Forest Systems
31 (2), e009, 13 pages (2022)
eISSN: 2171-9845
https://doi.org/10.5424/fs/2022312-18787
INIA-CSIC

RESEARCH ARTICLE

OPEN ACCESS

The effect of flowering stage on distribution modelling performance: A case study of *Acacia dealbata* using maximum entropy modelling and RPA images

Antonio Vazquez de la Cueva, Fernando Montes Pita and Isabel Aulló-Maestro

Centro de Investigación Forestal (INIA, CSIC). Carretera de A Coruña km 7.5, 28040-Madrid, Spain.

Abstract

Aim of study: To classify and validate the coverage of *Acacia dealbata* by stratifying its area into three different flowering stages using remotely piloted aircraft (RPA)-derived image orthomosaics.

Area of study: We selected three sites in the west of Ourense province (Galicia, Spain). This area is the eastern cluster of *A. dealbata* populations in Galicia.

Materials and methods: We used a multirotor RPA equipped with an RGB and a multispectral camera. The flights were carried out on 10th and 11th March 2020. We performed a visual interpretation of the RGB orthomosaics to identify the patches of *A. dealbata* in three different flowering stages. We then used a maximum entropy (MaxEnt) programme to estimate the probability of *A. dealbata* presence in each study site at each of the three flowering stages.

Main results: The performance of the MaxEnt models for the three flowering stages in each of the three study sites were acceptable in terms of ROC area under the curve (AUC) analyses the values of which ranged from 0.74 to 0.91, although in most cases was greater than 0.80, this being an improvement on the classification without stratification (AUC from 0.73 to 0.86).

Research highlights: Our approach has proven to be a valid procedure to identify patterns of species distributions at local scale. In general, the performance of the models improves when stratification into flowering stages is considered. Overall accuracy of the presence prediction maps ranged from 0.76 to 0.91, highlighting the suitability of this approach for monitoring the expansion of *A. dealbata.*

Additional key words: drone; vegetation cartography; invasive species; multispectral; Verín County

Abbreviations used: AUC (area under the curve); DLS (down-welling light sensor); GNSS (global navigation satellite systems); GSD (ground sample distance); MaxEnt (maximum entropy; NDVI (normalized difference vegetation index); OA (overall accuracy); PA (producer accuracy); RMS (root mean square); ROC (receiver operating characteristic); RPA (remotely piloted aircraft); SDM (species distribution modelling); SfM (structure from motion); UA (user accuracy).

Authors' contributions: AVC designed the general plan of the paper and performed the analyses. AVC and IAM carried out the RPA flights. The three authors wrote and revised the manuscript.

Citation: Vázquez de la Cueva, A; Montes Pita, F; Aulló-Maestro, I (2022). The effect of flowering stage on distribution modelling performance: A case study of *Acacia dealbat*a using Maximum Entropy Modelling and RPA images. Forest Systems, Volume 31, Issue 2, e009. https://doi.org/10.5424/fs/2022312-18787

Received: 24 Aug 2021. **Accepted:** 30 May 2022.

Funding agencies/institutions	Project / Grant
MINECO (formerly Ministry of Economy, Industry, and Competitiveness), Spain	Project 'FORESCHANGE': Influencia del régimen de perturbaciones y la gestion en el balance de carbono, estructura y dinámica de las masas forestales (I+D+i Retos, AGL2016-76769-C2-1-R)
	FPI grant to Isabel Aulló-Maestro (BES-2017-081606)

Competing interests: The authors have declared that no competing interests exist.

Correspondence should be addressed to Antonio Vázquez de la Cueva: vazquez@inia.csic.es

Introduction

Acacia dealbata is a naturalized tree of invasive behaviour that has expanded considerably in recent years in the North West of the Iberian Peninsula (Lorenzo *et al.,* 2010; Vázquez de la Cueva, 2014; Hernández *et al.,* 2014). It was ranked as the 'worst' invasive plant within the plants life-form using a generic impact scoring by Nentwig *et al.* (2018), posing a serious threat to a large number of communities and ecosystems, depending on the severity of the impact on the community type (Lorenzo *et al.,* 2012; Hernández *et al.,* 2014). The above- and belowground impacts

caused by invasions of *Acacia* species have been described in detail over recent decades (*e.g.* Souza-Alonso *et al.,* 2017). The expansion of *A. dealbata* populations is facilitated by certain characteristics, which make this species a very successful invader (*e.g.* Richardson *et al.,* 2015; Souza-Alonso *et al.,* 2017). Fire also plays an important role in the persistence of *Acacia* populations in other ecosystems (*e.g.* Ward *et al.,* 2014 in Australia). To control a highly invasive species such as *A. dealbata*, it is essential to understand the historical progression of the invasion within the landscape (Gouws & Shackleton, 2019).

Remotely piloted aircraft (RPA) and remote sensing techniques are revolutionizing environmental monitoring. This is because RPAs are already becoming cost-efficient tools for mapping and also because photogrammetry algorithms such as Structure from Motion (SfM) are becoming more accessible (Aasen *et al.,* 2018; Díaz-Delgado & Mücher, 2019; Gómez *et al.,* 2020). RPA missions are capable of improving the efficiency of acquisition of field data and make it feasible to obtain high-resolution imagery and three-dimensional data, which can be used for forest monitoring and assessing tree attributes (Goodbody *et al.,* 2017; Mohan *et al.,* 2017). Data captured by RPA and processed by SfM techniques can provide relatively accurate and timely forest inventory information at local scale (Puliti *et al.,* 2015, 2018).

RPA-derived products have frequently been used as an alternative to field sampling for invasive species mapping as a previous step to species distribution upscaling through satellite images (Kattenborn *et al.,* 2019; Martínez-Sánchez *et al.,* 2019). Several authors have reported that data acquisition costs using RPAs are lower than those of conventional methods such as extensive ground sampling, and therefore they might provide an effective tool to support invasive plant management based on early detection and regular monitoring (Lehmann *et al.,* 2017; de Sá *et al.,* 2018; Lopatin *et al.,* 2019).

The use of multispectral (and hyperspectral) data captured from RPAs and applied to invasive plants has also been the subject of research (Müllerová *et al.,* 2017; Papp *et al.,* 2021). It provides new possibilities for assessing plant traits (Potgieter *et al.,* 2017) or for monitoring the growth of cultivated plants (*e.g.* Zhang *et al.,* 2020). In the Iberian Peninsula, Große-Stoltenberg *et al.* (2016) used hyperspectral field radiometry as a tool to discriminate *Acacia* longifolia populations on the Atlantic coast. Other authors such as de Sá *et al.* (2017) used medium resolution satellite images to study the expansion of A. longifolia, also on the Portuguese Atlantic coast. Martins *et al.* (2016) applied medium resolution satellite images to map *A. dealbata* populations in central Portugal.

A. dealbata flowering occurs in winter, normally between February and March, depending on the site conditions. Populations of *A. dealbata* exhibit yellow blooms, which are considered useful for the spectral discrimination and classification of this species. However, there is little agreement in the literature as to whether this spectral feature enables more accurate mapping of populations. As regards *A. longifolia* in Portugal, de Sá *et al.* (2017) considered that Landsat image classifications were more accurate in fall than at the time of flowering (winter). In contrast, by comparing ASTER images from March and August, Martins *et al.* (2016) concluded that classification of *A. dealbata* populations improved during the flowering season. In addition, highly accurate mapping of *A. longifolia* was achieved using RPA images captured at two phenological stages: peak and off-peak flowering (de Sá *et al.,* 2018).

Species distribution models (SDM) can be used to estimate areas prone to invasion (Vicente *et al.,* 2016) and also to estimate the probability of presence, based on samples in a certain area where species presence is known. MaxEnt is a 'one-class' classifier based on maximum entropy (Philips *et al.,* 2006), which estimates the probability of presence based on the combination of presence-only samples of the target species and a set of independent variables (Phillips *et al.,* 2017). The final output is the relative likelihood distribution (between 0 and 1) of the target species or suitability. MaxEnt has been used to model and forecast species distributions (Monterroso *et al.,* 2009; Felicisimo *et al.,* 2012) as well as to model fire occurrence and burn severity (Fonseca *et al.,* 2016; Quintano *et al.,* 2019). In recent years MaxEnt has been used in combination with remote sensing techniques, providing reliable results (Mack *et al.,* 2016; Amici *et al.,* 2017; Skowronek *et al.,* 2017; Stenzel *et al.,* 2017; Lopatin *et al.,* 2019; Fernández-Manso & Quintano, 2020).

In this study, we hypothesise that the presence of *Acacia dealbata* can be modelled from RPA data using MaxEnt and that the accuracy of the models would increase through stratification based on different phenological stages. We have used data acquired in winter (mid-March, around peak flowering) in the studied region. We have considered three flowering stages in the photointerpretation and modelling of predicted distributions with the idea of addressing plant canopy variability that hampers the analysis of high-resolution image. Hence, in this approach we decided to separately map three different flowering stages identified in the field at the time of the flight. The three flowering stages were considered as different target species in the MaxEnt models.

The specific objectives were to: (i) assess the accuracy of *A. dealbata* SDM using MaxEnt and based on RGB interpretation and multispectral information derived from RPA flights; (ii) analyse the performance of MaxEnt, either stratifying the samples into three flowering stages identifiable in winter images or considering one single class without stratification.

Material and methods

Study area

The study was conducted in Verín County in the west of Ourense province (Galicia, NW Spain), where the easternmost cluster of *A. dealbata* populations in Galicia is located. We selected three sites with notable presence of *A. dealbata* populations. The three sites, Infesta, Salgueira and Cantera have been widely affected by fires in recent years, often with short rotation periods (Vázquez de la Cueva *et al.,* 2015), and most of the *Acacia* patches originated from perturbations. The three sites mostly comprise forest vegetation (Fig. 1).

The climate is temperate with dry, mild summers. Based on the standard climate maps of the Spanish Meteorological Agency (AEMET) for the period 1981-2010, the annual precipitation in Verín County was higher than 1000 mm, corresponding to 250 mm in spring, 75 mm in summer, and 350 mm in autumn and winter. Mean annual temperature is between 11 and 14 ºC, mean annual minimum temperature is between 9 and 11 ºC and mean annual maximum temperature between 16 and 19 ºC.

Verín County is located in a transition area between Oceanic and Mediterranean vegetation types. We have used the Spanish Forest Map (MFE25) to describe the forest vegetation of the area, dominated by conifers, deciduous species and *Acacia*. This map is based on visual photo-interpretation of aerial photographs from the yêars 2006-2007 and provides the basis for Fig. 1.

RPA data collection

We used a DJI Matrice 210 multirotor drone equipped with an RGB sensor (Zenmuse X4S) and a multispectral sensor RedEdge-M (MicaSense), a Down-welling Light Sensor (DLS) and an additional GPS module.

The flights were conducted on the 10th and 11th of March 2020, as close as possible to the local solar noon (Table 1). The area captured by the Zenmuse X4S RGB camera was larger than that captured by the RedEdge-M and the spatial resolution was also greater (5472 × 3078 *vs* 1280 × 960 pixels). The total area captured by the RGB and the RedEdge sensors for the three sites were 45.4 and 37.9 ha respectively for Site 1, 62.8 and 46.6 ha for Site 2, and 48.9 and 38.6 ha for Site 3 (Table 1). For both cameras we selected side and front overlap of 75% in accordance with the indications in the MicaSense bibliography. The altitude (from the home point) was set at 120 m (the maximum according to the Spanish regulations). At an altitude of 120 m, the theoretical ground sample distance (GSD) is 3.29 cm/pixel for the RGB images and approximately 8.3 cm/pixel for the multispectral images, although the

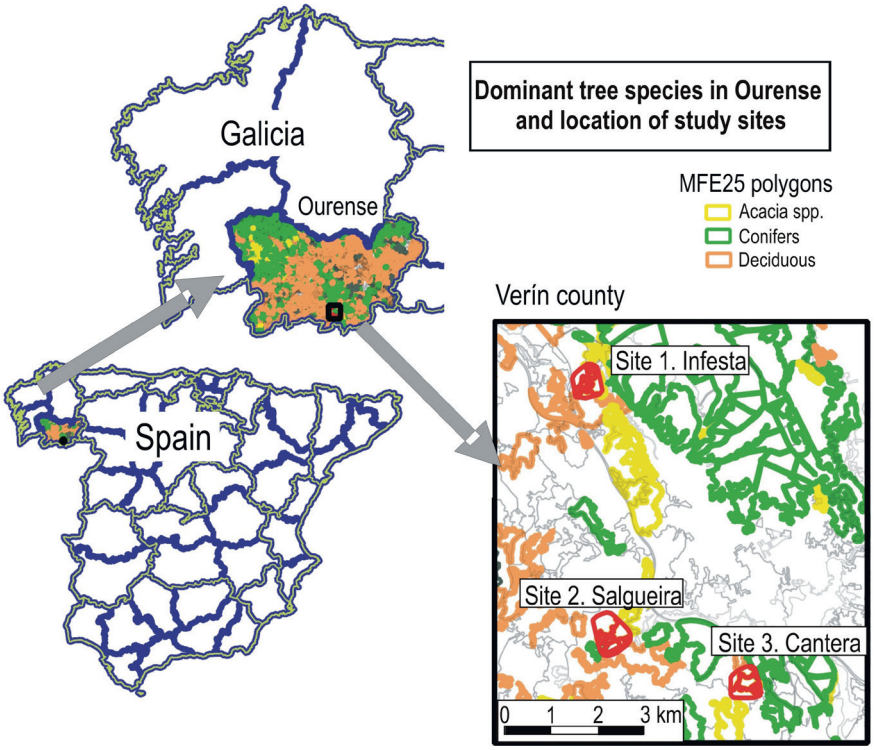

Figure 1. Location of the three study sites in Verín County, province of Ourense within the autonomous community of Galicia, Spain. The three study sites are displayed in red. On the three maps the colored polygons represent the dominant tree species according to the Forest Map of Spain at a scale of 1:25.000. Yellow indicates *Acacia* dominated polygons, green are conifers and orange is used for deciduous trees.

Table 1. Characteristics of the images acquired in March 2020 at the three study sites. The time of image capture was close to the local solar noon. The total area covered by each flight and sensor ranged from near 38 ha to more than 60 ha. The table also provides information of the number of individual images, the mean theoretical ground sample distance (GSD), the elevation ranges at each site, the number of 3D Ground Control Points (GCP) finally used and the root mean square (RMS) obtained in the six orthomosaics. The output pixel size was set to 5 cm for RGB and 10 cm for multispectral orthomosaics.

Date	Time (start UTC)	Sensor	Total size (ha)	Images captured		Elevation range		Georeferencing	
				#	Average GSD (cm)	min	max	3D GCP	RMS error (m)
Site 1. Infesta									
10 Mar	13:55	RGB	45.4	201	4.20	481	655	8	0.073
		RedEdge	37.9	1440	10.84			8	0.185
Site 2. Salgueira									
11 Mar	12:58	RGB	62.8	182	7.19	422	697	6	0.090
		RedEdge	46.6	1290	16.95			4	0.312
Site 3. Cantera									
11 Mar	14:53	RGB	48.9	197	4.91	459	628	6	0.123
		RedEdge	38.6	1385	12.57			7	0.246

elevation gradient led to a larger average GSD. The speed was set to 5 m/s and the flight duration was around 20 minutes at each study site (Table 1). Immediately before and after each flight, we took an image of the MicaSense calibrated reflectance panel to create reflectance-compensated outputs. During the three flights the weather was clear and sunny with no clouds and very low wind speed.

Prior to the RPA flight we deployed 6 to 8 ground markers at each of the three field sites, which were later identified in individual images. The centre of the mark was measured with an RTK (Real-Time Kinetics) GNSS (Global Navigation Satellite System) Reach RS2 GNSS Receiver (manufactured by Emlid Ltd). This GPS was mounted on a pole of known height and was connected via the internet using the protocol NTRIP to the nearest IGN (Instituto Geográfico Nacional) station (Ponferrada, at approximately 100 km) to give real-time position correction. The GPS was also connected to a mobile device through the Reach application for management and visualization. These points were used to improve orthomosaics accuracy.

RGB and multispectral data processing

We processed the RPA images using Pix4D software to obtain the orthomosaics of the RGB and the reflectance images in the five bands captured by the MicaSense-M camera. The size of the mosaics varied among the three sites and the sensors used. The RMS (root mean square) errors ranged from 0.073 to 0.123 m in RGB and from 0.185 to 0.246 m in multispectral data sets (Table 1). The output pixel size of the orthomosaics was set to 5 cm in the RGB orthomosaics and 10 cm in multispectral images.

During multispectral image processing the objective was to obtain reflectance values as a percentage of incoming to reflected light for each of the spectral windows considered. In MicaSense, the five bands are centred at 480 nm (blue), 560 nm (green), 670 nm (red), 720 nm (red edge) and 840 nm (near-infrared). The full width at half maximum is 20 nm (blue and green), 10 nm (red and red-edge), and 40 nm (near-infrared). Several of the pre-processing steps such as correcting images for dark pixels, vignette effects, exposure and gain are performed by Pix4D software. Reflectance values are based on EXIF information of the images, the DLS values and the calibration of the images using the reflectance panel albedo factors. These are obtained for each panel from the MicaSense web page (www.micasense.com). We also calculated the NDVI (normalized difference vegetation index) from red and near-infrared bands.

Identification and digitalization of *Acacia dealbata* patches

We performed a visual interpretation of the RGB orthomosaics to delimit the patches of *A. dealbata*. The interpretation was performed at a scale of 1:375 on-screen. Pictures taken the same day as the flight along with knowledge of the vegetation at the three study sites helped supporting visual interpretation. The minimum size of the digitized polygons was established at 5 m between borders and the distance between contiguous patches was also 5 m. In these images it was possible to differentiate different flowering stages of the trees based on the number of visible flowers (yellow) at the moment of the flight.

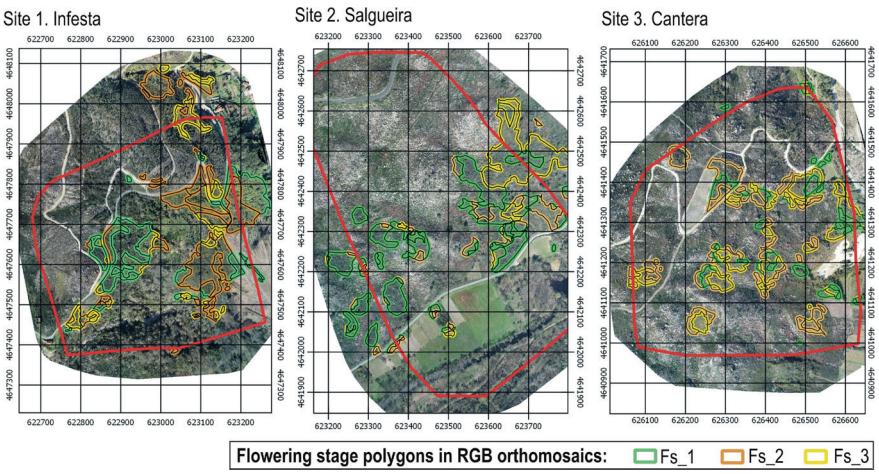

Figure 2. RGB orthomosaics derived from the early-March 2020 RPA flights at the three study sites. UTM coordinates grid spacing is 100 m. The outer red polygon delimits the common area for the RGB and the multispectral orthomosaics for which the MaxEnt analyses were performed. Flowering stage polygons are based on the visual interpretation performed on the RGB orthomosaics. The codes for the flowering stage (Fs_) patches ranged from 1 (patches with scattered flowers) to 3 (patches dominated by flowers).

These three types are: 1) *A. dealbata patches* with few or no flowers, 2) *A. dealbata* patches with sparse spots of flowers and 3) *A. dealbata* patches with a dense flower cover. The orthomosaics generated from the individual RGB pictures captured in the RPA flights are displayed in Fig. 2. Overlaying these images are the patches of *A. dealbata* classified into the three flowering stages (Fs_1, Fs_2 and Fs_3). The common area used finally in the MaxEnt analyses (red polygons in Fig. 2) is 25.0 ha for Site 1, 38.5 ha for Site 2 and 31.5 ha for Site 3.

Maximum entropy (MaxEnt) modelling procedures and validation

The MaxEnt programme estimates the probability of presence, based on the combination of presence-only samples of the target class (derived from visual photointerpretation in our approach) and a set of auxiliary variables (Phillips *et al.,* 2006, 2017). In our case, these auxiliary variables are the five reflectance bands (blue, green, red, red-edge and near-infrared) and the NDVI.

The training and testing data set for the MaxEnt algorithm were 200 points randomly selected from a 5 × 5 m point grid for each of the three flowering stages of *A. dealbata,* derived from the interpretation performed in the three RGB orthomosaics. MaxEnt creates a map of probability of the target class, estimated between 0 and 1. From the 200 random samples we used 75% of them to train MaxEnt, and 25% for testing purposes, running 500 iterations. The rest of the user-specified parameters were set to their default values (Phillips *et al.,* 2006, 2017). To estimate the importance of auxiliary variables in the model, we used a jackknife test: one variable was excluded at a time and MaxEnt was run with the remaining variables. We then ran the algorithm for each variable, one at a time. Finally, we ran the algorithm using all variables, as before. In this way, it is possible to evaluate the importance of each variable in the model. Another useful output provided by the MaxEnt software is the receiver operating characteristic (ROC) analysis, in which area under the curve (AUC) represents the model capability of adequately predicting presence (sensitivity) and absence (specificity).

The continuous MaxEnt output was converted into binary maps to build confusion matrices between the predicted classification and the photointerpreted classes using the points of the 5×5 meter grid overlaying each of the tree sites. We used a 0.6 probability threshold in accordance with the MaxEnt instructions and outputs: for the three sites this value is close to the thresholds proposed under the equal training sensitivity and specificity criteria (Phillips *et al.,* 2006, 2017). Finally, we calculated the Kappa index, the user accuracy (UA), the producer accuracy (PA), and the overall accuracy (OA) between the classes predicted by MaxEnt and the visual photointerpretation of the flowering stages.

Results

MaxEnt models performance

The AUC values for training samples ranged from 0.81 to 0.88 and from 0.74 to 0.91 in the test samples (Table 2). The ROC curves corresponding to these models are displayed in Fig. 3 for the training and test dataset. In all but two of the nine models, the AUC values were greater than 0.80, which indicates a good model performance. The

Table 2. Results for the twelve MaxEnt models run at the three study sites and for the three flowering stages defined. The table includes the gain (regularized for the Training data) and for Test data, the AUC for both datasets and the standard deviation (SD) of the models. Models are based on 200 random points: 75% for Training and 25 for Test. The models were obtained after 500 iterations for each of the twelve runs. Nine of the models are stratified by the flowering stage (*Fs_1, Fs-2* and *Fs_3*) while in three of them (All the Fs) the stratification was not considered.

Flowering stages	Training gain (Reg.)	Training AUC	Test gain	Test AUC	AUC SD
Site 1. Infesta					
Fs_1	0.48	0.81	0.47	0.74	0.044
Fs_2	0.47	0.82	0.66	0.82	0.026
Fs_3	0.75	0.87	0.68	0.82	0.032
All the Fs	0.43	0.80	0.45	0.73	0.033
Site 2. Salgueira					
Fs_1	0.85	0.86	1.28	0.91	0.014
Fs_2	0.79	0.86	0.87	0.85	0.019
Fs_3	0.84	0.88	1.02	0.88	0.018
All the Fs	0.79	0.86	0.93	0.86	0.024
Site 3. Cantera					
Fs_1	0.61	0.83	0.73	0.82	0.029
Fs_2	0.46	0.82	0.57	0.81	0.024
Fs_3	0.64	0.85	0.60	0.80	0.027
All the Fs	0.49	0.84	0.56	0.80	0.027

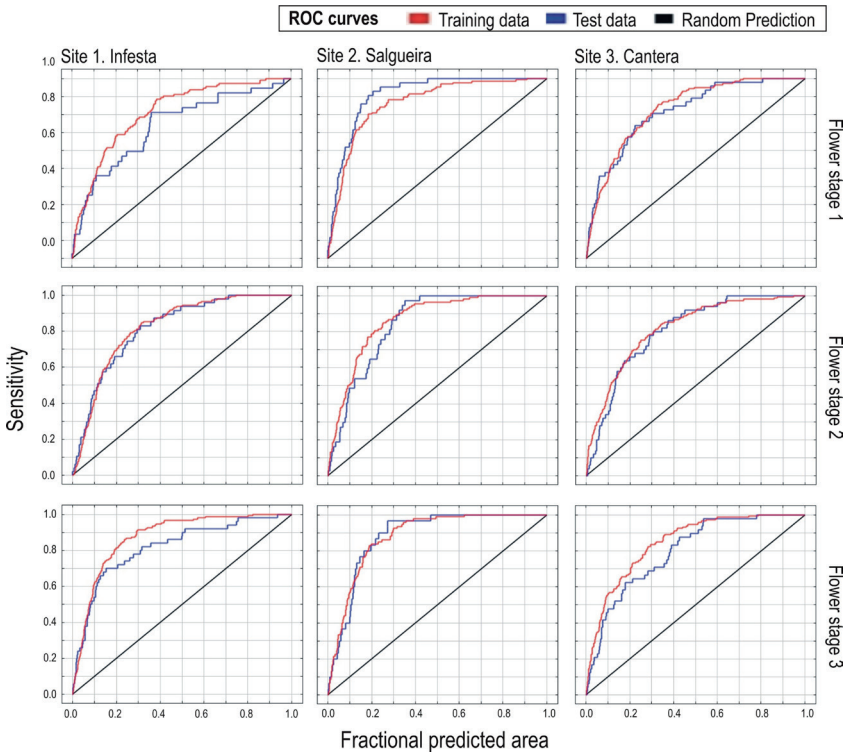

Figure 3. Receiver operating characteristic (ROC) curve for each of the three study sites and the three *Acacia dealbata* flowering stages considered as target in the MaxEnt models. Horizontal axis is Fractional predicted area (1 – Specificity) while vertical axis is Sensitivity (1 – Omission Rate) (Phillips *et al.,* 2016). This implies that the maximum achievable AUC is less than 1. The straight black line is the random prediction (AUC= 0.5) while the red curve is the AUC for the training data (75%) and the blue line is the AUC for the test data (25%) of the 200 random points used in each of the nine MaxEnt model runs. AUC values are shown in Table 2.

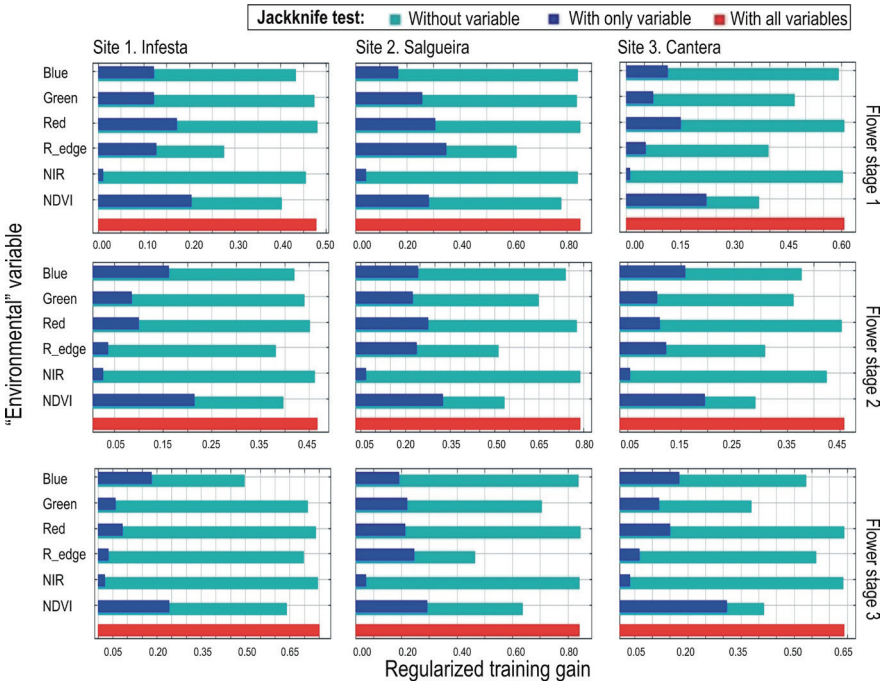

Figure 4. Results of the jackknife tests on the importance of the auxiliary (or "environmental") variables for the regularized training gain in MaxEnt models. Each plot represents the results for the three study sites and the three *A. dealbata* flowering stages considered as target class in each model. Variables used are reflectance bands and NDVI.

performance of the models when stratifying the three flowering stages was a little better than when using all the flowering stages together. At two of the three sites the median values of AUC for the three flowering stages were higher than when the models were run without distinguishing between stages (Table 2).

The importance of the auxiliary variables in each of the nine MaxEnt models according to the jackknife test is shown in Fig. 4. In eight of the nine models, the auxiliary variable with the highest gain was the NDVI. In three of the nine cases, the variable with the second highest gain was the red and in four cases the blue band. The variable with the lowest gain in all cases was the near-infrared reflectance, perhaps because the NDVI already accounted for its contribution.

The auxiliary variable that most reduces the gain when it is omitted is red-edge reflectance, which therefore appears to have the most information not present in the

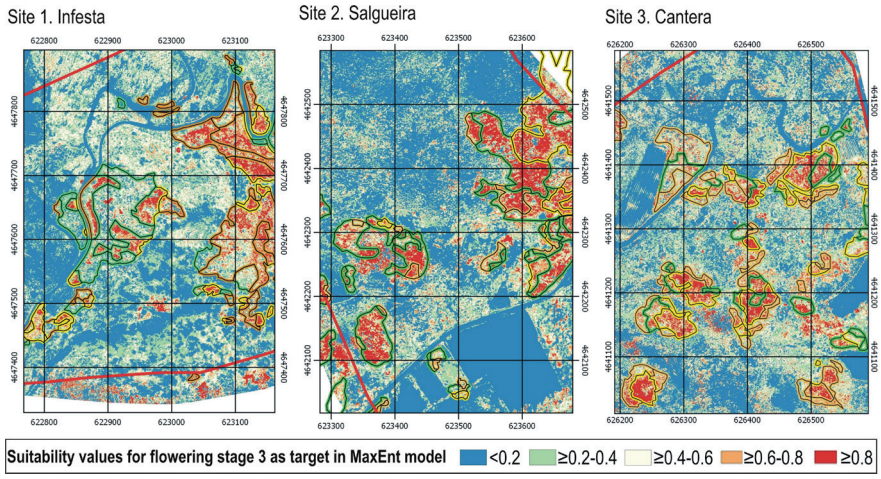

Figure 5. Classification maps derived from the continuous suitability values output of the MaxEnt models for parts of the three study sites and for flowering stage 3 as target "species". Red color indicates ≥ 0.8 and orange ≥ 0.6 and < 0.8. The threshold values used in the accuracy assessment is 0.6 (Table 4). Most of the pixels with orange or red values are within the digitalized polygons in accordance with the high overall accuracy obtained in the models.

Table 3. Mean values of the suitability output obtained for the final MaxEnt models within the digitalized polygons classified into the three flowering stages of *A. dealbata* as well as outside these polygons. For each of the three sites we show the values obtained considering each of the flowering stages as target for the model. The random points used in the analyses and in the confusion matrices are derived from a 5 × 5 m grid overlaid on the common area covered (see Fig. 2) by RGB and multispectral orthomosaics. Based on these points the percentage of each flowering stage at each site is shown along with the total area covered by *A. dealbata.*

Flowering stages (digit. polygons)	5×5 node points		MaxEnt target (mean)		
	(#)	(%)	*Fs_1*	*Fs_2*	*Fs_3*
Site 1. Infesta					
Fs_1	685	30%	0.67	0.55	0.36
Fs_2	967	43%	0.58	0.69	0.55
Fs_3	614	27%	0.53	0.75	0.69
Total area Acacia	2,266	23%	0.60	0.66	0.54
Area No Acacia	7,750		0.35	0.32	0.22
Site 2. Salgueira					
Fs_1	1,125	59%	0.68	0.67	0.68
Fs_2	350	18%	0.54	0.68	0.67
Fs_3	430	23%	0.54	0.65	0.67
Total area Acacia	1,905	12%	0.61	0.68	0.67
Area No Acacia	13,512		0.20	0.22	0.22
Site 3. Cantera					
Fs_1	433	23%	0.70	0.59	0.59
Fs_2	718	37%	0.51	0.66	0.56
Fs_3	769	40%	0.58	0.64	0.68
Total area Acacia	1,920	15%	0.60	0.63	0.58
Area No Acacia	10,684		0.29	0.36	0.27

other variables. This is true for the six MaxEnt models at Site 1 and 2. At Site 3, however, the NDVI and green reflectance variables were those that most reduced the gain when omitted, the former in two cases and the latter in one case. The regularized training gain values for all the variables are also shown in Table 2.

Flowering stage probability maps

Table 3 shows the total number of 5-m grid points at each site and the percentage of *A. dealbata* presence. These percentages are based on the visual photointerpretation (polygons in Figs. 2 and 5). At each of the sites, the most frequent flowering stage was one of the three considered: the Fs_2 at Site 1, the Fs_1 at Site 2 and the Fs_3 at 3 3. The mean probability values for the grid points located in the areas where the flowering stage coincides with the modelled class ranged from 0.66 to 0.70. In contrast, the mean probability values beyond the areas covered by *A. dealbata* (according to visual interpretation) ranged from 0.22 to 0.36.ning samples ranged from 0.81 to 0.88 and from 0.74 to 0.

Classification into binary maps and accuracy assessment

Table 4 shows the accuracy measures (PA, UA, OA and the Kappa index of agreement) derived from the confusion matrices. The OA values ranged from 0.76 to 0.91 with the 0.6 threshold value used. In general, for Sites 1 and 2, the OA percentage increased from the flowering stage 1 to 3 but not in the case of Site 2 (Salgueira). The Kappa index values ranged from 0.35 to 0.57 and for two of the sites were greater in flowering stage 3. PA, which represents the quality of the classification of field points (interpreted), and UA, which is the probability that the prediction represents reality, are shown for the different models in Table 4. PA and UA are greater for the absence class than for the presence class for the three flowering

Table 4. Confusion matrice parameters (producer, user and global accuracy and the Kappa index of agreement) for the 0.6 threshold of the suitability output values of MaxEnt models for the three study sites and the three flowering stages. The classification based MaxEnt suitability output values are the predicted values, while digitalized polygons are considered the ground truth datasets.

Flowering stages	Classified data[1]	Producer accuracy	User accuracy	Overall accuracy	Kappa index
Site 1. Infesta					
Fs_1	0	0.83	0.86		
	1	0.54	0.48		
	Total	0.68	0.67	0.76	0.35
Fs_2	0	0.91	0.90		
	1	0.66	0.68		
	Total	0.78	0.79	0.85	0.57
Fs_3	0	0.97	0.85		
	1	0.41	0.81		
	Total	0.69	0.83	0.85	0.46
Site 2. Salgueira					
Fs_1	0	0.96	0.94		
	1	0.55	0.64		
	Total	0.75	0.79	0.91	0.54
Fs_2	0	0.91	0.95		
	1	0.64	0.51		
	Total	0.78	0.73	0.88v	0.50
Fs_3	0	0.93	0.95		
	1	0.67	0.56		
	Total	0.80	0.76	0.89	0.55
Site 3. Cantera					
Fs_1	0	0.88	0.90	0.81	0.33
	1	0.47	0.41		
	Total	0.67	0.65		
Fs_2	0	0.81	0.93	0.78	0.35
	1	0.64	0.37		
	Total	0.73	0.65		
Fs_3	0	0.90	0.92	0.85	0.43
	1	0.55	0.50		
	Total	0.72	0.71		

[1] 0 (absence); 1 (presence)

stages and the three sites. In general, the total values of UA and PA are quite similar.

Discussion

Greater variability within plant canopies is a consequence of high-resolution images such as those captured by RPAs, in our case the resolution being 5 and 10 cm. At this spatial resolution, the canopy of *A. dealbata* is not spectrally homogeneous. At the time of the flight, at the peak of flowering, there were large differences due to the contrast between the yellow of the flowers and the green of the leaves. In this study we stratified the canopy of *A. dealbata* into three flowering stages in order to deal with this variability, the results obtained being generally better than without stratification. Thus, we reduced the variability in the spectral response of *A. dealbata* patches according to the flowering stage. In this sense, our approach of considering several flowering stages to train and validate classification algorithms could be of interest for future prospection and monitoring of *A. dealbata*.

The effect of shadows on the variability within a plant canopy has been discussed by several authors with regard

to very high (centimetres) pixel resolutions. Shadows are an important factor in these classifications (Adeline *et al.,* 2013; Lopatin *et al.,* 2019). It is also necessary to bear in mind the fact that images are captured with low sun angles (early-March). Although the three flights were performed close to noon in order to minimize shadows, these are apparent in the orthomosaics. The effect of shadows is another side effect of the high spatial resolution. In the images captured by RPA the shadows can affect data analysis due to the high reflectance variability of individual plants within the canopy (Mamaghani *et al.,* 2019). In this approach, no pre-processing was carried out to minimize these effects, which could affect predictions. Exposed and shadow pixels were included in the training data set, so variability due to shadowing was incorporated in the model uncertainty.

RPAs are contributing to increase spatial resolution in SDM, as well as becoming the providers of an intermediate scale between ground measurements and satellite/airborne image data. The results obtained with RPA are promising and could lead to the development of more cost-effective surveys (Puliti *et al.,* 2018). Decreasing costs through RPA-based surveying and open access to high-resolution satellite images open new avenues to high-resolution cartography of invasive plants. Invasive alien species are considered a risk for biodiversity, hence the need to inventory them for further control and eradication. The methodology used in this paper could contribute to addressing this need, allowing areas covered by the target species to be identified, even at different phenological stages.

The MaxEnt SDM approach has several advantages over other classifiers (Muñoz *et al.,* 2016; Fernández-Manso & Quintano, 2020): 1) MaxEnt is a non-parametric model; 2) from an operational point of view it is an attractive substitute for machine-learning-based classifiers, as presence samples alone are needed to train it; the delineation of training data of unwanted classes (i.e. non-target cover) is not required during modelling, which notably decreases pre-processing or sampling efforts (Lopatin *et al.,* 2019); and 3) its output, a continuous probability surface, has physical meaning, making it easy to interpret (Felicisimo *et al.,* 2012). In any case, the performance of the models obtained is comparable to those reported by other authors using binary classifications to map the occurrence of invasive species (Lehmann *et al.,* 2017; Piiroinen *et al.,* 2019). The AUC values obtained through the twelve MaxEnt models were in most cases higher than 0.80 and close to 0.90 in some models, which indicated excellent model performance (Muñoz *et al.,* 2016; Fernández-Manso & Quintano, 2020). AUC values between the three flowering stages were similar in each site.

Overall accuracy of the presence prediction maps obtained is quite acceptable, with values generally ranging between 0.80 and 0.90 for the three flowering stages and the three sites. Lehmann *et al.* (2017), using RPA in an approach to map invasion by *Acacia* mangium in a Brazilian Savanna ecosystem, reported an overall accuracy of 82.7% from the analysis of imagery. Several improvements could be made to obtain more accurate parameter values. One possible improvement, in addition to reducing the shadows effect, could be the use of filters to avoid the salt and pepper effect of the classifications (*e.g*. Fernández-Manso & Quintano, 2020; Gómez *et al.,* 2020). Other authors (*e.g*. Kattenborn *et al.,* 2019) have used MaxEnt with additional spectral, textural or canopy 3D structural predictors. In our approach, we have used the five spectral bands and only one spectral index (NDVI).

Forecasting invasions, modelling and managing ecosystems dominated by *A. dealbata* are challenging tasks that must be addressed given that climatic conditions, intensification of land uses and increased risk of perturbations such as fires, are increasing the likelihood of *Acacia* invasions into new areas (Hernández *et al.,* 2014, Souza-Alonso *et al.,* 2017). Future research should focus on the early detection and prevention of new *Acacia* invasions and on cost-effective and sustainable management of the novel ecosystems resulting from invasions (Marchante *et al.,* 2011). Problems associated with the invasiveness of non-native tree species are increasing rapidly worldwide, especially in areas with a long history of plantations (Richardson *et al.,* 2015). We cannot forget that the origin of *A. dealbata*, and also of other *Acacia* species that are currently invading diverse ecosystems, was their intentional plantation. In areas with a long presence of *A. dealbata* populations such as South Africa (Van Wilgen *et al.,* 2011), the specialists consider that even with increased budget spending, control would probably not be achieved under the forecasted less favourable climatic and social scenarios. For these reasons, a great deal of money has been and will continue to be wasted in the future on control of alien species invasions (Van Wilgen *et al.,* 2016).

In conclusion, plant canopy variability hampers the analysis of high-resolution image such as those acquired using RPA. By considering three flowering stages in the photointerpretation and modelling of predicted distributions the idea of this work was to address this variability, our results being better than those obtained without this stratification. The performance of the MaxEnt models for the three study sites and the three flowering stages was suitable for *A. dealbata* distribution assessment. This method has several advantages over other classifiers such as its non-parametric character, the need for presence-only samples and its continuous output. RPA high spatial resolution products may contribute to forest monitoring and management. In this regard, RPAs have become the providers of an intermediate scale between ground measurements and satellite/airborne image data, which could

lead to more cost-effective surveys of the distribution and structure of invasive alien species populations.

References

Aasen H, Honkavaara E, Lucieer A, Zarco-Tejada PJ, 2018. Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: A review of sensor technology, measurement procedures, and data correction workflows. Remote Sens 10: 1091. https://doi.org/10.3390/rs10071091

Adeline KRM, Chen M, Briottet X, Pang SK, Paparoditis N, 2013. Shadow detection in very high spatial resolution aerial images: A comparative study. ISPRS J Photogramm Remote Sens 80: 21-38. https://doi.org/10.1016/j.isprsjprs.2013.02.003

Amici V, Marcantonio M, La Porta N, Rocchini D, 2017. A multi-temporal approach in MaxEnt modelling: A new frontier for land use/land cover change detection. Ecol Inform 40: 40-49. https://doi.org/10.1016/j.ecoinf.2017.04.005

de Sá NC, Carvalho S, Castro P, Marchante E, Marchante H, 2017. Using Landsat Time Series to understand how management and disturbances influence the expansion of an invasive tree. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 10 (7): 3243-3253. https://doi.org/10.1109/JSTARS.2017.2673761

de Sá NC, Castro P, Carvalho S, Marchante E, López-Núñez FA, Marchante H, 2018. Mapping the flowering of an invasive plant using unmanned aerial vehicles: Is there potential for biocontrol monitoring? Front Plant Sci 9: 293. https://doi.org/10.3389/fpls.2018.00293

Díaz-Delgado R, Mücher S, 2019. Editorial of Special Issue "Drones for Biodiversity Conservation and Ecological Monitoring". Drones 3: 47. https://doi.org/10.3390/drones3020047

Felicísimo AM, Muñoz J, Mateo RG, Villalba CJ, 2012. Vulnerabilidad de la flora y vegetación españolas ante el cambio climático. Ecosistemas 21(3): 1-6.

Fernández-Manso A, Quintano C, 2020. A synergetic approach to burned area mapping using maximum entropy modeling trained with hyperspectral data and VIIRS hotspots. Remote Sens 12: 858. https://doi.org/10.3390/rs12050858

Fonseca MG, Aragao LEOC, Lima A, Shimabukuro YE, Arai E, Anderson LO, 2016. Modelling fire probability in the Brazilian Amazon using the maximum entropy method. Int J Wildland Fire 25: 955-969. https://doi.org/10.1071/WF15216

Gómez C, Goodbody TRH, Coops NC, Alvarez-Taboada F, Sanz-Ablanedo E, 2020. Forest ecosystem monitoring using Unmaned Aerial Systems. In: Unmanned aerial remote sensing. UAS for environmental applications; Green DR, Gregory BJ, Karachok A, (eds), pp: 173-196. CRC Press. ISBN 9781482246070. https://doi.org/10.1201/9780429172410-11

Goodbody TRH, Coops NC, Marshall PL, Tompalski P, Crawford P, 2017, Unmanned aerial systems for precision forest inventory purposes: A review and case study. The Forestry Chronicle 93(1): 71-81. https://doi.org/10.5558/tfc2017-012

Gouws AJ, Shackleton CM, 2019. A spatio-temporal, landscape perspective on *Acacia dealbata* invasions and broader land use and cover changes in the northern Eastern Cape, South Africa. Environ Monit Assess 191: 74. https://doi.org/10.1007/s10661-019-7204-y

Große-Stoltenberg A, Hellmann C, Werner C, Oldeland J, Thiele J, 2016. Evaluation of continuous VNIR-SWIR spectra versus narrowband hyperspectral indices to discriminate the invasive *Acacia* longifolia within a mediterranean dune ecosystem. Remote Sens 8: 334. https://doi.org/10.3390/rs8040334

Hernández L, Martínez-Fernández J, Cañellas I, Vázquez de la Cueva A, 2014. Assessing spatio-temporal rates, patterns and determinants of biological invasions in forest ecosystems. The case of *Acacia* species in NW Spain. Forest Ecol Manage 329: 206-213. https://doi.org/10.1016/j.foreco.2014.05.058

Kattenborn T, Lopatin J, Förster M, Braun AC, Fassnacht FE, 2019. UAV data as alternative to field sampling to map woody invasive species based on combined Sentinel-1 and Sentinel-2 data. Remote Sens Environ 227: 61-73. https://doi.org/10.1016/j.rse.2019.03.025

Lehmann JRK, Prinz T, Ziller SR, Thiele J, Heringer G, Meira-Neto JAA, Buttschardt TK, 2017. Open-source processing and analysis of aerial imagery acquired with a low-cost unmanned aerial system to support invasive plant management. Front Environ Sci 5: 44. https://doi.org/10.3389/fenvs.2017.00044

Lopatin J, Dolos K, Kattenborn T, Fassnacht FE, 2019. How canopy shadow affects invasive plant species classification in high spatial resolution remote sensing. Remote Sens Ecol Conserv 5(4): 302-317. https://doi.org/10.1002/rse2.109

Lorenzo P, González L, Reigosa MJ, 2010. The genus *Acacia* as invader: the characteristic case of *Acacia dealbata* Link in Europe. Ann For Sci 67 (1): 101. https://doi.org/10.1051/forest/2009082

Lorenzo P, Pazos-Malvido E, Rubido-Bará M, Reigosa MJ, González L, 2012. Invasion by the leguminous tree *Acacia dealbata* (Mimosaceae) reduces the native understorey plant species in different communities. Aust J Bot 60(8): 669-675. https://doi.org/10.1071/BT12036

Mack B, Roscher R, Stenzel S, Feilhauer H, Schmidtlein S, Waske B, 2016. Mapping raised bogs with an

iterative one-class classification approach. ISPRS J Photogramm Remote Sens 120: 53-64. https://doi.org/10.1016/j.isprsjprs.2016.07.008

Mamaghani B, Saunders MG, Salvagio C, 2019. Inherent reflectance variability of vegetation. Agriculture 9: 246. https://doi.org/10.3390/agriculture9110246

Marchante H, Freitas H, Hoffmann JH, 2011. Post-clearing recovery of coastal dunes invaded by *Acacia* longifolia: is duration of invasion relevant for management success? J Appl Ecol 48: 1295-1304. https://doi.org/10.1111/j.1365-2664.2011.02020.x

Martínez-Sánchez J, González-de-Santos LM, Novo A, González-Jorge H, 2019. UAV and satellite imagery applied to alien species mapping in NW Spain. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W13, 2019 ISPRS Geospatial Week 2019, 10-14 June 2019, Enschede, The Netherlands. https://doi.org/10.5194/isprs-archives-XLII-2-W13-455-2019

Martins F, Alegria C, Gil A, 2016. Mapping invasive alien *Acacia dealbata* Link using ASTER multispectral imagery: a case study in central-eastern of Portugal. Forest Syst 25(3): e078. https://doi.org/10.5424/fs/2016253-09248

Mohan M, Silva CS, Klauberg C, Jat P, Catts G, Cardil A *et al.,* 2017. individual tree detection from unmanned aerial vehicle (UAV) derived canopy height model in an open canopy mixed conifer forest. Forests 8: 340. https://doi.org/10.3390/f8090340

Monterroso, P, Brito JC, Ferreras P, Alves PC, 2009. Spatial ecology of the European wildcat in a Mediterranean ecosystem: Dealing with small radio-tracking datasets in species conservation. J Zool 279: 27-35. https://doi.org/10.1111/j.1469-7998.2009.00585.x

Müllerová J, Bartaloš T, Brůna J, Dvořák P, Vítková M, 2017. Unmanned aircraft in nature conservation: an example from plant invasions. Int J Rem Sens 38(8-10): 2177-2198. https://doi.org/10.1080/01431161.2016.1275059

Muñoz A, Santos X, Felicisimo AM, 2016. Local-scale models reveal ecological niche variability in amphibian and reptile communities from two contrasting biogeographic regions. PeerJ 4: e2405. https://doi.org/10.7717/peerj.2405

Nentwig W, Bacher S, Kumschick S, Pysek P, Vila M, 2018. More than "100 worst" alien species in Europe. Biol Invasions 20: 1611-1621. https://doi.org/10.1007/s10530-017-1651-6

Papp L, van Leeuwen B, Szilassi P, Tobak Z, Szatmári J, Árvai M *et al.,* 2021. Monitoring invasive plant species using hyperspectral remote sensing data. Land 10: 29. https://doi.org/10.3390/land10010029

Phillips SJ, Anderson RP, Schapire RE, 2006 Maximum entropy modeling of species geographic distributions. Ecol Model 190(3-4): 231-259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME, 2017. Opening the black box: an open-source release of Maxent. Ecography 40: 887-893. https://doi.org/10.1111/ecog.03049

Piiroinen R, Fassnacht FE, Heiskanen J, Maeda E, Mack B, Pellikka P, 2018. Invasive tree species detection in the Eastern Arc Mountains biodiversity hotspot using one class classification. Remote Sens Environ 218: 119-131. https://doi.org/10.1016/j.rse.2018.09.018

Potgieter AB, George-Jaeggli B, Chapman SC, Laws K, Suárez Cadavid LA, Wixted J *et al.,* 2017. Multi-spectral imaging from an unmanned aerial vehicle enables the assessment of seasonal leaf area dynamics of sorghum breeding lines. Front Plant Sci 8: 1532. https://doi.org/10.3389/fpls.2017.01532

Puliti S, Ørka HO, Gobakken T, Næsset E, 2015. Inventory of small forest areas using an unmanned aerial system. Remote Sens 7: 9632-9654. https://doi.org/10.3390/rs70809632

Puliti S, Talbot B, Astrup R, 2018. Tree-stump detection, segmentation, classification, and measurement using unmanned aerial vehicle (UAV) imagery. Forests 9: 102. https://doi.org/10.3390/f9030102

Quintano C, Fernández-Manso A, Calvo L, Roberts DA, 2019. Vegetation and soil fire damage analysis based on species distribution modelling trained with multispectral satellite data. Remote Sens 11: 1832. https://doi.org/10.3390/rs11151832

Richardson DM, David M, Le Roux JJ, Wilson JRU, 2015. Australian *Acacia*s as invasive species: lessons to be learnt from regions with long planting histories. Southern Forests 77: 31-39. https://doi.org/10.2989/20702620.2014.999305

Skowronek S, Ewald M, Isermann M, Van De Kerchove R, Lenoir J, Aerts R *et al.,* 2017. Mapping an invasive bryophyte species using hyperspectral remote sensing data. Biol Invasions 19: 239-254. https://doi.org/10.1007/s10530-016-1276-1

Souza-Alonso P, Rodríguez J, González L, Lorenzo P, 2017. Here to stay. Recent advances and perspectives about *Acacia* invasion in Mediterranean areas. Ann For Sci 74: 5. https://doi.org/10.1007/s13595-017-0651-0

Stenzel S, Fassnacht FE, Mack B, Schmidtlein S, 2017. Identification of high nature value grassland with remote sensing and minimal field data. Ecol Ind 74: 28-38. https://doi.org/10.1016/j.ecolind.2016.11.005

van Wilgen BW, Dyer C, Hoffmann JH, Ivey P, Le Maitre DC, Moore JL *et al.,* 2011. National-scale strategic approaches for managing introduced plants: insights from Australian acacias in South Africa. Divers Distrib 17: 1060-1075. https://doi.org/10.1111/j.1472-4642.2011.00785.x

van Wilgen BW, Fill JM, Baard J, Cheney C, Forsyth A, Kraaij T, 2016. Historical costs and projected future scenarios for the management of invasive alien plants in protected areas in the Cape Floristic Region. Biol Conserv 200: 168-177. https://doi.org/10.1016/j.biocon.2016.06.008

Vázquez de la Cueva A, 2014. Case studies of the expansion of *Acacia dealbata* in the valley of the river Mino (Galicia, Spain). Forest Syst 23(1): 3-14. https://doi.org/10.5424/fs/2014231-02531

Vázquez de la Cueva A, Climent JM, Casais Calo L, Quintana JR, 2015. Current and future estimates for the fire frequency and the fire rotation period in the main woodland types of peninsular Spain: a case-study approach. Forest Syst 24(2): e031. https://doi.org/10.5424/fs/2015242-06454

Vicente JR, Alagador D, Guerra C, Alonso JM, Kueffer C, Vaz AS *et al.,* 2016. Cost-effective monitoring of biological invasions under global change: a model-based framework. J Appl Ecol 53: 1317-1329. https://doi.org/10.1111/1365-2664.12631

Ward BG, Bragg TG, Hayes BA, 2014. Relationship between fire-return interval and mulga (*Acacia* aneura) regeneration in the Gibson Desert and Gascoyne-Murchison regions of Western Australia. Int J Wildland Fire 23: 394-402. https://doi.org/10.1071/WF13007

Zhang J, Wang C, Yang C, Xie T, Jiang Z, Hu T *et al.,* 2020. Assessing the effect of real spatial resolution of in situ UAV multispectral images on seedling rapeseed growth monitoring. Remote Sens 12: 1207. https://doi.org/10.3390/rs12071207