DE

EN

Wald5Dplus

An AI benchmark dataset for the combined spatial, spectral, polarimetric and temporal coverage of forest stands using Sentinel-1 & -2

In the **Wald5Dplus** project, a labelled reference dataset is generated for the use of AI methods in forest remote sensing. A newly developed method of image data fusion on **hypercomplex bases** [1] enables the space-saving, but information-preserving fusion of Sentinel-1 & -2 images in the spectral, polarimetric and temporal domain. In a recent review article [2], where existing approaches for image data fusion are categorized, the advantageous use of **hypercomplex bases** is especially highlighted. This fusion is demonstrated over the period of two years using three selected and representative forest areas, which have already been surveyed with high-resolution LiDAR and multispectral systems: *Bavarian Forest National Park*, *Kranzberger Forst near Freising*, and *Steigerwald* in Germany (Figure 1 and Table 1).

Table 1: Description of the study sites.

AOI	Geographic location	Area	Reference data
1) Bavarian Forest National Park	49° 15' N, 13° 15' E	25,000 *ha*	- Airborne LiDAR image including evaluation (2017)
2) Kranzberger Forst near Freising	48° 25' N, 11° 40' E	100 *ha*	- UAV LiDAR and multispectral data (2020) and preliminary study - UAV LiDAR and multispectral data (2023)
3) Steigerwald	49° 53' N, 10° 32' E	2,600 *ha*	- Airborne LiDAR image including evaluation (2015)

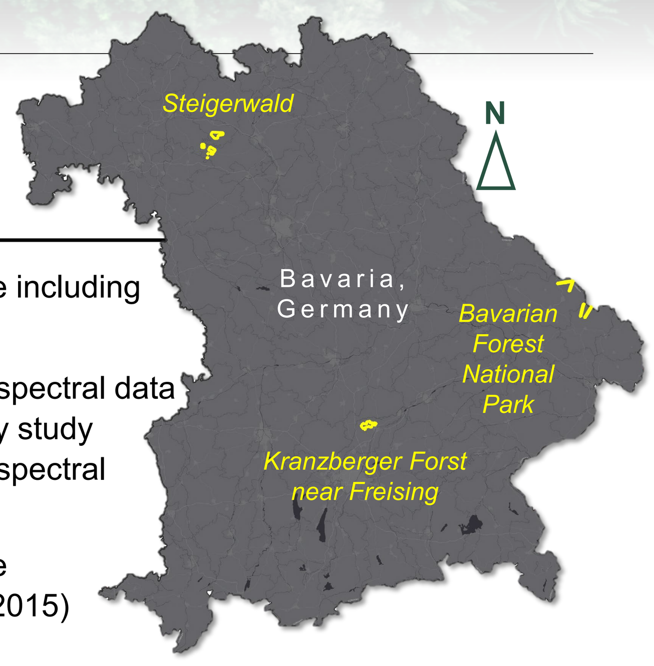

***Figure 1:** Location of the study sites.*

IMAGERY

WP 1 (Figure 2): A data stack of geocoded and radiometrically corrected *Sentinel-1 Single Look Complex (SLC)* and *Sentinel-2 MAJA* images is created over a one-year period: Sentinel-1 images are decomposed into polarimetric Kennaugh elements, geocoded onto a 10 m pixel grid, and radiometrically calibrated. These Kennaugh elements are comparable to those generated by the *MultiSAR* system at the German Aerospace Center [4]. Spectral Kennaugh elements are calculated from Sentinel-2, bringing both radar and optical systems onto a common basis [3]. These datasets can be further merged [1] or treated separately for comparability purposes. With a ground sampling distance of 10 m, sixty images per year are available. Two additional datasets before and after are added, resulting in a total of 64 images per year, optimal for calculating temporal Kennaugh elements. The resulting polarimetric or spectrometric temporal Kennaugh elements can be saved as *8-bit images* without significant information loss. This process is repeated for three study areas and two years each, resulting in six data batches, stored in Analysis Ready Data (ARD) Cubes. In Figure 3, a composition of the step-wise evolution, also in terms of information-gain, of the datasets till the polarimetric, spectrometric and temporal fusion is displayed featuring parts of AOI 2.

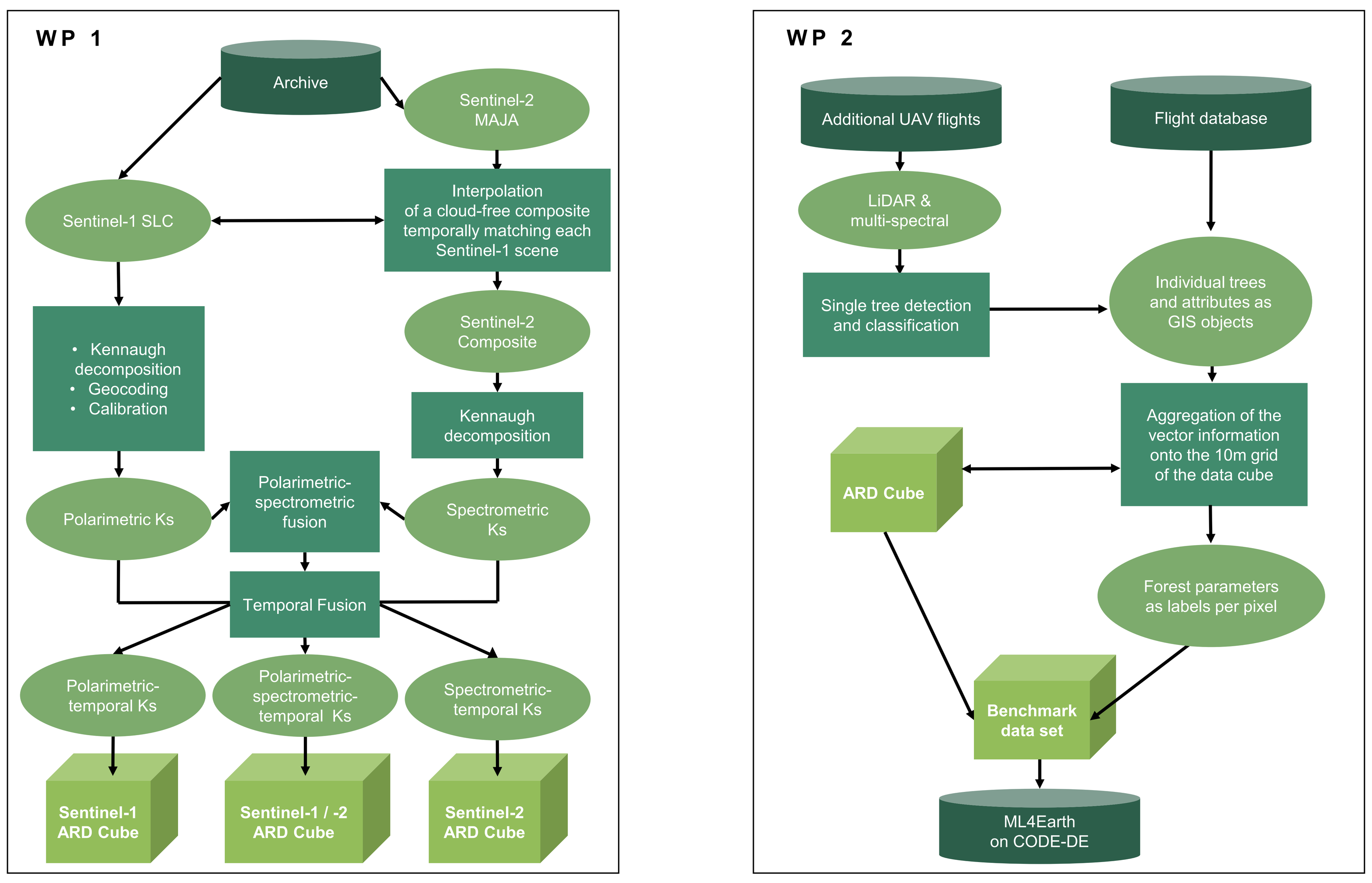

***Figure 2:** Description of the work packages, WP 1 and WP 2.*

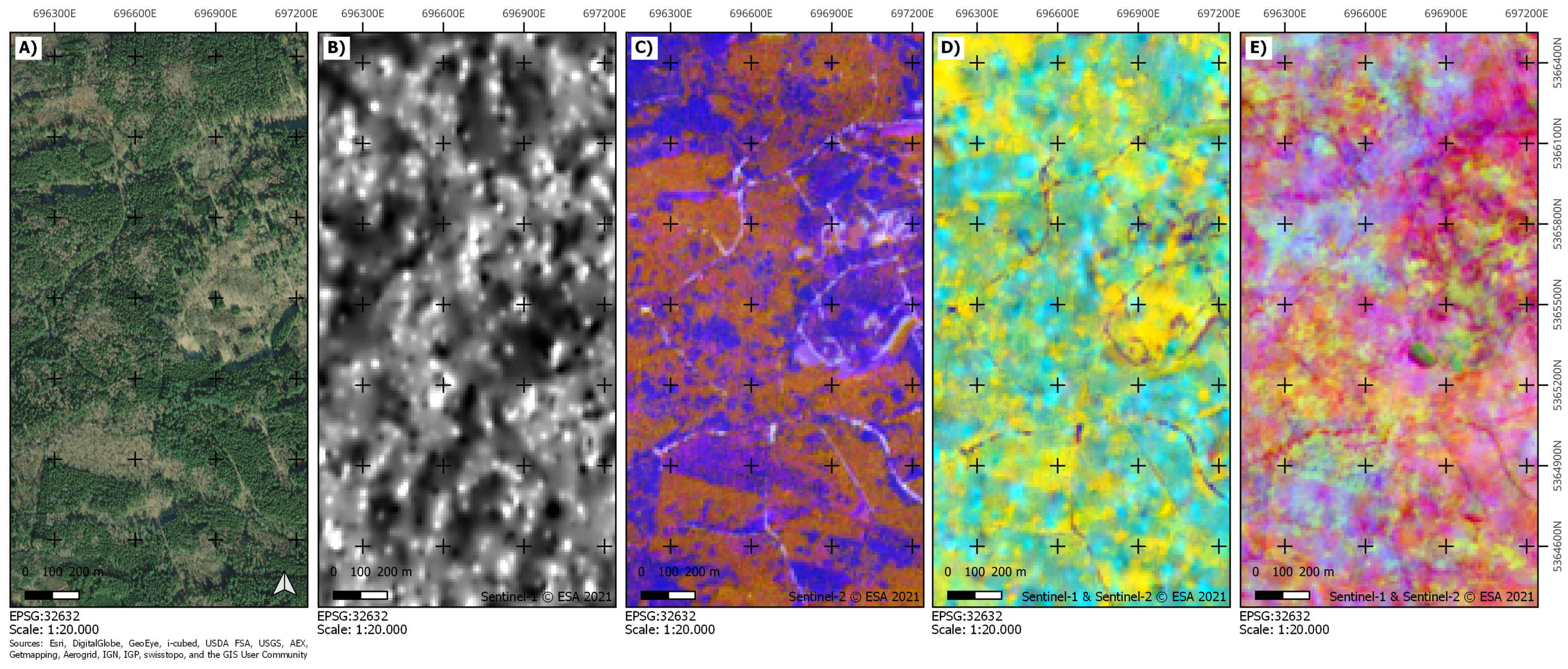

***Figure 3:** Composition displaying the step-wise evolution of the datasets in AOI 2, and for comparison purposes a satellite overview of the ground (A); Kennaugh element K0 of Sentinel-1 MultiSAR (2021-07-19) (B); Kennaugh elements of Sentinel-2 (2021-07-19) (C); Polarimetric and spectrometric fused dataset of Sentinel-1 & Sentinel-2 (2021-07-19) (D); and a polarimetric, spectrometric and temporally fused dataset over the whole period of 2021 (E).*

LABELS

WP 2 (Figure 2): A single tree detection method derives forest parameters such as tree type, crown base height, crown volume or height, and based on this, the stand composition and density etc., from a LiDAR point cloud and multispectral aerial acquisitions [5]. These tree parameters are aggregated onto the pixel grid of the fused satellite dataset using the average and percentage of the values. The parameters serve then as labels for the merged dataset. Thus, for each 10 m pixel in the Sentinel-1 stack, **256 polarimetric-temporal** and in the Sentinel-2 stack correspondingly **256 spectrometric-temporal Kennaugh elements** are available as features, which can be correlated by AI algorithms with the forest parameters. The labelled ARD Cube and the algorithms needed to create it as well as pre-trained AI classifiers will be made permanently available to the public via the ML4Earth project of the Technical University of Munich. In this way, even beginners in the application of machine learning methods are motivated to take their first steps. Furthermore, experts are encouraged to use the image fusion on **hypercomplex bases** for the preparation of their ARD Cubes.

REGRESSION

By combining multi-modal and multi-temporal satellite data, we can analyze vegetation changes and enhance our understanding of tree characteristics. Our UAV datasets provide precise reference information for tree attributes such as tree type, crown volume, tree height, and crown base height. To create a predictive model, we employ a random forest (RF) regression approach. The RF model correlates fused satellite imagery with the tree parameters derived from the UAV datasets. In AOI 1, over 350,000 trees with attributes are available [5]. This dataset is split into 70% training and 30% testing sets. The RF model is then trained on the training data and used to predict tree segment values for the testing data. Figure 4 displays the correlation of a polarimetrically, spectrometrically, and temporally fused Sentinel-1 and Sentinel-2 dataset over the period of one year (2021) for one subset in AOI 1. A feature importance test ranks individual features and bands based on their R^2 score. The robust correlation between tree parameters and the fused imagery is confirmed by an **R^2 score of 0.821** as demonstrated in Figure 4 (the horizontal axis shows the true values; the vertical axis denotes the predicted values; the red line marks the perfect fit). Onto the right of Figure 4, the feature importance, R^2 scores ranked per band. Additionally, Figure 5 displays an exemplary prediction of tree type (deciduous or coniferous) as well as the crown volume.

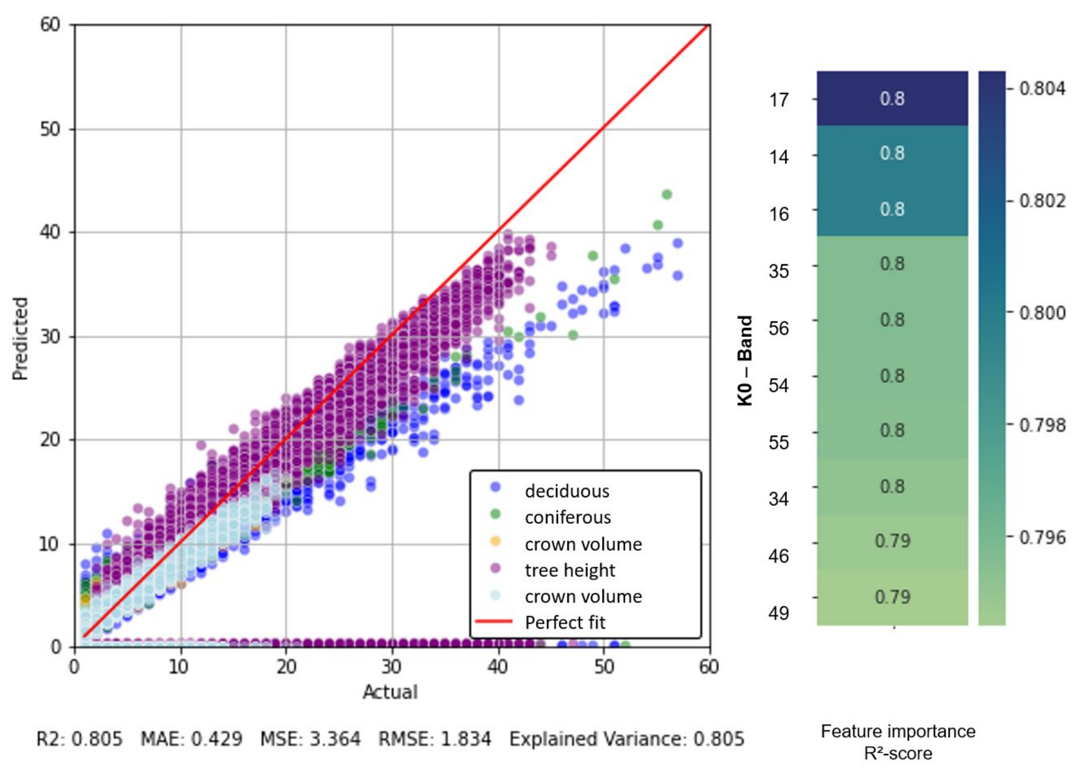

***Figure 4:** RF Regression of five forest parameters (l.) and a feature ranking of R² scores per Band (r.).*

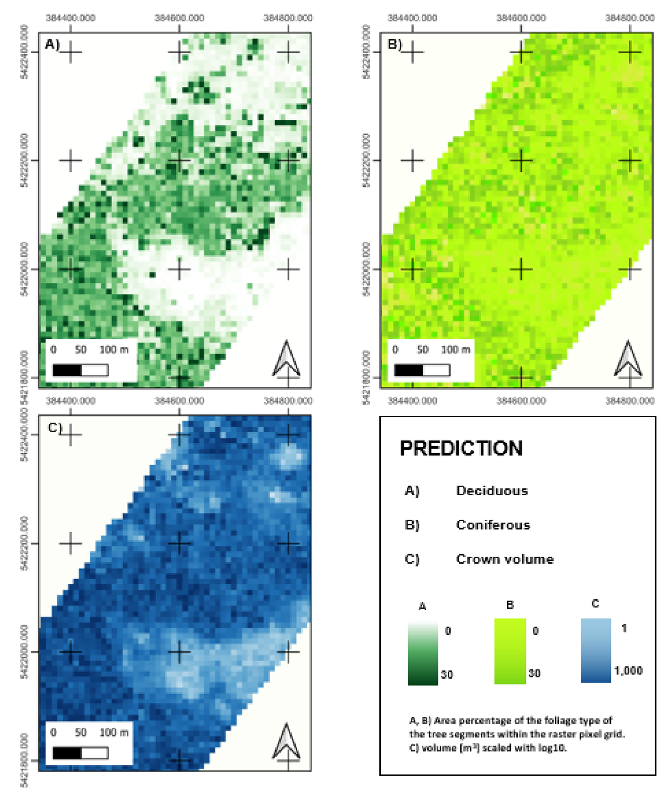

***Figure 5:** Prediction results of tree class and crown volume.*

SUMMARY

- Polarimetric, Spectrometric, and Temporal Kennaugh elements from merged Sentinel-1 and -2 data
- Single tree detection and classification from airborne LiDAR and multispectral cameras
- Analysis Ready Data Cubes with forest parameters as labels
- Expected availability by the end of 2023
- Provision of the benchmark data set and the algorithms via the ML4Earth platform

LITERATURE

[1] Schmitt, A.; Wendleder, A.; Kleynmans, R.; Hell, M.; Roth, A. and Hinz, S. (2020): Multi-Source and Multi-Temporal Image Fusion on Hypercomplex Bases. Remote Sensing, vol.12, pp. 943. DOI: 10.3390/rs12060943

[2] Zangl, R.; Hauser, S. and Schmitt, A. (2022): Guidelines for the Practical Use of Image Data Fusion in Remote Sensing, gis.Science, vol. 4, p.123-147.

[3] Schmitt, A. and Wendleder, A. (2018): SAR-sharpening in the Kennaugh framework applied to the fusion of multi-modal SAR and Optical Images. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., vol. IV-1, p. 133–140, DOI:10.5194/isprs-annals-IV-1-133-2018.

[4] Schmitt, A.; Wendleder, A; and Hinz, S. (2015): The Kennaugh element framework for multi-scale, multi-polarized, multi-temporal and multi-frequency SAR image preparation", ISPRS Journal of Photogrammetry and Remote Sensing, vol. 102, p. 122-139.

[5] Krzystek, P.; Serebryanyk, A.; Schnörr, C.; Červenka, J. and Heurich, M. (2020): Large-Scale Mapping of Tree Species and Dead Trees in Šumava National Park and Bavarian Forest National Park Using Lidar and Multispectral Imagery", *Remote Sens.*, vol. 12, pp. 661.

AUTHORS

Sarah Hauser M. Sc. (sarah.hauser@hm.edu)

Prof. Dr.-Ing. Andreas Schmitt (andreas.schmitt@hm.edu)

Prof. Dr.-Ing. Peter Krzystek (peter.krzystek@hm.edu)

HM

IAMLIS

Institut für Anwendungen des maschinellen Lernens und intelligenter Systeme

Gefördert durch:

Bundesministerium für Wirtschaft und Energie

aufgrund eines Beschlusses des Deutschen Bundestages

Neue Perspektiven der Erdbeobachtung