EDICIONES COMPLUTENSE

Contribution of free satellite time-series images to mapping plant communities in the Mediterranean Natura 2000 site: the example of Biguglia Pond in Corse (France)

Sébastien Rapinel[1] (iD), Clémence Rozo[1], Pauline Delbosc[2] (iD), Frédéric Bioret[2], Jan-Bernard Bouzillé[3] & Laurence Hubert-Moy[1]

Abstract. Mapping plant communities, which is essential to assess the conservation status of natural habitats, is currently based mainly on time-consuming field surveys without the use of satellite data. However, free image time-series with high spatial and temporal resolution have been available since 2015. This study assessed the contribution of Sentinel-2 time-series images to mapping the spatial distribution of 18 plant communities within a Natura 2000 site (1978 ha) located on the Mediterranean biogeographical region (Corsica, France). The method was based on random forest modeling of six Sentinel-2 images acquired from 26 February to 24 October 2017, which were calibrated and validated using a field vegetation map. The results showed that the 18 plant communities were modeled correctly, with 72% overall accuracy. The uncertainty map associated with the model indicated areas that required additional field observations.
Keywords: Coastal lagoon; Corsica; multispectral images; Sentinel-2.

Cartografía de las comunidades vegetales mediante imágenes gratuitas de series cronológicas de satélites en el estanque de Biguglia (Natura 2000, Córcega, Francia)

Resumen. La cartografía de las comunidades vegetales es esencial para evaluar el estado de conservación de los hábitats naturales. Actualmente, el mapeo de la comunidad de plantas se basa principalmente en estudios de campo que requieren mucho tiempo, sin tener en cuenta los datos satelitales. Sin embargo, desde 2015 se dispone de series temporales gratuitas de imágenes Sentinel-2 con una alta resolución espacial y temporal. El objetivo de este estudio era evaluar la contribución de las imágenes de las series temporales del Sentinel-2 a la modelización de la distribución espacial de 18 comunidades vegetales en un espacio Natura 2000 de 1978 hectáreas situado en la costa mediterránea (Córcega, Francia). El método se basó en un modelado forestal aleatorio de 6 imágenes Sentinel-2 adquiridas entre el 26 de febrero y el 24 de octubre de 2017, que fueron calibradas y validadas utilizando un mapa de vegetación. Los resultados mostraron que las 18 comunidades vegetales se modelaron correctamente con una precisión global del 72%. El mapa de incertidumbre asociado con el modelo indicaba las áreas que requerían observaciones de campo adicionales.
Palabras clave: Laguna costera; Córcega; imágenes multiespectrales; Sentinel-2.

Introduction

Mapping plant communities is a major challenge when monitoring biodiversity, especially in the framework of the European Union (EU) Habitats Directive (Attorre *et al*., 2018). Since plant communities indicate the conservation status of natural habitats (Rodwell *et al*., 2018), mapping and monitoring them is essential to assess impacts of management practices on plant biodiversity (Silva *et al*., 2019). Maps of plant communities are usually produced from time-consuming field observations supplemented by visual analysis of aerial photographs (Díez-Garretas *et al.*, 2019). These maps are flawed, however, because they (i) do not cover all landscapes but rather tend to focus on heritage sites (Campagnaro *et al*., 2019), (ii) are outdated because they are not updated frequently (EEA, 2015), and (iii)

include delimitation or characterization bias related to the map producer's interpretations (Ullerud *et al*., 2018).

Recent remote sensing data with very high spectral, spatial, or temporal resolution offer a promising alternative for mapping plant communities. Hyperspectral and superspectral (i.e. more than 10 bands) data have been commonly used to accurately map plant communities (Wang & Gamon, 2019). For example, 19 grassland plant communities in eastern Europe were spatially discriminated with 81% accuracy using an airborne image acquired in the visible and near-infrared region, with 128 spectral bands at 5 m spatial resolution (Burai *et al*., 2015). Fifteen salt grassland plant communities on the Atlantic coast were modeled with 95% accuracy from a WorldView-3 satellite data, with 16 spectral bands and a spatial resolution of 0.3 m (Collin *et al*., 2018). Unmanned aerial vehicles have also been used to acquire

[1] LETG UMR 6554 CNRS, Université de Rennes. Place du recteur Henri le Moal. F-35043 Rennes, France. Email: sebastien.rapinel@univ-rennes2.fr, sebastienrapinel@gmail.com, clemence.rozo@univ-rennes2.fr, laurence.moy@univ-rennes2.fr
[2] Brest University, EA 7462 Géoarchitecture. F-29200 Brest, France. Email: delboscpauline@gmail.com, frederic.bioret@univ-brest.fr
[3] ECOBIO UMR 6553 CNRS, Université de Rennes. Avenue du Général Leclerc. F-35042 Rennes, France. Email: jbbouzille@gmail.com

multispectral or hyperspectral data that successfully discriminated plant communities (Aasen *et al.*, 2018). For example, nine plant communities from grasslands and moors in Portugal were mapped with 89% accuracy from RGB drone data at a resolution of 0.1 m (Gonçalves *et al.*, 2016). In addition, several studies have shown that annual satellite time-series can be used to model plant communities accurately. For example, seven wet grassland plant communities in Germany were modeled with 85% accuracy using an annual time-series of either 13 TerraSAR-X SAR satellite data at a spatial resolution of three m or 6 RapidEye optical satellite data at a spatial resolution of 6 m (Schuster *et al.*, 2015). However, the use of these remote sensing data by botanists raises two related issues: (i) the high cost of acquiring the data, especially using airborne platforms (Turner *et al.*, 2015) and (ii) the high degree of expertise required to analyze the data –especially hyperspectral, SAR, and drone data– and to pre-process them before integrating them in a GIS (Borre *et al.*, 2011).

Among the remote sensing data available, Sentinel-2 satellite time-series are free of charge (e.g. from the Copernicus platform of the European Space Agency) and in a ready-to-use GIS format (i.e. geometrically and atmospherically corrected) (Hagolle *et al.*, 2015). Sentinel-2 sensors provide 13 spectral bands in the visible, near-, and mid-infrared spectra, with a spatial resolution of up to 10 m. Although these characteristics are useful for monitoring the phenology of natural vegetation (Vrieling *et al.*, 2018) or characterizing vegetation functional traits (Schauman *et al.*, 2018), the contribution of Sentinel-2 data to mapping plant communities remains poorly explored. However, a recent study revealed that seven wet grassland plant communities derived from unsupervised classification of vegetation plots were modeled with 78% accuracy using an annual Sentinel-2 time-series (Rapinel *et al.*, 2019). These initial results must be confirmed with further studies in other regions with other types of plant communities defined within the Prodromus framework (Reymann *et al.*, 2017).

The objective of this study was to evaluate the contribution of Sentinel-2 time-series images to mapping grassland, shrubland, tree communities and related directive habitats, in a Natura 2000 site in the Mediterranean biogeographical region.

Table 1. The main plant communities identified at the study site and their corresponding habitat codes (Annex 1 of the European Union Habitats Directive). The asterisk denotes a priority habitat. Plant communities are sorted by ecological unit.

Ecological unit		Syntaxa	Habitat code
Coastal lagoons	1.	*Ruppietum cirrhosae* Hocquette 1927 corr. Iversen 1934	1150*
Salt marshes	2.	*Suaedo maritimae-Salicornietum patulae* Brullo & Furnari 1976 Géhu & Géhu-Franck 1984	1310
	3.	*Arthrocnemo glauci-Salicornietum emerici* O. Bolos 1962 Géhu & Géhu-Franck 1978	1310
	4.	*Juncetum acuti* Molinier & Tallon 1969	1410
Semi-natural grasslands	5.	*Trifolio fragiferi-Cynodontion dactylonis* Braun-Blanq. & O. Bolos 1958	6420
	6.	*Lino biennis-Festucetum arundinaceae* Dubuis & Simmoneau ex B. Foucault 2012	6420
	7.	*Echio lycopsis-Galactitetum tomentosae* Molinier 1937	-
	8.	*Scirpetum compacto-littoralis* Braun-Blanq. (1931) 1952 Rivas Mart. *et al.* 1980	-
Reeds	9.	*Inulo crithmoidis-Phragmitetum australis* Gamisans 1992	-
	10.	*Kosteletzkyo pentacarpos-Phragmitetum australis* Gamisans 1992	-
	11.	*Phragmitetum australis* Schmale 1939	-
	12.	*Phragmitetum australis* Schmale 1939 *calystegietosum sepii* Gamisans 1992	-
Scrubs	13.	*Cisto salviifolii-Halimietum halimifolii* Géhu & Biondi 1994	2260
	14.	*Pulicario odoratae-Arbutetum unedonis* Allier & Lacoste 1980 Reymann *et al.* 2016 *quercetosum suberis* Gamisans 1991 Aurière & Lejour 2016	-
	15.	*Pruno spinosae-Rubion ulmifolii* O. Bolos 1954	-
Forests	16.	*Inulo crithmoidis-Tamaricetum africanae* Gamisans 1992	92D0
	17.	*Althaeo officinalis-Tamaricetum africanae* Gamisans 1992	92D0
	18.	*Angelico sylvestris-Alnetum glutinosae* Gamisans 2013	-

Materials and Methods

Study site

The Biguglia pond is a site of Community Importance (FR9410101 SCI) and covers 1978 ha along the eastern coast of Corsica, France (Figure 1). The topography is flat with elevation ranging between 0 and 3 meters above sea level. The site has a Mediterranean climate (mean annual precipitation 835 mm, mean annual temperature 16°C), with summer droughts from June to September (Delbosc, 2015). The site consists of a coastal lagoon enclosed by salt marshes, sand dunes, grasslands, thickets, and wooded groves. It is designated a Natura 2000 site because it contains several habitats of EU community importance, such as 1150 "Coastal lagoons", 1310 "*Salicornia* and other annuals colonizing mud and sand", 1410 "Mediterranean salt meadows (*Juncetalia maritimi*)", 2260 "*Cisto-Lavanduletalia* dune sclerophyllous scrubs", 6420 "Mediterranean tall humid grasslands of the *Molinio-*

Holoschoenion", and 92D0 "Southern riparian galleries and thickets (*Nerio-Tamaricetea* and *Securinegion tinctoriae*)". A management plan has been developed to conciliate conservation, agricultural practices (grazing, mowing), and recreation (educational activities). A field vegetation map was produced at 1:2500 scale (Figure 1) for reporting purposes for the Habitats Directive (Delbosc *et al.*, 2020). To this end, 118 vegetation plots were sampled from May to June 2014 according to the Braun-Blanquet method (Delbosc, 2015). Species taxonomy and nomenclature followed the "Flora Corsica" (Jeanmonod & Gamisans, 2013). Each vegetation plot was assigned to a plant community, following the Corsican Vegetation Prodromus (Reymann *et al.*, 2017). The extent of each plant community was mapped by interpreting aerial photographs acquired in 2013. Table 1 lists the plant communities identified and represented in the vegetation map with a spatial extent sufficiently large to be detected in satellite data with a spatial resolution of 10 m, along with their corresponding habitat code (Annex 1, Habitats Directive).

Spectral analyses

Sentinel-2 satellite data were selected from the Theia catalogue (theia.land.fr) among cloudless images from 2017 (Figure 1). Images were already orthorectified and atmospherically corrected using the MACCS-ATCOR Joint Algorithm (MAJA) (Hagolle *et al.*, 2015). Six images acquired from 26 February to 24 October 2017 were retained (Figure 2). For each image we used the 4 bands at 10 m spatial resolution acquired in the blue (b2: 497 nm), green (b3: 560 nm), red (b4: 664 nm), and near-infrared (b8: 835 nm) spectra, as well as the 6 bands at 20 m spatial resolution acquired in the red-edge (b5: 704 nm, b6: 740 nm, b7: 782 nm), near-infrared (b8a: 865 nm), and short-wave infrared (b11: 1614 nm, b12: 2202 nm) spectra. The three atmospheric bands (b1, b9 and b10) were discarded. The six spectral bands at 20 m spatial resolution were resampled to 10 m. As a result, the satellite data used as input for modeling plant communities included 60 variables (10 spectral bands × six dates) at 10 m spatial resolution.
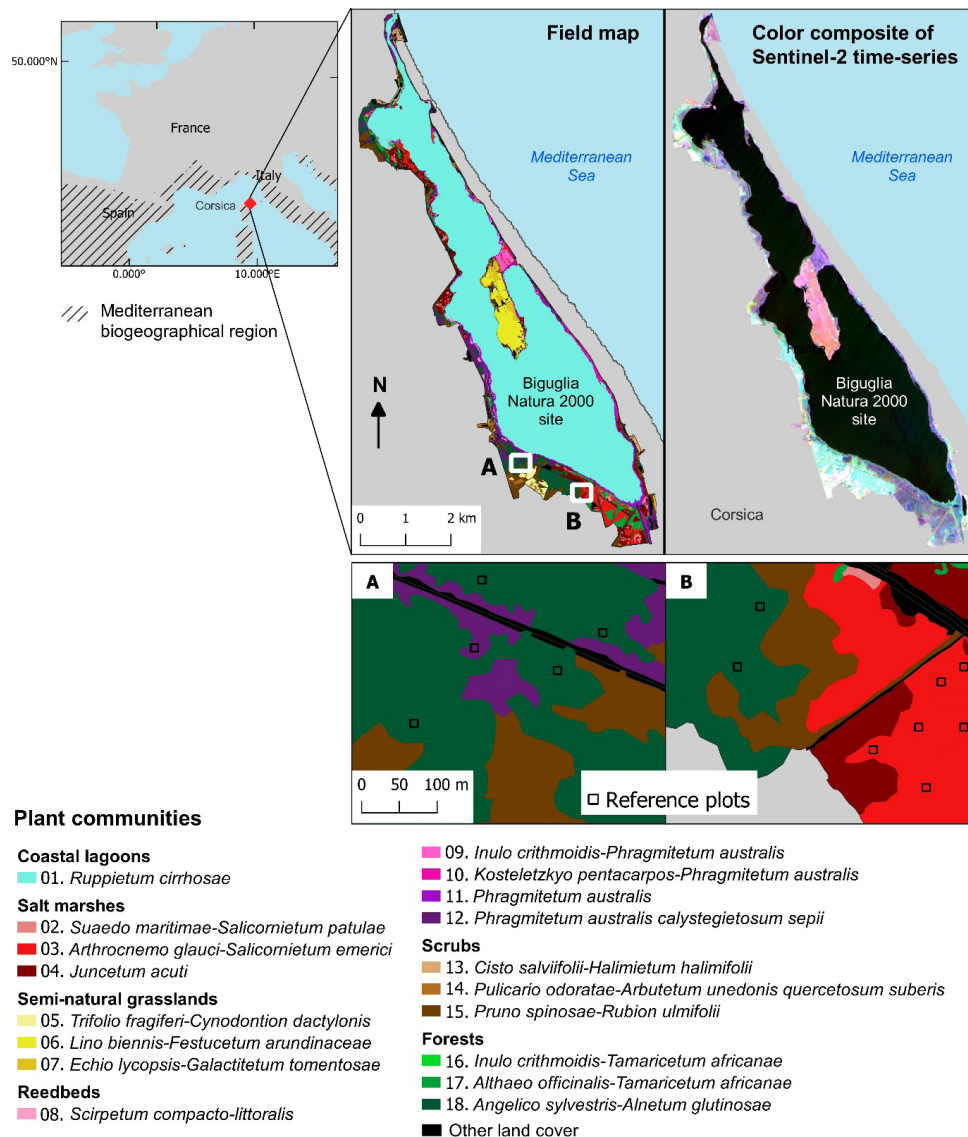

Figure 1. Study site location (left panel), visualization of field-based vegetation map (middle panel), and color composite of Sentinel-2 time-series (© ESA): blue = band 8, 26 February 2017, green = band 8, 17 May 2017 and red = band 8, 24 October 2017 (right panel).

The pairwise spectral distance between plant communities was calculated using the Jeffries-Matusita (JM) distance. The JM distance is commonly used in remote sensing science to measure the average distance between two classes (Richards, 2013), where the lowest value (0) indicates identical classes that are impossible to separate and the highest value (1.41) reflects two classes that are perfectly separable. A dimension reduction was performed using a principal component analysis to reduce the collinearity between the 60 spectral variables (10 bands × 6 dates). The first five components of the PCA (90.2% of the variance) were used to calculate the JM distance.

Plant communities were mapped using random forest (RF) modeling (Breiman, 2001). The RF model is widely used in remote sensing science because of its ability to manage a large number of variables, its low sensitivity to mislabeling errors, and its ability to map model uncertainty (Belgiu & Drăguţ, 2016). Reference plots were selected from the vector field vegetation map at 1:2500 scale to calibrate and validate the RF model. To this end, an expert botanist manually identified the Sentinel-2 pixels in GIS by overlapping a multi-temporal near-infrared Sentinel-2 color composite (26 February, 17 May, and 24 October) and Google Earth images with the contours of the vegetation map. The botanist selected or excluded pixels from the sample based on

uncertainties in (i) the spatial extent of polygons in the vegetation map, (ii) changes in the vegetation between the date of the field map and the year of Sentinel-2 image acquisition, and (iii) the spatial intermingling of plant communities not specified within a polygon in the vegetation map.

To increase the independence of the observations, sub-sampling was performed at a minimum distance of 20 m between two reference plots of the same plant community. A total of 1224 reference plots were collected for the 18 plant communities. The model was calibrated and validated using a 10-repeated 5-fold approach (Kuhn & Johnson, 2013). To assess model accuracy, we calculated the overall accuracy index, the Kappa index, and the F1 score per plant community (Congalton & Green, 2008). The F1 score is an accuracy metric specific to each class (i.e. plant community) that reflects a weighted average between precision (user's accuracy) and recall (producer's accuracy) calculated as follows:

$$\mathbf{F1} = 2 \times \frac{(precision \times recall)}{(precision + recall)}$$

Analyses were performed using R software (R Core Team, 2015) with the raster (Hijmans, 2015), rgdal (Bivand *et al.*, 2015), and caret (Kuhn, 2008) packages.
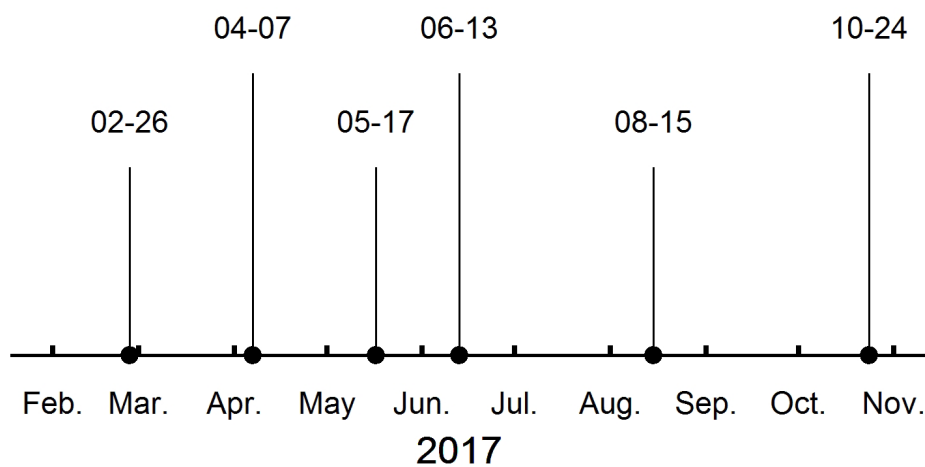

Figure 2. Acquisition dates of the six satellite Sentinel-2 images used in the study.

Results

The pairwise spectral separability between plant communities based on the JM distance is shown in Table 2. Overall, plant communities are spectrally well separated (median JM distance value = 1.37). Consistently, the spectral separability decreases with the increase of the similarity in floristic composition. For example, the spectral response of salt marshes plant community 2 (*Suaedo maritimae-Salicornietum patulae*) is very

distinct (JM distance = 1.41) from that of grasslands plant community 5 (*Trifolio fragiferi-Cynodontion dactylonis*) but closer (JD distance = 0.90) to that of salt marshes plant community 3 (*Arthrocnemo glauci-Salicornietum emerici*). Nevertheless, this relationship between spectral separability and floristic composition is not systematically confirmed: for example, spectral separability between salt marshes and reed beds plant communities is moderate (JM distances 0.89−1.25), possibly due to the frequent occurrence of open water.

Table 2. Pairwise spectral separability of the 18 plant communities based on JM distance. ID refers to the plant community code number (Table 1). The triangles outlined in black identify the main ecological units. Higher JM distance values are indicated in gray (≥1.2) and dark gray (≥1.3).

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1		1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41
2			0.90	0.92	1.41	1.41	1.41	1.07	1.22	1.20	1.09	1.25	1.36	1.35	1.32	1.26	1.23	1.40
3				0.85	1.41	1.41	1.41	0.92	1.12	1.09	0.89	1.22	1.38	1.33	1.29	1.12	1.11	1.40
4					1.40	1.41	1.41	1.14	1.25	1.23	1.09	1.22	1.34	1.23	1.20	1.17	1.09	1.37
5						1.39	1.34	1.39	1.41	1.41	1.40	1.32	1.41	1.34	1.24	1.38	1.41	1.41
6							1.00	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41
7								1.41	1.41	1.41	1.41	1.41	1.38	1.40	1.41	1.41	1.41	1.41
8									1.20	1.09	0.65	1.02	1.38	1.33	1.31	1.15	1.19	1.40
9										1.11	1.22	1.37	1.38	1.40	1.38	1.39	1.36	1.41
10											1.08	1.29	1.40	1.39	1.36	1.37	1.36	1.40
11												1.05	1.38	1.34	1.33	1.12	1.18	1.40
12													1.41	1.18	1.07	1.17	1.23	1.33
13														1.40	1.40	1.40	1.40	1.41
14															1.00	1.17	1.23	1.19
15																1.20	1.25	1.21
16																	0.84	1.38
17																		1.34
18																		

According to the confusion matrix between the 18 plant communities modeled from the Sentinel-2 time series and the reference plots (Table 3), modeling accuracy was satisfactory (overall accuracy 72%, Kappa index 0.70), but differed significantly among classes:

- Plant communities 1. *Ruppietum cirrhosae*, 5. *Trifolio fragiferi-Cynodontion dactylonis*, 6. *Lino biennis-Festucetum arundinaceae*, 13. *Cisto salviifolii-Halimietum halimifolii*, and 18. *Angelico sylvestris-Alnetum glutinosae* were modeled with high accuracy (F1 scores: 81–99).
- Plant communities 2. *Suaedo maritimae-Salicornietum patulae*, 4. *Juncetum acuti*, 10. *Kosteletzkyo pentacarpos-Phragmitetum australis*, 12. *Phragmitetum australis calystegietosum sepii*, 14. *Pulicario odoratae-Arbutetum unedonis*

quercetosum suberis, and 17. *Althaeo officinalis-Tamaricetum africanae* were modeled with good accuracy (F1 scores: 70–77).

- Plant communities 3. *Arthrocnemo glauci-Salicornietum emerici*, 7. *Echio lycopsis-Galactitetum tomentosae*, 8. *Scirpetum compacto-littoralis*, 9. *Inulo crithmoidis-Phragmitetum australis*, 11. *Phragmitetum australis*, and 15. *Pruno spinosae-Rubion ulmifolii* were modeled with moderate accuracy (F1 scores: 57–68).
- Plant community 16. *Inulo crithmoidis-Tamaricetum africanae* was modeled with low accuracy (F1 score: 31) due to its strong confusion with plant community 17. *Althaeo officinalis-Tamaricetum africanae*, since these two communities included *Tamarix africana*.

Table 3. Confusion matrix between the 18 plant communities modeled from the Sentinel-2 time-series (rows) and the reference plots (columns). Entries are percentages of counts. ID, plant community code number (Table 1); UA, User's Accuracy; Count, total number of reference plots used per plant community; PA, Producer's Accuracy. The squares outlined in grey identify the main ecological units.

ID	Reference																		
Mod.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	UA
1	100											1							98
2		63	4	2				3											83
3		18	73	10				3	9	1	3	1			2	7	3	1	64
4		15	11	79					13	1	6		6	5	7	4	3		63
5					90	1	4			1				2	6				85
6					5	91	47							1					73
7					3	8	49						6	2					70
8		1	1					52		2	4	4		1		5			70
9			3					8	59	4	3								65
10		1						4	9	73	7	3							76
11			4	3				17	9	13	65	7		1		11	1		58
12				2				8		4	8	72		2	11	4		3	72
13			1										88			2			84
14												2		72	10	3	1	1	83
15				3					2			6		5	52	1	1	2	64
16			1					1			1	1			1	21	1		61
17		3	2	3				4			5	3		3	6	41	89	5	68
18														6	6			88	87
Count	57	67	99	100	50	78	48	62	41	71	92	100	17	75	56	43	100	68	
PA	100	63	73	79	90	91	49	52	59	73	65	72	88	72	52	21	89	88	
F1 score	99	71	68	70	87	81	57	60	62	74	61	72	86	77	57	31	76	87	

Comparing maps of the plant community and RF model certainty to the very high spatial resolution Google Earth image (Figure 3, top) showed that the model reflected the fine-grained pattern of plant communities correctly, especially for grassland areas. In addition, scrub or forest plant communities were well identified, including those structured as hedges. Analysis of the certainty map indicated large disparities between areas of high certainty (≥ 0.8) in the middle of large patches of forest or grassland plant communities, and areas of low certainty (≤ 0.3) that corresponded in particular to transitions or a *continuum* between plant communities.
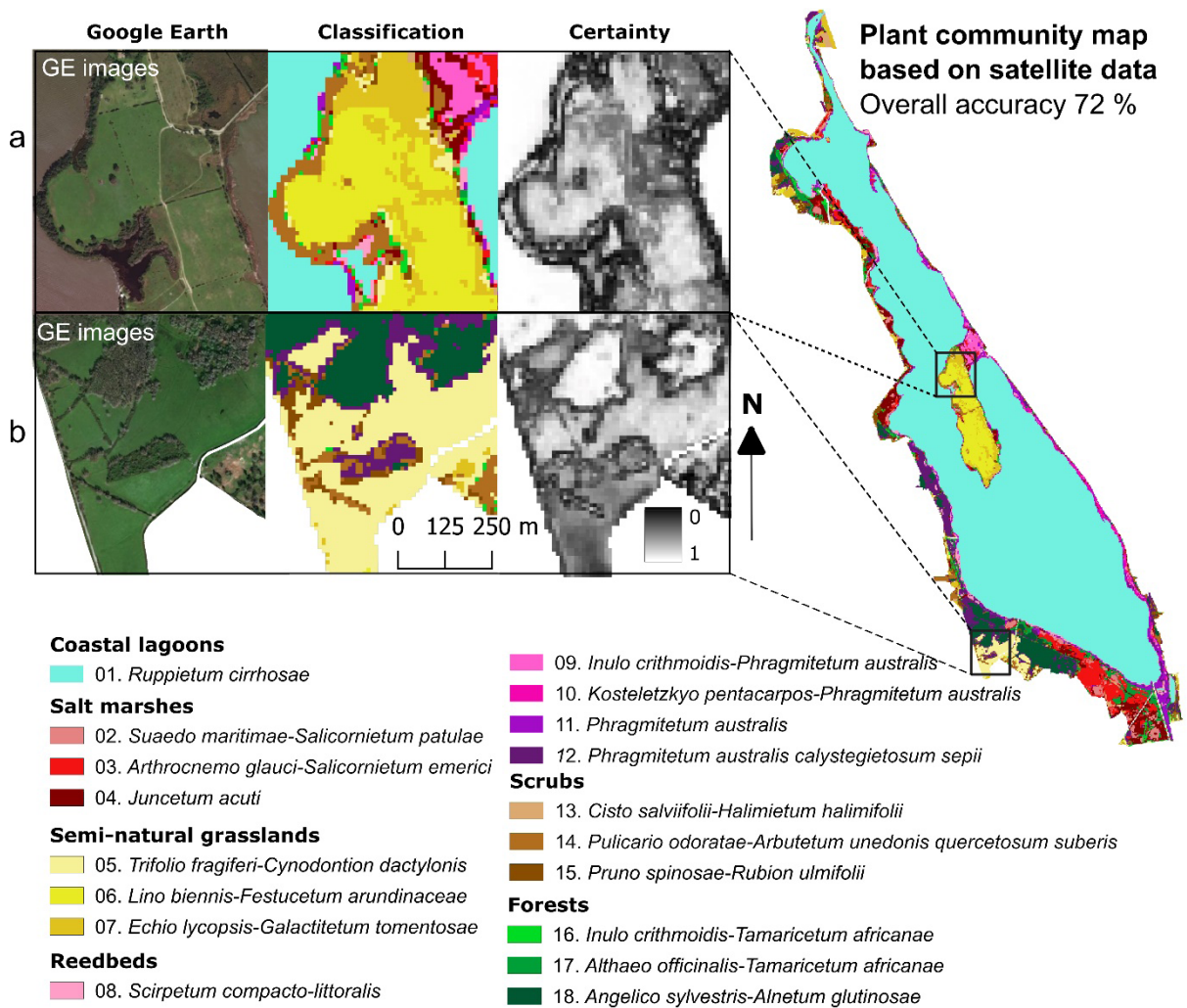
Coastal lagoons
- 01. *Ruppietum cirrhosae*

Salt marshes
- 02. *Suaedo maritimae-Salicornietum patulae*
- 03. *Arthrocnemo glauci-Salicornietum emerici*
- 04. *Juncetum acuti*

Semi-natural grasslands
- 05. *Trifolio fragiferi-Cynodontion dactylonis*
- 06. *Lino biennis-Festucetum arundinaceae*
- 07. *Echio lycopsis-Galactitetum tomentosae*

Reedbeds
- 08. *Scirpetum compacto-littoralis*

- 09. *Inulo crithmoidis-Phragmitetum australis*
- 10. *Kosteletzkyo pentacarpos-Phragmitetum australis*
- 11. *Phragmitetum australis*
- 12. *Phragmitetum australis calystegietosum sepii*

Scrubs
- 13. *Cisto salviifolii-Halimietum halimifolii*
- 14. *Pulicario odoratae-Arbutetum unedonis quercetosum suberis*
- 15. *Pruno spinosae-Rubion ulmifolii*

Forests
- 16. *Inulo crithmoidis-Tamaricetum africanae*
- 17. *Althaeo officinalis-Tamaricetum africanae*
- 18. *Angelico sylvestris-Alnetum glutinosae*

Figure 3. Maps of plant communities at the Biguglia site derived from modeling Sentinel-2 satellite data

Discussion

Plant communities in Natura 2000 sites located within the Mediterranean biogeographical region, which includes a wide diversity of ecological units, can be accurately modeled (overall accuracy 72%) using a random forest classification model and Sentinel-2 time-series. This overall accuracy was slightly lower than that obtained by Rapinel *et al.* (2019) for semi-natural Atlantic grasslands (78%), but the modeling process was more challenging due to the larger number of plant communities (18 vs. 7, respectively) and consequently more intermingled vegetation types and complex patterns. Although the high accuracy of modeling the amphibious *Ruppietum cirrhosae* plant community (F1 score 99) was not remarkable (since vegetation in coastal lagoons has a more distinct spectral signature than that in terrestrial plant communities) (MacKay *et al.*, 2009), some plant communities in the same ecological unit were discriminated well. For example, the *Trifolio fragiferi-Cynodontion dactylonis* and *Lino biennis-Festucetum arundinaceae* communities were modeled correctly (F1 score 87 and 81, respectively), even though both contained common species such as *Cynodon dactylon* and *Ranunculus sardous* and were associated with the same habitat (6420 "Mediterranean tall humid grasslands of

the *Molinio-Holoschoenion*"). The strong discrimination between these 2 plant communities can be explained by phenological differences observed among the 6 satellite data acquired in all 4 seasons of the same year (Figure 2). Recent studies indicate the potential of Sentinel-2 time series to discriminate fine-scale vegetation patterns based on phenological profiles (Vrieling *et al.*, 2018). Conversely, the model confused plant communities patterned by a common and covering species that strongly influenced the spectral signature, such as for the 2 forest plant communities (*Inulo crithmoidis-Tamaricetum africanae* and *Althaeo officinalis-Tamaricetum africanae*), which were physiognomically and structurally patterned by *Tamarix africana*, and the 4 reed bed plant communities (*Inulo crithmoidis-Phragmitetum australis*, *Kosteletzkyo pentacarpos-Phragmitetum australis*, *Phragmitetum australis*, and *Phragmitetum australis calystegietosum sepii*) with high coverage of *Phragmites australis*. Remote sensing data with higher spatial and spectral resolution are required to discriminate these plant communities more accurately (Corbane *et al.*, 2015).

This study aimed to evaluate the contribution of an annual Sentinel-2 satellite time series to model 18 plant communities. However, some plant communities could not be properly discriminated based only on the use

of spectral variables. The accuracy of the model could be increased by combining several types of variables (Fois *et al*., 2018) providing information on vegetation structure, micro-topography, micro-climate or soil characteristics. In this case, the potential envelope of each plant community could first be modelled based on environmental variables and then the realized plant community area could be predicted based on spectral variables derived from remote sensing data. As an example, this methodology was successfully applied by Álvarez-Martínez *et al*. (2018) to model 24 Natura 2000 habitats on a 120,000-ha site by combining variables derived from remote sensing data (1 Landsat 8 OLI image + 1 canopy height model derived from LiDAR data) with topographic and climatic variables.

Although botanists do not yet use remote sensing data widely (Borre *et al*., 2017), this study illustrates that the data and method used to model plant communities and related habitats do not present any major technical challenges. First, the pre-processed Sentinel-2 satellite data (Level 2A) can be downloaded free of charge from the Theia (theia.land.fr) or Copernicus hub (https://scihub.copernicus.eu) websites in a format that is ready-to-use in GIS. Second, satellite data can be analyzed using R software, which is commonly used in vegetation science (Dixon & Palmer, 2003; Zelený & Tichý, 2009). This method models plant communities with < 75% accuracy, which is insufficient to be considered a baseline vegetation map. However, combining it with the RF model uncertainty map provides a valuable tool for applying preferential sampling (De Cáceres *et al*., 2015) to new areas to prospect during future field missions (Zlinszky *et al*., 2014) and thus improves vegetation typology and mapping accuracy (Fanelli *et al*., 2005).

This mapping approach for plant communities has several advantages compared to approaches based on expert decision rules of vegetation physiognomy and environmental criteria (Adamo *et al*., 2016): (i) it highlights floristic variations within a habitat (Feilhauer *et al*., 2014); (ii) allows the conservation status of a habitat to be assessed (Angelini *et al*., 2018); and (iii) avoids one-to-many correspondences between a vegetation physiognomy and habitats (Tomaselli *et al*., 2016). The disadvantage of this approach, however, is the need for reference data at the syntaxonomic level, which unfortunately remains rare (Campagnaro *et al*., 2019).

The quality of the reference data is an important issue when modeling with satellite data (Maxwell *et al*., 2018). Ideally, the reference plots are selected from phytosociological records collected in the same year as the satellite data acquisitions (Rapinel *et al*., 2019). This was not the case in the current study because the archive vegetation plots (2 × 2 m) were smaller than the resolution of the Sentinel-2 satellite data (10 m). As a default, reference points were collected within the polygons of the field vegetation map, and although we adopted precautions, it inevitably resulted in biases related to (i) the 3-year difference between producing the field map (2014) and acquiring the Sentinel-2 data (2017) and (ii)

subjective delineation of plant community patches by the field map producer (Ullerud *et al*., 2018). In this study, the reference plots were manually selected, which was time consuming. While "pure" pixels can be identified and selected automatically, the detection of outliers due to geometric errors made in the field vegetation map remains challenging. Nevertheless, recent studies have shown that outliers could be detected based on their spectral characteristics (Raab *et al*., 2018; Halladin-Dąbrowska *et al*., 2020), which would significantly reduce the efforts required to select reference pixels. This suggests that the limitation for modeling plant communities is no longer access to satellite data but rather access to field reference data (Borre *et al*., 2017). This highlights the crucial importance for botanists to georeference vegetation plots accurately and, above all, to structure and disseminate them in databases such as the European Vegetation Archive (Chytrý *et al*., 2016) or VegFrance (Bonis and Bouzillé, 2012) so that they can be used to calibrate and validate models based on satellite data.

Although the results of modeling plant communities revealed heterogeneous patterns (Figure 3), the spatial resolution of Sentinel-2 satellite data (10 m) obviously remains too coarse to detect small vegetation patches (Rapinel *et al*., 2019). Consequently, we modeled only 18 plant communities. Thus, many plant communities that occur only occasionally but may nonetheless characterize habitats of community importance were excluded. Using satellite data with a higher spatial resolution could resolve this limitation. SPOT-6 images, which are available at low cost and with a spatial resolution of 1.5 m, offer an interesting alternative to Sentinel-2 images for sites where mapping fine-grained vegetation patterns is critical.

The map of plant communities derived from Sentinel-2 images can be used as a decision support tool for the conservation status assessment of natural habitats. For example, the presence of the plant community *Arthrocnemo glauci-Salicornietum emerici* —typical of Mediterranean salt meadows— in the Natura 2000 site of Biguglia is remarkable given its large extent (61 ha) and indicates a good conservation status of the habitat HD 1310 (*Salicornia* and other annuals colonising mud and sand) (Bensettiti *et al*., 2004). In addition, the uncertainty map provides information that can be used for implementing different management plans, since it highlights transition areas —or *continua*— between plant communities, which are rarely provided by field vegetation maps (Schmidtlein *et al*., 2007). These *continua* can be considered as biodiversity hot spots that have a high probability of change (Álvarez-Martínez *et al*., 2011). In the case of the Biguglia site that has a very low and flat micro topography, the *continua* are mainly structured by variations in soil moisture and salinity in relation to climatic hazards (i.e. fluvial inflow during precipitation and seawater inflow during storms) (Delbosc *et al*., 2020). Thus, these *continua* —highlighted by Sentinel-2 images— reflect the effective functioning and self-regulating of the lagoon ecosystem.

Remote sensing data are a tool with standardized and spatialized measurements over long time scales suitable for monitoring changes in biodiversity (Schrodt *et al.*, 2019), such as the Habitats Directive reporting. In our study, we mapped plant communities using one-year satellite time-series. In the perspective of longer-term, interannual monitoring, the major factors in the dynamics of natural vegetation that could be easily considered in the site of Biguglia due to the very high accuracy of annual maps derived from remote sensing data such Sentinel-2 images are fire hazard (García-Llamas *et al.*, 2019 and land use and cover changes (e.g. Álvarez-Martínez *et al.*, 2010; Mallinis *et al.*, 2011; Kallimanis *et al.*, 2015; Amici *et al.*, 2017). Conversely, changes in plant communities, for example in response to a decrease in grazing pressure, remain challenging to quantify not only because of the lower accuracy of annual maps (e.g. 72% in this study) — the changes observed being related both to real vegetation dynamics and to cumulative modelling errors (Álvarez-Martínez *et al.*, 2011)— but also because of the limited access to georeferenced historical field vegetation data that are needed to calibrate and validate the remote sensing-based models (Anderson, 2018). To overcome these methodological issues, several solutions can be considered, such as: i) the "time-first, space-latter" modelling approach (i.e. modelling of inter-annual time profiles rather than independent modelling by year) (Picoli *et al.*, 2018); ii) the consideration of uncertainty in the change detection (Álvarez-Martínez *et al.*, 2010); iii) the exploitation of vegetation archives (e.g. Chytrý *et al.*, 2016).

Conclusions

This study demonstrates that plant communities in Natura 2000 sites located within the Mediterranean biogeographical region, which includes a wide diversity of ecological units, were modeled accurately using a random forest classification model and Sentinel-2 satellite data time series. The approach, based on the use of free, ready-to-use images processed using R software, can be used by botanists. Although the plant community map derived from satellite data is not sufficiently accurate to be considered a reference document, combining it with a model certainty map is a useful tool for targeting areas that require additional field sampling. In perspective, SPOT-6 satellite time series analysis with very high spatial resolution is a promising approach for mapping fine-scale vegetation patterns.

Acknowledgements

References

Aasen, H., Honkavaara, E., Lucieer, A. & Zarco-Tejada, P.J. 2018. Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: a review of sensor technology, measurement procedures, and data correction workflows. Remote Sens.-Basel 10, 1091.

Adamo, M., Tarantino, C., Tomaselli, V., Veronico, G., Nagendra, H. & Blonda, P. 2016. Habitat mapping of coastal wetlands using expert knowledge and Earth observation data. J. Appl. Ecol. 53: 1521–1532.

Álvarez-Martínez, J.M., Stoorvogel, J.J., Suárez-Seoane, S., & Calabuig, E.D.L. 2010. Uncertainty analysis as a tool for refining land dynamics modelling on changing landscapes: a case study in a Spanish Natural Park. Landscape Ecol 25: 1385–1404. Doi:10.1007/s10980-010-9492-z

Álvarez-Martínez, J.M., Suárez-Seoane, S. & Calabuig, E.D.L. 2011. Modelling the risk of land cover change from environmental and socio-economic drivers in heterogeneous and changing landscapes: The role of uncertainty. Landscape Urban Plan. 101: 108–119.

Álvarez-Martínez, J.M., Jiménez-Alfaro, B., Barquín, J., Ondiviela, B., Recio, M., Silió-Calzada, A. & Juanes, J.A. 2018. Modelling the area of occupancy of habitat types with remote sensing. Methods Ecol. Evol. 9: 580–593.

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. Doi: 10.1016/j.ecoinf.2017.04.005

Anderson, C.B. 2018. Biodiversity monitoring, earth observations and the ecology of scale. Ecol. Lett. 21: 1572–1585.

Angelini, P., Chiarucci, A., Nascimbene, J., Cerabolini, B.E., Dalle Fratte, M. & Casella, L. 2018. Plant assemblages and conservation status of habitats of Community interest (Directive 92/43/EEC): definitions and concepts. Ecol. Quest. 29: 87–97.

Attorre, F., Pignatti, S., Spada, F., Casella, L. & Agrillo, E. 2018. Introduction: Vegetation science and the habitats directive: approaches and methodologies of a never-ending story. Rend. Lincei-Sci. Fis. 29: 233–235.

Belgiu, M. & Drăguţ, L. 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens.-Basel 114: 24–31.

Bensettiti, F., Bioret, F., Roland, J., & Lacoste, J.-P. (Eds.). 2004. Cahiers d'habitats Natura 2000. Connaissance et gestion des habitats et des espèces d'intérêt communautaire.Vol. 2: Habitats côtiers. La Documentation française, Paris.

Bonis, A. & Bouzillé, J.B. 2012. The project VegFrance: Towards a national vegetation database for France. Plant Sociol. 49: 97–99.

Borre, J.V., Paelinckx, D., Mücher, C.A., Kooistra, L., Haest, B., De Blust, G. & Schmidt, A.M. 2011. Integrating remote sensing in Natura 2000 habitat monitoring: Prospects on the way forward. J. Nat. Conserv. 19: 116–125.

Borre, J.V., Spanhove, T. & Haest, B. 2017. Towards a Mature Age of Remote Sensing for Natura 2000 Habitat Conservation: Poor Method Transferability as a Prime Obstacle. In: Díaz-Delgado, F., Lucas, R. & Hurford, C. (Eds). The Roles of Remote Sensing in Nature Conservation. Pp. 11–37. Springer, New York.

Breiman, L. 2001. Random Forests. Mach. Learn. 45: 5–32.

Burai, P., Deák, B., Valkó, O. & Tomor, T. 2015. Classification of Herbaceous Vegetation Using Airborne Hyperspectral Imagery. Remote Sens.-Basel 7: 2046–2066.

Campagnaro, T., Sitzia, T., Bridgewater, P., Evans, D. & Ellis, E.C. 2019. Half Earth or Whole Earth: What Can Natura 2000 Teach Us? BioScience 69: 117–124.

Chytrý, M., Hennekens, S.M., Jiménez-Alfaro, B., Knollová, I., Dengler, J., Jansen, F., Landucci, F., Schaminée, J.H., Aćić, S. & Agrillo, E. 2016. European Vegetation Archive (EVA): an integrated database of European vegetation plots. Appl. Veg. Sci. 19: 173–180.

Collin, A., Lambert, N. & Etienne, S. 2018. Satellite-based salt marsh elevation, vegetation height, and species composition mapping using the superspectral WorldView-3 imagery. Int. J. Remote Sens. 33(17): 5619–5637.

Congalton, R.G. & Green, K. 2008. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, 2nd ed. CRC Press, Boca Raton.

Corbane, C., Lang, S., Pipkins, K., Alleaume, S., Deshayes, M., García Millán, V.E., Strasser, T., Vanden Borre, J., Toon, S. & Michael, F. 2015. Remote sensing for mapping natural habitats and their conservation status – New opportunities and challenges. Int. J. Appl. Earth Obs. 37: 7–16.

De Cáceres, M., Chytrý, M., Agrillo, E., Attorre, F., Botta-Dukát, Z., Capelo, J., Czúcz, B., Dengler, J., Ewald, J., Faber-Langendoen, D., Feoli, E., Franklin, S.B., Gavilán, R.G., Gillet, F., Jansen, F., Jiménez-Alfaro, B., Krestov, P., Landucci, F., Lengyel, A., Loidi, J., Mucina, L., Peet, R.K., Roberts, D.W., Roleček, J., Schaminée, J.H.J., Schmidtlein, S., Theurillat, J.-P., Tichý, L., Walker, D.A., Wildi, O., Willner, W., Wiser, S.K. & Scheiner, S. 2015. A comparative framework for broad-scale plot-based vegetation classification. Appl. Veg. Sci. 18: 543–560.

Delbosc, P. 2015. Phytosociologie dynamico-caténale des végétations de la Corse: méthodologies typologique et cartographique. PhD thesis. Univ. Bretagne occidentale, 638p.

Delbosc, P., Bioret, F., Etourneau, S., Panaïotis, C., Piazza, C. & Poli, P. 2020. Typology and Mapping of Plant Landscape of Biguglia Pond: Application to Planning and Conservation Management. In: Delbosc, P., Bioret, F. & Panaïotis, C. (Eds.). Plant Landscape of Corsica: Typology and Mapping Plant Landscape of Cap Corse Region and Biguglia Pond. Springer Int. Publ., Cham. Doi: 10.1007/978-3-030-35776-4_3

Díez-Garretas, B., Comino, O., Pereña, J. & Asensi, A. 2019. Spatio-temporal changes (1956-2013) of coastal ecosystems in Southern Iberian Peninsula (Spain). Mediterr. Bot. 40: 111–119.

Dixon, P. & Palmer, M.W. 2003. VEGAN, a package of R functions for community ecology. J. Veg. Sci. 14: 927–930.

Fanelli, F., Bianco, M.P., D'Angeli, D., De Sanctis, M., Sauli, A., Testi, A. & Pignatti, S. 2005. Remote sensing in phytosociology: The map of vegetation of the Provincia of Rome. Ann. Bot. 5: 173–184.

Feilhauer, H., Dahlke, C., Doktor, D., Lausch, A., Schmidtlein, S., Schulz, G. & Stenzel, S. 2014. Mapping the local variability of Natura 2000 habitats with remote sensing. Appl. Veg. Sci. 17: 765–779.

Fois, M., Cuena-Lombraña, A., Fenu, G. & Bacchetta, G. 2018. Using species distribution models at local scale to guide the search of poorly known species: Review, methodological issues and future directions. Ecol. Model. 385: 124–132.

García-Llamas, P., Suárez-Seoane, S., Fernández-Guisuraga, J.M., Fernández-García, V., Fernández-Manso, A., Quintano, C., Taboada, A., Marcos, E. & Calvo, L. 2019. Evaluation and comparison of Landsat 8, Sentinel-2 and Deimos-1 remote sensing indices for assessing burn severity in Mediterranean fire-prone ecosystems. Int. J. Appl. Earth Obs. 80: 137–144. Doi: 10.1016/j.jag.2019.04.006

Gonçalves, J., Alves, P., Pôças, I., Marcos, B., Sousa-Silva, R., Lomba, Â. & Honrado, J.P. 2016. Exploring the spatiotemporal dynamics of habitat suitability to improve conservation management of a vulnerable plant species. Biodivers. Conserv. 25: 2867–2888.

Hagolle, O., Huc, M., Villa Pascual, D. & Dedieu, G. 2015. A Multi-Temporal and Multi-Spectral Method to Estimate Aerosol Optical Thickness over Land, for the Atmospheric Correction of FormoSat-2, LandSat, VENμS and Sentinel-2 Images. Remote Sens.-Basel 7: 2668–2691.

Halladin-Dąbrowska, A., Kania, A. & Kopeć, D. 2020. The t-SNE Algorithm as a Tool to Improve the Quality of Reference Data Used in Accurate Mapping of Heterogeneous Non-Forest Vegetation. Remote Sens.-Basel 12, 39.

Jeanmonod, D. & Gamisans, J. 2013. Flora Corsica, 2nd rev. and augmented edition. Bull. Soc. Bot. Cent.-Ouest 39: 1–1072.

Kallimanis, A.S., Touloumis, K., Tzanopoulos, J., Mazaris, A.D., Apostolopoulou, E., Stefanidou, S., Scott, A.V., Potts, S.G. & Pantis, J.D. 2015. Vegetation coverage change in the EU: patterns inside and outside Natura 2000 protected areas. Biodivers. Conserv. 24: 579–591.

Kuhn, M. 2008. Caret package. J. Stat. Softw. 28: 1–26.

Kuhn, M. & Johnson, K. 2013. Applied predictive modeling. Springer, New-York.

MacKay, H., Finlayson, C.M., Fernandez-Prieto, D., Davidson, N., Pritchard, D. & Rebelo, L.-M. 2009. The role of Earth Observation (EO) technologies in supporting implementation of the Ramsar Convention on Wetlands. J. Environ. Manage. 90: 2234–2242.

Mallinis, G., Emmanoloudis, D., Giannakopoulos, V., Maris, F. & Koutsias, N. 2011. Mapping and interpreting historical land cover/land use changes in a Natura 2000 site using earth observational data: The case of Nestos delta, Greece. Appl. Geogr. 31: 312–320. https://doi.org/10.1016/j.apgeog.2010.07.002

Maxwell, A.E., Warner, T.A. & Fang, F. 2018. Implementation of machine-learning classification in remote sensing: an applied review. Int. J. Remote Sens. 39: 2784–2817.

Picoli, M.C.A., Camara, G., Sanches, I., Simões, R., Carvalho, A., Maciel, A., Coutinho, A., Esquerdo, J., Antunes, J. & Begotti, R.A. 2018. Big earth observation time series analysis for monitoring Brazilian agriculture. ISPRS J. Photogramm. Remote Sens. 145: 328–339.

Raab, C., Stroh, H.G., Tonn, B., Meißner, M., Rohwer, N., Balkenhol, N. & Isselstein, J. 2018. Mapping semi-natural grassland communities using multi-temporal RapidEye remote sensing data. Int. J. Remote Sens. 39: 5638–5659.

Rapinel, S., Mony, C., Lecoq, L., Clément, B., Thomas, A. & Hubert-Moy, L. 2019. Evaluation of Sentinel-2 time-series for mapping floodplain grassland plant communities. Remote Sens. Environ. 223: 115–129.

Reymann, J., Panaïotis, C., Bioret, F., Bacchetta, G., Delage, A., Delbosc, P., Gamisans, J., Gauberville, C., Hugot, L. & O'Deye-Guizien, K. 2017. Prodrome des végétations de Corse. Doc. Phytosociol. 3: 1–175.

Richards, J.A. 2013. Remote Sensing Digital Image Analysis, 5th ed. Springer-Verlag, Berlin Heidelberg. Doi: 10.1007/978-3-642-30062-2

Rodwell, J.S., Evans, D. & Schaminée, J.H.J. 2018. Phytosociological relationships in European Union policy-related habitat classifications. Rend. Lincei-Sci. Fis. 29: 237–249.

Schauman, S., Verger, A., Filella, I. & Peñuelas, J. 2018. Characterisation of Functional-Trait Dynamics at High Spatial Resolution in a Mediterranean Forest from Sentinel-2 and Ground-Truth Data. Remote Sens.-Basel 10, 1874.

Schmidtlein, S., Zimmermann, P., Schüpferling, R., & Weiß, C. 2007. Mapping the floristic continuum: Ordination space position estimated from imaging spectroscopy. J. Veg. Sci. 18: 131–140. Doi: 10.1111/j.1654-1103.2007.tb02523.x

Schrodt, F., Santos, M.J., Bailey, J.J., & Field, R. 2019. Challenges and opportunities for biogeography-What can we still learn from von Humboldt? J. Biogeogr. 46: 1631–1642. Doi: 10.1111/jbi.13616

Schuster, C., Schmidt, T., Conrad, C., Kleinschmit, B. & Förster, M. 2015. Grassland habitat mapping by intra-annual time series analysis – Comparison of RapidEye and TerraSAR-X satellite data. Int. J. Appl. Earth Obs. 34: 25–34.

Silva, V., Catry, F.X., Fernandes, P.M., Rego, F.C., Paes, P., Nunes, L., Caperta, A.D., Sérgio, C. & Bugalho, M.N. 2019. Effects of grazing on plant composition, conservation status and ecosystem services of Natura 2000 shrub-grassland habitat types. Biodivers. Conserv. 28(5): 1205–1224.

Tomaselli, V., Adamo, M., Veronico, G., Sciandrello, S., Tarantino, C., Dimopoulos, P., Medagli, P., Nagendra, H. & Blonda, P. 2016. Definition and application of expert knowledge on vegetation pattern, phenology, and seasonality for habitat mapping, as exemplified in a Mediterranean coastal site. Plant Biosyst. 151(5): 887–899.

Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A.K., Szantoi, Z., Buchanan, G., Dech, S., Dwyer, J., Herold, M., Koh, L.P., Leimgruber, P., Taubenboeck, H., Wegmann, M., Wikelski, M. & Woodcock, C. 2015. Free and open-access satellite data are key to biodiversity conservation. Biol. Conserv. 182: 173–176.

Ullerud, H.A., Bryn, A., Halvorsen, R. & Hemsing, L. Ø. 2018. Consistency in land-cover mapping: Influence of field workers, spatial scale and classification system. Appl. Veg. Sci. 21: 278–288.

Vrieling, A., Meroni, M., Darvishzadeh, R., Skidmore, A.K., Wang, T., Zurita-Milla, R., Oosterbeek, K., O'Connor, B. & Paganini, M. 2018. Vegetation phenology from Sentinel-2 and field cameras for a Dutch barrier island. Remote Sens. Environ. 215: 517–529.

Wang, R. & Gamon, J.A. 2019. Remote sensing of terrestrial plant biodiversity. Remote Sens. Environ. 231: 111218. Doi: 10.1016/j.rse.2019.111218

Zelený, D. & Tichý, L. 2009. Linking JUICE and R: New developments in visualization of unconstrained ordination analysis. 18th Workshop of Eur. Veg. Survey. Rome.

Zlinszky, A., Schroiff, A., Kania, A., Deák, B., Mücke, W., Vári, Á., Székely, B. & Pfeifer, N. 2014. Categorizing Grassland Vegetation with Full-Waveform Airborne Laser Scanning: A Feasibility Study for Detecting Natura 2000 Habitat Types. Remote Sens.-Basel 6: 8056–8087.

Websites

Bivand, R., Keitt, T. & Rowlingson, B. 2015. Rgdal: Bindings for the Geospatial Data Abstraction Library. Retrieved from http://CRAN.R-project.org/package=rgdal

Hijmans, R.J. 2015. Raster: Geographic Data Analysis and Modeling. Retrieved from http://CRAN.R-project.org/package=raster

R Core Team. 2015. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/