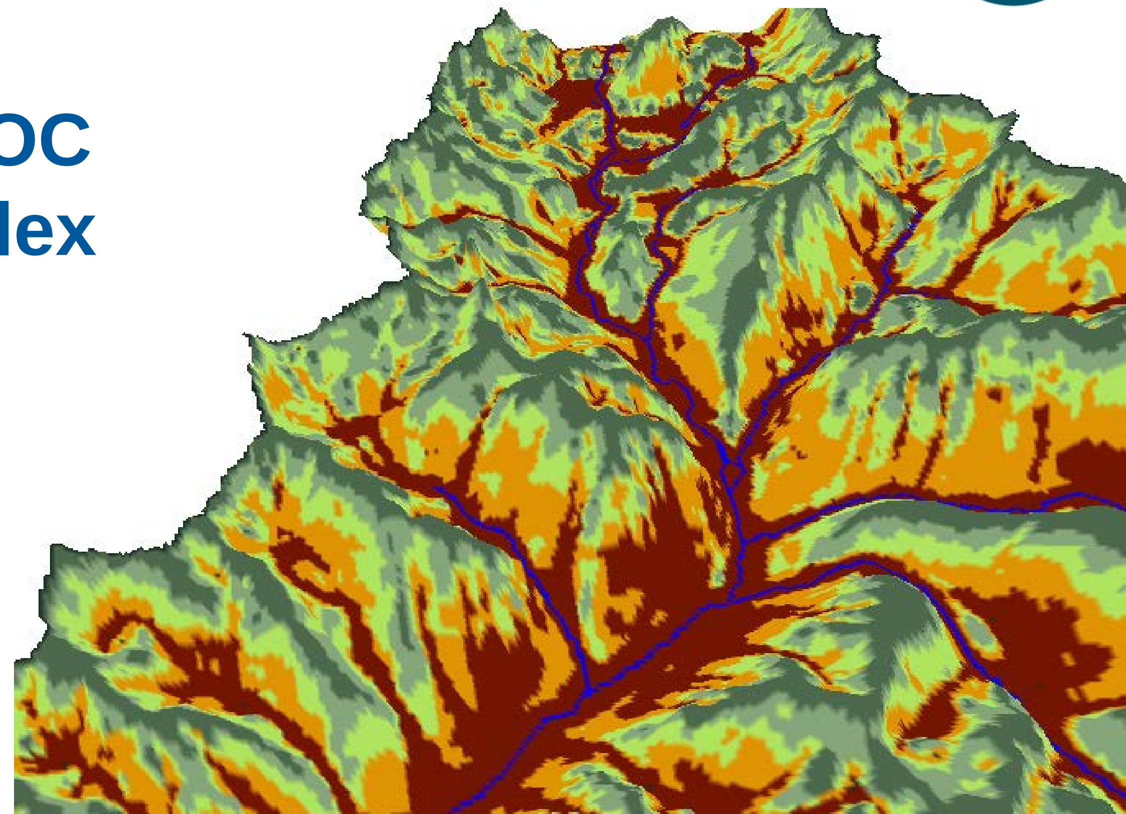
PLATFORM
FOR BIODIVERSITY AND
ECOSYSTEM MONITORING AND RESEARCH
IN SOUTH ECUADOR



Providing spatial SOC estimates for complex and remote soil-landscapes

Mareike Ließ

DSM Aarhus
27.06. - 01.07.2016

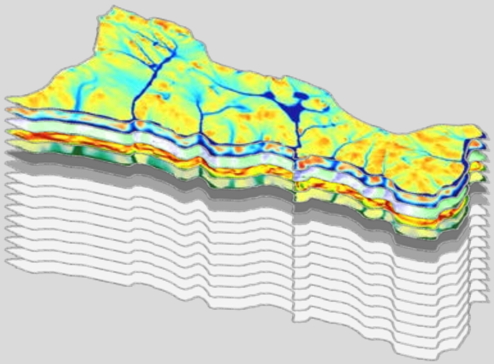
HELMHOLTZ
CENTRE FOR
ENVIRONMENTAL
RESEARCH – UFZ

Introduction

- Tropical forests are significant carbon sinks and their soil‘s carbon storage potential is immense
- Little is known about SOC stocks of tropical mountain areas whose complex soil-landscape and difficult accessibility pose a challenge to spatial analysis

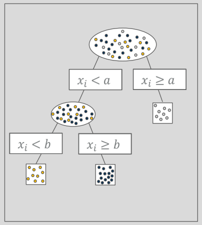
Mareike Ließ

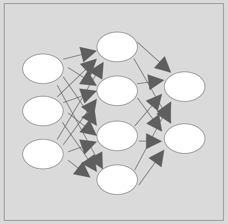
Predictors



- SCORPAN proxies: Landsat & DEM
- Search radii
- Selection procedures
- 236 predictors

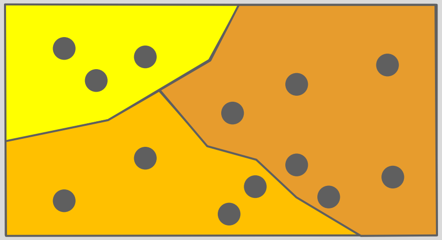
Machine Learning





tuning via grid search with repeated 10-fold CV (parallel proces.)

SOC data



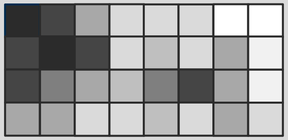
- Representative dataset/applicable sampling design
- Sampling cost
- Small size: 60

Validation

repeated 10-fold CV

Prediction

Position specific density functions



Mareike Ließ

Sampling design

In sampling for spatial prediction, the above all aim is to ensure that the **spatial variability of the soil is well-captured** without introducing any bias, while the design remains **feasible according to operational constraints**.

Adapt designs for areas of low accessbility:

1) Replace random points by accessible points in similar landscape positions

2) Adapt cLHS: quantile formation and point selection

3) Create landscape units by various objective methods => stratified random s.

4) Condense predictor space by factor analysis

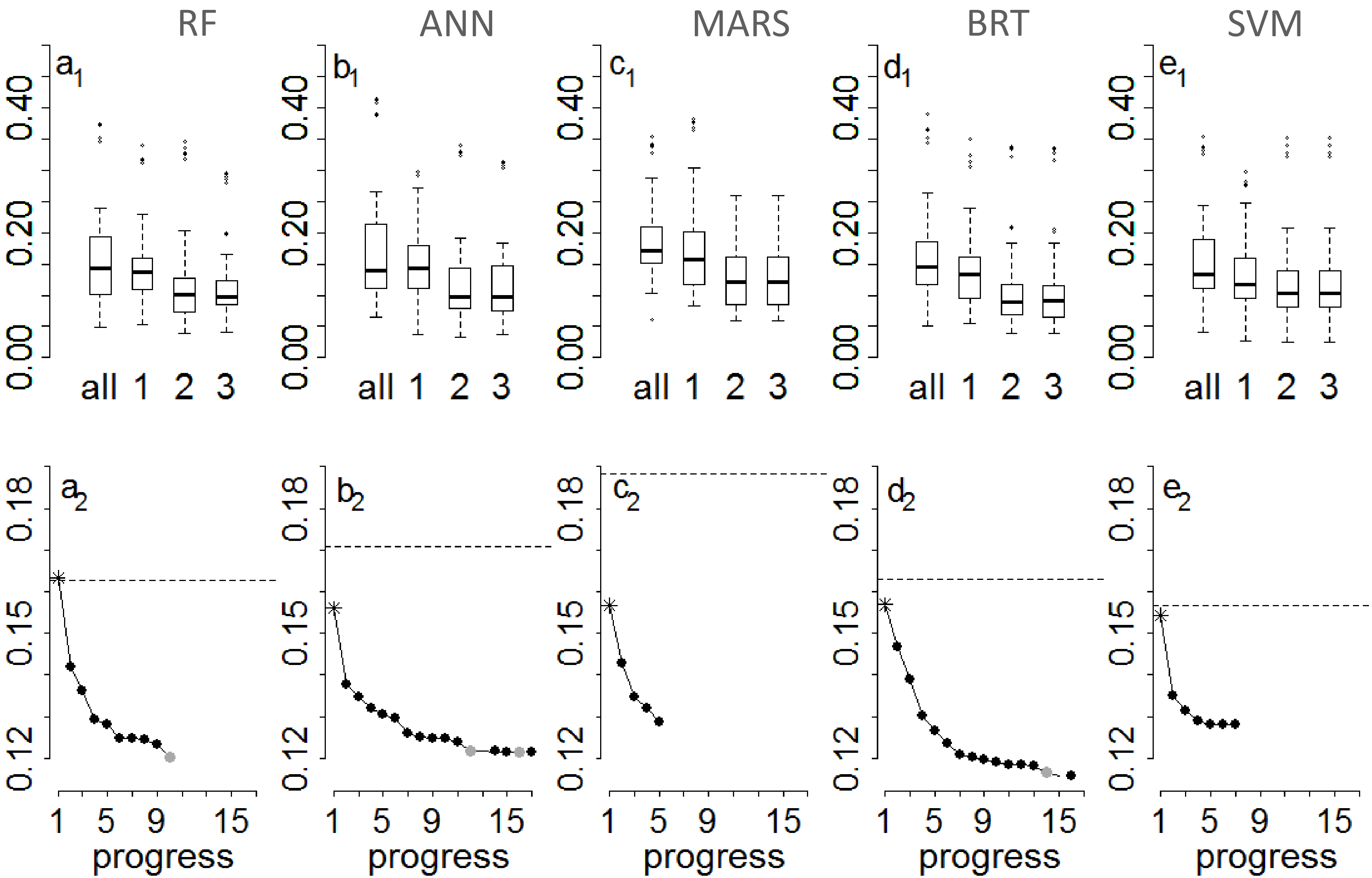
Predictor selection

- Best subset not possible due to the high number of predictors (236)

all all predictors

1. 10 single best predictors
2. Simple forward selection (non-linear correlation)
3. Three step forward selection procedure

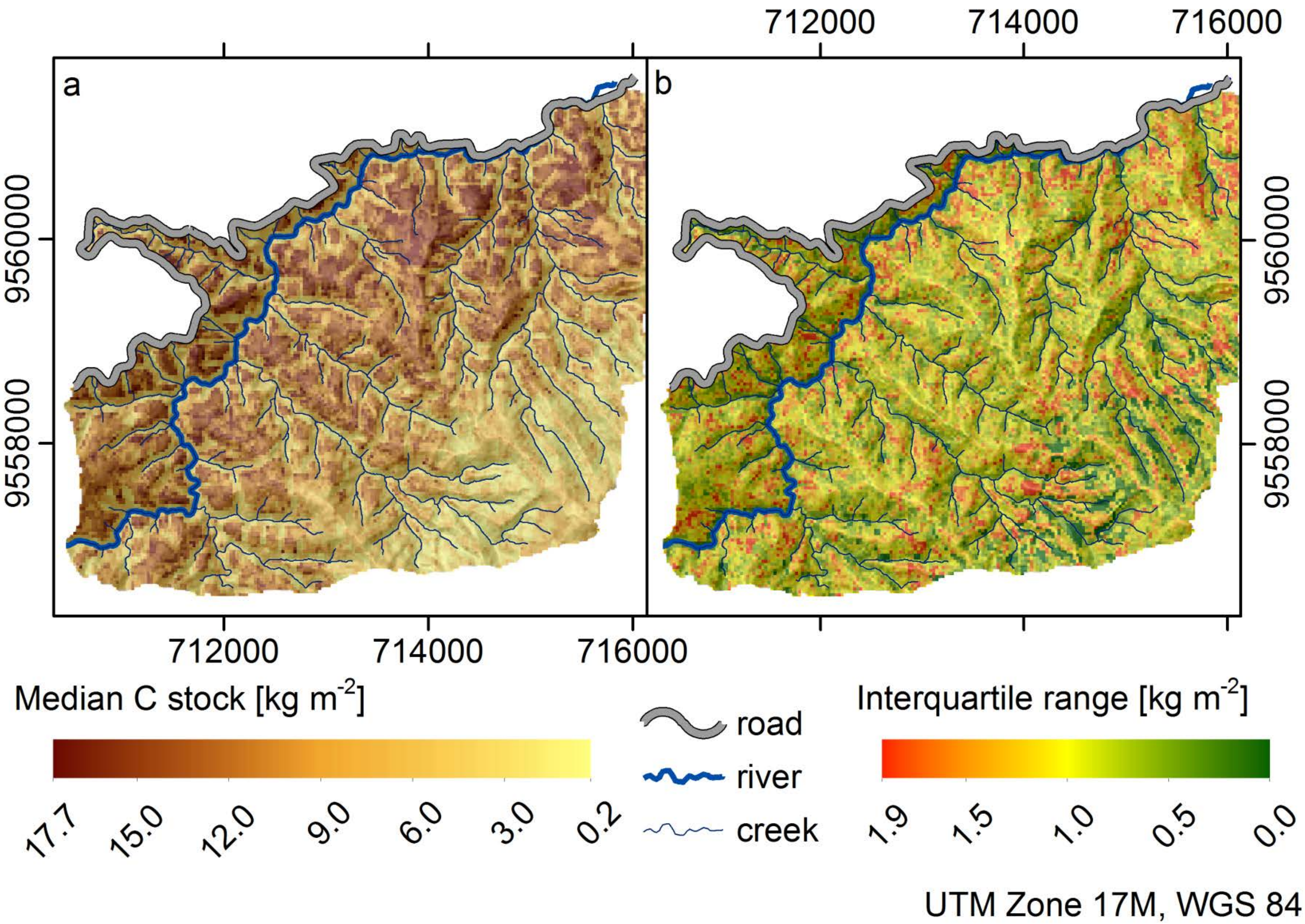
Predictor selection impact on predictive performance



Mareike Ließ

Spatial SOC stocks prediction

Position specific density functions



Mareike Ließ

Conclusions

- It is important to **make the most out of what you have** available: Predictors & soil data, ML, tuning parameter range and step size.
- In areas of complex structure and high sampling cost even more importance has to be given to the **sampling design**
- **Predictor selection** can reduce noise and improve predictive performance even for recursive partitioning methods
- The **position specific density functions** showed a higher predictive uncertainty in areas of high landslide frequency.

Research framework

Deutsche Forschungsgemeinschaft

PLATFORM FOR BIODIVERSITY AND ECOSYSTEM MONITORING AND RESEARCH IN SOUTH ECUADOR

DFG PAK 823-825 (www.topicalmountainforest.org): Research at global hotspot of biodiversity.

Science-directed development and implementation of suitable strategies and tools

(Basic research & Know-ledge transfer).

Ließ M (2015). Sampling for regression-based digital soil mapping: Closing the gap between statistical desires and operational applicability. Spatial statistics, 13: 106-122.

Ließ M, Schmidt J, Glaser B (2016). Improving the spatial prediction of soil organic carbon stocks in a complex tropical mountain landscape by methodological specifcations in machine learning approaches. PLOS ONE, 11(4).

Mareike Ließ