

BENCHMARK DATASET

For field phenotyping

Author: Lars Grygosch
l.grygosch@fz-juelich.de

This work was created as part of the NFDI consortium FAIRagro (www.fairagro.net). We gratefully acknowledge the financial support of the German Research Foundation (DFG) – project number 501899475.

Version 1.0 | Date: 20.12.2024

Table of Contents

| | |
|-------------------------------------|---|
| 1 Introduction..... | 1 |
| 2 Dataset Description..... | 2 |
| 2.1 HyPlant | 2 |
| 2.2 PhenoRob..... | 2 |
| 2.3 UAV Campus Klein-Altendorf..... | 3 |
| 2.4 BreedFACE..... | 3 |
| 3 Benchmark Metadata | 3 |

1 Introduction

Climate change, through shifts in temperature and precipitation patterns and more frequent extreme weather events, poses significant challenges to global crop production and food security (Davis, 2021). Understanding how plants respond to these changing conditions requires robust data and insights into both structural and functional traits of crops. Field phenotyping, which studies plant traits under realistic conditions, plays a critical role in addressing these challenges by leveraging advanced technologies such as spectral sensors, remote sensing, and machine learning. These advancements enable non-destructive, scalable measurements across spatial and temporal scales, providing insights into crop growth, photosynthetic processes, and yield potential.

However, the increasing complexity and volume of field phenotyping data demand solutions that address challenges in data standardization, integration, and metadata management. To meet this need, we present a benchmark dataset collection as part of the FAIRagro Consortium initiative. This benchmark adheres to FAIR principles—Findable, Accessible, Interoperable, and Reusable (Wilkinson et al., 2016)—and integrates diverse datasets enriched with detailed metadata. Spanning multiple temporal and spatial resolutions, sensor technologies, and research objectives, these datasets offer a comprehensive representation of the field phenotyping domain.

This benchmark aims to address data heterogeneity and support innovation in field phenotyping research by providing a valuable resource for understanding crop responses under dynamic environmental conditions. The following sections introduce the datasets and their contributions to advancing agricultural research. Subheading 1

[illegible]

2 Dataset Description

2.1 HyPlant

Author: Buffat, Jim (IBG-2: Plant Sciences, Forschungszentrum Jülich)

The dataset comprises solar-induced fluorescence (SIF) and associated radiance measurements collected between 2018 and 2023. It integrates data from the HyPlant airborne sensor system and the FLOX top-of-canopy reflectance systems, focusing on the retrieval of SIF and its relationship with plant photosynthetic activity. In addition to raw measurements, the dataset includes derived products generated using spectral fitting methods (SFM) and a neural network-based SFM model (SFMNN), facilitating advanced computational approaches to SIF retrieval.

The dataset has a total size of 1.6 TB, comprising 3,748 files across 16 distinct file format types.

- Data storage: [HyPlant Data](<https://datapub.fz-juelich.de/plantsciences/HyData-HyPlant-FLUO>)
- Metadata: [HyPlant Metadata](<https://doi.org/10.26165/JUELICH-DATA/QKJKPW>), including an enhanced MIAPPE-based .xlsx metadata file.
- ARC structure version: [ARC Structure](<https://git.nfdi4plants.org/fairagro-dataplant-deepdive/HyPlant>)

The dataset is published under the CC-BY 4.0 - "Attribution" license. Subheading 2

Standard text Standard text Standard text Standard text Standard text Standard text Standard text Standard text Standard text Standard text Standard text Standard text Standard text Standard text

2.2 PhenoRob

Author: Kraemer, Julie (IBG-2: Plant Sciences, Forschungszentrum Jülich)

The dataset, titled "Multi-scale field phenotyping of wheat-bean intercrops: Integrating spectral and agronomic datasets from a three-year trial", was collected during a spring wheat–faba bean intercropping experiment conducted at the Campus-Klein Altendorf (CKA) field site from 2021 to 2023. The trials were implemented using organic farming practices in a completely randomized design, comparing intercropping systems of three spring wheat cultivars with two faba bean cultivars and their respective sole crop controls.

The dataset integrates a heterogeneous collection of remote sensing and agronomic data, encompassing 6.3 GB across multiple file formats. Remote sensing parameters include high-resolution imagery, leaf and canopy reflectance, and both active and solar-induced fluorescence measurements, providing insights into vegetation performance and plant photosynthetic activity. Traditional agronomic measurements, such as biomass, leaf area, chlorophyll content, and grain yield, complement the remote sensing data to offer a comprehensive view of plant growth and productivity in intercropping systems.

The dataset is openly accessible under the following storage and metadata repositories:

- Data storage: [PhenoRob Data](<https://phenoroam.phenorob.de/geonetwork/srv/eng/catalog.search#/metadata/2388c9c0-7baa-4bf3-b258-b848143b2d0e>)
- Metadata: [PhenoRob Metadata](<https://data.fz-juelich.de/dataset.xhtml?persistentId=doi:10.26165/JUELICH-DATA/XWRNHR>)

The dataset is published under the CC-BY 4.0 - "Attribution" license. Subheading 3

2.3 UAV Campus Klein-Altendorf

Author: Warstat, Kevin (IBG-2: Plant Sciences, Forschungszentrum Jülich)

The dataset consists of RGB and multispectral UAV measurements of the PhenoRob central experiment conducted at Campus Klein-Altendorf in 2023. It contains raw image data, products derived using the structure-from-motion method, and additional materials such as README files, ground control point coordinates, and processing reports. The data is divided into 13 individual flight days throughout the vegetative period, with each day containing all necessary data to reproduce the resulting models. Key products include a digital elevation model (DEM) derived from RGB data and two orthomosaics, one with reflectance and the other with radiance information.

The dataset has a total size of approximately 309 GB.

- Data storage: [UAV Data](https://datapub.fz-juelich.de/plantsciences/PhenoRob_UAV_Data_2023/)
- Metadata: [UAV Metadata](<https://doi.org/10.26165/JUELICH-DATA/2BYDTH>)

The dataset is published under the CC-BY 4.0 - "Attribution" license, enabling open access and reuse with proper citation.

2.4 BreedFACE

Author: Knopf, Oliver (IBG-2: Plant Sciences, Forschungszentrum Jülich)

The dataset contains results from an experiment conducted as part of the BigBaking project, in which 10 wheat species were analyzed under an artificially increased CO₂ concentration. Measurements were carried out using advanced sensor systems, including the Light-Induced Fluorescence Transient (LIFT) device mounted on an autonomous platform (FieldSnake), as well as RGB and multispectral UAV data. Additional environmental sensor data is also included.

The dataset has a total size of approximately 4.2 MB and includes 19 `.csv` and 8 `.xlsx` files.

- Data storage and metadata: [BreedFACE Data and Metadata](<https://doi.org/10.26165/JUELICH-DATA/QJIY7C>)

The dataset is published under the CC0 - "Public Domain Dedication" license, enabling open access and reuse with proper citation.

3 Benchmark Metadata

All Datasets are linked in on DOI: <https://doi.org/10.26165/JUELICH-DATA/3F6OFJ>