**Test Summary**

This test is to check that the output of the brt model SDM experiment is accurate. The modeling script changes the default parameter arguments. Evaluation measures are included.

**Preconditions & Test Data-sets Required**

occur.csv

bkgd.csv

current climate layers (current.76to05 | bioclim01-bioclim19.tif)

**Test Steps**

1. Log in to the BCCVL
2. Select Experiments tab
3. Click new SDM Experiment
4. Enter “Phascolarctos cinereus brt model config2” as the name for this experiment.
5. Enter “Phascolarctos cinereus brt model with modified parameters, a current projection, and model evaluation” as the description of experiment
6. Click Next
7. Select Boosted Regression Trees under Species Distribution Model Production Algorithms
8. Expand configuration parameters
9. Check the default parameter are as follows:
   1. tree.complexity = 15
   2. learning.rate = 0.005
   3. bag.fraction = 1
   4. var.monotone = rep(1, length(predictors)) //19 bioclim variables
   5. n.folds = 10
   6. prev.stratify = TRUE
   7. family = “bernoulli”
   8. n.trees = 100
   9. max.trees = 25000
   10. tolerance.method = “auto”
   11. tolerance = 0.01
10. Click Next
11. Select Occurence Data for Phascolarctos cinereus
12. Click Next
13. Select Absence Data for Phascolarctos cinereus
14. Select Current climate layers for Australia, 2.5arcmin (~5km)
15. Click Next
16. Click start Experiment

**Expected output in files:**

1. model object.RData – binary model output file generate by model algorithm
2. current.tif – brt model projected onto current climate layers
3. combined.modelEvaluation.csv – table of all accuracy measures provided by dismo and biomod2 packages
4. brt.Rout – R output file, text file containing record of commands sent to R, generated automatically when using R CMD BATCH
5. AUC.png – Area Under the Receiver Operating Characteristic (ROC) Curve; threshold-independent plot of model predictive performance with test statistic value as figure title
6. \*\_response.png – response plot for each current climate layer used during model creation
7. biomod2\_like\_VariableImportance.csv – table of variable importance using biomod2’s procedure
8. maxent\_like\_VariableImportance.csv – table of variable importance using maxent’s permutation importance procedure
9. results.html – html file displaying AUC.png and combined.modelEvaluation.csv

**Additional checking for model configuration:**

Option 1 (requires R to be installed on tester’s machine):

1. Find the model.object.RData file and double click on it to load it into an R session
2. Do an ls() at the prompt to determine the name of the modeling object.
3. Perform the following commands:
   1. brt$gbm.call$tree.complexity
   2. brt$gbm.call$learning.rate
   3. brt$gbm.call$bag.fraction
   4. brt$gbm.call$var.monotone
   5. brt$gbm.call$cv.folds
   6. brt$gbm.call$prev.stratification
   7. brt$gbm.call$family
   8. brt$trees.fitted
   9. <not available>
   10. brt$gbm.call$tolerance.method
   11. brt$gbm.call$tolerance

Option 2 (requires the modeling script to set “brt.verbose = TRUE”)

1. Find the brt.Rout file and open in a text editor
2. Search for the following phrases:
   1. <not available>
   2. <not available>
   3. <not available>
   4. <not available>
   5. “creating 10 initial models”
   6. “folds are stratified by prevalence”
   7. “using a family of bernoulli”
   8. “models of 100 trees”
   9. <not available>
   10. <not available>
   11. <not available>

**Comments**

Test Step #11,12 Names likely to change when using Gerhard’s koala data

Test Step #12 should be the selection of individual climate layers, not ones already packaged

Option 1 test for model configuration is based on the model object being called “brt”. If the name returned by ls() is different, then ‘brt’ needs to replaced in the remaining commands with the right name. Also, for h) parameter, you will only be able to determine the n.trees by looking at the first value returned.