Detecting and distinguishing between apicultural plants using UAV multispectral imaging

Alexandros Papachristoforou [Equal first author, 1, 2], **Maria Prodromou** [Equal first author, 3, 4], **Diofantos Hadjimitsis** [3, 4], **Michalakis Christoforou** [Corresp. 4, 5]

1 Department of Food Science and Technology, School of Agriculture, Aristotle University of Thessaloniki, Thessaloniki, Greece

2 Department of Food Science and Nutrition, School of the Environment, University of the Aegean, Myrina, Greece

3 Department of Civil Engineering and Geomatics, Cyprus University of Technology, Limassol, Cyprus

4 Department of Environment and Climate, Eratosthenes Center of Excelence, Limassol, Cyprus

5 Department of Agricultural Science, Biotechnology and Food Science, Cyprus University of Technology, Limassol, Cyprus

Corresponding Author: Michalakis Christoforou
Email address: m.christoforou@cut.ac.cy

Detecting and distinguishing apicultural plants are important elements of the evaluation and quantification of potential honey production worldwide. Today, remote sensing can provide accurate plant distribution maps using rapid and efficient techniques. In the present study, a five-band multispectral UAV (unmanned aerial vehicle) was used in an established beekeeping area on Lemnos Island, Greece, for the collection of high-resolution images from three areas where *Thymus capitatus* and *Sarcopoterium spinosum* are present. Orthophotos of UAV bands for each area were used in combination with vegetation indices in the Google Earth Engine (GEE) platform, to classify the area occupied by the two plant species. From the five classifiers (Random Forest – RF, Gradient Tree Boost – GTB, Classification and Regression Trees – CART , Mahalanobis Minimum Distance – MMD and Support Vector Machine – SVM) in GEE, the RF gave the highest overall accuracy with a Kappa coefficient reaching 93.6%, 98.3%, 94.7%, and coefficient of 0.90, 0.97, 0.92 respectively for each case study. The training method used in the present study detected and discriminated the two plants with great accuracy and results were confirmed using 70% of the total score to train the GEE and 30% to assess the method's accuracy. Based on this study, identification and mapping of *Thymus capitatus* areas is possible and could help in the promotion and protection of this valuable species which, on many Greek Islands, is the sole foraging plant of honeybees.

1   **Detecting and distinguishing between apicultural plants using UAV multispectral imaging**

3   Alexandros Papachristoforou [1,2†], Maria Prodromou [3,4†], Diofantos Hadjimitsis[3,4] and Michalakis
4   Christoforou [4,5]*

6   [1] Department of Food Science and Technology, School of Agriculture, Aristotle University of
7   Thessaloniki, 54124 Thessaloniki, Greece
8   [2] Department of Food Science and Nutrition, School of the Environment, University of the Aegean,
9   Ierou Lochou 10 & Makrygianni, 81400 Myrina, Greece
10   [3] Remote Sensing and Geo-Environment Lab, Department of Civil Engineering and Geomatics,
11   Faculty of Engineering and Technology, Cyprus University of Technology, Limassol 3036,
12   Cyprus
13   [4] ERATOSTHENES Centre of Excellence, Limassol 3036, Cyprus
14   [5] Department of Agricultural Science, Biotechnology and Food Science, Cyprus University of
15   Technology, Limassol 3036, Cyprus

17   [†]Both authors contributed equally. Names in alphabetic order.
18   *Corresponding Author:
19   Michalakis Christoforou[4,5]
20   Anexartisias 46, Limassol 3036, Cyprus
21   Email address: m.christoforou@cut.ac.cy

Abstract

24 Detecting and distinguishing apicultural plants are important elements of the evaluation and
25 quantification of potential honey production worldwide. Today, remote sensing can provide
26 accurate plant distribution maps using rapid and efficient techniques. In the present study, a five-
27 band multispectral UAV (unmanned aerial vehicle) was used in an established beekeeping area on
28 Lemnos Island, Greece, for the collection of high-resolution images from three areas where
29 *Thymus capitatus* and *Sarcopoterium spinosum* are present. Orthophotos of UAV bands for each
30 area were used in combination with vegetation indices in the Google Earth Engine (GEE) platform,
31 to classify the area occupied by the two plant species. From the five classifiers (Random Forest –
32 RF, Gradient Tree Boost – GTB, Classification and Regression Trees – CART, Mahalanobis
33 Minimum Distance – MMD and Support Vector Machine – SVM) in GEE, the RF gave the highest
34 overall accuracy with a Kappa coefficient reaching 93.6%, 98.3%, 94.7%, and coefficient of 0.90,
35 0.97, 0.92 respectively for each case study. The training method used in the present study detected
36 and discriminated the two plants with great accuracy and results were confirmed using 70% of the
37 total score to train the GEE and 30% to assess the method's accuracy. Based on this study,
38 identification and mapping of *Thymus capitatus* areas is possible and could help in the promotion
39 and protection of this valuable species which, on many Greek Islands, is the sole foraging plant of
40 honeybees.
41

Introduction

42
43 In recent decades, large-scale plant and species identification looked to aerial multispectral sensors
44 because of their ability to provide high-resolution spectral and spatial data (*Lewis & Brown, 2021;*
45 *Paz-Kagan et al., 2019; Robinson et al., 2016*). Due to human impacts and climate change,
46 cultivated plants and native vegetation are stressed by various specific factors such as the increase
47 in temperature, extended droughts, salinity, overgrazing and fires (*Van den Elsen et al., 2020*).
48 This leads to negative effects on food production, the environmental equilibrium and biodiversity.
49 Apiculture with its associated pollination benefits is one of the most important agricultural
50 activities that can affect vegetation, crops production and the environment equilibrium. The
51 importance of the role of bees is being increasingly acknowledged as bees pollinate a significant
52 percentage of the world's crops (*Breeze et al., 2011; Khalifa et al., 2021*).
53 Remote sensing systems have become powerful, supporting novel applications in agriculture, but
54 applications in apiculture are, as yet, limited and focused mainly on tracking honeybee predators
55 (*Reynaud & Guérin-Lassous, 2017*) or honeybee congregation areas (*Carl et al, 2017*). However,
56 remote sensing could become a useful tool for beekeepers by providing vegetation maps indicating
57 the presence and flowering periods of plants that are highly attractive to bees (*Marnasidis et al.,*
58 *2021*) as well as indicating the presence of conventional crops (*Donkersley et al., 2014*). It can
59 alert beekeepers to the possibilities of transferring colonies to better, safer and more productive
60 locations. As bees' nutrition requirements are very specific (*Louargant et al., 2017*), detecting and
61 distinguishing apicultural plants from non-apicultural plants and crops is of great importance to
62 beekeepers.

63 In recent years, classification models used in combination with multispectral and/or hyperspectral
64 aerial images have been deployed to discriminate soil and vegetation types and within vegetation
65 types to distinguish monocotyledonous/dicotyledonous plants (*Papp et al., 2021*) as well as weeds
66 from crops (*Barrero & Perdomo, 2018; Senthilnath et al., 2017*). Classification algorithms and
67 models can separate materials and plants using low-altitude, high-resolution spatial images
68 (*Rasmussen et al., 2013*). Low-altitude images can be collected using aerial platforms such as
69 unmanned aerial vehicles (UAVs) equipped with high-resolution sensors able to detect small
70 objects such as Mediterranean thorny shrubs (*Henkin et al., 2014*) and measure crop growth by
71 capturing a range of spectral information (*Duan et al., 2021*). The popularity of UAV technologies
72 has increased rapidly in remote sensing recently because of i) the ability to collect frequently very
73 precise high-resolution spatial images (*Lu & He, 2017; Rango et al., 2006*), ii) the flexibility in
74 flight planning and the rapid acquisition of the images (*Torresan et al., 2017*) and iii) the low cost
75 of the equipment (*Anderson & Gaston, 2013*).
76 Detecting apicultural plants and discriminating them from non-apicultural plants is essential for
77 beekeepers, especially in the case of thorny shrubs with similar phenological cycles and close
78 morphological characteristics such as Thymus and Sarcopoterium species. *Thymus capitatus* (or
79 *Corydothymus capitatus*) is a small, compact perennial shrub growing to a height of 40-60 cm. It
80 is recognizable for its shape and the long violet flowers held in cone-shaped clusters at the end of
81 the stems. Although in many areas it is considered an evergreen plant, in the Greek Aegean Islands,
82 the abscission of all leaves occurs in autumn and results in a "dry" appearance of the shrub. *T.*
83 *capitatus* is considered the most valuable apicultural plant of the Aegean Islands. In many islands
84 (especially the Cyclades Islands complex) it is the only forage for honeybees; maintenance
85 depends exclusively on the blossom and nectar flow of this species. Furthermore, the superb taste
86 and aroma of thyme honey is highly appreciated by consumers, making it the most expensive
87 Greek honey.
88 *Sarcopoterium spinosum* (L.), is a spiny, evergreen rosaceous dwarf shrub, 30–60 cm in height,
89 with branches ending in dichotomous and leafless thorns. A combination of the plant's clonal and
90 sexual reproduction contributes to its long-term survival and dominance (*Hill et al., 1998*).
91 Because of its reproductive capability, *S. spinosum* is considered a dominant species in the
92 Mediterranean basin, occupying up to 96% of the total vegetation area in some areas (*Arianoutsou*
93 *& Margaris, 1982*). The *S. spinosum* can resprout and grow faster than the apicultural plants after
94 destruction, e.g., by fire (*Niphadkar & Nagendra, 2016*). It may then colonize land traditionally
95 covered by *Thymus* spp. and thereby have a negative impact on apiculture, especially for
96 beekeepers who produce and sell value-added honey labeled as "thymus honey".
97 Systematic monitoring could be of great importance for apiculture management to enable
98 beekeepers to detect and estimate the areas covered by *Thymus* species, to distinguish thyme and
99 other apicultural plants from non-apicultural plants and to take measures to respond to changes. It
100 is also mandatory to keep track of newer invasions (*Hestir et al., 2008*) of *S. spinosum* within wild
101 *T. capitatus* ecosystems and other apicultural plant species, especially after a wildfire.

102   Various methods are described in the literature for the classification of ground "features" such as
103   plants (weeds, bryophytes and invasive plants) (*Ishida, et al., 2018; Papp et al., 2021; Rasmussen*
104   *et al., 2013*). The methods referred in the literature for detection and identification of apicultural
105   plants (or invasive species/weeds) vary. Some include the Random Forest- RF *(Dmitriev et al.,*
106   *2022;, Große-Stoltenberg et al., 2016, Hill et al., 2017, Kattenborn et al., 2019, Sheffield et al.,*
107   *2022, Singh & Singh, 2022*), Classification and Regression Trees -CART (*Fariz et al., 2022,*
108   *Juraiza Ishak et al., 2008, Raj & Sharma, 2022, Traganos et al., 2018*), Support Vector Machine
109   -SVM *(Forster et al., 2017, Paz-Kagan et al., 2019, Skowronek et al., 2017*), Mahalanobis
110   Minimum Distance - MMD *(Sampedro & Mena, 2018, Yang & Everitt, 2010*) and Gradient Tree
111   Boost (*Sujud et al., 2021*). For comparative purposes, the majority of the studies have used multiple
112   supervised classification models (*Arasumani et al., 2021, Gašparovičová et al, 2022, Zhu et al.,*
113   *2022*) such as Support Vector Machines (SVMs), Artificial Neural Network (ANN), Gradient Tree
114   Boost (GTB) and Random Forest (RF) classifiers. They have been widely applied for the detection
115   and identification of plants, in combination with the use of UAV high-resolution aerial images
116   with (*Barrero & Perdomo, 2018; Bolch et al., 2021; Pretorius & Pretorius, 2015*). The accuracy
117   of the data retrieved from individual pixels (the spatial units in an image) are analyzed by the
118   spectral information they contain and compared with ground-truth information (*Google Earth*
119   *Engine, 2022*).
120   Through remote sensing techniques, the present study aims to detect, identify and distinguish *T.*
121   *capitatus* from the non-apicultural, fast-growing and aggressive plant *S. spinosum*. Remote sensing
122   will be used for preliminary mapping and classification of apicultural areas. We conducted the
123   present study on Lemnos (Greece), an island of great apicultural interest where expansion of *S.*
124   *spinosa* is increasing rapidly and replaces the apicultural vegetation, especially after wildfires
125   which pose a serious threat for the beekeeping community of the island. A UAV equipped with a
126   high-resolution multispectral camera was used, and pixel-based supervised classification
127   algorithms (Gratiend Tree Boost, CART, Random Forest, Mahalanobis Minimum Distance and
128   Support Vector Machine) were implemented on the UAV orthophotos and processed in the GEE
129   platform (*Google Earth Engine, 2022*) along with ground-truth information. The GEE platform
130   was used for image processing since it is low-cost, accessible and user-friendly. The
131   comprehensive review by *Yang et al.* (2022) proves the incorporation of artificial intelligence (AI)
132   models into to GEE with the most common AI models used for vegetation and/or crop mapping
133   are the RF, CART and SVM. For the evaluation of the AI models the Producer's Accuracy (PA),
134   User's Accuracy (UA), Overall Accuracy (OA) and kappa coefficient are mostly used.
135   Through the GEE it is possible to implement further supervised classification methods that also
136   provide deep learning methods, such as fully convolutional neural networks (FCNN), or deep
137   neural networks (DNN), but these algorithms require external training on the TensorFlow platform
138   (*Abadi et al., 2016; Orieschnig et al., 2021*). Our study focuses entirely on the use of open-access
139   solutions and for this, the traditional machine-learning algorithms were selected because those of
140   Google Cloud are billable and require much more training data than machine learning (*Yang et al.,*
141   *2018*).

142

Materials & Methods

143

144

Study area

The research was conducted in three experimental sites on Lemnos Island (Figure 1) which lies in the northern part of the Aegean Sea in Greece (SA1: 25º 19'10''E, 39º 48'43''N, SA2: 25º 18'39''E, 39º 48'24''N, SA3: 25º 19'5''E, 39º 48'50''N). The historical fire events of the same plots were identified through Fire Information for Resource Management System (FIRMS, 2021) and downloaded as active fire and thermal anomaly locations using MODIS and Visible Infrared Imaging Radiometer Suite (VIIRS) platforms to ensure that no fire was misreported at the specific plots.

The aerial and ground-truth measurements were collected on the 21st October 2020. During the autumn, *T. capitatus* is dry and can be recognized and distinguished from S. *spinosum* through visual observation.

The vegetation in Lemnos is mainly low, with shrubs, bushes, and flowers, and includes the Aegean Sea plants such as thyme, oregano, and other herbs (*Panitsa et al., 2003*). Because it is a volcanic island, there is no forest but there are extensive fertile plains. Lemnos is characterized by a Mediterranean climate with mild winters, and a significant characteristic are the strong winds, especially during August and winter.

161

UAV Images acquisition and processing

A DJI P4 multispectral drone was used for the missions on Lemnos Island. The P4 multispectral is a tetra-copter rotary-wing aircraft with vertical takeoff and landing and can hover at low altitudes. It is equipped with one RGB sensor for visible light imaging and five monochrome sensors – BLUE (450nm ± 16nm), GREEN (560nm ± 16nm), RED (650nm ± 16nm), RED EDGE (730nm ± 16nm) and NIR (840nm ± 26nm) – for multispectral imaging plus a spectral sunlight sensor on top of the aircraft to compensate images while detecting the solar irradiance in real time. Flight paths were designed using DJI GSpro software, and missions were conducted between 11:00 am and 14:00 pm to reduce shading and at 20m altitude, 7-8 m/s wind speed, 90∘ course angle, 70% front overlap ratio and 60% side overlap ratio. The P4 multispectral camera was set to capture images at regular intervals parallel to the aircraft's flightline. Three orthophotos were derived from 3138 images acquired using a multispectral UAV with a ground sampling resolution of 0.8 cm for the classification method.

175

Ground Reference Survey

For this study, the training sample plots were chosen from three different regions on Lemnos Island using the DJI P4 multispectral UAV and orthophotos created in DJI TERRA software and analyzed manually in ArcGIS Pro software. The samples were classified into three main types: Thyme, Sarcopoterium, and Others. Within each sample plot, a sub-plot was visually identified and used as a training data set for the three selected classes. Ground points were added using a GNSS sensor

182   and used as known targets. Orthophotos and ground point targets were added in the ArcGIS Pro
183   software for the creation of shapefile polygons required by the GEE.
184
185   **Data Processing**
186   For this study the GEE was used for the UAV image processing. The GEE is a planetary-scale
187   platform for scientific analysis and visualization of geospatial datasets. In this platform, the open-
188   source images acquired by several satellites are accessible and can be efficiently imported and
189   processed in the cloud without the necessity of downloading the data to local computers. Several
190   image-driven products and many remote sensing algorithms, including classification algorithms
191   and cloud masking methods, are also available on this platform (*Gorelick et al., 2017; Kumar &*
192   *Mutanga, 2018*).
193   The first stage focuses on the processing of drone images to create GeoTIFFs, the format required
194   by the GEE, for each case study. The second stage is the collection of the training sample for each
195   category. The third stage includes the image classification, and the final steps focus on the
196   products' accuracy assessment.
197
198   **Vegetation indices**
199   In this study, we tested how different drone band combinations, referring both to single spectral
200   bands and spectral indices can affect the classification results and which can be considered as the
201   best combination resulting in higher thematic accuracy. According to the literature, existing studies
202   have shown that the use of spectral indices can increase classification accuracy (*Phan et al., 2020;*
203   *Pratico et al., 2021; Yang et al., 2021*). The spectral response of apicultural plants was estimated
204   using six commonly used vegetation indices (Vis). They were calculated from multispectral
205   images using the equations presented in Table 1. The spectral indices used in this study were added
206   as a new layer (band) to the image derived from the UAV.
207
208   **Supervised Classification of UAV orthophotos**
209   For the implementation of the pixel-based supervised classifications, training samples were
210   collected for the three study areas. The samples were allocated to one of three categories: Thyme,
211   Sarcopoterium and Other. In the Other category, we included features such as bare soil, shadows,
212   dry and other vegetation types. Samples were imported into the GEE platform to undergo
213   supervised classification of the orthophotos for each study area, SA1 (Figure 2A), SA2 (Figure 2B
214   and SA3 (Figure 2C).
215   The GEE platform uses a Classifier package that provides supervised classification by several
216   different traditional Machine Learning (ML) algorithms, running in Earth Engine. In this study,
217   we compared the performance of Random Forest (RF), Classification of Regression Trees
218   (CART), Gradient Tree Boostbased (GTB), Mahalanobis minimum-distance (MMD), and Support
219   Vector Machine (SVM) for their reliability in species classification.
220   **CART** is a non-parametric tree-based machine learning predictive algorithm used for training data
221   to generate prediction models. CART algorithm is based on a decision tree classifier, which is

222 made by a set of training data. CART, developed by *Breiman* (1984), facilitates simple decision
223 making and regression analyses, and is used for training data to generate prediction models. CART
224 algorithm runs by separating nodes until it reaches the terminal nodes. CART is based on a decision
225 tree where the producer models are acquired by recursively partitioning the data space and fitting
226 the classification model in each part to predict its class. CART also uses the surrogate separator to
227 make the best use of data with missing values. The advantage of CART is its ability to deal with a
228 variety of types of data, producing a complex decision flowchart like a tree structure through
229 learning the input dataset *(Breiman et al. 2017; Nakhaei et al, 2023)*.
230 **CART**, as previously described in *Neetu & Ray, (2019)* is "an advanced technique based on
231 Decision Tree (DT) classifier, which is built from a set of training data. The advantage of CART
232 is that it is simple in understanding, visualizing and interpreting. Both numerical and categorical
233 data can handle in CART".
234 **RF** is a tree-predictor algorithm whereby each object depends on the values of a random vector
235 sampled independently and with the same distribution for all objects *(Breiman, 2001)*. The RF has
236 the advantage of combining multiple decision trees and using average or voting schemes to
237 calculate the final results. This contrasts with a single decision tree and helps to reduce the
238 instability and sensitivity of the model while increasing its predictive power *(Vorpahl et al., 2012)*.
239 Using the tree rule, the decision-making process is repeated at each internal node from the root
240 node until the given termination condition is satisfied *(Ließ, Glaser & Huwe, 2012)*.
241 **GTB** is a mass tree-based algorithm (*Friedman, 2002*) that classifies the train weak learners
242 assigning them to stronger weights by adding new objects which decrease prediction errors
243 (*Callens, et al., 2020; Ustuner et al., 2019*). A set of logical if-then conditions (instead of linear
244 relationships) from the classification method and tree regression are then used to predict and
245 classify the result. The GTB method provides the capabilities of two algorithms: a) regression tree
246 models that respond to the prediction where they describe the operators through optimal and
247 boosting binary separation, and b) an adaptive method that achieves the combination of many
248 simple models that helps to obtain higher performance. It is an additive model where the gradient
249 identifies the shortcomings of the previous models. The residuals of the current classifier are used
250 as input to the next one in which trees are built, thus capturing the variation in the data (*Elith,
251 Leathwick & Hastie, 2008*).
252 **The Mahalanobis distance classifier** is a non-parametric and distance-based classifier that
253 assumes that all class fluctuations are equal. All pixels are classified in the nearest class (training
254 sample) unless a distance limit is set, in which case some pixels may not be classified if they do
255 not meet the limit (*Hengst, 2020*). Specifically, the Mahalanobis distance classifier has three main
256 steps: 1) computation of the mean vector in each band based on the application of the training
257 sample data; 2) calculation of the spectral distance between the mean vector of each sample and
258 the measurement vector of the candidate pixel; and 3) the pixel with the smallest spectral distance
259 is classified into the nearest class (*Perumal & Bhaskaran 2010*).
260 **SVM,** proposed by *Cortes & Vapnik* (1995), is a widely used algorithm for classifying data. It
261 attempts to create hyperplanes that best divide a dataset into different classes. The best hyperplane

262 is one that maximizes the margin between the closest data points from each class, which are called
263 support vectors (*Bishop*, 2006). However, the method only works for datasets that can be separated
264 linearly. To handle non-linear datasets, a non-linear kernel function is used to transform the data
265 into a high-dimensional space where it becomes linearly separable. The most commonly used
266 kernels are polynomial and radial basis function kernels. The formula for linear SVM is an
267 optimization problem that has an objective function which is the minimization of $min_{w,\xi_i,b}$

268 $\left\{\frac{1}{2}||w||^2 + c\sum_{I=1}^{M}\xi_i\right\}$ with the constraint that

269
270 $y_i\left((\varphi(x_i),w+b\right) \geq 1 - \xi_i, \ \xi_i \geq 0, \ i = 1,...,N$

271
272 The y_i is the class label, the ξ_i presents the positive slack variables, the C is the regularization and
273 b is the bias term (*Oommen et al., 2008; Cortes & Vapnik, 1995*).
274

275 **Accuracy Assessment of image classification and Statistical analysis**
276 For the classification, 70% of the total sample was used for training the GEE classifiers while the
277 remaining 30% was used for evaluation purposes (accuracy assessment). The evaluation was
278 achieved based on confusion matrix statistics allowing the calculation of the user's accuracy,
279 producer's accuracy, Kappa value and F-score per classifier.
280 Confusion matrix is estimated by comparing reference and classified results per thematic category
281 (i.e., class). Producer's accuracy estimates how well the reference pixels per class are correctly
282 classified, while user's accuracy estimates the probability that a pixel classified into a class
283 represents the same (true) class using the reference sample (*Lee et al., 2016, Lu & Weng, 2007,*
284 *Praticò et al., 2021*). The Kappa value is a measure of the classification performance compared to
285 values assigned by chance (*Lu & Weng, 2007*). It can take values from 0 to 1, in which a Kappa
286 coefficient close to 0 indicates no agreement between the classified results and the reference (truth)
287 data, while a Kappa coefficient close to 1 indicates an agreement between classified and reference
288 (truth) data. Finally, the so-called F-score, known as the harmonic mean of recall and precision, is
289 another classification accuracy metric with the same meaning as producer's accuracy (PA) and
290 user's accuracy (UA) respectively (*Praticò et al., 2021*).
291
292 Kappa Coefficient equation used for the training of the GEE classifiers.

293 $$OA = \frac{1}{N}\sum_{i=j=1}^{n} C_{ij}$$

294

295 $$PA = \frac{C_{ij}}{\sum_{i=j=1}^{n} C_j}$$

296

$$UA = \frac{C_{ij}}{\sum\limits_{i=j=1}^{n} C_i}$$

298

$$F = 2 * \frac{PA * UA}{PA + UA}$$

300

$$Kappa\ Coefficient = \frac{N \sum\limits_{i=j=1}^{n} C_{ij} - \sum\limits_{i=j=1}^{n} C_i C_j}{N^2 - \sum\limits_{i=j=1}^{n} C_i C_j}$$

302

Where,

N	Number of rows in error matrix
C_{ij}	Number of observations in row and column
C_i	Total number of observations in row
C_j	Total number of observations in column
N	Total number of observations
PA	Producer's accuracy
UA	User's Accuracy

303
304

Results

306

Accuracy assessment

For detecting and distinguishing apicultural plants such as *T. capitatus* from non-apicultural plants, such as *S. spinosum*, we trained samples of *T. capitatus* and *S. spinosum* using the classification accuracy assessment derived from the confusion matrix. Different band compositions (DJi bands + spectral indices) were implemented for each case study, and the GEE provided the same training samples and validation data for five different classifiers.

The evaluation of the classification results accuracy for each classifier as well as for each band combination with spectral indices, the PA, UA and F-score parameters, were estimated for the three classes separately. The Total Accuracy estimation was derived from each classifier and each combination based on the OA and Kappa parameters. Figure 3 presents the results for each study

317   area where the stacked columns show the PA, UA and F-score parameters for each class separately.
318   PA, UA and F-score parameter give values from 0 to 1, so if the three parameters together are
319   close to the value three, they have achieved high values in the classification. When the score
320   approaches the value two on the Y axis, the classifier achieves the higher accuracy from the OA
321   and Kappa parameters.
322   For the SA1 study area, the RF and GTB classifiers presented the highest accuracy when using the
323   combinations *ALL BANDS + VEGETATION INDICES, VIS + NIR + RE and VIS + NIR bands,*
324   which presented the same behavior with kappa values = 0.97 and OA= 0.98 in all cases. The lowest
325   accuracy was presented by the SVM classifier when combined with the VIS classifier (Kappa=0.42
326   and OA=0.61) and with the *VIS + NIR, VIS + NIR + RE* where in both cases they scored
327   kappa=0.62 and OA=0.74 values. Low performance was also presented by the MMD classifier
328   with performance kappa=0.66 and OA=0.77.
329   For the SA2 study area, the highest accuracy was achieved using the RF classifier based on the
330   combinations *VIS + NIR + RE, VIS + NIR plus ALL BANDS + VEGETATION INDICES*, and GTB
331   with the combination *ALL BANDS + VEGETATION INDICES*, where they presented similar
332   behavior, with Kappa values=0.90 and OA=0.93 in all cases. The lowest accuracy was presented
333   by the SVM classifier mainly with the VIS combination (kappa=0.47 and OA=0.65). The MMD
334   classifier also showed low performance. In this case, the CART classifier also showed a relatively
335   low performance when the *VEGETATION INDICES* combination was used providing Kappa =
336   0.75 and OA = 0.83 values.
337   In the third study area (SA3), the highest accuracy was provided by GTB and RF classifiers when
338   combined with *ALL BANDS + VEGETATION INDICES, VIS + NIR + RE and VIS + NIR* with
339   Kappa = 0.92 and OA = 0.95 values. The lowest accuracy was presented by the SVM classifier
340   when combined with the VIS (Kappa=0.58 and OA=0.72). Low performance was achieved when
341   the SVM classifier was combined with the *VIS + NIR + RE and VIS + NIR*. The MMD classifier
342   also performed low accuracy (Kappa 0.68-0.73 and OA=0.79-0.82).
343

344   **Classification performance between Thyme, Sarcopoterium and Other in Google Earth**
345   **Engine**
346   For the three case studies, the Total Accuracy for the five classifiers (RF, GTB, CART, MMD and
347   SVM), was calculated using the average value from the Kappa and OA parameters. On average
348   the GTB and RF classifiers presented similar behavior with Kappa=0.91 and OA=0.94 values.
349   Also, the average of the CART classifier presented accuracies that belong to the "near-perfect
350   agreement" category with the average values for Kappa and OA amounting for 0.88 and 0.92,
351   respectively. The classifiers which presented the lowest average were SVM (Kappa = 0.70 and
352   OA=0.80) and MMD (Kappa=0.78 and OA=0.83). In addition, the average of the accuracy
353   (obtained from the various combinations of spectral channels and vegetation indices was
354   calculated with the combination of *ALL BANDS + VEGETATION INDICES*), scored the highest
355   accuracy (Kappa=0.90 and OA=0.93). The combinations VIS + NIR and VIS + NIR + RE gave
356   similar results with an average accuracy of Kappa=0.84 and OA=0.89, while the combinations

357 *VEGETATION INDICES* presented average accuracy of Kappa=0.82 and OA=0.88. The VIS
358 presented average accuracy of Kappa=0.78 and OA=0.85 providing the lowest returns on average
359 (Figure 3).
360 To sum up, the findings of the present study show that the GTB and RF classifiers together with
361 the *ALL BANDS + VEGETATION INDICES* combination is considered ideal for the classification
362 of beekeeping plants and more specifically for the separation of *T. capitatus* and *S. spinosum.*
363
364 **Detection and distinguishing maps**
365 The classification maps retrieved from GEE after the selection of the GTB classifier for the
366 detection and distinguishing of the two shrub species (Thyme and Sarcopoterium) are presented in
367 Figure 4. In the three sample plots (SA1, SA2 and SA3), the *T. capitatus* plants are shown in
368 purple, *S. spinosum* plants in yellow, and all the other vegetation species and bare soil are shown
369 in black. The roads are masked out to improve the accuracy of the results. Based on these maps
370 for each case study, the total area covered by each class was estimated. The area covered by each
371 species is shown in Figure 4. The evaluation of the total area for each class was for SA1: Thyme
372 = 0.21ha, Sarcopoterium = 0.10 ha and Other = 0.07 ha (Figure 4A), SA2: Thyme = 0.13 ha,
373 Sarcopoterium = 0.04 ha, other = 0.02 (Figure 4B), SA3: Thyme = 0.52 ha, Sarcopoterium = 0.26
374 ha and Other = 0.15 ha (Figure 4C). These figures were derived using the GTB classifier.
375 While running the Kappa coefficient algorithm, known sample points and combinations of UAV
376 bands and VIs were added to the algorithm until the Kappa coefficient reached 81% or higher. Of
377 the five classifiers used in this study (RF, GTB, CART, MMD and SVM), the algorithms which
378 are based on the tree-based decision give the best outcomes. The Gradient Tree Boost and the RF
379 classifier achieved the higher Overall Accuracy and Kappa coefficient value when used for the
380 three sample plots but also when used among all five UAV bands and the five VIs combinations.
381 Based on the results, for the first three classifiers with the best performances, a variable importance
382 chart was created. As shown in Figure 4, the variable importance shows that RF classifier uses all
383 the variables (UAV bands and VIs), in contrast to GTB and CART classifiers, in which the RVI
384 and CI-RE indices do not contribute to the prediction analysis. The variable importance plot is a
385 very useful tool to present the variables that can make accurate predictions when using the model.
386

Discussion

388 The present study aimed to investigate the feasibility of detecting and distinguishing apicultural
389 plants from non-apicultural plants using UAV aerial multispectral images through GEE platform
390 in traditional beekeeping areas, such as Lemnos Island. Results showed that such detection and
391 identification is possible. From the five classification algorithms (RF, CART, GTB, MMD and
392 SVM) used to train the classification model, RF had the highest accuracy values and this is
393 confirmed by *Pérez-Cutillas et al.* (2023) who found that among the non-parametric classification
394 methods the most frequently used algorithm was the Random Forest, with 31% of the cases used.
395 This study employs five different machine-learning algorithms because the existing literature is
396 contradictory in terms of the classification accuracy of various classifiers for vegetation species.

397   For example, *Elkind et al.* (2019) used Random Forest models on the GEE platform to classify
398   buffelgrass in Saguaro National Park with an average overall accuracy of 93%. However, other
399   studies have found different classifiers to have better performance. For example, *Sujud et al.*
400   (2021) found that the SVM classifier had the highest overall accuracy for cannabis detection, while
401   *Mustak et al.* (2019) found that SVM outperformed RF and CART for crop discrimination. In
402   contrast, *Wahap & Shafri* (2020) found that CART had the best accuracy for land cover monitoring
403   compared to RF and SVM.
404   The performance of the three classifiers along with the evaluation of the different UAV bands and
405   Vegetation Indices and their combination for the detection and distinguishing of *T. capitatus* and
406   *S. spinosum* were the main goals of this study. The resulting metrics presented high accuracy
407   values for all three classifiers, with RF providing the highest values for overall accuracy and Kappa
408   coefficients. The overall accuracy in distinguishing *T. capitatus* from *S. spinosum* when using RF
409   was OA = 98% and Kappa = 0.97 for SA1, OA = 93% and Kappa = 0.90 for SA2, and OA = 95%
410   and Kappa = 0.92 for SA3. As a result, it was possible to distinguish between the apicultural and
411   non-apicultural species.
412   In apiculture, UAVs have been used as a tool for monitoring drone congregation areas (DCAs –
413   where male honeybees congregate for mating) (*Carl et al., 2017*) to provide useful information for
414   honeybee queen-rearing and breeding. Furthermore, UAVs have been used for the tracking of
415   hornets *Vespa velutina* (*Reynaud & Guérin-Lassous, 2017*). *V. velutina* is an invasive species in
416   Europe (*Rortais et al., 2010*) and can cause severe losses in *Apis mellifera* colonies (*Arca et al.,*
417   *2014*). Being able to track such predators could help in the control of their spread and limit their
418   impact on honeybees. The use of UAVs equipped with multispectral sensors is nowadays used as
419   a tool for monitoring invasive plants (*Elkind et al., 2019*) and wild plant species in inaccessible
420   areas and with zero environmental disturbance (*Questad et al., 2022*). Success in identifying
421   Thyme plants, the dominant apicultural plant of Lemnos Island and most Greek Islands, is of great
422   importance for beekeeping in Greece. Even though *T. capitatus* is protected by Greek legislation,
423   many Thyme areas are destroyed every year (by farmers, herdsmen etc) on the assumption that
424   they are not Thyme areas but just areas of scrub vegetation. Fire services usually claim that they
425   cannot recognize the difference between a Thyme plant and other scrubs and, as a result, they
426   cannot take any action against the destruction of *T. capitatus* areas. Our work could be extended
427   to applications such as the mapping of *T. capitatus* areas in the entire country and lead to the
428   monitoring and protection of such areas. Furthermore, since competition between apicultural plants
429   and non-apicultural plants is intensified as a result of plant stress caused mainly by climate change
430   and wildfires (*Kosmas et al., 2015*), the identification and protection of Thyme areas will help
431   protect honeybee populations that are under threat. Previous research (*Vieira et al., 2013*),
432   investigated the consequences of plant extinctions on pollinators and produced prediction models
433   in which changes in plant diversity (especially the extinction of *T. capitatus*) could lead to the
434   coextinction of pollinators (including honeybees). The encroachment of Sarcopoterium in areas of
435   Thymus species is not easily detected or managed during the early invasion stages because of the
436   morphological similarity of the two species.

437   Remote detection and distinguishing of the two species using UAV images and machine learning
438   techniques might allow researchers, beekeepers, and state authorities to monitor the growth and
439   early expansion of *S. spinosum* over *T. capitatus* after a catastrophic fire events and provide real-
440   time data enabling timely control measures. Accurate estimation of *T. capitatus* distribution is
441   fundamental in maintaining and developing the quantity and quality of the Lemnian honey. It
442   would also benefit beekeepers who need to move their hives in search of *T. capitatus* blossom. In
443   a similar study, UAV aerial images have been used to estimate flower numbers and their potential
444   nectar yield, and to identify areas of *Robinia pseudoacacia* suitable for honeybee foraging (*Carl*
445   *et al., 2017*).
446   The continuing monitoring of historic Thymus lands will provide valuable environmental
447   information and enable estimates of Lemnian wild Thyme honey production.
448

Conclusions

450   The aim of the present study was to detect apicultural plants and to distinguish them from plant
451   species not suitable for apiculture using a pixel-based approach. It explored the capability of the
452   GEE cloud platform for apicultural plant species classification. For this purpose, *T. capitatus* and
453   *S. spinosum* were chosen and scanned using a five-band multispectral drone in three areas of
454   Lemnos Island.
455   From the five classification algorithms (RF, CART, GTB, MMD and SVM) used to train the
456   classification model, the RF and GTB classifiers in combination with the *ALL BANDS +*
457   *VEGETATION INDICES, VIS + NIR + RE* and */or VIS + NIR bands,* provided the highest accuracy
458   values as shown in the Kappa coefficient and overall accuracy, in all three study sites and manage
459   to detect and discriminate *T. capitatus* from *S. Spinosum*. As described in the literature, the RF and
460   GTB classifiers were effective for distinguishing the two species and thus should be used to
461   identify extensive areas of apicultural plants and to monitor bee habitats This research confirmed
462   that cloud computing architectures in combination with machine learning algorithms can very
463   accurately and quickly identify invasive species.
464   The use of a tetra-copter drone flying at a low altitude and equipped with a multispectral camera
465   can provide high resolution images. However, multiple flights are necessary. To overcome this
466   limitation, we propose the use of a vertical take-off and landing (VTOL) UAV with a high-
467   resolution coverage. According to *Awad (2018),* the multispectral images retrieved from satellites
468   do not provide complete separation of forest plant species and thus the hyperspectral sensor is
469   proposed as a better plant species separator due to its almost continuous spectra, covering spectral
470   details that might pass unnoticeable when using the multispectral sensor *(Adão et al., 2017).* A
471   limitation of hyperspectral sensors attached on a UAV is their high cost compared to a low-cost
472   multispectral sensor. A tetracopter UAV is low cost and suitable for small areas but costly for
473   extensive mapping because of the need for multiple flights. A further study will use of VTOL
474   UAV for the detection of other apicultural plants and to distinguish them from invasive plant
475   species in the beekeeping areas in Greece and other Mediterranean regions.
476

477 **Acknowledgements**

478 The authors acknowledge the Agricultural and Apicultural Cooperation of Lemnos Beekeepers
479 and especially their president, Dimitris Paleologos for supporting our field work. Furthermore, the
480 authors acknowledge the 'EXCELSIOR': ERATOSTHENES: EXcellence Research Centre for
481 Earth Surveillance and Space-Based Monitoring of the Environment H2020 Widespread Teaming
482 project (www.excelsior2020.eu). A.P. thanks Vita Bee Health for supporting his research. We also
483 thank Stephen Fleming for editing the manuscript.

484 **References**

485 **Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Zheng X. 2016**. TensorFlow: A system
486 for large-scale machine learning. In 12th {USENIX} Symposium on Operating Systems Design
487 and Implementation ({OSDI} 16) (pp. 265–283). *Savannah, GA: {USENIX} Association*

488 **Adão T, Hruška J, Pádua L, Bessa J, Peres E, Morais R, Sousa JJ. 2017.** Hyperspectral
489 imaging: A review on UAV-based sensors, data processing and applications for agriculture and
490 forestry. *Remote sensing*, **9**: 1110.

491 **Anderson K, Gaston KJ. 2013**. Lightweight unmanned aerial vehicles will revolutionize spatial
492 ecology. *Frontiers in Ecology and the Environment* **11**: 138-146.

493 **Arasumani M, Singh A, Bunyan M, Robin VV. 2021.** Testing the efficacy of hyperspectral
494 (AVIRIS-NG), multispectral (Sentinel-2) and radar (Sentinel-1) remote sensing images to detect
495 native and invasive non-native trees. *Biological Invasions* **23**: 2863–2879.

496 **Arianoutsou M, Margaris NS. 1982**. Phryganic (East Mediterranean) ecosystems and fire.
497 *Ecologia Mediterranea* **8**: 473-480.

498 **Arca M, Papachristoforou A, Mougel F, Rortais A, Monceau K, Bonnard O, Arnold G. 2014.**
499 Defensive behaviour of *Apis mellifera* against *Vespa velutina* in France: Testing whether European
500 honeybees can develop an effective collective defence against a new predator. *Behavioural*
501 *Processes* **106**: 122–129.

502 **Awad, MM. 2018**. Forest mapping: a comparison between hyperspectral and multispectral images
503 and technologies. *J. For. Res.* **29**, 1395–1405

504 **Barnes EM, Clarke TR, Richards SE. 2000**. Coincident detection of crop water stress, nitrogen
505 status and canopy density using ground based multispectral data. In: Robert PC, Rust RH, Larson
506 WE. (eds), Proceedings of the fifth international conference on precision agriculture., Madison,
507 WI, USA: American Society of Agronomy, Unpaginated CD.

508 **Barrero O, Perdomo SA. 2018**. RGB and multispectral UAV image fusion for Gramineae weed
509 detection in rice fields. *Precision Agric*ulture **19**: 809–822.

510 **Bishop CM. 2006.** Pattern Recognition and Machine Learning. Springer, New York, pp 738

511 **Bolch EA, Hestir EL, Khanna S. 2021**. Performance and feasibility of drone-mounted imaging
512 spectroscopy for invasive aquatic vegetation detection. *Remote Sensing* **13**: 582.

513 **Breeze TD, Bailey AP, Balcombe KG, Potts SG. 2011**. Pollination services in the UK: How
514 important are honeybees? Agriculture, Ecosystems and Environment **142**: 137–143.

515 **Breiman L. 2001**. Random forests. *Machine Learning* **45**: 5–32.

516 **Breiman L 1984.** Classification and Regression Trees. 1st Edition ed. *Chapman & Hall, New York*

517 **Callens A, Morichon D, Abadie S, Delpey M, Liquet B. 2020**. Using Random forest and
518 Gradient boosting trees to improve wave forecast at a specific location. *Applied Ocean Research*
519 **104**: 102339.
520 **Carl C, Landgraf D, Van der Maaten-Theunissen M, Biber P, Pretzsch H. 2017.** *Robinia*
521 *pseudoacacia* L. flower analyzed by using an Unmanned Aerial Vehicle (UAV). *Remote Sensing*
522 **9**: 1091.
523 **Cohen J. 1960.** Combined Local and Overall Bulcking of Rectangular Thin-Walled Tubular
524 Columns. University of Illinois at Urbana-Champaign ProQuest Dissertations Publishing.
525 **Cramp DC. 2017.** The use of an unmanned aerial vehicle (UAV) to investigate aspects of honey
526 bee drone congregation areas (DCAs). *Journal of Apicultural Research* **56**: 172-174.
527 **Cortes, C. and Vapnik, V. 1995**. Support-vector networks. *Machine learning*, *20* **(3),** pp.273-
528 297.
529 **da Silva SDP, Eugenio FC, Fantinel RA, Amaral L de P, dos Santos AR, Mallmann CL, dos**
530 **Santos FD, Pereira RS, Ruoso R. 2023.** Modeling and detection of invasive trees using UAV
531 image and machine learning in a subtropical forest in Brazil. *Ecological Informatics* 74:101989.
532 **Daughtry C. 2000**. Estimating corn leaf chlorophyll concentration from leaf and canopy
533 reflectance. *Remote Sensing of Environment* **74**: 229–239.
534 **Dmitriev PA, Kozlovsky BL, Kupriushkin DP, Dmitrieva AA, Rajput VD, Chokheli VA,**
535 **Tarik EP, Kapralova OA, Tokhtar VK, Minkina TM, Varduni TV. 2022.** Assessment of
536 invasive and weed species by hyperspectral imagery in agrocenoses ecosystem. *Remote Sensing*
537 **14**: 2442.
538 **Donkersley P, Rhodes G, Pickup RW, Jones KC, Wilson K. 2014.** Honeybee nutrition is linked
539 to landscape composition. *Ecology and Evol*ution **4:** 4195–206.
540 **Duan, B.; Fang, S.; Gong, Y.; Peng, Y.; Wu, X.; Zhu, R. 2021**. Remote estimation of grain yield
541 based on UAV data in different rice cultivars under contrasting climatic zone. *Field Crop.*
542 *Res. 267,* 108148–108166.
543 **Elkind K, Sankey TT, Munson SM, Aslan CE. 2019**. Invasive buffelgrass detection using high-
544 resolution satellite and UAV imagery on Google Earth Engine. *Remote Sensing of Ecology and*
545 *Conservation* **5**: 318-331.
546 **Fariz TR, Suhardono S, Sultan H, Rahmawati D, Arifah EZ. 2022.** Land cover mapping in
547 lake Rawa pening using Landsat 9 Imagery and Google Earth Engine, *Journal of Environmental*
548 *and Science Education* **2**: 1–6.
549 **FIRMS**. Available online: https://firms.modaps.eosdis.nasa.gov. Assessed on 14 September 2021
550 **Forster M, Schmidt T, Wolf R, Kleinschmit B, Fassnacht FE, Cabezas J, Kattenborn, T.**
551 **2017**. Detecting the spread of invasive species in central Chile with a Sentinel-2 time-series. In:
552 *9th International Workshop on the Analysis of Multitemporal Remote Sensing Images*
553 *(MultiTemp)*. IEEE: 1–4.
554 **Friedman JH. 2002**. Stochastic gradient boosting. *Computational Statistics and Data Anal*ysis
555 **38**: 367–378.
556 **Gašparovičová P, Ševčík M, David S. 2022.** The prediction of distribution of the invasive

557   Fallopia taxa in Slovakia. *Plants* **11**: 1484.

558   **Gitelson AA, Gritz Y, Merzlyak MN. 2003**. Relations between leaf chlorophyll content and
559   spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant
560   leaves. *Journal of Plant Physiology* **160**: 271-282.

561   **Gitelson AA, Viña A, Ciganda V, Rundquist DC, Arkebauer TJ. 2005**. Remote estimation of
562   canopy chlorophyll content in crops. *Geophysical Research Letters* **32**: 1–4.

563   **Google Earth Engine** (GEE). Available online: https://earthengine.google.com . Assessed on 20
564   May 2022.

565   **Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. 2017**. Google Earth
566   Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* **202**:
567   18–27.

568   **Große-Stoltenberg A, Hellmann C, Werner C, Oldeland J, Thiele J. 2016**. Evaluation of
569   continuous VNIR-SWIR spectra versus narrowband hyperspectral indices to discriminate the
570   invasive *Acacia longifolia* within a Mediterranean dune ecosystem. *Remote Sensing* **8**: 334.

571   **Hengst AM. 2020.** Sub-Annual to Annual Dynamics of Alaskan Ice-Marginal Lakes from
572   Automated Image Classification Using Google Earth Engine. Appalachian State University, pp.
573   46.

574   **Henkin Z, Rosenzweig T, Yaniv Z. 2014.** *Sarcopoterium spinosum.* In: *Medicinal and Aromatic*
575   *Plants of the Middle-East*, Medicinal and Aromatic Plants of the World 2, Yaniv, Z.; Dubai, N.
576   Eds.; Spiringer Science-Business Media Dordrecht, Netherlands, pp. 151-163.

577   **Hestir EL, Khanna S, Andrew ME, Santos MJ, Viers JH, Greenberg JA, Rajapakse SS, Ustin**
578   **SL. 2008.** Identification of invasive vegetation using hyperspectral remote sensing in the
579   California delta ecosystem. *Remote Sensing of Environment* **112**: 4034–4047.

580   **Hill DJ, Tarasoff C, Whitworth GE, Baron J, Bradshaw JL, Church JS. 2017**. Utility of
581   unmanned aerial vehicles for mapping invasive plant species: a case study on yellow flag iris (*Iris*
582   *pseudacorus L.*). *International Journal of Remote Sensing* **38**: 2083–2105.

583   **Hill J, Hostert P, Tsiourlis GM, Kasapidis P, Udelhoven T, Diemer C. 1998.** Monitoring
584   twenty years of in- creased grazing pressure on the Greek island of Crete with earth observation
585   satellites. *Journal of Arid Enviroments* **39**: 165–178.

586   **Ishak AJ, Tahir MDN, Hussain A, Mustafa MM. 2008.** Weed classification using Decision
587   Tree. in *2008 International Symposium on Information Technology*. IEEE: 1–5.

588   **Ishida T, Kurihara J, Viray FA, Namuco SB, Paringit EC, Perez GJ, Takahashi Y, Marciano**
589   **JJ. 2018.** A novel approach for vegetation classification using UAV-based hyperspectral imaging.
590   *Computers and Electronics in Agriculture* **144**: 80–85.

591   **Kattenborn T, Lopatin J, Förster M, Braun AC, Fassnacht FE. 2019.** UAV data as alternative
592   to field sampling to map woody invasive species based on combined Sentinel-1 and Sentinel-2
593   data. *Remote Sensing of Environment* **227**: 61–73.

594   **Khalifa SAM, Elshafiey EH, Shetaia AA, El-Wahed AAA, Algethami AF, Musharraf SG,**
595   **AlAjmi MF, Zhao C, Masry SHD, Abdel-Daim MM, Halabi MF, Kai G, Al Naggar Y, Bishr**

596 **M, Diab MAM, El-Seedi HR. 2021**. Overview of bee pollination and its economic value for crop
597 production. *Insects* **12**: 688.

598 **Kosmas C, Detsis V, Karamesouti M, Kounalaki K, Vassiliou P, Salvati L. 2015**. Exploring
599 long-term impact of grazing management on land degradation in the socio-ecological system of
600 Asteroussia Mountains, Greece. *Land* **4**: 541-559.

601 **Kumar L, Mutanga O. 2018**. Google Earth Engine applications since inception: Usage, trends,
602 and potential. *Remote Sensing* **10**: 1509.

603 **Lee JSH, Wich S, Widayati A, Koh LP. 2016.** Detecting industrial oil palm plantations on
604 Landsat images with Google Earth Engine. *Remote Sensing Applications: Society and*
605 *Environment* **4**: 219–224.

606 **Lewis HG, Brown MA. 2001**. Generalized confusion matrix for assessing area estimates from
607 remotely sensed data. *International Journal of Remote Sensing* **22**: 3223-3235.

608 **Louargant M, Villette S, Jones G, Vigneau N, Paoli J, Gée C. 2017**. Weed detection by UAV:
609 Simulation of the impact of spectral mixing in multispectral images. *Precision Agriculture* **18**:
610 932–951.

611 **Lu B, He Y. 2017.** Species classification using Unmanned Aerial Vehicle (UAV)-acquired high
612 spatial resolution imagery in a heterogeneous grassland. *ISPRS Journal of Photogrammetry and*
613 *Remote Sensing* **128**: 73-85.

614 **Lu D, Weng Q. 2007**. A survey of image classification methods and techniques for improving
615 classification performance. *International Journal of Remote Sensing* **28**: 823–870.

616 **Marnasidis S, Kantartzis A, Malesios C, Hatjina F, Arabatzis G, Verikouki E. 2021.** Mapping
617 priority areas for apiculture development with the use of Geographical Information Systems.
618 *Agriculture* **11**: 182.

619 **Mustak S, Uday G, Ramesh B, Praveen B. 2019.** Evaluation of the performance of SAR and
620 SAR-optical fused dataset for crop discrimination. *International Archives of the Photogrammetry,*
621 *Remote Sensing and Spatial Information Sciences - ISPRS Archives* 42:563–571.

622 **Nakhaei M, Tafreshi AM, Saadi T. 2023.** An evaluation of satellite precipitation downscaling
623 models using machine learning algorithms in Hashtgerd, Iran. *Modeling Earth Systems and*
624 *Environment*, 9: 1-15

625 **Neetu, Ray SS. 2019**. Exploring machine learning classification algorithms for crop classification
626 using Sentinel 2 data, *ISPRS,* International Archives of the Photogrammetry, Remote Sensing &
627 Spatial Information Sciences. *XLII-3/W6*: 573–578.

628 **Niphadkar M, Nagendra H. 2016.** Remote sensing of invasive plants: incorporating functional
629 traits into the picture. *International Journal of Remote Sensing* **37**: 3074–3085.

630 **Pal M, Mather PM. 2005.** Support vector machines for classification in remote sensing.
631 *International Journal of Remote Sensing* **26**: 1007–1011.

632 **Panitsa M, Snogerup B, Snogerup S, Tzanoudakis D. 2003**. Floristic investigation of Lemnos
633 Island (NE Aegean area, Greece). *Willdenowia* **33**: 79–105.

634 **Papp L, van Leeuwen B, Szilassi P, Tobak Z, Szatmári J, Árvai M, Mészáros J, Pásztor L.**
635 **2021**. Monitoring invasive plant species using hyperspectral remote sensing data. *Land* **10**: 29.

636 **Paz-Kagan T, Silver M, Panov N, Karnieli A. 2019.** Multispectral approach for identifying
637 invasive plant species based on flowering phenology characteristics. *Remote Sensing* **11**: 953.
638 **Pérez-Cutillas P**, **Pérez-Navarro A, Conesa-García C, Zema DA, Amado-Álvarez JP.** 2023.
639 What is going on within google earth engine? A systematic review and meta-analysis. *Remote*
640 *Sensing Applications: Society and Environment* **29**:100907.
641 **Perumal, K. and Bhaskaran, R., 2010.** Supervised classification performance of multispectral
642 images. *Journal of Computing* **2**, 124-129.
643 **Phan TN, Kuch V, Lehnert LW. 2020**. Land cover classification using Google Earth Engine and
644 random forest classifier—The role of image composition. *Remote Sensing* **12**, 2411.
645 **Praticò S, Solano F, Di Fazio S, Modica G. 2021a**. Machine learning classification of
646 Mediterranean forest habitats in Google Earth Engine based on seasonal Sentinel-2 time-series and
647 input image composition optimisation. *Remote Sensing* **13**: 586.
648 **Pretorius E, Pretorius R. 2015**. Improving the potential of pixel-based supervised classification
649 in the absence of quality ground truth data. *South African Journal of Geomatics* **4**: 250-263.
650 **Questad EJ, Antill M, Liu N, Stavros EN, Townsend PA, Bonfield S, Schimel DA. 2022**.
651 Camera-based method for collecting rapid vegetation data to support remote-sensing studies of
652 shrubland biodiversity. *Remote Sensing* **14**: 1933.
653 **Raj A, Sharma LK. 2022**. Assessment of land-use dynamics of the Aravalli range (India) using
654 integrated geospatial and CART approach. *Earth Science Informatics* **15**: 497–522.
655 **Rango A, Laliberte A, Steele C, Herrick JE, Bestelmeyer B, Schmugge T, Roanhorse A,**
656 **Jenkins V. 2006.** Using unmanned aerial vehicles for rangelands: current applications and future
657 potentials. *Environmental Practice* **8**:159-168.
658 **Rasmussen J, Nielsen J, Garcia-Ruiz F, Christensen S, Streibig J.C. 2013.** Potential uses of
659 small, unmanned aircraft systems (UAS) in weed research. *Weed Research* **53**: 242–248.
660 **Reynaud B, Guérin-Lassous I. 2016.** Design of a force-based controlled mobility on aerial
661 vehicles for pest management. *Ad Hoc Networks* **53**: 41–52.
662 **Robinson TP, Wardell-Johnson GW, Pracilio G, Brown C, Van Klinken RD. 2016.** Testing
663 the discrimination and detection limits of worldview-2 imagery on a challenging invasive plant
664 target. I*nternational Journal of Applied Earth Observation and Geoinformation.* **44**: 23–30.
665 **Rortais A, Villemant C, Gargominy O, Rome Q, Haxaire J, Papachristoforou A, Arnold G.**
666 **2010.** A new enemy of honeybees in Europe: The Asian hornet *Vespa velutina*. In: *Atlas of*
667 *Biodiversity Risks—From Europe to the Globe, From Stories to Maps*; Settele, J., Ed.; Pensoft:
668 Moscow, Russia, p. 11.
669 **Rouse JW, Haas RH, Schell JA, Deering DW, Harlan JC.** 1974. Monitoring the vernal
670 advancements and retrogradation of natural vegetation. *NASA/GSFC, Final Report, Greenbelt,*
671 *MD, USA*. 1–137.
672 **Senthilnath J, Kandukuri M, Dokania A, Ramesh KN. 2017**. Application of UAV imaging
673 platform for vegetation analysis based on spectral-spatial methods. *Computers and Electronics in*
674 *Agriculture* **140**: 8–24.
675 **Sheffield KJ, Clements D, Clune DJ, Constantine A, Dugdale TM. 2022.** Detection of aquatic
676 alligator weed (*Alternanthera philoxeroides*) from aerial imagery using random forest

677 classification. *Remote Sensing* **14**: 2674.

678 **Singh V, Singh D 2022.** Development of an approach for early weed detection with UAV
679 omagery. in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*
680 IEEE: 4879–4882.

681 **Skowronek S, Asner GP, Feilhauer H. 2017**. Performance of one-class classifiers for invasive
682 species mapping using airborne imaging spectroscopy. *Ecological Informatics* **37**: 66–76.

683 **Sujud L, Jaafar H, Haj Hassan, MA, Zurayk R. 2021**. Cannabis detection from optical and
684 RADAR data fusion: A comparative analysis of the SMILE machine learning algorithms in Google
685 Earth Engine. *Remote Sensing Applications: Society and Environment* **24**: 100639.

686 **Torresan C, Berton A, Carotenuto F, Di Gennaro SF, Gioli B, Matese A, Miglietta F, Vagnoli**
687 **C, Zaldei A, Wallace L.** 2017. Forestry applications of UAVs in Europe: A review. *International*
688 *Journal of Remote Sensing* 38: 2427-2447.

689 **Traganos D, Aggarwal B, Poursanidis D, Topouzelis K, Chrysoulakis N, Reinartz P. 2018.**
690 Towards global-scale seagrass mapping and monitoring using Sentinel-2 on Google Earth Engine:
691 The case study of the Aegean and Ionian Seas. *Remote Sensing* **10**: 1227.

692 **Tucker CJ. 1979**. Red and photographic infrared linear combinations for monitoring vegetation.
693 *Remote Sensing of Environment* **8**: 127-150.

694 **Ustuner M, Balik Sanli F. 2019**. Polarimetric target decompositions and light Gradient boosting
695 machine for crop classification: A comparative evaluation. *ISPRS International Journal of Geo-*
696 *Information* **8**: 97.

697 **Van den Elsen E, Stringer LC, De Ita C, Hessel R, Kéfi S, Schneider FD, Bautista S, Mayor**
698 **AG, Baudena M, Rietkerk M, Valdecantos A, Vallejo VR, Geeson N, Brandt CJ, Fleskens L,**
699 **Hemerik L, Panagos P, Valente S, Keizer JJ, Schwilch G, Jucker Riva M, Sietz D,**
700 **Christoforou M, Hadjimitsis DG, Papoutsa C, Quaranta G, Salvia R, Tsanis IK,**
701 **Daliakopoulos I, Claringbould H, de Ruiter PC. 2020**. Advances in understanding and
702 managing catastrophic ecosystem shifts in Mediterranean ecosystems. *Frontiers in Ecology and*
703 *Evolution* **8**: 561101.

704 **Vieira MC, Cianciaruso MV, Almeida-Neto M. 2013**. Plant-pollinator coextinctions and the loss
705 of plant functional and phylogenetic diversity. *PLoS ONE* **8**: e81242

706 **Yang C, Everitt JH. 2010**. Mapping three invasive weeds using airborne hyperspectral imagery.
707 *Ecological Informatics* **5**: 429–439.

708 **Yang L, Driscol J, Sarigai S, Wu Q, Chen H, Lippitt CD. 2022**. Google Earth Engine and
709 artificial intelligence (AI): A comprehensive review. *Remote Sensing* **14**: 3253.

710 **Yang Y, Yang D, Wang X, Zhang Z, Nawaz Z. 2021**. Testing accuracy of land cover
711 classification algorithms in the Qilian mountains based on GEE cloud platform. *Remote Sensing*
712 **13**: 5064.

713 **Zhu W, Ren, G, Wang J, Wang J, Hu Y, Lin Z, Li W, Zhao Y, Li S, Wang N. 2022.** Monitoring
714 the invasive plant *Spartina alterniflora* in Jiangsu coastal wetland using MRCNN and Long-Time
715 Series Landsat Data. *Remote Sensing* **14**: 2630.

716 **Wahap N., Shafri HZM. 2020.** Utilization of Google Earth Engine (GEE) for land cover

717   monitoring over Klang Valley, Malaysia. *IOP Conference Series: Earth and Environmental*
718   *Science* **540**: 012003.

Kappa Value interpretation

Table 1. Kappa Value interpretation, (*Cohen, 1960*)

1   Table 1. Kappa Value interpretation, (*Cohen, 1960*)

Kappa	Interpretation
0	No agreement
0,0 - 0,20	Slight agreement
0,21-0,40	Fair agreement
0,41-0,60	Moderate agreement
0,61-0,80	Substantial agreement
0,81-0.99	Near perfect agreement
1	Perfect agreement

Figure 1

Location of the experimental plots

Figure 1. The location of sample plots. (A) sample plots in southeast part of the island. (B) Lemnos Island and the experimental area. (C) the island of Lemnos in the northeast Aegean Sea, Greece. Source: Esri, Maxar, Earthstar Geographics, and the GIS User Community
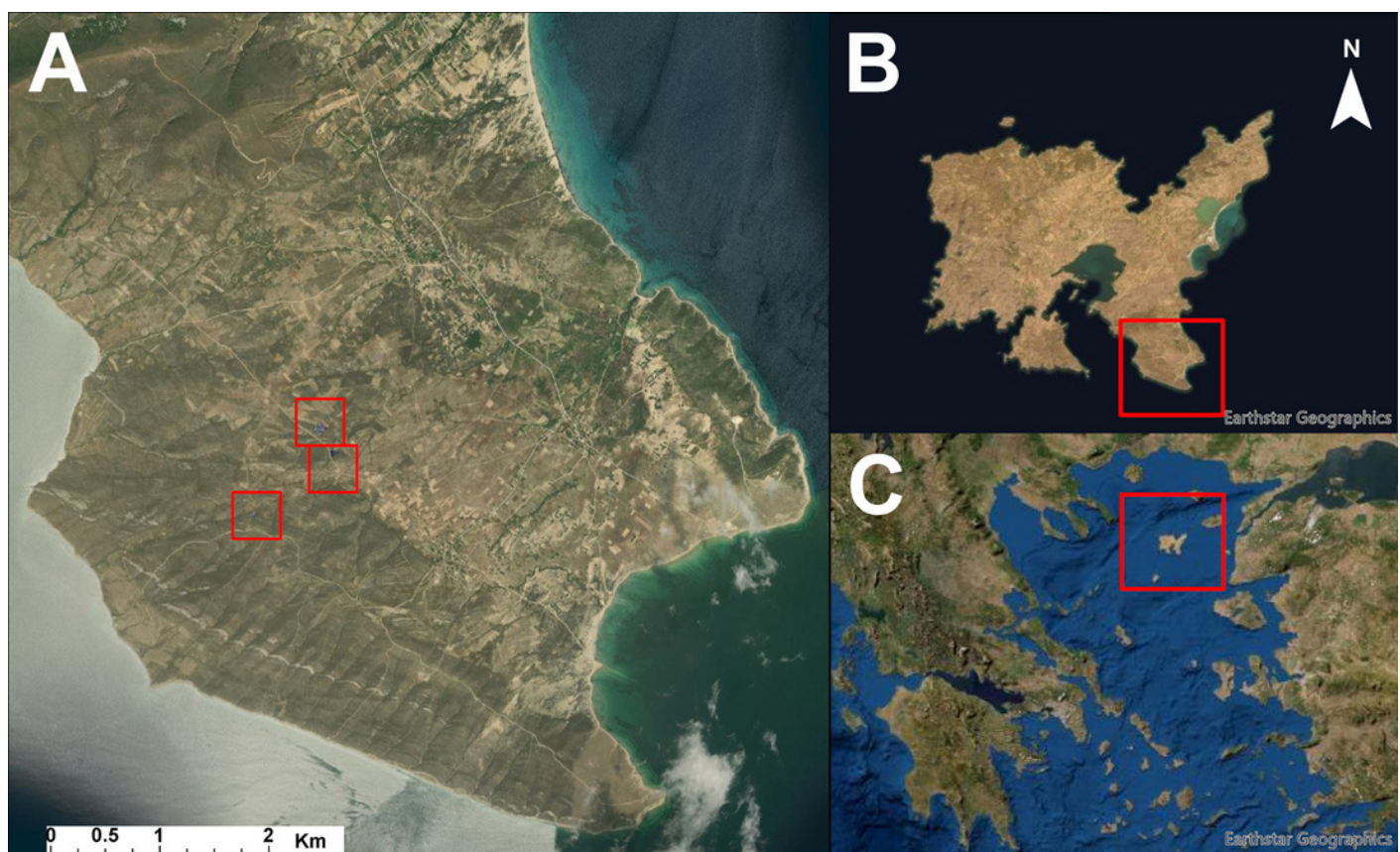
Figure 2

The three sample plots used for the detection and discrimination of plant species

Figure 2: The three sample plots used for the detection and discrimination of *Thymus capitatus* and *Sarcopoterium spinosum* species. *(A)* Study area 1. *(B)* Study area 2. *(C)* Study area 3.

A

N

AREA 1

Km
0   0.015   0.03

B

N

AREA 2

Km
0   0.01   0.02

C

N

AREA 3

Km
0   0.02   0.04
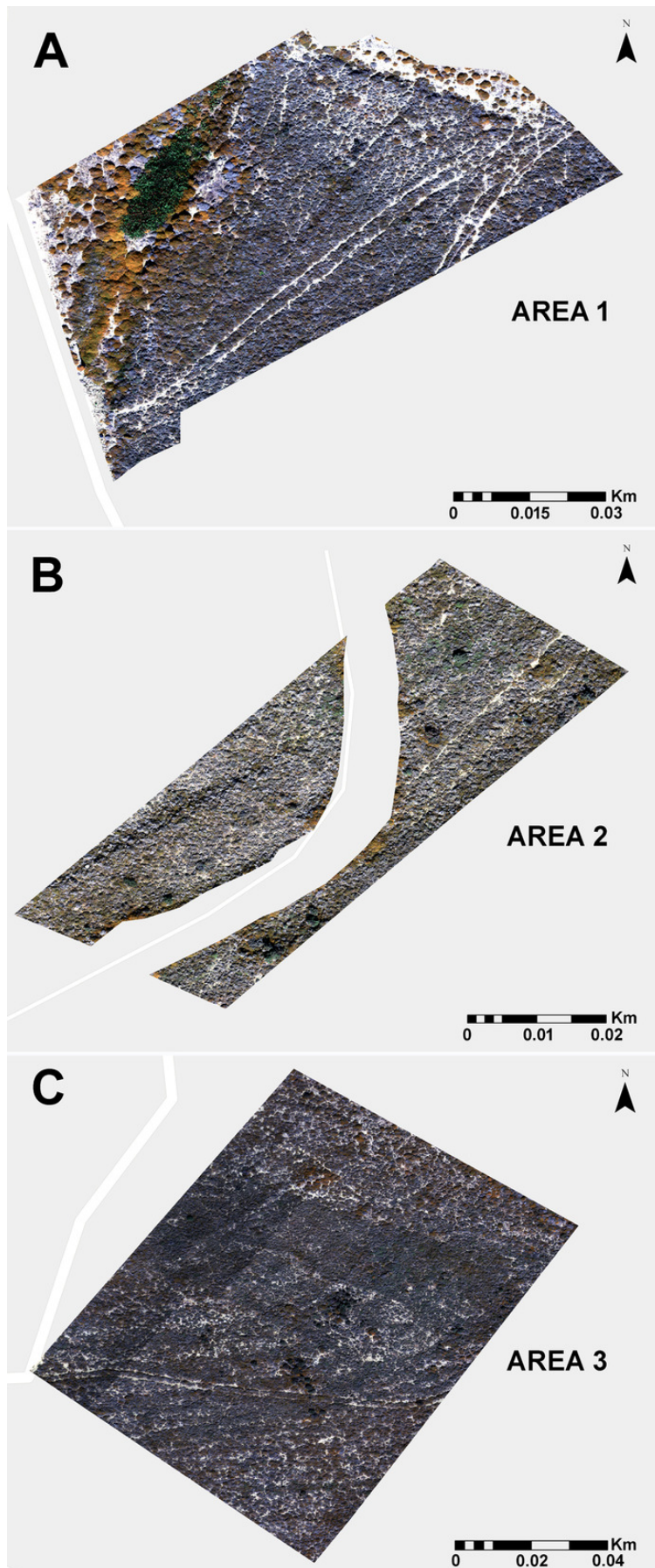
Figure 3

Classification performance of the five classifiers

Figure 3: Classification performance of the five classifiers (RFT: Random Forest, CART: Classification and Regression Trees, GTB: Gradient Tree Boost, SVM: Support Vector Machine, MMD: Mahalanobis Minimum Distance), in Google Earth Engine using the Producers Accuracy, Users Accuracy and f-score for each class separately (thyme-blue charts, Sarcopoterium – purple charts, and Other- grey charts) and the Overall accuracy and Kappa Coefficient with green charts for the three study sites (SA: Study Area). A, B, C: Stusy Area 1, 2, 3, correspondingly, VIS: Visible spectrum, NIR: Near Infrared, RE: Red Edge, VEG.INDICES: Vegiatation indices

A

SA 1 - Thyme

SA 1 - Sarcopoterium

SA 1 - Other

SA 1

B

SA 2 - Thyme

SA 2 - Sarcopoterium

SA 2 - Other

SA 2

C

SA 3 - Thyme

SA 3 - Sarcopoterium

SA 3 - Other

SA 3

Figure 4

Classification maps for the detection and distinguishing of the two species

Figure 4. Classification maps retrieved from GEE after the selection of the GTB classifier for the detection and distinguishing of the two shrub species. Evaluation of area covered by *Thymus capitatus* and *Sarcopoterium spinosum* in *(A)* study area 1, *(B)* study area 2 and *(C)* study area 3.