Executive Summary

Geog 560

Jake Kruse

## Random Forest Prediction of Deforestation in Para, Brazil

**Motivation & Goal**

Deforestation in the Brazilian Amazon has occurred over centuries, but it has dramatically increased in the last decade (Fearnside and Barbosa 2004). The state of Pará is one of the most problematic states in Brazil in terms of deforestation, and therefore assessing deforestation risk in Pará is crucial to succeed in preserving forest. Between 2001 and 2018, Pará lost an area of 79,000 km2 which is 37% of the area deforested in the Brazil Amazon for the same period (INPE, 2018). To better understand the variables of deforestation and to develop a model that predicts deforestation in that region, random forest classification was performed on a dataset for the state of Para.

**Required data**

Publicly available datasets PRODES and SICAR were used for the land classification, infrastructure, Human Development Index (HDI), and population. Specifically, the following predictors were used: Conservation areas, Distance to Dams (DistoDams), Distance to Rivers (DistoRivers), HDI, Population, Land Class (LandUse), and Distance to Roads (DistoRoads).

**Methods**

A random forest model was trained on 500 random sample points, and predictions were made for 1000 points. Both drew from the same raster stack, which included a land classification layer for the actual class information. Euclidean distance to features (rivers, roads, and dams) was calculated in ArcMap; shapefile to raster conversions were also done in ArcMap. All testing and training of the model was performed in R Studio.

**Results**

With the random forest model overall prediction accuracy was 74%, and incorrect classification was 26%.

**References**

Fearnside, P. (2017). Deforestation of the Brazilian Amazon. Oxford Research Encyclopedia of Environmental Science. doi:10.1093/acrefore/9780199389414.013.102