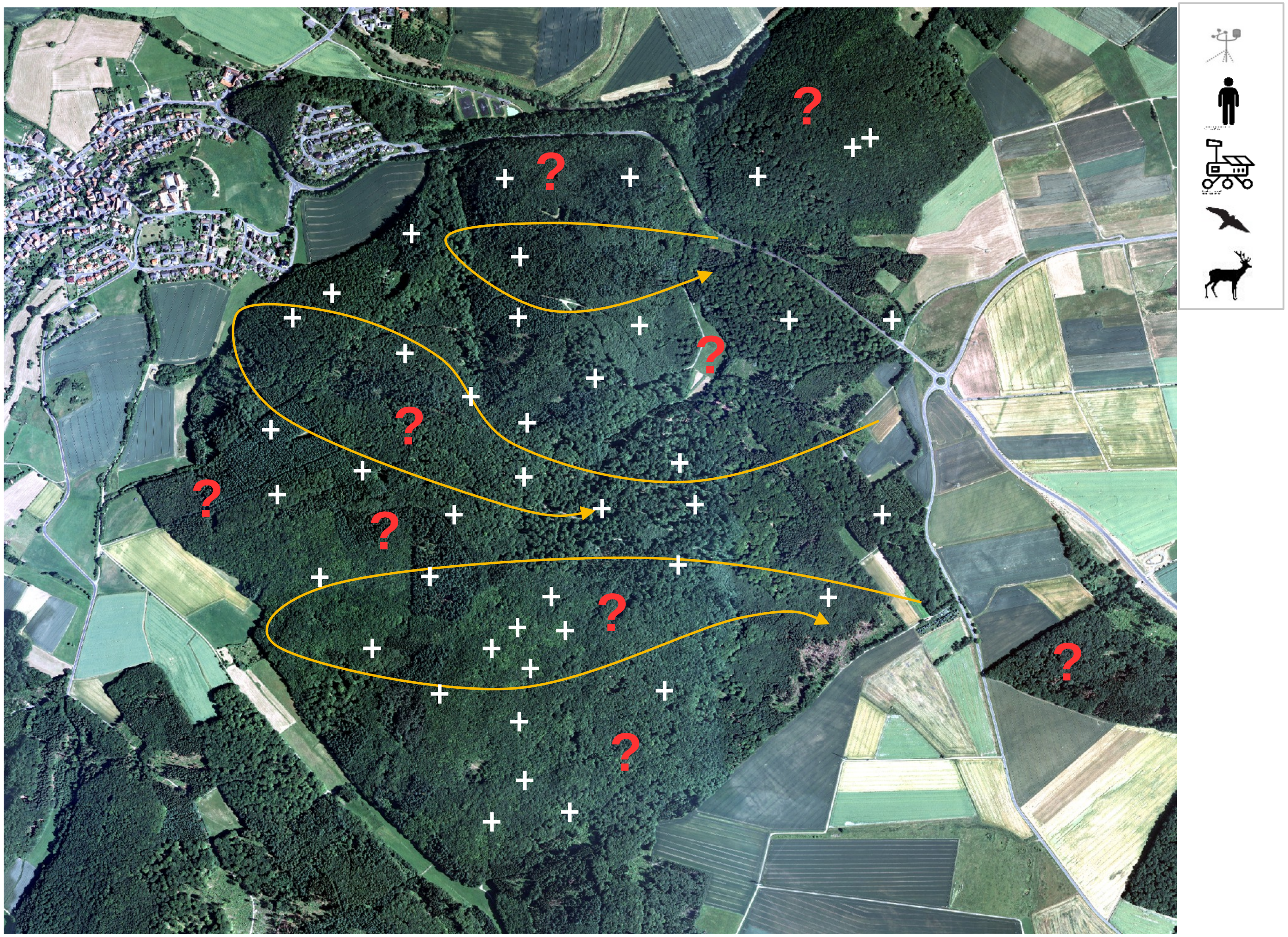
Machine learning as a tool to map the world?
Challenges and perspectives of spatial predictive mapping of the environment

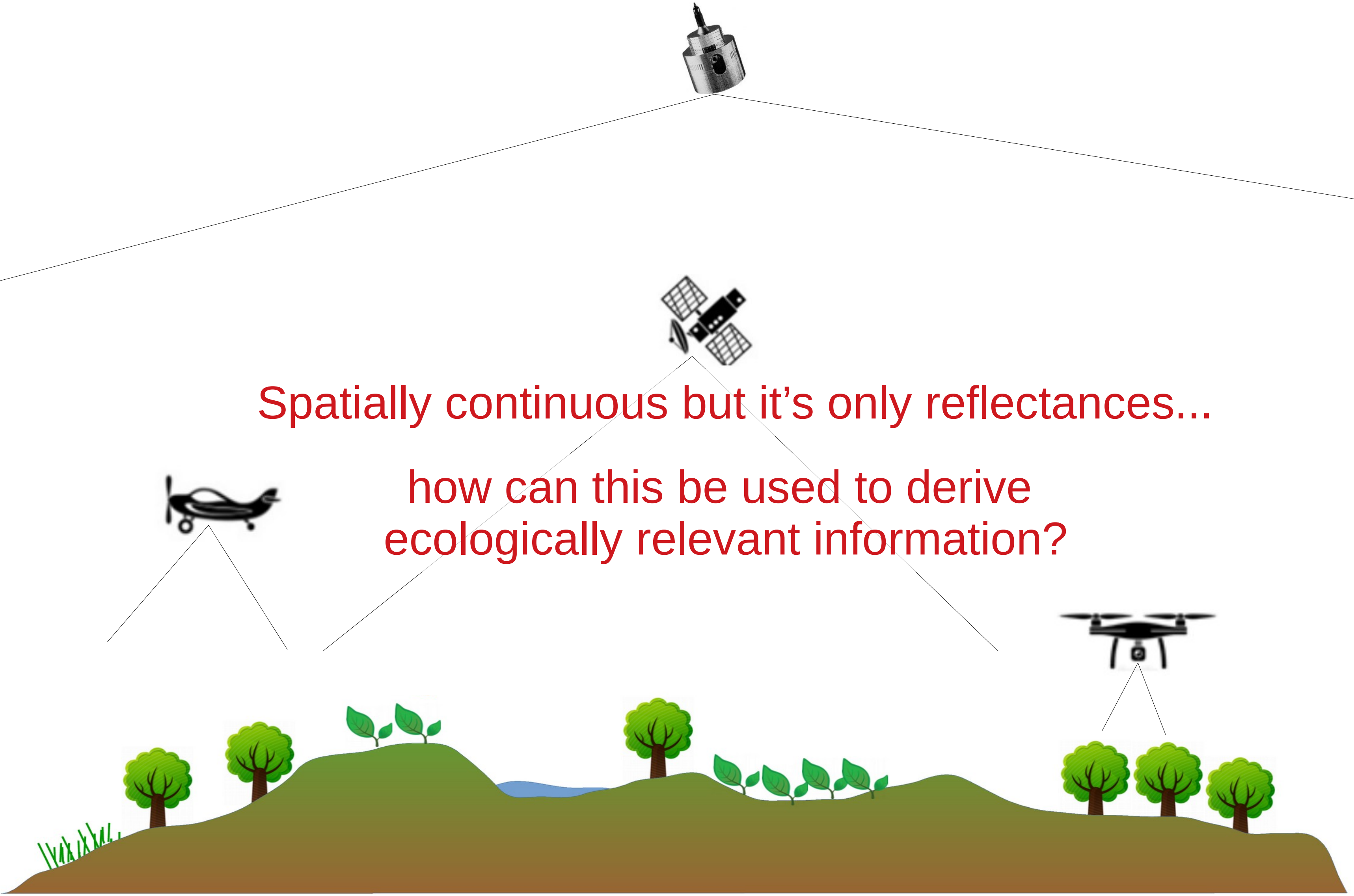
Hanna Meyer
Remote Sensing & Spatial Modelling,
Institute of Landscape Ecology, Universität Münster

Problem: From field observations to maps of ecosystem variables

Nature 4.0 | Sensing Biodiversity

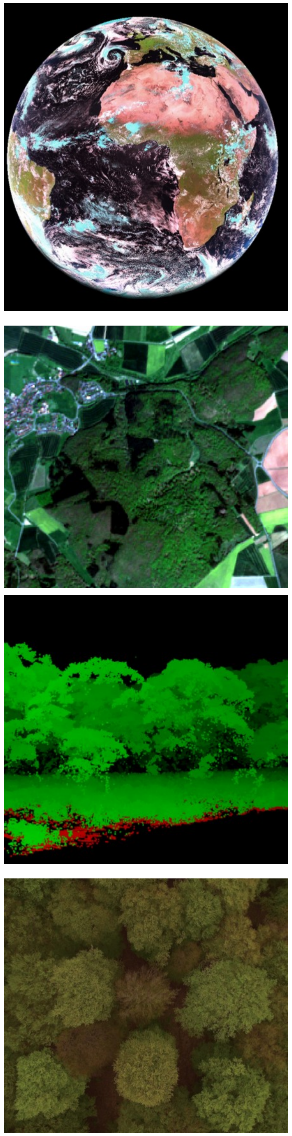


Remote Sensing to derive continuous information



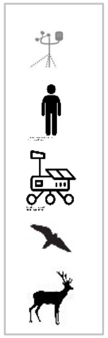
Spatially continuous but it’s only reflectances...

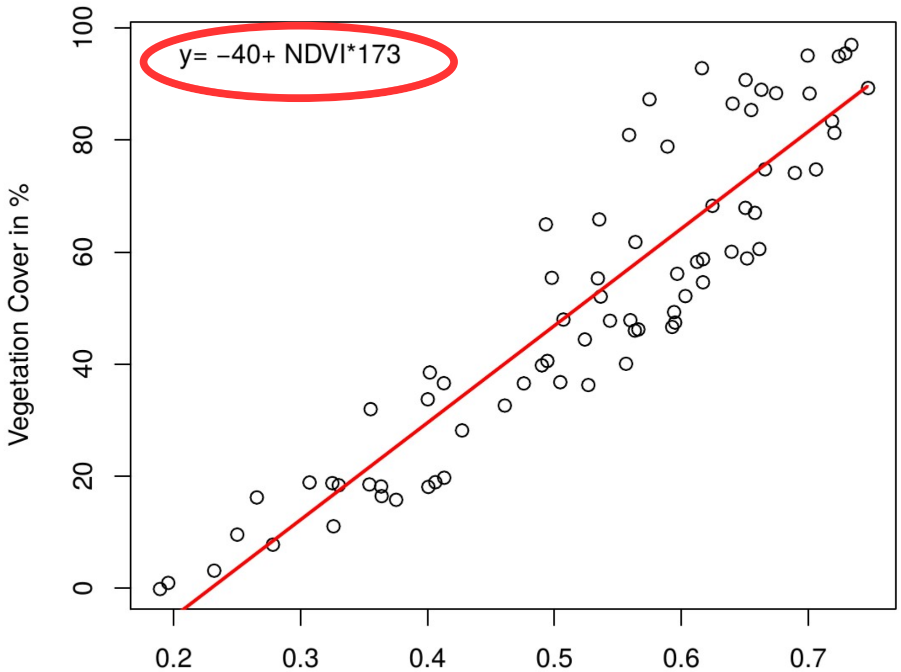
how can this be used to derive ecologically relevant information?



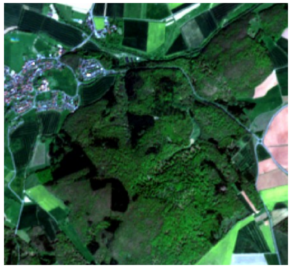
Statistical modelling

z.B. Vegetationsbedeckung aus Fernerkundungsdaten

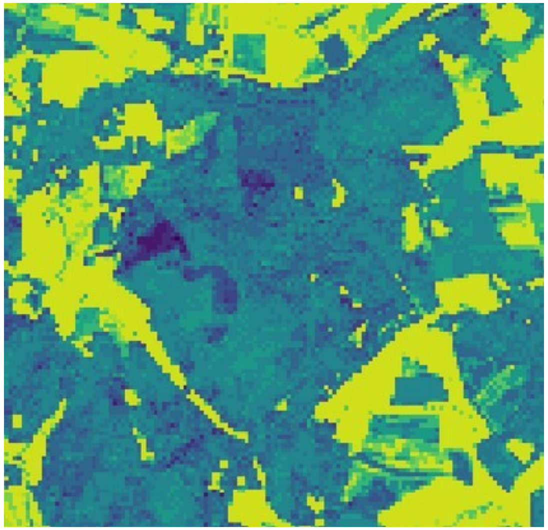




NDVI

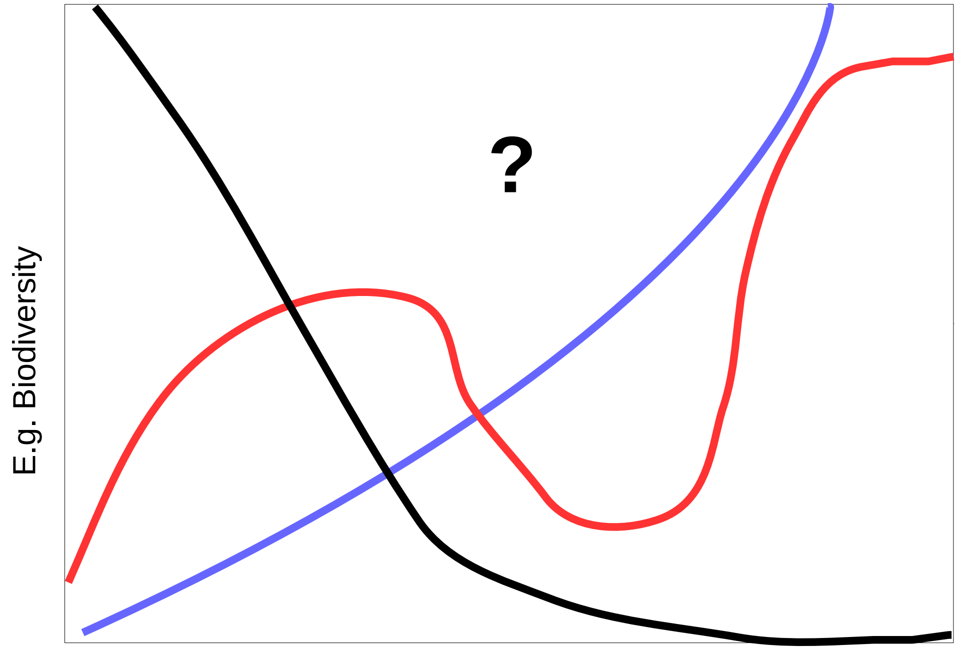


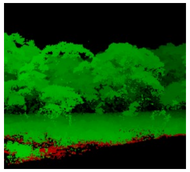
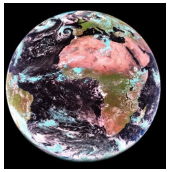
Modellierte Vegetationsbedeckung



Statistical modelling

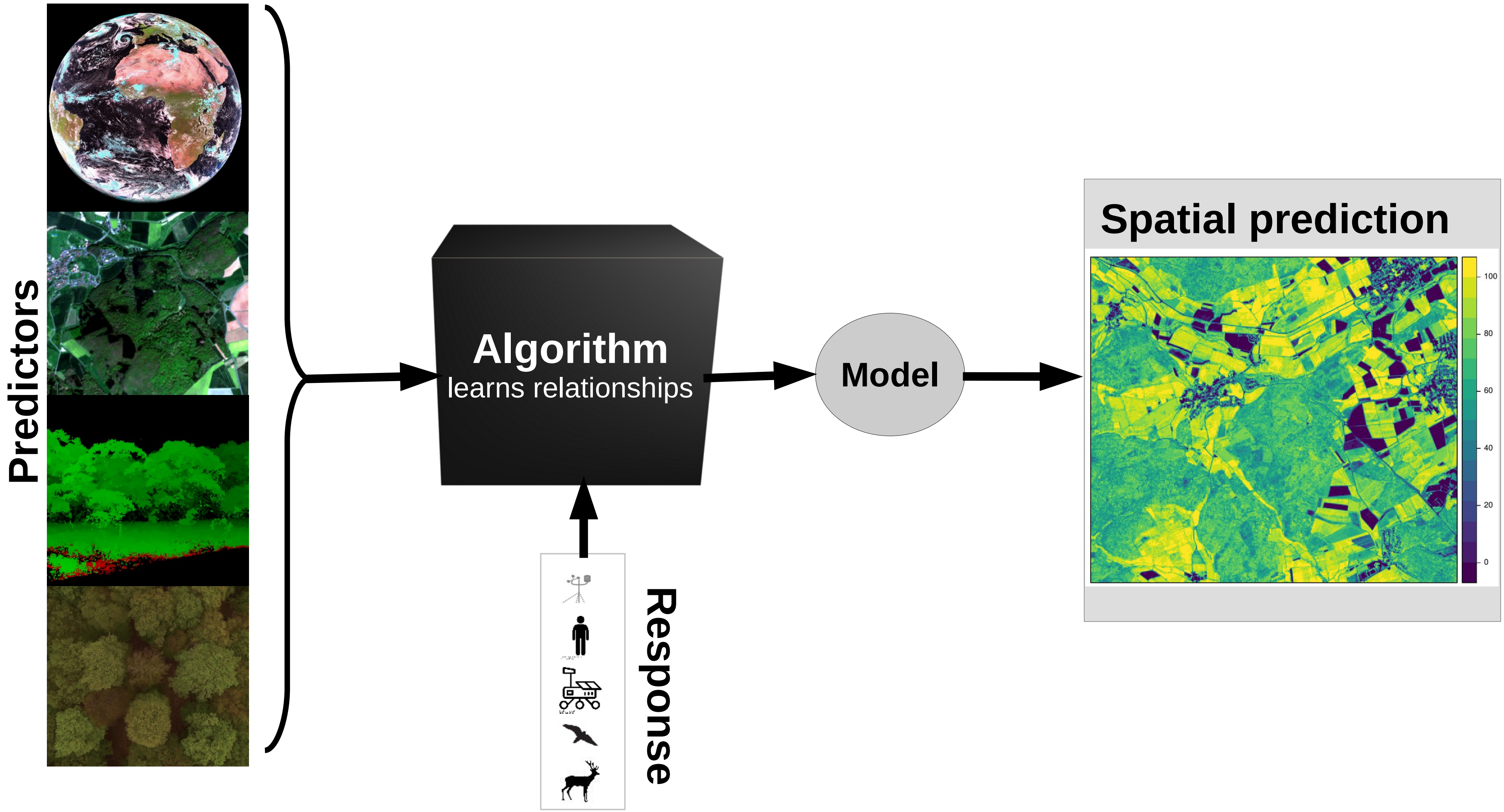
...aber was ist mit komplexeren ökologischen Variablen?



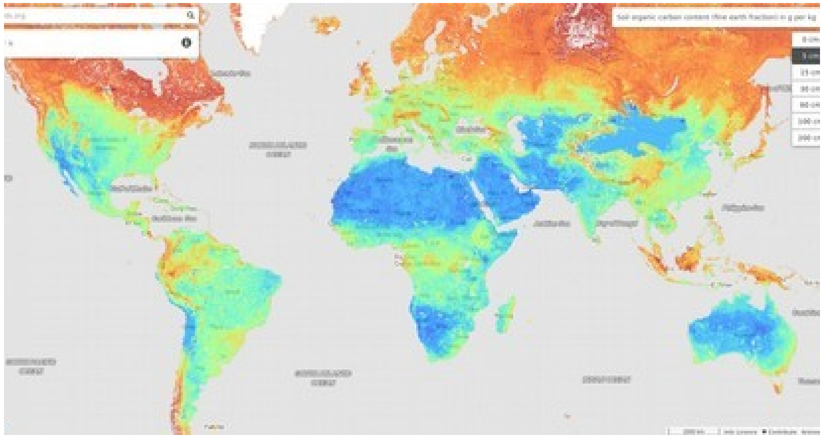
Wir brauchen Modelle die mit komplexen Zusammenhängen umgehen können!

From field observations to maps of ecosystem variables

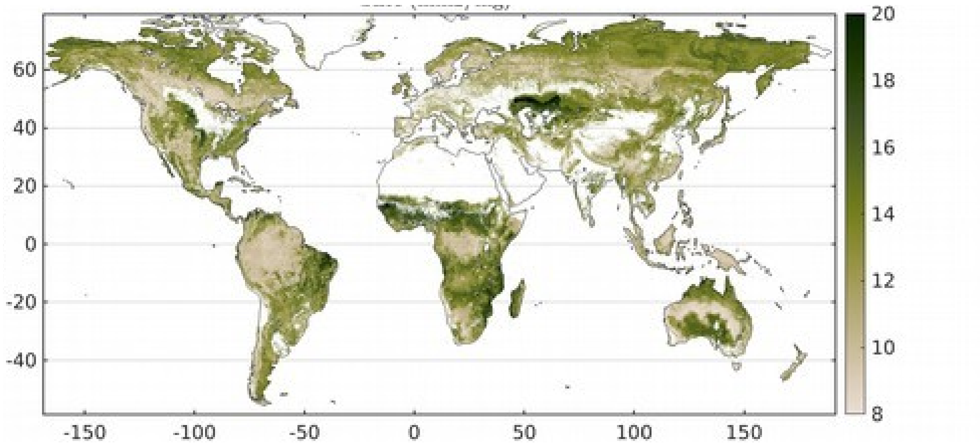


Global maps of ecosystem variables based on machine learning (a few examples)

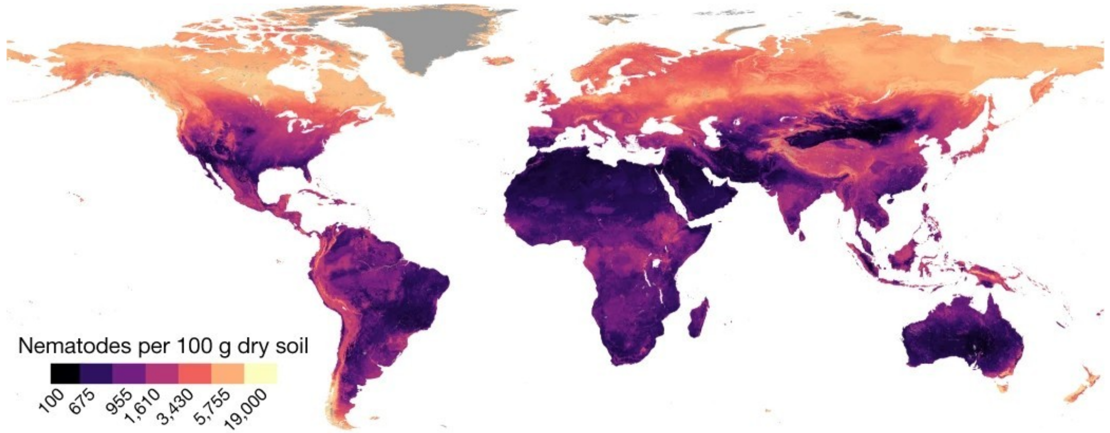
Soil organic carbon (Hengl et al., 2018)



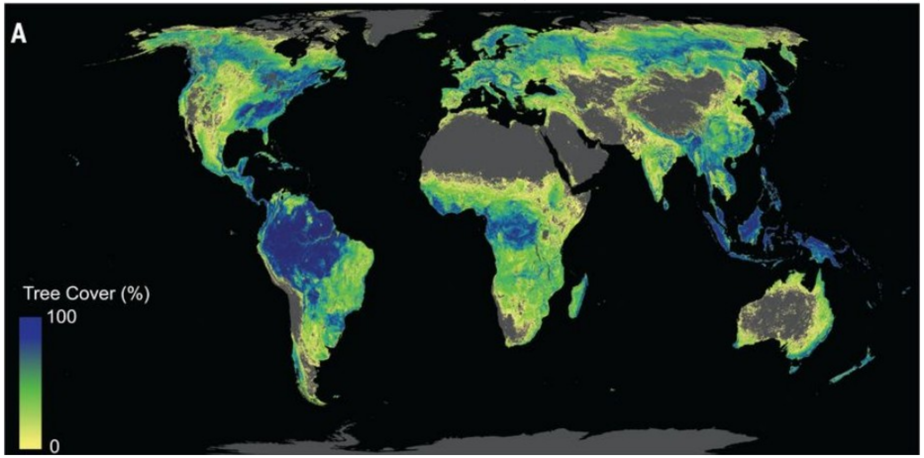
Specific Leaf Area (Moreno-Martínez et al., 2018)



Abundances of Nematodes (van den Hoogen et al., 2019)

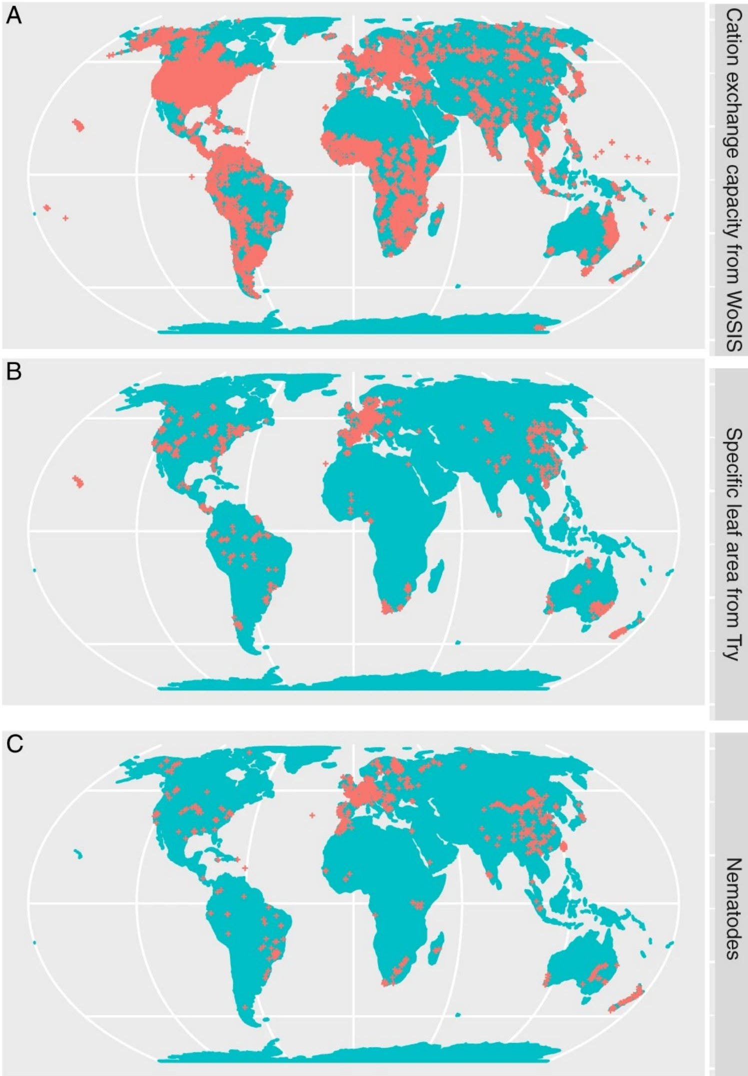


Potential forest cover (Bastin et al. 2019)



Machine learning as a tool to map everything ?

But the models are trained on very few and biased samples



Meyer & Pebesma (2022)

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

Four independent groups say the work overestimates the carbon-absorbing benefits of global forest restoration, but the authors insist their original estimates are accurate.

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

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

DEEP TROUBLE FOR DEEP LEARNING

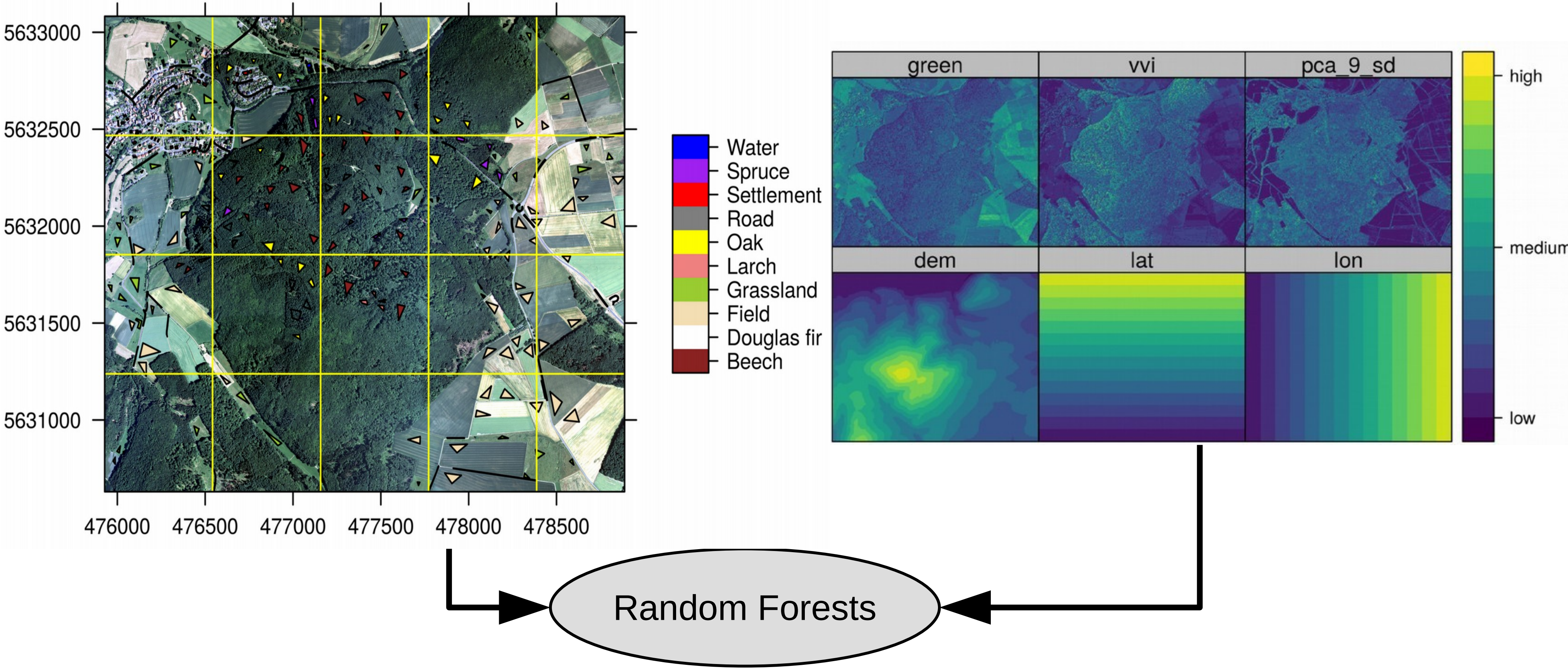
BY DOUGLAS HEAVEN

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

Example of a land cover classification

Aerial image overlayed by training sites

Example of predictors



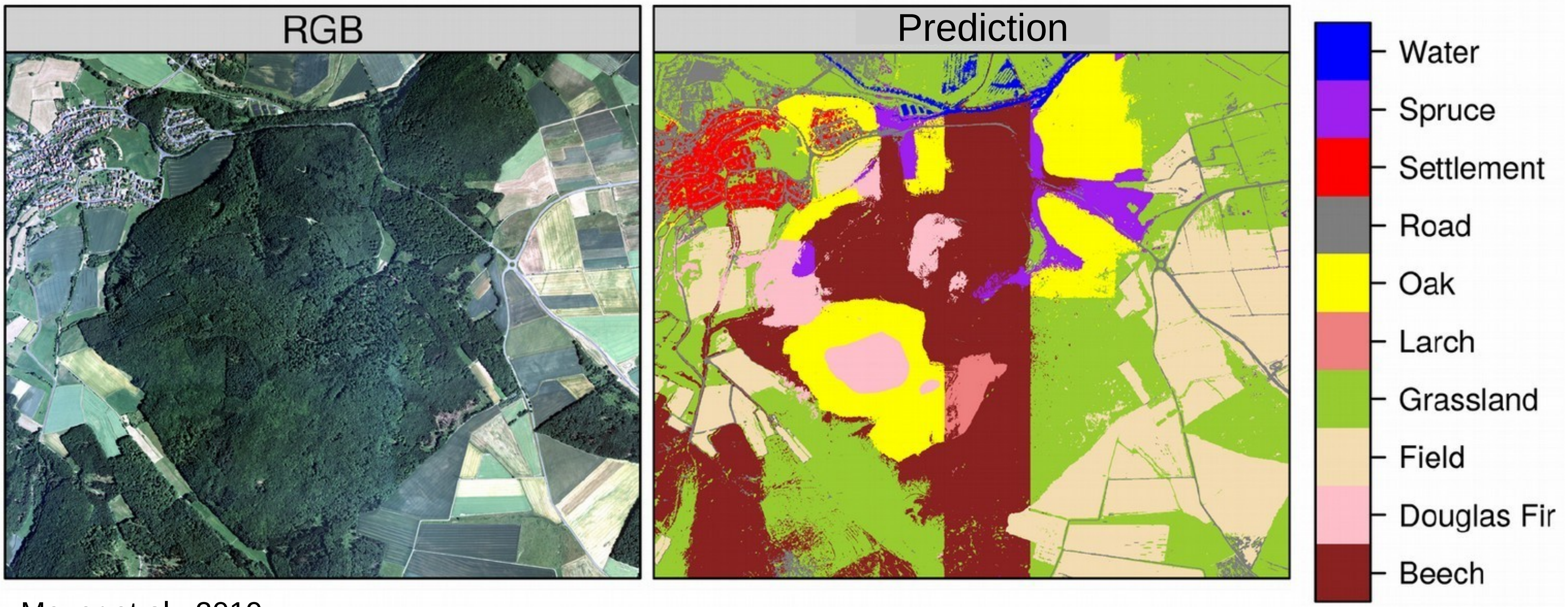
How well can we model land cover with this approach?

Performance assessment by the default validation strategy

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

Perfect prediction?

...but it doesn’t look like a perfect prediction

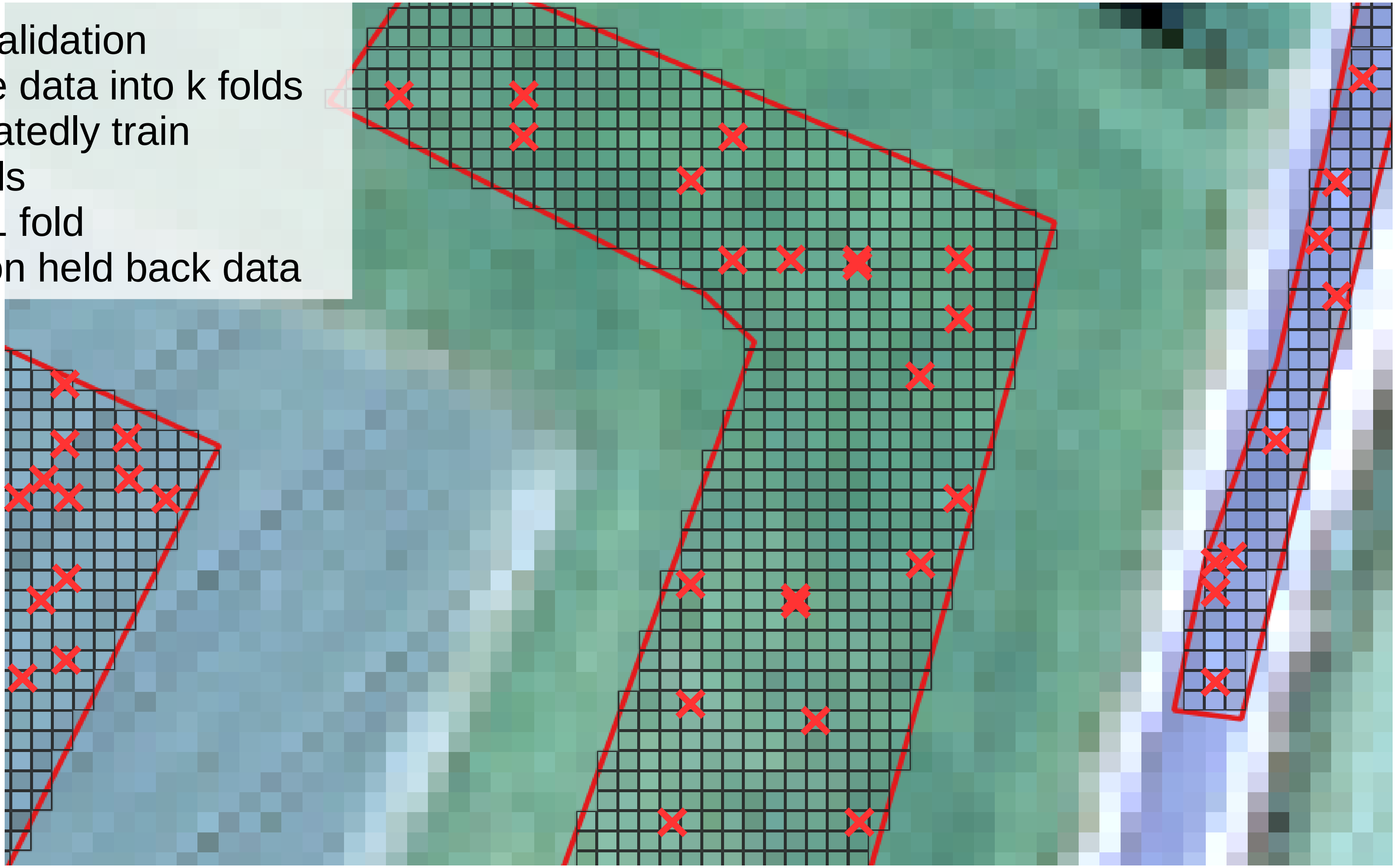


Meyer et al., 2019

But statistically it’s a perfect model.
How is this possible?

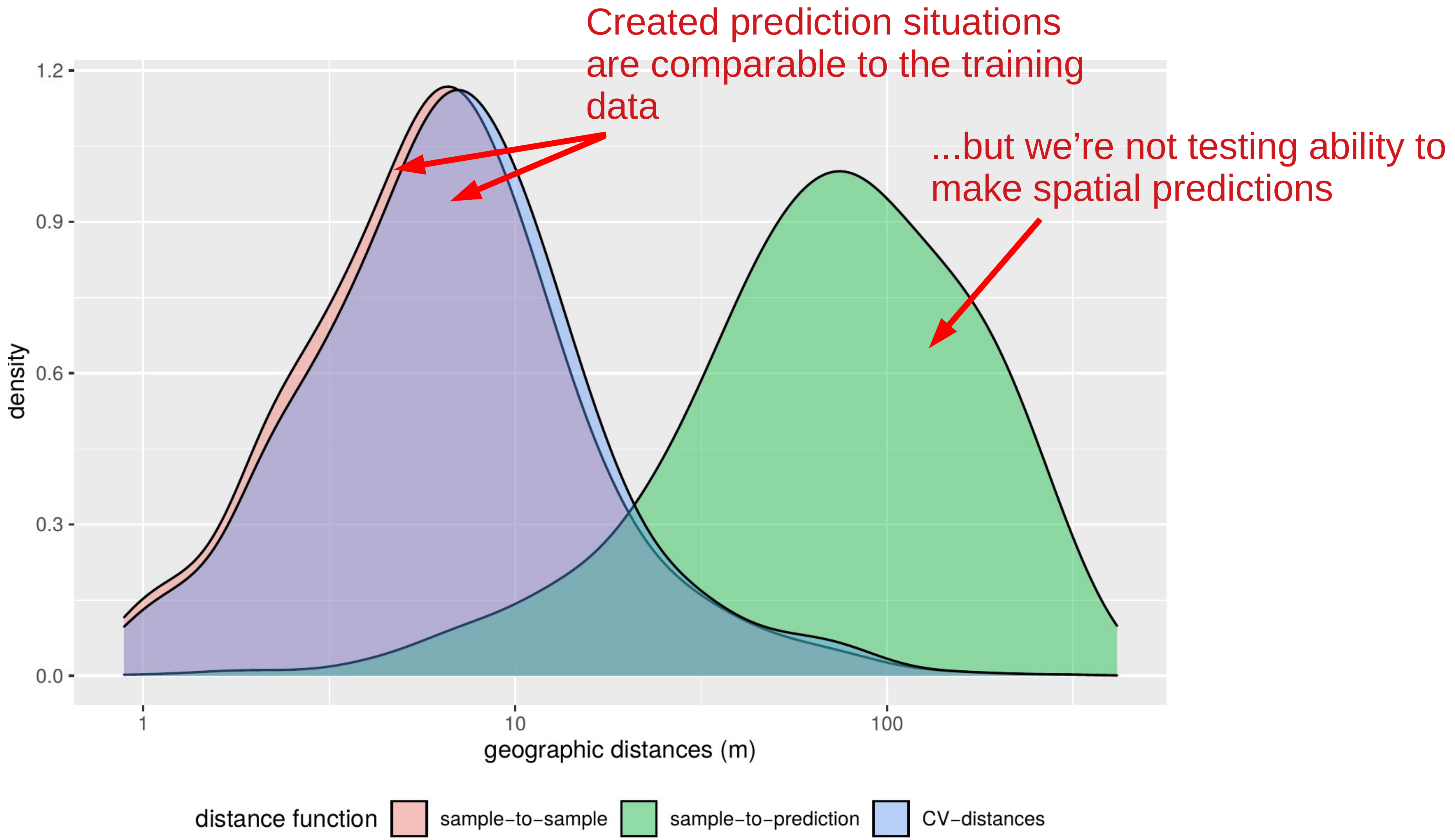
Assessment of performance by default random cross-validation

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



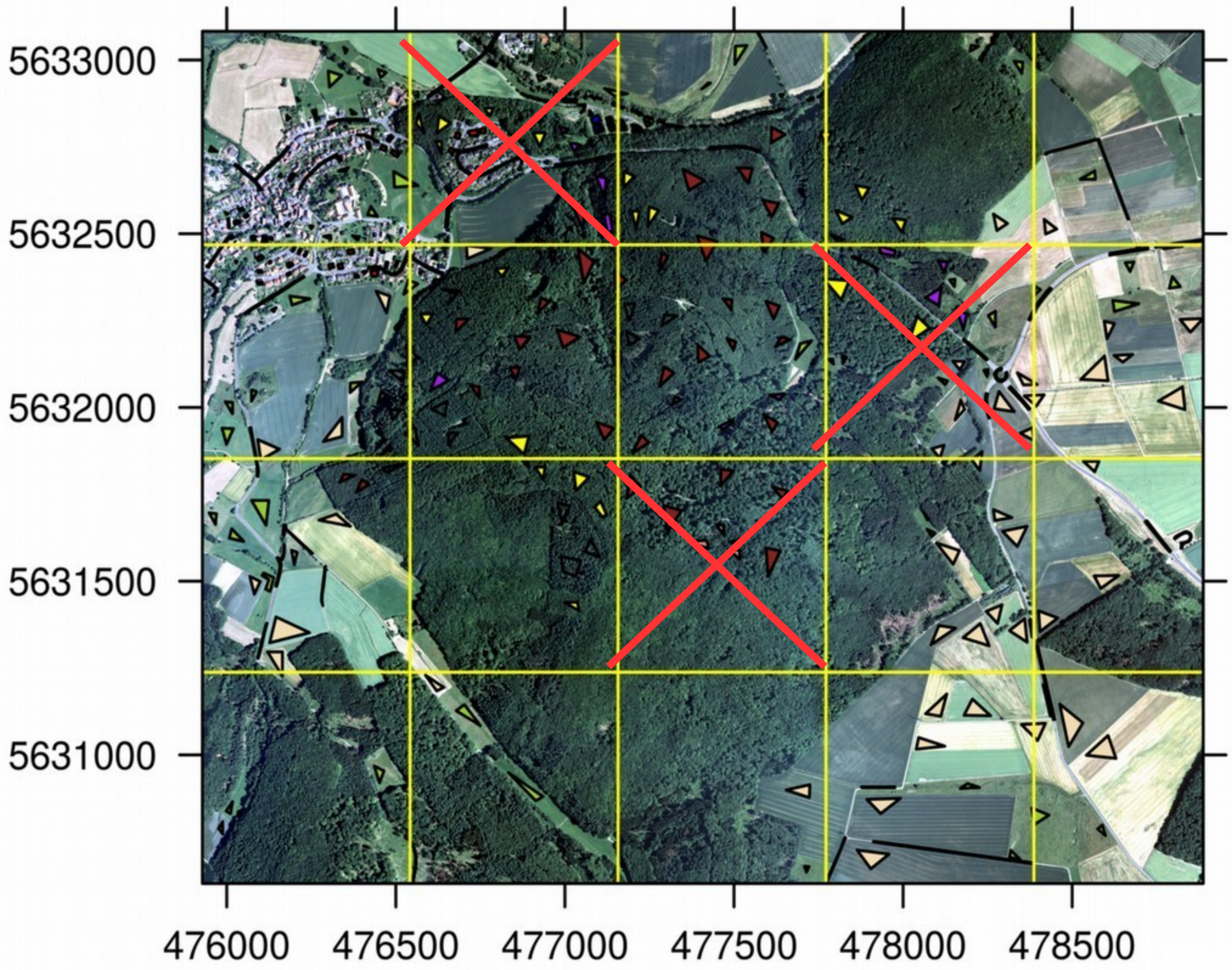
Answers question how well model performs on very similar locations

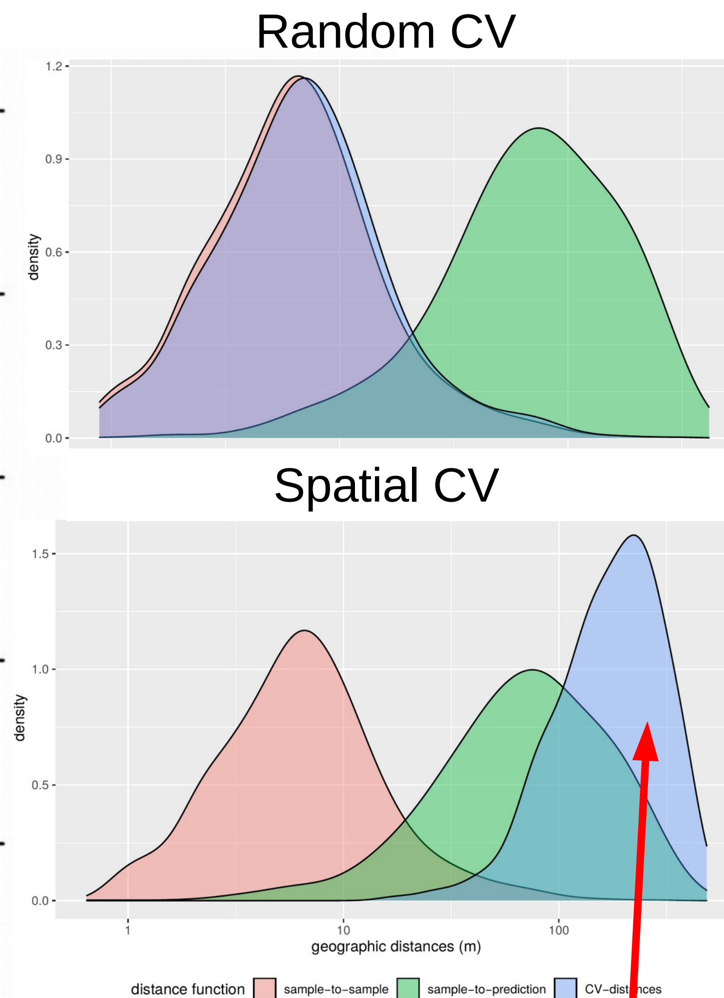
Assessment of performance by default random cross-validation



Assessment of spatial performance

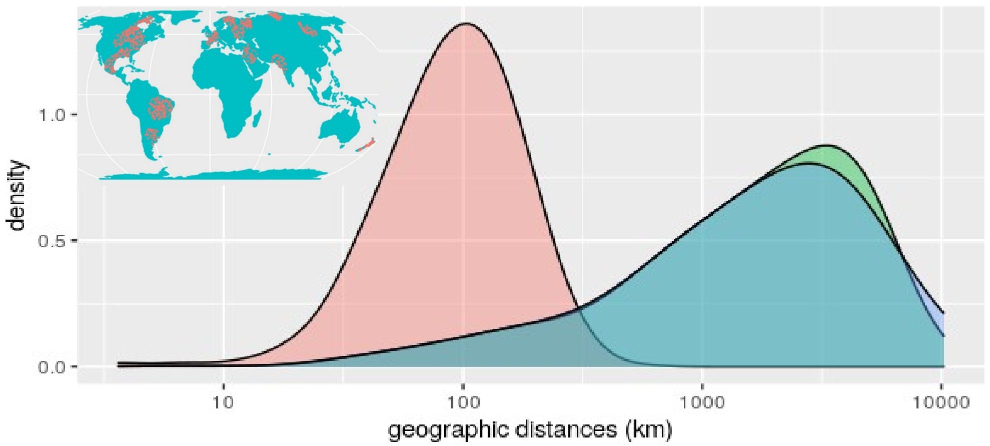
...But the aim is to fill the gaps between sampling locations!
Spatial cross-validation is required

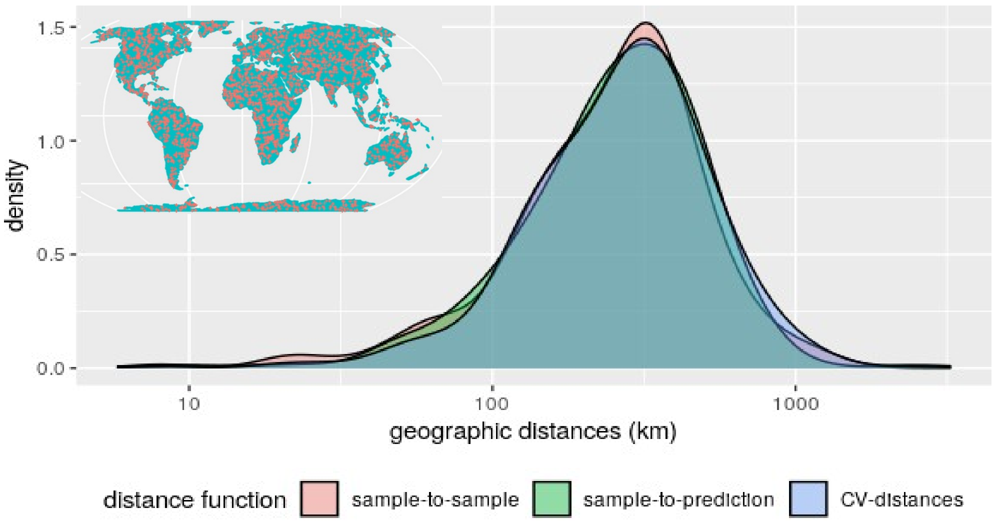




CV predictions
are harder than the prediction task

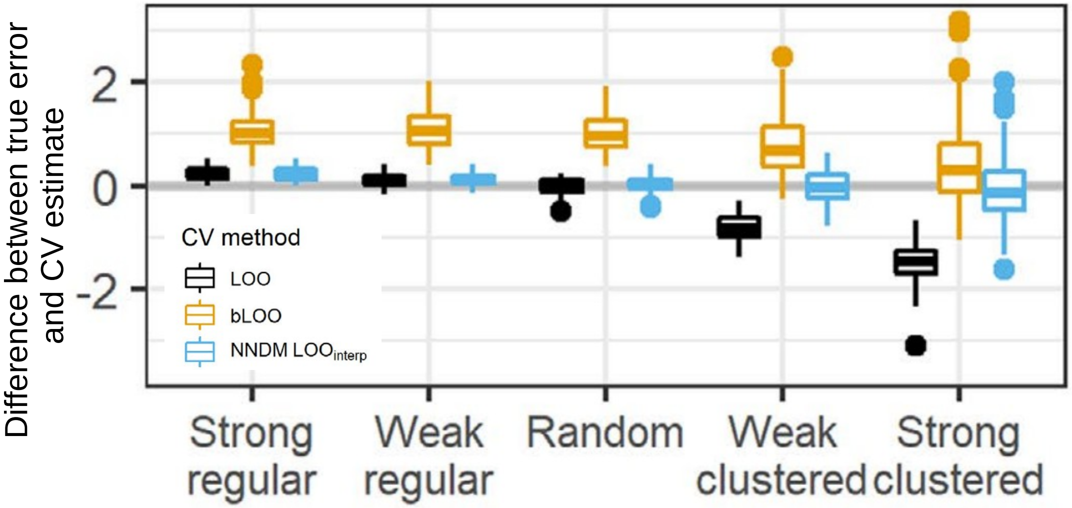
Suggestion of a nearest neighbor distance matching cross-validation





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

→ Suggestion of a “nearest neighbor distance matching CV” (Mila et al., 2022; Linnenbrink et al., 2024)



Mila et al., 2022

Assessment of spatial performance

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

- Standard validation procedures may lead to an overoptimistic view on prediction performance!
- Prediction situations created during CV need to resemble those encountered while predicting the map from the reference data

...but the relevance of spatial validation is still highly underestimated

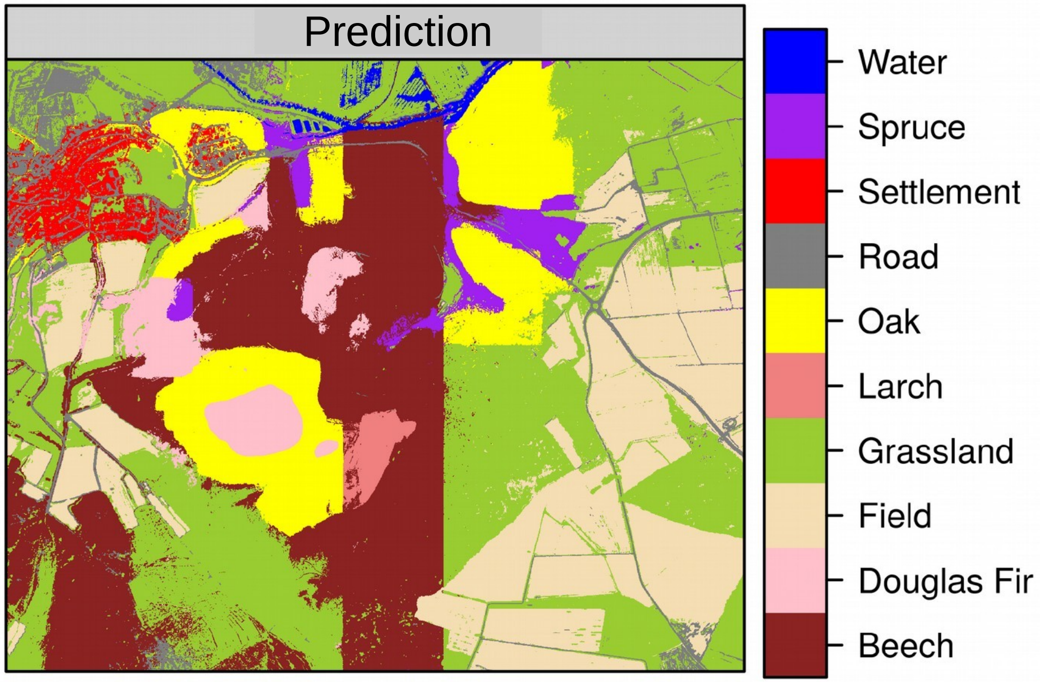
"*I am actually surprised to see the poor performance of your NN approach[...]. Typically with sufficient training data a NN approach can often reproduce the predicted variable very well even if the underlying reasons are unknown*"
(an editor from a high impact journal in the remote sensing community)

Data reproduction is not the same as data prediction!

Random cross-validation!

A suitable spatial cross-validation!

Spatial performance of models needs to be improved!



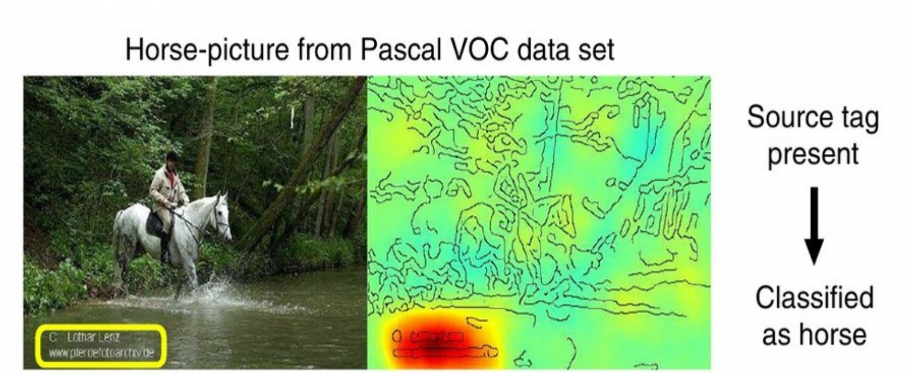
Where do these prediction patterns come from?

https://xkcd.com/1838/

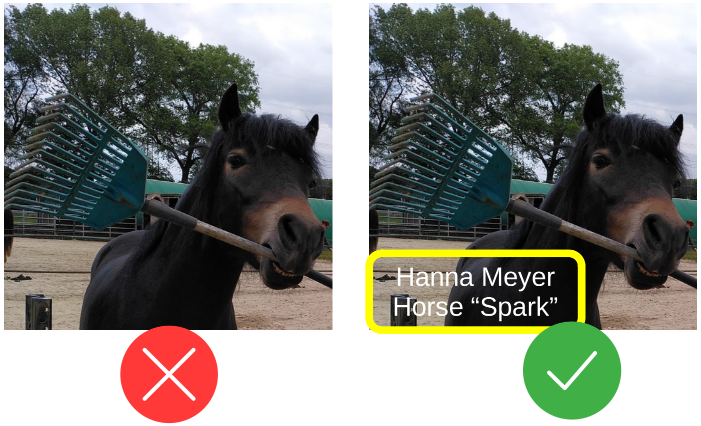
Problem: The model cannot well generalize...but why?

Is the model behaving like the “clever Hans” ?

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



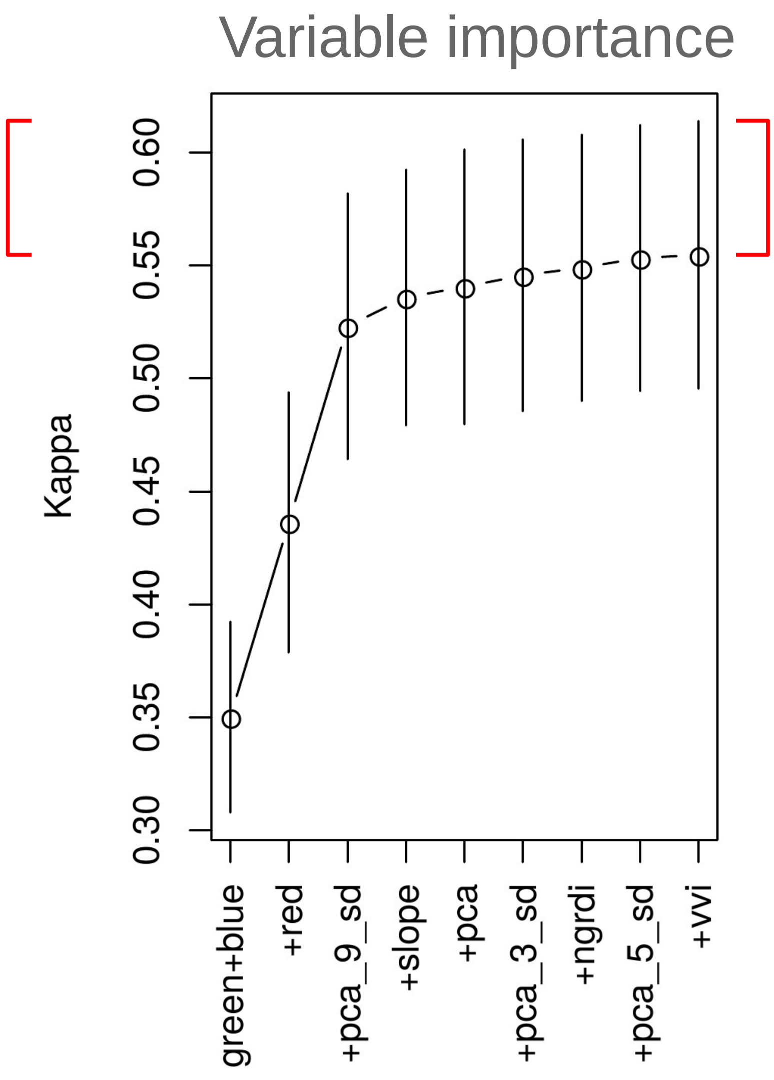
“Unmasking Clever Hans predictors and assessing what machines really learn” (Lapuschkin et al., 2019, Nature communications)



Horse?

Models that are not able to learn the scientifically meaningful relationships → not transferable!

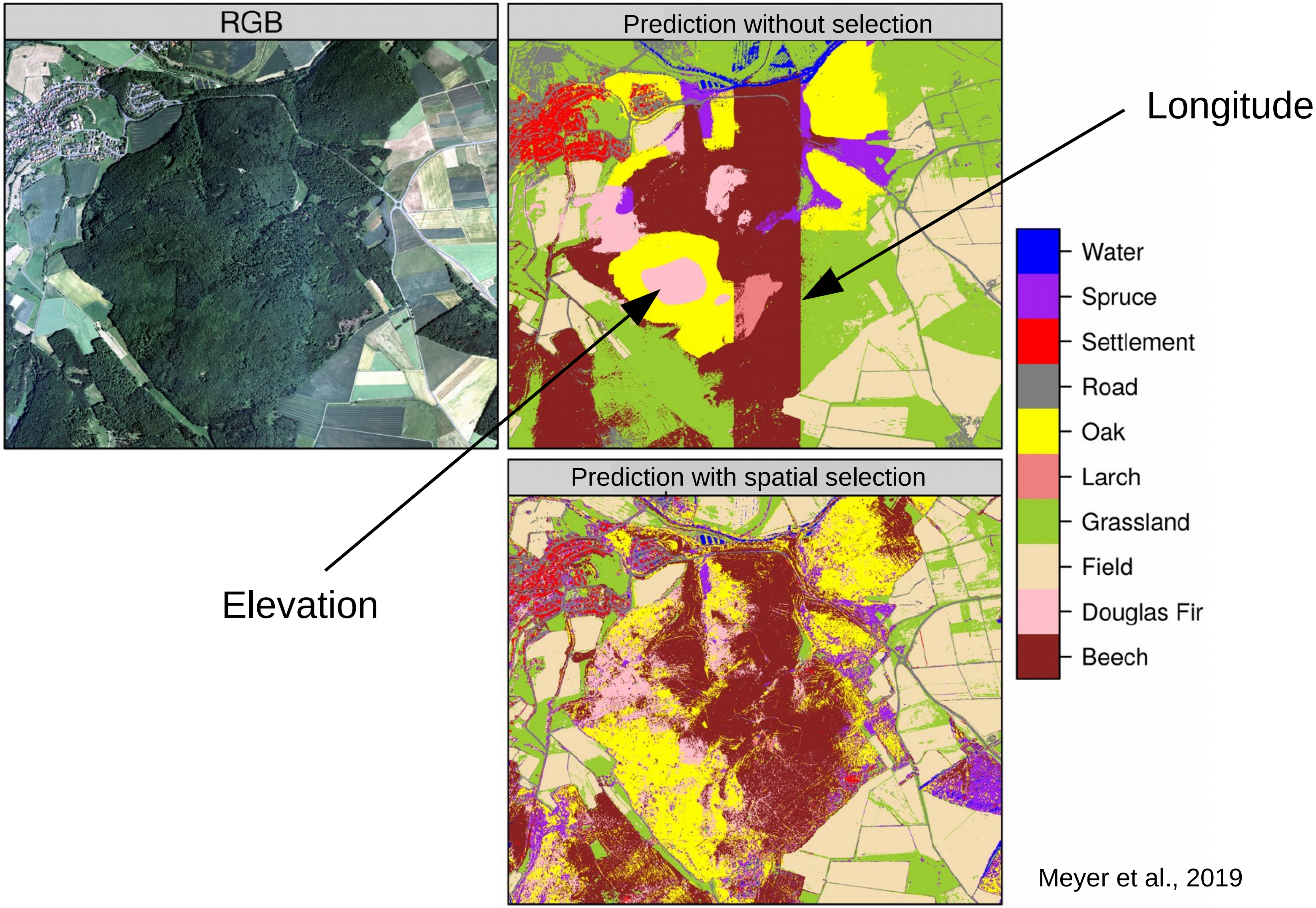
Unmasking “clever Hans predictors” to improve the model?



- Assumption: some of the variables are “clever Hans predictors”
- Removing those variables should improve the results
- Spatial variable selection required!

Implemented in our R package “CAST”

Unmasking “clever Hans predictors” to improve the model?



Meyer et al., 2019

Statistical performance of the spatial model

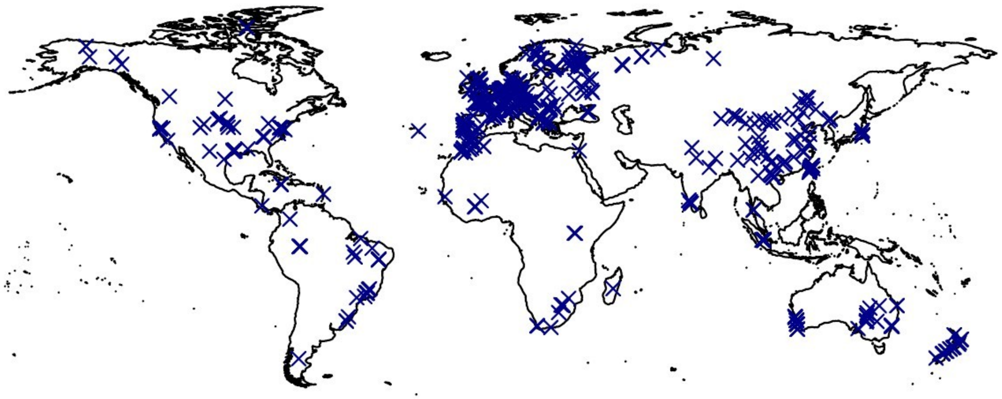
Variables	Validation	Accuracy	Kappa
all	random	>0.99	>0.99
all	spatial	**0.68**	**0.61**
selected by FFS spatial	spatial	**0.70**	**0.62**
selected by FFS spatial	random	0.78	0.82

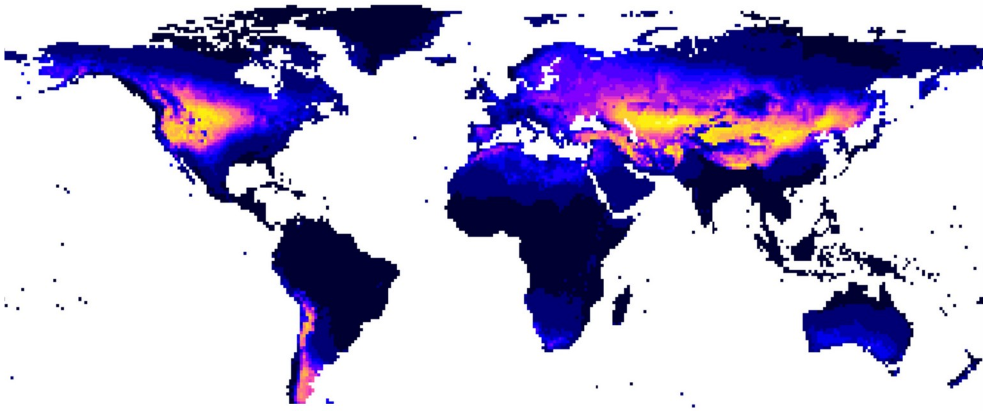
What we have learned so far...

- Cross-validation strategy affect:
  - Performance estimate
  - Selected hyperparameters
  - Variable selection
- Consequences of using an unsuitable CV:
  - Unreliable performance estimates
  - Models that can well reproduce but not necessarily predict ("clever Hans effect")
- Hence, CV strategies that fit the prediction task are required during model selection and validation!

But is this sufficient for reliable mapping ?

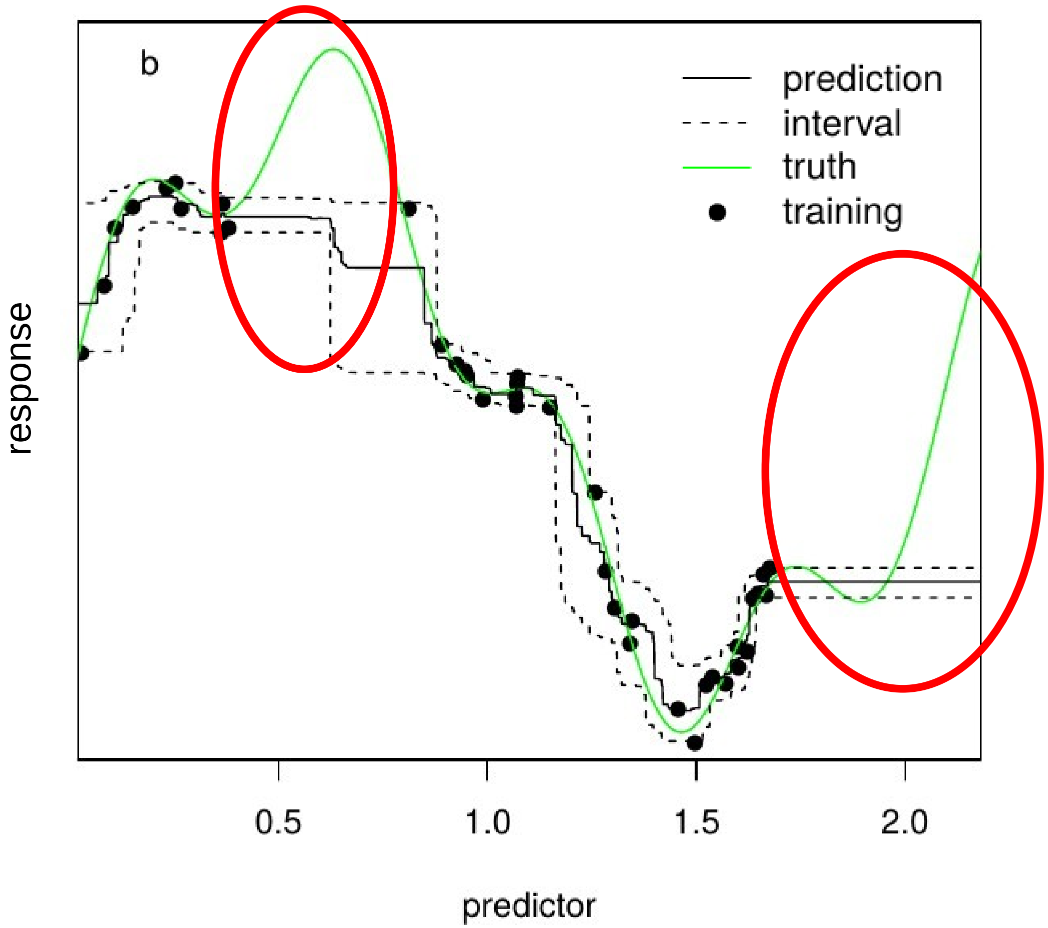
Is this sufficient for reliable mapping ?





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

Machine learning models are weak in extrapolations



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” is needed!

Suggestion: Area of Applicability (AOA)

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE | Open Access

Predicting into unknown space? Estimating the area of applicability of spatial prediction models

Hanna Meyer, Edzer Pebesma

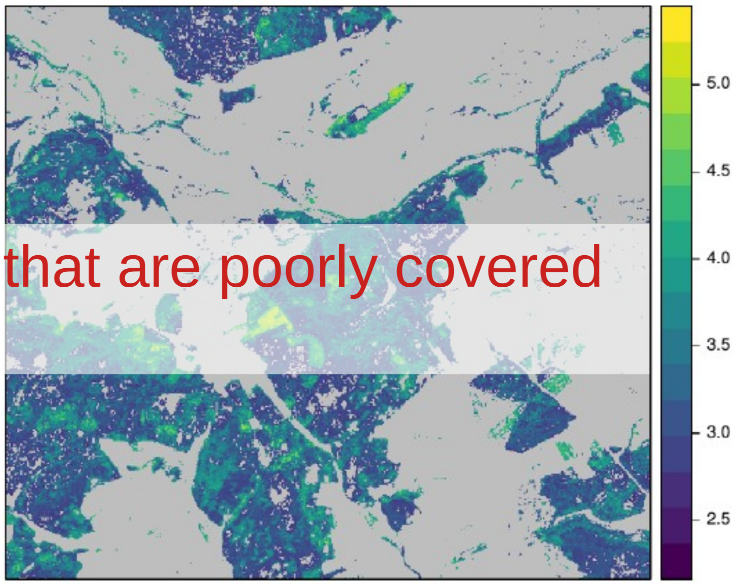
Based on distances in the predictor space, we try to derive the area...

- to which the model can be applied because it has been enabled to learn about relationships
- where the estimated performance holds

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

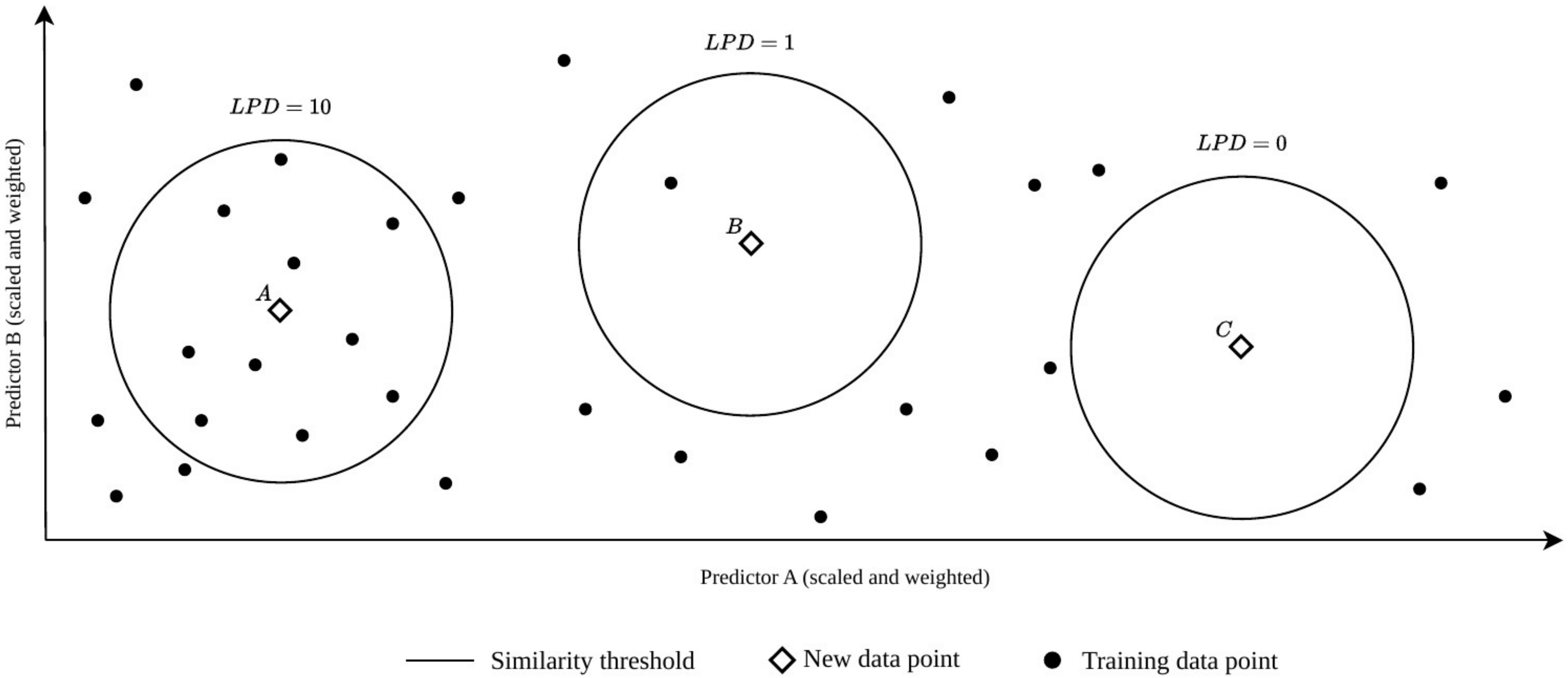
Dissimilarity Index based on distance to the nearest training point

Predictions limited to the AOA



...but the AOA does not differentiate between areas that are poorly covered and areas that are well covered by training data

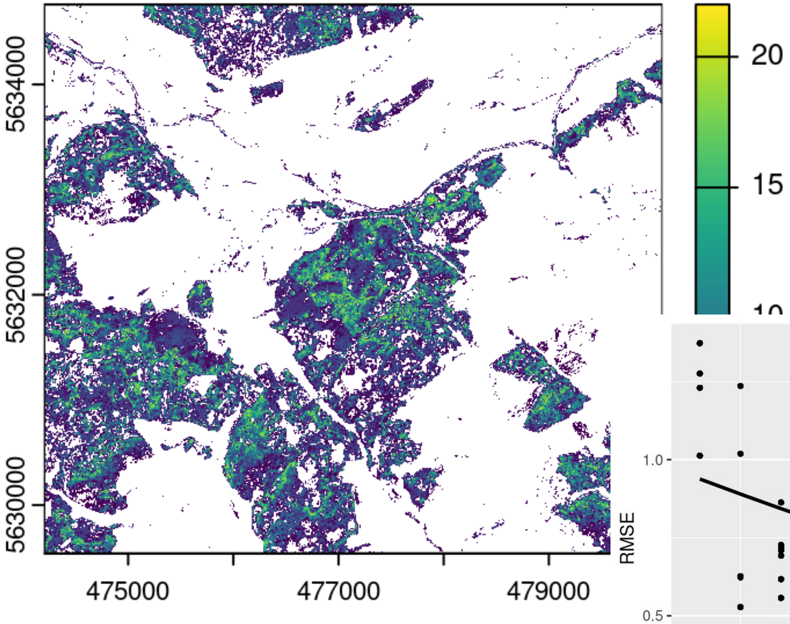
Suggestion of local training point densities (LPD)



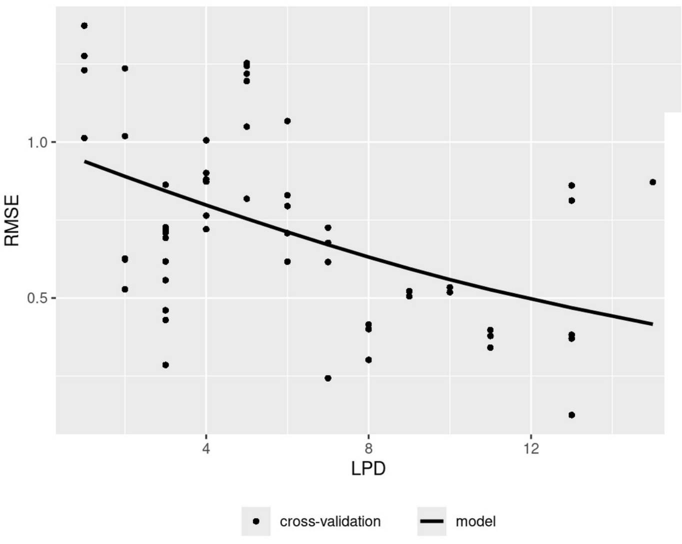
Schumacher et al., 2024

Densities in the predictor space as a measure for uncertainty?

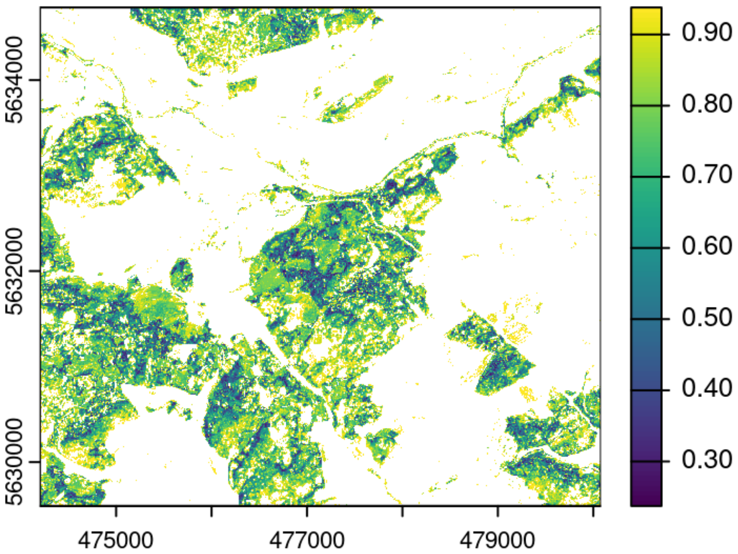
Local Point Density within AOA
(Schumacher et al., 2024)



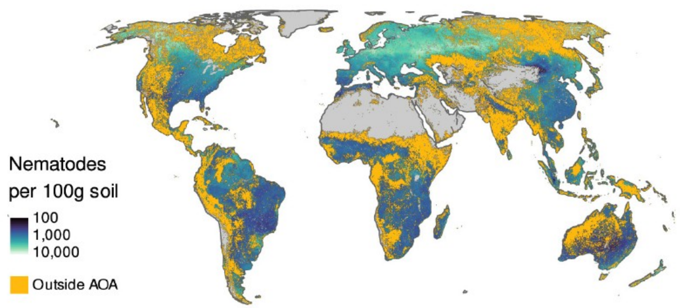
RMSE~LPD

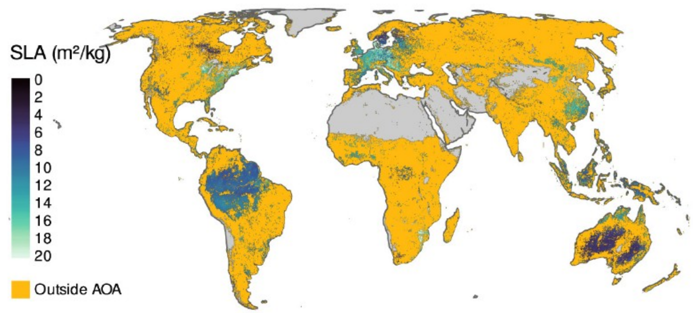


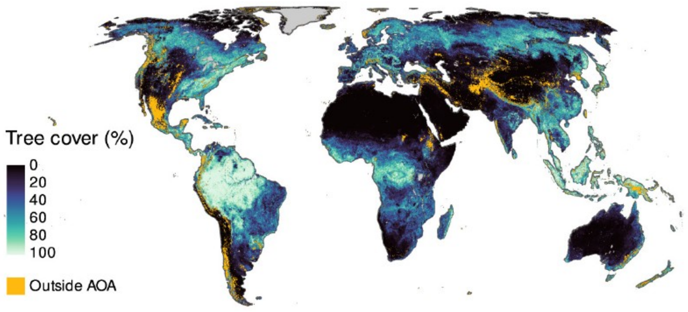
Estimated RMSE



Conclusion







Ludwig et al., 2023

- Results are not just maps; they support modeling, conservation, risk assessment, and more
- We, as map producers, must clearly communicate appropriate usage. Don’t leave interpretation to the user
- Identifying gaps isn't a flaw. It’s an opportunity to refine and improve future models towards more comprehensive monitoring of the environment

References

- Bastin et al. 2019: The global tree restoration potential. Science. Vol. 365, Issue 6448, pp. 76-79.
- Batjes, N. H., Ribeiro, E. & van Oostrum, A. Standardised soil profile datasupport global mapping and modelling (wosis snapshot 2019). Earth Syst. Sci. Data 12, 299–320 (2020).
- Hengl et al. (2017): SoilGrids250m: Global gridded soil information based on machine learning. PloS one 12(2): e0169748.
- Kattge, J. et al. TRY plant trait database – enhanced coverage and open access. Glob. Change Biol. 26, 119–188 (2020).
- Lapuschkin et al (2019): Unmasking Clever Hans predictors and assessing what machines really learn. Nature Communications volume 10.
- Linnenbrink, J, Milà, C, Ludwig, M, und Meyer, H. 2024. „kNNDM CV: k-fold nearest-neighbour distance matching cross-validation for map accuracy estimation.“ Geoscientific Model Development, Nr. 17 (15): 5897–5912. doi: 10.5194/gmd-17-5897-2024.
- Meyer H, Pebesma E. 2022. ‘Machine learning-based global maps of ecological variables and the challenge of assessing them.’ Nature Communications 13.
- Meyer H, Pebesma E (2021): Predicting into unknown space? Estimating the area of applicability of spatial prediction models. Methods in Ecology and Evolution.
- 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.
- Milà, C., Mateu, J., Pebesma, E. & Meyer, H. Nearest neighbour distance matching Leave-One-Out Cross-Validation for map validation. Methods in Ecology and Evolution. 00, 1–13 (2022).
- Moreno-Martinez, A. et al. A methodology to derive global maps of leaf traits using remote sensing and climate data. Remote Sens. Environ. 218, 69–88 (2018).
- Schramowski, P., Stammer, W., Teso, S. et al. (2020): Making deep neural networks right for the right scientific reasons by interacting with their explanations. Nat Mach Intell 2, 476–486.
- Schumacher, F, Knoth, C, Ludwig, M, und Meyer, H. 2024. „Estimation of local training data point densities to support the assessment of spatial prediction uncertainty.“ EGUsphere doi: 10.5194/egusphere-2024-2730.
- Van den Hoogen, J., Geisen, S., Routh, D. et al. (2019): Soil nematode abundance and functional group composition at a global scale. Nature 572, 194–198.