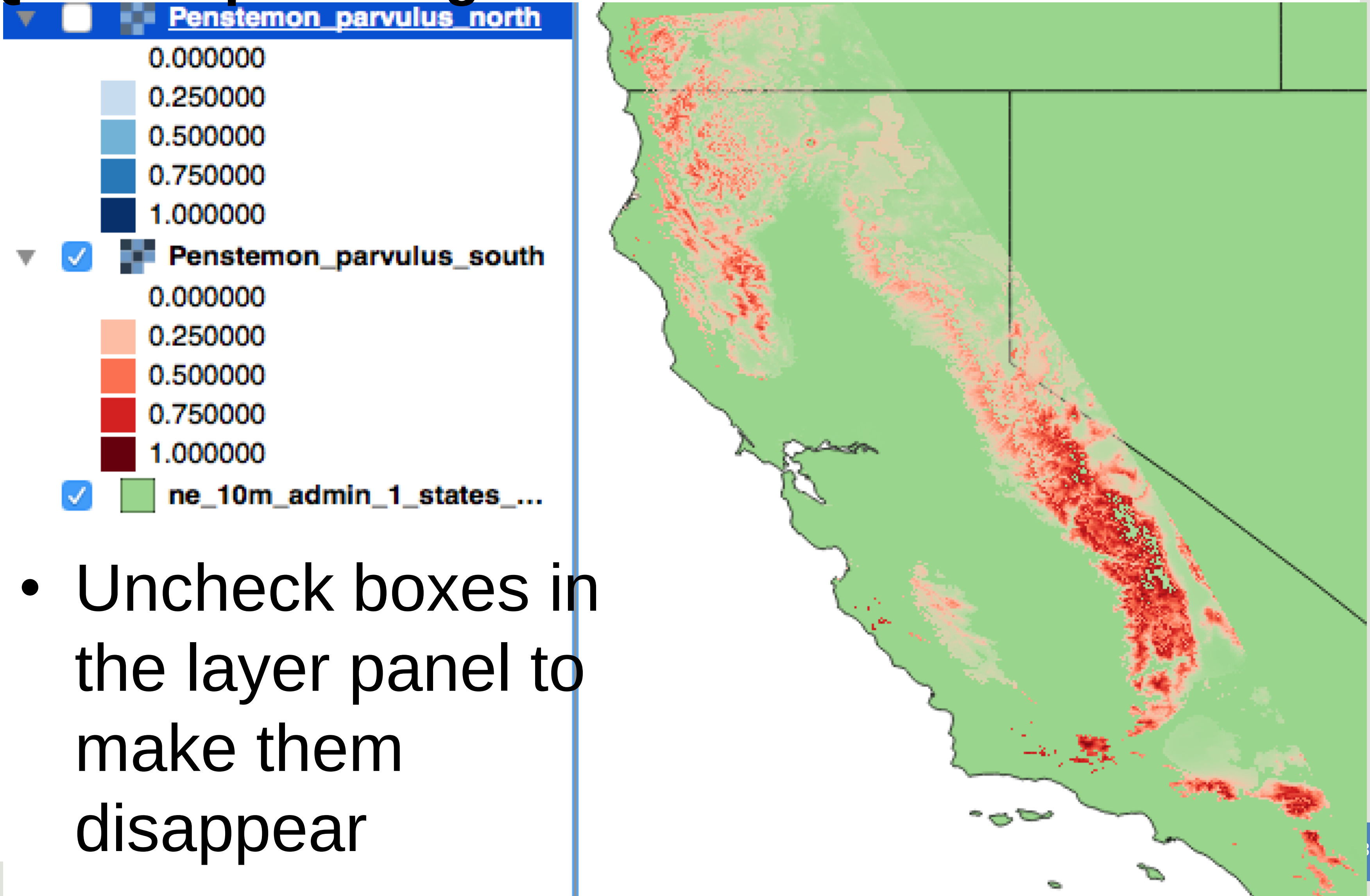
None 50% Full

QGIS Map making

- A little better
  - When there is a lot of overlap it is better to look at each separately



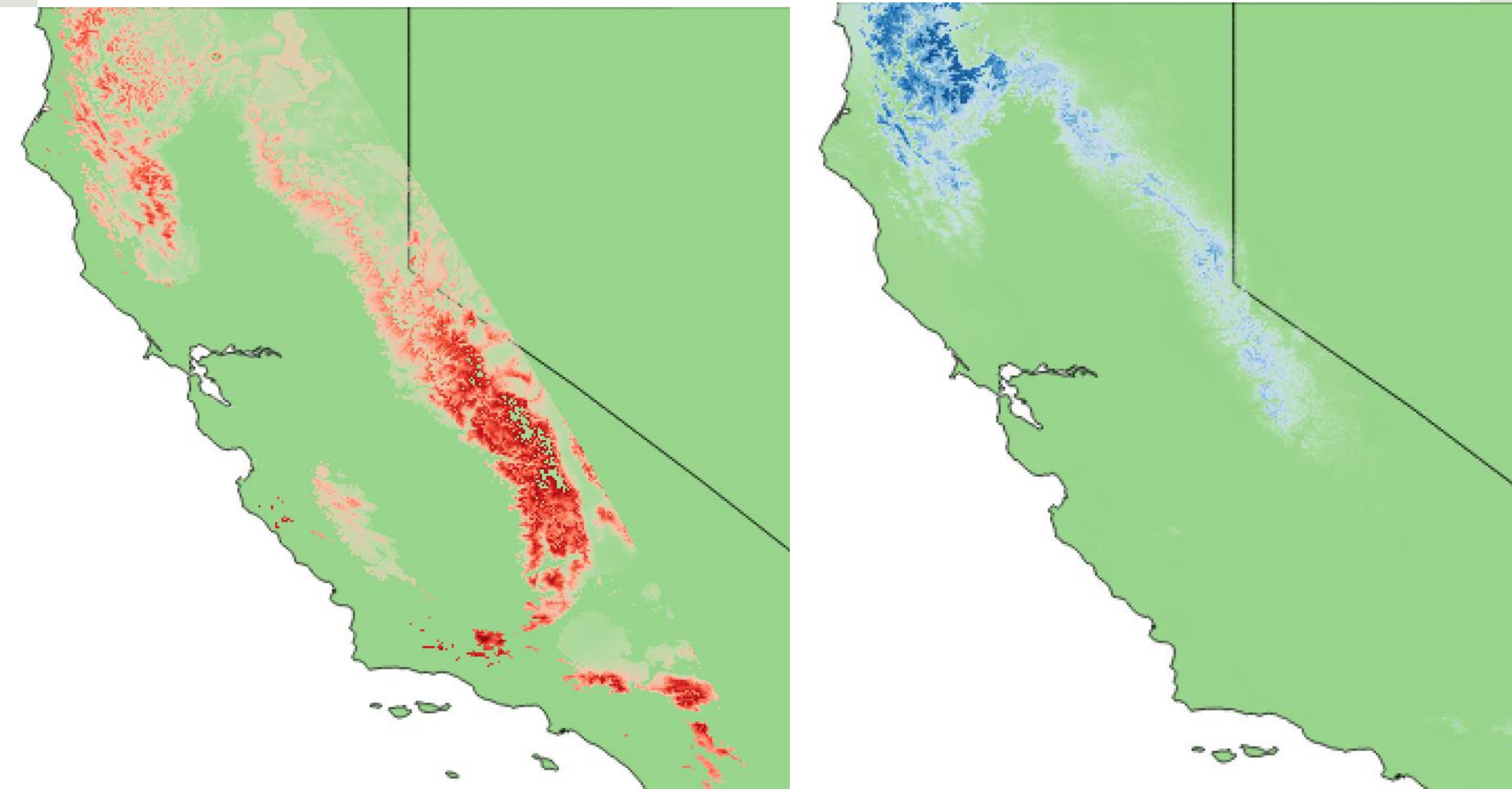
QGIS Map making



▼ ☐ Penstemon_parvulus_north
0.000000
🟦 0.250000
🟦 0.500000
🟦 0.750000
🟦 1.000000
▼ ☑ Penstemon_parvulus_south
0.000000
🟥 0.250000
🟥 0.500000
🟥 0.750000
🟥 1.000000
☑ 🟩 ne_10m_admin_1_states_...

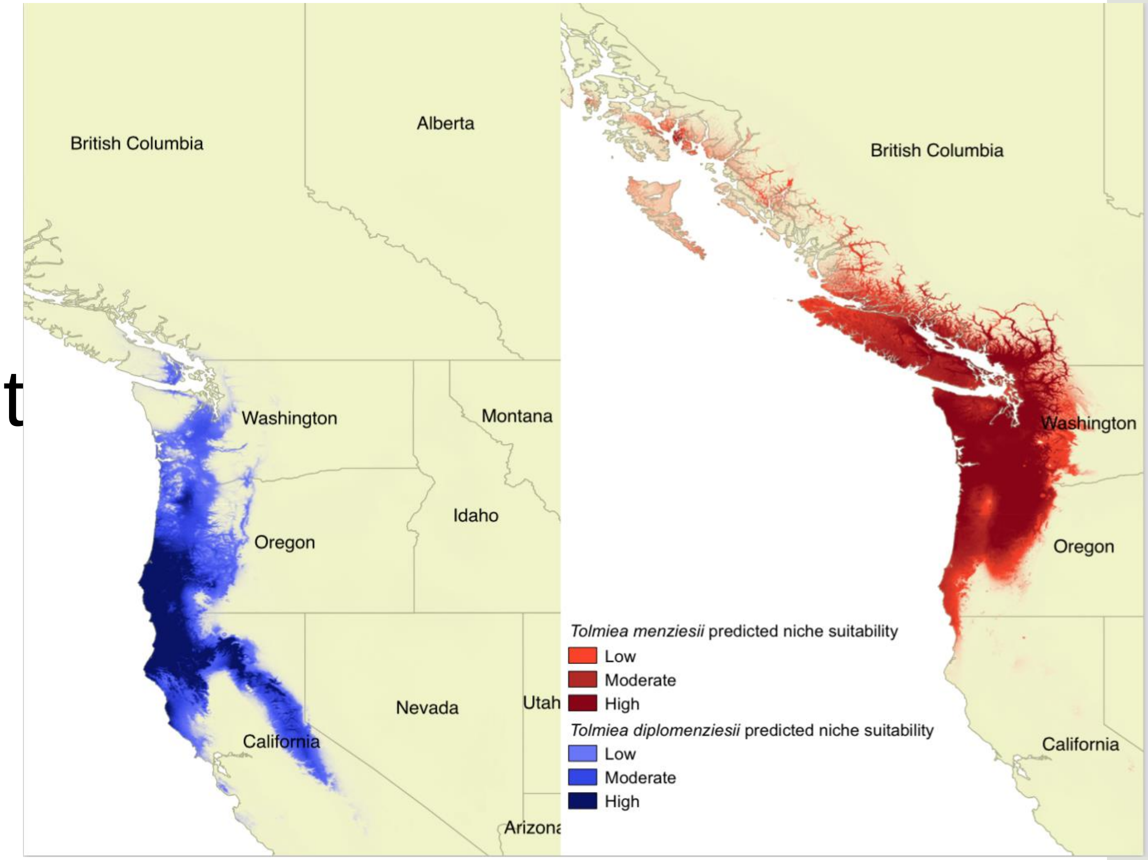
- Uncheck boxes in the layer panel to make them disappear

QGIS Map making



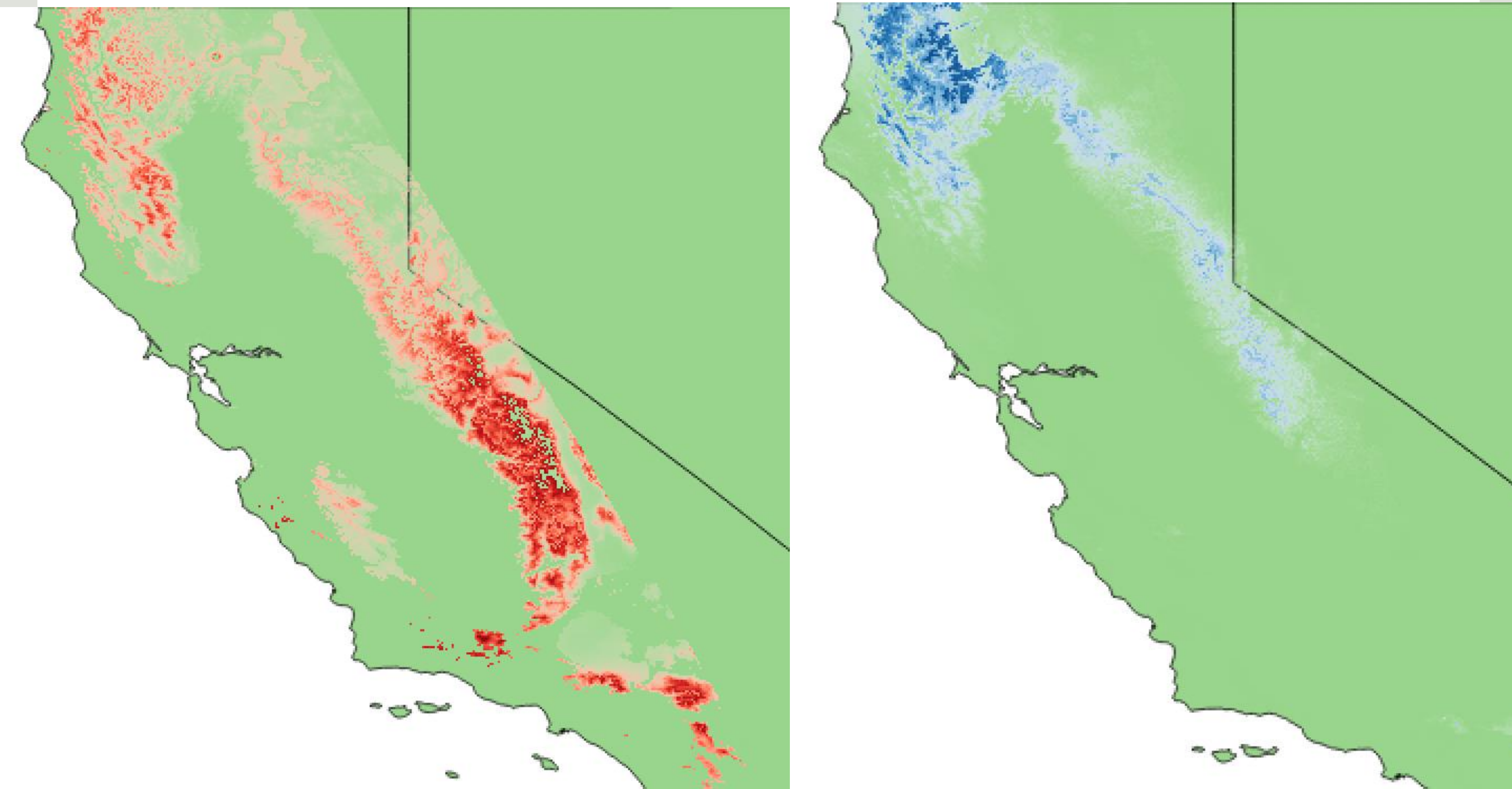
QGIS Map making

- Work with color classification break points and/or transparency settings
- Use the built in print composer to add in legends, etc
  - Beyond the scope of this workshop



Niche Overlap

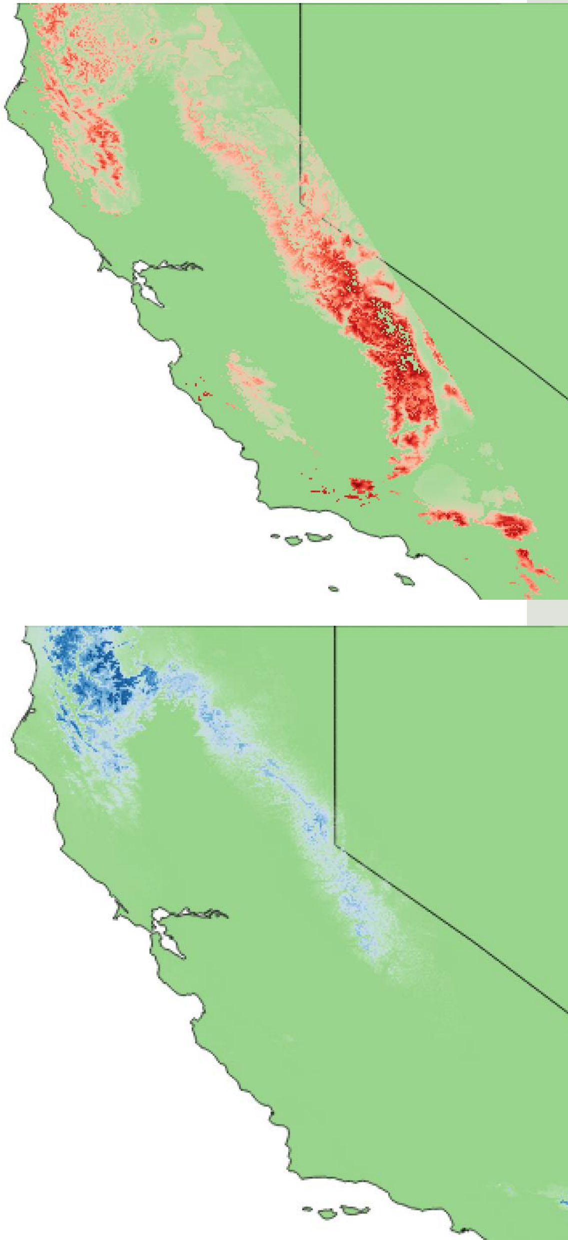
- These overlap, but how do we quantify it?



Niche Overlap

- These are similar, but how do we quantify similarity?

- Schoener’s D
  - Ranges from 0 (no overlap) to 1 (total overlap)
  - Takes into account the suitability scores



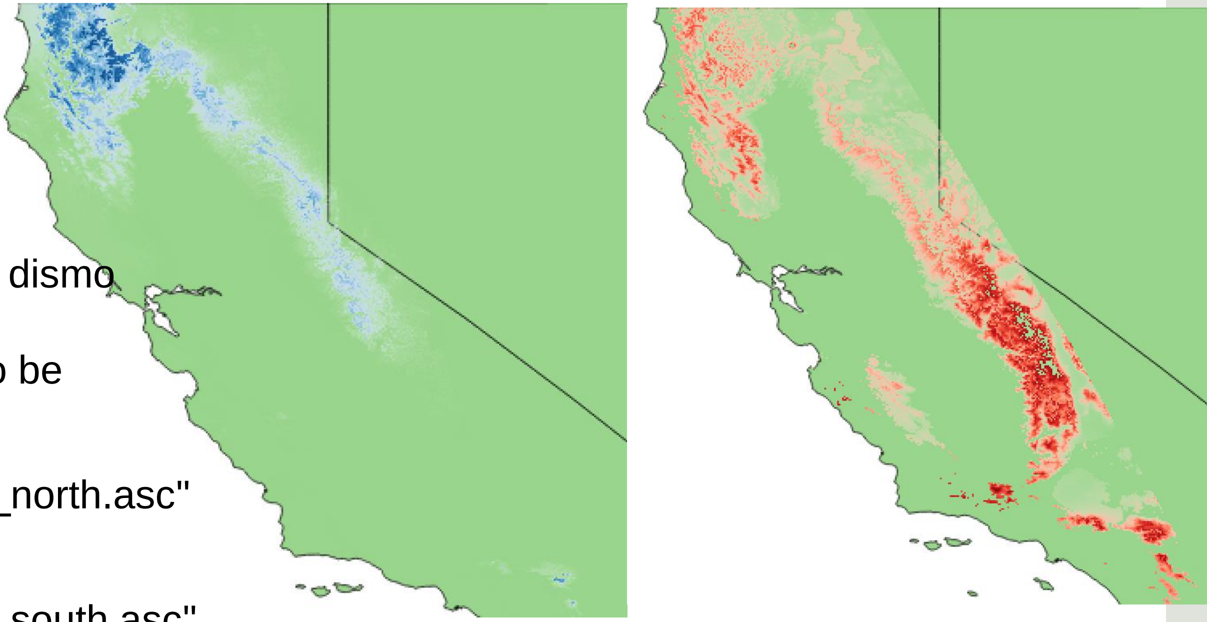
Niche overlap

- #load packages
- require(raster)
- require(dismo)

- #calculating niche overlap using dismo

- ###import two or more rasters to be compared
- x <- raster("~/Desktop/enm workshop/Penstemon_parvulus_north.asc")
- y <- raster("~/Desktop/enm workshop/Penstemon_parvulus_south.asc")

- #calculate niche overlap
- nicheOverlap(x, y, stat='D', mask=TRUE, checkNegatives=TRUE)



D = 0.4092094

Additional info

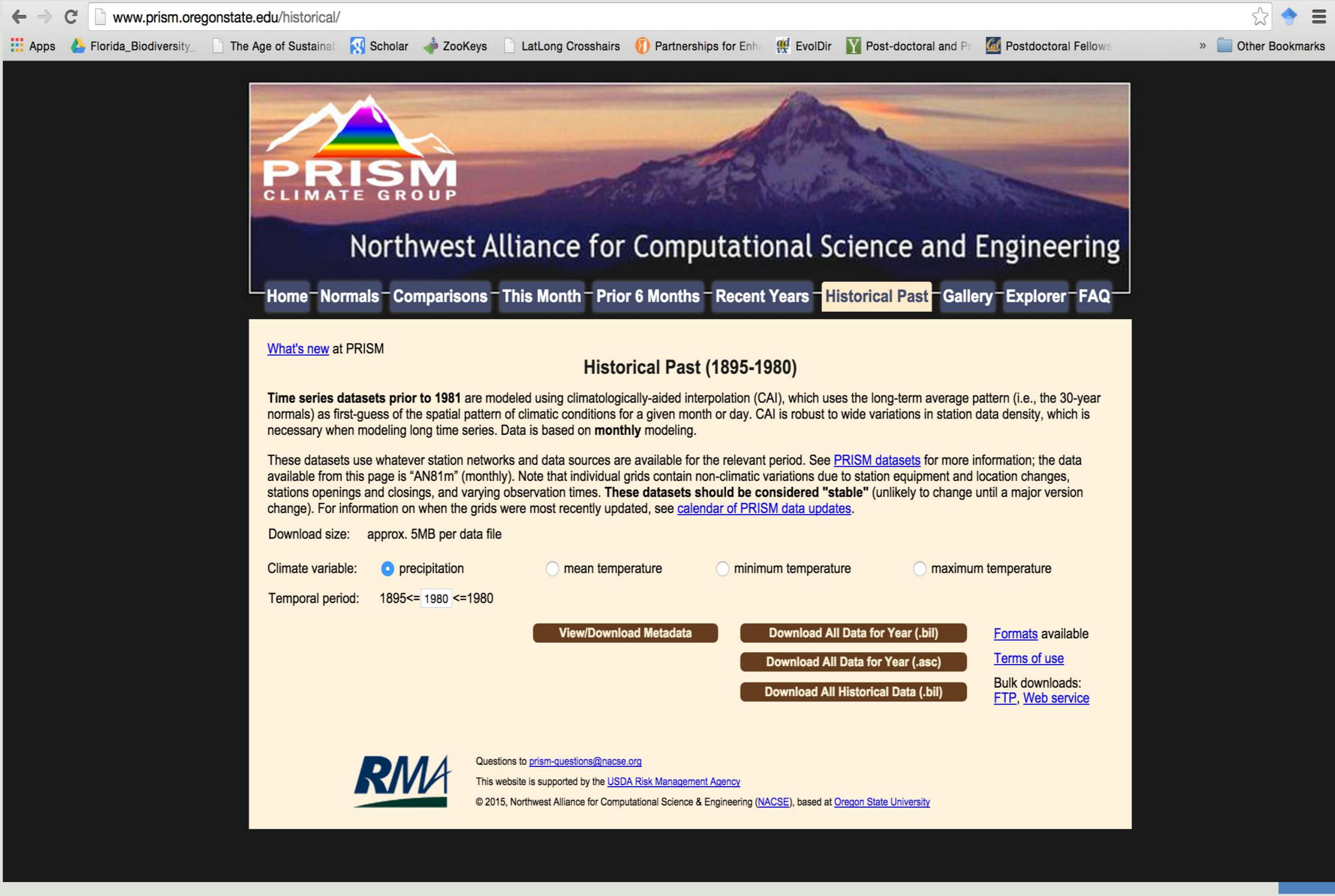
- Hypothesis testing using niche models and similarity
  - Are the models similar by chance and/or geographic autocorrelation?
    - The ENMtools manual provides a nice summary of these methods and references

    http://www.danwarren.net/enmtools/ENMTools_User_Manual%201.0.pdf

Using Yearly Climate Data for Niche Modeling

Charlotte Germain-Aubrey
July, 25th, 2015

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. All images used with permission or are free from copyright.



What's new at PRISM

Historical Past (1895-1980)

Time series datasets prior to 1981 are modeled using climatologically-aided interpolation (CAI), which uses the long-term average pattern (i.e., the 30-year normals) as first-guess of the spatial pattern of climatic conditions for a given month or day. CAI is robust to wide variations in station data density, which is necessary when modeling long time series. Data is based on **monthly** modeling.

These datasets use whatever station networks and data sources are available for the relevant period. See PRISM datasets for more information; the data available from this page is “AN81m” (monthly). Note that individual grids contain non-climatic variations due to station equipment and location changes, stations openings and closings, and varying observation times. **These datasets should be considered "stable"** (unlikely to change until a major version change). For information on when the grids were most recently updated, see calendar of PRISM data updates.

Download size: approx. 5MB per data file

Climate variable: ● precipitation ○ mean temperature ○ minimum temperature ○ maximum temperature

Temporal period: 1895<= 1980 <=1980

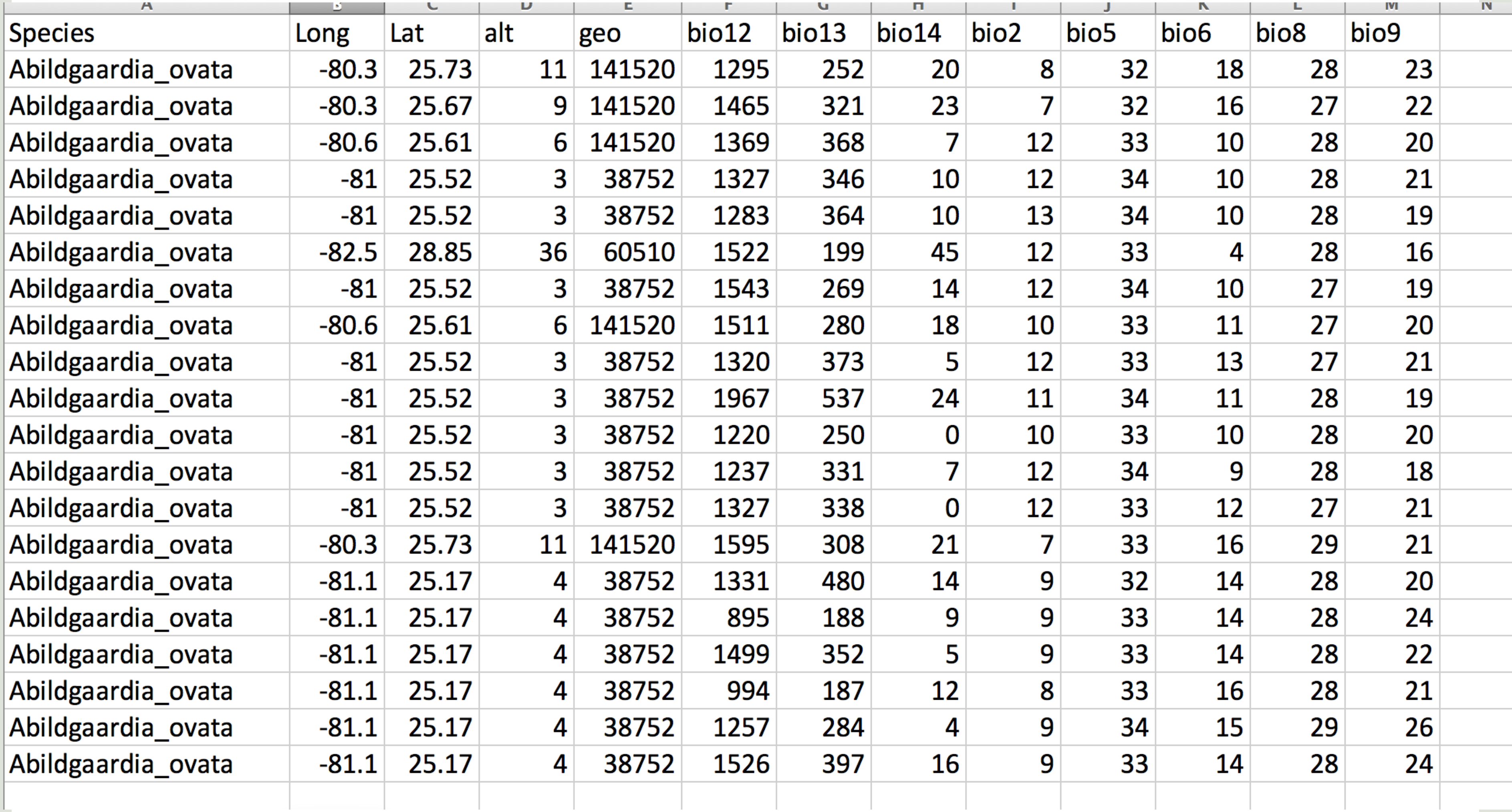
Questions to prism-questions@nacse.org

This website is supported by the USDA Risk Management Agency

© 2015, Northwest Alliance for Computational Science & Engineering (NACSE), based at Oregon State University

Modeling with historical data

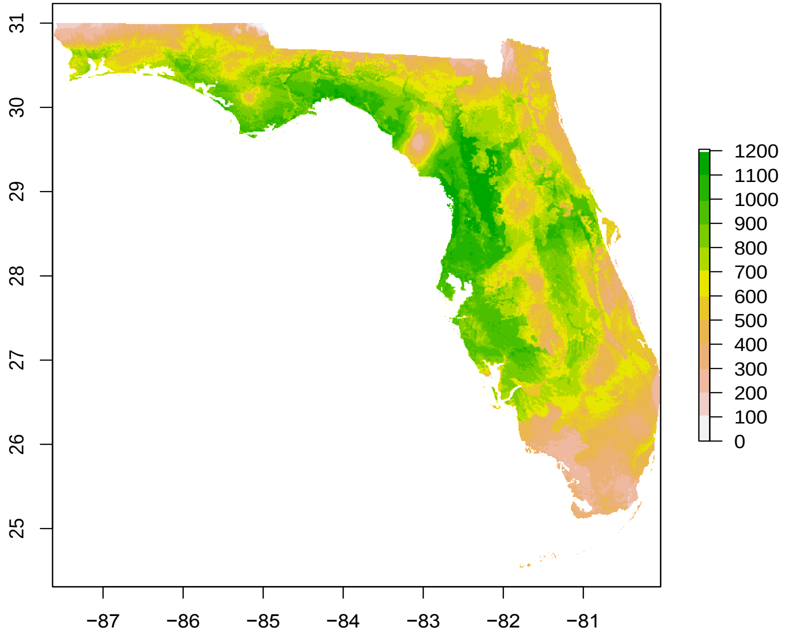
- Museum specimens are dated (at least a year)
- Museum data reflect distribution of plants through recent history
- Can use the climate conditions at the time of collection
- Use PRISM monthly data
- Reconstruct Bioclim layers for each year
- Run MaxEnt model



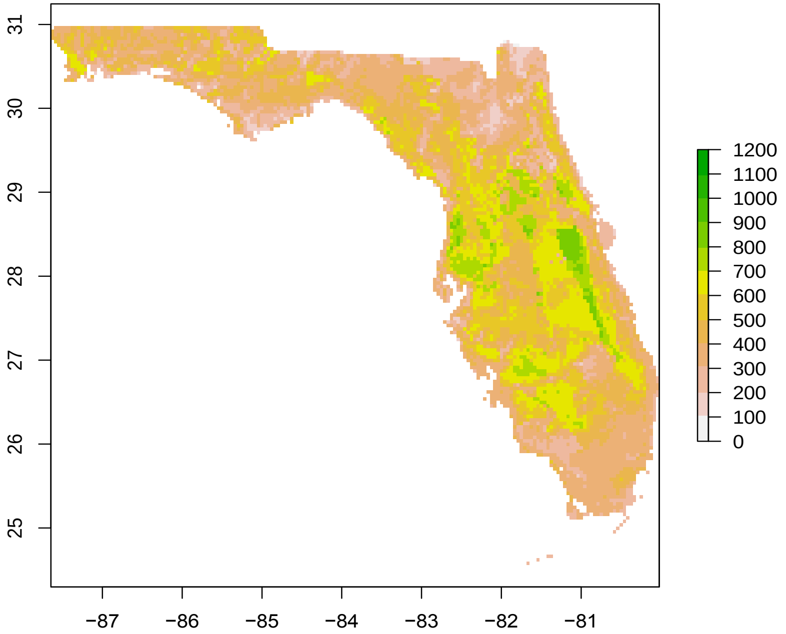
Species	Long	Lat	alt	geo	bio12	bio13	bio14	bio2	bio5	bio6	bio8	bio9
Abildgaardia_ovata	-80.3	25.73	11	141520	1295	252	20	8	32	18	28	23
Abildgaardia_ovata	-80.3	25.67	9	141520	1465	321	23	7	32	16	27	22
Abildgaardia_ovata	-80.6	25.61	6	141520	1369	368	7	12	33	10	28	20
Abildgaardia_ovata	-81	25.52	3	38752	1327	346	10	12	34	10	28	21
Abildgaardia_ovata	-81	25.52	3	38752	1283	364	10	13	34	10	28	19
Abildgaardia_ovata	-82.5	28.85	36	60510	1522	199	45	12	33	4	28	16
Abildgaardia_ovata	-81	25.52	3	38752	1543	269	14	12	34	10	27	19
Abildgaardia_ovata	-80.6	25.61	6	141520	1511	280	18	10	33	11	27	20
Abildgaardia_ovata	-81	25.52	3	38752	1320	373	5	12	33	13	27	21
Abildgaardia_ovata	-81	25.52	3	38752	1967	537	24	11	34	11	28	19
Abildgaardia_ovata	-81	25.52	3	38752	1220	250	0	10	33	10	28	20
Abildgaardia_ovata	-81	25.52	3	38752	1237	331	7	12	34	9	28	18
Abildgaardia_ovata	-81	25.52	3	38752	1327	338	0	12	33	12	27	21
Abildgaardia_ovata	-80.3	25.73	11	141520	1595	308	21	7	33	16	29	21
Abildgaardia_ovata	-81.1	25.17	4	38752	1331	480	14	9	32	14	28	20
Abildgaardia_ovata	-81.1	25.17	4	38752	895	188	9	9	33	14	28	24
Abildgaardia_ovata	-81.1	25.17	4	38752	1499	352	5	9	33	14	28	22
Abildgaardia_ovata	-81.1	25.17	4	38752	994	187	12	8	33	16	28	21
Abildgaardia_ovata	-81.1	25.17	4	38752	1257	284	4	9	34	15	29	26
Abildgaardia_ovata	-81.1	25.17	4	38752	1526	397	16	9	33	14	28	24

Alpha diversity

Using averaged climate data (old method)



Using yearly climate data (new method)



Min=131 max=604

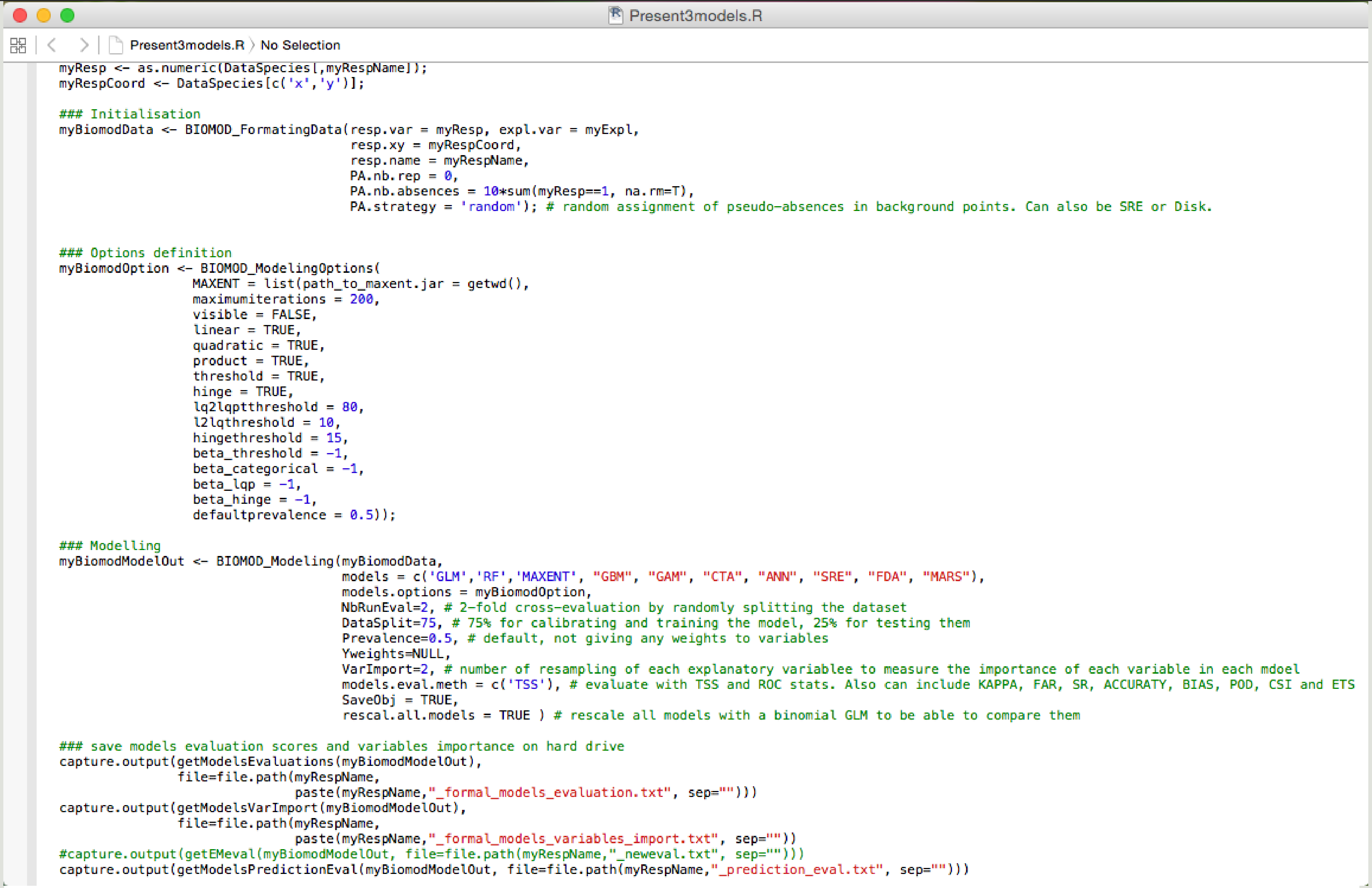
Other methods for niche modeling

Charlotte Germain-Aubrey

July, 25th, 2015

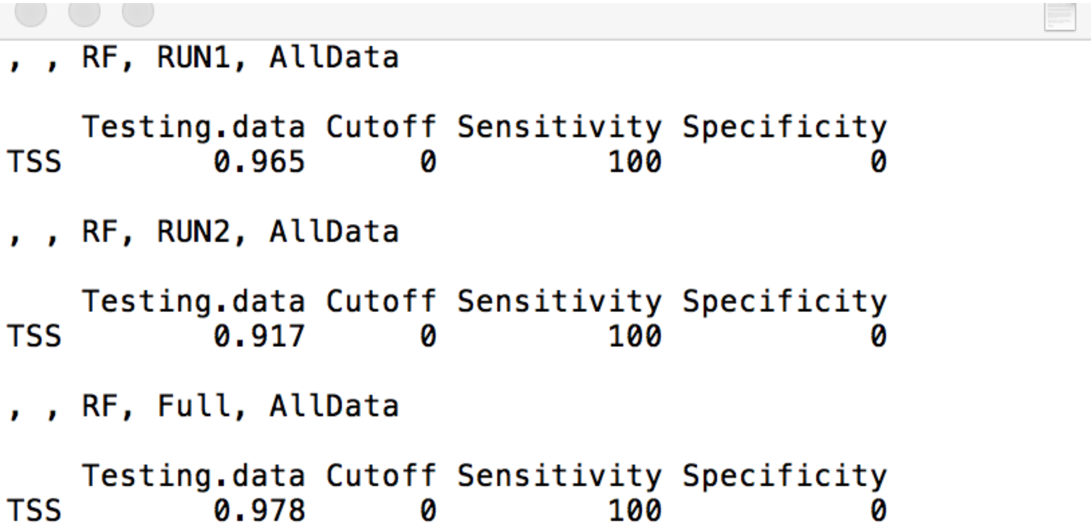
iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. All images used with permission or are free from copyright.

Biomod2



```r
myResp <- as.numeric(DataSpecies[,myRespName]);
myRespCoord <- DataSpecies[c('x','y')];

### Initialisation
myBiomodData <- BIOMOD_FormatingData(resp.var = myResp, expl.var = myExpl,
                                     resp.xy = myRespCoord,
                                     resp.name = myRespName,
                                     PA.nb.rep = 0,
                                     PA.nb.absences = 10*sum(myResp==1, na.rm=T),
                                     PA.strategy = 'random'); # random assignment of pseudo-absences in background points. Can also be SRE or Disk.


### Options definition
myBiomodOption <- BIOMOD_ModelingOptions(
                 MAXENT = list(path_to_maxent.jar = getwd(),
                 maximumiterations = 200,
                 visible = FALSE,
                 linear = TRUE,
                 quadratic = TRUE,
                 product = TRUE,
                 threshold = TRUE,
                 hinge = TRUE,
                 lq2lqptthreshold = 80,
                 l2lqthreshold = 10,
                 hingethreshold = 15,
                 beta_threshold = -1,
                 beta_categorical = -1,
                 beta_lqp = -1,
                 beta_hinge = -1,
                 defaultprevalence = 0.5));

### Modelling
myBiomodModelOut <- BIOMOD_Modeling(myBiomodData,
                                    models = c('GLM','RF','MAXENT', "GBM", "GAM", "CTA", "ANN", "SRE", "FDA", "MARS"),
                                    models.options = myBiomodOption,
                                    NbRunEval=2, # 2-fold cross-evaluation by randomly splitting the dataset
                                    DataSplit=75, # 75% for calibrating and training the model, 25% for testing them
                                    Prevalence=0.5, # default, not giving any weights to variables
                                    Yweights=NULL,
                                    VarImport=2, # number of resampling of each explanatory variablee to measure the importance of each variable in each mdoel
                                    models.eval.meth = c('TSS'), # evaluate with TSS and ROC stats. Also can include KAPPA, FAR, SR, ACCURATY, BIAS, POD, CSI and ETS
                                    SaveObj = TRUE,
                                    rescal.all.models = TRUE ) # rescale all models with a binomial GLM to be able to compare them

### save models evaluation scores and variables importance on hard drive
capture.output(getModelsEvaluations(myBiomodModelOut),
               file=file.path(myRespName,
                              paste(myRespName,"_formal_models_evaluation.txt", sep="")))
capture.output(getModelsVarImport(myBiomodModelOut),
               file=file.path(myRespName,
                              paste(myRespName,"_formal_models_variables_import.txt", sep=""))
#capture.output(getEMeval(myBiomodModelOut, file=file.path(myRespName,"_neweval.txt", sep="")))
capture.output(getModelsPredictionEval(myBiomodModelOut, file=file.path(myRespName,"_prediction_eval.txt", sep="")))
```

Biomod2



```
, , RF, RUN1, AllData

    Testing.data Cutoff Sensitivity Specificity
TSS        0.965      0         100           0

, , RF, RUN2, AllData

    Testing.data Cutoff Sensitivity Specificity
TSS        0.917      0         100           0

, , RF, Full, AllData

    Testing.data Cutoff Sensitivity Specificity
TSS        0.978      0         100           0
```

LifeMapper – QGIS plugin

- Live demonstration

Thank you!

www.idigbio.org

psoltis@flmnh.ufl.edu

facebook.com/iDigBio

twitter.com/iDigBio

vimeo.com/idigbio

idigbio.org/rss-feed.xml

webcal://www.idigbio.org/events-calendar/export.ics

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. All images used with permission or are free from copyright.