UNIVERSITÄT TRIER

ForestPulse

Mapping tree species fractions in temperate mixed forests using Sentinel-2 time series and synthetically mixed training data

David Klehr[a], Johannes Stoffels[b], Andreas Hill[c], Vu-Dong Pham[d], Sebastian van der Linden[d], David Frantz[a]

[a] Geoinformatics - Spatial Data Science, Trier University, Universitätsring 15, 54296 Trier, Germany
[b] Environmental Remote Sensing and Geoinformatics, Trier University, Universitätsring 15, 54286 Trier, Germany
[c] Forest Planning and Inventory, Forest Service Rhineland Palatinate (Germany), Rhein-Mosel-Str. 9, 56281 Emmelshausen, Germany
[d] Earth Observation and Geoinformation Science Lab, University of Greifswald, Friedrich-Ludwig-Jahn-Str. 16, 17489 Greifswald, Germany
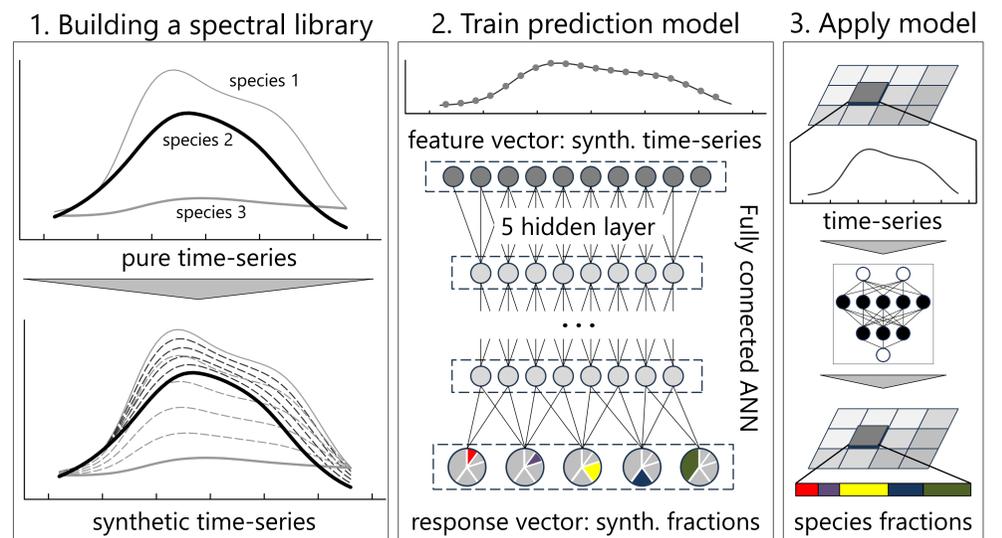
Motivation

Sustainable forest management and monitoring relies on **accurate information about tree species distribution**, which is difficult and costly to obtain through traditional field surveys. Remote sensing, particularly using dense Sentinel-2 time-series, offers a promising alternative by capturing **phenological differences between species**. Mapping mixed forest stands - especially including rare species - remains challenging due to the **limited availability of reference data**.

This study addresses this gap by exploring a regression-based unmixing approach (Okujeni et al. [1]) with synthetic data augmentation to map sub-pixel tree species mixtures using minimal reference information.
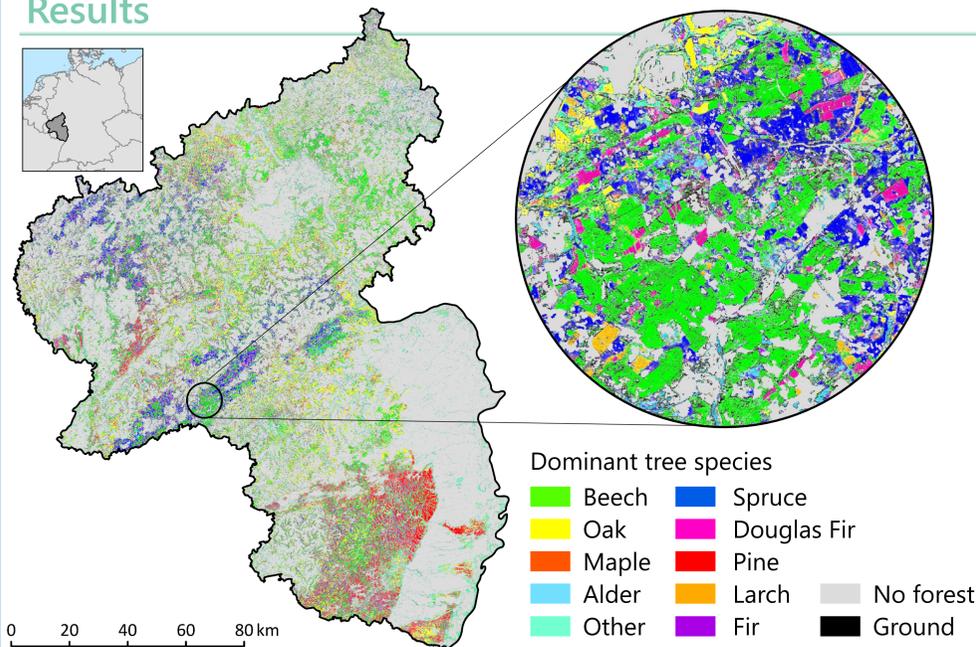
Data

In this study we utilized the **high spatial and temporal resolution** of **Sentinel-2** to reconstruct a gap-filled, **dense time-series** using a spline-based approach (Bolton et al. [2]) and user-defined function in FORCE [3].
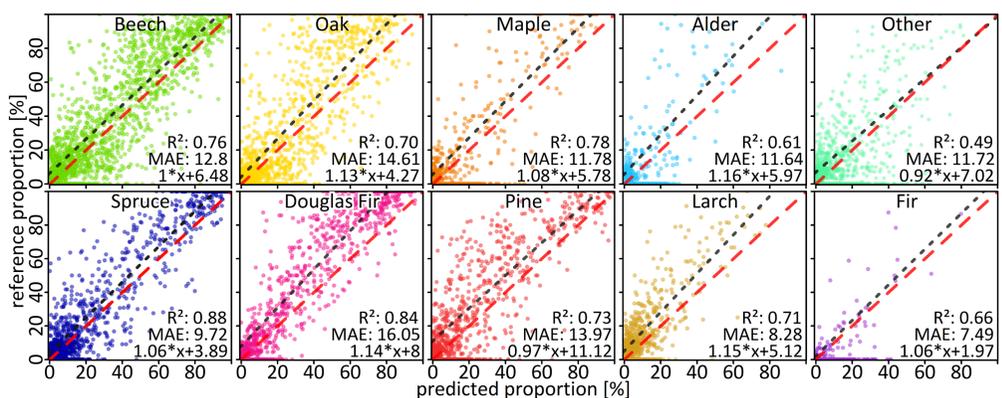
Methods

1. Building a spectral library
species 1
species 2
species 3
pure time-series

synthetic time-series

2. Train prediction model
feature vector: synth. time-series
5 hidden layer
Fully connected ANN
response vector: synth. fractions

3. Apply model
time-series
species fractions

Results

Dominant tree species
- Beech
- Oak
- Maple
- Alder
- Other
- Spruce
- Douglas Fir
- Pine
- Larch
- Fir
- No forest
- Ground

The unmixing was **performed and validated** for the federal state of Rhineland-Palatinate (Germany) as a proof-of-concept but has **potential for national application**.

The predicted species fractions are **validated** using polygon-based forest inventory data, which contain the **area occupied by each species** in every forest stand.

Proportions in **2,621 polygons** were validated, resulting in shown scatterplots.

Example validation inventory polygon

Beech: R²: 0.76, MAE: 12.8, 1*x+6.48
Oak: R²: 0.70, MAE: 14.61, 1.13*x+4.27
Maple: R²: 0.78, MAE: 11.78, 1.08*x+5.78
Alder: R²: 0.61, MAE: 11.64, 1.16*x+5.97
Other: R²: 0.49, MAE: 11.72, 0.92*x+7.02
Spruce: R²: 0.88, MAE: 9.72, 1.06*x+3.89
Douglas Fir: R²: 0.84, MAE: 16.05, 1.14*x+8
Pine: R²: 0.73, MAE: 13.97, 0.97*x+11.12
Larch: R²: 0.71, MAE: 8.28, 1.15*x+5.12
Fir: R²: 0.66, MAE: 7.49, 1.06*x+1.97

reference proportion [%]
predicted proportion [%]

Predicted proportions per species in all validation inventory polygons

Conclusion

The overreaching goal of this study was to **assess the potential of mapping sub-pixel tree species mixtures** in mixed temperate forests, with size-limitations of available reference information. Our proposed method shows:

- **Successful prediction** of tree species fractions on a stand forest stand level with **high accuracy**
- Polygon-based validation provides **robust results** but is not directly transferable to pixel level
- Ensemble approach **reduces model variability**, but is computationally intensive
- **Operationalization for forestry authorities** is possible due to a small required sample size and open data approach

Outlook

- **National-scale implementation** planned to test the method's scalability and operational potential across diverse forest regions.
- **Integration of NFI data** could improve species coverage and validation, enabling more reliable large-scale predictions.
- **Model refinement** aims to replace the ensemble with more efficient single-model architecture to reduce computational demands.

Sources

[1] Okujeni et al. 2013. Support vector regression and synthetically mixed training data for quantifying urban land cover. Remote Sensing of Environment 137, 184–197.

[2] Bolton et al. 2020. Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. Remote Sensing of Environment 240, 111685.

[3] Frantz 2019: FORCE - Landsat + Sentinel-2 Analysis Ready Data and Beyond. Remote Sensing 11.

GitHub

To the Publication

Download this poster

Contact:
M.Sc. David Klehr
klehr@uni-trier.de
https://www.researchgate.net/profile/David-Klehr