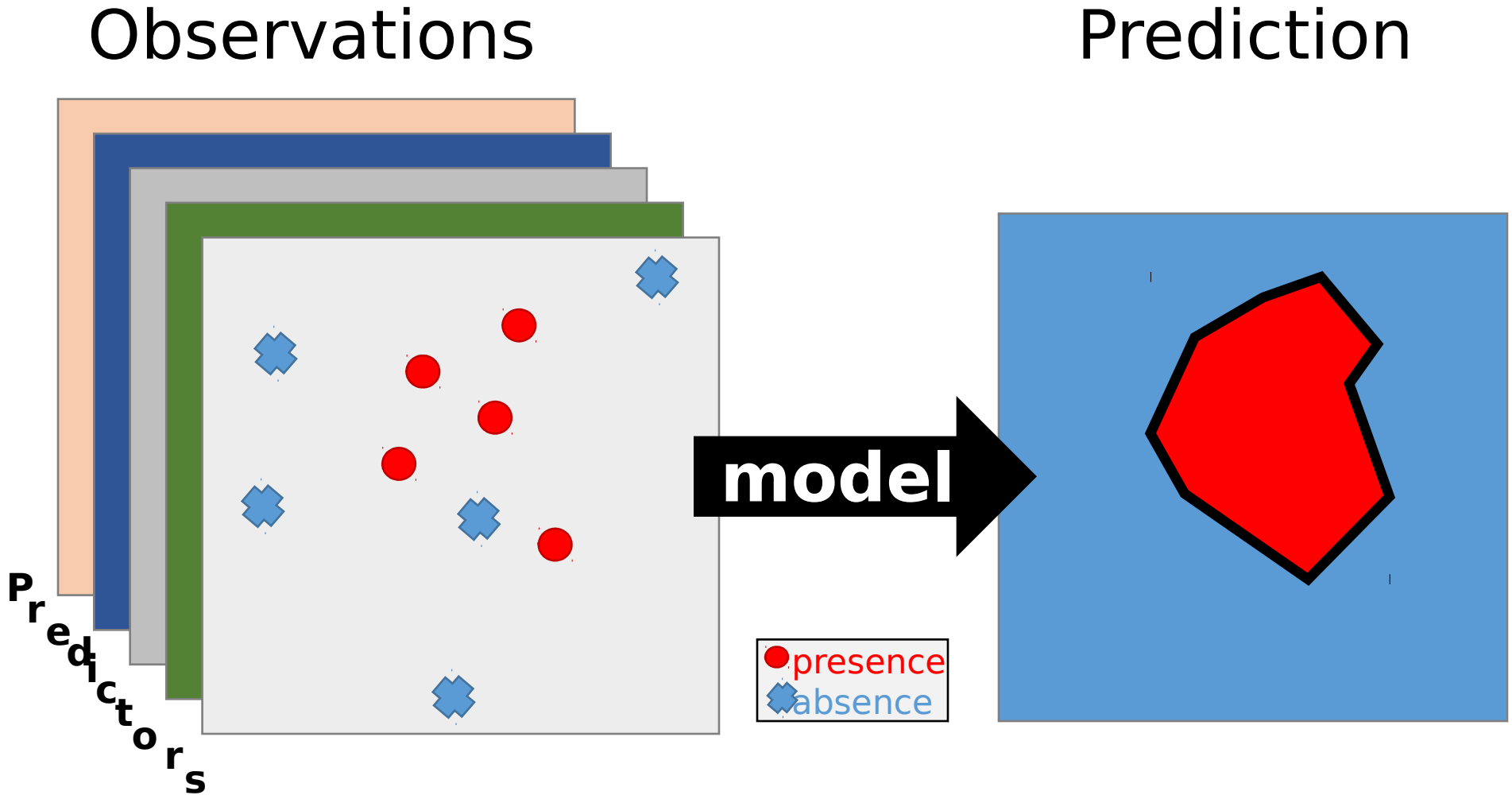
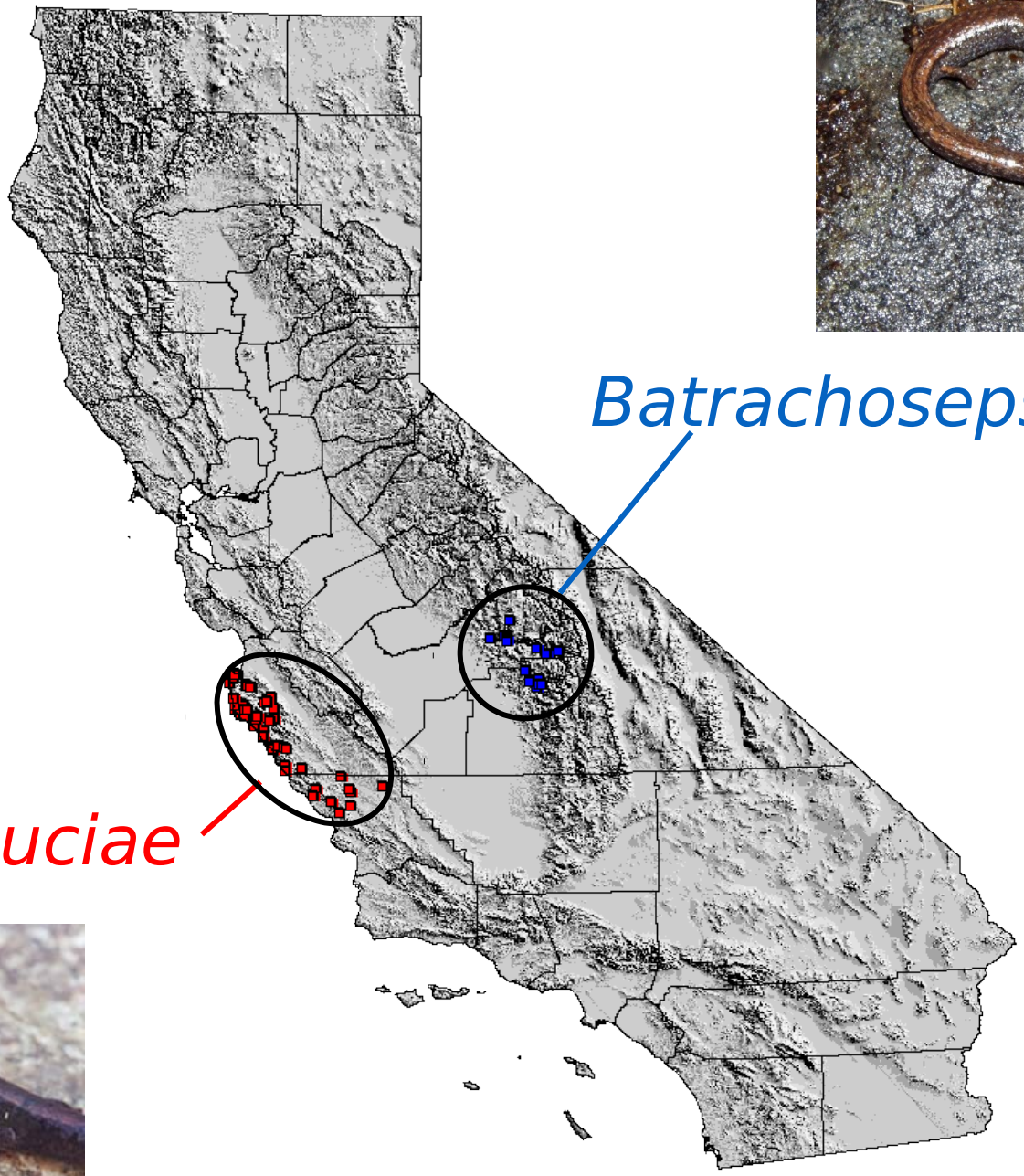
Spatial Distribution Models

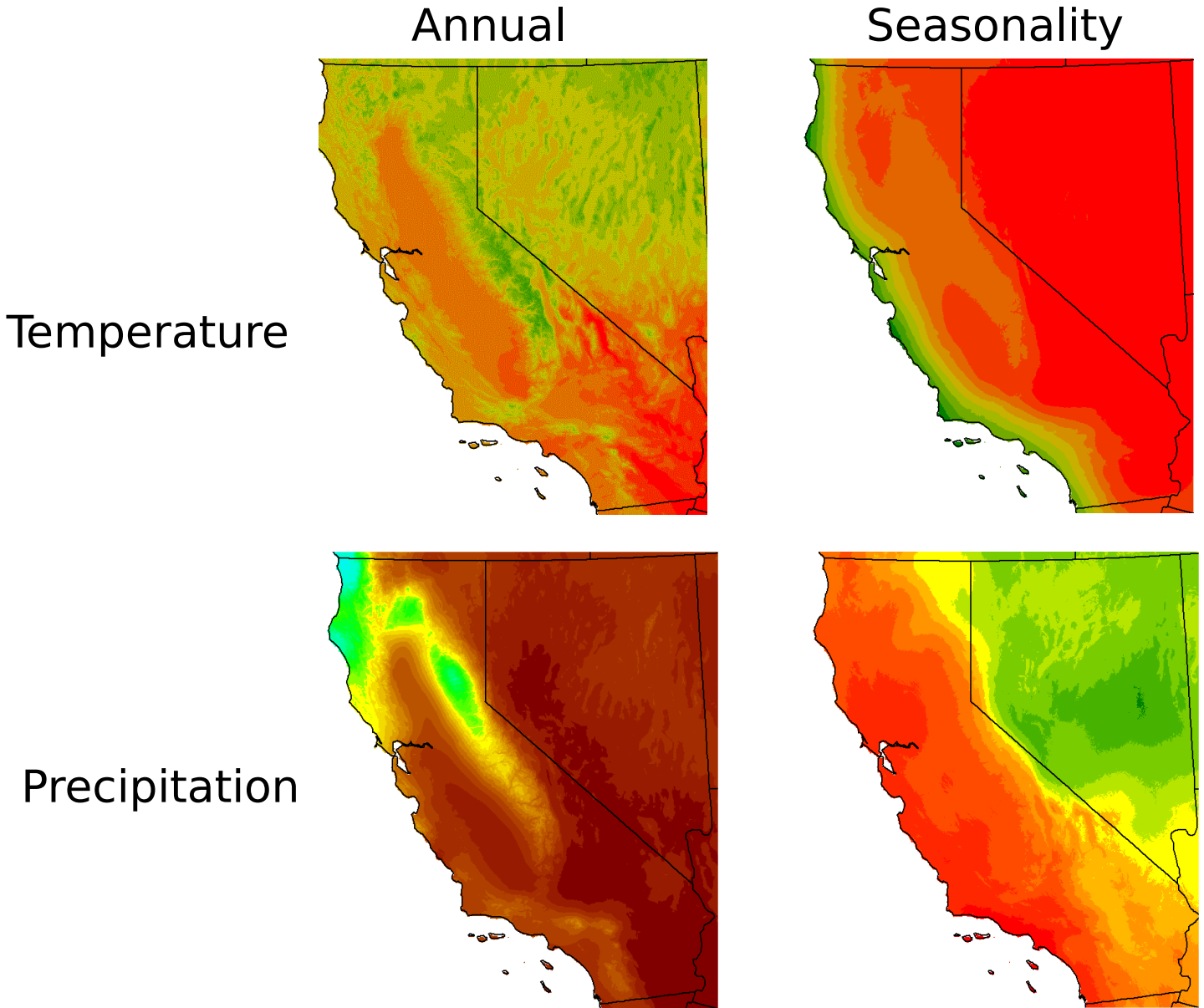


Batrachoseps regius



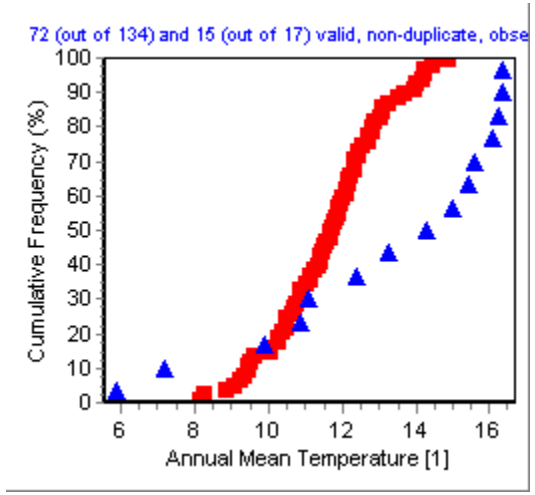
Batrachoseps luciae

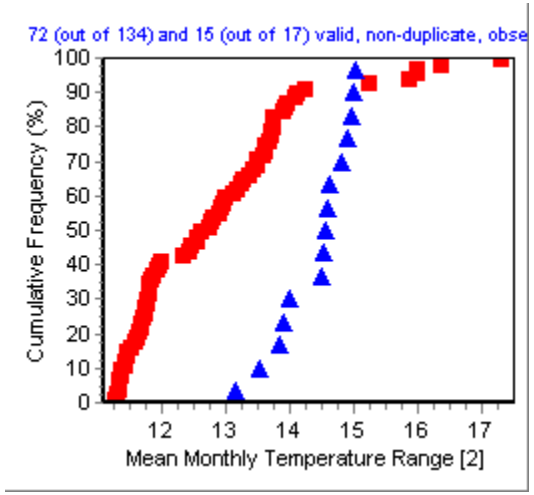
Environmental data

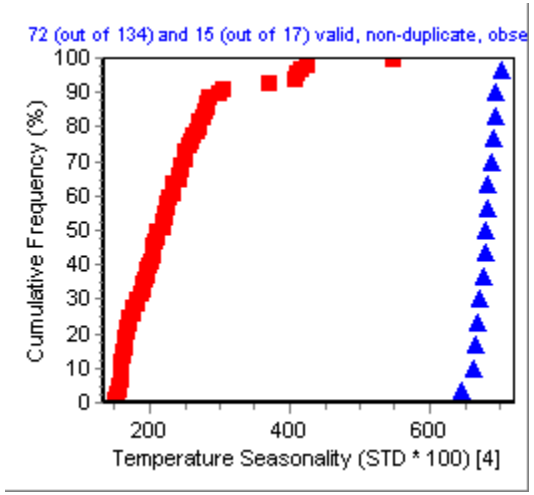


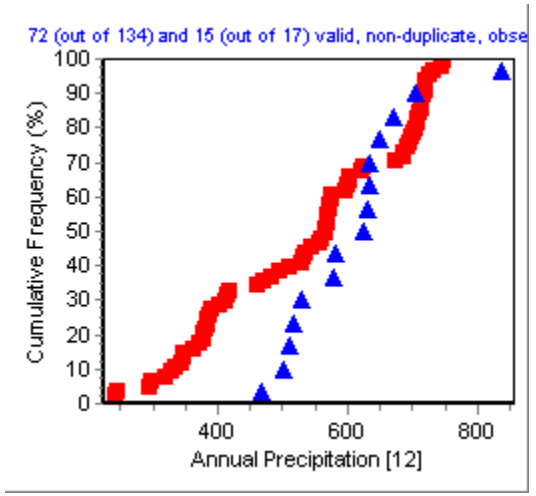
B. regius

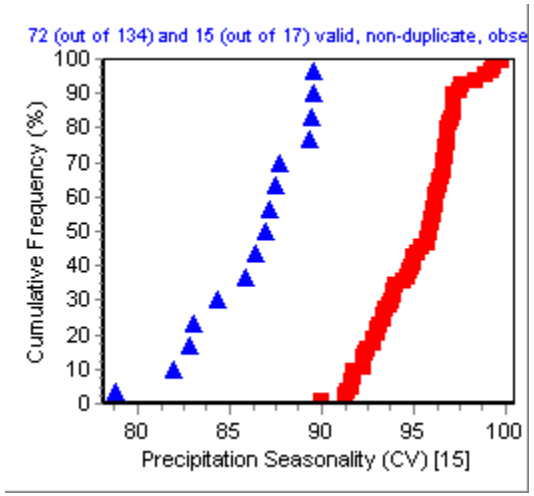
B. luciae

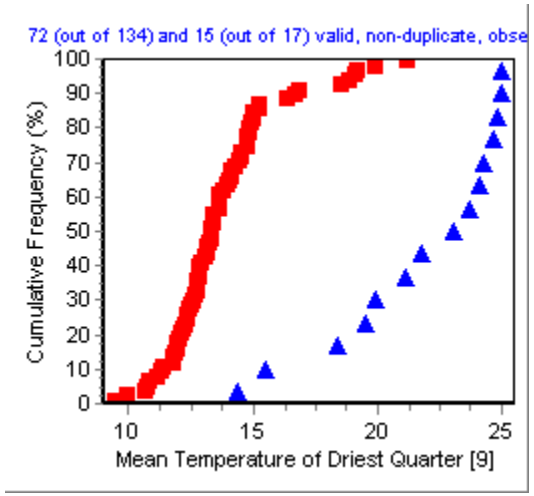


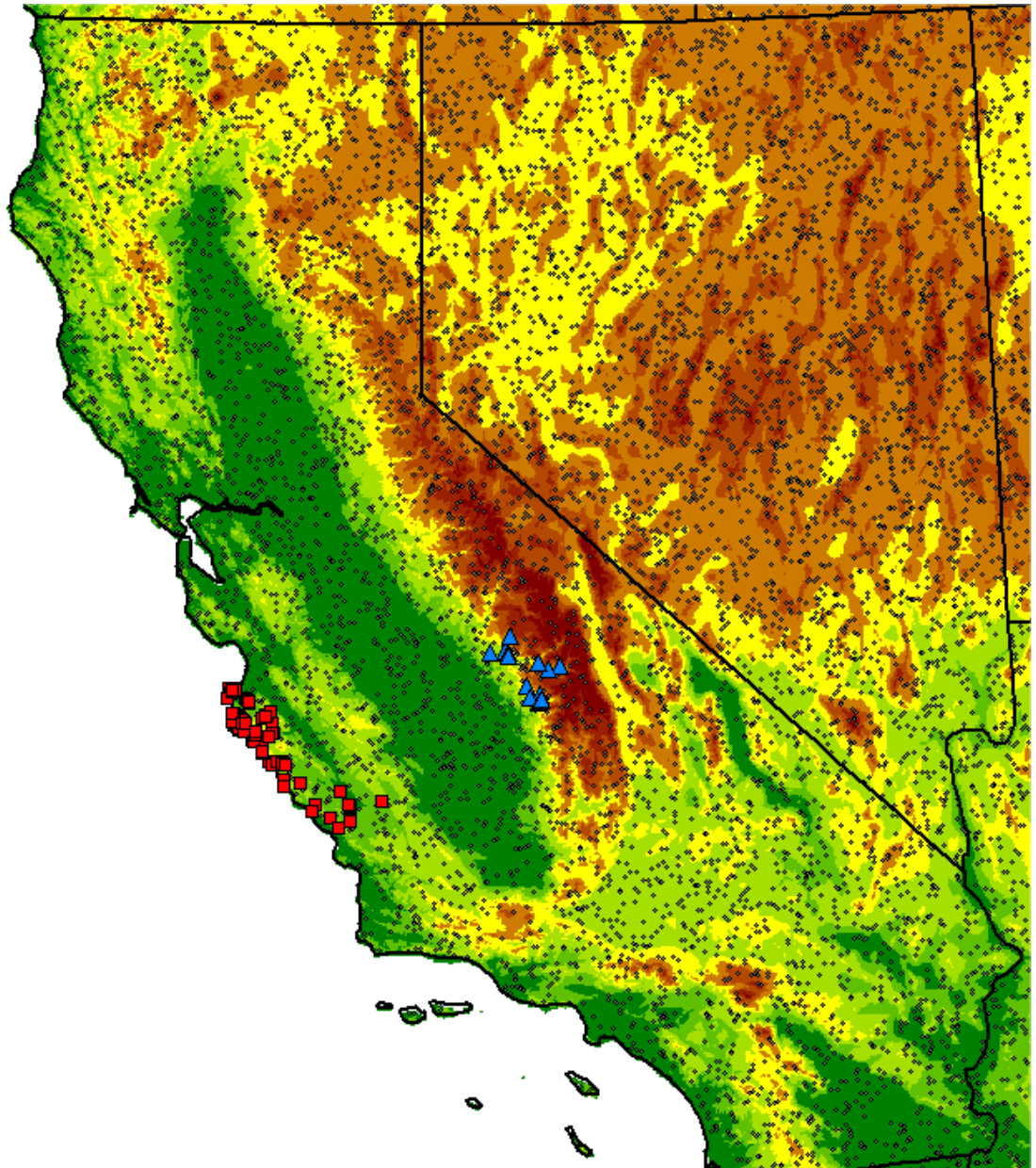


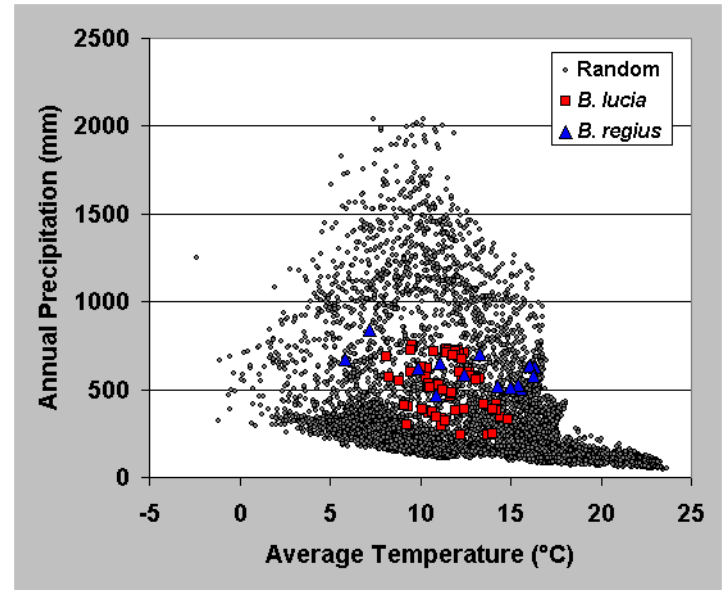


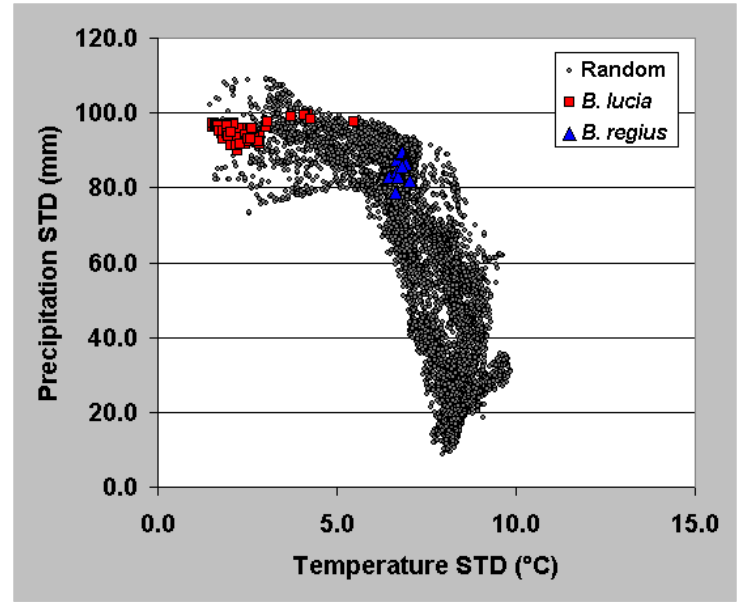








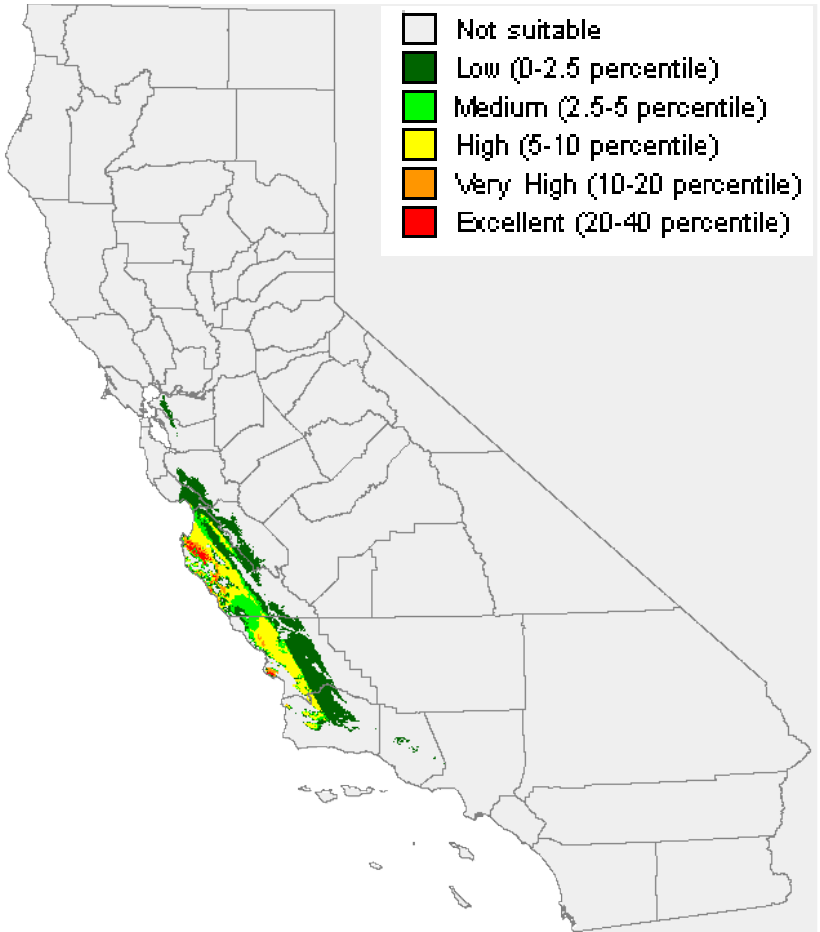

Random
B. lucia
B. regius
Annual Precipitation (mm)
2500
2000
1500
1000
500
0
-5
0
5
10
15
20
25
Average Temperature (°C)


Random
B. lucia
B. regius
Precipitation STD (mm)
120.0
100.0
80.0
60.0
40.0
20.0
0.0
0.0
5.0
10.0
15.0
Temperature STD (°C)

4. Predict

BIOCLIM

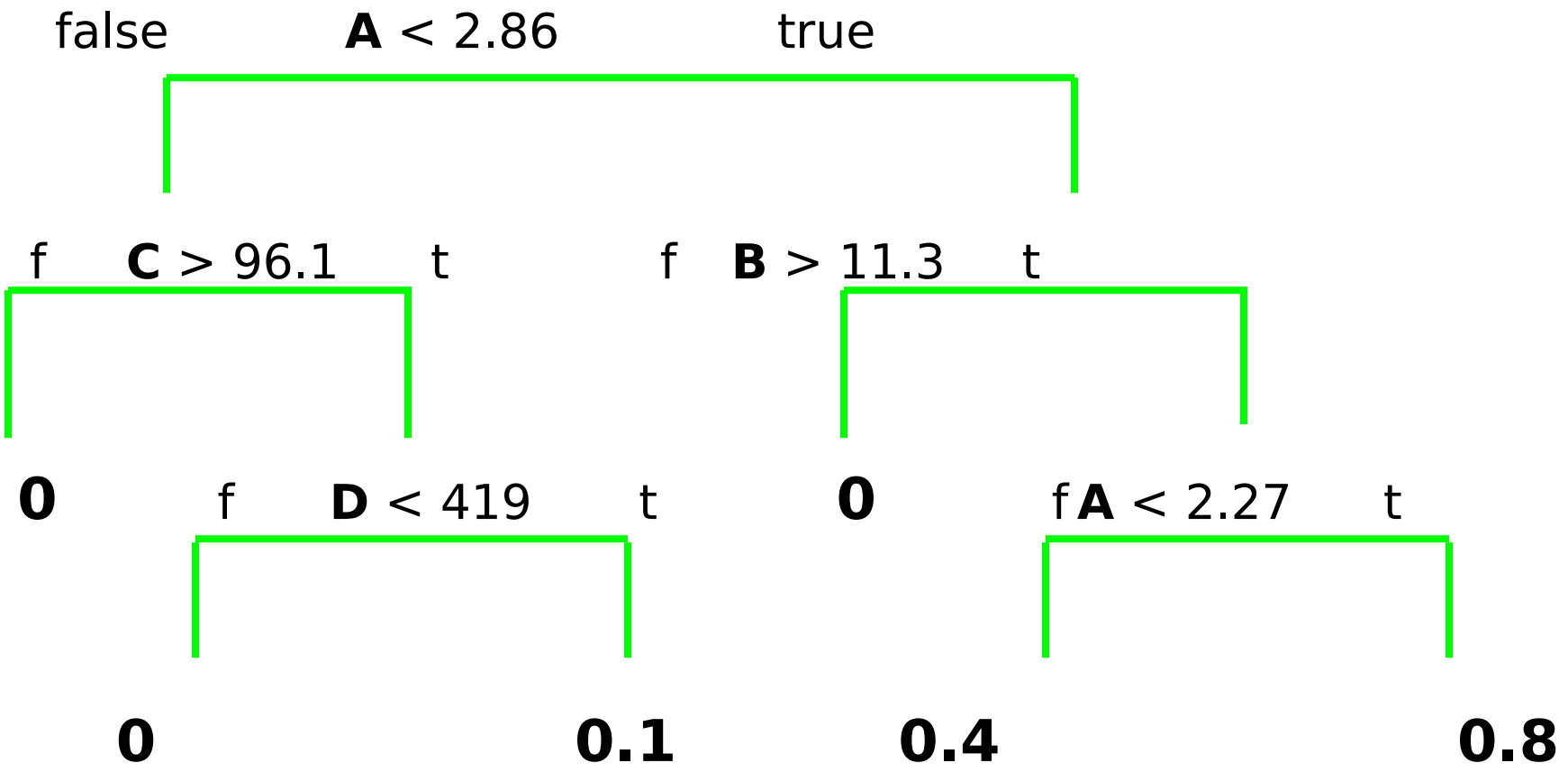
B. lucia



Regression Trees

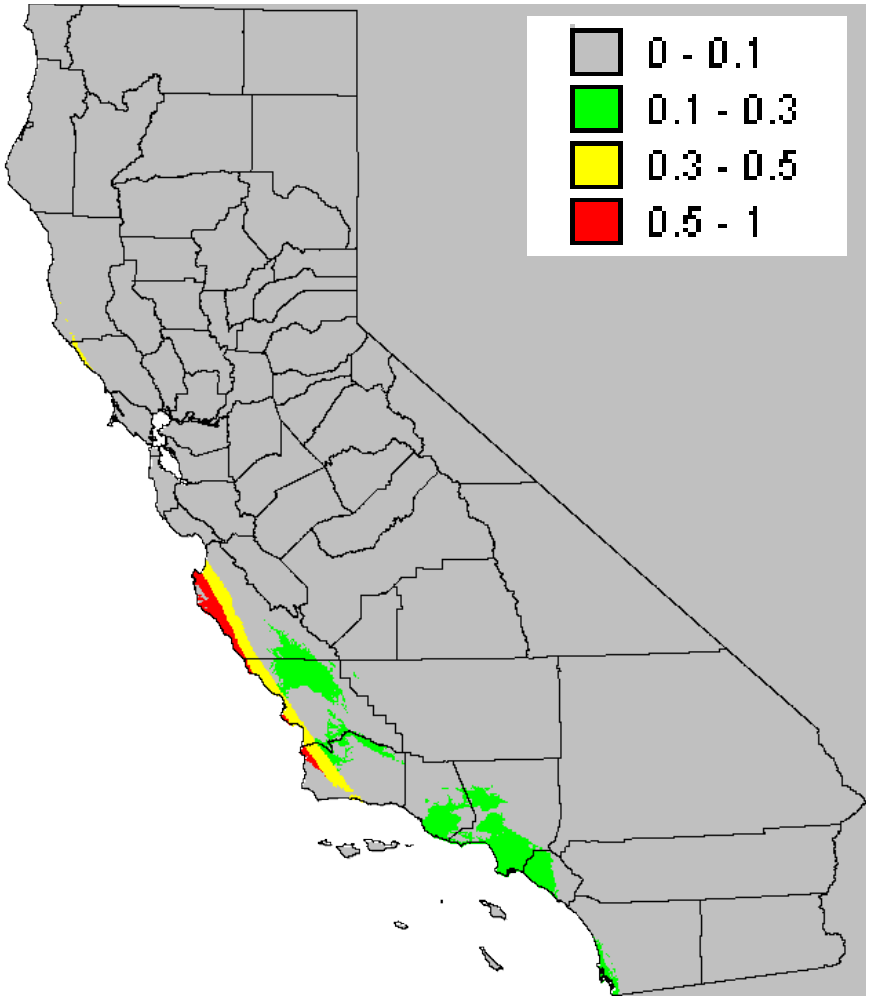
Presence / Absence

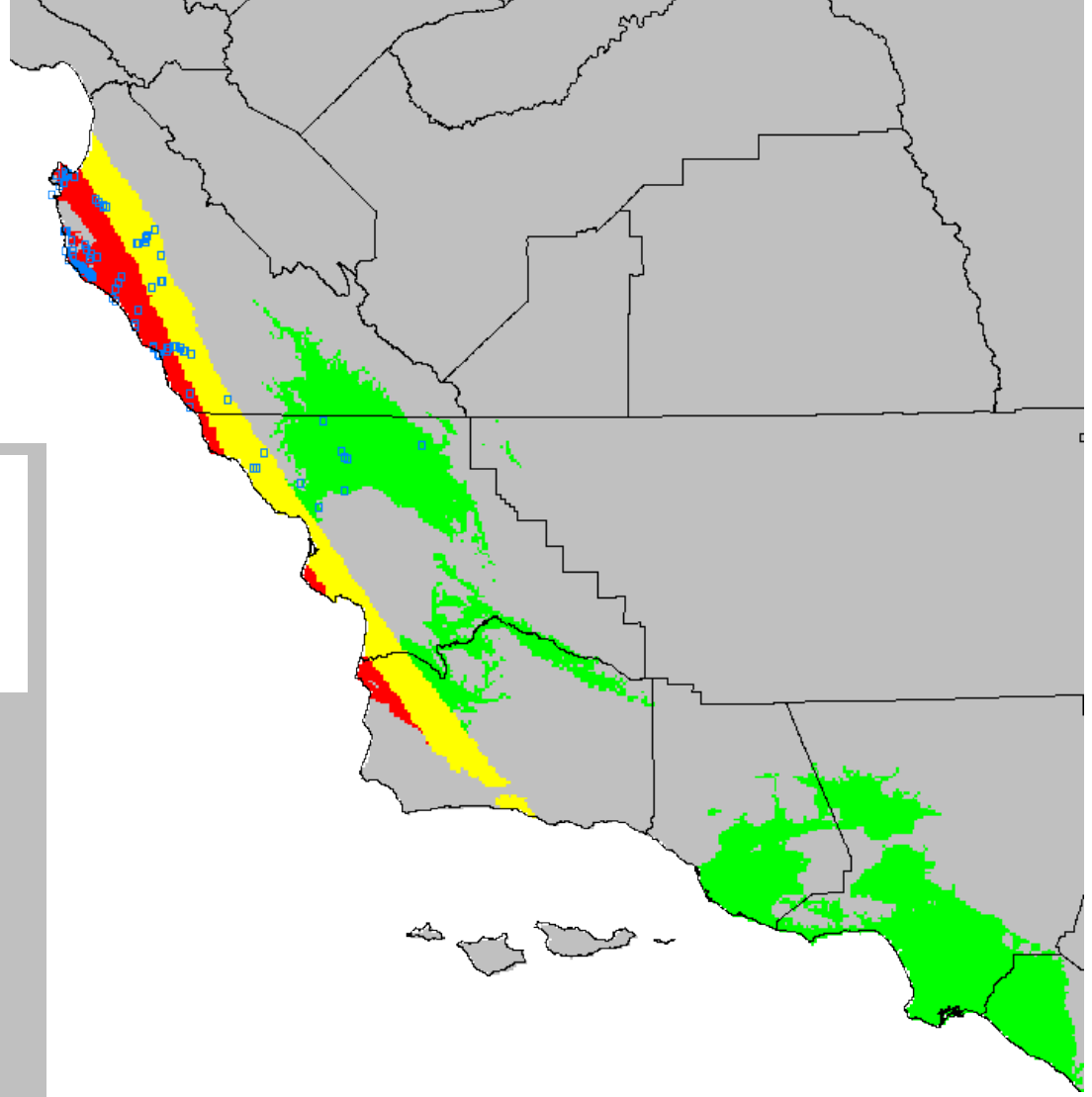
A - Temperature seasonality (STD)
B - Monthly Temperature range
C - Precipitation seasonality (CV)
D - Annual precipitation

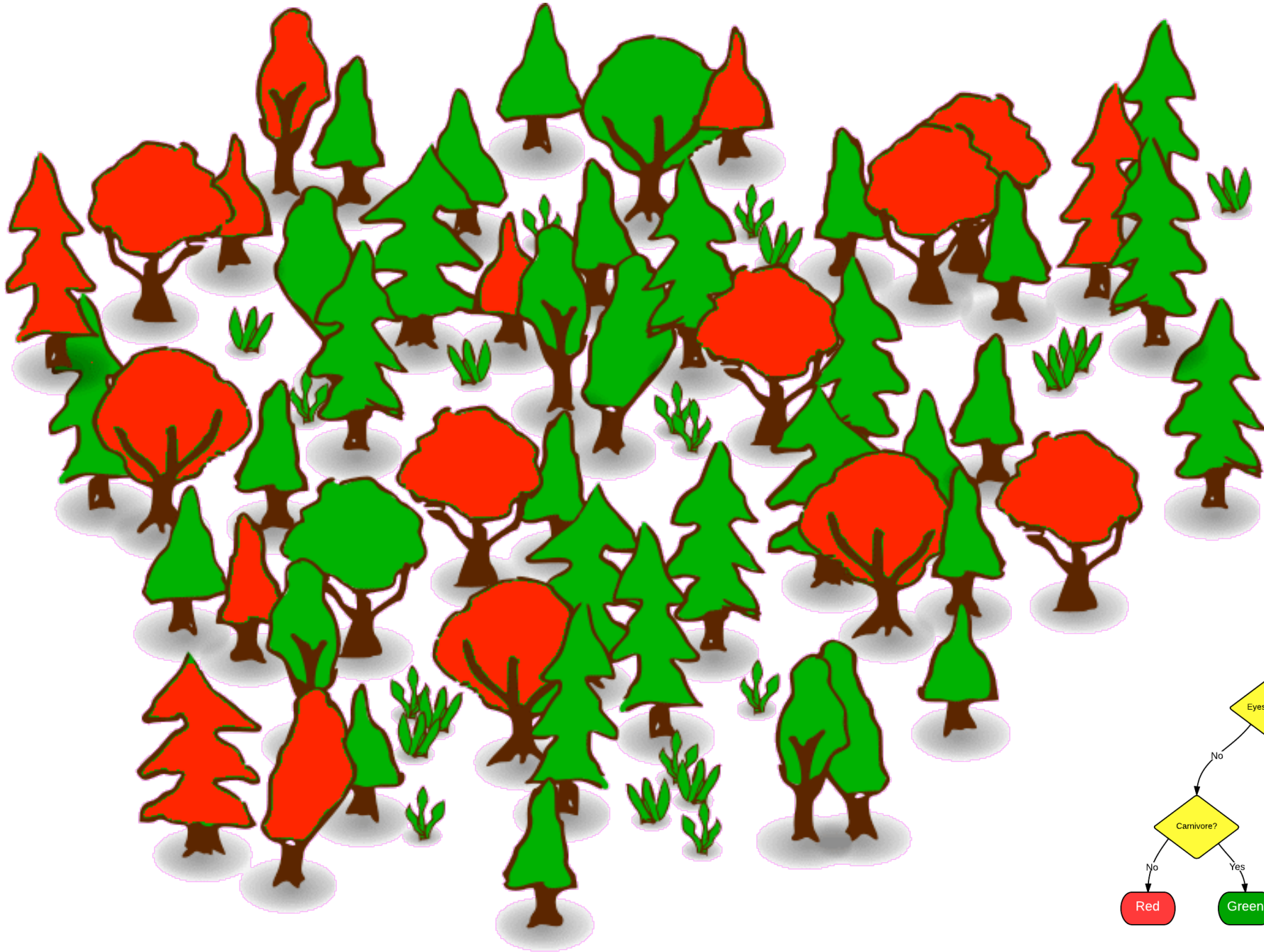


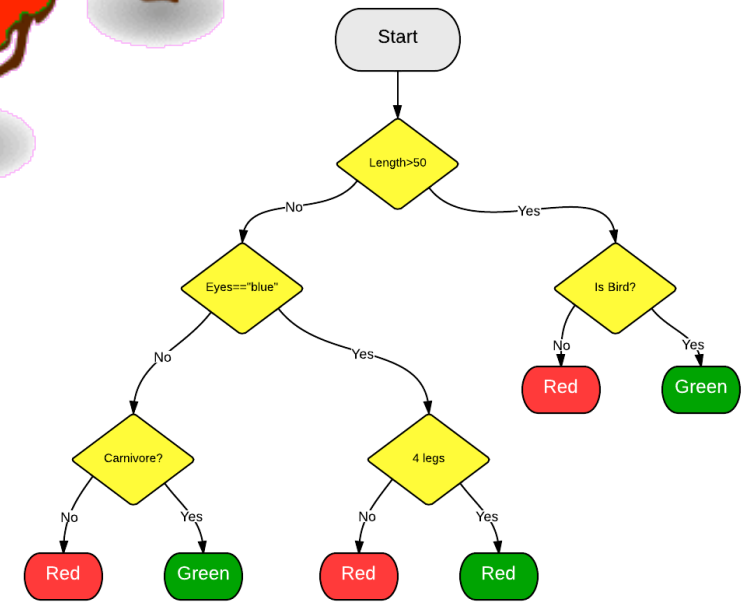
CART

B. lucia



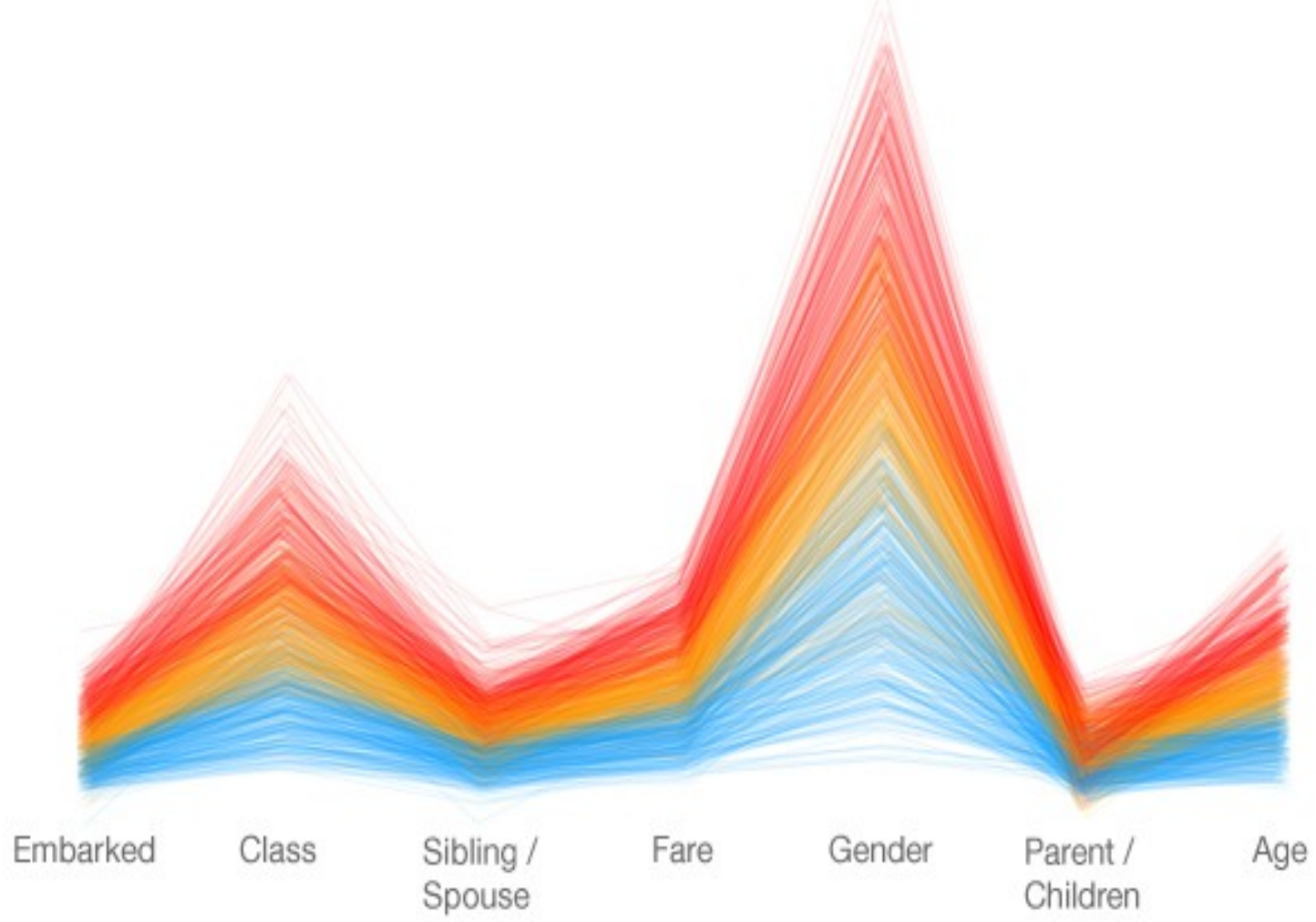




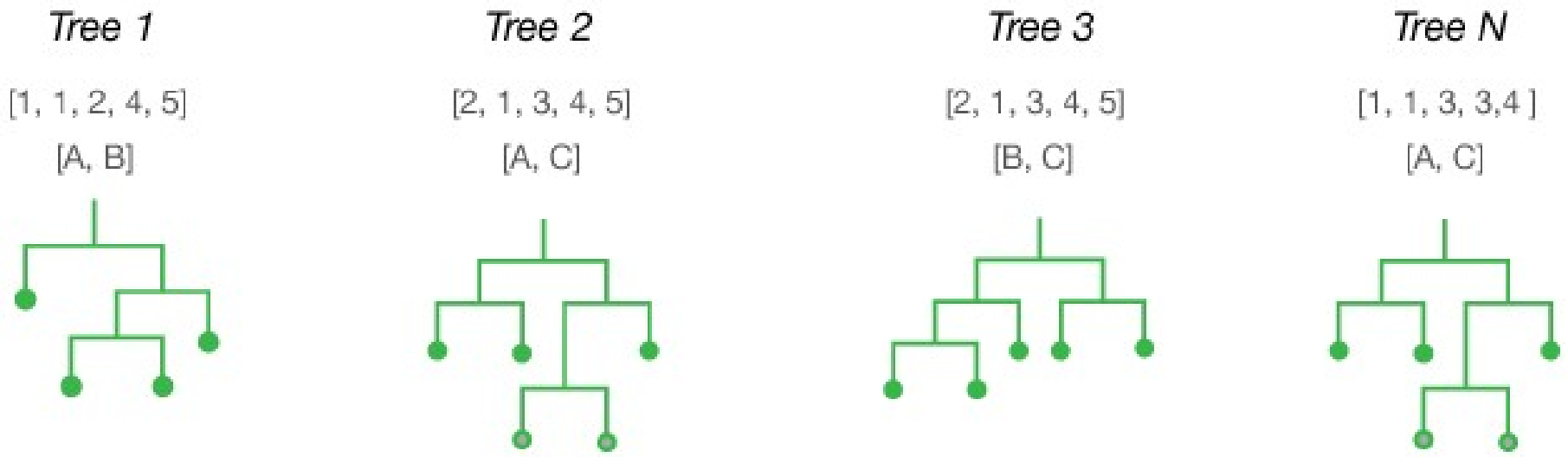

Start
Length>50
No
Yes
Eyes=="blue"
Is Bird?
No
Yes
No
Yes
Red
Green
Carnivore?
4 legs
No
Yes
No
Yes
Red
Green
Red
Red

Random Forest

Bagging



CART models offer a more interpretable output than compared to Random Forests. However, the relative importance of each predicator variable can be displayed. The graph above shows the results of 1000 different Random Forest models in regards to predicting survival on the Titantic. The relative importance of each predictor variable is displayed. The output is grouped into 3 clusters in order to display variation between the different models.



The diagram above assumes five observations [1, 2, 3, 4, 5] and three predictor variables [A, B, C]. It shows the construction of four simple (but different) trees. The observations have been sampled with replacement which means that some observations can occur more than once. In the scenario above, two predictor variables are used to grow each tree, rather than using the entire set of predictor variables.