Universität Zürich^UZH

Department of Geography

Using High-Resolution Satellite Data to Estimate Canopy Cover in Plot Experiments

ESS 511 Master's Thesis

Author: Timon Boos, 19-707-082

Supervised by: Dr. Helge Aasen (helge.aasen@agroscope.admin.ch)

Faculty representative: Prof. Dr. Alexander Damm

27.04.2025

ESS 511

Master's thesis

Zurich, April 25, 2025

Using High-Resolution Satellite Data to Estimate Canopy Cover in Plot Experiments

Author
Timon Boos
University of Zurich, 19-707-082

Supervised by
Dr. Helge Aasen
Earth Observation of Agroecosystems, Agroscope

Faculty representative
Prof. Dr. Alexander Damm
Remote Sensing Laboratories, Department of Geography, University of Zurich

ABSTRACT

Understanding and monitoring crop growth is crucial for addressing global food security challenges and promoting sustainable agricultural practices. Canopy cover (CC) is a key indicator of crop development, particularly during early growth stages when traditional metrics like the Leaf Area Index (LAI) are less reliable. While Unmanned Aerial Vehicles (UAVs) offer high-resolution phenotyping capabilities, their limited spatial and temporal coverage restricts scalability. In contrast, conventional satellite platforms provide broader coverage but lack the spatial and temporal resolution required for continuous plot-level monitoring. Recent advancements in satellite technology, such as the PlanetLabs SuperDove (PSB.SD) constellation, provide daily imagery at a 3-meter resolution. Despite their high spatial detail, these images face challenges related to radiometric stability, spatial co-registration accuracy, and quality masks. Addressing these limitations, this research investigates the use of such high-resolution multispectral satellite data from PSB.SD, integrated with UAV-derived RGB imagery for calibration, to bridge the gap between fine-scale UAV-based phenotyping and scalable satellite-based monitoring.

Across different agricultural sites in Switzerland, UAV-derived RGB data was acquired and processed using a two-stage segmentation method to effectively separate vegetated from non-vegetated pixels. These reference CC masks were then integrated with radiometrically and spatially preprocessed PSB.SD multispectral data to develop a training dataset for satellite-based CC models. Several modeling approaches were tested, including linear regression models based on vegetation indices (VIs), machine learning methods such as Random Forest (RF) and Histogram-Based Gradient Boosting Regressor (HGBR), and a feedforward neural network (NN). The results demonstrated strong model performance, with MSAVI (R^2 = 0.861), HGBR (R^2 = 0.919), and NN (R^2 = 0.923) showing high predictive accuracy across different crop species, sites, and growth stages. Cross-validation indicated that while linear VI-based models such as MSAVI perform adequately under data-scarce conditions, more complex machine learning and deep learning models yield superior performance when high-quality data are available. The models were successfully applied to a winter barley field experiment with varying fertilizer treatments, enabling time-series CC estimation across multiple 9 m wide plots. This application demonstrated the models' ability to capture treatment-specific canopy development, including early growth stage CC differences.

This research demonstrates the feasibility of using high-resolution multispectral satellite data for continuous, scalable, and non-invasive crop monitoring, supporting more efficient and sustainable agricultural practices. In light of these findings, future work should focus on expanding the training dataset to include a broader range of crops and farming systems, enhancing model generalizability. Incorporating crop-specific models and utilizing open data sources can enhance prediction accuracy and extend global applicability, advancing the development of satellite-based CC monitoring for sustainable agriculture.

DECLARATION ON AI USAGE

In the course of writing this thesis, I have utilized Artificial Intelligence (AI) tools, specifically ChatGPT 4.0 and Grammarly (as of April 25, 2025), for various aspects of the work. The AI was used as a valuable assistant in the following areas:

- **Text Editing**: ChatGPT was employed for editing text, including making suggestions for improvements, shortening sections, checking grammar, and refining phrasing and formulation. This assistance helped ensure clarity and comprehension of the thesis. Additionally, Grammarly was used to further refine grammar and punctuation, helping to ensure coherence and overall linguistic accuracy.
- **Coding Assistance**: ChatGPT was also utilized to support coding tasks. It helped explain complex code snippets from external sources, provided advice on debugging, and assisted in improving my own code. This proved to be a valuable tool in increasing efficiency and minimizing potential errors in the workflow.

In all cases, the AI was used as a supplementary tool to enhance my work and ensure the accuracy and quality of the content. The intellectual contributions and interpretations of the research are entirely my own, with the use of AI being strictly for technical support.
I hereby declare that the use of AI has been conducted in accordance with ethical guidelines and with full acknowledgment of the limitations of AI in academic research.

Timon Boos
25.04.2025

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Helge Aasen, my supervisor, for giving me the opportunity to undertake this thesis at Agroscope. I truly appreciate his enthusiasm and expertise for Earth Observation. His clear guidance, ongoing support, and valuable feedback played a key role in shaping both the direction and quality of my work.

I would also like to thank Prof. Dr. Alexander Damm, the faculty representative, for his valuable advice, constructive feedback, and prompt responses, which were greatly beneficial to this thesis.

My thanks extend to the EO Team at Agroscope for creating a collaborative and supportive environment. Their shared commitment to sustainable agriculture and the knowledge they provided were instrumental throughout my project. I would also like to acknowledge Séléne Ledain for her assistance with both administrative and technical matters. Additionally, I appreciate the collaboration and exchange of ideas with Matthias Diener from Agroscope on the ”Bio4Recycle” experiment, which enriched my research.

I am deeply grateful to Safiya Schwarz for her careful proofreading of my thesis, which helped ensure grammatical accuracy and improved its overall clarity.

On a personal note, I would like to thank my family and friends for their support, patience, and encouragement throughout my academic journey. Finally, I wish to acknowledge my peers, whose friendship and shared experiences contributed to a rewarding time at the University of Zurich.

ABBREVIATIONS

ANN	Artificial Neural Network
BRDF	Bidirectional Reflectance Distribution Function
CC	Canopy Cover
R^2	Coefficient of Determination
DL	Deep Learning
DEM	Digital Elevation Model
ELC	Empirical Line Correction
EWS	Eschikon Wheat Segmentation
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization
GSD	Ground Sampling Distance
HGBR	Histogram Gradient Boosting Regression
IoU	Intersection over Union
LAI	Leaf Area Index
ML	Machine Learning
MSE	Mean Squared Error
MSAVI	Modified Soil-Adjusted Vegetation Index
NBR2	Normalized Burn Ratio 2
NIR	Near-Infrared
nRMSE	Normalized Root Mean Squared Error
PS	PlanetScope
PSB.SD	PlanetScope SuperDove
PITs	Pseudo-Invariant Targets
RTMs	Radiative Transfer Models
RTK	Real-Time Kinematic
NRMSE	Range-Normalized RMSE
RF	Random Forest
RS	Remote Sensing
RMSE	Root Mean Squared Error
RGB	Red, Green, Blue
QA	Quality Assurance
SAVI	Soil-Adjusted Vegetation Index
SNP	Snapshot
SVM	Support Vector Machines
UAV	Unmanned Aerial Vehicle
VI	Vegetation Index

Contents

List of Figures

List of Tables

1 INTRODUCTION

1.1 The Need for Field-Based Phenotyping in Modern Agriculture

The increasing world population and climate change place immense pressure on the agriculture sector to become more efficient and sustainable. According to the Food and Agriculture Organization (FAO), agricultural production must increase by 70% by 2050 to meet the growing demand for sustainable, high-quality nutrition (Martos et al., 2021). To address these food safety and security challenges, continuous monitoring of crop varieties in phenotyping experiments is essential for tracking growth progression and understanding how distinct genotypes respond to environmental changes (Cobb et al., 2013; Saint Pierre et al., 2012).

While many phenotyping experiments are conducted under controlled environmental conditions in greenhouses, the insights gained in these settings often fail to translate effectively to real-world field applications (Araus and Cairns, 2014). This is primarily because controlled environments do not capture the full complexity of field conditions, where soil variability, climate fluctuations, and biotic/abiotic stressors interact and influence plant development (White et al., 2012). Therefore, field phenotyping, here defined as the quantitative assessment of a plant's phenotype under field conditions, has become increasingly important in modern breeding programs (Yang et al., 2017). It enables the accurate evaluation of trait expression in realistic environments, making it a key tool for crop improvement (Araus and Cairns, 2014; Yang et al., 2017). However, accurately and continuously measuring plant physiological responses to management practices in field conditions remains a substantial challenge, since these measurements are time-consuming, and often limited in spatial and temporal information content (Sankaran et al., 2015).

1.2 Remote Sensing Technologies for Scalable Crop Monitoring

To overcome these spatial and temporal limitations, the efficient use of remote sensing (RS) technologies has emerged as a promising solution to improve sustainable food production (Mulla, 2013; Martos et al., 2021). RS instruments mounted on satellites or unmanned aerial vehicles (UAVs) provide critical data on soil conditions, environmental factors, and crop health and development (Martos et al., 2021). High precision data from UAVs and frequently updated insights from satellites already play a pivotal role in precision agriculture by optimizing resource use and enhancing yields (Sishodia et al., 2020). Walter et al. (2015) demonstrated that image-based phenotyping techniques, traditionally employed in controlled environments, can be successfully adapted to field conditions, thereby enabling comprehensive characterization of plant performance. Similarly, Haghighattalab et al. (2016) demonstrated that low-cost UAVs can efficiently capture accurate, high-resolution phenotypic data, underscoring their potential for large-scale crop analysis under field conditions. This underscores the growing relevance of modern precision agriculture technologies in field phenotyping and their contribution to the advancement of sustainable agriculture (Zhang et al., 2002; Delgado et al., 2019).

Utilizing satellite-based RS for field phenotyping could offer a non-destructive and efficient method to gather data over vast geographical areas systematically, minimizing sampling bias and providing consistent, comparable information (Liaghat and Balasundram, 2010). However, for agricultural application with small field sizes, the coarse spatial ($> 10\,m$) and limited temporal resolution ($> 5\ days$) of traditional satellite imagery is insufficient to differentiate between field parcels and small-scaled spatial patterns (Wildhaber et al., 2023; Zhang et al., 2020). Satellite constellations with high spatial and temporal resolution have been limited in terms of spectral information, offering fewer and lower-quality spectral channels compared to Sentinel-2 or Landsat. Yet, the recent introduction of Super-Dove CubeSats (PSB.SD) from Planet Labs PBC (2024) has addressed this limitation. Since 2020, these new low-orbiting satellites provide multispectral data with eight reflective channels at 3 m spatial resolution. Additionally, the novel bands cover the Red Edge spectral region, which is critical for vegetation studies (Zhang et al., 2020). Despite their high spatial detail, PSB.SD images face challenges related to radiometric stability, spatial co-registration accuracy, and quality masks, which must be addressed to ensure effective monitoring activities (Frazier and Hemingway, 2021; Wildhaber et al., 2023). Compared to traditional platforms such as Landsat and Sentinel-2, the PSB.SD imagery has been described as not "analysis ready" upon download (Wildhaber et al., 2023). Thus, researchers must consider these variations and the evolving technologies to exploit the potential of PSB.SD imagery from Planet Labs PBC (2024) for agricultural monitoring applications (Frazier and Hemingway, 2021).

1.3 Estimating Crop Development with Key Physiological Indicators

Access to information on crop development throughout the growth season is essential to monitor the crop performance in phenotyping experiments (Brisco et al., 1998). Among the most commonly predicted plant physiology measurements from satellite imagery are canopy cover (CC) and Leaf Area Index (LAI), which are increasingly estimated using deep learning (DL) techniques (Victor et al., 2024). LAI serves as a key indicator of vegetation growth and health by assessing leaf coverage and leaf density (Lu et al., 2022; Hashimoto et al., 2023). Given LAI's significant influence on spectral variation, it emerges as a variable of interest (Berger et al., 2018; Verrelst et al., 2016). Radiative Transfer Models (RTMs) such as PROSAIL are commonly used to estimate LAI, allowing the retrieval of biophysical and biochemical variables from satellite-based RS data (Berger et al., 2018). Despite its widespread application, the PROSAIL model has limitations, particularly during the early vegetative stage when LAI is low. At this stage, background reflectance from soil and non-vegetative elements dominates the spectral signal, making accurate LAI retrieval challenging (Berger et al., 2018). To overcome these challenges, a recent study by Chen et al. (2024), on harmonized Landsat and Sentinel-2 data suggests that alternative vegetation indices (VIs), such as the Enhanced Vegetation Index (EVI) and VIs incorporating red-edge bands, provide superior performance in the early growth stages of crops. A study on winter wheat growth monitoring revealed that the Soil-Adjusted Vegetation Index (SAVI) and Modified Soil-Adjusted Vegetation Index (MSAVI) are more reliable for crop monitoring in early growth stages (Li and Chen, 2011). The described indices thus offer a more reliable representation of vegetation dynamics when LAI is still low, improving early-stage monitoring and precision in phenotyping applications. CC, often described as fractional vegetation cover, is defined as the fraction of ground area covered by vegetation (as cited in Chianucci et al., 2018, Walker and Tunstall 1981) and thus primarily captures the horizontal extent of vegetation. Since CC measures coverage and LAI accounts for both coverage and layering, their correlation is particularly strong when leaf overlap is low, making CC a reliable proxy for estimating LAI in the early growth stage (Nielsen et al., 2012). If the entire growth curve is to be assessed using satellite imagery, accurate capture of CC in the early growth stage is crucial, as it provides a foundational reference point for tracking biomass development throughout the season.

1.4 Spectral Approaches for Canopy Cover Estimation

Spectral RS techniques are widely used to characterize CC, with VIs serving as key predictors. Among these, the Normalized Difference Vegetation Index (NDVI) is one of the most frequently applied for CC estimation (Hatfield et al., 2008; Gao et al., 2020; Tenreiro et al., 2021). However, it has limitations, particularly with low vegetation density, due to its high sensitivity to soil optical properties, which can reduce accuracy in areas with significant soil background interference (Rondeaux et al., 1996; Montandon and Small, 2008). To address these limitations, various alternative VIs have been developed, including MSAVI, EVI, and Atmospherically Resistant Vegetation Index (ARVI) (Bannari et al., 1996; Gitelson, 2013). According to (Bannari et al., 1996), indices like NDVI and ARVI show lower performance in heterogeneous environments due to strong soil reflectance interference. ARVI, furthermore, often produces inconsistent results in practical applications. In contrast, MSAVI was specifically developed to minimize soil background effects, improving CC estimation in heterogeneous agricultural landscapes (Bannari et al., 1996). These findings highlight that while NDVI remains widely used, soil-adjusted and modified indices often provide greater accuracy, particularly in fields with bare soil exposure and varying canopy structures. While many crop-specific studies have reported a strong correlation between certain VIs and CC, Gao et al. (2020) highlighted significant variability in these relationships across and within different crop species. Similarly, Gitelson (2016) emphasized that predicting CC using NDVI varies considerably among crop types.

1.5 Segmentation Approaches for High-Resolution Canopy Cover Extraction

For images with greater spatial detail, segmentation techniques have become a key approach for estimating CC by differentiating vegetation from non-vegetation (Li et al., 2023). In this context, UAVs play a crucial role in data collection by capturing high-resolution imagery, which enables the generation of spatially continuous CC data that serve as quasi-ground truth for calibrating satellite-based models (Li et al., 2023). However, field-based segmentation methods face challenges due to varying illumination, heterogeneous backgrounds, and complex plant structures, making accurate CC estimation difficult (Yu et al., 2017). Traditional segmentation approaches often rely on thresholding techniques, where a fixed or adaptive threshold is set to separate vegetation from the background (Meyer and Neto, 2008; Wang et al., 2013). However, these methods struggle under fluctuating lighting conditions and diverse plant structures. To address these limitations, machine learning (ML)-based segmentation methods have been increasingly employed, leveraging their ability to learn complex patterns and improve classification accuracy (Zenkl et al., 2022). Zenkl et al. (2022) conducted an extensive evaluation of segmentation techniques for CC estimation with RGB (red, green, blue)-imagery, testing several ML models, including Decision Trees (Yu et al., 2017), Support Vector Machines (SVM) (Rico-Fernández et al., 2019), and RF classifiers (Sadeghi-Tehran et al., 2017). Their study utilized the Eschikon wheat segmentation (EWS) dataset, a dataset containing 190 labeled RGB images captured under diverse illumination and soil moisture conditions (Zenkl et al., 2022). Among the tested methods, Yu et al. (2017) proposed a two-stage classification approach, where an initial SVC was used for weather-state classification, followed by a second SVC for pixel-level segmentation. Rico-Fernández et al. (2019) further improved segmentation accuracy by integrating spatial context information within a 5×5 pixel window, transforming the data into the CIE-Luv color space before classification with an SVM. Meanwhile, Sadeghi-Tehran et al. (2017) employed an RF classifier trained on 21 color space features, achieving the most robust plant pixel detection across variable lighting conditions.

1.6 Bridging UAV and Satellite Data for Canopy Cover Estimation in Plot Experiments

While most satellite-based phenotyping studies focus on large experimental sites, research on smaller field trials has primarily concentrated on VIs. For instance, Sankaran et al. (2019) found a strong correlation (r up to 0.95) between NDVI from satellites and UAV multispectral imagery in $24\,\mathrm{m}^2$ maize plots. Similarly, Tattaris et al. (2016) demonstrated high correlations between satellite-based and UAV-based NDVI (r = 0.84) in wheat breeding trials, confirming the feasibility of high-resolution satellite imagery for plot-level phenotyping. While biomass and yield estimation is the primary focus of most studies, there is a notable gap in the use of satellite data for estimating CC, despite its significance for soil erosion (Zuazo and Pleguezuelo, 2008; Ruiz-Colmenero et al., 2013), crop development, and plant health (Bouman, 1995; Rasti et al., 2022). A meta-analysis on NDVI-based CC models conducted by Tenreiro et al. (2021) revealed that most existing approaches rely predominantly on in situ measurements rather than satellite-derived observations, limiting their applicability for larger scales. This finding underlines that the broader use of satellite data for CC estimation in agricultural contexts remains insufficiently explored. Moreover, while NDVI is the most commonly used index due to its simplicity, Tenreiro et al. (2021) suggest that alternative VIs may offer improved accuracy for CC estimation. These findings highlight an opportunity to expand the use of satellite imagery in CC modeling through the integration of alternative indices.

Complementing this, Chianucci et al. (2016) accentuate that recent advances in UAV technology and ML approaches hold strong potential for enhancing CC estimation beyond traditional index-based methods. Among the few studies applying modern learning techniques to satellite-based CC estimation, Bocco et al. (2012) demonstrates the potential of neural network-based models in the context of soybean crops. Given that model performance is highly sensitive to data characteristics and methodological design, a comprehensive comparison of linear indices, ML, and DL approaches remains largely unexplored. Therefore, I identify the systematic evaluation of different CC modeling approaches as a substantial research gap.

To bridge the identified gaps, I propose an integrated modeling framework that includes linear regression, ML, and DL approaches for satellite-based CC prediction, leveraging UAV-derived CC reference data exclusively for calibration and validation. However, this introduces new challenges such as cloud cover and potential data gaps during critical crop development stages. To ensure reliable CC estimation, satellite-derived predictions must be robust enough to account for data gaps and variations in spatial, spectral, and temporal resolutions. While multi-source satellite imagery can help mitigate these issues, integrating data from different missions remains complex due to inconsistencies in acquisition parameters (El Hajj, 2008; Sun et al., 2019; Yan et al., 2021). A more practical solution is to use multiple sensors from a single mission, such as PSB.SD, which provides consistent spatial and spectral characteristics across acquisitions, improving the stability and reliability of CC predictions (Planet Labs PBC, 2024).

Considering the effectiveness of low-cost UAVs for detailed CC mapping and the established utility of RS for monitoring, combining high-resolution RGB UAV data with continuous multispectral PSB.SD imagery provides a promising solution for continuous field monitoring, especially during the early stages of crop growth.

1.7 Research Aims and Objectives

This study aims to evaluate the potential of high-resolution satellite imagery for estimating CC in plot-scale agricultural field experiments. By integrating UAV-derived RGB imagery as a source of calibration data and leveraging multispectral satellite data from PSB.SD, this research seeks to bridge the gap between fine-scale UAV-based phenotyping and scalable RS technologies. Ultimately, this study seeks to improve the utility of satellite-based monitoring in small-plot field experimental setups.
The specific objectives of this study are:

1. **Application of an existing CC segmentation model** to UAV-derived RGB imagery to generate high-resolution reference data. This reference dataset, derived from multiple crops across several training sites in northern Switzerland, serves as a calibration source for training and validating satellite-based CC prediction models.

2. **Development and validation of CC prediction models** using PSB.SD multispectral satellite imagery. Leveraging UAV-derived CC data as a reference, various modeling approaches, including linear regression, ML, and DL, are used to predict CC and enable scalable time-series monitoring of crop canopy dynamics.

3. **Time series analysis of plot experiments** to monitor CC development under different (bio)fertilizer treatments. The developed CC prediction models are applied to an independent experimental site to assess their performance under real-world conditions.

2 EXPERIMENTAL DESIGN

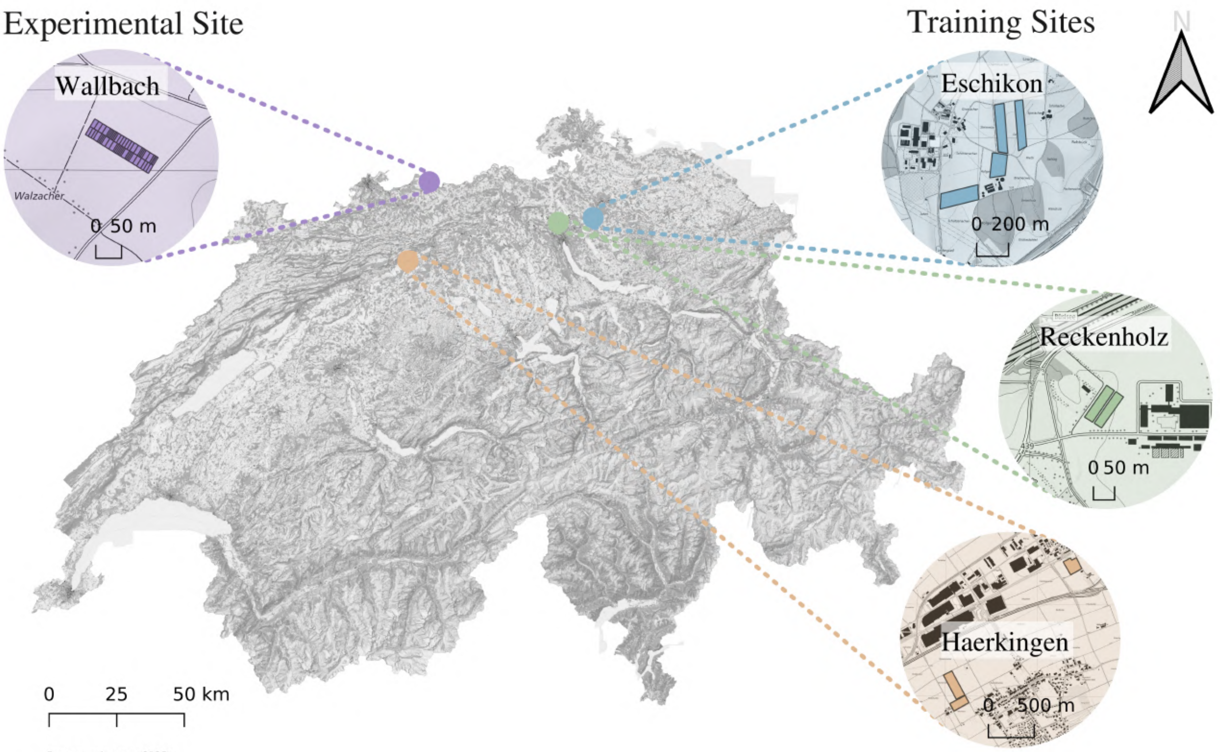
The study was conducted across multiple sites located in northern Switzerland, as illustrated in Figure 1. These sites were selected to represent a diverse range of soil properties, crop types, and management practices. The field sites were divided into training sites, used for model development, and an experimental site for independent model evaluation. For the training sites, both UAV imagery and satellite data were collected to develop and validate prediction models for CC. UAV imagery provides high spatial resolution and represents the closest available estimate to ”ground truth” data, and was therefore used as a proxy during model development. For the experimental site, only satellite data from the PSB.SD constellation is used for predictions. This enables a clear evaluation of the model’s performance and transferability in estimating CC under conditions independent of the training sites.

2.1 Training sites

The training of the CC models is based on data collected from three distinct locations in northern Switzerland: Eschikon, Reckenholz, and Haerkingen. The crops included at these training sites are sunflower, corn, and sugarbeet. A summary of the site-specific crop types, seeding dates, and coordinates is presented in Table 1. All study sites are classified as ”very good production suitability” (slopes $< 25\%$) based on the Digital Soil Suitability Map provided by BLW (2024a). According to the Crop Type Suitability Layer, the sites are suitable for different cropping systems, including arable farming, cereal cultivation, and forage production (BLW, 2024b).

Table 1. **Training Site Overview** with crop type, seeding date, and site coordinates (EPSG:32632).

Location	Site	Crop Type	Seeding Date	Coordinates (E, N)
Eschikon	Steinmueri1	Sunflower	08/04/2024	(476395, 5255065)
	Steinmueri3	Corn	13/05/2024	(476516 5255083)
	Laubisgruet2	Sugarbeet	08/04/2024	(476131, 5254637)
	Holgenbuehl	Corn	12/05/2024	(476400 5254834)
Reckenholz	Eingang	Sugarbeet	08/04/2024	(463378, 5252956)
Haerkingen	Altgraben	Corn	08/05/2024	(409367, 5240774)
	Chaeppeli	Corn	15/05/2024	(408108, 5239294)
	Neufeld4	Corn	15/05/2024	(408081 5239428)



Source: swisstopo (2025)

Figure 1. **Overview of Studysites** in Switzerland comprising one experimental site (Wallbach (purple)) and three training sites (Eschikon (blue), Reckenholz (green), Haerkingen (orange)). Colored dots mark site locations, with dashed lines linking to inset maps showing field extents as filled polygons on the background map: Landeskarte grau from © swisstopo (2025)

2.2 Experimental site

The experimental site is part of the ongoing ”Bio4Recycle” project and is located in northern Switzerland. It lies within a region classified as having very good production suitability (slope $< 25\%$) according to the Digital Soil Suitability Map, provided by BLW (2024a). The Crop Type Suitability Layer (BLW, 2024b) reveals that the site is well suited for cereal cultivation and forage production, making it comparable in terms of soil characteristics to the training sites included in this study. In the 2023/2024 growth season, winter barley is cultivated at this location, with a seeding date at the beginning of October 2023, shown in Table 2.

Table 2. Experimental Site Overview with crop type, seeding date, and site coordinates (EPSG:32632).

Location	Site	Crop Type	Seeding Date	Coordinates (E, N)
Wallbach	Bio4Recycle	Winter Barley	10/10/2023	(461324, 5268252)
	Bio4Recycle	Winter Wheat	07/10/2022	(461324, 5268252)
	Bio4Recycle	Corn	11/05/2022	(461324, 5268252)

Since 2018, the ”Bio4Recycle” trial of Bünemann-König and Mayer (2021) has focused on optimizing nitrogen use efficiency and assessing the climate impact of recycled fertilizers in organic farming systems. The current project phase investigates yield response, nitrogen utilization efficiency, loss potential, and climate effects of recycled fertilizers, including liquid and solid digestates, as outlined in Table 3 (Diener et al., 2024). An overview of the plot-wise treatments and their dimensions, as shown in the appendix (see Figure 35), is reproduced with permission from Agostini et al. (2024), with treatment plots typically measuring 9 m in width and 18 m in length. The overarching goal, according to Bünemann-König and Mayer (2021), is to support the parametrization of Agroscope’s nitrogen efficiency model using robust field data collected under organic farming conditions. Alongside plant and soil sampling, the ongoing monitoring also incorporates multispectral UAV imagery (Diener et al., 2024).

Table 3. Selected Fertilization Treatments in the ”Bio4Recycle” experimental site at Wallbach based on Bünemann-König and Mayer (2021).

Treatment	Description
Controls	
NON	No nitrogen (0-control)
MIN	Mineral nitrogen fertilizer (positive control)
Liquid Organic Fertilizers	
SLU	Cattle slurry
SLA	Digestate (agricultural biogas plant)
LID	Liquid digestate (industrial biogas plant)
Solid Organic Fertilizers	
SD	Solid digestate (industrial biogas plant)

To provide environmental context for interpreting crop development and model predictions, climatic and soil conditions at the Wallbach site are illustrated in Figure 2. The figure, reproduced with permission from Agostini et al. (2024), presents daily precipitation, air temperature at 2 m, soil temperature, and soil water content during the 2023–2024 period, along with key field management events such as fertilization, sowing, and soil preparation. Data on air temperature, precipitation, global radiation, and wind speed were obtained from nearby weather stations operated by MeteoSwiss

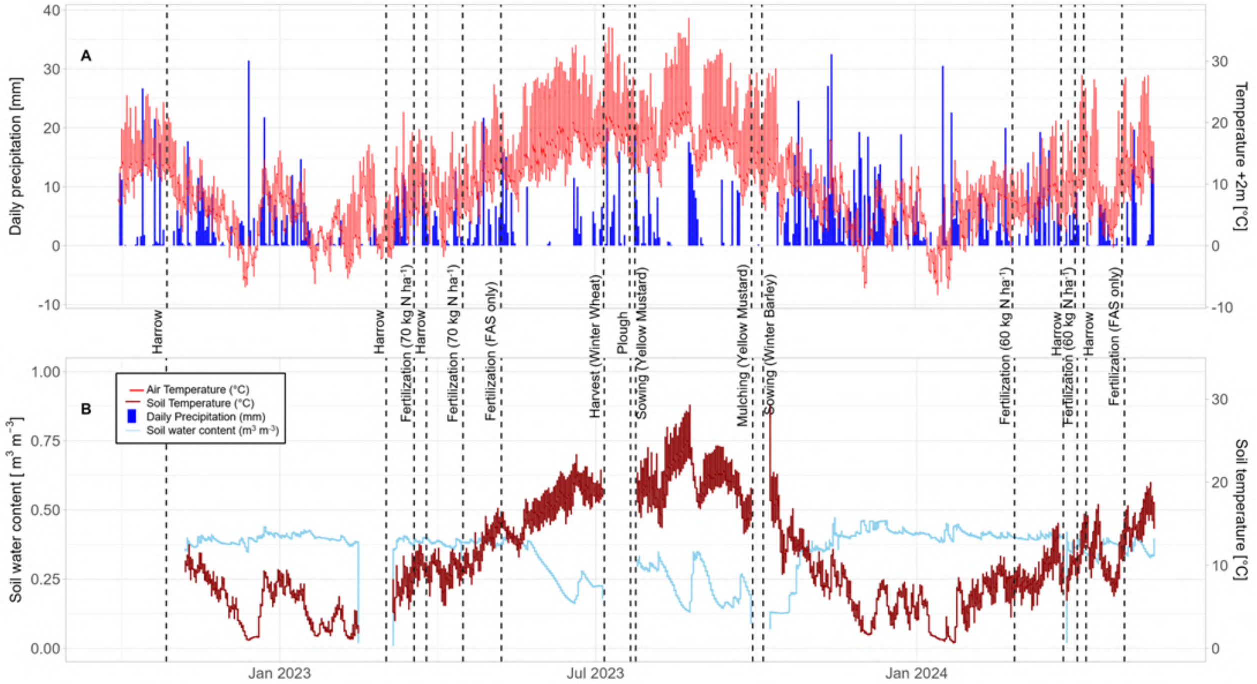


Figure 2. **Climatic and Soil Conditions at the "Bio4Recycle" Experimental Site** in Wallbach over the 2023–2024 period. Panel A shows daily precipitation (blue bars), air temperature at 2 meters (red line), and key management events such as harrowing, fertilization, sowing, and harvesting (vertical dashed lines). Panel B displays soil temperature (dark red) and soil water content (light blue), providing additional insight into soil moisture dynamics. Data on air temperature, precipitation, global radiation, and wind speed were obtained from nearby weather stations operated by MeteoSwiss (47°34'19.9" N, 7°52'40.5" E, 343 m a.s.l.) and Agrometeo (47°34'10.6" N, 7°53'54.4" E, 346 m a.s.l.). *Figure reproduced with permission from Agostini et al. (2024).*

(47°34'19.9" N, 7°52'40.5" E, 343 m a.s.l.) and Agrometeo (47°34'10.6" N, 7°53'54.4" E, 346 m a.s.l.). This contextual information supports the interpretation of temporal patterns in CC predictions and observed treatment-level differences.

3 METHODS

The methodological workflow of this study consists of three main stages: training data development, satellite model training, and model application. The Python environment for this analysis was based on version 3.11.9. First, UAV-based reference data were generated at selected training sites using a two-stage segmentation approach. A high-resolution CC model was trained on annotated image patches and applied to high-resolution RGB UAV imagery. These outputs were then processed and used to train a field-level CC model, which was applied to UAV mapping images. The resulting segmented CC data were used to construct a training dataset by spatially aligning UAV-derived CC estimates with PSB.SD satellite pixels and augmenting them with spectral bands, VIs, and LAI. Next, a series of satellite-based CC models, including linear regression, ML, and neural network approaches, were trained on this dataset. Model performance was evaluated using both stratified sampling and mission-based cross-validation. Finally, the trained models were applied to an independent experimental site, where plot-level CC predictions were generated and used to assess model transferability and treatment-specific responses across multiple crop seasons in a long-term field experiment.

3.1 Establishing Canopy Cover Reference Data with UAV RGB Data

To generate reliable CC reference data for model calibration, UAV-based RGB imagery was collected and processed using a multi-scale segmentation pipeline. This section details the UAV platform and mission planning, followed by the application of two complementary segmentation models: a high-resolution CC model and a field-level CC model. Together, these models enable accurate CC estimation across different spatial scales.

3.1.1 DJI Mavic 3E: Platform and Imaging Configuration

The DJI Mavic 3 Enterprise (Mavic 3E) is a compact, multi-rotor UAV designed for high-resolution spatial data collection across a range of scientific applications. The platform is equipped with an integrated dual-camera system that captures high-quality RGB imagery. It combines a high-resolution wide-angle sensor with a precision telephoto imaging unit, thereby supporting both aerial mapping and detail-oriented observations (DJI, 2022). The primary imaging sensor utilizes a 4/3-inch CMOS detector with a 20-megapixel resolution and a mechanical shutter, effectively minimizing motion blur and rolling shutter distortion during flight. This configuration is suitable for large-scale mapping workflow (DJI, 2022). The secondary camera consists of a 1/2-inch CMOS sensor with a lower native resolution, specifically engineered for zoom functionality. It supports a hybrid zoom system that combines both optical and digital zoom capabilities. The optical zoom enables lossless magnification of up to 7×, while the hybrid system extends the total zoom range to a maximum of 56× (DJI, 2022). To enhance the spatial accuracy of image geotagging, the Mavic 3E is equipped with an integrated Real-Time Kinematic (RTK) module. This system enables centimeter-level positioning precision by receiving correction data via the NTRIP (Networked Transport of RTCM via Internet Protocol) protocol, eliminating the need for a local base station (DJI, 2022). This functionality markedly improves the absolute accuracy of georeferenced outputs such as orthomosaics, while also reducing or eliminating the need for ground control points in post-processing workflows (DJI, 2022). In this study, both imaging systems were employed to combine broad-scale field coverage with high-detail snapshot (SNP) captures.

3.1.2 UAV Mission Planning and Image Acquisition

The UAV-based data collection for this study was conducted using a Mavic 3E, employing both SNP and mapping mode to capture high-resolution RGB imagery from a specific area on the surface and ensure large-area coverage, respectively. The required degree of detail and aerial extent were achieved while optimizing battery capacity using this approach. For the mapping missions, the UAV followed the predefined settings of the Mavic 3E's mapping mode at a flight altitude of 12 m. The overlap between successive images was set to 70-80%, ensuring sufficient data redundancy for accurate reconstruction. The UAV maintained a constant flight speed of 3 m/s throughout the acquisition process. The classification of the mapped area was set to farmland, and the illumination settings were configured to be automatic to adapt best to the varying light conditions during the flights. Under these parameters, the UAV achieved its highest possible ground sampling distance (GSD) of 32 mm/pixel.

Several zoom height variations were tested to maximize image quality in the SNP missions. The best qualitative outcome in terms of illumination, area coverage, and detail was achieved at a zoom factor of 10 with a flight altitude of 15 m. Under these conditions, each captured SNP image covered a ground-truth extent of approximately 2.16 meters × 1.62 meters, resulting in an area of $3.48\,\mathrm{m}^2$ per image. The UAV operated at a flight speed of 5 m/s, but to maximize image sharpness and minimize motion blur, it stopped momentarily for each image acquisition. The SNP file was generated by processing a 9 m raster grid to extract a flight path within the defined field extent. The centroids of each grid cell were then converted into point features, which were spatially filtered and sequentially connected to form a continuous path, as illustrated in Figure 36 of the appendix. This path was subsequently converted into a KML file and imported into `FlightHub2`, where gimbal tilt (-90°), flight height (15m), camera zoom (10), and filenames were manually set for mission execution.

Finally, for the mapping and SNP missions, the field extent and predefined flight paths, provided in KML format, were uploaded to the UAV console for each mission, respectively. Both missions were executed in automatic mode and in full compliance with the EASA A2 subcategory regulations. Due to the proximity of the Reckenholz fields to the airport, prior notification and approval from the airport authority were required for each flight. Following data acquisition, the mapping-mode images were processed using `DJI SmartFarm` software. The processing workflow involved merging the image datasets and generating orthomosaics, with the map resolution set to 100%.

3.1.3 Canopy Cover Segmentation Workflow

To segment vegetation from background in the UAV imagery, two complementary segmentation models were applied at different spatial scales. The first model, referred to as the high-resolution CC model, utilized the annotated EWS dataset (Zenkl et al., 2022) in combination with UAV-based high-resolution SNP imagery. The second model, referred to as the field-level CC model, made use of segmented SNP masks together with UAV-derived mapping images to estimate CC at the field scale. Together, these two models enabled a multi-scale segmentation approach that incorporated both detailed visual information and a broader spatial context.

High-Resolution Canopy Cover Model

First, a high-resolution CC model was established using UAV-based high-resolution SNP imagery and the EWS dataset. The publicly available EWS dataset, provided by Zenkl et al. (2022), consists of 190 manually selected and annotated image patches (350 × 350 pixels) of winter wheat genotypes in Eschikon, Switzerland. The corresponding RGB images were captured at a height of 3 m, with a GSD of 0.3mm/pixel. Images were collected throughout the growth season under varying ambient light and soil moisture conditions, representing a broad range of real-field environment conditions. This labeled dataset was used to train the high-resolution CC model proposed by Sadeghi-Tehran et al. (2017). The corresponding code for this model is freely available through Zenkl et al. (2022). A distinctive characteristic of the approach is the transformation of the original RGB images into various color spaces using the `skimage.color` module (Van Der Walt et al., 2014). These include Lab, Luv, HSV, HSI, YCbCr, and YUV, each capturing complementary aspects of brightness, chromaticity, and color contrast. The model architecture is based on an RF classifier, trained on 21 image features derived from multiple color spaces (Sadeghi-Tehran et al., 2017). Prior to training, all extracted features were normalized using `sklearn.preprocessing.StandardScaler` (Pedregosa et al., 2011) to ensure comparability across feature dimensions. The RF classifier was configured with a maximum tree depth of 95 (*max_depth=95*), 55 estimators (*n_estimators=55*), and a minimum of 6 samples per leaf (*min_samples_leaf=6*). All preprocessing steps and parameter settings followed the shared implementation provided by (Zenkl et al., 2022). Once the model was trained on the EWS dataset, the model was applied to predict CC from SNP imagery, as illustrated in Figure 3. According to their evaluation on the EWS dataset, Zenkl et al. (2022) reported a pixel-wise segmentation accuracy of 90% for the model proposed by Sadeghi-Tehran et al. (2017).

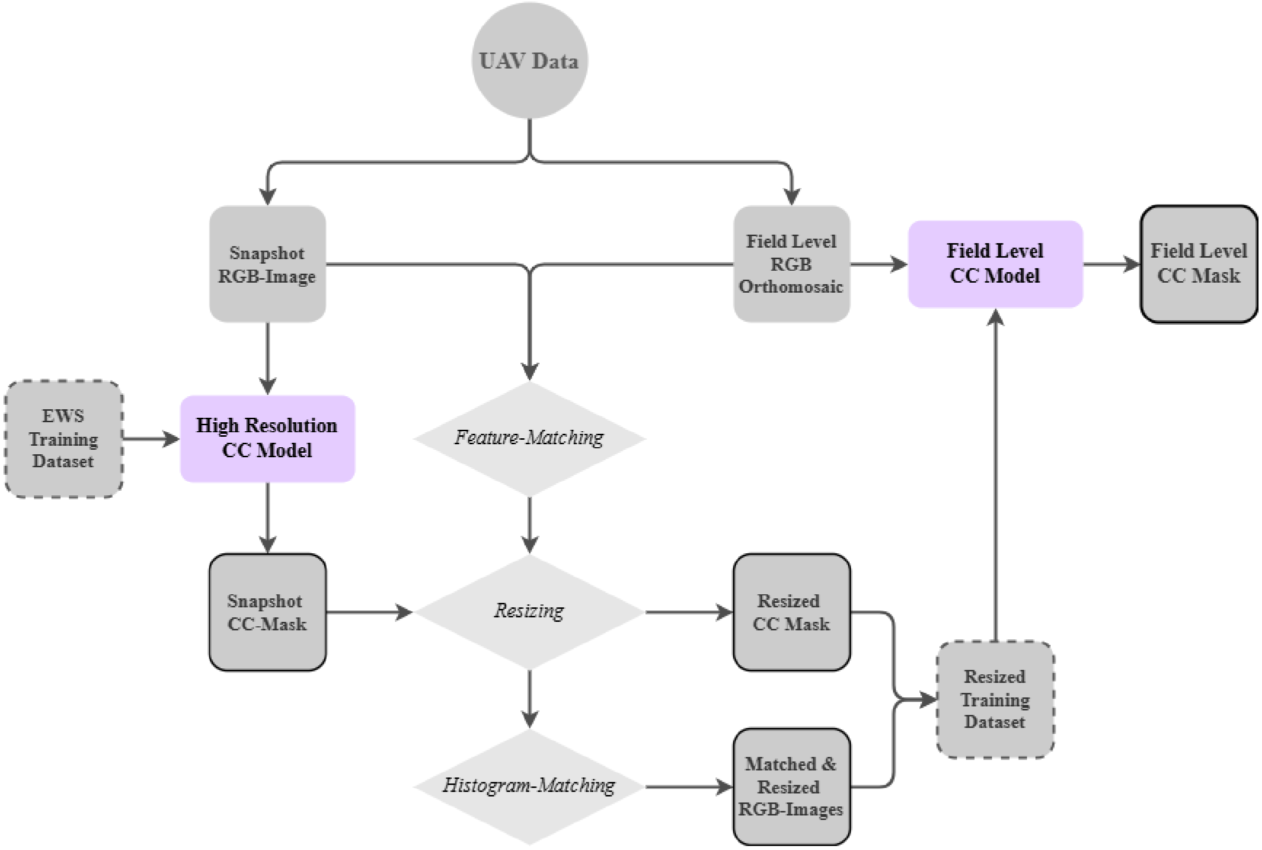


Figure 3. **Segmentation Pipeline for High-Resolution and Field-Level Canopy Cover Model**. This flowchart outlines the processing steps from UAV RGB input data to the generation of high-resolution and field-level canopy cover (CC) models, ultimately producing a field-level CC mask

Field-Level Canopy Cover Model

Following the segmentation of the high-resolution SNP images, additional preprocessing was required to prepare the SNP data for training a model suited to the lower-resolution mapping data. Since the SNP and mapping images were acquired using different sensors, geometric and radiometric matching was crucial for training a reliable and meaningful field-level CC model.

Spatial Alignment through Feature Matching

Given the high-resolution of the training data, several processing steps were necessary before it could be applied to the lower-resolution field-level mapping image (32mm/pixel GSD). To achieve this, the model was first applied to high-detail SNP images (0.3 mm/pixel GSD) from the SNP mission to predict plant pixels on the RGB images. The binary prediction masks, used to distinguish between plant and background pixels, provided a high-resolution reference for specific areas within the field. To align the SNP images with the mapping image, a Scale-Invariant Feature Transform (SIFT) (Lowe, 2004) was employed. This method allowed for the detection of distinctive key points across both images, ensuring accurate alignment despite variations in scale and rotation (Lowe, 2004). The algorithm was configured with three octave layers (*nOctaveLayers=3*), a contrast threshold of 0.04, an edge response threshold of 10, and a Gaussian smoothing factor of 1.6 (*sigma=1.6)* to ensure robust keypoint detection. Feature matching was performed using a ratio test threshold of 0.3 to retain only strong correspondences and a minimum of three valid matches (*MIN_MATCH_COUNT=3*) was required to compute the homography transformation. This approach enabled the scale- and rotation-invariant detection of common features, allowing for the precise extraction of the SNP image extent on the mapping image, as shown in Figure 37 in the appendix. In a few individual cases with high CC or strong winds, SIFT feature matching failed due to insufficient key points. In these instances, centered cropping was used as a fallback method, with a fixed window size applied around the estimated center location for cropping. In a next step, the RGB SNP images and binary masks were resized to match the dimensions of the corresponding cropped mapping image.

Radiometric Normalization using Histogram Matching

The spatially aligned SNP images were matched to the mapping image in terms of color distribution to correct for differences in illumination conditions caused by different sensors used for the two image types. To achieve this, histogram matching was applied to adjust the pixel intensity distribution of the SNP images, aligning them with the corresponding cropped mapping images. This process was carried out independently for each color channel (red, green, and blue) using a cumulative distribution function (CDF) mapping approach. The histogram of each image was computed, and the corresponding CDFs were derived to represent the cumulative pixel intensity distributions (see Figure 38 in the appendix). To align these distributions, an interpolation function was applied to map the CDF of the SNP image to the CDF of the cropped mapping image. This method preserved the local contrast and texture of the original images while achieving global color normalization across the two datasets, as shown in Figure 39 in the appendix. The resulting consistency in color representation was crucial for ensuring the applicability of the model trained on SNP images to the mapping images.

Field-Level CC Prediction from Aligned Training Data

To predict the CC of the field-level mapping images, the feature-matched, resized, and histogram-matched SNP images were used for training. These RGB images, along with their corresponding binary masks, served as input for the field-scale segmentation model, adapted from Sadeghi-Tehran et al. (2017) with the previously described parameters. By using training data from the same site and date, the reliability of the field-level segmentation model is enhanced further by training it within the same contextual conditions. To predict the CC of the field-level mapping images, the feature-matched, resized, and histogram-matched SNP images were used for training. These RGB images, along with their corresponding binary masks, served as input for the field-level CC model, adapted from Sadeghi-Tehran et al. (2017) with the previously described parameters. As part of this adaptation, the model pipeline was modified to handle GeoTIFF files instead of PNGs, ensuring compatibility with the geospatially referenced mapping data. Additionally, the input patch size was adjusted from 350×350 pixels to 256×256 pixels to better match the lower resolution of the training data derived from the mapping images. To improve computational efficiency, parallelization was implemented using Python's `ThreadPoolExecutor`, enabling concurrent processing of multiple field sites and subfolders. This approach significantly reduced the overall runtime by an estimated factor of 5 compared to sequential execution. Model validation was performed using an independent subset of the training data. The F1 score, precision, recall, IoU, and accuracy metrics, which were already implemented in the code of Sadeghi-Tehran et al. (2017), are reported in Table 20 in the appendix.

3.2 Canopy Cover Modeling Using Multispectral PlanetScope Imagery

Multispectral PlanetScope (PS) imagery from the SuperDove sensor was utilized to develop satellite-based CC models. The modeling process involved extensive preprocessing and data preparation steps, as illustrated in Figure 5. Finally, the section provides an overview of the evaluation methods used to assess the performance and accuracy of the models.

3.2.1 PlanetScope SuperDove: Sensor Overview

The PS constellation, operated by Planet Labs PBC (2024), consists of over 130 CubeSat 3U satellites that capture high-resolution, near-daily imagery of Earth's land surface at approximately 3 m spatial resolution. These satellites operate in a sun-synchronous orbit at an altitude of 475 km to 525 km with an inclination of 98° (Planet Labs PBC, 2025a). The repeat cycle is daily, with equator crossings occurring between 9:30 and 11:30 am (local solar time). The latest generation, SuperDove (SD), has been providing data since mid-March 2020 and records imagery across eight spectral bands, as listed in Table 4, including Coastal Blue, Blue, Green I, Green, Yellow, Red, Red Edge, and Near-Infrared (NIR) (Planet Labs PBC, 2025b).

Table 4. **Spectral Bands of SuperDove** and corresponding wavelength ranges (nm) of the sensor (Planet Labs PBC, 2025b).

Spectral Band	Wavelength Range (nm)
Coastal Blue	431 – 452
Blue	465 – 515
Green I	513 – 549
Green	547 – 583
Yellow	600 – 620
Red	650 – 680
Red Edge	697 – 713
NIR	845 – 885

This study utilizes the Surface Reflectance (*analytic_sr_udm2*) product, which provides atmospherically corrected bottom-of-atmosphere (BOA) reflectance values. The ”analytic_sr_udm2” product undergoes several preprocessing steps. These include radiometric calibration, which converts raw digital numbers into reflectance values, orthorectification, which corrects terrain distortions using a digital elevation model (DEM), and atmospheric correction, which mitigates scattering and absorption effects using the 6S RTM (Planet Labs PBC, 2025b). Additionally, Bidirectional Reflectance Distribution Function (BRDF) normalization adjusts for variations in illumination and sensor geometry, while Quality Assurance (QA) masking identifies and classifies clouds, haze, shadows, snow, and clear-sky pixels to improve data usability (Planet Labs PBC, 2025b). The downloaded ”analytic_sr_udm2” product is provided in 16-bit GeoTIFF format with accompanying metadata on atmospheric disturbances and confidence values (Planet Labs PBC, 2025b).

3.2.2 Satellite Data Preprocessing Workflow

Since challenges with PSB.SD imagery, such as unreliable quality masks, misalignment, and inconsistent radiometry, were pointed out by earlier research (Wildhaber et al., 2023; Frazier and Hemingway, 2021), additional processing steps were included. These contained quality masking, co-registration, and Empirical Line Correction (ELC) to improve both spatial and radiometric reliability.

Automated Data Acquisition via Planet API

To acquire PSB.SD imagery provided by Planet Labs PBC (2024) across the various study sites, an API-based Python script was implemented using the `requests` library (Kenneth Reitz, 2011) to interact with the PS API. This programmatic approach facilitated automated querying, ordering, and downloading of satellite imagery. A spatial filter was defined using a GeoJSON file to clip the extent of each study site. The corresponding geometry was incorporated into a ”GeometryFilter”, thereby restricting the imagery selection to scenes intersecting the area of interest. Further, a ”DateRangeFilter” was employed to confine image selection to the designated observation period. To mitigate atmospheric interference and exclude scenes with excessive cloud contamination, a ”CloudFilter” was applied, restricting total scene-level cloud coverage to a maximum of 60%. It is important to note that the PS API does not currently support cloud filtering specifically within the area of interest (AOI) when using automated scripting. Consequently, a relatively high cloud cover threshold was chosen, with additional cloud masking procedures applied during further processing steps. The spatial, temporal, and cloud cover filters were combined and ”PSScene” was selected as item type, which corresponds to standard PSB.SD imagery. Imagery was ordered using the ”analytic_8b_sr_udm” product bundle, which includes eight-band, atmospherically corrected surface reflectance data, along with associated confidence and cloud masking layers. Following scene selection, relevant scene IDs were extracted, and orders were placed via the PS Orders API (*https://api.planet.com/compute/ops/orders/v2*). Upon order completion, the resulting assets were downloaded.

Cloud and Artifact Masking

The masking process utilized masks provided with the metadata of the PSB.SD imagery. A combined mask was generated by identifying pixels affected by shadow (band 3), heavy haze (band 5), and cloud presence (band 6). Due to unreliable cloud masks, a confidence mask was additionally applied, discarding all pixels with confidence values below 90% (band 7). The final mask was obtained by merging these criteria, ensuring that only high-confidence, cloud-free pixels were retained for further analysis.

Co-registration to SwissImage

To address spatial misalignment of the PSB.SD imagery, a co-registration procedure was implemented using SwissImage orthophotos, which provide high spatial accuracy with a resolution of 10 cm (Federal Office of Topography (swisstopo), 2025). The Red band from the SwissImage dataset was selected as the reference band due to its high spatial contrast. For image alignment, the `AROSICS` (Automated and Robust Open-Source Image Co-Registration Software) library (Scheffler et al., 2017) was employed. `AROSICS` is a feature-based co-registration tool designed for RS applications, utilizing phase correlation for sub-pixel shift estimation in the frequency domain based on the Fourier shift theorem (Scheffler et al., 2017). Additionally, `AROSICS` supports the masking of areas affected by disturbances. This option was utilized to mask out high buildings, thereby avoiding potential misalignment artifacts caused by elevated structures. By ensuring that such regions were excluded from spatial shift detection, the robustness of the co-registration process was aimed to be improved.

Empirical Line Correction for Radiometric Harmonization

An ELC was applied to correct radiometric inconsistencies in satellite imagery. This method normalizes image radiometry by using reference targets with stable reflectance properties. In this study, Pseudo-Invariant Targets (PITs) were selected as reference areas for radiometric calibration. The selection of PITs was guided by several key criteria to ensure reliability and consistency. Ideally, targets should exhibit minimal spectral and radiometric variability over time, making them suitable for long-term calibration. Artificial surfaces such as flat rooftops, roads, and unused parking areas were prioritized due to their stable reflectance properties. Horse tracks, composed of sandy, non-vegetated surfaces, were also considered radiometrically stable. To reduce noise and mixed-pixel effects, selected areas needed to be large enough to cover multiple pixels. Furthermore, efforts were made to include surfaces with diverse reflectance characteristics, incorporating both bright and dark targets. While these criteria were aimed to be met as consistently as possible, practical constraints such as the availability of suitable surfaces sometimes required compromises. The selected PITs represent the most stable and appropriate reference targets available within the study sites. The final set of PITs used for ELC is listed in Table 5.

Table 5. **Selection of Pseudo-Invariant Targets** used for Empirical Line Correction (ELC), including the description of Pseudo-Invariant Targets (PIT) and their center locations in UTM coordinates (EPSG:32632) for each site location.

Location	PIT Description	Center Location (X, Y)
Eschikon	Horse track	(476161.81, 5255005.49)
	Junction	(476002.86, 5254781.05)
	Roof	(476389.35, 5254685.51)
Reckenholz	Junction	(463206.21, 5252866.39)
	Roof	(463623.44, 5252945.75)
	Roof	(463581.26, 5252874.19)
Haerkingen	Horse track	(408238.68, 5239326.56)
	Roof	(408781.21, 5240760.88)
	Parking	(408259.87, 5240158.00)
Wallbach	Crossing	(415958.68, 5268734.20)
	Crossing	(416519.48, 5268317.50)
	Horse track	(415888.75, 5268684.17)

For each of the selected PITs and satellite images, the spectral mean ($\mu_{i,b}$) was computed bandwise. PITs containing invalid pixels were excluded from further processing. A baseline mean reflectance value ($\bar{\mu}_b$) was established by averaging spectral measurements across all valid PITs over the entire time series, serving as a radiometrically stable reference. Correction factors ($CF_{i,b}$) were calculated for each band as:

$$CF_{i,b} = \frac{\mu_{i,b}}{\bar{\mu}_b} \tag{1}$$

Where:

- $CF_{i,b}$ is the **correction factor** for image i and band b,
- $\mu_{i,b}$ is the **mean** of band b in image i over the PIT area,

- $\bar{\mu}_b$ is the **baseline mean** for band b across all valid PITs over the entire timeseries.

The computed correction factors were applied at the pixel level to normalize radiometric variations across images. Each pixel value was divided by its corresponding correction factor, reducing values that exceeded the baseline while increasing those below it. The computed correction factors are summarized in Table 6 and range from 0.6 to 1.4 across all study sites. Notably, the maximum value of 1.4 is observed at multiple locations on the same date (Figures 40 and 42 in the appendix), reinforcing the robustness and consistency of the correction methodology. A specific characteristic was noted for the PIT ”Horse track” in Wallbach, where distinct reflectance properties were observed prior to 30.05.2022. Based on visual inspection, this shift is likely attributable to a change in the surface material of the horsetrack, most likely from wood chips or sand. To ensure consistency in reflectance conditions, this PIT was therefore only included in the analysis for data acquired after 30.05.2022.

By implementing ELC, radiometric inconsistencies were minimized, enhancing the comparability of SuperDove imagery across acquisition dates and improving the reliability of subsequent analyses.

Table 6. **Correction Factor Ranges** used in Empirical Line Correction (ELC). The table shows the minimum and maximum correction factors applied at each study location (Eschikon, Reckenholz, Haerkingen, Wallbach), with the overall correction factor range provided. Data © 2024 Planet Labs PBC.

Location	Minimum Correction Factor	Maximum Correction Factor
Eschikon	0.9	1.4
Reckenholz	0.8	1.4
Haerkingen	0.85	1.2
Wallbach	0.6	1.4
Overall Range	**0.6**	**1.4**

Temporal Data Availability and Coverage

PSB.SD imagery acquisition was consistent across most sites, with multiple cloud-filtered scenes available per month. However, despite the theoretical daily revisit frequency of the PSB.SD constellation, effective data availability was limited by cloud cover, resulting in several temporal gaps. Figure 4 illustrates the temporal distribution of available satellite imagery in relation to UAV observations across all study sites. Grey vertical bars represent PSB.SD acquisition dates that passed quality filtering. UAV campaign dates are categorized as either temporally overlapping (within ±2 days of a satellite acquisition) or non-overlapping. Temporal overlap was frequent at the Eschikon and Reckenholz sites, enabling the integration of UAV and satellite data for multi-scale analyses. In contrast, Haerkingen exhibited fewer overlaps, likely due to persistent cloud cover or fog, which reduced the number of usable satellite observations. These temporal gaps emphasize the need for interpolation methods to construct continuous time series and support downstream analyses.

Interpolation for Temporal Gap Filling

To address temporal gaps in PSB.SD imagery at the time of UAV data acquisition, a linear interpolation approach was implemented to estimate missing spectral values based on adjacent satellite observations. For UAV missions where a PSB.SD acquisition was available within ±2 days of the UAV observation, the nearest satellite-derived spectral

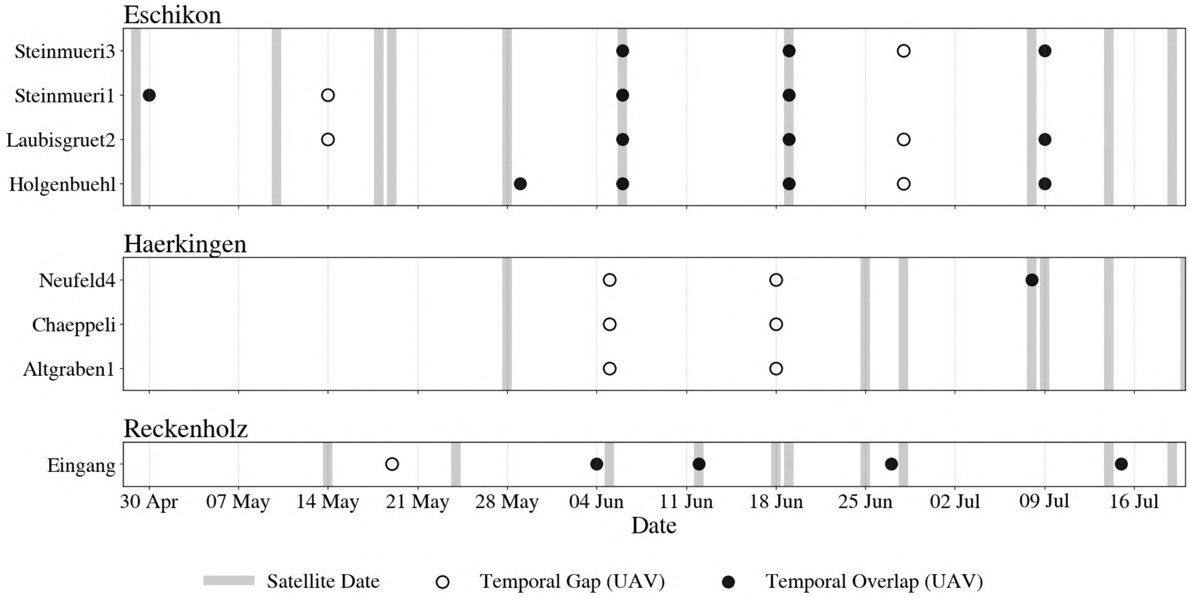


Figure 4. **Data Availability of UAV and Satellite Observations for Training Sites**. The figure illustrates the variability in data density across sites (Eschikon, Haerkingen, and Reckenholz) throughout the monitoring period (28 April 2024 to 20 July 2024). Gray vertical bars represent dates with available satellite data, while UAV data availability is shown as temporal overlap with satellite acquisitions (black circles) or temporal gaps (white circles). Data © 2024 Planet Labs PBC.

values were directly assigned without interpolation. For UAV data points lacking such temporal proximity, linear interpolation was performed using the two closest cloud-filtered PSB.SD acquisitions before and after the UAV flight date. Interpolation was performed at the pixel level across all eight spectral bands using a commonly applied approach, previously used for NDVI time series reconstruction by Pan et al. (2017). The reflectance value $\hat{R}(t)$ at the UAV acquisition time t was estimated using the following linear interpolation formula:

$$\hat{R}(t) = R_1 + \left(\frac{t - t_1}{t_2 - t_1}\right)(R_2 - R_1) \tag{2}$$

Where R_1 and R_2 are the surface reflectance values from the earlier and later satellite acquisitions at times t_1 and t_2, respectively. In addition to spectral values, satellite metadata such as orbit number, product type, and sensor ID were concatenated across the two bounding observations to retain acquisition context. Furthermore, only satellite acquisitions obtained after the seeding date were considered to avoid potential reflectance influences from preceding vegetation. This interpolation strategy enabled the construction of harmonized and continuous UAV–satellite time series, supporting robust temporal comparisons between PSB.SD imagery and UAV-derived reference observations.

3.2.3 Data Preparation for Satellite Canopy Cover Model

The preparation of a training data frame for the satellite-based CC model involves several key steps, including pixel-level dataset construction, aggregation of canopy cover from UAV data, derivation of VIs, and the estimation of LAI using an RTM.

Pixel-Level Dataset Construction

The processed satellite data was converted into a structured tabular format to prepare the data for satellite-based CC model development. Site-specific pixels were identified and extracted for each raster file. Each pixel was indexed by its row and column position within the raster grid and georeferenced using its spatial center coordinates (easting and northing). Metadata embedded in the raster filenames, such as acquisition date, satellite ID, orbit number, product type, and details on preprocessing steps, were parsed and appended to each pixel entry. To ensure full traceability across data sources, metadata linking each satellite observation to the corresponding UAV flight, including UAV acquisition date and flight ID, were also included. Additionally, crop type and field presence, distinguishing between pixels located entirely within field boundaries and those near or outside field edges, were assigned to each pixel.

Canopy Cover Aggregation from UAV Data

To include the UAV-based reference data, CC was quantified at the level of individual indexed PSB.SD pixels by aggregating the high-resolution UAV-derived canopy vectors. To estimate CC at the pixel level, the spatial overlap between each pixel footprint and the canopy polygons was determined using the `intersection` function from the `Shapely` library (Gillies et al., 2025), and the area of overlap was computed using its `area` method. The intersection area was then normalized by the total pixel area to derive a fractional CC value ranging from 0 to 1.

Vegetation Index Derivation

In addition to the eight bands, VIs derived from their combinations were also used as input for model development. A set of VIs was computed using reflectance data obtained from PSB.SD 8-band imagery. The reflectance values were originally scaled in the range of 0 to 10000. These were normalized to a [0, 1] range prior to any index computation to standardize the input data. The calculated indices include the NDVI (Rouse et al., 1973), EVI (Huete et al., 1997), and MSAVI (Qi et al., 1994). Each index was computed per pixel using established spectral relationships. NDVI was calculated as shown in Equation (3), while EVI was computed using Equation (4). MSAVI was used to minimize soil background effects and is defined in Equation (5).

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{Red}}{\mathrm{NIR} + \mathrm{Red}} \tag{3}$$

$$\mathrm{EVI} = 2.5 \times \frac{\mathrm{NIR} - \mathrm{Red}}{\mathrm{NIR} + 6 \times \mathrm{Red} - 7.5 \times \mathrm{Blue} + 1} \tag{4}$$

$$\mathrm{MSAVI} = \frac{2 \times \mathrm{NIR} + 1 - \sqrt{(2 \times \mathrm{NIR} + 1)^2 - 8 \times (\mathrm{NIR} - \mathrm{Red})}}{2} \tag{5}$$

Leaf Area Index Estimation Using Radiative Transfer Modeling

An RTM approach using the coupled PROSAIL model was employed to estimate LAI. PROSAIL integrates the PROSPECT model for leaf optical properties (Jacquemoud and Baret, 1990) and the SAIL model for canopy reflectance (Verhoef, 1984), enabling physically-based simulation of vegetation reflectance across spectral and geometric configurations (Jacquemoud et al., 2009).

To adapt the model, originally developed for Sentinel-2 data, for use with PSB.SD imagery, radiative transfer simulations were calibrated to the sensor's acquisition geometry and spectral response function. Specifically, the illumination azimuth angle (137.53°), illumination elevation angle (50.62°), and view azimuth angle (13.95°) were incorporated into the PROSAIL RTM. These parameters were extracted from PSB.SD scene metadata to ensure accurate representation of sun–sensor geometry during simulation. Furthermore, a Gaussian spectral response function was applied, and the reflectance data obtained from PSB.SD 8-band imagery was normalized to a [0, 1] range to ensure compatibility with the trained model. To further refine the forward simulations, a soil-informed hybrid retrieval approach was applied, following the unpublished method developed by Ledain et al. (2025). In this method, representative bare soil spectra were derived from Sentinel-2 imagery by Ledain et al. (2025) and colleagues using NDVI and Normalized Burn Ratio 2 (NBR2) thresholding within cultivated areas across Switzerland. These spectra were then clustered and interpolated to match the spectral resolution required by the PROSAIL model and subsequently incorporated as background inputs in the forward simulations. The resulting synthetic dataset, combining canopy parameters with realistic soil reflectance variability, was used to train an artificial neural network (ANN) that emulates PROSAIL inversion (Ledain et al., 2025). To ensure spectral consistency, the simulated spectra were resampled to match PSB.SD's spectral response functions. The trained model was then applied to normalized PSB.SD reflectance data to infer LAI.

The final PSB.SD-adapted LAI model demonstrated solid performance, with an R^2 of 0.757 and a normalized root mean squared error (nRMSE) of 0.149. The model predictions showed a clear linear relationship with observed LAI values up to approximately 3, followed by a typical saturation pattern at higher LAI levels as displayed in Figure 45 in the appendix.

3.2.4 Satellite Canopy Cover Model Development

The satellite CC model was developed using a pixel-level dataset comprising the eight multispectral bands from PSB.SD imagery, along with the derived VIs and LAI, as well as the UAV-based CC information. The objective was to train a model capable of predicting CC using only satellite inputs, allowing it to make accurate predictions at spatially independent sites.

To explore the effectiveness of different modeling strategies, both simple and advanced approaches were evaluated. This included the assessment of linear regression models based on selected VIs and LAI, as well as more complex ML and DL algorithms. Specifically, the performance of the RF, Histogram Gradient Boosting Regressor (HGBR), and Neural Network (NN) models was evaluated. To ensure comparability between the different approaches, a fixed random seed of 42 was used to ensure reproducibility. For the general model evaluation, the dataset was partitioned into training, validation, and test subsets using a stratified random split: 70% of the data was allocated to training, 20% to validation, and 10% to testing. Additionally, the model was tested iteratively on unseen sites to assess its ability to generalize to

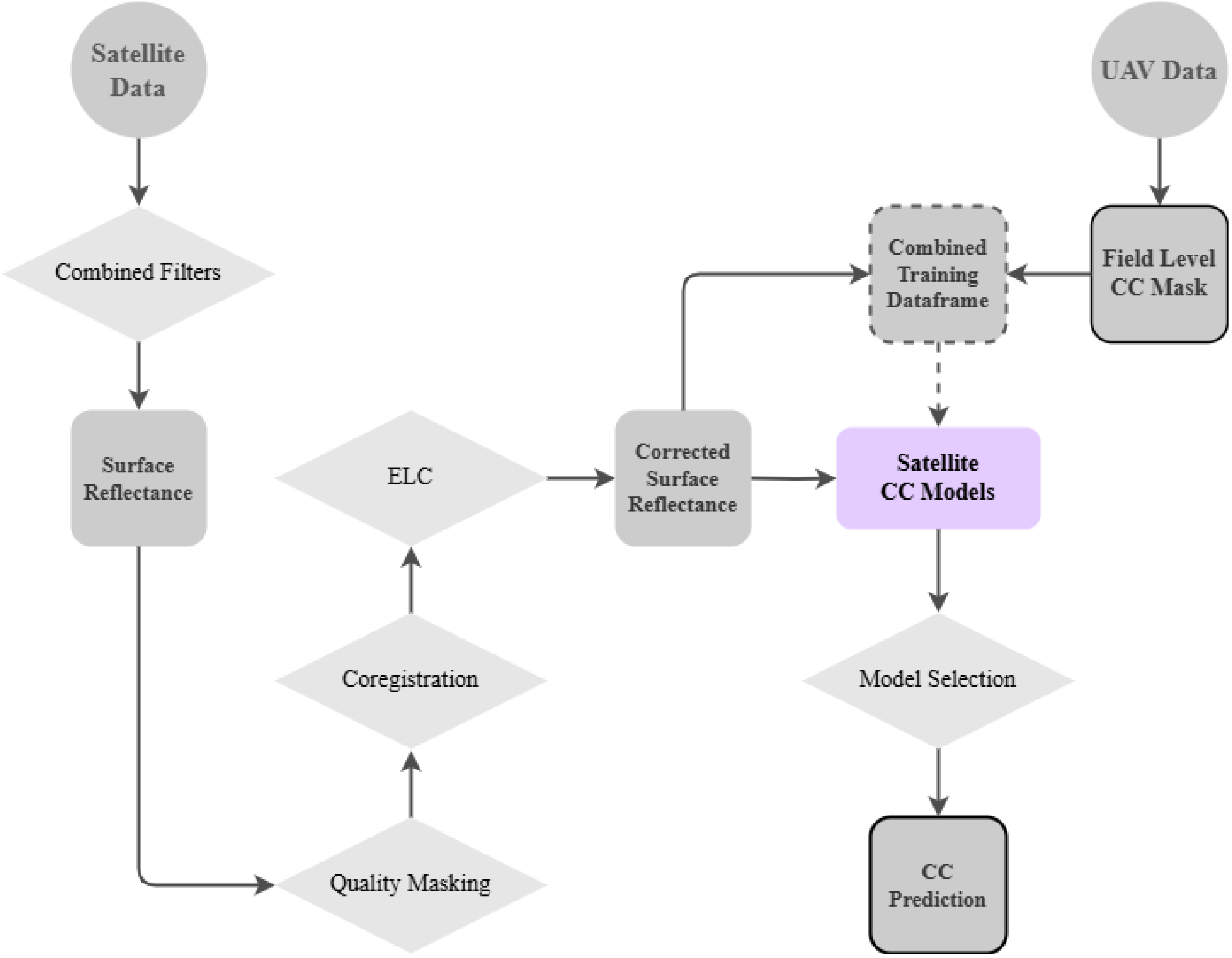


Figure 5. **Processing Pipeline for Satellite Canopy Cover Model**. This flowchart illustrates the processing pipeline for filtered satellite data, including quality masking, co-registration, and empirical line correction (ELC). The corrected surface reflectance is combined with UAV-based canopy cover (CC) masks to train various satellite CC models using the combined training data.

independent spatial and temporal contexts. In each iteration, the data from one flight was held out as the test set, while 80% of the remaining data were used for training. Model performance was assessed using three evaluation metrics: the coefficient of determination (R^2), root mean squared error (RMSE), and range-normalized RMSE (NRMSE), with the latter normalized by dividing the RMSE by the range of observed CC values. These metrics are defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{6}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{7}$$

$$\text{NRMSE} = \frac{\text{RMSE}}{y_{\text{max}} - y_{\text{min}}} \tag{8}$$

where y_i represents the observed CC values, $\hat{y}_i$ the predicted values, $\bar{y}$ the mean of the observed values, and y_{max} and y_{min} the maximum and minimum observed CC values, respectively.

Regression Models Based on Vegetation Indices and Leaf Area Index

A linear regression model was employed using either VIs or the LAI as input features. The UAV-based reference CC served as the response variable, while a single VI (NDVI, EVI, or MSAVI) or LAI was used as the predictor in each model instance. A standard linear regression model (Equation 9) was trained to predict CC values from the selected predictor:

$$y = ax + b + \varepsilon \tag{9}$$

where y denotes the predicted CC, x represents the selected VI or LAI feature, b is the intercept, a is the slope coefficient describing the linear relationship, and ε is the residual error term, assumed to be normally distributed with mean zero and constant variance.

Random Forest for Multispectral Canopy Cover Modeling

An RF regression model was trained on CC based on the eight normalized multispectral bands from PSB.SD imagery. The RF algorithm is an ensemble method that constructs multiple decision trees during training and outputs, which average the prediction of the individual trees, thus reducing overfitting and improving generalization (Pedregosa et al., 2011). The RF implementation was based on `scikit-learn's` `RandomForestRegressor` (Pedregosa et al., 2011). Model parameters were adapted and optimized to improve predictive accuracy while limiting model complexity. The number of trees in the ensemble was set to 200, and the maximum tree depth was restricted to 15, as summarized in Table 7.

Table 7. **Hyperparameters for the Random Forest Model**.

Parameter	Value
Number of trees	200
Maximum depth	15

Histogram-Based Gradient Boosting Regressor for Multispectral Canopy Cover Modeling

An HGBR was employed to model CC using the eight normalized multispectral bands from PSB.SD imagery. HGBR is a tree-based ensemble method optimized for efficiency and scalability, implementing gradient boosting over histogram-binned features to accelerate training and reduce memory usage (Pedregosa et al., 2011). This model is particularly well-suited for structured tabular data and offers strong predictive performance with built-in support for early stopping and regularization (Pedregosa et al., 2011). The model hyperparameters were adapted and optimized to achieve a balance between model complexity and generalization performance. The final configuration included a learning rate of 0.01, a maximum of 1000 boosting iterations, a tree depth limit of 15, and L2 regularization to mitigate overfitting. The model also employed early stopping based on a validation fraction of 10%. An overview of the selected hyperparameters is provided in Table 8.

Table 8. **Hyperparameters for the Histogram-Based Gradient Boosting Regressor**

Parameter	Value
Learning rate	0.01
Maximum iterations	1000
Maximum depth	15
Minimum samples per leaf	10
L2 regularization	0.1
Maximum leaf nodes	50
Validation fraction	0.1
Early stopping	Enabled

Neural Network for Multispectral Canopy Cover Modeling

An NN was developed to predict CC using the eight normalized multispectral bands from PSB.SD imagery. The network architecture consisted of an input layer corresponding to the number of input features, followed by three fully connected hidden layers, each with 256 neurons and `ReLU activation` functions. A final output layer produced the regression output. The model was implemented using the `PyTorch framework` (Paszke et al., 2019) and trained with the `Adam optimizer` and mean squared error (MSE) loss function. To prevent overfitting and ensure generalizability, early stopping was applied based on the validation loss, with a patience threshold of 50 epochs. The training process was carried out for a maximum of 1000 epochs, using a batch size of 256. In practice, training was typically stopped earlier once the validation loss plateaued, most commonly between 400 and 600 epochs. A summary of the NN configuration and training parameters is provided in Table 9.

Table 9. **Neural Network Architecture and Training Configuration**

Parameter	Value
Hidden layers	3
Neurons per layer	256
Activation function	ReLU
Loss function	Mean Squared Error (MSE)
Optimizer	Adam
Learning rate	0.001
Batch size	256
Max epochs	1000
Early stopping patience	50

3.2.5 Evaluation of Satellite-Based Canopy Cover Models

A two-tiered evaluation strategy was employed to evaluate the performance of the different modeling approaches, including linear regression, ML, and DL methods. This comprised overall performance testing based on randomly partitioned datasets and mission-specific cross validation designed to assess each model's capacity to generalize to temporally and geographically independent observations. Both evaluation stages were conducted using the pixel-level training datasets introduced in Section 3.2.3.

Baseline Evaluation Using Stratified Sampling

In the first stage of the evaluation, a stratified random split of the complete dataset was used to allocate 70% of the data for training, 20% for validation, and 10% for testing. This configuration provided a controlled setting to assess model performance under ideal conditions, where training data are assumed to be representative of the target domain. Performance was quantified using standard regression metrics, including the R^2, RMSE, and NRMSE, the latter of which normalizes the RMSE by the range of reference values to enable more robust comparisons between datasets of varying scales. The mean and median NRMSE values are summarized in Tables 11 and 12, respectively. For qualitative comparison, the predictions of the test set were plotted against the reference values and visualized as scatterplots (see Figures 10 - 12).

In addition to numerical evaluation, permutation-based feature importance analyses were conducted for the HGBR and NN models to assess the relevance of individual input variables. These analyses, shown in Figure 13, were implemented using the `permutation_importance` function from the `scikit-learn` library (Pedregosa et al., 2011).

Evaluating Generalization with Mission Cross Validation

The second stage of evaluation aimed to assess the spatial and temporal transferability of the modeling approaches. For this purpose, a mission cross validation framework was implemented, wherein data associated with a single acquisition mission were entirely excluded from training and validation and used solely for testing. This approach reflects practical application scenarios in which models must perform reliably in previously unseen environments. For each held-out mission, a separate model was trained on the remaining dataset and evaluated on the excluded mission. This enabled the assessment of model robustness under domain shift and provided insight into the extent to which each approach generalizes beyond its training distribution. Performance metrics were computed consistently with the overall testing phase, using RMSE and NRMSE. For qualitative comparison, the predictions were plotted against the reference values and visualized as scatterplots (see Figures 14 - 18).

Pixel-Level Validation of Canopy Cover Models Using UAV Reference

Spatial prediction accuracy at the pixel level was assessed by comparing cross-validated model outputs with UAV-based reference CC values for selected acquisition missions. For each pixel, the ratio between predicted and reference CC was calculated and log-transformed (base 10) to capture both over- and underestimations symmetrically. The spatial anomalies were visualized as heatmaps, depicting the CC (UAV) reference and the ratio of predicted CC to UAV-derived CC, which indicates deviations for the MSAVI-, HGBR-, and NN-based models. A diverging colormap centered at zero was used, with green shades representing overestimation and purple shades indicating underestimation, as shown in Figures 19 to 23.

Field-Scale Mean CC Comparison Using UAV Observations

To evaluate model performance at the field-level, mean CC predictions from the different modeling approaches were compared against UAV-derived mean CC reference values aggregated per acquisition mission. For each mission, the mean CC was computed from all pixels within the corresponding field, both for the UAV reference and for each cross-validated model prediction. Scatter plots were generated to visualize model agreement with the reference data, with individual points colored and shaped by crop type and location. Performance metrics, including the R^2, RMSE, and NRMSE, were calculated both overall and for each crop-location combination, as shown in Figures 24, 25, and 26.

3.3 Application on field experiments

The feasibility of the selected baseline models was assessed under independent real-world conditions by applying them to satellite-based time series data from the experimental site in Wallbach, covering the period from May 2, 2022, to June 28, 2024. Area-normalized CC values were computed for each plot by spatially intersecting raster-derived CC predictions with plot geometries. Raster data from each CC band (MSAVI, HGBR, NN, and UAV) were first vectorized into pixel-level polygons. For each plot defined in the shapefile, spatial intersections with these pixel polygons were calculated using the `intersection` function from the `shapely` library. Each raster value was weighted by the area of its intersection with the plot geometry, and the area-normalized CC was derived by dividing the sum of all weighted values by the total intersected area. This method ensured that partial pixel contributions along field boundaries were accurately represented in the final per-plot estimates. Plot-wise model predictions of CC were initially compared spatially to UAV-derived reference data acquired on April 25, 2024, providing a direct evaluation of prediction accuracy. Furthermore, interseasonal dynamics were analyzed by generating CC time series across a full crop rotation, including corn, winter wheat, and winter barley. This allowed for a comparative assessment of model performance across different treatment regimes of the ”Bio4Recycle” field phenotyping experiment.

Spatial Analysis of Plot Experiments

Plot-level predictions were visualized to assess the spatial distribution of predicted CC values at the experimental site. The analysis included CC estimates from three baseline models (MSAVI, HGBR, and NN) alongside UAV-based reference data collected on April 25 and 27, 2024, respectively. Three sets of spatial visualizations were generated to support this analysis. The first visualization, shown in Figure 27, presents the CC predictions from all models together with the UAV reference using a shared grayscale color scale. This allowed for a direct comparison of absolute CC values across treatments and models. The second set of maps, depicted in Figure 28, displays model-specific min-max normalized CC values on a 0 to 1 grayscale, enabling comparison of the internal spatial variability within each model's CC predictions. Finally, the plot-wise ratio between the predicted CC and UAV-based CC values was computed. The results of this comparison are visualized in Figure 29, using a diverging PRGn colormap centered at 1.0.

Inter-seasonal Timeseries Analysis

Time series plots were generated using area-normalized CC values derived from baseline model predictions to investigate treatment-wise CC dynamics across multiple seasons and crops. The analysis incorporated three modeling approaches (MSAVI, HGBR, and NN) and grouped treatments into six categories: NON, MIN, SLU, SD, SLA, and LID. For each model, CC predictions were averaged across all plots assigned to a given treatment group and visualized over time in Figure 30. The temporal scope spanned three growth seasons between 2022 and 2024, capturing a complete crop rotation of corn, winter wheat, and winter barley. To contextualize management interventions, key dates related to weed control and fertilizer applications, obtained from the ”Recycle4Bio” experimental records (Agostini et al., 2024), were overlaid as vertical dashed lines. Crop growth phases were further highlighted using shaded intervals to indicate the active growth periods in each season. To complement the season-wide overview, zoomed-in time series plots were generated for both the early and late growth phases of each crop. The early growth stage was defined as the period from seeding until CC reached its seasonal plateau, reflecting the phase of active vegetative growth (Figure 31). In contrast, the late growth stage encompassed the period from peak CC through to harvest (Figure 32). These zoomed plots enabled a more detailed evaluation of treatment differentiation within critical phenological stages.

4 RESULTS

This section presents the results of the CC modeling framework. As a first step, some segmentation outputs from the high-resolution and field-level CC model are presented. Then, the composition and distribution of the UAV-derived training dataset are described. The model selection outcome is outlined, and the performance of selected models under baseline conditions is compared. This is followed by a cross-validation analysis, in which model transferability is assessed through quantitative, qualitative, and spatial evaluations at pixel and field-levels. Finally, the models are applied to an independent experimental site, and their performance is evaluated through spatial and temporal analyses of CC.

4.1 Canopy Cover Segmentation

CC segmentation results are presented for both high-resolution and field-level CC models, highlighting their performance across crop types and spatial scales.

4.1.1 High-Resolution Canopy Cover

The high-resolution CC model demonstrated consistently high segmentation performance across all evaluated crops, as illustrated by the segmented RGB images in Figure 6. Fine structural details were accurately captured, including the aligned row patterns in sugarbeet and corn, as well as the broader, more clustered leaf structures in sunflowers. Notably, the model also performed well under visually complex conditions, such as the dense and irregular canopy of winter barley and the presence of weed interference in sugarbeet and sunflower plots. The proportion of misclassified pixels was very low, with only minor instances of noise such as lichens or green-tinged soil being incorrectly labeled as vegetation. On the other hand, some yellow plant compartments of winter barley are misclassified as soil.

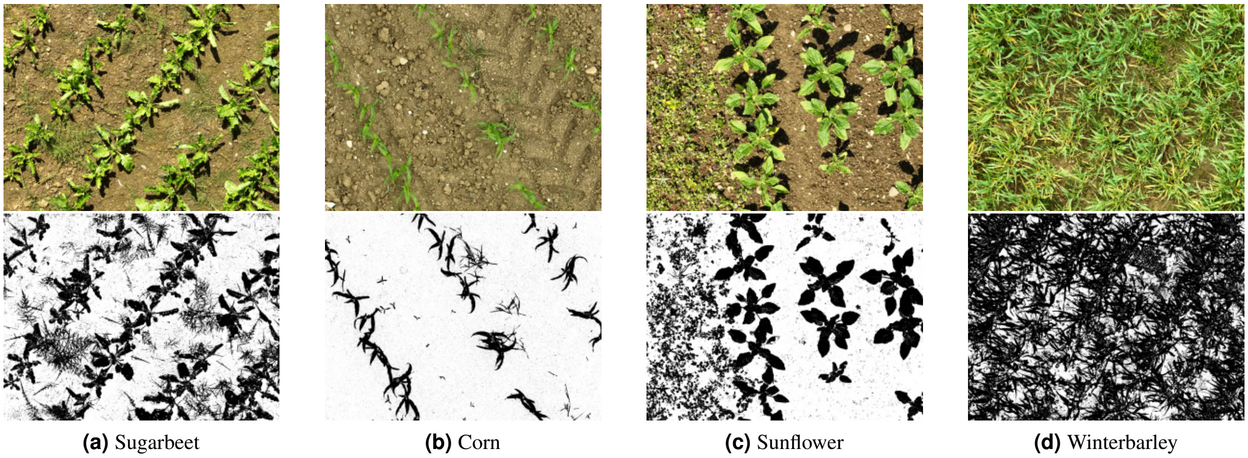
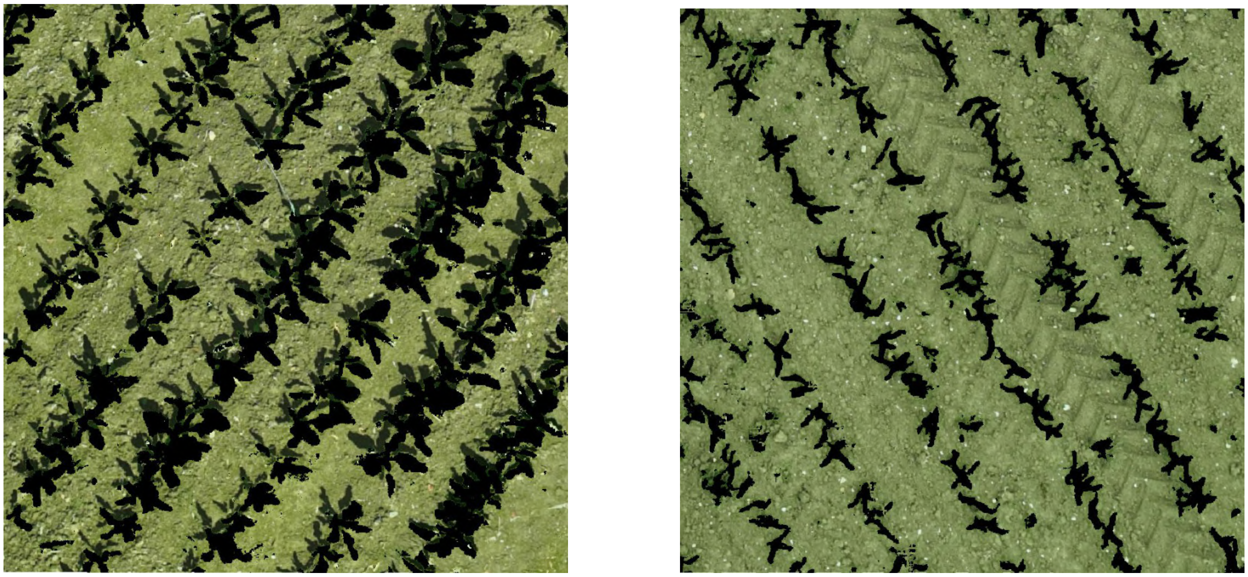

(a) Sugarbeet **(b)** Corn **(c)** Sunflower **(d)** Winterbarley

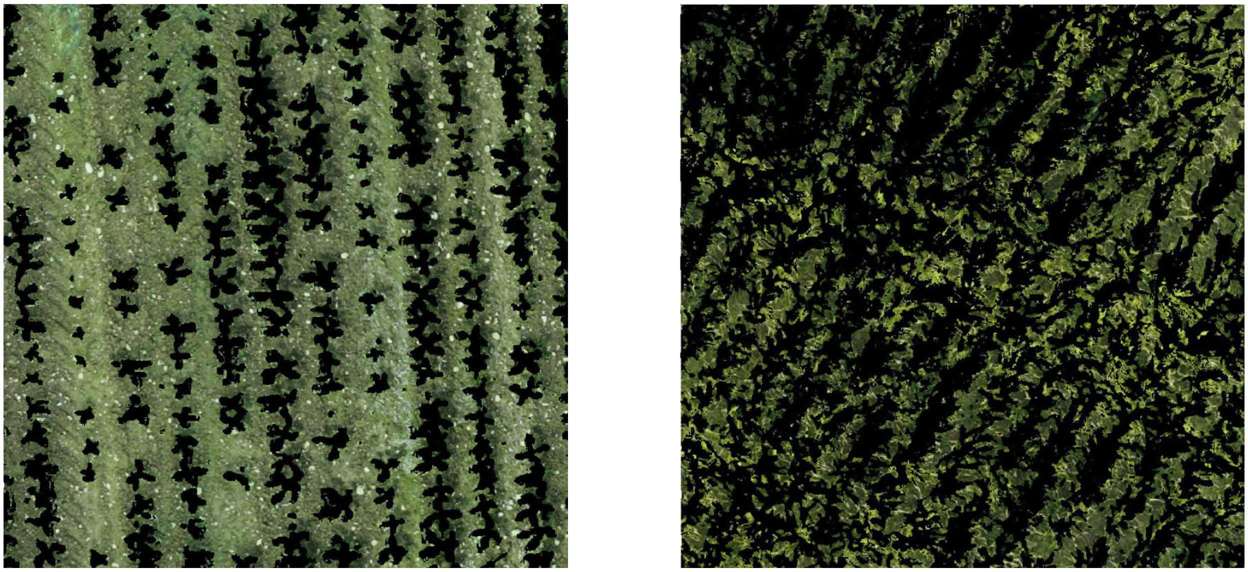
Figure 6. **Illustrations of High-Resolution Canopy Cover Segmentation**. UAV-derived snapshot (SNP) RGB images (top) and corresponding predicted canopy cover (CC) masks (bottom) for four crop types: (a) sugarbeet, (b) corn, (c) sunflower, and (d) winter barley. The masks were generated by the High-Resolution Canopy Cover Model, where black pixels indicate predicted vegetation and white pixels represent non-vegetation areas.

4.1.2 Field-Level Canopy Cover

The field-level CC model demonstrated strong performance across all evaluated crops, as illustrated in Figure 7. This robust performance of the site- and date-specific trained models is further substantiated by a mean overall accuracy of 0.9645, as reported in Table 20 in the appendix. Model validation revealed strong performance, with an overall accuracy of 96.5%, an F1 score of 0.91 for the plant class, and an Intersection over Union (IoU) of 0.83 for plant detection. Site-specific results are provided in the appendix (Table 20), showing consistent performance across different field conditions and locations. Distinct row and plant structures are clearly visible in all four crop types, demonstrating the model's ability to preserve spatial coherence even at lower resolution. In sugarbeet, corn, and sunflower plots, canopy structures are effectively distinguished from the surrounding soil, with minimal noise and high segmentation quality. In contrast, the model tends to underestimate canopy extent in winter barley (Figure 7d), which exhibits relatively high vegetation cover. This underestimation can be observed in areas with yellow-toned and thin plant parts that are not consistently classified as vegetation. The variation in model performance is also reflected in the mission-specific accuracies shown in Table 20 (appendix), where sugarbeet (0.9715), corn (0.9922), and sunflower (0.9549) display higher accuracy scores than winter barley (0.9412). Despite these differences, the model reliably captures the general canopy structure and the healthier vegetation components across all crops.



(a) Sugarbeet **(b)** Corn



(c) Sunflower **(d)** Winter barley

Figure 7. Illustrations of Field-Level UAV-Based Canopy Cover Mapping. UAV-based RGB mapping images overlaid with the predicted canopy cover (CC) masks from the Field-Level Canopy Cover Model for four crop types: (a) sugar beet, (b) corn, (c) sunflower, and (d) winter barley. Vegetation is shown in black, illustrating the model's performance in delineating canopy structure at the field scale.

4.2 Segmented Canopy Cover Training Dataframe

A large dataset consisting of aggregated UAV-derived CC data and preprocessed multispectral satellite observations was developed across multiple agricultural sites in Switzerland during the 2024 growth season. This pixel-level dataset served as the training data frame for the satellite-based CC modeling framework.

4.2.1 Data Availabilty

The training dataset comprises spatially distributed pixel-level observations of CC collected from UAV imagery across multiple agricultural sites. These data are linked with corresponding metadata, such as site location, crop type, field size, and acquisition date. Additionally, they are aligned with the spatial resolution of PSB.SD imagery, which is approximately 9 m^2 per pixel. A summary of the observed sites, crop types, number of annotated pixels, the total area covered, and UAV flight dates is are presented in Table 10. In total, the dataset comprises 31,429 pixels, which is equivalent to approximately 282,861 m^2, covering three crop types. Corn accounts for the largest share with 13,366 pixels, followed by sugarbeet which has 10,589 pixels, and sunflower which has 7878 pixels. Data were collected from field sites in Eschikon (20,031 pixels), Reckenholz (3,880 pixels), and Haerkingen (7,518 pixels). Field sizes vary considerably, ranging from 286 pixels (2,574 m^2) to 1,970 pixels (17,730 m^2), with Steinmueri1 representing the largest field and Chaeppeli the smallest, respectively.

Table 10. **Overview of UAV-based Data Collection Across Training Sites**. The table lists location names, specific field sites, associated crop types, and the total area covered by UAV imagery, expressed in $9\,\mathrm{m}^2$ pixel units. *UAV Acquisition Start* refers to the date and time at which each UAV mission began. Corresponding pixel counts are provided to reflect the temporal resolution of data acquisition at each site. The total covered area across all sites sums to 31,429 pixels, corresponding to $282,861\,\mathrm{m}^2$. Data © 2024 Planet Labs PBC.

Location	Site	Crop	Covered Area [in Pixels ($9\,\mathrm{m}^2$)]	Acquisition Start (UAV)
Eschikon	**Steinmueri1**	**Sunflower**	**7878**	
			2518	30.04.2024 12:44
			2500	14.05.2024 14:16
			1449	06.06.2024 14:19
			1411	19.06.2024 12:34
Eschikon	**Steinmueri3**	**Corn**	**5444**	
			1483	06.06.2024 14:19
			1385	19.06.2024 13:28
			1405	28.06.2024 10:56
			1171	09.07.2024 11:15
Eschikon	**Laubisgruet2**	**Sugarbeet**	**6709**	
			1408	14.05.2024 13:09
			1307	06.06.2024 13:07
			1333	19.06.2024 14:10
			1332	28.06.2024 11:58
			1329	09.07.2024 12:14
Reckenholz	**Eingang**	**Sugarbeet**	**3880**	
			881	19.05.2024 13:34
			825	04.06.2024 14:25
			830	12.06.2024 12:40
			515	27.06.2024 14:02
			829	15.07.2024 13:59
Haerkingen	**Chaeppeli**	**Corn**	**571**	
			283	05.06.2024 11:36

Continued on next page

Location	**Site**	**Crop**	**Covered Area [in Pixels ($9\,\mathrm{m}^2$)]**	**Flight Date (UAV)**
			288	18.06.2024 11:47
Haerkingen	**Neufeld4**	**Corn**	**4681**	
			1514	05.06.2024 12:09
			1582	18.06.2024 12:44
			1585	08.07.2024 11:36
Haerkingen	**Altgraben1**	**Corn**	**2670**	
			1319	05.06.2024 13:56
			1351	18.06.2024 10:53
		Total Covered Area	**31429**	

4.2.2 Data Distribution

To characterize the variability and representativeness of the training dataset, the distribution of CC values was analyzed across all crop types and acquisition sites. As illustrated in Figure 8, the cumulative distribution is skewed toward lower CC values, which reflects the study's focus on early-season vegetative development. The highest frequency of observations is concentrated between 0.0 and 0.05, corresponding to minimal CC. While this is consistent with the study's objectives, this range appears somewhat overrepresented and is a consequence of the increased number of UAV acquisitions conducted shortly after sowing. Among the three crop types, corn exhibits a relatively broad distribution of CC values, with a pronounced peak at low values and a gradual decline in frequency across mid to high CC intervals. Sugarbeet shows a similarly strong peak at low values but displays a more even spread across the 0.1–0.5 range. In contrast, sunflower presents a strong concentration of values within the 0–0.1 interval, with a notable reduction in sample frequency across the 0.1–0.4 range. Overall, the dataset is dominated by low CC values across all crops, with varying degrees of distributional spread observed at higher CC values.

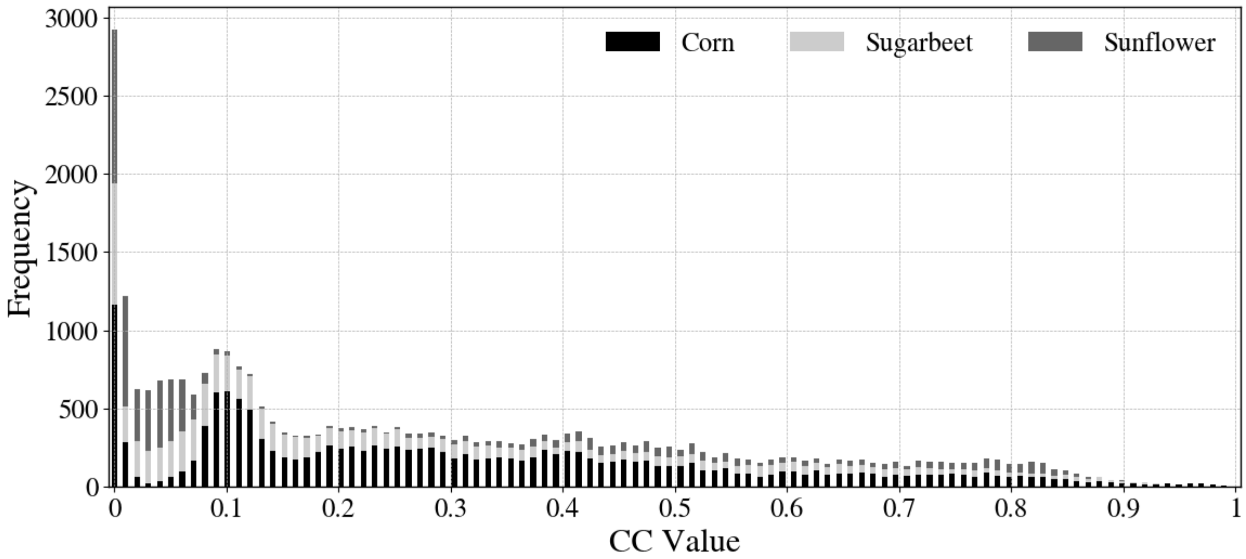
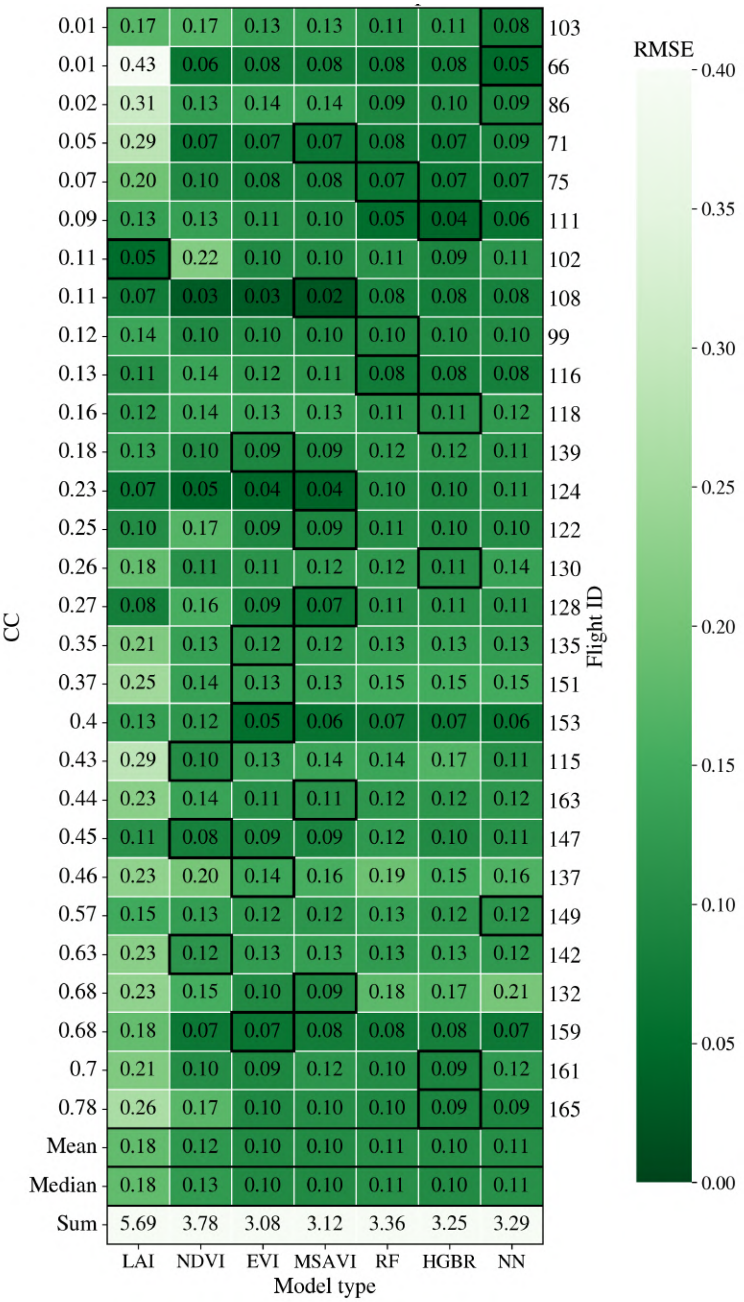


Figure 8. **Distribution of Crop-Specific Canopy Cover Values in the Training Dataset**. The histogram shows the frequency of canopy cover (CC) values ranging from 0 to 1 for different crop types, represented by distinct shades: corn (black), sugarbeet (light gray), and sunflower (dark gray).

4.3 Satellite Model Selection

Several modeling approaches were implemented to find the best modeling approach for the satellite CC model, including linear regression models using VIs (LAI, NDVI, EVI, MSAVI), ML algorithms (RF, HGBR), and a DL approach using a NN. From each category, the best-performing model was selected based on an error analysis of the mission cross-validation, with a particular focus on early growth stages. The overall performance comparison with RMSE across all models is summarized in Figure 9. Among the linear models, MSAVI was selected as it exhibited the lowest RMSE during early growth stages and was identified five times as the best-performing index for CC values below 0.3. Although EVI showed the lowest total RMSE sum across the full range of CC values, its performance was mainly superior at higher CC levels. By contrast, LAI showed the highest total RMSE by a large margin and consistently underperformed across all growth stages, making it the least suitable index. NDVI showed moderate performance, competing with other linear models primarily during mid to late growth stages but lacking reliability in the early season. For ML models, RF and HGBR showed similar behavior at early growth stages, but HGBR ultimately achieved a lower sum of RMSE values and was therefore selected for this study. The NN demonstrated stable performance across all growth stages and achieved the lowest RMSE for the three lowest mean CC bins, highlighting its strength in early canopy development phases. Based on these findings, MSAVI, HGBR, and NN were selected as the representative models for further analysis due to their consistently low RMSE values in early growth stages and overall robust performance across the CC spectrum.



CC	LAI	NDVI	EVI	MSAVI	RF	HGBR	NN	Flight ID
0.01	0.17	0.17	0.13	0.13	0.11	0.11	0.08	103
0.01	0.43	0.06	0.08	0.08	0.08	0.08	0.05	66
0.02	0.31	0.13	0.14	0.14	0.09	0.10	0.09	86
0.05	0.29	0.07	0.07	0.07	0.08	0.07	0.09	71
0.07	0.20	0.10	0.08	0.08	0.07	0.07	0.07	75
0.09	0.13	0.13	0.11	0.10	0.05	0.04	0.06	111
0.11	0.05	0.22	0.10	0.10	0.11	0.09	0.11	102
0.11	0.07	0.03	0.03	0.02	0.08	0.08	0.08	108
0.12	0.14	0.10	0.10	0.10	0.10	0.10	0.10	99
0.13	0.11	0.14	0.12	0.11	0.08	0.08	0.08	116
0.16	0.12	0.14	0.13	0.13	0.11	0.11	0.12	118
0.18	0.13	0.10	0.09	0.09	0.12	0.12	0.11	139
0.23	0.07	0.05	0.04	0.04	0.10	0.10	0.11	124
0.25	0.10	0.17	0.09	0.09	0.11	0.10	0.10	122
0.26	0.18	0.11	0.11	0.12	0.12	0.11	0.14	130
0.27	0.08	0.16	0.09	0.07	0.11	0.11	0.11	128
0.35	0.21	0.13	0.12	0.12	0.13	0.13	0.13	135
0.37	0.25	0.14	0.13	0.13	0.15	0.15	0.15	151
0.4	0.13	0.12	0.05	0.06	0.07	0.07	0.06	153
0.43	0.29	0.10	0.13	0.14	0.14	0.17	0.11	115
0.44	0.23	0.14	0.11	0.11	0.12	0.12	0.12	163
0.45	0.11	0.08	0.09	0.09	0.12	0.10	0.11	147
0.46	0.23	0.20	0.14	0.16	0.19	0.15	0.16	137
0.57	0.15	0.13	0.12	0.12	0.13	0.12	0.12	149
0.63	0.23	0.12	0.13	0.13	0.13	0.13	0.12	142
0.68	0.23	0.15	0.10	0.09	0.18	0.17	0.21	132
0.68	0.18	0.07	0.07	0.08	0.08	0.08	0.07	159
0.7	0.21	0.10	0.09	0.12	0.10	0.09	0.12	161
0.78	0.26	0.17	0.10	0.10	0.10	0.09	0.09	165
Mean	0.18	0.12	0.10	0.10	0.11	0.10	0.11	
Median	0.18	0.13	0.10	0.10	0.11	0.10	0.11	
Sum	5.69	3.78	3.08	3.12	3.36	3.25	3.29	

Model type

Figure 9. **Mission Cross-Validation Model Root Mean Square Error Comparison** across different UAV flight IDs with respective canopy cover (CC) mean. The heatmap displays root mean square error (RMSE) values for each model type (LAI, NDVI, EVI, MSAVI, RF, HGBR, and NN). Color intensity represents RMSE magnitude, with darker shades indicating lower errors. Summary statistics (mean, median, and sum) for each model type are provided at the bottom to support the overall performance comparison. Data © 2024 Planet Labs PBC.

4.4 Baseline Satellite Model Comparison

For the general model comparison, the performance of each model was evaluated based on a random split of the training dataset. The three modeling approaches, MSAVI, HGBR, and NN, are compared and presented in Figures 10 - 12. Each scatterplot shows predicted CC values plotted against UAV-derived reference data, with point densities color-coded to illustrate the distribution across the CC range. The MSAVI-based linear model (Figure 10), expressed in Equation 10, achieved an R^2 of 0.861 and an RMSE of 0.092. It performed comparably well in the mid to high CC range, exhibiting a clear linear relationship with the reference data. However, predictive performance declined at early growth stages (CC $<$ 0.2), where predicted CC values were widely spread and tended to deviate from the 1:1 line. This pattern highlights the model's limited ability to capture variability in early-stage canopy development.

$$CC_{\mathrm{MSAVI}} = -0.2456 + 1.5945 \times MSAVI \tag{10}$$

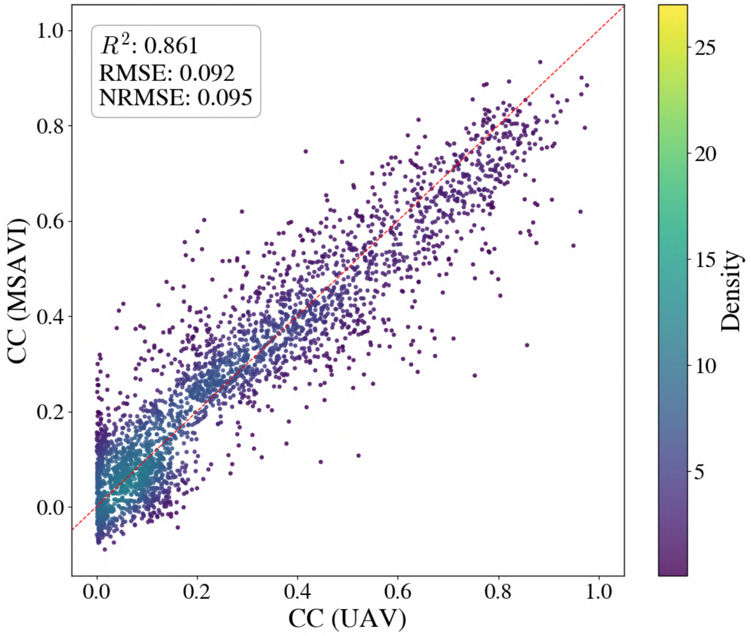


Figure 10. Performance of the Linear MSAVI Baseline Model. The scatter plot compares predicted CC values derived from the MSAVI index with reference UAV-based CC measurements. Point color represents data density, with higher density shown in yellow and lower density in purple. The red dashed line indicates the 1:1 reference line. Model performance metrics are displayed in the top-left corner: $R^2 = 0.861$, RMSE = 0.092, and NRMSE = 0.095. Data © 2024 Planet Labs PBC.

The HGBR model (Figure 11) showed improved accuracy, with an R^2 of 0.919 and an RMSE of 0.070. Predictions were more symmetrically distributed around the 1:1 line across the full CC spectrum. While the model slightly overestimated CC values at early growth stages, it still outperformed MSAVI substantially in the low CC range. An increasing spread of predicted values was observed with increasing CC, indicating some variability at higher canopy levels. As shown in Figure 13a, the model relied most heavily on the Red and NIR bands, while other spectral bands contributed only marginally to prediction outcomes.

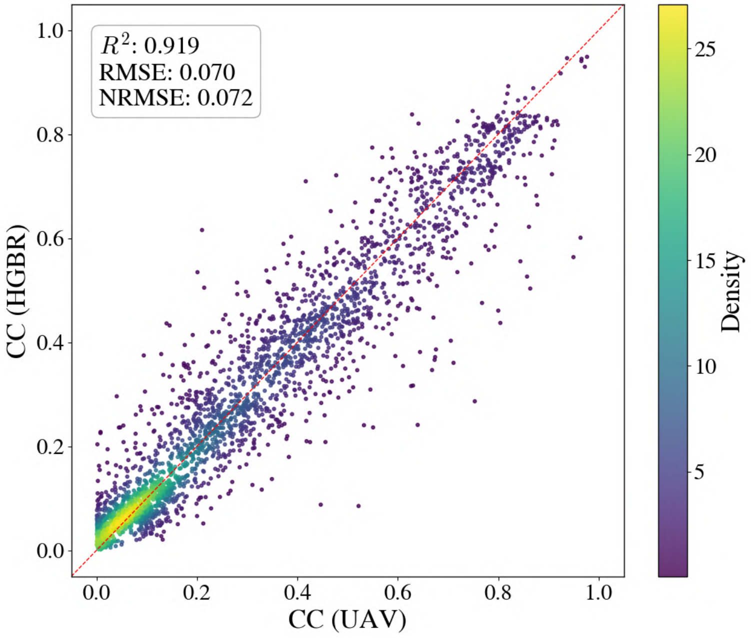


Figure 11. Performance of the Histogram Gradient Boosting Regressor Baseline Model. The scatter plot compares the Histogram Gradient Boosting Regressor (HGBR) model's canopy cover (CC) predictions with reference UAV-based CC measurements. Point color represents data density, with higher density shown in yellow and lower density in purple. The red dashed line indicates the 1:1 reference line. Model performance metrics are displayed in the top-left corner: $R^2 = 0.919$, RMSE = 0.070, and NRMSE = 0.072. Data © 2024 Planet Labs PBC.

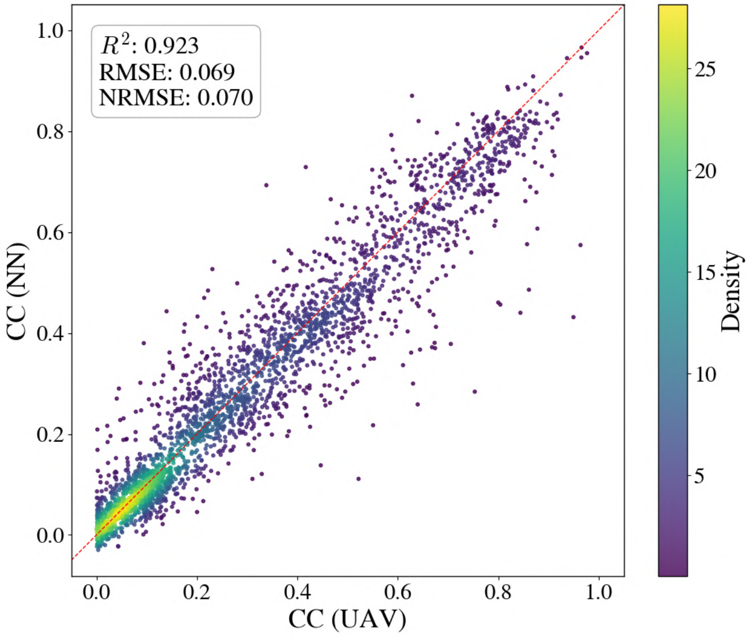
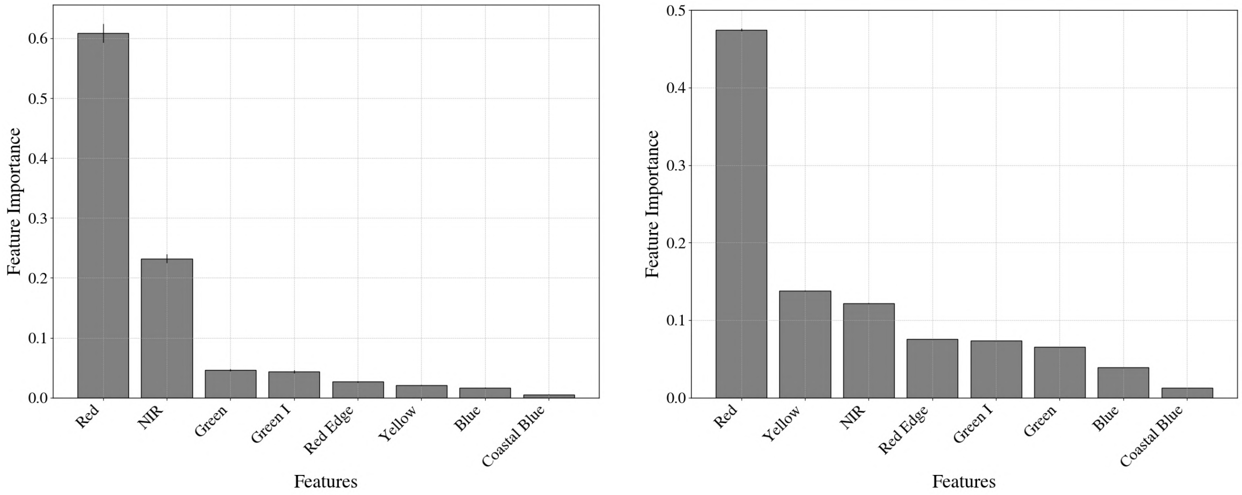


Figure 12. Performance of the Neural Network Baseline Model. The scatter plot compares the Neural Network (NN) model's canopy cover (CC) predictions with reference UAV-based CC measurements. Point color represents data density, with higher density shown in yellow and lower density in purple. The red dashed line indicates the 1:1 reference line. Model performance metrics are displayed in the top-left corner: $R^2 = 0.923$, RMSE = 0.069, and NRMSE = 0.070. Data © 2024 Planet Labs PBC.

The NN model (Figure 12) achieved the highest overall performance, with an R^2 of 0.923 and an RMSE of 0.069. Predictions were clustered along the 1:1 line across all CC values. Notably, the NN model achieved superior accuracy in the early growth phase (CC < 0.2), where other models performed less reliably. At higher CC values, the spread in predictions was comparable to that of the HGBR model. An examination of the feature importance in Figure 13b reveals that the Red band was by far the most influential, followed by moderate contributions from the Yellow and NIR bands. Notably, the Red Edge band exhibited a greater impact in the NN model compared to the HGBR model.



(a) Permutation feature importances (HGBR)

(b) Permutation feature importances (NN)

Figure 13. Permutation Feature Importances for Baseline Models. (a) Permutation feature importances for the Histogram Gradient Boosting Regressor (HGBR) model, and (b) Permutation feature importances for the Neural Network (NN) model. The bar charts display the mean importance of each feature, with the features sorted by their importance. Data © 2024 Planet Labs PBC.

In summary, while the MSAVI model showed acceptable performance in moderate to high CC conditions, it underperformed during the early stages. HGBR offered more balanced and accurate predictions across the full CC range, particularly at low values. The NN model outperformed both, demonstrating the highest accuracy across all growth stages.

4.5 Mission Cross Validation Satellite Model Comparison

A mission-based cross-validation strategy was employed to assess the transferability of the models across independent missions. In each iteration, the predicted mission was excluded from the training set and used solely for testing. Performance evaluation is based on the NRMSE, which was preferred over RMSE as it accounts for the scale of the target variable. In addition to the quantitative analysis, a qualitative and spatial visual evaluation was conducted to identify and consider special cases that may not be adequately captured by statistical metrics alone.

4.5.1 Quantitative Pixelwise Comparison

Table 11 presents the mean NRMSE values for three models (MSAVI, HGBR, and NN) applied to three crop types (corn, sugarbeet, and sunflower). The table also includes overall means across crops and models. Among the crops, sugarbeet showed the lowest average NRMSE (0.187), indicating consistently strong model performance. Sunflower followed closely with a mean NRMSE of 0.194. Corn exhibited the highest mean NRMSE (0.359), suggesting greater difficulty. The comparison with the median reveals that corn has the lowest median NRMSE across crops, indicating the presence of a few high-error cases that increased the mean. The model performance per crop shows minimal variation among models for corn. NN achieved the lowest mean NRMSE (0.353), followed by MSAVI (0.356) and HGBR (0.367). In contrast, sugarbeet showed lower NRMSE values but greater variation, with HGBR performing best (0.175), NN following (0.183), and MSAVI showing the highest error (0.203). For sunflower, NN again performed best (0.174), followed by MSAVI (0.191), while HGBR had the highest NRMSE (0.217). While NN outperformed the other models in terms of mean NRMSE, the comparison with the median values reveals an interesting contrast. MSAVI exhibits lower median NRMSEs than NN across all crops, suggesting that it performs well in the majority of cases. However, MSAVI's higher mean NRMSE indicates rare failure cases that inflate the average error. Comparing mean values of NRMSE across all crops, NN exhibits the lowest overall mean NRMSE (0.237), followed by MSAVI (0.250) and HGBR (0.253).

Table 11. Comparison of Baseline Models Performance with Mean Range-Normalized Root Mean Squared Error (NRMSE) across three baseline models (MSAVI, HGBR, and NN) and crop types. Crop-wise and model-wise NRMSE means are included to highlight performance differences both across crops and among the evaluated models. Data © 2024 Planet Labs PBC.

Crop	MSAVI	HGBR	NN	Crop Mean
Corn	0.356	0.367	0.353	0.359
Sugarbeet	0.203	0.175	0.183	0.187
Sunflower	0.191	0.217	0.174	0.194
Model Mean	0.250	0.253	0.237	–

Table 12. Comparison of Baseline Models Performance with Median Range-Normalized Root Mean Squared Error (NRMSE) across three baseline models (MSAVI, HGBR, and NN) and crop types. Data © 2024 Planet Labs PBC.

Crop	MSAVI	HGBR	NN
Corn	0.144	0.161	0.177
Sugarbeet	0.162	0.177	0.176
Sunflower	0.153	0.215	0.186

4.5.2 Qualitative Pixelwise Comparison

For the qualitative evaluation, a selection of representative cases was used to illustrate the strengths and weaknesses of the different models.

Early Growth Stage

Figure 15 displays model performance for an early growth phase (CC around 0.1) sugarbeet field, where crop-specific and site-specific training data was available. In this case, the MSAVI model performs the worst, significantly underestimating CC and even yielding negative values. The NN and especially the HGBR perform better, achieving RMSE values of 0.061 and 0.044, respectively. Furthermore, HGBR predictions align more closely with the 1:1 line, whereas NN results appear more scattered.

Figure 14 shows an early-season corn site where only limited and qualitatively poor training data were available. In

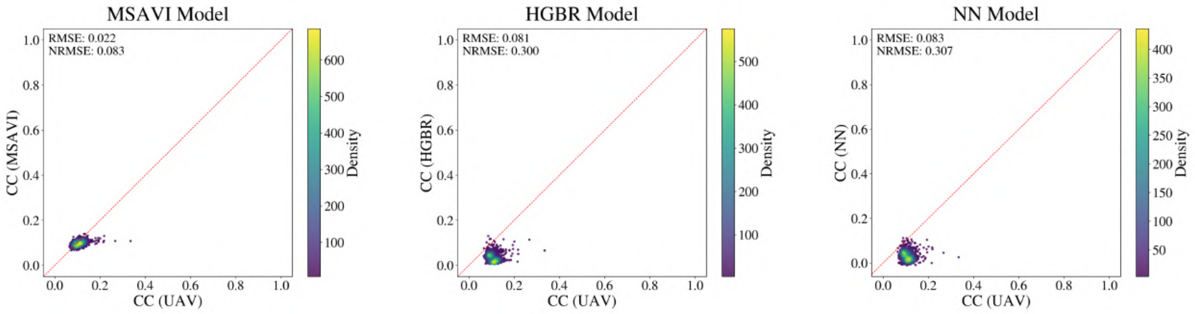


Figure 14. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Haerkingen, Altgraben1 Site (Corn) on June 5, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.022, NRMSE = 0.083; HGBR: RMSE = 0.081, NRMSE = 0.300; NN: RMSE = 0.083, NRMSE = 0.307. Each model was trained on 23,251 samples and tested on 1,144 samples as part of a mission-based cross-validation setup, with data from this flight held out entirely for testing. Data © 2024 Planet Labs PBC.

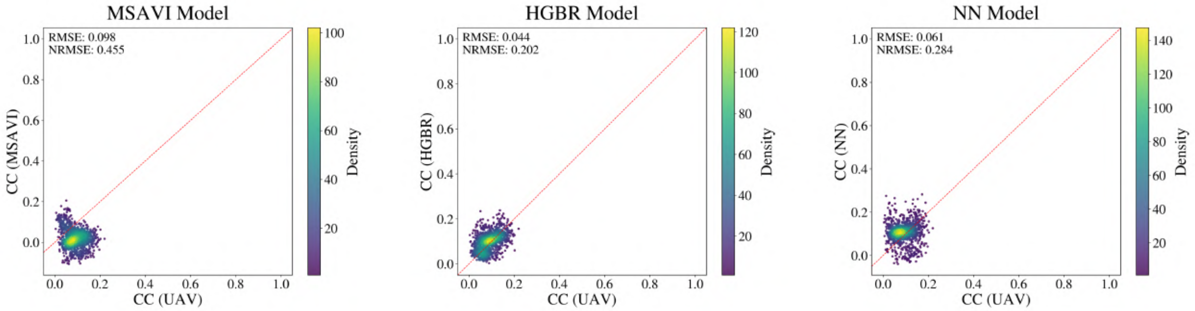


Figure 15. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Laubisgruet2 Site (Sugarbeet) on June 6, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.098, NRMSE = 0.455; HGBR: RMSE = 0.044, NRMSE = 0.202; NN: RMSE = 0.061, NRMSE = 0.284. Each model was trained on 23,232 samples and tested on 1,252 samples as part of a mission-based cross-validation setup, with data from this flight held out entirely for testing. Data © 2024 Planet Labs PBC.

this case, the MSAVI-based model performed best. It achieved an RMSE of 0.022, while both HGBR and NN had higher RMSE values around 0.08 and tended to underestimate CC. The majority of MSAVI predictions exhibit a linear relationship closely following the 1:1 line. In contrast, the HGBR and NN outputs show two distinct clusters of underestimated values, along with a generally wider spread in predictions.

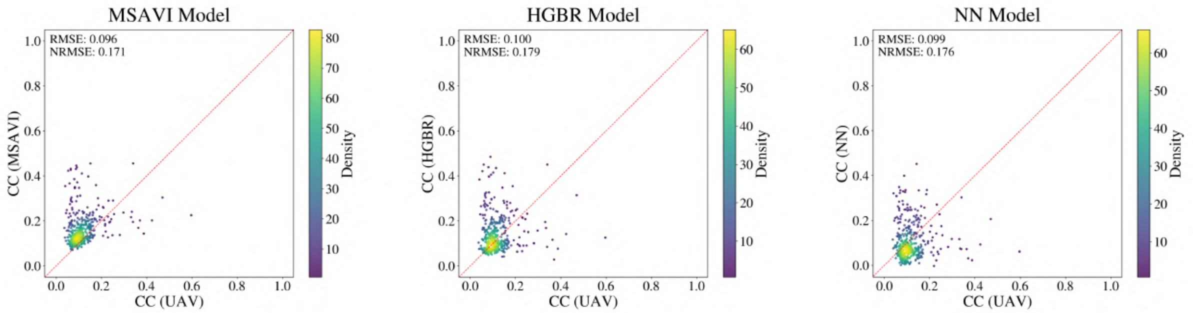


Figure 16. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Reckenholz, Eingang Site (Sugarbeet) on June 4, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.096, NRMSE = 0.171; HGBR: RMSE = 0.100, NRMSE = 0.179; NN: RMSE = 0.089, NRMSE = 0.176. Each model was trained on 23,088 samples and tested on 407 samples as part of a mission-based cross-validation setup, with data from this flight held out entirely for testing. Data © 2024 Planet Labs PBC.

At this early growth stage of sugarbeet (CC around 0.1), all three models demonstrate solid predictive performance (Figure 16), as indicated by RMSE values close to 0.1. The MSAVI model slightly overestimates the central cluster of observations, whereas the NN tends to underestimate it. Among the three, the HGBR model's cluster center aligns most closely with the 1:1 line in this low CC range, although it displays a broader vertical spread, suggesting greater variability in its predictions. In this example, a striking horizontal and especially vertical pattern is evident in the scatter plots, highlighting instances of systematic under- and overestimation.

In summary, across different crops and sites at the early growth stage, no model consistently over- or underestimates CC for specific conditions. MSAVI performs best when training data is limited, while HGBR and NN excel with larger datasets. Notably, the Reckenholz Eingang site exhibits the greatest variability in model predictions.

Later Growth stage

At later growth stages, the training dataset was generally more complete due to increased data availability (Figure 4), particularly for corn. This provides a valuable opportunity to assess model performance under more favorable conditions and compare outcomes across different crops.

For a late growth stage corn site (CC $\approx$ 0.7), Figure 17 demonstrates that the NN model achieves the highest predictive accuracy, with an RMSE of 0.07. In comparison, the linear MSAVI model exhibits a slight tendency to overestimate CC (RMSE = 0.078), while the HGBR model shows greater variability and yields the least accurate performance (RMSE = 0.082). Notably, the HGBR predictions display a distinct horizontally striped pattern, suggesting the presence of systematic prediction artifacts. This issue has been consistently observed across multiple missions, as displayed in Figures 56, 62, 67, 68, 70 in the appendix.

Sunflower was cultivated exclusively at the Steinmueri1 site, making it the only location in the dataset lacking crop-specific reference data under high CC conditions. As a result, predictions for this site represent a fully independent application of the models to an unseen and challenging scenario. Figure 18 highlights substantial differences in prediction performance across model types. The MSAVI-based model (RMSE = 0.094) produces estimates that closely follow the 1:1 line. In contrast, both the NN (RMSE = 0.206) and HGBR (RMSE = 0.166) models systematically overestimate CC, with their predictions exhibiting greater dispersion in the scatterplots.

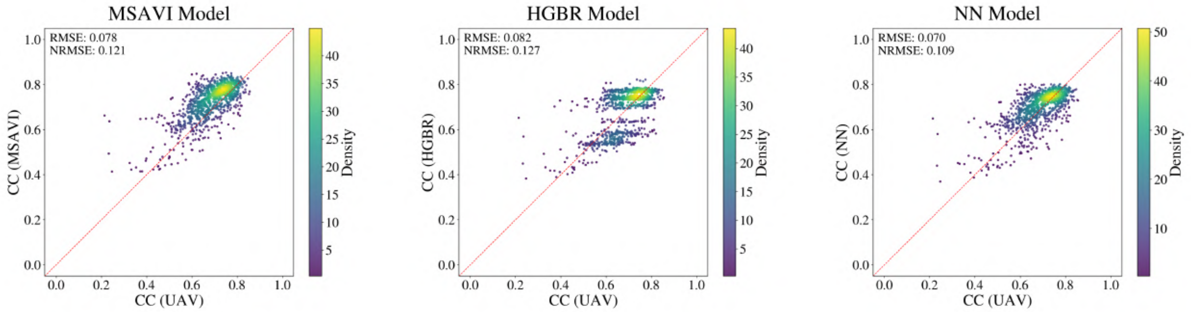


Figure 17. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Steinmueri3 Site (Corn) on July 9, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.078, NRMSE = 0.121; HGBR: RMSE = 0.082, NRMSE = 0.127; NN: RMSE = 0.070, NRMSE = 0.109. Each model was trained on 23,645 samples and tested on 1,055 samples as part of a mission-based cross-validation setup, with data from this flight held out entirely for testing. Data © 2024 Planet Labs PBC.

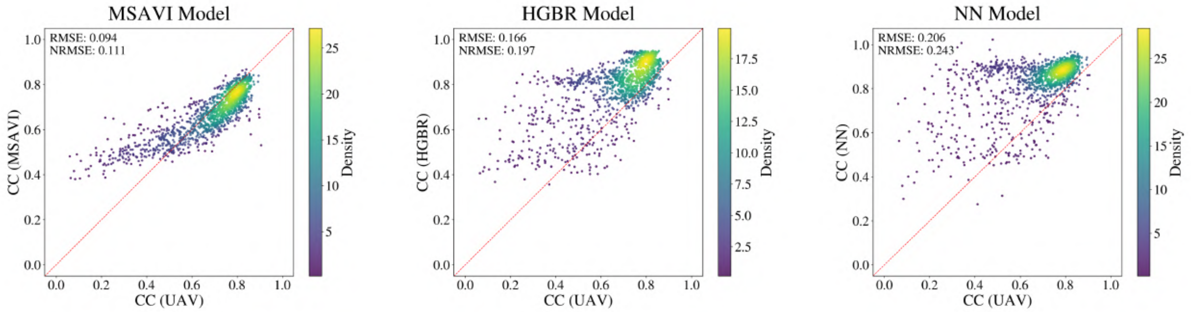


Figure 18. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Steinmueri1 Site (Sunflower) on June 19, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.094, NRMSE = 0.111; HGBR: RMSE = 0.166, NRMSE = 0.197; NN: RMSE = 0.206, NRMSE = 0.243. Each model was trained on 23,179 samples and tested on 1,318 samples as part of a mission-based cross-validation setup, with data from this flight held out entirely for testing. Data © 2024 Planet Labs PBC.

All in all, the models exhibit distinct characteristics across site- or crop-specific scenarios at later growth stages. The NN model delivers the best performance when sufficient training data is available, whereas the MSAVI-based model demonstrates stable and reliable behavior even without crop-specific training data. In contrast, the HGBR model shows the least consistent performance, with a tendency to produce horizontally striped artifacts, suggesting systematic prediction issues.

4.5.3 Spatial Comparison

To further explore spatial patterns in model predictions, I analyzed the pixel-wise deviation between model-predicted CC and UAV-derived reference data for the same missions presented in Section 4.5.2

Early Growth stage

Figures 19–23 show the ratio of model-predicted CC to UAV-observed CC for the MSAVI, HGBR, and NN models. A value of 1 (white) indicates perfect agreement, values below 1 (purple) denote underestimation, and values above 1 (green) signify overestimation. The UAV-based reference CC is visualized in inverted grayscale, where black corresponds to high CC.

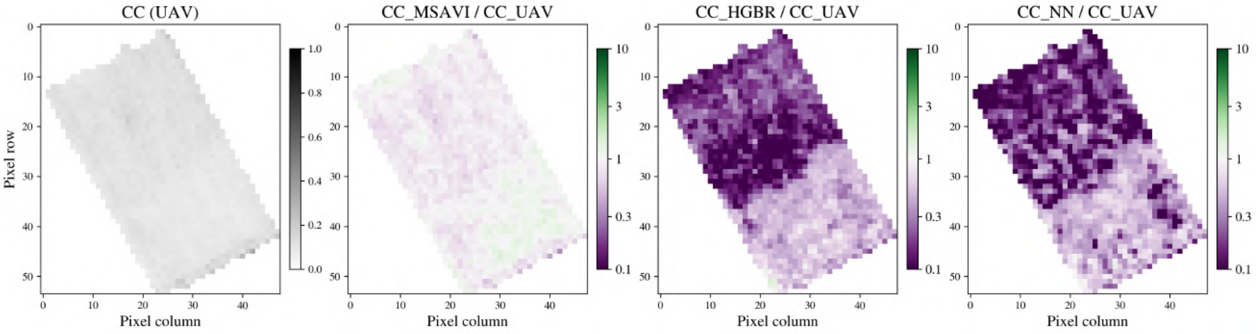


Figure 19. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Haerkingen, Altgraben1 Site (Corn) on June 5, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio ¿ 1), whereas purple tones indicate underestimation (ratio < 1). Data © 2024 Planet Labs PBC.

At Altgraben1 (Haerkingen), the UAV-derived CC is generally low and uniform, with slightly higher values observed in the upper portion of the field. The MSAVI-based model exhibits the most balanced performance. Its deviation map is predominantly bright, with only subtle patches of purple and green. The model slightly underestimates CC in the denser upper region and slightly overestimates it in the sparser lower part of the field. In contrast, both the HGBR and NN models show a tendency to systematically underestimate CC. The two rightmost images in Figure 19 reveal a distinct boundary separating the upper section, where CC is strongly underestimated, from the lower section, which exhibits smaller deviations from the reference. This pattern is more pronounced and spatially coherent in the HGBR output, while the NN predictions are more scattered. In this example, the differences between models are clearly distinguishable, and no artifacts attributable to satellite imagery are apparent.

Figure 20 shows that the UAV-derived CC for sugarbeet at Laubisgruet (Eschikon) is moderately high across the field, as indicated by the grayscale reference image (left). Some low CC management stripes and a localized patch of lower CC are visible around pixel row 20 and column 18. The MSAVI-based model shows substantial underestimation in both the left and right sections of the field, though these deviations do not appear spatially related to the underlying

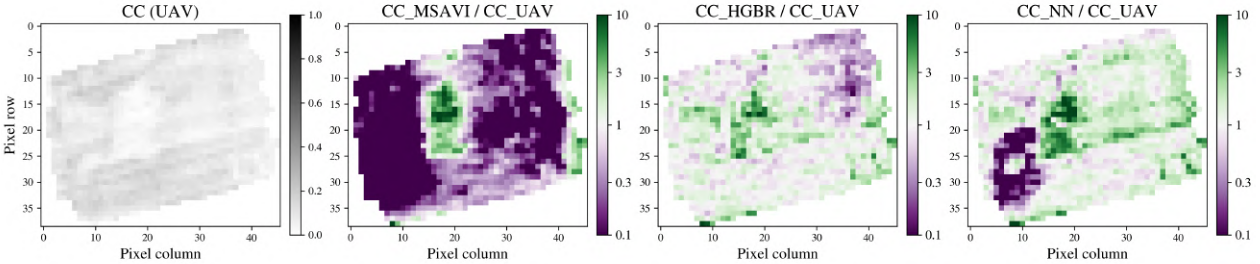


Figure 20. **Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Laubisgruet2 Site (Sugarbeet) on June 6, 2024**. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio > 1), whereas purple tones indicate underestimation (ratio < 1). Data © 2024 Planet Labs PBC.

CC distribution. A central region of strong overestimation stands out distinctly. This patch of overestimation also appears in the subplot of the HGBR model, although less pronounced. Generally, the HGBR model presents a mixed pattern of light purple and green deviations distributed across the field. While areas of overestimation can generally be associated with lower CC in the reference, the underestimation pattern appears more randomly distributed, lacking a clear relationship to the reference CC values. The NN model exhibits the strongest tendency to overestimate CC, with widespread green tones across the field. These overestimations are mostly concentrated in areas of low reference CC. A distinct circular region of strong underestimation is evident in the lower-left corner of the field, an artifact not observed in the outputs of the other two models. All three models display spatially inconsistent prediction behavior across the field. However, all outputs share two recurring spatial artifacts. The first is a localized patch of overestimation in the center of the field. The second is a narrow margin, approximately 1–2 pixels wide, of overestimation along the lower and left field boundaries.

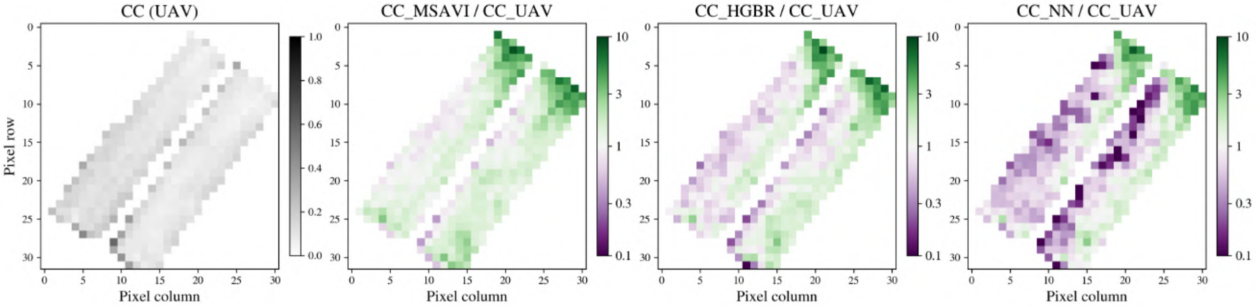


Figure 21. **Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Reckenholz, Eingang Site (Sugarbeet) on June 4, 2024**. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio > 1), whereas purple tones indicate underestimation (ratio < 1). Data © 2024 Planet Labs PBC.

At the Eingang site around Reckenholz, UAV-derived CC for sugarbeet was spatially uniform across the two parallel planting strips, as shown in the reference image (left) of Figure 21. The grayscale tones indicate consistent CC values between 0.1 and 0.2, with minimal variation between rows. A few higher CC pixels are scattered along the field edges. The MSAVI-based model displays a consistent tendency toward overestimation, particularly along the upper-right and right-edge regions of the field (green hues). While the deviations are generally moderate, the pattern appears systematic, with only a few isolated areas of slight underestimation (light purple) near the left edges. The HGBR model exhibits a similar overestimation trend in the upper regions but shows a more distinct spatial pattern. In particular, underestimations along the left-facing portions of the strips are more extensive and pronounced compared to MSAVI. Notably, the NN model produces the most distinct and structured pattern. A clear diagonal boundary separates areas of overestimation and underestimation, running from the upper left to the lower right corner of each planting strip. In addition, the NN predictions show stronger local deviations, indicating higher spatial variability compared to MSAVI and HGBR. All three models share a consistent overestimation artifact along the top edge of the field and a tendency to underestimate along the left margins. While the MSAVI model demonstrates the most spatially stable behavior, all models exhibit a mix of over- and underprediction despite the relatively uniform reference CC conditions.

Later Growth stage

At the Steinmueri3 site in Eschikon, CC for corn was observed at a relatively advanced growth stage. As shown in Figure 22, the UAV-derived reference image indicates generally high CC across the field, with values mostly above 0.7. Slight spatial variation is visible, characterized by vertical striping patterns.

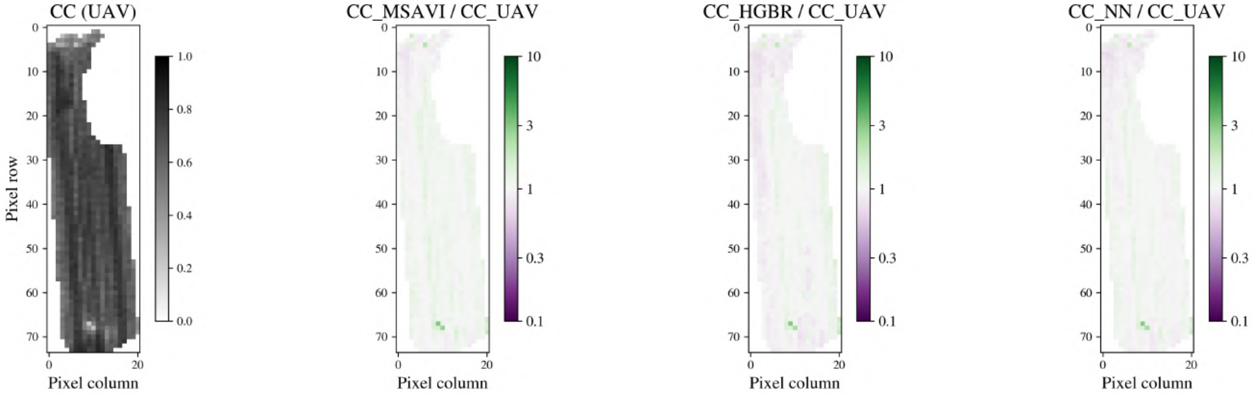


Figure 22. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Steinmueri3 Site (Corn) on July 9, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV reference CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio > 1), whereas purple tones indicate underestimation (ratio < 1). Data © 2024 Planet Labs PBC.

All three models exhibit similar prediction behavior in this high-CC scenario. In each case, the deviation maps are dominated by near-white tones, indicating close agreement with the UAV reference. Only minor patches of over- or underestimation appear, mainly aligned with the vertical stripe structures. The MSAVI model shows a very subtle tendency to overestimate CC in localized areas. In contrast, the HGBR model exhibits a slightly more pronounced overestimation pattern compared to the other models. The NN model demonstrates the most consistent prediction performance, with the lowest overall deviation from the reference. Overall, all three models show stable and spatially coherent prediction behavior. Compared to other sites, Steinmueri3 stands out for its uniform and accurate model performance, with only minimal systematic errors or spatial artifacts.

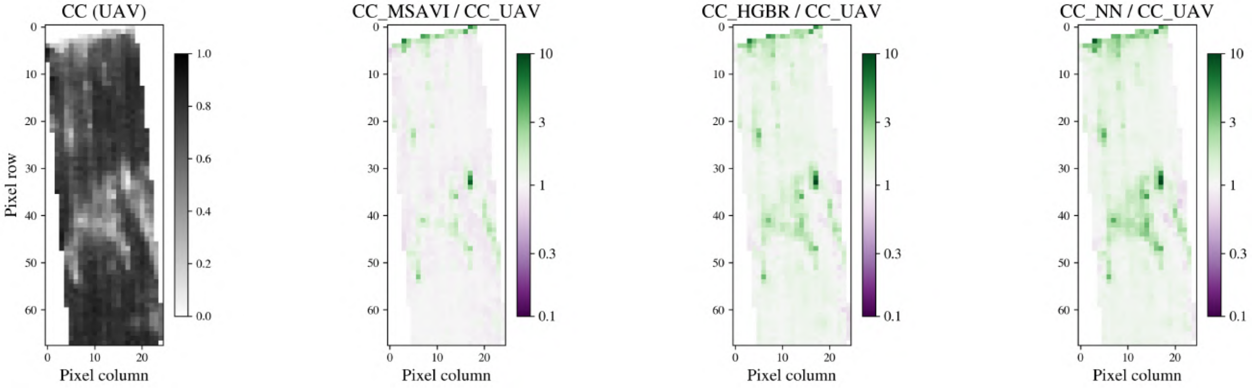


Figure 23. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Steinmueri1 Site (Sunflower) on June 19, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV reference CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio > 1), whereas purple tones indicate underestimation (ratio < 1). Data © 2024 Planet Labs PBC.

At the Steinmueri1 site in Eschikon, CC for sunflower was assessed under conditions representing an advanced growth stage. As shown in Figure 23, the UAV-derived reference map reveals generally high CC across the field, with most values around 0.8. Distinct spatial variability is present, with scattered patches of lower density in the lower central parts of the field. The MSAVI-based model shows relatively subtle deviations from the reference, with a general tendency towards underestimation. Overestimations are primarily located along the upper edge of the field and in a few pixels within the low-density regions. Both the HGBR and NN models exhibit a general overestimation tendency, with stronger deviations than MSAVI, particularly in the center areas with low CC and along the upper edge of the field. Overall, all three models perform well for this mission but fail to fully capture the lower CC observed in the sparse central region. Among them, the MSAVI model provides the closest approximation. The overestimation along the upper edge is a consistent artifact across all three models.

4.5.4 Field-Level Comparison

In this section, model performance is evaluated at the field-level by comparing the mean predicted CC for each mission against the corresponding reference mean values derived from UAV observations. This analysis provides insight into how well each model captures site-level variation across different crops and growth stages.

Overall, outcomes of all three models demonstrate strong agreement with the UAV-derived field-level means, as reflected in high coefficients of determination (R^2) and low error metrics. Specifically, MSAVI achieved an R^2 of 0.933, HGBR reached 0.934, and NN exhibited the highest agreement with an R^2 of 0.940. The RMSE values range from 0.057 to 0.060, while NRMSE values lie between 0.073 and 0.077, with the NN model consistently showing the lowest error across both metrics (see Tables 13, 14, and 15).

Model performance varies depending on crop type and site, as illustrated in the scatterplots of field-level means

(Figures 24, 25, and 26). Different crops are represented by distinct marker shapes, while sites are displayed by color. The MSAVI model performed particularly well for sunflower and corn at Eschikon, with R^2 values of 0.931 and 0.921, respectively. However, visual analysis reveals that the MSAVI model struggles to predict CC values at early growth stages(CC $<$ 0.2).

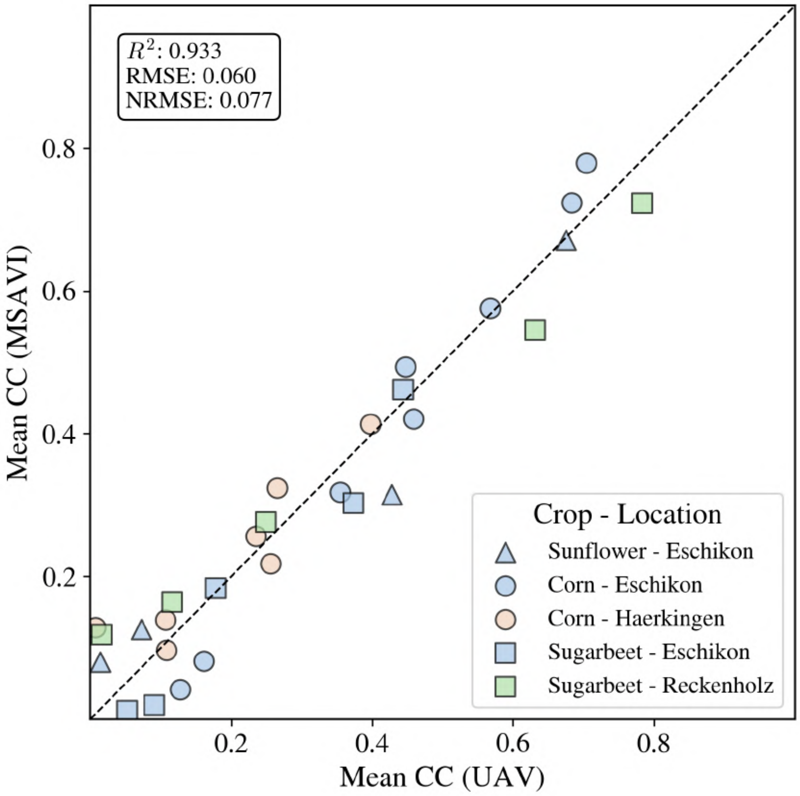


Figure 24. **Comparison of Field-Level Canopy Cover MSAVI Model Predictions to UAV Reference**. The scatter plot shows the mean canopy cover (CC) from UAV data on the x-axis and the mean CC derived from the MSAVI index on the y-axis for each location and crop. Crop types are represented by different shapes (Triangle = Sunflower, Circle = Corn, Square = Sugarbeet), and locations are represented by different colors (Blue = Eschikon, Orange = Haerkingen, Green = Reckenholz). The dashed line represents the 1:1 line, indicating perfect agreement between the model and reference. Model performance metrics are displayed in the top-left corner: $R^2 = 0.933$, RMSE = 0.060, and NRMSE = 0.077. Data © 2024 Planet Labs PBC.

Table 13. **Model Performance Metrics for Field-Level Canopy Cover Predicted by the MSAVI Model Compared to UAV-Reference**. The table shows the R^2, RMSE, and NRMSE for different crop types and locations. The mean performance metrics, calculated across all locations, are provided in bold. Data © 2024 Planet Labs PBC.

Crop - Location	R^2	RMSE	NRMSE
Sunflower - Eschikon	0.931	0.071	0.107
Corn - Eschikon	0.921	0.057	0.099
Corn - Haerkingen	0.790	0.055	0.142
Sugarbeet - Eschikon	0.900	0.049	0.125
Sugarbeet - Reckenholz	0.944	0.070	0.091
Mean	**0.933**	**0.060**	**0.077**

The HGBR model exhibited its strongest performance for corn at Eschikon (R^2 = 0.962) and sugarbeet at Reckenholz (R^2 = 0.970). However, performance declined for corn at Haerkingen (R^2 = 0.694). Visual patterns also indicate that for Eschikon, the predictive strength of the HGBR model increases with CC, thereby contributing significantly to the high R^2 value.

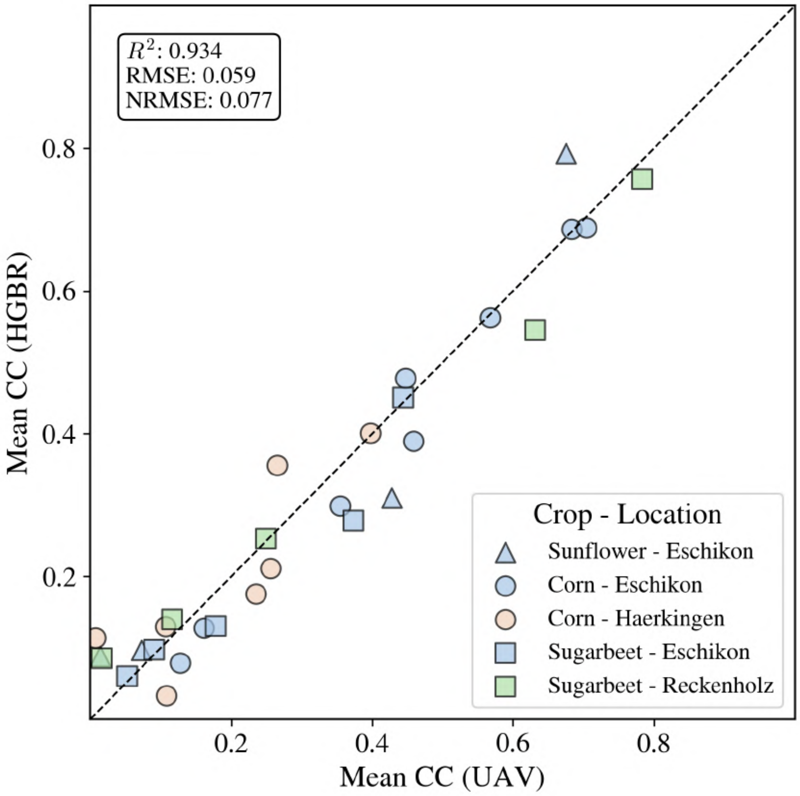


Figure 25. Comparison of Field-Level Canopy Cover HGBR Model Predictions to UAV Reference. The scatter plot shows the mean canopy cover (CC) from UAV data on the x-axis and the mean CC derived from the Histogram Gradient Boosting Regressor (HGBR) model on the y-axis for each location and crop. Crop types are represented by different shapes (Triangle = Sunflower, Circle = Corn, Square = Sugarbeet), and locations are represented by different colors (Blue = Eschikon, Orange = Haerkingen, Green = Reckenholz). The dashed line represents the 1:1 line, indicating perfect agreement between the model and reference. Model performance metrics are displayed in the top-left corner: $R^2 = 0.934$, RMSE = 0.059, and NRMSE = 0.077. Data © 2024 Planet Labs PBC.

Table 14. Model Performance Metrics for Field-Level Canopy Cover Predicted by the HGBR Model Compared to UAV-Reference. The table shows the R^2, RMSE, and NRMSE for different crop types and locations. The mean performance metrics, calculated across all locations, are provided in bold. Data © 2024 Planet Labs PBC.

Crop - Location	R^2	RMSE	NRMSE
Sunflower - Eschikon	0.884	0.092	0.139
Corn - Eschikon	0.962	0.040	0.069
Corn - Haerkingen	0.694	0.067	0.171
Sugarbeet - Eschikon	0.904	0.048	0.122
Sugarbeet - Reckenholz	0.970	0.052	0.068
Mean	**0.934**	**0.059**	**0.077**

The NN model demonstrated the highest overall performance, achieving the best R^2 (0.990) and lowest NRMSE (0.039) for sugarbeet at Reckenholz. Despite slightly lower performance in corn at Haerkingen (R^2 = 0.673, NRMSE = 0.177), the NN model consistently displayed strong predictive ability. The visual scatter analysis further supports this, showing predictions closely aligned with the 1:1 line across most sites.

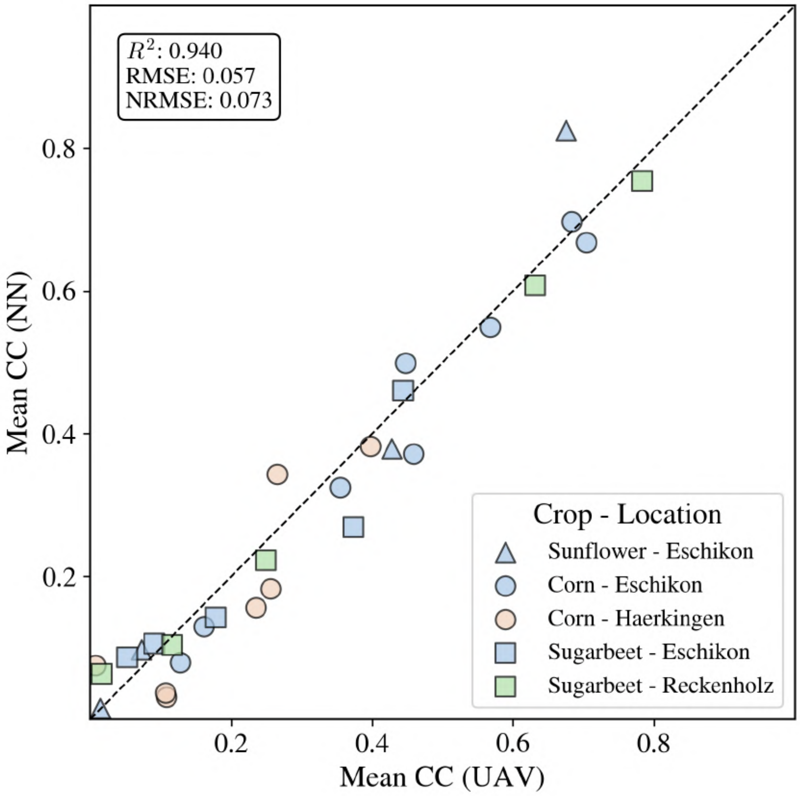


Figure 26. **Comparison of Field-Level Canopy Cover NN Model Predictions to UAV Reference**. The scatter plot shows the mean canopy cover (CC) from UAV data on the x-axis and the mean CC derived from the Neural Network (NN) model on the y-axis for each location and crop. Crop types are represented by different shapes (Triangle = Sunflower, Circle = Corn, Square = Sugarbeet), and locations are represented by different colors (Blue = Eschikon, Orange = Haerkingen, Green = Reckenholz). The dashed line represents the 1:1 line, indicating perfect agreement between the model and reference. Model performance metrics are displayed in the top-left corner: $R^2 = 0.940$, RMSE = 0.057, and NRMSE = 0.073. Data © 2024 Planet Labs PBC.

Table 15. **Model Performance Metrics for Field-Level Canopy Cover Predicted by the NN Model Compared to UAV-Reference**. The table shows the R^2, RMSE, and NRMSE for different crop types and locations. The mean performance metrics, calculated across all locations, are provided in bold. Data © 2024 Planet Labs PBC.

Crop - Location	R^2	RMSE	NRMSE
Sunflower - Eschikon	0.912	0.080	0.121
Corn - Eschikon	0.950	0.045	0.078
Corn - Haerkingen	0.673	0.069	0.177
Sugarbeet - Eschikon	0.885	0.053	0.134
Sugarbeet - Reckenholz	0.990	0.030	0.039
Mean	**0.940**	**0.057**	**0.073**

Among crop-specific differences, it is observed that sugarbeet is, on average, best predicted by the HGBR and NN models. In contrast, MSAVI proved to be most accurate in predicting sunflower. During the growth season, both HGBR and NN captured low CC values more effectively than MSAVI. On the other hand, MSAVI showed better performance under conditions with limited crop-specific training, such as in the case of high CC sunflower at Eschikon, where predictions from HGBR and NN deviated more substantially from the UAV reference. Certain missions, such as corn at Haerkingen during the early growth stage, consistently appear difficult for all models. However, the MSAVI model outperformed both HGBR and NN, achieving a higher R^2 value. All in all, the three models showed strong overall agreement with UAV-derived mean CC, with the NN model achieving the highest accuracy. However, performance varied by crop and site, with MSAVI excelling under limited training data, while HGBR and NN performed better in low CC scenarios.

4.6 Application on Experimental Site

To evaluate the feasibility of each model in an independent, real-world scenario, the three models were applied to predict CC based on satellite time series data collected between May 2, 2022, and June 28, 2024, at the experimental site in Wallbach. In the following analysis, these plot-wise predictions are first compared with UAV-derived reference data obtained on a specific date. Subsequently, the performance of the different models is assessed through a comprehensive time series analysis within a given crop rotation.

4.6.1 Spatial Analysis

The three models were successfully applied to predict satellite-based CC for the experimental site in Wallbach. Figures 27, 28, and 29 display the CC predictions, range-normalized CC and relative CC per plot, respectively, for the MSAVI, HGBR, and NN models. In each case, the leftmost subplot presents the UAV-derived reference data collected on April 25 2024.

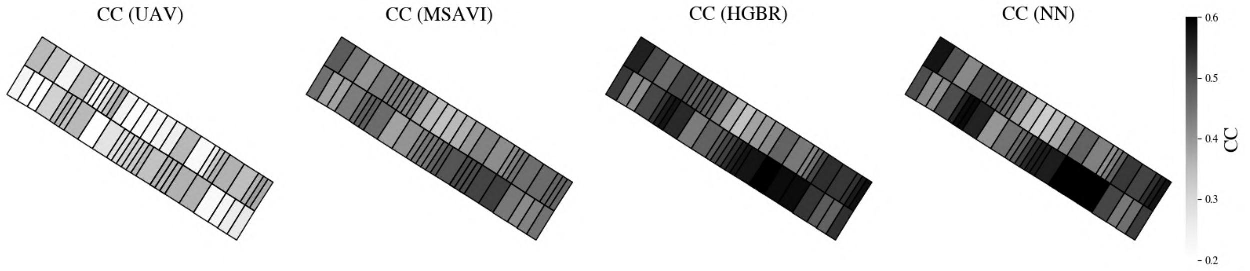


Figure 27. Canopy Cover Predictions from Baseline Models and UAV Reference for Experimental Site at Wallbach on 27.04.2024. The figure shows the plot-wise canopy cover (CC) UAV reference data (left) and the predictions from the three baseline models (MSAVI, HGBR, and NN). The color scale ranges from white (CC = 0.2) to black (CC = 0.6). Data © 2024 Planet Labs PBC.

The plot-wise CC predictions shown in Figure 27 reveal that, overall, all three models predict higher CC values than the UAV reference, as indicated by the generally darker shading in most plots. This observation is supported quantitatively by the treatment-level mean values summarized in Table 16. The UAV-derived mean CC across treatments is 0.296, while model-derived means are higher: 0.429 for MSAVI, 0.501 for HGBR, and 0.491 for NN. Despite this consistent overestimation, the spatial pattern of high and low CC plots broadly mirrors the UAV reference, suggesting that the models capture the overall CC structure of the field. However, the UAV reference indicates that NON treatment has the lowest CC mean (0.214) while all the models predict SD treatment as the lowest CC (mean = 0.416). The same pattern is visible for the maximal values, where the UAV reference reveals the highest values for MIN (0.355), and the models predict SLA slightly higher (mean = 0.509). Among the model outputs, the MSAVI-based predictions with a mean of 0.429 most closely resemble the UAV reference (0.296) in terms of magnitude but display a lower degree of contrast between plots. The HGBR model provides a more accentuated contrast, with predominantly higher predictions (0.501), especially in the dense CC plots. The NN model with a predicted CC mean of 0.491 follows a pattern similar to HGBR, but appears to distinguish high CC plots more clearly. One spatial inconsistency stands out in the fourth plot from the right in the lower row of Figure 27, where the UAV reference indicates low CC, while all models predict high CC. A similar pattern can be observed in the top left corner, where CC predictions are substantially higher than the UAV reference. In summary, while all models tend to overestimate CC compared to UAV measurements, they successfully replicate the general spatial distribution of CC. HGBR and NN offer stronger plot-level contrast, whereas MSAVI aligns more closely with the UAV reference in terms of overall magnitude.

Table 16. Mean Canopy Cover Values Derived from Baseline Models and UAV Reference for Experimental Site. The table provides the mean canopy cover (CC) values derived from UAV data and three different models (MSAVI, HGBR, and NN) for each fertilizer treatment in the plot experiment ”Bio4Recycle” at Wallbach on 27.04.2024. The treatment-wise mean is calculated across all models, excluding the UAV-derived values. The final row presents the overall mean CC for each method. Data © 2024 Planet Labs PBC.

Treatment	CC (UAV)	CC (MSAVI)	CC (HGBR)	CC (NN)	Mean (Treatment)
NON	0.214	0.411	0.468	0.440	**0.440**
SD	0.231	0.400	0.433	0.414	**0.416**
SLA	0.348	0.445	0.541	0.542	**0.509**
SLU	0.301	0.438	0.516	0.508	**0.487**
LID	0.326	0.436	0.521	0.505	**0.487**
MIN	0.355	0.444	0.529	0.536	**0.503**
Mean (Model)	**0.296**	**0.429**	**0.501**	**0.491**	**0.474**

To better understand the spatial distribution of CC, the plot-wise CC data were normalized using the model-specific minimum and maximum values and displayed in grayscale in Figure 28. The UAV reference exhibits the highest degree of contrast between plots, highlighting its superior spatial resolution. Nevertheless, all three models show distinct differences between plots that generally follow the spatial pattern observed in the reference data. For the smaller plots with a width of 3 m, the models were unable to capture the same level of detail as the UAV reference. Interestingly, when comparing the lower row of plots in the MSAVI subfigure to those in the HGBR and NN subfigures, MSAVI appears to capture small variations more accurately. The primary difference between the HGBR and NN models lies in their representation of high CC values. While both follow similar spatial trends, the NN model tends to be less discriminative in regions with dense CC. Overall, the normalized CC visualization confirms that the general spatial pattern is well captured by all three models. However, finer-scale variation in the smaller plots remains a challenge.

To identify areas with substantial disagreement between model predictions and the UAV reference, relative CC values were calculated and visualized using a divergent color scale, where higher green intensity indicates stronger

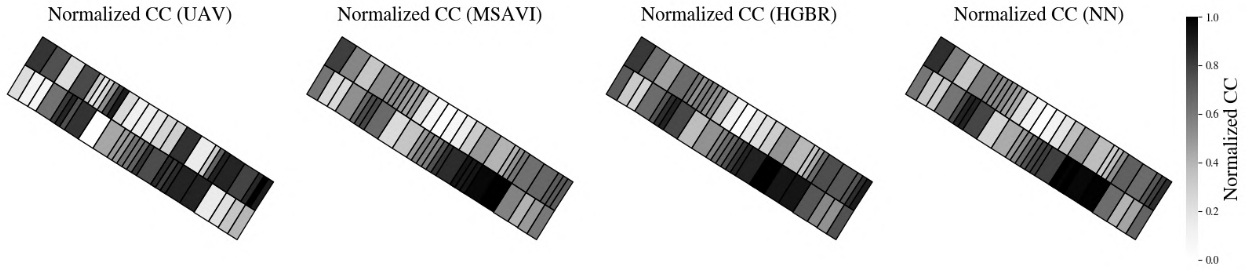


Figure 28. Normalized Canopy Cover Predictions from Baseline Models and UAV Reference for Experimental Site at Wallbach on 27.04.2024. The figure shows plot-wise range-normalized canopy cover (CC) UAV reference data (left) and the predictions from the three baseline models (MSAVI, HGBR, and NN). The color scale is normalized to a range of 0 (white) to 1 (black). Data © 2024 Planet Labs PBC.

overestimation. As illustrated in Figure 29, all models exhibit a general tendency to overestimate CC. Table 17 reveals the highest overestimation for the control plots (*NON*), with a mean relative value of 2.056 across all models. Notably, three plots located in the upper left corner of the field, corresponding to the sparsely vegetated ”SD” treatment, also display pronounced overestimation. In contrast, the lowest mean deviation from the UAV reference is found in the densely vegetated ”MIN” treatment (1.418). Among the models, HGBR shows the highest overall overestimation (1.738), while MSAVI reports the lowest (1.497).

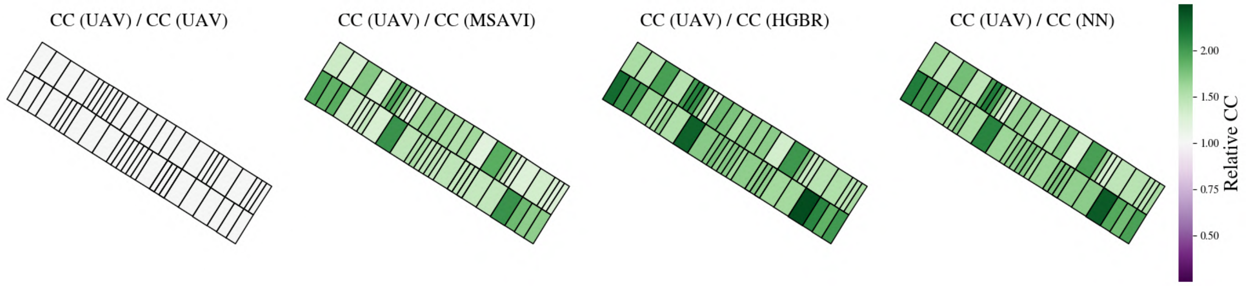


Figure 29. Relative Canopy Cover Predictions from Baseline Models Compared to UAV Reference for Experimental Site at Wallbach on 27.04.2024. The figure shows plot-wise relative canopy cover (CC) UAV reference data (left) and the predictions from the three baseline models (MSAVI, HGBR, and NN), with a color scale ranging from purple (underestimation) to dark green (overestimation). The color intensity represents the magnitude of relative CC values for each model's prediction compared to the UAV reference. Data © 2024 Planet Labs PBC.

Treatment	CC (MSAVI) / CC (UAV)	CC (HGBR) / CC (UAV)	CC (NN) / CC (UAV)	Mean (Treatment)
NON	1.923	2.188	2.058	**2.056**
SD	1.734	1.879	1.793	**1.802**
SLA	1.277	1.555	1.556	**1.463**
SLU	1.457	1.715	1.687	**1.620**
LID	1.339	1.600	1.549	**1.496**
MIN	1.251	1.491	1.511	**1.418**
Mean (Model)	**1.497**	**1.738**	**1.692**	**1.642**

Table 17. **Ratio of Model-Predicted Canopy Cover to UAV-Derived Values at Experimental Site**. The table provides the ratio of model-predicted canopy cover (CC) values to UAV-derived CC measurements for each fertilizer treatment in the plot experiment ”Bio4Recycle” at Wallbach on 27.04.2024. The table shows how much higher each model’s predictions are relative to UAV reference data. The treatment-wise mean represents the average ratio across all three models. The final row shows the average ratio per model across all treatments. Data © 2024 Planet Labs PBC.

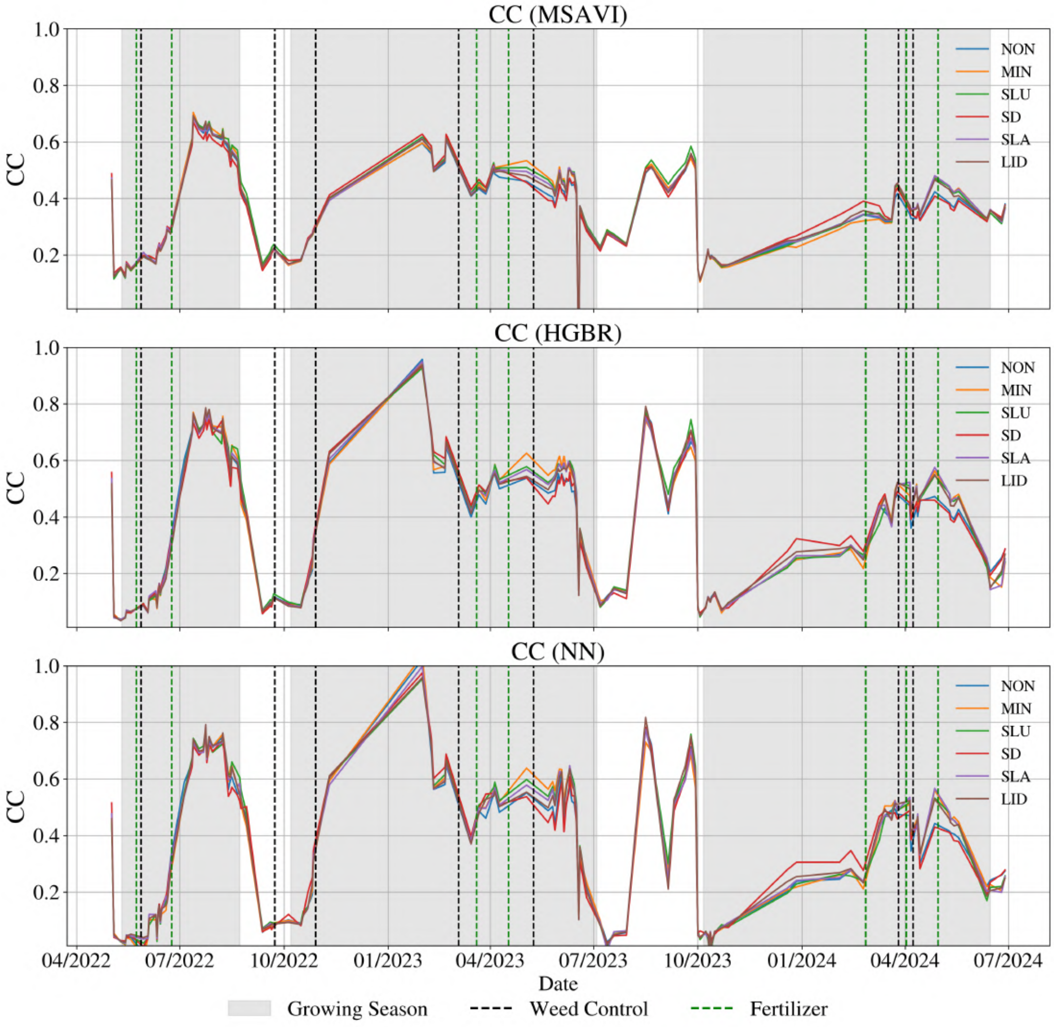


Figure 30. **Inter-Seasonal Canopy Cover Predictions from Baseline Models for Experimental Site** ”Bio4Recycle” at Wallbach. The figure shows the temporal evolution of predicted mean canopy cover (CC) values for different treatments, based on baseline models (MSAVI, HGBR, and NN), across three growth seasons (corn: 2022, winter wheat: 2022–2023, and winter barley: 2023–2024). Fertilizer treatments are color-coded as follows: NON (grey), MIN (orange), SLU (green), SD (red), SLA (purple), and LID (brown). Grey shading indicates the duration of the growth seasons, while vertical dashed lines mark the dates of weed control (black) and fertilizer applications (green). Data © 2024 Planet Labs PBC.

4.6.2 Inter-seasonal Timeseries Analysis

The interseasonal CC time series for the plot phenotyping field experiment ”Bio4Recycle” in Wallbach has been successfully derived from satellite data and is visualized in Figure 30. The three different models (MSAVI, HGBR, and NN) are displayed from top to bottom, each illustrating the mean CC across plots subjected to various fertilizer treatments over multiple growth seasons. The individual growth seasons (grey) correspond to corn (2022), winter wheat (2022–2023), and winter barley (2023–2024). In general, a clear seasonal pattern emerges, characterized by an initial increase in CC during early growth phases and followed by a gradual decrease towards the end of the growth season. Additionally, CC values noticeably decrease following weed control management, provided that the temporal resolution of the data is sufficient. The effect of fertilizer application is also noticable, with CC values of respective treatment groups increasing after a short time lag. When comparing the different crops, corn exhibits the highest CC during the key growth phase prior to harvest, whereas winter barley shows the lowest. On a model level, MSAVI shows less extreme predictions overall than HGBR and NN. MSAVI predictions reach lower peak and higher minimum CC values than the other models. In contrast, the NN model produces the most extreme values, with CC occasionally exceeding 1 and dropping below 0. This effect is most prominent for winter wheat in January 2023, where CC values have unrealistically high peak values. A similar, though less extreme, effect can also be observed in the HGBR model. Despite these two learning-based approaches showing more extreme values, they also reveal more differentiated results for the different treatments, especially in the periods before harvest in 2023 and 2024. Overall, while HGBR and NN offer greater sensitivity to treatment differences, MSAVI provides more stable and consistent inter-seasonal predictions.

4.6.3 Early Growth Timeseries Analysis

The analysis of the early growth season is done for winter barley in Figure 31, but similar results are visualized and appended in 97 and 98. All models exhibit a general trend of increasing CC with time. Thereby, the MSAVI model reveals a more stable prediction compared to HGBR and NN. Especially after the first application of fertilizer, the CC increases substantially for HGBR and NN. In contrast, the trend in MSAVI is less strong, with certain treatment plots even showing a decrease in CC values. The comparison of treatments displays similar patterns among all models. It is particularly noteworthy that the SD treatment shows the highest CC before the initial fertilizer application. On the other hand, the MIN treatment shows the lowest CC values in this time span. All the other treatments show similar CC values, with LID treatment being slightly higher. After fertilizer application, the range of CC between the different treatments is narrowing.

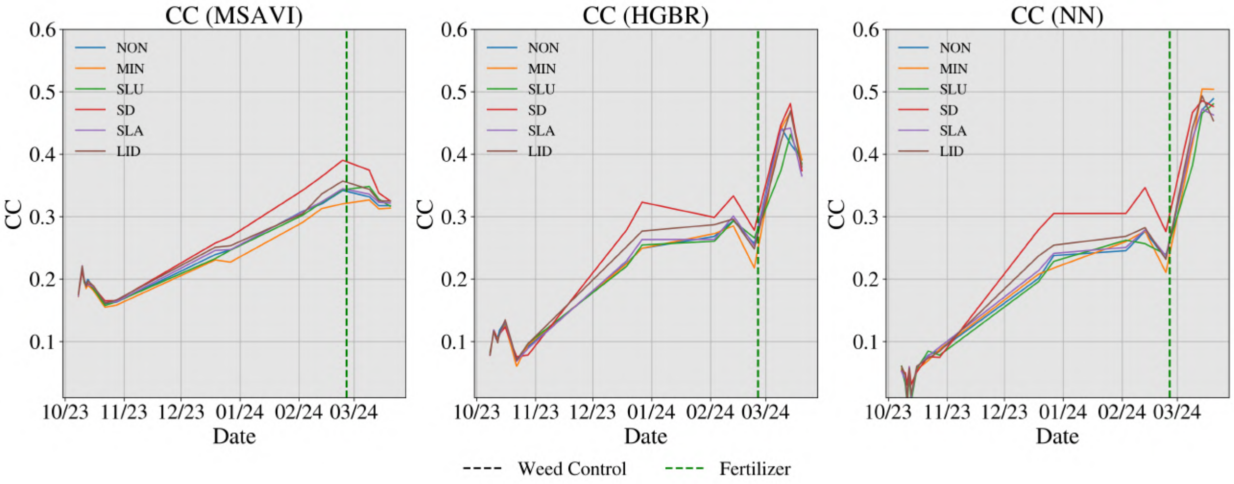


Figure 31. Early Growth Stage Canopy Cover Predictions from Baseline Models for Winter Barley at Experimental Site ”Bio4Recycle” in Wallbach. The figure shows the development of predicted mean canopy cover (CC) values for different treatments, based on baseline models (MSAVI, HGBR, and NN), between October 2023 and March 2024. Fertilizer treatments are color-coded as follows: NON (grey), MIN (orange), SLU (green), SD (red), SLA (purple), and LID (brown). Grey shading indicates the duration of the growth seasons, while vertical dashed lines mark the dates of weed control (black) and fertilizer applications (green). Data © 2024 Planet Labs PBC.

4.6.4 Late Growth Phase Timeseries Analysis

At the later growth stage, the CC predictions from the three models (MSAVI, HGBR, and NN) exhibit a consistent temporal pattern across treatments. Two distinct peaks are observed in Figure 32. The first occurs at the end of March, just before the initial weed control application, and the second toward the end of April, with slightly higher CC values. Following the first weed control event, a general decline in CC is observed across all treatments. This reduction is less pronounced after the second weed control, which occurs approximately one week after a fertilizer application. At both the beginning and end of the observation period, treatment curves lie closely together, while differences between treatments are most pronounced during the peak growth stages. The model comparison reveals that MSAVI predicts CC values ranging from approximately 0.3 to 0.5, showing stable and smooth curves. It also reflects the clearest visual response to weed control in the shape of its predictions. HGBR shows greater variability, with CC values ranging from 0.2 to 0.6 and a more fluctuating pattern over time. The NN model exhibits the highest degree of variability, especially at lower CC values, with a distinct drop in mid-May. Among the three models, NN displays the widest range of CC values at peak growth, also between 0.2 and 0.6. Regarding treatment performance, the untreated control (NON) and the SD treatment consistently display the lowest CC values. In April, the NON plots exhibit the lowest CC values, whereas in May, the SD plots show the minimum CC values. Across all models, the maximum CC for these two treatments falls between 0.42 and 0.47. The remaining treatments (MIN, SLU, SLA, and LID) show slightly higher CC values, with maxima ranging from 0.48 to 0.58. Among these, the SLU and the positive control MIN generally show the highest CC, while SLU and LID tend to exhibit slightly lower values during the late growth stage.

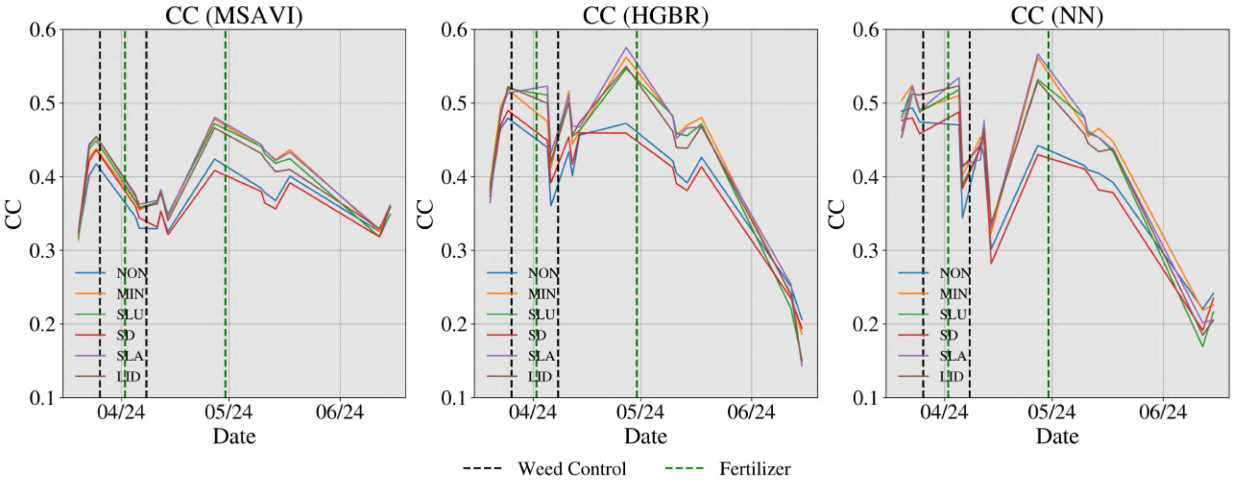


Figure 32. Late Growth Stage Canopy Cover Predictions from Baseline Models for Winter Barley at Experimental Site ”Bio4Recycle” in Wallbach. The figure shows the development of predicted mean CC values for different treatments, based on baseline models (MSAVI, HGBR, and NN), between March and June 2024. Fertilizer treatments are color-coded as follows: NON (grey), MIN (orange), SLU (green), SD (red), SLA (purple), and LID (brown). Grey shading indicates the duration of the growth seasons, while vertical dashed lines mark the dates of weed control (black) and fertilizer applications (green). Data © 2024 Planet Labs PBC.

5 DISCUSSION

5.1 Reflection on the Data Processing Pipeline

The following discussion examines the performance and limitations of the data processing pipeline, beginning with UAV-based acquisition and segmentation and proceeding to an assessment of the quality and reliability of the PSB.SD satellite data workflow.

5.1.1 UAV Data Acquisition and Processing

The results demonstrate that the model initially proposed by Sadeghi-Tehran et al. (2017) for high-resolution, handheld RGB imagery can be effectively scaled to field-level predictions using lower-resolution UAV-based RGB mapping data. This scalability was enabled by integrating SNP imagery, captured by a secondary UAV camera system, into the high-resolution CC model. The inclusion of these SNP images enhanced the transferability of the segmentation model across imaging systems and spatial resolutions. A fully automated workflow was implemented to harmonize the SNP and mapping images, combining feature-based geometric alignment and histogram matching. These preprocessing steps successfully transformed the UAV imagery into a reliable training dataset for the Field-Level CC Model. Both quantitative accuracy metrics and qualitative visual inspections of the segmentation masks indicated high performance on both scales. Notably, the field-level reference data, generated from imagery with a GSD of approximately 30 mm/pixel, provided detailed and accurate inputs, enabling robust model training of a satellite CC model with minimal labeling error.

Despite the overall success of the pipeline, certain limitations were encountered for feature matching in windy conditions, particularly when imaging dense canopies. In such cases, feature matching proved to be unreliable due to motion blur and insufficient stable key points. To address this, a fallback method was implemented using image center point coordinates to perform geometric alignment. While less precise, this approach proved sufficient under high CC conditions, where spectral and structural variations are minimal, and overall segmentation accuracy remained stable. Also, variation in field-level segmentation performance across crop types was observed. Models performed best on crops with well-separated leaf structures. In contrast, winter barley showed the lowest accuracy, with an overall model accuracy of 0.9412, as shown in Table 20 in the appendix. This was confirmed through visual assessment, which revealed reduced performance in detecting finer structures and vegetation with lower greenness, with a tendency to slightly underestimate CC. A notable counterpart to the RF segmentation approach used in this study is presented by Puig et al. (2025), who developed a convolutional neural network (CNN)-based model for CC prediction using 283000 RGB images collected under field conditions. High accuracy was achieved by their model ($R^2 = 0.98$), highlighting the potential of DL approaches for CC estimation. However, this method relied on a training dataset 1,000 times larger than mine and required substantially greater computational resources. In contrast, similarly strong results were achieved by the segmentation models presented here through color space transformation and mission-specific scaling, underscoring the strengths of the proposed pipeline. While NNs may offer slightly higher accuracy in scenarios requiring ultra-fine resolution, the findings presented here demonstrate that well-tuned ML models can yield highly accurate and generalizable results with significantly lower computational demands.

Overall, the developed model scaling pipeline from high-resolution UAV-based RGB imagery to field-level CC mapping performed robustly across a variety of crop types and conditions. Nonetheless, the complexity of the approach, particularly the geometric alignment of SNP reference data, may constrain its broader applicability. For future implementations or related studies, it is worth considering that comparable performance may be attainable using SNP imagery at lower zoom levels, thereby increasing field coverage and simplifying data acquisition without substantially compromising segmentation quality.

5.1.2 PlanetScope Satellite Imagery

PSB.SD satellite data was successfully acquired using an automated API-based script, which enabled efficient, large-scale querying and downloading of imagery across study sites and observation periods. However, the automation process was limited by the absence of AOI-specific cloud cover filtering, which reduced the precision of scene selection. Although the metadata provided global cloud cover estimates and a pixel-wise confidence layer, the associated cloud masks were unreliable in practice. While the confidence layer offered partial support in filtering low-quality pixels, manual inspection and filtering of specific scenes remained necessary to ensure data reliability. Large-scale spatial shifts were effectively corrected using an `AROSICS`-based co-registration workflow. Nevertheless, small-scale geometric mismatches persisted. These misalignments became apparent during the pixel-level model comparison (Section 4.5.3), which revealed localized discrepancies likely attributable to residual geometric shifts. Although these shift effects are challenging to quantify directly, they can subtly degrade model accuracy and are consistent with previously reported limitations in the geometric stability of CubeSat imagery, where shifts of up to 20 m have been observed (Houborg and McCabe, 2016, 2018). While not directly quantified in this study, the successful application of the method to a plot experiment with 6–9 m plots suggests that any geolocation shifts are substantially smaller than 20 m. From a radiometric perspective, the ELC workflow yielded stable results overall, as revealed by the visual diagnostics that were incorporated to ensure radiometric consistency over time. PITs were monitored for temporal stability, and significant deviations, such as the observed material change at the “Horse track” in Wallbach, were flagged and excluded when necessary. Still, the short-term fluctuations observed in the model predictions from the HGBR and NN models raise open questions. It remains unclear whether these variations reflect subtle residual inconsistencies in satellite reflectance or are artifacts introduced by model sensitivity due to minor spectral differences.

In summary, the PSB.SD processing workflow introduced greater uncertainty than the UAV-based approach, particularly with respect to spatial alignment. In spite of this, its successful application to the experimental site and the generation of interpretable, multi-season time series demonstrate the potential of PSB.SD data for operational field phenotyping. Although the workflow remains semi-automatic, it enables the use of PSB.SD imagery for CC monitoring if careful preprocessing and quality control are applied.

5.2 Assessment of the Training Dataset

The training dataset, comprising more than 30000 data points, used for satellite-based CC modeling has a resolution of 3m and captures a diverse range of crop types and field conditions, with data collected across multiple sites and growth stages. This diversity allowed the models to generalize effectively across site-specific, crop-specific, and phenological variability. A major strength of the dataset is its spatial and temporal coverage, which contributed to strong model performance across the included summer crops. However, several limitations were also observed during the application. The absence of winter crop data limits the accuracy of year-round model applications, as CC prediction values exceeding one were observed during winter at the experimental site. Additionally, early-season imagery at the Haerkingen site was affected by persistent cloud cover, resulting in large temporal gaps that required interpolation and potentially introducing uncertainty into early growth stage predictions. Furthermore, the dataset was skewed toward very low CC values (< 0.05) and exhibited an uneven distribution across crop types. This imbalance may have influenced the model’s learning behavior, particularly under low-vegetation conditions or for underrepresented crops. Despite these challenges, the dataset proved to be a solid foundation for modeling CC under realistic field conditions. It supported robust predictions across seasons, crop types, and locations, demonstrating its value for field phenotyping applications.

5.3 Satellite Model Performance and Interpretation

This section discusses the performance and behavior of the different models for CC estimation, including model selection, baseline comparisons, and detailed cross-validation across missions and spatial scales. The aim is to interpret the models’ strengths and limitations in relation to crop type, growth stage, and data quality.

5.3.1 Satellite Model Selection

The study identified three best-performing models for satellite-based CC estimation: MSAVI (linear model), HGBR (ML), and NN (DL). Each model achieved the lowest RMSE within its respective category and performed particularly well during early crop growth stages. MSAVI recorded the lowest cumulative RMSE across cross-validation and outperformed other VIs in low CC conditions. This is consistent with findings by Bannari et al. (1996), who reported that indices such as TSAVI, TSARVI, SAVI, and MSAVI are more resilient to soil background effects and better distinguish vegetation from bare soil in heterogeneous environments. Several other studies have also demonstrated the effective use of MSAVI in CC estimation (Anurogo et al., 2018; Mcgwire et al., 1991). Among the ML models, both RF and HGBR performed well during the early stages, but HGBR showed slightly better overall accuracy and was therefore selected for further analysis. In contrast, Raza et al. (2025) reported that RF outperformed Gradient Booosting (GB) in a wheat yield prediction study. Whereas their models employed VIs as input variables, the present study utilized the complete multispectral reflectance dataset. Differences in the complexity of input features and validation strategies may explain the discrepancy, but both results highlight the comparable performance of ensemble tree-based methods. The NN model demonstrated consistently high performance across all crop growth stages, with particular strength during early canopy development. These results underscore the potential of DL models to deliver accurate and generalizable predictions for CC estimation, as it was also highlighted in the case of soybean by Bocco et al. (2012).

5.3.2 Baseline Satellite Model

The baseline model comparison revealed differences in performance across the tested modeling approaches. The simple MSAVI-based linear model showed limited predictive capacity in the early growth stages (CC < 0.2), where it deviated substantially from the reference values. In contrast, the HGBR model achieved higher accuracy and more stable performance across the full CC range, although slight overestimation was observed at low CC. The NN model demonstrated the highest overall accuracy, particularly excelling in early growth stages. While MSAVI and HGBR models mainly relied on the Red and NIR bands, the NN model additionally incorporated information from the Yellow, Red Edge, and Green bands. All models captured the general CC patterns with an R^2 around 0.9, but the NN model consistently outperformed the others across all stages, achieving an R^2 of 0.923. The superior performance of HGBR and especially NN underlines the learning capacity of ML and DL approaches, offering the best trade-off between accuracy and robustness of the baseline model for a seen scenario. The findings are consistent with recent research by Bebie and Kyparissis (2024) on durum wheat yield estimation using Sentinel-2 data. In this study, index-based multiple linear regression models demonstrated moderate performance, with R^2 values ranging from 0.476 to 0.503. In contrast, ML models, such as RF and K-Nearest Neighbors (KNN), exhibited significantly higher accuracy, with R^2 values exceeding 0.94.

Although the linear MSAVI model exhibited slightly lower performance, it still provided comparable predictions. To contextualize the developed linear MSAVI model with previous studies that explored linear relationships between CC and VIs, I compare it with studies that have established relationships between NDVI and CC across various crop species, as reported in the meta-analysis by Tenreiro et al. (2021). To facilitate a comparison with the linear MSAVI model, as presented in Equation 10, the results from Tenreiro et al. (2021) were normalized to a 0-1 scale by dividing the original values by 100. These reproduced and normalized results of Tenreiro et al. (2021) are presented in Table 18.

When compared to the general model listed in Table 18, it becomes evident that a stronger correlation was achieved in the MSAVI-based linear model presented in this study ($a = 1.595$) than in the aggregated NDVI-based model ($a = 1.055$) reported by Tenreiro et al. (2021). This higher correlation can be attributed, at least in part, to the non-linearly higher NDVI values compared to MSAVI, as shown for the training data frame in Figure 44 in the appendix. Furthermore, this higher correlation may also be attributed to the increased sensitivity of MSAVI, as it was specifically developed to compensate for soil effects (Qi et al., 1994). Since the presented MSAVI model is predominantly based on the crop corn,

Table 18. **Linear NDVI-Canopy Cover Models**. The table provides the coefficients a and b for the linear models ($CC = a + bx$) relating NDVI to canopy cover (CC) for various acquisition and crop types. The models are based on different datasets, with the values of a and b divided by 100 as per the original study of Tenreiro et al. (2021). *Reproduced from Tenreiro et al. (2021).*

Crop	**a**	**b**	R^2
General	1.05427	-0.06501	0.71
Satellite	0.97088	0.04106	0.62
In-situ	1.12023	-0.13813	0.81
Corn	1.13956	-0.22268	0.78
Sunflower	0.7737	0.0158	0.69
Barley	0.93243	-0.0158	0.82
Winter-spring	0.96337	-0.02734	-
Spring-summer	0.98762	0.0156	-

a closer inspection of the NDVI-based corn model indicates that the closest slope (a) and similar intercept coefficients (b) were observed between the MSAVI model and the linear corn model, suggesting consistency with findings from prior field studies. Although R^2 strongly depends on the validation method used, the R^2 of the developed MSAVI model is consistently higher than any of the other NDVI-based models presented by Tenreiro et al. (2021).

To contextualize the findings from the developed baseline NN model, a comparison is made with the results of a soybean ground cover estimation study using MODIS and Landsat conducted by Bocco et al. (2012). It is noteworthy that while MODIS provides daily observations, it has a low spatial resolution (250 m). Conversely, Landsat offers higher spatial resolution (30 m) but at the cost of lower temporal frequency, with imagery (only) available at 16-day intervals (Bocco et al., 2012). Thus, it is compared at a much larger spatial scale while still employing a similar NN approach. For the soybean crop from a single study area, two models for MODIS (NN_MOD) and Landsat (NN_TM) were developed by Bocco et al. (2012). The NN_MOD model achieved a correlation of 0.904 and an RMSE of 0.112, while the NN_TM model showed an R^2 of 0.909 and an RMSE of 0.100 (Bocco et al., 2012). In comparison, the developed NN model, which includes multiple crops and training sites, outperformed these with an R^2 of 0.923 and an RMSE of 0.069. These findings emphasize the value of higher-resolution satellite data. Additionally, since Bocco et al. (2012) only used the NIR and Red bands, the results demonstrate that incorporating other spectral channels, such as Yellow, RedEdge, and Green, as shown in Figure 13(b), significantly contributes to better model performance.

5.3.3 Mission Cross Validation Satellite Model Comparison

Quantitative comparison

The quantitative pixel-wise comparison revealed several important trends in model performance. Among the three crop types, sugarbeet was the most accurately predicted, exhibiting the lowest average NRMSE, while corn proved to be the most challenging. The variation in model performance across crops can be ascribed to several factors. First, sugarbeet displays the most stable data distribution across all CC stages, as shown in Figure 8. Second, its predominantly horizontal growth structure leads to spectral changes that are more directly correlated with CC, unlike vertically structured crops such as corn and sunflower. Third, the reduced performance of corn is likely influenced by a subset of high-error cases, as indicated by the disparity between its low median and high mean NRMSE. This is further supported by mission-specific NRMSE values presented in Table 21, where corn-related missions with flight IDs 102 and 103 show NRMSE values exceeding 1, suggesting the presence of artifacts or anomalies in satellite imagery.

Across models, the NN consistently demonstrated the strongest performance, achieving the lowest overall mean NRMSE (0.237) and performing best on both sunflower and corn. In contrast, MSAVI exhibited lower median NRMSEs across all crops, suggesting reliable performance in typical cases but reduced robustness due to sensitivity to outliers. The superior mean performance of the NN model reflects its ability to generalize well across diverse conditions, including those affected by noise or anomalies, as previously discussed.

Qualitative comparison

The qualitative pixel-wise comparison across multiple sites and growth stages revealed distinct performance characteristics for each model, influenced strongly by crop type as well as the availability and quality of training data. Model performance was closely tied to data availability. The MSAVI-based model performed better compared to the others in scenarios where training data were sparse or of lower quality, such as for late-stage sunflower (Figure 18) or early-stage corn at Haerkingen (Figure 14), respectively. However, in cases where sufficient and high-quality data were available, such as sugarbeet, both the HGBR and NN models outperformed MSAVI, demonstrating their ability to utilize richer datasets effectively. At later growth stages, where data availability was generally more uniform, the NN model consistently achieved the highest predictive accuracy, reflecting its ability to learn from more complex patterns. The HGBR model showed slightly lower but comparable performance to the NN model. However, HGBR frequently exhibited horizontally striped prediction artifacts, indicative of systematic errors. These performance differences can be attributed to the architectural characteristics of each model and their interaction with data availability and quality. The MSAVI-based model, relying on predefined spectral relationships, performed robustly when training data were limited or noisy. Its simplicity and independence from large datasets explain its strong performance in early-stage or data-scarce scenarios. However, its linear structure limits its capacity to adapt in richer data contexts, where it was outperformed by more flexible models such as the HGBR. The HGBR captures non-linear patterns more effectively but is sensitive to spatial inconsistencies in satellite data. The horizontally striped artifacts observed in several sites likely stem from its use of histogram binning and sharp threshold-based tree splits. Small radiometric variations between adjacent image rows can lead to abrupt prediction jumps, which are amplified through its sequential decision structure. Lastly, the NN performs best with abundant, high-quality training data, making it particularly effective at later growth stages when CC distributions are more uniform. Its flexibility allows it to model complex relationships, as demonstrated in the late-stage corn and sugarbeet sites. However, this same flexibility can lead to overfitting in early growth stages or under sparse data conditions, as observed in the early-season corn example, thereby reducing its reliability in such scenarios.

In summary, each model's strengths and weaknesses align closely with its architectural design. The NN performs best in complex, data-rich environments, while MSAVI remains a robust and interpretable choice under sparse or unseen conditions.

Spatial Comparison

The spatial analysis of model predictions reveals that performance varies notably between missions. Spatial deviations between predicted and reference CC are most pronounced during early growth stages. In contrast, in later growth stages, all models exhibit highly consistent and accurate spatial predictions, with only minor deviations. Among the evaluated models, the MSAVI-based model consistently demonstrates the most spatially stable behavior across all sites.

Its deviation maps are generally balanced, showing only localized and moderate differences with minimal artifacts or structured error patterns. In comparison, the HGBR and NN models display more systematic spatial deviations. The HGBR model tends to underestimate CC in areas with denser vegetation and overestimates in sparser regions, often producing zonal patterns. The NN model shows even greater spatial variability in deviations. Across all models, several recurring spatial artifacts were identified. These include localized patches of overestimation in low-CC regions, narrow edge bands (1–2 pixels wide) of over- or underestimation, and striping patterns.

Many of these spatial patterns, such as striping, localized over- or underestimation, and edge effects, can be attributed to limitations in PS data quality. As previously reported in the literature, these sensors often struggle to capture fine-scale spatial variation, which can affect the accuracy of vegetation structure assessments (Wildhaber et al., 2023; Frazier and Hemingway, 2021; Houborg and McCabe, 2018). The non-uniform edge effects, with spatial shifts between over- and underestimation, suggest issues related to the spatial accuracy of the satellite imagery. Co-registration errors, which are also evident in several fields, are a well-documented issue associated with PS data (Wildhaber et al., 2023; Frazier and Hemingway, 2021; Houborg and McCabe, 2016, 2018). In addition to sensor-related artifacts, model-specific deviations are also evident. The high spatial variability of deviations observed in HGBR and NN predictions, particularly in large regions of extremely sparse or dense vegetation, can be attributed to low model sensitivity to less frequent data points. In summary, spatial analysis highlights both sensor-induced artifacts and model-specific biases. A precise spatial prediction depends not only on the robustness of the model but also on the quality and spatial accuracy of the input imagery. Notably, small spatial patterns are not consistently captured, showing the spatial limits of PS data. Furthermore, the observed inconsistencies in HGBR and NN predictions once again emphasize their strong dependence on the quality and representativeness of the training data, particularly under heterogeneous field conditions.

Field-Level Mean Comparison

The key findings from the mean comparison at field-level reveal that all three models, MSAVI, HGBR, and NN, demonstrated strong overall agreement with UAV-derived field-level CC estimates ($R^2 > 0.933$). While the NN model overall delivered the most accurate predictions across sites and growth stages, performance varied notably by crop type and location. The NNs showed strong performance when sufficient and well-distributed training data were available. In contrast, MSAVI showed better performance in unseen or underrepresented scenarios, likely due to its reliance on a simpler VI-based relationship that is less dependent on large training datasets. Sugarbeet emerged as the most accurately predicted crop, particularly for the HGBR and NN models. This can be attributed to its predominantly horizontal growth structure, which enhances canopy visibility and results in a more stable spectral signal in satellite imagery. The high prediction accuracy and strong performance of HGBR and NN at the Reckenholz site are further supported by the uniform field conditions, which align well with model architectures optimized for generalization. Additionally, the sugarbeet dataset is characterized by high availability and well-distributed training samples, likely facilitating effective model learning and contributing to robust performance.

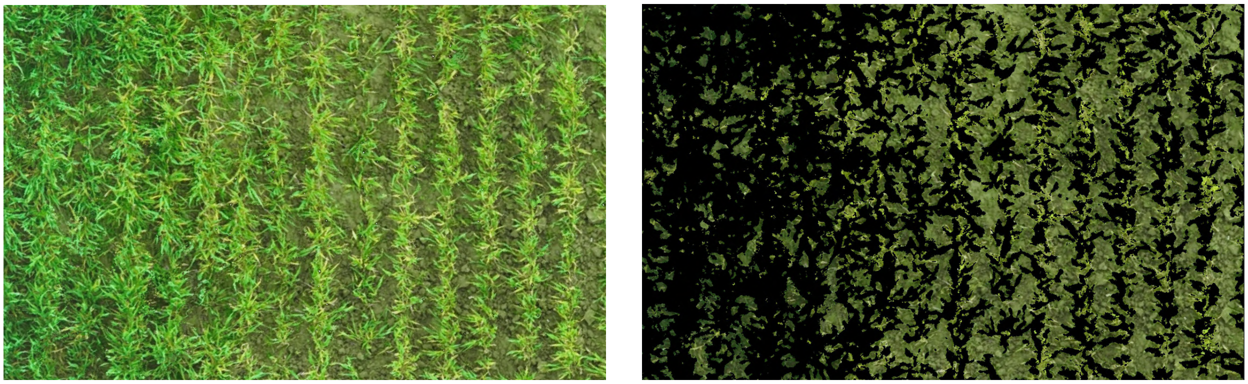
5.3.4 Application on Experimental Site

All three models were applied to satellite time series data from the experimental ”Bio4Recycle” site in Wallbach to evaluate their performance under field conditions and analyze the plot-wise effect of different fertilizer treatments.

Spatial Analysis

Model predictions were compared against UAV-derived CC data to assess spatial accuracy. A consistent overestimation of CC was observed across all models, with mean values predicted at 0.429 (MSAVI), 0.501 (HGBR), and 0.491 (NN), compared to a UAV-derived mean of 0.296. This overestimation aligns with the findings of Tenreiro et al. (2021), who observed in their meta-analysis that satellite-derived NDVI often overestimates CC at low NDVI values when compared to in-situ proximal sensing. Tenreiro et al. (2021) attribute this overestimation to the noise effect that disrupts reflected signals from sparse vegetation surfaces, as described by Todd and Hoffer (1998), cited in Tenreiro et al. (2021). Furthermore, the mentioned overestimation can be partially attributed to the underestimation of CC by the UAV method in the case of winter barley, as previously discussed in Section 4.1.2. However, these factors alone do not account for the magnitude of overestimation, which reaches approximately 40% above the reference values. A likely contributing factor is the absence of winter barley-specific data in the model training process. The satellite-based CC models were calibrated using data from crops such as sugarbeet, corn, and sunflower, which exhibit markedly different canopy architectures. In contrast to the broader and more open structures of these crops, winter barley features finer, denser vegetation that generates a less distinctive spectral signal. The observed overestimation can be confirmed through comparison with crop-specific linear NDVI models analyzed by Tenreiro et al. (2021). Tenreiro et al. (2021) report a flatter correlation line ($0.93x - 0.02$) for the barley model than for the general model ($1.05x - 0.07$). This further supports the findings of this study, suggesting that the CC of barley is overestimated when using a general model as compared to a model specifically developed for barley. Among the models, MSAVI produced predictions that were most closely aligned with the UAV reference, showing the lowest mean overestimation ratio. This supports the hypothesis that the ML-based models (HGBR and NN) were less effective in capturing the distinctive characteristics of winter barley, as they were trained on other crop types. Although HGBR and NN showed higher overestimation, both produced more pronounced spatial contrasts between plots, suggesting that the overestimation partly arose from model-related limitations rather than solely inherent issues in the satellite imagery.

Finer-scale variation, particularly in the smaller 3-meter-wide plots, was not well captured by any of the models, highlighting constraints imposed by the spatial resolution of the satellite data. While PSB.SD imagery was generally sufficient to differentiate CC at the 6-meter and especially 9-meter plot scale, notable mispredictions occurred in plots with low UAV-derived CC values that were adjacent to high-CC plots. This pattern suggests that signal blending from neighboring plots, spatial co-registration errors, or mixed-pixel effects may have influenced the model outputs, although these effects appear to occur at spatial scales much smaller than the 20 m reported by Houborg and McCabe (2016, 2018). However, not all instances of high overestimation in low-vegetated areas can be attributed solely to satellite-related effects. The treatment-level analysis revealed additional discrepancies, most notably that the greatest overestimation occurred in the NON treatment, while the lowest was observed in the MIN treatment. A visual comparison of these two treatments in Figure 33 supports this finding.



(a) Original RGB **(b)** Masked RGB

Figure 33. Comparison of Canopy Cover Segmentation for Different Treatments at Experimental Site. The figure compares canopy cover (CC) segmentation between the NON and MIN treatments at the experimental site ”Bio4Recycle” in Wallbach on 27.04.2024. Panel (a) shows the original RGB imagery from the UAV mapping mission, while panel (b) displays the masked RGB imagery with the Field-Level CC model applied. Black pixels in panel (b) indicate areas predicted as plant canopy. The left side of each image shows the MIN treatment, while the NON treatment is on the right.

The RGB image of the NON plot shows a distinctly yellowish hue, indicating unhealthy vegetation. In the corresponding segmented image, many of these yellowish plant components were not successfully classified as vegetation, an issue previously discussed in Section 4.1.2. In contrast, the segmentation of the MIN treatment plot appears more accurate, with only a small number of plant pixels left unclassified. These results suggest that the apparent model overestimation is at least partially attributable to an underestimation of CC in the UAV-derived reference, particularly in the poor fertilized treatments.

Timeseries Analysis

The seasonal pattern of CC at the experimental site was consistently captured by all models (MSAVI, HGBR, and NN), with CC increasing during early growth stages and CC decreasing as the season progressed. Among the three models, MSAVI delivered the most stable predictions, exhibiting less extreme values and consistent inter-seasonal trends. It effectively captured broad temporal dynamics, as well as management interventions such as weed control. This stable performance underscores MSAVI’s strength in producing reliable CC estimates across various temporal scales. However, its lower sensitivity to treatment-level differences is likely tied to its narrower dynamic range, which limits the contrast between treatments. In addition, the lack of predicted CC below 0.15 before seeding suggests a potential overestimation of CC in low-vegetation conditions. On the other hand, the HGBR model and especially the NN model showed greater sensitivity to different treatments. This responsiveness was most apparent following fertilization events, where both models reflected sharper increases in CC. Yet, this advantage came with increased variability and occasional biophysically implausible outputs. The NN model produced extreme predictions, with CC values exceeding one or dropping below 0, raising concerns about potential overfitting. A likely explanation for this behavior lies in the limited temporal scope of the training data. Both models were trained primarily on summer imagery and thus lacked exposure to winter conditions such as frost or snow. This absence may have impaired the models’ ability to generalize during colder periods. Supporting this, temperature data from the date of a notable prediction spike on February 7, 2023, shows a minimum of -2.3 °C, as shown in Figure 2, suggesting that frost may have affected vegetation reflectance and contributed to the observed out-of-range values.

Distinct differences in CC dynamics were also observed between fertilizer treatments, highlighting the successful application of the models to a plot-scale field experiment. During the early growth phase of winter barley, treatment-related differences in CC were clearly pronounced. The SD treatment consistently showed the highest CC values prior to fertilization, whereas the MIN treatment exhibited the lowest. Following fertilization, these differences diminished, suggesting a convergence in canopy development among treatments as the season progressed. In the late growth phase, the NON and SD treatments maintained the lowest CC values, whereas the MIN and SLU treatments reached the highest CC levels. These findings underscore the differential impact of fertilizer regimes on CC growth and emphasize the value of temporally well-resolved satellite imagery, rather than relying solely on end-of-season assessments.

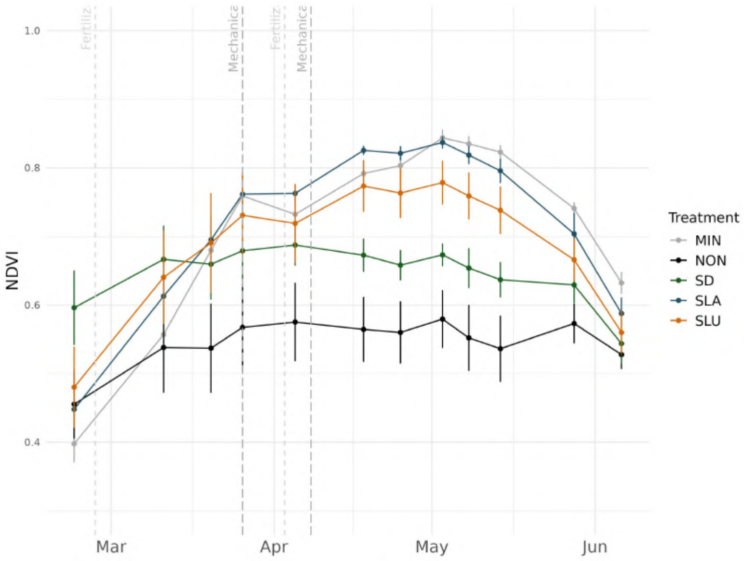


Figure 34. **UAV-Based NDVI Curves for Different Fertilizer Treatments at the Experimental Site** ”Bio4Recycle” at Wallbach for Winter Barley in 2024. The figure shows the mean NDVI values with standard deviation over time from multispectral UAV data of Diener et al. (2024). Treatments include MIN, NON, SD, SLA, and SLU. *Reproduced with permission from Diener et al. (2024).*

To complement the satellite-based results, independent NDVI time series derived from a multispectral Phantom UAV were included for comparison (Figure 34). The data was collected and visualized by Diener et al. (2024) as part of the ”Bio4Recycle” experiment in Wallbach. The figure, which shows NDVI trajectories across different fertilizer treatments, is reproduced from Diener et al. (2024). As mentioned earlier, NDVI has been widely used to estimate CC, although it is not the most suitable index for this purpose. Nonetheless, in this case, it serves as a relevant proxy for plant physiological activity and is therefore used to compare against the CC predictions from the satellite-based models. The NDVI data primarily covers the late growth stage and is thus mainly compared with model predictions from this period. However, a short section in early March also overlaps with the early growth phase. During this period, the SD treatment clearly shows the highest NDVI values, while the MIN treatment shows the lowest. Also, the other treatments rank consistently across both the NDVI curve (Figure 34) and the satellite-based CC predictions (Figure 31). Approximately two weeks after fertilizer application, a shift in treatment performance becomes visible. This inversion of treatment performance in NDVI rankings temporally aligns with the CC time series. In mid-April, MIN and SLA emerges as the best-performing treatments, which is also consistent with the observations from satellite-derived CC reported in Figure 32. Furthermore, as observed in that section, the untreated control (NON) and the SD treatment consistently show the lowest CC values during the late growth stage, mirroring the trends seen in the NDVI data. From a physiological perspective, this pattern is plausible. The SD treatment, which received solid fertilizer early in the season, initially outperforms the others, likely due to enhanced nutrient availability during the early growth phase. In contrast, the other plots had not yet received fertilizer at that time and therefore showed more similar and generally lower performance. Notably, the MIN treatment underperforms during the early growth stage compared to the organic treatments. This could be due to nitrogen depletion from the previous crop or may result from reduced soil mineralization caused by the continuous use of mineral fertilizers, which, according to Ren et al. (2020), can alter soil microbial dynamics. Long-term application of mineral nitrogen inputs has been shown to suppress microbial diversity and activity(Ren et al., 2020). The negative impact of nitrogen fertilizers on soil mineralization is further supported by the findings of Mahal et al. (2019) in a study on corn, where sustained nitrogen input led to decreased microbial biomass and lower enzymatic activity. The shift in treatment performance following the first fertilizer application and its further intensification after additional nutrient inputs is consistent with expectations, as treatments with high ammonium (NH_4^+) concentrations, such as MIN and SLA, performed well. This aligns with the findings of Wang et al. (2019), who reported positive effects of NH_4^+ when applied in combination with nitrate. Toward the end of the season, the MIN treatment, as the positive control, achieves the highest CC and NDVI values, which aligns with its role as a performance benchmark. Notably, the SLA treatment ultimately reached CC levels comparable to those of mineral fertilizer treatments, supporting its suggested potential as an effective organic alternative (Ehmann et al., 2018; Riva et al., 2016).

5.4 Future Directions

5.4.1 Dataset Expansion and Diversity

Future improvements in CC estimation models will greatly benefit from broader and more balanced training datasets. A key priority is the inclusion of a wider range of crop types, particularly fine-structured winter crops such as barley and wheat. These crops are currently absent from the training data but may introduce important structural and seasonal variations. Their integration is essential to enable year-round monitoring and for improving model generalizability across diverse cropping systems. For applications requiring high predictive accuracy, the development of crop-specific models should be prioritized. These models can benefit from tailored UAV reference datasets collected continuously throughout an entire year, allowing precise tracking of CC development across all phenological stages. In the context of long-term experiments focused on a single crop, reference data acquired in one season can serve as a stable foundation for training and validating models over multiple years. This approach minimizes the need for repeated reference data collection while still enabling reliable, crop-specific predictions over time. Furthermore, it is critical to balance the distribution of CC values in the training dataset, particularly by improving representation across mid to high CC ranges and underrepresented crop types. Addressing this imbalance will help mitigate model bias and support more robust learning across varying canopy structures and agroecological conditions.

5.4.2 Improved Labeling and Reference Data

To improve the assessment of sustainable and alternative farming practices, it is essential to include labeled data that reflect the full range of crop conditions typically observed in such systems. Unconventional management approaches, such as organic or reduced-input farming, often produce more heterogeneous vegetation patterns characterized by low greenness, irregular plant structures, and stressed canopies. To ensure robust segmentation performance under these challenging conditions, it is important to integrate labeled data from a broader spectrum of farming systems into the training datasets.

To improve the scalability and practicality of the segmentation workflow, the use of lower zoom level SNP imagery with broader spatial coverage offers a promising alternative to the dual-camera approach that relies on geometric matching between high-resolution SNP and mapping images. This alternative method maintains sufficient quality for segmentation tasks while significantly reducing the complexity of data acquisition and preprocessing. By simplifying the overall process, it enables wider adoption across diverse research and user communities.

5.4.3 Enhancing the Spatial and Radiometric Reliability of PlanetScope Imagery

Future efforts involving PSB.SD imagery should prioritize addressing the co-registration and radiometric stability challenges that became apparent in this study. Enhancing the co-registration of PSB.SD imagery is essential for achieving spatially accurate comparisons. In addition to refining the AROSICS-based approach proposed by Scheffler et al. (2017), the DL method proposed by Vakalopoulou et al. (2019) has demonstrated high potential in aligning satellite imagery. Therefore, DL approaches, such as the unsupervised CNN approach proposed by Vakalopoulou et al. (2019), should be considered for future improvements in image co-registration. Given the small GSD of PSB.SD data, in-field ground calibration using bright and dark reference objects, similar to PITs with known reflectance properties, may also serve as a complementary strategy to improve both geometric alignment and radiometric calibration. On the radiometric side, one accessible correction method is dark object subtraction (DOS), which compensates for atmospheric haze and scattering by subtracting the minimum reflectance observed in dark image regions (Chavez, 1988). This approach has been used effectively to correct for radiometric irregularities, as shown by Wicaksono and Lazuardi (2018) and cited in Frazier and Hemingway (2021). Another well-known approach involves radiometric normalization to well-calibrated external sensors such as Landsat-8 or Sentinel-2. This method typically uses a linear transformation to align PS reflectance values with those from reference sensors, thereby improving inter-sensor comparability and temporal consistency (Frazier and Hemingway, 2021). This approach has been successfully applied by Csillik et al. (2020), as cited in Frazier and Hemingway (2021), and further supported by Li et al. (2021), who implemented a two-stage normalization process based on Sentinel-2 reflectance data. Beyond these traditional methods, recent advances in ML may offer additional opportunities for radiometric harmonization. For instance, the Cubesat Enabled Spatio-Temporal Enhancement Method (CESTEM), developed by Houborg and McCabe (2018) and cited in Frazier and Hemingway (2021), applies non-parametric model trees to fuse PS, MODIS, and Landsat data into temporally consistent, high-resolution reflectance products. Similarly, the DL model presented by Latte and Lejeune (2020), also referenced in Frazier and Hemingway (2021), uses a residual CNN to normalize and sharpen PS imagery using Sentinel-2 as a spectral reference. These opportunities underscore the potential to enhance the future quality and consistency of PS imagery, particularly if PlanetLabs inherently applies reliable geometric and radiometric corrections as part of their processing to provide analysis-ready products.

5.4.4 Model Architecture and Learning Improvements

To further improve model performance, several directions are proposed in the area of model design. These include crop-specific modeling, hybrid approaches, spatial context integration, and uncertainty estimation. While the present study has demonstrated the feasibility of a general crop model for CC prediction, future work should also emphasize the development of crop-specific models for more precise and reliable outcomes. Since canopy structure and spectral characteristics vary substantially between crops, models trained on homogeneous crop types are likely to yield more accurate and interpretable results. In particular, crop-specific linear models based on VIs, such as MSAVI, offer a practical solution due to their minimal data requirements. These models do not require continuous training data across the full CC range and can be efficiently calibrated for specific crop types or crop groups. Beyond these simple models, hybrid modeling approaches that integrate VI-based predictions with outputs from ML or DL architectures could present a promising direction. By combining the biophysical relevance of VIs with the modeling flexibility of ML and DL methods, such approaches can enhance prediction accuracy across a range of phenological stages and crop types. Different VIs may contribute more informative signals at specific stages of canopy development, which can be effectively exploited in hybrid models. Additionally, incorporating spatial context into model architectures could further improve overall prediction robustness. Spatially aware techniques, such as CNNs, effectively capture multispectral features in a temporal and spatial context as shown in a 3D-CNN yield prediction model by Fernandez-Beltran et al. (2021), while the models with attention mechanisms can benefit from capturing the influence of neighboring pixels as shown by Nejad et al. (2024). Finally, to better assess the reliability of model outputs, uncertainty quantification should be integrated into future ML and DL models. Methods such as Bayesian Neural Networks (Blundell et al., 2015), Monte Carlo dropout (Gal and Ghahramani, 2016), and deep ensembles (Lakshminarayanan et al., 2017) provide scalable and effective ways to estimate predictive confidence. These techniques are instrumental for decision-making, where understanding the reliability of model predictions is critical.

5.4.5 Towards Open Data for Canopy Modeling

Future work should prioritize systematic benchmarking across various satellite platforms. This includes evaluating prediction performance using imagery from freely available sources, such as Sentinel-2, alongside commercial systems, such as PS. By assessing trade-offs between spatial resolution, revisit frequency, and prediction accuracy, researchers can identify the most suitable platforms for specific monitoring needs, such as continuous crop tracking, fine-scale phenotyping, or broad-scale agricultural assessments. Moreover, expanding the methodology to include publicly accessible, non-commercial satellite data is essential for promoting equitable access to crop monitoring tools. Open platforms such as Sentinel-2 help lower financial barriers and support independent research across economically diverse regions. In this context, CC modeling for plot-level experiments could particularly benefit from the enhanced spatial resolution of upcoming open-access missions.

6 CONCLUSION

This study bridges the gap between detailed field-level phenotyping and scalable satellite-based monitoring by integrating UAV-derived RGB imagery with multispectral PSB.SD satellite observations. A two-stage segmentation approach achieved high accuracy in separating vegetation from non-vegetation pixels under varying field and lighting conditions in UAV imagery. Combined with radiometric and spatial preprocessing of multispectral PSB.SD data, a reliable reference dataset was established for training satellite-based CC models. Various modeling approaches were tested and evaluated for predicting CC from satellite data, including linear regression models based on VIs (NDVI, EVI, MSAVI), ML methods such as RF and Histogram-Based Gradient Boosting Regressor (HGBR), and a feedforward neural network (NN). Among these, MSAVI (R^2 = 0.861), HGBR (R^2 = 0.919), and NN (R^2 = 0.923) demonstrated strong baseline performance, showing resilience across different crop species, locations, and growth stages. Cross-validation results indicate that while VIs such as MSAVI provide reasonable predictions under data-scarce conditions, more complex models outperform simple linear regression when high-quality and abundant data are available. Ultimately, the models were applied to winter barley in a field plot experiment with varying (bio)fertilizer treatments, enabling time-series CC estimation across multiple plots, each 9 m wide. This application demonstrated the models' capability to capture treatment-specific canopy development, including early-stage growth differences. In essence, this research demonstrates the feasibility and utility of high-resolution multispectral satellite data for plot-scale crop monitoring. The proposed approach enables continuous, scalable, and non-invasive phenotyping, supporting more efficient crop monitoring under field conditions and thereby contributing to the broader goal of sustainable food production. The MSAVI model shows promise for CC modeling in low-data scenarios, while the NN approach is found to be better suited for high-precision predictions in crop- and site-specific applications. Looking ahead, future work should prioritize expanding the training dataset to encompass a broader range of crop types and farming systems, thereby enhancing model generalizability across diverse agroecological environments. Incorporating DL architectures, which expand upon both temporal and spatial dimensions, can further enhance the accuracy and robustness of satellite-based CC predictions. Leveraging open data sources, such as Sentinel-2, can help reduce access barriers and foster global collaboration in the development of satellite-based CC monitoring, unlocking valuable insights that can drive advancements in sustainable agriculture.

7 REFERENCES

Agostini, L., Diener, M., Mayer, J., and Bünemann-König, E. (2024). Optimierung der Stickstoffausnutzung und Klimawirkung von Recyclingdüngern im Biolandbau. *FiBL, Agroscope*, unpublished data.

Anurogo, W., Lubis, M. Z., and Mufida, M. K. (2018). Modified Soil-Adjusted Vegetation Index In Multispectral Remote Sensing Data for Estimating Tree Canopy Cover Density at Rubber Plantation. *Journal of Geoscience, Engineering, Environment, and Technology*, 3(1):15.

Araus, J. L. and Cairns, J. E. (2014). Field high-throughput phenotyping: The new crop breeding frontier. *Trends in Plant Science*, 19(1):52–61.

Bannari, A., Huete, A. R., Morin, D., and Zagolski, F. (1996). Effets de la couleur et de la brillance du sol sur les indices de vegetation. *International Journal of Remote Sensing*, 17(10):1885–1906.

Bebie, M. and Kyparissis, A. (2024). A comparative analysis of machine learning and vegetation index-based modeling approaches for durum wheat yield assessment using Sentinel-2 imagery. In Neale, C. M. U., Maltese, A., Bostater, C. R., and Nichol, C., editors, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 13191 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 1319107.

Berger, K., Atzberger, C., Danner, M., D'Urso, G., Mauser, W., Vuolo, F., and Hank, T. (2018). Evaluation of the PROSAIL model capabilities for future hyperspectral model environments: A review study. *Remote Sensing*, 10(1).

Blundell, C., Cornebise, J., Kavukcuoglu, K., Com, W., and Deepmind, G. (2015). Weight Uncertainty in Neural Networks Daan Wierstra. In *32nd International Conference on Machine Learning*, Lille. JMLR: W&CP volume 37.

BLW, B. f. L. (2024a). Digitale Bodeneignungskarte der Schweiz - Kulturland. https://www.geo.admin.ch. https://data.geo.admin.ch/browser/index.html/collections/ch.blw.bodeneignung-kulturland. Accessed: October 23, 2024.

BLW, B. f. L. (2024b). Digitale Bodeneignungskarte der Schweiz - Kulturtyp. https://www.geo.admin.ch. https://data.geo.admin.ch/browser/index.html/collections/ch.blw.bodeneignung-kulturtyp?.language=en. Accessed: October 23, 2024.

Bocco, M., Ovando, G., Sayago, S., Willington, E., and Heredia, S. (2012). Estimating soybean ground cover from satellite images using neural-networks models. *International Journal of Remote Sensing*, 33(6):1717–1728.

Bouman, B. (1995). Crop modelling and remote sensing for yield prediction. *Netherlands Journal of Agricultural Science*, (43):143–161.

Brisco, B., Brown, R. J., Hirose, T., Mc Naim, H., and Staenz, K. (1998). Precision agriculture and the role of remote sensing: A review. *Canadian Journal of Remote Sensing*, 24(3):315–327.

Bünemann-König, E. and Mayer, J. (2021). Optimaler Einsatz von Recyclingdüngern im Biolandbau: Ertragswirkung und Stickstoffeffizienz. Technical report, FiBL, Agroscope.

Chavez, P. S. (1988). An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24(3):459–479.

Chen, R., Sun, L., Chen, Z., Wuyun, D., and Sun, Z. (2024). Early Identification of Corn and Soybean Using Crop Growth Curve Matching Method. *Agronomy*, 14(1):1–17.

Chianucci, F., Disperati, L., Guzzi, D., Bianchini, D., Nardino, V., Lastri, C., Rindinella, A., and Corona, P. (2016). Estimation of canopy attributes in beech forests using true colour digital images from a small fixed-wing UAV. *International Journal of Applied Earth Observation and Geoinformation*, 47:60–68.

Chianucci, F., Lucibelli, A., and Dell'Abate, M. T. (2018). Estimation of ground canopy cover in agricultural crops using downward-looking photography. *Biosystems Engineering*, 169:209–216.

Cobb, J. N., DeClerck, G., Greenberg, A., Clark, R., and McCouch, S. (2013). Next-generation phenotyping: Requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. *Theoretical and Applied Genetics*, 126(4):867–887.

Csillik, O., Kumar, P., and Asner, G. P. (2020). Challenges in estimating tropical forest canopy height from planet dove imagery. *Remote Sensing*, 12(7).

Delgado, J. A., Short, N. M., Roberts, D. P., and Vandenberg, B. (2019). Big Data Analysis for Sustainable Agriculture on a Geospatial Cloud Framework. *Frontiers in Sustainable Food Systems*, 3.

Diener, M., Agostini, L., Bünemann-König, E., and Mayer, J. (2024). Optimierung der Stickstoffausnutzung und Klimawirkung von Recyclingdüngern im Biolandbau. *FiBL, Agroscope*, unpublished data.

DJI (2022). DJI MAVIC 3E/3T - User Manual 2022.09. Technical report.

Ehmann, A., Thumm, U., and Lewandowski, I. (2018). Fertilizing Potential of Separated Biogas Digestates in Annual and Perennial Biomass Production Systems. *Frontiers in Sustainable Food Systems*, 2.

El Hajj, M. (2008). *Processing of time series of high spatial resolution satellite images through multi-source information fusion for monitoring agricultural practices: The case of sugarcane harvest on Reunion Island*. PhD thesis, AgroParisTech.

Federal Office of Topography (swisstopo) (2025). SwissImage 10 cm Imagery. https://www.swisstopo.admin.ch/de/orthobilder-swissimage-10-cm. Accessed: November 2, 2024.

Fernandez-Beltran, R., Baidar, T., Kang, J., and Pla, F. (2021). Rice-yield prediction with multi-temporal sentinel-2 data and 3D CNN: A case study in Nepal. *Remote Sensing*, 13(7).

Frazier, A. E. and Hemingway, B. L. (2021). A technical review of planet smallsat data: Practical considerations for processing and using planetscope imagery. *Remote Sensing*, 13(19).

Gal, Y. and Ghahramani, Z. (2016). Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. In *33rd International Conference on Machine Learning*, New York. University of Cambridge, JMLR: W&CP volume 48.

Gao, L., Wang, X., Johnson, B. A., Tian, Q., Wang, Y., Verrelst, J., Mu, X., and Gu, X. (2020). Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 159:364–377.

Gillies, S., van der Wel, C., den Bossche, J., Taves, M. W., Arnott, J., Ward, B. C., and others (2025). Shapely.

Gitelson, A. A. (2013). Remote estimation of crop fractional vegetation cover: The use of noise equivalent as an indicator of performance of vegetation indices. *International Journal of Remote Sensing*, 34(17):6054–6066.

Gitelson, A. A. (2016). Hyperspectral Vegetation Indices. In *Hyperspectral Remote Sensing of Vegetation*, chapter 15, pages 345–364. CRC Press.

Haghighattalab, A., González Pérez, L., Mondal, S., Singh, D., Schinstock, D., Rutkoski, J., Ortiz-Monasterio, I., Singh, R. P., Goodin, D., and Poland, J. (2016). Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries. *Plant Methods*, 12(1).

Hashimoto, N., Saito, Y., Yamamoto, S., Ishibashi, T., Ito, R., Maki, M., and Homma, K. (2023). Relationship between Leaf Area Index and Yield Components in Farmers' Paddy Fields. *AgriEngineering*, 5(4):1754–1765.

Hatfield, J. L., Gitelson, A. A., Schepers, J. S., and Walthall, C. L. (2008). Application of spectral remote sensing for agronomic decisions. *Agronomy Journal*, 100(3 SUPPL.).

Houborg, R. and McCabe, M. F. (2016). High-Resolution NDVI from planet's constellation of earth observing nano-satellites: A new data source for precision agriculture. *Remote Sensing*, 8(9).

Houborg, R. and McCabe, M. F. (2018). A Cubesat enabled Spatio-Temporal Enhancement Method (CESTEM) utilizing Planet, Landsat and MODIS data. *Remote Sensing of Environment*, 209(July 2017):211–226.

Huete, A. R., Liu, H. Q., Batchily, K., and van Leeuwen, W. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, 59(3):440–451.

Jacquemoud, S. and Baret, F. (1990). PROSPECT: A model of leaf optical properties spectra. *Remote Sensing of Environment*, 34(2):75–91.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., François, C., and Ustin, S. L. (2009). PROSPECT + SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment*, 113(SUPPL. 1).

Kenneth Reitz (2011). Requests: HTTP for Humans.

Lakshminarayanan, B., Pritzel, A., and Deepmind, C. B. (2017). Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles. In *31st Conference on Neural Information Processing Systems*. NIPS.

Latte, N. and Lejeune, P. (2020). PlanetScope radiometric normalization and sentinel-2 super-resolution (2.5 m): A straightforward spectral-spatial fusion of multi-satellite multi-sensor images using residual convolutional neural networks. *Remote Sensing*, 12(15).

Ledain, S., Gilgen, A., and Aasen, H. (2025). Enhanced large-scale winter wheat LAI retrieval from Sentinel-2: a soil-informed radiative transfer based approach. unpublished.

Li, L., Mu, X., Jiang, H., Chianucci, F., Hu, R., Song, W., Qi, J., Liu, S., Zhou, J., Chen, L., Huang, H., and Yan, G. (2023). Review of ground and aerial methods for vegetation cover fraction (fCover) and related quantities estimation: definitions, advances, challenges, and future perspectives. *ISPRS Journal of Photogrammetry and Remote Sensing*, 199:133–156.

Li, Z. and Chen, Z. (2011). Remote sensing indicators for crop growth monitoring at different scales. In *International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 4062–4065.

Li, Z., Scheffler, D., Coops, N. C., Leach, N., and Sachs, T. (2021). Towards analysis ready data of optical CubeSat images: Demonstrating a hierarchical normalization framework at a wetland site. *International Journal of Applied Earth Observation and Geoinformation*, 103.

Liaghat, S. and Balasundram, S. K. (2010). A review: The role of remote sensing in precision agriculture. *American Journal of Agricultural and Biological Science*, 5(1):50–55.

Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110.

Lu, Z., Deng, L., and Lu, H. (2022). An Improved LAI Estimation Method Incorporating with Growth Characteristics of Field-Grown Wheat. *Remote Sensing*, 14(16).

Mahal, N. K., Osterholz, W. R., Miguez, F. E., Poffenbarger, H. J., Sawyer, J. E., Olk, D. C., Archontoulis, S. V., and Castellano, M. J. (2019). Nitrogen fertilizer suppresses mineralization of soil organic matter in maize agroecosystems. *Frontiers in Ecology and Evolution*, 7(MAR).

Martos, V., Ahmad, A., Cartujo, P., and Ordoñez, J. (2021). Ensuring agricultural sustainability through remote sensing in the era of agriculture 5.0. *Applied Sciences (Switzerland)*, 11(13).

Mcgwire, K., Minor, T., and Fenstermaker, L. (1991). Hyperspectral Mixture Modeling for Quantifying Sparse Vegetation Cover in Arid Environments. *REMOTE SENS. ENVIRON*, 72:360–374.

Meyer, G. E. and Neto, J. C. (2008). Verification of color vegetation indices for automated crop imaging applications. *Computers and Electronics in Agriculture*, 63(2):282–293.

Montandon, L. M. and Small, E. E. (2008). The impact of soil reflectance on the quantification of the green vegetation fraction from NDVI. *Remote Sensing of Environment*, 112(4):1835–1845.

Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*, 114(4):358–371.

Nejad, S. M. M., Abbasi-Moghadam, D., Sharifi, A., and Tariq, A. (2024). Capsular attention Conv-LSTM network (CACN): A deep learning structure for crop yield estimation based on multispectral imagery. *European Journal of Agronomy*, 161.

Nielsen, D. C., Miceli-Garcia, J. J., and Lyon, D. J. (2012). Canopy cover and leaf area index relationships for wheat, triticale, and corn. *Agronomy Journal*, 104(6):1569–1573.

Pan, Z., Hu, Y., and Cao, B. (2017). Construction of smooth daily remote sensing time series data: a higher spatiotemporal resolution perspective. *Open Geospatial Data, Software and Standards*, 2(1).

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury Google, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Xamla, A. K., Yang, E., Devito, Z., Raison Nabla, M., Tejani, A., Chilamkurthy, S., Ai, Q., Steiner, B., Facebook, L. F., Facebook, J. B., and Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. Technical report.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85):2825–2830.

Planet Labs PBC (2024). Planet Application Program Interface: In Space for Life on Earth. https://api.planet.com Accessed: March 19, 2025.

Planet Labs PBC (2025a). PlanetScope - Earth Online. https://earth.esa.int/eogateway/missions/planetscope. Accessed: March 19, 2025.

Planet Labs PBC (2025b). PlanetScope Overview. https://developers.planet.com/docs/data/planetscope/. Accessed: March 19, 2025.

Puig, F., Gonzalez Perea, R., Daccache, A., Soriano, M. A., and Rodríguez Díaz, J. A. (2025). Convolutional neural networks for accurate estimation of canopy cover. *Smart Agricultural Technology*, 10.

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., and Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, 48(2):119–126.

Rasti, S., Bleakley, C. J., Holden, N. M., Whetton, R., Langton, D., and O'Hare, G. (2022). A survey of high resolution image processing techniques for cereal crop growth monitoring. *Information Processing in Agriculture*, 9(2):300–315.

Raza, A., Shahid, M. A., Zaman, M., Miao, Y., Huang, Y., Safdar, M., Maqbool, S., and Muhammad, N. E. (2025). Improving Wheat Yield Prediction with Multi-Source Remote Sensing Data and Machine Learning in Arid Regions. *Remote Sensing*, 17(5).

Ren, N., Wang, Y., Ye, Y., Zhao, Y., Huang, Y., Fu, W., and Chu, X. (2020). Effects of Continuous Nitrogen Fertilizer Application on the Diversity and Composition of Rhizosphere Soil Bacteria. *Frontiers in Microbiology*, 11.

Rico-Fernández, M. P., Rios-Cabrera, R., Castelán, M., and Guerrero-Reyes, H.-I. (2019). A contextualized approach for segmentation of foliage in different crop species. *Computers and Electronics in Agriculture*, 156:378–386.

Riva, C., Orzi, V., Carozzi, M., Acutis, M., Boccasile, G., Lonati, S., Tambone, F., D'Imporzano, G., and Adani, F. (2016). Short-term experiments in using digestate products as substitutes for mineral (N) fertilizer: Agronomic performance, odours, and ammonia emission impacts. *Science of the Total Environment*, 547:206–214.

Rondeaux, G., Steven, M., and Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, 55(2):95–107.

Rouse, J. W., Hass, R. H., Schell, J. A., and Deering, D. W. (1973). Monitoring vegetation systems in the great plains with ERTS. *the Third ERTS Symposium. NASA SP-351*, pages 309–317.

Ruiz-Colmenero, M., Bienes, R., Eldridge, D. J., and Marques, M. J. (2013). Vegetation cover reduces erosion and enhances soil organic carbon in a vineyard in the central Spain. *Catena*, 104:153–160.

Sadeghi-Tehran, P., Virlet, N., Sabermanesh, K., and Hawkesford, M. J. (2017). Multi-feature machine learning model for automatic segmentation of green fractional vegetation cover for high-throughput field phenotyping. *Plant Methods*, 13(1):1–16.

Saint Pierre, C., Crossa, J. L., Bonnett, D., Yamaguchi-Shinozaki, K., and Reynolds, M. P. (2012). Phenotyping transgenic wheat for drought resistance. *Journal of Experimental Botany*, 63(5):1799–1808.

Sankaran, S., Khot, L. R., and Carter, A. H. (2015). Field-based crop phenotyping: Multispectral aerial imaging for evaluation of winter wheat emergence and spring stand. *Computers and Electronics in Agriculture*, 118:372–379.

Sankaran, S., Quirós, J. J., and Miklas, P. N. (2019). Unmanned aerial system and satellite-based high resolution imagery for high-throughput phenotyping in dry bean. *Computers and Electronics in Agriculture*, 165(February):104965.

Scheffler, D., Hollstein, A., Diedrich, H., Segl, K., and Hostert, P. (2017). AROSICS: An automated and robust open-source image co-registration software for multi-sensor satellite data. *Remote Sensing*, 9(7).

Sishodia, R. P., Ray, R. L., and Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A review. *Remote Sensing*, 12(19):1–31.

Sun, C., Bian, Y., Zhou, T., and Pan, J. (2019). Using of multi-source and multi-temporal remote sensing data improves crop-type mapping in the subtropical agriculture region. *Sensors*, 19(10).

swisstopo (2025). Landeskarte grau. https://www.swisstopo.admin.ch. Accessed: April 2, 2025.

Tattaris, M., Reynolds, M. P., and Chapman, S. C. (2016). A direct comparison of remote sensing approaches for high-throughput phenotyping in plant breeding. *Frontiers in Plant Science*, 7(AUG2016).

Tenreiro, T. R., García-Vila, M., Gómez, J. A., Jiménez-Berni, J. A., and Fereres, E. (2021). Using NDVI for the assessment of canopy cover in agricultural crops within modelling research. *Computers and Electronics in Agriculture*, 182.

Todd, S. W. and Hoffer, R. M. (1998). Responses of spectral indices to variations in vegetation cover and soil background. *Photogrammetric Engineering and Remote Sensing*, 64(9):915 – 921.

Vakalopoulou, M., Christodoulidis, S., Sahasrabudhe, M., Mougiakakou, S., and Paragios, N. (2019). Image Registration of Satellite Imagery with Deep Convolutional Neural Networks. In *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 4939–4942.

Van Der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., and Yu, T. (2014). Scikit-image: Image processing in python. *PeerJ*, 2014(1).

Verhoef, W. (1984). Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. *Remote Sensing of Environment*, 16(2):125–141.

Verrelst, J., Rivera, J. P., Gitelson, A., Delegido, J., Moreno, J., and Camps-Valls, G. (2016). Spectral band selection for vegetation properties retrieval using Gaussian processes regression. *International Journal of Applied Earth Observation and Geoinformation*, 52:554–567.

Victor, B., Nibali, A., and He, Z. (2024). A Systematic Review of the Use of Deep Learning in Satellite Imagery for Agriculture. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 18:2297–2316.

Walter, A., Liebisch, F., and Hund, A. (2015). Plant phenotyping: From bean weighing to image analysis. *Plant Methods*, 11(1).

Wang, P., Wang, Z., Pan, Q., Sun, X., Chen, H., Chen, F., Yuan, L., and Mi, G. (2019). Increased biomass accumulation in maize grown in mixed nitrogen supply is mediated by auxin synthesis. *Journal of Experimental Botany*, 70(6):1859–1873.

Wang, Y., Cao, Z., Bai, X., Yu, Z., and Li, Y. (2013). An automatic detection method to the field wheat based on image processing. In *MIPPR 2013: Automatic Target Recognition and Navigation*, volume 8918, page 89180F. SPIE.

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., Feldmann, K. A., French, A. N., Heun, J. T., Hunsaker, D. J., Jenks, M. A., Kimball, B. A., Roth, R. L., Strand, R. J., Thorp, K. R., Wall, G. W., and Wang, G. (2012). Field-based phenomics for plant genetics research. *Field Crops Research*, 133:101–112.

Wicaksono, P. and Lazuardi, W. (2018). Assessment of PlanetScope images for benthic habitat and seagrass species mapping in a complex optically shallow water environment. *International Journal of Remote Sensing*, 39(17):5739–5765.

Wildhaber, S., Graf, L., and Aasen, H. (2023). Assessing the Potential of High-Resolution Satellite Constellations for Agricultural Monitoring. In *43. Wissenschaftlich-Technische Jahrestagung der DGPF*, pages 197–205, München.

Yan, S., Yao, X., Zhu, D., Liu, D., Zhang, L., Yu, G., Gao, B., Yang, J., and Yun, W. (2021). Large-scale crop mapping from multi-source optical satellite imageries using machine learning with discrete grids. *International Journal of*

Applied Earth Observation and Geoinformation, 103.

Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., Xu, B., Yang, X., Zhu, D., Zhang, X., Zhang, R., Feng, H., Zhao, X., Li, Z., Li, H., and Yang, H. (2017). Unmanned aerial vehicle remote sensing for field-based crop phenotyping: Current status and perspectives.

Yu, K., Kirchgessner, N., Grieder, C., Walter, A., and Hund, A. (2017). An image analysis pipeline for automated classification of imaging light conditions and for quantification of wheat canopy cover time series in field phenotyping. *Plant Methods*, 13(1).

Zenkl, R., Timofte, R., Kirchgessner, N., Roth, L., Hund, A., Van Gool, L., Walter, A., and Aasen, H. (2022). Outdoor Plant Segmentation With Deep Learning for High-Throughput Field Phenotyping on a Diverse Wheat Dataset. *Frontiers in Plant Science*, 12(January):1–19.

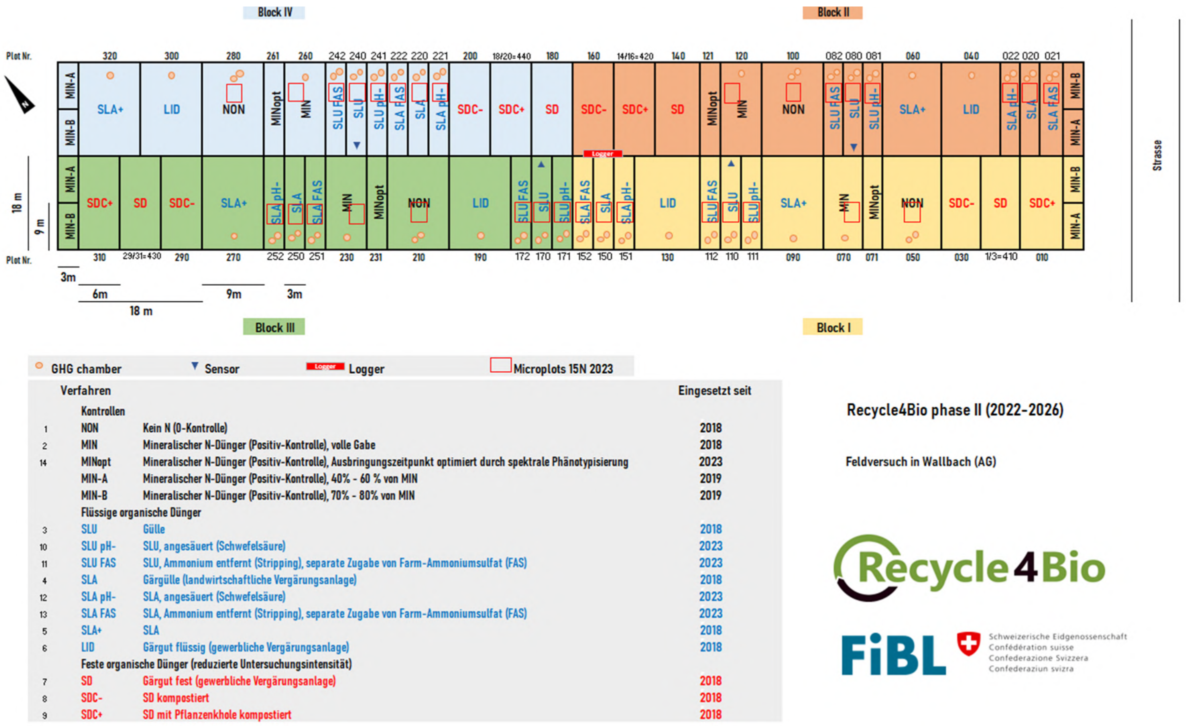
Zhang, C., Marzougui, A., and Sankaran, S. (2020). High-resolution satellite imagery applications in crop phenotyping: An overview. *Computers and Electronics in Agriculture*, 175(June):105584.

Zhang, N., Wang, M., and Wang, N. (2002). Precision agriculture - a worldwide overview. *Computers and Electronics in Agriculture*, 36:113–132.

Zuazo, V. H. D. and Pleguezuelo, C. R. R. (2008). Soil-erosion and runoff prevention by plant covers. A review. *Agronomy for Sustainable Development*, 28(1):65–86.

8 APPENDIX

Experimental design

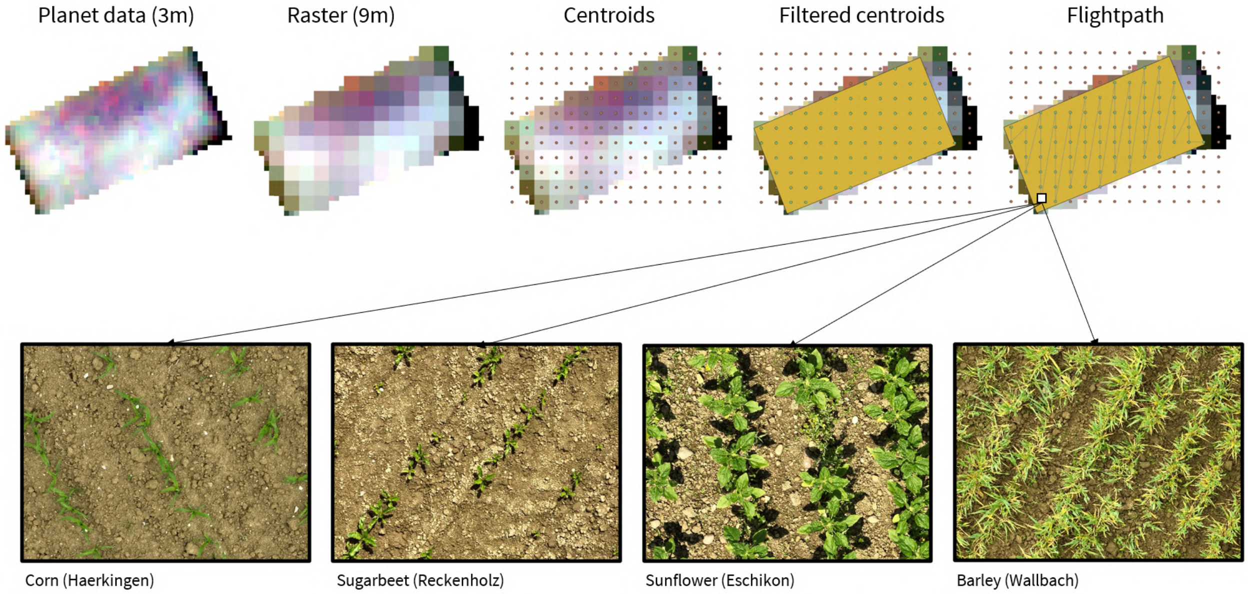


○ GHG chamber ▼ Sensor 🟥 Logger ☐ Microplots 15N 2023

	Verfahren		Eingesetzt seit
	Kontrollen		
1	NON	Kein N (0-Kontrolle)	2018
2	MIN	Mineralischer N-Dünger (Positiv-Kontrolle), volle Gabe	2018
14	MINopt	Mineralischer N-Dünger (Positiv-Kontrolle), Ausbringungszeitpunkt optimiert durch spektrale Phänotypisierung	2023
	MIN-A	Mineralischer N-Dünger (Positiv-Kontrolle), 40% - 60 % von MIN	2019
	MIN-B	Mineralischer N-Dünger (Positiv-Kontrolle), 70% - 80% von MIN	2019
	Flüssige organische Dünger		
3	SLU	Gülle	2018
10	SLU pH-	SLU, angesäuert (Schwefelsäure)	2023
11	SLU FAS	SLU, Ammonium entfernt (Stripping), separate Zugabe von Farm-Ammoniumsulfat (FAS)	2023
4	SLA	Gärgülle (landwirtschaftliche Vergärungsanlage)	2018
12	SLA pH-	SLA, angesäuert (Schwefelsäure)	2023
13	SLA FAS	SLA, Ammonium entfernt (Stripping), separate Zugabe von Farm-Ammoniumsulfat (FAS)	2023
5	SLA+	SLA	2018
6	LID	Gärgut flüssig (gewerbliche Vergärungsanlage)	2018
	Feste organische Dünger (reduzierte Untersuchungsintensität)		
7	SD	Gärgut fest (gewerbliche Vergärungsanlage)	2018
8	SDC-	SD kompostiert	2018
9	SDC+	SD mit Pflanzenkhole kompostiert	2018

Figure 35. **Plot-Wise Experimental Design of the ”Bio4Recycle” Site in Wallbach**, showing the spatial layout and treatment distribution across plots. *Reproduced with permission from Bünemann-König and Mayer (2021).*

UAV Mission Planning


Planet data (3m)
Raster (9m)
Centroids
Filtered centroids
Flightpath

Corn (Haerkingen)
Sugarbeet (Reckenholz)
Sunflower (Eschikon)
Barley (Wallbach)

Figure 36. **UAV Snapshot Mission Planning** The figure illustrates the process of the Snapshot (SNP) acquisition. The centroids of each grid cell of the 9m raster are being connected to form a continuous flight path. At each spatially filtered centroid, a UAV snapshot was acquired as illustrated in examples of various crops (Corn, Sugarbeet, Sunflower, and Barley). Data © 2024 Planet Labs PBC.

UAV Image Acquisition

Table 19. **UAV Flight Protocol**. The table summarizes the UAV flight protocol, listing key flight details such as the date, flight ID, location, site, flight mode, start location, weather, and wind conditions for each flight conducted. The table includes data for multiple flights spanning from April to July 2024, covering various sites including Eschikon, Haerkingen, Reckenholz, and Wallbach.

Date	FlightID	Location	Site	Mode	Start Location	Weather	Wind
20240425	061	Wallbach	Recycle	Mapping		variable	medium
20240425	060, 061	Wallbach	Recycle	Snapshot	47.561726,7.888348	variable	medium
20240430	065	Eschikon	Steinmueri1	Snapshot	47.449179,8.686255	sunny	medium
20240430	066	Eschikon	Steinmueri1	Mapping	47.449179,8.686260	sunny	medium
20240514	069	Eschikon	Laubisgruet2	Snapshot	47.445047,8.685014	sunny	medium
20240514	070, 071	Eschikon	Laubisgruet2	Mapping		sunny	medium
20240514	072	Eschikon	Steinmueri1	Snapshot	47.449887,8.686204	sunny	medium
20240514	075, 076	Eschikon	Steinmueri1	Mapping		sunny	medium
20240519	077	Reckenholz	Vedichtung	Snapshot	47.429934,8.520483	variable	medium
20240519	078	Reckenholz	Vedichtung	Mapping		variable	medium
20240519	083	Reckenholz	Autobahn	Snapshot	47.430411,8.515502	variable	medium
20240519	084	Reckenholz	Autobahn	Mapping		variable	medium
20240519	085	Reckenholz	Eingang	Snapshot	47.428717,8.515573	variable	medium
20240519	086	Reckenholz	Eingang	Mapping		variable	medium
20240519	087	Wallbach	Recycle	Mapping		variable	light
20240528	088	Wallbach	Recycle	Snapshot	47.561735,7.888537	variable	light
20240529	091	Eschikon	Holgenbuehl	Snapshot	47.447278,8.686492	cloudy	medium
20240529	092	Eschikon	Holgenbuehl	Mapping		cloudy	medium

Date	FlightID	Location	Site	Mode	Start Location	Weather	Wind
20240529	093, 094	Eschikon	Steinmueri1	Mapping		cloudy	medium
20240604	095	Reckenholz	Autobahn	Snapshot	47.430316,8.515598	variable	light
20240604	097	Reckenholz	Autobahn	Mapping		variable	light
20240604	098	Reckenholz	Eingang	Snapshot	47.429470,8.514512	variable	light
20240604	099, 100	Reckenholz	Eingang	Mapping		variable	light
20240605	109, 110	Haerkingen	Altgraben1	Snapshot	47.313250,7.800572	variable	light
20240605	101	Haerkingen	Chaeppeli	Snapshot	47.300958,7.785722	variable	light
20240605	102	Haerkingen	Chaeppeli	Mapping		variable	light
20240605	103	Haerkingen	Neufeld4	Mapping		variable	light
20240605	104, 106	Haerkingen	Neufeld4	Snapshot	47.301448,7.785202	variable	light
20240605	108	Haerkingen	Altgraben1	Mapping		variable	light
20240606	111	Eschikon	Laubisgruet2	Mapping		variable	light
20240606	112	Eschikon	Laubisgruet2	Snapshot	47.445158,8.685010	variable	light
20240606	114	Eschikon	Steinmueri1	Snapshot	47.449901,8.686218	sunny	no
20240606	115	Eschikon	Steinmueri1	Mapping		sunny	no
20240606	115	Eschikon	Steinmmueri3	Mapping		variable	light
20240606	117	Eschikon	Steinmmueri3	Snapshot	47.450031,8.687137	variable	light
20240606	118	Eschikon	Holgenbuehl	Mapping		variable	no
20240606	119	Eschikon	Holgenbuehl	Snapshot	47.447243,8.686632	variable	no
20240612	120	Reckenholz	Autobahn	Snapshot	47.430317,8.515600	sunny	light
20240612	121	Reckenholz	Autobahn	Mapping		sunny	light
20240612	122	Reckenholz	Eingang	Mapping		variable	light
20240612	123	Reckenholz	Eingang	Snapshot	47.429446,8.514527	variable	light
20240618	124	Haerkingen	Altgraben1	Mapping		sunny	light
20240618	125, 126	Haerkingen	Altgraben1	Snapshot	47.314371,7.801784	sunny	light
20240618	127	Haerkingen	Chaeppeli	Snapshot	47.301474,7.785336	sunny	light
20240618	128	Haerkingen	Chaeppeli	Mapping		sunny	light
20240618	129	Haerkingen	Neufeld4	Snapshot	47.301473,7.785335	variable	light
20240618	130	Haerkingen	Neufeld4	Mapping		variable	light
20240619	131	Eschikon	Steinmueri1	Snapshot	47.449999,8.686871	sunny	light
20240619	132, 133	Eschikon	Steinmueri1	Mapping		sunny	light
20240619	134	Eschikon	Steinmmueri3	Snapshot	47.449999,8.686872	sunny	light
20240619	135	Eschikon	Steinmmueri3	Mapping		sunny	light
20240619	136	Eschikon	Holgenbuehl	Snapshot	47.447207,8.686743	sunny	light
20240619	137	Eschikon	Holgenbuehl	Mapping		sunny	light
20240619	138	Eschikon	Laubisgruet2	Snapshot	47.445345,8.685015	sunny	light
20240619	139, 140	Eschikon	Laubisgruet2	Mapping		sunny	light
20240627	141	Reckenholz	Eingang	Snapshot	47.429074,8.514185	variable	medium
20240627	142	Reckenholz	Eingang	Mapping		variable	medium
20240627	143	Reckenholz	Autobahn	Snapshot	47.429712,8.514094	variable	medium
20240627	145	Reckenholz	Autobahn	Mapping		variable	medium
20240628	146	Eschikon	Steinmmueri3	Snapshot	47.449999,8.686871	sunny	light
20240628	147	Eschikon	Steinmmueri3	Mapping		sunny	light
20240628	148	Eschikon	Steinmueri1	Snapshot	47.447284,8.686491	sunny	light
20240628	149	Eschikon	Steinmueri1	Mapping		sunny	light
20240628	150	Eschikon	Laubisgruet2	Snapshot	47.445393,8.685019	variable	light
20240628	151	Eschikon	Laubisgruet2	Mapping		variable	light
20240708	152	Haerkingen	Neufeld4	Snapshot	47.301468,7.785389	sunny	light
20240708	153	Haerkingen	Neufeld4	Mapping		sunny	light
20240708	154	Haerkingen	Chaeppeli	Mapping		sunny	light
20240708	155	Haerkingen	Chaeppeli	Snapshot	47.301468,7.785390	sunny	light
20240708	156	Haerkingen	Altgraben1	Snapshot	47.314726,7.801551	sunny	light

Date	FlightID	Location	Site	Mode	Start Location	Weather	Wind
20240708	157	Haerkingen	Altgraben1	Mapping		sunny	light
20240709	158	Eschikon	Steinmueri3	Snapshot	47.450039,8.687144	sunny	light
20240709	159	Eschikon	Steinmueri3	Mapping		sunny	light
20240709	160	Eschikon	Holgenbuehl	Snapshot	47.447318,8.686472	sunny	light
20240709	161	Eschikon	Holgenbuehl	Mapping		sunny	light
20240709	162	Eschikon	Laubisgruet2	Snapshot	47.445133,8.685005	sunny	light
20240709	163	Eschikon	Laubisgruet2	Mapping		sunny	light
20240715	164	Reckenholz	Eingang	Snapshot	47.429061,8.514163	variable	no
20240715	165	Reckenholz	Eingang	Mapping		variable	no

Field-Level CC Model

Feature Matching

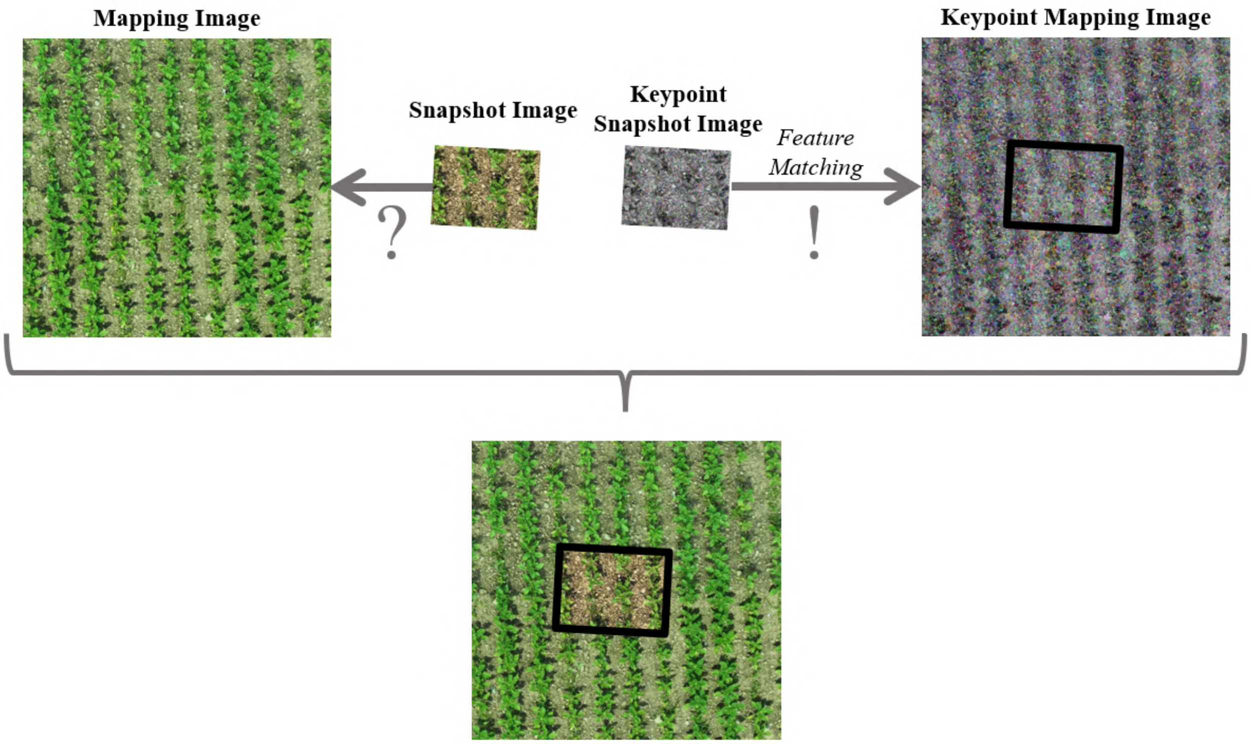
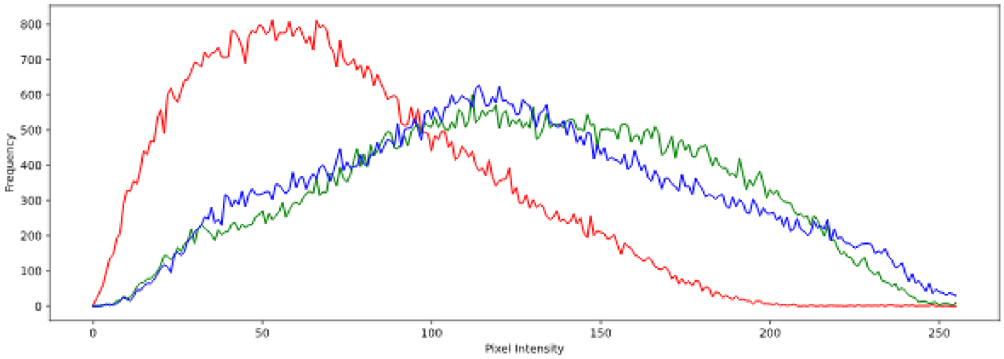
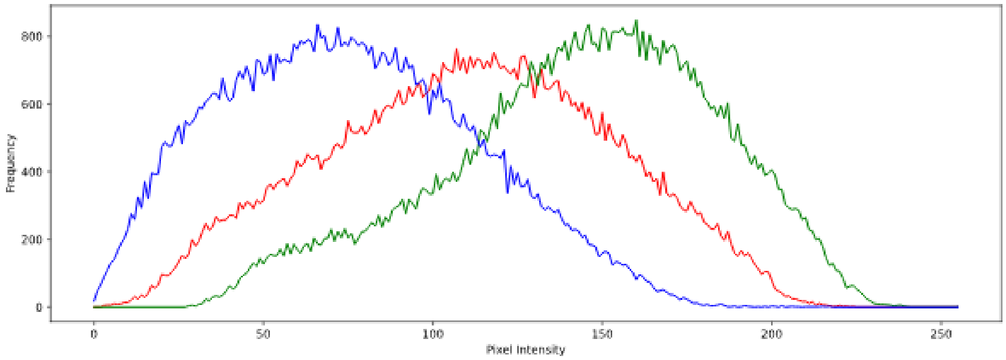


Figure 37. **Feature Matching Example** between a snapshot image and a mapping image (left). The process involves detecting keypoints in both images (right), followed by matching these keypoints to find the corresponding region (highlighted with a black frame) in the mapping image. The keypoints are highlighted in circles, with the size of the circle indicating the scale at which the keypoint was detected. The example is from Eschikon, Steinmueri1 (sunflower) on 06.06.2024.

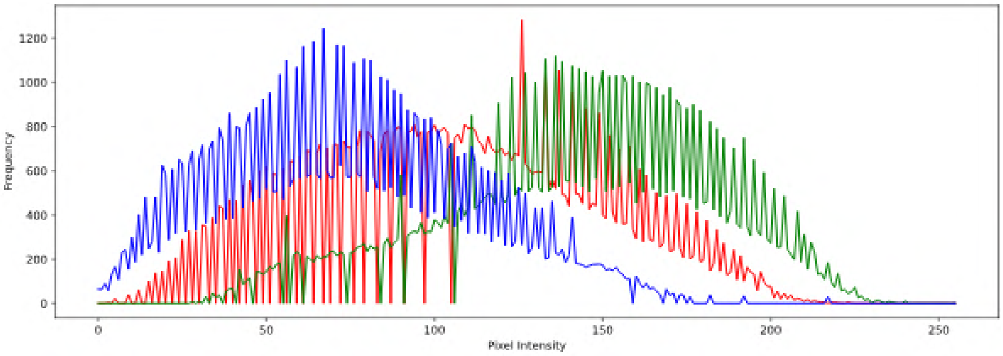
Histogram Matching



(a) Resized Snapshot Image

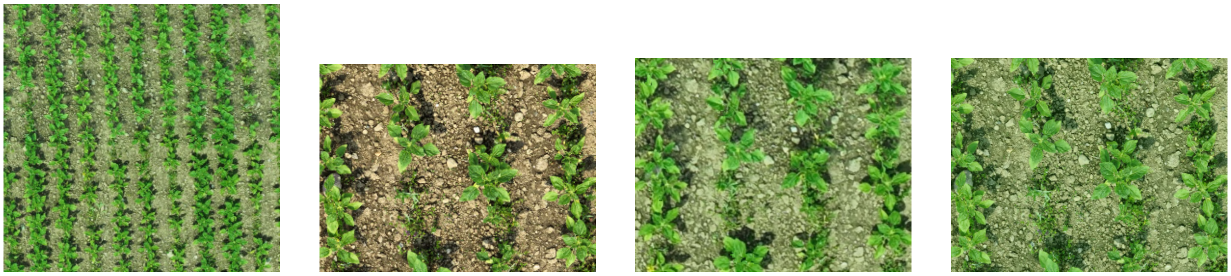


(b) Mapping Image



(c) Matched Snapshot Image

Figure 38. Histogram Matching Example showing the histograms of pixel intensities for the resized snapshot (SNP) image (a), low-resolution mapping image (b), and the matched SNP image (c). The red, green, and blue curves represent the frequency of the corresponding color channels (0-256).



(a) Mapping Image

(b) Snapshot Image

(c) Feature Matched Mapping Image

(d) Histogram Matched Snapshot Image

Figure 39. **Histogram Matching Pipeline Example** showing four steps of the histogram matching pipeline. The example is from Eschikon, Steinmüri1 (sunflower) on 06.06.2024. (a) The mapping image used for matching. (b) The original snapshot (SNP) image. (c) The feature-matched mapping image. (d) The histogram-matched SNP image, demonstrating the process of adjusting pixel intensity distributions to match the mapping image.

Table 20. Mission Cross-Validation Field-Level Canopy Cover Model Performance with all metrics (F1 score, precision, Intersection over Union (IoU), and overall accuracy) listed for the mission dates shown. The F1 score, precision, recall, IoU, and accuracy metrics, which were already implemented in the code of Sadeghi-Tehran et al. (2017), are used here to evaluate the model's performance. Accuracy is computed as the fraction of correct predictions over all samples. Data © 2024 Planet Labs PBC.

Location	Site	Date	Flight #	F1 Plants	F1 Soil	Precision Plants	Precision Soil	IoU Plants	IoU Soil	Accuracy
Wallbach	Recycle	20240425	061	0.9402	0.9278	0.9533	0.9316	0.8876	0.8665	0.9412
Eschikon	Steinmueri1	20240430	066	0.8606	0.9993	0.9165	0.9990	0.7580	0.9986	0.9986
Eschikon	Laubisgruet2	20240514	071	0.8787	0.9960	0.8592	0.9954	0.7905	0.9920	0.9927
Eschikon	Steinmueri1	20240514	075	0.9385	0.9973	0.9280	0.9976	0.8851	0.9945	0.9948
Reckenholz	Eingang	20240519	086	0.7301	0.9891	0.8977	0.9819	0.5960	0.9796	0.9832
Haerkingen	Chaeppeli	20240605	102	0.9294	0.9932	0.9691	0.9891	0.8698	0.9866	0.9877
Haerkingen	Neufeld4	20240605	103	0.7385	0.9922	0.7756	0.9887	0.6507	0.9850	0.9851
Haerkingen	Altgraben1	20240605	108	0.9517	0.9958	0.9477	0.9963	0.9087	0.9916	0.9922
Eschikon	Laubisgruet2	20240606	111	0.8894	0.9911	0.8841	0.9915	0.8075	0.9825	0.9840
Eschikon	Steinmueri1	20240606	115	0.9321	0.9646	0.9649	0.9519	0.8734	0.9319	0.9549
Eschikon	Steinmueri3	20240606	116	0.9269	0.9929	0.9444	0.9915	0.8672	0.9862	0.9882
Reckenholz	Eingang	20240604	099	0.8868	0.9567	0.9211	0.9708	0.8044	0.9247	0.9588
Reckenholz	Eingang	20240612	122	0.8695	0.9620	0.9418	0.9819	0.7769	0.9404	0.9715
Eschikon	Steinmueri3	20240619	135	0.9405	0.9756	0.9417	0.9766	0.8879	0.9525	0.9664
Eschikon	Laubisgruet2	20240619	139	0.9087	0.9789	0.9239	0.9761	0.8381	0.9590	0.9682
Eschikon	Steinmueri1	20240619	132	0.9474	0.8979	0.9528	0.9304	0.9043	0.8247	0.9496
Haerkingen	Chaeppeli	20240618	128	0.9489	0.9856	0.9453	0.9874	0.9032	0.9716	0.9777
Haerkingen	Altgraben1	20240618	124	0.9627	0.9905	0.9662	0.9902	0.9283	0.9812	0.9851
Haerkingen	Neufeld4	20240618	130	0.9450	0.9882	0.9573	0.9868	0.8972	0.9767	0.9813
Eschikon	Laubisgruet2	20240628	151	0.9404	0.9753	0.9443	0.9781	0.8895	0.9521	0.9693
Eschikon	Steinmueri3	20240628	147	0.9395	0.9574	0.9449	0.9564	0.8862	0.9186	0.9510
Haerkingen	Neufeld4	20240708	153	0.9414	0.9728	0.9515	0.9694	0.8897	0.9474	0.9637
Eschikon	Laubisgruet2	20240709	163	0.9141	0.9561	0.9547	0.9472	0.8451	0.9168	0.9482
Reckenholz	Eingang	20240715	165	0.9366	0.8516	0.9434	0.8735	0.8820	0.7593	0.9288
Overall				**0.9054**	**0.9655**	**0.9303**	**0.9638**	**0.8347**	**0.9440**	**0.9645**

Satellite Data Preprocessing

Empirical Line Correction

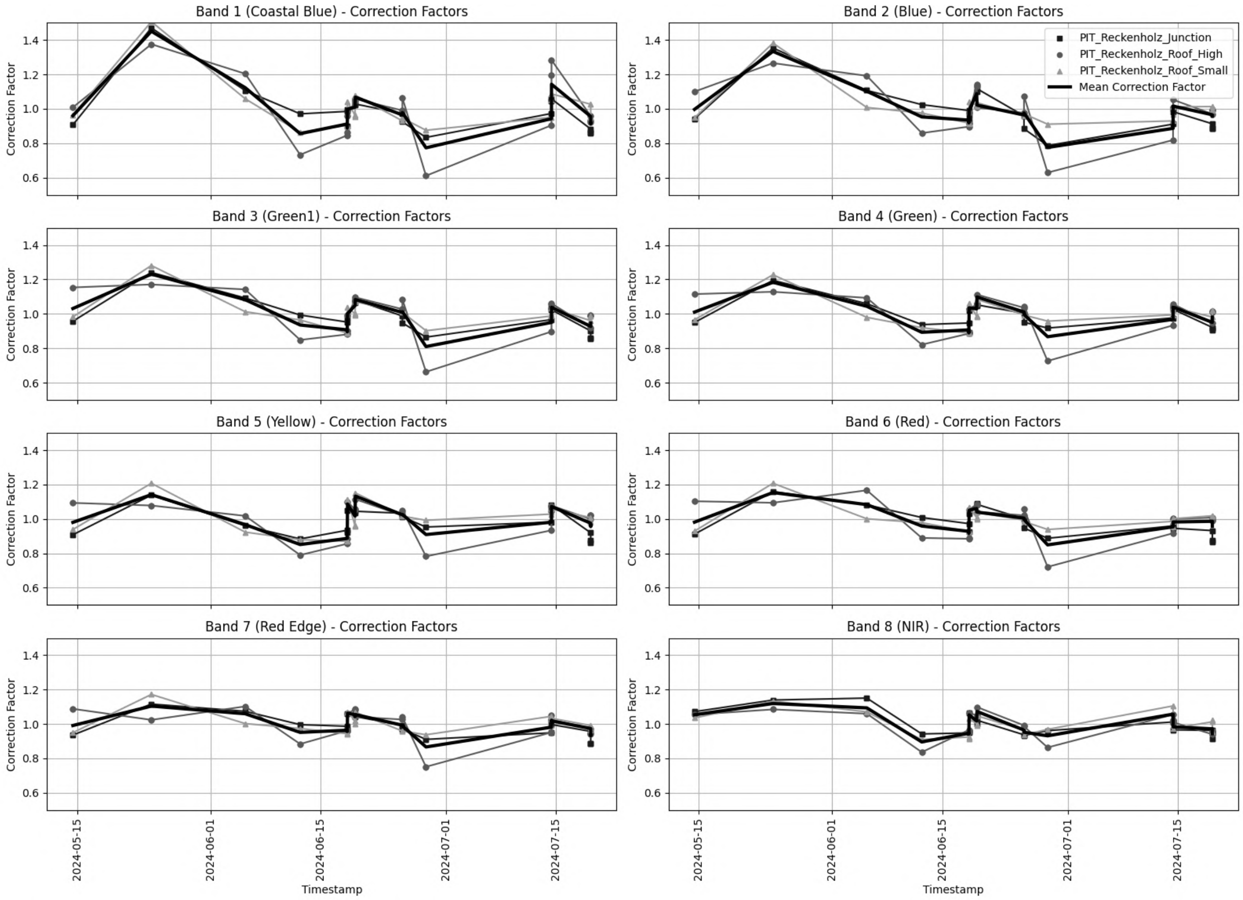


Figure 40. **Band-Specific Correction Factors for Reckenholz**. This figure shows the correction factors for each spectral band (Coastal Blue, Blue, Green1, Green, Yellow, Red, Red Edge, and NIR) from 2024. Data points represent correction factors for different Pseudo-Invariant Targets (PIT) at Reckenholz, with the mean correction factor shown in black. Data © 2024 Planet Labs PBC.

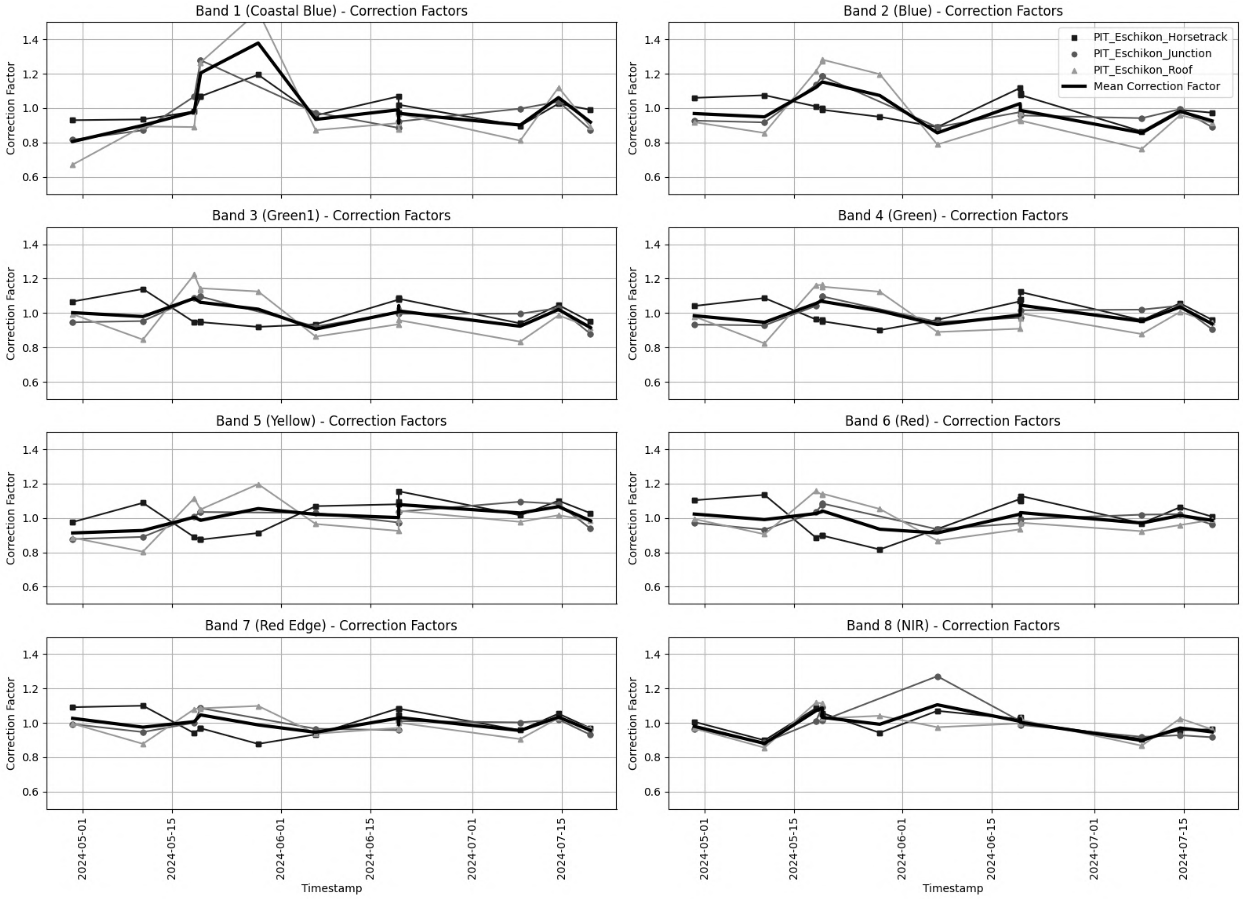


Figure 41. **Band-Specific Correction Factors for Eschikon**. This figure shows the correction factors for each spectral band (Coastal Blue, Blue, Green1, Green, Yellow, Red, Red Edge, and NIR) from 2024. Data points represent correction factors for different Pseudo-Invariant Targets (PIT) at Eschikon, with the mean correction factor shown in black. Data © 2024 Planet Labs PBC.

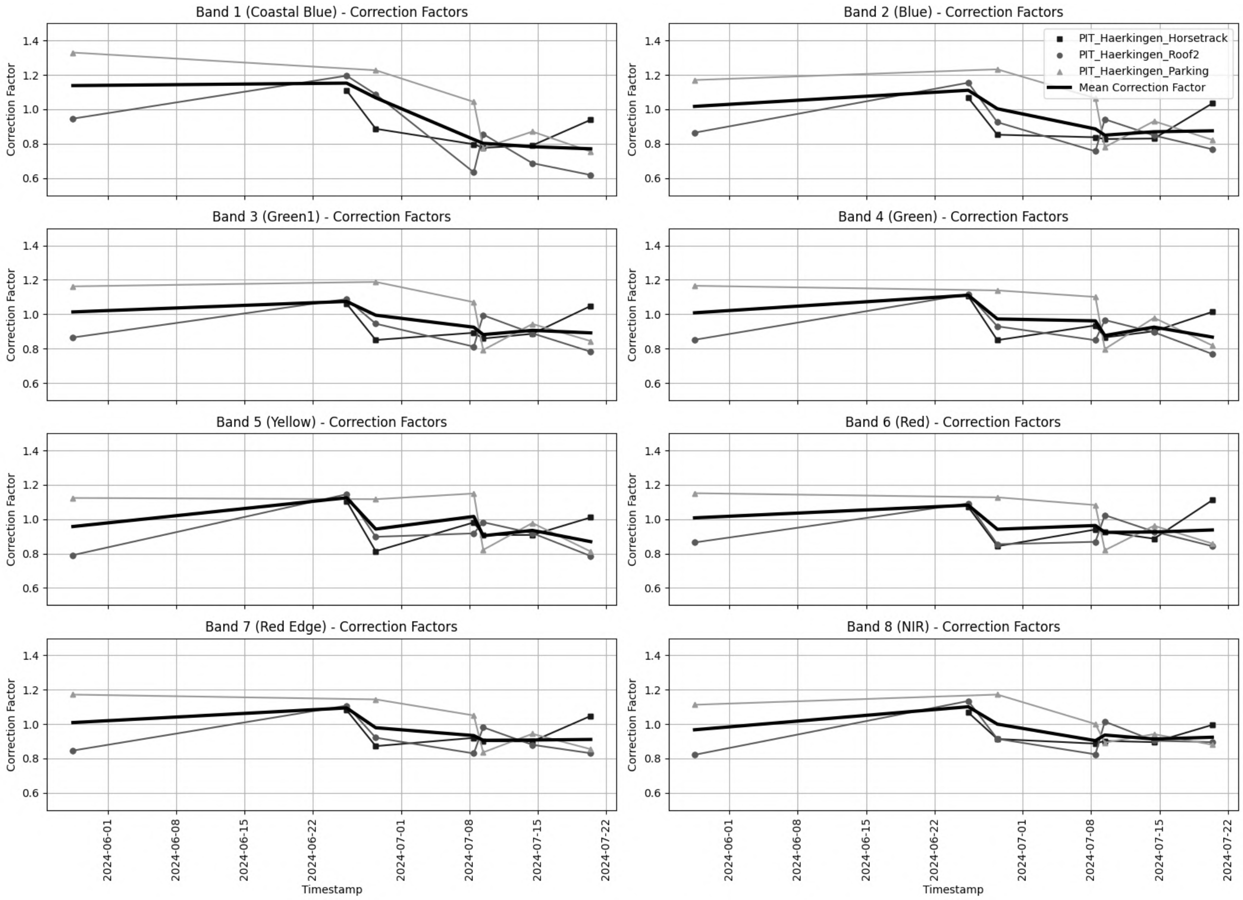


Figure 42. **Band-Specific Correction Factors for Haerkingen**. This figure shows the correction factors for each spectral band (Coastal Blue, Blue, Green1, Green, Yellow, Red, Red Edge, and NIR) from 2024. Data points represent correction factors for different Pseudo-Invariant Targets (PIT) at Haerkingen, with the mean correction factor shown in black. Data © 2024 Planet Labs PBC.

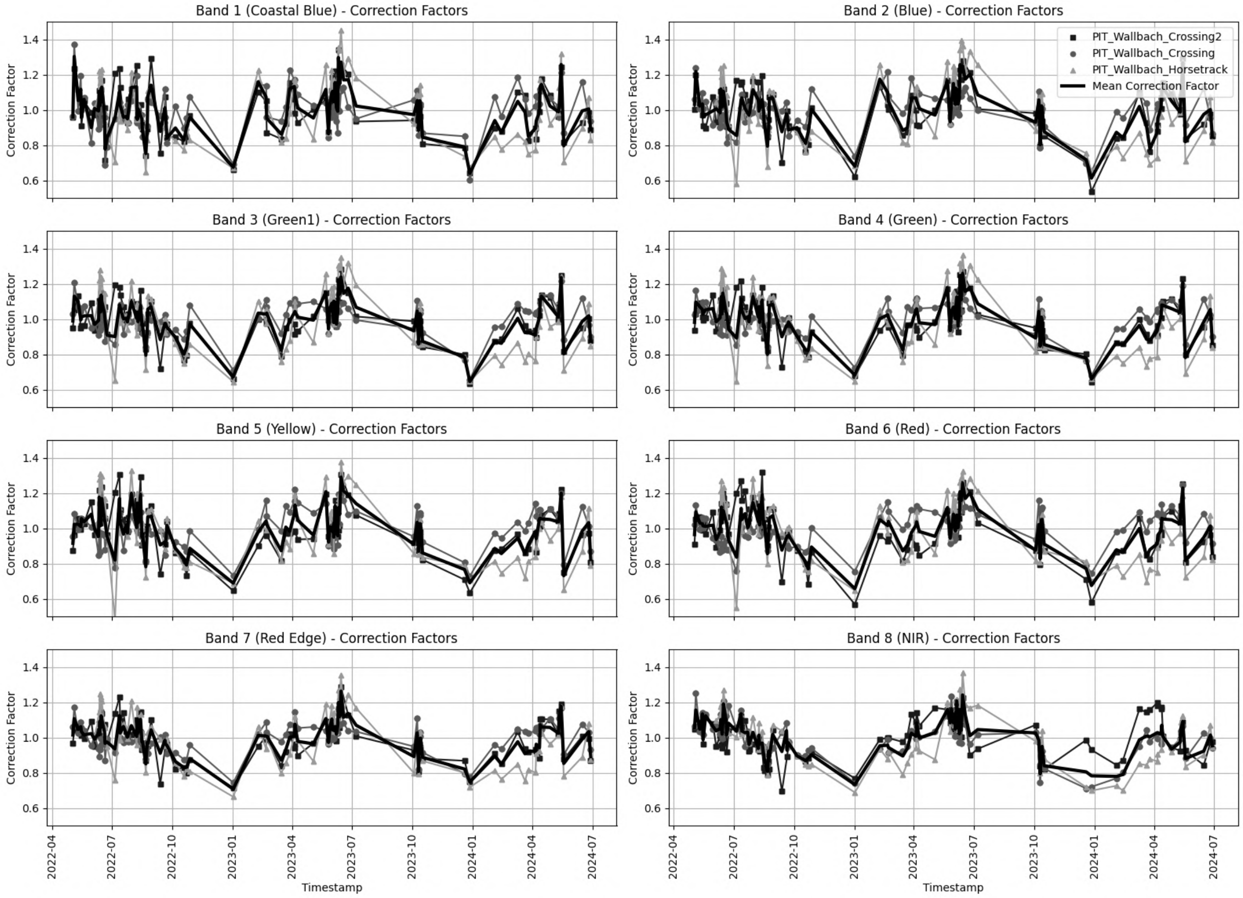


Figure 43. **Band-Specific Correction Factors for Wallbach**. This figure shows the correction factors for each spectral band (Coastal Blue, Blue, Green1, Green, Yellow, Red, Red Edge, and NIR) from 2022 to 2024. Data points represent correction factors for different Pseudo-Invariant Targets (PIT) at Wallbach, with the mean correction factor shown in black. Data © 2024 Planet Labs PBC.

MSAVI-NDVI-Relation in Training Dataframe

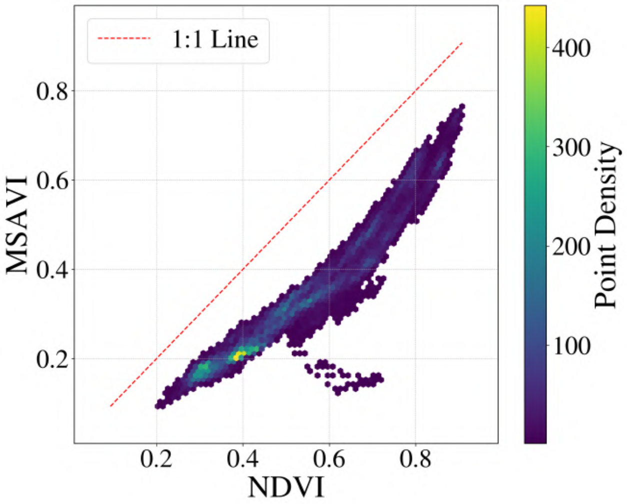


Figure 44. **Scatterplot of MSAVI Compared to NDVI Values in the Training Dataframe** The plot shows the relationship between the NDVI and MSAVI indices for multiple crop types (corn, sugarbeet, and sunflower) across various sites included in the training dataframe. The points are colored according to their density, with the 1:1 line represented in red. Data © 2024 Planet Labs PBC.

Leaf Area Index Estimation Using Radiative Transfer Modeling

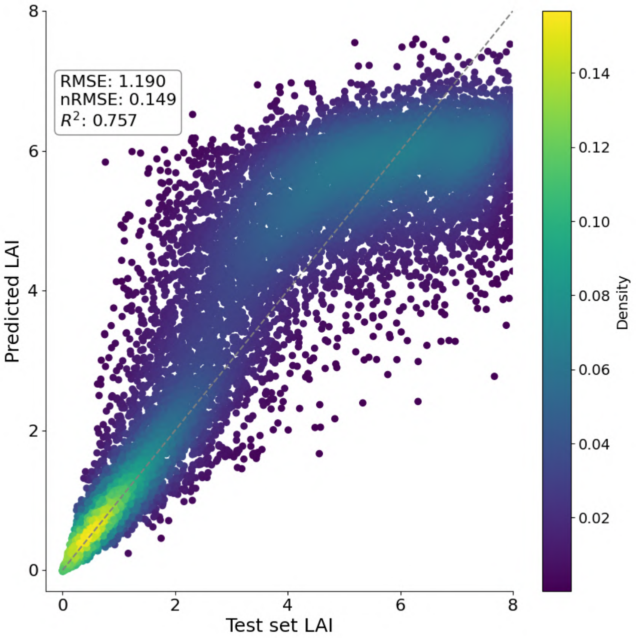
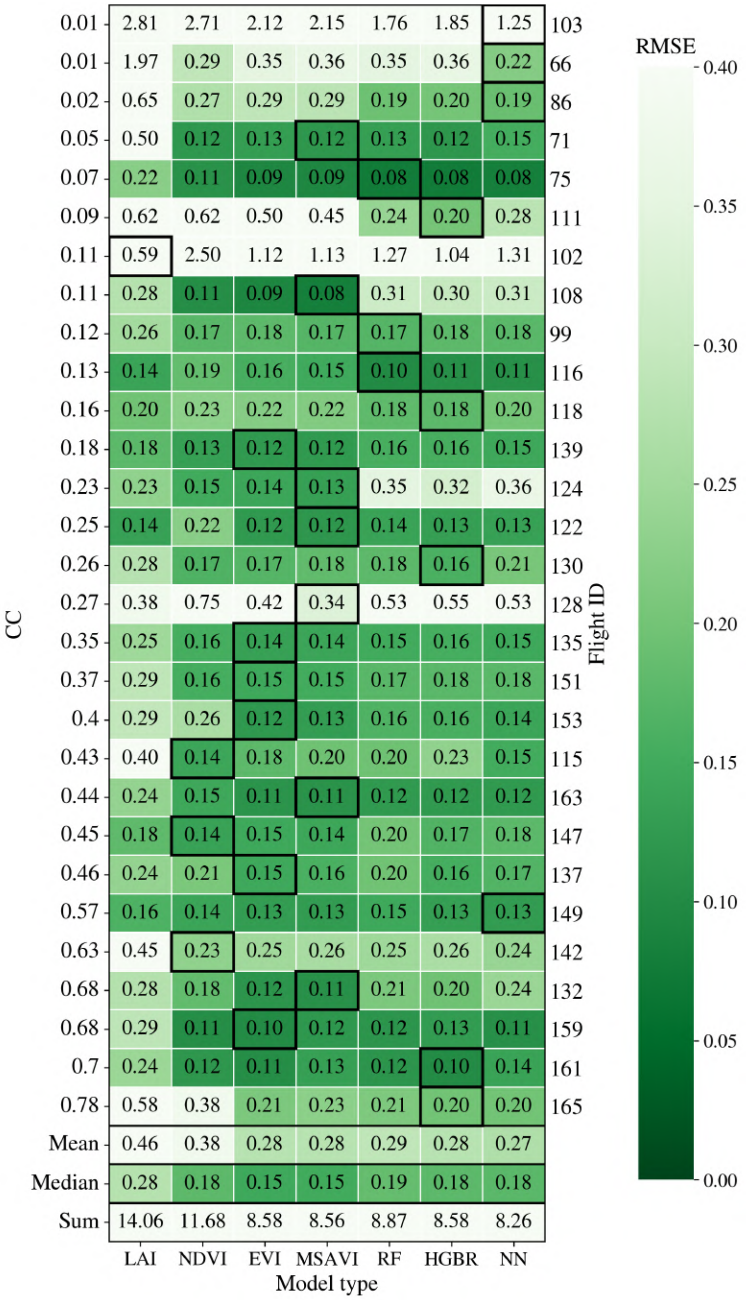


Figure 45. **LAI Model Predictions**. The figure shows the relative density of the model-based predictions of leaf area index (LAI) values, based on the soil-informed hybrid retrieval approach proposed by Ledain et al. (2025). Data © 2024 Planet Labs PBC.

Mission Cross Validation Satellite CC Model

Quantitative Comparison



CC	LAI	NDVI	EVI	MSAVI	RF	HGBR	NN	Flight ID
0.01	2.81	2.71	2.12	2.15	1.76	1.85	1.25	103
0.01	1.97	0.29	0.35	0.36	0.35	0.36	0.22	66
0.02	0.65	0.27	0.29	0.29	0.19	0.20	0.19	86
0.05	0.50	0.12	0.13	0.12	0.13	0.12	0.15	71
0.07	0.22	0.11	0.09	0.09	0.08	0.08	0.08	75
0.09	0.62	0.62	0.50	0.45	0.24	0.20	0.28	111
0.11	0.59	2.50	1.12	1.13	1.27	1.04	1.31	102
0.11	0.28	0.11	0.09	0.08	0.31	0.30	0.31	108
0.12	0.26	0.17	0.18	0.17	0.17	0.18	0.18	99
0.13	0.14	0.19	0.16	0.15	0.10	0.11	0.11	116
0.16	0.20	0.23	0.22	0.22	0.18	0.18	0.20	118
0.18	0.18	0.13	0.12	0.12	0.16	0.16	0.15	139
0.23	0.23	0.15	0.14	0.13	0.35	0.32	0.36	124
0.25	0.14	0.22	0.12	0.12	0.14	0.13	0.13	122
0.26	0.28	0.17	0.17	0.18	0.18	0.16	0.21	130
0.27	0.38	0.75	0.42	0.34	0.53	0.55	0.53	128
0.35	0.25	0.16	0.14	0.14	0.15	0.16	0.15	135
0.37	0.29	0.16	0.15	0.15	0.17	0.18	0.18	151
0.4	0.29	0.26	0.12	0.13	0.16	0.16	0.14	153
0.43	0.40	0.14	0.18	0.20	0.20	0.23	0.15	115
0.44	0.24	0.15	0.11	0.11	0.12	0.12	0.12	163
0.45	0.18	0.14	0.15	0.14	0.20	0.17	0.18	147
0.46	0.24	0.21	0.15	0.16	0.20	0.16	0.17	137
0.57	0.16	0.14	0.13	0.13	0.15	0.13	0.13	149
0.63	0.45	0.23	0.25	0.26	0.25	0.26	0.24	142
0.68	0.28	0.18	0.12	0.11	0.21	0.20	0.24	132
0.68	0.29	0.11	0.10	0.12	0.12	0.13	0.11	159
0.7	0.24	0.12	0.11	0.13	0.12	0.10	0.14	161
0.78	0.58	0.38	0.21	0.23	0.21	0.20	0.20	165
Mean	0.46	0.38	0.28	0.28	0.29	0.28	0.27	
Median	0.28	0.18	0.15	0.15	0.19	0.18	0.18	
Sum	14.06	11.68	8.58	8.56	8.87	8.58	8.26	

Figure 46. Mission Cross-Validation Model Range-Normalized Root Mean Square Error Comparison across different UAV flight IDs with respective canopy cover (CC) mean. The heatmap displays range-normalized root mean square error (NRMSE) values for each model type (LAI, NDVI, EVI, MSAVI, RF, HGBR, and NN). Color intensity represents RMSE magnitude, with darker shades indicating lower errors. Summary statistics (mean, median, and sum) for each model type are provided at the bottom to support the overall performance comparison. Data © 2024 Planet Labs PBC.

Table 21. **Mean and Median RMSE values for MSAVI, HGBR, and NN per crop** Data © 2024 Planet Labs PBC.

Crop	Statistic	MSAVI	HGBR	NN
Corn	Mean	0.098	0.103	0.107
	Median	0.113	0.101	0.109
Sugarbeet	Mean	0.106	0.102	0.103
	Median	0.101	0.100	0.101
Sunflower	Mean	0.099	0.120	0.109
	Median	0.089	0.123	0.090

Qualitative Comparison

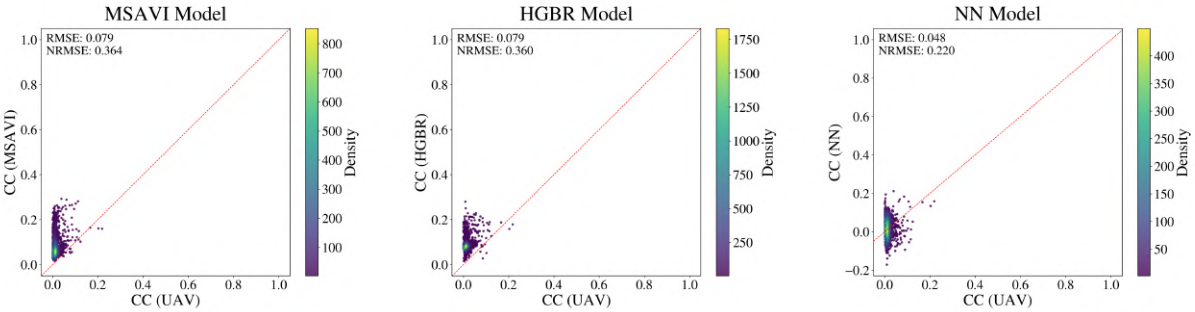


Figure 47. **Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Steinmueri1 Site (Sunflower) on April 30, 2024**. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.079, NRMSE = 0.364; HGBR: RMSE = 0.079, NRMSE = 0.360; NN: RMSE = 0.048, NRMSE = 0.220. Each model was trained on 22,639 samples and tested on 1,993 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

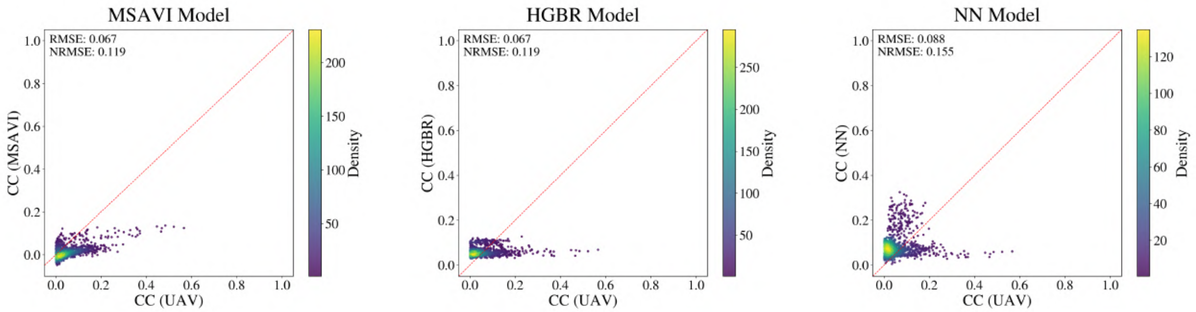


Figure 48. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Laubisgruet2 Site (Sugarbeet) on May 14, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.067, NRMSE = 0.119; HGBR: RMSE = 0.067, NRMSE = 0.119; NN: RMSE = 0.088, NRMSE = 0.155. Each model was trained on 23,205 samples and tested on 1,285 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

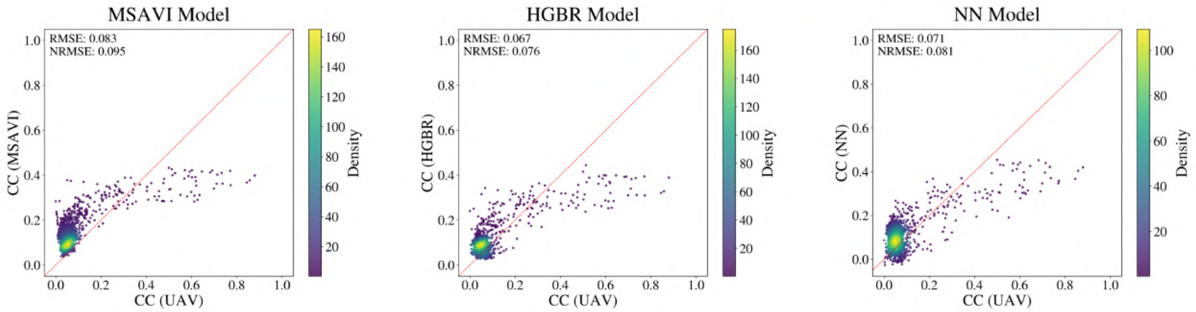


Figure 49. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Steinmueri1 Site (Sunflower) on May 14, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.083, NRMSE = 0.095; HGBR: RMSE = 0.067, NRMSE = 0.076; NN: RMSE = 0.071, NRMSE = 0.081. Each model was trained on 22,635 samples and tested on 1,998 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

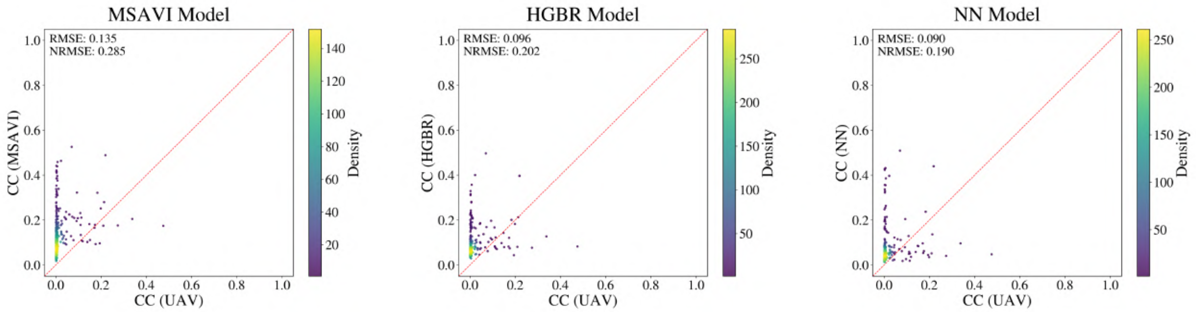


Figure 50. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Reckenholz, Eingang Site (Sugarbeet) on May 19, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.135, NRMSE = 0.285; HGBR: RMSE = 0.096, NRMSE = 0.202; NN: RMSE = 0.090, NRMSE = 0.190. Each model was trained on 23,892 samples and tested on 427 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

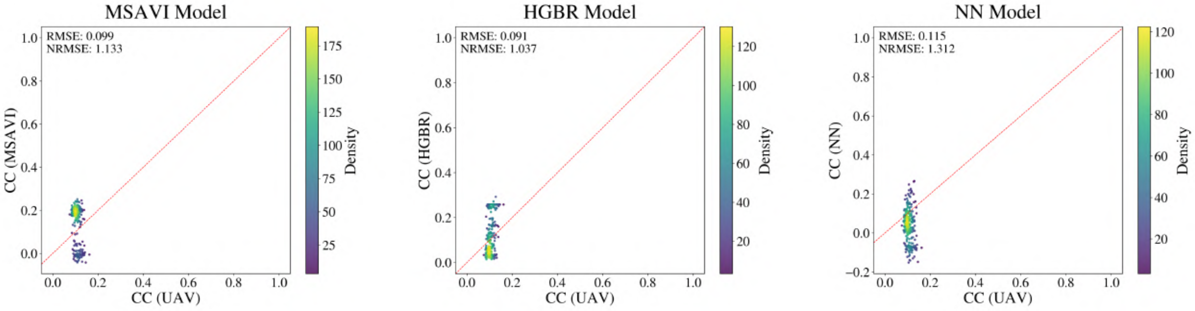


Figure 51. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Haerkingens, Chaeppeli Site (Corn) on June 5, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.099, NRMSE = 1.133; HGBR: RMSE = 0.091, NRMSE = 1.037; NN: RMSE = 0.115, NRMSE = 1.312. Each model was trained on 24,032 samples and tested on 251 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

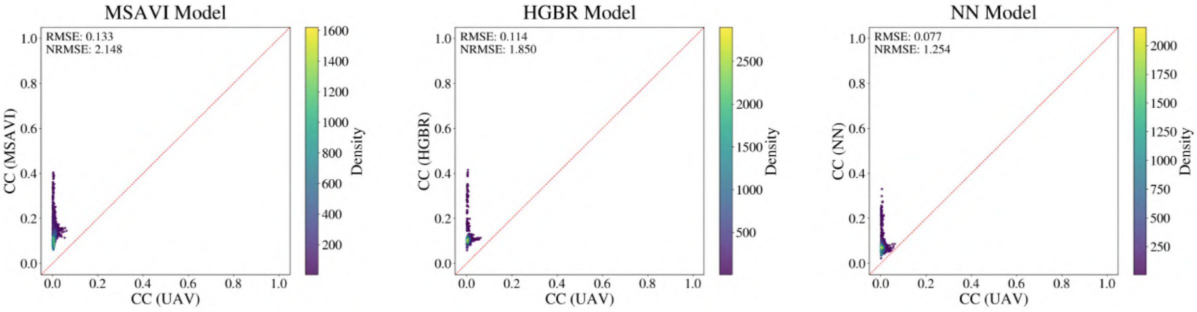


Figure 52. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Haerkingens, Neufeld4 Site (Corn) on June 5, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.133, NRMSE = 2.148; HGBR: RMSE = 0.114, NRMSE = 1.850; NN: RMSE = 0.077, NRMSE = 1.254. Each model was trained on 23,022 samples and tested on 1,514 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

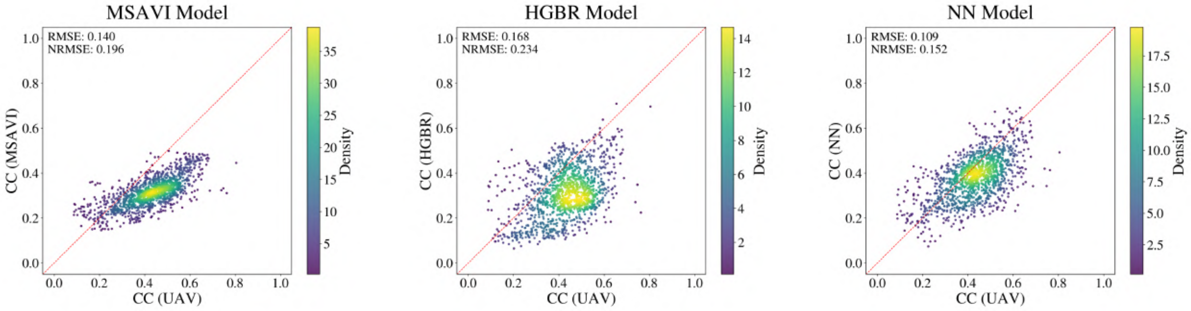


Figure 53. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Steinmueri1 Site (Sunflower) on June 6, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.140, NRMSE = 0.196; HGBR: RMSE = 0.168, NRMSE = 0.234; NN: RMSE = 0.109, NRMSE = 0.152. Each model was trained on 23,168 samples and tested on 1,331 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

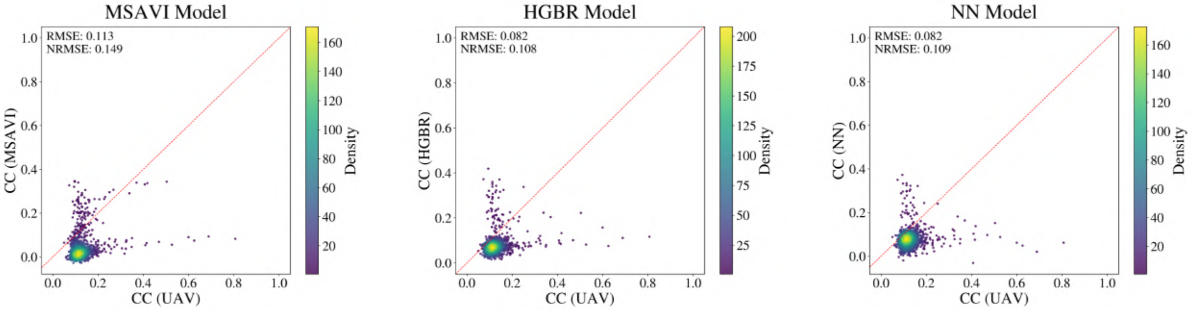


Figure 54. **Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Steinmueri3 Site (Corn) on June 6, 2024**. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.113, NRMSE = 0.149; HGBR: RMSE = 0.082, NRMSE = 0.108; NN: RMSE = 0.082, NRMSE = 0.109. Each model was trained on 23,244 samples and tested on 1,236 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

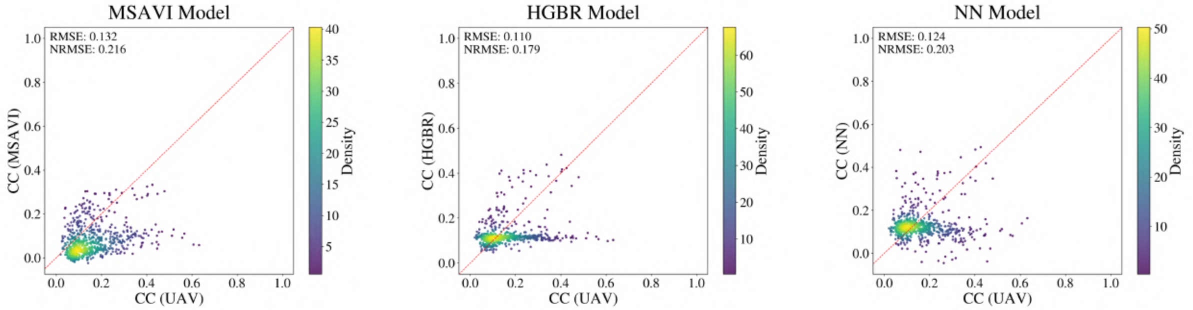


Figure 55. **Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Holgenbuehl Site (Corn) on June 6, 2024**. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.132, NRMSE = 0.216; HGBR: RMSE = 0.110, NRMSE = 0.179; NN: RMSE = 0.124, NRMSE = 0.203. Each model was trained on 23,708 samples and tested on 657 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

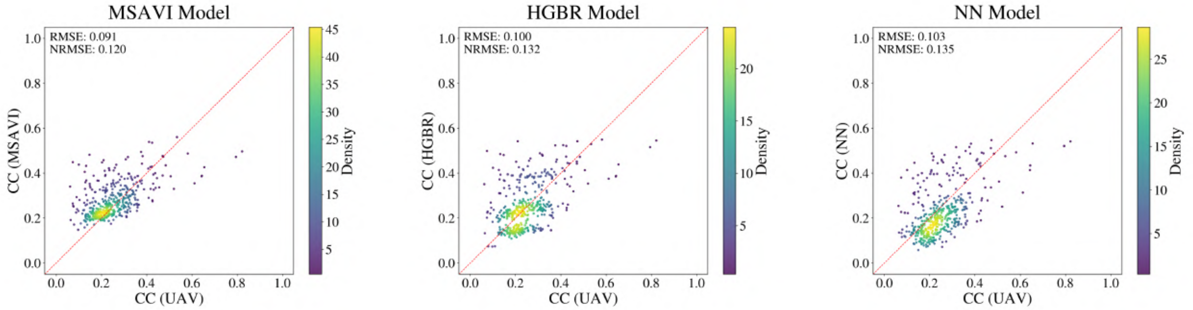


Figure 56. **Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Reckenholz, Eingang Site (Sugarbeet) on June 12, 2024**. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.091, NRMSE = 0.120; HGBR: RMSE = 0.100, NRMSE = 0.132; NN: RMSE = 0.103, NRMSE = 0.135. Each model was trained on 23,892 samples and tested on 426 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

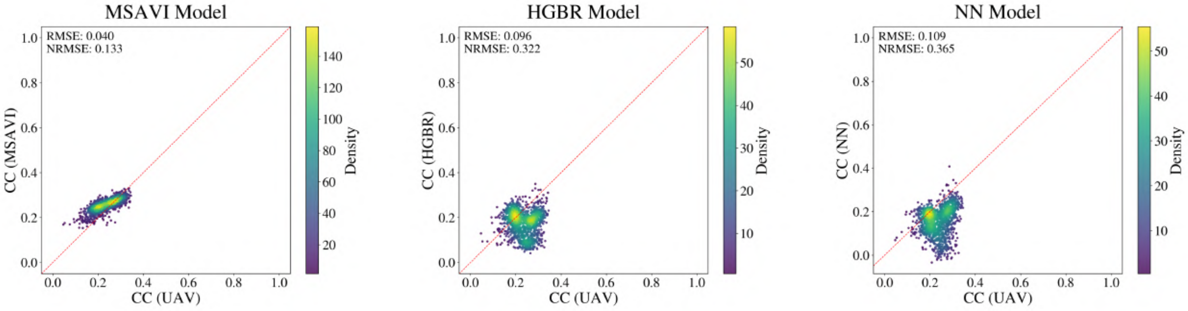


Figure 57. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Haerkingens, Altraben1 Site (Corn) on June 18, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.040, NRMSE = 0.133; HGBR: RMSE = 0.096, NRMSE = 0.322; NN: RMSE = 0.109, NRMSE = 0.365. Each model was trained on 23,155 samples and tested on 1,348 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

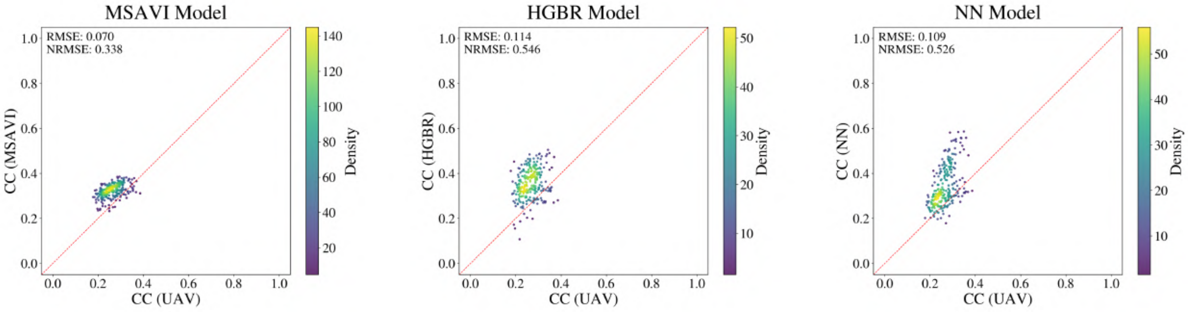


Figure 58. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Haerkingens, Chaeppeli Site (Corn) on June 18, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.070, NRMSE = 0.338; HGBR: RMSE = 0.114, NRMSE = 0.546; NN: RMSE = 0.109, NRMSE = 0.526. Each model was trained on 24,028 samples and tested on 257 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

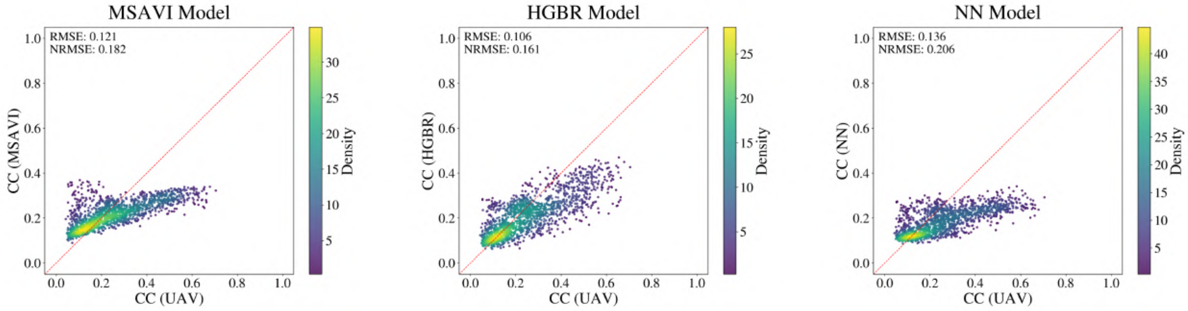


Figure 59. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Haerkingens, Neufeld4 Site (Corn) on June 18, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.121, NRMSE = 0.182; HGBR: RMSE = 0.106, NRMSE = 0.161; NN: RMSE = 0.136, NRMSE = 0.206. Each model was trained on 22,968 samples and tested on 1,582 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

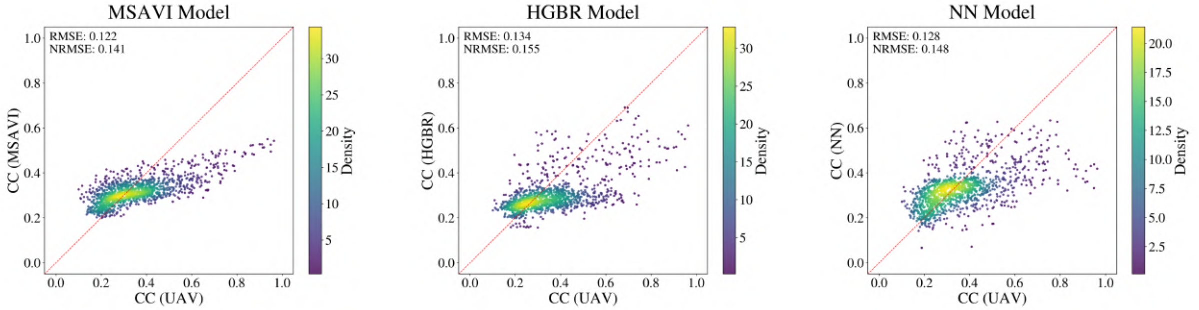


Figure 60. **Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Steinmueri3 Site (Corn) on June 19, 2024**. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.122, NRMSE = 0.141; HGBR: RMSE = 0.134, NRMSE = 0.155; NN: RMSE = 0.128, NRMSE = 0.148. Each model was trained on 23,264 samples and tested on 1,211 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

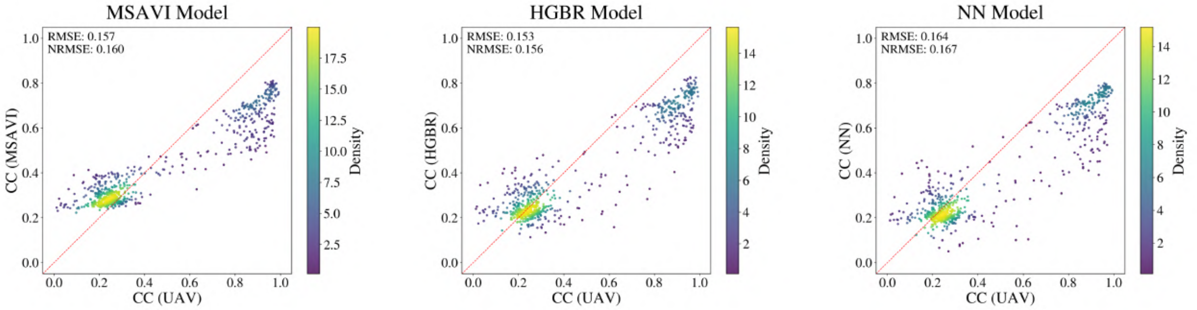


Figure 61. **Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Steinmueri3 Site (Corn) on June 19, 2024**. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.157, NRMSE = 0.160; HGBR: RMSE = 0.153, NRMSE = 0.156; NN: RMSE = 0.164, NRMSE = 0.167. Each model was trained on 23,264 samples and tested on 1,211 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

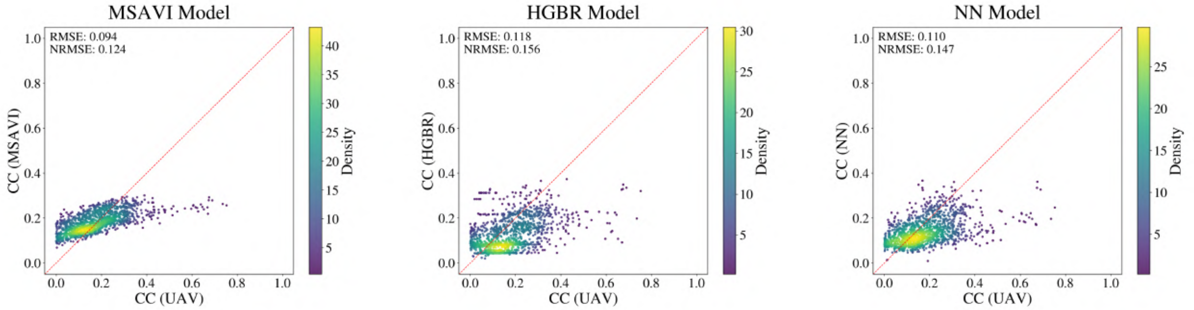


Figure 62. **Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Holgenbuehl Site (Corn) on June 19, 2024**. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.094, NRMSE = 0.124; HGBR: RMSE = 0.118, NRMSE = 0.156; NN: RMSE = 0.110, NRMSE = 0.147. Each model was trained on 23,706 samples and tested on 659 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

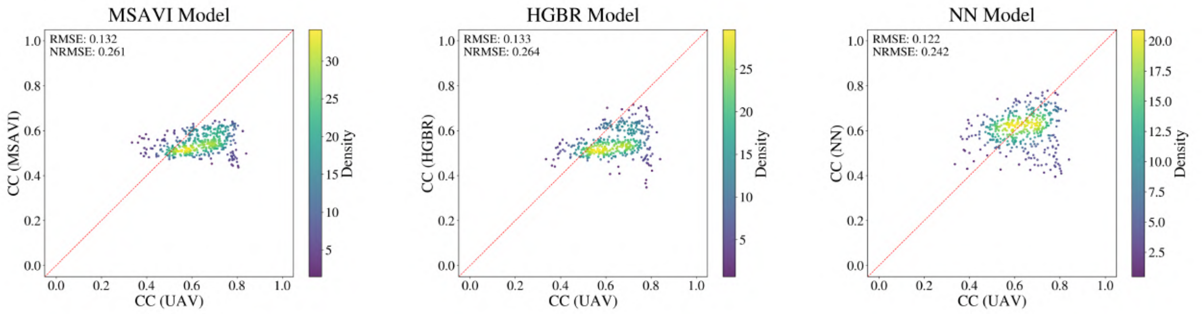


Figure 63. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Laubisgruet2 Site (Sugarbeet) on June 19, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.132, NRMSE = 0.261; HGBR: RMSE = 0.133, NRMSE = 0.264; NN: RMSE = 0.112, NRMSE = 0.242. Each model was trained on 23,220 samples and tested on 1,267 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

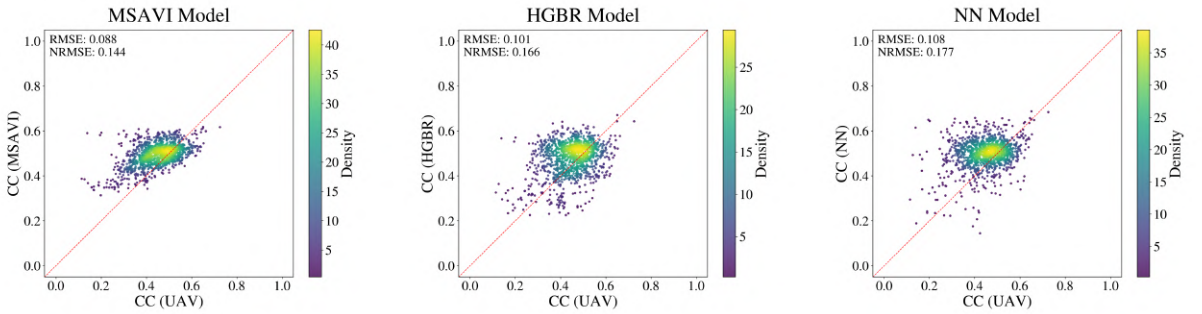


Figure 64. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Steinmueri3 Site (Corn) on June 28, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.088, NRMSE = 0.144; HGBR: RMSE = 0.101, NRMSE = 0.166; NN: RMSE = 0.108, NRMSE = 0.177. Each model was trained on 23,260 samples and tested on 1,216 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

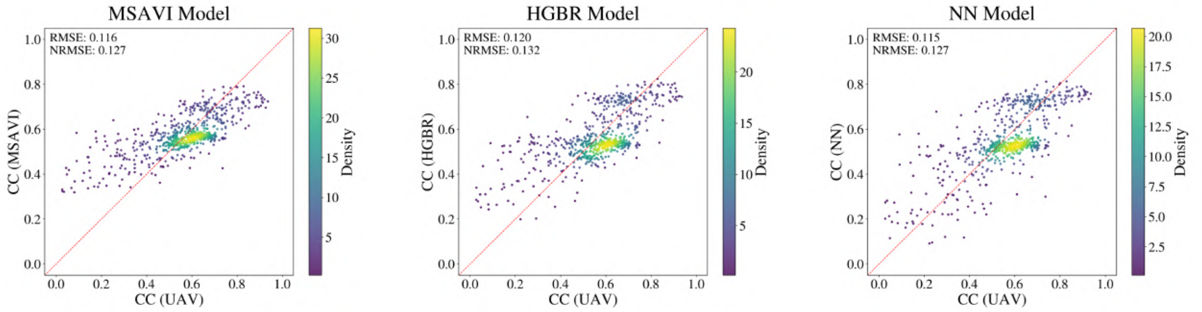


Figure 65. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Holgenbuehl Site (Corn) on June 28, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.116, NRMSE = 0.127; HGBR: RMSE = 0.120, NRMSE = 0.132; NN: RMSE = 0.115, NRMSE = 0.127. Each model was trained on 23,696 samples, validated on 5,924 samples, and tested on 672 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

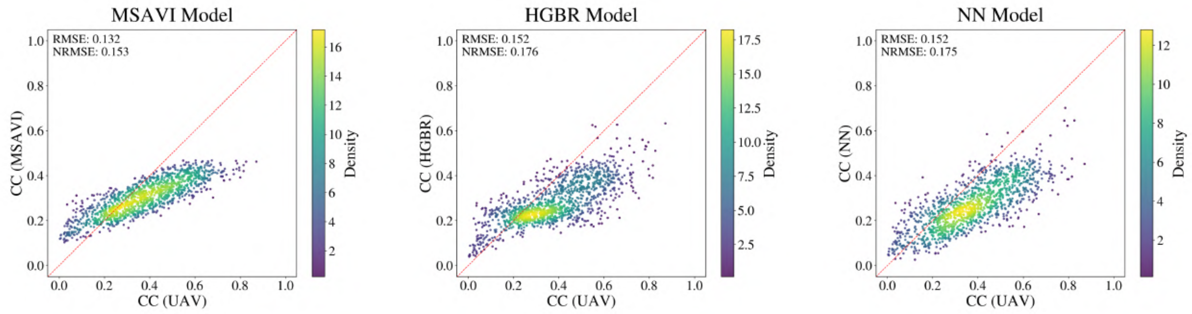


Figure 66. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Laubisgruet2 Site (Sugarbeet) on June 28, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.132, NRMSE = 0.153; HGBR: RMSE = 0.152, NRMSE = 0.176; NN: RMSE = 0.152, NRMSE = 0.175. Each model was trained on 23,219 samples and tested on 1,268 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

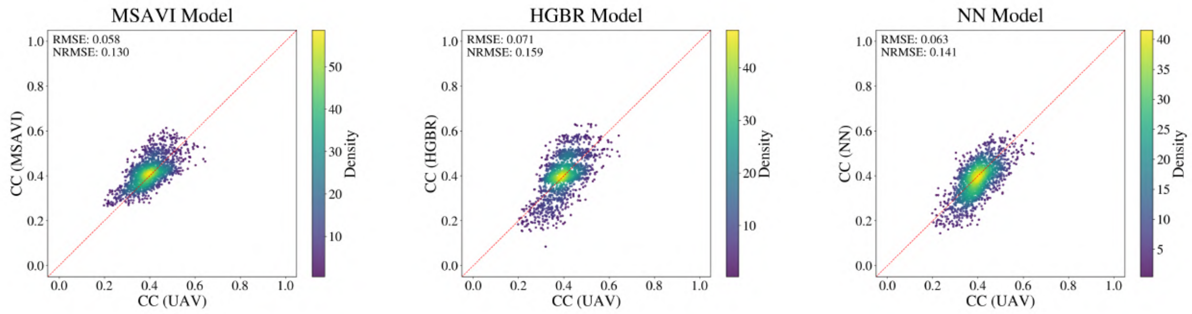


Figure 67. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Haerkinger, Neufeld4 Site (Corn) on July 8, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.058, NRMSE = 0.130; HGBR: RMSE = 0.071, NRMSE = 0.159; NN: RMSE = 0.063, NRMSE = 0.141. Each model was trained on 22,965 samples and tested on 1,585 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

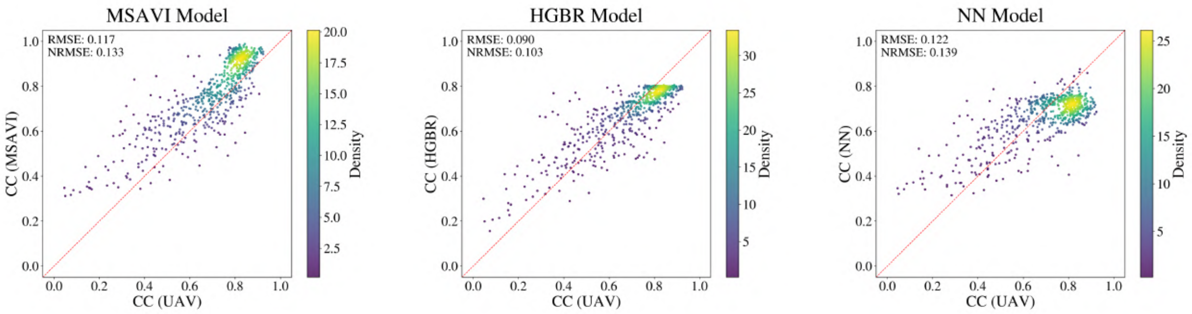


Figure 68. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Eschikon, Holgenbuehl Site (Corn) on July 9, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.117, NRMSE = 0.133; HGBR: RMSE = 0.090, NRMSE = 0.103; NN: RMSE = 0.122, NRMSE = 0.139. Each model was trained on 23,692 samples, validated on 5,923 samples, and tested on 677 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

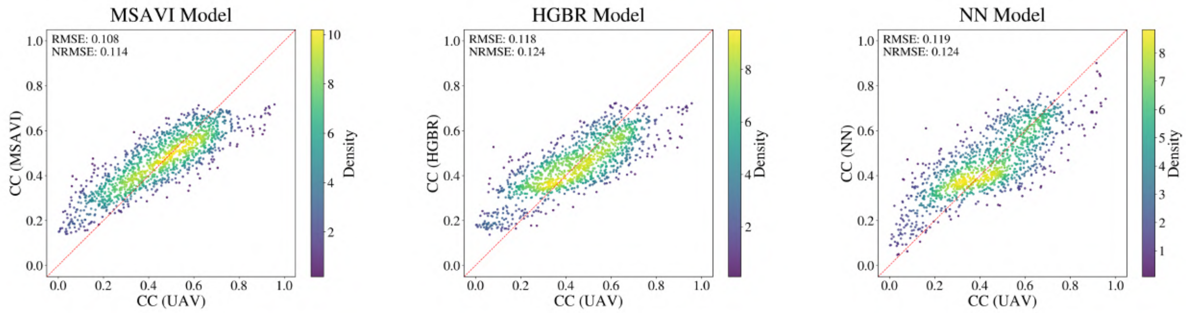


Figure 69. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Reckenholz, Eingang Site (Sugarbeet) on July 15, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.103, NRMSE = 0.227; HGBR: RMSE = 0.091, NRMSE = 0.199; NN: RMSE = 0.091, NRMSE = 0.199. Each model was trained on 23,894 samples and tested on 424 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

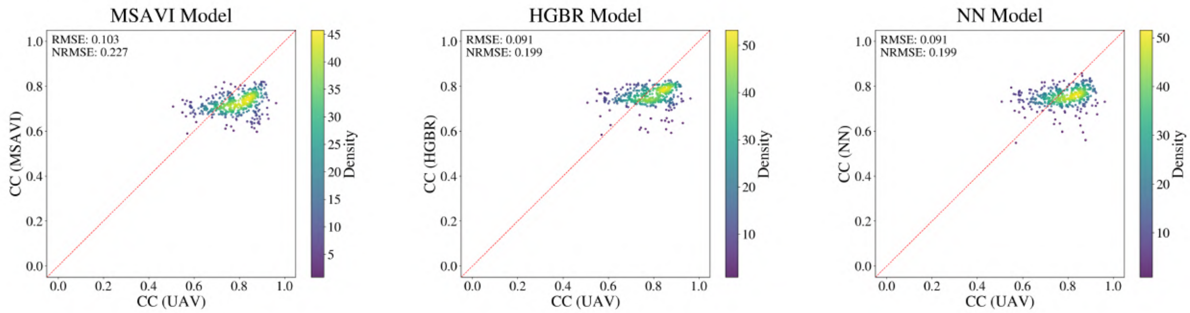


Figure 70. Mission Cross-Validation Model Comparison for Canopy Cover Prediction at the Reckenholz, Eingang Site (Sugarbeet) on July 15, 2024. The scatterplots display predicted vs. observed canopy cover (CC) values for the MSAVI (left), HGBR (center), and NN (right) models. Point color indicates data density, with higher densities shown in yellow and lower densities in purple. The red dashed line represents the 1:1 reference line. Model performance metrics for this flight are as follows: MSAVI: RMSE = 0.103, NRMSE = 0.227; HGBR: RMSE = 0.091, NRMSE = 0.199; NN: RMSE = 0.091, NRMSE = 0.199. Each model was trained on 23,220 samples and tested on 1,266 samples as part of a mission-based cross-validation setup. Data © 2024 Planet Labs PBC.

Spatial Comparison

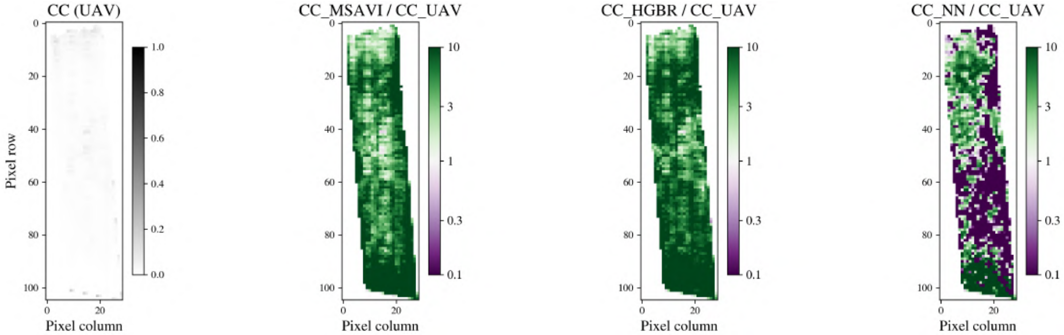


Figure 71. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Steinmueri1 Site (Sunflower) on April 30, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio > 1), whereas purple tones indicate underestimation (ratio < 1). Data © 2024 Planet Labs PBC.

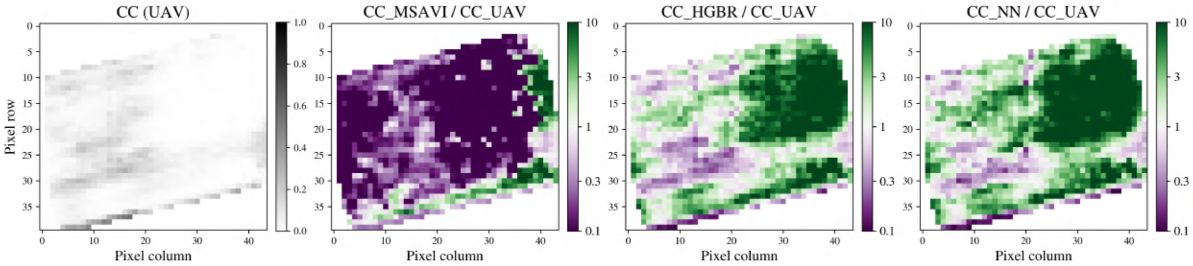


Figure 72. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Laubisgruet2 Site (Sugarbeet) on May 14, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio > 1), whereas purple tones indicate underestimation (ratio < 1). Data © 2024 Planet Labs PBC.

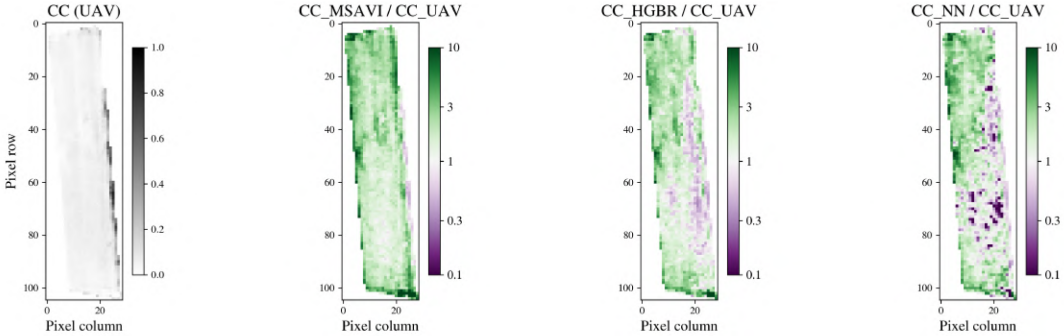


Figure 73. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Steinmueri1 Site (Sunflower) on May 14, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio > 1), whereas purple tones indicate underestimation (ratio < 1). Data © 2024 Planet Labs PBC.

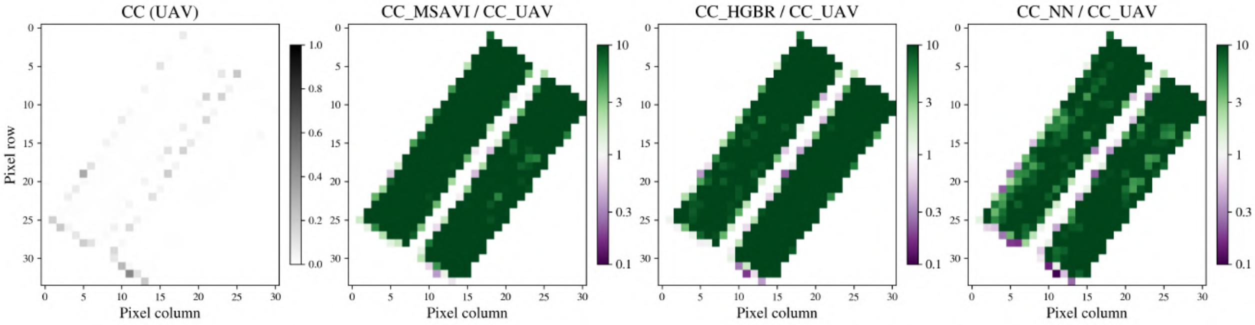


Figure 74. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Reckenholz, Eingang Site (Sugarbeet) on May 19, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio > 1), whereas purple tones indicate underestimation (ratio < 1). Data © 2024 Planet Labs PBC.

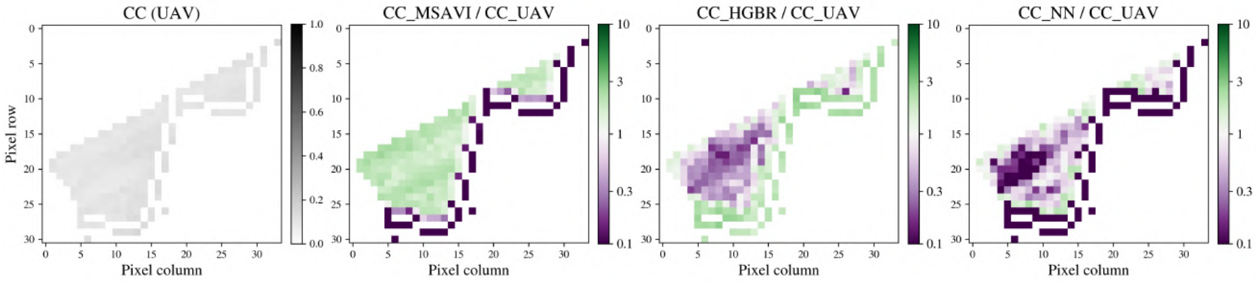


Figure 75. **Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Haerkingens, Chaeppeli Site (Corn) on June 5, 2024**. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

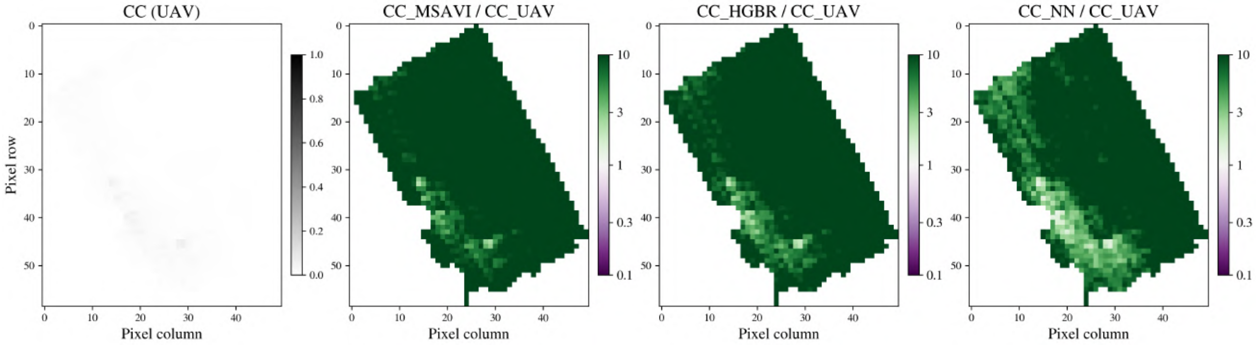


Figure 76. **Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Haerkingens, Neufeld4 Site (Corn) on June 5, 2024**. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

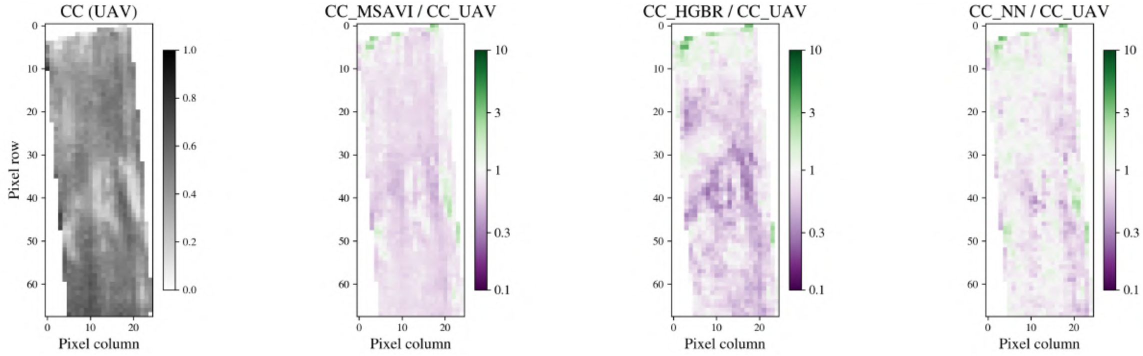


Figure 77. **Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Steinmueri1 Site (Sunflower) on June 6, 2024**. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

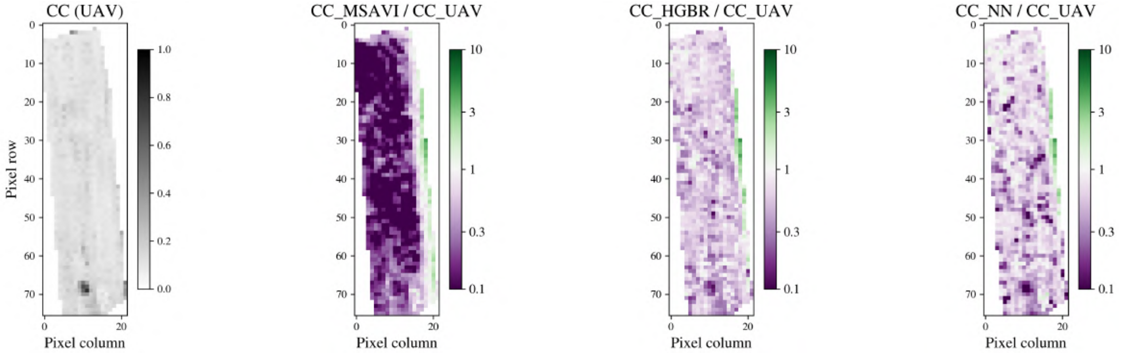


Figure 78. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Steinmueri3 Site (Corn) on June 6, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio > 1), whereas purple tones indicate underestimation (ratio < 1). Data © 2024 Planet Labs PBC.

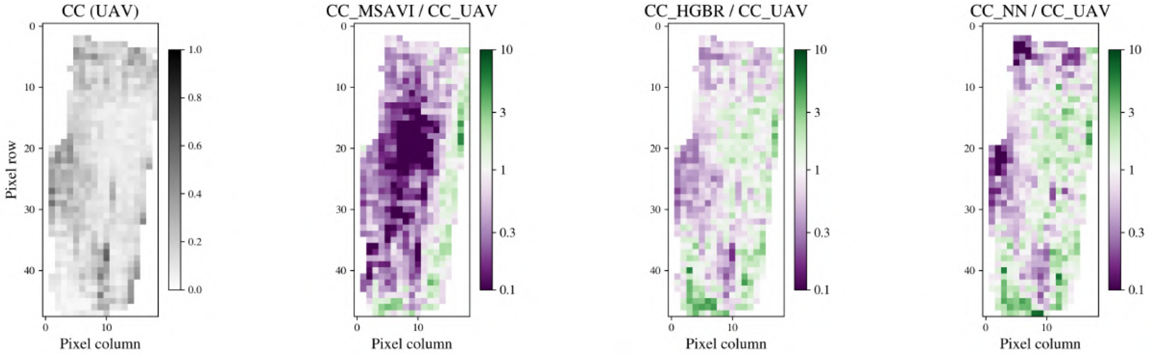


Figure 79. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Holgenbuehl Site (Corn) on June 6, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio > 1), whereas purple tones indicate underestimation (ratio < 1). Data © 2024 Planet Labs PBC.

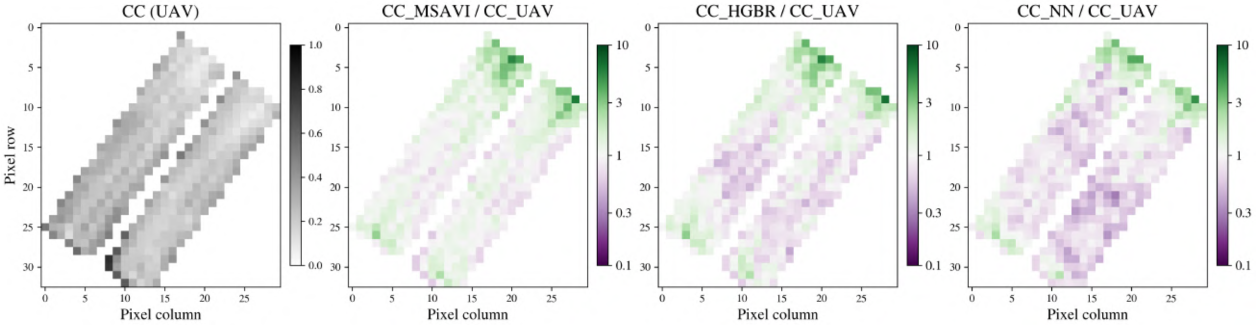


Figure 80. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Reckenholz, Eingang Site (Sugarbeet) on June 12, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio > 1), whereas purple tones indicate underestimation (ratio < 1). Data © 2024 Planet Labs PBC.

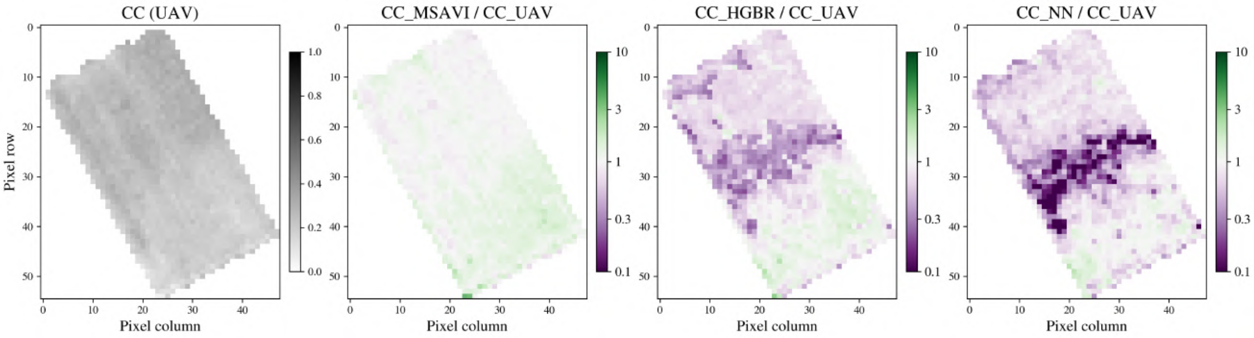


Figure 81. **Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Haerkingens, Altgraben1 Site (Corn) on June 18, 2024**. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

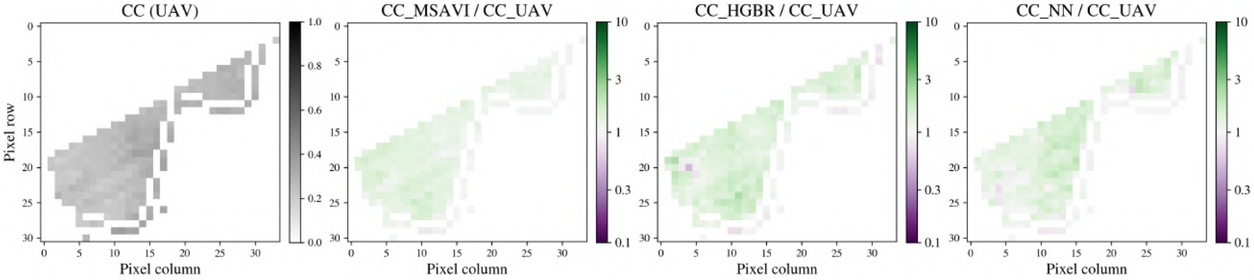


Figure 82. **Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Haerkingen, Chaeppeli Site (Corn) on June 18, 2024**. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

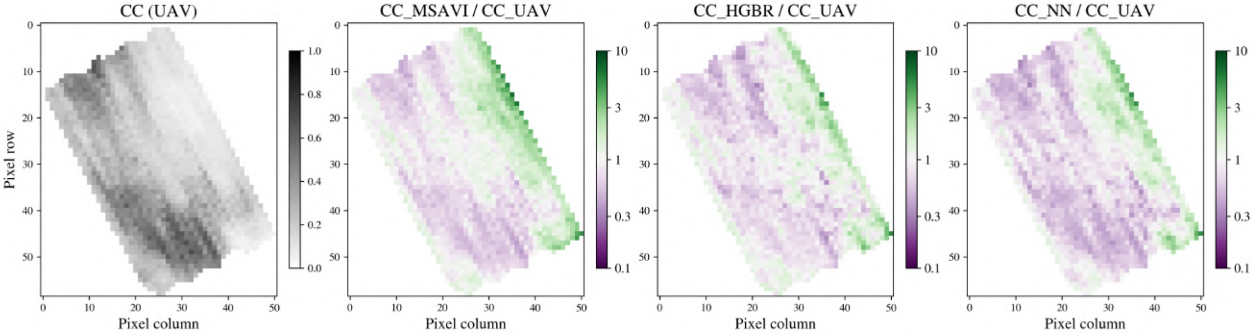


Figure 83. **Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Haerkingen, Neufeld4 Site (Corn) on June 18, 2024**. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

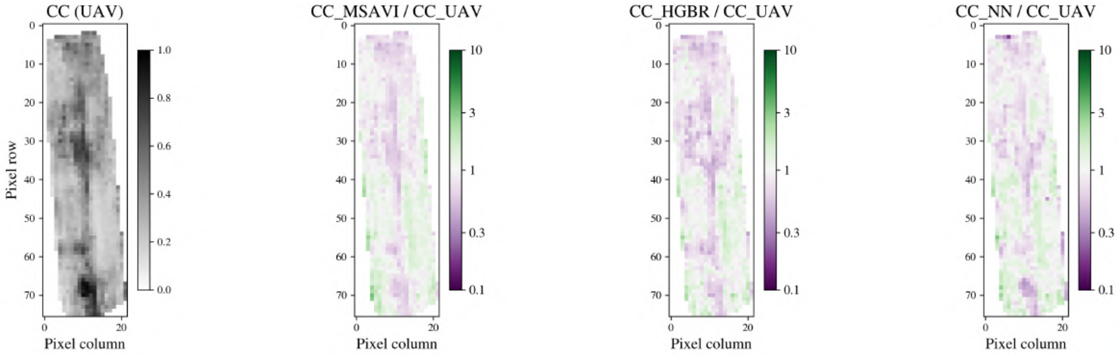


Figure 84. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Steinmueri3 Site (Corn) on June 19, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

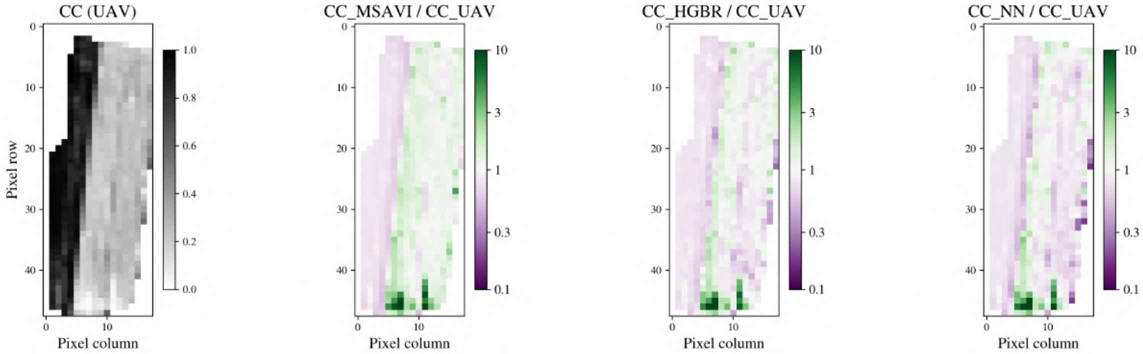


Figure 85. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Holgenbuehl Site (Corn) on June 19, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

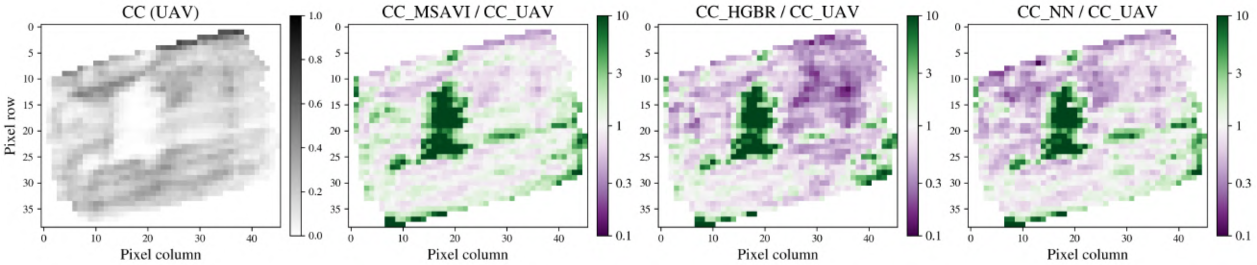


Figure 86. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Laubisgruet2 Site (Sugarbeet) on June 19, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

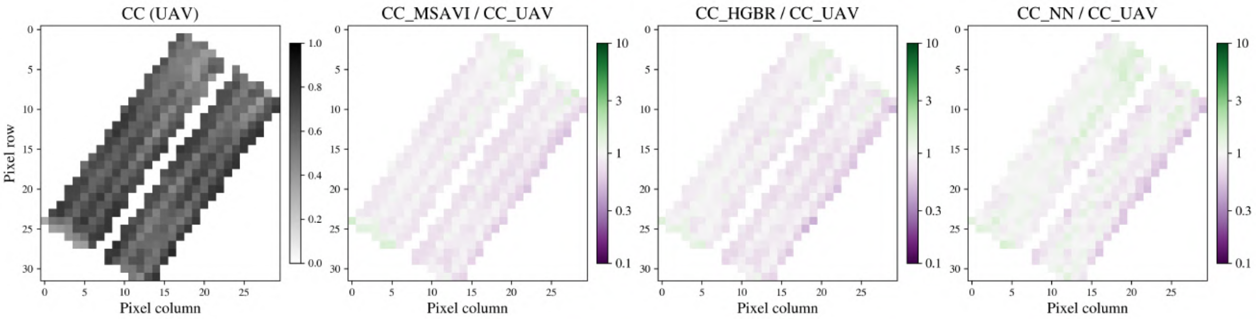


Figure 87. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Reckenholz, Eingang Site (Sugarbeet) on June 27, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

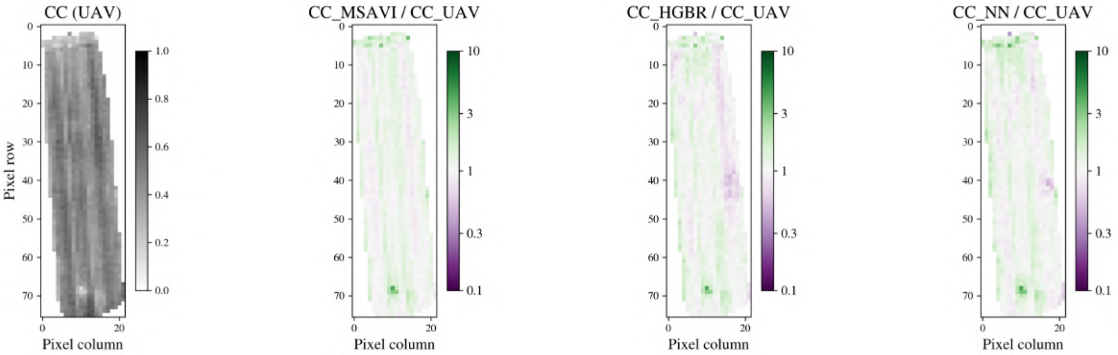


Figure 88. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Haerkingen, Neufeld4 Site (Corn) on July 8, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

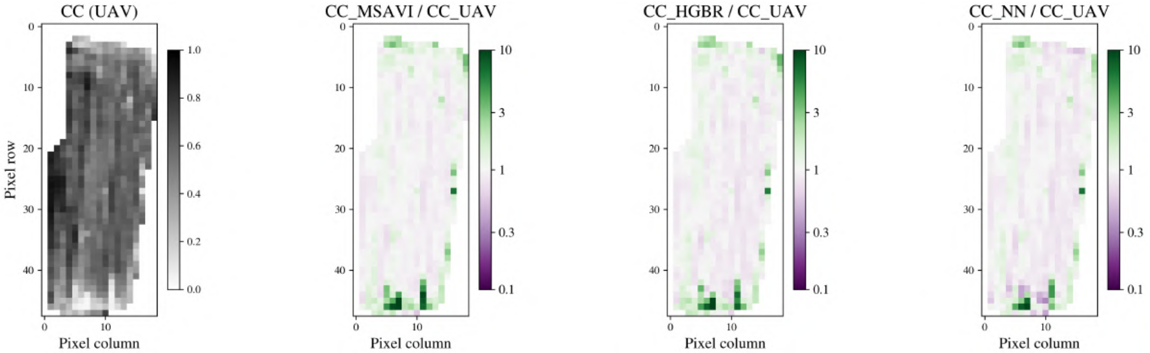


Figure 89. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Holgenbuehl Site (Corn) on June 28, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

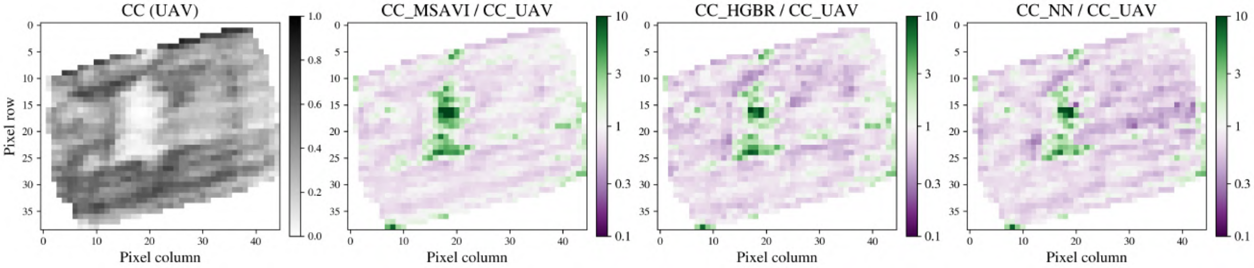


Figure 90. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Laubisgruet2 Site (Sugarbeet) on June 28, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

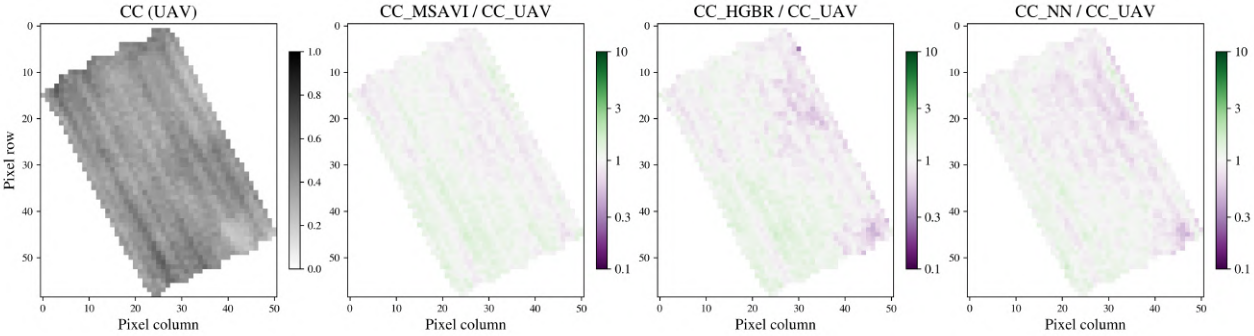


Figure 91. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Haerkingen, Neufeld4 Site (Corn) on July 8, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

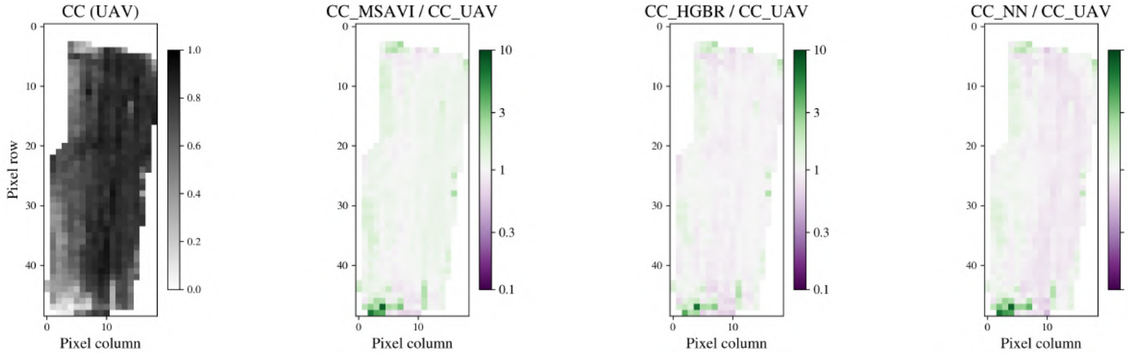


Figure 92. Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Holgenbuehl Site (Corn) on July 9, 2024. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

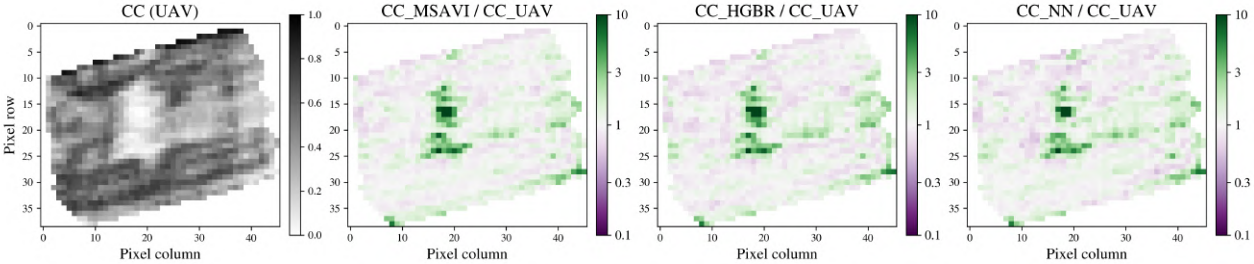


Figure 93. **Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Eschikon, Laubisgruet2 Site (Sugarbeet) on July 9, 2024**. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2024 Planet Labs PBC.

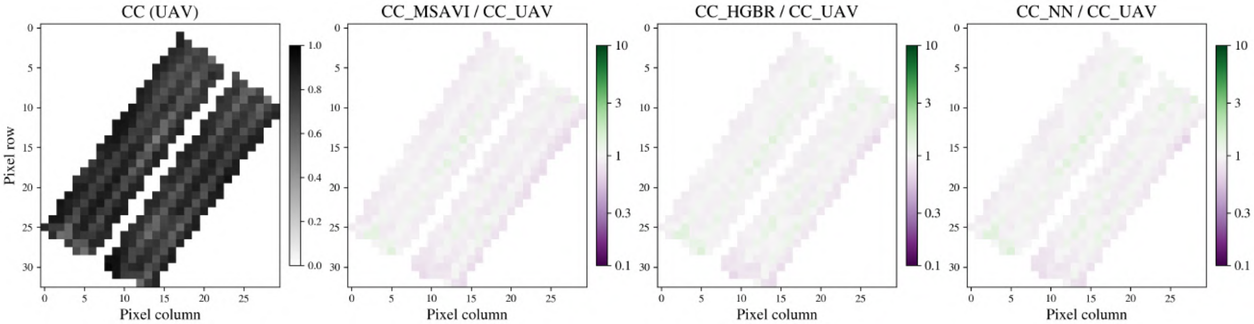


Figure 94. **Spatial Deviation of Canopy Cover Predictions from UAV-Reference Values at the Reckenholz, Eingang Site (Sugarbeet) on July 15, 2024**. The first panel shows the reference canopy cover (CC) map based on UAV data. The following three panels display the pixel-wise ratio of predicted CC to UAV-derived CC for the MSAVI, HGBR, and NN models, respectively. Values close to 1 indicate strong agreement, while deviations reflect model under- or overestimation. Green tones indicate overestimation (ratio $>$ 1), whereas purple tones indicate underestimation (ratio $<$ 1). Data © 2023 Planet Labs PBC.

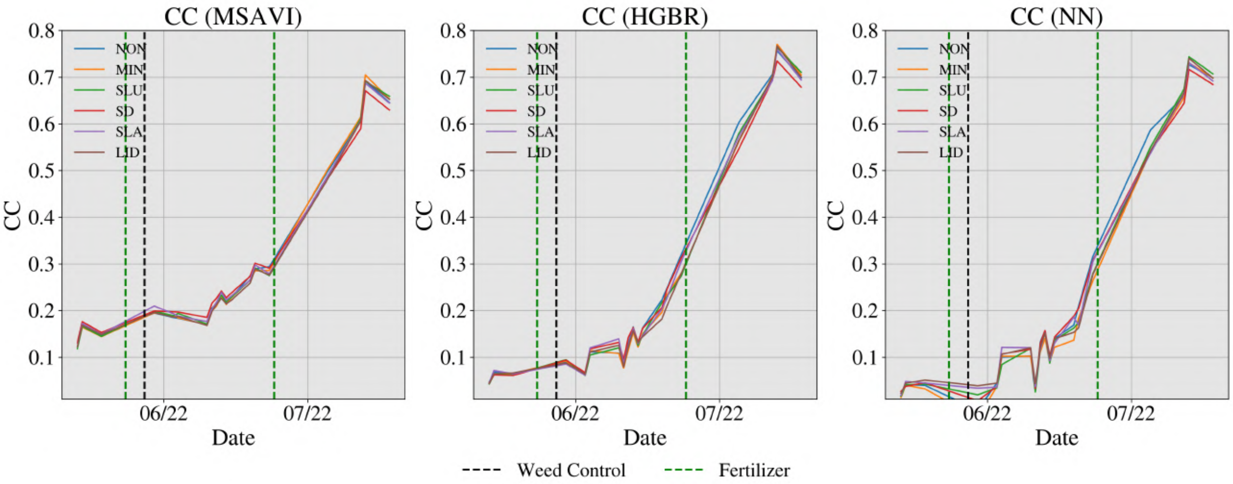


Figure 95. **Early Growth Stage Canopy Cover Predictions from Baseline Models for Corn at Experimental Site** ”Bio4Recycle” in Wallbach. The figure shows the development of predicted mean canopy cover (CC) values for different treatments, based on baseline models (MSAVI, HGBR, and NN), between May 2022 and July 2022. Fertilizer treatments are color-coded as follows: NON (grey), MIN (orange), SLU (green), SD (red), SLA (purple), and LID (brown). Grey shading indicates the duration of the growth seasons, while vertical dashed lines mark the dates of weed control (black) and fertilizer applications (green). Data © 2024 Planet Labs PBC.

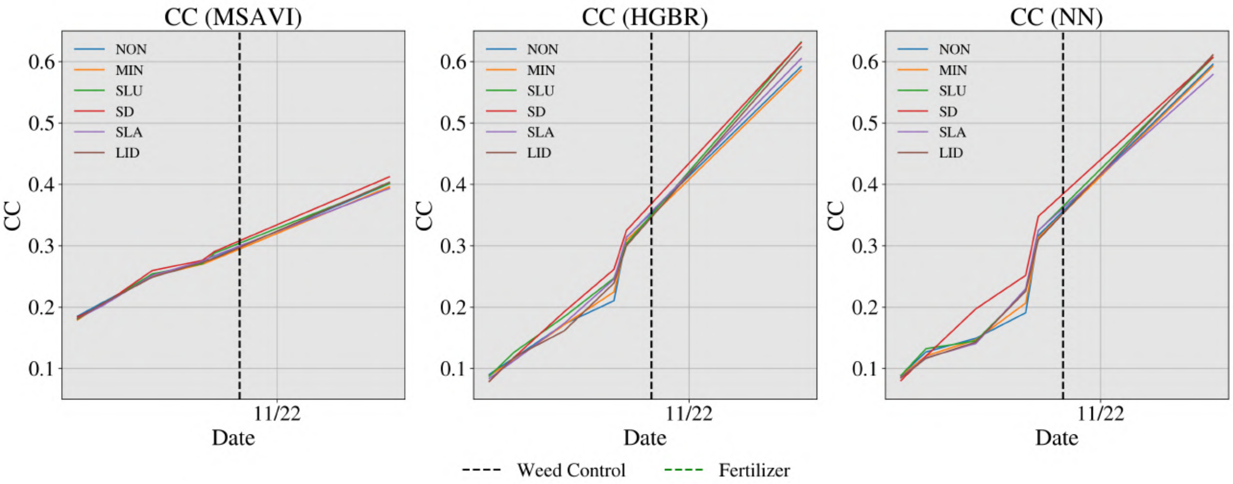


Figure 96. **Early Growth Stage Canopy Cover Predictions from Baseline Models for Winter Wheat at Experimental Site** ”Bio4Recycle” in Wallbach. The figure shows the development of predicted mean canopy cover (CC) values for different treatments, based on baseline models (MSAVI, HGBR, and NN), between October 2022 and November 2022. Fertilizer treatments are color-coded as follows: NON (grey), MIN (orange), SLU (green), SD (red), SLA (purple), and LID (brown). Grey shading indicates the duration of the growth seasons, while vertical dashed lines mark the dates of weed control (black) and fertilizer applications (green). Data © 2024 Planet Labs PBC.

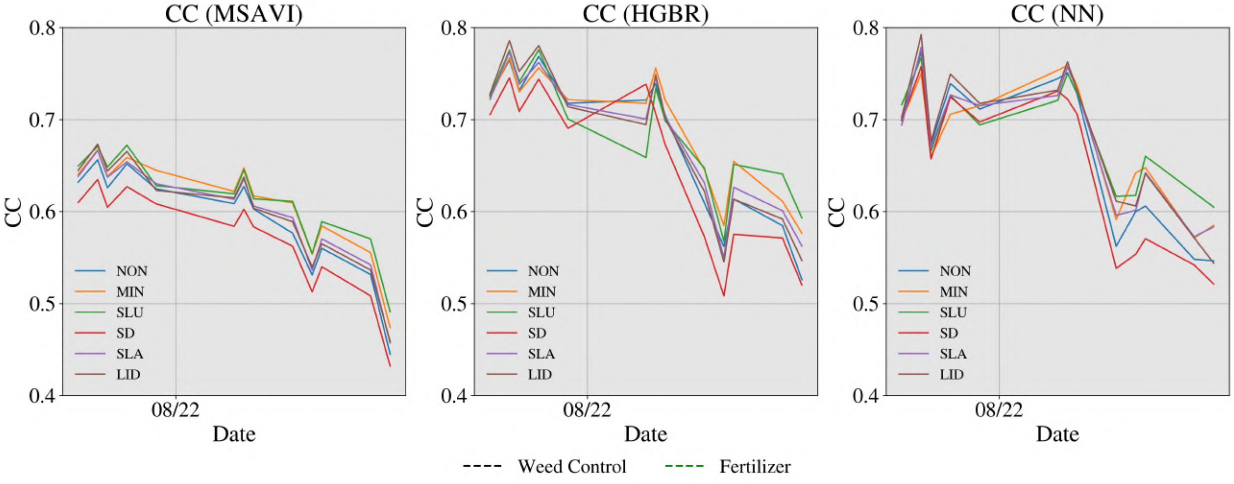


Figure 97. **Late Growth Stage Canopy Cover Predictions from Baseline Models for Corn at Experimental Site** ”Bio4Recycle” in Wallbach. The figure shows the development of predicted mean canopy cover (CC) values for different treatments, based on baseline models (MSAVI, HGBR, and NN), between July 2022 and August 2022. Fertilizer treatments are color-coded as follows: NON (grey), MIN (orange), SLU (green), SD (red), SLA (purple), and LID (brown). Grey shading indicates the duration of the growth seasons, while vertical dashed lines mark the dates of weed control (black) and fertilizer applications (green). Data © 2024 Planet Labs PBC.

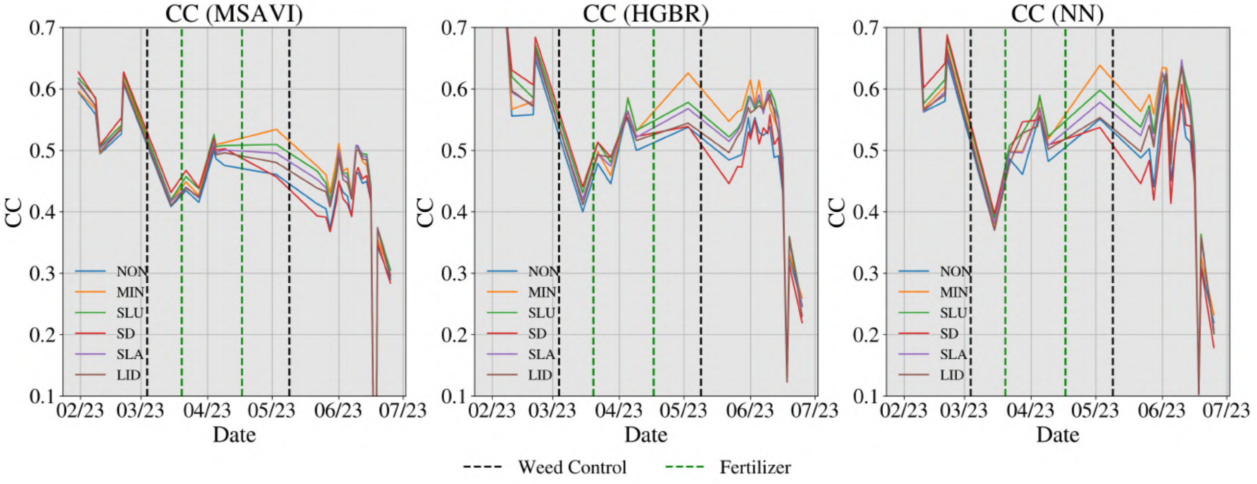


Figure 98. **Late Growth Stage Canopy Cover Predictions from Baseline Models for Winter Wheat at Experimental Site** ”Bio4Recycle” in Wallbach. The figure shows the development of predicted mean canopy cover (CC) values for different treatments, based on baseline models (MSAVI, HGBR, and NN), between January 2023 and July 2023. Fertilizer treatments are color-coded as follows: NON (grey), MIN (orange), SLU (green), SD (red), SLA (purple), and LID (brown). Grey shading indicates the duration of the growth seasons, while vertical dashed lines mark the dates of weed control (black) and fertilizer applications (green). Data © 2024 Planet Labs PBC.

PERSONAL DECLARATION

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

Date

April 25, 2025

Signature

Timon Boos