

Presentation of preliminary results for French Guiana

Camille Dezécache

Université des Antilles et de la Guyane - UMR Ecofog - CIRAD

camille.dezecache@ecofog.gf



May 3, 2015

Objectives of the presentation

- ▶ Methods
- ▶ Random Forest classifier
- ▶ Results concerning location of deforestation
- ▶ First hypothesis concerning the intensity of deforestation

Methods: Variables

- ▶ Explained variable
 - ▶ Deforestation between 2004-2008 (Hansen deforestation map)
- ▶ Explanatory variables
 - ▶ Distance to previous deforestation (2000-2004, Hansen)
 - ▶ Distance to roads
 - ▶ Distance to forest tracks
 - ▶ Distance to main cities
 - ▶ Distance to nearest creek
 - ▶ Protected areas
 - ▶ Permanent Forest Estate
 - ▶ Flooded areas
 - ▶ Greenstone belt
 - ▶ Mining autorizations (small scale) and mining titles (large scale)
 - ▶ Slope and elevation
 - ▶ Guiana Space Center

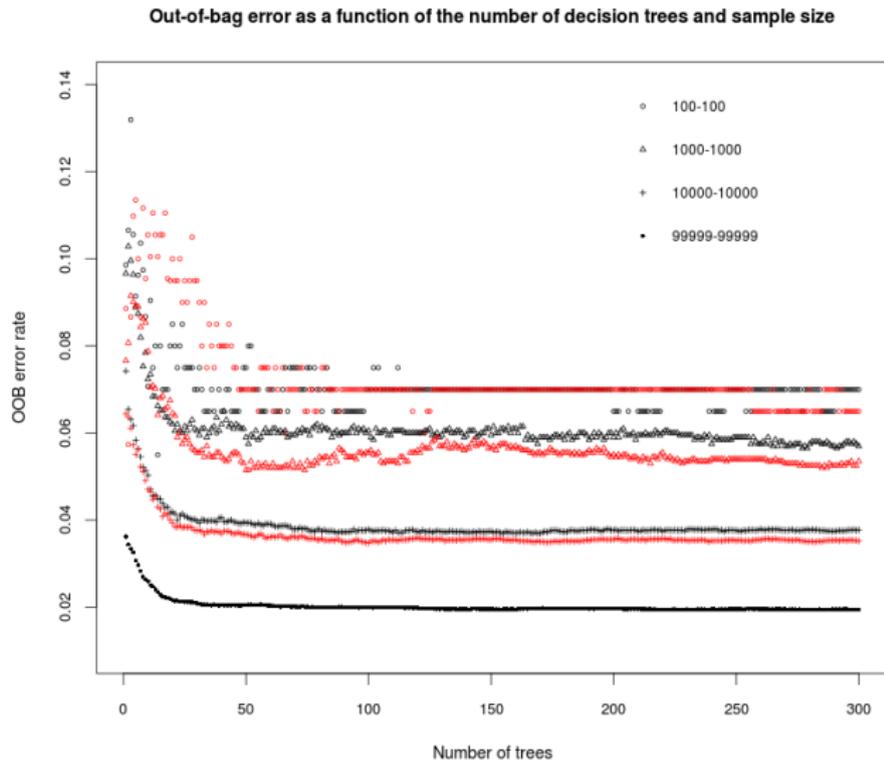
Methods: Sampling options

- ▶ Sampling all pixels
 - ▶ Impossible for technical reasons
- ▶ Sampling a minimal number of pixels at which we are sure at 95% to find the observed deforestation rate with low error (Vieilledent et al.)
 - ▶ Low deforestation implies that minimum sample is almost equal to total sample (see above)
 - ▶ Moreover model would fit the majoritary class (no deforestation) where we want to understand why drives deforestation
- ▶ Define an arbitrary sample
 - ▶ Take a sample equal to the number of deforested pixels per period + an equal number of not deforested pixels (penalize errors on no deforestation as well as deforestation classification)

Methods: Random Forest (Breiman, 2001) classification

- ▶ Classify each pixel based on multiple decision trees (forest)
- ▶ Each tree: subsample of total sample, and each node: sample of all explanatory variables, keep best one for splitting (random)
- ▶ Advantages:
 - ▶ Much better classification accuracy compared to logistic regression
 - ▶ Able to take into account multiple interactions
 - ▶ No overfitting
 - ▶ Allow to display relationship between each explanatory variable and the explained variable
 - ▶ Calculate an index of variable importance
- ▶ Disadvantage: not a statistical method

Results: Classification error rates

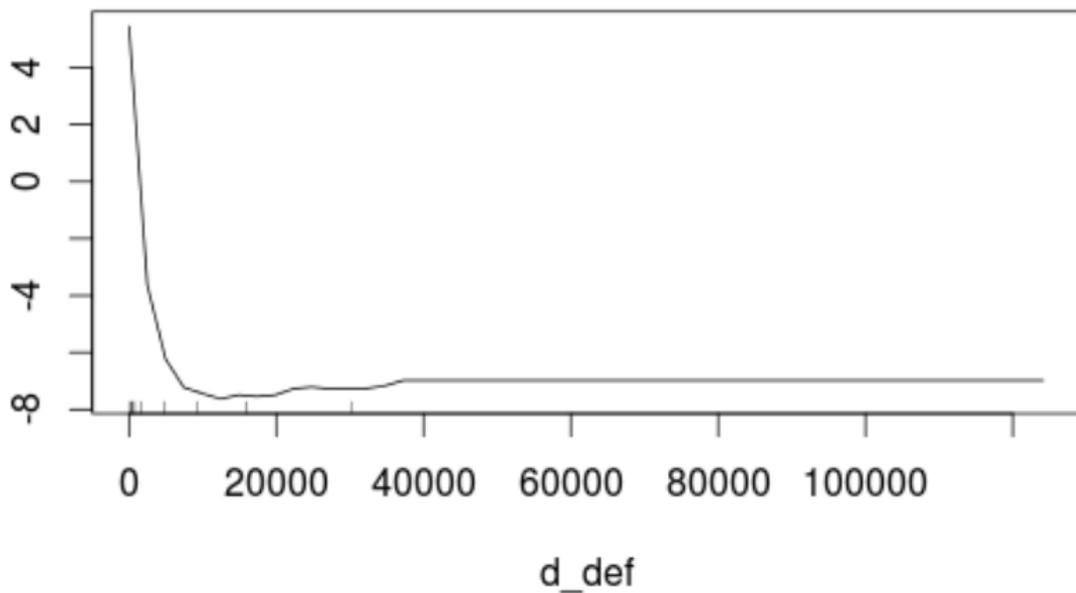


Results: Variable importance

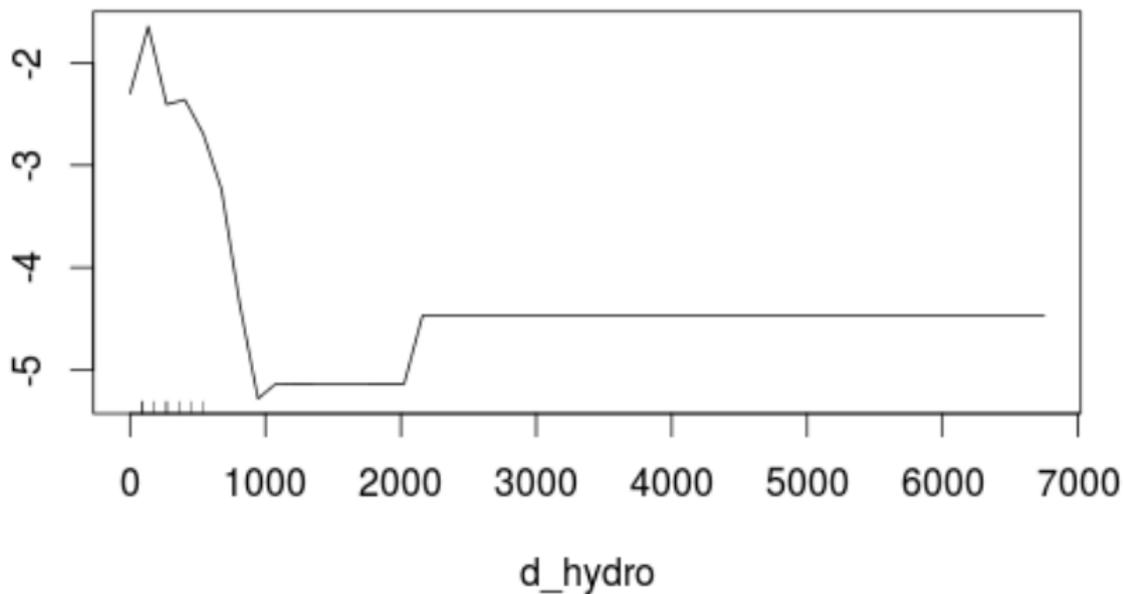
Mean decrease in accuracy: change in accuracy when a given variable is permuted (decreasing order of importance)

- ▶ Distance to previous deforestation
- ▶ Distance to nearest creek
- ▶ Slope
- ▶ Small scale mining authorization
- ▶ Distance to road

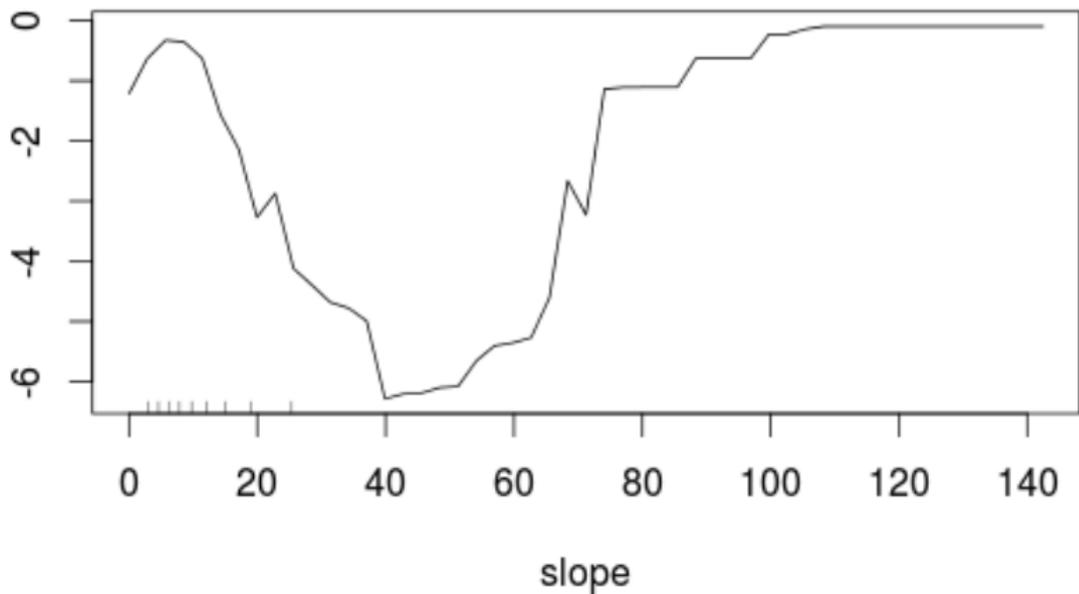
Partial Dependence on d_def



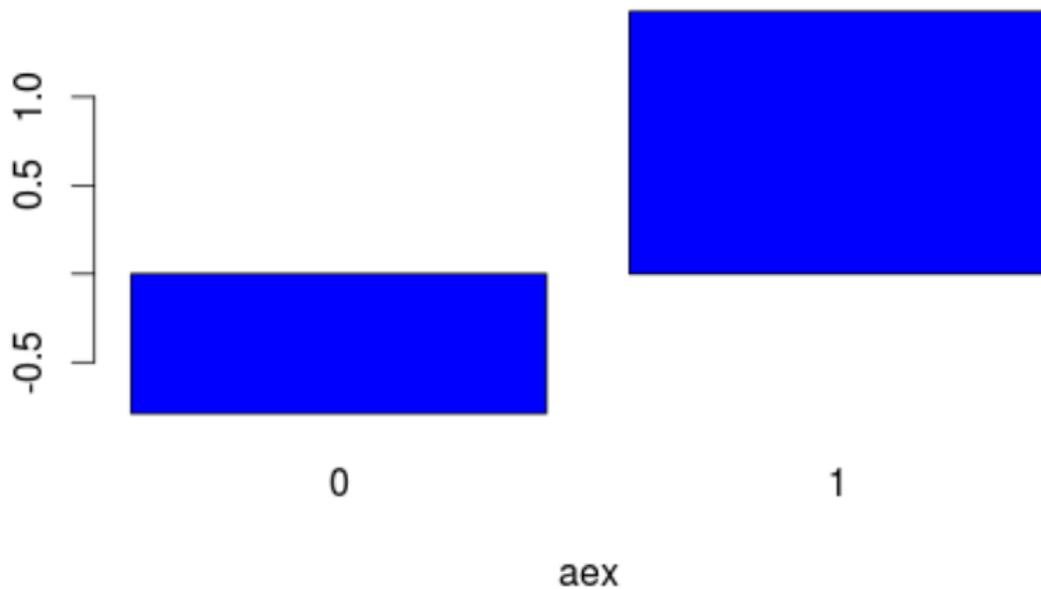
Partial Dependence on d_hydro



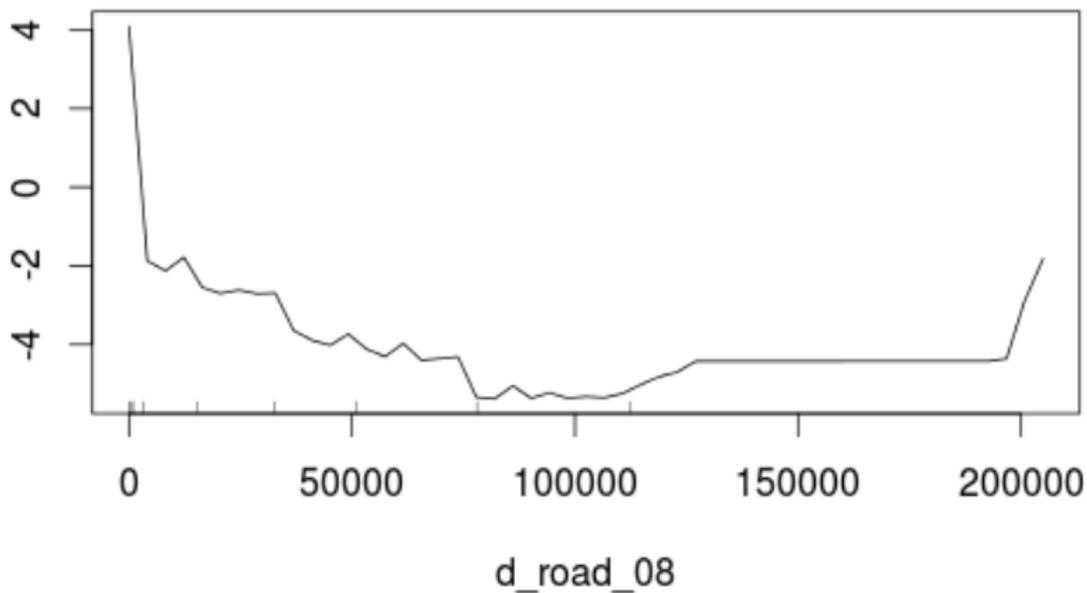
Partial Dependence on slope



Partial Dependence on aex



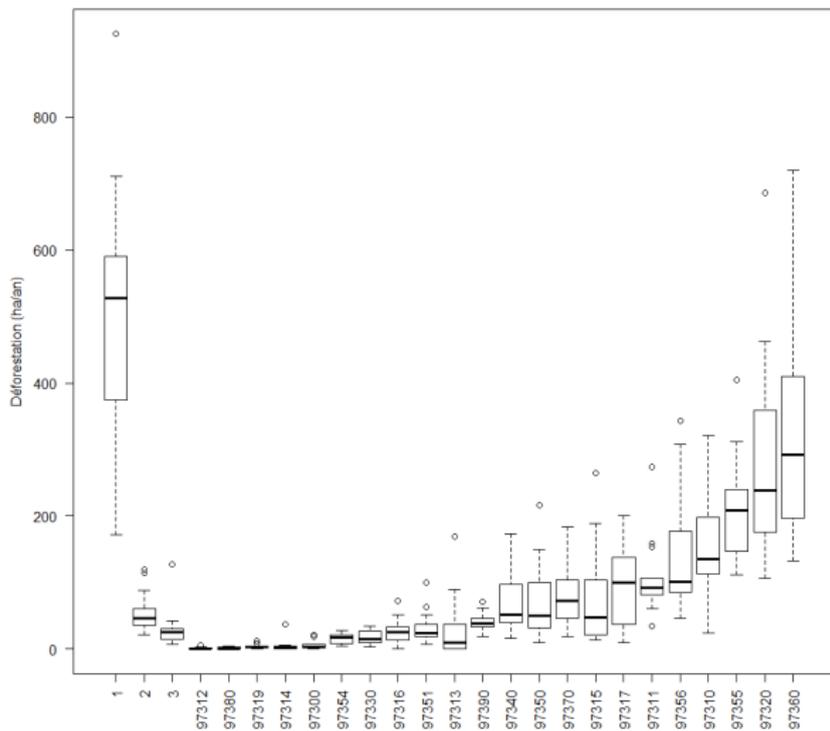
Partial Dependence on d_road_08



Results: Deforestation risk map

See QGis

Predicting the intensity of deforestation



Population growth and agricultural subsidies as major drivers of deforestation

See 3D plot in R Adding population density is impossible due to big differences in district areas. Considering population per area non-forested (excluding natural non-forested areas)?

Deforestation intensity in the Permanent Forest Estate and Gold-mining areas

- ▶ No model yet
- ▶ Permanent Forest Estate: tracks opening and wood production?
- ▶ Gold-mining: Gold prices and political/military intervention?

Conclusions, perspectives

- ▶ Population growth and agricultural subsidies possible drivers of deforestation in other countries?
- ▶ Forest Estate? Gold-mining areas?