Stanford University

EPFL

Design Project Final Presentation

Spatial data science to better understand anthropogenic drivers of Forest Carbon Storage

•••

EPFL students:
Aurélie de Jong, Raphael Renevey, Alma Vexina Wilkinson
External Supervisor : Charlotte Weil
EPFL professor: Athanasios Nenes

10.06.2020

Project Description

Motivation

- Carbon storage prediction is relevant for climate change research
- Some variables influencing carbon storage are not yet taken into account in prediction models

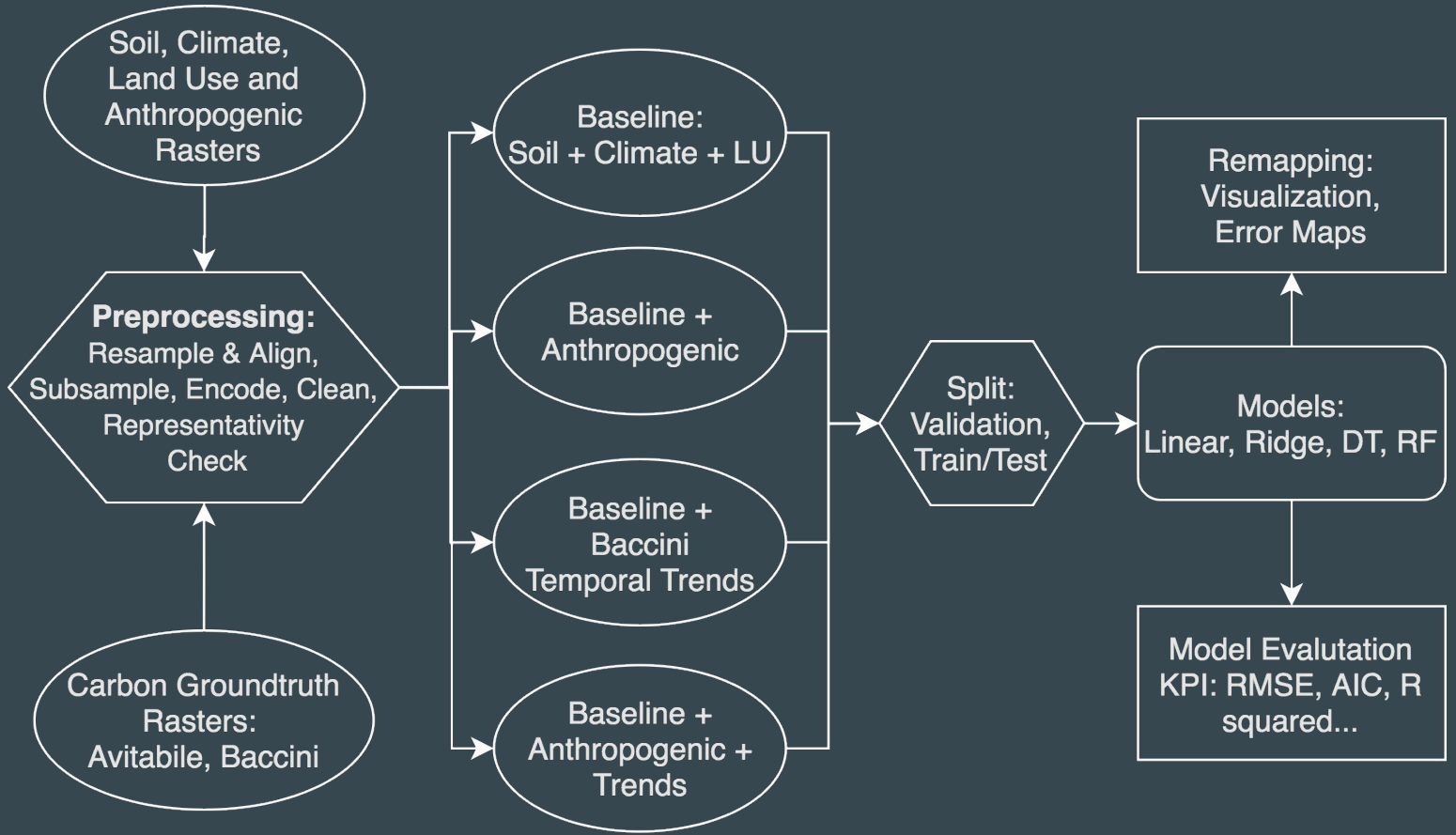
Research questions

- What are the anthropogenic variables impacting Carbon Storage the most?
- To what extent can these anthropogenic variables help improve the accuracy of models predicting Carbon Storage?

Approach

- Data Exploration of global, medium resolution datasets of the 2000s
- Model iteration for prediction of AGB density at a global scale using new variables

Modeling Pipeline



Baseline Inputs: Soil, Climate and Land Use

➔ **Soil : SoilGrids2019**

Texture, nutrients, pH, depth to bedrock, soil carbon content, water availability

➔ **Climate: Wordlclim version 2.1**

Variables linked with temperature and precipitation

➔ **Land Use : ESA 2014**

reclassification in 13 land use classes, such as Forest, Arable Land, Permanent Cropland, Transition Cropland or Artificial surface

data sources: WorldClim version 2.1 , SoilGRIDS 2019, ESA land use maps

Anthropogenic variables

- ➔ Analyzed variables: Nighttime Lights 2014 and 2015, GDP per capita, population density and 8 livestock datasets
- ➔ Strongest relation to AGB according to Pearson's R and Spearman's R:
  - ◆ Negative: Nighlights VIIRS 2014 and GDP
  - ◆ Positive: Chicken and Sheep
- ➔ For the other variables, near 0 correlation coefficients are observed. However, all correlations are statistically significant. Therefore, all variables are used as inputs in the model

data sources: VIIRS Nightlights, GDP, pop. density, livestock

Temporal trends from Baccini dataset

Baccini dataset includes Aboveground Biomass Density data from 2003 to 2018.

This, means it is possible to derive new model inputs based on an analysis of the story of each pixel in terms of carbon change through time, from 2003 to 2015.

Thus 3 temporal variables were created:

- ➔ land use trend
- ➔ carbon change trend
- ➔ land use and carbon trend combined.

Avitabile, Baccini models comparison

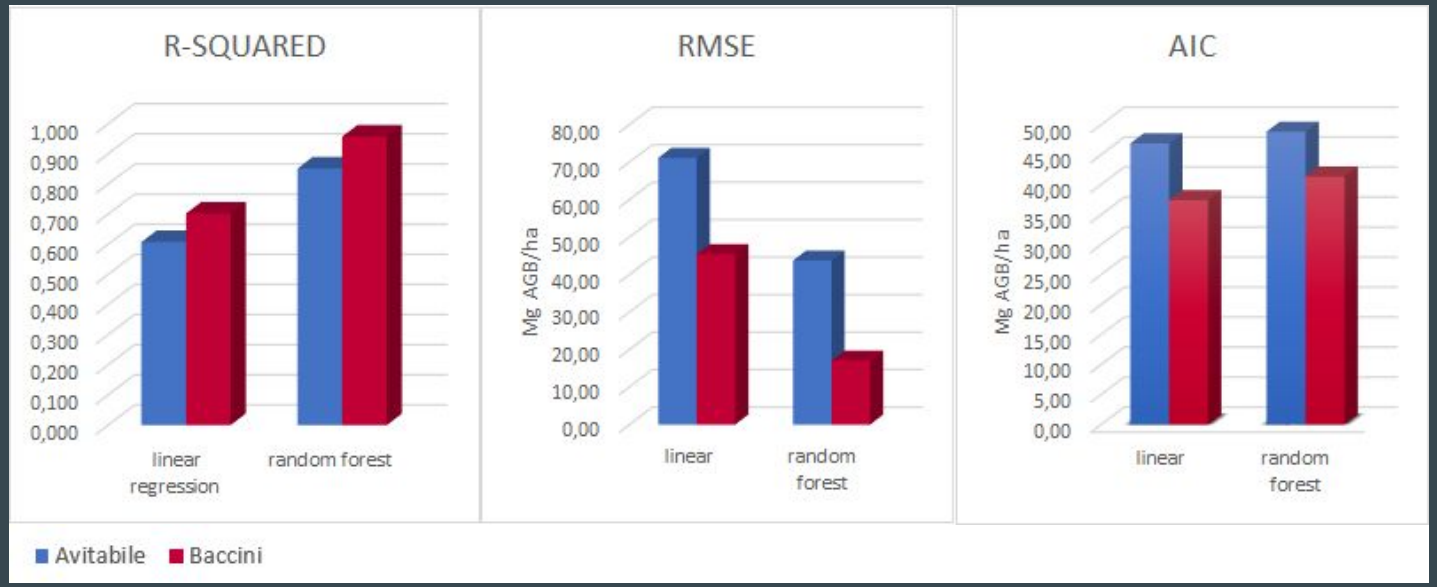
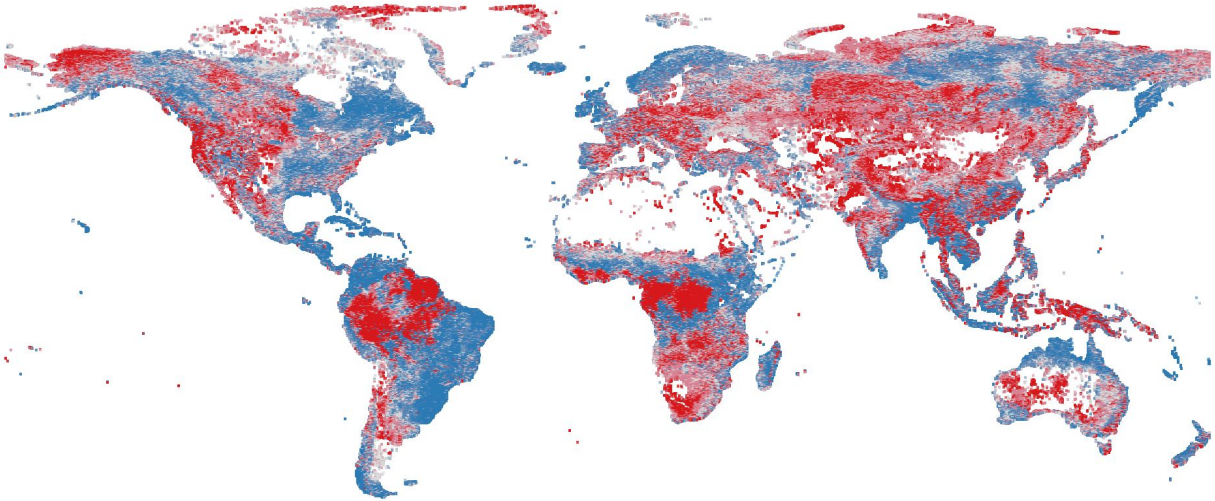
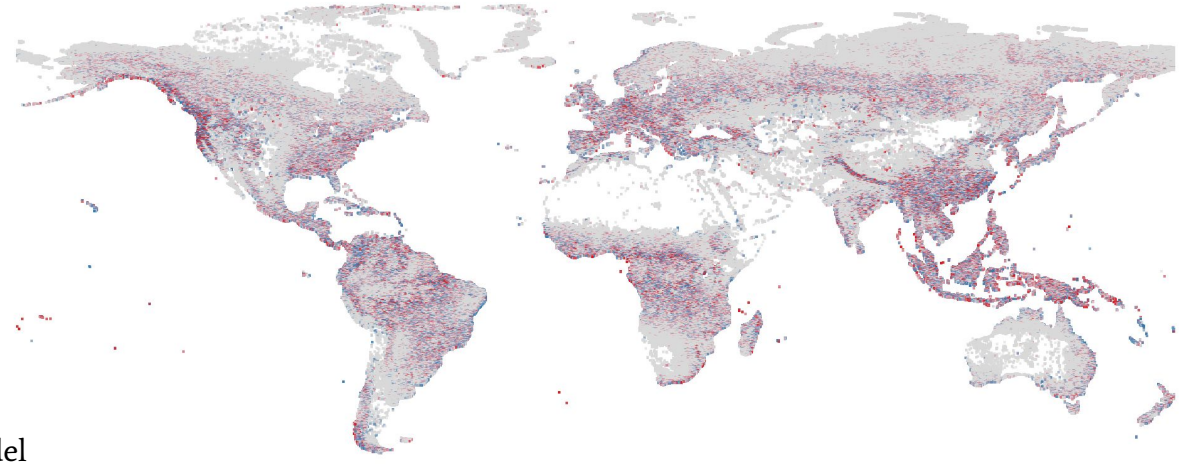


Figure 1: Results produced from linear regression and random forest baseline models run on Avitabile and Baccini as ground truth

Error maps Baccini



Baseline **Linear** Model



Baseline **Random Forest** Model

Error in AGB density [Mg/ha]

- -331 - -33
- -33 - -11
- -11 - 7
- 7 - 29
- 29 - 924

Baccini, model combinations comparison

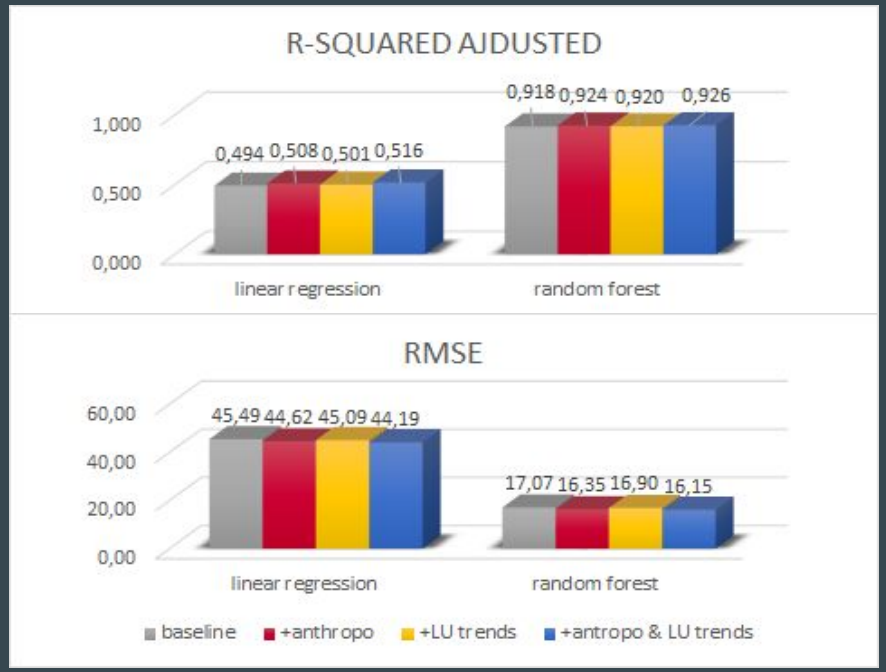


Figure 2: Results produced by linear regression and random forest based on different input variables with Baccini

Features importance

Number of variables of each type		Baccini							
		baseline		*+anthropogenic*		*+trends*		*+anthropogenic & trends*	
		LR	**RF**	**LR**	**RF**	**LR**	**RF**	**LR**	**RF**
	LU	10	1	9	1	0	1	0	1
	Climate	6	10	6	8	3	9	3	6
	Soil	0	5	0	3	0	4	0	3
	Anthropo	n/a	n/a	1	4	n/a	n/a	0	4
	Trends	n/a	n/a	n/a	n/a	13	2	13	2

Table 1: Results of the top 16 best features distribution ordered by features types produced by different model combinations

Linear regression:

- ➔ 3 climate variables stays: temp annual range, min temp of coldest month, max temp of warmest month
- ➔ anthropogenic: horse
- ➔ temporal trends: 13 variables

Random Forest:

- ➔ Forest is the only LU categories, mostly climate plus some soils (phh2o, bdod, soc, BDTICM, cfvo).
- ➔ anthropogenic factors (cattle, pig, sheep, GDP).
- ➔ temporal trends : reforestation from cropland and carbon gain

Anthropogenic impact on carbon

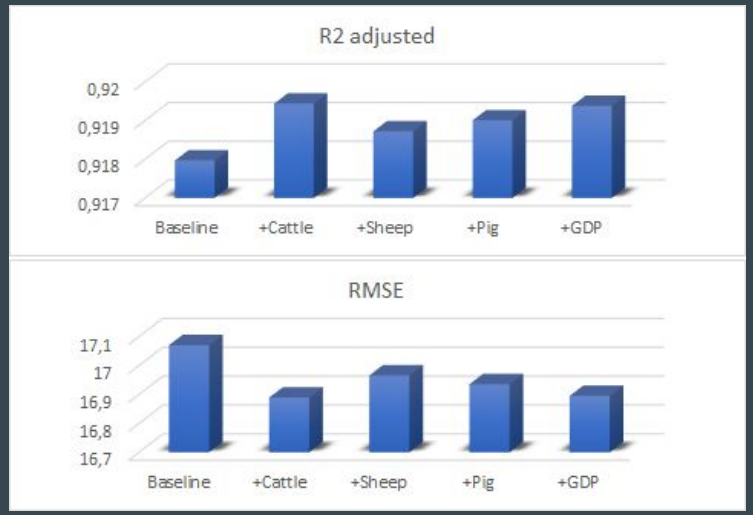


Figure 3: Results produced by random forest run on Baccini baseline with one of the most influencing anthropogenic factors

	AGB increase [%]	Q98% value	Q98% unit
Cattle	4.37	84	animal/ha
Sheep	7.69	66	animal/ha
Pig	4.38	108	animal/ha
GDP	1.54	189535	USD/capita

Table 2: Results produced by random forest on Baccini baseline with addition of one anthropogenic factor

Conclusion

- ➔ Some Anthropogenic variables do have predictive power on AGB density
- ➔ Temporal trends have great predictive power on AGB in the linear model
- ➔ To further develop this project:
  - ◆ Further research on temporal trends
  - ◆ Add more anthropogenic variables (Fire datasets or Atmospheric pollution maps)