ISPRS Journal of Photogrammetry and Remote Sensing 201 (2023) 54–66

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Detecting abandoned citrus crops using Sentinel-2 time series. A case study in the Comunitat Valenciana region (Spain)

Sergio Morell-Monzó [a,*], María-Teresa Sebastiá-Frasquet [b], Javier Estornell [c], Enrique Moltó [d]

[a] Universitat Politècnica de València, EPS Gandia, C/Paranimf, 1, 46730 Gandia, Spain
[b] Instituto de Investigación para la Gestión Integrada de Zonas Costeras, Universitat Politècnica de València, C/Paranimf, 1, 46730 Grau de Gandia, Spain
[c] Geo-Environmental Cartography and Remote Sensing Group, Universitat Politècnica de València, Camí de Vera s/n, 46022 Valencia, Spain
[d] Instituto Valenciano de Investigaciones Agrarias, 46113 Moncada, Spain

ARTICLE INFO

Keywords:
Sentinel-2
Time series
Crop monitoring
Agricultural land abandonment
Perennial crops
Citrus crops

ABSTRACT

Agricultural land abandonment (ALA) is becoming a growing phenomenon around the world that needs to be monitored and quantified. A massive abandonment of citrus orchards has been experienced in the last years in the Comunitat Valenciana (CV) region (Spain) driven by different socio-economic factors. Therefore, developing time and cost-efficient methods for monitoring ALA is a priority. Citrus are a perennial crop trees which make orchards have low spectral variation during the year. In the CV region, they are planted in relatively small parcels, thus creating a highly fragmented and heterogeneous landscape. This study proposes a machine learning-based classification framework that uses annual time series of spectral indices extracted from Sentinel-2 images to identify crop status at parcel level. The method is based on features extracted from the reconstructed OSAVI and NDMI time series used to train a Random Forest classifier. Then, a parcel-based classification is performed using the parcel boundaries and the probabilities of belonging to each category for the full pixels inside the boundaries. The research assessed the potential to identify three statuses of crops (non-productive, productive, and abandoned). Results on three different temporal and spatial datasets provided an overall accuracy ranging from 89 to 92 %, demonstrating the importance of multi-temporal data to identify the abandonment of perennial crops. Furthermore, we studied the ability of the model to be spatially and temporally transferred. Limitations to recall the abandoned parcels when using models trained in other areas or time periods are exposed, opening the way to model improvements.

1. Introduction

Agricultural land abandonment (ALA) is a growing phenomenon around the world that has received attention since the early 1990s. However, around 95 % of the papers were published in the last 15 years (Subedi et al., 2022). This indicates a growing concern in the scientific community. Agriculture represents more than half of the territory of the European Union (EU). Agriculture guarantees food supply, manages important natural resources, and supports the socio-economic development of rural areas (Terres et al., 2015). However, it has been estimated around 11 % of the agricultural area is at high risk of abandonment for the period 2015–2030 (Perpiña Castillo et al., 2018). ALA is a major obstacle to the sustainable development of many European regions (Leal Filho et al., 2016). Nevertheless, it will increase in Europe over the next decades (Terres et al., 2015). In this context, new time and cost-efficient methods to identify ALA are required to update inventories of land use, a first step to improve land management strategies. Temporal and spatial explicit information about ALA can help to support policy instruments to counteract or reverse the process and can also help implement adequate landscape management measures (Volpi et al., 2023).

The development of the Copernicus program by the European Space Agency (ESA) and the European Commission has contributed to great advances in remote sensing systems that have significantly improved agricultural monitoring (Phiri et al., 2020). ALA is particularly frequent in both mountainous areas and in highly fragmented landscapes (Czesak et al., 2021). In such areas, the spatial resolution of other commonly used sensors, such as MODIS (Moderate Resolution Imaging

* Corresponding author.
E-mail addresses: sermomon@upv.es (S. Morell-Monzó), mtsebastia@hma.upv.es (M.-T. Sebastiá-Frasquet), jaescre@cgf.upv.es (J. Estornell), molto_enr@gva.es (E. Moltó).

https://doi.org/10.1016/j.isprsjprs.2023.05.003
Received 29 November 2022; Received in revised form 15 April 2023; Accepted 4 May 2023
Available online 24 May 2023

Spectroradiometer) (500 m) or Landsat (30 m) is not enough to precisely monitor the small-sized parcels. The higher spatial resolution of Sentinel-2 (10–20 m) may improve these capabilities, which is particularly interesting in the Comunitat Valenciana (CV), where most citrus orchards are between 0.25 and 0.50 ha, especially in coastal areas. Morell-Monzó et al. (2020) used Sentinel-2 spectral single scene data in the first experience to identify the abandonment status of citrus crops in this region, obtaining an overall accuracy of around 77 %.

The information in the spectral domain of the Sentinel-2 data provides great opportunities to discriminate between types of crops with different spectral signatures (Asgarian et al., 2016). In parallel, information in the temporal domain is also important because it allows tracking the evolution of a crop along the seasons. The high temporal resolution of Sentinel-2 has been an opportunity to improve the identification and mapping of crop types and crop rotations (Vuolo et al., 2018). The inherently dynamic nature of crops and their spectral-temporal signatures make time series approaches particularly convenient (Gómez et al., 2016). However, the benefits of multi-temporal data for detecting the status of permanent evergreen crops, such as citrus, may be not so obvious. Unlike seasonal crops, the growth of citrus to productive age lasts about 5 years, and reaching maximum growth may last 15 years. In consequence, there is normally a single plantation event in each orchard that is repeated after more than fifteen years. The growth of trees is relatively slow until it reaches a mature state, when the vegetative growth is controlled by pruning. Harvesting usually does not produce important reflectance changes as in annual or bi-annual crops. It is important to remark that citrus cultivated in Europe are irrigated and most farmers remove weeds. When a productive parcel is abandoned, the evolution of the trees and the growth of spontaneous vegetation will be totally conditioned by the soil, temperature and rainfall regimes since irrigation is cut and no weeding practices are made. For all these reasons orchard spectral signatures of abandoned orchards change little and gradually in time, which makes detecting the abandonment status of citrus parcels especially challenging.

Spain is the largest citrus producer in Europe, the fourth in the world, and the first exporter of citrus for fresh consumption. Comunitat Valenciana region accounts for more than 50 % of the Spanish citrus production. A massive abandonment of citrus orchards has been experienced from 2000 to 2020, representing a 20 % decrease of the cultivated surface (MAPA, 2021). In this region the current estimated area dedicated to citrus 160,088 ha (MAPA, 2022). In this context, it is of priority interest to develop time and cost-efficient methods to identify and quantify citrus crop abandonment. Abandonment of citrus crops in the CV region is attributed to the small size of the exploitation units, which hinders mechanization and subsequent reduction of costs, while prices do not increase because supply and demand are unbalanced (Garcia Álvarez-Coque and Moltó, 2020). Furthermore, the emergence of more profitable land uses (e.g. urban uses) which means that many owners keep their parcels without any agricultural interest in the hope of future capital gains from urban development processes (Morell-Monzó and Garófano-Gómez, 2022). Finally, there are weaknesses in the structure, organization and commercial quality of the Valencian supply chain which are aggravated by unfavorable policies for the Valencian citrus sector and institutional failures both in the negotiation phase of the agreements and in controls at the points of entry into the EU (Compés et al., 2019).

The main objective of this study is to propose a method to automatically identify the status of citrus crops using Sentinel-2 data. For this purpose, a machine learning-based classification framework was developed using time series of spectral indices derived from Sentinel-2 to identify the crop status at parcel level. A comprehensive evaluation of the model's performance and the transferability of spectral-temporal signatures across regions and years inside the Comunitat Valenciana region (Spain) was assessed.

2. Data and methods

2.1. Study area

Our study area is in La Safor, a coastal region of 43,000 ha, approximately centered at (34°56′14″N, 0°8′42″E) in CV region. It was selected due to its large tradition of citrus production. Its landscape is characterized by two clearly differentiated structural elements. On the west, a mountainous forest area. On the east a plain alluvial zone and the coastline, which is highly populated and where most of the irrigated agricultural production is located. This plain was generated by successive sedimentary systems of Quaternary origin. It stands out for its flat relief, soils with high agronomic potential and high availability of water that have made it a historically agricultural area. The major crops in this region are citrus representing 95 % of cultivated area (Generalitat Valenciana, 2020). However, in recent years a large part of the farms has been abandoned. The agrarian structure is characterized by the small size of the farms (between 0.25 and 0.5 ha) resulting in a highly fragmented landscape.

This region has a typical Mediterranean climate with two distinct seasons, wet and dry (Gasith and Resh, 1999; González-Hidalgo et al., 2005). The characteristic rainfall is scarce and torrential and has an extremely high spatial and temporal variability being the raining seasons autumn and spring. In contrast, summer (June to August) is the dry period in this Mediterranean zone.

2.2. Map classes and reference data

This work proposes a classification based on three orchard status: non-productive (NP), productive (PR), and abandoned (AB) (Morell-Monzó et al., 2021 for exhaustive details). NP class includes bare soil orchards before tree planting and recently planted trees that have not reached productive age and are predominantly occupied by bare or low vegetated soil. Citrus trees reach productive age around 5 years after planting. Therefore, this category contains trees less than 5 years old. PR class includes orchards at productive age (>5 years). Finally, class AB includes abandoned orchards with physically evident signs of abandonment.

Major changes that occur in an abandoned orchard are the growth of spontaneous vegetation, loss of vigor of the trees, and, therefore, progressive loss of the typical planting patterns of fruit crops. However, growers may carry out management practices that generate different evolutions on abandoned orchards. In some cases, owners cut the tree canopies, then sprouts and small branches may appear. In other cases, owners carry out periodic pruning or weeding, limiting the height and density of spontaneous vegetation. In general, the lack of water leads to the progressive death of the trees at different rates.

Identifying AB parcels is essential for monitoring citrus land abandonment and differentiating PR from NP parcels is interesting to improve citrus yield estimations. Declared citrus parcels are georeferenced in the Land Parcel Identification System of Spain known as *Sistema de Información Geográfica de Parcelas Agrícolas* (SIGPAC). However, this database has no up-to-date information on crop status. The proposed classification framework was designed to complement the SIGPAC information and be applied to previously identified citrus parcels (see Fig. 1).

Seven ground truth datasets were obtained through: a) field inspection campaigns, b) images captured by unmanned aerial vehicles (UAV), and c) very-high resolution orthophoto photointerpretation. Field campaigns and UAV data were obtained in previous published studies (Morell-Monzó et al., 2020; Morell-Monzó et al., 2021), and for this research we enlarged the dataset with parcels classified by photointerpretation of very-high resolution orthophotos. Table 1 summarizes these ground truth sources.

The first dataset (OLV-19) consists of 240 parcels classified during a field campaign carried out between July 11 and 14, 2019 in the

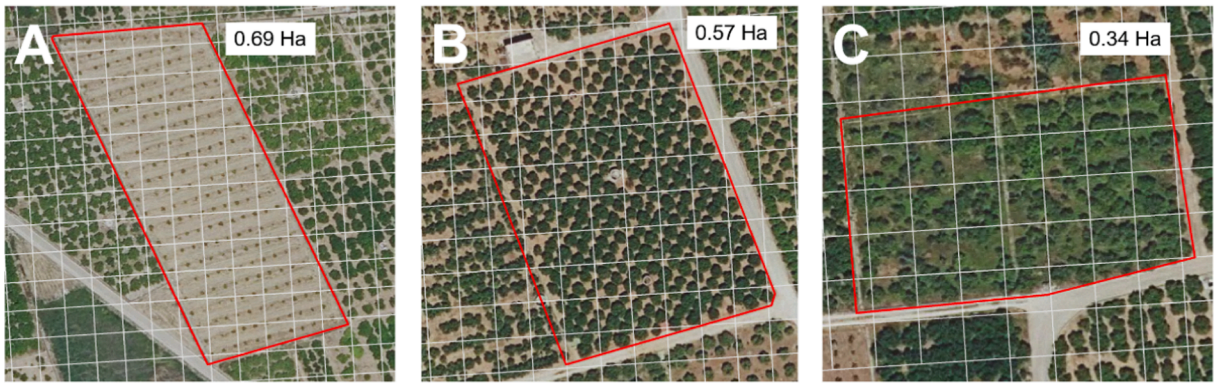

Fig. 1. Examples of the three types of crop status: non-productive (left), productive (center), and abandoned (right) parcels. The grey grid corresponds to the 10-meter Sentinel-2 pixels.

Table 1
Summary and characteristics of the ground truth data used.

Dataset	Number of parcels	Sampling method	Acquisition date (verification)	Sampled area (ha)	Average parcel size (ha)	Parcel size range (ha)	NP parcels ratio	PR parcels ratio	AB parcels ratio
OLV-19	240	Field campaigns	July 11–14, 2019 (May 14, 2019)	100.5	0.41	0.10–2.70	33.3 %	33.3 %	33.3 %
LSF-20	280	UAV photointerpretation	February 1–29, 2020 (May 6, 2020)	136.7	0.48	0.08–3.23	21.4 %	42.9 %	35.7 %
LSF-21	836	Field campaigns	March 9, 2021 (May 16, 2021)	368.2	0.44	0.07–7.32	22.5 %	44.0 %	33.5 %
PV-21	80	Photointerpretation	May 16, 2021	34.3	0.36	0.09–2.66	21.3 %	45.0 %	33.7 %
TV-21	80	Photointerpretation	May 15, 2021	36.6	0.46	0.07–1.90	12.6 %	63.7 %	23.7 %
BP-21	80	Photointerpretation	May 13, 2021	65.9	0.86	0.89–3.87	16.3 %	55.7 %	25.0 %
NL-21	80	Photointerpretation	July 22, 2021	39.5	0.49	0.12–1.84	20.0 %	55.1 %	24.9 %

municipality of Oliva (Fig. 2b). Parcels have an average size of 0.41 ha, summing up 100.5 ha representing a sampling effort of 0.85 % of the total area dedicated to citrus in the region of La Safor. They were selected through stratified random sampling with equal proportion for each category. This dataset contains 80 NP, 80 PR, and 80 AB parcels.

The second dataset (LSF-20) consists of 280 parcels classified by visual interpretation of 12 UAV photogrammetric flights carried out in February 2021 (Fig. 2c). Parcels were selected through a systematic and intensive sampling of the area covered by the UAV. Therefore, these samples were obtained through cluster sampling from the 12 areas overflown by the UAV. This dataset contains 60 NP (21.4 %), 120 PR (42.9 %), and 100 AB parcels (35.7 %). The parcels have an average size of 0.48 ha, summing up 136.7 ha representing a sampling effort of 1.15 % of the total area dedicated to citrus in the region of La Safor.

The third dataset (LSF-21) consists of 836 parcels classified during field campaigns carried out from March 9 to May 28, 2021 in the coastal area between Gandia and Oliva (Fig. 2d). Parcels have an average size of 0.44 ha, summing up 368.2 ha representing a sampling effort of 3.10 % of the total area dedicated to citrus in the region of La Safor. They were selected through a simple random sampling. This dataset contains 188 NP (22.5 %), 368 PR (44 %), and 280 AB parcels (33.5 %).

All three datasets were in the coastal area of La Safor, which is about 20 km long and 6 km wide, hereafter named the training area. To avoid possible errors caused by changes in land use between field campaigns and the acquisition date of the images, abandonment status was verified using high-resolution airborne images from the Valencian Cartographic Institute (ICV) https://idev.gva.es/va/, taken annually in May.

Additionally, 4 datasets were generated outside the training area (Fig. 2a). These datasets are located in the municipalities of Potries and Villalonga (PV-21), Tavernes de la Valldigna (TV-21), Benicull and Polinyà del Xuquer (BP-21), and Nules (NL-21). Datasets PV-1 and TV-21 are located in La Safor region, close to the training area, while datasets BP-21 and NL-21 are located far from the study area. Each of these datasets contains 80 parcels selected by simple random sampling and classified by photointerpretation of the 2021 orthophoto offered by ICV. Orthophotos from previous years and Street View images from Google Maps were also used for verification.

2.3. Sentinel-2 time series processing and feature extraction

Three annual collections of Sentinel-2A/B L2A bottom-of-atmosphere (BOA) reflectance images were acquired from the Google Earth Engine platform (Gorelick et al., 2017) generating annual data time series: a) from September 1st, 2018 to September 30th, 2019, representing the 2019 season, b) from September 1st, 2019 to September 30th, 2020, representing the 2020 season and c) from September 1st, 2020 to September 30th, 2021, representing the 2021 season. The time series consist of 65, 57 and 54 images, respectively. In our study area the revisit time of the Sentinel-2 constellation is 5 days. Images were projected in UTM/WGS84 30N and they were filtered to include only those with less than 75 % cloud percentage. The L2A product provides a scene classification band (SC). Pixels classified as no data, saturated or defective, cast shadows, cloud shadows, unclassified, cloud medium and high probability, and thin cirrus in the SC band were masked in all reflectance bands. These bands were resampled to 10 m resolution by pixel disaggregation of the lower resolution bands. Finally, two spectral indices were calculated using the bands B4, B8, and B11: the Optimized Soil Adjusted Vegetation Index (OSAVI) (Rondeaux et al., 1996) and the Normalized Difference Moisture Index (NDMI) (Gao, 1996). These spectral indices were selected based on our previous work (Morell-Monzó et al., 2022) where they obtained the best discriminating results among a total of 15 indices.

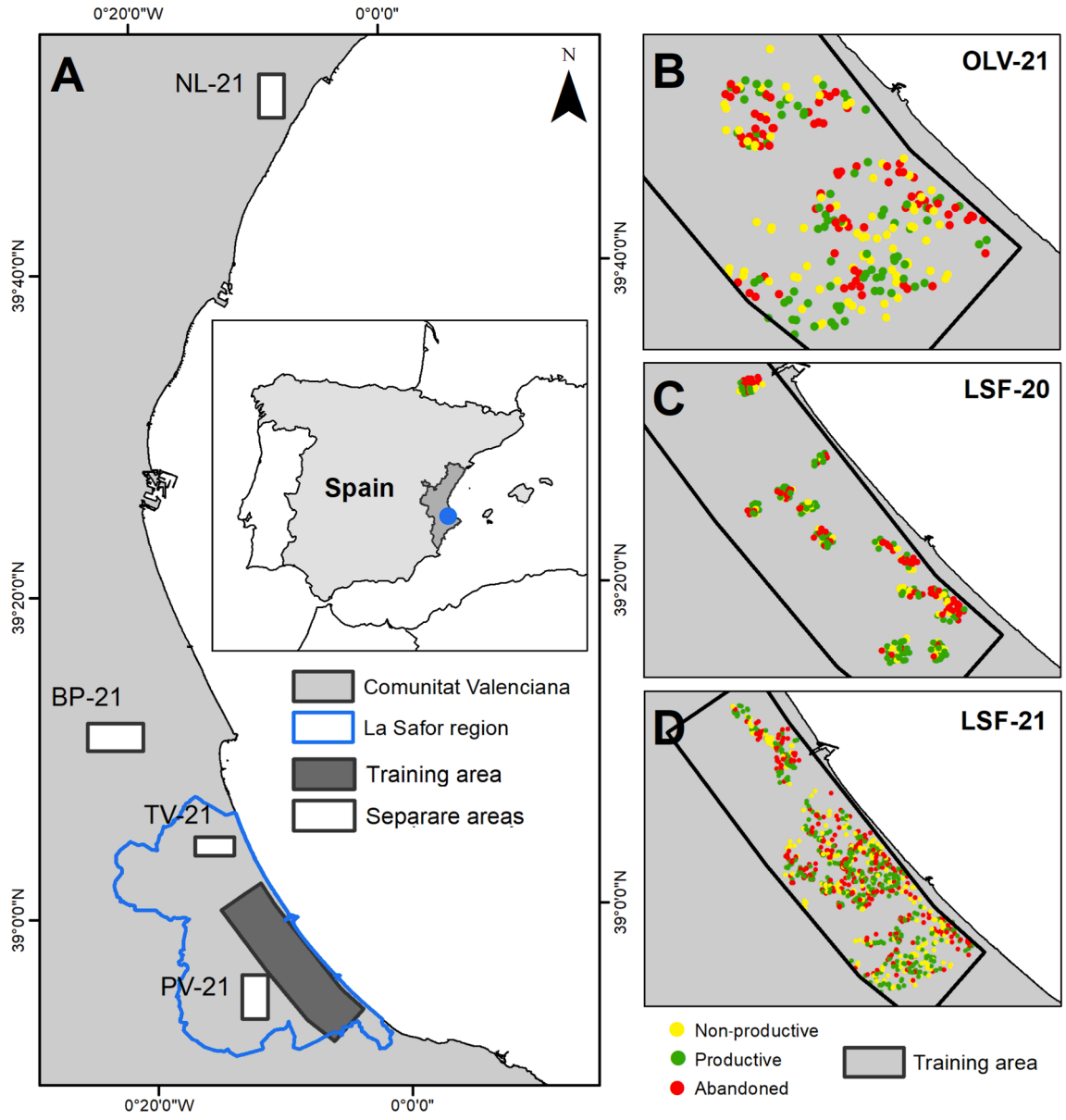

Fig. 2. Study area and ground truth datasets. (A) Study area in La Safor, Comunitat Valenciana region (Spain) and ground truth datasets. (B) OLV-19 dataset. (C) LSF-20 dataset. (D) LSF-21 dataset.

For Sentinel-2 images, OSAVI is calculated from the near-infrared band (B8 ~ 0.83 μm) and red band (B4 ~ 0.66 μm) as shown in Equation (1). The OSAVI is an optimization of the Soil Adjusted Vegetation Index (SAVI) proposed by Huete (1988). OSAVI is a soil-adjusted vegetation index optimized for agricultural monitoring which uses a soil adjustment coefficient (L), to minimize variation produced by soil background (Steven, 1998). OSAVI uses a value $L = 0.16$ proposed by Rondeaux et al. (1996). This index was used because in citrus parcels a significant part of the area is not covered by the tree canopy and the soil is exposed, so the influence of the soil on the spectral response may be important.

$$OSAVI = (1+L)\frac{(B8 - B4)}{(B8 + B4 + L)} \tag{1}$$

For the Sentinel-2 images, NDMI is calculated from the near-infrared band (B8 ~ 0.83 μm) and short-wave infrared band (B11 ~ 1.61 μm) as shown in Equation (2). The near-infrared spectral channel collects the reflectance of the internal structure of the leaf and the dry matter content of the leaf. At this wavelength the absorption produced by liquid water in the soil and vegetation is insignificant. In contrast, the short-wave infrared spectral (SWIR) channel is sensitive to soil and vegetation water content and to the internal mesophyll structure of the leaf. As a result, the NDMI is an indicator of the moisture content of soil and vegetation. This index was chosen because it can provide differential information between the PR and AB parcels due to the differences in vegetation and soil moisture because of the absence of irrigation in the abandoned parcels (Morell-Monzó et al., 2023). Furthermore, the NDMI is less sensitive to atmospheric effects than other spectral indices such as the Normalized Difference Vegetation Index (NDVI) (Gao, 1996).

$$NDMI = \frac{(B8 - B11)}{(B8 + B11)} \tag{2}$$

Both OSAVI and NDMI time series contained temporal gaps due to the removal of images (those having more than 75 % cloud coverage) and spatial gaps due to masked pixels. In order to create continuous time series with non-empty values every 5 days, those gaps were filled by linear interpolation expressed in Equation (3).

$$Y_i = Y_{i-1} + (Y_{i+1} - Y_{i-1})^* \left(\frac{t_i - t_{i-1}}{t_{i+1} - t_{i-1}} \right) \quad (3)$$

where Y_i is the interpolated pixel value, Y_{i-1} is the pixel value of the previous image, Y_{i+1} is the pixel value of the next image, t_i is the time value of the image containing the pixel to interpolate, and t_{i-1} and t_{i+1} are the time values of the previous and next images.

To reduce noise in the time series introduced by the undetected clouds, partial cloud shading or unfavorable atmospheric conditions, the original OSAVI and NDMI signals were smoothed using the Savitzky-Golay filter (Savitzky and Golay, 1964; Chen et al., 2004). This is a moving window filter in the time domain that removes outliers and smoothes data in the time series while preserving the trends. It is based on a simplified least square fitting convolution to smooth and compute derivatives of a set of consecutive values. The convolution can be understood as a weighted moving average filter, which uses weights given by a polynomial of a user-defined degree. When these weights are applied to the signal, they perform a least squares polynomial fit within the filter window (Chen et al., 2004). In this study, a moving window of 50 days and a third-order polynomial were used.

Features were extracted from the smoothed signals to be used as inputs to train a supervised classifier. They can be categorized as: a) central tendency statistics: mean (12 per year); b) dispersion statistics: range (2 per year) and standard deviation (2 per year); c) distribution shape statistics: skewness (1 per year) and kurtosis (1 per year) of each spectral index, providing a total of 36 features per pixel of the smoothed time series. The length of the intervals used for feature extraction was optimized through a search grid with values of 12, 6, 4, 3, 2, and 1 month. Finally, monthly intervals were used for the central tendency statistics, semi-annual intervals for the dispersion statistics, and annual intervals for the distribution shape statistics (Fig. 3).

2.4. Parcel-based classification

A parcel-based classification method was developed since this approach allows spatial indexing with agricultural databases (e.g., SIGPAC). The proposed method is described below:

1. Using the extracted features, a supervised classification model is trained using pixels as training samples. Only pixels completely covered by the parcels are used for training.
2. Then, the fitted model is used to predict the probability of each pixel being in each class (Fig. 4). Therefore, for each pixel of the image, a probability vector with length C is predicted, being C the number of classes. In this study the number of classes is 3.
3. Finally, parcel-based classification is performed. For each parcel, the average probability of each class is calculated using the full pixels inside the parcel boundaries (Fig. 4). Full pixels are those that have a coverage threshold of 100 %, however, the method iteratively reduces this threshold to 75 % and 50 % in case there are no full pixels inside the parcel boundaries. Coverage percentages were measured using the *exactextractr* R package. Finally, each parcel is assigned the class with the maximum average probability.

This novel approach was designed to deal with the characteristics of a highly fragmented agricultural landscape and differs from the traditional parcel-based classification strategy by applying majority voting. It should be outlined that in majority voting, a classification value is predicted for each pixel and then parcel is classified by applying the mode operator using the pixels whose centroid is inside the parcel. In the proposed method, the class probability is predicted for each pixel. Then, the probability values for each class are averaged using the pixels completely inside the parcel. Finally, each parcel is classified according to the class with the maximum probability, thus applying the argument of the maxima (argmax). In practice, the outcome of the proposed method is similar to the majority voting method. However, the use of full pixels is especially interesting in highly fragmented agricultural structures since it allows for avoiding edge effects. Additionally, this method returns the average probability of each class per parcel. This data, combined with the number of complete pixels in the parcel, can be employed to automatically filter parcels for manual class re-assignment, removal, or further specific processing.

The Random Forest (RF) method (Breiman, 2001) was used to build the pixel classifier. Although there is a growing trend toward the use of deep learning algorithms, RF remains as a benchmark algorithm for many remote sensing classification problems (Sheykhmousa et al., 2020). This is due to its good performance with small amounts of training data and its robustness against overfitting. RF requires tuning two key hyperparameters: the number of trees (*ntree*) that make up the forest and the number of randomly selected variables at each node split (*mtry*). RF models was trained with an increasing *ntree* up to 150 trees and the hyperparameter *mtry* was set to 6, corresponding to the square root of the number of variables. The out-of-bag samples were used to measure model convergence. To perform the inference, the RF must predict the probability of belonging to each class. This is achieved by calculating the proportion of trees voting for each class. In this work, the RF algorithm was applied from the *randomForest* R package (Liaw and Wiener, 2002).

2.5. Accuracy assessment

One of the greatest challenges in many machine learning applications is to build accurate and robust models that reduce recalibration efforts and ensure model transferability to never seen data distributions (Dimov, 2022). Three models for citrus crop status identification were trained using the OLV-19, LSF-20, and LSF-21 datasets. These models were validated using two cross-validation strategies. First, 4-fold cross-validation was performed using random splits without replacement. In each iteration, 3/4 of the parcels were used for training and 1/4 of the parcels were used for validation, ensuring balanced data sets with the same number of parcels in each class. Second, 4-fold spatial cross-validation was performed. This strategy was used to avoid an underestimation of the model error due to spatial autocorrelation of the samples, which is inherent in remote sensing data (Karasiak et al., 2021).

Random cross-validation is suitable when the data is randomly distributed. However, when the data is strongly aggregated, this strategy may underestimate the model error (Wadoux et al., 2021; Stock and Subramaniam, 2022). In our case, the LSF-20 dataset is clustered due to the sampling method through UAV flights. The other datasets were obtained through random sampling. However, the agricultural areas are

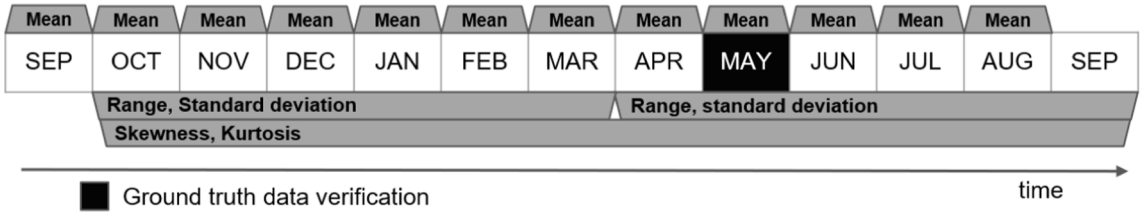

Fig. 3. Feature extraction by time intervals from Sentinel-2 OSAVI and NDMI smoothed time series.

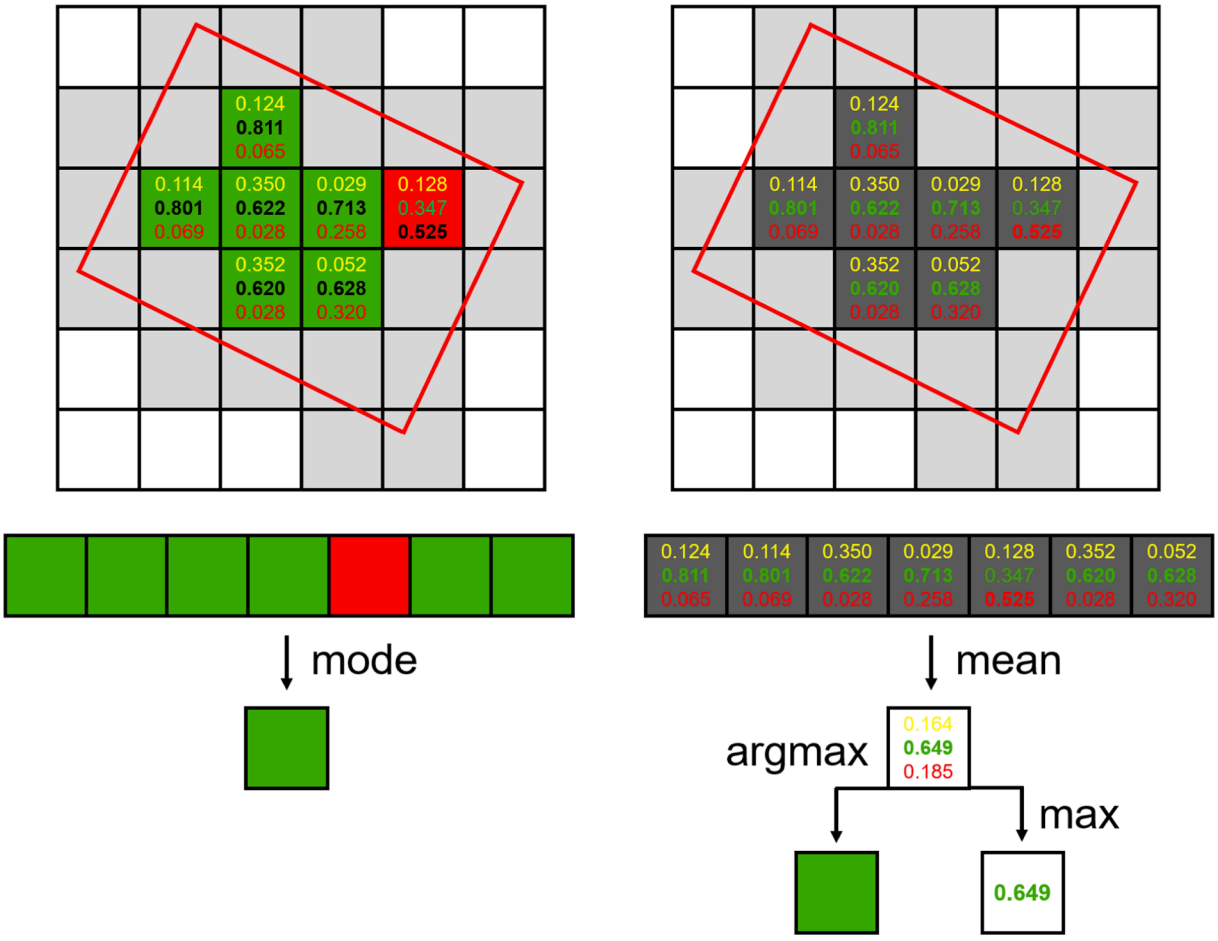

Fig. 4. Differences between traditional classification based on majority voting (left) and the proposed method based on maximum averaged probability (right).

not randomly distributed in the area, which causes a given degree of samples aggregation. These circumstances motivated the use of spatial cross-validation. Spatial partitions were generated using the k-Means method on the XY coordinates of the parcels. This procedure allows for generating data partitions with the maximum distance between groups, although in this case it cannot be ensured that the data is balanced.

To assess the model performance the following accuracy metrics were calculated: overall accuracy (OA) and its 95 % confidence intervals, precision (producer's accuracy by class), recall (user's accuracy by class) (Olofsson et al., 2014) and Cohen's kappa index (Cohen, 1960). Due to the small size of the agricultural parcels, relationships between model accuracy and parcel size were explored using the Kruskall-Wallis non-parametric test.

One of the biggest challenges for the operational implementation of a remote sensing-based monitoring system is to obtain robust spatial and temporal transferability. For this reason, the spatial and temporal transferability of the models was additionally evaluated with the aim of knowing the generalization capacity outside the training area and in years different from the training one. To assess spatial transferability, data from separate areas (PV-21, TV-21, BP-21 and, NL-21) were used. The model trained in the study area with LSF-21 data was used to evaluate its performance outside the study area without using additional data to calibrate the model. To evaluate the temporal transferability of the models, cross-validation of the models trained in the years 2019, 2020, and 2021 was carried out. Each model was evaluated in the two years other than the training year without using additional training samples.

2.6. *Influence of features on class separability*

The Jeffries-Matusita (JM) distance measures the separability between a pair of probability distributions. This method does not inform about the influence of a descriptor on the model performance, but rather collects the intrinsic structure of the variables. Unlike methods based on permutation or impurity (such as measures of the importance of RF features), this statistical method is not biased by collinearity or cardinality. The JM distance is widely used as a separability criterion for evaluation of classification results. In this study the JM distance defined in Richards and Jia (2006) was used, so it takes the range [0,2]. Class separability of features based on JM distances can be illustrated in radar charts.

3. Results

3.1. *Spectral-temporal signature*

This section shows the temporal behavior of the spectral indices for the three types of parcels studied. Fig. 5 shows the average time series of OSAVI and NDMI for each crop status. Regarding OSAVI, in average, NP parcels had lower values than PR and AB parcels along the three years of study, being easily separated. A periodical behavior is observed in the NP profile with a maximum between December–February and a minimum between June–August. This cyclical behavior corresponds to the growth of spontaneous vegetation in the cold and wet winter and its wilt in the hot and dry summer. The effect of spontaneous vegetation cycles is more evident in the NP parcels than in the others, since planted tree canopies represent only a small part of the surface. In contrast, PR and

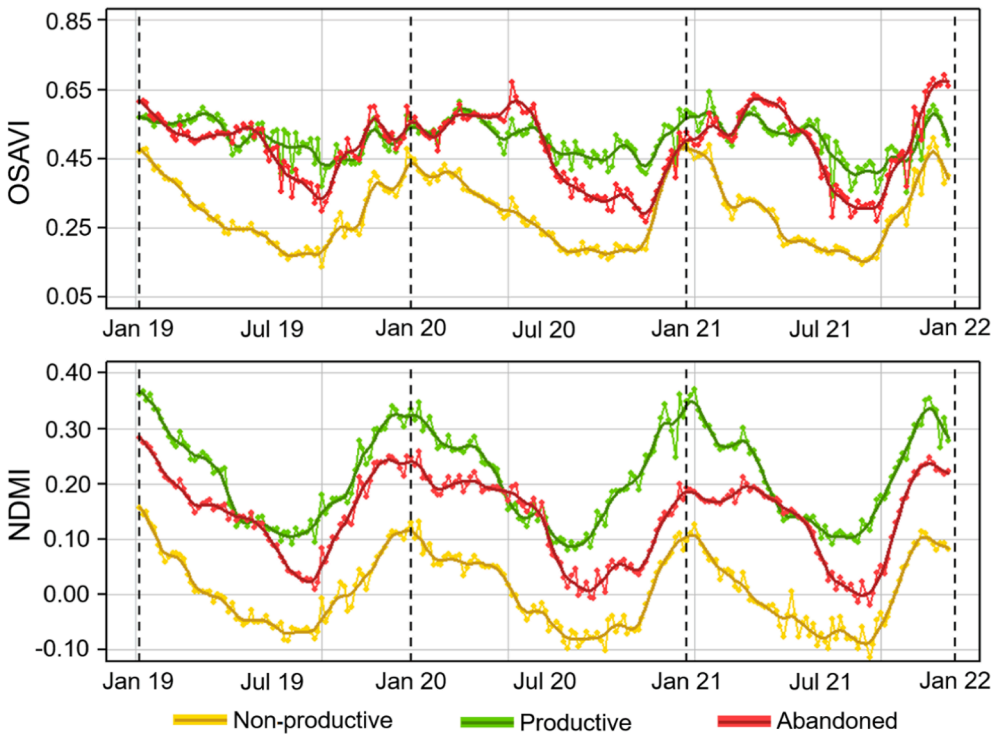

Fig. 5. Average temporal profile of OSAVI and NDMI for the three classes: non-productive, productive and abandoned. Dots joined by light color lines represent original values, dark color lines are the smoothed time series after applying the Savitzky-Golay filter. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

AB parcels had higher OSAVI values along the whole period of study and showed much lower separability. The OSAVI profile of the PR and AB classes also showed periodic but less defined behavior. Major differences between PR and AB were found in the three summers (June–August). In these seasons, AB parcels reduced OSAVI levels more than PR parcels.

Similar patterns are observed from NDMI signals, with maximum values in winter and lower values in summer. NP parcels have lower NDMI values than PR and AB parcels, showing a higher separability with PR and AB classes. NDMI differences between PR and AB parcels are greater than OSAVI. Based on these results, the major challenge is to separate the PR and AB classes. In this sense, NDMI produces greater differences between these classes. However, NDMI differences are severely reduced in May–June.

Fig. 6 shows OSAVI evolution of three sample parcels from 2018 to 2022. PR and AB parcels have higher OSAVI values (between 0.65 and 0.30). In contrast, the NP parcel, having a limited tree canopy and large areas of bare soil during most of the year, had lower values (between 0.25 and 0.05). The PR parcel has higher OSAVI values than the AB parcel. In addition, the PR parcel is more stable over time. Moreover, the effect of changing the status from AB to NP can be observed at Fig. 6a. Figure also shows the effect of the Savitzky-Golay filter. The filter smooths the original signal and reduces noise (sudden lower/higher OSAVI values), mainly caused by cloudy pixels not detected by the SC band of the Sentinel-2 L2A product.

3.2. Random Forest model

All RF models built during cross-validation converged before 150 decision trees (between 50 and 100 trees). A larger number of trees did not result in a reduction of OOB-error. The average OOB-error was reduced to 2.0, 2.1 and 3.7 % for the OLV-19, LSF-20 and LSF-21 datasets, respectively. Therefore, a higher OOB-error was obtained as the number of parcels in the dataset increased. In all three datasets, the NP class produced the lowest error (0.8–3.2 %), followed by the PR class (1.4–2.3 %), and the AB class (2.9–7.0 %).

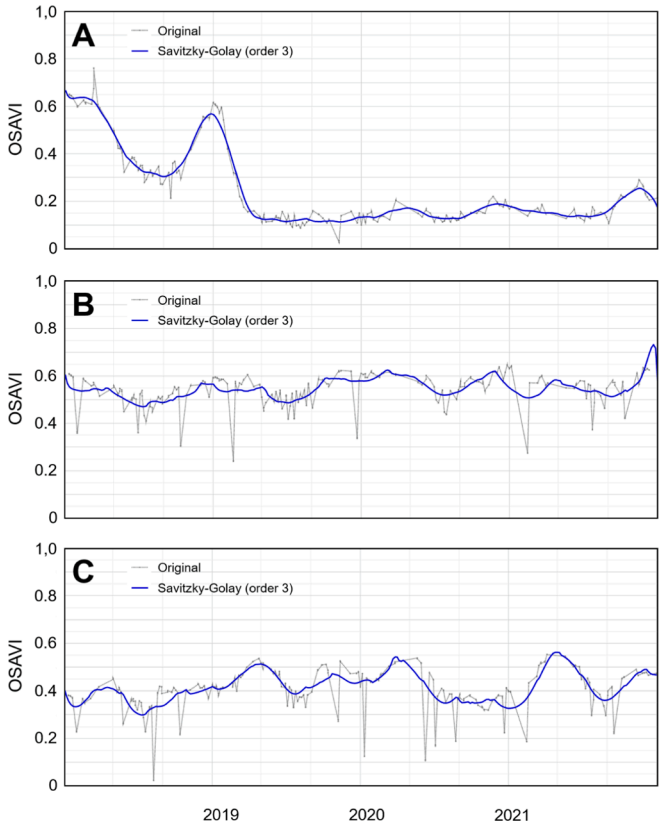

Fig. 6. OSAVI from 2018 to 2022 of three sample parcels: (A) non-productive parcel replanted at the beginning of 2019, (B) productive parcel, (C) abandoned parcel.

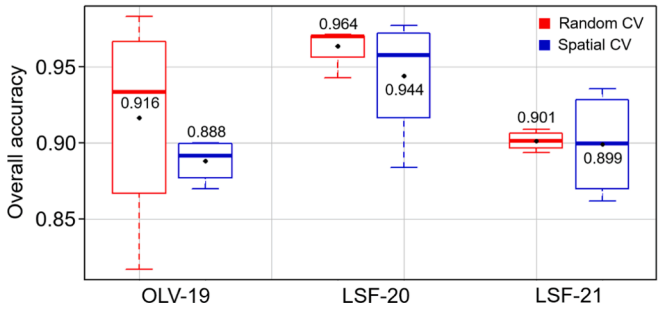

Fig. 7. Overall accuracy of the models obtained using 4-fold cross-validation for the OLV-19, LSF-20 and LSF-21 datasets. In red are shown the results obtained using the random cross-validation strategy and in blue those obtained using the spatial cross-validation strategy. Dots show the average overall accuracy. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Accuracy assessment

This section evaluates the performance of the classification models. Fig. 7 shows a summary of the OA obtained with each dataset and each validation strategy. Considering all the experiments, the average OA ranged from 0.89 to 0.96, with an average of 0.92. For the OLV-19 dataset, the OA obtained through random cross-validation was 0.92 ± 0.03 while through spatial cross-validation was 0.89 ± 0.04. For the LSF-20 dataset, the OA obtained through random cross-validation was 0.96 ± 0.02 while through spatial cross-validation was 0.94 ± 0.03. Finally, for the LSF-21 dataset, the OA obtained through random cross-validation was 0.90 ± 0.02 while through spatial cross-validation was 0.89 ± 0.02. All the above errors expressing a 95 % confidence interval.

OA was higher for the LSF-20 dataset. This fact agrees with our previous experiences (Morell-Monzó et al., 2023) and suggest that this dataset is less demanding than the rest. It can be attributed to the fact orchards were selected systematically, thus being types of abandonment more evident types of abandonment or a more advanced process of abandonment that made them easily detectable. Consequently, we propose the more conservative OA values (0.89–0.92). Additionally, small (up to 3 %) OA differences between random cross-validation and spatial cross-validation, were found. These differences were smaller in the larger dataset (LSF-21).

Below, the model performance by classes is studied. Table 2 shows precision and recall by classes of random and spatial cross-validation on the 3 datasets. The following average class accuracy was obtained: NP = 0.935 > PR = 0.921 > AB = 0.899. NP class was the class detected with more accuracy, while AB was the most difficult to detect. On the other hand, recall values were in general higher than 0.900, except for the AB class both in the OLV-19 and LSF-21 datasets. This last dataset produced an omission error of around 18 % for the AB class, which was mostly misclassified as PR. These results suggest a limitation of the model to detect some AB parcels.

Regarding the effect of parcel size in the results, Fig. 8 shows the size of correctly classified and misclassified parcels. In all datasets, the model showed higher accuracy on larger parcels. However, Kruskal-Wallis's analysis of variance showed no significant differences between the size of correctly classified and misclassified parcels.

Table 2
Cross-validation results of the Random Forest model trained on the OLV-19, LSF-20, and LSF-21 datasets.

Dataset	Performance	Random cross-validation	Spatial cross-validation
OLV-19	Overall accuracy (95 % C.I.)	0.916 ± 0.034	0.888 ± 0.039
	Cohen's Kappa	0.875	0.827
	NP precision	0.942	0.932
	PR precision	0.878	0.857
	AB precision	0.932	0.883
	NP recall	0.962	0.942
	PR recall	0.912	0.916
	AB recall	0.875	0.805
LSF-20	Overall accuracy (95 % C.I.)	0.964 ± 0.022	0.944 ± 0.027
	Cohen's Kappa	0.944	0.909
	NP precision	0.983	0.985
	PR precision	0.945	0.918
	AB precision	0.980	0.956
	NP recall	0.900	0.897
	PR recall	0.991	0.972
	AB recall	0.970	0.926
LSF-21	Overall accuracy (95 % C.I.)	0.901 ± 0.020	0.899 ± 0.020
	Cohen's Kappa	0.846	0.841
	NP precision	0.915	0.915
	PR precision	0.883	0.890
	AB precision	0.919	0.915
	NP recall	0.919	0.934
	PR recall	0.953	0.941
	AB recall	0.820	0.818

3.4. Spatial and temporal transferability

To assess spatial transferability, the model trained with the LSF-21 dataset was used to classify citrus crop status in parcels of the four separate areas (PV-21, TV-21, BP-21, and NL-21) without using additional training. Table 3 shows the classification results. OA ranged from 0.87 to 0.92. Highest accuracies were obtained in the NL-21 (0.92 ± 0.06) and PV-21 (0.92 ± 0.06) datasets, while TV-21 and BP-21 had the lowest (0.88 ± 0.07 and 0.87 ± 0.07 respectively). Average OA outside the study area was 0.898, which is only 1 % below the classification within the training area. These results show a good transferability of the model within the studied territory. However, the recall of the AB class was highly compromised when transferring the model out of the study area (0.700 and 0.579 for BP-21 and TV-2,1 respectively). This indicates that the model is not able to recall a large part of the abandoned parcels in these areas, suggesting some inability of the model to detect new abandonment patterns that are not well represented in the training dataset.

Temporal transferability of the trained models was cross-evaluated on models trained with the OLV-19, LSF-20, and LSF-21. The performance of each model was tested with data from years other than the training year. Table 4 shows the results of time transferability.

In the different cross-experiments OA ranged from 0.850 to 0.946, with an average of 0.884. This represents about 3.5 % reduction in performance when using the model out of the training year, which can be considered as acceptable. However, we should highlight that a reduction of the precision and recall of the PR and AB classes, leading to more confusion between the two categories can be observed. Fig. 9 shows a summary of the OA obtained by classifying each dataset. It compares the performance obtained using models trained in years other than the prediction year versus the cross-validation performance in the same year.

3.5. Feature importance on class separability

Fig. 10 shows 4 radar charts of JM distances, which are related to feature separability assessment. The upper part of the figure displays calculated pairwise JM distances of the 3 classes generated by the features previously categorized as central tendency statistics (12 average monthly values of OSAVI and NDMI). The lower part refers to JM values of the 6 features previously categorized as dispersion-related (range October–March, range April–September, standard deviation October–March, standard deviation April–September, annual kurtosis annual skewness of OSAVI an NDMI) for the 3 classes of abandonment status.

All JM values were in the range 0.1–1.6. Despite this apparently low

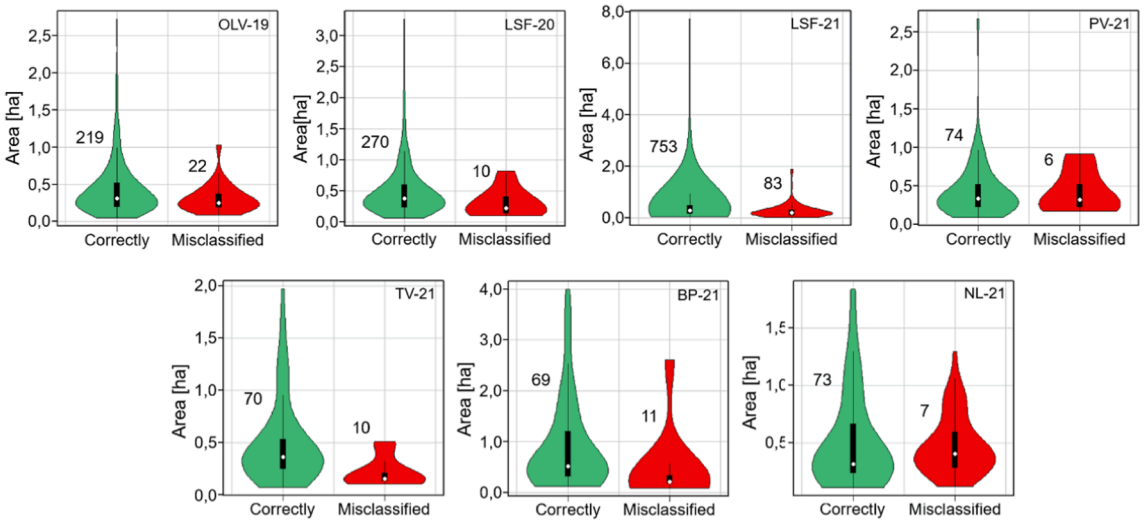

Fig. 8. Size of correctly classified and misclassified parcels in each data set.

Table 3
Spatial transferability results of the model trained with the LSF-21 dataset.

Dataset	Overall accuracy (95 % C.I.)	Cohen's Kappa	NP precision	PR precision	AB precision	NP recall	PR recall	AB recall
PV-21	0.924 ± 0.058	0.879	0.933	0.900	0.958	0.875	1.000	0.852
TV-21	0.875 ± 0.072	0.742	1.000	0.862	0.846	0.900	0.980	0.579
BP-21	0.870 ± 0.074	0.773	0.733	0.913	0.875	0.846	0.955	0.700
NL-21	0.922 ± 0.058	0.858	1.000	0.968	0.700	0.833	0.968	0.875
Average	0.899	0.813	0.916	0.910	0.845	0.863	0.975	0.752

Table 4
Temporal transferability. Cross-evaluation of models trained with the OLV-19, LSF20 and LSF-21 datasets.

Training dataset (year)	Tested with year:	Overall accuracy (95 % C. I.)	Cohen's Kappa	NP precision	PR precision	AB precision	NP recall	PR recall	AB recall
OLV-19 (2019)	2020	0.889 ± 0.037	0.829	0.983	0.979	0.776	0.983	0.775	0.970
	2021	0.851 ± 0.021	0.769	0.922	0.914	0.751	0.819	0.838	0.888
LSF-20 (2020)	2019	0.883 ± 0.040	0.825	0.962	0.772	0.966	0.950	0.975	0.725
	2021	0.883 ± 0.022	0.826	0.950	0.975	0.725	0.981	0.856	0.987
LSF-21 (2021)	2019	0.850 ± 0.045	0.775	0.917	0.733	0.980	0.926	0.962	0.625
	2020	0.946 ± 0.026	0.916	0.936	0.923	0.988	0.983	1.000	0.860

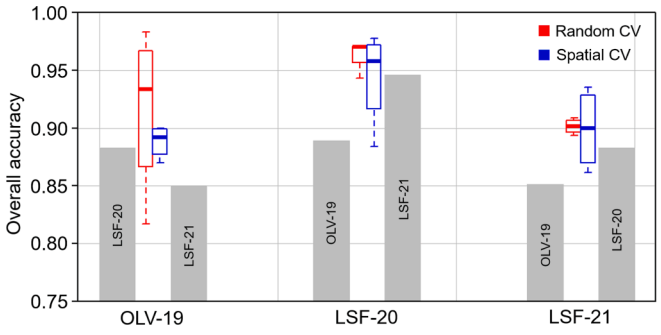

Fig. 9. Performance obtained in each dataset as a function of the model used. The gray bars show the overall accuracy obtained using models trained on data sets from other years and the boxplots show the overall accuracy obtained by cross-validation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

separability, the Random Forest classifier produced high classification accuracy, which reveals the importance of using a multitemporal approach and combined with a powerful machine learning classifier. Monthly averages of both OSAVI and NDMI showed the highest JM values. In fact, removing the rest of variables resulted only in a ~2 % decrease in model performance (data not shown).

Comparison NP-PR produced the highest JM distances for both OSAVI and NDMI features, and PR-AB the lowest. OSAVI derived features induced greater NP–PR and NP–AB separability. However, NDMI derived features produced greater PR–AB separability, being these the most difficult to separate, as observed in Fig. 10.

OSAVI derived features showed higher separability for NP–PR and NP–AB classes in April–August when spontaneous vegetation grows. PR–AB presented higher separability in July–August, which are the warmest and driest months, when spontaneous vegetation wilts. This is the typical lifecycle of *Oxallis pes-caprae*, one of the most extended weeds in the study area. The pattern particularly affects NP orchards, where tree canopies occupy a very small proportion of land. In any case, JM values remained lower than 0.5, which are commonly considered as very low.

Regarding NDMI derived features, separation of NP–PR classes showed a relatively stable pattern throughout the year, with JM values around 1. Separability of NP–AB classes increased in May–June and decreased in December–October. PR–AB separability was higher in September–November. This may be attributed to differences in humidity due to the absence of irrigation in the abandoned orchards. However, JM values of PR–AB remained low (less than 0.7).

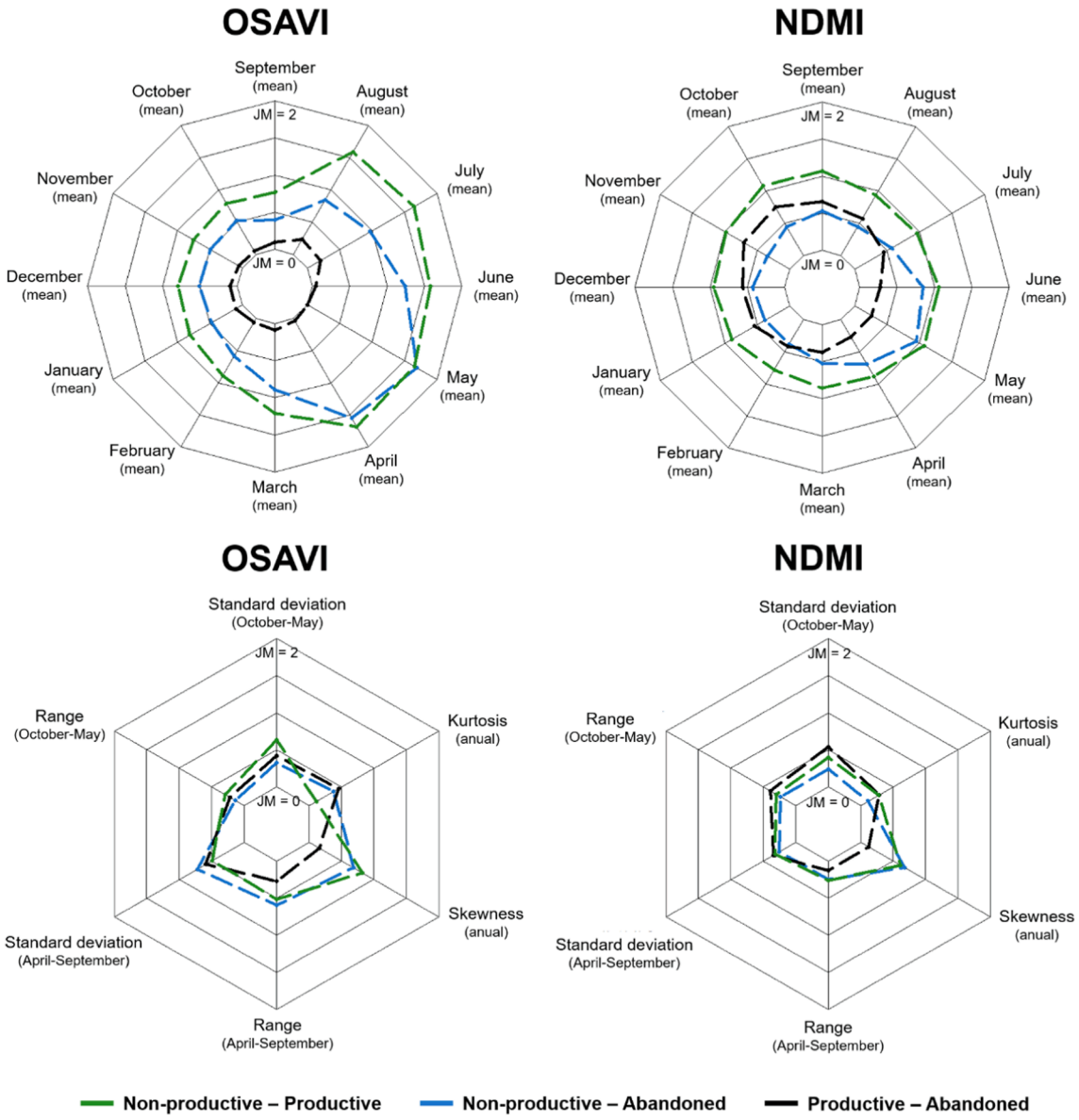

Fig. 10. Jeffries-Matusita separability for each variable studied and each pair of categories. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

This work studied the use of Sentinel-2 time series to identify the status of citrus crops in a highly fragmented agricultural land. The study develops a classification framework for processing OSAVI and NDMI time series extracted from Sentinel-2 L2A data to identify three types of parcels (non-productive, productive, and abandoned). Therefore, it is necessary to have a database of citrus parcels previously identified. In this work we used the Land Parcel Identification System of Spain. This database is updated on the basis of farmers' declarations, review of orthophotos by specialized technicians and field visits and automatic support procedures Amorós López et al., (2011). In other contexts, it would be necessary to have an updated map to identify citrus parcels. The proposed approach involves a supervised classification using the Random Forest algorithm and a novel strategy for classification at parcel level. Results on three different temporal and spatial datasets provided OA ranging from 89 to 92 %, which represents significant improvements of about 13 % over previous classification methods based on single image (Morell-Monzó et al., 2020). We attribute these improvements to the greater availability of images registered in a higher temporal period in which some spectral differences among AB and PR parcels can be amplified. Thus, the Random Forest algorithm can establish more cut-off points that are able to differentiate the two types of parcels. In this context, the presence of dry periods (summers) generates a greater impact on the vegetation abandoned due to lack of irrigation. Overall, the work demonstrates the potential of Sentinel-2 time series for monitoring abandonment of citrus parcels and highlights the importance of the time domain information to separate parcels with similar spectral responses. Despite the low spectral separability between the studied parcel types, Random Forest algorithm allowed to obtain the following average class accuracy: NP = 0.935 > PR = 0.921 > AB = 0.899. The use of powerful machine learning algorithms together with a representative set of training samples allows detecting land use/land cover classes without significant differences between them (Chaves et al., 2020).

To date there are some classes that have been little studied in land use and land cover mapping, such as abandoned perennial crops. With recent advances in sensor technologies, data management and data

analysis, several remote sensing options are available to the scientific community. However, the agricultural sector has yet to fully benefit from remote sensing technologies due to lack of knowledge on their sufficiency, appropriateness and techno-economic feasibility (Khanal et al., 2020). Several research studies highlight the importance of time series data for monitoring seasonal crops (Vuolo et al., 2018; Roumenina et al., 2015; Palchowdhuri et al., 2018), reporting on increasing accuracy when including additional multi-temporal information to Random Forest models. Our work confirms that this importance also extends to permanent, perennial crops. Today, with more advanced technologies, it is possible to explore the time sequence beyond the Red–NIR relationships to determine crop status. Researchers can explore different strategies and multi-source data relationships to analyze landscapes in a more detailed level. New remote sensing technologies enable us to investigate some classes that were not possible in the past. This study demonstrates the feasibility of identifying citrus crop abandonment using Sentinel-2 data.

To our knowledge there are few studies on the identification of abandoned perennial crops. Although a direct comparison of accuracy estimates is not possible, our results support those obtained in detecting abandonment in other perennial crops, like olives (Volpi et al., 2023) or in forest perennial tree species identification (Hemmerling et al., 2021). In contrast to these studies, our methodological approach uses two spectral indices related to two characteristics of the crop. These indices provide complementary information about vegetation vigor and moisture content. Both are key features to differentiate non-productive, productive and abandoned crops. In addition, our approach focuses on the feature extraction from the OSAVI and NDMI time series. This allows to reduce the dimensionality of the problem compared to classification approaches based on dense time series (e.g., Lambert et al., 2018; Vuolo et al., 2018) where all available images are used as model input variables. Finally, our approach uses the agricultural Land Parcel Identification System (LPIS) to produce a parcel-level classification using a strategy that minimizes edge effects. The method reports a single classification value per parcel and the probability of membership, which can be used as a measure of confidence in the assignment (Mitchell et al., 2008). The proposed approach could be used in other agricultural classification contexts (e.g. crop type classification) and other study areas where sufficient cloud-free Sentinel-2 imagery and an updated agricultural LPIS are available.

Since the launch of the Sentinel-2 constellation, less-used vegetation indices have been applied. In particular, the use of spectral indices involving red-edge and SWIR bands has increased as they can improve the classification performance of a wide range of land use/land cover classes. In addition, these bands enable the retrieval of biophysical parameters and vegetation monitoring (Chaves et al., 2020). Our approach used OSAVI and NDMI time series together. For the calculation of these indexes, we selected the bands with the maximum spatial resolution possible considering that this aspect can play an important role in our study area due to the reduced size of the parcels. In this context, we did not use the Sentinel-2 red-edge bands for OSAVI index, selecting the red and near-infrared bands with 10 m of spatial resolution. In contrast, we used SWIR (B11) for calculating NDMI index with a spatial resolution of 20 m as it is the best available alternative for studying the behavior of vegetation and soil. We assume that a higher spatial resolution in the SWIR for NDMI calculation could improve the classification performance.

Overall, this study provides promising results for monitoring the abandonment of citrus orchards in the Comunitat Valenciana region. Accuracies obtained allow the possibility of identifying general trends of land abandonment at medium to large scales, which may help to support policies for adequate landscape management. However, the results also showed limitations in recall of the AB parcels. This fact suggests that some types of abandonment are unidentifiable through our approach. Statistical analysis could not confirm significant differences between the size of correctly classified and misclassified parcels. However, Vajsová et al. (2020) pointed out other characteristics of the parcels not linked to size but shape (for instance, the ratio between the number of full pixels and the number of incomplete pixels within the parcel). More research is needed to associate the characteristics of the parcels and the types of abandonment that can be identified due to the particular resolution of Sentinel-2 images.

Agricultural management practices and environmental conditions influence landscape dynamics and could affect the classification accuracy. Crop management practices are similar throughout the CV region, since citrus are cultivated there for more than 70 years and farmers rely on the same sources of technical information (public advisors or private advisors from cooperatives or unions, information released by the regional authorities) and legislation. The two major differences are that some parcels are irrigated by drip irrigation and others by flood irrigation, and that some farmers allow a weed cover to maintain soil moisture and others are not. In some cases, drip irrigation may be accompanied by fertilization. This variability is represented in our study area. In the coastal areas of the CV, where most of the citrus crops are located, soil variability is high due to the geomorphology of these coastal systems (Viñals, 1995). In the coastal barrier the soils are sandy and less productive, while the alluvial plain is made up of sedimentary materials and the soils are more productive. There are also differences in soil characteristics in areas close to coastal wetlands with vegetation unique to these areas. In addition, the particular soil characteristics of each parcel may be different, as they can often be classified as technosols. It is common for farmers to have added material to the soil to improve its characteristics. Climatically, the south of the Iberian Southeast is included within the dry variant of the semi-arid Mediterranean climates with dry summer (Bs Köppen climate classification) with a gradation of rainfall from 300 mm in the northern coastal sector, to less than 150 mm in the southern zone (Gil-Guirado and Pérez-Morales, 2019). The rest of the Mediterranean coastal area is included in the warm temperate climate with dry and hot summers: typical Mediterranean climate (Csa Köppen climate classification). In this region, we can find areas such as the south of the Valencia province (this study area) with average annual precipitation above 800 mm. In Fig. 5, we can observe the difference in the NDMI between PR and AB class in an area with typically high annual precipitation. It is expected that in semi-arid areas the difference would be higher, and then better separation results could be obtained.

The model showed acceptable spatial and temporal transferability. When transferred the model outside the training area, an average OA of 0.89 was obtained, which represents around a 1 % reduction. However, transferring the model to other areas notably reduced the capacity to recall the AB parcels. Omission errors for the AB class raised to 42.1 % and 30.0 % in some cases. Volpi et al. (2023) studied the abandonment of olive groves in Tuscany (Italy) and obtained a poor recall of AB class. These results highlight the importance of having a database with a well-represented AB class. Abandonment is a process, not a status, that depends on management practices and environmental conditions that affect soil and vegetation. This makes AB a very heterogeneous land use across space whose characteristics depend on time elapsed since the start of abandonment, management practices, local weather conditions affecting growth rates, further modulated by site-specific soil conditions. These conditions affect the crop remaining in the orchard and the vegetative patterns of spontaneous species. This causes variations in the spectral-temporal profile of the AB crops across space. This may cause that intra-class variability is not well represented in the training dataset. From an operational perspective an annual monitoring system requires field data that are representative of the variability present in the study area. Therefore, it is necessary to better understand the influence of the spatial distribution and size of the training data set on model performance.

When transferring the model out of the training year, an average OA of 0.88 was achieved, which represents a 3.5 % reduction. A decrease in recall metric of the AB class was observed. When the model was transferred out of the training year the confusion between AB and PR classes

increased. The phenological profile of the AB parcels, which do not receive irrigation, is more sensitive to the weather of each particular year. The high variability of rainfall between years makes it difficult to obtain a good recall of the AB class when a model trained with data from a single year is transferred to other years. Such performance reductions may be attributed to a lack of correlation between the index values along the different years, which depend on agrometeorological conditions like soil fertility, temperature, and precipitation. This leads to variations in spectral-temporal profiles, which fluctuate on a yearly basis, and introduces uncertainties (Vuolo et al., 2018). Temporal transferability could be a limiting factor for annual monitoring of citrus crop abandonment, since it is impractical to collect consistent, timely, and high-quality reference samples to produce annual maps Therefore, it is necessary to develop efficient methods for generating high-quality reference samples from the year when ground truth data were collected (i.e., reference year) to another year (i.e., target year) (Ghorbanian et al., 2020). A good alternative to overcome the limitations of time transferability is the migration of training samples, which has been successfully used in several previous studies (e.g. Huang et al., 2020; Ghorbanian et al., 2020; Fekri et al., 2021).

Further research is required to improve accuracy and the recall of AB class. We suggest adding texture features able to detect planting patterns (also called planting frames) as these can be a key feature in identifying abandoned crops. However, this would require higher spatial resolution images, as we showed in past studies (Morell-Monzó et al., 2021). Texture information can be extracted from Very High-Resolution (VHR) airborne and satellite images. But airborne images may show limitations to transfer the model spatially and temporally due to differences in acquisition conditions (sensor, date, sun angle, atmospheric conditions, etc.). Image fusion techniques can optimally merge information from image sources of different time and spatial resolutions (Moltó, 2022). Future research should focus on developing methods to combine Sentinel-2 time series and VHR images in a transferable way. Furthermore, it is possible to add new information from synthetic aperture radar (i.e. form Sentinel-1 mission). Moreover, using longer time series or algorithms able to model sequential information (e.g. Long Short-Therm Memory or 1D Convolutional Neural Networks) could also improve model performance (Campos-Taberner et al., 2023; Mohammadi et al., 2023). Future research should also focus on improving spatial and temporal transferability. With respect to spatial transferability, it would be interesting to evaluate the impact of the spatial distribution of training samples on model performance, to develop an analysis of agro-climatic conditions and farm management practices affecting the model. To improve temporal transferability, a possible alternative would be the use of migrated samples to improve the temporal transferability of the model.

5. Conclusions

This work studied the use of Sentinel-2 time series to identify the status of citrus crops in a highly fragmented agricultural landscape in the Comunitat Valenciana region (Spain). A classification framework was developed to identify three types of parcels (non-productive, productive, and abandoned) using features extracted from the OSAVI and NDMI time series. The proposed approach involves a supervised classification using the Random Forest algorithm and a novel strategy for classification at parcel level. Results show the potential of the method to identify citrus crop status. Differences were found between the three types of parcels in the temporal profiles of OSAVI and NDMI associated with vegetation vigor and moisture content that allow the identification of the crop status. Results on three different temporal and spatial datasets provided overall accuracies ranging from 0.89 to 0.92. The model showed acceptable spatial and temporal transferability with an overall accuracy reduction of 1.0 % and 3.5 %, respectively. However, transferring the model to other areas notably reduced the capacity to recall the abandoned parcels. When transferring the model out of the training year there was more confusion between the productive and abandoned parcels. Overall, this study provides promising results for monitoring the abandonment of citrus orchards in the Comunitat Valenciana region. More research is needed to improve identification, generate more transferable models, and understand the abandonment patterns or parcel types that cannot be identified.

Funding

This research was partially funded by regional government of Spain, Generalitat Valenciana, within the framework of the research project AICO/2020/246. Funding for open access charge: CRUE-Universitat Politècnica de València.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Amorós López, J., Izquierdo Verdiguier, E., Gómez Chova, L., Muñoz Marí, J., Rodríguez Barreiro, J.Z., Camps Valls, G., Calpe Maravilla, J., 2011. Land cover classification of VHR airborne images for citrus grove identification. In: ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 66, Issue 1. Elsevier BV, pp. 115–123. doi: 10.1016/j.isprsjprs.2010.09.008.

Asgarian, A., Soffianian, A., Pourmanafi, S., 2016. Crop type mapping in a highly fragmented and heterogeneous agricultural landscape: a case of central Iran using multi-temporal Landsat 8 imagery. In: Computers and Electronics in Agriculture, Vol. 127. Elsevier BV, pp. 531–540. doi: 10.1016/j.compag.2016.07.019.

Breiman, L., 2001. In Machine Learning, Vol. 45, Issue 1. Springer Science and Business Media LLC, pp. 5–32. doi: 10.1023/a:1010933404324.

Campos-Taberner, M., Javier García-Haro, F., Martínez, B., Sánchez-Ruiz, S., Moreno-Martínez, Á., Camps-Valls, G., Amparo Gilabert, M., 2023. Land use classification over smallholding areas in the European Common Agricultural Policy framework. In: ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 197. Elsevier BV, pp. 320–334. doi: 10.1016/j.isprsjprs.2023.02.005.

Chaves, M.E.D., Picoli, M.C.A., Sanches, I.D., 2020. Recent Applications of Landsat 8/OLI and Sentinel-2/MSI for Land Use and Land Cover Mapping: A Systematic Review. Remote Sens. 12 (18), 3062. https://doi.org/10.3390/rs12183062.

Chen, J., Jönsson, Per., Tamura, M., Gu, Z., Matsushita, B., Eklundh, L., 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. In: Remote Sensing of Environment, Vol. 91, Issues 3–4. Elsevier BV, pp. 332–344. doi: 10.1016/j.rse.2004.03.014.

Cohen, J., 1960. A coefficient of agreement for nominal scales. In: Educational and Psychological Measurement, Vol. 20, Issue 1. SAGE Publications, pp. 37–46. doi: 10.1177/001316446002000104.

Compés, R., García, J.M., Martínez, V., 2019. La crisis citrícola en la Comunitat Valenciana y el acuerdo de asociación económica con el sur de África. XII Congreso Economía Agraria AEEA.

Czesak, B., Różycka-Czas, R., Salata, T., Dixon-Gough, R., Hernik, J., 2021. Determining the intangible: detecting land abandonment at local scale. In: Remote Sensing, Vol. 13, Issue 6. MDPI AG, p. 1166. doi: 10.3390/rs13061166.

Dimov, D., 2022. Classification of remote sensing time series and similarity metrics for crop type verification. In: Journal of Applied Remote Sensing, Vol. 16, Issue 02. SPIE-Intl Soc Optical Eng. doi: 10.1117/1.jrs.16.024519.

Fekri, E., Latifi, H., Amani, M., Zobeidinezhad, A., 2021. A training sample migration method for wetland mapping and monitoring using sentinel data in google earth engine. In: Remote Sensing, Vol. 13, Issue 20. MDPI AG, p. 4169. doi: 10.3390/rs13204169.

Gao, B., 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. In: Remote Sensing of Environment, Vol. 58, Issue 3. Elsevier BV, pp. 257–266. doi: 10.1016/s0034-4257(96)00067-3.

García Álvarez-Coque J.M., Moltó García, E. (coords.), 2020. Una hoja de ruta para la citricultura española. Cajamar Caja Rural. ISBN: 978-84-95531-49-0.

Gasith, A., Resh, V.H., 1999. Streams in the Mediterranean climate regions: abiotic influences and biotic responses to predictable seasonal events. Annu. Rev. Ecol. Syst. 30, 51–81.

Geleralitat Valenciana. Portal Estadístico de la Generalitat Valenciana. Fichas Municipales 2020. Available online: http://www.pegv.gva.es/es/fichas (accessed on 15 November 2022).

Ghorbanian, A., Kakooei, M., Amani, M., Mahdavi, S., Mohammadzadeh, A., Hasanlou, M., 2020. Improved land cover map of Iran using Sentinel imagery within Google Earth Engine and a novel automatic workflow for land cover classification using migrated training samples. In: ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 167. Elsevier BV, pp. 276–288. doi: 10.1016/j.isprsjprs.2020.07.013.

Gil-Guirado, S., Pérez-Morales, A., 2019. Variabilidad climática y patrones termopluviométricos en Murcia (1863–2017). Técnicas de análisis climático en un

contexto de cambio global. In: Investigaciones Geográficas, Issue 71. Universidad de Alicante Servicio de Publicaciones, p. 27. doi: 10.14198/ingeo2019.71.02.
Gómez, C., White, J.C., Wulder, M.A., 2016. Optical remotely sensed time series data for land cover classification: a review. In: ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 116. Elsevier BV, pp. 55–72. doi: 10.1016/j.isprsjprs.2016.03.008.
González-Hidalgo, J.C., de Luís Arrillaga, M., Peña Monné, J.L., 2005. Extreme rainfall events, climate variability and soil erosion. Some comments related to climate change in Mediterranean environments. Rev. C. & C. 19 (1–2), 49–62.
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. In: Remote Sensing of Environment, Vol. 202. Elsevier BV, pp. 18–27. doi: 10.1016/j.rse.2017.06.031.
Hemmerling, J., Pflugmacher, D., Hostert, P., 2021. Mapping temperate forest tree species using dense Sentinel-2 time series. In: Remote Sensing of Environment, Vol. 267. Elsevier BV, p. 112743. doi: 10.1016/j.rse.2021.112743.
Huang, H., Wang, J., Liu, C., Liang, L., Li, C., Gong, P., 2020. The migration of training samples towards dynamic global land cover mapping. In: ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 161. Elsevier BV, pp. 27–36. doi: 10.1016/j.isprsjprs.2020.01.010.
Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). In: Remote Sensing of Environment, Vol. 25, Issue 3. Elsevier BV, pp. 295–309. doi: 10.1016/0034-4257(88)90106-x.
Karasiak, N., Dejoux, J.-F., Monteil, C., Sheeren, D., 2021. Spatial dependence between training and test sets: another pitfall of classification accuracy assessment in remote sensing. In: Machine Learning, Vol. 111, Issue 7. Springer Science and Business Media LLC, pp. 2715–2740. doi: 10.1007/s10994-021-05972-1.
Khanal, S., KC, K., Fulton, J.P., Shearer, S., Ozkan, E., 2020. Remote sensing in agriculture—Accomplishments, limitations, and opportunities. Remote Sens. 12 (22) https://doi.org/10.3390/rs12223783. MDPI AG.
Lambert, M.-J., Traoré, P.C.S., Blaes, X., Baret, P., Defourny, P., 2018. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. In: Remote Sensing of Environment, Vol. 216. Elsevier BV, pp. 647–657. doi: 10.1016/j.rse.2018.06.036.
Leal Filho, W., Mandel, M., Al-Amin, A.Q., Feher, A., Chiappetta Jabbour, C.J., 2016. An assessment of the causes and consequences of agricultural land abandonment in Europe. In: International Journal of Sustainable Development & World Ecology, Vol. 24, Issue 6. Informa UK Limited, pp. 554–560. doi: 10.1080/13504509.2016.1240113.
Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 2(3), 18–22. https://cran.r-project.org/web/packages/randomForest/.
Ministerio de Agricultura, Pesca y Alimentación. ESYRCE: Encuesta Sobre Superficies y Rendimientos del año 2020. Ministerio de Agricultura, Pesca y Alimentación, Madrid, Spain, 2021. https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/esyrce/.
Ministerio de Agricultura, Pesca y Alimentación. ESYRCE: Encuesta Sobre Superficies y Rendimientos del año 2021; Ministerio de Agricultura, Pesca y Alimentación, Madrid, Spain, 2022. https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/esyrce/.
Mitchell, S.W., Remmel, T.K., Csillag, F., Wulder, M.A., 2008. Distance to second cluster as a measure of classification confidence. In: Remote Sensing of Environment, Vol. 112, Issue 5. Elsevier BV, pp. 2615–2626. doi: 10.1016/j.rse.2007.12.006.
Mohammadi, S., Belgiu, M., Stein, A., 2023. Improvement in crop mapping from satellite image time series by effectively supervising deep neural networks. In: ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 198. Elsevier BV, pp. 272–283. doi: 10.1016/j.isprsjprs.2023.03.007.
Moltó, E., 2022. Fusion of different image sources for improved monitoring of agricultural plots. In Sensors, Vol. 22, Issue 17. MDPI AG, p. 6642. doi: 10.3390/s22176642.
Morell-Monzó, S., Garófano-Gómez, V., 2022. Investigación del abandono de tierras en la comarca de La Safor (Comunitat Valenciana) utilizando datos Sentinel-2. CDR La Safor, Beniarjó, València.
Morell-Monzó, S., Estornell, J., Sebastiá-Frasquet, M.-T., 2020. Comparison of Sentinel-2 and high-resolution imagery for mapping land abandonment in fragmented areas. In: Remote Sensing, Vol. 12, Issue 12. MDPI AG, p. 2062. doi: 10.3390/rs12122062.
Morell-Monzó, S., Sebastiá-Frasquet, M.-T., Estornell, J., 2021. Land use classification of VHR images for mapping small-sized abandoned citrus plots by using spectral and textural information. In Remote Sensing, Vol. 13, Issue 4. MDPI AG, p. 681. doi: 10.3390/rs13040681.
Morell-Monzó, S., Estornell, J., Sebastiá-Frasquet, M.T., 2022. Clasificación del estado de parcelas de cítricos utilizando datosmultitemporales Sentinel-2. En Teledetección para una Agricultura Sostenible en la era del Big Data, Actas del XIX Congreso de la AsociaciónEspañola de Teledetección, Pamplona, pp. 35-38. 2022. Disponible en http://www.aet.org.es/?q=congresos.
Morell-Monzó, S., Estornell, J., Sebastiá-Frasquet, M.-T., 2023. Assessing the capabilities of high-resolution spectral, altimetric, and textural descriptors for mapping the status of citrus parcels. In: Computers and Electronics in Agriculture, Vol. 204. Elsevier BV, p. 107504. doi: 10.1016/j.compag.2022.107504.
Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. In: Remote Sensing of Environment, Vol. 148. Elsevier BV, pp. 42–57. doi: 10.1016/j.rse.2014.02.015.
Palchowdhuri, Y., Valcarce-Diñeiro, R., King, P., Sanabria-Soto, M., 2018. Classification of multi-temporal spectral indices for crop type mapping: a case study in Coalville, UK. In: The Journal of Agricultural Science, Vol. 156, Issue 1. Cambridge University Press (CUP), pp. 24–36. doi: 10.1017/s0021859617000879.
Perpiña Castillo, C., Kavalov, B., Diogo, V., Jacobs-Crisioni, C., Batista e Silva, F., Lavalle, C., 2018. Agricultural Land Abandonment in the EU within 2015-2030. JRC113718. European Commission.
Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V., Murayama, Y., Ranagalage, M., 2020. Sentinel-2 data for land cover/use mapping: a review. In: Remote Sensing, Vol. 12, Issue 14. MDPI AG, p. 2291. doi: 10.3390/rs12142291.
Richards, J.A., Jia, X., 2006. Remote Sensing Digital Image Analysis: An Introduction. Springer, Berlin, Germany.
Rondeaux, G., Steven, M., Baret, F., 1996. Optimization of soil-adjusted vegetation indices. In: Remote Sensing of Environment, Vol. 55, Issue 2. Elsevier BV, pp. 95–107. doi: 10.1016/0034-4257(95)00186-7.
Roumenina, E., Atzberger, C., Vassilev, V., Dimitrov, P., Kamenova, I., Banov, M., Filchev, L., Jelev, G., 2015. Single- and multi-date crop identification using PROBA-V 100 and 300 m S1 products on Zlatia Test Site, Bulgaria. In: Remote Sensing, Vol. 7, Issue 10. MDPI AG, pp. 13843–13862. doi: 10.3390/rs71013843.
Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36, 1627–1639.
Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., Homayouni, S., 2020. Support vector machine versus random forest for remote sensing image classification: a meta-analysis and systematic review. In: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 13. Institute of Electrical and Electronics Engineers (IEEE), pp. 6308–6325. doi: 10.1109/jstars.2020.3026724.
Steven, M.D., 1998. The sensitivity of the OSAVI vegetation index to observational parameters. In: Remote Sensing of Environment, Vol. 63(1). Elsevier BV, pp. 49–60. https://doi.org/10.1016/s0034-4257(97)00114-4.
Stock, A., Subramaniam, A., 2022. Iterative spatial leave-one-out cross-validation and gap-filling based data augmentation for supervised learning applications in marine remote sensing. In: GIScience & Remote Sensing, Vol. 59, Issue 1. Informa UK Limited, pp. 1281–1300. doi: 10.1080/15481603.2022.2107113.
Subedi, Y.R., Kristiansen, P., Cacho, O., 2022. Drivers and consequences of agricultural land abandonment and its reutilisation pathways: a systematic review. In: Environmental Development, Vol. 42. Elsevier BV, p. 100681. doi: 10.1016/j.envdev.2021.100681.
Terres, J.-M., Scacchiafichi, L.N., Wania, A., Ambar, M., Anguiano, E., Buckwell, A., Coppola, A., Gocht, A., Källström, H.N., Pointereau, P., Strijker, D., Visek, L., Vranken, L., Zobena, A., 2015. Farmland abandonment in Europe: identification of drivers and indicators, and development of a composite indicator of risk. In: Land Use Policy, Vol. 49. Elsevier BV, pp. 20–34. doi: 10.1016/j.landusepol.2015.06.009.
Vajsová, B., Fasbender, D., Wirnhardt, C., Lemajic, S., Devos, W., 2020. Assessing spatial limits of sentinel-2 data on arable crops in the context of checks by monitoring. In: Remote Sensing, Vol. 12, Issue 14. MDPI AG, p. 2195. doi: 10.3390/rs12142195.
Viñals, M.J., 1995. Secuencias Estratigráficas y Evolución Morfológica del Extremo Meridional del Golfo de Valencia (Cullera-Dénia). El Cuaternario del País Valenciano, 1st ed. Universitat de València-AEQUA, Valencia, Spain.
Volpi, I., Marchi, S., Petacchi, R., Hoxha, K., Guidotti, D., 2023. Detecting olive grove abandonment with Sentinel-2 and machine learning: the development of a web-based tool for land management. In: Smart Agricultural Technology, Vol. 3. Elsevier BV, p. 100068. doi: 10.1016/j.atech.2022.100068.
Vuolo, F., Neuwirth, M., Immitzer, M., Atzberger, C., Ng, W.-T., 2018. How much does multi-temporal Sentinel-2 data improve crop type classification? In: International Journal of Applied Earth Observation and Geoinformation, Vol. 72. Elsevier BV, pp. 122–130. doi: 10.1016/j.jag.2018.06.007.
Wadoux, A.M.J.-C., Heuvelink, G.B.M., de Bruin, S., Brus, D.J., 2021. Spatial cross-validation is not the right way to evaluate map accuracy. In: Ecological Modelling, Vol. 457. Elsevier BV, p. 109692. doi: 10.1016/j.ecolmodel.2021.109692.