HAWK Fakultät Ressourcenmanagement Göttingen

Collection of reference data in the context of remote sensing assisted forest monitoring

Paul Magdon

Geoinformation & Forest Planning
University of Applied Sciences and Arts (HAWK), Göttingen

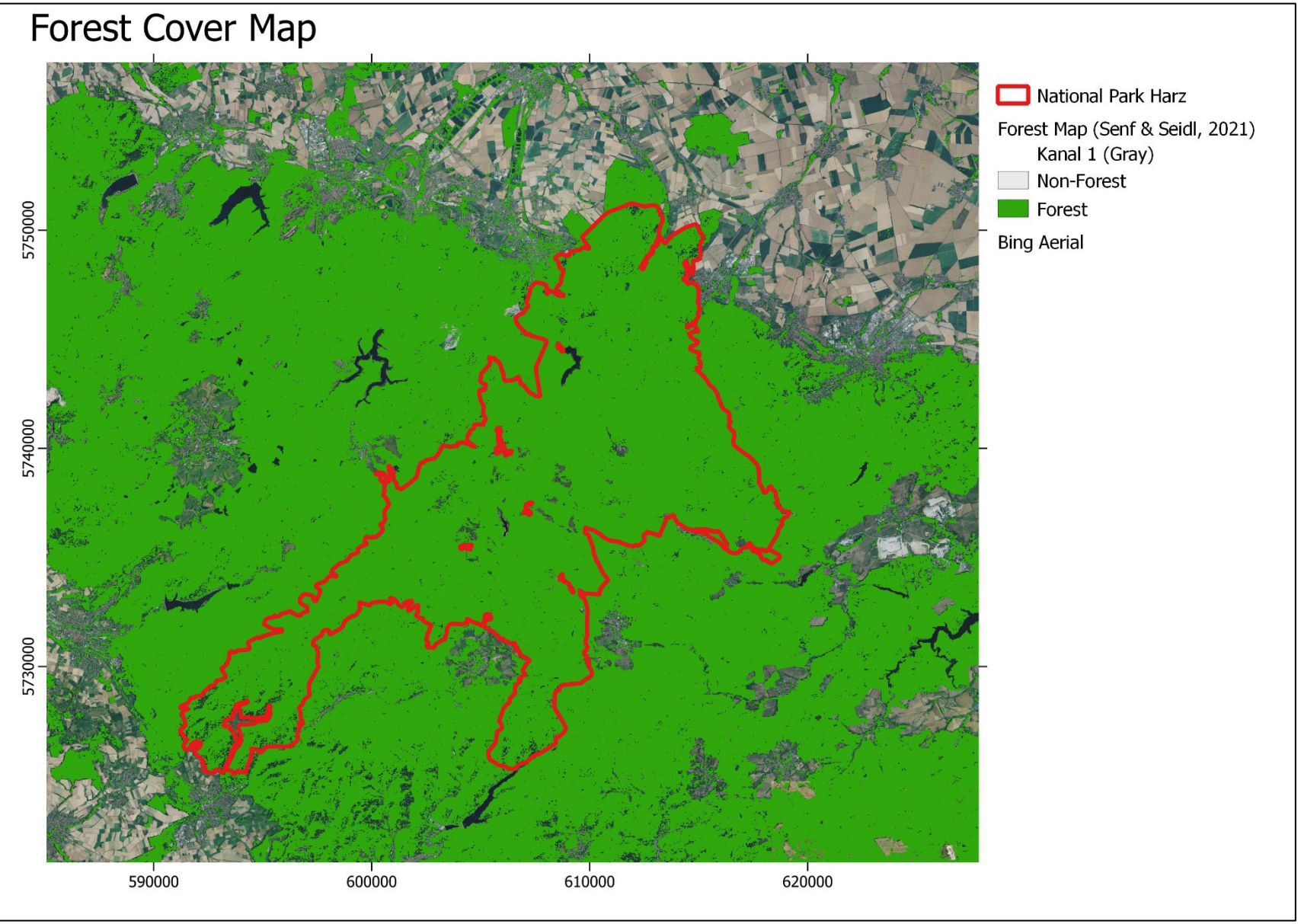
Content

- Why do we need reference data?
- What target variables are relevant in the context of forest monitoring?
- How can such reference data be measured in the field?
- Where to measure the reference data?
- What is the difference between the model and the map accuracy?
- Best practice guidelines for reference data collection and map validations

Tutorials:

1. Collecting reference data based on spectral variability
2. Design-based approaches for map validation of continuous maps

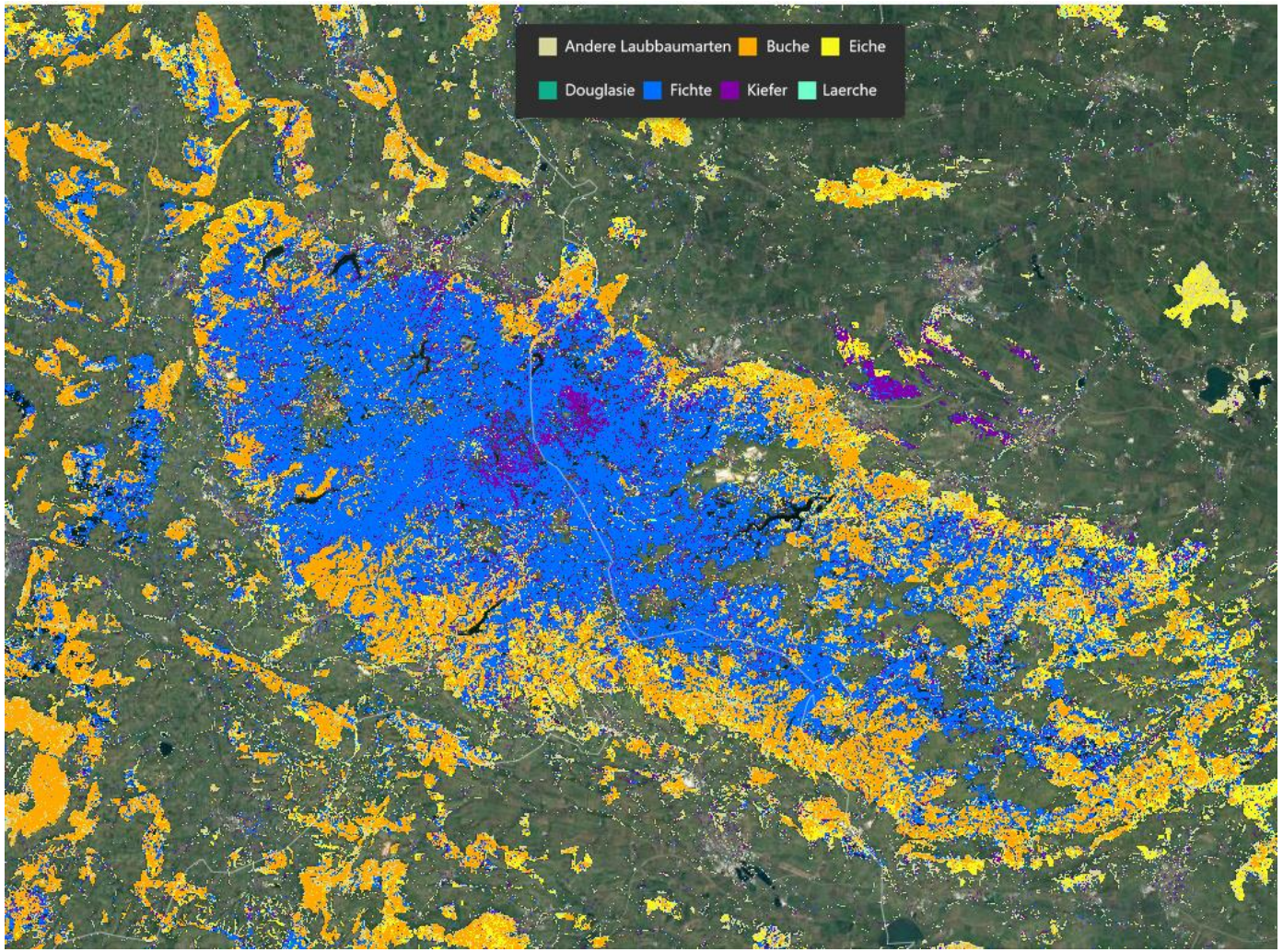
Examples of forest monitoring maps (Categorial)



- Forest cover maps are a basic forest monitoring product
- Creating a forest cover map requires a forest definition
- Depending on the forest definitions, high accuracies can be achieved for forest /non-forest classifications
- Important product for the monitoring of forest cover changes in the context of international environmental programs e.g. REDD+& CBD

Data source: Forest cover map based on time series from 1986-2018 (Senf & Seidl, 2021)

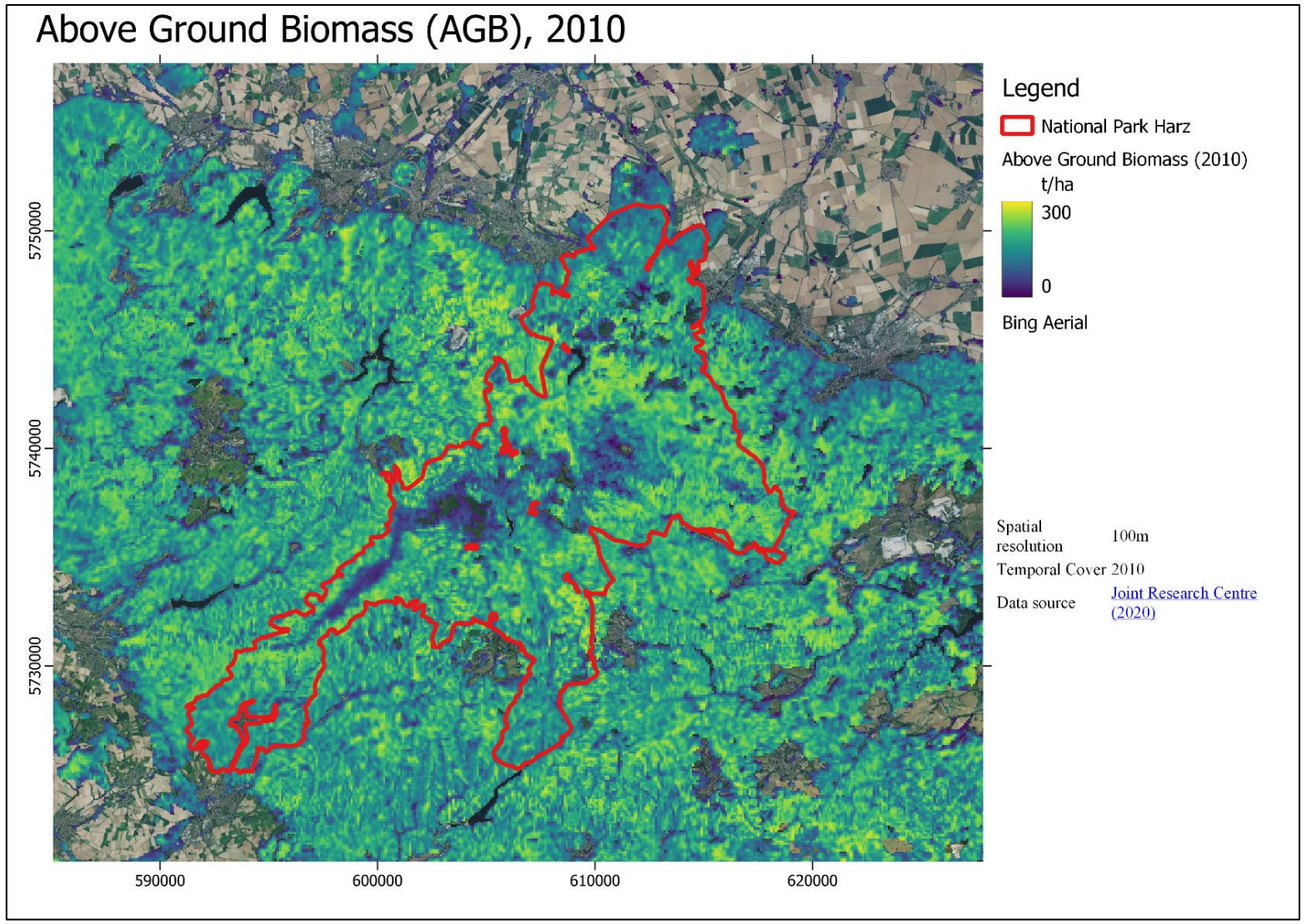
Tree species/Forest type maps



- Forest type and tree species maps are an important forest monitoring product
- Creating forest cover type maps requires a classification key of forest types
- Depending on the classification key/number of species or forest types the classification is quite challenging
- Tree species is one of the most studied variables in forest remote sensing applications
- Separating coniferous and broadleaf tree species often results in high accuracies
- Important product for forest management and nature conservation applications

Data source: Dominant tree species map for Germany, 2017 (Welle et al. 2022)

Examples of forest monitoring maps (Continuous)

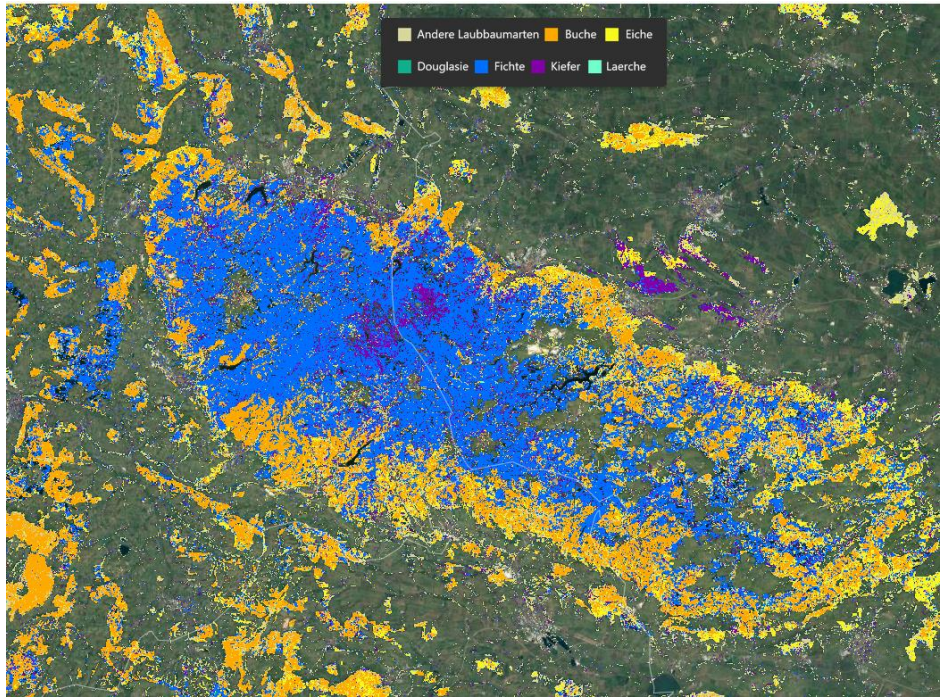


Data source: Above ground biomass, JRC(2020)

- Aboveground biomass provide valuable information on the carbon stock changes of forests -> Unit: t/ha
- Growing stock maps provide important information for forest managers
- Reporting Unit: m^3/ha
- Saturation effect for high biomass forests for optical sensors
- Important product in the context of results-based payment schemes for REDD

Introduction/Problem Statement

We are interested in maps depicting relevant environmental information.

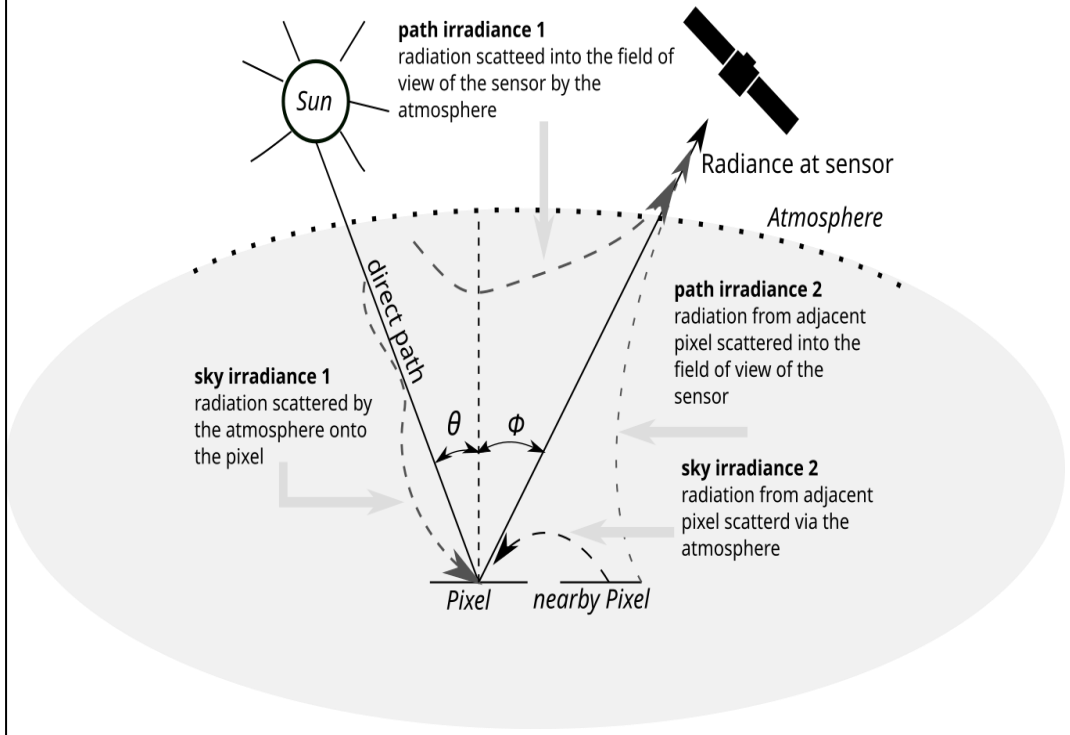


Source: RSS GmbH, https://www.remote-sensing-solutions.com/waldmonitor-deutschland/#baumartenkarte

MODEL

?

We observe reflected electromagnetic radiation.



Often, we cannot observe/measure the target variable directly but we need to predict the values using a model

Reference data

Training data

Used to build the classification model

Requirements:

- Should have high quality
- Need to cover all classes/the value range of the target variables to avoid extrapolation
- Should cover the feature space of the predictor variables

Validation data

Used to validate the model predictions

Requirements:

- Must be different from the training data
- Need to cover the variability of the area where it is applied (AoA)
- Need to be “representative” for the population

How to collect reference data in the context of forest monitoring?

Forest Inventories

How can we compile a reference data set for training a model?

- We need to have field observations at known locations
- But how can we observe/measure e. g. above ground biomass in the field?

Forest Inventory

Image: Magdon, P. (2014)

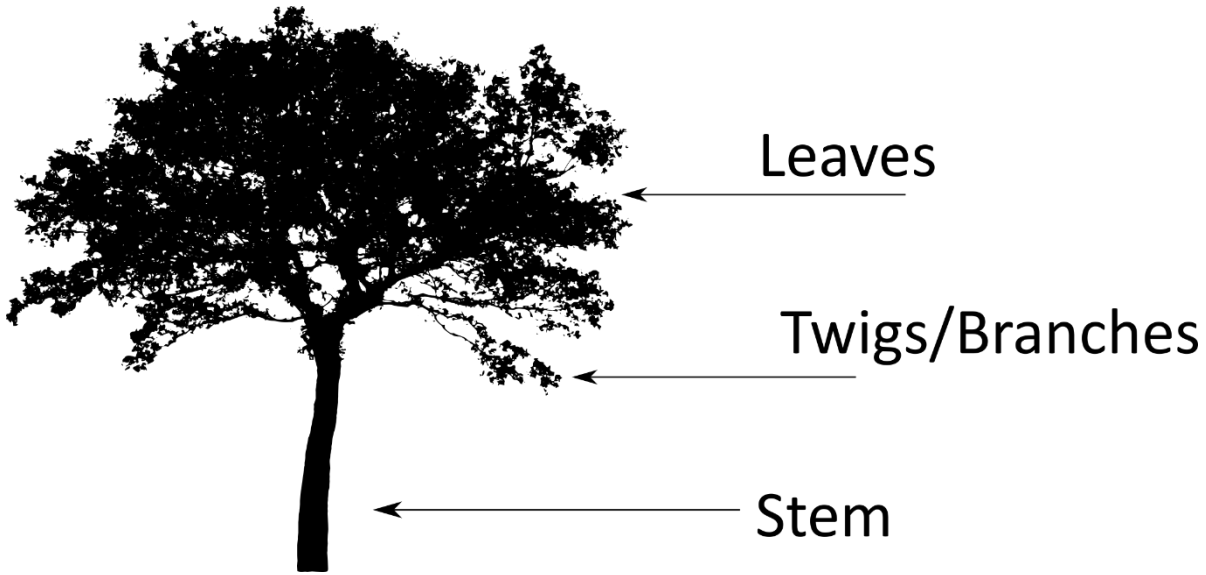
Scientific discipline:

- Forest mensuration, forest inventory and forest planning are key research areas in forest science and forest ecosystem analysis
- develop methods and models to collect quantitative information on forest ecosystems

Operational programs:

- Provide information on forest variables such as growing stock, biomass, forest cover or tree species
- Are the basis to implement and enforce sustainable forest management practices
- Support nature conservation programs by the provision of sound information on the effectiveness of conservation strategies

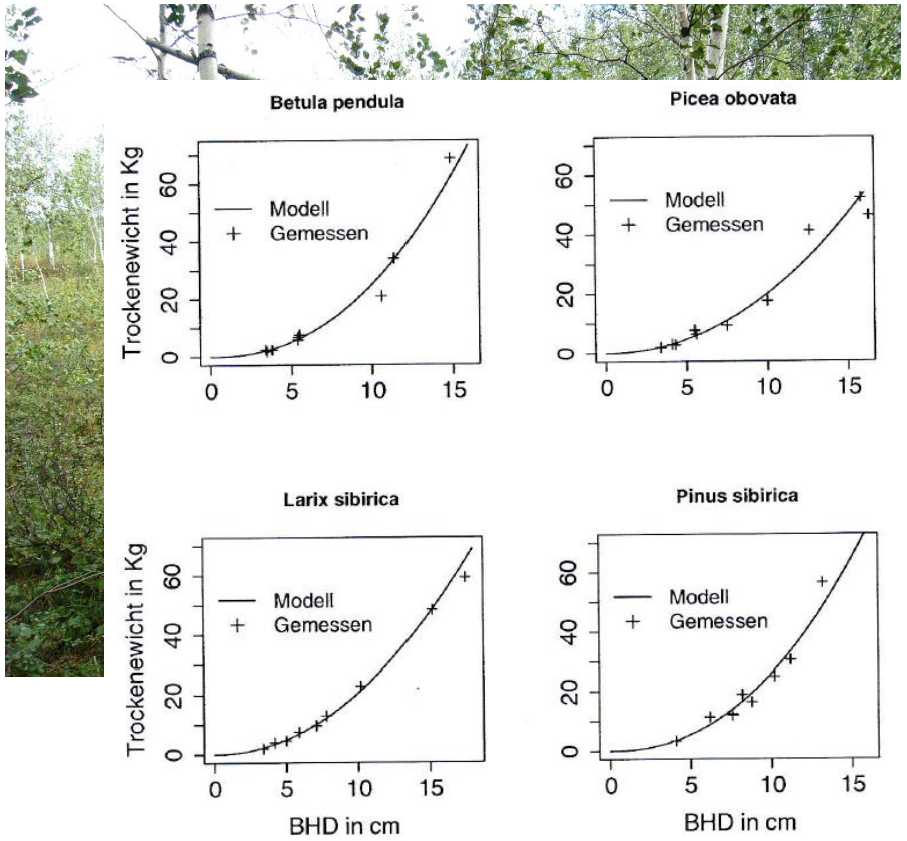
How to „measure“ above ground biomass (t/ha) in the field?



Models are used to predict the single tree biomass based on easy to measure variables like tree height and diameter at breast height (DBH):

- Allometric models, e. g. Chave et al. (2016)
- Conversion factors that convert stem volume to tree biomass

- Can only be measured destructively
- Very laborious and error prone

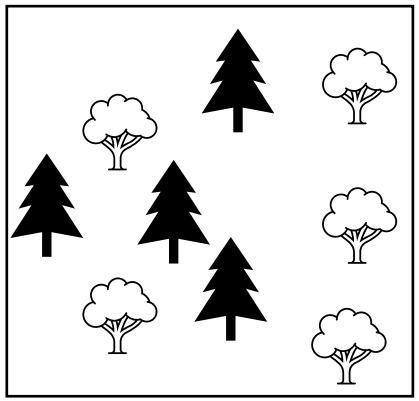


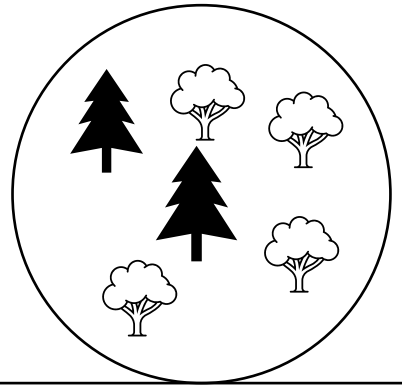
Source: Magdon, P. (2008)

How to select the trees to be measured?

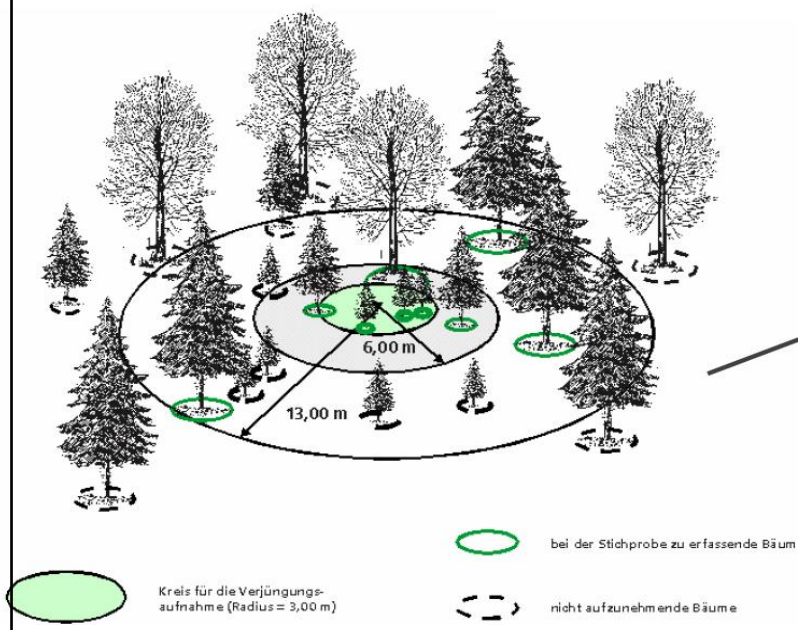
Plot design options:

Fixed area plots

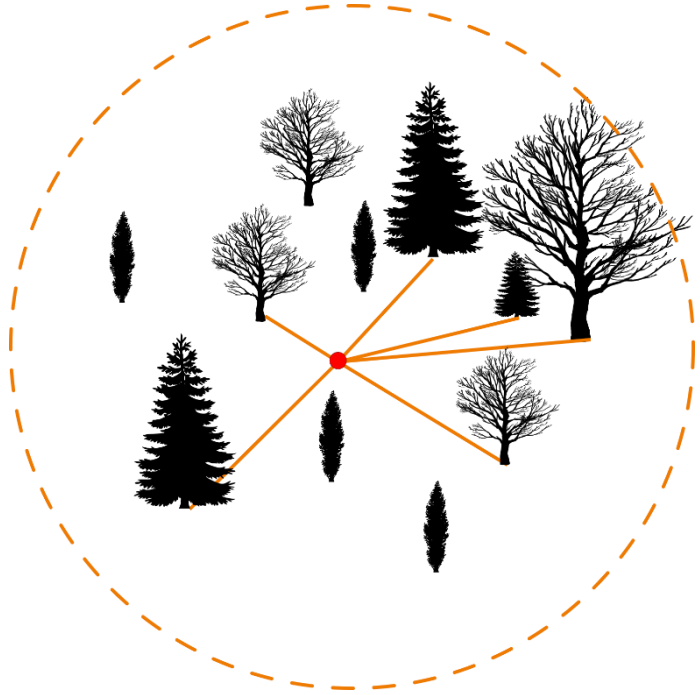




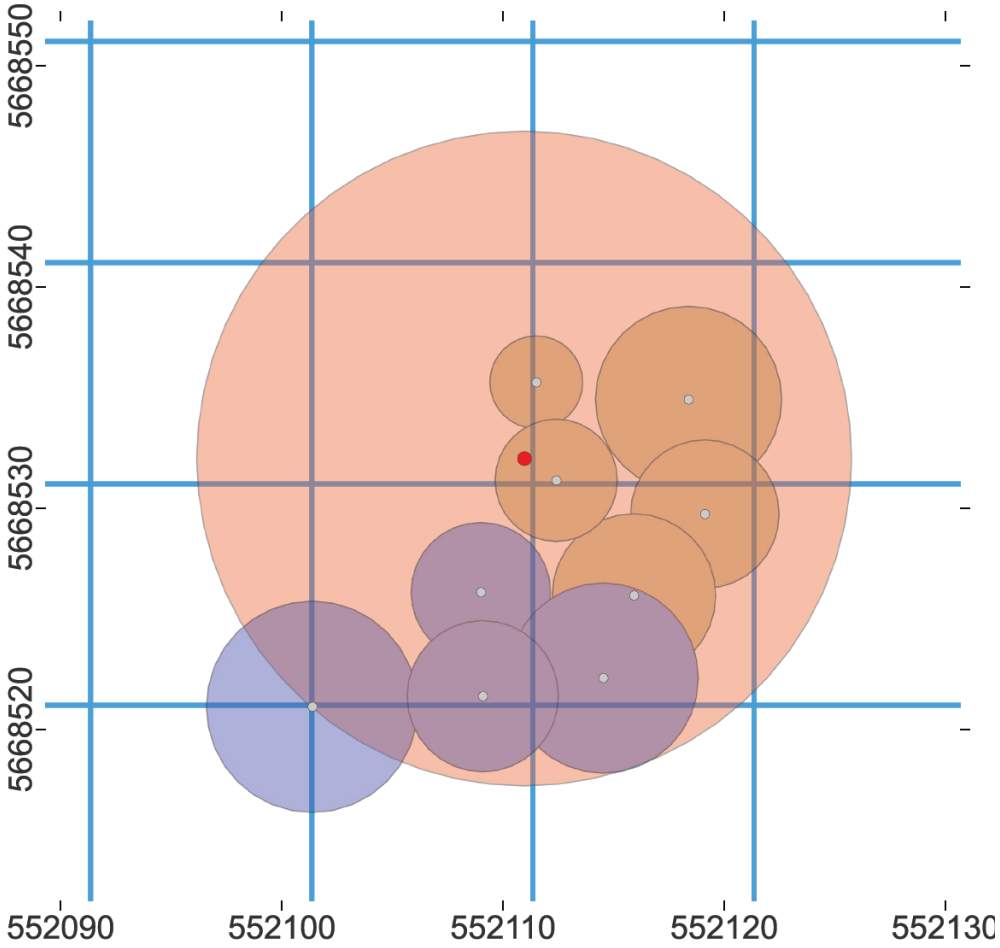
Concentric Plots



Angle count/ Bitterlich sampling



How to link the inventory information to the remote sensing data?



<u>Prerequisite:</u>
Location of the plot center needs to be known precisely -> differential GNSS devices need!

<u>Options:</u>

1. Extract all pixels that are within/touch the fixed area plot and assign a value of the target variable
2. Extract all pixels that are within/touch the fixed area plot calculate aggregated pixel values (*mean, median, sd, min, max, majority,* …) and assign the plot aggregate of the target variable
3. Calculate the aggregates an assign each pixel that is within/touches the fixed area plot similar values

Compiling a training data set

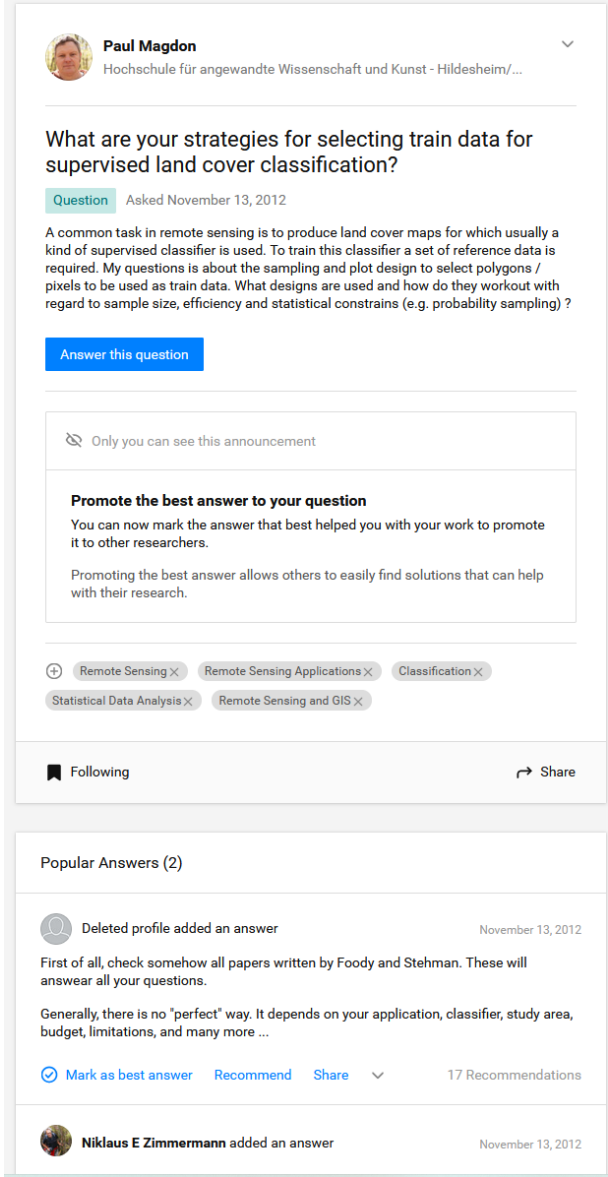
Where to collect data /Sampling strategies for training data collection?

Requirements towards the training data from a model building perspective:

- Should have high quality
- Need to cover the value range of the predictors
- Need to cover the value range of the target variables
- Should be collected efficiently

Strategies for selecting plot locations for training data collection:

- If information on the spatial distribution of the target variable is available e. g. from existing maps, we can locate plots in every class
- Without such information, we can group according to the pixel values e.g. with PCA & unsupervised clustering
- Additionally, we should check the range of the pixel values between training and image dataset



https://www.researchgate.net/post/What-are-your-strategies-for-selecting-train-data-for-supervised-land-cover-classification/1

References

Senf, C. and Seidl, R. (2021) Mapping the forest disturbance regimes of Europe. *Nature Sustainability*, 4, 63-70. https://doi.org/10.1038/s41893-020-00609-y

Welle, T.; Aschenbrenner, L.; Kuonath, K.; Kirmaier, S.; Franke, J. Mapping Dominant Tree Species of German Forests. *Remote Sens.* 2022, *14*, 3330. https://doi.org/10.3390/rs14143330

European Commission, Joint Research Centre (JRC) (2020): Forest Biomass Map of Europe. European Commission, Joint Research Centre (JRC) [Dataset] PID: http://data.europa.eu/89h/d1fdf7aa-df33-49af-b7d5-40d226ec0da3

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M.S., Delitti, W.B., Duque, A., Eid, T., Fearnside, P.M., Goodman, R.C., Henry, M., Martínez-Yrízar, A., Mugasha, W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A., Nogueira, E.M., Ortiz-Malavassi, E., Pélissier, R., Ploton, P., Ryan, C.M., Saldarriaga, J.G. and Vieilledent, G. (2014), Improved allometric models to estimate the aboveground biomass of tropical trees. Glob Change Biol, 20: 3177-3190. https://doi.org/10.1111/gcb.12629

Magdon, P. (2008): Erfassung und Analyse der räumlichen Variation von Kohlenstoffvorräten in einem sibirischen Wald-Tundra Ökosystem. , Masterthesis, Faculty of Forest Science and Forest Ecology, University of Göttingen, Germany.