**Ecol 8990**

**Assignment # 5**

**Due Fri Nov 17, 5 pm**

**Note: The work should be individual. Use R Markdown to complete the assignment. The Markdown (.Rmd) file itself is part of the assignment.**

Load the .cvs file “SNP\_TC\_change.csv” into a dataframe. The column names for this dataset are as follows:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | **Variable** | **Description** | | ID | Unique ID | | X | Eastings | | Y | Northings | | MAP | Mean Annual Precipitation (mm) | | DRY.RATIO | Dry:wet season rainfall | | FIRE | 2000-2013 fire frequency | | ELEV | Elevation (m) | | ELEV.SD | Elevation SD (m) | | SLOPE | Slope (°) | | CURVAT | Curvature | | WCI | Woody cover change index | |

The dataset is part of an analysis of tree cover change (WCI, an index of annualized rate of change in percent tree cover) in Serengeti as a function of a number of environmental covariates (MAP, FIRE, etc.). We also have spatial locations (X and Y). X and Y are in UTM coordinates and have large values (they are in m). To do this exercise I highly recommend that you rescale and center X and Y the way we have done with other variables in the past (e.g., MAP in one of our *glmer* examples). You can this by subtracting the mean and dividing by the standard deviation for each variable. You can do much of this assignment by drawing on our in-class ‘Spatial\_regression.R’ file.

**1**- Produce variograms for MAP and FIRE. Test for significant spatial autocorrelation using Moran’s I. What is the range of the spatial autocorrelation of the model residuals?

**2**- Fit two models relating woody cover change (WCI) to MAP and FIRE using *gls*, one with and one without spatial autocorrelation. Compare them with a Likelihood Ratio test.

**3**- Fit the same model again, but test alternative autocorrelation functions (Spherical, Exponential). Do some fit better than others? How do the semivariograms change?

**4**- Conduct a model selection exercise using *gls*. What is the top model? Examine the semivariogram of your top model using normalized residuals. Does your model do a good job accounting for autocorrelation?

**5**- Now rerun your models in an OLS framework by thinning your data set. You previously estimated the range of the semivariogram for one of your models. Thin your data using sample so that on average any two points are at least this distance apart. How might you do this?