POLITECNICO DI TORINO
Repository ISTITUZIONALE

The Use of Open Source Software for Monitoring Bee Diversity in Natural Systems: the Beems Project

Original
The Use of Open Source Software for Monitoring Bee Diversity in Natural Systems: the Beems Project / Dabove, P.; Di Pietra, V.. - In: INTERNATIONAL ARCHIVES OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES. - ISSN 2194-9034. - STAMPA. - XLVIII-4/W1-2022:(2022), pp. 111-118. (Intervento presentato al convegno Free and Open Source Software for Geospatial (FOSS4G) 2022 tenutosi a Firenze nel 22–28 August 2022) [10.5194/isprs-archives-xlviii-4-w1-2022-111-2022].

Availability:
This version is available at: 11583/2976920 since: 2023-03-14T10:31:26Z

Publisher:
Copernicus GmbH

Published
DOI:10.5194/isprs-archives-xlviii-4-w1-2022-111-2022

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
Common Ground Research Network postprint versione editoriale/Version of Record, con licenza CC by

(Article begins on next page)

24 April 2024

THE USE OF OPEN SOURCE SOFTWARE FOR MONITORING BEE DIVERSITY IN NATURAL SYSTEMS: THE BEEMS PROJECT

Paolo Dabove[1,*], Vincenzo Di Pietra[1]
[1]Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy – (paolo.dabove, vincenzo.dipietra)@polito.it

Commission IV, WG IV/4

KEY WORDS: Photogrammetry, supervised classification, segmentation, bees monitoring, natural systems.

ABSTRACT:

This work wants to highlight the results obtained during the BEEMS (Monitoring Bee Diversity in Natural System) project, which the main goal was to answer the following question: Which biotic and abiotic indicators of floral and nesting resources best reflect the diversity of bee species and community composition in the Israeli natural environment? The research was oriented towards the cost-effectiveness analysis of new aerial geomatics techniques and classical ground-based methods for collecting the indicators described above, based only on open-source software for data analysis. Two complementary study systems in central Israel have been considered: the Alexander Stream National Park, an area undergoing an ecological restoration project in a sandy ecosystem, and the Judean foothills area, to the South of Tel Aviv. In each study system, different surveys of bees, flowers, nesting substrates and soil, using classical field measurement methods have been conducted. Simultaneously, an integrated aero photogrammetric survey, acquiring different spectral responses of the land surface by means of Uncrewed Aerial Vehicle (UAV) imaging systems have been performed. The multispectral sensors have provided surface spectral response out of the visible spectrum, while the photogrammetric reconstruction has provided three-dimensional information. Thanks to Artificial Intelligence algorithms and the richness of the data acquired, a methodology for Land Cover Classification has been developed. The results obtained by ground surveys and advanced geomatics tools have been compared and overlapped. The results are promising and show a good fit between the two approaches, and high performance of the geomatics tools in providing valuable ecological data.

1. INTRODUCTION

One of the major factors in maintaining global human food supply is the pollination (Klein et al., 2007) in addition to the functional integrity of most terrestrial ecosystems (Ollerton et al., 2011). Wild pollinators have been defined as highly effective, often critical contributors to pollination services in natural and agricultural systems (Garibaldi et al., 2013), while colonies of the domesticated honey bee (Apis mellifera) offer their service mainly to crop pollination. Considering all wild pollinators, it is possible to define native bees as the most important pollinator group (Delaplane and Mayer 2000), because many bee communities usually provide more efficient and reliable pollination services than a single pollinator species (Blüthgen and Klein 2011) and are therefore of high conservation priority.

The monitoring of these species and the climate change analyses has been emphasized the importance of these species, highlighting the ongoing global declines in both wild and managed environments. This has raised the interest in native bee conservation and restoration and therefore in long-term monitoring programs. However, executing such monitoring programs is a major challenge, as they are labor-intensive, and require high expertise in the collection of bees and their subsequent taxonomic identification. To date, little has been done in developing efficient tools for monitoring bee communities.

The current prevailing approach for bee monitoring is site sampling oriented, time-consuming and labor-intensive upon collecting the required amount of data. Another approach is identifying, determining, and measuring which biotic and abiotic indicators of floral and nesting resources best reflect bee species diversity and community composition. Recent advances in the field of remote sensing have allowed performing such measurements parallelly to classical ground methods. The use of Radar technology (Milanesio et al., 2020), UAVs, and image analysis can provide a wide-scale, fast, data-rich digital-platform for developing cost-effective bee monitoring programs. These tools have been widely used for many research activities, starting from subfluvial springs' investigations (Aicardi et al., 2017), thermal analyses (Banding et al., 2012), forestry applications (Aicardi et al., 2016) and hyperspectral measurements (Weinmann et al., 2018). For this reason, the BEEMS project, a project of Scientific and Technological Cooperation between Italy and Israel (Scientific Track 2019), has focused the attention on the development of a technology-based approach for advanced bee community monitoring. The main idea is coupling photogrammetric tools, based on RGB images, with thermal and multispectral data, to develop a multi-scale and multi-temporal platform for monitoring bees and possibly other insect groups.

After this introduction, Sections 2 and 3 show the case study considered and the data acquisition, while section 4 describes the photogrammetric approach applied to this research activity. Section 5 is dedicated to showing the results and discussion, while section 6 concludes this article by proposing also some future steps for the next research activities.

2. METHODS

This section presents dataset acquisition, feature extraction, pre-processing and the structure of the prediction models, including the training procedure with respect to the case studies described in Section 3. Colour and spectral imagery can be used to infer physical and chemical characteristics of vegetation and soil, allowing to develop several models or forecasting, monitoring and management. Classifying the land use and the land cover through maps is one of the most important tasks for several actors.

* Corresponding author

The proposed method for land cover classification of Very High Resolution (VHR) imagery collected from Uncrewed Aerial Systems (UAS) exploits the benefits of Artificial Intelligence (AI) in extrapolating useful features from heterogeneous geospatial data and clustering geographic maps. In the state of the art, several algorithms of machine learning and image analysis are already implemented in open-source tools commonly applied in remote sensing and earth-observation. Very few works (e.g., Pontoglio et al., 2021) have tried to use, adapt and validate those algorithms on VHR images, in particular in photogrammetric products from UAS (orthomosaic and DSM). This section will provide information on the methods applied in the entire pipeline of map classifications, from the data acquisition, feature extraction, data preparation, pre-processing, training and set-up of the prediction model. Pixel-based and Object-based approaches are applied and compared addressing the benefits and drawbacks of each methodology. Particular attention is paid on the data preparation as well as on validation of the results.
The processing and the prediction model for land cover classification have been performed by exploiting the algorithms implemented in ORFEO ToolBox (OTB). OTB is an open-source project for processing satellite images, developed by CNSE (www.orfeo-toolbox.org, accessed 10 April 2019). The suite of libraries offered can be bound both in QGIS software (www.qgis.org, accessed 10 April 2019) and in a Python environment. Together with OTB, scikit-learn (https://scikit-learn.org/stable/) tools have been used in some steps of the analysis.

2.1 On-field and photogrammetric data acquisition

The Israeli research group, composed of ecologists and soil chemists, has carried out actions aimed at acquiring ecological indicators with classical field survey methods and analysing them statistically to model bee species diversity. In particular, the onfield exploratory survey of the region under study has allowed to define about 1000 samples to use as training and validation data. Those sample have been selected in each study case, assuring a good spatial distribution. As explained by Congalton and Green (Congalton and Green, 2019), to obtain a statistically consistent representation of the analyzed image, the accuracy assessment requires that an adequate number of samples should be collected per each land cover class. To this end, we considered it appropriate to select a total number of trainer polygons that did not deviate from 2% of the total number of segments of the entire image.
Regarding the photogrammetric data, multispectral images acquired from UAS flight serve as input information for deriving radiometric, spatial and three-dimensional information of the area through dense image matching and structure from motion. As for any aerial photogrammetric survey based on UAVs, a flight plan has been designed, assuring lateral and longitudinal overlaps within frames of 80 % in both directions. The flight elevation above the ground level, as well as the velocity of the quadricopter, was settled in order to assure a ground sample distance (GSD) of about 1.5 cm and to avoid motion blurring. The UAS was the same for each case study (DJI Matrice 200), equipped with a multispectral camera (Slantrange 4P+) which integrates multi-constellation and multi-frequency GNSS receivers that allow accurate positioning of the cameras' frames and the direct georeferencing of the photogrammetric block. The data processing has followed the classical photogrammetric pipeline, from feature extraction to image matching and aerial triangulation. The details regarding the georeferencing procedure and the residual errors on the GCP are reported in Section 3.
The only change between classical methods is related to the data preparation and, in particular, the reflectance calibration of the multispectral data. As is known, changes in sunlight conditions generate a strong variability in the spectral and colour response of the acquired surface unless proper calibrations are applied. Slantrange P4+ camera incorporates an on-board incident solar light measurement sensor, a range measurement and a geo-positioning sensor to provide necessary input data to the on-field calibration algorithm. For each reflectance data acquired in a specific discrete spectral frequency range with the photo sensor, the algorithm allows correlating the spectral measurements of the ambient at the given time of acquisition, as well as the incident angle of the light source. The algorithm, as well as the implementation in the commercial sensor, are presented in Figure 1. The Raw Slantrange images have been calibrated radiometrically, spectrally and geometrically by Slantview black box software and exported as multiple single-band .tif in 8-bit format. Some analysis has been made in order to evaluate the data missing behaviour of the calibration tool and assure the replicability of the analysis.

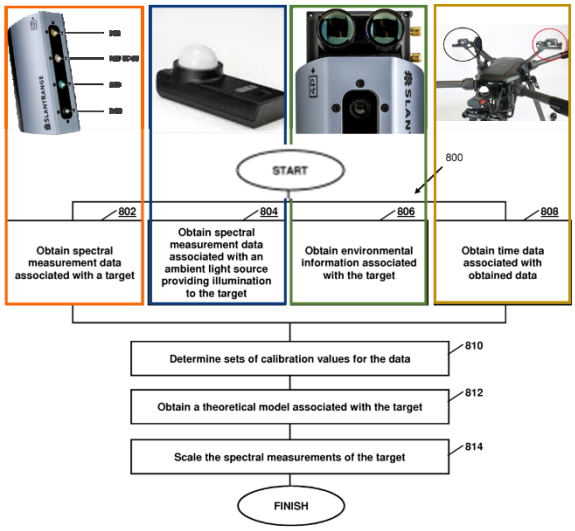

Figure 1. The workflow of the calibration procedure implemented by Slantrange and its relation to the sensor components.

2.2 Feature extraction and data preparation

The production of very high scale digital cartography allowed the extraction of the necessary data for training the proposed Artificial Intelligence model. These data were applied to two different approaches for automatic land cover classification. The first approach was based on unsupervised classification at the pixel level, while the second approach is based on object classification, i.e. vector polygons describing the boundaries of a real object. The algorithms operate differently on these two types of data. In fact, in the pixel-based approach, they are applied at the level of the single pixel, while in the object-oriented approach, they are applied to groups of homogenous pixels for a given feature. The implementation of all the training and validation phases of the proposed models was based on Python programming language using open libraries for data management (shapely, raster) and learning (sk-learn). The segmentation of the input data is fundamental in the approach in order to define the objects to be classified. Therefore, the OTB library was applied.
For object-oriented classification, we proceeded to apply automatic segmentation algorithms based on the analysis of multi-band spectral variability. In particular, the Large-Scale Mean-Shift (LSMS)-segmentation algorithm is used, producing a clustered image in which the pixels around a target pixel that

present similar behaviour from both the spatial and spectral points of view are grouped together. The LSMS-segmentation algorithm is a non-parametric and iterative clustering method introduced in 1975 by Fukunaga and Hostetler (Fukanaga and Hostetler 1975). It enables the performance of tile-wise segmentation of large VHR imagery, and the result is an artefact-free vector file in which each polygon corresponds to a segmented image containing the radiometric mean and variance of each band. It is based on four steps: LSMS-smoothing, LSMS-segmentation, LSMS-merging and LSMS-vectorization. The final step procedure vectorizes these clusters in order to obtain objects from which it is possible to compute the geometrical features (roundness, elongness, etc.). Finally, the operator associates a label to each vector for the dataset's generation.

The dataset preparation is a fundamental step in ML. Considering the aims of this work in using UAS photogrammetric products to estimate land cover, two different test sites with peculiar characteristics for the interested region were selected. In those areas, direct measurements were made (MSI, coordinates of ground points, indicators), and from them, meaningful synthetic features were extracted. The object-oriented approach was applied for the Alexander Stream National Park site while the pixel-based approach was applied on the Judean Foothills area.

Effective use of features as input data for a classification procedure can improve classification accuracy. Thus, the classification dataset was enriched with derivative features, namely, spectral-, textural-, and statistical-based features which can be descriptive of the object to classify. The definition of the ground cover classes has been assessed by the ecologist expert which have identified the most important indicator to monitoring bee diversity. The description of these classes is reported in Table 1. Knowing the classes and consequently identifying possible beneficial features, the raw MS images have been used to derive the aiding features. Table 2 reports the used measures for segmentation and classification. For each segment, the mean, the standard deviation, the median, the variance, the skewness and the kurtosis of the spectral, histogram and textural features were computed. The elevation is an important feature obtained by the photogrammetric procedure as well as the geometric-derived information. All those features have been stacked in an input data frame composed of thousands of samples/objects and hundreds of features. In image-based classification, independent features highly correlated with dependent features must be maintained, while independent features highly correlated between them must be eliminated. Therefore, once selected, they were subjected to importance and dimensionality reduction procedures in order to extract only the features with the highest weight for the given learning process. Firstly, the features were subjected to correlation analysis, reducing the features highly correlated in the function of a given threshold, then a feature importance analysis was addressed using the Gini criterion as described in Section 2.3.

Herbaceous plants		Plants that have non-woody steams, mostly annual plants	0
Dry herbaceous plants		Dead herbaceous plants	1
Woody plants		Perennial plants that have hard steams, like bushes and trees	2 (subclass 8 and 9)
Dry woody material		Dead woody material like broken brunches, dead steams	3
Bare soil		Exposed ground, free from rocks and plant litter	4
Leaves on ground		Leaf litter	5
Rocks		Rocks and stones	6
Moss		Moss	7
Trees		Subclass of Dry herbaceous plants mainly Eucalyptus/almond trees	8
Bushes		Subclass of Dry herbaceous plants	9

Table 1. Class labelling for LC classification based on ecologist expert directives.

Category	Features
Spectral-based	Color Index
Statistical-based	Variance
	Mean
	Skewness
	Kurtosis
Elevation	Digital Elevation Model
Geometric	Extension (Number of Pixels)
	Flat
	Roundness
	Longness
	Perimeter
Morphological	Edges sobel
	Harlick features
	Local statistics

Table 2. Features selected for the classification input dataset in object-based approach, computed using ORFEO toolbox.

The final dataset array has been prepared and each input feature was scaled using scikit-learn pre-processing MinMaxscaler (Pedregosa et al., 2011), which resizes each feature on the

minimum and maximum values. At this point, the dataset has been split into training and testing (70% training – 30% test). The results of the data preparation are a training and testing dataset arrays where the prediction variable constituted the Y dataset and the X is composed of the features, both sample-based and object-based.

2.3 Machine Learning algorithms and Validation

The ML algorithm that has provided the best results for the present task is the Random Forest classifier. Random forest is an automatic learning algorithm which consists of a set of decision trees that, thanks to a specific criterion, create several subsets of trees, each providing a classification result. The labels assigned to each input sample are assigned from the trees with the most votes. The RF classification was performed in a Python environment using Pandas, NumPy and Sklearn libraries. Random Forest classification model (Breiman, 2001) was trained using both sample and object datasets. The criterion used for the node splitting is the Gini (Hastie et al., 2009). The GINI gain is defined as the sum of impurity decreases from two nodes and the parent node. The GINI is calculated for each variable of the classifier. This parameter is a proxy of the importance of each feature within the model: the variables that have high GINI gain (so they have less impurity) are more important. Features that have GINI importance less than the median values of the feature importance values were excluded from the classification. The validation was performed in two steps. Firstly, the classical metrics to evaluate the model, then the comparison with the ground truth data acquired by the ecologist. The classification results were carried out using different validation metrics such as Precision, Recall, F1 score, etc.

3. CASE STUDY

The study involved two complementary study systems in central Israel (Figure 2), the Alexander Stream National Park, an area undergoing an ecological restoration project in a sandy ecosystem, and the Judean foothills area, to the South of Tel Aviv, which is characterized by an agro-natural landscape on vertisols.

Figure 2. The two areas investigated in this study in Israel.

In each study system, surveys of bees, flowers, nesting substrates and soil, using classical field measurement methods have been conducted. Simultaneously, an integrated aerophotogrammetric survey, acquiring different spectral responses of the land surface by means of UAV imaging systems has been performed. The acquisition of the indicators identified in the planning phase took place through several measurement campaigns conducted in the period between February 2020 and April 2020 located in two areas of interest in the Israeli territory. A total of 934 and 543 wild bees were collected in the two systems under study, respectively. From a geomatics point of view, 8 flights were carried out in the Alexander Stream National Park on 24 February 2020, acquiring approximately 65 GB of 8-bit multi-band images in tiff format. In the Judean foothills area, 11 flights were carried out on 26 February 2020, obtaining approximately 77 GB of tiff images. In addition, in order to obtain a correctly geo-referenced 3D model, a total of 54 Ground Control Points (GCPs) were acquired, of which 27 in Alexander Stream National Park and 27 in the Judean foothills, with a multi-frequency, multi-constellation GNSS geodetic receiver in RTK mode.

Figure 3. Instruments used during the survey's campaigns.

4. RESULTS AND DISCUSSION

4.1 Analysis of Slantrange on-field calibration

Once the collection is done, the imagery is uploaded to Slantrange calibration software, using the information from the ambient light irradiance sensor and the positioning, timing and attitude information of the camera drone and performing the reflectance calibration as well as the pixel crops to stack the 6 optical sensor images. As the algorithms are patented and not publicly available, the algorithm's performance has been analyzed in terms of the quantity of information extracted from the raw data. Table 3 and Table 4 show the image extraction results for both case studies.

time	Site name	Raw	Extracted	% extracted
T080621	Block 5	898	860	95.77
T082849	Block 5	273	130	47.62
T093532	Block 4	628	424	67.52
T095941	Block 4	372	248	66.67
T110657	Block 1-2	653	448	68.61
T113048	Block 1-2	478	293	61.30
T120149	Block 2-3	702	522	74.36
T122211	Block 2-3	454	309	68.06

Table 3. % of images extracted from the Slantrange calibration procedure for Alexander Stream National Park data acquisition performed on 24th February 2020.

time	Site name	Raw	Extracted	% extracted
T071450	Block 1	519	361	69.56
T080452	Block 1	502	278	55.38
T081754	Block 1	315	170	53.97
T084209	Block 1	390	138	35.38
T085536	Block 2	638	309	48.43
T091109	Block 2	597	283	47.40
T092900	Block 2	656	305	46.49

T094556	Block 2	618	248	40.13
T104605	Block 3	759	395	52.04
T110719	Block 3	592	248	41.89

Table 4. % of images extracted from the Slantrange calibration procedure for Judean Foothill data acquisition performed on 26th February 2020.

4.2 Photogrammetric results

The photogrammetric processing results are several digital products that contain radiometric, texture, spectral, and spatial information. These products are:

- Three-dimensional dense point clouds (DPC);
- Digital Surface Models (DSM);
- Orthomosaic map;

All these products are georeferenced and defined in WGS84-UTM zone 36N. The results are summarized in Table 5and Table 6 for Alexander Stream National Park and the Judean foothills.

Product	Param.	Dataset 1	Dataset 3	Dataset 4
Aerial images	n°	4446	3918	5022
DPC	pts n°	171888569	130377749	170992083
	run time	28 min	38 min	38 min
DSM	size	10,127x 13,513	11,317x 6,440	8,813x 12,936
	pixel	2,57 cm/pix	2,86 cm/pix	2,75 cm/pix
	run time	3 min	2 min	24min
Ortho-mosaic	band n°	6	6	6
	size	19,870x 27,002	22,623x 12,87	17,569x 25,85
	pixel	1,28 cm/pix	1,43 cm/px	1,38 cm/pix
	run time	27 min	15 min	43 min

Table 5. Summary of processing parameters and accuracy results of the Alexander Stream National Park site's digital products.

Digital Product	Param	Dataset 1	Dataset 2	Dataset 3
Aerial images	n°	1450	6846	4722
DPC	pts n°	98452320	194483854	82769132
	run time	37 min	3 hours	1 hours 40 min
DSM	size	9,595 x 10,824	13,979 x 13,852	7,042 x 8,467
	pixel	1,61 cm/pix	1,9 cm/px	2,01 xm/pix
	run time	2 min	27 min	2 min
Ortho-mosaic	band n°	6	6	6
	size	19,190 x 21,648	27,904 x 27,696	13,929 x 16,893
	pixel	0,08 cm/pix	0,09 cm/pix	1,01 cm/pix
	run time	22 min	3 hours	115 min

Table 6. Summary of processing parameters and accuracy results of the Judean foothills site's digital products.

4.3 Segmentation

Tuning the LSMS-segmentation algorithm is quite complex due to the original implementation, which was mainly focused on remote sensing images and the spatial variability of the objects defined in the class statement. Therefore, in order to define the most suitable parameters to segment the two study areas, several tests have been performed, and the results are compared. The LSMS-smoothing algorithm requires defining two parameters called spatial r and range r. The spatial r corresponds to the spatial radius. A higher value will lead to stronger smoothing and will also take more time. Range r is the spectral radius which means how pixels in the spatial radius and with close radiometry values will be averaged. A higher ranger will increase the smoothing effect. After that, the LSMS-segmentation application produces a labelled image where neighbour pixels whose range distance is below range radius ranger (and optionally spatial distance below spatial radius spatial r) will be grouped together into the same cluster. The parameters tested for tuning the algorithm are reported in Table 7. The final used parameters are highlighted in the table and represent the most satisfying parameters for segmenting our areas of interest. An example of orthomosaic segmentation is presented in Figure 4.

LSMS-Smoothing		
Test number	**Spatial r**	**Range r**
Test 1	10	50
Test 2	10	15
Test 3	5	15
LSMS-Segmentation		
Test number	**Spatial r**	**Range r**
Test 1	5	30
Test 2	5	15
Test 3	5	10
Test 4	5	2
Test 5	2	2
Test 6	2	1

Table 7. Spatial radius and range radius parameters tested to obtain the most effective segmentation result.

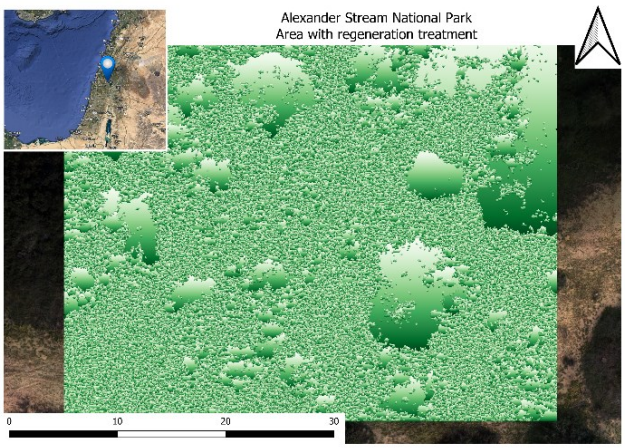

Figure 4. Example of LSMS-Segmentation algorithm result. The green objects are vector polygons of the Alexander Stream National Park subject to regeneration policies.

4.4 Feature extraction and dataset preparation

Identifying the features of greatest interest in describing the classes is necessary, which can lead to better classification accuracy. The features are the attributes of the pixels or objects derived from the digital images, from the multiple acquisition bands, their spectral response, their combination, and the textural, spatial, and shape information.

Using the implemented algorithms in OTB 41 features have been computed for the pixel-based sample and for object based-

sample. Then, QGIS statistical algorithm operating on vector polygon was applied in the Alexander stream objects, reaching a final number of 129 features. The dataset array has been composed as follows: 1011 rows and 41 columns for the pixel-based classification;

- 1053 rows and 129 columns for the object-based classification.

Staring from these datasets array, a Pearson correlation analysis has been performed, and the features with a correlation score of 0.9 have been dropped. The new dataset is composed of:

- 1011 rows and 21 columns for the pixel-based classification;
- 1053 rows and 47 columns for the object-based classification.

After this feature dropping, a feature importance analysis was addressed using the Gini criterion during the random forest training. Figure 5 shows the importance analysis for the 21 features used in pixel-based classification, while Figure 6 shows the importance of the 47 features used in object-based classification; the features that have an importance Gini score greater than 0.02 have been extracted, and the model has fitted on this new decreased dataset.

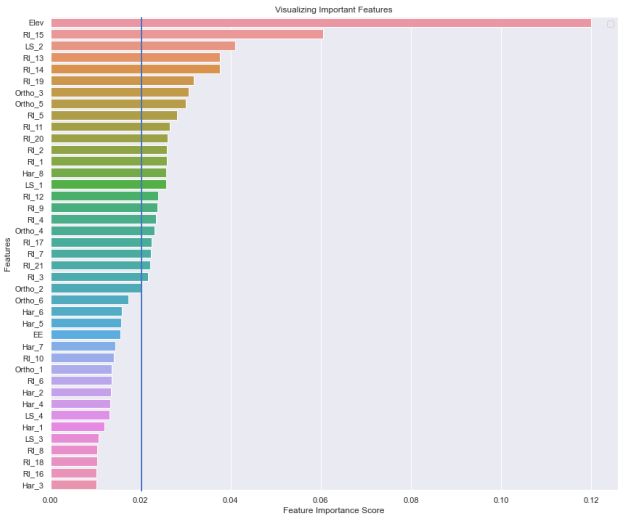

Figure 5. Importance analysis with Gini criterion for the pixel-based classification.

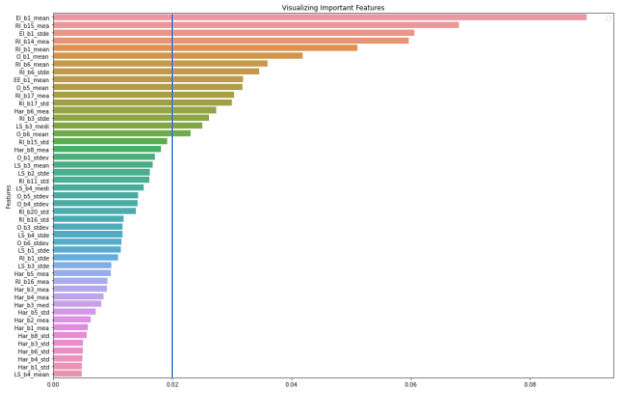

Figure 6. Importance analysis with Gini criterion for the object-based classification.

The input dataset in pixel-based classification was a 1011x21 matrix which was split into test (30% observations, 304) and training (70% observations, 707) datasets. The input dataset in pixel-based classification was a 1051x17 matrix which was split into test (30% observations, 316) and training (70% observations, 735) datasets.

4.5 Classification and accuracy results

The experiences described in previous works (Trisasongko et al., 2017) (Immtzer et al., 2016) show that the default values of the OTB parameters for training and classification processes seem to provide optimal results. For this reason, the default value parameters for the RF algorithm have been used, in particular, the maximum depth of the tree was set to 10, while the maximum number of trees in the forest was fixed to 1000. The class weight was placed equal to the parameter 'balanced_subsample' while the criterion parameter was settled equally to 'Gini'. The results of the validation for pixel-based and object-based approach are expressed in Table 8 and Table 9, which shows the accuracy metrics of the classification procedure, in particular the F1-score, Precision and Recall values of each class. Figure 7 and Figure 8 represent the confusion matrices. Generally, the quality of the classification is good. Indeed, the overall accuracy of the RF model is 0.92 for the sample-based approach and 0.79 for the object-based approach. The difference between the two approaches can be assumed to be related intrinsically to the object-based method, in which each vector polygon is expressed by its feature mean value and therefore injects some spectral response noise in the computation. Considering Table 8, the classification in pixel-based samples is well-performant in almost all classes, as demonstrated by the F1-score (0.78 as minimum value). Class 6, i.e. rocks and stones, has no predicted sample, which is due to the very low number of support data. The lowest accuracy is related to classes 4 and 6, bare soil and rocks, respectively. In this case, the number of false-positive increases probably due to the similarity in spectral response and the unbalanced number of objects supporting class 6.

class	precision	recall	f1-score	support
0	0.93	0.96	0.94	53
1	0.89	0.70	0.78	23
3	0.84	0.94	0.89	69
4	0.95	0.95	0.95	55
6	0.00	0.00	0.00	2
8	0.98	0.98	0.98	61
9	0.97	0.88	0.92	41
accuracy			0.92	304
macro avg	0.79	0.77	0.78	304
weighted avg	0.92	0.92	0.92	304

Table 8. Accuracy metrics calculated for pixel-based approach.

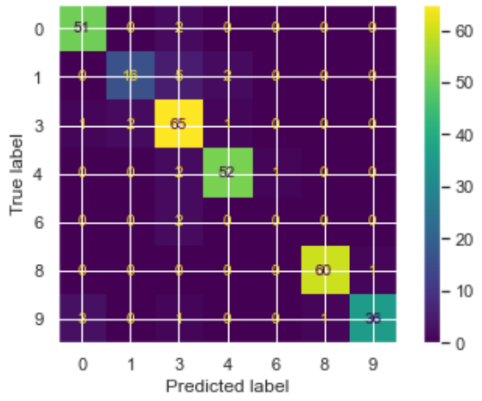

Figure 7. Normalized confusion matrix for pixel-based RF classification.

class	precision	recall	f1-score	support
0	0.94	0.98	0.96	52
1	0.86	0.80	0.83	15
3	0.93	0.94	0.94	70
4	0.57	0.48	0.52	66
6	0.20	0.26	0.23	31
8	0.98	1.00	0.99	49
9	1.00	0.94	0.97	33
accuracy			0.79	316
macro avg	0.78	0.77	0.78	316
weighted avg	0.80	0.79	0.79	316

Table 9. Accuracy metrics calculated for object-based approach.

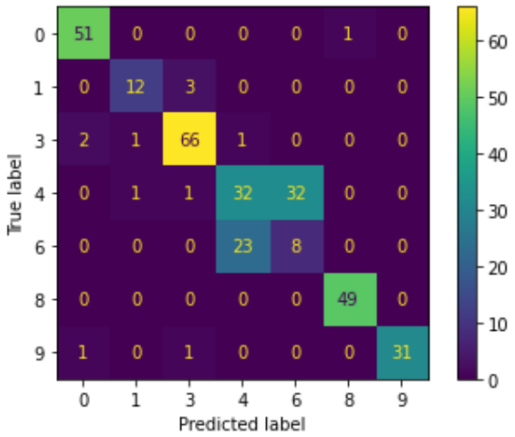

Figure 8. Normalized confusion matrix for object-based RF classification.

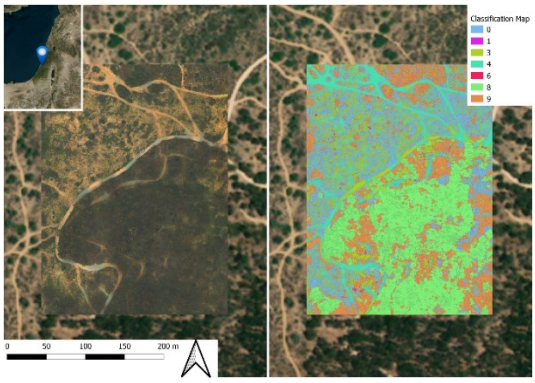

Figure 9. Extract of land cover map obtained with methodology developed by the geomatics group of the Politecnico di Torino.

code ***		0	1	2	3	4	5	6	8	9	Tot %
NA*	FS%**	22,20	4,20	17,80	3,40	53,67	15,67	0,0	n.a.	n.a.	116,9
	CL%	25,55	0,585	0,07	10,991	3,745	n.a.	0,0	0,32	58,802	100,0
	Δ%	-3,36	3,62	17,73	-7,59	49,92	n.a.	0,00	n.a.	n.a.	16,93
EU	FS%	32,73	1,20	0,07	5,20	0,33	63,13	0,0	n.a.	n.a.	102,7
	CL%	0,68	0,00	0,00	0,048	0,00	n.a.	0,0	89,44	9,83	100
	Δ%	32,05	1,20	0,07	5,15	0,33	n.a.	0,00	n.a.	n.a.	2,66
RG	FS%	30,27	0,53	0,00	3,00	49,53	16,67	0,0	n.a.	n.a.	100,0
	CL%	47,52	0,443	0,00	32,482	15,09	n.a.	0,047	0,00	4,39	100,0
	Δ%	-17,25	0,09	0,00	-29,48	34,45	n.a.	-0,05	n.a.	-n.a.	0,02
RS	FS%	26,07	1,40	0,00	1,27	65,27	6,00	0,0	n.a.	n.a.	100,0
	CL%	40,92	0,252	0,00	28,135	22,51	n.a.	n.a.	0,03	8,17	100,0
	Δ%%	-14,85	1,15	0,00	-26,87	42,76	n.a.	0,00	n.a.	n.a.	-0,02

Table 10. Land cover percentages of Block 1 identified in Alexander Stream National Park and discrepancies between the two methods implemented. *Natural (NA), Eucalyptus (EU), Regeneration (RG), Restoration (RS). **Field Sampling (FS), Classification results (CL), Discrepances (Δ%). ***Coding: Herbaceous plants (0), Woody plants (1), Dry woody materials (2), Bare soil (3), Leaves on the ground (4), Rocks (5), Threes (8), Bushes (9)

5. CONCLUSIONS

This work involved two complementary study systems in central Israel, the Alexander Stream National Park, an area undergoing an ecological restoration project in a sandy ecosystem, and the Judean foothills area, to the south of Tel Aviv, which is characterized by an agro-natural landscape on vertisols. A measurement campaign was conducted by the Israeli and Italian teams. We conducted surveys of bees, flowers, nesting substrates and soil in each study system, using classical field measurement methods. Simultaneously, we performed an integrated aerophotogrammetric survey, acquiring different spectral responses of the land surface through UAV imaging systems. The results of our analysis are summarised hereinafter. We found three focal habitat characteristics (indicators) shaping bee communities; the amount of bloom, perennial cover, and extent of bare ground (used by many bee species for nesting). VHR orthophotos and DSM have been produced with rigorous

photogrammetric methodology and with accurate georeferencing. The multispectral sensors have provided surface spectral response out of the visible spectrum, while the photogrammetric reconstruction has provided three-dimensional information. Thanks to Artificial Intelligence (AI) algorithms and the richness of the data acquired, a methodology for Land Cover Classification has been developed. The results obtained by ground surveys and advanced geomatics tools have been compared and overlapped. The results are promising and show a good fit between the two approaches, and the high performance of the geomatics tools in providing valuable ecological data.

ACKNOWLEDGEMENTS

The authors want to acknowledge the Israeli colleagues and partners of the BEEMS project for their help, competencies and passion, without which it had been impossible to reach the goal of this project.

REFERENCES

Aicardi, I., Chiabrando, F., Lingua, A., Noardo, F., Piras, M., Vigna, B., 2017. A methodology for acquisition and processing of thermal data acquired by UAVs: a test about subfluvial springs investigations. Geomatics, Natural Hazards and Risk, 8(1), 5–17.

Aicardi, I., Dabove, P., Lingua, A. M., Piras, M., 2016. Integration between TLS and UAV photogrammetry techniques for forestry applications. Iforest-Biogeosciences and Forestry, 10(1), 41.

Bartomeus, I., Ascher, J. S., Gibbs, J., Danforth, B. N., Wagner, D. L., Hedtke, S. M., Winfree, R., 2013. Historical changes in northeastern US bee pollinators related to shared ecological raits. Proceedings of the National Academy of Sciences, 110(12), 4656–4660.

Bendig, J., Bolten, A., Bareth, G., 2012. Introducing a low-cost mini-UAV for thermal-and multispectral-imaging. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci, 39, 345–349.

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.

Blüthgen, N., Klein, A.-M., 2011. Functional complementarity and specialisation: the role of biodiversity in plant–pollinator interactions. Basic and applied ecology, 12(4), 282–291.

Breiman, L., 2001. Random forests. Machine learning, 45(1), 5–32.

Congalton, R. G., Green, K., 2019. Assessing the accuracy of remotely sensed data: principles and practices. CRC press.

Delaplane, K. S., Mayer, D. F. et al., 2000. Crop pollination by bees. CABI publishing.

Fukunaga, K., Hostetler, L., 1975. The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Transactions on information theory, 21(1), 32–40.

Garibaldi, L. A., Steffan-Dewenter, I., Winfree, R., Aizen, M. A., Bommarco, R., Cunningham, S. A., Kremen, C., Carvalheiro, L. G., Harder, L. D., Afik, O. et al., 2013. Wild pollinators enhance fruit set of crops regardless of honey bee abundance. science, 339(6127), 1608–1611.

Geijzendorffer, I. R., Targetti, S., Schneider, M. K., Brus, D. J., Jeanneret, P., Jongman, R. H., Knotters, M., Viaggi, D., Angelova, S., Arndorfer, M. et al., 2016. How much would it cost to monitor farmland biodiversity in Europe? Journal of Applied Ecology, 53(1), 140–149.

Hastie, T., Tibshirani, R., Friedman, J., 2009. Random forests. The elements of statistical learning, Springer, 587–604.
Immitzer, M., Vuolo, F., Atzberger, C., 2016. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. Remote sensing, 8(3), 166.

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.

Mandelik, Y., Roll, U., Fleischer, A., 2010. Cost-efficiency of biodiversity indicators for Mediterranean ecosystems and the effects of socio-economic factors. Journal of applied ecology, 47(6), 1179–1188.

Milanesio, D., Bottigliero, S., Saccani, M., Maggiora, R., Viscardi, A., Gallesi, M. M., 2020. An harmonic radar prototype for insect tracking in harsh environments. 2020 IEEE International Radar Conference (RADAR), IEEE, 648–653.

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

Pisanty, G., Mandelik, Y., 2015. Profiling crop pollinators: life history traits predict habitat use and crop visitation by Mediterranean wild bees. Ecological Applications, 25(3), 742–752.

Pontoglio, E., Dabove, P., Grasso, N., Lingua, A. M., 2021. Automatic features detection in a fluvial environment through machine learning techniques based on UAVs multispectral data. Remote Sensing, 13(19), 3983.

Schmeller, D. S., Niemela, J., Bridgewater, P., 2017. The intergovernmental science-policy platform on biodiversity and ecosystem services (IPBES): getting involved. Biodiversity and Conservation, 26, 2271–2275.

Trisasongko, B. H., Panuju, D. R., Paull, D. J., Jia, X., Griffin, A. L., 2017. Comparing six pixel-wise classifiers for tropical rural land cover mapping using four forms of fully polarimetric SAR data. International Journal of Remote Sensing, 38(11), 3274–3293.

Van Rijn, I., Neeson, T. M., Mandelik, Y., 2015. Reliability and refinement of the higher taxa approach for bee richness and composition assessments. Ecological Applications, 25(1), 88–98.

Weinmann, M., Maier, P., Florath, J., Weidner, U., 2018. Investigations on the potential of hyperspectral and sentinel-2 data for land-cover / land-use classification. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences, 4(1).