ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

SCIENZE DELLA VITA, DELLA TERRA E DELL'AMBIENTE

Ciclo XXXVIII

Settore Concorsuale: 05/A1 – BOTANICA

Settore Scientifico Disciplinare: BIO/03 – BOTANICA AMBIENTALE E APPLICATA

ARTIFICIAL INTELLIGENCE FOR BIODIVERSITY: UNDERSTANDING THE LIVING WORLD THROUGH MACHINE VISION

Presentata da: Ludovico Leopoldo Chieffallo

Supervisor
Duccio Rocchini

PhD Coordinator
Barbara Cavalazzi

Co-supervisor
Michele Torresani

Esame finale anno 2026

To my parents

Abstract

The loss of biodiversity represents a major environmental emergency of our time, with profound repercussions on ecosystems and the services they provide. The global recognition of the need for quantitative, reliable, and spatially explicit indicators for ecological monitoring is now globally recognized. However, traditional field-based methods are constrained by high costs, limited spatial coverage, and poor temporal replicability. In this context, the integration of proximal remote sensing through unmanned aerial vehicles (UAVs) and artificial intelligence (AI) techniques offers new opportunities to more efficiently, standardized, and scalably analyze and understand biodiversity. The thesis Artificial Intelligence for Biodiversity: Understanding the Living World Through Machine Vision explores the combined use of UAV imagery and machine learning models to estimate key ecological parameters related to vegetation and pollinators. The research is articulated into three main studies: (I) the ability of RGB UAV imagery and classical machine learning algorithms to estimate floral cover and infer bee community metrics is evaluated; (II) a generalizable model capable of maintaining robust performance across different grassland ecosystems; (III) the third applies convolutional neural networks (U-Net) to multispectral imagery for the semantic segmentation of vegetation and the extraction of reproducible landscape metrics. Overall, this study demonstrates how artificial intelligence (AI) can bridge the gap between ecological detail and spatial coverage, enabling the automated derivation of ecological indicators at operational scales. The results of this study contribute to the definition of a reproducible and transparent approach to biodiversity monitoring, fostering the integration of advanced computational techniques into conservation and environmental management programs.

Extended Abstract The loss of biodiversity is one of the most pressing challenges facing contemporary science and society. The decline of plant and animal species, intensified by agricultural intensification, climate change, and habitat fragmentation, undermines essential ecosystem processes and reduces the resilience of socio-ecological systems. It is essential to develop quantitative, reliable, and spatially explicit indicators describing the state and dynamics of biodiversity to guide effective conservation and management policies. However, traditional survey methods based on direct observation and taxonomic identification face structural limitations that limit their scalability and long-term applicability. Recent advances in remote sensing and artificial intelligence (AI) have transformed the way we observe and analyze nature. Imagery acquired by unmanned aerial vehicles (UAVs) enables the collection of ultra–high-resolution spatial data, providing detailed insights into vegetation structure and composition. Artificial intelligence, particularly machine learning (ML) and deep learning (DL) techniques, allows the extraction of ecologically relevant information from such imagery, translating spectral, textural, and contextual patterns into quantitative indicators of landscape cover and configuration. Thus, the integration of UAVs and AI represents a promising frontier for more efficient, standardized, and reproducible ecological monitoring. The dissertation AI for Biodiversity: Understanding the Living World Through Machine Vision is situated within this framework, presenting a methodological progression across three studies that explore the application of AI to the understanding and quantification of plant biodiversity and pollinator communities.

Chapter I-UAV RGB imagery and machine learning for estimating floral cover as a proxy for bee communities

The first study serves as the proof of concept for the entire thesis. Field surveys of flowering plants and bee communities were conducted in 30 grasslands across southeastern Netherlands, characterized by different land-use intensities, alongside ultra–high-resolution RGB UAV flights. This study assessed whether RGB imagery could accurately estimate floral cover and indirectly infer bee abundance and diversity. Several ML algorithms (RF, SVM, and shallow neural networks) were tested to explore the influence of spatial resolution. The results showed

significant correlations between UAV-derived and *in situ* estimates, with RF and NNs achieving the best performance ($R^2 \approx 0.80$ for floral cover; $R^2 \approx 0.65$ for bee abundance). This study highlights the importance of spatial resolution and its potential transferability to automated biodiversity monitoring programs.

Chapter II-Toward a unified and generalizable model

The second study extends the first by proposing a single-model approach that automatically recognizes floral cover across different grassland ecosystems and subsequently infers bee community metrics. Several algorithms (RF, GBM, SVM, and NNET) were compared using a consistent validation protocol and an integrated multisite dataset. Gradient-boosting machines emerged as the most robust model for estimating floral cover, showing near-linear relationships with field estimates and positive correlations with bee richness and diversity. This approach reduces the overfitting typical of site-specific models while improving the efficiency, repeatability, and scalability of monitoring application. This study also emphasizes the importance of consistent sampling and annotation protocols over time and space and highlights the need for multispectral data to further enhance model generalization.

Chapter III: Segmentation of Deep Semantic Vegetation

The third study addresses a distinct yet complementary problem: the semantic segmentation of vegetation (vegetation vs. non-vegetation) from a five-band UAV multispectral orthomosaic (RGB, red-edge, NIR). We developed a fully reproducible end-to-end pipeline based on a U-Net architecture, implementing spatial separation between training, validation, and test datasets, spectral channel normalization, and comparison against classical NDVI- and hybrid NDVI+Brightness-based baselines. The U-Net achieved high accuracy ($F_1 \approx 0.946$; IoU ≈ 0.898), producing consistent, georeferenced estimates of vegetated cover ($\sim 16.4\%$) and improving patch delineation relative to threshold-based methods. All workflow steps, from pre-processing to landscape metrics computation, were documented to ensure full reproducibility in accordance with open science principles.

Synthesis and Contributions

Taken together, the three studies outline a coherent research trajectory: from the proof of concept linking floral cover to bee community metrics, through the development of generalizable ML models, to the application of DL for vegetation segmentation. The thesis contributes to three main fronts: (I) providing empirical evidence of the potential of RGB UAV imagery as ecological proxies; (II) proposing a unified and scalable framework for automated biodiversity monitoring; and (III) introducing a reproducible DL pipeline for vegetation mapping and spatial analysis. Overall, the research demonstrates that AI can serve as an operational tool for ecological science, combining scientific rigor, transparency, and reproducibility with the efficiency required for large-scale environmental monitoring.

Contents

1 Introduction

The loss of biodiversity represents one of the most pressing and complex challenges in life and environmental sciences (Kremen et al.,2002). Over the past few decades, a substantial body of research has consistently demonstrated a significant and rapid decline in biological diversity across various organizational levels(Kleijn et al.,2009; Potts et al.,2003), including genetic, species, and ecosystem levels, with a significant reduction in both plant and animal species abundance and distribution (Kleijn et al.,2015; Cardinale et al.,2012). This trend, which is particularly evident in intensively managed and fragmented agricultural landscapes, reflects the combined influence of anthropogenic factors such as land-use change, pollution, the introduction of alien species, and, increasingly, climate change. The deterioration of biodiversity is not merely an esthetic or cultural loss; it undermines fundamental ecosystem processes, including pollination, nutrient cycling, hydrological regulation, and primary productivity, upon which socio-ecological systems' stability and resilience depend. At the international level, the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) and the Convention on Biological Diversity (CBD) have repeatedly emphasized the urgent need for quantitative, reliable, and spatially explicit indicators to assess the state and dynamics of biodiversity, thus informing evidence-based conservation and management strategies (Díaz et al.,2015). However, the acquisition of high-quality ecological data remains one of the main operational bottlenecks. Traditional methods, based on in situ surveys and direct taxonomic identification, provide great accuracy but suffer from structural limitations: limited spatial coverage, costly and logistically demanding temporal replication, and heterogeneity among operators or protocols can introduce difficult-to-control biases. Consequently, ecological observations tend to be concentrated in a few study sites, while vast areas remain poorly monitored or entirely lacking updated information (Torresani et al.,2023;

Chieffallo et al.,2025). In response to these constraints, recent years have witnessed an unprecedented acceleration in the use of remote sensing and Earth observation as tools to complement and enhance ecological monitoring (Torresani et al.,2020; Tamburlin et al.,2021)
(Rocchini et al.,2021; Torresani et al.,2018; Rocchini et al.,2022). Although satellite platforms provide global data at regular intervals, their spatial resolution often fails to capture habitat- or community-scale dynamics (Magnússon et al.,2021; Rocchini et al.,2022; Torresani et al.,2022). Proximal remote sensing using unmanned aerial vehicles (UAVs) can acquire ultra-high-resolution imagery (on the order of a few centimeters or even millimeters) while precisely controlling flight altitude, overlap, and temporal frequency at an intermediate scale. These capabilities make UAVs a flexible and highly promising tool for mapping and monitoring complex habitats, allowing the quantification of parameters such as vegetation cover, flowering phenology, or the 3D structure of plant canopies(Cruzan et al.,2016; Alvarez-Taboada et al.,2017). The evolution of artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), has opened new perspectives for the automated analysis of ecological imagery (Krizhevsky et al.,2012). The integration of UAV-based observation and AI bridges the gap between fine-scale ecological detail and broad spatial coverage, enabling more standardized, repeatable, and cost-effective monitoring strategies. Within this framework, computer vision plays a central role: it transforms spectral, textural, and contextual patterns within orthophotos into ecologically meaningful information, turning pixels into quantitative indicators of cover, structure, and landscape configuration (Christin et al.,2019). In the field of environmental and applied ecology, this approach is situated within a rapidly evolving scientific context, in which the ability to combine traditional ecological knowledge with advanced computational models is becoming essential to describe and understand the complexity of the living world(Liang et al.,2018; Chen et al.,2020). The potential of computer vision unfolds along a methodological continuum, ranging from classical learning algorithms, such as random forest, support vector machines, gradient boosting, and shallow neural networks, which operate on engineered features (spectral indices, textural statistics, morphological metrics), to convolutional neural networks (CNNs)

capable of learning hierarchical representations directly from data, with particular effectiveness in classification and semantic segmentation tasks. These tools make it possible to automate the visual interpretation of ecological images, reducing the analyst's subjectivity while expanding the spatial and temporal scale of ecological observation (Li et al.,2023; Conrady et al.,2022). This dissertation, titled Artificial Intelligence for Biodiversity: Understanding the Living World Through Machine Vision, is situated precisely along this methodological gradient, exploring how artificial intelligence (AI) techniques can support the understanding and monitoring of plant biodiversity and associated organisms. This study presents a coherent progression of three main studies, conceived as a unified framework moving from classical ML algorithms to DNs for semantic segmentation. The first study aimed to validate the use of UAV RGB imagery to estimate floral cover, a key functional resource for pollinators, and to assess the extent to which this proxy can predict bee community abundance and diversity at the site scale. The second study expands this perspective by developing a unified and generalizable ML model capable of maintaining robust performance across different grassland ecosystems and providing more scalable and transferable estimates over time and space. Finally, the third study addresses a distinct yet complementary problem: the application of DL techniques to the semantic segmentation of vegetation (not flowers), showing how a multispectral DL pipeline can produce high-resolution and reproducible landscape indicators suitable for advanced ecological analyses. Taken together, the three studies delineate a methodological transition from point-based, labor-intensive observations toward the automated, scalable, and statistically traceable derivation of ecological indicators. This evolution responds to a central need in contemporary ecological research: reconciling the accuracy of field-based measurements with the capacity to systematically and repeatedly monitor large territories. The general objectives of this thesis are as follows:

(1) to develop and evaluate artificial intelligence methods for quantifying biodiversity-relevant resources and vegetation covers at operational scales;

(2) to verify the potential of unmanned aerial vehicle (UAV)-observable proxies to infer patterns associated with pollinator habitat quality;

(3) to build reproducible and transferable pipelines capable of integrating accuracy metrics, spatial indicators, and computational traceability. The main research questions addressed are as follows:
To what extent does the floral cover estimated from UAV imagery and ML classifiers correlate with bee abundance and diversity? Is it possible to train a single generalist model that maintains robust performance across different grassland ecosystems and environmental conditions? How can a deep segmentation approach applied to UAV multispectral mosaics improve the delineation of vegetated areas, their patch structure, and the analysis of landscape metrics compared to classical spectral baselines? The answers to these questions articulate the original contribution of this work along three complementary dimensions: (I) providing empirical and quantitative evidence of the UAV–ML–pollinator link; (II) developing a generalizable and standardized ML framework for monitoring floral resources; and (III) defining a complete, reproducible, and comparable DL pipeline for vegetation mapping and the spatial analysis of biodiversity.

1.1 Chapter I-UAV RGB imagery and machine learning for estimating floral cover as a proxy for bee communities

The first study serves as the proof of concept for the entire thesis. Field surveys on flowering plants and bee communities were conducted in parallel with UHR UAV RGB acquisitions across 30 grasslands in southeastern Netherlands. This study assessed whether orthomosaics derived from RGB imagery could accurately estimate floral cover and, by extension, infer bee abundance and diversity at the site level. Classical ML algorithms (Random Forest, shallow neural networks, and SVMs) were tested, while also exploring the effect of spatial resolution (original images at $\sim$ 0.5 cm and resampled versions at 1, 2, and 5 cm). Results showed significant relationships between UAV-based and field-based estimates for both floral cover and bee community metrics, with the highest performances achieved at finer resolutions and with RF and neural networks compared to SVMs (e.g., $R^2 \approx 0.80$ for floral cover with RF; $R^2 \approx 0.65$ for bee abundance with RF). The analysis also confirmed that lowering resolution tends to overestimate “flower” pixels due to sub-pixel mixing, with practical implications for flight planning and altitude selection. Methodologically, the chapter highlights three key aspects: (I) the importance of spatial resolution for the spectral and textural separation of flowers from vegetated backgrounds; (II) the ability of tree-based and neural models to capture robust nonlinear patterns between RGB features and flower presence; and (III) the potential transferability of this approach toward standardized monitoring programs, where UAV automation could act as a force multiplier for costly and logistically demanding protocols. The chapter concludes by outlining the limitations and future perspectives: the need for independent validations and STM tests, integration of 3D information (photogrammetry), and standardization of ingestion and classification pipelines for large areas. Chapter I establishes the operational link between UAV-observable ecosystem proxies (floral cover) and pollinator community metrics, clarifying the conditions, constraints, and algorithmic choices that maximize accuracy. This framework provides the empirical basis for addressing the second research question, i.e., developing more general and scalable models.

1.2 Chapter II-Toward a unified and generalizable model: systematic comparison of ML algorithms and indirect pollinator inference

The second study advances from the proof of concept to a more operational framework to overcome the need to train SSMs. Unlike the previous study, this study pursues a single model approach capable of automatically recognizing floral cover across multiple grassland ecosystems and subsequently inferring bee abundance and diversity. Several algorithms (RF, GBM, SVM, and NNET) were compared using an integrated multi-site dataset under a unified training and validation protocol. Within this framework, GBM emerged as the most robust overall model for estimating floral cover from UAV RGB imagery, yielding near 1:1 correlations with field-based estimates and positive relationships with bee community metrics (richness and diversity). The chapter also emphasizes the often-overlooked practical considerations: balancing accuracy and generalization, sensitivity to flight parameters and site heterogeneity, and the importance of consistent sampling and annotation protocols over time and space. From an ecological standpoint, the use of floral cover as an indirect indicator is grounded in well-known relationships between resource availability and bee community structure. This study demonstrates how an automatically derived, spatially explicit indicator can contribute to rapid habitat assessments in agricultural and semi-natural contexts by complementing (rather than replacing) expert-based taxonomic surveys. The unified approach reduces overfitting associated with per-site models, at the cost of a slight loss in local accuracy, while gaining efficiency, repeatability, and scalability for large-scale biodiversity monitoring programs. The chapter also outlines future development directions: improving spatial and radiometric resolution, integrating multispectral/hyperspectral data, and exploring DL architectures to enhance generalization across highly heterogeneous environments. In summary, the UAV + ML combination can be structured into a generalist framework, where a single model provides robust and transferable estimates of floral cover and, by extension, indirect indicators of pollinator habitat quality. This scalability forms the logical bridge to the third study, which addresses a distinct yet methodologically related challenge: deep mapping of vegetation (not flowers) using a reproducible DL pipeline comparable to classical spectral baselines.

1.3 Chapter III-Deep Semantic Segmentation of Vegetation Using U-Net and Landscape Metrics: A Reproducible Ecological Monitoring Pipeline

The third study did not focus on flower mapping. Instead, it addresses the semantic segmentation of vegetation (vegetation vs. non-vegetation) using a five-band unmanned aerial vehicle multispectral orthomosaic (RGB, red-edge, NIR). It develops a fully reproducible end-to-end workflow extending from radiometric and geometric pre-processing to U-Net training, vectorization, and LM computation. The approach adopts a rigorous spatial separation of training/validation/test sets with buffer zones to minimize spatial autocorrelation and applies spectral channel normalization. The

U-Net model, trained on expert-annotated orthophotos, is evaluated against classical baselines based on spectral indices (NDVI) and a hybrid NDVI+Brightness method, allowing explicit comparison between the DL pipeline and optimized threshold-based methods (e.g., NDVI ≥ 0.25; Brightness ≤ 0.13). In terms of performance, the U-Net achieved high accuracy on the test set (F1 ≈ 0.946; IoU ≈ 0.898) and estimated vegetated cover of $\sim$16.37% ($\approx$17,446 m^2) in the study area, whereas the NDVI baseline yielded $\sim$17.88%. The most pronounced differences were observed in patch geometry and complexity (edge density, shape indices), where deep segmentation produced sharper boundaries and reduced overestimation, which is typical of threshold-based methods. The pipeline outputs georeferenced products (binary masks, vector polygons, and landscape metrics such as patch density, edge density, largest patch index, and MSI/AWMSI), ready for GIS integration and standardized spatial ecological analyses. All steps and intermediate artifacts are archived to ensure full reproducibility and computational auditability.

The scientific contribution of this chapter is twofold. First, deep semantic segmentation applied to UAV multispectral data enables the derivation of spatial vegetation indicators with superior detail and consistency compared to spectral baselines, where threshold selection and ad hoc rules often introduce subjectivity and limit transferability. Second, it provides a transparent and well-documented pipeline from pre-processing to metric computation that can be retrained on different target classes (beyond vegetation) and transferred to other contexts, facilitating cross-study comparison and promoting open-science practices (open datasets and code). This directly addresses a recognized need within the community: establishing standardized, repeatable, and declarative workflows to move from UAV imagery to landscape-scale ecological indicators consistently.

References

[Alvarez-Taboada et al.,2017] Alvarez-Taboada, Flor, Claudio Paredes, and Julia Julián-Pelaz. "Mapping of the invasive species Hakea sericea using unmanned aerial vehicle (UAV) and WorldView-2 imagery and an object-oriented approach." Remote Sensing 9.9 (2017): 913.

[Agisoft LLC.,2023] Agisoft LLC. (2023). Agisoft Metashape (Versione 1.7.1) [Software]. `https://www.agisoft.com`

[Atanasov et al.,2024] Atanasov, Atanas Z., et al. "A Novel Algorithm to Detect White Flowering Honey Trees in Mixed Forest Ecosystems Using UAV-Based RGB Imaging." AgriEngineering 6.1 (2024): 95-112.

[Biesmeijer et al.,2006] Biesmeijer, Jacobus C., et al. "Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands." Science 313.5785 (2006): 351-354.

[Blaauw et al.,2014] Blaauw, Brett R., and Rufus Isaacs. "Flower plantings increase wild bee abundance and the pollination services provided to a pollination-dependent crop." Journal of Applied Ecology 51.4 (2014): 890-898.

[Breiman,1996] Breiman, L. Bagging predictors. Mach Learn 24, 123–140 (1996). https://doi.org/10.1007/BF00058655

[Breiman,2001] Breiman, L. Random Forests. Machine Learning 45, 5–32 (2001). https://doi.org/10.1023/A:1010933404324

[Capolupo et al.,2015] Capolupo, Alessandra, et al. "Estimating plant traits of grasslands from UAV-acquired hyperspectral images: a comparison of statistical approaches." ISPRS International Journal of Geo-Information 4.4 (2015): 2792-2820.

[Cardinale et al.,2012] Cardinale, Bradley J., et al. "Biodiversity loss and its impact on humanity." Nature 486.7401 (2012): 59-67.

[Cervantes et al.,2020] Cervantes, Jair, et al. "A comprehensive survey on support vector machine classification: Applications, challenges and trends." Neurocomputing 408 (2020): 189-215.

[Chen et al.,2020] Chen, Guobin, Zhiyong Jiang, and M. M. Kamruzzaman. "Radar remote sensing image retrieval algorithm based on improved Sobel operator." Journal of Visual Communication and Image Representation 71 (2020): 102720.

[Conrady et al.,2022] Conrady, Christopher R., et al. "Automated detection and classification of southern African Roman seabream using mask R-CNN." Ecological Informatics 69 (2022): 101593.

[Chieffallo et al.,2025] Chieffallo, Ludovico, et al. "Machine learning for biodiversity: UAV-based flower detection as an indirect proxy for bee abundance." Ecological Informatics (2025): 103346.

[Cristianini & Taylor,2013] Cristianini N, Shawe-Taylor J.(2013) Kernel-Induced Feature Spaces. In: An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press; 2000:26-51.

[Christin et al.,2019] Christin, Sylvain, Éric Hervet, and Nicolas Lecomte. "Applications for deep learning in ecology." Methods in Ecology and Evolution 10.10 (2019): 1632-1644.

[Cruzan et al.,2016] Cruzan, Mitchell B., et al. "Small unmanned aerial vehicles (micro-UAVs, drones) in plant ecology." Applications in plant sciences 4.9 (2016): 1600041.

[Cutler A. et al.,2012] Cutler, A., Cutler, D.R., Stevens, J.R. (2012). Random Forests. In: Zhang, C., Ma, Y. (eds) Ensemble Machine Learning. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-9326-7_5

[Dietterich,T.G.,2000] Dietterich, T.G. (2000). Ensemble Methods in Machine Learning. In: Multiple Classifier Systems. MCS 2000. Lecture

Notes in Computer Science, vol 1857. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45014-9_1

[Díaz et al.,2015] Díaz, Sandra, et al. "The IPBES Conceptual Framework—connecting nature and people." Current opinion in environmental sustainability 14 (2015): 1-16.

[Ebeling et al.,2008] Ebeling, Anne, et al. "How does plant richness affect pollinator richness and temporal stability of flower visits?." Oikos 117.12 (2008): 1808-1815.

[Falk et al.,2017] Falk, Steven, and R. Lewington. Veldgids bijen voor Nederland en Vlaanderen. 2017.

[Friedman,2002] Friedman, Jerome H. "Stochastic gradient boosting." Computational statistics & data analysis 38.4 (2002): 367-378.

[Fründ et al.,2010] Fründ, Jochen, Karl Eduard Linsenmair, and Nico Blüthgen. "Pollinator diversity and specialization in relation to flower diversity." Oikos 119.10 (2010): 1581-1590.

[Gallmann et al.,2022] Gallmann, Johannes, et al. "Flower mapping in grasslands with drones and deep learning." Frontiers in plant science 12 (2022): 774965.

[Gallai et al.,2009] Gallai, Nicola, et al. "Economic valuation of the vulnerability of world agriculture confronted with pollinator decline." Ecological economics 68.3 (2009): 810-821.

[Ghazoul,2005] Ghazoul, Jaboury. "Buzziness as usual? Questioning the global pollination crisis." Trends in ecology & evolution 20.7 (2005): 367-373.

[Ghisbain et al.,2023] Ghisbain, Guillaume, et al. "The new annotated checklist of the wild bees of Europe (Hymenoptera: Anthophila)." Zootaxa 5327.1 (2023): 1-147.

[Goodfellow et al.,2016] Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press, 2016.

[Guo et al.,2016] Guo, Qinghua, et al. "Perspectives and prospects of unmanned aerial vehicle in remote sensing monitoring of biodiversity." Biodiversity Science 24.11 (2016): 1267.

[Kaneko et al.,2014] Kaneko, Korehisa, and Seiich Nohara. "Review of effective vegetation mapping using the UAV (Unmanned Aerial Vehicle) method." Journal of Geographic Information System 6.06 (2014): 733.

[Kattenborn et al.,2020] Kattenborn, Teja, et al. "Convolutional Neural Networks accurately predict cover fractions of plant species and communities in Unmanned Aerial Vehicle imagery." Remote Sensing in Ecology and Conservation 6.4 (2020): 472-486.

[Kavzoglu et al.,2020] Kavzoglu, Taskin, Furkan Bilucan, and Alihan Teke. "Comparison of support vector machines, random forest and decision tree methods for classification of sentinel-2A image using different band combinations." 41st Asian Conference on Remote Sensing (ACRS 2020). Vol. 41. 2020.

[Kuhn M. & Johnson K.,2013] Kuhn, M., & Johnson, K. (2013). Applied Predictive Modeling. Springer New York, NY. https://doi.org/10.1007/978-1-4614-6849-3

[Kleijn et al.,2009] Kleijn, David, et al. "On the relationship between farmland biodiversity and land-use intensity in Europe." Proceedings of the royal society B: biological sciences 276.1658 (2009): 903-909.

[Kleijn et al.,2015] Kleijn, David, et al. "Delivery of crop pollination services is an insufficient argument for wild pollinator conservation." Nature communications 6.1 (2015): 7414.

[Klein et al.,2007] Klein, Alexandra-Maria, et al. "Importance of pollinators in changing landscapes for world crops." Proceedings of the royal society B: biological sciences 274.1608 (2007): 303-313.

[Kremen et al.,2002] Kremen, Claire, Neal M. Williams, and Robbin W. Thorp. "Crop pollination from native bees at risk from agricultural intensification." Proceedings of the National Academy of Sciences 99.26 (2002): 16812-16816.

[Krizhevsky et al.,2012] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems 25 (2012).

[LeCun et al.,2015] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436-444.

[Liu et al.,2018] Liu, Zelin, et al. "Application of machine-learning methods in forest ecology: recent progress and future challenges." Environmental Reviews 26.4 (2018): 339-350.

[Li et al.,2023] Li, Jiajun, et al. "Strawberry R-CNN: Recognition and counting model of strawberry based on improved faster R-CNN." Ecological Informatics 77 (2023): 102210.

[Liang et al.,2018] Liang, Xining, et al. "A real-time infrared image enhancement algorithm based on improved CLAHE." 2018 international conference on image and video processing, and artificial intelligence. Vol. 10836. SPIE, 2018.

[Liaw A. et al.,2022] Liaw, Andy, and Matthew Wiener. "Classification and regression by randomForest." R news 2.3 (2002): 18-22.

[Magnússon et al.,2021] Magnússon, Rúna Í., et al. "Shrub decline and expansion of wetland vegetation revealed by very high resolution land cover change detection in the Siberian lowland tundra." Science of the Total Environment 782 (2021): 146877.

[Moe et al.2020] Moe, Kyaw Thu, et al. "Comparing individual tree height information derived from field surveys, LiDAR and UAV-DAP for high-value timber species in Northern Japan." Forests 11.2 (2020): 223.

[Nair & Hinton,2010] Nair, V., & Hinton, G. E. (2010). Rectified Linear Units Improve Restricted Boltzmann Machines. Proceedings of the 27th International Conference on Machine Learning (ICML-10).

[Nasiri et al.,2022] Nasiri, Vahid, et al. "Modeling forest canopy cover: a synergistic use of Sentinel-2, aerial photogrammetry data, and machine learning." Remote Sensing 14.6 (2022): 1453.

[Natekin et al.,2013] Natekin, Alexey, and Alois Knoll. "Gradient boosting machines, a tutorial." Frontiers in neurorobotics 7 (2013): 21.

[Neumüller et al.,2021] Neumüller, Ulrich, et al. "Prolonged blooming season of flower plantings increases wild bee abundance and richness in agricultural landscapes." Biodiversity and Conservation 30 (2021): 3003-3021.

[Nieuwenhuijsen et al.,2016] Nieuwenhuijsen, Hans, and Theo Peeters, eds. Nederlandse bijen op naam brengen: deel 1. Stichting Jeugdbondsuitgeverij, 2016.

[Nieuwenhuijsen et al.,2020] Nieuwenhuijsen, Hans, Theo Peeters e Dominic Dijkshoorn, eds. Nederlandse bijen op naam brengen: deel 2. Stichting Jeugdbondsuitgeverij, 2020.

[Ollerton et al.,2011] Ollerton, Jeff, Rachael Winfree, and Sam Tarrant. "How many flowering plants are pollinated by animals?." Oikos 120.3 (2011): 321-326.

[Perrone et al.,2024] Perrone, Michela, et al. "“Flower power”: How flowering affects spectral diversity metrics and their relationship with plant diversity." Ecological Informatics 81 (2024): 102589.

[Pielou et al.,1966] Pielou, Evelyn Chrystalla. "Shannon's formula as a measure of specific diversity: its use and misuse." The American Naturalist 100.914 (1966): 463-465.

[Potts et al.,2003] Potts, Simon G., et al. "Linking bees and flowers: how do floral communities structure pollinator communities?." Ecology 84.10 (2003): 2628-2642.

[Potts et al.,2021] Potts, Simon, et al. "Proposal for an EU pollinator monitoring scheme." (2021).

[Prechelt,1998] Prechelt, L. (1998). Early Stopping - But When?. Neural Networks: Tricks of the Trade, 55-69.

[Randelovic,2020] RANDELOVIĆ, Predrag. "(2020). Prediction of soybean plant density using a machine learning model and vegetation indices extracted from RGB images taken with a UAV." Agronomy. España 10.8: 1108.

[Reed et al.,2003] Reed, Bradley C., Michael White, and Jesslyn F. Brown. "Remote sensing phenology." Phenology: an integrative environmental science (2003): 365-381.

[Redhead,John W. et al.,2016] Redhead, John W., et al. "Effects of habitat composition and landscape structure on worker foraging distances of five bumble bee species." Ecological applications 26.3 (2016): 726-739.

[Robinson and Sutherland,2002] Robinson, Robert A., and William J. Sutherland. "Post-war changes in arable farming and biodiversity in Great Britain." Journal of applied Ecology 39.1 (2002): 157-176.

[Rocchini et al.,2021] Rocchini, Duccio, et al. "rasterdiv—An Information Theory tailored R package for measuring ecosystem heterogeneity from space: To the origin and back." Methods in Ecology and Evolution 12.6 (2021): 1093-1102.

[Rocchini et al.,2022] Rocchini, Duccio, et al. "Double down on remote sensing for biodiversity estimation: a biological mindset." Community Ecology 23.3 (2022): 267-276.

[Rocchini et al.,2024] Rocchini, Duccio, et al. "Under the mantra:'Make use of colorblind friendly graphs'." Environmetrics (2024): e2877.

[Rodríguez et al.,2024] Rodríguez Ballesteros, Alberto, et al. "Towards acoustic monitoring of bees: wingbeat sounds are related to species and individual traits." Philosophical Transactions of the Royal Society B 379.1904 (2024): 20230111.

[Rumelhart et al.,1986] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533-536.

[Saunders et al.,1991] Saunders, Denis A., Richard J. Hobbs, and Chris R. Margules. "Biological consequences of ecosystem fragmentation: a review." Conservation biology 5.1 (1991): 18-32.

[Scheper et al.,2014] Scheper, Jeroen, et al. "Museum specimens reveal loss of pollen host plants as key factor driving wild bee decline in The Netherlands." Proceedings of the National Academy of Sciences 111.49 (2014): 17552-17557.

[Scheper et al.,2015] Scheper, Jeroen, et al. "Local and landscape-level floral resources explain effects of wildflower strips on wild bees across four European countries." Journal of Applied Ecology 52.5 (2015): 1165-1175.

[Scheper et al.,2021] Scheper, Jeroen, et al. "Attractiveness of sown wildflower strips to flower-visiting insects depends on seed mixture and establishment success." Basic and Applied Ecology 56 (2021): 401-415.

[Schölkopf & Smola,2018] Schölkopf, B., & Smola, A. J. (2018). Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. The MIT Press. https://doi.org/10.7551/mitpress/4175.001.0001

[Segre et al.,2023] Segre, Hila, et al. "Butterflies are not a robust bioindicator for assessing pollinator communities, but floral resources offer a promising way forward." Ecological Indicators 154 (2023): 110842.

[Sledevič et al.,2025] Sledevič, Tomyslav, et al. "Visual recognition of honeybee behavior patterns at the hive entrance." PloS one 20.2 (2025): e0318401.

[Srivastava et al.,2014] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research, 15(1), 1929-1958.

[Steffan-Dewenter et al.,2001] Steffan-Dewenter, I., and T. Tscharntke. "Succession of bee communities on fallows. Ecography (Cop.) 24, 83–93." (2001).

[Steffan-Dewenter et al.,2005] Steffan-Dewenter, Ingolf, Simon G. Potts, and Laurence Packer. "Pollinator diversity and crop pollination services are at risk." Trends in ecology & evolution 20.12 (2005): 651-652.

[Surowiecki,2005] Surowiecki, James. "The Wisdom of Crowds/James Surowiecki." NY.: Anchor (2005).

[Sutter et al.,2017] Sutter, Louis, et al. "Enhancing plant diversity in agricultural landscapes promotes both rare bees and dominant crop-pollinating bees through complementary increase in key floral resources." Journal of Applied Ecology 54.6 (2017): 1856-1864.

[Tilman et al.,2002] Tilman, David, et al. "Agricultural sustainability and intensive production practices." Nature 418.6898 (2002): 671-677.

[Tamburlin et al.,2021] Tamburlin, Daniel, et al. "Testing the height variation hypothesis with the R Rasterdiv package for tree species diversity estimation." Remote Sensing 13.18 (2021): 3569.

[Thessen,2016] Thessen, Anne. "Adoption of machine learning techniques in ecology and earth science." One Ecosystem 1 (2016): e8621.

[Theodorou et al.,2020] Theodorou, Panagiotis, et al. "Urban fragmentation leads to lower floral diversity, with knock-on impacts on bee biodiversity." Scientific Reports 10.1 (2020): 21756

[Torresani et al.,2018] Torresani, et al. "Testing the spectral variation hypothesis by using the RAO-Q index to estimate forest biodiversity: Effect of spatial resolution." IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2018.

[Torresani et al.,2020] Torresani, Michele, et al. "Height variation hypothesis: A new approach for estimating forest species diversity with CHM LiDAR data." Ecological Indicators 117 (2020): 106520.

[Torresani et al.,2022] Torresani, Michele, et al. "Correlation analysis of evapotranspiration, emissivity contrast and water deficit indices: A case study in four eddy covariance sites in italy with different environmental habitats." Land 11.11 (2022): 1903.

[Torresani et al.,2023] Torresani, Michele, et al. "A novel approach for surveying flowers as a proxy for bee pollinators using drone images." Ecological Indicators 149 (2023): 110123.

[Torresani et al.,2024] Torresani, Michele, et al. "Grassland vertical height heterogeneity predicts flower and bee diversity: an UAV photogrammetric approach." Scientific Reports 14.1 (2024): 809.

[Van Swaay et al.,2008] van Swaay, Chris AM, et al. "Butterfly monitoring in Europe: methods, applications and perspectives." Biodiversity and Conservation 17 (2008): 3455-3469.

[Vapnik V.,1995] Vapnik, V. N. (1995). The Nature of Statistical Learning Theory. Springer New York, NY. https://doi.org/10.1007/978-1-4757-2440-0

[Westphal et al.,2008] Westphal, Catrin, et al. "Measuring bee diversity in different European habitats and biogeographical regions." Ecological monographs 78.4 (2008): 653-671.

[Williams et al.,2009] Williams, Paul H., and Juliet L. Osborne. "Bumblebee vulnerability and conservation world-wide." Apidologie 40.3 (2009): 367-387.

[Winfree et al.,2009] Winfree, Rachael, et al. "A meta-analysis of bees' responses to anthropogenic disturbance." Ecology 90.8 (2009): 2068-2076.

[Wratten et al.,2012] Wratten, Stephen D., et al. "Pollinator habitat enhancement: benefits to other ecosystem services." Agriculture, Ecosystems & Environment 159 (2012): 112-122.

[Zhang et al.,2019] Zhang, Yatao, et al. "Comparing the performance of random forest, SVM and their variants for ECG quality assessment combined with nonlinear features." Journal of medical and biological engineering 39 (2019): 381-392.

2 Chapter I

A novel approach for surveying flowers as a proxy for bee pollinators using drone images

Michele Torresani [a,b,*], David Kleijn [c], Jan Peter Reinier de Vries [c], Harm Bartholomeus [d], Ludovico Chieffallo [b], Roberto Cazzolla Gatti [b], Vítězslav Moudrý[e], Daniele Da Re [f], Enrico Tomelleri [a], Duccio Rocchini [b,e]

[a] Free University of Bolzano/Bozen, Faculty of Agricultural, Environmental and Food Sciences,
Piazza Università 1, 39100 Bolzano/Bozen, Italy

[b] BIOME Lab, Department of Biological, Geological and Environmental Sciences,
Alma Mater Studiorum University of Bologna, via Irnerio 42, 40126 Bologna, Italy

[c] Plant Ecology and Nature Conservation Group, Wageningen University,
Droevendaalsesteeg 3a, Wageningen 6708PB, The Netherlands

[d] Laboratory of Geo-Information Science and Remote Sensing,
Wageningen University and Research, P.O. Box 47, 6700 AA Wageningen, The Netherlands

[e] Czech University of Life Sciences Prague, Faculty of Environmental Sciences,
Department of Spatial Sciences, Kamýcka 129, Praha–Suchdol 16500, Czech Republic

[f] Georges Lemaître Center for Earth and Climate Research, Earth and Life Institute,
UCLouvain, Louvain-la-Neuve, Belgium

2.1 Abstract

The abundance and diversity of plants and insects are important indicators of biodiversity, overall ecosystem health and agricultural production. Bees in particular are interesting indicators as they provide a key ecosystem service in many agricultural crops. Worldwide, habitat loss and fragmentation, agricultural intensification and climate change are important drivers of plant and bee decline (Potts et al.,2003; Kleijn et al.,2009). Monitoring of plants and bees is a crucial first step to safeguard their diversity and the services they provide but traditional in situ methods are time consuming and expensive. Remote sensing and Earth observation have the advantages that they can cover large areas and provides repeated, spatially continuous and standardized information. However, to date it has proven challenging to use these methods to assess small-scaled species-level biodiversity components with this approach. Here we surveyed bees and flowering plants using conventional field methods in 30 grasslands along a land-use intensity gradient in the Southeast of the Netherlands. We took RGB (true colored Red-Green–Blue) images using an Unmanned Aerial Vehicle (UAV) from the same fields and tested whether remote sensing can provide accurate assessments of flower cover and diversity and, by association, bee abundance and diversity. We explored the performance of different machine learning methods: Random Forest (RF), Neural Networks (NNET) and SupportVector Machine (SVM). To evaluate the effect of the spatial resolution on the accuracy of the estimates, we tested all approaches using images at the original spatial resolution (~ 0.5 cm) and re-sampled at 1 cm, 2 cm and 5 cm. We generally found significant relationships between UAV RGB derived estimates of flower cover and in situ estimates of flower cover and bee abundance and diversity. The highest resolution images generally resulted in the strongest relationships, with RF and NNET methods producing considerably better results than SVM methods (flower cover RF R2 = 0.8, NNET R2 = 0.79; bee abundance RF R2 = 0.65, NNET R2 = 0.54, bee species richness RF R2 = 0.62, NNET R2 = 0.52; bee species diversity RF R2 = 0.54, NNET R2 = 0.46). Our results suggest that methods based on the coupling of UAV imagery and machine learning methods can be developed into valuable tools for large-scale, standardized and cost-effective monitoring of flower cover and therefore of an important aspect of habitat quality for bees.

Keywords: bees, biodiversity, flowering plants, machine learning, monitoring, spatial resolution, UAV

2.2 Introduction

Worldwide, biodiversity is declining at unprecedented rates, threatening the persistence of many species and the benefits that humans derive from ecosystems (24). Since the second half of the 20th century, biodiversity in agricultural areas in particular has been decreasing drastically due to habitat fragmentation and intensification of farming practices (51; 45; 23) with negative consequences on ecosystem functioning (27). Bees are one of the best studied species groups in this respect. Studies show a marked decline in the distribution of populations at national scales and in abundance and richness of populations at the local scale (75; 14; 59; 2; 74). These have mainly been attributed to habitat loss and fragmentation, intensification of farming, climate change and loss of host plants (54). Bees provide essential pollination services that support maximum productivity of different 76% of the leading global food crops (25), and promote seed set of about 87% of wild plants globally (38). In fact, one-third of the worldwide food production take advantage directly or indirectly from this ecosystem service which is valued more than 150 billion euros a year worldwide (13; 12).

We now have a relatively good understanding of drivers of bee population decline and the effectiveness of mitigation measures. However, we still have a poor understanding of the actual extent of population decline itself, especially at large spatial scales, and how this is linked to the quality of their habitat. Unlike butterflies, that are being monitored in a standardized way by citizen scientists in more than 10 European countries (70), large-scale surveying of bees by laypersons is challenging because bee species are difficult to identify. Studies examining trends in bee abundance are therefore invariably small-scaled while large scale trends are based on bee distribution data (2), which generally underestimates population trends. There are now advanced proposals for standardized monitoring of pollinators in EU member states (40). Although this will vastly improve our grasp on pollinator trends, because *in-situ* monitoring is costly, a standardized monitoring program can inevitably cover only a small proportion of the relevant surface area. A complementary approach could be to infer trends in bee pollinators from trends in flower cover and species richness. Bees forage for nectar and pollen from flowers and rely exclusively on these resources for sustenance throughout their lifecycles. Flower cover and diversity have been linked to bee abundance and diversity in several studies (41; 61) and can therefore be considered an indicator of an important aspect of bee habitat quality. Assessing flower cover and diversity is less time-consuming than assessing bee abundance and species richness especially when this can be done in an automated fashion.

Recently, remote sensing and Earth observation has made significant progress with estimating various aspects of biodiversity in a standardized way at large spatial scales (68; 62; 48; 31; 49). Satellite remote sensing (hereafter SRS) has the advantage that it covers vast spatial scales, it can provide information on a variety of ecological characteristics such as vegetation distribution (e.g. vegetation dynamics (30; 67), plant phenology (44), habitat and environmental diversity (66; 47; 65), and plant size distributions (55; 69) over large areas. However, it has not yet been proven possible to accurately describe with SRS vegetation aspects that can only be measured at detailed scales of up to a few mm, such as the estimation of flower cover (6). The technology that has recently evolved around Unmanned Aerial Vehicles (UAVs), in particular photogrammetry based on structure-from-motion algorithms, has led to the development of highly detailed orthomosaics and 3D information over large areas at a relatively low cost with the accuracy of some mm's to cm's (13; 22). So far, studies have shown that it is possible to use UAV images to estimate vegetation properties such as vegetation diversity (16), species and plant community distribution (21), plant trait distribution (5) and the mapping and monitoring of invasive species (1).

In this paper we aim to assess flower cover as a proxy for bee abundance using images from a UAV. The fundamental step in estimating flower cover is distinguishing flower pixels from grass or soil pixels by means of differences in the spectral signatures (18). Different types of images can be used for this purpose. In our study, we decided to rely on RGB (Red-Green-Blue, true colored) images. These images, in comparison with multispectral/hyperspectral images (i.e. the collection of image layers that have been acquired each at a different wavelength band), have less power in characterizing the different spectral signatures of vegetation but they have a number of strengths that make them highly competitive in ecological classification analysis. RGB cameras are relatively cheap, readily available on the market and often come mounted as an integral part of commercial drones. Furthermore, RGB images require less pre-post processing analyses. This makes RGB cameras more user-friendly and allow for a more easy reproduction of analyses done by different operators or stakeholders (e.g., ecologists, farmers, researchers).

The spatial resolution of the images also plays an important role in the assessment of flower cover. Generally, fine spatial resolution offers more detailed information and reduces the mixed pixel issue (18). On the other hand, fine spatial resolution requires lower flight heights which limits the geographical area covered by the UAV in a single flight (20). In addition to spectral and spatial resolution issues, the choice of an adequate classification method is also significant for the generation of reliable classification results. Since more than two decades, different machine learning methods such Random Forest (RF), Neural Networks (NNET) and Support-vector machine (SVM) have been used extensively for classification with remote sensing data with good results (28). With the information obtained from UAV data, these models can be a powerful and efficient tool for the detailed characterization of the vegetation (42). For example, Guo et al. (17) estimated the chlorophyll content of maize with UAV RGB images using three different ML methods (RF, SVM and NNET) finding that RF performed the best for their purpose. Sandino et al. (50) made use of UAV images and an Extreme Gradient Boosting ML algorithm for monitoring of invasive grasses and vegetation in remote arid lands reaching a detection rates of around 96 %. A similar level of accuracy was found by combining convolutional neural network ML algorithms and RGB UAV images for the estimation of *Convallaria keiskei* patches (56).

The aim of this study is to test whether we can estimate flower cover using UAV RGB images and to assess whether flower cover can be used as a proxy for bee abundance and diversity. Additionally we explored to what extent the observed relationships were affected by using different machine learning algorithms (RF, NNET, SVM) and by using image data with different spatial resolution (0.5cm, 1cm, 2cm and 5cm). Our study system consists of grasslands situated in the Southeast of the Netherlands that vary strongly in management intensity and therefore represent a gradient in flower cover. From each grassland, *in-situ* data was collected within two days of obtaining the UAV images.

2.3 Materials and Methods

2.3.1 Study area

The study area (7 x 10 km) is located in the Southeast of the Netherlands (Fig. 1) and consists of a mosaic of different land use types including intensive agriculture, low-intensity farming and nature reserves. We selected 30 grasslands that covered a land use intensity gradient ranging from nutrient-poor, biodiversity-rich semi-natural grasslands to intensively fertilized grasslands for fodder production. The study sites are part of the experimental biodiversity area network of the EU Showcase project `https://showcase-project.eu/`. The study sites are situated on loess soils, colluvial clay deposits and locally lime-rich soils, and range in elevation from 70 to 171 m asl. We minimized spatial clustering of specific grassland types by selecting semi-natural, extensively used and intensively used grasslands from different parts of the study area.

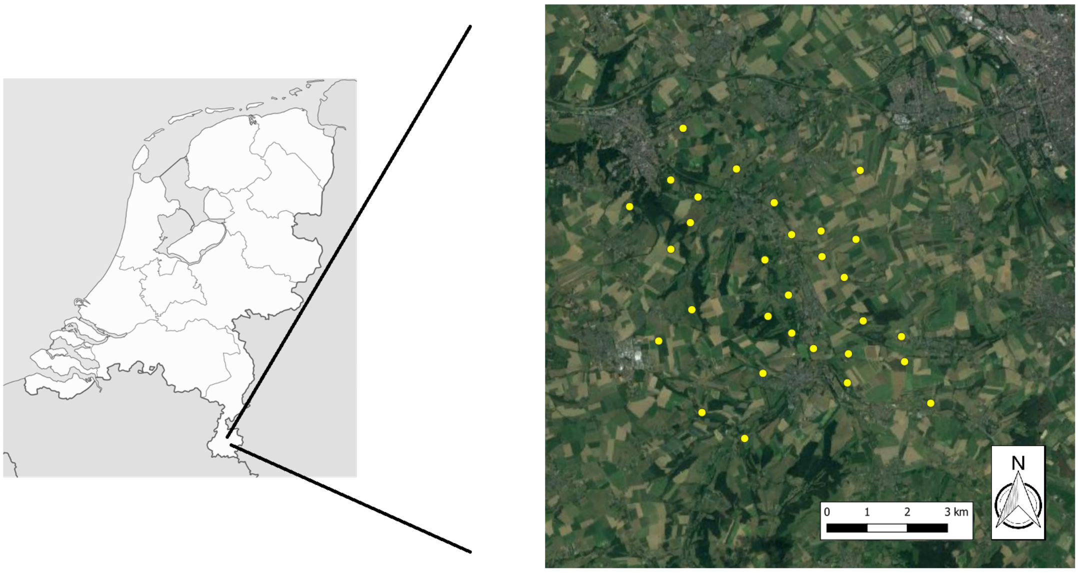

Figure 1: The study area located in the Southeast of the Netherlands. The 30 plots are indicated by yellow dots (Basemap: Google Earth map at August 2022).

2.3.2 Field data collection

A transect of 150 by 1 meter, subdivided in three parts of 50m, was set up in each grassland and was marked by plates clearly visible from drone imagery. The transects were placed from the edge to the center of the grassland, and led across elevational differences within the grassland if present, in order to represent heterogeneity within the grassland. To avoid sampling the same bee populations, adjacent transects were mostly separated by distances of >500 m (minimum 435 m). Studies have shown that although large-bodied bees such as bumblebees can forage at distances of a few kilometer, they mainly forage at short distances (mean distance about 250-550 m), while smaller wild bees forage at yet shorter distances (43). Both bees and flowers were surveyed along each transect. Wild bees as well as the honeybee *Apis mellifera* were counted by transect walks which is a standard method for studying plant-pollinator associations (72). All transects were surveyed by the same two observers who counted all bees up to a meter in front of them while slowly walking along the transect for 15 minutes, excluding handling time of caught specimens. Specimens were identified using keys to the Dutch Apidae (9; 36; 37). Distinct species could be identified in the field, while other specimens were collected and identified in the lab using stereo-microscopes and, in some cases, expert consultation. The counts were performed during May 12th-31st 2021, between 10 a.m. and 17 p.m., and under favourable weather conditions (dry, >50% sunny and at least 15 degrees Celsius with a wind speed <2 Beaufort). Bee data collected in the field were used to derive at plot level, bee abundance (the total number of observed bees), bee species richness (the number of unique species) and the Shannon's H index (39) as an indicator of bee diversity. In each transect, flowers were surveyed generally on the same day as the bee surveys following the method of Scheper et al.(52). For logistical reasons some grasslands were surveyed one or two days before or after the bee survey. The number of open flowers of a given species was counted. Which was then used to calculate flower cover per transect by multiplying counts with species-specific estimates of flower size and summing over all observed flowering plant species (52).

2.3.3 UAV Data Acquisition and Data Processing

The UAV data collection was carried out in parallel with the collection of the field data. The UAV model "DJI Matrice 210 RTK" was used to carry the RGB Zenmuse X5 camera (16.0 MP, 17.3 x 13.0 mm sensor) with an integrated RTK gps. The images were acquired with an overlapping rate of 80% in order to facilitate the creation of the final orthomosaic. All flights were performed at a height of approximately 20m above ground altitude.

We used the Agisoft Metashape Professional Edition software for UAV photogrammetric processing. The software offers a user-friendly workflow that combines algorithms based on structure-for-motion and stereo-matching for image alignment and reconstruction of the 3D image (32). In order to create the orthomosaic of the collected RGB images, 4 procedural stages have been followed: image alignment, dense point cloud assessment, development of the digital elevation model (DEM) and finally the building of the orthomosaic. In the first step, set with "high" accuracy, the software extract features within the images and match them in order to produce a sparse 3D point cloud. In this stage the software automatically detects the precise features of the "Ground Control Points" - extracting the GPS coordinates to each of them. We kept the "high" accuracy also in the building of the dense cloud stage. We used the Metashape default setting for building the DEM and the final orthomosaic that was successively exported as GeoTIFF with the higher spatial resolution (the mean spatial resolution of the 30 considered areas is about 0.5cm) to assess the role of spatial resolution in the estimation of the flower cover, we exported the orthomosaic also with a spatial resolution of 1cm, 2cm and 5cm.

2.3.4 Modelling of UAV image estimates of flower cover

Three common machine learning algorithms included in the caret R package (63) were used to assess the flower cover over the study areas: RF, SVM, NNET.

RF is an effective machine learning method based on a series of decision trees. Its algorithm uses bootstrapping to build a large number of different training subsets based on a randomly selected sample of the training samples and after constructing multiple decision trees voting is used to obtain the final prediction. Each decision tree returns a classification result for the samples not chosen as training samples. The final class prediction is determined by the largest number of votes given by the decision tree for that class (34; 19). SVM is a supervised learning model that uses learning algorithms to examine data for classification and a support vector regression. It is based on a kernel function (that helps to reflect similarity between data points and the cost loss function) in order to convert initial feature space into an N-dimensional space, searching later for a hyperplane to split classes (34). Recently, NNET models have become very popular due to the increase of computational capacity of computers and the higher availability of big data with which these models can be "trained". NNET are based on several layers of processors that are connected to each other attempting to recognize hidden relationships though a process that simulates the human brain (26).

The training and testing data were collected based on the manual interpretation of the orthomosaic images creating 70 polygons for each class (e.g. flowers vs grass/soil). Due to the different polygon sizes, the number of pixels for each class also differed. For this reason, in order to build a more statistically robust and more reliable strategy for classification a total of 750 points were randomly chosen within the polygons per class. Out of them, 70% were used as training sets, and 30% were used as testing sets. The performance accuracy is calculated based on the testing data-set to avoid the problem of overfitting.

We used a 10-fold cross-validation method to estimate the model accuracy (76). After building the confusion matrix, validity and reliability of the selected models was evaluated using the following metrics: Accuracy, Kappa, Precision, Sensitivity, Specificity, Negative Predictive Value , Positive Predictive Value , Balanced Accuracy, F1. Accuracy indicates the percentage of correct predictions. The sum of false negatives and true positives is divided by the total number of predictions. Kappa is a metric that compares an Observed Accuracy with an Expected Accuracy (random chance). It is used not only to evaluate a single classifier, but also to evaluate classifiers amongst themselves. In addition, it takes into account random chance (agreement with a random classifier), which generally means it is less misleading than simply using accuracy as a metric (an Observed Accuracy of 80% is a lot less impressive with an Expected Accuracy of 75% versus an Expected Accuracy of 50%). Precision identifies the accuracy with which the model predicted positive classes (number of true positive divided by the number of all positive results). Sensitivity is defined as the proportion of positive results out of the number of samples which were actually positive while Specificity is measured as the number of correct negative results divided by the total number of negatives. The positive predictive value is defined as the percent of predicted positives that are actually positive while the negative predictive value is defined as the percent of negative positives that are actually negative. The F1 is used to compare the performance of two classifiers, combining the precision and recall of a classifier into a single measure, considering their harmonic mean. Balanced accuracy is the arithmetic mean of sensitivity and specificity, and it is used when one of the target classes appears a lot more than the other (e.g. areas with a lot of grass and less flowers).

2.3.5 Assessment of the relationship between UAV image flower cover estimates and estimates based on *in-situ data*

We examined whether the estimated flower cover using UAV images was significantly correlated with *in-situ* estimated flower cover, bee abundance, species richness and diversity (Shannon's H) using simple linear regression analysis. Next, we compared the fit of these relationships when using different machine learning algorithms and different spatial resolutions of the UAV images.

2.4 Results

The *in-situ* surveys of flowering plants showed that the species with the largest flower cover across all transects were three Ranunculus species (*R. repens*, *R. acris* and *R. bulbosus*) and *Leucanthemum vulgare*, followed by *Trifolium pratense*, *Bellis perennis* and *Taraxacum officinale* (see in Appendix 1 a table with details on the in-situ flower survey).

These seven species represent 84% of the total flower cover recorded and are generally the most widespread flowering forbs in the studied grasslands. Examples of grasslands with contrasting flower cover can be found in Appendix 2.

Flower cover estimated by RGB UAV images was generally significantly positively related to flower cover estimates made in the field by trained observers (Fig. 2a and Appendix 3 for the linear regressions). The goodness of fit of the relationship varied considerably with the algorithm and spatial resolution used, however the accuracy of the relationships generally declined with a decrease in the spatial resolution of the image data. The only exception was the SVM algorithm that had a higher R^2 at 1 cm resolution than at 0.5 cm resolution. The flower cover estimates produced by RF and NNET algorithms both performed similar with generally high R^2 values, while the SVM algorithms performed much worse although it still produced significant relations with field estimated flower cover data. The accuracy, ranged from a robust high 0.8 for RF algorithm using images with a 0.5cm spatial resolution (linear regression shown in Fig. 5) to 0.31 for the SVM algorithm using images with a 5cm spatial resolution.

Bee abundance, species richness and diversity were also significantly and positively related to flower cover estimated by the RGB UAV images (Fig. 2b-c-d). The goodness of fit for the bee variables followed patterns that were very similar to those for flower cover with the R^2 generally declining with decreasing spatial resolution and RF and NNET algorithm performing considerably better than SVM algorithms. Interestingly, goodness of fit of the best models that used RGB UAV estimates (RF with images at 0.5cm spatial resolution) were considerably higher than those of the models that used flower cover estimates by field observers (Fig. 4). The best relationships with flower cover estimated from UAV RGB had R^2's of 0.65, 0.62 and 0.54 for bee abundance, bee species richness and diversity respectively.

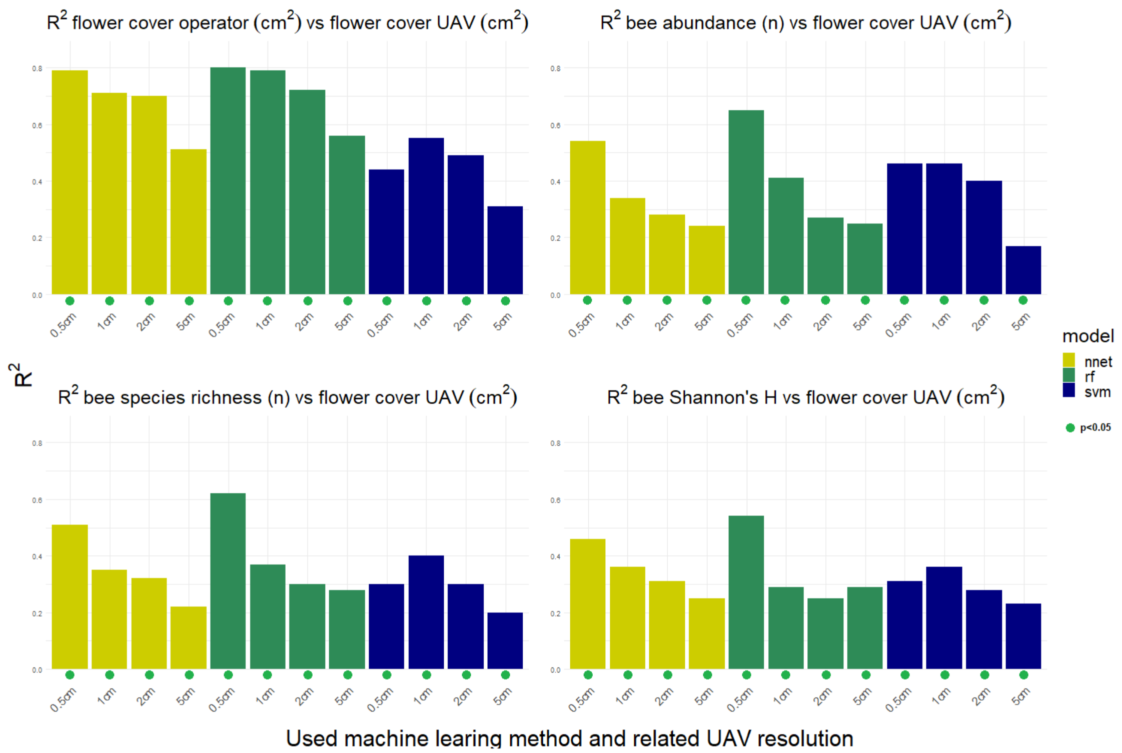

Figure 2: R^2 values derived from the linear regression between the field data (flower cover, bee abundance, bee species richness and bee Shannon's H diversity) and the flower cover estimated by RGB UAV images at different spatial resolution (0.5cm, 1cm, 2cm, 5cm) using different machine learning methods. Green dots show statistically significant ($p<0.05$) correlations.

Figure 3 shows as an example the visual flower cover estimation by the RF machine learning method at the different spatial resolutions (0.5cm, 1cm, 2cm, 5cm) in one of the 30 study areas (the background image for the four sub-plots is at 0.5cm resolution). It illustrates why the accuracy in the flower cover estimation is generally higher in images having the higher spatial resolution (0.5-1cm). With lower spatial resolution of the images (e.g. from 0.5 to 2-5cm) more pixels classified as flowers actually contain a mixture of flowers and green vegetation which leads to an overestimation of the flower cover.

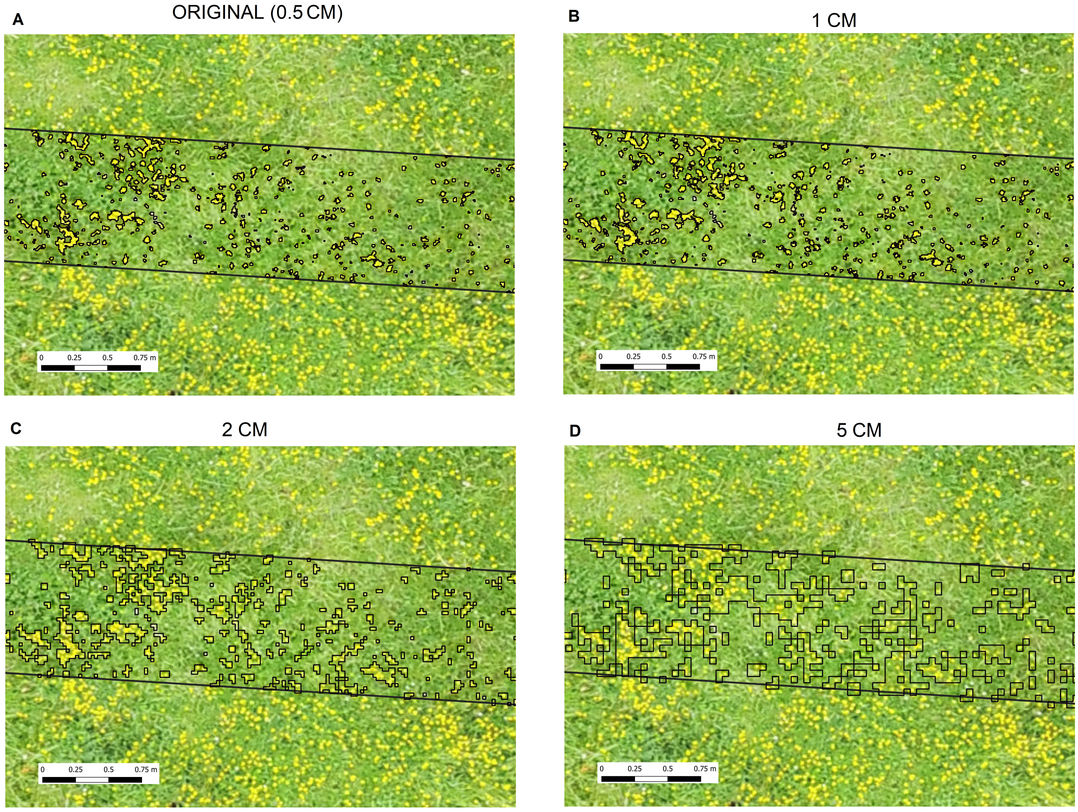

Figure 3: A visualisation of the results of the flower cover estimation by the RF machine learning methods at the different spatial resolution in one of the 30 study sites. The background image for the four sub-plots is at 0.5cm resolution.

The RGB UAV flower cover estimates of the best model was strongly positively related to flower cover as estimated by the observer in the field (Fig. 5). Interestingly, the goodness of fit of some models (RF in particular) that used RGB UAV (at 0.5cm) estimates to predict bee abundance, richness and diversity were higher than those of the models that used flower cover estimates by field observers (Fig. 4).

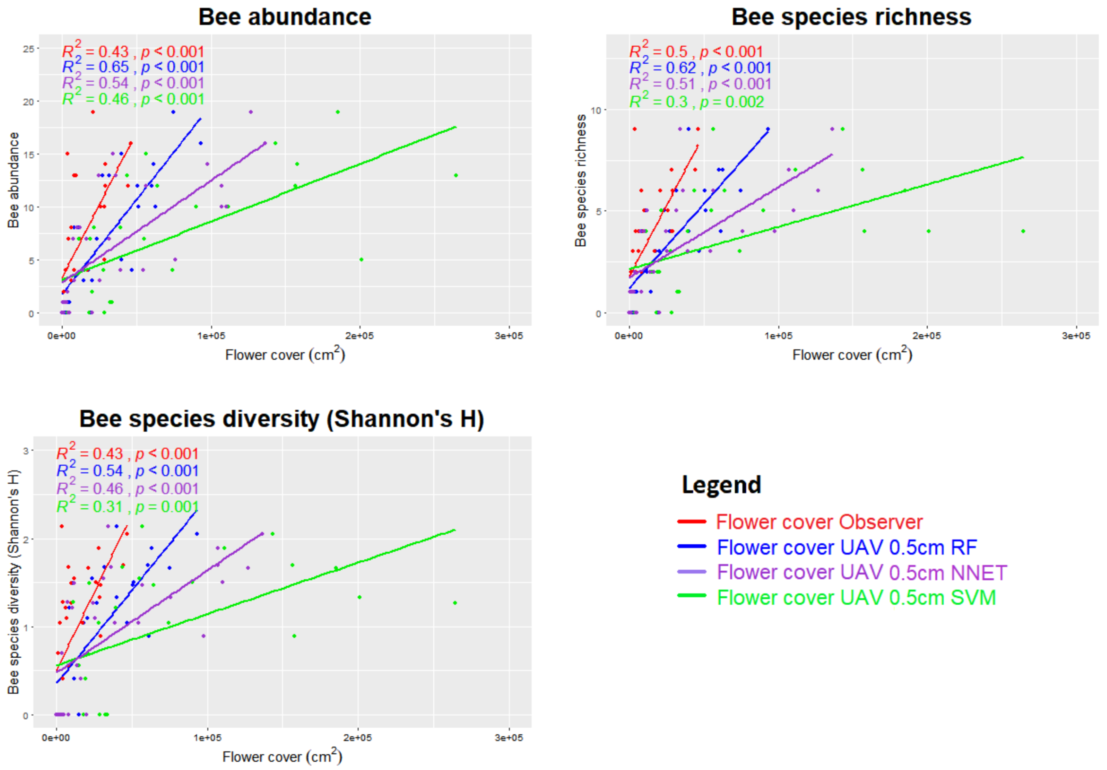

Figure 4: The comparison of the relationships between bee abundance, species richness and diversity (Shannon's H) and the flower cover estimated by the *in-situ* observations (Flower cover observer cm^2) and by the different machine learning models using RGB UAV images at the higher spatial resolution (0.5cm).

Figure 5 shows the relationship between the flower cover estimated by the *in-situ* observations and the flower cover derived from the best machine learning UAV model (RF 0.5cm). The relationship is positive and significant, with a R^2 value of 0.8.

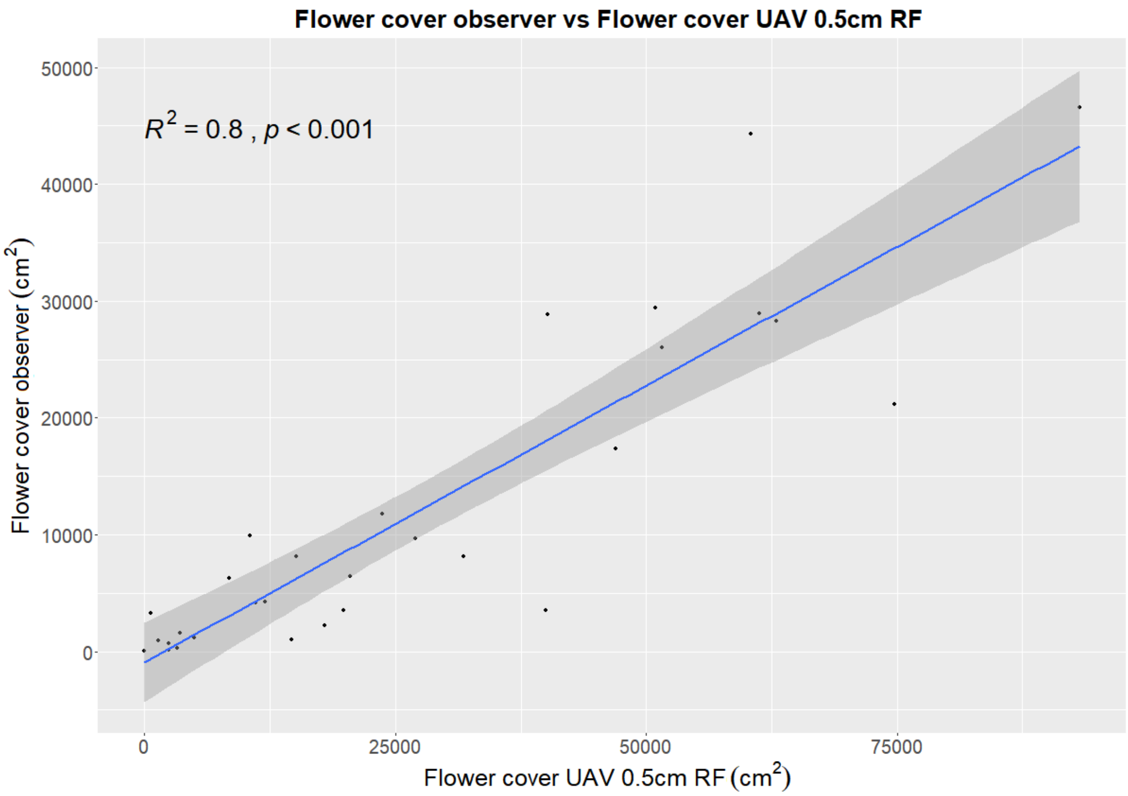

Figure 5: The relationship between the *in-situ* flower cover (Flower cover observer cm^2) and the flower cover estimated by the best machine learning UAV model (RF 0.5cm).

Figure 6 shows the performance of each machine learning model in the estimation of the flower cover using the images at different spatial resolution. The performance of the models (with the exception of the SVM model used with RGB images at 5cm of spatial resolution) at the different spatial resolution is high, ranging for all the parameters between 0.8 and 1.

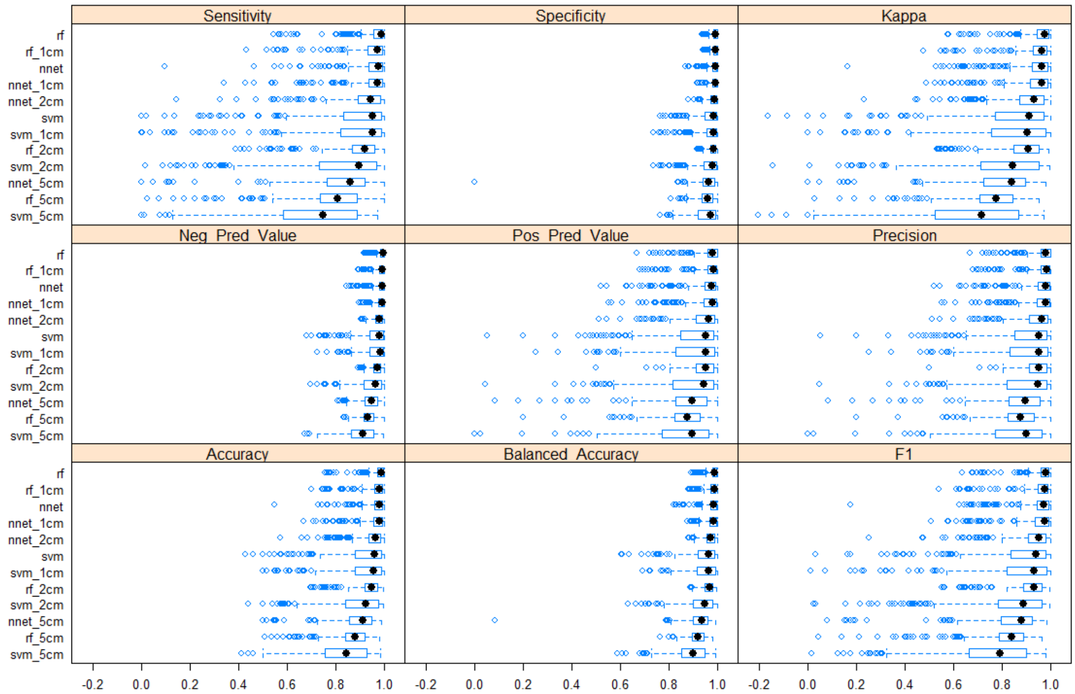

Figure 6: Performance of the machine learning models (Sensitivity, Specificity, Neg.Pred:VAlue, Pos Pred Value, Precision, Recall, Accuracy, Balanced Accuracy, F1, Kappa) at the different spatial resolution.

2.5 Discussion

In this paper we presented a novel approach for estimating flower cover, as an indicator of bee abundance and diversity in grassland ecosystems by using UAV RGB images and different machine learning methods. We find highly significant positive relationships between flower cover estimates obtained through UAV RGB images and machine learning algorithms and flower cover estimates obtained the traditional way, *in-situ* by local observers. We also find reliable relationships between UAV RGB image obtained flower estimates and *in-situ* bee abundance and diversity. This represents a proof of concept that imagery from UAVs can be used to reliably assess an important aspect of grassland quality for bees.

Overall, RF and NNET models trained with high-resolution images (<1cm) showed the highest R^2 between the predicted flower cover estimated and the *in-situ* observed flower cover, bee abundance, richness and Shannon's H. Goodness of fit metrics showed also a similar behaviour, with the highest scores observed for RF and NNET models trained with high-resolution images (<1cm). SVM showed generally lower scores of R^2 and goodness of fit metrics compared to the other two approaches. Notwithstanding the good performances of the RF and NNET models, we argue that finding the absolute best machine-learning methods could be a challenging quest, as their results are heavily influenced by the quality of the training datasets (see for instance the class overlap and class imbalance issues), their task (classification *vs* regression) and the tuning of the models' hyperparameters ((29), Chapt. 11.5.2), which we decided to keep as default. Instead, one of the main outcomes of this study is the influence of the spatial resolution of the optical RGB UAV images on the flower cover estimates. Our results showed that, the higher the spatial resolution of the RGB UAV images for the assessment of flower cover, the higher the accuracy (through the goodness of fit) with the *in-situ* flower cover and with the bee abundance, diversity and richness. This is probably largely the result of the fact that images with higher spatial resolution are better able to differentiate between flowers, grass or soil, offering higher information details and reducing the mixed pixel issue (18). Images with coarse spatial resolution result in mixed signal at pixel scale, integrating the spectral signature of various vegetative organisms (e.g. trees, shrubs, flowers), homogenizing the signal and causing difficulties to clearly identify boundaries between spatial entities (individuals, vegetation types, ecosystem types) (46; 57; 10). In our study, the weighted average flower size was 1.66 cm^2. Pixel sizes of up to 3.2 cm^2 (i.e. our two most detailed spatial resolution classes) would classify a pixel correctly if a full single flower would be in it, with larger pixels misclassifying. In our study, the mixed pixel issue most likely also led to an overestimation of the range in UAV RGB estimated flower cover because pixels with less than 50% flower cover are classified as 0% flower cover while pixels with more than 50% flower cover are classified as 100% flower cover. In flower-poor fields there will be disproportionately more pixels with less than 50% flower cover while in flower-rich fields there will be disproportionately more pixels with more than 50% flower cover. This probably explains why, in the case of the results shown in Figure 5, the range in RGB UAV estimated flower cover is twice as wide as the range in the *in-situ* collected flower cover data. A consequence of this is that care should be taken to use the relations from UAV RGB estimates of flower cover to predict effects of changing flower cover on bees. Because the slope of the relationship with bees is twice as low for RGB UAV flower cover compared to *in-situ* flower cover (2.3 times for bee abundance, and 2.1 times for bee species richness and diversity), the predicted effects on bees of improving flower cover will be twice as low as well.

With very few exceptions the accuracy of the relationships with *in-situ* flower estimates were higher than those with *in-situ* bee estimates. This is to be expected because the algorithms we used were trained to classify flower presence (Fig.2, sub-plot A). Bees are too small and inconspicuous to be detected by UAV RGB imagery and their abundance and species richness can only be inferred using such images because they generally show a very strong relationship with flower availability (61; 41). Previous research shows that abundance and species richness of bees are often strongly related to the cover and species richness of flowers. Bee species richness is generally best predicted by the floral species richness (41; 11; 35) while bee abundance by flower cover (3; 53). However, flower species richness and cover are mostly correlated and flower species richness has also been found to be related to bee abundance (60; 64) while flower cover has been shown to be related to bee species richness (8; 52). Which relationship predominates at a particular place and time is probably largely determined by the composition of both the local bee and the flowering plant community. Additional factors such as the quantity and quality of the reward offered by different flowers, the availability of nesting sites and landscape composition may additionally affect local bee abundance which probably explains the weaker correlation with UAV RGB imagery. It is perhaps surprising that the accuracies of the best models explaining bee variables that used RGB UAV estimates (RF with images at 0.5cm spatial resolution) were considerably higher than those of the models that used flower cover estimates by field observers (Fig. 4). The goodness of fit of the observed relationships between our *in-situ* collected flower and bee data varied between 0.43 and 0.5. This is not particularly low and in line with results from other empirical studies (e.g. 0.12-0.52 in Potts et al.(41)). Whether the UAV RGB image based estimates or observer based estimates better reflect reality is impossible to say, but the higher goodness-of-fit suggests that UAV RGB images capture some variation in flower cover that is relevant to bees and that was missed by our observers.

Our results contrast a bit with findings from other studies that attempted to relate flower estimates based on UAV RGB imagery to flower estimates from field observations and that show more mixed results (7). Smigaj and Gaulton (58) made use of RGB and multispectral images for the flower abundance estimation in a farmland ecosystem in northeast England. The authors found RGB imagery to be poorly related to *in-situ* estimates while multispectral images performed markedly better. This was explained by the fact that the multispectral images included the near infrared band that helped with the differentiation between flowers and woody parts of hedgerows (58). Also De Sá et al.

(7) found a poor correlation between the estimation of the number of flowers of an invasive shrub using RGB UAV images and field observations explaining this with the variability of the structure and phenology of the examined invasive plants and the variation of the different habitats where the species grows. An explanation for the good performance of the RGB images in our case might therefore be due to the relatively uniform herbaceous vegetation that we surveyed that lack components that have a reflectance similar to the flowers (identifiable with the red, green, and blue bands). Additional explanations could be that we collected our field data mostly on the same day as the UAV RGB images were made, which minimizes errors caused by the phenology of the flowering plants and the fact that the most dominant flowering plant species (e.g. *Ranunculus spp.*, *Taraxacum officinale*, *Leucanthemum vulgare*, *Bellis perennis*) had horizontal inflorescences that can be observed relatively accurately from aerial imagery.

The method presented in this paper opens up a wide range of uses beyond complementing the manual area estimation of flowers in grassland. Remote sensing approaches can increase the temporal and spatial efficiency of sampling methods, allowing large portions of land to be surveyed in less time, creating standardized and reproducible data. Optical UAV data could represent a potential answer to the lack of baseline ecological data for plant-pollinator interactions, as pointed out by the International Pollinator Initiative (73). UAV based methods could represent promising tools for monitoring habitat quality for bee pollinator communities because they are now affordable (15) and allow different operators (e.g., researchers, farmers, ecologists) to obtain high spatial resolution data from different sensors that can be carried simultaneously, covering large areas in a limited time. In time, they can be used with an "on-demand" approach that allows to capture specific stages of vegetation phenology (flowering time) in particular in areas characterized by high cloud cover (33; 7). However, a number of challenges need to be addressed before any UAV based methods can be deployed routinely and at large spatial scales. As stressed in this study, the spatial and spectral resolution of the UAV images together with the choice of the appropriate ML algorithm, play a fundamental role in the characterization of the vegetation. Further limitations include the high amount of time needed to collect and to process the data, especially when using imagery and dense point clouds at the millimeter level (71). Topography, meteorological condition and the right time-window can also influence negatively the image acquisition and the successively correlation analysis. When applying the methodology explained in our study, it must be clear whether relationships determined in one area or year can be generalized to other areas and/or years (4). If UAV based images indicate high flower cover, does that mean that a site is suitable for bees throughout the growing season or just during a few days or weeks before or after the day the image was obtained. Furthermore, we need to know whether UAV based images can also be used to survey flower cover and bee habitat quality in structurally more complex vegetation types and other colors of flowering plant species.

2.6 Conclusions

This paper shows proof-of-concept that flower cover, an important aspect of habitat quality for bees, in grasslands can be monitored using UAV RGB imagery which opens the way for large-scale, standardized and cost-effective monitoring strategies in one of the most widespread vegetation types globally. The results suggests that the higher the spatial resolution of the optical images, the higher the accuracy of the approach. Furthermore, the RF machine learning algorithms generally produced the most accurate relationships. The next steps will need to focus on operationalizing the approach, testing and validating with an independent data-set the proposed approach examining its reproducibility. Further analysis will focus on assessing the trade-off between UAV flight height and precision in order to understand how high the drone can fly in order to increase the flight area without losing accuracy in the assessment of flower cover. Through structure-from-motion technology UAV images could be used to characterize the 3D structure of different habitats in order assess how the habitat structure influences the bee diversity and abundance. Finally, an attempt could be made to standardize the method in order to obtain imagery from large areas and translate them automatically in 'flower cover data' and successively in bee information.

2.7 Author Contributions

L.C. conceived the study together with D.R. and M.T.; contributed to the design of the field surveys; performed data analysis and machine-learning modelling; developed the computational workflow; interpreted the results; and led the writing of the manuscript. M.T., D.K., J.P.R.d.V., H.B., V.M., D.D.R., and R.C.G. contributed to field data collection, ecological interpretation, and manuscript revision. D.R. supervised the study and contributed to conceptualization and editing.

2.8 Acknowledgments

We thank both the drone pilots, Berry Onderstal and Peter van der Zee, for their work in the field data campaigns.

We thank the Department of Innovation, Research, University and Museums of the Autonomous Province of Bozen/Bolzano for covering the Open Access publication costs.

2.9 Funding

This research received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 862480 (SHOWCASE). DDR is supported by the FRS-FNRS ASP Belgian grant.

References

[1] Alvarez-Taboada, F., Paredes, C., and Julián-Pelaz, J. (2017). Mapping of the invasive species hakea sericea using unmanned aerial vehicle (uav) and worldview-2 imagery and an object-oriented approach. *Remote Sensing*, 9(9):913.

[2] Biesmeijer, J. C., Roberts, S. P., Reemer, M., Ohlemuller, R., Edwards, M., Peeters, T., Schaffers, A., Potts, S. G., Kleukers, R., Thomas, C., et al. (2006). Parallel declines in pollinators and insect-pollinated plants in britain and the netherlands. *Science*, 313(5785):351–354.

[3] Blaauw, B. R. and Isaacs, R. (2014). Flower plantings increase wild bee abundance and the pollination services provided to a pollination-dependent crop. *Journal of Applied Ecology*, 51(4):890–898.

[4] Blasi, M., Bartomeus, I., Bommarco, R., Gagic, V., Garratt, M., Holzschuh, A., Kleijn, D., Lindström, S. A., Olsson, P., Polce, C., et al. (2021). Evaluating predictive performance of statistical models explaining wild bee abundance in a mass-flowering crop. *Ecography*, 44(4):525–536.

[5] Capolupo, A., Kooistra, L., Berendonk, C., Boccia, L., and Suomalainen, J. (2015). Estimating plant traits of grasslands from uav-acquired hyperspectral images: a

comparison of statistical approaches. *ISPRS International Journal of Geo-Information*, 4(4):2792–2820.

[6] Cruzan, M. B., Weinstein, B. G., Grasty, M. R., Kohrn, B. F., Hendrickson, E. C., Arredondo, T. M., and Thompson, P. G. (2016). Small unmanned aerial vehicles (micro-uavs, drones) in plant ecology. *Applications in plant sciences*, 4(9):1600041.

[7] De Sa, N. C., Castro, P., Carvalho, S., Marchante, E., López-Núñez, F. A., and Marchante, H. (2018). Mapping the flowering of an invasive plant using unmanned aerial vehicles: is there potential for biocontrol monitoring? *Frontiers in plant science*, 9:293.

[8] Ebeling, A., Klein, A.-M., Schumacher, J., Weisser, W. W., and Tscharntke, T. (2008). How does plant richness affect pollinator richness and temporal stability of flower visits? *Oikos*, 117(12):1808–1815.

[9] Falk, S. and Lewington, R. (2017). *Veldgids bijen voor Nederland en Vlaanderen.*

[10] Feilhauer, H., Zlinszky, A., Kania, A., Foody, G. M., Doktor, D., Lausch, A., and Schmidtlein, S. (2021). Let your maps be fuzzy!—class probabilities and floristic gradients as alternatives to crisp mapping for remote sensing of vegetation. *Remote Sensing in Ecology and Conservation*, 7(2):292–305.

[11] Fründ, J., Linsenmair, K. E., and Blüthgen, N. (2010). Pollinator diversity and specialization in relation to flower diversity. *Oikos*, 119(10):1581–1590.

[12] Gallai, N., Salles, J.-M., Settele, J., and Vaissière, B. E. (2009). Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. *Ecological economics*, 68(3):810–821.

[13] Gallmann, J., Schüpbach, B., Jacot, K., Albrecht, M., Winizki, J., Kirchgessner, N., and Aasen, H. (2021). Flower mapping in grasslands with drones and deep learning. *Frontiers in plant science*, 12.

[14] Ghazoul, J. (2005). Buzziness as usual? questioning the global pollination crisis. *Trends in ecology & evolution*, 20(7):367–373.

[15] Gonzales, D., Hempel de Ibarra, N., and Anderson, K. (2022). Remote sensing of floral resources for pollinators–new horizons from satellites to drones. *Frontiers in Ecology and Evolution*, 10.

[16] Guo, Q., Wu, F., Hu, T., Chen, L., Liu, J., Zhao, X., Gao, S., Pang, S., et al. (2016). Perspectives and prospects of unmanned aerial vehicle in remote sensing monitoring of biodiversity. *Biodiversity Science*, 24(11):1267–1278.

[17] Guo, Y., Yin, G., Sun, H., Wang, H., Chen, S., Senthilnath, J., Wang, J., and Fu, Y. (2020). Scaling effects on chlorophyll content estimations with rgb camera mounted on a uav platform using machine-learning methods. *Sensors*, 20(18):5130.

[18] Hu, P., Chapman, S. C., and Zheng, B. (2021). Coupling of machine learning methods to improve estimation of ground coverage from unmanned aerial vehicle (uav) imagery for high-throughput phenotyping of crops. *Functional Plant Biology*, 48(8):766–779.

[19] Immitzer, M., Atzberger, C., and Koukal, T. (2012). Tree species classification with random forest using very high spatial resolution 8-band worldview-2 satellite data. *Remote sensing*, 4(9):2661–2693.

[20] Jin, X., Liu, S., Baret, F., Hemerlé, M., and Comar, A. (2017). Estimates of plant density of wheat crops at emergence from very low altitude uav imagery. *Remote Sensing of Environment*, 198:105–114.

[21] Kaneko, K., Nohara, S., et al. (2014). Review of effective vegetation mapping using the uav (unmanned aerial vehicle) method. *Journal of Geographic Information System*, 6(06):733.

[22] Kattenborn, T., Eichel, J., Wiser, S., Burrows, L., Fassnacht, F. E., and Schmidtlein, S. (2020). Convolutional neural networks accurately predict cover fractions of plant species and communities in unmanned aerial vehicle imagery. *Remote Sensing in Ecology and Conservation*, 6(4):472–486.

[23] Kleijn, D., Kohler, F., Báldi, A., Batáry, P., Concepción, E., Clough, Y., Díaz, M., Gabriel, D., Holzschuh, A., Knop, E., et al. (2009). On the relationship between farmland biodiversity and land-use intensity in europe. *Proceedings of the royal society B: biological sciences*, 276(1658):903–909.

[24] Kleijn, D., Winfree, R., Bartomeus, I., Carvalheiro, L. G., Henry, M., Isaacs, R., Klein, A.-M., Kremen, C., M'gonigle, L. K., Rader, R., et al. (2015). Delivery of crop pollination services is an insufficient argument for wild pollinator conservation. *Nature communications*, 6(1):1–9.

[25] Klein, A.-M., Vaissière, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., and Tscharntke, T. (2007). Importance of pollinators in changing landscapes for world crops. *Proceedings of the royal society B: biological sciences*, 274(1608):303–313.

[26] Kreig, J. A., Parish, E., and Jager, H. I. (2021). Growing grasses in unprofitable areas of us midwest croplands could increase species richness. *Biological Conservation*, 261:109289.

[27] Kremen, C., Williams, N. M., and Thorp, R. W. (2002). Crop pollination from native bees at risk from agricultural intensification. *Proceedings of the National Academy of Sciences*, 99(26):16812–16816.

[28] Kwak, G.-H. and Park, N.-W. (2019). Impact of texture information on crop classification with machine learning and uav images. *Applied Sciences*, 9(4):643.

[29] Lovelace, R., Nowosad, J., and Muenchow, J. (2019). *Geocomputation with R*. CRC Press.

[30] Magnússon, R. Í., Limpens, J., Kleijn, D., van Huissteden, K., Maximov, T. C., Lobry, S., and Heijmans, M. M. (2021). Shrub decline and expansion of wetland vegetation revealed by very high resolution land cover change detection in the siberian lowland tundra. *Science of the Total Environment*, 782:146877.

[31] Michele, T., Duccio, R., Marc, Z., Ruth, S., and Giustino, T. (2018). Testing the spectral variation hypothesis by using the rao-q index to estimate forest biodiversity: Effect of spatial resolution. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 1183–1186. IEEE.

[32] Moe, K. T., Owari, T., Furuya, N., and Hiroshima, T. (2020). Comparing individual tree height information derived from field surveys, lidar and uav-dap for high-value timber species in northern japan. *Forests*, 11(2):223.

[33] Müllerová, J., Brna, J., Bartaloš, T., Dvořák, P., Vítková, M., and Pyšek, P. (2017). Timing is important: Unmanned aircraft vs. satellite imagery in plant invasion monitoring. *Frontiers in Plant Science*, 8:887.

[34] Nasiri, V., Darvishsefat, A. A., Arefi, H., Griess, V. C., Sadeghi, S. M. M., and Borz, S. A. (2022). Modeling forest canopy cover: A synergistic use of sentinel-2, aerial photogrammetry data, and machine learning. *Remote Sensing*, 14(6):1453.

[35] Neumüller, U., Burger, H., Schwenninger, H. R., Hopfenmüller, S., Krausch, S., Weiß, K., and Ayasse, M. (2021). Prolonged blooming season of flower plantings increases wild bee abundance and richness in agricultural landscapes. *Biodiversity and Conservation*, 30:3003–3021.

[36] Nieuwenhuijsen, H. and Peeters, T. (2015). *Nederlandse bijen op naam brengen. Deel 1. - Stichting Jeugdbondsuitgeverij, 's Graveland.*

[37] Nieuwenhuijsen, H., Peeters, T., and Dijkshoorn, D. (2020). *Nederlandse bijen op naam brengen. Deel 2. - Stichting Jeugdbondsuitgeverij, 's Graveland.*

[38] Ollerton, J., Winfree, R., and Tarrant, S. (2011). How many flowering plants are pollinated by animals? *Oikos*, 120(3):321–326.

[39] Pielou, E. C. (1966). Shannon's formula as a measure of specific diversity: its use and misuse. *The American Naturalist*, 100(914):463–465.

[40] Potts, S., Dauber, J., Hochkirch, A., Oteman, B., Roy, D., Ahrne, K., Biesmeijer, K., Breeze, T., Carvell, C., Ferreira, C., et al. (2021). Proposal for an eu pollinator monitoring scheme. *Publications Office of the European Union.*

[41] Potts, S. G., Vulliamy, B., Dafni, A., Ne'eman, G., and Willmer, P. (2003). Linking bees and flowers: how do floral communities structure pollinator communities? *Ecology*, 84(10):2628–2642.

[42] Ranđelović, P., Đorđević, V., Milić, S., Balešević-Tubić, S., Petrović, K., Miladinović, J., and Đukić, V. (2020). Prediction of soybean plant density using a machine learning model and vegetation indices extracted from rgb images taken with a uav. *Agronomy*, 10(8):1108.

[43] Redhead, J. W., Dreier, S., Bourke, A. F., Heard, M. S., Jordan, W. C., Sumner, S., Wang, J., and Carvell, C. (2016). Effects of habitat composition and landscape structure on worker foraging distances of five bumble bee species. *Ecological Applications*, 26(3):726–739.

[44] Reed, B. C., Schwartz, M. D., and Xiao, X. (2009). Remote sensing phenology. In *Phenology of ecosystem processes*, pages 231–246. Springer.

[45] Robinson, R. A. and Sutherland, W. J. (2002). Post-war changes in arable farming and biodiversity in great britain. *Journal of applied Ecology*, 39(1):157–176.

[46] Rocchini, D., Balkenhol, N., Carter, G. A., Foody, G. M., Gillespie, T. W., He, K. S., Kark, S., Levin, N., Lucas, K., Luoto, M., et al. (2010). Remotely sensed spectral heterogeneity as a proxy of species diversity: recent advances and open challenges. *Ecological Informatics*, 5(5):318–329.

[47] Rocchini, D., Santos, M. J., Ustin, S. L., Féret, J.-B., Asner, G. P., Beierkuhnlein, C., Dalponte, M., Feilhauer, H., Foody, G. M., Geller, G. N., et al. (2022a). The spectral species concept in living color. *Journal of Geophysical Research: Biogeosciences*, 127(9):e2022JG007026.

[48] Rocchini, D., Thouverai, E., Marcantonio, M., Iannacito, M., Da Re, D., Torresani, M., Bacaro, G., Bazzichetto, M., Bernardi, A., Foody, G. M., et al. (2021). rasterdiv—an information theory tailored r package for measuring ecosystem heterogeneity from space: To the origin and back. *Methods in ecology and evolution*, 12(6):1093–1102.

[49] Rocchini, D., Torresani, M., Beierkuhnlein, C., Feoli, E., Foody, G. M., Lenoir, J., Malavasi, M., Moudrỳ, V., Šímová, P., and Ricotta, C. (2022b). Double down on remote sensing for biodiversity estimation: a biological mindset. *Community Ecology*, 23(3):267–276.

[50] Sandino, J., Gonzalez, F., Mengersen, K., and Gaston, K. J. (2018). Uavs and machine learning revolutionising invasive grass and vegetation surveys in remote arid lands. *Sensors*, 18(2):605.

[51] Saunders, D. A., Hobbs, R. J., and Margules, C. R. (1991). Biological consequences of ecosystem fragmentation: a review. *Conservation biology*, 5(1):18–32.

[52] Scheper, J., Bommarco, R., Holzschuh, A., Potts, S. G., Riedinger, V., Roberts, S. P., Rundlöf, M., Smith, H. G., Steffan-Dewenter, I., Wickens, J. B., et al. (2015). Local and landscape-level floral resources explain effects of wildflower strips on wild bees across four european countries. *Journal of Applied Ecology*, 52(5):1165–1175.

[53] Scheper, J., Bukovinszky, T., Huigens, M. E., and Kleijn, D. (2021). Attractiveness of sown wildflower strips to flower-visiting insects depends on seed mixture and establishment success. *Basic and Applied Ecology*, 56:401–415.

[54] Scheper, J., Reemer, M., van Kats, R., Ozinga, W. A., van der Linden, G. T., Schaminée, J. H., Siepel, H., and Kleijn, D. (2014). Museum specimens reveal loss of pollen host plants as key factor driving wild bee decline in the netherlands. *Proceedings of the National Academy of Sciences*, 111(49):17552–17557.

[55] Sheeren, D., Fauvel, M., Josipović, V., Lopes, M., Planque, C., Willm, J., and Dejoux, J.-F. (2016). Tree species classification in temperate forests using formosat-2 satellite image time series. *Remote Sensing*, 8(9):734.

[56] Shirai, H., Tung, N. D. M., Kageyama, Y., Ishizawa, C., Nagamoto, D., Abe, K., Kojima, T., and Akisawa, M. (2020). Estimation of the number of convallaria keiskei's colonies using uav images based on a convolutional neural network. *IEEJ Transactions on Electrical and Electronic Engineering*, 15(10):1552–1554.

[57] Šímová, P., Moudrỳ, V., Komárek, J., Hrach, K., and Fortin, M.-J. (2019). Fine scale waterbody data improve prediction of waterbird occurrence despite coarse species data. *Ecography*, 42(3):511–520.

[58] Smigaj, M. and Gaulton, R. (2021). Capturing hedgerow structure and flowering abundance with uav remote sensing. *Remote Sensing in Ecology and Conservation*, 7(3):521–533.

[59] Steffan-Dewenter, I., Potts, S. G., and Packer, L. (2005). Pollinator diversity and crop pollination services are at risk. *Trends in ecology & evolution*, 20(12):651–652.

[60] Steffan-Dewenter, I. and Tscharntke, T. (2001). Succession of bee communities on fallows. *Ecography*, 24(1):83–93.

[61] Sutter, L., Jeanneret, P., Bartual, A. M., Bocci, G., and Albrecht, M. (2017). Enhancing plant diversity in agricultural landscapes promotes both rare bees and dominant crop-pollinating bees through complementary increase in key floral resources. *Journal of Applied Ecology*, 54(6):1856–1864.

[62] Tamburlin, D., Torresani, M., Tomelleri, E., Tonon, G., and Rocchini, D. (2021). Testing the height variation hypothesis with the r rasterdiv package for tree species diversity estimation. *Remote Sensing*, 13(18):3569.

[63] Team, R. C. (2014). R: A language and environment for statistical computing. r foundation for statistical computing, vienna, austria. 2013.

[64] Theodorou, P., Herbst, S.-C., Kahnt, B., Landaverde-Gonzalez, P., Baltz, L. M., Osterman, J., and Paxton, R. J. (2020). Urban fragmentation leads to lower floral diversity, with knock-on impacts on bee biodiversity. *Scientific Reports*, 10(1):1–11.

[65] Thouverai, E., Marcantonio, M., Lenoir, J., Galfré, M., Marchetto, E., Bacaro, G., Gatti, R. C., Da Re, D., Di Musciano, M., Furrer, R., et al. (2023). Integrals of life: Tracking ecosystem spatial heterogeneity from space through the area under the curve of the parametric rao's q index. *Ecological Complexity*, 52:101029.

[66] Torresani, M., Feilhauer, H., Rocchini, D., Féret, J.-B., Zebisch, M., and Tonon, G. (2021). Which optical traits enable an estimation of tree species diversity based on the spectral variation hypothesis? *Applied Vegetation Science*, 24(2):e12586.

[67] Torresani, M., Masiello, G., Vendrame, N., Gerosa, G., Falocchi, M., Tomelleri, E., Serio, C., Rocchini, D., and Zardi, D. (2022). Correlation analysis of evapotranspiration, emissivity contrast and water deficit indices: A case study in four eddy covariance sites in italy with different environmental habitats. *Land*, 11(11):1903.

[68] Torresani, M., Rocchini, D., Sonnenschein, R., Zebisch, M., Hauffe, H. C., Heym, M., Pretzsch, H., and Tonon, G. (2020). Height variation hypothesis: A new approach for estimating forest species diversity with chm lidar data. *Ecological Indicators*, 117:106520.

[69] van Lier, O. R., Fournier, R. A., Bradley, R. L., and Thiffault, N. (2009). A multi-resolution satellite imagery approach for large area mapping of ericaceous shrubs in northern quebec, canada. *International Journal of Applied Earth Observation and Geoinformation*, 11(5):334–343.

[70] van Swaay, C. A., Nowicki, P., Settele, J., and Van Strien, A. J. (2008). Butterfly monitoring in europe: methods, applications and perspectives. *Biodiversity and Conservation*, 17:3455–3469.

[71] Vanbrabant, Y., Delalieux, S., Tits, L., Pauly, K., Vandermaesen, J., and Somers, B. (2020). Pear flower cluster quantification using rgb drone imagery. *Agronomy*, 10(3):407.

[72] Westphal, C., Bommarco, R., Carré, G., Lamborn, E., Morison, N., Petanidou, T., Potts, S. G., Roberts, S. P., Szentgyörgyi, H., Tscheulin, T., et al. (2008). Measuring bee diversity in different european habitats and biogeographical regions. *Ecological monographs*, 78(4):653–671.

[73] Williams, I. H. (2003). The convention on biological diversity adopts the international pollinator initiative. *Bee World*, 84(1):27–31.

[74] Williams, P. H. and Osborne, J. L. (2009). Bumblebee vulnerability and conservation world-wide. *Apidologie*, 40(3):367–387.

[75] Wratten, S. D., Gillespie, M., Decourtye, A., Mader, E., and Desneux, N. (2012). Pollinator habitat enhancement: benefits to other ecosystem services. *Agriculture, Ecosystems & Environment*, 159:112–122.

[76] Yu, R., Luo, Y., Zhou, Q., Zhang, X., Wu, D., and Ren, L. (2021). A machine learning algorithm to detect pine wilt disease using uav-based hyperspectral imagery and lidar data at the tree level. *International Journal of Applied Earth Observation and Geoinformation*, 101:102363.

3 Chapter II

Machine learning for biodiversity: UAV-based flower detection as an indirect proxy for bee abundance

Ludovico Chieffallo[1*], Michele Torresani[2], Piero Zannini[1], Jan Peter Reinier de Vries[4], Marharyta Blaha[4,5], Alessio Monacchia[1,5], David Kleijn[6], Duccio Rocchini[1,3]

[1] BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum University of Bologna, via Irnerio 42, 40126 Bologna, Italy

[2] Free University of Bolzano/Bozen, Faculty of Agricultural, Environmental and Food Sciences,
Piazza Università / Universitätsplatz 1, 39100 Bolzano/Bozen, Italy

[3] Czech University of Life Sciences Prague, Faculty of Environmental Sciences, Department of Spatial Sciences, Kamýcka 129, Praha–Suchdol 16500, Czech Republic

[4] Center Agriculture Food Environment, University of Trento, S. Michele all'Adige, Italy

[5] Research and Innovation Centre, Fondazione Edmund Mach, S. Michele all'Adige, Italy

[6] Plant Ecology and Nature Conservation Group, Wageningen University, Droevendaalsesteeg 3a, Wageningen 6708PB, The Netherlands

3.1 Abstract

Pollination plays a crucial role in supporting agriculture and ecosystem functioning, making it an essential ecosystem service provided by bees and other insects. However, bee populations are increasingly threatened by habitat fragmentation, intensive agriculture, and climate change, among other threats. Improved monitoring of critical habitat factors, such as flower cover, is crucial to restore pollinator populations. Traditional approaches, such as field analyses, are often time-consuming and expensive, prompting the adoption of alternative methods to achieve greater efficiency and cost-effectiveness. Here, we introduce a novel approach that incorporates machine learning algorithms and optical images obtained from an unoccupied aerial vehicle (UAV). Using machine learning methods on RGB UAV imagery enabled us to estimate flower cover in UAV-monitored areas and make numerical inferences of wild bee pollinator abundance from those estimates.

Unlike our previous study, which relied on separate machine learning models for each study area, our new method develops a single model that can automatically and efficiently recognize flower cover in various grassland ecosystems to successively estimate bee abundance and diversity. In addiction to this main objective, we also sought to determine which machine learning model would perform this important task best. The machine learning models used, particularly the Gradient Boost Machine (GBM), highlighted the capability of UAV RGB images combined with artificial intelligence to predict flower cover over time, which was highly correlated with bee abundance and diversity.

This development represents an additional starting point for the use of machine learning and deep learning techniques in biodiversity studies within AES systems.

3.2 Introduction

Over the decades, global biodiversity has steadily decreased. This undermines a crucial resource that supports ecosystem processes, maintains ecological balance, and underpins a wide range of essential goods and services for human survival (Kleijn et al.,2015; Cardinale et al.,2012).

The most widespread threats to biodiversity are habitat fragmentation and intensive agriculture. Natural habitats have been turned into large monocultures with high fertilizer and pesticide use, drastically reducing the biological diversity of agricultural areas. Industrialization, increased use of pesticides in industry, and the inadequacy of current EU regulations are all contributing to these issues. Additionally, these anthropogenic pressures may exacerbate one another, further threatening bee populations at multiple spatial and temporal scales (Kremen et al.,2002).

Among the insects most vulnerable to such changes, wild bee pollinators are among the most studied (Winfree et al.,2009). Research highlights the catastrophic losses in both the abundance and diversity of bee populations, both nationally and locally (Wratten et al.,2012; Ghazoul,2005; Steffan-Dewenter et al.,2005; Biesmeijer et al.,2006; Williams et al.,2009). Habitat loss, intensive agriculture, climate change, and host plant reduction are among the identified major causes (Scheper et al.,2014). Pollinators contribute to the pollination and production of approximately 76% of the world's leading food crops (Klein et al.,2007) and contribute to seed set in an additional 87% of wild plants globally (Ollerton et al.,2011). Bees are the most economically and ecologically significant group. Overall, one-third of human food production relies directly on this ecosystem service, valued at over 150 billion euros annually worldwide (Gallmann et al.,2022; Gallai et al.,2009).

Despite an increased understanding of the causes of declines and of potential solutions, knowledge about the magnitude of the decline attributed to habitat quality remains limited (Torresani et al.,2023).

Unlike butterflies, for which citizen science monitoring networks have been established in Europe (Van Swaay et al.,2008), bees are notoriously difficult to monitor for species identification. Therefore, most studies of bee abundance occur at the local scale, and larger-scale analyses often rely on distribution data, which tends to be a poor proxy of population trends (Biesmeijer et al.,2006).

Recent European initiatives (EU Pollinators Initiative, EU Pollinator Monitoring Scheme, Strategy for Biodiversity 2030) (Potts et al.,2021) seek to address this problem through standardized monitoring programs. The high cost of field monitoring, largely due to the labour-intensive nature of data collection, the need for taxonomic expertise, and the logistical difficulties of accessing remote or heterogeneous landscapes, will inevitably restrict geographical coverage. Additional cost factors include the need for repeated site visits within a narrow time window to ensure temporal consistency across multiple areas, acquisition and maintenance of specialized sampling equipment, and time and coordination required to train observers and implement standardized protocols at scale.

Floral cover and diversity are indirect methods for estimating bee abundance and richness. This correlation has been confirmed in various studies because bees exclusively feed on nectar and pollen from flowers, and their life cycle depends on these resources

(Potts et al.,2003; Sutter et al.,2017; Segre et al.,2023). This method is substantially faster than counting and identifying individual bees and can be automated, allowing the development of fast and efficient protocols for studying floral cover.

In parallel to indirect floral-based approaches, recent advances in artificial intelligence have enabled direct monitoring of bee populations through both acoustic and visual methods. For instance, Rodríguez Ballesteros et al. (Rodríguez et al.,2024) demonstrated that analyzing bee wingbeat sounds can provide valuable insights into species identification and stress levels. Similarly, Sledevič et al. (Sledevič et al.,2025) developed a computer vision system based on YOLOv8 to automatically detect and classify the behaviors of honeybees at the hive entrance. Although these approaches offer promising avenues for direct bee monitoring, they often require specialized equipment and limited spatial coverage.
AI gained prominence in 2012 when the deep convolutional neural network AlexNet won the ImageNet competition (ILSVRC), outperforming other methods with improved accuracy (Krizhevsky et al.,2012). Machine learning or deep learning applications have been applied in several fields, including big and complex datasets (Christin et al.,2019). With ecological data becoming increasingly large and complex, AI and ML can address current and future challenges, such as large-scale monitoring.

Remote sensing has undergone significant development in standardized biodiversity studies across large spatial scales (Torresani et al.,2020; Tamburlin et al.,2021; Rocchini et al.,2021; Torresani et al.,2018; Rocchini et al.,2022). While satellite remote sensing can cover vast areas and provide essential biodiversity parameters (Magnússon et al.,2021; Rocchini et al.,2022; Torresani et al.,2022), it is insufficient in offering precise measurements for finer details, such as floral cover estimation (Cruzan et al.,2016). Remote sensing by unoccupied aerial vehicles (UAVs) plays a crucial role in this regard.

UAVs can carry high-resolution cameras that fly at low altitudes, thus providing detailed imagery with accuracy as high as a few millimeters (Gallmann et al.,2022; Kattenborn et al.,2020; Torresani et al.,2024). UAV imagery can be used to assess some vegetation characteristics, including diversity (Guo et al.,2016), species and community distributions (Kaneko et al.,2014), plant functional traits (Capolupo et al.,2015), and invasive species mapping and monitoring (Alvarez-Taboada et al.,2017).
More recently, UAV imagery has also proven effective for detecting flowering vegetation and supporting pollinator ecology. For example, Atanasov et al. (Atanasov et al.,2024) developed a UAV-RGB-based algorithm for identifying white-flowering honey trees in

mixed forest ecosystems, demonstrating its utility in apiculture and forest mapping. Perrone et al. (Perrone et al.,2024)investigated the effect of flowering on spectral diversity metrics derived from UAV multispectral imagery, reinforcing the connection between flowering phenology and plant diversity assessment. These recent studies reflect the increasing potential of UAVs in ecological monitoring. Building on this foundation, our study focuses on open-field, species-rich grasslands, applying machine learning models to UAV RGB imagery to scalable and standardized floral cover estimation. This method is particularly suited to support large-scale pollinator monitoring programs.

Although our current study focuses exclusively on quantifying floral cover rather than identifying individual plant species, species-level identification could offer further ecological insights. However, such identification typically requires higher spatial resolution and the integration of derived image features, such as CLAHE, Sobel, texture layers, and entropy filters (Liang et al.,2018; Chen et al.,2020), which can enhance model performance by capturing structural and textural floral traits. Although technically feasible, this step was intentionally omitted because it would exceed our project's scope and primary aim.

This study proposes an indirect yet scalable and efficient approach for estimating pollinator abundance based on floral cover mapping through UAV imagery and automatic flower classification using machine learning models. In this study, remote sensing methods using drones were applied to capture high-resolution images, while various machine learning algorithms were tested for flower classification. With a well-structured protocol, reliable results can be obtained from new data within a short timeframe. Therefore, we believe it is crucial to better understand and continue exploring the application of ML models in ecological contexts, with the broader aim of supporting pollinator conservation and contributing to the development of standardized indicators of ecosystem health in agricultural landscapes.

The most crucial aspect of this study was the selection of testing models. We tested the performance of Random Forest(RF), Support Vector Machine(SVM), Neural Network(NNET), and Gradient Boost Machine(GBM). Despite conflicting reports in the literature (Torresani et al.,2023) on the use of SVM, we decided to test it with the addition of numerous optimization parameters to develop a more thorough understanding of how well it would perform for our specific application. Merging these four models with high-quality high-resolution images has great potential to advance ecological monitoring because it would allow for faster processing in areas of study and an efficient vegetation characterization (Randelovic,2020).

3.3 Materials and methods

3.3.1 Study area

The study area is located in the southern part of the Netherlands (figure 7), specifically near the town of Gulpen, and spans approximately 7x10 km. A diverse array of land uses, ranging from intensive agriculture and low-intensity farming to natural reserves, can be observed within our study area. We considered 30 grasslands within a land use intensity gradient ranging from nutrient-poor, biodiversity-rich semi-natural grasslands to intensively fertilized grasslands for fodder production. Our areas are part of the experimental biodiversity area network of the "SHOWCASE" project and are situated on loess soils (soils formed by the accumulation of wind-blown silt), colluvial clay deposits, and locally on lime-rich soils, with an elevation range from 70 to 171 m above sea level.

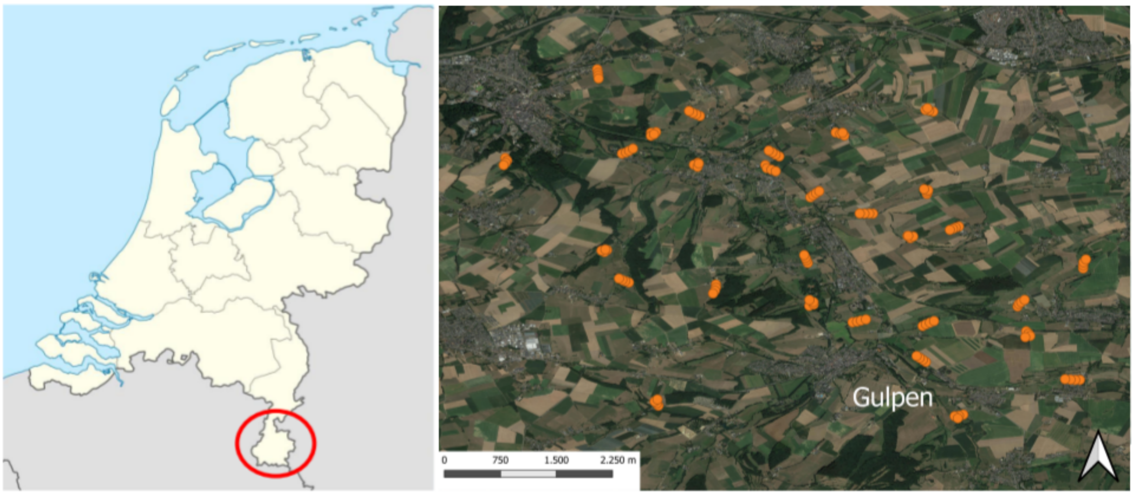

Figure 7: Study area in the southern region of the Netherlands, specifically in the Gulpen area.

3.3.2 Field data collection

A 150-m-long transect with a 1-m-wide buffer was defined in each of the 30 grasslands. Each transect was then divided into three sections of 50 m. The transect was created by placing markers in the field, clearly visible from the drone imagery. The transects were positioned from the edge to the center of the grassland and crossed elevational differences within the grassland, if present, to represent heterogeneity. The transects were spaced at a minimum distance of 435 m apart but, on average, more than 500 m apart to avoid repeat sampling of the same bee populations. Although some studies have indicated that some species, such as bumblebees, can forage over distances of several kilometers, the average foraging distance ranges between 250-550 meters, while smaller bees forage over even shorter distances (Redhead,John W. et al.,2016). Within the 30 transects, both bees and flowers were monitored to collect the necessary data for our project. Wild bees and the honeybee *Apis mellifera* were counted by walking along the transects, which is a standard method for studying plant-pollinator associations (Westphal et al.,2008). Only wild bees were included in the field analysis, whereas other bees were sampled and excluded from the analyses.

To mitigate potential issues related to human error by different operators, all transects were monitored by the same two trained and experienced observers. They counted all the bees using a butterfly net. The observers walked along the outlined transect for 15 min each for study standardization, excluding the time dedicated to handling and species identification. The samples were identified using the identification keys for Dutch Apidae (Falk et al.,2017; Nieuwenhuijsen et al.,2016; Nieuwenhuijsen et al.,2020). A significant portion of the samples could be analyzed in the field, while others required more detailed analysis.

Identification in the laboratory using stereomicroscopes and with a pollinator expert's assistance in some isolated cases. After observation, the bees were handled and collected in jars when necessary before proceeding with the survey. The time of day and weather can indeed influence observations; therefore, sampling was conducted under appropriate conditions and timings. Specifically, the fieldwork was conducted between April 19, 2022, and April 30, 2022, from 10:00 a.m. to 5:00 p.m., under favorable weather conditions (dry, >50% sunshine, and at least 15°C with wind speeds < 2 Beaufort) to ensure optimal sampling. Flower and bee surveys were conducted consistently within a two-day period. The abundance of bees (the total number of bees observed), bee species richness (number of unique species), and the Shannon diversity index (Pielou et al.,1966) were derived at the plot level.

$$H' = -\sum_{i=1}^{S} p_i \ln(p_i) \tag{1}$$

Shannon Index: *H' is the Shannon Index, S is the total number of species, p_i is the proportion of individuals belonging to species i, and* ln *is the natural logarithm. In ecological applications, Shannon's index captures both species richness and evenness. The index increases when individuals are more evenly distributed across species, reflecting a more balanced and diverse community. For example, a transect with two equally represented species will yield a higher H' value than a transect where one species dominates, even if both contain the same number of species. This reflects the idea that ecological diversity involves not only the presence of multiple species, but also the uniformity of their relative abundance.*

Flowers were sampled in each transect on the same day as the bees were monitored and sampled following the Scheper et al.'s methodology (Scheper et al.,2015). The number of open flowers in a given species was determined.

3.3.3 UAV data acquisition and processing

Two certified pilots affiliated with Wageningen University conducted the UAV flights. The flights of UAVs were simultaneously carried out using "DJI Matrice 210 RTK" to carry the RGB Zenmuse X5 camera (16.0 MP, 17.3 × 13.0 mm sensor) with an integrated RTK GPS. The images taken during the flight mission reached a ground sampling overlap of approximately 85%, which enabled the production of all the orthomosaics necessary for the successive phases of analysis. Flights were executed with the UAV flying at an altitude of 30 m above ground level.

For data preprocessing, the protocol outlined by Torresani et al. (2023) was followed (Torresani et al.,2023). UAV images were processed using Agisoft Metashape Professional Edition (Agisoft LLC.,2023),a photogrammetric software that applies structure-from-motion and stereo-matching algorithms for image alignment and 3D reconstruction. This workflow allowed us to generate high-resolution orthomosaics—geometrically corrected aerial images with uniform scale obtained by stitching together multiple overlapping photos and corrected for camera tilt and terrain distortion. Its intuitive workflow combines structure-from-motion and stereo-matching algorithms for image alignment and 3D image reconstruction (Moe et al.2020). The overall steps in the workflow included image alignment, dense point cloud creation and assessment, digital elevation model (DEM) development, and orthomosaic generation. In the first step with the "high" precision setting, the program analyzes and compares features within the images, creating a sparse 3D point cloud representation. During the same process, it automatically detects the details of the "Ground Control Points," downloading their GPS coordinates. It was also at this "high" level of detail that the denser point cloud was generated. The default settings of Metashape for DEM generation and final orthomosaic have been accepted. The latter were exported in GeoTIFF format, selecting the highest available spatial resolution, which is approximately 0.5 cm on average for the 30 examined areas. Once the orthomosaics were generated, they were imported into the QGIS software for further spatial processing and labeling. To ensure spatial consistency across all study sites, the orthomosaics were georeferenced using the WGS 84/UTM Zone 32 N coordinate system (EPSG: 32632). The labeling phase involved the manual creation of 150 flower and 150 grass polygons per transect (Fig.8). The number of polygons was arbitrarily chosen to maximize the amount of available training data for model development.

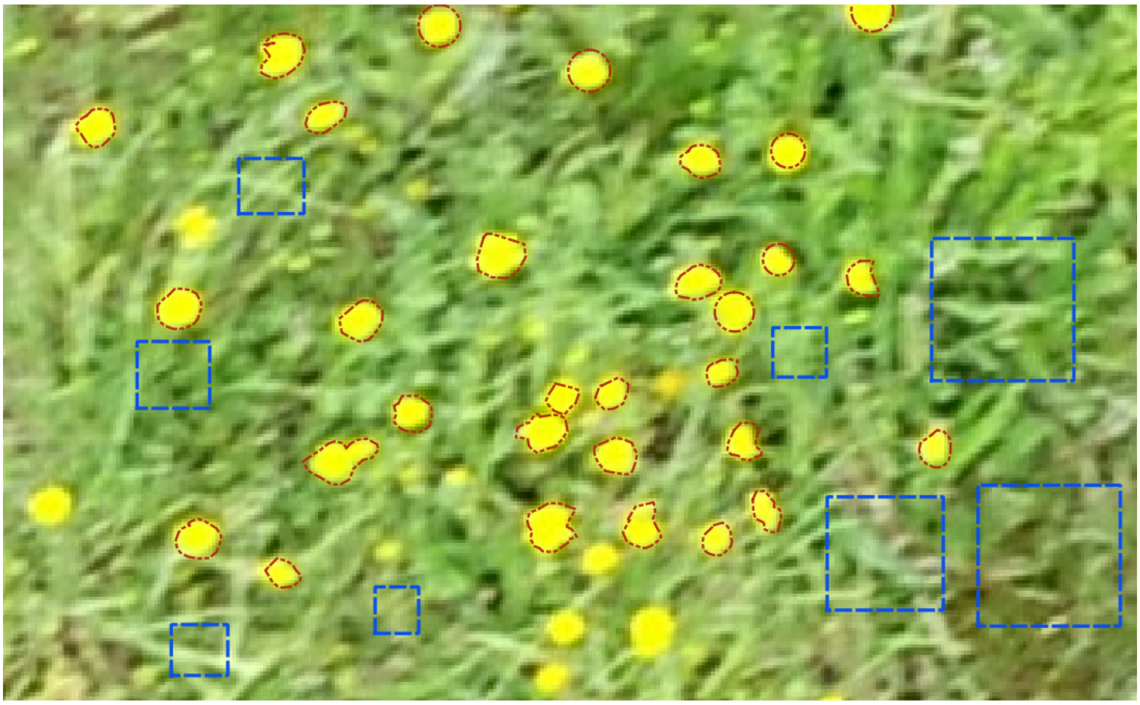

Figure 8: Zoomed-in section of a UAV image from one of the transects, showing manual annotations of vegetation elements. Polygons with red borders represent flowers, while polygons with blue dashed borders represent grass. The annotations were created using QGIS based on visual interpretation of the UAV imagery.

3.3.4 Machine learning methods

The choice of AI models in machine learning necessitated a conscious consideration of our initial data and objectives. The binary nature of our data aimed at discriminating between "flowers" and "non-flowers", and high dimensionality motivated us to go for a machine learning model that could reasonably serve our purpose. When selecting an appropriate ML model, several factors, such as the size and nature of the dataset, the complexity of the problem, the need for model interpretability, and the time available for training and prediction, are considered (Kuhn M. & Johnson K.,2013; Cutler A. et al.,2012). All models were created, trained, and tested using the R programming language.

The spectral bands used in our models were derived from RGB images acquired by the same UAV sensor under homogeneous environmental conditions. The bands are naturally correlated, a well-documented phenomenon in remote sensing applications. To ensure that this correlation did not negatively impact our models, we computed a correlation matrix and the Variance Inflation Factor (VIF). The results confirmed high correlation values among the RGB bands, ranging from 0.77 to 0.91, but a low Variance Inflation Factor (VIF) of 2.47. This indicates that multicollinearity is not a critical issue. All the RGB bands in the models were retained. This decision was further supported by the fact that models such as GBM and RF are inherently robust to correlated features, and SVM and NNET can tolerate moderate collinearity.

While previous studies (Torresani et al.,2023) developed machine learning models for each individual transect of the study area, this study used a more holistic approach where all transects from our study areas were integrated into the model design. This added a significant level of complexity to the project because the areas were highly heterogeneous and had differing characteristics. An exhaustive assessment of the most appropriate parameters and hyperparameters was performed to establish a more reliable, generalized, and flexible method for discriminating flower cover, which can then be correlated with bee abundance.

3.3.5 Random Forest (RF)

RF (Breiman,1996; Breiman,2001)because it is one of the most robust methods in high-dimensional and noisy data, which is quite normal in many UAV-based ecological studies. The ensemble approach of averaging multiple decision trees helps reduce overfitting, making it suitable for the complex task of distinguishing flower and non-flower areas in high-resolution drone imagery. This structure provides greater robustness and generalization capability than many other models (Breiman,2001). RF works by creating a collection of decision trees at the time of training: each tree is grown using a random sample of the training data, selected by replacement, a process known as bootstrapping (Liaw A. et al.,2022). Moreover, at each split of a tree node, only candidates for splitting are chosen at random from a subset of available variables. This randomization helps decorrelate the trees, reducing model variance with only a slight increase in bias. Once the trees have been built, the predicted class is determined by a majority vote across all trees (Nasiri et al.,2022; Kuhn M. & Johnson K.,2013; Cutler A. et al.,2012).

A RF algorithm was implemented to train the classification model using the ranger package. The model was then optimized via a grid search over key tuning parameters. The first important tune parameter is `mtry`: it regulates the number of variables that are randomly selected at each split. The value of `mtry` varied from 2 to the maximum, computed as the square root of the available variables minus one. Then, the `splitrule` tested with `Gini` impurity measure, which favors purer class separation, and the method of `extratrees` with robustness was improved by combining several split strategies. Finally, we explored the balance between model complexity and overfitting by tuning `min.node.size` from 1, 5, 10, and 15 as the minimum number of observations in each node.

Model performance was assessed by 10-fold cross-validation: the training set was divided into 10 sub-samples, one of which was chosen for validation and the rest for training in each iteration. The main metric of evaluation is represented by AUC-ROC, its most accurate measure of the model's ability to discriminate between the positive and negative classes. In doing so, this approach allowed for the testing of various combinations of parameters to find the best configuration that maximizes the predictive capability of the model, thereby enhancing classification performance between flowers and grass with effective handling of sensitivity and specificity trade-offs.

3.3.6 Gradient Boosting Machines (GBM)

GBM falls within the realm of ensemble learning techniques in machine learning, similar to RF, yet they belong to a distinct category of methods (Dietterich,T.G.,2000). While RF is categorized under Bagging (Bootstrap Aggregating) models, GBMs are situated within the Boosting methods, where each model aims to correct the errors made by preceding models. Meanwhile, in the case of training each model on random subsets of data, Boosting increases the weight of poorly classified instances, allowing the subsequent models focus more on those instances (Natekin et al.,2013). This was a sequential approach to enable GBM to iteratively refine the predictions, thus being particularly effective in capturing the variability of floral cover in such heterogeneous ecosystems of grasslands.

These techniques take their cue from the so-called "wisdom of crowds" (Surowiecki,2005), an important sociological factor in which the aggregation of independent, varied, and sometimes simplistic predictions leads to a globally superior result. The key for the GBM technique is to iteratively join many weak models, usually decision trees, each new model learning from previous ones where they went wrong (Natekin et al.,2013). The power of this approach lies in its ability to capture data complexities that could not be narrated through simpler models. Moreover, it is quite flexible and can be easily employed for both classification and regression tasks, thereby handling numerous data types (Friedman,2002).

GBM was chosen because it manages complex nonlinear interactions within high-dimensional ecological data. Its sequential learning approach captures the tiny variations in UAV imagery that distinguish the presence of flowers from background noise, whereby each corrects the previous models' errors. One of the key challenges of using GBMs is their natural tendency to overfit the first set of data they are presented with (Friedman,2002).

To avoid this problem, a lot of attention had to be paid to proper data management, namely, careful adjustment of training and test datasets. Cross-validation and regularization techniques were used to balance the risk of overfitting and `shrinkage`, which slows down the learning process while generalizing the model to unseen data. A GBM algorithm is applied to the `caret` R library to train this classification model. The final model was optimized through a grid search on several key parameters: the `interaction.depth` parameter, which prescribes the maximum depth of the trees in the model and thus determines the complexity of interactions among variables, was explored with values of 1, 3, and 5; the number of trees, expressed by the argument `n.trees`, representing the number of boosting iterations, was tested with 50, 100, and 150 trees; the shrinkage parameter, which accounts for the reduction of each tree's impact to prevent overfitting by tuning the learning rate, was varied from 0.01 to 0.1; and `n.minobsinnode`, which specifies the minimum number of observations required in a terminal node, was tested with 10 and 20 to balance model complexity and generalization, respectively. These parameters are selected based on their ability to balance model flexibility with the need to avoid overfitting in data with high feature variability, including UAV-derived imagery across diverse grasslands.

Performance was tested by 10-fold cross-validation: a robust approach in which the training set is divided into 10 subsets. Each subset is used for validation, and the other subsets are used for training. This technique reduces the model variance and ensures an intensive performance check.

3.3.7 Support Vector Machines (SVMs)

SVMs represent a powerful class of supervised ML classifiers that effectively operate on binary classification tasks (Vapnik V.,1995). SVMs stand out for their ability to effectively create a hyperplane or margin within high-dimensional space that effectively separates the two classes. This feature makes SVMs very applicable in various classification problems and on highly complicated and high-dimensional datasets, including those related to image recognition (Cervantes et al.,2020). The implementation of kernel functions further enhances the versatility and strength of SVMs. These functions allow SVMs to work in a transformed feature space without explicitly computing the data coordinates within that space (Cristianini & Taylor,2013). The use of the "kernel trick" enables SVMs to work upon an original space dataset that is not linearly separable because they map it into a higher-dimensional space in which the linear separation becomes possible. This approach enables SVMs to effectively handle complex nonlinear relationships between features (Schölkopf & Smola,2018). Moreover, SVMs are designed based on the principle of margin maximization. The algorithm maximizes the distance between the separating hyperplane and each class's nearest data points, which are referred to as support vectors. The greater the margin, the greater the model's generalization capability, reducing the risk of overfitting and yielding better prediction power against new data (Schölkopf & Smola,2018). The kernel function selection and tuning of SVM parameters, such as the regularization and kernel parameters, are crucial to the success of this model. The commonly used kernels include linear, polynomial, radial basis function (RBF), and sigmoid (Cristianini & Taylor,2013). Each kernel offers advantages and is suited to different types of data and relationships.

A linear kernel was selected for the proposed SVM. The training included 10-fold cross-validation with class probability estimates and the ROC metrics as evaluation criteria. Hyperparameter tuning was performed using a grid search over a range of regularization parameters, C, from 10^{-2} to 10^{2}. To ensure process integrity, the missing values were cleaned from both the training and test data before model training.

3.3.8 Neural Network (NNETs)

Among the most powerful models, NNETs are well-known in classification and regression problems, capturing intricate patterns in data (Goodfellow et al.,2016). A neural network is constructed from interconnected layers of nodes or neurons through which a multilayer architecture is used to process input data (LeCun et al.,2015).

Weight adjustment trains NNETs using the backpropagation technique to minimize gradient descent prediction errors (Rumelhart et al.,1986). They are composed of input, hidden, and output layers; nonlinear activation functions like ReLU or sigmoid make the network learn abstract features (Nair & Hinton,2010).

Regularization techniques, such as dropout and L2 regularization, prevent overfitting (Srivastava et al.,2014). During training, dropout randomly turns off neurons, whereas large weights are punished by L2 regularization. The early termination stops the training if the performance on a validation set starts stabilizing, thereby preventing overfitting (Prechelt,1998).

NNETs have applications in image recognition, binary data recognition, and many other areas due to their flexibility and scalability (Krizhevsky et al.,2012). Thus, neural networks can be characterized as a robust, flexible model with high-accuracy applications that can adapt to complex data-related challenges (Goodfellow et al.,2016).

Then, the Neural Network algorithm using the `nnet` method in the caret package was applied. The model was optimized by a grid search over critical hyperparameters: the `decay` parameter, which applies a regularization penalty to prevent overfitting by controlling weight size, was tested with values of 0.5, 0.01, and 0.0001; the `size` parameter, which defines the number of neurons in the hidden layer, was varied to 2, 4, and 8 to explore the model's different levels of complexity. The model's architecture was selected to balance flexibility and generalization capacity, considering the trade-offs between model accuracy and computational efficiency.

We employed 10-fold cross-validation to guide the training process. This means that the training set will be divided into 10 subsets, with one subset used for validation and the remaining subsets used in every training iteration. This gives a very robust performance estimate from the model and prevents overfitting.

Class weighting was performed in case of an imbalance between classes; the network used a weight of 10 for "flower" and 1 for "grass" to give proper importance to the minority classes. Then, the code takes advantage of parallel processing, speeding up the training process without sacrificing the accuracy of the results.

3.4 Results

The machine learning models employed for estimating floral cover demonstrated a high level of robustness. For instance, Figure 9 shows the models' evaluation, not only in terms of the accuracy, which has always been greater than 80%, but also regarding different performance metrics: precision, recall, F1 score, and AUC-ROC to provide a comprehensive overview.

The GBM model achieved the highest AUC-ROC value of 0.926, demonstrating its ability to balance the presence and background of flowers. This is reflected in its accuracy of 0.831, combined with a balanced accuracy of 0.785, which collectively indicate a reliable overall performance across various conditions. Furthermore, the high specificity of the model (0.835) suggests that it is effective in minimizing false positives, a crucial aspect in achieving precise flower counting.

The RF model was quite accurate at 0.97, and its precision was relatively good at 0.778, proving its effectiveness in accurately identifying the presence of flowers. However, the recall of the model is a relatively low at 0.403, indicating some difficulties in finding all flower instances. Nevertheless, this interaction between high accuracy and precision shows that it may be useful for cases where false positives are much more important than missed detections.

The SVM model achieved near-perfect specificity (0.999) and strong accuracy (0.965), with lower AUC-ROC and balanced accuracy scores (both 0.594), which may limit its reliability in diverse or complex backgrounds. This suggests that SVM, while highly selective, may struggle to effectively generalize across all test conditions, possibly missing some true flower instances.

The NNETs model was quite balanced, with high specificity (0.913) and strong performance on most metrics, such as an accuracy rate of 0.907 and an AUC-ROC score of 0.847. This shows its reliability and suggests that it can handle fluctuations in flower detection without the specific precision or specialization that might be obtained using the GBM or RF. We computed 95% confidence intervals for the AUC-ROC values using DeLong's method, a nonparametric approach specifically designed for estimating AUC, to provide additional support for evaluating model performance. Additionally, we applied a bootstrapping procedure with 2000 resamples to assess the variability of the AUC estimates. The confidence intervals for each model are as follows: GBM (0.9241–0.9298), NNET (0.8465–0.8574), RF (0.6975–0.7098), and SVM (0.5891–0.6052). Using these criteria, we conclude that the models, in particular GBM and RF, fulfilled the technical reliability standards, and their strengths fit the specific demands in floral cover estimation. The models achieved high accuracy and robustness, indicating their effectiveness in automated flower counting.

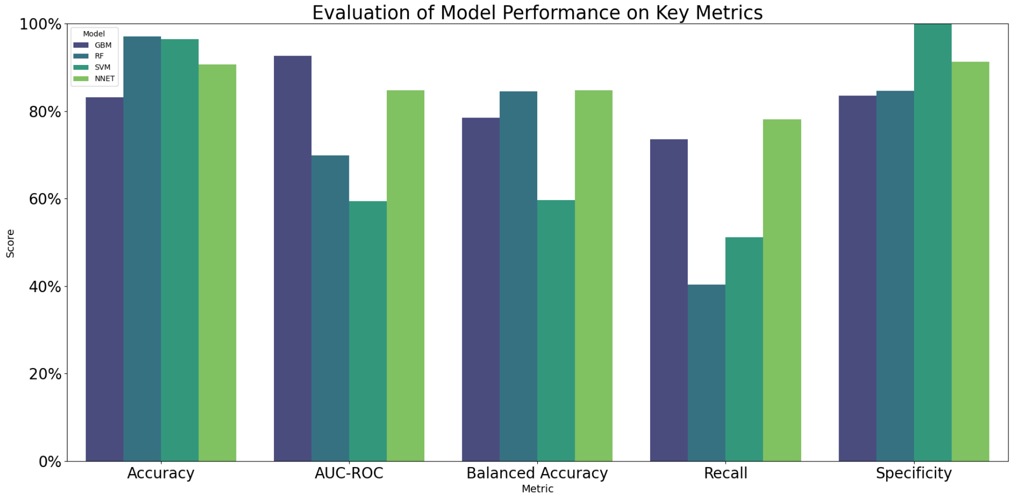

Figure 9: *Performance metrics for each model: the histograms display evaluation metrics—such as accuracy, specificity, recall and AUC-ROC—used to assess the effectiveness of the classification models in estimating floral cover. The models presented are, in order: Gradient Boosting Machine (GBM), Random Forest (RF), Support Vector Machine (SVM), and Neural Network (NNET), with GBM demonstrating the highest overall robustness. All metrics were derived from the confusion matrices generated during model validation. Colours were selected to be readable by individuals with colour vision deficiencies (Rocchini et al.,2024).*

The figure 10 shows an example of the classification workflow and how the GBM accurately and efficiently identified the flowers.

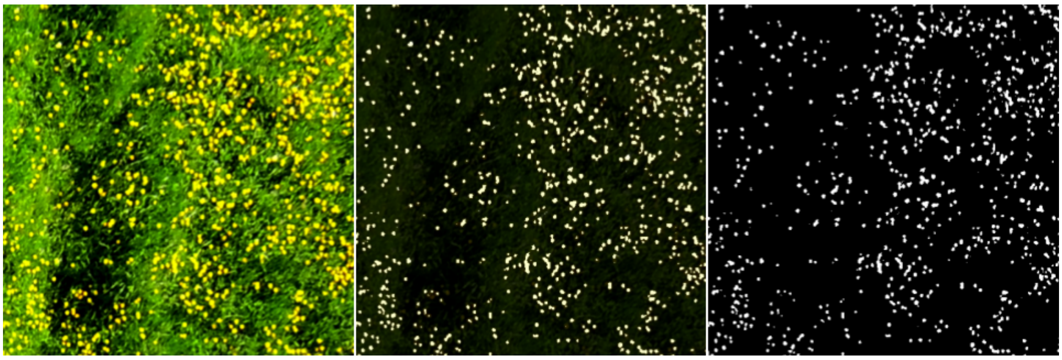

Figure 10: *Example of automatic flower classification in RGB images.*
(a) Original drone-acquired image showing herbaceous vegetation with yellow flowers.
(b) Overlay of the predicted mask on the original image, with pixels classified as "flowers" highlighted in white.
(c) Final binary mask resulting from the classification, where white pixels represent areas identified as flowers.
This process enables the automatic and spatially explicit estimation of floral cover, which is subsequently used as an indirect proxy to infer the potential presence of pollinators.

The floral cover estimates obtained from RGB UAV images from our models were generally positively correlated with those made by expert field observers 11. We displayed a subset of 20 sites in each figure to ensure visual clarity and avoid overlapping points that would reduce interpretability. The selection was made to ensure representativeness, and all models were applied to the full dataset. The site exclusion in the figures was solely for graphical purposes and does not affect the consistency of model comparison. The GBM model was the most performant among the four, achieving an R^2 correlation exceeding 70%. All models involved in this phase demonstrated a p-value < 0.05. Specifically, GBM provided accurate floral cover estimates, whereas RF and NNET tended to overestimate, whereas SVM significantly underestimated the results.

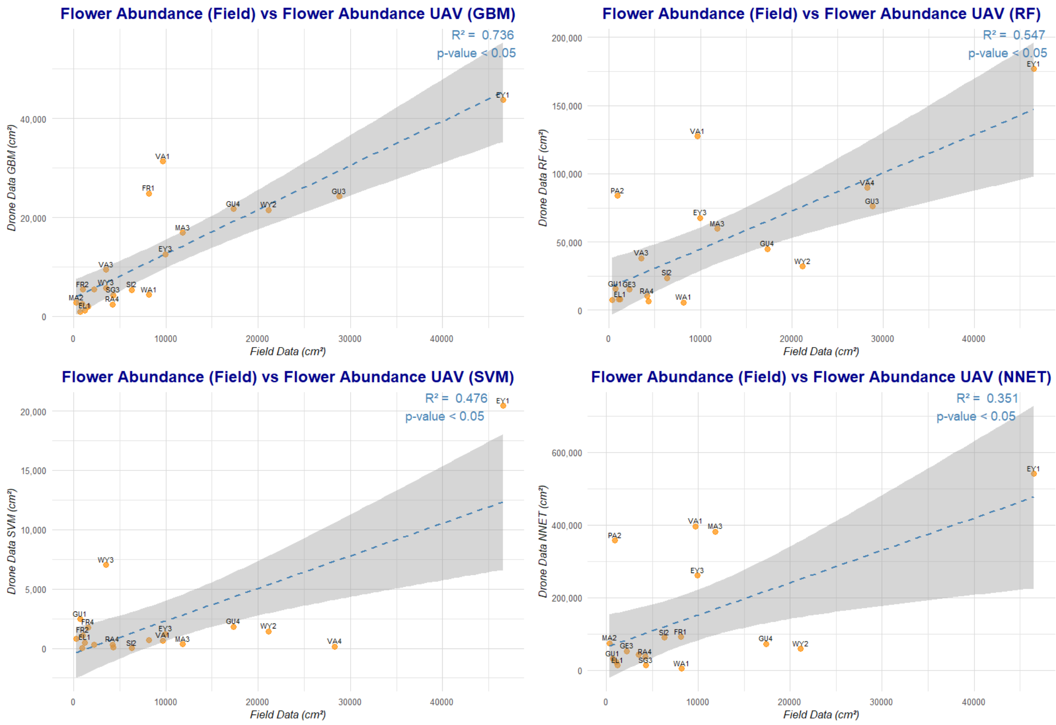

Figure 11: *Graphs of correlations between flower cover estimated in situ by botanical experts and flower cover estimated through machine learning models, using RGB images captured by UAV drones.*

Only the GBM and RF produced acceptable results regarding the correlation between bee abundance and flower area in graphs, with an R^2 correlation of 42% and 54%, respectively, and a p-value < 0.05.

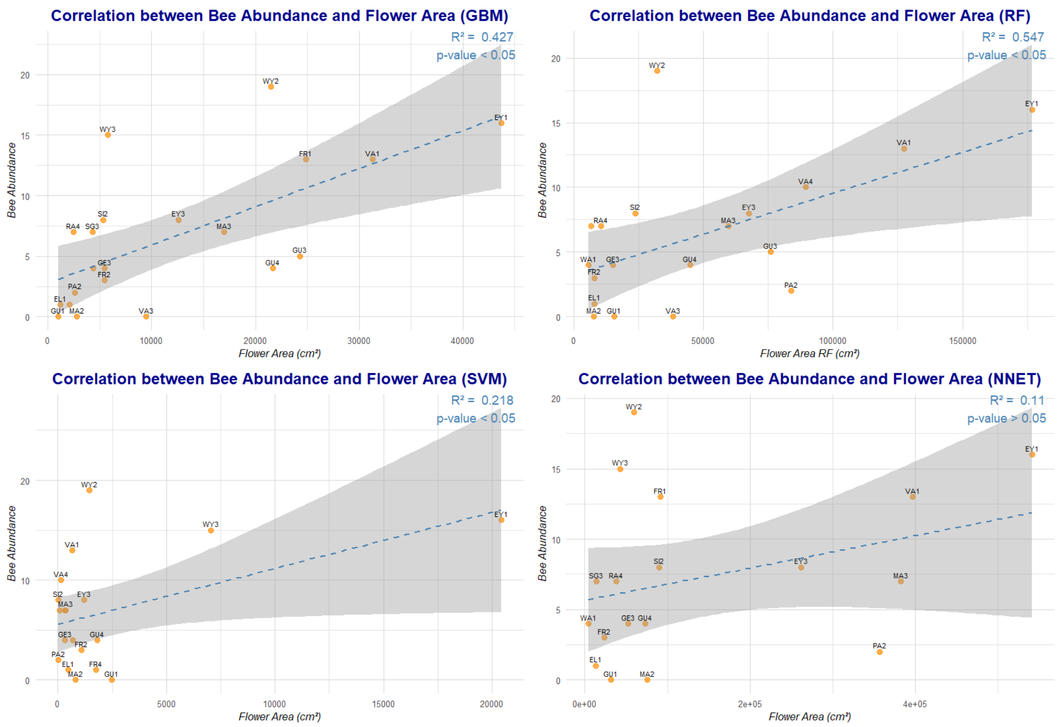

Figure 12: *Graphs of correlations between bee abundance estimated in situ by experts and flower area in cm^2 estimated through machine learning models trained on RGB images captured by UAV drones.*

For bee species richness, the highest results were obtained with GBM and RF, both presenting an R^2 close to 40% and 50% and a p-value < 0.05 (Graph 13).
This indicates that, in our case, decision tree models are the most effective for binary image classification tasks, specifically distinguishing between flower and non-flower.

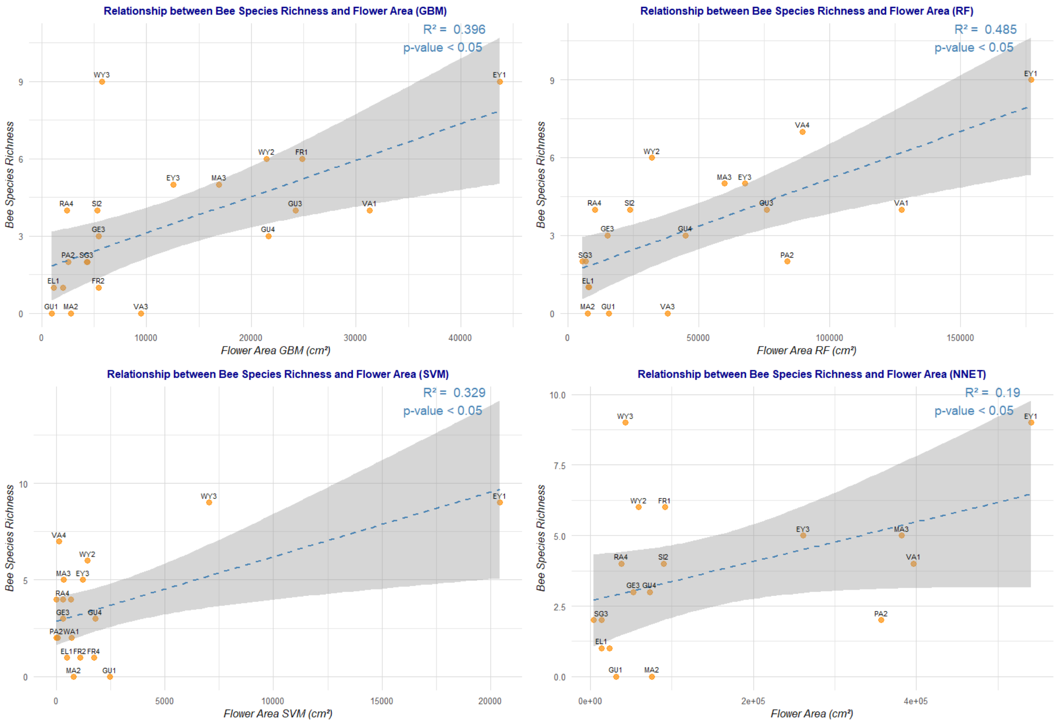

Figure 13: *Graphs of correlations between bee species richness and flower area in cm^2 estimated through machine learning models trained on RGB images captured by UAV drones.*

A correlation was ultimately found between the floral area in cm² estimated by our models and the Shannon index (Graph 14). The RF and GBM models were the most effective at achieving an average value correlation.

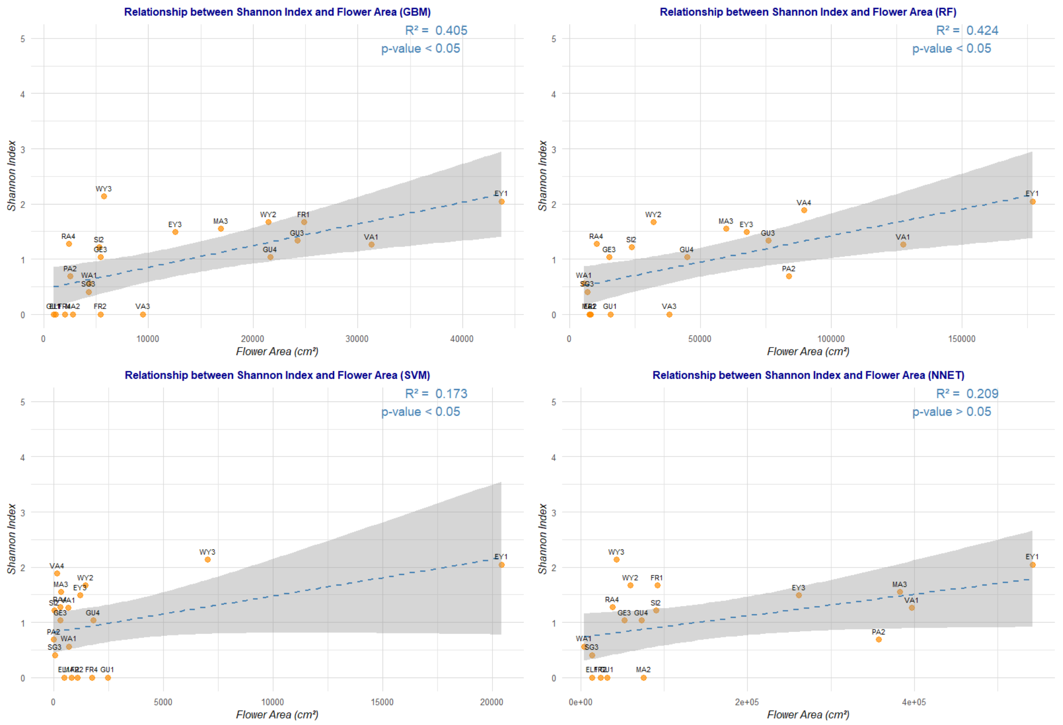

Figure 14: *Graphs of correlations between Shannon index and flower area in* cm^2 *estimated through machine learning models trained on RGB images captured by UAV drones.*

3.5 Discussion

Our project initially focused on identifying suitable ML models to achieve our objective. Models were trained using a consistent protocol across all study areas. This uniform protocol allows for quicker and more efficient analyses in the succeeding years and even for the expansion of study areas. In this study, we present a new, alternative, and complementary approach for estimating floral coverage correlated with field analyses. Using RGB aerial images from UAVs can accelerate the analysis process, providing key parameters for future studies. Additionally, UAV-based surveys in our study were conducted with minimal environmental disturbance. To avoid direct interaction with vegetation and pollinators, UAVs were launched from roads or non-sensitive areas. Furthermore, they operated at an altitude of 30 m, which has negligible effects on pollinator activity. This methodological choice allows for large-scale biodiversity monitoring while minimizing ecosystem disruption. Despite the large initial dataset and its heterogeneity, our analyses revealed highly significant and positive correlations between floral coverage estimates obtained through UAV RGB images, our ML methods, and traditional in situ estimates by expert botanists. In a previous study (Torresani et al.,2023), correlations were higher than in our current results due to the high specificity for each area, achieved by analyzing each area individually. However, we managed to maintain a higher level of model reliability by integrating all study areas within our dataset, allowing them to be generalized.

These results highlight the capability of this approach in ecological monitoring and biodiversity management, providing a fast and scalable approach for assessing floral coverage in large-scale conservation efforts. Positive correlations were identified between bee abundance, Shannon index, bee species richness, and floral area (in cm^2) estimated by our models. However, not all models showed optimal results with these parameters. This could be mainly due to the nature of our data. Despite its simplicity, the data's binary nature (flower/non-flower) presents limitations, such as the inability to integrate useful parameters to support the recognition of more complex patterns. Decision tree models, particularly GBM, demonstrated excellent performance in this context. We attribute this result to several factors: decision trees can leverage the simplicity of binary data due to their iterative nature on the most relevant features. Moreover, unlike SVM and NNET (Kavzoglu et al.,2020), which often requires advanced tuning with various data preprocessing steps, decision trees can account for nonlinear relationships in the data without complex transformation (Zhang et al.,2019). Furthermore, decision trees are less sensitive to noise in the data than SVM and NNET. If the dataset contains noise or anomalies, as in our case, decision trees can better manage them, offering more stable performance (Zhang et al.,2019).

Although models such as SVM and NNET demonstrate excellent performance on certain metrics—with SVM achieving almost 100% specificity—the GBM and RF models' actual capability in predicting floral cover proved superior. This difference can be attributed to the resilience and adaptability of tree-based algorithms in handling the variability present in the dataset, including factors such as lighting conditions and background characteristics. GBM performs better in generalizing complex scenarios because of its ability to effectively select relevant features and handle noise, as demonstrated by the balance between accuracy and recall metrics. This stability allowed for a more reliable flower count, reducing errors from false positives/negatives, and enabling high precision even in more challenging conditions. Moreover, the hierarchical structure of trees makes GBM particularly effective in handling high-resolution image datasets, where the capability to capture fine details is fundamental for accurately estimating floral cover. Owing to their ability to balance high accuracy with operational flexibility, these models not only offer solid performance in quantitative tests but also prove more suitable for fulfilling the specific needs of our application. Therefore, although other models may show high values on isolated metrics, GBM represents the best balance for our study. This fact, in addition to justifying its use in the current research, suggests that it may be the optimal choice for future applications in floral cover estimation and other complex ecological contexts, where reliability and scalability are essential.

There are several considerations regarding flower classification using our models. In our study, the average weighted size of the flowers was 1.66 cm^2. With the pixel size likely up to 3.2 cm^2, this would suggest that a pixel would have been correctly classified if it contained one whole flower, and larger pixels could easily have caused misclassification. In our study, the issue of mixed pixels likely affected the underestimation or overestimation of UAV RGB-estimated floral coverage. This may have occurred because pixels with less than 50% floral coverage were classified as having 0% coverage, whereas those with more than 50% coverage were classified as having 100%. However, this problem can be easily addressed by capturing more detailed resolution images. Flights at 30 m resulted in negligible flowers occupying only one pixel of the image in some cases, making classification difficult. Using images obtained from flights at lower altitudes, such as 15 m, we can incorporate more complex features, such as shape and floral formula, into our datasets. Our orthomosaics encompassed a significant unused area outside the analysis transects. We suggest decreasing the flight altitude, focusing more on the transect area to maintain the same flight time while obtaining more precise images. This allows for a much more accurate assessment and provides the opportunity to study not only floral cover but also diversity, enabling exploration of additional ecological and biological topics.

In our previous research (Torresani et al.,2023), we evaluated the effectiveness of using RGB aerial images from UAVs combined with appropriate ML methods by creating individual models for each analyzed area. This study confirms the effectiveness of this method and proposes a new protocol that focuses on creating unified models that encompass all considered areas. Working in this manner may result in a slight decrease in accuracy due to the increased number of variables. However, it ensures faster and more generalized analysis, which better manages overfitting compared to creating individual models for each area. Furthermore, the best model by Torresani et al. (Torresani et al.,2023) (RF) overestimated the floral coverage by approximately a factor of two compared to in situ analyses. In our study, however, the correlation between field analyses and those based on the combination of RGB UAV imagery and GBM was close to 1:1, despite the model being trained using all data simultaneously.

The accuracy of the correlation between our estimated floral coverage and the in situ floral coverage was higher than that of the correlation with in situ bee abundance. This was expected because the algorithms were trained to discriminate between the flower and non-flower variables. Bees are currently too small and difficult to detect in UAV RGB images, and their abundance and species diversity can only be indirectly inferred, as they generally show a strong relationship with flower availability (Sutter et al.,2017; Potts et al.,2003). Previous studies have shown that bee abundance and species diversity are strongly correlated with floral coverage and species diversity. Bee species diversity is generally better predicted by floral species diversity (Potts et al.,2003; Fründ et al.,2010; Neumüller et al.,2021), whereas bee abundance is more closely linked to floral coverage (Blaauw et al.,2014; Scheper et al.,2021). However, floral coverage and species diversity are often correlated, and floral species diversity is linked to bee abundance (Steffan-Dewenter et al.,2001; Theodorou et al.,2020), whereas floral coverage is associated with bee species diversity (Ebeling et al.,2008).

This study contributes to the growing body of research aimed at developing scalable, cost-effective methods for monitoring pollinators to support EUPoMS. Although our approach is indirect, it offers the potential for rapid assessments over large areas, particularly in agricultural landscapes where spatial coverage is limited by the logistical burden of fieldwork. Floral cover, as detected via UAV and automated classification, may serve as a proxy to identify changes in pollinator resource availability and help prioritize areas for more detailed species-level investigations. This method can complement rather than replace taxonomic sampling and contribute to reversing pollinator decline across Europe.

The method used in this study aims to explore the vast applications of ML models in the ecological and biological fields, areas that are still relatively new with infinite possibilities for exploration. Similar to other research that has employed machine learning to enhance and expedite ecological analyses (Thessen,2016; Liu et al.,2018), our study has identified current limitations during the image acquisition phase using four machine learning models, leading to the optimal planning of future UAV flight missions and the subsequent development of optimized multi-parameter models. We have recognized the enormous potential of these models, particularly GBM, which can be further developed by incorporating increasingly complex parameters, with the ultimate goal of creating a model capable of operating efficiently and semi-autonomously.

3.6 Conclusion

This study demonstrates the potential of combining UAV-acquired imagery with machine learning methods for large-scale floral coverage, a key factor in pollinator habitat quality. We offer a highly efficient and scalable approach that improves traditional biodiversity monitoring methods through automation and reproducibility using unmanned aerial vehicle (UAV) RGB imagery and sophisticated classification models.

Our findings indicate that decision tree-based models, particularly GBM, offer the most reliable performance for estimating floral cover from UAV imagery, exhibiting strong correlations with field observations. Although alternative models, such as SVM and NNET, yielded promising results in specific cases, they lacked the overall robustness required for generalized ecological applications. These findings underscore the importance of model selection when applying AI techniques to environmental studies.

This study has broader implications for pollinator conservation and ecological monitoring beyond its methodological contribution. Our approach can support conservation strategies, inform land management practices, and mitigate pollinator decline by enabling rapid and cost-effective floral resource assessments. Future research should focus on improving UAV imagery's spatial resolution, integrating multispectral and hyperspectral data, and exploring deep learning techniques to enhance model generalization across diverse ecosystems. This study aims to bridge the gap between remote sensing, AI, and ecological research to pave the way for more efficient, scalable, and data-driven biodiversity monitoring approaches. Continued advancements in these technologies are crucial for addressing the challenges of global biodiversity loss and supporting sustainable conservation initiatives.

3.7 Future Work and Project Directions

Further steps of our study include assessing the optimal flight parameters for capturing images with the best level of detail and continuing to develop our models in order to estimate floral cover and floral diversity in the study areas. Such models would, in fact, set the base for quite a few further studies and ecological applications.

Moreover, full exploitation of UAV capabilities will be ensured through the employment of simultaneous acquisitions from different spectral sensors. Beyond RGB sensors, images from UV and multispectral sensors will be integrated in order to explore and optimize their full potential.

We will explore new models of machine learning and deep learning with the intention of studying their performance to find out which model best fits our needs.

3.8 Author Contributions

L.C. contributed to the conceptualization of the study; designed and implemented the modelling framework; curated and harmonized multi-site datasets; performed all machine-learning analyses; interpreted the results; and led the writing of the manuscript. M.T. and D.R. contributed to study design and supervision. Co-authors contributed to data collection, annotation protocols, ecological interpretation, and manuscript revision.

3.9 Acknowledgments

We thank both the UAV pilots, Berry Onderstal and Peter van der Zee, for their work in the field data campaigns.
We thank 21 farmers and three nature managers (Staatsbosbeheer, Natuurmonumenten & Limburgs Landschap) for allowing data collection for this study on their fields.
We thank Peter Reinier de Vries, Brian Pater, Remco Ploeg, and Ivo Raemakers for their contribution to data collection and taxonomic identification during our study.

3.10 Funding

The study has received funding from the project SHOWCASE (SHOWCASing synergies between agriculture, biodiversity and Ecosystem services to help farmers capitalising on native biodiversity) within the European Union's Horizon 2020 Research and Innovation Programme (grant agreement No 862480). This publication reflects only the authors' view and the European Commission is not responsible for any use that may be made of the information it contains.

3.11 Data availability

We acknowledge the importance of immediate accessibility. However, the dataset used in this article is too large to be shared in its entirety (several terabytes). To ensure the replicability of our study while maintaining practical feasibility, we have provided a representative subset on Zenodo at `https://zenodo.org/records/14793347`.This repository also includes an ISO 19115 metadata file in XML format to ensure dataset reproducibility and traceability. For specific requests regarding access to the full dataset, we offer direct data transfer support. You can find all code about our machine learning models at this link `https://github.com/Ludovico-Chieffallo/Machine_Learning_for_biodiversity.git`

References

[Alvarez-Taboada et al.,2017] Alvarez-Taboada, Flor, Claudio Paredes, and Julia Julián-Pelaz. "Mapping of the invasive species Hakea sericea using unmanned aerial vehicle (UAV) and WorldView-2 imagery and an object-oriented approach." Remote Sensing 9.9 (2017): 913.

[Agisoft LLC.,2023] Agisoft LLC. (2023). Agisoft Metashape (Versione 1.7.1) [Software]. `https://www.agisoft.com`

[Atanasov et al.,2024] Atanasov, Atanas Z., et al. "A Novel Algorithm to Detect White Flowering Honey Trees in Mixed Forest Ecosystems Using UAV-Based RGB Imaging." AgriEngineering 6.1 (2024): 95-112.

[Biesmeijer et al.,2006] Biesmeijer, Jacobus C., et al. "Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands." Science 313.5785 (2006): 351-354.

[Blaauw et al.,2014] Blaauw, Brett R., and Rufus Isaacs. "Flower plantings increase wild bee abundance and the pollination services provided to a pollination-dependent crop." Journal of Applied Ecology 51.4 (2014): 890-898.

[Breiman,1996] Breiman, L. Bagging predictors. Mach Learn 24, 123–140 (1996). https://doi.org/10.1007/BF00058655

[Breiman,2001] Breiman, L. Random Forests. Machine Learning 45, 5–32 (2001). https://doi.org/10.1023/A:1010933404324

[Capolupo et al.,2015] Capolupo, Alessandra, et al. "Estimating plant traits of grasslands from UAV-acquired hyperspectral images: a comparison of statistical approaches." ISPRS International Journal of Geo-Information 4.4 (2015): 2792-2820.

[Cardinale et al.,2012] Cardinale, Bradley J., et al. "Biodiversity loss and its impact on humanity." Nature 486.7401 (2012): 59-67.

[Cervantes et al.,2020] Cervantes, Jair, et al. "A comprehensive survey on support vector machine classification: Applications, challenges and trends." Neurocomputing 408 (2020): 189-215.

[Chen et al.,2020] Chen, Guobin, Zhiyong Jiang, and M. M. Kamruzzaman. "Radar remote sensing image retrieval algorithm based on improved Sobel operator." Journal of Visual Communication and Image Representation 71 (2020): 102720.

[Cristianini & Taylor,2013] Cristianini N, Shawe-Taylor J.(2013) Kernel-Induced Feature Spaces. In: An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press; 2000:26-51.

[Christin et al.,2019] Christin, Sylvain, Éric Hervet, and Nicolas Lecomte. "Applications for deep learning in ecology." Methods in Ecology and Evolution 10.10 (2019): 1632-1644.

[Cruzan et al.,2016] Cruzan, Mitchell B., et al. "Small unmanned aerial vehicles (micro-UAVs, drones) in plant ecology." Applications in plant sciences 4.9 (2016): 1600041.

[Cutler A. et al.,2012] Cutler, A., Cutler, D.R., Stevens, J.R. (2012). Random Forests. In: Zhang, C., Ma, Y. (eds) Ensemble Machine Learning. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-9326-7_5

[Dietterich,T.G.,2000] Dietterich, T.G. (2000). Ensemble Methods in Machine Learning. In: Multiple Classifier Systems. MCS 2000. Lecture Notes in Computer Science, vol 1857. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45014-9_1

[Ebeling et al.,2008] Ebeling, Anne, et al. "How does plant richness affect pollinator richness and temporal stability of flower visits?." Oikos 117.12 (2008): 1808-1815.

[Falk et al.,2017] Falk, Steven, and R. Lewington. Veldgids bijen voor Nederland en Vlaanderen. 2017.

[Friedman,2002] Friedman, Jerome H. "Stochastic gradient boosting." Computational statistics & data analysis 38.4 (2002): 367-378.

[Fründ et al.,2010] Fründ, Jochen, Karl Eduard Linsenmair, and Nico Blüthgen. "Pollinator diversity and specialization in relation to flower diversity." Oikos 119.10 (2010): 1581-1590.

[Gallmann et al.,2022] Gallmann, Johannes, et al. "Flower mapping in grasslands with drones and deep learning." Frontiers in plant science 12 (2022): 774965.

[Gallai et al.,2009] Gallai, Nicola, et al. "Economic valuation of the vulnerability of world agriculture confronted with pollinator decline." Ecological economics 68.3 (2009): 810-821.

[Ghazoul,2005] Ghazoul, Jaboury. "Buzziness as usual? Questioning the global pollination crisis." Trends in ecology & evolution 20.7 (2005): 367-373.

[Ghisbain et al.,2023] Ghisbain, Guillaume, et al. "The new annotated checklist of the wild bees of Europe (Hymenoptera: Anthophila)." Zootaxa 5327.1 (2023): 1-147.

[Goodfellow et al.,2016] Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press, 2016.

[Guo et al.,2016] Guo, Qinghua, et al. "Perspectives and prospects of unmanned aerial vehicle in remote sensing monitoring of biodiversity." Biodiversity Science 24.11 (2016): 1267.

[Kaneko et al.,2014] Kaneko, Korehisa, and Seiich Nohara. "Review of effective vegetation mapping using the UAV (Unmanned Aerial Vehicle) method." Journal of Geographic Information System 6.06 (2014): 733.

[Kattenborn et al.,2020] Kattenborn, Teja, et al. "Convolutional Neural Networks accurately predict cover fractions of plant species and communities in Unmanned Aerial Vehicle imagery." Remote Sensing in Ecology and Conservation 6.4 (2020): 472-486.

[Kavzoglu et al.,2020] Kavzoglu, Taskin, Furkan Bilucan, and Alihan Teke. "Comparison of support vector machines, random forest and decision tree methods for classification of sentinel-2A image using different band combinations." 41st Asian Conference on Remote Sensing (ACRS 2020). Vol. 41. 2020.

[Kuhn M. & Johnson K.,2013] Kuhn, M., & Johnson, K. (2013). Applied Predictive Modeling. Springer New York, NY. https://doi.org/10.1007/978-1-4614-6849-3

[Kleijn et al.,2009] Kleijn, David, et al. "On the relationship between farmland biodiversity and land-use intensity in Europe." Proceedings of the royal society B: biological sciences 276.1658 (2009): 903-909.

[Kleijn et al.,2015] Kleijn, David, et al. "Delivery of crop pollination services is an insufficient argument for wild pollinator conservation." Nature communications 6.1 (2015): 7414.

[Klein et al.,2007] Klein, Alexandra-Maria, et al. "Importance of pollinators in changing landscapes for world crops." Proceedings of the royal society B: biological sciences 274.1608 (2007): 303-313.

[Kremen et al.,2002] Kremen, Claire, Neal M. Williams, and Robbin W. Thorp. "Crop pollination from native bees at risk from agricultural intensification." Proceedings of the National Academy of Sciences 99.26 (2002): 16812-16816.

[Krizhevsky et al.,2012] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems 25 (2012).

[LeCun et al.,2015] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436-444.

[Liu et al.,2018] Liu, Zelin, et al. "Application of machine-learning methods in forest ecology: recent progress and future challenges." Environmental Reviews 26.4 (2018): 339-350.

[Liang et al.,2018] Liang, Xining, et al. "A real-time infrared image enhancement algorithm based on improved CLAHE." 2018 international conference on image and video processing, and artificial intelligence. Vol. 10836. SPIE, 2018.

[Liaw A. et al.,2022] Liaw, Andy, and Matthew Wiener. "Classification and regression by randomForest." R news 2.3 (2002): 18-22.

[Magnússon et al.,2021] Magnússon, Rúna Í., et al. "Shrub decline and expansion of wetland vegetation revealed by very high resolution land cover change detection in the Siberian lowland tundra." Science of the Total Environment 782 (2021): 146877.

[Moe et al.2020] Moe, Kyaw Thu, et al. "Comparing individual tree height information derived from field surveys, LiDAR and UAV-DAP for high-value timber species in Northern Japan." Forests 11.2 (2020): 223.

[Nair & Hinton,2010] Nair, V., & Hinton, G. E. (2010). Rectified Linear Units Improve Restricted Boltzmann Machines. Proceedings of the 27th International Conference on Machine Learning (ICML-10).

[Nasiri et al.,2022] Nasiri, Vahid, et al. "Modeling forest canopy cover: a synergistic use of Sentinel-2, aerial photogrammetry data, and machine learning." Remote Sensing 14.6 (2022): 1453.

[Natekin et al.,2013] Natekin, Alexey, and Alois Knoll. "Gradient boosting machines, a tutorial." Frontiers in neurorobotics 7 (2013): 21.

[Neumüller et al.,2021] Neumüller, Ulrich, et al. "Prolonged blooming season of flower plantings increases wild bee abundance and richness in agricultural landscapes." Biodiversity and Conservation 30 (2021): 3003-3021.

[Nieuwenhuijsen et al.,2016] Nieuwenhuijsen, Hans, and Theo Peeters, eds. Nederlandse bijen op naam brengen: deel 1. Stichting Jeugdbondsuitgeverij, 2016.

[Nieuwenhuijsen et al.,2020] Nieuwenhuijsen, Hans, Theo Peeters e Dominic Dijkshoorn, eds. Nederlandse bijen op naam brengen: deel 2. Stichting Jeugdbondsuitgeverij, 2020.

[Ollerton et al.,2011] Ollerton, Jeff, Rachael Winfree, and Sam Tarrant. "How many flowering plants are pollinated by animals?." Oikos 120.3 (2011): 321-326.

[Perrone et al.,2024] Perrone, Michela, et al. "“Flower power”: How flowering affects spectral diversity metrics and their relationship with plant diversity." Ecological Informatics 81 (2024): 102589.

[Pielou et al.,1966] Pielou, Evelyn Chrystalla. "Shannon’s formula as a measure of specific diversity: its use and misuse." The American Naturalist 100.914 (1966): 463-465.

[Potts et al.,2003] Potts, Simon G., et al. "Linking bees and flowers: how do floral communities structure pollinator communities?." Ecology 84.10 (2003): 2628-2642.

[Potts et al.,2021] Potts, Simon, et al. "Proposal for an EU pollinator monitoring scheme." (2021).

[Prechelt,1998] Prechelt, L. (1998). Early Stopping - But When?. Neural Networks: Tricks of the Trade, 55-69.

[Randelovic,2020] RANDELOVIĆ, Predrag. "(2020). Prediction of soybean plant density using a machine learning model and vegetation indices extracted from RGB images taken with a UAV." Agronomy. España 10.8: 1108.

[Reed et al.,2003] Reed, Bradley C., Michael White, and Jesslyn F. Brown. "Remote sensing phenology." Phenology: an integrative environmental science (2003): 365-381.

[Redhead,John W. et al.,2016] Redhead, John W., et al. "Effects of habitat composition and landscape structure on worker foraging distances of five bumble bee species." Ecological applications 26.3 (2016): 726-739.

[Robinson and Sutherland,2002] Robinson, Robert A., and William J. Sutherland. "Post-war changes in arable farming and biodiversity in Great Britain." Journal of applied Ecology 39.1 (2002): 157-176.

[Rocchini et al.,2021] Rocchini, Duccio, et al. "rasterdiv—An Information Theory tailored R package for measuring ecosystem heterogeneity from space: To the origin and back." Methods in Ecology and Evolution 12.6 (2021): 1093-1102.

[Rocchini et al.,2022] Rocchini, Duccio, et al. "Double down on remote sensing for biodiversity estimation: a biological mindset." Community Ecology 23.3 (2022): 267-276.

[Rocchini et al.,2024] Rocchini, Duccio, et al. "Under the mantra:'Make use of colorblind friendly graphs'." Environmetrics (2024): e2877.

[Rodríguez et al.,2024] Rodríguez Ballesteros, Alberto, et al. "Towards acoustic monitoring of bees: wingbeat sounds are related to species and individual traits." Philosophical Transactions of the Royal Society B 379.1904 (2024): 20230111.

[Rumelhart et al.,1986] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533-536.

[Saunders et al.,1991] Saunders, Denis A., Richard J. Hobbs, and Chris R. Margules. "Biological consequences of ecosystem fragmentation: a review." Conservation biology 5.1 (1991): 18-32.

[Scheper et al.,2014] Scheper, Jeroen, et al. "Museum specimens reveal loss of pollen host plants as key factor driving wild bee decline in The Netherlands." Proceedings of the National Academy of Sciences 111.49 (2014): 17552-17557.

[Scheper et al.,2015] Scheper, Jeroen, et al. "Local and landscape-level floral resources explain effects of wildflower strips on wild bees across four European countries." Journal of Applied Ecology 52.5 (2015): 1165-1175.

[Scheper et al.,2021] Scheper, Jeroen, et al. "Attractiveness of sown wildflower strips to flower-visiting insects depends on seed mixture and establishment success." Basic and Applied Ecology 56 (2021): 401-415.

[Schölkopf & Smola,2018] Schölkopf, B., & Smola, A. J. (2018). Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. The MIT Press. https://doi.org/10.7551/mitpress/4175.001.0001

[Segre et al.,2023] Segre, Hila, et al. "Butterflies are not a robust bioindicator for assessing pollinator communities, but floral resources offer a promising way forward." Ecological Indicators 154 (2023): 110842.

[Sledevič et al.,2025] Sledevič, Tomyslav, et al. "Visual recognition of honeybee behavior patterns at the hive entrance." PloS one 20.2 (2025): e0318401.

[Srivastava et al.,2014] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research, 15(1), 1929-1958.

[Steffan-Dewenter et al.,2001] Steffan-Dewenter, I., and T. Tscharntke. "Succession of bee communities on fallows. Ecography (Cop.) 24, 83–93." (2001).

[Steffan-Dewenter et al.,2005] Steffan-Dewenter, Ingolf, Simon G. Potts, and Laurence Packer. "Pollinator diversity and crop pollination services are at risk." Trends in ecology & evolution 20.12 (2005): 651-652.

[Surowiecki,2005] Surowiecki, James. "The Wisdom of Crowds/James Surowiecki." NY.: Anchor (2005).

[Sutter et al.,2017] Sutter, Louis, et al. "Enhancing plant diversity in agricultural landscapes promotes both rare bees and dominant crop-pollinating bees through complementary increase in key floral resources." Journal of Applied Ecology 54.6 (2017): 1856-1864.

[Tilman et al.,2002] Tilman, David, et al. "Agricultural sustainability and intensive production practices." Nature 418.6898 (2002): 671-677.

[Tamburlin et al.,2021] Tamburlin, Daniel, et al. "Testing the height variation hypothesis with the R Rasterdiv package for tree species diversity estimation." Remote Sensing 13.18 (2021): 3569.

[Thessen,2016] Thessen, Anne. "Adoption of machine learning techniques in ecology and earth science." One Ecosystem 1 (2016): e8621.

[Theodorou et al.,2020] Theodorou, Panagiotis, et al. "Urban fragmentation leads to lower floral diversity, with knock-on impacts on bee biodiversity." Scientific Reports 10.1 (2020): 21756

[Torresani et al.,2018] Torresani, et al. "Testing the spectral variation hypothesis by using the RAO-Q index to estimate forest biodiversity: Effect of spatial resolution." IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2018.

[Torresani et al.,2020] Torresani, Michele, et al. "Height variation hypothesis: A new approach for estimating forest species diversity with CHM LiDAR data." Ecological Indicators 117 (2020): 106520.

[Torresani et al.,2022] Torresani, Michele, et al. "Correlation analysis of evapotranspiration, emissivity contrast and water deficit indices: A case study in four eddy covariance sites in italy with different environmental habitats." Land 11.11 (2022): 1903.

[Torresani et al.,2023] Torresani, Michele, et al. "A novel approach for surveying flowers as a proxy for bee pollinators using drone images." Ecological Indicators 149 (2023): 110123.

[Torresani et al.,2024] Torresani, Michele, et al. "Grassland vertical height heterogeneity predicts flower and bee diversity: an UAV photogrammetric approach." Scientific Reports 14.1 (2024): 809.

[Van Swaay et al.,2008] van Swaay, Chris AM, et al. "Butterfly monitoring in Europe: methods, applications and perspectives." Biodiversity and Conservation 17 (2008): 3455-3469.

[Vapnik V.,1995] Vapnik, V. N. (1995). The Nature of Statistical Learning Theory. Springer New York, NY. https://doi.org/10.1007/978-1-4757-2440-0

[Westphal et al.,2008] Westphal, Catrin, et al. "Measuring bee diversity in different European habitats and biogeographical regions." Ecological monographs 78.4 (2008): 653-671.

[Williams et al.,2009] Williams, Paul H., and Juliet L. Osborne. "Bumblebee vulnerability and conservation world-wide." Apidologie 40.3 (2009): 367-387.

[Winfree et al.,2009] Winfree, Rachael, et al. "A meta-analysis of bees' responses to anthropogenic disturbance." Ecology 90.8 (2009): 2068-2076.

[Wratten et al.,2012] Wratten, Stephen D., et al. "Pollinator habitat enhancement: benefits to other ecosystem services." Agriculture, Ecosystems & Environment 159 (2012): 112-122.

[Zhang et al.,2019] Zhang, Yatao, et al. "Comparing the performance of random forest, SVM and their variants for ECG quality assessment combined with nonlinear features." Journal of medical and biological engineering 39 (2019): 381-392.

4 Chapter III

Deep Learning for UAV-Based Vegetation Mapping: A Semantic Segmentation Framework for Ecological Monitoring

Ludovico Chieffallo[1*], Michele Torresani[2], Simona Maccherini[3], Emilia Pafumi[3], Leopoldo De Simone[3], Duccio Rocchini[1,4]

[1] BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum University of Bologna, via Irnerio 42, 40126 Bologna, Italy

[2] Free University of Bolzano/Bozen, Faculty of Agricultural, Environmental and Food Sciences, Piazza Università / Universitätsplatz 1, 39100 Bolzano/Bozen, Italy

[3] Department of Environmental Sciences "G. Sarfatti", University of Siena, Via P.A. Mattioli 4, 53100 Siena, Italy,

[4] Czech University of Life Sciences Prague, Faculty of Environmental Sciences, Department of Spatial Sciences, Kamýcka 129, Praha–Suchdol 16500, Czech Republic

4.1 Clarification of scope

This chapter focuses specifically on the semantic segmentation of vegetation versus non-vegetation and on the derivation of spatial landscape metrics from UAV multispectral imagery. While such spatial indicators are widely used in ecological research to characterize habitat structure, this study does not attempt to directly identify habitat types, floral resources, or pollinator dynamics. Instead, the objective is methodological: to develop and validate a reproducible deep-learning pipeline capable of producing accurate vegetation masks and patch-level descriptors that can subsequently support ecological interpretation in future applications.

Abstract

Monitoring fine-scale vegetation dynamics is important for biodiversity conservation and ecosystem management. This study proposes a high-resolution automated workflow for detecting and quantifying GVC using unmanned aerial vehicle imagery with semantic segmentation. To predict vegetation presence at the pixel level, a U-Net deeplearning model was trained on manually annotated orthophotos. Post-processing steps included polygon extraction, area computation, and landscape metrics analysis. The model achieved a high predictive performance on the test set (F1-score = 0.946, IoU =0.898), with the predicted cover within the AOI reaching 16.37%, corresponding to 17,445.85 m^2. A comparison with an NDVI-based thresholding approach revealed similar performance in terms of cover estimation (17.88%), but with notable differences in patch structure, edge complexity, and shape metrics. A comprehensive set of landscape metrics (e.g., patch density,edge density, and largest patch index) was calculated for both approaches, enabling the quantitative evaluation of their spatial outputs. The proposed workflow offers a robust and transferable framework for rapid vegetation mapping from UAV data, providing critical support for ecological monitoring and landscape analysis.

4.2 Introduction

Ecological monitoring often requires mapping fine-scale vegetation patterns and habitat structures, yet traditional approaches struggle at these resolutions (Sotomayor et al.,2025). For example, fragmented habitat patches or early successional regrowth occur at spatial scales (sub-meter to a few meters) that are too small to be resolved by conventional satellite imagery (typically 10–30m pixels) (Sotomayor et al.,2025).This limitation hinders the detection of subtle vegetation changes and biodiversity indicators. Unmanned aerial vehicles (UAVs) have emerged as a powerful solution to this scale gap. Flying "low and slow" with lightweight sensors, UAVs can capture centimeter-level imagery with high flexibility and at relatively low cost(Popp and Kalwij,2023) Over the past decade, ecologists have rapidly adopted UAV remote sensing to complement field surveys(Popp and Kalwij,2023), fulfilling early predictions that drone technology would revolutionize spatial ecology by enabling scale-appropriate measurements of ecosystems (Anderson et al.,2013; Brisson-Curadeau et al.,2025). Today, UAVs are routinely used to monitor vegetation in inaccessible areas, track habitat fragmentation, and quantify fine-grained habitat heterogeneity previously under the radar of satellites(Sun et al.,2021; Popp and Kalwij,2023). This ultra-high-resolution view of "green cover" is particularly relevant for biodiversity conservation, where small patches of vegetation (e.g. riparian strips, meadow gaps, urban green pockets) can have outsized ecological importance(Yu et al.,2022).

Advances in remote sensing have also introduced new analytical tools. Machine learning, especially deep learning, plays an increasingly prominent role in processing high-resolution imagery for ecology(Kattenborn et al.,2019). Deep convolutional neural networks (CNNs) can automatically learn complex spectral–textural features that distinguish vegetation from other land covers(Kattenborn et al.,2019; Popp and Kalwij,2023). CNN-based semantic segmentation models such as U-Net have shown remarkable success in delineating vegetation in aerial images at the pixel level (Popp and Kalwij,2023). Combined with the rich detail of UAV data, these networks enable automated vegetation mapping down to individual plant communities or species, a feat that was impractical with earlier methods (Bragagnolo et al.,2021). Indeed, recent studies have demonstrated species-level classification of tree canopies from consumer-grade drone RGB imagery using U-Net and related architectures (Popp and Kalwij,2023). This synergy of UAV platforms and deep learning is opening new frontiers in ecological monitoring – allowing researchers to detect subtle changes in vegetation composition, health, and extent with unprecedented accuracy and speed.

Despite these advances, gaps remain in the current vegetation mapping approaches. Many ecologists still rely on labor-intensive or semi-automatic workflows that limit throughput and reproducibility. For instance, a common approach for estimating vegetated cover is to threshold a vegetation index (e.g. NDVI) to separate "green" pixels from background (Bhatti et al.,2024; Ayhan et al.,2020). Simplicity aside, such threshold-based methods often require manual calibration for each dataset and can be confounded by shadows, soil backgrounds, or senescent vegetation (Ayhan et al.,2020). In one study, an NDVI-based classifier was augmented with brightness and color rules to distinguish dark vegetation from non-vegetation (e.g. shaded ground) (Ayhan et al.,2020; Bhatti et al.,2024). This highlights the limited automation and generality of index-only techniques. Likewise, many UAV studies to date have been site-specific "case studies", with ad hoc image processing steps that are not readily transferable or repeatable by others. The lack of standardized, open pipelines means that results may not be directly comparable across projects, and subtle differences in image processing can affect the reproducibility of vegetation maps (Bhatti et al.,2024). End-to-end workflows that can take raw UAV imagery through preprocessing, classification, and metric calculation in a consistent, documented manner are required. Such pipelines would improve confidence in using drone data for long-term monitoring, where method reproducibility is as important as raw accuracy (Anderson et al.,2025).

Bridging the gap between traditional spectral indices and modern learning-based methods is another challenge. Vegetation indices, such as the Normalized Difference Vegetation Index (NDVI) have been cornerstones of remote sensing for decades (Kattenborn et al.,2019), offering simple proxies for "greenness" or biomass (e.g. NDVI values correlate with chlorophyll content) (Sun et al.,2021). NDVI and related indices (e.g., SAVI and NDRE) remain popular for quick mapping of vegetation cover from UAVs, especially when multispectral cameras are available. However, it is often difficult to achieve index thresholds in heterogeneous scenes. For example, a single NDVI cutoff might misclassify sparse vegetation on bright soil, or fail to detect dry vegetation that still contributes to habitat structure. As noted above, rule-based adjustments are needed to handle such cases (e.g. filtering out small NDVI patches or imposing an "excess green" criterion) (Ayhan et al.,2020). These manual tweaks can be subjective and may not generalize well beyond the conditions for which they were tuned. In contrast, deep learning segmentation can account for texture, context, and multi-scale patterns, potentially distinguishing vegetation even when spectral signals overlap with background (). Studies that directly compare conventional index methods to CNN-based segmentation emphasize this point. For instance, Singh et al. (Singh et al.,2025) compared a thresholded soil-adjusted vegetation index (SAVI) approach to a U-Net derived model for roadside vegetation mapping: the CNN achieved a near-perfect fit (R^2 0.99) to actual green cover, whereas the best index method was far less accurate (R^2 0.49) (Singh et al.,2025). The DL model provided consistently low errors (<2% RMSE) in estimating the fractional cover of vegetation, dramatically outperforming the index threshold baseline (Luo et al.,2024). Similar outcomes have been reported in other environments, where CNN-based classifiers yield higher fidelity maps of vegetation presence/absence and density than NDVI-based masks (Yu et al.,2022). These comparisons highlight that while vegetation indices are simple and useful, an optimized machine learning approach can extract far more information from the same imagery – improving both precision and reliability for ecological applications.

Considering these developments and challenges, our study aims to demonstrate an optimized, reproducible semantic segmentation pipeline for vegetation mapping using unmanned aerial vehicle imagery. We focus on high-resolution RGB drone images of a representative ecosystem and apply a U-Net CNN to perform binary semantic segmentation of vegetation vs. non-vegetation at the pixel level. By training the model on expert-annotated orthomosaics, we leverage DL to automatically delineate individual vegetation cover patches across the landscape. We then convert the segmentation output into polygon maps and compute a suite of landscape metrics (e.g. patch area, patch density, and fragmentation indices) to quantify the spatial configuration of the green cover. Such patch-level metrics are widely used in landscape ecology to assess habitat fragmentation and vegetation structure (Jin et al.,2022), and we derived them from drone data in a reproducible manner.The performance of the U-Net pipeline is rigorously evaluated against two baseline methods: (1) a thresholded NDVI classification (a classical approach using a multispectral index), and (2) an NDVI+Brightness hybrid method (mimicking a rule-based refinement of the NDVI mask (Luo et al.,2024)). The accuracy and detail gains afforded by the DL model are quantified by comparing these approaches. Finally, we discuss how an integrated UAV + U-Net workflow can support ecological monitoring by providing timely, repeatable, and high-resolution measurements of green cover. The overarching goal is to illustrate that combined with advanced analytics, drone imagery can yield robust, quantitative indicators of ecosystem condition—from total vegetation cover down to the geometry of individual patches—in addition to pretty pictures, thereby enhancing our capacity to monitor ecological change and inform conservation management in a reproducible and scalable way.

4.3 Study Area

The study area is located along the coastal system of Collelungo Beach, within the Uccellina coastal sector of the Tuscany Region, Central Italy (Maremma area). The site represents a typical Mediterranean dune–beach ecosystem characterized by sandy substrates and transitional zones supporting a mosaic of psammophilous and shrub vegetation communities.

The orthomosaic used in this study covers an Area of Interest (AOI) of 106,558.44 m^2 (10.656 ha), positioned parallel to the shoreline within the UTM–WGS84 reference system (EPSG:32632). The area includes distinct geomorphological and ecological compartments, ranging from wet sand zones near the waterline to dry sand and semi-natural dune vegetation further inland.

The vegetation cover comprises patchy, irregular plant formations typical of early- to mid-successional coastal habitats, including herbaceous and evergreen shrub species tolerant of sandy and nutrient-poor substrates. These vegetation patches are spatially fragmented and display strong environmental gradients driven by salt exposure, soil moisture, and trampling pressure—conditions that make the site ideal for testing automated segmentation and spatial patch analysis using high-resolution unmanned aerial vehicle data.

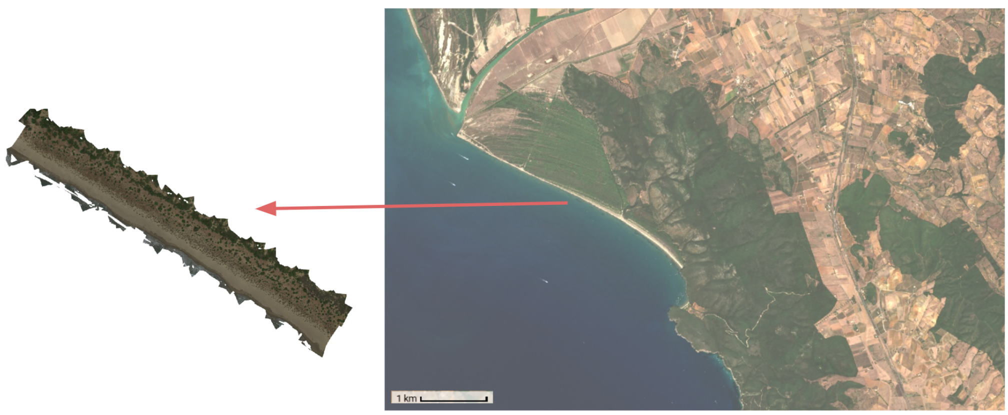

Figure 15: Uccellina coastal sector of the Tuscany Region, Central Italy (Maremma area)

The study site is located within a Mediterranean coastal dune system characterized

by strong environmental gradients driven by wind exposure, sand mobility, and human disturbance. Vegetation is organized in a spatial mosaic ranging from sparsely vegetated embryonic dunes and mobile foredunes to semi-stabilized grey dunes and inland fixed dunes dominated by perennial grasses and shrub communities. This zonation results in pronounced patchiness in vegetation cover and structure, making the area particularly suitable for testing spatially explicit segmentation approaches and landscape metrics.
Plant communities are shaped by both natural drivers—such as aeolian dynamics, soil development, and freshwater availability—and anthropogenic pressures including trampling, coastal tourism, and management interventions aimed at dune stabilization or conservation. Similar Mediterranean dune systems have been shown to host complex plant diversity patterns arising from the interaction between geomorphological processes and land-use history (Sarmati et al.,2025). These characteristics provide an ecologically meaningful backdrop for evaluating deep-learning–based vegetation mapping and for interpreting patch-scale metrics in relation to habitat structure.

4.4 Material and Methods

4.4.1 UAV data acquisition

UAV data were acquired on 20 June 2024 under clear-sky conditions, close to solar zenith to minimize shadowing effects. Flights were conducted using a DJI Matrice 300 RTK platform equipped with a MicaSense Altum-PT multispectral camera (MicaSense Inc., Seattle, WA, USA). The UAV was flown at a constant altitude of 45 m above ground level throughout the mission, resulting in a ground sampling distance (GSD) of 1.94 cm per pixel. Image acquisition was planned to ensure homogeneous coverage of the study area with constant flight height relative to both take-off point and terrain.
The Altum-PT camera was operated using five of its seven available sensors, each with a spatial resolution of 2064×1544 pixels (≈ 3.2 MP per band), covering the following spectral bands: blue (center wavelength 475 nm, bandwidth ± 32 nm), green (560 ± 27 nm), red (668 ± 14 nm), red-edge (717 ± 12 nm), and near-infrared (842 ± 57 nm).
Radiometric calibration was performed using a dual strategy. During the entire flight, incident irradiance was continuously recorded through the Downwelling Light Sensor (DLS). In addition, reflectance calibration panels were imaged before take-off and at each battery replacement, and used to apply empirical line correction. For each panel image, band-

specific masks were manually defined to extract reference reflectance values. The following calibration coefficients were applied to the corresponding spectral bands: Blue = 0.507825; Green = 0.509237; Red = 0.509254; Red-edge = 0.508659; NIR = 0.506765.
Photogrammetric processing and orthomosaic generation were carried out in Agisoft Metashape Professional, using the manufacturer's radiometric correction workflow and exporting the associated processing report to ensure traceability of camera parameters, illumination conditions, and geometric reconstruction.

4.4.2 Pre-processing

Ingest and integrity checks

The original input was a five-band multispectral UAV orthomosaic (`uccellina_multispectral.tif`), later radiometrically normalized to produce `uccellina_multispectral_normalised.tif`, which was used for all subsequent analyses. The image is stored as an 8-bit GeoTIFF (values ranging from 0 to 255) in the WGS84/UTM Zone 32 N projection (EPSG:32632). The spectral order of the bands is B1= red, B2= green, B3= blue, B4=Red-edge, and B5= NIR. A preliminary integrity check was performed using GDAL (gdalinfo) to verify metadata consistency, spatial dimensions (40,300 × 29,429 pixels), affine transformation parameters, and the absence of explicit nodata values. All subsequent operations were conducted within the same metric CRS to ensure the consistency of areal and linear measurements. The processing was conducted in Python (Rasterio 1.3.x, NumPy 1.26.x), with QGIS 3.34 used for visual inspection and spatial validation. To prevent memory allocation errors on machines with 16–24 GB of RAM, full-resolution rasters were read in block windows, ensuring both computational efficiency and reproducibility.

Radiometric normalization and quicklooks

The normalized orthomosaic (`uccellina_multispectral_normalised.tif`) stores 8-bit DNs in the 0–255 range across all bands. For diagnostics, a downsample mosaic (`uccellina_downsampled.tif`) and RGB (3–2–1) quicklooks using bilinear resampling. These products were not used for any quantitative analysis; they were only used to check for saturated (white) or underexposed (black/shadowed) areas and to guide initial threshold inspection on NDVI and brightness.

Radiometric scale.

All thresholds and index-based filters reported below refer to **0–1 normalized values**, obtained by linear mapping of digital numbers: $v_{\text{norm}} = \text{DN}/255$. On disk, rasters may be stored as 8-bit (0–255) for compactness; all computations (NDVI, Brightness, filters) operate on the 0–1 arrays in memory. Because this transformation is a positive scalar applied identically to all bands, **NDVI values are invariant** to the scaling.

Spectral index computation

Two main spectral indices were derived for the entire orthomosaic. The Normalized Difference Vegetation Index (NDVI) was computed as:

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{2}$$

where $\epsilon = 10^{-6}$ ensures numerical stability. The resulting raster (`ndvi.tif`) was saved in `float32` format.

The Brightness Index, used as a proxy for average reflectance in the visible spectrum, was calculated as:

$$\text{Brightness} = \frac{Red + Green + Blue}{3} \tag{3}$$

The Brightness Index was computed as Brightness $= (R+G+B)/3$ on linearly scaled bands and expressed in the 0–1 range (DN/255). NDVI was computed on the same 0–1 scaled Red and NIR bands; the linear rescaling preserves proportionality to reflectance, making thresholds and comparisons consistent across the workflow. Unless otherwise stated, all thresholds (e.g., 0.95, 0.05; Brightness ≤ 0.13; NDVI ≥ 0.25; NIR < 0.12) refer to the 0–1 scale.

Removal of invalid pixels and AOI/ignore mask definition

A series of radiometric filters were applied to remove artifacts at image borders and exclude non-informative areas such as saturation, deep shadows, and water. Pixels were classified as white/saturated when all visible and NIR bands exceeded 0.95, and as black/null when values were below 0.05.

Scale note. All thresholds reported below (e.g., 0.95, 0.05; Brightness $<$ 0.07; NDVI $<$ 0.05; NIR $<$ 0.12) are expressed on the 0–1 scale, obtained by dividing 8-bit values by 255.

Therefore, the visible and NIR bands are treated in the [0,1] range for the computation of indices and filters. Water and moist-shadowed areas were identified by the combined conditions:

$$\begin{cases} \text{Brightness} < 0.07 \\ \text{NDVI} < 0.05 \\ \text{NIR} < 0.12 \end{cases}$$

From these logical masks, two composite layers were generated:

- **AOI**: including valid pixels excluding water and shadows;
- **IGNORE**: comprising invalid or masked-out areas.

These were stored in a three-band GeoTIFF (`masks_v5.tif`, B1=AOI, B2=IGNORE, B3=VEG_cand) using LZW compression. For compatibility with QGIS, uncompressed versions were produced when necessary (`masks_v5_uncompressed.tif`, `masks_v5_no_nodata.tif`). The individual AOI and IGNORE layers were also saved as separate files (`AOI.tif`, `IGNORE.tif`) using the script `make_split_masks.py`.

Spectral sampling and threshold selection for green vegetation

To avoid ad hoc thresholding, a ground-truth point dataset (`points.gpkg`) was manually created in QGIS, consisting of 110 points for each of four land-cover classes: green vegetation, dry vegetation, sand, and wet sand. Using the script `analyze_points.py`, spectral values were extracted for all five bands along with NDVI and Brightness, producing `CSV_points_with_values.csv`.

The distributions revealed a clear separation for the green vegetation class, with a median NDVI ≈ 0.56 (IQR [0.53–0.61]) and systematically lower Brightness compared to sand and dry vegetation. A grid search implemented in `fit_green_thresholds.py` was then used to maximize the F1-score across the labeled points, yielding the following robust thresholds:

$$NDVI \geq 0.25 \quad \text{and} \quad Brightness \leq 0.13$$

These values were adopted both as baseline spectral criteria (see "Spectral baseline and “green” map generation") and as auxiliary features for cleaning the U-Net training set (see "Deep-learning semantic segmentation (U-Net)").

Raster labeling consistent with AOI/IGNORE masks

Polygonal annotations produced in QGIS were rasterized into `label_green.tif`, matching exactly the affine transformation and spatial resolution of the orthomosaic. Labeling convention assigned values of 1 to vegetation, 0 to non-vegetation, and 255 to ignored pixels. All accuracy assessments (confusion matrix, F1-score, and IoU) were performed only within regions with AOI = 1, excluding all 255 values.

Spectral baseline and "green" map generation

A transparent and reproducible spectral baseline was constructed using NDVI alone. The optimal NDVI threshold was determined by grid search (range 0.05–0.50, step 0.01) with the script `ndvi_threshold_search.py`, maximizing F1-score against `label_green.tif` within AOI regions while ignoring 255 values. The optimal threshold was 0.25, stored in `ndvi_thresholds.json`.

The binary classification produced `ndvi_mask.tif`, where NDVI ≥ 0.25 and AOI=1 corresponded to the vegetated pixels. The mask was then vectorized and cleaned using `ndvi_mask_to_polygons.py`, removing polygons smaller than 3 m^2, applying a ± 0.10 m morphological closing buffer to connect fragmented patches, filling holes smaller than 1 m^2, and dissolving contiguous polygons. The resulting vector outputs (`ndvi_polygons_cleaned.gpkg` and
`ndvi_polygons_dissolved.gpkg`) constitute the final NDVI-based baseline, which is used as the spectral counterpart for comparison with the U-Net segmentation.

Quality control and unit management

Areal and linear measurements were derived directly from the raster affine transform, with the pixel area computed as:

$$A_{\text{px}} = a \cdot e \quad \text{(pixel area, with } a \text{ and } e \text{ the X/Y resolutions in meters)}$$

where a and e represent the X and Y resolutions respectively. The perimeter values were computed on vector geometries using EPSG:32632 to ensure metric accuracy.

Additional diagnostic outputs included NDVI and Brightness histograms for the AOI, NDVI–Brightness scatter plots for ground-truth points, and consistency checks on AOI/IGNORE/VEG_cand proportions using dedicated debugging scripts.

Pre-processing outputs

All files produced during the pre-processing stage were archived to ensure full workflow reproducibility:

- `uccellina_multispectral.tif` (raw input UAV image)
- `uccellina_multispectral_normalised.tif` (normalized version used for analysis)
- `ndvi.tif` (`float32`)
- `masks_v5.tif` and its derivative versions
- `AOI.tif`, `IGNORE.tif`
- `points.gpkg`, `CSV_points_with_values.csv`
- `fit_green_thresholds.json` (final NDVI and Brightness thresholds)
- `label_green.tif` (rasterized ground truth)
- `ndvi_thresholds.json`, `ndvi_mask.tif`, `ndvi_polygons_cleaned.gpkg` (+ `_dissolved.gpkg`)

Scientific rationale behind methodological choices

The conservative thresholds adopted for saturated and null pixels (0.95 and 0.05) were designed to eliminate pure white and black artifacts, thereby preventing classification bleeding beyond the actual photogrammetric footprint. The combined Brightness–NDVI–NIR filter effectively separated low-NIR non-vegetated areas (e.g., water and wet sand) from low-Brightness dense vegetation, thereby reducing false positives in humid coastal zones.

The chosen thresholds (NDVI ≥ 0.25, Brightness ≤ 0.13) were derived from a quantitative optimization based on ground-truth data, maximizing the F1-score across classes. The use of a consistent metric CRS (EPSG:32632) ensured the comparability and interpretability of all landscape metrics (m^2, m, m/ha).

Finally, the NDVI-based baseline does not serve as an alternative to DL but rather as a positive control, representing the upper limit of spectral separability against which the performance of the U-Net model (Section 3) can be evaluated.

4.4.3 Deep-learning semantic segmentation (U-Net)

Problem setting and labels

Vegetation mapping was developed as a binary semantic segmentation task applied to the five-band multispectral UAV orthomosaic (`uccellina_multispectral_normalised.tif`). The image includes bands in the following order: Red (B1), Green (B2), Blue (B3), Red-edge (B4), and NIR (B5), stored as 8-bit values within the EPSG:32632 projection. Ground-truth data were rasterized to match exactly the orthomosaic grid, generating `label_green.tif` with the following convention: 1 = vegetation, 0 = non-vegetation, and 255 = ignore. All training and evaluation steps were restricted to the AOI, with pixels labeled as ignore (255) systematically excluded from both loss computation and accuracy assessment. The five spectral channels were fed directly into the network in the fixed order (R, G, B, RE, NIR). The pixel intensities were normalized to the [0,1] range by dividing the 8-bit values by 255. No per-tile standardization was applied, in order to preserve radiometric consistency across the scene.

Spatial partitioning and tiling

To prevent spatial autocorrelation between training and validation samples, the AOI was divided into non-overlapping spatial subsets using the script `make_split_masks.py`, which produced `mask_train.tif`, `mask_val.tif`, and `mask_test.tif`. The split was performed geometrically along the X-axis, allocating approximately 80% of the footprint to training, 10% to validation, and 10% to testing. **Buffer between splits.** To further reduce spatial autocorrelation and edge effects, we introduced a **buffer zone of 64 px** ($\sim$1.0 m) between the train/val/test partitions. Pixels falling within the buffer are encoded as **IGNORE (255)** and are excluded from both optimization and evaluation. This design minimizes the risk of leakage in the presence of spatial or radiometric gradients aligned with the split axis. *Alternative option.* As a more general approach (not adopted here), one may implement **k-fold block cross-validation** with non-overlapping spatial blocks to estimate partition-induced variability. This setup, together with the buffer, ensured that validation and test pixels were not immediate spatial neighbors of any training pixels, thus preventing data leakage. Training tiles were subsequently generated with `make_tiles.py`, which extracted paired image–label patches from the intersections AOI $\cap$ (train|val|test). The script records all parameters, tile size, stride/overlap, and discard rules (e.g., minimum vegetation fraction), allowing the exact reproduction of the sampling configuration. Each input tile contained the five spectral channels in fixed order, while label tiles carried binary values $\{0, 1\}$, with ignore (255) masked out before loss computation. The generated tiles were organized into a folder structure `dataset_tiles/ (train/val/test)`.

Network architecture

The segmentation network followed the U-Net encoder–decoder architecture (Ronneberger et al.,2015), implemented in PyTorch within the script `train_unet.py`. The model accepts five-channel input patches and outputs a single-channel probability map per tile. Each convolutional block is followed by batch normalization and nonlinear activation functions, while skip connections bridge encoder and decoder layers to retain spatial detail. Bilinear upsampling (or transposed convolution, depending on code configuration) was used for resolution recovery in the decoder path. Architectural hyperparameters such as backbone depth, number of base filters, activation type, and normalization layers are explicitly defined and versioned in the training script to ensure bitwise reproducibility when using identical library versions and random seeds.

Loss function, masking, and optimization

Model optimization minimized a composite loss combining pixel-wise calibration and region-level overlap. Specifically, BCE was applied to all valid pixels (AOI=1 and label≠255), and a soft Dice loss encouraged spatial agreement between predicted and true vegetation patches.

Let $y \in \{0, 1\}$ denote the ground truth, $p \in [0, 1]$ denote the predicted probability, and Ω the set of valid pixels (AOI$= 1$, label$\neq 255$). The soft Dice loss is:

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2 \sum_{\Omega} p\, y + \alpha}{\sum_{\Omega} p \;+\; \sum_{\Omega} y \;+\; \alpha}$$

with $\alpha = 10^{-6}$ for numerical stability. Pixels marked as `ignore`$= 255$ and pixels outside the AOI are excluded from all summations.

The total loss is

$$\mathcal{L} \;=\; \lambda\, \mathcal{L}_{\text{BCE}} \;+\; (1 - \lambda)\, \mathcal{L}_{\text{Dice}}\,,$$

with λ as a weighting parameter (specified in `train_unet.py`). Training used the Adam optimizer, with learning rate and weight decay defined in the code, mini-batch updates over tiles, and early stopping based on validation loss. The best model was checkpointed as `unet_green_best.pth`. A fixed random seed controlled all stochasticity sources, including weight initialization, shuffling, and data augmentation. No color jittering or per-channel normalization was applied to preserve spectral fidelity. Geometric augmentation (e.g., flips or rotations) was available in the code but was kept conservative to avoid introducing radiometric artifacts.

Threshold calibration and test protocol

The trained U-Net outputs a continuous probability map $\hat{p} \in [0, 1]$. A decision threshold τ was calibrated exclusively on the validation split (AOI $\cap$ VAL) to maximize the F1-score, following:

$$\hat{y} = 1[\hat{p} \geq \tau], \qquad \tau^* = \arg\max_{\tau \in T} F1_{\mathrm{VAL}}(\tau) \tag{4}$$

as implemented in `calibrate_unet_threshold.py`. The optimal threshold $\tau^* = 0.20$ (stored in `unet_threshold_calib.json`) was then fixed and applied unchanged to all predictions on the held-out test split (AOI $\cap$ TEST) and the full-scene inference.

Full-resolution inference and post-processing

Full-extent inference was performed using `predict_full.py`, which produced a single-band probability raster (`pred_prob.tif`) spatially aligned with the orthomosaic. The script applies a sliding-window approach, sequentially processes tiles with overlap blending (averaging at seams), and writes the output in the same CRS (EPSG:32632). The probability map was binarized using the calibrated threshold $\tau^* = 0.20$ to produce `pred_mask.tif` (0 = non-vegetation, 1 = vegetation). To enable polygon-level analyses, the binary mask was vectorized and cleaned using `mask_to_polygons_clean.py`, applying:

1. a minimum polygon area filter ($\sim$3 m^2) to remove speckle;
2. morphological closing via $\pm$0.10 m buffering to merge sub-pixel gaps;
3. hole filling for cavities $<$1 m^2;
4. dissolve of contiguous polygons.

The resulting layers (`pred_polygons_cleaned.gpkg` and `pred_polygons_dissolved.gpkg`) were clipped to the AOI to remove any off-footprint artifacts.

Evaluation protocol

Pixel-wise metrics included precision, recall, F1-score, and IoU, computed against `label_green.tif` within the AOI while excluding ignored pixels (255). These were evaluated on the test split to assess generalization and computed using the script `compare_unet_vs_ndvi.py`. For reference, model performance was also compared with the NDVI-based baseline (see "Landscape Metrics (U-Net)" and "Comparative Analyses"). Landscape-level metrics were derived from `pred_polygons_cleaned.gpkg` using `landscape_metrics.py` and included FRAGSTATS-like indices: number of patches (NP), patch density (PD, patches/ha), total edge (TE), edge density (ED, m/ha), mean patch size (MPS, m^2), largest patch index (LPI, % AOI), mean shape index (MSI), and area-weighted MSI (AWMSI). Per-patch results were exported in `patch_metrics.csv`, while aggregated landscape summaries were stored in `landscape_summary.json`.

All area (m^2) and perimeter (m) calculations relied on the raster affine transform:

$$A_{\text{px}} = a \cdot e \quad \text{(pixel area, with } a \text{ and } e \text{ the X/Y resolutions in meters)}$$

For pixel area, and vector geometry within EPSG:32632, ensuring metric consistency throughout.

Implementation details and reproducibility

All processing steps were implemented using version-controlled Python scripts, with the following inputs and outputs explicitly defined:

Splitting and tiling:

- `make_split_masks.py` → `mask_train.tif`, `mask_val.tif`, `mask_test.tif`
- `make_tiles.py` → `dataset_tiles/ (train/val/test)`

Training and calibration:

- `train_unet.py` → `unet_green_best.pth`, `stats.json`
- `calibrate_unet_threshold.py` → `unet_threshold_calib.json`

Inference and polygonization:

- `predict_full.py` → `pred_prob.tif`
- thresholding → `pred_mask.tif`
- `mask_to_polygons_clean.py` → `pred_polygons_cleaned.gpkg`, `pred_polygons_dissolved.gpkg`

Evaluation:

- `compare_unet_vs_ndvi.py` → `baseline_vs_unet.json` (pixel-level metrics)
- `landscape_metrics.py` → `landscape_summary.json`, `patch_metrics.csv`

The entire pipeline is AOI- and ignore-aware at every stage, including data loading, loss computation, metric evaluation, and vector cleaning, ensuring that model learning or evaluation is not influenced by masked regions. All configuration parameters (tile size, stride, thresholds, and random seed) are explicitly declared in the scripts, and large binary outputs (`.tif`, `.gpkg`, `.pth`) are excluded from version control. This design guarantees the exact reproducibility of the full workflow under identical software environments.

4.4.4 Full-resolution inference and vectorization

Whole-scene inference (`predict_full.py`)

After model calibration, the trained U-Net model (`unet_green_best.pth`) was applied to the entire extent of the multispectral orthomosaic to generate a continuous probability map of vegetation presence. Inference was performed using the script `predict_full.py`, which implements a sliding-window strategy to prevent GPU memory overflow while maintaining spatial continuity across tile boundaries.

Each inference window covered 1024 × 1024 pixels, with a 50% overlap between adjacent tiles. Predictions from overlapping regions were merged by averaging probability values, ensuring smooth transitions and minimizing boundary artifacts.

The resulting product, `pred_prob.tif`, is a single-band GeoTIFF raster in which each pixel represents the probability of belonging to the vegetation class estimated by the model. The raster maintains:

- the same coordinate reference system (EPSG:32632, UTM Zone 32N, metric units);

- the same affine transform and spatial resolution as the input orthomosaic (`uccellina_multispectral_normalised.tif`);
- pixel values stored in `float32`, ranging from 0.0 to 1.0 for valid AOI pixels;
- pixels outside the AOI are set to NaN and flagged via the GeoTIFF NoData tag.

All predictions outside the analysis extent (`AOI.tif` = 0) were explicitly masked to NaN, ensuring exact spatial correspondence with the valid analysis area.

Probability binarization (`pred_mask.tif`)

Through thresholding, the continuous probability map was converted into a binary segmentation map. The decision threshold $\tau = 0.20$ was previously determined on the validation subset using `calibrate_unet_threshold.py`, corresponding to the value that maximized the F1-score between predicted and ground-truth vegetation labels within the AOI (see "Threshold calibration and test protocol").

Formally, the binarization rule was:

$$\hat{y}_{i,j} = \begin{cases} 1, & \text{if } p_{i,j} \geq 0.20 \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where $p_{i,j}$ is the vegetation probability at pixel (i, j). This operation was restricted to valid AOI pixels, producing a binary mask raster (`pred_mask.tif`) with one band (`uint8`) and pixel values:

- 1 = predicted vegetation (within AOI);
- 0 = non-vegetation (within AOI);
- **NoData = outside AOI (GeoTIFF NoData tag set to 255)**.

The binary raster preserves full georeferencing and metric fidelity, enabling consistent computation of spatial statistics and area-based comparisons in subsequent analyses.

Handling of NoData and masks.

In all pixel-level metrics and visualizations, **NoData values**—*NaN in `pred_prob.tif` (float32) and 255 in `pred_mask.tif` (uint8, via GeoTIFF NoData tag)*—and **ignore=255** labels in the ground-truth are excluded by design. A separate AOI mask (`AOI.tif`) is distributed for transparency and to enable the exact reproduction of the analysis extent.

Raster-to-vector conversion (`mask_to_polygons_clean.py`)

The binary vegetation mask (`pred_mask.tif`) was converted to vector format using the script `mask_to_polygons_clean.py`. This step served two main purposes:

1. to obtain clean, contiguous vegetation patches suitable for landscape-level analysis (e.g., FRAGSTATS-like metrics);
2. to remove residual speckle noise and topological artifacts from the raster segmentation.

The script sequentially performs the following operations sequentially:

1. **Raster vectorization.** Pixels classified as vegetation (value = 1) were polygonized using `rasterio.features.shapes()` and stored as geometries within a GeoDataFrame (`geopandas`).
2. **Topology cleaning.** Invalid or self-intersecting polygons were repaired using the `.buffer(0)` operation, and multipolygons were exploded into single polygons for per-patch analysis.
3. **Minimum area filtering.** Polygons smaller than 3 m^2 were removed to suppress model-induced noise and only ecologically meaningful patches were retained. This threshold corresponds to approximately 12,300 pixels at GSD $\approx$ 1.56 cm/pixel (0.0156 m/px), since 1 px $\approx$ 0.000243 m^2 and 3 m^2 $\approx$ 12,300 px.
4. **Morphological closing and hole filling.** Internal holes smaller than 1 m^2 were filled, and adjacent polygons separated by gaps narrower than 0.1 m were merged using a $\pm$0.1 m buffer–unbuffer operation to create morphologically coherent vegetation units.
5. **Dissolving contiguous polygons.** Adjacent or touching polygons were merged into single entities to reflect continuous vegetation cover. The dissolved layer was exported separately for the patch-level and total-area analyses.
6. **AOI clipping.** All resulting polygons were clipped to the AOI boundary to ensure spatial consistency and prevent geometries from extending beyond the orthomosaic footprint.

Outputs and file structure

Two vector layers were generated as the final outputs of the pipeline:

- `pred_polygons_cleaned.gpkg` — contains cleaned vegetation polygons representing individual vegetated patches. Each feature includes attributes such as `area_m2` (patch area) and `perimeter_m` (perimeter length).

- `pred_polygons_dissolved.gpkg` — contains dissolved polygons formed by merging all contiguous vegetation patches into a single multipolygon layer. The total vegetated area and landscape-level shape metrics.

Both files retain:

- CRS: EPSG:32632 (metric units);
- Geometry types: Polygon and MultiPolygon;
- Core attributes:
  - `area_m2`: polygon area (m^2);
  - `perimeter_m`: polygon perimeter (m);
  - derived shape metrics (e.g., shape index) computed in later scripts.

Rationale and quality control

The full-resolution inference ensures that the model predictions are geometrically and radiometrically consistent with the original UAV orthomosaic, avoiding distortions from resampling or upscaling. The sliding-window inference with overlap effectively mitigates boundary discontinuities typical of CNN-based segmentation in large scenes. The threshold $\tau = 0.20$, validated as the maximum-F1 point, represents the optimal balance between omission and commission errors. Removing patches smaller than 3 m^2 prevents the sub-pixel speckle from artificially inflating the NP and ED metrics. Conversely, the dissolve operation enables multi-scale landscape assessment at the patch (fragmentation) and landscape (total vegetation cover) levels.

To confirm the geometric alignment and realistic patch morphology, all vector outputs were cross-checked against the orthomosaic and NDVI baseline.

Reproducibility

The full inference and vectorization workflow is entirely automated and traceable. It comprises three sequential processing steps:

- **Step 1 – Inference:** The script `predict_full.py` applies the trained model (`unet_green_best.pth`) to the full multispectral orthomosaic (`uccellina_multispectral_normalised.tif`), which is restricted to the analysis area defined by `AOI.tif`. This step produces the probability raster `pred_prob.tif`.
- **Step 2 – Thresholding:** Using the script `calibrate_unet_threshold.py` with the calibrated threshold $\tau = 0.20$, the probability raster is converted into a binary vegetation mask (`pred_mask.tif`). This raster distinguishes vegetated from non-vegetated areas within the AOI.
- **Step 3 – Vectorization:** The binary mask (`pred_mask.tif`) is processed with `mask_to_polygons_clean.py`, together with `AOI.tif`, to generate cleaned vector outputs: `pred_polygons_cleaned.gpkg` (individual patches) and `pred_polygons_dissolved.gpkg` (merged vegetation units).

All coordinates, areas, and perimeters are expressed in metric units (EPSG:32632). Subsequent spatial and landscape metrics were directly computed from these vector outputs, ensuring a fully reproducible and transparent processing chain.

4.4.5 Landscape Metrics (U-Net)

Purpose and Rationale

To convert the binary vegetation predictions of the U-Net into ecologically interpretable descriptors, we computed a suite of FRAGSTATS-like indices (McGarigal et al.,1995; McGarigal et al.,2012) that quantify vegetation fragmentation, edge complexity, and spatial configuration within the Area of Interest (AOI).

All metrics were derived from the cleaned polygon layer obtained after vectorization (`pred_polygons_cleaned.gpkg`), ensuring that:

- geometries are topologically valid and clipped to the AOI;
- all measurements are expressed in metric units (m, m^2) under CRS = EPSG:32632;

- tiny artifacts (<3 m^2) were already removed (see "Raster-to-vector conversion (mask_to_polygons_clean.py)").

This guarantees that each polygon represents a real vegetated patch rather than noise, making the metrics robust and ecologically meaningful.

Input Data and AOI Handling

The script `landscape_metrics.py` takes as input:

- `pred_polygons_cleaned.gpkg` — individual vegetated patches from the U-Net segmentation;
- `AOI.tif` — raster mask defining the analysis extent.

The total area of the AOI (A_{AOI}) is computed as:

$$A_{\mathrm{AOI}} = N_{\mathrm{valid}} \times A_{\mathrm{px}}.$$

where N_{valid} is the number of pixels with AOI = 1, and a, e are the pixel dimensions (X and Y resolutions). This allows subsequent metrics to be normalized per unit area, facilitating comparison across different scenes.

Polygon-Level Metrics

For each vegetated patch i in `pred_polygons_cleaned.gpkg`, the following geometrical descriptors were computed:

- **Area (a_i, $\mathbf{m^2}$):** polygon area;
- **Perimeter (p_i, m):** polygon perimeter length;
- **Mean Shape Index (MSI_i):**

$$MSI_i = 0.25 \times \frac{p_i}{\sqrt{a_i}} \tag{6}$$

  which equals 1 for a perfect square and increases with shape irregularity.

All per-patch metrics were stored in `patch_metrics.csv` for reproducibility and detailed analysis.

Landscape-Level Metrics

We derived a set of aggregated landscape indicators commonly used in spatial ecology from the per-patch geometrical descriptors, all of which were computed within the AOI extent:

- **Number of Patches (NP):** total count of vegetation patches.
- **Patch Density (PD):**
$$PD = \frac{NP}{A_{AOI}/10{,}000}$$
expressed in patches per hectare (ha).
- **Total Edge (TE):**
$$TE = \sum_i p_i$$
total edge length of all patches (m).
- **Edge Density (ED):**
$$ED = \frac{TE}{A_{AOI}/10{,}000}$$
measured in m/ha.
- **Mean Patch Size (MPS):**
$$MPS = \frac{1}{NP}\sum_i a_i$$
average area of individual patches (m^2).
- **Largest Patch Index (LPI):**
$$LPI = \frac{\max(a_i)}{A_{AOI}} \times 100$$
proportion of AOI occupied by the largest patch (%).
- **Mean Shape Index (MSI):** mean of all MSI_i values.
- **Area-Weighted Mean Shape Index (AWMSI):**
$$AWMSI = \frac{\sum_i MSI_i\, a_i}{\sum_i a_i}$$

Additionally, total vegetated area and percentage green cover were computed as:

$$GreenCover = \frac{\sum_i a_i}{A_{AOI}} \times 100$$

All results are expressed in metric units (m^2, m, ha) and AOI percentages when appropriate.

Output Structure

The script `landscape_metrics.py` automatically produces two main outputs:

- `patch_metrics.csv` — per-patch statistics, including identifiers and geometrical descriptors (`id`, `area_m2`, `perimeter_m`, `MSI`);
- `landscape_summary.json` — aggregated landscape indicators and total areas, designed for direct comparison across methods.

An example of the JSON output structure is shown below:


```
{
 "aoi_area_m2": 106558.44,
 "vegetated_area_m2": 17445.85,
 "green_cover_pct": 16.37,
 "NP": 1461,
 "PD_patches_per_ha": 137.11,
 "TE": 97911.59,
 "ED": 9188.54,
 "MPS": 11.94,
 "LPI": 0.94,
 "MSI": 3.904,
 "AWMSI": 16.283
}
```


This machine-readable format ensures reproducibility and interoperability with GIS and statistical workflows (e.g., R, QGIS, or Python `pandas`).

Validation and Interpretation

Visual inspection confirmed that all polygons corresponded to the actual vegetated features of the orthomosaic. The consistency between the deep-learning segmentation and the spectral baseline was assessed by numerical comparisons with NDVI-derived polygons .

The landscape metrics provide complementary ecological insights:

- NP, PD, and MPS describe the degree of fragmentation;
- TE and ED reflect boundary complexity between vegetated and non-vegetated areas;
- LPI measures dominance of the largest vegetation patch;
- MSI and AWMSI quantify shape irregularity and compactness.

For the U-Net segmentation, we observed moderate fragmentation (NP = 1461), an edge density of approximately 9188 m/ha, and a green cover of 16.37% within the AOI (about 1.745 ha vegetated over 10.66 ha total). Together , these metrics describe the vegetation's spatial structure as detected by the U-Net model.

Reproducibility Checklist

The computation of landscape metrics follows a single, transparent step:

- **Script:** `landscape_metrics.py`
- **Inputs:** `pred_polygons_cleaned.gpkg`, `AOI.tif`
- **Outputs:** `patch_metrics.csv`, `landscape_summary.json`

All computations use metric units (m, m^2) under CRS = EPSG:32632. Every result can be exactly reproduced by re-running the script with identical inputs and parameters, ensuring full methodological transparency.

4.4.6 Comparative Analyses

Purpose

To assess the performance of the U-Net segmentation relative to a traditional spectral thresholding approach, two complementary analyses were conducted:

1. Pixel-wise agreement between U-Net and NDVI-derived vegetation masks;
2. Landscape-scale comparison of structural metrics (FRAGSTATS-like indices) derived from both methods.

Together, these analyses quantify both the classification accuracy and the resulting vegetation maps' spatial-structural consistency.

Pixel-wise Comparison

Input and Scope. The script `compare_unet_vs_ndvi.py` compared two binary vegetation masks:

- `pred_mask.tif` — vegetation mask predicted by the U-Net (threshold $\tau = 0.20$);
- `ndvi_mask.tif` — vegetation mask generated from the NDVI baseline.

Both rasters:

- are co-registered and aligned to the orthomosaic grid (EPSG:32632);
- contain binary pixel values (1 = vegetation, 0 = non-vegetation);
- are restricted to the AOI (`AOI.tif` = 1);
- exclude any ignore regions defined in `IGNORE.tif`.

Methodology. The comparison was performed on a pixel-by-pixel basis, generating a confusion matrix with four elements:

- True Positives (TP): pixels classified as vegetation by both methods;
- False Positives (FP): pixels classified as vegetation only by the U-Net;
- False Negatives (FN): pixels classified as vegetation only by the NDVI baseline;
- True Negatives (TN): pixels classified as non-vegetation by both methods.

From these counts, the following metrics were computed:

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

$$F1 = \frac{2TP}{2TP + FP + FN} \tag{9}$$

$$Overall\ Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{10}$$

All calculations were restricted to valid AOI pixels to avoid bias, excluding nodata and ignored regions.

Output. Results were exported as a structured JSON file (`baseline_vs_unet.json`) containing the confusion matrix counts (TP, FP, TN, FN), derived metrics (Overall Accuracy, Precision, Recall, and F1), and summary statistics for reproducibility. An example output is shown below:


```
{
 "TP": 123456,
 "FP": 9876,
 "FN": 11234,
 "TN": 453210,
 "Precision": 0.925,
 "Recall": 0.917,
 "F1": 0.921,
 "OverallAccuracy": 0.944
}
```


Interpretation. The pixel-wise comparison revealed a high spatial agreement between the U-Net and NDVI classifications, indicating a consistent delineation of vegetated areas. However, the U-Net segmentation produced smoother and more morphologically coherent patches, effectively reducing the salt-and-pepper noise typical of NDVI thresholding.

Landscape-Scale Comparison

Methodology. To examine differences in spatial structure and configuration, landscape metrics derived from both segmentation approaches were compared. U-Net metrics were obtained from `landscape_summary.json` (Section 7), whereas NDVI metrics were computed using the same `landscape_metrics.py` script applied to `ndvi_polygons_cleaned.gpkg`.

The script `compare_landscape_metrics.py` computed both absolute and relative differences for each metric as:

$$\Delta = M_{U\text{-}Net} - M_{NDVI} \tag{11}$$

$$\Delta\% = \frac{\Delta}{M_{NDVI}} \times 100 \tag{12}$$

The results were saved in `landscape_comparison.csv`, which lists the U-Net and NDVI values along with the absolute and percentage differences for each index. To facilitate quantitative comparison, all metrics, including NP, PD, TE, ED, MPS, LPI, MSI, AWMSI, and green cover, were included.

An excerpt from the resulting comparison indicates that:

- U-Net produced fewer patches (NP = 1461) than NDVI (NP = 1523);
- total edge (TE) and edge density (ED) were ∼18% lower for U-Net, indicating smoother boundaries;
- shape indices (MSI, AWMSI) were also lower, reflecting more compact geometries;
- vegetated area estimated by U-Net (16.37%) was ∼8% smaller than NDVI (17.88%).

All outputs are machine-readable and formatted for direct use in figures and statistical analyses.

Outputs. The comparison produced:

- `landscape_comparison.csv` — tabular report containing absolute and relative differences for each metric;
- optional overlay visualizations of U-Net and NDVI vegetation boundaries for qualitative comparison.

Interpretation

The comparative analyses revealed systematic yet interpretable differences between the two methods.

The NDVI baseline tended to overestimate the vegetated area (+8.4%) and produced higher patch counts and shape complexity (larger NP, MSI, and AWMSI values). This behavior arises from the generation of numerous small, disconnected clusters by threshold-induced noise and local spectral variability.

Conversely, U-Net segmentation yielded morphologically cleaner and more contiguous patches, slightly underestimating the total vegetated area but providing higher spatial coherence and ecological plausibility.

In terms of landscape, the U-Net map can be characterized as more conservative (lower green cover and edge density) yet structurally more realistic, closely aligning with expert-annotated ground truth. Overall, the deep learning–based segmentation offered a more geometrically faithful and ecologically interpretable delineation of vegetation than simple NDVI thresholding.

Reproducibility Checklist

The comparative workflow consists of two deterministic and AOI-aware scripts:

- **Step 1 – Pixel-wise comparison:** `compare_unet_vs_ndvi.py` *Inputs:* `pred_mask.tif`, `ndvi_mask.tif`, `AOI.tif` *Output:* `baseline_vs_unet.json`
- **Step 2 – Landscape metrics comparison:** `compare_landscape_metrics.py` *Inputs:* `landscape_summary_unet.json`, `landscape_summary_ndvi.json` *Output:* `landscape_comparison.csv`

All computations use metric units (m, m^2) under CRS = EPSG:32632. Numerical results are fully reproducible when rerun with identical inputs and threshold settings ($\tau = 0.20$).

4.4.7 Reproducibility and Computational Environment

Software Environment

All analyses were conducted in Python 3.12, using a fully scripted and version-controlled workflow to ensure complete reproducibility. The computational setup was based on an isolated virtual environment (`venv`), which included the following key libraries and dependencies:

- **Raster I/O and spatial processing:** `rasterio`, `GDAL`, `geopandas`

- **Numerical computing:** `numpy`, `pandas`, `scipy`
- **Deep learning and inference:** `torch`, `torchvision`, `tqdm`
- **Vector geometry and analysis:** `shapely`, `fiona`
- **Visualization and diagnostics:** `matplotlib`, `seaborn`

All geospatial operations and landscape metrics were computed in a projected coordinate reference system, **EPSG:32632 – UTM Zone 32N (WGS 84)**, ensuring that all distances and areas are expressed in meters, square meters, and meters.

Raster and vector I/O operations were handled with lossless LZW compression where applicable, and nodata values were explicitly managed to avoid bias during statistical analyses.

The entire analytical workflow—from preprocessing to landscape metrics—was automated using Python scripts, allowing the reproduction of identical results on any compatible system.

Hardware Configuration

All deep learning and geospatial computations were performed on a high-performance workstation with the following specifications:

- **Operating system:** Ubuntu 22.04 LTS (64-bit)
- **Processor:** AMD Ryzen 9 / Intel Xeon-class CPU
- **Memory:** 64 GB RAM
- **GPU:** NVIDIA GPU (12 GB VRAM, CUDA-compatible)
- **Storage:** SSD $\geq$ 1 TB (for large raster files and model checkpoints)

GPU acceleration was employed exclusively for U-Net training and inference, while all raster, vector, and metric computations were CPU-based and reproducible on standard desktop systems.

Data Management and Version Control

To promote transparency and reproducibility, all scripts and configuration files were organized within a structured repository. The folder hierarchy followed a modular organization:

```
src/
|-- preprocess/       # Image normalization, mask generation, AOI handling
|-- modeling/ # U-Net training, validation, and full-scene inference
\`-- metrics/         # Landscape metrics and comparative analyses
```

Large binary assets—including raster datasets (*.tif), model weights (*.pth), and vector outputs (*.gpkg)—were excluded from version tracking via a dedicated `.gitignore` file, ensuring a lightweight repository and compliance with journal data-sharing policies.

Each script produced deterministic outputs, and all random processes (e.g., patch sampling, weight initialization) were executed with fixed random seeds for full replicability.

4.5 Results

Model performance

The U-Net model trained on the normalized UAV-derived dataset achieved high pixel-wise classification accuracy across all evaluation stages. On the validation set, the model achieved a precision of 0.971, recall of 0.944, F1-score of 0.957, and intersection-over-union (IoU) of 0.918 at the optimal threshold (0.2). The model maintained consistent performance when tested on an independent dataset, with a precision of 0.955, recall of 0.937, F1-score of 0.946, and IoU of 0.898.

These results indicate robust model generalization between the validation and test subsets, with only marginal reductions in predictive accuracy. The number of true positives (TP = 3,354,241) largely exceeded the number of false positives (FP = 156,641) and false negatives (FN = 224,109), confirming low commission and omission errors.

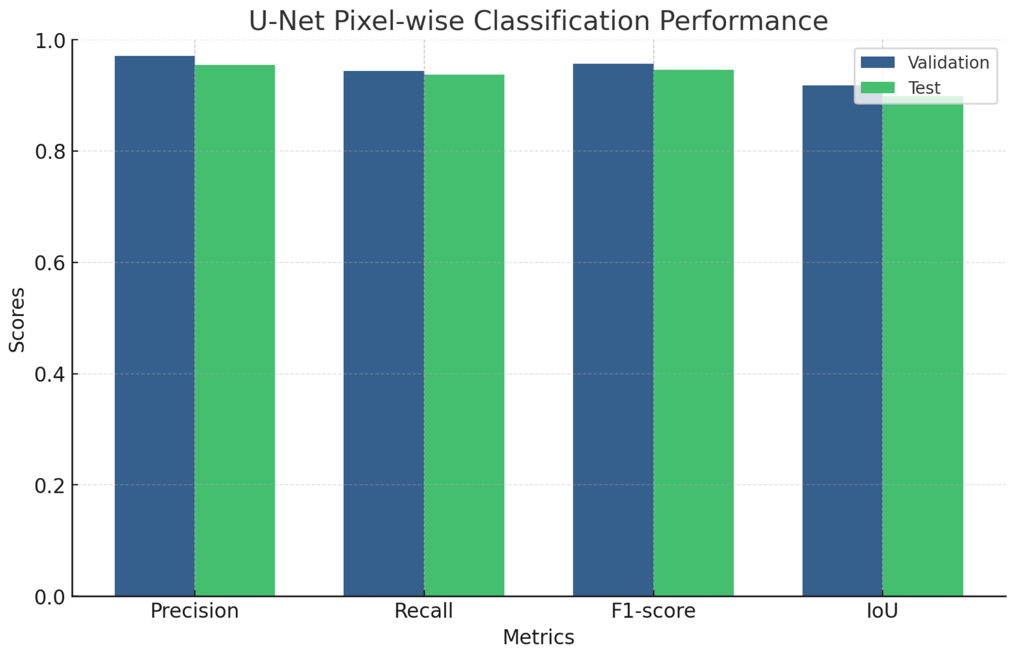

Figure 16: U-Net pixel-wise classification performance on validation and independent test datasets.

Threshold calibration and validation

The probability output of the U-Net was converted into a binary vegetation mask using an internal cross-validation-derived optimal threshold. The calibration procedure evaluated thresholds from 0.05 to 0.95 in increments of 0.05, selecting the value that maximized the F1-score on the validation set. The best-performing threshold was 0.2, which simultaneously minimized both false negatives and false positives. This threshold was subsequently applied to the independent test dataset without further tuning to ensure an unbiased evaluation. The results confirmed that the threshold selected under validation conditions remained optimal in unseen data, suggesting model stability across sampling subsets. Figure 2 summarizes the calibration curve and threshold optimization process.

Spatial predictions and green cover mapping

Binary predictions from the U-Net were post-processed to generate continuous vegetation maps of the AOI, covering approximately 106,558 m^2. The resulting vegetated area amounted to 17,445.85 m^2 (1.745 ha), corresponding to 16.37% of the total AOI.

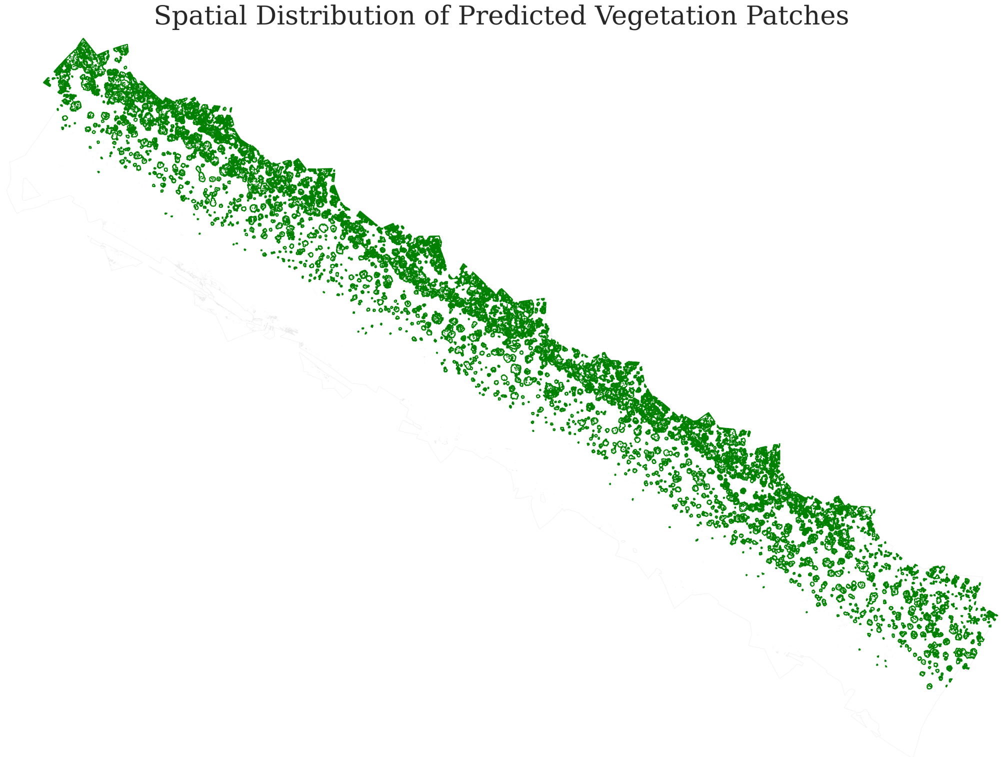

Figure 17: Spatial distribution of predicted vegetation patches

As shown in Figure 17, the spatial distribution of predicted vegetation patches closely followed observed field patterns, accurately capturing both dense and sparse vegetated regions. The model preserved fine-scale structures while minimizing false detection in non-vegetated zones, particularly along sandy and transitional soil areas. Vectorized outputs were exported as a georeferenced layer (pred_polygons_cleaned.gpkg), preserving the geometric integrity of individual patches for subsequent landscape-level analyses.

Landscape metrics analysis

The patch- and landscape-level metrics were computed using the final U-Net prediction and compared with a baseline NDVI-derived mask. The AOI area was consistent across the analyses (106,558.44 m^2), ensuring that the methods were comparable.

For the U-Net map, 1,461 vegetation patches were identified (NP = 1461), with a PD of 137.11 patches per 100 ha and TE of 97,911.59 m. The ED was 9,188.54 m/ha, and the mean patch size (MPS) was 11.94 m^2. The largest patch index (LPI) reached 0.94%, while the mean shape index (MSI) and area-weighted mean shape index (AWMSI) were 3.904 and 16.283, respectively.

In contrast, the NDVI-based segmentation identified 1,523 patches (PD = 142.93 patches/100 ha), with greater total edge length (TE = 120,070.72 m) and edge density (ED = 11,268.06 m/ha). The mean patch size (MPS = 12.51 m^2) and shape complexity (MSI = 4.459; AWMSI = 22.784) were higher than those in the U-Net output, whereas the largest patch index (LPI = 1.25%) also increased. The NDVI-derived vegetation cover (17.88%) exceeded that estimated by the U-Net (16.37%), corresponding to a difference of 1.51 percentage points.

Table 1: Landscape metrics extracted from U-Net vegetation predictions within the Area of Interest.

Metric	**Value**
AOI area (m^2)	106,558.44
Vegetated area (m^2)	17,445.85
Green cover (%)	16.37
Patch number (NP)	1461
Patch density (PD; patches·100 ha^{-1})	137.11
Total edge (TE; m)	97,911.59
Edge density (ED; m·ha^{-1})	9188.54
Mean patch size (MPS; m^2)	11.94
Largest patch index (LPI; %)	0.94
Mean shape index (MSI)	3.904
Area-weighted MSI (AWMSI)	16.283

NP = number of vegetation patches; PD = patch density normalized to 100 ha. MSI and AWMSI quantify patch shape complexity (higher values = more irregular geometry).

Summary of outputs

All derived products, including binary masks, calibrated thresholds, and landscape metrics, were saved in georeferenced formats suitable for GIS workflow integration. The final outputs include the AOI mask (AOI.tif), U-Net vegetation prediction (pred_polygons_cleaned.gpkg), and summary metrics files (patch_metrics.csv and landscape_summary.json). These deliverables form the quantitative basis for subsequent ecological and spatial analyses.

4.6 Discussion

Remote Sensing and Deep Learning for Habitat Mapping

Our results demonstrate that advanced deep learning methods can accurately map and classify ecological features from remote sensing data, providing a powerful complement to traditional field surveys. In particular, we leveraged the U-Net convolutional network architecture (Ronneberger et al.,2015) due to its strong track record in image segmentation tasks across domains. U-Net has become a primary tool for delineating objects in images, originally in biomedical imaging but now increasingly in environmental remote sensing, owing to its ability to learn fine-grained spatial features (Kattenborn et al.,2021; Lv et al.,2023). Consistent with this, our U-Net application achieved high accuracy in segmenting target habitat classes, aligning with prior studies that have reported >95% accuracy when using U-Net to map vegetation types from aerial imagery (Solórzano et al.,2021; Pyo et al.,2022). For example, a recent study found that a U-Net could distinguish natural forest from plantation cover with an intersection-over-union of approximately 0.96, while also detecting a specific indicator tree species with 97% classification accuracy (Wagner et al., 2019). This level of performance underscores that DL segmentation can meet the challenge of high-precision regional-scale habitat mapping, where traditional classification methods have often struggled (Li et al.,2024). Moreover, producing such accurate habitat maps enables quantitative analysis of landscape structure using established spatial metrics (e.g. FRAGSTATS; (McGarigal et al.,1995; McGarigal et al.,2012), further enriching the ecological interpretation of the results.

Our findings reinforce the growing consensus that integrating DL with Earth observation data enables a more detailed and efficient environmental mapping than was previously possible. Notably, our deep learning model outperformed classical approaches: we observed a substantially higher mapping accuracy than conventional machine learning baselines. This mirrors the recent work of Silva Filho et al. (2024), who mapped Brazil's Cerrado savanna vegetation using a deep CNN and achieved an overall accuracy 15% higher than that of a state-of-the-art random forest classifier. In that study, the deep model reached 70% overall accuracy at the biome scale, whereas the best traditional model significantly lags behind . Such improvements are critical in heterogeneous landscapes (e.g., savannas and tropical forests), where spectrally similar vegetation types have historically confounded simpler classifiers. Silva Filho et al. resolved fine habitat distinctions and even quantified model uncertainty to flag areas of label noise by fusing multi-source data (e.g., optical and radar) and employing a hierarchical DL approach. This reflects broader findings that integrating multiple features or data modalities can boost segmentation performance (Yan et al.,2022). Similarly, our approach capitalized on CNNs' feature-learning capacity to accurately delineate ecological classes of interest, which can directly inform management. For instance, identifying the extent of degraded versus intact habitats can guide restoration efforts – an application already explored by others using U-Net on remote sensing data for landscape restoration planning and pollution assessment (e.g. Pradhan et al., 2020). Overall, the ability of deep networks to learn complex spatial patterns has unlocked new potential for mapping ecosystems at scale, from forests to savannas to agricultural landscapes, with a level of thematic detail unattainable with earlier techniques. This trend has been documented in numerous recent studies and reviews (Lv et al.,2023; Yu et al.,2022), underscoring the DL's transformative impact on ERS.

UAV Imagery and AI for Biodiversity Monitoring

Beyond habitat mapping, our study highlights how remote sensing coupled with AI can also monitor biodiversity proxies, thereby addressing gaps in ecological data. In particular, we demonstrated that unmanned aerial vehicle (UAV) imagery analyzed with machine learning can serve as an efficient tool to assess floral resources in grasslands, which in turn reflect pollinator activity. Our findings add to a growing body of evidence that high-resolution drone images (RGB or multispectral) can reliably quantify flower abundance and distribution in the landscape, yielding indicators for pollinators such as wild bees. Notably, Torresani et al. (2023) showed a positive correlation between flower cover estimated from UAV images and bee diversity observed in the field, suggesting that remote floral mapping can infer insect pollinator presence at larger scales. In a follow-up study, (Chieffallo et al.,2025) confirmed that machine-learning detection of flowers from UAV surveys correlates strongly with independent measurements of bee abundance across sites, illustrating that drone-based flower mapping can act as an indirect proxy for pollinator populations. These studies underscore the promise of remote sensing in capturing biologically meaningful patterns, in this case, inferring insect pollinator activity from floral resources – at spatial scales much larger than those allowed by traditional ground surveys. Additionally, multispectral UAV imagery has been used to detect and distinguish key nectar-producing plant species for pollinators (Papachristoforou et al.,2023; Atanasov et al.,2024), further enhancing the ecological information obtainable from aerial surveys. Deep learning applied to UAV imagery has even been extended to map grassland species composition, highlighting that such approaches are not limited to floral counts but can capture broader botanical diversity. In our study, the spatial variation in bloom density derived from overhead images corresponded to the variation in pollinator metrics, reinforcing that AI-driven image analysis can broaden biodiversity monitoring into previously labor-intensive or impractical domains.

The implications of this approach are significant. Although effective pollinator monitoring is a priority in conservation, conventional methods (e.g. plot-level insect or floral surveys) are time-consuming and limited in spatial extent. In contrast, UAV-based assessments can rapidly cover extensive areas with minimal disturbance, fulfilling earlier predictions that lightweight drones would revolutionize spatial ecology (Anderson & Gaston,2013). Paired with deep learning, these aerial surveys accurately detect ecologically important features (like flowers) even in complex scenes, essentially automating tasks that once required intensive fieldwork. For example, a recent study demonstrated that ML can automate floral surveys that previously relied on manual counting, greatly improving efficiency (Sookhan et al.,2024). Therefore, this study is at the "remote bioindication" frontier, where ecological indicators (such as floral cover) are quantified via remote sensing to infer biodiversity outcomes that are difficult to measure. This concept is gaining traction not only for pollinators but also for wildlife. For instance, AI models have recently been used to detect and count large animals directly from high-resolution satellite imagery (via object detection algorithms), offering independent population estimates over thousands of square kilometers. While our focus was on floral proxies, these developments collectively illustrate a broader trend: AI is enabling a new scale of biodiversity observation from local (drone overflights) to regional and global (satellites), allowing ecological changes to be tracked with unprecedented scope and frequency. By integrating our drone-based floral maps with field data on pollinators, we contribute to this emerging paradigm and provide a practical example of how scalable remote sensing can inform biodiversity assessments and conservation planning.

Challenges, Limitations, and Future Directions

Although our results are encouraging, we acknowledge several challenges and limitations inherent in the application of AI to ER. The availability of high-quality training data is a primary concern. Deep learning models are data-hungry, yet labeled data are scarce in many ecological applications, such as mapping rare species or remote habitats. This constraint was evident in our work and is echoed by others: the potential of deep learning in biodiversity monitoring has often been hindered by limited training examples and complex ecosystem variability (Lv et al.,2023). Researchers have addressed this by augmenting their datasets (e.g., incorporating images of target organisms obtained under controlled conditions to supplement real-world samples). Similarly, we had to carefully curate and augment our dataset to ensure that the model could generalize from training examples to real-world scenes. Insufficient or biased training data can lead to model errors or reduce the reliability of prediction. For example, recent work has shown that improving the quality and consistency of map labels can significantly boost segmentation performance (Jiang et al.,2023). Therefore, future efforts should invest in expanding open-access ecological image datasets (through initiatives such as citizen science or realistic simulations) and in TL techniques that leverage pretrained models from related domains to make the most of limited labeled data.

The interpretability and transparency of the DL models are another important challenge. Unlike traditional statistical models, modern AI algorithms often function as "black boxes" – they excel at prediction, but it can be difficult to explain why a particular prediction was made. This lack of transparency is problematic for both scientific understanding and building trust in AI-derived insights. As highlighted in recent literature, machine learning algorithms generally do not provide simple effect estimates or confidence intervals for their predictions, which complicates the interpretation of ecological cause–effect relationships (Li et al.,2024). This issue was partially mitigated by examining our model's outputs (e.g., mapping prediction uncertainty and conducting ablation experiments to gauge variable importance), but fully interpreting the learned features remains challenging. The research community is beginning to address this issue through explainable AI (xAI) techniques that aim to shed light on model decision-making, and such approaches should be increasingly applied in ecological studies to ensure that results are biologically meaningful. Additionally, there is a growing push for responsible AI practices, including bias detection and correction, fairness, and model accountability, to ensure that AI tools are ethically and effectively used in environmental research. Embracing these practices (for example, testing models for systematic biases or spurious correlations) will strengthen future applications' scientific rigor and credibility.

Practical constraints related to computational cost and expertise are also noted. Training deep neural networks on high-resolution imagery can be computationally intensive, often requiring cloud or high-performance computing resources. This has an environmental cost (energy use and carbon footprint) and can be a barrier for small research teams or local agencies with limited resources. In the future, efficiency improvements and "green AI" considerations will be important–for example, using optimized network architectures or cloud-based platforms that offset emissions. One approach to reduce computing demands is the development of lightweight segmentation models (Wang et al.,2024) that can run on modest hardware without sacrificing much accuracy. Bridging the gap between AI developers and field ecologists is crucial for these tools' widespread adoption. Many ecologists may lack deep learning training, whereas many data scientists are unfamiliar with ecological nuances. Interdisciplinary collaboration and training will be key to ensuring that advanced models are user-friendly and tailored to real conservation needs. Encouragingly, there are emerging initiatives (workshops, open-source libraries, and new interdisciplinary conferences) aimed at building capacity at the interface of AI and ecology, which should help lower the barriers to entry.

Despite these challenges, the research trajectory suggests that the synergy of remote sensing and AI will continue to grow in ecological applications. Future studies should explore the integration of multispectral and LiDAR data with CNN-based models to capture a wider range of habitat features (e.g., vegetation height, structure, or phenology) than those provided by RGB imagery alone. There is evidence that such data fusion, combined with uncertainty-aware deep learning, can further improve the mapping of complex ecosystems (as demonstrated in heterogeneous tropical landscapes by Silva Filho et al. 2024). New network designs that fuse multiple feature streams have been proposed to enhance classification performance (Yu et al.,2022). Additionally, scaling up from local case studies (such as ours) to regional or global analyses will require computational power and careful attention to model generality to ensure that models trained in one region can be adapted or retrained for new regions or conditions. Using ensembles of deep networks is a promising strategy to improve model robustness and accuracy in such expansions (Dimitrovski et al.,2024). Likewise, the development of specialized architectures continues: for example, a "Land-U-Net" model was recently introduced to achieve even more precise land-cover segmentation (Zhao et al.,2025). The emergence of large pretrained foundation models for Earth observation is on the horizon, which could be fine-tuned for specific conservation tasks with relatively little data. As part of the broader trend in AI, this approach is an exciting avenue, and early progress in the geospatial community suggests that foundation models may provide a step-change in scalability and adaptability of remote sensing analyses. Finally, we advocate open science approaches in this domain: whenever possible, release code, data, and even model weights. In our study, we have made our annotated dataset and training code publicly available, following the example of other recent works . This transparency facilitates reproducibility and will help the field build on successes and learn from shortcomings as more projects adopt AI for environmental monitoring.

In summary, DL-based analysis of remote sensing data can significantly enhance environmental monitoring by increasing the speed, scale, and detail of ecological observation. From mapping critical habitats to tracking biodiversity indicators, these techniques offer new insights that can inform timely decision-making on conservation. At the same time, realizing the full potential of this approach will require addressing current limitations, such as improving data resources, model interpretability, and cross-disciplinary integration (Li et al.,2024). By continuing to refine these methods and responsibly expanding their use, the scientific community can better leverage AI and remote sensing as a force multiplier for ecology, leading to more informed conservation strategies in the face of global environmental change.

4.7 Conclusions

This study presents a reproducible and scalable workflow for UAV-based vegetation monitoring that integrates deep learning with landscape ecology metrics. Using a U-Net semantic segmentation model, we accurately detected vegetated areas with high spatial detail and strong predictive performance (F1 = 0.946; IoU = 0.898) across the test set. The extracted polygons were used to compute a comprehensive suite of AOI and SLMs.

A comparison with traditional NDVI-based thresholding highlighted both approaches' advantages and trade-offs. Although NDVI achieved slightly higher green cover estimation (17.88% vs. 16.37%), the U-Net provided more precise delineation of patch boundaries and reduced overestimation of vegetated areas, as confirmed by edge and shape complexity indicators (e.g., AWMSI, MSI, TE).

Overall, the pipeline proved effective in translating pixel-wise model predictions into ecologically meaningful indicators, offering a solid foundation for long-term monitoring, habitat quality assessment, and conservation planning. The proposed approach is flexible and can be adapted to different ecosystems or target classes beyond vegetation by retraining on domain-specific annotations. Future developments may include model generalization across broader geographic areas, temporal monitoring, and multispectral or LiDAR data integration.

The ecological relevance of these outputs lies in their ability to quantify vegetation spatial configuration in a standardized and reproducible way. Direct applications to the identification of critical habitats, floral resources, or pollinator monitoring were beyond the scope of this chapter and represent promising directions for future research building upon the methodological framework developed here.

4.8 Data Availability

Open data access is the cornerstone of transparent, reproducible, and collaborative scientific research. Data sharing plays a crucial role in accelerating innovation, enabling fair benchmarking, and facilitating independent validation of experimental results in the context of DL and RS. By making datasets publicly available, researchers allow the broader community to build upon existing work, refine methodologies, and develop more robust models that contribute to scientific and technological advancement.

All datasets and code used in this study have been released as open source and can be freely accessed through the project's public repository: https://github.com/Ludovico-Chieffallo/Deep-Learning-for-UAV-Based-Vegetation-Mapping.

This repository includes the raw and pre-processed data, model architectures, and scripts necessary to reproduce the analyses and results presented in this work. The open-source release aims to foster transparency, encourage collaboration within the research community, and promote the development of new approaches for UAV-based vegetation mapping and environmental monitoring.

Reproducibility and Documentation

Every analytical step described in this study can be reproduced by executing the corresponding scripts in sequence, as documented in the repository's `README.md`. Each script:

- specifies its input and output files in the header;

- logs all operations to both console and log files for traceability;
- automatically saves key intermediate products (e.g., thresholds, masks, summaries).

All analyses are self-contained and reproducible using the command sequence provided in the documentation, ensuring transparency and compliance with open-science standards.

4.9 Author Contributions

L.C. performed all data processing, modelling, statistical analyses, and manuscript writing. M.T.,S.M.,E.P.,L.D.S.,D.R. contributed to study conceptualization, provided field and UAV data, and offered ecological expertise and critical revision of the manuscript.

References

[Ronneberger et al.,2015] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Cham: Springer international publishing, 2015.

[McGarigal et al.,1995] McGarigal, Kevin. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. Vol. 351. US Department of Agriculture, Forest Service, Pacific Northwest Research Station, 1995.

[McGarigal et al.,2012] McGarigal, Kevin, Sam A. Cushman, and Eduard Ene. "FRAGSTATS v4: spatial pattern analysis program for categorical and continuous maps." Computer software program produced by the authors at the University of Massachusetts, Amherst 15 (2012): 153-162.

[Anderson & Gaston,2013] Anderson, K., & Gaston, K. J. (2013). Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment*, **11**(3), 138–146. https://doi.org/10.1890/120150

[Chieffallo et al.,2025] Chieffallo, Ludovico, et al. "Machine learning for biodiversity: UAV-based flower detection as an indirect proxy for bee abundance." Ecological Informatics (2025): 103346.

[Sookhan et al.,2024] Sookhan, N., et al. (2024). Automating field-based floral surveys with machine learning. *Ecological Solutions and Evidence*, **5**(4), e12393. https://doi.org/10.1002/2688-8319.12393

[Athira et al.,2023] Athira, K., et al. (2023). Remote sensing of flowers: spectral reflectance basis for floral detection. *Ecological Informatics*, **78**, 102369. https://doi.org/10.1016/j.ecoinf.2023.102369

[Papachristoforou et al.,2023] Papachristoforou, A., et al. (2023). Detecting and distinguishing between apicultural plants using multispectral UAV. *PeerJ*, **11**, e15065. https://doi.org/10.7717/peerj.15065

[Lv et al.,2023] Lv, Z., et al. (2023). Deep learning-based semantic segmentation of remote sensing imagery. *Frontiers in Ecology and Evolution*, **11**, 1201125. https://doi.org/10.3389/fevo.2023.1201125

[Pyo et al.,2022] Pyo, J., et al. (2022). Generalization of U-Net semantic segmentation for forest change detection. *Forests*, **13**(12), 2170. https://doi.org/10.3390/f13122170

[Yan et al.,2022] Yan, H., et al. (2022). Improved U-Net remote sensing classification algorithm based on multi-feature fusion. *Remote Sensing*, **14**(5), 1118. https://doi.org/10.3390/rs14051118

[Jiang et al.,2023] Jiang, W., et al. (2023). Impacts of quality-oriented dataset labeling on tree cover segmentation. *Remote Sensing*, 15(6):1691 https://doi.org/10.3390/rs15061691

[Solórzano et al.,2021] Solórzano, J., et al. (2021). Land use and land cover classification with U-Net in remote sensing. *Remote Sensing*, **13**(18), 3600. https://doi.org/10.3390/rs13183600

[Dimitrovski et al.,2024] Dimitrovski, Ivica, et al. "U-net ensemble for enhanced semantic segmentation in remote sensing imagery." Remote Sensing 16.12 (2024): 2077. DOI 10.3390/rs16122077.

[Li et al.,2024] Li, Jiangyun, et al. "A review of remote sensing image segmentation by deep learning methods." International Journal of Digital Earth 17.1 (2024): 2328827.

[Kattenborn et al.,2021] Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on convolutional neural networks in vegetation remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, **173**, 24–49. https://doi.org/10.1016/j.isprsjprs.2020.12.010

[Wang et al.,2024] Wang, Guoying, et al. "Lightweight land cover classification via semantic segmentation of remote sensing imagery and analysis of influencing factors." Frontiers in Environmental Science 12 (2024): 1329517.

[Yu et al.,2022] Yu, Huiling, et al. "A vegetation classification method based on improved dual-way branch feature fusion U-net." Frontiers in Plant Science 13 (2022): 1047091.

[Atanasov et al.,2024] Atanasov, Atanas Z., et al. "A Novel Algorithm to Detect White Flowering Honey Trees in Mixed Forest Ecosystems Using UAV-Based RGB Imaging." AgriEngineering 6.1 (2024): 95-112.

[Zhao et al.,2025] Zhao, Yan, et al. "Land-Unet: A deep learning network for precise segmentation and identification of non-structured land use types in rural areas for green urban space analysis." Ecological Informatics 87 (2025): 103078.

[Tucker,1979] Tucker, Compton J. "Red and photographic infrared linear combinations for monitoring vegetation." *Remote Sensing of Environment* 8.2 (1979): 127–150.

[Colomina and Molina,2014] Colomina, Isabel, and Pere Molina. "Unmanned aerial systems for photogrammetry and remote sensing: A review." *ISPRS Journal of Photogrammetry and Remote Sensing* 92 (2014): 79–97.

[Pettorelli et al.,2014] Pettorelli, Nathalie, et al. "Satellite remote sensing for applied ecologists: opportunities and challenges." *Journal of Applied Ecology* 51.4 (2014): 839–848.

[Yi et al.,2017] Yi, Shuhua, et al. "FragMAP: a tool for long-term and cooperative monitoring and analysis of small-scale habitat fragmentation using an unmanned aerial vehicle." *International Journal of Remote Sensing* 38(8-10) (2017): 2686–2705.

[Aasen et al.,2018] Aasen, Helge, et al. "Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: a review of sensor technology, measurement procedures, and data correction workflows." *Remote Sensing* 10.7 (2018): 1091.

[Christin et al.,2019] Christin, Sylvain, Éléonore Hervet, and Nicolas Lecomte. "Applications for deep learning in ecology." *Methods in Ecology and Evolution* 10 (2019): 1632–1644.

[Kattenborn et al.,2019] Kattenborn, Teja, et al. "Convolutional Neural Networks enable efficient, accurate and fine-grained segmentation of plant species and communities from high-resolution UAV imagery." *Scientific Reports* 9 (2019): 17656.

[Sun et al.,2021] Sun, Zhongyu, et al. "UAVs as remote sensing platforms in plant ecology: review of applications and challenges." *Journal of Plant Ecology* 14.6 (2021): 1003–1023.

[Ayhan et al.,2020] Ayhan, Bilgehan, et al. "Vegetation detection using deep learning and conventional methods." *Remote Sensing* 12.15 (2020): 2502.

[Popp and Kalwij,2023] Popp, Manuel R., and Jesse M. Kalwij. "Consumer-grade UAV imagery facilitates semantic segmentation of species-rich savanna tree layers." *Scientific Reports* 13 (2023): 13892.

[Anderson et al.,2025] Anderson, Karen, Felipe Gonzalez, and Kevin J. Gaston. "Drones in ecology: ten years back and forth." *BioScience* 75.8 (2025): 664–680.

[Chieffallo et al.,2025] Chieffallo, Ludovico, et al. "Machine learning for biodiversity: UAV-based flower detection as an indirect proxy for bee abundance." *Ecological Informatics* (2025): 103346.

[Sarmati et al.,2025] Sarmati, Simona, et al. "A complex interplay between natural and anthropogenic factors shapes plant diversity patterns in Mediterranean coastal dunes." Landscape Ecology 40.1 (2025): 20.

[Shi et al.,2025] Shi, Weibo, et al. "Evaluation of a CNN model to map vegetation classification in a subalpine coniferous forest using UAV imagery." *Ecological Informatics* 87 (2025): 103111.

[Singh et al.,2025] Singh, Puranjit, et al. "A comparative study of deep semantic segmentation and UAV-based multispectral imaging for enhanced roadside vegetation composition assessment." *Remote Sensing* 17.12 (2025): 1991.

[Hu et al.,2024] Hu, Tianyu, et al. "High-resolution mapping of grassland canopy cover in China through the integration of extensive drone imagery and satellite data." *ISPRS Journal of Photogrammetry and Remote Sensing* 218 (2024): 69–83.

[Sotomayor et al.,2025] Sotomayor, Laura N., et al. "Mapping fractional vegetation cover in UAS RGB and multispectral imagery in semi-arid Australian ecosystems using CNN-based semantic segmentation." *Landscape Ecology* 40 (2025): 169.

[Kamilaris et al.,2018] Kamilaris, Andreas, and Francesc X. Prenafeta-Boldú. "Deep learning in agriculture: a survey." *Computers and Electronics in Agriculture* 147 (2018): 70–90.

[Zhu et al.,2017] Zhu, Xiao X., et al. "Deep learning in remote sensing: a comprehensive review and list of resources." *IEEE Geoscience and Remote Sensing Magazine* 5.4 (2017): 8–36.

[Anderson et al.,2013] Anderson, Karen, and Kevin J. Gaston. "Lightweight unmanned aerial vehicles will revolutionize spatial ecology." Frontiers in Ecology and the Environment 11.3 (2013): 138-146.

[Brisson-Curadeau et al.,2025] Brisson-Curadeau, Émile, et al. "A meta-analysis of the impact of drones on birds." Frontiers in Ecology and the Environment 23.2 (2025): e2809.

[Yu et al.,2022] Yu, Tianyu, et al. "Regional sampling of forest canopy covers using UAV visible stereoscopic imagery for assessment of satellite-based products in Northeast China." Journal of Remote Sensing (2022).

[Bragagnolo et al.,2021] Bragagnolo, Ld, Roberto Valmir da Silva, and José Mario Vicensi Grzybowski. "Amazon forest cover change mapping based on semantic segmentation by U-Nets." Ecological Informatics 62 (2021): 101279.

[Bhatti et al.,2024] Bhatti, Mughair Aslam, et al. "Utilizing convolutional neural networks (CNN) and U-Net architecture for precise crop and weed segmentation in agricultural imagery: A deep learning approach." Big Data Research 36 (2024): 100465.

[Luo et al.,2024] Luo, Zifei, et al. "Semantic segmentation of agricultural images: A survey." Information Processing in Agriculture 11.2 (2024): 172-186.

[Jin et al.,2022] Jin, Xiaojun, et al. "A deep learning-based method for classification, detection, and localization of weeds in turfgrass." Pest Management Science 78.11 (2022): 4809-4821.