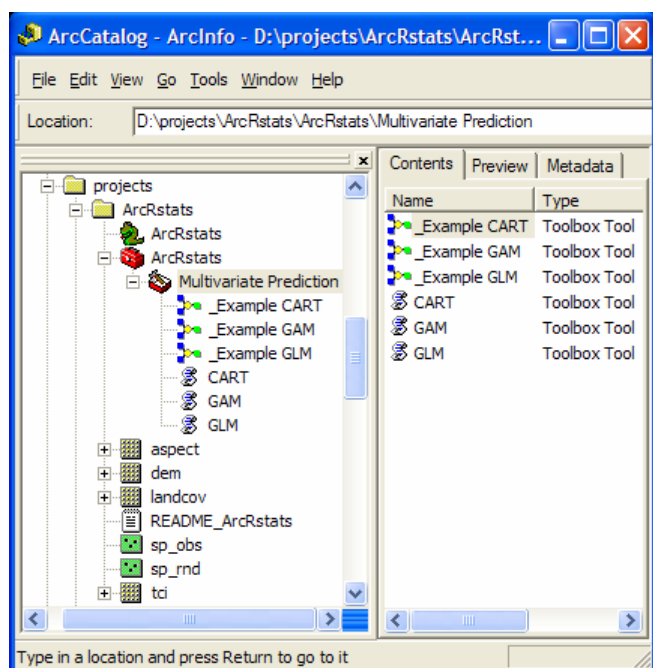


ArcRstats

Multivariate Modeling Script for ArcGIS

This script can be used with ArcGIS to produce predictive maps based on different techniques using the free and robust R statistical package:

- Generalized Linear Model (GLM)
- Generalized Additive Model (GAM)
- Classification and Regression Tree (CART)



This script can be used within the ESRI ModelBuilder environment. A basic script interface is included in the ArcRstats.tbx which can be viewed and used from ArcCatalog. Dragging the respective toolbox script (GLM, GAM or CART) into a new model allows you to connect the input points and rasters as well as defining the output prediction raster and sampling tables. Example data and model runs are included in the toolbox, which you can run by simply pressing the play button. Screenshots are available below in this pdf of these example models using the GLM, GAM and CART script interfaces within the ModelBuilder environment.

Models are inherently flawed. You are strongly advised to research the strengths and weaknesses of these different techniques as well as understanding their outputs, neither of which are explained with this tool. The help for each of the R commands should be consulted. See the out_md1.r file (where out_md1 is the name of your output prediction raster) to re-create the R session and try `help(glm)`, `help(gam)` or `help(rpart)` after loading the necessary package, which is `library(mgcv)` for GAM and `library(rpart)` for CART. Also look at the help for the `predict.glm`, `predict.gam` and `predict.rpart` functions. A good habitat modeling review paper for more background on these modeling techniques is:

Guisan, A., and N.E. Zimmermann. 2000, Predictive Habitat Distribution Models in Ecology: Ecol. Mod. 135 147-186.

Inputs / Outputs :

- in_absence - any point feature class
- in_presence - any point feature class
- in_rasters - on or more rasters
- out_md1 - output predicted raster

Requires :

- ArcGIS version 9 or higher (<http://www.esri.com>)
 - + Spatial Analyst extension
- R version 2 or higher (<http://www.r-project.org>)
 - + COM(D) Server for R (<http://cran.r-project.org/contrib/extra/dcom>)
- Python 2.1 or higher (<http://www.python.org>), included with ArcGIS 9 or higher
 - + win32com module (<http://starship.python.net/crew/mhammond>)

Term of Use:

This program is public domain under the GNU General Public License (www.gnu.org/copyleft/gpl.html). We provide this software with absolutely no warranty. If you use this, please cite with the following:

Best, B. D., S. Loarie, S. Qian, P. Halpin, D. Urban, 2005. ArcRstats - multivariate habitat modeling with ArcGIS and R statistical software. Available at <http://www.nicholas.duke.edu/geospatial/software>.

Authors:

Ben Best <bbest at duke dot edu>
Scott Loarie <srl6 at duke dot edu>
Song Qian <song at duke dot edu>
Patrick Halpin <phalpin at duke dot edu>
Dean Urban <deanu at duke dot edu>

Duke University Geospatial Analysis Program
<http://www.nicholas.duke.edu/geospatial>

Versions :

- 0.3 (2005-05-19):
 - fixed bug during GAM formula creation
- 0.2 (2005-05-08):
 - fixed projection of model output grid
 - changed from tempfile grid output to *.asc
 - cleaned up large variables from memory
 - validated _Example models so output not already generated in ModelBuilder
 - supplemented documentation
- 0.0.1 (2005-05-04):
 - first beta version

TODO : when feeding data to predict in R, remove unused coefficients from data frame (since NAs are excluded)

TODO : check in_absence/presence for raster or point locations, not just "dataset"

TODO : check for autodelete if exists in gp environment

TODO : add ROC to GLM/GAM...

TODO : work on other basic statistical tests, like T-tests, Moran's I stuff

TODO : failover R, launch R and capture real error using the R package "session"

TODO : check raster name length when creating reprojected (*_r) and integer (*_i) rasters

TODO : handle big grids with map algebra statement for GLM and CART

TODO : output error/deviance grids too

TODO : setup listserve for sending updates to users

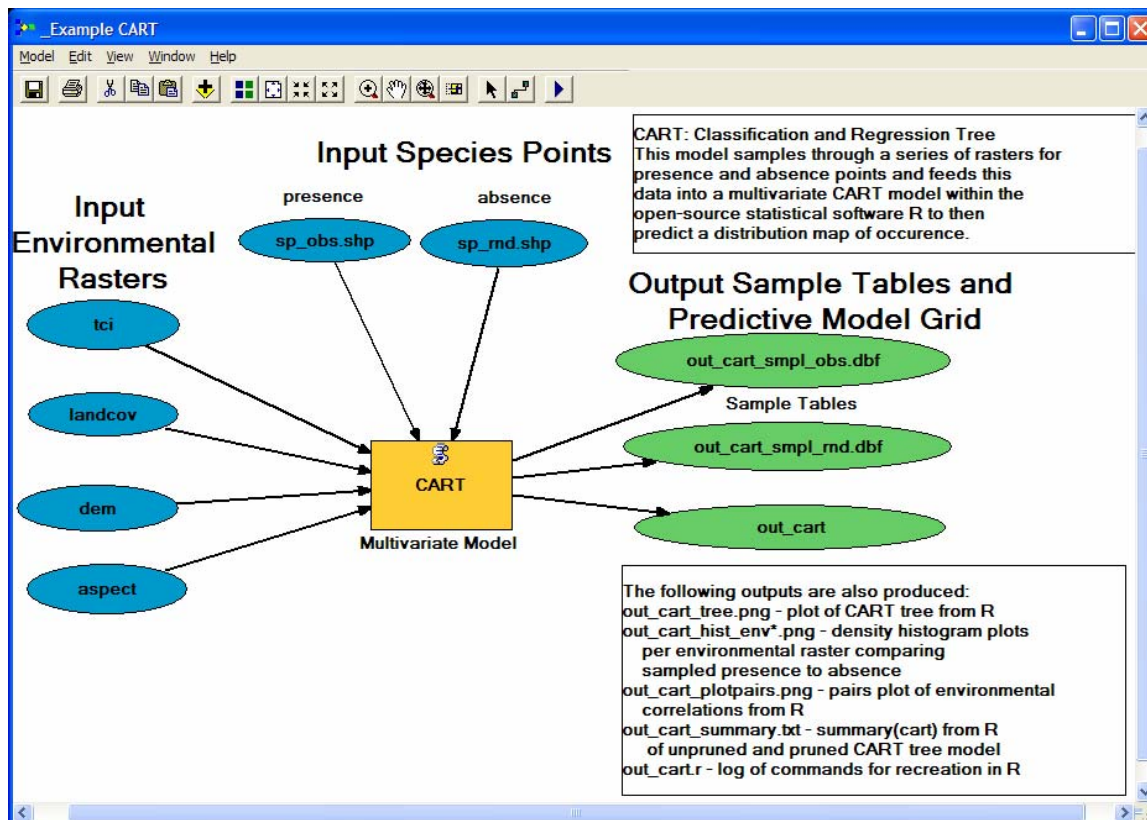


Figure 1. Example CART Model

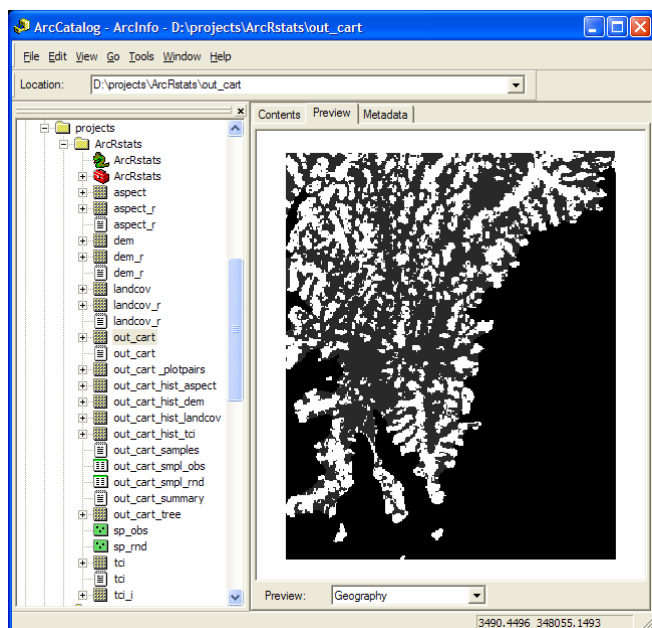


Figure 2. Output Prediction Grid from CART Example

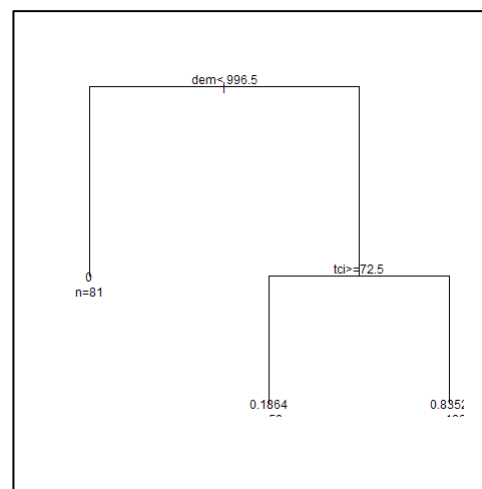


Figure 3. Ancillary Tree Plot for CART (*_tree.png)

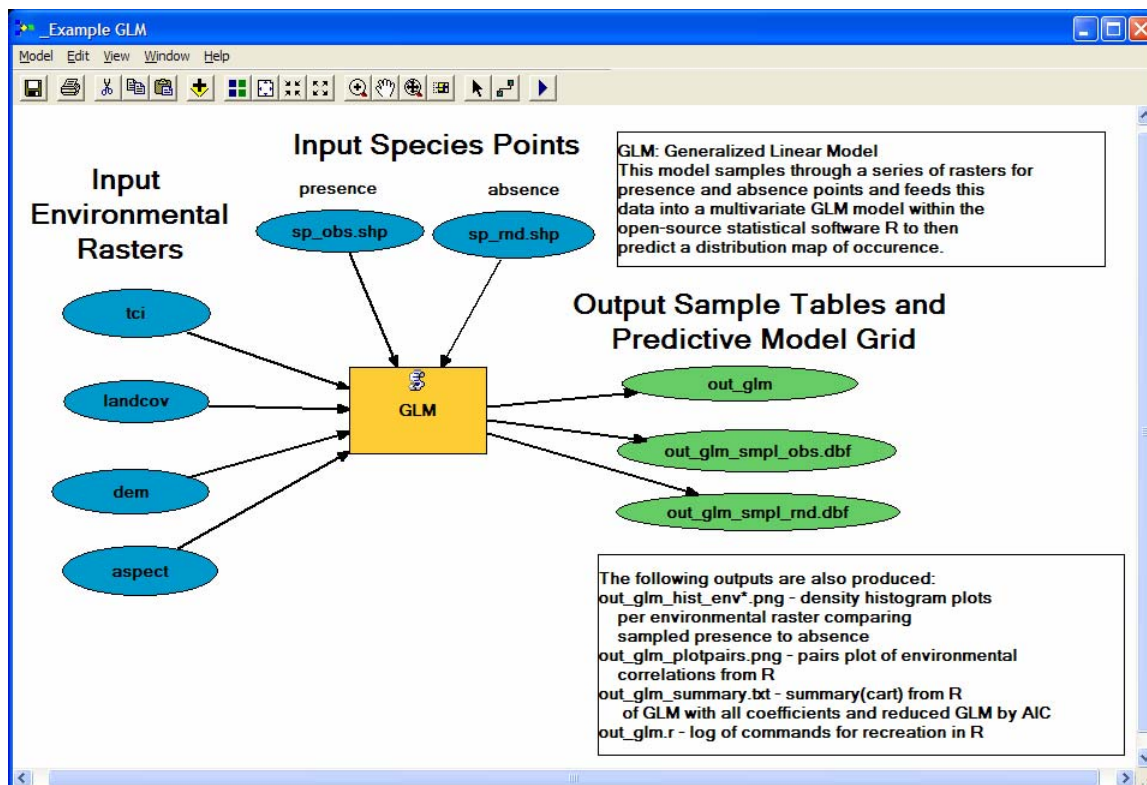


Figure 4. Example GLM Model

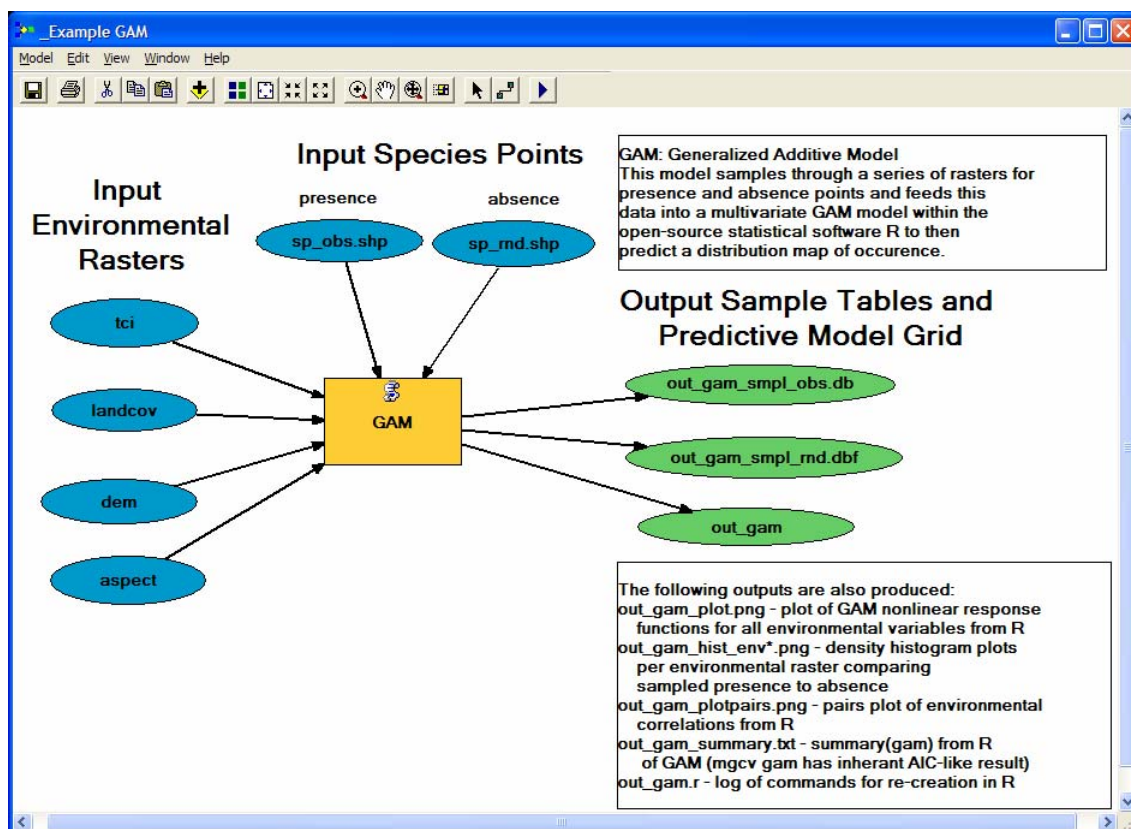


Figure 5. Example GAM Model

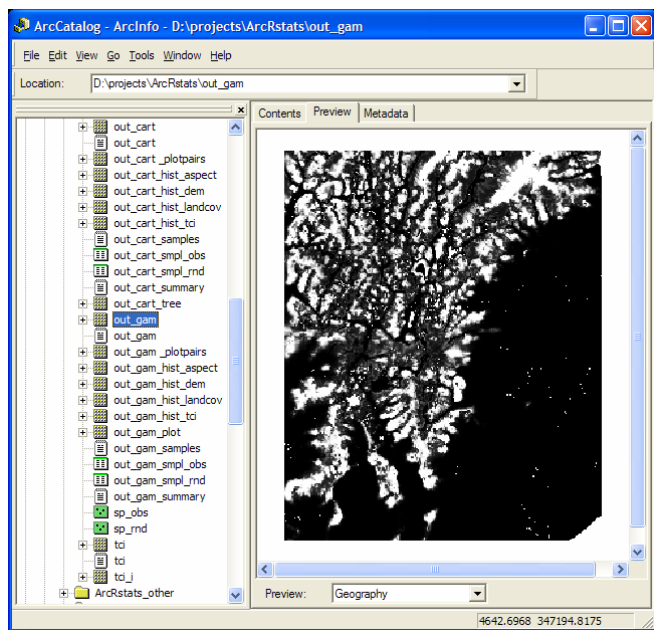


Figure 6. Output Prediction Grid from GAM Example

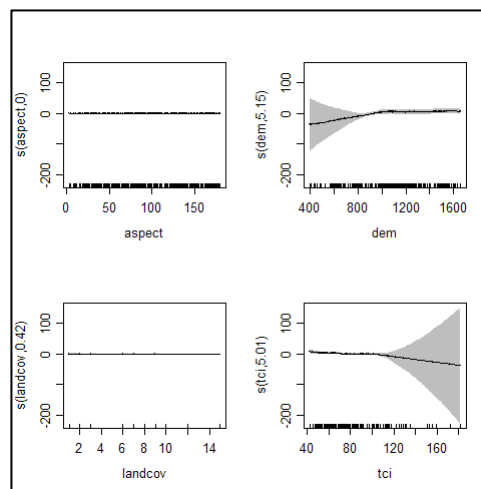


Figure 7. Ancillary GAM Plot for GAM (*_plot.png)

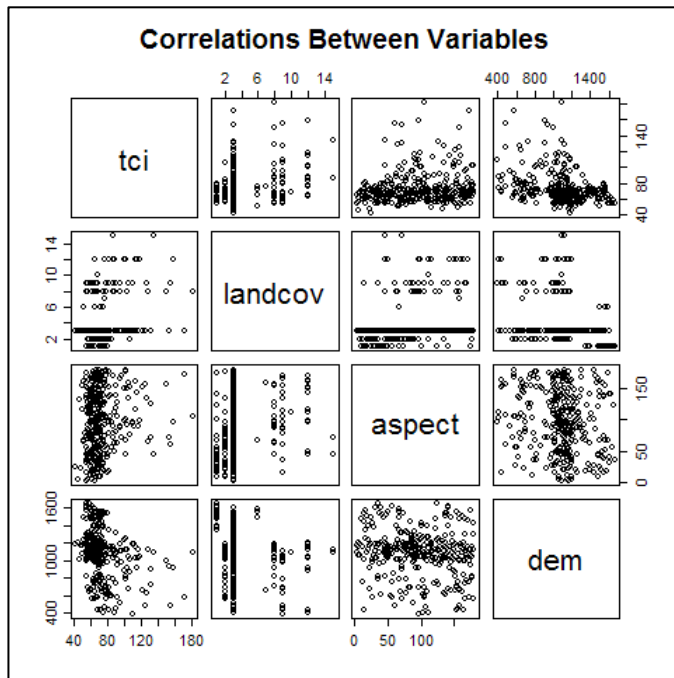


Figure 8. Ancillary Pairwise Plot for All Models (*_plotpairs.png)

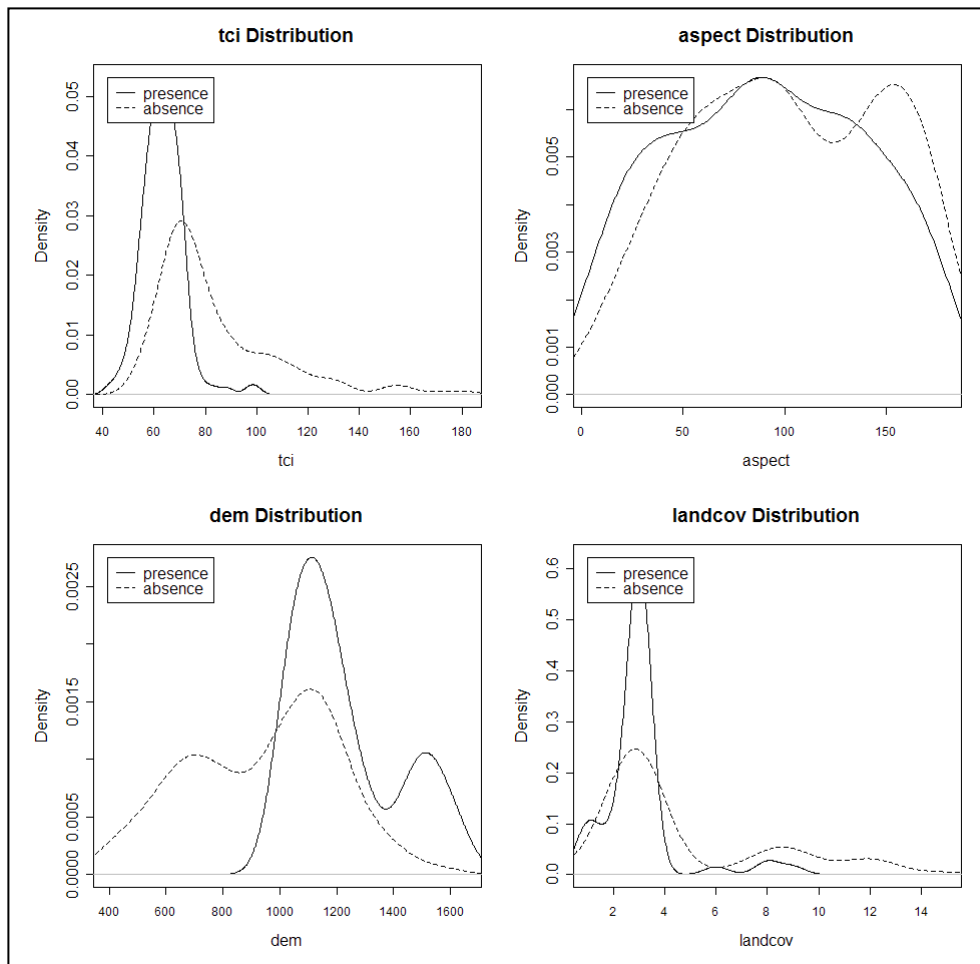


Figure 9. Ancillary Density Histogram Plots for All Models (*_hist_*env*.png)