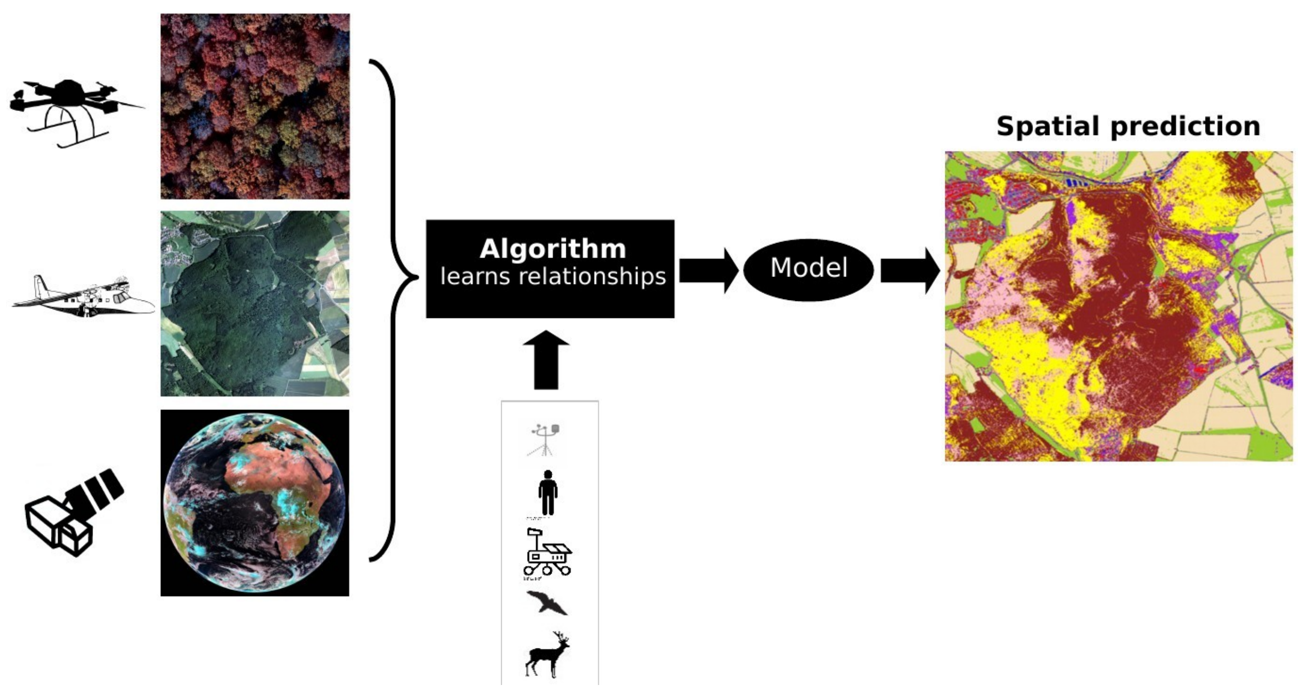
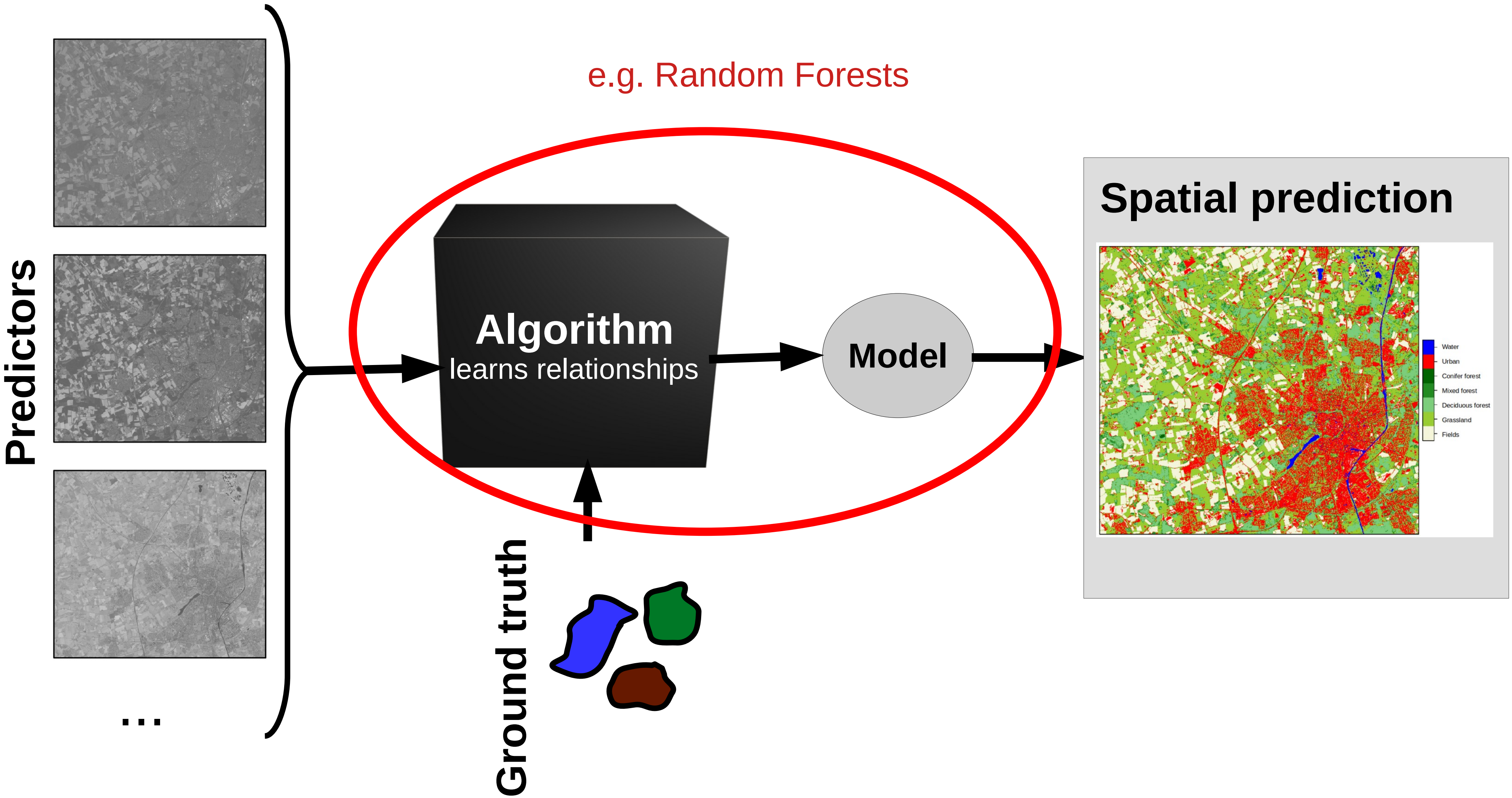
Remote sensing and machine learning:
Towards a spatio-temporal continuous monitoring of the environment

Hanna Meyer
Remote Sensing & Spatial Modelling,
Institute of Landscape Ecology, WWU Münster

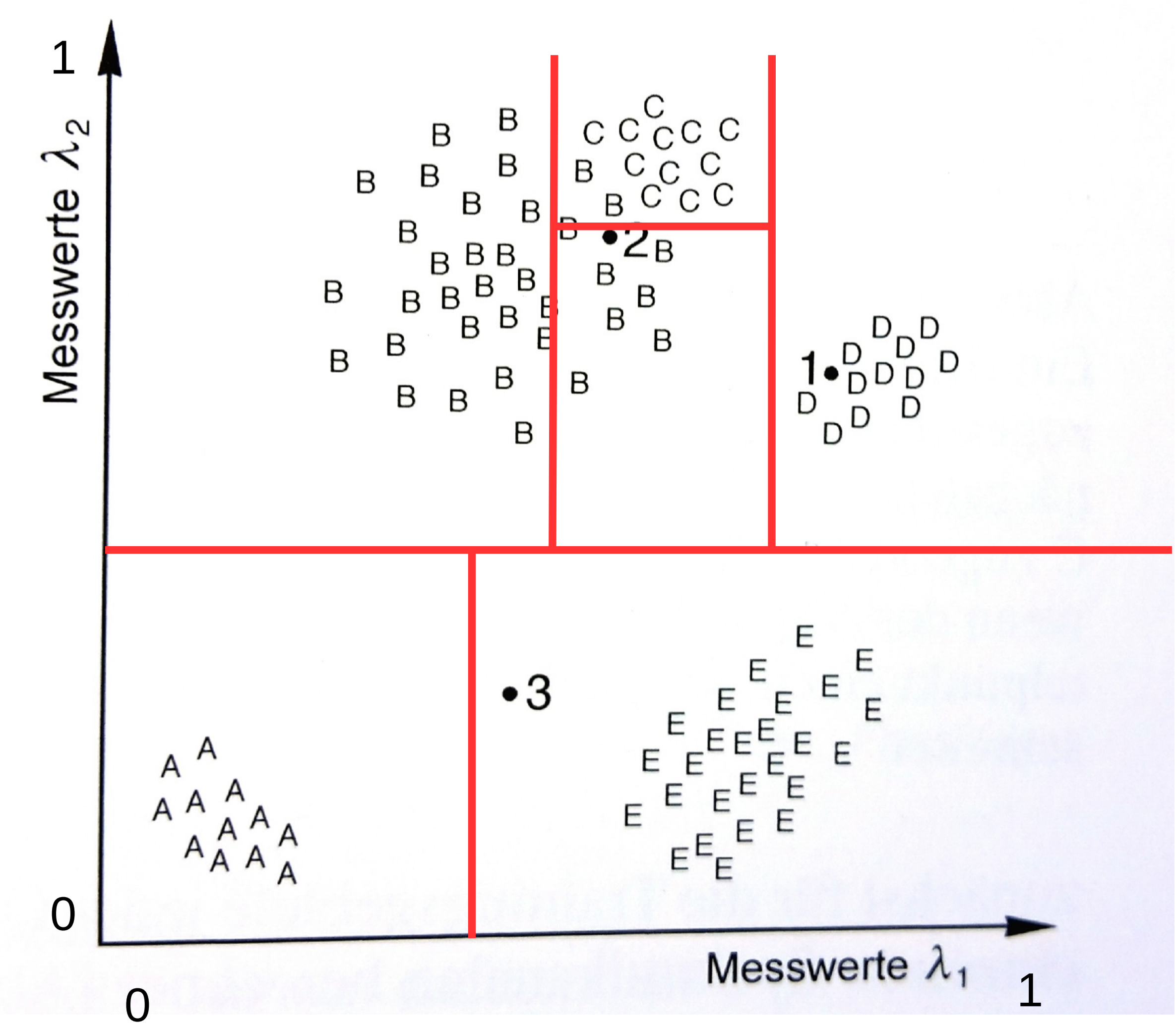
Part 4: Machine learning based land cover classification

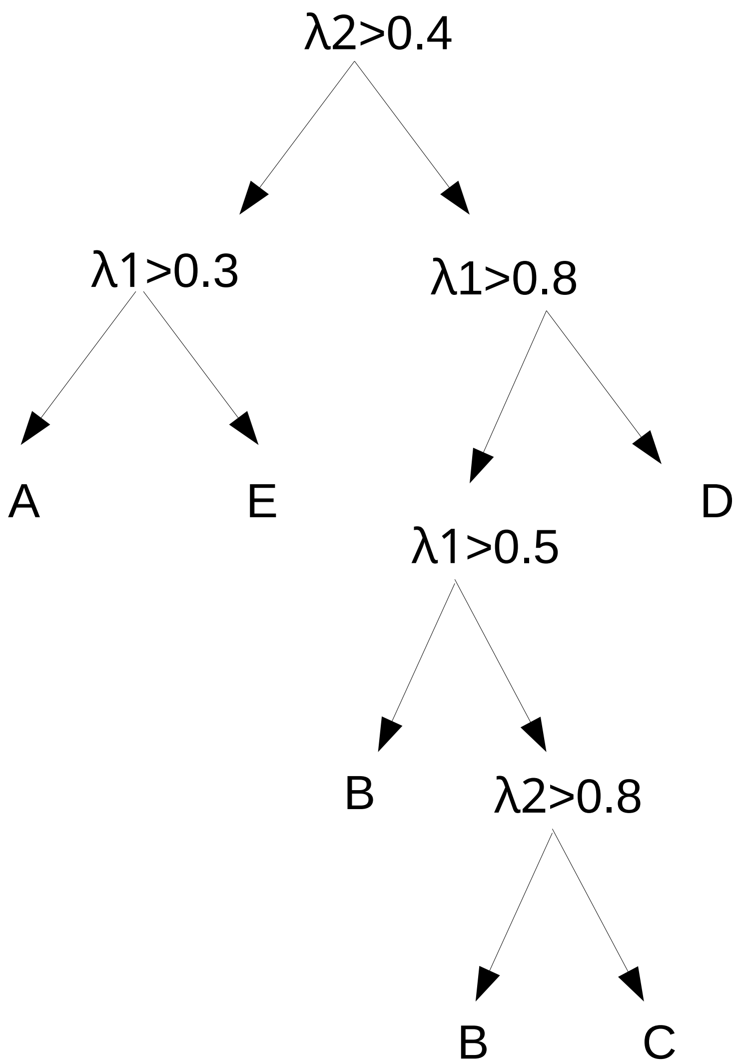


Model training

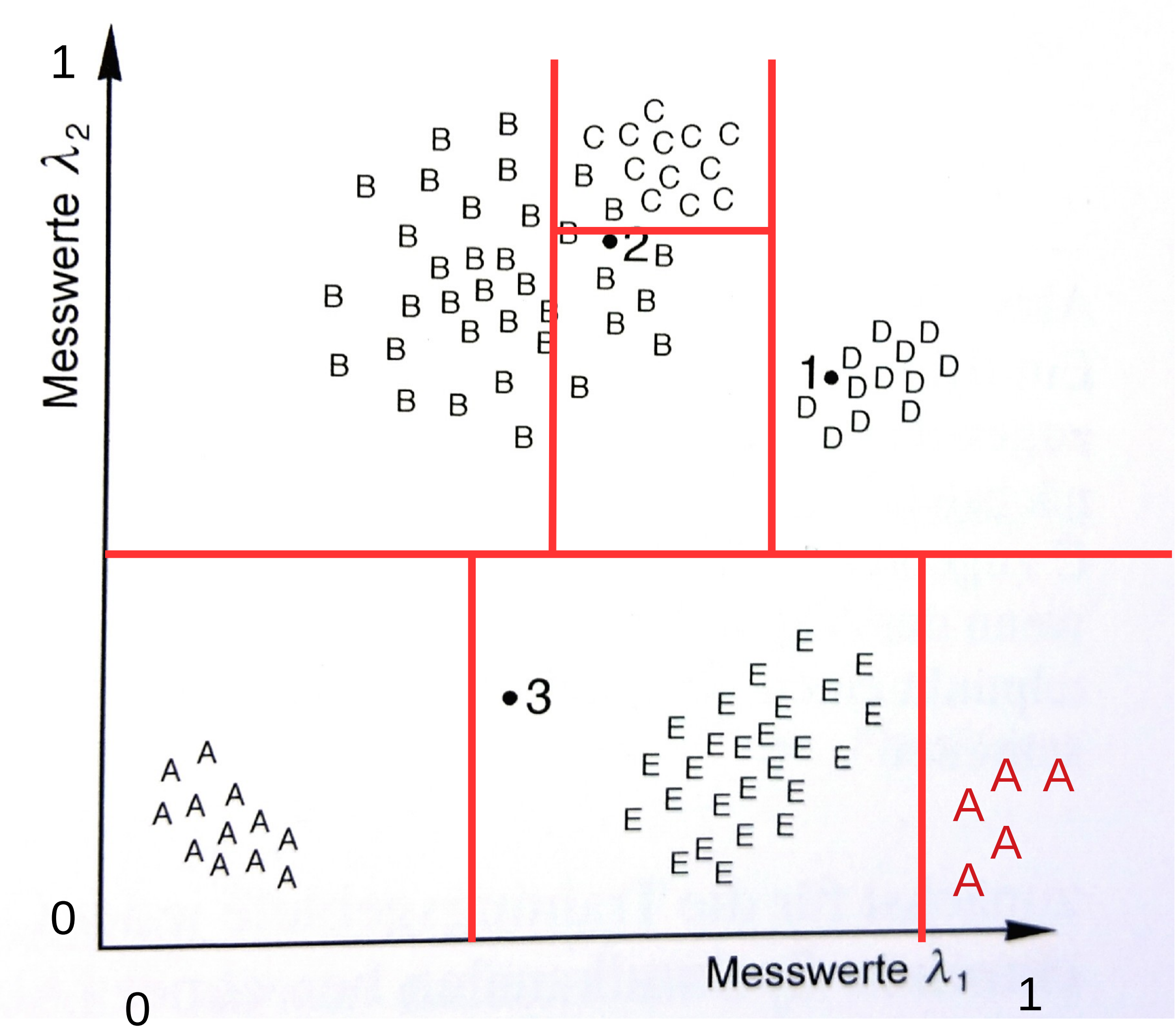


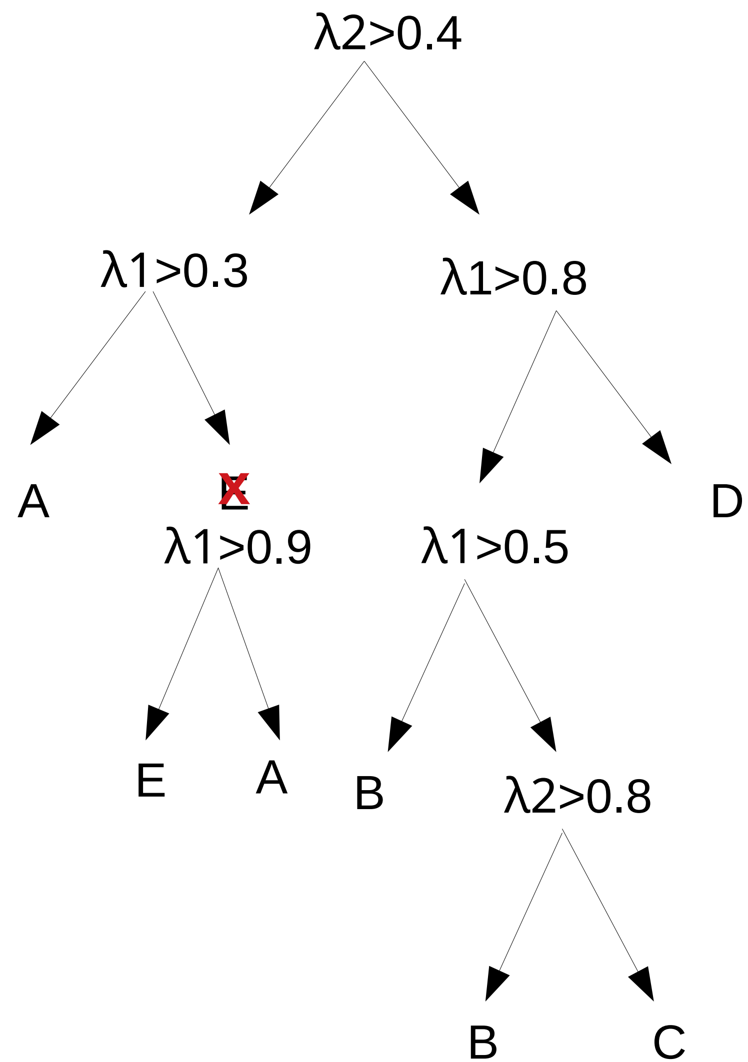
Basics of Random Forests: Classification trees



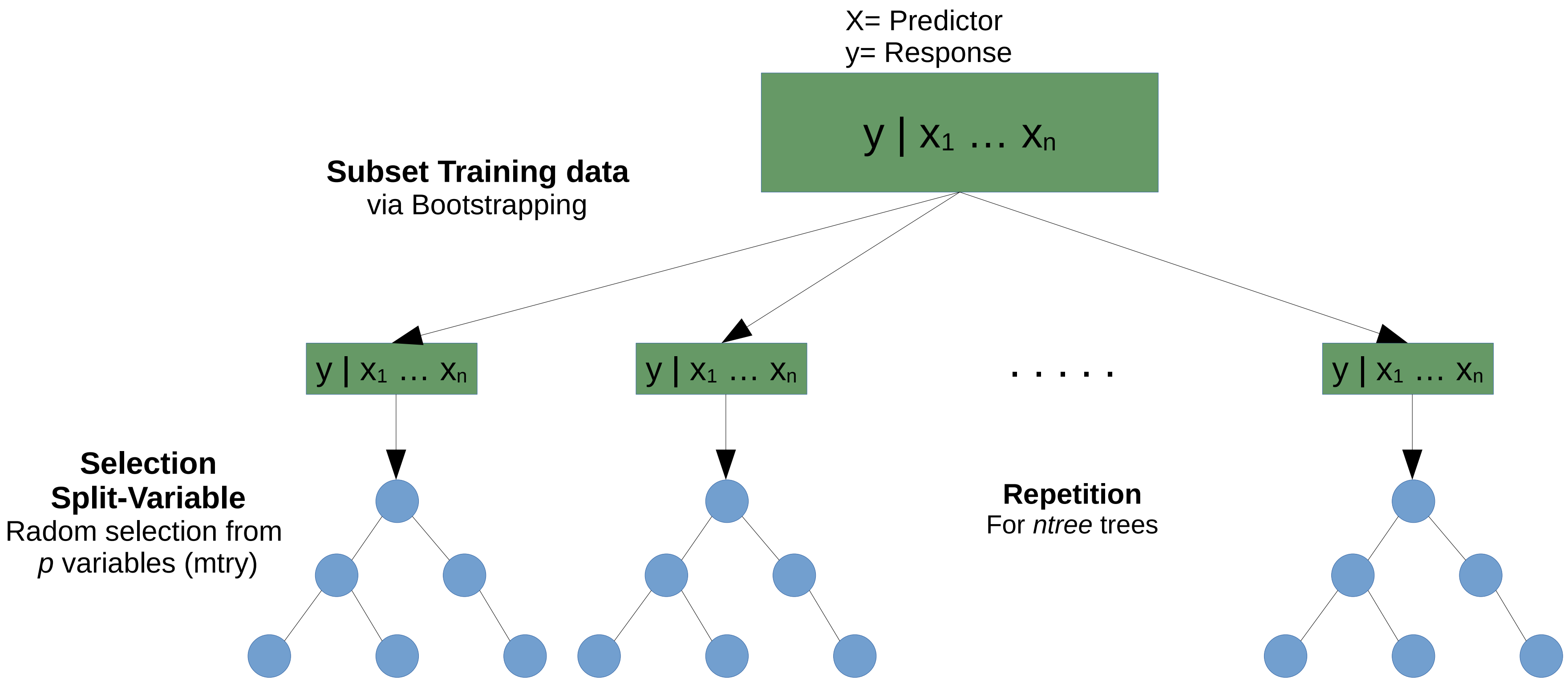


Basics of Random Forests: Classification trees

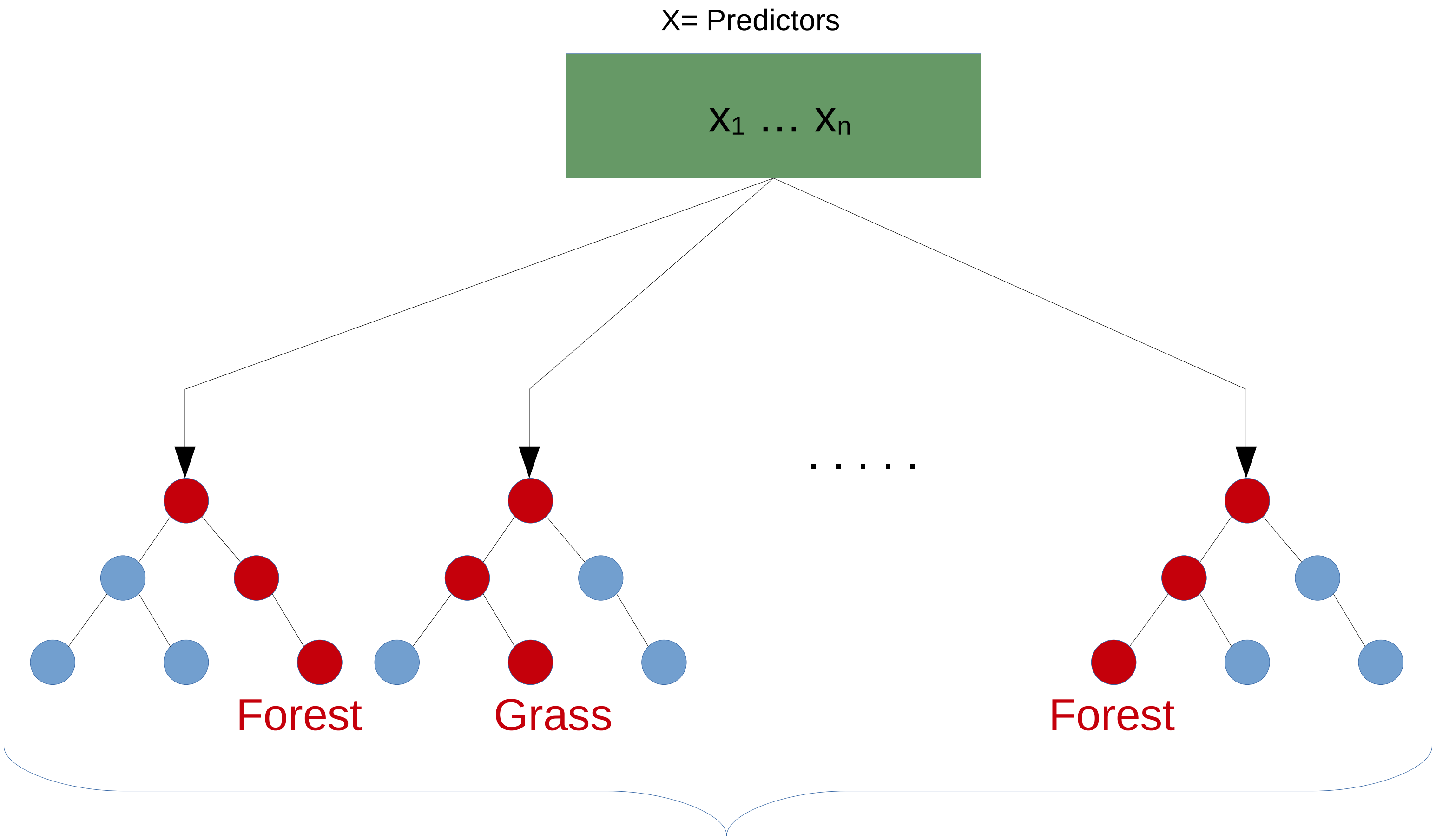




Random Forests: Training



Random Forests: Prediction

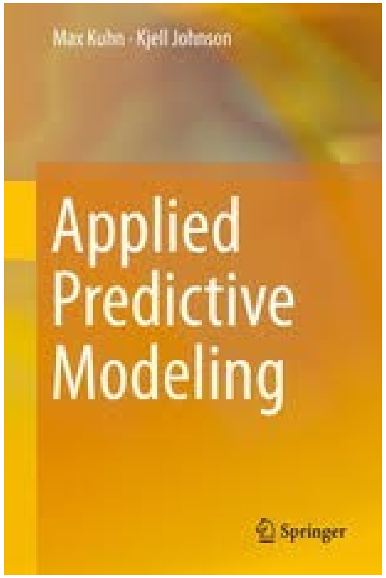


Decision via combination of all trees
(Classification: Majority; Regression: Mean)

Model training in R

- Many packages for different ML algorithms (e.g. Random Forests, Neural Networks, Support Vector Machines, ...)
- For classification and regression problems
- Wrapper packages
  - allowing access to many algorithms via a unified syntax
  - Supporting functionality for cross-validation etc.
  - Caret (Classification And REgression Training)
  - Mlr3 (Machine Learning in R)
  - Tidymodels

For today's session



Step 1: Model training

Predictors Response

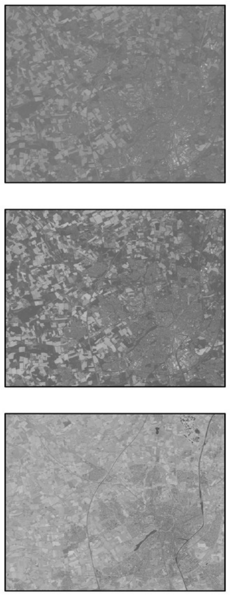
	B02	B03	B04	B08	...	Class
1	857	632	387	308		Water
2	848	633	389	312		Water
3	843	624	357	343		Water
4	854	630	360	333		Water
5	854	628	376	302		Water
6	859	615	364	350		Water

How to do it in R

```
library(caret)
model <- train(predictors,
               response,
               method="rf")
```

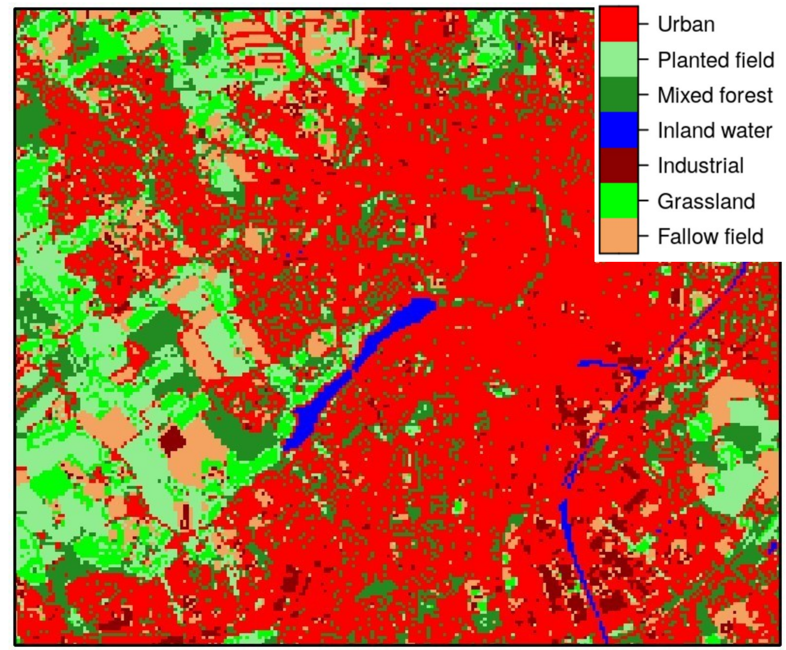
Random Forest used here as Machine learning algorithm

Step 2: Model prediction



...

+ trained model =

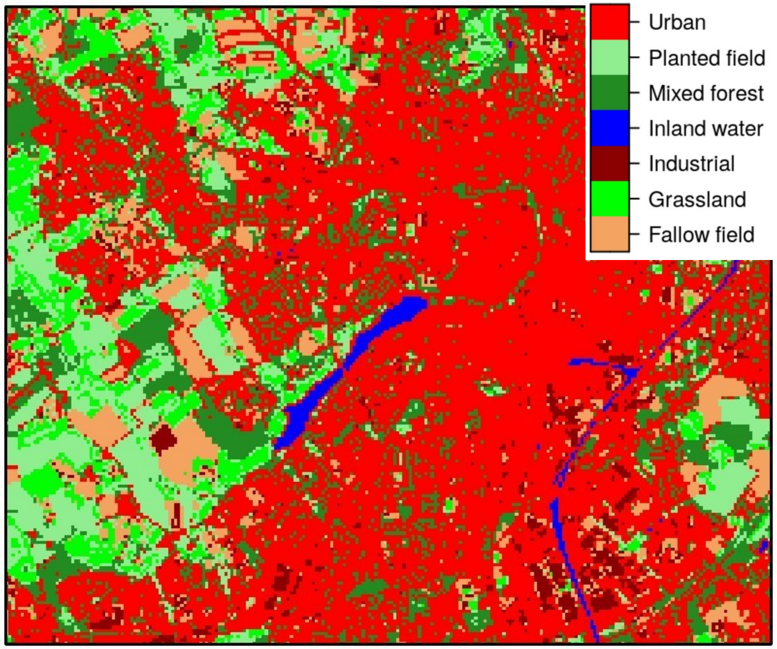


How to do it in R

```
library(terra)
pred_sp <- rast(predictors)
prediction <- predict(pred_sp,model,na.rm=TRUE)
```

Train a model and classify land cover for the entire scene based on the trained model

Map assessment



How good is the map?

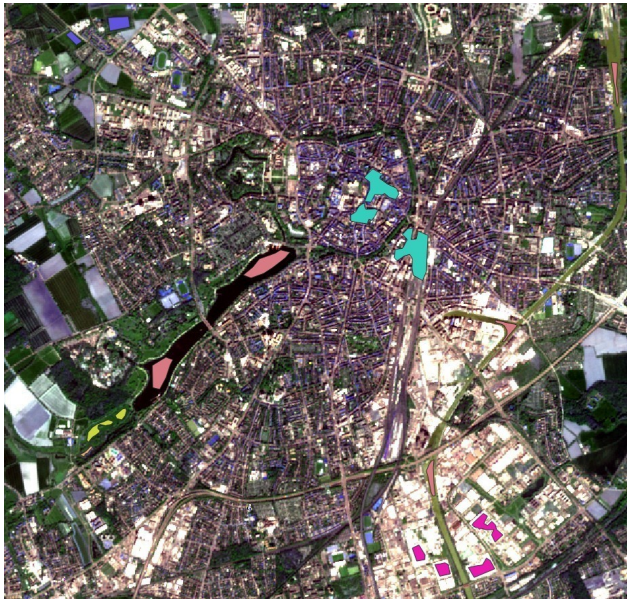
Map assessment – the ideal way: Probability sampling

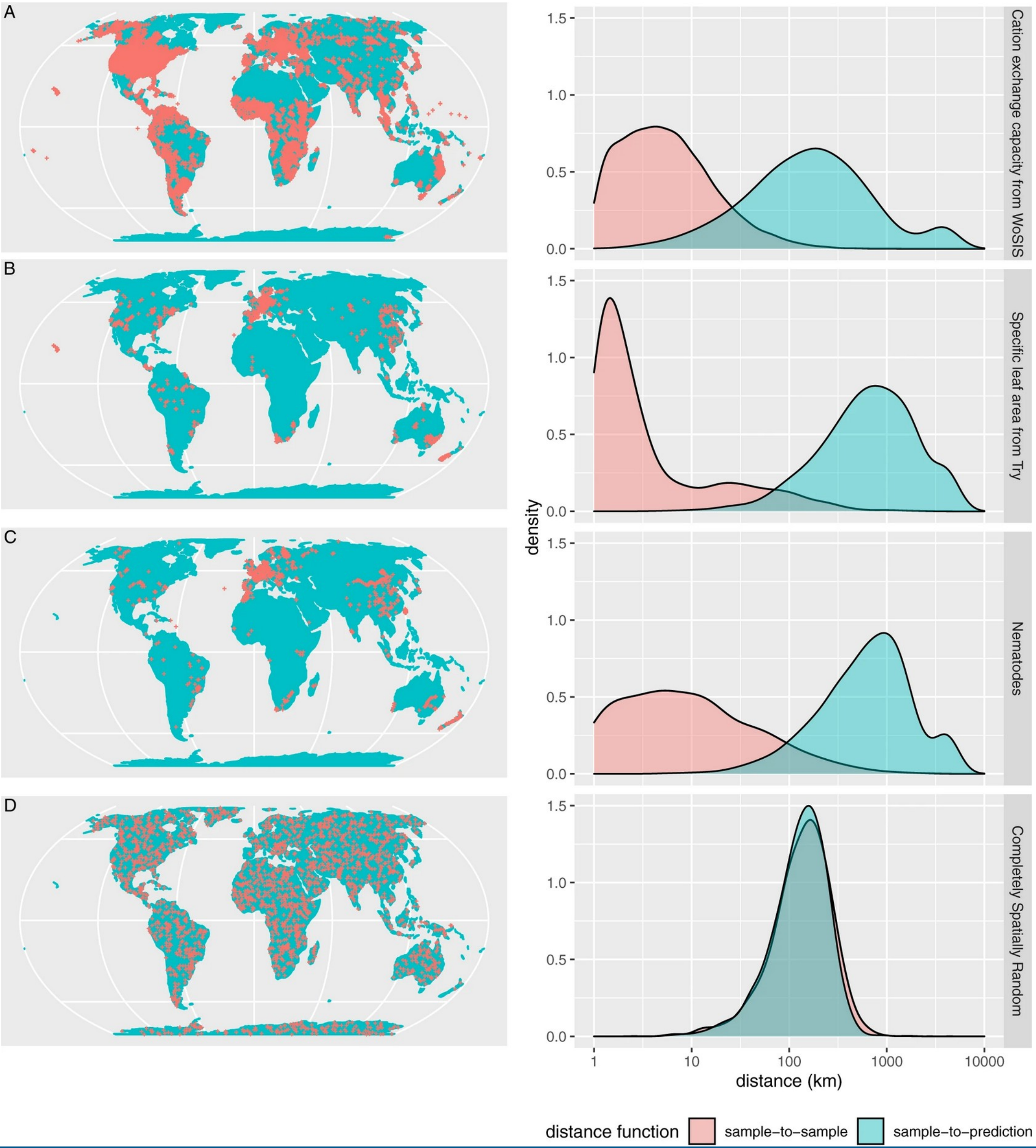
		Reference			
		A	B	C	*Sum*
Classifi cation	A	**300**	120	0	*420*
	B	20	**1500**	80	*1600*
	C	150	5	**20**	*175*
	Sum	*470*	*1625*	*100*	***2195***

Accuracy: (300+1500+20)/2195 = 83%

...but usually we don't have a probability sample

Example of reference data for ecological variables



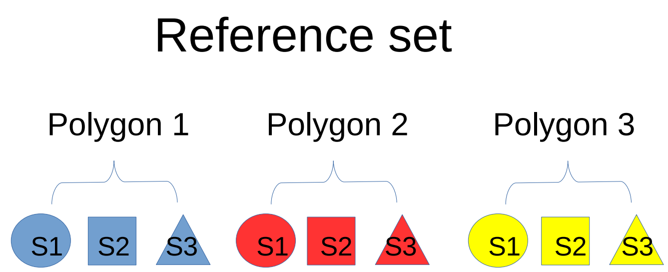


- Reference data are often extremely clustered
- Around areas that received intensive research

In Practice: Data are split into training an test sets. Multiple times: Cross-validation

Cross-validation to assess the model performance

Cross-validation
- Divide data into k folds
- Repeatedly train models
  on k-1 fold
- Test on held back data



Cross-validation to assess the model performance

```
> model
Random Forest

4042 samples
  13 predictor
  11 classes: 'Acker_bepflanzt', 'Fliessgewaesser', 'Gruenland', 'Industriegebiet', 'Laubwald',
 'Mischwald', 'Offenboden', 'See', 'Siedlung', 'Stadt', 'Verkehrswege'

No pre-processing
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 3636, 3636, 3638, 3640, 3638, 3639, ...
Resampling results across tuning parameters:

  mtry  Accuracy   Kappa
   2    0.8963344  0.8740098
   7    0.8960985  0.8737749
  13    0.8923831  0.8693102

Accuracy was used to select the optimal model using the
 largest value.
The final value used for the model was mtry = 2.
```

How to do it in R

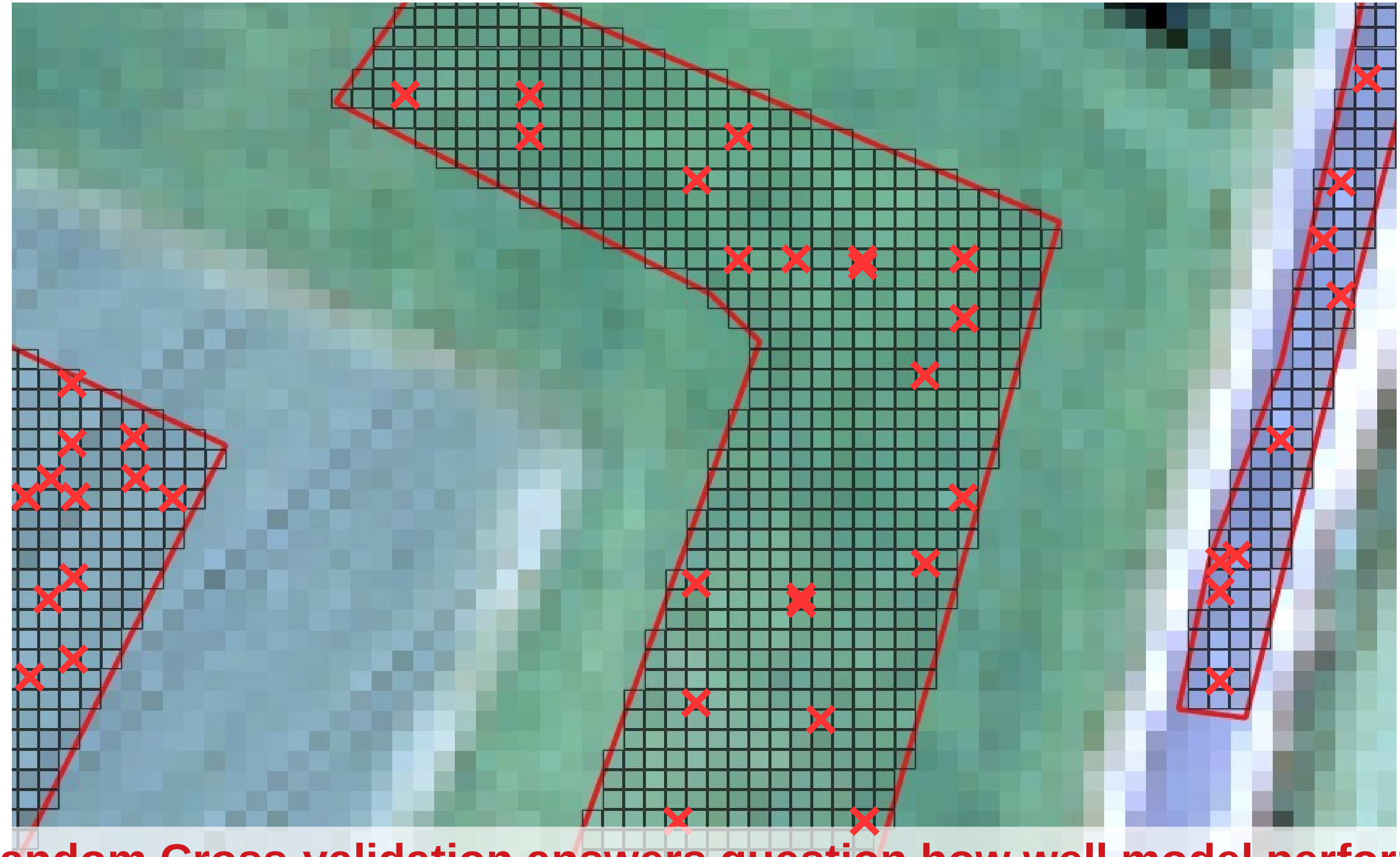
```r
ctrl <- trainControl(method="cv",number=10)

model <- train(trainDat[,predictors],
               trainDat$Label,
               method="rf",
               importance=TRUE,
               ntree=50,
               trControl=ctrl)
```

Typical example of results

Variables	Validation	Accuracy	Kappa
all	random	>0.99	>0.99

Perfect prediction?

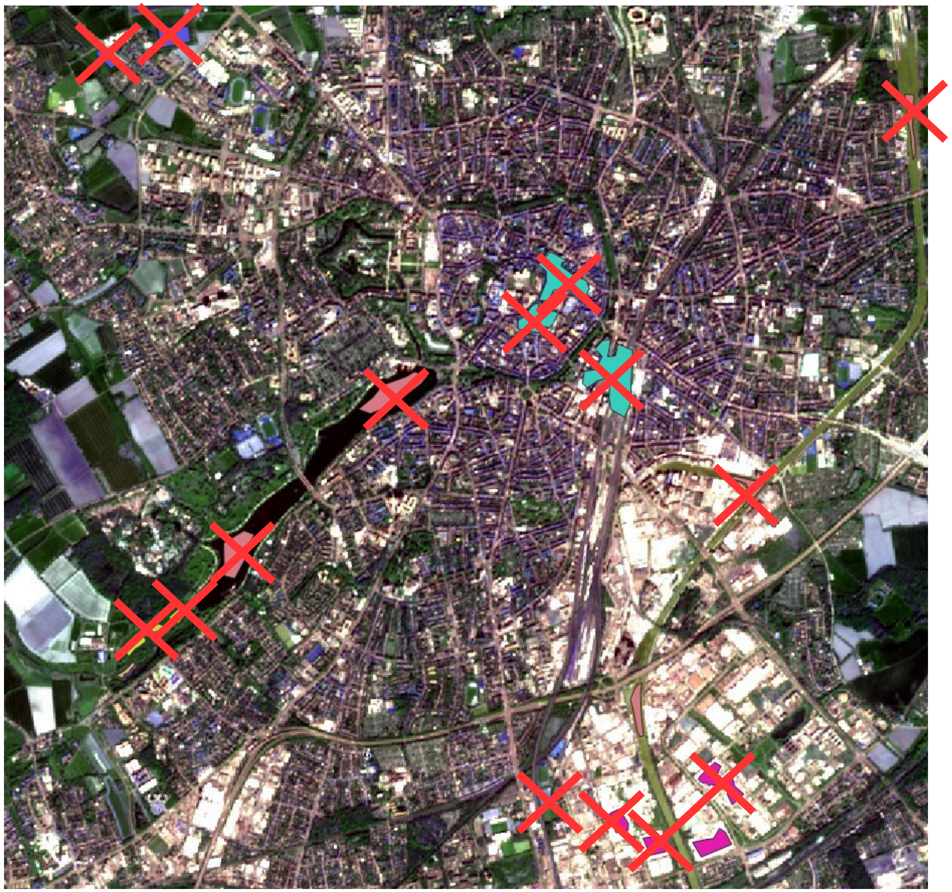
spatial dependencies need to be taken into account



Random Cross-validation answers question how well model performs on very similar locations

Spatial cross-validation (Option: Leave spatial block out)

Spatial cross-validation (Option: Leave group of polygons out)



How to do it in R

```
library(CAST)
```

"**C**aret **A**pplications for **S**tatio-**t**emporal models"

```
indices <- CreateSpacetimeFolds(trainDat,
                                spacevar = "PolygonID",
                                k=3,
                                class="Label")

ctrl <- trainControl(method="cv",
                     index = indices$index)

model <- train(predictors,
               response,
               method="rf",
               trControl=ctrl)
```

Spatial cross-validation: Example of results

Variables	Validation	Accuracy	Kappa
all	random	>0.99	>0.99
all	spatial	**0.68**	**0.61**

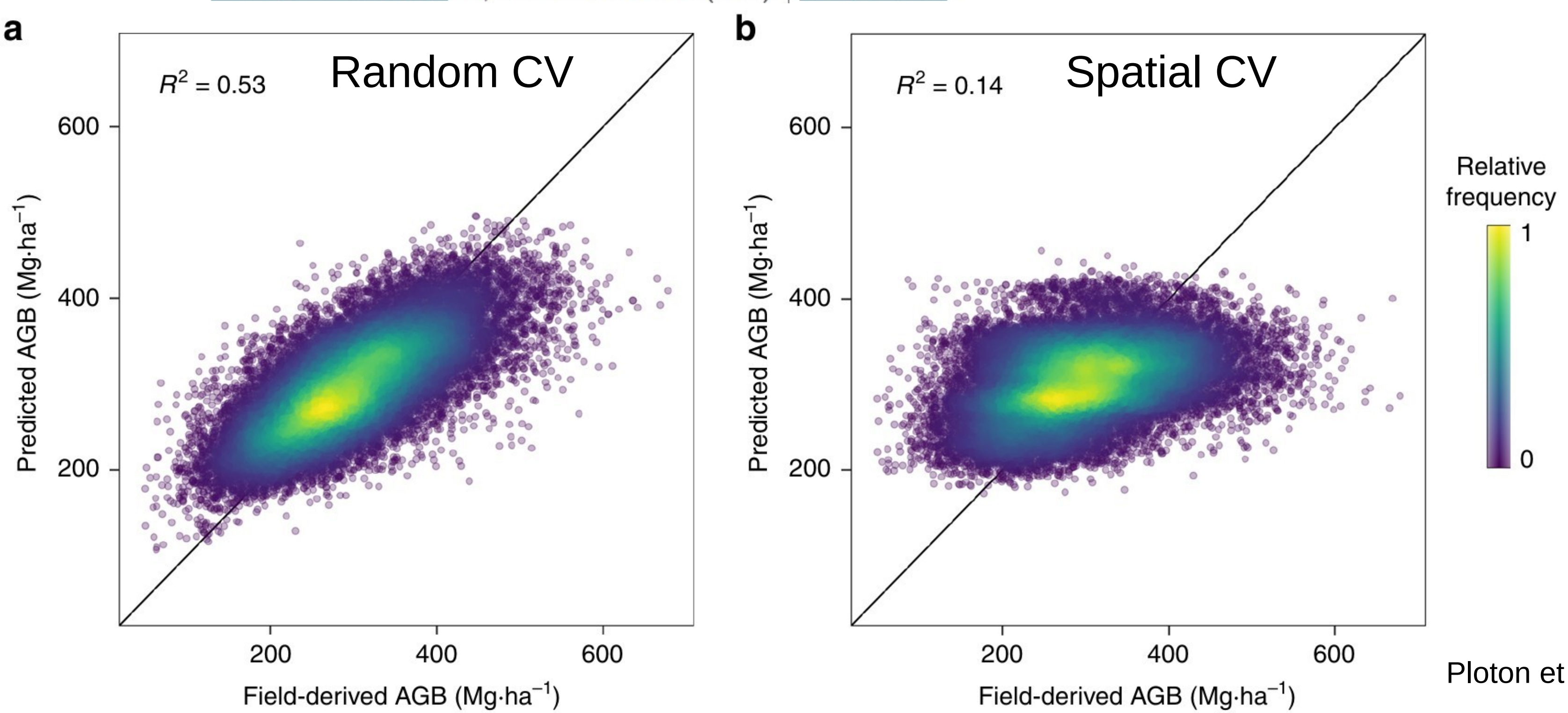
- Random cross-validation performance tells how well we can reproduce training data
- Spatial cross-validation tells us how well we can make predictions
- The spatial performance is usually lower, but this no reason to use the random performance!!!

Spatial cross-validation

Spatial validation reveals poor predictive performance of large-scale ecological mapping models

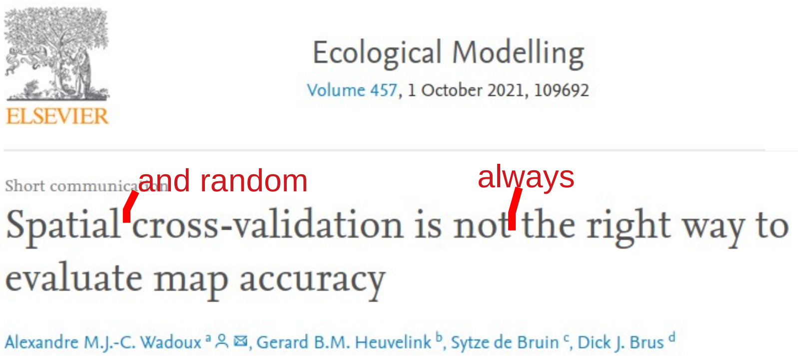
Pierre Ploton, Frédéric Mortier, Maxime Réjou-Méchain, Nicolas Barbier, Nicolas Picard, Vivien Rossi, Carsten Dormann, Guillaume Cornu, Gaëlle Viennois, Nicolas Bayol, Alexei Lyapustin, Sylvie Gourlet-Fleury & Raphaël Pélissier

Nature Communications **11**, Article number: 4540 (2020) | Cite this article



Ploton et al., 2020

...but spatial CV has also been blamed to be too pessimistic. Why ?


Ecological Modelling
Volume 457, 1 October 2021, 109692

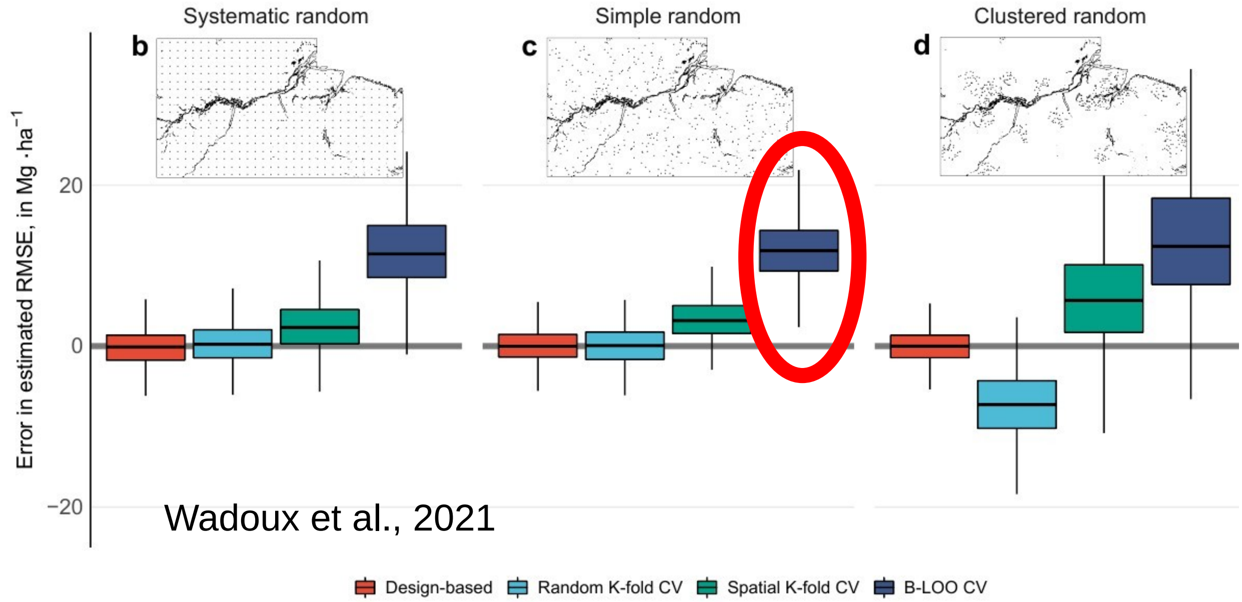
Short communication

and random — always

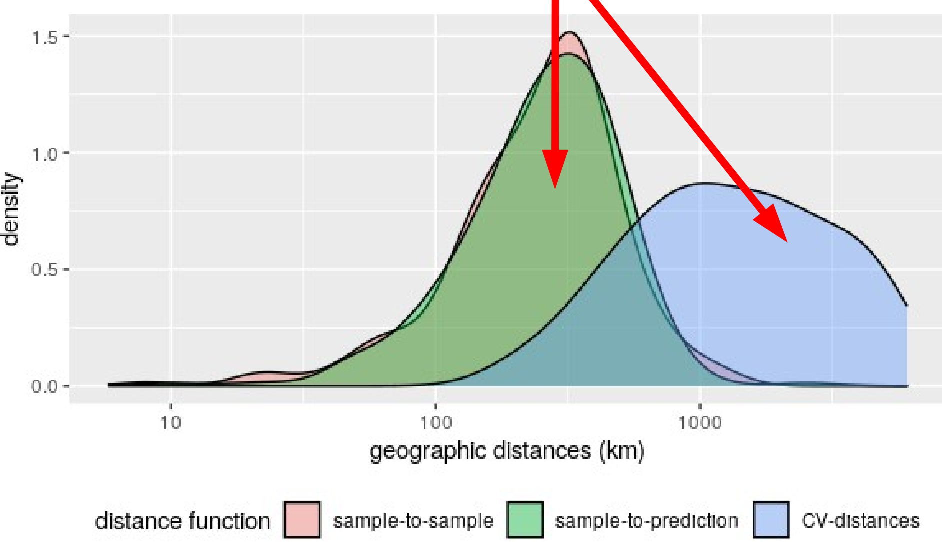
Spatial cross-validation is not the right way to evaluate map accuracy

Alexandre M.J.-C. Wadoux [a], Gerard B.M. Heuvelink [b], Sytze de Bruin [c], Dick J. Brus [d]

CV predictions are harder than the actual prediction task



Wadoux et al., 2021



It’s obviously NOT the right way here. Why?

Prediction situations created during CV need to resemble those encountered while predicting the global map from the reference data

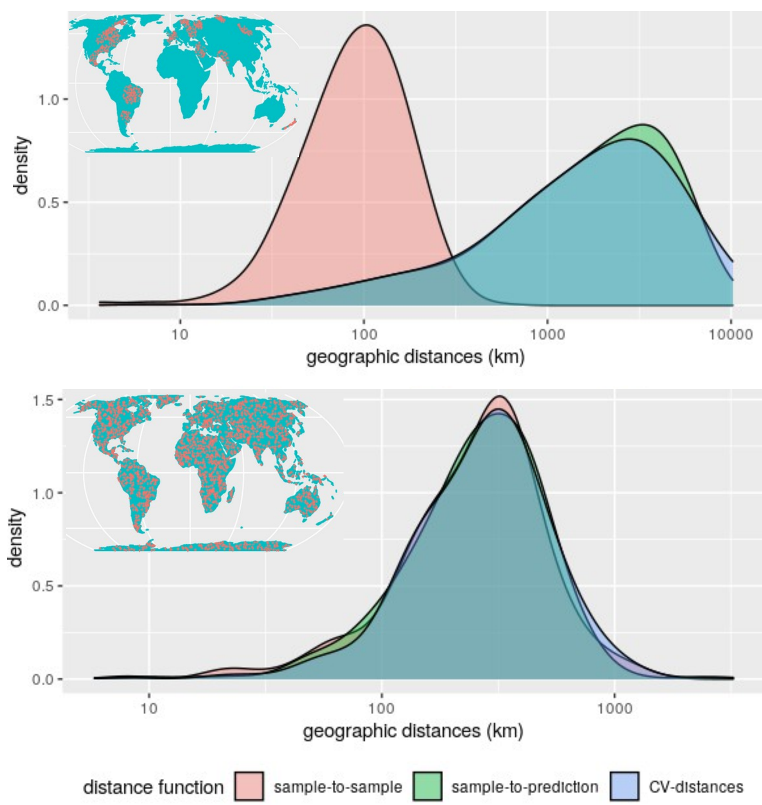
Suggestion of a nearest neighbor distance matching?

Received: 20 September 2021 | Accepted: 8 March 2022
DOI: 10.1111/2041-210X.13851
Methods in Ecology and Evolution
BRITISH ECOLOGICAL SOCIETY

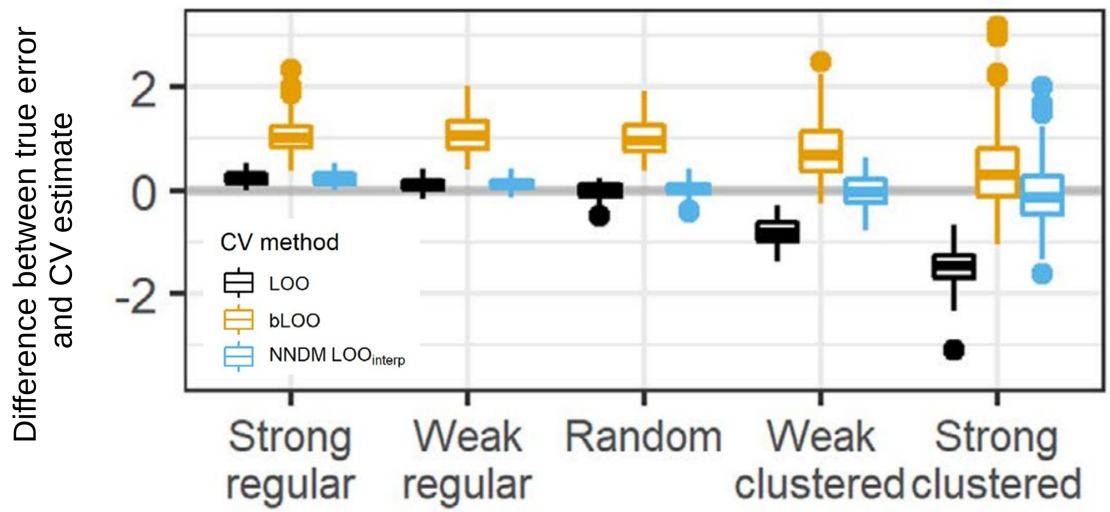
RESEARCH ARTICLE

Nearest neighbour distance matching Leave-One-Out Cross-Validation for map validation

Carles Milà[1] | Jorge Mateu[2] | Edzer Pebesma[3] | Hanna Meyer[4]



Aim: Prediction situations created during CV resemble those encountered while predicting the global map



Mila et al., 2022

Cross-validation

Train your model again and compare the validation statistics between a random and a spatial (leave polygon out?) CV

?

Cross validation for model tuning

Spatial cross-validation should be used for the entire modelling process

- For variable selection
- Hyperparameter tuning
- Error assessment

Usually improves spatial performance!

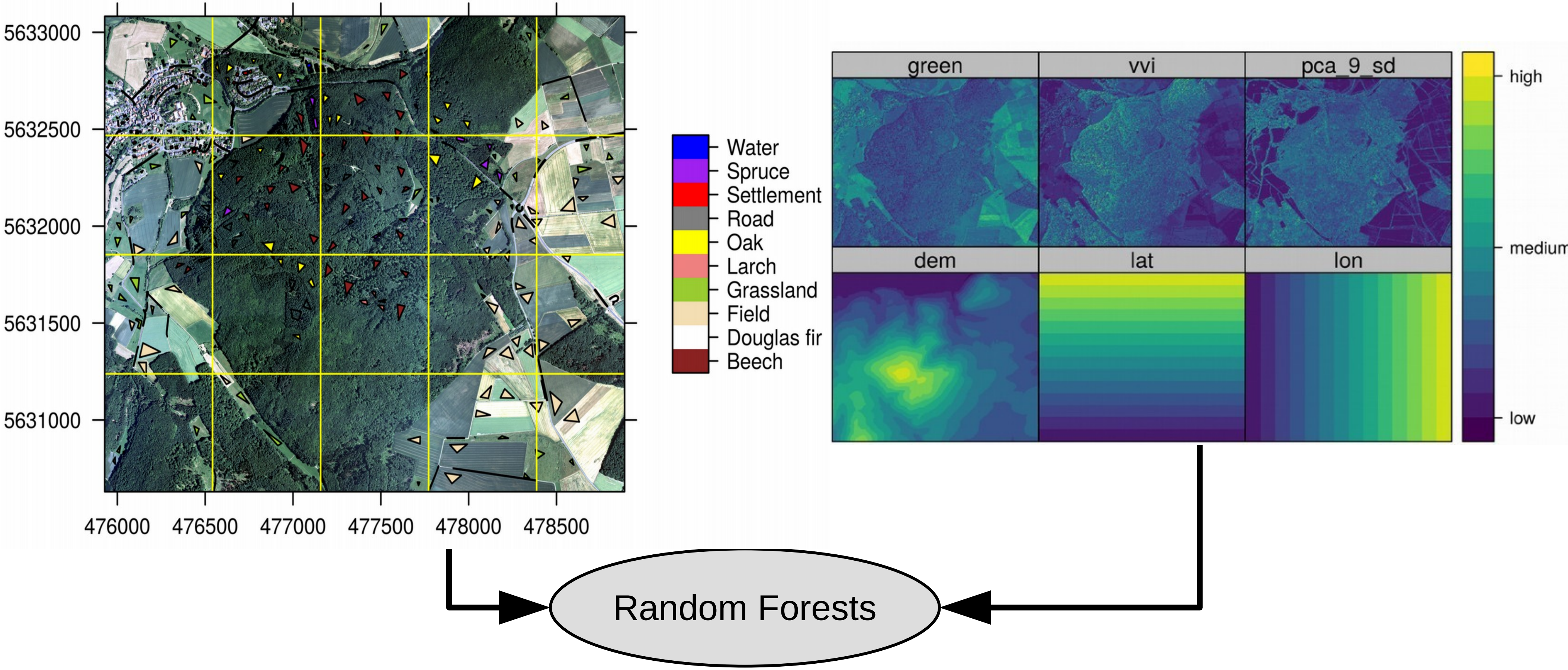
Further reading:

- **New Nearest neaighbor distance matching CV**: Mila C, Mateu J, Pebesma E, Meyer H. 2022. ‘Nearest neighbour distance matching leave-one-out cross-validation for map validation.’ Methods in Ecology and Evolution. doi: 10.1111/2041-210X.13851.
- **Spatial CV for variable selection**: Meyer H, Reudenbach C, Wöllauer S, Nauss T. 2019. ‘Importance of spatial predictor variable selection in machine learning applications – Moving from data reproduction to spatial prediction.’ Ecological Modelling 411: 108815. doi: 10.1016/j.ecolmodel.2019.108815.
- **Spatial CV for hyperparameter tuning**: Schratz P, Muenchow J, Iturritxa E, Richter J, Brenning A. 2019. ‘Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. ’ Ecological Modelling 406, 109-120. https://doi.org/10.1016/j.ecolmodel.2019.06.002.

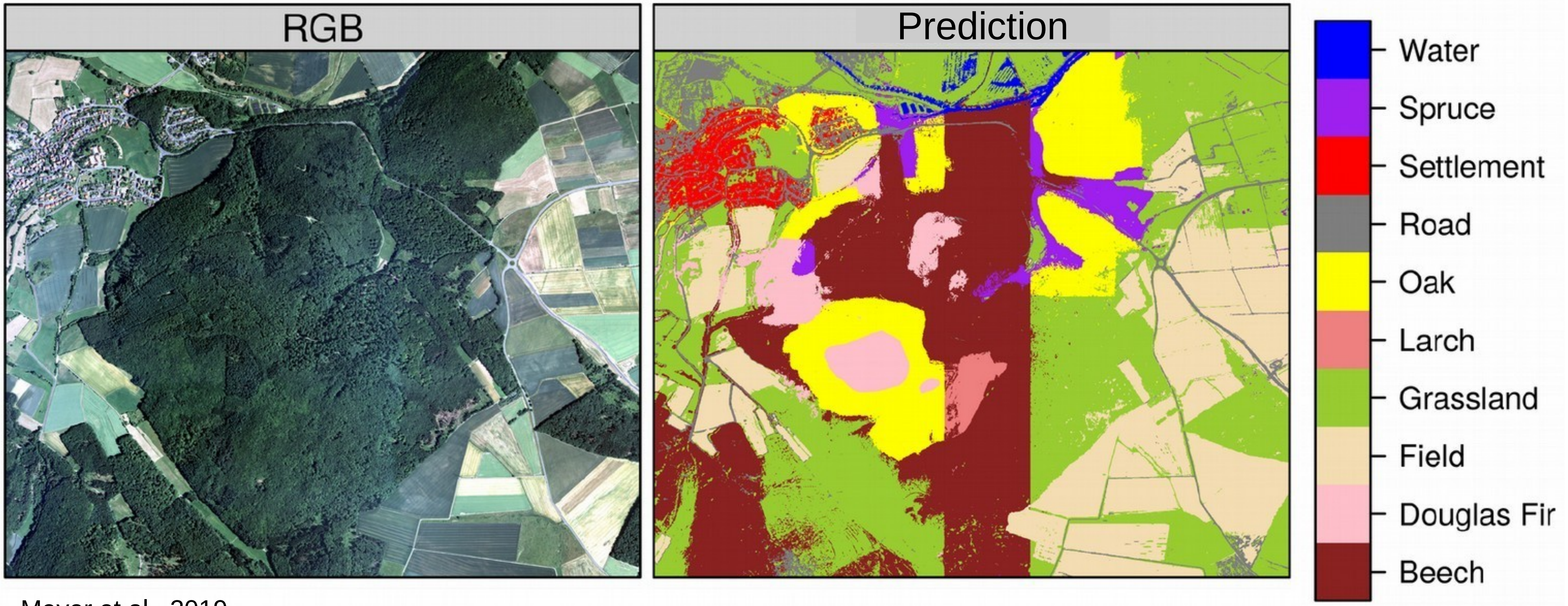
Spatial variable selection – Why ?

Aerial image overlayed by training sites

Example of predictors

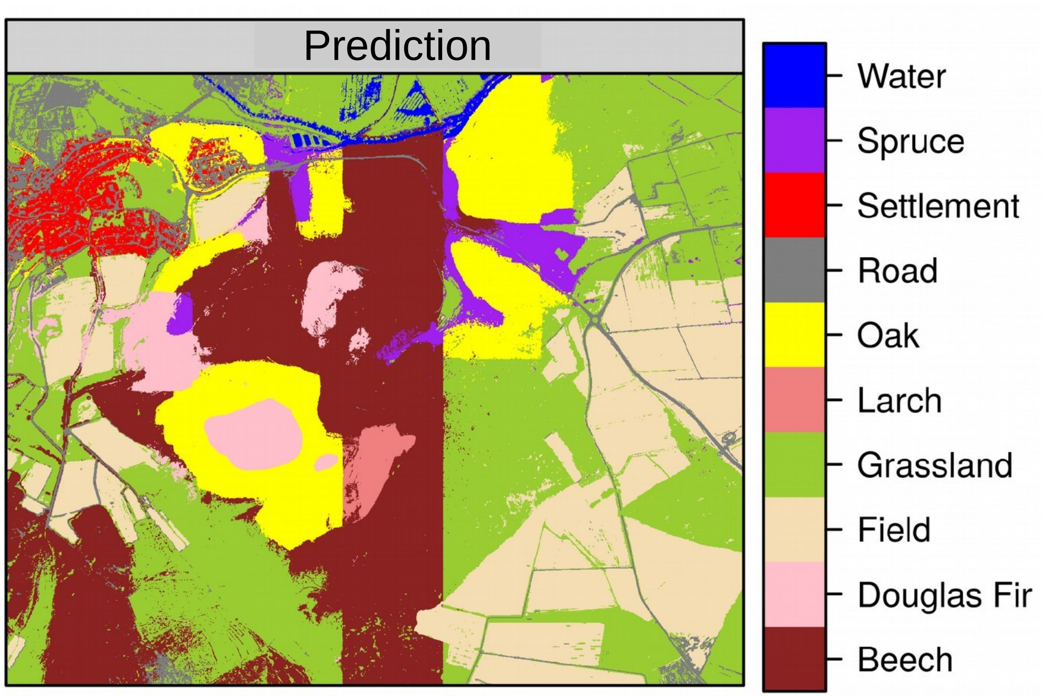


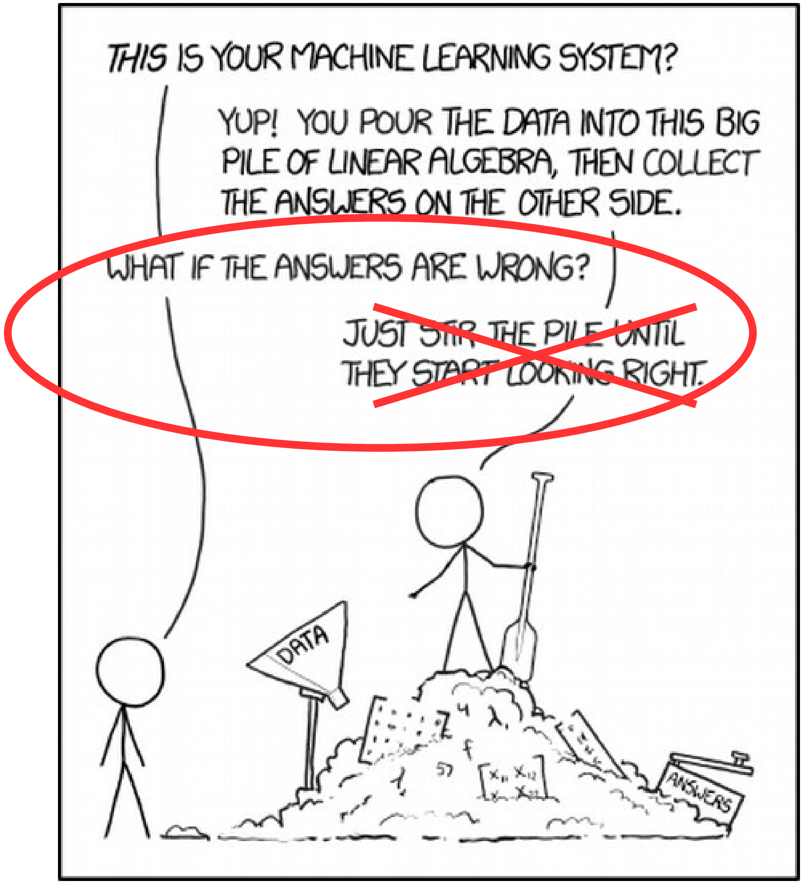
Spatial variable selection – Why ?



Meyer et al., 2019

Spatial performance of models needs to be improved!



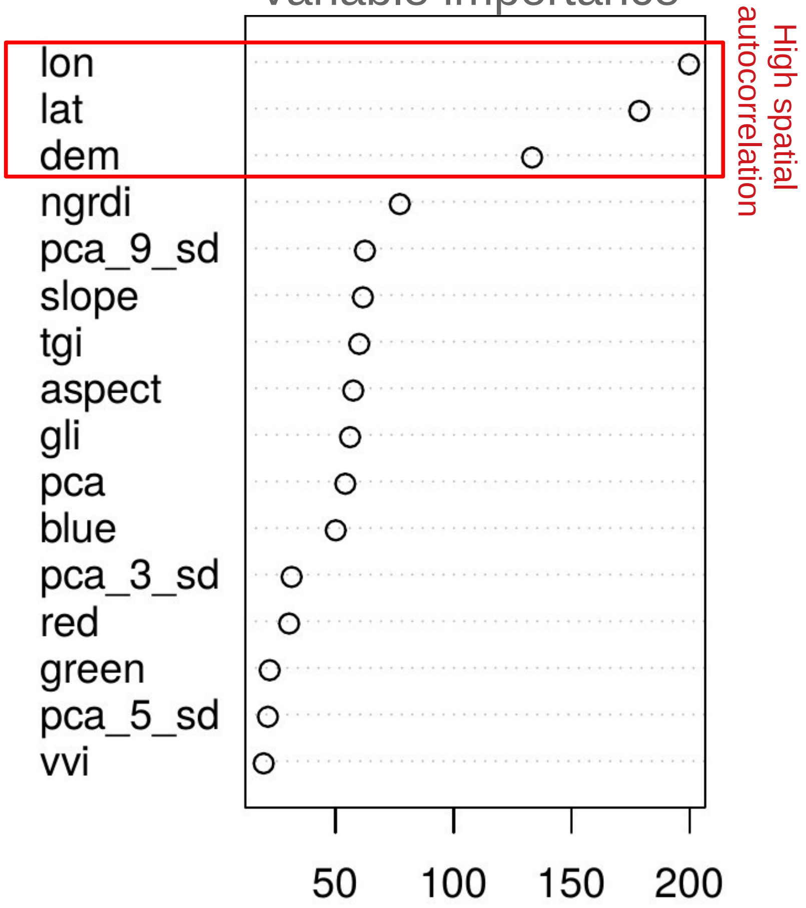


Where do these prediction patterns come from?

https://xkcd.com/1838/

An example of the “clever Hans effect” ?

Variable importance



Is the model behaving like the “clever Hans” ?

https://commons.wikimedia.org/wiki/File:Osten_und_Hans.jpg#/media/Datei:Osten_und_Hans.jpg

→ The models are not able to learn the scientifically correct relationships
→ not transferable

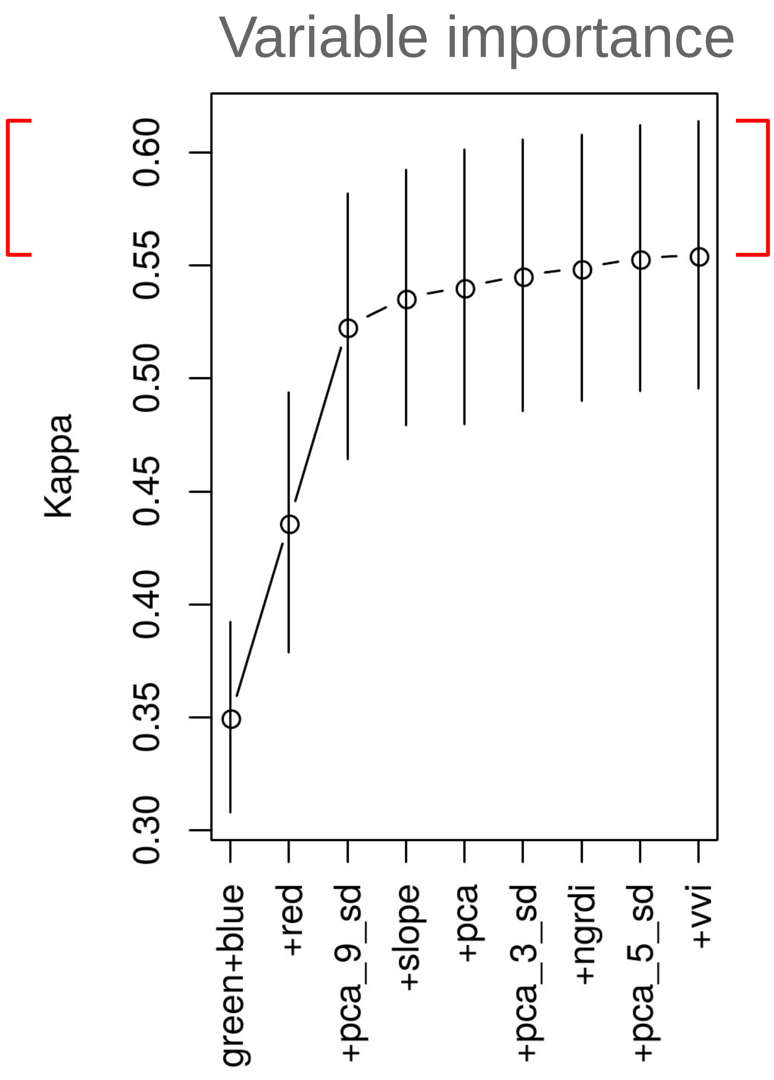
“Unmasking Clever Hans predictors and assessing what machines really learn”
(Lapuschkin et al., 2019, Nature communications)

Horse-picture from Pascal VOC data set

What is the information the algorithm uses to detect the horse?

Lapuschkin et al., 2019

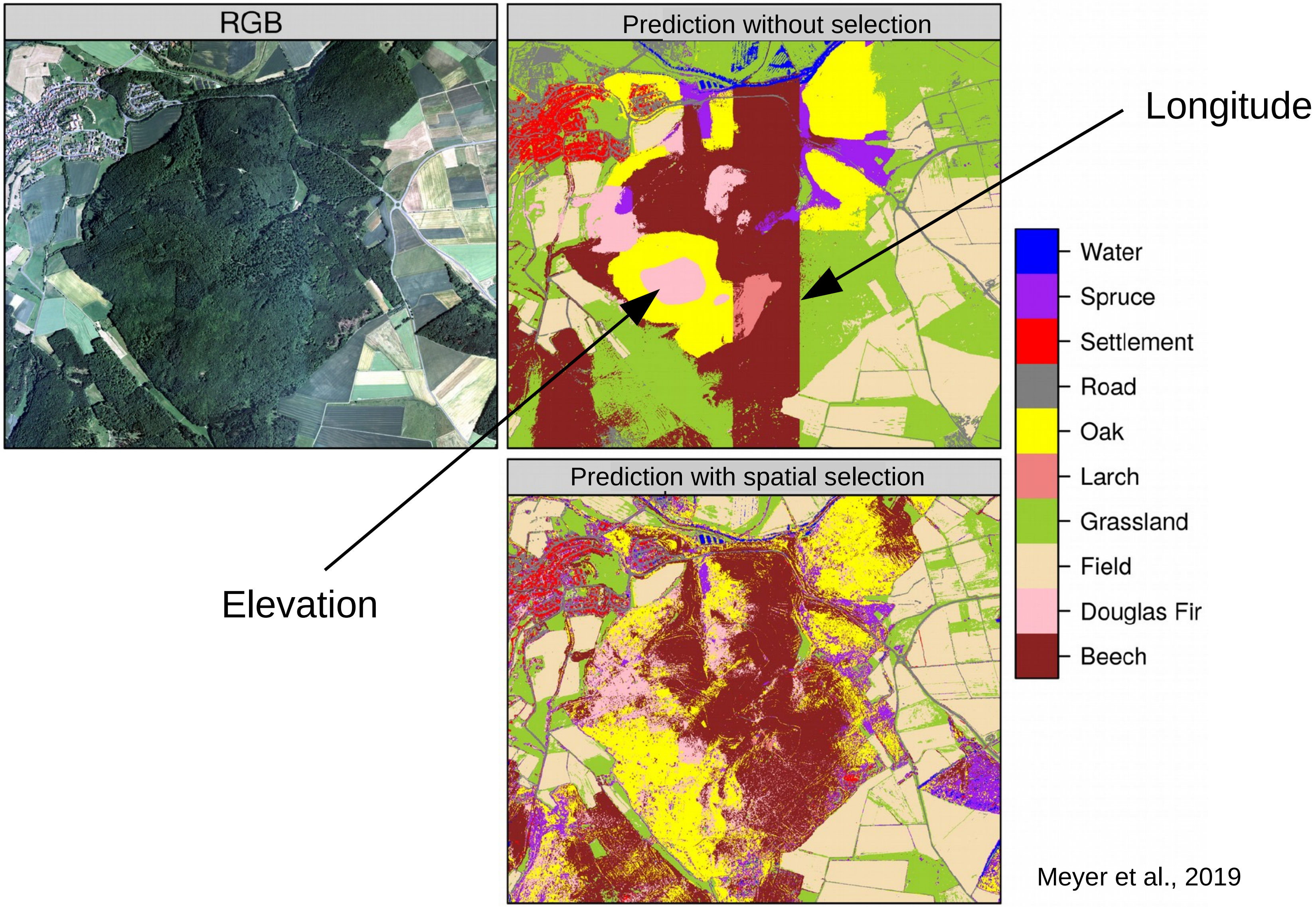
Unmasking "clever Hans predictors" to improve the model?



- Assumption: spatial autocorrelation leads to "clever Hans predictors"
- Removing those variables should improve the results
- Spatial variable selection required!

Implemented in R package "CAST"

Unmasking “clever Hans predictors” to improve the model?



Meyer et al., 2019

Use spatial variable selection to improve your model

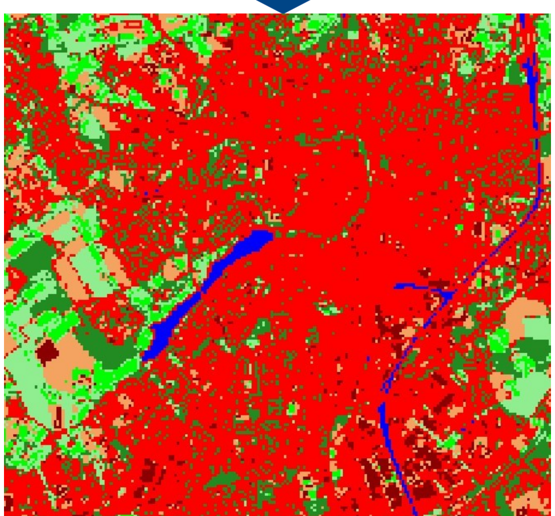
Re-train your model by using spatial variable selection (?CAST::ffs)

Now we have...

- A model trained to make spatial predictions
- Most reliable error estimates for the predictions

Is this sufficient for reliable mapping?

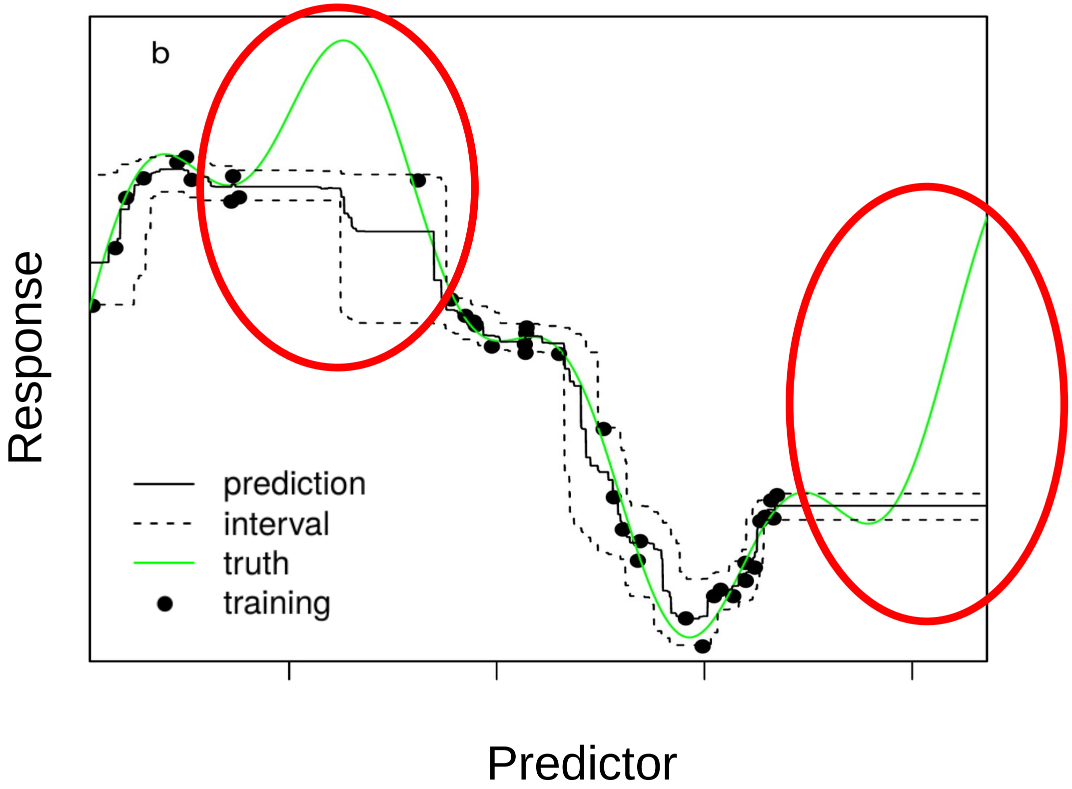
Estimating the area of applicability



Can we really apply the model to the entire area?

- Transfer to new space required
- New space might differ in spectral properties
- But what if the algorithm has never seen such properties?

Estimating the area of applicability

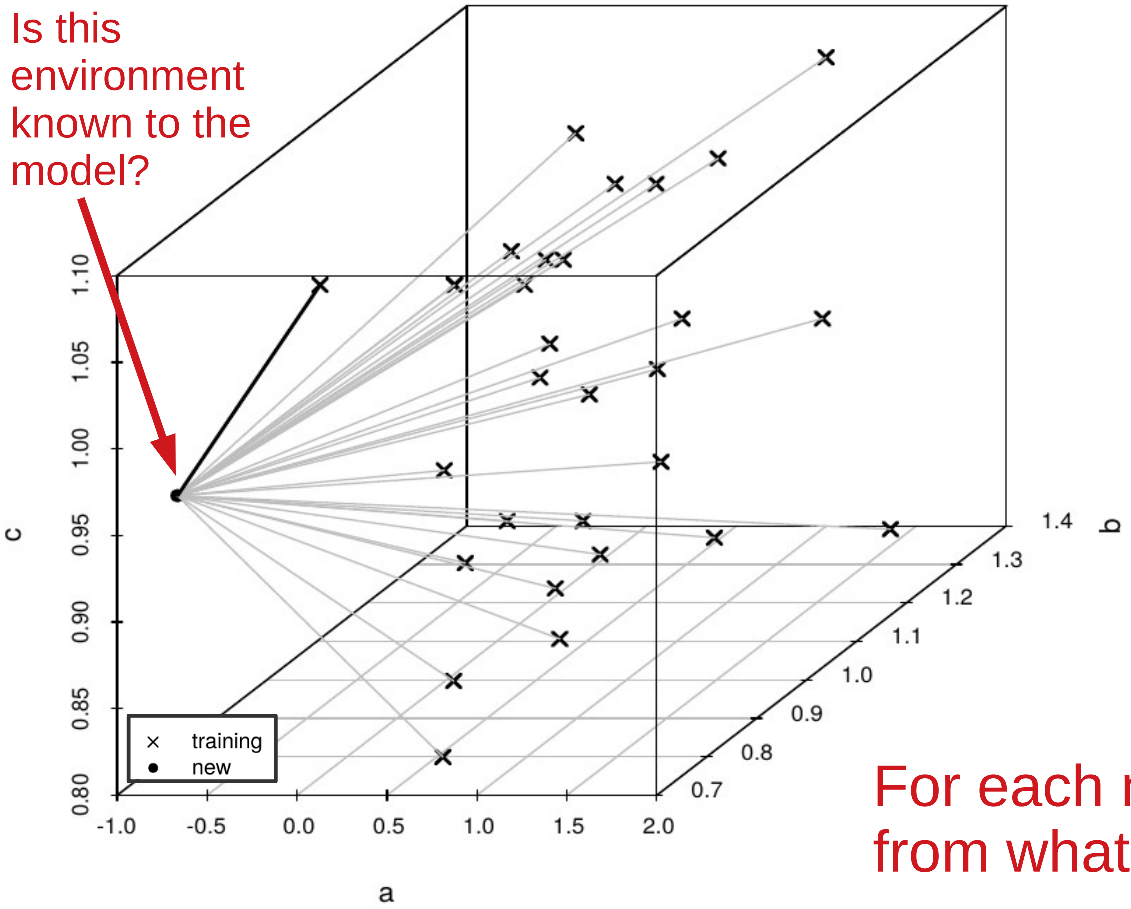


Meyer & Pebesma (2021)

- Machine learning can fit very complex relationships.
- But gaps in predictor space are problematic (the model has no knowledge about these areas!)
- A measure for “unknown space” is needed to estimate the area of applicability of a model!

Estimating the area of applicability

Is this environment known to the model?



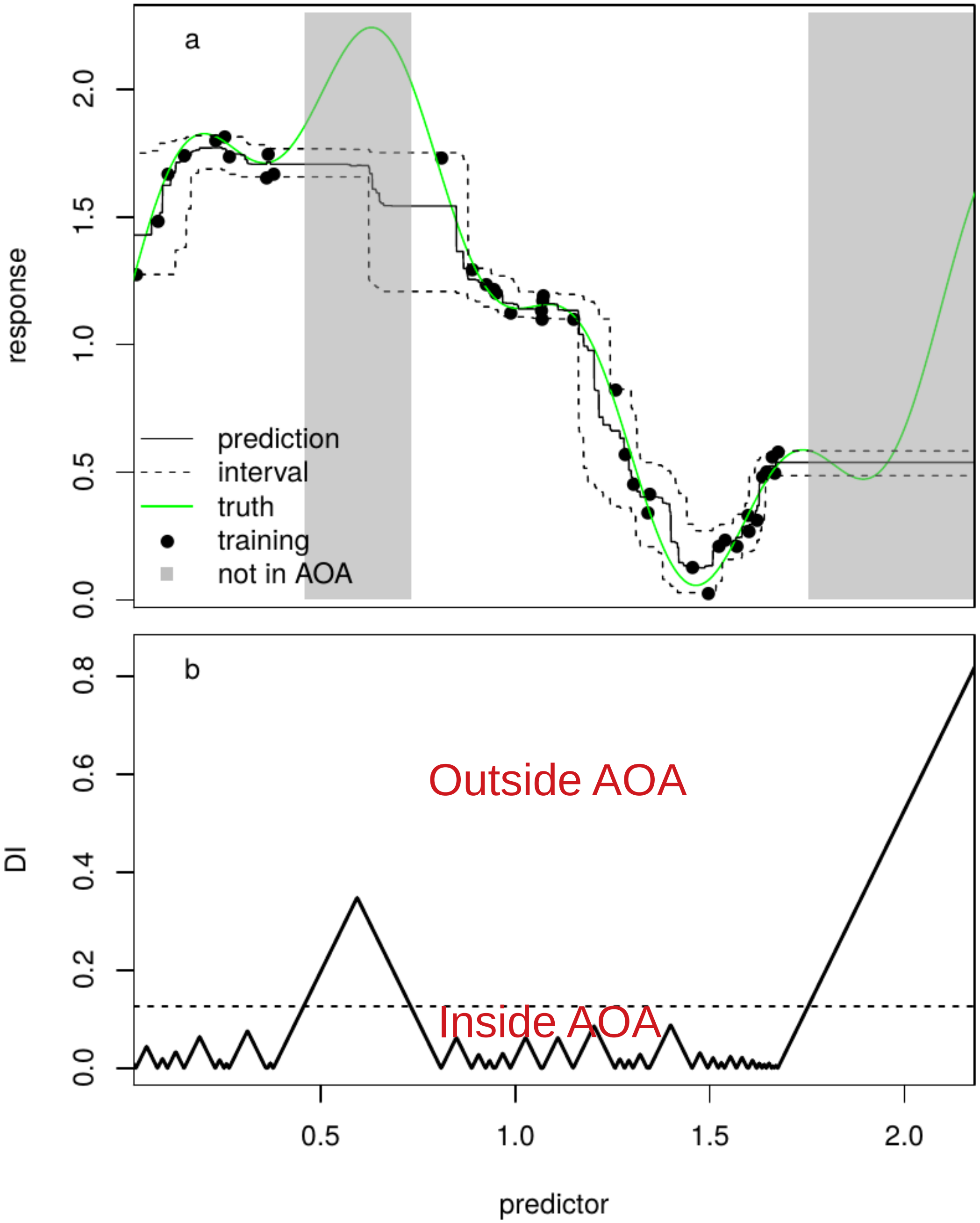
- Unknown space: Environmental conditions that are very different from the training locations
- Suggestion: Distance in (weighted) predictor space as measure for unknown space (Dissimilarity Index)

For each new location/pixel: how distant is it from what the algorithms has seen?

Estimating the area of applicability

- Definition: Area for which we enabled the model to learn about relationships; The area for which, on average, the cross-validation error of the model applies

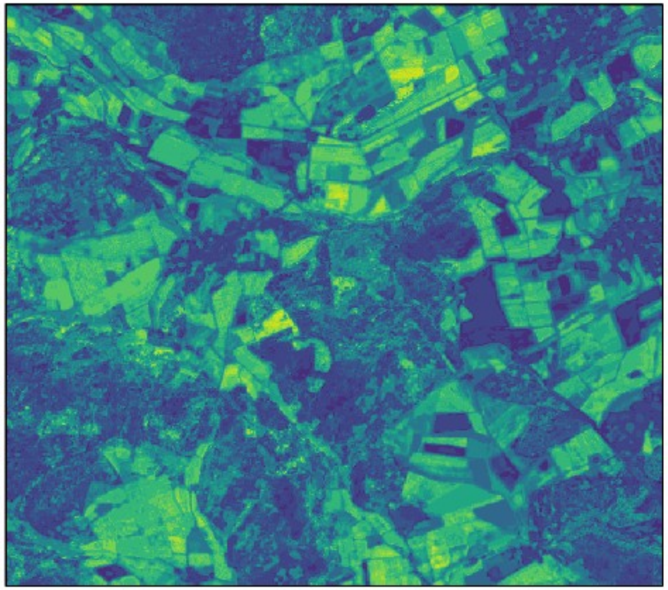
- Estimated using a threshold on the DI
- Threshold = maximum DI of cross-validated training data
- DI < threshold = inside AOA
  DI > threshold = outside AOA



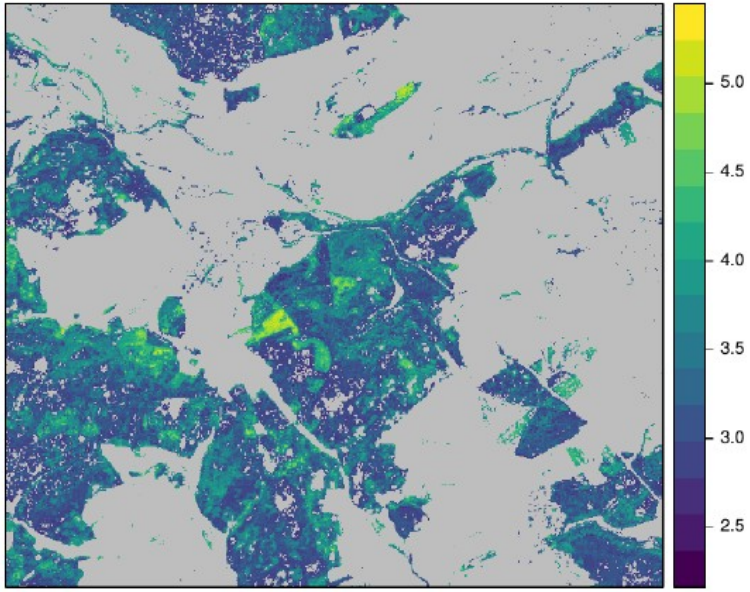
A very obvious and simple example

Sentinel-2 scene and training data points of leaf area index

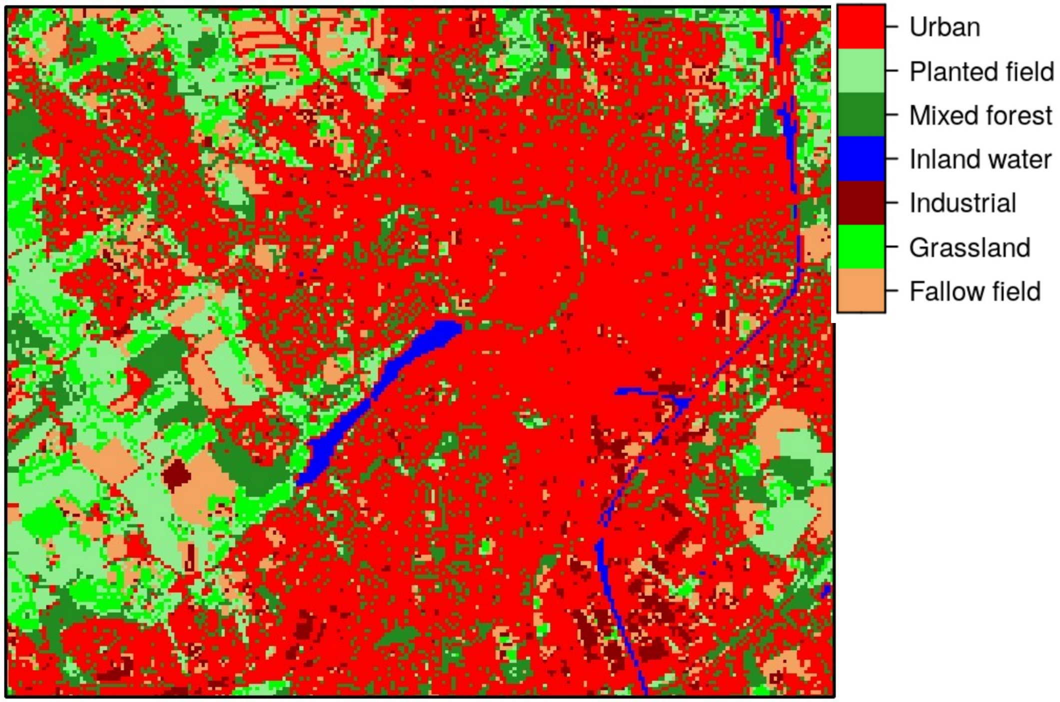
Predictions



Predictions limited to the AOA



Estimating the area of applicability



Technically we can make predictions for the entire area. But they only make sense if the predictor properties are known to the model

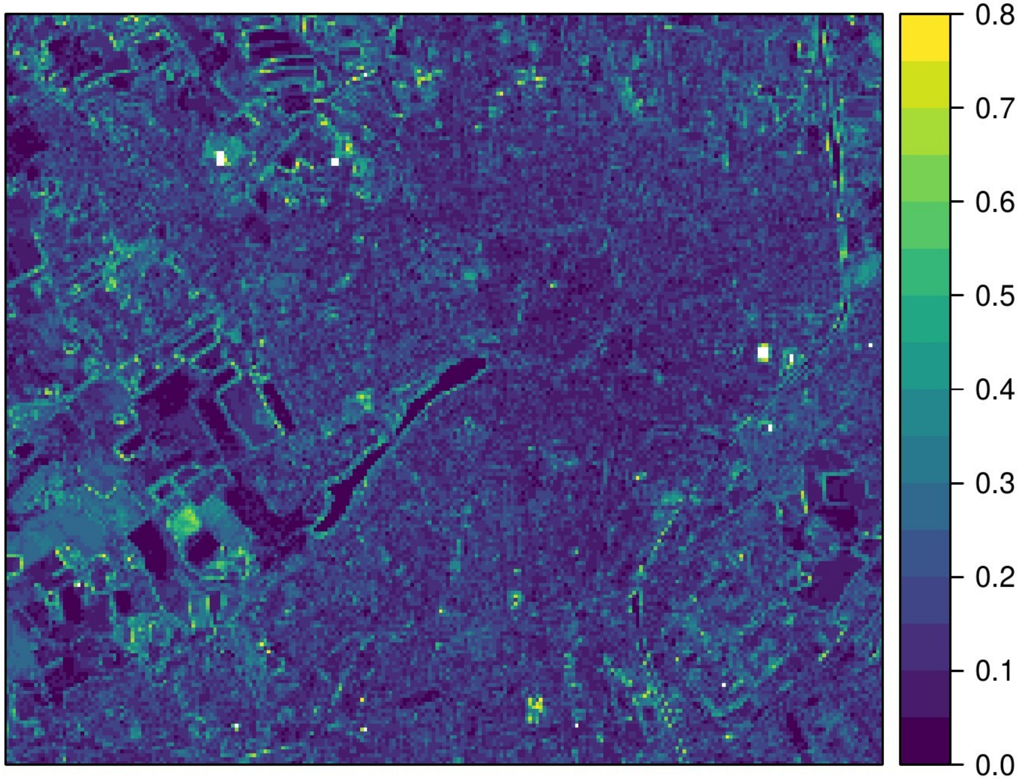
→ Estimate the area of applicability of the model

How to do it in R

```
AOA <- aoa(pred_sp,model)
```

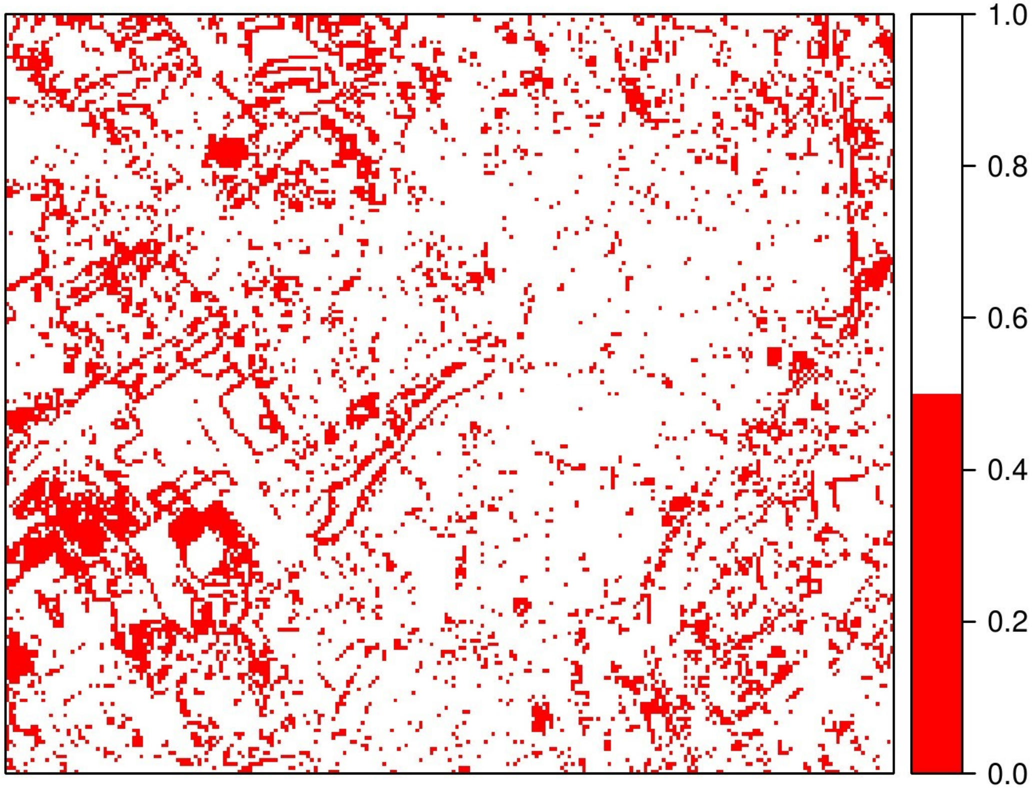
Estimating the area of applicability
Interpretation of the AOA

Dissimilarity Index



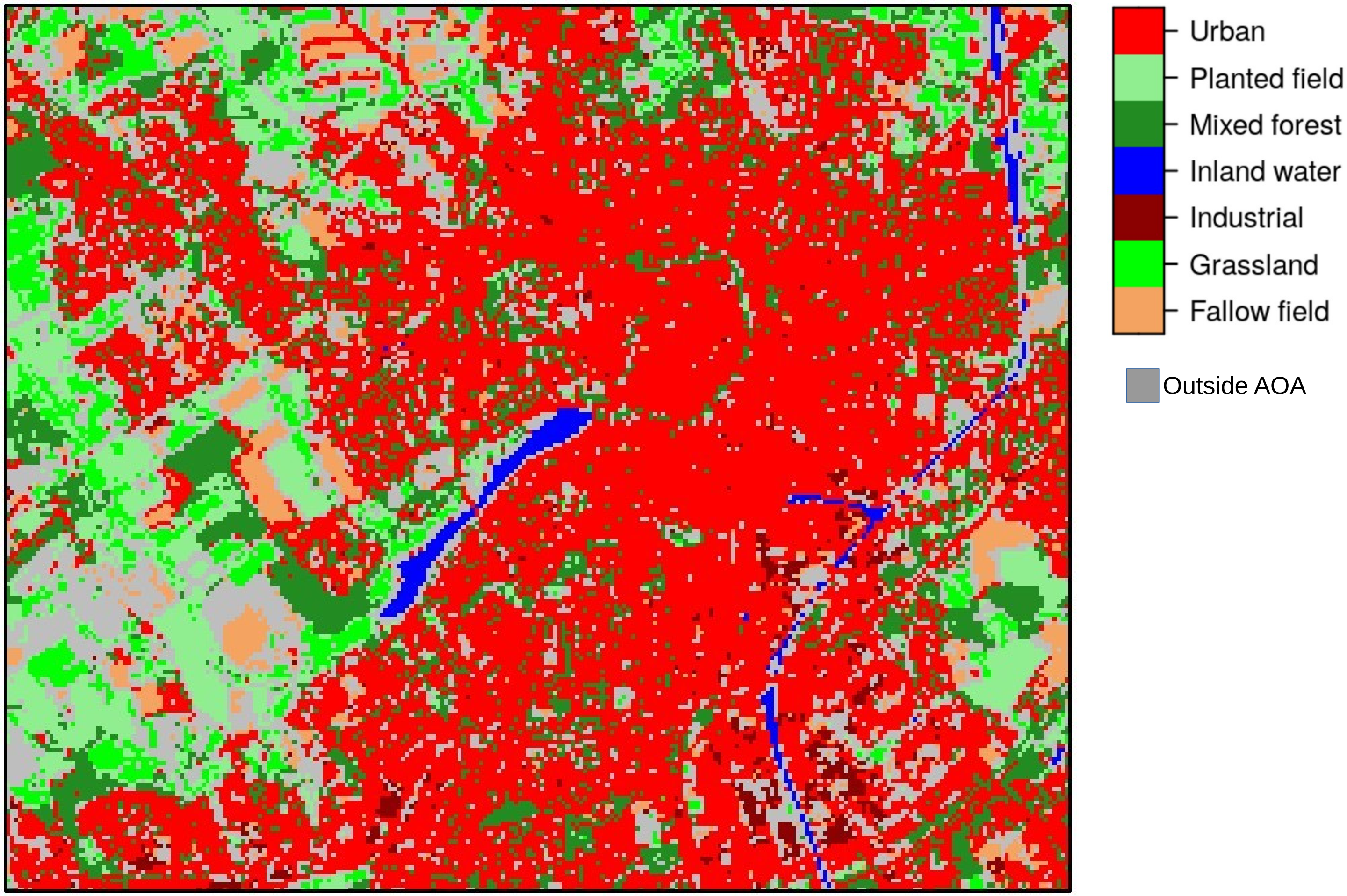
Can take values 0 to ∞

Area of Applicability



Can take values 0 = not in AOA
and 1= inside AOA

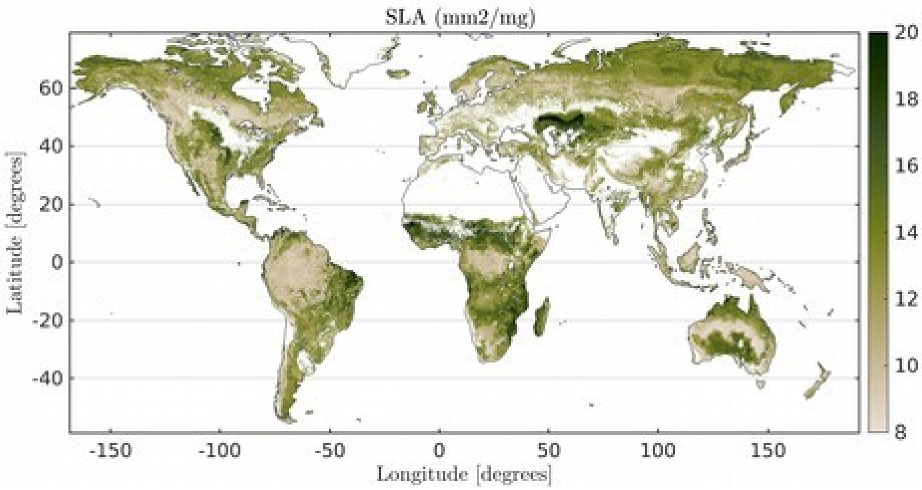
Estimating the area of applicability
Interpretation of the AOA



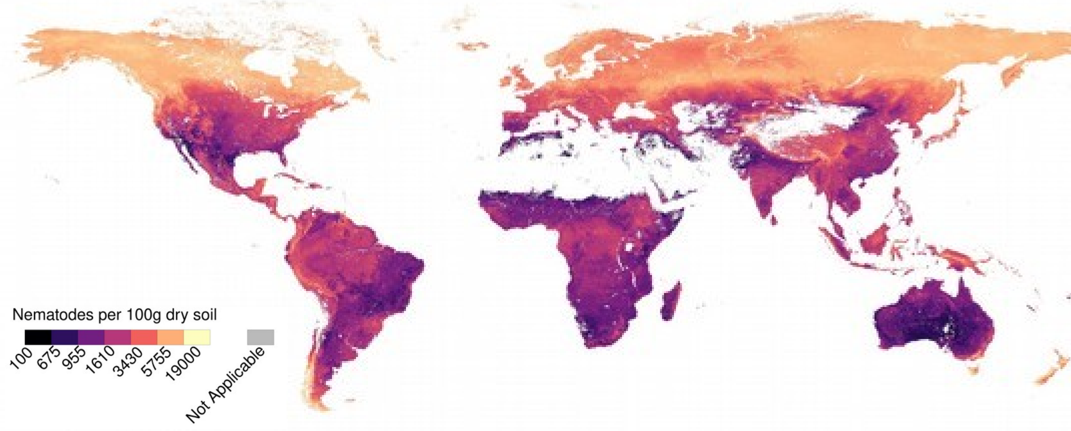
Estimation of the AOA

Estimate the Area of Applicability of your model

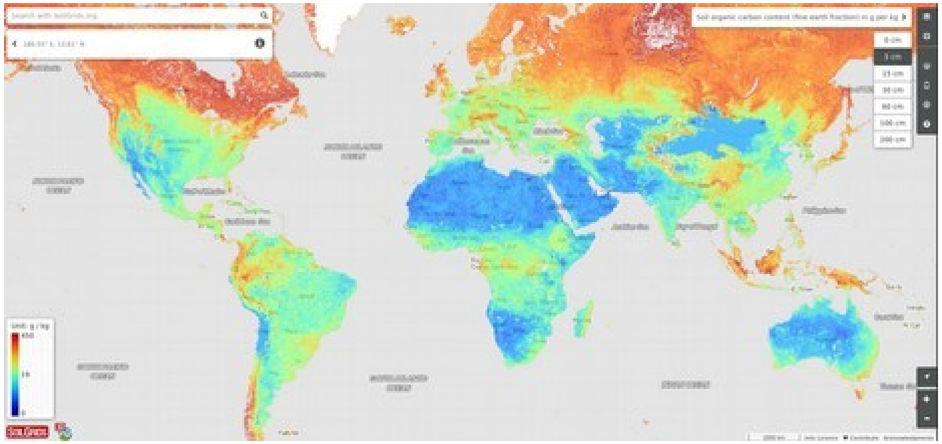
Global maps of ecosystem variables based on machine learning (a few examples)



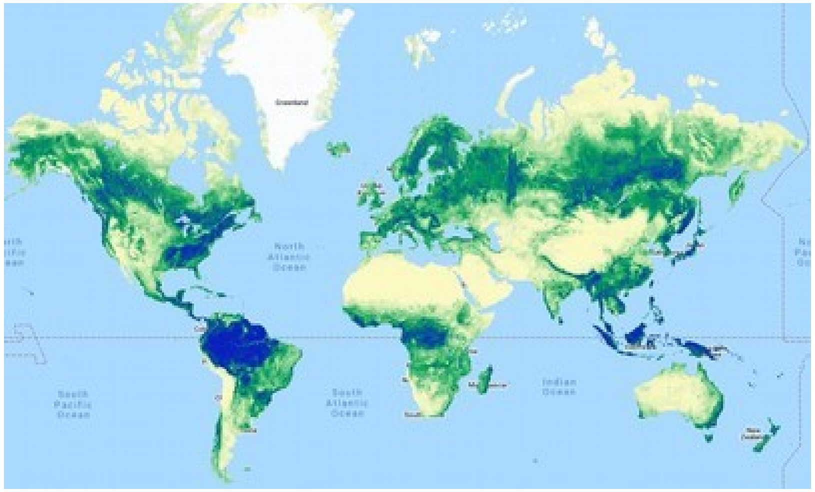
Moreno-Martínez et al., 2018



Based on van den Hoogen et al., 2019



Hengl et al., 2017



Bastin et al. 2019

...but there are increasingly doubts about the quality of these results

Wissenschaft

Wenn die KI daneben liegt

Welche Fehler drohen, wenn Forscher Wissenslücken per Computer schließen wollen, zeigen zwei aktuelle Klimastudien.

Von **Tin Fischer**

6. November 2019, 16:44 Uhr / Editiert am 9. November 2019, 17:42 Uhr / DIE ZEIT Nr. 46/2019, 7. November 2019 / 9 Kommentare

AUS DER ZEIT NR. 46/2019

DEEP TROUBLE FOR DEEP LEARNING

BY DOUGLAS HEAVEN

Nature 574, 163-166 (2019)

Home / News & Opinion

Researchers Find Flaws in High-Profile Study on Trees and Climate

Four independent groups say the work overestimates the c
global forest restoration, but the authors insist their origina

Oct 17, 2019

KATARINA ZIMMER

Comment | Published: 23 August 2021

Conservation needs to break free from global priority mapping

Carina Wyborn & Megan C. Evans

Nature Ecology & Evolution (2021) | Cite this article

Have we been too ambitious? When and why might the models fail?

Summary

- Accuracy assessment: prediction situations created during CV resemble those encountered during prediction
- For error assessment but also for model selection (tuning)
- Predictions should only be made for the AOA (accept gaps!?)
- We (= producers of the maps) are responsible for clearly indicating usage of maps, don’t leave it to the user.