Validated methods for testing reliability of landscape metrics based biodiversity indicators

Deliverable D1.4

27 March 2024

SHOWCASE
SHOWCASing synergies between agriculture, biodiversity and Ecosystem services to help farmers capitalising on native biodiversity

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 862480.

Prepared under contract from the European Commission
Grant agreement No. 862480
EU Horizon 2020 Research and Innovation action

Project acronym:	**SHOWCASE**
Project full title:	**SHOWCASing synergies between agriculture, biodiversity and Ecosystem services to help farmers capitalising on native biodiversity**
Start of the project:	November 2020
Duration:	60 months
Project coordinator:	Prof. David Kleijn Wageningen University
Deliverable title:	[Validated methods for testing reliability of landscape metrics based biodiversity indicators]
Deliverable n°:	D1.4
Nature of the deliverable:	Report
Dissemination level:	Public
WP responsible:	D1.4
Lead beneficiary:	UNIBO
Citation:	Torresani, M., Chieffallo, L. & Rocchini, D. (2023). *Validated methods for testing reliability of landscape metrics based biodiversity indicators*. Deliverable D1.4 EU Horizon 2020 SHOWCASE Project, Grant agreement No 862480.
Due date of deliverable:	36
Actual submission date:	37
Submission revision:	41

Deliverable status:

Version	Status	Date	Author(s)
1.0	Final	27-11-2023	Michele Torresani,Ludovico Chieffallo,Duccio Rocchini UNIBO/UNIBZ
2.0	Revision	27-03-2024	Michele Torresani,Ludovico Chieffallo,Duccio Rocchini UNIBO/UNIBZ

The content of this deliverable does not necessarily reflect the official opinions of the European Commission or other institutions of the European Union.

Contents

Abstract

The decline in global biodiversity, particularly in agricultural landscapes, poses a significant threat to ecosystems and human well-being. Bees, essential pollinators, have experienced a marked decline, impacting food security and ecosystem stability. To address the challenges of large-scale bee surveying, we explore, through a multi-faceted, multi-scale approach the relationship between habitat/landscape metrics de- rived from remote sensing information and pollinator diversity within different Experimental Biodiversity Areas (EBAs) in The Netherlands. At the local scale, we made use of UAV images to assess flower cover as a proxy for bee abundance and test vegetation grassland het- erogeneity as a proxy for flower and bee diversity. At the regional scale, we employed Sentinel-2 data for habitat classification to link habitat diversity with pollinator diversity. Our results indicate that RGB UAV images, coupled with machine learning algorithms like Random Forest and Neural Network, effectively estimate flower cover, showing positive correlations with in-situ bee abundance, species richness, and diversity. The choice of algorithm and spatial resolution plays a crucial role in capturing accurate ecological dynamics. Additionally, UAV-derived data on vegetation height heterogeneity, assessed through different indices and resolutions, exhibit positive correlations with flower and bee diversity, emphasizing the importance of habitat structure. At the regional scale, the Random Forest classification of forested areas using Sentinel-2 data reveals significant positive correlations with bee abundance and richness. This highlights the influence of forested habitats on shaping bee communities within EBAs. In summary, our task underscores how remote sensing technologies can propel ecological research forward, unraveling the intricate connections between habitat features and pollinator diversity. The fusion of local and regional data not only provides key insights into shaping conservation strategies but also enriches our comprehension of the intricate dynamics within the project EBAs. On a broader scale, our results aid in biodiversity assessment assisting stakeholders, (e.g., ecologists and farmers) for a better understand of biodiversity and ecological processes.

1 Introduction

The decline of global biodiversity poses a critical threat to both ecosystems and human well-being, with agricultural landscapes experiencing a particularly drastic reduction in biodiversity over the past decades [11, 23]. This decline, driven by factors like habitat fragmentation and intensified farming practices, has far-reaching consequences on ecosystem functioning [14]. Among the myriad species affected, bees, vital pollinators, have faced a marked decline in distribution, abundance, and richness at both national and local scales [30, 2]. This decline is attributed to habitat loss, fragmentation, climate change, and the loss of host plants [25]. Bees, in providing essential pollination services, play a pivotal role in the productivity of global food crops and the seed set of wild plants [12, 18]. This ecosystem service contributes significantly to worldwide food production, valued at over 150 billion euros annually [5, 6]. The decline of bee populations thus has profound implications for food security, biodiversity, and ecosystem stability. While mitigation measures and drivers of bee decline are increasingly understood, the actual extent of this decline, especially at large spatial scales, remains inadequately explored. Large-scale surveying of bees faces challenges due to the difficulty in species identification, making it less feasible for laypersons compared to standardized monitoring of butterflies [28]. Consequently, trends in bee abundance are often based on small-scale studies, underestimating population trends [2]. Standardized monitoring proposals in EU member states aim to address this gap [20]. However, due to the cost and coverage limitations of in-situ monitoring, there is a need for complementary approaches. One such approach involves inferring trends in bee pollinators from trends in flower cover and species richness.

Recent advancements in remote sensing and Earth observation offer promising avenues for large-scale biodiversity estimation [27, 22]. Satellite remote sensing (SRS) provides extensive coverage, ecological information for biodiversity estimation at broad scales [3], while Unmanned Aerial Vehicles (UAVs) equipped with optical cameras have emerged as a cost-effective solution for ecological purposes such as in the estimation of vegetation properties, species distribution, and invasive species mapping [7, 10, 1] at a local scale.

In this context, our task employs a multi-faceted and multi-scale approach, integrating data from diverse remote sensing platforms to advance our understanding of biodiversity indicators for pollinator diversity. Specifically, we leverage the potential of remote sensing data at different scales, including UAV optical data for local assessments and Sentinel-2 data for regional analysis in order to assess biodiversity indicators for pollinator diversity within the Dutch Experimental Biodiversity Areas (EBAs).

These are the objectives of our task computed at different scales using different remote sensing data:

• Local scale: assessing flower cover as a proxy for bee abundance using UAV images. In the first phase of our task, we focus on assessing flower cover as a proxy for bee abundance. The fundamental step involves utilizing images from a UAV and distinguishing flower pixels from non-floral surfaces through differences in spectral signatures [8]. We opted for RGB images, given their cost-effectiveness, ready availability, and user-friendly nature. RGB cameras, often integrated into commercial drones, provide a practical and easily reproducible approach for ecological classification analysis.

• Local scale: testing vegetation grassland heterogeneity as a proxy for flower diversity and pollinator diversity using UAV data. In this second stage, we aim to explore whether vegetation grassland heterogeneity, assessed through optical RGB images and based on structure from motion analysis, can serve as a proxy for flower diversity and, consequently, bee diversity and abundance.

• Regional scale: Habitat Classification using Sentinel-2 Data for Linking Habitat Diversity with Pollinator Diversity. The third objective involves utilizing satellite images, specifically from Sentinel-2 data, to classify different habitats (forest, agricultural areas, rivers, ponds, hedges). The goal is to establish connections between habitat diversity and pollinator diversity. This approach allows for a broader-scale analysis, providing valuable insights into the relationship between different landscapes and the presence of pollinators.

• Regional scale: we aim to evaluate the Spectral Variation Hypothesis for estimating butterfly biodiversity through Sentinel-2 satellite data. This concept hypothesizes that variability in reflectance known as spectral heterogeneity (SH) or "spectral variability" of an area is an expression of spatial ecosystem heterogeneity and therefore related to species diversity. The fourth goal involves identifying a potential link between spectral diversity—quantified by the Rao's Q heterogeneity index from Sentinel-2 imagery—and butterfly species richness, which is obtained from citizen science data. This method facilitates an analysis on a wider scale, offering significant insights into how environmental diversity influences butterfly distribution.

In summary, this task aims to adopt a multi-faceted and multi-scale approach, integrating data from diverse remote sensing platforms to advance our understanding of biodiversity indicators for pollinator diversity. The subsequent sections delve into the methodologies, results, and implications of each objective, contributing to the broader field of ecological research.

2 Methods

2.1 Study area

2.1.1 Bee biodiversity

For the analysis related to bee biodiversity, we focused on an area that covered approximately 70 km2, with elevations ranging from 70 to 171 meters above sea level, situated in the southeast of the Netherlands near the village of Gulpen (Fig. 1). Thirty grasslands, selected to encompass a spectrum of land use intensities—ranging from nutrient-poor, biodiversity-rich semi-natural grasslands to intensively fertilized areas—were chosen to assess the proposed approach. The deliberate selection of semi-natural, extensively utilized, and intensively managed grasslands from diverse regions aims to minimize the spatial clustering of distinct grassland types.

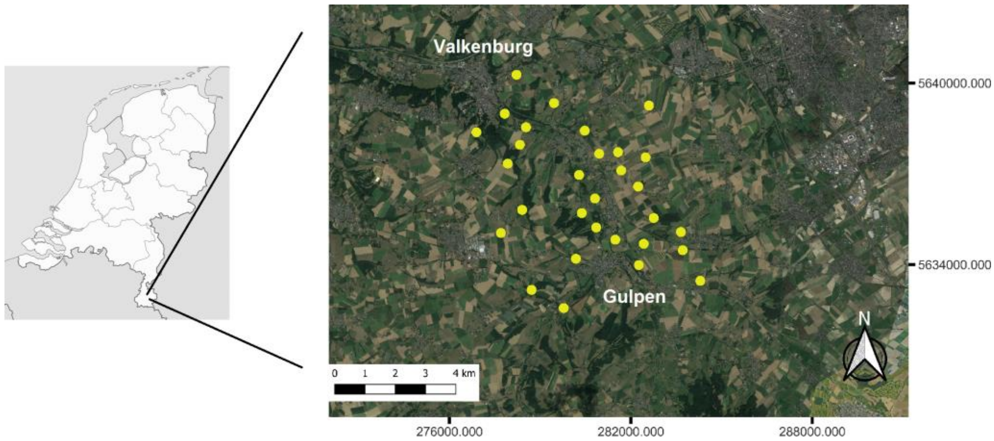

Figure 1: The study areas, situated in the Southeast of the Netherlands, are depicted in the map. Yellow dots mark the locations of the 30 transects within each study area (Basemap: Google Earth map as of August 2022).

2.1.2 Butterfly biodiversity

In the estimation of butterfly diversity, we focused on a bigger area, namely the whole Netherlands. The butterfly biodiversity data were collected within the European Butterfly Monitoring Scheme (eBMS), a citizen science network initiated by Butterfly Conservation Europe in April 2016. Its purpose is to consolidate data from various European country schemes into a unified database. The management of this database is overseen by the Natural Environment Research Council (operating through the Centre for Ecology and Hydrology) to support research efforts.

Figure 2: The location of all the plots where the field data were collected

2.2 Field data

2.2.1 Bee data

Each grassland was systematically surveyed using a transect measuring 150 by 1 meter, subdivided into three sections of 50 meters each (Figure 3). The transects were strategically positioned from the edge to the center of the grassland, traversing elevational differences when present, to capture the heterogeneity within the grassland. Distinctive plates, clearly visible in drone imagery, marked the transects. To minimize the likelihood of sampling the same bee populations, adjacent transects were predominantly separated by distances exceeding 500 meters (with a minimum separation of 435 meters). Studies have indicated that, although large-bodied bees like bumblebees can forage at distances of a few kilometers, their primary foraging range is within shorter distances (approximately 250-550 meters). Smaller wild bees tend to forage even closer [21]. Surveys of both bees and flowers were conducted along each transect. Transect walks, a standard method for studying plant-pollinator associations [29], were used to count wild bees and the honeybee (Apis mellifera). Two observers consistently conducted the surveys, counting all bees within a meter in front of them while walking along the transect for 15 minutes. This time excluded any handling time for caught specimens. For bee identification, specimens were either identified in the field using keys to the Dutch Apidae [4, 16, 17] or collected for later identification in the lab. Stereo-microscopes were employed for identification, and expert consultation was sought when necessary. The surveys were conducted between May 12th and May 31st, 2021, from 10 a.m. to 5 p.m. under favorable weather conditions (dry, >50% sunny, and at least 15 degrees Celsius, with wind speed <2 Beaufort). The field-collected bee data were utilized to derive plot-level metrics, including bee abundance (total observed bee count), bee species richness (count of unique species), and Shannon's H index [19] as an indicator of bee diversity. Flower surveys, following the method of Scheper et al. [24], were generally conducted on the same day as the bee surveys or, for logistical reasons, one or two days before or after. The number of open flowers of each species was recorded, and this information was used to calculate flower cover per transect. The calculation involved multiplying counts with species-specific estimates of flower size and summing over all observed flowering plant species [24].

Figure 3: Field data collection in the Dutch EBAs

2.2.2 Butterfly data

The field data collected, essential for the final comparison, are: length of the transect measured in meters (m), which includes only those transects of at least 300 m because the location of shorter transects might be inaccurate and short transects are more prone to local bias; total species richness (Nspp) (from raw data) and species richness of each monitoring scheme in a given climate region, proportion of species on transect relative to species richness of the region (Nspp). This standardized richness was used as an alternative richness variable. Transects had variable lengths, from 300 to 1000m.

2.3 Remote Sensing Data

2.3.1 Assessing Flower Cover as a Proxy for Bee Abundance using UAV Images

To achieve this goal at a local scale, we utilized UAV optical data. The UAV data collection, conducted concurrently with the field data collection, utilized the "DJI Matrice 210 RTK" UAV model equipped with the RGB Zenmuse X5 camera (16.0 MP, 17.3 x 13.0 mm sensor) featuring an integrated RTK GPS (Figure 4). Images were captured with an 80% overlapping rate to generate the final orthomosaic, with UAV flights executed at an altitude of approximately 20 meters above ground level. Employing the Agisoft Metashape Professional Edition software, a user-friendly workflow combining structure-from-motion and stereo-matching algorithms facilitated image alignment, dense point cloud assessment, digital elevation model (DEM) development, and orthomosaic construction. Key stages included feature extraction, sparse 3D point cloud creation, and automatic detection of Ground Control Points for precise georeferencing. The orthomosaic, exported as a GeoTIFF with a higher spatial resolution (around 0.5cm mean spatial resolution across 30 areas), played a crucial role in estimating flower cover. Additionally, variations in spatial resolution (1cm, 2cm, and 5cm) were explored to analyze their impact on flower cover estimation.

2.3.2 Testing Vegetation Grassland Heterogeneity as a Proxy for Flower Diversity and Pollinator Diversity using UAV Data

For this objective, conducted also at the local scale, the same UAV optical images of the previous objective were used. Conversely, the Agisoft Metashape Professional Edition software processed UAV images through three main stages: image alignment, dense point cloud creation, and digital elevation model inference. The procedures involved feature extraction, creating a sparse 3D point cloud, and automatically detecting Ground Control Point features for precise georeferencing. The dense point cloud, with a mean point density of 700 points/m2, was exported as a LAS file. Spatial resolutions for the Digital Surface Model (DSM) were derived at 10 cm, 25 cm, and 50 cm using the "dsmtin" algorithm. This algorithm utilized Delaunay triangulation to form a triangular irregular network (TIN), which was then rasterized to create the DSM. The Digital Terrain Model (DTM) and Canopy Height Model (CHM) were derived from the same point cloud, providing a comprehensive representation of terrain and vegetation structure.

Figure 4: UAV campaign in the Dutch EBA

2.3.3 Habitat Classification using Sentinel-2 Data for Linking Habitat Diversity with Pollinator Diversity

For the third objective, Sentinel-2 2A images (radiometrically calibrated and atmospherically corrected data, ensuring the accuracy of reflectance values), acquired through the Copernicus program were used. The selection of Sentinel-2 data was driven by its higher spatial resolution (compared to, for example, Landsat images), offering higher accuracy in our analysis. We focused on specific bands with a 10-meter spatial resolution, including the red, blue, green, and near-infrared (NIR) bands, to enhance spatial resolution for detailed habitat classification. Utilizing the calibrated and corrected Sentinel-2 2A images, a supervised classification approach was employed, leveraging machine learning algorithms (Random Forest) to discern and categorize different habitats based on their unique spectral signatures.

2.3.4 Testing the Spectral Variation Hypothesis for Butterfly biodiversity estimation

For this objective we tested the Spectral Variation Hypothesis for the estimation of butterfly diversity. In particular we assessed the spectral heterogeneity through the Rao's Q heterogeneity index, based on NDVI data derived from the Sentinel-2 images. Since our objective was to investigate the potential correlation between field data and landscape heterogeneity for the years 2018, 2019, and 2020 we used different Sentinel-2 images. We tested two different temporal approaches: one based on the use of a single image for the whole year (mean NDVI value from April to September) and one based on the use of four different images corresponding as the mean of the NDVI value for each season.

2.4 Analyses

2.4.1 Assessing Flower Cover as a Proxy for Bee Abundance using UAV Images

In order to derive the information of flower cover from the UAV RGB images, we applied three machine learning algorithms: Random Forest (RF), Support Vector Machine (SVM), and Neural Network (NNET)—using the caret R package [26].

Random Forest (RF) RF, based on decision trees, effectively classifies flower cover using bootstrapped training subsets and a voting mechanism [15, 9].

Support Vector Machine (SVM) SVM, a supervised learning model, classifies flower cover using a kernel function to separate classes in an N-dimensional space [15].

Neural Network (NNET) NNET, mimicking the human brain, captures relationships in data through connected processors [13]. Training and testing data, randomly selected within polygons, underwent a 10-fold cross-validation method. Performance metrics (Accuracy, Kappa, Precision, Sensitivity, Specificity, Negative Predictive Value, Positive Predictive Value, Balanced Accuracy, and F1) evaluated model validity and reliability. Successive correlation analyses explored the relationships between UAV image flower cover estimated with different machine learning methods at different spatial resolutions and *in-situ* data (flower cover, bee abundance, species richness, and diversity). This analysis informs the choice of algorithms and resolutions in capturing these relationships.

2.4.2 Testing Vegetation Grassland Heterogeneity as a Proxy for Flower Diversity and Pollinator Diversity using UAV Data

For the UAVs derived CHMs (at 10cm, 25cm and 50cm), the Vegetation Height Heterogeneity (HH) was calculated using the following heterogeneity indices.

Rao's Q Index: Originally devised by Rao and later recommended by Botta-Dukát for functional diversity in ecology, it was adapted as a heterogeneity index for remote sensing by Rocchini et al. The formula used is:

$$Q = \sum_{i,j=1}^{N} d_{ij} \times p_i \times p_j$$

where Q is Rao's Q index, p_I *and* p_j represents relative abundance, and d_{ij} is the distance between pixels i and j.

Coefficient of Variation (CV): Widely used in ecological studies, CV is calculated as $SD/\bar{x}$, where SD is the standard deviation, and x is the mean of pixel values.

Berger-Parker Index: A measure of dominance, calculated as n_{max}/N, where $n_{\max}$ is the abundance of the most dominant pixel and N is the total abun- dance.

Simpson's D Index: A diversity assessment measure applied to remote sensing data, calculated as $\sum_{i=1}^{n} p_i^2$ where D Is the Simpson index, n is the total number of pixels , and p_i is the relative abundance of a pixel value. The derived HH, calculated for each CHMs(at 10cm, 25cm, 50cm) using each heterogeneity indices was then correlated to vegetation species richness and bee species richness / abundance.

2.4.3 Habitat Classification using Sentinel-2 Data for Linking Habitat Diversity with Pollinator Diversity

For our third objective, we harnessed the potential of Sentinel-2 satellite images. Our aim was to classify diverse habitats within a 250-meter buffer around our central EBAs using a Random Forest classification approach. To train our classification model, we drew distinct polygons within the 250- meter buffer, representing four major classes: linear elements (e.g., linear forests, hedges), forests, urban areas, and agricultural fields. The model was trained using these polygons to recognize and classify the different land cover types. The classification process utilized a RF algorithm, a robust and versatile machine learning approach. This method leverages an ensemble of decision trees, each contributing to the final classification through a voting mechanism. Post-classification, we correlated the areas of each identified class with beediversity and abundance within each EBA. This analysis aims to unveil potential relationships between habitat types and the diversity and abundance of bee populations, providing valuable insights into the ecological dynamics of these areas.

2.4.4 Testing the Spectral Variation Hypothesis for Butterfly biodiversity estimation

Also in this case we made use of the Rao's Q index for the assessment of the Spectral Heterogeneity. To ensure accurate data analysis, the procedure of correlating SH and butterfly diversity was consistently applied across each year studied. The data for each year were divided according to transect length. The initial step involved analyzing all 300 m long transects with a corresponding 300 m buffer. This approach was systematically extended to longer transects in increments of 100 m, each time using a matching buffer size, up to transects of 1000 m length, which were analyzed using a 1000 m buffer.

3 Results

3.0.1 Assessing Flower Cover as a Proxy for Bee Abundance using UAV Images

RGB UAV images were leveraged to estimate flower cover, establishing positive correlations with field observations. Fig. 5 showcases the R^2 values derived from linear regressions between flower cover, bee abundance, species richness, and Shannon's H diversity, using different spatial resolutions (0.5cm, 1cm, 2cm, 5cm) and machine learning methods. The visual representation in Fig. 6 exemplifies RF machine learning's flower cover estimation at various spatial resolutions (0.5cm, 1cm, 2cm, 5cm) in a study area. Interestingly, the relationship's accuracy varied with algorithm and spatial resolution. SVM displayed a higher R^2 at 1 cm than at 0.5 cm resolution, contrary to other algorithms. RF and NNET algorithms consistently yielded high R^2 values, outperforming SVM. While SVM displayed lower performance, it still produced significant relations with field-estimated flower cover data. Accuracy ranged from a robust 0.8 for RF with a 0.5cm spatial resolution to 0.31 for SVM at 5cm resolution. The positive relationship extended to bee variables, with bee abundance, species richness, and diversity significantly and positively linked to flower cover estimated by RGB UAV images (Fig. 4b-c-d). Notably, RF and NNET algorithms consistently outperformed SVM, and the best models using RGB UAV estimates at 0.5cm spatial resolution exhibited higher goodness of fit than models relying on field observer data (Fig. 7). Specifically, the best relationships with flower cover estimated from UAV RGB had R^2 values of 0.65, 0.62, and 0.54 for bee abundance, species richness, and diversity, respectively. Further examination of the relationship between flower cover estimated by *in-situ* observations and the best machine learning UAV model (RF 0.5cm) revealed a positive and significant correlation with an R^2 value of 0.8 (Fig. 8). This trend extended tothe prediction of bee abundance, richness, and diversity, where RF models using RGB UAV estimates outperformed field observer data (Fig. 7). Assessing the performance of machine learning models across different spatial resolutions in estimating flower cover, Fig. 8 illustrates consistently high performance, except for the SVM model with RGB images at 5cm spatial resolution. In conclusion, the integration of RGB UAV images and machine learning algorithms proved effective in estimating flower cover and predicting its positive correlations with bee abundance, species richness, and diversity. Notably, RF and NNET algorithms demonstrated robust performance, surpassing traditional field observations in predicting ecological variables. The findings highlight the potential of remote sensing technologies for ecological studies and underscore the importance of spatial resolution in accurately capturing complex ecological dynamics.

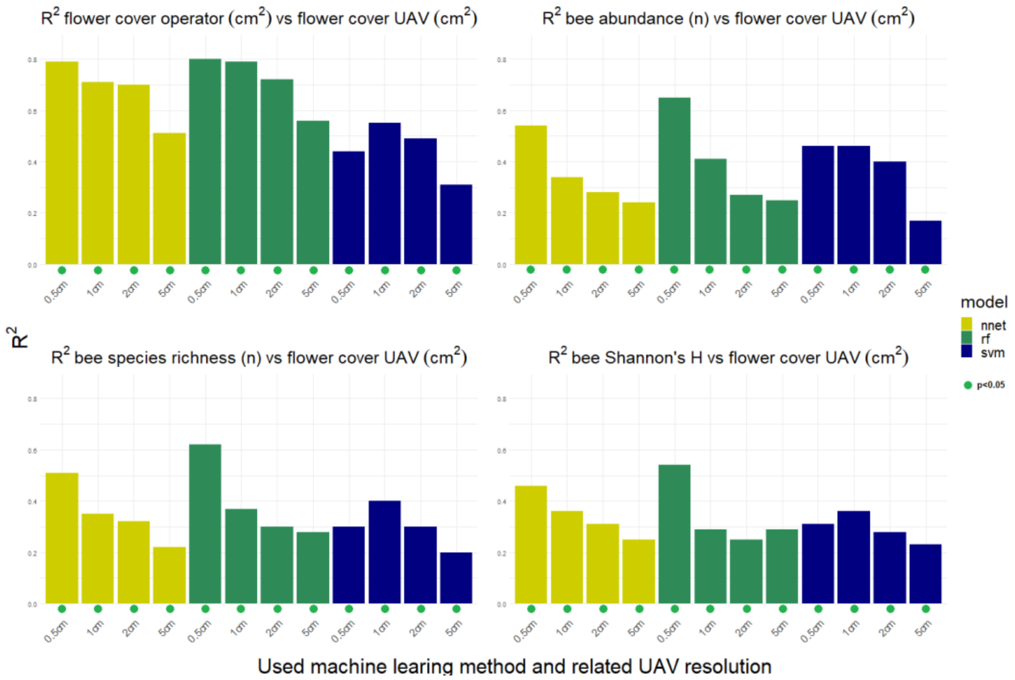

Figure 5: R^2 values derived from the linear regression between the field data (flower cover, bee abundance, bee species richness and bee Shannon's H diversity) and the flower cover estimated by RGB UAV images at different spatial resolution (0.5cm, 1cm, 2cm, 5cm) using different machine learning methods.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

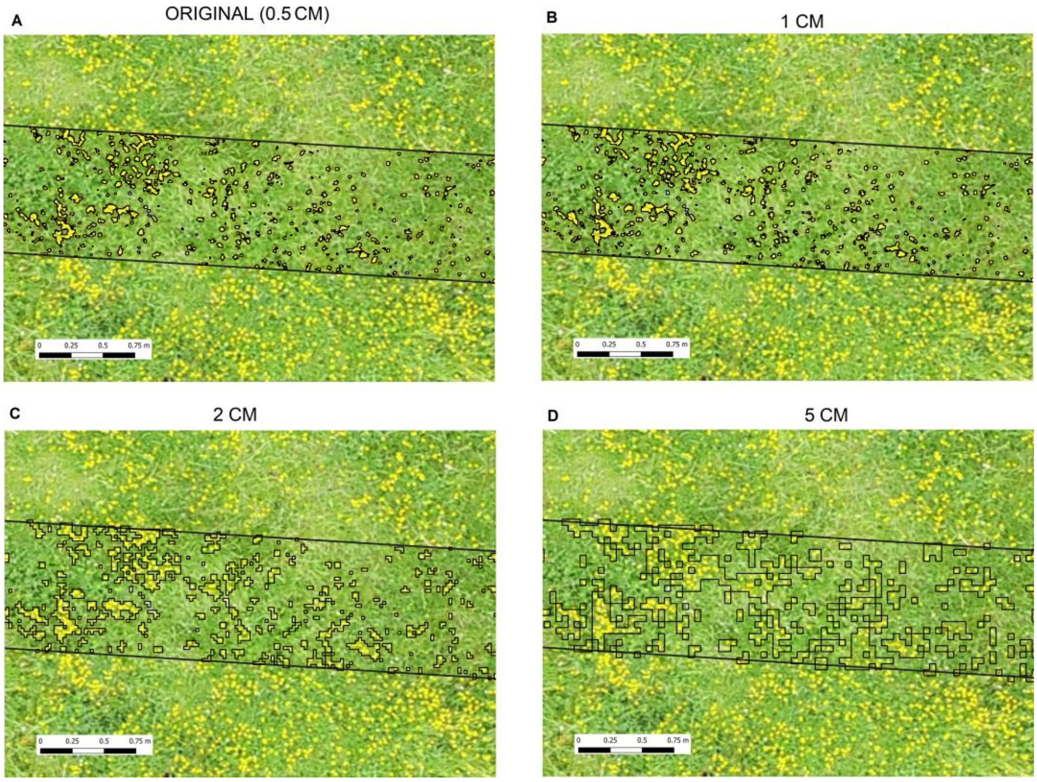

Figure 6: A visualisation of the results of the flower cover estimation by the RF machine learning methods at the different spatial resolution in one of the 30 study sites. The background image for the four sub-plots is at 0.5cm resolution.

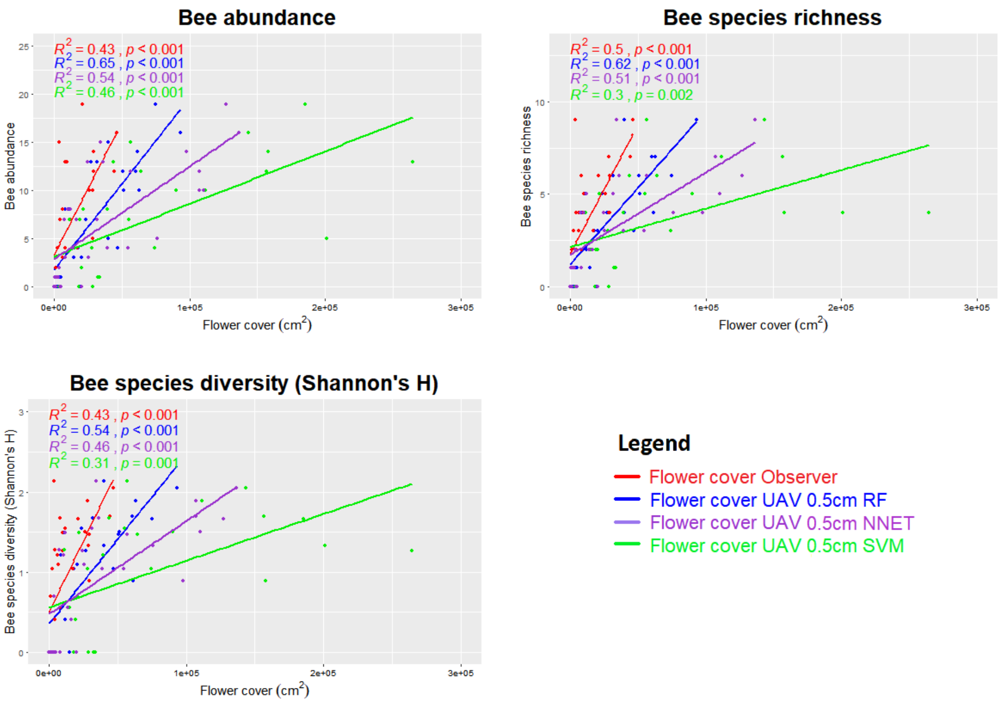

Figure 7: The comparison of the relationships between bee abundance, species richness and diversity (Shannon's H) and the flower cover estimated by the *in-situ* observations (Flower cover observer cm^2) and by the different machine learning models using RGB UAV images at the higher spatial resolution (0.5cm).

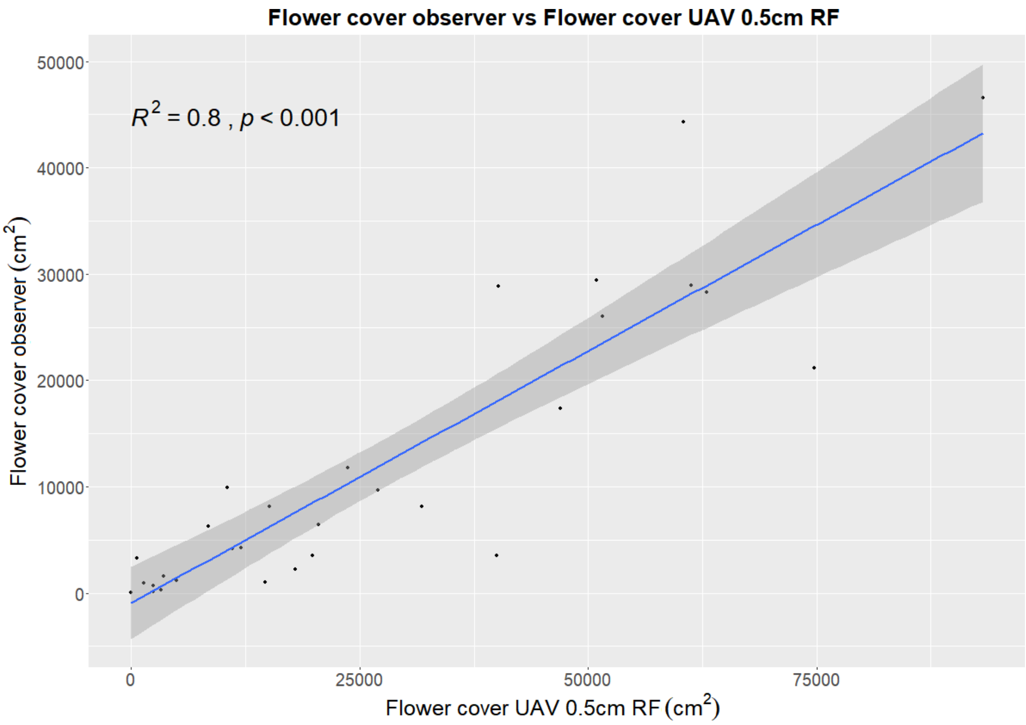

Figure 8: The relationship between the *in-situ* flower cover (Flower cover observer cm^2) and the flower cover estimated by the best machine learning UAV model (RF 0.5cm).

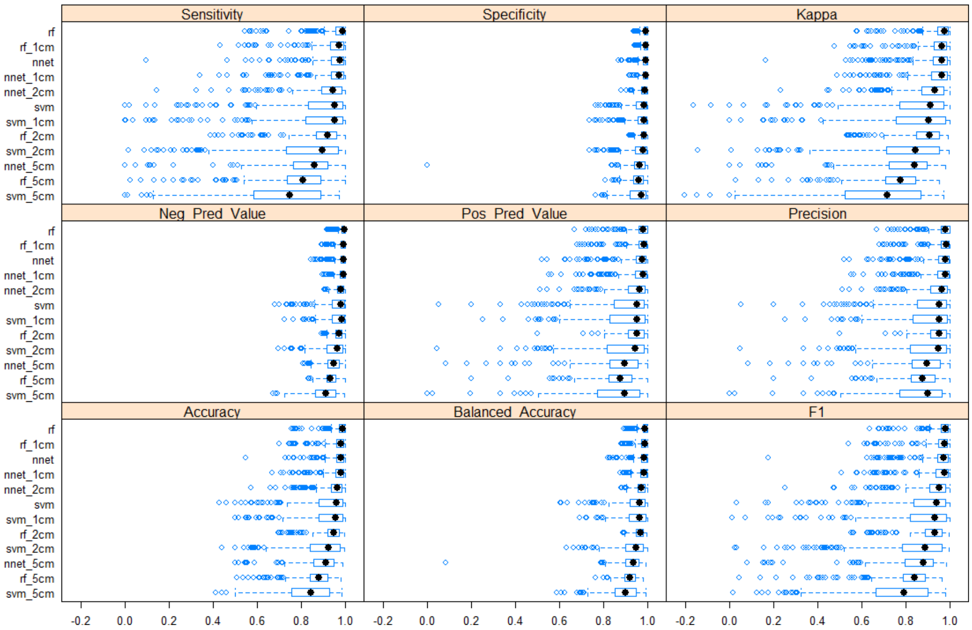

Figure 9: Performance of the machine learning models (Sensitivity, Specificity, Neg.Pred:VAlue, Pos Pred Value, Precision, Recall, Accuracy, Balanced Accuracy, F1, Kappa) at the different spatial resolution.

3.0.2 Testing Vegetation Grassland Heterogeneity as a Proxy for Flower Diversity and Pollinator Diversity using UAV Data

The correlation between flower diversity and HH calculated with various indices (Rao's Q, CV, Berger-Parker, and Simpson's D) from UAV photogrammetry at 10 cm, 25 cm, and 50 cm spatial resolutions is depicted in Figure10. Positive and significant correlations are observed, with the highest R^2 values associated with the Rao's Q index, ranging from 0.41 (10 cm resolution) to 0.44 (25 cm resolution). Figure 11 displays the correlation between bee abundance and HH calculated using the same heterogeneity indices and spatial resolutions. Positive and significant correlations exist, particularly with Rao's Q and Simpson's D indices, though R^2 values are generally lower compared to flower diversity correlations. The highest R^2 values occur when using the Rao's Q index, ranging from 0.31 (25 cm resolution) to 0.34 (50 cm resolution). Lastly, Figure 12 illustrates the correlation between bee diversity and HH calculated with the specified indices and resolutions. Positive correlations persist, with the Rao's Q index exhibiting the highest R^2 values. The Simpson's D index shows a comparatively modest correlation with HH.

These correlations are generally significant, except when calculated with the Berger-Parker index (at 10 cm and 50 cm CHM resolutions). In summary, the analysis reveals positive and significant correlations between flower diversity, bee abundance, and bee diversity with vegetation HH calculated using different indices and spatial resolutions, highlighting the importance of habitat structure in supporting biodiversity.

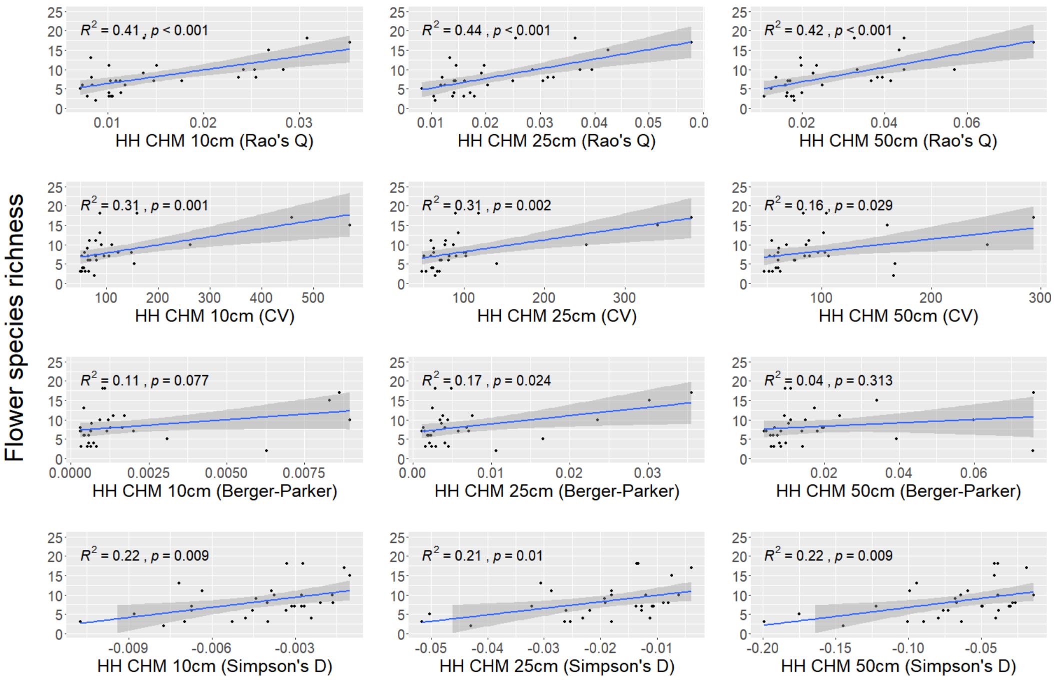

Figure 10: Correlation between the ground-based flower diversity and the HH calculated with the four heterogeneity indices (Rao's Q, CV, Berger-Parker and Simpson's D) derived from UAV CHM at 10 cm, 25 cm and 50 cm

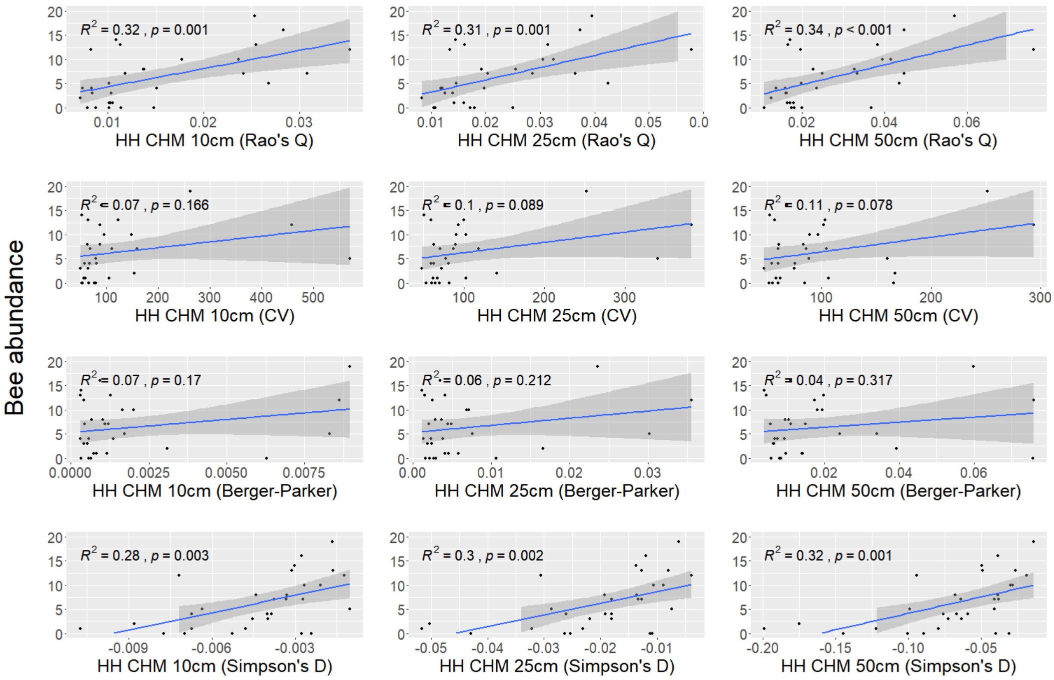

Figure 11: Correlation between ground-based bee abundance and HH calculated with the four heterogeneity indices (Rao's Q, CV, Berger-Parker, andSimpson's D) derived from UAV CHM at 10 cm, 25 cm, and 50 cm.

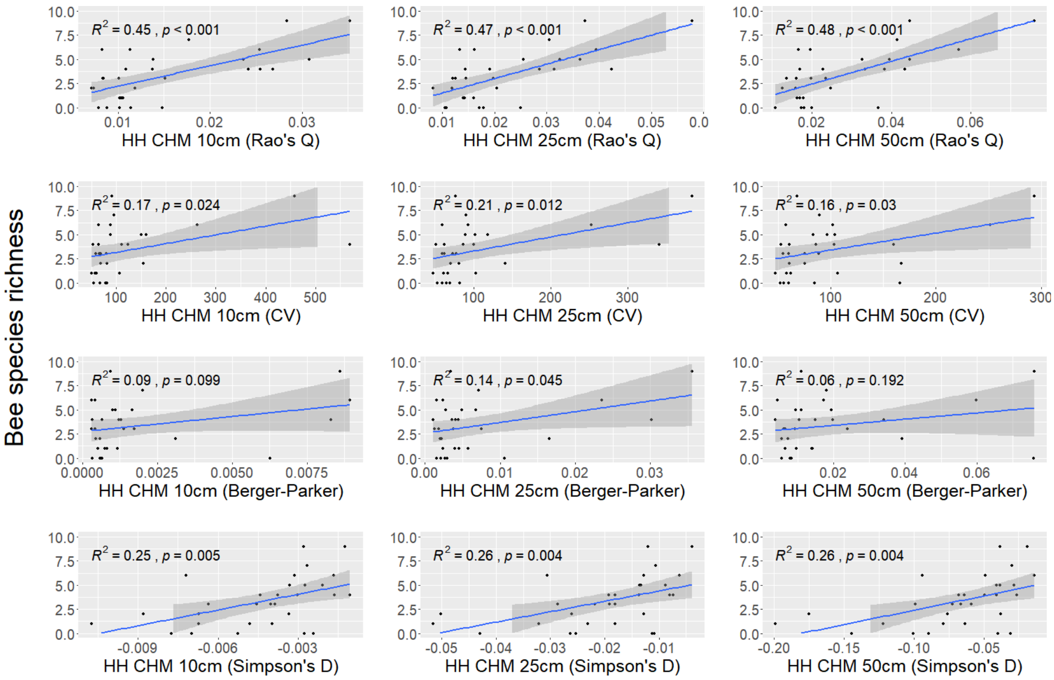

Figure 12: Correlation between ground-based bee diversity and HH calculated with the four heterogeneity indices (Rao's Q, CV, Berger-Parker, and Simpson's D) derived from UAV CHM at 10 cm, 25 cm, and 50 cm.

3.0.3 Habitat Classification using Sentinel-2 Data for Linking Habitat Diversity with Pollinator Diversity

In the ongoing exploration of various habitats, it is crucial to acknowledge the potential influence of confounding variables on the observed correlations. For instance, the analysis currently highlights a positive relationship between the extent of forested areas and both bee abundance and richness (Figure 14).

It is important to consider the possibility that areas rich in forests may also host the most flower-rich grasslands, contributing to the observed correlations. Therefore, future investigations will delve into landscape metrics and leverage remote sensing data and techniques (e.g. habitat classification from the Europe's land-cover map) to effectively assess and account for potential confounding variables, ensuring a comprehensive understanding of the factors shaping bee communities in different habitats.

4

Figure 13: The figure shows an example of the machine learning (Using the Random Forest Algorithms) based classification of the forest and agriculturalhabitats within one of the Dutch EBAs

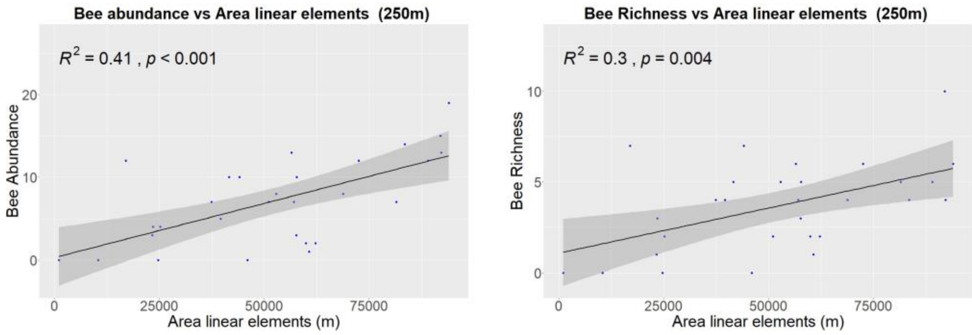

Figure 14: The figure shows the correlation between the area of forest areas (forest hedges, linear forests, small forests patches) derived from the machinelearning classification using Sentinel-2 data and the bee richness and abun- dance in the Dutch EBAs.

3.0.4 Testing the Spectral Variation Hypothesis for Butterfly biodiversity estimation

Figure 15-17 show the linear regression between the spectral heterogeneity calculated through the Rao's Q index on NDVI values derived from Sentinel-2 data for the three considered years.

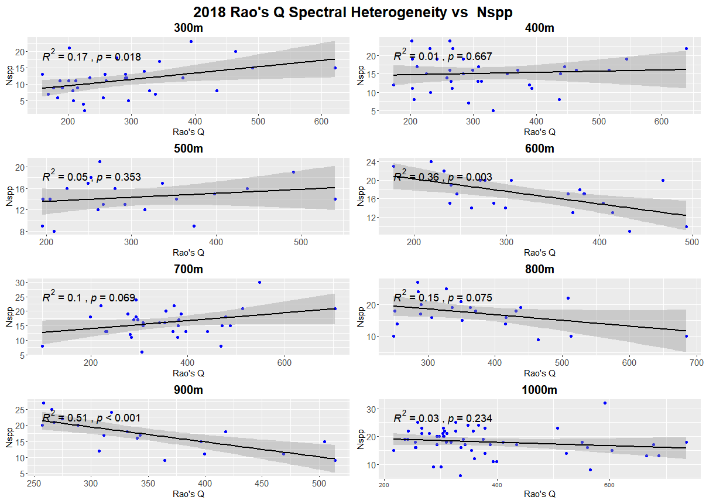

Figure 15: The figure shows the linear regressions between the spectral heterogeneity (Rao' s Q index) and the NSPP for the year 2018 for different buffers.

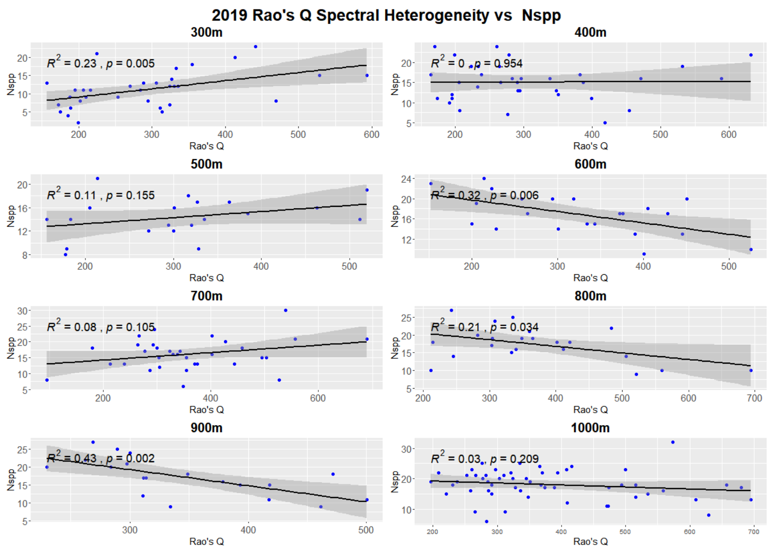

Figure 16: The figure shows the linear regressions between the spectral heterogeneity (Rao' s Q index) and the NSPP for the year 2019 for different buffers.

Across the three figures, each depicting scatter plots for the years 2018, 2019, and 2020 respectively, we observe the relationship between Rao's Q spectral heterogeneity and species richness (NSPP). Despite annual variations, a consistent trend across the years is the presence of some degree of positive correlation between spectral heterogeneity and NSPP, indicative of higher biodiversity in areas with greater spectral variation. Notably, the R^2 values fluctuate between years for corresponding transect lengths, suggesting a temporally dynamic effect. P-values associated with each plot reveal the statistical significance of these correlations, with many falling below the conventional significance threshold of 0.05, reinforcing the reliability of the observed trends. However, the significance does not uniformly translate across all transect lengths or years, hinting at complex ecological interactions.

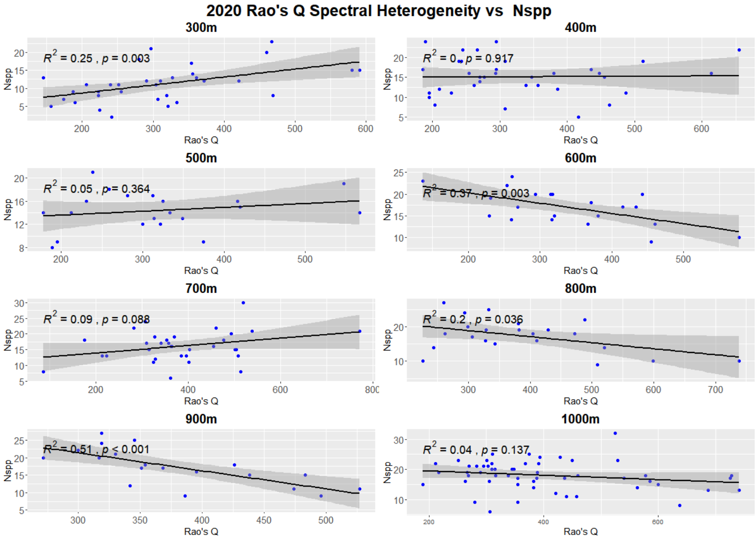

Figure 17: The figure shows the linear regressions between the spectral heterogeneity (Rao' s Q index) and the NSPP for the year 2020 for different buffers.

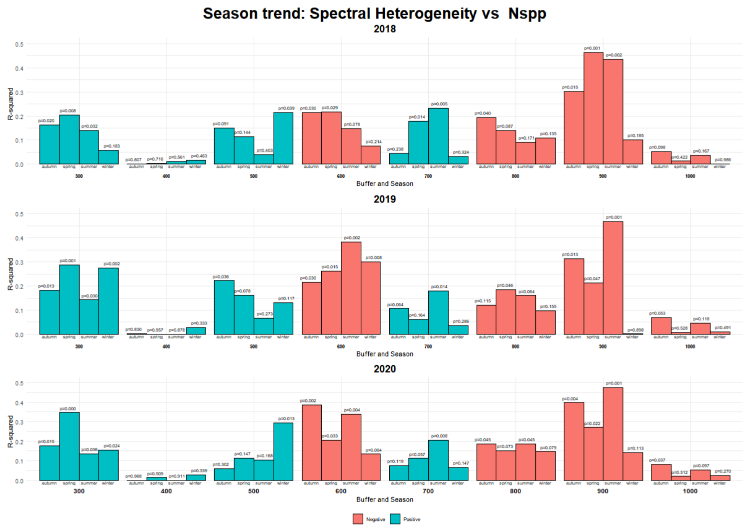

Figure 18: The figure shows the R² derived from the linear regressions between the spectral heterogeneity (Rao' s Q index) and the NSPP for the different years (2018, 2019 and 2020) for different buffers and for different seasons.

Figure 18 presents an insightful seasonal analysis of the correlation between spectral heterogeneity and NSPP over the course of three consecutive years. Each bar chart delineates the R^2 values for four distinct seasons, illustrating the ebb and flow of this relationship as influenced by the cyclical nature of ecological and phenological changes. It is evident that the strength of the correlation fluctuates with the seasons, which may be attributed to a myriad of factors, including vegetative growth patterns, seasonal climate variations, and other temporal ecological dynamics. The visualization of the R^2 value trends underscores significant seasonal impacts on spectral heterogeneity's predictive power regarding NSPP, a proxy for species richness.

4 Discussion

The utilization of remote sensing technologies, spanning from UAVs at the local scale to Sentinel-2 satellite data at the regional level, has emerged as a potent tool for advancing ecological studies, specifically in the domain of pollinator diversity within Dutch EBAs. Our multi-faceted approach encompasses various objectives, each shedding light on different facets of the intricate relationship between habitat characteristics and pollinator communities. At the local scale, the application of UAV optical data has proven particularly effective. In assessing flower cover as a proxy for bee abundance, our choice of RGB images coupled with machine learning algorithms such as Random Forest (RF), Support Vector Machine (SVM), and Neural Network (NNET), exhibited robust performance. The positive correlations observed between estimated flower cover and *in-situ* bee abundance, species richness, and diversity underscore the efficacy of this approach. Notably, RFand NNET algorithms outshone SVM, emphasizing the importance of algorithm selection. Furthermore, the exploration of different spatial resolutions (0.5cm, 1cm, 2cm, 5cm) revealed nuanced relationships, emphasizing thesignificance of spatial resolution in capturing ecological dynamics accurately. Similarly, in the second local-scale objective, the assessment of vegetation grassland heterogeneity as a proxy for flower and bee diversity, the useof UAV-derived data proved fruitful. The calculation of vegetation HH usingCHMs (derived from Structure from Motion UAV optical data) at varyingspatial resolutions (10 cm, 25 cm, 50 cm) demonstrated positive correlationswith both flower and bee diversity. Rao's Q index consistently yielded thehighest R^2 values, highlighting its efficacy in capturing the nuanced relationships within the ecosystem. This outcome reinforces the pivotal role ofhabitat structure, as represented by vegetation heterogeneity, in supporting and influencing pollinator diversity.

Moving to the regional scale, the application of Sentinel-2 satellite data for habitat classification showcased promising results. Specifically, the classification of forested areas using the Random Forest algorithm exhibited a positive correlation with bee abundance and richness. The R^2 values of 0.41 for bee abundance and 0.3 for richness underscore the significance of forested habitats in shaping bee communities. This machine learning-based classification approach provides a valuable tool for understanding the impact of different landscape elements on pollinator diversity within the broader context of EBAs.

According to our findings, the application of the Spectral Variation Hypothesis for estimating butterfly diversity at a regional scale seems ineffective. Despite thorough testing across various scales, buffer areas, and years, the hypothesis did not hold true in our study areas (within the whole Netherlands). One plausible explanation for this discrepancy could be the high level of human alteration in the landscape. The pronounced anthropogenic impact in the region may significantly influence spectral heterogeneity, thereby overshadowing the subtle spectral variations typically associated with butterfly biodiversity. Consequently, this human-dominated landscape presents a challenge for the SVH, suggesting that the method's applicability may be limited in heavily modified environments where human influence is the predominant factor affecting habitat heterogeneity.

In summary, our task integrates local and regional remote sensing data to unravel the intricate connections between habitat characteristics and pollinator diversity. The success of UAV optical data in estimating flower cover and assessing vegetation heterogeneity, coupled with the promising results from Sentinel-2 data in habitat classification, underscores the potential of remote sensing technologies in ecological research. On the other hand, the findings regarding the spectral variation hypothesis for butterfly diversity estimation yielded no positive outcomes. However, further analysis is necessary to refine our methodologies, consider additional environmental variables, and evaluate alternative hypotheses that may better capture the dynamics of butterfly populations. These efforts are essential for the advancement of remote sensing applications in biodiversity monitoring and contribute valuable insights that can inform conservation strategies and deepen our understanding of the complex interplay between habitats and pollinators in the project EBAs.

5 Supplementary Information

All the codes used to perform our analysis are stored in our public repository https://github.com/Ludovico-Chieffallo/Deliverable_1.4_Showcase/tree/main

References

1. Alvarez-Taboada, F., Paredes, C., & 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. Cruzan, M. B., Weinstein, B. G., Grasty, M. R., Kohrn, B. F., Hendrickson, E. C., Arredondo, T. M., & Thompson, P. G. (2016). Small unmanned aerial vehicles (micro-UAVs, drones) in plant ecology. Applications in Plant Sciences, 4(9), 1600041.
4. Falk, S., & Lewington, R. (2017). Veldgids bijen voor Nederland en Vlaanderen.
5. Gallai, N., Salles, J.-M., Settele, J., & Vaissière, B. E. (2009). Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. Ecological Economics, 68(3), 810-821.
6. Gallmann, J., Schüpbach, B., Jacot, K., Albrecht, M., Winizki, J., Kirchgessner, N., & Aasen, H. (2021). Flower mapping in grasslands with drones and deep learning. Frontiers in Plant Science, 12.
7. 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.
8. Hu, P., Chapman, S. C., & 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.
9. Immitzer, M., Atzberger, C., & 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.
10. 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.
11. 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.

12. Klein, A.-M., Vaissière, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., & Tscharntke, T. (2007). Importance of pollinators in changing landscapes for world crops. Proceedings of the Royal Society B: Biological Sciences, 274(1608), 303-313.
13. Kreig, J. A., Parish, E., & Jager, H. I. (2021). Growing grasses in unprofitable areas of US Midwest croplands could increase species richness. Biological Conservation, 261, 109289.
14. Kremen, C., Williams, N. M., & 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.
15. Nasiri, V., Darvishsefat, A. A., Arefi, H., Griess, V. C., Sadeghi, S. M. M., & 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.
16. Nieuwenhuijsen, H., & Peeters, T. (2015). Nederlandse bijen op naam brengen. Deel 1. - Stichting Jeugdbondsuitgeverij, 's Graveland.
17. Nieuwenhuijsen, H., Peeters, T., & Dijkshoorn, D. (2020). Nederlandse bijen op naam brengen. Deel 2. - Stichting Jeugdbondsuitgeverij, 's Graveland.
18. Ollerton, J., Winfree, R., & Tarrant, S. (2011). How many flowering plants are pollinated by animals? Oikos, 120(3), 321–326.
19. Pielou, E. C. (1966). Shannon's formula as a measure of specific diversity: its use and misuse. The American Naturalist, 100(914), 463–465.
20. 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.
21. Redhead, J. W., Dreier, S., Bourke, A. F., Heard, M. S., Jordan, W. C., Sumner, S., Wang, J., & 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.
22. 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.
23. Saunders, D. A., Hobbs, R. J., & Margules, C. R. (1991). Biological consequences of ecosystem fragmentation: a review. Conservation Biology, 5(1), 18–32.
24. 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.

25. Scheper, J., Reemer, M., van Kats, R., Ozinga, W. A., van der Linden, G. T., Schaminée, J. H., Siepel, H., & 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.

26. Team, R. C. (2013, 2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

27. Torresani, M., Rocchini, D., Sonnenschein, R., Zebisch, M., Hauffe, H. C., Heym, M., Pretzsch, H., & Tonon, G. (2020). Height variation hypothesis: A new approach for estimating forest species diversity with CHM LiDAR data. Ecological Indicators, 117, 106520.

28. Van Swaay, C. A., Nowicki, P., Settele, J., & Van Strien, A. J. (2008). Butterfly monitoring in Europe: methods, applications and perspectives. Biodiversity and Conservation, 17, 3455–3469.

29. 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.

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